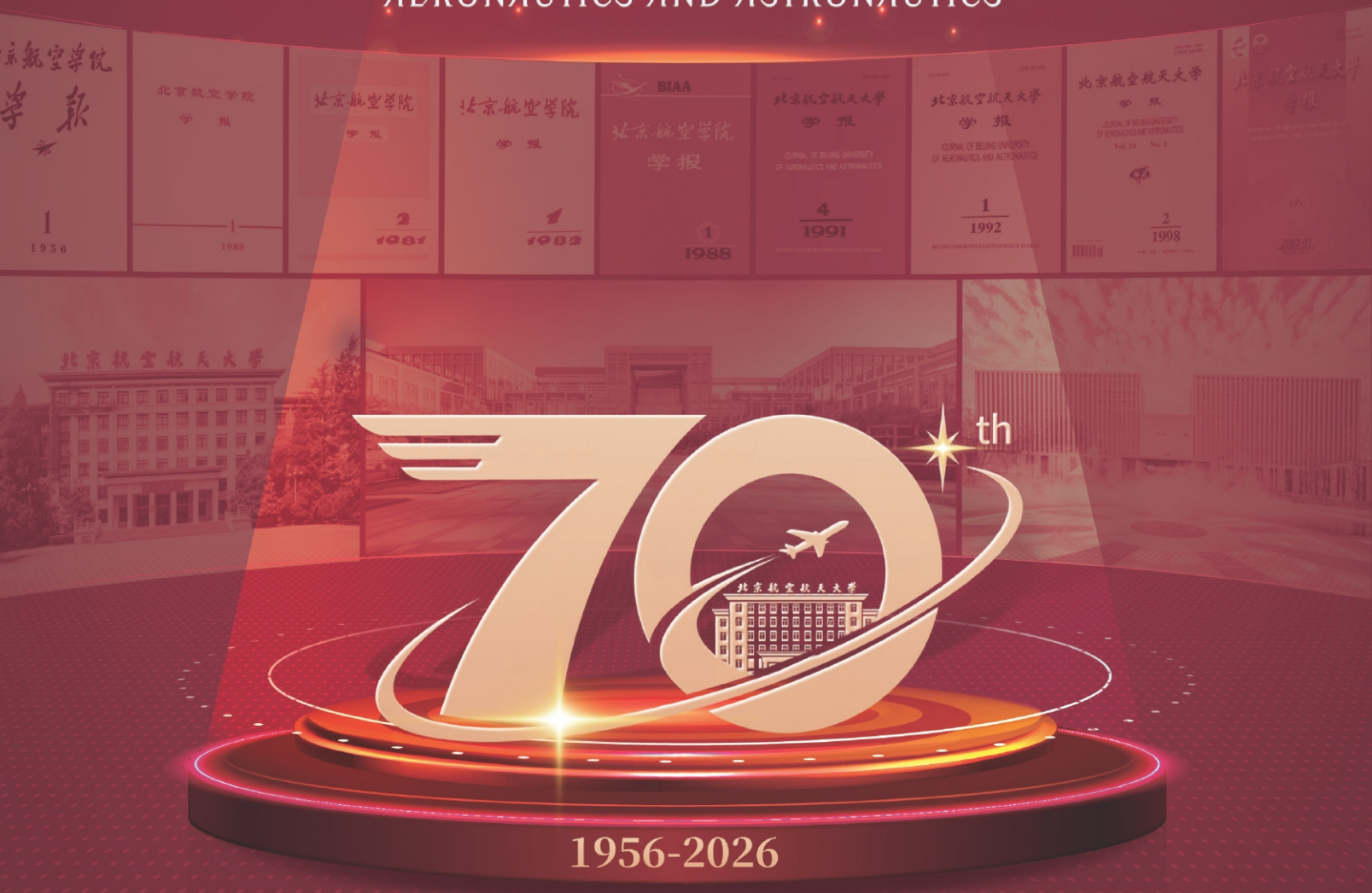
主管单位：中华人民共和国工业和信息化部
主办单位：北京航空航天大学

ISSN 1001-5965
CN 11-2625/V

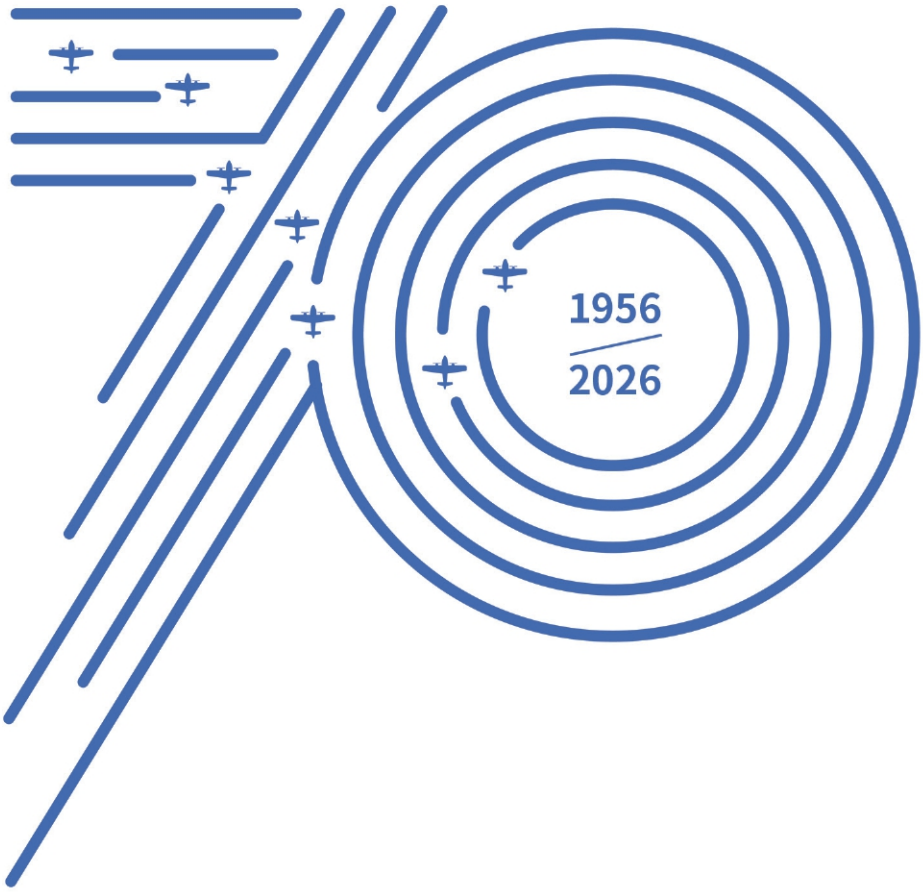
北京航空航天大學学报

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS



庆祝
北京航空航天大学学报
创刊70周年

10 | 2026
Vol.52



七十载空天报国
新时代逐梦一流

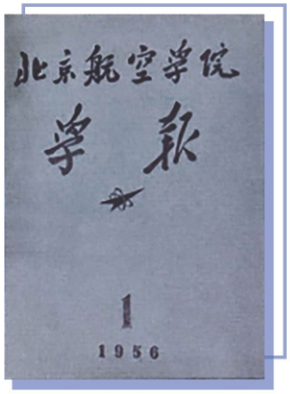
七秩耕耘创新不辍
空天报国再启征程

丙午桂月赵长禄

北京航空航天大学

校盛刊强

1956



1980



1981



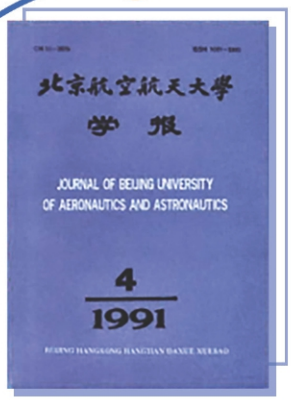
1982



1988



1991



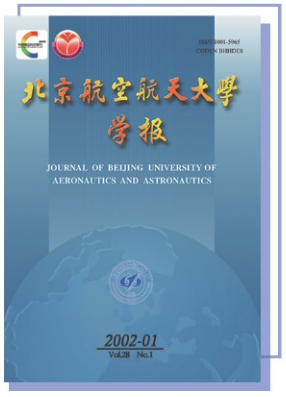
1992



1998



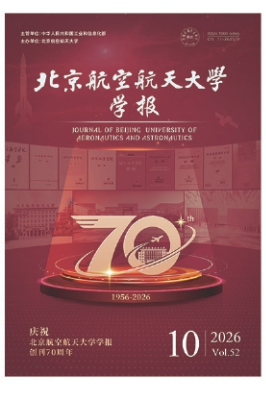
2002



2022



2026



北京航空航天大学学报

第 52 卷　第 10 期　（总第 404 期）　2026 年 10 月

目　　次

综述

强预冷发动机用紧凑快速强换热技术进展 …… 王一帆，李辉，陈一鸣，付超，邹正平（3237）
全固态电池特性建模与仿真研究进展 …… 赵立金，杨世春，刘国芳，曲婧瑶，孙旭东（3276）
热障涂层的智能化设计研究进展 …… 何健，李宏峰，康鹏，郑蕾，李圳，郭洪波（3298）
宽温域超弹性形状记忆合金设计优化研究进展 …… 钱明芳，尹钊，王桂松，张学习，耿林（3313）
镁合金半固态触变射铸技术研究进展 …… 姜巨福，张金泽，崔竞波，王迎，张晓东（3329）
氧化亚氮推进技术发展现状与趋势 …… 魏祥庚，谢远，路伟杰，刘雄峰，高玉闪，赵昊阳（3350）
燃料电池航空动力谱系——重塑零碳飞行新格局 ……
…… 李致朋，高子博，刘伟栋，彭飞扬，刘鸿，韩庆（3369）
人工智能赋能小行星抵近探测的研究现状及展望 ……
…… 尹继豪，魏晓东，崔林艳，吴文峻，张晓明，刘传凯（3396）
锂离子电池模组电热耦合灾变机制研究进展 …… 贺元骅，梁家鑫，李子桐，黄江（3408）
中国空间站延寿需求下机电产品在轨寿命再评估技术与展望 ……
…… 黄首清，夏侨丽，许庶，余溢方，刘守文，张兆霖（3426）
面向复杂飞行的气动/运动/控制耦合数值虚拟飞行研究综述 ……
…… 邵帅，梅亚飞，姚梦玉，高显忠，包磊（3437）
液流电池用压滤式电化学反应器传质优化研究进展 …… 刘君，韩治成，相艳，卢善富，王海宁（3453）
增程式飞机发展现状与展望 …… 冯诗愚，付子祺，刘冠男，台环宇，耿雷铭，缪日成（3471）
机载智能氧气系统关键技术综述 …… 卜雪琴，刘一鸣，林贵平，郁嘉，于坤洋（3482）
仿生沟槽减阻研究现状与进展 …… 王巍，崔申奥，陈家华，赵浩翔，张震（3496）
无人机视角下目标检测的 YOLO 系列算法研究进展 …… 陈金林，吴一全，苑玉彬（3509）
水下重力辅助惯性导航关键技术综述与展望 ……
…… 周召发，李易依，张志利，陈坡，边英磊，李新宇（3535）
氢燃料电池无人机研究进展与关键技术 …… 徐伟强，张炳鑫，祝赫，陈黎（3563）

论文

基于自适应加权优先虚拟自博弈的多机协同空战决策算法 ……
…… 周尧明，韩韬，姜雨菲，杨政，田艺枫（3580）
运载火箭地月转移入轨混合制导方法 …… 郭康，李惠峰，包为民（3592）
基于控制向量参数化的临近空间飞行器轨迹优化 …… 王远卓，代洪华，岳晓奎（3602）
考虑安全性评估的无人机集群鸽群优化避碰决策 …… 彭雅兰，段海滨，魏晨（3618）
一种高精度的星光折射天文导航观星方法 …… 詹先军，王新龙（3627）
基于在轨加注的载人登月任务分段式轨道设计 …… 李其玲，师鹏，陈勇（3635）
多模态生理信号驾驶舒适性机器学习模型构建与比较 ……
…… 姜维胜，周前祥，柳忠起，刘宇双，王家涛，靳萌萌（3647）
2020—2024 年火箭发射排放特性研究 …… 宋佳，杨宇，韩旺，杨立军（3658）
基于触发式批处理辨识的固定翼无人机自适应跟踪控制 …… 邸尚明，申佳军，刘烨晖，王薇（3667）

期刊基本参数：CN 11-2625/V * 1956 * m * A4 * 440 * zh * P * ￥50.00 * 400 * 27 * 2026-10

（**编　辑**　张　嵘　孙　芳　万玉敏　卞欢欢　王　茜　李亚泰）

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Vol. 52 No. 10 (Sum 404) October 2026

CONTENTS

Review

Progress in compact and rapid heat transfer technology for precooled engines
…… WANG Yifan, LI Hui, CHEN Yiming, FU Chao, ZOU Zhengping (3237)
Research progress on modeling and simulation of characteristics of all-solid-state batteries
…… ZHAO Lijin, YANG Shichun, LIU Guofang, QU Jingyao, SUN Xudong (3276)
Research progress on intelligent design of thermal barrier coatings
…… HE Jian, LI Hongfeng, KANG Peng, ZHENG Lei, LI Zhen, GUO Hongbo (3298)
Progress in design and optimization of wide-temperature-range superelastic shape memory alloys
…… QIAN Mingfang, YIN Zhao, WANG Guisong, ZHANG Xuexi, GENG Lin (3313)
Research progress in semi-solid thixomolding technology of magnesium alloys
…… JIANG Jufu, ZHANG Jinze, CUI Jingbo, WANG Ying, ZHANG Xiaodong (3329)
Current status and trend of nitrous oxide propulsion technology
…… WEI Xianggeng, XIE Yuan, LU Weijie, LIU Xiongfeng, GAO Yushan, ZHAO Haoyang (3350)
Fuel cells for aviation: a new paradigm for carbon-neutral flight
…… LI Zhipeng, GAO Zibo, LIU Weidong, PENG Feiyang, LIU Hong, HAN Qing (3369)
Artificial intelligence for asteroid proximity exploration research progress and future perspectives
…… YIN Jihao, WEI Xiaodong, CUI Linyan, WU Wenjun, ZHANG Xiaoming, LIU Chuankai (3396)
Progress in the investigation of electro-thermal coupled catastrophic mechanisms in lithium-ion battery modules
…… HE Yuanhua, LIANG Jiaxin, LI Zitong, HUANG Jiang (3408)
On-orbit life re-evaluation technology and prospects for electromechanical products in response to demand for China Space Station lifespan extension …… HUANG Shouqing, XIA Qiaoli, XU Shu, YU Yifang, LIU Shouwen, ZHANG Zhaolin (3426)
A review of numerical virtual flight studies for aerodynamic, motion and control coupling in complex flight conditions
…… SHAO Shuai, MEI Yafei, YAO Mengyu, GAO Xianzhong, BAO Lei (3437)
Research progress on mass transfer optimization of filter-press type electrochemical reactors for flow batteries
…… LIU Jun, HAN Zhicheng, XIANG Yan, LU Shanfu, WANG Haining (3453)
Developing status and prospect of range-extended aircraft
…… FENG Shiyu, FU Ziqi, LIU Guannan, TAI Huanyu, GENG Leiming, MIU Richeng (3471)
Review of key technologies of on-board intelligent oxygen system
…… BU Xueqin, LIU Yiming, LIN Guiping, YU Jia, YU Kunyang (3482)
Research status and progress of bionic groove drag reduction
…… WANG Wei, CUI Shen'ao, CHEN Jiahua, ZHAO Haoxiang, ZHANG Zhen (3496)
Research progress on YOLO series algorithms for target detection from UAV vision
…… CHEN Jinlin, WU Yiquan, YUAN Yubin (3509)
Overview and prospect of key technologies for underwater gravity-aided inertial navigation
…… ZHOU Zhaofa, LI Yiyi, ZHANG Zhili, CHEN Po, BIAN Yinglei, LI Xinyu (3535)
Research progress and key technologies of hydrogen fuel cell UAVs
…… XU Weiqiang, ZHANG Bingxin, ZHU He, CHEN Li (3563)

Article

A multi-aircraft cooperative air combat decision-making algorithm based on an adaptive weighted priority virtual self-play theory approach …… ZHOU Yaoming, HAN Tao, JIANG Yufei, YANG Zheng, TIAN Yifeng (3580)
Hybrid guidance method for launch vehicle trans-lunar injection
…… GUO Kang, LI Huifeng, BAO Weimin (3592)
Near-space vehicle trajectory optimization based on control vector parameterization
…… WANG Yuanzhuo, DAI Honghua, YUE Xiaokui (3602)
A safety-evaluation-driven UAV swarm collision avoidance decision-making based on pigeon-inspired optimization
…… PENG Yalan, DUAN Haibin, WEI Chen (3618)
A high-precision starlight refraction celestial navigation observation method
…… ZHAN Xianjun, WANG Xinlong (3627)
Phased orbit design for manned lunar landing mission with on-orbit refueling
…… LI Qiling, SHI Peng, CHEN Yong (3635)
Construction and comparison of machine learning models for driving comfort assessment using multimodal physiological signals
…… JIANG Weisheng, ZHOU Qianxiang, LIU Zhongqi, LIU Yushuang, WANG Jiatao, JIN Mengmeng (3647)
A study on emission characteristics of rocket launches (2020—2024)
…… SONG Jia, YANG Yu, HAN Wang, YANG Lijun (3658)
Batch-triggered identification based adaptive tracking control for fixed-wing UAV
…… DI Shangming, SHEN Jiajun, LIU Yehui, WANG Wei (3667)

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0518

强预冷发动机用紧凑快速强换热技术进展

王一帆[1]，李辉[2]，陈一鸣[1,3]，付超[1,3]，邹正平[1,3,*]

(1. 北京航空航天大学 航空发动机研究院，北京 102206；
2. 中国运载火箭技术研究院 临近空间物理重点实验室，100076；
3. 北京航空航天大学 航空发动机气动热力国防科技重点实验室，北京 100191)

摘 要： 紧凑快速强换热技术可在超紧凑空间、毫秒级极短时间内高效收集-转移发动机进气或内部大量热量，是决定强预冷发动机性能优劣与技术成败的核心关键，其内部极端工作环境及苛刻性能需求导致传统民用换热器的分析与设计方法难以适用、实验测试手段也难以满足其验证需求。因此，围绕强预冷发动机应用开展适用性强、针对性高的紧凑快速强换热技术研究，具有重要的学术价值与工程意义。基于此，从机理、技术与实验等多个维度，系统综述了紧凑快速强换热技术的研究进展，并分析了该技术在强预冷发动机应用中面临的主要挑战。现有研究已取得了广泛而深入的成果，包括超临界工质在微尺度通道与横掠管束中的快速强换热机理、快速仿真与强变物性工质相似性分析方法、不确定性与精细化设计方法及面向紧凑空间极端热环境的实验测试手段等。总体来看，在基础研究方面，亟须深入理解微小尺度复杂热环境下流动换热耦合机理，进一步完善紧凑快速强换热的高精度经验关联数据库；在分析与设计方面，应关注融合新兴手段的紧凑快速强换热分析及设计方法，以提升紧凑快速强换热在宽工作范围内的整体性能；在实验验证方面，有必要进一步发展适用于狭小空间复杂结构的精细化流场测试方法，并在真实装机匹配条件下开展全温全压实验验证。上述研究将为提升紧凑快速强换热技术成熟度、推动强预冷发动机技术发展与验证提供坚实支撑。

关 键 词： 强预冷发动机；紧凑快速强换热；紧凑式换热器；预冷器；微小尺度；流动换热
中图分类号： V236
文献标志码： A **文章编号：** 1001-5965(2026)10-3237-39

水平起降、可重复使用空天飞行器具备5倍声速以上工作速度、常规跑道起降等技术优势，将建立全新的低成本、快速响应全球航空新范式，已成为当前航空航天领域研究热点之一[1-4]。

动力系统是实现水平起飞至高马赫数工作的决定性因素，而现有空天动力方案受限于各自工作原理及适用范围，均难以完全满足全域高性能的动力需求。面对此类动力系统难题，协同融合的组合动力方案是一个重要发展路线，其中，拓展涡轮发动机工作速域是关键突破方向之一[5-6]。在低马赫数(0~3)飞行时，涡轮发动机是技术成熟、可靠性高、比冲高的理想动力形式；而在高马赫数(>3)时，涡轮发动机性能将急剧恶化且可靠性降低。当前，融合冲压发动机在高马赫数下性能优势的涡轮基组合循环发动机(turbine-based combined cycle, TBCC)是重要突破方向之一，但受限于模态转换、飞发一体化等关键技术瓶颈，尚未实现成熟的工程应用。此外，现有涡轮发动机在更高马赫数下高效可靠工作的另一核心难点是如何实现对进气来流高温中蕴含的热量进行高效可靠调控，以解决热力

收稿日期：2025-07-24；录用日期：2025-09-12；网络出版时间：2025-12-03 14:46
网络出版地址：link.cnki.net/urlid/11.2625.V.20251203.1139.002
基金项目：国家自然科学基金 (52406004)；先进航空动力创新工作站资助 (HKCX2024-01-001)
*通信作者. E-mail：zouzhengping@buaa.edu.cn

引用格式：王一帆，李辉，陈一鸣，等. 强预冷发动机用紧凑快速强换热技术进展 [J]. 北京航空航天大学学报，2026，52（10）：3237-3275. Wang Y F，Li H，Chen Y M，et al. Progress in compact and rapid heat transfer technology for precooled engines [J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（10）：3237-3275（in Chinese）.

循环效率急剧降低、部件难以冷却等性能及结构方面的问题[7-8]。紧凑快速强换热技术从空天动力面临的极端热问题出发，多采用具有极高的表面积与体积比的微小尺度通道换热器[9-10]，利用中间换热介质或燃料在有限空间、毫秒级极短时间内高效收集-转移发动机进气或内部大量热量。采用该技术一方面可解决高温进气来流高滞止温度对发动机产生的不利影响，另一方面可实现发动机内部热源-热沉间热能的高效转移转化，以提升发动机整体工作性能[11]。

由于紧凑快速强换热技术突出的技术潜力和广泛的应用前景，国内外自20世纪60年代开始提出了诸多强预冷发动机方案。国外代表方案包括俄罗斯深冷空气涡轮火箭发动机(deeply cooled air turborocket, ATRDC)[12]、日本膨胀循环空气涡轮火箭发动机(air-turboramjet with expander cycle, ATREX)[13]及预冷涡喷发动机(pre-cooled turbojet, PCTJ)[14]、英国反应发动机公司(Reaction Engines Limited, REL)协同式吸气火箭循环发动机“佩刀”SABRE[15-16]及Scimitar“弯刀”发动机[17]。国内代表方案包括北京航空航天大学吸气式强预冷发动机(hypersonic airbreathing pre-cooled engine, HAPE)[18]及预冷涡轮冲压组合发动机(precooled turbine-based combined cycle, PCTBCC)[5]、北京动力机械研究所多循环深度耦合预冷发动机(incorporated multi-cycle cooling engine, ICE)[19]及西安航天动力研究所预冷空气涡轮火箭发动机(precooling air turbo rocket engine, PATR)[20]等。考虑到上述方案展现出的巨大发展潜力，近年来各国开展了大量理论及实验研究工作，并在宽域高效总体[21-22]、高效换热[23]、宽域进排气[24-25]、高效燃烧[26]、叶轮机[27-28]等方面取得了多项关键技术突破。2025年7月，英国Frazer-Nash公司进一步公布了一项名为INVICTUS的空天飞行器研究计划[29]。该计划受到欧洲航天局高度重视及大力资助，吸纳了REL公司核心技术团队及其积累的大量强预冷技术研发成果，拟于2031年初以氢燃料吸气式强预冷发动机为动力完成飞行器研制及试飞，凸显了强预冷发动机方案在高超声速领域的重要应用价值。

纵观上述强预冷发动机方案发展历程及技术特点，结合已开展的大量关键技术研究及核心部件样机研制过程，不难看出，实现有限空间、有限质量约束下的快速强换热是决定强预冷发动机性能乃至成败的核心。例如，ATRDC发动机研制终止的关键原因之一是换热收益与质量代价无法平衡，其强预冷器质量超过整机40%以上且难以满足深冷的换热需求，致使预冷收益远低于代偿损失[30]；此外，研究指出强预冷器性能是决定强预冷发动机高马赫数性能的关键，强预冷器效能每提升1%，压气机压比将提升2%以上，而强预冷器质量每增加1%，将使整机推重比降低约0.4%(以SABRE4热力循环为例)[31]。

从发动机总体性能和强换热需求角度出发分析，强预冷发动机紧凑快速强换热技术与传统民用换热器在技术特征上的区别主要表现在以下方面：①瞬时大温降需求(10 ms级将约1 250 K降至约300～500 K传统航发工作温度)[32-33]；②低流动阻力需求(总压损失增加1%，则推力损失约1.1%以上)[31]；③高效-轻质-紧凑兼顾需求(目前地面换热设备单位质量较低，以典型管壳式换热器为例，其单位质量换热功率一般不超过10 kW/kg，紧凑度一般不超过300 m^2/m^3，而紧凑式换热器单位质量换热功率一般不超过20 kW/kg，但适用于强预冷发动机的强预冷器性能指标需达到不小于100 kW/kg及1 000 m^2/m^3)[34]。更进一步，强预冷器在发动机内部面临更为复杂严苛的工作环境，包括：极端的温度和压力条件(面向Ma=5工况的强预冷器高温侧温度约1 250 K，采用超临界氦作为中间换热介质工作压力约20 MPa)[35]；超强温度梯度(换热器内外侧存在数百K、进出口存在数百至1 000 K以上温差)[33,36]；工作环境受限且耦合复杂(与进气道、压气机等上下游部组件及整机结构布局匹配)[37]；换热介质热物理性质影响显著(跨临界点过程存在换热恶化或强化)[38]等。

1 本文内容及框架

复杂严峻的工作环境及苛刻的性能需求使得传统民用换热器技术难以适配强预冷发动机紧凑快速强换热器需要，传统换热器分析及设计方法、实验测试方法等方面无法直接应用，而已有实验结果难以有效指导分析及设计技术发展。为此，作为决定强预冷发动机能否实现的核心，有必要从机理-技术-实验多维度开展适用于强预冷发动机紧凑快速强换热技术的流动换热理论及实验工作。考虑到目前尚无文献针对该问题最新进展进行全面梳理，本文以适用于强预冷发动机应用需求的紧凑快速强换热技术为对象，从内部流动换热机理、分析及设计技术、实验测试技术及验证(见图1)多维度详细介绍了国内外及北京航空航天大学强预冷空天动力研究团队(以下简称“北航团队”)研究进展，总结当前不足并讨论了紧凑快速强换热技术的挑战与发展趋势，以期为空天动力相关技术发展提供重要思路。

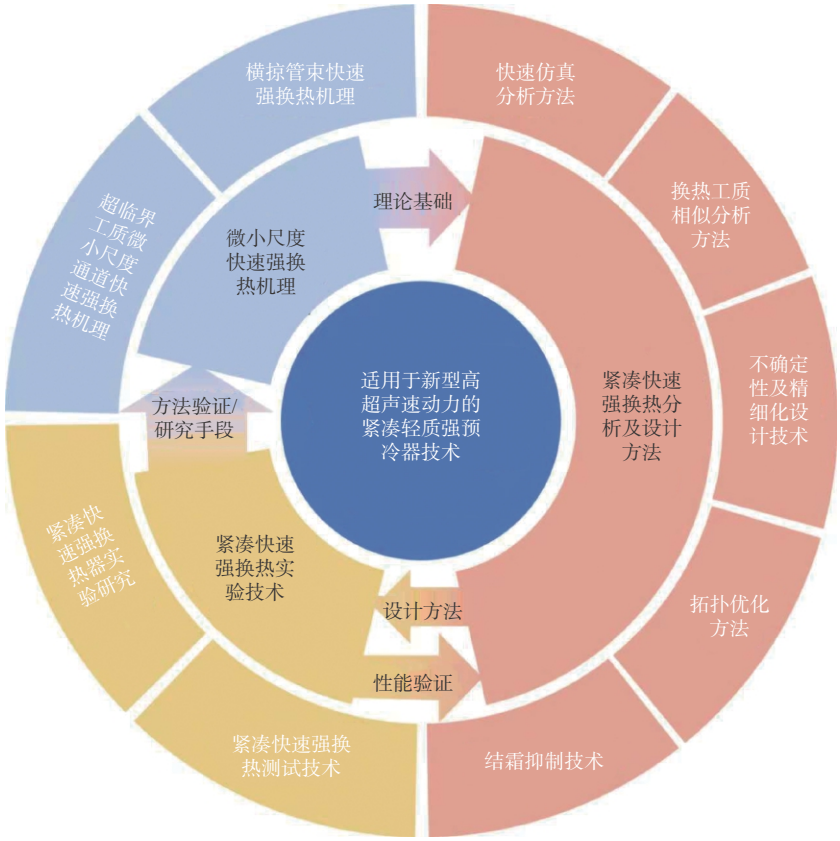


图 1　本文内容框架

Fig. 1　Outline of this article

2　微小尺度快速强换热机理

适用于强预冷发动机的紧凑快速强换热技术主要涉及 2 类换热器：针对进气预冷的微细管阵列结构强预冷器和针对发动机内部能量传递的印刷电路板结构强换热器（printed circuit heat exchanger, PCHE），主要形式及特征如图 2 所示。内部基本单元工作状态可归结为微小尺度流动换热这一物理过程。具体到强预冷发动机，除利用微小尺度通道以提升有限空间内换热能力，还涉及采用超临界流体作为换热工质以利用其高密度、高换热系数及低黏度等特性实现低阻强换热[39]。因此，根据紧凑快速强换热器主要形式，其机理主要涉及超临界工质微小尺度通道及空气横掠管束流动换热领域，本节将分类综述其研究进展。

2.1　超临界工质微小尺度通道快速强换热机理

根据 Kandlikar 和 Grande[40] 对通道尺度定义的准则（见表 1），强预冷发动机内微小尺度通道（微细管约 0.8～1.0 mm，蚀刻通道约 50～200 μm）主要集中于微小及微通道研究范畴。微小尺度通道带来的尺度效应是常规尺度经典流动和换热规律在微小尺度流动换热分析中出现偏差的重要原因，从表 2 中可以看出，随着尺度降低，表面力相对作用增强而体积力作用效果减弱，造成了与常规尺度不同的流动换热特性。而在强预冷发动机紧凑快速强换热技术中，由尺度效应引起的共性问题主要包括表面粗糙度和固体轴向导热问题。此外，强预冷发动机中的换热过程工质主要包括处于远离临界点的超临界氦及超临界煤油、存在跨临界过程（如液氢、液态甲烷等）的超临界压力工质[41]，这类工质的物性变化带来了与热力状态强耦合的换热强化/恶化及湍流结构耦合换热等问题。

在表面粗糙度影响方面，常规尺度表现为水力光滑的通道绝对粗糙度对于微小尺度通道而言则不可忽略，现有研究工作认为表面粗糙度对微小通道流动换热问题的影响主要集中在转捩、流阻和换热 3 方面。首先，Wang[42]、Yang[43-44] 和 Lin[45] 等通过大量实验指出，微小尺度通道内的转捩、流阻、换热受尺度降低和表面粗糙度共同作用（见图 3(a)、图 3(b)），其中，Lin 等[45] 的实验研究表明，较大的相对粗糙度造成了明显的提前转捩及流阻和换热的增强（见图 3(c)）。其次，为探究微小尺度下粗糙度对流动及换热的贡献，Kleinstreuer[46]、Guo[47] 及 Kadivar[48] 等分别提出了通过高斯函数和分形几何等方式生成表面粗糙度，通过微观流动分析指出粗糙表面产生的流动再循环及分离是粗糙

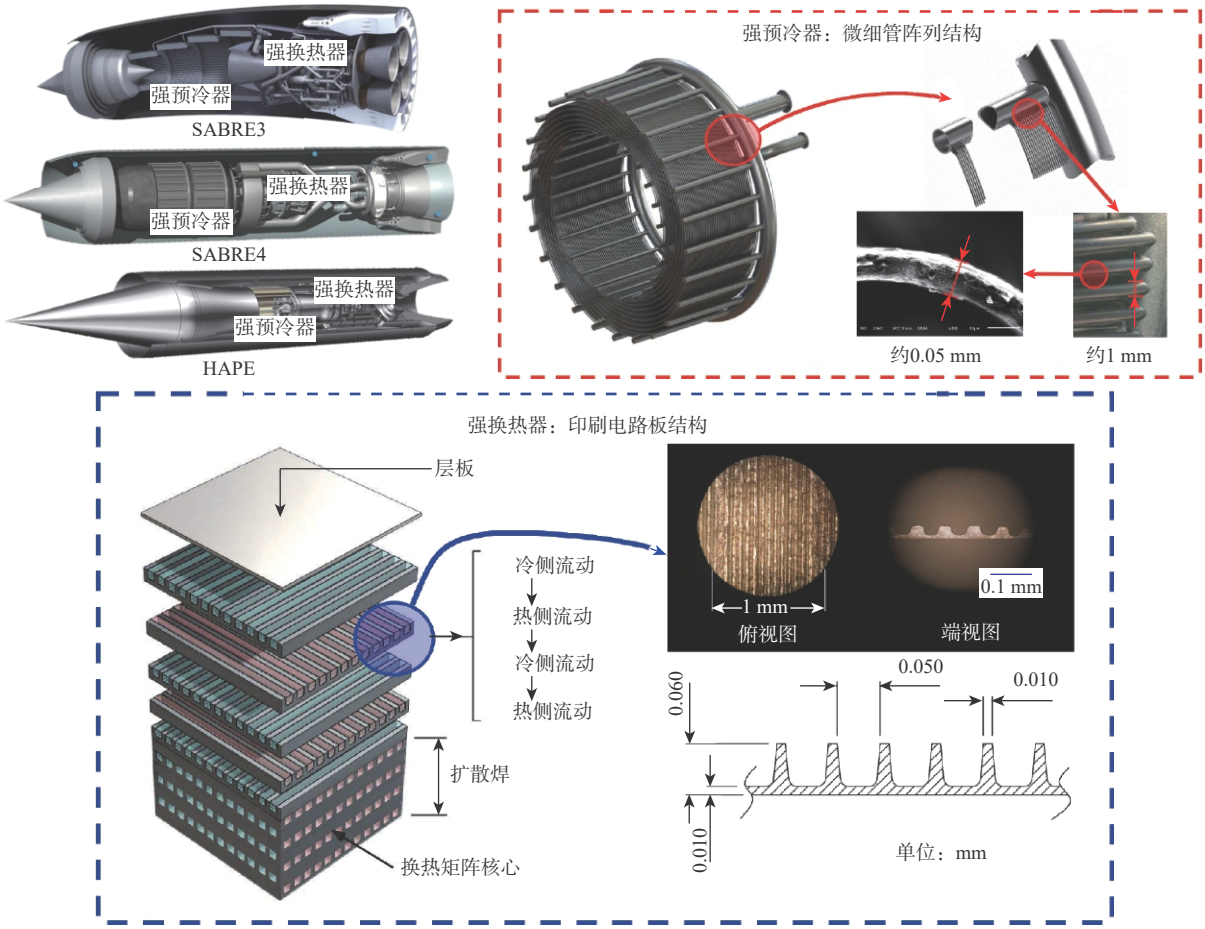


图 2 强预冷发动机中典型换热器形式

Fig. 2 Typical heat exchanger configurations in pre-cooling engines

表 1 Kandlikar 和 Grande 对通道尺度的定义[40]

Table 1 Kandlikar and Grande's definition of channel scale[40]

通道尺度类型	水力直径/μm
过渡尺度通道	0.1～10
微尺度通道	10～200
微小尺度通道	200～3 000
常规尺度通道	>3 000

表 2 不同作用力与尺度关系

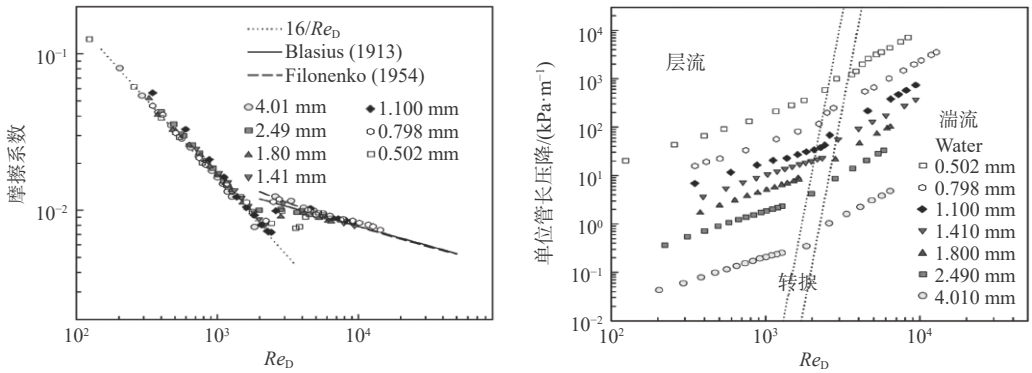
Table 2 Relationship between different forces and scales

作用力	尺度	作用力	尺度
电磁力	$\sim L^4$	惯性力	$\sim L^2$
离心力	$\sim L^4$	黏性力	$\sim L^1$
重力	$\sim L^3$	表面张力	$\sim L^1$
浮升力	$\sim L^3$	静电力	$\sim L^{-2}$

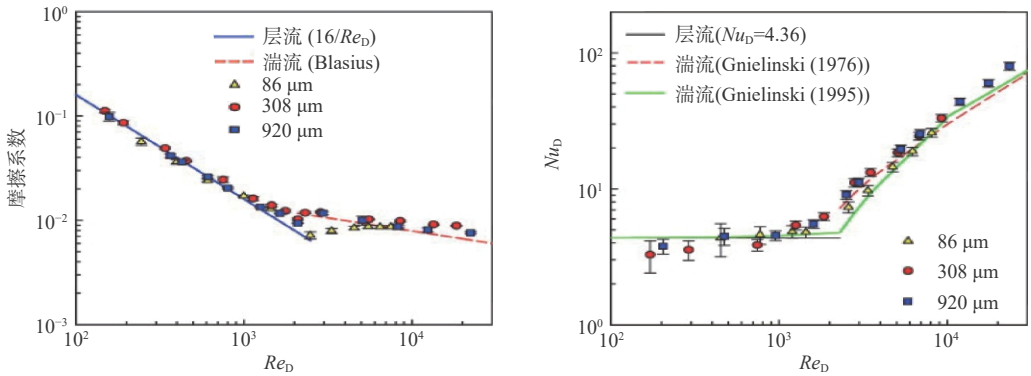
度对微小尺度流动阻力的重要影响因素。然而，在指导微小尺度流动设计的摩擦阻力方面，研究者开展了大量实验研究但仍存在争议[49-50]。Mala[51]、Li[52]等以微细光滑石英管和粗糙不锈钢管为对象开展了摩擦规律实验研究，表明微尺度通道的摩擦因子明显高于传统理论，但 Qu[53]、Sara[54] 等则指出传统管流理论可以有效预测 200～400 μm 级圆管内摩擦损失。因此，在强预冷发动机微小尺度管内流动分析方面，需仔细甄别规律适用条件及范围。

在固体轴向导热影响方面，在微小通道内部，该项因素对流动换热的影响占比增加。Maranzana[55]和 Kroeger[56] 等提出了无量纲参数 M 和无量纲导热系数 λ 用以衡量轴向导热影响，当此类判别参数达到阈值时，则认为轴向导热影响不可被忽略。Lelea[57]及 Moharana[58] 等的研究指出，当轴向导热产生较大影响时，换热器壁面温度分布显示出非线性变化，并偏离理想条件的线性变化规律（见图 4）。进一步地，Quintero[59]、Aminuddin[60]、Nellis[61] 等建立了换热单元内部轴向导热模型，针对轴向导热对换热器设计的影响及修正方法开展了广泛深入的研究工作，Baek 等[62] 则提出了降低轴向导热影响的调控及设计手段。

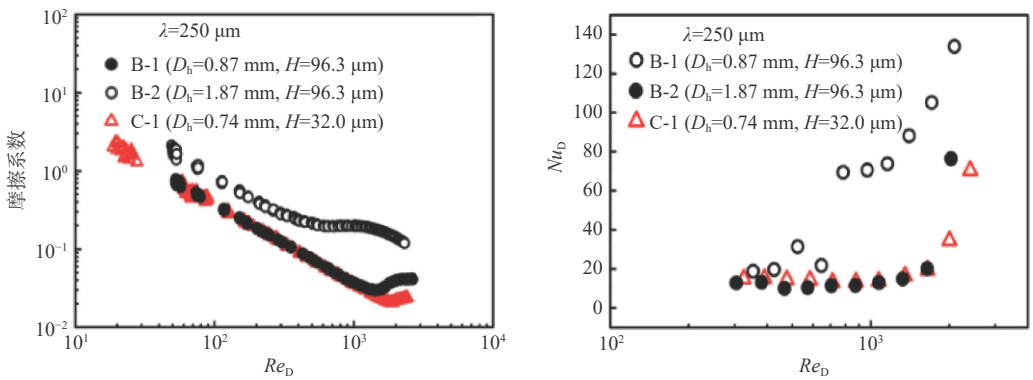
除上述由尺度效应带来的共性问题外，在强预冷发动机中存在强热流密度梯度的条件下，超临界压力换热介质（氢、甲烷等）跨临界时热物性的快速变化对流动换热带来了复杂影响，主要表现为换热强化和换热恶化 2 种形式，其发生一般是相伴的，



(a) 不同管径下水的摩擦系数及转捩雷诺数[43]



(b) 不同管径下空气的摩擦系数及*Nu*数[44]



(c) 相对粗糙度和管径对水的摩擦系数及*Nu*数的联合影响[45]

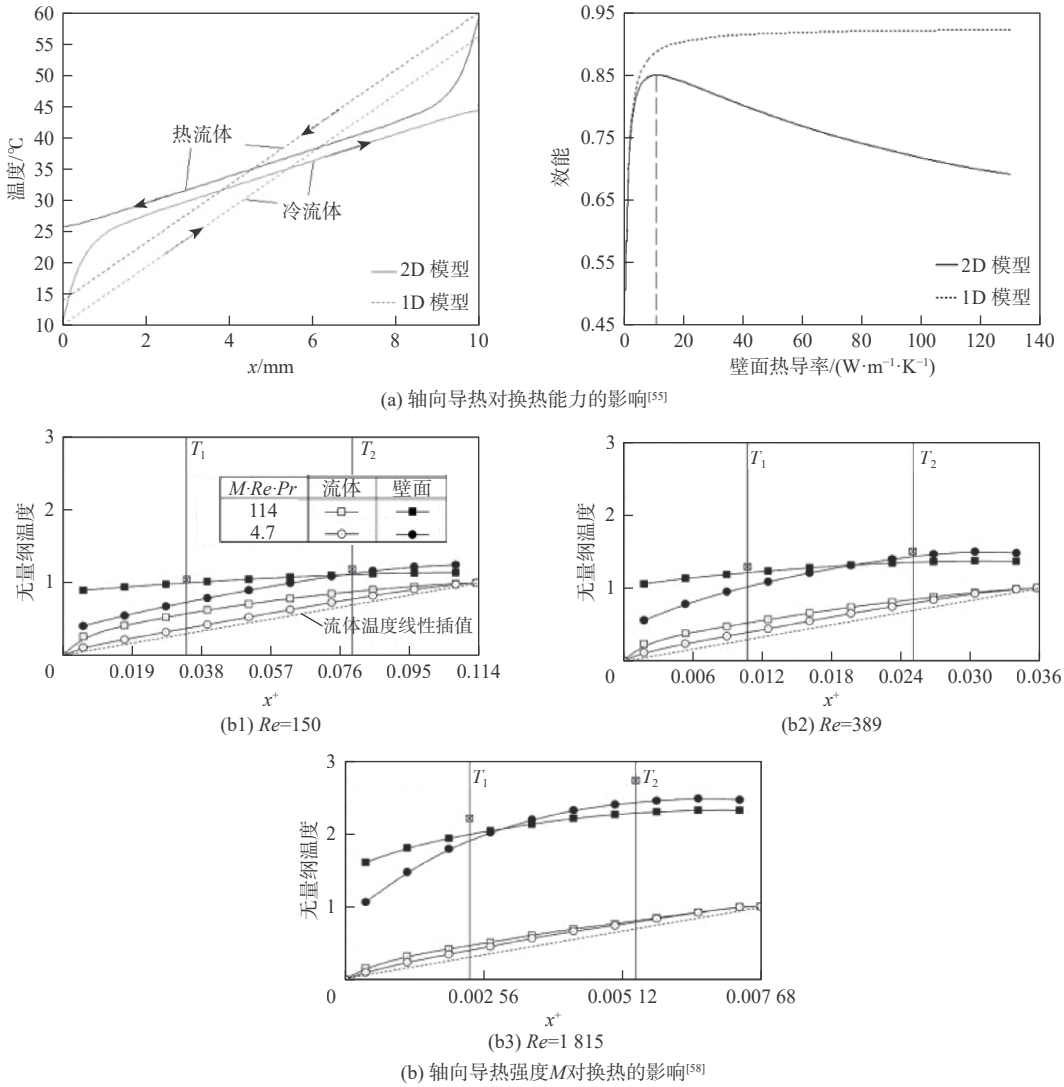
图 3　管径和表面粗糙度对管内流动的影响

Fig. 3　Effect of pipe diameter and surface roughness on flow within pipes

且部分情况下会从换热强化转变至换热恶化, 严重影响了强换热器可靠稳定工作的能力[63-64]。

在微小尺度管内换热恶化/强化问题方面, Dang[65]、Wang[66] 等以超临界 CO_2 为介质开展了微小尺度圆管流动换热实验研究, 结果表明, 在拟临界点附近换热获得强化, 且由于超临界流体的物性参数在流向及截面方向均存在较强梯度, 强变热物性、浮升力及其作用方向显著影响其对流换热能力。同样地, Gu[67]、Votta[68]、李辉[38] 等也在航空航天背景下以超临界压力低温甲烷为对象通过实验研究发现了强化换热现象, 并针对工程应用提出了修正方法。超临界压力下由流体浮升力和热膨胀加速主导的流动结构改变引起换热恶化是这一领域的一种主流学说[69-70]。Hall 和 Jackson[69] 为恶化机制提供了一种理论解释, 表明主导换热恶化的因素是管道剪切力应力分布及湍流结构的改变; 同样, Zhu 等[71] 指出换热恶化中浮升力影响湍流动能强度是影响壁面附近热量传输的重要因素; Liu 等[72] 则通过直接数值模拟(direct numerical simulation, DNS)手段研究指出换热过程中近壁面热加速效应对于湍流产生过程的抑制作用是换热恶化的原因之一。

此外, 在湍流结构耦合换热影响方面, 微小通道内部湍流结构耦合换热机理的核心在于强换热条件下近壁复杂湍流结构的改变, 因此, 理解近壁湍流结构的形成原因、演化规律及相互间的作用影响, 对于揭示微小尺度通道内超临界工质湍流流动与换热机理尤为重要。多项研究[73-74] 揭示了壁面加热对



(a) 轴向导热对换热能力的影响[55]

(b) 轴向导热强度M对换热的影响[58]

图 4 轴向导热对换热器性能的综合影响

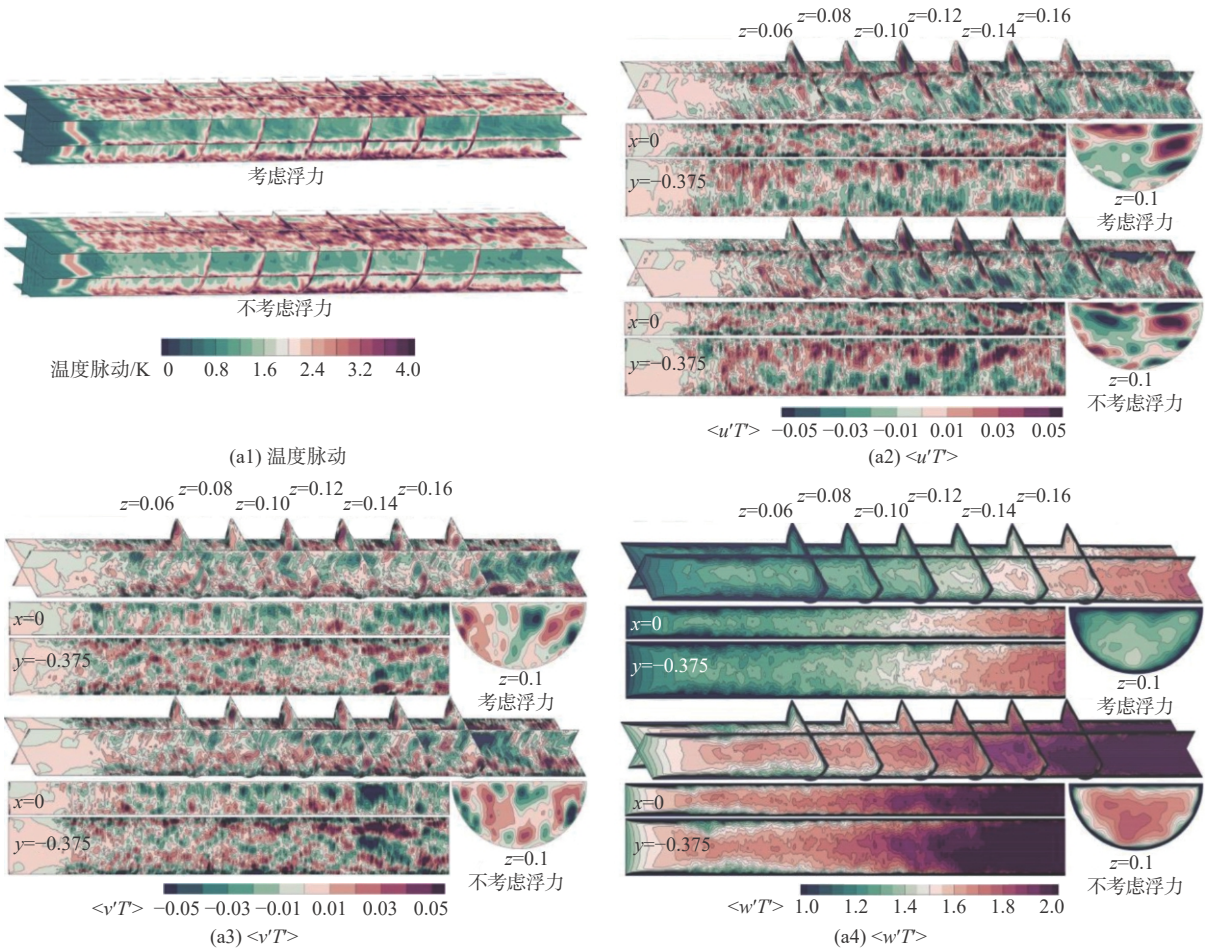
Fig. 4 Overall impact of axial heat transfer on heat exchanger performance

黏度随温度变化的层流剪切流线性稳定性的影响，进一步阐明了加热对转捩的影响。另外，关于湍流结构与换热耦合的研究聚焦于浮力和/或流向加速度的影响上(见图 5)，因为这些因素将导致湍流强度和湍流混合的衰减。例如，Kurganov 等[75]将换热恶化与速度和剪切应力分布的变化相关联，指出这是受负压梯度流动中浮力的影响；Sun[76]、Cao[77]等通过三维大涡模拟详细分析了浮力对亚临界和超临界流体换热的影响机制，证明了浮力能显著抑制由不可逆黏性耗散导致的能量损失，且削弱了湍流热传输，而忽略浮力将高估超临界流体换热能力；Cao 等[77]指出在浮力和热加速度作用下，湍流流动和换热受浮力和温度加速度主导。此外，针对发动机真实振动环境下微小尺度通道内部超临界工质强化流动换热问题，Zou 等[78]通过理论结合仿真分析手段，揭示了振动惯性生成项主导的特征旋涡结构和旋涡倾斜导致的湍流二次流是周期性振动激励促进超临界工质管流换热性能强化的关键机理，进一步发现了强密度分层的管流中通过施加横向振动激励来维持第三类普朗特二次流的新机制，可为强预冷发动机内部紧凑快速强换热器构建湍流换热提供的新思路。

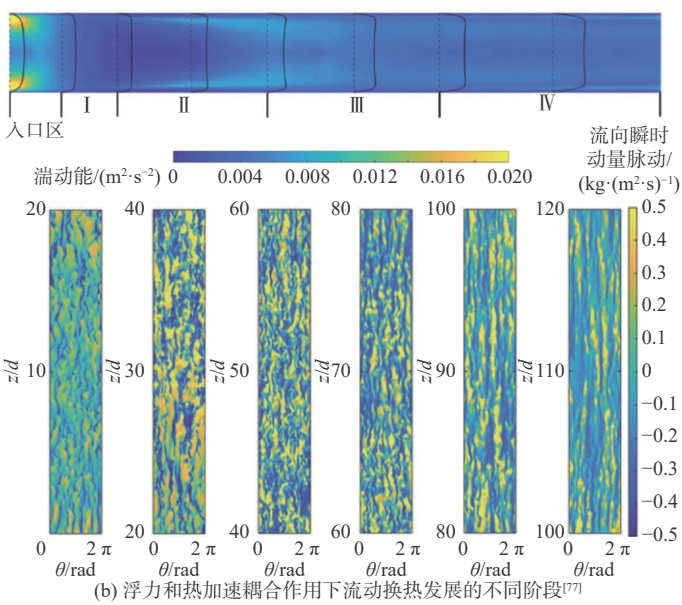
2.2 横掠管束快速强换热机理

对于微小尺度阵列管束形式强预冷器而言，空气通常以横掠管束的形式与管内冷却剂换热。深入理解壁面换热环境下圆柱管束的基本换热单元复杂流场中流动结构的演化规律、掌握流动结构与壁面换热间的相互作用机制，对于紧凑快速强换热技术研究具有重要意义。

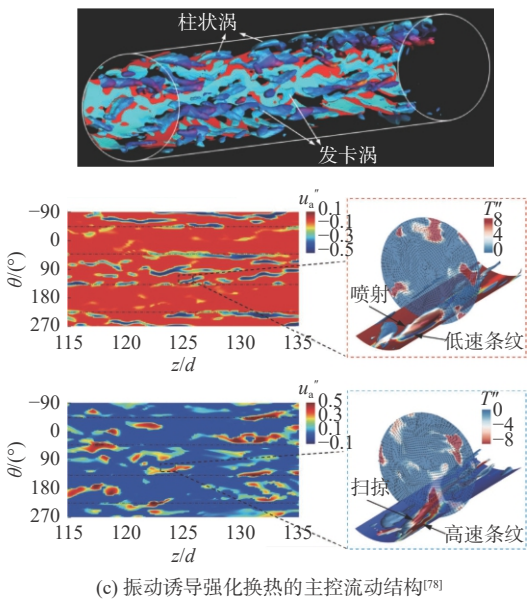
首先，针对叉排管束流动发展与演化现象，大量研究表明，当流向叉排管排数达到一定数量后，流动换热将进入充分发展阶段。在前排管束区能够观察到比较明显的较大尺度流向涡，而后排流场中的旋涡结构尺度变小，如图 6 所示[79]。Yoo 等[80]针对 6 排叉排管束的常温实验表明，第 2 排管和第 3 排管的平均 Nu 数分别比第 1 排管高出 30% 以上



(a) 考虑和不考虑浮力时的温度脉动和湍流热通量分布[76]



(b) 浮力和热加速耦合作用下流动换热发展的不同阶段[77]



(c) 振动诱导强化换热的主控流动结构[78]

图 5　湍流结构和换热耦合研究

Fig. 5　Study on coupling of turbulent structures and heat transfer

和 65% 以上, 前 3 排管管壁局部换热系数的周向分布规律不同, 但此后流动充分发展, 分布规律相似。Webber 和 Taylor[81] 在 SABRE 飞行包线中强预冷器的来流 *Re* 数范围(1 016～5 043)内, 针对 21 排叉排管矩阵开展实验, 结果表明, 叉排管束流向上第 16 排时的流动已经达到 95% 的充分发展程度[81], 因此, 第 16 排及以后的圆管换热特性可以代表强预冷器管束中的大多数管排(流向深度超过 150 排)。

在流动结构与壁面换热相互作用机制方面, 影

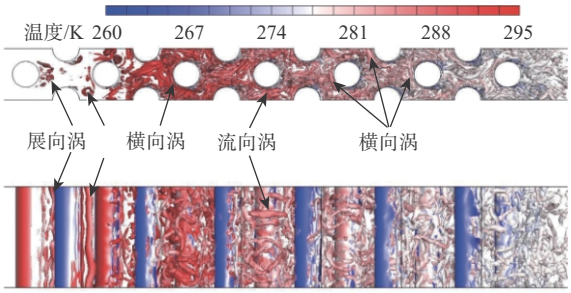


图6 某瞬态叉排圆柱流场中旋涡结构[79]

Fig. 6 Vortex structures in a transient flow field around a cylindrical fork[79]

响壁面换热变化的流动结构包括壁湍流相干结构和自由流湍流流动结构，研究学者对此开展了广泛的研究工作[82]。首先，相干结构在近壁区湍流动量输运现象中扮演重要角色[83]，Kong等[84]指出，在等壁温和等热流2种不同的壁面边界条件下，近壁区温度分布与流动结构分布均具有较强的相似性。具体而言，相干结构主要以成对出现反向旋转的流向涡对的形式存在[85-86]，在壁面被加热的情况下，流体下扫区外侧低温流体受到流向涡的诱导流向壁面，使壁面换热能力加强，出现低温条带，而在流体上扬区，由于流体在抵达此区域前已经和壁面换热而升温，导致壁面出现高温条带。与壁湍流相干结构相似，Hubble等[87]指出自由流湍流来流对壁面滞止区换热强化的物理机制也是旋转方向相反的涡对诱导引发的流体上扬和下扫，如图7所示。然而，上述2类不同尺度的流动结构对壁面换热能力的影响程度不同，存在主控流动结构，根据已有研究结果，壁面附近的速度和热流密度时序信号的总体变化趋势十分相似，即使信号各自的能谱均不存在典型的峰值频率，但二者的互相干系数在低频时具有较大值[88-89]，这表明影响壁面换热的主控流动结构主要为低频大尺度流动结构。而对于单管圆柱绕流，影响圆柱后侧壁面对流换热能力的主控流动结构为圆柱后方脱落的卡门涡及其诱导的二次涡[79]。

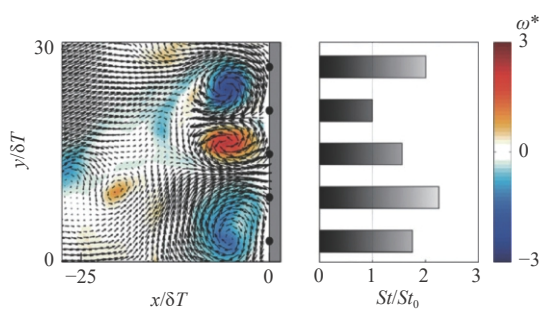


图7 上扬和下扫对壁面换热的影响[87]

Fig. 7 Effect of upward and downward sweeps on heat transfer across a wall[87]

其次，壁面换热对流动结构也会产生重要影响。在壁面绝热时，基于尾迹脱落频率、圆柱后方回流区长度、近尾迹区雷诺应力等圆柱绕流特征参数随来流 *Re* 数的变化规律，可将不可压缩圆柱绕流区分为7种流型[90]，而壁面换热会导致这些特征参数发生显著变化，并改变区分各流型的关键 *Re* 数[91-92]。在热壁面条件下，壁温越大，空气与壁面换热温差越大，相同 *Re* 数下的 *St* 数降低，也即涡脱落频率降低，区分各流型的关键 *Re* 数增大[93-94]，如图8所示。图中：T^*为壁温与自由来流温度比值。当壁温达到某一临界温度后，换热温差的增大导致 *Gr* 数增大，浮升力影响增强，脱落频率进一步大幅降低。随着壁温进一步提高，尾迹摆动及相应的脉动速度幅值减小，直到变为0[95]，针对适用于强预冷器问题的冷壁面被外流空气加热问题，北航团队[79]研究表明，影响流动熵产和换热熵产的主控流动结构均为相对于卡门涡脱落频率更高的小尺度结构，其中心频率约为卡门涡特征频率的7倍。此外，管壁壁温降低时，单管圆柱下游流体二阶统计量（黏性应力、雷诺应力和湍动能）均变大，分离剪切层的K-H不稳定性增强，回流区边界处湍流脉动对流向动量的横向输运能力变强，回流区长度变短。

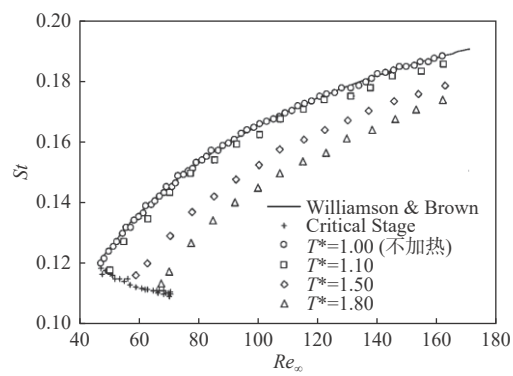
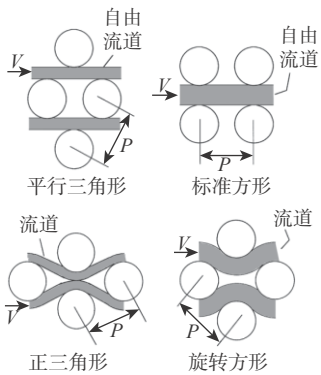


图8 不同壁温条件下 *St* 数随 *Re* 数的变化趋势[93]

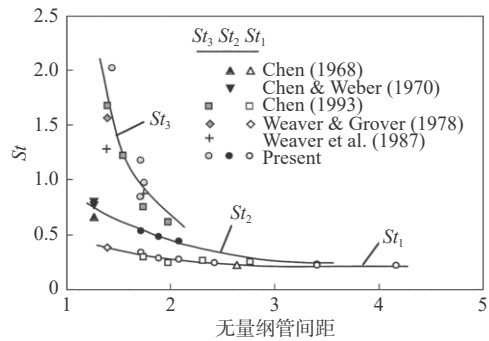
Fig. 8 Trends in the *St* number as a function of the Reynolds number under different wall temperature conditions[93]

而涡脱落频率与激振/共振也会在强预冷器这种圆管叉排阵列结构内产生显著影响。首先，这种排列方式在相同 *Re* 数条件下较顺排排列方式具备更佳的换热性能，但同时也会带来更大的流动损失[96-97]。Kim等[98]针对叉排管束的实验结果表明，当涡脱落频率接近流动结构的一阶固有频率并引发声学共振后，测得的流动压降明显高于Žukauskas经验关联式预测结果。进一步地，陈一鸣等[99]的大涡模拟结果表明，前排管束的流动结构为有规律的剪切层运动和尾涡脱落，管壁瞬时 *Nu* 数具有较为固定的波动频率，而后排管束的流动结构为无规律

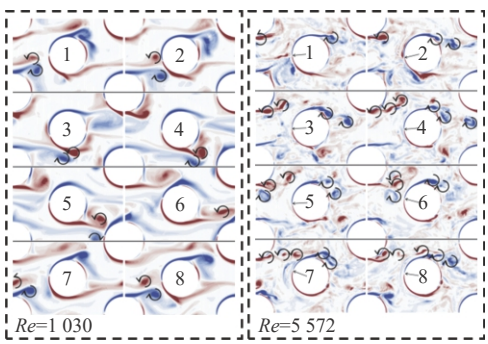
的小尺度旋涡结构, 流场表现出无序性, 无固定波动频率。Ziada 和 Oeng[100] 在正三角形叉排排列、常规尺度管束的实验研究中观察到了 3 个不同的涡脱落频率分量。Blevins[101] 归纳了不同文献中不同叉排排列方式的研究结果, 指出大部分研究均观察到了多个涡脱落频率, 并分析和归纳了平行三角形排列方式的管间距对 3 个涡脱落频率的影响, 如图 9(a)、(b) 所示; Linton 和 Thornber[102] 以 SABRE 强预冷器为背景, 采用 DNS 方法对管径 0.8 mm 的小尺度叉排管束开展研究, 结果表明, 当基于水力直径的雷诺数增长到约 2 500 时, 低频 *St* 数(St_1)开始出现, 此时的周期性频率特征不再完全由不同圆管尾迹之间的相互作用引起。低 *Re* 数下, 高频 *St* 数(St_2)由上游圆管涡流撞击引起的涡交替脱落产生, 与单管情况下由尾迹涡的相互作用引起所不同。高 *Re* 数下, 单管流动中导致剪切层破裂的不



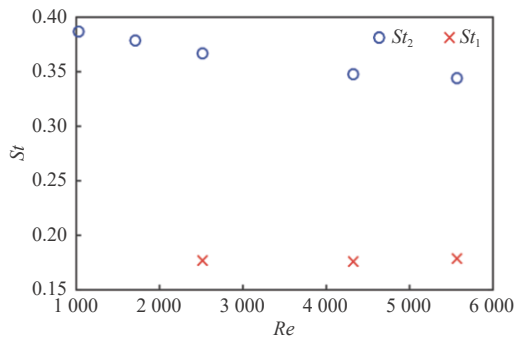
(a) 不同叉排排列方式[101]



(b) 平行三角形排列方式管间距对3个涡脱落频率的影响[101]



(c) 一个涡脱落周期内的展向涡量云图[102]



(d) St_1、St_2和*Re*数的关系[102]

图 9　不同排列方式及雷诺数对涡脱落频率的影响

Fig. 9　Effect of different arrangements and Reynolds numbers on frequency of vortex detachment

稳定性也在叉排管束剪切层中起作用, St_2 与剪切层高度活跃有关, St_1 与驻点的振荡有关, 如图 9(c)、图 9(d) 所示。一方面, 圆管的涡脱落会通过激振或共振影响换热器使用寿命, 甚至达到破坏性强度, 导致工作突然失效[101,103]; 另一方面, 北航团队[104] 的研究还表明, 来流压力变化频率达到流场固有频率时, 也将引起共振并导致换热性能发生剧烈振荡。为避免上述后果, 有效的解决办法是增加翅片或挡板, 提高流动结构的固有频率[98]。

此外, 虽然标准的叉排管束换热器已得到广泛应用, 但仍存在一定的换热性能提升空间, 也即圆管下游背风侧 *Nu* 数远小于迎风侧[105], 可进一步提高, 如图 10 所示。因此, 许多研究通过增加管周或管外翅片/涡流发生器[106-107]、增大换热表面粗糙度[108]、采用凸轮形管设计[109] 等方式, 来强化叉

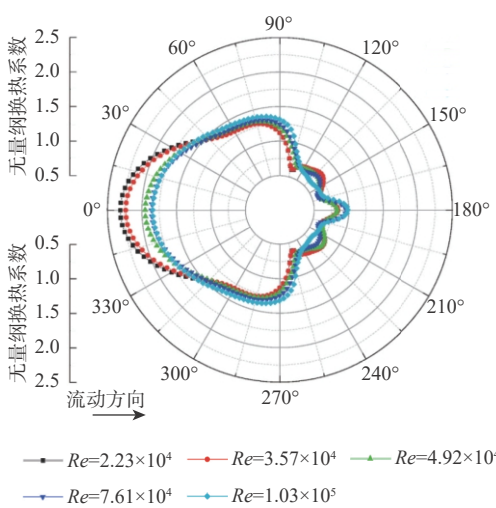


图 10　管壁无量纲换热系数周向分布[105]

Fig. 10　Circumferential distribution of dimensionless heat transfer coefficient on pipe wall[105]

排管束换热性能。例如，涡流发生器诱导的纵向涡可减小圆管下游换热恶化区或尾迹区，从而提高换热速率[97]。

3 紧凑快速强换热分析及设计方法

强预冷发动机的紧凑快速强换热器恶劣的工作环境和严苛的性能需求对其设计能力提出了极高要求。传统民用换热器设计方法在综合考虑极端工作环境、强物性变化、复杂非均匀流动换热方面，以及提升全域性能指标及鲁棒性等方面有所不足。本节将针对紧凑快速强换热器的气动热力分析及设计方法进行综述。

3.1 全尺寸强换热器快速仿真分析方法

强预冷发动机作为宽域动力系统，其流量范围动辄达到百公斤量级，致使强预冷器等换热器部件尺度最大将达到数十厘米至1 m以上，然而从图11可以看出，其基本换热单元在数十至数百微米级别，巨大的尺度差异致使当前虽然可以对基本换热单元单独开展仿真分析，但现有仿真能力难以完成真实发动机尺度全尺寸包含数以万计微细管束的强预冷器的建模仿真，因此，发展具备较高保真度的快速仿真计算方法，对于指导和校验紧凑快速强换热器设计具有重要意义。

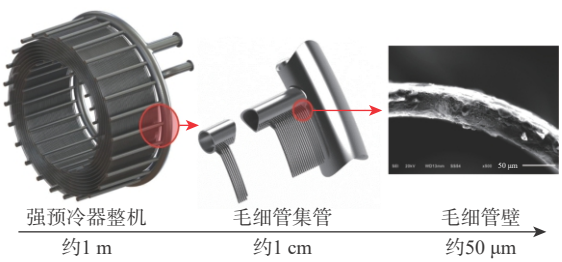


图11 强预冷器跨尺度特征[110]

Fig. 11 Cross-scale characteristics of pre-coolers[110]

近年来，基于虚拟源项模型的快速仿真分析方法被众多研究人员采纳并用于强预冷发动机中紧凑快速强换热器的研究，此类方法采用均匀结构网格替代复杂管束矩阵，并用数值预测替代实际物理作用，具有网格划分简单、计算速度快等显著优势，已在各类多管束或多通道换热器中得到广泛应用。具体而言，虚拟源项模型可细分为多孔介质模型、双单元换热模型、分布式体积热源法、均匀体积热源法等，仿真计算时，均只需对换热器所占空间划分均匀结构网格并附加动量源项和能量源项，简化示意图如图12所示。

其中，多孔介质模型是普遍认可的压降模拟准则，该模型将强预冷器的叉排管矩阵类比为多孔介质，通过经典的Darcy-Forchheimer方程[111]来描述流体流经多孔介质后的压降ΔP与平均流速

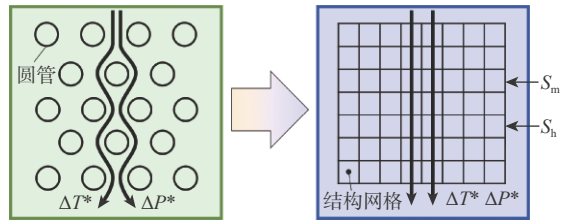


图12 虚拟源项简化示意图

Fig. 12 Schematic diagram of simplification of virtual source terms

U的近似关系。对于各向同性情况，其基本表达式如下：

$$\frac{\Delta P}{\Delta n}=C_1\mu U+C_2\frac{\rho U^2}{2} \tag{1}$$

式中：Δn为流向长度；μ为黏度；ρ为流体密度；C_1和C_2分别为黏性阻力系数和惯性阻力系数，通常以拟合实验数据、仿真数据或低维模型数据的方式获取。

将式(1)以动量源项$S_m=-\Delta P/\Delta n$的形式加入动量方程，即可模拟强预冷器对空气的流动阻力。考虑到高雷诺数下黏性项$C_1\mu U$的贡献较小[112]，Chung等[113]在PCHE强换热器的压降模拟中选择忽略黏性项。此外，考虑到强预冷器在轴向、周向和径向上具有较高的各向异性，Wei等[114]采用了各向异性多孔介质模型来评估空气总压损失。

同时，多孔介质模型还提供了热平衡和非热平衡2种状态下的换热模拟方法。对于适用性更强的非热平衡状态，需定义一个在空间上与多孔介质流体区重合的固体区，流体区和固体区仅在换热方面存在相互作用，简化的能量方程分别如式(2)和式(3)所示[115]：

$$\frac{\partial}{\partial t}(\gamma\rho_{\mathrm{f}}E_{\mathrm{f}})+\nabla\cdot\left[\boldsymbol{v}(\rho_{\mathrm{f}}E_{\mathrm{f}})\right]=\nabla\cdot(\gamma k_{\mathrm{f}}\nabla T)_{\mathrm{f}}+S_{h,\mathrm{f}} \tag{2}$$

$$\frac{\partial}{\partial t}\left[(1-\gamma)\rho_{\mathrm{s}}E_{\mathrm{s}}\right]=\nabla\cdot\left[(1-\gamma)k_{\mathrm{s}}\nabla T_{\mathrm{s}}\right]+S_{h,\mathrm{s}} \tag{3}$$

式中：下标f和s分别表示流体区和固体区；γ为孔隙率；E为总能量；k为导热系数；$\boldsymbol{v}$为速度矢量；$S_{h,\mathrm{f}}=h_{\mathrm{fs}}A_{\mathrm{fs}}(T_{\mathrm{s}}-T_{\mathrm{f}})$和$S_{h,\mathrm{s}}=-S_{h,\mathrm{f}}$分别为流体域和固体域的对流换热能量源项，$h_{\mathrm{fs}}$为对流换热系数，$A_{\mathrm{fs}}$为多孔介质面积密度。

一些学者在使用多孔介质模型开展研究时，对其适用性有更深入的思考和发展。Chen等[116]的研究表明，非热平衡态多孔介质模型可以准确预测具有大量微通道的PCHE强换热器的流动换热性能，而随着微通道数量减少引起流动不均匀性增加，预测相对误差会显著增加。Wu等[117]则指出上述两能量方程多孔介质模型无法反映管内流体的非均匀换热，因此，引入第三能量方程来计算固体内部

换热过程。在强预冷器这一应用背景下, 第三能量方程也即代表了内部冷却剂的换热过程。

此外, Ding 等[118]将双单元换热模型引入强预冷器换热模拟。热传递通过换热速率或换热单元数(number of transfer units, NTU)的形式求解[119]。Wang 等[120]采用分布式体积热源法实现强预冷器空气换热模拟, 基于强预冷器低维模型求解预冷区不同位置的目标温度, 并直接以体积热源$S_{h,\mathrm{cell}}$的形式加入各网格单元能量方程。深入分析此类简化仿真手段, 其核心难点是模型简化过度或验证说服力不足致使其可靠性缺乏有效保证。为此, 北航团队[121]进一步提出了一种可考虑空气轴向和径向非均匀性的分布式体积热源法。耦合流道的高维验证表明, 强预冷器空气温降及总压损失与真实模型的相对误差均在±1% 以内, 而网格量较真实模型降幅达 95% 以上。此外, 将该方法应用于燃料直接预冷发动机总体性能计算后发现, 强预冷器低维模型低估了空气总压损失及燃料预冷用量, 进而分别导致发动机净推力和燃烧当量比出现更大计算偏差。

$$S_{h,\mathrm{cell}}=\frac{\dot{m}_{\mathrm{cell}}c_p\left(T_{\mathrm{aim}}-T_{\mathrm{loc}}\right)}{V_{\mathrm{cell}}} \tag{4}$$

式中: $S_{h,\mathrm{cell}}$为网格体积热源项; $\dot{m}_{\mathrm{cell}}$为网格内质量流量; c_p为定压比热容; T为单元内温度, 下标 aim 和 loc 分别表示预测的目标数值和网格单元局部数值; V_{cell}为网格单元体积。

上述几种换热模拟方法均对能量源项或体积热源进行精细求解, 具有较好的物理可解释性。与此同时, 还有文献 [122-123] 以探讨强预冷器对进气道性能的影响为研究目标, 故仅通过给定整体的均匀体积热源$S_h=Q/V$实现基础换热模拟。表 3 归纳总结了上述所有换热器简化换热模拟方法的示意图和源项求解特点。

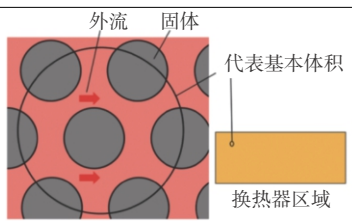
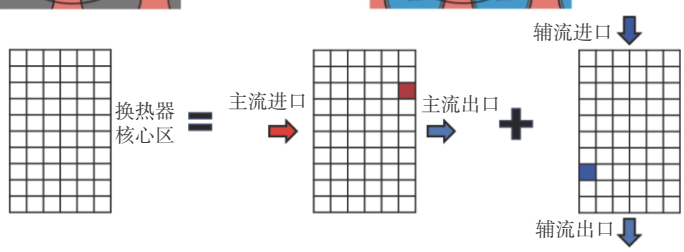
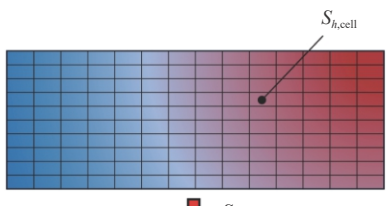
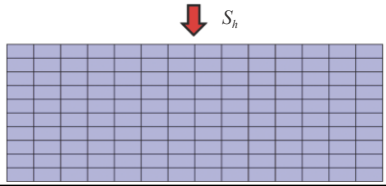
对于强预冷器的真实工程设计及分析过程来讲, 全部采用真实几何开展全尺寸仿真研究将受限于计算仿真能力, 短期内难以实现, 采用简化分析手段将成为主要手段之一。综上研究可以看出, 采用简化模型的核心在于保证真实流动换热模拟的高保真度, 因此, 在模型构建过程中应注重基于物理过程的模拟及合理化简化, 而在工程实践中还应注重与实验相结合的模型修正与完善。

3.2　强变物性换热工质相似分析方法

强预冷发动机内部紧凑快速强换热器工质常为以超临界氢、甲烷、煤油为代表的易燃易爆且强变物性工质, 在研究初期采用此类工质开展验证存在安全风险。为此, 研究人员常在研究初期基于安

表 3　换热器简化换热模拟方法总结

Table 3　Summary of simplified heat transfer simulation methods for heat exchanger

计算方法	示意图	源项特点
多孔介质模型		通过对流换热系数计算, 且流体和固体物性通过孔隙率加权
双单元换热模型		通过换热速率或NTU计算
分布式体积热源法		通过非均匀体积热源计算
均匀体积热源法		通过均匀整体体积热源计算

全的换热介质开展实验研究以降低实验难度，并采用相似分析方法得到真实换热介质流动换热特性。传统利用 *Re* 数和 *Pr* 数的相似原理在经典换热研究中已经得到广泛应用[124]，但上述强预冷常用工质物性沿流向及径向均存在较强梯度，且存在换热强化或恶化等换热异常现象，使现有基于常物性假设的对流换热相似准则不再适用，为此，研究人员针对这类强变物性工质在截面内的换热相似开展了针对性研究[125-127]。Pioro 等[128]指出可利用无量纲压力、无量纲温度及 *Re* 数实现工质间相似变换；Cheng 等[129]为提升相似准确性，将控制方程无量纲化结合保证热物性相似进而改进了相似准则；Zahlan 等[130]则基于超临界流体实验数据对无量纲参数幂进行了修正。而针对换热通道内流动相似问题，Ambrosini[131]提出了基于无量纲比焓的跨拟临界数和亚拟临界数方法，可以保证不同超临界流体的无量纲壁温比焓分布均具有较好的一致性，但 *Nu* 数存在一定差异。李辉[132]基于几何相似、热物性相似、动力相似和边界条件相似，围绕强预冷发动机常用燃料换热工质开展了强变物性影响下的对流换热相似方法研究，建立了可考虑强物性变化影响的对流换热相似准则及相似变换流程，验证表明该方法可保证整个通道内相同无量纲比焓截面处无量纲温度场及无量纲速度场相似，在验证工况中，2 种流体 *Nu* 数的相对偏差均小于±20%，为降低强预冷发动机内紧凑快速强换热研究实验难度及安全风险提供了重要方法支撑。

3.3 考虑不确定性及精细化的紧凑快速强换热设计技术

强预冷发动机内紧凑快速强换热器复杂的工作环境和严苛的性能要求对设计方法提出了更严格的要求。因此，针对性发展能综合考虑其复杂流动换热状态的精细化及不确定性设计方法对于保证其在强预冷发动机内部高效工作十分重要，本节综述了考虑精细化及不确定性的紧凑快速强换热设计技术。

3.3.1 紧凑快速强换热经验关联式

高精度的换热经验关联式是紧凑快速强换热器设计的基础，如前文指出，强预冷发动机主要涉及超临界介质微小尺度通道及空气横掠管束经验关联式 2 类。

在微小尺度管内流动换热方面，由于受到强物性变化及浮升力、热加速等作用的耦合影响，目前，针对不同工质、不同运行条件尚无统一的低维预测模型。北航团队[38,132]通过实验研究给出了计及近壁面物性修正的微细圆管内超临界甲烷换热关联式。

燃料裂解吸热是近年来高超声速动力主动冷却技术的一个重要研究方向。发生化学反应时，通道内剧烈的物性变化及化学反应与流动换热的交互作用带来了不同于常规流动换热的流动换热特性，常规经验关联式不能精确预测其流动换热特性。目前，相关研究相对有限，且主要集中在碳氢燃料领域。Kimura 等[133]首先提出了描述化学吸热效应的无量纲数，从而量化由化学吸热带来的换热增强效果；Yu 等[134]则通过碳氢燃料热裂解实验提出描述化学吸热量占比的无量纲参数，并针对 Dittus-Boelter 换热关联式提出修正；Xu 等[135]同样基于碳氢燃料的热裂解反应，引入了有效比热容和有效导热系数概念，对 Jackson 关联式[136]进行了拓展和修正；张议丰[137]则基于 Kimura 等[133]的工作，同样根据 Jackson 关联式[136]进行修正，进一步提出了碳氢燃料蒸汽重整过程的流动换热经验关联式。

此外，研究人员针对超临界氦、甲烷等工质开展了大量微小尺度通道换热研究并发展了相应经验关联式，如文献 [138-151] 所示，可为紧凑快速强换热设计提供依据。

此外，对于微细管叉排阵列形式强预冷器涉及的横掠叉排管束换热经验关联，国内外研究学者开展了大量的理论及实验研究[152-162]，其中，Žukauskas[153]详尽分析了空气横掠单管及管束的局部换热特征和平均换热性能，给出了最具代表性的经验关联式。然而，Žukauskas 公式在计算具有小直径、小管间距特征的微小尺度管束平均换热系数时是否适用仍存在争议[159-161]，为此研究学者在小管径叉排构型方面开展了大量理论与实验研究。其中，Bacellar 等[162]基于 CFD 手段获得了 0.5～2.0 mm 圆管叉排管束的流动及换热经验关联式，研究表明，与 90% 原始数据对比，该关联式的误差在 20% 范围内。另外，整理了微小管束相关换热经验关联式[152-158]，可为微细管叉排阵列形式强预冷器设计提供参考。

3.3.2 紧凑快速强换热器精细化设计方法

紧凑快速强预冷器设计过程主要包括基本形式选择、换热及压降设计。本节给出了强预冷器精细化设计过程构建方法。

首先，基本形式的选择主要取决于使用环境限制、边界条件约束下对于流动换热的综合考量与折中。例如，图 2 中给出了目前被认为流动换热综合性能优秀的一种强预冷器布局形式[110,163]，该布局是综合考虑强预冷器流阻及飞行阻力的折中方案。一方面，强预冷器为减小空气压降，应具有较大的迎风面积以降低摩擦损失。另一方面，飞行器若想

减小飞行阻力, 需尽可能地降低截面积。然而, 当强预冷器与飞行器迎风面重合, 即采用轴向进、出气方案时, 虽保证了较小的迎风阻力, 但强预冷器空气侧阻力过大, 而无法接受。因此, 将强预冷器设计为径向进气、轴向出气, 空气侧迎风面积的变大仅会增加部分长度而不会增加额外截面积, 有效实现了紧凑空间内阻力的折中。

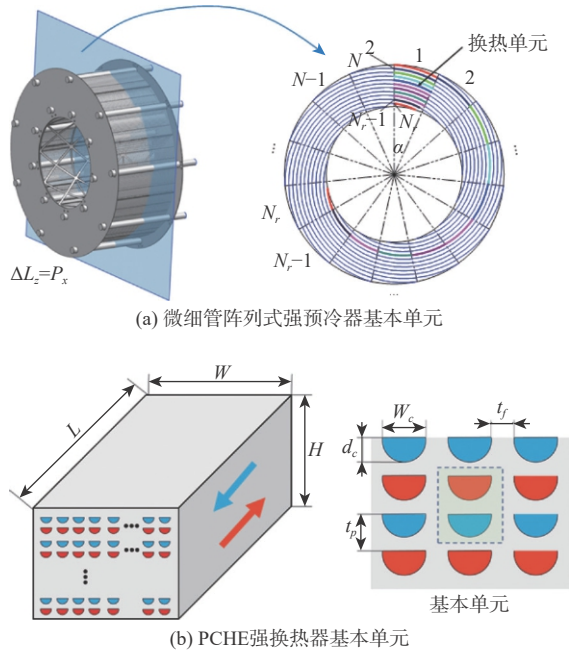
在基本形式确定的基础上, 紧凑快速强换热设计的核心是在受限空间内低阻高效换热的实现, 除上述高适用性及高准确度经验关联式作为基础外, 另一关键是准确地对强预冷器内部各位置流动换热的精确设计。为此, 北航团队[164]构建了基于多维耦合的紧凑快速强换热精细化设计方法, 耦合零维(0D)-二维(2D)各自优势, 可在较短时间内综合考虑强物性变化、强预冷器真实构型及非均匀来流等因素影响, 获得强预冷器内部温度、压力及换热系数等关键参数分布信息。

0D 方法是最为简单成熟的方法, 其将换热器视为一个整体, 利用有限温差法(logarithmic mean temperature difference, LMTD)、ε-NTU、P-NTU、ψ-P 和 P_1-P_2 等多种典型方法[124], 构建了换热功率、进出口参数及换热面积间的定量关系。对于换热介质物性变化较小、工作环境较为均匀的强换热器, 采用该方法可以在短时间内获取换热器基本性能。

而当强换热器介质物性及换热系数沿程剧烈变化时, 0D 设计方法由于其设计过程平均物性这一特点而导致无法捕捉内部物性剧烈变化下的换热特异状态, 因而换热器整体流动换热设计将存在较大偏差。Saeed 等[165]开展性能设计方法研究时发现, 整体换热评估方法会严重低估低雷诺数下 Nu 数, 最严重时可达 30% 以上。为解决这一问题, 研究人员基于离散化思想提出了一维(1D)/二维(2D)离散化设计分析方法, 如图 13 所示。

紧凑快速强换热器一般具有周期性特征, 基于此离散化最小周期单元开展设计分析, 则可在提高设计准确度的同时降低模型难度。例如, 针对 SABRE 发动机强预冷器这种典型结构, 在轴向参数均匀假定下, 轴向最小单元为保留管束叉排特征(2～3 排管束)的环形切片, 进一步该模型可沿周向划分为 N 个相同的扇形单元。若符合周向均匀性假设, 则每个扇形单元内的局部流动与换热过程完全一致, 即一根完整毛细管的 N_r 段短管可以依次拼接为单个扇形单元。

通过上述手段, 可以将强预冷器由原本的三维(3D)流动换热过程降低至径向 1D, 大幅缩减了流动换热设计难度, 设计计算流程如图 14 所示。设



(a) 微细管阵列式强预冷器基本单元

(b) PCHE强换热器基本单元

图 13　紧凑式快速强换热器基本单元

Fig. 13　Basic unit of compact fast heat exchanger

计人员在得到设计要求后(流体流量、总压损失等), 应首先假定强预冷器的宏观及细节几何参数(内外径、管间距等)。随后, 针对图 13(a) 中的 1 号离散单元, 假定冷/热流体出口热力参数以获得流体物性, 进而基于经验关联式获得初步计算结果并与猜测值对比。迭代至残差满足要求后, 通过边界条件对应关系赋予 2 号离散单元边界。重复上述流程即可获得 N_r 号离散单元的热力参数, 并与强预冷器初始设计条件对比, 迭代得到收敛解。最后, 若冷、热流体总压损失高于设计要求, 则应继续改变几何参数直至得到满足设计要求的设计方案。

强预冷发动机在真实工作时, 由于飞行状态变化、吸入外部高温气流等因素, 会不可避免地在强预冷器入口形成非均匀来流条件[120]。如图 15 所示, 为解决该问题, 北航团队[164]提出了一种考虑全周的非均匀来流的 2D 精细化设计方法, 在 Ma=4 来流条件下通过实验得到的换热功率预测误差均低于±10%(见图 15(b))[166-167], 在一定程度上验证了该方法在仿真快速计算性能方面的工程可用性。如图 15(a)所示, 该模型沿周向共划分了 N 个图 13 中的强预冷器扇形单元, 由于每个基本换热单元又包含 N_r 个径向扇形单元, 2D 离散设计模型中共包含 $N \times N_r$ 个强预冷器单元体。通过引入 2D 离散设计模型, 可捕捉进口周向不均匀条件下强预冷器内部热力参数的分布情况, 进而分析非均匀来流与换热过程的耦合作用关系。此外, 高远等[168]提出了一种针对强预冷器扇形单元的二维离散化模型, 可以捕捉强预冷器扇形单元内部各节点处的气动热

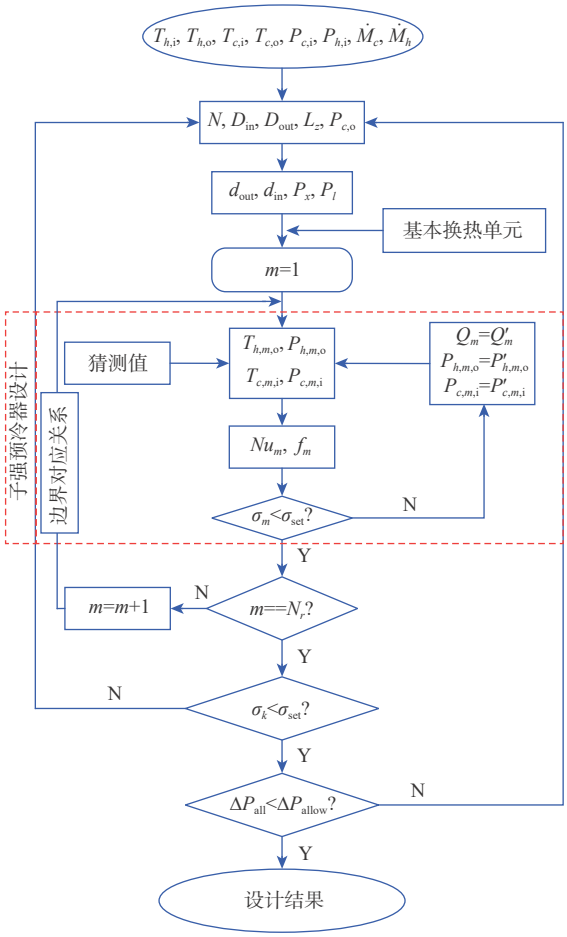


图 14 强预冷器 1D 离散化设计计算流程

Fig. 14 Calculation workflow for discrete design of 1D pre-cooler

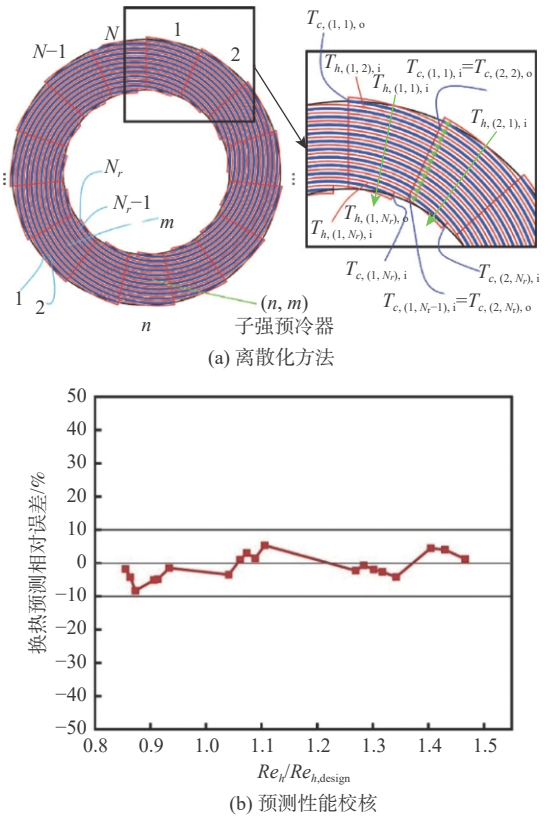


图 15 考虑全周非均匀的 2D 离散化模型[167]

Fig. 15 A 2D discretisation model accounting for non-uniformity across entire circumference[167]

力参数；北航团队[169]则提出了一种考虑化学反应机理的强预冷器 2D 离散化仿真分析方法，在上述基础上可以进一步考虑强预冷器内部流体热解等化学反应对流动换热的综合影响。此外，Hoopes 等[170]为降低非设计点性能评估复杂度，采用一种基于 1D 离散化结合 *Nu* 数修正的换热器无几何信息非设计工况工作特性评估方法，北航团队[171]通过实验校验了该方法良好的评估准确度。

基于该精细化设计方法，北航团队围绕强预冷发动机需求完成了超临界氦、超临界甲烷、超临界煤油、氢、液氮等多种结构形式紧凑快速强换热器设计，如图 16 所示。经实验验证，强预冷器换热功率、空气侧总压恢复系数及空气出口温度的实验值均与模型预测值吻合度高，其中，存在跨临界强变物性流动换热过程的超临界甲烷强预冷器实验结果与设计值最大相对偏差分别为 8.94%、–0.34% 和 –1.62%[132]，证明了该方法具有良好的设计精度。

上述 1D、2D 离散化方法除对螺旋线形式强预冷器适用外，在 PCHE 强换热器设计方法中也已经得到了广泛的应用[172-173]。但是，对于 PCHE 形式强换热而言，尺度效应带来的轴向导热问题及复杂结构下热阻是精细化设计需要考虑的另一重要问题[55-56]。常用方法是将轴向导热处理在 ε-NTU 法中处理为轴向导热数进行处理。在此基础上，李欢[171]提出了一种并联热阻模型，在考虑轴向导热的基础上，基于临界截面将 PCHE 半圆通道基本换热单元分割为 2 个子单元，如图 17 所示，而后分解为带

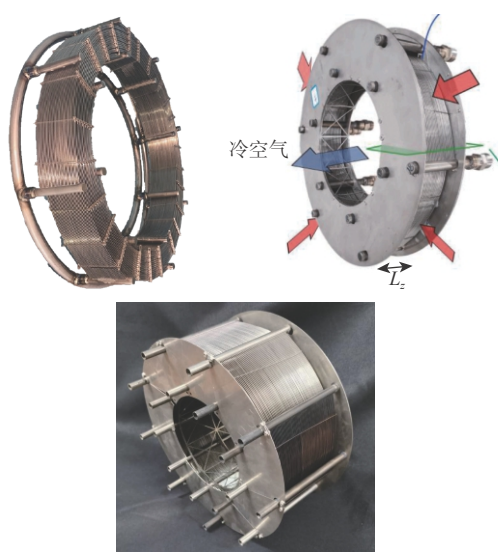


图 16 多种不同工质强预冷器

Fig. 16 Pre-coolers for various working fluids

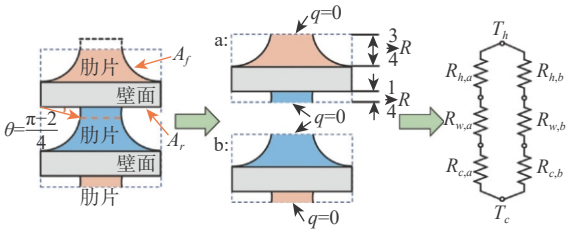


图 17　基本换热单元并联热阻模型[171]

Fig. 17　Parallel thermal resistance model for basic heat exchange units[171]

第三类边界条件一维导热问题求解, 提高了该类型换热器设计准确度。此外, Quintero[59]、Nellis[61] 等也建立了不同的换热器模型, 分别考虑了轴向和横向二维导热影响及轴向导热、外部热流和物性变化影响。

强预冷发动机中, PCHE 形式紧凑快速强换热器由于受发动机进气布局等影响较小, 且由于换热工质均为超临界压力流体, 可实现较小的流动损失, 因此, 可通过更为精细化的微小尺度异形通道设计实现强化换热效果。

首先, 在 PCHE 的通道横截面形状方面, 现有研究涉及半圆形、圆形、矩形、梯形、三角形等多种形式[174-175]。Lee 等[174] 针对半圆、矩形、梯形、圆形的 Z 形通道 PCHE 开展了效能、换热面积、摩擦系数等重要参数的对比, 结果显示, 矩形的换热性能最佳, 流动性能最差, 圆形的换热性能最差; Jeon 等[176] 针对半圆形、三角形、矩形 3 种直通道 PCHE 流动换热性能的研究表明, 在同等水力直径和质量流量条件下, PCHE 的通道横截面形状对流动换热性能的影响并不显著。总而言之, 在 PCHE 实际设计过程中, 需要结合具体工作环境、通道形式和加工能力进行考虑。目前, 半圆形的通道是综合性能及加工可实现性下最为常见的一种通道截面形状。

在增强换热方面, 研究人员还通过在通道内部增加扰流结构以增强换热。Aneesh 等[177] 在通道内增加了半球形凸起, 这种方式产生了更多漩涡, 增强换热的同时使得流动损失相应增大。当半球数量超过 2 个后, 换热能力变化减缓而压力损失不断增大。Han 等[178] 在 PCHE 通道中通过添加肋以增强换热, 研究表明, 一种离散分布的短肋条可使换热能力提升 19.3%～19.8%。此外, Zhao[179]、Ali[180] 等分别在 PCHE 通道内部添加不同形式的挡板以在增强换热的同时尽可能降低流动损失。

优化通道形式是 PCHE 精细化设计中的常见手段, 现有研究方案较为丰富, 大体可分为连续型和非连续型两大类, 一些典型方案示意图如图 18 所示。现有研究表明, 不同的通道形式往往需要在流动损失和换热能力 2 个方面进行协调, 而一些特殊的优化结构可以在换热能力基本不变的条件下得到更好的流动性能。

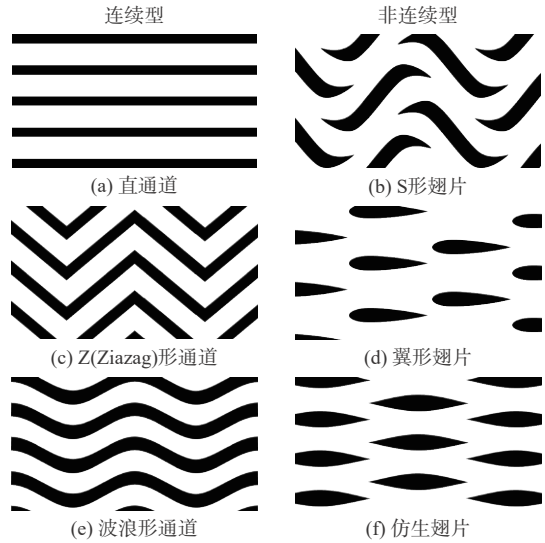


图 18　常见 PCHE 通道形式

Fig. 18　Common PCHE channel configurations

连续型通道中, 以直通道和 Z 形通道最具代表性, 同时还有正弦形、蛇形等多种形式[181-183]。相对而言, 直通道具有最低的压力损失, 而 Z 形通道通过转折处的局部湍流强化了换热性能, 较为简单的加工难度使其同样广受关注。Ngo[184] 和 Meshram[185] 等在各自研究范围内针对 Z 形通道的研究均得到相似结论, 即 Z 形通道的换热能力得到增强, 但流动损失大幅增加。其中, Meshram 等[185] 的研究表明, Z 形通道 *Nu* 数相比于直通道提升约 50%, 流动损失放大 4～5 倍。而围绕流动损失问题, 研究人员也通过通道优化改进方式开展了相关研究。Lee 等[186] 通过在转角处插入直通道或替换为波浪形通道抑制了流动分离, 削弱了流动损失; Xu 等[187] 提出了一种扭转式 Z 形通道, 通过 20°的扭转角实现了 13% 的换热性能强化并具有更低的流动损失。

在非连续型通道中, S 形翅片、正弦形翅片、翼形结构及仿生结构是精细化设计中被研究人员重点关注的方向。Ngo 等[184] 研究得出了 S 形翅片相对于 Z 形通道流动损失降低约 80%, 而换热能力仅下降 24%～34%; Saeed 等[188] 提出一种正弦形翅片方案以抑制流体的回流和分离问题, 相较于传统 Z 形通道压力损失降低了 2.4 倍, 而换热能力至多可提升 21%; Kim 等[189] 采用 NACA0020 翼型翅片, 在保证和 Z 形通道相似换热性能的前提下流动损失仅为 Z 形通道的 5%; 同时, Kim[190]、Li[191] 及 Cui[192] 等基于此通过对翼形肋安装形式角度进行精细化

设计及翼形改进，以进一步提升流动换热性能；此外，Xu 等[193]提出一种剑鱼仿生形翅片结构，Lu 等[194]研究表明，和 Z 形通道相比，剑鱼仿生翅片的换热能力提升了 15%～21%，而流动损失降低 33%。上述不同通道形式流动换热研究可为强预冷发动机内 PCHE 形式紧凑快速强换热器精细化设计提供重要参考。

此外，综合上述 2 种典型构型换热器设计方法，结合预冷发动机对于空气-氦强预冷器及氢-氦强换热器这 2 类不同类型换热器性能需求可以看出，两者设计理念的共性是尽可能实现高单位质量换热能力及高紧凑度，以降低在发动机安装环境下的质量及体积代偿。然而，由于强预冷器直接安装在发动机进气道及压气机之间，存在更为显著的工况多变及流动非均匀性问题，对上下游影响更为显著，因此，在设计优化方向选择中，应更为关注非均匀来流条件下高效换热及包线范围内流动换热，尤其是流动阻力问题。

3.3.3 紧凑快速强换热器不确定性设计方法

上述紧凑快速强换热设计均属于围绕一给定工况及名义设计参数的确定性过程。然而，真实条件下，换热器存在不确定性会对其实际工作性能产生影响[195]。Cho[196]前期研究指出，采用确定性设计的换热器在真实环境下实现设计热力性能的置信概率仅 50%。为此，研究人员已围绕换热器内部不确定性来源及影响开展了部分理论研究及设计方法改进工作。Haseler 等[197]研究了流体物性不确定性对不同种类换热过程的影响；Berryman[198]、Badar[199]等在换热器设计过程中引入随机模型，基于蒙特卡罗模拟给出设计方差，并开展了不确定性影响分析。

紧凑快速强换热器热力设计的不确定性因素主要包含随机不确定性和认知不确定性 2 类。其中，随机不确定性是强换热器及其环境的内在可变性，具体表现包括：换热单元直径及壁厚等几何加工误差、冷/热来流气动热力参数（温度、压力和流量）波动等；认知不确定性主要包括设计模拟简化及假设、经验关联式及物性计算模型等。这 2 方面问题共同作用，使紧凑快速强换热器在真实匹配装机环境下偏离设计工作状态，进而影响强预冷发动机可靠高效工作。

针对上述非确定性问题，北航团队发展了强预冷发动机紧凑快速强换热器非确定性设计方法[132,200]，分别构建了基本换热单元、来流热力参数等不确定性来源及分布模型，如图 19 所示（图中 μ' 为外径测量值的平均值）；进一步基于蒙特卡罗及 SOBOL 方法构建了紧凑快速强换热器分析及设计流程。基于此，北航团队针对超临界工质强预冷器不确定性设计展开了深入研究，探究了来流参数、工质物性及几何偏差等输入参数不确定度对强预冷器设计的影响，相关结果如图 20 所示。研究表明，强预冷器总换热量近似呈威布尔分布，且高

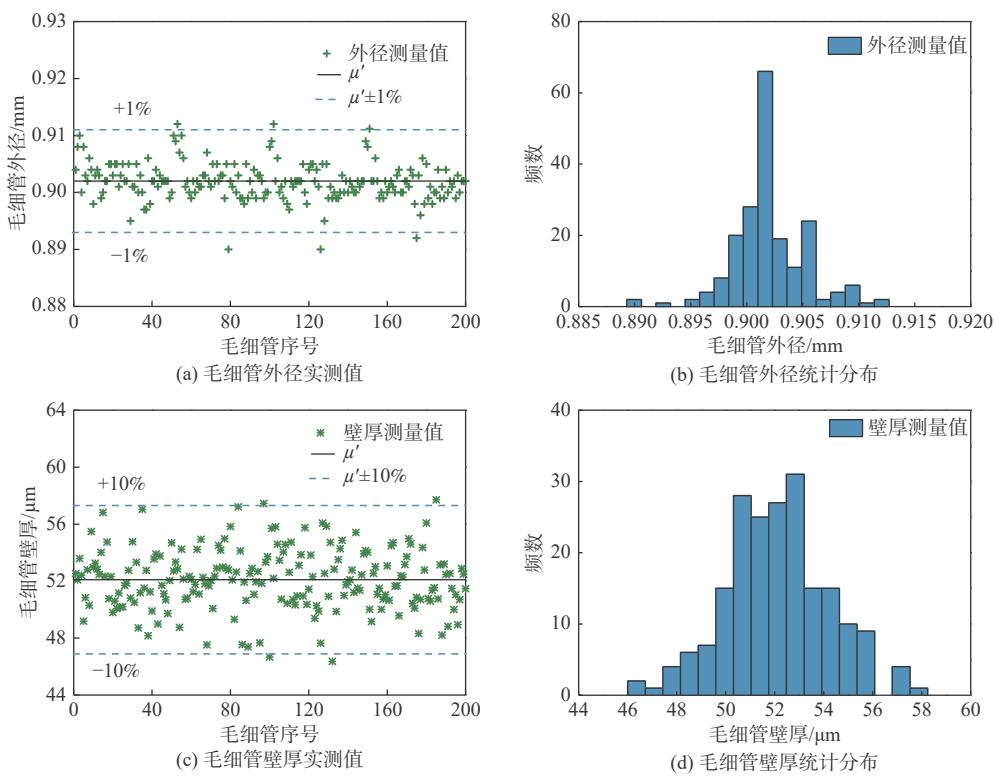


图 19 超薄壁毛细管几何偏差统计结果

Fig. 19 Statistical results on geometric deviations in ultra-thin-walled capillaries

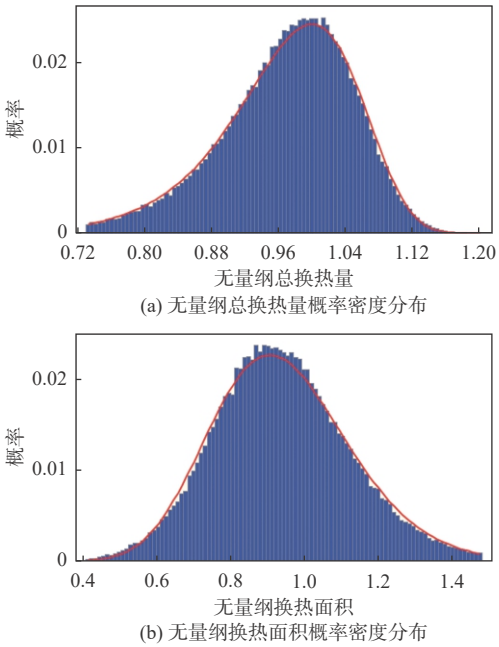


图 20　超临界氦强预冷器不确定性设计结果

Fig. 20　Design results for a supercritical helium pre-cooler with uncertainty analysis

温来流温度是影响其分布形式的主要因素。由于强预冷器所需换热面积同样受输入不确定性影响，为保证设计置信度达到 95%，换热面积的计算冗余应不低于 27.7%；而当设计置信度达到 99% 时，总换热面积冗余将达到 47.1%。

此外，北航团队[132]对紧凑快速强预冷器在发动机匹配环境下的实验结果开展了不确定性分析，如图 21 所示。结果表明，不同影响因素的敏感性大小不同，在相同不确定度情况下，强预冷器换热功率和空气出口温度对来流温度最敏感，而空气侧总压恢复系数对空气流量最为敏感，加工几何偏差的影响较小；且不确定度的分布形式是影响极端异常换热现象的重要因素，虽然不同分布形式下强预冷器工作性能统计均值与设计名义值均基本相同，但换热性能极端值及极端事件发生概率具有明显区别，将有可能直接影响强预冷器的工作安全性。

3.4　紧凑快速强换热拓扑优化方法

拓扑优化已成为换热器流动换热设计的重要方法之一，也可被借鉴用于强预冷发动机的紧凑快速强换热器设计之中。从理论上讲，拓扑优化具有最大的设计自由度，可实现远超传统设计方法的复杂换热器结构，当前变密度法和水平集法是最具代表性的参数化方法[201-202]，其基本思想及原理可见文献 [201-203] 所述，其中，变密度法得益于较好的收敛性及更低的计算成本，是目前应用最广泛的换热器结构拓扑优化方法[204-206]。近年来，围绕变密度法优化分析精度问题，Osher 等[201]提出水平集法以解决对流固边界敏感的换热器优化问题，Yaji 等[207]基于水平集法开展了二维和三维流动换热耦合的拓扑优化研究，而 Coffin 和 Maute[208]针对自然对流问

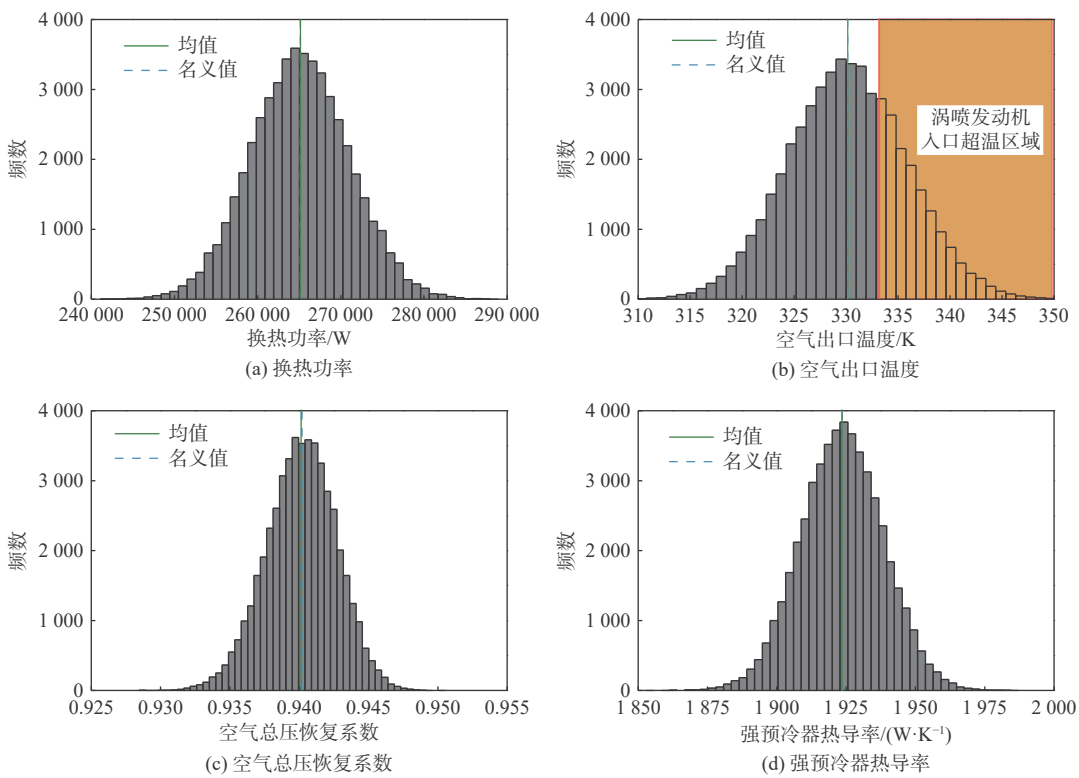


图 21　强预冷器关键性能参数实际分布直方图

Fig. 21　Histogram of actual distribution of key performance parameters for pre-cooler

题的稳态和瞬态工况开展了研究；进一步，以基函数或自定义函数的组合替代传统 Hamilton-Jacobi 控制方程成为水平集法研究的热点之一，Li 等[209]基于一种简单条状结构构建了一组由 8 个额外参数确定的水平集函数，并构建了一种具备良好优化能力的微小尺度通道换热器拓扑优化框架。在理论研究之外，研究人员开展了相关实验研究与验证，如图 22 所示。Li[210]、Qian[211]、Mo[212] 等均基于拓扑优化结果开展了换热器研制及实验验证工作，结果均验证了拓扑优化在增强换热、降低流动损失方面的有效性。

此外，伴随增材制造技术的发展，基于三周期极小曲面 (triply periodic minimal surface, TPMS) 的紧凑快速强换热器成为重要的研究方向[213]，常见的 TPMS 类型包括 IWP、Diamond、Gyroid、Primitive 等（见图 23）[214]。

Khalil[215]、Kaur[216] 等多个研究学者开展了 TPMS 换热器研究工作，结果表明，TPMS 结构具有更好的换热性能，因为孔隙率对这类换热器性能有着直接影响；Huang[217]、Men[218]、Shu[219] 等分别采用变密度法和水平集法开展了针对 TPMS 换热器的拓扑优化研究，并得到了换热和流阻的不同优化方向和结果。

同时，北航团队初步开展了 TPMS 换热器设计方法研究及实验验证（见图 24），结果验证了这种新型换热器在流动换热增强方面的优势；此外，Huang 等[220] 将 TPMS 思路应用在强预冷发动机强预冷器中，探讨了 TPMS 应用于强预冷发动机的可能性。

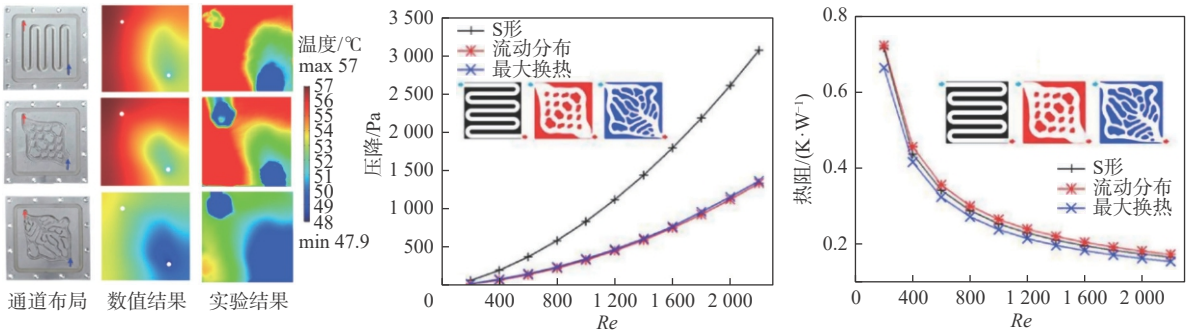
可以看出，针对换热器的拓扑优化研究方法在过去的 20 年间已经得到长足的发展，部分换热器拓扑优化工作越来越贴近工程实践[210-212]，且表现出强大的优化能力。TPMS 换热器和拓扑优化作为紧凑快速强换热器的研究前沿，在换热器优化设计领域具备相当的潜力。但仍需注意的是，这类新型紧凑快速强换热器结构复杂，其换热功重比、紧凑度等指标性能的提升将显著依赖于加工制造水平的进步。

3.5 强预冷器结霜抑制技术

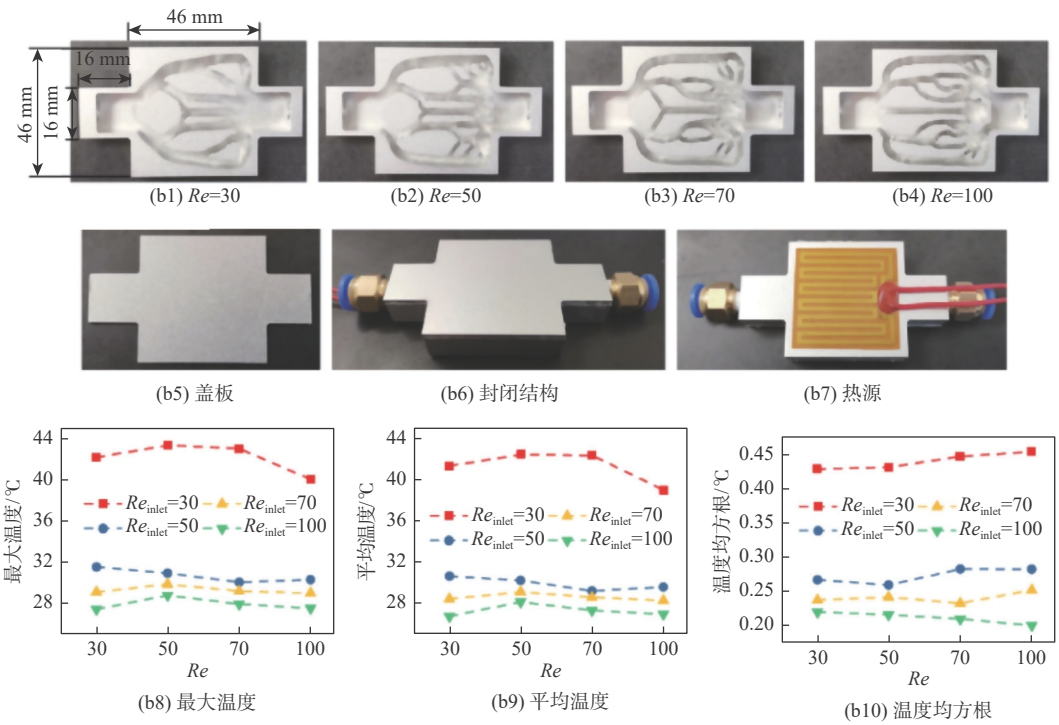
强预冷器在较低空域（约 12 km 以下）工作时，将面临管壁外侧狭窄通道内形成霜层或冰层问题，严重时将堵塞流道致使空气侧压力损失急剧增加且畸变加剧，此外霜层/冰层会显著增加热阻，使强预冷器换热性能急剧恶化，为此，有必要在强预冷器工作时采取一定措施对其进行抑霜/抑冰操作。目前，研究者针对强预冷器内部结霜问题开展了广泛的理论及实验研究工作[221-223]，提出的抑霜/抑冰措施主要包括 3 种：①在主流中掺混减轻结霜的抑霜剂（如甲醇、甲醛等工质）或低温工质以提前析出冰粒，通过改变霜层形成或理化性质以降低霜层对流动换热的影响；②采用微纳结构、表面涂层等技术改变表面结构，进而降低霜层附着力以抑制表面结霜；③采用喷射固体颗粒或脉冲空气，利用粒子撞击或空气泡撞击霜层以抑制附着。本节分别综述了上述强预冷器抑霜/抑冰手段进展。

首先，对于抑霜剂喷射掺混方法，英国 REL 公司提出了一种强预冷器甲醇喷射抑霜装置，如图 25 所示[224]。通过喷射甲醇作为抑霜剂，并通过一种集成在强预冷器微细管阵列内部的回收结构收集气流中的水-甲醇混合物，宣称可以去除空气中 99% 的水分，进一步缓解堵塞问题。基于此，英国 REL 公司搭建了抑霜技术验证风洞并开展了实验研究[36]，结果表明，在−80℃、100% 相对湿度情况下可持续工作 8 min，保证空气侧压力损失无明显变化。日本 JAXA 采用该技术也开展了大量的理论及实验研究[23]，其中，Sato 等[221] 通过缩尺模型实验对比分析了不同抑霜剂效果，结果表明，甲醇具有最佳的抑霜效果，对比工况下可将强预冷器工作时间由 20 s 提升至 300 s 以上，乙醇与异丙醇也具有较为良好的效果。北航团队[225] 从霜层形成过程出发，搭建了缩尺实验台并开展了霜层生长模式及影响规律研究，结果表明，强预冷器在低雷诺数及高雷诺数工况下结霜模式不同，较低雷诺数（$Re \leqslant 1\,500$）下，结霜可能会造成严重的管束变形，且表面润湿性也将显著影响结霜过程中的流动损失特性，如图 26 所示。在此基础上，北航团队联合上海交通大学团队[226-227] 完善了强预冷器抑霜剂喷射设计方法及参数选取准则，通过缩尺实验验证了甲醇喷射抑霜的有效性。此外，北京动力机械研究所[228]、大连理工大学[229] 等研究单位也开展了相关结霜与抑霜实验研究。

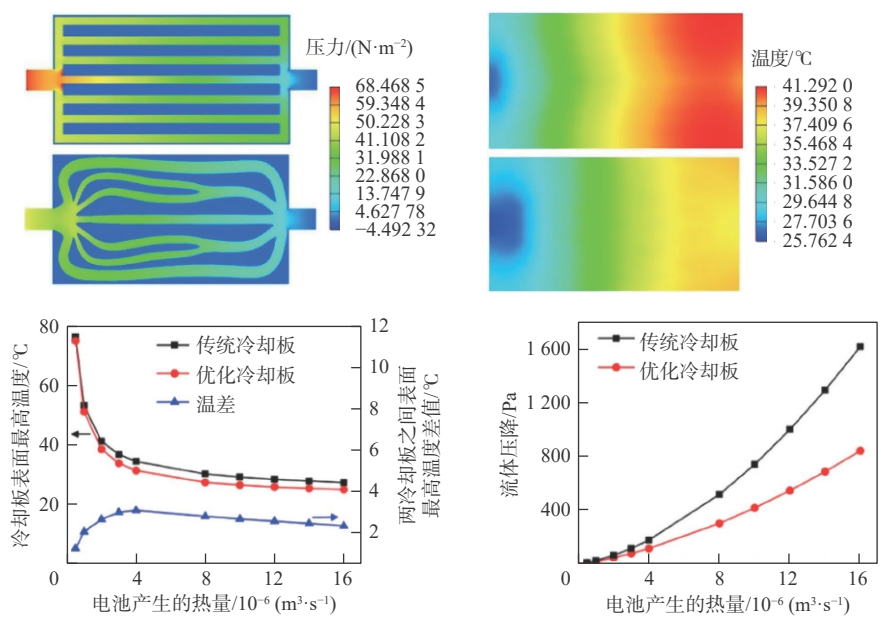
其次，基于表面微纳结构的抑霜/抑冰手段近年来越来越受到关注[230-232]，其降低霜-表面附着力、减小霜层厚度、延迟结霜时间等特点使其有望成为解决强预冷器结霜问题的手段之一。主流观点认为表面润湿性和表面结构是影响抑霜效果的主要因素。国防科技大学岳晓菲[222] 以强预冷器为研究背景，对疏水微纳结构表面开展了防冰抑霜性研究，提出了微-纳二级结构和高度规整、耐磨性强的单级结构是合适的防霜表面结构。东南大学、北京动力机械研究所[233] 则从涂层手段出发开展了强预冷器抑霜研究，结果表明，在−20℃ 下 ZnO 超疏水表



(a) 蛇形通道换热器拓扑优化的实验对比[210]



(b) 简单平板拓扑优化的实验对比[211]



(c) 平行通道换热器拓扑优化的仿真对比(通过实验校核仿真可靠性)[212]

图 22　一些拓扑优化结果的实验研究及性能对比

Fig. 22　Experimental study and performance comparison of some topological optimisation results

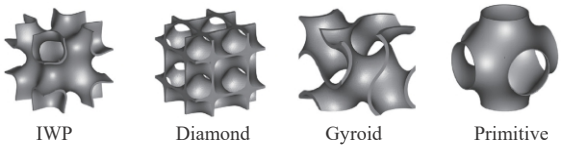
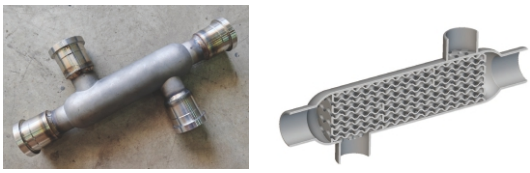


图23 常见的TPMS结构[214]

Fig. 23 Common TPMS configurations[214]



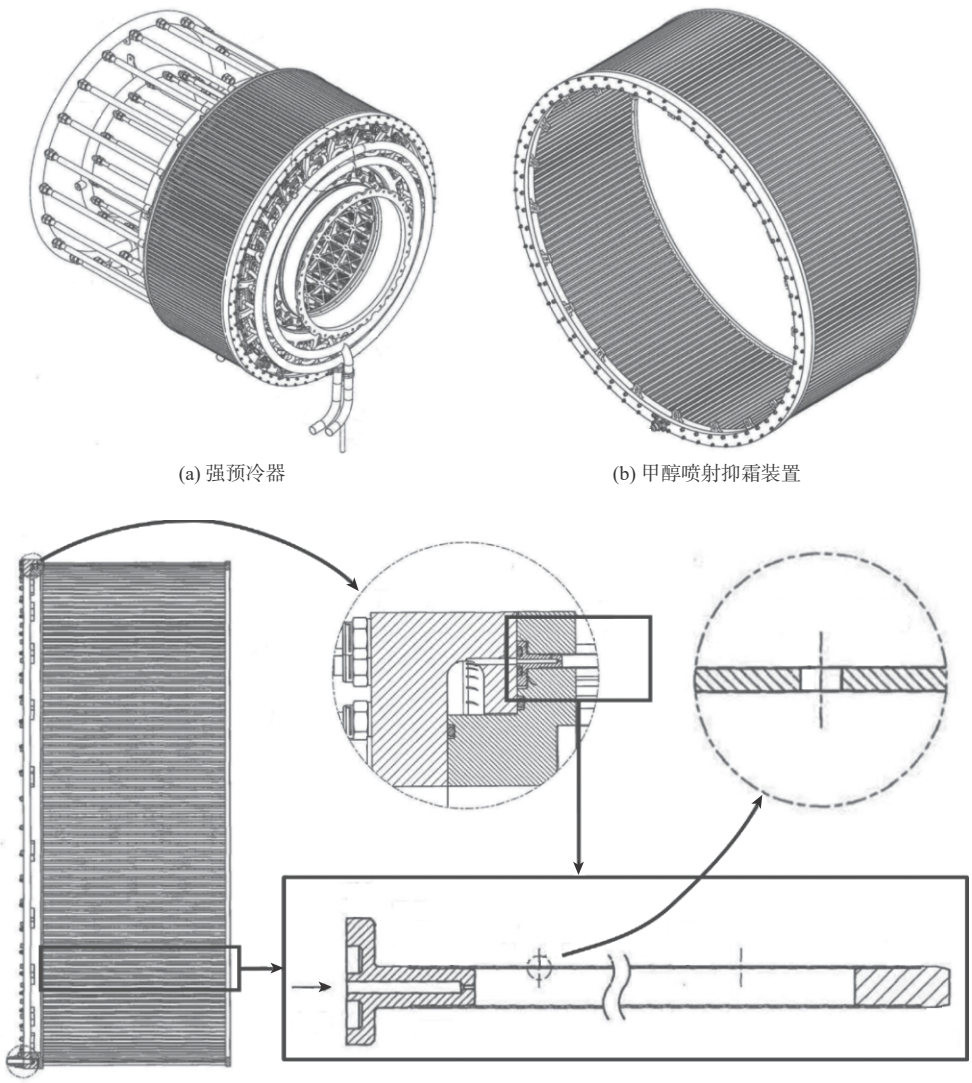
(a) 实物图 (b) 剖视图

图24 TPMS换热器

Fig. 24 TPMS heat exchanger

面可以延迟表面霜晶的形成长达7 min，结霜量减少了一半，在−150℃下超疏水表面仍有一定的抑制结霜的效果。此外，对于喷射固体颗粒或脉冲气流手段，Sonobe等[223]开展了实验研究(见图27)，结果表明，该方法也可有效防止霜层聚积，初步验证了该方法的有效性。

然而，综合分析上述3种方法，虽然在研究过程中均表现出明显的抑霜效果，但也存在一些有待突破的问题：①对于抑霜剂喷射方法而言，其工质一般为可燃性气体，在强预冷器高温工作环境下可能存在工质局部达到可燃极限而高温下提前燃烧，这给强预冷发动机可靠安全工作带来隐患；②对于表面结构及涂层而言，首先是有研究表明在长时间工作下抑霜效果一般，其次是强预冷器由钎焊加工而成，在密集阵列管束表面构造耐高温、抗氧化的表面结构或涂层存在一定难度；③喷射固体颗粒或脉冲气流手段在发动机平台增加额外引气、粒子装置，存在系统较为复杂问题，且脉冲气流有可能对下游压气机等部件稳定工作带来影响。

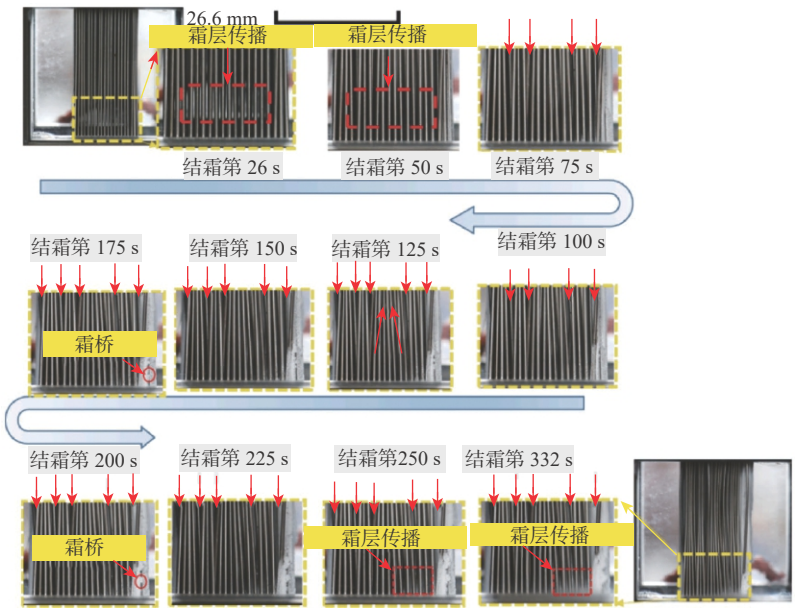


(a) 强预冷器 (b) 甲醇喷射抑霜装置

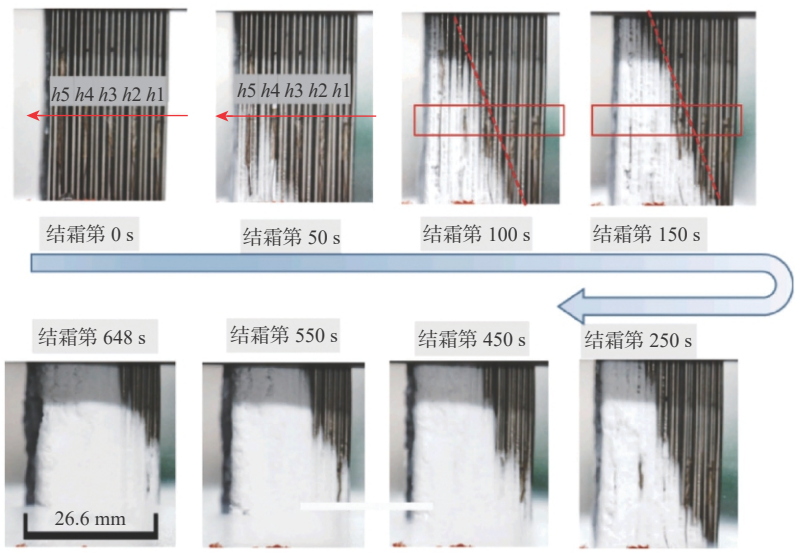
(c) 甲醇喷射抑霜装置细节

图25 强预冷器及甲醇喷射抑霜装置结构[224]

Fig. 25 Schematic diagram of pre-cooler and methanol injection anti-icing system[224]



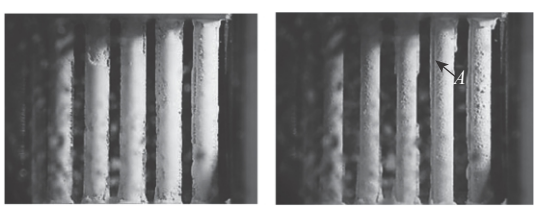
(a) 空气, *Re*=1 500, 表面无湿润



(b) 空气, *Re*=2 800, 表面湿润

图 26　霜层生长模式及影响规律研究[225]

Fig. 26　Study on growth patterns and influencing factors of frost layers[225]



(a) 除霜前　　(b) 除霜后

图 27　脉冲气流除霜前与除霜后对比[223]

Fig. 27　Comparison of pulsed airflow before and after defrosting[223]

4　紧凑快速强换热实验技术

实验手段是探究紧凑快速强换热内部真实流动换热现象的核心手段之一, 也是充分验证紧凑快速强换热器设计技术、掌握其真实工作特性的关键一步。然而, 强预冷发动机内部紧凑快速强换热器存在通道狭窄、高温、超低温、高压、参数畸变或强瞬变、振动及热疲劳等条件的多种组合, 对真实环境模拟实验及其参数测量提出了严峻挑战。本节针对复杂条件下紧凑快速强换热测试手段及实验验证现状进行了综述。

4.1　紧凑快速强换热测试技术

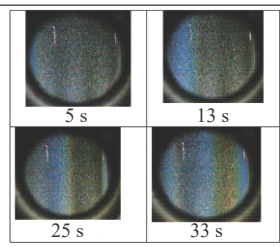
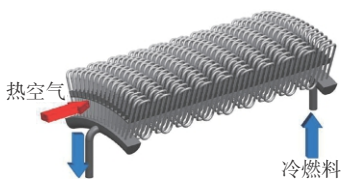
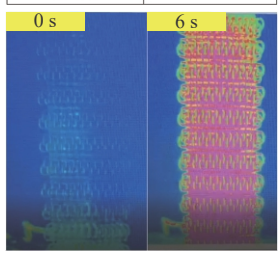
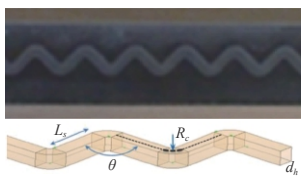
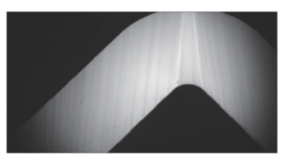
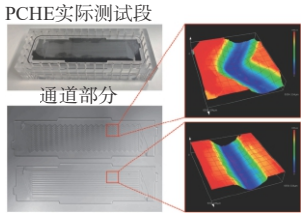
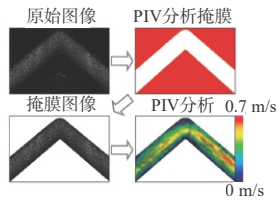
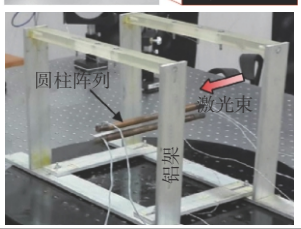
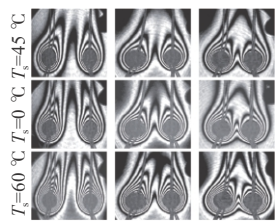
对于强预冷发动机内部紧凑快速强换热过程而言, 实验技术的核心之一在于紧凑空间复杂工况环境下对于流动状态及换热能力的高精度测量。然而, 传统接触式测量技术受限于流动通道狭窄且被测换热单元尺度较小而受到应用限制, 因此, 在微小尺度实验研究中, 受部分实验对象极紧凑结构限制,

实验过程仅可获取研究对象进出口的参数而忽略内部流场及温度场细节。而随着测量技术水平增加，进一步扩展了流动换热测试细节，其中，最典型的为接触式测量，即在微小通道试验件盖板上钻取小孔，通过引压或埋藏热电偶的方式获得沿程的压力和温度分布。例如，Murray 等[234]、北航团队[150,235-236]均基于此对紧凑快速强换热器基本换热单元流动换热机理开展了实验研究；此外，Chen 等[167]利用梳状探针获得了紧凑强预冷器进出口温度分布特征。

然而，随着流动换热研究的深入，传统接触式测量由于尺寸及结构等问题将限制空间分辨率及测量误差，已无法完全满足对内部细节的测试需求。为此，研究人员引入了多种非接触式测量手段，以期进一步提升流动换热测试能力，其中，主要手段包括液晶示温技术[81,237]、红外测温技术[238-239]、激光诱导荧光（laser induced fluorescence, LIF）技术[240-241]、粒子图像测速（particle image velocimetry, PIV）技术[242-243]、激光干涉测温[244-245]等，其基本原理与特点如表4所示。基于上述技术方法，研究人员开展了广泛的微小尺度流动换热研究。例如，Xu 等[246]利用红外技术测量了微通道热沉表面温度分布；Xia 等[242]利用 μ-PIV 流动显示技术研究微通道热沉内部流场；Eckert 等[247]等利用干涉成像方法开展了圆管横掠绕流流动换热研究；Fu 等[248]利用 PIV 对圆管壁面换热环境中流动与换热的交互作用机理开展了实验研究；REL 公司 Webber 联合布里斯托大学 Taylor[81]针对强预冷发动机内部紧凑快速强换热研究需求，利用液晶示温技术开展了管排内部圆柱流动换热研究。但需要注意的是，由于紧凑快速强换热器尺度小、结构紧凑、通道长径比大，上述对其内部进行流动换热细节现象捕捉的测试手段也难以直接应用，需进一步基于相似原理

表4 非接触式测量手段基本原理与特点

Table 4 Basic principles and characteristics of non-contact measurement methods

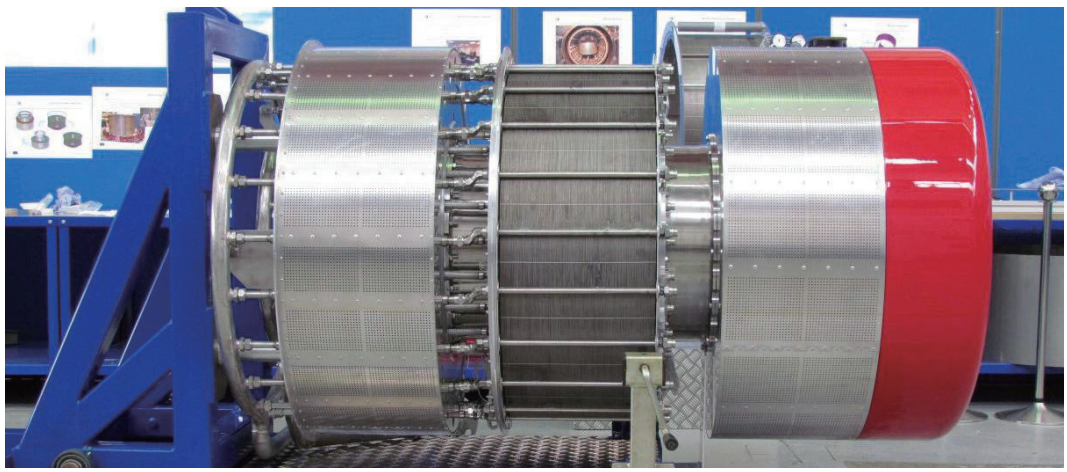
测量手段	被测换热管或微通道	测量结果	基本原理与特点
液晶示温[81]			①液晶分子排列对温度敏感，反射特定波长的光，通过颜色变化指示温度，可测量表面温度； ②空间分辨率高； ③测温范围有限
红外测温[239]			①基于热辐射原理，通过检测物体发出的红外辐射能量来测量温度； ②测量响应快； ③无法测量内部温度，易受环境干扰
激光诱导荧光[241]			①激光激发被测分子/原子的电子跃迁，通过所产生的荧光强度/波长获取温度； ②空间分辨率高； ③对被测样品及环境要求较高
粒子图像测速[243]			①捕捉含有示踪粒子的流体图像并计算粒子位移，进而获得流速； ②可实现全场瞬态测量； ③示踪粒子可能干扰流动
激光干涉测温[245]			①利用激光的相干性，将参考光与测量光叠加形成干涉条纹； ②可实现纳米级精度； ③对环境振动敏感

通过尺度缩放手段对实验进行设计。此外, 受限于紧凑快速强换热器通道狭窄、空间结构复杂, 对强预冷器通道狭窄空间内的流动进行多视角成像较为困难。近年来, 融合物理信息/机理的人工智能技术在复杂流场重构与可视化中的应用获得了大量关注[249-250], 如 Raissi[251]、Arzani[252]、Cai[253] 等开展的相关研究分别实现了流场及热边界条件反演, 可为今后紧凑快速强换热实验测试手段提供新的思路。

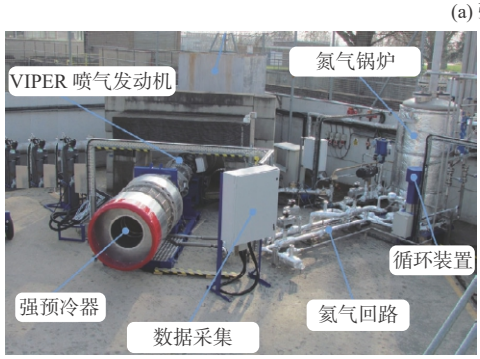
4.2　紧凑快速强换热器实验研究

紧凑快速强换热器由于工况多变、内部流动换热特征复杂, 上述机理级别实验研究虽然能够对强换热器设计及工作性能预测提供重要理论依据, 但其部件实际工作性能至少需要在样机级别、模拟真实复杂工作环境下进行综合验证。为此, 国内外研究机构围绕紧凑快速强换热器工作性能需求搭建了不同规模流动换热实验台, 并完成了多种构型、多种工质、不同工作条件下的实验研究, 本节围绕国内外开展的典型紧凑快速强换热器实验研究进行了综述。

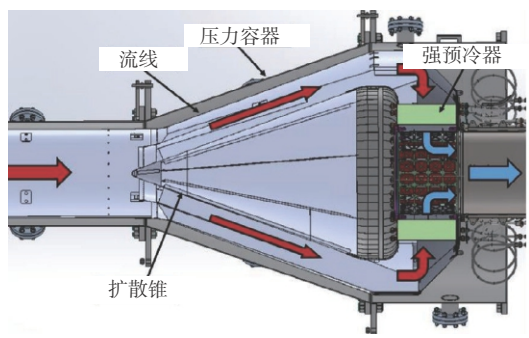
英国 REL 公司为验证其强预冷器性能, 分阶段完成了强预冷器常温及高温实验验证。第一阶段, 以实现低温高压超临界氦闭式循环和大气湿度条件为建设目标, REL 公司搭建了 B9 实验系统, 该系统由一台罗·罗公司 Viper 发动机、一套超临界氦闭式循环系统构成[254], 强预冷器实验件及联合 Viper 发动机的常温实验场景如图 28(a)、图 28(b)



(a) 强预冷器实验件



(b) 联合Viper常温实验



(c) 强预冷器高温实验装置



(d) *Ma*=5模拟工况高温实验平台

图 28　REL 开展强预冷器常温及高温实验[33,110]

Fig. 28　REL conducts ambient and high-temperature tests on high-efficiency pre-coolers[33,110]

所示[110]。其中，Viper发动机可提供约10 kg/s量级的空气流量，通过燃烧液态丁烷可实现发动机在约150 K进气条件下稳定工作，氦闭式循环系统可实现最低温度约140 K、稳定工作流量约1.3 kg/s、压力为152.0 kPa超临界氦持续供给；此外，该实验系统还配置了水滴喷注系统以实现不同湿度环境的模拟。至2014年，B9实验系统已承担了超过700次台架实验和300次强预冷器-发动机匹配实验，支撑了强换热器低温工作性能及整机环境下抑霜技术的实验验证工作[254]。

基于B9的设计运行经验，REL公司在美国国防高级研究计划局（Defense Advanced Research Projects Agency, DARPA）的支持下完成了强换热器高温性能验证循环实验系统的研制与建设工作[16]。该循环实验系统由TF2高温来流系统和GSE超临界氦闭式循环系统构成。TF2高温来流系统使用一台带有加力燃烧室的J79涡喷发动机作为高温气源，在J79发动机处于最大加力状态时能够实现*Ma*=5来流温度环境（约1 300 K）的模拟。GSE超临界氦闭式循环系统主要由超临界氦循环压缩机和大功率冷却系统组成，能够将温度高达950 K的氦冷却至室温附近，最大冷却能力达到5 MW，具备2 kg/s、20 MPa超临界氦的持续供给能力。该循环实验系统已成功承担了*Ma*=3.3和*Ma*=5来流温度条件下的强换热器高温考核实验[33]，在*Ma*=3.3来流总温模拟实验中，高温燃气流经强预冷器后温度降至约100 ℃，产生了近1.5 MW的换热功率，强预冷器样机实验效能达到99.8%；在图28（c）、图28（d）所示的*Ma*=5来流总温模拟实验中，REL公司成功验证强预冷器样机可在超高温来流条件下（约1 000 ℃）稳定工作，预估完成了100 kW/kg功重比的技术指标。

此外，日本JAXA为满足ATREX及PCTJ发动机对于紧凑快速强换热器性能需求，搭建了基于发动机匹配环境的地面实验验证平台，并开展了多种构型强预冷器在发动机匹配环境下实验验证。该试验台利用高压氦气驱动液氢或液氮冷剂（均处于超临界压力）进入强预冷器，以冷却发动机吸入的室温来流，同时配有甲醇喷射装置探究强预冷器结霜抑制效果。在1992—2003年间，JAXA针对图29（a）所示的ATREX发动机强预冷器开展了大量常温来流实验验证，实验工况空气流量可达约6 kg/s级，可将常温来流降至约200 K[13,23]。通过对多种构型强预冷器对比实验验证，得出所示的Type-Ⅱ型强预冷器具有良好的可靠性及空气总压恢复系数，但换热功率相较设计预期低15%～20%，在此基础上发展的Type-Ⅲ型强预冷器实现减重

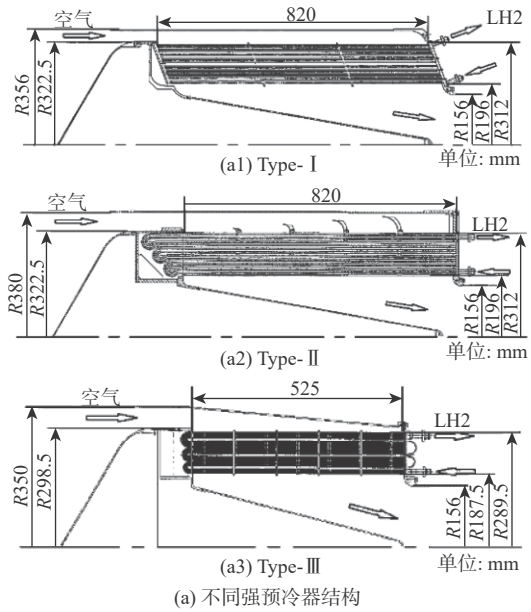
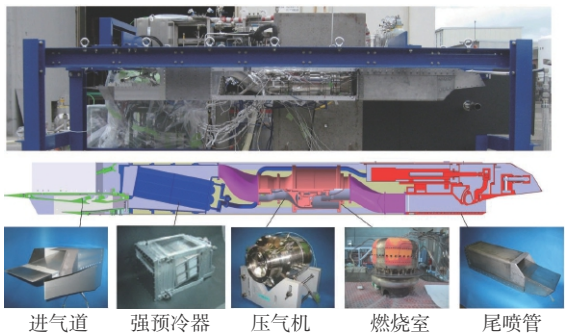


图29 不同强预冷器结构及总压损失

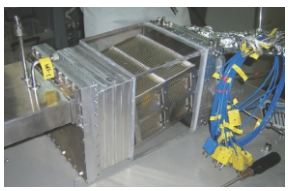
Fig. 29 Different pre-cooler configurations and total pressure drop

54.3%，但其压力损失却随着微细管表面结霜而急剧增加，远高于其他构型强预冷器，如图29（b）所示。

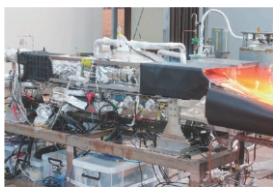
在上述基础上，围绕PCTJ构型，日本JAXA进一步为其飞行演示验证发动机S-engine（见图30（a）[255]）设计了缩尺矩形强预冷器样机（见图30（b），包含1 296根外径为2 mm的薄壁毛细管[256]）并开展了匹配条件下常温、高温实验验证。2015年，Taguchi等[257]利用图30所示风洞试验台探究了发动机整机环境下强预冷器的温降性能，以超临界氦为强预冷器冷剂，空气流量约为0.5 kg/s，试验结果表明，约800 K高温空气流经强预冷器后平均温度降至590 K（不同测点温度处于550～630 K之间），实现了210 K温降水平。2018年，Taguchi等[258]在JAXA角田航天中心完成了超临界氢预冷的S-engine发动机*Ma*=4风洞吹风试验，强预冷器样机入口温度同样约为800 K，但最低出口温度仅达到650 K，温降水平较低，研究人员给出2个可能原因：①强预冷器



(a) S-engine实物图及示意图[255]



(b) 矩形强预冷器[256]



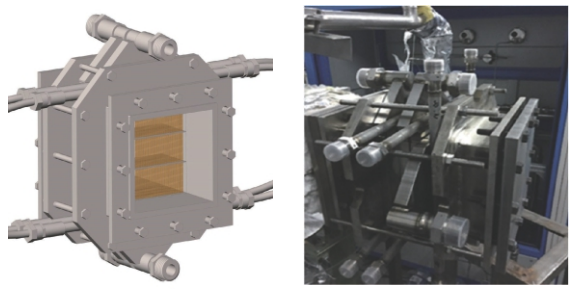
(c) 模拟*Ma*=4实验

图 30　日本 PCTJ 构型实验研究

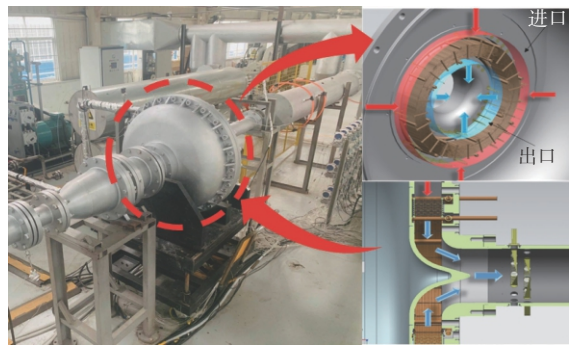
Fig. 30　Experimental study of PCTJ configuration in Japan

入口处安装的异物防护板降低了换热功率; ②液氢供应管道未完全预冷, 导致流入强预冷器样机内的冷剂并非液态, 且流量也未达到目标值。为进一步探究超临界氢强预冷器工作性能, Taguchi 等[37] 基于图 30 所示直连试验台完成了模拟 *Ma*=4 来流总温条件的发动机海平面静态实验, 如图 30(c)所示。实验中, 超临界氢入口温度和压力分别为 40 K 和 2 MPa, 空气来流平均温度达到 900 K。由于流动存在非均匀性, 强预冷器出口温度畸变范围为 390～470 K, 但平均出口温度仅为 430 K, 温降达到 470 K。此外, 高温来流在强预冷器内的压力损失保持在 10 kPa 左右, 总压损失为 3.6%。

在国内, 北航团队基于上述技术基础分阶段完成了多种构型紧凑快速强换热器样机研制工作, 如图 31 所示。进一步地, 为实现对紧凑快速强换热设计技术的综合验证, 团队分阶段完成了不同构型紧凑快速强换热器设计研制, 搭建了国内领先的综合实验验证平台, 并先后完成了多工质、多工况强预冷器性能实验验证。2017 年, 北航团队[166,235] 在国内首次针对微细管阵列结构强换热器模块开展了实验研究, 以图 31(a)所示的矩形换热模块为实验对象, 以超临界氦作为冷却介质, 完成了来流 800 K 的高温实验验证, 实现了 86.8 kW/h、温降 226.2 K 的既定技术指标, 证明了这种微细管阵列结构紧凑快速强换热器在强预冷发动机上应用的可行性。进一步, 为验证强预冷器工作性能, 团队在国内率先搭建了基于超临界氦闭式循环的强预冷器高温来流性能验证平台(见图 32(a)), 实现了 8 MPa 超临界氦闭式连续供给。针对图 31(b)中阿基米德螺线构型强预冷器原理样机, 在国内率先完成了模



(a) 微细管阵列结构强换热器模块



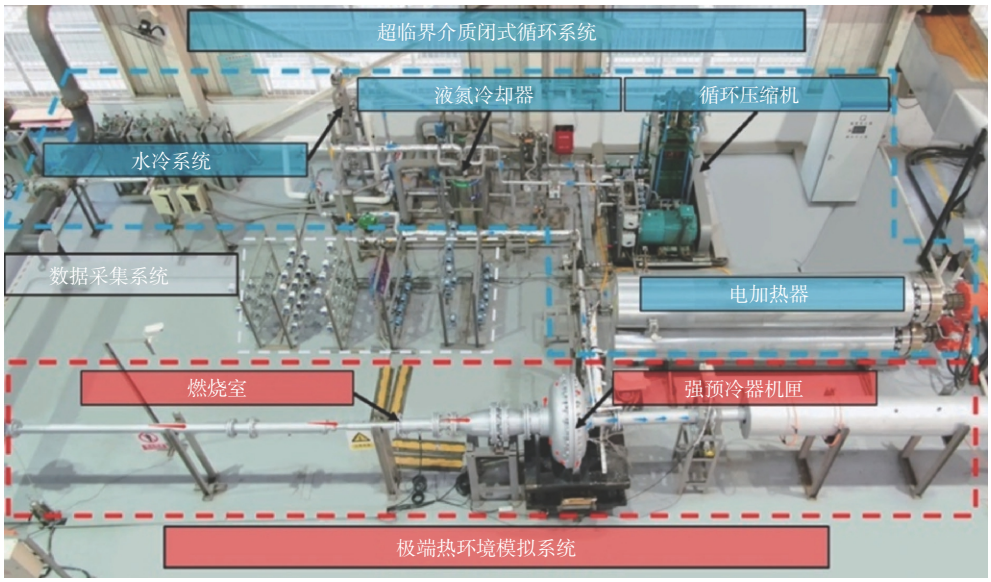
(b) 强预冷器原理样机

图 31　不同阶段试验样机

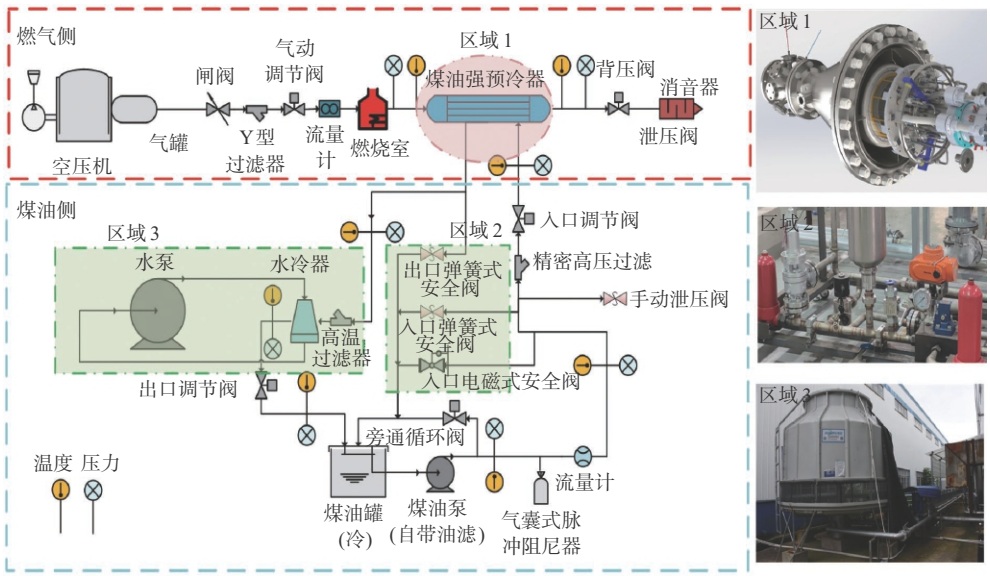
Fig. 31　Prototypes at different stages

拟 *Ma*=4+来流总温条件 1 800 s 长时实验验证, 实现了温降 600 K、换热功重比 100 kW/kg 的先进技术指标[167]。

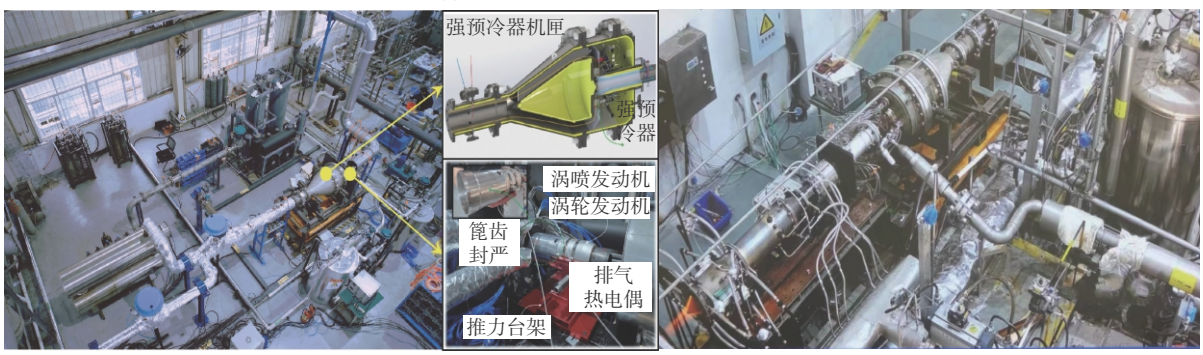
此外, 为验证紧凑快速强换热器对于传统航空燃气涡轮发动机扩包线的适用性, 团队完成了模拟 *Ma*=4 来流总温条件实验验证, 实验平台如图 32(b)所示。结果表明, 虽然煤油强预冷器实现了功重比约 112 kW/kg、总压恢复系数 0.93, 但受限于煤油有限热沉, 预冷所需燃料与发动机化学恰当燃烧所需煤油流量不平衡, 存在“热沉缺口”。进一步地, 为探究紧凑快速强换热器在匹配环境下工作特性, 北航团队于 2022 年搭建基于涡喷发动机匹配环境的强预冷器实验平台, 如图 32(c)所示, 利用安装在强预冷器后的涡喷发动机, 验证以氦、液氮及液态甲烷为冷却工质的强预冷器在复杂匹配环境下的流动换热性能[132,259]。结果表明, 采用液氮工质的强预冷器缩尺样机可实现将模拟 *Ma*=3 来流总温条件(约 600 K)降至约 338 K, 核心区换热功重比达到 60 kW/kg, 总温畸变指数 7.8%, 同时, 实验结果表明, 由于压气机抽吸作用影响, 将导致局部换热能力有所不同; 采用液态甲烷作为冷却工质的强预冷器匹配实验中, 空气来流平均温度高于 *Ma*=3.5 滞止温度(最大值达到 764.8 K), 实现了约 420 K 温降, 整机和核心换热区功重比的平均值分别达到 84.5 kW/kg 及 116.1 kW/kg。在充分掌握强预冷器匹配工作性能基础上, 北航团队完成了基于液态甲烷的强预冷涡轮发动机缩尺样机研制及地面实验



(a) 基于超临界氦闭式循环系统的强预冷器高温实验平台



(b) 超临界煤油强预冷器高温实验平台



(c) 强预冷器-涡喷发动机实验平台

(d) 强预冷涡轮发动机缩尺样机地面实验验证平台

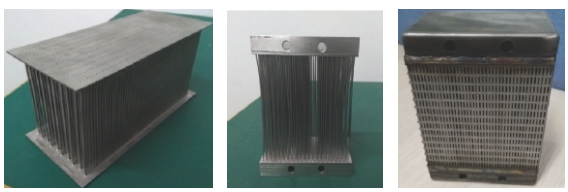
图 32 北航团队强预冷器研究实验平台

Fig. 32 Beihang University team's research and experimental platform for high-performance pre-coolers

验证平台建设，具备国际先进的强预冷发动机实验验证能力，如图 32(d)所示。团队完成国内首次模拟 Ma=3.7 来流总温条件下的基于液态甲烷的强预冷涡轮发动机地面直连实验验证，结果表明，液态甲烷工质强预冷器可实现约 480 K 温降以保证后端涡轮发动机正常工作，达到了国际领先水平。综上来看，团队在流动换热机理、模拟高马赫数进气高温环境及强预冷器匹配环境开展了大量实验验证，在对整体流动换热性能验证方面具有技术优势，但当前实验验证仍局限在直连实验台环境，缺

乏在真实飞行环境下对于强预冷器工作特性的考核验证，后续需进一步开展相关真实飞行环境下实验验证。

此外，根据公开文献，国内其他多个研究机构也同样开展了强预冷器性能实验研究。西安航天动力研究所[158]针对强预冷器样机开展了空气-空气流动换热实验，其中，高温侧来流 686 K，结果表明，经典公式对强预冷器换热预测偏差较大，最高可达 32.1%。北京动力机械研究所[19]完成了强预冷器 *Ma*=5 来流高温条件下换热实验，但实验工质尚未披露。国防科技大学纪佳正[260]通过实验对比分析了图 33 中 3D 打印管式强预冷器、钎焊管式强预冷器和板式强预冷器工作性能开展了研究，实验表明，钎焊管式强预冷器单位体积换热功率最大，可达到 13.8 MW/m^3。



(a) 3D打印管式　(b) 钎焊管式　(c) 扩散焊板式

图 33　不同形式紧凑式换热器[260]

Fig. 33　Different types of compact heat exchangers[260]

上述实验综合表明，当前适用于强预冷发动机的紧凑快速强换热器已完成了多种常用工质下高温来流及匹配环境实验验证，充分验证了紧凑快速强换热效果及拓展涡轮发动机工作包线、提升性能的可行性。

5　小　结

1）当前，围绕微小尺度通道流动换热、横掠管束快速强换热等机理问题，已开展了大量理论及实验研究。前者集中于由尺度效应带来的共性问题及超临界工质物性变化和相应的换热恶化/强化等强预冷背景性问题的研究；后者集中于叉排管束流动结构演化、涡脱落频率与激振/共振等问题研究。除微小尺度表面粗糙度摩阻规律等问题尚存在争议外，绝大部分研究结果均对紧凑快速强换热器基本单元设计优化具有重要指导价值。进一步地，在厘清流动换热机理基础上，需结合强预冷发动机实际工作环境及工质，进一步完善高精度经验关联数据库。

2）紧凑快速强换热器由于尺度跨度大、工质复杂且部分存在易燃易爆风险，其理论研究及验证过程中存在成本较高、风险较大等难题。基于多孔介质模型等简化模型的强换热器快速仿真分析方法及相似分析方法可有效降低前期研究成本及风险，是研究耦合及匹配工作环境下紧凑快速强换热性能的重要手段。

3）提升设计精度及复杂工程应用环境下实际工作性能是紧凑快速强换热设计的重要目标，当前所发展的基于 1D/2D 离散的精细化设计方法及考虑加工误差、参数波动等因素的不确定性设计方法可显著提升设计性能准确性及实际性能鲁棒性。此外，针对 PCHE、TPMS 等换热器形式的拓扑优化具有相当的潜力，但其在强预冷发动机的工程化应用仍显著依赖加工制造水平。

4）实验验证是综合检验紧凑快速强换热技术性能及其在强预冷发动机中适用性的重要手段，当前围绕不同工质、不同构型紧凑快速强换热器已开展了大量常温实验、模拟高马赫数工况的高温来流实验验证，充分验证了紧凑快速强换热技术在拓展涡轮发动机工作速域、支撑强预冷发动机技术发展的可行性。未来需推进真实装机匹配环境下全温全压实验验证、强预冷器高马赫数飞行实验验证等工作，以进一步提升紧凑快速强换热技术成熟度。

5）北航团队近年来围绕紧凑快速强换热技术开展了大量理论分析、方法研究及实验验证工作，取得了多项先进的技术成果。目前，已逐步进入技术成熟度提升及整机匹配验证阶段，为强预冷发动机技术发展及应用提供核心支撑。

6　结 束 语

本文详细介绍了国内外围绕紧凑快速强换热技术所开展的研究工作，主要包括微小尺度紧凑强换热机理、紧凑快速强换热分析及设计方法、紧凑快速强换热实验技术等方面，这些工作为以强预冷器、氢-氦/氦-氦换热器等强预冷发动机内部多工质换热器技术发展及研制提供了关键支撑，本节结合强预冷发动机对于复杂极端热环境下强换热需求探讨了这方面研究所面临的挑战及未来展望。

1）完善紧凑快速强换热高精度经验关联数据库。强预冷发动机内部强换热过程具有微小尺度、空间超紧凑、极大温度梯度、流动三维性强、强变物性等典型特征，导致传统针对常规尺度及工况发展的经验关联式与紧凑快速强换热器误差较大（前期研究表明基于传统经验关联式的计算结果与实验验证误差达到 25% 以上），将直接影响紧凑快速强换热器设计准确性，进而影响其在复杂环境下工作性能，为此有必要先从机理角度出发发展流动换热调控分析模型以指导精细化设计方向，进而通过

仿真分析及缩尺模型、全尺寸模型实验等手段，针对不同构型、不同工质等特征下紧凑快速强换热过程构建更为完善的流动换热数据库，形成工程可用的基于无量纲参数流动换热经验关联式，以指导紧凑快速强换热高水平设计。

2）探究真实发动机匹配非均匀环境下紧凑快速强换热器综合性能。强预冷器等紧凑快速强换热器在匹配环境下将受上下游部件（如压气机、进气道等）及安装环境影响，其非均匀及复杂压力梯度特征将显著影响紧凑快速强换热器内部流动均匀性，进而影响其流动换热性能，前期研究表明在进气道-强预冷器-压气机匹配环境下，一体化流道设计将显著影响强预冷器轴向不同位置空气流量分配，进而影响实际换热性能，因此，需着重关注真实匹配非均匀环境下流动换热现象及综合性能。

3）发展并完善耦合结构强度的紧凑快速强换热设计技术。由于强预冷发动机用紧凑快速强换热器追求极致的高功重比及紧凑度，其存在突出的薄壁微小通道和密集阵列特征，强温度梯度将带来显著的换热器变形热协调问题，反之，结构变形问题也将使真实流动换热过程有别于理想设计状态，这都将给紧凑快速强换热器实际运行过程带来严峻挑战。此外，对于超大长径比的微细管阵列结构而言，较高速气流将会带来显著的流致振动问题，对其流动换热及结构强度均会带来直接影响，应在设计及加工制造过程中重点关注。

4）关注融合人工智能的紧凑快速强换热分析及设计。随着紧凑快速强换热器工况复杂性提升，以及拓扑优化、TPMS等新手段、新构型的引入，其流动及结构复杂度将显著提升，虽然这种超越传统方法的优化手段和设计理念已经表现出强大的潜力，但同时带来的问题是换热器性能参数复杂度增加。人工智能方法在复杂换热器分析及优化中已展现出巨大的应用潜力，需进一步结合紧凑快速强换热实际需求开展相关手段及方法研究。

5）发展完善的紧凑快速强换热器实验方法及验证平台。模拟真实高温及流动环境的实验是进行紧凑快速强换热研究和获取关键数据来源最为重要的手段，而高精度实验测试方法和能够模拟真实匹配环境的实验平台是开展相关研究的核心支撑。当前虽已初步完成了紧凑快速强换热器在高温环境下性能实验验证，证实了紧凑快速强换热在空天动力方面重要的工程意义，但在紧凑快速强换热器真实装机匹配环境下全温全压实验验证、狭窄空间复杂结构流动换热精细化流场测试方法等方面仍存在不足，因此需进一步结合强预冷发动机实际工作条件及验证需求，开展更为深入的实验测试手段发展及条件建设。

参考文献（References）

[1] 邹正平, 刘火星, 唐海龙, 等. 高超声速航空发动机强预冷技术研究[J]. 航空学报, 2015, 36(8): 2544-2562.
Zou Z P, Liu H X, Tang H L, et al. Precooling technology study of hypersonic aeroengine[J]. Acta Aeronautica et Astronautica Sinica, 2015, 36(8): 2544-2562(in Chinese).

[2] Dong Z Y, Sun M B, Wang Z G, et al. Survey on key techniques of rocket-based combined-cycle engine in ejector mode[J]. Acta Astronautica, 2019, 164: 51-68.

[3] Sam L, Idithsaj P T, Nair P P, et al. Prospects for scramjet engines in reusable launch applications: a review[J]. International Journal of Hydrogen Energy, 2023, 48(92): 36094-36111.

[4] 李韦康, 张乐临, 申丽艳, 等. 预冷涡轮组合循环发动机及其关键技术研究综述[J]. 航空发动机, 2024, 50(4): 16-29.
Li W K, Zhang L L, Shen L Y, et al. Overview of precooled turbine-based combined cycle engine and its key technologies[J]. Aeroengine, 2024, 50(4): 16-29(in Chinese).

[5] 邹正平, 王一帆, 额日其太, 等. 高超声速强预冷航空发动机技术研究进展[J]. 航空发动机, 2021, 47(4): 8-21.
Zou Z P, Wang Y F, Eri Q T, et al. Research progress on hypersonic precooled airbreathing engine technology[J]. Aeroengine, 2021, 47(4): 8-21(in Chinese).

[6] 侯晓. 组合循环发动机技术研究进展[J]. 航空学报, 2023, 44(21): 28-40.
Hou X. Research progress in combined cycle engines[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(21): 28-40(in Chinese).

[7] Dai J, Zuo Q R. Key technologies for thermodynamic cycle of precooled engines: a review[J]. Acta Astronautica, 2020, 177: 299-312.

[8] 王一帆, 邹正平, 陈懋章. 高超声速强预冷发动机热力循环研究进展[J]. 航空学报, 2023, 44(21): 41-77.
Wang Y F, Zou Z P, Chen M Z. Progress in thermodynamic cycle research of hypersonic precooled engine[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(21): 41-77(in Chinese).

[9] Wang Z G, Wang Y, Zhang J Q, et al. Overview of the key technologies of combined cycle engine precooling systems and the advanced applications of micro-channel heat transfer[J]. Aerospace Science and Technology, 2014, 39: 31-39.

[10] Meng B, Wan M, Zhao R, et al. Micromanufacturing technologies of compact heat exchangers for hypersonic precooled airbreathing propulsion: a review[J]. Chinese Journal of Aeronautics, 2021, 34(2): 79-103.

[11] 邹正平, 王一帆, 杜鹏程, 等. 强预冷发动机新型热力循环布局及性能分析[J]. 火箭推进, 2021, 47(6): 62-75.
Zou Z P, Wang Y F, Du P C, et al. Thermodynamic performance analysis of anovel precooled airbreathing engine layout[J]. Journal of Rocket Propulsion, 2021, 47(6): 62-75(in Chinese).

[12] Rudakov A S, Balepin V V. Propulsion systems with air precooling for aerospaceplane[J]. SAE Technical Paper Series, 1991, 1: 911182.

[13] Sato T, Kobayashi H, Tanatsugu N, et al. Development study of the precooler of the ATREX engine: AIAA 2003-6985[R]. Reston:

AIAA, 2003.

[14] Taguchi H, Sato T, Kobayashi H, et al. Design study on a small pre-cooled turbojet engine for flight experiments: AIAA 2005-3419[R]. Reston: AIAA, 2005.

[15] Varvill R, Bond A. A comparison of propulsion concepts for SSTO reusable launchers[J]. Journal of the British Interplanetary Society, 2003, 56: 108-117.

[16] Feast S. The synergetic air-breathing rocket engine (SABRE) development status update[C]//Proceedings of the International Astronautical Congress. Glasgow: IAF, 2020.

[17] Jivraj F, Varvill R, Bond A, et al. The Scimitar precooled Mach 5 engine[C]//Proceedings of 2nd European Conference for Aerospace Sciences. Brussels: EUCASS, 2007: 1-10.

[18] Zou Z P, Wang Y F, Du P C, et al. A novel simplified precooled airbreathing engine cycle: thermodynamic performance and control law[J]. Energy Conversion and Management, 2022, 258: 115472.

[19] 赵聪,徐阳,张志刚.剑指“空天往返”毫秒降温 1 000℃[N].中国航天报,2021-12-30(2).
Zhao C, Xu Y, Zhang Z G. Aiming at "Aerospace Round-Trip": millisecond cooling of 1000 ℃[N]. China Space News, 2021-12-30(2)(in Chinese).

[20] 张蒙正, 刘典多, 马海波, 等. PATR 发动机关键技术与性能提升途径初探[J]. 推进技术, 2018, 39(9): 1921-1927.
Zhang M Z, Liu D D, Ma H B, et al. Preliminary analysis on critical technology and performance improvement of PATR engine[J]. Journal of Propulsion Technology, 2018, 39(9): 1921-1927(in Chinese).

[21] VILLACÉ V F. Simulation, design and analysis of air-breathing combined-cycle engines for high speed propulsion[D]. Madrid: Universidad Politecnica de Madrid, 2013.

[22] 玉选斐. 预冷吸气式组合推进系统热力循环及控制规律研究[D]. 哈尔滨: 哈尔滨工业大学, 2020.
Yu X F. Research on thermodynamic cycle and control law of precooled airbreathing propulsion system[D]. Harbin: Harbin Institute of Technology, 2020(in Chinese).

[23] Harada K, Tanatsugu N, Sato T. Development study of a precooler for the air-turboramjet expander-cycle engine[J]. Journal of Propulsion and Power, 2001, 17(6): 1233-1238.

[24] Wang C S, Eri Q T, Wang Y, et al. Multi-objective aerodynamic optimization of an axisymmetric variable-geometry inlet with a Mach 5 design point[J]. Aerospace Science and Technology, 2023, 136: 108189.

[25] Wang Y, Lin Y Z, Eri Q, et al. Flow and thrust characteristics of an expansion-deflection dual-bell nozzle[J]. Aerospace Science and Technology, 2022, 123: 107464.

[26] Bai N J, Fan W J, Zhang R C. A mixing enhancement mechanism for a hydrogen transverse jet coupled with a shear layer for gas turbine combustion[J]. Physics of Fluids, 2023, 35(4): 045111.

[27] Varvill R, Paniagua G, Kato H, et al. Design and testing of the contra-rotating turbine for the Scimitar precooled Mach 5 cruise engine[C]//Proceedings of the 59th International Astronautical Congress. Glasgow: IAF, 2008: 6531-6540.

[28] Chen Y M, Zou Z P, Fu C. A study on the similarity method for helium compressors[J]. Aerospace Science and Technology, 2019, 90: 115-126.

[29] Frazer-Nash. Pioneering new programme to realise lift-off for horizontal space launch[EB/OL]. (2025-07-15) [2025-07-16]. https://www.fnc.co.uk/discover-frazer-nash/news/pioneering-new-programme-to-realise-lift-off-for-horizontal-space-launch/.

[30] Balepin V. High speed propulsion cycles[J]. Advances on Propulsion Technology for High-Speed Aircraft, 2008, 2: 1-32.

[31] 王一帆. 高超声速预冷发动机数学模型及多构型热力循环性能研究[D]. 北京: 北京航空航天大学, 2022.
Wang Y F. Modeling method and thermodynamic analysis of multi-layout airbreathing precooled engine[D]. Beijing: Beihang University, 2022(in Chinese).

[32] Hempsell M, Bond A, Bond R, et al. Progress on the SKYLON and SABRE development programme[C]//Proceedings of the 62nd International Astronautical Congress. [S.l.]: IAC, 2011: 7519-7525.

[33] Varvill R, Duran I, Kirk A, et al. Sabre technology development status and update[C]//Proceedings of the 8th European Conference for Aeronautics and Space Sciences. [S.l.]: EUCASS, 2019.

[34] 吕多, 陆海鹰, 周建军, 等. 临近空间飞行器推进系统预冷器关键技术[J]. 航空学报, 2016, 37(增刊 1): 119-126.
Lü D, Lu H Y, Zhou J J, et al. Key technology for pre-cooler in near space vehicle propulsion system[J]. Acta Aeronautica et Astronautica Sinica, 2016, 37(Sup 1): 119-126(in Chinese).

[35] Hempsell M. Progress on the SKYLON and SABRE[C]//Proceedings of the International Astronautical Congress. Pairs: IAF, 2013, 11: 8427-8440.

[36] Varvill R. Heat exchanger development at Reaction Engines Ltd[J]. Acta Astronautica, 2010, 66(9-10): 1468-1474.

[37] Taguchi H, Harada K, Kobayashi H, et al. Mach 4 simulating experiment of pre-cooled turbojet engine using liquid hydrogen[J]. Aerospace, 2022, 9(1): 39.

[38] 李辉, 汝卓霖, 邹正平, 等. 微小尺度通道内超临界甲烷传热特性研究[J]. 南京航空航天大学学报 (自然科学版), 2021, 53(4): 513-520.
Li H, Ru Z L, Zou Z P, et al. Investigation on heat transfer characteristic of supercritical methane in a microtube[J]. Journal of Nanjing University of Aeronautics & Astronautics (Natural Science Edition), 2021, 53(4): 513-520(in Chinese).

[39] 郭楚微, 刘晓波. 预冷高超声速发动机发展及应用研究[J]. 飞航导弹, 2020(3): 92-96.
Guo C W, Liu X B. Development and application of precooled hypersonic engine[J]. Aerodynamic Missile Journal, 2020(3): 92-96 (in Chinese).

[40] Kandlikar S G, Grande W J. Evolution of microchannel flow passages: thermohydraulic performance and fabrication technology[J]. Heat Transfer Engineering, 2003, 24(1): 3-17.

[41] Schmidt E. Wärmetransport durch natürliche konvektion in stoffen Bei kritischem zustand[J]. International Journal of Heat and Mass Transfer, 1960, 1(1): 92-101.

[42] Wang B X, Peng X F. Experimental investigation on liquid forced-convection heat transfer through microchannels[J]. International Journal of Heat and Mass Transfer, 1994, 37: 73-82.

[43] Yang C Y. Friction characteristics of water, R-134a, and air in small tubes[J]. Microscale Thermophysical Engineering, 2003, 7(4): 335-348.

[44] Yang C Y, Chen C W, Lin T Y, et al. Heat transfer and friction

characteristics of air flow in microtubes[J]. Experimental Thermal and Fluid Science, 2012, 37: 12-18.

[45] Lin T Y, Kandlikar S G. An experimental investigation of structured roughness effect on heat transfer during single-phase liquid flow at microscale[J]. Journal of Heat Transfer, 2012, 134(10): 101701.

[46] Kleinstreuer C, Koo J. Computational analysis of wall roughness effects for liquid flow in micro-conduits[J]. Journal of Fluids Engineering, 2004, 126(1): 1-9.

[47] Guo L, Xu H J, Gong L. Influence of wall roughness models on fluid flow and heat transfer in microchannels[J]. Applied Thermal Engineering, 2015, 84: 399-408.

[48] Kadivar M, Tormey D, Mcgranaghan G. CFD of roughness effects on laminar heat transfer applied to additive manufactured minichannels[J]. Heat and Mass Transfer, 2024, 60(12): 1915-1929.

[49] Peng X F, Peterson G P, Wang B X. Frictional flow characteristics of water flowing through rectangular microchannels[J]. Experimental Heat Transfer, 1994, 7(4): 249-264.

[50] Salman B H, Mohammed H A, Munisamy K M, et al. Characteristics of heat transfer and fluid flow in microtube and microchannel using conventional fluids and nanofluids: a review[J]. Renewable and Sustainable Energy Reviews, 2013, 28: 848-880.

[51] Mala G M, Li D Q. Flow characteristics of water in microtubes[J]. International Journal of Heat and Fluid Flow, 1999, 20(2): 142-148.

[52] Li Z, He Y L, Tang G H, et al. Experimental and numerical studies of liquid flow and heat transfer in microtubes[J]. International Journal of Heat and Mass Transfer, 2007, 50(17-18): 3447-3460.

[53] Qu W L, Mudawar I. Experimental and numerical study of pressure drop and heat transfer in a single-phase micro-channel heat sink[J]. International Journal of Heat and Mass Transfer, 2002, 45(12): 2549-2565.

[54] Sara O N, Ergu B, Arzutug M E, et al. Experimental study of laminar forced convective mass transfer and pressure drop in microtubes[J]. International Journal of Thermal Sciences, 2009, 48(10): 1894-1900.

[55] Maranzana G, Perry I, Maillet D. Mini- and micro-channels: influence of axial conduction in the walls[J]. International Journal of Heat and Mass Transfer, 2004, 47(17-18): 3993-4004.

[56] Kroeger P G. Performance deterioration in high effectiveness heat exchangers due to axial heat conduction effects[C]//Proceedings of the 1966 Cryogenic Engineering Conference University of Colorado Engineering Research Center and Cryogenics Division NBS Institute for Materials Research Boulder. Berlin: Springer, 1967: 363-372.

[57] Lelea D. The heat transfer and fluid flow of a partially heated microchannel heat sink[J]. International Communications in Heat and Mass Transfer, 2009, 36(8): 794-798.

[58] Moharana M K, Agarwal G, Khandekar S. Axial conduction in single-phase simultaneously developing flow in a rectangular mini-channel array[J]. International Journal of Thermal Sciences, 2011, 50(6): 1001-1012.

[59] Quintero A E, Vera M. Laminar counterflow parallel-plate heat exchangers: an exact solution including axial and transverse wall conduction effects[J]. International Journal of Heat and Mass Transfer, 2017, 104: 1229-1245.

[60] Aminuddin M, Zubair S M. Analytical solutions to counter-flow heat exchanger subjected to external heat flux and axial conduction[J]. International Journal of Refrigeration, 2017, 74: 22-37.

[61] Nellis G F. A heat exchanger model that includes axial conduction, parasitic heat loads, and property variations[J]. Cryogenics, 2003, 43(9): 523-538.

[62] Baek S, Kim J H, Jeong S, et al. Development of highly effective cryogenic printed circuit heat exchanger (PCHE) with low axial conduction[J]. Cryogenics, 2012, 52(7-9): 366-374.

[63] Zhang Z, Zhao C R, Yang X T, et al. Influences of tube wall on the heat transfer and flow instability of various supercritical pressure fluids in a vertical tube[J]. Applied Thermal Engineering, 2019, 147: 242-250.

[64] Liu Z B, He Y L, Yang Y F, et al. Experimental study on heat transfer and pressure drop of supercritical CO_2 cooled in a large tube[J]. Applied Thermal Engineering, 2014, 70(1): 307-315.

[65] Dang C B, Hihara E. In-tube cooling heat transfer of supercritical carbon dioxide. Part 1. experimental measurement[J]. International Journal of Refrigeration, 2004, 27(7): 736-747.

[66] Wang L, Pan Y C, Der Lee J, et al. Experimental investigation in the local heat transfer of supercritical carbon dioxide in the uniformly heated horizontal miniature tubes[J]. International Journal of Heat and Mass Transfer, 2020, 159: 120136.

[67] Gu H F, Li H Z, Wang H J, et al. Experimental investigation on convective heat transfer from a horizontal miniature tube to methane at supercritical pressures[J]. Applied Thermal Engineering, 2013, 58(1-2): 490-498.

[68] Votta R, Battista F, Salvatore V, et al. Experimental investigation of transcritical methane flow in rocket engine cooling channel[J]. Applied Thermal Engineering, 2016, 101: 61-70.

[69] Hall W B, Jackson J D. Laminarization of a turbulent pipe flow by buoyancy forces[C]//Proceedings of the 11th ASME-AIChE National Heat Transfer Conference. Minneapolis: American Society of Mechanical Engineers, 1969.

[70] Bourke P J, Pulling D J, Gill L E, et al. Forced convective heat transfer to turbulent CO_2 in the supercritical region[J]. International Journal of Heat and Mass Transfer, 1970, 13(8): 1339-1348.

[71] Zhu J Q, Zhao C F, Cheng Z Y, et al. Experimental investigation on heat transfer of n-decane in a vertical square tube under supercritical pressure[J]. International Journal of Heat and Mass Transfer, 2019, 138: 631-639.

[72] Liu J M, Jin Y X, Zhao P H, et al. Analysis of heat transfer of supercritical water by direct numerical simulation of heated upward pipe flows[J]. International Journal of Thermal Sciences, 2019, 138: 206-218.

[73] Kral L D, Fasel H F. Direct numerical simulation of passive control of three-dimensional phenomena in boundary-layer transition using wall heating[J]. Journal of Fluid Mechanics, 1994, 264: 213-254.

[74] Wall D P, Wilson S K. The linear stability of channel flow of fluid with temperature-dependent viscosity[J]. Journal of Fluid Mechanics, 1996, 323: 107-132.

[75] Kurganov V A, Kaptil'Ny A G. Velocity and enthalpy fields and eddy diffusivities in a heated supercritical fluid flow[J]. Experimental Thermal and Fluid Science, 1992, 5(4): 465-478.

[76] Sun J, Xie G N, Jin P H, et al. Buoyancy effect on entropy generation and heat transfer performances of transcritical liquefied natural gas in a printed circuit heat exchanger[J]. International Journal of Thermal Sciences, 2025, 212: 109761.

[77] Cao Y L, Xu R N, Yan J J, et al. Direct numerical simulation of convective heat transfer of supercritical pressure in a vertical tube with buoyancy and thermal acceleration effects[J]. Journal of Fluid Mechanics, 2021, 927: A29.

[78] Zou Z P, Zhang J Q, Fu C, et al. Heat transfer enhancement mechanism of supercritical pressure fluids in circular pipes under transverse vibration[J]. International Journal of Heat and Mass Transfer, 2025, 250: 127259.

[79] 付超. 壁面传热环境中流动与换热的交互作用机理[D]. 北京: 北京航空航天大学, 2019.
Fu C. Interaction mechanism between flow and heat transfer under non-adiabatic wall conditions[D]. Beijing: Beihang University, 2019(in Chinese).

[80] Yoo S Y, Kwon H K, Kim J H. A study on heat transfer characteristics for staggered tube banks in cross-flow[J]. Journal of Mechanical Science and Technology, 2007, 21(3): 505-512.

[81] Webber H, Taylor N. Local heat transfer measurements inside a compact heat exchanger: AIAA 2010-4653[R]. Reston: AIAA, 2010.

[82] Xia K Q, Huang S D, Xie Y C, et al. Tuning heat transport via coherent structure manipulation: recent advances in thermal turbulence[J]. National Science Review, 2023, 10(6): nwad012.

[83] Robinson S K. Coherent motions in the turbulent boundary layer [J]. Annual Review of Fluid Mechanics, 1991, 23: 601-639.

[84] Kong H, Choi H, Lee J S. Direct numerical simulation of turbulent thermal boundary layers[J]. Physics of Fluids, 2000, 12(10): 2555-2568.

[85] Kline S J, Reynolds W C, Schraub F A, et al. The structure of turbulent boundary layers[J]. Journal of Fluid Mechanics, 1967, 30(4): 741-773.

[86] Townsend A A. The structure of turbulent shear flow[M]. 2nd ed. Cambridge: Cambridge University Press, 1976.

[87] Hubble D O, Vlachos P P, Diller T E. The role of large-scale vortical structures in transient convective heat transfer augmentation[J]. Journal of Fluid Mechanics, 2013, 718: 89-115.

[88] Moss R W, Oldfield M L G. Effect of free-stream turbulence on flat-plate heat flux signals: spectra and eddy transport velocities[J]. Journal of Turbomachinery, 1996, 118(3): 461-467.

[89] Nix A C, Diller T E, Ng W F. Experimental measurements and modeling of the effects of large-scale freestream turbulence on heat transfer[J]. Journal of Turbomachinery, 2007, 129(3): 542-550.

[90] Green S I. Fluid vortices[M]. Berlin: Springer, 1995.

[91] Küchemann D. Report on the IUTAM symposium on concentrated vortex motions in fluids[J]. Journal of Fluid Mechanics, 1965, 21(1): 1-20.

[92] Crum G F, Hanratty T J. Dissipation of a sheet of heated air in a turbulent flow[J]. Applied Scientific Research, Section A, 1966, 15(1): 177-195.

[93] Wang A B, Trávníček Z, Chia K C. On the relationship of effective Reynolds number and Strouhal number for the laminar vortex shedding of a heated circular cylinder[J]. Physics of Fluids, 2000, 12(6): 1401-1410.

[94] Sabanca M, Durst F. Flows past a tiny circular cylinder at high temperature ratios and slight compressible effects on the vortex shedding[J]. Physics of Fluids, 2003, 15(7): 1821-1829.

[95] Michaux-Leblond N, Bélorgey M. Near-wake behavior of a heated circular cylinder: viscosity-buoyancy duality[J]. Experimental Thermal and Fluid Science, 1997, 15(2): 91-100.

[96] Bhuiyan A A, Islam A S, Amin M R. Numerical prediction of laminar characteristics of fluid flow and heat transfer in finned-tube heat exchangers[J]. Innovative Systems Design and Engineering, 2011, 2(6): 1-12.

[97] Awais M, Bhuiyan A A. Enhancement of thermal and hydraulic performance of compact finned-tube heat exchanger using vortex generators (VGs): a parametric study[J]. International Journal of Thermal Sciences, 2019, 140: 154-166.

[98] Kim J, Choi B I, Kim K, et al. Acoustic resonance and pressure drop through a staggered tube bundle in a compact heat exchanger[J]. International Journal of Heat and Mass Transfer, 2015, 91: 63-67.

[99] 陈一鸣, 李泽鹏, 张俊强, 等. 紧凑式强预冷换热器叉排管束的大涡模拟[J]. 航空动力学报, 2021, 36(4): 701-712.
Chen Y M, Li Z P, Zhang J Q, et al. Large eddy simulation on the staggered tube bundle of the compact precooler[J]. Journal of Aerospace Power, 2021, 36(4): 701-712(in Chinese).

[100] Oeng Ö A, Ziada S. An in-depth study of vortex shedding, acoustic resonance and turbulent forces in normal triangle tube arrays[J]. Journal of Fluids and Structures, 1998, 12(6): 717-758.

[101] Blevins R D. Review of sound induced by vortex shedding from cylinders[J]. Journal of Sound and Vibration, 1984, 92(4): 455-470.

[102] Linton D, Thornber B. Direct numerical simulation of transitional flow in a staggered tube bundle[J]. Physics of Fluids, 2016, 28(2): 024111.

[103] Païdoussis M P. A review of flow-induced vibrations in reactors and reactor components[J]. Nuclear Engineering and Design, 1983, 74(1): 31-60.

[104] 姚李超, 付超, 张俊强, 等. 非定常来流压力下基于 DMD 方法的预冷器换热特性[J]. 航空动力学报, 2020, 35(10): 2064-2077.
Yao L C, Fu C, Zhang J Q, et al. Heat transfer performance of precooler under unsteady inflow pressure condition using dynamic mode decomposition method[J]. Journal of Aerospace Power, 2020, 35(10): 2064-2077(in Chinese).

[105] Jiang H T, Wang X S, Liu Y W. Numerical study of different *Pr* number medium turbulent cross flow in staggered tube bundle[J]. International Journal of Thermal Sciences, 2023, 187: 108206.

[106] Elmekawy A M N, Ibrahim A A, Shahin A M, et al. Performance enhancement for tube bank staggered configuration heat exchanger-CFD study[J]. Chemical Engineering and Processing-Process Intensification, 2021, 164: 108392.

[107] Hu D L, Zhang Q, Song K W, et al. Performance optimization of a wavy finned-tube heat exchanger with staggered curved vortex generators[J]. International Journal of Thermal Sciences, 2023, 183: 107830.

[108] Karali M A, Almohammadi B A, Bin Mahfouz A S, et al. Effect of surfaces roughness of a staggered tube bank in cross flow with air on heat transfer and pressure drop[J]. Case Studies in Thermal Engineering, 2023, 43: 102779.

[109] Moazezi A, Lavasani A M. The impact of spacing ratios on thermal-hydraulic performance of a cam-shaped tube in a four-tube staggered configuration[J]. International Communications in Heat and Mass Transfer, 2025, 164: 108869.

[110] Bond R, Varvill R, Bond A, et al. Pre-cooler design and development[C]//Space Propulsion 2012. Paris: 3AF, 2012.

[111] Lage J. The fundamental theory of flow through permeable media from Darcy to turbulence[M]//Ingham D B. Transport phenomena in porous media. Amsterdam: Elsevier, 1998: 1-30.

[112] Wu J S, Yu B M, Yun M J. A resistance model for flow through porous media[J]. Transport in Porous Media, 2008, 71(3): 331-343.

[113] Chung S S, Kim N, Jo H. Effect of flow distribution on heat transfer in printed circuit heat exchangers using conjugate heat transfer model with a modified porous media approach[J]. Applied Thermal Engineering, 2025, 266: 125540.

[114] Wei X, Shi X J, Ji H H, et al. A simplified numerical method based on anisotropic porous media model for the precooler of the hypersonic precooled engine[J]. Applied Thermal Engineering, 2024, 252: 123709.

[115] ANSYS Inc. ANSYS Fluent user's guide[Z]. Pittsburgh: ANSYS Inc. , 2021: 1137-1139.

[116] Chen W N, Ma Q Y, Liu X Y, et al. Adaptability analysis of flow and heat transfer multi-scale numerical method for printed circuit heat exchanger[J]. Energy, 2024, 311: 133349.

[117] Wu Y B, Xu R N, Wang C, et al. Novel three-equation porous media model for prediction of compact heat exchanger and its case study for microtube bundle precooler[J]. International Journal of Heat and Mass Transfer, 2025, 248: 127177.

[118] Ding W H, Eri Q, Kong B, et al. Numerical investigation of a compact tube heat exchanger for hypersonic pre-cooled aero-engine[J]. Applied Thermal Engineering, 2020, 170: 114977.

[119] ANSYS Inc. ANSYS Fluent theory guide[Z]. Pittsburgh: ANSYS Inc. , 2021: 202-204.

[120] Wang C S, Eri Q, Wang Y, et al. Novel three-dimensional simplified numerical simulation method of a compact precooler for hypersonic precooled air-breathing engine[J]. Applied Thermal Engineering, 2023, 230: 120827.

[121] He K, Li H, Wang Y F, et al. A novel simplified simulation method for the precooler based on the virtual source term: modeling, validation and application[J]. Applied Thermal Engineering, 2024, 254: 123933.

[122] 薛亮波, 孙波, 卓长飞, 等. 预冷对发动机进气道流动特性影响的数值模拟研究[J]. 推进技术, 2020, 41(6): 1227-1236.
Xue L B, Sun B, Zhuo C F, et al. Numerical simulation study of effects of precooling on flow characteristics of engine inlets[J]. Journal of Propulsion Technology, 2020, 41(6): 1227-1236(in Chinese).

[123] 王海, 孙波, 卓长飞, 等. 预冷发动机进气道节流特性数值研究[J]. 航空动力学报, 2021, 36(3): 553-563.
Wang H, Sun B, Zhuo C F, et al. Numerical investigation on throttle characteristics of precooled engine inlets[J]. Journal of Aerospace Power, 2021, 36(3): 553-563(in Chinese).

[124] 杨世铭, 陶文铨. 传热学[M]. 4 版. 北京: 高等教育出版社, 2006.
Yang S M, Tao W Q. Heat transfer[M]. 4th ed. Beijing: Higher Education Press, 2006(in Chinese).

[125] Wang H, Lu D G, Wang Y T. Discussion on the fluid-to-fluid scaling of heat transfer at supercritical pressures[J]. Annals of Nuclear Energy, 2018, 121: 108-117.

[126] Pucciarelli A, Ambrosini W. A successful general fluid-to-fluid similarity theory for heat transfer at supercritical pressure[J]. International Journal of Heat and Mass Transfer, 2020, 159: 120152.

[127] Azih C, Yaras M I. Similarity criteria for modeling mixed-convection heat transfer in ducted flows of supercritical fluids[J]. Journal of Heat Transfer, 2017, 139(12): 122501.

[128] Pioro I L, Duffey R B. Heat transfer and hydraulic resistance at supercritical pressures in power engineering applications[M]. New York: ASME Press, 2007.

[129] Cheng X, Liu X J, Gu H Y. Fluid-to-fluid scaling of heat transfer in circular tubes cooled with supercritical fluids[J]. Nuclear Engineering and Design, 2011, 241(2): 498-508.

[130] Zahlan H, Groeneveld D C, Tavoularis S. Fluid-to-fluid scaling for convective heat transfer in tubes at supercritical and high subcritical pressures[J]. International Journal of Heat and Mass Transfer, 2014, 73: 274-283.

[131] Ambrosini W. Discussion of similarity principles for fluid-to-fluid scaling of heat transfer behaviour at supercritical pressures[J]. Nuclear Engineering and Design, 2011, 241(12): 5149-5173.

[132] 李辉. 物性与热力状态强耦合的微小尺度换热器设计方法及验证[D]. 北京: 北京航空航天大学, 2023.
Li H. Design method and its validation for micro-scale heat exchangers with strong coupling between thermophysical properties and thermodynamic states[D]. Beijing: Beihang University, 2023 (in Chinese).

[133] Kimura H, Aritomi M. Method for evaluation of effective forced-convective heat transfer rate associated with chemical heat absorption of heated fluid mixture: Part II[J]. International Journal of Heat and Mass Transfer, 2011, 54(11-12): 2703-2712.

[134] Yu W L, Zhou W X, Jia Z J, et al. Characteristics of scramjet regenerative cooling with endothermic chemical reactions[J]. Acta Astronautica, 2022, 195: 1-11.

[135] Xu S, Feng Y, Zhang Y F, et al. A new modeling method to evaluate effective forced-convective heat transfer rate of pipe flow with endothermic chemical reactions[J]. Physics of Fluids, 2023, 35(10): 107130.

[136] Jackson J D. Fluid flow and convective heat transfer to fluids at supercritical pressure[J]. Nuclear Engineering and Design, 2013, 264: 24-40.

[137] 张议丰. 超临界碳氢燃料蒸汽重整反应与流动换热的耦合特性研究[D]. 哈尔滨: 哈尔滨工业大学, 2024.
Zhang Y F. Investigation on the coupling characteristics of steam reforming reaction and flow heat transfer for supercritical hydrocarbon fuel[D]. Harbin: Harbin Institute of Technology, 2024(in Chinese).

[138] Dittus F W. Heat transfer in automobile radiators of the tubular type[M]. Oakland: University of California Press, 1930: 443-461.

[139] Dittus F W, Boelter L M K. University of California (Berkley) publications on engineering, vol.2[M]. Berkley: University of California, 1930.

[140] Taylor M F. Prediction of friction and heat-transfer coefficients with large variations in fluid properties: NASA-TM-X-2145[R].

Washington, D. C. : NASA, 1970.

[141] Gnielinski V. New equations for heat and mass transfer in turbulent pipe and channel flow[J]. International Journal of Chemical Reactor Engineering, 1976, 16(2): 359-368.

[142] Nikitin K, Kato Y, Ngo L. Printed circuit heat exchanger thermal-hydraulic performance in supercritical CO_2 experimental loop [J]. International Journal of Refrigeration, 2006, 29(5): 807-814.

[143] Kim I H, No H C, Lee J I, et al. Thermal hydraulic performance analysis of the printed circuit heat exchanger using a helium test facility and CFD simulations[J]. Nuclear Engineering and Design, 2009, 239(11): 2399-2408.

[144] Deng H W, Zhu K, Xu G Q, et al. Heat transfer characteristics of RP-3 kerosene at supercritical pressure in a vertical circular tube[J]. Journal of Enhanced Heat Transfer, 2012, 19(5): 409-421.

[145] Seo J W, Kim Y H, Kim D, et al. Heat transfer and pressure drop characteristics in straight microchannel of printed circuit heat exchangers[J]. Entropy, 2015, 17(5): 3438-3457.

[146] Wen J, Huang H R, Fu Y C, et al. Heat transfer performance of aviation kerosene RP-3 flowing in a vertical helical tube at supercritical pressure[J]. Applied Thermal Engineering, 2017, 121: 853-862.

[147] Ren Z, Zhao C R, Jiang P X, et al. Investigation on local convection heat transfer of supercritical CO_2 during cooling in horizontal semicircular channels of printed circuit heat exchanger[J]. Applied Thermal Engineering, 2019, 157: 113697.

[148] Liu S H, Huang Y P, Wang J F, et al. Experimental study of thermal-hydraulic performance of a printed circuit heat exchanger with straight channels[J]. International Journal of Heat and Mass Transfer, 2020, 160: 120109.

[149] Katz A, Aakre S R, Anderson M H, et al. Experimental investigation of pressure drop and heat transfer in high temperature supercritical CO_2 and helium in a printed-circuit heat exchanger[J]. International Journal of Heat and Mass Transfer, 2021, 171: 121089.

[150] Zou Z P, Wang Y F, Li H, et al. Thermal-hydraulic characteristics of a PCHE with zigzag microchannel for hypersonic precooled aero-engines: an experimental study[J]. Experimental Heat Transfer, 2023, 36(7): 1012-1033.

[151] Fu Y C, Liu W T, Wang J, et al. Experimental investigation on heat transfer enhancement of supercritical pressure aviation kerosene in tubular laminar flow by vibration[J]. Applied Thermal Engineering, 2024, 257: 124206.

[152] Colburn A P. A method of correlating forced convection heat-transfer data and a comparison with fluid friction[J]. Transactions of the American Institute of Chemical Engineers, 1933, 29: 174-210.

[153] Žukauskas A. Heat transfer from tubes in crossflow[M]//Hartnett J P, Irvine Jr T F. Advances in heat transfer. Volume 8. Amsterdam: Elsevier, 1972: 93-160.

[154] Kays W M, London A L. Compact heat exchangers[M]. New York: McGraw-Hill, 1984.

[155] Chen C J, Wung T S. Finite analytic solution of convective heat transfer for tube arrays in crossflow: Part II: heat transfer analysis[J]. Journal of Heat Transfer, 1989, 111(3): 641-648.

[156] Khan W. Modeling of fluid flow and heat transfer for optimization of pin-fin heat sinks[D]. Waterloo: University of Waterloo, 2004.

[157] Khan W A, Culham J R, Yovanovich M M. Convection heat transfer from tube banks in crossflow: Analytical approach[J]. International Journal of Heat and Mass Transfer, 2006, 49(25-26): 4831-4838.

[158] Li N, Zhao Y, Wang H, et al. Thermal and hydraulic performance of a compact precooler with mini-tube bundles for aero-engine[J]. Applied Thermal Engineering, 2022, 200: 117656.

[159] 李钢, 马同玲, 刘洪涛, 等. Zukauskas 关系式用于微细管式预冷器适用性的评价与分析[J]. 推进技术, 2021, 42(12): 2754-2761.
Li G, Ma T L, Liu H T, et al. Evaluation and analysis of applicability of Zukauskas correlation to microtubular pre-cooler[J]. Journal of Propulsion Technology, 2021, 42(12): 2754-2761(in Chinese).

[160] Jia Y T, Xu X H, Li Y L, et al. Experimental studies on frost and defrost of fine tube bundles under coolant temperature between −20 and −58℃[J]. International Journal of Heat and Mass Transfer, 2018, 116: 617-620.

[161] 覃扬佳, 庄来鹤, 刘启航, 等. 横掠叉排管束传热特性的数值研究[J]. 航空动力学报, 2023, 38(11): 2718-2728.
Qin Y J, Zhuang L H, Liu Q H, et al. Numerical investigation on heat transfer characteristics of staggered tube bundles in cross flow[J]. Journal of Aerospace Power, 2023, 38(11): 2718-2728(in Chinese).

[162] Bacellar D, Aute V, Huang Z W, et al. Airside friction and heat transfer characteristics for staggered tube bundle in crossflow configuration with diameters from 0.5 mm to 2.0 mm[J]. International Journal of Heat and Mass Transfer, 2016, 98: 448-454.

[163] Murray J J, Guha A, Bond A. Overview of the development of heat exchangers for use in air-breathing propulsion pre-coolers[J]. Acta Astronautica, 1997, 41(11): 723-729.

[164] Li H, Zou Z P, Liu Y M. A refined design method for precoolers with consideration of multi-parameter variations based on low-dimensional analysis[J]. Chinese Journal of Aeronautics, 2022, 35(3): 329-344.

[165] Saeed M, Berrouk A S, Siddiqui M S, et al. Numerical investigation of thermal and hydraulic characteristics of sCO_2-water printed circuit heat exchangers with zigzag channels[J]. Energy Conversion and Management, 2020, 224: 113375.

[166] 陈一鸣. 紧凑快速换热器循环实验系统建设及实验研究[D]. 北京: 北京航空航天大学, 2021.
Chen Y M. Test platform and experimental test on compact heat exchanger[D]. Beijing: Beihang University, 2021(in Chinese).

[167] Chen Y M, Zou Z P, Liu H X, et al. Verification at Mach 4 heat conditions of an annular microtube-typed precooler for hypersonic precooled engines[J]. Applied Thermal Engineering, 2022, 201: 117742.

[168] 高远, 陈玉春, 史新兴. SABRE 预冷器计算模型及其在整机模型中的应用[J]. 推进技术, 2021, 42(11): 2485-2493.
Gao Y, Chen Y C, Shi X X. SABRE precooler calculation model and its application in engine model[J]. Journal of Propulsion Technology, 2021, 42(11): 2485-2493(in Chinese).

[169] Xu Y S, Wang Y F, Du P C, et al. Improve the thermal performance of precooler by enhancing the utilization of chemical heat sink: refined design method and characteristic analysis[J]. Applied Thermal Engineering, 2024, 254: 123853.

[170] Hoopes K, Sánchez D, Crespi F. A new method for modelling off-design performance of sCO_2 heat exchangers without specifying

detailed geometry[C]//Proceedings of the 5th Supercritical CO_2 Power Cycles Symposium. [S. l. : s. n.], 2016.

[171] 李欢. 高超声速强预冷发动机微小尺度换热器流动与换热实验研究[D]. 北京: 北京航空航天大学, 2021.
Li H. Experimental investigation on flow and heat transfer for micro-scale heat exchangers of hypersonic precooled engines[D]. Beijing: Beihang University, 2021(in Chinese).

[172] Bennett K, Chen Y T. Printed circuit heat exchanger performance analysis using non-uniform segmental design method[J]. Applied Thermal Engineering, 2019, 153: 69-84.

[173] Zhao Z C, Chen X D, Li S L, et al. Methodology of design and analysis on the thermal hydraulic performance of the cross-flow printed circuit heat exchanger[J]. International Journal of Heat and Mass Transfer, 2020, 156: 119756.

[174] Lee S M, Kim K Y. Comparative study on performance of a zigzag printed circuit heat exchanger with various channel shapes and configurations[J]. Heat and Mass Transfer, 2013, 49(7): 1021-1028.

[175] Gupta R, Geyer P E, Fletcher D F, et al. Thermohydraulic performance of a periodic trapezoidal channel with a triangular cross-section[J]. International Journal of Heat and Mass Transfer, 2008, 51(11-12): 2925-2929.

[176] Jeon S, Baik Y J, Byon C, et al. Thermal performance of heterogeneous PCHE for supercritical CO_2 energy cycle[J]. International Journal of Heat and Mass Transfer, 2016, 102: 867-876.

[177] Aneesh A M, Sharma A, Srivastava A, et al. Thermal-hydraulic characteristics and performance of 3D straight channel based printed circuit heat exchanger[J]. Applied Thermal Engineering, 2016, 98: 474-482.

[178] Han Z X, Guo J F, Huai X L. Theoretical analysis of a novel PCHE with enhanced rib structures for high-power supercritical CO_2 Brayton cycle system based on solar energy[J]. Energy, 2023, 270: 126928.

[179] Zhao J X, Liu M Y, Fei J J, et al. Numerical investigation of convective heat transfer enhancement mechanism in printed circuit heat exchangers with transversely corrugated channels[J]. Applied Thermal Engineering, 2025, 278: 127026.

[180] Ali N, Haque I, Alam T, et al. Numerical investigation on heat transfer and flow mechanism in microchannel heat sink having V shape ribs[J]. Case Studies in Thermal Engineering, 2025, 65: 105684.

[181] Yoon S H, No H C, Kang G B. Assessment of straight, zigzag, S-shape, and airfoil PCHEs for intermediate heat exchangers of HTGRs and SFRs[J]. Nuclear Engineering and Design, 2014, 270: 334-343.

[182] Gschwind P, Regele A, Kottke V. Sinusoidal wavy channels with Taylor-Goertler vortices[J]. Experimental Thermal and Fluid Science, 1995, 11(3): 270-275.

[183] Lee S M, Kim K Y. Optimization of zigzag flow channels of a printed circuit heat exchanger for nuclear power plant application[J]. Journal of Nuclear Science and Technology, 2012, 49(3): 343-351.

[184] Ngo T L, Kato Y, Nikitin K, et al. Heat transfer and pressure drop correlations of microchannel heat exchangers with S-shaped and zigzag fins for carbon dioxide cycles[J]. Experimental Thermal and Fluid Science, 2007, 32(2): 560-570.

[185] Meshram A, Jaiswal A K, Khivsara S D, et al. Modeling and analysis of a printed circuit heat exchanger for supercritical CO_2 power cycle applications[J]. Applied Thermal Engineering, 2016, 109: 861-870.

[186] Lee S Y, Park B G, Chung J T. Numerical studies on thermal hydraulic performance of zigzag-type printed circuit heat exchanger with inserted straight channels[J]. Applied Thermal Engineering, 2017, 123: 1434-1443.

[187] Xu T T, Zhao H X, Wang M, et al. Numerical study of thermal-hydraulic performance of a new spiral Z-type PCHE for supercritical CO_2 brayton cycle[J]. Energies, 2021, 14(15): 4417.

[188] Saeed M, Kim M H. Thermal-hydraulic analysis of sinusoidal fin-based printed circuit heat exchangers for supercritical CO_2 Brayton cycle[J]. Energy Conversion and Management, 2019, 193: 124-139.

[189] Kim D E, Kim M H, Cha J E, et al. Numerical investigation on thermal-hydraulic performance of new printed circuit heat exchanger model[J]. Nuclear Engineering and Design, 2008, 238(12): 3269-3276.

[190] Kim T H, Kwon J G, Yoon S H, et al. Numerical analysis of airfoil shaped fin performance in printed circuit heat exchanger in a supercritical carbon dioxide power cycle[J]. Nuclear Engineering and Design, 2015, 288: 110-118.

[191] Li Z, Lu D G, Wang Z C, et al. Analysis on flow and heat transfer performance of SCO_2 in airfoil channels with different fin angles of attack[J]. Energy, 2023, 282: 128600.

[192] Cui X Y, Guo J F, Huai X L, et al. Numerical study on novel airfoil fins for printed circuit heat exchanger using supercritical CO_2[J]. International Journal of Heat and Mass Transfer, 2018, 121: 354-366.

[193] Xu X Y, Ma T, Li L, et al. Optimization of fin arrangement and channel configuration in an airfoil fin PCHE for supercritical CO_2 cycle[J]. Applied Thermal Engineering, 2014, 70(1): 867-875.

[194] Lu Y M, Guo Z P, Gong Y, et al. Optimal study of swordfish fin microchannel heat exchanger for the next generation nuclear power conversion system of lead-based reactor[J]. Annals of Nuclear Energy, 2022, 165: 108679.

[195] James C A, Taylor R P, Hodge B K. The application of uncertainty analysis to cross-flow heat exchanger performance predictions[J]. Heat Transfer Engineering, 1995, 16(4): 50-62.

[196] Cho S M. Uncertainty analysis of heat exchanger thermal-hydraulic designs[J]. Heat Transfer Engineering, 1987, 8(2): 63-74.

[197] Haseler L E, Owen R G, Sardesai R G. The sensitivity of heat exchanger calculations to uncertainties in the physical properties of the process fluids[J]. Proceedings of the Institution of Mechanical Engineers, Part A: Power and Process Engineering, 1983, 197(3): 171-178.

[198] Berryman J E, Himmelblau D M. Influence of stochastic inputs and parameters on heat exchanger design[J]. Industrial & Engineering Chemistry Process Design and Development, 1973, 12(2): 165-171.

[199] Badar M A, Zubair S M, Sheikh A K. Uncertainty analysis of heat-exchanger thermal designs using the Monte Carlo simulation technique[J]. Energy, 1993, 18(8): 859-866.

[200] 武翼飞. 微小尺度流动换热及换热器不确定性设计研究[D]. 北京: 北京航空航天大学, 2016.

Wu Y F. Study on microscale flow with heat transfer and uncertainty design of heat exchangers[D]. Beijing: Beihang University, 2016(in Chinese).

[201] Osher S, Sethian J A. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations[J]. Journal of Computational Physics, 1988, 79(1): 12-49.

[202] Borrvall T, Petersson J. Topology optimization of fluids in Stokes flow[J]. International Journal for Numerical Methods in Fluids, 2003, 41(1): 77-107.

[203] Van Dijk N P, Maute K, Langelaar M, et al. Level-set methods for structural topology optimization: a review[J]. Structural and Multidisciplinary Optimization, 2013, 48(3): 437-472.

[204] Dede E M. Multiphysics topology optimization of heat transfer and fluid flow systems[C]//Proceedings of the COMSOL Users Conference. [S. l.: s. n.], 2009.

[205] Alexandersen J, Aage N, Andreasen C S, et al. Topology optimisation for natural convection problems[J]. International Journal for Numerical Methods in Fluids, 2014, 76(10): 699-721.

[206] Lee K. Topology optimization of convective cooling system designs[D]. Ann Arbor: University of Michigan, 2012.

[207] Yaji K, Yamada T, Yoshino M, et al. Topology optimization in thermal-fluid flow using the lattice Boltzmann method[J]. Journal of Computational Physics, 2016, 307(C): 355-377.

[208] Coffin P, Maute K. A level-set method for steady-state and transient natural convection problems[J]. Structural and Multidisciplinary Optimization, 2016, 53(5): 1047-1067.

[209] Li X H, Zhang L K, Li B T. Heat transfer augmentation in microchannel heat sink based on isogeometric topology optimization framework[J]. Applied Mathematical Modelling, 2022, 104: 163-187.

[210] Li H, Ding X H, Jing D L, et al. Experimental and numerical investigation of liquid-cooled heat sinks designed by topology optimization[J]. International Journal of Thermal Sciences, 2019, 146: 106065.

[211] Qian S H, Lou S X, Ge C L, et al. The influence of temperature dependent fluid properties on topology optimization of conjugate heat transfer[J]. International Journal of Thermal Sciences, 2022, 173: 107424.

[212] Mo X B, Zhi H, Xiao Y Z, et al. Topology optimization of cooling plates for battery thermal management[J]. International Journal of Heat and Mass Transfer, 2021, 178: 121612.

[213] Prussack B, Jentz I, Moreira T A, et al. Thermal and hydraulic performance of volumetrically heated triply periodic minimal surface heaters[J]. Applied Thermal Engineering, 2024, 248: 123291.

[214] Al-Ketan O, Rowshan R, Al-Rub R K A. Topology-mechanical property relationship of 3D printed strut, skeletal, and sheet based periodic metallic cellular materials[J]. Additive Manufacturing, 2018(19): 167-183.

[215] Khalil M, Ali M I H, Khan K A, et al. Forced convection heat transfer in heat sinks with topologies based on triply periodic minimal surfaces[J]. Case Studies in Thermal Engineering, 2022, 38: 102313.

[216] Kaur I, Singh P. Flow and thermal transport characteristics of triply-periodic minimal surface (TPMS)-based gyroid and Schwarz-P cellular materials[J]. Numerical Heat Transfer, Part A: Applications, 2021, 79(8): 553-569.

[217] Huang M Z, Sha W, Xiao M, et al. Multiscale topology optimization of cellular structures with high thermal conductivity and large convective surface area[J]. International Journal of Thermal Sciences, 2024, 201: 109053.

[218] Men Z C, Chen W J, Li Q H, et al. Topology optimization of the IWP triply periodic minimal surfaces (TPMS) heat sink based on porous media effective model[J]. International Journal of Heat and Mass Transfer, 2025, 240: 126657.

[219] Shu Z T, Zhao K, Li H, et al. Multi-scale topology optimization for graded hollow lattice structures with variable wall thickness[J]. Thin-Walled Structures, 2025, 213: 113274.

[220] Huang W S, Ning H Y, Tang G H. Investigation of high-compactness and high-efficiency TPMS precoolers for precooled aeroengine[J]. Energy, 2025, 319: 135014.

[221] Sato T, Tanatsugu N, Kobayashi H, et al. Countermeasures against the icing problem on the ATREX precooler[J]. Acta Astronautica, 2004, 54(9): 671-686.

[222] 岳晓菲. 宽速域吸气式发动机的空气预冷器表面疏水和抑冰机理研究[D]. 长沙: 国防科技大学, 2019.
Yue X F. Mechanism study on surface hydrophobicity and icing inhibition of pre-cooler in air-breathing engine of wide velocity range[D]. Changsha: National University of Defense Technology, 2019(in Chinese).

[223] Sonobe N, Fukiba K, Sato S, et al. Method for defrosting heat exchangers using an air-particle jet[J]. International Journal of Refrigeration, 2015, 60: 261-269.

[224] 艾伦·邦德, 理查德·瓦维尔. 热交换器: 202110296520[P]. 2021-07-30.
Bond A, Varvill R. Heat exchangers: 202110296520[P]. 2021-07-30(in Chinese).

[225] Tang Z F, Liu H X, Ji Z X, et al. Experimental study on frost growth patterns and surface wettability effects of precooler module[J]. Applied Thermal Engineering, 2024, 256: 124146.

[226] 韦宏, 杨学森, 李治达, 等. 无水甲醇质量比对紧凑型预冷器的结霜和抑霜性能的影响[J]. 航空动力学报, 2023, 38(3): 596-606.
Wei H, Yang X S, Li Z D, et al. Effects of mass ratio of anhydrous methanol on frosting and defrosting performance of compact precooler[J]. Journal of Aerospace Power, 2023, 38(3): 596-606(in Chinese).

[227] Wei H, Guo X F, Li Z D, et al. Experimental investigation on effects of freestream humidity on the frosting and defrosting characteristics of microtubule precooler[J]. Applied Thermal Engineering, 2022, 206: 118104.

[228] 刘国栋, 张志刚, 马同玲, 等. 深度预冷发动机地面试验结霜风险及应对方法浅析[C]//第六届空天动力联合会议暨中国航天第三专业信息网第四十二届技术交流会暨 2021 航空发动机技术发展高层论坛. 北京: 北京动力机械研究所, 2022: 85-93.
Liu G D, Zhang Z G, Ma T L, et al. A preliminary analysis of frosting risks and countermeasures in ground tests of deeply precooled engines[C]//Proceedings of the 6th Joint Conference on Aerospace Propulsion, the 42nd Technical Exchange Conference of the 3rd Professional Information Network of China Aerospace, and the 2021 High-Level Forum on Aero-Engine Technology. Beijing: Beijing Power Machinery Institute, 2022: 85-93(in Chinese).

[229] 张鹏腾. 微细管束换热器结霜与抑霜实验研究[D]. 大连: 大连理工大学, 2019.
Zhang P T. Experimental research of frost and defrost on fine tube bundle[D]. Dalian: Dalian University of Technology, 2019(in Chinese).

[230] Boreyko J B, Srijanto B R, Nguyen T D, et al. Dynamic defrosting on nanostructured superhydrophobic surfaces[J]. Langmuir, 2013, 29(30): 9516-9524.

[231] Wen M X, Wang L, Zhang M Q, et al. Antifogging and icing-delay properties of composite micro- and nanostructured surfaces[J]. ACS Applied Materials & Interfaces, 2014, 6(6): 3963-3968.

[232] Subramanyam S B, Kondrashov V, Rühe J, et al. Low ice adhesion on nano-textured superhydrophobic surfaces under supersaturated conditions[J]. ACS Applied Materials & Interfaces, 2016, 8(20): 12583-12587.

[233] 张友法, 张文文, 郑日恒, 等. 高超声速组合发动机预冷器抗结霜涂层技术研究[J]. 推进技术, 2017, 38(2): 463-470.
Zhang Y F, Zhang W W, Zheng R H, et al. Research of anti-frosting coating for pre-cooler of hypersonic combined propulsion[J]. Journal of Propulsion Technology, 2017, 38(2): 463-470(in Chinese).

[234] Murray J J, Hempsell C M, Bond A. An experimental precooler for airbreathing rocket engines[J]. Journal of the British Interplanetary Society, 2001, 54(5/6): 199-209.

[235] Lee H, Ma S, Chen Y M, et al. Experimental study on compact heat exchanger for hypersonic aero-engine[C]//Proceedings of the 21st AIAA International Space Planes and Hypersonics Technologies Conference. Reston: AIAA, 2017: 2333.

[236] Li H, Liu H X, Zou Z P. Experimental study and performance analysis of high-performance micro-channel heat exchanger for hypersonic precooled aero-engine[J]. Applied Thermal Engineering, 2021, 182: 116108.

[237] Yang Y H, Morini G L, Brandner J J. Experimental analysis of the influence of wall axial conduction on gas-to-gas micro heat exchanger effectiveness[J]. International Journal of Heat and Mass Transfer, 2014, 69: 17-25.

[238] Zhao S F, Wang W T, Shao T, et al. Mixing performance and drug nano-particle preparation inside slugs in a gas-liquid microchannel reactor[J]. Chemical Engineering Science, 2013, 100: 456-463.

[239] Liu Y L, Xu G Q, Fu Y C, et al. Airside pressure drop characteristics of three analogous serpentine tube heat exchangers considering heat transfer for aero-engine cooling[J]. Chinese Journal of Aeronautics, 2022, 35(12): 32-46.

[240] Waelchli S, Von Rohr P R. Two-phase flow characteristics in gas-liquid microreactors[J]. International Journal of Multiphase Flow, 2006, 32(7): 791-806.

[241] Shi H, Di Miceli Raimondi N, Cid E, et al. Temperature field acquisition by planar laser induced fluorescence using the two-color/two-dye technique for liquid flows in a millimetric zigzag channel[J]. Chemical Engineering Journal, 2021, 426: 131460.

[242] Xia G D, Chen Z, Cheng L X, et al. Micro-PIV visualization and numerical simulation of flow and heat transfer in three micro pin-fin heat sinks[J]. International Journal of Thermal Sciences, 2017, 119: 9-23.

[243] Lee Y, Lee S, Park H B, et al. PIV visualization and CFD analysis of thermal-hydraulic performance in printed circuit heat exchangers with zigzag channels[J]. Case Studies in Thermal Engineering, 2025, 73: 106446.

[244] Herráez J V, Belda R. A study of free convection in air around horizontal cylinders of different diameters based on holographic interferometry. Temperature field equations and heat transfer coefficients[J]. International Journal of Thermal Sciences, 2002, 41(3): 261-267.

[245] Narayan S, Singh A K, Srivastava A. Interferometric study of natural convection heat transfer phenomena around array of heated cylinders[J]. International Journal of Heat and Mass Transfer, 2017, 109: 278-292.

[246] Xu J L, Gan Y H, Zhang D C, et al. Microscale heat transfer enhancement using thermal boundary layer redeveloping concept[J]. International Journal of Heat and Mass Transfer, 2005, 48(9): 1662-1674.

[247] Eckert E R G, Soehngen E. Distribution of heat-transfer coefficients around circular cylinders in crossflow at Reynolds numbers from 20 to 500[J]. Journal of Fluids Engineering, 1952, 74(3): 343-346.

[248] Fu C, Zou Z P, Liu H X, et al. Experimental study on the flow and heat transfer mechanism of the pre-cooler in the hypersonic aero-engine: AIAA 2017-2417[R]. Reston: AIAA, 2017.

[249] Karniadakis G E, Kevrekidis I G, Lu L, et al. Physics-informed machine learning[J]. Nature Reviews Physics, 2021, 3(6): 422-440.

[250] 张伟伟, 寇家庆, 刘溢浪. 智能赋能流体力学展望[J]. 航空学报, 2021, 42(4): 20-65.
Zhang W W, Kou J Q, Liu Y L. Prospect of artificial intelligence empowered fluid mechanics[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(4): 20-65(in Chinese).

[251] Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations [J]. Journal of Computational Physics, 2019, 378: 686-707.

[252] Arzani A, Wang J X, D'Souza R M. Uncovering near-wall blood flow from sparse data with physics-informed neural networks[J]. Physics of Fluids, 2021, 33(7): 071905.

[253] Cai S Z, Mao Z P, Wang Z C, et al. Physics-informed neural networks (PINNs) for fluid mechanics: a review[J]. Acta Mechanica Sinica, 2021, 37(12): 1727-1738.

[254] Varvill R, Bond A. The SKYLON spaceplane-process to realisation[J]. Journal of the British Interplanetary Society, 2008, 61: 412-418.

[255] Kojima T, Kobayashi H, Taguchi H, et al. Design study of hypersonic components for precooled turbojet engine: AIAA 2008-2504[R]. Reston: AIAA, 2008.

[256] Sato T, Taguchi H, Kobayashi H, et al. Development study of pre-cooled-cycle hypersonic turbojet engine for flight demonstration [J]. Acta Astronautica, 2007, 61(1-6): 367-375.

[257] Taguchi H, Kobayashi H, Kojima T, et al. Performance evaluation of hypersonic pre-cooled turbojet engine: AIAA 2015-3593[R]. Reston: AIAA, 2015.

[258] Taguchi H, Hongoh M, Kojima T, et al. Mach 4 performance evaluation of hypersonic pre-cooled turbojet engine: AIAA 2018-5203[R]. Reston: AIAA, 2018.

[259] Li H, Zou Z P, Chen Y M, et al. Experimental insights into thermal

performance of a microtube precooler with drastic coolant properties variation and precooling impacts on turbojet engine operation[J]. Energy, 2023, 278: 127916.

[260] 纪佳正. 微小通道换热器设计及工作特性研究[D]. 长沙: 国防科学技术大学, 2016.

Ji J Z. Research on design and performance of mini-channel exchanger[D]. Changsha: National University of Defense Technology, 2016(in Chinese).

附录 A

表 A1　圆管（含微细管和常规尺度）或微通道内流体换热经验关联式

Table A1　Empirical correlations for fluid heat transfer in circular tubes (including mini- and mirco-tubes and conventional-scale tubes) or micro-channels

作者	方法	换热经验关联式	结构类型及尺寸	流体类型	适用范围
Dittus-Boelter[138-139]	实验	$Nu_b = 0.023Re_b^{0.8}Pr_b^n$ $n = \begin{cases} 0.3 & \text{冷却} \\ 0.4 & \text{加热} \end{cases}$	圆管	水	$L/d_i \geqslant 10$ $Re_b \geqslant 10\,000$ $0.7 \leqslant Pr_b \leqslant 160$
Taylor[140]	实验	$Nu_b = \begin{cases} 0.023Re_b^{0.8}Pr_b^{0.4} & \text{冷却} \\ 0.023Re_b^{0.8}Pr_b^{0.4}\left(\frac{T_w}{T_b}\right)^{-C} & \text{加热} \end{cases}$ $C = 0.57 - \frac{1.59d_i}{L}$	圆管	冷却: 空气、二氧化碳; 加热: 氢气、氦气、氮气、空气、二氧化碳	冷却: $0.25 < T_w/T_b < 1$ $4 < L/d_i < 52$ 加热: $1 < T_w/T_b < 23$ $2 < L/d_i < 52$
Gnielinski[141]	理论推导	$Nu_b = \frac{(f/8)(Re_b - 1\,000)Pr_b}{1 + 12.7\sqrt{f/8}(Pr_b^{2/3} - 1)}\left[1 + (d_i/l)^{2/3}\right]c_t$ $f = \left(1.82\lg Re_b - 1.64\right)^{-2}, c_t = (T_b/T_w)^{0.45}$	圆管		$2\,300 < Re_b < 10^6$ $0.7 < Pr_b < 10^5$
Nikitin等[142]	实验	$h_b = \begin{cases} 2.52Re_b^{0.681} & \text{加热} \\ 5.49Re_b^{0.625} & \text{冷却} \end{cases}$	半圆形截面、“Z”字形微通道 $d_{hot} = 1.9$ mm $\alpha_{Z,hot} = 115°$ $d_{cold} = 1.8$ mm $\alpha_{Z,cold} = 110°$	二氧化碳	$2\,800 < Re_b < 5\,800$ 加热 $6\,200 < Re_b < 12\,100$ 冷却
Kim等[143]	实验和仿真	$Nu_b = 4.065 + 0.003\,05Re_b$	半圆形截面、“Z”字形微通道 $d = 1.51$ mm $\alpha_Z = 150°$	氦	$350 < Re_b < 800$ $Pr_b = 0.66$ 1.5 MPa < P < 1.9 MPa
Jackson[136]	实验	$Nu_b = 0.018\,3Re_b^{0.82}Pr_b^{0.5}\left(\frac{\rho_w}{\rho_b}\right)^{0.3}\left(\frac{\bar{c}_p}{c_{p,b}}\right)^n$	水平圆管 $d_i = 19$ mm	水、二氧化碳	$Re_b \geqslant 8\,000$
Seo等[145]	实验	$Nu_b = 0.720\,3Re_b^{0.177\,5}Pr^{1/3}(\mu_b/\mu_w)^{0.14}$ 加热	半月形截面、直微通道宽0.8 mm, 高0.6 mm	水	$100 \leqslant Re_b \leqslant 850$
Liu等[148]	实验	$Nu_b = 0.122\,9Re_b^{0.602\,1}Pr_b^{0.3}\left(\frac{c_{p,w}}{c_{p,b}}\right)^{0.131}$	圆形截面、水平直微通道 $d = 1.87$ mm	二氧化碳	$3\,600 < Re_b < 36\,500$ $0.9 < Pr_b < 10.6$ 7.53 MPa < P < 11.97 MPa 299.55 K < T_b < 394.42 K 50.03 kg/(m^{2}s) < G < 780.11 kg/(m^{2}s)
Katz等[149]	实验和仿真	$Nu_b = \begin{cases} 0.026\,09Re_b^{0.876\,5}\overline{Pr}_b^{1/3} & CO_2 \\ 0.080\,72Re_b^{0.435\,9}\overline{Pr}_b^{1/3} & He \end{cases}$	近似半圆形截面、“Z”字形微通道 $d_h = 1$ mm $\alpha_Z = 106°$	二氧化碳、氦	$500 \leqslant Re_{b,CO_2} \leqslant 16\,000$ $P_{CO_2} = 20$ MPa 293 K < T_{CO_2} < 473 K $300 \leqslant Re_{b,He} \leqslant 3\,000$ $P_{He} = 4$ MPa 293 K < T_{He} < 823 K
北航团队[38,132]	实验	$Nu_b = 0.008\,9Re_b^{0.83}Pr_b^{0.021}\left(\frac{\rho_b}{\rho_w}\right)^{0.105}\left(\frac{c_{p,b}}{c_{p,w}}\right)^{0.003}$	水平圆管 $d_i = 0.8$ mm	甲烷	$5\,000 \leqslant Re_b \leqslant 30\,000$ 4.8 MPa < P < 7.5 MPa
北航团队[150]	实验	$Nu_b = 0.010\,3Re_b^{0.902\,2}Pr_b^{1/3}$	半椭圆形截面、“Z”字形微通道 $d_h = 291.6$ μm $\alpha_Z = 160°$	氦	$1\,350 \leqslant Re_b \leqslant 8\,200$ $P < 10$ MPa
Yu等[134]	实验	$Nu_b = 0.023Re_b^{0.8}Pr_b^{0.4}\exp[0.3(1-\overline{c_{p,b}}/\overline{c_{p,b,che}})]$	水平圆管 $d_i = 2$ mm	正癸烷, 甲基环已烷, 乙苯	$24\,000 \leqslant Re_b \leqslant 35\,000$ 50 kW/m^2 < q < 320 kW/m^2

注: 各关联式中, 下角标b指代流体体平均参数, 微通道问题中雷诺数 特征长度为水力直径。

表 A2 管外流体绕流叉排圆管对流换热经验关联式

Table A2 Empirical correlations for convective heat transfer of fluid cross flow over staggered circular tubes

文献	方法	经验关联式	结构类型及尺寸	流体类型	工况范围
Colburn[152]	实验	$Nu_{\mathrm{b}}=0.33Re_{\mathrm{b}}^{0.6}Pr_{\mathrm{b}}^{1/3}$	叉排圆管	空气	$10\leqslant Re_{\mathrm{b}}\leqslant 40\,000$ $N_{\mathrm{row}}\geqslant 10$
Žukauskas[153]	实验	$Nu_{\mathrm{b}}=\varepsilon_n c\left(\frac{s_T}{s_L}\right)^{k_1}Re_{\mathrm{b}}^{k_2}Pr_{\mathrm{b}}^{0.36}(Pr_{\mathrm{b}}/Pr_{\mathrm{w}})^{0.25}$	叉排圆管 $d_o=30$ mm	空气	$1\leqslant Re_{\mathrm{b}}\leqslant 2\times10^6$ $0.7\leqslant Pr_{\mathrm{b}}\leqslant 500$
Kays和London[154]	实验	$Nu_{\mathrm{b}}=C_hRe_{\mathrm{b}}^{0.6}Pr_{\mathrm{b}}^{1/3}$	叉排圆管	空气	$1.25<s_T<2.5$ $0.75<s_L<1.5$ $300<Re_{\mathrm{b}}<15\,000$
Chen和Wung[155]	仿真	$Nu_{\mathrm{b}}=0.78Re_{\mathrm{b}}^{0.45}Pr_{\mathrm{b}}^{0.38}$	叉排圆管	空气	$s_T=s_L=2d_o$ $40<Re_{\mathrm{b}}<800$ $0.1<Pr_{\mathrm{b}}<10$
Khan[156]	仿真及数值分析	$Nu_{\mathrm{b}}=\frac{0.61{s_T}^{0.091}{s_L}^{0.053}}{1-2\exp(-1.09s_L)}Re_{\mathrm{b}}^{1/2}Pr_{\mathrm{b}}^{1/3}$	叉排圆管 $1\ \mathrm{mm}\leqslant d_o\leqslant 3\ \mathrm{mm}$	空气	$1.05d_o\leqslant s_T\leqslant 3d_o$ $1.05d_o\leqslant s_L\leqslant 3d_o$ $10^3\leqslant Re_{\mathrm{b}}\leqslant 10^5$
Li等[158]	仿真	$Nu_{\mathrm{b}}=0.126Re_{\mathrm{b}}^{0.786}Pr_{\mathrm{b}}^{0.36}(Pr_{\mathrm{b}}/Pr_{\mathrm{w}})^{0.25}$	叉排圆管 $d_o=1$ mm	空气	$s_T=2d_o, s_L=2.5d_o$ $500\leqslant Re_{\mathrm{b}}\leqslant 1\,500$

Progress in compact and rapid heat transfer technology for precooled engines

WANG Yifan[1], LI Hui[2], CHEN Yiming[1, 3], FU Chao[1, 3], ZOU Zhengping[1, 3, *]

(1. Research Institute of Aero-Engine, Beihang University, Beijing 102206, China;
2. National Key Laboratory of Nearspace Physics, China Academy of Launch Vehicle Technology, Beijing 100076, China;
3. National Key Laboratory of Science and Technology on Aero-Engine and Aero-Thermodynamics, Beihang University, Beijing 100191, China)

Abstract: Compact and rapid heat transfer technology can, within ultra-confined spaces and on the order of milliseconds, efficiently capture and transport large amounts of heat from the intake or internal components. It is the core determinant of both performance and technical success for a precooled engine. Conventional civil use heat exchanger analysis, design procedures, and experimental testing approaches are insufficient for validation due to the harsh operating environment and demanding performance requirements inside such systems. Therefore, conducting highly targeted, application oriented research on compact and rapid heat transfer technology specifically for precooled engines carries significant academic and engineering value. In this paper, the research progress of compact and rapid heat transfer technology from multiple dimensions, including underlying mechanisms, technical approaches, and experimental methods, is systematically reviewed, and the principal challenges it faces when applied to precooled engines are analyzed. Existing studies have yielded extensive and in-depth results in several areas: the mechanisms of rapid and intensive heat transfer of supercritical working fluids in microscale channels and cross-flow tube bundles; rapid simulation and similarity analysis methods for fluids with sharply varying properties; refined and uncertainty quantification based design methodologies; and experimental testing techniques tailored for extreme thermal environments in confined spaces, etc. Overall, in terms of fundamental research, it is urgent to deepen our understanding of the coupled flow heat transfer mechanisms under complex thermal conditions at micro scales and to further enrich high accuracy empirical correlation databases for compact and rapid heat transfer. From an analysis and design perspective, attention should be paid to integrating emerging techniques into analysis and design methods to enhance overall performance across broad operating ranges. In terms of experimental validation, more sophisticated flow-field measurement techniques appropriate for intricate structures in confined spaces must be developed, and complete temperature and full pressure testing must be conducted under actual engine installation circumstances. These research efforts will provide a solid foundation for improving the maturity of compact and rapid heat transfer technology and for advancing and validating precooled engine technology.

Keywords: precooled engine; compact and rapid heat transfer; compact heat exchanger; precooler; microscopic scale; flow and heat transfer

Received: 2025-07-24; **Accepted**: 2025-09-12; **Published Online**: 2025-12-03 14:46

URL: link.cnki.net/urlid/11.2625.V.20251203.1139.002

Foundation items: National Natural Science Foundation of China (52406004); Supported by Advanced Aviation Propulsion Innovation Workstation (HKCX2024-01-001)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0322

全固态电池特性建模与仿真研究进展

赵立金[1,2]，杨世春[1,*]，刘国芳[2]，曲婧瑶[2]，孙旭东[2,3]

(1. 北京航空航天大学 交通科学与工程学院，北京 100191；2. 中国汽车工程学会，北京 100176；
3. 重庆大学 机械与运载工程学院，重庆 400044)

摘 要：全固态电池(ASSB)具备高比能和高安全的潜在特点，是全球各国政府及主流车企均在积极布局的下一代电池技术，但仍面临较大的研发和产业化挑战。通过开展全固态电池建模和仿真研究，可揭示其工作机制、掌握影响规律，进而设计满足指标需求、达到工作性能的全固态电池，缩短实验测试与开发周期。围绕全固态电池的分子动力学计算、关键材料及界面特性、电化学性能领域的建模和仿真工作进行了总结和分析，并重点对硫化物固态电解质材料微观结构建模、界面特性、离子输运机制、电解质间隙诱导应力和断裂等领域进行梳理，对全固态电池建模及仿真发展趋势进行展望，提出今后研究重点，期望可以促进全固态电池尽快实现产业化和应用。

关 键 词：全固态电池；建立模型；仿真分析；多物理场；固态电解质
中图分类号：TM912
文献标志码：A **文章编号**：1001-5965(2026)10-3276-22

高效安全的动力电池研发及产业化是保持新能源汽车产业竞争力的核心。经过多年发展，中国锂离子电池产业在全球形成较强竞争力并占据绝对优势。虽然液态动力电池的能量密度仍有发展空间，但运行安全性存在较大风险，其综合性能难以满足行业发展的高质量要求。以全固态电池(all-solid-state battery, ASSB)为代表的下一代电池，有望突破传统液态电池的瓶颈，具备兼顾高能量密度、长循环寿命、高安全性能、宽温度范围等综合特性的潜力，正在成为全球汽车强国铆力发展的下一代电池新赛道[1]。目前，全固态电池成为产学研焦点，多条技术路线齐头并进。全固态电池使用固体电极和固体电解质，不含液体电解质成分。

本文围绕全固态电池特性建模及仿真，总结和分析目前正极、负极、电解质材料微观模型建立的现状，并重点围绕锂枝晶生长、负极颗粒膨胀等重点领域探讨技术解决手段，总结基于多物理场耦合的全固态电池电芯的电化学性能和安全性能的建模和仿真。期望通过本文的梳理，能促进全固态电池的精确建模，并依托模型开展关键材料改性、电芯设计优化、系统集成创新，最终加速全固态电池产业化进程。

1 全固态电池基本原理及产业化进展

典型的全固态电池结构如图1所示，隔膜被替换为固态电解质，减少隔膜破损失效造成内短路引发热失控的风险。

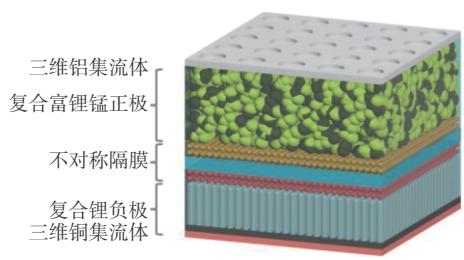


图1 全固态电池典型结构

Fig. 1 Typical structure of all-solid-state battery

收稿日期：2025-05-27；**录用日期**：2025-09-25；**网络出版时间**：2025-10-15 17:00
网络出版地址：link.cnki.net/urlid/11.2625.V.20251015.1343.001
基金项目：国家重点研发计划(2023YFB2408200)
* **通信作者**. E-mail：yangshichun@buaa.edu.cn

引用格式：赵立金，杨世春，刘国芳，等. 全固态电池特性建模与仿真研究进展[J]. 北京航空航天大学学报，2026，52(10)：3276-3297.
Zhao L J，Yang S C，Liu G F，et al. Research progress on modeling and simulation of characteristics of all-solid-state batteries[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3276-3297 (in Chinese).

日本丰田汽车、韩国三星 SDI 等硫化物基全固态电池代表企业，宣布开发出大容量软包电池样品，计划于 2026—2030 年完成量产及装车应用。美国 Quantum Scape 公司采用氧化物、硫化物 2 种电解质材料路线，开发的 2.2 Ah 实验室原型全固态电池能量密度推算可超过 400 Wh/kg，1 C 充放电循环 580 次，容量保持率达到 90%。中国多技术路线布局全固态电池研发和产业化，以硫化物电解质为主要产品路线的宁德时代、比亚迪等计划于 2026—2027 年完成小批量生产及装车应用；以氧化物和聚合物复合电解质为主要产品路线的北京卫蓝新能源、上汽清陶能源等计划于 2026—2027 年完成小批量生产及装车应用。目前，多家动力电池企业已开发出 10 Ah 级的高比能量全固态锂离子电池，能量密度可达 400 Wh/kg。整车企业固态电池产品布局如表 1 所示，电池企业固态电池产品布局如表 2 所示。

表 1　整车企业固态电池产品布局

Table 1　Layout of solid state battery products for vehicle enterprises

车企	搭载车型	电池供应商	装车进程
丰田汽车		自研	2025年8月，宣布突破硫化物固态电池耐久性瓶颈，计划2027—2028年量产装车
日产汽车		自研	2024年4月，宣布横滨工厂建成全固态电池试产线，计划2028年量产
宝马集团	i7	Solid Power	2025年5月，宣布与Solid Power联合开发20 Ah全固态电芯，启动原型车i7路测
大众集团		Quantum Scape	通过Quantum Scape投资推进固态电池商业化，2024年1月通过PowerCo耐久性指标验证；2025年9月，完成全球首次QS固态锂金属电池驱动电动车的实车演示
福特汽车		SK On	与SK On合作开发聚合物基固态电池，计划2026年装车测试
长安汽车		自研	预计2026年实现固态电池装车验证，2027年推进全固态电池逐步量产，能量密度达400 Wh/kg
东风汽车	奕派007	自研	2025年9月，表示全自主研发350 Wh/kg、1 C固态电池，2026年装车；第二代半固态电池量产装车（奕派007车型），支持4 C快充
上汽集团	MG4	上汽清陶能源	2025年在全新MG4上量产新一代固态电池，并计划于2027年推出首款全固态“光启电池”
广汽集团	昊铂	自研	2024年4月，宣布第三代海绵硅负极技术和高面容量固态正极技术，实现全固态电池能量密度达到400 Wh/kg以上，目标在2026年完成全固态电池开发，并首先搭载于昊铂车型
奇瑞汽车		自研	2024年宣布，自主研发的全固态电池，1 500 km续航，6 C充电效率，且将会在2026年上车，2027年开始量产
蔚来汽车	ES6/ET7	北京卫蓝新能源	搭载150 kWh半固态电池的ES6、ET7于2023年量产；2024年6月起，150 kWh电池包适配蔚来旗下所有车型

表 2　电池企业固态电池产品布局

Table 2　Layout of solid state battery products for battery enterprises

企业	电池类型	能量密度/(Wh·kg^{-1})	产业化进程
宁德时代	全固态电池	500	2025年5月，表示预计2027年有望实现小批量生产
比亚迪	硫化物全固态电池	400	2024年宣布，首款全固态电池将搭载于海豹EV车型；全固态电池中试完成，重庆璧山基地规划100 GWh产能
中创新航	全固态电池	430	自研的430 Wh/kg全固态电池已完成产线建设，并规划了从半固态到全固态的稳健推进路线，配套产线于2025年正式投用
国轩高科	全固态电池	350	2024年5月，推出金石全固态电池，比能量350 Wh/kg；2025年5月，推出G垣准固态电池，比能量300 Wh/kg；宣布首条全固态实验线贯通，该生产线100%自主开发，设计产能达0.2 GWh，良品率高达90%
蜂巢能源	半固态电池	300	2025年11月，半固态电池产线正式投产，第一代高镍半固态电池（能量密度270 Wh/kg）已实现小批量生产，首搭于欧洲豪华品牌乘用车；第二代高镍半固态电池（能量密度342 Wh/kg）开发完成，主要配套低空飞行器，已完成装机试飞测试
孚能科技	硫化物全固态电池	400～500	硫化物固态电池突破车规级验证，能量密度400 Wh/kg，续航超1 000 km，10 min快充80%，安全性能提升10倍，2026年小批量交付广汽埃安等车企
赣锋锂业	固态电池	400～500	2025年6月，表示已形成电池全链路布局，首款500 Wh/kg级10 Ah产品实现小批量量产，400 Wh/kg电池循环寿命突破800次并完成工程验证
鹏辉能源	固态电池	320	2025年8月，表示固态电池能量密度已由280 Wh/kg提升至320 Wh/kg
上汽清陶能源	半固态电池+全固态电池	368	国内首条1 GWh半固态电池产线已投产，能量密度368 Wh/kg，搭载上汽智己L6车型实现1 000 km续航，全固态电池计划2026年量产，能量密度目标500 Wh/kg
北京卫蓝新能源	半固态电池+全固态电池	360	150 kWh半固态电池包已交付蔚来ET7，充电倍率提升至3 C，2025年底，半固态电池已有产能规模达5.6 GWh；全固态电池能量密度达400 Wh/kg，计划2027年小批量装车
LG新能源	固态电池		计划开发包括固态电池在内的下一代电池，并将在2030年前推出成熟产品
三星SDI	全固态电池		新一代电池能量密度预计可达约900 Wh/L，有望在2027年前后应用于下一代Galaxy系列手机及电动交通工具中
SK On	半固态电池		一条半固态电池中试线正在运行，半固态电池样品将在2026年底实现生产
Quantum Scape	全固态电池	400	与大众合作推进固态电池量产，目标能量密度400 Wh/kg；2025年6月，宣布独特的Cobra隔膜工艺已顺利融入基线电池生产流程

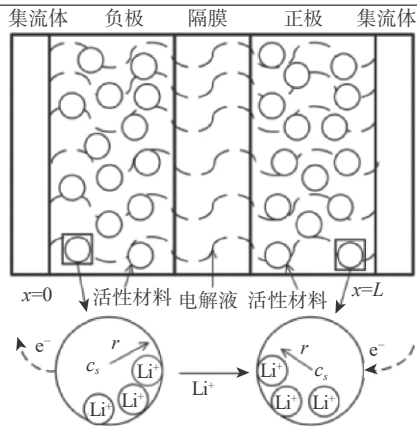
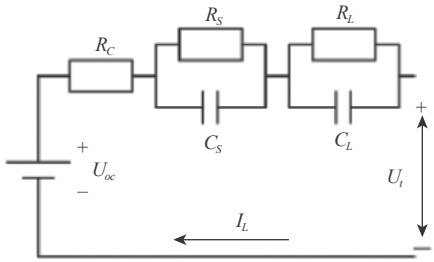
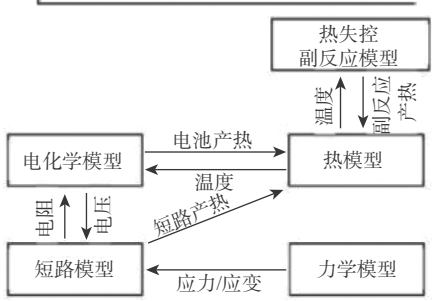
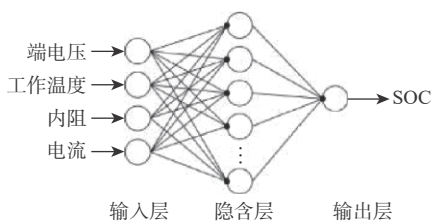
2 动力电池模型对比分析

目前，常用的锂电池模型主要有电化学模型、等效电路模型、神经网络模型和热耦合模型，对其进行分析具有重要的工程应用价值，如表3所示。

等效电路模型适合实时控制，电化学模型和热耦合模型用于深度机理分析，神经网络模型在数据驱动场景中表现优异。实际应用中，常采用混合模型（如等效电路+神经网络模型）来平衡精度与效率。

表3 锂电池常用模型总结

Table 3 Summary of commonly used models for lithium batteries

模型	基本介绍	示意图	优点	缺点	适用范围
电化学模型	基于电化学反应机理构建，如Newman伪二维（pseudo-two-dimensions, P2D）模型，通过偏微分方程描述锂离子在固/液相中的扩散、电荷转移和电解质迁移等过程，典型应用包括电极材料优化和充放电特性预测		①物理意义明确，能精确描述电池内部反应机理； ②适用于微观尺度（如活性颗粒）的优化设计	①计算复杂度高，需大量参数（如扩散系数、反应速率常数）； ②难以扩展到模组或系统级别	电极材料开发、高精度单体电芯性能仿真、基础电化学研究
等效电路模型	通过电路元件（电阻、电容、电感等）模拟电池外特性，如Rint、Thevenin、PNGV模型等，分数阶等效电路模型进一步引入非线性元件提升动态特性拟合能力		①结构简单，参数辨识容易（如通过脉冲测试提取欧姆内阻、极化电阻）； ②计算效率高，适用于实时状态估计，如SOC、SOH	①无法反映电池内部机理（如锂枝晶生长）； ②精度受限于模型阶数，高阶模型参数辨识困难	电池管理系统在线控制、电动汽车能量管理策略、快速仿真场景（如硬件在环测试）
热-电化学-力耦合模型	多物理场耦合模型，结合电化学、热力学和力学效应，模拟电池充放电过程中的电流分布、温度场变化及应力演变（如扩散诱导应力、热应力）		①全面分析电池性能与安全性（如热失控、结构形变）； ②支持固态电池等新型设计的优化	①计算资源消耗巨大，需高性能计算支持； ②模型参数获取复杂（如材料弹性模量、热膨胀系数）	电池热管理设计、固态电池应力分析、极端工况下的安全性评估
神经网络模型	数据驱动模型，通过大量实验数据训练网络（如长短期记忆网络、卷积神经网络）预测电池特性（如容量衰减、电压响应），无需依赖机理公式，适合处理非线性问题		①适应复杂工况（如动态负载、老化演变）； ②可融合多源数据（电压、温度、电流）提升预测精度	①依赖高质量数据集，数据获取成本高； ②模型泛化能力受限，外推性能较差	电池寿命预测、复杂动态工况下的状态估计、数据丰富的工业场景（如电池回收评估）

3 基于分子动力学计算的微观材料特性仿真研究

分子动力学方法是通过对体系中各个原子的运动方程进行时间积分来追踪原子运动轨迹的一种模拟方法[2-3]，分子动力学在传统锂电池建模仿真方面已取得了较好的进展，在全固态电池建模仿真领域正逐渐应用。Foroozan等[4]通过分子动力学方法研究了一种局限于玻璃纤维隔板的编织结构中的三维保形氧化石墨烯纳米片涂层，发现锂离子最初会吸附在亲锂的涂层上，通过缺陷部位扩散，这种延迟的锂转移消除了尖端效应。Deng等[5]使用密度泛函理论（density functional theory, DFT）和投影增广波方法进行离子相互作用，以获得不同化合

物弹性张量的完整描述。

微观尺度模型在锂电池设计方面已取得了较好的应用进展。Franco 等[6] 做了一系列工作，建立了模拟锂离子电池制造过程并预测其电化学性能的多尺度仿真平台，该仿真平台涵盖了电极介观结构生成、网格划分、有限元或有限体积及格子玻尔兹曼方法（lattice Boltzmann method, LBM）和离散元方法（discrete element method, DEM）的数值模拟仿真应用，采用粗粒化分子动力学（coarse grained molecular dynamics, CGMD）模型用于浆料制备和电极干燥过程介观结构的生成，从电极制备工艺追溯模拟电极介观的生成，考虑了活性材料（active material, AM）和碳-黏结剂相（carbon-binder domain, CBD），可以评估 CBD 的空间位置及其对锂离子传输效果对整体电化学性能的影响。Hao 等[7] 开发了动力学蒙特卡罗（kinetic Monte Carlo, KMC）模型来求解电化学公式，并将其耦合到有限元解中，其中，KMC 中的每个单元对应于一个有限元分析（finite element analysis, FEA）模型，以解释化学诱导的活性颗粒的体积变化。Yu 等[8] 利用光滑边界法（smooth boundary method, SBM）求解力学平衡，计算热膨胀和化学诱导变形引起的应力，展示了对热收缩和剥落产生的等效应力的代表性结果。Tian 等[9] 采用多尺度框架耦合 DFT 和相场模型，研究了具有微观结构和内部缺陷的固态电解质内部锂枝晶的萌生和生长，采用 DFT 计算方法对 4 种具有代表性的固态电解质进行表面电子态密度和内部缺陷中多余电子的分布研究，结果表明，表面电子态的存在作为多余电子的捕获机制，使表面缺陷处的锂金属沉积成为核[10]。从 DFT 结果中获得的表面多余电子浓度随后被纳入相场模型来研究枝晶演化的形态。

Ahmad 和 Viswanathan[11] 意识到 DFT 计算中与计算平衡态周围能量相关的不确定性，并通过量化这种不确定性来改进先前的工作，对 $Li_{10}GeP_2S_{12}$、Na_3PS_4、LiI 和 Na_3OCl 这 4 种化合物的能量集合进行计算，获得了具有量化标准差的力学性能的概率分布，研究表明，电池运行过程中的机械现象在很大程度上取决于固态电解质材料的选择。虽然 DFT 模拟可以帮助克服复杂实验表征的需要，但考虑到潜在固态电解质化合物的大量组合，仍然相对耗时。加速固态电解质机械设计的一个可能途径是利用数据科学方法，机器学习模型提供了一种方法，将大分子中的原子与其相邻的对应物并置，给出其键合和相互作用的表示。同样，机器学习可以很容易地对大分子的能量或低精度和高精度计算之间的差异进行预测。机器学习特别适用于寻找稳定无害的锂电池电解质，还提供了在计算和大数据分析中降低数量级的解决方案。Ahmad 等[12] 通过使用机器学习技术扫描超过 12 000 种无机固态电解质化合物的空间，确定具有最佳机械性能的化合物来抑制枝晶生长，从而扩展这项工作，但数据只针对化合物的结构特征进行训练，认为固态电解质的弹性特性是抑制枝晶的唯一主导因素。

分子动力学计算被广泛应用于全固态电池材料的微观特性仿真中，能够从原子尺度揭示离子扩散机制、界面稳定性及电极-电解质相互作用等关键物理化学过程，为材料性能优化提供理论支持。

4　全固态电池关键材料及界面特性建模与仿真

在全固态锂电池中，电极-电解质界面是区别于传统液体电解质电池的关键，存在着亟待解决的问题：固态锂电池在充放电时，在空间电荷区附近可能会形成锂耗尽层，这将显著增大其界面迁移电阻[13-14]；固-固界面在空间电荷分布的均匀性上也远不如固-液界面电荷分布的非均匀性，降低了全固态电池的循环稳定性[15-16]；固态电解质可能会发生氧化或还原反应，由此进一步降低全固态锂电池的循环稳定性[17]。这些问题会导致全固态电池的循环性能、寿命等受限[18]。传统显微成像、材料分析表征等方法能够辅助揭示全固态电池界面失效演变机制，但建模方法仍是揭示全固态电池界面演变机制的关键方法[19]。

4.1　全固态电池正极材料建模与仿真

高比容正极材料是全固态电池的重要组成部分，其中，机械效应将起到关键作用[20]。在传统正极材料之外，目前复合电极的概念也越来越热。复合电极是指一种包含多种成分的电极结构，不仅包含电极活性材料，还包含固体电解质和导电剂。针对电极的建模，主要方法包括多孔电极理论和离散化处理。

多孔电极理论具有明显优势，可以产生计算效率高的解决方案，同时仍然保持对微观结构多孔性较为准确的描述。基于多孔电极理论[21] 的电极材料建模，可以实现对电极复合特性的精准描述。Renganathan 等[22] 开发了一个全电池模型，将多孔电极理论用于 2 个电极，加强了传统的多孔电极理论，以解释应力对扩散的影响。Christensen[23] 进行了类似分析，通过假设电极中存在一定程度的预先存在的裂缝，包括断裂的作用。Mykhaylov 等[24] 考虑了使用多孔电极理论的全固态结构，不考虑机械

变形和应力，并为系统的性能实现了封闭形式或半封闭形式的解决方案。Ferguson 等[21] 基于详细的 Cahn-Hilliard 模型发展多孔电极理论，以解释活性颗粒中的相分离，通过考虑局部浓度与其体积平均量的变化，在化学势中引入应变的作用，通过分析，能够证明力学在抑制相分离和促进均匀填充方面的重要性。

多孔电极较为复杂，相关参数有几十个，为其建立完整的可映射的传输线模型（transmission line model, TLM）存在挑战。Zelič等[25] 从浓溶液理论出发，经推导与映射，演绎得到嵌入电极的完整可视 TLM，其响应与解析解保持一致，如图 2 所示。

在多孔电极理论基础上，需要开发有效的均质模型以分析固态复合材料正极的详细微观结构行为。通过计算效率高的模型，指导全固态电池的设计和开发，并确定下一代材料的前景分析。数据科学方法（机器学习等）很有可能在这些新的均质化工具的开发中发挥作用。

将微观结构或其代表性体积完全离散化，也能获得对电极的精准描述。Garcia 等[26] 开发了一种二维有限元描述详细的复合正极微观结构，其中，圆柱形颗粒耦合到隔膜和负极行为的简化模型，通过使用具有化学诱导应变的各向同性弹性活性颗粒来捕获机械变形。Xu 等[27] 进行了三维有限元模拟，以数值解析 $LiNi_xMn_yCo_zO_2$（NMC）复合正极的微观结构，配合加速器 X 射线断层扫描，用于微观结构建模并验证模型预测，如图 3 所示，在该方法中，研究人员仍需要了解微观结构的关键实验描述，如粒度和形状分布。Bielefeld 等[28] 使用复合正极微观结构的三维描述来研究其渗透特性。

机械性能及结构损伤对固态正极性能的影响至关重要，需要开发具有可能的活性粒子失效、固态电解质失效和两者分层的综合效应的模拟方法。对于活性颗粒，已经针对许多不同的活性材料开发了耦合化学-力学模型，并开发了用于嵌入电极的模型[29] 及用于转换电极的模型[30]，包括失效模式和损伤的各个方面[31]。Bucci 等[32] 关注约束诱导应力如何使用解析方法和数值有限元模拟来改变电池性能，研究结果表明，随着电解质矩阵刚度的变化，硅与锂的开路电位会改变，增加基质刚度会导致更大的压应力，从而产生改变开路电位和降低存储容量的效果。Al-Siraj 等[33] 研究了正极成分中活性物质的临界体积分数的存在，结果表明，固态电解质的有效扩散系数随活性粒子体积分数的增加呈线性衰减，在临界体积分数以上，固态电解质的有效扩散系数可以忽略不计，从而限制了性能。Bucci 等[34] 开展了第一个关于力学损伤在复合材料全固态正极中作用的定量研究，内聚区建模用于捕获复合正极循环期间固态电解质中的损伤。伴随着耦合化学力学理论框架的有限元实现，研究人员已在二维和三维模拟中证明了锂离子形态受到机械变形的严重影响的方式[35]。近年来的研究表明，柔顺型固体结构有利于防止分层，而刚性固体结构有利于延缓微裂。近期，新建立的模型已被证明可以准确地预测单个颗粒的电化学性能，以及机械应变影响相偏析形貌的方式[36]，甚至完全抑制[37]。

机械压力对锂离子电池正极材料的电化学性能影响主要体现在扩散速率和界面动力学 2 方面，具体影响机制如下：①压力对扩散速率的影响。Mussa 等[38] 发现机械压力会改变正极材料孔隙率和迂曲度，进而影响锂离子（Li^+）在多孔电极中的扩散速率，实验表明，随着压力增加，锂离子在正极的扩散阻抗显著上升，尤其在高压（如 1.98 MPa）下，扩散速率下降幅度可达 40% 以上，但适度压力（如 1.32 MPa）可优化扩散条件，减少活性物质脱落，延缓容量衰降。②压力对界面动力学的影响。正极与隔膜间的界面阻抗也会因压力变化而波动。实验数据显示，当压力从 0.66 MPa 增至 1.98 MPa 时，

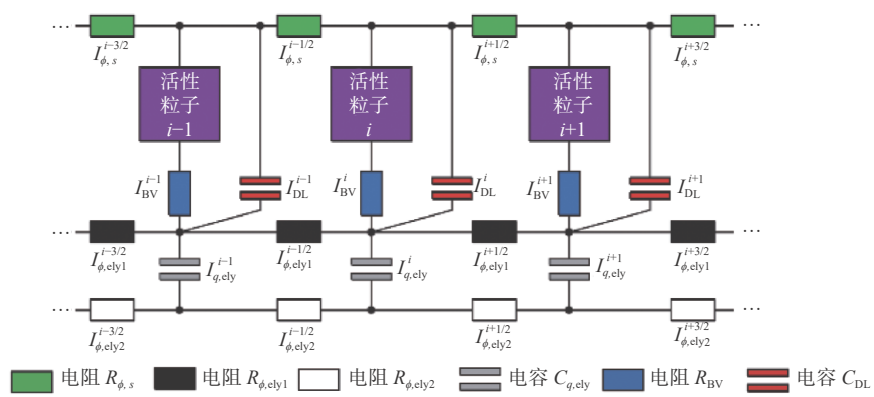


图 2 电极 TLM 模型[25]

Fig. 2 Electrode transmission line model[25]

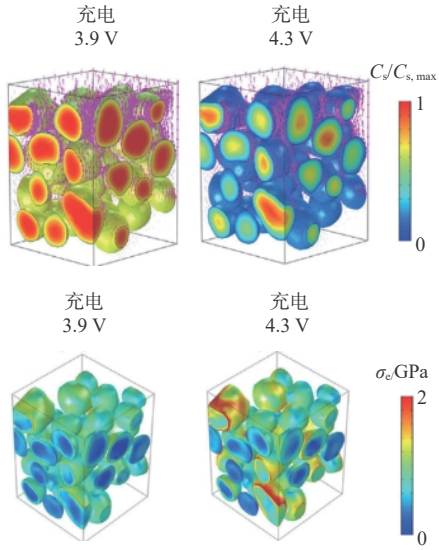


图 3　充电 3.9 V 和 4.3 V 复合 NMC 正极的三维数值模拟结果[27]

Fig. 3　Three-dimensional numerical simulation results of composite NMC positive electrode of charging 3.9 V and 4.3 V[27]

正极界面阻抗先降后升，最佳压力为 1.32 MPa。该现象与颗粒间接触面积变化及电解液渗透能力有关。

正极建模需重点考虑：①颗粒半径。压力改变颗粒形变程度，直接影响锂离子扩散路径长度。②孔隙率。压力导致孔隙结构塌陷，降低锂离子传输效率。③接触电阻。颗粒间接触紧密度影响电子传导效率[39]。

4.2　全固态电池负极材料建模与仿真

4.2.1　复合锂金属负极

复合锂金属负极材料具有最低的还原电位（−3.04 V）和超高理论比容量（3 860 mAh/g），远超传统的碳负极材料，但锂在负极和固态电解质界面的不稳定沉积可能会造成锂枝晶的形成和生长，降低电芯比能量的同时也带来安全风险。因此，了解锂的沉积/溶解机理、非均匀界面动力学下的不均匀沉积机制、锂枝晶的生长和蔓延机制具有重要意义。

锂离子在沉积过程中形态各异。基于主要结构、沉积机制及对锂电池性能的不同影响，锂枝晶的形态主要分为 3 种模式：针状、苔藓状和树状[40]，如图 4 所示。针状枝晶会在横向和纵向方向上同时生长，当枝晶达到一定长度时就会戳破隔膜，造成内部短路，严重时可能导致锂电池发生爆炸。因此，针状锂枝晶是导致锂离子电池性能下降的主要原因之一。相比针状枝晶，苔藓状枝晶有更高的表面积，其枝晶直径较小，需要消耗更多的锂以形成固体电解质界面膜（solid electrolyte interface, SEI），并且苔藓状枝晶在被氧化过程中更容易断裂形成死锂。树状枝晶是最常见的锂枝晶生长模式，与针

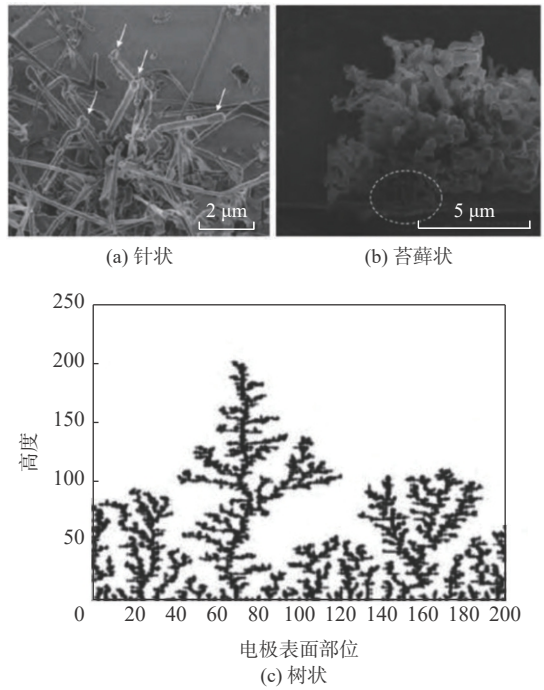


图 4　不同类型的锂枝晶生长形貌观测[40]

Fig. 4　Observation of growth morphology of different types of lithium dendrites[40]

状枝晶和苔藓状枝晶不同，其可以向各个方向生长，是锂枝晶模拟的重要研究方向之一，这种树状枝晶也会带来各种危害，最常见的有短路和锂容量损失等问题。

目前，枝晶生长问题仍无法解决，但相关研究已经发现了能够抑制枝晶生长的策略，包括双极脉冲电流电荷法、界面优化、电解质改进等，对仿真理论方法与枝晶生长和抑制机理提供了更多的参考[41]。最早的枝晶生长模型由 Barton 和 Bockris[42] 提出，他们研究了硝酸银电解质中球形银沉积的电解生长，用公式描述了纯扩散控制下枝晶生长模拟过程，而不考虑离子物种在运输中的迁移影响。随着 Chazalviel[43] 和 Brissot 等[44] 的工作，出现了枝晶生长的迁移限制传质公式。这些公式研究了高电压（10 V）下的金属电沉积，当电压的值达到限制电流密度的上限时，在该领域中电子不属于中性。枝晶起始的机制被规定为电极附近阴离子耗尽导致空间电荷积聚的直接结果。

研究人员围绕锂枝晶生长的描述做了大量工作。考虑力学与界面动力学的耦合作用，Monroe 和 Newman[45-46] 提出了一个扩展的电沉积框架，该框架结合了外部压力、弹性变形和黏性应力的影响，并引入了描述界面的热力学相。从动量平衡开始，考虑到电极-电解质界面上的平衡条件，建立了由表面张力、压力、偏差和黏性等局部变形引起的电化学电位差的关系，并首先提出了一个电化学枝晶

生长模型，将枝晶尖端曲率的影响纳入其枝晶生长反应动力学中。Akolkar[47]扩展了该模型，加入了浓度相关的扩散系数，并应用于液体电解质。Ganser等[48]开发了更通用的Butler-Volmer动力学耦合力学框架，影响界面反应的应力机械成分不必是静应力，而是在界面生长方向上的分解应力的测量，为模拟电沉积动力学中力学作用搭建了理论框架，为进一步模拟固态固体中枝晶生长创造了条件。Klinsmann等[49]用更完整的连续力学框架模拟了锂金属负极附近的锂化裂纹的进展，探讨了锂镀层在裂纹内形成的压力可以促进不稳定裂纹的扩展条件，也可以通过阻止锂的插入来阻碍不稳定裂纹的扩展，断裂扩展所需的临界压力随后被转化为等效的表面过电位状态，在该状态下，电池的安全运行得到了保证。Yu等[50]研究了微观结构特征在促进锂不均匀沉积和随后金属穿透固态电解质中的作用，假设晶界附近固态电解质的纳米级软化为锂的优先镀提供了一种机制，从而导致枝晶生长，考虑了固态电解质微观结构的异质性，在2个低能倾斜和扭转晶界上对具有代表性的固态电解质进行了分子动力学模拟，以测量电解质主体和晶界附近的弹性特性，与体相区域相比，在晶界处局部观察到弹性模量降低了50%。洪清和等[51]利用多物理面向对象仿真环境平台(multiphysics object oriented simulation environment, MOOSE)，使用相场模型建立了仿真对象与研究界面，发现材料参数各向异性数值大时枝晶生长慢，但会有更多晶核生长为枝晶。

考虑力学与电沉积动力学的耦合作用，Barai等[52]考虑了枝晶起始形态域，将锂金属的初始结构修改为无应力，去除界面处的外部牵引力，可实现从松弛状态对早期枝晶生长行为进行建模，研究表明，无论电解质的剪切模量如何，锂的初始松弛状态都可以在低电流密度下阻止枝晶生长，并提出电解液中的电流分布是枝晶生长的主导因素。Natsiavas等[53]制定了弹性预应力存在下电极/固态电解质界面生长的详细框架，采用线性稳定性分析方法，研究了Li-LiPON系统的平面生长，以及稳定沉积的条件，作为应力状态、弹性模量、沉积动力学和界面能的函数，证实了预应力的存在对电极-电解质界面具有稳定作用。Ahmad和Viswanathan[54]假设了各向同性固-固界面电沉积的一般稳定性准则，扩展了先前公式，解释锂金属密度变化对界面电沉积稳定性的影响，并引入了一个额外的参数-体积比，以说明电极中锂的密度变化对固态电解质的影响。Ding等[55]应用碱金属离子添加剂提出了锂枝晶的屏蔽机制，并对锂枝晶生长形貌进行了仿真，利用碱金属离子的有效还原电位低于低浓度锂离子的机制，使其在锂离子沉积过程中能够稳定生长，迫使后来的锂离子进一步沉积到负极凸起的邻近区域，该策略可有效抑制锂离子连续生长。

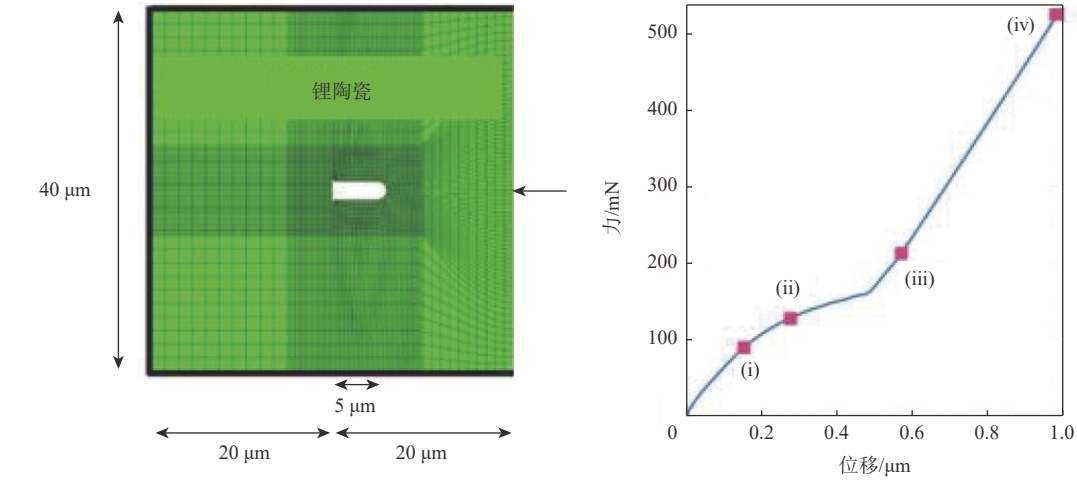
上述建模工作的共同关键特征是依赖于弹性接口。锂金属负极表现为塑性，具有较低的同源温度，其塑性响应已成为近期许多实验研究的焦点。将塑性变形对电沉积动力学和枝晶生长的影响纳入理论框架的发展，刺激了计算研究工作的新途径。Schultz[56]利用锂金属的实验应力-应变数据，采用各向同性硬化J_2塑性模型，但忽略了黏塑性效应。

虽然固态电解质的电化学和机械响应被认为是导致锂枝晶形成的主要因素，但锂电极的性质对于捕捉整体情况也很重要。Narayan和Anand[57]提出了锂的大变形弹-黏塑性模型，并根据纳米压痕数据计算出了模型中的材料参数，仿真结果表明，当锂和固态电解质相互机械推动时，锂枝晶可以被压平，锂可以渗透到固态电解质的腔中，如图5所示。为分析锂的流动在锂枝晶生长中的作用，Yang[58]引入了一个非牛顿流动模型来描述锂流动，并研究了锂球和锂晶须在电解质中的生长，结果表明，通过抑制循环诱导的应变能来缓解锂枝晶生长。

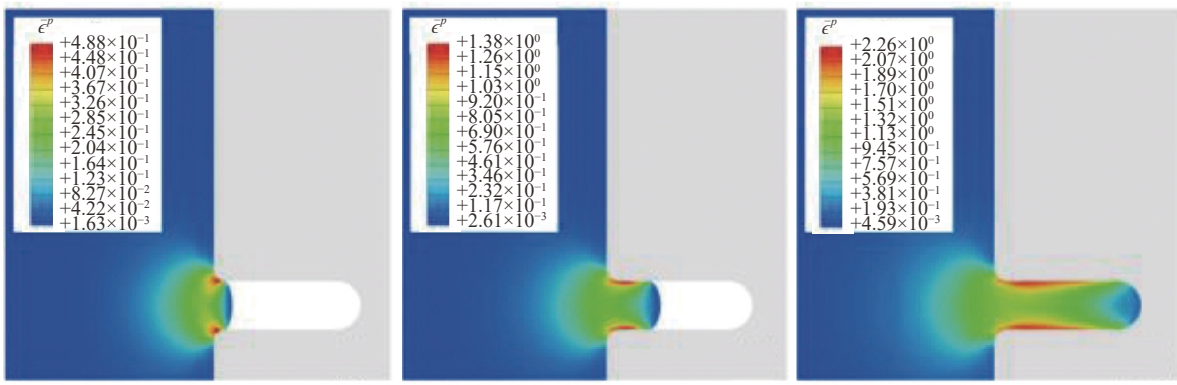
4.2.2 硅基负极材料

转换电极通常会发生非常大的弹性-非弹性变形，如已被广泛研究的硅基负极，其化学应变的数量级为100%[59]。目前，已经开发了连续统模型，该模型解释了伴随着物质扩散和化学应变的大弹性-非弹性变形[60]。对于一些简单的情况，已经开发了封闭形式的解析解。Bower等[61]开发了具有尖锐传播相边界的弹塑性球形颗粒经历扩散诱导变形问题的解析解。许多研究使用简化几何(即球形或径向对称)来简化为可以用数值方法解决的一维问题。Huang等[62]证明了在有无尖锐界面存在的情况下，应力在经历化学诱导的大弹塑性变形的球形颗粒中发展的方式。这些研究发现，由于非弹性变形和两相锂化的存在，表面上的环向应力可能是拉伸或压缩的，这对材料的断裂建模具有重要意义。

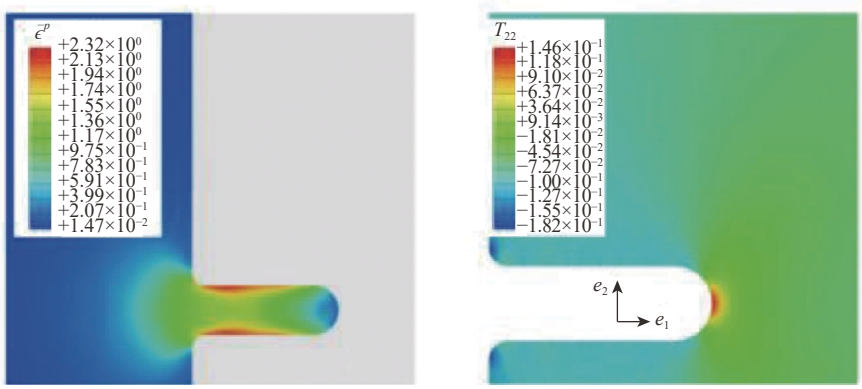
非晶硅因其作为高容量负极的潜力而得到了广泛研究。Anand[60]和Di Leo等[63]都校准了连续介质力学理论，并证明了预测简单非晶硅电极的电压与电荷行为及其力学行为的能力。Di Leo等[63]使用大变形弹塑性化学力学理论和有限元模拟来预测Wu等[64]实验实现的空心双壁纳米管的行为。Zhang等[65]使用基于变分的相场损伤模型来



(a) 仿真的几何模型　　(b) 加载曲线



(c) 初始裂纹尖端开始出现损伤集中 (i)　　(d) 损伤区域开始向前延伸 (ii)　　(e) 裂纹持续向前扩展 (iii)



(f) 裂纹完成扩展 (iv)　　(g) 竖直方向正应力分布 (iv)

图 5　不同加载速率下载荷与压痕深度的仿真对比[57]

Fig. 5　Simulation comparison of load and indentation depth at different loading rates [57]

模拟非晶和晶硅电极的扩散诱导塑性和断裂，采用有效的时间积分方案，成功模拟了硅纳米颗粒和纳米柱的断裂，并研究了能量释放率对电极损伤行为的影响。

4.3　全固态电池固态电解质材料建模与仿真

4.3.1　固态电解质材料微观结构建模

锂硫银锗矿型固态电解质是一种非常有前途的硫化物固态电解质，具有高离子电导率和良好的化学稳定性，其化学通式为 Li_6PS_5X(X=Cl,Br,I)[66]。混合卤化物银柱石 Li_6PS_5X(X=$Cl_{0.75}Br_{0.25}$,$Cl_{0.5}Br_{0.5}$,$Cl_{0.25}Br_{0.75}$)与单卤化物相相比，具有较高的离子电导率，不同结构形式的离子电导率和电子电导率如表 4 所示[67]。

De Klerk 等[68] 利用 DFT 分子动力学模拟，通过调整硫/卤化物比进一步提高了锂电导率，展示了由溶液工程合成方法制备的 $Li_{5.75}PS_{4.75}Cl_{1.25}$ 和 $Li_{5.5}PS_{4.5}Cl_{1.5}$ 的 X 射线衍射法(X-ray diffraction, XRD)尺度结果，主要成品为锂硫银锗矿型固态电解质，掺杂剂为 Li_3PO_4、Li_6PS_5X、少量的 Li_2S 和筛选过的 LiCl, EIS 结果显示，Cl-S 比增大，$Li_{5.75}PS_{4.75}Cl_{1.25}$ 和

表 4 不同结构锂硫银锗矿型固态电解质的离子电导率和电子电导率[67]

Table 4 Ionic and electronic conductivity of lithium sulfur silver germanite solid electrolytes with different structures[67]

化学式	离子电导率/(mS·cm^{-1})	电子电导率/(mS·cm^{-1})
$Li_{5.5}PS_{4.5}Cl_{1.5}$	3.9	1.4×10^{-5}
$Li_{5.75}PS_{4.75}Cl_{1.25}$	3.0	2.6×10^{-5}
Li_6PS_5Br	1.9	4.4×10^{-6}
$Li_6PS_5Cl_{0.25}Br_{0.75}$	3.4	1.1×10^{-5}
$Li_6PS_5Cl_{0.5}Br_{0.5}$	3.9	1.4×10^{-5}
$Li_6PS_5Cl_{0.75}Br_{0.25}$	3.2	3.7×10^{-6}
Li_6PS_5Cl	2.4	5.1×10^{-6}

$Li_{5.5}PS_{4.5}Cl_{1.5}$ 的离子电导率激增，离子电导率的增加和 XRD 标度表明，Cl 取代了 S，为锂硫银锗矿结构引入了锂离子空隙，从而提高了离子电导率。

Baktash 等[69] 基于人工智能技术，利用经验模态分解（empirical mode decomposition, EMD）和非平衡分子动力学模拟（non-equilibrium molecular dynamics, NEMD），在不同温度下对 $Li_5PS_4Cl_2$ 的扩散系数进行模拟，利用 EMD 和 NEMD 技术的精度，以及 Arrhenius 方程在 300 K 下的外推计算，对 AI-EMD 在 300, 600, 800, 1 000 K 下的 45, 40, 20, 20 ps 的模拟进行了评估，利用 EMD 模拟的结果计算了锂离子的均方位移（mean square displacement, MSD）、10 个模型的值及不同温度下的统计误差，并确保模拟时间（450～200 ps）足以区分锂离子的扩散。

假设 Arrhenius 行为，Baktash 等[70] 将无序构型电导率的经验值以正方形和三角形形式呈现，并报告了计算结果，通过对误差条的研究，计算出的拆装结构电导率的结果与 Yu 等[71] 报道的当代经验结果一致。因此，无序结构的电导率比纯 Li_6PS_5Cl 结构高 2 个数量级。相应地，轻微的结构变化（包括替换 104 个原子中的 4 个）给电导率带来了相当大的改变。使用 EMD 模拟可以调节 600 K 下有序和无序锂离子结构的扩散机制（见图 6）。比较 2 种方法，切换 Cl 和 S 离子的位置会导致无序，从而改变锂离子自由。

4.3.2 固态电解质界面特性

全固态电池存在固-固界面接触，界面两侧材料的低润湿性导致界面处产生接触面积损失，界面电阻显著增大，这对电池性能产生不利影响。因此，考虑界面接触面积损失对电池性能的准确模拟至关重要[72]。全固态电池中的复合电极主要由固态电解质包围的活性粒子组成。在循环充放电过程中，活性颗粒会随着锂离子的插入和提取而不断发生膨胀和收缩。在全固态电池中，活性粒子与固态电解质之间的界面处都伴随着高应力。诱导的高

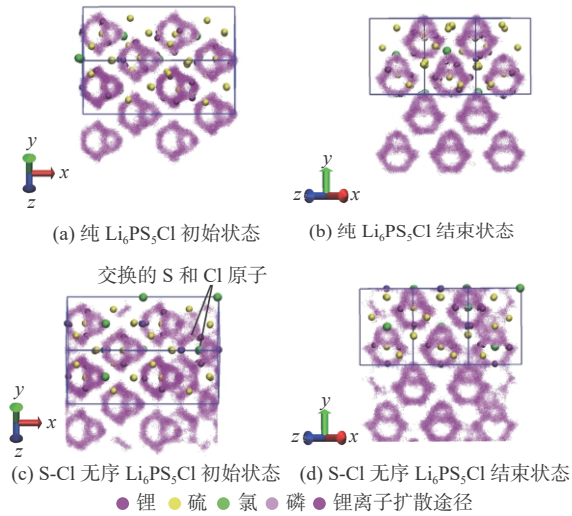


图 6 600 K 下观察到的 50 ps 的锂离子扩散轨迹[70]

Fig. 6 Observation of 50 ps lithium ion diffusion trajectories at 600 K[70]

应力可能导致活性颗粒和固态电解质的破碎及其之间界面的分层，如图 7 所示。图中：SPE 表示固体聚合物电解质，SCE 表示复合固体电解质，CE 表示复合正极，NCM 表示镍钴锰酸锂。Lewis 等[73] 指出，在固态电池中，不需要的中间相必须是离子导电的，同时是电子绝缘的。因为电子导电的中间相允许原始固态电解质不受控制的反应，因此，允许这种不需要的混合相的不稳定生长。在实际应用中，形成电子绝缘界面，新混合物的低离子电导率导致电池的高电阻，是固态电解质的一大挑战[74]。Han 等[75] 关于提高固态电解质电阻的实验工作如图 8 所示。图中：W/O ALD 表示未经过原子层沉积，SSE 表示固态电解质。在全固态电池中，锂的连续电镀和剥离伴随着循环过程中体积的膨胀，导致界面形成空洞和部分分层。这种局部分层失去接触减慢了锂离子在界面上的传输，显著增加了阻抗[76]。

当前研究多基于一维模型对全固态电池进行性能仿真，但一维模型无法直观体现颗粒位置导致的电池性能差异，同时一维模型基于均相结构假设，而实际上电极通常难以做到完全均质化。因此，研究颗粒尺寸及间距对电池性能的影响，必须建立精确的电池微观尺度模型[77]。为研究固态电解质孔隙率、裂纹形式及界面接触面积对固态电池的影响，郝增辉等[78] 利用电阻网络方法对固态电解质微观结构建模，建立了固态电池一维电化学耦合接触面积模型，构建了一维电化学与二维固态电解质电阻网络模型，并基于该物理模型进行了电化学阻抗谱仿真分析。Tian 等[79] 通过假设电流和锂浓度集中在接触区域，将不完美接触面积的影响纳入一维

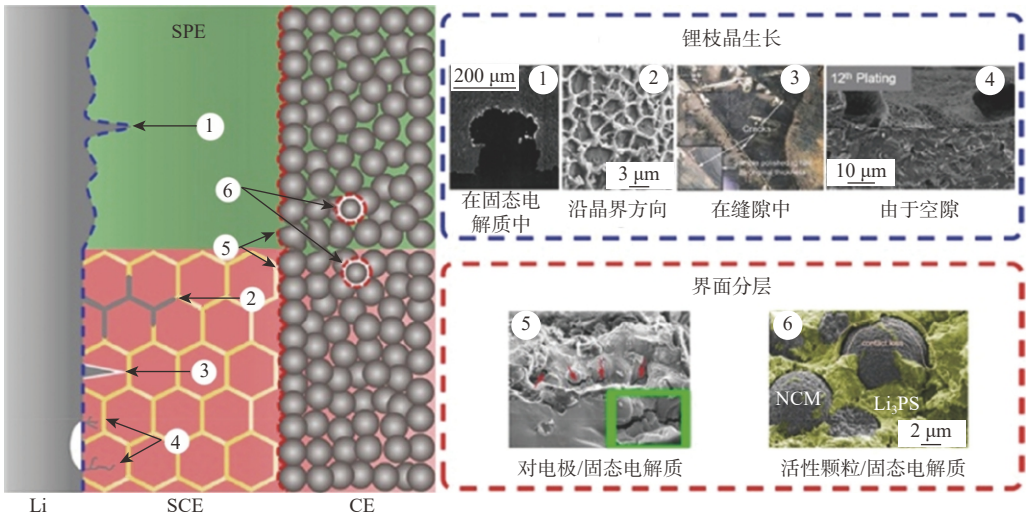


图 7　固态电池界面失效示意图

Fig. 7　Schematic diagram of interface failure of solid-state battery

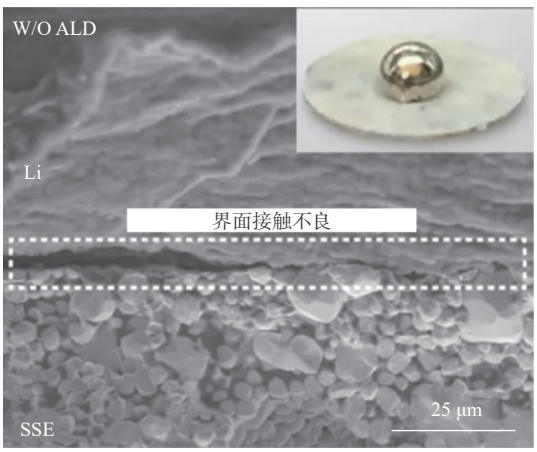


图 8　裂纹扩展程度与阻抗增加之间的直接相关性[75]

Fig. 8　Direct correlation between crack propagation degree and impedance increase[75]

Newman 电池模型，模拟了薄膜型 Li/LiPON/$LiCoO_2$ 全固态锂离子电池在不同速率和接触面积下的恒流放电过程，研究发现，在较低的截止电压下，相关性几乎呈线性，斜率为 1，而在较高的截止电压下，下降速率较慢，应用 Persson 的接触力学理论，计算了薄膜型和块状全固态锂离子电池的接触面积和压力，该模型用于评估应施加多少压力来恢复因接触面积损失而导致的容量下降。Javed 等[80] 利用有限元方法定量讨论界面电子/离子转移对固态电池正极放电特性的影响，研究了活性材料/活性材料、活性材料/固态电解质和固态电解质/固态电解质界面电化学过程产生的电化学电阻，以及与活性材料表面缓冲层涂层相关的电化学电阻，通过验证 $LiCoO_2$（LCO）、$LiNbO_3$ 和 $Li_{3.25}Ge_{0.25}P_{0.75}S_4$（LGPS）分别作为活性材料、涂层材料和固态电解质的实验结果，确定了电化学模型中使用的各种界面参数。Xiong 等[81] 通过间隔为 1 ms 的时间分辨电位测量来解释电化学反应和离子扩散动力学，揭示了经典的 Bulter-Volmer（B-V）方程偏离了电流密度、过电位和离子浓度之间的实际关系，建立了一个界面离子插层反应模型，考虑了锂离子（脱）插层的过剩驱动力，如图 9 所示，仿真结果表明，模型能够准确预测各种活性材料在颗粒和电极尺度上的充放电情况。

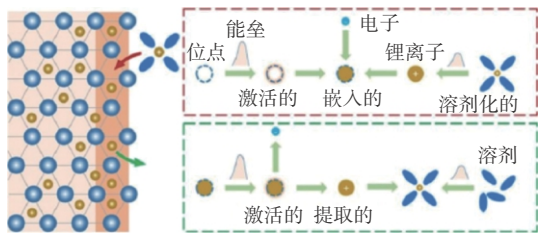


图 9　界面离子插层反应模型[81]

Fig. 9　Interface ion intercalation reaction model[81]

采用三维电极结构的模拟可以允许在计算中更准确地反映结构的效果。接触面积和锂扩散距离与活性电极材料颗粒的聚集程度有关，且这种聚集的影响在全固态电池中比在液态电解质电池中更显著。因此，应将活性电极材料颗粒聚集的影响纳入全固态电池的模拟中，以确保高精度。阮海辉等[82] 提出了一个将全固态电池的电化学过程和机械反应联系起来的模型，其中，空隙和裂缝的动力学是关键因素，模型仿真结果表明，空隙的增长是锂在界面和电极中不同扩散率导致分离的结果，引发了脱嵌锂过程中锂离子的非均匀分布，模拟揭示的裂纹和锂枝晶与实验观察非常相似，高堆积压力不能抑制裂纹和锂的渗透，但适当的侧向压应力可以防止硫化物固态电解质开裂，从而抑制锂枝晶的产生。

4.3.3　全固态电池离子输运机制

开发具有高锂离子导电率的新型固态电解质需

要了解其锂离子迁移机制。关于硫化物固态电解质中锂离子迁移机制的研究,按照单离子跃迁理论,Yang等[83]认为其具有较低的锂振动频率,从而导致晶格内的锂离子具有较大位移振幅和较低的迁移能垒,同时,其体心立方结构提供了较大的离子传输通道[84],表现出非常高的离子电导率。Mo等[85]认为这种电解质的快速扩散是通过多个离子协同迁移来实现的,当高能量位置的离子向低能量位置跃迁时,会将邻位低能量位置的离子迁移能垒降低,多离子协同迁移可以降低能垒,进而提升电解质的锂离子电导率。张田[86]以PEO、PEO+SiO_2、PEO+LGPS、PEO+LGPS+MoS_2这4种固态电解质为对象,采用分子动力学和DFT这2种方法,从锂离子扩散系数、界面处离子径向分布函数、PEO玻璃化转变温度、陶瓷颗粒表面态密度、异质结中平面平均差分电荷曲线和MoS_2对LGPS中锂离子结合能的影响等6个方面研究了复合电解质中锂离子传输特征及其背后原因。复合电极中的电荷输运限制与复合电极的离子/电子输运特性与复合正极内不同组分的比例及粒径、黏结剂与导电碳的种类、形貌及含量、混合/制备工艺等因素相关。近年来,由于模拟仿真技术的发展,相关研究开始增加,但目前还缺乏将制造参数与复合电极的微观结构和电化学性能相关联的计算模型,需要进一步结合电化学测量技术等开展定量研究[87]。

硫化物固态电解质因其高离子电导率和良好的机械性能,具有较大强度,可以防止锂凸出。“流体状”锂可以穿透强度大的固态电解质并使电池短路,这一违反直觉的事实也迫切需要一个相应的理论来理解这一现象。Tikekar等[88]假设离子输运不是反应,而是锂沉积的限制因素,当锂离子到达锂电极表面时,可以在现场立即还原,因此,沉积在锂表面的锂的数量是由化学电位梯度驱动的通量而不是B-V方程型反应控制的,基于该模型,发现通过在隔膜中固定一小部分阴离子并使用具有中等剪切模量的固态电解质,可以在电解质中实现稳定的电沉积。Mistry和Mukherjee[89]讨论了锂体输运和表面反应之间的竞争效应,发现固态电解质中的锂摩尔体积与锂负极中的锂摩尔体积之间的不匹配是导致电沉积不稳定的根本原因,而不是固态电解质的刚度。Ganser等[90]采用严格的电化学-力学描述体输运和界面反应动力学来预测界面稳定性,研究结果表明,电解质的刚度和输运性质对界面稳定性同样重要,应在电化学性能和力学性能之间找到一个正确的权衡,而不是无条件地增加固态电解质的刚度。

结合硫化物固态电解质离子输运机制的最新研究发现,离子输运机制的核心矛盾已从晶格扩散转向界面工程,技术聚焦于“掺杂+包覆+原子层沉积”综合方案。硫化物电解质通过晶体骨架中的三维锂离子扩散通道实现离子迁移,Song等研究发现,掺杂稀土元素可将离子电导率提高到10^{-2} S/cm级别,原因是稀土离子扩大了晶格间距,可降低锂离子迁移能垒[91]。但是,固-固接触界面缺陷会阻碍离子运输[92],环境湿度过大会使硫化物遇水分解产生硫化氢破坏晶体结构和阻塞离子通道[93],这2种因素都会导致离子传输速度下降。

4.3.4 固态电解质间隙诱导应力和断裂

固态电解质在运行中面临破碎的可能,除界面形成引起的应力及其对扩散和反应驱动力的影响外,破坏的一个关键机制是裂纹的形成和扩展。目前,文献中固态电解质破裂的现有模型基本没有。然而,有许多研究致力于电极颗粒中裂纹扩展的建模,这可以作为解决裂纹存在下固态电解质化学力学问题的基础。在力学界有一些模型,从中可以得出有用的界面生长公式,因为它们描述的物理现象与全固态电池中的界面反应有相似之处。这些现有公式与模拟全固态电池现象的主要区别在于控制每个反应的不同化学驱动力。

为捕捉固态电解质的整体力学行为,Ganser等[94]提出了热力学一致的宿主材料输运理论,使用亥姆霍兹能量公式,该公式考虑了物质输运、弹性变形和静电,扩展了Newman和Balsara[95]的浓溶液理论,以纳入力学效应,将公式专门用于二元电解质的情况。Bucci等[96]提出了类似的热力学一致处理方法,模拟了晶格中的离子传输。上述模型为模拟固体中的离子输运提供了基础,然而这些工作忽略了相间形成和生长的基本物理规则,即反应诱导的结构相变及与机械变形和应力相关的耦合。在固态电解质中,引起机械降解的是新形成的间相的反应引起的力学效应及其相关的化学和结构相变[97]。

固态电解质中间相离子电导率较低,还存在关键机械问题,特别是与导致新相的反应相关的体积膨胀,这种体积膨胀发生在受限的全固态结构中,会引入内应力,导致可能的断裂和降解。Tippens等[97]使用X射线断层扫描显示了$Li_{1.4}Al_{0.4}Ge_{1.6}(PO_4)_3$(LAGP)固体电解质中裂纹的形成,认为这些裂纹的存在和随后的接触损失导致了电池运行中的高阻抗,而不是界面本身的输运性质,同时还利用化学-力学连续体理论[98]研究了上述固态电池中的应力分布。

针对活性粒子的损伤，已经建立了相场损伤模型，这些模型可以将裂纹退化参数作为化学变量集的附加内部变量，同时考虑到裂纹的起裂和扩展。Miehe 等[99]提出了有限应变化学弹性理论与相场建模相结合来研究锂离子电极颗粒的断裂，正则化裂纹模型使用破坏表面的主应力准则和退化函数来破坏存储的能量和物质在裂纹面上的迁移率。Yamakawa 等[100]研究了断裂扩展对力学的影响，以及随后对 $LiCoO_2$ 多晶颗粒中锂浓度和相分离的影响，估计了晶粒尺寸和晶体取向对断裂程度的影响。Anand 等[101]开发了一种替代的基于热力学的损伤模型。Shi 等[102]表明，活性材料体积分数提高后，固态电解质相必须仍保持连续的离子传输网络，否则会出现离子渗流不足和活性材料利用率下降。Bucci 等[103]基于断裂内聚理论推导出了一个解析模型，并表示减小粒径无助于避免活性颗粒的界面分层，当活性颗粒的体积变化达到 7.5% 左右时，界面分层就会开始，顺应性固态电解质（弹性模量 E<25 GPa）可以适应体积变化并延迟界面分层的发生。Barai 等[104]建立了多尺度数学模型，将正极材料的摩尔体积变化、正极与 LLZO 的弹性性质、界面断裂阈值和界面反应动力学，与分层面积、阻抗增长和容量衰减联系起来，研究结果表明，减小活性材料的有效体积变化、降低固态电解质刚度及提高界面断裂强度，均有利于缓解正极/固态电解质界面分层。Rezaei 等[105]建立了一个更全面的化学-机械耦合内聚破裂模型，表明该模型可用于研究全固态电池中主要的降解机制，如活性颗粒内部的晶间破碎。Fathiannasab 等[106]基于复合电极的重构形貌建立了化学-力学模型，阐明了颗粒/电解质界面和空隙空间对锂化诱导应力演化的影响，发现空隙可以通过容纳活性颗粒的膨胀来缓解应力的形成，采用外压改善离子通路可能导致断裂，而刚度较低的固态电解质会使复合电极中的应力松弛，加剧活性颗粒的各向异性位移。Wan 等[107]研究表明，界面应力也会因界面处的强电场而增强。

4.3.5　固态电解质力学特性

在了解固态电解质力学行为的过程中，研究和模拟固态电解质本身的力学行为和性能十分有必要，因为电解质的弹性、塑性和断裂性能决定其如何响应枝晶生长、由界面形成引起的应力、裂纹扩展和其他相关的力学问题[108]。Wang 等[109]将 LGPS 与 $Li_4P_2S_7$ 和 $Li_3P_7S_{11}$ 共同研究，与 LGPS 的四边形结构不同，它们具有三斜结构，研究表明，LGPS 具有更高的弹性常数，LGPS 的 Pugh 模量比为 2.12，表明了其相对延展性。Yang 等[110]对 Li_3PS_4 的 2 个 β 相和 γ 相进行了详细研究，表明 γ 相是比 β 相更硬的材料，2 种材料均表现出相对的延展性，γ 相和 β 相的 Pugh 模量比分别为 2.26 和 2.89。

研究表明，锂/固态电解质界面特征在锂枝晶形成中起到关键作用，而不是固态电解质的体积性质。高剪切模量并不能保证无枝晶的电池体系；相反，增加锂/固态电解质界面的附着力可以帮助缓解枝晶问题，这可以通过电池层面的外部压力来实现。Zhang 等[111]确定了至少 20 MPa 的首选堆叠压力，可以减少空隙体积并保持相对较小的界面阻力。Tu 等[112]分析了影响界面稳定性的因素，如固态电解质中的堆压、界面电荷传递系数和传质速率，结果表明，低堆压（3～5 MPa）对界面沉积几乎没有影响，但低离子电导率和低界面面积比电阻会增加锂沉积的不均匀性。Verma 等[113]利用并开发了该模型来辨别外部堆叠压力对多晶/非晶固体电解质结构中界面不稳定性的影响，固态电解质的微观结构包括晶粒、晶界和孔洞，其特征是有效输运和力学参数，结果表明，外部压力、高温和低表面粗糙度可以实现稳定沉积。

预先存在的裂纹也可能作为缺陷点发挥作用，可能会促进锂枝晶在那里生长。Porz 等[114]发现固态电解质由于晶体表面细小的划痕而无法抵抗锂枝晶的传播，与单晶的剪切模量无关，锂离子会在晶粒表面楔开先前存在的缺陷，建立了锂填充裂纹扩展的一维电化学力学模型，并解释了填充锂产生的压力足够大，可以通过 Griffith 样裂纹扩展穿透脆性固态电解质，如图 10 所示。Barroso-Luque 等[115]建立了电沉积诱发塑性流动的数学模型，该模型表明在孤立缺陷处沉积产生的压力最大值由边界牵引势决定。Yuan 等[116]提出了用于裂纹扩展和锂枝晶生长的耦合电化学-机械相场模型，发现在 10 MPa 以上的堆积压力和较长且边缘、角度锐利的缺陷都能促进裂纹扩展和枝晶生长。同时，Yuan 等[117]还考虑了电池模型、力学模型、相场模型和短路模型，建立了基于物理的全耦合电化学-力学模型，模型显示，高充电率和固态电解质的低电导率会加速电

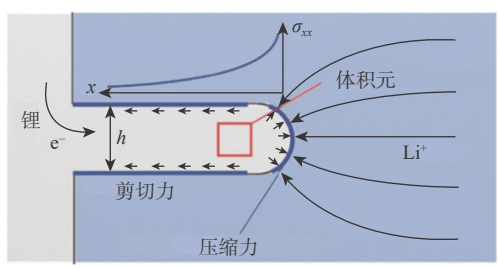


图 10　枝晶通过扩展裂纹扩展的加压裂纹示意图[114]

Fig. 10　Schematic diagram of pressurized cracks propagated by dendrites through crack propagation[114]

池的电化学失效，因为过电位驱动的应力促进了裂纹中的枝晶生长并渗透到固态电解质中。Shishvan等[118]提出了一种新的机制，将枝晶生长视为一种攀爬边位错，这种位错可以在没有过量锂+恒定化学势的情况下自发发生，进入枝晶尖端的锂+通量提供的自由能足以使枝晶生长，枝晶的生长速率也可以通过模型计算得到[119]。

多尺度建模方法在全固态锂电池中关键材料（正极、负极、固态电解质）及其界面特性的状态描述和性能分析中发挥了重要作用，可表征界面失效机制、材料微观结构演化、离子输运行为、应力与断裂行为等关键行为，为全固态电池的综合性能仿真分析提供了理论支撑。

5 全固态电池整体电化学性能建模与仿真

全固态电池作为一种典型的电化学储能系统，涵盖多种电化学反应、质量传递、电荷传递及热量传递等物理化学过程[120]。先进实验手段在一定程度上可以阐明电池材料对电池性能的作用机理，但存在单一变量难以控制、重复性操作多、试验周期长、试错成本高的局限性，特别是难以形成材料设计对电池评估的直观认识[121]。随着计算机技术的飞速发展，特别是基于有限元、有限体积、有限差分等数值仿真技术的日趋完善，建立数学物理模型并通过数值模拟求解已成为分析和解决复杂多物理场、开展产品预估、设计及优化的重要途径[122]。

5.1 全固态电池电化学特性仿真

传统锂电池采用P2D模型作为电化学仿真基础，可通过有效的微结构参数、活性物质的球坐标2个步骤将复杂的三维微结构分解为计算方便的二维系统，虽然无法直接应用于全固态电池，但也为全固态电池电化学模型的建立提供了理论依据。

Danilov等[123]首次建立了用于描述全固态电池体系的一维电化学模型，模型包含了界面电荷转移动力学、锂在电极中的扩散、锂离子在固态电解质中的扩散和迁移，但该模型并未深入探讨电极设计与电化学行为的关系。Zhang等[124]建立了一维全固态电池的数学方程用来对电池进行状态估计。Kazemi等[125]根据电化学实验结果构建了浓度相关的正极扩散系数，进一步提高了模型精度。李学鹏[126]采用Comsol建立全固态锂电池计算模型，仿真分析了全固态电池的充放电特性，研究了各种放电电流下的电压损失，仿真结果表明，较高的放电电流会使电极处的平衡浓度偏差较大，随着放电速率的增加，锂的浓度梯度变化变得更陡峭，电池能量耗尽前电压急剧下降以获得更高的放电率。

Ansah等[127]开展了基于有限元模拟的固态锂离子电池参数综合研究，提出了数值模型来研究一些参数以优化固态电池性能，该模型考虑了锂离子在电极、电解质中的扩散及固态电池内的电化学反应，结果发现，增加正极厚度会增加电池容量，而减少电解质厚度则会提高电池的容量，如图11所示。

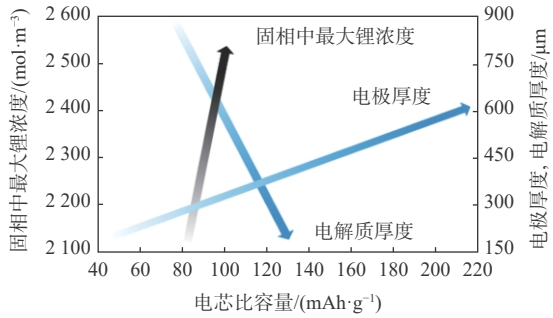


图11 不同设计参数与固态电池性能之间的关系[127]

Fig. 11 Relationship between different design parameters and performance of solid-state batteries[127]

5.2 全固态电池力学特性仿真

电池电化学行为通常会导致正极和负极发生体积膨胀或收缩，而体积变化引起的应力反过来会影响电化学行为，应力通常受电化学和热场影响。除之前提到的颗粒断裂、界面分层等机械降解，应力对锂离子的传输和电化学反应都有显著影响[128]。应力可以增强锂向颗粒内部的扩散，促进快速嵌入，这一现象对硫化物电解质全固态电池效应更为明显，在外部施加应力下，离子扩散和迁移速率会被改善[129]。

准确分析全固态电池中的离子传输对于提高和评估性能至关重要。对于硫化物全固态电池，需要耦合应力和电场模拟，因为电池是加压的，这会影响空隙空间，进而影响离子传输。Kodama等[130]提出了一种通过耦合非线性应力分析和电场分析来分析加压全固态电池离子电导率的新方法，非线性应力分析可以估算固体电解质的应力分布，发现源自孔隙的非线性应力-应变特性随着压力的增加而减小。

涉及力-电化学耦合的全固态电池高精度建模近年来备受关注，得到了较好的发展。Grazioli等[131-132]建立了固体电解质的力-电化学耦合模型，给出了均匀平面电池在恒流条件下的解析解，力学性能对电解质电导率有很大影响，电解质静电位预测值差异甚至高达38%，这表明在电池设计过程中应考虑应力和变形的影响。Ganser等[94]拓展了浓溶液理论，提出了基于亥姆霍兹能的固态电解质力学-电化学耦合理论，考虑了静电学、组分传输和非

线性弹性力学间的相互作用。Bistri 等[133] 制定并从数值模拟上实现了力学-电化学本构框架，该框架捕获了电极微观结构中复杂的化学-机械多粒子相互作用。Song 等[134] 提出了考虑物种扩散、界面转移动力学和机械变形的全固态锂离子薄膜电池力学-电化学耦合模型，该模型考虑了物种扩散、界面转移动力学和机械变形，将具有 $LiCoO_2$ 正极、石墨负极和聚合物固态电解质的全固态薄膜电池应用于模拟过程，以研究弯曲效应，结果表明，薄膜电池可以向正极弯曲以增强锂化能力，并降低电极/电解质界面处的应力。

电极或固态电解质的机械稳定性不佳会造成电池的电化学性能大幅下降，正极材料在锂离子脱嵌时会发生相变或晶格膨胀/收缩，体积效应导致正极材料与导电剂的界面在充放电过程中不断生成-破碎，消耗可迁移的锂离子，使电池容量下降；同时，正极材料在充放电时体积变化会造成其与导电剂和集流体发生剥离，使电池阻抗大幅上升[135]。Bucci 等[34] 利用力聚区模型对因正极材料体积变化导致全固态电池正极侧产生裂纹的行为进行模拟，如图 12 所示，证实了只有具备低断裂能和高体积变化的正极材料才会使全固态电池正极侧在充放电过程中产生裂纹，并使其在正极侧蔓延传播。图中，SE 表示固态电解质与碳混合物。

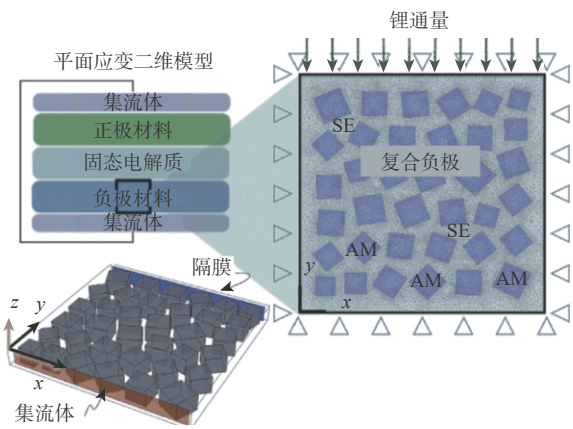


图 12　正极材料有限元模型的几何、离散化和边界条件[34]

Fig. 12　Geometry, discretization, and boundary conditions of finite element model for positive electrode materials[34]

5.3　基于多物理场耦合的全固态电池建模及性能仿真

目前，围绕全固态电池仿真已开展了一些研究工作，但其仿真建模基础理论仍相对薄弱，特别是基于传热、应力、电化学等多个物理场耦合建模方面还没有全面认识。电池系统在充放电时将同时发生多个物理化学过程，如电场驱动的迁移过程、浓度场驱动的扩散过程、电化学场驱动的电荷转移过程、温度场驱动的传热过程及应力场驱动的形变（应变）过程等，而以上过程又相互影响，互相制约。因此，有必要考虑全固态电池工作中经历的各种物理化学过程，建立考虑热-力-电化学耦合的精确仿真模型，进而实现电池性能的精准预测。在多物理场耦合仿真中，电压变量作为关键物理量，其处理方式需根据具体耦合机制和场景进行针对性设计。对于简单的系统，可以直接将 2 个物理场进行耦合，以关键参数作为耦合元素；对于复杂的系统，建议采用模块化建模，先单场验证，再逐步耦合。

1993 年，Doyle 和 Newman 等[136] 设计了 P2D 模型，通过简化厚度方向，采用分段线段的方式代表电池不同结构，从而以一维尺度进行分析，设定平行电极方向各部分属性一致，正负极颗粒是大小一致的球形，仅考虑颗粒内部径向的扩散等假设，能够准确地模拟出充放电过程中电荷和离子的传输。Fabre 等[137] 开发了 $Li/LiPON/LiCoO_2$ 全固态薄膜微电池的数学模型，其是等温、一维的，考虑了锂在正极中的扩散和电子迁移、锂离子在固体电解质中的迁移及电极/电解质界面处的电荷转移动力学，使用各种电化学技术，如恒电流间歇滴定技术和电化学阻抗谱，将输入参数的温度依赖性引入模型，从而预测微电池在不同温度下的运行情况。Pang 等[138] 建立了固态锂电池电化学-热耦合模型，比较了并联和串联 2 种单体连接方式的产热量及能量密度，串联电池的平均温升为 33.91 K，其焦耳热高达 146 W/m^2，而并联电池的平均温升仅为 18.51 K，其焦耳热可忽略不计，并且并联具有更高的能量密度，研究结果表明，全固态电池的生热特性与液态锂电池相似，电化学反应过程中会产生不可逆极化热与可逆熵热，进而由于温度不均匀性而产生热应力。在单个活性颗粒开发的理论框架的基础上，模拟全电池行为的第一个进展是考虑使用固体电极的半电池或全电池模型。Wan 等[107] 开发了一种半电池连续化学力学模型，包括对正极、负极、静电场和电化学反应的机械变形和应力的考虑。孙哲韬等[139] 基于 Comsol 软件对 LPON 基全固态电池体系进行了仿真，利用电流分布完成了固体电解质传质模型的建立，利用电极表面仿真负极的反应表面，再利用稀物质传递加上电极表面模块完成了整个电池电化学体系的构建，结合固体力学与固体传热进行多物理场耦合，完成了对体系应力及温度的仿真与监控，仿真结果表明，全固态电池体系的浓度梯度、应力预分布、传热传质过程的控速步骤与充放电过的体积变化等都会对电池的性能与安全管理产生不同的影响。Bower 等[140] 推导了控制锂离子半电池中变形、应力和电流的连续场方程和本

构方程，考虑了通过系统的质量输运、阳极和阴极的变形和应力、静电场及电极/电解质界面上的电化学反应，该模型用于预测一维半电池模型电池中应力和电势的变化，如图 13 所示。

为建立具有复合正极电极的全固态电池的电化学模型，张希等[141]创新性地考虑了 2 个固态电池属性：不完美的固-固接触界面和电极、电解质界面处的空间电荷层，为提高模型的准确性，引入依赖于活性物质中锂浓度的可变离子扩散系数，如图 14 所示，研究结论包括：①通过拉普拉斯变换与帕德近似方法获取了降阶传递函数，降低了模型的计算复杂性而不影响准确性；②研究采用实验数据，验证了不同工况下模型的预测能力，在 15 个工况范围内，模型与实验的平均均方根误差为 19.5 mV，弛豫结束时的平均误差为 2.8 mV；③通过一些关键参数的敏感性分析，发现正极颗粒内部的浓差极化和界面反应引起的过电势，是限制固态电池性能的因素。Fang 等[142]为深入了解电解质老化对电池衰退的影响机制，建立了一个连续体模型，用于定量分析锂电池在时间老化过程中的容量演变，模拟揭示了电解质老化引起的容量退化现象，还详细探讨了放电速率、工作温度、电解质中锂盐浓度及电解质厚度的影响。Bates 等[143]提出了一种模拟方法，用于研究材料和几何形状对锂离子固态电池性能的影响，可以轻松探索不同的材料特性和电池几何形状，以优化锂离子固态电池的性能，该仿真能够显示放电循环中任何时候的放电曲线、热通量、锂离子浓度、电子和固体锂。

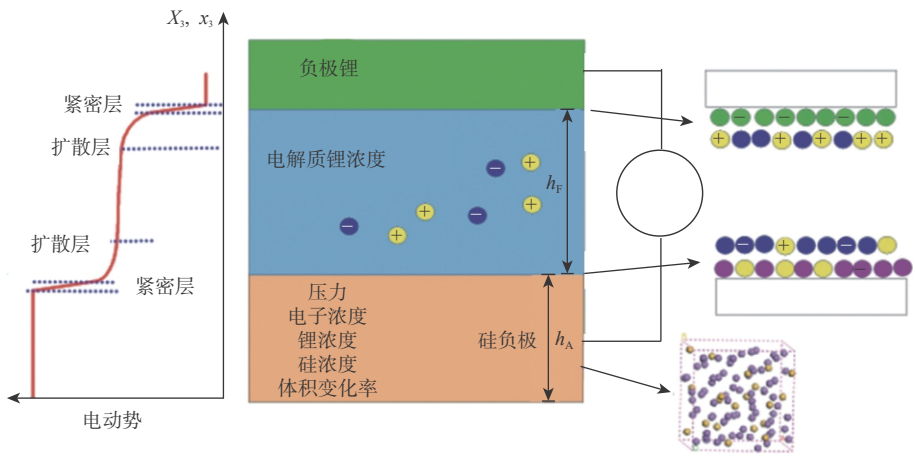


图 13 一种理想化的模型化锂离子半电池[140]

Fig. 13 An idealized model of a Li-ion half-cell[140]

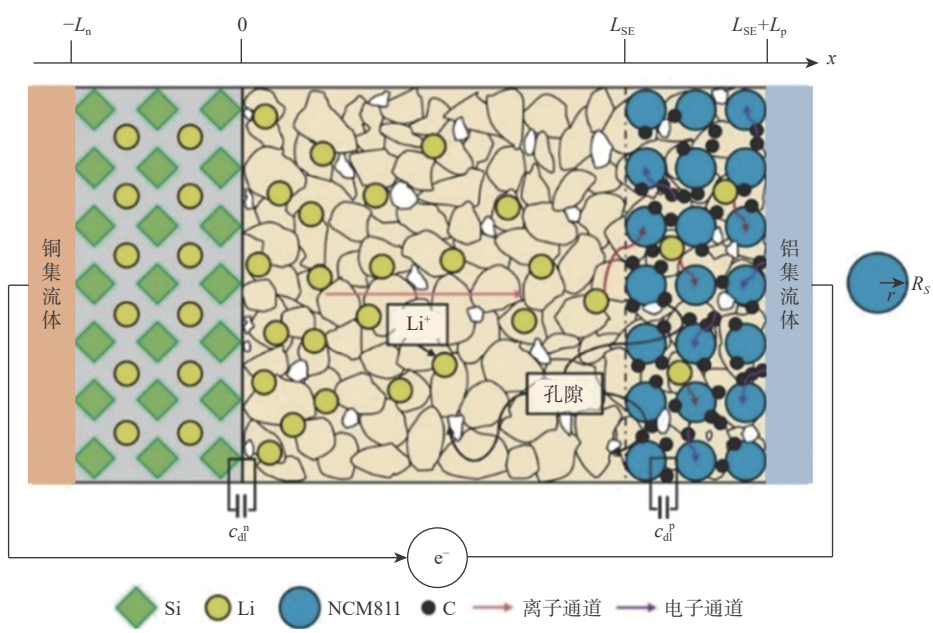


图 14 考虑不完美固-固接触界面与电极和电解质界面处的空间电荷层的全固态电池模型[141]

Fig. 14 A fully solid-state battery model considering space charge layer at imperfect solid solid contact interface and space charge layers at electrode and electrolyte interface[141]

三维全固态锂离子电池的演示一直是电池界众多研究人员的长期目标。Talin 等[144] 描述了使用与半导体器件工艺兼容的材料和薄膜沉积方法制造的三维全固态电池的实验测试和模拟，这些三维电池由硅微柱组成，电池层使用物理气相沉积顺序沉积在硅微柱上，三维电池的功率性能明显落后于类似制备的平面电池，使用有限元建模对实验结果进行分析表明，功率性能差的根源是三维电池的结构不均匀性，再加上正极中电解质离子电导率和扩散速率低，导致内部电流密度分布高度不均匀，正极利用率低。

全固态电池电芯层面的建模和仿真主要体现在电化学特性、力学特性，并且可以进行多物理场耦合来实现对实际工况的高精度模拟，通过系统分析不同模型在揭示材料设计、界面行为、应力响应及多场耦合机制中的应用与局限，可为全固态电池的性能优化与结构设计提供理论支撑与仿真依据。

6　结 束 语

本文总结了全固态电池特性建模和仿真的研究进展，重点聚焦于高比容量正极、复合锂金属负极、硅基负极材料、固态电解质的微观结构建模与仿真，全固态电池电化学性能、力学性能等，并在固态电解质材料微观结构、界面特性、离子输运机制、诱导应力和断裂、力学特性等方面开展了专题分析。现有研究成果有力支撑了对全固态电池工作机理的揭示和分析，并指导了全固态电池研发设计。

展望未来全固态电池建模与仿真工作，还有以下几个方面可以开展持续研究：

1）正极材料。复合正极将在全固态电池中发挥关键作用，对其耦合机械-化学行为的研究尚处于起步阶段，许多研究途径仍有待探索。

① 对实验相关的三维固态正极微观结构进行详细的连续体建模，并考虑微观结构的损伤和破坏行为对正极电化学性能和力学性能的影响。

② 在直接模拟实验相关微观结构的背景下，改进微观结构相关相的本构描述，提出模拟代表性体积单元的适当边界条件，对各向异性特性、热力学行为、非弹性变形等现象进一步揭示。

2）负极材料。复合锂金属负极仍面临锂枝晶生长、沉积不均匀等问题，通过下述手段可持续改善。

① 针对锂金属黏塑性变形和塑性流动对电沉积动力学的影响进行建模，基于锂金属本构理论模拟应力在枝晶生长区域附近发展的方式，揭示应力与生长动力学的耦合作用，并开展实验验证结合。

② 建立完全耦合的电化学-力学模型，模拟固态电解质中枝晶的形成和生长行为。

③ 与固态电解质微观结构协同考虑，描述锂枝晶生长及在固态电解质中传播的规律，重点是考虑固态电解质缺陷、裂纹和晶界的存在，这些缺陷、裂纹和晶界可以作为锂枝晶的成核和生长位点。

④ 为模拟锂负极/固态电解质界面上的空洞形成和接触损失，建立微观尺度的孔隙形成和生长机制的模型，并考虑宏观尺度的界面变形和形态演化，将空洞的形成和形态变化与界面处的阻抗增加联系起来。

3）固态电解质。固态电解质建模还处于起步阶段，特别是对负极/固态电解质界面形成的界面相进行建模对于理解应力产生、断裂和分层至关重要，目前还没有针对这种现象的有效模型。

① 从建模角度捕捉相间形成和演化现象仍然是固态电解质实现的挑战之一，需要建立能够适应化学诱导相变的扩散变形反应理论，以解决固态电解质中相间形成的问题。

② 针对固态电解质损伤和断裂行为的建模，该模型考虑能够以阻抗增加的形式捕获机械退化和电化学退化。

③ 在筛选高离子电导率、高可加工性、低成本的固态电解质时，使用完全耦合的化学-力学连续体模型作为基础，结合人工智能、机器学习等手段，将从完全耦合的化学-机械连续体模型中获得的结果作为特征添加到神经网络等数据科学方法的训练空间中，高效地训练筛选出合适的固态电解质材料。

④ 建立描述负极/固态电解质处形成空隙和失去接触行为的模型，以掌握相关材料特性、微观结构特征和操作要求，减少固态电解质失效。

⑤ 加强固态电解质中间相形成反应动力学理论和模型的研究，能够匹配化学诱导相变的扩散变形反应行为。

虽然针对全固态电池分子动力学计算、关键材料及界面特性、电化学性能等开展了不少研究工作，对其工作机理等进行了一定揭示，但目前全固态电池仍有大量科学问题和工程化问题尚未解决，距离大规模产业化还有不少时间。

为进一步提升全固态电池性能，应继续深入探索以下共性科学问题及工程化问题：

1）开展理论计算分析失效机制研究，可以理论计算分析电池机理，为原理分析提供强有力的辅助证据，也可以在没有强有力的表征方法的支持下提供相对可靠的预测分析手段。

2）开发新的表征方法，探究硫化物基全固态电池中存在的大量未明确的机理性问题。硫化物基

全固态锂电池由于硫化物电解质空气不稳定特性，使得其表征手段较为单一，需要丰富其表征手段。

3）在低N/P比下，提高正极、负极材料的利用率和性能，提高电池容量和稳定性。正极侧，正极混料中、正极与电解质层间引入稳定性更高的电解质，隔绝电解质与高电压活性材料直接接触，避免两者间的副反应发生。负极侧，改性锂金属负极表面，引入能稳定硫化物电解质层的缓冲物质，改善电解质和锂金属的稳定性。

4）合成具有优异空气稳定性且离子电导率较高的硫化物电解质。目前，硫化物固态电解质的合成路径还比较复杂，合成条件要求较高，需要开发简单高效、重复性高且适合大批量制备的合成方法，为其在全固态电池规模化应用提供保障。

5）通过改善现有工艺或开发新工艺制备机械性能优异且薄的电解质膜。结合电解质的高室温锂离子电导率，可采用液相法或干混法来制备超薄的电解质薄膜，兼顾高性能及高安全。

6）全固态电池生产工艺优化及专用设备开发。将全固态电池的组装工艺与当前的液态电池制备工艺相结合，基于当前电池组装技术，设计适合全固态电池的制备工艺和路线，对于降低固态电池成本和推动实用化进展意义重大。

参考文献（References）

[1] Zhao Q, Stalin S, Zhao C Z, et al. Designing solid-state electrolytes for safe, energy-dense batteries[J]. Nature Reviews Materials, 2020, 5(3): 229-252.

[2] Lindorff-Larsen K, Piana S, Dror R O, et al. How fast-folding proteins fold[J]. Science, 2011, 334(6055): 517-520.

[3] Hertig S, Latorraca N R, Dror R O. Revealing atomic-level mechanisms of protein allostery with molecular dynamics simulations [J]. PLoS Computational Biology, 2016, 12(6): e1004746.

[4] Foroozan T, Soto F A, Yurkiv V, et al. Synergistic effect of graphene oxide for impeding the dendritic plating of Li[J]. Advanced Functional Materials, 2018, 28(15): 1705917.

[5] Deng Z, Wang Z B, Chu I H, et al. Elastic properties of alkali superionic conductor electrolytes from first principles calculations[J]. Journal of the Electrochemical Society, 2016, 163(2): A67-A74.

[6] Ngandjong A C, Rucci A, Maiza M, et al. Multiscale simulation platform linking lithium ion battery electrode fabrication process with performance at the cell level[J]. The Journal of Physical Chemistry Letters, 2017, 8(23): 5966-5972.

[7] Hao F, Mukherjee P P. Mesoscale analysis of the electrolyte-electrode interface in all-solid-state Li-ion batteries[J]. Journal of the Electrochemical Society, 2018, 165(9): A1857-A1864.

[8] Yu H C, Taha D, Thompson T, et al. Deformation and stresses in solid-state composite battery cathodes[J]. Journal of Power Sources, 2019, 440: 227116.

[9] Tian H K, Liu Z, Ji Y Z, et al. Interfacial electronic properties dictate Li dendrite growth in solid electrolytes[J]. Chemistry of Materials, 2019, 31(18): 7351-7359.

[10] Tian H K, Xu B, Qi Y. Computational study of lithium nucleation tendency in $Li_7La_3Zr_2O_{12}$ (LLZO) and rational design of interlayer materials to prevent lithium dendrites[J]. Journal of Power Sources, 2018, 392: 79-86.

[11] Ahmad Z, Viswanathan V. Quantification of uncertainty in first-principles predicted mechanical properties of solids: application to solid ion conductors[J]. Physical Review B, 2016, 94(6): 064105.

[12] Ahmad Z, Xie T, Maheshwari C, et al. Machine learning enabled computational screening of inorganic solid electrolytes for suppression of dendrite formation in lithium metal anodes[J]. ACS Central Science, 2018, 4(8): 996-1006.

[13] Su S M, Ma J B, Zhao L, et al. Progress and perspective of the cathode/electrolyte interface construction in all-solid-state lithium batteries[J]. Carbon Energy, 2021, 3(6): 866-894.

[14] Liao Y L, Hu J K, Fu Z H, et al. Integrated interface configuration by in-situ interface chemistry enabling uniform lithium deposition in all-solid-state lithium metal batteries[J]. Journal of Energy Chemistry, 2023, 80: 458-465.

[15] Wang A N, Li J, Yi M Y, et al. Stable all-solid-state lithium metal batteries enabled by ultrathin LiF/Li_3Sb hybrid interface layer[J]. Energy Storage Materials, 2022, 49: 246-254.

[16] Umeshbabu E, Zheng B Z, Yang Y. Recent progress in all-solid-state Lithium-Sulfur batteries using high Li-ion conductive solid electrolytes[J]. Electrochemical Energy Reviews, 2019, 2(2): 199-230.

[17] Wang Y Z, Zhu G M, Zeglio E, et al. N-type organic electrochemical transistors with high transconductance and stability[J]. Chemistry of Materials, 2023, 35(2): 405-415.

[18] Díez-Pascual A M, Rahdar A. Graphene-based polymer composites for flexible electronic applications[J]. Micromachines, 2022, 13(7): 1123.

[19] Aïssa B, Ali A, El-Mellouhi F. Oxide and organic-inorganic halide perovskites with plasmonics for optoelectronic and energy applications: a contributive review[J]. Catalysts, 2021, 11(9): 1057.

[20] Koerver R, Aygün I, Leichtweiss T, et al. Capacity fade in solid-state batteries: interphase formation and chemomechanical processes in nickel-rich layered oxide cathodes and lithium thiophosphate solid electrolytes[J]. Chemistry of Materials, 2017, 29(13): 5574-5582.

[21] Ferguson T R, Bazant M Z. Nonequilibrium thermodynamics of porous electrodes[J]. Journal of the Electrochemical Society, 2012, 159(12): A1967-A1985.

[22] Renganathan S, Sikha G, Santhanagopalan S, et al. Theoretical analysis of stresses in a lithium ion cell[J]. Journal of the Electrochemical Society, 2010, 157(2): A155.

[23] Christensen J. Modeling diffusion-induced stress in Li-ion cells with porous electrodes[J]. Journal of the Electrochemical Society, 2010, 157(3): A366.

[24] Mykhaylov M, Ganser M, Klinsmann M, et al. An elementary 1-dimensional model for a solid state lithium-ion battery with a single ion conductor electrolyte and a lithium metal negative electrode[J]. Journal of the Mechanics and Physics of Solids, 2019,

123: 207-221.

[25] Zelič K, Katrašnik T, Gaberščček M. Derivation of transmission line model from the concentrated solution theory (CST) for porous electrodes[J]. Journal of the Electrochemical Society, 2021, 168(7): 070543.

[26] Garcia R E, Chiang Y M, Carter W C, et al. Microstructural modeling and design of rechargeable lithium-ion batteries[J]. Journal of the Electrochemical Society, 2005, 152(1): A255.

[27] Xu R, Yang Y, Yin F, et al. Heterogeneous damage in Li-ion batteries: experimental analysis and theoretical modeling[J]. Journal of the Mechanics and Physics of Solids, 2019, 129: 160-183.

[28] Bielefeld A, Weber D A, Janek J. Microstructural modeling of composite cathodes for all-solid-state batteries[J]. The Journal of Physical Chemistry C, 2019, 123(3): 1626-1634.

[29] Purkayastha R, Mcmeeking R. Stress due to the intercalation of lithium in cubic-shaped particles: a parameter study[J]. Meccanica, 2016, 51(12): 3081-3096.

[30] Bucci G, Nadimpalli S P V, Sethuraman V A, et al. Measurement and modeling of the mechanical and electrochemical response of amorphous Si thin film electrodes during cyclic lithiation[J]. Journal of the Mechanics and Physics of Solids, 2014, 62: 276-294.

[31] Lyu D D, Ren B, Li S F. Failure modes and mechanisms for rechargeable lithium-based batteries: a state-of-the-art review[J]. Acta Mechanica, 2019, 230(3): 701-727.

[32] Bucci G, Swamy T, Bishop S, et al. The effect of stress on battery-electrode capacity[J]. Journal of the Electrochemical Society, 2017, 164(4): A645-A654.

[33] Al-Siraj M, Stein P, Xu B X. The effect of morphology changes and mechanical stresses on the effective diffusivity of solid electrolyte for lithium ion batteries[J]. Journal of the Electrochemical Society, 2020, 167(2): 020535.

[34] Bucci G, Swamy T, Chiang Y M, et al. Modeling of internal mechanical failure of all-solid-state batteries during electrochemical cycling, and implications for battery design[J]. Journal of Materials Chemistry A, 2017, 5(36): 19422-19430.

[35] Zhang T, Kamlah M. Mechanically coupled phase-field modeling of microstructure evolution in sodium ion batteries particles of Na_xFePO_4[J]. Journal of the Electrochemical Society, 2020, 167(2): 020508.

[36] Cogswell D A, Bazant M Z. Size-dependent phase morphologies in $LiFePO_4$ battery particles[J]. Electrochemistry Communications, 2018, 95: 33-37.

[37] Bai P, Cogswell D A, Bazant M Z. Suppression of phase separation in $LiFePO_4$ nanoparticles during battery discharge[J]. Nano Letters, 2011, 11(11): 4890-4896.

[38] Mussa A S, Klett M, Lindbergh G, et al. Effects of external pressure on the performance and ageing of single-layer lithium-ion pouch cells[J]. Journal of Power Sources, 2018, 385: 18-26.

[39] Sauerteig D, Hanselmann N, Arzberger A, et al. Electrochemical-mechanical coupled modeling and parameterization of swelling and ionic transport in lithium-ion batteries[J]. Journal of Power Sources, 2018, 378: 235-247.

[40] Mcconohy G, Xu X, Cui T, et al. Mechanical regulation of lithium intrusion probability in garnet solid electrolytes[J]. Nature Energy, 2023, 8(3): 241-250.

[41] 孙富, 刘海君, 倪玲, 等. 锂枝晶的生长理论模型及其抑制方法[J]. 科学通报, 2021, 66(9): 1046-1056.
Sun F, Liu H J, Ni L, et al. Growth mechanisms of Li dendrite and the suppression strategies[J]. Chinese Science Bulletin, 2021, 66(9): 1046-1056(in Chinese).

[42] Barton J L, Bockris J O. The electrolytic growth of dendrites from ionic solutions[J]. Proceedings of the Royal Society of London Series A, Mathematical and Physical Sciences, 1962, 268(1335): 485-505.

[43] Chazalviel J N. Electrochemical aspects of the generation of ramified metallic electrodeposits[J]. Physical Review A, 1990, 42(12): 7355-7367.

[44] Brissot C, Rosso M, Chazalviel J N, et al. In situ study of dendritic growth inlithium/PEO-salt/lithium cells[J]. Electrochimica Acta, 1998, 43(10-11): 1569-1574.

[45] Monroe C, Newman J. The effect of interfacial deformation on electrodeposition kinetics[J]. Journal of the Electrochemical Society, 2004, 151(6): A880.

[46] Monroe C, Newman J. Dendrite growth in lithium/polymer systems[J]. Journal of the Electrochemical Society, 2003, 150(10): A1377.

[47] Akolkar R. Mathematical model of the dendritic growth during lithium electrodeposition[J]. Journal of Power Sources, 2013, 232: 23-28.

[48] Ganser M, Hildebrand F E, Klinsmann M, et al. An extended formulation of Butler-Volmer electrochemical reaction kinetics including the influence of mechanics[J]. Journal of the Electrochemical Society, 2019, 166(4): H167-H176.

[49] Klinsmann M, Hildebrand F E, Ganser M, et al. Dendritic cracking in solid electrolytes driven by lithium insertion[J]. Journal of Power Sources, 2019, 442: 227226.

[50] Yu S, Siegel D J. Grain boundary softening: a potential mechanism for lithium metal penetration through stiff solid electrolytes[J]. ACS Applied Materials & Interfaces, 2018, 10(44): 38151-38158.

[51] 洪清和, 耿晓彬, 徐波, 等. 固态电解质电池的枝晶生长仿真研究[J]. 武汉理工大学学报, 2023, 45(2): 81-87.
Hong Q H, Geng X B, Xu B, et al. Simulation of dendrite growth in solid electrolyte batteries[J]. Journal of Wuhan University of Technology, 2023, 45(2): 81-87(in Chinese) .

[52] Barai P, Higa K, Srinivasan V. Effect of initial state of lithium on the propensity for dendrite formation: a theoretical study[J]. Journal of the Electrochemical Society, 2017, 164(2): A180-A189.

[53] Natsiavas P P, Weinberg K, Rosato D, et al. Effect of prestress on the stability of electrode-electrolyte interfaces during charging in lithium batteries[J]. Journal of the Mechanics and Physics of Solids, 2016, 95: 92-111.

[54] Ahmad Z, Viswanathan V. Stability of electrodeposition at solid-solid interfaces and implications for metal anodes[J]. Physical Review Letters, 2017, 119(5): 056003.

[55] Ding F, Xu W, Graff G L, et al. Dendrite-free lithium deposition via self-healing electrostatic shield mechanism[J]. Journal of the American Chemical Society, 2013, 135(11): 4450-4456.

[56] Schultz R. Lithium: measurement of Young's modulus and yield strength: FERMILAB-TM-2191[R]. Batavia: Fermilab, 2002.

[57] Narayan S, Anand L. A large deformation elastic-viscoplastic model for lithium[J]. Extreme Mechanics Letters, 2018, 24: 21-29.

[58] Yang F Q. Modeling analysis for the growth of a Li sphere and Li whisker in a solid-state lithium metal battery[J]. Physical Chemistry Chemical Physics, 2020, 22(24): 13737-13745.

[59] Mcdowell M T, Xia S M, Zhu T. The mechanics of large-volume-change transformations in high-capacity battery materials[J]. Extreme Mechanics Letters, 2016, 9: 480-494.

[60] Anand L. A Cahn-Hilliard-type theory for species diffusion coupled with large elastic-plastic deformations[J]. Journal of the Mechanics and Physics of Solids, 2012, 60(12): 1983-2002.

[61] Bower A F, Guduru P R, Chason E. Analytical solutions for composition and stress in spherical elastic-plastic lithium-ion electrode particles containing a propagating phase boundary[J]. International Journal of Solids and Structures, 2015, 69-70: 328-342.

[62] Huang S, Fan F, Li J, et al. Stress generation during lithiation of high-capacity electrode particles in lithium ion batteries[J]. Acta Materialia, 2013, 61(12): 4354-4364.

[63] Di Leo C V, Rejovitzky E, Anand L. Diffusion-deformation theory for amorphous silicon anodes: the role of plastic deformation on electrochemical performance[J]. International Journal of Solids and Structures, 2015, 67-68: 283-296.

[64] Wu H, Chan G, Choi J W, et al. Stable cycling of double-walled silicon nanotube battery anodes through solid-electrolyte interphase control[J]. Nature Nanotechnology, 2012, 7(5): 310-315.

[65] Zhang X X, Krischok A, Linder C. A variational framework to model diffusion induced large plastic deformation and phase field fracture during initial two-phase lithiation of silicon electrodes[J]. Computer Methods in Applied Mechanics and Engineering, 2016, 312: 51-77.

[66] Wang Y, Richards W D, Ong S P, et al. Design principles for solid-state lithium superionic conductors[J]. Nature Materials, 2015, 14(10): 1026-1031.

[67] Zhou L D, Park K H, Sun X Q, et al. Solvent-engineered design of argyrodite Li_6PS_5X (X = Cl, Br, I) solid electrolytes with high ionic conductivity[J]. ACS Energy Letters, 2019, 4(1): 265-270.

[68] De Klerk N J J, Rosłoń I, Wagemaker M. Diffusion mechanism of Li argyrodite solid electrolytes for Li-ion batteries and prediction of optimized halogen doping: the effect of Li vacancies, halogens, and halogen disorder[J]. Chemistry of Materials, 2016, 28(21): 7955-7963.

[69] Baktash A, Reid J C, Yuan Q H, et al. Shaping the future of solid-state electrolytes through computational modeling[J]. Advanced Materials, 2020, 32(18): 1908041.

[70] Baktash A, Reid J C, Roman T, et al. Diffusion of lithium ions in lithium-argyrodite solid-state electrolytes[J]. NPJ Computational Materials, 2020, 6: 162.

[71] Yu C, Van Eijck L, Ganapathy S, et al. Synthesis, structure and electrochemical performance of the argyrodite Li_6PS_5Cl solid electrolyte for Li-ion solid state batteries[J]. Electrochimica Acta, 2016, 215: 93-99.

[72] He P P, Tang Y, Tan Z L, et al. Solid-state batteries encounter challenges regarding the interface involving lithium metal[J]. Nano Energy, 2024, 124: 109502.

[73] Lewis J A, Tippens J, Cortes F J Q, et al. Chemo-mechanical challenges in solid-state batteries[J]. Trends in Chemistry, 2019, 1(9): 845-857.

[74] Wenzel S, Randau S, Leichtweiß T, et al. Direct observation of the interfacial instability of the fast ionic conductor $Li_{10}GeP_2S_{12}$ at the lithium metal anode[J]. Chemistry of Materials, 2016, 28(7): 2400-2407.

[75] Han X G, Gong Y H, Fu K, et al. Negating interfacial impedance in garnet-based solid-state Li metal batteries[J]. Nature Materials, 2017, 16(5): 572-579.

[76] Tsai C L, Roddatis V, Chandran C V, et al. $Li_7La_3Zr_2O_{12}$ interface modification for Li dendrite prevention[J]. ACS Applied Materials & Interfaces, 2016, 8(16): 10617-10626.

[77] Rana M, Rudel Y, Heuer P, et al. Toward achieving high areal capacity in silicon-based solid-state battery anodes: what influences the rate-performance[J]. ACS Energy Letters, 2023, 8(7): 3196-3203.

[78] 郝增辉, 刘训良, 孟缘, 等. 电极界面微观结构对固态锂离子电池性能的影响[J]. 储能科学与技术, 2023, 12(7): 2095-2104.
Hao Z H, Liu X L, Meng Y, et al. Effect of electrode interface microstructure on the performance of solid-state lithium-ion battery[J]. Energy Storage Science and Technology, 2023, 12(7): 2095-2104(in Chinese) .

[79] Tian H K, Qi Y. Simulation of the effect of contact area loss in all-solid-state Li-ion batteries[J]. Journal of the Electrochemical Society, 2017, 164(11): E3512-E3521.

[80] Javed B, Koyama M. Multi-physics simulation of solid-state batteries with active material coating[J]. Journal of the Electrochemical Society, 2020, 167(2): 020521.

[81] Xiong R Y, Zhou M Y, Li L H, et al. Decoupled measurement and modeling of interface reaction kinetics of ion-intercalation battery electrodes[J]. Energy Storage Materials, 2023, 54: 836-844.

[82] Wang W, Wang J X, Lin C, et al. Modeling of void-mediated cracking and lithium penetration in all-solid-state batteries[J]. Advanced Functional Materials, 2023, 33(41): 2303484.

[83] Muy S, Bachman J C, Giordano L, et al. Tuning mobility and stability of lithium ion conductors based on lattice dynamics[J]. Energy & Environmental Science, 2018, 11(4): 850-859.

[84] Rodger A R, Kuwano J, West A R. Li^+ ion conducting γ solid solutions in the systems Li_4XO_4-Li_3YO_4: X=Si, Ge, Ti; Y=P, As, V; Li_4XO_4-$LiZO_2$: Z=Al, Ga, Cr and Li_4GeO_4-Li_2CaGeO_4[J]. Solid State Ionics, 1985, 15(3): 185-198.

[85] He X F, Zhu Y Z, Mo Y F. Origin of fast ion diffusion in superionic conductors[J]. Nature Communications, 2017, 8: 15893.

[86] 张田. 固态电池中电极电解质相关模拟计算[D]. 天津: 天津大学, 2022.
Zhang T. Correlated simulation calculation of electrode and electrolyte in solid state batteries[D]. Tianjin: Tianjin University, 2022 (in Chinese).

[87] Ren D S, Lu L G, Hua R, et al. Challenges and opportunities of practical sulfide-based all-solid-state batteries[J]. eTransportation, 2023, 18: 100272.

[88] Tikekar M D, Archer L A, Koch D L. Stabilizing electrodeposition in elastic solid electrolytes containing immobilized anions[J]. Science Advances, 2016, 2(7): e1600320.

[89] Mistry A, Mukherjee P P. Molar volume mismatch: a malefactor for irregular metallic electrodeposition with solid electrolytes[J].

Journal of the Electrochemical Society, 2020, 167(8): 082510.

[90] Ganser M, Hildebrand F E, Mcmeeking R M, et al. Stiffer is not necessarily better: requirements analysis for binary solid polymer electrolytes that ensure stable lithium metal electrodes[J]. Journal of the Electrochemical Society, 2020, 167(13): 130525.

[91] Song S S, He F, Xia Q, et al. Research advances in rare-earth-based solid electrolytes for all-solid-state batteries[J]. Small, 2025, 21(23): 2502008.

[92] Zhu X X, Wu J X, Lu J. Insight into inorganic solid-state electrolytes: ionic transport and failure mechanisms[J]. Advanced Functional Materials, 2024, 34(49): 2409547.

[93] Yang S M, Tang Y, Yao Y, et al. Sulfide electrolytes for all-solid-state sodium batteries: fundamentals and modification strategies[J]. Materials Horizons, 2025, 12(4): 1058-1083.

[94] Ganser M, Hildebrand F E, Kamlah M, et al. A finite strain electro-chemo-mechanical theory for ion transport with application to binary solid electrolytes[J]. Journal of the Mechanics and Physics of Solids, 2019, 125: 681-713.

[95] Newman J, Balsara N P. Electrochemical systems[M]. 4th ed. Hoboken: John Wiley & Sons, 2021.

[96] Bucci G, Chiang Y M, Carter W C. Formulation of the coupled electrochemical-mechanical boundary-value problem, with applications to transport of multiple charged species[J]. Acta Materialia, 2016, 104: 33-51.

[97] Tippens J, Miers J C, Afshar A, et al. Visualizing chemomechanical degradation of a solid-state battery electrolyte[J]. ACS Energy Letters, 2019, 4(6): 1475-1483.

[98] Mücke R, Finsterbusch M, Kaghazchi P, et al. Modelling electrochemical induced stresses in all-solid-state batteries: anisotropy effects in cathodes and cell design optimisation[J]. Journal of Power Sources, 2021, 489: 229430.

[99] Miehe C, Dal H, Schänzel L M, et al. A phase-field model for chemo-mechanical induced fracture in lithium-ion battery electrode particles[J]. International Journal for Numerical Methods in Engineering, 2016, 106(9): 683-711.

[100] Yamakawa S, Nagasako N, Yamasaki H, et al. Phase-field modeling of stress generation in polycrystalline $LiCoO_2$[J]. Solid State Ionics, 2018, 319: 209-217.

[101] Anand L, Mao Y W, Talamini B. On modeling fracture of ferritic steels due to hydrogen embrittlement[J]. Journal of the Mechanics and Physics of Solids, 2019, 122: 280-314.

[102] Shi T, Tu Q, Tian Y, et al. High active material loading in all-solid-state battery electrode via particle size optimization[J]. Advanced Energy Materials, 2020, 10(1): 1902881.

[103] Bucci G, Talamini B, Balakrishna A R, et al. Mechanical instability of electrode-electrolyte interfaces in solid-state batteries[J]. Physical Review Materials, 2018, 2(10): 105407.

[104] Barai P, Rojas T, Narayanan B, et al. Investigation of delamination-induced performance decay at the cathode/LLZO interface[J]. Chemistry of Materials, 2021, 33(14): 5527-5541.

[105] Rezaei S, Asheri A, Xu B X. A consistent framework for chemo-mechanical cohesive fracture and its application in solid-state batteries[J]. Journal of the Mechanics and Physics of Solids, 2021, 157: 104612.

[106] Fathiannasab H, Zhu L K, Chen Z W. Chemo-mechanical modeling of stress evolution in all-solid-state lithium-ion batteries using synchrotron transmission X-ray microscopy tomography[J]. Journal of Power Sources, 2021, 483: 229028.

[107] Wan T H, Ciucci F. Electro-chemo-mechanical modeling of solid-state batteries[J]. Electrochimica Acta, 2020, 331: 135355.

[108] Ke X Y, Wang Y, Ren G F, et al. Towards rational mechanical design of inorganic solid electrolytes for all-solid-state lithium ion batteries[J]. Energy Storage Materials, 2020, 26: 313-324.

[109] Wang Z Q, Wu M S, Liu G, et al. Elastic properties of new solid state electrolyte material $Li_{10}GeP_2S_{12}$: a study from first-principles calculations[J]. International Journal of Electrochemical Science, 2014, 9(2): 562-568.

[110] Yang Y H, Wu Q, Cui Y H, et al. Elastic properties, defect thermodynamics, electrochemical window, phase stability, and Li^+ mobility of Li_3PS_4: insights from first-principles calculations[J]. ACS Applied Materials & Interfaces, 2016, 8(38): 25229-25242.

[111] Zhang X, Wang Q J, Harrison K L, et al. Pressure-driven interface evolution in solid-state lithium metal batteries[J]. Cell Reports Physical Science, 2020, 1(2): 100012.

[112] Tu Q S, Barroso-Luque L, Shi T, et al. Electrodeposition and mechanical stability at lithium-solid electrolyte interface during plating in solid-state batteries[J]. Cell Reports Physical Science, 2020, 1(7): 100106.

[113] Verma A, Kawakami H, Wada H, et al. Microstructure and pressure-driven electrodeposition stability in solid-state batteries[J]. Cell Reports Physical Science, 2021, 2(1): 100301.

[114] Porz L, Swamy T, Sheldon B W, et al. Mechanism of lithium metal penetration through inorganic solid electrolytes[J]. Advanced Energy Materials, 2017, 7(20): 1701003.

[115] Barroso-Luque L, Tu Q S, Ceder G. An analysis of solid-state electrodeposition-induced metal plastic flow and predictions of stress states in solid ionic conductor defects[J]. Journal of the Electrochemical Society, 2020, 167(2): 020534.

[116] Yuan C H, Gao X, Jia Y K, et al. Coupled crack propagation and dendrite growth in solid electrolyte of all-solid-state battery[J]. Nano Energy, 2021, 86: 106057.

[117] Yuan C H, Lu W Q, Xu J. Unlocking the electrochemical-mechanical coupling behaviors of dendrite growth and crack propagation in all-solid-state batteries[J]. Advanced Energy Materials, 2021, 11(36): 2101807.

[118] Shishvan S S, Fleck N A, Mcmeeking R M, et al. Dendrites as climbing dislocations in ceramic electrolytes: Initiation of growth[J]. Journal of Power Sources, 2020, 456: 227989.

[119] Shishvan S S, Fleck N A, Mcmeeking R M, et al. Growth rate of lithium filaments in ceramic electrolytes[J]. Acta Materialia, 2020, 196: 444-455.

[120] 李建伟. 硫基固态电解质的改性及其全固态电池性能研究[D]. 济南: 山东大学, 2023.
Li J W. Modification of sulfide based solid electrolyte and its application in all-solid-state batteries[D]. Jinan: Shandong University, 2023(in Chinese).

[121] Sun F, Wang C, Osenberg M, et al. Clarifying the electro-chemo-mechanical coupling in $Li_{10}SnP_2S_{12}$ based all-solid-state batteries[J]. Advanced Energy Materials, 2022, 12(13): 2103714.

[122] 白洋. 全固态厚膜锂电池的研究[D]. 成都: 电子科技大学, 2022.

Bai Y. All solid-state thick film lithium battery[D]. Chengdu: University of Electronic Science and Technology of China, 2022(in Chinese).

[123] Danilov D, Niessen R A H, Notten P H L. Modeling all-solid-state Li-ion batteries[J]. Journal of the Electrochemical Society, 2011, 158(3): A215-A222.

[124] Zhang D, Tang S X, Couto L D, et al. PDE observer for all-solid-state batteries via an electrochemical model[C]//Proceedings of the IEEE Conference on Control Technology and Applications. Piscataway: IEEE Press, 2022: 51-56.

[125] Kazemi N, Danilov D L, Haverkate L, et al. Modeling of all-solid-state thin-film Li-ion batteries: accuracy improvement[J]. Solid State Ionics, 2019, 334: 111-116.

[126] 李学鹏. 基于多物理场全固态锂电池特性仿真研究[J]. 集成电路应用, 2019, 36(12): 4-6.
Li X P. Study on characteristics simulation of all solid state lithium battery based on multiphysics[J]. Application of IC, 2019, 36(12): 4-6(in Chinese).

[127] Ansah S, Shin N, Lee J S, et al. A comprehensive parametric study for solid-state lithium-ion battery through finite element simulation[J]. Electronic Materials Letters, 2021, 17(6): 532-542.

[128] Wang Q K, Shen J N, He Y J, et al. Design and management of lithium-ion batteries: a perspective from modeling, simulation, and optimization[J]. Chinese Physics B, 2020, 29(6): 068201.

[129] Wang Y, Liu J H, Chen T W, et al. Factors that affect volume change during electrochemical cycling in cathode materials for lithium ion batteries[J]. Physical Chemistry Chemical Physics, 2022, 24(4): 2167-2175.

[130] Kodama M, Horikawa N, Ohashi A, et al. Coupled nonlinear stress and electric field numerical simulation for all-solid-state lithium-ion batteries[J]. Journal of Power Sources Advances, 2021, 8: 100049.

[131] Grazioli D, Verners O, Zadin V, et al. Electrochemical-mechanical modeling of solid polymer electrolytes: impact of mechanical stresses on Li-ion battery performance[J]. Electrochimica Acta, 2019, 296: 1122-1141.

[132] Grazioli D, Zadin V, Brandell D, et al. Electrochemical-mechanical modeling of solid polymer electrolytes: Stress development and non-uniform electric current density in trench geometry microbatteries[J]. Electrochimica Acta, 2019, 296: 1142-1162.

[133] Bistri D, Di Leo C V. Modeling of chemo-mechanical multiparticle interactions in composite electrodes for liquid and solid-state Li-ion batteries[J]. Journal of the Electrochemical Society, 2021, 168(3): 030515.

[134] Song X, Lu Y J, Wang F H, et al. A coupled electro-chemo-mechanical model for all-solid-state thin film Li-ion batteries: the effects of bending on battery performances[J]. Journal of Power Sources, 2020, 452: 227803.

[135] Yamada M, Watanabe T, Gunji T, et al. Review of the design of current collectors for improving the battery performance in lithium-ion and post-lithium-ion batteries[J]. Electrochem, 2020, 1(2): 124-159.

[136] Doyle M, Fuller T F, Newman J. Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell[J]. Journal of the Electrochemical Society, 1993, 140(6): 1526-1533.

[137] Fabre S D, Guy-Bouyssou D, Bouillon P, et al. Charge/discharge simulation of an all-solid-state thin-film battery using a one-dimensional model[J]. Journal of the Electrochemical Society, 2012, 159(2): A104-A115.

[138] Pang M C, Wei Y X, Wang H Z, et al. Large-format bipolar and parallel solid-state lithium-metal cell stacks: a thermally coupled model-based comparative study[J]. Journal of the Electrochemical Society, 2020, 167(16): 160555.

[139] 孙哲韬, 何英杰, 陈邵杰, 等. 全固态锂金属电池多物理场耦合下的电化学过程仿真模拟[J]. 高等学校化学学报, 2021, 42(5): 1598-1609.
Sun Z T, He Y J, Chen S J, et al. Simulation of the electrochemistry process with the coupling of multiple physical fields for all-solid-state lithium batteries[J]. Chemical Journal of Chinese Universities, 2021, 42(5): 1598-1609(in Chinese).

[140] Bower A F, Guduru P R, Sethuraman V A. A finite strain model of stress, diffusion, plastic flow, and electrochemical reactions in a lithium-ion half-cell[J]. Journal of the Mechanics and Physics of Solids, 2011, 59(4): 804-828.

[141] 郭邦军, 贾理男, 张希. 全固态硫化物锂电池中 NCM 正极及其界面研究[J]. 化工学报, 2024, 75(3): 743-759.
Guo B J, Jia L N, Zhang X. A review of NCM cathode and interface characteristics in all-solid-state lithium-ion battery with sulfide electrolytes[J]. CIESC Journal, 2024, 75(3): 743-759 (in Chinese).

[142] Fang X S, He Y L, Fan X M, et al. Modeling and simulation in capacity degradation and control of all-solid-state lithium battery based on time-aging polymer electrolyte[J]. Polymers, 2021, 13(8): 1206.

[143] Bates A, Mukherjee S, Schuppert N, et al. Modeling and simulation of 2D lithium-ion solid state battery[J]. International Journal of Energy Research, 2015, 39(11): 1505-1518.

[144] Talin A A, Ruzmetov D, Kolmakov A, et al. Fabrication, testing, and simulation of all-solid-state three-dimensional Li-ion batteries[J]. ACS Applied Materials & Interfaces, 2016, 8(47): 32385-32391.

Research progress on modeling and simulation of characteristics of all-solid-state batteries

ZHAO Lijin[1, 2]，YANG Shichun[1, *]，LIU Guofang[2]，QU Jingyao[2]，SUN Xudong[2, 3]

(1. School of Transportation Science and Engineering，Beihang University，Beijing 100191，China；
2. China Society of Automotive Engineers，Beijing 100176，China；
3. College of Mechanical and Vehicle Engineering，Chongqing University，Chongqing 400044，China)

Abstract：Governments and large automakers worldwide are aggressively investing in and developing all-solid-state batteries (ASSBs) as a next-generation battery technology because of their promise to provide high specific energy and improved safety. However, significant challenges remain in terms of research and development as well as large-scale industrialization. Conducting modeling and simulation studies on ASSBs can help reveal their underlying operational mechanisms and clarify the factors influencing their performance. This, in turn, enables the design of ASSBs that meet targeted specifications and achieve desired operational performance, ultimately shortening the cycles of experimental testing and development. This paper provides a comprehensive review and analysis of modeling and simulation efforts in the areas of molecular dynamics calculations, key materials and interfacial characteristics, and electrochemical performance of ASSBs. Particular emphasis is placed on the microstructural modeling of sulfide-based solid electrolytes, interfacial behavior, ion transport mechanisms, stress generation, and fracture induced by the interphase between electrodes and electrolytes. The study also identifies possible research goals, describes future trends in ASSB modeling and simulation, and attempts to use these insights to hasten the industrialization and practical application of ASSBs.

Keywords：all-solid-state battery；model building；simulation analysis；multi physics field；solid electrolyte

Received：2025-05-27；**Accepted**：2025-09-25；**Published Online**：2025-10-15 17：00
URL：link.cnki.net/urlid/11.2625.V.20251015.1343.001
Foundation item：National Key Research and Development Program of China (2023YFB2408200)
* **Corresponding author**. E-mail：yangshichun@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0901

热障涂层的智能化设计研究进展

何健[1,2]，李宏峰[1,☆]，康鹏[1,3]，郑蕾[1,3]，李圳[1]，郭洪波[1,2,3,*]

(1. 北京航空航天大学 材料科学与工程学院，北京 100191；2. 高温结构材料与涂层技术工信部重点实验室，北京 100191；
3. 北京航空航天大学 天目山实验室，杭州 311115)

摘 要：随着航空发动机推重比与涡轮前进口温度的持续增高，热障涂层(TBC)已成为突破热端部件耐温极限、保障其长时服役可靠性的核心手段，但其服役行为涉及极度复杂的热-力-化(TMC)多场耦合与跨尺度损伤演化，使得传统单一尺度的实验与解决方法面临严峻挑战。基于此，系统梳理了TBC领域传统与先进建模、仿真及分析技术的研究进展，并拓展了人工智能(AI)技术的相关应用，旨在系统总结TBC服役行为的解析方法、涂层结构优化策略，并推动前瞻性维护策略的落地。涉及的技术方法包括TBC体系的热导率建模、热应力仿真、裂纹扩展分析及力学行为研究等，这些方法多依赖于多尺度建模和有限元分析等多样化手段。重点阐述了人工智能在TBC仿真中的应用，深入探究了支持向量机、随机森林、神经网络与深度学习等机器学习算法，以及数据驱动建模技术。研究结果表明，人工智能在多个方面具备革新TBC技术的潜力，涵盖材料与结构设计、性能优化与预测、状态监测、故障诊断与健康管理、剩余使用寿命(RUL)预测等方向，并能显著提升预测精度。随着智能化新材料设计技术的不断突破，新型TBC必将为下一代航空发动机突破耐温极限、实现更高效率与可靠性提供有力支撑。

关 键 词：热障涂层；智能化设计；结构设计；使用性能；机器学习
中图分类号：V23; TG174.4
文献标志码：A **文章编号**：1001-5965(2026)10-3298-15

新一代喷气发动机和燃气轮机是飞机、发电机及大型船舶的核心动力装置，其高温区叶片由镍基高温合金制成，需在超过1 200℃的极端环境下长期服役。为保护高温合金，普遍采用热障涂层(thermal barrier coating, TBC)系统与空气冷却技术[1-3]。TBC系统沉积于高温合金表面，温度超过1 300℃时，氧化钇稳定氧化锆(yttria-stabilized zirconia, YSZ)的相稳定性与抗损伤能力下降。新一代发动机的目标工作温度往往突破1 500℃，现阶段很难满足新一代发动机需求，部分核心部件均面临重大技术升级需求。因此，亟须开发高温稳定性的新型TBC材料[4-6]。

人工智能(artificial intelligence, AI)，尤其是机器学习与数据驱动建模方法，为从海量实验与模拟数据中挖掘关键信息、加速TBC研发提供了新途径。AI不仅可构建高精度预测模型，还能辅助多尺度结构优化，提升TBC隔热、力学与耐久性能。然而，数据质量、模型可解释性等问题仍构成技术瓶颈，制约着材料、结构、环境这3个系统耦合等。尽管如此，AI与TBC研究的深度融合前景广阔，有望为材料创新设计带来范式突破。

本文旨在系统梳理TBC建模、仿真与有限元分析的最新进展，涵盖传统与先进方法，并结合AI技术进行拓展分析。在此基础上，深入探讨AI在数据挖掘、成分优化与性能预测中的应用潜

收稿日期：2025-12-30；**录用日期**：2026-05-08；**网络出版时间**：2026-05-25 11:11
网络出版地址：link.cnki.net/urlid/11.2625.V.20260523.1228.001
基金项目：国家自然科学基金(U21B2052)
* **通信作者**. E-mail：guo.hongbo@buaa.edu.cn
☆ **共同一作**. E-mail：lihongfeng@buaa.edu.cn

引用格式：*何健，李宏峰，康鹏，等. 热障涂层的智能化设计研究进展[J]. 北京航空航天大学学报，2026，52(10)：3298-3312.*
He J，Li H F，Kang P，et al. Research progress on intelligent design of thermal barrier coatings[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3298-3312 (in Chinese).

力, 通过分析现有研究明确未来研发方向, 为新型 TBC 技术的发展提供理论支撑, 助力其在高温极端工况下实现更高性能与可靠性。

1　AI 在 TBC 研究上的应用

1.1　TBC 应用概述

TBC 系统沉积于高温合金表面, 晶格热导率通常低于 2W/(m·K), 如图 1 所示, 主要作用是为航空发动机及燃气轮机高温结构件提供热屏蔽层, 以提高其耐受温度和服役寿命。TBC 包括金属黏结层、因服役过程中金属黏结层氧化而形成的热生长氧化物(thermally grown oxide, TGO)层及陶瓷层。由于高温服役条件下 ZrO_2 会发生四方相转变为单斜相的相变, 通常造成约 3%～5% 的体积膨胀, 在 ZrO_2 中添加钇可有效抑制, 并且掺杂钇会在 TBC 材料中引入氧空位, 进一步降低其热导率。因此, 目前使用最多的 TBC 材料为 YSZ。由于立方相 YSZ 比四方相 YSZ 具有更大的晶胞体积, 使得氧离子更容易在其中迁移, 电导率也更高。在立方相的 YSZ 中, 8YSZ 具有最高的电导率[7]。由于新一代喷气式和燃气轮机的工作温度往往超过 1 500℃, YSZ 在其中已不再适用。当温度超过 1 300℃ 时, YSZ 的相稳定性下降, 抗损伤能力降低, 而且工作温度提高需要更低的热导率。因此, 亟须开发具有更优异高温稳定性和更低热导率的新型 TBC 材料。目前, 全球正广泛寻找能取代 YSZ 的新型高性能 TBC 候选材料[8-12]。

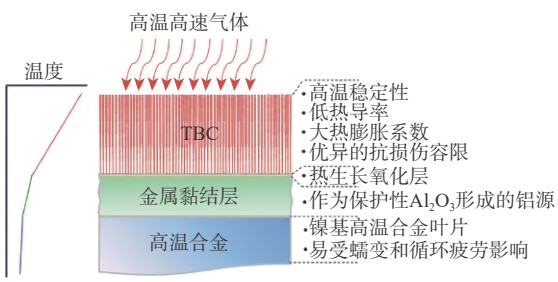


图 1　TBC 系统多层结构示意图[12]

Fig. 1　Schematic of multilayer of TBC system[12]

新型 TBC 材料的选择需满足两大核心要求: ①作为核心隔热层, 需在极端苛刻的热机械环境中服役, 必须具备高熔点、良好的热循环稳定性, 即在热循环过程中不会发生显著相变, 且具备优异的损伤耐受性、出色的抗高温氧化与腐蚀能力, 以及与金属基底相匹配的热膨胀系数; ②TBC 在实际服役过程中, 性能的准确评估仍是重大挑战, 因为其工作环境通常极端恶劣, 包括高温、高压、大温度梯度及强热机械和腐蚀载荷, 而这些复杂工况在实验室条件下难以完全复现。此外, 当前大多数研究仍主要集中在识别一种或几种具有较低热导率的复合氧化物上, 但这些材料的晶体结构对 TBC 应用的适配性尚未明确, 其低热导率的结构与化学特性也尚未被完全揭示, 相关热导率理论预测仍不清晰, 限制了 TBC 在不同行业工程中的应用[13], 如图 2 所示。

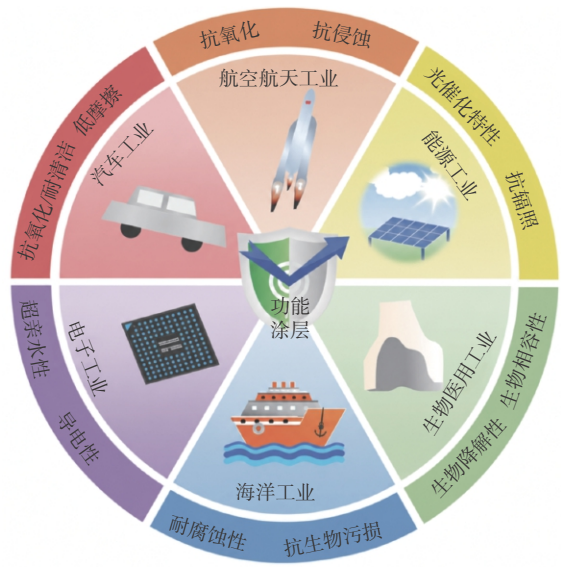


图 2　TBC 在多种工业应用中的特定保护或功能需求[13]

Fig. 2　Specific protection or function needs of TBC in versatile industrial applications[13]

1.2　AI 赋能新型 TBC 材料设计

如果能将 AI(如机器学习与数据驱动建模等)融入新型 TBC 材料的研究, 有望在涵盖材料性能、工艺参数及环境条件的海量数据中挖掘出具有重要价值的信息[14-16]。这种融合不仅有助于构建复杂、高精度的预测模型, 实现对不同工况下 TBC 行为的准确预判, 还能助力探索多尺度、多维度的结构与涂层构型, 从而加速研发兼具更优隔热性能、力学强度与长耐久性的理想 TBC。然而, 将 AI 技术应用于 TBC 研究仍面临诸多技术和方法上的挑战, 包括数据质量的可靠性、模型可解释性及计算复杂度等问题。这些挑战在一定程度上增加了对材料性能、微观结构特征、环境因素之间复杂相互作用机制进行精准捕捉的难度, 而这种多因素耦合的相互作用对 TBC 的性能表现及耐久性具有重要影响。尽管存在上述技术瓶颈, TBC 研究与 AI 技术的融合仍展现出广阔的发展前景, 为新型 TBC 材料的创新设计提供了新的思路和方法, 有助于开发在高温应用场景中具备更优隔热性能、耐久性和可靠性的新型 TBC 材料[17-19]。

在 TBC 研究中, AI 技术的应用主要体现在 2 种核心方法上: ①采用支持向量机(support vector machine, SVM)、随机森林(random forest, RF)等机

器学习算法，对大数据集进行分析，识别材料成分、微观结构特征与环境条件之间的潜在关联与规律，为设计更耐久的 TBC 材料提供科学依据。②借助循环神经网络（recurrent neural network, RNN）和长短期记忆（long short term memory, LSTM）网络等 AI 预测建模技术，用于预测不同使用条件下 TBC 的退化行为，进而制定前瞻性维护与失效缓解策略，防止涂层过早失效，提高服役寿命。在这一预测过程中，监督学习与无监督学习等机器学习技术发挥着关键作用[20-23]。SVM、决策树、神经网络等监督学习算法通常依赖于带标签的数据集，通过输入材料成分、微观结构特征及环境条件等参数，构建具有预测能力的模型，从而实现对 TBC 性能的模拟与预测。聚类算法等无监督学习算法则可在缺乏标签数据的情况下，对 TBC 数据集进行模式识别与关联分析，挖掘潜在信息并辅助构建预测模型。此外，基于概率图模型和卷积神经网络（convolutional neural networks, CNN）、RNN 等深度学习架构，能够有效捕捉 TBC 数据集中复杂的非线性关系与时间依赖性，进一步提升对材料行为预测的准确性与鲁棒性。粒子群优化（particle swarm optimization, PSO）和遗传算法（genetic algorithm, GA）等进化算法通过模拟自然选择过程，采用迭代优化方法对 TBC 材料成分进行系统性探索与选择。此类算法能够从大量可能的材料组合中筛选出具备理想性能特征的最优组合，如低导热性、优力学强度及化学稳定性，从而为设计具有优异综合性能的 TBC 提供科学依据[24-26]。

AI 驱动的预测分析技术能够深入解析微观结构特征、环境因素与性能指标之间的复杂相互作用，为针对特定应用需求定制的 TBC 系统设计提供指导。此外，基于 AI 的数字孪生技术和预测性维护算法，可实现对 TBC 状态的实时监测，为制定前瞻性维护策略提供高效、智能化的解决方案，减少非计划停机时间，提升系统的整体可用性与运行效率。前瞻性维护策略的实施可借助 AI 驱动的预测模型进一步优化。这些模型能够有效识别 TBC 在不同工况下的退化机制，支持在早期阶段采取干预措施，以降低失效风险并延长关键部件的使用寿命。同时，AI 赋能的诊断技术可通过分析传感器数据检测退化或损伤迹象，提升故障检测能力，保障涂覆 TBC 部件的可靠性与安全性[27-30]。

2 TBC 材料与结构的设计及应用

2.1 TBC 热物理性能预测

TBC 的热导率、热膨胀系数、定压比热容等核心热物理、热力学性能是涂层隔热效率、结构兼容性与高温稳定性的关键表征指标。其中，热导率直接决定 TBC 的隔热能力；热膨胀系数需与金属基体形成温域适配的梯度匹配，通过调控涂层成分与微观结构，缩小涂层与基体在全服役温度区间的热膨胀系数差异，缓解热循环过程中因热胀冷缩不同步产生的界面剪切应力与法向应力，从根源上抑制界面裂纹的萌生与扩展；而定压比、热容等热力学参数则影响涂层在极端温度波动下的服役寿命。上述性能的精准量化，是涂层配方设计、性能优化及工程应用的核心前提[31-33]。传统性能实验测试需制备多组样品、经历复杂高温环境模拟，耗时耗力且成本高昂，物理建模难以全面捕捉成分掺杂、微观结构演化等多因素耦合的非线性影响，泛化能力不足，难以适配新型 TBC 涂层的性能预测需求。机器学习等 AI 算法通过整合材料成分、微观结构特征、制备工艺参数等多维度数据，构建支持向量回归（support vector regression, SVR）、神经网络等模型，可实现性能与影响因素间非线性构效关系的精准映射，提升热物理、热力学性能预测的精度与效率，为新型高性能 TBC 的快速筛选与配方优化提供了智能化工具，加速了 TBC 技术的研发迭代进程[34-35]。

Zhu 等[36]通过 GA-SVR 算法预测 TBC 材料 $Dy_{1-x}Ta_{1-x}Zr_{2x}O_4$ 的热导率和热膨胀系数等热物理性质（见图 3），具有较高的效率，结果表明，该模型可实现对 TBC 热导率、热膨胀系数的高精度预测，2 项参数的预测值与实测值的决定系数 R^2 分别高达 0.998 和 0.999，均方误差（mean square error, MSE）

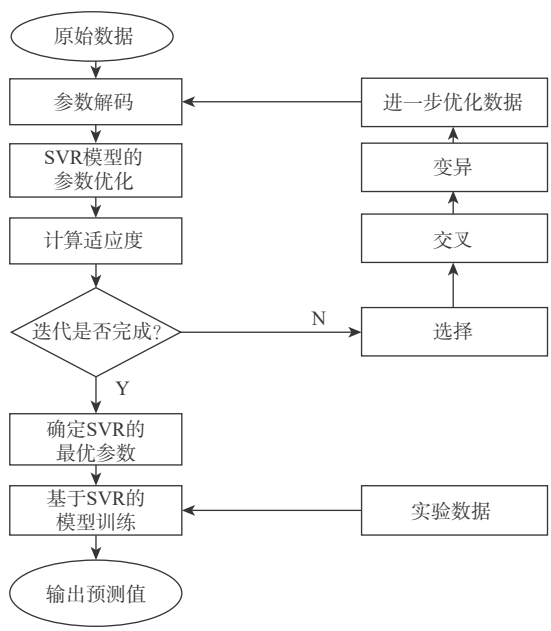


图 3 基于 GA 优化 SVR 的流程[36]

Fig. 3 Flow chart of GA optimization SVR[36]

仅为 0.000 52 和0.000 19, 远优于同类预测模型, 可为涂层热物理性能的正向设计提供精准的理论支撑。100～800℃ 时 $DyTaO_4$ 陶瓷热导率最低的 ZrO_2 掺杂浓度为 0.085～0.095, 200～1 200℃ 时 $DyTaO_4$ 陶瓷热膨胀系数最高的 ZrO_2 掺杂浓度为 0.11～0.13, 能够进一步降低其热导率并增加热膨胀系数。Chen 等[37]通过机器学习和高斯过程回归(Gaussian process regression, GPR)结合的方法预测 TBC 材料的晶格热导率, 先整理近 100 种无机化合物热导率实验值的基准数据集, 使用 Matminer 软件包生成并优化通用特征, 并在之前使用过的机器学习模型上用 GPR 模型训练数据, 得到的模型比传统的物理计算方法更快, 能够实现对无机材料热导率的快速准确预测, 训练后的模型不仅有助于合理设计和筛选新型材料, 还揭示了 TBC 材料中热传输行为的关键特征。

2.2　TBC 结构参数无损检测

TBC 的厚度、孔隙率及 TGO 厚度等结构参数直接决定隔热性能、力学稳定性与服役寿命, 是涂层设计优化、质量管控及服役状态评估的核心指标。传统无损检测方法(如热成像、超声检测)虽具备非破坏性优势, 但普遍存在人工依赖度高、复杂微观结构下特征提取困难、多参数耦合干扰导致精度不足、检测效率低等局限, 难以适配现代 TBC 技术对高分辨率、快速响应、智能化检测的需求。借助 AI 技术优化传统无损检测方法, 是突破上述局限、加速高性能新型 TBC 材料研发的有效途径。通过机器学习算法与深度学习架构的融合, 可实现检测信号的自动化特征提取、冗余信息剔除与关键参数精准映射, 同时通过 GA、PSO 算法等优化模型拓扑结构与初始权重, 提升模型收敛效率与泛化能力[38-40]。

2.2.1　厚度检测

随着化石燃料成本的上升, 热效率已成为新一代航空喷气发动机、燃气轮机热端部件设计与制造的必要考量。TBC 的厚度与均匀性直接决定设备隔热效果与热效率, 根据服役需求, 航空与能源领域热端部件所用 TBC 的典型厚度范围为 100～500 μm, 其厚度精准检测是保障涂层服役性能的核心前提。当前, TBC 厚度主流无损检测技术以脉冲与锁定红外热成像、太赫兹时域光谱(Terahertz time-domain spectroscopy, THz-TDS)为核心, 2 类技术的厚度测量本质上均为特征信号的处理与反演问题, 检测精度高度依赖信号处理技术的水平, 而传统解析方法存在精度低、对非规则厚度适应性差的局限, AI 技术为该瓶颈的突破提供了新路径。Halloua 等[38]围绕 TBC 厚度对热效率的关键影响, 提出结合神经网络与 GA 的混合神经算法, 利用脉冲与锁定红外热成像数据评估非规则 TBC 厚度, 该算法以热响应经傅里叶变换得到的相位特征训练神经网络, 通过 GA 搜索最优初始权重矩阵, 促使模型快速收敛, 结果显示, 该混合模型相比传统神经网络减少了迭代次数、提升了反演精度, 2 种热成像方法的厚度评估偏差均低于 5%, 为 TBC 厚度的高效无损检测提供了可靠方案, 也印证了基于 AI 的信号处理技术对涂层厚度精准测量的关键支撑作用, 为同类涂层参数的精准评估提供了算法参考。Halloua 等[39]利用该混合神经算法结合的方法, 通过激光脉冲热成像数据建立非破坏性且精准测定模型, 进行 TBC 厚度评价应用, 通过主成分分析(principal component analysis, PCA)对激光脉冲热成像采集的热响应数据进行预处理, 筛选出 5 个关键主成分以简化输入维度、剔除冗余信息, 根据 GA 优化神经网络的初始权重参数, 相较于标准神经网络模型, 有效提升其学习效率, 迭代次数显著减少、均方误差明显降低, 该混合方法对薄 TBC 的厚度评估偏差小于 3%, 不仅实现了非破坏性检测的技术目标, 其测量精度更优于 Zhang 等[40]的同类研究结论, 为 TBC 的质量管控提供了一种高效、精准的技术手段。

THz-TDS 技术是一种基于太赫兹波与物质相互作用的先进检测手段, 其独特的频段特性使其成为 TBC 智能化设计研究的新兴工具。相较于传统表征技术, THz-TDS 具有非破坏性、高相位灵敏度和宽频带覆盖的优势, 可反复检测涂层样品而不损伤结构, 适配设计中多轮参数验证的需求, 并且能精准表征 TBC 厚度、孔隙率、TGO 厚度等 TBC 关键设计参数, 结合机器学习等可实现复杂微观结构下的高效定量检测, 还具有检测效率高、适配多尺度结构, 支持设计中的实时参数反馈与快速迭代, 为涂层结构与性能的智能化优化提供了可靠的无损技术支撑, 具有从原子尺度到宏观性能的全链条分析能力。由于 TBC 厚度测量需确定太赫兹信号峰, 为减少人工干预、提升智能化测量能力、提高效率与精度, Sun 等[41]通过由太赫兹信号理论模型驱动的 LSTM 网络框架来定位信号峰, 该框架先自动定位, 再通过局部极值消除定位误差, 同时利用最大面积提取信号峰以计算在线折射率、消除变化对测量的影响, 经验证, 该方法与人工干预结果虽一致, 但效率大幅提升, 如 200 μm 厚样品测量仅需约 0.17 s, 而传统人工法需 140 s, 面对折射率更小的新涂层无需制备新标样, 仅需基于理论模型生成新

数据集，重新训练 LSTM 即可精准定位信号峰。同时，Sun 等[42]还通过新型双流太赫兹残差网络（Terahertz residual network, THz-ResNet）来实现制造阶段 TBC 的在线精准测量，如图 4 所示，前 2 个太赫兹波受散射影响更小，并且含足够的测量信息，设计双流结构来实现厚度测量，第 1 流通过分割层和倒数层提取折射率，第 2 流结合飞行时间生成涂层厚度，使网络操作符合物理规律，精度与效率较优，51 次厚度测量仅需 0.09 s。Yuan 等[43]针对 TBC 在高温环境下易减薄、剥落而需精准检测厚度的需求，提出 THz-TDS 与 PCA 法、GA 与极限学习机（extreme learning machine, ELM）相结合的混合算法的无损评估方法，该方法以实验室制备样本与实际涡轮叶片样本的太赫兹信号为基础，先通过 PCA 将信号维度从 80×4 096 降至 80×28，保留 99.9% 的信息以提升训练效率，再用 GA 优化 ELM 构建厚度预测模型，通过决定系数 R^2、平均绝对百分比误差（mean absolute percentage error, MAPE）、均方根误差（root mean square error, RMSE）及 8 折交叉验证评估模型性能，结果显示，PCA-GA-ELM 模型相关系数超过 0.98、误差小于 2.57，精度与泛化性均优于 PCA-ELM 模型，可满足实际检测的精度要求，为 TBC 厚度的高效无损评估提供了可靠技术路径。

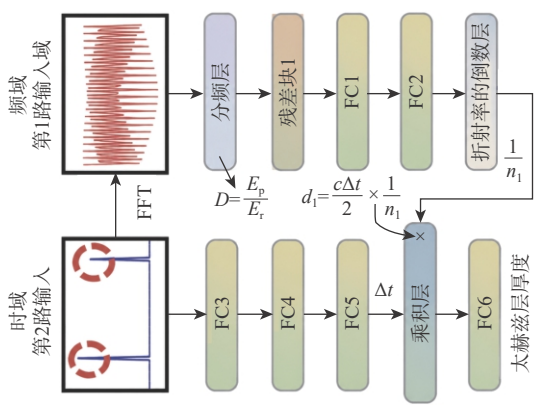


图 4 新型双流太赫兹残差网络框架[42]

Fig. 4 Framework of dual-stream THz-ResNet[42]

2.2.2 孔隙率检测

孔隙率是 TBC 的核心结构参数，其含量、形貌及空间分布直接决定涂层的隔热性能、力学稳定性与服役寿命。根据航空发动机、燃气轮机热端部件的服役需求，大气等离子喷涂（atomsphere plasma spray, APS）TBC 的最优孔隙率区间为 8%～15%，电子束物理气相沉积（electron beam-physical vapor deposition, EB-PVD）柱状晶 TBC 的最优孔隙率区间为 3%～8%，孔隙率的微小波动会引发涂层核心性能的显著变化。适量孔隙可通过增强声子散射降低涂层热导率，实验数据表明，TBC 孔隙率每提升 1%，1 200℃ 高温热导率可降低 0.05～0.09 W/(m·K)，15% 孔隙率可使涂层热导率较致密态降低约 50%，是提升设备热效率的重要手段；但孔隙率过高会严重劣化涂层力学性能，当孔隙率从 12% 升至 20% 时，涂层界面结合强度降幅可超 50%，1 100℃ 热循环服役寿命从约 1 200 次降至不足 300 次，孔隙率超过 25% 时涂层可能出现结构失效。同时，孔隙的形貌、尺寸与分布均匀性，同样会对涂层隔热效果与服役稳定性产生显著影响。因此，孔隙率的精准定量表征是 TBC 配方设计、生产质量管控及服役状态评估的关键环节，对检测技术提出了明确要求：需实现孔隙率体积分数的高精度定量检测，测量误差需控制在±0.5% 以内，同时需完成孔隙形貌、尺寸分布与空间均匀性的全维度表征，为涂层性能优化与全生命周期管控提供可靠的数据支撑[44-46]。

APS ZrO_2 基 TBC 的孔隙结构，可按形貌、尺寸与连通性分为球形闭孔、不规则连通孔、层间扁平裂纹状孔隙与界面孔隙 4 类，其孔隙率与孔隙特征直接决定涂层的隔热性能与服役稳定性。现有超声检测虽为 TBC 孔隙率原位无损检测的主流方法，但针对工程上厚度小于 300 μm 的超薄 TBC，存在回波信号时域混叠、复杂孔隙形貌导致特征提取困难、传统反演模型精度不足的核心瓶颈。针对这一问题，Zhang 等[47]提出了一种融合超声检测、时域数值模拟与 GA 优化反向传播神经网络（back propagation neural network, BPNN）的混合无损检测方法，实现了超薄 TBC 孔隙率的高精度定量表征，该研究以厚度 288 μm、孔隙率 5.71%～26.59% 的等离子喷涂 ZrO_2 基 TBC 为研究对象，通过时域数值模拟构建超声信号正向传播模型，提取与孔隙率强相关的超声反射系数振幅谱（ultrasonic reflection coefficient amplitude spectrum, URCAS）特征参数作为模型输入，结合扫描电子显微镜（scanning electron microscopy, SEM）原位表征获取的 132 组实测数据集完成模型训练与验证，并通过 GA 优化神经网络的初始权重与拓扑结构，有效提升了模型的收敛效率与孔隙率反演精度，该方法突破了传统超声检测针对超薄 TBC 孔隙率表征的技术瓶颈，为 TBC 的原位质量管控提供了新的技术路径，但仍存在对异质样品泛化能力不足、无法实现孔隙形貌全维度表征的局限，最终对 TBC 孔隙率的预测平均相对误差为 1.45%，最大相对误差不超过 6.2%，实现了具有复杂孔隙形貌的 TBC 孔隙率的自动化检测与高精度定量表征。GA-BPNN 为多孔材料的孔隙率表征提供了高效可靠的技术手段，也为 TBC 孔隙率

的智能化表征模型构建提供了可行的技术路径。Ma 等[48] 提出了基于 BPNN-GPR 的超声表征方法，通过 BPNN 结合 R^2>0.8 准则优化 URCAS 的特征参数，选取与孔隙率高度相关的特征参数 P_6、P_7、P_8、P_{10}、P_{12} 训练 GPR 以此来预测孔隙率，经等离子喷涂厚度 300 μm、孔隙率 1%～9% 的 ZrO_2 基 TBC 涂层的随机孔隙模型有限元仿真验证，该方法预测相对误差为 1.07%～7.62%，相比于单独使用 BPNN 精度提升 32%、单独使用 GPR 精度提升 48%，对复杂孔隙形貌的 TBC 孔隙率具备高适应性与高精准度，为孔隙率的高效精准表征提供了可靠技术路径。针对 TBC 孔隙率现有评估时破坏性强且精度不足的问题，Li 等[49] 提出太赫兹多特征融合结合机器学习的无损评估方案，先从 THz-TDS 数据中提取多特征参数，经相关性分析与 p 值检验筛选后完成特征融合，再采用蜣螂优化算法（dung beetle optimization, DBP）优化 RF 模型，结合 K 折交叉验证评估性能，实验结果显示，DBO-RF 模型的孔隙率预测 RMSE 为 1.802、平均绝对误差为 1.549、MAPE 为 8.362，决定系数达 0.912，预测精度较高。该方法可在一定程度上表征 TBC 孔隙率与孔结构特征，在非破坏性前提下实现孔隙率的评估，既能够有效监测 TBC 服役状态，也为航空材料的无损性能评估提供了可靠技术路径，推动了航空材料研究的发展。

2.2.3　TGO 厚度反演

TGO 是 TBC 服役过程中，黏结层高温选择性氧化原位形成的界面层，其厚度演化、微观结构与界面缺陷直接决定 TBC 系统的结构稳定性与服役寿命[50]。理想状态下，致密连续的 α-Al_2O_3 相 TGO 可有效阻挡氧扩散，是 TBC 的核心抗氧化屏障；而 TGO 过度生长会引发显著热膨胀失配应力，同时伴随亚稳相析出、界面凹凸度加剧、内部孔洞与界面楔形裂纹萌生等有害结构演化，大幅提升界面剥离与涂层剥落风险；若 TGO 无法实现连续致密、仿形均匀地完整覆盖黏结层表面，则难以形成有效的抗氧化扩散屏障，无法充分发挥对黏结层的高温防护作用[51]。现有研究已明确，TGO 的无损检测绝非单一的厚度测量，其核心检测目标包括 4 类[52]：①TGO 的平均厚度与厚度分布均匀性，YSZ-TBC 的安全服役临界厚度为 10 μm，界面凹凸幅值超过 2 μm 即会引发严重应力集中；②TGO 的相组成，需明确 α-Al_2O_3 相纯度与亚稳相、尖晶石相的占比，这类杂相会使 TGO 生长速率提升一个数量级以上；③TGO 内部的孔洞、微裂纹等缺陷，孔隙率超过 3% 会使涂层热循环寿命降低 40% 以上；④TGO 与陶瓷层、黏结层的界面脱黏、楔形裂纹等早期失效特征，这类缺陷是涂层大面积剥落的直接前兆。因此，上述核心指标的精准反演与全维度表征，是实现 TBC 失效预警、寿命评估及服役安全性保障的核心环节[53]。

TGO 是决定 TBC 服役寿命的核心结构，现有研究已明确，EB-PVD 与 APS 制备 YSZ-TBC 的 TGO 临界失效厚度约为 10 μm，超过该阈值后 TBC 进入快速失效阶段，20～30 μm 厚度的 TGO 已达到涂层失效判据，该区间的检测无实际工程价值，10 μm 以内薄 TGO 层的高精度无损检测才是工程应用的核心目标[54]。围绕 TGO 厚度的无损检测，现有技术已完成从 30 μm 以上定性检测到 20 μm 区间定量反演的技术跨越，近年来更是通过算法优化实现了 10～15 μm 区间 TGO 的初步检测[55]。但传统 TGO 厚度无损检测方法，面对 10 μm 以内的薄 TGO 层，仍存在核心技术瓶颈：检测信号易受陶瓷层基体与微观结构的强干扰，薄 TGO 的特征信号辨识度低、有效特征提取困难，最终导致厚度反演精度不足、检测效率低下，难以满足工程化原位实时监测的需求。针对该主要问题，未来需从高分辨率检测原理开发、多物理场正向模型构建、AI 智能信号处理 3 个方向突破，实现 10 μm 以内薄 TGO 的高精度、高效率检测，为 TBC 寿命预测提供核心支撑[56]。Luo 等[57] 针对 THz-TDS 检测时分辨率不足的问题，提出稳定小波转换（stationary wavelet transform, SWT）优化 BPNN 结合太赫兹技术的检测方案，如图 5 所示，先通过 COMSOL Multiphysics 工具模拟太赫兹信号反射并求解涂层厚度，采用 SWT 处理信号以提升精度，通过参数化扫描获取 66 组不同厚度的太赫兹脉冲成像（Terahertz pulse imaging, TPI）数据训练网络，用于预测 15 组 1～29 μm 的 TGO 厚度，结果显示，SWT 细化信号的回归系数达 0.92，与真实结果高度吻合，实现了 TGO 的超分辨厚度检测，为 TBC 各层厚度检测提供了关键技术支持，但仍需补充更多的实验数据以进一步优化算法的工程实用性。Li 等[58] 提出 THz-TDS 结合多尺度分析的方法，先通过时域有限差分法（finite difference time domain, FDTD）仿真获取 30 组不同 TGO 厚度的太

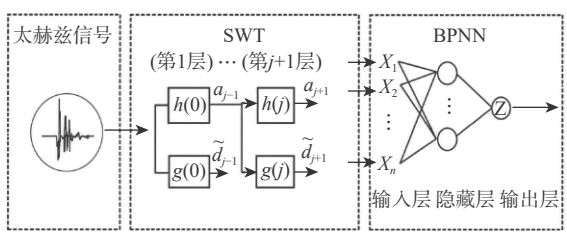


图 5　SWT-BPNN 的基本原理[57]

Fig. 5　Basic principle of SWT-BPNN[57]

赫兹光谱数据，基于 db4、sym3、haar、coif3 的 4 种小波基函数，经连续小波变换与最大池化提取多尺度特征参数，再以部分尺度数据开展多元线性回归、全尺度数据开展机器学习回归，实现 TGO 厚度反演，最终无量纲化特征参数随 TGO 厚度先升后降，随尺度变化差异显著；其中，haar 小波在 2 种模型中的回归系数分别达 0.976 3 与 0.984 0，预测精度最优。该方法证实了太赫兹多尺度分析在 TBC 评价中的应用潜力，虽然样本量较小需进一步验证，但为 TGO 厚度反演提供了可靠的技术方案与理论基础，可支撑 TBC 无损检测与科学研究。

2.3 TBC 缺陷、微观结构及服役应变表征

TBC 的核心服役性能由界面 TGO 的厚度与均匀性主导，在 TGO 仿形均匀、致密连续的前提下厚度越薄越有利；同时，陶瓷层的微观结构、内部裂纹及服役应变等复杂特征，也会协同决定涂层的力学稳定性、隔热耐久性与失效模式。微观结构不均会诱发应力集中，内部裂纹的萌生与扩展是涂层剥落的核心诱因，而服役应变的异常分布则预示着早期失效风险，因此，对这些特征的精准识别与预测是 TBC 设计优化、服役状态监测及寿命评估的关键。人工识别主观性强、效率低，难以实现微观结构的像素级量化，物理仿真难以全面耦合多因素，如温度、载荷、微观缺陷多种非线性影响，对裂纹扩展、应变演化的预测精度不足，且计算成本高昂。AI 对 TBC 复杂特征的高精度识别与预测，不仅摆脱了传统方法的主观性与低效率，更实现了从“事后检测”向“事前预警”的转变，为 TBC 全生命周期管理、失效机制解析及维护策略优化提供了智能化工具，有力保障其在航空航天、能源动力等高温领域的可靠服役。

Ning 等[59]提出弹性力学约束生成对抗网络（elasticity-mechanics-informed generative adversarial network, EMI-GAN）用于进行 CaO-MgO-Al_2O_3-SiO_2（CMAS）渗透和腐蚀下 TBC 的热应变预测，先以 pix2pix 网络为基础构建 EMI-GAN 模型，将变形协调条件、考虑热膨胀的胡克定律作为约束嵌入生成器，同时将约束损失、结构及应变分布输入判别器，基于单张高分辨率 SEM 结构图像的有限元结果构建训练数据库，以具有不同 CMAS 渗透时长与微观结构特征的试样 SEM 作为验证数据库，利用同一数据库分别训练 EMI-GAN 与常规 pix2pix 网络，对比两者生成的热应变结果，最终 EMI-GAN 的热应变预测精度比 pix2pix 网络提升约 10%，即使在训练数据较少的规模下仍高出 7%，其生成的应变分布更符合物理规律且噪声更低，应变集中位置与有限元结果一致，同时该模型在不同微观结构、温度范围的样本中泛化性良好，可扩展至多相结构，且小数据训练即可完成，大幅降低了真实数据库的构建成本。Shi 等[60]针对涡轮叶片 TBC 内部裂纹现有检测方法泛化能力差、精度低的问题，提出基于多尺度增强型快速区域卷积神经网络（multi scale enhanced-faster region-based convolutional neural network, MSE-Faster R-CNN）的检测方案，先以 VGG-19 为基准网络筛选裂纹候选区域并扩展以获取上下文信息，再通过新型 MSE-Faster R-CNN 检测器进行特征融合、候选区域状态分类及裂纹位置与长度测量，同时提出融合寿命预测模型估算 TBC 剩余寿命，该方案经不同尺度实验验证，内部裂纹定位精度达 0.898、检测精度达 0.806，在大、小尺度下均优于现有方法，可区分小型裂纹与孔隙，其融合寿命估算的 RMSE 为 2.7、计算耗时 1.63 s，所有检测条件下误差率均低于 15%。该方法为涡轮叶片 TBC 内部裂纹的精准无损评估及寿命预测、缺陷检测的设计提供参考。Shan 等[61]提出基于深度卷积神经网络（deep convolution neural network, DCNN）的微观结构自动识别方法来解决机器学习，实现 SEM 图像像素级识别，采集并标注 4 张含裂纹、等轴孔隙及未熔区的 TBC 的 SEM 图像来训练 DCNN，结果显示，训练后模型对一般 TBC 的 SEM 图像的像素识别精度达 97.0%，对未熔区较少的典型涂层精度提升至 98.5%，由于 YSZ 与非 YSZ、裂纹与等轴孔隙的误分类，部分微观结构类别存在识别误差，但少量 SEM 图像数据集即可完成模型训练，该方法为 APS 的 TBC 微观结构的定量表征提供了可行且高效的技术路径。热老化对 TBC 光学性能与孔隙结构产生影响，Blanchard 等[62]采用深度学习图像分割模型、压汞法等量化涂层孔隙，结合 1 450 K/1 111 h 的循环热处理与 FDTD 模型分析，最终气溶喷涂 TBC 的冷沉积工艺比热沉积工艺产生更多小孔隙，对应光散射更强，热处理使冷、热沉积样品的小孔隙占比分别降低 28.1%、13.5%，2 种涂层散射系数均降低约 20%，且直径低于 2 μm 的孔隙是 0.5～2.5 μm 波段光的主要散射体，其占比与散射系数呈线性关系，构建的 FDTD 模型与实验数据吻合良好，可通过修改涂层截面图预测劣化对光学性能的影响。

3 TBC 优化与性能预测

3.1 AI 驱动 TBC 工艺参数优化

APS、EB-PVD 的喷涂功率、粒子速度、沉积温度、气体压力，这些核心制备 TBC 工艺的参数直接调控 TBC 的微观结构，并进一步耦合影响涂层隔

热效率、力学强度、抗裂性等多维度核心性能。传统工艺参数选择多依赖经验试错或单目标优化，难以平衡多性能间的协同与制约关系，面对新材料体系或小样本实验场景，物理建模因缺乏充足数据支撑，泛化能力受限，导致参数优化周期长、试错成本高，且涂层性能一致性与服役稳定性难以保障。

316L 不锈钢基底具有潜在的储氢应用价值，为提升 316L/Al_2O_3 功能梯度层与 SiC 顶层涂层系统界面分层抗性，Abedini 等[63]优化功能梯度层与 SiC 顶层厚度等参数，以边缘界面分层为关键失效位置，先通过基于能量变化的有限元分析确定最大稳定裂纹长度的分层判据，又因传统热残余应力目标函数不适配本系统，推导 2 个新多目标函数，结合多目标粒子群优化(multi-objective particle swarm optimization, MOPSO)开展参数优化，该方法将最大稳定分层裂纹长度从传统方法的 7 μm 提升至 50 μm，优化后的分层抗性优于传统单目标 PSO，有效提升了涂层的界面分层抗性。APS 制备 Al_2O_3-13%TiO_2 涂层时，喷涂距离、喷枪速度等工艺参数对涂层硬度、孔隙率、空蚀抗性的影响复杂，传统参数选择难以平衡多性能目标。Szala 等[64]先通过人工神经网络(artificial neural network, ANN)构建模型，表征工艺参数与涂层性能的关联，再结合多目标遗传算法(multi-objective genetic algorithm, MOGA)开展优化以获得最大化硬度与空蚀抗性、最小化孔隙率的 TBC，对喷涂工艺参数进行优化，当喷涂距离为 86.9 mm、喷枪速度为 304 mm/s 时，侵蚀深度仅 10.1 μm，当喷涂距离约为 80 mm 时，孔隙率可低至 2.28%，此外，当喷涂距离为 90 mm、喷枪速度为 400 m/s 时，涂层表现出最高的空蚀抗性。Xu 等[65]通过高斯混合模型(Gaussian mixture model, GMM)将原始 8 组数据扩充至 400 组，再结合 SVR 构建参数与结合强度的关联模型，并用 Shapley 分析、灵敏度分析得到厚度与结合强度负相关，借助 PCA 优化参数，如图 6 所示，经验证，模型留一法交叉验证的相关系数从 0.690 提升至 0.990，在电流 617 A、电压 65 V 等优化参数下，涂层结合强度从 46.6 MPa 升至 55.5 MPa，提升超 19%，该方法既解决了小数据预测精度不足的问题，又实现了 APS 工艺参数的有效优化，为 YSZ 涂层的高性能设计提供了支撑。电沉积 Ni/Al_2O_3 纳米复合涂层的显微硬度受多工艺参数影响，传统方法很难精准预测。Dehestani 等[66]采用因子 D-最优设计实验，收集 105 次测试的温度、电流密度等电沉积参数与涂层显微硬度数据，构建人工神经网络-改进粒子群优化(artificial neural network modified particle swarm optimization, ANN-MPSO)混合模型，以多种误差指标验证模型性能，结果显示，隐藏层含 17 个神经元的 ANN-MPSO-7 模型最优，相关系数达 0.95，灵敏度分析指出电流密度、镀液 Al_2O_3 含量等是关键影响参数，残余压应力助力硬质涂层形成，该模型精准、低成本，能有效支撑 Ni/Al_2O_3 纳米复合涂层的显微硬度预测与工艺优化。APS 喷涂对 NiCrAlY 的高质量训练数据较少，传统建模既难高效关联工艺参数与飞行粒子行为，又耗费计算资源与时间。Zhu 等[67]采用深度迁移学习，先在 YSZ 粒子数据集预训练 CNN，再设计 6 种参数微调方案，用 NiCrAlY 的模拟数据优化模型，最终微调预训练 CNN 所有层及递减学习率训练损失最低、测试精度最高，滤波器可视化显示学习效果良好，且模型预测的粒子状态分布与设计目标值吻合度高，该方法不仅解决了数据不足的问题，还节省了计算资源与时间，同时有望用于新材料、新设备的 APS 喷涂优化。

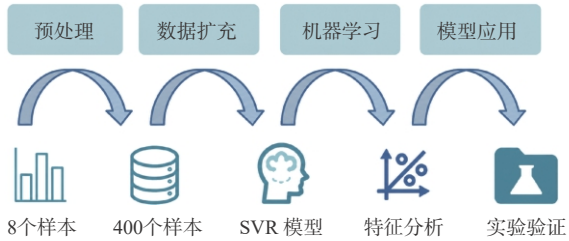


图 6　利用机器学习改善 YSZ 涂层结合强度的流程[65]

Fig. 6　Flowchart of using machine learning to improve bonding strength of YSZ coating[65]

3.2　AI 赋能 TBC 建模效率提升

计算流体力学(computational fluid dynamics, CFD)、有限元法(finite element method, FEM)等传统数值模型具有局限性，TBC 相关仿真中，CFD 与 FEM 虽能精准描述制备过程中的热、力、流体多场耦合及服役中的失效演化机制，但需通过复杂偏微分方程求解与高密度网格划分实现高精度模拟，导致单次计算耗时长达数小时至数天，且多参数迭代优化、实时工况仿真时计算成本呈指数级增长，难以适配 TBC 设计快速迭代、服役状态在线监测的工程需求。

作为新一代超高温 TBC 的核心制备技术，等离子体增强原子层沉积(plasma-enhanced atomic layer deposition, PEALD)制备 HfO_2 薄膜涂层时，与传统 TBC 制备工艺面临同样的工艺优化瓶颈。传统 3D-CFD 模型虽能精准表征沉积过程，却因单流速沉积轮廓需约 1 天导致计算耗时、成本高，难以高效探索最优操作策略，无法快速实现涂层性能的精准调控。Ding 等[68]先将 PEALD 反应器 3D-CFD 模

型简化为2D轴对称模型，再基于RNN构建数据驱动模型，利用CFD数据集训练模型以预测薄膜沉积轮廓，该模型与全3D-CFD结果的偏差不超过5%，且计算效率大幅提升，仅需几秒即可完成预测，可用于实时过程控制，随后利用该模型生成操作策略数据库，结合行业标准确定可行操作区域，再通过前驱体分析为每个半周期选定最优操作条件，同时该模型也适用于带喷头的PEALD反应器，为其他半导体制造过程提供建模参考，助力提升过程操作性能。Balasubramaniam等[35]提出仿真辅助AI方法，先通过多层传热模型生成低保真热剖面训练深度神经网络，再融合脉冲红外热成像的高保真热图谱优化模型，用于预测24～120 μm厚顶层涂层的热导率、比热容与厚度，结果显示，热导率、比热容预测误差仅1.71%、1.37%，厚度误差为0.81%～6.14%，精度良好。这种融合低保真仿真与高保真实验数据的方式，既解决了训练数据不足的问题，还能拓展至涂层分层检测、寿命预测，为航空部件的维护与寿命评估提供了高效支撑。含CMAS夹杂的陶瓷涂层等两相材料的热应力评估，传统依赖FEM但耗时极长，小样本下也难高效建模。Ning等[59]提出条件生成对抗网络(conditional generative adversarial network, CGAN)，先构建含CMAS夹杂的陶瓷涂层两相结构SEM合成图像，用灰度分割基体与夹杂，再通过图像恢复FEM计算等温过程的热应力，将仅8张完整结构及对应应力分布分割为局部图像，训练CGAN直接从SEM图像预测热应力，结果显示，模型与FEM应力的偏差在超98%区域小于6 MPa，仅少数热点偏差略大，且效率远优于FEM，FEM-CGAN耗时仅6 s，效率平均提升116倍，该方法仅需小样本即可训练，适配工程中热应力的快速评估。

3.3 AI实现TBC工艺、微观结构、性能的精准关联预测

通过AI模型打破TBC的工艺参数、微观结构、服役性能、寿命等关键指标评估的信息割裂，构建多维度、非线性的全链条跨尺度关联体系。涂层全链条中，APS喷涂功率、粒子速度、沉积温度等工艺参数直接调控孔隙率、孔裂比、晶粒尺寸、粒子温速等微观结构特征，而微观结构的优劣又进一步决定热边界电阻(thermal boundary resistance, TBR)、隔热效率、力学强度等服役性能与寿命可靠性，各环节存在复杂耦合关系，覆盖性能指标量化、结构参数表征、寿命精准评估等核心需求。传统方法多依赖单一环节的经验试错或局部建模，难以捕捉全链条的非线性关联，导致预测精度不足、各环节协同性差，无法满足TBC高性能定制与全生命周期管控的需求。AI技术的引入为全链条定量预测提供了核心支撑，实现从工艺参数到结构特征、从结构特征到服役性能、从服役性能演化到寿命评估的全链条定量预测。

APS飞行粒子的温度、速度与工艺参数间存在复杂非线性关系，且喷枪电极老化会改变这一关联，传统方法既难同时预测温速，也难兼顾老化的影响，Yu等[69]同步记录工艺参数、粒子温速与电极使用时间，采用RF与梯度提升(gradient boosting, GB)2种集成方法，结合直接嵌入、先差分平稳再嵌入2种时间序列策略，进行特征排序与预测，结果显示，RF比GB更适用，RF能同时预测粒子温速、考虑二者交互，而GB仅能分别预测，且将数据作为时间序列、先差分平稳后，预测效果更优，粒子特征也清晰地反映了电极老化的影响。APS喷涂8YSZ的工艺参数、微观结构与力学性能间存在复杂非线性关联，传统定性实验难以实现定量控制，小样本下精准建模也存在难度。Zhu等[70]将陶瓷粉末粒径、喷涂功率等作为输入，孔隙率、结合强度等作为输出，分别构建BPNN与花朵授粉算法(flower pollination algorithm, FPA)优化的ELM模型，并用k折交叉验证评估鲁棒性，结果显示，FPA-ELM模型的回归系数R>0.94、RMSE<2，预测精度与鲁棒性优于BPNN，该模型可实现涂层制备的定量控制，为TBC工艺优化、材料研发提供了新方法。Ye等[71]提出布谷鸟搜索算法优化SVM方法，以APS的粉末粒径、喷涂距离、功率为输入，孔隙率、孔裂比、最大费雷特直径等5种微观结构特征为输出，基于15个试样构建数据集训练模型，经验证该模型的相关系数R>95%，RMSE、平均绝对误差等指标均小于0.1，经过12个随机试样进一步验证，模型性能依然稳定，适配小样本回归预测需求，高效建立工艺参数与微观结构的关联，用于构建工艺参数-微观结构-服役性能的完整联系，助力TBC的服役性能提升与安全状态监测。

TBR是TBC的核心性能指标，但传统声学失配、扩散失配模型的预测精度不足，难以支撑开发。Zhan等[72]收集实验TBR数据作为训练集，将相关材料属性设为描述符，训练了包括GPR、SVR在内的机器学习模型，这些模型的预测精度显著优于传统模型，其中，GPR与SVR的表现更优，同时对比不同描述符集后发现，膜厚是影响TBR预测的关键参数。Gao等[73]提出Adap-Alex算法，基于CNN用3×3、步长1的小卷积核减少参数计算，同时调整感受野、步长等参数，结合sigmoid函数实现

自适应池化, 解决了传统 CNN 训练复杂 TBC 图像时的耗时、过拟合问题, 通过热震试验构建 TBC 数据集后验证, 该算法的平均训练时间比 VGG-Net 短 125 s、比 Alex-Net 短 685 s, 热振动 300 次时, 测试精度达 93%, 高于 Alex-Net 与 VGG-Net, 能高效精准判断 TBC 寿命, 有效提升了航空发动机叶片 TBC 寿命预测的效率与精度。Ma 等[74] 提出 M-GPR 算法, 通过 Morlet 小波分解结合超声反射幅值谱, 解耦含孔隙率信息的多尺度声学特征, 再以 GA 与交叉验证优化 GPR, 用提取的特征与孔隙率数据训练模型, 如图 7 所示, 同时构建随机孔隙模型(random pore model, RPM), 在孔隙率 1%～9% 的 ZrO_2 涂层 RPM 上仿真验证, 预测相对误差低至 0.60%、最高 6.28%, 精度比 BPNN 高 16%, 25 个训练 RPM 的预测误差低于 5.7%, 5 个测试 RPM 结果与实际孔隙率吻合度良好。

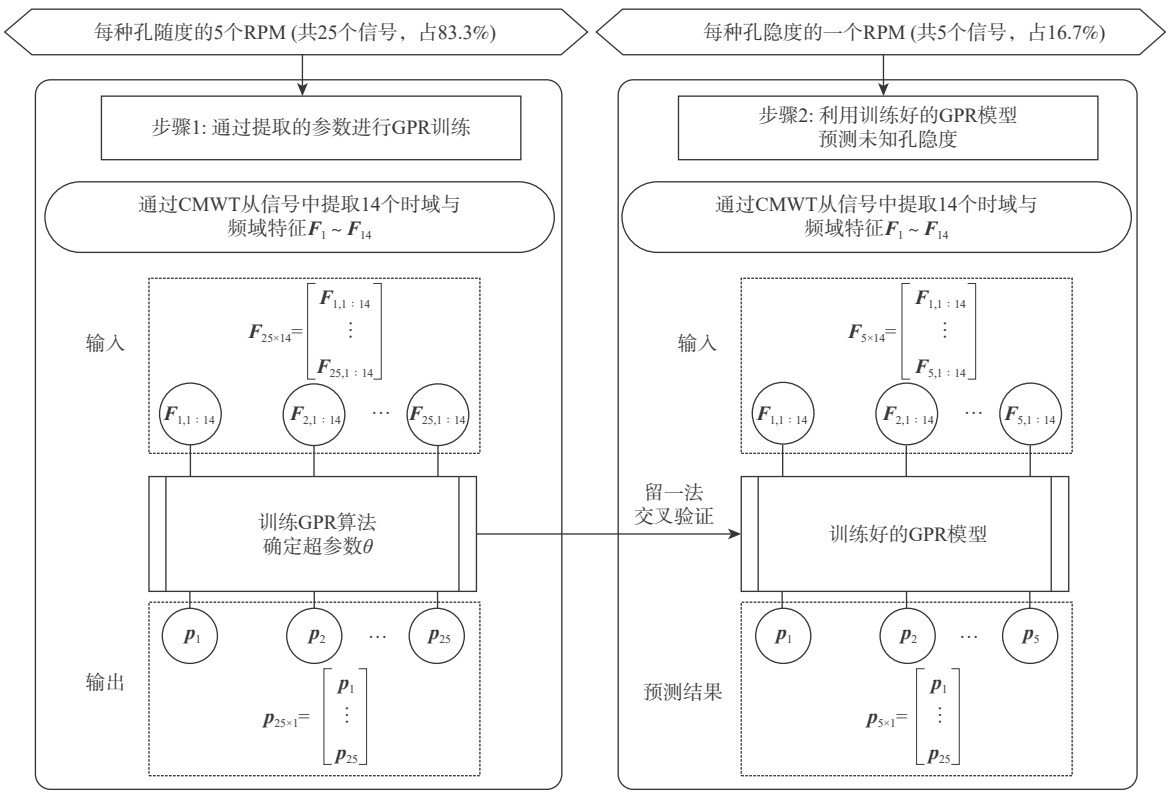


图 7　GPR 算法流程[74]

Fig. 7　Flow chart of GPR algorithm[74]

4　TBC 监测、健康管理和决策

TBC 作为航空发动机涡轮叶片等高温部件的核心防护材料, 其陶瓷层厚度、介电常数、黏结层电导率及结构完整性直接决定部件隔热性能与服役安全, 精准的状态监测、健康评估及决策支持是保障发动机可靠运行的关键。传统监测与健康管理方法面临挑战, 如多参数同步反演难度大、应力评估依赖 FEM 导致耗时冗长、复杂信号中结构参数提取效率低, 难以满足实时监测、快速失效预警的工程需求, 制约了 TBC 全生命周期安全管控能力的提升。传统方法难以同步反演陶瓷层、黏结层参数变化, 直接导致失效风险, 影响航空发动机的安全监测[75-78]。

Wang 等[79] 提出基于电磁/电容双模传感器的 BPNN 方法, 集电磁模式下不同激励频率的信号, 结合电容模式信号作为输入, 构建模型同步反演陶瓷层厚度、介电常数及黏结层电导率这 3 个关键参数, 实验验证显示, 所有参数的反演误差均小于 2%, 满足工业精度要求。Ning 等[80] 提出图像识别和 ANN 方法, 构建含 CMAS 的陶瓷顶层涂层两相图像, 识别边界后用平均坐标法优化, 再构建图像恢复有限元模型计算等温热应力, 提取结构参数与局部最大应力建立样本库, 训练 ANN 预测新样本的应力, 如图 8 所示, 结果显示, 新方法应力预测 RMSE 仅约 6%, 耗时不足 5 s, 而 FEM 耗时 265～382 s, 新方法快 77 倍, 还能精准定位椭圆最近点的高应力区, 该方法高效精准, 且应力与椭圆几何参数相关, 未来可拓展到其他两相工程结构的应力评估。Ye 等[81] 提出太赫兹光谱结合混合机器学习的方法, 先通过 FDTD 算法模拟 64 组结构参数对应的太赫兹时域与反射信号, 添加高斯白噪声模拟实

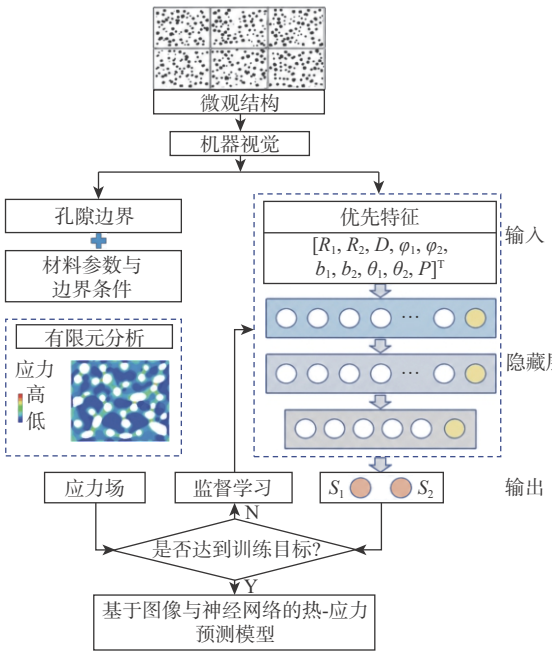


图 8 图像与基于 ANN 的预测模型架构[80]

Fig. 8 Architecture of image and ANN-based prediction model[80]

际测试环境后，用 haar 小波降噪处理，再通过将高维数据压缩至累计贡献率超 95% 的低维主成分实现 PCA 降维，减少冗余信息、提升效率，最终构建 ELM 模型，用 GA 优化得到 PCA-GA-ELM，用于预测涂层结构参数，该模型在预测性能与计算速度上均表现优异，且反射数据更适配建模。这种方法结合了太赫兹无损技术的精准性与机器学习的高效性，能可靠支撑多层 TBC 的结构健康监测，为航空发动机热端部件的安全保障提供了新手段。

5 结束语

TBC 智能化设计依托数据驱动与 AI，已在多维度实现系统性突破，推动研发模式转型，同时新一代高温设备对 TBC 的性能需求已从单一热防护向多功能集成升级，未来研发需聚焦结构功能一体化与多目标协同优化，AI 将成为突破核心瓶颈的关键支撑。未来 TBC 设计需重点向集成方向发展，兼顾抗 CMAS 侵蚀、耐高温腐蚀、抗热震性与雷达吸波隐身等功能，构建“隔热-防护-功能”一体化体系。然而，抗 CMAS 侵蚀、抗热震性与低热导率等目标之间常存在此消彼长的制约关系，需依托 AI 驱动的多目标优化算法，构建“工艺-结构-多性能”耦合模型，挖掘性能协同区间，实现多目标平衡设计。

AI 和 TBC 的核心发展方向包括 3 方面：①AI 融合高通量计算，优化材料热物理性能综合评价描述符，解决当前高通量筛选的核心痛点，加速新一代超高温 TBC 材料研发。②构建跨尺度多场 AI 模型，整合微观晶粒、介观梯度结构与宏观服役形态数据，精准解析热-力-电磁多场耦合机制，支撑一体化结构设计。③AI 和数字孪生技术，结合实时传感数据动态预测涂层性能演化，输出自适应修复与调控策略，实现服役全生命周期智能管控。

AI 技术为 TBC 的研发与应用带来了颠覆性的变革机遇，未来通过突破工程化应用的核心瓶颈，推动 TBC 从单一防护材料升级为多功能智能适配材料，有望攻克新一代超高温 TBC 研发的核心难题，为高推重比航空发动机、重型燃气轮机等高端设备的发展提供核心材料支撑。

参考文献（References）

[1] 徐惠彬, 宫声凯, 刘福顺. 航空发动机热障涂层材料体系的研究[J]. 航空学报, 2000, 21(1): 7-12.
Xu H B, Gong S K, Liu F S. Recent development in materials design of thermal barrier coatings for gas turbine[J]. Acta Aeronautica et Astronautica Sinica, 2000, 21(1): 7-12(in Chinese).

[2] 刘永泉, 刘太秋, 季路成. 航空发动机风扇/压气机技术发展的若干问题与思考[J]. 航空学报, 2015, 36(8): 2563-2576.
Liu Y Q, Liu T Q, Ji L C. Some problems and thoughts in the development of aero-engine fan/compressor[J]. Acta Aeronautica et Astronautica Sinica, 2015, 36(8): 2563-2576(in Chinese).

[3] Miller R A. Thermal barrier coatings for aircraft engines: history and directions[J]. Journal of Thermal Spray Technology, 1997, 6(1): 35-42.

[4] Darolia R. Thermal barrier coatings technology: critical review, progress update, remaining challenges and prospects[J]. International Materials Reviews, 2013, 58(6): 315-348.

[5] 郭洪波, 宫声凯, 徐惠彬. 先进航空发动机热障涂层技术研究进展[J]. 中国材料进展, 2009, 28(9): 18-26.
Guo H B, Gong S K, Xu H B. Progress in thermal barrier coatings for advanced aeroengines[J]. Materials China, 2009, 28(9): 18-26(in Chinese).

[6] Morelli S, Bursich S, Bolelli G, et al. Thermal conductivity and micromechanical properties of plasma-sprayed yttria-stabilized zirconia thermal barrier coatings[J]. Surface and Coatings Technology, 2025, 513: 132498.

[7] Rima A N, Rahman M A, Ferdous R, et al. DFT simulation to study the physical properties of ternary intermetallic materials ACuSb (A=Ca,Sr,Ba) for solar cell and TBC materials[J]. Computational Condensed Matter, 2024, 39: e00900.

[8] Liu N N, Dou R F, Wen Z, et al. Correction to: numerical and experimental studies on the in-situ measurement of thermal conductivity of the thermal barrier coating[J]. Journal of Thermal Spray Technology, 2025, 34(1): 477.

[9] Schulz U, Saruhan B, Fritscher K, et al. Review on advanced EB-PVD ceramic topcoats for TBC applications[J]. International Journal of Applied Ceramic Technology, 2004, 1(4): 302-315.

[10] Mohanraj R, Elangovan S. Incremental sheet metal forming of Ti-6Al-4V alloy for aerospace application[J]. Transactions of the Canadian Society for Mechanical Engineering, 2020, 44(1): 56-64.

[11] 郭洪波, 宫声凯, 徐惠彬. 梯度热障涂层的设计[J]. 航空学报, 2002, 23(5): 467-472.
Guo H B, Gong S K, Xu H B. Design of gradient thermal barrier coatings[J]. Acta Aeronautica et Astronautica Sinica, 2002, 23(5): 467-472(in Chinese).

[12] Liu B, Liu Y C, Zhu C H, et al. Advances on strategies for searching for next generation thermal barrier coating materials[J]. Journal of Materials Science & Technology, 2019, 35(5): 833-851.

[13] Xu K, Xiao X L, Wang L J, et al. Data-driven materials research and development for functional coatings[J]. Advanced Science, 2024, 11(42): 2405262.

[14] Jithin P V, Bitla Y, Patidar M M, et al. Structural, magnetic and electrical transport properties of the sol-gel derived $La_{1-x}Ca_xMnO_3$ ($0 \leqslant x \leqslant 0.3$) nanoparticles[J]. Materials Chemistry and Physics, 2023, 301: 127651.

[15] Kim H, Lee H, Son H, et al. Ab initio calculations of Nb-based MAX phases as bond coats for thermal barrier coatings[J]. Journal of Materials Research and Technology, 2024, 32: 2495-2501.

[16] Ahmed R, Islam M, Hossain M M, et al. A comprehensive first-principles insights into the physical properties of binary intermetallic Zr_3Ir compound[J]. Results in Materials, 2024, 21: 100518.

[17] 范宝安, 颜家保, 邱江华, 等. 水热法制备 8YSZ 纳米粉体及其烧结活性的研究[J]. 纳米技术, 2011, 1(1): 6-11.
Fan B A, Yan J B, Qiu J H, et al. Hydrothermal synthesis of 8YSZ nanopowders and their sinterabilities[J]. Hans Journal of Nanotechnology, 2011, 1(1): 6-11(in Chinese).

[18] Audigié P, Vande Put A, Malard B, et al. Analysis of void formation in Pt-rich γ-γ′ bond-coatings for TBC system application[J]. Corrosion Science, 2025, 249: 112785.

[19] Shahbazi H, Vakilifard H, Nair R B, et al. High entropy alloy bond coats for thermal barrier coatings: a review[J]. Journal of Thermal Spray Technology, 2024, 33(2): 430-446.

[20] Poliarus O, Morgiel J, Pomorska M, et al. Effect of HT cycling in air on microstructure of detonation Gun sprayed YSZ/NiCr coating [J]. Journal of Thermal Spray Technology, 2024, 33(1): 148-159.

[21] Ferrasse T, Costil S, Darut G, et al. Improved adhesion of plasma-sprayed ceramic coatings on textured ceramic substrates[C]//Proceedings from the International Thermal Spray Conference. Russel: ASM International, 2024, 84864: 356-364.

[22] Mia M H, Parvin F, Islam A K M A, et al. Ternary chalcogenides $NbIn{X}_2$ (X = S,Se): a comprehensive investigation of mechanical, electronic, vibrational, optical and thermophysical properties[J]. International Journal of Modern Physics B, 2025, 39(14): 2550118.

[23] Liu B S, Zhao J L, Chu K L, et al. Unraveling the atomic-scale mechanisms of hydrogen defects behavior in yttria-stabilized tetragonal zirconia by first principles calculation[J]. Journal of Advanced Ceramics, 2025, 14(7): 9221099.

[24] Qu D B, Cai Z H, Chen T Y, et al. An advanced methodology for nondestructive measurement of thermal barrier coating thickness via point cloud processing[J]. Journal of Thermal Spray Technology, 2024, 33(1): 51-70.

[25] 王一丹, 郭谦, 周琪杰, 等. 先进航空发动机高温功能涂层研究进展[J]. 航空材料学报, 2024, 44(5): 48-69.
Wang Y D, Guo Q, Zhou Q J, et al. Research progress in high temperature functional coatings for advanced aeroengines[J]. Journal of Aeronautical Materials, 2024, 44(5): 48-69(in Chinese).

[26] Küçük Y, Öge M. Elevated temperature wear behavior of FeCr slag coating as an alternative coating material for caster rolls[J]. Protection of Metals and Physical Chemistry of Surfaces, 2022, 58(1): 119-128.

[27] Che J W, Liu X Y, Wang X Z, et al. Ultralow oxygen ion diffusivity in pyrochlore-type $La_2(Zr_{0.7}Ce_{0.3})_2O_7$[J]. Journal of Materials Science & Technology, 2022, 102: 174-185.

[28] Sokolowski P, Kielczawa T, Musalek R, et al. The behavior of plasma sprayed thermal barrier coating with laser microtextured bond coat under high temperature testing[J]. Surface and Coatings Technology, 2023, 453: 129095.

[29] Wu Y S, Wang Y, Gildersleeve V E J, et al. Laser thermal gradient testing of air plasma sprayed multilayered, multi-material thermal barrier coatings[J]. International Journal of Applied Ceramic Technology, 2022, 19(5): 2632-2647.

[30] Kumar D, Pandey K N. Experimental investigations of sol-gel process parameters for wear reduction on thermal barrier coated AA2024 aluminum alloys with the use of Taguchi-based optimization[J]. Sādhanā, 2020, 45(1): 187.

[31] Mori K, Suga C, Hasegawa J. Development of elastic modulus measurement system for multilayer thermal barrier coatings at high temperatures by impact-excited vibration[J]. Journal of the Japanese Society for Experimental Mechanics, 2021, 21(3): 215-221.

[32] Abdelgawad A, Al-Athel K. Effect of TGO thickness, pores, and creep on the developed residual stresses in thermal barrier coatings under cyclic loading using SEM image-based finite element model [J]. Ceramics International, 2021, 47(14): 20064-20076.

[33] Yu C T, Liu H, Jiang C Y, et al. Modification of NiCoCrAlY with Pt: Part II. application in TBC with pure metastable tetragonal (t′) phase YSZ and thermal cycling behavior[J]. Journal of Materials Science & Technology, 2019, 35(3): 350-359.

[34] 郭洪波, 宫声凯, 徐惠彬. 新型高温/超高温热障涂层及制备技术研究进展[J]. 航空学报, 2014, 35(10): 2722-2732.
Guo H B, Gong S K, Xu H B. Research progress on new high/ultra-high temperature thermal barrier coatings and processing technologies[J]. Acta Aeronautica et Astronautica Sinica, 2014, 35(10): 2722-2732(in Chinese).

[35] Kunjl Purayil S K, Puthiyaveettil N, Chakravarthy V S, et al. Simulation-assisted AI for the evaluation of thermal barrier coatings using pulsed infrared thermography[J]. Journal of Applied Physics, 2022, 132(6): 064902.

[36] Zhu W J, Gan M D, Ye B, et al. Optimization of the thermophysical properties of the thermal barrier coating materials based on GA-SVR machine learning method: illustrated with ZrO_2 doped $DyTaO_4$ system[J]. Materials Research Express, 2021, 8(12): 125503.

[37] Chen L H, Tran H, Batra R, et al. Machine learning models for the lattice thermal conductivity prediction of inorganic materials[J]. Computational Materials Science, 2019, 170: 109155.

[38] Halloua H, Elhassnaoui A, Saifi A, et al. Neural networks and genetic algorithms for the evaluation of coatings thicknesses in thermal barriers by infrared thermography data[J]. Procedia Structural Integrity, 2017, 5: 997-1004.

[39] Halloua H, Saifi A, El-Amiri A, et al. The processing of active infrared thermography data by a hybrid neural algorithm for the eva-

luation of thermal barrier coating thicknesses[C]//Proceedings of the 6th International Renewable and Sustainable Energy Conference. Piscataway: IEEE Press, 2019: 1-6.

[40] Zhang J Y, Meng X B, Ma Y C. A new measurement method of coatings thickness based on lock-in thermography[J]. Infrared Physics & Technology, 2016, 76: 655-660.

[41] Sun F S, Fan M B, Cao B H, et al. Terahertz based thickness measurement of thermal barrier coatings using long short-term memory networks and local extrema[J]. IEEE Transactions on Industrial Informatics, 2022, 18(4): 2508-2517.

[42] Sun F S, Fan M B, Cao B H, et al. THzResNet: a physics-inspired two-stream residual network for thermal barrier coating thickness measurement[J]. IEEE Transactions on Industrial Informatics, 2023, 19(7): 8328-8339.

[43] Yuan B H, Wang W Z, Ye D D, et al. Nondestructive evaluation of thermal barrier coatings thickness using terahertz technique combined with PCA-GA-ELM algorithm[J]. Coatings, 2022, 12(3): 390.

[44] Wang Y, Cui S Y, Yu Q M, et al. The effect of pore structure on the erosion resistance of air plasma sprayed thermal barrier coatings on finite element simulation[J]. Scientific Reports, 2025, 15: 26032.

[45] 周洪, 李飞, 何博, 等. 热障涂层材料研究进展[J]. 材料导报, 2006, 20(10): 40-43.
Zhou H, Li F, He B, et al. Research progresses in materials for thermal barrier coatings[J]. Materials Reports, 2006, 20(10): 40-43 (in Chinese).

[46] Evans A G, Mumm D R, Hutchinson J W, et al. Mechanisms controlling the durability of thermal barrier coatings[J]. Progress in Materials Science, 2001, 46(5): 505-553.

[47] Zhang S X, Lv G L, Guo S F, et al. Porosity characterization of thermal barrier coatings by ultrasound with genetic algorithm back-propagation neural network[J]. Complexity, 2021, 2021: 8869928.

[48] Ma Z Y, Zhang W, Luo Z B, et al. Ultrasonic characterization of thermal barrier coatings porosity through BP neural network optimizing Gaussian process regression algorithm[J]. Ultrasonics, 2020, 100: 105981.

[49] Li R, Ye D, Zhang Q, et al. Nondestructive evaluation of thermal barrier coatings' porosity based on terahertz multi-feature fusion and a machine learning approach[J]. Applied Sciences, 2023, 13(15): 8988.

[50] 薛召露, 郭洪波, 宫声凯, 等. 新型热障涂层陶瓷隔热层材料[J]. 航空材料学报, 2018, 38(2): 10-20.
Xue Z L, Guo H B, Gong S K, et al. Novel ceramic materials for thermal barrier coatings[J]. Journal of Aeronautical Materials, 2018, 38(2): 10-20(in Chinese).

[51] 郑蕾, 郭洪波, 郭磊, 等. 新一代超高温热障涂层研究[J]. 航空材料学报, 2012, 32(6): 14-24.
Zheng L, Guo H B, Guo L, et al. New generation thermal barrier coatings forultrahigh temperature applications[J]. Journal of Aeronautical Materials, 2012, 32(6): 14-24(in Chinese).

[52] Unnikrishnakurup S, Dash J, Ray S, et al. Nondestructive evaluation of thermal barrier coating thickness degradation using pulsed IR thermography and THz-TDS measurements: a comparative study [J]. NDT & E International, 2020, 116: 102367.

[53] 王飞, 董会, 伊静, 等. 双层结构黏结层热生长氧化物近指数生长机理研究[J]. 表面技术, 2026, 55(1): 198-207.
Wang F, Dong H, Yi J, et al. Near-exponential growth mechanism of thermally grown oxide in double bond coat[J]. Surface Technology, 2026, 55(1): 198-207(in Chinese).

[54] Wang Z Y, Zhao M, Wu R F, et al. Non-destructive evaluation of thermally grown oxides in thermal barrier coatings using impedance spectroscopy[J]. Journal of the European Ceramic Society, 2019, 39(15): 5048-5058.

[55] Hutchinson J W, Evans A G. On the delamination of thermal barrier coatings in a thermal gradient[J]. Surface and Coatings Technology, 2002, 149(2-3): 179-184.

[56] 徐惠彬, 宫声凯, 蒋成保, 等. 特种功能材料中的固态相变及应用[J]. 中国材料进展, 2011, 30(9): 1-12.
Xu H B, Gong S K, Jiang C B, et al. Solid-state phase transformations and theirapplications in special functional materials[J]. Materials China, 2011, 30(9): 1-12(in Chinese).

[57] Luo M T, Zhong S C, Yao L G, et al. Thin thermally grown oxide thickness detection in thermal barrier coatings based on SWT-BP neural network algorithm and terahertz technology[J]. Applied Optics, 2020, 59(13): 4097.

[58] Li R, Ye D D, Xu J F, et al. Multi-scale analysis of terahertz time-domain spectroscopy for inversion of thermal growth oxide thickness in thermal barrier coatings[J]. Coatings, 2023, 13(7): 1294.

[59] Ning L Y, Cai Z W, Liu Y Z, et al. Conditional generative adversarial network driven approach for direct prediction of thermal stress based on two-phase material SEM images[J]. Ceramics International, 2021, 47(24): 34115-34126.

[60] Shi L C, Long Y, Wang Y Z, et al. Evaluation of internal cracks in turbine blade thermal barrier coating using enhanced multi-scale faster R-CNN model[J]. Applied Sciences, 2022, 12(13): 6446.

[61] Shan X, Huang T M, Luo L R, et al. Automatic recognition of microstructures of air-plasma-sprayed thermal barrier coatings using a deep convolutional neural network[J]. Coatings, 2023, 13(1): 29.

[62] Blanchard F, Kadi M J, Bousser E, et al. Effect of thermal ageing on the optical properties and pore structure of thermal barrier coatings[J]. Surface and Coatings Technology, 2023, 452: 129080.

[63] Abedini S, Dong C S, Davies I J. Multi-objective particle swarm optimisation of multilayer functionally graded coating systems for improved interfacial delamination resistance[J]. Materials Today Communications, 2020, 24: 101202.

[64] Szala M, Łatka L, Awtoniuk M, et al. Neural modelling of APS thermal spray process parameters for optimizing the hardness, porosity and cavitation erosion resistance of Al_2O_3-13 wt% TiO_2 coatings[J]. Processes, 2020, 8(12): 1544.

[65] Xu P C, Chen C, Chen S Z, et al. Machine learning-assisted design of yttria-stabilized zirconia thermal barrier coatings with high bonding strength[J]. ACS Omega, 2022, 7(24): 21052-21061.

[66] Dehestani M, Khayati G R, Sharafi S. An improved optimization model to predict the microhardness of Ni/Al_2O_3 nanocomposite coatings prepared by electrodeposition: a hybrid artificial neural network-modified particle swarm optimization approach[J]. Measurement, 2021, 179: 109423.

[67] Zhu J W, Wang X Z, Kou L Y, et al. Application of combined transfer learning and convolutional neural networks to optimize plasma spraying[J]. Applied Surface Science, 2021, 563: 150098.

[68] Ding Y Y, Zhang Y C, Chung H Y, et al. Machine learning-based

modeling and operation of plasma-enhanced atomic layer deposition of hafnium oxide thin films[J]. Computers & Chemical Engineering, 2021, 144: 107148.

[69] Yu K R, Cojocaru C V, Ilinca F, et al. Ensemble methods for APS in-flight particle temperature and velocity prediction considering torch electrodes ageing[J]. Journal of Thermal Spray Technology, 2023, 32(1): 175-187.

[70] Zhu H, Li D P, Yang M, et al. Prediction of microstructure and mechanical properties of atmospheric plasma-sprayed 8YSZ thermal barrier coatings using hybrid machine learning approaches [J]. Coatings, 2023, 13(3): 602.

[71] Ye D D, Wang W Z, Xu Z, et al. Prediction of thermal barrier coatings microstructural features based on support vector machine optimized by cuckoo search algorithm[J]. Coatings, 2020, 10(7): 704.

[72] Zhan T Z, Fang L, Xu Y B. Prediction of thermal boundary resistance by the machine learning method[J]. Scientific Reports, 2017, 7: 7109.

[73] Gao R P, Mao W, Wang Y R, et al. Intelligent life prediction of thermal barrier coating for aero engine blades[J]. Coatings, 2021, 11(8): 890.

[74] Ma Z Y, Sun L M, Chen Y J, et al. Ultrasonic prediction of thermal barrier coating porosity through multiscale-characteristic-based Gaussian process regression algorithm[J]. Applied Acoustics, 2022, 195: 108831.

[75] Liu Z Y, Yang L, Zhou Y C. A multiscale model integrating artificial neural networks for failure prediction in turbine blade coatings[J]. Surface and Coatings Technology, 2023, 457: 129218.

[76] Tan Z X, Zhong S S, Lin L. Trans-layer model learning: a hierarchical modeling strategy for real-time reliability evaluation of complex systems[J]. Reliability Engineering & System Safety, 2019, 182: 120-132.

[77] Liu H, Sun J Z, Lei S Y, et al. In-service aircraft engines turbine blades life prediction based on multi-modal operation and maintenance data[J]. Propulsion and Power Research, 2021, 10(4): 360-373.

[78] Bodenschatz C J, Saidi W A, Stokes J L, et al. Theoretical prediction of thermal expansion anisotropy for $Y_2Si_2O_7$ environmental barrier coatings using a deep neural network potential and comparison to experiment[J]. Materials, 2024, 17(2): 286.

[79] Wang W, Ren Y, Hu J F, et al. Simultaneous inversion method of thermal barrier coating parameters based on electromagnetic/capacitive dual-module sensor[J]. Review of Scientific Instruments, 2023, 94: 015002.

[80] Ning L Y, Cai Z W, Zhao X F, et al. Fast stress evaluation of the top coat of thermal barrier coatings under CaO-MgO-Al_2O_3-SiO_2 penetration based on image recognition and an artificial neural network[J]. Ceramics International, 2021, 47(13): 18252-18261.

[81] Ye D D, Wang W Z, Yin C D, et al. Pulsed Terahertz spectroscopy combined with hybrid machine learning approaches for structural health monitoring of multilayer thermal barrier coatings[J]. Optics Express, 2020, 28(23): 34875.

Research progress on intelligent design of thermal barrier coatings

HE Jian[1, 2], LI Hongfeng[1, ☆], KANG Peng[1, 3], ZHENG Lei[1, 3], LI Zhen[1], GUO Hongbo[1, 2, 3, *]

(1. School of Materials Science and Engineering, Beihang University, Beijing 100191, China;
2. Key Laboratory of High-Temperature Structural Materials and Protective Coatings (Ministry of Industry and Information Technology), Beijing 100191, China;
3. Tianmushan Laboratory, Beihang University, Hangzhou 311115, China)

Abstract: The sustained escalation in aero-engine thrust-to-weight ratios and turbine inlet temperatures has established thermal barrier coatings (TBCs) as an indispensable solution for surpassing the intrinsic temperature limits of hot-section components and ensuring their long-term service reliability. Nevertheless, the service response of TBCs involves exceedingly complex thermo-mechanical-chemical (TMC) multi-field coupling and cross-scale damage evolution, which poses severe fundamental challenges to conventional single-scale experimental and analytical methodologies. This review systematically summarizes the advances in conventional and advanced modeling, simulation, and analysis techniques for TBC and further extends the discussion to the application of artificial intelligence (AI) technologies. The study aims to provide a systematic overview of analysis methods for TBC service behavior and structural optimization strategies for coatings and to promote the implementation of proactive maintenance strategies. The technical methods covered include thermal conductivity modeling, thermal stress simulation, crack propagation analysis, and mechanical behavior studies of TBC systems, which predominantly rely on diverse techniques such as multiscale modeling and finite element analysis. The paper focuses on the use of AI in TBC simulation, including data-driven modeling approaches and machine learning algorithms such as support vector machines, random forests, neural networks, and deep learning. The findings indicate that AI holds the potential to revolutionize TBC technology in multiple respects, encompassing material and structural design, performance optimization and prediction, condition monitoring, fault diagnosis and health management, and remaining useful life (RUL) prediction, while significantly improving prediction accuracy. Novel TBCs are anticipated to offer robust support for next-generation aero-engines to transcend temperature limits and achieve improved efficiency and reliability with ongoing advancements in the intelligent design of new materials.

Keywords: thermal barrier coating; intelligent design; structural design; service performance; machine learning

Received: 2025-12-30; **Accepted**: 2026-05-08; **Published Online**: 2026-05-25 11:11
URL: link.cnki.net/urlid/11.2625.V.20260523.1228.001
Foundation item: National Natural Science Foundation of China (U21B2052)
* **Corresponding author**. E-mail: guo.hongbo@buaa.edu.cn
☆ **Co-first Author.** E-mail: lihongfeng@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0565

宽温域超弹性形状记忆合金设计优化研究进展

钱明芳[1,2]，尹钊[1]，王桂松[1]，张学习[1,3,*]，耿林[1,3,4]

(1. 哈尔滨工业大学 材料科学与工程学院，哈尔滨 150001；2. 哈尔滨工业大学 材料结构精密焊接与连接全国重点实验室，哈尔滨 150001；3. 哈尔滨工业大学 金属精密热加工国家级重点实验室，哈尔滨 150001；4. 哈尔滨工业大学 苏州研究院，苏州 215104)

摘 要：超弹性是形状记忆合金最重要的功能特性之一，其本质来源于应力诱导的可逆热弹性马氏体相变。超弹性温度窗口主要受合金化学成分与微观组织结构的协同影响。为此，系统回顾了近10年来具有宽温域超弹性特性的多类形状记忆合金体系的发展现状，重点比较了不同合金在超弹性温度窗口、可恢复应变大小及临界应力对温度的依赖性等关键性能参数方面的差异与特点。通过对比分析，为新型宽温域超弹性形状记忆合金的设计、制备与工程应用提供了理论支持与实践参考，并对该领域未来的发展方向进行了展望。

关 键 词：形状记忆合金；超弹性；宽温域；马氏体相变；组织调控

中图分类号：TB31; TB34; TB381

文献标志码：A **文章编号**：1001-5965(2026)10-3313-16

形状记忆合金(shape memory alloy, SMA)因其独特的形状记忆效应和超弹性特性，在航空航天、汽车、医疗器械等领域受到广泛关注[1-3]。SMA的超弹性是指材料在外力作用下发生可逆的大应变(通常4%～8%，甚至更高)，卸载后几乎完全恢复原始形状，而无需加热恢复的一种力学行为，其源于应力诱导马氏体相变(stress-induced martensitic transformation, SIMT)。施加外载时，母相转变为马氏体，产生显著的非线性变形；卸除外载后，马氏体因热力学不稳定性自发逆转变为奥氏体，从而实现形变恢复[4]。需要说明的是，本文所讨论的“超弹性(superelasticity)”是材料学和SMA研究领域普遍采用的术语，特指由应力诱导的可逆马氏体相变所导致的可恢复应变行为[4-5]，而非力学学科中多用于橡胶、高分子、软组织等高弹性体的“超弹性(hyperelasticity)”概念，高弹性体的超弹性主要来源于分子链的拉伸与收缩，通常可用超弹性本构模型来描述[6-7]。

根据应力-应变曲线的特征，SMA的超弹性可分为宽滞后型与窄滞后型2类。宽滞后超弹性通常伴随着一级马氏体相变[8-10]，具有明显的应力滞后，可在循环加卸载过程中形成应力-应变环，从而实现机械能耗散，可应用于被动阻尼和缓冲器件[11-13]。相比之下，窄滞后或近线性超弹性通常与应力诱导弱马氏体相变或连续马氏体相变有关[14-17]，其应力响应更平滑、滞后更小，适用于对灵敏度、精度、循环稳定性和能量效率要求较高的精密工程场景[16]。

考虑到SMA广泛的应用领域，其实际应用往往涉及多变或极端的温度环境。例如，民用汽车中发动机舱的温度可达233～398 K，驾驶舱区域通常处于233～358 K之间[18]。而在深空探测等极端场景下，SMA器件需面对更严峻的环境挑战，如火星地表温度波动范围为120～290 K，月球昼夜温差甚至高达100～400 K[19-20]。因此，为实现SMA在复杂工况下的功能响应，亟须开发具有宽温域超弹性的合金体系。调控材料的成分、相变特征、进行组织

收稿日期：2025-08-13；**录用日期**：2025-10-24；**网络出版时间**：2025-10-30 10:13

网络出版地址：link.cnki.net/urlid/11.2625.V.20251029.1512.002

基金项目：国家重点研发计划(2022YFB3805701)；国家自然科学基金(52371182,51701052)；黑龙江省自然科学基金(YQ2024E014)；黑龙江省头雁创新团队项目

* **通信作者**. E-mail：xxzhang@hit.edu.cn

引用格式：*钱明芳，尹钊，王桂松，等. 宽温域超弹性形状记忆合金设计优化研究进展[J]. 北京航空航天大学学报，2026，52(10)：3313-3328. Qian M F，Yin Z，Wang G S，et al. Progress in design and optimization of wide-temperature-range superelastic shape memory alloys[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3313-3328 (in Chinese).*

结构设计，已成为构建适应多温区环境功能器件的关键手段。

尽管近年来已有诸多宽温域超弹性 SMA 的研究工作，但目前仍缺乏系统综述不同合金体系在超弹性温度窗口调控方面的最新进展的文献。过去 10 年间，宽温域超弹性 SMA 的研究取得了长足发展，本文旨在系统总结上述研究成果，特别是围绕合金设计理念、组织调控手段、制备工艺路径等方面进行归纳与分析，为未来高性能宽温域超弹性 SMA 的研究提供理论指导与设计思路。需要说明的是，本文以超弹性温度窗口大于 100 K 作为宽温域超弹性 SMA 的评判标准，据此标准筛选了目前研究较为系统、代表性强的合金体系，包括 Ni-Ti 基、Cu 基、Fe 基、Ni 基、Co-Cr-Al-Si、Mg-Sc 及 Ti-Al-Cr 等 SMA，并对其超弹性温度窗口、微观组织结构演化机制与适用工况进行比较分析，力求系统呈现该领域研究进展。

1 SMA 宽温域超弹性产生的途径

马氏体相变是一种典型的一级相变，具有明显的温度和应力滞后，这种特性限制了超弹性在宽温域内的稳定表现。因此，为实现宽温域超弹性功能，有必要采取策略性设计以有效拓宽 SMA 的超弹性温度窗口。

通常而言，SMA 的超弹性温度窗口由 2 个界限决定，如图 1 所示。下限温度由逆马氏体相变结束温度 A_f 控制，当环境温度低于 A_f 时，材料超弹性行为迅速减弱甚至消失，转而表现为形状记忆效应。上限温度则由材料屈服强度 σ_s 决定，由 Clausius-Clapeyron 关系[4] 可知，SIMT 的临界应力 σ_c 随温度升高而增加，而 σ_s 则趋于下降。当二者相等时，即标志着达到超弹性的温度上限。超过此温度，塑性变形（如位错滑移）优先于 SIMT 发生，材料超弹性

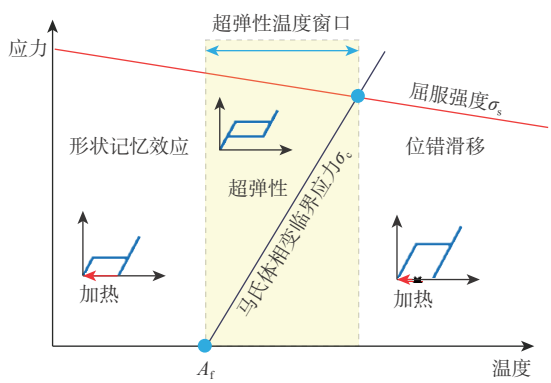


图 1 温度-应力空间中形状记忆效应和超弹性区域示意图（根据文献 [21] 修正）

Fig. 1 Schematic diagram of shape memory effect and superelastic region in temperature-stress space (modified based on Ref. [21])

迅速衰减[8]。因此，为在更高温度下获得超弹性，必须使材料具备高的屈服强度。综上，降低 A_f 的同时提高 σ_s，是实现宽温域超弹性的通用策略之一，适用于大多数 SMA。

另外，近年来发展出以组织调控获得应变玻璃（strain glass, STG）转变的超弹性合金，为拓宽超弹性温度窗口提供了另一种有效路径。STG 起源于合金中缺陷（点缺陷、位错、纳米析出相等）引入的随机应力场，其抑制了合金的一级马氏体相变，使得马氏体相变不能形成长程有序相，而是被“冻结”在局域纳米尺度上。STG 具有以下典型特征：①宏观晶体结构不随温度的变化而变化，体现在变温 X 射线衍射（X-ray diffraction, XRD）衍射峰保持不变和 DSC 相变峰消失；②纳米尺度有序，体现在透射电子显微镜（transmission electron microscope, TEM）下观察到纳米畴的存在；③零场冷却/场冷却（zero-field cooling/field cooling, ZFC/FC）实验中冷却时遍历性的连续损失；④符合 Vogel-Fulcher 关系的储能模量频散[22-25]。STG 合金表现出窄滞后、宽温域的独特超弹性特性[22-25]。其核心设计理念是在奥氏体基体中引入纳米尺度的马氏体胚（纳米畴），这些稳定存在的微结构单元在宽温度范围内可作为马氏体形核核心，在加载与卸载过程中发生可逆性长大与收缩。这种机制不仅显著降低了 σ_c 对温度的敏感性，还有效拓宽了超弹性响应的工作温区。此外，晶粒细化同样有助于降低 σ_c 的温度敏感性[26-27]，并提高屈服强度，从而进一步拓宽超弹性温度窗口。

综上所述，拓宽 SMA 的超弹性温度窗口可通过 3 条路径协同实现：①降低马氏体相变温度；②提升合金屈服强度；③降低 σ_c 对温度的敏感性。上述策略的结合有利于开发出性能优异的宽温域超弹性材料。

2 宽温域超弹性 SMA 研究进展

2.1 Ni-Ti 基宽温域超弹性 SMA

Ni-Ti 基 SMA 是目前研究最为广泛的 SMA 体系之一。为拓宽其超弹性温度窗口，研究者主要采用以下策略：①引入适量点缺陷以降低相变温度，并通过固溶强化提升屈服强度；②通过高密度点缺陷或位错诱导材料转变为 STG 合金，以降低 σ_c 的温度敏感性；③利用共格纳米析出相和位错双重机制强化基体，提高 σ_s；④构建复合结构以协同提高超弹性响应性能。通常，这些手段会组合应用以提高 Ni-Ti 基合金在更宽温域内的超弹性表现。

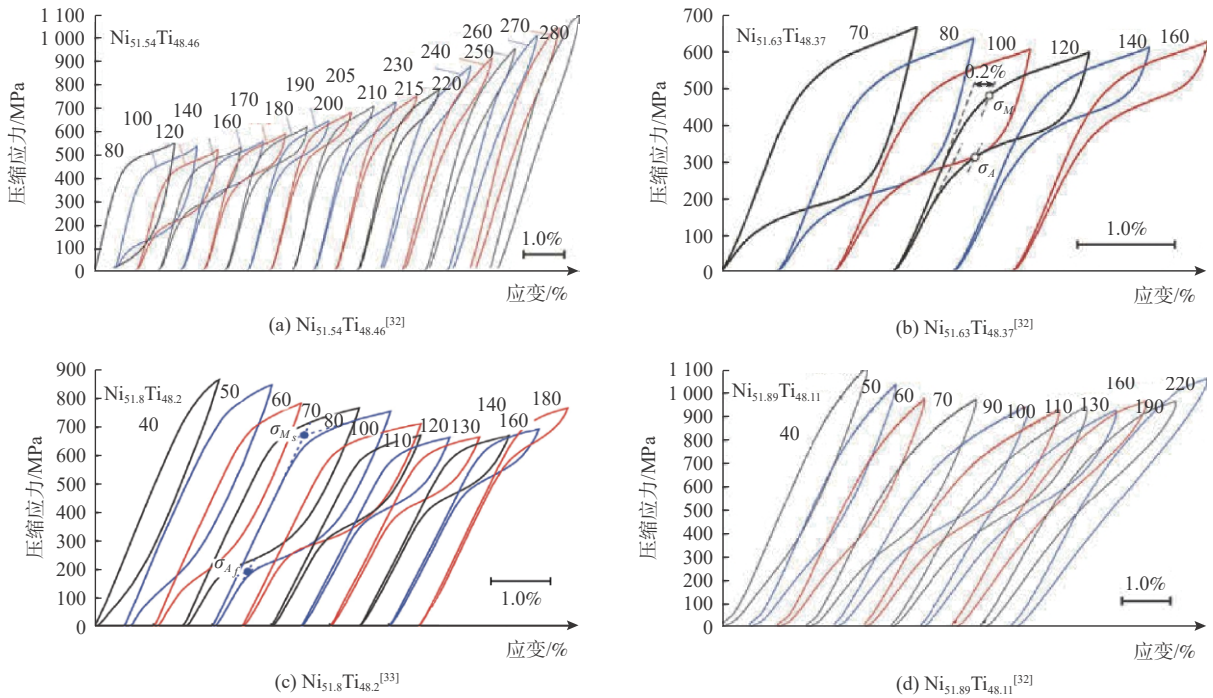
2.1.1 点缺陷掺杂与 STG 机制诱发宽温域超弹性

Ni-Ti 二元合金的相变温度对 Ni 含量极敏感。

在富 Ni 合金中, Ni 的摩尔分数每增加 0.01, 马氏体相变开始温度 M_s 将下降约 100 K[28-29]。然而, 由于富 Ni 合金在 673～773 K 时效处理过程中会析出 Ni_4Ti_3 相, 使基体 Ni 含量下降, 从而导致相变温度再次升高。因此, 为获得具有低相变温度的稳定组织, 通常采用 1 173～1 273 K 高温固溶处理与水淬工艺。这种通过引入点缺陷的方法降低了 Ni-Ti 合金的相变温度, 同时也通过固溶强化效应提高了合金的屈服强度。

当研究者向 Ni-Ti 合金中添加足量 Ni 元素后(Ni 的摩尔分数大于 0.51), 发现了一种新的应变状态, 即 STG[30]。STG 的特征是不发生热致马氏体相变及 B2 基体中存在可长大的马氏体纳米畴。随后, Wang 等[31] 证明, 尽管 STG 本身不会发生自发的长程有序马氏体相变, 但在额外外场(如应力场)的作用下, 仍能够克服局部动态势垒, 从而诱发长程有序的马氏体相变。

Niitsu 等[32-33] 对不同成分高掺杂量富镍 Ni-Ti 固溶态多晶合金进行了深入研究, 通过过量掺杂 Ni 元素使合金处于 STG 状态, 从而获得宽温域超弹性。图 2 为 4 种不同成分 STG 合金在不同温度下的压缩应力-应变曲线, 4 种合金均具有宽温域超弹性特性。图中: 数字表示温度值, 单位为 K。但是, 如图 2(a)所示, 当温度升高至接近室温时, 合金的残余应变随着温度升高而增加, 说明固溶态 Ni-Ti 合金在高温下的强度不够, 温度升高, 位错滑移加剧, 不利于获得较高温度下的超弹性。



(a) $Ni_{51.54}Ti_{48.46}$[32]　(b) $Ni_{51.63}Ti_{48.37}$[32]

(c) $Ni_{51.8}Ti_{48.2}$[33]　(d) $Ni_{51.89}Ti_{48.11}$[32]

图 2　固溶态多晶合金在不同温度下的压缩应力-应变曲线

Fig. 2　Compressive stress-strain curves of solid solution polycrystalline alloys at different temperatures

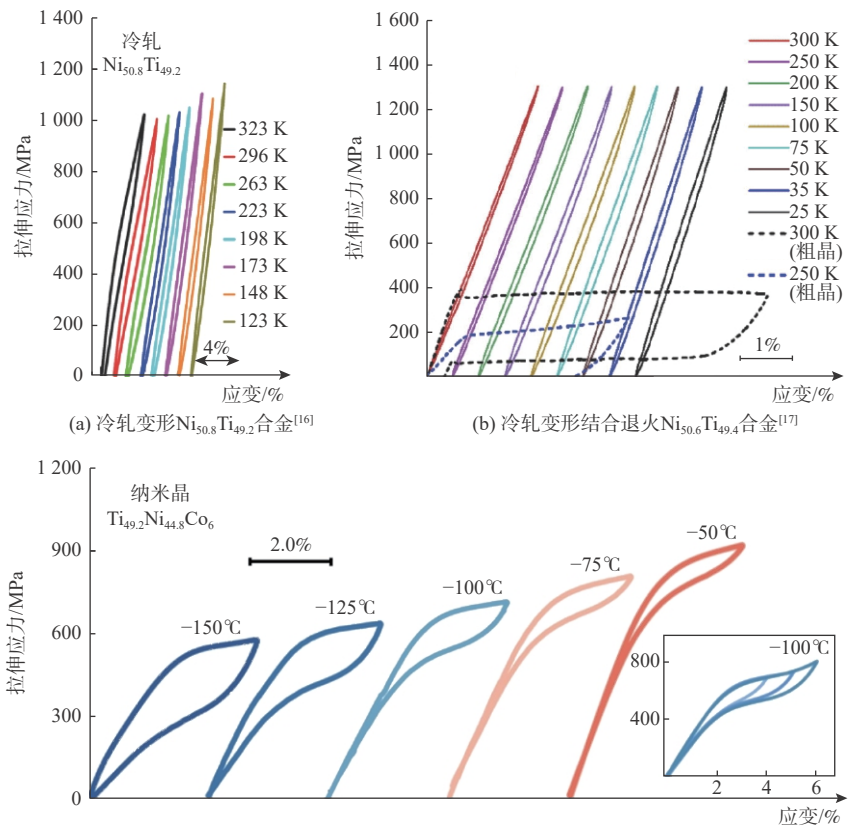
第三元素的添加同样可以降低 SMA 的相变温度, 如 Cu[34-35]、Fe[36-37]、Nb[38]、Sn[39-40]、V[41-42]、Co[43-44]、Cr[43]、Mn[43] 和 Hf[45-46] 等。2016 年, Kimura 等[34] 研究了 $Ti_{50-x}Ni_{40+x}Cu_{10}$ 合金(x = 0～3.0)的马氏体相变行为与超弹性, 发现样品中存在 B2→B19、B19→B19'、B2→I 和 I→B19 多种相变过程, 并且在不发生热致相变的 $Ti_{47.6}Ni_{42.4}Cu_{10}$ 合金中检测到宽温域超弹性特性, 温度窗口为 70～200 K, 但当温度低于 100 K 时, 超弹性应变无法完全恢复。最近, Yin 等[39] 报道了具有宽温域超弹性特性的 $Ni_{49.4}Ti_{50.6-x}Sn_x$(x= 2.5,3)三元合金, 当 Sn 掺杂含量摩尔分数>0.025 时, 合金处于 STG 状态, 其中, $Ni_{49.4}Ti_{47.6}Sn_3$ 合金在 83～298 K 具有约 5.5% 的压缩超弹性特性, 不过, 虽然通过固溶强化效果提高了合金屈服强度, 但高温下合金的超弹性性能仍然不佳。

上述报道中的材料均为固溶态 Ni-Ti 基 SMA, 且在低于室温的环境下展现出良好的宽温域超弹性特性, 但在高温条件下的超弹性表现仍存在明显不足。这主要归因于: 固溶态 Ni-Ti 基合金的屈服强度较低, 当温度升高, 容易发生位错滑移, 从而破坏 SIMT 机制, 导致超弹性性能迅速衰减。因此, 为进一步拓宽 Ni-Ti 基 SMA 的超弹性温度窗口, 亟须通过适当的热-机械处理手段对合金的组织结构进行优化调控, 提升合金的高温屈服强度与结构稳定性, 从而实现更宽温域内的稳定超弹性响应。

2.1.2 位错、析出相和纳米晶强化对宽温域超弹性的影响

冷变形是在Ni-Ti合金中引入位错等缺陷的一种有效方法，Ni-Ti基SMA中常见的冷变形方法有冷轧、冷拉拔、等通道挤压、高压扭转等。在冷变形过后，一般会进行473～723 K的短时退火，退火的目的除了使部分位错回复、恢复超弹性性能外，还有细化晶粒、引入纳米析出相或纳米马氏体畴，提高合金可恢复应变、提高疲劳性能的作用。

Liang等[15-16]研究了不同冷轧变形量$Ni_{50.8}Ti_{49.2}$合金的相变行为与应力应变响应，当冷轧变形量超过27%时，$Ni_{50.8}Ti_{49.2}$合金处于STG状态，基体中存在B19'马氏体纳米畴，在加卸载过程中，纳米畴的连续长大与收缩导致了合金具有近线性超弹性的特征，并且在123～323 K范围内获得了约4%的可恢复应变(见图3(a))，指出可以通过“缺陷工程”在SMA中诱导STG态从而获得优异超弹性。除了室温冷轧变形外，Zhao等[47]研究了液氮温度冷轧$Ni_{50.8}Ti_{49.2}$合金，对比室温冷轧NiTi合金，合金在马氏体状态下更容易产生位错等缺陷，这使得在液氮温度冷轧13%变形量即可获得STG合金，并在4.2～475 K具有宽温域超弹性特性。Ahadi等[17]报道了严重塑性变形$Ni_{50.6}Ti_{49.4}$合金的宽温域超弹性特性，合金组织由约5 nm的奥氏体晶粒、高密度位错及B19'马氏体纳米畴组成，降温过程中合金发生STG转变，最终获得了25～300 K范围的近线性超弹性，可恢复应变为1.5%～2%，如图3(b)所示。



(a) 冷轧变形$Ni_{50.8}Ti_{49.2}$合金[16]

(b) 冷轧变形结合退火$Ni_{50.6}Ti_{49.4}$合金[17]

(c) 冷轧变形结合退火$Ti_{49.2}Ni_{44.8}Co_6$合金[48]

图3 不同合金的拉伸应力-应变曲线

Fig. 3 Tensile stress-strain curves of different alloys

2023年，Dang等[48]报道了具有大可恢复应变(6%)的宽温域超弹性$Ti_{49.2}Ni_{44.8}Co_6$合金，并展现出较高的绝热温变(10.7 K)和循环稳定性，如图3(c)所示。合金通过45%冷轧变形结合673 K/1 h退火处理，最终得到纳米晶、残余位错、纳米Ni_4Ti_3析出相共存的组织结构。纳米晶的晶界、Co原子掺杂和残余位错引起的局部晶格畸变应力场稳定奥氏体母相，降低了相变温度。此外，纳米结构和纳米析出强化了基体，加载过程中有效抑制了位错滑移和形变孪生，使材料在加载与卸载过程中表现出较大的稳定功能响应。

除了冷轧外，还有研究者通过冷拔丝获得具有宽温域超弹性的Ni-Ti基SMA。马之远[49]通过冷拔工艺(变形率56%)制备了非晶Ni-Ti-Nb微丝，随后通过低温退火使非晶合金晶化形成纳米晶，与商用NiTi合金相比，晶粒尺寸为20 nm的NiTiNb微

丝具有超高的 σ_c, 所有温度下的 σ_c 均超过 1.4 GPa, 并且微丝在 77～333 K 的温度范围内具有超弹性。这种特性主要源于纳米尺度成分非均匀对相变热力学、动力学及屈服强度 3 方面的影响。Hu 等[50]设计了一种纳米晶 NiTi/NiTiNbFe 层状复合材料, 在宽温度范围内利用晶格应变匹配和纳米晶强化机制, 通过组合 2 种本身不具有宽温域超弹性的组分, 实现了宽温域超弹性特性(77～413 K), 这种复合机制为制备宽温域超弹性合金提供了另一种可行方法。

2.1.3　相变路径调控宽温域超弹性

一般来说, 673～723 K 退火是冷变形 NiTi 合金获得优异超弹性的典型手段。然而, 近年来还有研究者探究了更低温度(如 473～573 K)退火对合金超弹性特性的影响, 并取得了显著的成果。Liang 等[51]对 40% 冷轧变形 $Ni_{50.8}Ti_{49.2}$ 合金在 473 K 退火 2 h 后发现合金在降温过程中发生了两步 STG 转变, 即 B2→R 相 STG→B19'相 STG, 相较更高温度退火合金(发生一级马氏体相变), 这种两步 STG 转变合金在室温具有更好的可恢复应变与循环稳定性, 并在 119～329 K 具有宽温域超弹性, 这种特性源自 473 K 退火合金中 R 相纳米畴和 B19'纳米畴的同时存在。Qian 等[52]对 45% 冷轧变形 $Ni_{50}Ti_{50}$ 合金在 573 K 短时退火 2～10 min 后获得了一种交叉 STG 合金, 这种合金具有宽的工作温度窗口(93～453 K)及大的可恢复应变(4%～6%), 这种特性源于交叉 STG 合金独特的纳米组织演变路径, 冷却过程中, 这种合金内部发生 STG 的结晶过程, 即 STG 缓慢转变为马氏体, 随着温度的降低, 在 B2 母相基体内逐步出现尺寸在数纳米至数十纳米范围的 B19'马氏体结构, 这一过程贯穿于较宽的温度范围, 促进了大可恢复应变与宽工作温度范围的实现。

2024 年, Xu 等[53]报道了一种具有超高强度(1.8 GPa)、超低模量(10.5 GPa)及宽温域(123～423 K)超弹性的双种子 STG(dual seed STG, DS-STG)$Ni_{50.8}Ti_{49.2}$ 合金, 如图 4 所示, 通过三步法制备获得这种合金, 先将合金拉伸至 50% 伸长率, 随后在 573 K 短时退火 10 min, 获得双交叉 STG(dual crossover STG, DC-STG), 最终将退火的合金再次拉伸 12% 的变形获得 DS-STG, 这种 DS-STG 合金内部具有 R 相和 B19'相 2 类纳米晶种子结构, 该独特的 DS-STG 状态使合金具有类似聚合物的超低模量特性、宽温域超弹性、超过百万次的室温超弹性循环稳定性(5×10^6)。

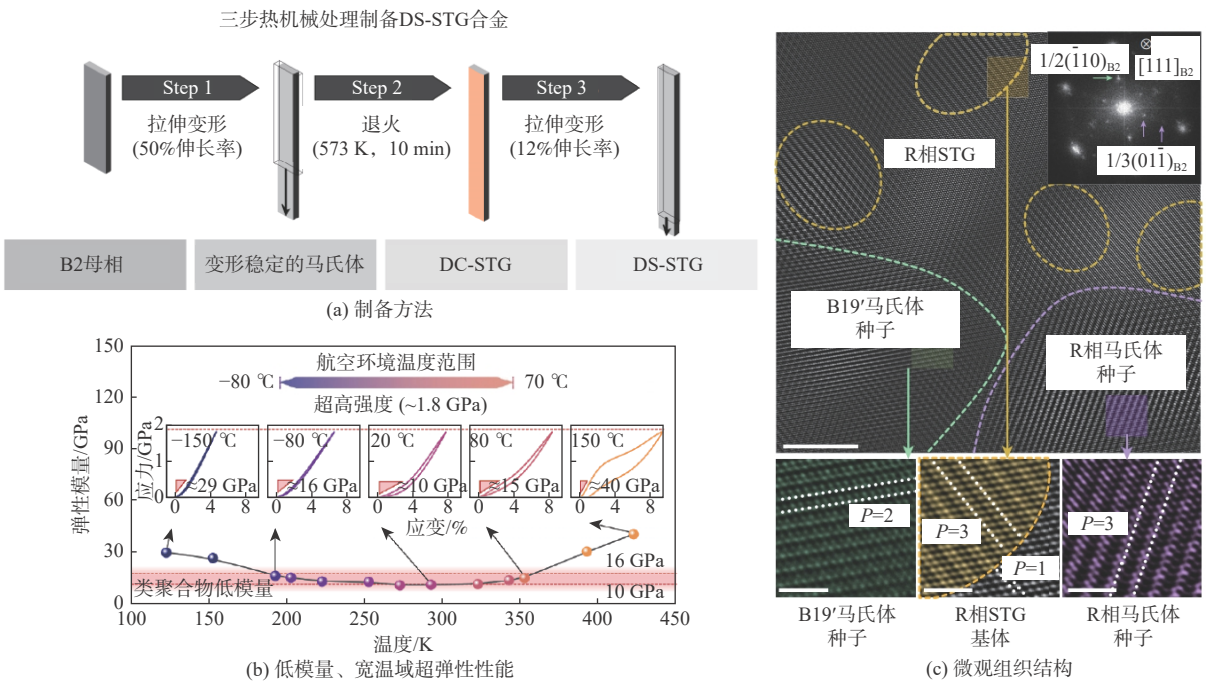


图 4　DS-STG 状态 $Ni_{50.8}Ti_{49.2}$ 合金的制备工艺、性能、组织图[53]

Fig. 4　Preparation process, properties, and microstructure of $Ni_{50.8}Ti_{49.2}$ alloy in DS-STG state[53]

2.1.4　小　结

综上所述, Ni-Ti 基 SMA 可通过多种结构调控手段实现宽温域超弹性, 主要包括点缺陷掺杂、位错引入、析出相强化及纳米晶组织设计等, 其本质是通过对合金微观组织的精准调控, 构建出有利于应力诱导相变和形变可逆性的多级结构, 从而实现热力学稳定性与力学响应性能的协同提升。此外, 基于异质结构的复合强化策略, 通过不同功能相的协同效应, 实现材料性能的“功能叠加”, 也是当前实现宽温域超弹性的一种发展方向。

表1对近年来已报道的Ni-Ti基宽温域超弹性合金的关键参数进行了系统归纳。需要指出的是，对于窄滞后型超弹性合金，由于其应力-应变曲线呈现近线性特征，难以明确界定 σ_c，在 $d\sigma_c/dT$ 斜率统计中不予包含。对于具有凹型 σ_c-T 的Ni-Ti基合金，分别给出了其在低温段与高温段的斜率数值，下文涉及其他合金体系的统计方式也遵循相同标准。

表1 宽温域超弹性Ni-Ti基SMA性能参数

Table 1 Performance parameters of wide temperature range superelastic Ni-Ti based SMA

成分	超弹性温区/K	可逆应变/%	$\frac{d\sigma_c}{dT}$ /(MPa·K^{-1})	加载方式	状态	参考文献
$Ni_{51.89}Ti_{48.11}$	40～220	3.0～4.0	−8.77/2.11	压缩	S/STG	[32]
$Ni_{51.8}Ti_{48.2}$	40～180	2.0～2.5	−2.71/2.50	压缩	S/STG	[33]
$Ni_{51.63}Ti_{48.37}$	70～160	～2.0	1.00	压缩	S/STG	[32]
$Ni_{51.54}Ti_{48.46}$	80～280	～2.0	3.95	压缩	S/STG	[32]
$Ni_{50.8}Ti_{49.2}$	123～323	3.5～4.0		拉伸	CD/STG	[16]
$Ni_{50.8}Ti_{49.2}$	4.2～475	2.5～5.0		拉伸	CD/STG	[47]
$Ni_{50.8}Ti_{49.2}$	119～329	4.0～6.0		拉伸	CD/STG	[51]
$Ni_{50.8}Ti_{49.2}$	123～423	4.8～8.6		拉伸	CD/STG	[53]
$Ni_{50.6}Ti_{49.4}$	25～300	1.75～2.2		拉伸	NG/STG	[17]
$Ni_{50}Ti_{50}$	93～453	4.0～6.0		拉伸	CD/STG	[52]
$Ti_{49.2}Ni_{44.8}Co_6$	123～223	～4.0	3.17	拉伸	CD/NG	[48]
$Ti_{47.6}Ni_{42.4}Cu_{10}$	70～200	～1.5	−3.25/1.96	压缩	S	[34]
$Ni_{49.4}Ti_{48.1}Sn_{2.5}$	113～298	3.6～5.0		压缩	S/STG	[39]
$Ni_{49.4}Ti_{47.6}Sn_3$	83～298	～5.5		压缩	S/STG	[39]
NiTiNb	77～333	～5.0	−2.39/3.19	拉伸	CD/NG	[49]
$Ni_{50.2}Ti_{49.8}$/$Ni_{46}Ti_{47}Nb_3Fe_4$	77～413	3.0～4.0	0.24/2.93	拉伸	CD/NG/CO	[50]

注：S表示固溶态，CD表示冷变形，NG表示纳米晶，CO表示复合材料。

2.2 Fe基宽温域超弹性SMA

SMA的超弹性行为源于其热弹性马氏体相变机制。在冷却过程中，马氏体相逐步扩展长大，而在加热过程中，则可逆地收缩回奥氏体相。这一相变的可逆性是实现可恢复形变的基础。然而，与Ni-Ti基合金不同，大多数Fe基SMA表现出非热弹性的马氏体相变特征，其相变过程通常伴随不可逆的微观结构，导致缺乏可逆应变行为，难以展现典型的超弹性性能。因此，构建热弹性马氏体相变机制是Fe基SMA实现宽温域超弹性的先决条件[54-58]。

研究表明，通过在Fe基合金母相中引入共格析出相，可在局部区域建立稳定的应力场，促使合金发生可逆的热弹性马氏体相变。这一机制为Fe基合金实现超弹性提供了理论基础，并为其宽温域超弹性设计提供了思路。

2.2.1 Fe-Mn-Al-Ni基宽温域超弹性SMA

1）共格析出相对超弹性的影响。对于三元Fe-Mn-Al合金，其本征马氏体相变属于非热弹性/半热弹性类型[59]。当掺杂Ni元素后，多晶Fe-Mn-Al-Ni合金中会形成共格析出相，促使合金发生热弹性马氏体相变，从而展现出超弹性特性[60]。Omori等[60]在 $Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$ 多晶合金中成功实现了超弹性，向Fe-Mn-Al合金中添加Ni后，在A2无序基体中形成了纳米级B2有序析出相，在冷却和加热过程中分别观察到马氏体片的生长与收缩，证实了热弹性马氏体相变的发生。如图5所示，与商用Ni-Ti合金相比（插图），多晶 $Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$ 合金表现出更优异的宽温域超弹性行为，其可恢复应变超过5%。该合金还呈现出一个独特特性：在宽温度范围内，σ_c 具有低的温度依赖性，即合金在77～513 K温度区间内，临界拉伸应力随温度变化的斜率仅为0.53 MPa/K，表明其在深空探测等极端环境应用中具有巨大潜力。

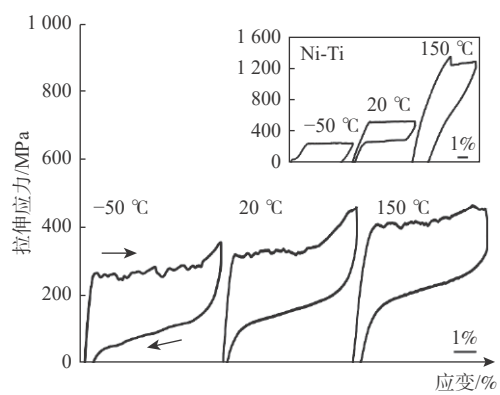


图5 多晶 $Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$ 合金的拉伸应力-应变曲线[60]

Fig. 5 Tensile stress-strain curves of polycrystalline $Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$ alloy[60]

此后, 共格析出相的尺寸、体积分数及成分对超弹性的影响得到了广泛研究。Tseng 等[55]指出析出相对材料超弹性具有显著调控作用, 在 $Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$ 合金中, 时效温度与时间会改变 B2 纳米析出相的尺寸、体积分数及化学成分, 进而影响相变温度、超弹性应变、σ_c 及滞后效应, 研究发现, 473 K 保温 3 h 的最优时效处理可形成 6～10 nm 的第二相析出颗粒, 使材料在室温下获得约 7.2% 的超弹性应变。需要注意, 虽然纳米级共格析出相有利于 Fe-Mn-Al-Ni-X 合金的超弹性性能, 但并非所有析出相都具有积极作用。例如, 在晶界处析出的 γ 相就会对 Fe-Mn-Al-Ni-X 合金的超弹性产生不利影响[61]。

在 Vollmer 等[61]的研究中, 通过用 Ti 替代 Al 元素, 成功降低了 Fe-Mn-Al-Ni 单晶合金对淬火工艺的敏感性, 并提升了其超弹性性能, 研究结果表明, Ti 元素的添加能显著抑制 γ 相的形成, 该相在其前期研究中被发现会形成锯齿状晶界形貌, 会对可逆相变过程产生不利影响[62], 最终, 经 523 K 时效 3 h 后, $Fe_{42}Mn_{34}Al_{15}Ni_{7.5}Ti_{1.5}$ 合金在 123～293 K 温域内展现出 658～719 MPa 的高 σ_c, 且应力滞后约达 400 MPa, 表明其具有优异的阻尼性能, 值得注意的是, 该合金的 σ_c 和应力滞后大小值达到前文所述 473 K 时效 3 h 的 $Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$ 合金[55]的 2 倍和 3 倍。

2020 年, Xia 等[19]报道了一种 Fe-Mn-Al-Cr-Ni 五元超弹性合金, 其在 10～400 K 的宽温域范围内表现出近乎零或负的 σ_c 温度依赖性, 研究者认为这一现象源于极小的熵变 ΔS 及 ΔS 的改变, 与四元合金类似[63], 该五元合金中析出的 β-NiAl 相对热弹性马氏体相变起关键作用。当 Cr 含量(摩尔分数)从 0.000 增至 0.075 时, σ_c 的温度依赖性由正转负。并且添加摩尔分数为 0.03 的 Cr 的合金, 其 σ_c-T 曲线平均斜率仅为 0.085 MPa/K(见图 6, 图中数字表示温度值, 单位为 K), 较传统 Fe-Mn-Al-Ni 合金(0.514 MPa/K)和商用 Ti-Ni 合金(5.87 MPa/K)降低了 1～2 个数量级, 这为深空探测用材料提供了新的选择方案。

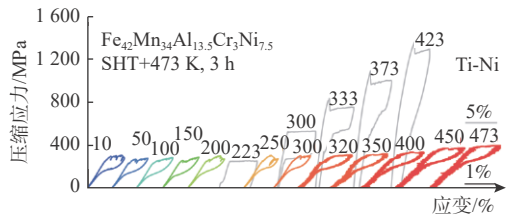
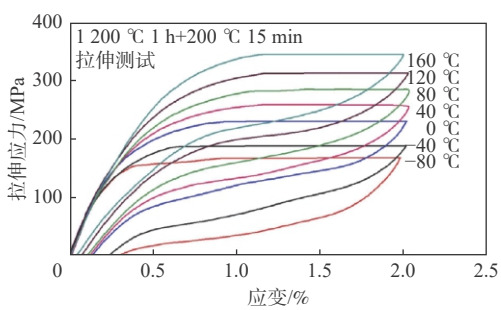


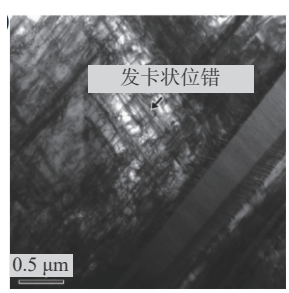
图 6　$Fe_{42}Mn_{34}Al_{13.5}Cr_3Ni_{7.5}$ 单晶的压缩应力-应变曲线[19]

Fig. 6　Compression stress-strain curves of $Fe_{42}Mn_{34}Al_{13.5}Cr_3Ni_{7.5}$ single crystal[19]

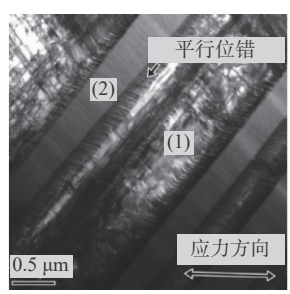
2) 加载方向与加载方式对超弹性的影响。Tseng 等[55,64]研究了 $Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$ 单晶沿 [1 0 0] 方向的拉伸与压缩超弹性, 发现在 193～433 K 的宽温域内, 合金在拉伸(0.51 MPa/K)和压缩(0.41 MPa/K)状态下均表现出较低的 σ_c 温度依赖性(见图 7(a)、图 7(d))。值得注意的是, 压缩测试中的样品展现出比拉伸测试更优的形状恢复性能。这种差异主要源于 2 个因素: ①拉伸样品在奥氏体基体中具有发卡状位错(见图 7(b)), 并且在奥氏体-马氏体界



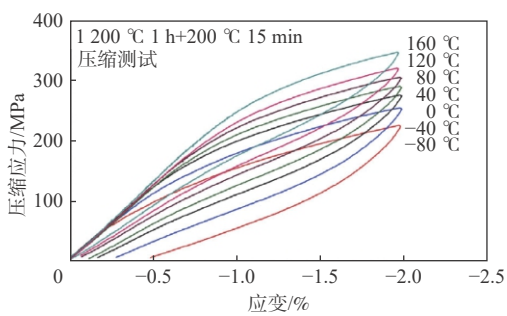
(a) $Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$ <1 0 0> 单晶拉伸应力-应变曲线



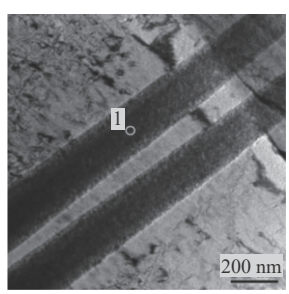
(b) 拉伸测试后的发卡状位错组织



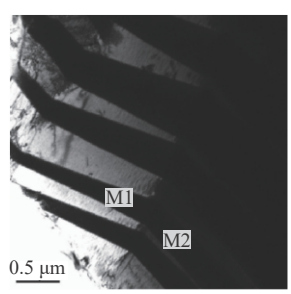
(c) 拉伸测试后的平行位错组织



(d) $Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$ <1 0 0> 单晶压缩应力-应变曲线



(e) 压缩测试后的平整界面组织



(f) 压缩测试后的变体组织

图 7　[1 0 0] 取向 $Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$ 单晶的超弹性响应及加载后的显微组织[64]

Fig. 7　Superelastic response and post-loading microstructures of a [1 0 0]-oriented $Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$ single crystal[64]

面上存在高密度平行位错(见图 7(c));而压缩样品的界面相对平整,奥氏体基体中未见发卡状位错(见图 7(e))。这些位错对马氏体相产生钉扎效应,导致拉伸样品的可恢复性降低。②拉伸样品仅激活一个马氏体变体(见图 7(c)),而压缩样品可激活 2 个变体(见图 7(f))。更多活性马氏体变体的存在有助于更好地协调相变应变,因此,压缩样品表现出更优异的形状恢复性能。

基于前期对 [1 0 0] 晶向 $Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$ 单晶的研究[64],研究者发现能够协调相变应变的活性马氏体变体数量显著受晶体取向或加载方向的影响。为深入探究晶体取向对该单晶合金超弹性的作用机制,随后 Tseng 研究团队又分别通过拉伸[65]和压缩[66]试验系统研究了单晶超弹性行为的取向依赖性,研究发现,在拉伸测试中,[1 0 0] 晶向的可恢复超弹性应变明显小于 [1 2 3] 晶向。这主要归因于 2 个微观结构特征:①奥氏体基体中形成高密度位错;②奥氏体-马氏体界面处产生发卡状位错,这些位错对马氏体相产生钉扎效应。与 [1 2 3] 取向样品相比,[1 0 0] 取向样品中马氏体变体的激活能力较弱,导致其超弹性响应较差。在压缩条件下,[1 0 0] 与 [1 1 1] 晶向的 σ_c 存在显著差异,这源于理论相变应变及其在 Clausius-Clapeyron 关系中作用的差异。其中,[1 0 0] 方向表现出最大的相变应变,而 [1 1 1] 方向的应变最小。对于 [1 2 3] 取向,理论预测应变与实验测得的相变应变相符,表明该取向存在孪晶马氏体形貌。尽管 $Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$ 单晶展现出显著的超弹性应变,但其可恢复性能较差,表现为较大的不可逆应变,这对循环性能产生不利影响。图 8(a)~图 8(c)分别展示了 [0 0 1]、[1 1 1] 和 [1 2 3] 取向固溶处理单晶的应力-应变曲线,图 8(d)则显示了 [1 2 3] 取向时效处理单晶的曲线。可以得出以下结论:①不同晶向的 σ_c 及 σ_c-T 曲线斜率存在显著差异,其中,[1 1 1] 取向的 σ_c 和 σ_c-T 曲线斜率最大;②经 473 K 时效 3 h 后,σ_c 升高,其对温度的依赖性降低(见图 8(e)、图 8(f))[66]。

2.2.2 Fe-Pd 宽温域超弹性 SMA

在 Fe-Pd 合金体系中,研究人员目前仅发现 2 种合金具有热弹性马氏体相变特性:有序结构的 $Fe_{75}Pd_{25}$ 合金和 $Fe_{70}Pd_{30}$ 合金[67]。研究表明,$Fe_{68.8}Pd_{31.2}$ 单晶在应力作用下能实现 6.5% 的大幅可逆应变,并伴随窄滞后特性,这主要归因于单马氏体变体晶格参数的应力调控连续变化[68]。

Xiao 等[69]报道了一种 $Fe_{68.8}Pd_{31.2}$ 单晶合金,该材料在宽温域范围内展现出卓越的超弹性性能,当沿 [0 0 1] 晶向施加压缩应力时,样品在 240~350 K 温区内表现出具有窄应力滞后的超弹性行为,研究者将其归因于母相 [0 0 1] 方向弹性模量的温度依赖性。另一项研究也报道了类似现象[70]。

2.2.3 小 结

综上所述,Fe-Mn-Al-Ni-X 合金体系是当前 Fe 基 SMA 中研究最为深入、应用前景最广的一类。该体系主要通过在无序基体中引入有序共格析出相,实现热弹性马氏体相变,从而赋予材料可逆应变能力。此外,加载方式(如拉伸或压缩)与晶粒取向对合金的超弹性行为也具有显著影响,是实现性能最优化的重要调控手段。虽然近年来也有其他类型 Fe 基宽温域超弹性 SMA 被报道(如 Fe-Pd 合金),但目前仅有少数材料在实验上被证实具备宽温域超弹性行为。因此,现阶段 Fe 基宽温域超弹性合金的研究仍集中在 Fe-Mn-Al-Ni-X 合金体系。表 2 总结了已报道的 Fe 基宽温域超弹性 SMA 的相关信息。

2.3 Cu 基宽温域超弹性 SMA

Cu 基 SMA 是另一类被广泛研究的超弹性材料。相较于 Ni-Ti 基和 Fe 基 SMA,Cu 基合金具有原材料成本低等优势。通过调控合金成分,尤其是 Cu、Al 及其他合金元素的比例,可有效降低其马氏体相变温度,从而拓宽超弹性温度窗口的下限,实现更宽温域范围内的稳定超弹性响应。目前,报道的宽温域超弹性 Cu 基 SMA 主要为 Cu-Al-Mn[8-10,71-72] 合金。

Wang 等[8]对不同成分的柱状晶 Cu-Al-Mn 合金进行了系统研究,其中,$Cu_{70.6}Al_{17.5}Mn_{11.9}$ 合金在 215~393 K 温区内表现出约 10% 的可逆超弹性应变(见图 9(a)),$Cu_{68.6}Al_{18.1}Mn_{13.3}$ 合金展现出更优异的宽温域超弹性,其超弹性温度窗口从 77 K 延伸至 353 K,特别是在 103~313 K 的温区内实现了 10% 的完全可逆应变(见图 9(b)),基于 M_s 与合金成分的实验数据建立了 M_s 与 Al、Mn 含量的线性关系式:$M_s = 1\,890 - 4\,210C_{Al} - 7\,800C_{Mn}$。其中,$C_{Al}$ 和 C_{Mn} 分别为 Al 和 Mn 摩尔分数数值。需要指出的是,由于制备加工条件对合金相变温度具有显著影响,不同研究者得出的线性关系式可能存在差异。

Niitsu 等[72]报道了一种具有竹节状粗晶组织的 $Cu_{68}Al_{17}Mn_{15}$ 合金,该合金在 77~273 K 内不发生热致马氏体相变,但表现出宽温域超弹性特性,值得注意的是,即使在 77 K 的低温测试条件下,该合金仍能实现约 8.5% 的可恢复应变,研究还发现,随着温度降低,σ_c 呈现下降趋势,且变化梯度在低温区域逐渐减小,研究者将其归因于低温下 ΔS 的减小。2018 年,该团队进一步报道了通过循环热处

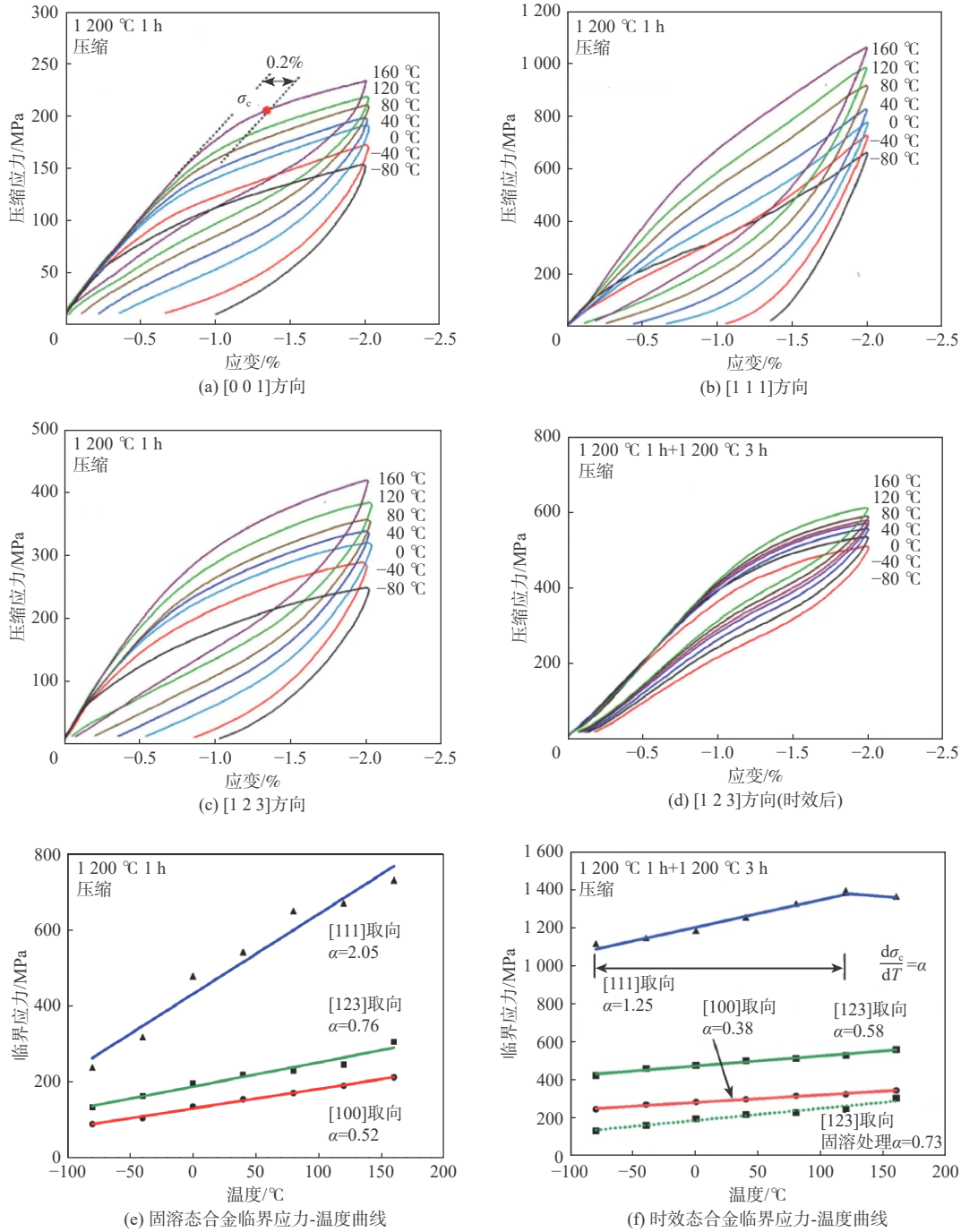


图 8　$Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$ 单晶沿不同晶向的压缩超弹性性能[66]

Fig. 8　Compressive superelastic properties of $Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$ single crystal along different crystal orientations[66]

理制备的 $Cu_{68}Al_{17}Mn_{15}$ 单晶[10] 在 4.2～160 K 具有宽温域超弹性, 所有温度下应力-应变曲线均表现出极小且一致的应力滞后(30 MPa), 并且在 4.2 K 的极低温条件下仍能保持超过 7% 的超弹性应变, 但超弹性应变未能达到理论应变值 9.1%[73], 这种差异可能源于 Mn 含量变化和温度共同导致的晶格参数改变[10]。

此外, Yuan 等[9] 研究了通过 Taylor 拔丝法制备的竹节晶 $Cu_{71.1}Al_{17.2}Mn_{11.7}$ 微丝(直径约 150 μm)的宽温域超弹性特性, 结果表明, 在 223 ～313 K 温区内可实现 8% 的固定应变超弹性, 曲线中出现的多重应力台阶现象, 源于不同取向晶粒诱发马氏体相变所需 σ_c 的差异, 这种独特的微丝结构使得应力滞后(σ_{hys})保持恒定值(～40 MPa), 且不受温度变化和外加应力的影响。

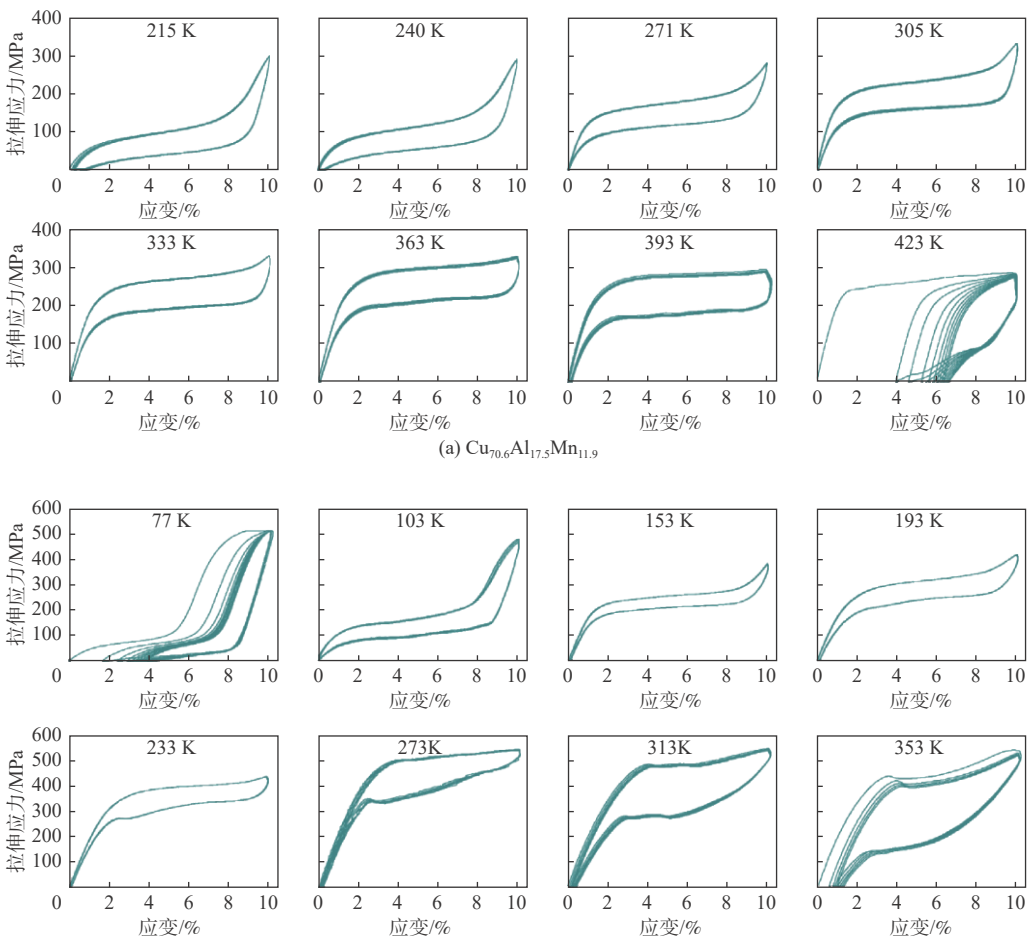
综上所述, 在 Cu 基 SMA 中, Cu-Al-Mn 合金因其宽温域超弹性特性成为研究较为广泛的体系, 其马氏体相变温度主要取决于合金成分。值得注意的是, 微观组织结构对 Cu 基 SMA 的超弹性行为具有显著影响: 具有单晶、柱状晶或竹节状结构的 Cu-Al-Mn 合金均展现出优异的宽温域超弹性性能。虽然 Cu-Al-Mn 合金中已报道大量宽温域超弹

表 2 宽温域超弹性 Fe 基 SMA 性能参数

Table 2 Performance parameters of wide temperature range superelastic Fe-based SMA

成分	超弹性温区/K	可逆应变/%	$\frac{d\sigma_c}{dT}$/(MPa·K^{-1})	加载方式	状态	参考文献
$Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$	223～423	～5.5	0.530	拉伸	PG	[60]
$Fe_{42}Mn_{34}Al_{15}Ni_{7.5}Ti_{1.5}$	123～293	～1.5	0.340	压缩	SG	[61]
$Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$	10～300	～1.3	0.510	压缩	SG	[19]
$Fe_{42}Mn_{34}Al_{13.5}Cr_3Ni_{7.5}$	10～473	1.1～2.3	0.085	压缩	SG	[19]
$Fe_{41}Mn_{34}Al_{12.5}Cr_5Ni_{7.5}$	10～300	1.2～1.4	−0.370	压缩	SG	[19]
$Fe_{38}Mn_{36}Al_{11}Cr_{7.5}Ni_{7.5}$	10～293	～1.2	−0.860	压缩	SG	[19]
$Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$	193～433	1.7～2.0	0.540	拉伸 [1 0 0]	SG	[64]
$Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$	193～433	1.5～2.0	0.410	压缩 [1 0 0]	SG	[64]
$Fe_{43.5}Mn_{34}Al_{15}Ni_{7.5}$	193～433	1.8～2.0	0.580	压缩 [1 2 3]	SG	[66]
$Fe_{68.8}Pd_{31.2}$	240～350	0.6～3.7		压缩 [0 0 1]	SG	[69]

注: PG表示多晶, SG表示单晶。



(a) $Cu_{70.6}Al_{17.5}Mn_{11.9}$

(b) $Cu_{68.6}Al_{18.1}Mn_{13.3}$

图 9 不同温度下合金的拉伸应力-应变曲线[8]

Fig. 9 Tensile stress-strain curves of alloys at different temperatures[8]

性案例, 但其他 Cu 基合金体系的相关研究仍较为有限。表 3 总结了已报道的 Cu 基宽温域超弹性 SMA 的相关信息。

2.4 其他宽温域超弹性 SMA

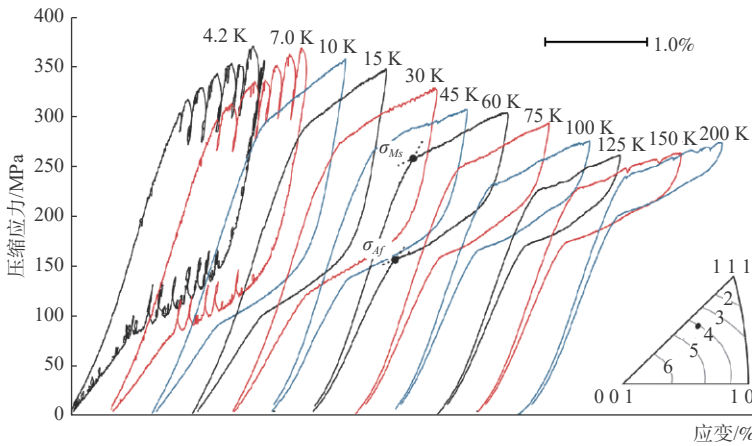
Niitsu 等[74] 报道了 $Ni_{45}Co_5Mn_{36}In_{14}$ 单晶在 200 K 以下低温区间具有完美超弹性特性(见图 10(a)), 发现在 4.2 K 和 7 K 的应力-应变曲线中存在锯齿状波动, 这可归因于动能存储与绝热局部热释放效应, 最重要的是, 研究者发现应力诱导相变与磁场诱导相变的滞回温度变化趋势具有一致性, 证明应力和磁场对马氏体相变的影响具有等价性。同年, Xu 等[75] 报道的铁磁 Heusler 合金 $Co_{51}Cr_{27}Ga_{11}Si_{11}$ 呈

表 3　宽温域超弹性 Cu 基 SMA 性能参数

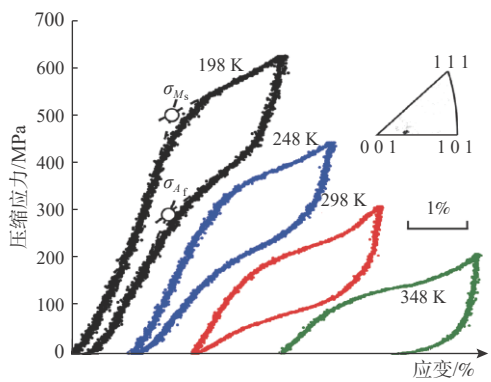
Table 3　Performance parameters of wide temperature range superelastic Cu-based SMAs

成分	超弹性温区/K	可逆应变/%	$\frac{d\sigma_c}{dT}$/(MPa·K^{-1})	加载方式	状态	参考文献
$Cu_{70.6}Al_{17.5}Mn_{11.9}$	215～393	～10.0	1.86	拉伸	CG	[8]
$Cu_{68.8}Al_{17.9}Mn_{13.3}$	77～353	6.0～10.0	2.00	拉伸	CG	[8]
$Cu_{68}Al_{17}Mn_{15}$	77～223	0.8～1.3	2.23	拉伸	BG	[72]
$Cu_{68}Al_{17}Mn_{15}$	4.2～160	～2.0	2.05	拉伸	SG	[10]
$Cu_{71.1}Al_{17.2}Mn_{11.7}$	223～313	～8.0	2.92	拉伸	BG	[9]

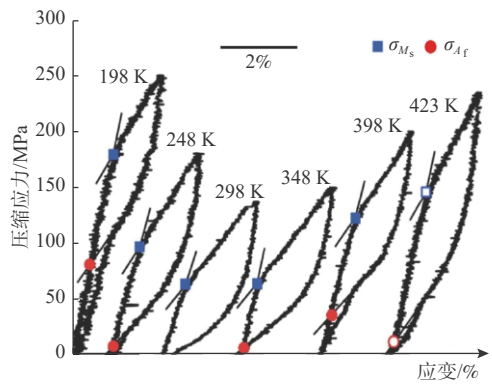
注: BG表示竹节晶织构, CG表示粗晶, SG表示单晶。



(a) 单晶$Ni_{45}Co_5Mn_{36}In_{14}$[74]



(b) 单晶$Co_{51}Cr_{27}Ga_{11}Si_{11}$[75]



(c) 主取向为[0 0 1]的$Co_{55.7}Cr_{23.3}Al_{10.5}Si_{10.5}$伪单晶[76]

图 10　不同温度下合金的压缩应力-应变曲线

Fig. 10　Compressive stress-strain curves of alloys at different temperatures

现出 σ_c 的反常温度依赖性(见图 10(b)), 认为这种异常现象主要源于母相的铁磁转变。后续研究中, Odaira 等[76] 用 Al 替代 Ga, 制备出具有 [0 0 1] 主取向的 $Co_{55.7}Cr_{23.3}Al_{10.5}Si_{10.5}$ 伪单晶, 其超弹性温度窗口扩展至 198～423 K(见图 10(c))。特别值得注意的是, 该合金 σ_c 随温度呈抛物线型变化, 这是由于母相磁效应导致 ΔS 从负值演变为正值所致。

2020 年, Chen 等[77] 报道了一种具有可变钴含量的无滞后超弹性 $Ni_{55-x}Co_xFe_{18}Ga_{27}$ 合金, 该独特性能被命名为“超临界弹性”(supercritical elasticity, SCE), 其中, Co20 成分的纤维样品不仅实现了高达 10% 的应变, 还展现出无滞回超弹性及极低的温度依赖性, 研究者认为, SCE 可能源于连续相变过程, 该现象与有序-无序晶体结构的原子级纠缠态密切相关。

STG 转变现象并非 Ni-Ti 基 SMA 所独有。2022 年, Liu 等[14] 在 Mg-Sc 合金中首次发现了 STG 转变现象, 当 Sc 含量(摩尔分数)为 0.195 时, 合金表现出典型的马氏体相变行为, 然而当 Sc 含量(摩尔分数)增至 0.213 时, 合金在冷却过程中发生 STG 转变, $Mg_{78.7}Sc_{21.3}$ STG 展现出独特的宽温域超弹性特性, 在 123～298 K 温区内呈现近线性的超弹性行为, 且疲劳寿命超过百万次循环, 特别值得注意的是, $Mg_{78.7}Sc_{21.3}$ STG 表现出 Elinvar 效应, 其弹性模

量在宽温域内保持温度不敏感性，研究者认为，该现象源于 STG 转变过程中适度的模量软化与持续弹性硬化相互抵消的作用机制。

最近，Song 等[78]提出了一种 $Ti_{75.25}Al_{20}Cr_{4.75}$ 合金，其兼具超宽温域超弹性（4.2～400 K）、低密度（4.36×10^3 kg/m^3）和高比强度（185×10^3 (Pa·m^3)/kg）的特点，为航天、极端低温与医疗器械提供了强有力的材料候选，在低温区域（<200 K），合金 σ_c 随温度降低而升高，呈现反常趋势，这种行为归因于 B2 母相剪切模量随温度降低而增强，即晶格剪切稳定性增强，从而推迟马氏体形核，研究者通过原位中子衍射解释了超弹性源于应力诱导有序体心立方 B2 相到有序正交 B19 相的马氏体转变，类似于 Ni-Ti-Cu 合金中观察到的马氏体转变[79]。表 4 总结了其他成分宽温域超弹性 SMA 的相关信息。

表 4　其他宽温域超弹性 SMA 性能参数

Table 4　Performance parameters of other wide temperature range superelastic SMA

成分	超弹性温区/K	可逆应变/%	$\frac{d\sigma_c}{dT}$/(MPa·K^{-1})	加载方式	状态	参考文献
$Ni_{45}Co_5Mn_{36}In_{14}$	4.2～200	～2.0	−0.83	压缩	SG	[74]
$Co_{51}Cr_{27}Ga_{11}Si_{11}$	198～298	～3.5	−3.18	压缩	SG	[75]
$Co_{55.7}Cr_{23.3}Al_{10.5}Si_{10.5}$	198～423	～2.5	−1.12/1.09	压缩 [0 0 1]	SG	[76]
$Ni_{35}Co_{20}Fe_{18}Ga_{27}$	123～423	～11.0		拉伸	SG	[77]
$Mg_{78.7}Sc_{21.3}$	123～298	～1.0		拉伸	STG	[14]
$Ti_{75.25}Al_{20}Cr_{4.75}$	4.2～400	～5.7	−2.13/2.49	拉伸	SG	[78]

注：SG表示单晶。

3　结束语

随着深空探测和极端环境工程的不断推进，传统结构材料已难以满足在超低温或温度波动环境下的高可靠性与高适应性需求。宽温域超弹性 SMA 因其可逆应变能力、良好的力学性能和适应性调控潜力，正成为面向未来航空航天、汽车工程等领域的新一代功能结构材料。

在航天探索领域，深空探测任务（如火星和月球探测）对材料提出了严苛挑战。例如，火星表面温度变化范围为 120～290 K，而月球昼夜温差甚至可达 100～400 K。如此极端且频繁波动的热环境，对材料的热稳定性与循环响应提出极高要求。具备宽温域超弹性的 SMA 可用于火星车轮胎柔性骨架，以适应多变地形并提供减震缓冲功能；也可集成于航天器内部的振动控制系统、热调控结构（如智能散热翅片、微致动器）或自修复组件中。在民用汽车方面，宽温域超弹性 SMA 不容易受四季环境温度变化而出现性能恶化，可用于智能车身结构与碰撞吸能元件。

从材料体系角度来看，Ni-Ti 基 SMA 是目前研究最成熟、应用最广的 SMA。该类合金具有良好的热-力性能耦合响应和较高的恢复应变，且加工工艺成熟、适于大规模工业生产。近年来，通过“缺陷工程”“多元合金化”“热-机械处理”等方法，其超弹性温度窗口进一步拓宽，可满足一定范围内的复杂应用需求。然而，其高密度（约 6.5 g/cm^3）在一定程度上限制了其在对质量极为敏感的空间航天器件中的应用潜力。

相比之下，Fe 基与 Cu 基 SMA 在低成本、制备简单等方面具有优势，部分 Fe-Mn-Al-Cr-Ni 单晶合金已实现高达 463 K 的超弹性温度窗口。然而，这类材料往往依赖于单晶或少晶结构，生产工艺复杂、成材率低，加之对微结构控制要求极高，限制了其在大尺寸或复杂构件中的推广应用。近年来兴起的轻质 SMA，如 Mg-Sc 和 Ti-Al-Cr 合金，则因其低密度而备受关注。轻质化特性使其在航空航天、无人机、微型驱动器等领域具备减重优势。然而，Mg-Sc 合金目前的可逆应变较低（约 1%），力学性能尚难满足工程实际负载需求；而 Ti-Al-Cr 类合金大多需单晶态下才能实现宽温超弹性，依赖复杂晶体生长过程，效率低、成本高，同样制约了其应用拓展。

综上，不同 SMA 体系在性能、工艺与应用适配性方面各有优劣。面向实际应用场景，必须综合考虑目标温度窗口、形变恢复能力、材料密度、可加工性、循环寿命及成本等因素，针对性选择最优合金体系。未来，跨体系复合设计（如 NiTi 基与轻质元素掺杂）、增材制造工艺结合微结构调控技术，将可能在实现高性能宽温域 SMA 材料的同时解决其产业化瓶颈。

基于上述讨论，本文展望了开发具有宽温域超弹性 SMA 的潜在途径：

1）轻量化的材料设计与性能优化。为实现宽温域超弹性 SMA 的实用化应用，轻量化与高性能

需同步推进。未来可借助材料基因组工程与机器学习方法，在合金成分设计上进行智能筛选与性能预测，探索基于轻元素（如 Ti、Mg）构建的新型 SMA。同时，可通过多元掺杂策略调控相变温度、抑制位错滑移、强化相界匹配，从而拓宽超弹性温度窗口。此外，“高熵合金”理念的引入也可能赋予合金更优异的力学一致性与结构稳定性，在保持较低密度的前提下实现高强度-高超弹性的协同。

2）结构调控机制创新与多尺度架构设计。微观结构调控是实现宽温域超弹性的核心手段。未来研究可重点围绕“晶界工程”“缺陷工程”与“纳米结构调控”等多尺度设计策略展开。通过构建多级层次结构（如纳米孪晶、纳米析出强化、纳米级成分波动、梯度结构等）同时提高材料的可恢复应变与超弹性温度窗口。此外，异相界面调控技术如半共格匹配、多重应变相互作用等，也可能在提高应变恢复性与疲劳寿命方面发挥关键作用。

3）制备工艺的创新与升级。增材制造为复杂构件与微结构的精准控制提供了可能。未来可进一步探索复合梯度结构构建、制备策略优化、多尺度热处理等工艺，以兼顾构形复杂性与微观性能统一性。

参考文献（References）

[1] Ma J, Karaman I. Expanding the repertoire of shape memory alloys [J]. Science, 2010, 327(5972): 1468-1469.

[2] Zhu J M, Wang D, Gao Y P, et al. Linear-superelastic metals by controlled strain release via nanoscale concentration-gradient engineering[J]. Materials Today, 2020, 33: 17-23.

[3] 李瀚宇, 周震涛, 唐江龙, 等. 形状记忆合金宽温区超弹性的研究进展[J]. 热加工工艺, 2021, 50(17): 1-5.
Li H Y, Zhou Z T, Tang J L, et al. Research progress on superelasticity of shape memory alloy at wide temperature zone[J]. Hot Working Technology, 2021, 50(17): 1-5(in Chinese).

[4] Otsuka K, Ren X. Physical metallurgy of Ti-Ni-based shape memory alloys[J]. Progress in Materials Science, 2005, 50(5): 511-678.

[5] 章涛, 顾嘉俊, 孔永华. 时效处理对 Ni-Ti 形状记忆合金马氏体相变行为与超弹性的影响[J]. 功能材料, 2025, 56(6): 6221-6228.
Zhang T, Gu J J, Kong Y H. Effect of aging treatment on martensitic transformation behavior and superelasticity of Ni-Ti shape memory alloys[J]. Journal of Functional Materials, 2025, 56(6): 6221-6228(in Chinese).

[6] 周知进, 黄雪值, 潘建林, 等. 铁氟龙材料超弹性本构模型拟合效果评估及有限元仿真验证[J]. 塑料科技, 2025, 53(7): 137-142.
Zhou Z J, Huang X Z, Pan J L, et al. Fitting effect evaluation of hyperelastic constitutive model of teflon material and validation by finite element simulation[J]. Plastics Science and Technology, 2025, 53(7): 137-142(in Chinese).

[7] 王文华, 涂春潮, 郭瑞毅, 等. 减振橡胶材料本构模型及其应用研究进展[J]. 材料工程, 2025, 53(9): 50-64.
Wang W H, Tu C C, Guo R Y, et al. Research progress in constitutive models and application of vibration-reduction rubber materials[J]. Journal of Materials Engineering, 2025, 53(9): 50-64(in Chinese).

[8] Wang H, Huang H Y, Su Y J. Tuning the operation temperature window of the elastocaloric effect in Cu-Al-Mn shape memory alloys by composition design[J]. Journal of Alloys and Compounds, 2020, 828: 154265.

[9] Yuan B, Zhong S J, Qian M F, et al. Elastocaloric effect in bamboo-grained $Cu_{71.1}Al_{17.2}Mn_{11.7}$ microwires[J]. Journal of Alloys and Compounds, 2021, 850: 156612.

[10] Niitsu K, Kimura Y, Omori T, et al. Cryogenic superelasticity with large elastocaloric effect[J]. NPG Asia Materials, 2018, 10(1): e457.

[11] Saedi S, Acar E, Raji H, et al. Energy damping in shape memory alloys: a review[J]. Journal of Alloys and Compounds, 2023, 956: 170286.

[12] Patil R A, Rane S B, Kumbhar S B. Investigation of the damping behavior of shape memory alloy-nitinol reinforced composite[J]. Engineering Research Express, 2022, 4(4): 045018.

[13] Casagrande L, Villa E, Nespoli A, et al. Innovative dampers as floor isolation systems for seismically-retrofit multi-storey critical facilities[J]. Engineering Structures, 2019, 201: 109772.

[14] Liu C, Ji Y C, Tang J X, et al. A lightweight strain glass alloy showing nearly temperature-independent low modulus and high strength[J]. Nature Materials, 2022, 21(9): 1003-1007.

[15] Liang Q L, Zhao S S, Liang C X, et al. Strain states and unique properties in cold-rolled TiNi shape memory alloys[J]. Acta Materialia, 2022, 231: 117890.

[16] Liang Q L, Wang D, Zhang J, et al. Novel B19′ strain glass with large recoverable strain[J]. Physical Review Materials, 2017, 1(3): 033608.

[17] Ahadi A, Kawasaki T, Harjo S, et al. Reversible elastocaloric effect at ultra-low temperatures in nanocrystalline shape memory alloys [J]. Acta Materialia, 2019, 165: 109-117.

[18] Leo D J, Weddle C, Naganathan G, et al. Vehicular applications of smart material systems[C]//Proceedings of the SPIE Conference on Smart Structures and Materials 1998: Industrial and Commercial Applications of Smart Structures Technologies. Bellingham: SPIE, 1998: 106.

[19] Xia J, Noguchi Y, Xu X, et al. Iron-based superelastic alloys with near-constant critical stress temperature dependence[J]. Science, 2020, 369(6505): 855-858.

[20] 梁新增, 关慰勉, 王星, 等. 形状记忆合金研究现状及其在航空领域的应用[J]. 稀有金属材料与工程, 2025, 54(6): 1641-1652.
Liang X Z, Guan W M, Wang X, et al. Research progress in shape memory alloys and their applications in aviation field[J]. Rare Metal Materials and Engineering, 2025, 54(6): 1641-1652(in Chinese).

[21] Otsuka K, Shimizu K. Pseudoelasticity and shape memory effects in alloys[J]. International Metals Reviews, 1986, 31(1): 93-114.

[22] Wang D, Ji Y C, Ren X B, et al. Strain glass state, strain glass transition, and controlled strain release[J]. Annual Review of Materials Research, 2022, 52: 159-187.

[23] Liu C, Ji Y C, Ren X B. Strain glass and novel properties[J]. Shape Memory and Superelasticity, 2019, 5(4): 299-312.

[24] 赵新青, 王凯, 吕超, 等. Ni-Ti 基合金的应变玻璃转变及其研究进展[J]. 航空材料学报, 2025, 45(1): 1-14.
Zhao X Q, Wang K, Lü C, et al. Progress in research on strain glass

transition in Ni-Ti based alloys[J]. Journal of Aeronautical Materials, 2025, 45(1): 1-14(in Chinese).

[25] 任帅, 纪元超, 王栋, 等. 应变玻璃超级铁性智能材料的进展与展望[J]. 前瞻科技, 2025, 4(1): 128-138.
Ren S, Ji Y C, Wang D, et al. Progress and prospect of strain glass super ferroic smart materials[J]. Science and Technology Foresinght, 2025, 4(1): 128-138(in Chinese).

[26] Wang T, Ma Z, Rao X, et al. Temperature-dependence of superelastic stress in nanocrystalline NiTi with complete transformation capability[J]. Intermetallics, 2020, 127: 106970.

[27] Ahadi A, Sun Q P. Stress hysteresis and temperature dependence of phase transition stress in nanostructured NiTi: effects of grain size [J]. Applied Physics Letters, 2013, 103(2): 021902.

[28] Tang W J. Thermodynamic study of the low-temperature phase B19′ and the martensitic transformation in near-equiatomic Ti-Ni shape memory alloys[J]. Metallurgical and Materials Transactions A, 1997, 28(3): 537-544.

[29] Frenzel J, George E P, Dlouhy A, et al. Influence of Ni on martensitic phase transformations in NiTi shape memory alloys[J]. Acta Materialia, 2010, 58(9): 3444-3458.

[30] Sarkar S, Ren X B, Otsuka K. Evidence for strain glass in the ferroelastic-martensitic system $Ti_{50-x}Ni_{50+x}$[J]. Physical Review Letters, 2005, 95(20): 205702.

[31] Wang Y, Ren X, Otsuka K, et al. Temperature-stress phase diagram of strain glass $Ti_{48.5}Ni_{51.5}$[J]. Acta Materialia, 2008, 56(12): 2885-2896.

[32] Niitsu K, Kimura Y, Kainuma R. Transformation entropy change and precursor phenomena in Ni-rich Ti-Ni shape memory alloys[J]. Journal of Materials Research, 2017, 32(20): 3822-3830.

[33] Niitsu K, Omori T, Kainuma R. Stress-induced transformation behaviors at low temperatures in Ti-51.8Ni (at. %) shape memory alloy[J]. Applied Physics Letters, 2013, 102(23): 231915.

[34] Kimura Y, Xu X, Niitsu K, et al. Martensitic transformations and superelastic behavior at low temperatures in $Ti_{50-x}Ni_{40+x}Cu_{10}$ shape memory alloys[J]. Materials Transactions, 2016, 57(3): 269-277.

[35] Tan S M, Miyazaki S, Ueki T, et al. Ti-content and annealing temperature dependence of transformation behavior of $Ti_XNi_{(92-X)}$ Cu_8 shape memory alloys[J]. Materials Science and Engineering: A, 1997, 230(1-2): 132-138.

[36] Brachet J C, Olier P, Brun G, et al. Superelasticity and impact properties of two shape memory alloys: $Ti_{50}Ni_{50}$ and $Ti_{50}Ni_{48}Fe_2$[J]. Le Journal de Physique IV, 1997, 7(C5): C5-561-C5-566.

[37] Zhang Y Q, Jiang S Y, Zhu X M, et al. Influence of Fe addition on phase transformation behavior of NiTi shape memory alloy[J]. Transactions of Nonferrous Metals Society of China, 2017, 27(7): 1580-1587.

[38] Piao M, Miyazaki S, Otsuka K, et al. Effects of Nb addition on the microstructure of Ti-Ni alloys[J]. Materials Transactions, JIM, 1992, 33(4): 337-345.

[39] Yin Z, Qian M F, Zhang X X, et al. Strain glass in Sn-doped NiTi shape memory alloys[J]. Applied Physics Letters, 2025, 126(25): 251901.

[40] Young A W, Torgerson T, Ley N A, et al. Effects of Sn addition on NiTi shape memory alloys[J]. Shape Memory and Superelasticity, 2019, 5(1): 125-135.

[41] 贺志荣, 王芳, 王永善, 等. V 和 Cr 对 Ti-Ni 超弹性合金相变和形变特性的影响[J]. 金属学报, 2007, 43(12): 1293-1296.
He Z R, Wang F, Wang Y S, et al. Effects of V and Cr on transformation and deformation characteristics of Ti-Ni superelastic alloy[J]. Acta Metallurgica Sinica, 2007, 43(12): 1293-1296(in Chinese).

[42] Lin H C, Yang C H, Lin M C, et al. Aging effect on a $Ti_{47.25}Ni_{48.75}V_4$ shape memory alloy[J]. Journal of Alloys and Compounds, 2008, 449(1-2): 119-124.

[43] Zhou Y M, Xue D Z, Ding X D, et al. Strain glass in doped $Ti_{50}(Ni_{50-x}D_x)$ (D=Co, Cr, Mn) alloys: implication for the generality of strain glass in defect-containing ferroelastic systems[J]. Acta Materialia, 2010, 58(16): 5433-5442.

[44] Jing R R, Liu F S. The influence of co addition on phase transformation behavior and mechanical properties of TiNi alloys[J]. Chinese Journal of Aeronautics, 2007, 20(2): 153-156.

[45] Benafan O, Bigelow G S, Garg A, et al. Viable low temperature shape memory alloys based on Ni-Ti-Hf formulations[J]. Scripta Materialia, 2019, 164: 115-120.

[46] Umale T, Salas D, Tomes B, et al. The effects of wide range of compositional changes on the martensitic transformation characteristics of NiTiHf shape memory alloys[J]. Scripta Materialia, 2019, 161: 78-83.

[47] Zhao S S, Liang Q L, Su Y T, et al. Cryogenic rolling induces quasi-linear superelasticity with high strength over a wide temperature range in TiNi shape memory alloys[J]. Scripta Materialia, 2024, 243: 115996.

[48] Dang P F, Zhang L, Zhou Y M, et al. Cryogenic superelasticity and elastocaloric effect in a nanostructured Ti-Ni-Co alloy[J]. Scripta Materialia, 2023, 236: 115638.

[49] 马之远. 宽温域高超弹应力 NiTi 基形状记忆合金研究 [D]. 北京: 中国石油大学 (北京), 2021.
Ma Z Y. Study of NiTi based shape memory alloywith high superelastic stress over widetemperature range[D]. Beijing: China University of Petroleum, Beijing, 2021(in Chinese).

[50] Hu X W, Liu Y N, Guo F M, et al. Enhanced superelasticity of nanocrystalline NiTi/NiTiNbFe laminar composite[J]. Journal of Alloys and Compounds, 2021, 853: 157309.

[51] Liang Q L, Wang D, Ding X D, et al. Two-step strain glass transition in NiTi shape memory alloy with unique properties[J]. Materials Research Letters, 2024, 12(9): 678-687.

[52] Qian Y, Wang W J, Wang Y, et al. Crossover strain glass alloy exhibiting large recoverable strain over a wide temperature range[J]. Acta Materialia, 2025, 283: 120533.

[53] Xu Z Z, Ji Y C, Liu C, et al. A polymer-like ultrahigh-strength metal alloy[J]. Nature, 2024, 633(8030): 575-581.

[54] L'Vov V A, Rudenko A A, Chernenko V A, et al. Stress-induced martensitic transformation and superelasticity of alloys: experiment and theory[J]. Materials Transactions, 2005, 46(4): 790-797.

[55] Tseng L W, Ma J, Hornbuckle B C, et al. The effect of precipitates on the superelastic response of [1 0 0] oriented FeMnAlNi single crystals under compression[J]. Acta Materialia, 2015, 97: 234-244.

[56] Li Z X, Zhang Y, Dong K, et al. Research progress of Fe-based superelastic alloys[J]. Crystals, 2022, 12(5): 602.

[57] Tanaka Y, Himuro Y, Kainuma R, et al. Ferrous polycrystalline shape-memory alloy showing huge superelasticity[J]. Science, 2010, 327(5972): 1488-1490.

[58] Sehitoglu H, Efstathiou C, Maier H J, et al. Hysteresis and deformation mechanisms of transforming FeNiCoTi[J]. Mechanics of Materials, 2006, 38(5-6): 538-550.

[59] Ando K, Omori T, Ohnuma I, et al. Ferromagnetic to weak-magnetic transition accompanied by bcc to FCC transformation in Fe-Mn-Al alloy[J]. Applied Physics Letters, 2009, 95(21): 212504.

[60] Omori T, Ando K, Okano M, et al. Superelastic effect in polycrystalline ferrous alloys[J]. Science, 2011, 333(6038): 68-71.

[61] Vollmer M, Krooß P, Karaman I, et al. On the effect of titanium on quenching sensitivity and pseudoelastic response in Fe-Mn-Al-Ni-base shape memory alloy[J]. Scripta Materialia, 2017, 126: 20-23.

[62] Vollmer M, Segel C, Krooß P, et al. On the effect of gamma phase formation on the pseudoelastic performance of polycrystalline Fe-Mn-Al-Ni shape memory alloys[J]. Scripta Materialia, 2015, 108: 23-26.

[63] Omori T, Nagasako M, Okano M, et al. Microstructure and martensitic transformation in the Fe-Mn-Al-Ni shape memory alloy with B2-type coherent fine particles[J]. Applied Physics Letters, 2012, 101(23): 231907.

[64] Tseng L W, Ma J, Wang S J, et al. Superelastic response of a single crystalline FeMnAlNi shape memory alloy under tension and compression[J]. Acta Materialia, 2015, 89: 374-383.

[65] Tseng L W, Ma J, Wang S J, et al. Effects of crystallographic orientation on the superelastic response of FeMnAlNi single crystals[J]. Scripta Materialia, 2016, 116: 147-151.

[66] Tseng L W, Ma J, Chumlyakov Y I, et al. Orientation dependence of superelasticity in FeMnAlNi single crystals under compression[J]. Scripta Materialia, 2019, 166: 48-52.

[67] Maki T, Kobayashi K, Minato M, et al. Thermoelastic martensite in an ausaged Fe-Ni-Ti-Co alloy[J]. Scripta Metallurgica, 1984, 18(10): 1105-1109.

[68] Xiao F, Fukuda T, Kakeshita T. On the physical nature of high reversible strain in Fe-Pd single crystals exhibiting lattice softening[J]. Acta Materialia, 2013, 61(11): 4044-4052.

[69] Xiao F, Fukuda T, Kakeshita T. Significant elastocaloric effect in a Fe-31.2Pd (at. %) single crystal[J]. Applied Physics Letters, 2013, 102(16): 161914.

[70] Xiao F, Fukuda T, Kakeshita T, et al. Elastocaloric effect by a weak first-order transformation associated with lattice softening in an Fe-31.2Pd (at. %) alloy[J]. Acta Materialia, 2015, 87: 8-14.

[71] Trehern W, Ozcan H, Franco B, et al. Exploring thermomechanical functionality of CuAlMn as an extreme low temperature shape memory alloy[J]. Materials Letters, 2022, 308: 131246.

[72] Niitsu K, Omori T, Kainuma R. Superelasticity at low temperatures in Cu-17Al-15Mn (at%) shape memory alloy[J]. Materials Transactions, 2011, 52(8): 1713-1715.

[73] Sutou Y, Omori T, Kainuma R, et al. Enhancement of superelasticity in Cu-Al-Mn-Ni shape-memory alloys by texture control[J]. Metallurgical and Materials Transactions A, 2002, 33(9): 2817-2824.

[74] Niitsu K, Xu X, Umetsu R Y, et al. Stress-induced transformations at low temperatures in a $Ni_{45}Co_5Mn_{36}In_{14}$ metamagnetic shape memory alloy[J]. Applied Physics Letters, 2013, 103(24): 242406.

[75] Xu X, Omori T, Nagasako M, et al. Cooling-induced shape memory effect and inverse temperature dependence of superelastic stress in $Co_2Cr(Ga,Si)$ ferromagnetic Heusler alloys[J]. Applied Physics Letters, 2013, 103(16): 164104.

[76] Odaira T, Xu X, Miyake A, et al. Thermal, magnetic field- and stress-induced transformation in Heusler-type Co-Cr-Al-Si shape memory alloys[J]. Scripta Materialia, 2018, 153: 35-39.

[77] Chen H Y, Wang Y D, Nie Z H, et al. Unprecedented non-hysteretic superelasticity of [0 0 1]-oriented NiCoFeGa single crystals [J]. Nature Materials, 2020, 19(7): 712-718.

[78] Song Y X, Xu S, Sato S, et al. A lightweight shape-memory alloy with superior temperature-fluctuation resistance[J]. Nature, 2025, 638(8052): 965-971.

[79] Wang W J, Ji Y C, Fang M X, et al. Reentrant strain glass transition in Ti-Ni-Cu shape memory alloy[J]. Acta Materialia, 2022, 226: 117618.

Progress in design and optimization of wide-temperature-range superelastic shape memory alloys

QIAN Mingfang[1, 2], YIN Zhao[1], WANG Guisong[1], ZHANG Xuexi[1, 3, *], GENG Lin[1, 3, 4]

(1. School of Materials Science and Engineering, Harbin Institute of Technology, Harbin 150001, China;
2. State Key Laboratory of Precision Welding and Joining of Materials and Structures, Harbin Institute of Technology, Harbin 150001, China;
3. State Key Laboratory of Metal Precision Hot Processing, Harbin Institute of Technology, Harbin 150001, China;
4. Suzhou Research Institute, Harbin Institute of Technology, Suzhou 215104, China)

Abstract: Superelasticity is one of the most important functional characteristics of shape memory alloys, which essentially originates from stress-induced reversible thermoelastic martensitic transformation. The superelastic working temperature range is mainly influenced by the synergistic effect of alloy chemical composition and microstructure. The development status of several shape memory alloy systems with broad temperature range superelastic properties over the last ten years is methodically reviewed in this article, with an emphasis on contrasting the features and variations of various alloys in important performance parameters like recoverable strain, temperature dependence of critical stress, and superelastic temperature window. This article uses comparative analysis to support the design, preparation, and engineering application of novel wide-temperature-range superelastic shape memory alloys theoretically and practically, as well as to offer opportunities for the field's future development.

Keywords: shape memory alloy; superelasticity; wide temperature range; martensitic transformation; microstructure regulation

Received: 2025-08-13; **Accepted**: 2025-10-24; **Published Online**: 2025-10-30 10:13

URL: link.cnki.net/urlid/11.2625.V.20251029.1512.002

Foundation items: National Key Research and Development Program of China (2022YFB3805701); National Natural Science Foundation of China (52371182,51701052); Heilongjiang Provincial Natural Science Foundation of China (YQ2024E014); Heilongjiang Touyan Innovation Team Program

* **Corresponding author**. E-mail: xxzhang@hit.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0584

镁合金半固态触变射铸技术研究进展

姜巨福[1,2,*]，张金泽[1,2]，崔竞波[1,2]，王迎[3]，张晓东[1,2]

(1. 哈尔滨工业大学 材料科学与工程学院，哈尔滨 150001；2. 哈尔滨工业大学 精密热成形全国重点实验室，哈尔滨 150001；
3. 哈尔滨工业大学 机电工程学院，哈尔滨 150001)

摘 要：镁合金因其低密度特性，在交通设施轻量化领域具有重要的战略意义，但传统铸造工艺面临缩孔、气孔及氧化等技术瓶颈，基于半固态加工原理的创新成形工艺——触变射铸技术，为有效突破上述瓶颈提供了可行途径。为此，系统阐述了触变射铸技术的基本原理，揭示了非枝晶结构的形成机制与半固态浆料的制备方法，并重点分析了该技术对AZ系及AM系镁合金综合性能的提升作用。目前，该技术已在消费电子及交通运输载具领域实现规模化应用。依托新一代大吨位成形设备，其应用领域逐步拓展至大型结构件制造领域，在新能源汽车产业中展现出突出的发展潜力。未来研究需重点关注新型半固态触变射铸镁合金的成分设计、成形过程的多物理场耦合模拟及大型构件缺陷控制机制等方面。

关 键 词：镁合金；半固态成形；触变射铸；力学性能；轻量化
中图分类号：TG146.22
文献标志码：A **文章编号**：1001-5965(2026)10-3329-21

镁及其合金作为轻质结构金属，具有低密度、高比强度、优异的减震特性及良好的铸造性能[1-2]，是汽车工业、航空航天、消费电子行业、运动设施及手持器械等领域理想的轻质结构解决方案[3-4]。然而，市场对高性能轻量化制件需求的持续增长，凸显了当前商用镁合金在力学性能与成形性能方面存在的瓶颈。突破这一瓶颈需同步优化合金成分并革新制备工艺[5-6]。现有材料热加工技术主要包括固态塑性成形(如锻造)和熔融凝固成形(如高压压铸)[7-8]。其中，前者因工艺流程复杂、生产周期较长，导致生产效率偏低，应用受到一定限制；后者作为当前的主流工艺，虽适用性广泛，但仍存在缩孔、缩松、充型过程中湍流导致气孔聚集、热应力升高及断裂敏感性增加等问题，制约了镁合金力学性能的进一步提升[9-12]。

半固态加工技术由Flemings教授[13]于20世纪70年代首次提出，其核心在于利用合金在固液相线之间的温度区间，形成固-液两相共存的半固态浆料，浆料中的固相呈现球状或退化枝晶形态，该独特形貌赋予材料触变性，其黏度在剪切力作用下显著降低，从而使得混合物能够高速压射进入模腔并保持层流状态。该技术有效突破了传统镁合金加工工艺的技术瓶颈。经过数十年的发展，已实现从实验室研究向产业化应用的跨越。该技术的优势包括：层流充型模式(雷诺数 Re<100[11])可有效抑制气体卷入，显著降低气孔与缩孔缺陷的倾向；特有的触变行为有助于获得晶粒细小、组织致密的微观结构，从而显著提升材料的力学性能[14]；半固态成形构件可进行后续热处理(如固溶处理和时效强化)，以进一步优化强度、硬度及断裂韧性等关键力学指标[15]。

基于浆料制备与成形过程的不同，半固态加工技术按成形方式可分为流变成形与触变成形2类[4,16-17]。流变成形通过将液态金属冷却至半固态温度区间，并结合机械搅拌、电磁搅拌等调控手

收稿日期：2025-08-21；**录用日期**：2025-10-10；**网络出版时间**：2025-11-06 11:48
网络出版地址：link.cnki.net/urlid/11.2625.V.20251105.1731.005
基金项目：国家自然科学基金(52431004)；东莞市重点领域研发项目(20241200300012)
*通信作者. E-mail：jiangjufu@hit.edu.cn

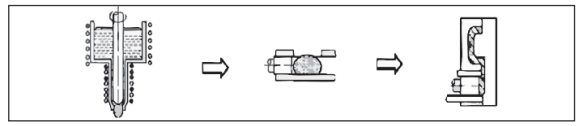
引用格式：姜巨福，张金泽，崔竞波，等. 镁合金半固态触变射铸技术研究进展[J]. 北京航空航天大学学报，2026，52(10)：3329-3349.
Jiang J F，Zhang J Z，Cui J B，et al. Research progress in semi-solid thixomolding technology of magnesium alloys[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3329-3349 (in Chinese).

段，制备具有非枝晶组织的半固态浆料并直接成形。该工艺虽具有流程短、能耗低的优势，但其工业化应用受到限制，主要由于镁合金浆料在转运过程中易发生氧化，导致夹杂缺陷产生，从而降低制件成品率，难以满足规模化生产对稳定性的需求。触变成形通过精确控制半固态浆料的制备与二次加热过程，以维持其非枝晶近球状组织的稳定性。其标准工艺流程包含以下关键步骤：首先，将半固态浆料快速冷却后分割；然后，对预制浆料重新加热至固-液两相区，并实施等温保温处理；最后，利用触变效应实现低应力充型[18-19]。在工业化应用上，触变成形更适合高固相率浆料的加工成形，以及对力学性能要求较高的结构件、厚壁件的精密成形。

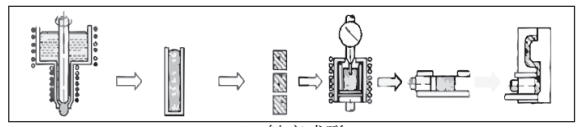
本文聚焦于镁合金触变射铸技术，系统探讨其非枝晶结构形成机制、浆料制备方法、材料性能及产业应用现状。该技术融合了注塑成型的密封环境控制原理与触变成形的快速填充特性[8,20]。浆料在高压下被注射进入精密模具完成成形。工艺全过程严格维持半固态相稳定，既避免了材料完全液化，又规避了传统镁合金熔炼所需的 SF_6 保护气氛（其温室效应为 CO_2 的 24 000 倍[20]），从而在源头上解决了氧化缺陷与环境危害的双重问题[18]。相较于常规半固态加工技术，触变射铸在继承其成形温度低、模具服役周期长、尺寸稳定性优异等共性优势的基础上，创新性地实现了工艺流程的密闭化与连续化，从而显著降低了能耗，并有效抑制了镁合金熔体氧化。

1 半固态浆料与其非枝晶结构形成机制

若按浆料制备与充型方式的差异划分，上述半固态加工可分为流变成形与触变成形 2 类，其工艺原理如图 1 所示。半固态浆料是由熔融液态金属与均匀分布的近球形固相颗粒组成的混合物，在特定温度区间内呈现触变特性与固态特征并存的物相状态[21]。高质量半固态浆料的制备是触变射铸工艺的关键基础，而非枝晶组织的形成与演变机制构成半固态成形技术的核心理论内容。根据现有研究，浆料中固体颗粒的形态可分为 3 类典型特征[22-23]：树枝状型、蔷薇状型及近球状型，如图 2 所示。这些形态差异反映了热力学因素（如冷却速率）与动力学因素（如剪切强度）的协同作用机制。值得关注的是，显微组织从枝晶形态向球形非枝晶结构的转变，能够显著减少铸造缺陷（如偏析、气孔等）的产生。因此，制备具有细小、球形非枝晶组织的浆料，对提升镁合金触变射铸构件的质量具有重要意义[24]。



(a) 流变成形



(b) 触变成形

图 1 半固态加工原理示意图[18]

Fig. 1 Schematic diagram of semi-solid forming[18]

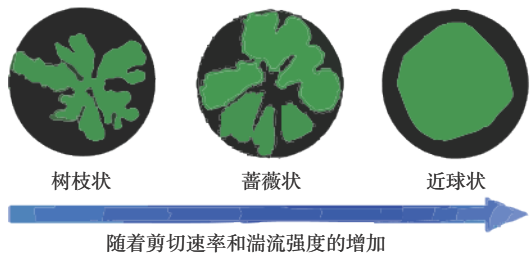


图 2 剪切速率和湍流强度对凝固过程中生长模式影响的示意图[25]

Fig. 2 Schematic diagram of influence of shear rate and turbulence intensity on growth patterns in solidification process[25]

半固态区域内存在 5 类相互耦合的动态机制：晶体形核与生长、溶质扩散、Ostwald 熟化、熔体流动及固相迁移。其中，熔体对流强度与冷却速率的协同作用主导着组织的演化过程；强对流（如电磁搅拌）与低速冷却相结合，可有效促进晶粒球形化，并赋予浆料显著的触变行为。这种特性在经历凝固-部分重熔循环后仍能维持球化组织的“记忆效应”，为半固态金属成形提供了重要的理论依据[23-24]。

目前普遍认为，浆料非枝晶结构的形成机制主要包括枝晶臂破碎[26]、枝晶臂根部重熔[27]及生长控制理论[28]等假说。本节将对这些机制进行综述，并介绍其在实际半固态工艺中的指导作用。

1.1 枝晶臂破碎与枝晶臂根部重熔理论

在早期研究中，Vogel 等[26]提出枝晶臂机械破碎机制，认为搅拌过程通过剪切力导致枝晶臂发生塑性弯曲并最终破裂。随着原位表征、多尺度模拟等新型研究技术的普及，半固态成形领域的诸多研究者采用多尺度分析方法追踪非枝晶组织的动态演变规律，为揭示浆料微观结构演化机制提供了新证据。Ruvalcaba 等[29]利用同步辐射 X 射线显微镜对 Al-Cu 合金（Cu 的质量分数为 0.2）定向凝固过程中的枝晶臂破碎现象进行了定量表征，其观测结果（见图 3）表明，溶质富集液膜导致三级枝晶尖端局部热过冷梯度降低，从而促进枝晶生长，同时，溶质

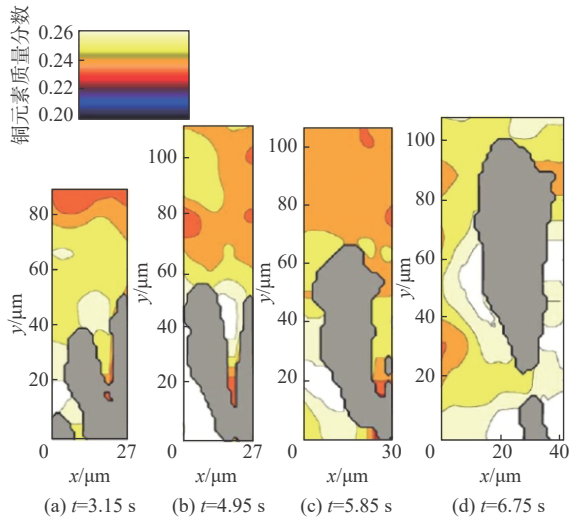


图 3　第 2 个三次枝晶臂脱落过程中溶质富集的演变[29]

Fig. 3　Evolution of solute-enrichment during fragmentation of a second tertiary dendrite arm[29]

的聚集打破平衡状态诱发局部重熔, 最终导致枝晶结构分裂。Hellawell 等[30] 提出, 在缺乏晶粒细化剂的情况下, 等轴晶形成依赖于初生枝晶破碎机制, 该机制源于局部重熔, 且熔体流动特性对扩散过程起决定性作用[31]。此外, 机械剪切效应亦被证实可促进结构解构[29,32]。

当前, 基于枝晶臂破碎及根部重熔机制的理论模型, 尚无法完整阐明熔体搅拌过程中断裂枝晶向蔷薇状或近球状形貌转变的具体路径。尽管如此, 基于上述假说, 已发展出多种可制备具有优良微观结构半固态浆料的工艺方法, 其中, 机械搅拌法和电磁搅拌法 (electromagnetic stirring, EMS) 等已成为当前镁合金半固态浆料工业化制备中的成熟技术。

机械搅拌法通过强制对流效应实现浆料热力学场与成分场的均质化分布, 导致枝晶发生剪切、变形或熔断, 最终生长形成球形颗粒[33]。该技术现已发展形成单螺杆与双螺杆两大工艺体系(见图 4)。凭借操作简便性与经济优势, 机械搅拌法已应用于镁合金半固态浆料的工业化制备。

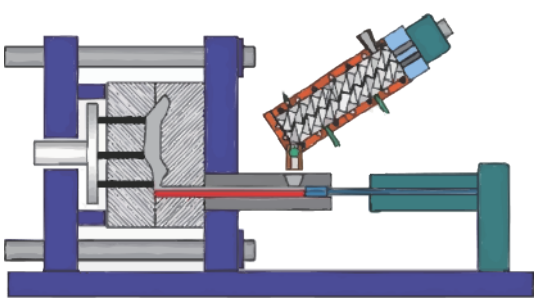


图 4　双螺杆流变成形工艺示意图[34]

Fig. 4　Schematic illustration of twin-screw rheomolding process[34]

电磁搅拌法是一种基于电磁感应原理的非接触式金属液流控技术(见图 5), 通过在熔融金属内部施加交变电磁场, 诱导产生感应电流与洛伦兹力的耦合作用, 从而实现对熔体的定向驱动[35]。该工艺体系具有 2 方面的核心优势: ①可有效维持熔体无氧化加工环境, 降低外源性杂质污染; ②独特的搅拌机制能够维持镁合金表面氧化膜的完整性, 抑制氧化物向熔体内部扩散迁移[36]。作为半固态浆料制备领域的成熟工艺, 电磁搅拌技术兼具工艺可靠性与工程适用性优势, 在规避机械搅拌法固有缺陷(如熔体污染、设备磨损等)的同时, 具有设备结构紧凑且工艺效率高的特点[37]。

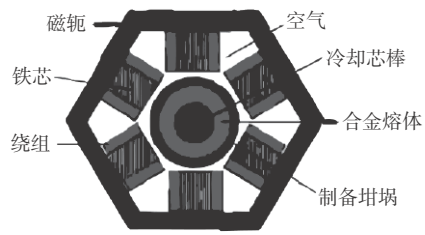


图 5　电磁搅拌装置结构示意图[35]

Fig. 5　Schematic diagram of electromagnetic stirring device[35]

1.2　生长控制理论及其他假说

随着研究的持续深入, 有学者提出球形颗粒源于强制对流条件下液相析出的原生固相颗粒生长, 而非机械破碎等机制, 即生长控制理论。Ji 等[38] 基于双螺杆剪切浆料的研究表明, 高剪切速率与强湍流条件促使 Sn-Pb 合金(Pb 的质量分数为 0.15)的晶体形貌由蔷薇状向等轴状转变, 在低剪切速率下, 层流中悬浮的颗粒发生旋转且枝晶臂产生弯曲, 从而形成玫瑰花状颗粒并持续生长, 而当剪切速率增至足够高时, 固液界面不断更新, 有助于球形颗粒的形成, 这一现象与溶质过饱和度的降低相关。Das 等[39] 利用蒙特卡罗法模拟界面演化机制显示, 在低剪切速率下, 蔷薇状生长由层流和界面失稳共同驱动, 而在高剪切速率下, 湍流引发的渗透性流动消除了边界层的溶质富集, 提高了界面稳定性, 从而促进了球形颗粒的形成。其他学者[31,39-41] 的研究也表明, 浆料中的固体颗粒形态(见图 6)经历了从枝晶结构经蔷薇状中间态向球形结构的转变图中: f_i 为晶粒形状因子。

除主流理论外, 一些学者还提出了其他解释非枝晶组织形成机制的假说。Fan 等[42] 通过实验提出了“有效持续形核”机制, 强对流使晶核球状扩展, 该现象源于对流场对温度场与溶质成分场的均匀化调控, 进而触发周期性循环, 在凝固初期, 组织演变则遵循 Ostwald 熟化机制驱动的粗化规律。张景

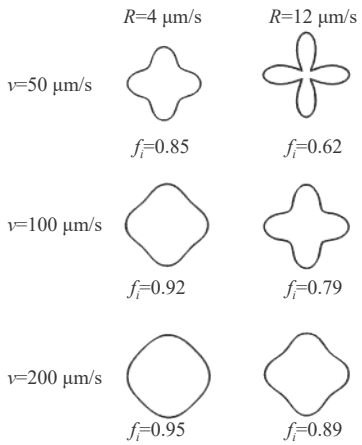


图6 凝固速率 R 与熔体流动速度 v 对晶粒形貌影响的模拟结果[40]

Fig. 6 Simulated grain morphologies under different solidification velocities R and liquid flow velocities v[40]

新团队[43]证实了电磁搅拌条件下“晶粒迁移-混合抑制”耦合机理的存在，该机理通过电磁场引发的熔体剧烈扰动与强制对流，有效提高了异质形核效率，并促进了晶粒多向均衡生长。另有研究者[44-45]提出自由晶核生成机制，认为临界密度的游离晶核通过浓度场叠加可增强固液界面稳定性，从而诱导各向同性生长。基于该自由晶核生成机制，研究者开发了液相线铸造、剪切低温浇注法等新型工艺[46-47]，这些方法可在凝固过程中直接诱导近球状晶的形成，无需依赖后续冷却中的枝晶熔断或晶粒增殖机制来生成非枝晶组织。

1.3 半固态浆料组织形成机制的总结

目前，学术界尚未建立关于半固态浆料非枝晶组织形成机制的统一理论框架。尽管枝晶臂断裂假说难以充分解释枝晶碎片向蔷薇状晶、近球状晶的转变行为，但其在工程应用层面展现出显著价值：该假说表述直观且易于工程转化，可有效指导流变浆料制备工艺的开发。关于塑性弯曲机制的学术争议，其核心焦点在于枝晶臂是否能够发生永久塑性形变。综合本节所引研究成果（见表1）可以看出，各类机制假说并非相互排斥，而是分别从不同角度解释了半固态浆料制备过程中微观组织演变的2个密切联系阶段：晶核的增殖过程与其后续的受限枝晶生长[48]。基于该观点分析相关假说可见，枝晶臂破碎与根部重熔理论有效阐释了形核强化机制，而生长受限机制和其他假说则为后续组织演变提供了更合理的理论支撑。

表1 半固态浆料组织形成机制的主要假说

Table 1 Major hypotheses on formation mechanism of semi-solid slurries microstructure

假说	特点及适用范围	主导因素	对实际生产的指导
枝晶臂破碎与枝晶臂根部重熔理论	应用广泛，解释了晶核增殖过程的机制，对实际生产有重要指导意义，但尚无法解释枝晶到蔷薇状晶、近球状晶的转变路径	破碎枝晶	机械搅拌法、电磁搅拌法等成熟制浆工艺
生长控制理论	解释了枝晶在剪切作用下，向蔷薇状晶、近球状晶的转变机制	剪切速率	双螺杆挤出机等
有效持续形核机制	半固态浆料微观结构主要依赖于湍流而不是剪切速率，湍流的作用提高了浆料的形核速率	形核速率	双螺杆剪切制浆、流变压铸等工艺
自由晶核生成机制	利用临界密度的游离晶核在凝固过程中直接诱导近球状晶形核	游离晶核	液相线铸造、剪切低温浇注法等新型工艺

2 触变射铸的工艺特点

在金属材料中，镁合金因加工过程中显著的氧化敏感性，其半固态浆料制备面临特殊挑战。触变射铸成形工艺创新性地融合了塑料注射成型与金属半固态加工技术优势，在单一工序中连续完成浆料制备与构件成形，有效解决镁合金氧化问题并显著缩短工艺流程，为大规模工业化生产提供创新解决方案。因此，该工艺不仅可在最终构件力学性能层面与流变压铸、高压压铸等方法进行对比，也可在浆料制备层面，与传统方法就浆料的微观组织参数（如晶粒尺寸及形状因子）进行比对。本节从传统触变成形制浆方法出发，引入触变射铸工艺的基本特性，系统阐述该工艺对传统方法在制浆方面的继承与创新，并重点从晶粒尺寸、形状因子等微观组织特征角度进行对比分析。

2.1 触变成形工艺常用制浆方法

应变诱发熔化活化 (strain-induced melt activation, SIMA) 技术由 Young 等[49]首次提出，该技术利用机械储能调控相变以实现显微组织重构，可在显著降低设备成本的同时优化合金组织（见图7）。与其他制浆工艺相比，SIMA 法在高固相分数半固态浆料制备中具有优势，是触变成形工艺常用的成熟制浆工艺。

SIMA 工艺主要包括2个阶段：第1阶段对铸态合金进行预变形处理，破碎铸态粗大枝晶结构，并在合金内部储存变形应变能；第2阶段为后续的半固态等温处理环节，依托预变形储存的应变能驱动动态再结晶，促使晶粒球化，最终制备得到具备良好触变性的半固态浆料[49]。在动力学层面，释放的应变

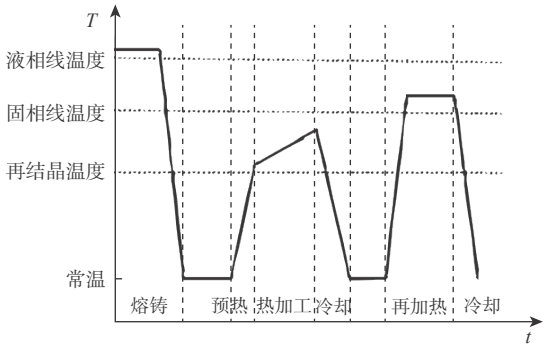


图 7　SIMA 法工艺路线图[50]

Fig. 7　Process map of SIMA[50]

储能驱动再结晶过程, 而界面能主导晶粒球化过程, 最终获得组织均匀、晶粒细小的球状显微结构[51]。SIMA 工艺具有工艺纯净度高(不会污染浆料、无需添加晶粒细化剂)、设备兼容性强(可利用常规热机械加工设备)、固相分数可控性优异(可通过温度参数调节)及经济性优越等优势, 是一种可自动化的半固态制浆工艺。

SIMA 工艺的核心机理在于通过引入临界应变能实现铸态组织重构。该机理使得工艺路径得以简化: 省略冷变形环节, 仅需在再结晶温度以上实施大应变热加工变形, 随后进行半固态重熔即可获得球状组织。后续发展的再结晶重熔 (recrystallisation and partial melting, RAP) 法及新 SIMA 法等, 本质上均属于 SIMA 法体系的工艺改良。

RAP 法由谢菲尔德大学 Kirkwood 团队[52]提出。该方法借鉴了 SIMA 法的基本原理, 但通过在亚再结晶温度以下实施预先温变形, 利用回复再结晶过程引入应变能; 同时, 通过协同调控等温参数(如温度与保温时间), 优化了组织球化效率, 降低了能耗, 并简化了工艺流程[53]。Jiang 等[54]对比了 RAP 与 SIMA 法处理 7075 铝合金的效果, 发现 2 种方法均涉及晶粒合并生长与 Ostwald 熟化的协同作用, 但 RAP 的起始球化温度更低(仅为 590℃, SIMA 法需达到 600℃), 球化效果更显著, 所得晶粒也相对更粗大。

然而, 由于镁合金的密排六方结构塑性变形能力较差, 传统 SIMA 法应用于镁合金时常导致晶粒细化不足和球化效果欠佳[5]。针对此问题, 哈尔滨工业大学罗守靖、姜巨福[55-57]等提出了改进型 SIMA 法, 即新 SIMA 法, 其核心创新在于将剧烈塑性变形技术与传统 SIMA 方案相整合[58-59]。图 8 展示了该方法制备的浆料显微组织。镁合金经过等径角挤压 (equal channel angular pressing, ECAE) 预变形, 局部重熔后可获得理想的半固态浆料。新 SIMA 工艺中, 随着等温温度提高, 浆料固相分数降低, 固相颗粒粗化和球化速率提高; 随着等温时间延长, 浆料固相分数逐渐降低, 固相颗粒尺寸增大。该工艺有效实现了微观组织的细化与球化程度的提升[60-61]。

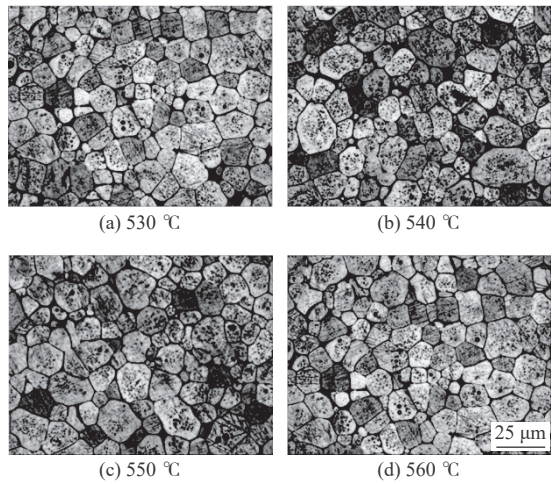


图 8　采用新 SIMA 法制备的 AZ80 合金半固态浆料在不同温度下的显微组织[62]

Fig. 8　Microstructure of AZ80 alloy semisolid billet fabricated by new SIMA at different temperatures[62]

与传统 SIMA 法制备的浆料相比, 新 SIMA 法制备的浆料中, 固相颗粒显著细化。随着保温时间的延长, 组织遵循 Ostwald 熟化机制发生粗化, 晶粒尺寸呈现先增大后减小的变化趋势, 形状因子趋近于 1[63-64]。然而, 新 SIMA 法仍面临一定局限性, 如 ECAE、多向锻造等预变形工艺存在成本高、效率低的问题, 限制了该技术在大规模工业化生产中的应用。综上所述, 新 SIMA 法可被视为一种有潜力的优化工艺, 用于制备高性能镁合金半固态浆料[55]。

2.2　触变射铸工艺介绍

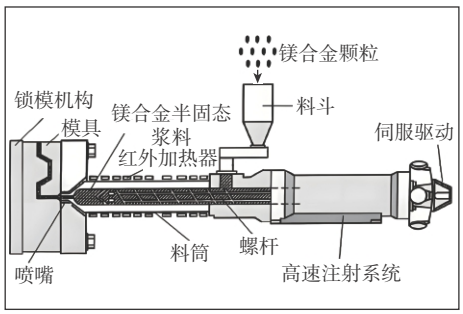
触变射铸作为半固态金属加工领域的革新性工艺, 其创新性在于将浆料制备与成形工艺进行全流程集成化设计, 显著提升了工艺兼容性与过程稳定性[65]。该工艺采用经机械破碎或快速凝固工艺预处理的专用镁合金颗粒作为原材料, 依托精密温度场调控系统, 在模具浇口区精确维持半固态浆料的流变状态, 使其固相分数可在较宽范围内灵活调控。

相较于传统半固态加工中浆料制备与成形工序分离的模式, 触变射铸工艺的原位浆料制备与同步成形一体化协同机制, 不仅有效避免了浆料转移环节可能引发的热力学不稳定性和氧化污染风险, 还赋予最终制件极低的孔隙率特征, 并通过优化微观组织均匀性显著提升了镁合金压铸件的综合力学性能[2,66-67]。

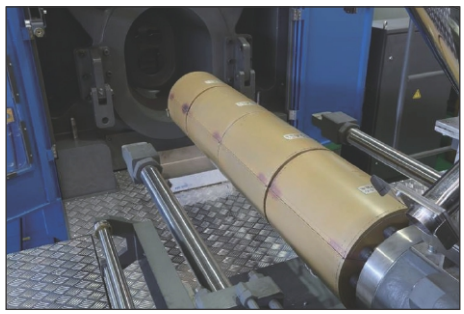
2.2.1　镁合金触变射铸的特性

如图 9 所示, 镁合金触变射铸设备的运行流程

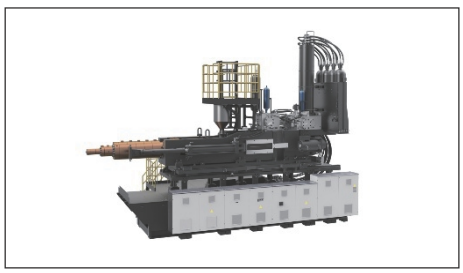
如下：常温镁合金颗粒或加工碎料经计量装置输送至温控料筒后端。料筒采用多区段温度控制，通常通过安装于料筒外侧的红外加热装置实现。金属碎屑在料筒中受热并经历超耐热高强材质螺杆剪切处理，达到精确控制的半固态相。初始阶段，金属浆料中的固相组分呈不规则几何形态；在螺杆旋转产生的机械剪切力作用下，这些组分经历连续的颗粒间相互作用，最终形成几何形态均匀的球形颗粒。这种颗粒形态的优化显著改善了浆料的流变特性，尤其是流动性能。最终，经处理的浆料经过自密封喷嘴加压注射进入模具型腔，实现零件的近净成形[8,18,68]。



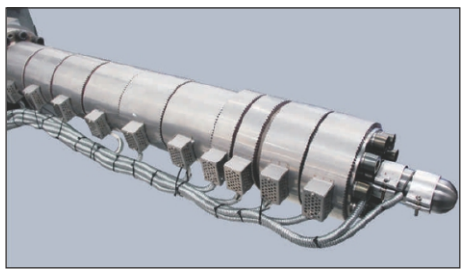
(a) 触变射铸成形设备整体示意图



(b) 红外加热器



(c) 超耐热高强材质螺杆



(d) 热作件料管组

图 9 镁合金触变射铸设备及其关键部件[2,69-70]

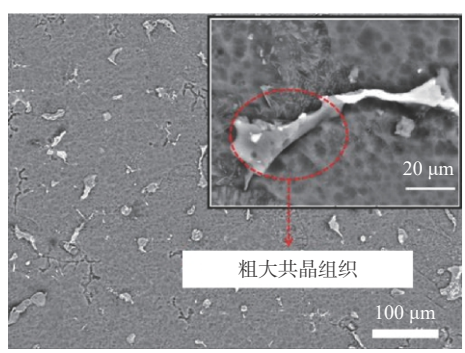
Fig. 9 Magnesium alloy thixomolding equipment and its key components[2,69-70]

Zhang 等[48]研究了 AZ91D 镁合金触变射铸过程中的微观结构与工艺参数之间的关联机制。触变注射成形工艺的关键参数包括料筒温度、螺杆转速、注射速度及模具温度，这些参数显著影响半固态浆料的充型行为与凝固过程，并直接决定最终制件的质量与力学性能，实验结果表明，随着料筒温度的升高，固相颗粒尺寸呈现减小趋势，但其形态特征基本保持不变，这种颗粒尺寸的细化效应与形态均匀性的提升均与螺杆转速呈正相关，该现象主要源于溶质传输动力学的增强效应，研究进一步证实，较低的注射速度和模具温度通过促进快速散热过程主导微观结构的细化。

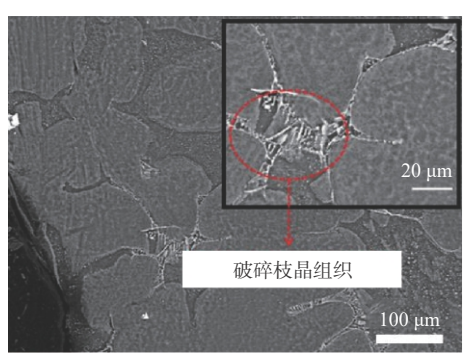
值得注意的是，不合理的工艺参数组合也会导致触变注射成形制件出现缩孔、气孔、撕裂等缺陷[71-73]。具体而言，高固相分数浆料在较低温度下注射时易产生更多缩孔；而在较高注射速度下成形时，固相分数较高的浆料会出现更多气孔，固相分数较低的浆料则容易出现撕裂缺陷。

2.2.2 触变射铸与传统半固态制浆法的对比

以 AZ91D 镁合金在机械加工切削处理前后为例（见图 10），在切削过程中，合金因刀面的挤压和摩擦作用而发生显著变形。随着应力的增加，



(a) 切削前



(b) 切削后

图 10 镁合金切削前后显微组织[74]

Fig. 10 Microstructure of magnesium alloy before and after cutting[74]

变形孪晶的数量上升; 在合金的显微组织中, 晶界因切削力作用而发生断裂, 导致 $Mg_{17}Al_{12}$ 相在晶界区域被挤压并破碎, 触变射铸工艺通过此过程向切削镁合金中引入了变形能。该工艺在制浆层面与传统的触变加工工艺存在一定的继承性, 但其创新性地在同一工序中完成浆料制备与构件成形, 使其能够获得组织更均匀、晶粒更细小的球状显微结构。

要实现组织与性能的精准关联, 必须对微观组织的几何特征进行系统解析与量化表征。在半固态构件显微分析中, 区别于传统铸件的枝晶臂间距测量方法, 其关键评价指标[75]包含表征固相晶粒尺寸的等效直径 d、反映形貌复杂度的形状因子 F_s, 公式为

$$d = \frac{1}{N}\sum_{i=1}^{N}\sqrt{\frac{4A_i}{\pi}} \tag{1}$$

$$F_s = \frac{1}{N}\sum_{i=1}^{N}\frac{P_i^2}{4\pi A_i} \tag{2}$$

式中: A_i为颗粒 i 的面积; N 为颗粒数量; P_i为颗粒 i 的周长。

固相的形态及尺寸等参数对浆料质量具有决定性影响, 当平均晶粒尺寸优化至 100 μm 以下时, 半固态浆料可获得满足流变成形要求的优异流变特性[76]。表 2 对比了触变射铸与传统半固态制浆法的固相晶粒尺寸及形状因子的指标。

相较于多数工程材料, 镁合金具有较高的 Hall-Petch 系数值, 表明其力学性能对晶粒尺寸的变化

表 2　国内外镁合金触变射铸与传统半固态制浆法的组织参数对比

Table 2　Comparison of microstructural parameters between domestic and international magnesium alloy thixomolding and traditional semi-solid slurry processing methods

镁合金成分	固相晶粒尺寸/μm	形状因子	工艺	参考文献
AZ91D	32.0	0.72	触变射铸	[67]
AM60	50.0		触变射铸	[77]
AZ91D	66.6		触变射铸	[78]
AZ91D	30.0		触变射铸	[79]
AZ91	10.5		触变射铸	[80]
AZ91	50.0		触变射铸	[81]
AZ91D	70.0		触变射铸	[82]
AZ91D	45.3	0.65	触变射铸	[83]
AZ61	60.0	0.70	半固态等温处理	[53]
AM60B	200.6	0.71	半固态等温处理	[63]
AZ80	62.0	0.71	半固态等温处理	[84]
Mg-10Gd-3Y-0.5Zr	90.0	0.70	半固态等温处理	[85]
Mg-3Nd-0.2Zn-0.4Zr	46.0	0.59	电磁搅拌	[86]
AZ91D	77.2	0.59	气泡诱导法	[87]
AZ91-2Ca	50.0	0.70	气泡诱导法	[88]
Mg-2.5Gd-1Zn	180.0	0.29	低频电磁搅拌	[37]
Mg-Gd-Zn	501.0	0.30	低频电磁搅拌	[89]
Mg-6Zn-1.4Y	32.0	0.76	超声振动法	[90]
Mg-11Gd-3Y-1Zn-0.2Zr	63.0		旋转热焓平衡装置法	[91]
AZ91	50.0	0.81	剪切低温浇注法	[92]
AZ91D	45.3	0.85	内部快速冷却搅拌工艺	[93]
AM60	65.0		自孕育法	[94]
AM60	100.0	0.70	再结晶重熔	[53]
Mg-8.20Gd-4.48Y-3.34Zn-0.36Zr	62.0		SIMA法	[95]
AZ80	12.5	0.89	新SIMA法	[62]
AM60B	15.1	0.88	新SIMA法	[63]
AZ80	15.6	0.87	新SIMA法	[96]

更为敏感；因此，开展镁合金晶粒细化工艺研究具有重要的应用价值[97-98]。表 2 所示数据表明，触变射铸工艺制备的构件具有晶粒尺寸细小且球化程度高(形状因子趋近于 1)的微观组织，这一显微特性是该技术能够成功制备具有复杂几何结构及薄壁特征的高性能铸件的关键因素。

3 触变射铸镁合金的组织与性能

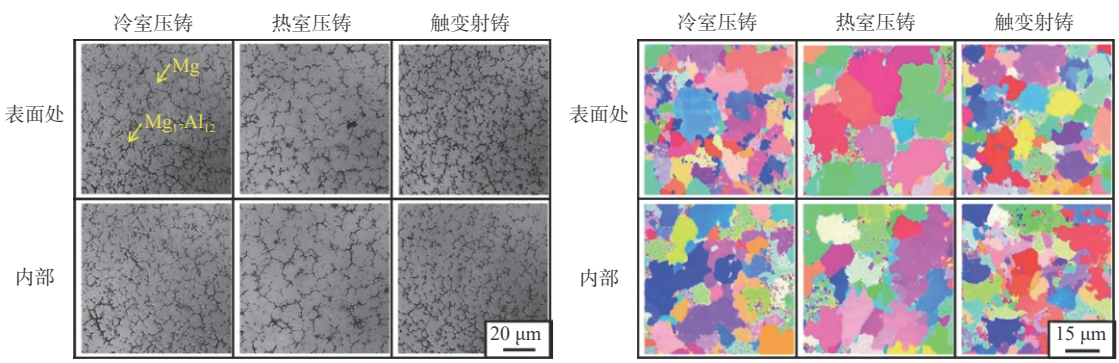
基于触变射铸工艺所诱发的独特微观组织演变，通过该工艺制备的镁合金在晶粒尺寸上相较于传统铸造产品显著细化，综合性能也得到明显提升。这种改善不仅体现在显微组织和力学性能方面，还在耐腐蚀性能、疲劳寿命及蠕变性能等多个维度指标上均有显著表现[20]。在当前的工业领域应用中，镁铝系商业镁合金仍占据主导地位。该类合金具备较宽的半固态温度区间、良好的充型能力、适中的强度及可接受的耐腐蚀性能[87]。其中，AZ91D 和 AM60B 是目前应用最为成熟的 2 种半固态成形镁合金材料。

如前文所述，与当前成熟的镁合金高压压铸相比，触变射铸镁合金凭借其抑制卷气的层流充型模式和优良的充型能力，形成了致密均匀的微观组织结构，有效消除了皮下气孔和密集缩孔缩松等缺陷。这一特点使得该工艺所制备的构件能够通过固溶时效热处理进一步显著提升力学性能[15]。近年来，为克服传统压铸工艺的局限性，发展出了如真空压铸、充氧压铸等特殊工艺，但这些方法仍普遍存在设备投入高、过程控制复杂和生产效率偏低等问题[99-100]。本节以 AZ 系和 AM 系镁合金为主要研究对象，以成熟的高压压铸工艺为主要参考对象，系统分析触变射铸工艺对镁合金组织与性能的影响及其内在机制，并对当前适用于该工艺的新型镁合金体系的学术研究进展进行综述。

3.1 触变射铸镁合金的组织

基于触变射铸工艺的独特属性，该工艺能够获得比传统铸造工艺更细小的晶粒尺寸，同时相较于高压压铸具有更低的孔隙率。此外，触变射铸成形件不会出现类似挤压或锻造成形件中常见的晶体学织构[101]。在触变射铸工艺的商业应用中，AZ 系镁合金应用最为广泛。其中，AZ91D 作为典型代表，凭借其卓越的铸造性能、力学特性、可以接受的耐蚀性，以及高达 171℃ 的凝固范围和显著较低的液相线，展现出优异的半固态成形性能[87,102]。

Okayasu 和 Fukui[20] 分析了触变射铸技术制备的 AZ91 合金的力学性能与断裂机制，并与冷室压铸 (cold-chamber die-casting, CD) 及热室压铸 (hot-chamber die-casting, HD) 试样进行对比。显微组织表征如图 11 所示，3 类铸造试样的相组成均以 α-Mg 基体相和共晶 $Mg_{17}Al_{12}$ 相为主。然而，触变射铸试样展现出显著不同的晶粒形貌：其 α-Mg 晶粒呈现均匀分布的细小球状特征，平均晶粒尺寸约为 10.5 μm；相比之下，热室压铸和冷室压铸试样的晶粒尺寸分别较触变射铸试样增大了 70% 和 20% 以上。背向散射电子衍射技术 (electron back scatter diffraction, EBSD) 反极图分析表明，晶格呈随机取向，但局部区域显示晶体学排列的一致性，即多晶粒聚集成大晶团。力学结果表明，触变射铸试样获得了 279 MPa 的极限抗拉强度和 192 MPa 的屈服强度，较冷室压铸试样分别提升约 10%，展现出显著的力学性能优势。



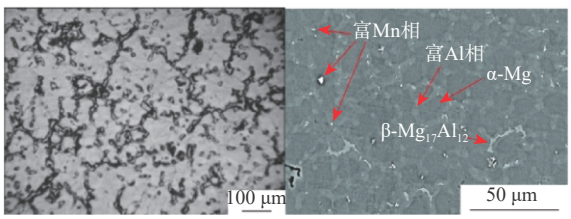
(a) 光学显微镜图像　(b) 背向散射电子衍射技术图像

图 11 冷室压铸、热室压铸和触变射铸试样的微观结构[20]

Fig. 11 Microstructures of cold-chamber die-casting, hot-chamber die-casting, and thixomolding samples[20]

在结构应用方面，当铝元素质量分数处于 0.05～0.06 范围时，AM 系镁合金的强度与延展性实现最佳平衡。因此，相较于 AZ 系镁合金，AM 系镁合金具有更优异的塑性变形能力[103-104]。Song 等[105] 通过拉伸试验系统研究了高压压铸制备的 AM50 合金的断裂机制，发现裂纹主要萌生于显微组织内部存在的缩松缺陷及微孔隙富集区域。基于 Jiang 等[63] 对铸态 AM60 镁合金的微观组织表征结果显

示, 其由平均晶粒尺寸达 320 μm 的粗大枝晶组成; 类似的组织特征在 AM50A 合金中同样存在(见图 12(a))[106]。作为对比, Anthony 等[107] 利用扫描电子显微镜 (scanning electron microscope, SEM) 背散射电子 (back scattered electron, BSE) 成像技术研究了触变射铸 AM60 的微观结构特征(见图 12(b))。在 BSE 图像中, α-Mg 基体呈现最暗的灰度强度, 而晶界附近可清晰分辨的富铝区域则表现为较亮区域, 灰度强度最高的颗粒对应于富锰金属间化合物。



(a) AM50A铸态镁合金的微观结构[106]　(b) 触变射铸AM60镁合金的SEM BSE图像[107]

图 12　常规铸态与触变射铸态 AM 系镁合金的微观组织特征

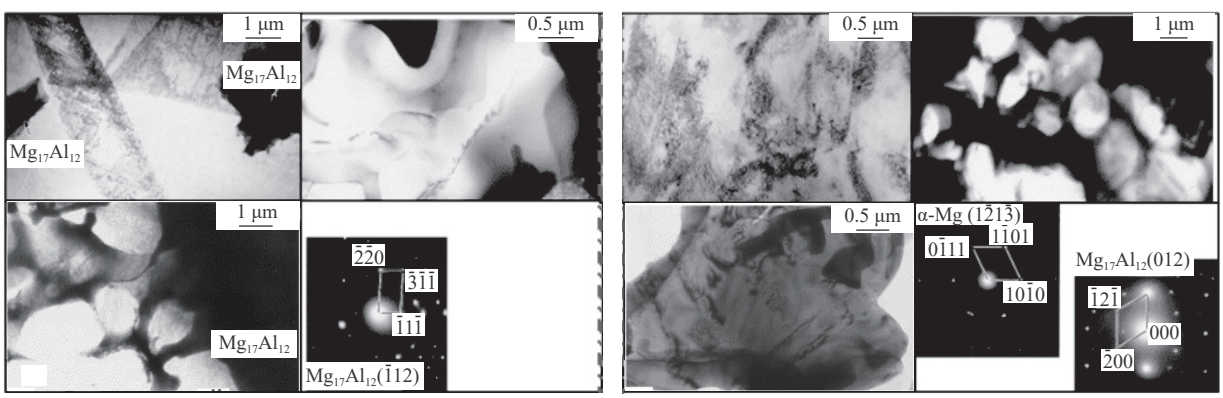
Fig. 12　Microstructure characteristics of AM series magnesium alloys prepared by conventional casting and thixomolding

3.2　触变射铸镁合金的性能

3.2.1　力学性能

由于采用触变射铸成形的镁合金可通过固溶时效热处理进一步显著提升力学性能, 目前已有大量研究致力于解析镁铝合金在不同热处理条件下的相变机制。固溶处理时, β-$Mg_{17}Al_{12}$ 金属间化合物在 α-Mg 基体中发生固溶溶解, 形成过饱和固溶体, 随后的时效处理则促使亚稳态 β 相粒子从过饱和基体中析出[108]。太原理工大学的 Yang 等[79] 对比了触变射铸和高压压铸 2 种成形方式对 AZ91D 合金中 $Mg_{17}Al_{12}$ 析出动力学的影响, 发现触变射铸成形样品的时效动力学明显快于高压压铸样品, 在 216℃ 时效时, 峰时效时间由 12 h 显著缩短至 5 h。

Czerwinski 等[67] 研究了采用触变射铸工艺制备的 AZ91D 镁合金, 其中, 未熔化的固相比例在 5%～60% 之间, 随着初生固相含量从 5% 增加至 43%, 共晶区 α-Mg 晶粒尺寸由 10 μm 细化至 5 μm, 拉伸性能测试表明, 当初生固相含量低于 20% 时, 合金抗拉强度和延伸率分别稳定在 240 MPa 和 4.5% 的水平, 当初生固相含量超过 20% 后, 力学性能呈现持续下降趋势。透射电子显微镜 (transmission electron microscopy, TEM) 图像表征(见图 13)显示, 当初生固相分数由 5% 增加至 43% 时, 显微组织的位错密度明显提升, 共晶组织相比低固相分数条件下更为细密, 部分 α-Mg 晶粒尺寸仅 1 μm。该研究指出, 在高初生固相含量条件下, 初生固相颗粒与共晶基体间的界面特性是决定触变射铸 AZ91D 合金力学性能的关键因素。值得注意的是, 大尺寸 $Mg_{17}Al_{12}$ 相也是镁铝系合金主要的裂纹源和重要的裂纹扩展路径。因此, $Mg_{17}Al_{12}$ 化合物的尺寸、分布和含量将显著影响材料的力学性能[109]。



(a) 初生固相分数为5%　(b) 初生固相分数为43%[67]

图 13　触变射铸镁合金试样的 TEM 图像[67]

Fig. 13　TEM images of thixomolding magnesium alloy specimens[67]

表 3 系统对比分析了国内外采用触变射铸技术制备的镁铝系镁合金的性能参数, 并与传统高压压铸及其他半固态成形工艺进行了性能对比。由于 AZ91D 合金中铝元素质量分数较高(约 0.09), 其 $Mg_{17}Al_{12}$ 相含量较多, 因而表现出更高的屈服强度, 但延伸率略低于 AM60B 合金, 后者含有更多初生 α-Mg 相和较少的硬脆金属间化合物 $Mg_{17}Al_{12}$ 相。实验结果表明, 与传统铸造技术及其他半固态成形工艺相比, 触变射铸镁合金在综合力学性能方面具备显著优势。

3.2.2　耐腐蚀性能

镁合金的腐蚀行为显著受其显微组织的影响, 当基体中的合金元素分布较为均匀时, 耐蚀性能得以提升。此外, 晶粒尺寸的差异同样对腐蚀速率产

表3 国内外半固态触变射铸Mg-Al系镁合金力学性能分析

Table 3 Analysis of mechanical properties of domestic and foreign semi-solid thixomolding Mg-Al series magnesium alloys

镁合金成分	屈服强度/MPa	抗拉强度/MPa	延伸率/%	成形方法	参考文献
AZ91D	169	299	8.0	触变射铸	[15]
AZ91D	192	279	10.0	触变射铸	[20]
AZ91D		240	4.5	触变射铸	[67]
AZ91D	180	230	4.4	触变射铸	[77]
AZ91D		245	3.9	触变射铸	[110]
AZ91D		285	7.6	触变射铸	[111]
AZ91D		253	4.3	触变射铸	[112]
AZ91D		170	3.0	触变射铸	[113]
AZ91D	170	273	7.2	触变射铸	[71]
AZ91D	158	256	6.0	触变射铸	[114]
AZ91D	145	248	7.4	流变射铸	[34]
AZ91D	138	225	4.5	流变压铸	[115]
AZ91D	144	200	2.5	高压压铸	[115]
AM60	127	192	5.0	触变射铸	[15]
AM50A	138	268	20.0	触变射铸	[22]
AM60B	147	278	18.8	触变射铸	[22]
AM60B	160	240	7.2	触变射铸	[77]
AM60	140	210	5.4	触变射铸	[116]
AM60B	146	269	11.0	触变射铸	[117]
AM50A	140	269	7.8	触变射铸	[118]
AM60B	102	155	8.0～9.0	高压压铸	[103]
AM60B		141	5.3	高压压铸	[119]
AM50A		103	5.5	高压压铸	[120]

生重要影响。研究针对AZ91合金的不同晶粒尺寸表明,腐蚀速率随着晶粒尺寸的减小而降低[121]。触变射铸工艺获得的致密、偏析较少且更细小的微观结构,有利于提升镁合金的耐腐蚀性。

Zander等[122]将高压压铸、双辊铸造(twin-roll casting, TRC)与触变射铸制备的AZ91D镁合金样品在5×10^{-4} mol/L NaCl溶液中浸泡约4.5天,以测试不同工艺下镁合金的耐腐蚀性能,结果如图14所示,与双辊铸造试样表面整体因腐蚀显著变色及高压压铸试样发生丝状腐蚀不同,触变射铸试样仅呈现轻微局部腐蚀特征,腐蚀试样横截面的SEM腐蚀形貌分析结果显示,双辊铸造材料发生平均深度约为10 μm的均匀腐蚀,高压压铸材料表现为平均渗透深度达30 μm的高度局部化腐蚀,相比之下,触变射铸材料仅显示可忽略的腐蚀损伤,平均渗透深度约为2 μm。该结果表明,触变射铸工艺可有效提升镁合金的耐腐蚀性能。

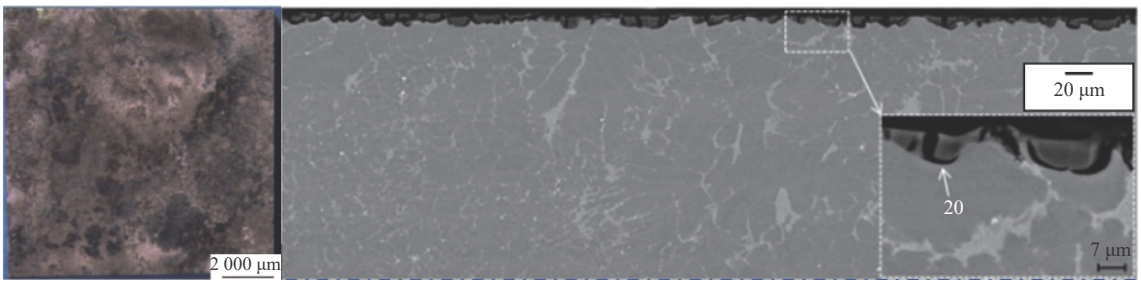
3.2.3 疲劳性能

在实际工程应用中,由疲劳引起的损伤累积及其引发的金属材料灾难性破坏是导致构件失效的主要机制之一。据统计,在安全关键型工程结构中,超过90%的构件失效可归因于疲劳行为。因此,疲劳抗力已成为衡量材料力学性能的一个关键指标[123]。触变射铸镁合金具有独特的微观组织和极低的孔隙率,因而其疲劳性能优于传统铸造工艺制备的镁合金。

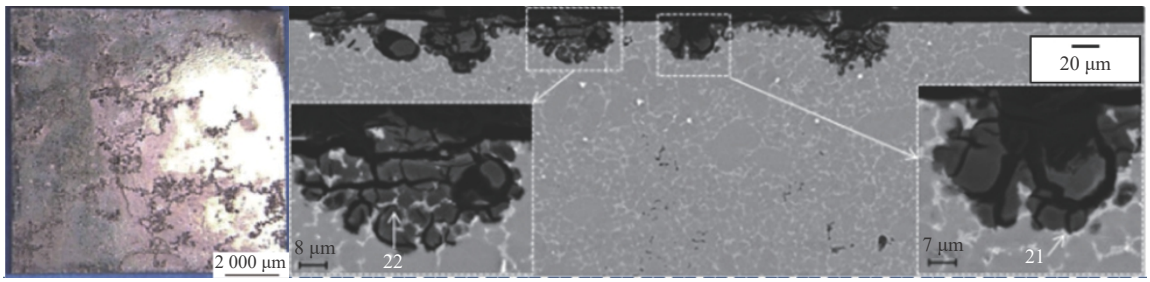
Patel等[124]对触变射铸工艺制备的AM60B镁合金进行研究发现,在应变速率为1×10^{-2} s^{-1}下进行测试,其极限抗拉强度为234 MPa,屈服强度为141 MPa,并且延伸率为9.1%,同时还评估了触变射铸AM60B镁合金的疲劳性能,如图15所示,相较于因存在强织构而导致滞回曲线高度不对称的挤压镁合金,触变射铸镁合金呈现几乎对称的滞回曲线,这一特征使得触变射铸AM60B在较低应变幅(0.2%～0.4%)下表现出更优的疲劳性能,此外,触变射铸合金的低气孔率,显著减小了实验数据的离散性。该研究者还在其他研究中指出[125],在低应变幅条件下,半固态成形的AZ91D和AM60B合金的疲劳寿命明显高于高压压铸制备的对应合金,这主要归因于其气孔缺陷尺寸相对较小。

3.3 其他新型半固态镁合金

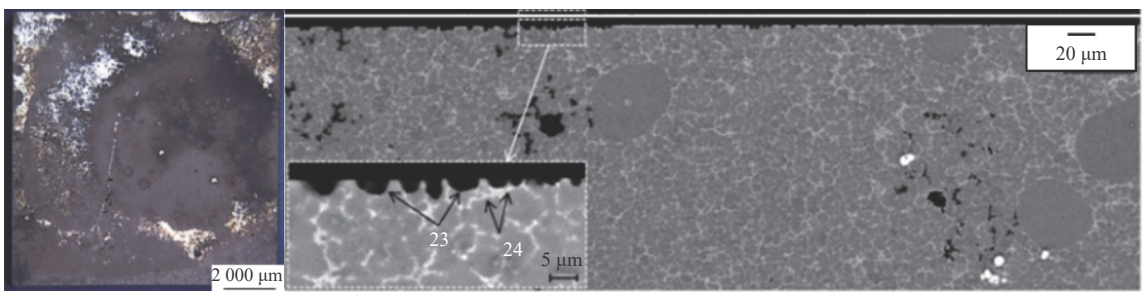
半固态镁合金体系的研究与应用目前主要集



(a) 双辊铸造



(b) 高压压铸



(c) 触变射铸

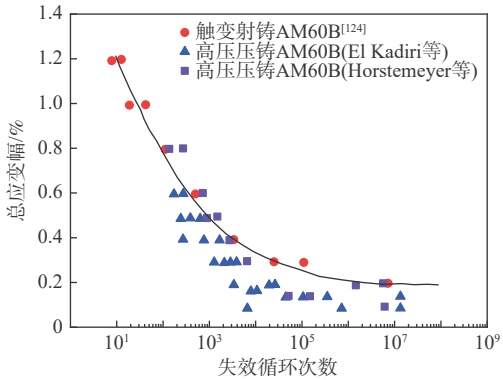
图 14　不同工艺的镁合金试样在室温下经 5×10^{-4} mol/L NaCl 溶液浸泡约 4.5 天后的顶面腐蚀形貌与横截面 SEM 图像[122]

Fig. 14　Top surface corrosion morphology and cross-sectional SEM images of magnesium alloy specimens with different processes after immersion in 5×10^{-4} mol/L NaCl solution at room temperature for approximately 4.5 days[122]

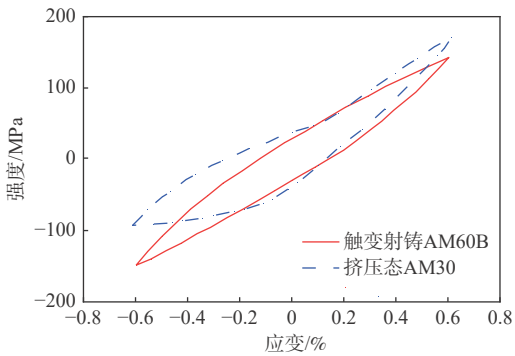
中于高 Al 元素含量的合金体系[66,101]。然而, 此类合金存在若干瓶颈问题, 难以满足工业界对轻量化金属日益增长的综合性能需求。其主要局限性体现在 2 方面: ①时效硬化响应不足。镁合金中的刚性金属间化合物相通过直接强化机制及改善成形性能来提升其力学性能, 其强化效果高度依赖于颗粒的形貌特征, 包括尺寸、与基体晶粒的晶体学取向关系及空间分布规律。通常, 细小弥散沉淀相主要提升材料强度, 而粗大颗粒则更侧重于抑制变形过程中的织构演变。然而, 高 Al 元素含量镁铝合金中的 $Mg_{17}Al_{12}$ 析出相, 因其与基体间的共格界面结构及其宏观尺寸特征(通常呈粗大晶粒形貌且相分布密度较低), 显著制约了合金的时效硬化响应, 导致其强度提升幅度普遍低于 30 MPa[126]。②高温蠕变抗力低。研究指出, $Mg_{17}Al_{12}$ 相熔点较低, 在高温下易发生粗化并促进晶界滑移, 这一机制导致现有商用镁铝合金普遍存在蠕变抗力不足的问题[127]。为突破现有技术瓶颈并推动触变射铸工艺的更广泛应用, 开发适用于该工艺且兼具优异综合性能的新型镁合金材料已成为本领域的重要研究方向。

在镁铝基合金体系中引入 Ca 元素, 不仅显著增强了材料的抗氧化性、阻燃特性及耐腐蚀能力, 还能有效抑制凝固过程中亚稳态 $Mg_{17}Al_{12}$(A12) 相的析出现象。尤为值得注意的是, 该合金体系(见图 16)在铸态及热处理状态下均呈现出独特的相结构演变特征, 形成了包含 Mg_2Ca(C14)、$(Mg,Al)_2Ca$(C36) 及 Al_2Ca(C15) 等含钙金属间化合物的新型显微组织。这些化合物具有高热稳定性和高熔点, 可显著增强镁合金在高温环境下的抗拉强度与抗蠕变性能[128]。

Tsukeda 等[129] 通过系统研究揭示了 Mg-Al-Ca 合金在触变射铸过程中的组织演化规律及其力学性能特征, 研究表明, 该合金的典型微观结构由等轴状初生 α-Mg 固溶体晶粒构成, 晶界处分布着由共晶 α 相与 $Mg_{17}Al_{12}$、Al_2Ca 等金属间化合物共同组成的连续网状结构, 这种独特的显微组织特征可显著增强合金的力学性能。但需要指出, 过量钙添加会因脆性相数量增加导致延伸率呈现显著下降趋势。



(a) 触变射铸与高压压铸AM60B镁合金的疲劳寿命对比[124]
(图中触变射铸AM60B数据引自文献[124]，高压压铸AM60B数据分别源自El Kadiri等和Horstemeyer等的研究结果，经文献[124]整理引用)



(b) 相同应变加载条件下(应变比为−1，应变速率为1×10^{-2} s^{-1}，总应变幅值为0.6%)触变射铸AM60B与挤压态AM30镁合金的疲劳滞后环对比[124]

图 15 触变射铸 AM60B 镁合金与不同工艺制备同系镁合金的疲劳性能对比[124]

Fig. 15 Comparison of fatigue properties between thixomolding AM60B magnesium alloy and counterpart Mg alloys fabricated by conventional processes[124]

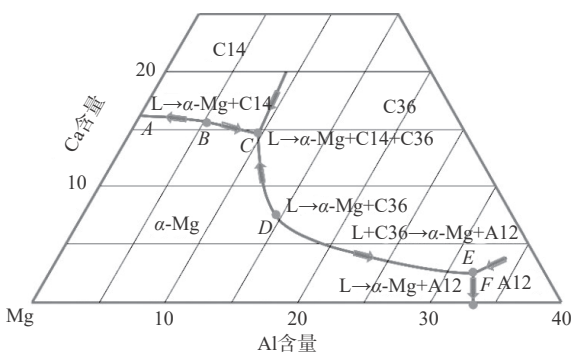
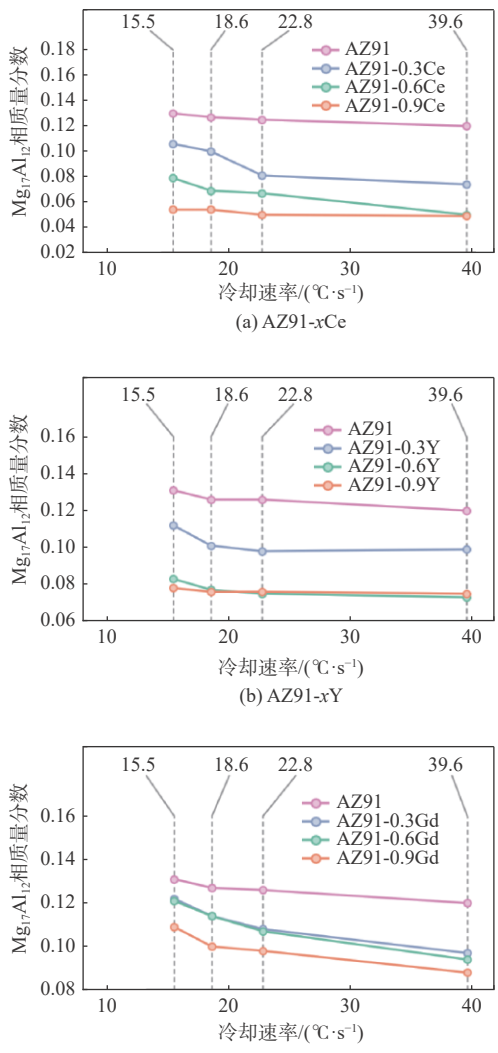


图 16 Mg-Al-Ca 合金体系富 Mg 角的液相线投影图[128]

Fig. 16 Phase line projection of Mg rich angle in Mg-Al-Ca alloy system[128]

除碱土金属钙外，稀土元素(rare-earth element, RE)也被证明是改善镁合金性能的有效途径。添加稀土元素可通过调控微观组织结构改善材料特性，从而有效提升镁合金综合性能[130]。Suzuki 等[131]系统研究了稀土元素对触变成形 Mg-Al-Ca 合金的影响机制，发现稀土元素的引入可促进 α-Mg 基体中纳米级析出相的形成，且这些析出相的特征参数(数量密度与尺寸分布)受热处理工艺参数的影响。适量稀土元素既可保持晶界连续网状结构特征，又能显著增强材料在蠕变过程中的微观组织稳定性。但需注意的是，过量添加会导致蠕变强度出现显著劣化现象，该现象主要源于 Al-RE 金属间化合物粗化引起的晶界覆盖度下降。在另一项研究中，Cai 等[132]研究了冷却速率和稀土元素 (Ce、Y、Gd) 对 AZ91合金中 $Mg_{17}Al_{12}$ 相的影响，实验结果如图 17 所示，冷却速率主要调控金属间化合物的形貌特征及其空间分布状态，但对第二相体积分数无明显调控作用，稀土元素的添加不仅能显著抑制 β 相形成，还可促进显微组织细化，其中，Ce 元素对 β 相体积分数的抑制作用最为显著。



(a) AZ91-xCe

(b) AZ91-xY

(c) AZ91-xGd

图 17 AZ91-xRE(Ce、Y、Gd, x = 0, 0.3, 0.6, 0.9) 合金中 $Mg_{17}Al_{12}$ 化合物的质量分数[132]

Fig. 17 Mass fraction of $Mg_{17}Al_{12}$ compound in AZ91-xRE (Ce, Y, Gd, x = 0, 0.3, 0.6, 0.9) alloys[132]

当前，针对适用于半固态触变射铸成形工艺的新型镁合金(如低合金化镁合金、镁稀土合金等)的公开研究报道还相对匮乏。随着材料基因工程的

快速发展, 基于高通量计算、高通量实验、材料数据库及机器学习辅助成分优化的集成研发模式, 有望为开发适用于该工艺的新型镁合金提供新途径, 推动此类新型镁合金在复杂结构部件制造中的规模化产业化应用[133]。

4　镁合金触变射铸的成形设备及技术的产业应用

4.1　镁合金触变射铸成形设备的发展

Thixomolding®技术由美国陶氏化学公司于 20 世纪 80 年代首次应用于工业生产。1991 年, 陶氏化学公司为此技术申请了美国专利并成立 Thixomat 公司, 以 “Thixomolding®” 为商标名称, 协同日本制钢所 (Japan Steel Works, JSW) 推进该工艺的商业化应用, 实现了镁合金半固态触变射铸成形技术较为成熟的工业化应用[8,15,134-135]。

半固态成形设备的核心技术突破主要于 2000—2010 年间在日本完成。期间, JSW 公司相继推出第二代和第三代改进机型, 确立了其在该领域的技术主导优势[136]。截至 2016 年, 中国大陆地区配置的半固态成形设备总数约为 170 台套。从设备参数分布看, 锁模力为 650 t 级的机型占据市场主导地位, 在国内外市场均保持约 80% 的装机占比。

在 21 世纪 20 年代之前, 中国国内的镁合金触变射铸设备主要应用于生产 3C 行业的具有薄壁结构的零部件, 特别是笔记本外壳等零件[135]。自 2020 年以来, 受碳中和战略推进、新能源汽车轻量化需求迫切及镁价阶段性回调等多重因素驱动, 镁合金替代铝合金构件的研究已成为材料工程领域的重要课题。然而, 彼时的镁合金半固态触变射铸设备仍存在锁模力吨位较低和熔体注射量有限等固有缺陷, 在汽车结构件制造领域, 特别是大型部件制造方面, 远不能满足产业界的工艺要求。因此, 高压压铸至今仍是制造车辆零部件的主流工艺[137]。为应对产业界的需求和工艺要求的挑战, 一些国内外的触变射铸成形设备制造商着手研发大吨位的成形设备, 陆续推出 3 000～4 000 t 级别锁模力的设备解决方案, 其成形能力可满足长度达 2 m 的大型零件生产需求[138]。设备相关参数如表 4 及图 18 所示。

表 4　国内外部分设备制造商旗下 3 000～4 000 t 级镁合金触变射铸成形机参数[70,136,139-143]

Table 4　Parameters of 3 000−4 000 t thixomolding machines from selected domestic and international equipment manufacturers[70,136,139-143]

设备厂家	型号	设备吨位/t	锁模机构	螺杆直径/mm	最大理论注射量/kg	已公布下一代设备吨位/t
日本制钢所	JLM3000-MGIIeL	3 000	三板式	130	8	
海天金属	HMG3000	3 000	三板式	150	10	5 000
伊之密	UN3200MGII	3 200	三板式	160	11	6 600
伯乐智能	MTX4000D	4 000	二板式	170	17	



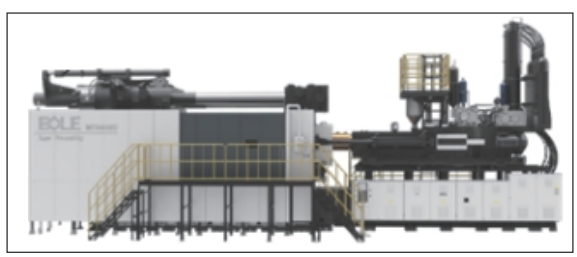
(a) JLM3000-MGIIeL



(b) HMG3000



(c) UN3200MGII



(d) MTX4000D

图 18　大型镁合金触变射铸成形设备[70,136,140-142]

Fig. 18　Large magnesium alloy thixomolding equipment[70,136,140-142]

近年来，镁合金半固态触变射铸成形技术在大型化设备研发方面取得了显著进展。国内设备制造商通过持续攻克多项关键技术。例如，伊之密在料管组采用特殊的耐高温钢材和热处理工艺，使其能够耐受高达650℃的高温，使用寿命可达40万次；伯乐智能为成形设备配备了螺杆直径170 mm、长度超过4 m的超耐热高强材质螺杆，Co、Mo元素质量分数达0.14以上，能够在650℃以上的高温环境中长期稳定运行。这些关键技术的突破为大型设备的研发提供了支撑，目前多家公司已推出下一代5 000 t级以上大型成形设备方案[70,140,143]。该技术体系为新能源汽车及消费电子领域大尺寸镁合金构件的制造提供了创新工艺路径，其典型应用场景涵盖车身轻量化组件（如多联屏背板、车内门板及仪表板骨架）及三电系统核心结构件（如驱动电机壳体、电池箱体及电控箱）[138]。此外，基于高压压铸设备的广泛应用背景，力劲公司开发了TPI镁合金半固态模组，该模组能够将传统压铸机改造为半固态触变铸造系统。这一创新使得镁合金压铸生产商能够在不更换现有设备的前提下，灵活切换高压压铸与半固态压铸工艺。具体而言，高压压铸与半固态压铸之间的转换可在8 h内完成，而不同镁合金（如从AZ91到AM60）的切换仅需1 h。通过这种低成本升级方式，原有压铸机可转型为高效的半固态触变铸造设备，从而优化资源利用并降低投资负担[144]。

总体而言，工艺设备的持续迭代升级，不仅有力推动了触变射铸技术在镁合金复杂薄壁构件批量化生产中的工程化应用，更为其未来拓展至大型车身结构件、车内结构件等应用领域奠定了坚实基础[136,140]。

4.2 镁合金触变射铸技术的产业应用

与成熟且应用广泛的压铸工艺相比，触变射铸成形的镁合金在实际生产中的优势主要体现在以下3个方面：①触变射铸成形的零部件具有更低的孔隙率和更优异的综合性能；②在能耗方面，由于触变射铸工艺的热熔温度较压铸工艺低80℃，无需一次性熔炼几百公斤镁合金原料，也无需保温，能耗降低60%，同时减小了对模具的热冲击，使模具寿命可延长2～4倍；③工艺生产流程更为简化，得益于良好的充型性能，近净成形制件仅需简单少量的机械加工即可完成生产[145]。这些优势共同使得镁合金触变射铸技术在近年来展现出广阔的市场前景。

触变射铸工艺杰出的薄壁成形能力（0.5～1 mm），以及其固态加工特性带来的能效与安全性优势，使其在3C领域具有独特适用性[82]。对于笔记本电脑面板、盖板等对表面光洁度与尺寸精度具有严苛要求的外观组件，触变射铸技术已成为成熟的工业化解决方案；该技术不仅突破了传统工艺对镁合金壁厚的限制，更能通过精确的温控系统与流变学参数优化，稳定实现高良品率的生产目标[146]。值得注意的是，该技术的应用已扩展至功能性结构件制造领域（见图19）。在游戏机壳体、医疗显示设备防护罩、人机交互界面组件（如鼠标、键盘）等典型产品中，该技术成功实现了复杂曲面构型的精密成形与薄壁加强筋的一体化铸造，显著提升了结构效率和空间利用率，为移动终端设备的微型化与轻量化设计提供了重要的技术支撑。

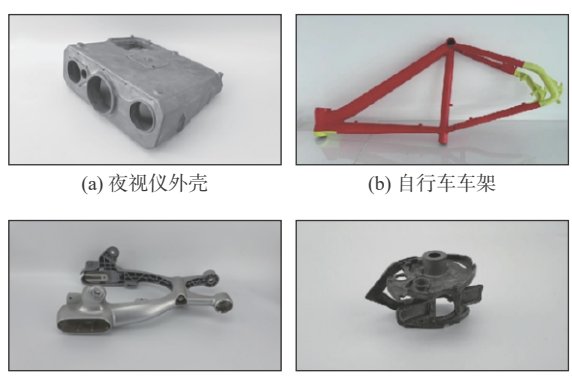

(a) 笔记本A面　(b) 笔记本键盘面　(c) 医疗屏外壳

(d) 游戏机外壳　(e) 鼠标骨架　(f) 键盘外壳

图19　消费电子结构件、外壳件[145,147]

Fig. 19　Structural components and housings for consumer electronics[145,147]

除3C领域外，镁合金触变射铸技术的应用范围正逐步拓展至多元化民用消费品领域，为产品轻量化提供了创新解决方案。如图20所示，以户外运动设施为例，其服役环境复杂：一方面，需承受高强度机械载荷及严苛环境条件；另一方面，受限于其复杂薄壁构件的结构特点。传统镁合金加工工艺因难以兼顾高力学性能要求与复杂结构成形需求而面临应用瓶颈。相比之下，触变射铸技术不仅成功


(a) 夜视仪外壳　(b) 自行车车架

(c) 渔线轮　(d) 摩托车前叉

图20　户外运动产品结构件、外壳件[145]

Fig. 20　Structural components and housings for outdoor sports products[145]

实现了构件的轻量化设计目标, 更通过改善材料微观组织结构显著提升了关键部件的服役可靠性。此外, 该技术特有的近净成形特性显著提高了材料利用率, 在降低生产成本的同时契合绿色制造理念。

近年来, 汽车工业正逐步成为镁合金半固态触变射铸技术的重要应用领域。在新能源汽车车体轻量化需求的驱动下, 市场对镁合金构件在车辆轻量化解决方案中的需求呈现显著增长态势[5]。该工艺在制备具有复杂几何构型的部件方面展现出独特优势, 尤其适用于薄壁件和高精度零部件的批量化生产。典型应用案例如图 21 所示, 包括方向盘骨架、中控仪表盘支架等。当前, 成形设备性能与工艺控制体系的持续优化正推动产业升级, 国内主要结构件供应商正加速实施跨行业产能转移战略, 逐步将其业务从消费电子领域延伸至汽车供应链体系。



(a) 方向盘骨架　(b) 越野车备胎架　(c) 中控显示屏支架



(d) 汽车GPS外壳　(e) 汽车座椅头枕支架　(f) 中控仪表盘支架

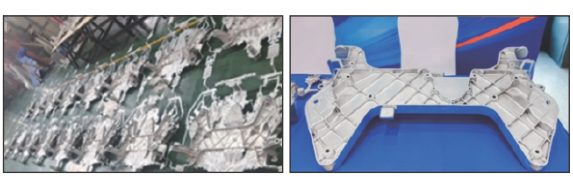
图 21　汽车结构件[68,145,147]

Fig. 21　Structural components for automobiles[68,145,147]

基于大型成形设备的技术突破, 镁合金半固态触变射铸技术显著突破了传统工艺的性能极限, 其应用范围逐步拓展至汽车关键结构部件制造领域(部分产品如图 22 所示)。依托该技术革新, 目前已成功制备出一系列创新构件, 包括电驱系统外壳、大尺寸镁合金轮毂、一体式副车架、减震塔总成及仪表板横梁支架等[143]。



(a) 镁合金电驱壳体　(b) 20寸级别大型轮毂



(c) 副车架1　(d) 副车架2

图 22　汽车关键大型结构件[140,142,148-149]

Fig. 22　Critical large structural components for automobiles[140,142,148-149]

5　结 束 语

相较于传统压铸工艺, 镁合金触变射铸成形技术能够实现近净成形结构件, 显著提高材料利用率并延长模具寿命。尤为关键的是, 该工艺制备的构件具有晶粒细小、孔隙率低、可靠性高的微观组织特征。这一显微组织特性使该技术能够成功制备具有复杂几何结构及薄壁特征的高性能铸件, 进而在 3C 产品外壳、汽车显示背板等精密构件制造中形成差异化竞争优势, 并逐步向汽车底盘结构件等关键安全部件领域拓展。

1) 在基础研究层面, 学者们已系统揭示了镁合金半固态浆料中非枝晶微观结构的演变机制。研究表明, 获得高质量浆料的关键在于精确控制合金熔体的冷却速率和流动状态。尽管关于枝晶破碎与球形颗粒形成的机理仍存在学术争议, 现有理论框架为浆料制备工艺优化提供了重要指导。实验数据表明, 触变射铸组织的形态均匀性(呈现球状或蔷薇状特征)与浆料流变性能及最终铸件力学性能呈显著正相关。以 AZ/AM 系合金为例, 其室温拉伸性能持续优于传统加工工艺制备的制品。然而, 针对含稀土镁合金等新型合金体系在触变射铸工艺下的浆料制备行为、晶粒细化机理及强化机制等方面的系统性研究数据仍显不足。

2) 从工业应用的技术经济性角度考量, 触变射铸成形设备的开发需要统筹协调生产效率、质量稳定性与成本效益的动态平衡。当前, 中国在大型化设备研制方面虽已取得阶段性突破, 但产业化进程中仍面临着工艺标准化体系不完善、缺陷形成机制不明确及规模化成本控制等关键瓶颈问题。针对大注射量镁合金半固态触变射铸工艺的研究仍处于技术攻关阶段, 其核心科学问题集中于以下方面的突破: 如何构建固相体积分数的高精度动态调控机制; 建立微观组织均匀性的多尺度工艺调控体系; 揭示凝固缺陷的跨尺度形成与演化机制; 系统研究成形过程中多物理场耦合作用机制并协同优化工艺参数。这些关键问题的解决亟须通过系统性理论探索与技术创新实现协同攻关。

随着全球轻量化需求的持续增长, 半固态触变射铸成形技术因其独特的工艺优势, 有望成为实现镁合金性能优化与成形缺陷控制的最佳工艺解决方案。未来研究的重点方向应围绕以下核心领域展开: 新型镁合金体系开发(涵盖稀土元素合金化与低合金化镁合金); 半固态浆料球状晶显微组织演化机理的系统阐释; 突破传统机械粉碎制浆模式

的创新工艺研发；基于材料基因工程的多物理场耦合数值模拟与人工智能预测模型的工艺优化体系构建。需特别指出，尽管新型半固态触变射铸镁合金的设计、制备及其成形相关基础研究尚处于起步阶段，但通过整合高通量计算模拟与机器学习驱动的合金成分优化策略，并结合我国在半固态触变射铸设备自主化研发与大型化制备方面的技术积累，可显著推进镁合金复杂构件近净成形工艺的产业化进程，从而为交通设施轻量化战略的实施提供有力的技术支撑。

参考文献（References）

[1] 王渠东. 镁合金及其成形技术[M]. 北京: 机械工业出版社, 2017: 5.
Wang Q D. Magnesium alloys and their forming technologies[M]. Beijing: China Machine Press, 2017: 5(in Chinese).

[2] Powell B R, Krajewski P E, Luo A A. Magnesium alloys for lightweight powertrains and automotive structures[M]//Mallick P K. Materials, design and manufacturing for lightweight vehicles. 2nd ed. Cambridge: Woodhead Publishing, 2021: 125-186.

[3] 邓辉, 郭学锋, 黄丹, 等. 镁合金半固态成形技术的研究和发展[J]. 热加工工艺, 2011, 40(23): 9-13.
Deng H, Guo X F, Huang D, et al. Research and development of semi-solid processing of magnesium alloy[J]. Hot Working Technology, 2011, 40(23): 9-13(in Chinese).

[4] 潘帅, 李强, 于宝义, 等. 镁合金半固态研究进展[J]. 稀有金属材料与工程, 2019, 48(7): 2379-2385.
Pan S, Li Q, Yu B Y, et al. Research progress of Mg alloy semi-solid forming[J]. Rare Metal Materials and Engineering, 2019, 48(7): 2379-2385(in Chinese).

[5] D'Errico F, Rivolta B, Gerosa R, et al. Thixomolded magnesium alloys: strategic product innovation in automobiles[J]. JOM, 2008, 60(11): 70-76.

[6] Fang X G, Lü S L, Zhao L, et al. Microstructure and mechanical properties of a novel Mg-RE-Zn-Y alloy fabricated by rheo-squeeze casting[J]. Materials & Design, 2016, 94: 353-359.

[7] She J, Chen J, Xiong X M, et al. Research advances of magnesium and magnesium alloys globally in 2023[J]. Journal of Magnesium and Alloys, 2024, 12(9): 3441-3475.

[8] Czerwinski F. Magnesium research in Canada: highlights of the last two decades[J]. Journal of Magnesium and Alloys, 2023, 11(10): 3484-3510.

[9] Janudom S, Wannasin J, Basem J, et al. Characterization of flow behavior of semi-solid slurries containing low solid fractions in high-pressure die casting[J]. Acta Materialia, 2013, 61(16): 6267-6275.

[10] Luo A A. Magnesium casting technology for structural applications[J]. Journal of Magnesium and Alloys, 2013, 1(1): 2-22.

[11] Michaeli W, Rettemeier G. Thixomoulding injection moulding of semi-solid magnesium alloys[J]. Journal of Polymer Engineering, 2004, 24(1-3): 49-64.

[12] Wu G H, Wang C L, Sun M, et al. Recent developments and applications on high-performance cast magnesium rare-earth alloys[J]. Journal of Magnesium and Alloys, 2021, 9(1): 1-20.

[13] Spencer D B, Mehrabian R, Flemings M C. Rheological behavior of Sn-15 pct Pb in the crystallization range[J]. Metallurgical Transactions, 1972, 3(7): 1925-1932.

[14] Czerwinski F. Magnesium injection molding[M]. Berlin: Springer, 2008: 81.

[15] Decker R, Lebeau S, Wilson B, et al. Thixomolding® at 25 years [J]. Solid State Phenomena, 2016, 256: 3-8.

[16] 徐平, 杨昆. 金属半固态成型发展现状及展望[J]. 世界科技研究与发展, 2010, 32(1): 77-79.
Xu P, Yang K. The development and prospects of semi-solid metal forming[J]. World Sci-Tech R & D, 2010, 32(1): 77-79(in Chinese).

[17] Kirkwood D H. Semisolid metal processing[J]. International Materials Reviews, 1994, 39(5): 173-189.

[18] Lebeau S E, Walukas M W, Decker R F, et al. Evaluation of thixomolded, magnesium alloy components for structural automotive applications[C]//Proceedings of the SAE 2004 World Congress & Exhibition. Warrendale: SAE International, 2004.

[19] 肖冠菲, 姜巨福, 王迎, 等. 高熔点合金半固态加工技术研究进展[J]. 塑性工程学报, 2019, 26(2): 62-68.
Xiao G F, Jiang J F, Wang Y, et al. Research progress of semi-solid processing technology of high-melting alloy[J]. Journal of Plasticity Engineering, 2019, 26(2): 62-68(in Chinese).

[20] Okayasu M, Fukui T. A study of the mechanical properties of a Mg-Al-Zn alloy (AZ91) produced via thixomolding[J]. International Journal of Material Forming, 2021, 14(2): 271-280.

[21] 罗守靖, 田文彤, 李金平. 21世纪最具发展前景的近净成形技术——半固态加工[J]. 特种铸造及有色合金, 2001(S1): 175-180.
Luo S J, Tian W T, Li J P. The most promising near-net forming technology in 21 Century—semi-solid processing[J]. Special Casting and Nonferrous Alloys, 2001(S1): 175-180(in Chinese).

[22] Kiuchi M, Kopp R. Mushy/semi-solid metal forming technology–present and future[J]. CIRP Annals, 2002, 51(2): 653-670.

[23] Flemings M C. Behavior of metal alloys in the semisolid state[J]. Metallurgical Transactions B, 1991, 22(3): 269-293.

[24] Flemings M C, Martinez R A. Principles of microstructural formation in semi-solid metal processing[J]. Solid State Phenomena, 2006, 116-117: 1-8.

[25] Chang Z Y, Su N, Wu Y J, et al. Semisolid rheoforming of magnesium alloys: a review[J]. Materials & Design, 2020, 195: 108990.

[26] Vogel A, Cantor B. Stability of a spherical particle growing from a stirred melt[J]. Journal of Crystal Growth, 1977, 37(3): 309-316.

[27] Doherty R D, Lee H I, Feest E A. Microstructure of stir-cast metals[J]. Materials Science and Engineering, 1984, 65(1): 181-189.

[28] Apaydin N, Prabhakar K V, Doherty R D. Special grain boundaries in rheocast Al-Mg[J]. Materials Science and Engineering, 1980, 46(2): 145-150.

[29] Ruvalcaba D, Mathiesen R H, Eskin D G, et al. In situ observations of dendritic fragmentation due to local solute-enrichment during directional solidification of an aluminum alloy[J]. Acta Materialia, 2007, 55(13): 4287-4292.

[30] Hellawell A, Liu S, Lu S Z. Dendrite fragmentation and the effects of fluid flow in castings[J]. JOM, 1997, 49(3): 18-20.

[31] Li T, Lin X, Huang W. Morphological evolution during solidification under stirring[J]. Acta Materialia, 2006, 54(18): 4815-4824.

[32] Paradies C J, Smith R N, Glicksman M E. The influence of con-

vection during solidification on fragmentation of the mushy zone of a model alloy[J]. Metallurgical and Materials Transactions A, 1997, 28(3): 875-883.

[33] Behnamfard S, Mousavian R T, Khosroshahi R A, et al. A comparison between hot-rolling process and twin-screw rheo-extrusion process for fabrication of aluminum matrix nanocomposite[J]. Materials Science and Engineering: A, 2019, 760: 152-157.

[34] Fan Z. Development of the rheo-diecasting process for magnesium alloys[J]. Materials Science and Engineering: A, 2005, 413: 72-78.

[35] 毛卫民. 半固态金属浆料先进制备技术的研究进展[J]. 铸造, 2012(8): 839-855.
Mao W M. Research progress of advanced preparation technology for semi-solid metal slurry[J]. Foundry, 2012(8): 839-855(in Chinese).

[36] Vives C. Elaboration of semisolid alloys by means of new electromagnetic rheocasting processes[J]. Metallurgical Transactions B, 1992, 23(2): 189-206.

[37] Wang C L, Chen A T, Zhang L, et al. Preparation of an Mg-Gd-Zn alloy semisolid slurry by low frequency electro-magnetic stirring[J]. Materials & Design, 2015, 84: 53-63.

[38] Ji S, Fan Z. Solidification behavior of Sn-15 wt pct Pb alloy under a high shear rate and high intensity of turbulence during semisolid processing[J]. Metallurgical and Materials Transactions A, 2002, 33(11): 3511-3520.

[39] Das A, Ji S, Fan Z. Morphological development of solidification structures under forced fluid flow: a Monte-Carlo simulation[J]. Acta Materialia, 2002, 50(18): 4571-4585.

[40] Wu S S, Wu X P, Xiao Z H. A model of growth morphology for semi-solid metals[J]. Acta Materialia, 2004, 52(12): 3519-3524.

[41] Qin R S, Wallach E R. Phase-field simulation of semisolid metal processing under conditions of laminar and turbulent flow[J]. Materials Science and Engineering: A, 2003, 357(1-2): 45-54.

[42] Fan Z, Liu G. Solidification behaviour of AZ91D alloy under intensive forced convection in the RDC process[J]. Acta Materialia, 2005, 53(16): 4345-4357.

[43] 张景新, 张奎, 刘国钧, 等. 电磁搅拌制备半固态材料非枝晶组织的形成机制[J]. 中国有色金属学报, 2000, 10(4): 511.
Zhang J X, Zhang K, Liu G J, et al. Formation mechanism of non-dendritic structure in semi-solid metals produced by ES process[J]. The Chinese Journal of Nonferrous Metals, 2000, 10(4): 511(in Chinese).

[44] Martinez R A, Flemings M C. Evolution of particle morphology in semisolid processing[J]. Metallurgical and Materials Transactions A, 2005, 36(8): 2205-2210.

[45] 赵春芳, 马光全, 颜泽华, 等. 浇注温度对近液相线挤压铸造 AZ91D 镁合金组织及性能的影响[J]. 特种铸造及有色合金, 2024, 44(1): 105-110.
Zhao C F, Ma G Q, Yan Z H, et al. Effects of pouring temperature on microstructure and mechanical property of near-liquidus squeeze casting AZ91D alloy[J]. Special Casting and Nonferrous Alloys, 2024, 44(1): 105-110(in Chinese).

[46] 王娜, 周志敏, 路贵民. 近液相线铸造过程中近球形 α 相的形成机理[J]. 材料科学与工艺, 2011, 19(2): 139-143.
Wang N, Zhou Z M, Lu G M. Formation mechanism of sphere-like α phase in near-liquidus casting[J]. Materials Science and Technology, 2011, 19(2): 139-143(in Chinese).

[47] 郭洪民, 杨湘杰. LSPSF 流变铸造工艺及其应用[J]. 特种铸造及有色合金, 2008(S1): 355-359.
Guo H M, Yang X J. Overview of LSPSF rheocasting process and its applications[J]. Special Casting and Nonferrous Alloys, 2008(S1): 355-359(in Chinese).

[48] Zhang Y F, Liu Y B, Zhang Q Q, et al. Microstructural evolution of thixomolded AZ91D magnesium alloy with process parameters variation[J]. Materials Science and Engineering: A, 2007, 444(1-2): 251-256.

[49] Young K P, Kyonka C P, Courtois J A. Fine grained metal composition: US4415374[P]. 1983-11-15.

[50] Zhang L, Liu Y B, Cao Z Y, et al. Effects of isothermal process parameters on the microstructure of semisolid AZ91D alloy produced by SIMA[J]. Journal of Materials Processing Technology, 2009, 209(2): 792-797.

[51] 程远胜, 卢华领, 陈善荣, 等. 高稀土 Mg-Gd 合金半固态等温处理过程中微观组织演变[J]. 特种铸造及有色合金, 2019, 39(3): 238-242.
Cheng Y S, Lu H L, Chen S R, et al. Microstructure evolution of high rare earth Mg-Gd alloy during semi-solid isothermal treatment process[J]. Special Casting and Nonferrous Alloys, 2019, 39(3): 238-242(in Chinese).

[52] Kirkwood D H, Sellars C M, Elias-Boyed L G. Thixotropic materials: US5133811[P]. 1992-07-28.

[53] Chen Q, Zhao Z D, Chen G, et al. Effect of accumulative plastic deformation on generation of spheroidal structure, thixoformability and mechanical properties of large-size AM60 magnesium alloy[J]. Journal of Alloys and Compounds, 2015, 632: 190-200.

[54] Jiang J F, Wang Y, Xiao G F, et al. Comparison of microstructural evolution of 7075 aluminum alloy fabricated by SIMA and RAP[J]. Journal of Materials Processing Technology, 2016, 238: 361-372.

[55] Jiang J F, Luo S J. Reheating microstructure of refined AZ91D magnesium alloy in semi-solid state[J]. Transactions of Nonferrous Metals Society of China, 2004, 14(6): 1074-1081.

[56] Luo S J, Chen Q, Zhao Z D. Effects of processing parameters on the microstructure of ECAE-formed AZ91D magnesium alloy in the semi-solid state[J]. Journal of Alloys and Compounds, 2009, 477(1-2): 602-607.

[57] Jiang J F, Wang Y, Luo S J. Application of equal channel angular extrusion to semi-solid processing of magnesium alloy[J]. Materials Characterization, 2007, 58(2): 190-196.

[58] Fu J L, Jiang H J, Wang K K. Influence of processing parameters on microstructural evolution and tensile properties for 7075 Al alloy prepared by an ECAP-based SIMA process[J]. Acta Metallurgica Sinica (English Letters), 2018, 31(4): 337-350.

[59] Kim W J, Jeong H G, Jeong H T. Achieving high strength and high ductility in magnesium alloys using severe plastic deformation combined with low-temperature aging[J]. Scripta Materialia, 2009, 61(11): 1040-1043.

[60] Yang B W, Wang Y, Gao M Q, et al. Microstructural evolution and strengthening mechanism of Al-Mg alloys with fine grains processed by accumulative continuous extrusion forming[J]. Journal of Materials Science & Technology, 2022, 128: 195-204.

[61] Mohammadi A, Enikeev N A, Murashkin M Y, et al. Developing age-hardenable Al-Zr alloy by ultra-severe plastic deformation: significance of supersaturation, segregation and precipitation on hardening and electrical conductivity[J]. Acta Materialia, 2021, 203: 116503.

[62] Jiang J F, Wang Y, Du Z M, et al. Microstructure and properties of AZ80 alloy semisolid billets fabricated by new strain induced melt activated method[J]. Transactions of Nonferrous Metals Society of China, 2012, 22: s422-s427.

[63] Jiang J F, Wang Y, Qu J J, et al. Microstructure evolution of AM60 magnesium alloy semisolid slurry prepared by new SIMA[J]. Journal of Alloys and Compounds, 2010, 497(1-2): 62-67.

[64] Jiang J F, Lin X, Wang Y, et al. Microstructural evolution of AZ61 magnesium alloy predeformed by ECAE during semisolid isothermal treatment[J]. Transactions of Nonferrous Metals Society of China, 2012, 22(3): 555-563.

[65] Czerwinski F. On the generation of thixotropic structures during melting of Mg-9%Al-1%Zn alloy[J]. Acta Materialia, 2002, 50(12): 3267-3283.

[66] Czerwinski F. Magnesium alloy particulates for thixomolding applications manufactured by rapid solidification[J]. Materials Science and Engineering: A, 2004, 367(1-2): 261-271.

[67] Czerwinski F, Zielinska-Lipiec A, Pinet P J, et al. Correlating the microstructure and tensile properties of a thixomolded AZ91D magnesium alloy[J]. Acta Materialia, 2001, 49(7): 1225-1235.

[68] Zirps M. How thixomolding is redefining electric vehicle production[EB/OL]. (2024-01-15). [2025-08-02]. https://www.lightmetalage.com/news/industry-news/magnesium/how-thixomolding-is-redefining-electric-vehicle-production/.

[69] 伊之密. 半固态镁合金注射成型机[EB/OL]. [2025-08-02]. https://www.yizumi.com/tms/thixomolding-machine.
YIZUMI. Thixomolding magnesium machine[EB/OL]. [2025-08-02]. https://www.yizumi.com/tms/thixomolding-machine(in Chinese).

[70] 海天金属. 镁合金注射成型机[EB/OL]. [2025-08-02]. https://haitian-dc.com/products/hmg-series/.
Haitian Die Casting. Thixomolding magnesium machine[EB/OL]. [2025-08-02]. https://haitian-dc.com/products/hmg-series/(in Chinese).

[71] Gu L D, Shang X Q, Wang J, et al. Effect of hierarchical cell structure and internal pores on mechanical properties of thixomolded AZ91D magnesium alloy[J]. Transactions of Nonferrous Metals Society of China, 2025, 35(3): 749-764.

[72] Qin Z H, Kang N, Zhang F Y, et al. Role of defects on the high cycle fatigue behavior of selective laser melted Al-Mg-Sc-Zr alloy[J]. International Journal of Fracture, 2022, 235(1): 129-143.

[73] Murakami Y, Miwa K J, Omura N, et al. Effects of injection speed and fraction solid on tensile strength in semisolid injection molding of AZ91D magnesium alloy[J]. Materials Transactions, 2012, 53(10): 1775-1781.

[74] 胡宇阳. 热处理工艺对半固态注射成形镁合金组织及性能研究[D]. 西安: 西安科技大学, 2023.
Hu Y Y. Effect of heat treatment process on microstructure and performance of semi-solid injection molded magnesium alloy[D]. Xi'an: Xi'an University of Science and Technology, 2023(in Chinese).

[75] Fu J L, Wang K K, Li X W, et al. Microstructure evolution and thixoforming behavior of 7075 aluminum alloy in the semi-solid state prepared by RAP method[J]. International Journal of Minerals, Metallurgy, and Materials, 2016, 23(12): 1404-1415.

[76] 徐升亮, 郭洪民, 杨湘杰. 金属半固态浆料非枝晶微观组织形成机制与表征[J]. 精密成形工程, 2024, 16(7): 109-123.
Xu S L, Guo H M, Yang X J. Formation mechanism and characterization of non-dendritic microstructure in metal semi-solid slurry [J]. Journal of Netshape Forming Engineering, 2024, 16(7): 109-123 (in Chinese).

[77] Patel H A, Chen D L, Bhole S D, et al. Microstructure and tensile properties of thixomolded magnesium alloys[J]. Journal of Alloys and Compounds, 2010, 496(1-2): 140-148.

[78] Buzolin R H, Volovitch P, Maltseva A, et al. Thixomolded AZ91D and MRI153M magnesium alloys and their enhanced corrosion resistance[J]. Materials and Corrosion, 2020, 71(3): 339-351.

[79] Yang L J, Wei Y H, Hou L F. Microstructure evolution of thixomolding AZ91D magnesium alloy during heat treatment[J]. Journal of Materials Science, 2010, 45(13): 3626-3634.

[80] Sadayappan K, Vassos M. Evaluation of a thixomolded magnesium alloy component for automotive application[C]//Proceedings of the SAE 2010 World Congress & Exhibition. Warrendale: SAE International, 2010.

[81] Otarawanna S, Gourlay C M, Laukli H I, et al. Agglomeration and bending of equiaxed crystals during solidification of hypoeutectic Al and Mg alloys[J]. Acta Materialia, 2010, 58(1): 261-271.

[82] Luo A A, Sachdev A K, Apelian D. Alloy development and process innovations for light metals casting[J]. Journal of Materials Processing Technology, 2022, 306: 117606.

[83] Czerwinski F, Zielinska-Lipiec A. The microstructure evolution during semisolid molding of a creep-resistant Mg-5Al-2Sr alloy [J]. Acta Materialia, 2005, 53(12): 3433-3444.

[84] Chen G, Zhang S, Zhang H M, et al. Controlling liquid segregation of semi-solid AZ80 magnesium alloy by back pressure thixo-extruding[J]. Journal of Materials Processing Technology, 2018, 259: 88-95.

[85] Wu G H, Zhang Y, Liu W C, et al. Microstructure evolution of semi-solid Mg-10Gd-3Y-0.5Zr alloy during isothermal heat treatment[J]. Journal of Magnesium and Alloys, 2013, 1(1): 39-46.

[86] Chen Y S, Zhang L, Liu W C, et al. Preparation of Mg-Nd-Zn-(Zr) alloys semisolid slurry by electromagnetic stirring[J]. Materials & Design, 2016, 95: 398-409.

[87] Zhang Y, Wu G H, Liu W C, et al. Effects of processing parameters on microstructure of semi-solid slurry of AZ91D magnesium alloy prepared by gas bubbling[J]. Transactions of Nonferrous Metals Society of China, 2015, 25(7): 2181-2187.

[88] Zhang Y, Wu G H, Liu W C, et al. Preparation and rheo-squeeze casting of semi-solid AZ91–2 wt% Ca magnesium alloy by gas bubbling process[J]. Journal of Materials Research, 2015, 30(6): 825-832.

[89] Wang C L, Wu G H, Sun M, et al. Formation of non-dendritic microstructures in preparation of semi-solid Mg-RE alloys slurries: roles of RE content and cooling rate[J]. Journal of Materials Processing Technology, 2020, 279: 116545.

[90] Fang X G, Wu S S, Lü S L, et al. Microstructure evolution and mechanical properties of quasicrystal-reinforced Mg-Zn-Y alloy subjected to ultrasonic vibration[J]. Materials Science and Engineering: A, 2017, 679: 372-378.

[91] Chang Z Y, Deng Q C, Lan Q, et al. Microstructure and mechanical properties of Mg-Gd-Y-Zn-Zr alloy prepared by rheo-diecasting[J]. Materials Science and Engineering: A, 2022, 848: 143287.

[92] Guo H M, Zhang A S, Hu B, et al. Refining microstructure of AZ91 magnesium alloy by introducing limited angular oscillation during initial stage of solidification[J]. Materials Science and Engineering: A, 2012, 532: 221-229.

[93] Yang X J, Yu A S, Zheng B, et al. Microstructure evolution of a semisolid magnesium alloy slurry obtained via an internal rapid cooling stirring process (IRCSP)[J]. Metallurgical and Materials Transactions B, 2020, 51(6): 2895-2904.

[94] Xing B, Li Y D, Ma Y, et al. Microstructure of partially remelted billet of AM60 alloy prepared with self-inoculation method[J]. Transactions of Nonferrous Metals Society of China, 2010, 20(9): 1622-1629.

[95] Meng Y, Chen Q, Sugiyama S, et al. Effects of reheating and subsequent rapid cooling on microstructural evolution and semisolid forming behaviors of extruded Mg-8.20Gd-4.48Y-3.34Zn-0.36Zr alloy[J]. Journal of Materials Processing Technology, 2017, 247: 192-203.

[96] Jiang J F, Wang Y, Qu J J, et al. Preparation and thixoforging of semisolid billet of AZ80 magnesium alloy[J]. Transactions of Nonferrous Metals Society of China, 2010, 20(9): 1731-1736.

[97] Meng Z Y, Wang C, Hua Z M, et al. Achieving extraordinary thermal stability of fine-grained structure in a dilute magnesium alloy[J]. Materials Research Letters, 2022, 10(12): 797-804.

[98] Pan H C, Kang R, Li J R, et al. Mechanistic investigation of a low-alloy Mg-Ca-based extrusion alloy with high strength-ductility synergy[J]. Acta Materialia, 2020, 186: 278-290.

[99] 蒋文明, 樊自田. 铝/镁合金特种精密铸造技术的研究进展及发展趋势[J]. 金属加工 (热加工), 2023(2): 1-9.
Jiang W M, Fan Z T. Research progress and development trend of special precision casting technology for Al alloy and Mg alloy[J]. MW Metal Forming, 2023(2): 1-9(in Chinese).

[100] 贾海龙, 丁家源, 查敏, 等. 镁合金压力下凝固技术研究进展[J]. 当代化工研究, 2022(20): 148-151.
Jia H L, Ding J Y, Zha M, et al. Current research progress on solidification techniques under pressure for magnesium alloys[J]. Modern Chemical Research, 2022(20): 148-151(in Chinese).

[101] Berman T D, Pollock T M, Jones J W. Microstructure and texture through thixomolding and thermomechanical processing and the role of $Mg_{17}Al_{12}$ particles[J]. Metallurgical and Materials Transactions A, 2016, 47(6): 3125-3136.

[102] Li Z X, Li D J, Zhou W K, et al. Characterization on the formation of porosity and tensile properties prediction in die casting Mg alloys[J]. Journal of Magnesium and Alloys, 2022, 10(7): 1857-1867.

[103] Jiang J F, Chen G, Wang Y, et al. Microstructure and mechanical properties of thin-wall and high-rib parts of AM60B Mg alloy formed by double control forming and die casting under the optimal conditions[J]. Journal of Alloys and Compounds, 2013, 552: 44-54.

[104] Decker R, Lebeau S, Berman T, et al. Thermomechanical processing of thixomolded alloys[C]//Proceedings of the Magnesium Technology 2017. Berlin: Springer, 2017: 235-243.

[105] Song J, Xiong S M, Li M, et al. In situ observation of tensile deformation of high-pressure die-cast specimens of AM50 alloy[J]. Materials Science and Engineering: A, 2009, 520(1-2): 197-201.

[106] Zhao Z D, Chen Q, Huang S H, et al. Microstructure and tensile properties of AM50A magnesium alloy prepared by recrystallisation and partial melting process[J]. Transactions of Nonferrous Metals Society of China, 2010, 20(9): 1630-1637.

[107] Anthony B T, Dowdell B G, Miller V M. Evolution of the intermetallic particle distribution in thixomolded magnesium alloys [C]//Proceedings of the Magnesium Technology 2019. Berlin: Springer, 2019: 193-197.

[108] 杨丽景, 卫英慧, 侯利锋, 等. 触变成型镁合金铸件热处理显微组织的演变[J]. 材料热处理学报, 2010, 31(4): 100-104.
Yang L J, Wei Y H, Hou L F, et al. Effect of heat treatment on microstructure of thixoformed AZ91D magnesium alloy[J]. Transactions of Materials and Heat Treatment, 2010, 31(4): 100-104(in Chinese) .

[109] Zhang Y, Li X P, Sun S P, et al. Microstructure evolution and mechanical properties of rheo-squeeze casting AZ91-Ca alloy during heat treatment[J]. China Foundry, 2017, 14(6): 485-491.

[110] 张婷, 赵宇宏, 陈利文, 等. 触变注射成形法制备石墨烯纳米片增强镁基复合材料[J]. 金属学报, 2019, 55(5): 638-646.
Zhang T, Zhao Y H, Chen L W, et al. Graphene nanoplatelets reinforced magnesium matrix composites fabricated by thixomolding[J]. Acta Metallurgica Sinica, 2019, 55(5): 638-646(in Chinese).

[111] Zhang Y F, Liu Y B, Cao Z Y, et al. Mechanical properties of thixomolded AZ91D magnesium alloy[J]. Journal of Materials Processing Technology, 2009, 209(3): 1375-1384.

[112] Chen L W, Zhao Y H, Li M X, et al. Reinforced AZ91D magnesium alloy with thixomolding process facilitated dispersion of graphene nanoplatelets and enhanced interfacial interactions[J]. Materials Science and Engineering: A, 2021, 804: 140793.

[113] Czerwinski F, Pinet P J, Overbeeke J. The influence of primary solid content on the tensile properties of a thixomolded AZ91D magnesium alloy[M]//Hryn J N. Magnesium technology 2001. Warrendale: TMS, 2001.

[114] Decker R. The technology and commercialization of thixomolding®[J]. Solid State Phenomena, 2012, 192-193: 47-57.

[115] Yang L Q, Kang Y L, Zhang F, et al. Rheo-diecasting of AZ91D magnesium alloy by taper barrel rheomoulding process[J]. Transactions of Nonferrous Metals Society of China, 2010, 20(6): 966-972.

[116] Mansoor B, Mukherjee S, Ghosh A. Microstructure and porosity in thixomolded Mg alloys and minimizing adverse effects on formability[J]. Materials Science and Engineering: A, 2009, 512(1-2): 10-18.

[117] Czerwinski F. The generation of Mg-Al-Zn alloys by semisolid state mixing of particulate precursors[J]. Acta Materialia, 2004, 52(17): 5057-5069.

[118] Lebeau S, Decker R. Microstructural design of thixomolded magnesium alloys[C]//Proceedings of the 5th International Conference on Semi-solid Processing of Alloys and Composites. Zurich: Trans

Tech Publications, 1998: 301-306.

[119] Jiang J F, Wang Y, Li Y F, et al. Microstructure and mechanical properties of the motorcycle cylinder body of AM60B magnesium alloy formed by combining die casting and forging[J]. Materials & Design, 2012, 37: 202-210.

[120] Jiang J F, Wang Y, Qu J J. Effect of process parameters on microstructure and properties of AM50A magnesium alloy parts formed by double control forming[J]. Transactions of Nonferrous Metals Society of China, 2014, 24(2): 321-333.

[121] 高志恒. 镁合金的腐蚀特性及防护技术[J]. 表面技术, 2016, 45(3): 169-177.
Gao Z H. Corrosion characteristics and protection techniques for Mg alloys[J]. Surface Technology, 2016, 45(3): 169-177(in Chinese).

[122] Zander D, Schnatterer C. The influence of manufacturing processes on the microstructure and corrosion of the AZ91D magnesium alloy evaluated using a computational image analysis[J]. Corrosion Science, 2015, 98: 291-303.

[123] Zhang Y, He C Y, Yu Q, et al. Nacre-like surface nanolaminates enhance fatigue resistance of pure titanium[J]. Nature Communications, 2024, 15: 6917.

[124] Patel H A, Chen D L, Bhole S D, et al. Low cycle fatigue behavior of a semi-solid processed AM60B magnesium alloy[J]. Materials & Design, 2013, 49: 456-464.

[125] Patel H A, Chen D L, Bhole S D, et al. Cyclic deformation and twinning in a semi-solid processed AZ91D magnesium alloy[J]. Materials Science and Engineering: A, 2010, 528(1): 208-219.

[126] Cepeda-Jiménez C M, Castillo-Rodríguez M, Pérez-Prado M T. Origin of the low precipitation hardening in magnesium alloys[J]. Acta Materialia, 2019, 165: 164-176.

[127] Amberger D, Eisenlohr P, Göken M. Microstructural evolution during creep of Ca-containing AZ91[J]. Materials Science and Engineering: A, 2009, 510-511: 398-402.

[128] 张明全, 徐超, 张旭, 等. Ca、Al 质量比对 Mg-Al-Ca-Mn 合金组织和性能的影响[J]. 特种铸造及有色合金, 2023, 43(3): 296-299.
Zhang M Q, Xu C, Zhang X, et al. Effects of Ca,Al mass ratio on microstructure and properties of Mg-Al-Ca-Mn alloy[J]. Special Casting and Nonferrous Alloys, 2023, 43(3): 296-299(in Chinese).

[129] Tsukeda T, Maehara A, Saito K, et al. Mechanical properties and microstructure of heat-resistant Mg-Al-Ca alloys formed by thixomolding[C]//Proceedings of the Magnesium Technology 2000. New York: Wiley, 2000: 395-402.

[130] Prasad S V S, Prasad S B, Verma K, et al. The role and significance of magnesium in modern day research-a review[J]. Journal of Magnesium and Alloys, 2022, 10(1): 1-61.

[131] Suzuki A, Saddock N D, Jones J W, et al. Solidification paths and eutectic intermetallic phases in Mg-Al-Ca ternary alloys[J]. Acta Materialia, 2005, 53(9): 2823-2834.

[132] Cai H S, Wang Z Z, Liu L, et al. Regulation mechanism of cooling rate and RE (Ce,Y,Gd) on $Mg_{17}Al_{12}$ in AZ91 alloy and it's role in fracture process[J]. Journal of Materials Research and Technology, 2022, 19: 3930-3941.

[133] 宿彦京, 付华栋, 白洋, 等. 中国材料基因工程研究进展[J]. 金属学报, 2020, 56(10): 1313-1323.
Su Y J, Fu H D, Bai Y, et al. Progress in materials genome engineering in China[J]. Acta Metallurgica Sinica, 2020, 56(10): 1313-1323(in Chinese).

[134] Zhu Y B, Midson S P. The status of magnesium injection molding in China[J]. Solid State Phenomena, 2019, 285: 436-440.

[135] Liu B, Yang J, Zhang X Y, et al. Development and application of magnesium alloy parts for automotive OEMs: a review[J]. Journal of Magnesium and Alloys, 2023, 11(1): 15-47.

[136] The Japan Steel Works. Technical review[EB/OL]. (2023-11-01) [2025-08-02]. https://www.jsw.co.jp/ja/technology/technical_review/2023/.

[137] 查敏, 顾焘, 马品奎, 等. 轻合金大型一体化结构部件压铸成形技术研究进展[J]. 特种铸造及有色合金, 2024, 44(8): 1009-1022.
Zha M, Gu T, Ma P K, et al. Research progress in die-casting technology for large-scale integrated structural components of light alloy[J]. Special Casting and Nonferrous Alloys, 2024, 44(8): 1009-1022(in Chinese).

[138] 压铸周刊. 全球首台 3000T 镁合金半固态双射台注射成型机投产 [EB/OL]. (2024-12-29) [2025-08-02]. https://www.yzweekly.com/xyzd/qiche/9738.htm.
Yzweekly. The world's first 3000-ton magnesium alloy semi-solid dual-shot injection molding machine commissioned[EB/OL]. (2024-12-29)[2025-08-02].https://www.yzweekly.com/xyzd/qiche/9738.htm(in Chinese).

[139] The Japan Steel Works. Magnesium injection molding machine [EB/OL]. [2025-08-02]. https://www.jsw.co.jp/en/product/business/molding_machine/mm_1600/.

[140] 伊之密. 伊之密镁合金半固态设备, 助力开启新能源汽车 20 万吨级镁应用[EB/OL]. (2024-11-29) [2025-08-02]. https://www.yizumi.com/news/product/tms/tms-20241129.
YIZUMI. YIZUMI's semi-solid magnesium alloy equipment drives 200,000-ton-level magnesium application in new energy vehicles[EB/OL]. (2024-11-29) [2025-08-02]. https://www.yizumi.com/news/product/tms/tms-20241129(in Chinese).

[141] 伯乐智能. 高能镁合金半固态 注射成型机[EB/OL]. [2025-08-02]. https://www.bole-machinery.com/contents/10/6.html.
BOLE. High energy magnesium alloy semi-solid injection molding machine[EB/OL]. [2025-08-02]. https://www.bole-machinery.com/contents/10/6.html(in Chinese).

[142] 伯乐智能. 4 000 t 半固态镁合金注射成型机启运, 伯乐智能助力合作伙伴技术升级[EB/OL]. (2025-01-16) [2025-08-02]. https://www.bole-machinery.com/contents/17/38.html.
BOLE. 4 000 t semi-solid magnesium alloy injection molding machine starts shipping, Bole Intelligent helps partners to upgrade technology[EB/OL]. (2025-01-16) [2025-08-02]. https://www.bole-machinery.com/contents/17/38.html(in Chinese).

[143] 伊之密. 上海压铸展 Day2 | 伊之密 6 600 T 半固态镁合金注射成型机成功交付, 开启大型一体化镁合金成型时代[EB/OL]. (2025-07-17) [2025-08-02]. https://www.yizumi.com/news/product/tms/CHINA-DIECASTING-20250717.
YIZUMI. Day 2 of Shanghai die casting exhibition|Yizumi's 6 600 T semi-solid magnesium alloy injection molding machine successfully delivered, marking the beginning of the era of large-scale integrated magnesium alloy forming[EB/OL]. (2025-07-17) [2025-08-02]. https://www.yizumi.com/news/product/tms/CHINA-DIECASTING-20250717(in Chinese).

[144] 冯维彦, 杨锐. TPI 模组——镁合金半固态成型的一种新工艺

[C]//重庆市铸造年会. 重庆: 重庆市机械工程学会, 2025: 357-362.
Feng W Y, Yang R. TPI module—a new process for semi-solid forming of magnesium alloys[C]//Proceedings of the Chongqing Foundry Annual Conference. Chongqing: Chongqing Society of Mechanical Engineering, 2025: 357-362(in Chinese).

[145] 宜镁泰. 镁合金制品[EB/OL]. [2025-08-02]. http://www.jsymt.com.cn/list/74_1.html.
EMT. Magnesium alloy products[EB/OL]. [2025-08-02]. http://www.jsymt.com.cn/list/74_1.html(in Chinese).

[146] 陶永亮, 任登安, 赵勇, 等. 镁合金半固态注射成型在笔记本电脑上应用[J]. 模具制造, 2022, 22(2): 48-51.
Tao Y L, Ren D A, Zhao Y, et al. Application of magnesium alloy semi-solid injection molding in notebook computer industry[J]. Die & Mould Manufacture, 2022, 22(2): 48-51(in Chinese).

[147] 伊之密. 内饰件[EB/OL]. [2025-08-02]. https://www.yizumi.com/tms/interior-parts.
YIZUMI. Interior parts[EB/OL]. [2025-08-02]. https://www.yizumi.com/tms/interior-parts(in Chinese).

[148] 伯乐智能. 伯乐智能携手产学研联盟全球首例 20 寸镁合金半固态轮毂, 引领新能源汽车轻量化革命[EB/OL]. (2025-04-03) [2025-08-02]. https://www.bole-machinery.com/contents/17/56.html.
BOLE. Bole Intelligent and industry-university-research alliance jointly developed the world's first 20-inch magnesium alloy semi-solid wheel hub, leading the lightweight revolution of new energy vehicles[EB/OL]. (2025-04-03) [2025-08-02]. https://www.bole-machinery.com/contents/17/56.html(in Chinese).

[149] 压铸周刊. 4 000 t 超大型镁合金半固态注射成型机[EB/OL]. (2024-05-23) [2025-08-02]. https://www.yzweekly.com/cd2016/9144.htm.
Yzweekly. 4 000-ton ultra-large-scale magnesium alloy semi-solid injection molding machine[EB/OL]. (2024-05-23) [2025-08-02]. https://www.yzweekly.com/cd2016/9144.htm(in Chinese).

Research progress in semi-solid thixomolding technology of magnesium alloys

JIANG Jufu[1, 2, *], ZHANG Jinze[1, 2], CUI Jingbo[1, 2], WANG Ying[3], ZHANG Xiaodong[1, 2]

(1. School of Materials Science and Engineering, Harbin Institute of Technology, Harbin 150001, China;
2. National Key Laboratory of Precision Hot Forming, Harbin Institute of Technology, Harbin 150001, China;
3. School of Mechatronics Engineering, Harbin Institute of Technology, Harbin 150001, China)

Abstract: Because of their low density, magnesium alloys are strategically valuable for lightweight transportation equipment. Nevertheless, conventional casting processes encounter technical challenges, including shrinkage porosity, gas porosity, and oxidation. Thixomolding—an innovative forming process based on semi-solid processing principles-provides a viable approach to overcome these limitations. This work systematically elucidates the fundamental mechanisms of thixomolding, reveals the formation principles of non-dendritic microstructures and preparation methodologies for semi-solid slurries, and comprehensively analyzes the technology's enhancement effects on the comprehensive properties of AZ-series and AM-series magnesium alloys. Currently, this technology has achieved industrial-scale implementation in consumer electronics and transportation equipment sectors. Supported by next-generation higher-tonnage forming systems, its applications are progressively expanding to large-scale structural component manufacturing, demonstrating exceptional development potential within the new energy vehicle industry. Among other crucial areas, future research should focus on defect control mechanisms for large-scale components, multi-physics coupled modeling of the forming process, and composition design of innovative thixomolded magnesium alloys.

Keywords: magnesium alloy; semi-solid forming; thixomolding; mechanical properties; lightweight

Received: 2025-08-21; **Accepted**: 2025-10-10; **Published Online**: 2025-11-06 11:48

URL: link.cnki.net/urlid/11.2625.V.20251105.1731.005

Foundation items: National Natural Science Foundation of China (52431004); Dongguan City Key Field Research and Development Projects (20241200300012)

* **Corresponding author**. E-mail: jiangjufu@hit.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2026.0063

氧化亚氮推进技术发展现状与趋势

魏祥庚[1,*]，谢远[1]，路伟杰[1]，刘雄峰[1]，高玉闪[2]，赵昊阳[1]
(1. 西北工业大学 固体推进全国重点实验室，西安 710072；2. 西安航天动力研究所，西安 710100)

摘 要：氧化亚氮(N_2O)作为一种无毒绿色、稳定安全的自增压推进剂，广泛应用于单组元和双组元火箭发动机，尤其在姿轨控动力领域具有重要的研究价值与应用潜力。对几类氧化亚氮推进系统的发展历程与研究现状进行综述，总结氧化亚氮推进系统的特点，指出推动其工程应用的关键技术，并进一步展望了氧化亚氮推进技术在姿轨控动力方向的未来研究重点，以期为航天姿轨控动力系统的创新与工程实践提供参考。

关 键 词：姿轨控发动机；绿色推进剂；氧化亚氮；关键技术
中图分类号：V511
文献标志码：A **文章编号**：1001-5965(2026)10-3350-19

姿轨控动力系统是航天器系统的重要组成部分，其主要作用是实现航天器在轨姿态精确控制与轨道调整。目前，国内外已应用多种推力级别的姿轨控发动机完成航天任务[1-3]。当前工程应用中多采用传统化学推进剂(如肼类)，虽技术成熟，但仍面临毒性高、储存条件苛刻、环境影响较大等问题。随着未来空间任务向长寿命、高机动、绿色环保及高可靠性方向发展，对姿轨控发动机提出了更严苛的要求，包括更紧凑的结构尺寸、更高的比冲、更多次的启动能力、更长的在轨服役时间，以及推进剂的低毒化和环境友好性。在此背景下，氧化亚氮(N_2O)因其绿色无毒、饱和蒸汽压高、价格低廉、安全稳定等优点而展现出显著优势，可用于设计自增压系统以大幅降低系统结构质量，有望在航天器姿轨控发动机中得到广泛应用[4]。

近几年，氧化亚氮被成功应用于多款发动机，推力等级包含1～1 068 N不等，用于CubeSats[5]、大型行星际转移级[6]、火星着陆器和空间站[7]等一系列项目。国外黎明宇航公司(Dawn Aerospace)采用氧化亚氮/丙烯作为推进剂，研制了适用于多型卫星的推进系统。该推进系统已应用于AstroForge与OrbAstro(深空小行星采矿卫星)、Pixxel(高光谱成像星座)及Lynk(卫星直连地面移动通信星座)等任务[5]。脉冲公司(Impulse Space)采用氧化亚氮/乙烷推进剂，开发了用于轨道运载任务的Mira航天器，该航天器搭载空间监视有效载荷，并纳入美国太空部队“战术响应空间”(TacRS)计划[6]。近期，该公司与相对论公司(Relativity Space)合作，联合研制一款基于氧化亚氮推进剂的火星着陆器[7]。与此同时，浩瀚太空公司(Vast Space)亦采用氧化亚氮/乙烷作为推进剂，研制了Orbiter轨道转移飞行器，该飞行器还可作为Launcher轻型运载火箭的第三级使用[8]。此外，氧化亚氮还发展出了电加热、催化分解、单组元等多种推进形式[4]，展现出了不俗的发展潜力。

本文通过梳理氧化亚氮发动机的发展历程，提炼研究重点及关键技术，总结氧化亚氮推进技术发展方向，以期对姿轨控发动机的技术发展提供借鉴。

收稿日期：2026-01-27；**录用日期**：2026-04-24；**网络出版时间**：2026-05-15 14:10
网络出版地址：link.cnki.net/urlid/11.2625.V.20260515.1009.001
基金项目：陕西省自然科学基础研究计划(2021JC-14)
*通信作者. E-mail：realysnow@nwpu.edu.cn

引用格式：魏祥庚，谢远，路伟杰，等. 氧化亚氮推进技术发展现状与趋势[J]. 北京航空航天大学学报，2026，52(10)：3350-3368.
Wei X G，Xie Y，Lu W J，et al. Current status and trend of nitrous oxide propulsion technology[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3350-3368 (in Chinese).

1　基于氧化亚氮的动力系统主要研究进展

目前, 氧化亚氮推进剂主要应用于以下几类推进系统: 电加热式、催化分解式、复合推进剂单组元式、双组元液体推进系统。此外, 其在固液混合推进系统中也有应用。根据对现有氧化亚氮发动机技术的梳理(如表 1 所示), 德国、法国、美国、日本等国均已开展相关研究。本文聚焦液体火箭发动机方向, 因此, 固液混合推进系统不在讨论范围内, 后续将重点综述上述几类液体推进系统。

表 1　氧化亚氮发动机研究概况

Table 1　Research overview of nitrous oxide engines

年份	研究机构	推进剂组成	发动机推力/N	应用
2020	德国-航空航天中心[4]	N_2O/乙烷	1, 22	
2000～2010	法国-机电一体化研究中心[9]	N_2O/PE	1 000	探空火箭
2020	新西兰-黎明宇航公司[5]	N_2O/丙烯	0.49～1.35, 6.1～16.7	B1和B20推进器, 卫星推进
2000～2010	美国-斯坦福大学[10]	N_2O/石蜡		“游隼”探空火箭
2010～2020	美国-桑迪亚国家实验室[11]	N_2O/乙醇	577	空间推进
2010～2020	美国-Firester公司[12]	NOFBX	450	推进动力-火星单级入轨火箭
2000～2010	美国-阿拉巴马大学	N_2O/丙烷	200	空间推进
2000～2010	美国-阿拉巴马大学[13]	N_2O/乙醇	1	卫星推进
2000～2010	美国-鲁坦及美国缩尺复合体公司[14]	N_2O/HTPB	73 500	“太空船一号”载人亚轨道飞行器
2010～2020	美国-斯卡尔德复合材料公司和英国-维珍银河公司[15]	N_2O/HTPB	310 000	“太空船二号”载人亚轨道飞行器
2020	美国-脉冲公司[16]	N_2O/乙烷	26	空间推进
2020	美国-脉冲公司[17]	N_2O/乙烷	800	空间推进
2010～2020	日本-宇宙航空研究开发机构JAXA[18]	N_2O/乙醇	700	高超声速飞行器火箭验证平台
2010～2020	日本-宇宙航空研究开发机构JAXA[19]	N_2O/乙醇	2 000	上面级主发动机
2010～2020	日本-宇宙航空研究开发机构JAXA[19]	N_2O/乙醇	20	系统推进
2010～2020	英国-AEL公司[20]	N_2O/乙醇(预混)	600	
2000～2010	英国-萨里大学[21]	N_2O	5.00×10^{-4}	“UoSAT-12”卫星推进
2000～2010	中国-北京航空航天大学[22]	N_2O/HTPB	7 000	“北航二号”探空火箭
2000～2010	中国-北京航空航天大学[23]	N_2O	1	卫星推进
2020～2023	中国-西北工业大学[24]	N_2O/煤油	500	轨控定推力发动机
2023～2025	中国-西北工业大学[25]	N_2O/煤油	80～400	轨控变推力发动机
2024	中国-北京交通大学[26]	N_2O/煤油		
2025	中国-北京理工大学[27]	N_2O/乙烷		
2009	中国-国防科学技术大学[28]	N_2O/乙醇		

注: HTPB为端羟基聚丁二烯(hydroxyl-terminated polybutadiene); PE为聚乙烯(polyethylene))。

1.1　氧化亚氮电加热推进系统

氧化亚氮电加热推进系统利用其高饱和蒸汽压特性实现自增压供应, 通过电阻加热至约 800 ℃后从喷管排出, 系统原理如图 1 所示[21]。该推力器不仅具备氮气推进系统响应快、结构简单、成本低的特点, 而且结构体积更小、比冲性能更高, 可达 148 s。该系统于 1999 年在 UoSAT-12 小型卫星上成功应用。该卫星同时配备了氧化亚氮与氮气 2 种推进方式, 其中, 氧化亚氮虽仅占总推进剂质量的 1/3, 却贡献了 2/3 的总冲量。由此可见, 氧化亚氮电加热推进系统在性能上表现出显著优势。

1.2　氧化亚氮催化分解推进系统

氧化亚氮催化分解推进系统的比冲为 160～190 s, 高于氮气冷气推进系统与氧化亚氮电加热推进系统的比冲(148 s)。该系统具备自增压能力, 结构简单, 非常适用于微小卫星的动力系统。氧化亚氮流经催化床, 受热催化分解为氮气与氧气, 之后经过喷管加速喷出形成推力[28], 催化分解推进器工作原理示意如图 2 所示。

氧化亚氮热分解反应的化学方程式为

$$N_2O \rightarrow N_2 + 0.5O_2 + 1\,864\ \text{kJ/kg}$$

该反应活化能较高, 约为 250 kJ/mol, 热分解温度不低于 520 ℃。加入催化剂可显著降低反应活化能、促进分解过程, 如图 3 所示, 图中, E_a 为活化能, ΔH_r° 为反应焓。目前常用的催化剂在约 200 ℃即可表现出良好的催化活性。

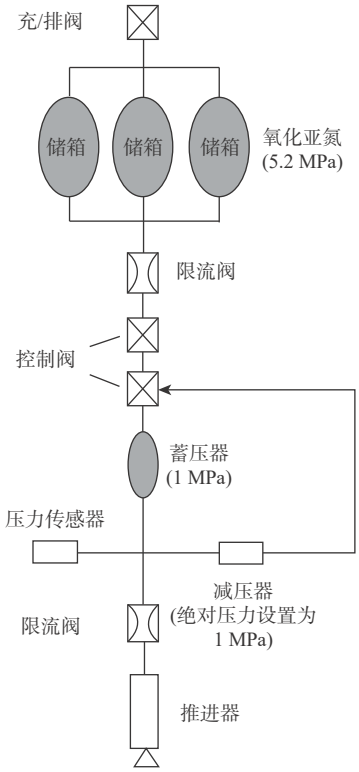


图1 UoSAT-12氧化亚氮系统原理图[21]

Fig. 1 Schematic diagram of nitrous oxide system on UoSAT-12 [21]

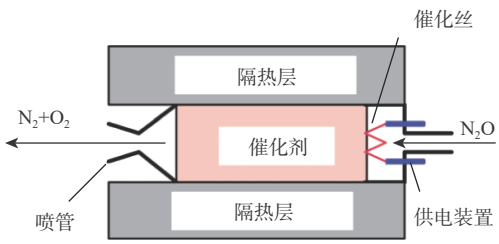


图2 氧化亚氮催化分解推进器工作原理示意图[29]

Fig. 2 Schematic diagram of catalytic decomposition thruster for nitrous oxide [29]

研究表明，推力器热损失、催化床的床载与床长设计及其与推力器的匹配程度等因素，均对系统性能具有重要影响[30-33]。可采取以下措施提升催化分解推进器的加热性能：采用将加热器埋在催化剂中的内加热方式提高加热效率；优化推力器结构与尺寸以减少预热能量损失；在长时间停机后重启时，适当增加加热功率以缩短点火延迟时间[34-35]。

1.3 氧化亚氮复合推进剂单组元推进系统

液态氧化亚氮对大多数有机物及烃类燃料具有良好的溶解能力，所形成的混合物表现出良好的相容性与稳定性。研究人员通过将烃类、醇类或氨类燃料溶解于氧化亚氮，并经特殊工艺制备出氧化亚氮基复合推进剂（nitrous oxide fuel blend, NOFBX）。该类推进剂兼具双组元推进剂的高比冲优势与单组元推进剂的系统结构简单的特性[36]。

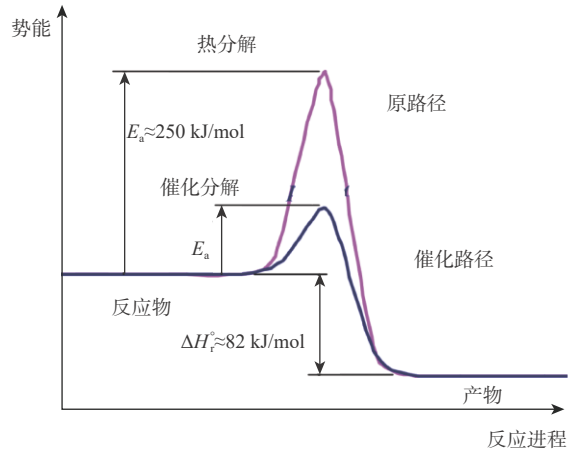


图3 氧化亚氮催化分解[29]

Fig. 3 Nitrous oxide catalytic decomposition [29]

2000年，美国Firestar公司为了完成美国国家航空航天局（National Aeronautics and Space Administration, NASA）火星技术，应用项目中单级入轨火箭（single-stage-to-orbit, SSTO）携带火星样本返回任务率先提出NOFBX[12]，于2012年成功研制出氧化亚氮和C_2烃（含有2个碳原子的简单有机碳氢化合物）推进剂组合的445 N级发动机，并在美国国家空间站完成真空点火实验[37]，如图4所示。除C_2烃之外，甲烷、丙烷及乙醇[38]等也可以作为NOFBX中的燃料。

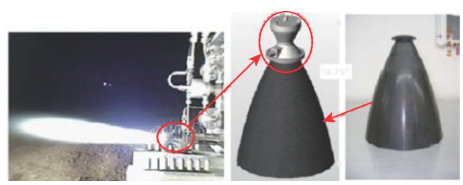


图4 NOFBX™445 N燃烧室地面点火测试（燃烧室压强为2.06 MPa）[37]

Fig. 4 Ground ignition test of the NOFBX™ 445 N (combustion chamber with a chamber pressure of 2.06 MPa)[37]

中国关于NOFBX的研究虽起步较晚，但目前已掌握了公斤级的氧化亚氮复合推进剂的制备技术，并成功开展了小推力点火实验[39-40]。同时，NOFBX基础燃烧特性的研究也取得不少进展，主要集中在推进剂的反应机理模型、着火延迟时间、自着火特性等方面[41-46]。

单组元复合推进剂制备困难，为简化研究难度，德国宇航中心（Deutsches Zentrum für Luft- und Raumfahrt e.V., DLR）[47]选择气态N_2O与气态乙烯（C_2H_4）预混制成NOFBX，并进行了水冷点火实验，如图5所示。NOFBX不仅可用于火箭发动机，也有学者尝试将其用于爆轰发动机。Bangalore等[48]分别进行了氧化亚氮与乙烯、乙炔、乙炔-氧混合物等复合推进剂的爆燃转爆轰等实验研究，发现在低

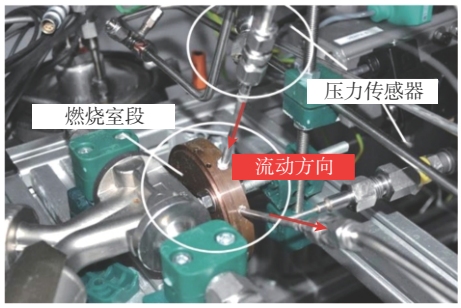


图 5　带水冷的 N_2O/C_2H_4 测试装置[47]

Fig. 5　N_2O/C_2H_4 test setup with cooling channel[47]

推进剂预混压强下, 复合推进剂的点火延迟时间会随预混压强上升而减少, 减少的速度较快, 且氧化亚氮在高温下的分解会使复合推进剂产生更高的爆轰压强。

研究表明, 与常规双组元推进系统相比, NOFBX 作为一种单组元推进剂, 具备自增压特性, 可显著降低推进系统的整体质量并简化结构。其在能量特性方面表现突出, 真空比冲可达 325 s, 密度比冲约为 2.5×10^5 (kg·s)m^3, 且燃烧过程平稳。此外, NOFBX 还具有无毒环保、推重比高(铝制推力室可达 100:1)、可重复起动、压强调节范围宽、可采用火花塞点火、与材料相容性好等优点, 展现出显著的发展潜力[49]。

尽管 NOFBX 性能优异, 但仍存在若干关键技术问题有待解决。首先是回火问题。NOFBX 饱和蒸汽压较高, 常温下易气化, 且气相燃烧速率极快, 在点火阶段易发生回火, 可能导致供应系统爆炸, 构成严重安全隐患。美国“机载发射辅助进入空间”项目在使用 NOFB37(氧化亚氮质量分数 70%的 N_2O/C_2H_2 复合推进剂)进行静态点火测试时, 4 次试验均发生回火爆炸, 最终导致项目中止[50]。

其次是冷却问题。在主动冷却过程中, 若 NOFBX 温度变化过大, 可能导致推进剂组分分离, 引发燃料裂解积碳或管路烧蚀等问题[51]。喷注器面板也易烧蚀, 在 Mayer 等[38] 的实验中就曾观察到因长时间工作导致的喷注器面板烧蚀现象。

最后是供给稳定性问题。尽管 NOFBX 具备自增压能力, 但在供给过程中, 储箱内压强变化会引起推进剂理化性质与组分的显著波动[52]。实际应用中常需通过精确控制储箱压强或温度以减缓该变化, 然而, 当前调控精度仍不足, 限制了其工程应用的可靠性。

综上所述, NOFBX 发动机具有比冲高、燃烧效率良好和供给系统简单等突出优势。若能在发动机设计过程中合理选择推进剂组分, 并系统解决回火、长效冷却与稳定供应等关键问题, 将有效提升发动机的稳定性与安全性, 使 NOFBX 有望成为小推力姿轨控发动机系统的优选推进剂。

1.4　氧化亚氮双组元推进系统

氧化亚氮在双组元推进系统中也有广泛的应用, 目前, 已开展与甲烷、乙烷、乙烷-乙烯(混合燃料)、丙烷、丙烯、乙醇、煤油等多种燃料组合的研究。

氧化亚氮/丙烷(nitrous oxide/propane, NOP)是一种无毒推进剂组合, 其化学反应活性较低, 可在太空环境中长期储存且不易降解, 有助于降低系统运营成本。在环境温度下, 氧化亚氮的饱和蒸汽压约为 5.1 MPa, 丙烷约为 0.76 MPa, 因此, 高压氧化亚氮可同时作为氧化剂和丙烷的加压气体, 实现自增压供应。N_2O 的自增压供应可采用 2 种方式: 从储箱底部以两相流形式供应[53-54], 或从储箱上部以气态形式供应[55]。Smith 等[53] 首次开展了 NOP 推进剂的地面论证实验, 采用火炬点火方式, 实现了平稳的点火与燃烧过程, 燃烧室压强范围覆盖 0.96～2.10 MPa, 验证了氧化亚氮自增压供应在液体火箭发动机中的可行性。在此基础上, Tiliakos 等[54] 为评估催化分解点火方案的适用性, 改进了 Smith[53] 所采用发动机的头部结构如图 6 所示, 发现燃烧室内压强振荡会导致燃烧效率下降, 而高流量、高流速的两相氧化亚氮流会加剧该不稳定现象。王栋等[55] 设计了图 7 所示的火炬式点火器, 研究了气态氧化亚氮分别与气态及液态丙烷的点火特性, 结果表明, 余氧系数对点火延迟时间具有显著影响。

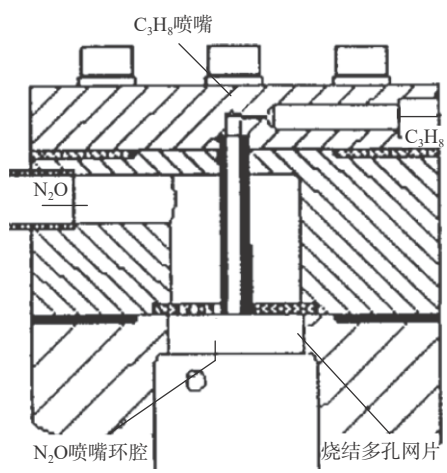
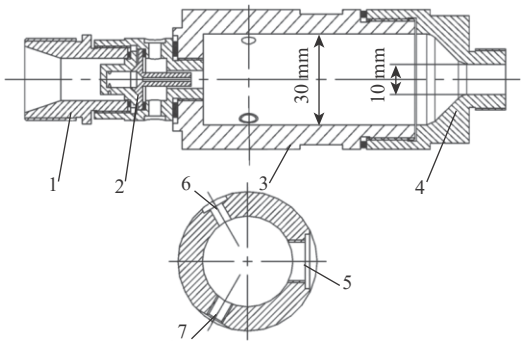


图 6　NOP 发动机喷嘴设计[54]

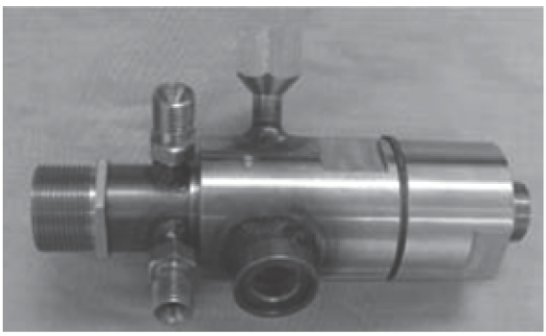
Fig. 6　Injector design for NOP rocket[54]

氧化亚氮/乙醇(nitrous oxide/ethanol, NOEL)是另一种常见的推进剂组合, 其比冲性能略低于 NOP 发动机, 但燃烧稳定性和燃烧效率更高, 更适用于较大推力的发动机, 如小型运载火箭的上面级发动机, 因此, NOEL 通常采用携带高压气源进行挤压式供应。日本宇宙航空研究开发机构(Japan



1. 直通接头；2. 同轴旋流式喷嘴；3. 燃烧室；4. 喷管；5. 火花塞基座；6. 压力传感器测孔；7. 温度传感器测孔。

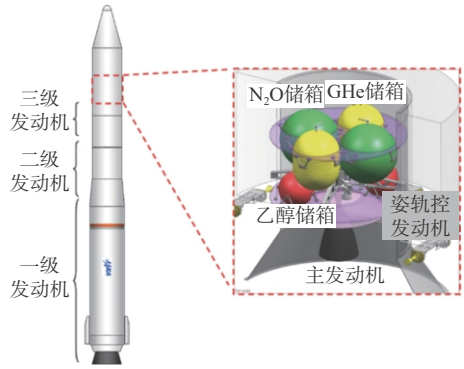
(a) 结构示意图



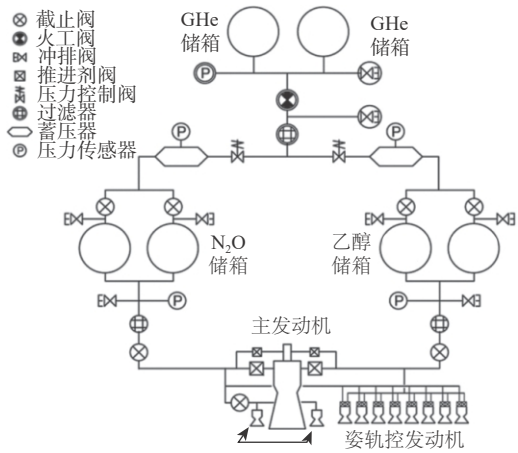
(b) 实物图

图 7 NOP 火炬点火器[55]

Fig. 7 NOP torch igniter schematic[55]



(a) 搭载NOEL推进系统的Epsilon原型机



(b) 系统原理图

图 8 配备 NOEL 推进系统的原型 Epsilon 运载火箭及其系统图[19]

Fig. 8 Prototype Epsilon launch vehicle equipped with NOEL propulsion system and its system diagram[19]

Aerospace Exploration Agency, JAXA)从 2003 年开始相关研究，进行了 NOEL 发动机点火实验[19]，该发动机计划用于 Epsilon 运载火箭的第三级，运载火箭和推进系统如图 8 所示，发动机海平面比冲为 225 s，海平面推力为 1.64 kN，室压为2 MPa。实验表明，氧化亚氮的喷射温度与挤压压强对燃烧稳定性有较大影响。Phillip 等[11]对不同混合比工况下的地面比冲进行了测试，测得结果介于 250～260 s 之间，与 NASA 化学平衡计算程序 CEA(chemical equilibrium with applications)的理论预测值吻合良好。邓德涛等[28]通过实验发现，NOEL 发动机可在宽范围余氧系数下稳定工作，且在余氧系数为 0.25～0.35 时燃烧效率较高。上述研究均采用火炬点火器实现点火。

除以上研究外，研究人员针对甲烷[56]、乙烷[27,57]、乙烷/乙烯(两者各占 50%)[58]、丙烯[5]等燃料进行了研究，其中，氧化亚氮/丙烯发动机已经实现了工程应用，如黎明宇航公司(Dawn Aerospace)研制的 B20 发动机，该型发动机已于 2021 年作为 D-Orbit 公司的 ION 卫星运载器的轨控发动机成功完成空间验证，实现了运载器的连续变轨，从而在不同轨道高度下释放不同的卫星[5]。2023 年 11 月 11 日，脉冲公司(Impulse Space)研发的 Mira 平台搭载 SpaceX 猎鹰 9 号火箭进入太空，在 Mira 平台上搭载着 8 个 26 N 级氧化亚氮/乙烷发动机[16]。

近年来，西北工业大学魏祥庚等采用煤油作为燃料，对氧化亚氮/煤油双组元推进系统开展了系统研究。在一项专利中详细阐述了该型发动机动力系统的工作原理[59]，并成功通过点火实验验证了自增压供应系统的能力与热补偿效果[25]。此外，团队还针对变推力工况下的复合冷却方案开展了实验研究，结果表明，该冷却策略在变推力条件下仍能有效维持发动机热平衡，验证了其工程可行性[24]。

总之，目前已开展多种推进剂组合的发动机工作特性研究。对比各推进剂组合及现阶段常用的双组元推进剂组合的性能，如表 2 所示，发现氧化亚氮推进剂组合燃温较高而比冲略低，在各种氧化亚氮推进剂组合中，氧化亚氮/乙炔比冲最高，氧化亚氮/乙烯稍次之，NOEL 比冲最低。然而，乙烯、乙炔安全性较差，其中，乙烯安全性高于乙炔[4]。在选择与氧化亚氮组合的燃料时，不仅要考虑比冲性能，还要综合考虑燃料的理化性质。对比几种常用燃料的饱和蒸汽压(如图 9 所示)，乙烯的饱和蒸汽压较高，接近氧化亚氮的饱和蒸汽压，氧化亚氮/

表 2　氧化亚氮双组元推进剂及其他推进剂的理论计算性能

Table 2　Calculated theoretical performance for nitrous oxide bipropellant and others

推进剂组合	燃烧温度/K	余氧系数α	特征速度C^*/(m·s^{-1})	燃气相对分子质量M_c/(kg·mol^{-1})	比冲I_{sp}/(m·s^{-1})
N_2O/C_2H_5OH	3 219.3	0.8	1 560.5	25.65	2 460.7
N_2O/CH_4	3 229.6	0.7	1 622.7	23.67	2 550.5
N_2O/C_2H_6	3 302.7	0.7	1 615.8	24.38	2 538.4
N_2O/C_2H_4	3 400.1	0.6	1 643.5	24.10	2 573.4
N_2O/C_2H_2	3 585.2	0.4	1 732.3	22.41	2 685.3
N_2O/C_3H_8	3 320.9	0.7	1 609.8	24.69	2 528.5
N_2O_4/MMH	3 384.3	0.6	1 730.6	20.40	2 725.7
LOX/LCH4	3 405.9	0.7	1 835.6	19.91	2 918.3
LOX/JP-10	3 631.0	0.7	1 755.5	22.98	2 777.4
LOX/LH2	2 694.7	0.4	2 417.1	8.95	3 794.6

注: 燃烧室压强约为7.09 MPa; 喷管出口及环境压强均为101 325 Pa; 计算结果为最佳比冲; 模型采用冻结流。燃气按理想气体处理, 燃烧为绝热过程, 喷管膨胀为等熵流动膨胀。

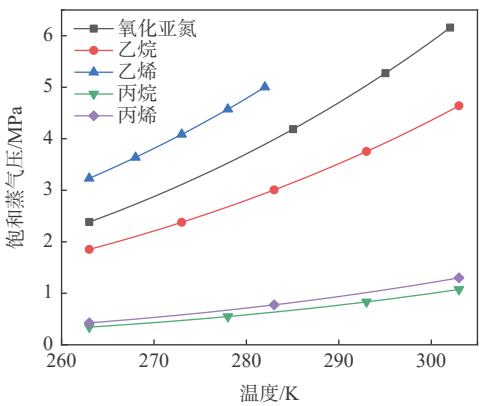


图 9　氧化亚氮、乙烷、乙烯、丙烷和丙烯的饱和蒸汽压

Fig. 9　Saturated vapor pressures of nitrous oxide, ethane, ethylene, propane and propylene

乙烯双高饱和蒸汽压的推进剂组合可以增加发动机工作时的稳定性, 被视为非常有潜力的自增压推进剂组合[54]。

在空间推进系统中, 与常规常温双组元推进剂(如 MMH/NTO)相比, 氧化亚氮基双组元推进剂具备多项优势: 绿色无毒、可自增压供应、系统结构简化、使用维护成本低, 且便于在轨长期贮存。未来, 采用该类推进剂的液体火箭发动机在姿轨控领域具有广阔的应用前景。

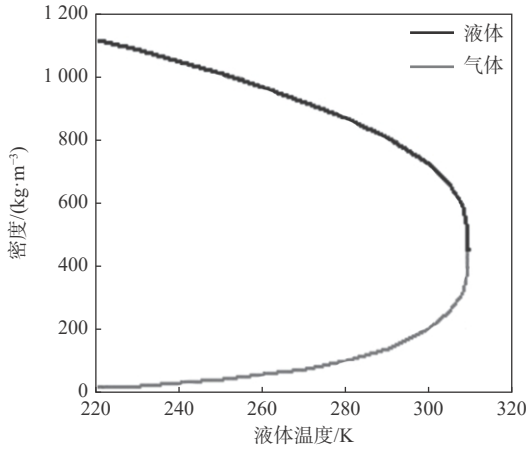
综上, 氧化亚氮推进系统的选择, 本质是在性能、复杂度、可靠性和成本之间权衡: 电加热式适用于追求简单与可靠的小型任务场景, 如立方星姿态控制; 催化分解式适用于需要中等性能且系统复杂度有限的小卫星推进场景, 但需重点关注催化剂寿命; 复合单组元式适用于对性能要求较高、同时希望降低系统复杂性的任务场景, 如上面级, 是一种有潜力的方案; 双组元系统则适用于主推进、需要高比冲和具备推力调节能力的任务场景, 如小型发射器, 是当前较为成熟方案。

2　氧化亚氮推进系统关键技术

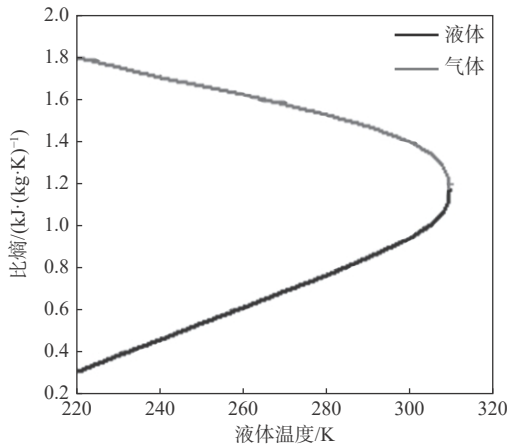
第 1 节总结了氧化亚氮基推进系统的主要研究进展, 可以看出, 尽管该推进剂优势显著, 但目前大多数研究仍处于地面试验阶段。为促进其应用, 梳理了亟待解决的关键技术问题, 主要包括: 氧化亚氮自增压模型和供应稳定控制技术、氧化亚氮催化分解技术、氧化亚氮发动机性能提升技术、氧化亚氮两相流喷注雾化技术、氧化亚氮发动机热防护技术及氧化亚氮发动机点火技术等内容。针对上述关键技术, 本文进一步提出了具体的研究思路, 为氧化亚氮推进技术发展提供参考。

2.1　氧化亚氮自增压模型和供应稳定控制技术

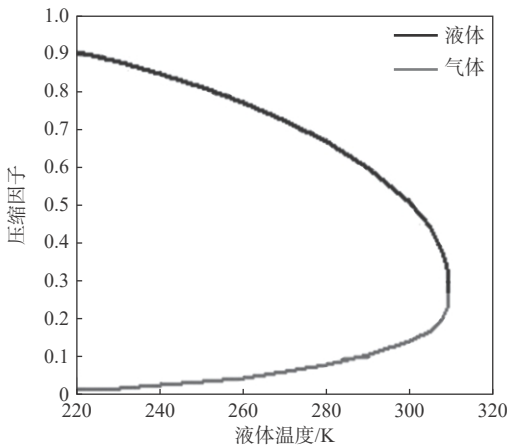
对于液体火箭发动机而言, 供给系统的质量通常占整个推进系统总质量的 25%～40%, 且复杂的供给系统往往会降低推进系统的可靠性, 而氧化亚氮自增压供给系统结构简单、质量较轻, 极大地降低了使用、维护成本[60-61]。氧化亚氮本身物性特殊, 其饱和蒸汽压、密度、比熵、压缩因子等物性参数随温度变化较大[62](饱和蒸汽压见图 9, 其余参数见图 10)。在 300 K 时, N_2O 的气相压缩因子为 0.512 0, 与压缩因子为 1 的理想气体偏差较大, 使理想气体模型不再满足; N_2O 的液相压缩因子为 0.143 4, 也不满足不可压液体(压缩因子为 0)的假设, 这给 N_2O 的建模带来了困难。且自增压排放过程中, 储箱内会发生复杂相变现象, N_2O 蒸发吸热会导致储箱内温度下降, 氧化亚氮的蒸发速率又与温度相关, 温度下降后蒸发速率即随之下降, 两者的耦合作用导致储箱内压强进一步下降, 在一些特殊工况下, 甚至会出现由于快速相变导致的闪蒸现象[63], 影响流量的稳定供应, 因此, 亟须建立准确的氧化亚氮自增压供给数学模型来预示储箱状态, 并采取合理的措施



(a) 密度



(b) 比熵



(c) 压缩因子

图 10 温度对氧化亚氮密度、比熵、压缩因子的影响[62]

Fig. 10 Influence of temperature on density, specific entropy, and compression factor of nitrous oxide[62]

实现推进剂流量稳定、高利用率供应。

氧化亚氮自增压供应预示研究主要采用计算流体力学(computational fluid dynamics, CFD)数值模拟和理论计算2种方法。鉴于储箱排放需要采用瞬态模拟,计算量偏大,目前该研究主要借助理论计算集总参数模型来开展。集总参数模型根据排放过程中是否处于平衡状态,分为“平衡模型”和“非平衡模型”。“平衡模型”是指建模过程中假设储箱内气液两相一直处于平衡状态,且温度均匀分布,该模型适用于排放初期及排放流量较小的工况。Whitmore等[62]基于“平衡模型”,并假设排空过程中,储箱内剩余氧化亚氮与已排出氧化亚氮的总熵保持守恒,即等于初始状态下氧化亚氮的总熵。在此假设基础上,利用两相热力学性质表,根据当前的质量与熵确定系统状态,并据此计算排空过程中储箱内氧化亚氮的压强、温度及质量等参数的演化规律。Fernandez[64]对比了理想气体方程(Raolt's Law)和非理想气体状态方程(Peng-Robinson)模型预测结果,非理想气体模型和实验结果更加相符。排放过程中储箱内的一部分液相会气化来补充排空的体积,受重力场或加速度及来自壁面的热流影响,气相部分会出现明显的温度分层现象。当氧化亚氮排放流量较大时,液相来不及转变为气相,导致储箱压强下降,处于非平衡状态。此时若采用平衡模型会产生较大的误差,因此,研究人员发展了“非平衡模型”。Dyer等[65]将气液平衡条件引入Span-Wagner状态方程,提出了新的多参数状态方程,开发了非平衡流动模型,计算结果与实验结果较为一致。Zilliac等[66]针对储箱内温度分层现象,提出了推进剂分区蒸发传热传质模型,利用该模型进行了氧化亚氮的蒸发速率计算,较为精确地预测储箱内氧化亚氮压强、温度等参数变化规律。Kardaś等[67]提出一种储箱内部流体动力学模型,将流体视为两相混合物,也能预测氧化亚氮供给过程储箱内参数的变化情况。宋长青等[68-69]基于Zilliac等[66]的三区假设,引入Peng-Robinson状态方程,建立了储箱自增压供给过程模型,并利用相关实验结果修正了经验参数。

研究人员针对NOFBX推进剂也开展了相关自增压建模研究。杨学森等[52]通过集总参数模型和CFD这2种方法对NOFBX推进剂的储箱自增压特性、系统流动特性及管路气化特性开展了研究,如图11所示,发现Helmholtz自由能方程可较好地描述流动特性,集总参数模型和CFD计算结果比较吻合。

在氧化亚氮自增压供给过程中,储箱内会出现压强、温度下降的现象,影响推进剂流量的稳定供应。尤其是当剩余液相氧化亚氮较少时,储箱压强过小,会导致发动机无法工作,进而造成推进剂利用率下降。为保持流量稳定并提高推进剂的利用率,国内外研究人员提出了多种自增压稳定供应方案。

为保持储箱压强稳定,提高推进剂的利用率,最简单的办法是对氧化亚氮直接进行加热,使其温

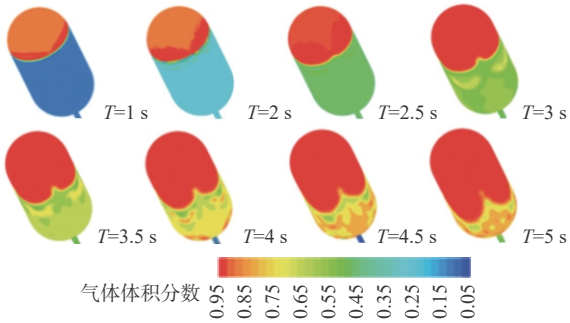
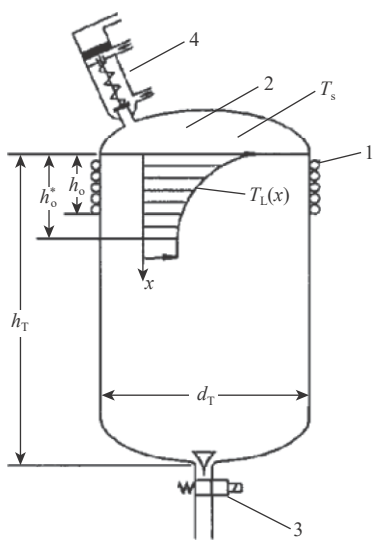


图 11　NOFBX 在储箱内流动数值模拟[52]

Fig. 11　Numerical simulation of NOFBX flowing out of a storage tank[52]

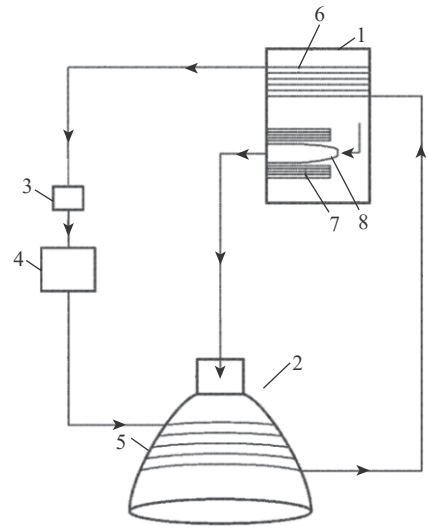
度保持稳定。可在储箱内部设置加热器来进行温度补偿[70], 在加热时要避免出现局部温度过高引起氧化亚氮分解。也有学者通过安装在储箱外表面的电加热器对推进剂进行温度控制, 如 Kozlov 等[71]提出一种用于低温液体推进剂预热的自增压方案, 使储箱在无增压气的情况下仍然有较高的初始压强, 原理如图 12 所示, 虽然在供应过程中未能稳定压强, 然而提高了推进剂的利用率。龚亮[72] 针对落压式系统的氧化亚氮储箱, 对比了对整个储箱壁面加热和只对液相壁面的加热效果, 并采用数值方法研究了不同加热功率值的稳压效果, 发现仅对液相壁面加热和提高加热功率能更好地提高氧化亚氮气态供应时的自增压性能。这种加热方式结构简单, 加热功率调节方便, 但当储箱出口推进剂的流量较大时, 所需的加热功率很大, 对供能系统提出较高的要求。为减小对供能系统的要求, 有学者提出利用燃烧室的热量对储箱进行加热[73], 系统如图 13 所示。冷却剂(非氧化亚氮)流经燃烧室的冷却通



1. 加热器; 2. 气相; 3. 推进剂主阀; 4. 安全阀。

图 12　通过加热实现自增压储箱稳压方案[71]

Fig. 12　Self-pressurized tanks pressure stabilizing scheme achieved through heating [71]



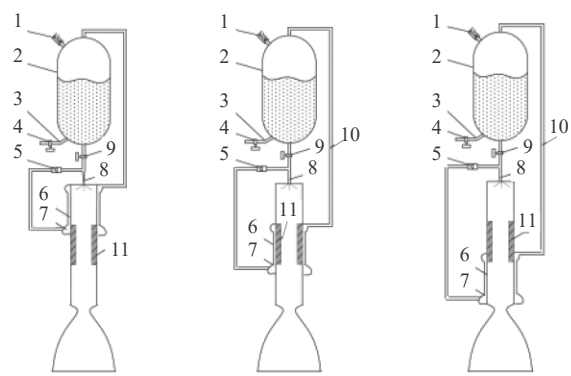
1. 推进剂贮箱; 2. 推力室; 3. 冷却剂驱动泵; 4. 冷却剂贮箱; 5~7. 换热器; 8. 贮箱出口。

图 13　冷却液循环方案示意图[73]

Fig. 13　Schematic diagram of coolant circulation scheme[73]

道, 吸收燃烧室的热量, 升温后的冷却剂在泵的作用下流经储箱表面的通道, 实现对氧化亚氮加热, 并再次进入冷却通道。这种加热方式利用燃烧室的热源降低了之前方案对“外加能源”的需求, 但冷却剂和泵增加了系统的质量和系统复杂度。

也有学者提出将氧化亚氮作为冷却剂, 经过再生冷却后再流入储箱, 从而避免引入专用冷却系统, 降低了系统的质量和复杂度。王正凯[74] 提出将再生冷却与储箱热补偿相结合的自增压方案, 在固液发动机前燃室布置再生冷却通道, 使氧化亚氮蒸发为蒸气后输送回储箱增压的方案, 如图 14 所示。Yeul 和方杰等[75-76] 设计了类似的再生热补偿系统。这种加热方式结构简单, 在对储箱加压的同时, 在一定程度上也缓解燃烧室高温烧蚀。然而,



(a) 冷却通道布置在前燃烧室　(b) 冷却通道布置在装药段　(c) 冷却通道布置在后燃烧室

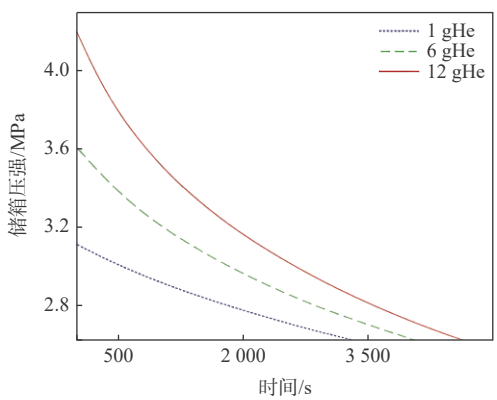
1. 安全阀; 2. 贮箱; 3. 加注管路; 4. 加注阀; 5. 单向阀; 6. 冷却通道; 7. 集液环; 8. 供应管路; 9. 主阀; 10. 增压管路; 11. 固体药柱。

图 14　基于再生冷却的自增压供应系统示意图[74]

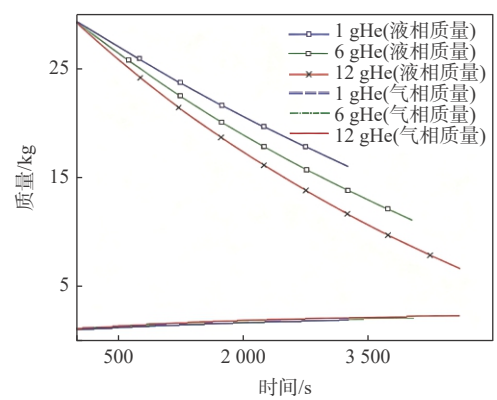
Fig. 14　Schematic diagram of a self-pressurized supply system based on regenerative cooling[74]

推进剂在冷却通道内的相态变化复杂，冷却通道设计要综合考虑推进剂流量、冷却位置、室压等因素，还要严格控制冷却通道内推进剂的温度在其安全温度以下。王正凯[74]通过数值模拟对冷却通道进行了深入研究，发现在亚临界压强下，通道内干涸点位置的温度会突升，应尽量将冷却通道设置在燃烧室热流密度较小的地方或将氧化亚氮主动冷却作为冷却壁面的辅助措施。杨中旭[24]对再生冷却过程进行数值模拟，发现液态氧化亚氮在进入冷却通道后，在壁面热流的影响下温度很快升高并发生相变，形成贴壁的高温气态氧化亚氮区域，随后气态区与液态区换热，使液态区温度不断升高，气态区域的范围也不断扩展。

另外，也有学者在氧化亚氮储箱内添加高压气体来提高利用率，增加自增压工作时长。早在2009年Fernandez[64]使用高压氦气对氧化亚氮储箱进行增压，提高了挤出效率，并建立了相关模型。龚亮[72]对比了有无氦气增压辅助对自增压特性的影响，发现当氦气量增加时，N_2O储箱初始工作压强提高，自增压工作时间增加，储箱内液态N_2O剩余质量减少，但气态N_2O剩余质量没有太大变化，结果如图15所示。此外，氦气量增加使储箱自增压性能提高的程度有限，可加注的氦气量存在极大值。Karabeyoglu[77]提出了将氧化亚氮和氧气的混合物作为火箭发动机的氧化剂，既可以提高氧化亚氮的自增压性能，又能提高氧化剂的能量特性。对此，张灿宇等[78]利用Peng-Robinson状态方程研究了氧化亚氮和氧气二元组分气-液相平衡，在范德华方程的基础上改善修正，提高了计算饱和蒸气压、饱和液相密度的准确性。综上，通过辅助气体增加起始压强可以延长自增压工作时间。



(a) 储箱压强



(b) 储箱内各相质量

图15 不同氦气量对储箱自增压工作特性的影响[72]

Fig. 15 Effect of different helium quantities on self-pressurization performance of storage tanks[72]

综上，氧化亚氮凭借其独特的自增压特性，可显著简化发动机供给系统结构，然而，其复杂的理化性质与相变行为也大幅增加了自增压过程的建模难度。现有研究表明，在集总参数模型中采用非平衡模型、Peng-Robinson非理想气体状态方程或Span-Wagner状态方程、推进剂分区蒸发传热传质模型等假设模型可以提高自增压预测精度；随着高精度模型的不断完善，结合数值模拟，将进一步增强系统仿真的可靠性。针对自增压过程中储箱压强下降的稳压需求，研究人员提出了多种调控方案：直接加热氧化亚氮需较大功率，适用于小流量工况；利用冷却剂回收燃烧室热量对储箱加热可降低功耗，但对冷却通道设计提出更高要求；高压气体辅助增压可在一定程度上延长使用时间，但改善效果有限。综合现有方案，目前尚未形成高效简洁的氧化亚氮稳压供应方法。最具潜力的技术路径是将经再生冷却后的氧化亚氮引回储箱用于增压，而该方案的实现需要在氧化亚氮冷却通道设计、冷却剂流量分配及储箱工作稳定性等关键方面开展更为深入的研究。

2.2 氧化亚氮催化分解技术

氧化亚氮催化分解技术在氧化亚氮推进领域具有重要的应用价值。首先，氧化亚氮作为单组元推进剂可直接用于卫星的姿态控制、位置保持、纳米卫星推进及推力矢量控制等任务。其次，该技术可作为点火装置提供一种可靠、结构简单且可重复启动的点火源，并可用于液体火箭发动机点火等功能。此外，氧化亚氮催化分解还可用于航天生命支持系统，因其分解产物为适宜呼吸的氮氧混合气体。氧化亚氮催化分解装置的基本工作原理如图16所示[79]。

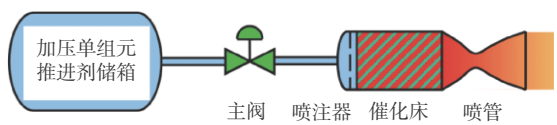


图16 氧化亚氮催化分解装置基本原理[79]

Fig. 16 Basic principle of catalytic decomposition device for nitrous oxide[79]

氧化亚氮催化分解技术的难点在于缩短催化剂的启动时间与提高催化分解效率。目前, 催化剂通常需借助外部热源预热至特定温度才能激活, 从室温升至活化温度通常需要数分钟, 降低了发动机的响应速度[29]。此外, 常用的催化剂多数难以耐受高温环境。实验表明, 氧化亚氮分解后的气体温度可超过 1 500 ℃[29], 在持续高温作用下催化剂易发生板结, 从而导致分解效率显著下降。

为缩短氧化亚氮催化床的启动时间, 当前研究主要聚焦于催化剂的优选、引入辅助气体以降低预热温度, 优化加热方式等方面。合适的催化剂能够有效降低反应活化能, 提高反应速率, 并降低预热温度。Zakirov 等[29] 评估了包括 NiO、铱基催化剂 Shell-405 和 LCH-201 在内的 50 余种催化剂, 发现铱基催化剂 Shell-405 能够将氧化亚氮的催化分解温度显著降低至 523.15 K, 且分解产物温度可达 1 773.15 K 以上。该团队还研制出一种以氧化钛包覆的氧化铝为载体的负载型铑(Rh)催化剂, 可在 673 K 触发分解并支持 5 次以上的重复启动, 表现出良好的稳定性。研究表明, 在氧化亚氮中添加少量特定的气体有助于提高氧化亚氮的化学反应动力学, 降低预热温度。Balasubramanyam 等[80] 为提高 NOP 预混气催化效果, 使用具有铱基 Shell-405 和铂的混合催化剂床, 在丙烷中添加 10% 丙烯作为燃料, 使发动机的起燃温度降低了约 25%, 在 423 K 时便观察到氧化亚氮与微量丙烷/丙烯混合物的完全热分解, 但高温分解气会导致催化剂烧结失活。Scherson 等[81] 将氧燃比为 30 的氧化亚氮/甲烷混合物流过贵金属/氧化铝催化剂, 在约 333 K 床温下, 10 s 内催化剂温度超 1 623 K; 而无甲烷的催化分解测试, 需至少 3 倍加热功率。Tiliakos 等[54] 也证实通入微量的烃(如丙烷或丙烯)可以降低催化剂预热温度至 366 K。综上, 通过添加特定气体可以降低催化剂预热温度, 从而减少加热能耗, 并缩短点火延迟时间。合适的预热方式会缩短催化床加热时间。目前常用的预热方式主要包括外热式和内热式 2 种。外热式通常通过在反应器外壁缠绕电阻丝进行加热, 如图 17 所示[30], 图中, T_{CB} 为燃烧室前端温度, T_C 为燃烧室温度, L 为催化床长度; 内热式则采用置于催化床内部的电热塞实现直接加热, 如图 18 所示[33]。此外, Hitch 等[82] 提出了逆流式换热反应器结构, 如图 19 所示。该反应器使用柱状电加热器将其外表面涂覆的催化剂加热至约 520 K, 通入的液态氧化亚氮在催化剂作用下发生分解, 反应释放的热量进一步促使未分解的液态推进剂蒸发。该方案有效利用了分解反应热, 同时

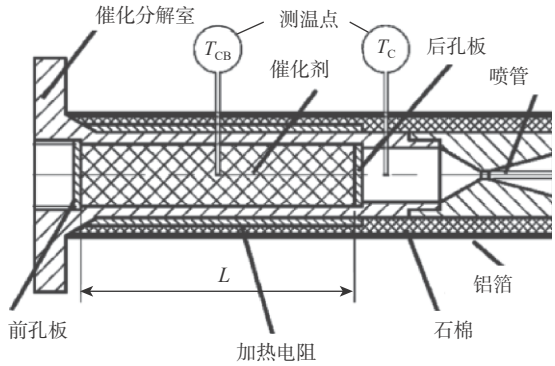
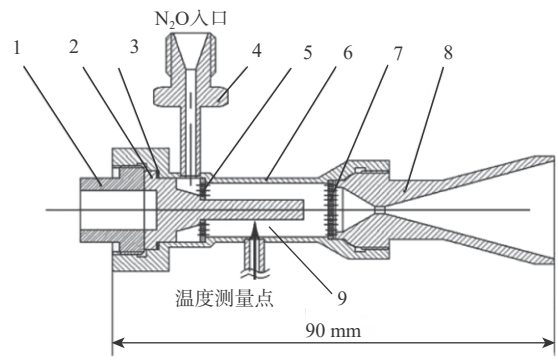


图 17　氧化亚氮微推力推进器方案[30]

Fig. 17　Schematic of N_2O micro-thruster[30]



1. 堵头; 2. 加热器; 3. 垫圈; 4. N_2O入口接管嘴; 5. 前垫圈; 6. 壳体; 7. 后垫圈; 8. 喷管; 9. 催化剂。

图 18　N_2O 单组元推进器原理图[33]

Fig. 18　Schematic of N_2O monopropellant thruster[33]

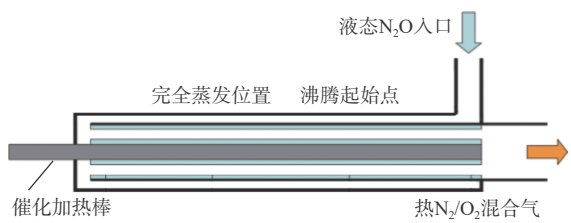


图 19　氧化亚氮催化换热器/反应器[82]

Fig. 19　Catalytic N_2O heat exchanger/reactor [82]

增大了氧化亚氮与催化剂的接触面积, 从而显著提高了换热与分解效率。

催化床的尺寸及氧化亚氮与催化床接触时的相态是影响催化分解效率的重要因素。在确定催化剂种类后, 需要对催化床的尺寸进行优化。Lohner 等[79] 采用金属氧化物与陶瓷材料作为催化剂, 发现催化床长度对分解性能影响显著。龚亮[72] 研究了催化床最优尺寸问题, 指出在低流量条件下催化剂利用率低、热损失大, 导致比冲下降; 高流量时则因分解不完全, 同样造成性能衰减; 仅在特定流量范围内, 比冲随流量增加而提升。李茂等[30] 在亚牛级微推进系统上开展的真空实验进一步验证了上述结论。氧化亚氮在催化剂作用下进行催化分解时, 气体和液体状态的氧化亚氮理论上都可以

参与反应。不过在实际应用场景中，更多是气态氧化亚氮参与催化分解反应。因为气态物质分子间距大、活性较高，在与催化剂接触时，能够更充分地发生反应，反应速率相对较快。液态氧化亚氮若要参与反应，往往需要先将其转化为气态，如通过加热使其气化，以满足催化分解的高效性需求。目前研究中，只有少数学者[82]针对液态氧化亚氮开展了催化分解研究。

综上，目前已开展了较多的氧化亚氮催化分解技术研究，研制出了推力从毫牛级到百牛级的各种原理样机。然而，该技术仍面临预热时间长、启动温度高及催化剂工作寿命短等问题。为进一步提升性能，未来的装置可采用高温抗氧化材料、使用3D打印等新技术制造更大接触面积的催化床[83]；寻求可以进一步提高氧化亚氮的化学反应动力学的辅助催化气体；发展适合液态氧化亚氮催化分解的催化剂，可进一步缩小发动机的尺寸和质量，提高密度比冲。

2.3 氧化亚氮发动机性能提升技术

如表2所示，与液氧/甲烷、液氧/液氢及液氧/煤油等常用推进剂组合相比，氧化亚氮基推进剂比冲性能相对较低[77]。为提升其能量特性，研究人员采用多种方法提高氧化亚氮双组元发动机的比冲性能。

目前研究主要集中在添加氧制作混合氧化剂Nytrox，不同于将燃料和氧化亚氮制做为单组元推进剂NOFBX，Nytrox是将氧和氧化亚氮互溶混合，旨在制备无需深低温储存的高密度自增压氧化剂；在Nytrox的组分中，氧起到增加比冲与密度比冲的作用，而氧化亚氮则弥补了液氧的操作温度低、稳定性与可调节能力差等缺点。在较低温度时，氧化剂为纯液氧。随着温度的升高，两种物质可以互溶为Nytrox氧化剂。温度继续升高，Nytrox的氧含量降低，最后在接近室温(298 K)的温度下，者不能互溶，为纯氧化亚氮。由于Nytrox的使用温度高于纯液氧而低于纯氧化亚氮，因此，Nytrox中的氧将缓慢蒸发至气相起增压作用，而氧化亚氮则主要存在于液相中，增加了Nytrox的密度，使Nytrox同样具备自增压能力，但其自增压能力低于氧化亚氮。

氧化亚氮、氧气和Nytrox这3种氧化剂的特性对比如表3所示[77]。可以看出，Nytrox混合氧化剂具有诸多优势。与N_2O相比，Nytrox混合氧化剂的比冲性能更高，凝固点更低，与火星表面的温度完美匹配，适合执行火星表面空间任务；液氧的沸点为−183 ℃，Nytrox氧化剂在−80 ℃(或更高的温度)便可以使用，可以使用复合罐，降低了对储箱的使用要求，且系统总重比液氧推进系统更轻；能够通过部分自加压方式供应更高密度的推进剂，有效降低成本；Nytrox混合物的温度和压强可以单独控制，可以根据任务需求灵活调整储箱状态。

表3 O_2、N_2O和Nytrox作为氧化剂的特性对比[77]

Table 3 Comparison of pure O_2, N_2O and Nytrox as oxidizers[77]

氧化剂	比冲性能	密度	密度比冲	化学稳定性	毒性	可存储性	自增压能力	气相燃烧	自燃性	易操作性	材料成本	化学相容性	性能调节能力	发动机稳定性/燃烧效率	综合安全性
O_2	5	4	4	5	5	1	1	1	1	3	5	5	1	2	3
N_2O	3	2	1	4	4	5	5	3	1	5	4	5	2	5	2
Nytrox	4	4	3	4	4	3	3	5	1	4	4	5	5	4	5

注：5为性能最佳，1为性能最差。

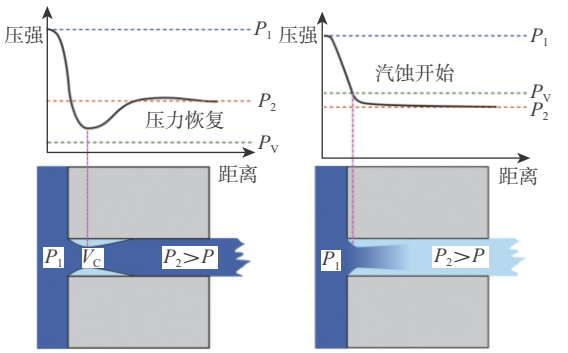
现阶段已经有研究人员提出将Nytrox用于小型运载火箭、火星上升器等动力系统。Karabeyoglu等[84]设计了Nytrox/添加20%铝的石蜡固液混合发动机以取代猎户座运载火箭的纯固体上面级发动机。也有学者对Nytrox供应过程进行了研究，张灿宇等[78]利用Peng-Robinson状态方程在不同的温度条件下对氧化亚氮-氧气(气-液相)平衡状态进行了计算，并对密度进行了分析，得到了温度与压强变化时储箱内气液两相密度的变化趋势。

综上，Nytrox能量性能优异且兼具自增压能力，相较于纯液氧推进系统有显著的结构质量优势和低成本的使用温度，相较于纯氧化亚氮推进系统则有性能优势。然而，Nytrox工作过程中组分变化较大，供应过程还不清晰，需进一步开展相关研究。目前，Nytrox主要在固液混合发动机使用，未来可尝试在氧化亚氮双组元发动机上使用，提升其比冲。为提升氧化亚氮发动机比冲，还可以使用高能燃料。氧化亚氮含氧量较少，生成产物中一般会有大量的氮气，可通过在生成物中添加镁粉等金属粒子的方式，使氮气和其进一步反应，用来提升发动机性能。

2.4 氧化亚氮两相流喷注雾化技术

在氧化亚氮自增压供应系统中，储箱挤压压强通常接近其饱和蒸汽压。当氧化亚氮在喷嘴内加速膨胀时，局部静压会低于饱和蒸汽压，引发闪蒸现象。不同饱和蒸汽压液体在喷嘴中的流动特性

如图 20 所示[65]。此外, 由于燃烧室压强通常远低于氧化亚氮的饱和蒸汽压, 喷嘴出口处的压强骤变会导致剧烈相变, 产生爆炸式雾化, 甚至可能因吸热导致局部“结冰”而造成点火困难[85]。喷嘴内外的复杂相变行为, 加之氧化亚氮物性对温度变化敏感, 显著增加了喷嘴设计的难度, 影响发动机工作稳定性。因此, 深入研究氧化亚氮两相流喷注与雾化机理, 对指导喷嘴优化设计具有重要意义。



(a) 高饱和蒸汽压　　(b) 液体在喷嘴中的流动示意图

图 20　低饱和蒸汽压[65]

Fig. 20　Low saturated vapor pressure[65]

喷嘴结构参数和喷嘴类型对氧化亚氮喷注雾化过程影响明显。Waxman 等[86] 通过实验发现, 直流喷嘴进口几何形状对雾化特性影响不大, 但喷嘴的直径和长径比对雾化效果起着决定性的作用, 且撞击式喷嘴可以显著改善雾化性能和推进剂在燃烧室内的分布。

氧化亚氮的温度也是影响其喷注雾化的重要参数。通过改变氧化亚氮的温度从而改变饱和蒸气压, 实现氧化亚氮在喷嘴内流动状态的改变, 从而调整发动机工作特性。NASA 研究中心开发的游隼单级探空火箭采用了氧化亚氮/石蜡固液混合发动机, 研究发现氧化亚氮喷注器内两相流加剧了发动机内的燃烧不稳定, 如图 21 所示[87]。Karabeyoglu 等[88] 采取减少氧化亚氮的蒸发时间、提高喷嘴喷注压降等措施改善系统低频耦合不稳定。Waxman 等[89] 通过增加氧化亚氮的初始供给温度, 提升了其初始饱和蒸汽压, 使燃烧室室压低于该饱和蒸汽压的 80%。试验结果表明, 该方法能够有效减弱室压振动的幅值和频率, 如图 22 所示。未来可从以下几个方面抑制不稳定燃烧: 促进氧化剂快速蒸发, 缩短蒸发延迟时间; 减少管路供

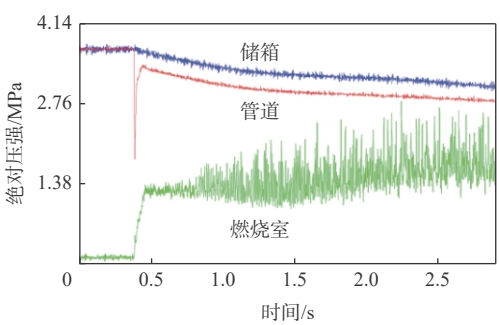
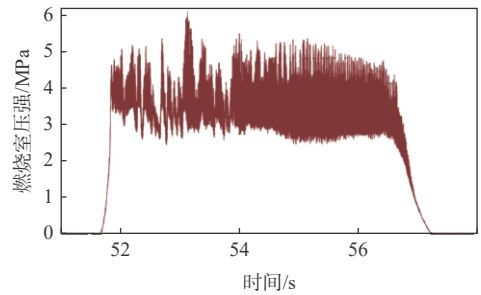
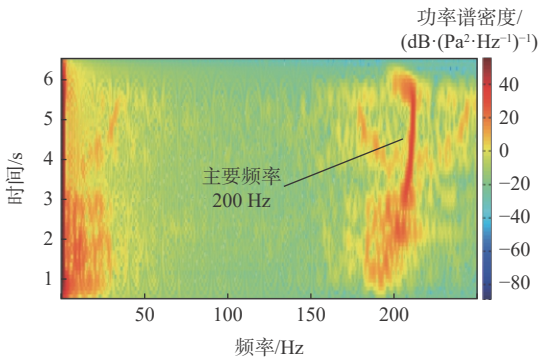


图 21　氧化亚氮/石蜡固液混合动力发动机中典型的低频和高频耦合燃烧不稳定性[87]

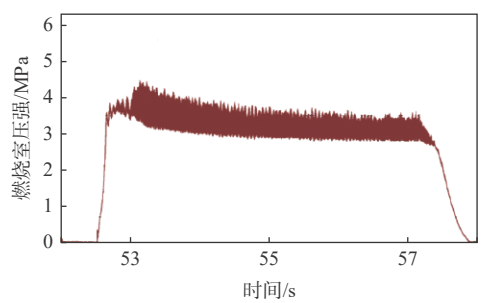
Fig. 21　Typical low frequency and high frequency coupled combustion instability in a nitrous oxide/paraffin solid-liquid hybrid engine[87]



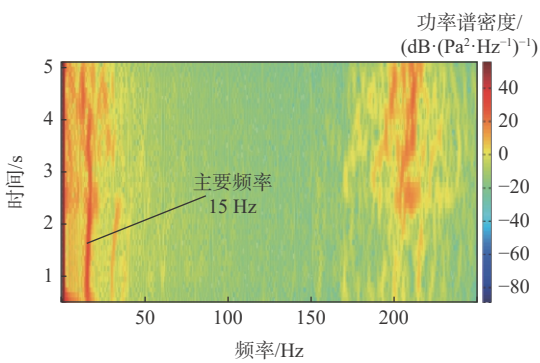
(a) 氧化亚氮初始温度约13.5 ℃(286.7 K)时的燃烧室压强曲线



(b) 氧化亚氮初始温度约13.5 ℃(286.7 K)时的燃烧室室压频率



(c) 氧化亚氮初始温度约21 ℃(294.2 K)时的燃烧室室压曲线



(d) 氧化亚氮初始温度约21 ℃(294.2 K)时的燃烧室室压频率

图 22　氧化亚氮造成的发动机系统的耦合不稳定[89]

Fig. 22　Coupling instability of engine system caused by nitrous oxide[89]

给充填时间;修改喷嘴设计以增加喷注压降。

综上,氧化亚氮的两相流喷注雾化过程复杂,并极大地影响发动机工作稳定性,然而,目前其喷注雾化过程研究还不足,尤其是氧化亚氮在喷嘴处的剧烈闪蒸现象,需深入开展氧化亚氮自增压流动相变研究,对自增压供给系统特性分析与氧化亚氮自增压火箭发动机设计有重要意义。

2.5 氧化亚氮发动机热防护技术

与常规推进剂相比,氧化亚氮热容低(T=300 K, P=5.88 MPa, C_p≈2.83 kJ/(kg·K)),在高温条件下易发生相变,其作为氧化剂的火箭发动机又具有更高燃温的特点,因此,氧化亚氮热防护更加困难。为应对这一挑战,可根据实际任务需求,灵活采取氧化亚氮气态、液态或超临界状态(supercritical N_2O, SCN_2O, T_c=309.58 K, P_c=7.254 MPa)等冷却方案。

目前,多数氧化亚氮发动机采用气态或液态氧化亚氮进行再生冷却,以有效控制燃烧室壁面温度。Berg 等[90]研究了黎明宇航公司(Dawn Aerospace)的B1发动机(推力为0.5 N),该发动机采用自增压气态氧化亚氮供应。研究发现,在辐射冷却条件下,壁面稳态温度高于所有再生冷却方案;而在采用螺旋形冷却通道时,其冷却性能优于冷却夹套和直线型通道,最高壁温降低22%,且低于材料允许的最高工作温度,壁面温度对比如图23所示。Gallo 等[91]研究了氧化亚氮固液混合发动机石墨喉衬的烧蚀情况,如图24所示,结果表明,在喉衬外部实施液态氧化亚氮再生冷却可显著降低喉部温度,减轻高温燃气对材料的侵蚀,从而在多次点火试验中保持结构完整性。在实际应用中,由于氧化亚氮自身冷却能力有限,常将其与燃料共同用于冷却。脉冲公司的氧化亚氮/乙烷发动机在燃烧室和喉部位置使用冷却技术使其最长工作时间长达62 s[16],如图25所示。此外,该公司800 N级发动机使用氧化亚氮对燃烧室进行再生冷却,并利用燃料对喷管进行冷却[17],如图26所示。

除了气态、液态氧化亚氮用来冷却,SCN_2O也有很大潜力用于冷却。Kumar 等[92]为解决SSTO

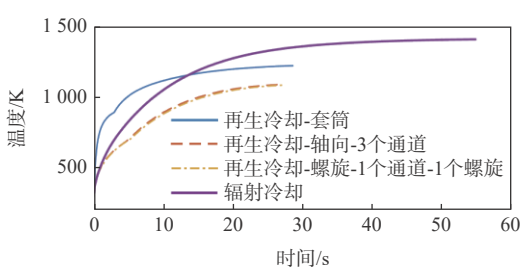
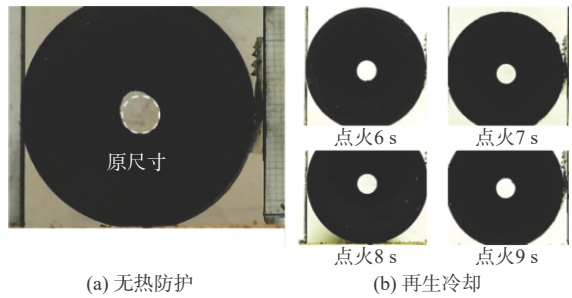


图23 4种不同冷却方式下最高壁温随燃烧时间的变化[90]

Fig. 23 Maximum wall temperature as function of burn time under four different cooling methods[90]



(a) 无热防护 (b) 再生冷却

图24 热测试后无热保护和有热保护的发动机喉部比较[91]

Fig. 24 Comparison of nozzle throats without thermal protection and with thermal protection after thermal test[91]

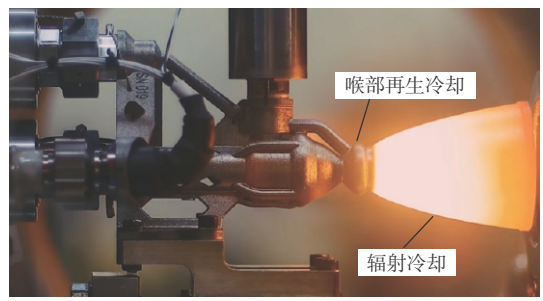


图25 脉冲空间公司研发的22 N级氧化亚氮/乙烷发动机[16]

Fig. 25 22 N levels of nitric oxide/ethane engine developed by Impulse Space company[16]



图26 脉冲空间公司的800 N级发动机[17]

Fig. 26 Class 800 N engine developed by Impulse Space[17]

系统上混合塞式喷管喉部高温烧蚀问题,对比了气态、液态N_2O和SCN_2O的物理性质,发现SCN_2O的比热值和对流传热系数比液态和气态N_2O高得多,证明了SCN_2O作为冷却剂的可行性。

也有学者采用发汗冷却的方式实现对氧化亚氮发动机的热防护。例如,Tiliakos 等[54]在NOP发动机中将25%的氧化亚氮用于燃烧室的发汗冷却,也保证了发动机的正常工作。杨中旭[24]采用复合冷却方案,推力室圆柱段采用再生冷却,喷管采用复合材料,同时在喷管收缩段布置液膜冷却喷孔,从而使液态氧化亚氮自推力室圆柱段末端进入再生冷却通道与液膜喷孔,而再生冷却通道出口的氧化亚氮经集液腔与喷注器等结构以气态喷注燃烧,通过数值模拟验证了方案的可行性。

综上所述,现有研究已分别采用气态与液态氧

化亚氮对发动机实施冷却, 并取得了良好的热防护效果。SCN_2O 因其热物理性质上的优势, 如更高的传热系数, 在发动机热管理中展现出显著潜力, 值得在未来开展更系统、深入的研究。此外, 将氧化亚氮与燃料相结合的复合冷却策略, 有望通过协同互补进一步提升冷却效率, 是该领域一个重要的发展方向。

2.6　氧化亚氮发动机点火技术

氧化亚氮在喷嘴处易发生剧烈闪蒸, 其复杂的相态变化增加了点火难度。在氧化亚氮双组元发动机中, 常用的点火方式包括电点火和火炬点火。电点火在小推力发动机中广泛应用[45,90], 具有良好的重复启动能力, 但随着流量的增大, 点火难度显著提高, 仅在特定余氧系数范围内方可实现可靠点火。火炬点火也广泛应用于氧化亚氮双组元发动机, 推进剂组合包括气/液(NOEL 火炬[93])方案和气/气方案(氧化亚氮/氢气火炬[94])。火炬点火具有结构简单、所需点火流量小、能量高、工作范围宽及可靠性强等优点, 主燃烧室的高压环境还可进一步提升其工作稳定性和重复启动能力。

当氧化亚氮双组元发动机以 Nytrox 作为氧化剂时, 液相推进剂之间的点火相对容易。然而, 对于气相 Nytrox 而言, 其点火所需能量比纯氧化亚氮高出 4～5 个数量级[77]。这一特性虽提升了 Nytrox 的安全性, 但在设计 Nytrox 发动机时, 却需要确保点火过程中尽量有液相 Nytrox 参与。

氧化亚氮液体双组元发动机在点火时的困难主要体现在需要点火能量高且点火延迟时间长等方面。Werling 等[95] 使用氧化亚氮/乙烯混合推进剂进行实验, 捕捉到了点火过程, 测量了点火延迟时间, 点火过程如图 27 所示。王栋等[55] 采用 NOP 组合, 发现余氧系数明显影响点火延迟时间。

对 NOFBX 的点火过程也开展不少研究, 主要

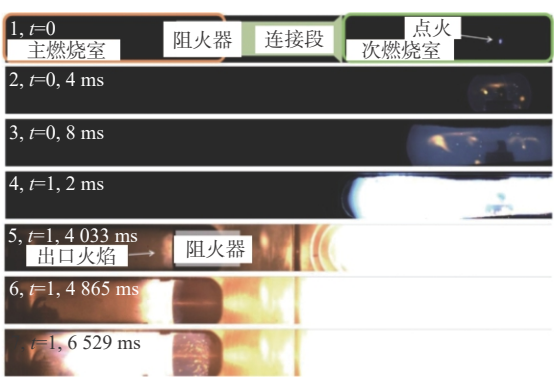


图 27　氧化亚氮/乙烯火焰传播机制的高速图像(火花塞点火)[95]

Fig. 27　High speed images of N_2O/C_2H_4 flame propagation mechanism(ignition at spark plug)[95]

集中在 N_2O 与小分子碳氢燃料(C_1、C_2), 特别是 N_2O/C_2H_4 的着火和燃烧特性。Yang 等[41] 建立了包含 129 种组分、900 步基元反应的乙烯基 NOFBX 燃烧化学反应机理模型和层流火焰传播模型, 模型计算结果与实验结果较为吻合。段志强和郑东等[42-43] 针对氧化亚氮和 C_2 烃, 以现有碳氢化物与氮氧化物的反应机理为基础, 发展了包含 52 组分、325 步基元反应的 NOFBX 燃烧化学反应机理模型, 得到的结果可以在宽范围内较为准确地预测推进剂的点火延迟时间及层流火焰速度, 有望在 NOFBX 发动机的多维燃烧数值模拟中得到广泛应用。魏豪等[44] 基于准稳态假设分析建立了 21 种组分, 18 步基元反应的 N_2O/C_2H_4 化学反应动力学简化模型, 通过 CFD 仿真计算分析了不同参数对 NOFBX 推进剂燃烧特性的影响, 并针对推进剂的再生冷却特性进行了计算。杨猛等[45-46] 利用加热的快速压缩机研究了氩气稀释下甲烷/氧化亚氮($CH_4/N_2O/Ar$)推进剂的自着火特性, 总结了混合气体的着火延迟时间预测模型。

综上, 在设计氧化亚氮发动机点火装置时, 需综合考虑推进剂供应状态、流量及推进剂组合等因素。在后续的研究中, 还需针对如何减小点火延迟时间、如何增加点火的成功率、如何在大流量条件下实现可靠电点火、多次点火及低功率电点火等问题进行重点突破。

3　结 束 语

氧化亚氮推进系统具有操作维护简便、绿色无毒、成本低廉、常温可长期贮存及易于获取等突出优势, 已在部分微纳卫星及空间推进器中得到实际应用。目前大多数发动机型号仍处于关键技术攻关阶段, 然而氧化亚氮推进系统基础理论与关键技术还需要从以下几方面进行加强和深化:

1) 高精度自增压控制研究。氧化亚氮自增压优势明显且受环境影响致储箱压强波动, 需构建自增压模型并研制检测调控系统, 保障储箱压强稳定与流量精确控制。

2) 低温预热、长寿命高效催化分解技术研究。当前催化剂存在起燃温度高、催化效率有限和寿命较短的问题, 需研发新型催化材料及助催化方法, 显著提升催化活性和稳定性, 以支持可靠多次点火。

3) 发动机性能提升研究。可从氧化亚氮添加氧制作混合氧化剂 Nytrox、高能燃料及添加金属颗粒和生成物进一步反应 3 个角度开展深入研究。

4) 两相流动喷注雾化研究。定量研究自增压液相氧化亚氮排放的流动相变, 明晰物性变化对发

动机工作稳定性的影响。

5）高可靠性热防护研究。因氧化亚氮发动机氧燃比高、燃温高，常规燃料主动冷却方式不适用，需开展液态或超临界态氧化亚氮再生冷却、氧燃复合冷却等多种方式热防护研究。

6）低点火延迟、高安全性的可靠点火技术研究。应通过推进剂配方优化、新型点火方法开发、供应与点火系统集成改进等措施，全面提升点火安全性和响应速度，保障发动机快速可靠启动。

随着这些理论和关键技术的不断突破，氧化亚氮发动机将会在绿色无毒高性能姿轨控动力领域发挥更大的作用，支撑未来大航天时代发展。

参考文献（References）

[1] 赵婷, 陈夏超, 杨成虎, 等. 面向高轨卫星的液体轨控发动机研制进展[J]. 火箭推进, 2018, 44(1): 1-7.
Zhao T, Chen X C, Yang C H, et al. Developing progress of liquid bipropellant orbit control engine for high orbit satellite[J]. Journal of Rocket Propulsion, 2018, 44(1): 1-7(in Chinese).

[2] 刘志泉, 叶超, 林庆国. 微小型双组元姿控发动机技术研究[J]. 火箭推进, 2014, 40(4): 1-6.
Liu Z Q, Ye C, Lin Q G. Study on miniature liquid bipropellant attitude control rocket engine[J]. Journal of Rocket Propulsion, 2014, 40(4): 1-6(in Chinese).

[3] Jang Y H, Lee K H. A development trend study of bipropellant rocket engine for orbit transfer and attitude control of satellite[J]. Journal of the Korean Society of Propulsion Engineers, 2015, 19(1): 50-60.

[4] Werling L, Lauck F, Dobusch J, et al. Mono-and bipropellants based on nitrous oxide and fuels: state of the art, recent developments in academia and industry, and a special focus on DLR's activities[C]//Proceedings of the Recent Advancements in Green Propulsion: Green Propellants for Micropropulsion Systems. Berlin: Springer, 2024: 161-186.

[5] Dawn Aerospace. In-space propulsion[EB/OL]. (2025-05-15)[2026-04-29]. https://www.dawnaerospace.com/green-propulsion#anchors-mido9m7e7.

[6] Impulse Space. Impulse Space selected for $34.5M contract by Space Systems Command in support of VICTUS SURGO and VICTUS SALO missions[EB/OL]. (2024-10-03)[2026-04-29]. https://www.impulsespace.com/updates/impulse-space-selected-for-345m-contract-by-space-systems-command.

[7] Relativity Space. Impulse Space and Relativity Space announce first commercial mission to Mars[EB/OL]. (2022-07-19)[2026-04-29]. https://www.relativityspace.com/press-release/2022/7/19/impulse-space-and-relativity-space-announce-first-commercial-mission-to-mars.

[8] Vast. Vast acquires Launcher in quest to build artificial gravity space stations[EB/OL]. (2023-02-21)[2026-04-29]. https://www.vastspace.com/updates/vast-acquires-launcher-to-accelerate-growth.

[9] Bristeau P J, Petit N. Trajectory estimation for a hybrid rocket [C]//Proceedings of the AIAA Guidance, Navigation, and Control Conference. Reston: AIAA, 2009.

[10] Zilliac G, Waxman B, Doran E, et al. Peregrine hybrid rocket motor ground test results[C]//Proceedings of the 48th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit. Reston: AIAA, 2012.

[11] Phillip J, Youngblood S, Grubelich M, et al. Development and testing of a nitrous-oxide/ethanol bi-propellant rocket engine[C]//Proceedings of the 52nd AIAA/SAE/ASEE Joint Propulsion Conference. Reston: AIAA, 2016.

[12] Mungas G, Fisher D, Vozoff J, et al. NOFBX™ single stage to Orbit Mars Ascent Vehicle[C]//Proceedings of the IEEE Aerospace Conference. Piscataway: IEEE Press, 2012: 1-11.

[13] Kakami A, Egawa T, Yamamoto N, et al. Plasma-assisted combustion of N_2O/ethanol propellant for space propulsion[C]//Proceedings of the 46th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit. Reston: AIAA, 2010.

[14] 杨威, 毛成立, 戴祖明. 混合发动机的研究现状及应用前景[J]. 火箭推进, 2006, 32(3): 35-39.
Yang W, Mao C L, Dai Z M. Investigation on status and prospective application of hybrid motors[J]. Journal of Rocket Propulsion, 2006, 32(3): 35-39(in Chinese).

[15] Menshenina Y, Crawleyb E. Model-based concept framework for suborbital human spaceflight missions[C]//Proceedings of the 69th International Astronautical Congress. Paris: International Astronautical Federation, 2018: 11597.

[16] Impulse Space. Mira-the agile spacecraft[EB/OL]. (2025-04-15)[2026-04-29]. https://www.impulsespace.com/mira.

[17] SatNow. Hello, solar system! introducing impulse space propulsion [EB/OL]. (2025-09-10)[2026-04-29]. https://www.satnow.com/videos/details/575-hello-solar-system-introducing-impulse-space-propulsion.

[18] Tokudome S, Yagishita T, Habu H, et al. Experimental study of an N_2O/ethanol propulsion system[C]//Proceedings of the 43rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit. Reston: AIAA, 2007.

[19] Tokudome S, Goto K, Yagishita T, et al. An experimental study of a nitrous oxide/ethanol (NOEL) propulsion system[C]//Proceedings of the AIAA Propulsion and Energy 2019 Forum. Reston: AIAA, 2019.

[20] Waugh I, Moore E, Macfarlanej, et al. Testing of anovel nitrous-oxide and ethanol fuel blend[C]//Proceedings of the Space Propulsion Conference. Paris: European Space Agency, 2018: 14-18.

[21] Gibbon D, Baker A, Coxhill I, et al. The development of a family of resistojet thruster propulsion systems for small spacecraft[C]//Proceedings of the 17th Annual AIAA/USU Conference on Small Satellites. Reston: AIAA, 2003.

[22] 蔡国飙. 固液混合火箭发动机技术综述与展望[J]. 推进技术, 2012, 33(6): 831-839.
Cai G B. Development and application of hybrid rocket motor technology: overview and prospect[J]. Journal of Propulsion Technology, 2012, 33(6): 831-839(in Chinese).

[23] 孙威, 方杰, 吴靖, 等. 亚牛级 N_2O 单组元推力器研制[C]//中国化学会第五届全国化学推进剂学术会议论文集. 北京: 中国化学会, 2011: 131-135.
Sun W, Fang J, Wu J, et al. Development of a sub-Newton N_2O

monopropellant thruster[C]//Proceedings of the 5th National Conference on Chemical Propellants, Chinese Chemical Society. Beijing: Chinese Chemical Society, 2011: 131-135(in Chinese).

[24] 杨中旭. 氧化亚氮煤油火箭发动机燃烧及热防护技术研究[D]. 西安: 西北工业大学, 2024.
Yang Z X. Research on combustion and thermal protection technology of nitrous oxide/kerosene rocket motor[D]. Xi'an: Northwestern Polytechnical University, 2024(in Chinese).

[25] 张峻铭. 氧化亚氮自增压供应特性研究[D]. 西安: 西北工业大学, 2024.
Zhang J M. Investigation of the characteristics of nitrous oxide self-pressurized supply[D]. Xi'an: Northwestern Polytechnical University, 2024(in Chinese).

[26] 王志杰. 煤油/氧化亚氮空间发动机燃烧与再生冷却换热特性研究[D]. 北京: 北京交通大学, 2024.
Wang Z J. Study on the heat transfer characteristics of combustion and regeneration cooling of kerosene/nitrous oxide space engines [D]. Beijing: Beijing Jiaotong University, 2024(in Chinese).

[27] 刘轩坤, 苏泽, 戴磊, 等. 液态氧化亚氮/乙烷发动机设计及试验研究[C]//第九届空天动力联合会议暨中国航天第三专业信息网第四十五届技术交流会论文集. 贵阳: 中国空天动力联合会, 浙大宁波理工学院, 2025: 142-146.
Liu X K, Su Z, Dai L, et al. Design and experimental study of a liquid nitrous oxide/ethane engine[C]//Proceedings of the 9th Joint Conference on Aerospace Propulsion & the 45th Technical Exchange Conference of the 3rd Professional Information Network of China Aerospace. Guiyang: China Aerospace Propulsion Federation, NingboTech University, 2025: 142-146(in Chinese).

[28] 邓德涛, 沈赤兵, 李清廉, 等. N_2O/酒精发动机试验研究[C]//第二届高超声速科技学术会议会议日程及摘要集. 北京: 中国力学学会, 2009.
Deng D T, Shen C B, Li Q L, et al. Experimental study on N_2O/alcohol engine[C]//Proceedings of the Program and Abstracts of the 2nd Hypersonic Science and Technology Academic Conference. Beijing: Chinese Society of Theoretical and Applied Mechanics, 2009(in Chinese).

[29] Zakirov V, Richardson G, Sweeting M. Surrey research update on N_2O catalytic decomposition for space applications[C]//Proceedings of the 37th Joint Propulsion Conference and Exhibit. Reston: AIAA, 2001.

[30] 李茂, 韩乐, 方杰, 等. N_2O 单组元微推力器高空模拟实验研究[J]. 推进技术, 2011, 32(3): 307-311.
Li M, Han L, Fang J, et al. Simulated altitude experimental investigation for N_2O monopropellant micro-thruster[J]. Journal of Propulsion Technology, 2011, 32(3): 307-311(in Chinese).

[31] 方杰, 田辉, 蔡国飙. N_2O 单组元微推进系统及其喷管流场的初步研究[J]. 推进技术, 2005, 26(6): 495-498.
Fang J, Tian H, Cai G B. Preliminary study on a N_2O micro propulsion system and its nozzle flow field[J]. Journal of Propulsion Technology, 2005, 26(6): 495-498(in Chinese).

[32] 韩乐, 李茂, 方杰, 等. N_2O 单组元微推力器真空试验设计与分析[J]. 航空动力学报, 2011, 26(10): 2358-2363.
Han Y, Li M, Fang J, et al. Vacuum experiment study on a N_2O monopropellant micro-thruster[J]. Journal of Aerospace Power, 2011, 26(10): 2358-2363(in Chinese).

[33] 吴靖, 孙威, 蔡国飙. 内加热式 N_2O 单组元推力器预热过程仿真与试验[J]. 航空动力学报, 2013, 28(3): 556-560.
Wu J, Sun W, Cai G B. Simulation and experiment for preheating process of N_2O monopropellant thruster with inner-heater[J]. Journal of Aerospace Power, 2013, 28(3): 556-560(in Chinese).

[34] 霍雪亮, 韦迪, 李路明. N_2O 单组元推进器预热过程的建模[J]. 推进技术, 2005, 26(1): 80-83.
Huo X L, Wei D, Li L M. Modeling of N_2O monopropellant thruster pre-heating[J]. Journal of Propulsion Technology, 2005, 26(1): 80-83 (in Chinese).

[35] 万科, 李路明, 韦迪, 等. N_2O 混合火箭发动机的催化点火研究[J]. 推进技术, 2007, 28(1): 1-3.
Wan K, Li L M, Wei D, et al. Experimental investigation on catalytic ignition of N_2O hybrid rocket motor[J]. Journal of Propulsion Technology, 2007, 28(1): 1-3(in Chinese).

[36] 逄凯, 李新艳, 李效斯, 等. NOFBX 推进剂技术的发展概况[J]. 兵器装备工程学报, 2023, 44(10): 147-158.
Pang K, Li X Y, Li X S, et al. Review of the development of NOFBX propellant technology[J]. Journal of Ordnance Equipment Engineering, 2023, 44(10): 147-158(in Chinese).

[37] Mungas G, Vozoff M, Rishikof B. NOFBX: a new non-toxic, green propulsion technology with high performance and low cost[C]//Proceedings of the 63 International Astronautical Congress. Paris: International Astronautical Federation, 2012: 1-5.

[38] Mayer A, Wieling W, Watts A, et al. European fuel blend development for in-space propulsion[C]//Proceedings of the Space Propulsion Conference. Paris: European Space Agency, 2018: 14-18.

[39] 朱成财, 韩伟, 于忻立, 等. 氧化亚氮基单元复合推进剂技术研究述评[J]. 火箭推进, 2016, 42(2): 79-85.
Zhu C C, Han W, Yu X L, et al. Review of nitrous-oxide-based composite monopropellants technology[J]. Journal of Rocket Propulsion, 2016, 42(2): 79-85(in Chinese).

[40] 曾祥敏, 张玉刚, 蒋榕培, 等. $N_2O/C_2H_4/CO_2$ 预混气体火焰传播及爆炸特性的试验研究[J]. 火炸药学报, 2018, 41(5): 501-505.
Zeng X M, Zhang Y G, Jiang R P, et al. Experimental investigation of flame propagation and explosion properties of premixed gases $N_2O/C_2H_4/CO_2$[J]. Chinese Journal of Explosives & Propellants, 2018, 41(5): 501-505(in Chinese).

[41] Yang X S, Hong X, Dong W. Investigation on self-pressurization and ignition performance of nitrous oxide fuel blend ethylene thruster[J]. Aerospace Science and Technology, 2018, 82-83: 161-171.

[42] 段志强, 郑东, 周斌. N_2O-C_2 烃类燃料推进剂燃烧化学反应机理与动力学分析[J]. 火箭推进, 2021, 47(3): 43-51.
Duan Z Q, Zheng D, Zhou B. Chemical reaction mechanism and kinetic analysis of N_2O: C_2 hydrocarbons propellant[J]. Journal of Rocket Propulsion, 2021, 47(3): 43-51(in Chinese).

[43] 郑东, 熊鹏飞, 钟北京. NOFBX 新型绿色推进剂燃烧化学反应动力学模型[J]. 物理化学学报, 2019, 35(11): 1241-1247.
Zheng D, Xiong P F, Zhong B J. Chemical kinetic model for the combustion of the green propellant of the nitrous oxide fuel blend[J]. Acta Physico-Chimica Sinica, 2019, 35(11): 1241-1247(in Chinese).

[44] 魏豪, 李国岫, 虞育松. 氧化亚氮基单组元绿色无毒推进剂燃烧特性仿真研究[C]//中国化学会第八届全国化学推进剂学术会议. 北京: 中国化学会, 2017: 101-109.

Wei H, Li G X, Yu Y S. Simulation study on combustion characteristics of nitrous oxide green non-toxic propellant[C]//Proceedings of the 8th National Chemical Inpropellant Conference of Chinese Chemical Society. Beijing: Chinese Chemical Society, 2017: 101-109 (in Chinese).

[45] 杨猛, 丁晓倩, 余涛, 等. 甲烷/氧化亚氮绿色推进剂自着火特性实验及动力学[J]. 化工学报, 2025(3): 1221-1229.
Yang M, Ding X Q, Yu T, et al. Experimental and kinetic studies for the ignition characteristic of the green propellant of methane/nitrous oxide[J]. CIESC Journal, 2025(3): 1221-1229(in Chinese).

[46] Yang M, Yang Y, Liao C Y, et al. The auto-ignition boundary of ethylene/nitrous oxide as a promising monopropellant[J]. Combustion and Flame, 2020, 221: 64-73.

[47] Werling L K, Hochheimer B, Baral A L, et al. Experimental and numerical analysis of the heat flux occurring in a nitrous oxide/ethene green propellant combustion demonstrator[C]//Proceedings of the 51st AIAA/SAE/ASEE Joint Propulsion Conference. Reston: AIAA, 2015.

[48] Bangalore Venkatesh P, Meyer S E, Bane S P M, et al. Deflagration-to-detonation transition in nitrous oxide/oxygen-fuel mixtures for propulsion[J]. Journal of Propulsion and Power, 2019, 35(5): 944-952.

[49] Taylor R. Safety and performance advantages of nitrous oxide fuel blends (NOFBX) propellants for manned and unmanned spaceflight applications[C]//Proceedings of the 5th IAASS Conference 'A Safer Space for a Safer World'. Paris: European Space Agency, 2012: 67.

[50] 贾平, 李辉. DARPA 取消"机载发射辅助空间进入"项目简析[J]. 中国航天, 2016(1): 33-35.
Jia P, Li H. Analysis of DARPA's cancellation of "airborne launch assisted space entry" project[J]. Aerospace China, 2016(1): 33-35 (in Chinese).

[51] Mayer A, Waugh I, Poucet M, et al. European fuel blend development final report: TNO 2018 R10640[R]. Rijswijk: NTO, 2018: 14-16.

[52] 杨学森, 洪鑫, 王子模, 等. 新型自增压单元推进剂流动特性研究[C]//中国航天第三专业信息网第三十七届技术交流会暨第一届空天动力联合会议论文集. 西安: 中国航天第三专业信息网, 2016: 449-456.
Yang X S, Hong X, Wang Z M, et al. Study on flow characteristics of a novel self-pressurizing monopropellant[C]//Proceedings of the 37th Technical Exchange Conference of the 3rd Professional Information Network of China Aerospace & the 1st Joint Conference on Aerospace Propulsion. Xi'an: The 3rd Professional Information Network of China Aerospace, 2016: 449-456(in Chinese).

[53] Smith N, Moser M, Charles Kopicz J Jr, et al. Nitrous oxide/propane rocket test results[C]//Proceedings of the 36th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit. Reston: AIAA, 2000.

[54] Tiliakos N, Tyll J, Herdy R, et al. Development and testing of a nitrous oxide/propane rocket engine[C]//Proceedings of the 37th Joint Propulsion Conference and Exhibit. Reston: AIAA, 2001.

[55] 王栋, 郭红杰, 梁国柱, 等. N_2O/C_3H_8 点火器初步实验[J]. 北京航空航天大学学报, 2010, 36(10): 1162-1165.
Wang D, Guo H J, Liang G Z, et al. Preliminary experimental investigation on nitrous oxide/propane igniter[J]. Journal of Beijing University of Aeronautics and Astronautics, 2010, 36(10): 1162-1165 (in Chinese).

[56] Hageman M D, Knadler M. Detonation cell size of methane with nitrous oxide[C]//Proceedings of the AIAA Propulsion and Energy 2020 Forum. Reston: AIAA, 2020.

[57] Perry M, Hunt A, Christian S. Ground test development of 800lbf nitrous oxide-ethane bipropellant engine[C]//Proceedings of the AIAA Propulsion and Energy 2020 Forum. Reston: AIAA, 2020.

[58] Browne E, Rasmussen J, Marchese A, et al. Modeling single-channel and dual-channel regenerative cooling systems for an ethylene/ethane/nitrous oxide liquid fuel rocket engine[C]//Proceedings of the AIAA Scitech 2021 Forum. Reston: AIAA, 2021.

[59] 魏祥庚, 宋泽林, 张峻铭, 等. 一种氧化亚氮煤油双组元液体火箭发动机动力系统: CN114776479B[P]. 2023-04-28.
Wei X G, Song Z L, Zhang J M, et al. A nitrous oxide/kerosene bipropellant liquid rocket engine propulsion system: CN114776479B [P]. 2023-04-28(in Chinese).

[60] Huzel D K, Huang D H. Modern engineering for design of liquid-propellant rocket engines[M]. Reston: AIAA, 1992.

[61] Ewig R. Vapor pressurization (VaPak) systems history, concepts, and applications: Corpus ID 198920145[R]. Renton: Holder Consulting Group, 2009, 4: 1-14.

[62] Whitmore S A, Chandler S N. Engineering model for self-pressurizing saturated-N_2O-propellant feed systems[J]. Journal of Propulsion and Power, 2010, 26(4): 706-714.

[63] Sargent D H, Tran Y. Liquid nitrous oxide BLEVE tests[C]//Proceedings of the 49th AIAA/ASME/SAE/ASEE Joint Propulsion-Conference. Reston: AIAA, 2013.

[64] Fernandez M M. Propellant tank pressurization modeling for a hybrid rocket[M]. Rochester: Rochester Institute of Technology, 2009.

[65] Dyer J, Zilliac G, Sadhwani A, et al. Modeling feed system flow physics for self-pressurizing propellants[C]//Proceedings of the 43rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit. Reston: AIAA, 2007.

[66] Zilliac G, Karabeyoglu M. Modeling of propellant tank pressurization[C]//Proceedings of the 41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit. Reston: AIAA, 2005.

[67] Kardaś D, Szymborski J. Development of a novel model for emptying of a self-pressurising nitrous oxide tank[J]. Journal of Physics: Conference Series, 2021, 1781(1): 012001.

[68] 宋长青, 徐万武, 陈健. N_2O 贮箱自增压模拟[J]. 航空动力学报, 2016, 31(4): 1019-1024.
Song C Q, Xu W W, Chen J. Self-pressurization modeling of nitrous oxide tank[J]. Journal of Aerospace Power, 2016, 31(4): 1019-1024(in Chinese).

[69] 宋长青. 氧化亚氮自增压研究[D]. 长沙: 国防科学技术大学, 2014: 14-33.
Song C Q. Investigation of a nitrous oxide self-pressurized tank[D]. Changsha: National University of Defense Technology, 2014: 14-33 (in Chinese).

[70] 朱浩, 郭子豪, 张源俊, 等. 氧化亚氮温控增压系统: CN107829842A[P]. 2018-03-23.
Zhu H, Guo Z H, Zhang Y J, et al. Temperature-controlled pressurization system for nitrous oxide: CN107829842A[P]. 2018-03-23 (in Chinese).

[71] Kozlov A, Hinckel J, Comiran A, et al. A self-pressurization sys-

tem for cryogenic propellant tanks[C]//Proceedings of the 32nd Joint Propulsion Conference and Exhibit. Reston: AIAA, 1996.

[72] 龚亮. 一氧化二氮单组元发动机系统方案设计及仿真研究[D]. 长沙: 国防科学技术大学, 2013: 35-65.
Gong L. Design and simulation study on nitrous oxide monopropellant engine system[D]. Changsha: National University of Defense Technology, 2013: 35-65(in Chinese).

[73] Greason J K, Delong D L, Jones D B. System and method for cooling rocket engines: US7784269[P]. 2010-08-31.

[74] 王正凯. 基于再生冷却的自增压方案研究[D]. 长沙: 国防科学技术大学, 2015.
Wang Z K. Research on self-pressurization scheme based on regenerative cooling[D]. Changsha: National University of Defense Technology, 2015(in Chinese).

[75] Yeul S, Bae J. System and method for cooling engines: KR20120015402[P]. 2012-02-21.

[76] 方杰, 李开阳, 孙冰, 等. 实现氧化亚氮自增压稳定供应的再生热补偿系统及方法: CN111734554B[P]. 2020-11-20.
Fang J, Li K Y, Sun B, et al. Regenerative thermal compensation system and method for achieving stable self-pressurized supply of nitrous oxide: CN111734554B[P]. 2020-11-20(in Chinese).

[77] Karabeyoglu M A. Nitrous oxide and oxygen mixtures (nytrox) as oxidizers for rocket propulsion applications[J]. Journal of Propulsion and Power, 2014, 30(3): 696-706.

[78] 张灿宇, 胡松启, 刘林林, 等. 氧化亚氮和氧气二元组分气-液相平衡研究[J]. 上海航天, 2019, 36(S1): 104-108.
Zhang C Y, Hu S Q, Liu L L, et al. Study of gas-liquid equilibrium for nitrous oxide and oxygen[J]. Aerospace Shanghai (Chinese & English), 2019, 36(S1): 104-108(in Chinese).

[79] Lohner K, Dyer J, Doran E, et al. Design and development of a subscale nitrous oxide monopropellant gas generator[C]//Proceedings of the 43rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit. Reston: AIAA, 2007.

[80] Balasubramanyam M, Moser M, Sharp D. Catalytic ignition of nitrous oxide with propane/propylene mixtures for rocket motors [C]//Proceedings of the 41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit. Reston: AIAA, 2005.

[81] Scherson Y, Lohner K, Lariviere B, et al. A monopropellant gas generator based on N_2O decomposition for “green” propulsion and power generation[C]//Proceedings of the 45th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit. Reston: AIAA, 2009.

[82] Hitch B, Wickham D. Design of a catalytic nitrous oxide decomposition reactor[C]//Proceedings of the 46th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit. Reston: AIAA, 2010.

[83] Orth M R, Pandey U, Pourpoint T L. Preparation for pressure drop measurement in monolithic 3D-printed catalyst beds[C]//Proceedings of the AIAA Propulsion and Energy 2021 Forum. Reston: AIAA, 2021.

[84] Karabeyoglu A, Stevens J, Geyzel D, et al. High performance hybrid upper stage motor[C]//Proceedings of the 47th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit. Reston: AIAA, 2011.

[85] 罗盟, 王铁岩, 张青松, 等. 低温流体闪蒸喷雾研究[J]. 宇航总体技术, 2021, 5(6): 27-34.
Luo M, Wang T Y, Zhang Q S, et al. Investigation of cryogenic flashing sprays[J]. Astronautical Systems Engineering Technology, 2021, 5(6): 27-34(in Chinese).

[86] Waxman B, Cantwell B, Zilliac G. Effects of injector design and impingement techniques on the atomization of self-pressurizing oxidizers[C]//Proceedings of the 48th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit. Reston: AIAA, 2012.

[87] Waxman B, Beckwith R, Tybor F, et al. Paraffin and nitrous oxide hybrid rocket as a Mars ascent vehicle demonstrator[C]//Proceedings of the AIAA SPACE 2010 Conference & Exposition. Reston: AIAA, 2010.

[88] Karabeyoglu A, Stevens J, Cantwell B. Investigation of feed system coupled low frequency combustion instabilities in hybrid rockets[C]//Proceedings of the 43rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit. Reston: AIAA, 2007.

[89] Waxman B S, Zimmerman J E, Cantwell B, et al. Effects of injector design on combustion stability in hybrid rockets using self-pressurizing oxidizers[C]//Proceedings of the 50th AIAA/ASME/SAE/ASEE Joint Propulsion Conference. Reston: AIAA, 2014.

[90] Berg P M V D, Botchu J, Hermsen R J G. Investigation of thermal behaviour of additively manufactured green bi-propellant thrusters in cubesat applications using transient thermal modelling[C]//Proceedings of the 72nd International Astronautical Congress. Paris: IAF, 2021: 1-10.

[91] Gallo G, Miyahara Y, Kamps L, et al. Regenerative cooling in hybrid rocket engines based on self-pressurized liquid nitrous oxide[J]. Applied Thermal Engineering, 2024, 254: 123928.

[92] Kumar A, Saini V, Usurumarti P R, et al. Feasibility studies on the cooling of hybrid rocket nozzles using supercritical nitrous oxide[J]. Energy Procedia, 2017, 109: 338-345.

[93] 刘盛田, 胡兴伟, 柳琪, 等. 一氧化二氮/乙醇燃气发生器试验研究[J]. 舰船科学技术, 2008, 30(S2): 223-226.
Liu S T, Hu X W, Liu Q, et al. The experimental studies of gas generator based on nitrous oxide and ethanol[J]. Ship Science and Technology, 2008, 30(S2): 223-226(in Chinese).

[94] 刘盛田, 柳琪, 高磊, 等. N_2O/H2 富燃火炬式点火器研究[J]. 舰船科学技术, 2008, 30(2): 94-97.
Liu S T, Liu Q, Gao L, et al. Study on N_2O/H2 fuel-rich spark igniter[J]. Ship Science and Technology, 2008, 30(2): 94-97(in Chinese).

[95] Werling L K, Hörger T, Manassis K, et al. Nitrous oxide fuels blends: research on premixed monopropellants at the German aerospace center (DLR) since 2014[C]//Proceedings of the AIAA Propulsion and Energy 2020 Forum. Reston: AIAA, 2020.

Current status and trend of nitrous oxide propulsion technology

WEI Xianggeng[1,*], XIE Yuan[1], LU Weijie[1], LIU Xiongfeng[1], GAO Yushan[2], ZHAO Haoyang[1]

(1. National Key Laboratory of Solid Rocket Propulsion, Northwestern Polytechnical University, Xi'an 710072, China;
2. Xi'an Aerospace Propulsion Institute, Xi'an 710100, China)

Abstract: Both monopropellant and bipropellant rocket engines frequently employ nitrous oxide (N_2O) as a non-toxic, stable, and secure self-pressurizing propellant. It holds significant research value and application potential, especially in the field of attitude and orbit control propulsion. In addition to outlining research priorities and development recommendations for N_2O propulsion technology in the field of attitude and orbit control, this paper reviews the development history and current research status of several types of nitrous oxide-based propulsion systems, summarizes their features and advantages, and identifies the key technologies facilitating their engineering application. The aim is to provide valuable references for the innovation and engineering practice of aerospace attitude and orbit control propulsion systems.

Keywords: attitude and orbit control engine; green propellant; nitrous oxide; critical technology

Received: 2026-01-27; **Accepted**: 2026-04-24; **Published Online**: 2026-05-15 14:10
URL: link.cnki.net/urlid/11.2625.V.20260515.1009.001
Foundation item: Shaanxi Provincial Natural Science Basic Research Program (2021JC-14)
* **Corresponding author**. E-mail: realysnow@nwpu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0796

燃料电池航空动力谱系——重塑零碳飞行新格局

李致朋[1,*]，高子博[1]，刘伟栋[1]，彭飞扬[1]，刘鸿[1]，韩庆[2]

(1. 西北工业大学 柔性电子研究院，西安 710072；2. 西北工业大学 航空学院，西安 710072)

摘 要：在全球航空业加速推进碳中和的背景下，燃料电池航空推进系统凭借高能效和零碳排放等优势，成为替代传统燃油动力的关键方向之一。该研究聚焦质子交换膜燃料电池(PEMFC)与固体氧化物燃料电池(SOFC)，系统梳理了两类燃料电池的工作原理、系统架构、关键技术瓶颈及适用场景，并从能量载体类型、功率特性、使用寿命及环境效益等多个方面，与锂电推进系统进行了比较分析。结果表明，PEMFC更适合用于短途及支线航空，而SOFC则在长航时任务中展现出显著潜力。尽管目前在储氢、热管理、系统集成与成本方面仍存在挑战，燃料电池航空动力系统通过结构优化与混合能源架构的持续发展，预期将在续航能力与减排效果等方面获得显著提升，有望成为支撑绿色航空转型的重要技术途径。

关 键 词：燃料电池；航空；固体氧化物燃料电池；质子交换膜燃料电池；可持续航空；混合动力系统

中图分类号：V237; TM911.4

文献标志码：A **文章编号**：1001-5965(2026)10-3369-27

在全球加速推进碳中和的背景下，航空业作为碳排放大户，年排放CO_2约9.3×10^{11} kg，占全球交通运输领域碳排放的12%，面临着深度脱碳的严峻压力[1]。为应对此挑战，国际民航组织(International Civil Aviation Organization, ICAO)推出了“国际航空碳抵消与减排计划”(Carbon Offsetting and Reduction Scheme for International Aviation, CORSIA)，旨在推动行业采用可持续航空燃料(sustainable aviation fuel, SAF)、引入先进低碳推进技术，并优化航线管理[2]。然而，当前主流减排路径仍存在显著局限：SAF全球使用比例不足1%，面临规模化应用瓶颈[1]；纯电池推进系统虽技术成熟，但其能量密度(约300 Wh/kg)难以满足高航程与大载重需求[3]。

因此，发展与应用新型替代推进系统已成为行业迫切需求。在此背景下，燃料电池航空推进系统凭借其高能效、零碳排放和燃料灵活性等优势，被视为替代传统燃油动力的重要技术方向[4]。质子交换膜燃料电池(proton exchange membrane fuel cell, PEMFC)与固体氧化物燃料电池(solid oxide fuel cell, SOFC)是2种最具代表性的技术路线。PEMFC系统效率通常可达50%以上，具有快速启动、高功率密度等特点[5]；SOFC本身发电效率可达65%以上，其与燃气轮机(gas turbine, GT)构成的混合系统SOFC-GT热效率可进一步提高至80%左右[6-7]。此外，燃料电池能量密度可达1 000～1 500 Wh/kg，远高于锂电池，能更好地适配航空业对质量的敏感性[8]。

目前，基于PEMFC-锂电池的混合动力系统已被应用于小型无人机与短途通勤飞机[3]。而SOFC凭借其高温工作特性与多燃料兼容性，在大型无人机和干线客机中展现出潜力[9]。尽管在储氢、热管理、系统集成和成本等方面仍面临挑战，但通过结

收稿日期：2025-11-17；录用日期：2025-11-28；网络出版时间：2025-12-17 19:07
网络出版地址：link.cnki.net/urlid/11.2625.V.20251217.1313.001
基金项目：中央高校基本科研业务费专项资金(D5000230358)；陕西省重点研发计划(2024GX-YBXM-456)
*通信作者. E-mail：iamzpli@nwpu.edu.cn

引用格式：李致朋，高子博，刘伟栋，等. 燃料电池航空动力谱系——重塑零碳飞行新格局[J]. 北京航空航天大学学报，2026，52(10)：3369-3395. Li Z P，Gao Z B，Liu W D，et al. Fuel cells for aviation: a new paradigm for carbon-neutral flight[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3369-3395 (in Chinese).

构优化与发展混合能源架构，燃料电池航空动力系统的续航能力与减排效果有望获得显著提升，从而成为支撑绿色航空转型的重要技术途径。

本文旨在系统分析PEMFC与SOFC的工作原理、系统架构、关键技术及适用场景，并从能量载体、功率特性、寿命及环境效益等维度，与锂电推进系统进行对比。此外，鉴于航空动力系统的设计与应用场景紧密相关，本文将进一步界定各项技术挑战的具体应用边界，从而阐明PEMFC与SOFC在从城市空运到干线飞行的任务谱系中，各自的适配路径与核心瓶颈。

1 燃料电池航空动力系统

燃料电池凭借其高能效、低排放和高度可扩展的特性，被视为推动航空业绿色转型的关键技术。燃料电池航空动力系统是实现这一目标的核心，其性能与可靠性直接决定了飞行器的可行性。该系统绝非简单的燃料电池电堆，而是一个高度集成的复杂工程体系，由电堆、燃料供给、氧化剂供给、热管理、水管理及功率控制等关键子系统精密耦合而成。PEMFC动力系统和SOFC动力系统架构分别如图1和图2所示。深刻理解这些子系统的基本构成、功能及其相互作用，是设计和优化整个动力系统的基石。

1.1 燃料电池电堆

燃料电池动力系统的核心是燃料电池电堆，其是实现化学能向电能高效转换的关键单元，目前，主要的技术路线包括PEMFC和SOFC，无论哪种技术路线，其电堆均由多个单电池串联堆叠而成，以满足不同功率需求，如图3所示。PEMFC的单电池由双极板、膜电极组件（membrane electrode assembly，MEA）和密封件构成。其中，双极板负责气体分配、电流收集并提供结构支撑；MEA则由质子交换膜、催化剂层和气体扩散层组成。质子交换膜仅允许质子通过，构建质子传导通道；催化剂层以贵金属铂为主，用于降低反应活化能；气体扩散层由碳布制成，用于支撑催化剂层及传输气体和电子[10]。而SOFC的单电池则由致密的氧化物陶瓷电解质、多孔的阴阳极和连接件组成，通常采用电解质支撑或阳极支撑结构。其高达600～1 000 ℃的运行温度，能显著提升电解质离子传导速率与电极反应活性，这一特性使SOFC无需使用贵金属催化剂，从而大幅降低了成本。金属或陶瓷连接件负责电流收集和单电池串联[11]。2种电堆均通过端板、拉杆或壳体提供机械压紧力，保障组件间良好的电接触与气密性，防止气体泄漏和短路现象[12-13]。

1.2 燃料储存与供给系统

燃料存储与供给系统的核心功能在于持续、稳定且安全地向燃料电池电堆提供符合其严格运行参数要求的燃料气体。在纯氢供给系统中，核心存储设备为高压储氢容器。氢气通过多级减压装置逐步降低压力至电堆所需的入口水平。精密流量控制阀依据电堆功率需求，精确调节氢气供给流量[12]。系统集成了压力传感器、温度传感器、吹扫阀、安全泄放阀、紧急切断阀等安全装置与监测元件，并通过专用氢气管路确保密封性。为提高燃料利用率，通常配备氢气循环泵或引射器，将电堆阳极出口尾气中的未反应氢气循环输送至电堆入口[13]。对于PEMFC，燃料气体通常需要加湿处理，以防电堆内的质子交换膜干燥，从而确保质子传导效率。加湿器通常集成在循环回路或主供气管路中，用于将燃料气体的湿度提升至电堆要求的水平。SOFC使用纯氢燃料时，燃料管路中通常需注入体积分数为3%～10%的水蒸气来抑制阳极镍颗粒的粗化（Ostwald熟化）过程，维持其微观结构稳定性，水蒸气供给一般通过蒸汽发生器或尾气循环模块实现[14]。当系统使用碳氢燃料（如汽油、柴油或天然气）时，需集成燃料重整子系统。液态燃料由燃料泵增压输送，气态燃料则由压力调节器稳定供给。燃料首先进入脱硫单元，该单元通常采用氧化锌脱硫剂吸附脱除硫化物，以防止下游催化剂中毒。净化后的燃料随后在重整反应器转化为主要成分为氢气和CO的富氢合成气。合成气中的CO需进一步净化处理，通过水气变换反应器，利用催化反应将CO体积分数大幅降低至1%以下。对于要求严格的PEMFC，还需通过CO深度净化设备降至低于10 ppm（12.5 mg/m^3）级别，以避免电堆催化剂中毒[15]。净化后的富氢气体流经换热器，并通过包含控制阀的配气系统精确调节其流量与温度后供给电堆。同样，对于PEMFC电堆，此最终供给的富氢合成气通常也需要进行加湿处理，以满足其燃料侧对湿度的严格要求。若SOFC直接使用碳氢燃料，则需维持高水碳比防止积碳，同样通过蒸汽发生器或尾气循环模块供给水蒸气，以防止重整过程及阳极内部发生积碳[16]。

1.3 氧化剂供给系统

氧化剂供给系统的核心任务是向电堆阴极提供足量的氧气（通常来自空气）并维持适当压力。对于PEMFC，通常需要较高的工作压力来维持质子交换膜的水合状态和中低温运行，因此，多采用效率较高的离心式空压机或压比可调的螺杆式空压机。空压机是辅助系统中耗能最大的部件，其能

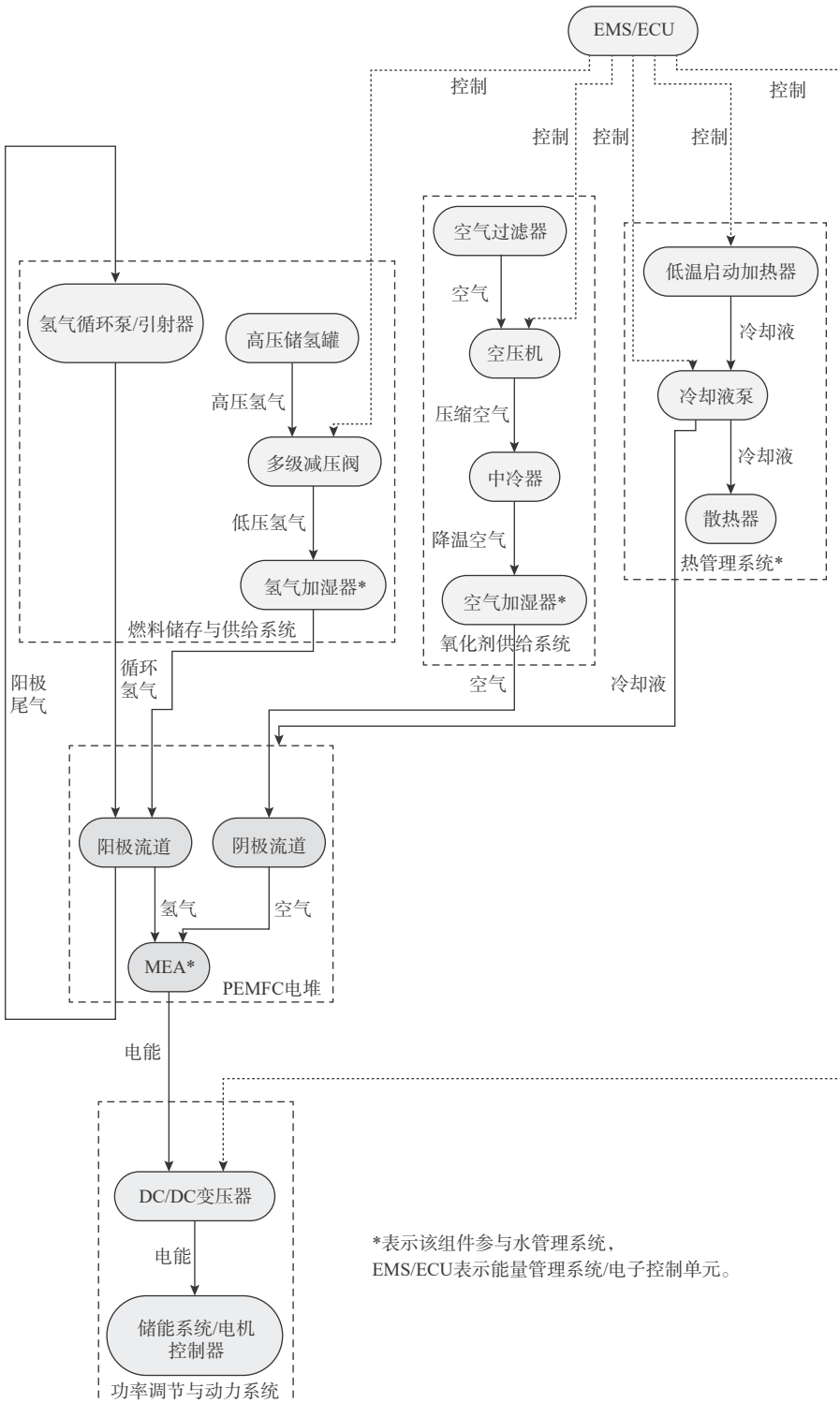


图 1　PEMFC 动力系统架构

Fig. 1　PEMFC power system architecture

耗最高可达电堆输出功率的 30%[17]。相比之下, SOFC 在高温下运行, 阴极侧压力要求较低, 接近常压, 因此, 多选用结构更简单且能耗更低的鼓风机或罗茨风机。净化后的气流由空气过滤器处理, 流量计负责实时监测其流量[18]。对于 PEMFC, 压缩后的高温空气需要经过中冷器降温, 并由加湿器调节湿度以确保质子交换膜充分水合。SOFC 则利用空气预热装置回收尾气余热来提升进气温度, 通常无需额外加湿。背压阀用于稳定整个系统的工作压力。温度传感器、压力传感器和湿度传感器等关键传感器部署在压缩机出口、热交换器、电堆入口及背压阀等关键位置, 提供实时状态反馈。整个系统在闭环控制策略下运行, 依据传感器数据和负载需求, 动态调节空压机或风机的转速、背压阀的开度、加湿量或预热功率等参数, 确保为电堆持续稳定地提供压力、流量、温度、湿度及洁净度均符合要求的氧化剂[19]。

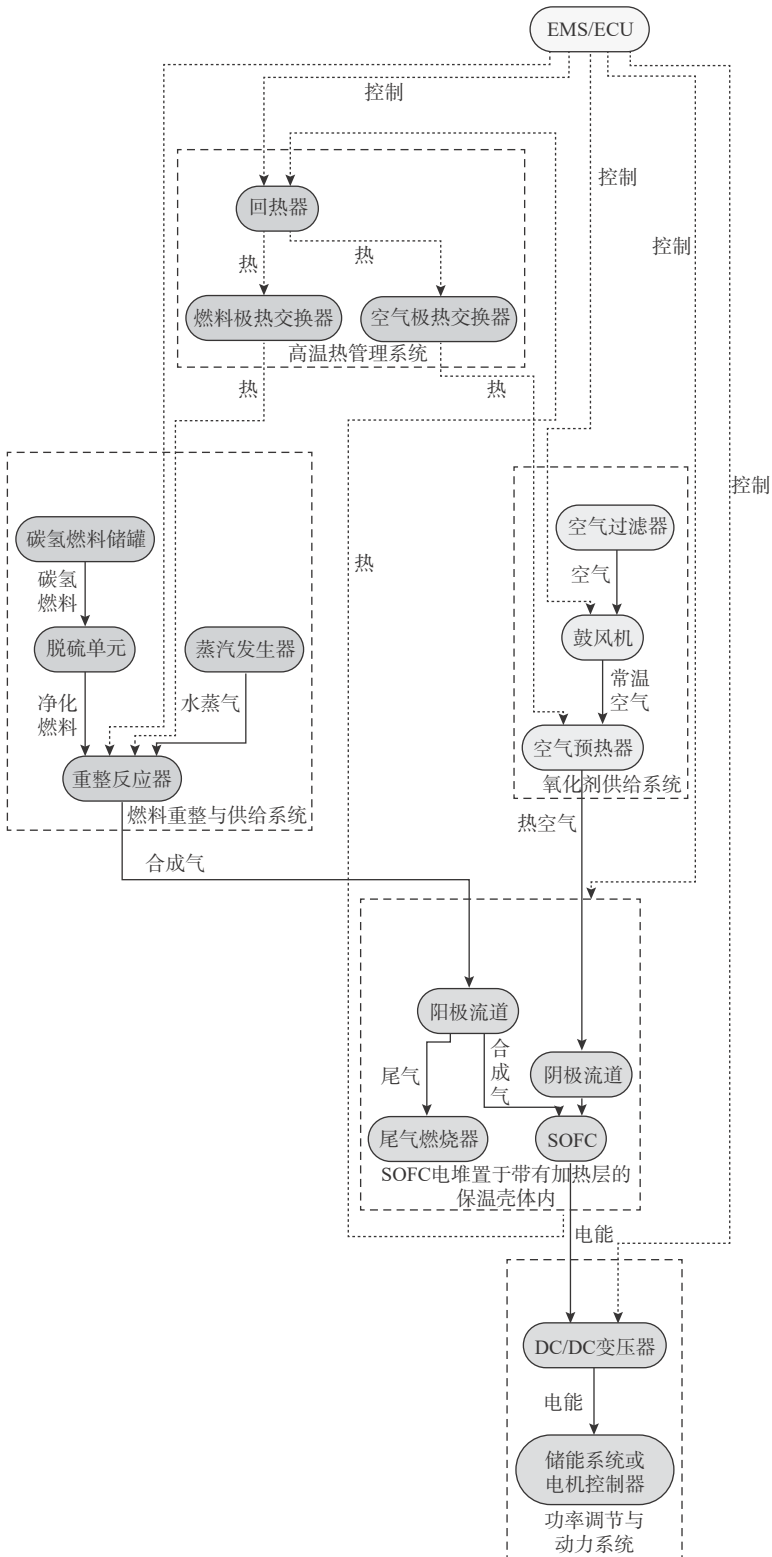
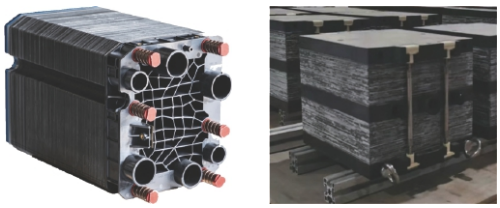


图2 SOFC动力系统架构

Fig. 2 SOFC power system architecture

1.4 热管理系统

热管理系统是保障燃料电池电堆高效、可靠及长期稳定运行的核心子系统，其设计策略因电池类型固有的工作温度与电化学特性差异而迥异。在PEMFC中，较低的工作温度区间（60～80 ℃）结合其实际运行中存在的输出电压损失，导致约40%～60%的能量以废热形式释放。加之其运行温度逼近膜脱水临界点且关键材料耐热性有限，实现电堆内部高效且均匀的热量散失成为热管理的首要目标，以防止性能衰减或材料失效[20]。为此，PEMFC普遍采用液冷方案，核心在于由循环泵驱动冷却介质流经精密集成于双极板内部的冷却流道，直接带



(a) AIR LIQUIDE生产的 20 kW PEMFC电堆　　(b) 索福人生产的 5 kW SOFC电堆

图 3　PEMFC 与 SOFC 电堆实物图

Fig. 3　Photographs of PEMFC and SOFC stacks

走反应热量。该系统还包含散热器、膨胀水箱、维持冷却液绝缘性的去离子装置、用于低温启动的电加热器、多点温度传感器及根据实时工况动态调节冷却强度的控制系统。SOFC 系统则呈现截然不同的热管理逻辑, 其运行在 600～1 000 ℃ 的极高温度下, 实际系统效率通常在 40%～60% 之间, 在此高温下产生的大量废热具有极高的品质[21]。因此, SOFC 热管理的核心挑战在于利用高性能保温材料严密隔绝电堆, 以最小化热损失并维持运行温度。同时, 通过燃烧器处理未反应燃料产生高温排气, 并借助空气预热器、燃料预热器、蒸汽发生器(为燃料重整提供蒸汽)及利用废热驱动的燃料重整器等热回收装置, 深度利用电堆与排气中的高品位余热, 以提升系统总能效。对于发电应用, 该热能还可进一步驱动蒸汽轮机实现联合循环。整个高温系统的稳定运行, 高度依赖复杂控制系统对热平衡的精准管理[22]。

1.5　水管理系统

水管理系统是保障 PEMFC 稳定运行的核心子系统, 通过协调电池内部水的生成传递与相态转换过程, 解决质子交换膜充分水合维持高质子传导率与阴极液态水积聚阻碍反应气体传输之间的固有矛盾[23]。该系统的实现深度集成于电堆相关子系统。膜内水的平衡通过两种相反的传输过程实现。在电渗拖拽效应下, 水随质子由阳极向阴极迁移; 与此同时, 阴极侧的水则依赖膜内水含量(λ, H_2O/SO_3H)梯度向阳极反向扩散, 对阳极侧形成补水。阳极侧水合保障直接由燃料供给系统集成于管路的加湿装置实现, 该加湿器向氢气注入水分以补偿电渗拖拽引发的水损失。阴极侧湿度调控则由氧化剂供给系统配置的空气加湿器执行, 该单元协同中冷器温度调节功能精准抑制过量冷凝水生成[24]。液态水管理依托三重机制协同作用, 紧贴催化剂层的气体扩散层经疏水改性在微孔域构建毛细压力梯度, 持续驱动液态水脱离催化剂活性位点。双极板表面精密加工的蛇形或交指构型流道利用高速气体流动剪切力剥离电极界面附着液滴, 并将游离液体定向输送至集水区[25]。最终由电堆外壳低点集水区配置的排水阀实时响应流道进出口压差传感器信号, 主动清除积聚液体。热管理系统通过冷却液温度场调控水的饱和蒸气压, 既联动氧化剂供给系统中冷器功率防止低温工况水淹恶化, 又联动燃料供给系统燃料温度规避高温膜脱水风险。闭环控制策略基于湿度传感器、压差传感器、温度传感器及单电池电压监测网络反馈, 动态优化反应气体化学计量比、系统背压设定值及双路加湿单元的输出量[26]。

而 SOFC 运行于 600～1 000 ℃ 高温环境, 生成水以蒸汽相态随尾气自然排出, 无液态水管理需求, 系统复杂性显著降低。

1.6　功率调节与动力系统

燃料电池动力系统的功率调节与控制系统是实现电堆与驱动负载之间高效能量传递与动态协调的关键环节。该系统主要由电力转换单元和控制系统两大部分构成, 并通过集成化的功能设计, 确保动力系统稳定、高效、安全地运行。

电力转换单元负责处理电能形式, 以满足负载需求。其中, DC/DC 转换器将电堆随工况动态变化的输出电压进行宽范围升压或降压, 稳定至动力总线的标称电压, 为驱动电机控制器或高压储能系统提供稳定直流电。若负载采用交流电机, 则需通过 DC/AC 逆变器将直流电转换为频率、幅值及相位精准可控的交流电。

控制系统包括专用控制器、多传感器网络及执行机构, 构成一个闭环架构。其持续采集并融合电堆单体电压/电流、温度分布梯度、反应气体流量压力及冷却回路参数等关键数据。通过执行分层控制策略, 以实现空压机转速、氢气供给阀开度、冷却液流量及功率转换器工作点的精细化动态协调并管理包括冷热启动、恒功率运行及负载跟随等多种模式的无扰切换。依托实时诊断触发涵盖过压限流、低温吹扫与气体饥饿防护等的多级安全保护机制, 并运用自适应氧化学计量比控制及温度-压力协同寻优等策略, 最大化系统净输出功率。该系统的控制精度、动态响应速度及鲁棒性, 直接决定了动力系统的输出稳定性、瞬态负荷适应能力、故障容错水平及电堆整体耐久性[27-28]。

2　PEMFC 航空动力系统

2.1　PEMFC 技术原理与特点

PEMFC 是一种基于电化学原理、直接将氢气的化学能转化为电能的装置。其核心在于利用质子交换膜(如 Nafion®膜)在 60～80 ℃ 的低温环境

下传导质子，从而实现高效发电[29]。其工作原理如图4所示，在阳极上，氢气在催化剂作用下氧化分解为质子(H^+)和电子(e^-)。质子通过膜内的水合通道迁移至阴极，而电子则经由外电路形成电流做功。最终在阴极上，氧气与抵达的质子及电子结合生成水并释放能量。这一独特的低温运行机制，既奠定了PEMFC高效率和快速响应的基础，也使其运行严重依赖膜的良好水合状态。因此，水热管理的平衡至关重要。

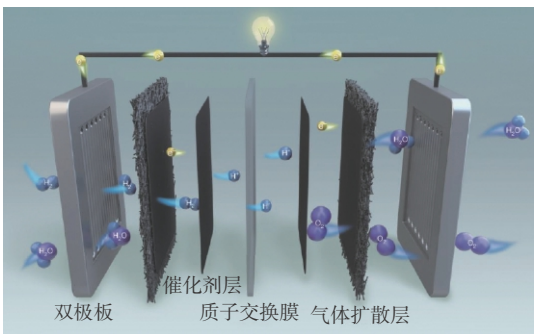


图4 PEMFC工作原理图[30]

Fig. 4 PEMFC schematic diagram[30]

基于该原理，PEMFC展现出高功率密度、高能量效率及快速响应等核心优势。目前，量产的功率密度已达4.4 kW/L(堆级，含端板)，2030年路线图目标指向6 kW/L以上，这得益于纳米级催化剂的应用和结构的创新[31]。通过直接的电化学反应，其电能转换效率高达50%～60%；在热电联产(combined heat and power, CHP)模式下，通过回收利用系统产生的低品位废热(如用于供暖或预热)，其综合能源利用率可超过80%，远超传统热机[32]。此外，低温运行使其具备数十秒内快速启动和对负载变化毫秒级响应的能力，特别适合动态应用场景[33-34]。

然而，该技术也面临固有挑战。一方面，其对燃料纯度极为敏感，微量的CO等杂质即可导致催化剂严重中毒，通常要求氢气纯度高于99.97%，或需开发抗毒催化剂[35]。另一方面，水热管理极具复杂性，必须在运行中精准平衡膜的湿度，既要防止脱水导致电阻增大，又要避免水淹阻碍气体传输，这对系统控制提出了持续性的高要求。这些基本特点与挑战共同构成了PEMFC技术应用于航空领域的出发点与基础。

2.2 PEMFC系统的核心优势

PEMFC技术凭借其独特的工作特性，在航空动力领域展现出显著的多维优势，为其应用奠定了坚实基础。其核心竞争力源于4个方面，即经过验证的技术成熟度、卓越的动态响应特性、优异的功率性能和卓越的环境兼容性。

首先，技术成熟度高，可靠性经过验证，大幅降低了航空应用的开发风险。PEMFC并非停留在实验室阶段的概念性技术，而是在汽车与固定式发电领域已实现大规模商业化应用的成熟平台。丰田Mirai、本田Clarity等量产车型的长期稳定运行，积累了海量的工程经验与宝贵的耐久性数据，其电堆寿命突破1万小时已成为行业标杆。这种经过实践检验的可靠性，为技术移植到航空领域提供了强有力的保障。一个典型的成功案例是H2Fly的HY4验证机项目，为其提供PEMFC系统的Hydrogenics公司同样为汽车提供燃料电池动力系统，如图5所示。该型飞机成功获得了斯图加特机场所在地航空监管机构的试飞许可，这充分表明，利用成熟技术极大缩短了概念到原型的研发周期[36-37]。



(a) 搭载Hydrogenics公司燃料电池系统的公交车



(b) HY4验证机搭载的PEMFC系统

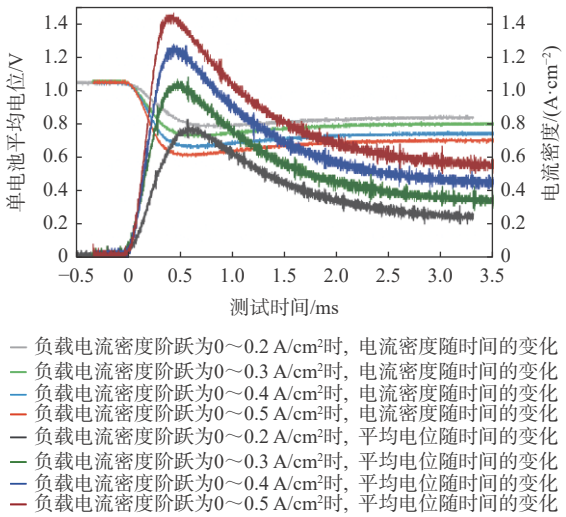
图5 PEMFC动力系统从车用领域向航空领域的迁移应用

Fig. 5 Migration of PEMFC power systems from automotive to aviation applications

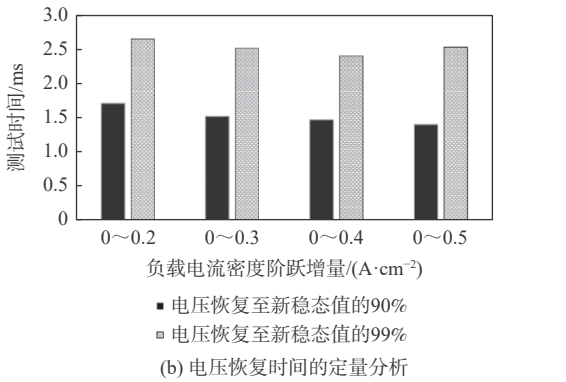
其次，卓越的动态响应能力是PEMFC适应航空复杂工况的关键特性。系统具备极快的启动速度与功率调节能力，即使在−30 ℃的低温环境下，亦可在30 s内完成冷启动，并在数分钟内达到额定功率[38]。面对负载的剧烈波动，如飞机起飞阶段功率需求的阶跃上升，PEMFC展现出优异的瞬态响应特性。在美国国家航空航天局(National Aeronautics and Space Adiministration, NASA)的测试中，当电流密度在0～0.5 A/cm^{2}范围内发生阶跃变化时，端电压在3 ms内即可从初始跌落恢复至对应新稳态值的90% [39]，如图6所示。该性能显著优于存在惯性延迟的燃气涡轮发动机，尤其适用于无人机频繁起降及载人飞机爬升加速等对动力响应要求严苛的场景。

此外，在功率密度方面，PEMFC展现出可持续竞争力。当前航空级系统的功率密度已达1.5 kW/kg(系统级，不含储氢设备)，通过纳米级Pt-Co合金催化剂优化电化学活性面积，以及使用新型高温质子交换膜，未来功率密度有望达到3 kW/kg(系统级，不含储氢设备)，逐步接近传统涡桨发动机水平[40-41]。

最后，在环境兼容性方面，当采用可再生能源制取的绿氢为燃料时，系统全生命周期可实现近零碳排放。此特性完全契合欧盟“净零航空”战略目标，不仅助力航空业深度脱碳，更能有效规避日



(a) 负载阶跃下的动态响应过程



(b) 电压恢复时间的定量分析

图 6　PEMFC 的毫秒级瞬态响应性能[39]

Fig. 6　Millisecond-scale transient response performance of PEMFC[39]

益严格的碳边境调节机制带来的贸易壁垒, 为运营商创造长期合规性与市场优势。

2.3　航空 PEMFC 系统设计关键

将 PEMFC 技术应用于航空领域, 必须解决一系列极端环境下的工程挑战。这些挑战的严峻程度与技术侧重点因飞行器的运营场景而异。本节聚焦于氢气存储、水热管理、高空空气供给及系统集成与轻量化四大关键技术瓶颈。

2.3.1　氢气储存方案

氢气储存是实现氢能航空的核心, 是所有氢动力航空器共同面临的基础性挑战。储氢技术的选择, 直接关系到飞行器的有效载荷与航程能力, 而不同应用场景(如小型无人机与大型客机)对能量密度、安全性和技术成熟度的需求侧重点截然不同。当前, 各技术路线均面临如何平衡高能量密度需求与储氢系统自重、体积及安全性之间的紧张关系。不同应用场景需依据其航程与规模, 做出不同的技术取舍。

70 MPa 高压气态储氢技术成熟度高, 适配现有加氢基础设施。但其发展受限于一个基本物理约束: 储氢量与压力成正比。为确保高压安全, 所需的容器壁厚与质量会随压力非线性增长。这导致系统质量储氢密度提升存在瓶颈, 使其更适用于对储氢密度要求不高的小型无人机或中短期内的实验型客机, 无法应用于大型远程氢能飞机[42]。此外, 氢气极易爆炸, 必须配备严密的泄漏监测、压力泄放和碰撞防护等安全措施, 这些措施不可避免地提高了系统质量和复杂度[32]。目前, 该领域的技术进展主要聚焦于通过先进复合材料和优化结构设计来减轻储罐自重, 属于工程层面的渐进式改进。

液氢存储方案能提供更高的能量密度, 是未来支线及干线飞机实现长航程的必然选择。研究表明, 采用液氢方案可使飞机最大起飞质量较传统方案降低 3.5%～11%[43]。然而, 液氢需要维持−253°C 的超低温环境, 面临着蒸发损失(每日质量损失率约 3%～5%)、材料低温脆化及系统复杂性等严峻挑战[43]。因此, 未来的技术发展必然指向高性能真空多层绝热结构、低漏热连接技术及耐低温材料和部件的深度研发, 以构建一个高效、可靠且轻量化的机载液氢存储与管理系统。

固态储氢技术凭借其高体积储氢密度与本质安全特性, 被视为极具潜力的氢储运方案。其根本瓶颈在于储氢材料的热力学与动力学性质, 目前, 具备实用性的主要是 AB_5、AB_2 型等金属氢化物, 但其质量储氢率通常<2.5%。而高容量的轻质材料如镁基合金(MgH_2, 7.6%)、配位氢化物(如 $NaAlH_4$, 5.6%)面临着放氢温度高、动力学缓慢及循环性能差等关键技术瓶颈[44]。当前, 固态储氢已在分布式储能等场景中得到示范应用, 但其系统质量储氢密度与快速响应能力距航空动力系统的大规模商用要求仍有显著差距。因此, 该领域的技术突破有赖于材料科学的创新, 重点在于降低储氢材料反应势垒、增强结构稳定性, 并同步开发与之匹配的高效传热传质储氢罐设计[44]。

2.3.2　水热管理

PEMFC 低温运行的特性使其水热管理在航空极端环境下变得尤为关键与复杂。其根本矛盾源于膜电极对精确水热环境的稳态需求与飞行全剖面中温度、压力剧烈动态变化之间的持续冲突。

在热管理方面, 挑战的核心是 PEMFC 产生的低品位废热与机上热需求(如客舱供暖、进气预热)在品位和时序上的不匹配。为提升系统能量利用效率, 通常需要设计紧凑式换热器回收电堆废热。然而, 热源与负荷在品位和时序上的不一致常导致供需失衡与能量损失[45-46]。因此, 未来的技术路径需超越简单的热回收, 转向基于热需求预测的智能

热流分配系统，并探索集成相变材料的缓冲单元，以实现能量的高效管理与利用。

结冰风险是航空PEMFC在低温环境下面临的独特挑战，对需在高空或寒冷地区运行的飞机尤为严峻，贯穿于启动与运行全程[47]。其本质是反应生成水在低温驱动下发生的相变失控。在冷启动阶段，阴极生成水易在气体扩散层孔隙中凝结并冻结，直接阻塞气体传质通道，导致冰堵。为应对该风险，需采取多级防护策略。在材料层面通过聚四氟乙烯（polytetrafluoroethylene, PTFE）改性气体扩散层，降低材料表面张力，以延缓结冰进程[48]。在系统层面，则需依赖电加热膜提供快速解冻能力，并执行阳极吹扫策略，利用干燥气体在启动和关闭阶段清除电堆内残留水分。此外，极端低温会威胁冷却回路的安全性，需添加低电导燃料电池专用防冻液。但工质流动特性的变化可能影响散热效率，需重新优化冷却系统设计。

PEMFC的水管理同样是低温运行的关键瓶颈，其核心矛盾在于水的分布与传输平衡。低温会削弱膜的水合能力，若湿度不足，将导致质子交换膜脱水，显著增大电阻，而阴极过量的液态水又会引发水淹，阻碍氧气传输[49]。这构成了膜脱水需求与电极防水淹要求之间的精细权衡。解决这一矛盾，必须通过高精度湿度传感器实时监测，并结合动态流量与湿度控制算法，精准调节反应气体的供给，以实现水在生成、电渗迁移与排出之间的动态平衡。这种精细化控制显著提升了系统的复杂性与技术难度。

2.3.3 高空空气供给系统

空气供给系统是PEMFC系统维持阴极侧氧化剂供应的关键，其性能直接决定电堆的输出能力。其核心矛盾在于满足电堆阴极反应所需的高质量氧化剂与为此在高空低密度、含污染大气环境中付出的巨大辅助功耗代价之间的权衡。

在支干线客机及大型无人机巡航阶段（通常高于6 000 m），随着飞行高度增加，大气密度与压力持续下降（高度每升高1 000 m，空气密度降低约11%），这使阴极测氧分压不足。为维持电堆高效运行，空压机必须提升压比或流量，导致其功耗急剧增加，有时可占电堆输出功率的30%以上[17]。因此，补偿性能衰减与抑制寄生功耗激增的矛盾成为制约系统的净输出效率与高空适用性的关键。在该背景下，离心式空压机凭借其高功率密度和快速动态响应（转速可达10^5 r/min以上），成为应对高空低压环境的首选方案[50]。提升空压机整体效率的关键在于优化其高速永磁同步电机的效率与流道气动设计[51]。在系统层面，则需通过采用更高压比（如采用多级压缩中间冷却技术）或引入涡轮增压技术等具备排气能量回收的复合增压方案，实现能量的梯级利用，从而系统性地提升高空工作条件下的总效率[52]。

对于在万米高空巡航的干线客机，其PEMFC的空气供给系统面临高空大气环境中的臭氧与悬浮颗粒物带来的独特挑战。高空臭氧具有强氧化性，促使阴极铂催化剂发生溶解，溶解后的铂物种迁移并在膜内再沉积，导致电化学活性面积减少并阻碍质子传导。同时，臭氧会攻击气体扩散层中的PTFE，削弱其疏水功能，从而影响水管理效率。此外，臭氧还可能诱发自由基反应，加速聚合物电解质膜的化学降解与碳载体腐蚀，共同导致电池性能的不可逆衰减[53]。大气中的细颗粒物则可能堵塞气体扩散层孔隙，阻碍反应气体传输，导致浓差极化加剧。

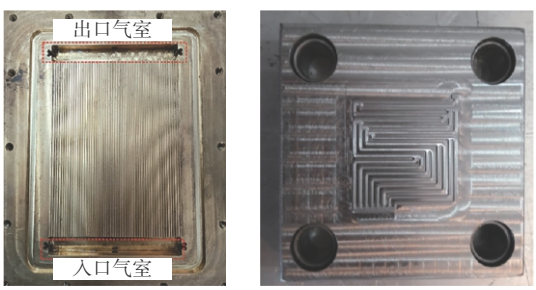
上述问题构成了进气净化需求与最小化附加流阻之间的矛盾。提高过滤精度与层级虽能有效去除污染物，但也会增加进气压降，反向增加空压机负载，降低系统净输出[54]。解决这一矛盾的核心在于发展低阻高效的多层净化方案。系统需配备集成式多级过滤装置。首层采用粗效滤网拦截大粒径颗粒，中间层布置高性能活性碳纤维毡（比表面积达1 500～2 000 m^2/g），通过物理吸附与催化氧化作用高效去除臭氧，末级采用纳米纤维膜（孔径为0.1～1 μm）实现亚微米级颗粒的精滤[55]。为优化流阻性能，需依托计算流体力学仿真，对滤材结构形态与流道布局进行设计优化，在保证污染物截留率的前提下，最大限度降低进气阻力。同时，可集成实时压差与污染物浓度监测模块，依据环境空气质量动态调整过滤策略，从而在保障电堆长期稳定运行的同时，维持系统整体运行效率。

2.3.4 系统集成与轻量化

PEMFC动力系统在飞行器上的成功应用最终取决于能否通过高度集成的设计，在满足极致轻量化要求的同时，确保其高可靠性与可维护性。这一系统级工程所面临的根本矛盾在于，飞行器对动力系统性能、质量、体积的极限追求，往往与系统的工程可靠性、可维护性及全生命周期成本产生直接冲突。化解这一矛盾，需要从部件创新与系统架构2个层面协同推进。

在部件层面，轻量化目标的实现主要依赖于微纳尺度结构设计与先进材料的应用，以突破传统技术条件下的性能与质量瓶颈。其中，微通道技术发挥着关键作用。例如，在热管理方面，采用微通道

的冷却板(通道宽度通常为 20～200 μm, 见图 7(a)), 可利用其巨大的比表面积与流动沸腾传热机制, 可在有限体积内高效排除电堆产生的热量。并通过内置孔板或挡板等优化结构保障冷却工质分布的均匀性, 对维持电堆温度一致性具有关键作用。同样, 在反应气加湿方面, 微通道式换热/加湿器依托类似的强化传热传质原理, 在实现高效加湿与热交换的同时, 显著缩减部件体积, 较传统设计可减少 50% 以上。与此同时, 材料与结构创新是轻量化的另一重要支柱。采用超薄金属双极板替代传统石墨板, 并借助化学蚀刻或激光加工技术在其表面加工精密流道, 如图 7(b) 所示, 可大幅降低电堆核心部分的质量。发展一体化气体扩散层设计, 如通过激光雕刻在碳纸基材上直接制备出集成波形流场与微通道脊结构。这种设计不仅强化了反应气体的传输与液态水的排出, 有助于缓解高电流密度下的水淹现象, 更能降低系统流阻, 从而间接减少辅助系统的功耗[56]。



(a) 微流道冷却板[57]　　(b) 微流道双极板[58]

图 7　基于微通道的轻量化燃料电池部件[57-58]

Fig. 7　Microchannel-based components for fuel cell lightweighting[57-58]

然而, 若极致的轻量化与高度集成设计缺乏系统层面的顶层设计考量, 往往会以牺牲可维护性和升级潜力为代价。因此, 在系统架构层面引入模块化设计理念, 是平衡上述矛盾、满足航空严苛运营要求的关键路径。模块化通过将整个系统按功能划分为相对独立的子模块(如电堆、空气供应、热管理、控制等), 并定义标准的机械、电气与流体接口, 实现了物理与功能的解耦。这种架构对于维护窗口短、系统复杂度高的支线及干线客机尤为重要。在实际执行中, 采用快速连接器与法兰式固定结构可实现关键模块的快速拆装, 大幅提升维护效率。在开发流程上, 坚持“先独立验证, 后系统集成”的策略, 通过预先完成各模块的测试以确保其接口兼容性与功能完整性, 从而从源头控制系统级风险, 奠定高可靠性的基础。

综上所述, PEMFC 动力系统凭借其低温快速启动、毫秒级动态响应及高功率密度等特性, 在航空脱碳领域展现出独特的应用价值。其核心竞争力不仅源于兼容绿氢的零排放潜力、模块化架构带来的高维护性, 更得益于已在汽车领域得到验证的技术成熟度。然而, 如前述各子系统分析所述, 将其成功应用于航空器仍面临严峻挑战。储氢系统需在有限质量与体积内实现超高能量密度; 水热管理必须在动态环境中维持膜电极精准的水合与温度平衡; 空气供给系统面临补偿高空低压、净化污染物与控制寄生功耗之间的艰难权衡; 而系统集成则需在极致轻量化与高可靠性、可维护性之间寻求最优解。未来的技术发展, 必须超越对单一问题的应对, 转而沿着轻量化/高密度储氢、智能水-热协同管理、高效低阻空气供给及模块化集成等系统性路径进行协同攻关, 才能使 PEMFC 动力系统在无人机及城市空中交通领域实现规模化应用。

3　SOFC 航空动力系统

3.1　SOFC 技术原理与特点

SOFC 是一种将燃料化学能直接、高效转化为电能的全固态能量转换装置, 主要由阳极、阴极和电解质 3 部分组成。其核心特征在于高温运行(通常为 600～1 000 ℃)。该温度下, 电解质材料的离子电导率显著提升, 阴阳极的电化学反应阻力显著降低。这一特性是成就其超高效率与燃料灵活性的基石, 也决定了其技术范式更适用于对效率极为敏感、而对快速启动要求相对宽松的中远程航空动力系统。

其工作原理如图 8 所示。以氢气作为燃料气为例, 氢气先吸附于阳极表面, 再经多孔阳极扩散至阳极与电解质的界面。同时, 空气里的氧气在阴极吸附, 变成氧离子, 随后在化学势驱动下, 经氧离子传导的电解质扩散至阳极与电解质的界面, 与氢气发生反应, 生成水并释放电子。电子经外电路回到阴

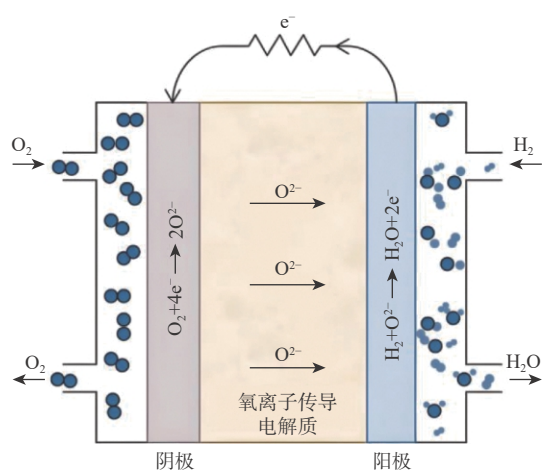


图 8　SOFC 工作原理图

Fig. 8　SOFC schematic diagram

极，从而实现电池的闭合回路和能量转化[59]。该过程不受卡诺循环的限制，因而具有很高的理论效率。

SOFC因其独特的高温运行特性，展现出若干显著优势。首先，高温环境显著提升了SOFC对燃料杂质的耐受性。在典型Ni基阳极上，CO的吸附作用相对较弱，且CO可直接参与电化学反应，发生电化学氧化生成CO_2，从而避免了其在催化剂活性位点上的过度积累[60]。这一特性与低温PEMFC中CO极易导致铂基催化剂中毒的现象形成鲜明对比。其次，高温条件下硫化物（如H_2S）在Ni基阳极表面的吸附能较低，赋予SOFC对较低体积分数（<0.002%）硫化物的较强耐受性。即使在0.005% H_2S气氛中发生中毒，移除H_2S后，其电化学性能也能在高温运行中恢复，这一特性显著优于PEMFC[61-62]。同时，SOFC的高温运行为碳氢燃料（如航空煤油、生物质衍生燃料、甲醇等）的原位重整提供了必要条件。这些燃料可在阳极催化层内部或其邻近区域，通过蒸汽重整（steam reforming, SR）或部分氧化重整（partial oxidation reforming, POX）等反应，直接转化为富含氢气和CO的合成气，随即参与阳极电化学反应。此过程有效降低了对复杂外部燃料预处理系统的依赖，简化了系统架构[63]。此外，SOFC运行产生的高温排气（通常高于600 ℃）蕴含高品位的热能。这使其排放的余热非常适合于构建高效的CHP系统，或用于驱动底部循环（如GT），通过构建联合循环系统将热能进一步转化为电能，从而显著提升整个能源系统的总效率[64]。

3.2 SOFC系统的多燃料适应性

SOFC在航空动力领域的战略价值核心在于其多燃料适应性，这一特性显著区别于低温燃料电池技术（如PEMFC）。其根本技术基础源于SOFC的高温运行特性，高温条件为碳氢燃料及氢载体的高效重整提供了必需的热力学与动力学环境，同时结合阳极对重整产物的本征耐受性，共同构成了SOFC燃料灵活性的核心支撑。

3.2.1 燃料重整机制

将复杂燃料转化为可直接参与电化学氧化的合成气（H_2+CO）的重整方法包括SR、POX和自热重整（autothermal reforming, ATR）3种。

SR是一种强吸热反应，碳氢燃料与水蒸气在催化剂（如Ni基催化剂）作用下转化为富氢合成气。该技术具有最高的产氢率（>70%），尤其适合甲烷及轻质燃料[65]。其核心优势在于最大化电效率。然而，强吸热特性需依赖外部热源，导致热管理系统复杂化。同时存在启动速度慢、系统传热传质复杂、热惯性大、动态响应滞后、难以匹配飞行瞬态工况等问题[66]。此外，反应器低温区易因Boudouard反应引发积碳，需维持高水碳比（S/C>3.0）缓解[67]。这样高的水碳比要求意味着需携带大量反应水，使其在航空动力系统中的应用面临巨大的挑战。

POX是碳氢燃料与不足量的氧气反应发生的不完全燃烧反应，生成以CO和H_2为主的合成气，具有反应产热高、启动速度快、产氢速率高的特点[68]。其优势在于无需外部持续供热，能够快速响应，系统传热结构简单，特别适合航空动力动态工况[69]。

ATR通过耦合吸热的SR与放热的POX反应，并精确调控水蒸气和氧气的比例，实现反应体系的热中性。该过程反应条件相对温和，产氢量较高[70]。ATR兼具较高的制氢效率、良好的动态响应能力及对复杂燃料的适应性等优势。尽管存在反应器结构复杂、质量增加及动态调节精度要求高等挑战，其仍是长航时航空器氢能系统中具有竞争力的技术选项[71]。

表1概括了以汽油、柴油、航空煤油等典型重烃燃料为原料时，3种重整路径的性能特征。根据催化剂种类、燃料组分和操作参数不同，具体数值会有所不同，但技术路线的特性趋势基本一致。在实际航空动力系统设计中，需根据具体燃料类型、任务剖面和平台要求进行详细权衡与优化。

表1 重烃燃料3种重整路径的性能对比[71-73]

Table 1 Performance comparison of three reforming pathways for heavy hydrocarbon fuels[71-73]

路径	外部供热需求	典型反应温度/℃	氢气体积分数	H_2/CO体积比	启动时间/min	动态响应	系统质量与体积	适用场景
SR	是	750～900	0.70～0.80	≈3～4	>20	分钟级	大	固定式
POX	否	1 100～1 200	0.35～0.45	≈1～1.5	<1	秒级	紧凑	紧凑设备
ATR	否	900～1 150	0.40～0.55	≈2～2.5	10～20	<1 min	略大于POX	车载/航空

3.2.2 SOFC适配的燃料类型

SOFC能够兼容从传统化石衍生燃料到零碳可持续能源的广泛谱系，这一特性为长航程航空动力系统的低碳化转型提供了关键路径。具体适用的燃料类型包括但不限于下列燃料。

氢气是SOFC最清洁的燃料选择。然而，如2.3.1节所述，氢气的低体积能量密度对航空应用构成显著挑战。目前，高压气态储氢瓶（如25～35 MPa

碳纤维缠绕钢瓶)质量占总质量超过 98%, 且存在着氢脆的问题, 严重影响了储氢设备的安全与寿命[74]。低温液态储氢液化过程耗能极大, 对容器绝热性能要求高, 设备投资成本高, 因此, 仅在大规模存储, 长距离运输中具有一定的优势[75]。正是这些储氢瓶颈, 反向凸显了 SOFC 系统燃料灵活性的战略价值。

相比之下, 碳氢燃料, 包括汽油、柴油、航空煤油、液化天然气、SAF 及醇类燃料, 凭借其成熟的供应链和较高的能量密度, 成为航空领域当前及低碳化转型中期更现实的选择。SOFC 可直接利用或通过内部重整技术转化以下碳氢燃料。

汽油主要成分为 C_4～C_{12} 的烃类, 是轻型内燃机的常用燃料[76]。符合国 VI 标准[77](国家市场监督管理总局2016 年 12 月 2 日发布)的汽油具有超低硫含量(0.000 3% ～0.000 5%), 使其适用于作为 SOFC 的燃料。在 SOFC 应用中, 汽油既可直接输送至阳极作为燃料使用[78], 也可通过集成于电堆内部或外部的重整器进行 ATR 后供给利用[79]。

柴油中碳氢化合物的 H/C 摩尔比范围为 1.8～2.3, 其含碳量主要分布在 C_{10}～C_{20} 区间, 尤以 C_{16} 组分为主[72]。相较于其他碳氢燃料, 柴油具有较高的密度, 因而单位体积氢含量较高, 但其硫含量也显著更高, 如−10 号军用柴油的硫质量分数超过 0.003% [80]。鉴于柴油相对分子质量更大、硫含量更高的特性, 研究中通常采用外置重整器方案, 通过重整将柴油转化为合成气, 再供给至 SOFC 利用[81-83]。

航空煤油主要包括民用型(如 Jet A-1、Jet A 和 RP-3)和军用型(如 JP-8)2 类。以美国标准航空煤油 Jet A 为例, 其主要成分为 C_8～C_{16} 的烃类混合物, H/C 摩尔比范围为 1.6～2.0[84]。值得注意的是, 航空煤油通常含有较高的硫含量。Xu 等的研究综合了文献报道数据, 指出其平均值可达 0.071 4% [71]。这种高硫特性对 SOFC 系统的设计构成了显著挑战。因此, 开发高效紧凑的脱硫反应器是提升 SOFC 系统航空煤油兼容性的关键环节[85]。

天然气是固定式 SOFC 系统最广泛使用的燃料[86]。液化天然气是天然气经干燥、脱酸处理并在约−163 °C(110 K)以下低温液化制得的燃料, 其主要成分为甲烷, 体积分数通常超过 75%。与航空汽油相比, 储存相同能量时, 液化天然气储罐体积需增大 47.41%, 但燃料质量可减轻 16%[87]。虽然将天然气直接通入 SOFC 阳极进行电化学反应在技术上是可行的[88], 但目前更成熟的技术路径是通过重整器将天然气转化为富氢合成气[86]。

生物质类 SAF 原料来源广泛, 生产工艺多样。动植物油脂和废弃油脂可通过催化加氢或高温水热解技术转化为航空燃料; 糖类和木质纤维素则主要经发酵或生物催化过程转化为醇类, 并进一步提纯; 而城市垃圾和生物质废弃物则可先气化为合成气, 再通过催化合成工艺生产航空煤油。得益于其与现有航空基础设施的高度兼容性, 大规模应用 SAF 被视为当前航空运输业实现深度脱碳最具可行性的方案之一[89]。根据美国 ASTM D7566 标准[90], SAF 的硫质量分数低于 0.001 5%。这一显著的低硫特性使其成为 SOFC 驱动航空器更为理想的燃料来源。

甲醇等液态醇类燃料在储运安全性和基础设施兼容性方面具有优势[91]。其分子结构简单, 重整温度(如甲醇约 200～300 °C)显著低于长链碳氢化合物, 有助于简化系统设计并提升启动响应速度[92]。此外, 甲醇可由可再生能源电解水制取的氢气与碳捕集获得的 CO_2 合成, 实现全生命周期的低碳或零碳排放[93]。甲醇可直接用作 SOFC 的燃料, 有利于电堆小型化设计。然而, 相较于氢气燃料, 其更复杂的电化学反应动力学特性往往导致 SOFC 的功率密度和耐久性降低 [94]。

氨是全球产量最大的无机化工产品之一, 2021 年总产量约为 1.9×10^{11} kg, 其含氢质量密度高(17.7%), 是 SOFC 的理想燃料。得益于易液化的特性(−33℃常压或 1 MPa 压力常温下即可液化), 液氨可通过车船或管道便捷运输, 储运基础设施完善且制备工艺成熟[95]。更重要的是, 氨可通过可再生能源电解水制氢、再与空气中的氮气合成, 实现“绿氨”生产[96]。作为 SOFC 燃料, 氨既可直接输送至阳极原位分解, 也可在外部重整器中裂解产氢后供给[97]。研究表明, 直接利用氨的 SOFC 运行状态稳定性良好[98]。

图 9 直观展示了各类候选燃料在航空场景下的储能特性。传统液态碳氢燃料(如汽油、柴油和航空煤油)仍以最高的体积能量密度占据航程优势, 这解释了为何它们是目前大型航空器的首选。氢气虽质量能量密度出众, 却因储运体积庞大而受限。甲醇、乙醇等绿色合成燃料则介于两者之间, 为兼顾航程与储运可行性提供了折中且现实的解决方案。

3.2.3　燃料灵活性对航空的意义

燃料灵活性作为 SOFC 航空动力系统的核心特征, 为航空业向零排放转型带来了超越单一燃料解决方案的革命性优势, 其意义在航程、载重与运营成本最为敏感的支线及干线航空领域体现得最为充分和深远。

首先, 多燃料兼容性极大提升了燃料灵活性与运营韧性。传统航空业长期依赖航空燃料, 受国际

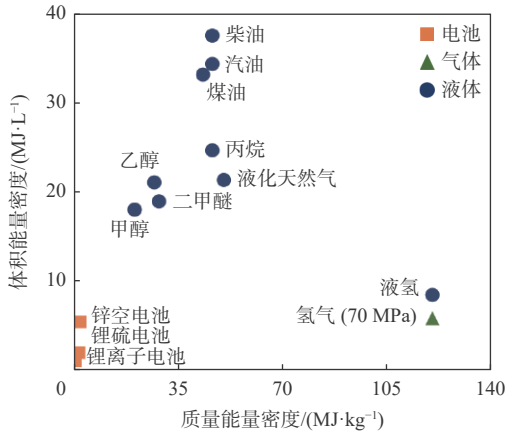


图 9 多元燃料与电池体系的能量密度对比[99]

Fig. 9 Comparation of energy densities among multi-fuel and battery systems[99]

原油价格变动、汇率波动影响大，如图 10 所示，供应链脆弱性高[100]。SOFC 系统允许航空公司根据航线、机场的燃料可用性、成本及可持续性等级，灵活选用多种传统化石燃料或氢气、SAF、甲醇和氨等可持续燃料，显著降低了对单一燃料供应链的依赖风险，增强了应对价格波动和地缘政治冲击的能源安全和运营韧性。

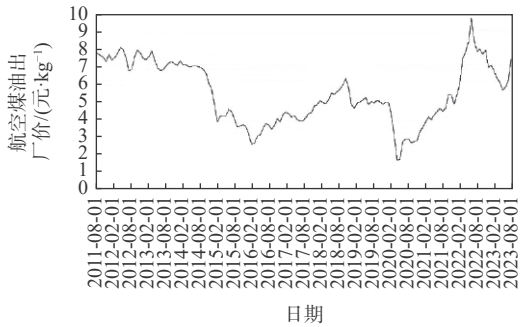


图 10 中国航油出厂价[101]

Fig. 10 Ex-works price of aviation fuel in China[101]

其次，多燃料兼容性，特别是对多种传统化石燃料的支持，实现了与现有基础设施的最大化兼容。建设全新的、大规模的氢或氨燃料机场基础设施成本高昂、周期漫长，是零排放航空早期部署的主要障碍[102]。而 SOFC 动力飞机能够最大程度地复用全球机场现有的燃油储存、运输管道和加注设施。这一特性显著降低了初期部署的门槛和成本，加速了 SOFC 动力飞机市场应用和验证。

在性能层面，多燃料优势直接关联到突破体积能量密度瓶颈并释放航程潜力。航空器对质量和体积极其敏感。气态氢或需要复杂绝热储罐的液态氢，其有效体积能量密度远逊于传统航油。而 SOFC 兼容的液态碳氢燃料（如 SAF、甲醇）或氨则具备显著优势，SAF 体积能量密度接近传统航油，甲醇和氨也远高于氢气方案。结合 SOFC 本身高达 60% 的电化学转化效率，使用这些高体积能量密度燃料的 SOFC 系统，为开发具有实用化、商业化中远程航程的纯电动或混合电动飞机提供了现实路径，成功突破了当前电池技术的能量密度限制。

最后，多燃料兼容性为航空业规划了一条清晰、平稳且低风险的脱碳过渡路径。航空脱碳涉及复杂的供应链转型和技术迭代，难以一蹴而就。SOFC 系统允许航空公司从使用现有成熟燃料起步（如传统化石燃料或低比例混合 SAF），随着可持续燃料供应链的成熟和绿氢/绿氨成本的下降，逐步过渡到使用 100% SAF 或零碳氢气/氨气。这种燃料来源的可切换性使航空公司、机场和燃料供应商能够根据自身发展节奏和市场条件分阶段调整，无需对飞机或基础设施进行颠覆性改造，从而构建了一条风险可控、投资可持续的清晰路径，是实现航空业长期净零排放目标的战略关键。

3.3 航空 SOFC 系统设计的关键与挑战

将 SOFC 技术成功应用于航空动力系统，面临着航空器对高功重比、快速响应、极端可靠性和燃料适应性的严苛要求。对于执行中远程任务的飞行器而言，其动力系统必须在长航时、高强度巡航状态下保持极致的效率与可靠性，这使下述挑战成为必须攻克的核心工程技术瓶颈。

3.3.1 燃料处理

SOFC 的多燃料适应性是其战略优势，但利用多元化燃料带来的运行灵活性同时，也引入了燃料中特定化学成分引发的性能衰减及系统复杂度的增加。不同燃料因其化学特性差异，对系统耐久性构成了独特挑战。

氨燃料具有含氢量高、零碳含量的优势。直接使用氨燃料会导致运行功率密度较纯氢基准下降 10%～15%。在较低的运行温度下（500～650 ℃），性能衰减可能超过 30%，并伴随着镍基阳极与氨反应生成脆性 Ni_3N 相的风险，严重影响电堆寿命与稳定性[103]。未来需要对应开发抗氮化的高性能新型阳极材料，或优化运行温度窗口以规避有害反应。

烃类燃料（如汽油和柴油）具有来源广泛、能量密度高的优点。然而，直接使用烃类燃料的系统因燃料中存在的芳香烃和烯烃易发生积碳[71]。研究表明，采用 ATR 的柴油 SOFC 系统在运行 40 h 后就因积碳导致燃料利用率由 100% 降低至 90%，重整效率由 65% 降低至 45%[104]。解决方案在于催化剂创新，即设计具有高抗积碳能力的结构化重整催化剂。通过活性组分的选择、载体改性，以及构建特殊的微观形貌，从化学本质上抑制有害副反应，以保证长航时下的耐久性[105]。

对于含硫量高的航空煤油, 还需要对燃料进行深度脱硫预处理。当前可行的机载脱硫技术包括吸附式脱硫与加氢脱硫。其中, 吸附式脱硫虽无需消耗氢气且能在常温常压下实现深度脱硫, 但其硫吸附容量限制了燃料处理通量。即使在最佳工况下, 每克吸附剂仅能处理几毫升燃料[85,106]。加氢脱硫虽能处理大量燃料, 但需持续消耗重整产生的氢气, 直接降低净输出功率。同时, 其必需的高压操作环境(3～7 MPa)与低压重整器及电堆之间存在巨大的压力失配, 增加了系统复杂度与寄生功耗[107]。因此, 技术发展需双线推进。对于规模较小的 SOFC 辅助电源, 对应研发超高硫容的低温吸附材料。而对于更大规模的 SOFC 动力系统, 需要开发新型加氢脱硫工艺, 同时开发与 SOFC 系统更好的压力集成方案。

综上, SOFC 燃料处理的技术发展路径, 核心是面向具体燃料的催化剂与工艺创新, 旨在化解燃料化学复杂性带来的性能与可靠性衰减, 从而真正发挥其多燃料适应的优势。

3.3.2　高温热惯性

航空应用对动力系统的轻量化、快速启动与瞬态响应能力提出了极限要求。传统用于固定式发电的平板式 SOFC 通常升温缓慢, 耗时数十分钟甚至数小时, 难以满足航空器快速启动和应对动态负载变化的需求[108]。这揭示了一个贯穿 SOFC 系统的核心物理矛盾: 高温运行所必须的高热容量, 与航空快速响应要求的低热惯性之间存在根本冲突。破解这一核心矛盾, 需从材料到系统架构实施多层级协同创新。

从材料体系上进行根本革新, 是降低热惯性的终极路径。通过采用质子导体电解质(如掺杂的 $BaZrO_3$ 或共掺杂 $Ba(Ce,Zr)O_3$), 可将 SOFC 的高效工作温度窗口从 800～1 000 °C 显著下移至 350～600 °C[109]。工作温度的降低直接意味着电堆热容的减少。这样既能缩短加热时间, 又能增强其在反复热循环下的耐受性[110], 从而从物理本源上缓解了高温热惯性矛盾。尽管该技术目前仍在攻关其长期化学稳定性与高功率密度输出等挑战, 但其代表了解决 SOFC 动态特性问题的根本方向。

部件级减惯涉及重整器、电堆、热交换器与保温壳体等关键组件, 其核心是通过降低质量与材料热容实现系统热惯性的下降。以重整器为例, 核心是发展微通道反应器。其能够提供极高的比表面积, 并因尺度效应极大增强反应物与催化表面的传质效率。Jin 等的研究表明, 在相同燃料供给速率下, 采用微通道结构的重整器, 其甲烷转化率可比传统反应器提升 71.9%～86.6%[111]。这种反应速率的跃升, 意味着重整器的处理通量密度大幅提高, 使满足相同功率需求所需的反应器体积与质量得以显著减小, 直接带来了热惯性的降低, 为系统实现秒级动态响应奠定了基础。

热管理是主动应对高热惯性的直接手段, 其核心任务是对系统内有限热流进行智能调度, 在启动和运行阶段面对着不同的挑战。

在启动阶段, 需要高效地将电堆、重整器等核心部件从环境温度迅速加热至工作点。为界定航空 SOFC 系统的快速启动时间, 明确的量化目标至关重要。美国能源部(United States Department of Energy, DOE)为 SOFC 辅助动力系统设定的快速启动目标为 30～45 min[112]。为实现该目标, 技术关键在于集成高功率辅助加热器, 直接对电堆和重整器的核心区域进行快速预热[113]。Özece 等的研究即通过将电加热器集成在热交换器下游, 将启动时间由 4 610 s 大幅减少至 3 130 s, 这不但验证了该路径的有效性, 也清晰标定了当前技术水平与目标之间的距离[114]。

进入运行阶段, 热管理的根本目的是精确维持电堆内部的温度分布均匀性, 以避免局部过热或过冷引发的热应力与性能衰减[115]。为达成此目的并大幅提升全系统能量效率, 必须构建深度热集成网络, 将系统内部热量充分应用于重整反应预热、驱动底部循环(如微型 GT[116])及环控除冰等多元场景。该策略的核心价值在于, 通过将“废热”转化为系统内快速循环的“有效热”, 构建了一个高度自洽的内部热循环, 从而显著降低了对独立、缓慢且低效的外部热源的依赖。这不仅提供了关键的热流调配柔性, 更通过为富余热量创造智能去路, 极大地增强了系统应对剧烈热波动的内在稳健性。然而, 这种高效集成也使电堆、动力、环控等子系统深度耦合, 构成了一个多目标、强干扰的复杂控制难题[117]。

为从根本上规避 SOFC 的热惯性短板, 必须在系统架构层面进行创新。核心路径是通过引入高功率电池或超级电容器作为辅助电源, 在 SOFC 启动期间或负载发生瞬态变化时提供瞬时功率支撑, 待 SOFC 达到稳定运行状态后再进行平滑过渡, 从而将 SOFC 本体的动态响应要求从秒级放宽至分钟级[118]。驾驭这种复杂的混合系统需要发展基于模型的智能控制算法, 根据实时飞行状态和燃料特性, 动态优化重整器工况(温度、空速、水碳比)与电堆运行参数, 并在电堆和辅助电源之间进行能量分配。在保障 SOFC 系统平稳运行下, 最大限度满

足航空动力系统的瞬变需求。

综上所述，通过材料本源创新降低工作温度、部件结构革新减少热容、热管理主动调度控制热流，以及混合架构解耦规避动态需求，形成了一套从本源到系统、逐层递进的综合解决方案，以应对 SOFC 航空应用中最核心的高温热惯性矛盾。

3.3.3 可靠性与系统集成挑战

SOFC 动力系统要达到商用航空数万飞行小时与起降循环的可靠性要求，必须攻克由高温和复杂性带来的集成难题。这一挑战主要集中于高温密封与系统集成 2 个相互关联的维度。

高温密封是保障 SOFC 系统在严酷航空环境下长期可靠运行的基础。密封材料必须在 600～1 000 °C 的高温下，长期暴露于氧化性（空气侧）和还原性（燃料侧）气氛中稳定工作，并能承受频繁的热循环考验。这对于需经历数万次起降循环的商用飞机而言，是保证其全生命周期内气密性与安全性的基础，其要求远高于地面固定式发电系统。因此，航空级密封绝非简单的材料选型问题，而是一个涉及热-机械-化学多场耦合的界面科学难题。理想的密封材料必须具备与相邻部件（如连接体、电解质）高度匹配的热膨胀系数（coefficient of thermal expansion, CTE），以缓解循环热应力；在运行温度下需具有合适的黏度与润湿性，以确保封接完整并能弥合微裂纹；同时，必须在复杂气氛中保持长期化学稳定性，避免与相邻材料发生有害反应或自身发生晶相转变[119]。

系统复杂性与高度集成则将可靠性挑战放大至系统层级。与依赖纯氢、系统相对简单的 PEMFC 相比，SOFC 系统包含了燃料处理单元、高温气体管理系统、复杂的排气热回收装置及可能的混合动力子系统（如 SOFC-锂电池或 SOFC-GT）等多个组成部分，其架构明显更为复杂[120]。这种复杂性在航空领域被极致轻量化的要求进一步放大，要求每个子系统都必须在极小的体积与质量约束下实现极高的功能密度与可靠性[121]。这带来了双重工程压力：一方面，子系统数量的增加意味着故障概率的叠加，对系统级冗余设计与故障容错策略提出了前所未有的要求；另一方面，各子系统间强烈的热、质、电耦合使系统成为一个多变量、强干扰的复杂控制对象，任何单一参数的扰动都可能通过耦合链被放大，威胁全局稳定。驾驭这种复杂性，实现从功能集成到稳健集成的跨越，是 SOFC 航空动力系统实用化的核心考验。

综上分析，SOFC 凭借其高温运行所实现的高效能量直接转化与卓越的多燃料适应性，为航空动力系统的深度脱碳提供了一条革命性的路径。其核心价值在于能够突破传统电池的能量密度瓶颈，并为利用现有航空燃料基础设施向 SAF 乃至零碳燃料（如氢、氨）过渡提供了技术兼容性。然而，将其成功集成于航空平台，必须直面并攻克一系列严峻工程挑战：在燃料处理方面，需解决特定燃料引发的性能衰减与重整催化剂失活问题，实现轻量化、高可靠性的机载燃料预处理与重整；在系统动态特性方面，必须破解高温运行所必需的高热容量与航空应用要求的低热惯性、快速响应之间的根本矛盾，这需要通过材料体系革新、部件级结构优化、主动式热管理策略及混合动力架构等多层级协同创新来系统解决；在长期可靠性方面，则需攻克能耐受数万次热循环与复杂气氛的高温密封这一界面科学难题，并驾驭由燃料处理、电堆、热回收及混合子系统高度集成所带来的多变量、强耦合的复杂系统控制与鲁棒性挑战。尽管前路艰巨，但通过在新材料、核心部件与系统集成层面的持续突破，SOFC 技术有望在未来中远程绿色航空的动力谱系中扮演关键角色，展现出重要的战略价值。

4 燃料电池与锂电池航空动力系统的技术范式对比

航空动力系统的技术演进正处在多元化探索的十字路口。燃料电池与锂电池作为两条核心的电动化技术路径，并非简单的竞争替代关系，而是基于其物理化学本质，演化出了截然不同的技术范式与应用边界。为系统呈现不同动力技术的应用边界，本节首先在图 11 中综合构建了一个基于任务航程需求与系统功率动态需求的二维适配性图谱。该图谱是对后续 4.1 节与 4.2 节核心结论的可

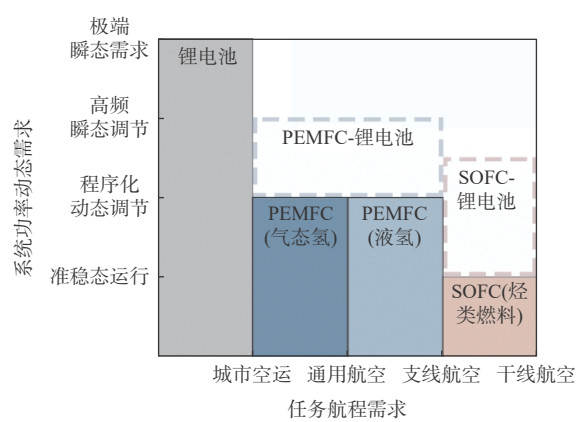


图 11 燃料电池与锂电池航空动力系统的技术范式谱系及其任务适配性分析

Fig. 11 Spectrum of technological paradigms and mission adaptability analysis for fuel cell and lithium battery aviation powertrains

视化总结, 其横轴映射了 4.1 节所剖析的能量载体范式如何根本性地决定航程潜力, 而纵轴则对应了 4.2 节将揭示的功率架构与动态响应极限如何制约技术对任务动态性的适应能力。进而, 衰减机理(4.3 节)与安全环境特性(4.4 节)将作为关键的约束条件, 深入探讨各技术在图示的适配区域内是否具有实现商业化所必需的耐久性、可靠性及合规性, 从而完整刻画从技术匹配到应用落地的全要素图景。

4.1　能量载体范式与航程特性

航空动力系统设计首要与核心的决策在于对能量载体的选择, 这直接导致了封闭式与开放式 2 种系统范式的分野, 并从根本上锚定了飞行器的航程能力、载荷裕度及运营模式, 衍生出截然不同的技术发展路径。

锂电池系统本质上是一个封闭的储能系统。其能量存储于电极材料的晶格结构之中, 通过锂离子的嵌入和脱出反应实现电能的储存与释放。其能量密度的理论上限由正负极材料的理论比容量和工作电压窗口严格限定, 当前锂离子电池的工程可用能量密度已逼近 250～300 Wh/kg 的阶段性瓶颈[122-124]。任何能量密度的提升都需在材料体系上进行革新, 且不可避免地与循环寿命、充放电倍率性能及热安全性等关键指标相互权衡, 面临巨大的技术挑战。这种物理本质决定了其航程与电池质量呈近似线性但边际效益递减的关系。增加电池质量虽能延长航程, 但也会同步增加飞行器的结构质量和飞行中的寄生能耗, 导致航程的净收益递减。因此, 锂电池动力系统的设计哲学是在有限的能量密度边界内进行极致的优化, 通过提升电芯能量密度、优化 pack 集成效率、减轻结构质量等手段挖掘每一分潜力, 这使其天然更适配于对质量高度敏感、航程需求相对有限的短途高频飞行任务, 如城市空中交通或轻型飞机。

反观燃料电池系统, 其核心是一个开放的能量转换装置[125]。其将存储于外部燃料中的化学能通过电化学反应直接转化为电能, 其系统总能量密度理论上限取决于燃料本身的质量能量密度及燃料储存系统的质量与体积效率。燃料电池电堆及其辅助子系统可以视为一个相对固定的基础质量, 而航程则与所携带的燃料质量成正比。系统总质量为基础质量与携带燃料质量之和。航程虽可以通过增加燃料携带量来延长, 但其收益并非线性, 因此, 定义一个航程盈亏平衡点至关重要。该点是指, 所增加燃料带来的航程收益足以抵消因其自身质量引起的额外能耗与间接结构增重。一旦设计越过此临界点, 增加燃料带来的航程收益将变得极为显著。这种“基础质量 + 可变燃料”的模型, 使其在追求长航时的应用场景中展现出无可比拟的潜力。此外, 其在补能方式上采用“加注”而非“充电”, 能够在分钟级别完成燃料补充, 极大地贴近了传统航空的运营节律, 对于提升飞机日利用率和高繁忙度机场的周转效率具有革命性意义。尽管目前储氢技术仍面临质量、体积或蒸发损失等挑战, 但其技术发展路径清晰, 旨在不断降低“基础质量”并提升储氢效率, 从而持续释放其对长航程、大载重任务的支撑能力。

综上, 能量载体的本质差异决定了两者不同的航程逻辑与发展范式。锂电池在封闭系统内“精耕细作”, 使其天然适配于对质量高度敏感、航程需求相对有限的短途高频飞行任务, 如城市空中交通。而燃料电池则通过开放系统寻求突破, 为其在支线及干线航空领域支撑长航程、大载重任务奠定了基础。这一根本分野为它们在航空动力谱系中处于各自不可替代的战略定位, 如图 11 所示。

4.2　功率特性与混合动力系统架构

飞行器在起降、爬升、巡航及突发机动等多变的任务剖面中, 对其动力系统的功率响应特性提出了多维度的严苛要求。燃料电池与锂电池因能量转换机理的根本性差异, 在该维度上展现出迥异的动态响应能力, 并直接塑造了截然不同的系统架构演化路径。

锂电池的功率输出基于固相离子嵌入/脱出反应, 具有高功率密度(最大放电倍率下超过 3 kW/kg, 单电池级)与毫秒级响应速度的显著优势, 能近乎瞬时地满足飞行器的瞬态大功率需求[126-128]。然而, 高倍率放电会诱发电池内部急剧的欧姆热与反应热积累, 这不仅加速电极活性材料衰减、固态电解质界面(solid electrolyte interphase, SEI)膜持续增厚, 更潜藏着热失控的安全风险。因此, 锂电池动力系统的设计核心在于实施精细而保守的功率管理策略, 通过精确规划功率剖面、严格限制峰值功率持续时间, 在性能、寿命与安全性之间寻求最佳权衡。

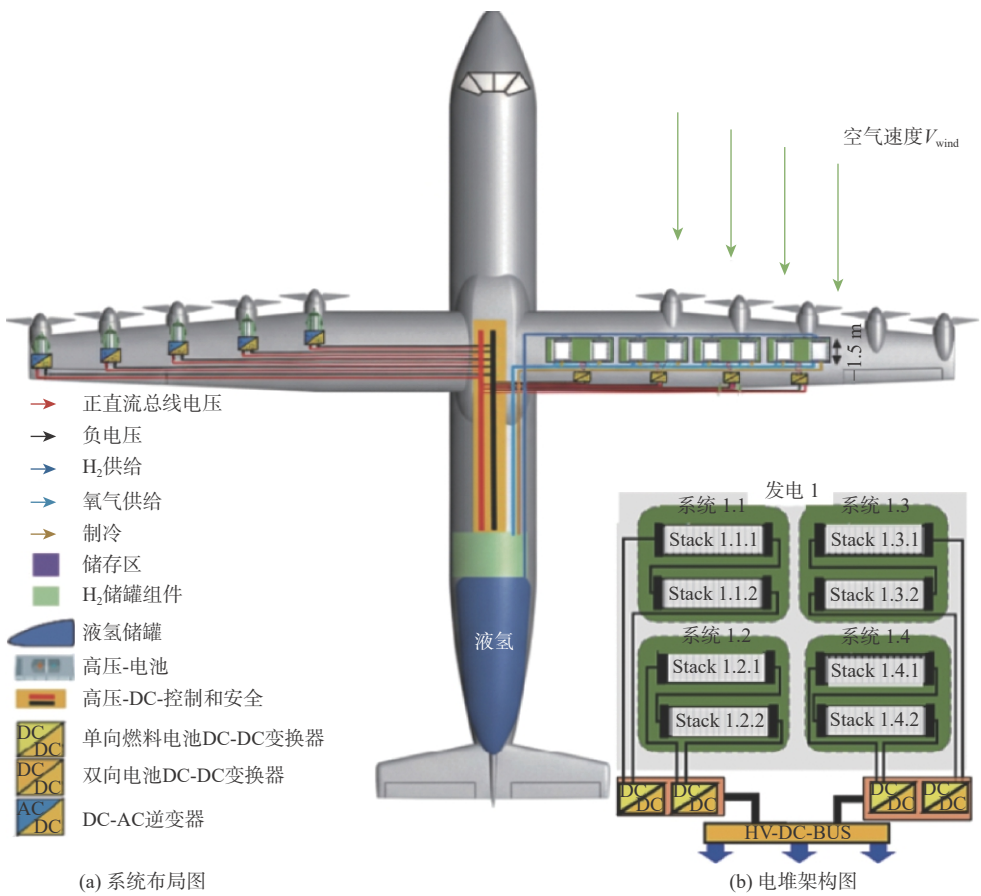
PEMFC 的动态响应主要受限于其空气供应系统的机械惯性(如空压机转速调节)与电化学反应动力学, 其功率响应通常存在秒级延迟(约 1～10 s)[50]。这种“慢速”的特性使其难以独立应对频繁变化的负载需求。

SOFC 面临的动态响应挑战源于其工作原理。高达 600～1 000 ℃ 的运行温度带来了巨大的热惯性, 使电堆的温度变化极为缓慢。实验研究表明,

即使对于功率仅为100 W的SOFC短堆，在经历从15～25 A的中等幅度电流阶跃后，其输出电压也需要5～25 s才能重新稳定，而即使在最优变载模式下，其内部温度场达到新的稳态也需要约300 s[129]。对于功率更大，热耦合更复杂的系统，这种惯性效应将进一步放大。针对5 kW级SOFC系统的模拟显示，集成热管理后，由电流或燃料扰动引发的热平衡恢复时间将延长至300～800 s[130]。由此可以合理推断，当航空推进用电堆的功率规模扩大至数百千瓦乃至兆瓦级（跨数个数量级）时，其质量、热容量及内部传热路径将显著增加，热惯性可能呈数量级增加，功率调节过程可能需要数分钟至数十分钟。任何超过热力学限制的快速功率调节都会产生巨大的热应力，导致陶瓷电解质、电极及连接体等关键脆性部件因CTE不匹配而开裂、剥落，造成不可逆的机械损伤与性能永久性衰减[131-133]。

然而，将PEMFC与SOFC的上述动态响应特性置于航空推进系统的实际应用场景中评估，其与真实需求之间存在显著差距。中国民用航空局颁布的《航空发动机适航规定》要求航空发动机必须能在不超过5 s的时间内，完成从低功率状态到95%额定起飞推力的瞬变[134]。该条款为航空推进系统的瞬态响应能力设定了必须满足的法律与技术门槛。与此强制性标准进行对比，PEMFC的秒级延迟已接近其极限，而SOFC的数分钟至数十分钟响应时间与标准要求存在数量级差距。因此，单一燃料电池系统在动态响应上均无法独立满足航空适航的严苛要求。这种源自合规性的约束决定了引入高功率密度、毫秒级响应的锂电池构成“燃料电池-锂电池”系统，由锂电池作为功率缓冲单元，提供起飞、爬升及机动所需的瞬态高功率，是满足航空推进瞬态性能需求的必然且唯一可行的技术路径。

图12展示了一种混合动力飞机燃料电池动力系统集成方案。对于PEMFC系统，通常采用功率混合型架构。在此架构中，燃料电池作为高效、稳定的主能量源，而锂电池则作为功率缓冲单元。其能量管理策略（energy management strategy, EMS）的核心在于功率分配，根据预设规则平滑燃料电池的功率输出，使其工作在高效率区间，由锂电池来削峰填谷。这种架构的本质是功率协同与性能优化。尽管PEMFC的秒级动态响应与锂电池的毫秒级响应存在差距，但两者处于同一数量级的时间尺度，在航空推进5 s瞬变要求的框架内具备协同



(a) 系统布局图　　(b) 电堆架构图

图12　混合动力飞机集成的燃料电池动力系统[135]

Fig. 12　Fuel-cell powertrain integrated in a hybrid-electric aircraft[135]

工作的基础。因此, 其设计哲学的核心是在满足动态需求的前提下, 追求系统整体的效率最大化、寿命延长与运行经济性(氢耗最低)。这决定了其 EMS 的复杂性主要集中于实时、多目标的功率分配优化算法, 如等效氢耗最小化策略、模型预测控制等, 旨在使 PEMFC 尽可能规避有害的动态负载波动, 从而优化其工作点[136]。总而言之, PEMFC 混合系统的设计是一种在既定性能边界内寻求经济性与可靠性最佳平衡的优化范式。

对于 SOFC 系统, 则需要采用更深层次的能量-功率结构型混动架构, 其控制逻辑的复杂性与设计哲学同 PEMFC 系统存在根本差异。SOFC 的动态调节过程与飞行器毫秒级的电功率需求之间存在巨大的时间尺度差异。这种物理根源决定了其混合架构的最高设计准则不是优化, 而是功能使能与安全隔离。SOFC 的角色被严格限定在恒定或极其缓慢变化的工况下提供背景基载功率, 而锂电池则必须升级为应对所有动态负载的主功率执行器。因此, 其 EMS 的核心目标是不惜一切代价维持电堆的热稳定性, 避免任何可能引发灾难性热机械应力的快速功率变化。这导致其控制逻辑上升为多时间尺度的、强制的解耦与预测控制。必须基于长期(如任务剖面)的功率预测, 为 SOFC 规划一条极其缓慢、平滑的功率设定曲线, 而将所有瞬时波动完全交由锂电池管理[137]。因此, SOFC 混合系统的设计是一种由物理安全边界强制的、以电堆安全为绝对前提的使能范式。

由此可见, PEMFC 与 SOFC 混合架构的分野, 远非简单的动态响应性能之别, 而是从物理约束(时间尺度)、首要设计目标(优化/安全), 到系统角色定义(协同/解耦)及控制哲学(实时分配/预测规划)的、贯穿技术范式各层次的本质性差异。这种差异, 是两者基础电化学原理与材料特性作用于复杂工程系统的必然结果, 也直接决定了两者在图 11 所示谱系的纵轴上具有不同的拓展极限。

4.3　衰减机理与寿命管理

动力系统的耐久性是决定其全生命周期经济性与可靠性的核心指标。锂电池与燃料电池的寿命衰减源于其内部材料体系在工作环境下发生的不可逆物理化学变化, 但两者的衰减机理、主导因素及可管理性存在根本差异。

锂电池的衰减是一个持续且累积性的不可逆过程, 主要表现为循环老化与日历老化。循环老化源于高倍率充放电、深放电及工作温度不当所加剧的副反应, 包括由于负极 SEI 膜的持续增厚, 不断消耗活性锂离子与电解液, 导致的容量持续衰减[138]; 正极活性材料的结构相变、晶格坍塌及金属离子溶出[139]; 以及在极端情况下锂枝晶的析出, 可能刺穿隔膜引发内短路[140]。日历老化即使在搁置状态下也会发生, 与环境温度、电池荷电状态(state of charge, SOC)密切相关, 主要表现为电解液的氧化分解与界面稳定性的缓慢退化[141]。电池管理系统(battery management system, BMS)的策略核心在于通过优化充放电区间、控制温度、限制倍率来延缓这些老化进程, 但无法逆转已发生的衰减。

PEMFC 的衰减机理复杂且与运行工况紧密耦合, 但其衰减在一定程度上是可管理与可部分缓解的。其主要衰减模式包括催化剂衰减[142]、碳载体腐蚀[143]及质子交换膜退化[144]。催化剂衰减表现为铂纳米颗粒的溶解、迁移和奥斯特瓦尔德熟化过程, 导致电化学活性面积持续减小[145]。碳载体腐蚀主要发生在高电位和启停工况下, 阴极碳载体发生电化学氧化, 造成催化剂流失和孔隙结构坍塌[146]。质子交换膜的退化则包括化学降解和机械损伤两方面, 前者源于自由基攻击导致的聚合物链断裂, 后者由于湿循环引发的溶胀-收缩应力所致。

SOFC 的衰减主要由高温环境下材料的持续退化所驱动, 这构成了其长期耐久性的核心挑战。在长期高温运行过程中, 阳极镍催化剂颗粒会通过表面扩散和奥斯特瓦尔德熟化过程不断粗化长大, 导致三相界面长度减少, 阳极活化极化显著增大, 从而降低电池输出性能[147]。当系统采用含铬合金作为连接体材料时, 挥发性六价铬物质会在阴极沉积, 对氧还原反应的活性位点产生毒化作用, 进一步加剧性能衰减[148]。同时, 密封材料在长期高温服役条件下会发生蠕变和脆化, 加之其与连接体、电解质之间 CTE 的不匹配, 容易引发密封失效和气体交叉问题[149]。更严重的是, 频繁的热循环或不均匀的温度分布会产生显著的热机械应力, 可能导致陶瓷电解质开裂、电极涂层剥落等不可逆的结构损伤。值得注意的是, SOFC 的寿命表现高度依赖于维持稳定的高温运行状态及实施极其缓慢、可控的负荷变化, 以最大限度地降低热应力带来的负面影响。与 PEMFC 不同, SOFC 的衰减过程通常难以通过在线手段进行恢复, 其寿命提升的关键路径在于开发新一代中温材料体系、具有抗铬毒特性的阴极材料, 以及展现更强热机械鲁棒性的电池结构设计。

综上所述, 三者在寿命维度呈现出不同的图景。锂电池的衰减是持续且不可逆的, 管理策略的核心在于延缓; PEMFC 的衰减是可干预的, 通过系统控制策略和维护手段有望部分恢复; 而 SOFC 的

寿命则与热历史强绑定，其衰减是累积且难以逆转的，设计与管理的第一要务是维持温度稳定，避免一切可能带来机械损伤的剧烈操作。这种衰减机理的本质差异，从根本上决定了其适航验证策略与维护周期的制定，进而影响其在高频次短途航线与低频次长航程航线中的经济性与适用性。

4.4 安全与环境特性

航空动力系统的安全性设计与环境效益评估，需建立在对不同技术路线风险本源与全生命周期影响的深刻理解之上。燃料电池与锂电池在此维度上呈现出截然不同的风险图谱与环境悖论，直接决定了其适航路径与可持续性认证的焦点。

在安全性方面，2种技术的风险本源截然不同。锂电池的风险本质源于内短路引发的链式热失控[150]。电池内部隔膜局部失效导致正负极直接接触，引发剧烈放热反应，进而点燃有机电解液并导致相邻电芯发生热蔓延。这一过程具有突发性、快速蔓延性及难以扑救的特点，其机制可以通过图13所示的链式反应来解释。其防护体系采用深度防御策略，依赖高精度BMS实现早期故障预警，并通过电芯级热屏障、模组级隔热及系统级高效冷却等多重措施物理阻断热扩散。

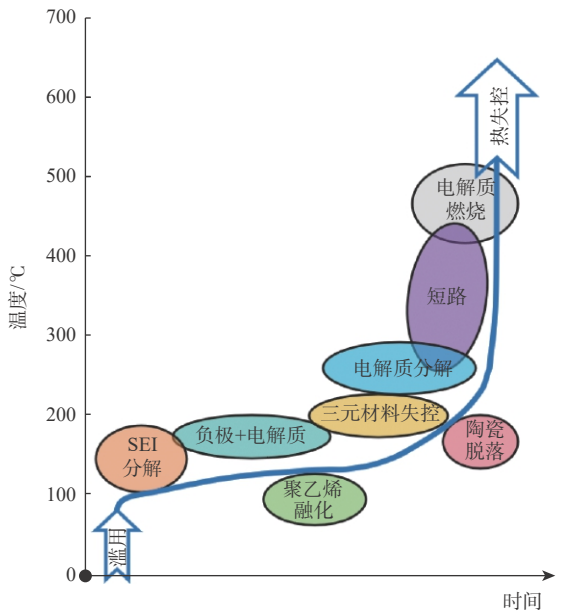


图13 热-温度-反应回路

Fig. 13 Heat-temperature-reaction loop

与之相比，燃料电池本体的电化学反应完全可控，其核心风险并非来自反应本身。对于采用纯氢燃料的PEMFC，风险集中于氢气泄漏后的燃烧与爆炸。氢气具有极宽的爆炸极限范围、较低的点燃能量和极强的扩散能力，其安全性高度依赖于“故障-安全”设计理念，即采用高度集成、冗余的氢气管路系统，配备多级快速切断阀、基于氢敏传感器的实时泄漏监测系统、碰撞感知与自动泄压装置，并通过精心设计的气体流场避免氢气的局部聚集。其安全设计旨在确保即使发生单一部件失效，也能将氢气泄漏的风险降至最低。而SOFC面临的首要风险来自600～1 000 ℃高温运行环境带来的热危害，包括高温部件可能造成的灼伤风险及特殊工况下的热失控可能。此外，当使用氨、甲醇等替代燃料时，还需要考虑这些燃料特有的毒性和腐蚀性风险，要求系统具备相应的防护措施。

在环境表现方面，必须采用“Well-to-Wake”全生命周期视角进行客观评估。锂电池在使用阶段是真正的零排放，但其碳足迹和环境代价大幅转移至上游的电池材料开采、冶炼、生产制造环节及下游的报废回收阶段。正极材料(如镍、钴、锂)的开采具有高能耗、高水耗和生态破坏性，电池生产过程也属于高碳排环节。其环保优势的兑现，最终依赖于绿色电网和闭环回收技术的成熟。

首先，燃料电池的环境正当性则完全依赖于燃料的来源。当使用由可再生能源电解水产生的绿氢时，其全生命周期可实现近零碳排放，展现出巨大的环境优势。然而，若使用当前主流化石能源重整制得的灰氢或蓝氢为燃料，其全生命周期的碳足迹可能甚至高于直接使用航空煤油的传统发动机。对于可使用SAF或生物甲醇的SOFC系统，其碳中和程度则取决于这些燃料的原料来源与生产工艺。因此，燃料电池技术的环境效益并非其固有属性，而是一个与整个国家或地区的能源结构和氢能战略紧密绑定的外部变量。图14展示了当前阶段氢气从生产到利用的全链条。

综上所述，安全与环境并非静态的优劣对比，而是一个动态的系统工程。安全性要求根据不同的风险本源，设计截然不同的防护体系。环境效益则要求超越使用阶段的局部排放，从能源系统的全局视角进行审视。这2种技术的真正清洁与安全未来共同依赖于绿色能源基础设施的完善、高标准全生命周期碳核算体系的建立及闭环材料循环经济的实现。

本节从能量载体与功率特性2个根本范式出发，系统阐释了燃料电池与锂电池在航程与动态响应能力上的本质分野，并在图11呈现的任务适配性图谱中对其互补的战略定位进行了可视化总结。然而，一项动力技术要从理论匹配走向商业应用，还必须通过衰减机理与安全环境特性这两大可行性维度的严格评估。前者决定了系统在长期运行中的经济性与可靠性边界，后者则关乎其全生命周期的环境合规性与社会接受度。因此，锂电池、

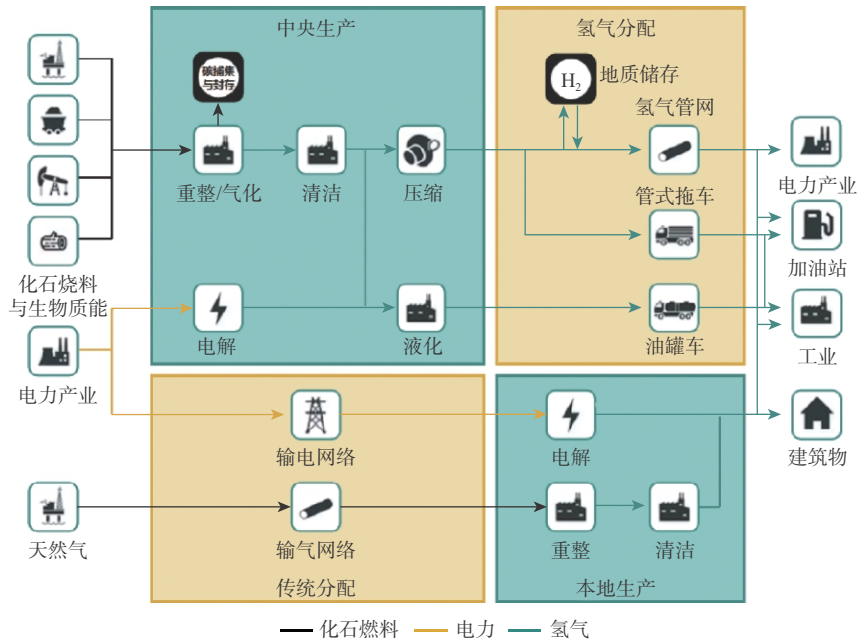


图 14　氢气全产业链示意图

Fig. 14　Schematic of the full hydrogen value chain

燃料电池及其混合架构在绿色航空未来的角色是由核心性能边界、长期运行耐久性及全周期可持续性这三大支柱共同定义的。它们在此多维评估框架下, 构成了互补共生的技术谱系, 为航空业的深度脱碳提供了多元且坚实的解决方案。

5　燃料电池在航空领域的应用案例与发展路径

在上文对 PEMFC 与 SOFC 的技术原理、系统架构及挑战进行深入剖析, 并从技术范式角度将其与锂电池系统进行全面对比的基础上, 燃料电池航空动力系统的研究已逐步从理论探索走向工程实践。当前, 全球多个创新项目通过原型机试飞与示范运行, 验证了燃料电池在不同航空平台中的应用潜力, 同时也揭示了从实验室走向商业化所需跨越的鸿沟。本节将结合代表性案例, 分析燃料电池在航空领域的实际应用成效, 并据此展望其未来发展的可能路径与关键挑战。

5.1　典型案例剖析

燃料电池在航空领域的应用呈现出多样化的技术路线和平台适配性。以下案例分别体现了 PEMFC 和 SOFC 系统的当前发展水平及其对应的应用场景, 如图 15 所示。

ZeroAvia 公司的 HyFlyer II 项目代表了 PEMFC 技术向支线航空领域拓展的积极尝试。如图 15(a) 所示, 该公司开发了 ZA600 氢电推进系统, 基于 19 座道尼尔 228 飞机平台进行飞行测试, 将 ZA600 安装在左翼替代原有的霍尼韦尔 TPE-331 涡桨发



(a) HyFlyer II原型机　(b) HY4验证机

(c) Prodrone无人机　(d) Stalker VXE无人机

图 15　燃料电池航空示范应用

Fig. 15　Demonstration of fuel cell applications in aviation

动机。该混合动力系统集成了额定功率为 400 kW 的 PEMFC 电堆(由 PowerCell 提供)和锂电池组, 通过电机提供峰值功率为 660 kW 的动力输出[151]。这一设计实践了 4.2 节所述的功率混合型架构, 其中, 燃料电池作为高效巡航的基载电源, 锂电池则负责满足高功率需求阶段(如起飞和爬升)的能量补充, 有效弥补了 PEMFC 动态响应相对滞后的不足。该项目目前实现了约 556 km 的设计航程, 标志着 PEMFC 系统正向更大载重、更远航程的航空平台迈进。然而, 其所采用的高压气态储氢方案也面临着能量密度限制的挑战, 这正是当前行业积极探索液氢储运技术的直接原因。

在液氢储运技术验证方面, H2FLY 公司的 HY4 四座验证机取得了突破性进展。如图 15(b) 所示, 该机采用独特的双机身设计, 中央动力舱集成了 80 kW 的电动机。项目最初采用 35 MPa 高压气态

储氢系统，储氢量为9 kg，航程为750 km。后续升级为液氢储罐系统后，储氢量大幅提升至24 kg，航程相应扩展至1 500 km[152]。这一成果不仅证明了深冷液氢技术在航空应用中的可行性，也彰显了其在提升航程方面的显著优势。液氢极高的质量能量密度为破解航空器储氢难题提供了切实可行的解决方案，但同时也带来了绝热、液化耗能高和蒸发损失控制等新的技术挑战，这些问题的解决将成为决定液氢技术在航空领域大规模应用的关键。

日本Prodrone公司开发的液化石油气燃料无人机展现了SOFC多燃料适应性的独特优势。如图15(c)所示，这款工业级无人机集成了新型平板式SOFC电堆，能够在电极内部直接将液化石油气转化为合成气进行发电，省去了外部重整器，显著提高了系统功率密度。采用SOFC-锂电池混合架构后，系统实现了超过1 h的连续飞行能力，且可通过更换市售液化石油气罐在1 min内完成燃料补充[153]。这一创新设计充分体现了SOFC燃料灵活性的工程价值，使其能够摆脱对氢气基础设施的依赖，在工业巡检、长航时监视等特殊应用场景中展现出独特的实用性。同时，混合动力架构的应用也是对SOFC动态响应较慢这一固有问题的有效解决方案，为SOFC系统在航空领域的应用提供了重要参考。

在长航时军事应用领域，洛克希德·马丁公司开发的Stalker VXE无人机将SOFC系统的性能发挥到了新的高度。该无人机采用SOFC-锂电池混合动力系统，以丙烷为燃料，在搭载2.5 kg任务载荷时续航时间可达8 h[154]。更令人瞩目的是，其原型机通过外置燃料箱扩容后，如图15(d)所示，实现了连续飞行超过39 h的卓越表现，创下了5～25 kg级无人机续航时间的世界纪录[155]。这一成就不仅验证了SOFC系统在高能量转换效率方面的优势，也展示了其在军事侦察、持久监视等特殊场景下的应用价值。该项目的成功依赖于先进的热管理、系统集成和燃料处理技术，特别是在应对高温密封、热循环耐受性等挑战方面取得了重要进展，代表了当前SOFC系统集成技术的最高水平。

5.2 发展路径与未来挑战

当前原型机与无人机项目的成功，为燃料电池航空动力系统勾勒出了一条从特定场景验证逐步走向大规模商业化的发展路径。近期，技术应用将聚焦于小型无人机与通勤飞机市场，依托PEMFC的快速响应与SOFC的燃料灵活性，在工业巡检、区域物流等细分领域率先实现商业化突破。随着液氢储运等关键技术的成熟，中期目标将瞄准支线航空领域，推动系统向更大功率、更高集成度发展。而远期愿景则是进军单通道干线客机市场，这无疑依赖于革命性新材料和新架构的突破，最终实现航空动力的深度脱碳[156]。

然而，实现这一蓝图绝非易事，诸多挑战横亘于前。首当其冲的是技术成熟度问题，将实验室的高性能转化为满足航空严苛要求的可靠性，仍需跨越巨大鸿沟。极致提升系统的功重比与功率密度，需要对整个动力系统进行前所未有的轻量化和集成化设计，其难度远超其他应用场景。同时，航空器数万小时的耐久性与可靠性要求，对燃料电池系统，特别是SOFC的抗热震疲劳性能、启停循环寿命及长期运行下的材料衰减等方面提出了极其苛刻的考验。此外，应对飞行中复杂多变工况的高智能控制策略，也是决定系统能否稳定高效运行的关键。

超越技术本身，绿色氢能供应链的缺失构成了另一个基础性挑战。燃料电池的环境效益本质上取决于燃料来源，而当前绿氢的产能、成本及储运技术远未满足航空规模应用的需求。这导致了一个典型的发展困境：航空业等待氢能基础设施的完善，而能源领域则期待航空应用带来明确的需求信号。破解这一“鸡与蛋”的困局，需要跨行业的深度协同与顶层设计。

最后，所有这些挑战都汇聚于商业化的经济性门槛。在产业发展初期，高昂的研发与制造成本使燃料电池飞机难以与传统机型竞争。其商业化进程在相当长时间内仍需依赖碳定价机制和政策支持来弥补成本劣势，直到技术突破和规模化效应带来成本的实质性下降。这条发展路径上的每个环节都充满挑战，但也正是这些挑战，定义了未来研究的方向与价值所在。

6 结束语

燃料电池航空动力系统作为航空业迈向碳中和的重要技术路径，展现出广阔的应用前景和独特的技术价值。本文围绕PEMFC与SOFC两类主流技术，系统分析了其工作原理、系统架构、关键挑战及适用场景，并从能量载体、功率响应、寿命机制、安全与环境等多维度，与锂电池推进系统进行了对比。研究表明，锂电池、PEMFC和SOFC在任务谱系上呈现出明确的互补协同关系：锂电池凭借其高功率密度与快速响应能力，在城市空中交通领域占据主导；PEMFC则凭借其快速启动、较高功率密度和技术成熟度，在通用航空及短途支线领域展现出强大的应用潜力；而SOFC依托其超高效率与

卓越的多燃料适应性, 成为未来中远程支线及干线航空最具竞争力的技术方案。

展望未来, 燃料电池航空动力系统仍处于从实验室走向工程应用的关键阶段。下一步的发展应持续聚焦于高温材料、抗毒催化剂、轻量化储氢、智能热管理、高效空气系统等核心技术的突破, 不断提升系统的功重比、可靠性和环境适应性。具体而言, PEMFC 技术需重点攻克适用于小型飞机的低成本、轻量化、高耐久性系统集成; 而 SOFC 技术则需着力解决适用于大型飞机的功率变化与燃料重整等核心瓶颈。

在系统层面, 混合架构仍是重要发展方向。通过将燃料电池与高功率储能单元有机结合, 可有效弥补燃料电池在动态响应方面的不足, 实现功率与能量的优化分配, 提升系统整体性能。同时, 推动电堆-重整器-热管理-功率调节等多子系统的高度集成与智能协同, 也是实现航空级动力系统轻量化、高可靠性的必由之路。

值得强调的是, 燃料电池的环境效益高度依赖于氢源和其他燃料的全生命周期绿色化。未来需依托可再生能源制备绿氢、绿氨或 SAF, 构建“能源-燃料-动力”一体化的零碳航空能源体系。这不仅需要技术突破, 更依赖政策支持、产业链协同和跨学科合作, 共同推动标准建立、场景验证和示范运行。

总之, 燃料电池航空动力系统正处于技术快速演进与应用场景不断拓展的历史机遇期。尽管在储氢、耐久性、系统控制等方面仍面临诸多挑战, 但其在实现长航时、低排放飞行方面的潜力已得到初步验证。通过持续的技术创新、系统优化与产研协同, 锂电池、PEMFC 与 SOFC 将共同构成覆盖从城市空运到干线飞行的完整绿色航空动力谱系, 为实现航空业深度脱碳提供坚实且多元化的技术支撑。

参考文献（References）

[1] Kroyan Y, Wojcieszyk M, Kaario O, et al. Modeling the impact of sustainable aviation fuel properties on end-use performance and emissions in aircraft jet engines[J]. Energy, 2022, 255: 124470.

[2] International Civil Aviation Organization. Carbon offsetting and reduction scheme for international aviation (CORSIA)[EB/OL]. (2025-05-15)[2025-09-19]. https://www.icao.int/CORSIA.

[3] 张毅伟, 薛健敏. 新能源飞机动力系统的发展现状与趋势[J]. 军民两用技术与产品, 2025(4): 22-27.
Zhang Y W, Xue J M. Development of power systems for new energy aircraft[J]. Dual Use Technologies & Products, 2025(4): 22-27 (in Chinese).

[4] 国旭涛, 蔡洁聪, 韩高岩, 等. 分布式能源技术与发展现状[J]. 分布式能源, 2019, 4(1): 52-59.
Guo X T, Cai J C, Han G Y, et al. Technologies and development status for distributed energy resources[J]. Distributed Energy, 2019, 4(1): 52-59(in Chinese).

[5] 童鑫, 熊哲, 高新宇, 等. 质子交换膜燃料电池研究现状及发展[J]. 硅酸盐通报, 2022, 41(9): 3243-3258.
Tong X, Xiong Z, Gao X Y, et al. Research status and development of proton exchange membrane fuel cell[J]. Bulletin of the Chinese Ceramic Society, 2022, 41(9): 3243-3258(in Chinese).

[6] 陈烁烁. 固体氧化物燃料电池产业的发展现状及展望[J]. 陶瓷学报, 2020, 41(5): 627-632.
Chen S S. Development and prospects of solid oxide fuel cell industry[J]. Journal of Ceramics, 2020, 41(5): 627-632(in Chinese).

[7] Chan S H, Ho H K, Tian Y. Modelling of simple hybrid solid oxide fuel cell and gas turbine power plant[J]. Journal of Power Sources, 2002, 109(1): 111-120.

[8] Sugano K, Mair S, Ganti-Agrawal S, et al. Sodium-air fuel cell for high energy density and low-cost electric power[J]. Joule, 2025, 9(6): 101962.

[9] 秦江, 姬志行, 郭发福, 等. 航空用燃料电池及混合电推进系统发展综述[J]. 推进技术, 2022, 43(7): 1-18.
Qin J, Ji Z X, Guo F F, et al. Review of aviation fuel cell and hybrid electric propulsion systems[J]. Journal of Propulsion Technology, 2022, 43(7): 1-18(in Chinese).

[10] 王梦琦. 质子交换膜燃料电池的研究进展与前景微探[J]. 当代化工研究, 2020(12): 8-10.
Wang M Q. Research progress and prospect of PEMFC[J]. Modern Chemical Research, 2020(12): 8-10(in Chinese).

[11] 陈汝妍, 高九涛, 高圆, 等. 管式固体氧化物燃料电池技术进展[J]. 硅酸盐学报, 2024, 52(6): 1900-1916.
Chen R Y, Gao J T, Gao Y, et al. Development of tubular solid oxide fuel cells technology[J]. Journal of the Chinese Ceramic Society, 2024, 52(6): 1900-1916(in Chinese).

[12] Chen X B, Zhong J, Wei J Z. Research on modeling and simulation of fuel cell bus power system based on forward-backward energy flow balance method[J]. Science Progress, 2022, 105: 00368504211064476.

[13] Zhang J M, Yan F W, Du C Q, et al. A multi-feature fusion method for life prediction of automotive proton exchange membrane fuel cell based on TCN-GRU[J]. Materials, 2024, 17(19): 4713.

[14] Wang F K, Han S R, Zhang Y L, et al. Constructing rapid water vapor transport channels within mixed matrix membranes based on two-dimensional mesoporous nanosheets[J]. Communications Chemistry, 2022, 5: 65.

[15] Li C J, Cheng K L, Ma S S, et al. Performance analysis of solid oxide fuel cell/piston engine hybrid system for aviation[J]. Applied Thermal Engineering, 2022, 214: 118797.

[16] Wade F, Roche R, Chailan A, et al. Optimal sizing and energy management of an integrated energy system coupling a hydrogen-fueled gas turbine with storage for power-to-power and hydrogen supply[J]. International Journal of Hydrogen Energy, 2025, 135: 31-47.

[17] Mancino A N, Menale C, Vellucci F, et al. PEM fuel cell applications in road transport[J]. Energies, 2023, 16(17): 6129.

[18] Arjmandi M, Fattahi M, Motevassel M, et al. Evaluating algorithms of decision tree, support vector machine and regression

for anode side catalyst data in proton exchange membrane water electrolysis[J]. Scientific Reports, 2023, 13: 20309.

[19] 马智文, 曾怡达, 李伦. 大功率 PEMFC 空气供给系统建模与实验验证[J]. 化工学报, 2016, 5(5): 2109-2116.
Ma Z W, Zeng Y D, Li L. Modeling and experimental verification of air supply system in large power PEMFC[J]. CIESC Journal, 2016, 5(5): 2109-2116(in Chinese).

[20] Fang R R, Li Z K, Zhang X H, et al. Spreading and drying dynamics of water drop on hot surface of superwicking Ti-6Al-4V alloy material fabricated by femtosecond laser[J]. Nanomaterials, 2021, 11(4): 899.

[21] Yakubu A U, Zhao J H, Jiang Q, et al. A comprehensive review of primary cooling techniques and thermal management strategies for polymer electrolyte membrane fuel cells PEMFC[J]. Heliyon, 2024, 10(19): e38556.

[22] Zahedi R, Forootan M M, Ahmadi R, et al. Exergy-economic assessment of a hybrid power, cooling and heating generation system based on SOFC[J]. Heliyon, 2023, 9(5): e16164.

[23] Wang H, Yang G G, Li S A, et al. Pore-scale modeling of liquid water transport in compressed gas diffusion layer of proton exchange membrane fuel cells considering fiber anisotropy[J]. Membranes, 2023, 13(6): 559.

[24] Hsu C H, Pan Z B, Qu H T, et al. Green synthesis of nitrogen-doped multiporous carbons for oxygen reduction reaction using water-caltrop shells and eggshell waste[J]. RSC Advances, 2021, 11(26): 15738-15747.

[25] Tellez-Cruz M M, Escorihuela J, Solorza-Feria O, et al. Proton exchange membrane fuel cells (PEMFCs): advances and challenges[J]. Polymers, 2021, 13(18): 3064.

[26] Rofiansyah W, Zalianty F R, Ahmad La Ito F, et al. IoT-based control and monitoring system for hydroponic plant growth using image processing and mobile applications[J]. PeerJ Computer Science, 2025, 11: e2763.

[27] 娄凤强. 燃料电池并网发电系统前端 DC/DC 变换器研究[D]. 济南: 山东大学, 2011.
Lou F Q. Study of the front-ended DC/DC converter in grid-connected fuel cell generation power system[D]. Jinan: Shandong University, 2011(in Chinese).

[28] Liu X F, Song E S, Zhang L M, et al. Design and implementation for the state time-delay and input saturation compensator of gas turbine aero-engine control system[J]. Energy, 2024, 288: 129934.

[29] Luo Y Q, Jiao K. Cold start of proton exchange membrane fuel cell[J]. Progress in Energy and Combustion Science, 2018, 64: 29-61.

[30] Jiao K, Xuan J, Du Q, et al. Designing the next generation of proton-exchange membrane fuel cells[J]. Nature, 2021, 595(7867): 361-369.

[31] Ahmed S, Beauger C, Zada A, et al. Recent advancements in designing high-performance proton exchange membrane fuel cells: a comprehensive review[J]. Applied Energy, 2025, 390: 125753.

[32] Majlan E H, Rohendi D, Daud W R W, et al. Electrode for proton exchange membrane fuel cells: a review[J]. Renewable and Sustainable Energy Reviews, 2018, 89: 117-134.

[33] Tajiri K, Tabuchi Y, Wang C Y. Isothermal cold start of polymer electrolyte fuel cells[J]. Journal of the Electrochemical Society, 2007, 154(2): B147.

[34] Chacko C, Ramasamy R, Kim S, et al. Characteristic behavior of polymer electrolyte fuel cell resistance during cold start[J]. Journal of the Electrochemical Society, 2008, 155(11): B1145.

[35] 国家市场监督管理总局, 国家标准化管理委员会. 质子交换膜燃料电池汽车用燃料 氢气: GB/T 37244—2018[S]. 北京: 中国标准出版社, 2018.
State Administration for Market Regulation, Standardization Administration of the People's Republic of China. Fuel specification for proton exchange membrane fuel cell vehicles—Hydrogen: GB/T 37244—2018[S]. Beijing: Standards Press of China, 2018(in Chinese).

[36] Borras J. HY4 passenger plane sets hydrogen altitude record [EB/OL]. (2022-04-23)[2025-11-10]. https://cleantechnica.com/2022/04/23/hy4-passenger-plane-sets-hydrogen-altitude-record/.

[37] DLR. Zero-emission air transport-first flight of four-seat passenger aircraft Hy4 [EB/OL]. (2016-09-29)[2025-11-10]. https://www.dlr.de/en/latest/news/2016/20160929_zero-emission-air-transport-first-flight-of-four-seat-passenger-aircraft-hy4_19469.

[38] Liu P C, Xu S C. A review of low-temperature proton exchange membrane fuel cell degradation caused by repeated freezing start[J]. International Journal of Hydrogen Energy, 2023, 48(22): 8216-8246.

[39] Smith P J, Bennett W R, Jakupca I J. Proton exchange membrane fuel cell transient load response: NASA/TM-20210018094[R]. Washington, D. C. : NASA, 2021.

[40] ZeroAvia. ZeroAvia-ZA600-Caravan-Whitepaper-2023[R]. Washington, D. C. : NASA, 2023.

[41] PowerCell Group. PowerCell and ZeroAvia to collaborate on next generation fuel cell stack development[EB/OL]. (2024-10-29)[2025-09-21]. https://powercellgroup.com/press-releases/powercell-and-zeroavia-to-collaborate-on-next-generation-fuel-cell-stack-development/.

[42] Mazzoni F, Accardo G, Biga R, et al. Hydrogen storage system design: case studies for airborne application[C]//Proceedings of the 34th ICAS Congress. Bonn: The International Council of the Aeronautical Sciences, 2024.

[43] 朱斌, 杨少柒, 郭梁, 等. 机载液氢燃料储供系统发展现状与展望[J]. 航空学报, 2025, 46(16): 123-141.
Zhu B, Yang S Q, Guo L, et al. Development and prospects of onborne liquid hydrogen fuel storage and supply system[J]. Acta Aeronautica et Astronautica Sinica, 2025, 46(16): 123-141(in Chinese).

[44] 张晓飞, 蒋利军, 叶建华, 等. 固态储氢技术的研究进展[J]. 太阳能学报, 2022, 43(6): 345-354.
Zhang X F, Jiang L J, Ye J H, et al. Research progress of solid-state hydrogen storage teachnology[J]. Acta Energiae Solaris Sinica, 2022, 43(6): 345-354(in Chinese).

[45] Gao R H, Cai B A, Zhao Y Q, et al. Advanced performance analysis and life cycle assessment of a coupled system: MSR, LT-PEMFC and spray flash desalination[J]. Energy, 2025, 323: 135859.

[46] Yu S, Fan Y, Shi Z R, et al. Innovative heat management method and metaheuristic algorithm optimized power supply-demand balance for PEMFC-ASHP-CHP system[J]. Applied Energy, 2024, 361(C): 122806.

[47] Khandelwal M, Mench M M. Direct measurement of through-plane thermal conductivity and contact resistance in fuel cell materials[J]. Journal of Power Sources, 2006, 161(2): 1106-1115.

[48] Ji W C, Lin R. Strategy optimization of proton exchange membrane fuel cell cold start[J]. Chinese Science Bulletin, 2022, 67(19): 2241-2257.

[49] González-Espasandín Ó, Leo T J, Raso M A, et al. Direct methanol fuel cell (DMFC) and H_2 proton exchange membrane fuel (PEMFC/H2) cell performance under atmospheric flight conditions of unmanned aerial vehicles[J]. Renewable Energy, 2019, 130: 762-773.

[50] 欧凯, 胡皓文, 吴雨衡, 等. 载运工具用燃料电池空气压缩机技术综述[J]. 交通运输工程学报, 2025, 25(1): 66-93.
Ou K, Hu H W, Wu Y H, et al. Review on technologies of fuel cell air compressors for vehicles[J]. Journal of Traffic and Transportation Engineering, 2025, 25(1): 66-93(in Chinese).

[51] 潘涛. 空压机用高速永磁同步电动机的设计及损耗分析[D]. 沈阳: 沈阳工业大学, 2022.
Pan T. Design and loss analysis of high-speed permanent magnet synchronous motor for air compressor[D]. Shenyang: Shenyang University of Technology, 2022(in Chinese).

[52] 赵冬冬, 华志广, 梁艳, 等. 飞机燃料电池的离心空气压缩机特性研究[J]. 航空科学技术, 2017, 28(2): 64-68.
Zhao D D, Hua Z G, Liang Y, et al. Centrifugal air compressor characteristic research on the fuel cells in the aircraft[J]. Aeronautical Science & Technology, 2017, 28(2): 64-68(in Chinese).

[53] Franck-Lacaze L, Bonnet C, Besse S, et al. Effects of ozone on the performance of a polymer electrolyte membrane fuel cell[J]. Fuel Cells, 2009, 9(5): 562-569.

[54] Guo X J, Xu B, Ma Z S, et al. Performance analysis based on sustainability exergy indicators of high-temperature proton exchange membrane fuel cell[J]. International Journal of Molecular Sciences, 2022, 23(17): 10111.

[55] Lyu C X, Zhao P, Xie J, et al. Electrospinning of nanofibrous membrane and its applications in air filtration: a review[J]. Nanomaterials, 2021, 11(6): 1501.

[56] He C, Wen Q L, Ning F D, et al. A new integrated GDL with wavy channel and tunneled rib for high power density PEMFC at low back pressure and wide humidity[J]. Advanced Science, 2023, 10(28): 2302928.

[57] Li Y Y, He J C, Luo X L, et al. Experimental study on flow boiling-based micro-channel configurations for the PEMFC cooling[J]. Thermal Science and Engineering Progress, 2023, 46: 102210.

[58] Khatir F A, Barzegari M M, Ghadikolaee H T. Optimal geometry selection for the microchannel in bipolar plates of PEM fuel cells from a multi-criteria decision analysis point of view[J]. International Journal on Interactive Design and Manufacturing (IJIDeM), 2025, 19(4): 2789-2800.

[59] 韩敏芳, 杨翠柏, 王玉倩, 等. 固体氧化物燃料电池[J]. 物理, 2004(2): 90-94.
Han M F, Yang C B, Wang Y Q, et al. Solid oxide fuel cell[J]. Physics, 2004(2): 90-94(in Chinese).

[60] 陈志远, 王丽君, 李福燊, 等. 碳基燃料 SOFC 的 Ni-YSZ 阳极过程动力学[J]. 电源技术, 2014, 38(7): 1377-1379.
Chen Z Y, Wang L J, Li F S, et al. Process kinetics of SOFCs Ni-YSZ anode using carbon based fuels[J]. Chinese Journal of Power Sources, 2014, 38(7): 1377-1379(in Chinese).

[61] Khan M S, Lee S B, Song R H, et al. Fundamental mechanisms involved in the degradation of nickel-yttria stabilized zirconia (Ni-YSZ) anode during solid oxide fuel cells operation: a review[J]. Ceramics International, 2016, 42(1): 35-48.

[62] 李一倩, 李敬威, 吕喆. 固体氧化物燃料电池阳极的硫毒化与再生活化[J]. 硅酸盐学报, 2021, 49(1): 126-135.
Li Y Q, Li J W, Lyu Z. Poisoning and activation of anode for solid oxide fuel cell[J]. Journal of the Chinese Ceramic Society, 2021, 49(1): 126-135(in Chinese).

[63] 乔金硕, 孙克宁, 张乃庆, 等. 固体氧化物燃料电池燃料重整技术研究进展[J]. 化工进展, 2004, 23(11): 1189-1194.
Qiao J S, Sun K N, Zhang N Q, et al. Process of fuel reforming technology for SOFC'S[J]. Chemical Industry and Engineering Progress, 2004, 23(11): 1189-1194(in Chinese).

[64] 吴雨泽, 王宇旸, 范红途. 固体氧化物燃料电池 (SOFC) 系统的研究现状[J]. 能源研究与利用, 2019(1): 40-46.
Wu Y Z, Wang Y Y, Fan H T. Research status of solid oxide fuel cell (SOFC) system[J]. Energy Research & Utilization, 2019(1): 40-46(in Chinese).

[65] 张莉, 邢耀华, 钟杰, 等. 1kW SOFC-CHP 系统用催化燃烧耦合蒸汽重整反应器的实验研究[J]. 化工学报, 2016, 67(2): 557-562.
Zhang L, Xing Y H, Zhong J, et al. Experimental study on reactor integrating catalytic combustion and steam reforming for 1 kW SOFC-CHP[J]. CIESC Journal, 2016, 67(2): 557-562(in Chinese).

[66] Reuse P, Renken A, Haas-Santo K, et al. Hydrogen production for fuel cell application in an autothermal micro-channel reactor[J]. Chemical Engineering Journal, 2004, 101(1-3): 133-141.

[67] Laosiripojana N, Assabumrungrat S. Catalytic steam reforming of methane, methanol, and ethanol over Ni/YSZ: the possible use of these fuels in internal reforming SOFC[J]. Journal of Power Sources, 2007, 163(2): 943-951.

[68] 牛海涛. 二甲醚部分氧化重整制氢催化剂作用的研究[D]. 锦州: 辽宁工业大学, 2014.
Niu H T. Study on roles of catalysts in hydrogen production by partial oxidation and reforming of DME[D]. Jinzhou: Liaoning University of Technology, 2014(in Chinese).

[69] Edwards N, Ellis S R, Frost J C, et al. On-board hydrogen generation for transport applications: the HotSpot™ methanol processor [J]. Journal of Power Sources, 1998, 71(1-2): 123-128.

[70] 张伟, 李世安, 刘聪敏, 等. 车载甲醇重整微型反应器研究进展[J]. 现代化工, 2019, 39(7): 27-32.
Zhang W, Li S A, Liu C M, et al. Research progress in vehicle-mounted methanol reforming microreactor[J]. Modern Chemical Industry, 2019, 39(7): 27-32(in Chinese).

[71] Xu X H, Li P W, Shen Y S. Small-scale reforming of diesel and jet fuels to make hydrogen and syngas for fuel cells: a review[J]. Applied Energy, 2013, 108: 202-217.

[72] 袁斌, 潘建欣, 王傲, 等. 燃料电池用柴油重整制氢技术现状与展望[J]. 化工进展, 2020, 39(S1): 107-115.
Yuan B, Pan J X, Wang A, et al. Research progress and perspective of diesel reforming to hydrogen production for fuel cell applications[J]. Chemical Industry and Engineering Progress, 2020, 39(S1): 107-115(in Chinese).

[73] Samsun R C, Prawitz M, Tschauder A, et al. An autothermal reforming system for diesel and jet fuel with quick start-up capability[J]. International Journal of Hydrogen Energy, 2019, 44(51): 27749-27764.

[74] 曹军文, 覃祥富, 耿嘎, 等. 氢气储运技术的发展现状与展望[J]. 石油学报 (石油加工), 2021, 37(6): 1461-1478.
Cao J W, Qin X F, Geng G, et al. Current status and prospects of hydrogen storage and transportation technology[J]. Acta Petrolei Sinica (Petroleum Processing Section), 2021, 37(6): 1461-1478(in Chinese).

[75] Staykov A, Yamabe J, Somerday B P. Effect of hydrogen gas impurities on the hydrogen dissociation on iron surface[J]. International Journal of Quantum Chemistry, 2014, 114(10): 626-635.

[76] 刘志军. 车用汽油及其添加剂综述[J]. 化工设计通讯, 2018, 44(1): 233.
Liu Z J. Review of automotive gasoline and its additives[J]. Chemical Engineering Design Communications, 2018, 44(1): 233(in Chinese).

[77] 中华人民共和国国家质量监督检验检疫总局, 中国国家标准化管理委员会. 车用柴油: GB 19147—2016[S]. 北京: 中国标准出版社, 2016-12-02.
General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Automotive diesel fuel: GB 19147—2016[S]. Beijing: China Standards Press, 2016-12-02(in Chinese).

[78] Hou X X, Marin-Flores O, Kwon B W, et al. Gasoline-fueled solid oxide fuel cell using MoO_2-based anode[J]. Journal of Power Sources, 2014, 268: 546-549.

[79] Kale G R, Kulkarni B D. An alternative process for gasoline fuel processors[J]. International Journal of Hydrogen Energy, 2011, 36(3): 2118-2127.

[80] 王昊, 梁卫忠, 周雨霏. 柴油制氢催化工艺开发[J]. 齐鲁石油化工, 2025, 53(1): 5-10.
Wang H, Liang W Z, Zhou Y F. Development of catalytic process for hydrogen production form diesel[J]. Qilu Petrochemical Technology, 2025, 53(1): 5-10(in Chinese).

[81] Höber M, Königshofer B, Boškoski P, et al. Diesel operated combined heat and power SOFC system: SOFC characterization and proof of concept[J]. Journal of Power Sources, 2023, 585: 233635.

[82] Lawrence J, Boltze M. Auxiliary power unit based on a solid oxide fuel cell and fuelled with diesel[J]. Journal of Power Sources, 2006, 154(2): 479-488.

[83] Shilov V A, Potemkin D I, Burmatova M A, et al. Development of diesel fuel processor and optimisation of the thermal circuit of a power plant based on sofc with a power of 1 kW[J]. Chemistry for Sustainable Development, 2024(3): 387-398.

[84] Edwards T, Maurice L Q. Surrogate mixtures to represent complex aviation and rocket fuels[J]. Journal of Propulsion and Power, 2001, 17(2): 461-466.

[85] Panthi D, Du Y H, Feng H, et al. Onboard miniature jet fuel desulfurizer for mobile fuel cell power systems[J]. Energy Technology, 2024, 12(6): 2400008.

[86] 阮鹏, 林梓荣, 白帆飞, 等. 天然气水蒸气重整与 SOFC 耦合应用[J]. 新能源进展, 2023, 11(1): 29-37.
Ruan P, Lin Z R, Bai F F, et al. Coupling of steam reforming in natural gas SOFC system[J]. Advances in New and Renewable Energy, 2023, 11(1): 29-37(in Chinese).

[87] 刘爱虢, 朱悦, 李鹏超, 等. 应用 LNG 的通用飞机动力系统[J]. 沈阳航空航天大学学报, 2016, 33(6): 28-34.
Liu A G, Zhu Y, Li P C, et al. The LNG power system for general aircraft[J]. Journal of Shenyang Aerospace University, 2016, 33(6): 28-34(in Chinese).

[88] 刘军民, 廖世军. 天然气直接固体氧化物燃料电池研究进展[J]. 天然气工业, 2005, 25(11): 107-110.
Liu J M, Liao S J. Advance of study on solid oxide fuel cells directly from gas[J]. Natural Gas Industry, 2005, 25(11): 107-110(in Chinese).

[89] 李蕴, 田涛. 可持续航空燃料技术经济分析进展与启示[C]//第七届中国航空科学技术大会论文集. 北京: 中国航空学会, 2024: 782-788.
Li Y, Tian T. Progress and insights on techno-economic analysis of sustainable aviation fuel[C]//Proceedings of the 7th China Aeronautical Science and Technology Conference. Beijing: CSAA, 2024: 782-788(in Chinese).

[90] ASTM International. Standard specification for aviation turbine fuel containing synthesized hydrocarbons: ASTM D7566[S]. West Conshohocken: ASTM International, 2024-03-01.

[91] Badwal S P S, Giddey S, Kulkarni A, et al. Direct ethanol fuel cells for transport and stationary applications—a comprehensive review[J]. Applied Energy, 2015, 145: 80-103.

[92] 李季铭, 杨扬, 朱恂, 等. 甲醇水蒸汽重整制氢装置及系统研究进展[J]. 中国电机工程学报, 2024, 44(18): 7276-7292.
Li J M, Yang Y, Zhu X, et al. Review on hydrogen production reactor and system of methanol steam reforming[J]. Proceedings of the CSEE, 2024, 44(18): 7276-7292(in Chinese).

[93] 姚春德, 姚安仁. 甲醇燃料的应用现状及其展望[J]. 汽车安全与节能学报, 2023, 14(5): 521-535.
Yao C D, Yao A R. Review on methanol as fuel for engines and its future prospect[J]. Journal of Automotive Safety and Engergy, 2023, 14(5): 521-535(in Chinese).

[94] Yang B C, Koo J, Shin J W, et al. Direct alcohol-fueled low-temperature solid oxide fuel cells: a review[J]. Energy Technology, 2019, 7(1): 5-19.

[95] 邓俊彦. 氨氢联动促进氢能源的发展[J]. 科技资讯, 2024, 22(1): 158-161.
Deng J Y. Development of hydrogen energy promoted by ammonia-hydrogen linkage[J]. Science & Technology Information, 2024, 22(1): 158-161(in Chinese).

[96] 刘书贤, 杨霞, 李文霄, 等. 绿氨制取与燃料电池利用技术研究进展[J]. 能源化工, 2025(1): 1-10.
Liu S X, Yang X, Li W X, et al. Research progress of green ammonia production and fuel cell utilization technology[J]. Energy Chemical Industry, 2025(1): 1-10(in Chinese).

[97] 倪士栋, 魏胜利, 杜振华, 等. 氨燃料固体氧化物燃料电池性能研究进展[J]. 精细化工, 2024, 41(4): 750-760.
Ni S D, Wei S L, Du Z H, et al. Research progress on performance of ammonia solid oxide fuel cells[J]. Fine Chemicals, 2024, 41(4): 750-760(in Chinese).

[98] Kishimoto M, Muroyama H, Suzuki S, et al. Development of

1 kW-class ammonia-fueled solid oxide fuel cell stack[J]. Fuel Cells, 2020, 20(1): 80-88.

[99] Liu H, Qin J, Xiu X Y, et al. Comparative study of fuel types on solid oxide fuel cell-gas turbine hybrid system for electric propulsion aircraft[J]. Fuel, 2023, 347: 128426.

[100] 邢睿. 汇率、油价对国内航空公司股价影响的研究[D]. 兰州: 兰州大学, 2023.
Xing R. An empirical study on the effect of exchange rate and oil price on the stock price of domestic airlines[D]. Lanzhou: Lanzhou University, 2023(in Chinese).

[101] 王姿涵, 黄建伟. 基于国内出厂价的航油价格预测与波动风险测度研究[J]. 物流科技, 2024, 47(24): 104-111.
Wang Z H, Huang J W. Research on jet fuel price forecast and volatility risk measurement based on domestic ex-factory price[J]. Logistics Sci-Tech, 2024, 47(24): 104-111(in Chinese).

[102] 王翔宇, 刘向雷, 韩枫. 氢能航空产业配套分析[J]. 航空动力, 2023(3): 50-53.
Wang X Y, Liu X L, Han F. Analysis to the industrial supporting of hydrogen powered aviation[J]. Aerospace Power, 2023(3): 50-53 (in Chinese).

[103] Rathore S S, Biswas S, Fini D, et al. Direct ammonia solid-oxide fuel cells: a review of progress and prospects[J]. International Journal of Hydrogen Energy, 2021, 46(71): 35365-35384.

[104] Yoon S, Kang I, Bae J. Effects of ethylene on carbon formation in diesel autothermal reforming[J]. International Journal of Hydrogen Energy, 2008, 33(18): 4780-4788.

[105] Meng H, Zhang J, Yang Y S. Recent status in catalyst modification strategies for hydrogen production from ethanol steam reforming[J]. ChemCatChem, 2023, 15(17): e202300733.

[106] Xu X H, Zhang S Y, Li P W, et al. Adsorptive desulfurization of liquid Jet-A fuel at ambient conditions with an improved adsorbent for on-board fuel treatment for SOFC applications[J]. Fuel Processing Technology, 2014, 124: 140-146.

[107] Wang Y, Pasel J, Peters R. Hydrodesulfurization process with presaturation using reformate for application in a 5 kW fuel cell system[J]. Fuel Processing Technology, 2014, 127: 59-65.

[108] Kobayashi Y, Kosaka K, Tomida K, et al. Start-up characteristics of segmented-in-series tubular SOFC power modules improved by catalytic combustion at cathodes[J]. Fuel Cells, 2014, 14(6): 1028-1035.

[109] Chun O, Jamshaid F, Khan M Z, et al. Advances in low-temperature solid oxide fuel cells: an explanatory review[J]. Journal of Power Sources, 2024, 610: 234719.

[110] Wang Y Q, Shi J X, Gu X, et al. Toward mobility of solid oxide fuel cells[J]. Progress in Energy and Combustion Science, 2024, 102: 101141.

[111] Jin M H, Lee C B, Lee D W, et al. Microchannel methane steam reformers with improved heat transfer efficiency and their long-term stability[J]. Fuel, 2016, 176: 86-92.

[112] Simopoulos G N. Solid oxide fuel cell development for auxiliary power in heavy-duty vehicle applications[C]//Proceedings of the DOE Hydrogen Program Annual Progress Report. Washington, D. C. : Department of Energy, 2005: 1115-1124.

[113] Mirahmadi A, Valefi K. Study on preheating techniques for start-up of tubular solid oxide fuel cells[J]. Journal of Fuel Cell Science and Technology, 2011, 8(6): 061008.

[114] Özece B, Şenyüz S, Tokdemir İ K, et al. Model-based analysis of electrical heater implementation in solid oxide fuel cell systems for improved heat-up performance[C]//Proceedings of the IEEE 19th International Conference on Compatibility, Power Electronics and Power Engineering. Piscataway: IEEE Press, 2025: 1-4.

[115] 杨博然, 陈志光, 秦朝葵. 固体氧化物燃料电池 (SOFC) 外围热管理系统研究进展[J]. 城市燃气, 2019(3): 4-10.
Yang B R, Chen Z G, Qin C K. Progress of SOFC balance of plant(BOP)[J]. Urban Gas, 2019(3): 4-10(in Chinese).

[116] Chen Z Y, Liang F L, Mao J K, et al. Optimization design of SOFC-GT hybrid power system for aviation application[J]. Energies, 2024, 17(15): 3681.

[117] Majewski A J, Bossel U, Steinberger-Wilckens R. Catalytic reforming system suitable for transportation applications[J]. Fuel Cells, 2018, 18(4): 535-542.

[118] Shen M H. Solid oxide fuel cell-lithium battery hybrid power generation system energy management: a review[J]. International Journal of Hydrogen Energy, 2021, 46(65): 32974-32994.

[119] 朱庆山, 彭练, 黄文来, 等. 固体氧化物燃料电池密封材料的研究现状与发展趋势[J]. 无机材料学报, 2006, 21(2): 284-290.
Zhu Q S, Peng L, Huang W L, et al. Present status and development trends of sealing materials for solid oxide fuel cells[J]. Journal of Inorganic Materials, 2006, 21(2): 284-290(in Chinese).

[120] 梁爽, 毛军逵, 贺振宗, 等. 基于神经网络的航空 SOFC-GT 混合动力系统功率匹配多目标优化研究[J]. 燃气涡轮试验与研究, 2025, 38(2): 119-132.
Liang S, Mao J K, He Z Z, et al. Multi-objective optimization of power matching for aviation SOFC-GT hybrid power system based on neural network[J]. Gas Turbine Experiment and Research, 2025, 38(2): 119-132(in Chinese).

[121] 张钰凡, 马睿, 张羽翔, 等. 航空燃料电池推进系统能量管理策略研究综述[J]. 中国电机工程学报, 2023, 43(13): 5012-5025.
Zhang Y F, Ma R, Zhang Y X, et al. A review on the energy management strategy for aviation fuel cell propulsion system[J]. Proceedings of the CSEE, 2023, 43(13): 5012-5025(in Chinese).

[122] 孙创超, 陈立新, 范修林. 锂硫电池用高效大孔催化正极[J]. 材料导报, 2021, 35(1): 1001-1002.
Sun C C, Chen L X, Fan X L. An efficient macroporous catalytic cathode for Li-S battery[J]. Materials Reports, 2021, 35(1): 1001-1002(in Chinese).

[123] Staub J W, Dzara M J, Crompton K R, et al. Balancing irreversible capacity in germanium nanoparticle: HE5050 lithium ion batteries for improved energy densities exceeding 250 Wh/kg[J]. Electrochimica Acta, 2020, 332: 135455.

[124] Kwasi-Effah C C, Rabczuk T. Dimensional analysis and modelling of energy density of lithium-ion battery[J]. Journal of Energy Storage, 2018, 18: 308-315.

[125] Ellis M W, Von Spakovsky M R, Nelson D J. Fuel cell systems: efficient, flexible energy conversion for the 21st century[J]. Proceedings of the IEEE, 2001, 89(12): 1808-1818.

[126] Great Power. Great power hosts transport batteries launch event [EB/OL]. (2024-12-18)[2026-06-29]. https://www.greatpower.net/en/news/323.html.

[127] 西北工业大学. Amprius 发布适用于电动航空的超高功率高能

量锂离子电池-飞机电推进技术工业和信息化部重点实验室[EB/OL]. (2023-08-19)[2025-12-02]. https://epal.nwpu.edu.cn/info/1034/1238.htm.
Northwestern Polytechnical University. Amprius unveils ultra-high power and high-energy lithium-ion battery for electric aviation-key laboratory of aircraft electric propulsion technology, ministry of industry and information technology[EB/OL]. (2023-08-19)[2025-12-02]. https://epal.nwpu.edu.cn/info/1034/1238.htm(in Chinese).

[128] Xiao C W, Wang B, Zhao D, et al. Comprehensive investigation on lithium batteries for electric and hybrid-electric unmanned aerial vehicle applications[J]. Thermal Science and Engineering Progress, 2023, 38: 101677.

[129] 马秦慧, 陈钢, 范光涛, 等. 多工况下固体氧化物燃料电池短堆变载特性研究[J]. 热力发电, 2025, 54(3): 12-21.
Ma Q H, Chen G, Fan G T, et al. Study on variable load characteristics of a solid oxide fuel cell stack under multiple operating conditions[J]. Thermal Power Generation, 2025, 54(3): 12-21(in Chinese).

[130] Li X R, Zhao Q B, Chong D T, et al. Dynamic behavior analysis of solid oxide fuel cell system with auxiliary components under variable operating conditions[J]. International Journal of Hydrogen Energy, 2025, 146: 150001.

[131] Xiang Y, Da Y L, Zhong Z, et al. Thermo-mechanical stress analyses of solid oxide fuel cell anode based on three-dimensional microstructure reconstruction[J]. International Journal of Hydrogen Energy, 2020, 45(38): 19791-19800.

[132] Zhang X Q, Yu S Q, Wang M K, et al. Thermal stress analysis at the interface of cathode and electrolyte in solid oxide fuel cells[J]. International Communications in Heat and Mass Transfer, 2020, 118: 104831.

[133] Xu M, Li T S, Yang M, et al. Solid oxide fuel cell interconnect design optimization considering the thermal stresses[J]. Science Bulletin, 2016, 61(17): 1333-1344.

[134] 中国民用航空局. 航空发动机适航规定: CCAR-33-R2 [S]. 北京: 中国民用航空局, 2016.
Civil Aviation Administration of China. Airworthiness regulations for aero-engines: CCAR-33-R2 [S]. Beijing: Civil Aviation Administration of China, 2016(in Chinese).

[135] Meindl M, Forster F, Eckardt B, et al. Fuel Cell systems in hybrid-electric aircraft[C]//Proceedings of the European Fuel Cell Forum. Lucerne: European Fuel Cell Forum, 2025.

[136] Lyu X Q, Qu Y, Wang Y D, et al. A comprehensive review on hybrid power system for PEMFC-HEV: issues and strategies[J]. Energy Conversion and Management, 2018, 171: 1273-1291.

[137] Tsai A, Tucker D, Clippinger D. Simultaneous turbine speed regulation and fuel cell airflow tracking of a SOFC/GT hybrid plant with the use of airflow bypass valves[C]//Proceedings of the ASME 9th International Conference on Fuel Cell Science, Engineering and Technology. New York: ASME, 2011: 365-374.

[138] Zhao Q, Stalin S, Archer L A. Stabilizing metal battery anodes through the design of solid electrolyte interphases[J]. Joule, 2021, 5(5): 1119-1142.

[139] Yan P F, Nie A M, Zheng J M, et al. Evolution of lattice structure and chemical composition of the surface reconstruction layer in $Li_{1.2}Ni_{0.2}Mn_{0.6}O_2$ cathode material for lithium ion batteries[J]. Nano Letters, 2015, 15(1): 514-522.

[140] Liu L S, Feng X N, Zhang M X, et al. Comparative study on substitute triggering approaches for internal short circuit in lithium-ion batteries[J]. Applied Energy, 2020, 259: 114143.

[141] Di Prima P, Dessantis D, Versaci D, et al. Understanding calendar aging degradation in cylindrical lithium-ion cell: a novel pseudo-4-dimensional electrochemical-thermal model[J]. Applied Energy, 2025, 377: 124640.

[142] Chung C G, Kim L, Sung Y W, et al. Degradation mechanism of electrocatalyst during long-term operation of PEMFC[J]. International Journal of Hydrogen Energy, 2009, 34(21): 8974-8981.

[143] Fan L X, Zhao J J, Luo X B, et al. Comparison of the performance and degradation mechanism of PEMFC with Pt/C and Pt black catalyst[J]. International Journal of Hydrogen Energy, 2022, 47(8): 5418-5428.

[144] Singh R, Sui P C, Wong K H, et al. Modeling the effect of chemical membrane degradation on PEMFC performance[J]. Journal of the Electrochemical Society, 2018, 165(6): F3328-F3336.

[145] Yu K, Li C Z, Xie J, et al. Understanding the degradation mechanisms of Pt electrocatalysts in PEMFCs by combining 2D and 3D identical location TEM[J]. Nano Letters, 2023, 23(5): 1858-1864.

[146] Strandberg L, Shokhen V, Skoglundh M, et al. Carbon support corrosion in PEMFCs followed by identical location electron microscopy[J]. ACS Catalysis, 2024, 14(11): 8494-8504.

[147] Hoffrogge P W, Schneider D, Wankmüller F, et al. Performance estimation by multiphase-field simulations and transmission-line modeling of nickel coarsening in FIB-SEM reconstructed Ni-YSZ SOFC anodes I: influence of wetting angle[J]. Journal of Power Sources, 2023, 570: 233031.

[148] Li J, Zhang Y J, Zhao K, et al. An investigation of Cr vaporization from SUS430 metallic interconnect and deposition on (La, Sr)MnO3 cathode of intermediate temperature solid oxide fuel cells[EB/OL]. (2023-07-14)[2025-09-22]. https://webofscience.clarivate.cn/wos/woscc/full-record/WOS:001021227400001.

[149] Govindaraju N, Liu W N, Sun X, et al. A modeling study on the thermomechanical behavior of glass-ceramic and self-healing glass seals at elevated temperatures[J]. Journal of Power Sources, 2009, 190(2): 476-484.

[150] Papas P. Meeting the challenge of mitigating Li-ion battery fires for aviation[J]. Applications in Energy and Combustion Science, 2024, 20: 100286.

[151] 韩玉琪. 2023年航空氢动力进展[J]. 航空动力, 2024(1): 38-42.
Han Y Q. Progress of hydrogen powered aviation in 2023[J]. Aerospace Power, 2024(1): 38-42(in Chinese).

[152] 纪宇晗. 氢能航空时代加速来临: H2FLY完成全球首次液氢电动飞机试飞[EB/OL]. (2023-11-17) [2025-09-08]. https://ep.cannews.com.cn/publish/zghkb7/html/4994/node_209857.html.
Ji Y H. Aviation hydrogen era accelerates: H2FLY completes world's first liquid hydrogen-electric aircraft test flight[EB/OL]. (2023-11-17)[2025-09-08]. https://ep.cannews.com.cn/publish/zghkb7/html/4994/node_209857.html(in Chinese).

[153] National Institute of Advanced Industrial Science and Technology. Development and demonstration of long-duration flight with world's first solid oxide fuel cell drone[EB/OL]. (2020-06-15)[2025-09-08]. https://www.aist.go.jp/aist_j/press_release/pr2020/

pr20200615/pr20200615.html.

[154] Lockheed Martin Corporation. Stalker XE: fixed-wing VTOL-capable small UAS product sheet[EB/OL]. (2019-10-25) [2025-09-08]. https://www.lockheedmartin.com/content/dam/lockheed-martin/aero/documents/Stalker/Stalker%20XE_Product_2019.pdf.

[155] Gamble D E. World record duration flight of group 2 unmanned aircraft with VTOL and hybrid propulsion system using solid oxide fuel cell[C]//Proceedings of the AIAA SciTech 2023 Forum. Reston: AIAA. 2023.

[156] 韩玉琪. 2024 年航空氢动力进展[J]. 航空动力, 2025(1): 44-49.
Han Y Q. Progress of hydrogen powered aviation in 2024[J]. Aerospace Power, 2025(1): 44-49(in Chinese).

Fuel cells for aviation: a new paradigm for carbon-neutral flight

LI Zhipeng[1,*], GAO Zibo[1], LIU Weidong[1], PENG Feiyang[1], LIU Hong[1], HAN Qing[2]

(1. Institute of Flexible Electronics, Northwestern Polytechnical University, Xi'an 710072, China;
2. School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: Due to their great energy efficiency and zero carbon emissions, fuel cell aircraft propulsion systems have become a crucial alternative to conventional fossil fuel-based power in the context of the aviation industry's increasing attempts to become carbon neutral. This study focused on proton exchange membrane fuel cell (PEMFC) and solid oxide fuel cell (SOFC), systematically analyzing their working principles, system architectures, key technological challenges, and applicable scenarios. The research also compared these systems with lithium battery propulsion systems from multiple perspectives, including energy carrier type, power characteristics, service life, and environmental benefits. The findings reveal that while SOFC have great promise for long-endurance missions, PEMFC are better suited for short-range aircraft applications. Although challenges remain in hydrogen storage, thermal management, system integration, and cost, continued advances in structural optimization and hybrid energy architectures are expected to significantly improve the endurance and emission reduction performance of fuel cell-based aviation power systems. These developments support their role as a critical technological pathway toward green aviation.

Keywords: fuel cell; aviation; solid oxide fuel cell; proton exchange membrane fuel cell; sustainable aviation; hybrid powertrains

Received: 2025-11-17; **Accepted**: 2025-11-28; **Published Online**: 2025-12-17 19:07

URL: link.cnki.net/urlid/11.2625.V.20251217.1313.001

Foundation items: The Fundamental Research Funds for the Central Universities (D5000230358); Shaanxi Province Key Research and Development Plan (2024GX-YBXM-456)

* **Corresponding author**. E-mail: iamzpli@nwpu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0724

人工智能赋能小行星抵近探测的研究现状及展望

尹继豪[1,*]，魏晓东[1]，崔林艳[1]，吴文峻[2]，张晓明[3]，刘传凯[4]

(1. 北京航空航天大学 宇航学院，北京 100191；2. 北京航空航天大学 人工智能研究院，北京 100191；3. 中国科学院国家天文台，北京 100101；4. 北京航天飞行控制中心，北京 100094)

摘　　要：小行星已成为深空探测领域的主要目标之一。围绕小行星抵近探测安全、高效处置的需求，综述了主要航天强国面向小行星抵近探测的人工智能(AI)技术应用及设想。从深空环境智能感知、智能规划及决策2个方面梳理了当前的研究进展，分析了当前人工智能技术应用于小行星抵近探测任务中面临的模型可靠性难以度量、未知环境适应性不足、星载计算资源受限等难点与挑战。结合中国未来的小行星抵近探测任务和行星防御系统构建任务，探讨和展望小行星抵近探测中人工智能技术的未来发展方向与应用趋势。

关　键　词：小行星探测；行星防御；人工智能；智能感知；智能决策

中图分类号：P185.7; V57

文献标志码：A　　**文章编号**：1001-5965(2026)10-3396-12

自20世纪90年代以来，针对小行星开展探测的手段从绕飞探测发展到附着采样，再到最近的双星重定向实验(double asteroid redirection test, DART)中的精准撞击。中国于2025年实施“天问二号”任务，首次对近地小行星2016HO3开展伴飞探测并取样返回，之后前往主带彗星311P开展伴飞探测。面向小行星探测目标距离远、轨道不确定、认知有限带来的挑战，各国在小行星探测无人化、自主化与智能化的方向上持续发力。

人工智能是一门研究和开发模拟、延伸和扩展人类智能的理论、方法、技术及应用系统的学科，借助计算机和某些感知与执行单元完成某些与人类智能有关的复杂功能[1]。人工智能技术是实现深空探测器自主化的一种技术途径，其应用将推动探测器局部自主能力的实现和提升[2]。

本文将概述各航天“强国”小行星抵近探测的实施及人工智能技术应用及规划，分析并总结相关的人工智能方法和面临的难点与挑战。在此基础上，结合中国未来的小行星探测任务和行星防御系统构建目标，探讨和展望小行星抵近探测中人工智能技术的未来发展方向与应用趋势。

1　小行星探测概述

小行星保留了太阳系早期形成和演化历史信息，还可能蕴含着地球生命起源的重要线索，是研究太阳系的起源和演化历史、探索生命起源的重要途径。同时，由于小行星蕴含丰富的稀有矿物，也是未来太空矿物资源开发的主要目标。另一方面，轨道与地球公转轨道相交的近地小行星撞击地球，是地球安全与人类发展长期面临的重大潜在威胁。恐龙灭绝的小行星撞击学说被提出以来，小行星撞击风险已经引起了国际社会的广泛担忧，被认为是“可能造成人类灭亡的重大灾害”之一。

自20世纪90年代开始，以小行星的科学探测或在轨处置为目标的抵近探测活动日益增多，已成

收稿日期：2024-10-08；**录用日期**：2024-11-07；**网络出版时间**：2024-11-29 10:47
网络出版地址：link.cnki.net/urlid/11.2625.V.20241129.1003.001
基金项目：国家科技重大专项(2022ZD0117402)
*通信作者. E-mail：jihaoyin@buaa.edu.cn

引用格式：尹继豪，魏晓东，崔林艳，等. 人工智能赋能小行星抵近探测的研究现状及展望[J]. 北京航空航天大学学报，2026，52(10)：3396-3407. Yin J H，Wei X D，Cui L Y，et al. Artificial intelligence for asteroid proximity exploration research progress and future perspectives[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3396-3407 (in Chinese).

为深空探测领域备受关注的热点，是当前各航天强国的重点发展目标之一[3]。至今为止，世界各国共实施了 13 次小行星抵近探测活动，其中，21 世纪已实施 9 次，任务及实施时间如图 1 所示，日本、美国、欧洲等国家和机构处于领先地位，具有代表性的包括对龙宫小行星进行采样并返回的隼鸟 2 号任务、首次行星防御技术和动能撞击技术演示的双小行星重定向测试任务等。小行星抵近探测任务的实施推动了高精度自主导航、弱引力天体表面作业、复杂过程高可靠自主规划与决策等深空探测关键技术发展，同时也是先进航天技术的试验场：随着多种小行星智能探测的创新构想和技术探索不断涌现，人工智能技术在深空探测领域的应用潜力日益显著。近 20 年来，不断有融合人工智能理论的小行星抵近探测设想被提出，涉及飞行动力学[4-5]、智能机器人、集群智能等多个领域。

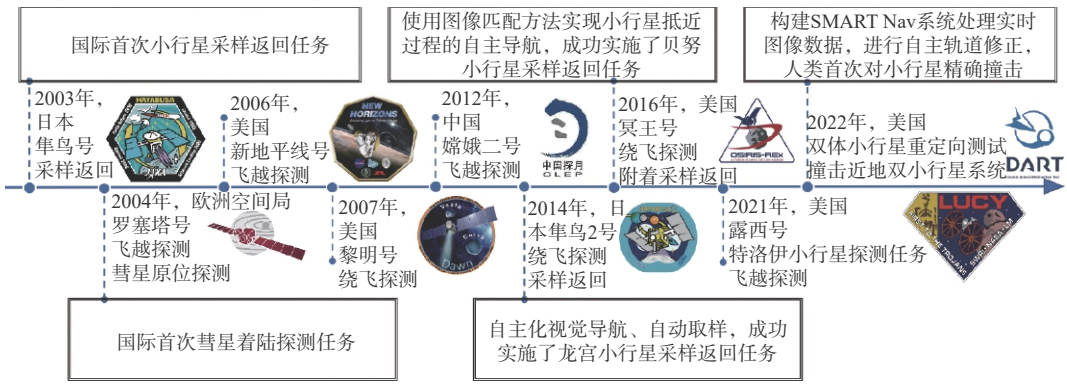


图 1　21 世纪已实施的小行星抵近探测活动

Fig. 1　Asteroids proximity exploration missions implemented in the 21th century

中国首次于 2012 年嫦娥二号再拓展任务中对 4179 号小行星实施了飞越探测[6]，于 2025 年开展“天问 2 号”任务，计划通过一次发射，实现从近地小行星 2016 HO3 采样返回地球，之后前往主带彗星 311P 开展伴飞探测任务；后续中国还将开展首次小行星防御实验任务，为获取小行星精细数据、深度认知小行星物理特性、掌握小行星运行规律提供了机遇和平台。

小行星附着采样-返回任务典型过程如图 2 所示，包含发射前的地面观测/监测和发射后的转移、接近、交会、环绕、附着采样等抵近探测阶段及采样后的返回阶段，区别于月球、火星等大天体，小行星抵近探测具有目标距离远、轨道不确定、对象认知有限等特点，现有探测主要依靠天地反复交互迭代的方式，逐步逼近探测目标和开展复杂探测任务，使探测周期很长、探测效率不高。

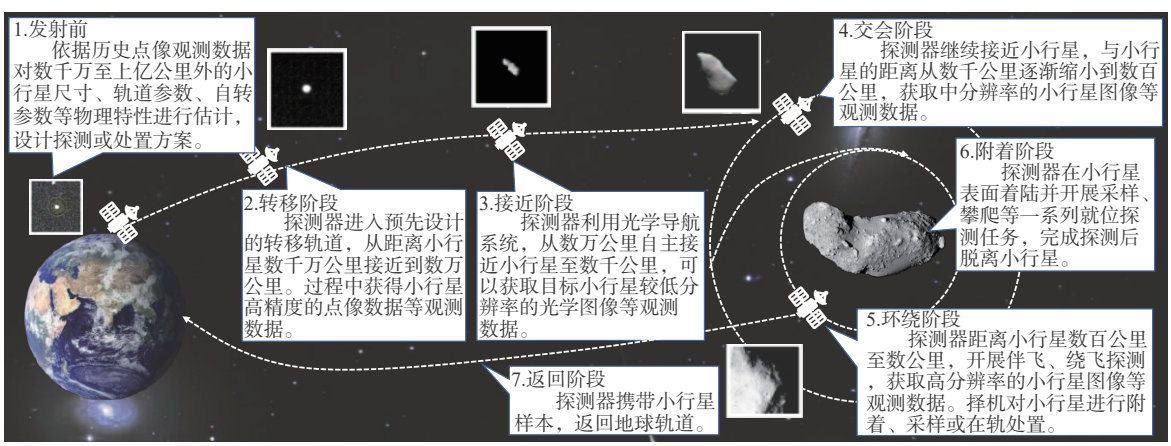


图 2　小行星采样返回任务典型过程

Fig. 2　Typical procedure for an asteroid sample-return mission

具体来说，小行星具有“微引力、不确定”[7]的特点：一方面，由于“微引力”，探测器难以建立环绕轨道开展全方位详细探查，也不能被小行星重力吸引实现安全着陆，使小行星抵近与附着呈现不稳定状态，增大了任务开展的难度；另一方面，依靠地面观测仅能获得小行星的基本轨道参数和少量信息，对小行星体积大小、地形地貌、运动特性、引力场等特性难以事先把握，带来认识上的“不确

定”。针对上述问题，发展小行星抵近探测人工智能技术是具有潜力的解决方案。人工智能技术使探测器系统具备与人类相似的智能感知及灵活应变能力，克服“微引力、不确定”的小行星环境中动态不确定和信息不完备等不利条件，提升任务优化的效率和决策水平[8]，有利于详查阶段的探测效率提升和绕飞/伴飞、攀爬、采样等高动态任务的开展。

本文阐述了主要航天强国面向小行星抵近探测的人工智能技术应用及设想，并对面向小行星抵近探测任务的人工智能探索性研究进行了综述和总结，同时，对其面临的难题与挑战及发展趋势进行了展望，并就未来可能的实际应用给出建议，可为相关研究工作提供参考。

2 各航天强国面向小行星抵近探测的智能技术及设想

近年来，人工智能在多种应用上彰显出巨大应用潜力[9-10]，并在众多新任务、新问题上体现出广泛的适应性，因此，人工智能技术在深空探测领域的应用潜力日益显著。国内外人工智能技术在空间科学和工程领域的应用处在较为初级的阶段，特别是赋能器上自主感知与管理的应用较少，然而，多种小行星智能探测的创新构想和技术探索不断涌现，受到了各航天强国的广泛关注。在中国即将实施的“天问二号”任务中，也指出将突破智能控制、弱引力天体绕飞及取样等关键技术[11]，采用智能自主技术应对小行星不确定的目标特性，实现小行星三维重构的测量与建模、小行星接近过程的自主信息识别与自主导航、小行星采样的自主操作等[12]。

2.1 美国国家航空航天局

以美欧为代表的航天大国很早就开展了近地小行星的观测和防御计划，整合大量地基望远镜组成近地天体观测计划、国际小行星预警网等全球性的小行星地面观测网，筛选出具有潜在科学探测价值或撞击风险的近地小行星，进而开展抵近附着探测或撞击重定向。美国国家航空航天局 (National Aeronautics and Space Administration, NASA) 已经支持并逐步部署多项应用于卡特林那巡天系统(Catalina Sky Survey, CSS)、美国小行星陆地撞击持续报警系统(Asteroid Terrestrial-impact Last Alert System, Atlas)检测和预警小行星的人工智能模型[13]，大幅提升了从每晚数十太字节的数据中提取小行星目标的自动化水平。

NASA 在 OSIRIS-Rex 任务中采用自然路标导航方式进行自主导航，通过将数字地形模型中的特征与实时导航图像匹配，估计探测器相对于目标小行星的位置和速度。具体来说，在开展绕飞/伴飞探测过程中，探测器收集小行星表面图像，生成小行星数字地形模型和用于地标导航的地图(maps for landmark navigation, MLNs)。抵近过程中，根据飞行轨迹和相机参数估计导航相机的可见范围，并进一步与该范围内小行星表面图像、数字地形模型进行配准，从而估计探测器相对于小行星的运动状态。NASA 认为，由于 Bennu 小行星表面以碎石堆为主、地形崎岖危险，基于自然路标特征的自主导航技术对于接触即离(touch-and-go, TAG)模式的采样任务成功实施非常重要。

NASA 还提出了多种利用人工智能技术进行小行星抵近或附着探测的构想，其中，智能化任务调度及规划是 NASA 关注的重点方向。NASA 的 JPL 实验室人工智能小组目前正研制自动调度和规划环境(automated scheduling and planning environment, ASPEN)系统[14]，该系统旨在降低运营成本和提高航空航天系统的自主性。ASPEN 系统利用人工智能技术将高级任务目标转化为有效的行为指令，主要应用于深空探测器、行星巡视器等。NASA 的 JPL 实验室与欧洲空间局合作构建的自动调度和规划组件(automated scheduling and planning environment-rosetta SGS scheduling component, ASPEN-RSSC)[15]，已应用于罗赛塔小行星号探测器的中期任务规划，该系统在 2014 年 1 月罗赛塔号脱离休眠后投入使用，解决了其所携带的 11 个科学仪器极其复杂的探测任务规划问题，总共成功制定了 35 个中期任务规划。

NASA 提出的另一项曾引起广泛影响的科学设想是利用自主微纳卫星技术集群(autonomous nano technology swarm, ANTS)[16]执行小行星勘探任务。如图 3 所示，该构想拟利用大量的微型、异构、携带不同载荷的智能探测器，自主协同开展针对主带小行星的抵近探测，可实现联合导航、协同的小行

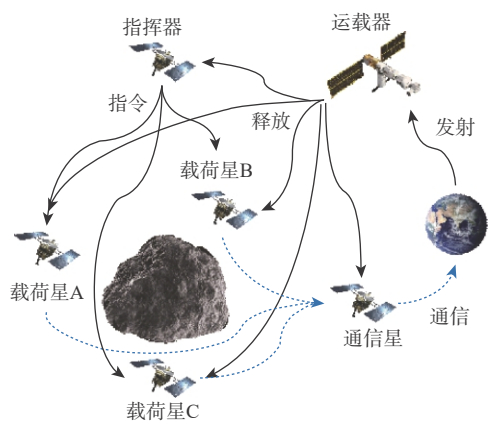


图 3 ANTS 实施小行星探测构想

Fig. 3 Concept of ANTS for asteroids exploration

星特征描述, 在无需先验形状信息的情况下, 自主探测并估计小行星重力场、3D 形状和旋转运动等物理特征, 实现人工参与度的降低和时间效率的提升。然而, 截至目前仍鲜有该框架下的相关任务实施。

2.2　日本宇宙航空研究开发机构

日本宇宙航空研究开发机构(Japan Aerospace Exploration Agency, JAXA)在隼鸟 2 号探测任务中重视自主化技术, 提出了自主化实施视觉导航、自动取样方法。在附着采样过程中, 利用可部署的人工标记, 设计了归一化二值化阈值等图像处理方法, 将隼鸟 2 号获取光学图像中的人工地标与小行星背景进行区分, 并保证了太阳角度较小等恶劣光照条件下的鲁棒性, 使在自主导航和控制过程中探测器能够准确地检测和跟踪位于小行星表面的人工地标, 在隼鸟 2 号实施小行星表面采样任务过程中起到关键作用[17]。

JAXA 计划在 2026 年实施火星卫星探测(Martian Moon eXploration, MMX)任务, 从火卫一取样返回, 面临小天体的近距离操作、样本回收等挑战。在该任务中, JAXA 提出将人工智能视为应对未来的小天体探测及其他深空探测任务、实现自主化探测的重要手段[18], 目前, 其重点研究领域包括精准着陆附着技术、具有障碍物规避的可靠着陆附着方案、在崎岖地形上的安全着陆附着机制等。其中, 为不同地形外观下深空探测器高效准确的自主视觉定位, JAXA 利用现有立体视觉里程计(visual odometer, VO)和同步定位与地图构建(simultaneous localization and mapping, SLAM)技术, 研发了适应深空天体表面的视觉特征检测框架和稀疏特征点的探测器运动估计等方法[19], 以实现不同地形外观下的高效和准确的视觉定位。针对现有“Move-Wait”遥操作模式在深空环境下通信效率受限的问题, JAXA 开发了智能路径规划系统, 检测可能导致探测车沿着可能与障碍物发生碰撞的轨迹[20], 并自主使用观测数据对航迹进行优化[21], 以实现探测器的高效和安全探索。

2.3　欧洲空间局

欧洲空间局于 2024 年 10 月 7 日成功发射了 Hera 号探测器, 旨在进一步推进人工智能在小行星探测中的应用[22], 其重要目标之一是在深空环境中进行自主导航和控制系统的测试及验证。Hera 探测器针对 DART 撞击的 Didymos 双小行星系统开展抵近探测, 并释放 Juventas 号和 Milani 号立方星在 Dimorphos 小行星上附着。其自主导航和控制系统能够使探测器在没有地面控制的情况下自主完成复杂的飞行任务, 从而解决深空通信延迟和实时决策需求的矛盾, 任务中将验证多传感器融合、自主视觉导航、自主特征地形跟踪等关键智能技术。

在多传感器融合方面, Hera 探测器将搭载类似于 GPU 的专用处理设备, 利用机器视觉算法融合小行星成像相机、星敏感器、激光高度计、热红外相机和惯性测量单元等多种传感器数据, 来构建小行星环境特性模型, 并实时提供探测器的位置信息和姿态数据。在自主视觉导航方面, Hera 探测器将利用基于高分辨率网格(high-resolution net, HRNet)开发的深度学习识别技术, 识别双小行星系统中较大的 Didymain 小行星, 并将其作为参考点进行导航[23]。当探测器抵近至 Didymain 小行星的距离小于 8 km 时, 探测器将验证自主特征地形跟踪技术, 采用机器视觉方法在不同图像中对巨石和陨石坑等表面特征进行匹配, 同时, 融合热红外相机和惯性测量单元数据, 实现探测器的航迹推算和飞越小行星夜间面时的决策应变。

已实施的小行星抵近探测任务基于详查数据构建精确的小行星环境模型, 基于数理方法设计自主系统使其具备做出自主选择的能力, 为小行星抵近探测的地形识别与导航、附着和采样区域选取及高动态任务中的决策应变等提供支持。同时, 包括中国在内的各主要航天强国及机构密切在关注如何推进人工智能技术在小行星抵近探测任务中的应用, 进一步提升探测器的环境感知、规划决策能力。以下针对小行星抵近探测的智能感知和智能决策 2 个方面的研究展开论述。

3　面向小行星抵近探测的智能感知技术

探测器要安全有效地执行科学与工程任务, 需要感知小行星附近环境, 智能感知技术通过实现任务周期内感知效能的提升, 支持探测器的规划和决策。目前, 面向小行星抵近过程的智能感知主要包括小行星形貌智能重建、小行星表面地物识别和小行星引力场智能重建等。

3.1　小行星形貌智能重建

小行星形貌的重建和识别是影响抵近和附着策略的重要因素。在 DART 任务中, “意大利光立方体卫星”伴随撞击器发射, 用于 Dimorphos 小行星形状建模和撞击地点规划[24]。该任务中主要的形状建模技术是立体光斜测量 (stereophotoclinometry, SPC) 方法。SPC[25] 方法由 JPL 实验室提出, 通过测量 2 个或多个拍摄位置不同的影像中同一地面点的像点位置, 推导出该地面点的三维坐标。该方法具有空间精度高的优点, 在小行星探测任务的感知

中起到了重要作用，还被应用于隼鸟2号探测龙宫小行星[26]、冥王号探测贝努小行星[27]、罗塞塔号探测67P号彗星[28]等历次任务中。但SPC方法使用每个地面点多组不同光照、不同角度的观测图像进行运算，要求大量人工参与制定探测器伴飞轨迹。为降低SPC对人工参与的需求，Brandonisio等[29]提出结合人工智能方法，利用光照入射角度、成像距离等参数评估小行星表面的图像质量，并基于深度强化学习方法产生在探测周期、推力等约束下的伴飞路径。

近期提出的神经辐射场、符号距离场（signed distance function, SDF）等隐式三维表征方法在三维重建任务中受到广泛关注。Chen等[30]提出利用隐式三维表征方法对小行星三维形状建模，并在糸川和龙宫小行星数据上对三维重建效果进行了实验验证。Chen等[31]使用SDF方法从小行星的多视角姿态图像重建3D表面模型，使用可学习的外观特征适应不同光照条件输入，处理小行星图像中常见的不同太阳高度和曝光条件变化。值得注意的是，目前，基于隐式三维表征的方法可以替代传统三维重建流程的表面重建部分，但无法进行图像位姿估计。由于小行星的表面通常包含弱纹理区域，在图像位姿估计过程中必须使用大量手动选择的控制点，需要一定程度的人工干预。未来的工作仍需对图像位姿估计开展进一步研究，以进一步增强小行星三维形貌重建过程的自主程度。

3.2 小行星表面地物识别

在小行星表面地物的识别方面，目前的研究主要关注陨石坑[32]、岩石等目标，服务于附着阶段的危险规避、科学探测的位置选取等，是提升任务安全性和环境适应性的重要手段[33]。深度学习模型是目前识别、分割等任务中主流和最优的方法，因此，深度学习模型在小行星表面地物识别上的应用受到研究者的广泛关注。

Pugliatti等[34]提出利用卷积神经网络对小行星、彗星等不规则的小天体进行辨识，并建立了包括谷神星、贝努小行星、4179号小行星、67P号彗星等11个小天体的数据集，准确率达到了98.52%。该工作验证了卷积神经网络等深度学习方法在小行星数据上相比传统方法具有突出的优势。Latorre等[35]利用月面数据学习陨石坑的特征，设计了基于U-Net的全卷积神经网络陨石坑识别模型，同时结合迁移学习方法，仅使用少量谷神星的数据进行微调，实现了接近80%的陨石坑识别准确率。

小行星表面地物识别的目标是分析表面环境、提高任务实施安全性。在小行星表面地物进行识别和分类基础上，Caroselli等[36]将Didymos小行星双星系统作为仿真场景，构建了在小行星表面软着陆操作过程的图像数据集，利用卷积神经网络识别其中的岩石，并根据岩石分布评估附着安全性，实现了软着陆过程中图像附着安全性的自主化评估。

3.3 小行星引力场智能重建

开展小行星抵近探测时，引力场的重建是克服弱引力场环境，实现抵近和接近过程轨道设计、表面附着和采样的关键，其精细重建主要依赖于探测器在绕飞、伴飞过程中的摄动量，或绕飞、伴飞过程中观测到的漂浮物轨迹推算。早期方法通常假设目标小行星具有均匀的密度，再通过对形状积分的方法估计小行星附近的引力场[37]。然而，密度均匀的假设并非广泛适用于组成成分、形成过程多样的小行星，因此，轨道确定过程中测量的引力场通常与形状模型积分产生的引力场不同。另一种常见的方法是利用球谐函数表征引力场[38]，球谐函数在布里渊球外对引力场有良好的拟合效果，然而，在形状距离球体差别较大的小行星上并不适用。

由于神经网络出色的拟合能力，近年来在小行星和其他天体的引力场重建上得到了广泛关注。Turgut[39]提出将反向传播神经网络（backpropagation artificial neural network, BPANN）作为引力场建模的工具，通过训练人工神经网络从数据中学习三维位置和重力加速度之间的映射，在不需要有关引力场和天体质量分布的先验知识实现了地球引力场的重建。这种范式被推广到其他模型上，如Gao等[40]利用高斯过程回归模型反演重力场和三维位置之间的直接映射关系；Furfaro等[41]采用贝叶斯优化的方法训练单层前馈神经网络，实现不规则小天体引力场建模；Cheng等[42]利用包含全连接层的深度前馈网络重建小行星的不规则引力场等。这些工作利用数据驱动的方法，使用神经网络拟合小行星的引力场，借助大量仿真数据实现了快速、准确的引力估计，但在小行星抵近探测的场景中，受探测器机动能力和任务周期的限制，可获取的轨道偏移量等支撑引力场建模的数据较有限，完全基于数据驱动的深度学习方法可能面临模型过拟合、泛化性能差等困难。

针对数据驱动的人工智能方法对数据量需求高的问题，Raissi等[43]提出了物理信息引导的神经网络（physics-informed neural network, PINN），通过将基础物理方程等先验知识引入神经网络的损失函数中，驱动网络在符合物理规律的范围内学习数据的映射关系，实现了数据驱动与物理知识的同步

驱动。Martin 等[44]将这一思想引入到天体引力场重建的问题中, 利用引力场的拉普拉斯方程作为知识约束, 使用多层感知机实现了天体引力场的重建。此外, 也有研究者使用 Hopfield 神经网络对物理场特性进行建模, 将引力场的特性作为先验知识并减少对数据量的需求。Pasquale 等[45]提出基于 Hopfield 神经网络进行球谐系数的在线调优来重建小行星的重力场, Zhao 等[46]基于旋转质量偶极子模型简化模型参数, 使用少量参数实现对细长形态小行星重力场的表征。

4　小行星抵近探测的智能决策

小行星抵近探测任务过程复杂、探测模式多样, 包含接近阶段、伴飞/绕飞阶段、抵近附着阶段等多个过程。面向复杂的小行星抵近探测过程, 在认识上“不确定”的深空环境中, 现有探测任务主要采用地面事先规划、器上自主实施相结合的模式, 实施过程中需进行多次调整。如何根据环境和任务规划探测行为, 提升探测器智能化水平, 实现小行星抵近过程中的在线执行决策, 从而提升探测安全性, 并缩短探测任务周期, 是当前各国重要的探索方向。

4.1　面向绕飞、详查阶段的智能决策技术

面向绕飞、详查阶段的智能决策技术主要关注探测器的自主导航及航迹规划, 以实现器上自主决策和控制。在 DART 任务中, 器上自主决策系统得到了较好的应用, 突破了器上自主图像处理、目标检测和精确制导等技术, 提升了探测器的精确自主控制水平[47]。撞击器上搭载了基于预设规则进行决策的自主小型机动实时导航 (SMART Nav) 系统, 在撞击前 245 min 开始进行自主制导[48]。SMART Nav 的作用过程如图 4 所示, 通过光学图像识别 Didymos 和 Dimorphos 组成的双星系统, 并在撞击前 73 min 基于目标的亮度和像素数量自动识别出双星系统中较暗、较小的撞击目标 Dimorphos 小行星, 随后开始自主地向 Dimorphos 调整航向。SMART Nav 直到撞击前 2.5 min 仍在持续进行制导, 最终达到了约 15 m 的撞击精度。

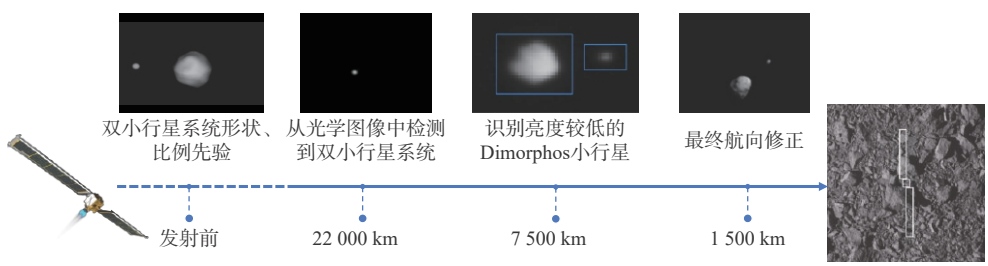


图 4　SMART Nav 基于光学影像实现对 Dimorphos 小行星撞击的引导

Fig. 4　SMART Nav system utilizes optical imaging for navigation to Dimorphos

相较于 DART, 其他小行星抵近探测任务过程更加复杂。除交会段的自主导航外, 还涉及伴飞/绕飞的详查轨道规划等复杂过程的自主决策控制。通过规划合适的轨道, 探测器在小行星周围进行多次接近和飞越, 从而获取不同角度上的高分辨率图像、光谱等数据。小行星伴飞/绕飞轨道规划需要考虑小行星的形状、质量分布、引力场和探测器的机动能力、燃料限制等多种因素, 同时, 小行星自转速度较快且自转轴不稳定, 使绕飞探测轨道动力学问题存在较强的非线性, 导致绕飞轨道的设计与分析非常困难。因此, 传统规划方法的难点在于需要建立复杂的模型和算法。

为提升小行星绕飞详查的效率, 一部分工作借用了主动 SLAM 的思想, 使用探测器智能探索环境并进行测绘。Pesce 等[49]提出了定量表征某一轨道下小行星测绘完整性和准确性的计算方法, 从而实现了小行星绕飞轨道的选择。在此基础上, Piccinin 等[50]提出利用深度 Q 学习设计小行星绕飞轨道上的图像采集策略。进一步地, Chan 等[51]将小行星的绕飞探测规约为一个部分可观察马尔可夫决策过程, 构建了小行星绕飞探测的模拟环境, 使用深度 Q 学习框架训练决策器, 综合考虑太阳方位角、探测器方位角、图像质量、限制探测器至目标的距离等因素实现伴飞轨迹规划。Piccinin 等[52]提出通过测绘生成小行星的三维点云质量评价规划结果, 并使用强化学习模型获得产生最高三维点云质量的轨迹。另一类方法考虑小行星引力较小, 难以获得卫星围绕目标小行星飞行的轨道, 因此, 将小行星的伴飞问题看作对空间非合作目标的跟踪问题。例如, Zhou 等[53]使用目标跟踪算法 SiamRPN[54]跟踪图像中的小行星等目标并判断其方位, 使用深度 Q 学习训练智能体, 引导航天器接近小行星等空间非合作目标。

然而, 目前的智能导航及航迹规划方案研究多从理想情况出发, 部分或完全忽略了小行星抵近探测过程中复杂的工程和安全性约束, 距离实际应用

仍有较远距离。

4.2 小行星附着采样智能决策技术

小行星表面引力微弱、逃逸速度小，同时，由于小行星的形状和引力场不规则、事先仅能对表面动力学环境部分特性进行估计[55]，实现精准软着陆和附着面临着技术挑战。目前的工作主要从动力学或视觉导航的角度出发，进行附着位置的选取和附着过程的精细控制。

4.2.1 结合动力学的方法

从动力学出发的工作通过对引力场及小行星表面的建模和分析，进行小行星附着过程的实时优化控制。经典的控制方法采用同伦法[56]、凸优化方法[57]等进行最优附着轨迹的设计。例如，崔平远等[58]在小行星自旋速度、引力场等动力学参数存在着较大不确性的情况下，根据名义的动力学参数设计探测器离轨速度，将探测器引向目标附着位置，并在最终接近段，采用势函数制导法确定每次制动的时刻及制动速度脉冲的大小、方向，实现小行星附着过程的自主实时优化控制。

初期，基于人工智能的方法主要引入深度学习模型描述小行星的引力场，在此基础上进行附着过程的优化。例如，Song等[59]利用深度神经网络估计不规则小行星的重力和重力梯度，再以燃料最优的优化目标进行了附着过程的优化。Gaudet等[60]将强化元学习用于小行星抵近操作的自适应制导、导航与控制，利用光学导引角和激光测距等观测量直接生成发动机推力指令，并在六自由度仿真中验证了其在未知小行星动力学环境、执行机构故障和传感器偏差等条件下完成着陆机动的鲁棒性。

近期的工作则进一步采用端到端的强化学习方法，实现弱引力场下的附着过程控制。Qi等[61]设计了四足机器人小行星附着一体化姿态控制框架，训练以根据重力分布和地形信息作为输入的智能体，实现了附着过程中探测器相对地表姿态的控制。

4.2.2 基于视觉导航的方法

从视觉导航出发的方法侧重于关注附着位置的可达性和安全性，通过相邻图像估计探测器的运动信息，进而对探测器的轨迹进行规划和控制[33]。部分工作基于边缘特征、几何特征、自然路标特征[62]（如陨石坑[63]等）实现高精度的附着过程引导，或通过对岩石的模板匹配[64]、地表高程图和崎岖度的计算[65]评估安全附着区域。深度学习等人工智能方法进一步提升了对地形的理解能力。

Ravani等[66]提出通过深度学习将附着点的规划问题转换为安全附着位置的检测问题，使用Mask CNN检测小行星表面的安全附着位置。Mancini等[67]设计了基于孪生卷积神经网络的地形相对导航方法，使用深度学习模型检测预先设定的自然路标位置，克服了附着过程中因图像光照条件、尺度和旋转等因素变化而难以检测预设路标的问题。

面向小行星抵近探测任务的伴飞绕飞、着陆附着、攀爬采样等高动态操作高效、精准实施的要求，人工智能模型的灵活应变特性将有助于减少与地面通信和调整次数，提升探测效率，但当前研究仍以技术探索为主，距离实际场景和应用有较大差距。类似于小行星抵近探测场景的智能感知技术，小行星抵近探测的智能决策方法也面临数据稀缺、质量低等问题。同时，从动力学角度出发的方法和从视觉、激光雷达等感知数据出发的方法“泾渭分明”，尚未出现实现综合考虑多种环境感知输入的一体化智能规划决策方法。此外，基于强化学习的方法虽然逐渐受到关注，但当前研究中对探测器智能体构建的动作空间和观测空间通常维度较低，距离实际任务和真实场景差距较远，智能体从仿真环境迁移到真实环境将产生泛化性的问题。

5 小行星探测中人工智能技术的发展趋势

近年来，以机器学习为代表的人工智能技术得到了快速发展，在计算机视觉、机器人等领域获得了日益广泛的成功应用，为构建新的航天器操控模式、发展智能化深空探测体系等提供了理论基础。同时，后续小行星和其他深空探测任务以对更深更远更广阔空间的探测和资源开发利用为主体，附着、攀爬、设施部署等高动态任务中对探测器自主能力，尤其是灵活应变和适应能力也提出了更高更迫切的需求。如表1所示，未来需要重点关注的人工智能发展方向可能包括：

1）大规模预训练模型。当前大量天文观测数据驱动了人工智能方法应用于小行星检测识别、定轨、特性反演，形成具有优越性能的专用模型，例如，以残差网络（residual network, ResNet）、VGG网络、

表1 小行星探测中人工智能技术的发展趋势

Table 1 Evolutionary trends of AI techniques for asteroids exploration

技术领域	大规模预训练模型	可信人工智能	内嵌数理知识的人工智能	具身化智能
发展方向	实现多样数据和任务的协同促进，形成小行星特性反演的通用智能范式	实现人工智能模型的稳定可靠、增强人工智能使用信心	引入数理规律先验知识，实现可控、可信的环境感知	支持智能模型向真实物理世界迁移，实现智能模型决策和应变能力提升
实现目标	提升智能模型性能	增强智能方法可靠性	提高深空未知环境下的适应性	提高深空未知环境下的适应性

密集连接卷积网络(densely connected convolutional network, DenseNet)等卷积神经网络模型作为骨干网络处理光学观测数据的智能方法[68-71]、以循环神经网络(recurrent neural network, RNN)、长短期记忆网络(long short-term memory, LSTM)等模型处理元数据序列的模型[72-73]、以统计学习方法进行小行星聚类的模型[74-75] 等, 已经在单一设备的特定任务上取得了优异的表现。然而, 如何综合利用来源于不同观测系统的海量、多模态、无标注数据; 如何挖掘不同模态数据、多种物理特性间的关联, 实现多样数据和任务的协同促进, 最终形成小行星特性反演的通用智能范式, 是尚未解决的问题。

近期以来, 以 Transformer 架构为基础的大规模预训练模型快速发展, 其智能涌现现象带来的超越预期的泛化性, 出现了 ChatGPT、Segment Anything[76] 等自然语言处理和图像处理模型, 并迅速扩展到自动驾驶、气象预测[77]、数学推理等领域。利用大规模预训练模型适应多模态输入数据的优势, 结合其优越的特征挖掘能力和泛化性, 是形成天文领域通用人工智能、支持小行星或其他空间目标特性智能反演的通用范式的途径之一。

2) 可信人工智能。人工智能模型的可靠性、稳定性、可解释性是其能否应用于深空探测领域的核心问题。由于深度学习具备高度复杂性和不确定性, 无法直观理解决策背后的判断依据和原因, 从而可能引发不确定的风险, 如对抗攻击相关研究[78] 表明, 少量关键特征信息的改变就有可能改变人工智能的判断结果, 这与重大深空探测任务的高可靠性、高成功率要求相违背。因此, 实现人工智能模型的稳定可靠、增强人工智能使用信心是深空探测领域应用的重要议题。

当前, 人工智能的稳定性仍面临较大挑战, 针对人工智能模型可解释性的研究仍处在初级阶段。随着对人工智能模型可解释性研究的深入, 未来将对人工智能模型如何利用数据特征进行预测和判断, 从而刻画人工智能模型对特征重要性的描述方法、特征与决策的关联机制, 使人工智能模型的决策过程更加可解释、可控制, 有助于其在深空探测等高可靠性要求领域的运用。此外, 针对小行星这一深空目标, 随着综合考虑其轨道特性、表面形貌、引力场分布等多方面特性的数字或半实物仿真环境逐步完善, 是协同人工智能可解释性理论增强智能决策模型可信度的重要途径。当前学术界已通过实验在可信人工智能方向上取得一定成果[79-80], 但尚未形成理论体系支撑。未来在该领域的研究将不断增多, 人工智能模型应用于小行星抵近探测的可行性将逐步得到验证。

3) 内嵌数理知识的人工智能。由于地面数据分辨率较低、已有抵近探测数据量有限, 目前小行星物理特性反演方法对小行星内部结构等复杂物理特性认知能力有限, 尤其是针对中国深空探测任务规划目标的了解有限。同时, 尽管将其他小行星的知识利用迁移学习方法推广到中国探测目标上是一种可能的方法, 但小行星具有多种可能构型, 其物理特性迥异, 人工智能模型不具备基于现有的数据建立起具有泛化性和普适的小行星特性认知的理论基础。因此, 在未来抵近探测任务中获取有限数据的条件下, 如何快速、可靠地逐步精细反演小行星物理特性、建立小行星形貌模型, 是需要重点关注的问题。

实现有限数据条件下物理特性精细反演的途径通常是引入数理规律等先验知识。近些年来, 内嵌数理知识的人工智能模型引起广泛关注[81], 其应用从偏微分方程求解扩展到地球、月球等大天体引力场的反演[44]。内嵌数理知识的人工智能方法可以有效聚合神经网络模型的特征提取能力和数理规律的先验约束, 且由于神经网络的推理速度仅与网络本身有关, 内嵌数理知识的人工智能模型在物理场高精度表达这一场景下具有超越传统方法的时间和内存效率。此外, 引入数理规律约束是引导神经网络产生可信、可控结果的一种可行途径。同时, 在人工智能模型中嵌入数理知识是将环境和工程约束条件引入智能模型的有效手段。因此, 内嵌数理知识的人工智能模型用于小行星特性反演和深空抵近探测的环境感知是值得关注和探索的研究方向之一。

4) 具身化智能。具身智能强调智能体在物理世界中与环境进行交互的能力, 包括感知环境、理解环境和在环境中做出决策和行动。其与非具身智能模型的区别在于智能体通过与环境交互, 从真实的物理或虚拟的仿真数字空间中学习, 并产生理解环境并适应和改造环境的能力。因此, 智能决策模型如何适应环境、适应任务、适配探测器计算效能, 是具身智能研究的主要问题。具身化智能研究强化学习、形态学习等领域, 目前的具身化智能研究主要聚焦在机器人抓取、视觉认知等场景。面向仿真环境与真实物理世界间的鸿沟, 已有具身化智能通过构建包含多项复杂任务和动作空间高维度智能体[82]、形态可变的深度进化强化学习智能体[83]、优化仿真环境[84-85] 等方式, 评估智能体在仿真空间中适应环境的能力, 并为最终向开放、复杂的真实世界迁移做准备。

具身化智能的研究有助于结合探测器自身形

态和载荷，指导智能决策模型的适配，进而逐步演化出复杂多变环境下更强的决策和感知能力。在未来探测任务中，探测器及其搭载的功能载荷有望构成具有高维度动作空间的智能体，在认知有限的深空环境中执行多项复杂任务。现有的探索包括利用视觉信息进行小天体表面形貌的自主感知[86]、对探测器自身状态认知及异常检测[87]等。随着具身智能受到业界关注，为具身智能的学习提供环境的具身仿真技术及平台也在蓬勃发展。目前，已涌现了一批物理高保真、对象可交互、环境可演进的仿真模拟器[88]。结合逐步抵近探测数据构建精细化仿真环境，在地面同步开展具身化智能相关验证实验，为探测策略的制定提供辅助，是验证人工智能模型感知和灵活应变能力、为探测任务智能决策提供补充的有益尝试。

6 结束语

本文介绍了当前小行星探测中人工智能方法的研究现状，总结了现有人工智能方法在小行星探测领域面临的问题，并对人工智能在小行星探测领域的趋势与方向进行了展望。

面向小行星探测中任务种类多、协调要求高、执行难度大等挑战，人工智能技术可能彻底改变空间探索技术格局，在自主应变、交互时效、环境适应等技术难题上产生变革性影响。从发展现状来看，人工智能赋能小行星探测是多学科交叉的先进研究领域，具有重大科学意义和应用潜力，但涉及的理念门类众多、核心理论与创新性技术仍在持续发展和完善。

小行星结构独特、人类认知有限，探测过程存在极大不确定性；同时探测器与地面通信延时大，地面无法全程掌握探测器运行状态并实时提供支持。根据小行星抵近探测任务特点和需求，人工智能赋能小行星探测的核心问题是实现小行星环境的智能认知、探测过程的智能决策。利用人工智能方法提升小行星环境感知的精确性与广泛性，增强对复杂环境的全方位认知能力，最终完成复杂的决策与控制任务，提升小行星探测任务效益与安全性。

目前，人工智能技术在小行星探测和其他深空探测领域的应用方兴未艾，中国应加快构建人工智能赋能小行星探测技术路线图，制定技术实现具体途径；在小行星物理特性智能反演和环境认知环节上先应用和验证人工智能技术，在不影响现有探测开展的情况下进行技术验证，进一步推动中国空间科学与空间技术发展。

参考文献（References）

[1] 于登云, 张哲, 泮斌峰, 等. 深空探测人工智能技术研究与展望[J]. 深空探测学报, 2020, 7(1): 11-23.
Yu D Y, Zhang Z, Pan B F, et al. Development and trend of artificial intelligent in deep space exploration[J]. Journal of Deep Space Exploration, 2020, 7(1): 11-23(in Chinese).

[2] 叶培建, 孟林智, 马继楠, 等. 深空探测人工智能技术应用及发展建议[J]. 深空探测学报, 2019, 6(4): 303-316.
Ye P J, Meng L Z, Ma J N, et al. Suggestions on artificial intelligence technology application and development in deep space exploration[J]. Journal of Deep Space Exploration, 2019, 6(4): 303-316 (in Chinese).

[3] 叶培建, 彭兢. 深空探测与我国深空探测展望[J]. 中国工程科学, 2006, 8(10): 13-18.
Ye P J, Peng J. Deep space exploration and its prospect in China[J]. Engineering Science, 2006, 8(10): 13-18(in Chinese).

[4] 李俊峰, 曾祥远. 不规则小行星引力场内的飞行动力学[J]. 力学进展, 2017, 47(1): 429-451.
Li J F, Zeng X Y. Flight dynamics in the gravitational fields of irregular asteroids[J]. Advances in Mechanics, 2017, 47(1): 429-451 (in Chinese).

[5] Izzo D, Märtens M, Pan B F. A survey on artificial intelligence trends in spacecraft guidance dynamics and control[J]. Astrodynamics, 2019, 3(4): 287-299.

[6] 乔栋, 黄江川, 崔平远, 等. 嫦娥二号卫星飞越探测小行星的目标选择[J]. 中国科学: 技术科学, 2013, 43(6): 602-608.
Qiao D, Huang J C, Cui P Y, et al. Target selection of Chang'e-2 satellite flying over to detect asteroids[J]. Scientia Sinica (Technologica), 2013, 43(6): 602-608(in Chinese).

[7] 张荣桥, 黄江川, 赫荣伟, 等. 小行星探测发展综述[J]. 深空探测学报, 2019(5): 417-423.
Zhang R Q, Huang J C, He R W, et al. The development overview of asteroid exploration[J]. Journal of Deep Space Exploration, 2019(5): 417-423(in Chinese).

[8] 程林, 蒋方华, 李俊峰. 深度学习在飞行器动力学与控制中的应用研究综述[J]. 力学与实践, 2020, 42(3): 267-276.
Cheng L, Jiang F H, Li J F. A review on the applications of deep learning in aircraft dynamics and control[J]. Mechanics in Engineering, 2020, 42(3): 267-276(in Chinese).

[9] 郑志明, 吕金虎, 韦卫, 等. 精准智能理论: 面向复杂动态对象的人工智能[J]. 中国科学: 信息科学, 2021, 51(4): 678-690.
Zheng Z M, Lyu J H, Wei W, et al. Refined intelligence theory: artificial intelligence regarding complex dynamic objects[J]. Scientia Sinica (Informationis), 2021, 51(4): 678-690(in Chinese).

[10] 张钹, 朱军, 苏航. 迈向第三代人工智能[J]. 中国科学 (信息科学), 2020, 50(9): 1281-1302.
Zhang B, Zhu J, Su H. Toward the third generation of artificial intelligence[J]. Science in China (Information Sciences), 2020, 50(9): 1281-1302(in Chinese).

[11] 国家国防科工局. 国家国防科工局: 务实推动国防科技工业高质量发展[EB/OL]. (2023-08-22) [2024-09-25]. https://www.sastind.gov.cn/n10086200/n10100131/c10347911/content.html.
State Administration of Science, Technology and Industry for National Defense. Pragmatically promoting high-quality development

of National Defense Science and Technology Industry[EB/OL]. (2023-08-22) [2024-09-25]. https://www.sastind.gov.cn/n10086200/n10100131/c10347911/content.html(in Chinese).

[12] 张熇. 小行星探测任务的信息技术挑战[EB/OL]. (2021-11-24) [2024-09-25]. http://academics.casad.cas.cn/xsjl/kxyjsqylt/2021/d122c/202111/t20211124_4979344.html.
Zhang H. Information technology challenges in asteroid exploration missions[EB/OL]. (2021-11-24)[2024-09-25]. http://academics.casad.cas.cn/xsjl/kxyjsqylt/2021/d122c/202111/t20211124_4979344.html(in Chinese).

[13] International Astronomical Union. Minor planet center: latest published data [EB/OL]. (2023-04-01)[2024-09-25]. https://www.minorplanetcenter.net/mpc/summary.

[14] Smith B, Sherwood R, Govindjee A, et al. Representing spacecraft mission planning knowledge in ASPEN[C]//Proceedings of the Artificial Intelligence Planning Systems Workshop on Knowledge Acquisition. Palo Alto: AAAI Press, 1998.

[15] Chien S A, Rabideau G, Tran D Q, et al. Activity-based scheduling of science campaigns for the Rosetta orbiter[J]. Journal of Aerospace Information Systems, 2021, 18(10): 711-727.

[16] Clark P, Rilee M, Curtis S, et al. BEES for ANTS: space mission applications for the autonomous NanoTechnology swarm[C]//Proceedings of the AIAA 1st Intelligent Systems Technical Conference. Reston: AIAA, 2004.

[17] Ogawa N, Terui F, Mimasu Y, et al. Image-based autonomous navigation of Hayabusa2 using artificial landmarks: the design and brief in-flight results of the first landing on asteroid Ryugu[J]. Astrodynamics, 2020, 4(2): 89-103.

[18] Doyle R, Kubota T, Picard M, et al. Recent research and development activities on space robotics and AI[J]. Advanced Robotics, 2021, 35(21-22): 1244-1264.

[19] Cocaud C, Kubota T. Development of an intelligent simulator with SLAM functions for visual autonomous landing on small celestial bodies[J]. Journal of Advanced Computational Intelligence and Intelligent Informatics, 2011, 15(8): 1167-1174.

[20] Otsu K, Kubota T. Energy-aware terrain analysis for mobile robot exploration[C]//Proceedings of the Field and Service Robotics: Results of the 10th International Conference. Berlin: Springer, 2016: 373-388.

[21] Ishigami G, Otsuki M, Kubota T. Range-dependent terrain mapping and multipath planning using cylindrical coordinates for a planetary exploration rover[J]. Journal of Field Robotics, 2013, 30(4): 536-551.

[22] Michel P, Küppers M, Bagatin A C, et al. The ESA Hera mission: detailed characterization of the DART impact outcome and of the binary asteroid (65803) Didymos[J]. The Planetary Science Journal, 2022, 3(7): 160.

[23] Kaluthantrige A, Feng J L, Gil-Fernández J. CNN-based image processing algorithm for autonomous optical navigation of Hera mission to the binary asteroid Didymos[J]. Acta Astronautica, 2023, 211: 60-75.

[24] Daly R T, Ernst C M, Barnouin O S, et al. Shape modeling of dimorphos for the double asteroid redirection test (DART)[J]. The Planetary Science Journal, 2022, 3(9): 207.

[25] Gaskell R W, Barnouin-Jha O S, Scheeres D J, et al. Characterizing and navigating small bodies with imaging data[J]. Meteoritics & Planetary Science, 2008, 43(6): 1049-1061.

[26] Hirata N, Hirata N, Tanaka S, et al. Initial results of shape modeling on the asteroid Ryugu from observations by Hayabusa2 for landing site selection[C]//Proceedings of the AAS/Division for Planetary Sciences Meeting Abstracts#50. Washington, D. C.: AAS/Division for Planetary Sciences, 2018, 50: 501.05.

[27] Barnouin O S, Daly M G, Palmer E E, et al. Digital terrain mapping by the OSIRIS-REx mission[J]. Planetary and Space Science, 2020, 180: 104764.

[28] Gaskell R W, Jorda L, Palmer E, et al. Comet 67P/CG: preliminary shape and topography from SPC[C]//Proceedings of the AAS/Division for Planetary Sciences Meeting Abstracts. Washington, D. C.: AAS/Division for Planetary Sciences #46, 2014, 46: 209.04.

[29] Brandonisio A, Lavagna M, Guzzetti D. Reinforcement learning for uncooperative space objects smart imaging path-planning[J]. The Journal of the Astronautical Sciences, 2021, 68(4): 1145-1169.

[30] Chen H, Hu X Y, Willner K, et al. Neural implicit shape modeling for small planetary bodies from multi-view images using a mask-based classification sampling strategy[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2024, 212: 122-145.

[31] Chen S H, Wu B, Li H L, et al. Asteroid-NeRF: a deep-learning method for 3D surface reconstruction of asteroids[J]. Astronomy & Astrophysics, 2024, 687: A278.

[32] 冯军华, 崔祜涛, 崔平远, 等. 行星表面陨石坑检测与匹配方法[J]. 航空学报, 2010, 31(9): 1858-1863.
Feng J H, Cui H T, Cui P Y, et al. Autonomous crater detection and matching on planetary surface[J]. Acta Aeronautica et Astronautica Sinica, 2010, 31(9): 1858-1863(in Chinese).

[33] 崔平远, 贾贺, 朱圣英, 等. 小天体光学导航特征识别与提取研究进展[J]. 宇航学报, 2020, 41(7): 880-888.
Cui P Y, Jia H, Zhu S Y, et al. Research progress on optical navigation feature recognition and extraction technologies for small body exploration[J]. Journal of Astronautics, 2020, 41(7): 880-888 (in Chinese).

[34] Pugliatti M, Topputo F. Small-body shape recognition with convolutional neural network and comparison with explicit features based method[J]. Advances in the Astronautical Sciences, 2021, 175: 2539-2558.

[35] Latorre F, Spiller D, Sasidharan S T, et al. Transfer learning for real-time crater detection on asteroids using a fully convolutional neural network[J]. Icarus, 2023, 394: 115434.

[36] Caroselli E, Belien F, Falke A, et al. Deep learning-based passive hazard detection for asteroid landing in unexplored environment[C]//Proceedings of the 44th Annual American Astronautical Society Guidance, Navigation, and Control Conference. Berlin: Springer, 2024: 319-334.

[37] Werner R A. Spherical harmonic coefficients for the potential of a constant-density polyhedron[J]. Computers & Geosciences, 1997, 23(10): 1071-1077.

[38] Werner R A, Scheeres D J. Exterior gravitation of a polyhedron derived and compared with harmonic and mascon gravitation representations of asteroid 4769 Castalia[J]. Celestial Mechanics and Dynamical Astronomy, 1996, 65(3): 313-344.

[39] Turgut B. Application of back propagation artificial neural networks for gravity field modelling[J]. Acta Montanistica Slovaca, 2016, 21(3): 200-207.

[40] Gao A, Liao W T. Efficient gravity field modeling method for small

bodies based on Gaussian process regression[J]. Acta Astronautica, 2019, 157: 73-91.

[41] Furfaro R, Barocco R, Linares R, et al. Modeling irregular small bodies gravity field via extreme learning machines and Bayesian optimization[J]. Advances in Space Research, 2021, 67(1): 617-638.

[42] Cheng L, Wang Z B, Song Y, et al. Real-time optimal control for irregular asteroid landings using deep neural networks[J]. Acta Astronautica, 2020, 170: 66-79.

[43] Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations[J]. Journal of Computational Physics, 2019, 378: 686-707.

[44] Martin J, Schaub H. Physics-informed neural networks for gravity field modeling of the Earth and Moon[J]. Celestial Mechanics and Dynamical Astronomy, 2022, 134(2): 13.

[45] Pasquale A, Silvestrini S, Capannolo A, et al. Small bodies non-uniform gravity field on-board learning through hopfield neural networks[J]. Planetary and Space Science, 2022, 212: 105425.

[46] Zhao Y J, Yang H W, Li S, et al. On-board modeling of gravity fields of elongated asteroids using Hopfield neural networks[J]. Astrodynamics, 2023, 7(1): 101-114.

[47] Adams E, O'Shaughnessy D, Reinhart M, et al. Double asteroid redirection test: the Earth strikes back[C]//Proceedings of the IEEE Aerospace Conference. Piscataway: IEEE Press, 2019: 1-11.

[48] Daly R T, Ernst C M, Barnouin O S, et al. Successful kinetic impact into an asteroid for planetary defence[J]. Nature, 2023, 616(7957): 443-447.

[49] Pesce V, Agha-Mohammadi A A, Lavagna M. Autonomous navigation & mapping of small bodies[C]//Proceedings of the IEEE Aerospace Conference. Piscataway: IEEE Press, 2018: 1-10.

[50] Piccinin M, Lunghi P, Lavagna M. Deep reinforcement learning-based policy for autonomous imaging planning of small celestial bodies mapping[J]. Aerospace Science and Technology, 2022, 120: 107224.

[51] Chan D M, Agha-Mohammadi A A. Autonomous imaging and mapping of small bodies using deep reinforcement learning[C]//Proceedings of the IEEE Aerospace Conference. Piscataway: IEEE Press, 2019: 1-12.

[52] Piccinin M, Lavagna M R. Deep reinforcement learning approach for small bodies shape reconstruction enhancement[C]//Proceedings of the AIAA Scitech 2020 Forum. Reston: AIAA, 2020: 1909.

[53] Zhou D, Sun G H, Lei W X, et al. Space noncooperative object active tracking with deep reinforcement learning[J]. IEEE Transactions on Aerospace and Electronic Systems, 2022, 58(6): 4902-4916.

[54] Li B, Yan J J, Wu W, et al. High performance visual tracking with Siamese Region proposal network[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 8971-8980.

[55] 张永隆, 闫翛然, 李俊峰, 等. 小行星 2016 HO3 表面动力学环境分析及采样挑战[J]. 力学与实践, 2021, 43(5): 674-679.
Zhang Y L, Yan X R, Li J F, et al. The analysis of surface dynamical environments and sampling difficulties for asteroid 2016 Ho3[J]. Mechanics in Engineering, 2021, 43(5): 674-679(in Chinese).

[56] Jiang F H, Baoyin H, Li J F. Practical techniques for low-thrust trajectory optimization with homotopic approach[J]. Journal of Guidance, Control, and Dynamics, 2012, 35(1): 245-258.

[57] Pinson R, Lu P. Rapid generation of optimal asteroid powered descent trajectories via convex optimization[C]//Proceedings of the AAS/AIAA Astrodynamics Specialist Conference. San Diego: Univelt, 2015: 2655-2672.

[58] 崔平远, 朱圣英, 崔祜涛. 小天体自主软着陆脉冲机动控制方法研究[J]. 宇航学报, 2008, 29(2): 511-516.
Cui P Y, Zhu S Y, Cui H T. Autonomous impulse maneuver control method for soft landing on small bodies[J]. Journal of Astronautics, 2008, 29(2): 511-516(in Chinese).

[59] Song Y, Cheng L, Gong S. Fast estimation of gravitational field of irregular asteroids based on deep neural network and its application [J]. Advances in the Astronautical Sciences, 2019, 168: 1721-1735.

[60] Gaudet B, Linares R, Furfaro R. Seeker based adaptive guidance via reinforcement meta-learning applied to asteroid close proximity operations[EB/OL]. (2019-07-13)[2026-06-28]. https://arxiv.org/abs/1907.06098.

[61] Qi J, Gao H B, Yu H T, et al. Integrated attitude and landing control for quadruped robots in asteroid landing mission scenarios using reinforcement learning[J]. Acta Astronautica, 2023, 204: 599-610.

[62] Xiao Y, Li S, Wang G Z, et al. Deep learning prediction frame matching algorithm of small celestial navigation landmarks[J]. Journal of Deep Space Exploration, 2022, 9(4): 400-406(in Chinese).

[63] Tian Y, Yu M. A novel crater recognition based visual navigation approach for asteroid precise pin-point landing[J]. Aerospace Science and Technology, 2017, 70: 1-9.

[64] Yu M, Cui H T. Robust hazard matching approach for visual navigation application in planetary landing[J]. Aerospace Science and Technology, 2015, 47: 378-387.

[65] Johnson A, Montgomery J, Matthies L. Vision guided landing of an autonomous helicopter in hazardous terrain[C]//Proceedings of the IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2006: 3966-3971.

[66] Ravani K, Mathavaraj S, Padhi R. Site detection for autonomous soft-landing on asteroids using deep learning[J]. Transactions of the Indian National Academy of Engineering, 2021, 6(2): 365-375.

[67] Mancini P, Cannici M, Matteucci M. Deep learning for asteroids autonomous terrain relative navigation[J]. Advances in Space Research, 2023, 71(9): 3748-3760.

[68] Cazeneuve D, Marchis F, Blaclard G, et al. ODNet: a convolutional neural network for asteroid occultation detection[J]. The Astronomical Journal, 2023, 165(1): 11.

[69] Kaluthantrige A, Feng J, Gil-Fernández J. Pseudorange measurement and Sun phase angle estimation using CNN-based image processing algorithm for Hera mission[C]//Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference. Hawaii: Maui Economic Development Board, 2022: 371-377.

[70] Rabeendran A C, Denneau L. A two-stage deep learning detection classifier for the ATLAS asteroid survey[EB/OL]. (2021-01-22)[2026-06-28]. https://arxiv.org/abs/2101.08912.

[71] Carruba V, Aljbaae S, Caritá G, et al. Optimization of artificial neural networks models applied to the identification of images of asteroids' resonant arguments[J]. Celestial Mechanics and Dynamical Astronomy, 2022, 134(6): 59.

[72] Naul B, Bloom J S, PÉRez F, et al. A recurrent neural network for classification of unevenly sampled variable stars[J]. Nature Astronomy, 2018, 2(2): 151-155.

[73] Carruba V, Aljbaae S, Domingos R C, et al. Machine learning applied to asteroid dynamics[J]. Celestial Mechanics and Dynamical Astronomy, 2022, 134(4): 36.

[74] Zappalà V, Bendjoya P, Cellino A, et al. Asteroid families: search of a 12, 487-asteroid sample using two different clustering techniques[J]. Icarus, 1995, 116(2): 291-314.

[75] Zappala V, Cellino A, Farinella P, et al. Asteroid families. I-Identification by hierarchical clustering and reliability assessment[J]. The Astronomical Journal, 1990, 100: 2030.

[76] Kirillov A, Mintun E, Ravi N, et al. Segment anything[EB/OL]. (2023-04-05)[2024-09-20]. https://arxiv.org/abs/2304.02643.

[77] Bi K F, Xie L X, Zhang H H, et al. Pangu-weather: a 3D high-resolution model for fast and accurate global weather forecast[EB/OL]. (2022-11-03)[2024-09-20]. https://arxiv.org/abs/2211.02556.

[78] Deng Y P, Karam L J. Universal adversarial attack via enhanced projected gradient descent[C]//Proceedings of the IEEE International Conference on Image Processing. Piscataway: IEEE Press, 2020: 1241-1245.

[79] Liu T L, Tao D C, Song M L, et al. Algorithm-dependent generalization bounds for multi-task learning[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(2): 227-241.

[80] He F X, Liu T L, Tao D C. Control batch size and learning rate to generalize well: theoretical and empirical evidence[C]//Proceedings of the 33rd Conference on Neural Information Processing Systems. Red Hook: Curran Associates, Inc. , 2019: 1141-1150.

[81] Karniadakis G E, Kevrekidis I G, Lu L, et al. Physics-informed machine learning[J]. Nature Reviews Physics, 2021, 3: 422-440.

[82] Srivastava S, Li C S, Lingelbach M, et al. BEHAVIOR: benchmark for everyday household activities in virtual, interactive, and ecological environments[EB/OL]. (2021-08-06)[2024-09-20]. https://arxiv.org/abs/2108.03332.

[83] Gupta A, Savarese S, Ganguli S, et al. Embodied intelligence via learning and evolution[J]. Nature Communications, 2021, 12: 5721.

[84] 杜德嵩, 宋以拓, 刘延芳, 等. 面向空间航天器机器人学习算法研究的高保真仿真平台[J]. 空间控制技术与应用, 2023, 49(3): 10-17.
Du D S, Song Y T, Liu Y F, et al. A high-fidelity simulation environment for spacecrafts with robot learning algorithms[J]. Aerospace Control and Application, 2023, 49(3): 10-17(in Chinese).

[85] 姚文龙, 刘毅, 邵巍, 等. 小天体地形动态模拟与着陆视景仿真系统设计[J]. 深空探测学报 (中英文), 2022, 9(4): 427-437.
Yao W L, Liu Y, Shao W, et al. Design of terrain dynamic simulation and landing view simulation system for small celestial bodies[J]. Journal of Deep Space Exploration, 2022, 9(4): 427-437 (in Chinese).

[86] Wei X D, Cui L Y, Liu C K, et al. Integrating neural radiance fields with deep reinforcement learning for autonomous mapping of small bodies[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 5642717.

[87] Song Y N, Li Z, Zhang Z Y, et al. SSC-KA: a framework of space situational knowledge acquisition for status-cognition of satellites[J]. IEEE Access, 2022, 10: 101750-101761.

[88] Liu Y, Chen W X, Bai Y J, et al. Aligning cyber space with physical world: a comprehensive survey on embodied AI[EB/OL]. (2024-07-09)[2024-09-20]. https://arxiv.org/abs/2407.06886.

Artificial intelligence for asteroid proximity exploration research progress and future perspectives

YIN Jihao[1,*], WEI Xiaodong[1], CUI Linyan[1], WU Wenjun[2], ZHANG Xiaoming[3], LIU Chuankai[4]

(1. School of Astronautics, Beihang University, Beijing 100191, China;
2. Institute of Artificial Intelligence, Beihang University, Beijing 100191, China;
3. National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China;
4. Beijing Aerospace Control Center, Beijing 100094, China)

Abstract: One of the main goals of deep space exploration is asteroid exploration, which has great scientific significance. Faced with the challenges of long distances and limited knowledge of asteroid exploration targets, the rapidly evolving technology of artificial intelligence (AI) has emerged as a promising complement to existing space mission implementation paradigms. In addition to introducing the current state of research on the application of AI methods in asteroid exploration, this paper analyzes the methods of AI systems used in space exploration missions and asteroid ground observation systems. It also summarizes the challenges that these methods face. It also proposes key technologies with developmental potential for asteroid exploration.

Keywords: asteroid exploration; planetary defense; artificial intelligence; intelligent sensing; intelligent control

Received: 2024-10-08; **Accepted**: 2024-11-07; **Published Online**: 2024-11-29 10:47

URL: link.cnki.net/urlid/11.2625.V.20241129.1003.001

Foundation item: National Science and Technology Major Project (2022ZD0117402)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0559

锂离子电池模组电热耦合灾变机制研究进展

贺元骅，梁家鑫，李子桐，黄江*
(中国民用航空飞行学院 民机火灾科学与安全工程四川省重点实验室，广汉 618307)

摘 要：锂离子电池的热灾害问题一直是其应用过程中重大安全威胁之一。锂离子电池系统热灾害存在界面热稳定性衰退、单体电池内短路热失控、模组乃至系统热扩散的灾变逻辑过程。电与热交互的耦合灾变则既是表征锂离子电池系统热灾害逻辑的重要维度，又是治理锂离子电池系统热灾害的关键突破口。针对锂离子电池系统长周期、多尺度下的电热耦合灾变问题，归纳总结了单体电池固体电解质界面(SEI)膜分解阶段、隔膜融化阶段，以及模组热扩散阶段蕴含的自产热、焦耳热及热传导等热行为对电行为的作用机制，梳理了过充、过放及老化等电行为在上述阶段的产热及传热行为，概括了单体及模组最具代表性的电-热耦合模型。从微观与宏观多角度探索电热行为的联动机制，并对未来锂离子电池模组电热耦合领域进行了展望，旨在为锂离子电池模组的设计与运维提供参考。

关 键 词：锂离子电池；电热耦合；SEI膜分解；隔膜熔化；热扩散
中图分类号：TM912; V242.2
文献标志码：A **文章编号**：1001-5965(2026)10-3408-18

锂离子电池作为先进能源代表[1]，其使用仍存在较大的热灾害风险。各滥用条件下，锂离子电池极易发生一系列连锁放热反应，继而形成热失控，最终酿致锂离子电池模组热扩散[2]。2011—2023年，全球共发生700多起电动汽车起火事件及75起储能电站失火事件，在各应用领域中，运行状态占事故发生总状态的75%以上[3]。从锂离子电池模组热灾害发生发展过程来看，其有一个由单体电池固体电解质界面(solid electrolyte interphase, SEI)膜分解自产热[4]，到单体电池隔膜熔化内短路[5]，再到模组热扩散[6]的3个恶化过程，且每个阶段都有伴随的热行为和电行为。因此，国际国内锂离子电池系统安全标准也从热行为和电行为2个方面推进锂离子电池模组热灾害的治理，联合国《Global technical regulation on the electric vehicle safety》[7]、中国《电动汽车用动力蓄电池安全要求》[8]、《电力储能用锂离子电池》[9]等从温升、电压等维度界定动力、储能领域电池和系统的安全性。学者们从其每个阶段诸多电热耦合特征进行分析，以揭示锂离子电池模组灾变生成演化机制。本文拟依据锂离子电池模组各阶段电热参数动态变化特征，分析其电热灾变的作用机制及转化机理，归纳总结迄今单体及模组多尺度下所采用的电-热耦合模型，以厘清锂离子电池模组电热耦合灾变机制研究发展趋势。

1 SEI膜分解阶段电热耦合

锂离子电池模组热灾害发端于单体电池热滥用、电滥用和机械滥用[10]，最早的灾变标志为SEI膜分解自产热[11]，进而显著地影响到单体电池的电行为，一定程度上表现为电热耦合的变动情形[12]。受热滥用的锂离子电池自产热形成多路热源的叠

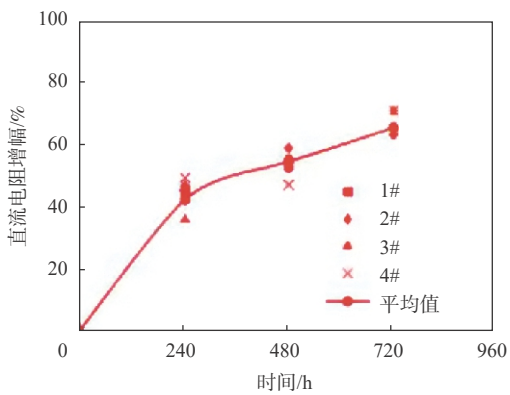
收稿日期：2024-07-26；录用日期：2024-08-30；网络出版时间：2024-09-24 12:31
网络出版地址：link.cnki.net/urlid/11.2625.V.20240924.1014.001
基金项目：国家自然科学基金(U2333210)；中央高校基本科研业务费专项资金(24CAFUC10161,24CAFUC03076)
*通信作者. E-mail：jianghuangscu@163.com

引用格式：贺元骅，梁家鑫，李子桐，等. 锂离子电池模组电热耦合灾变机制研究进展[J]. 北京航空航天大学学报，2026，52(10)：3408-3425. He Y H，Liang J X，Li Z T，et al. Progress in the investigation of electro-thermal coupled catastrophic mechanisms in lithium-ion battery modules[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3408-3425 (in Chinese).

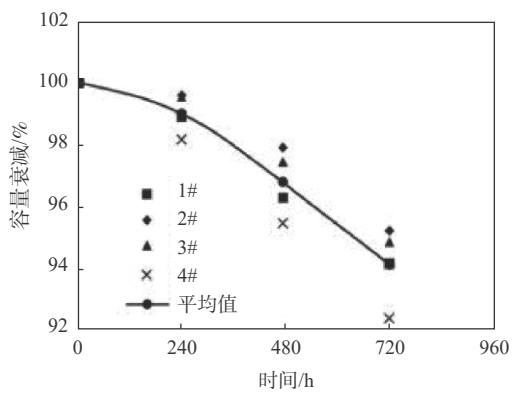
加, 进而引发电行为的恶化。而电池电行为的恶化, 又会加剧热行为的恶化。学者们聚焦于锂离子电池滥用热及自产热与电阻、电压等电行为, 以及过放、过充温升及多充放电倍率产热占比等主题, 系统地研究了其热行为和电行为的关系。

1.1　滥用热及自产热阶段电行为

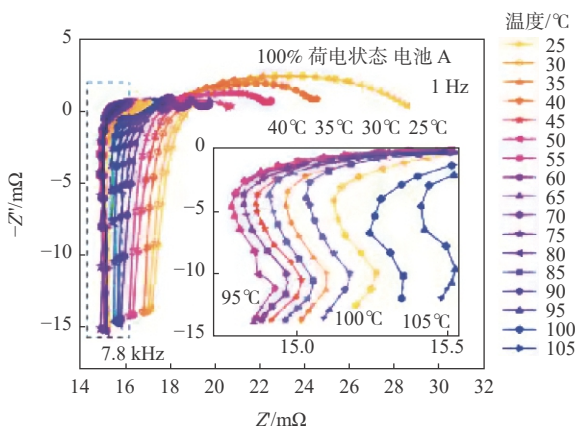
滥用热及锂离子电池的自产热会改变电池内部材料特性, 进而影响电行为, 最终造成电行为的时序性异常。Wang 等[13] 研究三元电池高温存储时长与电性能衰减的关系, 发现在 80 ℃ 环境存储 240 h、720 h 会引发石墨负极材料从铜箔上脱落, 导致电阻显著增加、正极晶格结构受损、容量大幅衰减, 如图 1(a)所示, 直流电阻分别提高 42.1%、65.6%, 如图 1(b)所示, 容量损失 0.9%、5.8%。Sun 等[14] 研究 70 ℃ 高温环境下电阻的降解机理, 指出电解液渗入负极产生的晶间裂纹使 SEI 膜厚度增加、面积扩大, 导致电阻 R_{SEI} 骤增。Ding 等[15] 结合应力-应变曲线、晶体结构表征及形貌观察等手段分析高温下电阻微观变化, 发现 90 ℃ 时隔膜大量孔隙关闭且随温度升高隔膜出现软化现象, 进而揭示高温可直接作用于隔膜使锂离子电池内阻增加及电流均一性恶化。Mendoza-Hernandez 等[16] 研究自产热阶段不同荷电状态(state of charge, SOC)下钴酸锂电池直流电阻的变化幅值, 由于自产热导致隔膜出现闭孔, 阻碍锂离子的扩散与脱嵌, 在 135 ℃ 时, 各 SOC 下的直流电阻均提高 1 000 倍左右。Li 等[17] 监测不同 SOC 三元电池滥用热造成的交流阻抗波动情况, 发现 100% SOC、50% SOC 三元电池在 25～90 ℃ 期间, 随温度增高电池界面阻抗降低导致电化学阻抗降低, 电化学阻抗呈下降趋势, 如图 1(c)所示, 图中, $-Z''$为阻抗虚部, Z'为阻抗实部。在 90 ～105 ℃ 时, 随温度进一步升高隔膜出现闭孔、电解液发生汽化, 进而导致电化学阻抗升高, 电化学阻抗呈上升趋势, 如图 1(d)所示。可见, 受滥用热和自产热影响, 三元电池电阻呈现先减小后上升的趋势。



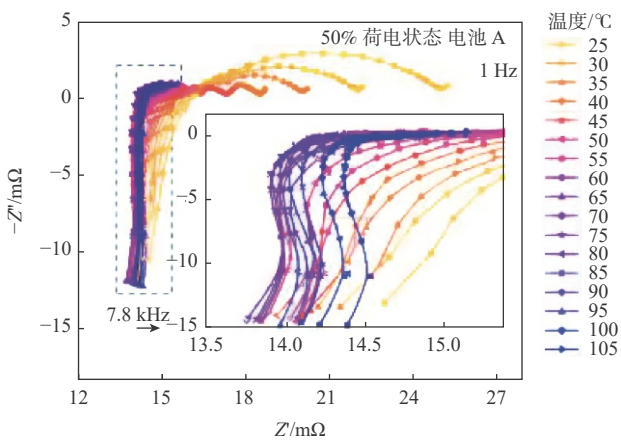
(a) 高温存储时长与电阻增幅关系曲线[13]



(b) 高温存储时长与容量衰减关系曲线[13]



(c) 100% SOC 各温度梯度下 Nyquist 曲线[17]



(d) 50% SOC 各温度梯度下 Nyquist 曲线[17]

图 1　滥用热及自产热阶段电行为变化

Fig. 1　Changes in electrical behavior during heat abuse and self-generated heat stages

学者研究发现, 滥用热及锂离子电池自产热与电压之间呈负相关。Liu 等[18] 研究了三元电池在绝热环境中温度对电压的阶段性影响, 发现因温升导致正极结构发生微观变化进而引发电压出现较小波动, 温度从 25 ℃ 升至 76.3 ℃, 电压未下降, 温度升至 137.4 ℃ 时, 电压下降 4.8% 左右, 温度升至

149.9 ℃时，电压下降10%左右。Li等[19]探究多低气压梯度下温升对电压的影响，也发现温升导致SEI分解与重塑进而诱发电压波动。Zhou等[20]分别研究滥用热及自产热作用下电压的波动幅度，滥用热导致电压下降11%，自产热导致电压下降16%，因自产热导致的锂离子活性物质减少是造成电压大幅下降的主要因素。Gao等[21]分别采用单侧加热与双侧加热的方式量化研究电压的陡降幅度，锂离子电池单侧加热因热失控周期较长不能快速引发内部热扩散，电压下降28%后出现回弹现象。而双侧加热因热失控周期缩短，电压无明显变动。Zhu等[22]研究热滥用对磷酸铁锂电池电压的影响，当温度从19.2 ℃升至75.1 ℃时，电压未发生改变。而当温度从75.1 ℃升至122.6 ℃时，因高温导致电池内阻减小、电压微幅上升，电压增长0.6%。Qi等[23]研究了绝热环境下三元电池电压的变化规律，发现锂离子电池自产热导致隔膜出现轻度熔化，造成电极层间压力不均，电压由4.051 V瞬时下降2.551 V后，再迅速恢复并维持3.972 V，直至隔膜熔化大面积短路电压骤降为0 V。可见，滥用热及锂离子电池自产热初期对锂离子电池电压影响不明显，但总体呈下降趋势。

1.2 过充、过放及充放电倍率产热行为

过充、过放是极端的电行为，耦合锂离子电池电量、电压会产生异常的热行为。过充则是通过诱发电解液发生氧化反应，导致电池表面温度开始缓慢上升[24]。随着过充程度加深，SEI膜将逐渐增厚，正极也因异常脱锂导致结构坍塌，进而加剧电池内部副反应产热[25]。同时，Ji等[26]还指出相较于小倍率过充，大倍率过充存在较大的极化现象，可诱发大量极化热，提升热失控孕育阶段风险指数。过放的产热机制与过充不同，过放可使SEI膜发生分解，进而逐步恶化电池内部的产热行为[27]。在深度过放作用下，正极表面铜沉积增厚导致离子扩散阻抗增大，进一步提高电池内部的产热速率，且产热速率与过放深度呈正相关[28]。此外，锂离子电池在充放电过程中伴随着可逆热、欧姆热和极化热，且随放电倍率增大产热加剧[29]。为定量评估各类电行为在锂离子电池产热层面的作用强度，表1总结了近几年在过充、过放及充放电倍率产热行为的研究进展。

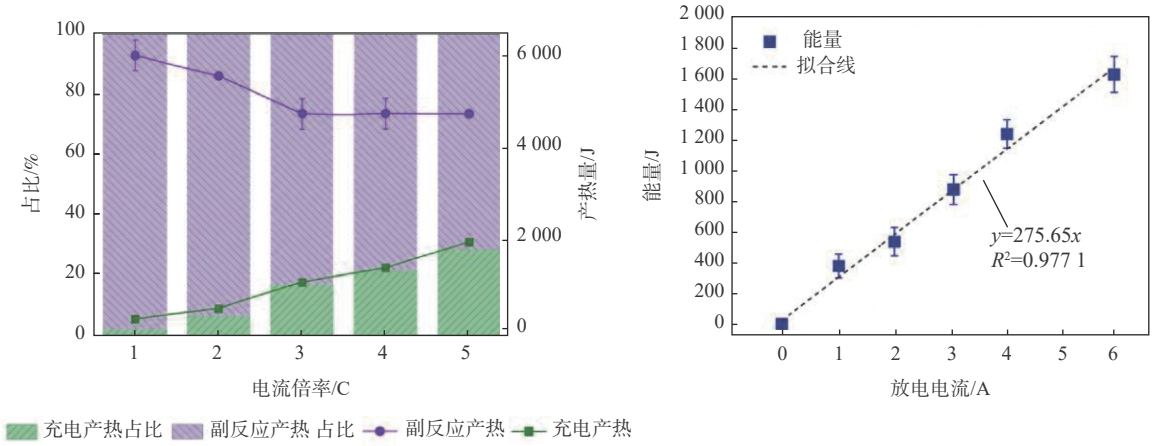
表1 过充、过放及充放电倍率产热行为的研究情况

Table 1 Research on thermal behavior of overcharging, overdischarging and charge/discharge rates

文献	型号	电行为	充放电倍率/C	截止电压/容量/V	温升/℃
Wang等[30]	NCM523	过充	0.5	5.292	19.80
Liu等[31]	$LiMn_2O_4$	过充	0.5	5.184	17.92
Zhang等[32]	LFP	过充	0.5	5.218	18.15
Ouyang等[28]	$LiCoO_2$	过放	1.0	0.200	9.30
Ouyang等[33]	NCM523	过放	1.0	0.200	6.70
Wang等[34]	NCM523	放电	0.5	2.750	3.30
Liu等[35]	$LiCoO_2$	放电	1.0	4.450	16.50
Panchal等[36]	LFP	放电	3.0	2.000	16.80
Dong等[37]	NCM523	充电	3.0	4.200	15.00

通过对比表1中各类电行为的产热数据，发现过充对锂离子电池温升的影响最为显著，其次是大倍率充放电，而过放对温升的影响相对较小，其主要是由于过充在破坏电池内部材料结构的同时，还可长时间积聚热量，大倍率充放电则可加剧欧姆热及极化热。此外，相较于单一充放电的产热行为，锂离子电池在热滥用与充放电耦合作用下，因其隔膜会在高温环境下出现闭孔现象，进而阻碍离子的迁移，加剧充放电时的产热[38]。为此，Hu等[39-40]分别量化研究热滥用耦合放电倍率的产热量及热滥用耦合充电倍率的产热量，发现放电倍率为7 C时，因欧姆热呈指数型上升，电化热占整个热失控总热量的40%。充电倍率为5 C时，因可诱发较大的极化热，电化热占比为28.1%，结果如图2(a)所示。Li等[41]还研究了18650电池在热滥用与放电电流耦合作用下的产热特征，发现当放电倍率为6 A时，电池产热1.6 kJ，池体表面温度提升34 ℃，结果如图2(b)所示。

综上所述，电能与热能的耦合灾变是锂离子电池热失控前的重要表征。滥用热及自产热诱发SEI膜的分解与重塑、负极产气正极释氧及电解液的分解等连锁放热反应，进而导致电压在一定范围内波动，电阻先减小后增大，单体电池容量损失及系统电气失衡等现象；而过充、过放及充放电倍率的变化，会恶化离子的迁移速度和电池内部电化学反应的强度，锂离子电池会伴随可逆热和极化热的大幅增加，进而加剧滥用热和自产热锂离子电池的产热行为。



(a) 各充电倍率在自产热阶段产热占比[40]　　(b) 各放电电流在自产热阶段产热量[41]

图 2　热滥用与充放电倍率耦合作用下的电化热占比

Fig. 2　Electrochemical heat share for thermal abuse coupled with charge/discharge multiplicity

2　隔膜熔化阶段电热耦合

锂离子电池模组中单体电池自产热恶化趋势若未得到有效缓解, 必然会引发电池内部结构进一步衰变, 进而引发一系列更剧烈的放热反应, 最终导致电池热失控[42]。学者们聚焦锂离子电池焦耳热及氧化还原反应热与内短路电行为、过热部位电行为, 以及动态过充与产热[43]、静态过充与产热[44]、静态过放与产热[45] 及内短路的焦耳热等主题, 较深入地研究了热行为与电行为的关系。

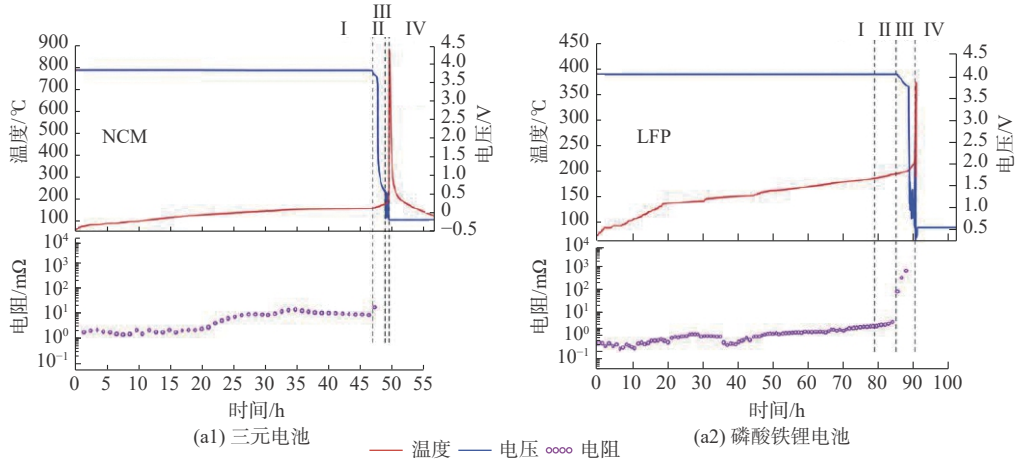
2.1　焦耳热及氧化还原热阶段电行为

锂离子电池内短路及热失控时伴随焦耳热及正负极氧化还原热, 上述 2 种热量会导致锂离子电池电行为呈现巨大变化, Ren 等[46] 量化研究了三元电池焦耳热产生后电行为的变动幅度, 发现因隔膜熔化导致正负极短接进而产生大量焦耳热, 电压瞬态陡降为 0 V、电阻增大了 700 倍。Tang 等[47] 分析了三元电池热失控时电性能的变化, 发现由于高温破坏隔膜微孔降低锂离子的传输速率, 进而导致电池内阻急剧增加。当锂离子电池温度升至 215.24 ℃时, 电池内部发生短路, 电压骤降为 0 V, 且电阻增大了 850 倍。Yang 等[48] 研究绝热环境下磷酸铁锂电池焦耳热对电阻的异常影响, 内短路导致正负极短接诱发焦耳热, 致使电阻瞬时增大 1 000 倍左右。

Tang 等[49] 利用液氮急冻热失控的磷酸铁锂电池, 拆解电池内部材料, 通过扫描电镜表征, 指出热失控时刻隔膜熔化与正极发生黏结, 电解液燃烧消耗造成电阻骤增。Zheng 等[50] 对比三元与磷酸铁锂电池热失控时电行为的异常变动如图 3(a)所示, 发现三元电池温度升至 173.4 ℃、磷酸铁锂电池温度升至 212.2 ℃时, 电压骤降为 0 V, 三元电池电阻增至 1.3 Ω, 磷酸铁锂电池电阻增至 1.2 Ω。其差异的原因在于三元电池高温导致内部化学反应剧烈, 并有大量氧气释放, 隔膜出现熔化, 低沸点溶剂蒸发, 更易诱发热失控。Huang 等[51] 研究锂离子电池受热部位差异与电行为的关系, 发现底部、侧面受热时, 其电压下降时刻与热失控触发时刻同步。电池大面受热时, 压降时刻晚于热失控触发时刻。这是因为大面受热沿厚度方向存在较大的温度梯度, 发生短路的电芯数量有限, 延缓电压的瞬时下降。此外, Liu 等[52] 研究使用聚对苯二甲酸乙二醇酯(polyethylene terephthalate, PET)/陶瓷隔膜的三元电池热失控、内短路时刻温度与电行为的关系, 量化分析了 2 个时刻对应的电压、电阻的变动幅值, 表征了 PET/陶瓷隔膜电池在各温度梯度对电性能的作用机制, 发现当温度升至 231 ℃时, 电池虽发生热失控但未发生内短路, 此时电阻增长 4.3 倍、电压维持在 2 V 以上。当温度升至 700 ℃时, 电池发生内短路, 此时电压骤降至 0 V, 电阻增长 106.6 倍, 如图 3(b)所示。因为 PET/陶瓷隔膜耐热性能优异, 热失控时 231 ℃的高温仅使此隔膜出现微弱收缩, 电压电阻变动较小。当温度升至 450 ℃, PET/陶瓷降解隔膜大面积塌缩, 电性能受损严重, 如图 3(c)所示。

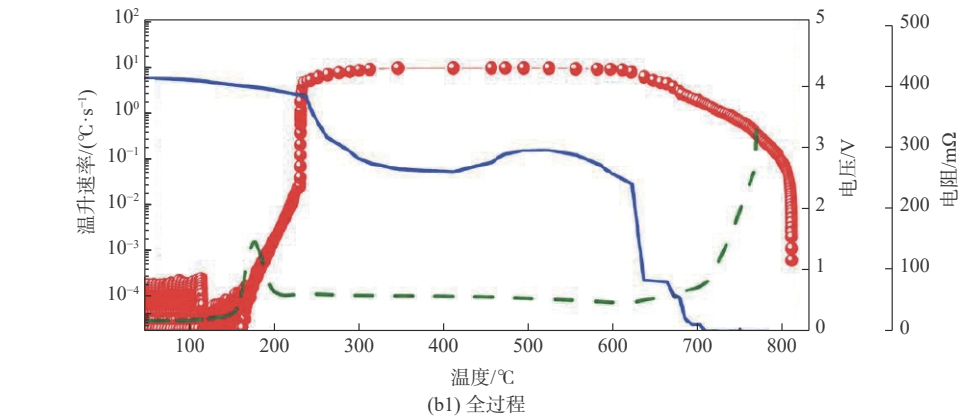
2.2　过充、过放及内短路产热行为

过充、过放电行为破坏正负极结构, 加快正极分解及正负极氧化还原反应, 进而影响热失控触发温度与峰值温度。Mao 等[53] 量化研究静态过充下正负极产热占比, 过充至 120% SOC 时, 大部分热量来自正极, 因为过充导致正极晶体结构发生改变。过充至 140% SOC 时, 大部分热量来自负极, 因为过充导致负极表面锂沉积与电解质反应。同时, Mao 等[54] 还研究了不同倍率下动态过充对三元电池热失控温度的影响, 如图 4(a)所示, 发现当充

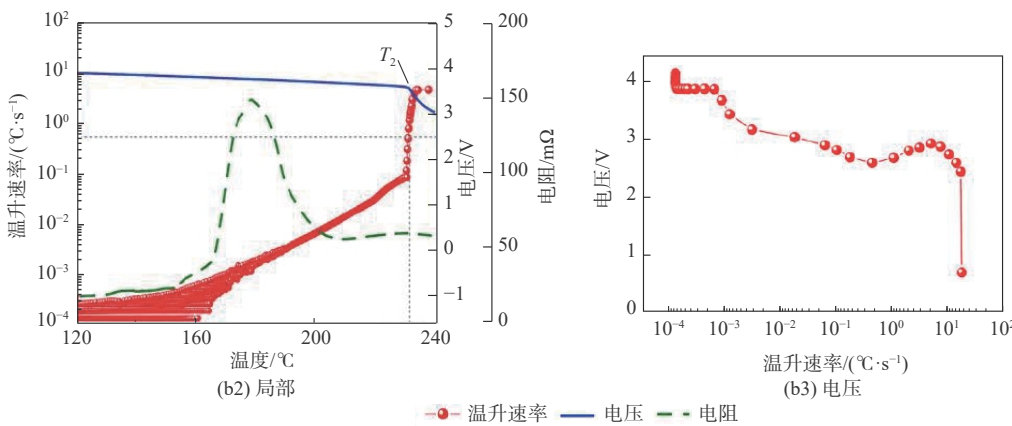


(a1) 三元电池 (a2) 磷酸铁锂电池

(a) 三元电池、磷酸铁锂电池的电压及电阻变化曲线[50]

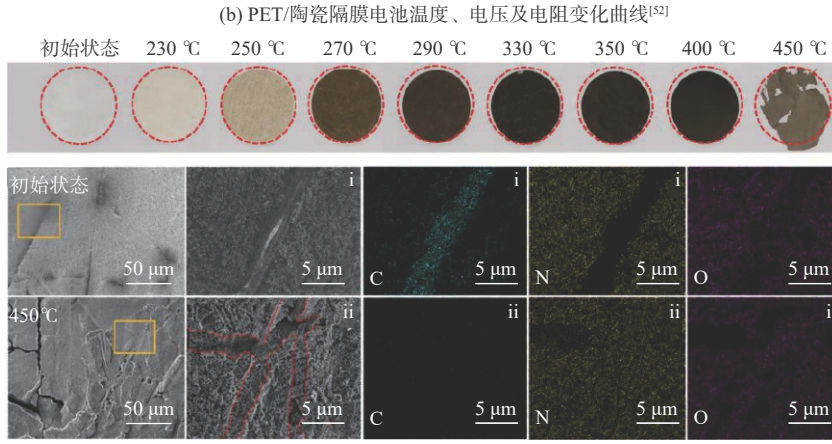


(b1) 全过程



(b2) 局部 (b3) 电压

(b) PET/陶瓷隔膜电池温度、电压及电阻变化曲线[52]



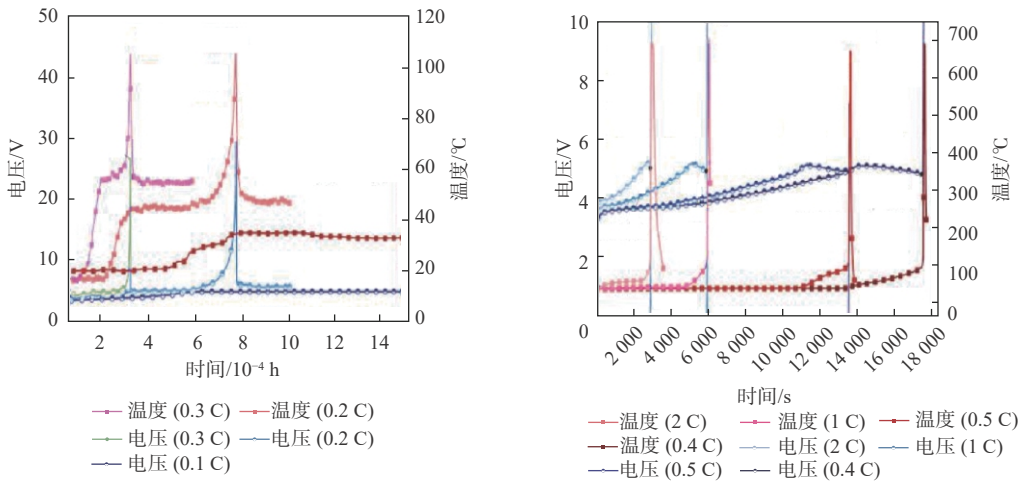
(c) PET/陶瓷隔膜各温度梯度下形貌表征[52]

图 3 焦耳热及氧化还原热阶段电行为变化

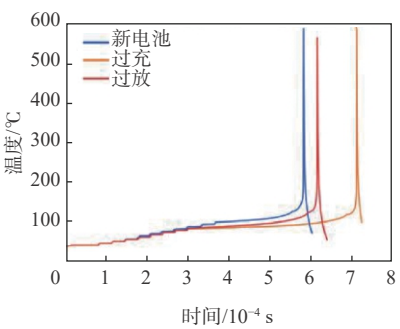
Fig. 3 Changes in electrical behavior during the Joule heat and redox heat stages

电倍率＜0.3 C 时, 电池温度最高升至 115.5 ℃, 未触发热失控。当充电倍率＞0.3 C 时, 由于大倍率充电引入大量可逆热, 进而提高电池温升速率, 触发热失控。Zhang 等[55] 对比静态过充电池和新电池的温升变化, 如图 4(b)所示, 相较于新电池, 轻微过充电池热失控触发温度降低了 1.6 ℃, 峰值温度提高了 1.7 ℃, 最大温升速率下降了 229 ℃/s, 指出轻微过充可加剧负极表面锂沉积现象, 进而加速副反应进程, 如图 4(c)所示。过放亦可破坏电池内部形貌结构, 进而作用于热失控阶段, Li 等[56] 研究过放程度对三元电池温升的影响, 发现过放至 1.5 V 比过放至 0.5 V, 热失控触发温度高出 1.4 ℃, 即深度过放更易触发热失控, 因为深度过放导致正极表面出现铜沉积现象, 如图 4(c)所示, 引发更多副反应, 从而降低结构稳定性。

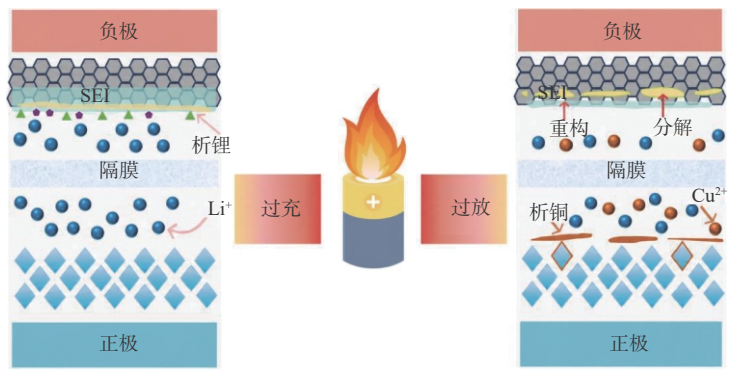
锂离子电池内短路所产生的焦耳热可直接诱发电解液与正负极、正极与负极间无障碍的放热反应, Liu 等[57] 研究由形状记忆合金触发三元电池各层级间内短路所引发的焦耳热大小, 如图 4(d)所示, 因短路热量主要为焦耳热及欧姆热, 发现在铝



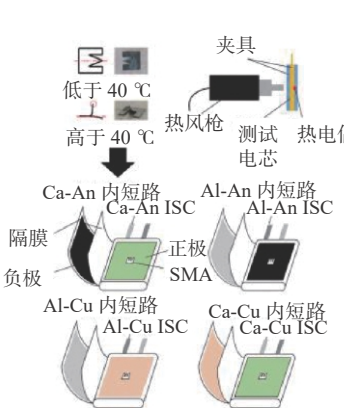
(a) 不同倍率过充下的温升曲线[54]



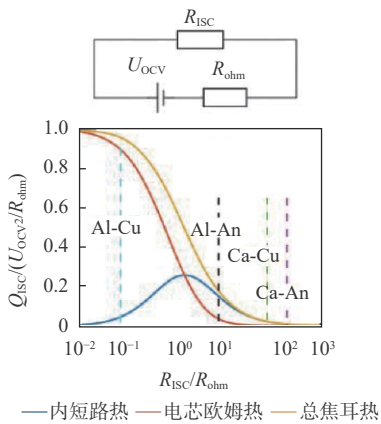
(b) 静态过充电池和新电池的温升曲线[55]



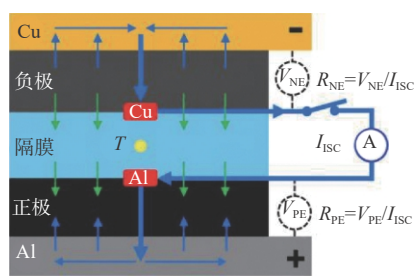
(c) 过充、过放电池内部界面反应



(d) 诱发内短路原理示意图[57]



(e) 各层级间焦耳热与欧姆热的比值[57]



(f) 基于外电路测量短路电流原理示意图[58]

图 4　过充、过放及内短路的产热行为及机理

Fig. 4　Heat generation behavior and mechanism of overcharge, overdischarge and internal short circuit

集流体和负极之间短路，该点短路电阻与欧姆电阻近似，焦耳热达到峰值，结果图4(e)所示，再因短路时良好的导电性及恶劣的导热性导致层级间焦耳热恶化，焦耳热比欧姆热更易堆积热量。同时，Long等[58]在正负极间插入铝铜垫片与外部电流构成回路，如图4(f)所示，原位测量三元电池铝集流体和负极间内短路电流及焦耳热功率大小，发现内短路电流瞬时升至70 A，且焦耳热功率为160 W。Ma等[59]基于短路电气模型分析各阶段焦耳热的成因及范围，发现当短路电阻＜初始电阻，初始电阻产生的欧姆热致电池表面温度均匀上升。当短路电阻=初始电阻，内短路处温度达到1 165 K峰值。当短路电阻＞初始电阻，短路电阻产生的焦耳热致使电池表面温度局部上升。Xin等[60]分别研究磷酸铁锂电池及三元电池焦耳热与热失控的关系，发现三元电池的焦耳热可骤升电池整体温度，随即触发热失控。磷酸铁锂电池的焦耳热只能提升短路处的温度不会立即引发热失控。

综上所述，热量主要作用于隔膜，进而触发一系列热灾害末端的放热反应，导致热失控阶段电压骤降为0 V，电阻骤升700～1 000倍。同时，电性能尽失伴随内部焦耳热的产生，且该热量与短路面积及电池容量呈正相关；过充将导致负极表面出现锂沉积，而过放将导致正极表面出现铜沉积，两者在一定程度上均可破坏正负极及隔膜结构的稳定性，直观表现为热失控触发温度降低，但两者在活性物质的增减层面存在差异，进而导致热失控峰值温度相差较大。当内短路触发位置位于铝集流体-负极层间且短路电阻等于初始电阻时，所产生的焦耳热最具威胁。

3 模组热扩散阶段电热耦合

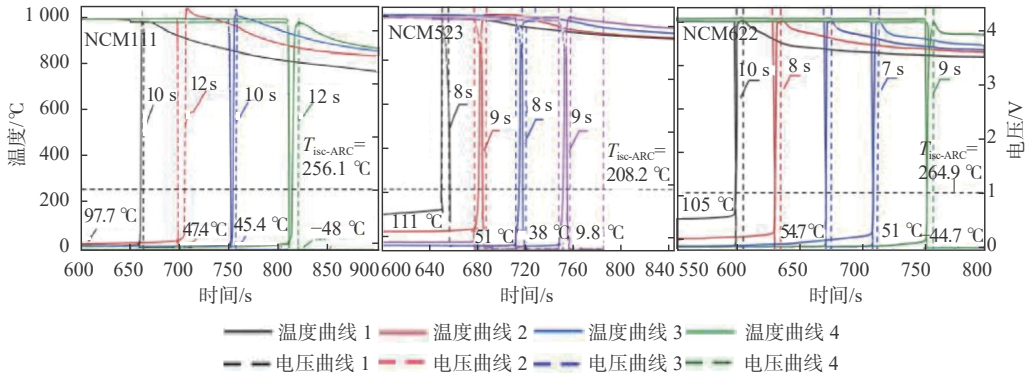
锂离子电池模组中单体电池发生热失控，所产热量若未有效处置，将引发相邻电池进入热失控状态，进而造成模组热扩散现象[61]。锂离子电池模组在工程应用中，通常以串联、并联及串并联的方式组成[62]，瞬时高温致使整个模组电性能受损。学者们聚焦于热扩散阶段各类模组电压、过热短路电流、老化梯度传热、动态过充传热及短路电流转化热等主题，从时间序列和空间尺度研究了其热行为与电行为的关系。

3.1 热扩散阶段的电行为

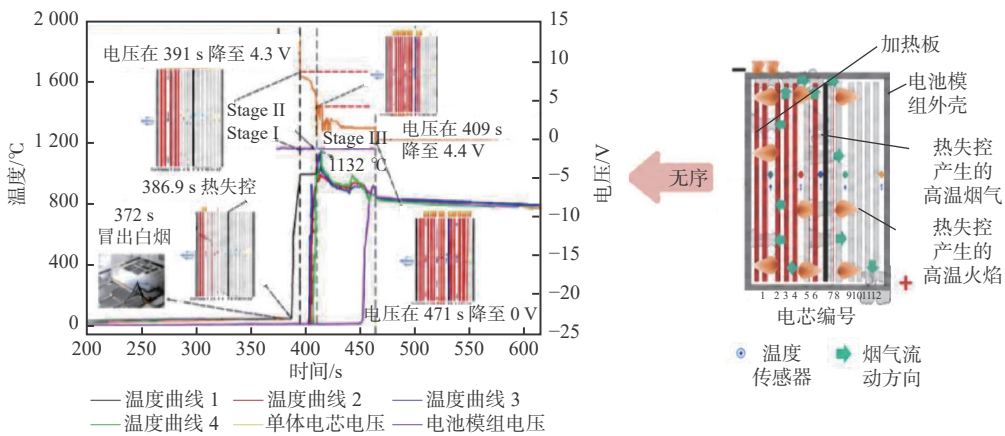
模组热扩散对电行为的影响极为显著，Wang等[63]对比研究NCM523、NCM622、NCM111这3类模组发生热扩散时内短路与热失控触发的时间间隔(t_{int1}～t_{int4})，T_{isc1}～T_{isc4}分别为其内短路温度，如图5(a)所示，发现电池正极镍含量越高，材料活性越高，正极极易分解产氧迫使隔膜熔化，同时，加剧正负极氧化还原反应的进程，致使NCM622模组和NCM523模组内短路与热失控触发的时间间隔比NCM111模组内短路与热失控触发的时间间隔缩短了3 s左右。Li等[64]将4并3串的三元电池模组发生热扩散时压降划分为3个阶段(Stage Ⅰ～Stage Ⅲ)，分析模组热扩散后期热量积聚引发射流火，导致扩散呈无序性，进而延长压降所需的时间，如图5(b)所示，第2阶段和第3阶段压降所需时间分别为18 s和64 s，远大于第1阶段压降所需时间。Schöberl等[65]研究未进行电连接的磷酸铁锂模组热扩散时各单体电池电压的波幅，如图5(c)所示，发现所有电池电压波幅一致，并将单体电池电性能受损过程划分为3个阶段。由于大容量磷酸铁锂电池内部由2个电芯并联而成，第1个电芯热失控时，电池内部产生较大的短路电流，导致电压由3.65 V下降至3 V。第1个电芯热失控后，电阻升高了0.1～2 Ω，内部短路电流减小，电压由3 V恢复至3.65 V。第2个电芯热失控时，电池电性能完全受损电压骤降至0 V。Li等[66]研究6串1并的NCM622模组热扩散时模组整体的压降趋势，随着热扩散的有序进行，模组共出现5个压降台阶。前4个台阶的压降幅度均为4.2 V左右，第5个台阶的压降幅度在8.3 V左右。串联模组由于存在相邻电池热传导、火焰热辐射等多种传热路径，其热扩散后期热量急剧陡升，因此，第5个台阶出现第5、6电池同时短路，不同于第1～4个电池次第短路现象。此外，Liu等[67]量化研究6并的NCM111模组热扩散时内部短路电流的大小，发现并联模组热扩散时存在环流现象，引发未热失控电池瞬时放电，已短路电池瞬时充电，模组内最大的短路电流为25.7 A，且短路电池的瞬时充电容量为11 mAh，进而得出短路电流大小与未热失控电池的数量及容量呈正相关的结论。为控制模组热扩散的作用范围，通常会在模组内设计主动隔热以阻断热失控传播，Zhou等[68]对比了有无气凝胶阻隔下4并的磷酸铁锂模组热扩散时的压降幅度，如图5(d)所示，发现无气凝胶阻隔下，模组内4块电池均触发热失控，且模组电压骤降为0 V。加入气凝胶阻隔后，仅靠近热源的电池触发热失控，因其他电池未受到损伤，且模组内短路电流持续时间也较短，最终导致模组电压可维持在3.33 V。

3.2 老化、过充及内短路传热行为

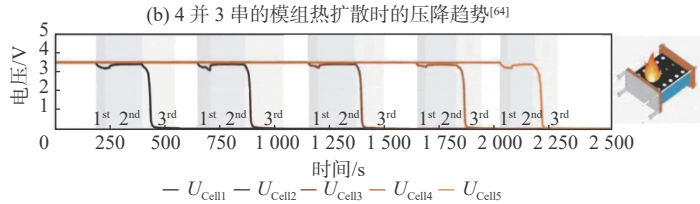
锂离子电池模组过度老化及过充，一定程度上加剧热扩散行为。Zhao等[69]量化不同老化程度下三元串联模组、并联模组及串并联模组电池间的平



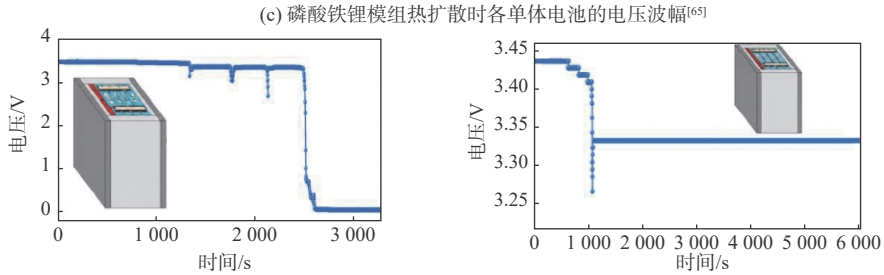
(a) NCM111、NCM523、NCM622模组内短路时刻与热失控触发时刻的时间间隔[63]



(b) 4 并 3 串的模组热扩散时的压降趋势[64]



(c) 磷酸铁锂模组热扩散时各单体电池的电压波幅[65]



(d1) 无气凝胶　　(d2) 有气凝胶

(d) 有无气凝胶阻隔的并联模组热扩散时的电压波幅[68]

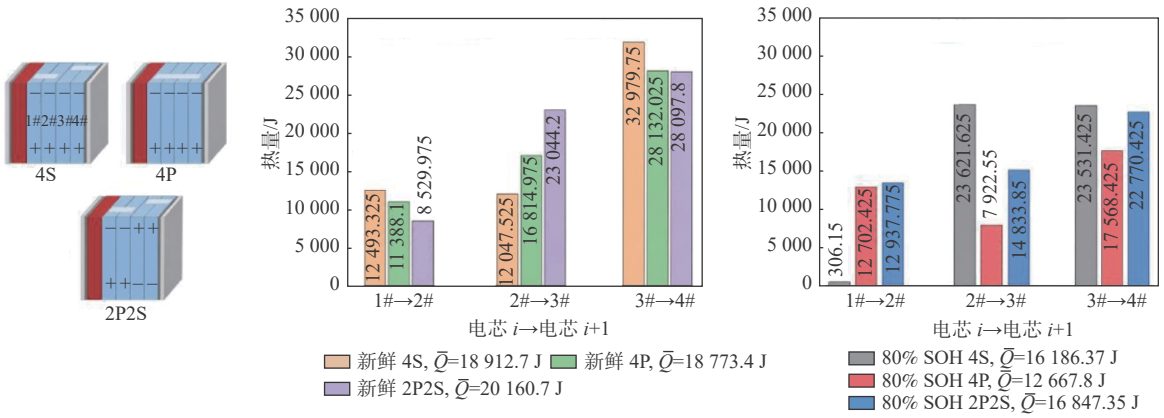
图 5　模组热扩散阶段电压变化

Fig. 5　Voltage change during the thermal runaway propagation

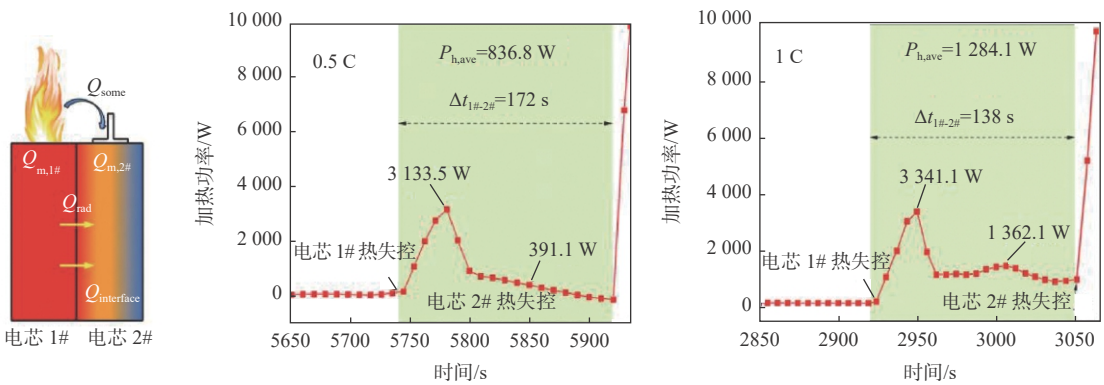
均传热, 因为老化过程所形成的锂枝晶影响电池内部结构的稳定性, 导致热失控孕育阶段温升显著, 致使触发热失控所需的能量减少, 结果如图 6(a)所示, 发现 100% 健康状态(state of health, SOH)的串联模组, 电池间热量传递的平均值为 18 912.7 J, 100% SOH 的并联模组, 电池间热量传递的平均值为 18 773.4 J, 100% SOH 的串并联模组, 电池间热量传递的平均值为 20 160.7 J。而 80% SOH 的串联模组, 电池间热量传递的平均值为 16 186.37 J, 80% SOH 的并联模组, 电池间热量传递的平均值为 12 667.8 J, 80% SOH 的串并联模组, 电池间热量传递的平均值为 16 847.35 J。同时, Han 等[70] 也量化不同老化程度下未进行电连接的三元模组电池间的平均传热, 发现 90% SOH 的模组比 100% SOH 的模组电池间热量传递的平均值提高 2 106 J, 指出老化程度是威胁模组安全性的重要因素。对比以上两位学者的研究结果, Zhao 等[69] 发现电池间热量传递的平均值与老化程度呈负相关, 而 Han 等[70] 发现电池间热量传递的平均值与老化程度呈正相关, 其主要原因是 Zhao 等[69] 所研究的对象为串并联模

组，热扩散时存在多条传热路径及短路产热，进而缩短热扩散时长，并减少电池间热量传递的功率。Huang 等[71]量化研究未进行电连接的三元模组在不同倍率下过充电池间的热扩散功率，因为大倍率过充加剧正负极氧化还原反应，导致热失控峰值温度骤升，模组温均性恶化，如图 6(b)所示，0.5 C 过充时平均热扩散功率为 836.8 W，热扩散时长为 172 s，1 C 过充时平均热扩散功率为 1 284.1 W，热扩散时长为 138 s。此外，Liu 等[72]量化研究 4 并的磷酸铁锂模组中短路所引发的电化热大小，由于模组内的瞬时短路电流可迅速转化为焦耳热，模组内部的电化热功率维持在 600～900 W。Zhou 等[73]量化研究并联模组内短路时焦耳热的大小，发现单体电池短路时，焦耳热的介入可加剧负极与电解质反应及正负极氧化还原反应，并提高电池内部热量的扩散速率如图 6(c)所示，当焦耳热为 29.5 kJ 时，短路电池极耳附近处的温度可提高 50 ℃ 以上。

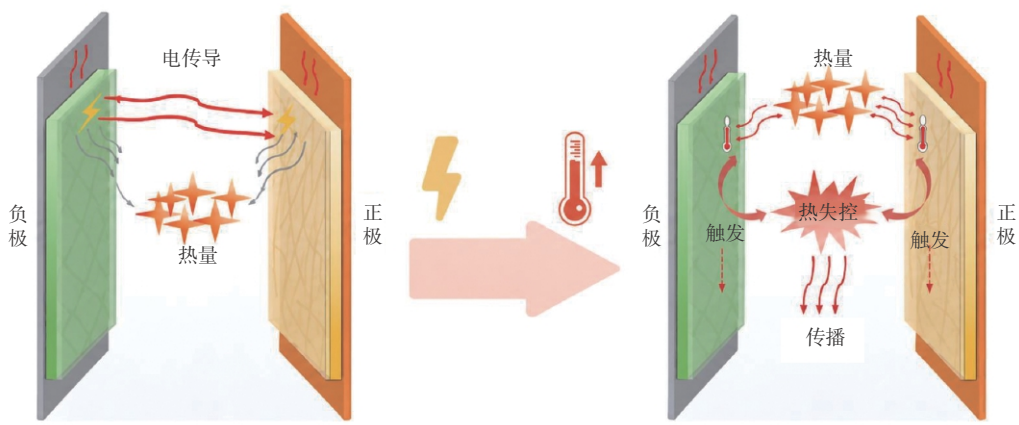
综上所述，锂离子电池模组热扩散主要通过有序或无序传播形成隔膜熔化临界温度点，锂离子电池模组整体电性能因而受到严重影响。串联模组



(a) 各类模组热失控内部传热量[69]



(b) 过充模组热失控内部热扩散功率[71]



(c) 并联模组中短路电池内部热扩散机理[73]

图 6 老化、过充及内短路的传热行为

Fig. 6 Heat transfer behavior of aging, overcharging and internal short circuit

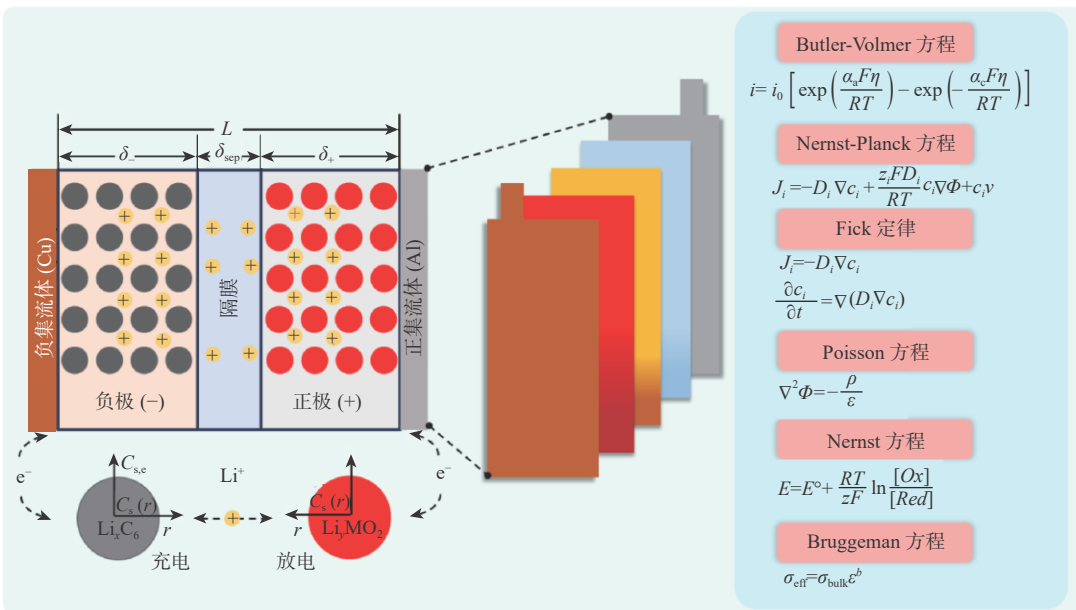
电压呈阶梯式下降, 并联模组电压呈现骤降骤升情形, 其电量、短路电流呈现整体下降趋势; 锂离子电池老化所生成的锂枝晶致使模组在运行时存在较大的温度梯度, 进而降低模组热稳定性。锂离子电池串联模组过充将诱发焦耳热[74], 并联模组内短路瞬时过充亦可诱发焦耳热。

4　多尺度电-热耦合模型

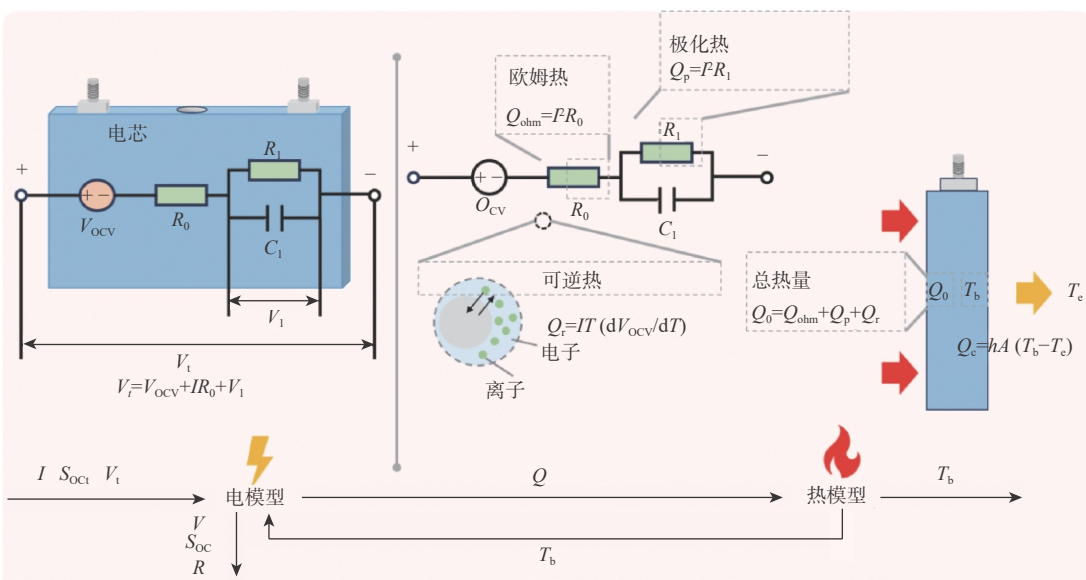
锂离子电池模组热灾害各阶段、各维度电热耦合灾变进程是近年研究重点。学者们分别从微观尺度和宏观尺度建立电行为和热行为的映射关系, 聚焦单体内部界面热源分布、单体老化诱因鉴定、全周期老化温升估算、模组短路产热量化、模组充电策略优化等, 构建有针对性的电-热耦合模型, 较全面地研究了电热耦合机理及能量转化机制。

4.1　单体电-热耦合模型

锂离子电池模组热灾害电热耦合最早发端于单体电池, 因此, 单体电池电热耦合研究是近年来的焦点之一。Newman 等[75] 基于 Butler-Volmer 方程、Nernst-Planck 方程及 Fick 定律等理论构建电化学模型, 如图 7(a)所示, 揭示锂离子的电极反应动力学、液相扩散过程及固相扩散过程等, 为探索充放电状态下电池内部产热反应奠定基础, 图中, $C_{s,e}$ 为固相锂电子浓度, r 为粒子半径。在此基础上, Doyle 等[76-77] 引入热变量改进前述的电化学模型, 建立热效应和电化学的耦合机制, 并研发了电



(a) 电化学模型



(b) Bernardi 产热模型

图 7　锂离子电池电-热耦合基础模型

Fig. 7　Basic model of electro-thermal coupling in lithium-ion battery

化学-热模型，可较好地揭示电池热源生成与分布、性能与寿命预测等问题。同时，依据电化学模型输出的SOC及终端电压，Bernardi等[78]提出了产热方程量化研究电池总产热，如图7(b)所示，图中，V_{ocv}为开路电压，R_0为欧姆内阻，R_1为极化内阻。可见，电行为与热行为之间存在相互影响、相互转化的定量模型关系[79]。

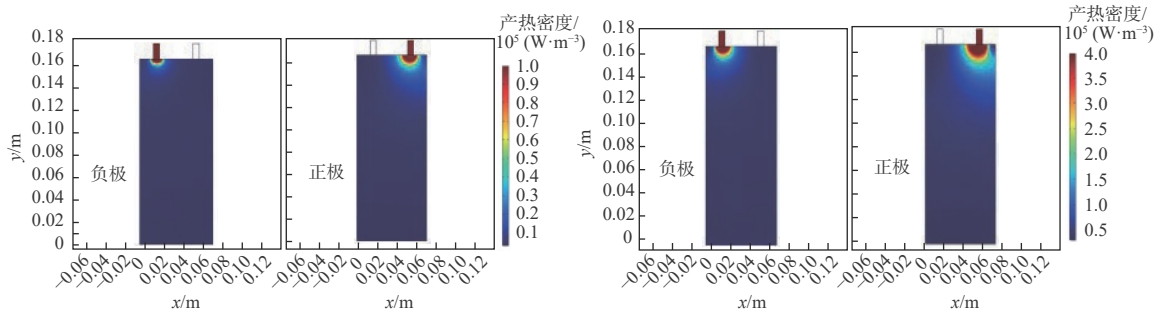
在单体电池产热分布层面，Ren等[80]基于COMSOL Multiphysics构建电化学-热模型量化研究放电正负集流体的产热密度，铝的电导率低于铜，因此，铝质正集流体聚集大量欧姆热，在1 C和3 C放电时，负集流体的最大产热密度分别为1.22×10^5 W/m^{3}和1.06×10^6 W/m^3，正集流体的最大产热密度分别为3.04×10^5 W/m^{3}和2.64×10^6 W/m^3，可见正集流体产热密度高出负集流体的2.5倍，结果如图8(a)所示。Lin等[81]构建三维电化学-热模型，如图8(b)所示，对比集总热源和非均匀热源下热量分布，发现集总热源模型下热量主要集中于电池大面中心，非均匀热源模型下热量主要集中于电池大面中心及正极极耳处。可见，集总热源模型虽具备较高的计算效率，但无法解决非均匀温度场引发的电池老化非均匀性的问题，而非均匀热源模型可对电池最高温差进行捕捉，准确反映热源分布。同时，Sarkar等[82]基于三维电化学-热模型研究不同放电倍率下电池内部的主导热源，由于放电倍率提高引发过电位陡升极化现象加剧，而且SOC降低内阻增大诱发大量欧姆热，致使极化热和欧姆热成为主要热源。在电池性能寿命研究层面，Fan等[83]构建电化学-老化-热耦合模型用于分析低温下长期循环的老化轨迹，发现低温环境所诱发的负极镀锂现象是加速电池寿命衰减的主要原因。如图8(c)所示，随着温度降低，循环圈数逐渐减小，−20 ℃低温环境下0.5 C和1 C仅能循环8圈。Xi等[84]也使用电化学-老化-热耦合的多物理场、多尺度模型，研究25 ℃和45 ℃环境下三元电池的循环寿命，发现45 ℃环境将导致极化电位增加镀锂现象加剧。因为25 ℃环境下，SOH下降至60%时循环次数大于2 500圈，因镀锂造成容量损失为44.8%，45 ℃高温环境下，SOH下降至60%时循环次数小于2 000圈，因镀锂造成容量损失为54.8 %。此外，Shen等[85]对电化学-热耦合模型进行改进，融合遗忘因子最小二乘法及平方根立方卡尔曼滤波算法构建的模型，如图8(d)所示，利用该模型对宽温域(−20～40 ℃)、全周期(80% SOH～100% SOH)的电池温度进行估算，所得结果的最大误差小于1.10 ℃，表明该模型具备一定的可行性及鲁棒性。Liu等[86]构建电化学-热耦合内短路模型，如图8(e)所示，对比研究机械滥用、电滥用及热滥用等多种内短路触发机制下的电热耦合行为，经验证，该模型不仅具备精准的内短路电压结果预测能力，而且还可有效预测内短路时的热行为。

4.2 模组电-热耦合模型

并联模组热灾害会引发短路电流现象，其电热耦合是近年来研究的重点之一。并联模组发生热扩散时，会在未短路电池与已短路电池间构成回路，进而产生强大的短路电流。由于连接组件及短路电池等存在电阻，致使短路电流转化为热量[87]，依图9所建立的电-热耦合模型，可对其产热进行量化。

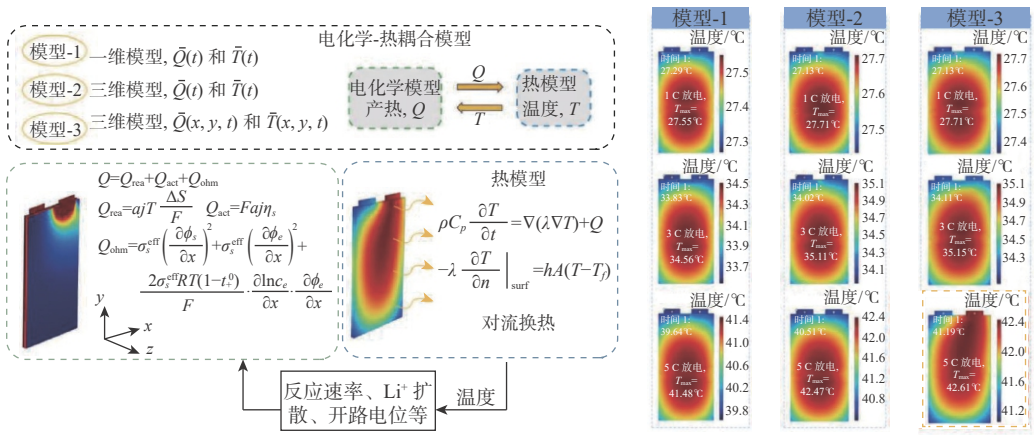
Zhou等[68]基于理想等效电路模型量化400 Ah的并联模组短路时内部焦耳热，发现未短路电池向短路电池传递1.83 Ah的电量，导致短路电池热失控时SOC为103.65%，且该电量可转化为22.43 kJ的焦耳热，致使热失控电池表面温度提高14.6 ℃。Niu等[88]量化研究并联模组发生内短路时未受损电池SOC的波动幅值及所释放的热量，发现在等效电路作用下，未受损电池SOC下降88 %，释放5.1 kJ的热量，占总电能的18%。Xu等[89]构建串并联模组的等效电路模型，如图10(a)所示，基于该模型量化研究3并4串模组热扩散时内部电化热，发现电化热致使池体表面温度提高28.2 ℃，远小于化学反应的峰值温度940 ℃，热量贡献率较小。此外，Gao等[90]以恒定短路电阻为边界条件构建瞬态等效电路模型，如图10(b)所示，基于该模型获取短路电池数量与整体电阻、短路电流及产热量间的拟合曲线，发现短路电流与正常电池SOC呈正相关，与已短路电池数量呈负相关；并联模组因电流分布不均，将滋生镀锂等副反应，威胁模组热安全。为此，Yang等[91]在电化学-热耦合模型基础上制定过电位镀锂充电策略，建立并联模组的电流闭环控制框架，如图10(c)所示，发现使用该策略过电位维持在0 V可有效减弱负极表面镀锂现象。

综上所述，单体电-热耦合模型应用较为广泛的为电化学-热模型。该模型不仅可精准捕捉单体电池内部热源分布及老化轨迹，从内部各界面反应微观揭示电热耦合机理，而且，可与最小二乘法及卡尔曼滤波算法等契合，精准获取多工况、长周期的电热耦合反应。但电化学-热模型因其复杂性及低效性限制，其在电池系统中的应用主要面向运行场景下的机理分析[92]；模组电-热耦合模型中应用较为广泛的为并联模组的等效电路模型，该模型不仅可评估并联模组内部短路产热及剩余电量，而且可分析并联模组充电电流不均衡的问题。但当前的等

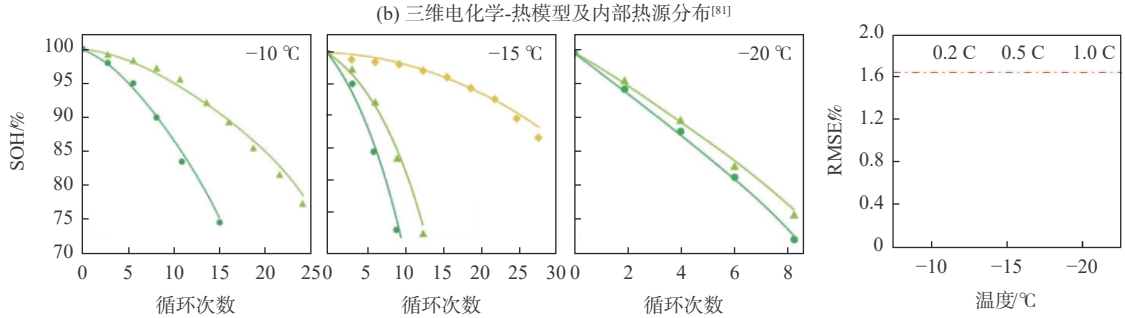


(a1) 1C　(a2) 3C

(a) 1 C 和 3 C 放电正负极流体产热密度[80]



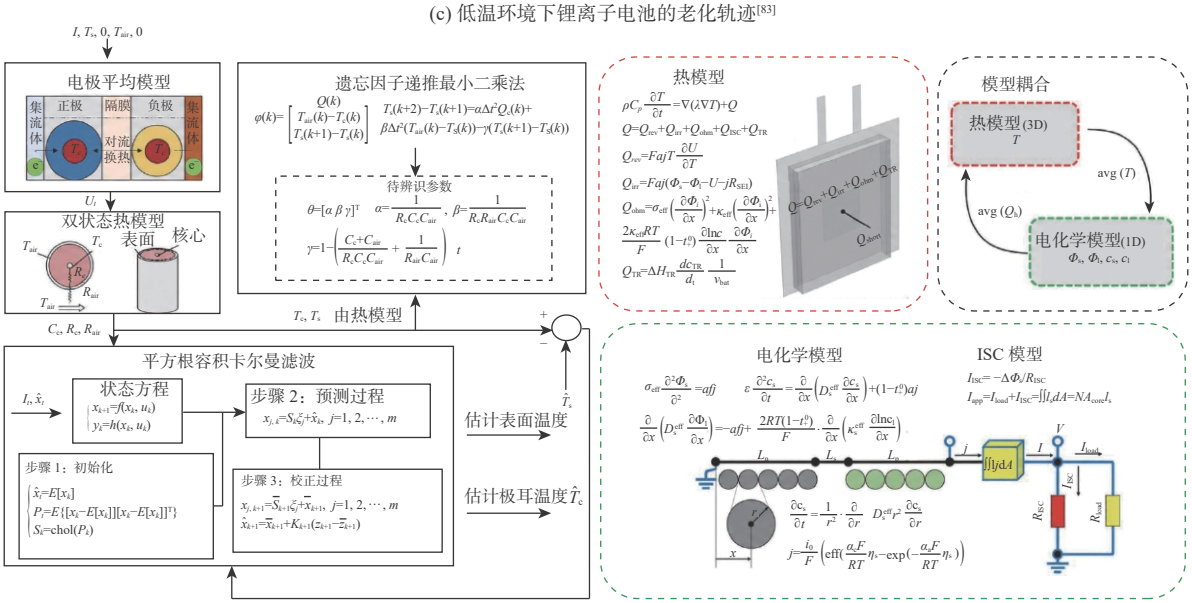
(b) 三维电化学-热模型及内部热源分布[81]



(c1) SoH　(c2) RMSE

实验−0.2 C　实验−0.5 C　实验−1.0 C
仿真−0.2 C　仿真−0.5 C　仿真−1.0 C

(c) 低温环境下锂离子电池的老化轨迹[83]



(d) 耦合 SRCKF 的电化学-热模型[85]　(e) 电化学-热耦合内短路模型[86]

图 8　电化学-热模型

Fig. 8　Electrochemical-thermal modeling

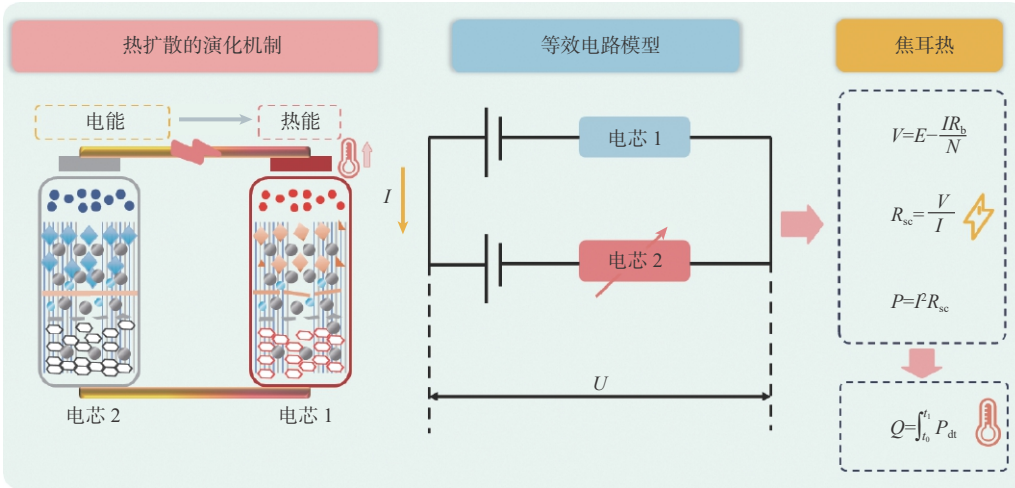
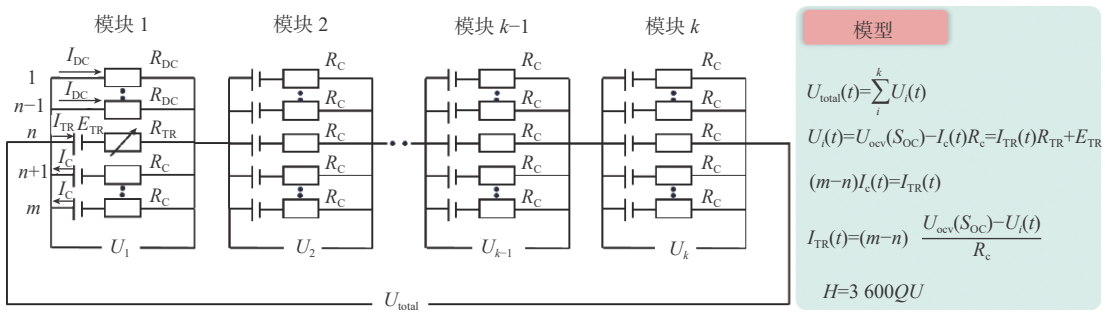
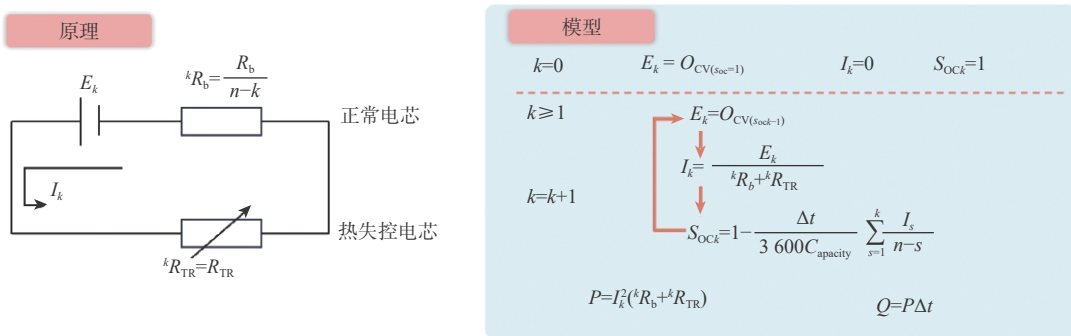


图 9 并联模组电-热耦合模型

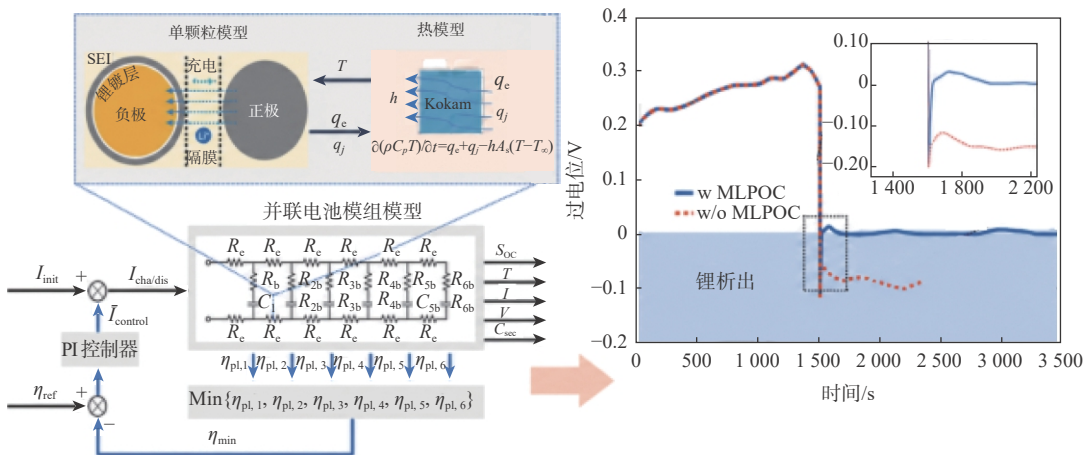
Fig. 9 Electro-thermal coupling model of the parallel module



(a) 串并联模组等效电路模型[89]



(b) 瞬态等效电路模型[90]



(c) 电化学-热耦合过电位镀锂策略的并联模组[91]

图 10 并联模组等效电路模型

Fig. 10 Equivalent circuit model of the parallel module

效电路模型过于理想化, 未能使用短路电流及模组电阻的动态数据, 导致其热量计算存在严重偏差。

5　结 束 语

综上, 锂离子电池从单体到模组各阶段、多维度的电行为与热行为耦合及多尺度电-热耦合模型研究取得了以下 4 个方面的进展:

1) SEI 膜分解使锂离子电池模组进入自产热阶段, 电压、电阻及容量等随即出现衰变, 热量对于电阻的衰变影响尤为显著; 而不同程度的过充、过放等电行为的介入将提升产热量。在此阶段, 标志着电热耦合行为呈现逐步恶化的态势。

2) 隔膜熔化使锂离子电池模组进入内短路阶段, 电压及电量骤降为 0, 电阻升至初始电阻的近 1 000 倍; 而过充、过放及内短路电阻将会降低热失控的触发温度, 提升热失控的峰值温度。在该阶段, 电热之间异常联动尤为显著。

3) 热量传递未阻断使锂离子电池模组进入热扩散阶段, 串并联模组电压总体呈下降趋势, 并联模组短路电流骤升至高达 1 000 A; 而老化、过充及内短路可降低模组内部温均性, 加速热扩散进程。在该阶段, 电热耦合灾变呈大尺度蔓延。

4) 锂离子电池单体的电化学-热模型主要是从电行为及热行为进行二维微观表征, 揭示单体电池内部电热耦合机理; 锂离子电池模组电-热耦合模型主要聚焦并联模组, 采用等效电路模型评价产热及电量, 揭示模组电热动态转换机制。

根据上述研究进展及未来的发展方向, 锂离子电池模组电热耦合行为研究还应从以下几个方面强化:

1) 强化锂离子电池模组运行状态下电热耦合行为研究。当前电热耦合研究多聚焦开路状态下锂离子电池单体及模组温度、电压及交流阻抗, 不能反映运行状态的电热耦合行为。而模组运行状态的温度、电压、电量及唯有的直流电阻才可有效表征运行状态下的电热耦合行为。因此, 亟待强化运行状态下自产热、焦耳热及热传导与电压、电阻及电量之间异常行为研究。

2) 强化热滥用诱导下内短路焦耳热的量化研究。当前焦耳热的量化研究以机械滥用为主, 无法诱导出 SEI 膜分解阶段产热。而热滥用可诱发锂离子电池自产热, 进而引发连锁放热反应, 导致隔膜熔化生成焦耳热, 更贴近锂离子电池模组运行场景。因此, 亟待强化构建热滥用下 SEI 膜分解自产热与内短路焦耳热的函数关系, 量化研究内短路阶段焦耳热的产生位置、影响因素及作用范围。

3) 优化锂离子电池并联模组电热耦合监测指标。当前并联模组的研究多聚焦于电压与温度, 而缺失直流电阻及短路电流, 致使内短路电化热量化研究存在较大的误差, 难以支撑并联模组在工程应用中的风险评估体系。因此, 亟待基于并联模组的电路结构有针对性地引入直流电阻及短路电流监测指标, 开展耦合电压、电流及温度的电热灾变机制研究。同时, 还应依据并联模组内短路电流的大小, 选取熔断装置并在电池间添加气凝胶, 切断电热间的异常联动, 阻断热失控传播。

4) 优化自产热阶段电-热耦合灾变模型研究。当前电化学-热模型、电化学-短路-热模型、理想等效电路模型等电热耦合灾变模型研究较为成熟, 但上述模型不能适用于 SEI 膜分解阶段自产热量化研究, 缺乏 SEI 膜分解阶段电热间的作用机理及两者的转换范围。因此, 亟须构建 SEI 膜分解阶段的电-热耦合模型, 分析各种电池材料在该阶段电热耦合作用下的产热机理。

参考文献（References）

[1] Dixit M, Bisht A, Essehli R, et al. Lithium-ion battery power performance assessment for the climb step of an electric vertical takeoff and landing (eVTOL) application[J]. ACS Energy Letters, 2024, 9(3): 934-940.

[2] Feng X N, Ouyang M G, Liu X, et al. Thermal runaway mechanism of lithium ion battery for electric vehicles: a review[J]. Energy Storage Materials, 2018, 10: 246-267.

[3] Peng R Q, Kong D P, Ping P, et al. Thermal runaway modeling of lithium-ion batteries at different scales: recent advances and perspectives[J]. Energy Storage Materials, 2024, 69: 103417.

[4] Song L F, Wang S P, Jia Z Z, et al. A comprehensive investigation of thermal runaway critical temperature and energy for lithium iron phosphate batteries[J]. Journal of Energy Storage, 2024, 86: 111162.

[5] Yuan C H, Wang L B, Yin S, et al. Generalized separator failure criteria for internal short circuit of lithium-ion battery[J]. Journal of Power Sources, 2020, 467: 228360.

[6] Underwrites Laboratories. Standard for safety test method for evaluating thermal runaway fire propagation in battery energy storage systems: UL 9540A—2018[S]. Northbrook: Underwriters Laboratories, 2018-01-31.

[7] UN Global Technical Regulation. Global technical regulation on the electrical vehicle safety: UN GTR No. 20[S]. Geneva: United Nations, 2018-03-14.

[8] 国家市场监督管理总局, 国家标准化管理委员会. 电动汽车用动力蓄电池安全要求: GB 38031—2020[S]. 北京: 中国标准出版社, 2020-05-12.
State Administration for Market Regulation, Standardization Administration of the People's Republic of China. Electric vehicles traction battery safety requirements: GB 38031—2020[S]. Beijing: Standards Press of China, 2020-05-12(in Chinese).

[9] 国家市场监督管理总局, 国家标准化管理委员会. 电力储能用锂离子电池: GB/T 36276—2023[S]. 北京: 中国标准出版社, 2023-12-28.
State Administration for Market Regulation, Standardization Administration of the People's Republic of China. Lithium ion battery for electrical energy storage: GB/T 36276—2023[S]. Beijing: Standards Press of China, 2023-12-28(in Chinese).

[10] An Z J, Shi T L, Du X Z, et al. Experimental study on the internal short circuit and failure mechanism of lithium-ion batteries under mechanical abuse conditions[J]. Journal of Energy Storage, 2024, 89: 111819.

[11] Zhang Y, Cheng S Y, Mei W X, et al. Understanding of thermal runaway mechanism of $LiFePO_4$ battery in-depth by three-level analysis[J]. Applied Energy, 2023, 336(C): 120822.

[12] Qi C, Yan H T, Yang J, et al. Study on the characteristics of thermal runaway expansion force of $LiNi_{0.5}Co_{0.2}Mn_{0.3}O_2$/graphite lithium-ion batteries with different SOCs[J]. Electrochimica Acta, 2024, 495: 144448.

[13] Wang Z, Pang P P, Ma Z, et al. Performance degradation of lithium-ion batteries with $LiNi_{0.33}Co_{0.33}Mn_{0.33}O_2$ cathodes during long-term, high-temperature storage: behaviors and mechanism[J]. ChemElectroChem, 2021, 8(2): 403-410.

[14] Sun Y Q, Zhang S T, Qi J, et al. Calendar degradation of Li-ion batteries under high storage temperature based on electrochemical impedance spectroscopy[J]. Microelectronics Reliability, 2021, 126: 114316.

[15] Ding L, Zhang C, Wu T, et al. Effect of temperature on compression behavior of polypropylene separator used for lithium-ion battery[J]. Journal of Power Sources, 2020, 466: 228300.

[16] Mendoza-Hernandez O S, Ishikawa H, Nishikawa Y, et al. Cathode material comparison of thermal runaway behavior of Li-ion cells at different state of charges including over charge[J]. Journal of Power Sources, 2015, 280: 499-504.

[17] Li Y X, Jiang L H, Zhang N J, et al. Early warning method for thermal runaway of lithium-ion batteries under thermal abuse condition based on online electrochemical impedance monitoring[J]. Journal of Energy Chemistry, 2024, 92: 74-86.

[18] Liu S Q, Ma T Y, Wei Z, et al. Study about thermal runaway behavior of high specific energy density Li-ion batteries in a low state of charge[J]. Journal of Energy Chemistry, 2021, 52: 20-27.

[19] Li Y W, Jiang L H, Huang Z H, et al. Pressure effect on the thermal runaway behaviors of lithium-ion battery in confined space[J]. Fire Technology, 2023, 59(3): 1137-1155.

[20] Zhou H W, Fear C, Parekh M, et al. The role of separator thermal stability in safety characteristics of lithium-ion batteries[J]. Journal of the Electrochemical Society, 2022, 169(9): 090521.

[21] Gao F, Kai Y, Wang C J, et al. Experimental study on thermal runaway of $LiFePO_4$/C battery under heating condition[J]. IOP Conference Series: Earth and Environmental Science, 2020, 546(4): 042025.

[22] Zhu X Q, Sun Z W, Wang Z P, et al. Thermal runaway in commercial lithium-ion cells under overheating condition and the safety assessment method: effects of SoCs, cathode materials and packaging forms[J]. Journal of Energy Storage, 2023, 68: 107768.

[23] Qi X P, Liu B X, Pang J, et al. Unveiling micro internal short circuit mechanism in a 60 Ah high-energy-density Li-ion pouch cell[J]. Nano Energy, 2021, 84: 105908.

[24] Zhu Y L, Wang C J, Gao F, et al. Rupture and combustion characteristics of lithium-ion battery under overcharge[J]. Journal of Energy Storage, 2021, 38: 102571.

[25] Zhang L, Huang L W, Zhang Z S, et al. Degradation characteristics investigation for lithium-ion cells with NCA cathode during overcharging[J]. Applied Energy, 2022, 327: 120026.

[26] Ji C W, Zhang S Q, Wang B, et al. Study on thermal safety of the overcharged lithium-ion battery[J]. Fire Technology, 2023, 59(3): 1089-1114.

[27] Flügel M, Kasper M, Pfeifer C, et al. Cu dissolution during overdischarge of Li-ion cells to 0 V: a post-mortem study[J]. Journal of the Electrochemical Society, 2021, 168(2): 020506.

[28] Ouyang D X, Weng J W, Chen M Y, et al. Experimental analysis on the degradation behavior of overdischarged lithium-ion battery combined with the effect of high-temperature environment[J]. International Journal of Energy Research, 2020, 44(1): 229-241.

[29] Mei W X, Li H, Zhao C P, et al. Numerical study on thermal characteristics comparison between charge and discharge process for lithium ion battery[J]. International Journal of Heat and Mass Transfer, 2020, 162: 120319.

[30] Wang J F, Li Y H, Liu F, et al. A comparative study of overcharge thermal runaway force-electrical-thermal characteristics and safety assessment of lithium batteries with different cathode materials[J]. Applied Thermal Engineering, 2024, 256: 124092.

[31] Liu J L, Wang Z R, Bai J L, et al. Heat generation and thermal runaway mechanisms induced by overcharging of aged lithium-ion battery[J]. Applied Thermal Engineering, 2022, 212: 118565.

[32] Zhang Y, Li S Y, Mao B B, et al. A multi-level early warning strategy for the $LiFePO_4$ battery thermal runaway induced by overcharge[J]. Applied Energy, 2023, 347: 121375.

[33] Ouyang D X, Chen M Y, Liu J H, et al. Investigation of a commercial lithium-ion battery under overcharge/over-discharge failure conditions[J]. RSC Advances, 2018, 8(58): 33414-33424.

[34] Wang X X, Zhang Y J, Ni H J, et al. Influence of different ambient temperatures on the discharge performance of square ternary lithium-ion batteries[J]. Energies, 2022, 15(15): 5348.

[35] Liu S Z, Chen J J, Zhang C, et al. Experimental study on lithium-ion cell characteristics at different discharge rates[J]. Journal of Energy Storage, 2022, 45: 103418.

[36] Panchal S, Mathew M, Fraser R, et al. Electrochemical thermal modeling and experimental measurements of 18650 cylindrical lithium-ion battery during discharge cycle for an EV[J]. Applied Thermal Engineering, 2018, 135: 123-132.

[37] Dong T, Peng P, Jiang F M. Numerical modeling and analysis of the thermal behavior of NCM lithium-ion batteries subjected to very high C-rate discharge/charge operations[J]. International Journal of Heat and Mass Transfer, 2018, 117: 261-272.

[38] 何春汕, 王子阳, 姚斌. 不同放电功率下的储能用磷酸铁锂电池热失控特性实验研究[J]. 储能科学与技术, 2024, 13(3): 981-989.
He C S, Wang Z Y, Yao B. Experimental study of the thermal runaway characteristics of lithium iron phosphate batteries for energy storage under various discharge powers[J]. Energy Storage Science and Technology, 2024, 13(3): 981-989(in Chinese).

[39] Hu J, Liu T, Wang X S. Effect of discharge operation on thermal runaway incubation process of lithium-ion battery: an experimental study[J]. Process Safety and Environmental Protection, 2024, 185: 25-35.

[40] Hu J, Liu T, Wang X S, et al. Investigation on thermal runaway of 18, 650 lithium ion battery under thermal abuse coupled with charging[J]. Journal of Energy Storage, 2022, 51: 104482.

[41] Li L, Ju X Y, Zhou X D, et al. Experimental study on thermal runaway process of 18650 lithium-ion battery under different discharge currents[J]. Materials, 2021, 14(16): 4740.

[42] Feng X N, Ren D S, He X M, et al. Mitigating thermal runaway of lithium-ion batteries[J]. Joule, 2020, 4(4): 743-770.

[43] Wang Z P, Yuan J, Zhu X Q, et al. Overcharge-to-thermal-runaway behavior and safety assessment of commercial lithium-ion cells with different cathode materials: a comparison study[J]. Journal of Energy Chemistry, 2021, 55: 484-498.

[44] Xie S, Gong Y Z, Ping X K, et al. Effect of overcharge on the electrochemical and thermal safety behaviors of $LiNi_{0.5}Mn_{0.3}Co_{0.2}O_2$/graphite lithium-ion batteries[J]. Journal of Energy Storage, 2022, 46: 103829.

[45] Juarez-Robles D, Vyas A A, Fear C, et al. Overdischarge and aging analytics of Li-ion cells[J]. Journal of the Electrochemical Society, 2020, 167(9): 090558.

[46] Ren D S, Feng X N, Liu L S, et al. Investigating the relationship between internal short circuit and thermal runaway of lithium-ion batteries under thermal abuse condition[J]. Energy Storage Materials, 2021, 34: 563-573.

[47] Tang X, Wei X Z, Zhang H N, et al. Experimental and modeling analysis of thermal runaway for $LiNi_{0.5}Mn_{0.3}Co_{0.2}O_2$/graphite pouch cell under adiabatic condition[J]. International Journal of Energy Research, 2021, 45(7): 10667-10681.

[48] Yang Z G, Gu Z J, Tao Q Y, et al. Early warning technology for common characteristic resistances of lithium-ion batteries with thermal runaway[J]. Chemical Communications, 2024, 60(1): 87-90.

[49] Tang X, Zhang G X, Wang X Y, et al. Investigating the critical characteristics of thermal runaway process for $LiFePO_4$/graphite batteries by a ceased segmented method[J]. iScience, 2021, 24(10): 103088.

[50] Zheng S Q, Wang L, Feng X N, et al. Probing the heat sources during thermal runaway process by thermal analysis of different battery chemistries[J]. Journal of Power Sources, 2018, 378: 527-536.

[51] Huang Z H, Yu Y, Duan Q L, et al. Heating position effect on internal thermal runaway propagation in large-format lithium iron phosphate battery[J]. Applied Energy, 2022, 325(C): 119778.

[52] Liu X, Ren D S, Hsu H, et al. Thermal runaway of lithium-ion batteries without internal short circuit[J]. Joule, 2018, 2(10): 2047-2064.

[53] Mao N, Zhang T, Wang Z, et al. Revealing the thermal stability and component heat contribution ratio of overcharged lithium-ion batteries during thermal runaway[J]. Energy, 2023, 263: 125786.

[54] Mao N, Zhang T, Wang Z R, et al. A systematic investigation of internal physical and chemical changes of lithium-ion batteries during overcharge[J]. Journal of Power Sources, 2022, 518: 230767.

[55] Zhang G X, Wei X Z, Chen S Q, et al. Revealing the impact of slight electrical abuse on the thermal safety characteristics for lithium-ion batteries[J]. ACS Applied Energy Materials, 2021, 4(11): 12858-12870.

[56] Li L, Zhou X D, Ju X Y, et al. Comprehensive analysis on aging behavior and safety performance of $LiNi_xCoyMn_zO_2$/graphite batteries after slight over-discharge cycle[J]. Applied Thermal Engineering, 2023, 225: 120172.

[57] Liu L S, Feng X N, Rahe C, et al. Internal short circuit evaluation and corresponding failure mode analysis for lithium-ion batteries[J]. Journal of Energy Chemistry, 2021, 61: 269-280.

[58] Long M K, Liu S Y, Zhang G S. A novel method for simultaneous triggering and in situ sensing of internal short circuit in lithium-ion cells[J]. Energy Advances, 2023, 2(12): 2018-2028.

[59] Ma J, Huang P F, Bai Z H. Numerical study on the critical characteristics of internal short-circuit hot spot in lithium-ion batteries[J]. Process Safety and Environmental Protection, 2024, 185: 1373-1384.

[60] Xin Y D, Liu C C, Li N, et al. In-situ monitoring of multiple signals evolution behaviour for commercial lithium-ion batteries during internal short circuit[J]. Applied Energy, 2023, 350: 121754.

[61] Börger A, Mertens J, Wenzl H. Thermal runaway and thermal runaway propagation in batteries: what do we talk about?[J]. Journal of Energy Storage, 2019, 24: 100649.

[62] Lyu P Z, Liu X J, Liu C Z, et al. The influence of tab overheating on thermal runaway propagation of pouch-type lithium-ion battery module with different tab connections[J]. International Journal of Heat and Mass Transfer, 2023, 211: 124279.

[63] Wang H B, Du Z M, Rui X Y, et al. A comparative analysis on thermal runaway behavior of Li (Ni Co Mn) O_2 battery with different nickel contents at cell and module level[J]. Journal of Hazardous Materials, 2020, 393: 122361.

[64] Li H X, Gao Q, Ling J J, et al. Thermal runaway propagation and combustion characteristics of enclosed $LiNi_{0.8}Co_{0.1}Mn_{0.1}O_2$ pouch battery modules in an open environment[J]. Journal of Energy Storage, 2024, 85: 110877.

[65] Schöberl J, Ank M, Schreiber M, et al. Thermal runaway propagation in automotive lithium-ion batteries with NMC-811 and LFP cathodes: safety requirements and impact on system integration[J]. eTransportation, 2024, 19: 100305.

[66] Li H X, Gao Q, Wang Y. Experimental investigation of the thermal runaway propagation characteristics and thermal failure prediction parameters of six-cell lithium-ion battery modules[J]. Energies, 2023, 16(13): 5172.

[67] Liu T, Hu J, Tao C F, et al. Effect of parallel connection on 18650-type lithium ion battery thermal runaway propagation and active cooling prevention with water mist[J]. Applied Thermal Engineering, 2021, 184: 116291.

[68] Zhou Z Z, Zhou X D, Wang B X, et al. Experimentally exploring thermal runaway propagation and prevention in the prismatic lithium-ion battery with different connections[J]. Process Safety and Environmental Protection, 2022, 164: 517-527.

[69] Zhao L Y, Han Z X, Guo W, et al. An experimental study on thermal runaway propagation over cyclic aging lithium-ion battery modules with different electrical connections[J]. Journal of Energy Storage, 2024, 89: 111823.

[70] Han Z X, Zhao L Y, Zhao J J, et al. An experimental study on the thermal runaway propagation of cycling aged lithium-ion battery modules[J]. Fire, 2024, 7(4): 119.

[71] Huang Z H, Liu J L, Zhai H J, et al. Experimental investigation on the characteristics of thermal runaway and its propagation of large-format lithium ion batteries under overcharging and overheating conditions[J]. Energy, 2021, 233: 121103.

[72] Liu L, Lin C J, Yang P X, et al. Experimental study on thermal runaway characteristic parameters of parallel battery modules[J]. International Journal of Energy Research, 2022, 46(4): 4994-5005.

[73] Zhou Z Z, Li M Y, Zhou X D, et al. Investigating thermal runaway characteristics and trigger mechanism of the parallel lithium-ion battery[J]. Applied Energy, 2023, 349: 121690.

[74] Sun L, Wei C, Guo D L, et al. Comparative study on thermal runaway characteristics of lithium iron phosphate battery modules under different overcharge conditions[J]. Fire Technology, 2020, 56(4): 1555-1574.

[75] Newman J, Tiedemann W. Porous-electrode theory with battery applications[J]. AIChE Journal, 1975, 21(1): 25-41.

[76] Doyle M, Newman J. The use of mathematical modeling in the design of lithium/polymer battery systems[J]. Electrochimica Acta, 1995, 40(13-14): 2191-2196.

[77] Doyle M, Fuller T F, Newman J. Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell[J]. Journal of the Electrochemical Society, 1993, 140(6): 1526-1533.

[78] Bernardi D, Pawlikowski E, Newman J. A general energy balance for battery systems[J]. Journal of the Electrochemical Society, 1985, 132(1): 5-12.

[79] 郭佳涌. 动力电池模组电—热特性及过热触发热失控机理研究[D]. 北京: 北京交通大学, 2023.
Guo J Y. Study on electro-thermal characteristics and overheat triggering thermal runaway mechanism of power battery module[D]. Beijing: Beijing Jiaotong University, 2023(in Chinese).

[80] Ren H L, Jia L, Dang C, et al. An electrochemical-thermal coupling model for heat generation analysis of prismatic lithium battery[J]. Journal of Energy Storage, 2022, 50: 104277.

[81] Lin X W, Zhou Z F, Zhu X G, et al. Non-uniform thermal characteristics investigation of three-dimensional electrochemical-thermal coupled model for pouch lithium-ion battery[J]. Journal of Cleaner Production, 2023, 417: 137912.

[82] Sarkar S, Amin M T, El-Halwagi M M, et al. Thermal behavior of $LiFePO_4$ battery at faster C-rates and lower ambient temperatures[J]. Process Safety and Environmental Protection, 2024, 186: 118-133.

[83] Fan W J, Zhu J G, Qiao D D, et al. Low-temperature charging strategy optimization based on electrochemical-aging-thermal coupling model[J]. IEEE Transactions on Transportation Electrification, 2025, 11(1): 901-911.

[84] Xi R Y, Mu Z Z, Ma Z Y, et al. Lifetime prediction of rechargeable lithium-ion battery using multi-physics and multiscale model[J]. Journal of Power Sources, 2024, 608: 234622.

[85] Shen J W, Zhang Z, Chen Z, et al. Temperature estimation of multiple places for lithium-ion batteries based on improved electrochemical thermal modeling[J]. IEEE Transactions on Transportation Electrification, 2025, 11(1): 382-392.

[86] Liu L S, Feng X N, Zhang M X, et al. Comparative study on substitute triggering approaches for internal short circuit in lithium-ion batteries[J]. Applied Energy, 2020, 259: 114143.

[87] Huang Z H, Zhao C P, Li H, et al. Experimental study on thermal runaway and its propagation in the large format lithium ion battery module with two electrical connection modes[J]. Energy, 2020, 205: 117906.

[88] Niu H C, Chen C X, Ji D, et al. Thermal-runaway propagation over a linear cylindrical battery module[J]. Fire Technology, 2020, 56(6): 2491-2507.

[89] Xu C S, Zhang F S, Feng X N, et al. Experimental study on thermal runaway propagation of lithium-ion battery modules with different parallel-series hybrid connections[J]. Journal of Cleaner Production, 2021, 284: 124749.

[90] Gao S, Feng X N, Lu L G, et al. An experimental and analytical study of thermal runaway propagation in a large format lithium ion battery module with NCM pouch-cells in parallel[J]. International Journal of Heat and Mass Transfer, 2019, 135: 93-103.

[91] Yang S C, Gao X L, Li Y L, et al. Minimum lithium plating overpotential control based charging strategy for parallel battery module prevents side reactions[J]. Journal of Power Sources, 2021, 494: 229772.

[92] Wang G Q, Ping P, Kong D P, et al. Advances and challenges in thermal runaway modeling of lithium-ion batteries[J]. The Innovation, 2024, 5(4): 100624.

Progress in the investigation of electro-thermal coupled catastrophic mechanisms in lithium-ion battery modules

HE Yuanhua，LIANG Jiaxin，LI Zitong，HUANG Jiang*

(Civil Aircraft Fire Science and Safety Engineering Key Laboratory of Sichuan Province，Civil Aviation Flight University of China，Guanghan 618307，China)

Abstract：In the many applications of lithium-ion batteries, the problem of thermal failure presents a significant safety risk. This failure is characterized by a sequence of events: the degradation of interface thermal stability, the onset of thermal runaway triggered by internal short circuits within a cell, and the subsequent thermal propagation at both the module and system levels. The coupled electro-thermal catastrophe serves as both a critical dimension in characterizing the thermal failure logic of lithium-ion battery systems and a key breakthrough in managing these thermal failures. This paper provides a thorough analysis of the thermal behaviors, such as self-generated heat, Joule heating, and heat conduction, that affect electrical behaviors during the solid electrolyte interphase (SEI) layer decomposition of a single cell, the separator melting, and heat propagation at the module level for electro-thermal coupling catastrophe in lithium-ion battery systems over long cycles and across multiple scales. The heat generation and heat transfer behaviors associated with electrical phenomena are also analyzed, such as overcharging, overdischarging and battery aging. The paper also reviews the most prominent electro-thermal coupling models for both individual cells and modules, providing a synthesis of the mechanisms that govern the electro-thermal interactions in lithium-ion batteries. By exploring these interactions from both microscopic and macroscopic viewpoints, the study aims to shed light on the future advancements in the management of electro-thermal coupling in battery modules. The goal of this research is to improve the safety and dependability of lithium-ion battery systems through design, operation, and maintenance.

Keywords：lithium-ion battery；electro-thermal coupling；SEI layer decomposition；separator melting；thermal runaway propagation

Received：2024-07-26；**Accepted**：2024-08-30；**Published Online**：2024-09-24 12：31

URL：link.cnki.net/urlid/11.2625.V.20240924.1014.001

Foundation items：National Natural Science Foundation of China (U2333210)； The Fundamental Research Funds for the Central Universities (24CAFUC10161,24CAFUC03076)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0799

中国空间站延寿需求下机电产品在轨寿命再评估技术与展望

黄首清[1,2]，夏侨丽[3]，许庶[1,2,*]，余溢方[1,2]，刘守文[1,2]，张兆霖[1,2]

(1. 北京卫星环境工程研究所，北京 100094；2. 可靠性与环境工程技术重点实验室，北京 100094；
3. 北京空间飞行器总体设计部，北京 100094)

摘 要：针对中国空间站延寿对准确定量开展机电产品在轨寿命再评估的迫切需求，在借鉴国际空间站寿命评估与延寿成功经验和中国空间站研制阶段寿命评估技术的基础上，较为系统地综述了空间站延寿需求下机电产品在轨寿命再评估技术，并给出了技术发展趋势和建议。该技术包括以地面试验数据为主和综合天地数据两类，涵盖寿命评估结果常规修正、综合动作和退化效应的寿命试验再评估、寿命试验复杂动态剖面再设计、多层级加速寿命试验与评估、综合辐照环境下寿命再评估、天地多源数据再筛选和寿命再评估、智能故障诊断和退化预测、基于天地数据迁移学习的寿命评估、基于在轨试验的寿命再评估9项具体技术。研究体现了中国空间站进入长期在轨运营阶段后，对机电产品寿命再评估工作的新思考和新理念，将服务中国空间站延寿这一长期的系统工程。

关 键 词：中国空间站；机电产品；在轨寿命再评估；地面试验数据；在轨遥测数据
中图分类号：V41
文献标志码：A **文章编号**：1001-5965(2026)10-3426-11

2022年底，中国空间站全面建成并进入长期在轨运营阶段[1]，主要任务为支持航天员乘组长期在轨健康生活、顺利开展各项载荷试验与科学研究。保障空间站平台安全稳定运行并逐步延长使用寿命，是中国空间站发挥国家太空实验室长期效益的基本要求[2]。考虑到长期载人和在轨可维修的鲜明特点[3-4]，空间站延寿比其他航天器需求更加突出、技术更加复杂、实施更加慎重[5-6]。国际空间站自1998年发射首个舱段至今已超过25年，最初设计寿命为15年，历经3次延寿/延寿决策，分别延寿至2020年、2024年和2030年[7]。从国际空间站的运营经验看，准确定量的在轨寿命评估是决策和实施空间站延寿的关键技术和前提，对评估在轨剩余寿命、减少计划外维修、支撑维修决策、规划寿命保障物资意义重大，是空间站运营阶段一项必须长期坚持的系统性工作[8-9]。此外，在轨寿命评估是研制阶段寿命评估工作的延续和拓展，具有典型的“再评估”特点，如国际空间站针对泵、电源、控制力矩陀螺、二氧化碳去除装置、水处理装置等产品，在初始评估的基础上，采用贝叶斯方法常态化更新平均故障间隔时间 (mean time between failure, MTBF) 和可靠度结果，为国际空间站运维和延寿提供支撑[10]。

机电产品在轨寿命再评估是空间站寿命评估面临的核心问题之一，这也是由机电产品的特点决定的。典型的空间站机电产品包括控制力矩陀螺、

收稿日期：2024-11-07；录用日期：2025-01-10；网络出版时间：2025-02-11 12:30
网络出版地址：link.cnki.net/urlid/11.2625.V.20250211.1056.001
基金项目：可靠性与环境工程技术重点实验室稳定支持科研项目 (614200420230201)
* 通信作者. E-mail：xushu@buaa.edu.cn

引用格式：黄首清，夏侨丽，许庶，等. 中国空间站延寿需求下机电产品在轨寿命再评估技术与展望[J]. 北京航空航天大学学报，2026，52（10）：3426-3436. Huang S Q，Xia Q L，Xu S，et al. On-orbit life re-evaluation technology and prospects for electromechanical products in response to demand for China Space Station lifespan extension[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（10）：3426-3436（in Chinese）.

机械臂关节、压气机、对日定向装置、驱动机构、热控流体回路泵、风机、尿旋转蒸馏装置、温控阀、舱外电动工具。很多机电产品具有技术复杂、运动工况多、寿命与可靠性要求高的特点[11], 在轨长期工作发生老化[12]、疲劳[13]、磨损[14]、蠕变[15]等退化乃至失效的概率较高, 还要承受特殊的空间辐射环境[16], 直接影响空间站的安全可靠运行。此外, 影响空间站机电产品寿命的因素很多, 包括产品寿命特性、空间环境、使用条件, 对寿命再评估提出重大挑战。考虑到实际工作特点并重点关注耗损型失效, 本文不讨论空间站燃料耗尽和空间碎片对机电产品寿命的影响。

直接研究空间站机电产品在轨寿命再评估的文献较少。文献 [17-18] 聚焦国际空间站在轨失效控制力矩陀螺的事后分析, 分析了失效前遥测数据的早期异常征兆, 作为反面案例佐证了在轨寿命评估的重要性和可行性。Li 等[19] 提出基于蒙特卡罗仿真和充足数据支撑下的空间站寿命预测模型。余溢方等[20] 基于 4 种神经网络模型和在轨遥测数据提出空间站流体回路泵智能故障诊断方法。刘凯等[21] 提出高加速寿命试验 (highly accelerated life test, HALT)-高加速应力筛选 (highly accelerated stress screen, HASS)-加速退化试验 (accelerated degradation test, ADT) 的高加速试验集成方法并建议在空间站有效载荷领域实施。地面寿命试验、加速寿命试验是空间站寿命验证与评估主要的研究手段, 且主要针对单机或部件开展, 如舱外电动工具[22]、风机轴承[23]、发动机[24]、空间挠性作动器[25]、锁紧装置[26]。基于失效物理的寿命评估方法常用于锂电池[27]、机械部件[28] 等物理模型和失效机理较为明确的产品, 具有良好的可解释性。近年来, 物理信息和数据混合驱动的剩余寿命评估方法[29] 受到广泛关注, 已应用于发动机、液压活塞泵、电液作动器等多种机电产品, 将有助于提升航天器产品寿命评估准确度。

针对其他航天器, 寿命评估的研究和实践较为丰富, 可为空间站机电产品在轨寿命再评估提供借鉴。杜善亮等[30] 提出一种单机-系统分层级、随机失效与性能退化相结合的航天器寿命评估方法。刘志全等[31] 在充分调研国内外资料的基础上, 提出航天器机构加速寿命试验技术在工程应用中存在的问题和解决思路。朱炜等[32] 提出包含寿命分析、寿命设计、寿命试验和寿命评估的航天器寿命工作流程。黄首清等[33-34] 设计了加速寿命试验设计与评估软件 ALT511, 提高试验设计与分析的规范性和效率。Sankararaman 等[35] 针对航天产品提出基于逆向一次二阶矩模型的剩余寿命预测方法, 以量化不确定性并提高预测准确性。

充分的数据支撑是准确开展空间站机电产品在轨寿命再评估的重要基础[10,36], 这也是国际空间站寿命评估和延寿的重要经验之一。皇甫素伟等[37] 在多个系列航天器的寿命与可靠性数据库支持下, 开展了氢镍蓄电池组、太阳电池阵、推力器、数管计算机等航天器产品的寿命与可靠性评估, 为开展空间站机电产品在轨寿命再评估提供了有益经验。

综上所述, 当前中国空间站机电产品寿命评估和相关研究主要基于研制阶段数据开展。随着不断积累在轨遥测数据、持续推进地面试验研究、逐渐深入对产品寿命特性认识, 寿命评估工作和结论将分阶段完善和更新, 越来越鲜明地体现“再评估”的特点。本文在充分总结空间站研制阶段机电产品寿命评估技术的基础上, 针对天地差异性和产品在轨工作特性, 梳理面向延寿的机电产品在轨使用寿命再评估的思路, 从以地面试验数据为主和综合天地数据两方面综述了在轨寿命再评估技术, 并给出相关技术发展趋势和建议。

1　中国空间站研制阶段机电产品寿命评估技术

如图 1 所示, 中国空间站研制阶段机电产品典型的寿命评估技术包括半定量寿命评估、定量寿命评估、多源数据融合评估、部件-组件两层级寿命试验验证、加速寿命试验评估, 以核心舱为例, 相关技术应用涉及环控、热控、推进、电源等重要分系统。

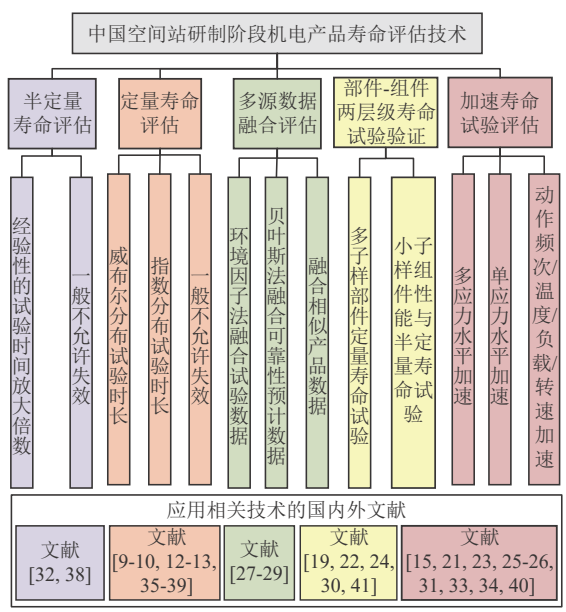


图 1　中国空间站研制阶段机电产品寿命评估技术

Fig. 1　Life evaluation technology for electromechanical products during China Space Station development phase

1.1 半定量寿命评估

本文中半定量寿命评估指不考虑可靠度的寿命评估，可验证机电产品以一定裕度满足寿命指标，但不能给出寿命评估结论成立的概率。一些寿命末期可靠性要求不高的机电产品可采用半定量寿命评估方法，基本思路是将寿命指标乘以试验时间放大倍数作为试验时长。试验时间放大倍数体现满足寿命指标的裕度，一般由工程经验确定，空间站产品通常取1.5。研制阶段，通过寿命试验且未出现失效验证研制阶段寿命指标。空间站机电产品寿命特征量通常为时间、转数、循环数、动作次数等。

1.2 定量寿命评估

本文定量寿命评估指考虑可靠度的寿命评估，可验证机电产品是否以一定概率满足寿命指标，考虑到寿命指标一般为下限值，通常还会约定置信度。机电产品的寿命一般可用威布尔分布来描述，还有一些机电产品的寿命满足指数分布，试验后需进行假设检验进一步确认寿命分布模型。以威布尔分布为例，各试验件试验时长[38]由下式确定：

$$L_{\mathrm{R}}=L_{\mathrm{l}}\left[\frac{\ln(1-\gamma)}{N\ln R_{\mathrm{L}}}\right]^{1/m} \tag{1}$$

式中：L_{R}为试验时长；L_{l}为寿命指标；γ为置信度；N为试验件数量，一般要求不少于2个；R_{L}为寿命末期可靠度下限；m为形状参数，当失效数≥3时通过分析失效数据确定，当失效数＜3时取经验值。m的经验值在实际应用中参考值为：新研单机取1.5，改进单机取2，沿用单机取3。通过寿命试验且未出现失效即验证了研制阶段寿命指标。

1.3 多源数据融合评估

研制阶段没有在轨数据、寿命试验数据也可能不充分，多源数据融合是解决定量寿命评估数据不足难题的重要手段之一。环境因子折算法、相似产品法、退化外推法、贝叶斯法是空间站研制阶段常用的多源数据融合方法。环境因子折算法指将多类环境试验时长乘以环境因子折算到寿命试验时长中[39]，环境因子一般由工程经验确定，如鉴定级热真空试验取13、验收级热真空试验取8、鉴定级随机振动试验取100、验收级随机振动试验取50。考虑到空间站的高可靠性要求，环境因子取值比一般航天器[39]更加保守。相似产品法指利用与待评估产品相似的已有成熟产品的寿命数据来估计待评估产品寿命的方法。相似产品应在结构、性能、任务剖面等方面与待评估产品存在相似性，这样可将相似产品在轨服役时长等寿命数据计入待评估产品寿命试验时长中。退化外推法指通过分析产品在寿命试验中性能参数的变化趋势，来预测未来某个时间点的产品状态或失效时间，可将失效时间（也称作伪寿命）和1次失效分别计入寿命试验总时长和失效数中，更新寿命评估结果。型号中应用的贝叶斯法与经典的贝叶斯模型有区别，主要利用该方法的思想，将试验前可靠性预计数据折算到寿命试验时长中[30,39]。

1.4 部件-组件两层级寿命试验验证

为解决复杂机电产品验证试验周期长、产能限制无法多子样投入的问题，空间站研制阶段选取决定寿命的关键部件，采取部件-组件两层级寿命试验验证方法。首先，综合寿命分析、危险分析、概率风险分析、故障模式影响分析、故障树分析的结果，识别系统风险点和产品寿命特征，确定关键部件。其次，针对部件层级策划多子样统计试验，验证关键部件的寿命，并积累多子样部件寿命数据。接着，开展组件试验，验证各部件工作的匹配性和整体功能性能指标的符合性。通过多子样部件定量寿命试验与小子样组件性能与半定量寿命试验相结合的策略，实现了验证时机前移，并显著控制了寿命验证成本。空间站控制力矩陀螺、太阳翼、驱动机构等重要产品采用上述技术，与基于统计模型直接开展组件层级定量寿命评估相比，试验成本减少了一个数量级以上，并实现了验证时机前移。

1.5 加速寿命试验评估

加速寿命试验是快速验证机电产品长寿命指标的重要手段。按照加速应力分类，空间站机电产品研制阶段开展的加速寿命试验主要包括转速加速、动作频次加速、温度加速、负载加速。转速加速产品/部件包括导电环、某敏感器转动装置；动作频次加速产品/部件包括关节、阀门；温度加速产品/部件包括某机构钢丝绳、密封圈；负载加速产品/部件包括风机轴承。从加速应力水平数的角度看，钢丝绳和密封圈等产品/部件开展了多加速应力水平（温度）加速寿命试验并外推正常温度下的寿命，其加速因子为试验后通过计算得到。而有些机电产品受限于研制周期等限制，仅开展了单应力水平加速寿命试验，加速因子在试验前由经验或文献确定。

1.6 效果和启示

研制阶段机电产品寿命评估技术对空间站在轨稳定运行发挥了重要作用，初步验证了中国空间站建造完成后至少运行10年的寿命指标[1]。但由于研制时间和数据有限，以及认知局限性，面向20年乃至30年延寿应用需求，针对天地差异性、在轨遥测数据持续补充、在轨飞行状态信息和经验不断积累等新情况，未来还需要在失效机理代表

性、数据完备性、试验条件有效性、数据分析合理性方面开展进一步研究。

2　面向延寿的机电产品在轨寿命再评估技术

中国空间站在轨运营阶段应在研制阶段大量寿命评估工作的基础上，充分利用持续增长在轨遥测数据，围绕失效机理代表性、数据完备性、试验条件有效性、数据分析合理性 4 个方面开展机电产品在轨寿命再评估工作，并应特别注重基础数据积累和分析、退化规律和失效机理研究及创新方法应用。典型的在轨寿命再评估技术及其适用的典型产品如图 2 所示，以可能对空间站寿命和运行风险产生较大影响的部分重要产品为例，从数据来源角度可将提出的空间站机电产品在轨寿命再评估技术分为以地面试验数据为主和综合天地数据两类，共 9 项具体技术。例如，对于缺少在轨遥测数据的阀门、密封件、舱外电动工具等产品，其寿命再评估仍以地面数据为主，结合历史试验数据和补充开展的地面试验数据开展寿命再评估；对于流体回路泵、控制力矩陀螺等可监测在轨状态的产品，着重开展综合天地数据的寿命再评估。

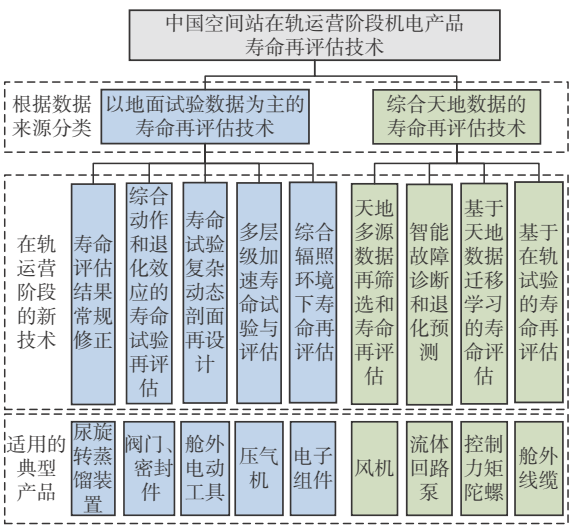


图 2　中国空间站在轨运营阶段机电产品寿命再评估技术

Fig. 2　Life re-evaluation technology for electromechanical products during China Space Station on-orbit operation phase

2.1　寿命评估结果常规修正

1）1:1 寿命试验结果修正。在完成研制阶段寿命指标验证且入轨运行后，一些机电产品地面试验仍在进行，寿命评估结果也应相应更新。对于半定量寿命评估，将最新的试验时长除以试验时间放大倍数即为寿命再评估结果。对于定量寿命评估，以寿命服从威布尔分布为例，将延寿预期寿命指标代入式(1)即可更新试验时长，确定无失效前提下的试验截尾时刻。如果研制阶段试验件状态已发生明显变化，则应投产新的试验件重新开展试验。

此外，考虑到部分试验将做到完全失效，本文给出有失效时寿命评估方法。以威布尔分布为例，再评估可靠寿命t_{R2}由下式确定[39]：

$$t_{R2}=\left[-\frac{2\sum_{i=1}^{N}l_i^m\ln R_L}{\chi_\gamma^2(2r+2)}\right]^{1/m} \tag{2}$$

式中：r为失效数；l_i为第i个试验件试验时长；$\chi_\gamma^2(2r+2)$为自由度为$2r+2$的卡方分布下侧γ分位数。当失效数≥3 时，形状参数 m 可由极大似然估计法计算。

未出现失效但有稳定退化的试验件，可根据退化建模得到伪寿命值计入试验时长，并等效 1 次失效，进而根据式(2)计算可靠寿命。

2）加速寿命试验结果修正。由于缺少正常应力水平的真实遥测数据，研制阶段加速寿命试验确定的正常应力水平常常比较保守，在获得在轨遥测数据后，应对正常应力水平进行修正，进而修正加速因子和加速寿命评估结果，通常会获得更长更真实的寿命评估结果，如图 3 所示。具体到转速加速、动作频次加速、温度加速、负载加速 4 类加速寿命试验，根据遥测数据修正正常应力水平下的转速、动作频次、温度、负载，进而修正加速寿命评估结果。

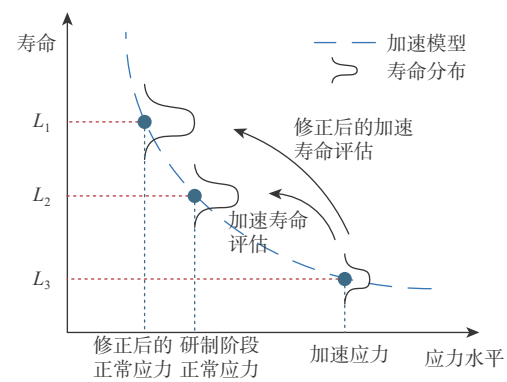


图 3　加速寿命评估结果常规修正示意图

Fig. 3　Schematic diagram of routine corrections to accelerated life evaluation results

2.2　综合动作和退化效应的寿命试验再评估

温控阀、电磁阀等机电产品在研制阶段主要验证了动作次数指标，然而，并未充分验证阀芯材料的老化和长期蠕变效应、弹簧部件的长期蠕变效应，可能导致寿命评估结果较为激进。根据温控阀、电磁阀的在轨实际动作频次及介质温度，可设计综合动作和退化效应的寿命试验剖面，典型剖面如图 4 所示。又如某舱外电动工具约 3 个月工作 1 次，单次工作约 8 h，在轨舱内长期贮存与短期舱外工作

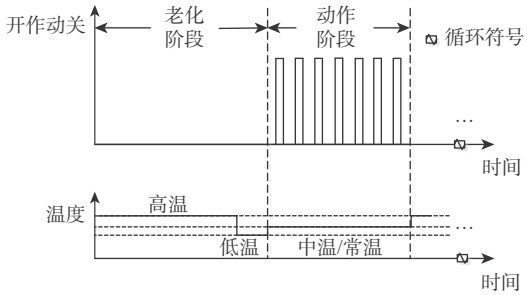


图 4 综合动作和退化效应的典型寿命试验剖面

Fig. 4 Typical life test profile combining action and degradation effects

状态交替进行，寿命试验剖面也应尽可能符合实际情况。开关按钮、锂电池等产品也有类似的长期在轨贮存状态，存在按钮弹簧蠕变、锂电池老化的风险。如果老化阶段过长，也可设计加速寿命试验，文献 [40] 针对空间站开关按钮、锂电池长期在轨贮存、多次短时工作的实际情况，开展了综合动作和退化效应的寿命试验，修正了寿命评估结果。

2.3 寿命试验复杂动态剖面再设计

由于难以确切预知在轨实际环境和工况，研制阶段设计的有些寿命试验剖面做了理想化的简化和剪裁。主要表现在用环境试验剖面代替寿命试验剖面、不符合实际的恒定或稳态应力，这些试验剖面难以充分体现在轨实际环境和工况的典型特征。如航天员出舱活动用的舱外电动工具，研制阶段针对锂电池在放电深度 50% 条件下进行了 1 000 次循环充放电寿命试验。而出舱活动中锂电池的实际充放电剖面要复杂得多，如图 5 所示的典型寿命试验剖面，包括恒流充电、恒压充电、待机放电、多次的工作放电。设计这种反映实际在轨复杂动态工况的寿命试验剖面，并充分考虑典型应力类型、应力大小、作用时机、持续时长等要素[42]，有助于提高试验有效性。考虑到航天器领域尚无设计寿命试验剖面的专用标准，对于温度应力和电应力，可

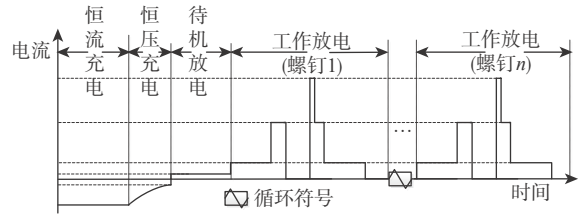


图 5 舱外电动工具锂电池典型的复杂动态寿命试验剖面

Fig. 5 Typical complex dynamic life test profile for lithium-ion batteries of extravehicular power tools

借鉴 GJB 899A—2009 附录 B[43] 的方法设计寿命试验剖面。

2.4 多层级加速寿命试验与评估

该技术在部件-组件两层级寿命试验验证技术的基础上结合加速寿命试验的特点提出。研制阶段加速寿命试验的主要目的是验证寿命指标，而在轨运营阶段开展加速寿命试验还应重点关注退化规律、加速规律及机理，更多体现技术增量和创新。多层级加速寿命试验与评估技术可以很好地满足在轨运营阶段的新要求。如图 6 所示，多层级加速寿命试验是指根据产品的组成结构和工作原理，在材料、部件和组件多个层级选取有代表性的试验对象开展加速寿命试验。材料级与部件级试验件具有成本低、样本量大的优势，可设计多组加速应力水平开展试验，获得更充分的试验数据，摸清主要失效机理的失效时间分布规律、加速规律和加速因子，从而在组件级以很少的试验件和试验时长代价获得更全面和可信的寿命评估结论，还可为产品设计改进提供直接技术支撑[41]。多层级加速寿命试验与评估是解决小子样长寿命多失效机理空间站产品加速寿命评估难题并有效支撑延寿设计的有效途径之一。

2.5 综合辐照环境下寿命再评估

针对空间站部分关键电子产品及机电产品中的电子组件，研制阶段已开展总剂量试验等辐照试

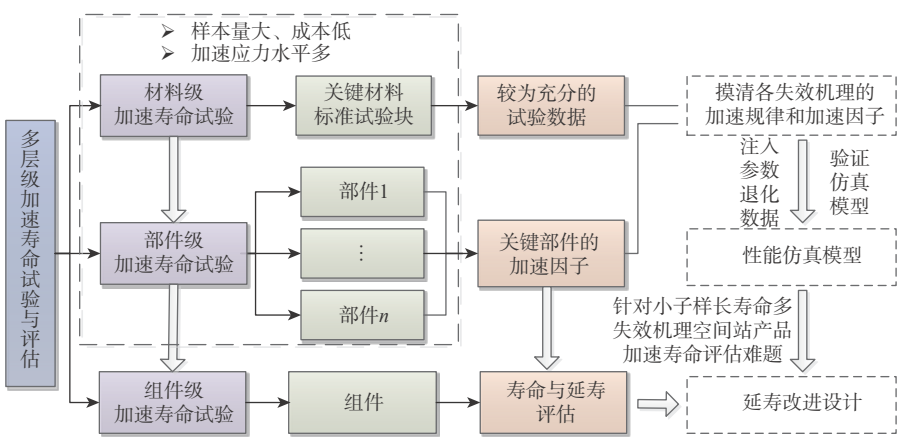


图 6 多层级加速寿命试验与评估示意图

Fig. 6 Schematic diagram of multi-level accelerated life test and evaluation

验验证寿命。以总剂量试验为例，通常以钴源作为辐照源进行 γ 射线辐照[44]。然而，真实空间环境包含宽能谱的电子、质子、重离子等多种辐射粒子，单一种类粒子辐照试验及多种粒子辐照串行试验难以模拟不同粒子辐射之间的协同效应[45]，可能造成欠试验或过试验，影响寿命评估结论。空间环境模拟技术的进步是解决这一问题的重要手段，2024 年正式运行的国家重大基础科技设施“空间环境地面模拟装置”（见图 7）具备真空、高低温、宽能谱电子/质子/重离子辐照、高准直太阳及紫外电磁辐照等空间综合环境模拟能力[46]，为开展空间站机电产品综合辐照环境下寿命再评估提供了技术支撑。

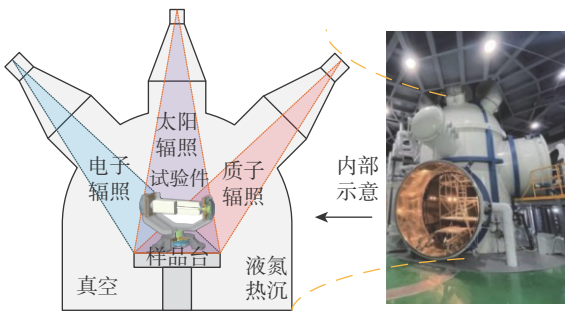


图 7　空间环境地面模拟装置综合辐照模拟舱

Fig. 7　Integrated irradiation simulation chamber of the space environment simulation rescearch infrustructure

2.6　天地多源数据再筛选和寿命再评估

从数据可信度和价值方面分析，在轨运营阶段可用于产品寿命评估的多源数据可分为 4 类：第 1 类是在轨遥测数据，可信度和价值最高；第 2 类是未加速的地面寿命试验数据，可信度和价值较高；第 3 类主要包括退化外推数据、加速寿命试验数据和相似产品数据，可信度和价值较低；第 4 类主要包括环境因子折算数据和可靠性预计数据，可信度和价值最低。在轨运营阶段的寿命再评估应优先用第 1 类和第 2 类数据，谨慎使用第 3 类数据，建议舍弃第 4 类数据。随着在轨运行时长不断增加，应逐渐增加第 1 类数据的占比，舍弃第 4 类数据，并相应减少第 3 类乃至第 2 类数据的占比，动态或分阶段地更新寿命再评估结论。

2.7　智能故障诊断和退化预测

空间站很多机电产品健康状态种类多，且与多个性能参数耦合相关，传统的基于阈值超差和逻辑规则的故障诊断技术在适应性方面难以满足实际应用需求。图 8 给出一种已在空间站流体回路泵应用的智能故障诊断技术[20]，包括输入训练数据集、训练人工智能模型、判断健康状态类型 3 个步骤。本文训练数据集是表征温度、出口压力、流

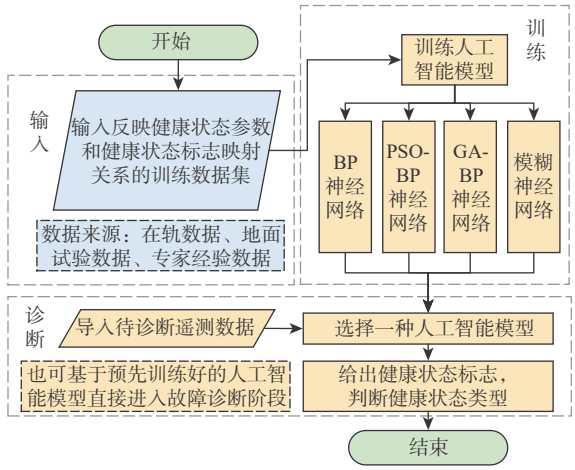


图 8　智能故障诊断技术示意图[20]

Fig. 8　Schematic diagram of intelligent fault diagnosis technology[20]

量、转速多个健康状态参数和健康状态标志（1～6 共 6 种）映射关系典型数据的集合，数据来源于在轨数据、地面试验数据和专家经验数据。通过反向传播（back propagation, BP）神经网络、粒子群优化的 BP（particle swarm optimization-BP, PSO-BP）神经网络、遗传算法优化的 BP（genetic algorithm-BP, GA-BP）神经网络、模糊神经网络等人工智能模型学习隐藏在数据背后的映射关系，可以泛化预测训练数据集以外的健康状态参数对应的健康状态。随着在轨运营时间和经验的不断积累，可不断更新和优化训练数据集，获得适应性更强、准确性更高的故障诊断结果。

图 8 所示故障诊断方法虽然可用于发现早期故障，但在故障预警方面仍存在不足，而退化预测技术可实现剩余寿命预警功能。图 9 展示了空间站流体回路泵退化预测软件界面，具备导入数据、降噪和特征提取、单特征退化建模与预测、多特征退化建模与预测等功能。基于合适的数据特征、足

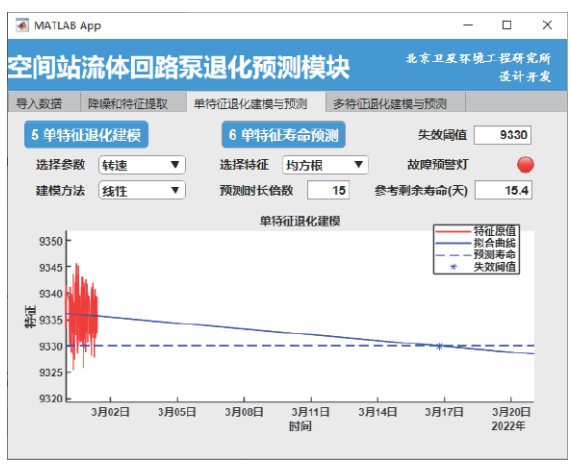


图 9　空间站流体回路泵退化预测软件界面

Fig. 9　Space station fluid loop pump degradation prediction software interface

够的数据时长，辅以必要的数据降噪、数据降维算法，利用合适的退化模型可取得更好的退化预测效果。软件中的退化模型包括线性、抛物线、指数、对数、BP神经网络、高斯回归。

2.8 基于天地数据迁移学习的寿命评估

如图10所示，地面数据缺点是不够真实，优点在于传感器多、采样率高、工况多、失效数据多；而在轨遥测数据优点是真实，但传感器少、采样率低、工况少、失效数据少。基于天地数据迁移学习的寿命评估可以充分发挥两类数据优势，利用其关联性和差异性，可有效提高故障诊断、退化预测、寿命评估的效率和准确性。可用于空间站产品天地数据迁移学习的方法[47-49]包括基于参数的迁移学习、基于实例的迁移学习和基于映射的迁移学习。

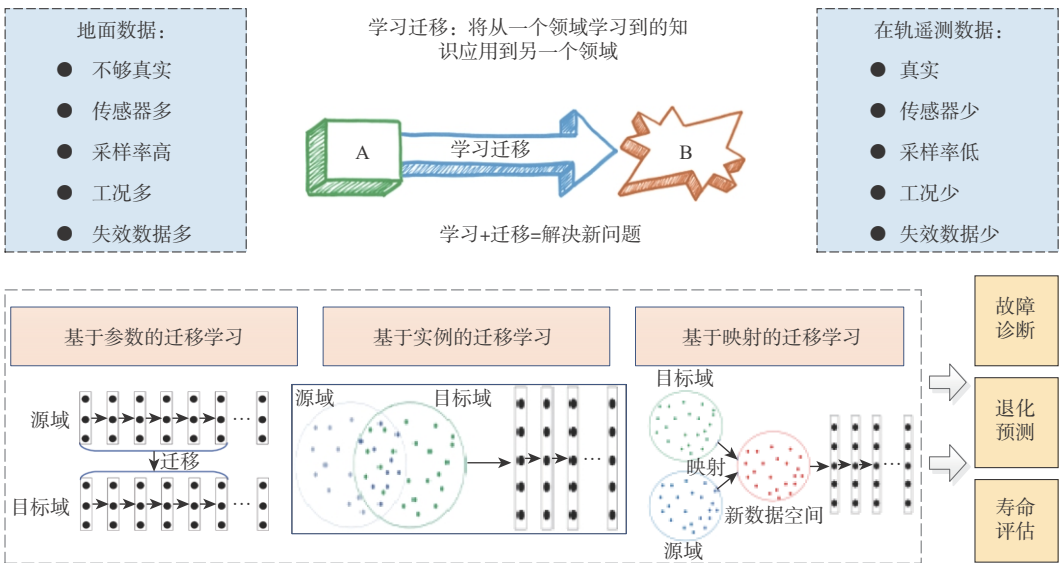


图10 基于天地数据迁移学习的寿命评估思路

Fig. 10 Lifetime evaluation approach based on transfer learning between on-orbit and ground data

对于空间站机电产品，以故障诊断场景下应用神经网络模型为例，简要介绍3种天地迁移学习方法。基于参数的迁移学习指重用或部分重用通过地面试验数据预先训练好的神经网络，包括网络参数和结构，将其迁移到用于处理在轨遥测数据的神经网络中。基于实例的迁移学习指将地面数据以一定权重融入在轨数据中，共同构建用于故障诊断的训练数据集。基于映射的迁移学习指将天地数据共同映射到一个新的数据空间，构建训练数据集。经映射后的地面数据更加真实可信，而经映射后的在轨遥测数据更加丰富多样。有条件的情况下，为完善训练数据集，可补充开展注入故障产品的地面试验，以获得多参数、多位置、多工况、多健康状态的高频采样数据，更好地保证天地数据融合的效果。

2.9 基于在轨试验的寿命再评估

对于无在轨遥测数据的机电产品及其组件，在轨试验是准确评估在轨寿命的重要技术途径之一，当地面难以真实模拟在轨真实空间环境时更是如此，且空间站本身就是开展在轨工程试验的重要平台[50]。站在寿命评估的角度，空间站在轨试验具有地面试验难以比拟的优势，可实现在轨原位测试、下行试验件、在轨比测、舱外暴露等操作。在轨原位测试可以弥补无在轨遥测数据的不足，并可通过航天员完成较为复杂的在轨测试和分析工作。下行试验件可获得最真实的退化或失效产品，借助地面上先进测试分析设备完成退化或失效分析。在轨比测指在轨完成不同设计、不同工艺、不同工作位置机电产品的对比测试，对比获得退化或寿命数据，给出不同设计、工艺、工作位置寿命优劣的结论。舱外暴露指利用空间站舱外暴露平台或舱外载荷适配器开展产品舱外暴露试验，在真实的空间辐射环境下考核产品寿命。此外，空间多粒子连续能谱综合辐射环境和持续的微重力环境是地面不能真实模拟的，凸显在轨试验的价值。但在轨试验应充分考虑质量、功耗、尺寸、航天员操作、物资和数据上下行、时间、经费等资源需求，以及对空间站可靠性和航天员安全性的影响。

3 发展趋势与建议

1）在轨维修和保障有限支持下的寿命动态评估。将在轨维修和保障能力及其影响要素模型化，充分考虑产品维修性设计、天地备品备件情况、航天员维修时机、在轨维修和保障约束等因素的综合作用和动态特性，并融合到寿命评估模型中，实现寿命动态评估，是需要突破的关键技术和重要的发

展趋势。由于寿命动态评估技术难度大、模型复杂, 还需要面向工程应用进行必要简化, 最终形成具有良好可操作性的方法、流程和标准。

2) 在轨原位自主和智能寿命评估。在寿命评估领域, 在轨原位自主完成寿命评估, 及时发现影响寿命和可靠性的潜在问题很有必要, 可支持空间站及时调整运行状态、采取应急措施或更新维护计划, 以延长使用寿命。在轨原位自主寿命评估需要基于在轨和历史积累的多源海量数据, 并考虑空间环境、任务工况、产品自然退化等多种复杂因素, 在大数据和人工智能技术支持下, 科学、准确、自主地给出和更新寿命评估结论。

3) 多源数据和失效物理综合驱动的寿命评估。多源数据驱动和失效物理驱动是高度互补的两类寿命评估模型, 不断完善两类模型并加强综合应用是重要的技术发展趋势。一方面, 要加强失效物理模型和数据驱动模型预测结果的相互校验和修正, 使两类模型更加符合实际情况。另一方面, 要加强模型融合方法研究, 可借助贝叶斯模型、可解释神经网络模型、迁移学习模型等实现两类模型融合。加强跨学科合作和技术创新也是推动这一领域发展的关键。

4) 机电产品空间辐射效应量化认知和预警。在空间站机电产品寿命评估领域, 对空间辐射效应的认识远不如对传统的磨损、疲劳等失效机理认识地深入和量化。空间辐射对机电产品的损伤主要作用在材料层级和微观尺度, 但影响却表现在单机和宏观层面。此外, 空间辐射引起的故障, 尤其是单粒子效应等瞬时效应受空间环境的影响很大, 如南大西洋异常区上空单粒子事件概率显著高于其他区域。因此, 建议加强空间辐射环境对典型机电产品环境效应传递机制、微观缺陷对宏观性能的跨尺度影响机理、空间辐射环境-效应-故障一体化分析和预警等研究。

5) 寿命评估软件工具应用。寿命评估软件工具可为地面试验设计与寿命评估、遥测数据分析、天地数据融合提供重要技术手段, 提高分析效率和规范性、方便研制数据管理、降低技术门槛、促进新技术推广应用。此外, 实现不同来源、不同类型寿命评估软件工具的数据交互、协同工作和统一管理有利于形成技术合力, 支撑全面、准确、及时地开展空间站寿命评估。建议将智能故障诊断和退化预测、天地数据融合相关的寿命评估软件工具或功能集成到中国数字空间站平台, 借助数字空间站强大的数字孪生、仿真计算、数据交互和协同工作能力, 形成空间站在轨寿命的实时智能评估能力。

4　结 束 语

在充分借鉴国际空间站寿命评估与延寿成功经验和系统总结中国空间站研制阶段寿命评估技术的基础上, 较为系统地综述了空间站延寿需求下机电产品在轨寿命再评估技术, 包括总体思路和寿命评估结果常规修正、天地多源数据再筛选和寿命再评估等 9 项具体技术。体现了中国空间站进入长期在轨运营阶段后, 在不断积累在轨遥测数据、持续推进地面试验研究、逐渐深入对产品寿命特性认识背景下, 对机电产品寿命再评估工作的新认识和新思考, 将服务中国空间站延寿这一长期的系统工程。建议未来加强在轨维修和保障有限支持下的寿命动态评估、在轨原位自主和智能寿命评估、多源数据和失效物理综合驱动的寿命评估、机电产品空间辐射效应量化认知和预警、寿命评估软件工具应用方面的研究和实践。

参考文献 (References)

[1] 夏侨丽, 侯永青, 李伟. 载人航天器安全性、可靠性、维修性技术实践及发展展望[J]. 航天器工程, 2022(6): 86-93.
Xia Q L, Hou Y Q, Li W. Practice and prospects of safety, reliability, maintainability technologies of manned spacecraft[J]. Spacecraft Engineering, 2022(6): 86-93(in Chinese).

[2] 杨宏, 周昊澄. 空间实验室的风险识别与控制方法[J]. 宇航学报, 2019, 40(3): 295-301.
Yang H, Zhou H C. Technology of space lab risk identification and control[J]. Journal of Astronautics, 2019, 40(3): 295-301(in Chinese).

[3] Cecil A J, Robert L P, Raymond S, et al. Extending the international space station life and operability[C]//Proceedings of the Space Communications, Spacecraft Communications, Command and Tracking. Reston: AIAA, 2012.

[4] 刁常堃, 王昊, 孟令子, 等. 航天员舱外维修操作时间的量化评估[J]. 北京航空航天大学学报, 2026, 52(5): 1487-1495.
Diao C K, Wang H, Meng L Z, et al. Quantitative assessment of extravehicular maintenance operation time for astronauts[J]. Journal of Beijing University of Aeronautics and Astronautics, 2026, 52(5): 1487-1495(in Chinese).

[5] 张伟, 侯永青, 夏侨丽, 等. 空间站维修设计与实践[J]. 中国科学: 技术科学, 2022, 52(9): 1332-1344.
Zhang W, Hou Y Q, Xia Q L, et al. Maintenance design and practice for the China Space Station[J]. Scientia Sinica (Technologica), 2022, 52(9): 1332-1344(in Chinese).

[6] Li W, Hou Y Q, An F J, et al. Safety design for the China space station[J]. Space: Science & Technology, 2023, 3: 89.

[7] 肖武平, 王功波, 张蕊, 等. “国际空间站”延寿情况发展研究[J]. 中国航天, 2024(4): 13-16.
Xiao W P, Wang G B, Zhang R, et al. Research on the development of the life extension of the international space station[J]. Aerospace China, 2024(4): 13-16(in Chinese).

[8] 陈露, 张桅, 李孝鹏, 等. NASA国际空间站的安全性、可靠性、维修性及寿命评估[J]. 质量与可靠性, 2020(3): 40-44.
Chen L, Zhang W, Li X P, et al. Safety, reliability, maintainability and life assessment of NASA international space station[J]. Quality and Reliability, 2020(3): 40-44(in Chinese).

[9] Shen L, Mcveigh P, Ziglar M, et al. Extending ISS life beyond 2030[C]//Proceedings of the 75th International Astronautical Congress. Paris: International Astronautical Federation, 2024: 1-12.

[10] Oakley S S, Warren J, Bivens A, et al. International Space Station (ISS) -ongoing assessments for life extension appear to be supported [EB/OL]. (2011-04-11)[2024-08-12]. https://www.gao.gov/products/gao-11-519r.

[11] 陈云霞, 王聪. 产品高可靠长寿命设计与试验技术[J]. 强度与环境, 2023, 50(6): 90-96.
Chen Y X, Wang C. High reliability and long lifetime design and test technology of products[J]. Structure & Environment Engineering, 2023, 50(6): 90-96(in Chinese).

[12] 王莉娜, 史纪军, 李征, 等. 长服役期橡胶密封性能分析与寿命评估[J]. 航天器环境工程, 2022, 39(5): 475-481.
Wang L N, Shi J J, Li Z, et al. Performance analysis and lifetime evaluation of rubber seal for long term service[J]. Spacecraft Environment Engineering, 2022, 39(5): 475-481(in Chinese).

[13] 叶田园, 庞贺伟, 周元子, 等. 巡天任务复杂工况下超静平台作动器疲劳寿命评估[J]. 宇航学报, 2022, 43(10): 1378-1388.
Ye T Y, Pang H W, Zhou Y Z, et al. Fatigue life evaluation of ultra-quiet platform actuator for complex astronomy survey missions[J]. Journal of Astronautics, 2022, 43(10): 1378-1388(in Chinese).

[14] 孙远航, 余建波, 刘贤军, 等. 小子样下导电滑环磨损失效仿真的可靠性评估研究[J]. 宇航学报, 2020, 41(5): 624-632.
Sun Y H, Yu J B, Liu X J, et al. Reliability evaluation of wear failure simulation of conductive slip ring under small sample[J]. Journal of Astronautics, 2020, 41(5): 624-632(in Chinese).

[15] 张兆霖, 杨勇, 黄首清, 等. 空间站电动工具开关按钮性能退化特性加速寿命试验与可靠性评估[J]. 航天器环境工程, 2024, 41(2): 234-243.
Zhang Z L, Yang Y, Huang S Q, et al. Accelerated life test and reliability assessment for degradation characteristics of power tool switch buttons onboard China's Space Station[J]. Spacecraft Environment Engineering, 2024, 41(2): 234-243(in Chinese).

[16] Steve K. Space radiation effects on spacecraft materials and avionics systems: no epic challenge here, but ignoring this can lead to loss of mission and spacecraft[EB/OL]. (2012-06-26)[2024-11-05]. https://ntrs.nasa.gov/api/citations/20120011865/downloads/20120011865.pdf.

[17] Burt R R, Loffi R W. Failure analysis of International Space Station control moment gyro[C]//Proceedings of the 10th European Space Mechanisms and Tribology Symposium. Noordwijk: ESA Publications Division, 2003: 13-25.

[18] Gurrisi C, Seidel R E, Dickerson S, et al. Space station control moment gyroscope lessons learned[C]//Proceedings of the 40th Aerospace Mechanisms Symposium. Washington, D. C. : NASA , 2010: 161-176.

[19] Li X P, Huang H Z, Li F Q, et al. Remaining useful life prediction model of the space station[J]. Eksploatacja i Niezawodność-Maintenance and Reliability, 2019, 21(3): 501-510.

[20] 余溢方, 黄首清, 刘庆海, 等. 航天器流体回路泵智能故障诊断方法[J]. 航天器工程, 2024, 33(3): 61-67.
Yu Y F, Huang S Q, Liu Q H, et al. Intelligent fault diagnosis method for fluid loop pumps in spacecraft[J]. Spacecraft Engineering, 2024, 33(3): 61-67(in Chinese).

[21] 刘凯, 吕从民, 党炜, 等. 可靠性高加速试验技术及其在我国空间站应用领域实施的总体思路[J]. 载人航天, 2017, 23(2): 222-227.
Liu K, Lyu C M, Dang W, et al. Highly accelerated reliability testing and thoughts of its implementation on China manned space station[J]. Manned Spaceflight, 2017, 23(2): 222-227(in Chinese).

[22] Zhang Z L, Huang S Q, Li F Y, et al. Experimental study on multi-stress failure characteristics of space station power tools motor[C]//Proceedings of the Global Reliability and Prognostics and Health Management Conference. Piscataway: IEEE Press, 2024: 1-5.

[23] 黄首清, 代巍, 姚泽民, 等. 两种工程化的航天器用滚动轴承加速寿命试验方法[J]. 航天器环境工程, 2021, 38(4): 413-419.
Huang S Q, Dai W, Yao Z M, et al. Two engineering methods for accelerated life test of aerospace rolling bearings[J]. Spacecraft Environment Engineering, 2021, 38(4): 413-419(in Chinese).

[24] Kamath U, Kuznetsov S, Spencer V, et al. Service life extension of the propulsion system of long-term manned orbital stations[C]//Proceedings of the AIAA Propulsion and Energy Forum and Exposition. Reston: AIAA, 2014.

[25] 叶田园, 刘鑫, 崔赫旻, 等. 空间挠性作动器的可靠性验证试验方法[J]. 中国空间科学技术 (中英文), 2024, 44(2): 118-124.
Ye T Y, Liu X, Cui C M, et al. Reliability verification test method of space flexible actuator[J]. Chinese Space Science and Technology, 2024, 44(2): 118-124(in Chinese).

[26] 叶田园, 庞贺伟, 周元子, 等. 空间可重复锁紧装置可靠性评估[J]. 哈尔滨工业大学学报, 2023, 55(4): 115-121.
Ye T Y, Pang H W, Zhou Y Z, et al. Reliability evaluation of space reusable locking device[J]. Journal of Harbin Institute of Technology, 2023, 55(4): 115-121(in Chinese).

[27] Xu Y W, Kohtz S, Boakye J, et al. Physics-informed machine learning for reliability and systems safety applications: state of the art and challenges[J]. Reliability Engineering & System Safety, 2023, 230: 108900.

[28] Hu Y W, Liu S J, Lu H T, et al. Remaining useful life model and assessment of mechanical products: a brief review and a note on the state space model method[J]. Chinese Journal of Mechanical Engineering, 2019, 32(1): 15.

[29] Li H Q, Zhang Z X, Li T M, et al. A review on physics-informed data-driven remaining useful life prediction: challenges and opportunities[J]. Mechanical Systems and Signal Processing, 2024, 209: 111120.

[30] 杜善亮, 禹志, 周艾玉, 等. 航天器寿命预测与评估方法在嫦娥五号上的应用[J]. 上海航天 (中英文), 2022, 39(S1): 74-80.
Du S L, Yu Z, Zhou A Y, et al. Application of a prediction and evaluation method for Chang'e-5[J]. Aerospace Shanghai (Chinese& English), 2022, 39(S1): 74-80(in Chinese).

[31] 刘志全, 李新立, 遇今. 长寿命航天器机构的加速寿命试验方法[J]. 中国空间科学技术, 2008, 28(4): 65.
Liu Z Q, Li X L, Yu J. Methods of accelerated life tests for long-life spacecraft mechanisms[J]. Chinese Space Science and Technology,

2008, 28(4): 65(in Chinese).

[32] 朱炜, 张国斌, 陈凤熹, 等. 长寿命卫星寿命工作方案研究[J]. 质量与可靠性, 2017(1): 22-26.
Zhu W, Zhang G B, Chen F X, et al. Research on long-life satellite life work scheme[J]. Quality and Reliability, 2017(1): 22-26(in Chinese).

[33] 黄首清, 郭振伟, 刘庆海, 等. 加速寿命试验设计与评估软件 ALT511 及其应用 (一)[J]. 航天器环境工程, 2023, 40(5): 541-549.
Huang S Q, Guo Z W, Liu Q H, et al. Accelerated life test design and evaluation software ALT511and its application(Ⅰ)[J]. Spacecraft Environment Engineering, 2023, 40(5): 541-549(in Chinese).

[34] 黄首清, 刘庆海, 李芳勇, 等. 加速寿命试验设计与评估软件 ALT511 及其应用 (二)[J]. 航天器环境工程, 2023, 40(6): 641-649.
Huang S Q, Liu Q H, Li F Y, et al. Accelerated life test design and evaluation software ALT511 and its application (Ⅱ)[J]. Spacecraft Environment Engineering, 2023, 40(6): 641-649(in Chinese).

[35] Sankararaman S, Goebel K. Uncertainty quantification in remaining useful life of aerospace components using state space models and inverse FORM[C]//Proceedings of the 54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. Reston: AIAA, 2013.

[36] Sharma G, Rai R N. Reliability modeling and analysis of environmental control and life support systems of space stations: a literature survey[J]. Acta Astronautica, 2019, 155: 238-246.

[37] 皇甫素伟, 文明, 李毅, 等. 卫星在轨寿命与可靠性数据库体系建设及应用[J]. 质量与可靠性, 2018(3): 27-31.
Huangfu S W, Wen M, Li Y, et al. Construction and applications of the lifetime and reliability database architecture for on-orbit satellites[J]. Quality and Reliability, 2018(3): 27-31(in Chinese).

[38] 周正伐. 可靠性工程技术问答 200 例[M]. 北京: 中国宇航出版社, 2011.
Zhou Z F. 200 questions and answers of reliability engineering technology[M]. Beijing: China Astronautic Publishing House, 2011(in Chinese).

[39] 朱炜. 航天产品可靠性评估方法[M]. 北京: 中国宇航出版社, 2022.
Zhu W. Reliability evaluation method of aerospace products[M]. Beijing: China Astronautic Publishing House, 2022(in Chinese).

[40] 张兆霖. 舱外电动工具多应力失效边界域与寿命预测研究[D]. 北京: 中国空间技术研究院, 2024.
Zhang Z L. Research on multi-stress failure envelope and life prediction of extravehicular power tools[D]. Beijing: China Academy of Space Technology, 2024(in Chinese).

[41] Zhang L W, Zheng H M, Huang S Q, et al. A life test of ultrasonic motors under different torque loads and the analysis of the characteristics of wearing surfaces[J]. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 2020, 234(5): 770-777.

[42] 陈云霞, 金毅. 机械产品寿命设计与试验技术[M]. 北京: 国防工业出版社, 2022.
Chen Y X, Jin Y. Life design and test technology of mechanical product[M]. Beijing: National Defense Industry Press, 2022(in Chinese).

[43] 中国人民解放军总装备部电子信息基础部. 可靠性鉴定和验收试验: GJB 899A—2009[S]. 北京: 中国人民解放军总装备部, 2009: 118-243.
Chinese People's Liberation Army General Armament Department Electronic Information Basic Department. Reliability testing for qualification and production acceptance: GJB 899A—2009[S]. Beijing: Chinese People's Liberation Army General Armament Department, 2009: 118-243(in Chinese).

[44] 蒋海青, 何伟荣, 蔡锡明, 等. 钴源与加速器辐照效应及辐照加工试验综合平台[J]. 辐射研究与辐射工艺学报, 2022, 40(1): 106-111.
Jiang H Q, He W R, Cai X M, et al. Integrated 60Co and electron accelerator platform for irradiation effect and irradiation processing tests[J]. Journal of Radiation Research and Radiation Processing, 2022, 40(1): 106-111(in Chinese).

[45] Bansal N, Arora S. Exploring the impact of gamma rays and electron beam irradiation on physico-mechanical properties of polymers & polymer composites: a comprehensive review[J]. Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 2024, 549: 165297.

[46] Harbin Institute of Technology. SESRI: bringing space to earth[EB/OL]. (2022-10-15)[2024-11-05]. https://www.nature.com/articles/d42473-020-00198-6.

[47] Xiang G, Tian K. Spacecraft intelligent fault diagnosis under variable working conditions via Wasserstein distance-based deep adversarial transfer learning[J]. International Journal of Aerospace Engineering, 2021, 2021(1): 6099818.

[48] Baireddy S, Desai S R, Mathieson J L, et al. Spacecraft time-series anomaly detection using transfer learning[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE Press, 2021: 1951-1960.

[49] Chen A, Xie Y C, Wang Y, et al. Knowledge graph-based image recognition transfer learning method for on-orbit service manipulation[J]. Space: Science & Technology, 2021, 2021: 9807452.

[50] 王正义, 尹钊, 杨金禄, 等. 载人空间站工程航天技术试验领域项目遴选程序初探[J]. 航天器环境工程, 2020, 37(5): 452-457.
Wang Z Y, Yin Z, Yang J L, et al. Procedures of project selection of aerospace technology experiment for manned space station[J]. Spacecraft Environment Engineering, 2020, 37(5): 452-457(in Chinese).

On-orbit life re-evaluation technology and prospects for electromechanical products in response to demand for China Space Station lifespan extension

HUANG Shouqing[1, 2], XIA Qiaoli[3], XU Shu[1, 2, *], YU Yifang[1, 2], LIU Shouwen[1, 2], ZHANG Zhaolin[1, 2]

(1. Beijing Institute of Spacecraft Environment Engineering, Beijing 100094, China;
2. National Key Laboratory of Reliability and Environmental Engineering Technology, Beijing 100094, China;
3. Beijing Institute of Spacecraft System Engineering, Beijng 100094, China)

Abstract: This paper systematically reviews an on-orbit life re-evaluation technology specifically designed for electromechanical products in response to the urgent need for quantitatively re-evaluating the on-orbit life of electromechanical products amid the lifespan extension of China Space Station. It does this by drawing on successful experiences from International Space Station lifetime evaluations and extensions as well as methods used during the development phase of China Space Station. Furthermore, we present technical development trends and offer pertinent recommendations. The suggested technology falls into two main categories: those that combine on-orbit and ground telemetry data and those that are mostly based on ground test data. It encompasses 9 specific techniques, among which are the traditional correction of lifetime evaluation outcomes, the re-evaluation of lifetime tests factoring in both integrated operations and degradation effects, complex dynamic profile re-design for lifetime tests, multi-level accelerated lifetime testing and evaluation, the re-evaluation of lifetime in integrated irradiation environments, the re-screening and re-evaluation of multi-source data stemming from both ground and space, intelligent fault diagnosis and degradation prediction, lifetime evaluation leveraging transfer learning between ground and on-orbit telemetry data, and the re-evaluation of lifetime based on on-orbit testing. This technology embodies innovative considerations and concepts for the re-evaluation of electromechanical product lifetimes as the China Space Station embarks on its protracted on-orbit operational phase. It stands poised to bolster the long-term and systematic endeavor aimed at extending the operational lifespan of the China Space Station.

Keywords: China Space Station; electromechanical products; on-orbit life re-evaluation; ground test data; on-orbit telemetry data

Received: 2024-11-07; **Accepted**: 2025-01-10; **Published Online**: 2025-02-11 12:30

URL: link.cnki.net/urlid/11.2625.V.20250211.1056.001

Foundation item: Stable Support Scientific Research Project of National Key Laboratory of Reliability and Environmental Engineering Technology (614200420230201)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2026.0116

面向复杂飞行的气动/运动/控制耦合数值虚拟飞行研究综述

邵帅*，梅亚飞，姚梦玉，高显忠，包磊
(国防科技大学试验训练基地，西安 710106)

摘 要：数值虚拟飞行(NVF)技术能够高保真地模拟复杂飞行环境下飞行器的动态响应特性，是飞行器控制律设计与飞行品质评估的关键技术。系统综述了气动、运动和控制耦合的NVF技术，重点讨论了该耦合方法在复杂飞行问题中的应用。梳理了NVF的基本框架与技术流程，重点阐述了动网格技术从单一变形向动态混合网格体系的演进历程，及其在保障部件相对运动与计算稳定性中的关键作用，并指出机器学习在网格质量控制中的赋能潜力。针对非定常流场求解，探讨了气动、运动与控制的耦合方式，介绍了从松耦合到紧耦合闭环求解的技术演变，指出紧耦合方法可显著提升对大迎角、过失速等高动态飞行的仿真精度。详细阐述了该耦合方法在高超声速多场耦合、变构飞行、柔性仿生等三大场景中的应用实践。面向未来，在人工智能技术的赋能中，NVF将进一步向高效统一、全系统多场耦合的方向发展，加速飞行器的研制模式向“仿真驱动、数智试验”转型。

关 键 词：复杂飞行；多体分离；柔性变构型；多物理场耦合；非定常流动
中图分类号：V211.3
文献标志码：A **文章编号**：1001-5965(2026)10-3437-16

现代空战环境愈发复杂，传统有人战机与固定翼无人机亟需高敏捷复杂机动能力应对复杂高威胁场景。变构飞行器、柔性仿生飞行器的快速发展，也不断提升着控制系统的设计难度。这些设备在开展整机过失速机动、仿生翅/翼扑动、躯体连续形变、机体折叠倾转等多模式运动与构型变换时，将与流体环境产生强耦合作用，形成旋涡、流动分离等复杂流态，衍生出建模难度极大的非定常非线性流体作用力，给其动力学建模仿真与飞行控制策略设计带来严峻挑战。传统的流体作用力静态数据库和降阶模型、纵向与横航向拆分的动力学建模方式，已无法完全满足现代飞行器对高难度复杂运动控制的需求[1]。以非定常数值模拟为核心，开展气动/运动/控制等多学科耦合的一体化数值方法——数值虚拟飞行(numerical virtual flight, NVF)技术，为此类非线性、强耦合问题的高保真仿真提供了有效途径。

通过高保真的计算流体动力学(computational fluid dynamics, CFD)数值仿真，NVF技术能够在复杂飞行条件下，实现气动、运动及控制系统的耦合分析，精确预测飞行器的气动载荷、动态响应特性，从而支撑部分高性能战机在过失速状态下完成如“J-Turn”等复杂机动的控制律设计[2]。相比传统实装飞行试验方式，该技术为飞行器控制律设计提供了经济、安全、高效、可靠的虚拟验证手段。

收稿日期：2026-03-02；**录用日期**：2026-05-08；**网络出版时间**：2026-06-05 18:40
网络出版地址：link.cnki.net/urlid/11.2625.V.20260605.1251.001
基金项目：陕西省自然科学基础研究计划一般项目(2024JC-YBQN-0030)；国防科技大学自主创新科学基金项目(25-ZZCX-JDZ-66)；国防科技大学青年自主创新科学基金项目(ZK25-106)
* **通信作者**. E-mail：shaoshuai_ss@nudt.edu.cn

引用格式：邵帅，梅亚飞，姚梦玉，等. 面向复杂飞行的气动/运动/控制耦合数值虚拟飞行研究综述[J]. 北京航空航天大学学报，2026，52(10)：3437-3452. Shao S，Mei Y F，Yao M Y，et al. A review of numerical virtual flight studies for aerodynamic, motion and control coupling in complex flight conditions[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3437-3452 (in Chinese).

随着计算能力的提升和数值方法的发展，NVF技术已由早期的离线分析发展为紧耦合实时闭环的求解框架。通过气动、运动与控制系统在统一时域内的协同求解，NVF在处理战机复杂机动飞行、飞行器跨声速和高超声速飞行多物理场耦合、多体分离、变构/仿生飞行器运动控制等问题上，展现出极大应用潜力。本文综述了气动/运动/控制耦合的NVF技术，讨论了当前技术的进展与面临的挑战，分析了其在复杂飞行任务中的应用场景，并展望了未来的发展趋势。

1 NVF技术的发展脉络

NVF技术最早源于飞行器的飞行仿真需求，主要依赖批量飞行试验、风洞试验或CFD仿真获得气动力和力矩数据，构建气动数据库，为基于飞行动力学模型的飞行性能分析和控制仿真提供必要输入[3-4]。然而，随着飞行器设计复杂性和控制精度要求的提升，传统的离线耦合方法在超机动、变构飞行、柔性仿生等复杂飞行场景中，无法有效刻画强非线性效应、迟滞现象及多物理场耦合效应[5]。此外，传统风洞试验和飞行试验在飞行工况覆盖范围、试验安全性、数据获取全面性及成本等方面存在明显局限。因此，单纯依赖传统的试验手段和静态气动数据库已经无法满足现代飞行器的控制系统设计需求[6]。

为弥补这些不足，虚拟飞行试验（virtual flight test, VFT）技术应运而生，通过在风洞试验中引入飞行器的运动自由度和控制律评估能力，在地面试验环境中提供了动态响应预测和控制系统验证的有效手段[7]。随着计算能力和多学科耦合算法的快速发展，由CFD取代风洞试验，提供飞行器运动状态下的气动力和力矩，VFT演进为NVF。这将飞行仿真向前推进一步，由离线分析模式向实时紧耦合的闭环求解阶段演进。这一演进不仅使气动、运动和控制系统能够在统一的时域内协同求解，还显著提升了对多场耦合等复杂飞行物理问题的预测能力，满足了飞行器研发对高精度、全面性能评估的需求[8]。为系统对比不同飞行特性预测方法的特点，本文总结了各方法的应用场景与主要局限，如表1所示。

表1 不同飞行试验与NVF方法的应用与局限对比

Table 1 Comparison of applications and limitations between different flight tests and NVF methods

方法	应用	局限
风洞试验[1,9-10]	用于获取典型构型在不同来流工况下的流场结构与气动特性，为气动机理分析和模型验证提供试验依据	工况覆盖有限，难以同时复现高马赫数、大雷诺数、气动力热耦合等极端飞行状态，多体分离与构型快速变化模拟能力受限
虚拟飞行试验[1,7,11]	在地面风洞试验环境中引入多自由度模型运动控制能力，用于动态响应和控制系统鲁棒性验证	受风洞试验条件限制，对强非线性气动/运动耦合效应、高超声速力热耦合效应复现能力不足，飞行工况全包线覆盖能力不足
数值虚拟飞行-离线耦合阶段[2-4,10]	基于CFD仿真结果构建气动数据库或代理模型，为飞行仿真提供气动输入	气动、运动与控制数据无法实时反馈，对强非线性、迟滞效应及极端机动场景预测能力有限
数值虚拟飞行-紧耦合阶段[12-16]	在统一时域内耦合气动、运动与控制系统求解，可用于复杂机动、变构飞行及多场耦合问题的仿真与验证	计算成本高，对动网格质量、时间推进精度与耦合算法稳定性要求严格，工程化应用仍依赖高性能计算支撑

紧耦合的NVF计算要求气动计算、运动求解与控制律解算需要在同一时间步内同步推进，这对计算网格的动态更新能力提出了极高要求。对于大迎角机动、快速姿态变化及构型调整等复杂工况，计算网格不仅需随飞行器完成大幅度的平移和旋转，还必须保持足够的网格质量，避免网格畸变引发数值误差或计算发散。因此，动网格生成与更新技术成为NVF的基础技术之一，其核心在于保持近壁面网格质量的同时，支持大范围运动、构型变化及多体相对运动[8,17]。从计算实现角度看，动网格的质量直接决定了紧耦合NVF的可行性、数值稳定性和工程适用性。如表2所示，NVF技术中的动网格技术经历了从刚性运动处理、连续网格变形到重叠/嵌套网格，再到变形-重构耦合动态混合网格的演进。

按网格类型分，动网格方法分为针对结构网格和非结构/混合网格两类[18]方法。结构网格计算效率高，适用于刚体运动和多体相对运动等问题；非结构网格/混合网格通过网格变形与局部重构，适合处理大幅机动、复杂外形及构型变化等问题[19]。动网格技术的应用使非定常CFD能够直接参与飞行器机动过程仿真，准确捕捉大迎角下流动分离、漩涡结构演化及气动力时变特性[20]。在此基础上，将动网格、非定常CFD、刚体动力学（rigid body dynamics, RBD）方程求解及控制律耦合，可在统一框架内实现气动-运动-控制的闭环求解，为复杂机动飞行提供了更可靠的数值分析手段[14]。

NVF技术的应用领域持续拓展，研究对象已不限于传统刚体构型飞行器。近年来，变构飞行器、柔性仿生飞行器及仿生水下航行器等已成为航空、航天、航海领域的研究前沿。该类系统依赖构型的连续变化、柔性结构的大变形或周期性扑动来实现推进与机动[21]。其运动过程伴随着强非定常流动、显著的结构/机体响应及复杂的控制反馈[22]。上述

表 2　NVF 中动网格技术的发展历程与演进方向

Table 2　Development history and evolutionary direction of dynamic mesh technology in NVF

时间历程	处理方式	代表方法
20世纪90年代	刚性运动与接口处理	刚性运动网格, 滑移网格, 多参考系, 多块结构网格接口处理
21世纪初期	连续网格变形方法	代数插值类(超限插值、反距离加权), 物理类(弹簧法, 线性弹性体法, 拉普拉斯光顺法), 全局插值类(径向基函数、Delaunay图映射位移传播)
21世纪中后期	重叠/嵌套网格体系	重叠网格, 嵌套网格, 孔洞切割, 供体搜索, 守恒插值; 背景网格与浸入边界法(笛卡儿网格、切割单元、浸入边界法、自适应网格加密)
21世纪10年代	动态混合网格(工程主流)	变形+局部重构, Delaunay映射, 局部拓扑操作(翻转/分裂/合并) 近壁面网格质量控制(边界层层状网格保持、各向异性网格保持, 网格质量监测与重构触发)
2018年-至今	智能化与多物理场一致网格	智能方法(机器学习网格质量评估、图神经网络网格判别、强化学习驱动自适应网格); 多物理场自适应(h/r/p自适应、流-固-热耦合一致更新、柔性变构型与仿生系统)

系统的流体动力特性、结构响应与控制行为在时域内高度耦合, 显著削弱了定常或低阶模型在性能预测和控制验证中的适用性[23]。基于动网格的非定常计算方法和气动-运动-控制紧耦合 NVF 框架为这些系统的运动机理分析、性能评估及控制策略设计提供了高保真且统一的数值试验手段[24]。

随着高性能计算和数据驱动方法的发展, NVF 技术正从离线分析工具逐步演化为支撑飞行器方案设计、控制律优化和虚拟验证的综合数字化试验平台, 为未来高机动、强耦合飞行任务提供了坚实的技术支撑。

2　动网格技术研究进展

动网格技术作为 NVF 的核心技术基础, 近年来已取得显著进展。张来平等[25] 在 2014年对基于结构网格与非结构/混合网格的两大技术路线进行了全面综述, 系统总结了动网格技术的发展现状, 并指出未来的研究趋势应集中于动态混合网格技术的工程化应用与智能化发展。随着飞行器设计复杂度的增加和对高精度流动分析需求的提升, 动网格技术逐步从传统的简单网格变形方法, 向更高效、更精确的动态混合网格方法过渡, 成为 NVF 技术发展的关键支撑。

2.1　近壁面与复杂构型的网格生成技术

在非定常流动数值模拟中, 近壁面区域伴随强剪切作用, 形成边界层、激波、流动分离和再附着等复杂物理现象, 这对网格的几何质量和分辨率提出了更高的要求。特别是在大迎角机动、舵面偏转等问题中, 计算网格不仅需要在飞行器运动或构型变形过程中实时更新, 还必须确保边界层网格的正交性和层状结构的完整性, 以避免网格畸变造成非物理解, 从而确保数值计算的稳定性与准确性。

为应对这一挑战, 近壁面动网格技术的发展显得尤为重要。目前, 近壁面动网格技术主要包括以下几种方法: 基于物理模型的线性弹性体法与拉普拉斯光滑法[26], 常用于中等幅度、连续变化的壁面运动场景, 通过对边界节点的平滑处理, 有效维持网格的正交性与层状结构, 在小范围几何变形下表现出较好的数值稳定性; 基于代数插值的超限插值(transfinite interpolation, TFI)方法[27], 通常用于结构化或多块结构网格体系, 适用于几何构型与网格拓扑连接关系保持不变、仅发生舵面偏转、部件俯仰等小幅刚体姿态变化的问题, 具有较高计算效率; 在处理更为复杂的网格变形问题时, 径向基函数(radial basis function, RBF)方法[28] 通过构建边界节点与内部节点之间的全局插值关系, 实现位移场的平滑传播, 具有几何适应性强、对网格拓扑依赖性低等优势, 特别适用于近壁面大变形及复杂构型演化问题。孙岩等[29] 提出了一种基于 RBF 插值的增强型动网格技术, 利用 RBF 的空间插值能力, 恢复和传递 Delaunay 背景网格在物面大变形时应有的运动特性, 如图 1 所示, 避免近壁面网格的严重挤压, 生成了 NACA 0012 翼型弯曲、NLR 7301 两段翼偏转及 M6 机翼扭转等大变形中高质量近壁面网格。该团队在后续的研究中[30] 进一步提出了 RBF-TFI 混合结构动网格技术, 通过精简 RBF 基点控制结构网格块棱线节点的变形, 再利用 TFI 方法高效重构面网格与体网格, 将该技术应用于 F6 翼身组合体风洞静气动弹性修正, 如图 2 所示。

随着多体分离、变构、柔性仿生等极端变形问

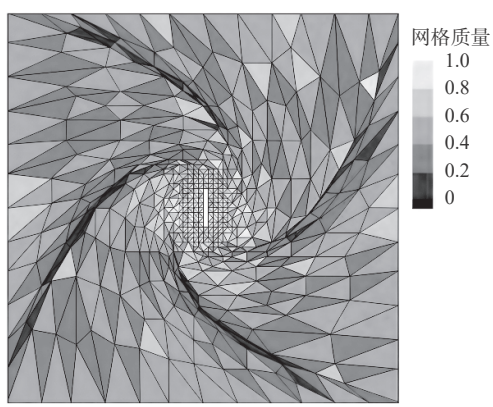


图 1　矩形旋转角度为 90°时变形网格质量分布[29]

Fig. 1　Deformation quality distribution of the mesh at a 90° rotation angle of the rectangle[29]

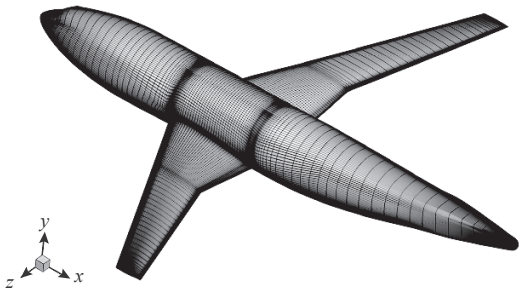


图2 F6翼身组合体外形和表面网格[30]

Fig. 2 F6 wing-body assembly geometry and surface mesh[30]

题的深入研究，连续网格变形与局部重构相结合的混合更新策略逐步成为满足高精度空间离散需求的重要方案。该方法通过在网格质量指标超限时触发局部重构，从而在保证计算效率的同时优化网格质量。该策略的有效实施依赖于网格质量的实时监控与控制机制，这推动了网格质量控制技术的系统化发展，使动网格技术应用更加广泛。

2.2 重叠网格与网格质量控制技术

在多体运动问题的网格生成技术中，尤其是导弹分离、副油箱抛撒、直升机旋翼与机身干扰等外流场应用场景，已形成以重叠网格和嵌套网格为核心的技术路线。该方法采用背景网格覆盖整个计算域，背景网格通常采用非结构网格或笛卡儿网格，以保证生成效率和几何适应性[31]；运动部件的近壁区域则采用结构化或半结构化网格以精确捕捉边界层流动；外围过渡区域由非结构网格填充[32]。不同网格组件之间通过“挖洞”操作形成重叠区域，并建立插值关系来实现流场信息的有效传递，从而确保多物体相对运动的过程中，流场的动态变化被准确捕捉。段旭鹏等[33]提出一种基于Delaunay背景网格插值与局部网格重构的动态混合网格生成方法。该方法将运动轨迹、动网格生成与非定常流场的耦合计算相结合，有效解决了复杂多体分离的数值仿真问题。在重叠区的处理中，Gillebaart等[34]采用智能控制点选取与边界误差修正相结合的双重策略，显著提高了重叠区网格处理的精度与效率。常兴华等[35]建立了并行环境下重叠网格与RBF变形网格相结合的动态混合网格生成技术，并通过鱼游（见图3）与战斗机分离投放（见图4）算例，验证了该动态混合网格技术及集成算法的实用性。中国空气动力研究与发展中心[36]发布的国家数值风洞工程指出，对于移动范围广、分体数量多的多体分离问题，混合重叠网格技术是现阶段和未来多体分离数值仿真中网格技术发展的主要趋势。在工程实践中，重叠网格通常作为动态混合网格体系的重要组成，与网格变形、局部重构协同使

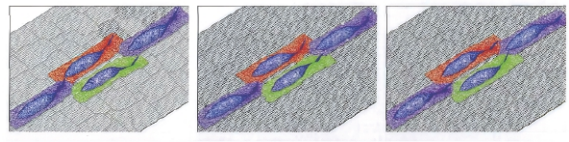


图3 鱼群游动过程中的动态混合网格[35]

Fig. 3 Dynamic mixed mesh for fish school movement[35]



图4 战斗机外挂物分离投放过程的动态混合网格[36]

Fig. 4 Dynamic hybrid mesh for separation and release process of fighter external stores[36]

用，以应对多体分离过程中大位移与复杂拓扑变化带来的挑战。

随着网格更新策略由单纯变形向“变形+重构”混合模式演进，网格质量监测与控制体系成为确保数值模拟精度与稳定性的关键环节。在自动网格生成过程中，二维网格质量通常采用长细比、单元内角、偏斜度及雅可比矩阵等指标进行评估，三维四面体网格则采用体积偏斜度和体积塌陷率等指标进行度量[37]。Sorgente等[38]针对网格质量评估指标的综述表明，现代网格质量评估体系已发展为涵盖局部网格几何与全局网格拓扑的多维度体系，关注长细比等几何指标、雅可比矩阵条件数等代数指标，以及网格与计算域的几何一致性、单元尺寸的分布均匀性等全局质量指标。该综述指出，不存在适用于所有场景的普适最优指标，指标选择应依据具体应用需求而定。

针对网格形变过程中的质量退化问题，研究者提出了多种主动干预策略。一方面，通过网格光顺与变形方法延缓质量劣化。线性弹性体法[26]在网格轻度畸变时触发网格光顺，通过对节点位移的平滑分配延缓质量下降。Secco等[39]提出了一种基于逆距离加权算法的网格变形方法，将空间网格节点视为独立点云，实现较大边界位移条件下仍能维持合理的单元形态。当网格变形无法解决严重单元畸变时，引入局部重构策略。李绍磊等[40]针对飞行器多体分离数值仿真中非结构网格框架下局部网格重构问题，提出了一种基于单元相邻关系的重构区域构造方法，该方法能够精确计算分离物的运动轨迹与姿态。杨宇辰[41]建立了一套完整的自适应网格技术体系，通过发展自适应笛卡儿网格的壁面函数-浸没边界法（wall fuction-immersed boundary method, WF-IBM）、自适应贴体网格的动态加密技

术, 以及并行混合网格自动装配方法, 有效解决了可压缩流动中激波、旋涡等特征的精确捕捉难题。

综上所述, 当前“变形+重构”混合策略的重叠网格和嵌套网格技术, 通过多维度质量监测体系与系统级鲁棒性控制机制的协同作用, 为高精度非定常流动仿真提供了坚实的技术支撑。

2.3　机器学习在动网格中的应用与挑战

随着人工智能技术的发展, 机器学习方法逐渐应用于动网格生成与质量控制领域。在处理大变形和复杂几何形状等问题时, 机器学习技术通过数据驱动的方式提升了网格质量评估、自动化调整和优化能力。Haubner 等[42] 针对流固耦合与形状优化中的网格退化问题, 提出了混合偏微分方程-神经网络(partial differential equations- neural network, PDE-NN)的架构, 通过学习生成理想网格运动的算子, 成功应用于流固耦合基准测试中。刘翰林等[43] 结合神经网络与阵面推进法, 提出一种基于法线方向的节点筛选算法, 该算法优化了神经网络对网格数据的处理量, 提高了模型的计算效率, 尤其在复杂几何结构和大变形问题中表现出较强的适应性。

在网格质量评估领域, 传统方法主要依赖单元几何特征, 但往往忽略了拓扑结构对数值误差的综合影响。近年来, 图神经网络(graph neural network, GNN)因其天然契合网格图结构的特性, 在网格质量评估中得到广泛应用。Wang 等[44] 基于 GNN 构建了网格质量评估模型, 并在 NACA-Market、NACA 6510 等公开数据集上进行训练与验证。研究表明, 该方法能够高精度预测雅可比矩阵条件数等复杂质量指标, 并能够识别由局部拓扑连接不当引起的数值不稳定区域。该方法在局部几何度量和全局质量场感知之间建立了重要联系, 为网格质量评估提供了新的视角。

在自适应网格生成方面, Foucart 等[45] 将自适应网格细化与部分可观测马尔可夫决策过程相结合, 利用深度强化学习直接从数值模拟中学习局部优化策略。该方法避免了对高保真解和预置数据集的依赖, 还展现出较强的泛化能力, 能够在多种偏微分方程和离散格式下实现精度与效率的平衡, 提升了自适应网格技术在复杂流场中的应用范围和效率。

如表 3 所示, 基于机器学习的方法在网格质量评估和自适应调整中具有更高的效率, 并能减少对人工经验的依赖。尽管如此, 机器学习方法在大变形等复杂工程问题中的适用性尚未得到全面验证, 其可靠性仍然受限于特定的训练数据和应用场景。因此, 尽管机器学习方法在有限问题中表现优异, 但仍需与传统的物理模型结合, 才能在实际工程中取得更广泛的应用。

表 3　传统动网格方法与机器学习方法在网格更新中的对比

Table 3　Comparison of traditional dynamic mesh methods and machine learning methods in mesh updates

要素	传统动网格方法	基于机器学习的方法
方法	根据几何关系和物理模型直接计算网格运动	通过已有数据学习网格变化或质量特征
作用对象	网格节点位移、单元变形和局部重构	网格质量、变形趋势或更新策略
稳定性与守恒性	可以明确控制, 结果可预测	本身不保证, 需要额外约束
应对大变形和多体运动	通过混合变形、重构或重叠网格实现	通过模式识别给出辅助判断
对新工况的适应性	规则固定, 但行为稳定	依赖训练数据, 需重新训练
当前应用情况	已用于工程问题 NVF 仿真	主要用于标模算例验证

经过数十年的发展, 动网格技术已由单一的运动适配方法演变为支撑多场、多体耦合仿真的核心基础技术。在基础方法层面, RBF 类方法可有效应对近壁面大变形问题, 重叠网格方法广泛应用于多体分离问题。在质量控制层面, 多维度监控与响应机制已用于保障网格质量稳定性。在前沿探索方面, 机器学习等人工智能技术正逐步融入质量评估与策略优化环节, 推动动网格技术向更自主、更高效的方向演进。随着计算能力的提升与算法的持续创新, 机器学习有望在动网格质量控制与自适应调整中发挥更加关键的作用, 为复杂几何外形、大变形、多体分离等问题提供更加可靠的技术支持。

3　基于动网格的非定常计算方法研究进展

基于动网格的非定常计算方法是求解运动或变形边界条件下非定常流场的核心技术, 广泛应用于飞行器动态仿真领域。非定常计算方法的研究主要围绕以下方面展开: 提高时间和空间离散格式的精度, 以及探索机器学习等新兴技术在计算效率提升中的应用。

3.1　时间离散格式

时间离散格式直接影响非定常流动模拟精度与计算效率, 其发展的核心是保证数值稳定性的同

时提高时间推进效率。目前，主流的时间离散方法可分为显式格式与隐式格式。显式方法以经典龙格-库塔(Runge-Kutta, RK)方法为代表，利用当前时刻的流场信息直接外推下一时刻流场状态，无需构建和求解大型线性方程组，因此，单步计算效率高、存储需求低。然而，显式格式受限于苛刻的稳定性条件和动量插值的高计算成本，在真实工程问题中应用较少[46]。隐式方法如向后差分格式，通过联立多个时间层流场控制方程，构建大型稀疏线性系统进行求解，数值稳定性强，对时间步长的限制较为宽松，适用于长时间演化流动问题及刚性较大的非定常流动问题的求解[47]。双时间步法通过在物理时间层内引入伪时间迭代，将非定常问题转化为一系列定常子问题逐步求解。在伪时间迭代中，通常采用下上对称高斯-赛德尔(lower-upper symmetric Gauss-Seidel, LU-SGS)或广义最小残差(generalized minimal residual, GMRES)等格式作为迭代器以加速收敛，这些方法在较低马赫数下刚体运动与流体相互作用问题的模拟中具有显著效率优势[48]。

3.2 空间离散格式

空间离散方法决定了控制方程在计算域中的近似精度及其对复杂几何的适应能力，其发展经历了从单一方法到混合框架的演进，主要包括有限差分法、有限体积法和有限元法三类主流方法。有限差分法基于泰勒展开对控制方程中的导数项进行离散，数学形式简洁，但通常依赖结构化网格。由于复杂几何外形下生成高质量结构化网格存在较大困难，该方法目前主要应用于基础研究和规则区域流动模拟[49]。有限体积法基于积分型控制方程，通过对单元界面通量进行数值积分，保证离散格式的守恒性，对网格类型适应性强，已成为动网格非定常模拟中应用最为广泛的方法[50]。在飞行器机动[51]、多体分离[52]等具有大范围运动边界的复杂流动问题中表现优异。有限元法采用分段连续多项式逼近解函数，对复杂几何边界拟合能力强，便于实施局部网格自适应。该方法在流固耦合[53]、结构振动[54]及边界层流动[55]等模拟中具有独特优势，常作为有限体积法的补充手段应用于多场耦合问题。为提升空间分辨能力并有效捕捉非定常流动中的多尺度结构，近年来，通量重构法和间断伽辽金法(discontinuous Galerkin, DG)等高阶精度数值格式得到了广泛关注。通量重构法[56]融合了有限体积法的守恒性与有限元的高精度逼近能力，通过在单元界面引入通量修正函数，实现整体守恒与局部高精度模拟。该方法适用于结构网格与非结构网格，具有良好的并行效率，已在气动噪声[57]、湍流模拟[58]等复杂非定常问题中得到应用。间断伽辽金法允许单元间解存在间断，结合高次多项式基函数，在激波等强间断流动中具有高分辨率优势[59]。近年来，间断伽辽金法与动网格技术的结合，在飞行器颤振、动态失速[60]等涉及运动边界的高精度非定常模拟中取得了显著进展。

3.3 机器学习在非定常计算中的应用与挑战

机器学习技术为非定常计算提供了新型优化路径，主要体现在降阶模型构建和计算参数优化方面。传统CFD模拟计算成本高昂，难以支撑大规模复杂仿真问题；机器学习方法通过学习输入参数与流场响应之间的映射关系，能够在一定程度上替代部分CFD求解过程。研究表明，机器学习在网格自适应调整、流场特征识别及时间步长与网格更新频率的动态优化方面，展现出提升仿真效率与精度的潜力。Yan等[61]构建了基于机器学习的降阶模型，通过建立输入参数与流场响应之间的非线性映射关系，实现了对部分CFD求解过程的有效替代，显著提升了气动数据获取与建模效率。此外，机器学习方法可根据流场特征自适应调整关键数值参数，如时间步长与网格更新频率。李彩云等[59]提出一种基于反向传播神经网络(back propagation neural network, BPNN)与移动网格偏微分方程法(moving mesh partial differential equation, MMPDE)的网格自适应方法，如图5所示。该方法利用训练后的神经网络预测流场中的间断量，并据此驱动网格进行一次性自适应调整，从而在保证网格质量的同时，显著提高了非定常计算的效率。

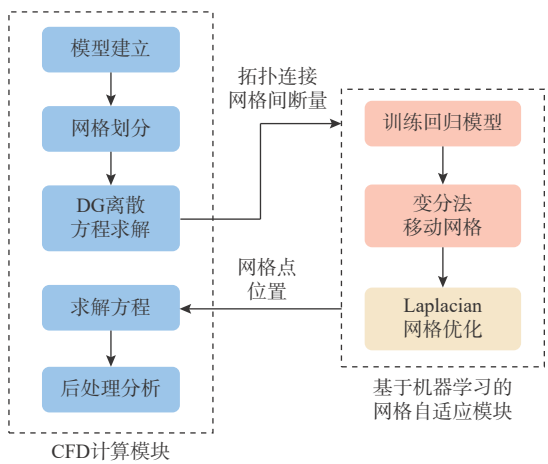


图5 网格自适应算法迭代的整体流程[59]

Fig. 5 Overall process of iterating the grid adaptive algorithm[59]

随着动网格上非定常计算方法的持续发展，该领域的研究重点已从单一的运动适配逐步演进为支持多场、多体、多尺度耦合仿真。在此过程中，时间离散格式与空间离散方法的改进，以及耦合求解策略的创新，为复杂飞行系统的仿真提供了更为

精准和高效的工具。与此同时, 机器学习作为一项新兴技术, 正在逐步融入基于动网格的非定常计算框架, 推动数值模拟方法向智能化与自主化方向演进。然而, 尽管机器学习技术具有极大的潜力, 其广泛应用仍受数据质量、模型泛化能力和计算等因素的制约。因此, 如何在未来研究中进一步完善智能算法与传统数值方法的深度融合, 仍是一个重要课题。

4　气动/运动/控制耦合计算与置信度评估方法研究进展

动网格技术与非定常计算方法构成了 NVF 技术的基础, 而气动/运动/控制耦合计算方法则作为其核心, 近年来实现了从算法改进到工程应用的重要跨越。随着计算能力的提升与算法的逐步成熟, 气动、运动与控制系统的耦合计算逐渐实现了从离线分步耦合到实时闭环紧耦合的转变, 为复杂飞行器的高精度仿真及多物理场耦合问题提供了有效的技术支撑。

4.1　离线分步耦合与实时闭环紧耦合方法

早期离线耦合方法采用“气动计算-运动分析”的分步执行模式, 首先, 通过 CFD 求解欧拉方程或雷诺平均 Navier-Stokes 方程获取非定常气动数据, 建立气动数据库; 然后, 将其输入飞行动力学模型进行动力学仿真。该方法在常规飞行状态下的运动特性分析中取得了良好效果, 如自旋稳定弹丸的动态稳定性导数预测[4]。然而, 由于无法实时反馈流场与运动、控制之间的动态交互, 该方法难以准确仿真大迎角、超机动等强非线性工况。对于涉及运动边界的非定常问题, CFD-RBD/计算结构动力学(computational structural dynamics, CSD)-飞行控制系统(flight control system, FCS)方程的耦合求解中, 耦合求解策略对模拟精度与稳定性具有决定性影响[62]。

不同耦合方式的数据交互过程如图 6 所示。松耦合方法将流场计算与运动求解分开推进, 在每个时间步结束后两者进行一次数据交换, 通过这种交错时间推进方式获得系统的最终解。中国空气动力研究与发展中心的达兴亚等[63] 基于结构网格, 采用改进型松耦合方法对方形截面弹和某窄条翼导弹开展了无控自由俯仰/摇滚、无动力情况下的舵面控制迎角保持与变迎角机动等简化模型的数值模拟。紧耦合方法通过隐式方式在每个物理时间步内构造多个子迭代步, 并在每个子迭代步中进行数据交换[64]。尽管计算成本较高, 但该方法在应对复杂问题时具有更强的鲁棒性。同时, 结合牛顿迭代法及高效预处理技术, 其计算效率已得到显著提升[65]。Yan 等[61] 在张来平团队[25] 提出的 NVF 框架基础上, 采用动态混合网格与重叠网格技术实现飞行器整体运动与舵面偏转, 在任意欧拉-拉格朗日(arbitrary lagrangian-eulerian, ALE)框架下求解非定常雷诺平均 Navier-Stokes方程, 实现了气动、运动与控制三者的紧耦合求解。在该耦合框架下, Yan 等[61] 完成了多自由度纵向机动仿真, 能够有效捕捉大迎角条件下的非定常气动特性与运动响应, 验证了紧耦合计算方法在强非线性机动问题中的可行性, 如图 7 所示, 图中, C_p 为表面压力系数。

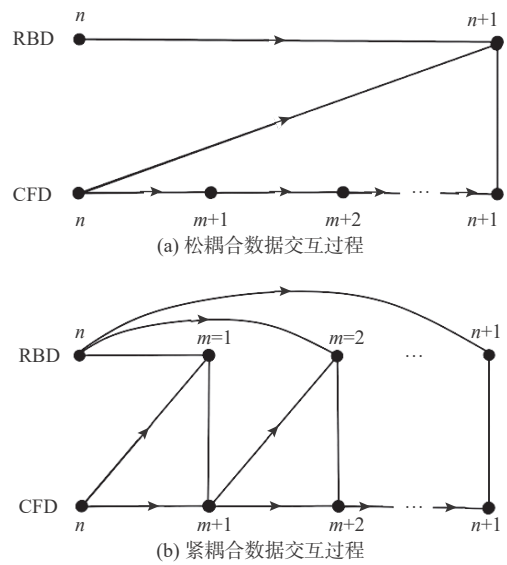


图 6　松耦合及紧耦合数据交互过程

Fig. 6　Loosely coupled and tightly coupled data exchange processes

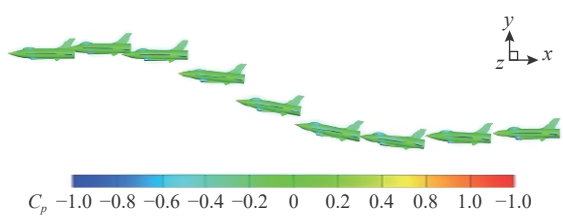


图 7　战斗机非定常机动飞行过程的位置和姿态图[61]

Fig. 7　Position and attitude diagram of fighter aircraft during unsteady maneuver flight[61]

与 Yan 等[61] 相似, 李怀璐等[66] 在大迎角机动飞行仿真中, 在训练控制律阶段, 融合多源飞行数据, 构建约束嵌入物理模型的集成神经网络框架, 以替代耗时较高的 CFD 计算, 实现飞行力学与气动、控制的一体化耦合。王建华等[67] 通过建立气动舵偏角与飞行器三维运动状态之间的直接关系, 在耦合方法上进一步推进, 提出了一种面向高超声速飞行器俯冲段的六自由度制导与控制一体化低阶建模方法, 实现了不依赖高保真流场在线求解的

强机动仿真。上述研究侧重于系统级气动-运动-控制耦合建模与快速仿真分析,提供了一条降阶模型大规模控制律训练与CFD高保真在线求解预测的工程实现路径。

近年来,多学科耦合已由单一刚体运动逐步扩展至包含柔性结构变形的流固耦合问题,Gong等[68]针对柔性翼构型,基于双向流固耦合方法与动网格技术,研究了柔性翼在非定常流动作用下的气动载荷与结构大变形响应,揭示了柔性结构变形对气动力特性及非定常流动机理的影响,为柔性飞行器的高保真飞行仿真提供了方法参考。

综上所述,六自由度耦合计算技术在动态混合网格与非定常CFD的支撑下,已从低阶气动-运动-控制一体化建模发展到高保真数值虚拟飞行闭环仿真,并进一步将研究对象拓展至包含柔性结构变形的复杂构型,从而为复杂机动飞行的多学科耦合仿真提供了重要的技术基础。

4.2 NVF仿真置信度评估

随着气动/运动/控制耦合计算方法在复杂飞行任务中的应用日益深入,仿真结果的可信度评估已成为工程实践中的关键问题。NVF仿真置信度评估的核心框架是验证与确认:验证旨在回答“是否正确地求解了方程”,确认则关注“是否求解了正确的方程”[69]。当前,NVF仿真的置信度评估主要依赖于试验数据的对比确认。风洞VFT通过允许模型在风洞中绕质心自由旋转并测量其姿态响应,为仿真结果提供了关键的验证基准。赵忠良等[70]在2.4 m跨声速风洞开展了某典型导弹模型的VFT研究,初步形成了高机动导弹气动/飞行力学/飞行控制一体化试验模拟能力。随着精细化风洞试验数据的累积,中国空气动力学研究发展中心已建立了CFD验证与确认数据库平台[71],用于存储管理风洞试验、飞行试验及数值计算产生的标准算例数据,为仿真置信度的评估提供数据支撑。

仿真与试验结果之间通常存在偏差,其主要来源可归纳为以下4类:①网格离散误差。其源于连续方程离散为代数方程时的空间近似[72]。网格尺寸过大会抹平边界层、激波等大梯度流动特征[73];长细比或偏斜度过高会降低插值精度[74];近壁面y+不达标影响边界层流体预测的准确性。②计算方法误差。其来自数值算法本身的离散与求解近似[75]。低阶格式截断误差大,容易抹平流动细节[76];时间步长过大则会降低时间分辨率,失去对一些流动结构时间演化过程的捕捉;插值格式与梯度重构方法粗糙造成界面通量计算偏差[77];边界条件近似引入系统误差;收敛准则过松导致迭代残差累积。③试验约束与测量误差。该类误差源于风洞VFT中固有的条件限制与硬件局限[78]。平动或转动自由度的约束使模型运动偏离真实飞行状态;支撑或悬吊系统产生附加力矩,干扰模型转动响应;支撑结构对流场形成物理干扰;传感器精度限制引入数据采集偏差;风洞的洞壁扰动难以完全消除,导致测量结果波动[78]。④物理模型和参数误差。该类误差源于仿真模型的假设简化及输入参数的不确定性[79]。大迎角机动飞行中存在漩涡与流动分离等复杂现象,超出了雷诺平均Navier-Stokes方程简化假设的适用范围,导致分离区域大小预测偏差;为降低耦合求解复杂度,仿真模型常忽略柔性变形、气动弹性、气动热等实际存在的物理效应[80];初始条件与边界条件的设定偏差,以及多输入参数不确定性的叠加,导致仿真结果偏离试验测量值。

针对上述误差来源,可采取以下改进措施与评估方法,以提升仿真置信度:针对网格离散误差,可手动调整边界层网格,确保近壁面y+满足所用湍流模型的要求;控制单元长细比与偏斜度在合理范围内;在计算资源允许的条件下,优先采用结构网格或高质量混合网格;对流动分离、激波等大梯度区域实施局部网格加密,并开展网格无关性验证[81],在涉及动网格的仿真中,还需设置网格质量实时监测机制,当单元畸变超过设定阈值时,触发局部网格自动重构[82]。针对计算方法误差,空间离散应优先采用二阶及以上精度的格式,对流项推荐选用低耗散格式以更好地分辨旋涡等流动结构[83];湍流模型数值离散应采用与主流场一致的高阶格式,以减少离散误差对湍流预测的附加干扰[84];时间步长的选取须满足CFL条件,并对瞬态问题开展时间步长敏感性分析[85],同时,应严格设定收敛准则,避免残差累积引入系统偏差。针对测量与约束误差,可采用虚拟约束替代刚性支撑,以减轻平动自由度的限制[86];对轴承摩擦等机械阻尼进行事先标定,并在数据处理环节引入补偿修正;优化支撑结构的气动外形,降低其对流场的物理干扰[87];选用高精度传感器并进行多次重复测量以评估随机误差;在风洞试验过程中增加环境监测与信号滤波处理,以抑制洞壁扰动与测量噪声的影响[88]。针对物理模型与参数误差,通过实测手段获取更准确的物性参数;根据实际流动特征选择合适的物理模型,避免为降低计算成本而过度简化关键物理过程[89];采用不确定性量化方法,系统评估输入参数波动对仿真输出的影响,对关键参数开展全局灵敏度分析,识别主

导误差源并优先校准[90]。上述评估步骤与改进措施共同构成了从误差识别到修正验证的闭环流程，是提升 NVF 置信度的可靠路径。

5　气动/运动/控制耦合问题应用场景拓展

上文系统梳理了动网格技术、基于动网格的非定常计算方法及气动/运动/控制耦合计算方法的研究进展。本节将结合实际应用，进一步分析上述数值方法在复杂飞行器系统中的应用场景与技术挑战，重点阐述其在多物理场耦合、多体分离与变构飞行，以及仿生飞行/航行等前沿领域的应用潜力。

5.1　热-流-固多物理场耦合下的气动/运动/控制问题

在跨声速和高超声速飞行仿真领域，气动弹性、气动热效应与飞行控制的耦合已经成为研究热点之一。随着飞行速度达到马赫数 Ma=3，气动加热可使结构表面温度升至 350 ℃，铝合金结构强度已难以满足要求。当飞行速度超过马赫数 Ma=5，激波-边界层相互作用主导的极端气动加热对结构稳定性与控制鲁棒性提出了更高要求。为精确预测上述多场耦合效应，研究者构建了可压缩雷诺平均 Navier-Stokes 方程、固体传热方程与动力学方程的耦合求解框架，并采用混合插值算法实现了界面数据的高精度传递，为高超声速飞行器的热防护设计、结构振动分析与控制系统响应评估提供了有效的数值手段。

贾洲侠等[91]针对高超声速飞行器的热防护设计需求，建立了可压缩雷诺平均 Navier-Stokes 方程与结构传热方程的耦合框架，采用分区迭代方法实现流场与结构热响应之间的界面数据传递，显著提升了结构热响应的预测精度，为飞行器的热防护设计提供了有力支撑。Wu 等[92]研究了高超声速充气再入柔性返回舱在极端气动力热载荷下的热气动弹性特性，发现气动加热会显著增大结构热应力、改变其固有振动特性。同时，高超声速激波-边界层强干扰在迎风面诱发的非定常大尺度分离涡发生周期性脱落，诱发柔性结构的剧烈受迫振动。

Bai 等[93]提出高超声速全动舵面热-流-固耦合分析方法。通过有限体积法完成气动热仿真，以 RBF 插值传递温度场，更新刚度矩阵开展热模态分析，再经线性颤振计算，量化气动加热对舵面模态与颤振特性的影响，为高超声速舵面热气动弹性设计提供数值支撑。在高速飞行器设计方面，Wang 等[94]针对大后掠三角翼构型高速飞行器如图 8 所示，采用热-流-固耦合仿真，研究了热效应对机翼颤振特性的影响。Guo 等[95]开发了针对高速飞行器全动舵面的热颤振预测方法，揭示了气动加热对结构模态与颤振特性的影响机制，并发现同步算法与传统分区算法在监测点的温差最高可达 8%。

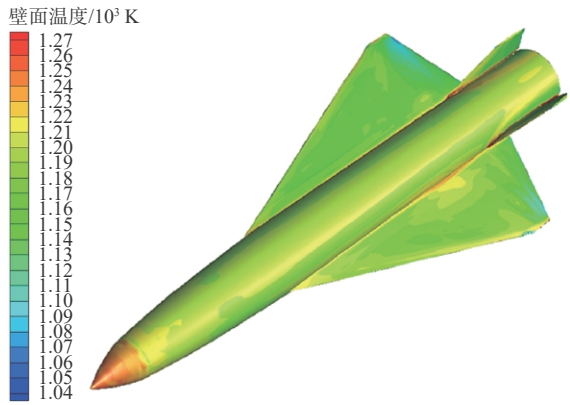


图 8　大后掠角三角翼壁面温度分布[94]

Fig. 8　Temperature distribution on the wing surface of a high-sweep-angle delta wing[94]

5.2　多体分离与变构飞行器的气动/运动/控制问题

多体分离与变构飞行器均表现出显著的构型时变性特征。该类问题涉及飞行器气动性能、结构响应、RBD 和控制系统之间的复杂交互作用，因而同样需要在统一的气动/运动/控制耦合框架下进行建模与求解。

在多体分离问题的求解中，重叠网格技术通过在背景网格与部件网格之间建立守恒插值关系，实现多体相对运动过程中流场信息的有效传递，为导弹分离、副油箱抛撒等典型应用场景的耦合仿真提供了支撑。王晓鹏[96]基于重叠网格技术开展了气动与动力学耦合仿真，预测了空空导弹与载机分离过程中的轨迹和姿态变化规律，揭示了两者之间的气动耦合特性。李孝伟等[97]耦合 Euler 方程与六自由度动力学方程，仿真了外挂物从飞行器分离后的绕流特性和相对运动，并通过动态嵌套网格技术显著提高了计算效率。

在变构型飞行器的姿态控制研究方面，安效民等[98]以变后掠机翼为对象，采用 ALE 格式 Euler 方程求解非定常流场，发展了滑移网格方法处理变后掠过程中的网格运动，建立了 CFD/CSD 耦合气动弹性模型。分析表明，机翼扭转变形导致局部迎角最大减小约 0.1°，变后掠运动的非定常效应使压力系数与固定构型静气动弹性的差距超过 10%。刘昊东等[99]针对变形过程中的姿态稳定问题，设计了增量非线性动态逆控制器（incremental nonlinear dynamic inversion, INDI），并结合扩张状态观测器（extended state observer, ESO）进行干扰补偿，通过增量形式降低对精确模型的依赖，有效抑制了变形

过程中的非线性与外部干扰。该方法为变构型飞行器的控制系统设计提供了有效方案，如图9所示，该增量非线性动态逆控制架构将包含期望迎角α_c、侧滑角β_c、航迹角γ_c的指令向量$\boldsymbol{y}_c$解算为角速率指令$\boldsymbol{\omega}_c$，经微分跟踪器输出z_{12}并与反馈值比较得到误差e_2，通过非线性控制律$L(e_2)$计算控制增量$\Delta\boldsymbol{u}$，再与上一时刻控制量$\boldsymbol{u}_0$叠加生成总控制量$\boldsymbol{u}=\boldsymbol{u}_0+\Delta\boldsymbol{u}$，驱动飞行器六自由度模型。上述研究表明，变构型飞行器的气动/结构/运动/控制耦合涉及多个层面的复杂交互：气动/结构/运动耦合决定变形过程中的载荷分布与运动响应，控制系统则需在此基础上实现姿态稳定。

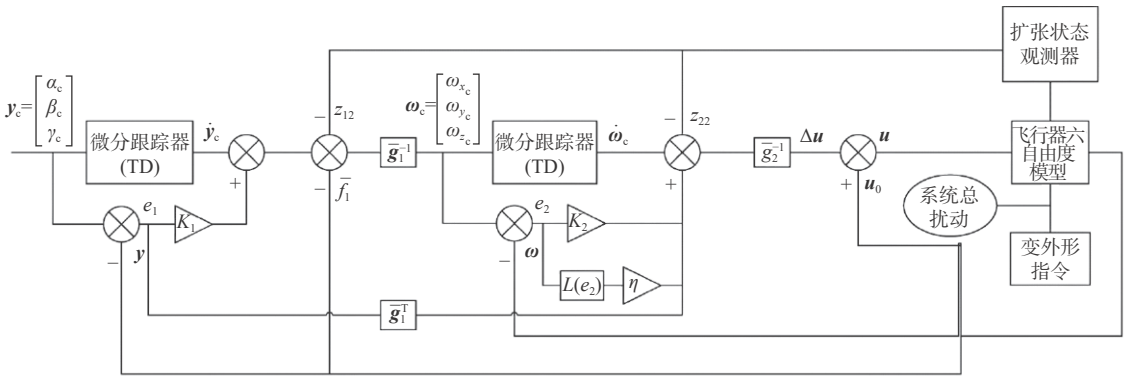


图9 INDI+ESO控制框架示意图[99]

Fig. 9 INDI+ESO control framework schematic[99]

近年来，智能控制方法在飞行器控制系统的实时补偿与参数自优化方面取得了显著进展，如通过人工神经网络优化比例-积分-微分（proportional-integral-derivative, PID）控制参数[100]、利用强化学习实现动力学系统参数的自适应调整等[101]，智能控制方法能够在极端飞行状态下快速调整飞行器的控制策略，并实现更加精确的飞行控制。此外，数据驱动的自整定策略[102]和扰动补偿方法[103]在复杂飞行环境中展现出巨大的应用潜力，尤其能够大幅度提升极端工况下控制系统的响应能力。然而，智能控制方法与高保真耦合计算框架的深度融合仍面临诸多挑战，两者的协同发展将是未来研究的重要方向。

5.3 仿生飞行器的气动/运动/控制耦合仿真问题

仿生飞行器通过模仿自然界中鸟类/昆虫的翼/翅的扑动，实现高效飞行。该类系统的核心特征是“形变-运动-流场-控制”高度一体化，其推进力与操纵力矩来源于机体整体和仿生翼/翅的连续大形变、周期性运动，而非传统的舵面控制。这一根本差异使仿生飞行器的数值仿真面临极大的技术挑战。

在动网格技术方面，基于RBF的网格变形策略被广泛用于维持扑翼大位移运动过程中的网格质量，在三维扑翼气动与柔性变形仿真中表现出良好的稳定性与计算效率[104]。仿真结果精准捕捉了翼面上的前缘涡结构，揭示了柔性机翼能够有效稳定前缘涡、延迟失速发生，并通过优化力矢量方向，显著提升扑翼推进效率。这一发现不仅为仿生飞行器气动性能优化提供了理论支持，还为复杂动力学系统的设计与控制提供了实践指导。Addo-Akoto等[105]通过对昆虫膜翅的研究，验证了柔性膜翅对前缘涡结构的稳定作用，并指出通过优化柔性膜翅分布能够延迟失速，提升整体气动效率。Li等[106]提出了三维变分气动弹性框架，专门用于处理柔性多体系统在不可压缩湍流中的流固耦合响应问题。研究表明，CFD-CSD耦合框架能够精准捕捉柔性膜翅扑动过程中的流场与结构相互作用，为仿生飞行器的动力学分析提供了有效工具。

智能控制方法在仿生飞行器的应用中也取得了显著进展。Xue等[107]提出一种基于深度强化学习的膜翅动力学控制器，用于实现大黄蜂在六自由度大扰动下的稳定悬停。该研究通过耦合CFD数据驱动气动模型与六自由度动力学模型，构建了高保真仿真环境，采用Soft Actor-Critic算法优化动作控制策略，在强扰动条件下实现了快速稳定控制，为仿生扑翼飞行器的控制器设计提供了新思路，同时，也为未来更复杂飞行任务的控制优化提供了可行的路径。

气动/运动/控制的耦合仿真已有效支撑了多个仿生飞行器的工程落地。Chen等[108]发明的“小隼2.0”扑翼飞行器成功实现了可调节的仿鸟扑-扫掠-折叠（flapping-sweeping- folding, FSF）运动模式，并通过调节扫掠幅度有效控制了升力和俯仰力矩，最终实现了基于该原理的仿鸟自主起飞，如图10所示。针对蜻蜓等仿昆虫微型飞行器，刘笑尘等[109]采用双向流固耦合与重叠网格技术进行数值仿真，

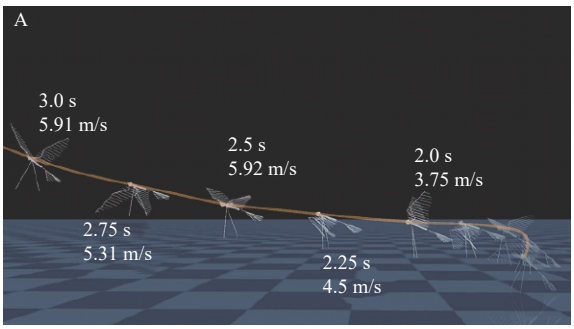


图 10　复杂扑动规律扑翼飞行器运动仿真[108]

Fig. 10　Dynamic simulation of flapping-wing aircrafts with complex flapping regimes[108]

研究表明, 沿弦向优化的非均匀柔性分布能显著提升蜻蜓前翼的气动性能。

综上所述, 双向流固耦合、动网格、非定常气动模拟与智能控制等技术的融合应用, 为新型仿生飞行器的设计与控制提供了关键支撑。随着这些技术的持续发展与深度融合, 将推动新一代高机动、低能耗、低噪声智能仿生飞行器的发展。

6　结束语

近十年来, 随着高性能计算、动网格技术、CFD 非定常算法、多学科耦合仿真方法的持续发展, NVF 技术已从理论探索进入工程应用阶段。基于结构网格、非结构/混合网格实现的动网格技术有效支撑了舵面偏转、复杂变构、多体相对运动、仿生扑动飞行等复杂系统的数值仿真需求。基于动网格的非定常计算方法在空间离散、时间推进与耦合求解 3 个层面均已形成较为完备的技术闭环。气动/运动/控制耦合计算则实现了从离线分步松耦合向实时紧耦合闭环的关键跨越, 能够在统一时域内刻画强非线性与多学科耦合效应。机器学习方法的引入, 进一步提升了仿真过程中网格自适应优化能力、气动模型构建效率及控制策略鲁棒性。上述技术的融合发展有力推动了 NVF 仿真结果与试验数据之间一致性的不断提升, 标志着飞行器的研制模式加速向“仿真驱动、数智试验”转型。面向未来, NVF 将进一步向高效统一、全系统多场耦合的方向发展, 具体方向包括:

1) 柔性仿生的应用牵引。未来柔性变构型飞行器、仿生飞行器及水下航行器等复杂运动控制需求, 将促使 NVF 在非定常气动代理模型高效构建和高保真度 CFD 气动计算之间取得平衡, 高效高保真地完成复杂运动模式仿真, 支撑控制律设计。

2) 机器学习的深度融合。机器学习方法将进一步赋能 NVF 的各个环节, 逐步发展为网格自适应变形与加密、仿真参数不确定性分析与优化, 以及非线性系统控制律设计等任务的核心工具, 从而系统性地提升 NVF 的仿真效率和结果置信度。

3) 多场耦合的加速演进。随着对飞行物理的认识深入, 涵盖气动、结构、热、动力学、控制等更多物理过程的耦合建模与一体化仿真, 将更加真实地反映飞行过程, 进一步支撑飞行机理的深入探究和飞行控制律的改进完善。

4) 数智试验的平台建设。融合 NVF 与机器学习技术, 打造数智试验场, 将有效支撑新一代飞行器开展数字试飞, 推动飞行器试验的全面数字化转型, 显著提升飞行器研发效率。

参考文献（References）

[1] Fu J Q, Shi Z W, Gong Z, et al. Virtual flight test techniques to predict a blended-wing-body aircraft in-flight departure characteristics[J]. Chinese Journal of Aeronautics, 2022, 35(1): 215-225.

[2] Salas M D. Digital flight: the last CFD aeronautical grand challenge[J]. Journal of Scientific Computing, 2006, 28(2): 479-505.

[3] E Q, Yang G W, Li F W. Numerical solution for complete aircraft's subsonic-transonic-supersonic flow at high angles of attack[J]. Computer Methods in Applied Mechanics and Engineering, 1997, 148(1-2): 11-21.

[4] Despirito J, Silton S I, Weinacht P. Navier-Stokes predictions of dynamic stability derivatives: evaluation of steady-state methods [J]. Journal of Spacecraft and Rockets, 2009, 46(6): 1142-1154.

[5] Mi B G, Zhan H, Lu S S. An extended unsteady aerodynamic model at high angles of attack[J]. Aerospace Science and Technology, 2018, 77: 788-801.

[6] Persova M G, Soloveichik Y G, Belov V K, et al. Modeling of aerodynamic heat flux and thermoelastic behavior of nose caps of hypersonic vehicles[J]. Acta Astronautica, 2017, 136: 312-331.

[7] Huang M, Wang Z W. A review of wind tunnel based virtual flight testing techniques for evaluation of flight control systems[J]. International Journal of Aerospace Engineering, 2015, 2015(1): 672423.

[8] 陈龙, 伍贻兆, 夏健. CFD/CSD 紧耦合及新型动网格方法在气动弹性模拟中的应用[J]. 计算力学学报, 2011, 28(5): 807-812.
Chen L, Wu Y Z, Xia J. CFD/CSD closely coupled and new dynamic grid method in application of aeroelastic simulation[J]. Chinese Journal of Computational Mechanics, 2011, 28(5): 807-812(in Chinese).

[9] Sutrisno, Rohmat T A, Wibowo S B, et al. Vortex dynamics study of the canard deflection angles' influence on the sukhoi su-30-like model to improve stall delays at high AoA[J]. Aerospace, 2019, 6(2): 12.

[10] Thibert J J, Arnal D. A review of ONERA aerodynamic research in support of a future supersonic transport aircraft[J]. Progress in Aerospace Sciences, 2000, 36(8): 581-627.

[11] Lee D K, Han J H. Safety-guaranteed flight test environment for micro air vehicles[J]. AIAA Journal, 2015, 54(3): 1018-1029.

[12] Schütte A, Einarsson G, Raichle A, et al. Numerical simulation of maneuvering aircraft by aerodynamic, flight mechanics and structural mechanics coupling[J]. Journal of Aircraft, 2009, 46(1): 53-64.

[13] 杨云军, 周伟江, 崔尔杰. 耦合时间精度对模拟飞行器自由运动特性的影响[J]. 计算物理, 2007, 24(1): 42-48.
Yang Y J, Zhou W J, Cui E J. Influence of coupling time accuracy on the simulation of aircraft free movement[J]. Chinese Journal of Computational Physics, 2007, 24(1): 42-48(in Chinese).

[14] Guo Z Z, Dai G, Yang H, et al. Unsteady flow simulation of a variable-sweep morphing aircraft coupled with flight control system[J]. International Journal of Modern Physics B, 2020, 34(14-16): 2040073.

[15] Liang Y M, Li G N, Xu M, et al. An intelligent control method based on artificial neural network for numerical flight simulation of the basic finner projectile with pitching maneuver[J]. Defence Technology, 2024, 32: 663-674.

[16] Wang N H, Ma R, Chang X H, et al. Numerical virtual flight simulation of quasi-cobra maneuver of a fighter aircraft[J]. Journal of Aircraft, 2020, 58(1): 138-152.

[17] 张来平, 段旭鹏, 常兴华, 等. 基于 Delaunay 背景网格插值和局部网格重构的变形体动态混合网格生成技术[J]. 空气动力学学报, 2009, 27(1): 32-40.
Zhang L P, Duan X P, Chang X H, et al. A hybrid dynamic grid generation technique for morphing bodies based on Delaunay graph and local remeshing[J]. Acta Aerodynamica Sinica, 2009, 27(1): 32-40(in Chinese).

[18] 张来平, 邓小刚, 张涵信. 动网格生成技术及非定常计算方法进展综述[J]. 力学进展, 2010, 40(4): 424-447.
Zhang L P, Deng X G, Zhang H X. Reviews of moving grid generation techniques and numerical methods for unsteady flow[J]. Advances in Mechanics, 2010, 40(4): 424-447(in Chinese).

[19] Liu X Q, Qin N, Xia H. Fast dynamic grid deformation based on Delaunay graph mapping[J]. Journal of Computational Physics, 2006, 211(2): 405-423.

[20] Chaderjian N M, Schiff L B. Numerical simulation of forced and free-to-roll delta-wing motions[J]. Journal of Aircraft, 1996, 33(1): 93-99.

[21] Woods B K S, Dayyani I, Friswell M I. Fluid/structure-interaction analysis of the fish-bone-active-camber morphing concept[J]. Journal of Aircraft, 2015, 52(1): 307-319.

[22] Ouyang Y, Gu Y S, Kou X P, et al. Active flutter suppression of wing with morphing flap[J]. Aerospace Science and Technology, 2021, 110: 106457.

[23] Fonzi N, Brunton S L, Fasel U. Data-driven nonlinear aeroelastic models of morphing wings for control[J]. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2020, 476(2239): 20200079.

[24] Yuan W X, Sandhu R, Poirel D. Fully coupled aeroelastic analyses of wing flutter towards application to complex aircraft configurations[J]. Journal of Aerospace Engineering, 2021, 34(2): 04020117.

[25] 张来平, 马戎, 常兴华, 等. 虚拟飞行中气动、运动和控制耦合的数值模拟技术[J]. 力学进展, 2014, 44(1): 376-417.
Zhang L P, Ma R, Chang X H, et al. Review of aerodynamics/kinematics/flight-control coupling methods in virtual flight simulations[J]. Advances in Mechanics, 2014, 44(1): 376-417(in Chinese).

[26] 谭伟伟, 刘峰博, 张一帆. 基于线性弹性体的网格变形方法研究[J]. 航空计算技术, 2017, 47(2): 37-40.
Tan W W, Liu F B, Zhang Y F. Study on linear elastic solid based grid deformation method[J]. Aeronautical Computing Technique, 2017, 47(2): 37-40(in Chinese).

[27] 张兵, 韩景龙. 带旋转修正的弹簧-TFI 混合动网格方法[J]. 航空学报, 2011, 32(10): 1815-1823.
Zhang B, Han J L. Spring-TFI hybrid dynamic mesh method with rotation correction[J]. Acta Aeronautica et Astronautica Sinica, 2011, 32(10): 1815-1823(in Chinese).

[28] 张伟伟, 高传强, 叶正寅. 气动弹性计算中网格变形方法研究进展[J]. 航空学报, 2014, 35(2): 303-319.
Zhang W W, Gao C Q, Ye Z Y. Research progress on mesh deformation method in computational aeroelasticity[J]. Acta Aeronautica et Astronautica Sinica, 2014, 35(2): 303-319(in Chinese).

[29] 孙岩, 邓小刚, 王光学, 等. 基于径向基函数改进的 Delaunay 图映射动网格方法[J]. 航空学报, 2014, 35(3): 727-735.
Sun Y, Deng X G, Wang G X, et al. Improvement on delaunay graph mapping dynamic grid method based on radial basis functions[J]. Acta Aeronautica et Astronautica Sinica, 2014, 35(3): 727-735(in Chinese).

[30] 孙岩, 邓小刚, 王运涛, 等. RBFTFI 结构动网格技术在风洞静气动弹性修正中的应用[J]. 工程力学, 2014, 31(10): 228-233.
Sun Y, Deng X G, Wang Y T, et al. Application of structural dynamic grid method based on RBFTFI on wind tunnel static aeroelastic modification[J]. Engineering Mechanics, 2014, 31(10): 228-233(in Chinese).

[31] Wang Z J, Parthasarathy V. A fully automated Chimera methodology for multiple moving body problems[J]. International Journal for Numerical Methods in Fluids, 2000, 33(7): 919-938.

[32] Nakahashi K, Togashi F, Sharov D. Intergrid-boundary definition method for overset unstructured grid approach[J]. AIAA Journal, 2000, 38(11): 2077-2084.

[33] 段旭鹏, 常兴华, 张来平. 基于动态混合网格的多体分离数值模拟方法[J]. 空气动力学学报, 2011, 29(4): 447-452.
Duan X P, Chang X H, Zhang L P. A CFD-and-6DOF-coupled solver for multiple moving object problems based on dynamic hybrid grids[J]. Acta Aerodynamica Sinica, 2011, 29(4): 447-452(in Chinese).

[34] Gillebaart T, Blom D S, Van Zuijlen A H, et al. Adaptive radial basis function mesh deformation using data reduction[J]. Journal of Computational Physics, 2016, 321: 997-1025.

[35] 常兴华, 王年华, 马戎, 等. 并行重叠/变形混合网格生成技术及其应用[J]. 气体物理, 2019, 4(6): 12-21.
Chang X H, Wang N H, Ma R, et al. Dynamic hybrid mesh generator coupled with overset and deformation in parallel environment[J]. Physics of Gases, 2019, 4(6): 12-21(in Chinese).

[36] 陈坚强. 国家数值风洞 (NNW) 工程关键技术研究进展[J]. 中国科学 (技术科学), 2021, 51(11): 1326-1347.
Chen J Q. Advances in the key technologies of Chinese national numerical windtunnel project[J]. Scientia Sinica (Technologica), 2021, 51(11): 1326-1347(in Chinese).

[37] 李海峰, 吴冀川, 刘建波, 等. 有限元网格剖分与网格质量判定指标[J]. 中国机械工程, 2012, 23(3): 368-377.
Li H F, Wu J C, Liu J B, et al. Finite element mesh generation and decision criteria of mesh quality[J]. China Mechanical Engineering, 2012, 23(3): 368-377(in Chinese).

[38] Sorgente T, Biasotti S, Manzini G, et al. A survey of indicators for

mesh quality assessment[J]. Computer Graphics Forum, 2023, 42(2): 461-483.

[39] Secco N R, Kenway G K W, He P, et al. Efficient mesh generation and deformation for aerodynamic shape optimization[J]. AIAA Journal, 2021, 59(4): 1151-1168.

[40] 李绍磊, 郑建靖, 尚萌萌. 基于非结构网格的飞行器多体分离数值模拟[J]. 计算机应用, 2016, 36(6): 1741-1744.
Li S L, Zheng J J, Shang M M. Numerical simulation of flight vehicle multi-body separation based on unstructured mesh[J]. Journal of Computer Applications, 2016, 36(6): 1741-1744(in Chinese).

[41] 杨宇辰. 可压缩流动模拟的自适应网格方法研究[D]. 南京: 南京航空航天大学, 2023.
Yang Y C. Research on adaptive mesh method for compressible flow simulation[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2023(in Chinese).

[42] Haubner J, Hellan O, Zeinhofer M, et al. Learning mesh motion techniques with application to fluid-structure interaction[J]. Computer Methods in Applied Mechanics and Engineering, 2024, 424: 116890.

[43] 刘翰林, 崔会敏, 张珍, 等. 基于 BP 算法的二维阵面推进网格划分方法的优化与应用[J]. 计算力学学报, 2024, 41(4): 626-633.
Liu H L, Cui H M, Zhang Z, et al. Optimization and application of two-dimensional advancing grid generation method based on BP algorithm[J]. Chinese Journal of Computational Mechanics, 2024, 41(4): 626-633(in Chinese).

[44] Wang Z C, Chen X H, Li T J, et al. Evaluating mesh quality with graph neural networks[J]. Engineering with Computers, 2022, 38(5): 4663-4673.

[45] Foucart C, Charous A, Lermusiaux P F J. Deep reinforcement learning for adaptive mesh refinement[J]. Journal of Computational Physics, 2023, 491: 112381.

[46] 万嘉伟. 龙格-库塔法数值求解基于有限体积的不可压 Navier-Stokes 方程和流固耦合问题[D]. 成都: 西南交通大学, 2020.
Wan J W. The numerical solution of finite-volume based incompressible navier-stokes equations and fluid-structure interaction problems by Runge-Kutta methods[D]. Chengdu: Southwest Jiaotong University, 2020(in Chinese).

[47] 张乐, 章定国. 基于向后差分法求解多体系统动力学微分-代数方程组的双循环隐式积分方法[J]. 机械工程学报, 2016, 52(7): 79-87.
Zhang L, Zhang D G. Two-loop implicit integration method based on backward differential formulation for differential-algebraic equations of multibody system dynamics[J]. Journal of Mechanical Engineering, 2016, 52(7): 79-87(in Chinese).

[48] 贡伊明, 张伟伟, 刘溢浪. 非定常求解的内迭代初值对计算效率的影响研究[J]. 西北工业大学学报, 2016, 34(1): 11-17.
Gong Y M, Zhang W W, Liu Y L. Researching how initial value of internal iteration impacts on computational efficiency in unsteady flow solving[J]. Journal of Northwestern Polytechnical University, 2016, 34(1): 11-17(in Chinese).

[49] 马晓永, 艾俊强, 宋笔锋. 有限差分法在机翼气动外形优化仿真中的应用[C]//第 14 届中国系统仿真技术及其应用学术年会论文集. 武汉: 科研出版社, 2012: 214-218.
Ma X Y, Ai J Q, Song B F. Application of finite difference method in wing aerodynamic shape optimization simulation[C]//Proceedings of the 14th China Annual Conference on System Simulation Technology and Its Application. Wuhan: Scientific Research Publishing Inc. , 2012: 214-218(in Chinese).

[50] Abalakin I, Dervieux A, Kozubskaya T. High accuracy finite volume method for solving nonlinear aeroacoustics problems on unstructured meshes[J]. Chinese Journal of Aeronautics, 2006, 19(2): 97-104.

[51] 席柯, 阎超, 黄宇. 带反馈控制的基本带翼外形的耦合气动/运动数值模拟研究[J]. 空气动力学学报, 2016, 34(2): 182-189.
Xi K, Yan C, Huang Y. Coupled CFD/RBD modelingfor a basic finner projectile with control[J]. Acta Aerodynamica Sinica, 2016, 34(2): 182-189(in Chinese).

[52] 李少伟, 张元静, 周丹杰. 支撑约束的并联飞行器多体分离安全性数值研究[C]//第九届全国流体力学学术会议论文集. 北京: 中国力学学会, 2016: 376.
Li S W, Zhang Y J, Zhou D J. Numerical study on the safety of multi-body separation of parallel aircraft with support constraints[C]//Proceedings of the 9th National Conference on Fluid Mechanics. Beijing: The Chinese Society of Theoretical and Applied Mechanics, 2016: 376(in Chinese).

[53] Deng Y, Jiao Z X, Xu Y Z. Frequency-domain analysis of fluid-structure interaction in aircraft hydraulic pipeline systems: numerical and experimental studies[J]. Journal of Zhejiang University-Science A (Applied Physics & Engineering), 2024, 25(8): 605-618.

[54] 白玉柔, 王永锋, 马艳红, 等. 大涵道比风扇转子双套齿连接稳健性及其力学影响[J]. 航空动力学报, 2025, 40(7): 20240124.
Bai Y R, Wang Y F, Ma Y H, et al. Robustness and mechanical effects of double set tooth connection for large bypass ratio fan rotor[J]. Journal of Aerospace Power, 2025, 40(7): 20240124(in Chinese).

[55] 黄国. 高超声速环境下缝隙热环境的数值模拟研究[D]. 北京: 北京交通大学, 2017.
Huang G. Numerical simulation for heating environment of gap in hypersonic flow[D]. Beijing: Beijing Jiaotong University, 2017(in Chinese).

[56] 于要杰, 刘锋, 高超, 等. 一种基于块雅可比迭代的高阶 FR 格式隐式方法[J]. 力学学报, 2021(6): 1586-1598.
Yu Y J, Liu F, Gao C, et al. An implicit block jacobi approach for high-order flux reconstruction method[J]. Chinese Journal of Theoretical and Applied Mechanics, 2021(6): 1586-1598(in Chinese).

[57] 孙岩. 共轴刚性旋翼直升机流场和噪声高精度模拟及质量射流控制研究[D]. 南京: 南京航空航天大学, 2022.
Sun Y. High-accuracy simulation of flowfield, noise and TAMI control for rigid coaxial rotor helicopter[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2022(in Chinese).

[58] 郭曼丽, 王园丁, 谭俊杰, 等. N-S 方程耦合雷诺平均湍流模型无网格算法研究[J]. 南京理工大学学报 (自然科学版), 2019, 43(5): 548-555.
Guo M L, Wang Y D, Tan J J, et al. Meshless method coupled with different Reynolds average turbulence model for solving Navier-Stokes equation[J]. Journal of Nanjing University of Science and Technology, 2019, 43(5): 548-555(in Chinese).

[59] 李彩云, 安慰, 刘学军, 等. 基于机器学习的非定常流场网格自适应[J]. 空气动力学学报, 2023, 41(6): 89-101.
Li C Y, An W, Liu X J, et al. Mesh adaptation for unsteady flow field based on machine learning[J]. Acta Aerodynamica Sinica,

2023, 41(6): 89-101(in Chinese).

[60] 赵明, 丁秋实, 曹京硕, 等. 基于间断伽辽金方法的波形前缘动态失速流场仿真分析[C]//第十二届全国流体力学学术会议论文集. 北京: 中国力学学会, 2022: 865.
Zhao M, Ding Q S, Cao J S, et al. Simulation analysis of dynamic stall flow field on wavy leading edge based on discontinuous Galerkin method[C]//Proceedings of the 12th National Conference on Fluid Mechanics. Beijing: The Chinese Society of Theoretical and Applied Mechanics, 2022: 865(in Chinese).

[61] Yan L, Chang X H, Wang N H, et al. Numerical virtual flight investigation for longitudinal maneuver of a generic fighter based on machine learning[J]. Physics of Fluids, 2024, 36(7): 077131.

[62] 刘艳, 白俊强, 华俊, 等. 基于 RBF 插值技术的 CFD/CSD 非线性耦合分析方法研究[J]. 计算力学学报, 2014, 31(1): 120-127.
Liu Y, Bai J Q, Hua J, et al. A approach to CFD/CSD non-linear coupling based on RBF interpolation technology[J]. Chinese Journal of Computational Mechanics, 2014, 31(1): 120-127(in Chinese).

[63] 达兴亚, 陶洋, 赵忠良. 基于预估校正和嵌套网格的虚拟飞行数值模拟[J]. 航空学报, 2012, 33(6): 977-983.
Da X Y, Tao Y, Zhao Z L. Numerical simulation of virtual flight based on predictioncorrection coupling method and chimera grid[J]. Acta Aeronautica et Astronautica Sinica, 2012, 33(6): 977-983(in Chinese).

[64] 马戎, 常兴华, 赫新, 等. 流动/运动松耦合与紧耦合计算方法及稳定性分析[J]. 气体物理, 2016, 1(6): 36-49.
Ma R, Chang X H, He X, et al. Loose and strong coupling methods for flow/kinematics coupled simulations and stability analysis[J]. Physics of Gases, 2016, 1(6): 36-49(in Chinese).

[65] 王博, 刘勇. 新型无铰式旋翼气弹综合分析的全本征方程方法[J]. 南京航空航天大学学报, 2016, 48(2): 261-267.
Wang B, Liu Y. Comprehensive aeroelasticity analysis on new hingeless rotor using fully intrinsic equations[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2016, 48(2): 261-267(in Chinese).

[66] 李怀璐, 王旭, 王霄, 等. 大迎角机动飞行的气动力建模与飞行仿真[J]. 航空学报, 2023, 44(19): 78-95.
Li H L, Wang X, Wang X, et al. Aerodynamic modeling and flight simulation of maneuver flight at high angle of attack[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(19): 78-95(in Chinese).

[67] 王建华, 刘鲁华, 王鹏, 等. 高超声速飞行器俯冲段制导控制一体化设计方法[J]. 航空学报, 2017, 38(3): 320328.
Wang J H, Liu L H, Wang P, et al. Integrated guidance and control scheme for hypersonic vehicles in dive phase[J]. Acta Aeronautica et Astronautica Sinica, 2017, 38(3): 320328(in Chinese).

[68] Gong Q W, Li Z D, He R Y, et al. Numerical study of flexible wing aerodynamic characteristics of bionic flapping-wing vehicle in Martian atmosphere[J]. Journal of Physics: Conference Series, 2025, 3078(1): 012001.

[69] Guide. Guide for the verification and validation of computational fluid dynamics simulations[M]. Reston: AIAA, 1998.

[70] 赵忠良, 吴军强, 李浩, 等. 2.4 m 跨声速风洞虚拟飞行试验技术研究[J]. 航空学报, 2016, 37(2): 504-512.
Zhao Z L, Wu J Q, Li H, et al. Investigation of virtual flight testing technique based on 2.4 m transonic wind tunnel[J]. Acta Aeronautica et Astronautica Sinica, 2016, 37(2): 504-512(in Chinese).

[71] 赵炜, 陈江涛, 肖维, 等. 国家数值风洞 (NNW) 验证与确认系统关键技术研究进展[J]. 空气动力学学报, 2020, 38(6): 1165-1172.
Zhao W, Chen J T, Xiao W, et al. Advances in the key technologies of verification and validation system of national numerical windtunnel project[J]. Acta Aerodynamica Sinica, 2020, 38(6): 1165-1172(in Chinese).

[72] Babuska I, Oden J T. Verification and validation in computational engineering and science: basic concepts[J]. Computer Methods in Applied Mechanics and Engineering, 2004, 193(36-38): 4057-4066.

[73] Celik I B, Ghia U, Roache P J, et al. Procedure for estimation and reporting of uncertainty due to discretization in CFD applications[J]. Journal of Fluids Engineering, 2008, 130(7): 078001.

[74] Calhoun D, Leveque R J. An accuracy study of mesh refinement on mapped grids[C]//Adaptive Mesh Refinement-Theory and Applications. Berlin: Springer, 2006: 91-101.

[75] Roy C J. Review of code and solution verification procedures for computational simulation[J]. Journal of Computational Physics, 2005, 205(1): 131-156.

[76] 周毓麟, 李德元. 非定常流体力学数值方法的若干问题[J]. 数学进展, 1981, 10(1): 48-56.
Zhou Y L, Li D Y. Some problems of numerical methods of unsteady fluid mechanics[J]. Advances in Mathematics, 1981, 10(1): 48-56(in Chinese).

[77] 高智. 数值摄动算法及其 CFD 格式[J]. 力学进展, 2010, 40(6): 607-633.
Gao Z. Numerical perturbation algorithm and its CFD schemes[J]. Advances in Mechanics, 2010, 40(6): 607-633(in Chinese).

[78] 郭林亮, 祝明红, 孔鹏, 等. 风洞虚拟飞行模型机与原型机动力学特性分析[J]. 航空学报, 2016, 37(8): 2583-2593.
Guo L L, Zhu M H, Kong P, et al. Analysis of dynamic characteristics between prototype aircraft and scaled-model of virtual flight test in wind tunnel[J]. Acta Aeronautica et Astronautica Sinica, 2016, 37(8): 2583-2593(in Chinese).

[79] Da Ronch A, Panzeri M, Drofelnik J, et al. Sensitivity and calibration of turbulence model in the presence of epistemic uncertainties[J]. CEAS Aeronautical Journal, 2020, 11(1): 33-47.

[80] 宁小磊. 数字化试验仿真模型验证方法验证分辨率研究[J]. 西北工业大学学报, 2025, 43(4): 831-842.
Ning X L. Study on validation resolution of digitized test simulation model validation method[J]. Journal of Northwestern Polytechnical University, 2025, 43(4): 831-842(in Chinese).

[81] Eça L, Vaz G, Hoekstra M, et al. Assessment of discretization uncertainty estimators based on grid refinement studies[J]. Journal of Verification, Validation and Uncertainty Quantification, 2021, 6(4): 041006.

[82] 张斌, 杨涛, 丰志伟, 等. 网格质量反馈的弹性体动网格改进[J]. 国防科技大学学报, 2018, 40(1): 10-16.
Zhang B, Yang T, Feng Z W, et al. Improvement of elasticity-based dynamic grid method based on mesh quality feedback[J]. Journal of National University of Defense Technology, 2018, 40(1): 10-16(in Chinese).

[83] Hao J X, Wang Q. A class of finite difference schemes with adaptive controllable dispersion and low dissipation for compressible turbulence[J]. Fluids, 2025, 10(2): 43.

[84] Liang T, Shyy W, Fu L. Efficient arbitrary-high-order TENO

schemes with local adaptive dissipation for compressible flow simulation on unstructured meshes[J]. Journal of Scientific Computing, 2025, 104(1): 1.

[85] Wang D D, Du Y M, Jin Y, et al. Comparative study on predicting turbulent kinetic energy budget using high-order upwind scheme and non-dissipative central scheme[J]. Advances in Aerodynamics, 2024, 6(1): 24.

[86] 吴惠松, 林麒, 柳汀, 等. 风洞虚拟飞行试验模型绳系并联支撑机构[J]. 航空学报, 2022, 43(8): 308-320.
Wu H S, Lin Q, Liu T, et al. Wire-driven parallel suspension mechanism of virtual flight test model in wind tunnel[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(8): 308-320(in Chinese).

[87] 林麒, 何聪, 林北辰, 等. 用于旋成体飞行器模型风洞虚拟飞行试验的支撑装置和试验系统: CN114778064B[P]. 2024-06-04.
Lin Q, He C, Lin B C, et al.Support device and test system for wind tunnel virtual flight test of axisymmetric aircraft model: CN114778064B[P]. 2024-06-04(in Chinese).

[88] 杜洪亮, 张伟. 冰风洞虚拟仪器环境参数测试系统设计实现[J]. 测控技术, 2016, 35(3): 8-12.
Du H L, Zhang W. Design and realization of virtual instrument environment parameters measurement system for ice wind tunnel[J]. Measurement & Control Technology, 2016, 35(3): 8-12(in Chinese).

[89] Shin E Y, Howland M F. Accelerated Bayesian calibration and uncertainty quantification of RANS turbulence model parameters for stratified atmospheric boundary layer flows[J]. Boundary-Layer Meteorology, 2025, 192(1): 3.

[90] Xue Y L, Feng Y L, Zheng X J. Uncertainty quantification of compressibility corrections on SA turbulence model for high Mach number flows[J]. Aerospace Science and Technology, 2025, 162: 110184.

[91] 贾洲侠, 吴振强, 吴建国, 等. 飞行器气动热与结构传热双向耦合研究[J]. 强度与环境, 2019, 46(6): 16-23.
Jia Z X, Wu Z Q, Wu J G, et al. Study on two-way coupled fluid-structure-thermal analysis for hypersonic vehicles[J]. Structure & Environment Engineering, 2019, 46(6): 16-23(in Chinese).

[92] Wu J, Zhang Z, Hou A P, et al. Thermal aeroelastic characteristics of inflatable reentry vehicle experiment (IRVE) in hypersonic flow[J]. International Journal of Aerospace Engineering, 2021, 2021(1): 6673818.

[93] Bai Y G, Cao L X, Xu X L, et al. Thermal-aero-elastic analysis of a typical all-moving control surface[J]. Thermal Science and Engineering Progress, 2022, 31: 101297.

[94] Wang W J, Qian W, Bai Y G, et al. Numerical studies on the thermal-fluid-structure coupling analysis method of hypersonic flight vehicle[J]. Thermal Science and Engineering Progress, 2023, 40: 101792.

[95] Guo T Q, Shen E N, Lu Z L, et al. Thermal flutter prediction at trajectory points of a hypersonic vehicle based on aerothermal synchronization algorithm[J]. Aerospace Science and Technology, 2019, 94: 105381.

[96] 王晓鹏. 导弹弹射发射分离轨迹的数值研究[J]. 弹箭与制导学报, 2012, 32(3): 69-72.
Wang X P. The numerical investigation on separation track of missile ejected from aircraft[J]. Journal of Projectiles, Rockets, Missiles and Guidance, 2012, 32(3): 69-72(in Chinese).

[97] 李孝伟, 范绪箕. 基于动态嵌套网格的飞行器外挂物投放的数值模拟[J]. 空气动力学学报, 2004(1): 114-117.
Li X W, Fan X J. Simulation of the release of store based on the moving chimera grid technique[J]. Acta Aerodynamica Sinica, 2004(1): 114-117(in Chinese).

[98] 安效民, 高昕, 屈佑文, 等. 变后掠翼的气动弹性特性分析[J]. 力学学报, 2025, 57(12): 3019-3031.
An X M, Gao X, Qu Y W, et al. Analysis of aeroelastic characteristics of a variable-sweep wing[J]. Chinese Journal of Theoretical and Applied Mechanics, 2025, 57(12): 3019-3031(in Chinese).

[99] 刘昊东, 杨建东, 赵文文, 等. 变构型飞行器增量非线性动态逆姿态控制律设计[J]. 战术导弹技术, 2024(4): 83-91.
Liu H D, Yang J D, Zhao W W, et al. Incremental nonlinear dynamic inversion attitude control law design for morphing aircraft[J]. Tactical Missile Technology, 2024(4): 83-91(in Chinese).

[100] 李墨吟, 马泽远, 周建平, 等. 基于神经网络的变后掠翼飞行器自适应控制方法研究[J]. 弹箭与制导学报, 2021, 41(5): 73-77.
Li M Y, Ma Z Y, Zhou J P, et al. Research on adaptive control method of variable-sweep wing aircraft based on neural network[J]. Journal of Projectiles, Rockets, Missiles and Guidance, 2021, 41(5): 73-77(in Chinese).

[101] Li C X, Wu H N. A dynamic neural network-based control method using reinforcement learning for nonlinear parameter-varying system with application to morphing aircraft[J]. IEEE Transactions on Neural Networks and Learning Systems, 2025, 99: 1-13.

[102] Cheng X M, Wang Y X, Zhao J, et al. Data-driven adaptive hybrid-based control for some tiltrotor UAVs in the low-speed flight[J]. Aerospace Science and Technology, 2026, 170: 111513.

[103] Yu J S, Yu J J. Neural learning control of fighter aircraft at high-angle-of-attack maneuvers with online recorded data and disturbance observer[J]. IEEE Access, 2026, 14: 39702-39712.

[104] Deng S H, Xiao T H, VAN Oudheusden B, et al. A dynamic mesh strategy applied to the simulation of flapping wings[J]. International Journal for Numerical Methods in Engineering, 2016, 106(8): 664-680.

[105] Addo-Akoto R, Han J S, Han J H. Roles of wing flexibility and kinematics in flapping wing aerodynamics[J]. Journal of Fluids and Structures, 2021, 104: 103317.

[106] Li G, Law Y Z, Jaiman R K. A novel 3D variational aeroelastic framework for flexible multibody dynamics: application to bat-like flapping dynamics[J]. Computers & Fluids, 2019, 180: 96-116.

[107] Xue Y J, Cai X F, Xu R, et al. Wing kinematics-based flight control strategy in insect-inspired flight systems: deep reinforcement learning gives solutions and inspires controller design in flapping MAVs[J]. Biomimetics, 2023, 8(3): 295.

[108] Chen A, Song B F, Liu K, et al. Flapping-wing robot achieves bird-style self-takeoff by adopting reconfigurable mechanisms[J]. Science Advances, 2025, 11(36): eadx0465.

[109] 刘笑尘, 何心怡, 何国毅, 等. 蜻蜓非均匀柔性前翼扑动飞行的气动性能研究[J]. 力学学报, 2022, 54(4): 991-1003.
Liu X C, He X Y, He G Y, et al. The aerodynamic performance of non-uniform flexible dragonfly's forewing in flapping flight[J]. Chinese Journal of Theoretical and Applied Mechanics, 2022, 54(4): 991-1003(in Chinese).

A review of numerical virtual flight studies for aerodynamic, motion and control coupling in complex flight conditions

SHAO Shuai*，MEI Yafei，YAO Mengyu，GAO Xianzhong，BAO Lei

(National University of Defense Technology Test Center，Xi'an 710106，China)

Abstract：High-fidelity simulation of an aircraft's dynamic reaction characteristics under challenging flight conditions is made possible by numerical virtual flight (NVF) technology, which is essential for designing control laws and assessing flight quality. This paper systematically reviews the NVF technology coupled with aerodynamics, motion, and control, with an emphasis on the application of this coupling method in complex flight problems. This paper outlines the basic framework and technical workflow of NVF, elaborating on the evolution of dynamic mesh technology from single deformation to dynamic hybrid mesh systems, as well as its critical role in ensuring relative motion of components and computational stability, while also pointing out the enabling potential of machine learning in mesh quality control. Regarding unsteady flow field solutions, this paper discusses the coupling modes of aerodynamics, motion, and control, introduces the technical evolution from loose coupling to tightly coupled closed-loop solutions, and highlights that the tightly coupled method can significantly improve the simulation accuracy for high-dynamic flight states such as high angle of attack and post-stall. This paper elaborates on the application practices of this coupling method in three major scenarios: hypersonic multi-field coupling, morphing flight, and flexible bionics. NVF will advance toward effective, unified, and full-system multi-field coupling in the future thanks to artificial intelligence technologies, hastening the shift in aircraft development paradigms toward simulation-driven, digital-intelligent testing.

Keywords：complex flight；multi-body separation；flexible morphing；multi-physics coupling；unsteady flow

Received：2026-03-02；**Accepted**：2026-05-08；**Published Online**：2026-06-05 18：40

URL：link.cnki.net/urlid/11.2625.V.20260605.1251.001

Foundation items：General Program of Shaanxi Provincial Natural Science Basic Research Plan (2024JC-YBQN-0030); Independent Innovation Science Foundation Project of National University of Defense Technology (25-ZZCX-JDZ-66); Youth Independent Innovation Science Foundation Project of National University of Defense Technology (ZK25-106)

* **Corresponding author**. E-mail：shaoshuai_ss@nudt.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2026.0133

液流电池用压滤式电化学反应器传质优化研究进展

刘君[1,2]，韩治成[2]，相艳[1,*]，卢善富[1]，王海宁[1]

(1. 北京航空航天大学 能源与动力工程学院 仿生能源材料与器件北京市重点实验室，北京 100191；
2. 中国恩菲工程技术有限公司，北京 100038)

摘 要：压滤式电化学反应器因结构简单、模块化程度高、槽电压低等优势，成为液流电池主流反应器构型，并在储能电池、电化学水处理、合成化工等多个领域呈现出较广阔的应用前景。其中，全钒液流电池是商业化程度最高、研究体系最为成熟的代表，但其性能提升仍受限于传质效率不足所引发的浓差极化加剧、电压效率偏低及长期循环稳定性衰减等瓶颈问题，直接影响系统成本的控制。为此，主要以全钒液流电池为对象开展分析与讨论，从核心组件与运行系统2个角度，系统梳理电极、流场、隔膜等结构传质优化研究进展，分析流量、温度、电流密度、电解液浓度及罐堆比例等运行参数对传质过程的调控规律，明确各组件传质特性与电池整体性能的耦合关系，提出组件结构与系统运行协同的传质改进方向。从工程应用视角指出，全钒液流电池传质优化需在反应器组件改良的基础上，同步考量电堆尺寸放大效应、罐堆匹配、运行参数调控及系统能耗平衡，旨在为压滤式液流电池反应器的工程化设计与性能提升提供理论参考和技术参考。

关 键 词：压滤式电化学反应器；钒液流电池；传质优化；结构设计；运行参数
中图分类号：TM911
文献标志码：A **文章编号**：1001-5965(2026)10-3453-18

压滤式电化学反应器是由平行板极板、电极、绝缘隔板、支撑框架等单元模块化重叠，经加压密封组装而成的电化学装置[1]。该构型的反应器具备电场分布均匀、槽电压低、功率与容量可独立设计的结构优势，已广泛应用于液流电池储能[2-3]、电化学废水处理[4]、电化学合成[5]、电解水制氢[6]等领域。在当前商用液流电池领域，压滤式电化学反应器已成为主流结构[7]。例如，中国科学院大连化学物理研究所钒液流电池研究团队文献[8-10]表明，压滤式反应器作为主流的堆栈设计，已被应用于高功率钒液流电池领域，有助于实现快速的电解液流动并提升能量效率；中国石油大学徐春明团队研究的铁铬液流电池关键材料和组件也采用了压滤式反应器结构[11]；北京航空航天大学相艳团队在催化电极改性[12-13]、膜改性[14]、多金属氧酸盐电解液材料[15-16]等研究中，压滤式电化学反应器均是主要的评价设备。压滤式电化学反应器赋予电池诸多特性，如循环寿命长[8]、运行安全性高[17-18]、充放电过程解耦[19]等，是液流电池实现工程化应用的核心载体。

液流电池按活性物质体系可分为全钒、铁铬、锌溴等类型，其中，钒液流电池因活性物质氧化还原可逆性好、电解液稳定性高、无跨金属离子污染等优势，成为目前唯一实现商业化应用的液流电池技术[20-21]，铁铬、锌溴等液流电池则多处于实验室研究与初步示范阶段，尚未形成成熟的工程化应用体系，亟待电化学工程理论的系统性指导。钒液流电池系统主要由电堆、能量存储罐、循环泵组、管道系统、功率转换设备及能量管理系统构成。能量效率、容量保持率与循环稳定性是液流电池的核心性能指标，而传质效果通过不同物理-电化学耦合

收稿日期：2026-03-09；录用日期：2026-07-10；网络出版时间：2026-07-30 17:59
网络出版地址：link.cnki.net/urlid/11.2625.V.20260730.1343.001
*通信作者. E-mail：xiangy@buaa.edu.cn

引用格式：刘君，韩治成，相艳，等. 液流电池用压滤式电化学反应器传质优化研究进展[J]. 北京航空航天大学学报，2026，52(10)：3453-3470. Liu J，Han Z C，Xiang Y，et al. Research progress on mass transfer optimization of filter-press type electrochemical reactors for flow batteries[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3453-3470 (in Chinese).

机制，对三者均产生深刻影响[22]。

传质对能量效率的影响：传质能力直接决定了浓度过电位的大小。优化的传质设计（如提高对流强度）能加速活性物质从电解液主体向电极表面的输送，有效降低浓差极化，从而提升电压效率，最终提高电池的能量效率。但需注意，过度强化传质会显著增加泵功损耗，因此，传质优化的核心在于寻求“电化学性能提升”与“系统能耗增加”之间的最佳平衡点[23-24]。

传质对容量保持率的影响：传质的空间均匀性是决定活性物质利用率的关键。不良的传质会导致电极局部出现“传质死区”（反应物供应不足的区域），使该区域活性物质无法参与反应，导致有效容量损失；同时，局部反应物耗尽会加剧析氢、活性物质不可逆沉淀等副反应，进一步加速容量衰减。因此，提升传质均匀性是保障高容量保持率的重要前提[25-26]。

传质对循环稳定性的影响：循环稳定性受多重因素制约，传质均匀性是其中关键的调控因素之一，而非唯一决定因素。具体而言，不均匀的传质会引发电极表面局部电流密度分布不均，产生热点区域，加速该处电极材料（如碳毡纤维）的氧化腐蚀和结构坍塌，同时，加剧隔膜两侧的活性物质浓度梯度，促进交叉污染。通过优化流场设计和电极结构来改善传质均匀性，可从源头上缓解这些局部衰减诱因，从而显著延长电池的循环寿命[27]。

1 传质优化多维度评价

为推动液流电池的规模化发展，亟需从反应器组件设计、运行调控和系统工程约束等层面开展针对传质优化的系统性研究，表1归纳了从多个维度进行传质优化的评价指标及其物理意义和对应的优化对象[21,24,26,28-46]。

表1 液流电池传质优化的多维度评价指标、物理意义及优化对象

Table 1 Multidimensional evaluation criteria, physical significance and optimisation targets for mass transfer optimisation in flow batteries

评价维度	核心评价指标	主要物理意义	对应优化对象
电极孔道与界面传质[28-31]	孔隙率、渗透率、有效扩散系数、曲折度、比表面积、润湿性	表征电解液进入多孔电极、活性物质扩散及电极/电解液界面反应物补给能力。极限电流密度、传质系数为该类本征参数的综合外在表现。	电极厚度、孔径尺寸、孔隙率、压缩率、纤维直径、微纳结构、活性位点
流体流动与空间均匀性[26,32-34]	速度分布、浓度/荷电状态（state of charge, SOC）分布、电流密度分布、肋下流速、流动均匀性因子、死区比例	表征流场对电极内反应物补给均匀性、局部贫液区形成和反应分布一致性的调控能力	流道结构选型与几何参数优化
隔膜选择性传质[35-38]	离子电导率、钒离子渗透率、离子选择性、膜阻、跨膜迁移通量	表征隔膜在离子传导、活性离子阻隔和交叉污染控制之间的平衡能力	离子交换膜、复合膜、多孔隔膜、选择性传输通道
运行参数与传质匹配[21,24,39-41]	流量、温度、电解液浓度、SOC、充放电电流密度	表征运行条件通过影响电解液物性（如黏度）及对流强度，对反应物补给速率、扩散能力和浓差过电位的调控作用	流量工作窗口、温度区间、活性物浓度配比、SOC运行范围、充放电制度等
系统级传质与工程约束[42-46]	多堆流量分配偏差、堆间温度偏差、旁路电流、泵基能耗占比、系统净效率	表征由单电池/单电堆放大至多堆串、并联系统后，因管路结构、热质耦合与电耦合等引发的流量不均衡、电和热管理困难及寄生能耗增加等工程约束问题。	多堆并联拓扑、歧管管路设计、管路阻力匹配、热管理策略、旁路电流控制、泵送系统

尽管不同液流电池体系所用电解液与活性物质存在差异，但从多孔电极内反应物补给、流场均匀化、隔膜选择性传输及传质-压降-效率权衡等底层问题看，各类液流电池具有一定共通性[17-19,22]。鉴于全钒液流电池商业化程度较高、研究基础较为系统，本文以其为研究对象，对压滤式电化学反应器的传质优化策略进行归纳，同时，讨论其对其他液流电池及液流式电氧化/还原反应器的借鉴意义。

2 传质结构参数优化

液流电池传质控制的本质是反应物由主体电解液向电极反应界面补给不足导致的浓度过电势问题，其核心解决思路为通过关键组件的结构优化提升反应物可达性，降低局部传质阻力，并改善浓度分布均匀性。本节按照“组件结构特征-传质瓶颈-优化策略-工程约束与发展方向”的逻辑展开。电极部分重点讨论孔隙率、厚度、压缩率和微观结构对有效反应面积与电解液渗透的影响；流场部分重点比较不同流道对电极对流、浓度均匀性和压降的调控作用；隔膜部分则聚焦质子选择性传输、钒离子阻隔及长期稳定性之间的平衡。

2.1 电极结构

电极是钒液流电池电化学反应的核心区域，电化学反应主要发生在电极孔道电极与电解液的界面，其结构同时决定电解液浸润、活性物质扩散、电子传导和界面反应动力学。理想电极需具备亲

水性优异[47]、电催化活性高[48]、界面传质能力强[49]的三重耦合特征, 以实现物质/电荷传递与界面电化学反应的协同进行。目前, 商用钒液流电池电极主要采用石墨毡或碳毡材料[50], 但其纤维表面光滑、本征润湿性有限、活性位点不足, 难以满足高电流密度下的快速反应和均匀传质的需求[48,51]。因此, 电极优化主要从宏观结构调控与微观界面改性 2 个层面展开。宏观层面聚焦于碳基毡的孔隙率、厚度、孔径尺寸及压缩率, 微观层面则通过表面修饰、构建纳米级孔道与活性位点阵列[52-53], 强化电极与电解液的界面作用及离子传输能力。

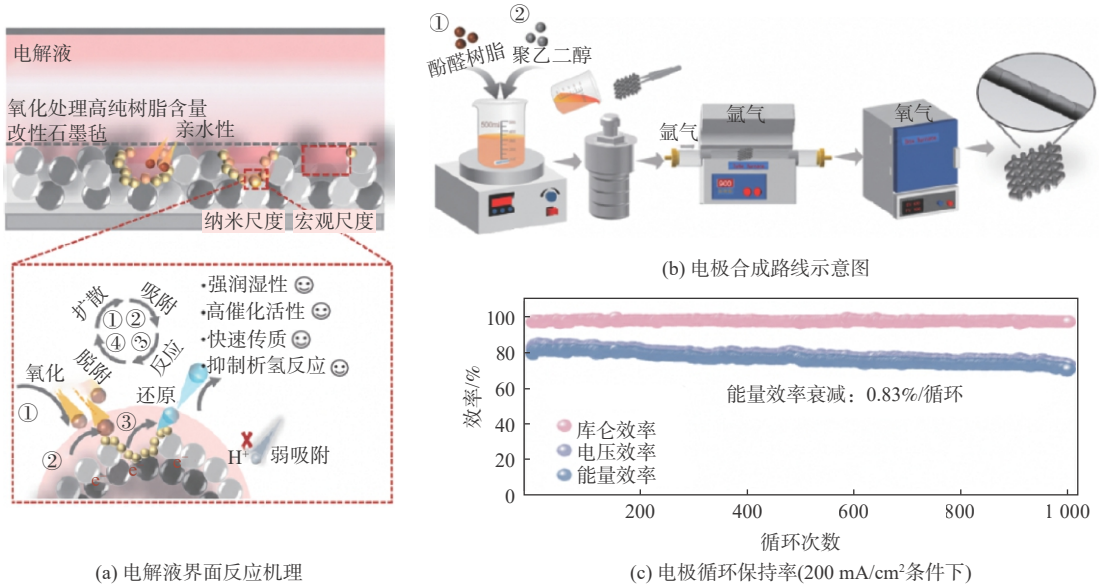
电极组装压缩率是宏观结构优化的重要考量因素, 适量的压缩率可降低电极与隔膜、集流体之间的接触电阻, 提升传质效率, 但过度压缩会降低电解液的渗透率, 增加传质阻力, 同时, 导致电池压降上升[54]。因此, 电极优化结构不能单纯追求低接触电阻或高比表面积, 而应在厚度、孔隙率与组装压缩率之间建立平衡, 实现接触电阻与传质阻力的协同降低[30]。孔隙率与厚度直接影响电极内部液相渗透、有效反应深度和电流密度分布, 是决定多孔电极传质性能的核心宏观参数。

Alphonse 等[55] 研究了孔隙率沿厚度方向变化对电池性能的影响, 当电极孔隙率沿厚度方向在 64%～94% 范围内变化时, 孔隙率下降会导致电流密度分布不均并呈阶梯式降低, 电池电压随之下降; 当孔隙率低于 64% 时, 电极渗透性急剧变差, 传质阻力大幅增加, 电池电压出现骤降。高孔隙率电极可形成厚电极层, 允许更多电解液渗透, 有效增强传质过程, 提升电池电压。Yaji 等[56] 通过数值模拟与拓扑优化方法, 研究了蛇形流道下电极厚度与孔隙率的耦合影响, 确定钒液流电池电极的最佳厚度为 3 mm、最佳孔隙率为 90%。香港科技大学 He 等[57] 通过三维数值模拟揭示了全钒液流电池电极孔隙率、纤维直径与流场的耦合优化机制, 提出梯度孔隙率设计(膜侧低、流场侧高)及大纤维直径配合表面改性的策略, 可同时降低浓差极化与压降, 显著提升大尺度电池性能。高孔隙率(95%)可最小化浓差极化, 提升大尺度电池孔隙率(80% 以上)可解决反应集中在流场侧分布不均的问题, 大纤维直径与高比表面积改性可以同时提高渗透率和降低活化损失。Ali 等[58] 通过建立二维等温物理模型, 推导了电极比表面积和渗透率与孔隙率及纤维直径的物理关联式, 系统分析表明, 当孔隙率控制在 0.8～0.92 且纤维直径为 40～70 μm 时, 可在降低泵送能耗的同时, 有效抑制电化学过电位。Martinez Lopez 等[59] 通过二维数值模型揭示了电极压缩比(20%～50%)和电解液流量(10～40 mL/min)对全钒液流电池性能的影响规律, 50% 压缩比使反应速率提升 9.2%、流速提升 12.7%, 但过电位增加; 低流量(10 mL/min)虽增强浓度梯度, 但导致过电位升高, 需权衡优化。

梯度化电极与定向排布电极为解决“反应面积与传质阻力”矛盾提供了新的结构思路。南方科技大学赵天寿团队[28] 研究发现, 梯度化电极可使孔隙率与渗透性能由隔膜侧向流道侧逐步提升, 从而增强活性物质从流道区域向膜侧界面的输运; 同时, 比表面积由流道侧向隔膜侧逐渐增大, 可保障核心反应区域拥有充足的活性位点。Sun 等[60] 通过定向排布碳纤维降低电极曲折度、缩短离子迁移路径, 并强化垂直膜面方向和肋下区域的对流补给, 使电流密度分布更加均匀。由此可见, 电极厚度、梯度孔结构、孔隙率和纤维曲折度的最优值并非固定常数, 而是随流场形式、流量、电流密度和温度等运行条件变化的结构-运行耦合结果。微观界面改性的核心目标是在不显著增加流动阻力的前提下, 提高电极与电解液界面可及性、润湿性和局部反应活性。微米级通道有利于电解液对流渗透, 纳米级孔道则为离子提供短程快速扩散路径并增加反应界面。同时, 含氧/含氮官能团[61]、碳纳米结构或金属氧化物[62] 等改性可改善界面电子结构和反应动力学。微纳结构改性并不等同于单纯增加活性位点, 其工程价值在于同时提升界面反应速率、反应物补给能力和结构稳定性。另外, 电极材料在长期循环中, 表面含氧/含氮官能团可能发生脱落或重构, 微纳结构可能出现堵塞、塌陷或机械损伤, 从而降低电解液浸润性和反应界面可及性。因此, 电极结构设计应兼顾高比表面积、快速传质与结构耐久性, 避免仅追求初始活性而牺牲长期稳定性。天津大学 Wang 等[31] 通过酚醛树脂涂层与聚乙二醇造孔刻蚀策略, 精准调控热氧化刻蚀过程制备微纳米分级多孔石墨毡电极, 改性后的电极具有更高的开路孔隙率、吸水率与压缩率, 不仅有利于电解液渗透并到达反应界面, 还能提升电极柔韧性, 减少对隔膜的机械损坏; 该电极在 200 mA/cm^2 高电流密度下的能量效率达 80.41%, 经 1 000 次循环后, 容量保持率仍在 90% 以上, 从而实现双重功能结构, 如图 1 所示。

2.2　流场优化

压滤式液流电池的流道作为极板上电解液的传输与分配的主要通道, 表征流场对电极内反应物补给均匀性、局部贫液区形成和反应分布一致性的调控能力。目前, 全钒液流电池的主流流道类型包



(a) 电解液界面反应机理 (b) 电极合成路线示意图 (c) 电极循环保持率(200 mA/cm²条件下)

图 1 氧化处理的高纯树脂含量改性石墨毡制备、性能及反应机理[31]

Fig. 1 Preparation, properties and reaction mechanisms of high-purity, oxidised modified graphite felt [31]

括平行流道、叉指流道、蛇形流道及多种异形流道。平行流道压降低，但电解液难以渗透进入电极内部，易造成传质不均；蛇形流道可形成较强肋下对流，但压降较高；叉指流道可强化穿流传质但依赖电极渗透率；多场协同、仿生和人工智能(artificial intelligence, AI)优化流道则试图在传质均匀性、压降和可制造性之间取得综合平衡[26,63]。从传质机理看，流场优化主要通过 3 条路径发挥作用：①改善入口至出口方向的流量分配，降低局部低浓度区；②增强肋下或电极厚度方向的对流渗透，缩短活性物质到达反应界面的路径；③在提高传质的同时，控制压降和循环泵功。因此，流场优劣不能脱离电极厚度、孔隙率、流量、电流密度和目标系统效率进行判断。

传统蛇形流道是应用较广的平流流道类型，其优势在于长流道和相邻通道压差可诱导肋下横向渗透，从而增强局部反应物补给；但其不足也较突出，即肋下对流分布不均。弯道附近可能因压力梯度不足形成传质死区，而非弯道连接区域又可能因较大横向压差形成局部过反应，导致反应速率热点与低反应区并存，影响钒液流电池电压效率。米兰理工大学 Messaggi 等[32]通过三维计算流体动力学(computational fluid dynamics, CFD)模型与实验验证揭示了蛇形流场(serpentine flow field, SFF)肋下区域的旁路对流机制，明确了流场设计、电极形貌与电化学性能之间的强耦合关系，为蛇形流道的局部强化设计提供了理论依据。

针对蛇形流道的缺陷，目前的优化思路主要为通过流道形状、流道变径、设置堵块等方式重构电极内的压力分布，平衡肋下流动压力梯度，增强电解液的横向渗透与分布均匀性。北京化工大学 Wang 等[64]通过三维数值模型系统对比了半圆形、矩形和三角形流道截面设计，发现半圆形流道凭借更高的雷诺数和优异的水动力学特性，显著增强了电解液渗透与活性物质分布均匀性(浓度均匀性提升 15.4%)，从而有效降低过电位，实现最低充电电压和最高放电电压。Lu 等[65]提出旋转 SFF 的创新设计，通过旋转变径的流道布局，强制电解液在电极厚度方向与横向实现深度渗透，该流场在低流速(2 mL/s)下的压损介于传统 SFF 与叉指流场之间，实际能效比传统 SFF 提升 4.2%～12.6%，兼具 SFF 低压降与叉指流场强制对流的双重优势。Karami 等[66]通过降低通道高度与在低压力梯度区域设置全堵塞阵列优化蛇形流道，在弯道处引入 100% 高度堵塞块、中点处设置 80% 堵塞块，有效增强了肋下对流强度，使电极中平面流速均匀性指数从 0.68 提升至 0.802，电池净功率密度提升 56.24%(达 276.3 mW/cm²)，如图 2 所示。SFF 的传质效率依赖电解液在肋下的深度渗透，其性能对电极厚度具有较高敏感性，设计时需实现流道结构与电极厚度的匹配[67]。香港科技大学赵天寿团队提出了一种沿流路梯度设计的 SFF[68]，通过将流道深度从入口到出口线性减小(优化梯度为 25%)来加快流速并增强下游对流传质，从而显著改善全钒液流电池的反应物分布均匀性，使能量效率和电解液利用率分别提升 5.0% 和 27.7%。天津大学 Wang 等[33]通过三维多物理场耦合模型阐明了旋转 SFF 几何参数(通道分数为 0.25、深度为 1 mm)对肋下对流的调控机制，使钒液流电池在 150 mA/cm² 和 40 mL/min 条件下达到 91.4% 的泵基效率，为高功率密度流场设计提供了

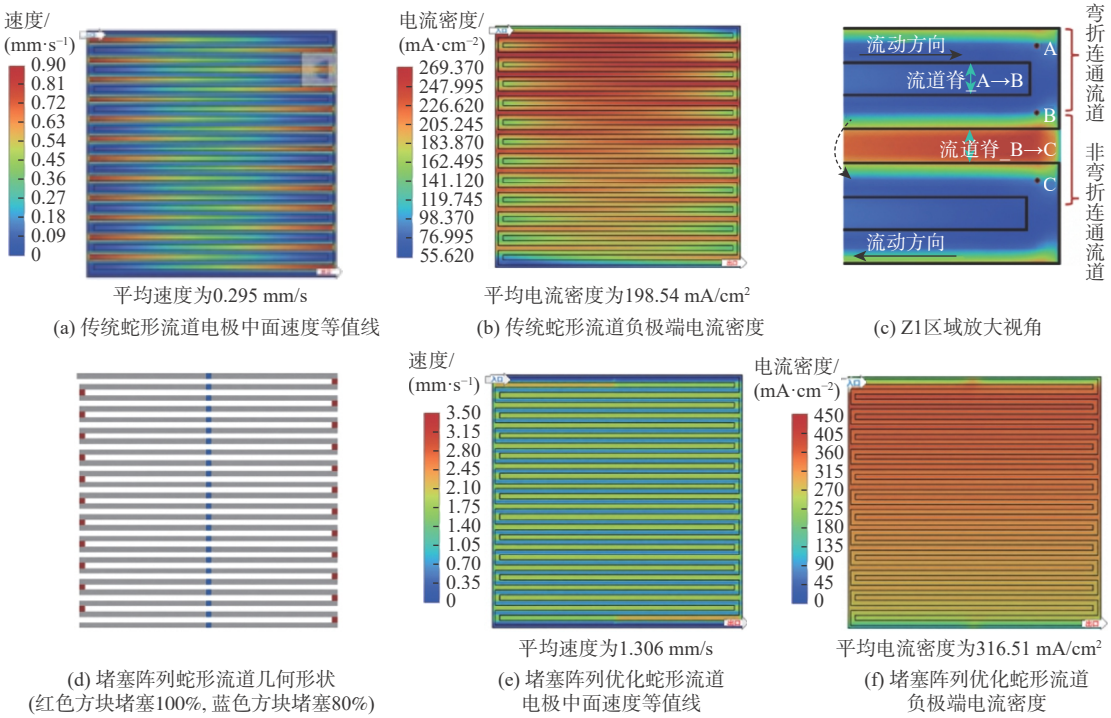


图 2　全堵塞阵列优化蛇形流道(工况流量 Q=20 mL/min)[66]

Fig. 2　Fully-blocked array optimised serpentine flow channels (Q=20 mL/min)[66]

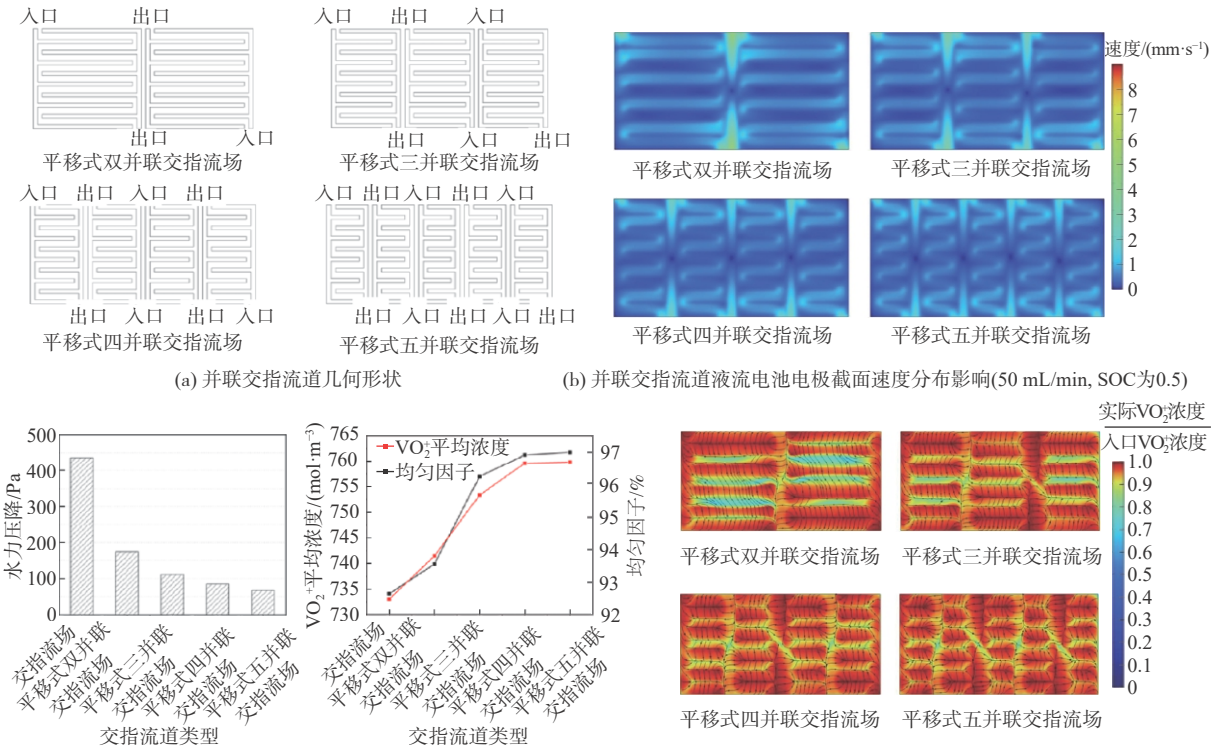
理论指导。蛇形流道对于高流速工况及厚电极结构展现出独特的适配性[21,30]，这是因为电解液主要通过扩散和少量渗透进入电极，由于厚电极自身的结构特性，并不需要电解液以强制对流的方式深度穿透电极来完成反应，蛇形流道的长且曲折的路径设计能够很好地适应这一需求。蛇形流道在工程层面主要约束来自面积放大后带来的局部热点与死区增多、压降大及泵耗高等问题。

由图 2 可见，堵塞阵列并非简单增加流动阻力，而是通过在低压力梯度区域重构局部流线，使原本集中于主流道附近的电解液向肋下区域渗透，从而改善电极中面速度分布和负极端电流密度分布。该结果说明，SFF 的优化重点不仅是延长流道或提高流量，而是在局部传质薄弱区域建立适度的横向压力梯度，使热点区和死区趋于均衡。

叉指流道的结构特征为进出口不直接相连，电解液经分流流道后需穿透电极才能汇流流出，是典型的穿流流道类型，该结构可增强厚度方向对流传质，尤其在低流量、高电流密度或薄电极条件下有助于缓解浓差极化，但其效果并非必然优于蛇形流道，而是取决于电极厚度、孔隙率、渗透率、压缩率和流量分配。若中部区域供液不足，局部反应速率仍可能降低[63]。Ishitobi 等[69] 搭建了二维稳态全电池模型，实现叉指流道与 0.2 mm 薄碳纸电极耦合的跨平面方向可视化，并提出通过实验数据拟合边界层传质系数关联式，为叉指流道与薄电极的匹配设计提供了方法。相较于蛇形流道，叉指流道的系统压降更低，从工程化系统角度，其能更好地实现泵功与电解液分布均匀性的平衡，更适合工业级钒液流电池的应用。

针对叉指流道中部传质薄弱的问题，优化方向主要为通过流道结构重构提升中部区域的流量与对流强度，在不增大压降的前提下强化中部区域反应物补给。吉林大学杨帆[70] 设计多组并联叉指流道，通过增加流道并联数量，有效改善了传统叉指流道电极中部低浓度区域大面积聚集的问题，使活性物质低浓度区域仅出现在分支出口流道单侧的肋下电极，且随着并联数量增加，低浓度区域更分散，浓度梯度更低；研究还发现，并联组数越多，系统压降越小，当并联四组时，VO_2^+平均浓度与分布均匀因子趋于平衡，实现传质效率与压降的最优匹配，如图 3 所示。并联交指结构通过分散入口与出口支路，削弱了传统交指流场中部低浓度区的大面积集聚，使VO_2^+浓度分布更加均匀；同时，多并联支路缩短了单支路流动路径，有利于降低水力压降。该结果表明，叉指流场的优化应重点关注分支数量、支路长度和出口布置对局部穿流强度的影响，以在强制对流和泵功之间取得平衡。香港科技大学赵天寿团队提出了一种分叉交指流场设计[71]，通过分级分流策略将电解液均匀分配至分支通道，使钒氧化还原液流电池压降降低 45%、泵基电压效率提升至 73.10%(250 mA/cm² 时)，且随活性面积增大优势更显著，展现出优异的规模化应用潜力。

多场协同流场的设计思想来源于场协同原理，



(a) 并联交指流道几何形状

(b) 并联交指流道液流电池电极截面速度分布影响(50 mL/min, SOC为0.5)

(c) 并联交指流道液流电池水力压降与电极内VO_2^+浓度分布均匀系数(50 mL/min, SOC为0.5)

(d) 并联交指流道液流电池截面VO_2^+浓度分布与流线分布

图 3 并联交指流道几何及性能参数[70]

Fig. 3 Geometry and performance parameters of parallel interdigitated-flow channels[70]

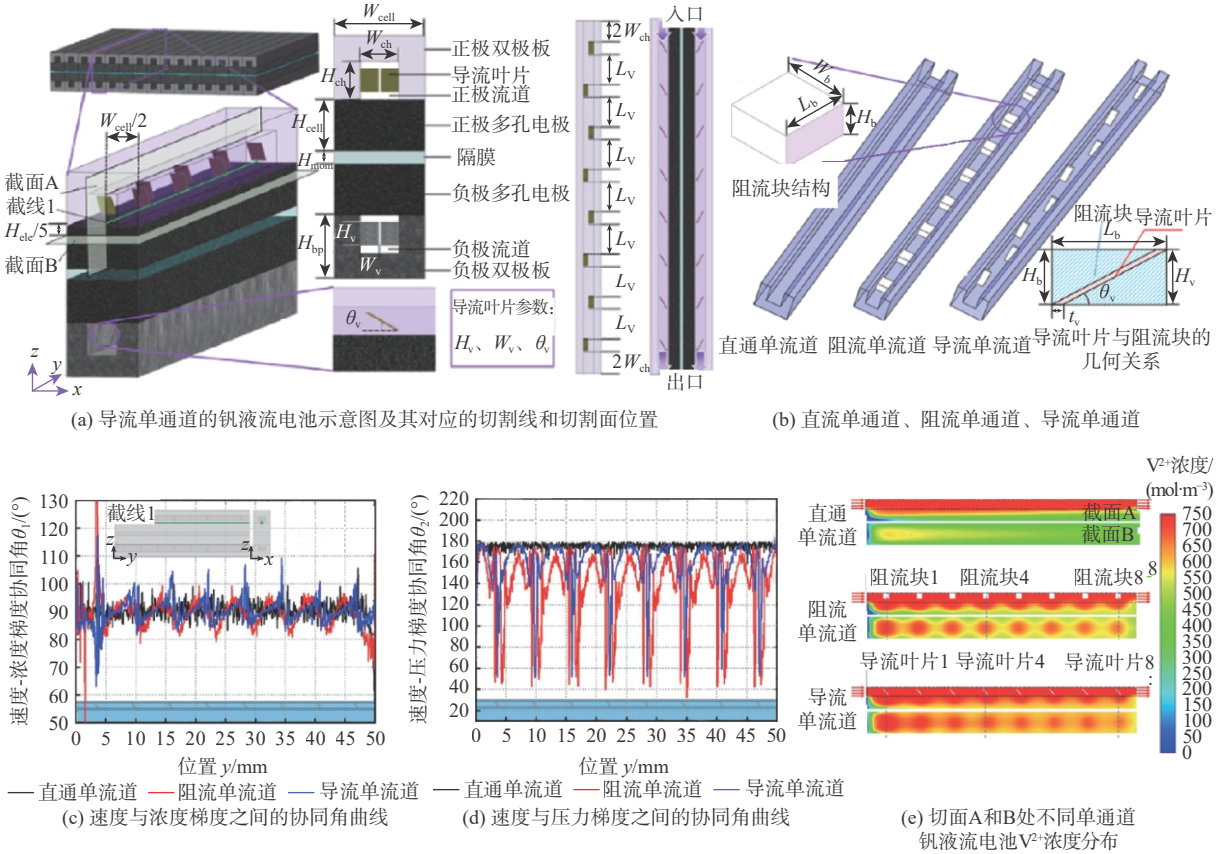
其核心不是单纯提高流速，而是改善速度场、浓度场、压力场之间的匹配关系，实现传质增强与压力损失降低的协同[34]。对于反应物消耗导致的浓度梯度，电解液速度方向应尽可能指向低浓度反应区，即与数学意义上的浓度梯度方向趋于相反，或与扩散通量方向趋于一致，以增强活性物质补给，并改善浓度分布均匀性；同时，速度方向应尽可能沿压力下降方向流动，使压力能更多用于有效输运，减少局部回流、滞流和无效压降。因而，多场协同流道的核心目标是在提高速度-浓度协同程度的同时，降低无效流动阻力，实现传质强化与泵功降低的协同优化。

北京理工大学 Cheng 等[34]设计的导流单通道的速度-浓度协同角比传统流道提升 2.7°，速度-压力协同角仅损失 0.11°，系统压力损失降低 11.73%，电极内 V^{2+}浓度提升 32.34%；该设计在实现速度-浓度协同提升(+1°～5°)的同时，几乎不损害速度-压力协同(−0.05°～0.08°)，实现了“降损增效”的优化目标，如图 4 所示。多场协同的设计思路也被应用于平行流道优化，如通过设计波纹流道调控电解液的流动速度、压力与流动角度，实现平行流道的传质强化[26]。多场协同流场适合指导流场设计，但工程化落地必须经过“结构简化、可制造性验证、压降/泵功核算、大面积电池验证和电堆级验证”。

针对传统流道局部浓度梯度大、浓差极化严重、电池效率与稳定性偏低的问题，研究者突破蛇形、叉指、平行流道的基本框架，提出仿生[67]、双螺旋[72]、AI 拓扑优化[63]等异形流场。该类流场的共同目标是通过分级分流、曲线导流或算法筛选，提高电解液覆盖率，降低局部贫液区，并优化压降分布。仿生流场借鉴植物叶脉等天然输运网络，利用主干-分支通道实现多尺度分配；AI 优化流场则通过高通量生成、性能预测和实验筛选提高设计效率。

Huang 等[67]借鉴植物叶脉的营养输送机制，设计了主通道添加圆形障碍物的仿生叶脉流场，仿生叶脉通过主干-分支通道模拟植物叶脉分级输运，使电解液更充分地覆盖电极表面和出口区域；主通道圆形障碍物进一步诱导电解液进入分支流道，显著提升了电解液在电极表面的分布均匀性，减少局部浓度贫化。研究表明，该仿生叶脉流场的能量效率最高，极化损失最小，电极内 V^{2+}平均浓度高于传统 SFF，尤其在电极出口区域的浓度分布更均匀，如图 5 所示。仿生流场的优势主要体现在分布均匀性和低浓差极化，而其工程应用仍需关注加工复杂度、密封可靠性和大面积放大时的流量一致性。

AI 驱动设计是近年来异形流场开发手段的前沿研究范式，可有效突破人工设计的局限性，实现流场结构的高通量筛选与优化，提高液流电池性能。



(a) 导流单通道的钒液流电池示意图及其对应的切割线和切割面位置　(b) 直流单通道、阻流单通道、导流单通道

(c) 速度与浓度梯度之间的协同角曲线　(d) 速度与压力梯度之间的协同角曲线　(e) 切面A和B处不同单通道钒液流电池V^{2+}浓度分布

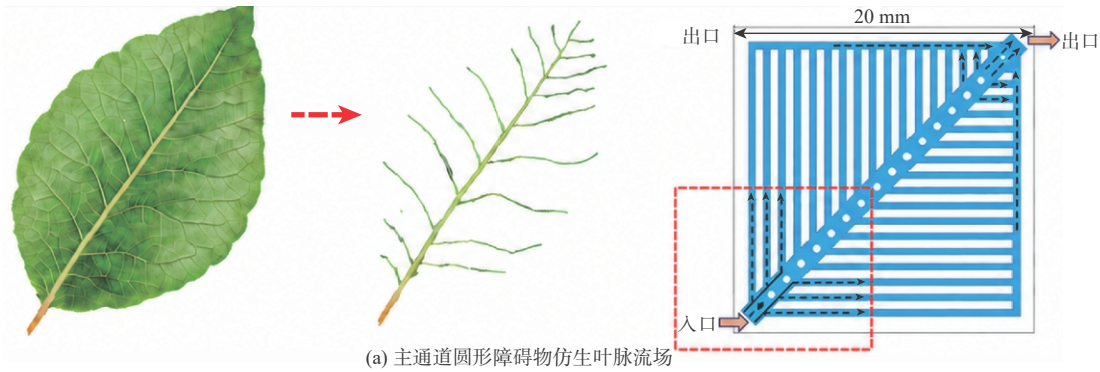
图 4　多场协同导流流道设计与液流电池研究结果[34]

Fig. 4　Research findings on the design of multi-channel co-flow channels and liquid-flow batteries[34]

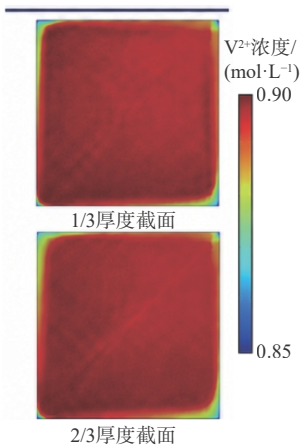
中国石油大学徐春明院士团队提出了融合梯度提升机器学习与专用大语言模型的液流电池 AI 驱动设计方法，通过数据-机理双驱动建模实现了铁铬液流电池性能的精准预测、关键影响因素识别与智能优化[73-76]。香港城市大学 Zheng 等[63] 开展钒液流电池流场的端到端 AI 驱动设计，构建流场库生成算法与误差＜1.5% 的卷积神经网络（convolutional neural network, CNN）性能预测器，通过算法探索了 11 564 种异形流场结构，从中筛选出 8 种最优结构。其中，HTS01 为第一轮筛选得分最高且专家投票排序最高的代表流场，可视为该研究中的综合最优候选结构。实验验证表明，HTS01 在 50～200 mA/cm^2 范围内均表现出高于传统 SFF 和叉指流场的能量效率，其优势主要来自更强的肋下对流、更高的反应物分布均匀性和更低的浓差极化。同时，HTS04 在电解液利用率方面表现突出，在 100 mA/cm^2 下电解液利用率达到 78.3%，较传统 SFF 提高 10.8%。因此，该研究表明，AI 筛选流场的优化目标并非单纯提高流速，而是在传质均匀性、压降和电解液利用率之间实现综合权衡。如图 6 所示。AI 驱动设计的优势在于将“结构生成-多物理场仿真-性能预测-实验验证”连接为闭环，可在大量候选流场中筛选兼具高极限电流密度、较高能量效率和较低压降的结构。该方法尤其适用于传统经验设计难以覆盖的异形流场空间，但其预测可靠性仍依赖训练数据质量、边界条件一致性和后续实验验证。

不同流道结构在不同电极和运行状态下有不同优势，当前研究多为单独优化流道或电极，极少研究将流道与多孔电极耦合进行协同。电极-流场耦合机制体现为多孔电极孔径分布、孔隙率、各向异性和压缩状态与流场几何结构、肋宽比、流动方向之间的相互制约，两者共同决定电解液分布、肋下传质、局部反应速率和压降。未来流场优化应从单独改变流道形状转向“流场-电极-流量-电流密度”的协同设计。

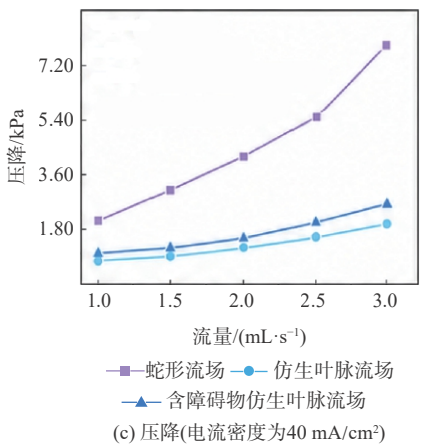
基于表 2 的对比结果，流场优化不能仅以局部传质增强程度作为唯一判据，而须在传质强化、压降损失、泵送能耗、可制造性及工况适应性之间进行综合权衡。从表 2 可看出：平行流场虽压降最低、泵耗最小，但肋下与远端区域传质不足，适用于对泵耗敏感但功率密度要求不高的场景；SFF 传质能力最强，却以显著泵送损耗为代价，适宜追求高功率密度的流经型设计；叉指流场通过强制穿透电极实现传质增强，但对电极渗透率高度敏感，其压降特性随电极压缩率和厚度变化显著；多场协同与仿生优化流场则在均匀性与压降之间展现出较好



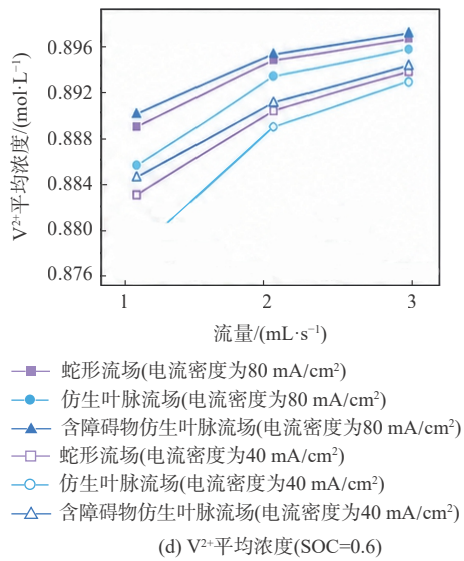
(a) 主通道圆形障碍物仿生叶脉流场



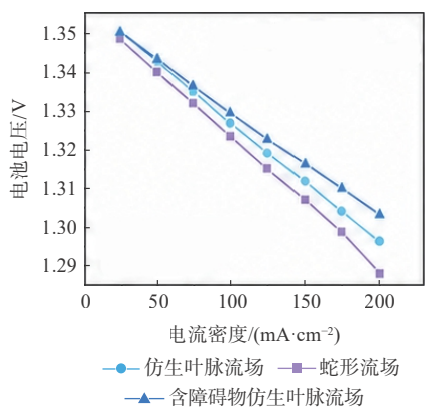
(b) 2 mL/s含障碍物仿生叶脉流场V^{2+}浓度分布



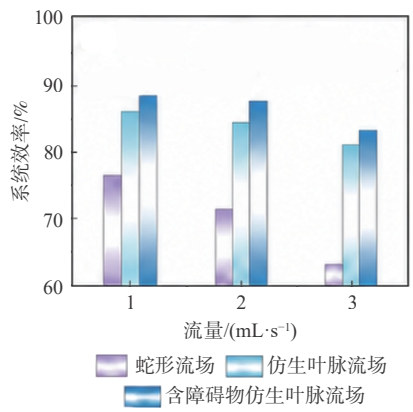
(c) 压降(电流密度为40 mA/cm²)



(d) V^{2+}平均浓度(SOC=0.6)



(e) 极化曲线(SOC=0.6, 流量为3 mL/s)



(f) 系统效率(电流密度为40 mA/cm²)

图5 仿生叶脉流场设计及钒液流电池计算结果[67]

Fig. 5 Design of biomimetic leaf-vein flow fields and computational results for vanadium flow batteries[67]

的平衡潜力，但结构复杂度提升对加工精度和放大一致性提出更高要求。因此，对于工程化电堆，较优方案通常并非传质最强或压降最低的单一极端结构，而是在目标电流密度与允许泵耗约束下，实现浓度均匀性、压降水平与电极利用率综合最优的流场-电极耦合设计。

2.3 隔膜传质

隔膜是全钒液流电池的核心功能组件，其作用是实现正负极电解液中质子的快速选择性传输，同时，抑制钒离子跨膜渗透与串液。理想隔膜应在高质子电导率、低钒离子渗透率、良好化学稳定性、机械强度和较低成本之间取得平衡[77]。对于工程化应用，隔膜不仅影响欧姆阻抗和库仑效率，还会通过跨膜迁移、水迁移和长期降解影响容量保持率与系统寿命。

在钒液流电池充放电过程中，正负极钒离子的透膜迁移会导致电解液交叉污染，不仅降低电池库仑效率与容量，还可能引发正负极短路，是电池性能衰减的重要原因。目前，商用钒液流电池所采用的隔膜仍以全氟磺酸质子交换膜为主[78]，其质子传导性

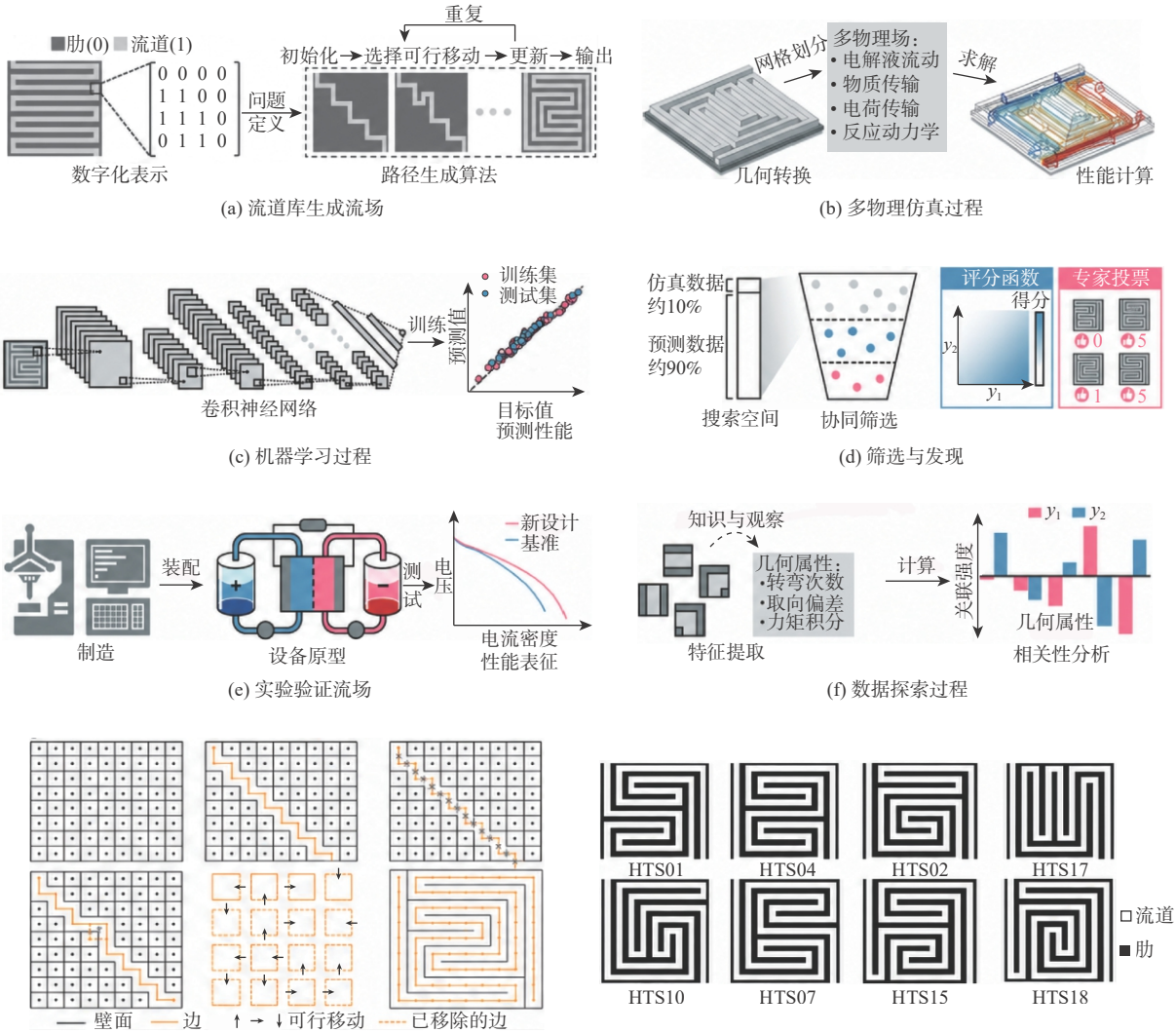


图 6　钒液流电池流场的端到端 AI 驱动设计[63]

Fig. 6　End-to-end AI-driven design of flow fields for vanadium flow batteries[63]

表 2　不同流场类型的传质特性、压降与适用场景对比

Table 2　A comparison of mass transfer characteristics, pressure drop and suitable applications for different flow field types

流场类型	传质增强机制	压降/泵耗特征	适用工况与电极匹配	局限性与优化方向
平行流场[57,74-75]	主要依赖通道内对流、扩散传输和有限肋下渗透对电极补给	压降最低、泵耗小	适合低-中等电流密度、高渗透率或较薄电极；大面积电堆需重点解决各通道间的均流问题	肋下和远端区域易传质不足，可通过波纹结构、导流板或多场协同结构改善
SFF[32-33,65-66]	长流道形成明显沿程压差，可增强局部横向渗透和肋下传质	压降较高，泵功代价明显	适合较高流速、高功率密度或与较厚电极匹配的流经型设计	易出现“热点”与“死区”交替，可通过变截面流道、堵块阵列、旋转或梯度深度设计优化
叉指流场[32,69,71,76]	进出口非直接连通，迫使电解液穿透电极厚度方向，缩短传质路径，并增强对流补给	压降取决于电极厚度、压缩率和渗透率；优化后可低于高阻SFF	适合高电流密度、低流速或薄/中等厚度高渗透电极；需与电极孔结构匹配	中部流量不足时易形成低浓度区，可采用并联交指、分叉交指等方式改善
多场协同流场[33-34,64,68]	通过速度场、浓度场和压力场的协同匹配调节速度-浓度梯度和速度-压力梯度的协同关系，实现低阻强化传质	可在提升浓度均匀性的同时，降低压降	适合对传质和泵耗均敏感的工程化应用场景，尤其在中高电流密度下优势明显	结构设计参数较多，需结合CFD数值模拟和实验验证，应兼顾加工可行性
仿生/双螺旋/异形流场[26,56,67,72]	通过分级分流、曲线流道增强电极覆盖面积和肋下对流传质的均匀性	压降一般控制在中等水平，但受结构复杂度和加工精度影响较大	适合高功率密度、复杂边界条件和大面积电堆的结构筛选	需关注加工成本、密封可靠性、模型泛化能力、大面积放大一致性和长期稳定性

与化学稳定性优异，但存在成本高、钒离子渗透性较高的缺陷；磺化聚醚醚酮膜[79]、磺化聚醚砜膜[80]、磺化聚酰亚胺膜[81]及多孔纳滤膜等新型隔膜虽在成本或离子选择性上具有优势，但在化学稳定性、质子传导性或机械强度等方面仍存在不足，难以满足工程化应用的要求。

隔膜的质子传质过程主要依赖亲水膨胀后形成的质子传输通道，其质子传导机制主要包括 Grotthus 传输[82]、Vehicular 传输[83]与表面传输[84] 3 种：Grotthus 传输是质子通过结构扩散实现的快速传输；Vehicular 传输是质子以 $H_5O_2^+$、$H_9O_4^+$等水合离子为载体，在隔膜通道内的对流扩散传输；表面传输是质子与阴离子结合后，在隔膜亲水通道壁（如含磺酸基树脂的−SO_3−基团）上的界面传输。在含水量较高的体系中，Grotthuss 传输与 Vehicular 传输是质子传导的主要方式。

隔膜传质优化的核心是构建高选择性、高传导性的质子传输通道，在保证钒离子低渗透率的前提下，提升质子传导效率[35,85]。目前的研究主要围绕膜材料的化学结构设计、亲水/疏水微相分离、复合填料、离子筛分层和多孔结构调控展开。通过上述方法实现隔膜质子传导能力与离子选择透过性的双重提升，满足高电流密度下的液流电池传质需求[36,86-87]。高质子电导率与高离子选择性往往存在权衡：过度提高亲水通道连通性可能增加钒离子渗透，而过度增强阻隔能力又可能提高膜阻。而且隔膜在长期氧化和溶胀过程中可能出现质子通道结构变化、钒离子渗透增加或机械强度下降，从而降低库仑效率和容量保持率[27]。因此，隔膜设计应兼顾电压效率、库仑效率、容量保持率和长期化学/机械稳定性。

中国科学院大连化学物理研究所 Zhai 等[88]通过磺化 UiO-66-NH_2 与磺化聚醚醚酮膜复合，构建连续质子通道并作为钒离子屏障，实现质子传导率提升 63%、钒渗透率降低 83%，显著提高钒液流电池性能。Aziz 等[86]通过在 Nafion 掺杂 $TiZrO_4$ 纳米管形成超高选择性通道，质子传导率达 75.9 mS/cm，H^+/VO^{2+}选择性提升 23 倍，电池能量效率达 83.9%、循环耐久性优秀。还可以通过增强尺寸筛选与选择性来制备多孔膜。中国科学院大连化学物理研究所 Zhou 等[89]定制多孔聚苯并咪唑膜，通过高效质子通道和多重阻挡系统，实现高质子传导与强钒离子阻挡，适用于高倍率钒液流电池；并通过 Nafion/聚偏氟乙烯双孔结构膜调控微/宏孔分布，质子传导率为 49.1 mS/cm，钒渗透显著抑制，电池库仑效率为 96.1%、能量效率为 83.8%[37]。还有一些其他策略，如设置层状交联结构，日本国立材料科学研究所 Sha'rani 等[90]通过在 Nafion 表面层组装聚乙烯亚胺与氧化石墨烯，实现质子传导与钒阻挡的优异平衡，提升电池整体性能。

3 运行参数对传质过程的影响

全钒液流电池的传质过程不仅受电极、隔膜、流道等核心组件结构的固有特性影响。也与流量、温度、电流密度、电解液浓度及罐堆比例等运行参数密切相关。与结构参数相比，运行参数具有在线调节性，能够改变电解液物性、反应物补给速率、活性物质消耗速率和系统能耗。可通过调控电解液流动特性、离子扩散速率与电化学反应速率，直接影响电池内部的传质效率与浓差极化程度[76]。从工程化应用视角，传质优化并非单一的结构改良，而是需要根据组件结构特性设计适配的运行参数，实现结构与运行的协同调控，在强化传质的同时，降低系统能耗，提升电池整体性能。本节系统分析各关键运行参数对钒液流电池传质过程的影响规律与调控策略。

3.1 流 量

电解液流量是调控液流电池传质过程的核心运行参数，直接影响电极表面的电解液更新速率、主体液相与反应界面的浓度差及局部浓差极化程度。文献 [29,91] 发现，在膜电极组件、流场结构及工作电流密度保持不变的前提下，钒液流电池通常表现出明显的临界流速特性。当电解液流速低于该临界值时，提高流速能够显著强化电解液在电极区域内的更新与补给，有效缓解电极界面的浓差极化现象，使充电过程电压降低，放电过程电压提升，进而显著改善电池效率。而当流速超过临界值后，电极表面活性物质浓度已趋于稳定饱和，电池电压曲线基本重合，此时，继续提高流速对电池整体性能的改善作用不再明显。Toja 等[25]通过 CFD 模拟和实验揭示了电解液速度均匀性对钒液流电池流场放大的关键作用，通过优化通道几何设计出“增强均匀性流场”，将电极内速度异质性指数从 19.0% 降至 3.2%，使电流密度提升 49%（达 654 mA/cm^2），并在完整电池中使能量效率提升 4.1% 和电解液利用率提升 10.4%。

从系统能效角度看，电解液流量提高通常伴随压降和泵功增加。当流速超过临界值后，继续提升流速会导致循环泵功大幅增加，而电池性能无明显改善，最终造成系统净效率降低。此外，流速还会影响电解液的跨膜渗透与钒离子扩散：低流速可减少钒离子跨膜扩散导致的电解液体积损失，但浓差过电位显著增大；高流速虽能降低浓差过电位，但会加速电解液的透膜渗透。因此，电解液流量的优化需进行多目标调控，纽约州立大学 Khaki 等[92]建立了流速与电池容量/电压损失的显式数学关系，揭示了电解液黏度差是膜渗透的关键驱动因素，为流量的精准调控提供了理论模型。从流场与流量的匹配角度，低流量工况下，电解液主要依靠扩散传

输，电极厚度方向分布不均，流场的传质强化作用有限；高流量工况下，需要流场的导流作用增强对流，实现电解液在电极内的均匀分布，提升电池极限电流密度[74]。澳大利亚西澳大学 Wang 等[93]提出一种非对称变流量控制策略，基于达西定律和 SOC-黏度关联模型动态优化正负极流量比，实验验证表明，该策略可显著抑制钒氧化还原液流电池的电解液迁移（84.48%），提升放电容量（13.77%），并减缓容量衰减（6.42%），有效兼顾了系统效率与长期容量保持率。

3.2　温　度

钒液流电池的运行温度易受季节、地域及实际运行工况的影响，可通过改变离子扩散速率、电解液黏度、电极反应动力学及隔膜传质特性，从多维度影响电池内部的传质过程与整体电化学性能。清华大学 Xi 等[94]研究发现，在−20～50 ℃的宽温度范围内，钒液流电池的库仑效率与电压效率呈现相反的温度变化趋势：库仑效率随温度升高而降低，主要原因是温度升高会加剧钒离子的跨膜扩散，增加电荷损失；电压效率随温度升高而增加，核心是因为温度升高可显著降低电解液的欧姆电阻与电极的极化电阻，提升离子传质与电化学反应效率。

温度对电池内阻的调控具有明显的阶段性特征：低温条件下，极化电阻占内阻主导地位，是制约电池性能的核心因素；高温条件下，欧姆电阻与极化电阻共同控制电池性能[95]。研究发现，钒液流电池极化电阻和欧姆电阻与温度呈负相关，在低温下，极化电阻和活化能显著增加，导致容量利用率下降，高温条件下，交叉污染和极化加剧[96]。基于温度对传质与反应的影响规律，钒液流电池的温度适配运行策略为：低温工况下，采用低电流密度运行，减少极化电阻引发的浓差极化与电压损失；高温工况下，需同步增强电流密度与流场传质的协同，在提升反应速率的同时，通过流场强化电解液对流，抑制高温下的钒离子跨膜扩散[97]。

3.3　电解液浓度

电解液浓度是决定钒液流电池能量密度的核心参数，同时，对传质过程与电化学反应具有显著调控作用。随着钒电解液浓度升高，电极表面的活性离子供给量增加，反应过程中的浓差极化与欧姆损耗降低，电池电压显著提升[98-99]。但电解液浓度的提升也存在固有瓶颈，正负极的传质与反应限制成为主要制约因素，负极的动力学限制源于 V^{2+}/V^{3+} 电对的电子转移速率慢，导致高过电位；正极的传质限制则因 VO_2^+/VO^{2+} 电对在碳基毡电极空隙中的扩散阻力大，尤其在高电流密度与电极出口区域，易出现反应物耗竭，引发严重的浓差极化。

液流电池向高能量密度发展的必然趋势是提升电解液浓度，但浓度升高会导致电解液黏度增加，还有可能在运行过程中析出固相[100]，影响电池容量，是进一步开展研究的方向。不同价态的钒在运行条件下会存在固态析出的条件，影响电池容量。麻省理工学院 Barton 等[39]通过实验独立调控电解液黏度，建立了 Sherwood 数、Peclet 数、Schmidt 数的普适关联式，实现了从材料物性到电池性能的定量预测。该研究系统解耦了黏度对传质、欧姆、动力学过程的独立影响，建立了黏度-传质-性能的定量关系，发现当电解液黏度从 1.5 mPa·s 增至 4.6 mPa·s 时，传质系数下降 30%～40%，为高浓度电解液的传质优化提供了核心理论依据。

3.4　电流密度

电流密度是钒液流电池运行过程中重要的电化学工况参数，其本质上反映单位电极面积上的反应强度和活性物质消耗速率[21,40]。与流量、温度和电解液浓度等主要影响反应物供给能力的参数不同，电流密度更多体现为传质需求端的驱动变量——其升高意味着单位时间内电极反应界面对活性物质的消耗速率增加，若电解液流量、流场分布和电极孔道传输能力不能同步匹配，则电极局部区域容易出现反应物贫化，导致浓差极化增大、放电电压降低、充电电压升高，最终降低电压效率和能量效率[40,98,101]。可见，电流密度的可调范围受限于“传质供给端”的实际能力，两者构成运行优化的基本矛盾，电流密度不宜孤立优化，而应与流量、SOC、电解液浓度、温度及电极/流场结构协同匹配。在较低电流密度下，传质限制通常不明显，过高流量带来的性能增益有限；而在高电流密度下，反应物消耗速率增加，流量不足或局部流动不均会迅速放大传质限制[24,41]。以上表明，液流电池运行优化的关键不是单纯提高或降低电流密度，而是在目标功率密度、浓差极化、泵功损失和系统能量效率之间确定合理的电流密度-流量匹配区间[24]。这本质上是一个“运行窗口寻优”问题。

进一步地，电流密度调控还需要考虑充放电过程中的传质差异。Kim 等[40]通过比较多孔电极不同区域的电位分布，对液流电池运行过程中的局部传质状态进行了表征。研究表明，在充电初始低 SOC 阶段，电极不同区域之间的电位偏差较大，说明局部反应物补给和电极内传质状态尚不均匀。此时，若直接采用较高电流密度，容易放大局部浓差极化，并可能增加析氢、析氧等副反应风险。因

此，在充电初期可适当降低电流密度，待电解液物性和电极内传质过程趋于稳定后，再逐步提高电流密度。相比之下，放电过程中的局部电位分布整体更均匀，但其对电流密度变化更为敏感。在放电过程中，当电流密度加倍时，电极不同区域之间的电压偏差增加161.7%，说明高电流密度会迅速放大反应物消耗速率与传质补给能力之间的不匹配。因此，放电过程应根据SOC、流量和电极/流场结构确定合理负载，使电流密度低于传质允许范围，避免浓差极化加剧和电压效率下降。

4 结构与运行参数协同传质优化

液流电池传质优化受多重因素耦合控制，并非单一结构参数或运行参数所能独立决定。这些因素涵盖电极结构（厚度、孔隙率、渗透率、压缩率）、流场构型、隔膜特性（厚度、选择性），及流量、电流密度、SOC窗口、电解液浓度等运行条件[19,22,75]，如表3所示。其中，电极厚度、孔隙率、渗透率和压缩率决定了电解液进入多孔电极的难易程度、有效反应面积和局部传质路径[25-26,32-33,68]；流场结构决定电解液在电极表面的分配方式、肋下对流强度和局部贫液区域；隔膜厚度与选择性则影响质子传导、钒离子迁移和容量衰减[35-37,87]。与此同时，流量决定反应物补给速率，电流密度决定反应物消耗速率，SOC和浓度差影响主体浓度、浓度梯度和跨膜迁移驱动力[21,24,39-41]。因此，结构参数的最优值并不是固定不变的，而是随运行流量、电流密度、SOC区间和系统压降约束发生变化[33,39,75,95,97]。若仅优化电极或流场结构，而不考虑运行条件匹配，可能出现局部反应物补给不足、浓差极化加剧、压降升高或泵功损失增加等问题[23-24,102]。

从控制变量角度看，结构与运行参数协同传质优化的核心在于建立“结构供给能力-运行传质需求-系统能耗约束”之间的匹配关系[21-22,24,41,75]。对于电极结构，应在有效反应面积、传质路径长度、孔道渗透能力和接触电阻之间取得平衡，并根据目标电流密度和循环流量确定合适的厚度、孔隙率和压缩率[29-30,54,75]；对于流场结构，应结合流量和电流密度评价其反应物分配能力、浓度均匀性、压降和泵功代价[26,32-34]，而不能仅在单一流量条件下比较流场优劣；对于隔膜结构，则需在质子电导率、离子选择性、机械稳定性和SOC运行窗口之间优化[35-37,79,87]。由此可见，液流电池传质优化应从单一组件改进转向电极-流场-隔膜-运行参数的系统协同设计，以实现高电流密度下低浓差极化、低泵功损失、高电解液利用率和长循环稳定运行的综合目标[27,42-43,93]。

表3 结构与运行参数协同传质优化的控制变量分析

Table 3 Analysis of control variables for co-optimisation of mass transfer based on structural and operational parameters

协同对象	控制变量分析结论	工程优化方向
电极厚度-流量、电流密度[29-30,75]	电极厚度增加可提高有效反应面积，但会延长反应物扩散路径。低流量或高电流密度下，过厚电极内部容易出现反应物补给不足和浓差极化；较高流量或强制穿流流场可提高厚电极利用率	根据目标电流密度和供液能力确定电极厚度，在反应面积增加与传质路径延长之间寻求最优平衡
电极孔隙率/渗透率-流量[28-29,75]	高孔隙率和高渗透率有利于电解液进入电极并降低压降，但可能降低单位体积反应面积；低孔隙率可增加反应界面，但在高流量下可能引起较大流阻和局部传质受限	建立孔隙率、渗透率与循环流量的匹配关系，在反应面积、传质能力和泵功损耗之间进行综合优化
电极压缩率-流量、电流密度[30,54]	适度压缩可降低接触电阻并改善电极与双极板接触，但过度压缩会降低孔隙率和渗透率，导致流动阻力增加和局部传质不足，尤其在较低流量或较高电流密度下更易放大浓差极化	选择兼顾低接触电阻和高渗透率的压缩率，避免单纯追求低欧姆阻抗而牺牲传质性能
流场结构-流量、电流密度[26,32-34]	流场结构决定电解液分布和局部传质路径，流量决定反应物补给强度，电流密度决定反应物消耗速率。不同流场在流量-电流密度平面上的优势区间各异，低流量或高电流密度下，例如，叉指等强制对流型流场的肋下传质优势更突出，但泵耗代价增加	根据目标电流密度、允许泵功选择流场结构，在相同流量或相同泵功条件下比较不同流场的适用性，最终以系统净效率作为优选判据
隔膜厚度/选择性-SOC、浓度差、电流密度[35-36,38]	高质子电导率有利于降低欧姆损失，但若钒离子渗透率过高，会加剧交叉污染和容量衰减；高SOC差、浓度差和高电流密度会放大膜传质与迁移问题	优化膜电导率、离子选择性和机械稳定性，并控制SOC窗口和浓度差，兼顾电压效率、库仑效率和容量保持率

5 系统级传质优化及工程约束

5.1 罐堆比例

罐堆比例（储液罐与电堆内电解液的体积比）通过影响电解液不同SOC的活性物质浓度分布，间接影响钒液流电池的传质过程与电池容量。在充放电过程中，储液罐与电堆内的电解液中不同SOC的活性物质浓度会持续变化并产生差异，这种浓度差并非由电压损失或电流损失引发，而是源于储罐中的电解液未及时输入电堆、活性物质未被充分利用，直接造成电池的可用容量损失；同时，储液罐内部的电解液混合损失会进一步降低活性物质利用率与系统效率[103]。

罐堆之间的活性物质浓度差与电流密度密切

相关, 电流密度越高, 浓度差越显著[104]: 当电流密度为 10 mA/cm^2 时, 正极储罐与电堆的四价钒离子浓度几乎无差异; 当电流密度升至 60 mA/cm^2 时, 两者的浓度差明显增大。罐堆比例直接影响电解液利用率, 伊利诺伊大学 Nemani 等的研究[105]表明, 当罐堆体积比为 1 时, 电解液利用率仅为 60%; 当体积比提升至 20 时, 电解液利用率可达 84%。目前, 提升电解液利用率的方法除了优化罐堆比例外, 还可通过储液罐的非混合型设计、基于传质特性的动态电流优化调控等方式, 减少罐堆之间的浓度差与混合损失。

5.2　堆间电解液分布与旁路电流控制

在大规模钒液流储能系统中, 多个电堆通常通过管路并联或串并联运行, 传质问题由单电池内部扩展为“管路-电堆-电解液储罐”的系统级分配问题。如表 4 所示, 多堆并联时, 各支路的管路阻力、堆内压降和温度状态(温度影响电解液黏度)若不一致, 会导致电解液流量分配不均, 使部分电堆处于低流量工况, 加剧浓差极化[44,106-107]。同时, 温度分布不均会改变电解液黏度、扩散系数和反应速率, 进一步放大热-流-电多场耦合效应的空间差异[44,106]。

表 4　系统级传质问题、性能影响及优化策略

Table 4　System-level mass transfer issues, performance implications and optimisation strategies

系统级问题	对传质和性能的影响	优化思路
多堆并联流量分配不均[23-24,102]	局部低流量电堆易出现浓差极化和容量利用率下降; 过量流量电堆则增加泵耗, 造成堆间性能离散	优化主管/支管阻力匹配, 采用对称或等阻管路布局, 并增加分堆流量监测
热-流-电耦合不均[32,75,106]	温度分布不均会引起黏度、扩散系数和反应速率的空间差异, 也可能加速副反应、跨膜迁移和材料老化, 并放大局部传质差异	建立温度-流量-电流密度联动控制, 保证电堆温度和SOC分布均匀
泵耗与电堆效率权衡[23,41,102]	高流量可降低浓差极化, 但超过临界流量后系统净效率下降	以泵基效率或系统净效率为目标, 避免单纯追求单电池电压提升
旁路电流[45-46]	共用歧管和电解液导电通道会形成旁路电流, 造成电流分布不均和副反应加剧	通过歧管电阻设计、流道等阻化、分段集流和优化堆内流量分配缓解
长期沉积、污染与材料老化[27,100]	杂质沉积、膜降解、孔道堵塞或电极表面官能团衰减会降低有效传质面积, 并增加压降	开展电解液净化、材料耐久设计、运行窗口控制和循环后传质复测

旁路电流是压滤式电堆系统中需要关注的另一类耦合问题[45-46]。其主要来源于相邻单电池或电堆之间通过共用歧管或电解液通道形成的离子导电路径, 直接后果是有效工作电流降低、局部电流分布偏离设计值, 并可能诱发局部副反应和传质负荷不均。

值得指出的是, 上述系统级问题的解决通常是以增加泵送损耗为代价。高流量虽然能够改善传质并降低浓差极化, 但会增加压降和泵功, 因此, 实际运行中通常需要以系统净效率或泵基效率为目标, 在传质强化与能耗代价之间进行综合权衡, 确定最优流量区间[41,102]。

6　结　论

本文从工程化应用视角系统综述了压滤式电化学反应器在钒液流电池中的传质优化研究进展, 围绕电极结构、流场结构、隔膜传质、运行参数调控、结构-运行参数协同及系统级工程约束进行了归纳整理。在此基础上, 进一步阐明了不同组件传质特性和系统设计与电池效率、容量利用率和长期稳定性之间的耦合关系, 形成以下主要认识:

1) 电极传质优化需实现宏观结构与微观改性的协同。宏观层面需在孔隙率、厚度、组装压缩率之间建立平衡, 微观层面通过构建微纳米分级多孔结构、表面官能团、活性位点构筑提高润湿性、反应界面可及性和界面反应速率, 从而降低浓差极化, 并提高电极利用率。

2) 流场优化的核心是在电极内浓度分布均匀性、肋下对流强度和系统压降之间的综合权衡。目前, 已从传统流道的局部改进(变径、堵块、并联)向多场协同设计、仿生设计及 AI 驱动的异形流场设计发展, 但工程应用仍需关注压降控制、泵功、加工复杂度和大面积放大效应等实际问题。

3) 隔膜传质优化的关键是构建高选择性、高传导性且稳定的离子传输通道。在保证钒离子低渗透率的前提下提升质子传导效率, 可同时改善库仑效率、电压效率和容量保持率; 但膜材料设计仍需兼顾化学稳定性、机械强度、长期运行条件下的结构稳定性。

4) 运行参数对传质过程具有显著调控作用。流量决定反应物补给速率, 电流密度决定反应物消耗速率, 温度和浓度改变电解液物性与反应动力学, SOC 影响主体浓度和跨膜迁移驱动力。因此, 运行优化应从单一参数调节转向流量-电流密度-SOC-温度-浓度的协同控制, 在降低浓差极化的同时, 控制泵功损失。

5) 工程化传质优化需由单电池尺度拓展至电堆和系统尺度。多堆并联流量分配、热-流-电耦

合、旁路电流、罐堆混合、泵耗与效率权衡及长循环材料衰减均会改变电堆局部传质状态和系统净效率，因此，需要建立面向电堆一致性、系统效率和寿命稳定性的多目标优化方法。

从全钒液流电池相关研究中总结出的相关方法（如流场均匀化、电极孔结构调控、传质-压降协同优化、传质-反应动力学匹配及膜传质路径设计等）可为其他液流电池体系提供反应器设计开发思路。其他液流电池在具体应用时，应结合不同体系的活性物质物化性质、反应动力学、副反应、沉积/络合行为和膜选择性要求进行适配。液流电池用压滤式电化学反应器的传质优化是一项系统工程，未来研究与工程化应用可重点从以下方面展开：

1）开展多组件协同的传质优化研究。突破单一组件优化的局限性，建立电极、流场和隔膜之间的结构匹配设计方法，考虑孔结构、流场分配、界面反应和膜传输之间的耦合效应，构建多组件协同的传质强化体系。

2）强化工程化系统集成考量。传质优化需在提升电池性能的基础上，同步兼顾电堆紧凑性、罐堆匹配性、循环泵功、热管理和旁路电流控制等系统因素，建立“性能-能耗-体积-寿命”的多目标优化模型。

3）推动低成本、高稳定性的传质优化技术研发。针对高端碳基电极、全氟磺酸膜和复杂异形流场成本较高的问题，应开发低成本电极改性方法、非氟或低氟高选择性膜材料及易加工、低压降、适合放大的高效流场结构。

4）深化多物理场耦合的传质理论研究。结合CFD、电化学动力学、分子模拟和数据驱动方法，揭示微纳界面传质、多孔电极内对流-扩散-反应耦合及电堆尺度热-流-电耦合规律，为结构设计和运行控制提供定量依据。

5）建立液流电池传质优化的标准化评价体系。制定涵盖传质效率、浓差极化、能耗损失、成本效益的综合评价框架，为不同优化方案的对比与工程化选型提供依据。

从储能工程的规模化发展视角，压滤式液流电池反应器的传质优化最终需实现组件高性能、系统低能耗、运行长寿命和成本可接受的统一。未来需通过材料研发、结构设计、系统集成与运行调控的协同创新，突破传质控制的核心瓶颈，为液流电池在大规模电化学储能领域的应用提供理论和技术支撑。

参考文献（References）

[1] Trinidad P, Walsh F C, Gilroy D. Conversion expressions for electrochemical reactors which operate under mass transport controlled reaction conditions. Part I: batch reactor, PFR and CSTR[J]. International Journal of Engineering Education, 1998, 14(6): 431-441.

[2] 杨智菲. 全钒液流电池流场结构优化及催化剂原位电沉积研究[D]. 成都: 西南交通大学, 2021.
Yang Z F. Research on flow field structure optimization and in-situ catalyst electrodeposition of all vanadium redox flow battery [D]. Chengdu: Southwest Jiaotong University, 2021(in Chinese).

[3] 刘柏辰. 面向高可用能量容量的全钒液流电池结构设计及运行策略优化研究[D]. 杭州: 浙江大学, 2020.
Liu B C. Architecture design and operation strategy of vanadium redox flow batteries for enhancing the effective energy capacities[D]. Hangzhou: Zhejiang University, 2020(in Chinese).

[4] 诸旻. 有机废水电催化处理工艺方法研究[D]. 苏州: 苏州科技学院, 2022.
Zhu M. Study on electrocatalytic treatment process of organic wastewater[D]. Suzhou: Suzhou University of Science and Technology, 2022(in Chinese).

[5] Wiebe A, Gieshoff T, Möhle S, et al. Electrifying organic synthesis[J]. Angewandte Chemie International Edition, 2018, 57(20): 5594-5619.

[6] Kumar S S, Himabindu V. Hydrogen production by PEM water electrolysis-A review[J]. Materials Science for Energy Technologies, 2019, 2(3): 442-454.

[7] Perry S C, Ponce de León C, Walsh F C. Review: the design, performance and continuing development of electrochemical reactors for clean electrosynthesis[J]. Journal of the Electrochemical Society, 2020, 167(15): 155525.

[8] Zhao P, Zhang H M, Zhou H T, et al. Characteristics and performance of 10kW class all-vanadium redox-flow battery stack[J]. Journal of Power Sources, 2006, 162(2): 1416-1420.

[9] Zhang H. Flow battery technology[M]//Handbook of clean energy systems. Chichester: John Wiley & Sons, 2015: 1-12.

[10] Ding C, Zhang H M, Li X F, et al. Vanadium flow battery for energy storage: prospects and challenges[J]. The Journal of Physical Chemistry Letters, 2013, 4(8): 1281-1294.

[11] Niu Y C, Gao Q T, Zhao R F, et al. A high current density and long cycle life iron-chromium redox flow battery electrolyte[J]. Green Chemical Engineering, 2026, 7(1): 61-69.

[12] Liu Y P, Guo C, Wu G F, et al. Uniformly dispersed bismuth metal nano catalyst modified carbon cloth electrode for iron-chromium flow battery[J]. Nano Research Energy, 2025, 4(1): e9120135.

[13] Xie X, Xiang Y, Daoud W A. MoO_3-deposited graphite felt for high-performance vanadium redox flow batteries[J]. ACS Applied Energy Materials, 2020, 3(11): 10463-10476.

[14] Wu C X, Lu S F, Wang H N, et al. Endowing nonionic membrane with superior ionic selectivity by using g-C_3N_4 for redox flow battery[J]. Journal of Power Sources, 2023, 564: 232857.

[15] Liu Z X, Wu X R, Xiao X, et al. Electrochemical dendrite management via voltage-controlled rearrangement[J]. National Science Review, 2025, 12(4): nwaf013.

[16] Liu Y Y, Zhang J L, Lu S F, et al. Polyoxometalate-based electrolyte materials in redox flow batteries: current trends and emerging opportunities[J]. Materials Reports: Energy, 2022, 2(2): 100094.

[17] Ashok A, Kumar A. A comprehensive review of metal-based redox flow batteries: progress and perspectives[J]. Green Chemistry Letters and Reviews, 2024, 17(1): 2302834.

[18] Rahman M M, Islam M M, Shafiullah M, et al. Advancing grid integration with redox flow batteries: an engineering and economic review[J]. Green Chemistry Letters and Reviews, 2025, 18: 2507333.

[19] Sánchez-Díez E, Ventosa E, Guarnieri M, et al. Redox flow batteries: status and perspective towards sustainable stationary energy storage[J]. Journal of Power Sources, 2021, 481: 228804.

[20] Pletcher D, Green R A, Brown R C D. Flow electrolysis cells for the synthetic organic chemistry laboratory[J]. Chemical Reviews, 2018, 118(9): 4573-4591.

[21] Kim D K, Yoon S J, Lee J, et al. Parametric study and flow rate optimization of all-vanadium redox flow batteries[J]. Applied Energy, 2018, 228: 891-901.

[22] 赵天寿, 蒋浩然, 李文甲. 流体电池的化学工程科学问题[J]. 中国科学基金, 2023, 37(2): 170-177.
Zhao T S, Jiang H R, Li W J. Chemical engineering scientific issue in flow cells[J]. Bulletin of National Natural Science Foundation of China, 2023, 37(2): 170-177(in Chinese).

[23] Tang A, Bao J, Skyllas-Kazacos M. Studies on pressure losses and flow rate optimization in vanadium redox flow battery[J]. Journal of Power Sources, 2014, 248: 154-162.

[24] Ma X K, Zhang H M, Sun C X, et al. An optimal strategy of electrolyte flow rate for vanadium redox flow battery[J]. Journal of Power Sources, 2012, 203: 153-158.

[25] Toja F, Oriani S S, Casalegno A, et al. Elucidating the role of electrolyte velocity homogeneity for scaling up flow fields in vanadium redox flow batteries[J]. Chemical Engineering Journal, 2025, 526: 171201.

[26] Wan S B, Jiang H R, Guo Z X, et al. Machine learning-assisted design of flow fields for redox flow batteries[J]. Energy & Environmental Science, 2022, 15(7): 2874-2888.

[27] Yuan X Z, Song C J, Platt A, et al. A review of all-vanadium redox flow battery durability: degradation mechanisms and mitigation strategies[J]. International Journal of Energy Research, 2019, 43(13): 6599-6638.

[28] Jiang H R, Zhang B W, Sun J, et al. A gradient porous electrode with balanced transport properties and active surface areas for vanadium redox flow batteries[J]. Journal of Power Sources, 2019, 440: 227159.

[29] Ali E, Kwon H, Choi J, et al. A numerical study of electrode thickness and porosity effects in all vanadium redox flow batteries[J]. Journal of Energy Storage, 2020, 28: 101208.

[30] Prasanna M M, Jayanti S. Effect of electrode thickness and compression on the performance of a flow-through mode vanadium redox flow battery cell[J]. Journal of Power Sources, 2025, 631: 236203.

[31] Wang Q Z, Shan X Y, Liu H C, et al. Mass transfer in micro-nano porous electrodes: a crucial role in optimizing vanadium redox flow battery performance[J]. Journal of Colloid and Interface Science, 2026, 705: 139465.

[32] Messaggi M, Canzi P, Mereu R, et al. Analysis of flow field design on vanadium redox flow battery performance: Development of 3D computational fluid dynamic model and experimental validation[J]. Applied Energy, 2018, 228: 1057-1070.

[33] Wang Z M, Su R H, Jiang H R, et al. Understanding and enhancing the under-rib convection for flow-field structured vanadium redox flow batteries[J]. International Journal of Heat and Mass Transfer, 2024, 230: 125789.

[34] Cheng Q, Li M J, Wang R L, et al. Design and optimization of guide flow channel for vanadium redox flow battery based on the multi-field synergy[J]. Journal of Power Sources, 2025, 650: 237526.

[35] Huynh T T K, Yang T, P S N, et al. Construction of high-performance membranes for vanadium redox flow batteries: challenges, development, and perspectives[J]. Nano-Micro Letters, 2025, 17(1): 260.

[36] Zhang D Z, Xin L, Xia Y S, et al. Advanced Nafion hybrid membranes with fast proton transport channels toward high-performance vanadium redox flow battery[J]. Journal of Membrane Science, 2021, 624: 119047.

[37] Wang F R, Zhang Z H, Jiang F J. Dual-porous structured membrane for ion-selection in vanadium flow battery[J]. Journal of Power Sources, 2021, 506: 230234.

[38] Lulay F, Weidlich C, Valtiner M, et al. Membrane degradation in redox flow batteries[J]. Green Chemistry Letters and Reviews, 2023, 16: 2274529.

[39] Barton J L, Milshtein J D, Hinricher J J, et al. Quantifying the impact of viscosity on mass-transfer coefficients in redox flow batteries[J]. Journal of Power Sources, 2018, 399: 133-143.

[40] Kim J, Park H. Mass transfer in flow batteries characterized by comparison of electrical potentials according to regions of porous electrodes[J]. Journal of Energy Storage, 2022, 55: 105556.

[41] Mccloy R, Li Y F, Bao J, et al. Electrolyte flow rate control for vanadium redox flow batteries using the linear parameter varying framework[J]. Journal of Process Control, 2022, 115: 36-47.

[42] Rashitov I, Voropay A, Tsepilov G, et al. Vanadium redox flow battery stack balancing to increase depth of discharge using forced flow attenuation[J]. Batteries, 2023, 9(9): 464.

[43] Zhang B W, Lei Y, Bai B F, et al. A two-dimensional mathematical model for vanadium redox flow battery stacks incorporating nonuniform electrolyte distribution in the flow frame[J]. Applied Thermal Engineering, 2019, 151: 495-505.

[44] Chen F Y, Gao H, Chen H, et al. Evaluation of thermal behaviors for the multi-stack vanadium flow battery module[J]. Journal of Energy Storage, 2020, 27: 101081.

[45] Tang A, McCann J, Bao J, et al. Investigation of the effect of shunt current on battery efficiency and stack temperature in vanadium redox flow battery[J]. Journal of Power Sources, 2013, 242: 349-356.

[46] Wandschneider F T, Röhm S, Fischer P, et al. A multi-stack simulation of shunt currents in vanadium redox flow batteries[J]. Journal of Power Sources, 2014, 261: 64-74.

[47] Yang Y, Wang X Y, Wang Y J, et al. Functional nano-carbon layer decorated carbon felt electrode for vanadium redox flow batteries[J]. Journal of Energy Chemistry, 2025, 106: 735-741.

[48] Liu Y, Meng H M, Wan K, et al. Nanoneedle array structure optimization-induced mass transfer in all-vanadium flow batteries[J]. Electrochimica Acta, 2025, 525: 146089.

[49] 吴雄伟, 刘俊, 谢浩, 等. 全钒液流电池碳电极材料的研究进展[J]. 中国科学 (化学), 2014, 44(8): 1280-1288.
Wu X W, Liu J, Xie H, et al. Latest advance on carbon electrode materials for all vanadium redox flow battery[J]. Scientia Sinica

Chimica, 2014, 44(8): 1280-1288(in Chinese).

[50] Köble K, Ershov A, Duan K J, et al. Insights into the hydrogen evolution reaction in vanadium redox flow batteries: a synchrotron radiation based X-ray imaging study[J]. Journal of Energy Chemistry, 2024, 91: 132-144.

[51] 张文泽, 吴贤文, 王玉娥, 等. 全钒液流电池电极材料的研究进展[J]. 吉首大学学报 (自然科学版), 2016, 37(4): 61-66.
Zhang W Z, Wu X W, Wang Y E, et al. Research progress on the electrode material of vanadium redox flow batteries[J]. Journal of Jishou University (Natural Sciences Edition), 2016, 37(4): 61-66 (in Chinese).

[52] Chen X R, Li L, Jiang Y Q, et al. Manipulating the local electronic structure microenvironment at the MXene interface to achieve efficient anode for vanadium redox flow battery[J]. Journal of Energy Chemistry, 2025, 104: 118-126.

[53] Mu R, Zhang W, Shi Q W, et al. Engineering carbon electrodes for high-efficiency vanadium redox flow battery: multi-faceted approaches to electrocatalytic interfaces[J]. Journal of Energy Storage, 2025, 132: 117923.

[54] Wang Q, Qu Z G, Jiang Z Y, et al. Experimental study on the performance of a vanadium redox flow battery with non-uniformly compressed carbon felt electrode[J]. Applied Energy, 2018, 213: 293-305.

[55] Alphonse P J, Taş M, Elden G. Novel electrode design having gradually increasing porosity in a vanadium redox flow battery[J]. Fuel, 2023, 333: 126198.

[56] Yaji K, Yamasaki S, Tsushima S, et al. Topology optimization for the design of flow fields in a redox flow battery[J]. Structural and Multidisciplinary Optimization, 2018, 57(2): 535-546.

[57] He Q J, Yu J, Guo Z X, et al. Modeling of vanadium redox flow battery and electrode optimization with different flow fields[J]. E-Prime-Advances in Electrical Engineering, Electronics and Energy, 2021, 1: 100001.

[58] Ali A, Ramadesigan V, Monder D S. Modelling and simulation based impact analysis of electrode parameters on the performance of vanadium redox flow batteries[J]. Journal of Power Sources, 2025, 653: 237659.

[59] Martinez Lopez J, Aramendia I, Fernandez-Gamiz U, et al. Computational modeling of a 2D vanadium redox flow battery cell[J]. Jom, 2024, 76(1): 130-140.

[60] Sun J, Jiang H R, Zhang B W, et al. Towards uniform distributions of reactants via the aligned electrode design for vanadium redox flow batteries[J]. Applied Energy, 2020, 259: 114198.

[61] Wang Q Z, Shan X Y, Hou B Y, et al. Construction of kapok-like microstructure electrode for high power density vanadium redox flow batteries[J]. Journal of Energy Storage, 2025, 115: 115992.

[62] Busacca C, Di Blasi O, Briguglio N, et al. Electrochemical performance investigation of electrospun urchin-like V_2O_3-CNF composite nanostructure for vanadium redox flow battery[J]. Electrochimica Acta, 2017, 230: 174-180.

[63] Zheng Q, Shi X Y, Cai Y Z, et al. Artificial intelligence-empowered modeling and management of flow batteries: a mini-review[J]. Future Batteries, 2025, 7: 100107.

[64] Wang F Z, Xiao G Z, Chu F M. Mass transfer enhancement in electrode and battery performance optimization of all-vanadium flow based on channel section reconstruction[J]. Chemical Engineering Journal, 2023, 451: 138619.

[65] Lu M Y, Deng Y M, Yang W W, et al. A novel rotary serpentine flow field with improved electrolyte penetration and species distribution for vanadium redox flow battery[J]. Electrochimica Acta, 2020, 361: 137089.

[66] Karami T, Zangeneh M R. Performance enhancement of vanadium redox flow battery by flow field modification: channel height reduction and channel blockage[J]. Future Batteries, 2025, 7: 100086.

[67] Huang Z B, Liu Y L, Xie X, et al. Design and optimization of a novel flow field structure to improve the comprehensive performance of vanadium redox flow batteries[J]. Journal of Power Sources, 2025, 640: 236736.

[68] Pan L M, Sun J, Qi H H, et al. Along-flow-path gradient flow field enabling uniform distributions of reactants for redox flow batteries[J]. Journal of Power Sources, 2023, 570: 233012.

[69] Ishitobi H, Saito J, Sugawara S, et al. Visualized cell characteristics by a two-dimensional model of vanadium redox flow battery with interdigitated channel and thin active electrode[J]. Electrochimica Acta, 2019, 313: 513-522.

[70] 杨帆. 面向流动传质的全钒液流电池流道与电极改进研究[D]. 长春: 吉林大学, 2023.
Yang F. Study on improvement of flow channel and electrode of all-vanadium flow battery for flow and mass transfer[D]. Changchun: Jilin University, 2023(in Chinese).

[71] Guo Z X, Ren J Y, Sun J, et al. A bifurcate interdigitated flow field with high performance but significantly reduced pumping work for scale-up of redox flow batteries[J]. Journal of Power Sources, 2023, 564: 232757.

[72] Chai Y W, Qu D W, Fan L Y, et al. A double-spiral flow channel of vanadium redox flow batteries for enhancing mass transfer and reducing pressure drop[J]. Journal of Energy Storage, 2024, 78: 110278.

[73] 刘子玉, 姜泽坤, 邱伟, 等. 人工智能在长时液流电池储能中的应用: 性能优化和大模型[J]. 储能科学与技术, 2024, 13(9): 2871-2883.
Liu Z Y, Jiang Z K, Qiu W, et al. Application of artificial intelligence in long-duration redox flow batteries storage systems[J]. Energy Storage Science and Technology, 2024, 13(9): 2871-2883(in Chinese).

[74] Xu Q, Zhao T S, Zhang C. Performance of a vanadium redox flow battery with and without flow fields[J]. Electrochimica Acta, 2014, 142: 61-67.

[75] Muñoz-Perales V, García-Salaberri P Á, Mularczyk A, et al. Investigating the coupled influence of flow fields and porous electrodes on redox flow battery performance[J]. Journal of Power Sources, 2023, 586: 233420.

[76] Yang T F, Zheng L Z, Teng L T, et al. Performance analysis of vanadium redox flow battery with interdigitated flow channel[J]. Thermal Science and Engineering Progress, 2024, 47: 102360.

[77] Swain S, Brahma K, Nayak R, et al. Recent advances in optimized polybenzimidazole-based membranes for vanadium redox flow battery applications[J]. Journal of Energy Storage, 2025, 129: 117372.

[78] Hemanth Kumar K, Gangasalam A, Chen Y S. Membrane technologies for vanadium redox flow and lithium-ion batteries: advances, challenges, and future perspectives: a review[J]. Journal of Energy Storage, 2025, 113: 115560.

[79] Xi J Y, Li Z H, Yu L H, et al. Effect of degree of sulfonation and casting solvent on sulfonated poly(ether ether ketone) membrane for vanadium redox flow battery[J]. Journal of Power Sources, 2015, 285: 195-204.

[80] Zhang Y X, Zheng L Y, Liu B, et al. Sulfonated polysulfone proton exchange membrane influenced by a varied sulfonation degree for vanadium redox flow battery[J]. Journal of Membrane Science, 2019, 584: 173-180.

[81] Zhang Y P, Li J C, Zhang H, et al. Sulfonated polyimide membranes with different non-sulfonated diamines for vanadium redox battery applications[J]. Electrochimica Acta, 2014, 150: 114-122.

[82] Kreuer K D. Proton conductivity: materials and applications[J]. Chemistry of Materials, 1996, 8(3): 610-641.

[83] Kreuer K D, Rabenau A, Weppner W. Vehicle mechanism, a new model for the interpretation of the conductivity of fast proton conductors[J]. Angewandte Chemie International Edition in English, 1982, 21(3): 208-209.

[84] Paul D K, Mccreery R, Karan K. Proton transport property in supported nafion nanothin films by electrochemical impedance spectroscopy[J]. Journal of the Electrochemical Society, 2014, 161(14): F1395-F1402.

[85] Zhai S X, Lu Z R, Ai Y N, et al. Highly selective proton exchange membranes for vanadium redox flow batteries enabled by the incorporation of water-insoluble phosphotungstic acid-metal organic framework nanohybrids[J]. Journal of Membrane Science, 2022, 645: 120214.

[86] Aziz M A, Han D B, Shanmugam S. Ultrahigh proton/vanadium selective and durable nafion/$TiZrO_4$ composite membrane for high-performance all-vanadium redox flow batteries[J]. ACS Sustainable Chemistry & Engineering, 2021, 9(33): 11041-11051.

[87] Yang Y, Wang Q G, Xiong S Z, et al. Research progress on optimized membranes for vanadium redox flow batteries[J]. Inorganic Chemistry Frontiers, 2024, 11(14): 4049-4079.

[88] Zhai S X, Jia X Y, Lu Z R, et al. Highly ion selective composite proton exchange membranes for vanadium redox flow batteries by the incorporation of UiO-66-NH2 threaded with ion conducting polymers[J]. Journal of Membrane Science, 2022, 662: 121003.

[89] Zhou Q, Zhang X H, Zhang Y B, et al. Tailored porous PBI membranes featuring efficient proton conduction pathways and multi-barrier systems for high-rate vanadium flow batteries[J]. Journal of Membrane Science, 2025, 719: 123769.

[90] Sha'rani S S, Nasef M M, Jusoh N W C, et al. A highly-selective layer-by-layer membrane modified with polyethylenimine and graphene oxide for vanadium redox flow battery[J]. Science and Technology of Advanced Materials, 2024, 25: 2300697.

[91] Yang F, Qu D W, Chai Y W, et al. Development of three-dimensional model for the analysis of the mass transport in vanadium redox flow batteries[J]. International Journal of Hydrogen Energy, 2022, 47(64): 27358-27373.

[92] Khaki B, Das P. Voltage loss and capacity fade reduction in vanadium redox battery by electrolyte flow control[J]. Electrochimica Acta, 2022, 405: 139842.

[93] Wang S J, Zhou Y N, Li Y, et al. Asymmetric variable flow-rate control enhances capacity and efficiency in vanadium redox flow batteries[J]. Journal of Energy Storage, 2025, 140: 118954.

[94] Xi J Y, Xiao S B, Yu L H, et al. Broad temperature adaptability of vanadium redox flow battery: part 2: cell research[J]. Electrochimica Acta, 2016, 191: 695-704.

[95] Zhang C, Zhao T S, Xu Q, et al. Effects of operating temperature on the performance of vanadium redox flow batteries[J]. Applied Energy, 2015, 155: 349-353.

[96] Tian W X, Du H, Wang J Z, et al. A review of electrolyte additives in vanadium redox flow batteries[J]. Materials, 2023, 16(13): 4582.

[97] Han L K, Chen H, Cheng X D, et al. Numerical analysis and optimization of flow rate for vanadium flow battery incorporating temperature effect[J]. Batteries, 2023, 9(6): 312.

[98] Messaggi M, Rabissi C, Gambaro C, et al. Investigation of vanadium redox flow batteries performance through locally-resolved polarisation curves and impedance spectroscopy: insight into the effects of electrolyte, flow field geometry and electrode thickness[J]. Journal of Power Sources, 2020, 449: 227588.

[99] Cecchetti M, Casalegno A, Zago M. Local potential measurement through reference electrodes in vanadium redox flow batteries: evaluation of overpotentials and electrolytes imbalance[J]. Journal of Power Sources, 2018, 400: 218-224.

[100] Carvalho W M Jr, Cassayre L, Quaranta D, et al. Stability of highly supersaturated vanadium electrolyte solution and characterization of precipitated phases for vanadium redox flow battery[J]. Journal of Energy Chemistry, 2021, 61: 436-445.

[101] Wang Z Y, Li Y J, Ren J Y, et al. A response time-based method to operando decouple the polarizations in redox flow batteries[J]. Cell Reports Physical Science, 2023, 4(5): 101395.

[102] Guarnieri M, Trovò A, Picano F. Enhancing the efficiency of kW-class vanadium redox flow batteries by flow factor modulation: an experimental method[J]. Applied Energy, 2020, 262: 114532.

[103] 王志文, 叶强. 液流电池系统储液罐中电解液的混合损失及导流策略[J]. 储能科学与技术, 2023, 12(4): 1148-1157.
Wang Z W, Ye Q. Investigation of the mixing loss and guiding strategy of the electrolyte flow in the tanks of a redox flow battery system[J]. Energy Storage Science and Technology, 2023, 12(4): 1148-1157(in Chinese).

[104] Lee H J, Choi N H, Kim H. Analysis of concentration polarization using UV-visible spectrophotometry in a vanadium redox flow battery[J]. Journal of the Electrochemical Society, 2014, 161(9): A1291-A1296.

[105] Nemani V P, Smith K C. Uncovering the role of flow rate in redox-active polymer flow batteries: simulation of reaction distributions with simultaneous mixing in tanks[J]. Electrochimica Acta, 2017, 247: 475-485.

[106] Gundlapalli R, Jayanti S. Comparative study of kilowatt-scale vanadium redox flow battery stacks designed with serpentine flow fields and split manifolds[J]. Batteries, 2021, 7(2): 30.

[107] Trovò A, Marini G, Sutto A, et al. Standby thermal model of a vanadium redox flow battery stack with crossover and shunt-current effects[J]. Applied Energy, 2019, 240: 893-906.

Research progress on mass transfer optimization of filter-press type electrochemical reactors for flow batteries

LIU Jun[1, 2], HAN Zhicheng[2], XIANG Yan[1, *], LU Shanfu[1], WANG Haining[1]

(1. Beijing Key Laboratory of Bionic Energy Materials and Devices, School of Energy and Power Engineering, Beihang University, Beijing 100191, China;
2. China ENFI Engineering Corporation, Beijing 100038, China)

Abstract: Filter-press electrochemical reactors have become the dominant reactor configuration for flow batteries owing to their simple structure, high modularity, and low cell voltage. They also show broad application potential in energy storage batteries, electrochemical water treatment, and electrosynthetic chemical processes. Among various flow battery systems, the all-vanadium redox flow battery represents the most commercially advanced and extensively studied technology. However, inadequate mass transfer continues to limit further performance development since it increases concentration polarization, reduces voltage efficiency, speeds up long-term performance degradation, and eventually raises system costs. This review focuses primarily on all-vanadium redox flow batteries and discusses mass-transfer enhancement from the perspectives of both key reactor components and system operation. Recent advances in the mass-transfer optimization of electrodes, flow fields, membranes, and other core components are systematically summarized. The effects of operating parameters, including flow rate, temperature, current density, electrolyte concentration, and tank-to-stack ratio, on mass-transfer behavior are analyzed. Based on this, coordinated optimization solutions including reactor structure and system operation are suggested, and the coupling relationship between component-level mass-transfer characteristics and total battery performance is elucidated. From an engineering perspective, mass-transfer optimization in all-vanadium redox flow batteries should not be limited to component modification alone. Instead, it should also account for stack scale-up effects, tank-stack matching, operating-parameter regulation, and the balance between electrochemical performance and system energy consumption. This review aims to provide theoretical guidance and technical references for the engineering design and performance improvement of filter-press flow battery reactors.

Keywords: filter-press type electrochemical reactor; vanadium flow battery; mass transfer optimization; structural design; operating parameters

Received: 2026-03-09; **Accepted**: 2026-07-10; **Published Online**: 2026-07-30 17:59
URL: link.cnki.net/urlid/11.2625.V.20260730.1343.001
* **Corresponding author**. E-mail: xiangy@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0363

增程式飞机发展现状与展望

冯诗愚[1,*]，付子祺[1]，刘冠男[2]，台环宇[1]，耿雷铭[3]，缪日成[1]

(1. 南京航空航天大学 航空学院，南京 210016；2. 西安交通大学 化学工程与技术学院，西安 710049；
3. 中航通飞华南飞机工业有限公司，珠海 519000)

摘 要：全电飞机受限于锂电池能量密度难以满足长航程需求，而增程式飞机通过“电主油辅”架构平衡了环保性与续航里程，成为航空低碳化的重要过渡方案。基于此，聚焦增程式技术路径，系统性对比分析5种典型混合动力架构，明确了增程式定义与工作原理；梳理了国内外代表性增程式机型的技术路线；深入探讨了支撑其发展的高功率密度永磁电机、智能能量管理系统、高效热管理系统及高压局部放电监测与防护技术的进展与挑战。明晰了增程式技术的当前格局与核心壁垒，为全电飞机长航程适应性技术迭代提供了重要依据，并可为航空产业低碳转型战略制定提供参考。

关 键 词：增程式飞机；混合动力飞机；全电飞机；永磁电机；能量管理系统；热管理；局部放电
中图分类号：V272
文献标志码：A **文章编号**：1001-5965(2026)10-3471-11

低碳转型已成为全球工业可持续发展的必由之路。尽管航空领域CO_2排放只占全球人为排放总量的2%～3%，但因其废气组分复杂并直接排放至高空，造成大气环境严重污染，加之排放量持续攀升，故被列为全球减排重点领域之一。

全电飞机完全依赖电能驱动推进系统，具有零排放、高效率、低噪声等显著优点，其代替传统燃油飞机已成为航空业低碳转型布局下的必然趋势。目前，全电飞机商用电池以锂电为主，最高能量密度目前可达到700 Wh/kg，远低于燃油，故其航程受到巨大限制[1]。为解决全电飞机的续航难题，增程式飞机采用“电主油辅”混合动力架构，通过燃油燃烧驱动涡轮带动发电机发电，增加全电飞机航程，以适应区域物流等更多场景。增程式飞机具有低碳环保和长航程特点，可以兼容现有机场燃料基础设施，通过能量管理优化实现系统能量效率最大化，在高能电池技术突破前提供技术过渡。

本文深度解析增程式飞机发展现状，对比5种混合动力飞机动力架构，提出增程式飞机的狭义定义，阐述其工作原理，并整理国内外增程式飞机机型；深入总结高功率密度电机、能量管理系统和热管理系统等核心技术，并针对电气系统安全问题进行讨论；对增程式飞机的未来发展做出展望。

1 混合动力技术与增程式飞机

1.1 混合动力技术

全电推进系统中，电动机与电力电子变换器响应速度快、控制精度高，使其整体效率优于传统燃油动力系统。然而，多数全电飞机，如德国Lilium公司[2]的Lilium Jet和美国Bye Aerospace公司[3]的eFlyer 2等续航仅约300 km，远低于燃油飞机。为突破续航瓶颈并提升效率，国内外学者结合全电与传统燃油动力特点，提出混合动力架构，并开展了一系列研究。

收稿日期：2025-06-14；**录用日期**：2025-07-17；**网络出版时间**：2025-09-23 13:04
网络出版地址：link.cnki.net/urlid/11.2625.V.20250923.1120.001
基金项目：南京航空航天大学2024年研究生科研与实践创新计划(xcxjh20240112)
*通信作者. E-mail：shiyuf@nuaa.edu.cn

引用格式：冯诗愚，付子祺，刘冠男，等. 增程式飞机发展现状与展望[J]. 北京航空航天大学学报，2026，52(10)：3471-3481.
Feng S Y，Fu Z Q，Liu G N，et al. Developing status and prospect of range-extended aircraft[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3471-3481 (in Chinese).

基于能量传递路径的拓扑差异，目前，混合动力飞机动力架构主要有5种，如图1所示[4]。图1(a)为串联动力架构，采用解耦式能量传递路径：内燃机驱动发电机产生电能，经电力电子转换器调配后，直接驱动推进电机或为储能电池充电。其核心特征在于推进系统仅由电动机单一路径驱动，实现动力源与推进器的隔离[5]。架构变体包含氢燃料电池方案，以储氢系统替代内燃机作为增程发电单元，达到零排放要求，但其需应对低温储氢环境与质子交换膜退化等技术挑战[6]。

图1(b)为并联动力架构，通过机械耦合机构实现电动机与内燃机的功率叠加输出，两者既可独立驱动推进器，也可协同工作[7]。图1(c)为混联动力架构，融合串、并联架构特性，可在飞行期间更改拓扑结构以变换动力架构。内燃机输出动力可经机械路径直接驱动推进器，亦可分流至发电机转化为电力[8]。图1(d)为涡电动力架构，涡轮发动机直接驱动发电机为电动机供电，无需电池组，但因缺乏电池缓冲，需增设飞轮等惯性装置来平滑功率波动[9]。图1(e)为部分涡电动力架构，在保留电驱动推进路径的基础上，增设涡轮发动机机械推进形成混合推力网络。涡轮机组同时承担机械推进与发电双重功能，通过功率分流传动装置实现机械能与电能的比例调节[10]。表1列出了各混合动力架构的优缺点。

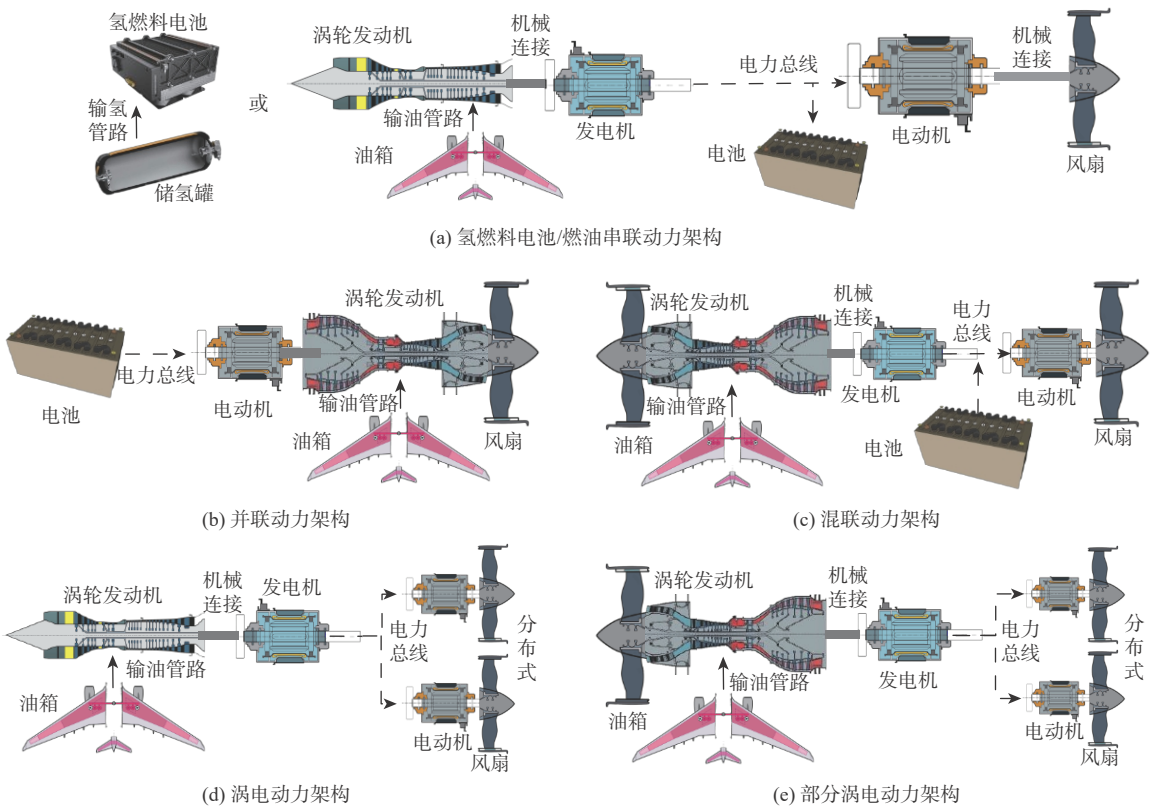


图1 混合动力飞机动力架构示意图[4]

Fig. 1 Schematic diagram of the power architecture of a hybrid power aircraft[4]

表1 混合动力架构优缺点总结

Table 1 Summary of the advantages and disadvantages of hybrid power architectures

动力架构类型	优点	缺点
串联动力架构	内燃机高效；氢燃料支持；控制简单	能量损耗大；设备增重
并联动力架构	快速供能；独立工作可靠	控制逻辑复杂；策略设计难
混联动力架构	适应任务；动态调配能量	质量大幅上升；切换难控；优化复杂
涡电动力架构	结构简单轻；直接转换效率高	需缓冲装置；排放不减；低负载效率低
部分涡电动力架构	优化能量分配	系统复杂；控制难；排放不减；推力不均衡

串联动力架构在动力传输方式上与其他类型有着根本区别：其内燃机或涡轮引擎仅作为发电机使用，产生的电能通过电动机单独驱动推进器，彻底切断了传统机械传动结构。这种设计完全通过电力线路传输动力，发动机无需随飞行状态变化调整转速，始终保持在最高效的工作区间。相比之

下，并联和混联动力架构仍保留发动机直接驱动螺旋桨的机械结构，需要协调机械与电力 2 种不同特性的能量传输。而涡电动力架构虽然也采用纯电驱动，但完全依赖实时发电且不带储能电池，无法应对突发功率需求。串联动力架构则配备储能电池作为电力缓冲，既解决了瞬时动力波动问题，又提供了紧急情况下的备用能源。

1.2　增程式飞机的概念及工作原理

航空领域并没有对增程式飞机的明确定义，而中国国标[11]明确规范了增程式汽车的定义，即纯电模式具备全动力性能，亏电时启动辅助供电装置补能，且该装置与驱动系统无机械连接。故本文所讨论的增程式飞机定义为：在纯电模式下具备全动力性能，当机载储能系统无法满足航程要求时，启动机载辅助供电装置为动力系统提供电能，以延长航程的一种电动飞机，且该机载辅助供电装置与动力系统无机械连接。

以 E-Fan X[12]动力架构为例，如图 2 所示，增程式飞机的动力系统由燃料发电系统、储能系统、能量管理系统、电驱动系统及热管理系统 5 大核心构成[10,13]。其中，燃料发电系统通过化学能-电能的持续转换实现空中自主供电，储能系统依托高能量密度电池组保障全工况功率稳定输出，能量管理系统基于飞行阶段特征动态优化供能策略，电驱动系统将电能转化为推进器机械动力，热管理系统则通过分级散热机制维持设备热安全边界[14-16]。

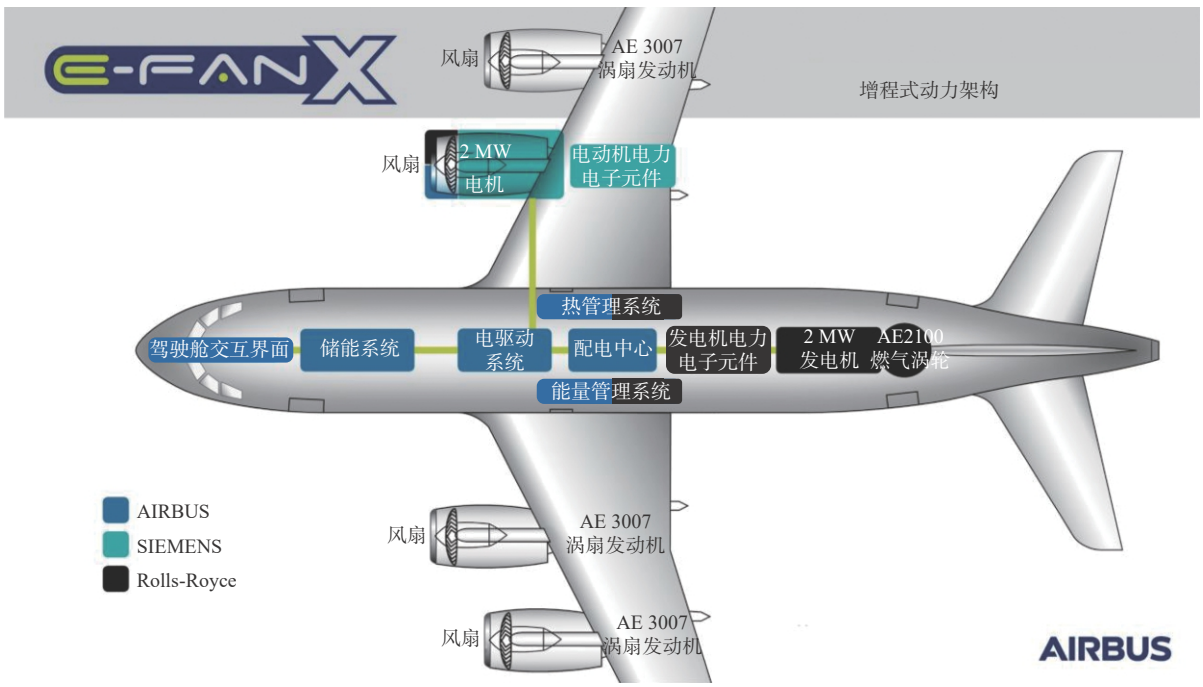


图 2　E-Fan X 动力架构示意图[12]

Fig. 2　Schematic diagram of the power architecture of the E-Fan X[12]

增程式飞机动力系统具备全电与混合动力双模式运行能力：全电模式下，储能系统独立承担供能任务，实现零排放运行，在高功率需求阶段提供瞬时动力支撑，并通过能量回收提升系统效率[17]；混合动力模式下，燃料发电系统与储能系统功率互补，通过航段自适应算法动态调配能量流以扩展航程。2 种模式依托能量管理系统实现无缝切换，使飞机在环保短途飞行与混合动力长途任务间灵活转换，平衡环境友好性与任务适应性[18]。

2　国内外增程式飞机研究现状

基于增程式架构的能量解耦特性与工作原理，全球航空界已展开多路径技术验证。本节聚焦国内外代表性机型，系统梳理如 Zunum Aero ZA10、锐翔 RX1E-A 等项目的动力系统配置与试飞进展，国内外增程式飞机机型如图 3 所示，其关键特性对比如表 2 所示[12,19-26]。通过横向对比揭示不同技术路线的设计特点与适航挑战，为后续关键技术迭代提供依据。

E-Fan X 验证机基于 BAE146 支线客机平台，将一台涡扇发动机替换为西门子研制的 2 MW 电动机，由罗罗 2.5 MW 发电机为主，辅以机腹锂电池协同供电，其渲染图如图 3(a) 所示[27]。2020 年 3 月完成气动风洞试验，验证了混合动力系统气动兼容性[12]。

Zunum Aero ZA10 的渲染图如图 3(b) 所示，设计航程达 1 126 km。动力系统采用燃气涡轮发动机作为增程器，与高比能电池组联合供电，驱动 2 台 500 kW 电动涵道风扇。机型最高巡航速度为



(a) E-Fan X (b) Zunum Aero ZA10 (c) Electra Aero EL-9

(d) Nexus 6HX (e) 锐翔RX1E-A (f) DA36 E-Star

(g) eDA40 (h) Panthera hybrid (i) e-Genius

图 3 国内外增程式飞机机型[12,19-26]

Fig. 3 Domestic and foreign extended-range aircraft[12,19-26]

表 2 国内外典型增程式飞机关键特性对比[12,19-26]

Table 2 Comparison of key characteristics of typical extended-range aircraft at home and abroad[12,19-26]

机型	类型	座位数	关键动力参数	航程性能	项目状态
E-Fan X	支线验证机	100	发电机: 2.5 MW, 电机: 2 MW		项目终止(2020年)
Zunum Aero ZA10	支线客机	12	增程器: 500 kW, 电机: 2×500 kW	增程后: 1 126 km	地面试飞(2019年)
Electra Aero EL-9	eSTOL	9	电机: 8×150 kW	纯电: 611 km, 增程后: 2 037 km	已获订单, 计划首飞(2027年)
Nexus 6HX	eVTOL	5	增程器: 600 kW, 电机: 6×100 kW	增程后: 240 km	项目终止(2020年)
锐翔 RX1E-A	轻型飞机	2	增程器: 30 kW, 电机: 30 kW	增程后: 300 km+	已交付(2025年)
DA36 E-Star	滑翔机改装	2	发电机: 30 kW, 电机: 70 kW	增程后: 900 km(第二代)	第二代完成首飞 (2013年)
eDA40	轻型飞机	4	发电机: 110 kW, 电机: 2×80 kW	增程后: 1 333 km	验证飞行 (2018年)
Panthera hybrid	轻型飞机	4	发电机: 100 kW, 电机: 145 kW	增程后: 1 200 km	已试飞验证 (2021年)
e-Genius	验证机	2	电机: 80 kW	纯电: 400 km, 增程后: 1 280 km	已完成载人飞行测试 (2021年)

547 km/h, 并具备短距起降能力[19]。

Electra Aero 公司基于 EL-2 Goldfinch 验证机开发了 EL-9 混合动力电动短距起降飞行器(electric short takeoff and landing, eSTOL)飞机, 其渲染图如图 3(c) 所示。该机集成 8 台分布式电驱风扇, 燃气涡轮可在飞行中为锂电池充电, 同时, 电驱风扇利用吹气增升技术实现 46 m 内起降。EL-9 已通过 IFR 适航认证并配备智能飞控系统, 已获全球 52 家运营商 2 100 架订单, 计划 2027 年首飞、2029 年商业运营[20]。

Bell 公司 Nexus 6HX 混合动力电动垂直起降飞行器(electric vertical takeoff and landing, eVTOL)飞机采用涡轮增程式架构, 600 kW 涡轮轴发动机驱动发电机, 供电给6 组可倾转涵道风扇及储能电池, 其渲染图如图 3(d)所示。验证显示, 2 722 kg 最大起飞质量下, 巡航速度达 241 km/h、航程为 240 km, 续航能力为纯电平台的 6 倍。已通过 Uber Elevate 完成适航验证[21]。

锐翔 RX1E 系列作为中国首款自主知识产权电动轻型运动飞机, 由辽宁通用航空研究院研制, 其实物图如图 3(e) 所示。升级版 RX1E-A 沿用基础型设计, 新增模块化增程器(可选燃气涡轮/内燃机), 低电量时可空中充电或直供电动机。该增程系统使航程提升至基础型的 2.3 倍, 满足 300 km 短途通勤与扩展训练需求[22]。

DA36 E-Star 基于 HK36 滑翔机构建, 集成 30 kW 转子发电机、70 kW 永磁电机及锂电池组, 依托动态算法优化能量管理, 其实物图如图 3(f) 所示。起飞阶段电池峰值输出达 110 kW, 巡航时发动机稳定在 2 400 r/min 高效发电, 滑翔阶段回收动能[23]。

eDA40 混合动力采用 110 kW 燃油发电机, 两侧集成电驱单元及 12 kWh 电池, 其实物图如图 3(g)

所示。纯电驱动下可飞行 30 min, 增程模式下可延长至 5 h。起飞阶段双电机联合输出 150 kW 峰值推力, 下降反桨回收能量, 实现 1 150 kg 最大起飞质量(载荷为 344 kg)下 263 km/h 巡航与 1 333 km 航程, 燃油效率较传统机型提升 40%[24]。

Pipistrel 公司基于 Panthera 纯电四座飞机加载增程子系统, 开发了 145 kW 级油电混合动力系统, 其实物图如图 3(h) 所示。2021 年验证显示, 巡航阶段燃油机组保持低于 230 g/kWh 燃油消耗率运行并动态补能, 航程较纯电提升 3 倍, 起降噪声低于 65 dB (经过 A 计权网络滤波后测得的分贝值)[25]。

斯图加特大学的 e-Genius 混合动力验证机 2021 年完成载人飞行测试, 其实物图如图 3(i) 所示。该机型以 80 kW 永磁电机配合螺旋桨为主推进单元, 在纯电驱动下航程可达 400 km, 集成微型燃油发电机组后, 可实现 40 kWh 锂电池与燃油机组互补供能。燃油机组锁定 2 400～2 600 r/min 高效区间运行, 通过 dSPACE 控制器实现 2.9 L/100 km 等效油耗(较传统系统降 67%)[26]。在 750 kg 级电动滑翔机牵引验证中, 系统航程较纯电提升 320%。智能飞控通过航迹优化降低 15%～22% 巡航能耗, 其“按需补能”逻辑已被纳入 EASA CS-23 修订草案, 为通航低碳转型提供工程化路径[28]。

3　增程式飞机关键部件及技术

3.1　高功率密度的发电机和电动机

相较于工业电机比功率的 0.1～0.5 kW/kg、汽车驱动电机的 1～3 kW/kg, 航空电机需突破 10 kW/kg 且功率达 1～3 MW[29]。此外, 航空电机还需应对高温、高频振动等极端工况。目前, 主流电机按电磁拓扑结构可分为永磁同步电机、绕线同步电机、感应电机、开关磁阻电机和高温超导电机[30]。永磁同步电机通过永磁体自励磁场, 兼具高功率密度、低维护成本等优势; 绕线同步电机通过主动励磁调节实现恒速恒频输出, 但功率密度较低[31-33]; 高温超导电机基于超导材料突破兆瓦级功率密度, 有望成为航空电机首选, 但仍受限于材料与工艺瓶颈[34]。感应电机和开关磁阻电机结构简单、环境适应性强, 适用于中小型增程器、eVTOL 启动器, 但瞬时扭矩不足且效率较低[35-40]。如表 3 所示[41], 不同电机在功率密度、鲁棒性及成本上各有侧重, 应用领域亦不同。

表 3　电机各项性能对比[41]

Table 3　Comparison of various performance of the motor[41]

电机	功率密度	效率	功率因数	扭矩脉动	鲁棒性	高速性能	热管理	可操控性	经济性	应用场景
高温超导电机	+	++	+	0	--	-	--	+	--	推进器、发电机
永磁同步电机	+	++	+	0	0	0	0	+	0	推进器、发电机
绕线同步电机	0	-	+	0	-	-	-	++	-	推进器、发电机、启动器
开关磁阻电机	0	+	-	-	++	++	+	+	++	发电机、启动器
感应电机	-	-	-	0	+	+	0	+	++	启动器

注: “+” 表示优势; “0” 表示中等; “-” 表示劣势。

为满足增程式飞机对超高比功率的要求, 国内外多数机型选用永磁同步电机, 既作为增程发电机, 也作为主动力电机。在永磁同步电机的技术突破上, Mecrow 等[42] 改进了四相容错电机, 采用 Halbach 阵列和燃油浸没冷却, 实现 15 000 r/(min·16 kW) 高功率密度驱动, 单相故障下仍可运行, 且温升低于 25 ℃。Zhang 等[43] 提出新型 Halbach 五相电机, 优化磁化角与齿宽增强三次谐波以提升转矩密度及效率。Elloumi 等[44] 设计新型辐条式内置电机, 结合机械弱磁装置, 通过模型与有限元验证可降低铜损并扩展 2 倍速域。

在轻量化材料应用上, Koch 等[45] 提出碳纤维-不锈钢混合轴(干法缠绕+离心铸造)与软磁复合材料(soft magnetic composite, SMC)注塑磁芯 2 种减重方案, 通过了 127 000 件/年的量产验证, 且可减重 50%。Zhou 等[46] 设计了一种碳纤维增强塑料高速永磁电机护套, 经板状与圆柱试样验证其环向强度满足设计指标, 并减轻了整体质量。

在永磁电机的应用上, 空客与劳斯莱斯联合研制的 E-Fan X 项目采用 2.5 MW 永磁电机技术, 开创兆瓦级航空电驱工程化先例[47]; 欧洲航空防务与航天公司 (European Aeronautic Defence and Space Company, EADS) 与西门子合作开发的 5 000 r/(min·65 kW) 电机应用于 Diamond 滑翔机, 实现油电无缝切换及 22% 巡航效率提升[48]; 美国国家航空航天局 (National Aeronautics and Space Administration, NASA) 研制的 60 000 r/(min·3 kW) 超高速电机采用陶瓷轴承与无铁芯绕组, 利用飞轮储能, 能量转换效率达到 97%, 并刷新航空电机转速记录[49]。斯洛文尼亚 Emrax 348 系列外转子轴向磁通电机如图 4

所示，在额定功率为210 kW、额定转速为4 000 r/min条件下实现10 kW/kg功率密度，已应用于Apis EA2滑翔机、Electric Taurus等机型，并适配车辆牵引、增程发电等多场景[50-51]。

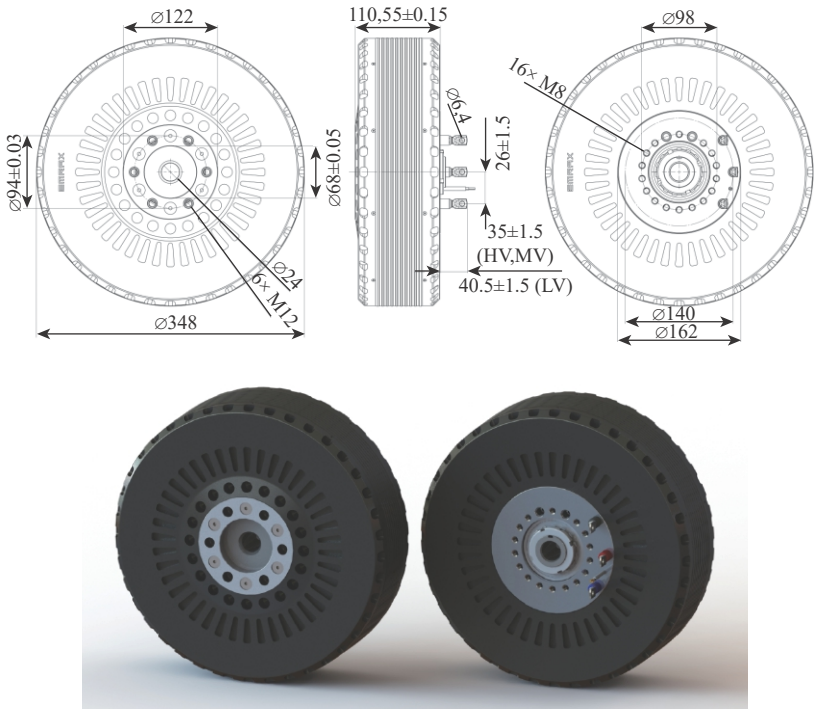


图4 Emrax 348电机尺寸及实物图(单位：mm)[51]

Fig. 4 Emrax 348 motor size and physical drawing (Unit: mm)[51]

永磁同步电机因其高功率密度与高效率，被广泛应用于增程式飞机动力系统。但目前仍面临永磁体在高温环境下易退磁及高速运转时对转子结构和动平衡的要求苛刻等问题。未来研发需结合碳纤维壳体等轻量化材料与紧凑设计，同时，克服高温、高频振动等极端工况挑战[52-54]。

3.2 能量管理系统

能量管理系统的动态协调能力直接关联航程经济性。增程式飞机电力系统架构如图5所示，硬件部分主要由整流器、逆变器和双向直流转换器组成，其拓扑架构决定能量损耗与动态响应特性。三相有源整流器可将机载发电机输出的交流电转换为直流电，其核心功能包括稳定母线电压、优化发电机侧功率因数、确保电流低谐波[55]。三相逆变器控制推进电机，多采用六开关逆变器，在高压化趋势下，多电平逆变器因高耐压、低谐波等优势逐渐成为增程式飞机的主流选择[56]。双向直流转换器负责协调电池、负载与电容间的功率，多采用双有源桥结构。传统系统(28 V/270 V)转向高压(540 V/1 kV)仍需较长时间，过渡过程中可借助多端口转换器适配多电压需求[57]。

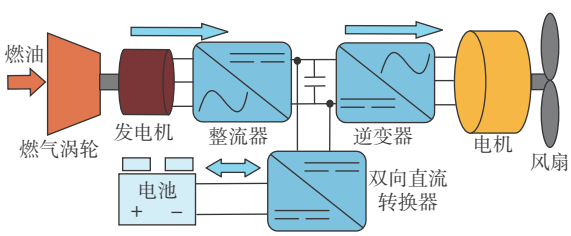


图5 增程式飞机电力系统架构

Fig. 5 Extended-range aircraft electric power system architecture

增程式飞机能量管理策略可分为非因果与因果2类：前者旨在全局优化，后者聚焦实时局部优化。非因果策略需要详略的全局信息，实施难度较大，但仍是理论研究的核心方向[58]。动态规划是一种常用的非因果策略，因直接求解计算量过大，常通过凸化处理来简化计算流程，例如，Murgovski团队[59]通过电池能量变量替换、约束松弛等方法将混合整数问题转化为凸规划，结合启发式决策实现电池尺寸与能量管理联合优化，其解与动态规划结果高度吻合。

因果策略包含规则控制与优化控制2类，在规则控制方法中：基于确定性规则的方法，如Hung团队[60]提出的理想操作线(ideal operating line, IOL)控制，虽能降低无人机燃料消耗，但存在鲁棒性差与最优性缺失问题；基于模糊规则的控制方法则可通过近似推理提升抗干扰能力，如Lee团队[61]在混合动力客车中验证其转矩分配优势，但缺乏理论最优性证明。而Karunarathne等[62]通过融合神经网络

提升模糊规则下的自适应能力。

基于优化的因果策略因理论支撑优势备受关注, 典型理论包括庞特里亚金极大值原理与等效消耗最小化策略。Elbert 团队[63]基于庞特里亚金极大值原理建立功率阈值驱动的引擎开关策略, 在计算效率上显著优于动态规划。Njoya Motapon 团队[64]在 14 kW 燃料电池飞机实验中验证等效消耗最小化策略的燃油经济性最优; Donateo 等[65]将等效消耗最小化策略应用于转子发动机无人机的能量管理, 以较低计算成本优化任务燃油消耗, 并确保电池能量高效利用。Zhai 等[66]提出了一种融合模糊逻辑的自适应能量管理策略 (fuzzy adaptive equivalent consumption minimization strategy, F-AECMS), 结合预测推力调节与电池荷电状态 (state of charge, SOC) 的闭环控制, 能够在多种飞行工况下稳定 SOC 并降低燃油消耗。Palaia 等[67]面向早期飞机设计阶段, 提出了基于任务划分的能量管理方法, 将飞行任务简化为爬升、巡航与下降 3 段, 并通过分段优化动态分配发动机与电机功率输出, 有效提升燃油效率与任务执行性能。

执行所有信息已知的特定任务时, 优先采用非因果控制, 而当飞行任务随飞行环境变化时, 则需要采用因果控制实时更新控制参数。对于非因果控制, 由于其方程的非线性, 未来研究应聚焦于凸化方法来简化计算, 并尽可能减小误差; 对于因果控制, 则要解决强烈随机外部扰动使实时优化难以求解的问题。

3.3　热管理

增程式飞机的动力系统运行中会产生大量热量, 高效可靠的热管理系统对保障性能、延长部件寿命及防止热安全事故至关重要[68]。其设计需综合考虑材料、流体热交换机制与整机热管理, 满足航空领域严苛的轻量化、高可靠性和环境适应性要求。

材料方面, 高性能导热与储热材料是基础。优化后的不同结构热管(管式、平板式、脉动式)结合相变材料, 能将电池组最高温度控制在 50 ℃ 内, 优于传统冷却[69]。Zhao 等[70]通过电极优化, 使 5 C 放电下电池内部温度梯度降低 68%, 减缓容量衰减。Coutinho 等[71]证实膨胀石墨-石蜡复合相变材料可使电池峰值温度降低 9.4 ℃。

流体热交换方面, 风冷结构简单轻便, 但换热能力有限, 适用于低功率或辅助散热, 优化风道与气流组织是关键。Hong 等[72]设计的带二次通风口平行空冷系统将电池组平行流道间温差降低 60%, 压差波动减小 67%。凭借高比热容和导热性, 液冷成为主流方案, 尤适于电池包和电机。Liu 等[73]实验显示, 沉浸冷却下电机运行近 3 000 s 后绕组温度稳定在 130 ℃ 以下, 显著优于传统油套冷却的 180 ℃。

针对整机热管理, Potamiti 等[74]对比发现, 集中式液冷系统整体湿重为 158.22 kg, 分散式液冷系统因含多个独立回路, 质量增加 32.2%, 但具有模块化和冗余优势。整体而言, 集中式液冷系统结构更紧凑轻量, 适合对质量敏感的航空应用; 分散式液冷系统则在冗余安全性和局部温控上更具优势。冷却方式上, 串联易致下游过热, 而并联需更复杂的流量控制, 但温度分布更均衡[75]。

增程式飞机热管理技术未来将采用水基 Fe_2O_3 纳米流体强化导热、蒙皮换热器、增材微通道等新兴散热方案, 通过多学科优化实现热管理与气动设计协同, 以驱动系统综合能效提升。

3.4　高压电子元器件的局部放电问题

在增程式飞机中, 除热管理系统需应对的热失控风险外, 高压电子元器件在高空低气压与高温环境下引发的局部放电问题同样构成严峻的电气安全威胁与潜在失效风险。局部放电作为关键的老化与故障前兆, 其演化机制与识别手段成为研究焦点。局部放电通常分为气隙放电、沿面放电与电晕放电, 其中, 气隙放电对气压变化最敏感, 沿面放电受湿污耦合作用主导, 电晕放电则常因振动调制而增强辐射干扰[76]。图 6 为某 35 kV 交联聚乙烯 (cross-linked polyethylene, XLPE) 电缆接头因局部放电引起的绝缘层击穿与屏蔽层腐蚀现象由外而内地显示了局部放电对电缆造成的损伤[77]。

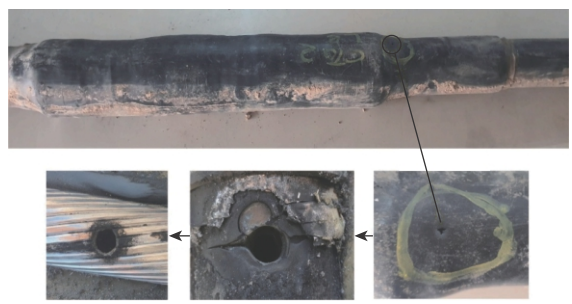


图 6　XLPE 电缆 35 kV 条件下失效实例[77]

Fig. 6　XLPE cable failure example under 35 kV condition[77]

Hu 等[78]通过油纸界面模型研究发现, 低压条件下沿面放电起始电压显著降低, 放电更易沿弱界面扩展。Katsuta 等[79]在超高压电缆试验中指出, 接头区域是局放高发部位, 且在稀薄空气中信号易受干扰。Bartnikas[80]总结了电晕放电的相位分布与频率特性, 为相位分辨局部放电 (phase-resolved partial discharge, PRPD) 分类提供了基础参考。

在新材料防治方面, Abdel-Gawad 等[81]采用 TiO_2 纳米颗粒改性聚氯乙烯 (polyvinyl chloride,

PVC), 发现质量分数为3%的纳米填料可使放电起始电压提升22%, 同时, 脉冲数量显著下降, 表明电子俘获机制有效增强了局放抑制能力。Yan等[82]通过大气等离子体处理SiO_2颗粒, 制备出具有更高耐局部放电能力和击穿强度的环氧树脂复合材料, 即使局部放电密度略有上升, 材料整体可靠性仍显著优于纯环氧树脂。

在检测与预警方面, Wester提出衰减交流法用于电缆绝缘评估, 适用于中长距离系统的非破坏性检测[83]。Gulski等[84-85]进一步开发了基于振荡波测试的在线检测平台, 结合高频电流互感器/高频电流传感器(high-frequency current Transformer, HFCT)、超高频(ultra-high frequency, UHF)等传感器布设方案, 并引入深度学习算法以提升多类局放的识别准确性。

综上所述, 局部放电的波形、幅值、频率与相位等参数可有效反映放电位置与类型, PRPD图谱仍是核心诊断工具。未来发展应聚焦于纳米复合绝缘材料的界面强化技术、多物理场耦合的老化模型与寿命预测方法及多源信号(电-声-光)融合的智能在线诊断与预警系统, 以支撑增程式飞机等复杂平台的绝缘安全保障。

4 结束语

1) 增程式飞机作为燃油飞机向全电飞机过渡的重要桥梁, 展现了显著的技术潜力与发展空间。本文系统梳理了增程式飞机发展脉络、工作原理及关键技术, 重点分析了增程式架构在能量解耦、效率提升及能量管理等方面的优势, 明确了其在中、短途通勤、支线物流及城市空中交通中的应用潜力。同时, 总结了高功率密度电机、能量分配策略、热管理方法及局部放电原理、监测、防护等核心技术的发展现状, 并指出了高功重比发电系统、轻量化材料等方面的关键技术挑战, 为后续增程式飞机的发展及关键部件技术研发提供了参考依据。

2) 增程式飞机技术的深化发展需攻克高功重比发电系统与轻量化材料等核心瓶颈, 同时, 重构适应混合动力的适航标准体系。通过多学科协同设计优化动力链、热管理及机体结构的系统集成, 结合智能能量管理算法, 可充分释放航程效能优势。短途通勤与支线物流将成为技术落地的突破口, 而规模化商用依赖关键部件工程化突破与绿色航空产业链的协同支撑。随着低碳航空动力的技术迭代, 该架构有望在效率提升与环境友好的双重维度上克服关键技术难题, 为全电式飞机降本增效提供创新解决方案。

参考文献(References)

[1] Li Q, Yang Y, Yu X Q, et al. A 700 $W\cdot h\cdot kg^{-1}$ rechargeable pouch type lithium battery[J]. Chinese Physics Letters, 2023, 40(4): 048201.

[2] Lilium. The lilium jet[EB/OL]. [2025-05-09]. https://lilium.com/.

[3] Bye Aerospace. eFlyer aircraft-Bye Aerospace[EB/OL]. [2025-05-09]. https://byeaerospace.com/.

[4] Roboam X. A review of powertrain electrification for greener aircraft[J]. Energies, 2023, 16(19): 6831.

[5] Silva F A. Modern electric, hybrid electric, and fuel cell vehicles, third edition[J]. IEEE Industrial Electronics Magazine, 2018, 12(4): 46-48.

[6] Peters A. This 19-seater plane is the first ever to run on a hydrogen-electric engine[EB/OL]. (2023-01-19)[2025-05-09]. https://www.fastcompany.com/90810924/zeroavia-hydrogen-electric-powered-plane-completes-first-flight/.

[7] Merical K, Beechner T, Yelvington P. Hybrid-electric, heavy-fuel propulsion system for small unmanned aircraft[J]. SAE International Journal of Aerospace, 2014, 7(1): 126-134.

[8] Thauvin J. Exploring the design space for a hybrid-electric regional aircraft with multidisciplinary design optimisation methods[D]. Toulouse: Institute National Polytechnique de Toulouse, 2018.

[9] Welstead J, Felder J L. Conceptual design of a single-aisle turbo-electric commercial transport with fuselage boundary layer ingestion[C]//Proceedings of the 54th AIAA Aerospace Sciences Meeting. Reston: AIAA, 2016: 1027.

[10] Wall T J, Meyer R. A survey of hybrid electric propulsion for aircraft[C]//Proceedings of the 53rd AIAA/SAE/ASEE Joint Propulsion Conference. Reston: AIAA, 2017: 4700.

[11] 国家质量监督检验检疫总局, 中国国家标准化管理委员会. 电动汽车术语: GB/T 19596—2017[S]. 北京: 中国标准出版社, 2017.
General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Terminology of electric vehicles: GB/T 19596—2017[S]. Beijing: Standards Press of China, 2017(in Chinese).

[12] Airbus. E-Fan X[EB/OL]. [2025-05-09]. https://www.airbus.com/en/innovation/low-carbon-aviation/hybrid-and-electric-flight/e-fan-x.

[13] Friedrich C, Robertson P A. Hybrid-electric propulsion for aircraft[J]. Journal of Aircraft, 2015, 52(1): 176-189.

[14] Kim D H. Progress in distributed electric propulsion vehicles and technologies[C]//Proceedings of the 2020 AIAA/IEEE Electric Aircraft Technologies Symposium. Piscataway: IEEE, 2020: 1-11.

[15] Thapa N, Ram S, Kumar S, et al. All electric aircraft: a reality on its way[J]. Materials Today: Proceedings, 2021, 43: 175-182.

[16] Wall T J, Meyer R T. Hybrid electric aircraft switched model optimal control[J]. Journal of Propulsion and Power, 2020, 36(4): 488-497.

[17] Yang B N, Lou F Y, Key N L. Conceptual design of a 10-passenger thin-haul electric aircraft[C]//Proceedings of the 2020 AIAA/IEEE Electric Aircraft Technologies Symposium. Piscataway: IEEE Press, 2020: 1-18.

[18] 李龙, 王鹏. 混合电推进系统关键技术研究进展[J]. 燃气涡轮试验与研究, 2021, 34(3): 49-53.
Li L, Wang P. Development of key technologies for hybrid electric propulsion system[J]. Gas Turbine Experiment and Research, 2021, 34(3): 49-53(in Chinese).

[19] Zunum Aero. Bringing you electric air travel out to a thousand miles[EB/OL]. [2025-05-09]. https://zunum.aero/.

[20] Electra. Electra reveals design for EL9 ultra short hybrid-electric aircraft[EB/OL]. (2024-11-13)[2025-05-09]. https://www.electra.aero/news/electra-reveals-design-for-el9-ultra-short-hybrid-electric-aircraft.

[21] Bell. Bell nexus 6HX[EB/OL]. [2025-05-09]. https://evtol.news/bell-air-taxi.

[22] 辽宁锐翔通用飞机制造有限公司. RX1E-A[EB/OL]. [2025-05-09]. http://www.rgac.net.cn/c/RX1E-A.html.
Liaoning Ruixiang General Aircraft Manufacturing Co. , Ltd. RX1E-A[EB/OL]. [2025-05-09]. http://www.rgac.net.cn/c/RX1E-A.html(in Chinese).

[23] Diamond Aircraft. Diamond aircraft proudly presents the world's first serial hybrid electric aircraft "DA36 E-Star" [EB/OL]. (2011-06-23)[2025-05-09].https://www.diamondaircraft.com/en/about-diamond/newsroom/news/article/diamond-aircraft-proudly-presents-the-worlds-first-serial-hybrid-electric-aircraft-da36-e-star/.

[24] Diamond Aircraft. eDA40—all electric aircraft[EB/OL]. [2025-04-28]. https://www.diamondaircraft.com/en/flight-school-solution/aircraft/eda40/.

[25] Pipistrel. Panthera hybrid-electric aircraft[EB/OL]. [2025-04-28]. https://www.pipistrel-aircraft.com/products/panthera/.

[26] University of Stuttgart. Hybrid propulsion system for e-Genius aircraft[EB/OL]. (2015-08-10)[2025-05-09]. https://www.dspace.com/ko/kor/home/applicationfields/stories/uni-stuttgart-e-genius.cfm.

[27] Airbus. Airbus, Rolls-Royce, and Siemens team up for electric future partnership launches E-Fan X hybrid-electric flight demonstrator[EB/OL]. (2017-11-28)[2025-05-09]. https://www.airbus.com/en/newsroom/press-releases/2017-11-airbus-rolls-royce-and-siemens-team-up-for-electric-future.

[28] University of Stuttgart. e-Genius is flying with hybrid propulsion for the first time[EB/OL]. [2025-05-09]. https://www.uni-stuttgart.de/en/university/news/all/e-Genius-is-flying-with-hybrid-propulsion-for-the-first-time/.

[29] Qu R H, Lipo T A. Dual-rotor, radial-flux, toroidally-wound, permanent-magnet machines[C]//Proceedings of the Conference Record of the 2002 IEEE Industry Applications Conference. 37th IAS Annual Meeting. Piscataway: IEEE Press, 2002: 1281-1288.

[30] Zhang X L, Haran K S. High-specific-power electric machines for electrified transportation applications-technology options[C]//Proceedings of the 2016 IEEE Energy Conversion Congress and Exposition. Piscataway: IEEE Press, 2017: 1-8.

[31] Wang Y L, Nuzzo S, Zhang H, et al. Challenges and opportunities for wound field synchronous generators in future more electric aircraft[J]. IEEE Transactions on Transportation Electrification, 2020, 6(4): 1466-1477.

[32] Andrade L, Tenning C. Design of Boeing 777 electric system[J]. IEEE Aerospace and Electronic Systems Magazine, 1992, 7(7): 4-11.

[33] Ferreira C A, Richter E. Detailed design of a 250 kW switched reluctance starter/generator for an aircraft engine[J]. SAE Transactions, 1993, 102: 289-300.

[34] Masson P J, Breschi M, Tixador P, et al. Design of HTS axial flux motor for aircraft propulsion[J]. IEEE Transactions on Applied Superconductivity, 2007, 17(2): 1533-1536.

[35] Schramm A, Gerling D. Researches on the suitability of switched reluctance machines and permanent magnet machines for specific aerospace applications demanding fault tolerance[C]//Proceedings of the International Symposium on Power Electronics, Electrical Drives, Automation and Motion, 2006. Piscataway: IEEE Press, 2006: 56-60.

[36] Garcia A, Cusido I, Rosero J A, et al. Reliable electro-mechanical actuators in aircraft[J]. IEEE Aerospace and Electronic Systems Magazine, 2008, 23(8): 19-25.

[37] Villani M, Tursini M, Fabri G, et al. Multi-phase fault tolerant drives for aircraft applications[C]//Proceedings of the Electrical Systems for Aircraft, Railway and Ship Propulsion. Piscataway: IEEE Press, 2010: 1-6.

[38] Villani M, Tursini M, Fabri G, et al. High reliability permanent magnet brushless motor drive for aircraft application[J]. IEEE Transactions on Industrial Electronics, 2012, 59(5): 2073-2081.

[39] Liu Y G, Pillay P. Improved torque performance of switched reluctance machines by reducing the mutual saturation effect[J]. IEEE Transactions on Energy Conversion, 2004, 19(2): 251-257.

[40] Van Der Geest M, Polinder H, Ferreira J A, et al. Machine selection and initial design of an aerospace starter/generator[C]//Proceedings of the 2013 International Electric Machines & Drives Conference. Piscataway: IEEE Press, 2013: 196-203.

[41] Sayed E, Abdalmagid M, Pietrini G, et al. Review of electric machines in more-/hybrid-/turbo-electric aircraft[J]. IEEE Transactions on Transportation Electrification, 2021, 7(4): 2976-3005.

[42] Mecrow B C, Jack A G, Atkinson D J, et al. Design and testing of a four-phase fault-tolerant permanent-magnet machine for an engine fuel pump[J]. IEEE Transactions on Energy Conversion, 2004, 19(4): 671-678.

[43] Zhang L F, Wang K, Sun H Y, et al. Multiphase PM machines with halbach array considering third harmonic flux density[J]. IEEE Transactions on Industrial Electronics, 2019, 66(12): 9184-9193.

[44] Elloumi N, Bortolozzi M, Masmoudi A, et al. Numerical and analytical approaches to the modeling of a spoke type IPM machine with enhanced flux weakening capability[J]. IEEE Transactions on Industry Applications, 2019, 55(5): 4702-4714.

[45] Koch S F, Peter M, Fleischer J. Lightweight design and manufacturing of composites for high-performance electric motors[J]. Procedia CIRP, 2017, 66: 283-288.

[46] Zhou Y, Tian L, Gao S H, et al. Design and testing of CFRP sleeve for a high-speed permanent magnet synchronous motor with surface-mounted rotor[J]. Journal of Mechanical Science and Technology, 2021, 35(1): 221-230.

[47] Nøland J K, Leandro M, Suul J A, et al. High-power machines and starter-generator topologies for more electric aircraft: a technology outlook[J]. IEEE Access, 2020, 8: 130104-130123.

[48] El-Refaie A, Osama M. High specific power electrical machines: a system perspective[J]. CES Transactions on Electrical Machines

and Systems, 2019, 3(1): 88-93.

[49] Tenconi A, Vaschetto S, Vigliani A. Electrical machines for high-speed applications: design considerations and tradeoffs[J]. IEEE Transactions on Industrial Electronics, 2014, 61(6): 3022-3029.

[50] Kiss G M, Vajda I. Magnet leakage flux of axial flux PM synchronous machines for EV[C]//Proceedings of the 2014 IEEE International Electric Vehicle Conference. Piscataway: IEEE Press, 2015: 1-5.

[51] Emrax. Emrax 348 axial flux electric motor specifications (400 kW | 1000 Nm) [EB/OL]. [2025-05-09]. https://emrax.com/products/emrax-348/.

[52] Yoon A, Yi X, Martin J, et al. A high-speed, high-frequency, air-core PM machine for aircraft application[C]//Proceedings of the 2016 IEEE Power and Energy Conference at Illinois. Piscataway: IEEE Press, 2016: 1-4.

[53] Huang S, Luo J, Leonardi F, et al. A general approach to sizing and power density equations for comparison of electrical machines [C]//Proceedings of the Conference Record of the 1996 IEEE Industry Applications Conference Thirty-First IAS Annual Meeting. Piscataway: IEEE Press, 2002: 836-842.

[54] Huang S R, Luo J, Leonardi F, et al. A comparison of power density for axial flux machines based on general purpose sizing equations[J]. IEEE Transactions on Energy Conversion, 1999, 14(2): 185-192.

[55] Singh B, Singh B N, Chandra A, et al. A review of three-phase improved power quality AC-DC converters[J]. IEEE Transactions on Industrial Electronics, 2004, 51(3): 641-660.

[56] Benzaquen J, He J B, Mirafzal B. Toward more electric powertrains in aircraft: technical challenges and advancements[J]. CES Transactions on Electrical Machines and Systems, 2021, 5(3): 177-193.

[57] Liu R, Xu L, Kang Y L, et al. Decoupled TAB converter with energy storage system for HVDC power system of more electric aircraft[J]. The Journal of Engineering, 2018, 2018(13): 593-602.

[58] Zhang P, Yan F W, Du C Q. A comprehensive analysis of energy management strategies for hybrid electric vehicles based on bibliometrics[J]. Renewable and Sustainable Energy Reviews, 2015, 48: 88-104.

[59] Murgovski N, Johannesson L, Sjöberg J. Convex modeling of energy buffers in power control applications[J]. IFAC Proceedings Volumes, 2012, 45(30): 92-99.

[60] Hung J Y, Gonzalez L F. On parallel hybrid-electric propulsion system for unmanned aerial vehicles[J]. Progress in Aerospace Sciences, 2012, 51: 1-17.

[61] Lee H D, Sul S K. Fuzzy-logic-based torque control strategy for parallel-type hybrid electric vehicle[J]. IEEE Transactions on Industrial Electronics, 1998, 45(4): 625-632.

[62] Karunarathne L, Economou J T, Knowles K. Adaptive neuro fuzzy inference system-based intelligent power management strategies for fuel cell/battery driven unmanned electric aerial vehicle[J]. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 2010, 224(1): 77-88.

[63] Elbert P, Nüesch T, Ritter A, et al. Engine on/off control for the energy management of a serial hybrid electric bus via convex optimization[J]. IEEE Transactions on Vehicular Technology, 2014, 63(8): 3549-3559.

[64] Njoya Motapon S, Dessaint L A, Al-Haddad K. A comparative study of energy management schemes for a fuel-cell hybrid emergency power system of more-electric aircraft[J]. IEEE Transactions on Industrial Electronics, 2014, 61(3):1320-1334.

[65] Donateo T, De Pascalis C L, Ficarella A. An application of the ECMS strategy to a wankel hybrid electric UAV[J]. MATEC Web of Conferences, 2019, 304: 03010.

[66] Zhai Y W, Wang X Y, Zhang H W. A fuzzy-adaptive ECMS based energy management strategy for series hybrid electric aircraft [C]//Proceedings of the ASME Turbo Expo 2024: Turbomachinery Technical Conference and Exposition. New York: ASME, 2024: V005T06A039.

[67] Palaia G, ABU Salem K, Carrera E. Efficient methodology for power management optimization of hybrid-electric aircraft[J]. Aerospace, 2025, 12(3): 230.

[68] Freeman J, Osterkamp P, Green M, et al. Challenges and opportunities for electric aircraft thermal management[J]. Aircraft Engineering and Aerospace Technology, 2014, 86(6): 519-524.

[69] Weragoda D M, Tian G H, Burkitbayev A, et al. A comprehensive review on heat pipe based battery thermal management systems[J]. Applied Thermal Engineering, 2023, 224: 120070.

[70] Zhao Y, Diaz L B, Patel Y, et al. How to cool lithium ion batteries: optimising cell design using a thermally coupled model[J]. Journal of the Electrochemical Society, 2019, 166(13): A2849-A2859.

[71] Coutinho M, Bento D, Souza A, et al. A review on the recent developments in thermal management systems for hybrid-electric aircraft[J]. Applied Thermal Engineering, 2023, 227: 120427.

[72] Hong S H, Zhang X Q, Chen K, et al. Design of flow configuration for parallel air-cooled battery thermal management system with secondary vent[J]. International Journal of Heat and Mass Transfer, 2018, 116: 1204-1212.

[73] Liu Y Q, Cao J W, Song Y C, et al. Analysis of the immersion cooling of electric motors for hybrid aircraft[J]. Process Safety and Environmental Protection, 2023, 178: 695-705.

[74] Potamiti M, Gkoutzamanis V G, Kalfas A I. Thermal management system design for a series hybrid-electric propulsion architecture[J]. The Aeronautical Journal, 2024, 128(1325): 1532-1555.

[75] Asli M, König P, Sharma D, et al. Thermal management challenges in hybrid-electric propulsion aircraft[J]. Progress in Aerospace Sciences, 2024, 144: 100967.

[76] 张本栋, 江军, 李治, 等. 面向未来多电飞机的低气压下局部放电[J]. 航空学报, 2022, 43(7): 268-278.
Zhang B D, Jiang J, Li Z, et al. Partial discharge characteristics of future more electric aircraft under low air pressure[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(7): 268-278(in Chinese).

[77] Shahsavarian T, Pan Y, Zhang Z S, et al. A review of knowledge-based defect identification via PRPD patterns in high voltage apparatus[J]. IEEE Access, 2021, 9: 77705-77728.

[78] Hu G C, Wu G N, Yu R, et al. The influence of pressure on the discharge along oil-paper interface under AC stress[J]. Energies, 2019, 12(10): 1920.

[79] Katsuta G, Toya A, Endoh T, et al. Development of a new detection method of partial discharge for EHV long-distance active cable line[J]. Electrical Engineering in Japan, 1992, 112(7): 77-91.

[80] Bartnikas R. Detection of partial discharges (Corona) in electrical

apparatus[J]. IEEE Transactions on Electrical Insulation, 1990, 25(1): 111-124.

[81] Abdel-Gawad N M K, El Dein A Z, Mansour D A, et al. PVC nanocomposites for cable insulation with enhanced dielectric properties, partial discharge resistance and mechanical performance[J]. High Voltage, 2020, 5(4): 463-471.

[82] Yan W, Phung B T, Han Z J, et al. Partial discharge characteristics of epoxy resin-based nanocomposites fabricated with atomospheric plasma treated SiO_2 nanoparticles[C]//Proceedings of 2011 International Symposium on Electrical Insulating Materials. Piscataway: IEEE Press, 2014: 353-357.

[83] Wester F J. Condition assessment of power cables using partial discharge diagnosis at damped AC voltage[D]. Delft: Delft University of Technology, 2004.

[84] Gulski E, Wester F J, Smit J J, et al. PD detection in power cables using oscillating wave test system[C]//Proceedings of the Conference Record of the 2000 IEEE International Symposium on Electrical Insulation. Piscataway: IEEE Press, 2002: 444-447.

[85] Gulski E, Smit J J, Seitz P N, et al. On-site PD diagnostics of power cables using oscillating wave test system[C]//Proceedings of the 11th International Symposium on High Voltage Engineering. London: IET, 1999, 5: 5.112-5.115.

Developing status and prospect of range-extended aircraft

FENG Shiyu[1,*], FU Ziqi[1], LIU Guannan[2], TAI Huanyu[1], GENG Leiming[3], MIU Richeng[1]

(1. College of Aerospace Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;
2. School of Chemical Engineering and Technology, Xi'an Jiaotong University, Xi'an 710049, China;
3. AVIC General Huanan Aircraft Industry Co., Ltd., Zhuhai 519000, China)

Abstract: All-electric aircraft face limitations in achieving long range due to the energy density constraints of current lithium batteries. Range-extended aircraft, employing an "electric-dominant, fuel-supplement" powertrain architecture, offer a balance between environmental benefits and extended range, positioning them as a crucial transitional solution for aviation decarbonization. This review focuses on the range-extended technology pathway. It provides an overview of the technological roadmaps adopted by representative range-extended aircraft models both domestically and internationally; it compares and analyzes five typical hybrid powertrain architectures, elucidating the definition and operational principles of the range-extended concept. It explores the advances and challenges of key enabling technologies that are driving its development, including high-power-density permanent magnet motors, intelligent energy management systems, efficient thermal management systems, and high-voltage partial discharge monitoring and protection technologies. This review clarifies the current landscape and core barriers of range-extended technology. It provides essential insights to inform the technological evolution towards long-range adaptability for all-electric aircraft and serves as a valuable reference for formulating strategies aimed at the low-carbon transition of the aviation industry.

Keywords: range-extended aircraft; hybrid power aircraft; all-electric aircraft; permanent magnet motor; energy management system; thermal management; partial discharge

Received: 2025-06-14; **Accepted**: 2025-07-17; **Published Online**: 2025-09-23 13:04

URL: link.cnki.net/urlid/11.2625.V.20250923.1120.001

Foundation item: 2024 Graduate Research and Practice Innovation Program of Nanjing University of Aeronautics and Astronautics (xcxjh20240112)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0615

机载智能氧气系统关键技术综述

卜雪琴[1,*]，刘一鸣[1]，林贵平[1,2]，郁嘉[1]，于坤洋[3]

(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 国际创新学院，杭州 311115；
3. 中国人民解放军 93126 部队，北京 100083)

摘　　要：面向高性能战斗机对精准生命保障的需求，具备生理监测、状态预测和主动干预功能的机载智能氧气系统已成为下一代航空生命保障系统的重要发展方向。系统梳理机载氧气系统的发展历程及国内外智能化研究进展，重点分析电子式氧气调节、飞行员生理监测和生理状态预测3项关键技术的应用潜力及工程化制约因素。分析表明：电子式氧气调节器为将生理信息接入供氧控制闭环提供了工程基础，但复杂飞行环境下的传感可靠性、个体化状态预测、干预策略匹配、闭环稳定性和飞行验证仍是制约其工程应用的关键问题。未来应加强多模态生理信息融合、机理与数据融合建模及分级试验验证，推动飞行员生理状态、座舱环境与供氧设备的协同感知和闭环调控。

关　键　词：氧气系统；生命保障系统；智能控制；战斗机；人机系统；飞行员
中图分类号：V244.1; R852.8
文献标志码：A　　**文章编号：**1001-5965(2026)10-3482-14

机载氧气系统承担战斗机飞行员生理供氧和加压防护任务，其为飞行员提供相对独立的安全呼吸环境，以保障其持续正常呼吸[1]。随着高性能战斗机的研制与列装，飞行员面临更为严苛的低气压、低氧分压和高敏捷过载等环境条件[2]，这些环境因素会对影响飞行作战任务的完成，机载氧气系统也应同步发展，以保障战斗机飞行员在极限生理范围内安全执行飞行任务的能力。在传统的战斗机生命保障体系中，飞行安全依赖于飞行员处于正常生理状态，飞行员一旦出现生理不适甚至失能状况，极容易导致不可挽回的灾难性后果。虽然现有军用标准严格约束了在不同高度和过载下机载氧气系统的供气成分和供气压力，但现有控制逻辑始终未将飞行员生理状态引入氧气系统调控中，目前，飞行员呼吸紊乱、呼吸困难、缺氧甚至意识丧失等生理事件及其导致的严重事故依然频发[3]。

先进战斗机生命保障系统的研发不可忽视对飞行员生理状态的感知与干预，并将其引入氧气系统控制闭环。综合考虑飞行员生理状态与飞行状态，构建具有监测、预测和干预功能的智能氧气系统是下一代生命保障系统的发展趋势。实现智能供氧需要3项关键技术：飞行员生理监测技术、生理状态预测技术和电子式氧气调节技术。本文介绍机载智能氧气系统的研究背景，回顾机载氧气系统发展历史与国内外研究现状，分析飞行员生理监测技术、生理状态预测技术和电子式氧气调节技术3项关键技术的研究进展，并提出机载智能氧气系统可能的实现途径和设想。

1　机载智能氧气系统的研究背景

1.1　氧气系统相关生理事件

战斗机飞行员依赖氧气系统实现持续正常呼吸，人体对呼吸气体的需求主要包括充足的通气量和适宜的气体组分2个方面，飞行员通气参数与人

收稿日期：2024-08-23；**录用日期：**2024-11-29；**网络出版时间：**2024-12-24 14:26
网络出版地址：link.cnki.net/urlid/11.2625.V.20241224.0936.001
基金项目：北京航空航天大学沈元学院卓越研究基金 (230123207)
*通信作者. E-mail：buxueqin@buaa.edu.cn

引用格式：卜雪琴，刘一鸣，林贵平，等. 机载智能氧气系统关键技术综述 [J]. 北京航空航天大学学报，2026，52（10）：3482-3495.
Bu X Q，Liu Y M，Lin G P，et al. Review of key technologies of on-board intelligent oxygen system[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（10）：3482-3495 （in Chinese）.

体生理特征不匹配制约飞行员生命保障系统的精准控制。

美军飞行员生理事件频发。2010 年 11 月 16 日, F-22 飞机因机载制氧故障导致飞行员丧失意识, 继而发生机毁人亡事件, 并促使 F-22 全面停飞[4]。F-35A、T-45、F/A-18 等机型飞行员也频繁遭遇因机载氧气系统引起的生理事件。根据美国对 T-45 和 F/A-18 飞机生理事件的调查, 2012—2017 年 T-45 飞机氧气系统失效共发生 52 起, 如图 1 所示。F/A-18 生理事件在 2011 年发生 31 起, 2016 年发生 125 起, 2017 年的前 5 个月出现 52 起, 生理事件数量呈现逐年上涨趋势。据美国生理事件团队调查, F/A-18 飞机生理事件中有 41% 与呼吸气体输送系统有关[5]。

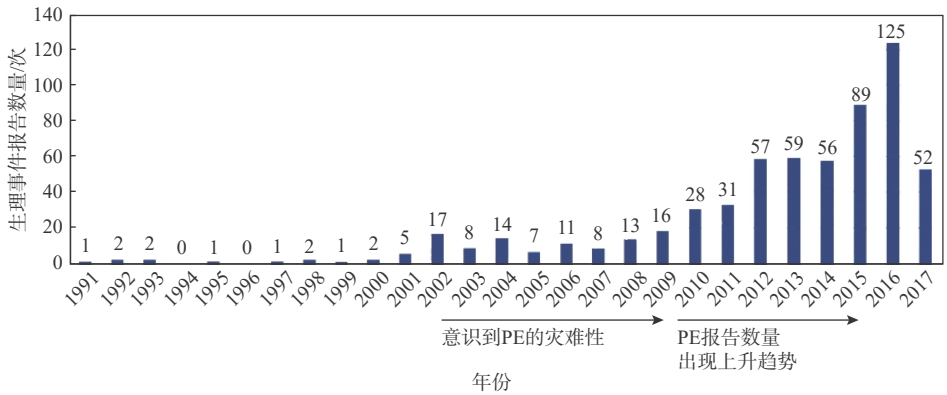


图 1　F/A-18, EA-18G 生理事件统计报告

Fig. 1　Physiological episode reports of F/A-18, EA-18G

中国某型飞机也多次出现缺氧生理事件。2017 年, 某型飞机的一名飞行员在 17 000 m 高空因缺氧昏迷, 下降到 11 000 m 后才恢复意识; 2021 和 2022 年, 某型飞机飞行员在巡航高度飞行时出现缺氧症状。据统计, 中国因高空缺氧引起的事故症候约占 6%～17%, 因高过载意识丧失引起的飞行事故约占严重飞行事故总数的 30%～40%。据悉, 飞行员呼吸紊乱、呼吸体验不佳, 甚至因供氧量不足而被迫摘下面罩的呼吸困难事件也时有发生。

1.2　航空供氧防护愈加重要

随着先进高性能战斗机的研制与列装, 超机动与长航时的飞行任务对飞行员生理耐受能力提出新的挑战。航空环境因素的突变能引起一系列的生理效应, 强烈的环境刺激可能会使人的生理反应过于剧烈, 导致飞行失能, 并影响飞行操纵。严苛的环境也致使人体产生病理变化, 或导致功能性影响与器质性损害。高性能战斗机的生命保障系统应保障飞行员具备在生理极限范围内进行空对空战斗的能力。

飞行员在进行复杂的高机动飞行(翻滚、旋转和回转等)时, 严苛的力学环境对循环系统、呼吸系统、视觉功能、脑功能等产生不利影响。高敏捷过载可能会诱发意识丧失、视觉障碍、心律失常和呼吸困难等症状。现代高性能战斗机所产生的正加速度+Gz 的特点是: 加载速率快、高过载值、长作用时间, 并可反复出现[6]。当表征脑部供血情况的眼水平动脉压过低时, 甚至会引发严重的意识丧失(G-induced loss of consciousness, G-LOC), 易导致机毁人亡的严重后果。据报道, 美国空军飞行员 G-LOC 的发生率为 25%, 海军飞行员发生率为 14%[7]。1982—1990 年, 美国空军发生 18 起归因于 G-LOC 的飞行事故, 造成 14 人死亡[8]。

飞行员的呼吸体验是由人体与氧气系统的匹配程度决定的。不恰当的通气参数会造成肺部局部过度通气, 致使局部应力和应变升高, 破坏肺的结构和功能, 并出现一系列炎症反应[9]。同时, 吸入气体的组分应保持在适宜的范围内, 呼吸中枢根据血液中二氧化碳、氧和氢离子水平调节通气[10], 吸入氧气体积分数过低易造成高空缺氧, 长时间吸入高体积分数氧气会造成吸收性肺不张等症状。

加压呼吸是提升飞行员过载耐受能力, 防止呼吸抑制的重要防护手段, 其作用机制是通过提高胸膜腔内压, 利用循环系统传递的压力使脑部血管的血压同步升高, 从而改善脑部供血。然而, 目前加压呼吸压力制度通常采用随过载线性梯度增加的加压方式, 更严苛的过载条件对航空供氧防护提出了新要求。

1.3　飞行员生理态势感知缺乏

为保障飞行员的安全, 并最大限度地保持其认知能力, 应该对飞行员进行监控和辅助性干预[10]。实际上, 现有的人机系统较少能实现飞行员实时监控。

“态势感知”的概念最早在军事领域被提出, Endsley 在 1988 年国际人因工程年会上提出其概

念，并将其定义为：在特定的时间和空间条件下，对环境因素的感知、理解及其未来发展趋势的预测[11]。在智能化时代，战场形势瞬息万变，单纯依靠人脑有限的认知能力难以处理包含环境信息和飞行员生理信息在内的庞大数据。在 F-22 疑似缺氧事故后，美国空军开展了机组穿戴式生理感知系统（aircrew mounted physiologic sensing system，AMPSS）计划[12-13]，通过实时监测飞行员的生理参数，寻找有效防止 F-22 和 F-35 飞行员缺氧的手段。

感知是对脑电、眨眼率、血氧饱和度、心电和肌电等信号进行监测，并通过解算得到相关生理参数。理解则是依据航空威胁环境及其相关的人体生理变化，通过对重要生理参数（如脑电和血氧饱和度等）的组合分类研究，建立生理参数与航空威胁环境之间的关联性，并对所有控制设备获取的原始数据进行分析，去除冗余信息、合并相似信息并修正错误信息，从而得到规范化的生理信号数据集。预测是基于飞行员生理指标建立数学模型，对不同的航空威胁（疲劳、G-LOC 等）的征兆或当前状态进行表征，预测飞行员随飞行环境变化过程中的状态变化趋势。

2 机载氧气系统发展历程

机载氧气系统主要由氧源、监控指示部件、供气调节部件和个体防护部件组成。随着战斗机的迭代升级，氧源经历了由气氧、液氧向机载分子筛制氧的发展，解决了长航时连续供氧的问题；供气调节部件经历了从连续供氧到肺式供氧，从机械式调节到电子式调节的发展，不断提升了飞行员呼吸舒适性与呼吸安全性。

自二战期间德国首先发明肺式氧气调节器以来，肺式供氧已成为氧气系统的主要供氧方式。机械式氧气调节器经过数十年的发展，已经满足基本供氧防护要求，目前，仍广泛应用于多型飞机，如图 2 所示，但仍存在以下不足：



F-22呼吸调节抗荷阀　F-15 CRU-98/A　F/A-18胸配式

图 2　机械式氧气调节器

Fig. 2　Mechanical oxygen regulator

1）调节性能与结构尺寸相关，并受到结构尺寸限制。膜片产生的驱动力需要通过杠杆等机构放大和改变方向，肺式机构复杂。

2）供氧能力与工作性能受入口压力影响显著，尤其是以分子筛氧气浓缩器为主氧源的机载分子筛制氧系统，受限于发动机引气和工作状态，产品气压力不满足调节器入口压力要求，会出现频繁切换备用氧源的现象。

3）呼吸过程中的加压供氧控制需要余压控制机构、余压接通机构等元件共同作用，加压供氧控制精度较差。

4）受限于其供氧调节机制，吸气阻力与调节器设计存在很大关系，吸气阻力大，容易产生呼吸疲劳。

随着机载电子技术的迅速发展及高性能战斗机对氧气系统性能要求的提高，自从法国 Air Liquide 公司于 1995 年研制出世界第一台电子式氧气调节器[14]，如图 3 所示，以传感器反馈、控制器调节、电机驱动肺式活门的电子式氧气调节器开始蓬勃发展。电子式氧气调节器最大特点在于压力感应和肺式活门调节方式实现了电子化。与机械式氧气调节器相比，电子式氧气调节器的压力感应不再依赖膜盒结构，而是通过压力传感器感受微压差；肺式活门的调节也不再是单纯的机械传动，而是利用微电机位移来控制活门开度。采用电信号来控制活门具有动态响应快、可控性强的优点，并为生理信号引入氧气调节器控制闭环提供了可能。

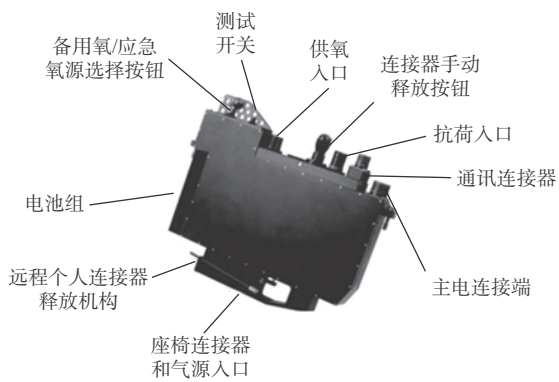


图 3　法国 Air Liquide 电子式氧气调节器

Fig. 3　France Air Liquide electronic oxygen regulator

为避免过载和气压波动引起的缺血与缺氧发生，飞行员生理状态的监测愈发得到重视。20 世纪 80 年代，美军开展过载保护系统的研究，该系统监测飞行员脑电、血氧饱和度等参数，在飞行员出现 G-LOC 现象时进行提醒，虽有成果，但缺少可行且可靠的过载意识丧失预警指标和方法。90 年代，美国开展了技术研讨，近红外光谱技术和脑电磁层析成像技术为研究严重飞行错觉和过载失能时的飞行员脑功能监测提供了可能，呼吸流率和面部脉搏的监测也为过载意识丧失的监测和预警提供了技术手段。

自 2011 年以来, 美军研制了多种飞行员生理监测装置, 包括衣物式、指套式、面罩式等类型, 其发展历程可分为数据采集、告警、主动防护 3 个阶段。2011 年, 针对 F-22 缺氧事件, 美国开发了战术空勤人员生理与认知监测 (tactical aircrew physiological and cognitive monitoring, TAPCOM) 系统, 在面罩上加装气体监测传感器, 监测气体压力、氧气流量、面罩内氧气和二氧化碳体积分数等, 该系统仅用于数据采集。

2015 年, 以色列 Elbit Systems 公司研制了 Canary 飞行员生理监测系统, 利用嵌入头盔的传感器实现对心率、血流量、血氧饱和度等指标的远程监测, 如图 4 所示。头盔前缘的包边条可监测心率、血流量和血氧饱和度等。该系统综合飞行员的生理状态、头部和躯干的姿势及飞机的过载、速度、高度和姿态, 并利用机器学习方法实现自动监测飞行员的生理状态, 为飞行员提供缺氧的早期预警。该系统在 F-16、F-18 和 T-38 战斗机上进行了飞行验证。

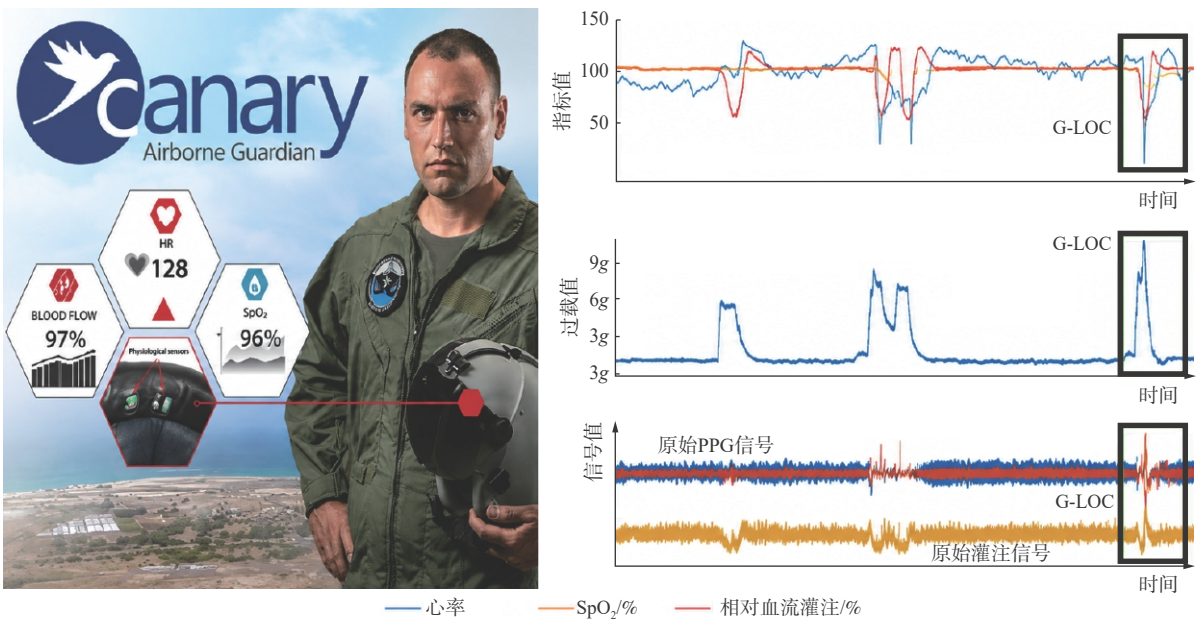


图 4　以色列 Elbit Systems 飞行员生理监测系统

Fig. 4　Israel Elbit Systems health detection device for pilots

2016 年, 芬兰 Insta 公司开发了飞行员呼吸监测 (Insta pilot's breath air monitor, IPBAM) 系统, 如图 5 所示。IPBAM 可持续监测飞行员的呼吸气体和座舱压力, 提供实时警报, 通过视觉、触觉信号提醒飞行员采取行动。该系统还记录飞行数据, 用于分析危险情况和改进飞机系统性能。目前, 该设备已被包括芬兰和奥地利在内的多个国家采用[15]。



图 5　芬兰 IPBAM

Fig. 5　Finland IPBAM

2016 年, 英国 Cobham 公司率先开展下一代智能氧气系统的研制。Cobham 已经完成了为美国空军航天医学院设计的呼气传感器的内部开发, 以支持机组人员安装的生理传感系统计划, 如图 6 所示, 用于实时采集飞行中的氧气系统、驾驶舱环境和飞行员生理数据, 帮助确定生理事件的根本原因。呼吸传感器系统包括吸气传感器模块和呼气传感器模块, 在飞行中实时监测氧气系统、驾驶舱环境和飞行员生理数据, 吸气数据用于判定供氧是否充足, 以及是否由于驾驶舱压力损失引起类似缺氧症状。呼气传感器能够探测飞行员呼吸是否正



图 6　英国 Cobham VigilOX 氧气系统

Fig. 6　UK Cobham VigilOX oxygen system

常，或是否因换气过度而导致缺氧症状。此外，Cobham设计研发了世界上第一款可根据动态生理需求和飞行状态自动调节供氧的传感器驱动型呼吸调节器VigilOX，该系统可测量一系列关键参数，包括飞行员吸入氧体积分数、呼吸压力和流速，以更好地了解人机系统各组成要素对无法解释的生理事件的相互作用[16]。在飞行员自主生命保障系统中，Cobham公司计划整合Canary与VigilOX，协同采集飞行员生理状态数据。最新的SmartFLO电子调节器在VigilOX的基础上新增预测和保护功能，使用实时传感器数据使SmartFLO调节器根据机舱压力和生理线索自动调节供氧流量。

2018年，美国Spotlight Labs研发了耳后生理监测装置SPYDR（standalone performance yielding deliberate risk）人体性能及环境监测系统，通过在飞行员头盔耳罩内置多种传感器，实时监测生理指标（如血氧饱和度、心率、脑部血流等）和环境参数（如加速度、座舱压力、温度等），并提供缺氧告警。同时，其无需与飞行器进行深度集成，可轻松安装在所有头盔上，不影响头盔的质量或平衡。其采用内置电池，支持长时间运行，已被广泛测试并部署于多种军事机型中，包括T-6、F-15、F/A-18等。其有效降低了因缺氧导致的事故风险，被认为是解决飞行员生理问题的突破性工具[17]。

2020年，Cobham公司提出下一代氧气系统——自动响应氧气系统（auto response guided oxygen system, ARGOS），包含从气源到面罩的整个氧气系统。其可与VigilOX协同工作，根据飞行员生理数据和飞行条件实时调整供氧体积分数，确保精准供氧。

自中国引进原苏联航空氧气系统并应用于第一代战斗机以来，中国氧气系统便走上了一条先仿制再自行研发的路线。经过中国科技人员的多年努力，目前已取得了显著进展。

20世纪50年代，中国从前苏联引进了13台低压舱，开始了高空飞行生命保障研究。1960年，贾司光研究员与北京航空学院（现北京航空航天大学）合作，对从前苏联引进的MiG-19和MiG-21飞机的氧气设备进行了消化分析和实验评定，并为多种机型制定了高空氧气生理防护方案。20世纪60～70年代，中国引进了原苏联的KKO-1和KKO-3氧气系统，仿制了YX-1、YX-1A、YX-31和YX-2氧气系统，并大量应用于第二代战斗机。进入新世纪，在中国科技人员的努力下，中国已先后发展了YX-3、YX-4、YX-5、YX-6、YX-7、YX-8、YX-9等多型氧气系统，初步形成了中国独特的航空氧气设备体系。

对于新型的电子式氧气调节器，北京航空航天大学的潘锐[14]、盛勇[18]均对相关部件进行了建模及仿真研究。上海航空测控技术研究所的孙灿飞等[19]对电子式氧气调节器的控制技术进行了相关的探索及研究，提出一套基于模糊控制理论对步进电机进行控制的氧气调节器控制方案，取得了一定的研究成果。到目前为止，中国已经拥有了一系列的航空氧气系统，其中，较为典型的氧气系统为YX-4、YX-9、氧-11和间接式氧气系统。近年来，为适应中国新型号和不同机种飞机的需求，中国已经开展了新型氧气系统、间接式氧气系统与电子式氧气系统相结合的研发工作，并取得了较大进展。

目前，中国现有氧气系统尚未将飞行员生理状态引入控制闭环，新一代智能氧气系统仍处于研制初级阶段。在飞行员生理监测方面，中国研发的新型飞行头盔生理参数检测装置原理样机实现了部分生理指标的监测，其利用现有飞行设备，可在飞行结束返回地面后对监测数据进行分析，以评价飞行员飞行中的生理状况[20]。同时，为某型飞机配套研制的耳罩式和面罩式生理监测设备目前在试验阶段，具备血氧饱和度、耳脉搏和氧分压参数采集、存储功能。

综上所述，机载氧气系统经百余年发展，其供氧方式已实现从机械式到电子式的跨越，并逐渐向智能化发展。为满足更严苛的飞行任务需要，单纯依据过载和高度制定的供氧策略已暴露出诸多问题，将飞行员生理纳入机载氧气系统的控制闭环成为未来发展方向。目前，实现智能供氧需要具备以下3个关键技术：飞行员生理监测技术、生理状态预测技术和电子式氧气调节技术，以实现飞行员状态、座舱环境、供氧设备三者协同联动。

在飞行员生理监测技术方面，中国与国外都开展了原理机的设计及试验，但在生理指标选取、生理状态评价方法等方面仍未形成“金标准”。对于飞行活动中的生理状态预测技术研究较少，而其他日常场景下的预测理论研究为飞行员生理状态预测提供了方法借鉴。对于智能电子式氧气调节技术，智能供氧控制律是联系飞行员与氧气系统的核心，经感知、分析、理解后的飞行员生理信号应融入现有的控制策略中以适配飞行员生理需求。

3 飞行员生理监测技术

飞行员生理监测技术是采用无创、连续、精准的生理监测技术获取心率、血压、血氧饱和度、呼吸频率、潮气末二氧化碳分压等生理信号。通过综合多项敏感生理指标及飞机机动状态，该技术能在

飞行过程中实时识别飞行员的生理状况, 帮助预测并预防潜在风险。

相比于地面环境, 飞行环境对飞行员的生理信号监测提出 3 个主要挑战:

1) 航空环境复杂。振动、温度变化、高过载、电磁干扰和飞行员运动干扰等使监测系统的精度和可靠性面临考验。

2) 数据质量要求。当前监测系统多基于实验室研究, 无法完全模拟飞行环境的实际复杂性, 数据分析需要结合多种传感器信息, 避免因数据质量问题导致误判。

3) 装机应用要求。飞行员使用的生理监测设备应具备较好的佩戴舒适性和稳定性。

3.1 生理指标筛选

3.1.1 氧合指标

1) 局部脑氧饱和度 (regional cerebral oxygen saturation, rSO_2)。急性高空缺氧和高过载会导致脑组织氧合水平降低[3,21-22], 在低压舱和离心机实验中, rSO_2 随着高度或载荷的上升而降低, 当 rSO_2 下降 6%～7% 时, 受试者发生意识丧失或失能, rSO_2 低于阈值的时间越长[23], 失能时间就越长。血氧饱和度传感器的常规安装位置有手腕、手指、耳垂、耳道或前额等身体部位, 位于颞骨乳突上方、耳后耳廓和发际之间的区域能更快地收集有关血液氧饱和度重要生理变化的数据[24]。

2) 脑组织氧合血红蛋白 (oxyhemoglobin, HbO_2) 含量。缺氧与+Gz 暴露导致脑组织去氧血红蛋白含量升高, HbO_2 含量降低。Smith[25] 和 Eiken[26] 等发现, 前额叶皮质氧合水平的降低最为显著, 前额叶皮质 HbO_2 是预测 G-LOC 的有效指标。除此之外, Kurihara 等[27] 认为, 额叶皮质组织氧合指数 (tissue oxygenation index, TOI)也是评价过载耐力的有效指标, 其研究结果显示, G-LOC 发生时, 飞行员前额叶皮质的 TOI 降低了 15%。不足之处是, 该指标对+Gz 的响应存在较大个体差异, 用来预测 G-LOC 时准确度有限。

3.1.2 血流动力学指标

1) 眼水平动脉血压。人的意识状态与头部供血状况有着密切联系。在持续性载荷的作用下, 血液将沿着惯性力方向流动, 头部眼水平位置会发生明显的血压降低, 导致脑组织缺血缺氧, 引发意识丧失。

眼水平动脉血压低于 50 mmHg 时, 周边视力开始丧失, 当低于 20 mmHg 时, 眼水平动脉血压便无法克服眼内压, 从而造成视网膜缺血, 引起视觉丧失与黑视[28]。因此, 监测头眼水平的血供有助于掌握飞行员在高载荷机动时的意识状态。

2) 耳脉搏。耳脉搏是客观评估飞行员过载耐力和意识状态的有效指标, 高过载期间, 耳脉搏波呈现周期性波动, 与主观的中心/周边视野消失相比, 通过耳脉搏信号可更加客观地评估监测头眼部供血情况[29-31]。

3.1.3 肺通气指标

肺通气指标也是常见的生理监测指标, 监测时通过呼吸传感器获得呼吸信号, 提取出能反映人体呼吸状态的物理量, 如肺通气量 $\dot{V}_E$、吸气量 V_I、呼气量 V_E、呼气流率峰值 R_{PEF}、吸气流率峰值 R_{PIF}、吸气间期 t_I、呼气间期 t_E、呼吸频率 f 和呼气末二氧化碳分压 P_{etCO_2} 等。

国家军用标准《飞行员肺通气参数》(GJB 306A—97)[32] 规定了静态肺通气、用力呼吸、佩戴供氧面罩和加压呼吸等状态下的肺通气参数。《飞行员飞行中肺通气参数》(GJB 305—87)[33] 规定了军用飞机飞行员飞行过程中的肺通气参数范围。

3.1.4 心电指标

飞行环境对心血管系统有显著影响, 如压力变化、高过载、缺氧和精神压力, 这些因素可能诱发心脏异常, 因此, 心电图(electrocardiogram, ECG)监测对保障飞行员的安全和健康至关重要。适用于飞行员的 ECG 监测设备应尽量避免监测过程对飞行员的干扰, 在满足监测要求的前提下, 应尽量减少 ECG 导联数量[34]。常用的ECG 指标有: 心率、心率变异性(heart rate variability, HRV)、QRS 波群(QRS complex, 由 Q 波、R 波和 S 波组成)、RR 间期(R-R interval, 相邻 2 次 R 波峰之间的时间间隔)、QT 间期(QT interval, 自 Q 波起点至 T 波终点)等。

心肌缺氧与过载会诱发心动过速、室性早搏等心律失常症状[35-36], 其影响可体现在 ECG 的 QRS 波形上[37]。与无症状者相比, G-LOC 前兆症状者峰值心率更低, QT 间期更长, QT/RR 比值更高, 这些参数在预测 G-LOC 前兆症状的出现时显示出较高的准确性, 在该研究样本和判别方法下, 最大心率的预测准确率为 62.16%, 而 QT/RR 比值的预测准确率达到了 83.78%[38]。同时, ECG 指标除了与心血管代偿控制机制有关, 还受到心理状态的显著影响, 飞行员心理应激水平的差异也将导致心血管反应发生变化[39]。在实际飞行应用中, ECG 指标需要综合多项指标进行评估, 以提高诊断的特异性和敏感性, 减少误判的可能性[40]。

3.1.5 脑电指标

脑电信号 (electroencephalography, EEG) 是大脑神经元活动产生的电信号, 通过放置在头皮上或直

接植入大脑的电极进行采集，其波形、波幅、频率和相位等特征可以直接反映大脑功能。

1）波幅。在缺氧、昏迷、睡眠及麻醉等意识水平降低状态下，EEG常可出现高幅慢波等典型变化[41-43]。

2）各频带功率谱。有国内外研究揭示，在进行特技飞行操作时，脑电图的各频带平均频率显著升高，特别是α、β频带的功率百分比呈现出显著增长的趋势，并与载荷的增加呈正相关[43-44]。EEG的频率特征也与G-LOC有一定相关性[45]，在G-LOC发生前，β频带的功率呈现出持续下降的趋势。G-LOC发生时，β频带的功率向频率较低的θ和δ频带转移，同时，还观察到低频频带功率显著增强[46]。

3）事件相关电位 (event-related potential, ERP)。ERP是一种特殊的脑诱发电位，反映大脑对特定事件或刺激的感知、处理和反应，是评定大脑功能状态的有效指标[47]。ERP通过对刺激后的脑电反应进行时间锁定分析，从而以毫秒级时间分辨率表征大脑的认知加工过程。

不同的ERP成分（如P300、N200）与不同的认知功能（如注意力、记忆、冲突监测）相关联。例如，当出现急性缺氧时，P300振幅降低，潜伏期延长，大脑出现认知功能受损[48-49]。理论上，ERP在G-LOC的鉴定中具有潜在的应用价值，然而，测量稳定性和准确性受限，其在这些特定环境下的实际应用仍存在挑战[50]。

3.2 飞行员生理状态识别

通过采集生理信号可获得生理系统的部分可观测体征信息，然而，人体生理系统是世界上最复杂的系统之一，单一生理指标往往难以完全刻画人体生理状态，复杂高阶信息往往隐含在多类别的生理信息中。因此，为获取飞行员实时状态并对机载氧气系统供气效果进行合理评价，需要对多种生理信号进行有效解读与整合，研发适用于飞行场景的多源生理状态识别算法。这里主要介绍2种可行的方法：基于专家系统的综合呼吸指数和多模态生理状态识别。

3.2.1 综合呼吸指数

综合呼吸指数是一种用于评估呼吸系统功能状态的综合指标，其通过融合多个生理参数来提供全面的呼吸功能评价，被广泛用于临床监测和评估患者的呼吸状况。在临床医学领域，常对患者的动脉血氧饱和度、呼吸频率、心率等生理信号进行实时监测，临床医生则对这些以数字或波形形式呈现的信息进行整合和解读。

2016年，Ronen等[51]将动脉血氧饱和度、呼吸频率、脉搏和潮气末二氧化碳分压综合为一个介于1～10之间的单一数值——综合呼吸指数，以描述当时通气和氧合的充分性，同时，该算法结合了专家的临床意见，通过与专家对临床情景的评分进行比较，对该指数的有效性进行验证，如表1所示。针对飞行场景，也可参考以上方式，在飞行员地面试验中采集飞行员生理数据，综合医学专家与飞行员主观量表制定综合呼吸指数的算法，在飞行期间实时计算综合呼吸指数，以此进行状态识别并开展后续工作。

表1 综合呼吸指数状态描述

Table 1 Integrated pulmonary index status descriptors

综合呼吸指数	人体状态
10	正常
8～9	在正常范围内
7	临近正常范围，需要注意
5～6	需要注意，或许需要干预
3～4	需要干预
1～2	需要马上干预

3.2.2 多模态生理状态识别算法

利用不同类别的生理信号相互补充，对互补信息进行融合处理，能够有效地提高飞行员生理状态的识别效果。1992年，Roscoe[52]指出心率、HRV和呼吸频率3个变量可在模拟和真实飞行中用于评估飞行员的唤醒调节能力。1993年，Jorna[53]同样得出结论，心率、HRV等心血管指标不仅有助于识别不同精神状态，也有助于观察飞行员在模拟和实际飞行中对工作量变化的动态反应。2019年，Charles等[50]指出没有单一测量方法能有效区分精神负荷，但认为生理测量方法确实表征飞行员状态。近年来，随着机器学习领域的蓬勃发展，使用多模态深度学习网络对飞行员的生理状态进行分类检测为提高状态识别精度提供了可能。多项研究[54-56]采用多模态方法来评估真实和模拟飞行中的精神状态。2022年，Masters等[10]应用多模态方法来预测飞行员的心理负荷，表明了该研究方向具有明确的发展潜力。

4 飞行员生理状态预测技术

飞行员生理状态预测是依据实时监测的飞行员生理信息与环境信息，综合当前与历史生理状态数据，预测飞行员随飞行环境变化的状态变化趋势，提前预警可能出现的航空威胁（缺氧、呼吸紊乱、G-LOC等）。根据预测方法不同，对飞行员生理的预测可以大致分为2种：基于生理模型的方法和基于机器学习的数据驱动方法。

4.1　基于生理模型的生理信号预测

生理模型是基于对生理机理的认知, 运用生理知识建立人体生理系统模型, 模拟人体内心血管系统、呼吸系统、自主神经系统等系统的作用, 仿真血氧代谢、血液循环和呼吸等生物过程。

人体根据功能可划分为不同的生理系统, 飞行过程中受影响较显著, 并可能降低飞行员作业效能的主要是心血管系统和呼吸系统。两者皆受神经中枢系统的调控, 例如, 呼吸的节律性是由延髓呼吸中枢控制的, 呼吸调节系统对体液中二氧化碳分压、氧分压及氢离子水平的变化高度敏感, 这些化学刺激可通过中枢和外周化学感受器反馈至呼吸中枢, 从而调节通气水平。

基于生理模型的生理信号预测的主要缺点是: 模型准确度依赖于具有个体差异性的参数, 并需要先验知识来调整模型参数, 模型的普适性、鲁棒性和计算效率存在问题。

4.1.1　呼吸系统模型

呼吸过程分为外呼吸、内呼吸和气体在血液中的运转 3 个环节, 其作用是为身体组织输送氧气并带走二氧化碳。呼吸系统模型一般由呼吸力学模型、气体交换模型和呼吸中枢模型组成。

1946 年, Gray[57] 应用数学模型来探究人体呼吸系统的化学控制机制, 开创了该领域研究。1983 年, Fincham 等[58] 研究了肺通气的机械过程及其机理、呼吸流量、容积和压强等的动态变化规律, 提出呼吸器官的弹性特性和流阻特性是呼吸力学问题的出发点和归宿, 在呼吸力学研究中居重要位置。2000 年, 仇安琪等[59] 建立了包含心脏的人体呼吸系统模型, 能仿真模拟低氧状态和高碳酸状态下人体中氧气和二氧化碳的动态变化过程。

4.1.2　心血管系统模型

人体心血管系统具有几何形态复杂、流动介质成分多样、流动非定常性、管壁顺应性等特点[60], 相较其他工程领域的流动问题求解具有更大难度。近几十年来, 得益于计算机技术的飞速发展和流体运动微分方程数值求解方法的不断进步, 计算流体力学方法逐渐应用于心血管系统的生物力学研究[61]。目前, 心血管系统模型按照尺度分为三维模型、二维模型、一维模型、集中参数模型和跨尺度耦合模型等, 如表 2[62-77] 所示, 这些模型在针对不同尺度、不同部位的仿真问题中发挥了重要作用。

表 2　心血管系统模型对比[62-77]

Table 2　Comparison of models of the cardiovascular system[62-77]

心血管系统模型	概述	适用情况	局限
三维模型[62-63]	考虑血管三维几何结构与力学性能, 血液被视为非牛顿、不可压缩的稳定层流流体	研究血管几何形状对血流动力学的影响	高计算成本, 空间尺度<1m
二维模型[64-65]	利用圆柱几何特性简化血管模型, 将血管视为二维管道	研究管道流径向流动参数变化	较高计算成本, 难以扩展到整个血管系统
一维模型[64,66]	血管模型只依赖于轴向坐标, 将血管等效为一维管道	研究血管系统中波传播现象的理想工具	没有摆脱空间依赖性, 不适用于对血管系统的整体分析
集中参数模型[67-71]	基于血管流体网络和电路网络间的相似性建立, 血管系统可简化为集中参数模型	模拟更大范围的血管系统行为, 方便考虑自我调节、代谢动力学, 运算成本较低	不能反映血管三维流动细节, 选参和调参困难
跨尺度耦合模型[72-77]	利用不同模型各自特点分别仿真循环系统的不同部分	利用集中参数模型为局部血管模型提供合理的边界条件	耦合交界面的处理

与重力生理学有关的心血管系统研究伴随着美苏太空竞赛而增多, Jaron 等[67] 利用心血管系统模型探讨了+Gz 加速度下心血管系统的反应, 并讨论了抗荷服和抗荷动作的效果。Davis 等[68] 提出了经典的集总参数模型, 用于仿真心脏在内的循环系统特征。Mukkamala 等[65] 提出一种心血管系统识别方法, 用于分析无创测量心率、动脉血压和瞬时肺容量的搏动变化。Frolov 等[69] 采用临床方法识别人体生理参数, 推导出含有体循环和肺循环的集总参数模型, 其模型应用于临床治疗心血管疾病的决策。Li 等[70] 建立了心脏四腔室、全身动静脉、典型全身器官循环和部分肺循环室在内的集总参数模型, 并在模型中考虑了自主调节机制, 研究了人体心血管系统在姿势变化和+Gz 加速度下的生理反应。

4.2　基于数据驱动的生理时间序列预测

采集到的飞行员生理数据实际上是多组按时间次序排列且相互关联的生理时间序列, 对生理时间序列进行分析, 能够挖掘数据中蕴含的有价值的信息, 能够揭示生理状态发展的内在规律, 以更好地对生理状态进行识别、预测和调控。“预测”作为时间序列分析的主要研究内容, 其基本思想是利用时间序列的当前值和历史值, 建立能够精确反映时间序列中所包含的动态依存关系的数学模型, 以精确地找出系统的内在特性并以此对序列的未来趋势进行预报。

神经网络特有的非线性处理能力、自学习、自适应、自组织能力和容错功能, 使其在模式识别、智能控制、组合优化和预测等领域获得了成功应用, 被广泛应用于心率、呼吸等生理信号的预测。

飞行员生理状态预测是机器学习领域典型的时间序列预测问题，在传感器采集到大量生理数据后，可以基于这些数据进行建模，并依据提取的特征判断飞行员状态。这些算法能够学习生理系统的复杂性和非线性特征，并实现对个体生理状态的准确预测。

研究者提出多种预测模型，包括自回归移动平均模型[78]、谐波模型[79]、扩展卡尔曼滤波器[80]、支持向量回归[81]、相关向量机[82]、神经网络[83]、长短期记忆网络[84]等。Hong等[85]基于生命体征数据库(vital signs database, VitalDB)，采用深度学习模型，利用光电容积脉搏波(photoplethysmography, PPG)波形对收缩压、舒张压和平均血压变化进行分类，有效实现了对对象的血压实时监测。Reiss等[86]为提高基于PPG信号的连续心率预测方法的通用能力，引入了新的用于日常生活活动条件下运动补偿和心率估计的PPG数据集(PPG dataset for motion compensation and heart rate estimation in daily life activities, PPG-DaLiA)对模型进行训练，使模型更适用于真实生活中的多种活动场景。Tan等[87]针对传统循环神经网络(recurrent neural network, RNN)方法在长期窗口(600 ms)泛化性能下降的问题，提出一种更适用于长期实时预测的呼吸预测长短期Transformer(long short-term Transformer, LSTformer)模型。从生理信号中提取有用的特征是提高预测准确性的关键。研究者正在研究如何从ECG[87-90]、PPG[91-93]等信号中提取呼吸相关的特征，并使用特征选择技术来提高模型的性能。

除一次性离线建模完成预测模型方法外，在线时间序列预测方法能随着时间序列数据的在线更新，预测模型也同时进行在线更新和优化，能满足飞行员生理状态的动态和实时预测。结合现有研究，在线时间序列预测方法主要可分为以下几类，即重新建模方法、动态神经网络方法、在线支持向量回归方法和核自适应滤波等方法[94]。例如，Wang等[95]提出面向随机卷积节点网络的在线序列学习框架(online sequential learning framework for random convolution nodes, OS-fRCN)，通过序列更新模型参数以适应呼吸运动的时变性。呼吸模式具有明显的个体差异，因此，预测模型需兼顾个体适应性和跨个体泛化能力。

5 基于生理感知的电子式氧气调节技术

智能供氧控制律是联系飞行员与氧气系统的核心，经感知、分析、理解后的飞行员生理信号将和环境信号一起用于调控氧气系统以适配飞行员生理需求，如图7所示。当识别或预测出飞行员缺氧时，智能氧气系统应主动提高氧气体积分数或实施加压供气以改善飞行员氧合状态。

5.1 电子式氧气调节

电子式氧气调节器作为氧气系统的核心部件，

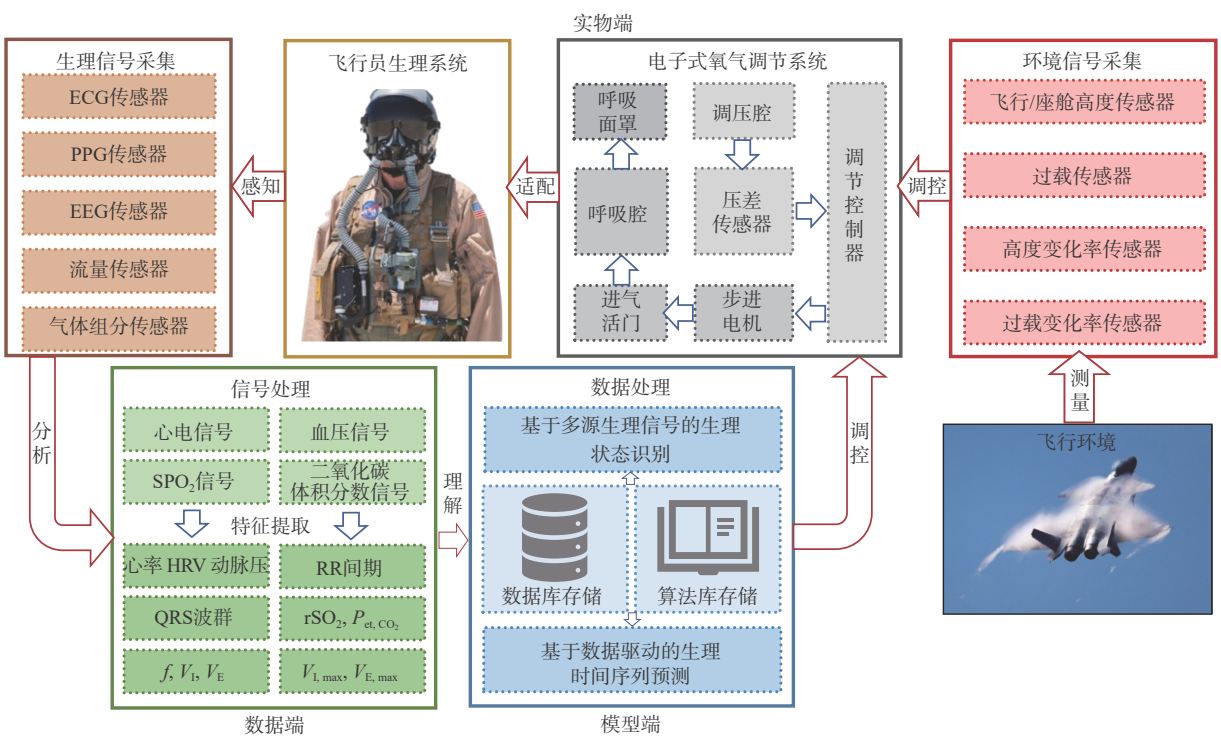


图7 机载智能氧气系统闭环控制逻辑

Fig. 7 Closed-loop control logic for an intelligent on-borne oxygen system

调节方式直接影响整个系统的供氧效率, 其信号采集、处理和驱动的工作方式为生理信号引入调节控制闭环提供了可能。

电子式氧气调节器相比于传统氧气调节器具有动态响应快、可控性强的优点。其压力感应不再是通过膜盒, 而是通过压力传感器感受微压差信号。采集到信号后, 在电子设备中进行分析、合成、变换、滤波等一系列处理, 继而采用电信号控制微电机等执行元件, 控制其作动, 从而实现调节的目的。

与机械式氧气调节器相比, 电子式氧气调节器的优点主要体现在以下几个方面:

1) 结构简便。机械肺式机构结构复杂, 膜片产生的驱动力需要通过杠杆等机构放大和改变方向, 产品体积和质量大, 调节精度差, 产品之间性能一致性较差, 调试工作复杂, 故障模式多。电子肺式机构由电机产生驱动力, 直接控制肺式活门, 结构简便、响应快、控制精度高。

2) 加压控制电子化。电子式氧气调节器根据高度传感器信号, 在软件中动态计算加压值, 能够实现高空供氧加压值的精准控制。

3) 灵活的控制策略。机械式氧气调节器的控制特性在结构设计完成后就已固定, 在不同的工作条件下存在一定的性能损失。此外, 由于影响控制特性的参数较多, 且存在复杂的耦合关系, 控制特性调整较为困难。电子式氧气调节器控制特性可直接通过控制律参数进行快速调整, 能够适应不同的工作条件。

4) 数据采集与显示。机械式氧气调节器只能通过压力表和示流器显示当前状态, 电子式氧气调节器依靠传感器进行参数反馈, 可以将系统状态参数集中反映在电控面板上并全程记录。若出现故障, 飞行过程参数还可作为故障诊断的重要依据。

采用电子设备来控制供氧量主要有 4 个控制节点: ①在 8 000 m 以下, 当空气活门打开时供氧; ②在 8 000～12 000 m, 空气活门关闭后供应纯氧; ③在 12 000 m 以上, 按生理需求进行加压供氧; ④过载飞行时, 过载超过 3g～3.5g 便启动加压呼吸, 对抗过载导致的脑缺血缺氧。常规的缺氧、过载和低气压防护按照既定的压力曲线进行, 智能化防护以现有供气制度为参考, 根据飞行员防护需求进行实时调整。

对于电子式调节设备来说, 中央控制器对采集信息进行分析处理, 并输出控制信号驱动执行机构, 从而实现供氧调节。因此, 对数据的采集直接影响到后续分析与控制。

5.2　智能供氧控制律

控制律是电控肺式活门和供氧抗荷活门等执行机构的控制指令, 是电子式氧气调节器实现供氧调节的关键。智能供氧控制律的设计应在现有电子式氧气调节器控制逻辑的基础上, 引入飞行员生理状态感知信号, 形成具有“智能预测、主动防护”的多信号协同控制规律, 基于生理和性能指标要求精确控制氧气体积分数和流量, 使系统满足安全余压供氧、高空加压供氧和抗荷加压供氧等不同工作模式的需要, 如图 7 所示。

电子式氧气调节器的最佳控制目标是: 供氧流量能够满足肺通气量的需求, 且能随肺通气量的变化而改变; 供氧压力和体积分数能够随座舱高度与生理状态的变化而实时调整; 避免出现灌气或呼吸困难的现象, 尽量不增加飞行员的呼吸负荷。

针对智能电子式供氧控制系统而言, 基本输入包括呼吸腔压力、飞行过载、抗荷服压力、座舱高度和生理信号。如果考虑到飞机的快速爬升和下降等机动过程, 座舱高度的变化率也是一个需要考虑的输入参数; 输出包括电控肺式机构电机作动信号、电控抗荷活门电机作动信号; 判断依据为不同座舱环境下输入信号与生理设定值的比较。

6　结 束 语

满足飞行员生理需求是生命保障系统设计的核心, 然而, 在严酷的飞行环境中, 飞行员的安全性和舒适性面临诸多挑战。传统座舱往往缺乏对飞行员生理状态的实时监测, 无法及时预测可能出现的飞行隐患并采取干预措施。随着未来战场环境的复杂化, 飞行员的操纵负担显著增加, 强烈的压力容易影响其生理和心理状态, 从而限制飞行操控决策的准确性, 甚至可能导致失能或丧失操纵能力等严重后果。

人体生理系统是一个动态且相互关联的复杂耦合系统, 而供氧防护系统和抗荷防护系统的研发通常分别基于呼吸生理学及血流动力学, 且彼此独立开展, 实际上, 过载和缺氧对呼吸系统和心血管系统的影响相互耦合。因此, 为满足高性能战斗机对飞行员日益增长的生命保障需求, 先进战斗机的机载氧气系统在研发过程中必须重视飞行员的生理特征, 实现对飞行员生理状态的准确感知, 将飞行员生理状态纳入控制闭环。

综合国内外研究, 智能氧气系统应是未来氧气系统的发展方向, 目前, 现有机载氧气系统及相关研究存在以下不足:

1) 战斗机飞行员依赖氧气系统实现持续正常呼吸, 通气参数与人体生理效应的匹配是氧气系统设计的基础;

2) 传统座舱存在飞行员生理监测缺失的问题,

缺乏有效的生理指标监测体系，使飞行安全仍在较大程度上依赖飞行员的自主感知和处置。实时监测飞行员的生理状态，及时预测可能带来的飞行隐患并及时干预，是智能氧气系统的基本要求；

3）人体各生理系统动态关联、相互作用，构成复杂耦合系统，然而，供氧防护系统研发应针对过载和缺氧提供可靠防护，确保呼吸系统和心血管系统功能正常并协调运行。

将飞行员生理引入闭环控制，具有生理监测、预测和干预功能的智能氧气系统是下一代生命保障系统的发展方向。实现智能供氧需要具备以下3个关键技术：飞行员生理监测技术、生理状态预测技术和电子式氧气调节技术。通过飞行员生理监测系统采集关键生理数据，通过智能算法识别并预测生理状态，将生理信号引入氧气系统控制闭环，依照实时生理需求和飞机状况自动调节氧气供给，以实现人体状态、座舱环境和供氧设备3个组成部分的协同联系。

目前，国外已围绕智能机载氧气集成系统开展了大量理论研究、上机测试与型号迭代，整体研究进度领先于中国，中国这一领域仍处于初步阶段，极其缺乏必要的理论与实验系统研究。聚焦到具体技术，国内外均已开展生理监测原理机的设计及试验，但还未形成生理指标选取与生理状态评价方法的“金标准”；针对飞行活动中的生理状态预测技术研究较少，而其他场景下的时序预测研究为飞行员生理预测提供了参考；智能供氧控制律是联系飞行员与氧气系统的核心，构建具有“智能预测、主动防护”能力的多信号协同供氧控制律是未来的重要研究课题。

参考文献（References）

[1] 肖华军. 航空供氧防护装备生理学[M]. 北京: 军事医学科学出版社, 2005.
Xiao H J. Physiology of aviation oxygen protective equipment[M]. Beijing: Military Medical Science Press, 2005(in Chinese).

[2] 胡文东, 文治洪. 航空航天医学全书: 航空航天医学工程学[M]. 西安: 第四军医大学出版社, 2013: 1-17.
Hu W D, Wen Z H. Aerospace medicine books: aerospace medical engineering[M]. Xian: Fourth Military Medical University Press, 2013, 1-17(in Chinese).

[3] Cragg C H, Kennedy K D, Shelton M B, et al. Understanding pilot breathing—a case study in systems engineering: NASA NESC-RP-18-01320[R]. Washington, D. C.: NASA, 2021: 10-28.

[4] 封文春, 朱永峰, 周慧红. F-22缺氧问题的分析与思考[J]. 航空科学技术, 2013, 24(1): 23-26.
Feng W C, Zhu Y F, Zhou H H. Analysis and thinking of the problem of F-22 oxygen system[J]. Aeronautical Science & Technology, 2013, 24(1): 23-26(in Chinese).

[5] Commander, U. S. Pacific Fleet. Comprehensive review physiological episodes of T-45 and F/A-18 aircraft[R]. Pearl Harbor: U. S. Navy, 2017: 1-56.

[6] 孙喜庆, 姜世忠. 航空航天生物动力学[M]. 西安: 第四军医大学出版社, 2013: 26-31.
Sun X Q, Jiang S Z. Aerospace biodynamics[M]. Xian: Fourth Military Medical University Press, 2013, 26-31(in Chinese).

[7] 耿喜臣, 颜桂定, 金朝. 航空加速度生理学的研究与应用[J]. 航空军医, 2004, 32(5): 189-196.
Geng X C, Yan G D, Jin Z. Research and application of aviation acceleration physiology[J]. Flight Surgeon, 2004, 32(5): 189-196(in Chinese).

[8] Lyons T J, Harding R, Freeman J, et al. G-induced loss of consciousness accidents: USAF experience 1982—1990[J]. Aviation Space and Environmental Medicine, 1992, 63(1): 60-66.

[9] 刘天亚. 与呼吸机相关性肺损伤的预防相关的呼吸力学问题的研究[D]. 北京: 北京航空航天大学, 2019: 1-28.
Liu T Y. Studies of respiratory mechanics on prevention of ventilator-associated lung injury[D]. Beijing: Beihang University, 2019: 1-28 (in Chinese).

[10] Masters M, Schulte A. Physiological sensor fusion for real-time pilot workload prediction and overload prevention[C]//Proceedings of the AIAA SCITECH 2022 Forum. Reston: AIAA, 2022: 2344.

[11] Endsley M R. Design and evaluation for situation awareness enhancement[C]//Proceedings of the Human Factors Society 32nd Annual Meeting. Santa Monica: Human Factors Society, 1988, 32: 97-101.

[12] NASA Engineering and Safety Center. Pilot breathing assessment volume I: NESC-RP-18-01320, Vol. 1[R]. Washington, D. C.: NASA, 2020: 27-519.

[13] NASA Engineering and Safety Center. Pilot breathing assessment volume II: NESC-RP-18-01320, Vol. 2[R]. Washington, D. C.: NASA, 2020: 1-260.

[14] 潘锐. 电子式氧调器设计及实验研究[D]. 北京: 北京航空航天大学, 2021.
Pan R. Design and experiment research of the electronic oxygen regulator[D]. Beijing: Beihang University, 2021(in Chinese).

[15] INSTA. Austria selects Insta pilot’s breath air monitor for its Eurofighter fleet[EB/OL]. (2023-09-02)[2024-08-20]. https://defence-industry.eu.

[16] Shaw D M, Harrell J W. Integrating physiological monitoring systems in military aviation: a brief narrative review of its importance, opportunities, and risks[J]. Ergonomics, 2023, 66(12): 2242-2254.

[17] Spotlight Labs. Launches hypoxia sensor device for USAF pilots[EB/OL]. (2019-01-30)[2024-08-20]. https://www.airforce-technology.com.

[18] 盛勇. 电子式供氧抗荷调节器的仿真研究[D]. 北京: 北京航空航天大学, 2011.
Sheng Y. Simulation for the electronic oxygen supply and anti-G regulator[D]. Beijing: Beihang University, 2011(in Chinese).

[19] 孙灿飞, 蔡元友, 龙海军. 电子式氧气调节器中步进电机模糊控制技术研究[J]. 测控技术, 2013, 32(4): 78-81.
Sun C F, Cai Y Y, Long H J. Research on stepping motor fuzzy control technology application in the aircraft electronic oxygen regulator[J]. Measurement & Control Technology, 2013, 32(4): 78-81 (in Chinese).

[20] 吕晓东, 耿斌, 郭云, 等. 飞行头盔生理参数检测装置: CN102657523A[P]. 2012-09-12.

Lyu X D, Geng B, Guo Y, et al. Physiological parameter detection device for flying helmet: CN102657523A[P]. 2012-09-12(in Chinese).

[21] Ryoo H C, Hrebien L, Shender B S. Noninvasive monitoring of human consciousness by near-infrared spectroscopy (NIRS) during high +Gz stress[J]. Biomedical Sciences Instrumentation, 2002, 38: 1-7.

[22] Ryoo H C, Sun H H, Shender B S, et al. Consciousness monitoring using near-infrared spectroscopy (NIRS) during high +Gz exposures[J]. Medical Engineering & Physics, 2004, 26(9): 745-753.

[23] Benni P B, Li J K J, Chen B, et al. NIRS monitoring of pilots subjected to +Gz acceleration and G-induced loss of consciousness (G-LOC)[J]. Advances in Experimental Medicine and Biology, 2003, 530: 371-379.

[24] Tamura T. Current progress of photoplethysmography and SPO2 for health monitoring[J]. Biomedical Engineering Letters, 2019, 9(1): 21-36.

[25] Smith C, Goswami N, Robinson R, et al. The relationship between brain cortical activity and brain oxygenation in the prefrontal cortex during hypergravity exposure[J]. Journal of Applied Physiology, 2013, 114(7): 905-910.

[26] Eiken O, Keramidas M E, Taylor N A S, et al. Intraocular pressure and cerebral oxygenation during prolonged headward acceleration [J]. European Journal of Applied Physiology, 2017, 117(1): 61-72.

[27] Kurihara K, Kikukawa A, Kobayashi A, et al. Frontal cortical oxygenation changes during gravity-induced loss of consciousness in humans: a near-infrared spatially resolved spectroscopic study[J]. Journal of Applied Physiology, 2007, 103(4): 1326-1331.

[28] Moore T W, Jaron D. Cardiovascular model for studying circulatory impairment under acceleration[J]. IEEE Engineering in Medicine and Biology Magazine, 1991, 10(1): 37-40.

[29] Ghamari M. A review on wearable photoplethysmography sensors and their potential future applications in health care[J]. International Journal of Biosensors & Bioelectronics, 2018, 4(4): 195-202.

[30] Buick F, Wood E H, Pecaric M, et al. Methods for measuring physiological responses and protection in man exposed to high +Gz[C]//AGARD Lecture Series. Neuilly sur Seine: AGARD, 1995: 8.1-8.15.

[31] Bradke B S, Miller T A, Everman B. Photoplethysmography behind the ear outperforms electrocardiogram for cardiovascular monitoring in dynamic environments[J]. Sensors, 2021, 21(13): 4543.

[32] 国防科学技术工业委员会. 飞行员肺通气参数: GJB 306A—97[S]. 北京: 中国标准出版社, 1997.
Commission for Science, Technology and Industry for National Defense. Pilot's parameters of pulmonary ventilation: GJB 306A—97[S]. Beijing: Standards Press of China, 1997(in Chinese).

[33] 国防科学技术工业委员会. 飞行员飞行中肺通气参数: GJB 305—87[S]. 北京: 中国标准出版社, 1987.
Commission for Science, Technology and Industry for National Defense. Pilot's parameters of pulmonary ventilation during flight: GJB 305—87[S]. Beijing: Standards Press of China, 1987(in Chinese).

[34] 刘爱东, 于群. 飞行员多生理参数采集模块嵌入式设计[J]. 信息技术, 2020(2): 108-111.
Liu A D, Yu Q. Embedded design of pilot multi-physiological parameter acquisition[J]. Information Technology, 2020(2): 108-111(in Chinese).

[35] Whinnery A M, Whinnery J E, Hickman J R. High +Gz centrifuge training: the electrocardiographic response to +Gz-induced loss of consciousness[J]. Aviation, Space, and Environmental Medicine, 1990, 61(7): 609-614.

[36] Whinnery A, Whinnery J. The electrocardiographic response of females to centrifuge +Gz stress[J]. Aviation, Space and Environmental Medicine, 1990, 61(11): 1046-1051.

[37] 陆霞, 初旭, 周杰, 等. 离心机高 G 训练中的心电图改变[J]. 航天医学与医学工程, 1997, 8(1): 62-64.
Lu X, Chu X, Zhou J, et al. ECG changes during high G centrifuge training[J]. Space Medicine & Medical Engineering, 1997, 8(1): 62-64 (in Chinese).

[38] 吴斌, 薛月英, 由广兴, 等. 加速度致意识丧失前驱反应特征及预警方法研究[J]. 空军医学杂志, 2011, 27(1): 28-32.
Wu B, Xue Y Y, You G X, et al. Study on prodromic responses and method for warning acceleration induced loss of consciousness[J]. Medical Journal of Air Force, 2011, 27(1): 28-32(in Chinese).

[39] Zawadzka-Bartczak E K, Kopka L H. Cardiac arrhythmias during aerobatic flight and its simulation on a centrifuge[J]. Aviation, Space, and Environmental Medicine, 2011, 82(6): 599-603.

[40] 常巍, 彭智勇, 李川涛, 等. 高过载失能生理预警指标和告警技术研究进展[J]. 海军军医大学学报, 2022, 43(2): 194-200.
Chang W, Peng Z Y, Li C T, et al. Physiological warning indexes and alerting techniques for G-induced loss of consciousness: research progress[J]. Academic Journal of Naval Medical University, 2022, 43(2): 194-200(in Chinese).

[41] 李毅峰, 张涛, 邓略, 等. 高重力加速度及其模拟条件下脑电变化研究综述[J]. 北京生物医学工程, 2013, 32(6): 647-654.
Li Y F, Zhang T, Deng L, et al. Survey of electroencephalograph change under real and simulated high gravity acceleration[J]. Beijing Biomedical Engineering, 2013, 32(6): 647-654(in Chinese).

[42] Whinnery J E, Burton R R, Boll P A, et al. Characterization of the resulting incapacitation following unexpected +Gz-induced loss of consciousness[J]. Aviation, Space, and Environmental Medicine, 1987, 58(7): 631-636.

[43] Wilson G F, Reis G A, Tripp L D. EEG correlates of G-induced loss of consciousness[J]. Aviation, Space, and Environmental Medicine, 2005, 76(1): 19-27.

[44] 金璋瑞. 飞行活动脑电图的特点[J]. 中国应用生理学杂志, 1986(3): 66-68.
Jin Z R. Characteristics of EEG of pilots during flight[J]. Chinese Journal of Applied Physiology, 1986(3): 66-68(in Chinese).

[45] Li Y F, Yang J J, Zhang L H, et al. Research of EEG change feature under +Gz acceleration[J]. Computers in Industry, 2015, 70: 144-152.

[46] Lewis N G, Mcgovern J B, Miller J C, et al. EEG indices of G-induced loss of consciousness(G-LOC)[C]//AGARD Lecture Series. Neuilly sur Seine: AGARD, 1988: 29.

[47] Galvez-Pol A, Calvo-Merino B, Forster B. Revealing the body in the brain: an ERP method to examine sensorimotor activity during visual perception of body-related information[J]. Cortex, 2020, 125: 332-344.

[48] Han W Q, Hu W D, Dong M Q, et al. Cerebral hemodynamics and brain functional activity during lower body negative pressure[J]. Aviation, Space, and Environmental Medicine, 2009, 80(8): 698-702.

[49] 韩文强, 胡文东, 马瑞山. +Gz 作用下人体事件相关电位 P300 及工效学指标的改变[J]. 中国临床康复, 2004, 8(1): 98-100.

Han W Q, Hu W D, Ma R S. Changes of event related potential P300 and ergonomics indices by the effect of +Gz[J]. Chinese Journal of Clinical Rehabilitation, 2004, 8(1): 98-100(in Chinese).

[50] Charles R L, Nixon J. Measuring mental workload using physiological measures: a systematic review[J]. Applied Ergonomics, 2019, 74: 221-232.

[51] Ronen M, Weissbrod R, Overdyk F J, et al. Smart respiratory monitoring: clinical development and validation of the IPI™ (Integrated Pulmonary Index) algorithm[J]. Journal of Clinical Monitoring and Computing, 2017, 31(2): 435-442.

[52] Roscoe A H. Assessing pilot workload. Why measure heart rate, HRV and respiration[J]. Biological Psychology, 1992, 34(2-3): 259-287.

[53] Jorna P G A M. Heart rate and workload variations in actual and simulated flight[J]. Ergonomics, 1993, 36(9): 1043-1054.

[54] Dahlstrom N, Nahlinder S, Wilson G F, et al. Recording of psychophysiological data during aerobatic training[J]. The International Journal of Aviation Psychology, 2011, 21(2): 105-122.

[55] Scannella S, Peysakhovich V, Ehrig F, et al. Assessment of ocular and physiological metrics to discriminate flight phases in real light aircraft[J]. Human Factors, 2018, 60(7): 922-935.

[56] Dehais F, Dupres A, Di Flumeri G, et al. Monitoring pilot's cognitive fatigue with engagement features in simulated and actual flight conditions using an hybrid fNIRS-EEG passive BCI[C]//Proceedings of the 2018 IEEE International Conference on Systems, Man, and Cybernetics. Piscataway: IEEE Press, 2019: 544-549.

[57] Gray J S. The multiple factor theory of the control of respiratory ventilation[J]. Science, 1946, 103(2687): 739-744.

[58] Fincham W, Tehrani F T. On the regulation of cardiac output and cerebral blood flow[J]. Journal of Biomedical Engineering, 1983, 5(1): 73-75.

[59] 仇安琪, 白净. 人体呼吸系统数学模型[J]. 北京生物医学工程, 2000, 19(1): 6-13.
Qiu A Q, Bai J. A mathematical model of human respiratory system[J]. Beijing Biomedical Engineering, 2000, 19(1): 6-13(in Chinese).

[60] 刘大为. 临床血流动力学[M]. 北京: 人民卫生出版社, 2013: 67-80.
Liu D W. Clinical hemodynamics[M]. Beijing: People's Medical Publishing House, 2013: 67-80(in Chinese).

[61] Hoskins P R, Lawford P V, Doyle B J. Cardiovascular biomechanics[M]. Berlin: Springer, 2017.

[62] Brown A G, Shi Y B, Marzo A, et al. Accuracy vs. computational time: translating aortic simulations to the clinic[J]. Journal of Biomechanics, 2012, 45(3): 516-523.

[63] Pfaller M R, Pham J, Wilson N M, et al. On the periodicity of cardiovascular fluid dynamics simulations[J]. Annals of Biomedical Engineering, 2021, 49(12): 3574-3592.

[64] Shi Y B, Lawford P, Hose R. Review of zero-D and 1-D models of blood flow in the cardiovascular system[J]. BioMedical Engineering OnLine, 2011, 10(1): 33.

[65] Mukkamala R, Cohen R J. A forward model-based validation of cardiovascular system identification[J]. American Journal of Physiology-Heart and Circulatory Physiology, 2001, 281(6): H2714-H2730.

[66] Hughes A D, Parker K H, Davies J E. Waves in arteries: a review of wave intensity analysis in the systemic and coronary circulations[J]. Artery Research, 2008, 2(2): 51-59.

[67] Jaron D, Moore T W, Bai J. Cardiovascular responses to acceleration stress: a computer simulation[J]. Proceedings of the IEEE, 1988, 76(6): 700-707.

[68] Davis T L, Mark R G. Teaching physiology through simulation of hemodynamics[C]//Proceedings of the Computers in Cardiology. Piscataway: IEEE Press, 2002: 649-652.

[69] Frolov S V, Sindeev S V, Lischouk V A, et al. A lumped parameter model of cardiovascular system with pulsating heart for diagnostic studies[J]. Journal of Mechanics in Medicine and Biology, 2017, 17(3): 1750056.

[70] Li W P, Wang B T, Wang Y W, et al. Variations of human cerebral and ocular blood flow during exposure to multi-axial accelerations[J]. Medical & Biological Engineering & Computing, 2022, 60(2): 471-486.

[71] Snyder M F, Rideout V C. Computer simulation studies of the venous circulation[J]. IEEE Transactions on Biomedical Engineering, 1969, BME-16(4): 325-334.

[72] Li S, Kitade H, Ishida S, et al. Multiscale modeling of human cerebrovasculature: a hybrid approach using image-based geometry and a mathematical algorithm[J]. PLoS Computational Biology, 2020, 16(6): e1007943.

[73] Sedeh S N, Toghraie D. Computational hemodynamics investigation of the heat transfer of blood flow in different geometries of the patient's body on different scales[J]. Journal of Thermal Analysis and Calorimetry, 2022, 147(10): 5761-5769.

[74] Qureshi M U, Vaughan G D A, Sainsbury C, et al. Numerical simulation of blood flow and pressure drop in the pulmonary arterial and venous circulation[J]. Biomechanics and Modeling in Mechanobiology, 2014, 13(5): 1137-1154.

[75] Garber L, Khodaei S, Keshavarz-Motamed Z. The critical role of lumped parameter models in patient-specific cardiovascular simulations[J]. Archives of Computational Methods in Engineering, 2022, 29(5): 2977-3000.

[76] Caforio F, Augustin C M, Alastruey J, et al. A coupling strategy for a first 3D-1D model of the cardiovascular system to study the effects of pulse wave propagation on cardiac function[J]. Computational Mechanics, 2022, 70(4): 703-722.

[77] Li B, Wang W X, Mao B Y, et al. Long-term hemodynamic mechanism of enhanced external counter pulsation in the treatment of coronary heart disease: a geometric multiscale simulation[J]. Medical & Biological Engineering & Computing, 2019, 57(11): 2417-2433.

[78] Mccall K C, Jeraj R. Dual-component model of respiratory motion based on the periodic autoregressive moving average (periodic ARMA) method[J]. Physics in Medicine and Biology, 2007, 52(12): 3455-3466.

[79] Akino Y, Shiomi H, Sumida I, et al. Impacts of respiratory phase shifts on motion-tracking accuracy of the CyberKnife Synchrony™ Respiratory Tracking System[J]. Medical Physics, 2019, 46(9): 3757-3766.

[80] Putra D, Haas O C L, Mills J A, et al. A multiple model approach to respiratory motion prediction for real-time IGRT[J]. Physics in Medicine and Biology, 2008, 53(6): 1651-1663.

[81] Vedam S S, Keall P J, Docef A, et al. Predicting respiratory motion for four-dimensional radiotherapy[J]. Medical Physics, 2004, 31(8): 2274-2283.

[82] Dürichen R, Wissel T, Schweikard A. Controlling motion prediction errors in radiotherapy with relevance vector machines[J]. Inter-

national Journal of Computer Assisted Radiology and Surgery, 2015, 10(4): 363-371.

[83] Murphy M J, Pokhrel D. Optimization of an adaptive neural network to predict breathing[J]. Medical Physics, 2009, 36(1): 40-47.

[84] Lin H, Shi C Y, Wang B, et al. Towards real-time respiratory motion prediction based on long short-term memory neural networks[J]. Physics in Medicine & Biology, 2019, 64(8): 085010.

[85] Hong J Y, Nandi M, Jin W W, et al. Using photoplethysmography to detect real-time blood pressure changes with a calibration-free deep learning model[EB/OL]. (2024-07-03)[2024-08-20]. https://arxiv.org/abs/2407.03274.

[86] Reiss A, Indlekofer I, Schmidt P, et al. Deep PPG: large-scale heart rate estimation with convolutional neural networks[J]. Sensors, 2019, 19(14): 3079.

[87] Tan M, Peng H X, Liang X K, et al. LSTformer: long short-term transformer for real time respiratory prediction[J]. IEEE Journal of Biomedical and Health Informatics, 2022, 26(10): 5247-5257.

[88] Moody G B, Mark R G, Zoccola A, et al. Derivation of respiratory signals from multilead ECGs[J]. IEEE Computers in Cardiology, 1985, 12: 113-116.

[89] Penzel T, Mcnames J, De Chazal P, et al. Systematic comparison of different algorithms for apnoea detection based on electrocardiogram recordings[J]. Medical and Biological Engineering and Computing, 2002, 40(4): 402-407.

[90] Sharma H, Sharma K K, Bhagat O L. Respiratory rate extraction from single-lead ECG using homomorphic filtering[J]. Computers in Biology and Medicine, 2015, 59(C): 80-86.

[91] Orphanidou C. Derivation of respiration rate from ambulatory ECG and PPG using ensemble empirical mode decomposition: comparison and fusion[J]. Computers in Biology and Medicine, 2017, 81: 45-54.

[92] Lindberg L G, Ugnell H, Öberg P Å. Monitoring of respiratory and heart rates using a fibre-optic sensor[J]. Medical and Biological Engineering and Computing, 1992, 30(5): 533-537.

[93] Johansson A. Neural network for photoplethysmographic respiratory rate monitoring[J]. Medical and Biological Engineering and Computing, 2003, 41(3): 242-248.

[94] 伍雪冬. 非线性时间序列在线预测建模与仿真[M]. 北京: 国防工业出版社, 2015: 10-14.
Wu X D. Modeling and simulation of online prediction of nonlinear time series[M]. Beijing: National Defense Industry Press, 2015: 10-14 (in Chinese).

[95] Wang Y B, Yu Z B, Sivanagaraja T, et al. Fast and accurate online sequential learning of respiratory motion with random convolution nodes for radiotherapy applications[J]. Applied Soft Computing, 2020, 95: 106528.

Review of key technologies of on-board intelligent oxygen system

BU Xueqin[1,*], LIU Yiming[1], LIN Guiping[1,2], YU Jia[1], YU Kunyang[3]

(1. School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China;
2. International Innovation Institute, Beihang University, Hangzhou 311115, China;
3. 93126 Troops of the PLA, Beijing 100083, China)

Abstract: To meet the growing demand for precision protection in advanced high-performance fighter aircraft and the increasing life-support requirements of pilots, intelligent oxygen systems integrating physiological monitoring, state prediction, and active intervention are expected to become a key component of next-generation life-support systems. Accordingly, the evolution of airborne oxygen systems and recent advances in intelligent oxygen technologies worldwide were systematically reviewed. Particular attention was given to the application potential and major limitations of three enabling technologies: electronic oxygen regulation, multimodal physiological monitoring, and physiological state prediction. Current evidence indicates that electronic oxygen regulators provide the engineering foundation for incorporating physiological information into the oxygen-supply control loop. However, sensor reliability under complex flight conditions, individualized prediction, appropriate matching of intervention strategies to physiological risks, closed-loop stability, and in-flight validation remain major barriers to practical implementation. Future research should integrate multimodal physiological information, hybrid physiology-data modeling, and staged experimental validation to achieve coordinated regulation of pilot physiological status, cockpit environment, and oxygen-supply equipment.

Keywords: oxygen systems; life support systems; intelligent control; fighter aircraft; man machine systems; pilots

Received: 2024-08-23; **Accepted**: 2024-11-29; **Published Online**: 2024-12-24 14:26

URL: link.cnki.net/urlid/11.2625.V.20241224.0936.001

Foundation item: Outstanding Research Project of Shen Yuan Honors College, BUAA (230123207)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0898

仿生沟槽减阻研究现状与进展

王巍[1,2,*]，崔申奥[1]，陈家华[1]，赵浩翔[1]，张震[1]
(1. 沈阳航空航天大学 民用航空学院，沈阳 110136；2. 辽宁通用航空研究院，沈阳 110136)

摘 要：基于鲨鱼皮脊状结构的沟槽减阻技术具有简单易行、能源消耗率低等优点，受到科学家广泛关注和研究，在海洋船舶、航空航天、管道运输等领域展现出巨大潜力。基于此，对近20年来仿生沟槽减阻技术的研究现状与进展进行梳理。综述3种被广泛认可的减阻机理，补充其他减阻机理；归纳总结沟槽的布置方式及位置、无量纲尺寸、偏航角等对减阻性能的影响，进而对沟槽形状减阻设计与优化及加工工艺的研究现状进行综述；简述仿生沟槽在多个领域中的实际应用，对仿生沟槽减阻的研究进行总结与展望。

关 键 词：仿生沟槽；减阻机理；鲨鱼表面；减阻率；湍流
中图分类号：Q811; O355
文献标志码：A **文章编号**：1001-5965(2026)10-3496-13

随着全球能源危机的日益严峻，如何有效减少流动过程中的能量损耗成为科研界和工业界共同关注的焦点。在航空航天、船舶运输、管道流体输送等多个领域，物体表面的流动阻力会直接影响能源的利用效率。例如：空客A340-300飞机在飞行过程中，其表面与空气之间产生的摩擦阻力占总阻力的48%[1]；在船舶航行中，表面摩擦力约占其行驶总阻力的50%左右，而高速核潜艇的这一数值则高达70%左右；天然气管道输送过程中，泵站的动力几乎全部用于克服表面摩擦力，极大降低了能源利用效率[2]。因此，开展表面减阻研究显得至关重要。

在减阻技术的早期探索中，研究者普遍遵循着一个直观的假设，即减小物体表面的粗糙度可以降低摩擦阻力，从而实现减阻[3]。这一观点基于光滑表面与流体之间更少的接触面积和更平滑的流动路径。然而，随着科学研究的不断深入，特别是仿生沟槽减阻的突破，这一传统观点受到了挑战。1967年，乌克兰基辅水动力学研究所的摩科洛夫在研究涡屏蔽时，正式提出了“riblet”一词，并推测沟槽表面能够降低水动力阻力。1970年，Johnson的死鲨鱼水下拖曳实验发现死鲨鱼在水中的阻力远高于海豚,他将鲨鱼的高阻力归因于其粗糙的表皮。然而，Reif的研究否定了Johnson的结论，1982年，Reif[4]发现当鲨鱼快速游动时，其表皮会展开不规则的脊状结构，并推测脊状结构可以稳定边界层，使鲨鱼体表具有较好的低阻特性。1984年，Walsh等[5]简化了鲨鱼表皮的脊状结构，并开展了仿生沟槽结构在湍流减阻技术中的应用研究，研究结果表明，这种纵向排列的沟槽结构能够有效降低物体表面的流动阻力，揭示了表面微观结构对流体流动特性的重要影响。

沟槽减阻技术作为一种被动减阻方法[6]，其最显著的特点在于不依赖于任何外部能量输入，仅通过精细设计并优化物体表面的微观结构即沟槽结构，便能有效降低物体受到的流动阻力。这种独特的减阻机制不仅极大地简化了减阻系统的复杂性，还降低了实施成本，被认为是高效、广泛应用的减阻技术之一[7]。因此，本文对沟槽减阻技术的研究现状与进展进行了详细的综述，并对仿生沟槽减阻未来的研究方向进行了展望。

收稿日期：2024-12-19；**录用日期**：2025-02-14；**网络出版时间**：2025-03-10 11:29
网络出版地址：link.cnki.net/urlid/11.2625.V.20250310.0924.001
基金项目：航空科学基金(2020Z006054002)；辽宁省自然科学基金(2023-MS-243)；辽宁省教育厅项目(20240109)
* **通信作者**. E-mail：wangwei7832@163.com

引用格式：王巍，崔申奥，陈家华，等. 仿生沟槽减阻研究现状与进展[J]. 北京航空航天大学学报，2026，52(10)：3496-3508.
Wang W，Cui S A，Chen J H，et al. Research status and progress of bionic groove drag reduction[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3496-3508 (in Chinese).

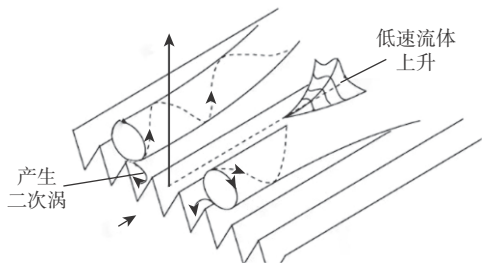
1　沟槽减阻机理

1.1　二次涡群论

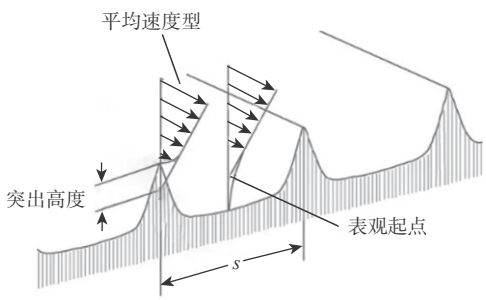
二次涡群论[8]由 Bacher 等在 1986 年提出, 纵向沟槽表面的“反向旋转涡对”与沟槽顶部相互作用, 在沟槽内部产生二次涡结构。二次涡的存在和发展能够削弱流向涡的强度, 使更多的低速流体保留在沟槽内部, 从而阻碍低速条带被抬升。二次涡可抑制流向涡的展向运动, 进一步削弱湍流的猝发频率和强度[9]。Lee 等[10]的流动可视化实验观察到沟槽结构能抬升流向涡, 且沟槽谷底存在大量的低速流体。Choi 等[11]通过数值模拟分析了减阻状态下的沟槽, 表明二次涡减少了湍流脉动引起的能量损失。Robinson[12]的风洞试验发现沟槽表面的低速条带数目明显少于光滑表面。李山等[13]通过分析沟槽表面的展向涡量云图, 证实了沟槽表面可以抑制展向涡旋运动, 从而削弱“反向旋转涡对”的强度, 减小摩擦阻力。二次涡群论示意图如图 1(a)[14]所示。

1.2　突出高度论

突出高度论[15]是 Bechert 等基于黏性理论及前人研究在 1989 年提出的较为完善的纵向沟槽减阻机理[16]。突出高度指沟槽尖端到沟槽等价光滑表面(各个质点的速度梯度为 0 的表面)的距离[17], 如图 1(b)[15]所示。沟槽将流体底层的低速条带保留在沟槽底部, 相当于增厚了黏性底层, 从而使壁面处剪切应力降低, 实现减阻。李新华等[18]认为减阻沟槽具备使湍流动能“活跃”的流体趋于“安静”的能力。Bushnell 等[19]证实了 V 形沟槽的谷底保留了低速流体, 减小了黏性阻力。Ozkan 等[20]借助粒子图像测速法测量了沟槽流场的速度分布, 发现当流动阻力减小时, 黏性底层厚度增加, Cui 等[21]对沟槽表面湍流边界层特性和黎润恒等[22]对沟槽面圆管湍流减阻的研究也得出了相同结论。



(a) 二次涡群论[14]



(b) 突出高度论[15]

图 1　沟槽减阻机理

Fig. 1　Drag reduction mechanism of groove

1.3　微型空气轴承论

横向沟槽的减阻机理主要为微型空气轴承论, 在 1996 年由潘家正[23]提出, 在湍流边界层布置有间隔的横向沟槽, 流体在沟槽内部产生涡旋, 该涡旋被理解为“微空气轴承”, 使流体与沟槽表面的滑动摩擦转变为滚动摩擦, 同时, 涡旋也会产生反向剪切应力, 从而达到减小摩擦阻力的效果[23]。吴正人等[24]分析了沟槽对风机运行特性的影响, 表明沟槽谷底保留的低速气体形成了二次涡, 其作用类似于“滚动轴承”, 使近壁区的湍流动能明显下降, 减缓了湍流的耗散。Wang[25]、Ding[26]等通过数值模拟验证了“微型空气轴承”减阻机理。

研究人员通过大量的相关实验和分析, 解释了沟槽减阻的部分现象[27]。Rastegari 等[28]在 2015 年提出滑移速度论, 利用壁面摩擦阻力与湍流统计量之间的定量关系式, 引入了滑移速度这一变量。在修正后的公式中, 滑移速度的大小与壁面摩擦阻力呈负相关。即随着滑移速度的增大, 壁面摩擦阻力逐渐减小。但该理论还存在一些局限性, 需要在未来的研究中进一步完善壁面滑移速度与沟槽尺寸的关系模型。

1.4　减阻机理研究未来发展趋势

二维沟槽的减阻机理研究足够成熟, 但在部分情况下并不适用于三维沟槽[29]。这意味着在三维沟槽的应用场景中, 无法简单照搬二维沟槽的减阻理论来解释和优化其减阻性能。三维沟槽因具有更复杂的空间结构和流体流动特性, 与二维沟槽有着本质区别, 且三维沟槽的减阻机理并未达成普遍共识[30]。鉴于此, 有必要加大对三维沟槽减阻机理的研究力度, 深入剖析流体与沟槽表面的相互作用机制, 揭示其独特的减阻规律, 为三维沟槽在航空航天、海洋船舶、管道运输等领域的高效应用提供坚实的理论支撑。

2　沟槽减阻特性

2.1　仿生沟槽结构

仿生沟槽是受鲨鱼皮启发而研究出的一种减阻技术, 仿生沟槽减阻研究路线如图 2 所示。20 世纪 70 年代, 美国兰利研究中心发现快速流动的鲨鱼

表皮存在着排列规则的脊状结构，鲨鱼皮脊状结构的长度通常为100～200 μm，宽度为50～100 μm[31]，能够优化鲨鱼体表流体边界层的流动状态，抑制和延迟湍流的发生，其研究团队对仿生沟槽结构进行了大量理论及实验研究，表明仿生沟槽能够降低表面阻力。鲨鱼皮脊状结构如图3[32]所示。

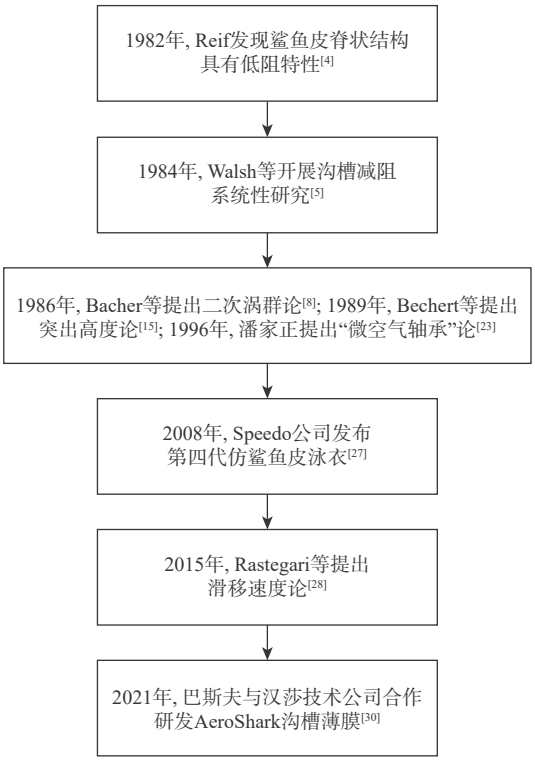


图2 仿生沟槽减阻研究路线

Fig. 2 Bionic groove drag reduction research roadmap

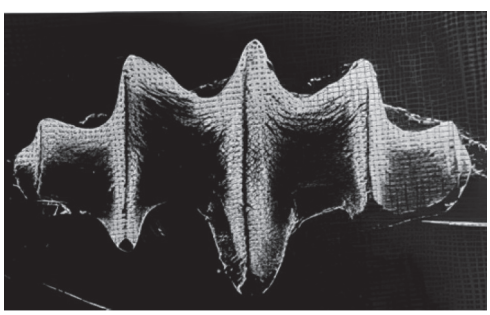


图3 鲨鱼皮脊状结构[32]

Fig. 3 Shark skin ridge structure[32]

鲨鱼皮脊状结构较为复杂，为便于研究，研究人员将鲨鱼皮脊状结构简化为V形、梯形、矩形等二维沟槽[33]。沟槽的几何尺寸通常采用Walsh提出的无量纲参数，主要考虑沟槽的高度 h、宽度 s 和间距 t[34]。其中，关于V形沟槽的研究较多，Walsh教授首次证明了V形沟槽减阻效果最好，并指出V形沟槽能够限制沟谷区域的湍流脉动、减少流体微团间的动量交换，从而降低黏性阻力[35]。但目前沟槽最佳减阻外形存在争议，李永成等[36]的数值模拟表明，雷诺数 Re=5×10^5 时，V形沟槽的剪切应力明显小于同尺寸的U形和矩形沟槽，因此，减阻性能最佳。丛茜等[37]采用有限体积法分析了V形、半圆形和矩形3种沟槽的减阻性能，雷诺数 Re=6×10^5 时，减阻率分别为3.2%、9.1%、9.7%，通过分析三者雷诺切应力得出，V沟槽表面所受雷诺切应力最大，因此，横向脉动最大，对流向脉动动能的传递功能最强，流体对壁面的作用力最大，壁面摩擦阻力最大，减阻效果最差；而矩形沟槽表面雷诺切应力最小，同理可得流体对壁面的作用力最小，壁面总阻力最小，减阻效果最好。杜淑雅等[38]基于大涡模拟计算流体动力学方法，研究了V形与梯形沟槽的减阻特性，雷诺数 Re=5 600 时，通过分析2种沟槽的阻力特性得出，V形沟槽顶部附近的阻力特性决定了沟槽减阻效果。而梯形沟槽则相反，主要是沟槽底部附近的阻力特性决定了沟槽减阻效果。研究结果显示，V形沟槽仅底部减阻效果优于梯形沟槽，梯形沟槽全局减阻效果更好。简化二维沟槽示意图如图4[5]所示。

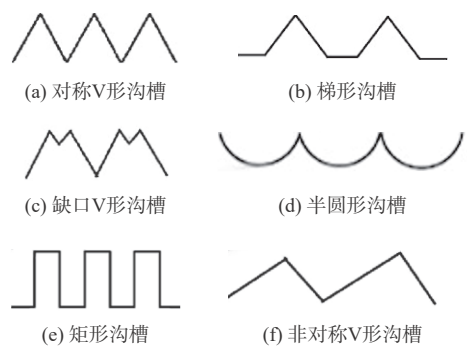


图4 二维沟槽[5]

Fig. 4 Two-dimensional groove[5]

2.2 沟槽布置方式和位置减阻特性

沟槽布置方式包括顺流向的纵向沟槽[39]和垂直流向的横向沟槽[40]。纵向沟槽受鲨鱼皮脊状结构启发[41]，海豚的真皮嵴(如图5所示)也为横向沟槽提供了生物原型[42]。梁在潮[43]通过沟槽表面边界层的流动显示实验得出，纵向沟槽通过改变低速条带的数量和间距控制减阻效果，横向沟槽通过改变低速条带的长度控制减阻效果。Sharma等[44]在等比缩小的鱼雷模型上布置纵向沟槽，采用动量平衡法确定鱼雷阻力，实验结果显示，在4 m/s来流速度下，沟槽可有效降低阻力系数达10.5%。Lunghi等[45]探索了沟槽结构船尾减阻的潜力，大涡模拟结果显示，雷诺数 Re=9.6×10^4 时，多个横向沟槽的存在使船尾阻力降低了23.2%。对于纵向沟槽而言，雷诺数对减阻效果影响不大，在不同的雷诺数下，纵向沟槽都能在一定程度上降低壁面摩擦阻力，保持较为稳定的减阻性能。

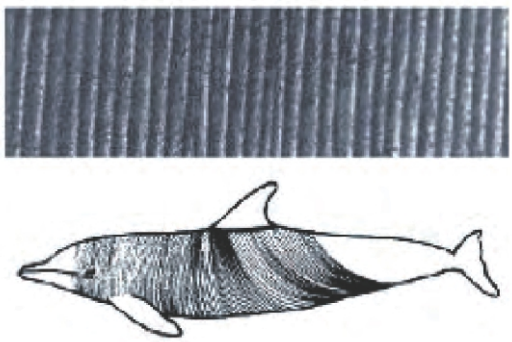
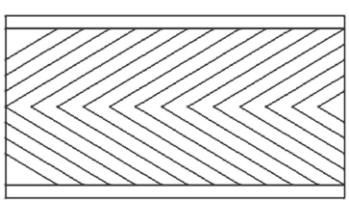


图 5　海豚真皮嵴结构[42]

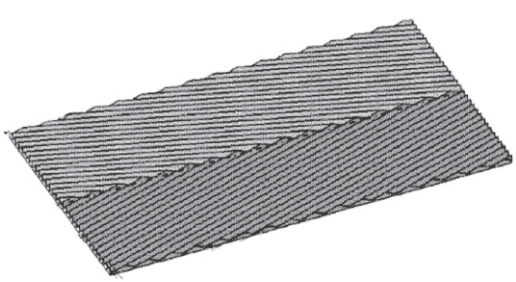
Fig. 5　Dolphin dermal ridge structure[42]

收敛沟槽指流体在沟槽中流动时, 流道逐渐变窄, 使流体有汇聚、集中的趋势; 发散沟槽则相反, 沟槽流道逐渐变宽, 流体在其中流动时会逐渐分散, 流体有扩散、分散的趋势。Chen 等[46] 通过电镜观察鸽子次级羽毛的微观结构, 提出了一种具有光滑边缘的发散沟槽, 如图 6(a) 所示, 将其与传统纵向沟槽减阻效果对比, 实验结果表明, 这种沟槽的减阻率可达 16%, 高于传统纵向沟槽。崔光耀等[47] 利用图像测速技术, 研究了收敛和发散沟槽表面的湍流边界层流动特性, 结果表明, 发散沟槽可降低摩擦阻力、雷诺应力, 具有较好的减阻能力, 而收敛沟槽 (如图 6(b) 所示) 则促进了流向大尺度相干结构的形成, 不利于减阻。

沟槽布置的位置也会对减阻产生影响, 戎瑞等[48] 分别在翼型的前段和后段布置沟槽, 翼型涡量云图如图 7 所示, 翼型前段布置沟槽时, 翼型边界层分离提前发生, 流动分离也提前结束, 略微减弱分离程度。翼型后段布置沟槽时, 不仅推迟了边界层分离点, 而且提前结束边界层的分离区域, 使流体在翼型表面出现较小的分离区域, 减小阻力系数。



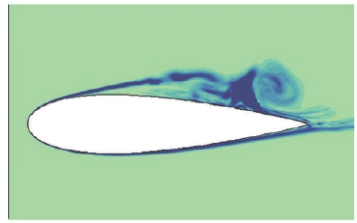
(a) 发散沟槽[46]



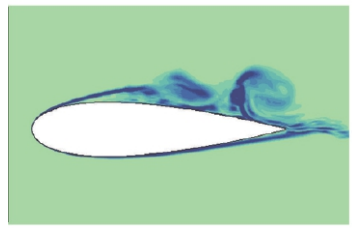
(b) 收敛沟槽[47]

图 6　发散与收敛沟槽

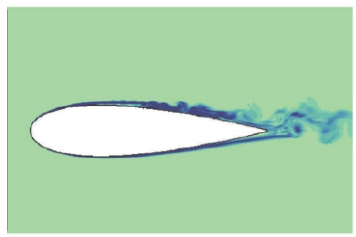
Fig. 6　Convergence and divergence grooves



(a) 光滑翼型



(b) 前段布置沟槽



(c) 后段布置沟槽

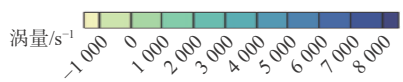


图 7　来流速度为 12 m/s 时翼型涡量云图[48]

Fig. 7　Vorticity contours for the airfoil at a freestream velocity of 12 m/s[48]

Guo 等[49] 在翼型吸力面布置沟槽, 实验结果表明, 布置沟槽翼型的阻力系数均小于光滑翼型, 且在翼型尾缘布置沟槽的减阻效果优于前缘。张子良等[50] 对翼型湍流区域沟槽的减阻性能展开数值模拟, 结果表明, 沟槽布置在压力面和吸力面时, 阻力系数最大可降低 3.07%。沟槽仅布置在吸力面时, 阻力系数最大可降低 2.55%, 两者情况下翼型的压差阻力系数和黏性阻力系数均有减小。而沟槽仅布置在压力面时, 几乎不产生减阻效果, 在有些迎角下甚至出现了压差阻力系数增加的现象。基于上述流动分析表明, 沟槽减阻是由于边界层湍流转捩, 导致分离点延迟引起的, 流动分离是产生压差阻力的重要原因之一, 沟槽减阻的一个重要途径就是通过影响流动分离来减小压差阻力。

2.3　沟槽高度与宽度减阻特性

美国国家航空航天局 (National Aeronautics and Space Administration, NASA) 的 Walsh 等[5] 针对简化的纵向沟槽进行了系统性研究, 发现沟槽的高度和宽度对减阻效果存在显著影响, V 形沟槽的无量纲高度 $h^+ \leqslant 25$ 和无量纲宽度 $s^+ \leqslant 30$ 时具有减阻功能, 在 $s^+=15$ 和 $h^+=10$ 时取得了 8% 的减阻。Li 等[51]

的平板沟槽减阻研究发现，沟槽的减阻率随深度的增加先增大后减小；在每个雷诺数下，存在最大减阻的最优槽深和零减阻率对应的最大槽深。袁一平等[52]在风力机翼型吸力面布置沟槽进行数值模拟，结果表明，h^+和s^+为 10～18.5 时，最大减阻率可达 11.934%。El-Samni 等[53]研究得出，沟槽s^+为 18 时可取得最佳减阻率为 11%，当s^+大于 30 时会导致阻力增加。沟槽减阻研究选取的无量纲尺寸多在 10～30 之间，沟槽的高度要保证沟槽尖端处于湍流边界层黏性底层和过渡区的下部，才能有效改变流动。

徐琰等[54]指出在设计微沟槽时，最佳h^+为 13.86，当宽高比高于 1.5 时，沟槽内部会产生局部高强度湍流区域，导致减阻效果减弱。郭飞飞等[55]的数值模拟表明，V 型沟槽的宽高比为 1 时减阻效果最佳，这与黄桥高等[56]得出的结论相同。Quintavalla 等[57]的沟槽圆柱绕流实验发现，宽高比小的沟槽同样可以减少流动阻力，并表明，通过改变沟槽的宽度、高度并减小沟槽总面积对降低临界雷诺数有显著的效果，与最小阻力系数也有较强的对应关系。Cui 等[58]的数值模拟表明，宽高比小的半圆形沟槽适用于低雷诺数的流动，宽高比大的半圆形沟槽适用于高雷诺数的流动。谭惠丰等[59]指出沟槽的宽高比对减阻效果有显著影响，在适合的宽高比下，沟槽能够形成充满整个沟槽的规则涡提高减阻率，不同宽高比沟槽的内部流线形状如图 8 所示。

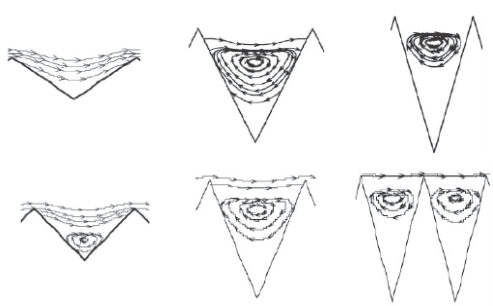


图 8 不同宽高比沟槽内部流线形状[59]

Fig. 8 Internal streamline shapes of grooves with different aspect ratios[59]

沟槽高度与宽度对减阻的影响主要通过改变涡旋结构，合适的沟槽高度与宽度可以使流体在沟槽内形成稳定的小尺度涡旋，这些涡旋能够与主流中的大尺度涡旋相互作用，打乱大涡结构，降低涡旋的强度和尺度，进而减小阻力。

2.4 沟槽间距减阻特性

针对沟槽间距，Mohammadi 等[60]对层流状态下的 4 种沟槽进行数值分析，结果表明，每一类沟槽都存在一个最优的沟槽间距，从而获得最大的减阻效果。刘占一等[61]指出有间距的沟槽表面减阻效果优于无间距的沟槽表面，与无间距沟槽表面相比，有间距的沟槽表面能够抑制压差阻力的产生，当间距与沟槽尺寸相当时，减阻效果最佳。潘光等[62]的水下回转体减阻研究表明，与光滑表面相比，有间距的沟槽表面所受的压差阻力略有增加，但摩擦阻力显著降低，并指出当沟槽内旋涡的最大速度临近外流场流速时，能够获得最佳的减阻效果。Martin 等[63]对矩形沟槽的流动分析表明，当沟槽间距变化范围为t^+= 0.08～0.50 时，减阻率基本保持不变，约为 9%。在较大的间距下，矩形沟槽波谷宽度减小，波峰处的表面积增大，这些变化将产生较弱的减阻效果，并最终导致阻力增加。窦如宏[64]利用数值仿真得出，沟槽间距影响沟槽压差阻力和黏性阻力在总阻力中的占比，沟槽间距过大会导致黏性阻力增加量大于压差阻力减小量，反而会出现增阻的现象。陈璠等[65]发现沟槽间距的差异对横向沟槽的影响较大，增大间距沟槽，导致沿程压力损失增加，适用于相对高雷诺数流动。Qiu 等[66]指出在特定的沟槽截面形状下，沟槽间距的变化在小尺度范围 (t^+ = 0.08～0.50) 内对阻力的影响并不显著。然而，在较大尺度范围 (t^+约为 3.5～24.3) 内，沟槽间距的变化会引起减阻性能的显著变化。Wu 等[67]发现沟槽间距过小会导致涡相互作用，增加阻力。因此，推测最优沟槽间距可以实现表面剪切应力最小且沟槽内涡旋互不影响。

沟槽间距对减阻的影响主要体现在对流向涡的抬升上，合适的沟槽间距将抬升流向涡，通过阻止较大尺度的湍流与沟槽表面的大部分区域相互作用，从而减小沟槽表面的摩擦阻力。不同沟槽间距对流向涡的抬升程度如图 9[66]所示。

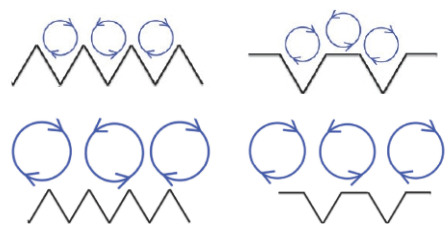


图 9 不同沟槽间距对流向涡的抬升程度[66]

Fig. 9 Lifting degree of different groove spacing on streamwise vortex[66]

2.5 沟槽偏航角减阻特性

偏航角为沟槽与自由流水平方向的夹角。1984 年，Walsh 等[5]发现当偏航角β小于 15 °时，平板沟槽的减阻效果几乎不受不影响，而当偏航角大于 30 °时，沟槽出现增阻现象，图 10 为不同偏航角下 V 形沟槽的减阻特性[5]。Zhang 等[68]研究了在不同偏航角下翼展沟槽表面的减阻效果，结果表明，在不同

偏航角下, 沟槽诱导的摩擦阻力减小趋势相似, 但沟槽表面引起的压差阻力增加趋势较为复杂, 表明不同偏航角下的减阻率并不是线性变化。王晋军等[69]对沟槽面三角翼的阻力特性进行了研究, 发现在机翼失速后, 偏航角对沟槽减阻效果几乎不产生影响。Li 等[70]的大涡模拟研究发现, 随着偏航角的增大, 沟槽降低黏性阻力的能力增强, 然而, 偏航角的增大会增加横向流动, 导致沟槽迎风面出现高压区, 背风面出现低压区, 从而增大压差阻力, 不同偏航角下的压力系数分布如图 11[70] 所示。

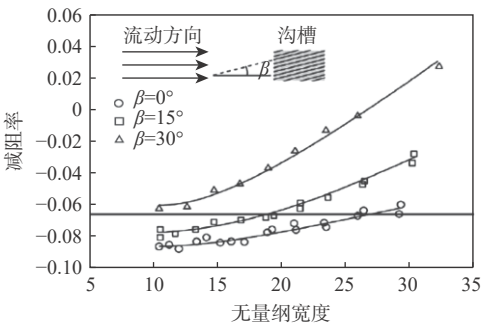
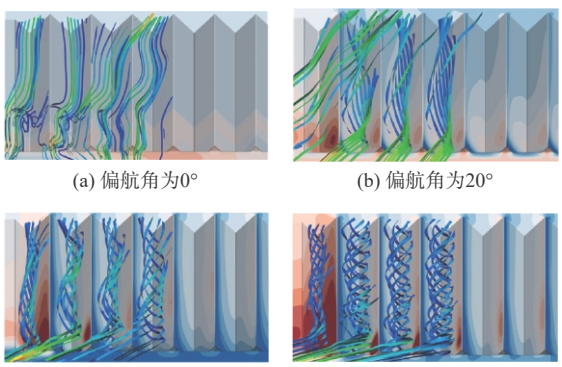


图 10　几何参数 h/s≈0.62 时不同偏航角下 V 形沟槽的减阻率[5]

Fig. 10　Drag reduction rate of V-groove under different yaw angles when geometric parameter h/s≈0.62[5]



(a) 偏航角为0°　(b) 偏航角为20°

(c) 偏航角为30°　(d) 偏航角为40°

图 11　不同偏航角下压力系数分布[70]

Fig. 11　Pressure coefficient distribution under different yaw angles[70]

偏航角对沟槽减阻效果的影响主要体现在摩擦阻力和压差阻力的动态平衡上。小偏航角下, 沟槽能有效降低摩擦阻力且压差阻力增加较小, 减阻效果显著; 但随着偏航角增大, 压差阻力的增加逐渐甚至超过摩擦阻力的降低, 导致减阻效果下降甚至出现增阻。

3　沟槽形状设计与优化

3.1　二维沟槽形状设计

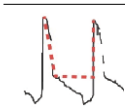
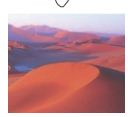
沟槽源于仿生学[71], 自然界生物的减阻方式可以作为沟槽设计的灵感来源[72]。减阻沟槽的设计已从早期对鲨鱼皮脊状结构的简单模仿, 发展为广泛借鉴具备减阻特性的生物体表结构。Wang 等[73]利用白光干涉仪扫描鸽子羽毛的表面形貌, 根据羽毛表面发现的不对称微结构, 设计出类羽毛结构的梯形沟槽, 通过风洞试验得出, 该沟槽适用的速度范围更广, 平板模型在 30～75 m/s 的速度范围内均取得不错的减阻效果 (在 60 m/s 速度下减阻率约为 19%)。Zheng 等[74]根据虎鲸皮脊状结构确定 11 个横向间距相同的插值点, 利用数值计算软件通过 3 次样条插值法生成仿生沟槽曲线, 在 2～12 m/s 的速度范围内, 沟槽表面具有显著的减阻效果, 最大摩擦阻力减阻率可达 26.91%, 总减阻率可达 9.63%。Song 等[75]根据新月形沙丘提出一种沟槽表面, 通过数值模拟得出, 在流速为 10～30 m/s 的低速气流中, 该沟槽表面最多可降低 33.63% 的摩擦阻力。程拼拼等[76]发现鲨鱼皮脊状结构是多尺度的, 根据这一发现在 V 型沟槽顶部两侧增加小尺度三角形突起, 设计出一种二级沟槽, 在一定雷诺数范围内, 与初始 V 形沟槽相比, 二级 V 形沟槽的减阻率提高了 3.47%。Wang 等[77]对矩形沟槽进行了改进, 通过增加二级沟槽提高了“二次涡”的复杂程度, 使沟槽内反向流动引起的反向剪切应力增大, 表面摩擦阻力减小, 最大减阻率可达 18.76%, 而常规矩形表面的最大减阻率仅为 13.61%。二维仿生沟槽减阻对比如表 1[73-77] 所示。

3.2　三维沟槽形状设计

三维沟槽的截面形状在不同位置会有所变化, 不再局限于平面内的简单形状。He 等[78]等根据比目鱼鳞片特征设计出的仿生沟槽可实现双向减阻,

表 1　二维仿生沟槽减阻对比

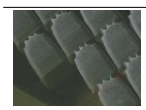
Table 1　Comparison of two-dimensional bionic groove drag reduction

形状	创新点	减阻差异	参考文献
	根据羽毛表面的不对称微结构, 设计出类羽毛结构的梯形沟槽	在30～75 m/s速度范围内, 最大减阻率为19%	[73]
	根据虎鲸皮脊状结构, 利用数值计算软件生成的仿生沟槽	在2～12 m/s的低速范围内, 总减阻率为9.63%	[74]
	根据新月形沙丘提出的沟槽表面	在10～30 m/s速度范围, 可降低33.63%的摩擦阻力	[75]
	在V型沟槽顶部两侧增加小尺度三角形突起	最大减阻率为8.07%, 可提高3.47%的减阻率	[76]
	通过增加二级沟槽提高了“二次涡”的复杂程度	最大减阻率可达18.76%, 减阻率提高5.15%	[77]

在正向流动中可降低19%表面摩擦阻力，在反向流动中减阻率可达4.2%。秦立果等[79]对鲨鱼皮的脊状结构进行简化，选用矩形沟槽作为对比，结果表明，简化脊状结构能降低近壁区域的湍流动能，减阻效果优于矩形沟槽。当入口速度为5 m/s时，简化脊状结构的减阻率可达17.86%。徐征等[80]在仿三维鲨鱼皮脊状结构的基础上，覆盖尺度更加微小的二级沟槽结构，风速为150 km/h和160 km/h的风洞试验结果表明，该沟槽能够增加“二次涡”的回旋路径，减小黏性阻力，减阻率可达10.3%。Wu等[81]制备出仿泥鳅鳞片沟槽，水洞实验结果表明，在低速条件下与光滑表面相比，仿生表面的减阻率在9.42%～17.25%之间，且该仿生表面可以改善疏水性，能进一步提高减阻效果。三维仿生沟槽减阻对比如表2[78-81]所示。

表2 三维仿生沟槽减阻对比

Table 2 Comparison of three-dimensional bionic groove drag reduction

形状	创新点	减阻差异	参考文献
	仿比目鱼鳞片沟槽，可实现双向减阻	正向流动可降低19%表面摩擦阻力，反向流动减阻率为4.2%	[78]
	简化三维鲨鱼皮脊状结构	当入口速度为5 m/s时，减阻率可达17.86%	[79]
	三维脊状结构增加二级沟槽	最大减阻率为10.3%	[80]
	仿泥鳅鳞片	低速减阻率为9.42%～17.25%，且具有疏水性	[81]

3.3 沟槽形状优化

常规沟槽形状优化基本采用尺寸参数(如深度、宽度、间距)优化规则形状，通过计算流体动力学逐个验证预设形状的减阻效果，依赖经验或试错法，易陷入局部最优解。而利用算法优化可推动沟槽形状优化从“经验驱动”到“数据-机理双驱动”的范式转变，可实现沟槽形状最优解的全局搜索。Abdollahzadeh等[82]对比了矩形、半圆形和梯形3种沟槽的减阻性能，在确定矩形沟槽减阻效果最佳后，采用遗传算法作为优化工具，并结合流体体积法，使优化后的矩形沟槽可降低约22%的阻力。宋居正等[83]将自由变形法与遗传算法结合对V形沟槽进行优化，沟槽壁面优化前后，黏性阻力与压差阻力占比变化较小，优化后的沟槽表面减阻率提高了3.7%。潘光等[84]利用iSIGHT软件及遗传算法优化沟槽高度，通过150步迭代获得了最优解，优化后高度$h=0.017$ mm时，减阻效果最佳。李超群等[85]基于神经网络的沟槽壁面优化研究显示，通过动态调整沟槽形状，最大减阻效果可达17.41%，且能降低湍流强度和雷诺切应力。Mohammadi等[86]借助一种仅包含少数傅里叶变换的简化几何模型，能够精准地表征出适用于层流减阻的最佳沟槽形状。

3.4 沟槽形状设计与优化发展趋势

大自然中的生物表面为减阻沟槽形状的优化提供了丰富的仿生灵感，如鲨鱼皮的盾鳞结构。未来研究将更深入地探索生物表面的减阻机制，并将其应用于沟槽形状设计中。模仿生物表面的沟槽在实际应用中可能面临加工复杂性，同时，需要确保仿生形状在不同流体环境下的稳定性和减阻效果。随着研究的深入，三维减阻沟槽形状逐渐受到关注。三维形状可以更全面地控制流体在3个方向上的流动，为特殊流体传输场景提供独特的解决方案。未来，将进一步探索三维形状沟槽的设计方法和优化策略。

沟槽形状优化方面，未来可引入人工智能和机器学习技术，通过构建智能优化算法，快速筛选出最优的沟槽形状和参数组合。在实际应用中，减阻往往不是唯一的目标。多目标优化下的减阻沟槽形状设计成为必然趋势，通过建立多目标优化模型，综合考虑多个因素来寻找最优解。多目标优化模型的建立和求解过程较为复杂，需要借助先进的优化算法和计算工具；同时，如何平衡各个目标之间的权重也是一大挑战。

4 沟槽加工及应用

4.1 沟槽加工

沟槽的加工方法丰富多样，每种方法都有其独特的优势与适用场景，可分为增材制造、减材制造和等材制造3类，其优缺点如表3所示。

非金属材料在沟槽增材制造中应用广泛。Wen等[87]利用3D打印技术将仿生盾鳞嵌入到柔性

表3 加工制造方式优缺点

Table 3 Advantages and disadvantages of manufacturing methods

制造方式	优点	缺点
增材制造	设计自由度、精度较高，是实验室原型开发和小批量高性能部件制造的理想选择，尤其适用于复杂三维沟槽	耐久性与环境适应性限制了大规模推广
减材制造	适用于金属材料高精度、高强度需求场景，而且加工效率高，适合大规模生产	难以加工三维沟槽，仅适合规则形状
等材制造	适合规则沟槽的大规模生产，滚压工艺可实现沟槽表面强化与耐久性提升，且材料利用率极高，接近100%	沟槽参数改变需更换模具，复杂沟槽需定制高精度滚压轮，经济性低

薄膜上, 得到了柔性的三维仿生鲨鱼皮。实验结果显示, 该仿生鲨鱼皮可提高船桨 6.6% 的移动速度。Stenzel 等[88] 提出一种可大面积生产沟槽涂层的新工艺, 在单个步骤中就可实现涂覆应用、修饰和部分固化等流程, 涂层材料由纳米复合材料组成, 使涂层具有良好的耐磨性和风化稳定性。Zhao 等[89] 研究了一种精确复制鲨鱼皮表面的真空铸造方法, 在水箱中进行了减阻试验。实验结果表明, 鲨鱼皮复制膜具有明显的减阻作用, 最大减阻可达 18.6%。Chen 等[90] 采用紫外光固化技术将鲨鱼皮的表面形态转移到自由曲面上, 实验表明, 生物复制的沟槽高度和宽度的表面误差均小于 3%, 这表明复制精度足够高, 可以准确地传递生物表面形貌。非金属增材制造具有设计自由度与精度较高的优势, 是实验室原型开发和小批量高性能部件制造的理想选择, 尤其适用于复杂三维沟槽 (如仿鲨鱼皮脊状结构), 但其耐久性与环境适应性限制了大规模推广。

减材制造属于传统的去除材料加工方法, 主要通过激光烧蚀、磨削等工艺, 在金属表面加工沟槽。Kaakkunen 等[91] 利用激光烧蚀在机翼表面加工出微米级的梯形沟槽, 风洞试验研究表明, 沟槽表面可以快速降低壁面剪切应力。Hockauf 等[92] 提出一种多层刀具的磨削工艺, 用于制造仿鲨鱼鳞片结构。通过油道和叶栅风洞中的流体动力学研究, 证实平板沟槽结构可以减少高达 4% 的表面摩擦阻力。Huang 等[93] 结合激光加工的优点, 采用激光辅助砂带磨削加工 U 型沟槽, 利用白光干涉仪对其进行了分析, 确保误差在 10% 以内, 该工艺能显著提高表面特征的完整性。减材制造适用于金属材料高精度、高强度需求场景, 且加工效率高, 适合大规模生产, 但难以加工三维沟槽, 仅适合规则形状。

等材制造通过塑性变形而非添加或去除材料的方式加工沟槽, 属于典型的“无屑加工”, 在减阻微沟槽制造中具有独特优势, 但也存在一定的局限性。Luo 等[94] 采用溅射和光刻工艺去除鳞片背面的楔形角, 采用生物滚压复制法获得连续逼真的鲨鱼皮表面。此外, 由于光刻和轧制保证了良好的工艺效果, 使沟槽横向几乎没有成型误差, 在标尺高度方向上的成型精度可达约 90%。等材制造优势突出, 适合规则沟槽的大规模生产, 滚压工艺可实现沟槽表面强化与耐久性提升, 且材料利用率极高, 接近 100%。其缺点也很明显, 沟槽参数改变需更换模具, 复杂沟槽需定制高精度滚压轮, 经济性低。

4.2　沟槽应用

4.2.1　航空航天领域应用

仿生沟槽通常应用于飞机机翼、机身表面及发动机叶片等部位, 主要通过沟槽薄膜实现。不改变民用飞机的气动布局, 降低压差阻力的减阻效果相对有限, 这使降低飞机摩擦阻力显的更为重要, 因此, 沟槽表面在高速气流和复杂的飞行环境, 需要提供较为稳定的摩擦减阻效果。空客公司在 A320 试验机表面 70% 的区域布置沟槽薄膜, 机身的摩擦阻力降低了 6%, 可减少 1%～2% 的燃油消耗, 表明沟槽薄膜在大型客机上应用具有一定的减阻和节能效果[95]。NASA 兰利研究中心在 Learjet 喷气式飞机上布置 V 形沟槽并进行飞行试验, 在马赫数为 0.3～0.7 之间取得了 6% 的减阻效果[96]。2021 年, 巴斯夫与汉莎技术公司合作研发了一款新型沟槽减阻薄膜 AeroShark, 汉莎货运航空从 2022 年开始为其波音 777 货机机队安装该款沟槽薄膜, 预计每年可节省约 4×10^5 kg 燃油和超过 1.2×10^6 kg 二氧化碳[30]。沟槽应用如图 12 所示。

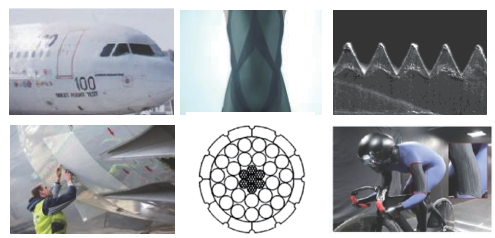


图 12　沟槽应用

Fig. 12　Groove application

4.2.2　海洋船舶领域应用

船舶在海洋中航行时除受表面摩擦阻力外, 生物污损也会增加船舶航行中受到的阻力。研究表明: 污泥会使海洋船舶的阻力增加 1%～2%; 杂草污垢能增加 10% 的阻力; 而硬壳层污垢 (如藤壶), 甚至会导致船舶阻力上升高达 40%[97]。因此, 海洋船舶所需沟槽应具有抗污减阻功能。相比快速航行, 船舶表面生物积累更容易在静态和低速条件下发生, 海洋中贝壳多数情况处于静止状态, 可以作为防污仿生对象[66]。

2008 年, Speedo 公司发布了第 4 代鲨鱼皮泳衣, 该泳衣整体性能比第 3 代提升了 5%, 拥有极轻、低阻、无皱褶等诸多优点, 第 4 代鲨鱼皮泳衣帮助比赛选手创造了 14 项世界纪录[27]。2010 年, BMW ORACLE Racing 的高速赛艇表面覆盖了由 3M 公司提供的定制乙烯沟槽薄膜, 该沟槽薄膜可以被动地为高速船舶提供减阻效果, 且不会增加质量, 极大提高了赛艇在水中移动的效率[98]。Sharma 等[44] 选用美国海军型号为 USN MK 13 (2019 年) 的鱼雷, 实验表明, 在其尾端布置仿生沟槽, 可降低 11.7% 的阻力。目前, 沟槽表面可用于赛艇、小型水下航行器和鱼雷上。对于大型海洋船舶, 可优先开发螺

旋桨局部沟槽使用，沟槽涂层是较为理想的规模化应用方法。

4.2.3 管道运输领域应用

管道运输的流体种类多样，物理性质各异。针对石油、天然气、水等不同流体，仿生沟槽需采用不同的设计策略。例如，对于高黏度的石油，沟槽设计需更注重流体的层流状态，对于低黏度的天然气，则需关注湍流控制。而且在管道中使用机器加工沟槽或铺设沟槽薄膜将造成较高的经济成本，可通过划痕技术在管道内加工沟槽，实验证实该方法可降低大于 5% 的阻力[99]。Bechert 等通过划痕技术在管道内制造沟槽表面，在一条长达 16 km 的天然气管线测试中取得了 10% 的减阻效果[15]。Bixler 等[100]在矩形管道中通过结合沟槽与纳米涂层，在湍流条件下实现最大 34% 的压降降低。Luo 等[94]将沟槽减阻技术引入天然气管道运输，进行空气和天然气的减阻实验，结果表明，减阻率均能达到 8% 以上。

4.2.4 其他领域应用

Zheng 等[101]将纵向沟槽织物应用于赛道自行车运动员的上臂，在不同速度下对骑行人体模型进行风洞试验，当流速大于 17 m/s 时，上臂的纵向沟槽织物可以降低 7% 的风阻。Eguchi 等[102]研制了一种带有半圆沟槽的新型架空电线 LP810，试验表明，新研制的电线 LP810 在实际使用中能够有效减少风阻。Wang 等[103]基于沟槽结构设计了一种仿生深松铲，进行了田间试验，结果表明，仿生结构对土壤的导向作用能够减小接触面积，破坏密封作用，减小黏附力和摩擦阻力。

4.3 沟槽加工发展趋势

从工业应用的角度来看，大规模生产具有减阻沟槽结构的部件是一个重要需求。但目前的制造工艺在大规模生产时面临成本高昂和效率低下的问题。例如，一些先进的微纳制造工艺虽然能够制造出高质量的沟槽结构，但设备昂贵、生产周期长，限制了沟槽减阻技术在大规模工业生产中的应用。3D 打印技术利于制造复杂结构，而微纳制造技术则能实现高精度制造。通过结合两者优势，开发出既能制造高精度沟槽结构又能满足复杂形状和大规模生产需求的工艺。利用 3D 打印技术构建沟槽结构的宏观框架，再用微纳制造技术进行精细加工，实现多技术融合，打破单一工艺局限，推动沟槽减阻技术的应用进展。

5 结束语

本文主要结论如下：

1）纵向沟槽减阻机理为二次涡群论与突出高度论，横向沟槽减阻机理为微型空气轴承论。

2）沟槽的布置方式及位置、无量纲尺寸、偏航角等因素对减阻性能有着不可忽视的影响，在沟槽减阻的应用与研究中，综合考虑上述因素可以获得更好的减阻效果。

3）根据低阻生物的表面特征，提出多种减阻性能良好的仿生沟槽。利用算法优化沟槽外形突破了传统试错和经验设计的局限性，能够实现对沟槽最佳减阻外形的全面搜索。

4）沟槽加工方式主要分为增材制造、减材制造和等材制造 3 类，各有优缺点。沟槽表面在多领域初步实现应用，但受限于加工成本、耐久性和性能稳定性等问题，并未实现大规模应用。

沟槽减阻研究展望如下：

1）考虑到实际应用中多物理场的相互作用，应开展多物理场耦合模拟研究。以航空发动机减阻为例，需考虑高温高速气流对沟槽结构的热应力影响及沟槽结构对气流流动和散热的影响。通过多物理场耦合模拟，全面评估沟槽结构在复杂工况下的性能。

2）建立沟槽的减阻性能退化模型，通过实验和模拟相结合的方法收集性能数据。基于这些数据建立统计学或物理机制的减阻性能退化模型，有助于优化沟槽结构设计，提高其性能稳定性。

3）为适应不同的流体环境和工况，自适应减阻沟槽形状成为研究热点，利用智能材料使沟槽形状能够自动调整。未来，随着智能材料性能的提升和控制技术的完善，有望实现更精准、快速响应的自适应沟槽形状，应用范围也将进一步拓展到更多复杂恶劣的流体环境。

参考文献（References）

[1] Ao M R, Wang M C, Zhu F L. Investigation of the turbulent drag reduction mechanism of a kind of microstructure on riblet surface[J]. Micromachines, 2021, 12(1): 59.

[2] Liu G J, Yuan Z C, Qiu Z Z, et al. A brief review of bio-inspired surface technology and application toward underwater drag reduction[J]. Ocean Engineering, 2020, 199: 106962.

[3] Luo Y H, Yuan L, Li J H, et al. Boundary layer drag reduction research hypotheses derived from bio-inspired surface and recent advanced applications[J]. Micron, 2015, 79: 59-73.

[4] Reif W E. Morphogenesis and function of the squamation in sharks[J]. Neues Jahrbuch für Geologie und Palä ontologie-Abhandlungen, 1982, 164(1-2): 172-183.

[5] Walsh M, Lindemann A. Optimization and application of riblets for turbulent drag reduction[C]//Proceedings of the 22nd Aerospace Sciences Meeting. Reston: AIAA, 1984: 347.

[6] Walsh M J. Riblets as a viscous drag reduction technique[J]. AIAA Journal, 1983, 21(4): 485-486.

[7] Sharma V, Dutta S. Experimental investigation of flow character-

istics over a surface-modified structured torpedo model[J]. Journal of Wind Engineering and Industrial Aerodynamics, 2022, 230: 105196.

[8] Bacher E V, Smith C R. Turbulent boundary-layer modification by surface riblets[J]. AIAA Journal, 1986, 24(8): 1382-1385.

[9] Wang J S, Fan Y T, Ge J N, et al. Effect of streamwise vortex induced by streamwise grooves on drag reduction and heat transfer performance[J]. International Journal of Thermal Sciences, 2023, 186: 108117.

[10] Lee S J, Lee S H. Flow field analysis of a turbulent boundary layer over a riblet surface[J]. Experiments in Fluids, 2001, 30(2): 153-166.

[11] Choi H, Moin P, Kim J. Direct numerical simulation of turbulent flow over riblets[J]. Journal of Fluid Mechanics, 1993, 255: 503-539.

[12] Robinson S K. Coherent motions in the turbulent boundary layer [J]. Annual Review of Fluid Mechanics, 1991, 23: 601-639.

[13] 李山, 杨绍琼, 姜楠. 沟槽面湍流边界层减阻的 TRPIV 测量[J]. 力学学报, 2013, 45(2): 183-192.
Li S, Yang S Q, Jiang N. Trpiv measurement of drag-reduction in the turbulent boundary layer over riblets plate[J]. Chinese Journal of Theoretical and Applied Mechanics, 2013, 45(2): 183-192(in Chinese).

[14] Bacher E, Smith C. A combined visualization-anemometry study of the turbulent drag reducing mechanisms of triangular micro-groove surface modifications[C]//Proceedings of the Shear Flow Control Conference. Reston: AIAA, 1985: 548.

[15] Bechert D W, Bartenwerfer M. The viscous flow on surfaces with longitudinal riblets[J]. Journal of Fluid Mechanics, 1989, 206: 105.

[16] Gallagher J, Thomas A. Turbulent boundary layer characteristics over streamwise grooves[C]//Proceedings of the 2nd Applied Aerodynamics Conference. Reston: AIAA, 1964: 2185.

[17] Hooshmand A, Youngs R A, Wallace J M, et al. An experimental study of changes in the structure of a turbulent boundary layer due to surface geometry changes[C]//Proceedings of the 21st Aerospace Sciences Meeting. Reston: AIAA, 1983: 230.

[18] 李新华, 董守平, 赵志勇, 等. 减阻沟槽表面近壁湍流相干运动的实验研究[J]. 山东大学学报 (理学版), 2006, 41(2): 106-110.
Li X H, Dong S P, Zhao Z Y, et al. Experiment on coherent motions in the turbulent near-wall region over a drag-reducing grooved surface[J]. Journal of Shandong University (Natural Science), 2006, 41(2): 106-110(in Chinese).

[19] Bushnell D M, Hefner J N. Viscous drag reduction in boundary layers[EB/OL]. (1991-01-01)[2024-12-09]. https://doi.org/10.2514/4.865978.

[20] Ozkan G M, Elsinga G E, Breugem W P, et al. Stereoscopic PIV measurements of flow over a riblet surface at high Reynolds number[J]. Experimental Thermal and Fluid Science, 2021, 120: 110246.

[21] Cui G Y, Pan C, Wu D, et al. Effect of drag reducing riblet surface on coherent structure in turbulent boundary layer[J]. Chinese Journal of Aeronautics, 2019, 32(11): 2433-2442.

[22] 黎润恒, 赵成璧, 唐友宏, 等. 三角形沟槽面圆管湍流减阻的大涡模拟数值研究[J]. 科学技术与工程, 2013, 13(8): 2021-2026.
Li R H, Zhao C B, Tang Y H, et al. Large eddy simulation on turbulent drag reduction over triangle riblet surfaces of pipes[J]. Science Technology and Engineering, 2013, 13(8): 2021-2026(in Chinese).

[23] 潘家正. 湍流减阻新概念的实验探索[J]. 空气动力学学报, 1996, 14(3): 304-310.
Pan J Z. The experimental approach to drag reduction of the transverse ribbons on turbulent flow[J]. Acta Aerodynamica Sinica, 1996, 14(3): 304-310(in Chinese).

[24] 吴正人, 郝晓飞, 戎瑞, 等. 离心风机叶片脊状结构减阻特性的三维数值分析[J]. 中国电机工程学报, 2014, 34(11): 1815-1821.
Wu Z R, Hao X F, Rong R, et al. Three-dimensional numerical analysis on drag reduction characteristics of the riblet structure for centrifugal fan blades[J]. Proceedings of the Chinese Society for Electrical Engineering, 2014, 34(11): 1815-1821(in Chinese).

[25] Wang J S, Zhuang Y, Liu X L. The flow and heat transfer characteristics in a rectangular channel with miniature cuboid dimples[J]. International Communications in Heat and Mass Transfer, 2021, 126: 105474.

[26] Ding L X, Shi W P, Luo H W. Numerical simulation of viscous flow over non-smooth surfaces[J]. Computers & Mathematics with Applications, 2011, 61(12): 3703-3710.

[27] Lee K H, Yu K, Al Ba'ba'a H, et al. Sharkskin-inspired magneto-active reconfigurable acoustic metamaterials[J]. Research, 2020, 2020: 4825185.

[28] Rastegari A, Akhavan R. On the mechanism of turbulent drag reduction with super-hydrophobic surfaces[J]. Journal of Fluid Mechanics, 2015, 773: R4.

[29] 刘明杰, 吴青山, 严昊, 等. 仿生减阻表面的进展与挑战[J]. 北京航空航天大学学报, 2022, 48(9): 1782-1790.
Liu M J, Wu Q S, Yan H, et al. Progress and challenges of bionic drag reduction surfaces[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1782-1790(in Chinese).

[30] Chen D K, Liu X L, Cui X X, et al. Research progress and development trend of the drag reduction inspired by fish skin[J]. Progress in Organic Coatings, 2023, 182: 107613.

[31] Jung Y C, Bhushan B. Biomimetic structures for fluid drag reduction in laminar and turbulent flows[J]. Journal of Physics: Condensed Matter, 2010, 22(3): 035104.

[32] Boomsma A, Sotiropoulos F. Direct numerical simulation of sharkskin denticles in turbulent channel flow[J]. Physics of Fluids, 2016, 28(3): 035106.

[33] Martin S, Bhushan B. Modeling and optimization of shark-inspired riblet geometries for low drag applications[J]. Journal of Colloid and Interface Science, 2016, 474: 206-215.

[34] Walsh M. Turbulent boundary layer drag reduction using riblets[C]//Proceedings of the 20th Aerospace Sciences Meeting. Reston: AIAA, 1982: 169.

[35] Endo T, Himeno R. Direct numerical simulation of turbulent flow over a compliant surface[J]. Journal of Turbulence, 2002, 3: N7.

[36] 李永成, 张华. 平板仿生沟槽表面减阻性能数值模拟研究[J]. 舰船科学技术, 2023, 45(9): 28-31.
Li Y C, Zhang H. Numerical simulation on drag reduction characteristics of bionic surface[J]. Ship Science and Technology, 2023, 45(9): 28-31(in Chinese).

[37] 丛茜, 封云, 任露泉. 仿生非光滑沟槽形状对减阻效果的影响[J]. 水动力学研究与进展 (A 辑), 2006, 21(2): 232-238.

Cong Q, Feng Y, Ren L Q. Affecting of riblets shape of non-smooth surface on drag reduction[J]. Chinese Journal of Hydrodynamics, 2006, 21(2): 232-238(in Chinese).

[38] 杜淑雅, 桑为民, 庞润. 基于数值模拟的2种条纹沟槽减阻特性对比分析[J]. 西北工业大学学报, 2022, 40(2): 261-270.
Du S Y, Sang W M, Pang R. Comparative analysis of drag reduction characteristics of two-type riblets based on numerical simulation[J]. Journal of Northwestern Polytechnical University, 2022, 40(2): 261-270(in Chinese).

[39] Huang C H, Liu D J, Wei J J. Experimental study on drag reduction performance of surfactant flow in longitudinal grooved channels[J]. Chemical Engineering Science, 2016, 152: 267-279.

[40] Wu Z R, Yang Y F, Liu M, et al. Analysis of the influence of transverse groove structure on the flow of a flat-plate surface based on Liutex parameters[J]. Engineering Applications of Computational Fluid Mechanics, 2021, 15(1): 1282-1297.

[41] Gray J. Studies in animal locomotion[J]. Journal of Experimental Biology, 1936, 13(2): 192-199.

[42] Lang A W, Jones E M, Afroz F. Separation control over a grooved surface inspired by dolphin skin[J]. Bioinspiration & Biomimetics, 2017, 12(2): 026005.

[43] 梁在潮. 工程湍流[M]. 武汉: 华中理工大学出版社, 1999: 271-272.
Liang Z C. Engineering turbulence[M]. Wuhan: Huazhong University of Science and Technology Press, 1999: 271-272(in Chinese).

[44] Sharma V, Dutta S. Experimental analysis of surface-modified structured torpedo model[C]//Proceedings of the Fluid Mechanics and Fluid Power. Berlin: Springer, 2023.

[45] Lunghi G, Pasqualetto E, Mariotti A, et al. Drag reduction in a spanwise-infinite boat-tailed body through multiple transverse grooves[J]. Computers & Fluids, 2024, 275: 106250.

[46] Chen H W, Rao F G, Shang X P, et al. Biomimetic drag reduction study on herringbone riblets of bird feather[J]. Journal of Bionic Engineering, 2013, 10(3): 341-349.

[47] 崔光耀, 潘翀, 高琪, 等. 沟槽方向对湍流边界层流动结构影响的实验研究[J]. 力学学报, 2017, 49(6): 1201-1212.
Cui G Y, Pan C, Gao Q, et al. Flow structure in the turbulent boundary layer over directional riblets surfaces[J]. Chinese Journal of Theoretical and Applied Mechanics, 2017, 49(6): 1201-1212(in Chinese).

[48] 戎瑞, 刘顺超, 王松岭, 等. 脊状结构对翼型边界层分离特性影响大涡模拟研究[J]. 空气动力学学报, 2018(2): 350-356.
Rong R, Liu S C, Wang S L, et al. Study on effects of riblet structure on airfoil boundary layer separation characteristics by LES[J]. Acta Aerodynamica Sinica, 2018(2): 350-356(in Chinese).

[49] Guo C Y, Wu Y Y, Han Y, et al. Experimental study on influence of sharkskin denticles structure on the hydrodynamic performance of airfoil[J]. Ocean Engineering, 2023, 271: 113756.

[50] 张子良, 张明明. 仿生减阻翼型的气动性能[J]. 航空动力学报, 2021, 36(8): 1740-1748.
Zhang Z L, Zhang M M. Aerodynamic performance for bionic drag-reducing airfoil[J]. Journal of Aerospace Power, 2021, 36(8): 1740-1748(in Chinese).

[51] Li Z P, He L, Zheng Y X. Quasi-analytical solution of optimum and maximum depth of transverse V-groove for drag reduction at different Reynolds numbers[J]. Symmetry, 2022, 14(2): 342.

[52] 袁一平, 杨华, 石亚丽, 等. 风力机专用翼型表面微沟槽减阻特性研究[J]. 工程热物理学报, 2018, 39(6): 1258-1266.
Yuan Y P, Yang H, Shi Y L, et al. Study on drag reduction characteristics of airfoil for wind turbine with microgrooves on surface[J]. Journal of Engineering Thermophysics, 2018, 39(6): 1258-1266(in Chinese).

[53] El-Samni O A, Chun H H, Yoon H S. Drag reduction of turbulent flow over thin rectangular riblets[J]. International Journal of Engineering Science, 2007, 45(2-8): 436-454.

[54] 徐琰, 张臣, 汪子轩. 局部近似平面V型沟槽减阻特性数值模拟研究[J]. 航空制造技术, 2021, 64(15): 86-99.
Xu Y, Zhang C, Wang Z X. Simulation analysis on drag reduction performance of characteristic parameters of V-groove on local approximate plane[J]. Aeronautical Manufacturing Technology, 2021, 64(15): 86-99(in Chinese).

[55] 郭飞飞, 王珂心, 苏慧明, 等. 表面微结构的尺寸参数对其减阻性能的影响[J]. 机械设计与研究, 2023, 39(1): 112-117.
Guo F F, Wang K X, Su H M, et al. Effect of the surface microstructure size parameters on its drag reduction performance[J]. Machine Design and Research, 2023, 39(1): 112-117(in Chinese).

[56] 黄桥高, 潘光, 胡海豹, 等. 回转体表面不同尺寸脊状结构的减阻数值仿真[J]. 火力与指挥控制, 2011, 36(3): 35-38.
Huang Q G, Pan G, Hu H B, et al. Numerical simulation research on drag reduction over riblet structure with different size of gyroidal object[J]. Fire Control & Command Control, 2011, 36(3): 35-38(in Chinese).

[57] Quintavalla S J, Angilella A J, Smits A J. Drag reduction on grooved cylinders in the critical Reynolds number regime[J]. Experimental Thermal and Fluid Science, 2013, 48: 15-18.

[58] Cui J, Fu Y B. A numerical study on pressure drop in microchannel flow with different bionic micro-grooved surfaces[J]. Journal of Bionic Engineering, 2012, 9(1): 99-109.

[59] 谭惠丰, 康敬天, 卫剑征, 等. 三角形微沟槽飞艇蒙皮表面的流场分析[J]. 哈尔滨工业大学学报, 2014, 46(7): 32-35.
Tan H F, Kang J T, Wei J Z, et al. Flow field analysis of Micro-V shape riblets airship surface[J]. Journal of Harbin Institute of Technology, 2014, 46(7): 32-35(in Chinese).

[60] Mohammadi A, Floryan J M. Numerical analysis of laminar-drag-reducing grooves[J]. Journal of Fluids Engineering, 2015, 137(4): 041201.

[61] 刘占一, 胡海豹, 宋保维, 等. 不同间隔脊状表面的减阻数值仿真研究[J]. 系统仿真学报, 2009, 21(19): 6025-6028.
Liu Z Y, Hu H B, Song B W, et al. Numerical simulation research about riblet surface with different spacing[J]. Journal of System Simulation, 2009, 21(19): 6025-6028(in Chinese).

[62] 潘光, 黄桥高, 胡海豹, 等. 回转体表面不同间隔脊状结构的减阻数值仿真研究[J]. 空气动力学学报, 2010, 28(3): 267-271.
Pan G, Huang Q G, Hu H B, et al. Numerical simulation research about the drag reduction over riblet structure with different space of gyroidal object[J]. Acta Aerodynamica Sinica, 2010, 28(3): 267-271(in Chinese).

[63] Martin S, Bhushan B. Fluid flow analysis of a shark-inspired microstructure[J]. Journal of Fluid Mechanics, 2014, 756: 5-29.

[64] 窦如宏. 仿生非光滑表面减阻特性数值研究[D]. 青岛: 青岛科技

大学, 2020: 75-76.
Dou R H. Numerical study of bionic non-smooth surface for drag reduction[D]. Qingdao: Qingdao University of Science & Technology, 2020: 75-76(in Chinese).

[65] 陈璠, 徐朋飞. “仿生学”沟槽减阻仿真分析及机理研究[J]. 航空发动机, 2021, 47(2): 28-32.
Chen F, Xu P F. Simulation analysis and mechanism study on drag reduction of “bionics” groove[J]. Aeroengine, 2021, 47(2): 28-32 (in Chinese).

[66] Qiu H Y, Chauhan K, Lei C W. A numerical study of drag reduction performance of simplified shell surface microstructures[J]. Ocean Engineering, 2020, 217: 107916.

[67] Wu Z R, Li S G, Liu M, et al. Numerical research on the turbulent drag reduction mechanism of a transverse groove structure on an airfoil blade[J]. Engineering Applications of Computational Fluid Mechanics, 2019, 13(1): 1024-1035.

[68] Zhang Y F, Yan C Y, Chen H X, et al. Study of riblet drag reduction for an infinite span wing with different sweep angles[J]. Chinese Journal of Aeronautics, 2020, 33(12): 3125-3137.

[69] 王晋军, 李亚臣. 沟槽面三角翼减阻特性实验研究[J]. 空气动力学学报, 2001, 19(3): 283-287.
Wang J J, Li Y C. Experimental studies on the drag reduction characteristics for flow over delta wings with riblets surfaces[J]. Acta Aerodynamica Sinica, 2001, 19(3): 283-287(in Chinese).

[70] Li Z P, He L, Zuo Y R, et al. Numerical study on the influence of riblet yaw-angle on drag reduction rate with different immersion heights[J]. Results in Engineering, 2024, 22: 102118.

[71] Hossain K R, Wu J Y, Xu X, et al. Tribological bioinspired interfaces for 3D printing[J]. Tribology International, 2023, 188: 108904.

[72] Kumar V, Verma R, Sharma V S, et al. Recent progresses in superhydrophobicity and micro-texturing for engineering applications[J]. Surface Topography: Metrology and Properties, 2021, 9(4): 043003.

[73] Wang Y Q, Wei Y J, Weng D, et al. Aerodynamic drag reduction by the trapezoid spanwise groove inspired by pigeon feathers[J]. Energies, 2023, 16(5): 2379.

[74] Zheng S H, Liang X, Li J Y, et al. Drag reduction using bionic groove surface for underwater vehicles[J]. Frontiers in Bioengineering and Biotechnology, 2023, 11: 1223691.

[75] Song X W, Zhang M X, Lin P Z. Skin friction reduction characteristics of nonsmooth surfaces inspired by the shapes of barchan dunes[J]. Mathematical Problems in Engineering, 2017, 2017(1): 6212605.

[76] 程拼拼, 蒋成刚, 吴承伟. 仿生二级微沟槽表面减阻特性数值模拟[J]. 中国科技论文, 2014, 9(8): 940-943.
Cheng P P, Jiang C G, Wu C W. Numerical simulation of drag reduction characteristics in bionic secondary micro-grooved surface[J]. China Sciencepaper, 2014, 9(8): 940-943(in Chinese).

[77] Wang L Y, Wang C, Wang S Y, et al. Design and analysis of micro-nano scale nested-grooved surface structure for drag reduction based on ‘Vortex-Driven Design’[J]. European Journal of Mechanics-B, 2021, 85: 335-350.

[78] He X X, Liu Y H, Zhan H Y, et al. Bidirectional underwater drag reduction on bionic flounder two-tier structural surfaces[J]. Biomimetics, 2023, 8(1): 116.

[79] 秦立果, 孙红江, 杨浩, 等. 鲨鱼皮盾鳞结构减阻性能数值模拟[J]. 中国科技论文, 2019, 14(12): 1271-1277.
Qin L G, Sun H J, Yang H, et al. Numerical simulation on drag reduction performance of shark skin with denticle structures[J]. China Sciencepaper, 2019, 14(12): 1271-1277(in Chinese).

[80] 徐征, 刘日, 王天昊, 等. 仿生鲨鱼皮复合微纳减风阻结构的仿真与制备[J]. 实验流体力学, 2024, 38(2): 107-114.
Xu Z, Liu R, Wang T H, et al. Simulation and fabrication of bionic sharkskincomposite micro-nano wind resistance reduction structure[J]. Journal of Experiments in Fluid Mechanics, 2024, 38(2): 107-114(in Chinese).

[81] Wu L Y, Luo G H, He F F, et al. Bionic research on Paramisgurnus dabryanus scales for drag reduction[J]. RSC Advances, 2022, 12(34): 22226-22235.

[82] Abdollahzadeh M J, Moosavi A. Optimization of microgrooves for water-solid drag reduction using genetic algorithm[J]. Journal of Ocean Engineering and Marine Energy, 2020, 6(3): 221-242.

[83] 宋居正, 李耕耘. 基于遗传算法的横向减阻沟槽优化及机理[J]. 航空动力学报, 2024, 39(5): 226-233.
Song J Z, Li G Y. Shape optimization and mechanism of transverse groo[J]. Journal of Aerospace Power, 2024, 39(5): 226-233 (in Chinese).

[84] 潘光, 武昊, 胡海豹, 等. 脊状结构减阻优化设计方法[J]. 火力与指挥控制, 2011, 36(6): 131-134.
Pan G, Wu H, Hu H B, et al. Optimization design method for drag reduction over riblet structure[J]. Fire Control & Command Control, 2011, 36(6): 131-134(in Chinese).

[85] 李超群, 唐硕, 李易, 等. 基于神经网络的减阻沟槽壁面形状优化[J]. 航空动力学报, 2022, 37(3): 639-648.
Li C Q, Tang S, Li Y, et al. Sub-optimization of riblet shape based on neural networks[J]. Journal of Aerospace Power, 2022, 37(3): 639-648(in Chinese).

[86] Mohammadi A, Floryan J M. Groove optimization for drag reduction[J]. Physics of Fluids, 2013, 25(11): 113601.

[87] Wen L, Weaver J C, Lauder G V. Biomimetic shark skin: design, fabrication and hydrodynamic function[J]. Journal of Experimental Biology, 2014, 217(10): 1656-1666.

[88] Stenzel V, Wilke Y, Hage W. Drag-reducing paints for the reduction of fuel consumption in aviation and shipping[J]. Progress in Organic Coatings, 2011, 70(4): 224-229.

[89] Zhao D Y, Huang Z P, Wang M J, et al. Vacuum casting replication of micro-riblets on shark skin for drag-reducing applications[J]. Journal of Materials Processing Technology, 2012, 212(1): 198-202.

[90] Chen H W, Zhang X, Ma L X, et al. Investigation on large-area fabrication of vivid shark skin with superior surface functions[J]. Applied Surface Science, 2014, 316: 124-131.

[91] Kaakkunen J J J, Tiainen J, Jaatinen-Värri A, et al. Nanosecond laser ablation of the trapezoidal structures for turbomachinery applications[J]. Procedia Manufacturing, 2018, 25: 435-442.

[92] Hockauf R, Asadi E, Denkena B, et al. Grinding of riblets with “beaver tooth” multi-layer tools[J]. Procedia CIRP, 2018, 71: 155-159.

[93] Huang Y, Wu Y, He Y, et al. Experimental study on laser assisted

belt grinding of shark-skin bio-inspired riblets under different surface roughness[J]. Procedia CIRP, 2020, 89: 277-281.

[94] Luo Y H, Zhang D Y. Investigation on fabricating continuous vivid sharkskin surface by bio-replicated rolling method[J]. Applied Surface Science, 2013, 282: 370-375.

[95] 沈洪, 任浩东, 李海东. 面向功能表面的激光仿生制造技术研究进展[J]. 空天防御, 2023, 6(2): 12-22.
Shen H, Ren H D, Li H D. Laser bionic manufacturing technology for functional surface[J]. Air & Space Defense, 2023, 6(2): 12-22 (in Chinese).

[96] Walsh M J, Sellers W L, Mcginley C B. Riblet drag at flight conditions[J]. Journal of Aircraft, 1989, 26(6): 570-575.

[97] Willsher J. The effect of biocide free foul release systems on vessel performance[R]. London: International Paint Ltd. , 2008.

[98] Viswanath P R. Aircraft viscous drag reduction using riblets[J]. Progress in Aerospace Sciences, 2002, 38(6-7): 571-600.

[99] Enyutin G V, Lashkov Y A, Samoilova N V. Drag reduction in riblet-lined pipes[J]. Fluid Dynamics, 1995, 30(1): 45-48.

[100] Bixler G D, Bhushan B. Shark skin inspired low-drag microstructured surfaces in closed channel flow[J]. Journal of Colloid and Interface Science, 2013, 393: 384-396.

[101] Zheng C T, Zhou P, Mao J Q, et al. Experimental investigations on the effects of longitudinal groove fabric for drag reductions on a track cyclist[J]. Experimental Thermal and Fluid Science, 2024, 156: 111223.

[102] Eguchi Y, Kikuchi N, Kawabata K, et al. Drag reduction mechanism and aerodynamic characteristics of a newly developed overhead electric wire[J]. Journal of Wind Engineering and Industrial Aerodynamics, 2002, 90(4-5): 293-304.

[103] Wang Y M, Li N, Ma Y H, et al. Field experiments evaluating a biomimetic shark-inspired (BioS) subsoiler for tillage resistance reduction[J]. Soil and Tillage Research, 2020, 196: 104432.

Research status and progress of bionic groove drag reduction

WANG Wei[1, 2, *], CUI Shen'ao[1], CHEN Jiahua[1], ZHAO Haoxiang[1], ZHANG Zhen[1]

(1. School of Civil Aviation, Shenyang Aerospace University, Shenyang 110136, China;
2. Liaoning General Aviation Academy, Shenyang 110136, China)

Abstract: The groove drag reduction technology based on the shark skin ridge structure has been widely concerned and studied by scientists because of its simple and easy operation and low energy consumption rate. It has shown great potential in marine ships, aerospace, pipeline transportation and other fields. The state of research and advancements in bionic groove drag reduction technologies during the previous two decades are examined. Firstly, three widely recognized drag reduction mechanisms are reviewed and other drag reduction mechanisms are supplemented. The design and optimization of the groove shape drag reduction as well as the current state of processing technology research are then discussed, along with a summary of the effects of the groove's position and arrangement, dimensionless size, and yaw angle on the drag reduction performance. Finally, the practical application of bionic grooves in many fields is briefly described, and the research on drag reduction of bionic grooves is summarized and prospected.

Keywords: bionic groove; drag reduction mechanism; shark surface; drag reduction rate; turbulent flow

Received: 2024-12-19; **Accepted**: 2025-02-14; **Published Online**: 2025-03-10 11:29

URL: link.cnki.net/urlid/11.2625.V.20250310.0924.001

Foundation items: Aeronautical Science Foundation of China (2020Z006054002); Natural Science Foundation of Liaoning Province (2023-MS-243); Project of the Education Department of Liaoning Province (20240109)

* **Corresponding author**. E-mail: wangwei7832@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0420

无人机视角下目标检测的YOLO系列算法研究进展

陈金林，吴一全*，苑玉彬
(南京航空航天大学 电子信息工程学院，南京 211106)

摘 要：在诸多基于深度学习的目标检测算法中，YOLO系列算法因具有检测速度快、检测精度高、适应性强等优点，成为无人机(UAV)航拍图像目标检测的优先选择。UAV航拍图像与视频复杂多变的特点使该视角下YOLO系列目标检测算法成为一个具有挑战性的难题，因此，有必要将近期UAV视角下目标检测的YOLO系列算法研究与应用成果进行全面总结与分析。归纳10种YOLO基础版本的发展、联系、优化特点，重新绘制YOLO系列结构图，指出各种YOLO版本优化策略的作用与特点。综述在UAV目标检测应用领域中YOLO的优化方法数据训练前的预处理、训练与推理优化、注意力模块的应用、卷积优化、检测头及损失函数的优化；梳理UAV视角下的应用数据集及数据扩展方法。通过实验对UAV视角下数据集上YOLO系列算法的性能及卷积块注意力模块(CBAM)对YOLO不同改进的作用进行评估。指出了UAV视角下目标检测YOLO系列算法的研究还存在的一些问题与可能发展趋势。

关 键 词：无人机系统；目标检测；深度学习；YOLO系列算法；数据集；性能对比
中图分类号：V279; TP391.41
文献标志码：A **文章编号**：1001-5965(2026)10-3509-26

无人机(unmanned aerial vehicle, UAV)起飞灵活，操作简单，受地形障碍物影响较小，在恶劣环境下也具有较强的适应性，这使其可作为一种获取低空高分辨率遥感图像与视频的专用遥感系统平台。与传统平台相比，UAV航拍成本低、效率高，且能够弥补由于环境、时间、成本等限制造成的信息损失。UAV采集的图像、视频与音频等数据无论是内容上还是时间上都具有丰富的信息量。UAV通过搭载的摄像机与智能系统，能够实现实时处理拍摄的图像与视频，使其具有目标检测、识别、定位与追踪等功能[1]。

现阶段目标检测已经成为UAV应用领域的一项重要技术，尤其是UAV与深度学习的紧密融合，使UAV航拍目标的智能检测与识别技术开始逐步应用到各个行业中[2]。UAV视角下的目标检测在军事侦察、精准打击、智慧农林业、交通监控、电力巡检、灾害监测、搜索救援等军事与民用领域发挥着越来越重要的作用[3]。

UAV的飞行姿态大体分为起飞、悬停、巡航和降落，图像与视频的拍摄通常处于悬停与巡航状态下[4]。受UAV的飞行稳定性、载荷能力、飞行高度、巡航能力与飞行环境等因素的影响，UAV视角下的目标物体多数呈现出尺度多样、数量多、目标易变形、背景和场景复杂等特点，可提取的单个目标特征信息较少，这使目标检测变得更加困难[5-6]。UAV机载遥感传感器需要实时准确地调整飞行速度、飞行高度、飞行姿态、航拍角度，使目标物体始终处于UAV的监控范围内，因此，优秀的目标检测算法可以促进UAV飞行控制技术的进步。

现阶段UAV目标检测算法有基于两阶段的快速区域卷积神经网络(fast region-based convolutional neural network, FAST-RCNN)[7]、级联区域卷积神

收稿日期：2024-06-12；**录用日期**：2024-07-26；**网络出版时间**：2024-08-26 11:10
网络出版地址：link.cnki.net/urlid/11.2625.V.20240824.1220.001
基金项目：国家自然科学基金(61573183)；安徽省高校科研计划项目(2022AH051584)
* **通信作者**. E-mail：nuaaimage@163.com

引用格式：陈金林，吴一全，苑玉彬. 无人机视角下目标检测的YOLO系列算法研究进展[J]. 北京航空航天大学学报，2026，52(10)：3509-3534. Chen J L，Wu Y Q，Yuan Y B. Research progress on YOLO series algorithms for target detection from UAV vision[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3509-3534 (in Chinese).

经网络(cascade region-based convolutional neural network, Cascade RCNN)[8]、基于区域的全卷积网络(region-based fully convolutional network, R-FCN)[9]等;基于一阶段目标检测算法有YOLO(you only look once)[10]、单次多框检测器(single shot multi-box detector, SSD)[11]、RetinaNet[12]和全卷积单阶段目标检测(fully convolutional one-stage object detection, FCOS)[13]等。在众多的UAV目标检测模型中,YOLO具有版本多、相对准确度高、检测速度快、适应性强、网络结构简单等优点,使其备受关注。

近几年,该研究内容的相关综述较多,但多数仅从YOLO系列算法,或者侧重于UAV视角下目标检测模型角度论述。文献[14]于2022年介绍了YOLOv2~YOLOv4及各种改进,并对比相关改进算法的性能及优缺点。文献[15]于2022年对当时8个版本的YOLO网络结构、损失函数与应用场景进行了梳理。文献[16]于2023年梳理了YOLOv1~YOLOv7算法发展历程及改进研究,综述了典型数据集、评价指标整体框架和发展历程,重点论述了在输入阶段、特征提取阶段及预测阶段3个方面的改进方法,也阐述了YOLO算法一些应用领域。文献[17]于2024年全面综述了当前YOLO的多个不同版本,从YOLO的发展历程、网络结构、改进特点、性能对比等方面进行论述。上述几篇综述文献的主要贡献是对YOLO细致的总结与分析,可以帮助研究人员更清晰、全面地认识YOLO算法及其改进方法,但没有侧重分析UAV视角下YOLO目标检测系列算法的改进与应用。

在UAV视角下目标检测研究方面,文献[18]于2022年全面分析了基于深度学习的UAV视角下目标检测与跟踪方法的研究进展和前景,较详细地分析了基于UAV目标检测的一些数据集。文献[19]于2024年对UAV视角下的目标特点、数据集、目标检测算法及目标检测算法的性能做了全面概括。文献[20]于2024年综述了UAV航拍影像目标检测与语义分割方法的发展历程,对各种网络模型进行分类,分别总结了其改进策略、应用场景、贡献和局限性,同时,梳理了近些年UAV航拍影像的数据集。文献[21]于2023年综述了基于深度学习的UAV航拍视频多目标检测与跟踪算法,总结了该领域的最新进展,包括多目标检测、多目标跟踪,分别阐述了2类算法的原理并分析其优缺点。上述文献的主要贡献是综述了UAV航拍影像目标检测、语义分割、目标跟踪算法,并对后继相关研究的可能方向做出了合理的展望。这些算法也包括了YOLO目标检测算法,但这些综述在UAV视角下YOLO目标检测的改进策略、改进原因、改进优势分析较少。

近些年来,在UAV视角下目标检测的众多算法中,基于YOLO的研究与应用成果较多、发展快,许多UAV视觉目标检测模型是基于YOLO开发的。为应对UAV目标检测应用的行业需求,需对YOLO系列算法进行优化与改进。虽然文献[22]于2020年综述了当时UAV目标识别和检测的YOLO系列算法研究进展,并进行了YOLO应用算法的性能分析。文献[23]于2022年综述了UAV视角下目标检测的YOLOv2~YOLOv4算法。上述算法综述的YOLO版本偏少,仅经过几年的发展,YOLO就出现了多个版本,且在UAV视角下目标检测任务的应用中发挥着重要作用。因此,为进一步研究UAV视角下YOLO目标检测算法,有必要全面总结该领域的研究成果,分析算法的优化与改进特点。

本文对UAV视角下YOLO目标检测改进算法做更系统、更深入的总结。概述10种YOLO目标检测基础版本的发展历程及其改进方法;分别综述UAV视角下YOLO目标检测应用数据集的采集、扩展、增强等;总结UAV目标检测的YOLO算法改进,包括特征提取模块、检测头模块、注意力机制、模型轻量化优化、损失函数等;从UAV视角下YOLO目标检测方法和数据集2个维度阐述了该领域当前存在的有关问题,并对下一步的可能研究趋势进行了展望。

1 YOLO目标检测版本的发展历程

自YOLO问世以来,仅经过不到10年的发展,已经出现了16个版本。其中,YOLOv1~YOLOv9与YOLOX在UAV应用领域是较有影响力的目标检测算法,主要体现YOLO的基本设计思想、创新特点、应用研究,图1是10种版本的发展时间线及其改进基础。除了本文列举的版本以外,YOLO还有很多优秀的改进算法,且被应用到各个领域。

1.1 YOLO的10种版本

1.1.1 YOLO目标检测算法

YOLO是在2015年被提出的一阶段目标检测算法,其主干网络基于GoogLeNet,网络结构如图2所示,其中,k为每个网格单元预测的边界框数量,c为目标类别数。特征提取模块将输入图像划分为等大的特征网格。检测头模块中每个特征网格预测同一类别的边界框和不同类别的置信度。每个边界框的预测由5个参数组成:目标的置信度分数、网格单元的边界框中心坐标(x、y)、高度和宽度。检测头可选择性地经过非极大值抑制(non-maximum suppression, NMS)来消除多余的边界框,直到找出与真实框损失最小的边界框。

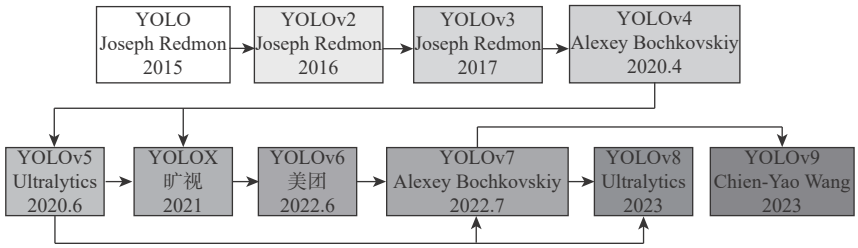


图 1　10 种 YOLO 版本发展时间线

Fig. 1　10 YOLO versions development timelines

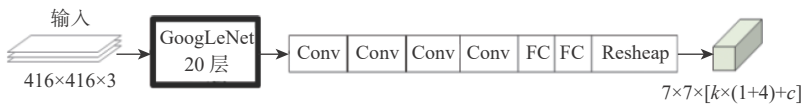


图 2　YOLOv1 检测网络结构

Fig. 2　YOLOv1 detection network structure

1.1.2　YOLOv2

YOLOv2[24] 网络结构如图 3 所示, 算法主要改进有 3 个方面: 锚框机制、特征提取、特征融合。YOLOv2 提前设定固定高宽的锚框, 使每个网格单元都有多个维度来预测目标。在训练过程中, 模型学习的是锚框和真实框之间的偏移; 在推理阶段, 模型学习锚框的偏移, 再逆推回最终的真实框预测值。YOLOv2 应用轻量级网络 Darknet-19 替代了 GoogLeNet, 改进了卷积层模块 CBL(convolution + batch normalization + leaky ReLU), 使用步幅为 2 的卷积层替代池化层实现特征图的降采样, 这样可以有效地降低因池化层引起的低层级特征损失。批归一化(batch normalization, BN)层可以使学到的数据达到同分布, 较大地提高训练速度。YOLOv2 将图像细粒度特征提取模块添加到转移层, 使浅层特征与深层特征相融合, 从而提升了模型对小目标的检测能力。

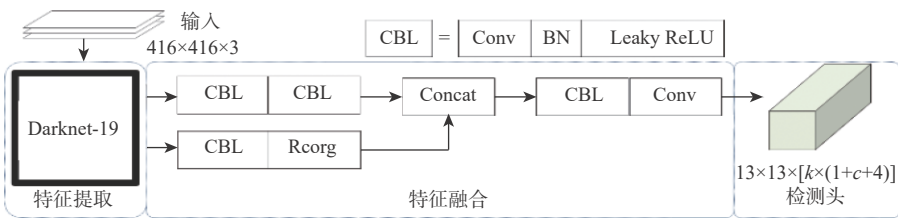


图 3　YOLOv2 检测网络结构

Fig. 3　YOLOv2 detection network structure

1.1.3　YOLOv3

YOLOv3[25] 网络结构如图 4 所示, 为提高特征提取能力、加深网络深度, 特征提取模块改进为堆叠多个残差模块的 Darknet53。图中相连的深色框表示自上而下, 浅色框表示自下而上, 后续不再赘述。为实现多尺度目标检测、提升小目标检测的性能, YOLOv3 在特征融合模块中使用了特征金字塔网络 (feature pyramid network, FPN), 如图 5 所示。FPN 强化了浅层网络特征信息, 使浅层特征图局部信息更易被提取到。模型采用了 3 个不同尺度检测大小的目标, 有效提高了多尺度目标检测能力。

1.1.4　YOLOv4

YOLOv4[26] 于 2020 年被提出, 其网络结构如图 6 所示。YOLO 发展到此时已经形成了比较完善的模型架构: 特征提取、特征融合、检测头。YOLOv4 的特征提取应用跨阶段部分连接(cross stage partial, CSP)模块[27] 替换了 YOLOv3 的残差模块。CSP 通过跨层连接进行合并, 增强了卷积神经网络的学习能力, 减少了计算量的同时保证了更高

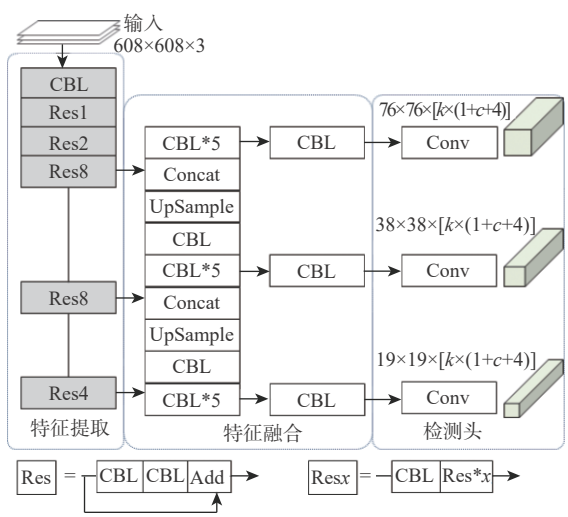


图 4　YOLOv3 检测网络结构

Fig. 4　YOLOv3 detection network structure

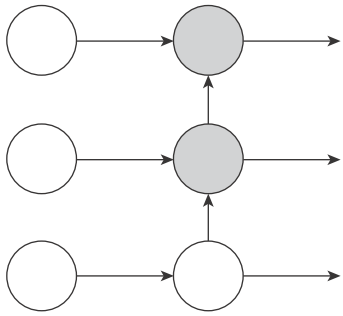


图 5 FPN 网络结构

Fig. 5 FPN network structure

的检测准确率。YOLOv4 在特征融合前新增了空间金字塔池化 (spatial pyramid pooling, SPP) 网络，如图 7 所示，运用多尺度的池化拼接丰富了目标特征，使每个特征图能包含多个感受野信息，突出了目标物体的轮廓信息。

在特征融合中，添加了一个信息自浅层向深层融合的特征金字塔路径聚合网络 (path aggregation network, PAN)，增强了特征图对不同感受野信息的感知，显著提升了模型的性能，如图 8 所示。为提高检测召回率，YOLOv4 改进了锚框偏移机制，通

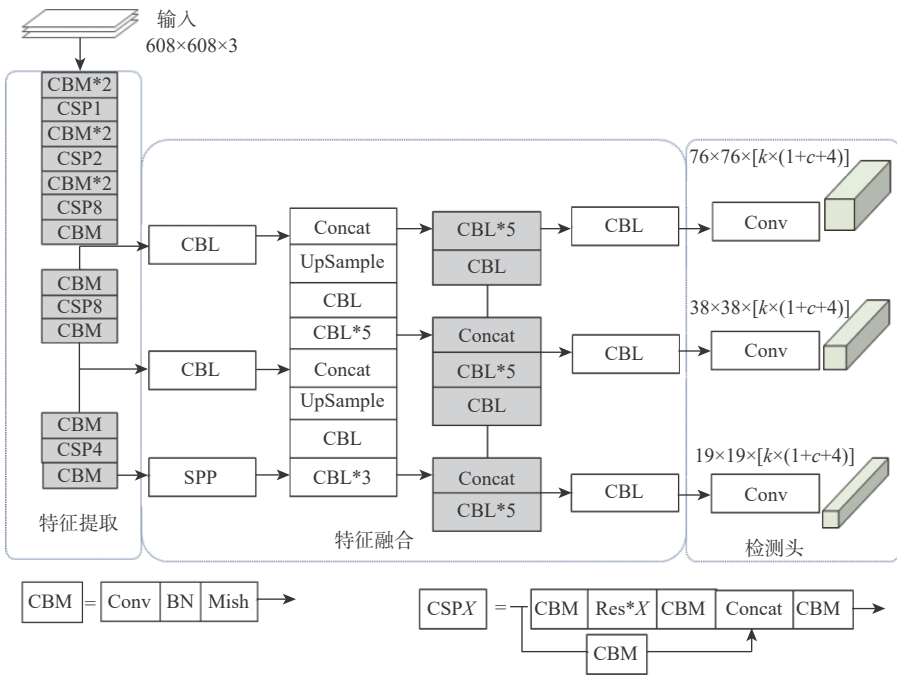


图 6 YOLOv4 检测网络结构

Fig. 6 YOLOv4 detection network structure

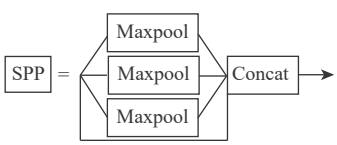


图 7 SPP 网络结构

Fig. 7 SPP network structure

过中心点偏移数值扩展，使正样本更多，增强了准确性。在训练阶段，对输入的数据进行了 Masaic 增强处理，丰富了物体的背景与小目标特征信息。

1.1.5 YOLOv5

Ultralytics 公司于 2020 年公开发布的YOLOv5 网络结构如图 9 所示。YOLOv5 的特征提取模块包括 Foucs 模块、Conv 模块、C3 模块、快速空间金字塔池化(spatial pyramid pooling fast, SPPF)模块[28]。Foucs 模块对图片进行切片操作，实现特征的无损下采样。采用 C3 模块替换 YOLOv4 的 CSP 模块，解决了大型卷积神经网络框架特征提取中梯度信息重复问题。SPPF 模块可以将尺度不固定的特征图转化为统一尺度，经过局部特征与全局特征相融合，丰富了特征图的表达能力，有利于检测复杂多目标图像。在获得数据集的锚框时，YOLOv5 采用了自动学习的方式，提高了模型对遮挡物体的检测，最终模型的召回率得到了提高。为满足对目标检测的不同需求，YOLOv5 的开发团队为其设计了 4 种不同尺寸的版本。

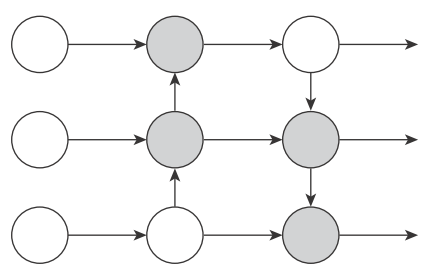


图 8 FPN 与 PAN 的融合网络结构

Fig. 8 Fusion network structure of FPN and PAN

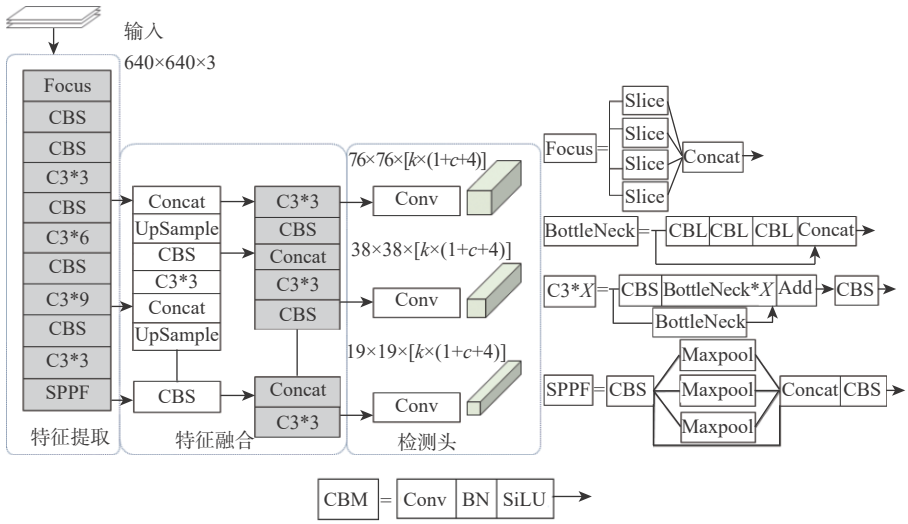


图 9　YOLOv5 检测网络结构

Fig. 9　YOLOv5 detection network structure

1.1.6　YOLOX

YOLOX[29] 的网络结构如图 10 所示。较早期的 YOLO 检测头是使用一层卷积同时完成分类和定位 2 个任务的耦合头。YOLOX 的解耦头是将分类和定位分为 2 个不同的卷积来完成, 如图 11 所示。这种做法虽然会增加模型的参数量, 但提高了目标检测的精度, 后期的版本多采用了解耦头结构。为解决因固定锚框导致训练过程中正负样本不均衡的问题, YOLOX 采用无锚框的方法, 直接预测网格偏移量及预测框的高度和宽度。YOLOX 选用简化最优传输分配(simplified optimal transport assignment, SimOTA)作为正负样本匹配策略, 解决了特殊情况下的样本分配不均衡问题, 但该方法增加了额外的训练时间。

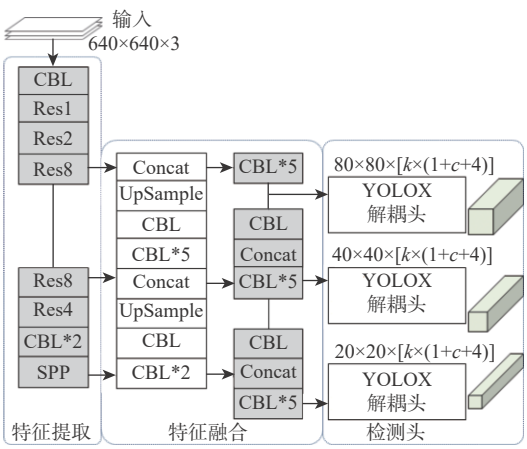


图 10　YOLOX 检测网络结构

Fig. 10　YOLOX detection network structure

1.1.7　YOLOv6

YOLOv6[30] 是美团视觉于 2022 年研发的, 其网络结构如图 12 所示, 其中, SConv 表示对称卷积模块。YOLOv6 的特征提取将重参数化模块(reparameterized block, RepBlock)作为小型网络的构建块, 对于大型模型使用跨阶段部分堆叠重参数化(cross stage partial stacked reparameterized, CSPStackRep)模块。为减少计算量并取得更低推理延迟, YOLOv6 采用混合通道策略重新设计了更高效的解耦头结构。为获得更多高质量的正样本, YOLOv6 引入了任务对齐的单阶段目标检测(task-aligned one-stage object detection, TOOD)动态分配正负样本策略, 并结合无锚框进一步提高检测精度。

1.1.8　YOLOv7

2022 年 YOLOv4 开发团队提出 YOLOv7[31] 的网络结构, 如图 13 所示, 其中, SPPCSPC 为空间金字塔池化与跨阶段局部网络。在特征提取与特征融合 2 个部分使用高效层聚合网络(efficient layer aggregation network, ELAN)模块, 其中, ELAN_H 表示改进 ELAN。ELAN 模块通过控制最短和最长的梯度路径, 使网络能够学习到更多的特征, 让更深的网络可以更加高效地学习和收敛, 模块结构如图 14 所示。在训练过程中, YOLOv7 的检测头引入了辅助头, 通过增加训练成本提升精度, 但不影

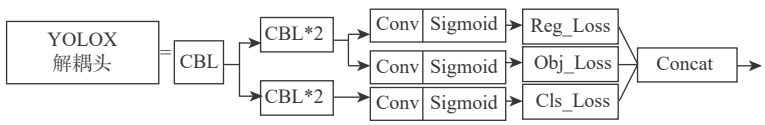


图 11　YOLOX 解耦头

Fig. 11　Decoupled head of YOLOX

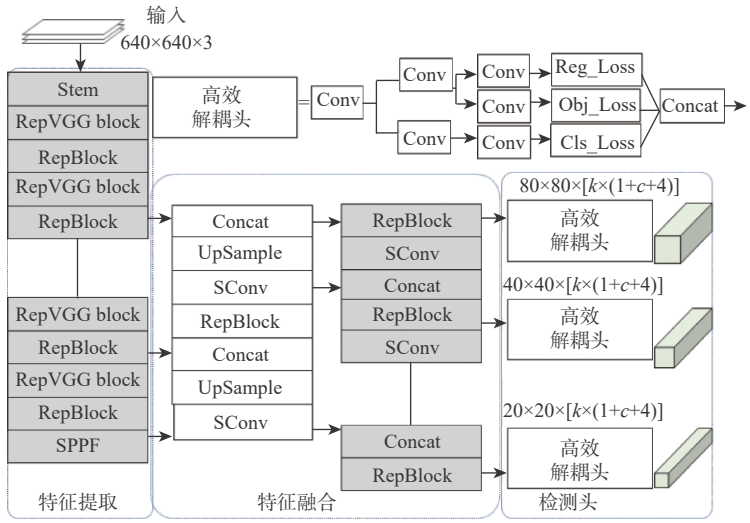


图 12 YOLOv6 检测网络结构

Fig. 12 YOLOv6 detection network structure

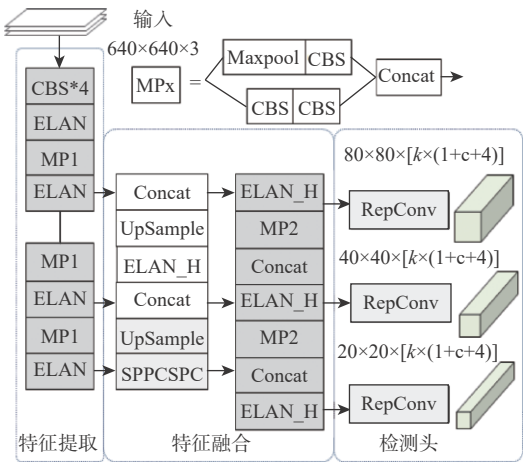


图 13 YOLOv7 检测网络结构

Fig. 13 YOLOv7 detection network structure

响推理的时间，辅助头结构如图 15 所示。YOLOv7 的检测头经过重参数化卷积（reparameterized convolution, RepConv）调整通道数，再使用 1×1 卷积预测坐标、类别、置信度，训练时有 3 个分支相加输出，推理时会将分支的参数重参数化到主分支上，这种重参数化技术在推理阶段将多个计算模块合并为一个模块，提高模型的效率和性能。

1.1.9 YOLOv8

Ultralytics 公司 2023 年发布了 YOLOv8，其继承了早期版本优良的改进策略，如无锚框、解耦头部、训练数据增强等，检测网络结构如图 16 所示。YOLOv8 将 YOLOv5 的 C3 模块替换成拥有更多的残差连接的 C2f 模块，提高了检测的精度。YOLOv8

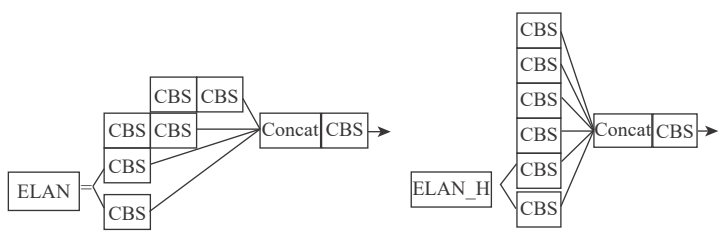


图 14 ELAN 网络结构

Fig. 14 ELAN network structure

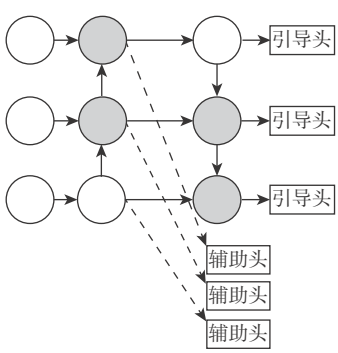


图 15 辅助检测头结构

Fig. 15 Auxiliary detection head structure

改变了交并比 (intersection over union, IOU) 匹配和单边比例的分配方式，运用分类与回归的分数作为加权分数选择正负样本。同时，YOLOv8 提供了更加丰富的超参数和模型结构选项，使其在应用中可以更加方便地进行模型调整和优化。模型开发团队提供了更加丰富的超参数和模型结构选项，使 YOLOv8 在应用中可以更加方便地进行模型调整和优化。

1.1.10 YOLOv9

YOLOv9[32] 由中国台湾 Academia、台北科技大

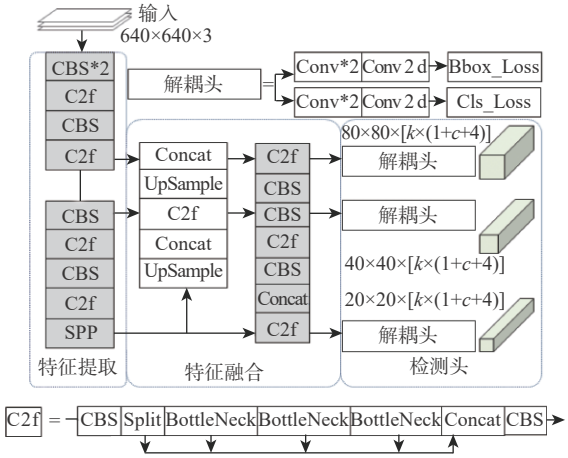


图 16　YOLOv8 检测网络结构

Fig. 16　YOLOv8 detection network structure

学等机构于 2023 联合开发, 网络结构如图 17 所示。研究者在 YOLOv9 中应用可编程梯度信息(programmable gradient information, PGI)学习目标特征。PGI 主要包括 3 个部分: 主分支用于推理的架构, 不需要额外的推理成本; 辅助可逆分支是为了解决网络加深造成的信息瓶颈问题; 多级辅助信息旨在处理深度监督带来的误差累积问题。在 YOLOv9 中, 广义高效层聚合网络(generalized efficient layer aggregation network, GELAN)使用传统的卷积实现了更好的参数利用, 该卷积兼顾了模型复杂度、准确性和推理速度。

除了上述 YOLO 系列算法的网络结构、训练策略、主要改进以外, YOLO 不同的版本在激活函数、损失

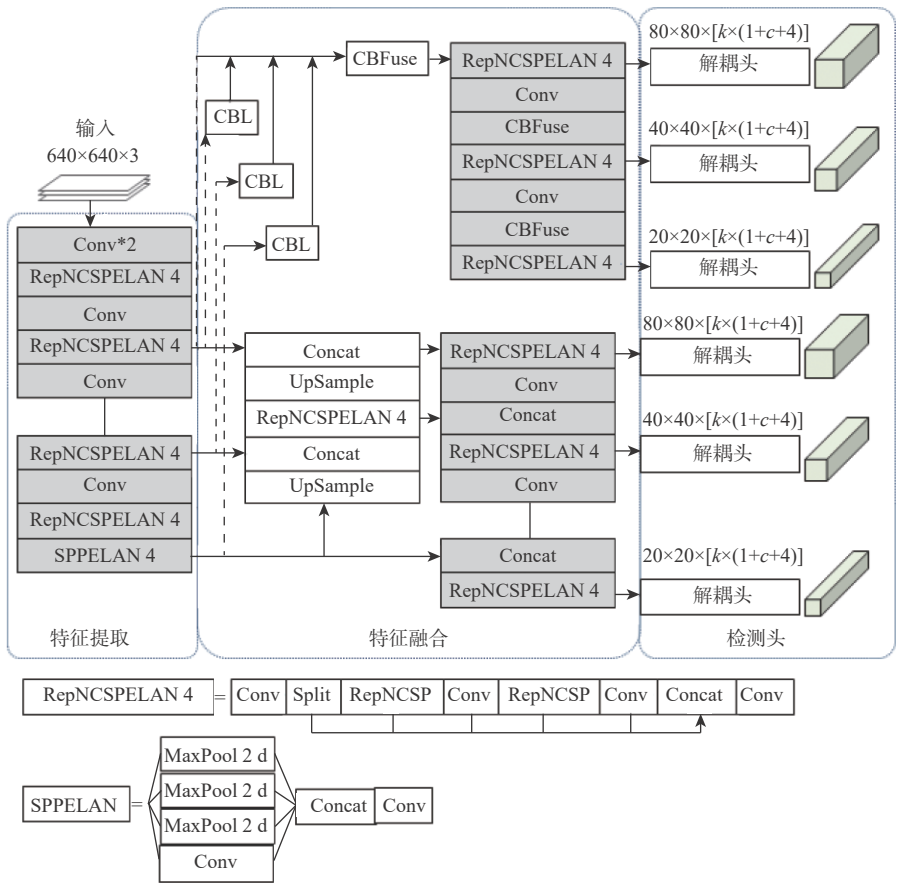


图 17　YOLOv9 检测网络结构

Fig. 17　YOLOv9 detection network structure

函数、Anchor 机制等方面也略有变化, 如表 1 所示。

1.2　YOLO 的性能

目前, YOLO 目标检测算法的性能评价指标大体可分为 2 类: 检测能力指标、算法复杂度指标。检测能力指标中有检测准确性指标与检测速度指标, 具体指标项目如表 2 所示, 其中, 样本中被正确识别的正样本数称为真阳性 (true positives, TP), 记为 n_{TP}; 样本中被正确识别的负样本数称为真阴性 (true negatives, TN), 记为 n_{TN}; 样本中被错误识别为负样本的正样本数称为假阴性 (false negatives, FN), 记为 n_{FN}; 样本中错误识别为正样本的负样本数为假阳性 (false positives, FP), 记为 n_{FP}。

在模型的训练过程中, 正样本帮助模型学习如何准确地定位和分类目标物体, 而负样本则负责模型学习如何区分目标物体和背景, 增强了模型的鲁棒性和准确性。在性能评价中, 正样本和负样本作为计算目标检测模型准确性的指标基础。正负样本的划分直接影响模型检测能力指标的优劣。

表1 YOLO不同的优化方案

Table 1 YOLO different optimization schemes

YOLO	锚框机制	正负样本分配	检测头	激活函数	损失函数		
					Cls_Loss	Obj_Loss	Bbox_Loss
YOLO	无锚框	Max IOU	耦合头	Leaky ReLU	均方差	均方差	IOU
YOLOv2	锚框	Max IOU	耦合头	Leaky ReLU	均方差	均方差	IOU
YOLOv3	锚框	Max IOU	耦合头	Leaky ReLU	2元交叉熵	2元交叉熵	IOU
YOLOv4	锚框	Max IOU	耦合头	Leaky ReLU、Mish	2元交叉熵	2元交叉熵	DIOU
YOLOv5	锚框	跨网格搜索	耦合头	Leaky ReLU	2元交叉熵	2元交叉熵	CIOU
YOLOX	无锚框	SimOTA	解耦头	SiLU	2元交叉熵	2元交叉熵	CIOU
YOLOv6	无锚框	TOOD	解耦头	SiLU	2元交叉熵	2元交叉熵	CIOU
YOLOv7	无锚框	跨网格搜索	解耦头	SiLU	2元交叉熵	2元交叉熵	CIOU
YOLOv8	无锚框	TOOD	解耦头	SiLU	2元交叉熵	VFL	CIOU
YOLOv9	无锚框	TOOD	解耦头	SiLU	2元交叉熵	VFL	CIOU

注: DIOU和CIOU分别为距离交并比和完全交并比。

表2 YOLO性能评价指标

Table 2 Performance evaluation indicators of YOLO

类别	指标	公式
准确性	召回率R	$n_{\mathrm{TP}}/(n_{\mathrm{TP}}+n_{\mathrm{FP}})$
	精确率P	$n_{\mathrm{TP}}/(n_{\mathrm{TP}}+n_{\mathrm{FN}})$
	平均精度P_k	$\int P(R)\mathrm{d}R$
	平均精度均值	$\frac{1}{K}\sum_{k=1}^{K}P_k$
速度	帧率	

表2中参数量与浮点运算量的计算主要包含卷积层和全连接(fully connected, FC)层,卷积层的参数量P_{J}和浮点运算量F_{J}计算式分别为

$$P_{\mathrm{J}}=\sum_{i=1}^{N}P_i \tag{1}$$

$$P_i=c_{\mathrm{in},i}(h_i w_i c_{\mathrm{out},i}+1) \tag{2}$$

$$F_{\mathrm{J}}=\sum_{i=1}^{N}F_i \tag{3}$$

$$F_i=P_i H_i W_i \tag{4}$$

式中: $c_{\mathrm{in},i}$、h_i、w_i、$c_{\mathrm{out},i}$分别为第i个卷积的输入通道数、核长、核宽、输出通道数; H_i和W_i分别为输出特征图的长度和宽度; N为卷积核个数。

FC层的参数量P_{Q}和浮点运算量F_{Q}计算式分别为

$$P_{\mathrm{Q}}=\sum_{j=1}^{M}(I_j O_j+O_j) \tag{5}$$

$$F_{\mathrm{Q}}=2\sum_{j=1}^{M}(I_j O_j) \tag{6}$$

式中: I_j和O_j分别为第j个FC层输入和输出神经元个数; M为FC层个数。

YOLO是由数据驱动的深度学习算法,因此,训练与测试数据集对性能的影响是不可忽略的。表3

表3 YOLO性能对比[31-34]

Table 3 YOLO performance comparison[31-34]

版本	框架	数据集	图像尺寸/像素	骨干网络	mAP/%	帧率/(帧·s^{-1})	参数量	浮点运算量/次
YOLO	Darknet	VOC	448	GoogLeNet	63.4	45	2.629×10^8	3.37×10^{10}
YOLOv2	Darknet	VOC	416	Darknet24	76.8	67	4.87×10^7	3.06×10^{10}
YOLOv3	Darknet	COCO	608	Darknet53	33.0	78	6.3×10^7	1.18×10^{11}
YOLOv4	Darknet	COCO	608	CSPDarknet53	43.5	65	2.76×10^7	5.20×10^{10}
YOLOv5s	Pytorch	COCO	640	CSPDarknet	37.4	95	7.2×10^6	1.65×10^{10}
YOLOXs	Pytorch	COCO	640	CSPDarknet	40.5	156	9.0×10^6	2.68×10^{10}
YOLOv6	Pytorch	COCO	640	EfficientRep	43.1	109	1.72×10^7	4.42×10^{10}
YOLOv7s	Pytorch	COCO	640	ELANblock	51.4	161	6.03×10^6	1.32×10^{10}
YOLOv8s	Pytorch	COCO	640	CSPDarknet	53.0	101	1.12×10^7	2.86×10^{10}
YOLOv9	Pytorch	COCO	640	ELANblock	56.0		7.1×10^6	2.64×10^{10}
SSD	Pytorch	COCO		ResNet101+FPN	33.2	35.5		
ConterNet	Pytorch	COCO		Hourglass104+FPN	45.1	1.42		
RetinaNet	Pytorch	COCO		ResNet101+FPN	39.1	5.4		
FCOS	Pytorch	COCO		ResNet101+FPN	43.2	60		

给出了 10 种 YOLO 版本的部分性能指标[31-34]。对比可以发现, 经过几代版本的改进, YOLO 目标检测算法的准确性、推理速度、参数量等性能都得到了不同程度的优化; 在同等的 Pytorch 框架下, 相比较于同类的一阶段目标检测算法, 较高版本的 YOLO 目标检测准确性、速度是比较有优势的。

2　UAV 视角下 YOLO 目标检测数据集与预处理

UAV 视角下的图像与视频因采集设备、采集方式、现场环境等因素的影响, 呈现出多样性的特点, 使 UAV 视角下 YOLO 目标检测模型的研究与应用更依赖于特定应用数据集与任务需求。研究人员在运用 YOLO 算法开发 UAV 视角下目标检测应用模型的过程中, 为提高模型对目标的检测能力, 需花费较多的精力收集、处理相关数据集以适应应用需求。

2.1　目标检测数据集

将 UAV 视角下目标检测模型部署到复杂且未知的现实场景时不仅要考虑模型的运算速度、检测准确性, 还要兼顾模型对目标物体的泛化能力。数据集中图像质量、大小、数量直接影响到模型检测的运算速度、准确性与泛化能力, 因此, 一个优秀深度学习模型数据集中的图像应该具有代表性、多样性、完整性和可靠性[33]。在基于 UAV 视角下的 YOLO 目标检测模型的研究与应用中, 常使用一些公共数据集, 如视觉对象类别(visual object classes, VOC)数据集[34]、微软常见对象上下文(microsoft common objects in context, COCO)数据集[35]、航空图像目标检测数据集(dataset for object detection in aerial images, DOTA)[36]、无人机视觉(vision meets drone, VisDrone)数据集[37]、光学遥感图像目标检测(object detection in optical remote sensing images, DIOR)数据集[38]、无人机基准目标检测与跟踪(unmanned aerial vehicle benchmark object detection and tracking, UAVDT)数据集、无人机 123 目标跟踪(unmanned aerial vehicle 123, UAV123)数据集、奥尔巴尼大学检测与跟踪(university at Albany detection and tracking, UA-DETRAC)数据集, 表 4 给出了这些公共数据集的性质及特点。DIOR、DOTA 是航空视角下的数据集, 其中包含大量小目标、丰富的场景, 可以提高检测模型的性能与泛化能力。然而, 这些数据集图像的拍摄视角是垂直地面的, 且图像背景色彩多以灰色、黑色、深绿色为主。因此, 这些数据集与 UAV 视角下的场景有所区别, 不利于 UAV 视角下 YOLO 检测模型的实验论证。

上述公共数据集可以用于模型预训练、测试、验证, 提高模型的泛化能力[39]。与自然拍摄视角不同, UAV 航拍图像由于视野更宽, 可能包含更多的目标物体, 且物体的尺寸比自然图像中小得多[38]。UAV 视角下的常用数据集包括:

表 4　常用公共数据集

Table 4　Common public datasets

类别	大小/(像素×像素)	类别	数量	拍摄视角
VOC	300×500	20	33 000	自然视角
COCO	600×800	80	330 000	自然视角
DIOR	800×800	20	23 463	航空视角
DOTA	800×4 000	15	11 268	航空视角
VisDrone	2 000×1 500	12	10 209	UAV视角
UAVDT	1 080×540	8	83 000	UAV视角
UAV123	1 280×720	12	110 000	UAV视角
UA-DETRAC	960×540	4	140 000	UAV视角
Staford Drone	1 417×2 019	10	19 000	UAV视角
AU-AIR	1 920×1 080	9	32 000	UAV视角

1) VisDrone 数据集, 包含 288 个视频片段, 由 261 908 帧视频和 10 209 幅静态图像组成。

2) UAVDT 数据集, 包含 100 个视频序列, 80 000 多帧, 大约 2 700 辆汽车。

3) UAV123 数据集, 由阿卜杜拉国王科技大学于 2016 年发布, 共包含 123 个视频序列和 110 000 多帧图像。

4) UA-DETRAC 数据集, 有 24 个不同场景的视频, 超过 140 000 帧, 包括不同天气环境下的车辆图像。

5) Staford Drone 数据集, 通过 UAV 收集斯坦福大学校园的航拍图像数据集, 含约 60 个航拍视频, 包含 10 种目标类型。

6) Aarhus 大学航空影像与遥感 (Aarhus university-aerial imagery and remote sensing, AU-AIR) 数据集包含 8 段视频, 共计 3 282 个提取帧。

这些公共数据集中的图像都是从 UAV 搭载的相机拍摄的视频中按帧采集的, 视频的数量少, 因此, 很多的图像差异性不大, 场景相似的图像较多, 这些相似图像不利于目标检测模型的应用与实验。

表 5 给出了部分 YOLOv5～YOLOv9 的 4 个不同版本检测模型在上述 4 种 UAV 视角下数据集上的性能指标。可以看出, 版本越高, 检测性能平均精度均值(mean average precision, mAP)也越高。每次实验数据从 UAV 视角下数据集中随机选取, 其中, 训练集数据是 4 000 张图像, 验证集为 1 000 张。

实验配制如表 6 所示, 本文后续实验也使用了相同的配制。除公共数据集外, 各个行业的具体需求催生了很多特定应用数据集。表 7 中展示了近

表5 YOLO检测模型在UAV视角下的实验性能

Table 5 Experimental performance of the YOLO detection model under UAV vision

算法	数据集	mAP/%
YOLOv5s	VisDrone2019	34.5
YOLOv5s	UA-DETRAC	40.3
YOLOv5s	UAVDT	57.8
YOLOv7s	VisDrone2019	38.9
YOLOv7s	UA-DETRAC	45.9
YOLOv7s	UAVDT	60.1
YOLOv8s	VisDrone2019	39.8
YOLOv8s	UA-DETRAC	45.1
YOLOv8s	UAVDT	65.1
YOLOv9c	VisDrone2019	45.3
YOLOv9c	UA-DETRAC	54.6
YOLOv9c	UAVDT	66.3

表6 实验运行环境

Table 6 Experimental environment configuration

云服务器	GPU	内存/GB	Python	CUDA
https://www.mistgpu.com/	RTX3080 Ti 12 GB	30	3.8	11.3

些年来基于UAV视角下YOLO目标检测模型使用的训练与测试集[38,40-50]。

2.2 UAV视角下目标检测数据集的扩充

UAV目标检测应用数据集中的图像多是从视频中采集,且过滤了特征差异较小的连续帧。这就导致数据集的数据量小、多样性不足、目标数量较少、背景多样性差。为提升目标检测模型的检测准确性,增强目标检测模型的泛化能力,可以应用各种扩充方法增加数据量。扩充数据集增加了数据量,但场景相似程度也相对较高,而YOLO系列算法将对于图像的特征提取与特征融合相结合,有利于相似图像特征的区分,因此,许多数据扩展方法的出现提高了YOLO目标检测算法的检测准确性。

扩充数据集的方式主要有2种:在已有的数据集上对图像进行各种变换处理;使用不同的拍摄设备以多种形式收集图像与视频。

2.2.1 基于图像变换的扩充

Andrew等[51]在研究奶牛检测的YOLOv2算法过程中,对UAV获取的图像进行随机合成,使扩充后的数据集包含11 384个合成和非合成图像。

Luo等[52]将1 978张UAV航拍图像应用随机旋转、镜像翻转、增加噪声与颜色变换方法扩充了4倍,再使用YOLOv3算法进行车辆检测模型训练与测试。Han等[40]使用UAV搭载的数码相机拍摄200张尺寸较大的钢结构表面裂纹图像构成原始数据集。由于钢结构表面裂缝区域在原始图像中的比例相对较小,很难对裂缝识别,因此,利用超像素分割方法对UAV图像进行预处理,将原始数据集扩充为4 000幅图像用于YOLOv3算法的训练与测试。

鲍文霞等[42]研究UAV自主着陆系统时,运用二维码图像作为着陆识别标志,标志识别算法使用YOLOv5m。为增加数据集的泛化性并提高检测准确性,在拍摄过程中,通过改变同一二维码的拍摄光照、摆放场景来丰富数据集。

2.2.2 基于数据收集的扩充

现阶段UAV可搭载的拍摄设备有数码相机、红外相机、多光谱相机、全景相机等,如图18所示。成像技术有可见光成像、红外成像、多光谱成像、全景摄像头成像等,其中,可见光成像是UAV最常使用的。

表7 UAV视角下的应用数据集

Table 7 Application datasets of UAV vision

文献	图像大小/(像素×像素)	类别	数量	成像特点	UAV拍摄高度/m
[41]	5 472×3 648	地面控制点	1 745	可见光	40～220
[40]	5 472×3 078	钢结构裂纹	200	可见光	10～25
[42]	640×640	快速响应码	3 260	可见光	<12
[42]	800×600	小麦	300	可见光	<3
[43]	4 000×3 000	输电线路	2 600	可见光	50～260
[44]	1 330×1 000	枯死树	3 084	可见光	75～100
[45]	500×500	松材线虫病	4 208	可见光	<525
[46]	1 152×864	绝缘子	860	可见光	<100
[38]	640×512	行人、车辆	2 898	红外热图	60～130
[47]	1 920×1 680	行人	8 416	LED光源	15～65
[48]	960×5 184	松材线虫病	4 500	多光谱图像	100～350
[49]	800×800	行人	7 092	鱼眼摄像头	<6
[50]	2 448×2 048	伪装目标	2 215	偏振成像	10



(a) 数码相机　(b) 红外相机　(c) 多光谱相机　(d) 全景相机

图 18　UAV 可搭载的拍摄设备

Fig. 18　UAV capable shooting equipment

不同的拍摄设备丰富了 UAV 视角下图像的表现形式, 使图像具有多种模式。因此, 目标检测模型要适应不同模态的图像数据。为适应不同类别、不同模态的 UAV 图像数据集, 通常 YOLO 检测模型会有所改进。

1) 红外成像。与传统光学图像相比, 红外热图像具有空间分辨率低、信噪比小、纹理特征少的特点, 可以在白天和晚上拍摄。在夜间、极端天气条件下, 因为光线太暗、物体遮挡、飞行高度太高, 相机很难捕捉到目标物体, 因此, 基于可见光条件下采集的数据集被应用于训练目标检测模型, 通常性能表现不佳。

Hinzmann 等[53] 将可见光图与热红外图像结合训练改进的 YOLOv3, 最终降低了 UAV 检测模型对行人的错误识别率, 提高夜间检测性能。

UAV 在夜间或在恶劣天气飞行中检测目标物体时, 易产生误检、漏检。为解决该问题, Wang 等[44] 使用红外图像集训练 YOLOv5s 改进版, 提升了模型对夜间目标的检测能力。

Mokayed 等[54] 使用蛋白化技术对 UAV 航拍的晴朗天气条件下车辆图像进行变换, 用以模拟雪、雨、雾等天气条件下的车辆。结合蛋白化车辆图像与原图像训练 YOLOv5 与 YOLOv8, 最终使 2 个模型在恶劣天气检测车辆的准确性分别提高了 30% 和 17%。

2) LED 光源成像。为提高夜间检测性能, 也可以使用特殊光源来增强图像中目标物体的特征信息。硅基金黄光 LED 光源只含有长波长的红光和黄光成分, 在雨、雾、沙尘、雾霾等恶劣天气条件下具有较好的穿透能力, 这使物体即使在恶劣的天气条件下也能保持较高的清晰度和亮度。因此, Zhu 等[47] 在 UAV 上搭载硅基金黄光 LED 光源, 扩充了原始训练集, 设计了 UAV 视角下的 YOLOv3 行人检测模型, 实验验证了该模型在夜间环境下检测行人具有显著优势。

3) 多光谱成像。单一的基于可见光图像或基于多光谱图像的识别方法, 会受到同物异谱、同谱异物现象的影响, 导致各种误检和漏检, 影响识别精度。多光谱相机可以获取多光谱图像, 包括更多的光谱波段, 如绿色、红色、近红外和红外边缘图像信息。多光谱成像将特征光谱信息划分为多个光谱波段, 可以从多个光谱通道获取更多信息, 揭示隐藏在较窄光谱范围内的特征。Qin 等[48] 在 UAV 上搭载了多光谱相机和可见光相机拍摄松木树林, 通过将可见光图像与多光谱图像相结合构成数据集。由于健康树和感染树在红色波长区域表现出最显著的差异, 因此, 多光谱成像更容易区分目标物体与背景。基于该数据集提出了一种松树萎蔫病 YOLOv5 检测模型, 能够排除树种、植被、光秃秃的地面和阴影的干扰, 准确地检测到受感染的松树。

4) 全景摄像头成像。UAV 在着陆时要避免着陆路径上的潜在危险, 而全景摄像头可为 UAV 在选择着陆点时提供比较安全的视角, 但全景摄像头采集的鱼眼图存在着陆点目标变形与扭曲, 增加了目标检测难度。Thuan 等[49] 改进了目标检测 YOLOv7 算法, 在 COCO 数据集进行预训练, 在 UAV 全景数据集上对模型微调, 微调前对全景图像做旋转处理增加训练图像数量, 最终提高了模型对着陆点检测的性能。

5) 偏振光成像。偏振成像不仅可以获得目标的强度信息, 而且可以获取附加的偏振信息[50]。偏振可以自适应增强图像中目标的对比度, 同时, 背景杂波会导致成像性能严重下降, 利用偏振对图像中的背景信息进行抑制, 可以有效地保留目标信息。当目标与周围环境相似时, 基于可见光图像的检测方法难以区分。为提高模型对伪装目标的检测能力, 沈英等[55] 在 YOLOX 的特征提取部分添加 3D 卷积, 提高了偏振特征提取能力, 引入特征增强模块和跨层级特征聚合网络, 将含有 UAV 拍摄的伪装目标偏振图像作为模型的输入, 最终改进的 YOLOX 在识别伪装目标检测中的 mAP 提高了 11.0%。

上述 UAV 视角下数据扩充方法的主要特点与优势如表 8[40-42,45-50,52-56] 所示。

2.3　UAV 视角下目标检测训练图像的预处理

虽然 UAV 视角下目标检测模型的数据集有很多, 在训练模型过程中, 这些数据集的大小仍然是有限的。第 2.2 节的方法可以增加数据量, 但还不足以解决问题。在模型训练前期, 对现有数据进行预处理成为提高模型检测能力的主要方法之一。训练中数据预处理能生成与现有数据同分布的训练输入数据, 用以解决数据量少、数据场景单一的问题, 提高模型泛化能力。现有的图像预处理方法有针对单幅图像的, 也可以利用多幅图像样本生成新的图像以丰富目标与场景性能, 其预处理作用如表 9[41,53,57-61] 所示。

表8 UAV视角下的数据扩充方法特性

Table 8 Advantages of data augmentation methods from UAV vision

文献	扩充方法	优势	作用
[52]	随机合成	增加数据	提高泛化能力
[41]	随机旋转、镜像翻转、增加噪声与颜色变换方法	增加数据、改变目标形态、增加背景	提高泛化能力与检测准确性
[40]	超像素分割方法	突出目标	提高检测准确性
[42]	改变拍摄光照、地点	丰富背景	提高泛化能力与检测准确性
[45-50,53-56]	改变成像模式	丰富目标背景	提高泛化能力;减少漏检、误检
[48]	多光谱成像	增加数据、突出目标	减少漏检、误检

表9 图像预处理方法特性

Table 9 Characteristics of image preprocessing methods

文献	方法	作用
[41,53,57]	数据增强	增强模型鲁棒性、提高准确性和召回率
[58]	SAHI	提高小目标的可见性
[59]	生成对抗网络	突出目标特征
[60]	图像模式变换	提高模型检测准确率、召回率
[61]	图像压缩	提升检测速度

在YOLOv4～YOLOv8版本的训练过程中都有数据增强模块,数据增强在早期UAV目标检测的应用已经出现过。

2019年,Andrew等[51]随机对训练图像进行不同增强处理,包括水平和垂直翻转、裁剪、仿射变换、高斯模糊、平均或中位数模糊、噪声添加、背景变化、对比度变化等任意组合,再使用扩充数据集训练YOLOv2。该方法是在小幅度改变数据特征的情况下,允许有限的数据产生更多的目标前景与背景,使模型对UAV视角下图像具有更强的鲁棒性。赵爽等[56]为解决UAV航拍图像下的多尺度车辆检测问题,将大尺寸航拍图像切分成若干有重叠区域的小图像块输入YOLOv2进行训练与检测,使检测模型相比原YOLOv2检测算法的平均精度(average precision, AP)提高了约8%。

Wang等[58]应用改进的生成对抗网络来融合可见光和红外图像,增强了UAV图像中行人目标信息,在此基础上训练YOLOv3行人检测模型,模型的召回率与准确性都有所提高。

为提高网络检测速度,可以使用压缩后的图像,但压缩后的图像必然会损失一部分目标物体的特征,降低识别的准确性,尤其是针对图像中的小目标物体。Coletta等[61]在ImageNet数据集上训练YOLOv4目标检测模型任务时,提出了以模型检测精度为变量的图像JPEG压缩级别函数。模型根据性能的要求预先设定好对应的JPEG压缩级别,不仅可以满足精度的需求,也可以实现检测速度的提升。在该理论基础上,设计了UAV视角下YOLOv4野生动物快速检测模型。

鲍文霞等[42]分析了UAV视角下的麦穗分布稠密、重叠现象严重、背景信息复杂等特点,基于Retinex算法[57]提出一种图像反射光增强方法,弱化入射图像,可以增强图像中的目标区域特征,提高了YOLOv5改进算法检测物体的准确性和召回率。

Wang等[36]利用切片辅助超推理(slicing aided hyper inference, SAHI)算法对训练数据集进行预处理,使小目标具有相对较大的可见性。同时,使用数据增强策略增加小物体的训练权值,提高模型对小物体信息的感知能力。Tian等[35]在研究UAV视角下小目标检测问题过程中,基于VisDrone数据集,比较了SAHI对于YOLOv5与YOLOX目标检测模型性能提升的作用。说明了SAHI可以大幅提高目标检测模型的性能,也指出了运用该方法,YOLOv5目标检测模型性能优于YOLOX。

Chen等[60]用自适应裁剪算法对高分辨率训练数据集进行增强,用以提高目标检测模型对小物体的检测能力。HSV色彩空间的亮度在区分饱和度或色调信息方面能起到重要的作用,同时,HSV图像模式比RGB提供了更多的可感知信息。Wang等[36]在目标检测模型训练过程中将图像由RGB转成HSV,有效提高了目标物体的检测准确率和召回率。改进后的YOLOX目标检测模型在VisDrone数据集的检测性能指标mAP0.5提高了8%。

2.4 模型的训练与推理

YOLO是应用数据训练出来的模型,再应用于实际任务进行推理,因此,训练策略可以提高模型的性能。深度学习模型通常假设训练数据和测试数据来自同一领域,并遵循相同的分布。然而,这种假设在现实场景中并不总是成立,如自动驾驶、UAV跟踪识别任务中。UAV视角下目标检测模型的实际应用研究中,训练数据和测试数据之间普遍存在着场景分布不同,这种差异性必将影响模型的检测能力[62]。YOLO系列算法中很多模型的训练与推理策略可以有效地解决这些问题,如迁移学习、多尺度学习、多分辨率训练。

将学习到的预训练模型泛化到未知的目标领域, 用预训练模型中初始化良好的权重作为基础, 在特定的训练集中对其微调, 可以提高目标检测能力。陈蕊等[63]先在仿真数据集上训练 YOLOv5 目标检测模型, 随后利用真实数据集对模型进行微调训练, 使 UAV 视角下目标检测的性能提升了 2.7%。鲍文霞等[42]在麦穗检测模型训练过程采用迁移学习, 先用田间采集小麦图像数据集对基于 Transformer 预测头的 YOLO(Transformer prediction head-YOLO, TPH-YOLO)模型进行训练得到预训练模型, 再冻结 TPH-YOLO 骨干网络预训练模型参数, 使用 UAV 采集的小麦图像数据集对模型进行参数更新和优化训练。该种训练方法在保证运算速度的前提下提高了检测精度, TPH-YOLO 的 mAP 比 YOLOv5 提高 4.1%。

使用 UAV 摄像机获取的图像分辨率通常较大, 且不同摄像机获得的图像分辨率也不同。因此, 可使用一些大型公共数据集进行目标检测模型的预训练, 这些数据一般比 UAV 视角下图像分辨率小, 例如, COCO 数据集的分辨率是 640 像素×360 像素。如果数据集中包含的图像大小是单一的, 则模型提取的特征大小也相应固定。为增强目标检测模型的鲁棒性, 需要增加多尺度数据集训练来弥补[64]。

Rančić等[65]认为在训练中使用多尺度图像能够使模型更容易学习到目标物体的差异特征, 保证模型能够学习到目标物体的不同尺度特征。魏湧明等[66]先采用 ImageNet 数据集对 Darknet-19 进行预训练, 再采用低分辨率的 UAV 航拍图像分类数据集, 微调 Darknet-19, 使网络调整每层权重以适应高分辨率输入。经过多尺度训练的 YOLOv2 网络对不同尺度的检测数据集具有很强的适应性。Keles 等[67]也提出了相似的结论, 并提出了 UAV 视角下鹿种群识别模型 YOLOv4-tiny, 为能够达到较好的检测效果, 建议将训练时的图像分辨率设为 512 像素×512 像素与 608 像素×608 像素。

UAV 视角下的目标检测模型多数是面向视频的, 视频的每一帧是一个静态图像, 视频则是由这些静态图像按照一定的帧顺序播放而成, 因此, 目标检测时可以逐帧推理或间隔帧推理。目标物体可能会出现在连续几帧图像内, 因此, 视频内目标检测与定位可以应用多帧图像的相近信息。Afifi 等[68]根据当前帧中包含物体的位置, 提出下一帧中的裁剪位置。将彼此靠近的框合并在一起, 形成一个单独的裁剪。通过仅在包含对象的区域位置进行推理, 可减少对图像背景部分的计算推理。该做法不仅可以提高检测的准确性, 而且可以提高推理速度。Yavariabdi 等[69]设计了一种多尺度 UAV 检测卷积神经网络 YOLOv3-tiny, 该架构应用 Inception Black[70]从输入视频中提取目标物体的局部和全局特征。每 6 帧检测一次, 将检测到的物体作用于中间帧, 提高了视频检测的能力。

3　UAV 视角下 YOLO 网络结构优化

YOLO 的版本多、可扩展性、易移植性使其更适用于 UAV 视角下目标检测的不同任务需求。在 UAV 视角下目标检测各个应用领域, 为适应不同场景与行业需求, 会选择不同的 YOLO 版本, 并基于该 YOLO 进行网络结构优化来提高 UAV 目标检测模型的性能。

3.1　特征提取优化策略

YOLO 目标检测算法的性能极大地依赖于主干网络对图像中目标物体特征的提取与融合能力, 因此, 在研究 UAV 视角下 YOLO 目标检测模型的过程中, 主干网络的优化成为主要改进方向。UAV 视角下图像中小目标特征信息少, 大多数现有算法的小目标识别与定位精度都不高[71]。

3.1.1　特征提取模块优化

为深化目标特征的提取, Liu 等[72]运用时空交互模块替代 YOLOv5 的 C3 模块, 并在 PAN 结构中引入递归门控卷积, 通过浅层信息挖掘和特征融合来提高小目标检测性能。改进后的 UAV 视角下 YOLO 目标检测模型在 VisDrone 数据集上的检测 mAP 较原来提高了 3.4%。陈范凯等[73]利用 C2f 模块替换 YOLOv5 的所有 C3 模块。优化后的目标检测模型 YOLOv5s_UAV 在增加少量参数的条件下, 较大地提升了模型的识别能力。该目标检测模型相较于 YOLOv5s 在数据集 VisDrone 上的 mAP 提高了 10.7%。冒国韬等[74]提出一种基于多尺度特征提取模块的 UAV 图像目标检测模型 多尺度自适应 YOLO(multi-scale adaptive YOLO, MSA-YOLO)。多尺度特征提取模块如图 19 所示, 其中, 4 个卷积运用不同大小的卷积核, 在 YOLOv5 的特征提取模块中应用多尺度卷积方式来获取不同尺度的特征信息。该目标检测模型在 VisDrone 数据集上的 mAP 较 YOLOv5 提升了 3.8%。

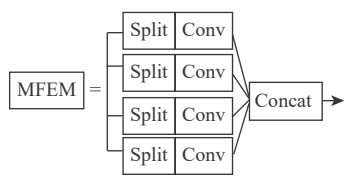


图 19　多尺度特征提取模块

Fig. 19　Multi-scale feature extraction module

注意力机制的核心思想是模仿人类视觉系统的选择性感知机制，对输入图像的特征权重进行动态调整，使模型集中在图像中最重要的区域，并抑制特征不相关部分[75]。

在YOLO目标检测模型中增加注意力模块必定增加模型的大小，但可以提高模型的目标检测能力。Yang等[76]改进了YOLOv4中的跨阶段局部X（cross stage partial X, CSPX）模块，在CSPX模块中增加了通道注意力机制压缩与激励网络（squeeze and excitation networks, SENet）与空间注意力机制空间变换网络（spatial Transformer networks, STNet）相结合的双注意力网络（dual attention network, DANet）结构。改进后的YOLOv4在识别UAV视角下输电线路绝缘子、鸟巢、间隔器和冲击锤较原YOLOv4模型的mAP值提高了2%。

为提高模型对相似目标的区分能力，减少物体的检测错误，李校林等[43]提出并行选择性注意力机制（parallel selective attention mechanism, PSAM），动态融合了特征的空间信息和通道信息。将这些模块增加到YOLOv5的特征提取中，虽然模型检测速度下降了15.0%，但检测精度提高了11.5%。Thuan等[49]为提高目标检测模型对图像特征的跨通道信息、位置信息的感知能力，将卷积块注意力模块（convolutional block attention module, CBAM）和坐标注意力（coordiate attention, CA）机制相结合，再应用于YOLOv5的特征提取模块；为减少模型参数量、提高检测速度，运用GhostNet[77]改进了C3，改进后的YOLOv5目标检测模型在自制UAV视角下松材线虫树数据集上的mAP提高到0.98%，与原YOLOv5模型相比，检测速度提高了1倍。奉志强等[78]为提高UAV目标检测模型对小目标特征的提取能力，提出一种融合了空间注意力与通道注意力的空间通道注意力模块（space channel attention module, SCAM），该模块相比于CBAM参数量更少，更关注小目标空间维度特征。在YOLOv5的C3后增加SCAM，改进后目标检测模型在VisDrone数据集上的mAP增加了12.1%，参数量只增加了约1×10^4。

Wang等[44]将高效的动态稀疏关注机制BiFormer[79]集成到YOLOv8骨干网络模型的C2f之前，提高了模型对特征图中关键信息的关注，进一步优化了检测性能。在VisDrone数据上训练后，模型的mAP增加了2.2%。

3.1.2 特征融合模块优化

张宏鸣等[80]在YOLOv3的特征融合模块后加入了STNet用于增强玉米幼苗的空间位置特征，提高UAV视角下目标检测模型对玉米幼苗的检测能力。

林文树等[81]将跨阶段部分网络（cross stage partial network, CSPNet）应用于YOLOv4模型的特征融合模块，且在计算卷积前加入SENet以增加感受野信息改变网络的深度、宽度，最终实现模型的缩放，提高了检测精度。改进的YOLOv4-CSP模型在UAV视角下松树木数据集上检测病虫害树的正确率为97.5%，相比于YOLOv4的平均检测正确率提高3.4%，检测速度提高了1.5倍。

Zhang等[82]为突出风电叶片表面长条缺陷的特征提取能力，在YOLOv5特征提取模型中增加了CBAM，改进的目标检测模型在自制的UAV视角下风电叶片数据集上的mAP提高了9.5%，参数量增加了17.3%。

Wang等[36]在YOLOX的PAN模块增加了超轻量级子空间注意力模块（ultra-lightweight subspace attention module, ULSAM），增强了不同感受野特征图之间信息的相关性，抑制背景信息，有效提升了Neck模块对小目标特征的融合，增强了网络局部感知目标的能力。

Zhao等[83]在YOLOv7的最大池化卷积（max pooling convolution, MPConv）层后增加了无参数的注意力机制（parameter-free attention mechanism, PFAM），突出浅层网络中的目标特征，弱化不相关信息。改进的YOLOv7目标检测模型被应用于UAV视角下海面人员、船只和漂浮物检测时，模型能有效地检测出海面上的小型目标。Zhao等[84]为提升UAV视角下小目标物体的检测能力，在YOLOv7的PAN模块中加入CBAM注意力机制，并引入Swin Transform模块[85]，改进后的目标检测模型在VisDrone数据集上的mAP提高了9.5%，但参数量增加了11.6%。

3.1.3 特征提取与特征融合的混合优化

丛玉华等[86]改进了YOLOv5的主干网络，在CSP模块后增加了CBAM注意力机制，同时，在Neck中使用加权双向FPN，如图20所示，同一尺寸的特征图上增加了跳跃连接，使特征融合时能够获得更多的信息。上采样和下采样之后均引入CBAM模块增加模型的收敛速率，同时，增强了不同目标的特征。韩俊等[87]将混合域注意力模块增加到YOLOv5的主干网络中，该模块关注特征图通道间的关系与空间位置信息。改进后的YOLOv5目标检测模型在VisDrone数据集上训练后，相比原模型的mAP提高了13.8%。李子豪等[88]为提高UAV对密集小目标检测的准确性，提出了一种自适应协同注意力机制（adaptive coordinate attention module, ACAM）嵌入到YOLOv5的主干网络中。该模块将输入特征通过并联的空间和通道注意力

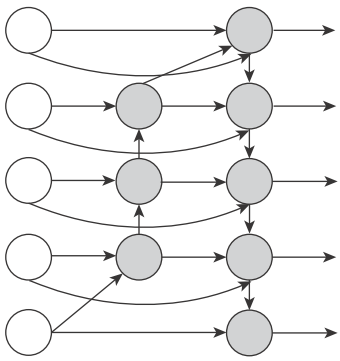


图 20　加权双向 FPN 结构

Fig. 20　Weighted bidirectional FPN structure

模块, 再将 2 个注意力权重自适应加权得到最终的注意力权重, 增强了特征的细节信息和语义信息。在 VisDrone 数据集上训练的改进 YOLOv5 与原模型相比, mAP 提升 1.5%。

表 10 归纳了在 YOLO 目标检测中增加注意力机制后的改进算法较改进前目标检测精度提升情况[36,44,48,76,78,80,82-84,86-89]。表中数据说明, 增加注意力机制可以有效提高 UAV 视角下目标检测的准确性。

在特征提取与特征融合模块中增加注意力机制或应用注意力机制替换原模型中某一个模块是 2 种重要的 YOLO 检测模型改进方法。为验证这些方法对于 YOLO 检测模型性能的改进能力, 共做了 12 次的实验, 实验中应用 CBAM 改进原 YOLO 检测模型, 实验数据集是从 VisDrone2019 中随机选取的 5 000 张图像, 训练与测试数据量为 4∶1, 结果可参考表 11。由表 11 的 mAP 指标可知, 在 3 种 YOLO 检测模型中增加注意力机制要优于替换原模型某个模块, 但增加注意力机制对于模型的大小会有一定的提高。因此, 本节中很多文献应用注意力机制优化 YOLO 检测模型时, 要结合自身的应用需求与 UAV 视角下数据集特点选择适当的被替换模块、增加位置、优化改进注意力机制或多种方法的融合。

3.2　YOLO 检测头结构优化策略

YOLO 采用多尺度检测可以提高不同大小目标的检测能力, 使用 UAV 航拍收集到的数据集通常包含许多非常小的目标, 使现有的网络检测非常困难。通常可以增加 1 个或 2 个预测头, 使其具有更小的感受野。因此, YOLO 的 FPN-PAN 多头结构可以缓解对象尺度变化的影响, 降低小目标检测网络的优化难度, 减少严重的物体尺度方差带来的负面影响。文献 [37,90-91] 在 UAV 视角下目标检测模型采用 4 个检测输出的检测头结构如图 21 所示; 文献 [81] 提出了有 5 个检测输出的检测头结构。

表 10　UAV 视角下增加注意力机制的 YOLO 算法性能提升度

Table 10　Performance improvement of YOLO algorithm with an added attention mechanism from UAV vision

算法	数据集	AP提升量/%	注意力机制	位置
YOLOv4[76]	高压输电线数据集	2	DANet	Backbone
YOLOv4[82]	虫害数据集	3.4	SENet	Backbone
YOLOv5[78]	VisDrone	11.5	PSAM	Backbone
YOLOv8[44]	VisDrone	2.2	BiFormer	Backbone
YOLOv3[80]	幼苗数据集	0.1	SAM	Neck
YOLO3[89]	棕榈树数据集	1.06	CBAM	Neck
YOLOv5[82]	风叶缺陷	4	CBAM	Neck
YOLOX[36]	VisDrone	5	ULSAM	Neck
YOLOv7[83]	ODv2	7	PFAM	Neck
YOLOv7[84]	VisDrone	9.5	CBAM	Neck
YOLOv5s[86]	VisDrone	2	CBAM	Backbone、Neck
YOLOv5[87]	VisDrone	7.9	SENet	Backbone、Nec
YOLOv5[78]	VisDrone	4	SCAM	Backbone、Neck
YOLOv5[48]	松木数据集	5	CBAM	Backbone、Neck
YOLOv5[88]	VisDrone	1.5	ACAM	Backbone、Neck

表 11　CBAM 对 YOLO 算法性能影响

Table 11　Impact of CBAM on the performance of YOLO algorithm

算法	位置	置入方式	mAP/%
YOLOv5s	Backbone+Neck	替换C3	31.6
YOLOv5s	Backbone	替换C3	32.4
YOLOv5s	Neck	替换C3	33.8
YOLOv5s	Backbone	增加	37.6
YOLOv7s	Backbone+Neck	替换ELAN	32.7
YOLOv7s	Backbone	替换ELAN	37.9
YOLOv7s	Neck	替换ELAN	39.0
YOLOv7s	Backbone	增加	39.8
YOLOv8s	Backbone+Neck	替换C2f	33.8
YOLOv8s	Backbone	替换C2f	34.9
YOLOv8s	Neck	替换C2f	38.6
YOLOv8s	Backbone	增加	40.1

多个检测头会增加模型的计算成本、增加模型尺寸, 不利于模型的部署。如果目标检测模型在不受小目标的影响下, 可以减少检测头, 使模型尺寸变小[89]。Wu 等[92] 设计了轻型 YOLOv4-tiny 用于 UAV 在高空观测火焰和烟雾场景。模型采用了 13 像素×13 像素和 26 像素×26 像素这 2 个尺度 FPN-PAN 双头结构, 如图 22 所示。

自 YOLOX 版本后, 检测头多采用解耦头, 可能会导致 2 个任务之间的不对称, 即分类得分最高的样本和位置得分最高的样本, 其定位分数可能相差较大, 导致目标检测中潜在的不准确[33]。

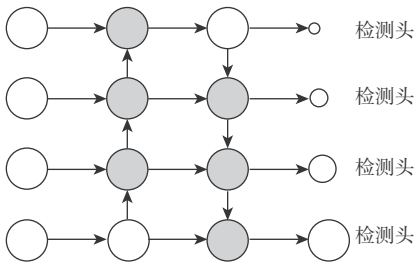


图 21 FPN-PAN 4 头结构

Fig. 21 Four headed structure of FPN-PAN

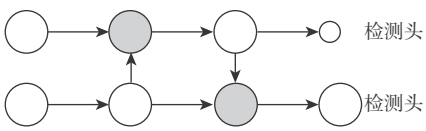


图 22 FPN-PAN 双头结构

Fig. 22 Double headed structure of FPN-PAN

吴明杰等[93]改进了 YOLOX 的解耦头，直接输出目标边界框的位置信息及目标类别的概率分布。应用该解耦头改进 YOLOv5，在 VisDrone 数据集上训练后，mAP 提高了 8.8%。陈旭等[94]提出一个新的解耦头模块替换 YOLOv5 的检测耦合头，新的解耦头可分阶段地执行不同的任务，第 1 阶段完成回归任务，第 2 阶段借助跨阶段卷积模块协助完成分类任务，防止因检测头特征共用引起回归与分类任务的混淆，提高了细类别的分类精度。

在 UAV 视角下，图像中目标长宽比与检测模型中的边界框长宽比可能达不到有效近似，从而使得目标检测模型精度下降。为解决该问题，Zhou 等[95]设计了基于位置编码与像素重排的轻量化检测头 PixSF-Head，如图 23 所示，其中，IFCOUS 为特征图的 FCOUS 结构逆运算。以 PixSF-Head 为基础重新设计了 YOLOX 的辅助头与引导头，其中，引导头是一个蒸馏层，使用辅助头中的标签分配结果替换引导头的标签分配结果，以提高引导头整体的表达能力，而辅助头只在训练过程中执行，没有推理成本。改进后的 YOLOX 参数量增加了 6×10^5，模型在 VisDrone 数据集上训练后的 mAP 较原版本提高了 9.8%。

Zhao 等[96]设计了一种新的重参数化对齐(reparameterization align, RepA)解耦头，通过分类和回归之间信息连接来对齐 2 个任务，可以有效缓解 2 个分支任务的不对齐，最终获得更高的检测精度。改进后的 YOLOX-s 在 COCO 数据集上的 mAP 比原版本提高了 2.3%，参数量只增加了 6×10^4。

为解决 UAV 航拍图像中交通目标检测时存在的水平包围框与目标真实轮廓贴合度较差、目标的水平包围框重叠度高导致相互抑制的问题，曲优等[97]提出一种基于锚框变换的单阶段旋转目标检测模型，改进 YOLOv3 的检测头，应用 2 次回归将水平锚框变换为旋转锚框，表示旋转检测框是 5 个参数张量，即 (x, y, w, h, θ)。在原有的边界框上增加一个角度参数。

3.3 UAV 视角下 YOLO 网络结构尺寸优化

为提高 YOLO 运算速度，减小模型的尺寸，通常可使用网络结构量化、网络剪枝、替换卷积、知识蒸馏和轻量级网络等。

3.3.1 YOLO 网络结构的量化

网络结构量化是指减少模型参数的代码长度。Wang 等[98]提出一种低比特量化方法来构建高效的单级检测模型，该方法是基于单边直方图的激活量化方法，通过最小化均方误差量化模型改进了 YOLOv5s，将该模型参数量化为 8 位时。在 COCO 数据集进行测试，平均精度值几乎没有下降，改进的模型计算速度显著加快，量化后模型的尺寸从 5.76×10^7 降低到 1.44×10^7。

3.3.2 基于 BN 层的网络剪枝

网络剪枝是指对网络结构进行剪枝，其主要目的是调整模型的一些属性，生成不同尺度的模型，以满足不同推理速度与部署的需要。YOLO 多个版本卷积后使用了 BN 层，如下所示：

$$\hat{x}_i=\frac{x_i-\mu}{\sqrt{\sigma^2+\varepsilon}}\gamma+\beta \tag{7}$$

式中：x_i为输入样本；μ为样本均值；σ^2为样本方差，$\varepsilon>0$为一个极小项，γ和β为需要学习的参数，$\hat{x}_i$为输出样本。

BN 层虽然起到了积极的作用，但增加了网络层数，占用了大量的计算时间。针对 BN 层缺点，Liu 等[99]提出了一种简单有效的通道剪枝方法。直接采用 BN 层尺度因子γ作为剪枝指标，如果一个通道包含了对目标分类很重要的信息，其对应的γ系数受到损失函数的刺激而变大；如果一个通道包含与分类无关的信息，则在损失函数的影响下，其系数γ会变小。因此，根据剪枝率对具有较小γ的卷积通道进行相应数量的剪枝，最终得到一个紧凑的网络。

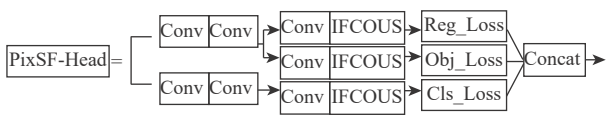


图 23 PixSF-Head 结构

Fig. 23 PixSF-Head structure

Shao 等[100]运用这种网络剪枝方法改进了 YOLOv3 模型, 获得更紧凑、参数更少的目标检测模型, 在此基础上设计了航空红外 YOLOv3(aerial infrared YOLOv3, AIR-YOLOv3)目标检测算法。在自制的红外航拍数据集上训练后, AIR-YOLOv3 与 YOLOv3 相比, mAP 仅下降了 1.7%, 但体积压缩接近 35.5%, 计算量减少了约 67%。

Qin 等[89]在运用 YOLOv3 算法检测 UAV 视角下单一目标棕榈树的过程中, 为提高检测速度, 减小模型尺寸, 也使用上述剪枝网络改进了 YOLOv3 模型。

Zhang 等[82]在检测 UAV 视角下风电叶片缺陷的研究过程中, 为降低模型的复杂性和参数量, 应用同样的 BN 尺度因子剪枝网络, 改进的 YOLOv5 检测模型推理计算少, 检测速度快。上述层剪枝与通道剪枝方法可以减少参数量, 但经多次剪枝后, 剪掉的层与通道可能也包含有用的特征信息, 影响检测准确性。因此, 侯鑫等[101]在训练中将迭代稀疏层剪枝与迭代稀疏通道剪枝相结合对模型进行剪枝处理, 剪枝后再通过微调来补偿剪枝带来的精度损失, 最终使模型的参数量更少、精度损失更小。

解决 BN 层缺点另一种有效的方法是 Xia 等[90]提出的卷积层和 BN 层融合, 即将卷积层和 BN 层融合为一个卷积层。如下所示:

$$\hat{x} = wx_i + b \tag{8}$$

$$y_i = \frac{\hat{x}_i - \mu}{\sqrt{\sigma^2 + \varepsilon}}\gamma + \beta = w_{\text{new}}x_i + b_{\text{new}} \tag{9}$$

$$w_{\text{new}} = \frac{w\gamma}{\sqrt{\text{Var}[\hat{x}] + b + \varepsilon}} \tag{10}$$

$$b_{\text{new}} = \frac{b - E[\hat{x}]}{\sqrt{\text{Var}[\hat{x}] + b + \varepsilon}} + \beta \tag{11}$$

式中: w为原始卷积权重; b为原始卷积层的偏置参数; $E[\hat{x}]$和 $\text{Var}[\hat{x}]$分别为均值和方差; w_{new}和b_{new}分别为新卷积层的权重和偏置参数。

融合后的网络结构可以删除 BN 层, 并使用新的权值和偏置进行卷积运算, 从而获得与原来相同的批归一化效果。这种将卷积层和 BN 层融合的方法在 YOLOX、YOLOv7 中也都有应用。

3.3.3　基于卷积的网络剪枝

深度学习模型大量地使用了卷积运算, 卷积运算增加了模型的尺寸与运算速度, 选择不同的卷积也可以改变 YOLO 检测模型的尺寸。每个卷积核对应于一个通道, 可以通过拆分空间维度和通道深度的相关性, 减少卷积计算所需要的参数量[102]。

为减少目标检测模型的参数量, 彭继慎等[103]使用深度可分离卷积(depthwise separable convolution, DSConv)改进了 YOLOv4 的 CBL 模块, 改进的目标检测模型 D-YOLOv4 检测速度提高了约 15%, 模型大小减少约 7.3%。

Wu 等[92]在研究 UAV 视角下火焰和烟雾区域检测任务过程中, 为减少算法的参数量和计算量, 使用 DSConv 替换 YOLOv4 特征融合模块中的标准卷积, 改进后的 YOLOv4 算法在自制火焰和烟雾数据集上较原算法的 mAP 增加了 3.14%, 参数量约为原算法的 10.1%。

Sharma 等[104]使用八度卷积替换了 YOLOv4 的标准卷积, 且在特征提取模块中加入了 CBAM 模块, 改进的目标检测模型提高了复杂背景下交通标志的检测速度、准确率。在自制的 UAV 视角下交通标志数据集上训练后, 模型检测的平均精度较 YOLOv4 提高了 3.5%, 检测速度提高到 12.5 帧/s。

陈卫彪等[105]使用 DSConv 替代 YOLOv5 标准卷积, 虽然网络层数增加了, 但整个网络的体积及参数量却降低了。改进的 YOLOv5 目标检测模型在 VisDrone 数据集上训练后, 模型的大小只有原来的 12.5%。

陈旭等[94]运用混合残差空洞卷积模块优化了 YOLOv5 骨干网络结构, 该卷积模块与普通卷积相比, 具有较少的计算量与较大的特征感受野。在 VisDrone 数据集上训练后的改进 YOLOv5 目标检测模型与原模型比较, 检测平均精度提升了 1.4%, 推理速度提升了 21.4%。

DSConv 在运算过程中会损失一部分特征信息, 从而降低模型的准确性, 而分组混洗卷积(group shuffle convolution, GSConv)结合了标准卷积和 DSConv 的优点。因此, Saydirasulovich 等[106]在 YOLOv8 模型中采用了 GSconv 与 DSConv 并行处理, 该方法没有完全切断通道间的连接, 而是选择在很大程度上保留这些连接, 从而保证了模型的检测准确性。在自定义的 UAV 视角下野火烟雾数据集训练后模型对野火烟雾目标检测 AP 提高了 4.3%。

3.3.4　基于知识蒸馏的轻量化

网络结构量化、网络剪枝、卷积替换对于模型的检测精度有一定的影响, 因此, 张可等[107]采用知识蒸馏对经过剪枝后的 YOLOv4 网络进行微调, 如图 24 所示。使用知识蒸馏可以进一步实现 1.6%的精度增益, 这表明直接训练学生网络时, 学生网络不可避免地会学习到部分冗余特征, 而利用教师网络指导训练可以在一定程度上起到强迫作用, 从而使学生网络学习到更多的重要知识。引入知识蒸馏后的 YOLOv4 模型有助于损失的降低, 加快网

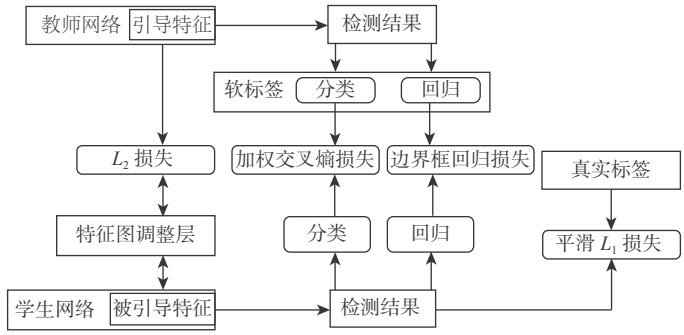


图 24 知识蒸馏架构

Fig. 24 Knowledge distillation structure

络收敛速度。改进后的YOLOv4模型在UAV视角下防振锤缺陷检测mAP提高了165.6%，视频检测速度由204帧/s提高到了217帧/s。

3.3.5 基于轻量级网络的优化

YOLO的骨干网络是算法核心部分，其对检测模型的检测速度、模型体积影响很大，因此，轻量级的骨干网络可以减小YOLO模型的体积，提高检测速度。

姜尚洁等[45]在研究UAV视角下玉米幼苗实时检测模型时，先对玉米幼苗图像进行增强处理；再使用MobileNetV3[108]改进YOLOv3的特征提取模块；检测头使用2个尺度检测玉米幼苗，改进后的目标检测模型特征增强YOLO(feature enhancement YOLO, FE-YOLO)在自制的玉米幼苗数据集上的计算量和推理时间较原YOLOv3模型分别提高了7.91%和33.76%。

罗潇等[46]将UAV拍摄的输电高压线的绝缘子分为正常与缺陷2类。考虑到目标检测任务中类别较少，选用ResNet18作为主干网络改进YOLOv3，能够避免网络出现过拟合，同时，对特征融合模块的特征金字塔执行上采样操作，将其与主干网络中的不同尺度的残差块进行融合，实现绝缘子检测的实时与快速处理。

为提取更多的目标有效特征，提高检测的精度与速度，黄丽明等[109]应用MobleNetV2[110]替换了YOLOv4的CSPDarkNet53，其中，模型的卷积使用DSConv。优化后的模型参数更少、结构尺寸变小，改进的YOLOv4在UAV视角下松木数据集上训练后，虽然检测精度增加不大，但模型的参数量仅为原模型的20%，模型检测速度提高了29.1%。

丛玉华等[86]为提升检测能力与检测速度，使用CSPDarkNet53-tiny替换YOLOv4的CSPDarkNet53，如图25所示。检测头使用通道缩减解耦头结构，改进的YOLOv4在VisDrone数据集上训练后，模型的推理时间较原模型减少了8 ms，mAP提升了7%。表12给出了UAV视角下一些YOLO目标检测模型的网络结构优化方法与作用[98,101-103,105-108,110-114]。



图 25 CSPDarkNet53-tiny网络结构

Fig. 25 CSPDarkNet53-tiny network structure

表 12 网络结构尺寸优化与作用

Table 12 Optimization and role of network structure size

文献	方法	方法作用
[98]	参数量化	减小模型大小
[101-103]	剪枝BN层	剪枝层、剪枝通道大小
[105-108]	替换卷积	剪枝卷积参数量
[107]	知识蒸馏	提高推理速度
[110-114]	替换小型网络结构	减小模型大小、提高推理速度

甄然等[111]使用单个ConvNeXt Block结构[112]替换YOLOv7主干网络的ELAN模块，提出一种更加轻量化的主干网络，改进后的主干网络既节省计算资源，又提升了模型检测精度。

3.4 UAV视角下YOLO目标检测模型的损失函数

在UAV拍摄的图像中，很多目标尺寸较小，这些小目标在图像中所占面积比例远小于背景，而YOLO新的版本会侧重于对小目标检测能力的提高，使YOLO成为UAV视角下目标检测模型的主要选择。小目标物体特征信息少导致了数据集中正、负样本的不平衡，因此，损失函数成为UAV视角下YOLO目标检测模型的主要改进之一。

Yu等[113]为增强YOLOv4在复杂背景下检测裂缝的能力，利用焦损失(focal loss, FL)作为置信度损失函数。FL是一个动态缩放的交叉熵损失，通过动态缩放因子，可以自适应地降低训练过程中易区分样本的权重，从而聚焦难区分的样本，可以解决模型中目标与背景分类不平衡和正负样本分配不平衡的问题。FL求解如下：

$$\varphi_{\mathrm{FL}}=\begin{cases}-\alpha(1-y')^{\eta}\ln y' & y=1\\ -(1-\alpha)y'^{\eta}\ln(1-y') & y=0\end{cases} \tag{12}$$

式中：$y=1$为正样本，$y=0$为负样本；y'为激活函数

的输出; α因子用于平衡正负样本比例不均匀; η因子的作用是减少易分类样本的损失, 使损失函数更加关注难分类和错误分类的样本。

引入 FL 替换损失函数后, 改进 YOLOv4 的 mAP 比原模型提高了 6.6%。

陈范凯等[73]在研究 UAV 视角下绝缘子检测问题中, 由于绝缘子目标尺寸较小, 导致数据集中正、负样本的不平衡, 采用 FL 代替 YOLOv4 置信度损失中的标准交叉熵函数, 将类别损失函数替换为平衡交叉熵损失函数, 真实框回归损失函数保持不变。改进后模型的绝缘子检测平均准确率有了明显提高。

为提高 UAV 遥感图像的小目标检测性能, 张上等[114]在式 (11) 的基础上进行了改进, 得到分类得分函数变焦损失 (varifocal loss, VFL):

$$\varphi_{\mathrm{VFL}}=\begin{cases}-q[q\ln p+(1-q)\ln(1-q)] & q>0\\ -\alpha p^{\eta}\ln(1-p) & q=0\end{cases} \quad (13)$$

式中: p为目标分数的分类预测值; q为分类条件。

若为目标类, 则将正样本类的q值设置为预测框和真实框之间的 IOU 值, 否则设置为 0; 若为背景类, 则将q值设置为 0。通过p^{η}缩放因子对负样本进行了处理, 突出正样本的贡献。将 VFL 应用于 YOLOv5 的分类损失函数, 在 VisDrone 数据集上训练后模型检测 mAP 提高了 1.4%。

在目标检测中, IOU 是预测框和真实框之间的 IOU, 其决定了预测框的选取, 因此, IOU 对于检测准确性起到了非常重要的作用。

Xia 等[90]应用广义交并比 (generalized intersection over union, GIOU) 对 YOLOv3 的损失函数做了改进, 在 COCO 数据集上训练改进后的 YOLOv3 较原模型的 mAP 提高了约 2.5%。

谢椿辉等[115]采用压缩卷积层与通道数改进解耦头替换了 YOLOv5 耦合头, 在提升精度的同时, 尽量减少模块带来的推理延时, 同时, 使用 Alpha-IoU[116]预测框位置的损失, 增强在数据较少及存在数据噪声的情况下模型检测效果的鲁棒性。改进的 YOLO 算法在 VisDrone 数据集上训练后, 浮点运算量增加了约 13%, 推理时间减少了 0.7 ms。当检测框与真实框不相交时, GIOU 仍然可以计算梯度, 梯度不会为零, 因此, GIOU 在一定程度上克服了 IOU 作为损失函数的缺点。

吴明杰等[93]为解决 YOLOv5 中完整交并比损失 (complete intersection over union loss, CIOU Loss) 计算回归结果不准确的问题, 使用聚焦式高效交并比 (focal and efficient intersection over union, Focal-EIOU) 损失函数, 将预测框的宽高拆分, 分别与最小外界框的宽高作差值运算, 从而提高模型对小型目标的检测精度。使用 Focal-EIOU 的改进检测模型与原 YOLOv5 相比, 虽然在浮点运算量的大小上有小幅度增加, 但在 VisDrone 数据集上的检测精度提高了 6.1%。

甄然等[111]用 SIOU (SCYLLA-IoU)[117]替换 YOLOv7 模型中的 CIOU, SIOU 加入角度向量来描述纵横比相同时预测框的坐标损失, 在训练阶段实现了更快的收敛, 且在推理方面具有更好的性能。改进的 YOLOv7 检测模型在 UAV 图像数据集中训练后, 相较于原模型的检测精度变化较小, 但参数量减少了 23.6%, 运算速度降为原来一半。

表 13 总结了这些方法的特点与作用[90,93,111,113-114,116], 可为后续 YOLO 目标检测模型的改进研究提供有价值的参考。

表 13　损失函数的优化方法特点

Table 13　Characteristics of optimization methods for loss functions

文献	解决问题	改进方法	作用	缺点
[113]	正负样本分配不平衡	FL	自适应地区分样本的权重	增加运算量
[114]	提高小目标检测能力	VFL	突出正样本的贡献	增加运算量
[116]	提升检测速度	Alpha-IOU	减少梯度为−1的风险, 增加收敛速度	增加运算量
[90]	提升检测精度	GIOU	减少梯度为0的风险	增加数据量
[111]	提升检测速度	SIOU	加入角度向量来描述纵横比, 增加收敛速度	精度没有得到有效提高
[93]	提高小目标检测能力	Focal-EIOU	拆分预测框的宽高	降低了推理速度

3.5　UAV 视角下 YOLO 目标检测模型的性能评价

模型训练过程中, 主要参数包括训练轮次、优化器中的学习率动量参数、权值衰减 (weight decay, WD)、初始学习率 (initial learning rate, ILR); 及实验使用的设备参数。表 14 给出了上述一部分 UAV 视角下 YOLO 目标检测模型实验性能[34-36,38,49,74,76,78,82,84,94-95,101,103,106,113]。可以看出, 在公共数据集上测试的目标检测准确性不是很高, 但在特定的应用数据集上对单个目标检测 mAP 平均约为 94%, 这说明现阶段的 YOLO 应用于 UAV 视角下目标检测模型中的目标检测准确性相对较好。

表14 UAV视角下YOLO目标检测模型的实验性能

Table 14 Experimental performance of YOLO object detection model from UAV vision

模型	数据集	mAP/%	mAP0.5/%	参数量	epoch	动量	WD	ILR	设备	帧率/(帧·s^{-1})
YOLOv4[38]	COCO	48	66		500	0.9	0.000 5	0.001	RTX2080Ti	19.4
YOLOv5[34]	COCO	45	65	2.48×10^7	300	0.9	0.000 5	0.01	TX1080Ti	82.2
YOLOX[95]	COCO	43	61	1.45×10^6	300	0.9	0.000 5	0.01	RTX3090	27.7
YOLOv3[103]	VisDrone	27		6.15×10^7	300	0.9	0.000 5	0.002 5	GTX1080Ti	139.7
YOLOv4[94]	VisDrone	32	65	6.17×10^6	300	0.9	0.000 5	0.01	2070Super	18.0
YOLOv5[74]	VisDrone	35		1.08×10^8					RTX3060	146.0
YOLOv7[78]	VisDrone	36	53	7.97×10^7	400	0.9	0.000 5	0.01	TESLA P100	
YOLOv3[101]	行人	91		2.28×10^8	500	0.9	0.000 5	0.001	GTX1660Ti	51.5
YOLOv3[36]	棕榈树	92		8.13×10^7	100	0.9	0.000 5	0.001	GTX1060	26.9
YOLOv4[76]	输电线路	90			300	0.949	0.000 5	0.001	RTX3090	50.0
YOLOv4[84]	松树病虫害	96		2.01×10^8	175	0.9	0.1	0.005	i7_10750H	4.2
YOLOv4[113]	钢裂纹	97		1.16×10^7	500	0.9	0.000 5	0.001	RTX1080Ti	8.4
YOLOv5[82]	风电叶片	95		3.78×10^7	300	0.9	0.000 5	0.003 2	GTX1650	41.3
YOLOv5[49]	松线虫树	98	98	4.67×10^7	500	0.9	0.000 5	0.01	RTX3060	67.0
YOLOv5[35]	枣椰树	92	90		300	0.9	0.000 5	0.001	RTX2080Ti	11.3
YOLOv8[106]	烟雾	92	85		200	0.9	0.000 5	0.001	GTX1080Ti	167.0

4 结束语

本文总结了YOLO系列算法的发展历程、相互关系、主要的优化策略,且运用统一流程绘制了其框架结构;梳理了近几年来UAV视角下目标检测YOLO系列算法的研究与应用成果,重点分析了目标检测模型的优化策略。从YOLO系列算法与UAV视角下目标检测应用2个角度进行了总结,有助于将YOLO算法的理论与应用进行比较;分析了UAV视角下目标检测YOLO算法的应用数据集、数据集扩充方法、预处理方法。尽管YOLO的版本越来越多,不同版本性能较之前有一定的提升,但仍无法满足UAV视角下的多元化任务需求。通过梳理,能为YOLO算法与UAV视角下目标检测技术的研究提供有价值的参考。

目前,UAV目标检测技术发展迅速、应用广泛,但其检测感知能力还不足以满足广泛应用与行业需求,还需要性能更为优越的YOLO目标检测模型与优化策略。根据当前UAV视角下的YOLO目标检测模型研究与应用现状,做出如下问题总结与研究展望:

1)数量更大、场景更丰富的UAV目标检测公共数据集。现有的UAV目标检测公共数据集中图像数量少,目标类别少,场景不够丰富,且YOLO系列算法在一些公共数据上的性能表现不佳,在VisDrone数据集上,YOLO目标检测的一些改进版本性能mAP的平均值约为0.35,而YOLO的10种基础版本在COCO数据集上mAP平均值约为0.43。使用这些公共数据集改进模型的测试与验证,很难充分检测模型的性能。如果使用迁移训练策略训练UAV目标检测模型,在这些公共数据集训练得到的低精度模型参数还需要更多地优化才能满足需求。因此,UAV视角下目标检测公共数据集的制作与完善需要研究人员或组织机构给予足够的关注与投入。

2)UAV视角下多目标检测YOLO算法。在UAV视角下YOLO目标检测应用分析中,多数应用是针对单一目标检测,检测精度较高。然而,多目标物体检测的精度往往不会很高,目标类别越多,检测模型平均精度会越低,因此,基于UAV视角下YOLO目标检测模型的多目标检测、多目标跟踪的能力有待进一步提升。

3)UAV视角下YOLO检测模型的多模态数据处理。由上述分析和总结可知,UAV采集的图像与视频易受到光源、环境、天气、设备等干扰,使图像数据具有同物异谱、同谱异物、遮挡、噪声干扰等多模态。因此,UAV视角下YOLO目标检测模型的改进要适应于不同模态的数据,确保目标检测的准确性与实时性。

4)轻量级UAV视角下YOLO算法与检测速度提升。提高小目标检测精度,通常在原有模型基础上增加更多的提取小目标特征的模块,这必将增加模型的复杂度与检测速度,因此,很多YOLO检测模型改进是权衡检测速度和检测精度。后续UAV视角下YOLO目标检测模型研究之一是寻求模型复杂度与检测精度的平衡。

5) YOLO 的特征信息共享机制。在 UAV 视角下 YOLO 目标检测模型使用的数据集中, 多数图像是在同一段视频中按帧采集, 相邻帧间的图像信息可能会有相关性。在目标检测过程中, 图像间的时间与空间信息共享可以提高目标检测模型的检测性能与检测速度。如何在 UAV 目标检测模型中应用图像空间或时间的共享信息来提高检测性能也是改进的研究方向。

6) 小目标检测能力提升。YOLO 现有的小目标检测性能难以满足各种现实应用, 如车辆无人驾驶系统、UAV 自动导航巡航系统, 这些应用对于小目标检测的准确性、实时性要求较高。从上述分析可见, UAV 视角下 YOLO 目标检测模型主要改进之一是提升模型对小目标的检测能力。因此, 今后 UAV 视角下 YOLO 检测模型的研究还会延续小目标检测性能的提升。

7) 多种特征提取技术融合。单纯依赖于 YOLO 解决 UAV 视角下目标检测任务可能达不到应用需求, 为提高模型的检测准确性与减小模型, 需将多种特征提取与融合技术和 YOLO 相结合。这些技术必将给 YOLO 的改进提供更好的研究方向。

8) UAV 视角下目标检测的多种感知技术融合。对目标的检测与识别仅仅运用图像识别技术可能达不到预期的效果, 因此, 可以结合目标所发出的声音, 使用音频识别技术提高对目标的定位、识别能力。

9) UAV 视角下旋转目标检测。YOLO 的基础版目标检测主要是找出物体在图像中的真实框, 真实框是一种矩形框, 矩形框的边平行于图像的边界。但实际上, UAV 在飞行过程中的姿态不一定能保证拍摄到的目标始终保持方向不变, 因此, YOLO 算法检测到的图像矩形框通常大于物体本身。用旋转后矩形框可能会更好地贴近目标, 有利于计算机视觉测量、目标定位、目标分类等任务。因此, 旋转矩形框检测正是当前 YOLO 目标检测的主要改进研究之一。

参考文献 (References)

[1] 蔡志浩, 陈文军, 赵江, 等. 基于动态视觉传感器的无人机目标检测与避障[J]. 北京航空航天大学学报, 2024, 50(1): 144-153.
Cai Z H, Chen W J, Zhao J, et al. Object detection and obstacle avoidance based on dynamic vision sensor for UAV[J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(1): 144-153(in Chinese).

[2] 薛康, 刘琨, 江俊峰, 等. 基于 YOLOv5s 模型的光纤振动传感事件精准检测研究[J]. 光学学报, 2023, 43(2): 0228001.
Xue K, Liu K, Jiang J F, et al. Optical fiber vibration sensing detection with high accuracy based on YOLOv5s model[J]. Acta Optica Sinica, 2023, 43(2): 0228001(in Chinese).

[3] 刘莉, 曹潇, 张晓辉, 等. 轻小型太阳能/氢能无人机发展综述[J]. 航空学报, 2020, 41(3): 1-28.
Liu L, Cao X, Zhang X H, et al. Review of development of light and small scale solar/hydrogen powered unmanned aerial vehicles[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(3): 1-28 (in Chinese).

[4] 张继贤, 刘飞, 王坚. 轻小型无人机测绘遥感系统研究进展[J]. 遥感学报, 2021, 25(3): 708-724.
Zhang J X, Liu F, Wang J. Review of the light-weighted and small UAV system for aerial photography and remote sensing[J]. Journal of Remote Sensing, 2021, 25(3): 708-724(in Chinese).

[5] 刘芳, 孙亚楠. 基于自适应融合网络的无人机目标跟踪算法[J]. 航空学报, 2022, 43(7): 359-369.
Liu F, Sun Y N. UAV target tracking algorithm based on adaptive fusion network[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(7): 359-369(in Chinese).

[6] 张博恒, 柴栋栋, 孟令博, 等. 基于多重检测的无人机抗遮挡目标跟踪算法[J]. 北京航空航天大学学报, 2023, 49(9): 2442-2454.
Zhang B H, Chai D D, Meng L B, et al. Anti-occlusion target tracking algorithm of UAV based on multiple detection[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(9): 2442-2454(in Chinese).

[7] Liu J H, Wang Z P, Wu Y P, et al. An improved faster R-CNN for UAV-based catenary support device inspection[J]. International Journal of Software Engineering and Knowledge Engineering, 2020, 30(7): 941-959.

[8] 刘英杰, 杨风暴, 胡鹏. 基于 Cascade R-CNN 的并行特征金字塔网络无人机航拍图像目标检测算法[J]. 激光与光电子学进展, 2020, 57(20): 201505.
Liu Y J, Yang F B, Hu P. Parallel FPN algorithm based on Cascade R-CNN for object detection from UAV aerial images[J]. Laser & Optoelectronics Progress, 2020, 57(20): 201505(in Chinese).

[9] Peng X, Zhong X G, Zhao C, et al. A UAV-based machine vision method for bridge crack recognition and width quantification through hybrid feature learning[J]. Construction and Building Materials, 2021, 299: 123896.

[10] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognitio. Piscataway: IEEE Press, 2016: 779-788.

[11] Li Y D, Dong H, Li H G, et al. Multi-block SSD based on small object detection for UAV railway scene surveillance[J]. Chinese Journal of Aeronautics, 2020, 33(6): 1747-1755.

[12] Liu J, Jia R, Li W, et al. High precision detection algorithm based on improved RetinaNet for defect recognition of transmission lines[J]. Energy Reports, 2020, 6: 2430-2440.

[13] Kannadaguli P. FCOS based human detection system using thermal imaging for UAV based surveillance applications[C]//Proceedings of the 2020 IEEE Bombay Section Signature Conference. Piscataway: IEEE Press, 2021: 79-83.

[14] 耿创, 宋品德, 曹立佳. YOLO 算法在目标检测中的研究进展[J]. 兵器装备工程学报, 2022, 43(9): 162-173.
Geng C, Song P D, Cao L J. Research progress of YOLO algorithm in target detection[J]. Journal of Ordnance Equipment En-

gineering, 2022, 43(9): 162-173(in Chinese).

[15] 邵延华, 张铎, 楚红雨, 等. 基于深度学习的YOLO目标检测综述[J]. 电子与信息学报, 2022, 44(10): 3697-3708.
Shao Y H, Zhang D, Chu H Y, et al. A review of YOLO object detection based on deep learning[J]. Journal of Electronics & Information Technology, 2022, 44(10): 3697-3708(in Chinese).

[16] 王琳毅, 白静, 李文静, 等. YOLO系列目标检测算法研究进展[J]. 计算机工程与应用, 2023, 59(14): 15-29.
Wang L Y, Bai J, Li W J, et al. Research progress of YOLO series target detection algorithms[J]. Computer Engineering and Applications, 2023, 59(14): 15-29(in Chinese).

[17] Terven J, Cordova-Esparza D. A comprehensive review of YOLO architectures in computer vision: from YOLOv1 to YOLOv8 and YOLO-NAS[EB/OL]. (2024-02-04)[2024-06-10]. https://arxiv.org/abs/2304.00501.

[18] Wu X, Li W, Hong D F, et al. Deep learning for unmanned aerial vehicle-based object detection and tracking: a survey[J]. IEEE Geoscience and Remote Sensing Magazine, 2022, 10(1): 91-124.

[19] 冷佳旭, 莫梦竟成, 周应华, 等. 无人机视角下的目标检测研究进展[J]. 中国图象图形学报, 2023, 28(9): 2563-2586.
Leng J X, Mo M J C, Zhou Y H, et al. Recent advances in drone-view object detection[J]. Journal of Image and Graphics, 2023, 28(9): 2563-2586(in Chinese).

[20] 罗旭东, 吴一全, 陈金林. 无人机航拍影像目标检测与语义分割的深度学习方法研究进展[J]. 航空学报, 2024, 45(6): 235-264.
Luo X D, Wu Y Q, Chen J L. Research progress on deep learning methods for object detection and semantic segmentation in UAV aerial images[J]. Acta Aeronautica et Astronautica Sinica, 2024, 45(6): 235-264(in Chinese).

[21] 苑玉彬, 吴一全, 赵朗月, 等. 基于深度学习的无人机航拍视频多目标检测与跟踪研究进展[J]. 航空学报, 2023, 44(18): 1-31.
Yuan Y B, Wu Y Q, Zhao L Y, et al. Research progress of UAV aerial video multi-object detection and tracking based on deep learning[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(18): 1-31(in Chinese).

[22] 邢姗姗, 赵文龙. 基于YOLO系列算法的复杂场景下无人机目标检测研究综述[J]. 计算机应用研究, 2020, 37(增刊2): 28-30.
Xing S S, Zhao W L. Review of UAV target detection in complex scenarios based on YOLO series algorithms[J]. Application Research of Computers, 2020, 37(Sup 2): 28-30(in Chinese).

[23] 杨欣, 王刚, 李椋, 等. 基于深度卷积神经网络的小型民用无人机检测研究进展[J]. 红外技术, 2022, 44(11): 1119-1131.
Yang X, Wang G, Li L, et al. Civil drone detection based on deep convolutional neural networks: a survey[J]. Infrared Technology, 2022, 44(11): 1119-1131(in Chinese).

[24] 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.

[25] Redmon J, Farhadi A. YOLOv3: an incremental improvement[EB/OL]. (2018-04-08)[2024-06-10]. https://arxiv.org/abs/1804.02767.

[26] Bochkovskiy A, Wang C Y, Liao H M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2024-06-10]. https://arxiv.org/abs/2004.10934.

[27] Wang C Y, Mark Liao H Y, Wu Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]//Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE Press, 2020: 1571-1580.

[28] 陈宇梁, 董绍江, 孙世政, 等. 改进YOLOv5s的弱光水下生物目标检测算法[J]. 北京航空航天大学学报, 2024, 50(2): 499-507.
Chen Y L, Dong S J, Sun S Z, et al. Improved YOLOv5s low-light underwater biological target detection algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(2): 499-507(in Chinese).

[29] Ge Z, Liu S T, Wang F, et al. YOLOX: exceeding YOLO series in 2021[EB/OL]. (2021-08-06)[2024-06-10]. https://arxiv.org/abs/2107.08430.

[30] Li C Y, Li L L, Jiang H L, et al. YOLOv6: a single-stage object detection framework for industrial applications[EB/OL]. (2022-09-07)[2024-06-10]. https://arxiv.org/abs/2209.02976.

[31] Wang C Y, Bochkovskiy A, Liao H M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[EB/OL].(2022-07-06)[2024-06-10].https://arxiv.org/abs/2207.02696.

[32] Wang C Y, Yeh I H, Liao H M. YOLOv9: learning what you want to learn using programmable gradient information[EB/OL]. (2024-02-29)[2024-06-10]. https://arxiv.org/abs/2402.13616.

[33] 纪守领, 杜天宇, 邓水光, 等. 深度学习模型鲁棒性研究综述[J]. 计算机学报, 2022, 45(1): 190-206.
Ji S L, Du T Y, Deng S G, et al. Robustness certification research on deep learning models: a survey[J]. Chinese Journal of Computers, 2022, 45(1): 190-206(in Chinese).

[34] Wang C X, Wang H R. Cascaded feature fusion with multi-level self-attention mechanism for object detection[J]. Pattern Recognition, 2023, 138: 109377.

[35] Tian G Y, Liu J R, Yang W Y. A dual neural network for object detection in UAV images[J]. Neurocomputing, 2021, 443: 292-301.

[36] Wang X, He N, Hong C, et al. Improved YOLOX-X based UAV aerial photography object detection algorithm[J]. Image and Vision Computing, 2023, 135: 104697.

[37] Nepal U, Eslamiat H. Comparing YOLOv3, YOLOv4 and YOLOv5 for autonomous landing spot detection in faulty UAVs[J]. Sensors, 2022, 22(2): 1-15.

[38] Suo J S, Wang T Y, Zhang X Z, et al. HIT-UAV: a high-altitude infrared thermal dataset for Unmanned Aerial Vehicle-based object detection[EB/OL]. (2023-03-31)[2024-06-10]. https://arxiv.org/abs/2204.03245.

[39] Zhan W, Sun C F, Wang M C, et al. An improved YOLOv5 real-time detection method for small objects captured by UAV[J]. Soft Computing, 2022, 26(1): 361-373.

[40] Han Q H, Liu X, Xu J. Detection and location of steel structure surface cracks based on unmanned aerial vehicle images[J]. Journal of Building Engineering, 2022, 50: 104098.

[41] Cheng C X, Yang J, Wang C, et al. Automatic detection of aerial survey ground control points based on YOLOv5-OBB[EB/OL]. (2023-03-06)[2024-06-10]. https://arxiv.org/abs/2303.03041.

[42] 鲍文霞, 谢文杰, 胡根生, 等. 基于TPH-YOLO的无人机图像麦穗计数方法[J]. 农业工程学报, 2023, 39(1): 155-161.
Bao W X, Xie W J, Hu G S, et al. Wheat ear counting method in UAV images based on TPH-YOLO[J]. Transactions of the Chinese Society of Agricultural Engineering, 2023, 39(1): 155-161

(in Chinese).

[43] 李校林, 刘大东, 刘鑫满, 等. 改进 YOLOv5 的无人机航拍图像目标检测算法[J]. 计算机工程与应用, 2024, 60(11): 204-214.
Li X L, Liu D D, Liu X M, et al. Target detection algorithm of UAV aerial image based on improved YOLOv5[J]. Computer Engineering and Applications, 2024, 60(11): 204-214(in Chinese).

[44] Wang G, Chen Y F, An P, et al. UAV-YOLOv8: a small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios[J]. Sensors, 2023, 23(16): 7190.

[45] 姜尚洁, 罗斌, 贺鹏, 等. 利用无人机多源影像检测车辆速度[J]. 测绘学报, 2018, 47(9): 1228-1237.
Jiang S J, Luo B, He P, et al. Vehicle speed detection by multi-source Images from UAV[J]. Acta Geodaetica et Cartog raphica Sinica, 2018, 47(9): 1228-1237(in Chinese).

[46] 罗潇, 於锋, 彭勇. 基于深度学习的无人机电网巡检缺陷检测研究[J]. 电力系统保护与控制, 2022, 50(10): 132-139.
Luo X, Yu F, Peng Y. UAV power grid inspection defect detection based on deep learning[J]. Power System Protection and Control, 2022, 50(10): 132-139(in Chinese).

[47] Zhu L, Xiong J H, Xiong F, et al. YOLO-Drone: airborne real-time detection of dense small objects from high-altitude perspective[EB/OL].(2023-10-11)[2024-06-10].https://arxiv.org/abs/2304.06925.

[48] Qin B X, Sun F G, Shen W X, et al. Deep learning-based pine nematode trees' identification using multispectral and visible UAV imagery[J]. Drones, 2023, 7(3): 183.

[49] Thuan P N, Westerlund T, Queralta J P. Vision-based safe autonomous UAV docking with panoramic sensors[EB/OL]. (2023-10-06)[2024-06-10]. https://arxiv.org/abs/2305.16008.

[50] Zhang J C, Shao J B, Chen J L, et al. Polarization image fusion with self-learned fusion strategy[J]. Pattern Recognition, 2021, 118: 108045.

[51] Andrew W, Greatwood C, Burghardt T. Aerial animal biometrics: individual Friesian cattle recovery and visual identification via an autonomous UAV with onboard deep inference[C]//Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2020: 237-243.

[52] Luo X, Tian X Y, Zhang H J, et al. Fast automatic vehicle detection in UAV images using convolutional neural networks[J]. Remote Sensing, 2020, 12(12): 1994.

[53] Hinzmann T, Stegemann T, Cadena C, et al. Deep learning-based human detection for UAVs with optical and infrared cameras: system and experiments[EB/OL]. (2020-08-10)[2024-06-10]. https://arxiv.org/abs/2008.04197.

[54] Mokayed H, Nayebiastaneh A, De K, et al. Nordic vehicle dataset (NVD): performance of vehicle detectors using newly captured NVD from UAV in different snowy weather conditions[EB/OL]. (2023-04-27)[2024-06-10]. https://arxiv.org/abs/2304.14466.

[55] 沈英, 刘贤财, 王舒, 等. 基于偏振编码图像的低空伪装目标实时检测[J]. 兵工学报, 2024, 45(5): 1374-1383.
Shen Y, Liu X C, Wang S, et al. Real-time detection of low-altitude camouflaged targets based on polarization encoded images[J]. Acta Armamentarii, 2024, 45(5): 1374-1383(in Chinese).

[56] 赵爽, 黄怀玉, 胡一鸣, 等. 基于深度学习的无人机航拍车辆检测[J]. 计算机应用, 2019, 39(S2): 91-96.
Zhao S, Huang H Y, Hu Y M, et al. Vehicle detection in satellite imagery based on deep learning[J]. Journal of Computer Applications, 2019, 39(S2): 91-96(in Chinese).

[57] Sun J, Yang Y, He X F, et al. Northern maize leaf blight detection under complex field environment based on deep learning[J]. IEEE Access, 2020, 8: 33679-33688.

[58] Wang C Q, Luo D, Liu Y, et al. Near-surface pedestrian detection method based on deep learning for UAVs in low illumination environments[J]. Optical Engineering, 2022, 61(2): 023103.

[59] Cui Q Q, Liu M, Huang X Y, et al. Coarse-to-fine visual autonomous unmanned aerial vehicle landing on a moving platform[J]. Biomimetic Intelligence and Robotics, 2023, 3(1): 100088.

[60] Chen Z W, Cao L J, Wang Q H. YOLOv5-based vehicle detection method for high-resolution UAV images[J]. Mobile Information Systems, 2022, 2022(1): 1828848.

[61] Coletta A, Giorgi F, Maselli G, et al. A2-UAV: application-aware content and network optimization of edge-assisted UAV systems[EB/OL]. (2023-07-24)[2024-06-10]. https://arxiv.org/abs/2301.06363.

[62] Koh P W, Sagawa S, Marklund H, et al. WILDS: a benchmark of in-the-wild distribution shifts[EB/OL]. (2021-07-16)[2024-06-10]. https://arxiv.org/abs/2012.07421.

[63] 陈蕊, 郑华飞, 蒋鸿宇, 等. 结合仿真迁移学习和自适应融合的无人机小目标检测[J]. 小型微型计算机系统, 2023, 44(8): 1743-1749.
Chen R, Zheng H F, Jiang H Y, et al. Combination of simulation-based transfer learning and adaptive fusion for UAV small object detection[J]. Journal of Chinese Computer Systems, 2023, 44(8): 1743-1749(in Chinese).

[64] Singh B, Davis L S. An analysis of scale invariance in object detection-SNIP[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 3578-3587.

[65] Rančić K, Blagojević B, Bezdan A, et al. Animal detection and counting from UAV images using convolutional neural networks[J]. Drones, 2023, 7(3): 179.

[66] 魏湧明, 全吉成, 侯宇青阳. 基于 YOLOv2 的无人机航拍图像定位研究[J]. 激光与光电子学进展, 2017, 54(11): 111002.1-111002.10.
Wei Y M, Quan J C, Hou Y Q Y. Aerial image location of unmanned aerial vehicle based on YOLOv2[J]. Laser & Optoelectronics Progress, 2017, 54(11): 111002.1-111002.10(in Chinese).

[67] Keles M C, Salmanoglu B, Guzel M S, et al. Evaluation of YOLO models with sliced inference for small object detection[EB/OL]. (2022-03-09)[2024-06-10]. https://arxiv.org/abs/2203.04799.

[68] Afifi M, Ali Y, Amer K, et al. Robust real-time pedestrian detection in aerial imagery on jetson TX2[EB/OL]. (2019-05-16)[2024-06-10]. https://arxiv.org/abs/1905.06653.

[69] Yavariabdi A, Kusetogullari H, Celik T, et al. FastUAV-NET: a multi-UAV detection algorithm for embedded platforms[J]. Electronics, 2021, 10(6): 1-9.

[70] Misbah M, Khan M U, Yang Z H, et al. TF-net: deep learning empowered tiny feature network for night-time UAV detection[EB/OL]. (2022-11-29)[2024-06-10]. https://arxiv.org/abs/2211.16317.

[71] 王兆魁, 方青云, 韩大鹏. 成像卫星在轨智能处理技术研究进展[J]. 宇航学报, 2022, 43(3): 259-270.

Wang Z K, Fang Q Y, Han D P. Research progress of on-orbit artificial intelligence technology for imaging satellite[J]. Journal of Astronautics, 2022, 43(3): 259-270(in Chinese).

[72] Liu H Y, Duan X H, Lou H T, et al. Improved GBS-YOLOv5 algorithm based on YOLOv5 applied to UAV intelligent traffic[J]. Scientific Reports, 2023, 13: 9577.

[73] 陈范凯, 李士心. 改进 YOLOv5 的无人机目标检测算法[J]. 计算机工程与应用, 2023, 59(18): 218-225.
Chen F K, Li S X. UAV target detection algorithm with improved YOLOv5[J]. Computer Engineering and Applications, 2023, 59(18): 218-225(in Chinese).

[74] 冒国韬, 邓天民, 于楠晶. 基于多尺度分割注意力的无人机航拍图像目标检测算法[J]. 航空学报, 2023, 44(5): 268-278.
Mao G T, Deng T M, Yu N J. Object detection in UAV images based on multi-scale split attention[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(5): 268-278(in Chinese).

[75] 祁宣豪, 智敏. 图像处理中注意力机制综述[J]. 计算机科学与探索, 2024, 18(2): 345-362.
Qi X H, Zhi M. Review of attention mechanisms in image processing[J]. Journal of Frontiers of Computer Science & Technology, 2024, 18(2): 345-362(in Chinese).

[76] Yang Z, Xu X F, Wang K K, et al. Multitarget detection of transmission lines based on DANet and YOLOv4[J]. Scientific Programming, 2021, 2021(1): 6235452.

[77] Han K, Wang Y H, Xu C, et al. GhostNets on heterogeneous devices via cheap operations[J]. International Journal of Computer Vision, 2022, 130(4): 1050-1069.

[78] 奉志强, 谢志军, 包正伟, 等. 基于改进 YOLOv5 的无人机实时密集小物体检测算法[J]. 航空学报, 2023, 44(7): 246-260.
Feng Z Q, Xie Z J, Bao Z W, et al. Real-time dense small object detection algorithm for UAV based on improved YOLOv5[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(7): 246-260(in Chinese).

[79] Zhu L, Wang X J, Ke Z H, et al. BiFormer: vision transformer with bi-level routing attention[EB/OL]. (2023-03-15)[2024-06-10]. https://arxiv.org/abs/2303.08810.

[80] 张宏鸣, 付振宇, 韩文霆, 等. 基于改进 YOLO 的玉米幼苗株数获取方法[J]. 农业机械学报, 2021, 52(4): 221-229.
Zhang H M, Fu Z Y, Han W T, et al. Detection method of maize seedlings number based on improved YOLO[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(4): 221-229(in Chinese).

[81] 林文树, 张金生, 何乃磊. 基于改进 YOLOv4 的落叶松毛虫侵害树木实时检测方法[J]. 农业机械学报, 2023, 54(4): 304-312.
Lin W S, Zhang J S, He N L. Real-time detection method of dendrolimus superans-infested Larix gmelinii trees based on improved YOLOv4[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(4): 304-312(in Chinese).

[82] Zhang R, Wen C B. SOD-YOLO: a small target defect detection algorithm for wind turbine blades based on improved YOLOv5[J]. Advanced Theory and Simulations, 2022, 5(7): 2100631.

[83] Zhao H Y, Zhang H P, Zhao Y Y. YOLOv7-sea: object detection of maritime UAV images based on improved YOLOv7[C]//Proceedings of the 2023 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops. Piscataway: IEEE Press, 2023: 233-238.

[84] Zhao L L, Zhu M L. MS-YOLOv7: YOLOv7 based on multi-scale for object detection on UAV aerial photography[J]. Drones, 2023, 7(3): 171-188.

[85] Liu Z, Lin Y T, Cao Y, et al. Swin Transformer: hierarchical vision Transformer using shifted windows[C]//Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 9992-10002.

[86] 丛玉华, 何啸, 邢长达, 等. 基于无人机的轻量化小目标检测网络[J]. 弹箭与制导学报, 2022, 42(6): 6-12.
Cong Y H, He X, Xing C D, et al. Lightweight small target detection network based on UAV[J]. Journal of Projectiles, Rockets, Missiles and Guidance, 2022, 42(6): 6-12(in Chinese).

[87] 韩俊, 袁小平, 王准, 等. 基于 YOLOv5s 的无人机密集小目标检测算法[J]. 浙江大学学报 (工学版), 2023, 57(6): 1224-1233.
Han J, Yuan X P, Wang Z, et al. UAV dense small target detection algorithm based on YOLOv5s[J]. Journal of Zhejiang University (Engineering Science), 2023, 57(6): 1224-1233(in Chinese).

[88] 李子豪, 王正平, 贺云涛. 基于自适应协同注意力机制的航拍密集小目标检测算法[J]. 航空学报, 2023, 44(13): 239-249.
Li Z H, Wang Z P, He Y T. Aerial-photography dense small target detection algorithm based on adaptive cooperative attention mechanism[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(13): 239-249(in Chinese).

[89] Qin Z W, Wang W S, Dammer K H, et al. A real-time low-cost artificial intelligence system for autonomous spraying in palm plantations[EB/OL]. (2021-03-06)[2024-06-10]. https://arxiv.org/abs/2103.04132.

[90] Xia Y, Ye G X, Yan S S, et al. Application research of fast UAV aerial photography object detection and recognition based on improved YOLOv3[J]. Journal of Physics: Conference Series, 2020, 1550(3): 032075.

[91] Liu X H, Wang C B, Liu L. Research on pedestrian detection model and compression technology for UAV images[J]. Sensors, 2022, 22(23): 9171.

[92] Wu R N, Hua C C, Ding W L, et al. Flame and smoke detection algorithm for UAV based on improved YOLOv4-tiny[C]//Proceedings of the PRICAI 2021: Trends in Artificial Intelligence. Berlin: Springer, 2021: 226-238.

[93] 吴明杰, 云利军, 陈载清, 等. 改进 YOLOv5s 的无人机视角下小目标检测算法[J]. 计算机工程与应用, 2024, 60(2): 191-199.
Wu M J, Yun L J, Chen Z Q, et al. Improved YOLOv5s small object detection algorithm in UAV view[J]. Computer Engineering and Applications, 2024, 60(2): 191-199(in Chinese).

[94] 陈旭, 彭冬亮, 谷雨. 基于改进 YOLOv5s 的无人机图像实时目标检测[J]. 光电工程, 2022, 49(3): 210372.
Chen X, Peng D L, Gu Y. Real-time object detection for UAV images based on improved YOLOv5s[J]. Opto-Electronic Engineering, 2022, 49(3): 210372(in Chinese).

[95] Zhou W, Min X L, Hu R, et al. FasterX: real-time object detection based on edge GPUs for UAV applications[EB/OL]. (2022-09-07)[2024-06-10]. https://arxiv.org/abs/2209.03157.

[96] Zhao Z P, He C, Zhao G M, et al. RA-YOLOX: re-parameterization align decoupled head and novel label assignment scheme based on YOLOX[J]. Pattern Recognition, 2023, 140: 109579.

[97] 曲优, 李文辉. 基于锚框变换的单阶段旋转目标检测方法[J]. 吉

林大学学报 (工学版), 2022, 52(1): 162-173.
Qu Y, Li W H. Single-stage rotated object detection network based on anchor transformation[J]. Journal of Jilin University (Engineering and Technology Edition), 2022, 52(1): 162-173(in Chinese).

[98] Wang M Z, Sun H X, Shi J, et al. Q-YOLO: efficient inference for real-time object detection[EB/OL]. (2023-07-01)[2024-06-10]. https://arxiv.org/abs/2307.04816.

[99] Liu Z, Li J G, Shen Z Q, et al. Learning efficient convolutional networks through network slimming[C]//Proceedings of the 2017 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2755-2763.

[100] Shao Y H, Zhang X P, Chu H Y, et al. AIR-YOLOv3: aerial infrared pedestrian detection via an improved YOLOv3 with network pruning[J]. Applied Sciences, 2022, 12(7): 3627.

[101] 侯鑫, 曲国远, 魏大洲, 等. 基于迭代稀疏训练的轻量化无人机目标检测算法[J]. 计算机研究与发展, 2022, 59(4): 882-893.
Hou X, Qu G Y, Wei D Z, et al. A lightweight UAV object detection algorithm based on iterative sparse training[J]. Journal of Computer Research and Development, 2022, 59(4): 882-893(in Chinese).

[102] Guo X J, Li S Y, Yu J K, et al. PFLD: a practical facial landmark detector[EB/OL]. (2019-03-03)[2024-06-10]. https://arxiv.org/abs/1902.10859.

[103] 彭继慎, 孙礼鑫, 王凯, 等. 基于模型压缩的 ED-YOLO 电力巡检无人机避障目标检测算法[J]. 仪器仪表学报, 2021, 42(10): 161-170.
Peng J S, Sun L X, Wang K, et al. ED-YOLO power inspection UAV obstacle avoidance target detection algorithm based on model compression[J]. Chinese Journal of Scientific Instrument, 2021, 42(10): 161-170(in Chinese).

[104] Sharma V K, Dhiman P, Rout R K. Improved traffic sign recognition algorithm based on YOLOv4-tiny[J]. Journal of Visual Communication and Image Representation, 2023, 91: 103774.

[105] 陈卫彪, 贾小军, 朱响斌, 等. 基于 DSM-YOLOv5 的无人机航拍图像目标检测[J]. 计算机工程与应用, 2023, 59(18): 226-233.
Chen W B, Jia X J, Zhu X B, et al. Target detection for UAV image based on DSM-YOLOv5[J]. Computer Engineering and Applications, 2023, 59(18): 226-233(in Chinese).

[106] Saydirasulovich S N, Mukhiddinov M, Djuraev O, et al. An improved wildfire smoke detection based on YOLOv8 and UAV images[J]. Sensors, 2023, 23(20): 8374.

[107] 张可, 乐全明, 黄文礼, 等. 联合剪枝和知识蒸馏的 PCSA-YOLOs 防振锤缺陷检测算法[J]. 电工电能新技术, 2022, 41(11): 59-69.
Zhang K, Le Q M, Huang W L, et al. PCSA-YOLOs vibration damper defect detection algorithm combined with pruning and knowledge distillation[J]. Advanced Technology of Electrical Engineering and Energy, 2022, 41(11): 59-69(in Chinese).

[108] Howard A, Sandler M, Chen B, et al. Searching for MobileNetV3[C]//Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2019: 1314-1324.

[109] 黄丽明, 王懿祥, 徐琪, 等. 采用 YOLO 算法和无人机影像的松材线虫病异常变色木识别[J]. 农业工程学报, 2021, 37(14): 197-203.
Huang L M, Wang Y X, Xu Q, et al. Recognition of abnormally discolored trees caused by pine wilt disease using YOLO algorithm and UAV images[J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37(14): 197-203(in Chinese).

[110] Sandler M, Howard A, Zhu M L, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 4510-4520.

[111] 甄然, 刘雨涵, 孟凡华, 等. 基于改进 YOLOv7 的低空飞行物目标检测方法[J]. 无线电工程, 2024, 54(3): 633-643.
Zhen R, Liu Y H, Meng F H, et al. Low altitude flying target detection method based on improved YOLOv7[J]. Radio Engineering, 2024, 54(3): 633-643(in Chinese).

[112] Liu Z, Mao H Z, Wu C Y, et al. A ConvNet for the 2020s[C]//Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2022: 11966-11976.

[113] Yu Z W, Shen Y G, Shen C K. A real-time detection approach for bridge cracks based on YOLOv4-FPM[J]. Automation in Construction, 2021, 122: 103514.

[114] 张上, 张岳, 王恒涛, 等. 轻量化无人机遥感图像小目标检测算法[J]. 无线电工程, 2023, 53(10): 2329-2336.
Zhang S, Zhang Y, Wang H T, et al. Lightweight small target detection algorithm in UAV remote sensing image[J]. Radio Engineering, 2023, 53(10): 2329-2336(in Chinese).

[115] 谢椿辉, 吴金明, 徐怀宇. 改进 YOLOv5 的无人机影像小目标检测算法[J]. 计算机工程与应用, 2023, 59(9): 198-206.
Xie C H, Wu J M, Xu H Y. Small object detection algorithm based on improved YOLOv5 in UAV image[J]. Computer Engineering and Applications, 2023, 59(9): 198-206(in Chinese).

[116] He J B, Erfani S, Ma X J, et al. Alpha-IoU: a family of power intersection over union losses for bounding box regression[EB/OL]. (2022-06-22)[2024-06-10]. https://arxiv.org/abs/2110.13675.

[117] Elfwing S, Uchibe E, Doya K. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning[J]. Neural Networks, 2018, 107: 3-11.

Research progress on YOLO series algorithms for target detection from UAV vision

CHEN Jinlin，WU Yiquan*，YUAN Yubin

(College of Electronic and Information Engineering，Nanjing University of Aeronautics and Astronautics，Nanjing 211106，China)

Abstract：The you only look once (YOLO) series of deep learning-based object identification algorithms has several advantages, including great adaptability, high detection accuracy, and quick detection speed. Make it the preferred choice for object detection in drone aerial images. Due to the complex and ever-changing nature of aerial images and videos captured by drones, the YOLO series object detection algorithms from the UAV vision have become a challenging problem. Therefore, it is necessary to comprehensively summarize and analyze the recent research and application achievements of YOLO series algorithms for object detection from the UAV vision. Firstly, the YOLO structural diagram is redrawn and the development, connection, and optimization features of the 10 fundamental versions of YOLO are summarized. The roles and objectives of various YOLO optimization strategies are pointed out, indicating the direction for further theoretical research. Secondly, the optimization methods of YOLO in the application field of drone object detection are summarized. This includes preprocessing before data training, optimization of training and inference, application of attention modules, convolutional optimization, detection heads, and optimization of loss functions. At the same time, some application datasets and data expansion methods from the UAV vision are sorted out. Through studies, the effectiveness of YOLO series algorithms on the UAV vision dataset and the influence of convolutional block attention module (CBAM) on various YOLO improvements are assessed. Finally, it is pointed out that there are still some problems and challenges in the research of YOLO series algorithms for object detection from the UAV vision. This article looks forward to possible development trends and research directions in this field.

Keywords：UAV system；object detection；deep learning；YOLO series algorithms；datasets；performance comparison

Received：2024-06-12；**Accepted**：2024-07-26；**Published Online**：2024-08-26 11：10

URL：link.cnki.net/urlid/11.2625.V.20240824.1220.001

Foundation items：National Natural Science Foundation of China (61573183); Anhui Provincial University Research Plan Project (2022AH051584)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2026.0255

水下重力辅助惯性导航关键技术综述与展望

周召发[1]，李易依[1]，张志利[1,*]，陈坡[2]，边英磊[2]，李新宇[1]
(1. 火箭军工程大学 智能与控制学院，西安 710025；2. 中国人民解放军 96416 部队，洛阳 471000)

摘 要：水下高精度自主导航是海洋探索与开发的关键核心技术。惯性导航系统 (INS) 的定位误差随航行时间累积发散，需要外部信息辅助修正。地球重力场作为地球固有属性，兼具稳定性、自主性与无源性，是理想的信息源，由此发展而来的重力辅助惯性导航技术具有重要的研究价值。在分析系统基本原理与技术架构的基础上，围绕重力基准图构建、重力场适配性分析、重力匹配算法、航迹规划四大关键技术，系统梳理了多源数据融合、网格化插值、特征评估与适配区提取、序列相关匹配、非线性滤波、智能路径优化等主流方法的研究进展，深入剖析了当前技术发展中尚未解决的共性问题。针对现阶段研究存在的技术瓶颈展望未来研究发展方向，旨在为后续构建高精度、高可靠性和高自主性的水下重力辅助惯性导航系统 (GAINS) 提供理论参考与技术指引。

关 键 词：水下自主导航；重力辅助惯性导航系统；重力基准图；适配性分析；重力匹配算法；航迹规划
中图分类号：U666.12; P223.7
文献标志码：A **文章编号**：1001-5965(2026)10-3535-28

近年来，随着大国博弈愈演愈烈，各国之间的资源争夺战已然进入白热化阶段，海洋安全显得愈发重要，其中，海洋探索和海洋资源开发成为关乎国家安全与发展的重要战略。水下航行器作为执行各类海洋任务不可或缺的载体设备，对于具备高精度、隐蔽性和长航时的水下导航需求显得十分迫切[1-2]。水下导航定位技术旨在为提供水下运载体提供精准、可靠的定位信息，是海洋开发利用、资源勘探及海域综合管控等任务顺利实施的关键基础。

然而，水下环境的复杂性与传统导航手段的局限性，使水下导航技术的发展面临诸多挑战，如何实现水下高精度导航定位一直是世界性技术难题[3-4]。当前，在水下环境中稳定全时接收全球卫星导航系统(global navigation satellite system, GNSS)信号非常困难，惯性导航系统(inertial navigation system, INS)因其自主性、连续性和全天候等特点，契合水下导航对隐蔽性和可靠性的要求，已成为水下自主定位的主要导航方式[5-6]。惯导系统是一种航位推算系统，利用传感器解算真实位置、速度和姿态，能够实现短时高精度定位，但惯导解算误差会随时间不断累积，存在长航时导航定位精度低的缺陷，因此，需要依靠外部信息辅助修正[7-8]。目前，常用的辅助 INS 的水下自主导航技术主要包括地形辅助导航、地磁辅助导航和重力辅助导航技术等。其中，地形辅助导航需要用声呐主动发射声波测量海底的轮廓，属于有源导航模式[9-10]；地磁辅助导航性能受限于地磁传感器精度和不可预测的磁干扰等外部因素[11-12]；而地球重力场作为地球相对稳定的重要固有属性，不易受气候、海浪等不确定复杂环境的影响而发生变化，具有良好的时空分布特征，且无需向外部发射或接收信号，因此，地球重力场数据可作为辅助惯性导航的可靠数据源，已被

收稿日期：2026-05-01；录用日期：2026-07-17；网络出版时间：2026-07-23 10:11
网络出版地址：link.cnki.net/urlid/11.2625.V.20260723.0944.001
*通信作者. E-mail：zzl202@hhu.edu.cn

引用格式：周召发，李易依，张志利，等. 水下重力辅助惯性导航关键技术综述与展望 [J]. 北京航空航天大学学报，2026，52（10）：3535-3562. Zhou Z F，Li Y Y，Zhang Z L，et al. Overview and prospect of key technologies for underwater gravity-aided inertial navigation [J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（10）：3535-3562 （in Chinese）.

应用于重力辅助惯性导航系统(gravity-aided inertial navigation system, GAINS)[13-14]。该系统利用重力匹配算法,将载体上搭载的高精度重力传感器所采集的航迹重力信息,与重力基准图中的背景信息进行比对,依据契合程度确定最佳匹配点,从而校正INS积累的定位误差。核心组成部分包括INS、高精度重力测量传感器、重力基准图及重力匹配算法。该技术兼具自主性、隐蔽性和稳定性,具有重要的研究价值和广泛的应用前景[15]。

近年来,精密测量科学的飞速发展和测高卫星数据处理方法的渐趋成熟为重力辅助导航技术的发展和应用奠定了坚实的基础。重力基准图构建、重力场适配性分析、重力匹配算法和航迹规划算法等关键技术共同构成了重力辅助导航从数据支撑、区域识别、核心定位到路径优化的完整技术链条。本文对上述四方面关键技术的研究现状进行梳理,并分析当前存在的不足,根据实际应用需求展望未来发展趋势,期望对该领域未来的研究重点、研究方向及可行的研究方法提供有价值的指引导向。

1 重力基准图构建技术

图1和图2分别展示了重力辅助导航系统的基本原理和完整技术链条。其中,重力基准图作为重力匹配定位的空间参照和数据基础,其分辨率、精度和覆盖范围直接决定了后续匹配结果的可靠性。因此,构建高空间分辨率和高精度的全球海洋重力场基准图是提升水下重力匹配辅助导航性能的关键。当前获取海洋重力异常的技术手段主要有卫星测高数据反演、航空重力测量和船载重力测量[10]。船载重力测量精度高,但覆盖稀疏,一般只能用于区域重力测量[16];航空重力测量兼顾分辨率与效率,但成本高昂,难以大范围推广;卫星重力测量可实现全球覆盖,但其空间分辨率相对较低[17]。当前卫星测高技术可获取全球大部分海域的重力异常信息,主要模型包括美国斯克里普斯海洋研究所Sandwell团队发布的S&S系列[18-20]、丹麦科技大学Andersen团队发布的DTU系列[21-24]。最新的S&S系列模型为V33.1版本(2026年1月发布),其在部分海域精度优于2×10^{-5} m/s^2,被国际学术界公认为当前分辨率最高、噪声控制最优的全球海洋重力场模型之一。该模型可在线公开获取,但覆盖范围严格限定于开放海洋及冰缘区,不包括格陵兰和南极大陆的永久冰盖内部。DTU系列最新版本为DTU21(2023年发布),该模型将重力场反演范围扩展至南北纬88°,显著改善了格陵兰和南极周边冰缘海域的数据覆盖,弥补了传统模型在高纬度冰缘区的覆盖空白,代表了当前基于多源卫星测高数据融合的全球重力场建模前沿。然而,该数据集目前尚未正式对外公开[25]。尽管上述两类模型的最新成果已将网格分辨率提升至1′×1′,但卫星测高数据在近海岸地区受海浪干扰、雷达回波失真等因素制约,精度显著下降,难以直接满足高精度水下重力匹配导航对近岸区域重力场精细表征的需求。如何有效融合多源、多尺度、多精度的重力数据,以及如何通过插值算法生成连续、光滑且符合重力场物理特性的网格化基准图,是重力基准图构建中2个核心技术环节。分别对多源重力数据融合方法和重力数据网格化插值算法的研究现状进行梳理,主要方法如图3所示,并指出当前存在的不足。

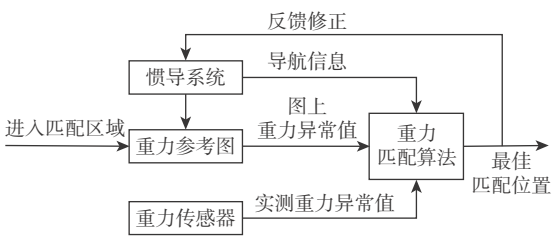


图1 GAINS工作原理

Fig. 1 Working principle of GAINS

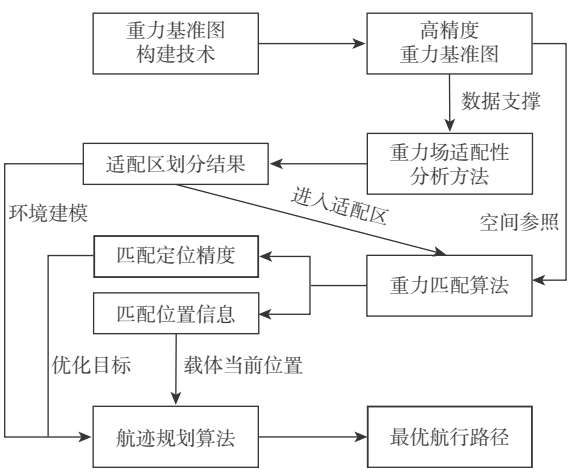


图2 GAINS关键技术

Fig. 2 Key technologies of GAINS

1.1 多源重力数据融合方法

多源重力数据的有效融合依赖于数据预处理与时空基准统一等基础工作的支撑。在此基础上,现有的融合方法可归纳为空间统计法、谱分解法和物理/参数化法3类,近年来,深度学习方法也展现出巨大潜力。

1.1.1 多源重力数据预处理及时空基准统一

不同来源的重力数据在获取时间、测量仪器、参考基准与处理流程等方面存在显著差异,若不经预处理直接融合,系统偏差和噪声干扰将严重降低融合结果的可靠性与一致性。因此,在数据融合之

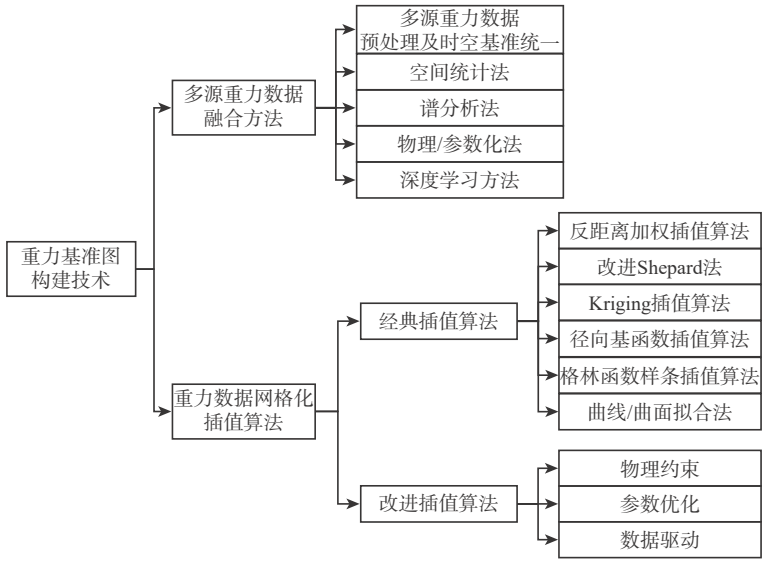


图 3　重力基准图构建核心技术及主要方法

Fig. 3　Core technologies and main methods for constructing gravity reference maps

前, 系统开展预处理和基准统一是重力基准图构建的基础前提。

卫星测高是获取全球海洋重力场信息的主导技术手段。国外已形成成熟的工业化处理体系, 美国斯克里普斯海洋研究所 Sandwell 团队在迭代更新 S&S 系列全球海洋重力模型过程中, 针对 ERS-1、Geosat、Envisat、CryoSat-2 及 Jason-1 等多代测高卫星的原始波形数据, 采用双参数波形重定算法进行重定, 有效抑制了海况引起的测距噪声, 显著提升了测距精度。在数据处理中, 团队基于各卫星任务机构提供的标准轨道及环境改正产品, 以 EGM2008 地球重力场模型为长波参考场, 采用移去-恢复法反演得到全球 1′×1′分辨率的海面重力异常格网数据[26-27]。丹麦科技大学 Andersen 团队在迭代 DTU 系列海洋重力模型(DTU17)的过程中, 持续优化卫星测高数据处理策略, 采用窄峰重定器对 CryoSat-2 波形进行重定, 以提升海冰覆盖区的数据质量; 将潮汐模型由 GOT4.8 更新为更高分辨率的 FES2014, 有效解决了粗分辨率潮汐模型在近岸区域导致大量数据被剔除的问题, 从而显著提升了海岸带及浅水区域卫星重力数据的有效利用率[22]。国内学者提出了一系列相关算法, 聂琳娟等[28] 指出需系统完成干对流层、湿对流层、电离层、潮汐及海况偏差等环境误差改正, 并利用波形重跟踪技术提高近岸测高数据的有效性。黄谟涛等[29] 提出将传统的交叉点平差整体解法简化为两步处理法, 解决了多代卫星测高数据之间的不协调问题。周瑞宸[30] 利用移去-恢复法反演海洋垂直重力梯度, 有效压制了长波信号误差。张思慧等[31] 提出了基于联合变分自编码器与门控循环单元网络的组合模型, 显著提升了卫星重力梯度数据粗差识别的效率与精度。

航空重力数据受飞机振动、颠簸等动态噪声干扰, 预处理的核心是兼顾精度与分辨率的信号分离与噪声压制。柳林涛等[32] 构造了一种由一个参数化的高斯函数与一个尺度化的香农函数构成的函数滤波器, 应用于航空重力测量数据滤波处理, 得到了分辨率为 7 km 的航空重力测量结果; 黄谟涛等[33] 提出滤波尺度自适应调整与零相关准则约束相协同的海空重力测量有色噪声分离和补偿方案, 将航空机载重力测线交叉点符合中误差从 3.57×10^{-5} m/s^2 降至 2.28×10^{-5} m/s^2, 分辨率提升 50%。王皓等[34] 将最小曲率方法用于航空重力数据噪声压制, 采用松弛迭代技术, 在实测数据的测试中, 外符合精度可以达到$\pm1.60\times10^{-6}$ m/s^2。熊盛青等[35] 提出融合卡尔曼滤波与有限冲激响应(finite impulse response, FIR)滤波的方法, 兼顾动态环境下的噪声适应性和频域选择性。

针对船载重力数据预处理, 粗差剔除和系统误差补偿是船载数据预处理的研究热点。Guo 等[36] 提出一种考虑时间误差影响的多年代船载重力异常联合再处理方法, 通过粗差剔除、高斯滤波、交叉点条件平差及引入时变重力项的混合多项式误差补偿, 成功将菲律宾海测线交叉点不符值标准差从 1.346×10^{-4} m/s^2 降至 4.30×10^{-5} m/s^2。郭金运等[37] 提出一种多源船载重力异常数据的联合再处理方法, 通过粗差剔除、长波误差二次项改正、交叉点条件平差及混合多项式系统误差补偿, 成功将墨西哥湾船载重力数据的交叉点不符值均方根从 1.210×10^{-4} m/s^2 降至 3.70×10^{-5} m/s^2, 与 SIO V32.1 模型的残差均方根从 6.70×10^{-5} m/s^2 降至 3.91×10^{-5} m/s^2, 显著

提升了数据质量。Abulkhair等[38]针对埃及海域17个航次的船载重力数据，采用加权最小二乘平差与条件筛选相结合的方法进行质量优化，经处理后，交叉点不符值标准差由调整前的5.10×10^{-5} m/s^{2}进一步降至4.62×10^{-5} m/s^2，符合一致性提升9.41%。Bako等[39]通过交叉点平差、2-sigma粗差剔除及留一法交叉验证精化尼日利亚沿海船载重力数据，并与DTU21GRA卫星测高重力模型经最小二乘配置(least-squares collocation, LSC)融合，使船测与模型差异的均值和标准差分别从4.30×10^{-6} m/s^{2}和3.14×10^{-5} m/s^{2}改善至-2.00×10^{-7} m/s^{2}和2.69×10^{-5} m/s^2。

完成单类数据预处理后，卫星、船载、航空重力数据仍存在坐标、高程、时间三类基准差异，当前国内外针对多源重力数据时空基准统一已建立较为完善的理论体系与技术方法。在坐标基准方面，学界普遍采用WGS84、CGCS2000或ITRF系列地心参考框架，通过布尔莎七参数模型可实现不同坐标系统间的严密转换[40]，并借助全球定位系统(global positioning system, GPS)精密单点定位或差分定位将传感器位置统一至同一空间基准[41]。在高程基准方面，不同数据源分别基于参考椭球面、平均海平面等不同起算面，现有研究围绕全球重力场模型(如EGM2008、EIGEN-6C4)与国家高程基准展开，通过高阶重力场模型、GNSS水准联测、海面地形改正及剩余地形改正等方法，实现了大地高、正常高、海平面高的系统归算，初步满足了陆海空重力数据垂直统一需求[42-44]。在时间基准方面，以协调世界时(coordinated universal time, UTC)为核心标准，结合GNSS硬件秒脉冲(pulse per second, PPS)硬同步、时序拟合软同步、交叉点时间平差及高精度授时技术，有效修正了卫星、航空、船载观测平台的时钟漂移、授时偏差与数据失配问题，使多源重力数据达到毫秒乃至亚毫秒级的时间同步精度[45-48]。

上述预处理与时空基准统一工作显著提升了多源数据的一致性与融合适用性，为后续重力场建模提供了可靠的数据基础。

1.1.2 空间统计法

该类方法将重力场视为随机信号场，利用空间自相关性推估未知点。LSC法最为典型，其通过协方差函数融合多源重力数据[49-53]。但LSC存在三方面不足：①需大量实测数据构建准确先验协方差；②协方差模型选择具有盲目性；③大样本下协方差矩阵易病态[54]。针对LSC法的协方差病态性问题，黄谟涛等[55]引入Tikhonov正则化方法，建立了相应的正则化配置模型，提高了配置解的精度和稳定性；高新兵等[56]利用基于逐级余差思想的点质量模型，有效解决了LSC存在的病态问题。针对依经验确定噪声协方差阵的局限性，孙文等[57]将Tikhonov正则化方法及方差分量估计引入LSC模型，解决了信号与误差的方差分量不一致问题及协方差阵病态性问题。

1.1.3 谱分解法

该类方法充分利用多源重力数据在频域上的互补特性。经典的移去-恢复方法利用全球重力场模型作为参考，从航空或舰载重力观测中减去其长波长分量，从而分离出中高频残余信号，用于局部高分辨率重力场建模[58]。吴怿昊等[59]基于此框架和部分欧洲地区的地面重力、船载重力及航空重力数据，建立了欧洲局部重力场模型。由于小波变换可以将数据分解为多个频带，按可靠性选择性融合系数以提升数据重建效果，因此，有学者将其应用于多源重力数据的融合，吴健生等[60]使用二维小波变换融合南海多源重力位场数据，Bolkas等[61]利用小波分解融合了卫星测高、航空与地面重力数据，Bai等[62]将卫星测高与船测重力异常分别进行离散小波分解，根据频率特性采用不同系数融合策略，再经逆变换获得融合异常图，Ma等[63]提出了一种基于小波变换和窗口加权的方法，综合利用船载重力数据和卫星重力数据，生成了兼具高精度、高分辨率和广覆盖的海洋重力场数据。此外，Chen等[64]提出了一种基于频域优势波段融合的方法，首先，通过谱分析确定卫星测高模型与船测数据的截止波长，然后，分别对卫星模型进行低通滤波、对船测数据进行高通滤波，再将滤波结果相加，从而在保留卫星数据中长波信息的同时，有效融入船测数据的短波信号。

1.1.4 物理/参数化法

该类方法通过物理元或数学基函数显式建模重力场，其中，物理元具有代表性的是点质量方法[65]，数学基函数主要包括球谐函数[66]和径向基函数(radial basis function, RBF)[67]。点质量方法源于解算大地测量边值问题的Bjerhammar换置理论，具有计算模型结构简单、使用方便的特点[68]，但因受到边界面和边界值粗糙化或奇异性的影响，观测值的向下解析延拓欠适定，使该方法的数值阶段过程失稳[69]。高新兵等[56]使用点质量方法对澳大利亚某区域的航空和地面重力测量数据进行融合，实验结果表明，点质量模型理论能够有效解决LSC存在的病态问题。黄谟涛等[70]将Tikhonov正则化方法引入点质量模型解算，有效抑制了病态系数矩阵小奇异值放大噪声对点质量解的污染，提高了解算结果

的精度和稳定性。成怡[71]通过对球谐函数的分析, 并在利用残差重力异常计算球谐函数系数改正项的基础上, 提出了修正局部重力场的球谐函数系数的多源重力数据融合算法。但该方法计算规则化的球谐系数过程繁琐, 且在融合过程中会丢失部分高频信息。RBF 方法可利用数量较少的基函数精确逼近重力场, 计算效率远高于球谐函数方法和 LSC 法[72], 被广泛用于局部重力场建模[73-77]。

1.1.5　深度学习方法

近年来, 数据驱动的深度学习方法被引入重力数据融合与增强任务。通过学习从低分辨率或稀疏输入到高分辨率输出的非线性映射, 神经网络可绕开统计或显式物理建模[78-79]。多层感知机(multilayer perceptron, MLP)[80]、卷积神经网络(convolutional neural network, CNN)[81]、编码器-解码器架构[82]等已被用于重力场超分辨率、去噪与插值任务。Xiao 等[83]提出了一种基于 MLP 神经网络的卫星测高重力异常精化方法, 通过融合船测重力、大地水准面高、水深及沉积厚度等多源数据, 在墨西哥湾划分近岸与远海区域进行建模, 构建了分辨率达 1′×1′的 MLP_GRA 模型, 显著优于 SIO_V32.1 等现有模型, 验证了 MLP 方法在非线性拟合和多源数据融合中的有效性。然而, 该方法缺乏显式空间特征提取, 没有物理可解释性[84]。孙明智等[85]针对海洋重力异常与多源地球物理数据之间的复杂函数关系, 结合移去-恢复法提出了基于多源数据和 CNN 的海洋重力异常预测方法, 提升了模型精度。Li 等[86]提出了一种基于 CNN 的多源卫星测高垂直偏差融合方法, 利用船载重力作为训练标签, 有效融合了多颗测高卫星的垂直偏差数据, 提升了海洋重力场模型精度与一致性。为进一步提高模型性能, Jia 等[87]提出了一种基于半监督对偶回归学习的海洋重力场超分辨率重建框架, 通过联合利用稀疏高精度的船载重力数据与广泛覆盖分辨率较低的卫星测高重力数据, 在少量标注样本和大量未标注样本上同时学习前向与后向映射, 并引入循环一致性约束与注意力机制, 有效缓解了高分辨率船测数据稀缺的瓶颈, 为构建高精度、高分辨率海洋重力基准图提供了新的技术路径。Jia 等[88]提出了物理信息约束的多通道 Swin Transformer 框架, 通过在损失函数中嵌入基于位场理论的谱域物理约束, 有效恢复了短波重力信号, 自适应多域多尺度特征融合模块改善了异源数据集成效果, 在南海实现了高分辨率重力异常场重建。

深度学习方法的优势在精度与数据效率上尤为突出。MLP 模型融合多源地球物理数据后, 所得重力异常模型的标准差较传统卫星测高模型(SIO_V32.1、DTU21GRA)降低 4.0×10^{-6} ～6.0×10^{-6} m/s^2[83]; CNN 方法在南海区域相较于传统方法, 标准差从 5.06×10^{-5} m/s^2 降至 3.21×10^{-5} m/s^2, 精度提升约 36.56%, 且超过 92% 的误差控制在 $\pm5\times10^{-5}$ m/s^2 以内[86]; 引入物理约束的 Swin Transformer 进一步将均方根误差降至 3.09×10^{-5} m/s^2, 较 SIO V32.1 模型提升 47.72%[88]; 而半监督双回归学习框架仅需 50% 的标注样本即可取得优于全监督传统模型的效果, 显著降低了对高成本船测数据的依赖[87]。上述结果充分验证了深度学习方法在构建高精度海洋重力基准图中的显著优势与广阔应用前景。

1.1.6　多源数据融合中的不确定性传播

重力基准图的可靠性不仅取决于融合精度, 更依赖于融合结果不确定性的准确量化。围绕这一问题, 上文综述的四类方法均开展了不同程度的探索, 但各自存在明显不足。

在空间统计法中, LSC 建立了最完整的协方差传递理论框架, 能够通过输入数据的协方差阵推估融合格网的方差矩阵, 提供逐点精度指标。黄谟涛等[55]引入 Tikhonov 正则化改善协方差矩阵病态性, 高新兵等[56]通过条件数分析量化了配置法协方差矩阵的病态程度, 揭示了不确定性在矩阵求逆过程中被急剧放大的风险, 孙文等[57]引入方差分量估计解决信号与误差方差因子不一致问题。然而, 正则化在压制噪声放大的同时扭曲了协方差的原始传播路径, 输出的方差估计既非无偏也非最优; 方差分量估计的有效性高度依赖先验协方差模型的准确性, 而多源重力数据的误差结构远非单一参数化模型所能刻画。因此, 协方差矩阵病态与先验失配使输出方差与真实误差之间仍存在难以量化的偏差。

在谱分解法中, 研究者对不确定性问题的关注主要集中在频域截断和滤波操作对融合结果的影响分析方面, 这些分析并非严格意义上的不确定性传播建模。各输入数据在不同频带上的误差协方差在融合过程中未被显式建模与传递, 该类方法不输出融合后网格点的误差协方差, 波长截断阈值选择所引入的主观不确定性也缺乏量化手段。

物理/参数化法通过物理元或数学基函数显式建模重力场, 其核心目标是通过最小二乘或正则化求解一组确定性系数, 进而重构出单一最优解。黄谟涛等[70]虽然通过 Tikhonov 正则化提高了点质量模型解算的数值稳定性, 但未能将正则化过程中的偏差与方差变化量化为最终的格网精度指标。成怡[71]提出的基于球谐函数系数改正的融合算法虽

能利用残差信息修正长波分量,但各次迭代之间的误差传播关系并未被记录和输出。该类方法的共同缺陷在于:输入数据各自具有不同的测量噪声谱和空间相关尺度,而参数化求解过程将其视为等权的观测值或仅凭经验赋权,既无法在数学上严格传递各输入源的协方差信息,也无法区分融合结果在不同空间位置上的可靠性差异。

深度学习方法显著提高了融合精度,但关于不确定性量化的研究尚处于起步阶段。该类方法将多源数据融合视为端到端的确定性函数逼近问题,其间没有显式地建模输入数据各自的测量噪声协方差,也没有任何机制将这些输入端的已知不确定性通过前向传播映射为输出端的预测方差。Jia 等[87]在提出的半监督对偶回归学习框架中引入“经验不确定性量化”的概念,通过独立检查点残差的统计分布来表征全局预测不确定性,但这属于事后误差估计,而非建立从输入数据协方差到输出格网方差的解析传递链条,因此,也无法向下游匹配算法传递完整的方差-协方差结构。

尽管多源重力数据融合技术取得了长足进步,现有方法仍面临若干挑战:① 统计法与参数化法的先验依赖与数值问题。LSC、点质量等方法对协方差先验或正则化参数敏感,在大样本或数据稀疏区域易出现矩阵病态、精度下降,且基函数配置缺乏普适性。②谱分解法的频域截断与边界效应。移去-恢复、小波变换等方法计算效率高,但在海底地形复杂区域易产生边界伪影[87],波长截断和小波基的选择对结果影响显著。③深度学习方法的数据瓶颈与可解释性不足。高精度船测重力数据获取成本高、覆盖有限,导致标签稀缺、模型泛化能力弱;现有研究多忽略地球物理约束,融合结果缺乏不确定性评估,制约了在实际导航等安全关键场景的部署。④现有研究普遍将提高融合精度作为核心目标,大都忽略了融合结果的可靠性评估问题。

1.2 重力数据网格化插值算法

在 GAINS 中,重力基准图通常按照以经纬度划分的网格点形式存储在计算机内。然而因测量条件的限制,实际获取的重力数据是分布不规则的离散点,因此,为便于实际应用,需要利用网格化插值算法对离散数据进行格网化。另一方面,在重力匹配导航过程中,通过查图获取 INS 解算位置在基准图中的重力异常值是不可或缺的关键步骤,而惯导解算位置并不一定位于重力基准图格网点上,该过程会不可避免地引入格网重力查询误差,且误差值会随重力基准图分辨率的降低而增大。此外,在实测重力数据缺失或不可获取时,数据插值技术同样至关重要。因此,需要用插值算法对重力基准图进行重构,在降低减小格网化误差对数据精度影响的同时提升其空间分辨率。

目前,常用的算法主要有反距离加权(inverse distance weighting, IDW)插值算法[89]、改进 Shepard 法[90]、Kriging 插值算法[91]、RBF 插值算法[92]、格林函数样条插值算法[93-94]和曲线/曲面拟合法等[95]。不同的算法有着各自的适用范围和特点。已有大量学者对不同格网插值算法进行实验比较和分析[96-100]。目前,常用插值算法的优劣势总结情况如表 1 所示。可知,不同插值算法各有其适用边界。在重力变化剧烈区,推荐优先选用 RBF 或改进 Shepard 法以保留高频细节;在大范围重力变化平缓区,推荐采用 IDW 或曲面拟合法以兼顾效率;在数据稀疏区,Kriging 插值算法因提供误差估计而更具可信度;在物理约束强、精度要求高的精细建模场景,格林函数样条插值算法因二阶光滑性和物理可解释性具有独特优势。

近年来,众多学者在经典插值算法基础上,针对重力基准图构建中的精度、效率及适应性等问题,提出了一系列改进算法。一类是在传统插值算法基础上进行物理约束或参数优化的传统改进算

表 1 不同插值算法性能对比

Table 1 Comparison of performance of different interpolation algorithms

插值算法	优势	劣势
IDW插值算法	原理简单,易于编程实现;计算效率高,适合大规模数据	对权重系数敏感,主观性强;易产生“牛眼”效应,结果不够平滑;无法估计插值误差
改进Shepard算法	继承IDW简单性,同时减少“牛眼”效应;通过半径搜索或权重修正,提高了局部拟合精度;计算量低于Kriging等统计方法	对搜索半径、拟合度等参数敏感;对数据边界区域的处理效果较差;不提供误差估计
Kriging插值算法	提供最优无偏估计,可给出插值误差;充分利用数据空间相关性,理论严密;可结合各向异性、趋势项等复杂模型	计算复杂,耗时长;需拟合变异函数,参数选择敏感;数据非平稳时效果下降
RBF插值算法	精确插值,光滑性好;适合不规则分布数据,灵活性高;多种核函数可选适应不同场特征	计算量大,速度慢;对核函数和光滑因子敏感;边缘区域误差较大
格林函数样条插值算法	物理意义明确;插值结果二阶光滑,适合连续物理场;可融合边界条件,自然处理数据稀疏区	格林函数构造复杂,涉及奇异性处理;大规模计算困难;对测量噪声敏感,易过拟合
曲线/曲面拟合法	整体拟合,趋势表达清晰,抑制噪声;参数少,模型稳定,易于解释;计算量小,适合大范围趋势提取	局部细节描述能力弱,易平滑掉真实重力异常;阶次选择主观,低阶欠拟合,高阶易振荡

法, 另一类是引入数据驱动模型的新型改进算法。在传统改进算法方面, While 等[101] 率先将广义采样扩展理论引入重力与重力梯度数据插值, 通过在频域组合多个欠采样数据集实现无混叠重建。刘繁明等[102] 利用小波分析感知重力场空间频率, 以基准图小波细节为转移概率在重力异常高频区域加密测点, 在保持精度的同时减少测点数量并降低误差极值。文超斌等[103] 提出基于二维高斯样条插值的精化导航模型, 通过优化局部支撑参数逼近局部重力异常场。任强强等[104] 基于层次分析法定义“优选度”系数, 实现了 Kriging、改进 Shepard、IDW 及 RBF 等插值算法的快速量化遴选。曾敏等[105] 借助 EGM2008 模型计算模型重力异常, 对残差进行 IDW 插值, 大幅提高了海域格网化精度。在新型改进算法方面, Sarzeaud 等[106] 提出改进 Kohonen 自组织神经网络插值, 通过无监督竞争学习使节点自适应表征数据空间分布, 有效降低了计算成本和网格化误差。余雅文等[107] 将长短期记忆(long short-term memory, LSTM)循环神经网络引入重力异常推估, 验证其有限数据下的推估精度优于 Kriging 法, 加入高程约束可进一步提升精度, 但运算效率较低。黄子炎等[108] 提出 Kmeans-RBF 神经网络, 利用神经网络的复杂非线性映射学习能力进行推估建模, 并在模型训练和推估时加入地形数据作为物理控制, 有效提高了重力格网化数据精度。Mao 等[109] 设计 BP-Bagging 集成插值算法, 通过自助聚合多个反向传播(back propagation, BP)神经网络提升卫星测高重力背景场分辨率, 在西太平洋海域实验中精度显著优于传统插值算法。谢心和等[110] 提出基于粒子群优化(particle swarm optimization, PSO)算法优化 BP 神经网络的重力数据插值算法, 实验结果表明, 该方法的插值效果优于普通 BP 和 Kriging 法, 加入高程约束可进一步提升精度, 但运算时间较长。谢心和等[111] 进一步提出一种基于 Legendre 多项式神经网络的插值算法, 根据多项式逼近理论, 将一组 Legendre 正交多项式作为隐含层神经元的传递函数, 再以其加权和函数实现区域重力数据的拟合和重力背景场的构建。Feng 等[112] 提出基于 Boosting 框架的 BP-Boosting算法, 以 BP 神经网络为基模型, 通过多轮迭代误差校正与 Bootstrap 重采样构建强预测器, 在 3 个海域的实验中均方根误差和平均误差优于传统及多种集成算法, 但在地形高度复杂或含噪声数据场景下效果欠佳。於青凤等[113] 从生成式建模角度提出一种基于扩散模型的网格化方法, 通过模拟重力数据的扩散和去噪过程, 学习重力数据的概率密度函数, 从而生成符合全局统计分布的重力异常值, 在稀疏及大面积缺失区域仍能保持平滑过渡与结构连续性。近年来, 基于深度卷积神经网络(deep convolutional neural network, DCNN)的超分辨率方法为重力基准图分辨率提升提供了新的技术路径。Zhang 等[114] 提出了一种非对称双分支网络, 通过 CNN 分支提取局部细节特征、Swin Transformer 分支捕捉长距离空间相关性, 并设计了多尺度边缘特征增强模块与跨尺度特征重建模块, 同时引入空频域联合损失函数以保持重力场的物理一致性。实验表明, 该方法在太平洋海域将重力异常参考图从 2′分辨率重建至 1′时, 均方根误差较传统 Kriging 插值算法、RBF 插值及 SRCNN、SwinIR 等深度算法分别降低 27.98%～14.29%; 在后续重力匹配导航仿真中, 其绝对位置误差较稀疏 2′分辨率参考图降低 29.2%, 较 Kriging 插值算法降低 14.5%, 且优于其他对比算法, 验证了 DCNN 超分辨率技术能够同时提升重力基准图的网格分辨率和精度。

综合上述文献, 现有插值算法主要存在以下共性问题: ① 高频局部特征刻画能力有限。现有插值算法普遍在重力场变化剧烈的高频区域产生较大误差。传统空间插值算法本质上是基于空间自相关或平滑性假设, 难以捕捉局部短波长异常。即使采用机器学习方法, 若训练数据在高频区分布稀疏, 模型也难以学习到有效的局部映射关系。②多源辅助信息利用不足, 物理约束薄弱。多数插值算法仅依赖重力异常值本身及其空间坐标, 对地形、重力梯度、多源卫星测高数据、全球重力场模型等丰富的地球物理信息利用不充分。③ 深度学习方法在稀疏区域泛化能力差, 计算负担重。基于神经网络(如 BP、LSTM、扩散模型)的方法对训练样本的数量和质量高度敏感, 在小样本条件下容易过拟合或欠拟合。同时, 深度模型训练时间长, 参数规模大。

2　重力场适配性分析方法

重力匹配算法输出的位置精度与载体所处匹配区域的重力场特征直接相关, 重力场的局部可匹配性称为导航适配性, 适配性分析旨在定量评估不同重力场的导航适配性。重力变化越明显的区域, 即适配区, 重力匹配算法的匹配精度越高; 相反, 若匹配区域重力场过于平缓, 即非适配区, 重力匹配算法的匹配精度会大幅降低甚至导致误匹配。因此, 该技术直接关系到匹配定位的可靠性, 是保障导航可靠性的前提条件。近年来, 国内学者对重力场的导航适配性分析先后开展了大量的研究工作,

涉及重力场特征参数的选取、重力场综合特征参数构建及导航适配区提取算法等关键技术，主要方法如图4所示。

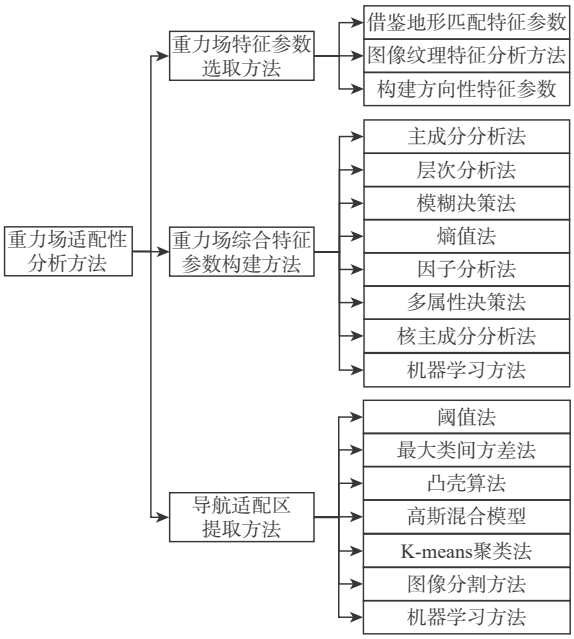


图4 重力场适配性分析关键技术及主要方法

Fig. 4 Key technologies and main methods for gravity field adaptability analysis

2.1 重力场特征参数选取方法

在选取重力场特征参数方面，大部分学者借鉴了地形匹配的思想，将重力场粗糙度、标准差、相关系数、坡度、偏态系数、熵等指标作为重力场特征参数[115-116]；有的学者借鉴数字图像处理的方法，Liu等[117]将重力图转换为灰度图，使用灰度直方图复杂度、Sobel算子和灰度共生矩阵等图像纹理特征分析方法，分析重力图的整体和局部特征，作为反映重力图导航适配性的特征指标。有的学者结合重力场的特性构建了新的重力特征指标，马越原等[118]利用信息熵具有能够整合多种统计参数且算法计算量小的特点，提出了一种基于特征参数信息熵的重力辅助导航适配区的选择方法。以上方法大部分基于区域内重力场的整体变化情况构造特征参数，未充分考虑匹配航迹方向对适配性分析结果的影响。对于同一区域，航行方向不同时，实测重力序列与背景图的匹配效果可能存在显著差异。这会导致利用现有方法提取的适配区内，可能存在某些方向匹配效果不佳的区域；或被划分为非适配区的区域中，也可能存在某些方向具有良好匹配潜力的局部范围。为此，部分学者的相关研究聚焦于构造体现重力场变化方向性的特征参数。Wang等[119]基于分形几何理论，通过分析三维表面上重力异常序列的频谱特征，提出了各向同性系数这一新的特征参数，体现了重力匹配区的方向性。相较于传统标量特征（如标准差、变异系数等），引入各向同性系数后，匹配区域的方向匹配率波动幅度显著减小，最大与最小方向匹配率之差从0.789降至0.200以内，表明方向性特征能有效提升匹配方向的一致性。Wang等[120]进一步提出一种新特征参数——重力异常和向量差，通过对各方向上的重力异常差值向量进行向量求和，得到局部变化最剧烈的方向，相较于传统的标量参数，该向量参数能够获取方向上的匹配适宜度，所选匹配区域的平均匹配率较基于传统标量特征的阈值法提高6%～7%，最低方向匹配率从约57%提升至76%。Ma等[121]提出基于重力纹理特征的方向适应性分析方法。通过灰度梯度共生矩阵与多方向Sobel算子提取特征，提出方向匹配率评价指标，量化区域多方向适配能力。该方法在4个主航行方向（0°、45°、90°、135°）上分别实现了95.05%、94.84%、93.56%和94.77%的匹配率，而传统标量特征不能区分适配区内不同方向的匹配精度。Wang等[122]提出了一种基于曲率几何属性融合重力变异性与方向一致性的特征构建思想，构建了一个能够全面评估匹配区域适用性的新特征——曲率方向特征，并与多个传统特征参数融合，提高了匹配区域选择的鲁棒性和导航路径的连续性。所提曲率方向特征选出的最优区域在6个测试方向上的平均匹配误差为67.92 m，相较于标准差和信息熵特征组合，匹配误差降低约69.7%；相较于粗糙度和相关系数特征组合，降低约22.6%。Zhang等[123]进一步深化了方向特性的分析，设计了3个新的重力方向特征参数：重力剖面相似性、方向梯度、方向梯度集中度，并结合LM-BP神经网络构建了适配性分析模型，相较于基于传统标量特征的阈值法，强匹配区域的召回率从30%～40%提升至70%～77%，且精度超过90%。表2综合对比了不同体现重力场变化方向性的特征参数的适用场景、优缺点、计算复杂度。以上特征选取方法存在一定局限性，一方面，没有对多个特征指标之间的独立性进行判别，如重力场标准差、粗糙度和坡度等本质上都是反映局部区域内重力变化大小的指标，冗余的信息易导致后续多指标融合效果不佳；另一方面，没有筛选出与导航适配性强相关的特征作为评估指标，甚至过分相信弱相关特征导致评估结果的误判，徒增导航适配性分析的复杂度。部分研究从不同角度引入了方向性特征，显著提升了对重力场空间异质性的描述能力，但特征参数的构造仍带有较强的主观性或经验性。

2.2 重力场综合特征参数构建方法

在实际应用中，完成重力场特征参数选取后，需要进一步利用信息融合方法将多种重力特征参

表 2　典型重力场方向性特征参数对比分析

Table 2　Comparison and analysis of classical directional characteristic parameters

方向性特征参数	适用场景	优点	缺点	计算复杂度
各向同性系数[119]	评估区域整体的方向敏感程度, 判断适配区是否在各方向上均适用; 适用于适配区的初筛与方向性定性评估	首次将方向性定量纳入适配性评估; 系数接近1时表明区域方向适应性好; 理论基础严密, 具有明确的物理意义	仅反映方向一致性, 无法度量重力变化强度, 需与其他特征联合使用; 计算涉及频域变换, 较为复杂	$O(n\log_2 n)$ (n为滑动窗口总数量)
重力异常和向量差[120]	识别区域重力变化的主导方向; 适合作为航迹规划中方向选择的参考依据; 在重力变化各向异性明显的区域效果突出	可获取方向上的匹配度, 克服标量参数无法反映方向性的局限; 模值设计基于圆概率误差与3σ准则, 理论依据充分	在重力变化剧烈但方向相互抵消的区域, 模值可能偏小; 矢量和方向对窗口形状敏感	$O(n)$
方向匹配率[121]	评估区域在多方向上的综合适配能力; 适合需要全面评估区域各方向匹配性能的场景, 尤其是航向不确定时的适配区选取	多方向梯度特征提取全面; CNN分类避免了主观阈值设定; 实现方向适配性的定量评估	灰度级量化参数和方向数选择主观; 特征计算涉及图像变换, 复杂度较高	$O(nd_S L^2)$ (d_S为Sobel算子的方向数量, L为灰度量化级数)
曲率方向特征[122]	兼顾区域重力变化幅度与方向一致性的综合评估; 适合对适配区进行方向鲁棒性筛选	曲率特征直接反映曲面几何特性, 物理意义清晰; 配合K-means聚类与形态学闭运算, 兼顾区域连通性	需计算主曲率与高斯曲率, 对重力数据平滑性敏感; 方向一致性指标仅基于主曲率方向, 未覆盖所有航向	$O(nW^2)$ (W为滑动窗口边长)
重力剖面相似性[123]	评估沿特定航向的匹配可行性, 直接反映轨迹序列的独特性, 适配于航迹规划中航向优选	直接反映沿航向的重力序列独特性, 对匹配算法的实际效果预测更准确	剖面长度参数与权重参数需人工设定, 对结果敏感	$O(nW^3)$
方向梯度集中度[123]	评估特定航行方向上的重力变化强度与一致性, 识别重力场方向性各向异性的区域	可量化任意给定航向的适配性; 正交方向负相关特性可验证特征合理性	需针对每个航向单独计算, 计算量随方向数量线性增长; Prewitt算子的方向分辨率有限	$O(nd_p W^2)$ (d_p为Prewitt算子的方向数量)

数进行有效融合。该类方法借助多指标综合评价技术, 将不同维度的重力特征加以融合, 从而构建能够综合反映重力导航区域适配性能的特征参数。目前, 在各领域应用较为广泛的多指标综合评价方法有主成分分析法[124-125]、层次分析法[115,120,126-127]、模糊决策法[128]、熵值法[118,129]、因子分析法[130]、多属性决策法[131-132]、核主成分分析法[133]等多种方法。有的学者引入了机器学习领域的方法, Ma 等[121]结合并行 CNN 对统计特征与纹理特征进行融合训练, 提出了方向匹配率的概念。以上方法通过引入各种特征融合方法, 构建重力场综合特征参数, 但在计算重力场特征时大都基于一定范围区域内的整体重力异常值, 体现的是区域的平均导航适配程度, 忽略了重力图的匹配定位效果存在节局部差异的问题, 不仅会因特征被平滑带来重力场特征粗略不精细的问题, 还容易为后续提取导航适配区出现不连续的问题埋下隐患。此外, 大范围重力场内适配区的面积并非统一, 不同区域因重力场空间特征变化的差异, 其适配范围大小不一。有些区域仅在小范围内具备良好的可匹配性, 而另一些区域则在较大范围内均能提供可靠的匹配结果。然而, 现有特征构建方法多采用固定尺度的分析窗口, 未能充分考虑这种适配区面积的多尺度差异, 导致大尺度窗口可能遗漏小范围高适配区域, 而小尺度窗口则难以全面评估大范围适配区的整体特征。

2.3　导航适配区提取方法

在提取导航适配区方面, 最传统的方法是阈值法[134-137], 即通过比较各种重力特征参数与其对应阈值, 作为适配区提取准则。Wang 等[138]利用最大类间方差准则实现匹配区域的筛选, 减少了人工干预与经验判断的依赖。马旭东等[139]采用了地形分析中的等值线技术对重力场特征区域进行分割计算, 提取区域边缘, 并进行聚类处理, 利用凸壳算法提取这些区域的凸壳范围线, 获得重力适配区。张勇刚等[133]将高斯混合模型(Gaussian mixture model, GMM)算法应用于适配区提取, 通过计算每个局部区域属于适配区的概率对特征矩阵进行聚类, 从而实现对重力适配区的准确划分。相关学者还借鉴了图像、机器学习等领域的适配区分割方法。Wang 等[140]提出了基于重力梯度高分辨率特性的频域匹配区域选择方法, 通过小波变换提取重力场高频信息, 利用形态学图像纹理分割方法提取密集纹理区域作为匹配区域。吴华玲等[141]引入水平集理论, 采用一种基于 Mumford-Shah 模型的全局化水平集的图像分割方法对重力场可匹配区域进行分割。Wang 等[142]采用 K-means 聚类分析方法将重力场背景图分为强适配区、弱适配区、非适配区三类区域。有的学者使用支持向量机(support vector classifier, SVC)作为分类器, 把重力场分为适配区和非适配区[143-144]。Wang 等[145]在此基础上提出构建多层支持向量分类器, 结合层次分类方法和遗传算法优化参数, 实现适配区的划分。Zhang 等[123]用 LM-BP 神经网络替代了传统的阈值法或聚类法, 自动学习并量化重力特征与匹配精度之间的复

杂非线性关系，实现了对不同航向角下重力匹配适应性的准确、快速评估。然而，阈值法中参数阈值的确定过于依赖人的主观经验，且重力场区域变化时阈值需重新标定；最大类间方差隐含直方图双峰假设，对实际多模态或单峰分布的重力场分割效果不佳；GMM和图像分割方法引入了额外的主观参数，算法性能对预设参数值敏感；K-means聚类方法依赖欧氏距离与球形簇假设，忽略了各类别的方差差异与概率特性，对噪声和野值比较敏感；利用支持向量机和神经网络等机器学习类方法提取导航适配区时，需要预标定的训练数据，这部分依赖重力匹配算法或人工设定阈值，在数据泛化、实时性和环境适应性等方面仍有待提升。

近年来，部分学者认为适配性分析的最终目标并非仅仅划分适配区与非适配区，而是建立重力场局部变化特征与配定位精度之间的定量映射关系。基于这一认识，他们开始尝试突破传统的定性分区模式，推动适配性分析向可定量化估计精度的方向发展。Wang等[146]提出了区域匹配能力(matching capability of areas, MCA)的概念，定义为在给定系统条件下，匹配结果落在允许误差范围内的概率。将重力场梯度、地图分辨率、传感器精度等作为先验信息，计算在给定允许误差范围内成功匹配的概率，建立了重力场特征与匹配成功概率之间的定量关系。进一步，Wang等[147]提出了一种基于粒子滤波(particle filter, PF)的匹配精度解析量化方法，通过建立估计总体匹配偏差(estimated overall matching biases, EOMBs)的解析方程，将重力异常图误差、重力仪测量噪声、INS漂移、初始位置误差等因素纳入统一框架，实现了匹配位置偏差的定量估计。上述研究初步实现了从“定性适配分区”向“定量精度估计”的跨越，为适配性分析提供了可量化的指标。

除上述方法层面的局限外，从工程转化的角度，仍需直面以下更深层次的共性挑战：①适配性分析对匹配策略与航迹规划的量化指导能力不足。现有适配性研究多止于特征参数计算与适配区划分，较少建立适配性指标与可达匹配精度之间的定量映射关系，导致分析结果难以直接为匹配参数调整或航迹规划提供量化约束，制约了适配性分析从理论评估向工程决策的转化。②现有适配性分级体系难以满足差异化导航任务需求。现有研究多将适配区简化为“适配/非适配”二元划分，该分类方式过于粗论，难以精细刻画不同区域适配性的差异性。虽有部分研究已尝试采用聚类分析、多属性决策等方法将适配区划分为强、弱、非适配3个等级，但分级粒度、分级依据、等级阈值等关键问题仍缺乏统一的理论框架，分级结果亦未与实际可达定位精度建立定量关联，难以针对不同精度需求的工程任务提供差异化指导。③尚未建立面向具体匹配算法的适配性评价体系。不同重力匹配算法对重力场特征的需求存在差异，同一区域对不同匹配算法的适配性可能不同。工程实际中，不同型号水下航行器搭载的匹配算法各异，因此，亟需突破“静态区域划分”的模式，建立面向具体算法的适配性评价体系。这一思路在地磁匹配领域已有初步探索[148]，对重力匹配适配性分析具有借鉴意义。

3 重力匹配算法

在重力辅助惯性导航系统中，重力匹配算法是核心技术之一[149]。该技术综合利用INS导航信息、预先存储的重力基准图和实时重力异常测量值，以实现运载体的位置估计，其匹配算法的性能直接决定了水下导航的精度、效率和可靠性[13,150]。根据所使用理论框架的不同，现有的重力匹配算法大致可分为3类：①基于相关性分析理论的序列匹配算法；②以贝叶斯非线性滤波理论框架为基础的实时匹配算法；③多算法融合的匹配算法。主要算法如图5所示。

3.1 基于相关性分析理论的序列匹配算法

序列匹配算法的代表性算法有地形轮廓匹配(terrain contour matching, TERCOM)算法[151]和迭代最近等值线点(iterated closest contour point, ICCP)算法[152]。

TERCOM算法是由美国E System公司于20世纪70年代开始研制、90年代成功应用于“战斧”巡航导弹的一种导航技术[153]。该算法的核心原理是：地球陆地表面任何地点的地理坐标，均可由其周围地域的等高线或地貌特征唯一确定。在重力辅助导航系统中，其作用是在重力基准图中找到一条航迹序列，该序列平行于导航系统指示的航迹，且基准图中对应的候选序列与实测重力异常序列具有最强的相关性，实现对INS航迹的修正。TERCOM算法在重力特征明显的区域可靠性良好，对初始位置误差不敏感；但相关性分析要在获得一定时间长度的重力异常序列后才能进行，属于断续的批相关处理技术，因而实时性较差；对航向误差较敏感，在定位过程中不能机动飞行；随着搜索范围的增大，算法的计算量会大幅增加。针对这些局限性，研究者们从不同角度提出了多种改进方案。Han等[154]通过最短路径算法实现逐点匹配，结合多参数评价机制，改善了算法的实时性与方向误差适应性；魏

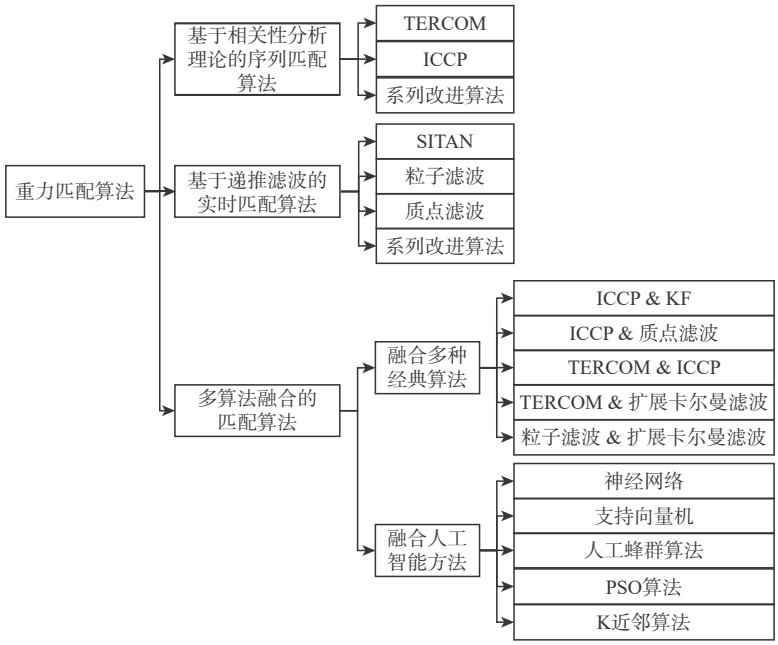


图 5　重力匹配算法分类及主要方法

Fig. 5　Classification and main methods of gravity matching algorithms

二虎等[155]进一步加入旋转与尺度变换, 通过四参数仿射变换有效补偿了 INS 的轨迹形变; Chen 等[156]则将 PSO 与仿射变换结合, 构建了两级匹配框架, 在保证精度的同时大幅降低了计算时间。近年来, 涌现出多种以动态搜索策略为核心的改进算法。这些算法聚焦于传统 TERCOM 算法存在的匹配效率低和“域外失配”问题, 通过设计精巧的预匹配或迭代机制来大幅缩减搜索空间。例如, Zhao 等提出的域中心自适应转移匹配方法(domain-center adaptive-transfer matching method, DAMM)[13], 通过水平和垂直线上的预匹配实现域中心的快速迁移, 避免了全网格遍历; 其进一步提出的交叉线自适应域匹配(cross-line adaptive domain matching, CADM)算法[15]则融合了十字线粗匹配与自适应域精匹配的双阶段搜索模式, 显著提升了匹配速度。此外, 该团队提出了迭代双十字域中心转移匹配(iterative double-cross domain-center transfer-matching, ID2TM)方法, 引入迭代双十字线粗匹配与局部小域精匹配机制, 并结合稀疏采样与重叠点消除策略, 使匹配效率较传统 TERCOM 提升了约 84%[3]。

ICCP 算法的思想源于计算机视觉中的迭代最近点(iterative closest point, ICP)算法, 起初是用来纠正图像对准问题, 随着地形匹配导航的发展, 逐渐用于各类匹配导航领域。ICCP 算法的核心原理是通过迭代优化, 将存在误差的惯导指示航迹逐步匹配到预存的重力场等值线图上, 从而估计出真实航迹。首先, 算法对当前估计航迹上的每个点, 在其测量值对应的等值线上搜索几何距离最近的点作为目标点; 然后, 求解一个刚性变换, 使整个估计点集与这些目标点集之间的欧氏距离平方和最小; 最后, 将求得的变换应用于当前航迹得到更新后的航迹, 并重复上述步骤直到收敛。这一过程本质上是在交替满足等值线约束和航迹的刚性形状约束, 通过不断拉近航迹与等值线的距离, 最终实现高精度的匹配修正。ICCP 算法具有精度高, 易于实现等优势, 但该算法也属于批处理方法, 因此, 实时性较差; 当惯导初始误差较大时, 算法迭代过程容易陷入局部最优解, 其定位精度难以保证; 由于该算法是利用重力测量值在重力背景图中寻找最近等值线点, 其匹配性能对参考图的分辨率和精度高度敏感。针对传统 ICCP 算法的局限性, 很多学者提出了一系列改进措施。童余德等[157]通过固定初始序列长度并采用滑动窗口局部搜索, 实现了单点迭代匹配, 显著提升了算法的实时性与搜索速度; 蒋东方等[158]利用 BFGS 拟牛顿方法在局部连续背景场中精确搜索最近等值点, 解决了离散网格下等值点定位受限于分辨率的问题, 提高了匹配精度; 白文平等[159]借鉴移位寄存器原理改变采样结构以增强实时性, 并引入“二次迭代”与加权平均处理, 进一步提升了匹配结果的可靠性; 王向磊等[160]同样基于移位寄存器原理实现了实时匹配, 并通过多结果加权平均有效削弱了随机误差; 蔡龙飞等[161]结合固定序列长度与滑动窗搜索, 改善了算法的实时性与运算速度; 针对水下重力辅助导航中实时 ICCP 算法存在的匹配序列长度固定、最近轮廓点搜索效率低、定位精度受限等问题, Liu 等[162]提出 OMSL-

ICCP优化算法。该算法以黄金分割搜索自适应确定当前航迹最优匹配序列长度，借助豪斯多夫距离缩减最近轮廓点搜索范围，并引入引力场算法构建适应度函数完成精匹配寻优，有效提升不同噪声、惯导误差与重力图分辨率条件下的实时定位精度。

尽管大量学者研究了一系列改进算法，一定程度上提升了算法的精度和实时性，但这类序列匹配算法本质上属于批处理方法，运动载体必须航行一段距离获取足够的重力数据后才能完成匹配运算，因此，不是严格意义上的实时处理方法。这一缺陷对于作战应用是不可接受的，很难满足常态化海上作战保障需求，更无法满足应急作战实时化保障需求[4]。

3.2 基于递推滤波的实时匹配算法

为满足实际应用中系统修正的实时性，目前，国外研究者主要将实时滤波算法作为研究重点。基于递推滤波的实时匹配算法大致分为基于线性化的近似解法和基于随机采样的滤波算法。基于扩展卡尔曼滤波器(extended Kalman filter, EKF)的桑迪亚地形辅助导航(Sandia inertial terrain aided navigation, SITAN)算法[163]属于线性化近似算法，随机采样算法主要是PF[164]和质点滤波(point mass filter, PMF)[165]。

SITAN算法具有实时性好、工作性能稳定、允许载体较大机动性等优点。然而，该算法对初始误差精度要求较高，且对非线性观测模型的线性化过程较为敏感，非线性函数局部线性化带来的误差容易造成SITAN算法匹配结果发散[166]。Yuan等[167]针对初始INS误差较大时传统SITAN易发散的问题，提出了等高线匹配与SITAN相结合的改进算法，先利用INS短时高精度特性进行等高线初始匹配，以降低初始位置误差，再采用SITAN进行精匹配，仿真表明，在初始误差为800 m时仍能有效匹配。董翠军[168]提出了相关SITAN+降权迭代匹配算法，构建了基于序列递推的SITAN算法模型，并将序列相关技术和降权迭代技术融入其中，提高算法的精度和鲁棒性。Deng等[169]从滤波非线性逼近能力和噪声鲁棒性角度出发，将无迹卡尔曼滤波(unscented Kalman filter, UKF)引入SITAN以减小线性化误差，同时，针对水下过程噪声不确定和量测噪声受污染的问题，提出自适应鲁棒UKF匹配算法，通过引入自适应因子与鲁棒函数提升系统鲁棒性。欧阳明达等[170]提出一种应用抗差估计SITAN算法的水下重力匹配导航方法，通过设置调节因子，降低异常点状态矩阵和增益矩阵等对下一位置点的影响来改善导航位置精度，增强可靠性。王傲明等[171]提出了一种基于自适应并行EKF的SITAN匹配算法，通过自适应因子调节状态预测信息的权重，抑制滤波发散，提高定位精度。付林威等[172]提出一种基于抗差自适应SITAN算法的水下重力匹配导航方法，通过构建自适应因子和抗差因子来调节状态预测信息与量测信息对滤波贡献的权重，有效抑制粗差和预测信息异常的影响，提高匹配算法的稳定性和可靠性；为进一步提高水下重力匹配算法的定位精度和稳健性，该团队将容积卡尔曼滤波(cubature Kalman filter, CKF)应用到水下重力匹配惯性导航中，同时，引入抗差估计和自适应因子，提出了基于抗差自适应CKF的水下重力匹配导航SITAN算法[173]。毛宁等[174]针对水下复杂环境中量测噪声统计特性易受污染的问题，提出自适应鲁棒SITAN算法，通过比较新息协方差矩阵递推值与计算值差异，设计新型补偿因子，实现对量测噪声协方差与滤波增益的同步自适应调整，有效抑制量测异常与滤波发散。

PF算法是一种基于蒙特卡罗理论的方法，该理论将传统贝叶斯滤波中的后验概率分布近似为从后验分布中随机采样的粒子的统计学分布。与SITAN算法相比，PF算法无需对系统进行线性化，通过数值方法直接求解无解析解的贝叶斯后验估计，更适用于非线性非高斯的动态系统，且具有良好的实时性，但存在粒子退化、样本贫化和重要性分布函数构建不准确等问题。刘繁明等[175]将差分进化引入粒子的滤波重采样过程，通过群体差异进行全局搜索，实现粒子分布优化，使粒子集提高对系统状态变量后验概率的逼近程度。Wang等[176]提出了一种基于PF的矢量匹配方法，将相邻INS采样点之间的位置相关性加入匹配算法中，以解决某些匹配区域匹配结果不稳定的问题，后来对粒子的范围进行改进优化[177]。Zhao等[150]提出基于重力测量特征的改进PF算法，通过引入航向角方向测量来优化粒子权重，显著降低了重力测量误差和背景图相似度的影响。鲁智谦[178]提出了一种基于模拟退火重采样及最优重要性的差分PF算法，引入模拟退火算法改进重采样过程，在粒子群发生贫化后重新调整粒子分布与权重分布，保障粒子多样性；设计基于最优重要性的采样更新机制，大大增加了粒子的多样性，从根源上减少重采样的发生。

PMF是一种基于贝叶斯估计的递推滤波方法，通过将连续状态空间离散化为有限网格点并递推其后验概率，能够在无需线性化观测模型的前提下处理重力匹配导航中的非线性、非高斯问题。该方法的优势在于对初始位置误差具有较强的鲁棒性，

可自然融合 INS 短时高精度位移约束, 且避免了 EKF 因线性化引入的误差。Han 等[179] 首次将 PMF 应用于重力匹配导航, 利用 PMF 在网格空间上直接递归计算位置后验概率, 解决了传统 SITAN 算法线性化误差大、鲁棒性差的问题。然而, PMF 存在平坦重力区易失配、粒子退化、计算复杂度随网格分辨率指数增长等劣势。针对上述问题, Han 等[180] 提出一种基于 PMF 与相似变换的两阶段重力匹配算法, 先利用 PMF 完成粗匹配, 快速抑制大初始位置误差, 通过相似变换施加空间约束, 对 PMF 结果精匹配, 解决了 PMF 算法在平坦重力区易失配的问题。为缓解粒子退化, 蔡体菁等[181] 提出基于模拟退火重采样的 PMF 重力匹配算法, 概率性保留小权重粒子并动态生成新粒子, 从而保证粒子权重与粒子空间分布的多样性。Wang 等[182] 提出观测差分 PMF, 通过对重力观测值做差分消除重力仪长期漂移, 并建立差分贝叶斯递推模型, 依据重力图误差与局部梯度比值自适应调整网格分辨率, 在保证优化精度的同时大幅提升计算效率, 解决固定分辨率精度与效率难以兼顾的问题。

这类基于递推滤波的实时匹配算法在处理非线性问题时有着较为出色的性能, 各类改进算法在精度、鲁棒性及计算量等方面均有所改进。然而, 在重力场特征不明显的区域或当重力测量噪声较大时, 滤波过程极有可能发散, 算法鲁棒性仍有待进一步提高。

3.3　多算法融合的匹配算法

针对单一算法的局限性, 一些学者研究了大量多算法融合的匹配算法, 联合不同算法的性能优势, 提高导航系统的定位精度。部分学者开展了联合多种经典匹配算法的相关研究, Liu 等[183] 联合使用 Kalman 滤波与 ICCP 算法, 提高了匹配的速度和效率。Wei 等[184] 提出了一种融合多点相关与降权迭代的算法, 将 TERCOM 的多点相关思想引入 EKF 框架, 构建包含多个连续点位置、速度和加速度的状态向量, 有效抑制了滤波发散。Han 等[185] 提出了 PMF 与 ICCP 相结合的两步策略, 利用 PMF 进行粗匹配以克服 ICCP 对大初始位置误差的敏感性。邹嘉盛等[186] 提出以 TERCOM 为主、ICCP 为辅的联合匹配策略, 通过对比惯导、TERCOM 和 ICCP 这 3 种方法在相邻采样点间的航距, 识别 TERCOM 匹配结果中的异常点, 并用 ICCP 对应匹配点进行替换, 降低了整体匹配轨迹的平均误差。Gao 等[187] 提出了一种并行多重自适应决策重力导航方法, 将重力异常观测值、PF 和最大相关法引入 EKF 的观测方程, 提高了重力辅助导航的定位精度。张锦柏等[188] 提出了粗精匹配结合的算法, 先利用惯导误差特性约束等值点搜索范围, 并采用 TERCOM 类方法获得粗匹配轨迹, 再通过 ICCP 算法对轨迹进行旋转与尺度修正, 从而提升匹配精度与效率。

除以上融合多种经典匹配算法的研究外, 部分学者将人工智能领域的多类方法引入到了重力匹配算法中。程力等先后将神经网络[189] 和 SVC[190] 应用于重力辅助导航, 有效提高了定位精度。黄鹏等[191] 提出一种改进的 BP 神经网络学习算法, 有效加快了搜索最近等值点的速度, 更好地满足重力辅助导航的实时性要求。高伟等[192] 将人工蜂群算法引入重力匹配搜索过程, 并利用多普勒测速仪速度信息作为约束条件, 同时引入平均 Hausdorff 距离对匹配结果进行筛选, 有效降低了误配率。Dai 等[193] 提出了一种基于仿射变换的人工蜂群算法, 在初始化及进化过程中引入尺度、旋转和平移变换, 并利用 INS 短时高精度位移特性约束仿射变换参数, 提高了算法在大初始误差下的匹配精度。邹嘉盛等[194] 提出了一种基于约束条件的重力匹配导航算法, 引入惯导提供的轨迹方位和航距信息双重约束, 将满足约束条件的有效轨迹输入概率神经网络完成最终匹配, 大大提高了重力匹配的效率。Zou 等[14] 分析了 PSO 算法应用于重力匹配导航的可行性, 并提出一种改进 PSO 筛选迭代算法, 定位精度和匹配效率显著提升。欧阳明达等[195] 联合 PSO 改进 PF 算法, 将初始航迹匹配误差作为平移参数, 修正惯性导航整体航迹, 削弱了初始配准误差对后续 PF 的影响, 提升了 PF 算法的匹配导航精度。Gao 等[196] 提出了基于 K 近邻的重力匹配算法, 在搜索范围内建立样本数据集, 通过欧氏距离选取 K 个近邻标签, 并根据样本空间位置设计距离加权方法, 同时, 引入航速约束对连续匹配结果进行递推修正, 显著提升了连续匹配的稳定性与轨迹平滑度。多算法融合的重力匹配导航算法在提升精度的同时, 也引入了更高的系统复杂性、参数敏感性和计算资源需求, 仍有待进一步深入研究。为定量评估不同多算法融合的匹配算法的性能, 表 3 对比分析了几种典型融合方法的精度、实时性、参数敏感性、适用场景等。

三类重力匹配算法在实时性、初始误差容忍度、重力特征敏感性、重力测量噪声鲁棒性及计算复杂度等方面各有优劣。为便于根据实际任务需求快速选择合适算法, 表 4 从上述维度对序列匹配、实时滤波及多算法融合方法进行了归纳对比, 并给出了各类算法的使用场景。

表3 典型多算法融合的匹配算法对比分析

Table 3 Comparison and analysis of representative multi-algorithm fusion matching algorithms

多算法融合	匹配精度	实时性	参数敏感性	适用场景
TERCOM+EKF[184]	匹配精度依赖重力测量误差，1×10^{-5} m/s^{2}的重力测量误差的平均定位误差为674.128 m，5×10^{-5} m/s^{2}的重力测量误差的平均定位误差为4 379.98 m	基于EKF的递推估计，可实时工作	对重力场粗糙度、初始误差、线性化误差及重力仪噪声均敏感	适用于中等粗糙度区域、重力仪精度较高和实时性要求较高的场景
TERCOM+ICCP[186]	TERCOM出现明显误匹配且ICCP匹配结果较好时，平均定位误差由1 998.308 m降至870.44 m	批处理算法，实时性较差	对TERCOM初始匹配质量和ICCP匹配精度敏感	适用于修正TERCOM误匹配的实时性要求不高的场景
PF+EKF[187]	全程平均定位误差达620.42 m；适配区的均方根误差：东向为281.504 m，北向为272.244 m	并行计算增加负载，但满足实时要求	依赖预计算阈值，但决策树自动选择方法降低调参难度	适用于各类重力场；存在较大测量噪声的长航时、连续水下导航场景
ICCP+BP神经网络[191]	与ICCP一致（未量化位置误差，仅提升计算速度）	相较于ICCP，迭代次数降82%，时间降低81%	对网络结构、学习率等超参数敏感，但文章未进行相关分析	适用于实时性要求高、计算资源受限的嵌入式平台
PF+PSO算法[195]	横穿等高线轨迹的平均定位误差为1 140.832 m，平行于等高线轨迹的平均定位误差为594.492 m	采用滑动窗口递推，满足实时导航需求	缓解了标准PSO算法对初始参数敏感的固有缺陷，但依赖数字重力图的精度和重力仪实测数据的质量	适用于长航时重力辅助惯性导航场景，但不适用于重力场变化平缓或背景图分辨率与精度不足的区域
引入K近邻算法[196]	平均定位误差均小于800.064 m	每匹配点平均耗时5.87 s，可满足实时导航需求	对K值、搜索半径敏感，但抗噪性优于平均平方差相关算法	适用于存在较大测量噪声或重力场变化较平缓等困难条件的长航时、大航程航行任务

表4 三类重力匹配算法的性能对比与适用场景

Table 4 Comparison of performance and suitable applications of three types of gravity matching algorithms

对比维度	序列匹配	实时滤波	多算法融合
实时性	属于批处理方法，需累积一定航迹长度后才能完成匹配，实时性较差	可逐点实时输出定位结果，实时性较好	实时性中等，通常包含批处理与递推两阶段，计算延迟较高
初始误差容忍度	TERCOM对初始误差不敏感；ICCP在惯导初始误差较大时，容易陷入局部最优解	对初始误差容忍度较低	对初始误差容忍度较高，常通过粗匹配阶段压制大初始误差
重力特征敏感性	对重力图分辨率和精度敏感，平坦区易误匹配	在平坦区易发散，易导致误匹配	相对鲁棒，可通过多算法互补缓解平坦区失配问题
重力测量噪声鲁棒性	对噪声较敏感，但长序列可在一定程度上抑制噪声影响	当重力测量噪声较大时，滤波过程极有可能发散	较好，多源信息融合可抑制噪声干扰
计算复杂度	计算量与搜索范围相关，会随搜索范围增大急剧增加	SITAN低，PF中等，PMF高（随网格分辨率指数增长）	高，需同时运行多个算法
适用场景	重力特征明显、初始位置误差大、对实时性要求不高的长航程任务	实时性要求高、机动航行、实时修正、初始误差较小的场景	高精度要求、允许较高计算资源、需兼顾大初始误差与复杂噪声的环境

综合上述分析，当前重力匹配算法面临以下挑战：①序列匹配算法本质为批处理方法，需累积一定航迹长度才能匹配，实时性差。同时，该类算法的匹配精度受多项因素制约：航向误差越大，航迹形变越严重，匹配精度越低；初始位置误差越大，搜索范围越大，误匹配概率越高；重力基准图的空间分辨率直接影响格网点上重力异常值的表征精度，匹配精度随基准图分辨率的降低而下降。②实时滤波算法虽具备实时处理能力，但在重力场特征不明显或测量噪声较大时易发散。SITAN对初始误差及非线性线性化误差敏感；PF存在粒子退化与样本贫化问题；PMF在平坦重力区易失配，且计算复杂度随分辨率指数增长。③多算法融合的匹配算法在重力场特征较为丰富、单一算法因初始误差过大或重力测量噪声过大而失效时可显著提升定位精度。然而，当载体进入重力场特征平缓区域时，各基础算法的量测信息同时趋于匮乏，融合结构难以提取有效判别信息。此外，该类方法大都设计复杂且调参繁琐，如何在不显著增加系统复杂度的前提下自适应地融合多算法优势，仍是有待深入研究的问题。

4 航迹规划算法

水下运载体在实际航行过程中必然交替经过重力变化显著的适配区与变化平缓的非适配区。然而，即使在同一匹配区采用相同的匹配算法，不同的航迹规划方案所获得的重力匹配效果也可能存在显著差异。这意味着，重力辅助导航的定位精度不仅取决于区域自身的重力场特征，更与水下运载体的航行轨迹密切相关。合理的航迹规划能够使水下运载体尽可能始终处于重力适配区内，有效避开非适配区，甚至在重力异常变化不显著的区域中也能获得较好的匹配效果。因此，航迹规划成为重力辅助导航系统中的另一项关键技术，其目标是

为水下运载体从起点到终点搜索出一条合理、便捷、高效的优化路径, 在尽量规避非适配区的同时缩短航行距离与时间, 从而提升导航效率、定位精度与作业安全性。

从问题本质上看, 水下重力辅助导航的最优航迹规划可转化为以重力非适配区为障碍的全局路径规划问题。围绕这一核心问题, 相关研究在算法架构、环境建模与优化目标 3 个维度上不断演进, 可划分 3 个发展阶段: 经典启发式图搜索、智能优化与混合算法阶段和自适应与理论最优规划阶段, 主要方法如图 6 所示。

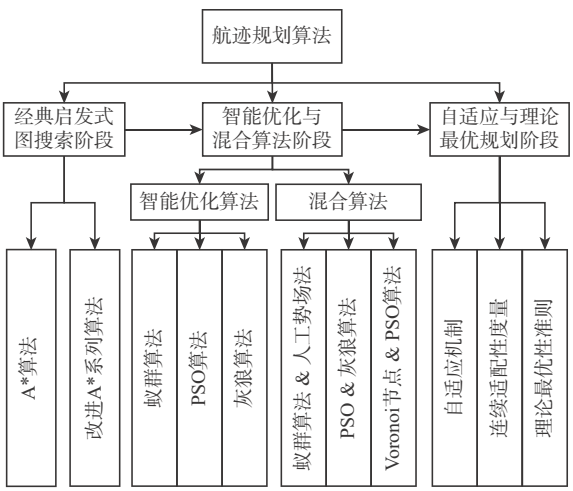


图 6　航迹规划算法发展阶段及主要方法

Fig. 6　Development stages and main methods of route planning algorithms

4.1　经典启发式图搜索阶段

早期研究主要依托 A*算法等图搜索方法, 通过设计启发函数将重力特征纳入路径代价, 引导航行器优先通过重力信息丰富的区域。陆泽橡等[197]将重力梯度信息引入 A*算法的代价函数, 以重力坡度作为导航启发信息, 并采用模糊自适应机制动态调整阈值。该方法使规划路径偏向重力变化剧烈的区域, 提升了匹配可行性。Feng 等[198]在此基础上, 利用熵值法将标准差、粗糙度、坡度、相关系数等多种重力统计特征融合为综合特征参数, 以此改进 A*的代价函数, 使规划路径更倾向于通过重力信息丰富的区域, 显著提高了匹配率。为进一步提升规划路径的实际可执行性, Shi 等[199]在 A*算法中引入了水下航行器的运动约束, 并采用贝塞尔曲线对折线路径进行平滑处理, 使输出路径更符合实际航行器的机动能力。欧阳明达等[200]设计了一种改进 A*算法, 通过前后向对比分析, 对单个路径节点逐个进行筛选过滤, 有效减少了冗余航向调整, 实现了路径的平滑优化, 提高航迹规划的合理性。张锦柏等[201]将 A*算法的启发函数改进为以对角线距离与曼哈顿距离复合的启发函数, 有效减少了航迹点冗余拐角和航程, 并对航迹点做去冗余与平滑处理, 最终使水下航行器沿重力特征显著的海域航行。为解决纯全局静态规划无法应对水下长航时未知动态障碍的问题, Zhao 等[202]提出了融合改进 A*与改进动态窗口法的全局动态航迹规划算法, 在全局静态规划中, 将基于密度估计的重力指数与航向指数融入 A*代价函数, 同时, 采用可变搜索步长优化搜索方式, 提升重力匹配精度并减少路径偏转; 在局部动态规划中, 将动态窗口法运动模型升级为地理坐标系下的高精度模型, 适配水下航行器运动特性以应对未知动态障碍。该方法初步实现了全局适配性规划与局部动态避障的协同。

从定量结果来看, 相较于未规划航迹, 陆泽橡等[197]将规划航迹的重力匹配平均定位误差降低 81.3%, Feng 等[198]使匹配率提高约 82%; 相较于传统 A*算法, 欧阳明达等[200]使折线节点数减少 65%、路径总长度缩短约 8.9%, 张锦柏等[201]使冗余拐角减少 46%、航程缩减 6.64%; 相较于传统路径, Zhao 等[202]使 2 个测试区域的匹配定位误差分别降低 28% 和 30%。上述数据表明, 该阶段方法能够在计算负担增加有限的条件下, 实现从最短路径向重力信息最优路径的初步演进。

该阶段方法结构简单、计算效率高, 从优化目标来看, 以最短路径为主目标, 辅以重力适配性作为惩罚项或启发项, 属于单目标或弱多目标优化。环境建模采用固定阈值将重力场二值化为适配区和非适配区, 非适配区被直接视为障碍物规避, 尚未充分利用重力场的连续渐变特性。

4.2　智能优化与混合算法阶段

随着智能优化算法的发展, 研究者开始将蚁群算法、PSO 算法、灰狼算法等引入航迹规划, 以处理非线性和多约束的路径优化问题。这些方法通常将路径节点作为决策变量, 通过群体迭代搜索全局最优路径。张驰等[203]提出了蚁群-势场融合算法, 通过在传统、蚁群算法中引入人工势场法重新构建了启发函数, 同时, 采用最大-最小蚁群系统改进信息素更新规则。该方法有效避免了蚁群算法的局部最优与“早熟”收敛问题。孙玮萱等[204]提出了 PSO 灰狼混合算法, 在灰狼算法中融入 PSO 算法的个体最优位置记忆机制, 重构了速度和位置更新函数。该算法保留了灰狼算法平衡全局探索与局部开发的能力, 同时, 通过 PSO 算的记忆机制提升了收敛速度, 显著提升了潜器的水下导航效率。Wang 等[205]提出了基于 Voronoi 节点的逐步跟随 PSO 算法, 采用目标驱动的种群分步初始化提升

初始路径质量，利用障碍 Voronoi 节点引导粒子搜索方向以增加有效路径、加快收敛，设计振荡惯性权重平衡全局与局部搜索能力，显著提升路径规划的精度、效率与鲁棒性。值得注意的是，该算法将传统单一航程最短目标，扩展为航程代价、威胁代价、深度安全代价与路径平滑代价的加权组合，通过线性加权法将多目标优化问题转化为单目标优化问题，属于典型的多目标协同优化。

从定量结果来看，张驰等[203]提出的蚁群-势场算法在简单环境下相较于传统蚁群算法，航程缩短22.7%，计算时间减少 56.8%；孙玮萱等[204]提出的PSO 灰狼算法在复杂环境下相较于标准灰狼算法，迭代次数减少 82.43%、航程缩短 2.11%；相较于标准 PSO 算法，迭代次数减少 67.50%、航程缩短10.99%；Wang 等[205]提出的基于 Voronoi 节点逐步跟随 PSO 算法相较于传统随机初始化方法，初始路径质量平均提升 72% 以上；相较于标准 PSO 算法，平均适应度提升 28.13%。上述数据表明，该阶段方法在路径长度、收敛速度和全局寻优能力上全面优于传统智能优化算法，实现了从重力信息引导向智能化全局优化的跨越。

该阶段方法仍存在明显局限：算法对惯性权重、学习因子、信息素挥发系数等关键参数高度敏感，鲁棒性依赖人工调参；环境建模多采用硬阈值二值化划分或简单连续代价函数，未充分考虑重力场特征的方向性。

4.3 自适应与理论最优规划阶段

为解决参数敏感性和环境适应性问题，近年来，研究者开始引入自适应机制、连续适配性度量与理论最优性准则，推动航迹规划向无参数、自调整、精度驱动方向演进。Wang 等[206]针对传统航迹方法依赖经验阈值划分适配区、难以全局最优规划导航精度最优路径的问题，从估计理论出发，提出了一种不依赖适配区划分的路径规划方法。将后验克拉美罗界作为 A*算法的代价函数，以惯导位置误差方差为启发函数，构建以导航定位精度最优为目标的航迹规划方法；该方法无需将非适配区视为硬障碍，可在任意海域规划路径，为水下长航时高精度导航提供理论最优航迹方案。这一工作标志着航迹规划从启发式规则向理论最优性的转变。Zhang 等[207]针对传统航迹规划算法收敛慢、参数敏感、路径平滑性差且未充分融合重力多特征的问题，提出了自适应精英蚁群优化算法。该算法包含 2 项核心创新：①精英强化机制，对最优路径额外增加信息素浓度，加速收敛；②基于重力场空间变异性的动态参数自调整策略，使信息素浓度因子、启发因子和蒸发率能够根据搜索反馈自动变化。在环境建模方面，摒弃了硬阈值划分，采用模糊统计与熵权法，将多种重力特征合成为连续的综合适配性指标以划分适配区。该算法路径更短、转向更少、收敛更快，无需人工调参且环境适应性强，显著提升复杂水下环境中重力辅助导航的航迹规划性能。算法定量指标显示，Wang 等[206]实现了航迹规划在重力特征平缓区域的总均方根误差相较于直线最短路径降低约 50%，相较于基于传统阈值法降低约 55%；Zhang 等[207]提出的算法性能全面优于经典蚁群算法、精英蚁群算法及自适应蚁群算法等传统算法，规划航迹长度最多可缩短 19%，航迹转向次数减少 40%，收敛速度提升 95%。

该阶段方法的核心特征是：环境建模从硬阈值二值化发展为连续适配性度量；优化目标从多目标加权升级为直接面向导航定位精度的优化；算法参数从人工设定演进为动态自调整。这些进展使航迹规划能够真正适应复杂多变的实际海洋环境。

尽管上述方法在重力辅助导航航迹规划领域取得了显著进展，但受制于问题本身的复杂性与技术发展阶段的局限，仍存在以下共性局限：①环境建模多采用硬阈值二值化。丢失了重力场适配性的连续渐变信息，在适配区分布零散或全局适配性偏低的区域，易导致路径不可达或规划失败。②规划空间局限于二维水平面。重力场的垂直变化同样蕴含丰富的定位信息，忽视这一维度限制了导航性能的进一步提升。③对重力场方向性特征挖掘不足。同一区域在不同航行方向上的匹配增益存在显著差异，各向同性的代价函数无法引导航行器选择最优航向。④动态适应能力弱。多数智能算法的参数依赖人工设定或离线调优，无法根据环境变化自适应调整，导致在复杂动态场景下性能下降。

5 未来发展趋势

通过分析水下 GAINS 中重力基准图构建、适配性分析、匹配算法及航迹规划四大关键技术的相关研究，发现现有研究虽已取得长足进展，但仍面临若干问题，据此推测未来的研究与发展将围绕以下方向重点突破。

5.1 重力基准图构建

高精度、高分辨率的重力基准图是提升 GAINS 性能的根本。针对当前研究存在的不足，未来研究应聚焦以下方向。

1）融合物理信息与深度学习。纯数据驱动的深度学习方法面临物理可解释性差、泛化能力不足的问题。未来应着力发展物理信息神经网络，将地

球物理定律作为约束条件嵌入神经网络损失函数或网络结构; 利用全球重力场模型提供长波背景趋势, 使模型学习中高频细节, 减少对密集实测样本的依赖, 提升复杂区域的重建精度与可解释性。

2) 深化多源数据智能融合机制。不同来源的重力数据在分辨率、精度与覆盖范围上差异显著, 现有固定权重或频带截断的融合策略难以实现最优互补, 且高精度实测数据覆盖稀疏, 传统监督模型在稀缺区域泛化能力不足。未来应发展融合智能学习与自适应加权的数据融合机制: 一方面, 借助掩码自编码器、对比学习、循环一致性等自监督机制提取数据深层结构特征, 再结合少量标注样本进行半监督微调, 改善近海及数据空白区的融合与插值精度; 另一方面, 设计跨模态注意力融合网络, 以卫星测高数据为查询, 以船测数据为键值, 通过注意力机制动态提取不同来源数据中的互补信息, 同时, 引入门控融合机制, 由网络学习各数据源在空间位置上的置信度权重, 实现高质量数据主导、低质量数据抑制的自适应融合。

5.2　重力场适配性分析

适配性分析作为导航决策的前置核心依据, 未来研究内容将从粗略分区向多尺度、带方向、智能化、可定量的导航适配性评估演进。

1) 强化方向敏感性与尺度自适应。已有研究注意到重力场的方向性差异, 但尚未形成系统的方向适配性度量体系。未来将重点发展带航向的适配性评估方法, 给出最佳航行方向。同时, 借助多尺度分析技术自动识别不同尺寸、零散分布的适配区, 避免大尺度平滑导致小尺寸高适配区域丢失。

2) 发展智能适配区划分方法。针对传统适配区提取方法依赖主观参数和泛化能力不足的局限, 未来应发展数据驱动与弱监督、自监督学习相结合的智能适配区划分方法, 可借鉴计算机视觉领域的先进模型架构。例如, 可引入检测 Transformer (detection Trasformer, DETR)及其变体, 将适配区划分建模为端到端的集合预测问题, 利用 Transformer 的自注意力机制捕捉重力场特征图中长距离依赖关系, 直接输出适配区边界框; 对于形状不规则、分布零散的适配区, 则可借鉴 Mask2Former 的掩码注意力机制, 实现像素级精细分割。此外, 可结合对比学习框架, 基于适配区与非适配区的特征对比构造正负样本对, 以提升模型在有限标注样本下的特征判别能力, 降低对大量人工标注数据的依赖。

3) 面向匹配算法评估适配性、推动算法工程转化。突破现有适配性分析方法尚未考虑匹配算法特点开展适配区划分的局限, 建立面向具体匹配算法的适配性评价体系, 实现从“静态区域划分”走向“算法导向评价”。建立基于真实重力数据的开源基准数据集与评估标准, 推动适配性分析方法从仿真验证走向可复现、可对比的实际工程验证。

4) 融合适配性分析与匹配、规划。传统适配性分析大都停留在重力场特征提取与区域划分的局限, 建议下一步加大对如何建立重力变化特征与匹配精度定量关系的方法研究, 强化由区域重力场特征直接预测匹配定位误差。并将预测的定位精度作为约束条件加入重力匹配算法与航迹规划流程, 使匹配策略可依据区域适配性自适应调整, 航迹规划能够优先选择高精度导航区域, 为水下长航时高精度自主导航提供量化支撑。

5.3　重力匹配算法

针对目前重力匹配算法存在的主要问题, 未来重力匹配导航算法可从以下几个方面进行改进。

1) 序列匹配由批处理向在线实时演进。突破 TERCOM、ICCP 等算法批处理模式, 可引入滑动窗口、动态序列长度、自适应搜索域和仿射变换等策略, 以克服实时性差、航向敏感、大初始误差易发散等问题, 实现边行进边匹配的逐点实时修正。

2) 滤波匹配突破非线性非高斯约束。针对 EKF、UKF、CKF、PF、PMF 等滤波框架, 可发展抗差估计、自适应噪声修正、观测差分、优化重采样等方法, 抑制粒子退化、滤波发散及在重力平坦区的失配问题。考虑重点研究无线性化假设、带约束的滤波技术, 提升强噪声和弱重力特征环境下的定位可靠性。

3) 多维信息深度融合。从传统重力异常标量匹配拓展至重力梯度、矢量特征匹配, 从二维平面匹配延伸至三维空间匹配, 同时可将深度网络作为特征提取器, 将原始重力序列映射到高维空间进行相似性度量, 以充分挖掘重力场多维度观测信息, 提升弱适配区的可匹配性和定位精度。

4) 轻量化自适应多算法融合框架。针对当前多算法融合设计复杂、参数敏感、计算负担重的问题, 未来可设计轻量级算法选择门控网络, 以局部重力特征为输入, 输出各类匹配算法的匹配成功率, 实现不同匹配算法根据当前区域适配性自适应切换或加权融合; 同时, 可引入知识蒸馏技术, 将多算法融合模型的知识迁移至轻量化学生网络, 满足嵌入式实时计算约束。

5.4　航迹规划算法

航迹规划是实现重力辅助导航由被动匹配向主动引导跨越的核心环节, 未来将突破传统路径规

划的局限，转向以构建连续适配场、覆盖三维空间、设计方向敏感的代价函数、支持在线重规划的自主闭环规划模式。

1）从二值障碍到连续适配场的环境建模。改变将非适配区简单视为障碍的粗糙建模方式，将重力场适配性、地形障碍、航行安全、隐蔽性、能耗等因素统一建模为连续适配代价场，实现全海域可规划、非适配区可优化通过，避免因适配区分布零散导致的路径不可达问题，提高路径的可行性和适应性。

2）从二维平面到三维空间的规划拓展。突破传统二维平面路径规划的局限，充分挖掘重力场垂向梯度与三维空间变化特征，设计三维航迹规划方法。

3）从各向同性到方向敏感的代价函数设计。针对重力场在不同航向上的可匹配性差异，引入方向适配性指标，构造与航行方向相关的路径代价以规划最优航行方向，引导载体沿重力变化最显著的方向通过匹配区域。

4）从离线静态规划到在线自适应重规划。克服多数智能算法参数离线设定、规划路径固定的局限，发展基于实时匹配反馈的在线重规划能力。可采用滚动时域优化框架，在每个规划周期内，根据实时获取的重力测量值与基准图之间的匹配定位误差，动态调整后续航迹点，并以此向前滚动推进，使航行器始终趋近高适配区域。同时，将匹配定位误差作为反馈信息，修正规划模型中的适配性置信度，可引入深度Q网络（deep Q-network, DQN）、近端策略优化（proximal policy optimization, PPO）等深度强化学习方法，将匹配定位误差设计为奖励信号，训练端到端的重规划策略，使航行器在航行过程中自主决策航向，形成“规划-执行-评估-修正”的闭环。图7展示了完整的在线自适应重规划闭环流程。

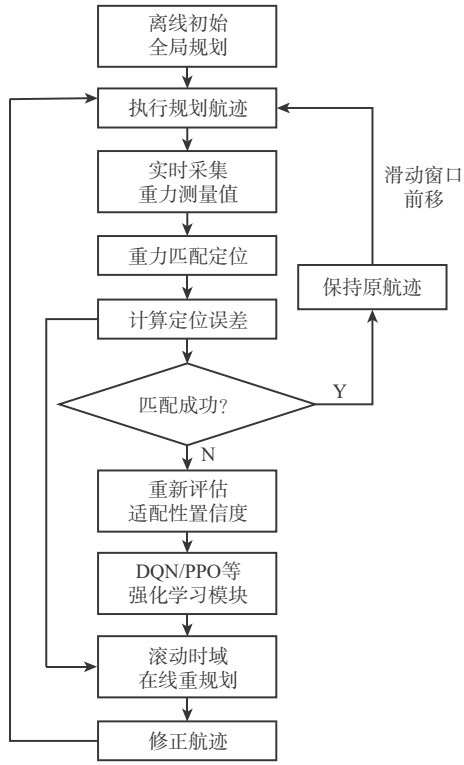


图7 在线自适应重规划闭环流程图

Fig. 7 Online adaptive reprogramming closed-loop flowchart

5.5 系统集成与工程验证

当前研究多聚焦于单项技术的理论改进与仿真验证，实际工程化应用仍面临多技术协同、实时计算、环境适应性等系统性挑战。未来应重点突破以下工程化问题：

1）多技术协同与一体化框架。重力基准图构建、适配性分析、匹配算法与航迹规划四大模块并非独立单元，其性能高度耦合。未来应发展一体化协同优化框架：将重力基准图的重构误差、适配性分析的不确定性、匹配算法的定位误差及航迹规划的代价函数纳入统一的目标函数，通过多目标优化或端到端联合学习实现全局最优，而非各模块的简单串联。例如，可将适配性分析结果作为匹配算法的先验信息，同时，将匹配误差反馈至适配性置信度更新，还可将重力基准图的重构误差纳入匹配定位的目标函数，或利用匹配误差BP修正重力基准图重建模型，形成跨模块闭环。

2）实时嵌入式实现。水下航行器有限的计算、存储与功耗资源对算法实时性提出了较高要求。未来需开展轻量化模型设计与嵌入式部署研究，针对深度学习模块，采用模型量化、知识蒸馏、网络剪枝等方法降低参数量和计算量，使其能在数字信号处理器（digital signal processor, DSP）、现场可编程门阵列（field-programmable gate array, FPGA）或嵌入式图形处理器（graphics processing unit, GPU）上实时运行；针对匹配算法，需优化其计算结构，降低在线递推的复杂度，满足水下实时导航的硬约束。

3）海试验证与环境适应性评估。当前绝大多数研究停留在仿真阶段，缺乏真实海洋环境下的闭环验证。未来应建立标准化的海试数据集与评估基准，包含不同海域、不同海况及不同航行模式的实测重力数据、惯导数据及高精度参考轨迹，供各类算法进行横向对比。同时，需开展长航时的海试验证，采集定位误差、匹配成功率、计算负载等指标，系统评估算法在真实动态环境下的稳定性、鲁棒性及误差漂移特性。

4）从仿真到实装的渐进式验证路线。建议

建立“仿真-半物理仿真-海试”的逐级验证体系。仿真阶段验证算法在理想条件下的性能边界; 半物理仿真阶段将真实惯导、重力仪数据注入算法模型, 模拟运载体动力学特性与传感器噪声; 海试阶段则在典型目标海域开展长航时任务, 全面评估系统在真实复杂环境下的稳定性、精度与可靠性。通过多层级、多场景的工程验证, 推动重力辅助惯性导航技术从实验室原理样机走向实用化设备部署。

6 结　论

本文围绕水下 GAINS 的四大关键技术开展系统综述。研究现状表明: 基准图构建已形成以卫星测高为主、多源融合与网格化插值协同精化的技术框架; 适配性分析已建立以特征融合与分区提取为核心的评估方法; 匹配算法已形成序列匹配、实时滤波与多算法融合并行发展的格局; 航迹规划已实现从被动规避向主动引导的初步跨越。

然而, 各技术方向仍面临 4 个层面共性瓶颈。数据支撑层面, 多源重力数据的精度、分辨率与覆盖范围不一致, 深度融合机制尚不完善; 环境表征层面, 重力场适配性分析尚未建立面向具体匹配算法的适配性评价体系, 与匹配精度的定量映射不足; 算法效能层面, 匹配算法难以兼顾实时性与鲁棒性, 多算法融合存在复杂度高与参数敏感性问题, 航迹规划缺乏与实时匹配的闭环反馈, 离线静态模式难以适应复杂海洋环境; 在工程应用层面, 四项关键技术长期割裂发展, 缺乏一体化协同框架, 轻量化与嵌入式部署尚存瓶颈, 制约了系统从仿真向设备的工程转化。

展望未来, 应重点突破以下方向: 发展融合物理约束与深度学习的基准图构建方法, 建立输入误差到输出不确定性的完整传递链条; 构建面向具体匹配算法和多尺度自适应能力的适配性评价体系, 实现从定性分区向定量精度估计的跨越; 发展轻量化自适应多匹配算法融合框架, 拓展多维重力特征输入与深度特征提取, 突破弱特征区匹配精度与实时性的制约; 构建基于连续适配代价场的三维航迹规划方法, 引入在线闭环重规划策略, 实现从离线静态规划向自适应决策的跨越; 建立多模块协同的一体化系统集成框架与轻量化部署体系, 推动系统从理论向实装的工程转化。

参考文献（References）

[1] 王博, 周明龙. 水下重力辅助导航适配区选取的研究进展[J]. 导航定位学报, 2020, 8(3): 32-39.
Wang B, Zhou M L. Perspective on matching area selection technology for underwater gravity aided navigation[J]. Journal of Navigation and Positioning, 2020, 8(3): 32-39(in Chinese).

[2] 郑伟, 付腾达, 李钊伟, 等. 潜器水下重力匹配导航研究进展[J]. 科学技术与工程, 2022, 22(18): 7721-7734.
Zheng W, Fu T D, Li Z W, et al. Progress of undersea gravity matching navigation[J]. Science Technology and Engineering, 2022, 22(18): 7721-7734(in Chinese).

[3] Zhao S J, Dou Z Y, Zhu H Z, et al. ID2TM: a novel iterative double-cross domain-center transfer-matching method for underwater gravity-aided navigation[J]. IEEE Internet of Things Journal, 2025, 12(11): 15415-15423.

[4] 黄谟涛, 欧阳永忠, 边少锋, 等. 美俄水下重力辅助惯性导航技术发展分析与思考[J]. 武汉大学学报 (信息科学版), 2024, 49(11): 1977-1991.
Huang M T, Ouyang Y Z, Bian S F, et al. Analysis and reflections on the development of underwater gravity-aided inertial navigation technology in the United States and Russia[J]. Geomatics and Information Science of Wuhan University, 2024, 49(11): 1977-1991 (in Chinese).

[5] Wang D, Wang B, Huang H Q, et al. A multisensor fusion method based on strict velocity for underwater navigation system[J]. IEEE Sensors Journal, 2023, 23(16): 18587-18598.

[6] Mao N, Li A, Xu J N, et al. Adaptive gravity-aided inertial navigation based on characteristic analysis of marine gravity anomaly from satellite altimetry[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 5920813.

[7] Sahoo A, Dwivedy S K, Robi P S. Advancements in the field of autonomous underwater vehicle[J]. Ocean Engineering, 2019, 181: 145-160.

[8] Karmozdi A, Hashemi M, Salarieh H, et al. INS-DVL navigation improvement using rotational motion dynamic model of AUV[J]. IEEE Sensors Journal, 2020, 20(23): 14329-14336.

[9] 孙明琦. 基于水下地貌的匹配导航算法研究[D]. 哈尔滨: 哈尔滨工程大学, 2018.
Sun M Q. Research on matching navigation algorithm based on underwater topography[D]. Harbin: Harbin Engineering University, 2018(in Chinese).

[10] 李姗姗. 水下重力辅助惯性导航的理论与方法研究[D]. 郑州: 解放军信息工程大学, 2010.
Li S S. Research on theory and method of underwater gravity aided inertial navigation[D]. Zhengzhou: PLA Information Engineering University, 2010(in Chinese).

[11] 张文君, 马腾, 李晔, 等. 水下地磁导航方法综述[J]. 哈尔滨工程大学学报, 2025, 46(12): 2583-2595.
Zhang W J, Ma T, Li Y, et al. A summary of underwater geomagnetic navigation methods[J]. Journal of Harbin Engineering University, 2025, 46(12): 2583-2595(in Chinese).

[12] 马明珠. 水下地磁辅助导航匹配算法研究[D]. 南京: 东南大学, 2019.
Ma M Z. Research on underwater geomagnetic assisted navigation matching algorithm[D]. Nanjing: Southeast University, 2019(in Chinese).

[13] Zhao S J, Zheng W, Li Z W, et al. Improving matching efficiency and out-of-domain positioning reliability of underwater gravity matching navigation based on a novel domain-center adaptive-

transfer matching method[J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 1001811.

[14] Zou J S, Cai T J. Improved particle swarm optimization screening iterative algorithm in gravity matching navigation[J]. IEEE Sensors Journal, 2022, 22(21): 20866-20876.

[15] Zhao S J, Zheng W, Li Z W, et al. A novel cross-line adaptive domain matching algorithm for underwater gravity aided navigation[J]. IEEE Geoscience and Remote Sensing Letters, 2024, 21: 7503605.

[16] 吴庭涛, 郑伟, 尹文杰, 等. 地球卫星重力场模型及其应用研究进展[J]. 科学技术与工程, 2020, 20(25): 10117-10132.
Wu T T, Zheng W, Yin W J, et al. Review of earth's gravity field model and application from satellite[J]. Science Technology and Engineering, 2020, 20(25): 10117-10132(in Chinese).

[17] 张驰. 重力辅助惯性导航的航迹规划与图像匹配方法研究[D]. 郑州: 战略支援部队信息工程大学, 2020.
Zhang C. Research on track planning and image matching methods of gravity-aided inertial navigation[D]. Zhengzhou: Information Engineering University, 2020(in Chinese).

[18] Sandwell D T, Smith W H F. Marine gravity anomaly from Geosat and ERS 1 satellite altimetry[J]. Journal of Geophysical Research: Solid Earth, 1997, 102(B5): 10039-10054.

[19] Sandwell D T, Müller R D, Smith W H F, et al. New global marine gravity model from CryoSat-2 and Jason-1 reveals buried tectonic structure[J]. Science, 2014, 346(6205): 65-67.

[20] Sandwell D T, Harper H, Tozer B, et al. Gravity field recovery from geodetic altimeter missions[J]. Advances in Space Research, 2021, 68(2): 1059-1072.

[21] Andersen O B, Knudsen P, Kenyon S, et al. Global and Arctic marine gravity field from recent satellite altimetry (DTU13)[C]// Proceedings of the 76th EAGE Conference and Exhibition. Houten: EAGE, 2014: 1-5.

[22] Andersen O B, Knudsen P. The DTU17 global marine gravity field: first validation results[C]//Proceedings of the Fiducial Reference Measurements for Altimetry. Berlin: Springer, 2020: 83-87.

[23] Zhang S J, Zhou R S, Jia Y J, et al. Performance of HaiYang-2 altimetric data in marine gravity research and a new global marine gravity model NSOAS22[J]. Remote Sensing, 2022, 14(17): 4322.

[24] Zhu C C, Guo J Y, Yuan J J, et al. SDUST2021GRA: global marine gravity anomaly model recovered from Ka-band and Ku-band satellite altimeter data[J]. Earth System Science Data, 2022, 14(10): 4589-4606.

[25] 刘乾坤, 陈兆钰, 许可. 卫星测高重力与船测重力协同融合的全球高分辨率海洋重力场模型[J]. 空间科学学报, 2026, 46(3): 850-858.
Liu Q K, Chen Z Y, Xu K. Global high-resolution marine gravity field model fused from altimetry-derived gravity and shipborne gravity[J]. Chinese Journal of Space Science, 2026, 46(3): 850-858 (in Chinese).

[26] Sandwell D, Garcia E, Soofi K, et al. Toward 1-mGal accuracy in global marine gravity from CryoSat-2, Envisat, and Jason-1[J]. The Leading Edge, 2013, 32(8): 892-899.

[27] Sandwell D T, Smith W H F. Global marine gravity from retracked Geosat and ERS-1 altimetry: ridge segmentation versus spreading rate[J]. Journal of Geophysical Research: Solid Earth, 2009, 114(B1): 2008JB006008.

[28] 聂琳娟, 超能芳, 金涛勇, 等. 卫星测高反演海洋重力异常的精度分析[J]. 测绘科学, 2016, 41(9): 1-6.
Nie L J, Chao N F, Jin T Y, et al. Accuracy analysis of marine gravity anomaly inferred by satellite altimetry data[J]. Science of Surveying and Mapping, 2016, 41(9): 1-6(in Chinese).

[29] 黄谟涛, 王瑞, 翟国君, 等. 多代卫星测高数据联合平差及重力场反演[J]. 武汉大学学报 (信息科学版), 2007, 32(11): 988-993.
Huang M T, Wang R, Zhai G J, et al. Integrated data processing for multi-satellite missions and recovery of marine gravity field[J]. Geomatics and Information Science of Wunan University, 2007, 32(11): 988-993(in Chinese).

[30] 周瑞宸. 基于卫星测高的海洋垂直重力梯度反演研究[D]. 青岛: 山东科技大学, 2024: 28-30.
Zhou R C. Study on inversion of vertical gravity gradient of ocean based on satellite altimetry[D]. Qingdao: Shandong University of Science and Technology, 2024: 28-30(in Chinese).

[31] 张思慧, 吴云龙, 张毅, 等. 一种基于联合变分自编码器的卫星重力数据粗差探测方法研究[J]. 武汉大学学报 (信息科学版), 2024, 49(6): 986-995.
Zhang S H, Wu Y L, Zhang Y, et al. A gross error detection method of satellite gravity data based on joint variational autoencoder[J]. Geomatics and Information Science of Wuhan University, 2024, 49(6): 986-995(in Chinese).

[32] 柳林涛, 梁星辉, 郎骏健, 等. 改进的函数滤波器及其在航空重力测量中的应用[J]. 测绘学报, 2023, 52(5): 706-713.
Liu L T, Liang X H, Lang J J, et al. Improved function filter and its application in airborne gravimetry[J]. Acta Geodaetica et Cartographica Sinica, 2023, 52(5): 706-713(in Chinese).

[33] 黄谟涛, 欧阳永忠, 邓凯亮, 等. 兼顾精度与分辨需求的海空重力测量误差分离与补偿方法[J]. 地球物理学报, 2018, 61(8): 3160-3169.
Huang M T, Ouyang Y Z, Deng K L, et al. An error separation and compensation method for marine-airborne gravity measurement balancing accuracy and resolution requirements[J]. Chinese Journal of Geophysics, 2018, 61(8): 3160-3169(in Chinese).

[34] 王皓, 熊盛青, 何涛, 等. 最小曲率模型的航空重力数据噪声压制方法研究[J]. 地球物理学报, 2024, 67(10): 3959-3972.
Wang H, Xiong S Q, He T, et al. Noise suppression of airborne gravity data with minimum curvature model[J]. Chinese Journal of Geophysics, 2024, 67(10): 3959-3972(in Chinese).

[35] 熊盛青, 王冠鑫, 罗锋, 等. 一种基于 K-F 滤波的航空重力数据的处理方法及系统: CN202510209565.9[P]. 2025-09-09.
Xiong S Q, Wang G X, Luo F, et al. A method and system for processing airborne gravity data based on K-F filtering: CN202510209565.9[P]. 2025-09-09(in Chinese).

[36] Guo J Y, Zhang H F, Zhang H Y, et al. Joint reprocessing method of multivintage shipborne gravity anomalies considering temporal error effects: a case study of the Philippine sea[J]. Geophysical Journal International, 2025, 242(3): 276.

[37] 郭金运, 张鸿飞, 李真, 等. 多源船载重力异常数据的联合再处理: 以墨西哥湾为例[J]. 武汉大学学报 (信息科学版), 2025, 50(7): 1277-1290.
Guo J Y, Zhang H F, Li Z, et al. Joint reprocessing of multi-sources shipborne gravity anomalies: a case study of gulf of Mex-

ico[J]. Geomatics and Information Science of Wuhan University, 2025, 50(7): 1277-1290(in Chinese).

[38] Abulkhair M, Zaki A, Elhadary A, et al. Optimizing shipborne gravity data around Egypt for enhanced geomatics applications[J]. The Egyptian Journal of Remote Sensing and Space Sciences, 2026, 29(1): 104-119.

[39] Bako M, Kusche J. Evaluation and homogenization of a marine gravity database from shipborne and satellite altimetry-derived gravity data over the coastal region of Nigeria[J]. Journal of Applied Geodesy, 2025, 19(2): 175-189.

[40] 张敬伟. 布尔莎模型坐标转换适用范围及精度分析[J]. 测绘与空间地理信息, 2013, 36(1): 175-176.
Zhang J W. The using range and precision analysis for bursa model coordinate conversion[J]. Geomatics & Spatial Information Technology, 2013, 36(1): 175-176(in Chinese).

[41] 欧阳永忠. 海空重力测量数据处理关键技术研究[D]. 武汉: 武汉大学, 2013.
Ouyang Y Z. On key technologies of data processing for air-sea gravity surveys[D]. Wuhan: Wuhan University, 2013(in Chinese).

[42] 束蝉方, 李斐, 张利明. 基于 EGM2008 重力场模型的局部高程基准统一[J]. 地球物理学进展, 2011, 26(2): 438-442.
Shu C F, Li F, Zhang L M. Local height datum unification using EGM2008[J]. Progress in Geophysics, 2011, 26(2): 438-442(in Chinese).

[43] 劳浩瑜, 张兴福, 孙炜程, 等. 多源重力场模型融合方法与精度分析[J]. 地球物理学进展, 2025, 40(5): 1905-1918.
Lao H Y, Zhang X F, Sun W C, et al. Combination methods and precision analysis of multi-source global geopotential models[J]. Progress in Geophysics, 2025, 40(5): 1905-1918(in Chinese).

[44] 刘鑫宇, 李姗姗, 张盼盼, 等. 顾及剩余地形影响的局部高程基准统一方法[J]. 武汉大学学报 (信息科学版), 2025, 50(5): 898-906.
Liu X Y, Li S S, Zhang P P, et al. Local vertical datum unification method considering effect of residual terrain[J]. Geomatics and Information Science of Wuhan University, 2025, 50(5): 898-906(in Chinese).

[45] 徐卫明, 舒嵘. 利用 GPS 接收机实现多传感器同步工作的两种方法[J]. 遥感技术与应用, 2003, 18(6): 404-406.
Xu W M, Shu R. Two methods of multi-sensor synchronization with GPS receiver[J]. Remote Sensing Technology and Application, 2003, 18(6): 404-406(in Chinese).

[46] Jiang M F, Xu K, Liu Y L, et al. Estimating the time tag bias of HY-2A radar altimeter and its application to dual-frequency ionosphere correction[J]. Marine Geodesy, 2017, 40(6): 361-377.

[47] 吴涛, 朱建良, 周维. 基于补偿 GNSS 滞后的实时外推时间配准方法[J]. 淮阴工学院学报, 2022, 31(1): 54-59.
Wu T, Zhu J L, Zhou W. Real-time extrapolation time registration method based on compensating GNSS lag[J]. Journal of Huaiyin Institute of Technology, 2022, 31(1): 54-59(in Chinese).

[48] Fan X, Guo J Y, Zhang H Y, et al. A two-step method of crossover adjustment for satellite altimeter data[J]. Advances in Space Research, 2025, 75(1): 219-232.

[49] 刘晓刚, 闫志闯, 孙文, 等. 确定地球重力场模型的最小二乘配置法与调和分析法的精度评析[J]. 地球物理学进展, 2014, 29(1): 46-50.
Liu X G, Yan Z C, Sun W, et al. Precision evaluation and analysis of least squares collocation method and spherical harmonic analysis method in the determination of the Earth's gravity field model[J]. Progress in Geophysics, 2014, 29(1): 46-50(in Chinese).

[50] 潘宗鹏, 王伟, 祖安然. 多源重力场信息融合处理方法[J]. 测绘科学技术学报, 2013, 30(3): 236-240.
Pan Z P, Wang W, Zu A R. Fusion methods of multi-source gravity field information[J]. Journal of Geomatics Science and Technology, 2013, 30(3): 236-240(in Chinese).

[51] 翟振和, 孙中苗. 基于最小二乘配置的航空和地面重力数据融合处理[J]. 测绘科学与工程, 2008(1): 10-14.
Zhai Z H, Sun Z M. Combining airborne gravity data and terrestrial gravity data based on least squares collocation[J]. Geomatics Science and Engineering, 2008(1): 10-14(in Chinese).

[52] Ophaug V, Gerlach C. On the equivalence of spherical splines with least-squares collocation and Stokes's formula for regional geoid computation[J]. Journal of Geodesy, 2017, 91(11): 1367-1382.

[53] Märdla S, Ellmann A, Ågren J, et al. Regional geoid computation by least squares modified Hotine's formula with additive corrections[J]. Journal of Geodesy, 2018, 92(3): 253-270.

[54] 李芳明. 水下重力辅助定位及重力图导航适配性研究[D]. 哈尔滨: 哈尔滨工程大学, 2020.
Li F M. Research on underwater gravity aided positioning and gravity map navigability analysis[D]. Harbin: Harbin Engineering University, 2020(in Chinese).

[55] 黄谟涛, 欧阳永忠, 翟国君, 等. 融合多源重力数据的 Tikhonov 正则化配置法[J]. 海洋测绘, 2013, 33(3): 6-12.
Huang M T, Ouyang Y Z, Zhai G J, et al. Tikhonov regularization collocation for multi-source gravity data fusion processing[J]. Hydrographic Surveying and Charting, 2013, 33(3): 6-12(in Chinese).

[56] 高新兵, 李姗姗, 李海, 等. 点质量模型与最小二乘配置在多源重力数据融合中的应用[J]. 大地测量与地球动力学, 2013, 33(1): 145-149.
Gao X B, Li S S, Li H, et al. Application of point mass model and least square collocation in multi-source gravity data fusion[J]. Journal of Geodesy and Geodynamics, 2013, 33(1): 145-149(in Chinese).

[57] 孙文, 吴晓平, 王庆宾, 等. 基于方差分量估计的正则化配置法及其在多源重力数据融合中的应用[J]. 武汉大学学报 (信息科学版), 2016, 41(8): 1087-1092.
Sun W, Wu X P, Wang Q B, et al. Normalized collocation based on variance component estimate and its application in multi-source gravity data fusion[J]. Geomatics and Information Science of Wuhan University, 2016, 41(8): 1087-1092(in Chinese).

[58] Sampietro D, Capponi M, Mansi A H, et al. Space-Wise approach for airborne gravity data modelling[J]. Journal of Geodesy, 2017, 91(5): 535-545.

[59] 吴怿昊, 罗志才. 融合多源数据的局部重力场建模中线性化误差的研究[J]. 大地测量与地球动力学, 2016, 36(6): 476-480.
Wu Y H, Luo Z C. The research on linearization errors in regional gravity field modeling based on heterogeneous data sets[J]. Journal of Geodesy and Geodynamics, 2016, 36(6): 476-480(in Chinese).

[60] 吴健生, 刘苗. 基于小波的位场数据融合[J]. 同济大学学报 (自然科学版), 2008, 36(8): 1133-1137.
Wu J S, Liu M. Potential field data fusion based on wavelet multiresolution decomposition[J]. Journal of Tongji University (Natural

Science), 2008, 36(8): 1133-1137(in Chinese).

[61] Bolkas D, Fotopoulos G, Braun A. Comparison and fusion of satellite, airborne, and terrestrial gravity field data using wavelet decomposition[J]. Journal of Surveying Engineering, 2016, 142(2): 04015010.

[62] Bai Y L, Dong D D, Wu S G, et al. A wavelet transformation approach for multi-source gravity fusion: applications and uncertainty tests[J]. Journal of Applied Geophysics, 2016, 128: 18-30.

[63] Ma L, Song H B, Bai Y L, et al. Gravity data fusion using wavelet transform and window weighting: a case study in the Ross Sea of Antarctica[J]. Acta Geophysica, 2025, 73(1): 187-201.

[64] Chen X, Kong X X, Zhou R S, et al. Fusion of altimetry-derived model and ship-borne data in preparation of high-resolution marine gravity determination[J]. Geophysical Journal International, 2024, 236(3): 1262-1274.

[65] Lin M, Denker H, Müller J. Regional gravity field modeling using free-positioned point masses[J]. Studia Geophysica et Geodaetica, 2014, 58(2): 207-226.

[66] Kern M. An analysis of the combination and downward continuation of satellite, airborne and terrestrial gravity data[D]. Calgary: University of Calgary, 2003.

[67] Tenzer R, Klees R. The choice of the spherical radial basis functions in local gravity field modeling[J]. Studia Geophysica et Geodaetica, 2008, 52(3): 287-304.

[68] 黄谟涛, 翟国君, 管铮, 等. 海洋重力场测定及其应用[M]. 北京: 测绘出版社, 2005: 266-296.
Huang M T, Zhai G J, Guan Z, et al. Determination and application of marine gravity field[M]. Beijing: Sino Maps Press, 2005: 266-296(in Chinese).

[69] 李建成. 地球重力场逼近理论与中国 2000 似大地水准面的确定[M]. 武汉: 武汉大学出版社, 2003: 30-37.
Li J C. Approximation theory of earth gravity field and determination of China 2000 quasi-geoid[M]. Wuhan: Wuhan University Press, 2003: 30-37(in Chinese).

[70] 黄谟涛, 欧阳永忠, 刘敏, 等. 融合海域多源重力数据的正则化点质量方法[J]. 武汉大学学报 (信息科学版), 2015, 40(2): 170-175.
Huang M T, Ouyang Y Z, Liu M, et al. Regularization of point-mass model for multi-source gravity data fusion processing[J]. Geomatics and Information Science of Wuhan University, 2015, 40(2): 170-175(in Chinese).

[71] 成怡. 多源海洋重力数据融合技术研究[D]. 哈尔滨: 哈尔滨工程大学, 2008.
Cheng Y. A study on heterogeneous marine gravity data combination technology[D]. Harbin: Harbin Engineering University, 2008 (in Chinese).

[72] Wittwer T. Regional gravity field modelling with radial basis functions[D]. Delft: Delft University of Technology, 2009.

[73] Klees R, Tenzer R, Prutkin I, et al. A data-driven approach to local gravity field modelling using spherical radial basis functions[J]. Journal of Geodesy, 2008, 82(8): 457-471.

[74] Bentel K, Schmidt M. Combining different types of gravity observations in regional gravity modeling in spherical radial basis functions[C]//Proceedings of the VIII Hotine-Marussi Symposium on Mathematical Geodesy. Berlin: Springer, 2015: 115-120.

[75] 吴怿昊, 罗志才, 周波阳. 基于泊松小波径向基函数融合多源数据的局部重力场建模[J]. 地球物理学报, 2016, 59(3): 852-864.
Wu Y H, Luo Z C, Zhou B Y. Regional gravity modeling based on heterogeneous data sets by using Poisson wavelets radial basis functions[J]. Chinese Journal of Geophysics, 2016, 59(3): 852-864(in Chinese).

[76] Mahbuby H, Safari A, Foroughi I. Local gravity field modeling using spherical radial basis functions and a genetic algorithm[J]. Comptes Rendus Géoscience, 2017, 349(3): 106-113.

[77] 吴怿昊, 何秀凤, 罗志才, 等. 基于航空重力梯度数据重建高分辨率局部重力场的径向基函数方法[J]. 地球物理学报, 2024, 67(5): 1913-1926.
Wu Y H, He X F, Luo Z C, et al. The approach for high-resolution local gravity field recovery from airborne gravity gradient data based on radial basis functions[J]. Chinese Journal of Geophysics, 2024, 67(5): 1913-1926(in Chinese).

[78] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 1646-1654.

[79] Deudon M, Kalaitzis A, Goytom I, et al. HighRes-net: recursive fusion for multi-frame super-resolution of satellite imagery [EB/OL]. (2020-02-15)[2026-04-05]. https://doi.org/10.48550/arXiv.2002.06460.

[80] Zhu C C, Guo J Y, Yuan J J, et al. Refining altimeter-derived gravity anomaly model from shipborne gravity by multi-layer perceptron neural network: a case in the South China Sea[J]. Remote Sensing, 2021, 13(4): 607.

[81] Alaofin O, Zhang Y, Sharma J, et al. Cross-modality super-resolution of satellite gravity data for geophysical exploration[C]//Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Piscataway: IEEE Press, 2022: 7539-7542.

[82] 王逸宸, 柳林涛, 许厚泽. 利用卷积自编码器重建含噪重力数据[J]. 武汉大学学报 (信息科学版), 2022, 47(4): 543-550.
Wang Y C, Liu L T, Xu H Z. Noisy gravity data reconstruction using the convolutional autoencoder[J]. Geomatics and Information Science of Wuhan University, 2022, 47(4): 543-550(in Chinese).

[83] Xiao C J, Guo J Y, Zhu C C, et al. Refining satellite Altimetry-Derived gravity anomaly model with shipborne gravity using multilayer perceptron neural networks[J]. Scientific Reports, 2025, 15: 19467.

[84] Zhang Y. Deep learning based super resolution of gravity data for geophysical exploration[D]. Baton Rouge: Louisiana State University, 2021.

[85] 孙明智, 冯伟, 安德超, 等. 基于多源数据和卷积神经网络的海洋重力异常反演[J]. 地球物理学报, 2025, 68(11): 4145-4156.
Sun M Z, Feng W, An D C, et al. Marine gravity anomaly recovery based on multi-source data and convolutional neural network [J]. Chinese Journal of Geophysics, 2025, 68(11): 4145-4156(in Chinese).

[86] Li Q Q, Zhai Z H, Bao L F, et al. A convolutional neural network to optimize multi-mission satellite altimeter fusion for improving the marine gravity field[J]. Earth, Planets and Space, 2024, 76(1): 129.

[87] Jia B D, Sun J, Geng X F, et al. Super-resolution reconstruction of gravity data using semi-supervised dual regression learning[J]. Re-

mote Sensing, 2026, 18(3): 453.

[88] Jia R C, Yang Y C, Wang Q B, et al. Multi-source geodetic data fusion using a physically informed swin transformer for high-resolution gravity field recovery: a case study of the South China Sea[J]. Journal of Marine Science and Engineering, 2026, 14(4): 403.

[89] Reyes O, Cano A, Fardoun H M, et al. A locally weighted learning method based on a data gravitation model for multi-target regression[J]. International Journal of Computational Intelligence Systems, 2018, 11(1): 282-295.

[90] 徐遵义, 姜玉祥, 赵亮, 等. 改进的 Shepard 算法及其在重力异常插值中的应用[J]. 武汉大学学报 (信息科学版), 2010, 35(4): 477-480.
Xu Z Y, Jiang Y X, Zhao L, et al. Improved Shepard method and its application in gravity field data interpolation[J]. Geomatics and Information Science of Wuhan University, 2010, 35(4): 477-480 (in Chinese).

[91] Hansen R O. Interpretive gridding by anisotropic Kriging[J]. Geophysics, 1993, 58(10): 1491-1497.

[92] Wright G B. Radial basis function interpolation: numerical and analytical developments[D]. Boulder: University of Colorado, 2003.

[93] Wessel P. A general-purpose Green's function-based interpolator[J]. Computers & Geosciences, 2009, 35(6): 1247-1254.

[94] Wessel P, Bercovici D. Interpolation with splines in tension: a Green's function approach[J]. Mathematical Geology, 1998, 30(1): 77-93.

[95] 韩雨蓉. 水下导航重力匹配算法研究[D]. 北京: 北京理工大学, 2017.
Han Y R. Research on gravity matching algorithm for underwater navigation[D]. Beijing: Beijing Institute of Technology, 2017(in Chinese).

[96] 吴太旗, 黄谟涛, 欧阳永忠, 等. 高精度海洋重力异常格网插值技术研究[J]. 测绘科学, 2008, 33(5): 70-72.
Wu T Q, Huang M T, Ouyang Y Z, et al. The study of high precision interpolation technology in marine gravity anomaly[J]. Science of Surveying and Mapping, 2008, 33(5): 70-72(in Chinese).

[97] 刘晓刚, 吴晓平, 王宝军, 等. 卫星重力梯度数据的网格化方法研究[J]. 大地测量与地球动力学, 2010, 30(6): 60-65.
Liu X G, Wu X P, Wang B J, et al. Study on gridding methods of satellite gravity gradient data[J]. Journal of Geodesy and Geodynamics, 2010, 30(6): 60-65(in Chinese).

[98] 李振海, 汪海洪. 重力数据网格化方法比较[J]. 大地测量与地球动力学, 2010, 30(1): 140-144.
Li Z H, Wang H H. Comparison among methods for gravity data gridding[J]. Journal of Geodesy and Geodynamics, 2010, 30(1): 140-144(in Chinese).

[99] 江东, 王庆宾, 王凯. 重力场精化的空间插值方法研究[C]//中国测绘学会学术年会会议论文集. 北京: 中国测绘学会, 2010: 181-185.
Jiang D, Wang Q B, Wang K. Research on spatial interpolation methods for gravity field refinement[C]//Proceedings of the China Society of Geodesy, Photogrammetry and Cartography Academic Annual Conference. Beijing: China Society of Geodesy, Photogrammetry and Cartography, 2010: 181-185(in Chinese).

[100] 周文健, 高春峰, 魏国, 等. 空间插值算法在海洋重力基准图构建中的适用性分析[C]//中国惯性技术学会科技工作者研讨会会议论文集——新型惯性元件与先进导航技术. 北京: 中国惯性技术学会, 2023: 133-142.
Zhou W J, Gao C F, Wei G, et al. Applicability analysis of spatial interpolation algorithms in marine gravity reference map construction[C]//Proceedings of the Science and Technology Workers Symposium of Chinese Society of Inertial Technology-New Inertial Components and Advanced Navigation Technology. Beijing: Chinese Society of Inertial Technology, 2023: 133-142(in Chinese).

[101] While J, Biegert E, Jackson A. Interpolation of gravity and gravity gradient data by using the generalized sampling expansion: theory[J]. Geophysics, 2008, 73(2): I11-I21.

[102] 刘繁明, 姚剑奇, 荆心. 由小波与随机采样提升重力场插值精度的策略[J]. 大地测量与地球动力学, 2015, 35(6): 948-952.
Liu F M, Yao J Q, Jing X. A strategy for upgrading the interpolation accuracy of gravity map based on wavelet analysis[J]. Journal of Geodesy and Geodynamics, 2015, 35(6): 948-952(in Chinese).

[103] 文超斌, 王跃钢, 郭志斌, 等. 重力辅助导航高斯插值精化算法[J]. 测绘学报, 2015, 44(1): 13-18.
Wen C B, Wang Y G, Guo Z B, et al. Gravity aided navigation precise algorithm with Gauss spline interpolation[J]. Acta Geodaetica et Cartographica Sinica, 2015, 44(1): 13-18(in Chinese).

[104] 任强强, 王跃钢, 文超斌, 等. 局部连续场下重力基准图插值优选算法[J]. 大地测量与地球动力学, 2016, 36(3): 198-201.
Ren Q Q, Wang Y G, Wen C B, et al. Interpolation optimization algorithm for gravity reference map in local anomaly fields[J]. Journal of Geodesy and Geodynamics, 2016, 36(3): 198-201(in Chinese).

[105] 曾敏, 汪振华, 邓凯亮, 等. 借助 EGM2008 的海域重力异常反距离加权插值算法[J]. 海洋测绘, 2018, 38(2): 5-7.
Zeng M, Wang Z H, Deng K L, et al. An inverse distance interpolation algorithm of marine gravity anomaly uniting EGM2008 gravity field model[J]. Hydrographic Surveying and Charting, 2018, 38(2): 5-7(in Chinese).

[106] Sarzeaud O, LeQuentrec-Lalancette M F, Rouxel D. Optimal interpolation of gravity maps using a modified neural network[J]. Mathematical Geosciences, 2009, 41(4): 379-395.

[107] 佘雅文, 付广裕. 基于循环神经网络的重力异常数据推估研究[J]. 大地测量与地球动力学, 2021, 41(3): 234-237.
She Y W, Fu G Y. Estimation of gravity anomaly data based on recurrent neural network[J]. Journal of Geodesy and Geodynamics, 2021, 41(3): 234-237(in Chinese).

[108] 黄子炎, 王庆宾, 李锁柱, 等. 利用 Kmeans-RBF 神经网络提高重力格网化数据精度[J]. 海洋测绘, 2021, 41(4): 43-47.
Huang Z Y, Wang Q B, Li S Z, et al. Using Kmeans-RBF neural network to improve the accuracy of grid gravity data[J]. Hydrographic Surveying and Charting, 2021, 41(4): 43-47(in Chinese).

[109] Mao N, Li A, Xu J N, et al. Marine gravity anomaly from satellite altimetry: interpolation and matching navigation[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 5918919.

[110] 谢心和, 赵东明, 刘长青. 一种基于 PSO-BP 神经网络的重力数据插值方法[J]. 测绘与空间地理信息, 2023, 46(11): 26-29.
Xie X H, Zhao D M, Liu C Q. A gravity data interpolation method based on PSO-BP neural network[J]. Geomatics & Spatial Information Technology, 2023, 46(11): 26-29(in Chinese).

[111] 谢心和, 赵东明, 刘长青. Legendre 多项式神经网络在重力数据插值中的应用[J]. 测绘与空间地理信息, 2023, 46(12): 19-23.
Xie X H, Zhao D M, Liu C Q. Application of Legendre polynomial-neural network in gravity data interpolation[J]. Geomatics & Spatial Information Technology, 2023, 46(12): 19-23(in Chinese).

[112] Feng T, Li T J. High-resolution reconstruction of gravity reference map via BP-boosting-driven interpolation methods[C]//Proceedings of the 3rd International Conference on Computer Graphics and Image Processing. Piscataway: IEEE Press, 2025: 39-44.

[113] 於青凤, 熊杰, 龚永琪. 基于扩散模型的重力数据网格化方法[J]. 地质论评, 2025, 99(6): 1-11.
Yu Q F, Xiong J, Gong Y Q. Gridding method of gravity data based on diffusion model[J]. Geological Review, 2025, 99(6): 1-11 (in Chinese).

[114] Zhang S Q, Gao L N, Zhou F H, et al. Asymmetric dual-branch network for gravity anomaly reference map super-resolution in underwater gravity matching navigation[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74: 8512416.

[115] 李艳, 刘繁明. 重力图的可导航性分析[J]. 应用科技, 2010(12): 60-63.
Li Y, Liu F M. Navigability analysis of gravity map[J]. Applied Science and Technology, 2010(12): 60-63(in Chinese).

[116] 李姗姗, 吴晓平, 张传定, 等. 我国重力场新的统计特征参数的计算分析[J]. 地球物理学报, 2010, 53(5): 1099-1108.
Li S S, Wu X P, Zhang C D, et al. Calculation and analysis of the new statistical character parameters of gravity field in China[J]. Chinese Journal of Geophysics, 2010, 53(5): 1099-1108(in Chinese).

[117] Liu F M, Li F M, Jing X. Navigability analysis of local gravity map with projection pursuit-based selection method by using gravitation field algorithm[J]. IEEE Access, 2019, 7: 75873-75889.

[118] 马越原, 欧阳永忠, 黄谟涛, 等. 基于重力场特征参数信息熵的适配区选择方法[J]. 中国惯性技术学报, 2016, 24(6): 763-768.
Ma Y Y, Ouyang Y Z, Huang M T, et al. Selection method for gravity-field matchable area based on information entropy of characteristic parameters[J]. Journal of Chinese Inertial Technology, 2016, 24(6): 763-768(in Chinese).

[119] Wang B, Zhu Y W, Deng Z H, et al. The gravity matching area selection criteria for underwater gravity-aided navigation application based on the comprehensive characteristic parameter[J]. IEEE/ASME Transactions on Mechatronics, 2016, 21(6): 2935-2943.

[120] Wang B, Zhou M L, Deng Z H, et al. Sum vector-difference-based matching area selection method for underwater gravity-aided navigation[J]. IEEE Access, 2019, 7: 123616-123624.

[121] Ma Z X, Wang B, Wang X Y, et al. Gravitational texture feature based directional adaptability analysis method for gravity-aided navigation[J]. IEEE/ASME Transactions on Mechatronics, 2025, 30(6): 6397-6406.

[122] Wang M C, Zheng W, Zhao S J, et al. A novel curvature directional feature-based matching area clustering method for underwater gravity-aided navigation[J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74: 8514710.

[123] Zhang W Z, Wang Y, Zhao S W, et al. A gravity matching adaptability analysis method based on gravity directional characteristics[J]. IEEE Sensors Journal, 2026, 26(7): 10731-10744.

[124] Xiong L, Li K H, Tang J Q, et al. Research on matching area selection criteria for gravity gradient navigation based on principal component analysis and analytic hierarchy process[C]//Proceedings of the Ninth International Symposium on Multispectral Image Processing and Pattern Recognition. Bellingham: SPIE, 2015, 9815: 981519.

[125] 李钊伟, 郑伟, 房静, 等. 基于新型主成分加权平均归一化法优选水下重力匹配导航适配区[J]. 地球物理学报, 2019, 62(9): 3269-3278.
Li Z W, Zheng W, Fang J, et al. Optimizing suitability area of underwater gravity matching navigation based on a new principal component weighted average normalization method[J]. Chinese Journal of Geophysics, 2019, 62(9): 3269-3278(in Chinese).

[126] 张辰, 马林, 张桂敏, 等. 一种新的重力辅助导航海区适配性分析方法[C]//中国惯性技术学会第七届学术年会会议论文集. 北京: 中国惯性技术学会, 2015: 130-135.
Zhang C, Ma L, Zhang G M, et al. A new method for sea area suitability analysis of gravity-aided navigation[C]//Proceedings of the 7th Academic Annual Conference of Chinese Society of Inertial Technology. Beijing: Chinese Society of Inertial Technology, 2015: 130-135(in Chinese).

[127] 蔡体菁, 陈鑫巍. 基于层次分析法的重力匹配区域选择准则[J]. 中国惯性技术学报, 2013, 21(1): 93-96.
Cai T J, Chen X W. Selection criterion based on analytic hierarchy process for matching region in gravity aided INS[J]. Journal of Chinese Inertial Technology, 2013, 21(1): 93-96(in Chinese).

[128] 马越原, 欧阳永忠, 曲政豪, 等. 基于模糊决策理论的重力辅助导航区域选择方法[J]. 海洋测绘, 2016, 36(6): 7-10.
Ma Y Y, Ouyang Y Z, Qu Z H, et al. Selection criteria of gravity aided navigation area based on fuzzy decision theory[J]. Hydrographic Surveying and Charting, 2016, 36(6): 7-10(in Chinese).

[129] 徐晓苏, 汤郡郡, 张涛, 等. 基于熵值法赋权灰色关联决策的地形辅助导航适配区选择[J]. 中国惯性技术学报, 2015, 23(2): 201-206.
Xu X S, Tang J J, Zhang T, et al. Selection for matching area in terrain aided navigation based on entropy-weighted grey correlation decision-making[J]. Journal of Chinese Inertial Technology, 2015, 23(2): 201-206(in Chinese).

[130] 宫京, 张崇猛, 周贤高, 等. 一种基于因子分析的重力适配区域选取方法[J]. 中国惯性技术学报, 2019, 27(6): 732-737.
Gong J, Zhang C M, Zhou X G, et al. Selection method of gravity adaptive region based on factor analysis[J]. Journal of Chinese Inertial Technology, 2019, 27(6): 732-737(in Chinese).

[131] 肖云, 张锦柏, 曹杰, 等. 多属性决策理论的重力匹配导航适配性分析[J]. 武汉大学学报 (信息科学版), 2023, 48(7): 1089-1099.
Xiao Y, Zhang J B, Cao J, et al. Suitability analysis of gravity matching navigation based on multiple attribute decision making theory[J]. Geomatics and Information Science of Wuhan University, 2023, 48(7): 1089-1099(in Chinese).

[132] 陈垲宁, 肖云, 张锦柏, 等. 水下重力匹配导航适配性评价方法比较研究[J]. 大地测量与地球动力学, 2024, 44(7): 737-743.
Chen K N, Xiao Y, Zhang J B, et al. Comparison of adaptive evaluation methods of underwater gravity matching navigation[J]. Journal of Geodesy and Geodynamics, 2024, 44(7): 737-743(in Chinese).

[133] 张勇刚, 张硕, 李宁, 等. 基于深度特征探索的重力适配区选择算法[J]. 中国惯性技术学报, 2025, 33(6): 546-558.
Zhang Y G, Zhang S, Li N, et al. Selection algorithm for gravity matching region based on deep feature exploration[J]. Journal of Chinese Inertial Technology, 2025, 33(6): 546-558(in Chinese).

[134] Tang J Q, Xiong L, Li K H, et al. Research on matching area selection based on multiple features fusion of full tensor gravity gradient[C]//Proceedings of the Automatic Target Recognition and Navigation. Bellingham: SPIE, 2015, 9812: 98120F-5-98120F-6.

[135] 程力, 张雅杰, 蔡体菁. 重力辅助导航匹配区域选择准则[J]. 中国惯性技术学报, 2007, 15(5): 559-563.
Cheng L, Zhang Y J, Cai T J. Selection criterion for matching area in gravity aided navigation[J]. Journal of Chinese Inertial Technology, 2007, 15(5): 559-563(in Chinese).

[136] Li K H, Xiong L, Cheng L, et al. The research of matching area selection criterion for gravity gradient aided navigation[C]//Proceedings of the Pattern Recognition. Berlin: Springer, 2014: 21-30.

[137] 徐辰龙, 黄玉龙, 王钊, 等. 基于 Haar 小波变换的重力场方向性适配区选取方法[J]. 中国惯性技术学报, 2024, 32(9): 907-917.
Xu C L, Huang Y L, Wang Z, et al. Gravity field directionality matching area selection method based on Haar wavelet transform[J]. Journal of Chinese Inertial Technology, 2024, 32(9): 907-917 (in Chinese).

[138] Wang C L, Wang B, Deng Z H, et al. A co-occurrence matrix-based matching area selection algorithm for underwater gravity-aided inertial navigation[J]. IET Radar, Sonar & Navigation, 2021, 15(3): 250-260.

[139] 马旭东, 闫利, 乔永杰. 重力异常辅助惯性导航中匹配区提取方法[J]. 测绘科学技术学报, 2009, 26(6): 399-402.
Ma X D, Yan L, Qiao Y J. Extraction method of matching area in gravity anomaly aided inertial navigation[J]. Journal of Geomatics Science and Technology, 2009, 26(6): 399-402(in Chinese).

[140] Wang B, Li T J, Deng Z H, et al. Wavelet transform based morphological matching area selection for underwater gravity gradient-aided navigation[J]. IEEE Transactions on Vehicular Technology, 2023, 72(3): 3015-3024.

[141] 吴华玲, 徐效波, 刘波. 重力场可匹配区域选择研究[J]. 测绘科学, 2012, 37(2): 14-16.
Wu H L, Xu X B, Liu B. Research on gravity suitable matching area segmentation[J]. Science of Surveying and Mapping, 2012, 37(2): 14-16(in Chinese).

[142] Wang C L, Wang B, Deng Z H, et al. A delaunay triangulation-based matching area selection algorithm for underwater gravity-aided inertial navigation[J]. IEEE/ASME Transactions on Mechatronics, 2021, 26(2): 908-917.

[143] Zhao X, Zheng W, Xu K K, et al. Optimizing the matching area for underwater gravity matching navigation based on a new gravity field feature parameters selection method[J]. Remote Sensing, 2024, 16(12): 2202.

[144] Wang S Q, Zheng W, Li Z W. Optimizing matching area for underwater gravity-aided inertial navigation based on the convolution slop parameter-support vector machine combined method[J]. Remote Sensing, 2021, 13(19): 3940.

[145] Wang B, Ma Z X, Huang L, et al. A filtered-marine map-based matching method for gravity-aided navigation of underwater vehicles[J]. IEEE/ASME Transactions on Mechatronics, 2022, 27(6): 4507-4517.

[146] Wang B, Cai T J, Fang K. The quantification method of matching capability of areas in gravity-aided inertial navigation[J]. IEEE Sensors Journal, 2022, 22(21): 20958-20967.

[147] Wang B, Zhang Z H, Deng Z H, et al. An analytic quantification method of matching accuracy based on particle filter in gravity-assisted inertial navigation[J]. IEEE Sensors Journal, 2025, 25(10): 18543-18552.

[148] 汪金花, 郭云飞, 郭立稳, 等. 基于 GRPM 井下定位的 MPMD 匹配算法的试验研究[J]. 煤炭学报, 2019, 44(4): 1274-1282.
Wang J H, Guo Y F, Guo L W, et al. Experimental research of MPMD matching algorithm based on the GRPM underground positioning[J]. Journal of China Coal Society, 2019, 44(4): 1274-1282 (in Chinese).

[149] Liu H, Wu L, Bao L F, et al. Gravity matching navigation algorithm based on multiscale search and Hadamard transformed difference[J]. ISA Transactions, 2022, 128: 409-422.

[150] Zhao S W, Xiao X, Wang Y, et al. An improved particle filter based on gravity measurement feature in gravity-aided inertial navigation system[J]. IEEE Sensors Journal, 2023, 23(2): 1423-1435.

[151] Golden J P. Terrain contour matching (TERCOM): a cruise missile guidance aid[C]//Proceedings of the Image Processing for Missile Guidance. Bellingham: SPIE, 1980, 238: 10-18.

[152] Behzad K P, Behrooz K P. Vehicle localization on gravity maps[C]//Proceedings of the SPIE Proceedings. Bellingham: SPIE, 1999, 3693: 182-191.

[153] 闫利, 崔晨风, 吴华玲. 基于 TERCOM 算法的重力匹配[J]. 武汉大学学报 (信息科学版), 2009, 34(3): 261-264.
Yan L, Cui C F, Wu H L. A gravity matching algorithm based on TERCOM[J]. Geomatics and Information Science of Wuhan University, 2009, 34(3): 261-264(in Chinese).

[154] Han Y R, Wang B, Deng Z H, et al. An improved TERCOM-based algorithm for gravity-aided navigation[J]. IEEE Sensors Journal, 2016, 16(8): 2537-2544.

[155] 魏二虎, 董翠军, 刘建栋, 等. 改进 TERCOM 算法用于重力场辅助惯性导航[J]. 测绘地理信息, 2017, 42(6): 29-31.
Wei E H, Dong C J, Liu J D, et al. An improved TERCOM algorithm for gravity-aided inertial navigation system[J]. Journal of Geomatics, 2017, 42(6): 29-31(in Chinese).

[156] Chen M L, Wei G, Zhou W J, et al. Gravity matching navigation algorithm with affine transformation based on particle swarm optimization and validation[J]. Measurement, 2026, 268: 120439.

[157] 童余德, 边少锋, 蒋东方, 等. 实时 ICCP 算法重力匹配仿真[J]. 中国惯性技术学报, 2011, 19(3): 340-343.
Tong Y D, Bian S F, Jiang D F, et al. Gravity matching simulation of real-time ICCP algorithm[J]. Journal of Chinese Inertial Technology, 2011, 19(3): 340-343(in Chinese).

[158] 蒋东方, 童余德, 边少锋, 等. ICCP 重力匹配算法在局部连续背景场中的实现[J]. 武汉大学学报 (信息科学版), 2012, 37(10): 1203-1206.
Jiang D F, Tong Y D, Bian S F, et al. The study on ICCP algorithm for gravity matching based on local continuous field[J]. Geomatics and Information Science of Wuhan University, 2012, 37(10): 1203-1206(in Chinese).

[159] 白文平, 王志刚. ICCP重力匹配辅助导航算法研究及改进[J]. 计算机仿真, 2013, 30(6): 15-19.
Bai W P, Wang Z G. Researching and improving of ICCP algorithm for gravity aided inertial navigation[J]. Computer Simulation, 2013, 30(6): 15-19(in Chinese).

[160] 王向磊, 苏牡丹, 刘培根, 等. 利用改进的ICCP算法辅助导航[J]. 测绘科学, 2013, 38(1): 36-39.
Wang X L, Su M D, Liu P G, et al. Application of improved ICCP algorithm in gravity matching aided navigation[J]. Science of Surveying and Mapping, 2013, 38(1): 36-39(in Chinese).

[161] 蔡龙飞, 郑彤. 基于ICCP算法的重力匹配仿真[J]. 舰船电子工程, 2016, 36(12): 109-112.
Cai L F, Zheng T. Gravity matching simulation of real-time ICCP algorithm[J]. Ship Electronic Engineering, 2016, 36(12): 109-112 (in Chinese).

[162] Liu F M, Li F M, Lin N, et al. Gravity aided positioning based on real-time ICCP with optimized matching sequence length[J]. IEEE Access, 2019, 7: 97440-97456.

[163] di Massa D E, Stewart W K. Terrain-relative navigation for autonomous underwater vehicles[C]//Proceedings of the Oceans' 97 MTS/IEEE Conference Proceedings. Piscataway: IEEE Press, 1997: 541-546.

[164] Doucet A, Freitas N, Gordon N. Sequential Monte Carlo methods in practice[M]. Berlin: Springer, 2001.

[165] Simandl M, Kralovec J, Soderstrom T. Anticipative grid design in point-mass approach to nonlinear state estimation[J]. IEEE Transactions on Automatic Control, 2002, 47(4): 699-702.

[166] 彭富清. 海洋重力辅助导航方法及应用[D]. 郑州: 解放军信息工程大学, 2009.
Peng F Q. Marine gravity aided navigation methods with their applications[D]. Zhengzhou: PLA Information Engineering University, 2009(in Chinese).

[167] Yuan G N, Zhang H W, Yuan K F, et al. Improved SITAN algorithm in the application of aided inertial navigation[C]//Proceedings of the International Conference on Measurement, Information and Control. Piscataway: IEEE Press, 2012: 922-926.

[168] 董翠军. 水下重力场辅助惯性导航匹配算法与适配区选择研究[D]. 武汉: 武汉大学, 2017.
Dong C J. Study on matching algorithm and selection of suitable matching area for underwater gravity-aided inertial navigation system[D]. Wuhan: Wuhan University, 2017(in Chinese).

[169] Deng Z H, Li C, Yin L J, et al. An adaptive robust unscented Kalman filter based matching algorithm for underwater gravity aided navigation[C]//Proceedings of the IEEE CSAA Guidance, Navigation and Control Conference. Piscataway: IEEE Press, 2018: 1-8.

[170] 欧阳明达, 孙艺轩, 邝英才, 等. 应用抗差估计SITAN算法的水下重力匹配导航方法[J]. 中国惯性技术学报, 2021, 29(2): 214-220.
Ouyang M D, Sun Y X, Kuang Y C, et al. Underwater gravity matching navigation method of SITAN algorithm with robust estimation[J]. Journal of Chinese Inertial Technology, 2021, 29(2): 214-220(in Chinese).

[171] 王傲明, 李姗姗, 李新星, 等. 基于自适应并行扩展卡尔曼滤波的SITAN匹配算法[J]. 中国惯性技术学报, 2022, 30(1): 81-88.
Wang A M, Li S S, Li X X, et al. SITAN matching algorithm based on adaptive parallel extended Kalman filter[J]. Journal of Chinese Inertial Technology, 2022, 30(1): 81-88(in Chinese).

[172] 付林威, 赵东明, 付林. 基于抗差自适应SITAN算法的水下重力匹配导航方法[J]. 大地测量与地球动力学, 2023, 43(8): 820-825.
Fu L W, Zhao D M, Fu L. Underwater gravity matching navigation method based on robust adaptive SITAN algorithm[J]. Journal of Geodesy and Geodynamics, 2023, 43(8): 820-825(in Chinese).

[173] 付林威, 赵东明, 范雕, 等. 基于抗差自适应CKF的水下重力匹配导航SITAN算法[J]. 海洋测绘, 2024, 44(2): 5-8.
Fu L W, Zhao D M, Fan D, et al. SITAN algorithm for underwater gravity matching navigation based on robust adaptive CKF[J]. Hydrographic Surveying and Charting, 2024, 44(2): 5-8(in Chinese).

[174] 毛宁, 李安, 许江宁, 等. 惯性/重力组合导航可观测性分析及鲁棒滤波方法[J]. 武汉大学学报 (信息科学版), 2024, 49(11): 2113-2121.
Mao N, Li A, Xu J N, et al. Observability analysis and robust fusion algorithms of INS/gravity integrated navigation[J]. Geomatics and Information Science of Wuhan University, 2024, 49(11): 2113-2121(in Chinese).

[175] 刘繁明, 唐英丽. 差分进化粒子滤波在惯性/重力组合导航中的应用研究[J]. 应用科技, 2015, 42(4): 15-19.
Liu F M, Tang Y L. Application of the particle filter in INS/gravity integrated navigation based on differential evolution[J]. Applied Science and Technology, 2015, 42(4): 15-19(in Chinese).

[176] Wang B, Yu L, Deng Z H, et al. A particle filter-based matching algorithm with gravity sample vector for underwater gravity aided navigation[J]. IEEE/ASME Transactions on Mechatronics, 2016, 21(3): 1399-1408.

[177] Wang B, Zhu J W, Ma Z X, et al. Improved particle filter-based matching method with gravity sample vector for underwater gravity-aided navigation[J]. IEEE Transactions on Industrial Electronics, 2021, 68(6): 5206-5216.

[178] 鲁智谦. 重力辅助惯性导航定位算法研究[D]. 南京: 东南大学, 2024.
Lu Z Q. Research on gravity-assisted inertial navigation and positioning algorithm[D]. Nanjing: Southeast University, 2024(in Chinese).

[179] Han Y R, Wang B, Deng Z H, et al. Point mass filter based matching algorithm in gravity aided underwater navigation[J]. Journal of Systems Engineering and Electronics, 2018, 29(1): 152-159.

[180] Han Y R, Wang B, Deng Z H, et al. A matching algorithm based on the nonlinear filter and similarity transformation for gravity-aided underwater navigation[J]. IEEE/ASME Transactions on Mechatronics, 2018, 23(2): 646-654.

[181] 蔡体菁, 鲁智谦, 高帅鹏. 基于模拟退火重采样的点质量滤波重力匹配方法[J]. 光学与光电技术, 2024, 22(3): 1-7.
Cai T J, Lu Z Q, Gao S P. Point mass filtering gravity matching method based on simulated annealing resampling[J]. Optics & Optoelectronic Technology, 2024, 22(3): 1-7(in Chinese).

[182] Wang B, Cai T J, Fang K. Observation-differenced point mass filter in gravity-aided inertial navigation[J]. IEEE/ASME Transactions on Mechatronics, 2024, 29(2): 1601-1606.

[183] Liu M Q, Wang B, Deng Z H, et al. Improved ICCP algorithm and its application in gravity matching aided inertial navigation system[C]//Proceedings of the 33rd Chinese Control Conference. Piscataway: IEEE Press, 2014: 562-567.

[184] Wei E H, Dong C J, Yang Y L, et al. A robust solution of integrated SITAN with TERCOM algorithm: weight-reducing iteration technique for underwater vehicles' gravity-aided inertial navigation system: a robust solution of integrated SITAN with TERCOM[J]. Navigation, 2017, 64(1): 111-122.

[185] Han Y R, Wang B, Deng Z H, et al. A combined matching algorithm for underwater gravity-aided navigation[J]. IEEE/ASME Transactions on Mechatronics, 2018, 23(1): 233-241.

[186] 邹嘉盛, 肖云, 孙爱斌, 等. 利用 TERCOM 与 ICCP 进行联合重力匹配导航[J]. 导航定位与授时, 2021, 8(1): 115-124.
Zou J S, Xiao Y, Sun A B, et al. Gravity matching navigation technology based on integration of TERCOM and ICCP[J]. Navigation Positioning and Timing, 2021, 8(1): 115-124(in Chinese).

[187] Gao S P, Cai T J. Parallel multiple methods with adaptative decision making for gravity-aided navigation[J]. Journal of Marine Science and Engineering, 2023, 11(8): 1624.

[188] 张锦柏, 肖云, 陈垲宁, 等. 一种粗精匹配结合的重力辅助惯性导航算法[J]. 地球物理学报, 2025, 68(3): 788-801.
Zhang J B, Xiao Y, Chen K N, et al. A gravity aided inertial navigation algorithm based on the rough/fine matching combination[J]. Chinese Journal of Geophysics, 2025, 68(3): 788-801(in Chinese).

[189] 程力, 蔡体菁. 一种模式识别神经网络重力匹配算法[J]. 中国惯性技术学报, 2007, 15(4): 418-422.
Cheng L, Cai T J. Gravity matching algorithm using pattern recognition neural network[J]. Journal of Chinese Inertial Technology, 2007, 15(4): 418-422(in Chinese).

[190] 程力, 蔡体菁. 基于支持向量机的重力匹配算法[J]. 系统仿真学报, 2008, 20(21): 5953-5956.
Cheng L, Cai T J. Gravity matching algorithm based on support vector machine[J]. Journal of System Simulation, 2008, 20(21): 5953-5956(in Chinese).

[191] 黄鹏, 成怡. 改进的 BP 神经网络算法在航迹匹配中的应用[J]. 计算机工程, 2011, 37(11): 218-219.
Huang P, Cheng Y. Application of improved back propagation neural network algorithm in track matching[J]. Computer Engineering, 2011, 37(11): 218-219(in Chinese).

[192] 高伟, 赵博, 周广涛. 基于带约束人工蜂群算法和平均 Hausdorff 距离的重力匹配方法[J]. 传感技术学报, 2014, 27(1): 74-78.
Gao W, Zhao B, Zhou G T. Gravity matching method based on artificial bee colony algorithm with restriction and MHD[J]. Chinese Journal of Sensors and Actuators, 2014, 27(1): 74-78(in Chinese).

[193] Dai T, Miao L J, Shao H J, et al. Solving gravity anomaly matching problem under large initial errors in gravity aided navigation by using an affine transformation based artificial bee colony algorithm[J]. Frontiers in Neurorobotics, 2019, 13: 19.

[194] 邹嘉盛, 肖云, 孟宁, 等. 一种基于约束条件的重力匹配导航算法[J]. 武汉大学学报 (信息科学版), 2020, 45(10): 1570-1577.
Zou J S, Xiao Y, Meng N, et al. A gravity matching navigation algorithm based on constraint[J]. Geomatics and Information Science of Wuhan University, 2020, 45(10): 1570-1577(in Chinese).

[195] 欧阳明达, 杨元喜. 改进粒子滤波的水下重力匹配导航仿真与实验[J]. 哈尔滨工程大学学报, 2022, 43(10): 1514-1521.
Ouyang M D, Yang Y X. Underwater gravity matching navigation simulation and experiment based on the improved particle filtering algorithm[J]. Journal of Harbin Engineering University, 2022, 43(10): 1514-1521(in Chinese).

[196] Gao S P, Cai T J, Fang K. Gravity-matching algorithm based on K-nearest neighbor[J]. Sensors, 2022, 22(12): 4454.

[197] 陆泽橡, 蔡体菁. 基于 A*算法的重力辅助导航航迹规划[J]. 中国惯性技术学报, 2010, 18(5): 556-560.
Lu Z Y, Cai T J. Gravity aided navigation path plan research based on A* algorithm[J]. Journal of Chinese Inertial Technology, 2010, 18(5): 556-560(in Chinese).

[198] Feng X Q, Wang B, Deng Z H, et al. Internal path planning method for underwater vehicle with gravity aided navigation[C]//Proceedings of the 35th Chinese Control Conference. Piscataway: IEEE Press, 2016: 5537-5541.

[199] Shi L, Zhang W Z, Cheng Y, et al. Gravity aided inertial navigation path planning algorithm based on underwater vehicle constraint and bezier curve[C]//Proceedings of the Chinese Control Conference. Piscataway: IEEE Press, 2019: 3857-3862.

[200] 欧阳明达, 马越原. 基于改进 A*算法的水下重力匹配导航路径规划[J]. 地球物理学报, 2020, 63(12): 4361-4368.
Ouyang M D, Ma Y Y. Path planning for gravity aided navigation based on improved A* algorithm[J]. Chinese Journal of Geophysics, 2020, 63(12): 4361-4368(in Chinese).

[201] 张锦柏, 肖云, 陈垲宁. 多属性决策与改进 A*算法规划重力匹配导航航迹[J]. 测绘科学, 2023, 48(9): 38-48.
Zhang J B, Xiao Y, Chen K N. Multi-attribute decision theory and improved A* algorithm for planning gravity matching navigation tracks[J]. Science of Surveying and Mapping, 2023, 48(9): 38-48(in Chinese).

[202] Zhao S W, Shi L, Zhang W Z, et al. Global dynamic path-planning algorithm in gravity-aided inertial navigation system[J]. IET Signal Processing, 2021, 15(8): 510-520.

[203] 张驰, 李姗姗, 史颜俊, 等. 蚁群-势场算法在水下重力辅助导航航迹规划中的应用[J]. 测绘学报, 2020, 49(7): 865-873.
Zhang C, Li S S, Shi Y J, et al. Application of ant colony-potential field algorithm in underwater gravity matching navigation track planning[J]. Acta Geodaetica et Cartographica Sinica, 2020, 49(7): 865-873(in Chinese).

[204] 孙玮萱, 肖云, 张锦柏, 等. 一种重力匹配导航路径规划新算法: 粒子群优化灰狼算法[J]. 大地测量与地球动力学, 2025, 45(9): 922-928.
Sun W X, Xiao Y, Zhang J B, et al. A new algorithm for gravity matching navigation path planning: particle swarm optimization and grey wolf optimization[J]. Journal of Geodesy and Geodynamics, 2025, 45(9): 922-928(in Chinese).

[205] Wang M C, Zheng W, Zhao S J, et al. Voronoi node stepwise following particle swarm optimization for path planning of underwater gravity aided navigation[J]. Ocean Engineering, 2025, 336: 121767.

[206] Wang B, Cai T J. Path planning method for underwater gravity-aided inertial navigation based on PCRB[J]. Journal of Marine Science and Engineering, 2023, 11(5): 993.

[207] Zhang J B, Zhou H, Xiao Y, et al. Adaptive elite ant colony optimization for track planning in gravity-aided navigation with multi-feature fusion[J]. Defence Technology, 2026, 59: 112-126.

Overview and prospect of key technologies for underwater gravity-aided inertial navigation

ZHOU Zhaofa[1], LI Yiyi[1], ZHANG Zhili[1,*], CHEN Po[2], BIAN Yinglei[2], LI Xinyu[1]

(1. School of Intelligence and Control, Rocket Force University of Engineering, Xi'an 710025, China;
2. The 96416th Troop of the Chinese People's Liberation Army, Luoyang 471000, China)

Abstract: Underwater high-precision autonomous navigation is a core enabling technology for ocean exploration and development. The positional error of an inertial navigation system (INS) builds up and diverges with time, requiring outside data to be corrected. As an inherent property of the Earth, the gravity field offers stability, autonomy, and passivity, making it an ideal information source. Consequently, gravity-aided inertial navigation technology holds significant research value. This paper systematically reviews the research progress of four important technologies: trajectory planning, gravity matching algorithms, gravity reference map construction, and gravity field adaptability analysis. It is based on an analysis of the system's technical architecture and fundamental principles. Mainstream methods, such as multi-source data fusion, gridded interpolation, feature evaluation and matching area extraction, sequential correlation matching, nonlinear filtering, and intelligent path optimization, are comprehensively summarized. Furthermore, common unresolved issues in current technological development are thoroughly examined. Finally, in view of the technical bottlenecks at this stage, future research directions are discussed, aiming to provide theoretical references and technical guidance for the subsequent development of a high-precision, high-reliability, and high-autonomy underwater gravity-aided inertial navigation system (GAINS).

Keywords: underwater autonomous navigation; gravity-aided inertial navigation system; gravity reference map; adaptability analysis; gravity matching algorithm; trajectory planning

Received: 2026-05-01; **Accepted**: 2026-07-17; **Published Online**: 2026-07-23 10:11
URL: link.cnki.net/urlid/11.2625.V.20260723.0944.001

* **Corresponding author**. E-mail: zzl202@hhu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2026.0440

氢燃料电池无人机研究进展与关键技术

徐伟强[1,2,*]，张炳鑫[2]，祝赫[1]，陈黎[3]

(1. 天目山实验室，杭州 311115；2. 北京航空航天大学 航空科学与工程学院，北京 100191；
3. 西安交通大学 能源与动力工程学院，西安 710049)

摘 要：在低空经济崛起与绿色航空发展的背景下，氢燃料电池混合动力无人机凭借高能量密度和零排放等优势，成为突破传统锂电无人机续航瓶颈的关键路径。基于此，对氢动力无人机的发展及其涉及的关键技术进行总结和梳理，并对氢燃料电池混合动力无人机未来的发展进行合理展望，旨在为相关研究提供参考。对氢动力航空器的发展历程和技术演替、氢燃料电池动力系统的优势以及氢燃料电池无人机的发展现状进行概述；对氢燃料电池混合动力无人机涉及的氢燃料电池技术、多源混合能源系统技术、机载储氢技术、能量管理策略以及热管理策略进行系统梳理；结合氢能无人机各关键技术的研究现状及其实际应用中所面临的问题，对未来的发展方向进行合理展望。

关 键 词：无人机；氢燃料电池；混合动力；机载储氢；能量管理；热管理
中图分类号：V279
文献标志码：A **文章编号**：1001-5965(2026)10-3563-17

随着我国“十四五”规划的圆满收官及“十五五”规划的全面开启，低空经济已经成为我国发展的重点关注领域。《中华人民共和国国民经济和社会发展第十五个五年规划纲要》明确将“推进低空经济健康有序发展”“加快大载重固定翼无人机、长航时垂直起降航空器等新型低空装备研制”列为重要任务。这表明，以固定翼无人机、旋翼无人机等为代表的飞行平台，将成为低空经济的核心载体。这些无人机平台凭借其高机动性、优秀的任务适应性以及多元化的起降能力，在物流配送、应急救援、精准农业、线路巡检、航拍乃至军事等领域都有着广泛的应用[1-8]。然而，日益复杂的任务场景对无人机的作业半径、续航时间、载荷能力以及全天候工作能力提出了越来越高的要求。传统以化石燃料为主的动力系统虽然能提供较长的航时，但其高噪声、高排放特性给环境带来的负面影响也难以忽视[9]。为响应我国“双碳”目标及绿色航空装备发展愿景，转向电能驱动是最佳选择。

目前，主流的电能动力系统往往以锂电池为能量载体，尽管其具有零排放、低噪声以及高动态响应等优势，但受限于较低的能量密度，严重制约了无人机的续航时间与任务效能，无法满足当前作业环境对无人机提出的要求，上述不足在功耗敏感的电动旋翼平台上表现得尤为显著。因此，寻找一种兼具高能量密度与环境友好特性的新型电能载体，已成为突破无人机续航瓶颈的关键。在此背景下，氢能凭借其高能量密度与零排放特质，成为无人机动力系统转型的理想选择。

氢能作为一种零碳排放且来源广泛的清洁能源，契合世界各国的减碳、零碳政策，是全球能源转型的重要方向，将成为人类用来代替化石燃料的最佳选择之一。根据国际能源署(International Energy

收稿日期：2026-06-29；**录用日期**：2026-08-21；**网络出版时间**：2026-08-28 16:38
网络出版地址：link.cnki.net/urlid/11.2625.V.20260828.1516.003
基金项目：天目山实验室关键领域科研项目 (TK-2024-D-011, TK-2025-D-003)
* **通信作者**. E-mail：09430@buaa.edu.cn

引用格式：徐伟强，张炳鑫，祝赫，等. 氢燃料电池无人机研究进展与关键技术 [J]. 北京航空航天大学学报，2026，52（10）：3563-3579.
Xu W Q，Zhang B X，Zhu H，et al. Research progress and key technologies of hydrogen fuel cell UAVs [J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（10）：3563-3579（in Chinese）.

Agency, IEA)的数据,预计到2030年,世界氢能源生产能力将达到140～190 GW[10],这将为发展氢动力无人机提供充足的能源供应。最重要的是,氢的质量能量密度高达120 MJ/kg,使用氢能源设计的燃料电池混合动力系统能量密度为250～540 Wh/kg[11],与锂电池(130～200 Wh/kg)相比,氢能源的引入能够提供至少2倍于相同质量的锂电池所能提供的续航时间[12]。这一特性为破解电动无人机续航难题提供了极具潜力的解决方案。另外,氢燃料电池在具备高能量密度的同时实现了真正的零排放,其反应过程中唯一的产物是水,对环境完全无害。

同时,近年来我国对发展氢能航空器的政策支持也在不断增加,在国家大力发展新质生产力、推进“双碳”目标与低空经济战略的时代背景下,氢能与无人机产业的深度融合正成为科技创新的重要前沿。2020年6月发布的《无人机用氢燃料电池发电系统》(GB/T 38954—2020),统一了无人机氢燃料电池相关的要求[13];2022年3月发布的《氢能产业发展中长期规划(2021—2035年)》,要求积极探索燃料电池在船舶、航空器等领域的应用,推动大型氢能航空器研发[14];2023年10月发布的《绿色航空制造业发展纲要(2023—2035年)》,明确布局氢能航空关键技术研发及商业化运营模式探索[15];2026年3月,工业和信息化部、财政部、国家发展改革委联合印发《工业和信息化部 财政部 国家发展改革委关于开展氢能综合应用试点工作的通知》,明确将“航空器”列为氢能六大创新应用场景之一,单个试点城市群4年最高可获16亿元奖励资金,氢能无人机单架最高可获8万元专项补贴[16]。这标志着氢能无人机已从技术验证阶段正式迈入国家战略支持的规模化应用新阶段。这些都为氢燃料电池无人机的技术研发与应用提供了坚实的政策支持。

在此背景下,本文立足于“十五五”期间对绿色、长航时航空装备的研制需求以及氢能航空创新应用的政策导向,旨在对氢燃料电池无人机的发展现状及其涉及的氢燃料电池动力系统技术、机载储氢技术、能量管理与热管理策略等进行系统性的梳理与总结,以期为相关研究提供有效参考。

1 氢能源无人机发展现状

1.1 氢动力航空器的发展历程与技术演进

随着我国“双碳”目标的推进以及低空经济的兴起,氢动力航空技术经历了从大型固定翼验证机到各型无人机、从直接燃烧供能到电化学供能转化的深刻演进。

人类对氢动力飞行器的探索始于对高能量密度燃料的渴求。最早的将氢能源应用于航空领域的方案可以追溯到20世纪前期,德国Heinkel公司将气态氢用于其设计的涡轮喷气发动机作为燃料进行了相关测试[17]。虽然此次探索中氢能源并没有推动航空器飞向天空,却为后来的探索打开了思路。1988年,苏联成功试飞了全球首架以液氢为燃料的验证机Tu-155,如图1所示。虽然其采用的是氢内燃机动力系统,但它首次在工程层面证明了机载液氢储存与输送的可行性[18]。



图1 Tu-155验证机[19]

Fig. 1 Tu-155 demonstrator[19]

进入21世纪,质子交换膜燃料电池(proton exchange membrane fuel cell, PEMFC)技术的成熟为航空动力系统的小型化、静音化提供了新路径。2008年,波音的燃料电池演示机采用氢燃料电池作为飞机的动力系统完成了巡航飞行[18],标志着氢能动力正式进入以燃料电池为主的新阶段。到这一阶段,对于氢动力航空器的研究仍然集中于固定翼平台,科研人员进行了大量的尝试来验证氢燃料电池作为航空器动力系统的可行性。

近年来,随着无人机产业与低空经济的爆发,氢燃料电池技术自然地向无人机领域迁移,并衍生出2条清晰的应用脉络:①在固定翼无人机上,科研人员依托其高升阻比的天然优势,利用高能量密度的氢燃料电池系统实现更长时间的持续监测、遥感、侦察等任务,极大地拓展了无人机的任务半径;②针对城市空中交通以及物流配送等场景,氢动力系统正成为多旋翼及电动垂直起降飞行器(electric vertical take-off and landing, eVTOL)破解锂电“续航瓶颈”的关键。面对垂直起降与悬停带来的高功耗需求,氢燃料电池凭借高能量密度与快速补能优势,显著提升了这类飞行平台的作业效能。目前,氢燃料电池系统已在固定翼长航时无人机、多旋翼无人机及新型eVTOL等多类飞行平台上展开了广泛的应用与探索。

1.2 氢燃料电池动力系统的优势

在工业级无人机领域,作业效能直接受限于动力系统的性能。传统的化石燃料动力虽然能量密度高,但其高噪声、高振动以及碳排放问题,既不符

合全球绿色环保的发展趋势, 也限制了无人机在如自然保护区、城市静音区等环境敏感区域的作业能力。相比之下, 电动无人机具有低噪声、低振动和运行过程零排放等特点, 同时能够简化机械结构并降低维护成本, 已成为当前无人机动力系统的重要发展方向。作为 2 类典型的电动动力方案, 锂电池与氢燃料电池在应用中具有不同的性能特点, 下面从 5 个方面进行对比分析:

1) 氢燃料电池动力系统的能量密度可达到传统锂电池的 2 倍以上, 可显著延长无人机续航时间, 扩大任务半径, 使其更好地应对现代复杂的任务环境。

2) 氢燃料电池在产生电能的同时也伴随热量释放, 这有利于维持动力系统工作温度, 减缓低温环境对动力系统的不利影响, 使得氢燃料电池在低温环境下也具有较好的工作适应性。

3) 相较于循环寿命有限、容量衰减显著的传统高能量密度锂电池, 氢燃料电池系统在耐久性与长期运行稳定性方面具备显著优势。随着累计飞行时间增加, 锂电池循环衰减及更换所带来的折旧成本逐步累积, 而氢燃料电池较长的使用寿命有利于降低单位飞行时间的成本, 使其长期综合运营成本低于纯电系统。

4) 氢燃料电池动力系统可以在数分钟内实现氢燃料的快速充装, 相较于锂电池充电能够显著缩短补能时间, 减少连续作业过程中的停机时间, 从而提高设备利用率和作业效率。

5) 氢燃料电池动力系统保留了传统锂电池的清洁、低噪声、无碳排放的优点, 保证了无人机的可作业范围, 同时其工作过程无机械振动的特性也降低了很多精密机载设备搭载使用的难度。

综上, 氢燃料电池无人机在续航能力、低温适应性、耐久性、快速补能以及长期运行稳定性与经济性等方面都具有明显优势。

1.3　氢燃料电池无人机国内外发展现状与应用

目前, 主流的氢动力无人机从构型上主要分为固定翼、多旋翼、复合翼无人机 3 类。其中, 复合翼又可以视为固定翼的一个特殊子类。

固定翼无人机依靠机翼产生的气动升力维持飞行, 其长航时特性与氢燃料电池的高能量密度特性具有极高的物理契合度。由于其巡航阶段仅需维持较低功率运行即可维持升力, 该类构型能将氢能的化学能高效转化为推进功, 从而实现超长时续航与更大的作业半径。凭借在广域覆盖与超长航时方面的绝对优势, 固定翼无人机非常契合对载荷与续航能力要求较高且对起降场地限制相对宽松的任务场景, 如大纵深区域的边境巡防、大面积海域的海事监察与搜救以及长距离油气管道与特高压电网的巡检等。

多旋翼无人机常通过多个旋翼产生的向下推力实现垂直起降与悬停, 具备极高的机动灵活性与定点悬停精度。引入氢动力系统后, 有效缓解了传统锂电多旋翼“载重小、续航短”的痛点, 将作业时间由数十分钟提升至数小时量级。凭借垂直起降以及操作灵活的特性, 多旋翼无人机非常适用于城市狭窄街巷、山区复杂地形等狭小的作业空间, 具体应用如低空拍摄、城市物流配送、城市应急救援等。

复合翼融合了多旋翼的垂直起降能力与固定翼的高效巡航能力, 通常采用“多旋翼+固定翼”的复合气动布局, 这使其可满足复杂地形起降需求和长距离高效巡航, 拓宽了使用场景, 但是技术复杂度高、成本相对昂贵、维护难度较大。复合翼构型能够适用于起降条件受限但作业半径较大的场景, 如海岛间的物流运输与物资投送(无需平整跑道)、山区与丛林的地理测绘与资源勘探及大面积林区的防火巡查等。

下面本文按照构型作为分类标准, 简要介绍氢动力无人机的代表性机型。

1.3.1　固定翼/复合翼无人机

固定翼无人机凭借其长航时、大载荷的优势, 是氢能源无人机发展的重点平台, 在军民 2 个领域中都发挥着相当重要的作用[20]。多年来, 国内外研究人员在氢动力固定翼无人机方向进行了大量的探索和实践。

美国的 Ion Tiger 无人机搭载氢燃料电池动力实现了长时间续航, 2009 年, 其首次试飞, 在携带约 2.27 kg 有效载荷的情况下连续飞行 26 h, 创下了燃料电池动力飞机的非官方续航纪录[21-22]。2012 年, 同济大学与上海奥科赛飞机有限公司共同研发了“飞跃一号”氢动力固定翼无人机, 该飞机翼展 5 m, 起飞重量 20 kg, 有效载荷 1 kg, 巡航速度 30 km/h, 续航时间 2 h, 使用一个 1 kW 的 PEMFC 作为动力[23]。2014 年, 武汉众宇动力系统科技有限公司研发出“天行者”燃料电池无人机, 首飞实现了 12 h 的不间断飞行[24]。2025 年 4 月 20 日, 中国民用航空飞行学院多个二级学院团队联合研发的 FC-30 型氢燃料电池全工况增程式复合翼无人机, 在成都国家级民用无人驾驶航空试验基地成功完成了首次全流程飞行测试, 该机型的最长续航时间可达 4 h, 具备多次起降、长时间悬停及多次悬停的功能, 成功实现了固定翼模态长航程与多旋翼模态长

航时的兼顾[25]。哈尔滨工业大学重庆研究院研发的氢燃料电池复合翼无人机“青鸥-30B”搭载35 MPa高压储氢罐，续航时间长达12 h，最远可飞行900 km，且凭借其自主研发的氢燃料电池系统，获得了高达60%的能量转化效率。目前，“青鸥-30B”已广泛应用于电网巡检、森林巡检、边防巡检等场景，并成功在海拔3 800 m以上的高原完成严苛测试[26]。部分典型机型如图2所示。

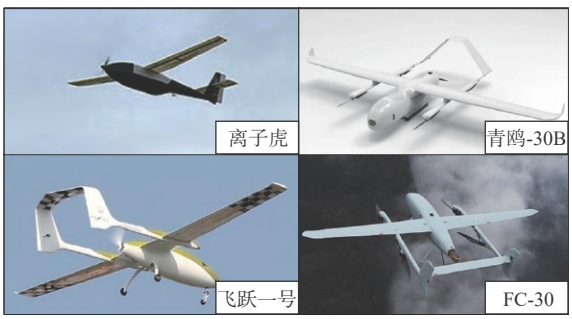


图2 固定翼/复合翼氢动力无人机代表机型[23,25,27]

Fig. 2 Representative hydrogen-powered fixed-wing/hybrid-wing UAV models[23,25,27]

1.3.2 多旋翼无人机

近年来，旋翼无人机在生活中的应用愈来愈广泛，尤其是多旋翼无人机，其凭借起降方便、灵活机动等优势，深受各个领域的青睐。国内外都开展了大量研究，并取得了许多成果，下面介绍其中的一些氢动力旋翼机成果。

我国在氢动力旋翼机领域起步较晚但发展势头迅猛。2023年，由浙江氢航科技有限公司设计的氢旋4号无人机顺利完成了−22 ℃下的低温巡检工作[28]。2024年8月，北京航空航天大学天目山实验室研发的“天目山一号”百公里级氢动力长航程多旋翼无人机首飞成功[23]。随着关键技术不断突破，“天目山一号”于2025年11月16日在北京航空航天大学杭州国际校园上空成功完成了不间断飞行188.605 km，总续航4 h17 min55 s的飞行挑战，创造了全新的世界纪录，并得到了吉尼斯世界纪录的官方认证[29]。同时，“天目山一号”还在−40 ℃的根河市进行了极寒环境下的飞行测试，证明了其良好的环境适应能力[30]。由陕西同尘和光低温科技有限公司和哈尔滨工业大学重庆研究院共同设计的“青鹞-10”氢动力多旋翼无人机在2025年完成了液氢能源的首次试飞，实现了超过3.5 h的连续飞行[31]。2026年，由氢鹏科技(浙江)有限公司设计的XC02与XC04这2款氢动力无人机在我国黑龙江省漠河市近−40 ℃的极寒环境中，圆满完成飞行测试，均实现了超150 min的连续稳定飞行[32]。国内氢动力旋翼无人机代表机型如图3所示。

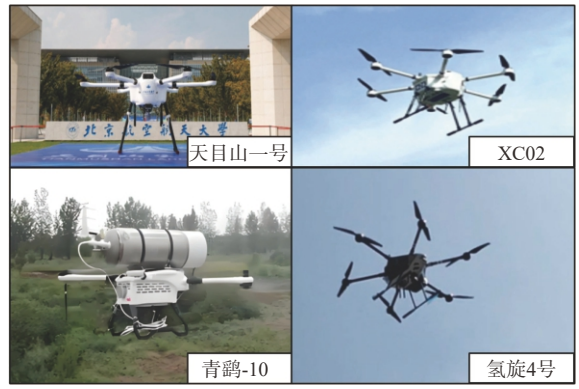


图3 氢动力旋翼无人机国内代表机型[28,30-32]

Fig. 3 Representative domestic hydrogen-powered rotary-wing UAV models[28,30-32]

国际上，各国对旋翼无人机技术的发展都十分重视。美国Flight-Wave Aerospace System公司研制的Jupiter-H2四旋翼无人机携带3 L压缩气态氢作为燃料，且在飞行过程中完全由燃料电池供电[23]。另外，美国Alaka'i Technologies公司推出了一款名为Skai的氢动力eVTOL，其采用六转子分布式动力设计，旨在解决未来城市空中的大载荷载人通勤任务[33]。韩国在氢动力无人机领域发展也十分迅速。韩国斗山集团旗下的斗山创新利用自研的氢燃料电池电源模块，设计出多款成熟的旋翼无人机平台，其中DT30N在负载后有效续航仍可达到2 h以上，且中间无需更换电池，大大提高了作业效率，尤其是需要低速精密测绘的领域[34]。

2 氢燃料电池动力系统与储氢技术

2.1 氢燃料电池混合动力系统架构

2.1.1 氢燃料电池技术

燃料电池常基于电化学原理将燃料中的化学能直接转化为电能，而不需要通过燃烧来获取能量，这使得此类装置在运行时噪声、振动都很微弱，且具有更高的效率。常用的燃料电池根据电解质种类不同可以分为以下几种：PEMFC、碱性燃料电池(alkaline fuel cell, AFC)、磷酸燃料电池(phosphoric acid fuel cell, PAFC)、熔融碳酸盐燃料电池(molten carbonate fuel cell, MCFC)、固体氧化物燃料电池(solid oxide fuel cell, SOFC)等。其中，相较于其他燃料电池而言，PEMFC具有质量轻、功率密度高、工作温度低、启动速度快等优点，是目前氢动力无人机中应用最为广泛的燃料电池。经典的PEMFC结构以及工作原理如图4所示，其由阴极和阳极两侧的流道、气体扩散层、催化层以及质子交换膜组成[35]。

在工作时，PEMFC中氢气经过阳极流道扩散

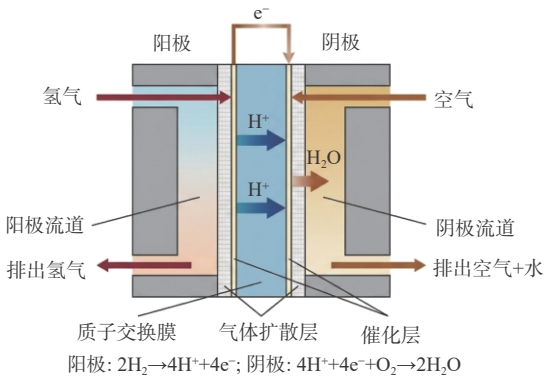


图 4　PEMFC 工作原理

Fig. 4　PEMFC working principle

至气体扩散层, 最后进入催化层, 在催化剂的作用下失去电子, 并将质子通过质子交换膜传递至阴极侧。氧气经过阴极流道扩散至气体扩散层, 最后进入催化层, 在催化剂的影响下, 与通过质子交换膜的质子以及来自外部电路的电子结合生成水, 从而在阴极和阳极之间形成持续的电流以驱动工作设备。

通常情况下, PEMFC 的阴极和阳极流道被设计在同一块导电板的正反两面, 故而常称该导电板为双极板[36]。另外, 气体扩散层、催化层和质子交换膜三者构成了电池的膜电极组件(membrane electrode assembly, MEA)[37]。在实际工程应用中, 通过“双极板—膜电极—双极板”的交替堆叠, 可将多个单体电池串联形成电池堆, 从而显著提升其整体输出电压与功率密度。然而, 这也使得传统的石墨双极板在电池系统总重中占到了相当大的比例。

为了更加契合航空领域轻量化与高功重比的要求, 双极板材料和结构的优化已成为氢动力无人机燃料电池的重要研究方向之一。近年来, 金属双极板凭借其在电堆轻量化方面的优势受到广泛关注。研究表明, 同等条件下, 与石墨双极板 PEMFC 电池堆相比, 金属双极板堆输出电压高 4%, 质量减轻 31%[38]。但是, 金属双极板在运行中也面临着电化学腐蚀、金属离子析出等问题, 为此高抗腐蚀、低接触电阻的表面涂层工艺成为金属双极板工程应用的关键技术瓶颈。目前, 前沿的这一类型产品也有尝试使用钛基双极板, 其比传统的不锈钢极板更轻。另外, 由树脂基体掺杂石墨粉、碳纳米管等导电填料构成的复合材料双极板凭借轻量化、优异的抗震性与高韧性, 也成为重要的发展方向, 但其目前仍存在导电性和厚度之间的博弈, 且内阻相较于金属双极板偏高, 使得系统功率密度受限。

除了双极板材料与结构的优化, 氢燃料电池热管理也是氢动力无人机亟需突破的核心技术之一, 其直接影响 PEMFC 的性能与寿命。由于热力学与电化学不可逆损失的限制, PEMFC 在将氢能转化为电能的过程中不可避免地会产生大量热量, (PEMFC 的典型能量流图如图 5 所示[39]), 如果不能及时将这部分能量排出, 电池温度将持续升高, 加速膜内水分蒸发, 影响膜电极水合, 导致膜干, 从而使膜电阻率急剧上升、欧姆损耗增大, 进而引发电池输出电压下降、功率衰减; 严重时还会造成膜局部热点、裂纹甚至穿孔, 导致反应气体交叉渗透, 造成不可逆的性能恶化与寿命缩短[40]。为此, 冷却系统对于燃料电池而言就显得尤为重要。本节从氢燃料电池技术的角度, 对热管理进行简要描述, 详细内容在热管理章节中详细展开。

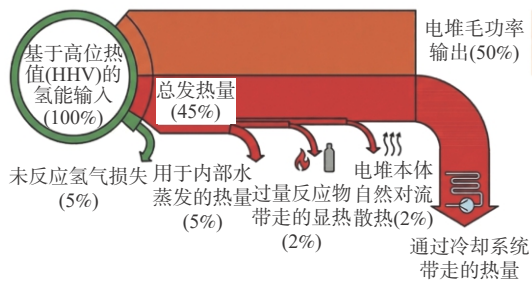


图 5　PEMFC 的典型能量流图[39]

Fig. 5　Typical energy flow diagram of PEMFC[39]

总体上, PEMFC 可分为液冷型和空冷型[41] 2 种。其中, 与液冷型 PEMFC 相比, 空冷型 PEMFC 通过空气吹扫为电堆降温, 无需复杂的辅助系统, 从而使整个系统更加简洁、轻便。根据结构的不同, 空冷型 PEMFC 又可进一步分为 3 类: 面冷型、边缘冷却型[42-43] 和开放阴极型[44]。其中, 开放阴极型直接将阴极流道暴露在大气中, 使得风扇在抽吸空气为电堆降温的同时, 也为电池提供了阴极反应所需的氧气。开放阴极 PEMFC 结构简单、功率密度高、噪声低, 是目前 3 种空冷型 PEMFC 中应用最广泛、最成熟的类型。典型的开放阴极空冷燃料电池动力系统如图 6 所示[45]。

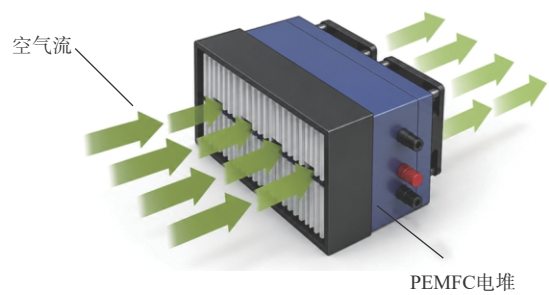


图 6　开放阴极质子交换膜燃料电池[45]

Fig. 6　Open-cathode PEMFC[45]

然而, 也正是因为直接将阴极流道暴露在外界环境中, 开放阴极型氢燃料电池的环境适应能力受到限制。而对于工业级、军工级、全天候作业的氢

动力无人机而言，常常要面临复杂多变的任务环境，如空气湿度的变化、海岛的盐雾、沙尘天气的固体微粒及工业区大气中的 SO_x 和 NO_x 等，这些都会对电池的性能造成损伤，最终导致电池寿命降低[46-47]。因此，如何提升燃料电池对复杂恶劣环境的适应能力成为亟需解决的难题。近年来，阴极闭合式电堆通过主动供气和可控热管理，为提升系统的环境适应性与寿命提供了新的技术路径。近期，由中国科学院大连化学物理研究所研发的阴极闭合式风冷电堆比功率达 1 970 W/kg，面功率密度达 1.15 W/cm^2，具备高功率密度、强环境适应性、长寿命等优势，经鉴定总体技术达到国际领先水平[48]。

同时，氢动力无人机日益拓展的作业环境也使得其不可避免地面临严寒天气或高海拔地区的极低环境温度，而冰点以下生成水会冻结堵塞孔隙与流道，导致电池启动失败或膜电极损伤。因此，低温冷启动技术也成为氢动力无人机燃料电池亟需突破的核心技术之一。下面对目前氢动力无人机上 PEMFC 的一些冷启动方案进行简要介绍。

首先，使用外部设备辅助加热的方案，典型如空气预热、端板与绝缘板加热。前者利用外部加热设备预热进入 PEMFC 的空气，从而提升电堆温度。加热空气能快速提高 MEA 的温度，有利于促发电化学反应[49]。然而，在无人机高海拔地区作业时，由于空气较为稀薄、热传导系数降低，该方法的能量利用效率会受到较大限制。后者则直接对 PEMFC 电堆的端板和绝缘板进行加热，从而使系统整体温度上升[49]。另外，通过增大启动电流也能实现系统的电化学自加热，但该方案需要对电流进行精准控制以免损坏系统中的元件[50]。相比之下，近年兴起的交流激励（alternating current, AC）阻抗加热技术通过内部焦耳热升温，加热更均匀且不易损伤组件[51]。根据 PEMFC 电堆内各组件之间的热性能差异，通过分阶段加热控制，优先加热热容较小的组件（如 MEA）以促发反应，再逐步加热热容较大的组件（如端板），能够在降低启动能耗的同时实现快速启动[52]。值得注意的是，停机后吹扫对冷启动也非常重要。该方法在系统关机后，利用干燥空气或氢气大流量吹扫电堆内部，将残余水分彻底排出，从根本上消除结冰隐患。

总体而言，面向氢动力无人机轻量化、高功重比以及复杂环境作业的发展需求，氢燃料电池仍存在若干亟待突破的关键技术问题。首先，需要进一步优化双极板材料与结构，在实现轻量化的同时兼顾导电性与耐腐蚀性；其次，需要提升燃料电池的热管理能力，以保证电堆在高功率密度运行条件下维持适宜的工作温度与稳定性能；同时，针对复杂恶劣的工作环境，需进一步增强燃料电池的环境适应能力；此外，面向严寒、高海拔等低温作业场景，快速、可靠的冷启动技术同样需要进一步研究。未来，如何在轻量化、高功率密度、环境适应性与运行可靠性之间实现协同优化，将是推动氢燃料电池进一步满足无人机工程应用需求的重要研究方向。

2.1.2 多源混合能源系统技术

氢燃料电池动力固然有着众多优点，但当其应用于对瞬时功率要求较高的任务[53]时性能往往受限。这种局限性在不同机型上表现各异：固定翼无人机在起飞爬升阶段需要极大的峰值功率，而进入巡航阶段后需求则显著降低；多旋翼无人机则在垂直起降、紧急规避或频繁的姿态调整中对动力系统提出较大的瞬时功率需求。综上所述，氢燃料电池更加适合长航时巡航这种需要长时间放电的任务，而不适合有瞬时大功率需求的任务场景。由此可见，独立的燃料电池动力系统难以兼顾现代高性能无人机严苛的双重要求。因此，构建多源混合能源系统（hybrid energy system, HES）已经成为解决氢燃料电池[54-55]缺点的有效技术路线。

氢燃料电池无人机的多源混合能源系统通常将氢燃料电池和一个或多个辅助电源相结合，其核心优势在于辅助电源可以提供峰值功率，进而在一定程度上将氢燃料电池与动态负载解耦，使得燃料电池能够在更稳定、更高效的输出范围内持续运行，有利于提高系统整体能量利用率和可靠性。从动力系统架构的角度来看，目前氢燃料电池无人机的主流配置是串联混合动力架构，如图 7 所示[56]。

在这种架构中，燃料电池系统与辅助电源通过电耦合方式连接到公共高压直流母线上。其中，燃料电池充当机载动力单元，提供主要能量；能量管理系统（energy management system, EMS）作为核心控制器，通过 DC/DC 转换器调节各能源的功率流向。电子调速器（electronic speed controller, ESC）直接从这一稳定的直流母线获取电能，并调制电流以驱动无刷电机和螺旋桨。这种配置实现了机械推进与电化学发电的解耦，使得燃料电池无论螺旋桨转速如何波动，都能在相对稳定的工况点运行。

混合能源系统的性能直接取决于所选的辅助电源类型。目前，最常见的 2 种类型是锂电池和超级电容器。近年来，传统液态锂电池在功率与能量密度上已逼近瓶颈。为了突破性能限制，固态锂电池逐渐成为研究热点，其用固态的电解质替代了易燃的液态电解质，避免了挤压/高温时泄漏起火的可能，安全性较优，且能兼容容量更大的锂金属负极，

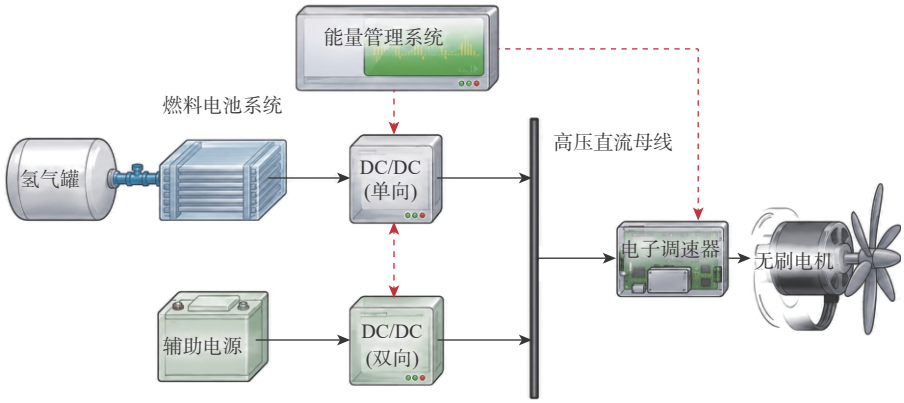


图 7　串联混合动力架构[56]

Fig. 7　Series hybrid power architecture[56]

能量密度更高, 即便是作为过渡方案的半固态锂电池, 能量密度也能够达到 350 Wh/kg 左右, 目前半固态锂电池已开始走向产业化[57-58]。与此同时, 传统对称型超级电容器虽具备高功率特性, 却因能量密度过低(通常＜10 Wh/kg)难以满足系统轻量化要求[59]。因此, 兼具超级电容高功率与锂电池高能量特性的锂离子超级电容器, 正成为无人机混合能源系统中的研究热点[60]。另外, 也有研究在探索太阳能、氢燃料电池、锂电池三能一体的混合动力。

另外, 不同构型无人机的混合能源系统配置应根据其任务功率需求进行差异化匹配。以巡航为主的固定翼无人机, 可由燃料电池主要承担持续功率需求, 辅助电源用于补偿短时峰值负载; 多旋翼无人机则需要提高辅助电源的功率配置能力, 以满足高功率和快速负载变化需求; 对于复合翼无人机, 可按照“燃料电池满足巡航持续功率、辅助电源补充垂直起降及过渡阶段峰值功率”的原则进行配置。

2.2　机载储氢技术

对于氢动力无人机而言, 高效且具备高动态响应能力的动力系统是其应对复杂任务环境的基础; 同时, 足量的能源储备则是提升其连续作业能力、拓展工作半径的关键。目前, 机载储氢技术主要分为三大类: 高压气态储氢、低温液态储氢以及基于固体载体的固态储氢。

2.2.1　高压气态储氢技术

高压气态储氢是指将氢气压缩至高压状态并储存于高强度气瓶中, 该技术的关键在于高压气瓶的结构强度和耐久性。目前, 根据储氢瓶材料可以将储氢瓶分为 5 种类型, 这 5 种类型氢气瓶的数据如表 1 所示[61]。其中, Ⅳ型瓶和处于前沿发展阶段的Ⅴ型瓶均属于复合材料储氢瓶, 其在轻量化的同时具有更高的质量储氢密度。

表 1　5 种高压储氢瓶的性能对比[61]

Table 1　Performance comparison of 5 types high-pressure hydrogen storage cylinders[61]

罐体类型	材料组成	最大内压/MPa	质量储氢密度/%
Ⅰ	金属储瓶	<20	1
Ⅱ	纤维复合材料包覆的金属瓶	<30	2
Ⅲ	带金属衬垫的复合瓶	35～70	3.9～5
Ⅳ	带有聚合物内衬的复合瓶	70	5
Ⅴ	无内衬复合瓶	>70	7.5

氢动力无人机的起飞重量受电机推力的严格限制, 而采用轻量化复合材料的储氢瓶则能在保证氢燃料储量不变的情况下, 显著降低储氢系统的质量, 从而为各无人机搭载的各种传感器或任务载荷留出更多余量。此外, 为确保气瓶的耐久性, 材料的抗氢脆性[62-63]和耐腐蚀性也是高压气瓶需要重点关注的问题。故而, 质量效率高且无氢脆现象的Ⅳ型瓶以及未来的Ⅴ型瓶更适合用于氢能源无人机[59]。除了瓶身材质的选择外, 通过提高气瓶储氢压力可在相同储氢量下减小气瓶容积与质量, 这也是实现无人机储氢系统轻量化的重要途径。目前, 氢动力无人机的储氢系统多采用 35～45 MPa 的氢气, 70 MPa 的Ⅳ型瓶技术在航空领域的应用尚处于探索阶段。此外, 减压阀门作为高压气态储氢技术的相关部件, 也对无人机轻量化做出了一定的贡献。传统工业或车载减压阀多采用不锈钢材质且结构冗余, 导致体积与质量偏大, 难以满足无人机严格的轻量化需求。近年来, 通过引入新型材料, 并结合阀体结构集成设计, 无人机专用减压阀的质量得到了明显降低。

总体而言, 压缩气态储氢凭借系统结构相对简单、技术成熟度较高以及标准体系较为完善等优势, 已成为目前工程应用中最为成熟的储氢方式[64]。受无人机起飞重量和机载空间的限制, 该技术未来的突破重点主要集中在储氢瓶轻量化和更高压力

等级储氢系统的航空应用。一方面，应进一步优化储氢瓶的材料与结构，在保证强度、耐久性和安全性的同时降低气瓶质量；另一方面，还需进一步研究70 MPa级储氢技术在航空领域的适用性，从而进一步提高氢动力无人机的机载储氢能力和续航性能。

2.2.2 低温液态储氢技术

低温液态储氢技术的核心思想是：将氢气冷却至−253 ℃以下，使其转变为液氢以储存于特制的超低温容器中。与高压气态储氢技术中使用的氢气相比，液氢密度更大，可达70.8 g/L[65]，这使得在相同的质量条件下其所需的储存空间更小。故而，液氢成为突破氢动力无人机续航瓶颈的关键路径。Stroman等的飞行实验表明[66]，在同等能源系统质量下，液氢系统提供的飞行航时相比高压气态储氢系统提升了约85%，充分展现了其在长航时飞行任务中的潜力。

尽管该技术相较于高压气态储氢具有更高的能量密度，但在实际使用中也存在一些问题。首先，保存液氢所需的超低温环境对储罐的温度控制能力提出了严苛的要求。当储罐的低温环境出现问题而导致液氢蒸发损耗，就会显著影响无人机的工作效能。为此，在实际应用中，液氢储罐会使用主动温度控制和被动温度控制2种技术。被动温度控制常常借助绝热层实现，常见的设计方案包括喷涂泡沫绝热和多层材料绝热，其中，后者的性能相对更加可靠[61]。同时，为了满足氢动力无人机轻量化的需求，近年来部分新型储罐的设计采用了薄壁承压结构与优化的轻量化支撑结构，此外，石墨烯改性气凝胶和真空绝热板等超轻隔热材料已处于实验室阶段[59]。主动温度控制主要利用低温制冷机来实现。在具体应用中，系统通过传感器实时监测储罐内的压力与温度，据此评估热渗透强度。当热渗透加剧时，控制系统会主动调高制冷机功率进行补偿。这种机制能够确保在作业过程中，储罐内的液氢除燃料电池正常消耗之外，不会因储罐绝热材料性能衰减或环境波动而发生不必要的损耗。

另外，液氢罐的循环使用次数受限于材料老化和结构疲劳。在液氢罐循环充装和排放液氢的过程中，罐体材料受到温度变化引起的热应力影响，疲劳损伤逐渐积累，最终可能引发微裂纹和泄漏[67]。此外，在实际使用中，这类储罐每次充装时，需要进行严格的预冷与吹扫工序，这不仅增加了操作的复杂度，还可能导致部分液氢损耗。为此，最新的研究与工程实践，如中国航天科技集团有限公司第六研究院的移动式液氢加注撬及液氢加氢机，开始采用闭式加注与泵压过冷技术，如图8所示。该技术通过构建闭环加注回路，工艺流程中无需向空排放，极大地降低了加注过程的蒸发损耗，同时通过单路连接降低了操作复杂度，提升了加注效率[68-69]。



图8 移动式液氢加注撬及液氢加氢机[69]

Fig. 8 Mobile liquid hydrogen refueling skid and liquid hydrogen dispenser[69]

综上所述，低温液态储氢技术凭借高能量密度，成为突破氢动力无人机续航瓶颈的关键方案之一。然而，高昂的储罐制造成本与绝热系统带来的额外载荷，使其更适用于载荷能力更强的中大型无人机。此外，尽管加注工艺持续优化，但全球液氢供应链与基础设施建设尚处于起步阶段，这也限制了其大规模的商业化应用。同为液态储运，新兴的液氨路线能够通过机载氨裂解制氢供给燃料电池，并且凭借更高的体积储氢密度、更低的储运要求（常温低压），以及更优的安全性，被研判为超长航时重载无人机的长期最优/终极平衡点技术路线。

2.2.3 固态储氢技术

固态储氢技术是指通过物理或化学手段使氢气以固态的形式被储存起来。固态储氢依据储氢机理不同，主流方法可以分为2种：一种是氢气分子通过物理吸附作用与储氢材料结合；另一种是氢通过化学键与其他成分发生化学结合形成稳定的氢化物[70]，使用时则通过热解或化学反应释放氢气以供能。相比于高压气态和低温液态储氢，固态储氢不仅展现出更高的体积储氢密度，而且在使用上更加安全。该技术的关键在于高效地实现氢气与其他材料的结合与释放。因此，固态储氢相关研究主要聚焦于储氢材料的物理化学性质[71]。

常用的物理吸附型储氢材料主要包括碳基材料、硅基材料、金属有机框架材料[72]以及其他具有高比表面积的多孔材料，氢分子通过范德瓦耳斯力与这些材料进行物理结合。但这种方式受限于材料性能和规模化制备仍处于研发阶段。

化学储氢材料主要包括配位氢化物材料和金

属基合金材料。其中, 最常用的是硼氢化钠。韩国科研人员开发了一种以硼氢化钠为储氢材料的无人机氢气发生器[73-74], 实验数据表明, 该氢气发生器的质量比能量和体积比能量分别为 739.1 Wh/kg 和 272.8 Wh/L, 且具备稳定的产氢速率。但其在投入实际应用的过程中也出现了诸多问题, 一方面, 在硼氢化钠体系中所使用的催化剂寿命较短, 以至于每次飞行都要重新装填, 这增加了使用成本, 限制了其商业应用潜力, 另一方面, 该体系在制氢过程中产生的副产物有可能堵塞燃料电池的燃料进气口, 进而影响整个电池系统的性能[75]。

与配位氢化物材料相比, 低压金属储氢合金材料通过可逆吸放氢反应实现固态储氢, 无需复杂的水解过程。该类材料在安全性与环境适应性方面具有优势, 但普遍面临合金密度大、系统质量偏高以及对部分关键金属元素依赖较强等问题, 限制了其在对载重与成本敏感的领域的进一步推广。

近年来, 镁基储氢材料凭借较高的质量储氢密度与原料来源广泛的优势也受到较多关注, 但其放氢温度偏高、动力学缓慢, 仍有待进一步优化。除此之外, 气-固氢负离子电池等新兴概念尝试将气态氢、固态储氢与电化学储能相耦合, 代表了一种具有前瞻性的技术路径, 目前尚处于原理验证与原型探索阶段[76]。

总体而言, 固态储氢具有体积储氢密度高、安全性较好等优势, 在氢动力无人机领域具有较大的应用潜力, 但目前整体仍处于研究和探索阶段。该技术未来的突破重点集中在高性能储氢材料开发和储放氢性能优化。一方面, 应进一步提高物理吸附材料的储氢能力并改善其规模化制备水平; 另一方面, 针对化学储氢和金属基储氢材料在催化剂寿命、副产物处理、系统质量、放氢温度等方面的不同限制, 还需进一步优化材料组成与反应过程, 从而实现氢气的高效、稳定储放, 推动固态储氢技术在无人机中的实际应用。

3　能量管理策略

设计的多源混合能源系统通过采用 2 种或多种互补的动力源, 能够使得氢动力无人机更好地面对实际作业中飞行工况复杂多变、功率需求非线性波动的问题。然而, 在此过程中, 必须通过能量管理策略对燃料电池与辅助电源的功率输出进行实时协调、动态分配与优化控制, 才能保障氢动力无人机高效、可靠、长航时运行。氢燃料电池无人机混合能源系统的能量管理逻辑框架如图 9 所示。

目前, 主流的能量管理策略有三大类: 基于规则的策略、基于优化的策略以及基于机器学习的策略。下面对这 3 类方法分别进行介绍。

3.1　基于规则的能量管理策略

基于规则的能量管理策略与大量的工程经验

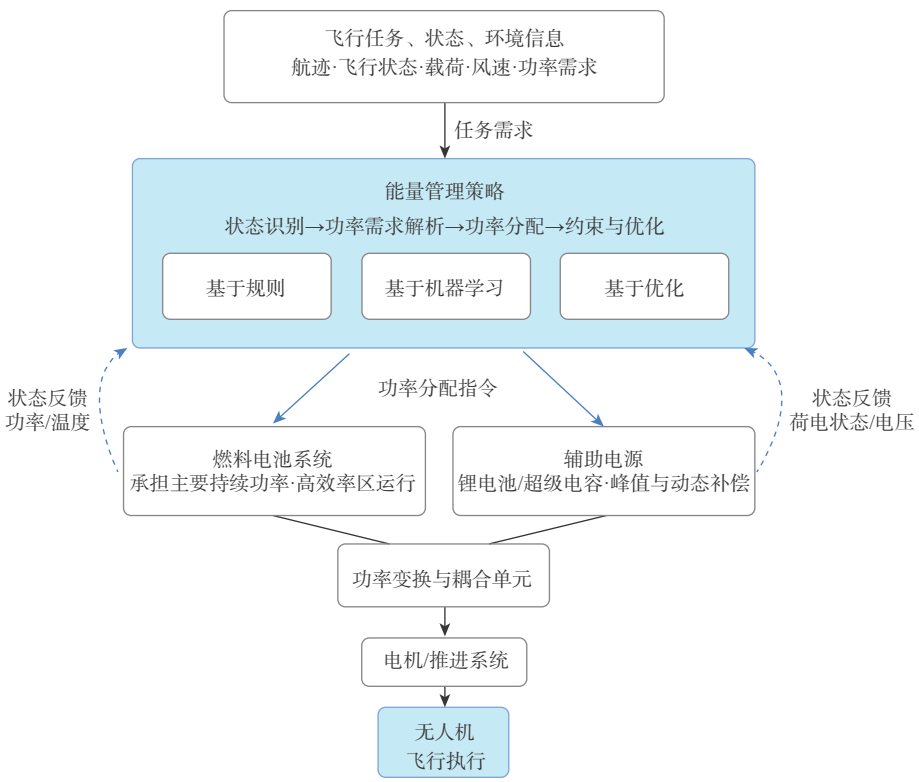


图 9　混合能源系统能量管理总体框架

Fig. 9　Overall energy management framework for hybrid energy system

息息相关[77]。在多年的工程实践中，面对各种飞行情况的合理应对措施被总结成规则，使得无人机在实景工作环境中，能够通过查询自身烧录的规则库得到对实时情况的合理应对措施。这些经验性的总结往往会呈现为“如果……那么……”的逻辑结构。例如，“如果无人机将保持巡航飞行，那么就增大燃料电池的供电比例，同时限制辅助电源的输出。”总体而言，基于规则的能量管理策略可以分为2类：确定性规则控制和模糊逻辑控制（fuzzy logic control, FLC）。

在工程应用中，确定性规则控制凭借控制结构简单且实时性强等优点成为主要的控制方法。此类方法的核心思想是通过对实时状态变量与固定阈值的比对，实现对系统能量的瞬时分配。确定性规则控制的研究众多，如Yan等[78]利用基于确定性规则的能量管理策略，使得在相同飞行剖面下，飞行器续航时间延长16.5 min，系统总重减少302.8 g，运行成本降低41.58美元，取得了显著的性能提升。Yang等[79]利用基于状态机的能量管理策略来实现确定性规则控制，在不同的电池荷电状态（state of charge, SOC）条件下进行了多次飞行试验，结果证实，该策略优化了功率分配并增强了飞行器的动态性能。

模糊逻辑控制的核心是一套模拟人类经验的“IF-THEN”规则库，其并不直接处理各种传感器传回的精确数字，而是通过模糊化过程将复杂的系统状态（如电压、电量）转化为“高、中、低”等模糊概念，从而实现更灵活的决策。模糊逻辑控制凭借卓越的非线性处理能力和抗干扰特性，在燃料电池能量管理中得到了广泛的应用。Zhang等[80]提出了一种在线模糊逻辑控制策略，相比于被动控制以及状态机策略，设计的系统在降低氢气消耗和提高电池使用效率方面有明显提升。Shen等[81]利用模糊逻辑控制精确地将燃料电池的工作点维持在最有效的区域内，在多种标准工况下进行实验，证明了该方法的优越性。

3.2 基于优化的能量管理策略

基于优化的能量管理策略通常涉及代价函数的建立。该类策略将确定燃料电池与辅助电源之间最优功率分配的过程转化为最小化代价函数的过程。通常情况下，该代价函数代表氢气消耗量、系统运行成本或者电池降解程度。此外，为了将系统的实时状态和飞行任务的执行情况也纳入代价函数，还需要引入额外的度量指标，如燃料电池堆内部的温度变化、电力稳定性指标、燃料电池的启停循环次数等。

在实际应用中，能量管理策略常通过动态规划中的加权惩罚项或模型预测控制（model predictive control, MPC）中的实时反馈将上述指标整合到代价函数中。这使得能量管理策略能够预测组件的健康状况，并据此调整无人机飞行任务剖面。相比基于规则的方法，基于优化的策略通常更加高效，因为其利用数学优化算法定义了明确的目标函数，在电池SOC和组件功率限制等约束条件下，能够统筹兼顾系统能耗、燃料电池耐久性、辅助电源降解程度及运行成本等。

总体上，根据优化范围和应用场景的不同，基于优化的策略主要分为2类：全局优化和实时优化[82]。全局优化策略旨在已知的完整任务周期内寻求理论上的最优解，其代表性算法包括动态规划（dynamic programming, DP）和庞特里亚金最小值原理（Pontryagin’s minimum principle, PMP）[83]。这类方法通常需要预先掌握全部运行工况的信息，且计算量巨大。因此，主要用于离线分析，其计算结果往往作为理论基准，用以评估其他在线实时策略的性能及表现。实时优化则要在缺少全局信息的情况下，尽可能地去逼近全局最优解。其中，2种最核心的算法是等效耗能最小策略（equivalent consumption minimization strategy, ECMS）[84]和模型预测控制[85]。

基于优化的能量管理策略也有着广泛的应用。Bradley[86]采用动态规划方法设计了无人机的能量管理策略。戴月领等[87]则采用模型预测能源管理策略，并验证了所提模型的有效性。

3.3 基于机器学习的能量管理策略

近年来，人工智能相关技术飞速进步，推动了智能能量管理策略的发展。这类策略主要包括神经网络（neural network, NN）、强化学习（reinforcement learning, RL）以及深度强化学习（deep reinforcement learning, DRL）等。下面对这3类方法分别展开介绍。

基于神经网络的策略主要采用有监督的学习模型。使用大量飞行数据对模型进行训练，使其能够拟合出一种复杂的映射函数，该函数能够将实时工况特征映射为混合能源系统的功率分配指令。然而，这类能源管理策略不可避免地受到神经网络弊端的影响，即性能高度依赖于训练数据的质量和规模[88]。相比之下，强化学习的核心思想是让智能体在奖励的引导下，通过与“环境”不断地交互、试错来自主学习最优的能量管理策略。这种智能体的训练过程中不需要事先标注好的“高质量”决策数据。而深度强化学习结合了深度神经网络强大

的表征能力, 显著提升了智能体处理高维状态-动作空间问题的潜力。总体而言, 这些智能控制策略都具备自学习与自适应能力, 能更好地应对混合动力旋翼无人机等具有高度非线性和动态特性的系统能量管理问题, 具有广阔的应用前景。

综上所述, 图 10 从策略类型、代表方法、主要优势及局限性等方面, 对氢燃料电池无人机常用能量管理策略进行了归纳和比较。

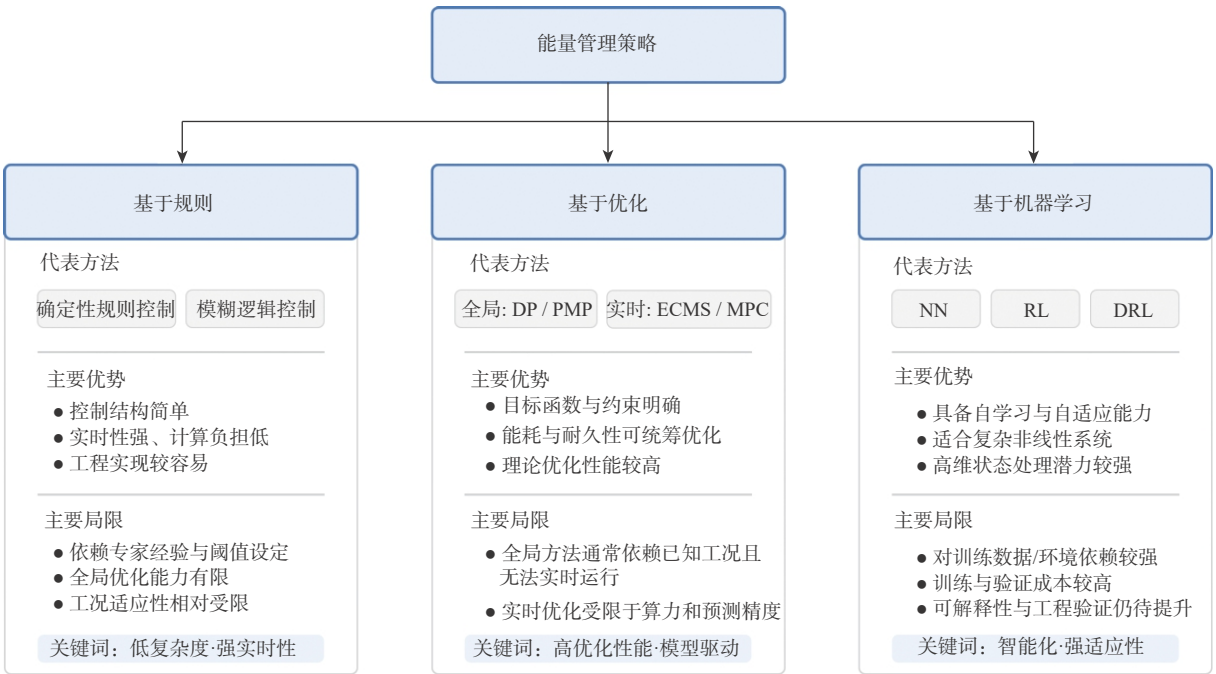


图 10　氢燃料电池无人机能量管理策略的分类及主要特征

Fig. 10　Classification and key characteristics of energy management strategies for hydrogen fuel cell UAVs

3.4　能量管理策略评价指标

为客观衡量能量管理策略的优劣, 需从能耗、能源利用效率、动态响应特性、算法复杂度、实时性、鲁棒性、适应性等维度建立评价体系。常见的评价指标如表 2 所示。

表 2　能量管理策略常见评价指标

Table 2　Common evaluation metrics for energy management strategies

评价维度	常用具体评价指标
氢耗	瞬时氢气消耗率/任务周期总氢耗
能源利用效率	燃料电池系统效率
动态响应特性	功率跟踪精度、响应延时、电压波动范围
算法实时性与计算开销	单步计算时间、训练时间/收敛速度(学习类策略)
鲁棒性与适应性	扰动及工况变化下主要性能指标的变化情况

3.5　未来发展方向

面向未来轻量化、高可靠、高效率、大载重氢动力无人机的发展需求, 能量管理技术可沿以下方向发展:

1) 大系统的深度耦合。将能量管理与飞行控制、任务规划联动, 根据航迹、风速、载荷预调度功率, 提升全局性能。

2) 轻量化智能算法。目前的许多管理策略依然需要一定量的数据作为前处理/学习, 因此未来可以聚焦以更少数据量实现等效控制精度, 降低算法对数据预处理与学习资源的依赖, 简化机载计算需求。

3) 多机协同的能量管理策略。面向无人机集群作业, 实现机间功率分配、充加氢调度与任务协同。

4) 标准化与集成化。形成面向不同构型、不同功率等级的通用能量管理策略框架, 加速工程化与产业化应用。

4　热管理策略

热管理对 PEMFC 的安全和工作效率至关重要。PEMFC 最佳工作温度通常在 80 ℃ 左右[89]。如果工作温度过高, 会导致电池内部的质子交换膜干燥、性能下降并加速极板劣化, 而温度过低, 则会抑制化学反应动力学。此外, 温度分布不均也会直接影响反应物分布、电流密度等, 进而干扰 PEMFC 的整体性能。然而, 受制于热力学定律, 任何动力装置的能量转换效率都不可能达到 100%, 必然有相当数额的能量被耗散掉, 以热能的形式释放。因此, 对于以 PEMFC 为动力装置的无人机而言, 高效的热管理显得尤为关键。本节以 PEMFC 为例, 从散热技术以及温度控制策略 2 个方面介绍氢能无

人机的热管理技术。

4.1 散热技术

PEMFC常用的散热方式包括液冷和空冷2种，二者在换热效率、系统复杂性及飞行器适配性上各有优劣。

液冷系统的优势是极高的换热效率和热容量，使得其非常适用于大功率PEMFC系统。在实际运行时，液冷系统依靠在冷却管道中内循环的冷却液（如高纯度去离子水、乙二醇水混合物、丙二醇水混合物），带走电池电堆产生的热量。根据接触方式的不同，液冷可分为直接接触式（浸入式）与间接接触式。直接接触式液冷系统设计简单、热阻低，且能有效抑制热失控，但对电池模块的密封要求比较严格；间接接触式液冷系统则往往通过额外的结构设计（如冷却板）实现电堆和冷却剂之间的传热，这虽增加了系统的复杂性和热阻，但具有更大的设计空间。尽管液冷散热系统的效果显著，但其所需的泵、管道及冷却液却增加了系统的复杂性和使用成本，同时增加了系统发生泄漏的风险[89]。

与液冷散热系统相反，空冷散热系统的优势是结构简单、设计轻便，无需额外的泵或管道，这使得其更加适用于对体积和载荷敏感的任务载体[90]。空冷系统通过增大阴极空气流量或设计独立的冷却通道，利用空气的自然对流或强制对流同时实现反应供氧与电堆散热。虽然空冷的换热效率低于液冷，但对于一些中小型无人机上搭载的低功率PEMFC系统而言，空冷也能够满足电池系统的散热需求[91]。

值得一提的是，针对不同机型，空冷散热系统的应用形式也存在一些差异。在无人机高速飞行过程中（如固定翼平飞），散热系统可以利用动压将空气吸入机体内部用以冷却电堆[92]。而当无人机停放于地面、低速飞行或者空中悬停（旋翼无人机）时，则要利用额外的装置（如风扇）来吸入冷却空气。旋翼无人机由于常处于悬停或低速飞行状态，故而对强制对流（即额外装置）的依赖更强，相对的固定翼飞机则更多地利用飞行时的来流进行散热。

此外，对于散热系统而言，燃料电池双极板材料的选择和优化也对电池的热管理有很大影响。高导热率材料能够有效传导电堆内部的热量，提高散热效率并减少局部热点的形成[93]。为了满足氢动力无人机轻量化与高导热率的双重需求，具有极薄厚度和高导热率的金属双极板（如表面改性的钛板或不锈钢板）正逐步取代传统石墨双极板。

总体而言，氢燃料电池无人机散热技术需要在散热能力、系统质量及辅助功耗之间进行权衡，同时要根据不同机型和飞行工况合理选择散热方式。未来需要突破的核心技术主要集中在轻量化、高效低功散热系统的设计，要通过优化进排气布局、液冷回路以及双极板材料与结构等手段，提高电堆散热效率。

4.2 温度控制策略

温度控制技术主要在燃料电池进入常规工作区间后发挥作用，其指的是通过实时监测电堆内部或出口的温度信号，利用特定的控制算法（如比例积分、模型预测控制、模糊逻辑等），动态调节冷却介质（空气或冷却液）的流量与流向，从而使电堆在不同负载需求和环境干扰下，维持在适宜的工作温度范围，以平衡水分管理、电化学反应速率和系统寿命的一套逻辑体系。高效的温度控制策略能够确保系统在稳态和动态负载波动下稳定运行，从而提高反应效率并延长燃料电池系统的使用寿命[94]。由于无人机在飞行过程中面临阵风干扰、爬升等高动态负载波动，且跨高度层飞行时会遭遇气压与温度的骤变，因此对温度控制策略的自适应能力和响应速度提出了极高要求。近年来，国内外学者围绕温度控制策略进行了大量探索。

Yuan等[95]提出了一种基于三维数值分析的多节点控制模型，通过有限状态机与比例积分算法协同调节冷却空气的流向与流速，实验结果显示，该方案成功将电堆内部温差控制在0.5 ℃以内，显著提高了电堆的性能。Wang等[96]提出了一种基于主动温度控制的空冷PEMFC实时功率优化策略，设计了一种改进的温度扰动观察算法，从而能够实时获取最佳参考温度，此外，其还采用基于超螺旋算法的控制器，实现了在多变环境条件下对目标温度的精确跟踪，实验结果显示，在相对较高的负载下，输出功率提升超过4%，验证了该策略的有效性。同时，近年来的温控策略研究正从单一的“温度跟踪”向“热-电-水-气”多物理场协同控制演进[97]。研究人员开始引入动态规划与机器学习算法，将电堆温度与内部膜湿度、风扇/水泵能耗进行全局耦合求解[98]。这类数据驱动的智能控制策略能够降低对精确电堆物理模型的依赖，同时可结合无人机实时飞行数据进行自适应调整。

总体而言，氢燃料电池无人机温度控制策略需要综合考虑控制精度、动态响应能力、自适应能力等。未来的研究主要集中于复杂飞行工况下的快速自适应温度控制与多变量协同控制等方面。

5 结束语

本文系统综述了氢燃料电池混合动力无人机

的研究进展与关键技术。首先，介绍了氢能源航空器的发展历程，阐述了氢燃料电池作为无人机动力源的独特优势，以及国内外典型氢动力无人机型号。然后，对氢燃料电池混合动力无人机涉及的关键技术进行了介绍，包括燃料电池混合动力系统、机载储氢技术、能量管理策略以及热管理技术。总体而言，氢燃料电池混合动力无人机技术顺应“绿色”的时代要求、受到国家政策支持，正处于关键技术攻关与早期应用示范并行的关键时期。展望未来，氢燃料电池混合动力无人机将沿着“轻量化、高可靠性、高效率、大载重”的路径继续演进，具体体现在以下方向：

1）核心电堆组件的迭代与轻量化。散热与载重的权衡是航空工程的永恒命题。燃料电池堆技术将从目前主流的石墨双极板向更薄、功率密度更高的金属双极板过渡，更长远的目标则是实现闭式电堆设计。

2）混动系统的深度集成与控制联动。未来的混合动力系统将不再是独立的能量源，而是追求极高的可靠性与冗余度。一方面，需开发具备工业级稳定性的氢电混动集成方案；另一方面，必须实现动力系统与飞行控制系统的深度联动。

3）应用场景向“重载运输”拓展。目前以“天目山一号”为代表的多旋翼巡检类飞机已在电力、管网等领域实现初步应用。下一步的重点将转向载重型飞行器，实现从数十千克到数百千克级的运载能力跨越。随之而来的是对 eVTOL 及支线客机等更高级别航空器的探索。

4）控制系统的多模态智能化。未来的氢动力无人机将部署大语言模型与视觉语言模型，实现自然语言交互下达任务以及高精度视觉感知，从而大幅降低用户的操作门槛。

5）热管理系统的多物理场协同。未来氢动力无人机热管理将融合多算法调控技术，依托多物理场耦合机制，实现工况自适应温控，抑制电池温差，适配复杂飞行环境，全面提升电堆运行稳定性与使用寿命。

参考文献（References）

[1] Arat H T, Sürer M G. Experimental investigation of fuel cell usage on an air vehicle’s hybrid propulsion system[J]. International Journal of Hydrogen Energy, 2020, 45(49): 26370-26378.

[2] Floreano D, Wood R J. Science, technology and the future of small autonomous drones[J]. Nature, 2015, 521(7553): 460-466.

[3] Mademlis I, Mygdalis V, Nikolaidis N, et al. High-level multiple-UAV cinematography tools for covering outdoor events[J]. IEEE Transactions on Broadcasting, 2019, 65(3): 627-635.

[4] Ahmed F, Mohanta J C, Keshari A, et al. Recent advances in unmanned aerial vehicles: a review[J]. Arabian Journal for Science and Engineering, 2022, 47(7): 7963-7984.

[5] Luo H, Zhang P, Wang J J, et al. Traffic patrolling routing problem with drones in an urban road system[J]. Sensors, 2019, 19(23): 5164.

[6] Humpe A. Bridge inspection with an off-the-shelf 360° camera drone[J]. Drones, 2020, 4(4): 67.

[7] Velusamy P, Rajendran S, Mahendran R K, et al. Unmanned aerial vehicles (UAV) in precision agriculture: applications and challenges[J]. Energies, 2022, 15(1): 217.

[8] Phang S K, Chiang T H A, Happonen A, et al. From satellite to UAV-based remote sensing: a review on precision agriculture[J]. IEEE Access, 2023, 11: 127057-127076.

[9] 张扬军, 彭杰, 钱煜平, 等. 氢能航空的关键技术与挑战[J]. 航空动力, 2021(1): 20-23.
Zhang Y J, Peng J, Qian Y P, et al. Key technologies and challenges of hydrogen powered aviation[J]. Aerospace Power, 2021(1): 20-23(in Chinese).

[10] Tiwari S, Pekris M J, Doherty J J. A review of liquid hydrogen aircraft and propulsion technologies[J]. International Journal of Hydrogen Energy, 2024, 57: 1174-1196.

[11] Apeland J, Pavlou D G, Hemmingsen T. Sensitivity study of design parameters for a fuel cell powered multirotor drone[J]. Journal of Intelligent & Robotic Systems, 2021, 102(1): 6.

[12] Depcik C, Cassady T, Collicott B, et al. Comparison of lithium ion batteries, hydrogen fueled combustion engines, and a hydrogen fuel cell in powering a small unmanned aerial vehicle[J]. Energy Conversion and Management, 2020, 207: 112514.

[13] 国家市场监督管理总局, 国家标准化管理委员会. 无人机用氢燃料电池发电系统: GB/T 38954—2020[S]. 北京: 中国标准出版社, 2020.
State Administration for Market Regulation, Standardization Administration of China. Hydrogen fuel cell power system for unmanned aerial vehicles: GB/T 38954—2020[S]. Beijing: Standards Press of China, 2020(in Chinese).

[14] 国家发展改革委, 国家能源局. 氢能产业发展中长期规划 (2021—2035 年)[EB/OL]. (2022-03-23)[2026-08-09]. https://zfxxgk.nea.gov.cn/1310525630_16479984022991n.pdf.
National Development and Reform Commission, National Energy Administration. Medium- and long-term plan for the development of hydrogen energy industry (2021–2035)[EB/OL]. (2022-03-23)[2026-08-09]. https://zfxxgk.nea.gov.cn/1310525630_16479984022991n.pdf(in Chinese).

[15] 工业和信息化部, 科学技术部, 财政部, 等. 绿色航空制造业发展纲要 (2023—2035 年)[EB/OL]. (203-10-01)[2026-08-09]. https://www.gov.cn/zhengce/zhengceku/202310/content_6908243.htm.
Ministry of Industry and Information Technology, Ministry of Science and Technology, Ministry of Finance, et al. Outline for the development of green aviation manufacturing industry (2023—2035)[EB/OL]. (203-10-01)[2026-08-09]. https://www.gov.cn/zhengce/zhengceku/202310/content_6908243.htm(in Chinese).

[16] 工业和信息化部, 财政部, 国家发展改革委. 工业和信息化部 财政部 国家发展改革委关于开展氢能综合应用试点工作的通知[EB/OL]. (2026-03-06)[2026-08-09]. https://www.gov.cn/zhengce/zhengceku/202603/content_7063704.htm.
Ministry of Industry and Information Technology, Ministry of Fin-

ance, National Development and Reform Commission. Notice of the Ministry of Industry and Information Technology, the Ministry of Finance and the National Development and Reform Commission on launching pilot projects for comprehensive hydrogen energy applications[EB/OL]. (2026-03-06)[2026-08-09]. https://www.gov.cn/zhengce/zhengceku/202603/content_7063704.htm(in Chinese).

[17] Haglind F, Hasselrot A, Singh R. Potential of reducing the environmental impact of aviation by using hydrogen Part I: background, prospects and challenges[J]. The Aeronautical Journal, 2006, 110(1110): 533-540.

[18] Adler E J, Martins J R R A. Hydrogen-powered aircraft: fundamental concepts, key technologies, and environmental impacts[J]. Progress in Aerospace Sciences, 2023, 141: 100922.

[19] Rao A G, Yin F J, Werij H. Energy transition in aviation: the role of cryogenic fuels[J]. Aerospace, 2020, 7(12): 181.

[20] 向锦武, 阚梓, 邵浩原, 等. 长航时无人机关键技术研究进展[J]. 哈尔滨工业大学学报, 2020, 52(6): 57-77.
Xiang J W, Kan Z, Shao H Y, et al. A review of key technologies for long-endurance unmanned aerial vehicle[J]. Journal of Harbin Institute of Technology, 2020, 52(6): 57-77(in Chinese).

[21] Capurso T, Stefanizzi M, Torresi M, et al. Perspective of the role of hydrogen in the 21st century energy transition[J]. Energy Conversion and Management, 2022, 251: 114898.

[22] Swider-Lyons K, Schuette M, Stroman R, et al. Liquid hydrogen fuel system for small unmanned air vehicles[C]//51st AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition. Reston: AIAA, 2013: 467.

[23] 向锦武, 马凯, 阚梓, 等. 氢能源无人机关键技术研究进展[J]. 航空学报, 2025, 46(5): 531229.
Xiang J W, Ma K, Kan Z, et al. Review of key technologies for hydrogen powered unmanned aerial vehicles[J]. Acta Aeronautica et Astronautica Sinica, 2025, 46(5): 531229(in Chinese).

[24] 刘莉, 曹潇, 张晓辉, 等. 轻小型太阳能/氢能无人机发展综述[J]. 航空学报, 2020, 41(3): 1-28.
Liu L, Cao X, Zhang X H, et al. Review of development of light and small scale solar/hydrogen powered unmanned aerial vehicles[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(3): 1-28(in Chinese).

[25] 古力尼格尔·迪力夏提, 安斯奇. 中飞院氢能源无人机首飞成功![EB/OL]. (2025-04-22)[2026-05-20]. https://news.cafuc.edu.cn/info/1012/72035.htm.
Gulinigeer D, An S. Successful maiden flight of CAFUC hydrogen-powered UAV![EB/OL]. (2026-04-22)[2026-05-20]. https://news.cafuc.edu.cn/info/1012/72035.htm(in Chinese).

[26] 吴媛. 哈工大重庆研究院氢能无人机获重庆电视台报道 长航时“黑科技”赋能重庆低空经济[EB/OL]. (2026-03-17)[2026-05-20]. https://cri.hit.edu.cn/2026/0317/c15728a388402/page.htm.
Wu Y. Hydrogen-powered UAV of HIT Chongqing Research Institute reported by Chongqing TV: long-endurance "black technology" empowers Chongqing's low-altitude economy[EB/OL]. (2026-03-17)[2026-05-20]. https://cri.hit.edu.cn/2026/0317/c15728a388402/page.htm(in Chinese).

[27] 哈尔滨工业大学重庆研究院. “青鸥-30B”亮相湖南 完成跨区域实战首秀[EB/OL]. (2025-08-15)[2026-08-09]. https://cri.hit.edu.cn/2025/0815/c15728a376436/page.htm.
Harbin Institute of Technology Chongqing Research Institute. "Qingou-30B" debuts in Hunan, completing its first cross-regional field mission[EB/OL]. (2025-08-15)[2026-08-09]. https://cri.hit.edu.cn/2025/0815/c15728a376436/page.htm(in Chinese).

[28] 中国日报中文网氢航科技: 黑龙江、湖南百万国网订单双交付, 氢旋 4 号低温电力巡检再获佳绩![EB/OL]. (2025-09-16)[2026-05-11]. https://www.hydrogencraft.com/archives/bai-wan-guo-wang-ding-dan-shuang-jiao-fu.
Hydrogencraft Technology: dual delivery of millions of State Grid orders in Heilongjiang and Hunan, Qingxuan 4 achieves great success in low-temperature power inspection![EB/OL]. (2025-09-16)[2026-05-11]. https://www.hydrogencraft.com/archives/bai-wan-guo-wang-ding-dan-shuang-jiao-fu(in Chinese).

[29] 氢鹏科技. 188.6 公里!氢鹏科技“天目山一号”刷新吉尼斯纪录, 中国氢动力无人机领跑世界[EB/OL]. (2025-12-11)[2026-05-20]. https://www.hydronepower.com/doc_30436659.html.
Hepeng Technology (Zhejiang) Co. , Ltd. 188.6 km! Hydrone Power's "Tianmushan No. 1" breaks Guinness World Record, Chinese hydrogen-powered UAVs lead the world[EB/OL]. (2025-12-11)[2026-05-20]. https://www.hydronepower.com/doc_30436659.html(in Chinese).

[30] 天目山实验室. “天目山一号”全球首款百公里级氢动力长航程多旋翼无人机首飞成功[EB/OL]. (2024-08-26)[2026-05-20]. https://www.tmslab.cn/info/1004/5111.htm.
Tianmushan Laboratory. Successful maiden flight of "Tianmushan No. 1", the world's first 100-km class hydrogen-powered long-endurance multi-rotor UAV[EB/OL]. (2024-08-26)[2026-05-20]. https://www.tmslab.cn/info/1004/5111.htm(in Chinese).

[31] 人民网. “青鹞-10”氢动力多旋翼无人机完成液氢能源首次试飞[EB/OL]. (2025-09-23)[2026-05-20]. http://cq.people.com.cn/n2/2025/0923/c367647-41361062.html.
People's Daily Online. "Qingyao-10" hydrogen-powered multi-rotor UAV completes maiden flight using liquid hydrogen energy [EB/OL]. (2025-09-23)[2026-05-20]. http://cq.people.com.cn/n2/2025/0923/c367647-41361062.html(in Chinese).

[32] 氢鹏科技. 漠河极寒认证!氢鹏科技氢动力无人机全系通过实战考验[EB/OL]. (2026-02-26)[2026-05-20]. https://www.hydronepower.com/doc_30711294.html.
Hepeng Technology (Zhejiang) Co. , Ltd. Mohe extreme cold certification! The full series of Hydrone Power's hydrogen-powered UAVs passes field tests[EB/OL]. (2026-02-26)[2026-05-20]. https://www.hydronepower.com/doc_30711294.html(in Chinese).

[33] Alaka'i Technologies. Alaka'i Technologies launches world's first hydrogen-powered air mobility system[EB/OL]. (2019-05-31)[2026-05-20]. https://www.businesswirechina.com/zh/news/40741.html.

[34] Doosan Mobility Innovation. IDG leads 27 billion KRW investment in Doosan Mobility Innovation[EB/OL]. (2023-11-13)[2026-04-02]. https://www.doosanmobility.com/cn/about-us/news-events/news/21000612.

[35] Sharaf O Z, Orhan M F. An overview of fuel cell technology: fundamentals and applications[J]. Renewable and Sustainable Energy Reviews, 2014, 32: 810-853.

[36] Tang A, Crisci L, Bonville L, et al. An overview of bipolar plates in proton exchange membrane fuel cells[J]. Journal of Renewable and Sustainable Energy, 2021, 13(2): 022701.

[37] Liu C Y, Sung C C. A review of the performance and analysis of proton exchange membrane fuel cell membrane electrode assemblies[J]. Journal of Power Sources, 2012, 220: 348-353.

[38] Yin C, Hua S Y, Nie W, et al. Comparative study on air-cooled fuel cell stacks with metal and graphite bipolar plate designs for unmanned aerial vehicles[J]. ETransportation, 2024, 21: 100344.

[39] Nguyen H Q, Shabani B. Proton exchange membrane fuel cells heat recovery opportunities for combined heating/cooling and power applications[J]. Energy Conversion and Management, 2020, 204: 112328.

[40] Zhang G S, Kandlikar S G. A critical review of cooling techniques in proton exchange membrane fuel cell stacks[J]. International Journal of Hydrogen Energy, 2012, 37(3): 2412-2429.

[41] Ramezanizadeh M, Alhuyi Nazari M, Hossein Ahmadi M, et al. A review on the approaches applied for cooling fuel cells[J]. International Journal of Heat and Mass Transfer, 2019, 139: 517-525.

[42] Flückiger R, Tiefenauer A, Ruge M, et al. Thermal analysis and optimization of a portable, edge-air-cooled PEFC stack[J]. Journal of Power Sources, 2007, 172(1): 324-333.

[43] Zakhvatkin L, Schechter A, Buri E, et al. Edge cooling of a fuel cell during aerial missions by ambient air[J]. Micromachines, 2021, 12(11): 1432.

[44] Kurnia J C, Chaedir B A, Sasmito A P, et al. Progress on open cathode proton exchange membrane fuel cell: performance, designs, challenges and future directions[J]. Applied Energy, 2021, 283: 116359.

[45] Shen Z N, Liu S Q, Zhu W, et al. A review on key technologies and developments of hydrogen fuel cell multi-rotor drones[J]. Energies, 2024, 17(16): 4193.

[46] Cheng X, Shi Z, Glass N, et al. A review of PEM hydrogen fuel cell contamination: impacts, mechanisms, and mitigation[J]. Journal of Power Sources, 2007, 165(2): 739-756.

[47] Lamard M, Auvity B, Buttin P, et al. Impact of NaCl spray on the durability of PEMFC single cells and stacks in marine environment[J]. Journal of the Electrochemical Society, 2023, 170(2): 024504.

[48] 杨毅. 大连化物所成功研制高比功率阴极闭合式风冷电堆并通过科技成果鉴定[EB/OL]. (2026-05-10)[2026-06-27]. https://www.chinanews.com/gn/2026/05-10/10618707.shtml.
Yang Y. High specific power cathode-closed air-cooled fuel cell stack developed by Dalian Institute of Chemical Physics passes scientific and technological achievement appraisal[EB/OL]. (2026-05-10)[2026-06-27]. https://www.chinanews.com/gn/2026/05-10/10618707.shtml(in Chinese).

[49] Zhan Z G, Yuan C, Hu Z R, et al. Experimental study on different preheating methods for the cold-start of PEMFC stacks[J]. Energy, 2018, 162: 1029-1040.

[50] Lei L, He P, He P, et al. Numerical research on the cold start-up strategy of a PEMFC stack from −30℃[J]. Journal of Thermal Science, 2023, 32(3): 898-910.

[51] Kurz T, Küfner F, Gerteisen D. Heating of low and high temperature PEM fuel cells with alternating current[J]. Fuel Cells, 2018, 18(3): 326-334.

[52] Pan W T, Li P, Gan Q Q, et al. Thermal stability analysis of cold start processes in PEM fuel cells[J]. Applied Energy, 2020, 261: 114430.

[53] Özbek E, Yalin G, Ekici S, et al. Evaluation of design methodology, limitations, and iterations of a hydrogen fuelled hybrid fuel cell mini UAV[J]. Energy, 2020, 213: 118757.

[54] Erdinc O, Uzunoglu M. Recent trends in PEM fuel cell-powered hybrid systems: investigation of application areas, design architectures and energy management approaches[J]. Renewable and Sustainable Energy Reviews, 2010, 14(9): 2874-2884.

[55] Vural B, Dusmez S, Uzunoglu M, et al. Fuel consumption comparison of different battery/ultracapacitor hybridization topologies for fuel-cell vehicles on a test bench[J]. IEEE Journal of Emerging and Selected Topics in Power Electronics, 2014, 2(3): 552-561.

[56] Yang X B, Pei X. Hybrid system for powering unmanned aerial vehicles: demonstration and study cases[M]//Pioro l. Hybrid technologies for power generation. Amsterdam: Elsevier, 2022: 439-473.

[57] 新华网. 固态、液态、半固态……一文了解锂电池[EB/OL]. (2025-10-23)[2026-05-20]. https://www.xinhuanet.com/finance/20251023/053b7908c3304ee294f177c53bf0ec97/c.html.
Xinhuanet. Solid, liquid, semi-solid… an article to understand lithium batteries[EB/OL]. (2025-10-23)[2026-05-20]. https://www.xinhuanet.com/finance/20251023/053b7908c3304ee294f177c53bf0ec97/c.html(in Chinese).

[58] 毛志亮, 彭顺东. 深圳首飞!氢能+固态电池混动无人机续航飙升[EB/OL]. (2025-12-31)[2026-05-20]. https://www.cinn.cn/difang/guangdong/2025/12-31/71eBL7Jr.html.
Mao Z L, Peng S D. Maiden flight in Shenzhen! Hydrogen energy + solid-state battery hybrid UAV sees surging endurance[EB/OL]. (2025-12-31)[2026-05-20]. https://www.cinn.cn/difang/guangdong/2025/12-31/71eBL7Jr.html(in Chinese).

[59] 高怡宁, 颉文庆, 方淳, 等. 燃料电池无人机关键技术与混合能量架构[J]. 航空学报, 2026, 47(11): 168-194.
Gao Y N, Xie W Q, Fang C, et al. Key technologies of fuel cell unmanned aerial vehicles and hybrid energy architecture[J]. Acta Aeronautica et Astronautica Sinica, 2026, 47(11): 168-194(in Chinese).

[60] Kang K M, Ko Y S, Lee Y S, et al. The operation method of hybrid power supply system combining lithium polymer battery and supercapacitor for industrial drones[J]. Energies, 2023, 16(22): 7552.

[61] Gollnow M. Passenger aircraft towards zero emission with hydrogen and fuel cells[D]. Hambury: Hamburg University of Applied Sciences, 2022: 23-26.

[62] Chen Y Z, Zhao S L, Ma H J, et al. Analysis of hydrogen embrittlement on aluminum alloys for vehicle-mounted hydrogen storage tanks: a review[J]. Metals, 2021, 11(8): 1303.

[63] Niste V B, Tanaka H, Ratoi M, et al. WS_2 nanoadditized lubricant for applications affected by hydrogen embrittlement[J]. RSC Advances, 2015, 5(51): 40678-40687.

[64] Tarhan C, Ali Çil M. A study on hydrogen, the clean energy of the future: hydrogen storage methods[J]. Journal of Energy Storage, 2021, 40: 102676.

[65] Moradi R, Groth K M. Hydrogen storage and delivery: review of the state of the art technologies and risk and reliability analysis[J]. International Journal of Hydrogen Energy, 2019, 44(23): 12254-12269.

[66] Stroman R O, Schuette M W, Swider-Lyons K, et al. Liquid hydrogen fuel system design and demonstration in a small long endurance air vehicle[J]. International Journal of Hydrogen Energy, 2014,

39(21): 11279-11290.

[67] Kang D, Yun S, Kim B K. Review of the liquid hydrogen storage tank and insulation system for the high-power locomotive[J]. Energies, 2022, 15(12): 4357.

[68] 王婧禹, 王凯. 航天科技六院氢能四大新品重磅发布[EB/OL]. (2026-03-20)[2026-05-21]. https://www.spacechina.com/n25/n2014789/n2014809/c4570205/content.html.
Wang J Y, Wang K. Sixth Academy of CASC officially releases four major new hydrogen energy products[EB/OL]. (2026-03-20)[2026-05-21]. https://www.spacechina.com/n25/n2014789/n2014809/c4570205/content.html(in Chinese).

[69] 航天节能环保. 化“氢”为力, 随“撬”而动 | 中国航天 Aero-S10™ 移动式液氢加注撬, 开启灵活加注新纪元[EB/OL]. (2026-03-31)[2026-05-21]. https://mp.weixin.qq.com/s?__biz=Mzg3ODQ2NDE2MA==&mid=2247487752&idx=1&sn=52c7845733e2de55f77b4fe63aef5f5c&chksm=ced5c3192012557bbac96b97789c764ee28f40d65e9510cf98ddac7ff3a74b04674b842b4b7e#rd.
Hangtian Energy Conservation and Environmental Protection. Transforming "hydrogen" into power, moving with the "skid" | China Aerospace Aero-S10™ mobile liquid hydrogen refueling skid opens a new era of flexible refueling[EB/OL]. (2026-03-31) [2026-05-21]. https://mp.weixin.qq.com/s?__biz=Mzg3ODQ2NDE2MA==&mid=2247487752&idx=1&sn=52c7845733e2de55f77b4fe63aef5f5c&chksm=ced5c3192012557bbac96b97789c764ee28f40d65e9510cf98ddac7ff3a74b04674b842b4b7e#rd(in Chinese).

[70] Osborn W, Markmaitree T, Shaw L L, et al. Solid-state hydrogen storage: storage capacity, thermodynamics, and kinetics[J]. JOM, 2009, 61(4): 45-51.

[71] Khafidz N Z A K, Yaakob Z, Lim K L, et al. The kinetics of light-weight solid-state hydrogen storage materials: a review[J]. International Journal of Hydrogen Energy, 2016, 41(30): 13131-13151.

[72] Thomas K M. Adsorption and desorption of hydrogen on metal-organic framework materials for storage applications: comparison with other nanoporous materials[J]. Dalton Transactions, 2009(9): 1487.

[73] Kim H, Oh T H, Kwon S. Simple catalyst bed sizing of a NaBH 4 hydrogen generator with fast startup for small unmanned aerial vehicles[J]. International Journal of Hydrogen Energy, 2016, 41(2): 1018-1026.

[74] Kwon S M, Kim M J, Kang S, et al. Development of a high-storage-density hydrogen generator using solid-state $NaBH_4$ as a hydrogen source for unmanned aerial vehicles[J]. Applied Energy, 2019, 251: 113331.

[75] Lapeña-Rey N, Blanco J A, Ferreyra E, et al. A fuel cell powered unmanned aerial vehicle for low altitude surveillance missions[J]. International Journal of Hydrogen Energy, 2017, 42(10): 6926-6940.

[76] Wang S S, Zhang W J, Yang L P, et al. A gas-solid hydride ion battery and a hydrogen-electricity co-storage system[J]. Joule, 2026: 102475.

[77] Zhang F Q, Wang L H, Coskun S, et al. Energy management strategies for hybrid electric vehicles: review, classification, comparison, and outlook[J]. Energies, 2020, 13(13): 3352.

[78] Yan Y Z, Wang B, Wang C H, et al. Novel synergistic framework design and optimization for hydrogen-electric hybrid power system in unmanned aerial vehicle applications[J]. IEEE Transactions on Transportation Electrification, 2025, 11(1): 337-347.

[79] Yang Z, Lei T, Lin Z C, et al. The testing platform of hybrid electric power system for a fuel cell unmanned aerial vehicle[C]//2018 IEEE International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles & International Transportation Electrification Conference (ESARS-ITEC). Piscataway: IEEE Press, 2018: 1-8.

[80] Zhang X H, Liu L, Dai Y L, et al. Experimental investigation on the online fuzzy energy management of hybrid fuel cell/battery power system for UAVs[J]. International Journal of Hydrogen Energy, 2018, 43(21): 10094-10103.

[81] Shen Y, Cui P F, Wang X C, et al. Variable structure battery-based fuel cell hybrid power system and its incremental fuzzy logic energy management strategy[J]. International Journal of Hydrogen Energy, 2020, 45(21): 12130-12142.

[82] Waseem M, Amir M, Lakshmi G S, et al. Fuel cell-based hybrid electric vehicles: an integrated review of current status, key challenges, recommended policies, and future prospects[J]. Green Energy and Intelligent Transportation, 2023, 2(6): 100121.

[83] Sánchez M, Delprat S, Hofman T. Energy management of hybrid vehicles with state constraints: a penalty and implicit Hamiltonian minimization approach[J]. Applied Energy, 2020, 260: 114149.

[84] Lei Z Z, Qin D T, Hou L L, et al. An adaptive equivalent consumption minimization strategy for plug-in hybrid electric vehicles based on traffic information[J]. Energy, 2020, 190: 116409.

[85] Vazquez S, Leon J I, Franquelo L G, et al. Model predictive control: a review of its applications in power electronics[J]. IEEE Industrial Electronics Magazine, 2014, 8(1): 16-31.

[86] Bradley T H. Modeling, design and energy management of fuel cell systems for aircraft[D]. Atlanta: Georgia Institute of Technology, 2008.

[87] 戴月领, 刘莉, 张晓辉. 燃料电池无人机动力系统半实物仿真[J]. 北京航空航天大学学报, 2020, 46(2): 439-446.
Dai Y L, Liu L, Zhang X H. Hardware-in-the-loop simulation of fuel cell UAV power system[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(2): 439-446(in Chinese).

[88] Muñoz P M, Correa G, Gaudiano M E, et al. Energy management control design for fuel cell hybrid electric vehicles using neural networks[J]. International Journal of Hydrogen Energy, 2017, 42(48): 28932-28944.

[89] Chen Q, Zhang G B, Zhang X Z, et al. Thermal management of polymer electrolyte membrane fuel cells: a review of cooling methods, material properties, and durability[J]. Applied Energy, 2021, 286: 116496.

[90] Afshari E, Asghari S, Jahantigh N, et al. Recent advancements and prospects of thermal management strategies in polymer electrolyte membrane (PEM) fuel cells[M]//Aloui F, Varuvel E G, Sonthalia A, ed. Handbook of thermal management systems. Amsterdam: Elsevier, 2023: 417-440.

[91] Meng H R, Yu X X, Luo X B, et al. Modelling and operation characteristics of air-cooled PEMFC with metallic bipolar plate used in unmanned aerial vehicle[J]. Energy, 2024, 300: 131559.

[92] van Heerden A S J, Judt D M, Jafari S, et al. Aircraft thermal management: practices, technology, system architectures, future challenges, and opportunities[J]. Progress in Aerospace Sciences, 2022,

128: 100767.

[93] Li Q H, Liu Z Q, Sun Y, et al. A review on temperature control of proton exchange membrane fuel cells[J]. Processes, 2021, 9(2): 235.

[94] He L G, Yang Y Y, Zhang Y, et al. A review of thermal management of proton exchange membrane fuel cell systems[J]. Journal of Renewable and Sustainable Energy, 2023, 15: 012703.

[95] Yuan W W, Ou K, Kim Y B. Thermal management for an air coolant system of a proton exchange membrane fuel cell using heat distribution optimization[J]. Applied Thermal Engineering, 2020, 167: 114715.

[96] Wang Y X, Chen Q, Zhang J, et al. Real-time power optimization for an air-coolant proton exchange membrane fuel cell based on active temperature control[J]. Energy, 2021, 220: 119497.

[97] Li J W, Li Y P, Yu T. An optimal coordinated proton exchange membrane fuel cell heat management method based on large-scale multi-agent deep reinforcement learning[J]. Energy Reports, 2021, 7: 6054-6068.

[98] 赵洪山, 潘思潮, 马利波, 等. 基于分类回放双延迟贝叶斯深度确定性策略梯度的燃料电池温度控制[J]. 电工技术学报, 2024, 39(13): 4240-4256.
Zhao H S, Pan S C, Ma L B, et al. Control of fuel cell temperature based on classified replay twin delayed Bayesian deep deterministic policy gradient[J]. Transactions of China Electrotechnical Society, 2024, 39(13): 4240-4256(in Chinese).

Research progress and key technologies of hydrogen fuel cell UAVs

XU Weiqiang[1, 2, *], ZHANG Bingxin[2], ZHU He[1], CHEN Li[3]

(1. Tianmushan Laboratory, Hangzhou 311115, China;
2. School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China;
3. School of Energy and Power Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: Due to their high energy density and zero emissions, hydrogen fuel cell UAVs have become a vital means of overcoming the endurance limitations of traditional lithium-battery-powered UAVs in the context of a growing low-altitude economy and the development of green aviation. This paper provides a comprehensive review and systematic categorization of the development and key technologies of hydrogen-fuel cell UAVs, along with a reasonable outlook on their future evolution, to serve as a valuable reference for related research. In particular, the paper starts with a summary of the scientific advancements and developmental history of hydrogen-powered aircraft, the inherent benefits of hydrogen fuel cell powertrain systems, and the state of the art in hydrogen fuel cell unmanned aerial vehicles. Subsequently, a systematic analysis is conducted on the core technologies involved in hydrogen fuel cell UAVs, including hydrogen fuel cell technology, multi-source hybrid energy systems, airborne hydrogen storage technology, energy management strategies, and thermal management strategies. Lastly, a summary of the study's significant findings is provided, and future research directions are indicated by matching the state of these important technologies' research with the real-world difficulties they face.

Keywords: unmanned aerial vehicles; hydrogen fuel cells; hybrid power; on-board hydrogen storage; energy management; thermal management

Received: 2026-06-29; **Accepted**: 2026-08-21; **Published Online**: 2026-08-28 16:38

URL: link.cnki.net/urlid/11.2625.V.20260828.1516.003

Foundation items: Tianmushan Laboratory Key Research Project (TK-2024-D-011,TK-2025-D-003)

*** Corresponding author**. E-mail: 09430@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2026.0377

基于自适应加权优先虚拟自博弈的多机协同空战决策算法

周尧明[1,*]，韩韬[1]，姜雨菲[1]，杨政[2]，田艺枫[2]

(1. 北京航空航天大学 航空科学与工程学院，北京 100191；2. 中国电子科技集团公司第二十九研究所，成都 610036)

摘 要：针对多机协同超视距(BVR)空战自博弈训练中收敛速度慢、策略循环振荡的问题，提出了一种结合优先虚拟自博弈(PFSP)与自适应胜局加权机制的多机协同空战决策算法。在多智能体强化学习框架下构建4v4超视距空战仿真对抗环境；引入Elo机制动态评估策略能力，实现历史对手策略的非均匀采样，打破策略循环；同时，设计基于胜率的损失函数加权机制，在初期高价值样本稀缺情况下，引导梯度更新方向。实验结果表明，所提算法在第200个训练回合便逼近平衡状态，较传统FSP算法的收敛效率提升约55%，且中后期的博弈振荡幅值相对较小。进一步采用消融实验证实了损失函数加权机制的加速收敛效果。采用PFSP算法的自博弈智能体策略在面对专家对手时，能够取得70.3%的胜率，展现出强大的空战决策能力。

关 键 词：空战决策；自博弈；强化学习；胜局加权；多机协同

中图分类号：V279

文献标志码：A **文章编号**：1001-5965(2026)10-3580-12

制空权是现代信息化战争中的关键，掌握制空权在一定程度上意味着获得了战争的主动权[1]，随着智能技术的日趋成熟，空战正朝着分布化、协同化、无人化、智能化的方向发展，与传统空战相比，未来空战的战争形态和面貌正在发生深刻变化。其中，多架高性能战斗机可在数据链的支持下协调配合，进行超视距(beyond-visual-range, BVR)空战。超视距空战也被称为中远距空战，是指基于中远距空空导弹，交战双方飞行员在目视观察范围以外进行的空战，与近距离空战相比，主要区别在于作战距离和作战方式[2]。随着战场环境的复杂化，飞行员面临态势感知压力激增与决策窗口缩短的挑战，因此，将无人机与先进的自主智能决策技术相结合具有重要意义。

近年来，深度强化学习(deep reinforcement learning, DRL)凭借其在处理高维连续状态空间和非线性决策问题上的卓越表现，在空战决策领域得到了广泛应用[3-5]。相比于传统的专家经验决策算法，DRL能够通过与环境的交互反馈获得奖励，不断自主优化策略。当前，较为常用的DRL算法包括深度Q网络(deep Q-network, DQN)、近端策略优化(proximal policy optimization, PPO)、深度确定性策略梯度(deep deterministic policy gradient, DDPG)等。

Li等[6]针对信息缺失问题，将Transformer网络引入DDPG算法的Actor-Critic框架，有效提升了部分可观测环境下的决策胜率与稳定性。Yang等[7]提出了耦合奖励机制，通过将角度与距离奖励相乘而非传统加权相加，使二者相互制约，在动态场景中展现出更优的收敛速度与鲁棒性。Wang等[8]提出将双向长短期记忆网络(bidirectional long short-term memory, BiLSTM)与多头自注意力机制融合

收稿日期：2026-06-09；录用日期：2026-07-17；网络出版时间：2026-08-06 14:42
网络出版地址：link.cnki.net/urlid/11.2625.V.20260806.1112.002
*通信作者. E-mail：zhouyaoming@buaa.edu.cn

引用格式：周尧明，韩韬，姜雨菲，等. 基于自适应加权优先虚拟自博弈的多机协同空战决策算法[J]. 北京航空航天大学学报，2026，52(10)：3580-3591. Zhou Y M，Han T，Jiang Y F，et al. A multi-aircraft cooperative air combat decision-making algorithm based on an adaptive weighted priority virtual self-play theory approach[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3580-3591 (in Chinese).

以增强态势感知, 实现了在复杂对抗中更高的平均回报与胜率。Mei 等[9] 提出一种双层分层决策框架, 利用门控循环单元(gated recurrent unit, GRU)的网络记忆能力补偿部分观测信息, 并结合自博弈课程学习, 显著提升了信息缺失环境下的决策鲁棒性与胜率。Zhang 等[10] 构建了融合规则、子目标与动态态势评估的复合奖励范式, 有效平衡了奖励信号的引导, 提升了训练效率与最终性能。Zhu 等[11] 通过引入中断机制与分层动作空间, 使智能体能够快速响应突发威胁并组合复杂机动, 结合高保真分布式仿真, 实现了对专家规则系统的战术超越。Gao 等[12] 通过改进 TD3 算法并引入场景迁移训练, 大幅提升了训练效率, 所生成的近距机动策略与飞行员经验高度吻合。Shi 等[13] 提出了融合态势感知、决策与控制的完整框架, 其创新的优先航向轮询随机观测权重(priority heading polling random observation weight, PHP-ROW)方法在六自由度模型上有效解决了超视距博弈中机动性与稳定性的平衡难题。

自博弈机制可以与 DRL 结合, 在不需要人类先验知识设计专家对手的条件下, 通过“左右互搏”的方式不断进化, 从而自发涌现出超越人类预设逻辑的高级战术[14]。Li 等[15] 将并行自博弈与软演员-评论家(soft Actor-Critic, SAC)算法结合, 通过多环境中的样本共享和策略共享提升了无人机机动决策的训练效率与泛化能力; Kong 等[16] 面向双机编队近距空战构建分层决策架构, 并利用竞争式自博弈联合训练主策略与子策略, 增强了策略的对称性与鲁棒性。在复杂观测与高保真验证方面, Tasbas 等[17] 将竞争式自博弈与状态堆叠结合, 提升了智能体在传感器噪声条件下的胜率与稳健性; 单圣哲等[14] 采用基于群落的自博弈训练连续动作空战智能体, 使其相对单智能体训练模型取得 70% 以上胜率, 并涌现出类似人类“单环/双环”的战术。

尽管上述研究验证了自博弈在生成强对手、促进战术涌现和提升策略鲁棒性方面的潜力, 现有工作仍主要集中于近距一对一或双机编队场景, 对多机协同超视距空战关注不足。将自博弈机制直接应用于多机超视距空战环境时, 仍面临以下 2 个问题: ①对手策略持续变化使训练环境呈现非平稳性, 超视距空战策略库中往往存在类似“剪刀-石头-布”的循环克制关系, 导致传统自博弈在采样对战对手时容易陷入性能振荡, 无法保证策略能力持续进化, 甚至出现灾难性遗忘; ②自博弈要求在无人类专家作为对手指导下自发进化, 导致在训练初期智能体成长极为缓慢, 高价值胜利样本较为稀缺, 易被大量的无意义经验样本所淹没, 最终使得收敛缓慢甚至停滞。

针对上述问题, 本文在多智能体 PPO (multi-agent PPO, MAPPO)算法的基础上, 引入优先虚拟自博弈(priority fictitious self-play, PFSP)机制, 并设计了优势经验损失加权的算法改进, 建立了无人机自博弈决策模型。本文主要创新如下:

1) 基于 Elo 机制的 PFSP。本文摒弃了传统的均匀采样策略, 在策略池中引入 Elo 评分系统与优先选择机制, 通过动态评估历史策略的绝对实力, PFSP 能够在“探索”与“利用”之间实现自适应平衡, 打破策略循环, 保证模型能力的持续进化。

2) 自适应胜率优势经验损失加权的 MAPPO 算法。本文提出了一种新型的梯度加权策略, 根据当前对局胜率动态调整不同样本在损失函数中的权重, 该机制能够显著增强高价值经验对策略更新的指引作用, 大幅提升样本利用率并加速收敛。

1　超视距空战环境建模

为真实刻画超视距空战中飞机的机动特性, 本文采用六自由度刚体动力学模型对飞机进行建模。该模型能够同时描述飞机在三维空间中的平动与转动行为, 相较于三自由度飞行运动, 六自由度这种更高维度的机动模型能够更准确地模拟无人机的姿态变化, 更接近无人机的真实飞行[18]。

1.1　六自由度无人机机动模型

本文动作控制采用标准飞行控制分层结构, 上层模块将期望高度、期望速度、期望航向角转换为姿态角和推力指令, 下层控制模块再将姿态角与推力转换为舵面偏角与油门, 实现了飞行器稳定控制。本文假设无人机为左右对称的理想刚体, 实验中所处空域的重力加速度、空气密度和无人机自身质量始终保持不变。图 1 为飞机模型示例。

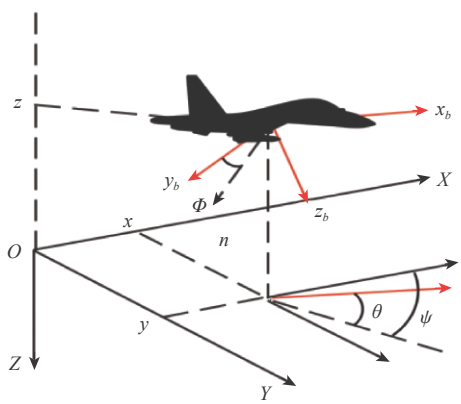


图 1　飞机模型

Fig. 1　Aircraft model

六自由度飞机的状态量可由如下非线性运动方程控制：

$$\begin{cases}\dot{x}=V\cos\gamma_{\mathrm{f}}\cos\chi\\ \dot{y}=V\cos\gamma_{\mathrm{f}}\sin\chi\\ \dot{z}=-V\sin\gamma_{\mathrm{f}}\\ \dot{\Phi}=p+\tan\theta(u\cos\Phi+q\sin\Phi)\\ \dot{\theta}=q\cos\Phi-u\sin\Phi\\ \dot{\psi}=(u\cos\Phi+q\sin\Phi)/\cos\theta\\ \dot{V}=\dfrac{F_T\cos\alpha-F_D}{m}-g\sin\gamma_{\mathrm{f}}\\ \dot{\gamma}_{\mathrm{f}}=\dfrac{(L+F_T\sin\alpha)\cos u}{mV}-\dfrac{g}{V}\cos\gamma_{\mathrm{f}}\end{cases}\tag{1}$$

式中：V为飞机的速度；$[x,y,z]$为飞机在三维空间中的坐标；α为飞机的迎角；γ_{f}为飞机的航迹倾角；χ为飞机的航迹偏角；$[\Phi,\theta,\psi]$分别表示飞机的滚转角、俯仰角和偏航角；$[p,q,u]$分别表示滚转角速度、俯仰角速度、偏航角速度；F_T为推力；F_D为阻力；g为重力加速度；m为飞行器质量；L为升力。

1.2 机载雷达建模

为描述超视距空战中无人机对目标的搜索与跟踪过程，本文建立机载雷达探测模型。以一架我方战机和一架敌方战机为例，位置坐标分别为$\boldsymbol{P}_A=(x_A,y_A,z_A)$和$\boldsymbol{P}_B=(x_B,y_B,z_B)$，可以得到敌方战机$B$相对于我方战机$A$的相对位置$\boldsymbol{P}_{AB}=(x_B-x_A,y_B-y_A,z_B-z_A)$。

1）雷达距离探测约束。由此可以得出二者之间的距离为

$$d_{AB}=\|\boldsymbol{P}_{AB}\|=\sqrt{(x_B-x_A)^2+(y_B-y_A)^2+(z_B-z_A)^2}\tag{2}$$

本文中，机载火控雷达采用基本雷达方程进行建模：

$$d_{\max}=\left[\frac{P_{\mathrm{tx}}G^2\xi}{(4\pi)^3P_{\min}}\right]^{1/4}\tag{3}$$

式中：P_{tx}为雷达发射功率；G为雷达天线的增益；ξ为目标雷达截面积；$P_{\min}$为雷达最小可检测信号；$d_{\max}$为雷达对于雷达截面积为ξ的目标最大作用距离。

则得到目标满足距离探测条件：

$$d_{AB}\leqslant d_{\max}\tag{4}$$

2）雷达方位角探测约束。设本机速度方向单位矢量为

$$\boldsymbol{e}_v=\frac{\boldsymbol{V}_A}{\|\boldsymbol{V}_A\|}\tag{5}$$

目标相对方向单位矢量为

$$\boldsymbol{e}_r=\frac{\boldsymbol{P}_{AB}}{\|\boldsymbol{P}_{AB}\|}\tag{6}$$

可以得到目标相对方位角为

$$\lambda_{AB}=\arccos(\boldsymbol{e}_v\boldsymbol{e}_r)\tag{7}$$

则得到目标满足雷达方位角探测条件：

$$|\lambda_{AB}|\leqslant\lambda_{\max}\tag{8}$$

式中：$\lambda_{\max}$为雷达水平搜索半视场角。

3）雷达俯仰角探测约束。目标相对俯仰角定义为

$$\theta_{AB}=\arctan\left(\frac{z_B-z_A}{\sqrt{(x_B-x_A)^2+(y_B-y_A)^2}}\right)\tag{9}$$

设雷达俯仰搜索半视场角为$\theta_{\max}$，则得到目标满足雷达俯仰角探测条件：

$$|\theta_{AB}|\leqslant\theta_{\max}\tag{10}$$

综合距离、方位和俯仰约束，目标探测判定函数定义为

$$I_{AB}=\begin{cases}1 & d_{AB}\leqslant d_{\max},|\lambda_{AB}|\leqslant\lambda_{\max},|\theta_{AB}|\leqslant\theta_{\max}\\ 0 & \text{其他}\end{cases}\tag{11}$$

$I_{AB}=1$表示成功探测目标，$I_{AB}=0$表示不能探测到目标。当目标满足探测条件时，本机可获得目标的距离、方位及速度等局部信息，并更新目标态势信息；否则，目标信息置为空。

1.3 导弹交战规则模型

1.3.1 导弹运动学建模

本文选用AIM-120D空空导弹并建立运动学模型。建立以导弹重心O_M为原点的弹体坐标系$O_Mx_My_Mz_M$。其中，O_Mx_M位于导弹对称平面内，平行于弹身轴线，指向弹头方向；O_Mz_M轴位于导弹的对称面内，垂直于O_Mx_M轴且方向向下；O_My_M轴垂直于战机的对称面，遵循右手定则方向向右。

导弹的运动学模型可表示为

$$\begin{cases}\dot{x}_M=V_M\cos\sigma\cos\varepsilon\\ \dot{y}_M=V_M\cos\sigma\sin\varepsilon\\ \dot{z}_M=-V_M\sin\sigma\end{cases}\tag{12}$$

式中：(x_M,y_M,z_M)为导弹的空间坐标；V_M为导弹的速度大小；σ和ε分别为导弹的弹道倾角和弹道偏角，其中，弹道倾角σ为弹体轴O_Mx_M与水平面$O_gx_gy_g$之间的夹角，规定弹头上扬时为正，弹道偏角ε为弹体轴O_Mx_M在水平面$O_gx_gy_g$上的投影与O_gx_g轴之间的夹角，规定导弹右偏航时为正。

1.3.2 基于专家规则的导弹发射决策模型

在超视距空战仿真建模中，导弹发射不仅依赖于底层的导弹动力学包线，更受到战术准则、战场态势感知及火力资源配置等多重约束。本文基于专家系统框架，构建了一套多维度的导弹发射逻辑模型。该模型在无人机进入攻击阶段后，通过对感知目标列表进行条件过滤与动态分配，最终输出开火指令及导弹协同参数。

1) 传感器跟踪精度约束。导弹的火控解算高度依赖于机载传感器(雷达/红外搜索与跟踪(infrared search and track, IRST)系统)对目标的跟踪精度。设定目标的距离误差标准差为σ_{R}, 方位角误差标准差为σ_{Az}, 俯仰角误差标准差为σ_{El}。只有当$(\sigma_{\mathrm{R}} \leqslant \sigma_{\mathrm{R0}}) \cap (\sigma_{\mathrm{Az}} \leqslant \sigma_{\mathrm{Az0}}) \cap (\sigma_{\mathrm{El}} \leqslant \sigma_{\mathrm{El0}})$时, 才允许发射导弹, 其中, σ_{R0}、σ_{Az0}、σ_{El0}分别为对应导弹允许的最大误差阈值。

2) 离轴角发射约束。为确保导弹发射后中继数据链的稳定传输, 对发射时刻的离轴角λ进行约束, 离轴角定义为我方载机机体纵轴(机头指向)与弹目视线之间的夹角。根据专家经验, 设置最大离轴角为30°。当$\lambda \leqslant 30°$时, 认为敌机在机载火控雷达的照射扇区或导弹导引头的最大搜索视场内, 允许导弹发射。

3) 防误伤机制。当敌我机群发生近距离混战时, 主动雷达弹易发生误伤。遍历我方无人机的位置并计算目标与最近我方战机的距离, 如果距离过近, 则为避免误伤友机, 禁止发射。

4) 飞机姿态与发射时间间隔。导弹的超视距有效射程高度依赖于其发射时的初始动力学条件, 当载机处于过大负俯仰角状态下发射时, 导弹弹道将迅速指向低空高密度大气区域, 导致空气动力阻力急剧增加, 动能过快耗散, 进而压缩导弹的动力学杀伤包线。定义发射最低俯仰角$\theta_{\min} = -5°$, 以保证导弹获得平飞或跃升初始姿态。

此外, 对最小发射时间间隔做出限制, 预定2枚导弹之间的最小时间间隔为0.5 s, 以符合实际空战场景。

2　基于优先虚拟自博弈的无人机集群博弈对抗算法

在多无人机集群超视距对抗任务中, 智能体之间的博弈关系呈现出典型的强对抗性与非平稳性特征。随着策略的持续更新, 环境动态分布不断变化, 使得强化学习过程面临显著的训练不稳定问题。此外, 超视距空战奖励稀疏, 加剧了策略收敛的难度。本文以4v4超视距空战为例, 设计了基于Elo评分的PFSP机制, 并引入胜率驱动的自适应损失加权算法, 引导梯度更新方向。

2.1　基于 Elo 评分的优先虚拟自博弈机制

在多智能体强化学习中, 传统的虚拟自博弈(fictitious self-play, FSP)通常采用均匀采样的方式从历史策略池中抽取对手进行训练。然而, 这种方式忽略了不同历史策略对当前智能体能力提升的差异性, 容易导致训练效率低下。为解决这一问题, 本文引入了PFSP机制。该机制的核心思想是: 通过对策略池中的历史策略进行能力评估, 并基于评估结果进行非均匀采样, 从而优先选择对当前策略具有较高学习价值的对手参与训练。为量化策略库中各策略的真实水平并计算抽取概率, 本文引入了动态Elo等级分系统。

2.1.1　Elo 评分机制

Elo等级分机制的核心原理是: 为每个参与博弈的策略赋予一个初始分数, 并通过与其他策略的实际对弈结果(胜、负、平)来动态调整该分数。分数的变动幅度不仅取决于实际对局结果, 还取决于双方对弈前的实力差距。本文将策略池中策略的初始Elo评分统一设置为1 200分。对于训练过程中达到入池条件的新策略, 其Elo分数由其在进入策略池前的对抗评估过程决定。具体而言, 候选策略在满足入池条件之前, 需要与策略池中的已有策略进行多轮博弈, 并根据每轮对抗结果按照Elo更新公式持续调整自身评分。当候选策略经过评估达到预设入池标准后, 将其当前累计更新后的Elo评分作为加入策略池时的初始评分。该设计能够保证新加入策略的评分与其实际训练表现保持一致, 其入池的Elo评分能够反映该策略相对于已有策略的实际表现。

设当前策略与对手策略的Elo评分分别为E_i和E_j, 则当前策略的期望胜率为

$$P = \frac{1}{1 + 10^{(E_j - E_i)/400}} \tag{13}$$

根据对战结果, 对Elo评分进行如下更新:

$$E_{i'} = E_i + k_{\mathrm{Elo}}(s_i - P) \tag{14}$$

式中: s_i表示实际对战结果, 胜出时$s_i = 1$, 平手时$s_i = 0.5$, 失败时$s_i = 0$; k_{Elo}为更新系数, 用于控制评分变化幅度。在训练初期, 为提高评分更新的敏感性, 采用较大的k_{Elo}值; 随着训练进行, 逐步减小k_{Elo}, 以提升评分稳定性。具体而言, 本文采用随训练轮次递减的动态更新策略, 使Elo评分能够兼顾收敛速度与稳定性。

2.1.2　多局对抗下的 Elo 更新

由于单次对抗结果具有较大的随机性, 在实际训练中, 每轮策略评估由多局对战构成。为降低偶然性对评分的影响, 本文采用多局累积更新的方式对Elo评分进行修正, 即在多次对战结果基础上迭代更新评分, 从而获得更加稳定的策略能力估计。

$$\begin{cases} E_1 = E_0 + k_{\mathrm{Elo}}(s_0 - P_0) \\ E_2 = E_1 + k_{\mathrm{Elo}}(s_1 - P_1) \\ \quad \vdots \\ E_n = E_{n-1} + k_{\mathrm{Elo}}(s_{n-1} - P_{n-1}) \end{cases} \tag{15}$$

该机制能够有效平滑单局对战中的随机波动，使得Elo评分更具鲁棒性。

2.1.3 基于Elo的优先对手采样机制

在获取了策略池中所有对手的Elo评分后，PFSP机制需要依据分数制定非均匀的对手采样分布。本文采用基于Softmax函数的概率选取方法，使得得分越高的强大对手被抽取的概率越高。策略池中第i个策略被选为对手的概率P_i^{opp}计算如下：

$$P_i^{\text{opp}} = \mathrm{e}^{(E_i/\tau)} \Big/ \sum_j \mathrm{e}^{(E_j/\tau)} \tag{16}$$

式中：τ为温度系数，用于调节抽取策略的探索与利用平衡。当τ取值较大时，指数分布趋于平缓，各个策略被抽取的概率趋于均等，系统倾向于进行全局探索；当τ取值较小时，系统呈现出“赢者通吃”的贪婪特性，几乎只采样当前策略池中Elo得分最高的精英策略。

综合考虑训练的多样性与难度进阶，本文在实验中将温度系数设置为$\tau = 300$，以确保智能体在主要挑战高水平对手的同时，仍保留与历史不同水平策略交手的概率，有效克服神经网络的灾难性遗忘。

2.2 基于自适应胜局信号的策略损失加权机制的MAPPO算法

多智能体空战对抗任务具有高维连续状态空间、强对抗性及多智能体耦合决策等特点。在该类场景中，智能体不仅需要完成个体机动控制（如速度、高度与航向调节），还需在编队协同与敌我对抗中进行策略博弈，使得环境呈现出显著的非平稳性与部分可观测性。

为了形式化描述该过程，本文将多智能体空战对抗任务建模为一个去中心化部分可观测马尔可夫决策过程。该过程可以由一个元组$\langle S, A, O, R, p_{\text{tr}}, n, \gamma \rangle$表示。其中，$A$为各智能体共享的动作空间；$o_t^i = O(s_t; i)$表示智能体$i$在时刻$t$对于环境状态$s_t \in S$的局部观测；$r_t = R(s_t, a_t)$表示各智能体共享的奖励函数；$p_{\text{tr}}$为状态转移方程，描述了在联合动作$a_t$下环境状态的变化；$\gamma \in [0,1)$为折扣因子；$n$为智能体数量。

针对上述问题，本文采用MAPPO算法作为基础强化学习框架。MAPPO算法基于集中训练、分散执行范式，在训练阶段利用全局信息构建价值函数，在执行阶段仅依赖局部观测进行决策，兼顾了多智能体协同建模能力与实际部署可行性。

然而，在多智能体空战对抗任务中，强化学习仍面临显著的稀疏奖励问题。在训练早期阶段，由于智能体策略尚未收敛，胜局出现频率较低，导致：①有效策略样本（胜局经验）占比较小；②胜局经验在梯度更新中被大量无效样本稀释；③PFSP机制导致强化学习要同时针对策略池中多个策略进行更新，学习过程收敛缓慢。

传统方法通常通过建立经验回放池或奖励重塑缓解该问题，但对于基于on-policy的MAPPO算法，直接使用经验回放池会破坏策略分布一致性。因此，本文引入一种基于胜局信号的策略损失加权机制，在不改变采样分布的前提下，提高关键样本对策略更新的贡献。

2.2.1 MAPPO算法框架

MAPPO算法采用Actor-Critic结构，其核心由分散执行的策略网络与集中式价值网络构成。

每个智能体基于自身局部观测s_i输出动作分布$a_i \sim \pi_\vartheta(a_i|s_i)$。

在训练阶段，引入全局状态S构建集中式价值函数$V_\phi(S)$，该结构能够融合敌我双方所有智能体的状态信息，从而提高价值估计精度。

在MAPPO算法中，策略更新基于如下裁剪目标函数：

$$L^{\text{policy}}(\vartheta) = -E_t\left[\min\left(\rho_t(\vartheta)A_t, \text{clip}(\rho_t(\vartheta), 1-\epsilon, 1+\epsilon)A_t\right)\right] \tag{17}$$

式中：$\rho_t(\vartheta) = \pi_\vartheta(a_t|s_t)/\pi_{\vartheta\text{old}}(a_t|s_t)$为重要性采样比，$\vartheta$为神经网络策略；$A_t$为优势函数；$\epsilon$为裁剪系数。该目标函数本质上通过优势函数$A_t$决定策略更新方向，通过裁剪机制限制更新幅度。

2.2.2 自适应胜局加权机制

为提升胜局样本在策略更新中的影响力，本文引入样本权重w_t对策略损失进行加权，如下：

$$L_{\text{weight}}^{\text{policy}}(\vartheta) = -E_t[w_t \cdot \min(r_t(\vartheta)A_t, \text{clip}(r_t(\vartheta), 1-\epsilon, 1+\epsilon)A_t)] \tag{18}$$

式中：权重w_t根据样本t所属对局结果自适应确定，若$t \in$胜局，则$w_t = w_{\text{win}}(w_{\text{win}} \geqslant 1)$，否则，$w_t = 1$。

该加权机制等价于在梯度层面对胜局样本进行放大：

$$\nabla_\vartheta L_{\text{weight}}^{\text{policy}}(\vartheta) \propto w_t A_t \nabla_\vartheta \ln \pi_\vartheta(a_t|s_t) \tag{19}$$

需要指出的是，本文的胜局损失加权机制并未改变PPO算法中重要性采样比值及clip约束过程，而是在裁剪目标函数计算完成后对不同样本的损失贡献进行调整。PPO算法中的clip机制主要限制新旧策略概率比值的变化范围，以避免单次策略更新幅度过大。而本文引入的权重w_t仅作用于样本损失尺度，对梯度贡献进行重新分配，不影响$r_t(\vartheta)$的计算及裁剪边界。因此，本文提出的自适应胜局加权机制在保证on-policy特性的前提下，增强了胜局经验对策略更新的权重。

2.2.3　动态权重调节策略

为避免固定权重导致过拟合或训练不稳定, 本文进一步引入基于胜率的自适应权重调节机制。设当前训练阶段的胜率为r_{current}, 目标胜率为r_{target}, 定义比值为

$$\eta = \frac{r_{\text{current}}}{r_{\text{target}}} \tag{20}$$

则胜局样本权重w_t定义为

$$w_t = \begin{cases} 1 & r_{\text{current}} \geqslant r_{\text{target}} \\ 1 + (1-\eta)(w_{\max} - 1) & r_{\text{current}} < r_{\text{target}} \end{cases} \tag{21}$$

式中: $w_{\max}$为策略损失加权机制的最大权重, 本文中$w_{\max}$设置为 1.5, 将单个胜局样本的最大梯度放大限制在 50%, 以平衡胜局信号增强与策略更新稳定性。通过引入动态策略损失加权机制, 在低胜率的训练前期放大了胜局经验的策略损失权重, 在高胜率阶段回归标准策略避免过拟合, 最终提高了训练效率和策略收敛速度。

2.2.4　加权机制对策略更新的影响分析

1) 策略梯度估计。设 PPO 算法裁剪目标中的单样本目标项为

$$f_t(\vartheta) = \min\left(\rho_t(\vartheta)A_t, \text{clip}(\rho_t(\vartheta), 1-\epsilon, 1+\epsilon)A_t\right) \tag{22}$$

对于包含B_{batch}个样本的训练批次, 标准 MAPPO 算法的策略梯度估计为

$$\hat{\boldsymbol{g}} = \frac{1}{B_{\text{batch}}} \sum_{t=1}^{B_{\text{batch}}} \nabla_\vartheta f_t(\vartheta) \tag{23}$$

引入胜局权重后, 策略梯度估计变为

$$\hat{\boldsymbol{g}}_w = \frac{1}{B_{\text{batch}}} \sum_{t=1}^{B_{\text{batch}}} w_t \nabla_\vartheta f_t(\vartheta) \tag{24}$$

在每次参数更新过程中, 样本权重由对局结果和更新前统计得到的胜率确定, 因此, 可视为与当前梯度计算相独立的常数, 不产生额外的权重求导项。从梯度期望角度, 有

$$E[\hat{\boldsymbol{g}}_w] = E[\hat{\boldsymbol{g}}] + E[(w_t - 1)\nabla_\vartheta f_t(\vartheta)] \tag{25}$$

第 2 项表示胜局经验产生的附加更新方向。因此, 该方法仍使用当前策略采集的轨迹数据, 保持了数据采集的 on-policy 属性, 但改变了原始 PPO 算法目标中不同轨迹的相对贡献, 从而形成面向胜局经验的重加权策略梯度估计。

2) 训练稳定性分析。自适应加权机制在增强有效信号的同时, 也可能放大胜局样本梯度, 从而增大训练早期的梯度波动。令

$$\boldsymbol{g}_t = \nabla_\vartheta f_t(\vartheta) \tag{26}$$

在近似独立采样条件下, 加权梯度估计的方差满足:

$$\text{Var}(\hat{\boldsymbol{g}}_w) = \frac{1}{B_{\text{batch}}} \text{Var}(w_t \boldsymbol{g}_t) \tag{27}$$

由于$1 \leqslant w_t \leqslant w_{\max}$, 其二阶矩满足:

$$E[\|w_t \boldsymbol{g}_t\|^2] \leqslant w_{\max}^2 E[\|\boldsymbol{g}_t\|^2] \tag{28}$$

由此分析可知, 过大的$w_{\max}$可能导致梯度方差和单次策略更新幅度增加。但本文通过以下 3 种方式控制该影响: ①将权重限制在有限区间$[1, w_{\max}]$, 避免少量胜局样本主导整个批次更新; ②PPO 算法裁剪项限制新旧策略概率比, 降低单次更新过大导致策略退化的风险; ③随着当前策略胜率接近目标胜率, 胜局权重逐步衰减至 1, 使算法在训练中后期回归标准 MAPPO 算法更新。因此, 该机制在可控的方差增幅下提高稀缺胜局信号的有效性, 并通过动态调节避免持续放大带来的训练振荡。

3) 探索能力分析。从探索与利用关系来看, 胜局加权提高了已获得成功回报的动作序列对策略更新的影响, 因此, 会使算法在训练初期更加倾向于利用已发现的有效战术。但当胜率提高后, 胜局权重逐渐下降至 1, 避免算法长期偏向少数早期成功策略, 同时 PFSP 机制通过从历史策略池中采样不同能力水平的对手, 使智能体持续面对多样化对抗策略, 从对手分布层面维持策略探索。因此, 动态胜局加权主要在训练初期增强有效经验利用, 随着胜局权重逐渐下降及多样化对手出现, 探索能力逐渐上升。

整体 PFSP 的伪代码如下。

输入: 初始策略π_0; 空策略池$M = \{\pi_0\}$; 初始 Elo 评分$E_0 = 1\,200$; 最大训练轮数T; 每轮训练的步数N; 温度系数$\tau = 300$; 目标胜率r_{target}; 最大损失权重$w_{\max}$。

输出: 训练收敛后的最优策略π^*。

1.初始化, 设置当前训练智能体$\pi_{\text{curr}} \leftarrow \pi_0$。

2. For 训练轮次$t = 1, 2, \cdots, T$ do

3.优先对手采样(PFSP)。计算策略池M中各个历史策略π_i的采样概率: $P_i^{\text{opp}} = \mathrm{e}^{(E_i/\tau)} \Big/ \sum_j \mathrm{e}^{(E_j/\tau)}$, 根据分布随机抽取对手策略$\pi_{\text{opp}}$。

4.多局对抗评估, 初始化本局累计得分E_0。

For 步数$n = 1, 2, \cdots, N$ do:

在 4v4 超视距空战环境中运行π_{curr}与π_{opp}的对局, 获取对局结果$s_n \in \{1, 0.5, 0\}$(胜、平、负), 收集训练轨迹数据D_n用于策略强化学习更新。

End For

5.自适应胜局信号加权更新。计算当前胜率$r_{\text{current}} = N_{\text{win}}/N$, 计算胜率比$\eta = r_{\text{current}}/r_{\text{target}}$。

计算本次训练中胜局样本权重：$w_t=\begin{cases}1 & r_{\text{current}}\geqslant r_{\text{target}}\\ 1+(1-\eta)(w_{\max}-1) & r_{\text{current}}<r_{\text{target}}\end{cases}$。利用加权目标函数更新策略参数$\vartheta$：$L^{\text{policy}}_{\text{weight}}(\vartheta)=-E_t\left[w_t\cdot\min(\rho_t(\vartheta)A_t,\ \text{clip}(\rho_t(\vartheta),1-\epsilon,1+\epsilon)A_t)\right]$。

6.动态Elo评分更新。计算当前策略相对于对手的期望胜率：$P=\dfrac{1}{1+10^{(E_{\text{opp}}-E_{\text{curr}})/400}}$。根据训练进程调整更新系数$k_{\text{Elo}}$。

For $n=1,2,\cdots,N$:

$E_{\text{curr}}\leftarrow E_{\text{curr}}+k_{\text{Elo}}(s_i-P)$

End For

If 当前胜率$r_{\text{current}}\geqslant$目标胜率$r_{\text{target}}$ And 入池间隔$G_{\text{current}}\geqslant$目标入池间隔$G_{\text{target}}$, then:

将当前策略π_{curr}加入策略池M，并记录当前的E_{curr}

End For

Return 最优策略π^*

整体的代码框架如图2所示。

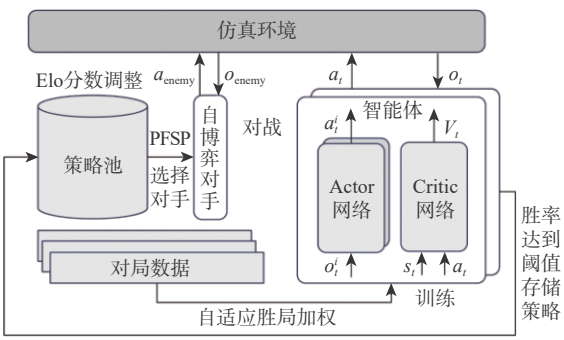


图2 代码框架

Fig. 2 Code framework

3 状态空间与奖励函数设置

3.1 状态空间

根据多机协同超视距空战集中式训练分布式执行架构的特点，分别设计了全局状态空间和局部状态空间，全局状态空间作为集中式训练过程中价值函数的输入，而各智能体的决策只能基于各自的局部观测信息实现。以4v4协同超视距空战为例，我方为蓝方，包括蓝1号机、蓝2号机、蓝3号机、蓝4号机；敌方为红方，包括红1号机、红2号机、红3号机、红4号机。在自博弈情况下，红蓝双方均视为智能体。

3.1.1 全局状态空间设计

本文全局状态空间S设计为8×424大小，每个维度是针对一架飞机的全局状态空间。以蓝1号机为例，其全局状态空间S_1如表1所示。

3.1.2 局部状态空间设计

局部状态空间o设计为8×236大小，每个维度也是针对一架飞机的局部状态空间。以蓝1号机为例，其局部状态空间o_1如表2所示。

3.2 动作空间

本文设计智能体i的动作空间为$a_i=\{h_i,v_i,\psi_i\}$，其中，h_i为期望高度，v_i为期望速度大小，ψ_i为期望航向角。h_i取值为3 000,4 000,5 000,6 000,7 000,8 000,9 000,10 000,11 000,12 000,13 000 m，v_i取值为200,250,300,350,400 m/s，ψ_i取值为0°,45°,90°,135°,180°,225°,270°,315°。动作空间中的期望动作虽为离散值，但战机通过上述期望高度、期望速度、期望角度组成动作对，可以实现在仿真作战环境的自由机动。这是因为战机跟踪期望动作需要时间，因此，即使上层动作指令离散，实际飞行轨迹仍具有连续性和较强的机动灵活性。

3.3 奖励函数

奖励设计是空战中至关重要的一环。本文设计了较为细致的结果奖励和状态奖励，更好地引导无人机根据实际态势进行正确的空战决策，提升无人机作战能力。

3.3.1 状态奖励

状态奖励的目的是缓解只有事件奖励时存在

表1 全局状态空间

Table 1 Global state space

状态分量	描述	维数
本机(蓝1号机)	位置信息，包括x坐标、y坐标、z坐标；	3
	角度信息，包括航向角、俯仰角、滚转角；	3
	速度信息，包括速度大小、北向速度分量、地向速度分量、东向速度分量；	4
	加速度信息，包括北向加速度分量、地向加速度分量、东向加速度分量	3
友机(蓝2号机，蓝3号机，蓝4号机)	与本机类似，但状态信息是相对于本机的相对数据	3×13
敌机(红1号机，红2号机，红3号机，红4号机)	与本机类似，但状态信息是相对于本机的相对数据	4×13
蓝方导弹信息(预留16枚导弹)	相对位置信息(x坐标、y坐标、z坐标)、相对速度信息(速度大小、北向速度分量、地向速度分量、东向速度分量)	16×7
红方导弹信息(预留16枚导弹)	与蓝方导弹信息类似	16×7
导弹告警信息(预留16枚导弹)	方位角、俯仰角(相对值)	16×2
导弹制导信息(预留16枚导弹)	标识符、方位角、俯仰角、制导步数	16×4

表 2　局部状态空间

Table 2　Local state space

状态分量	描述	维数
本机(蓝1号机)	位置信息, 包括x坐标、y坐标、z坐标;	3
	角度信息, 包括航向角、俯仰角、滚转角;	3
	速度信息, 包括速度大小;	1
	加速度信息, 包括北向加速度分量、地向加速度分量、东向加速度分量	3
友机(蓝2号机, 蓝3号机, 蓝4号机)	与本机类似, 但状态信息是相对于本机的相对数据	3×10
敌机(红1号机, 红2号机, 红3号机, 红4号机)	只包括位置信息(x坐标、y坐标、z坐标)、航向角、速度	4×5
蓝方导弹信息(预留8枚导弹)	只包括位置信息(x坐标、y坐标、z坐标)、航向角、速度	8×5
红方导弹信息(预留8枚导弹)	与蓝方导弹信息类似	8×5
导弹告警信息(预留16枚导弹)	方位角、俯仰角(相对值)	16×2
导弹制导信息(预留16枚导弹)	标识符、方位角、俯仰角、制导步数	16×4

的奖励稀疏问题, 引导无人机决策学到进攻、躲弹等决策, 并更多地产生与目标对抗的经验。

1) 探测敌机奖励。为增强智能体的态势感知能力与主动搜索能力, 本文设计了敌方目标探测奖励项。当敌方目标进入己方传感器(如机载雷达)的有效探测范围时, 给予正奖励, 从而引导智能体学习主动占据有利观测位置, 并与敌机产生更多的对抗经验。

具体而言, 设智能体的位置向量$\boldsymbol{P}_i=(x_i,y_i,z_i)$, 敌机的位置向量$\boldsymbol{P}_t=(x_t,y_t,z_t)$, 计算敌机相对我方智能体的方位角$\lambda_{xoy}$和俯仰角$\theta_{xoz}$。当满足以下条件时: $\lambda_{xoy}\in\left[-45°,45°\right]$, $\theta_{xoz}\in[-180°,180°]$, 智能体将得到探测正奖励$r_{\text{detect}}=0.000\,1$。

2) 敌方导弹接近负奖励。为提升智能体在空战环境中的生存能力, 本文在奖励函数中引入了导弹威胁感知惩罚项, 用于表征来袭导弹对自身平台的威胁程度。当敌方导弹进入一定攻击条件范围内时, 对智能体施加负奖励, 以引导其学习规避导弹的机动策略。

具体而言, 设智能体的位置向量$\boldsymbol{P}_i=(x_i,y_i,z_i)$, 导弹的位置向量$\boldsymbol{P}_m=(x_m,y_m,z_m)$, 二者之间的相对位置向量$\boldsymbol{d}=\boldsymbol{P}_i-\boldsymbol{P}_m$, 对应距离$D=\|\boldsymbol{d}\|$, 同时根据导弹的姿态信息计算其速度方向向量(即导弹航向)$\boldsymbol{v}_m$, 进一步定义导弹攻击角为导弹航向与目标方向之间的夹角$\eta_m=\arccos\left(\dfrac{\boldsymbol{d}\cdot\boldsymbol{v}_m}{\|\boldsymbol{d}\|\cdot\|\boldsymbol{v}_m\|}\right)$。当同时满足以下条件时: $|\eta_m|<30°$, $D<18\,\text{km}$, 在该条件下, 智能体将得到一个固定的负奖励: $r_{\text{missile_threat}}=-0.003$。

3) 高度奖励。为引导智能体在不同作战距离下占据合理的高度区间, 从而获得潜在的能量优势与战术机动空间, 本文设计了一种基于敌我距离的自适应高度奖励机制。

首先, 在所有敌方目标中选取与己方距离最近的目标作为主要参考对象, 并计算智能体与其三维距离$d_{\min}=\min\limits_i d_i$。

其次, 为使智能体在不同交战阶段采取不同的高度策略, 定义期望高度h^*并设置距离分段函数:

$$h^*=\begin{cases}h_{\min} & d_{\min}\leqslant d_{\min}^{\text{low}}\\ h_{\max} & d_{\min}\geqslant d_{\min}^{\text{high}}\\ h_{\min}+\dfrac{h_{\max}-h_{\min}}{d_{\min}^{\text{high}}-d_{\min}^{\text{low}}}\left(d_{\min}-d_{\min}^{\text{low}}\right) & \text{其他}\end{cases}\tag{29}$$

式中: $d_{\min}^{\text{low}}=60\,\text{km}$, $d_{\min}^{\text{high}}=100\,\text{km}$, $h_{\min}=6\,\text{km}$, $h_{\max}=10\,\text{km}$。由此可计算距离不同敌人时的期望高度h^*。

定义当前智能体的高度为h_i, 则高度偏差$\Delta h=|h_i-h^*|$。在容差范围: $\Delta h\leqslant\delta_h$, 则给予线性衰减奖励:

$$r_{\text{height}}=\lambda_{\text{height}}(1-\Delta h/\delta_h)\tag{30}$$

否则为 0。其中, 高度容差$\delta_h=2\,\text{km}$, 高度奖励系数$\lambda_{\text{height}}=0.000\,1$。由此可以确定智能体获得的高度正奖励$r_{\text{height}}$。

4) 制导奖励。为提升智能体对己方导弹的态势感知能力与持续制导能力, 本文设计了一种导弹制导奖励机制。

具体而言, 设智能体的位置向量$\boldsymbol{P}_i=(x_i,y_i,z_i)$, 我方已发射导弹的位置向量$\boldsymbol{P}_m=(x_m,y_m,z_m)$, 计算每枚导弹相对我方智能体的方位角$\lambda_{xoy}$和俯仰角$\theta_{xoz}$。当满足以下条件时: $\lambda_{xoy}\in\left[-60°,60°\right]$, $\theta_{xoz}\in\left[-60°,60°\right]$, 导弹位于该前向扇区内时, 认为其处于可制导状态, 智能体将得到探测正奖励$r_{\text{ego}}=0.001\,667$。

5) 置尾逃逸奖励。为提高智能体在空战中的生存能力, 引入尾部逃逸奖励机制。当雷达探测到敌方导弹接近时, 通过计算得到敌方导弹相对飞机的方位角λ_{xoy}, 若导弹位于飞机后方, 即$|\lambda_{xoy}|>90°$,

则认为飞机处于置尾逃逸状态，智能体将得到置尾逃逸正奖励$r_{\text{tail_escape}}=0.005$。

3.3.2 事件奖励

事件奖励通过给超视距空战中涉及的标志性事件单独设计奖励，从而引导无人机在攻击目标的同时，能够躲避导弹，生存下来。主要事件和对应的奖励如表3所示。

表3 事件奖励

Table 3 Event rewards

事件	奖励值
我方飞机发射一枚导弹r_{shoot}	+2
击落一架敌机r_{kill}	+10
我方一架飞机被击落r_{out}	−10
导弹告警消失且我方飞机存活$r_{\text{threat_gone}}$	+0.25

事件奖励和状态奖励构成了总体的奖励函数，每个时间步计算的奖励函数如下：

$$r=r_{\text{detect}}+r_{\text{missile_threat}}+r_{\text{height}}+r_{\text{ego}}+r_{\text{tail_escape}}+r_{\text{shoot}}+r_{\text{kill}}+r_{\text{out}}+r_{\text{threat_gone}} \tag{31}$$

本文奖励函数融合了目标探测、导弹规避、能量管理和持续制导等超视距空战领域先验知识，旨在缓解训练初期奖励稀疏的问题，加快智能体策略的形成。但奖励设计仍具有一定的人工经验依赖，未来将进一步研究更加稀疏和自主的奖励学习方法，以提升算法的探索能力及跨场景泛化性能。

4 仿真结果

为验证本文提出的多机协同超视距空战自博弈决策框架的有效性，在4v4对抗场景下开展仿真实验。首先，通过对比PFSP机制与标准FSP机制，验证优先采样策略在提升博弈稳定性方面的优越性；其次，通过对自适应胜率损失加权机制进行消融实验，量化分析该机制对算法收敛效率的贡献。此外，还引入基于规则的专家对手，旨在评估纯自博弈训练智能体的泛化对抗能力，并对仿真对局中自主涌现的典型战术机动进行深入剖析。

4.1 仿真环境设置

本文在仿真实验环境中设置红蓝双方，我方为蓝机，敌方为红机。自博弈红蓝双方神经网络初始设置均采用随机初始化参数的空白策略，设置多进程并行实验进程组数为4组，实验中每次对战最大时长为1 200 s，智能体控制间隔为0.4 s。对战双方初始距离为140 km，飞机初始高度为9 000 m，初始速度为300 m/s，蓝方飞机初始航向为90°，红方飞机初始航向为270°。在单场对局中，任意一方4架飞机全部阵亡，或者达到最大仿真对战时长，本局对战结束。仿真实验中，敌我双方位置轮流调换，消除地理位置偏差。

此外，多机协同超视距空战自博弈算法的相关参数如表4所示。

表4 算法参数

Table 4 Algorithm parameters

参数分类	参数	取值
神经网络参数	隐藏层神经元数量	256
	隐藏层数量	2
	Actor网络学习率	0.000 5
	Critic 网络学习率	0.000 5
	PPO裁剪参数	0.2
	折扣因子	0.995
	批数量	1
自博弈参数	策略池容量	5
	策略入库胜率阈值	0.6
	策略入库间隔阈值	5
	Elo初始评分	1 200
改进机制参数	策略损失加权机制的最大权重	1.5

4.2 对比实验与消融实验

设计仿真4v4超视距空战。分别进行PFSP自博弈算法与FSP自博弈算法对比实验，以及增加胜局加权机制的PFSP算法与仅PFSP算法的消融实验。以训练时自博弈对局胜利作为评价标准，分析其收敛情况。在对称自博弈演化框架下，博弈双方策略的均衡状态（即逼近平衡状态）表现为胜率曲线在50%基准线附近的收敛与微幅振荡，如图3所示。

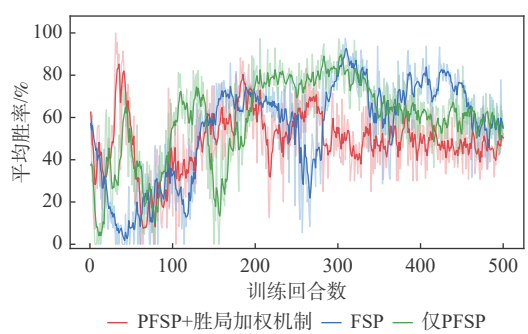


图3 算法对比实验胜率

Fig. 3 Win rates in algorithm comparative experiments

4.2.1 与基准FSP算法的对比实验

首先，从收敛效率维度，本文提出的“PFSP+胜局加权机制”表现出显著的加速收敛优势。在经历训练初期（前150个训练回合）因策略快速试错引起的剧烈激战期后，红线在第200个训练回合数附近便率先达到50%胜率，随后经历一段时间的振荡后进入稳定期，在50%胜率上下微幅波动。相比之

下, 基准 FSP 算法由于缺乏对手针对性筛选与有效的梯度反馈, 其策略更新路径呈现明显的随机性, 直到第 450 个回合附近才初步具备收敛到均衡线的迹象。实验结果表明, 本文算法使算法逼近均衡状态所需的训练回合数缩短了约 55%, 显著降低了多智能体博弈训练的计算成本。

其次, 在博弈稳定性纬度, 标准 FSP 算法在整个训练中后期(200～450 回合)表现出严重的“策略循环振荡”现象, 其平均胜率在 40%～90% 之间产生较大幅度的周期性起伏。这表明在多机协同超视距空战的复杂高动态环境中, 传统 FSP 算法易受到策略库中部分过时或劣质策略的干扰, 导致智能体陷入由于对手改变而频繁推倒重建的策略振荡。而本文算法在 250 个训练回合数后, 其曲线不仅拟合 50% 的平衡状态线, 且其振荡幅值较小, 展现出较好的策略稳定性。

4.2.2　与仅使用 PFSP 算法的消融实验

如图 3 中绿线(仅 PFSP 机制)所示, 当移除胜局加权机制后, 智能体虽然在早期能够凭借优先自博弈提供的合理对手, 避免进入策略循环振荡, 但在训练中期(200～300 回合)由于缺乏合适的梯度指引, 模型总体学习速率较低, 直至 350 个回合后才逐步稳定向均衡线靠拢。对比红线(本文算法)可知, 胜局加权机制大幅度加速了收敛速率, 引导智能体更早地逼近平衡状态。

4.3　对战专家对手实验

为了验证这组收敛后的本文算法模型在面对未知的外部对手时的真实作战能力, 进一步引入基于规则的专家对手进行泛化对抗评估。该专家对手具备“进攻-防御”双重战术态势切换能力, 其核心设计理念为“先发制敌、射后俯冲制导后防御”。具体而言, 专家对手在远距阶段保持雷达常开探测和加速接敌, 当敌机进入杀伤包线后, 优先爬升抢占有利态势, 并实施先手打击; 导弹发射后, 根据导弹制导需求朝向敌方目标飞行, 并通过俯冲、偏置制导提高自身生存能力。当接收到导弹告警时, 策略优先切换至防御机动, 通过转弯背向导弹方向降低高度躲避导弹。

选择训练收敛时的智能体分别对战专家对手。以 3 种算法达到 500 回合时的智能体策略, 分别与专家对手对战 300 场, 得到胜率效果如表 5 所示。

分析可知, 本文提出的结合胜局加权机制的 PFSP 算法在各项空战指标中相对于 FSP 算法和仅 PFSP 算法均有较大领先优势。在与专家对手的 300 场对抗中, 本文算法最终取得了 70.3% 的平均胜率, 大幅度领先 FSP 算法取得的 42.7% 平均胜率和仅 PFSP 算法的 53.3% 平均胜率。同时, 本文算法的平均战损比达到了较为优秀的 1.8:1, 场均我方存活数为 1.9 架, 是 FSP 算法的 2.4 倍。战损比和我方场均存活数表明, 结合胜局加权机制的 PFSP 算法训练出了具有较高决策能力的智能体, 能够在优先保证自身生存的前提下, 通过高超的战术规划击败对手。

表 5　对战专家对手胜率

Table 5　Win rates against expert opponents

决策算法	对战专家胜率/%	战损比(敌: 我)	我方场均存活数/架	平均对局时长/s
PFSP+胜局加权机制	70.3	1.8:1	1.9	455.1
FSP	42.7	0.9:1	0.8	654.2
仅PFSP	53.3	1.2:1	1.3	545.5

4.4　战术分析

以强化学习自博弈对战专家对手为例, 展示多机协同超视距空战决策算法涌现的空战战术。蓝方采用自博弈收敛时的智能体策略, 红方为专家对手, 对战双方初始距离为 140 km, 飞机初始高度为 9 000 m, 初始速度为 300 m/s, 蓝方飞机初始航向为 90°, 红方飞机初始航向为 270°。在单场对局中, 任意一方 4 架飞机全部阵亡, 或者达到最大仿真对战时长, 本局对战结束。

空战场景中红蓝双方对抗三维轨迹如图 4 和图 5 所示。

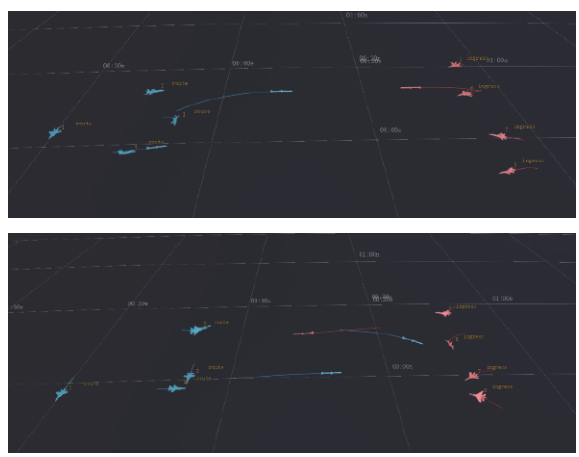


图 4　对战初期仿真表现

Fig. 4　Simulation performance in the initial phase of combat

在对战初期, 蓝 3 号机先视先射发射导弹, 蓝 2 号机和蓝 4 号机采用钳形攻势, 背对飞行, 拉开队内间距。蓝 3 号机完成导弹发射后, 俯冲侧向对敌, 采取防御机动, 此时重新转向敌方的蓝 2 号机和蓝 4 号机对蓝 3 号机发射的导弹进行接力制导。

进一步, 蓝 2 号机和蓝 4 号机同时爬升, 发射高抛弹使用火力压制敌方飞机。第 1 波态势弹有效压制敌方飞机的能量, 敌方红 1 号机和红 3 号机

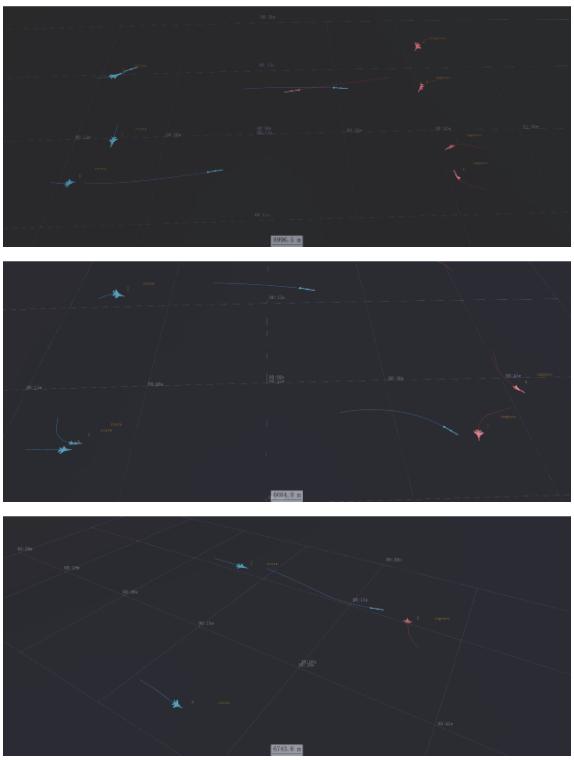


图5 对战中后期仿真表现

Fig. 5 Simulation performance in mid-to-late phases of combat

躲避导弹后，在低空重新进攻，蓝2号机和蓝4号机继续发射导弹进行第2波进攻。第2波进攻距离近，且蓝方占据高度优势，成功命中敌机。进一步对红2号机进行持续压制，通过发射导弹逼迫红方飞机侧向面对己方，从而使得红方飞机导弹失去制导，无法命中我方飞机，最终全歼敌方飞机。

5 结 论

本文针对多机协同超视距空战场景下自博弈训练较慢、策略循环振荡问题，提出了结合胜局加权机制的PFSP算法，通过理论建模和仿真验证，得出如下结论：

1）结合胜局加权算法的PFSP在训练收敛速率、收敛稳定性方面均显著优于基准FSP算法，收敛效率提升约55%。

2）使用本文算法训练的智能体对战专家对手时，能够取得70.3%的胜率，显著优于基准算法42.7%的胜率，展现出强大的空战决策能力。

3）仿真战术分析表明本文所提出的自博弈算法具有优异的表现效果，涌现出了钳形攻势、侧翼包抄等符合人类认知的专家级作战战术。

需要指出的是，本文采用离散化的期望高度、期望速度和期望航向动作，能够有效压缩动作搜索空间，降低多机空战决策中的探索难度，但也可能限制智能体对飞行控制量的精细调节能力，从而损失部分精细机动潜力。未来将进一步研究连续动作空间建模方法，以提升智能体在复杂空战场景下的精细机动与控制能力。

参考文献（References）

[1] 樊会涛, 闫俊. 空战体系的演变及发展趋势[J]. 航空学报, 2022, 43(10): 527397.
Fan H T, Yan J. Evolution and development trend of air combat system[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(10): 527397(in Chinese).

[2] 孙智孝, 杨晟琦, 朴海音, 等. 未来智能空战发展综述[J]. 航空学报, 2021, 42(8): 525799.
Sun Z X, Yang S Q, Piao H Y, et al. A survey of air combat artificial intelligence[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(8): 525799(in Chinese).

[3] 罗越群, 丁达理, 谭目来, 等. 无人作战飞机自主机动决策方法综述[J]. 航空学报, 2025, 46(7): 130884.
Luo Y Q, Ding D L, Tan M L, et al. A review of autonomous maneuver decision methods for unmanned combat aerial vehicle[J]. Acta Aeronautica et Astronautica Sinica, 2025, 46(7): 130884(in Chinese).

[4] 李波, 黄晶益, 万开方, 等. 基于深度强化学习的无人机系统应用研究综述[J]. 战术导弹技术, 2023(1): 58-68.
Li B, Huang J Y, Wan K F, et al. A review of research on the application of UAV system based on deep reinforcement learning[J]. Tactical Missile Technology, 2023(1): 58-68(in Chinese).

[5] 方思雨, 张振华, 牛经龙, 等. 多无人机空战机动决策中的深度强化学习: 实践中的关键技术综述与未来展望[J]. 航空兵器, 2025(2): 19-33.
Fang S Y, Zhang Z H, Niu J L, et al. Deep reinforcement learning in multi-UAV air combat maneuver decision-making: a review of key techniques in practice and future prospects[J]. Aero Weaponry, 2025(2): 19-33(in Chinese).

[6] Li W T, Fang F, Peng D L, et al. An intelligent maneuver decision-making approach for air combat based on deep reinforcement learning and transformer networks[J]. Entropy, 2024, 26(12): 1036.

[7] Yang J, Wang L P, Han J L, et al. An air combat maneuver decision-making approach using coupled reward in deep reinforcement learning[J]. Complex & Intelligent Systems, 2025, 11(8): 364.

[8] Wang H M, Zhou Z F, Jiang J Z, et al. Autonomous air combat maneuver decision-making based on PPO-BWDA[J]. IEEE Access, 2024, 12: 119116-119132.

[9] Mei J R, Li G, Huang H S. Deep reinforcement-learning-based air-combat-maneuver generation framework[J]. Mathematics, 2024, 12(19): 3020.

[10] Zhang X, Dong W H, Zhang P, et al. Research on the reward design method for deep reinforcement learning in WVR air combat[J]. IEEE Access, 2024, 12: 182693-182707.

[11] Zhu J Y, Kuang M C, Zhou W Q, et al. Mastering air combat game with deep reinforcement learning[J]. Defence Technology, 2024, 34: 295-312.

[12] Gao X Z, Zhang Y, Wang B L, et al. The optimal strategies of maneuver decision in air combat of UCAV based on the improved TD3

algorithm[J]. Drones, 2024, 8(9): 501.

[13] Shi G Q, Cao Y, Wang D H, et al. Towards intelligent unmanned adversarial games: a reinforcement learning framework with the PHP-ROW method[J]. Drones, 2025, 9(5): 331.

[14] 单圣哲, 张伟伟. 基于自博弈深度强化学习的空战智能决策方法[J]. 航空学报, 2024, 45(4): 328723.
Shan S Z, Zhang W W. Air combat intelligent decision-making method based on self-play and deep reinforcement learning[J]. Acta Aeronautica et Astronautica Sinica, 2024, 45(4): 328723(in Chinese).

[15] Li B, Huang J Y, Bai S X, et al. Autonomous air combat decision-making of UAV based on parallel self-play reinforcement learning[J]. CAAI Transactions on Intelligence Technology, 2023, 8(1): 64-81.

[16] Kong W R, Zhou D Y, Zhou Y, et al. Hierarchical reinforcement learning from competitive self-play for dual-aircraft formation air combat[J]. Journal of Computational Design and Engineering, 2023, 10(2): 830-859.

[17] Tasbas A S, Sahin S O, Üre N K. Reinforcement learning based self-play and state stacking techniques for noisy air combat environment: AIAA 2023-1077[R]. Reston: AIAA, 2023.

[18] Yang M C, Shan S Z, Zhang W W. Decision-making and confrontation in close-range air combat based on reinforcement learning[J]. Chinese Journal of Aeronautics, 2025, 38(9): 103526.

A multi-aircraft cooperative air combat decision-making algorithm based on an adaptive weighted priority virtual self-play theory approach

ZHOU Yaoming[1,*], HAN Tao[1], JIANG Yufei[1], YANG Zheng[2], TIAN Yifeng[2]

(1. School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China;
2. The 29th Research Institute, China Electronics Technology Group Corporation, Chengdu 610036, China)

Abstract: This study suggests an approach that combines prioritized fictitious self-play (PFSP) with an adaptive winning-game weighting mechanism to address the delayed convergence and policy oscillation in self-play training for multi-aircraft cooperative beyond-visual-range (BVR) air combat. Firstly, a 4v4 BVR air combat simulation environment is established under a multi-agent reinforcement learning framework. Second, in order to break the policy cycle, the Elo method is developed to conduct non-uniform sampling of past opponents and dynamically assess policy capabilities. Meanwhile, a win-rate-based loss function weighting mechanism is designed to guide gradient updates when early high-value samples are scarce. According to experiments, the suggested approach greatly lessens later-stage oscillation and converges to the Nash equilibrium by the 200th episode, increasing convergence efficiency by 55% over conventional FSP. The weighting mechanism's accelerating impact is further supported by ablation investigations. Lastly, agents educated using PFSP have strong air combat decision-making ability, winning 70.3% of the time against skilled opponents.

Keywords: air combat decision-making; self-play; reinforcement learning; winning-game weighting; multi-aircraft cooperation

Received: 2026-06-09; **Accepted**: 2026-07-17; **Published Online**: 2026-08-06 14:42
URL: link.cnki.net/urlid/11.2625.V.20260806.1112.002
* **Corresponding author**. E-mail: zhouyaoming@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0353

运载火箭地月转移入轨混合制导方法

郭康[1, 2]，李惠峰[1, 2, *]，包为民[3]

(1. 北京航空航天大学 宇航学院，北京 100191；2. 北京航空航天大学 天地往返高效运输技术全国重点实验室，北京 100191；
3. 中国航天科技集团有限公司，北京 100048)

摘 要：针对小推重比地月转移入轨任务中长飞行时间和强不确定性对运载火箭闭环制导方法带来的挑战，提出一种混合使用二阶锥规划和邻近最优的制导方法。建立考虑多源不确定性的地月转移入轨问题模型，并提出一种基于实时轨迹规划的混合制导策略。实现一种基于高斯伪谱离散与二阶锥规划的实时轨迹规划方法，综合应用协态映射定理和二阶锥规划对偶输出，能够高精度地估计协态量，为邻近最优方法提供必要基础。提出一种基于邻近最优的轨迹校正方法，在不确定性导致火箭偏离轨迹时，能够为二阶锥规划提供高质量初始参考。开展数值实验，在一个地月转移入轨任务中，与现有制导方法对比，所提制导方法对地月转移入轨任务具有更强的适应性和最优性，混合制导方法的平均计算时间降低30.67%，每制导周期的平均算时间为75.353 ms，为地月转移入轨提供了一条新且可行的技术路线。

关 键 词：运载火箭制导；地月转移入轨；实时轨迹规划；二阶锥规划；邻近最优
中图分类号：V448.13
文献标志码：A **文章编号**：1001-5965(2026)10-3592-10

月球是中国深空探测的第一站，也是未来开展行星探测和小天体探测的前哨[1]。月球探测是目前世界各国航天工程的重要目标，相关计划包括中国探月工程四期、国际月球科研站、美国Artemis计划等。开展月球探测的必要条件是能够可靠、高效地进入地月转移轨道。与近地入轨不同，地月转移入轨需要更大的速度增量，在小推重比条件下飞行时间显著增加。例如，美国空间发射系统(space launch system, SLS)Block 1的地月转移入轨时间可达数百秒，地心转角可达45°[2]。受地球重力场影响，火箭运动方程将具有更强的非线性，严重违背现有上升段制导方法的基础假设，对制导方法的可靠性和推进剂最优性提出挑战。

火箭上升段制导是一个经典的工程问题，现有方法主要分为3类：迭代制导方法、动力显式制导(powered explicit guidance, PEG)方法、闭环轨迹优化制导方法。迭代制导方法发展自美国国家航空航天局(national aeronautics and space administration, NASA)为土星五号运载火箭研制的迭代制导模式(iterative guidance mode, IGM)[3]。

IGM是一种路径自适应的制导方法，在火箭偏离标称轨迹时，基于线性程序角、常值重力加速度等简化假设规划未来飞行轨迹。迭代制导方法对任务和不确定性具有较强适应能力。中国部分长征系列运载火箭亦采用了迭代制导方法，在长征二号F遥八任务中得到了首次验证[4]。国内外学者针对迭代制导开展了大量的改进工作[5-8]。例如，文献[5]提出一种向量形式的迭代制导方法，将火箭推力方向作为控制量，与程序角作为控制量对比，其控制形式更加简洁。文献[6]提出一种能够控制入轨姿态的迭代制导方法，将迭代制导中的线性程序角扩展至二次形式，能够保障火箭发动机关机时位置、速

收稿日期：2025-06-06；**录用日期**：2025-12-05；**网络出版时间**：2025-12-17 10:49
网络出版地址：link.cnki.net/urlid/11.2625.V.20251216.1600.002
基金项目：国家自然科学基金(92471204)
* **通信作者**. E-mail：lihuifeng@buaa.edu.cn

引用格式：郭康，李惠峰，包为民. 运载火箭地月转移入轨混合制导方法[J]. 北京航空航天大学学报，2026，52(10)：3592-3601.
Guo K，Li H F，Bao W M. Hybrid guidance method for launch vehicle trans-lunar injection[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3592-3601 (in Chinese).

度及姿态约束的同时满足。文献 [7] 提出一种改进的迭代制导方法, 在迭代制导中不加简化地使用最优控制形式与横截条件, 提高了迭代制导的推进剂最优性, 并提出基于高斯-勒让德积分的推力积分数值计算方法, 提高了制导算法的积分精度。目前, 迭代制导的主要应用场景是大推重比近地轨道入轨任务。在小推重比长飞行时间任务中, 迭代制导的线性程序角、常值重力加速度等基础假设不再成立。若使用迭代制导直接入轨地月转移轨道, 制导性能可能显著下降[9]。除近地轨道入轨任务, 学者们还将迭代制导方法扩展应用于深空探测等对自主性要求较高的复杂场景[10]。例如, 文献 [11] 系统分析了月面上升与常规运载火箭上升的特性差异, 提出一种专用于月面上升的迭代制导算法, 并针对垂直上升段与迭代制导段之间存在的显著姿态角差异问题, 设计了一种最优轨迹衔接方法。此外, 文献 [12] 面向火星环境的高不确定性特点, 开发了一种具有自适应能力的火星上升迭代制导方法, 实现了在复杂火星环境下的高精度入轨控制。

PEG 方法是航天飞机应用的上升与轨道转移制导方法, 基于线性正切推力等简化假设, 以半解析方式规划飞行轨迹[13-15]。与 IGM 相比, PEG 具有更强的任务适应性, 兼顾了航天飞机的大推重比上升与小推重比轨道转移任务。然而, PEG 在小推重比地月转移入轨中的应用同样受限。例如, 文献 [16] 研究了 PEG 在 SLS Block 1 中的应用, 针对小推重比入轨地月转移轨道时 PEG 的性能严重下降甚至迭代发散现象, 提出推力方向变化限幅等改进方法, 增强了 PEG 对该工况的适应能力。在实际深空探测任务中, 嫦娥五号探测器的月面上升与着陆任务均采用了 PEG 方法[17]。该方案通过引入预报策略, 能够在 PEG 算法正式启动前预测飞行器状态并预迭代制导参数, 从而实现对姿态角的提前预报, 最终保证了飞行器能够及时完成姿态调整以满足轨道匹配的精确要求。

闭环轨迹优化制导方法通过数值方法优化火箭轨迹作为制导指令, 根据轨迹优化方法的不同, 可进一步分为间接法与直接法。间接法通过推导火箭轨迹优化问题的必要条件, 形成两点边值问题并求解的方式规划轨迹, 如文献 [18] 提出的 OPGUID (optimal guidance) 与 Switch 制导方法。直接法将火箭轨迹优化问题转化为数学规划问题并求解, 是目前广泛使用的火箭任务重构方法[19-25]。然而, 现有文献中的轨迹优化往往只进行一次, 随后以开环的方式使用其规划结果, 缺乏对不确定性的适应能力。

本文主要创新包括以下 3 点:

1) 提出一种混合使用二阶锥规划和邻近最优的制导策略。在多源不确定性导致火箭持续偏离优化轨迹、二阶锥规划收敛慢时, 基于邻近最优的一阶必要条件修正火箭飞行轨迹, 作为二阶锥规划优化的初始猜想。与现有方法对比, 混合制导方法解决了多源不确定性下闭环轨迹优化收敛慢的问题, 提升了计算效率。

2) 提出一种基于高斯伪谱法与二阶锥规划的实时轨迹优化方法。与现有方法对比, 主要增量在于使用高斯伪谱法的协态映射定理和二阶锥规划的对偶输出, 实现了对协态量的高精度估计, 为邻近最优方法提供了可靠支撑。

3) 在典型地月转移入轨任务中, 定量对比了本文制导方法与现有迭代制导方法, 说明了本文制导方法在适应性、推进剂最优性等方面的优势, 明确了应用场景。

1　地月转移入轨制导问题

1.1　多源不确定性下的火箭运动方程

考虑运载火箭在大气层外的质心运动, 在地心惯性坐标系中建立火箭运动方程。状态量包括质心位置$\boldsymbol{r}$、质心速度$\boldsymbol{V}$和质量m, 控制量包括推力方向$\boldsymbol{u}$, 则火箭的质心运动可描述为

$$\dot{\boldsymbol{r}}(t)=\boldsymbol{V}(t) \tag{1}$$

$$\dot{\boldsymbol{V}}(t)=\boldsymbol{g}[\boldsymbol{r}(t)]+\frac{(1+\delta_T)T}{m(t)}(\boldsymbol{I}+\boldsymbol{\delta}_u)\boldsymbol{u}(t) \tag{2}$$

$$\dot{m}(t)=-(1+\delta_R)R \tag{3}$$

式中: t为时间; $\boldsymbol{g}$为重力加速度; $\boldsymbol{I}$为 3 维单位矩阵; T为推力; R为推进剂秒耗量; δ_T、$\boldsymbol{\delta}_u$和δ_R为本文考虑的 3 种不确定性: 推力大小偏差、推力方向偏差、秒耗量偏差。

推力方向$\boldsymbol{u}$受到单位幅值约束:

$$\|\boldsymbol{u}(t)\|=1 \tag{4}$$

$\boldsymbol{\delta}_u$为三维矩阵, 满足以下约束关系:

$$\|\boldsymbol{I}+\boldsymbol{\delta}_u\|=1 \tag{5}$$

在火箭飞行过程中, 多源不确定性δ_T、$\boldsymbol{\delta}_u$和δ_R将使火箭持续偏离标称轨迹。在缺少不确定性模型、难以准确辨识δ_T、$\boldsymbol{\delta}_u$和δ_R的条件下, 实时轨迹规划是提高入轨精度和推进剂最优性的有效方法: 通过在每个制导周期中重新规划未来的最优轨迹, 能够避免跟踪标称轨迹带来的性能损失。该类方法的关键是如何在有限的计算时间(制导周期)内可靠地规划火箭飞行轨迹。

1.2　地月转移入轨条件

地月转移入轨的任务目标是抵达一个给定的地月转移轨道, 由其半长轴a_{ref}、偏心率e_{ref}、轨道倾角i_{ref}、升交点赤经Ω_{ref}、近地点幅角ω_{ref}确定。与近

地轨道入轨任务不同的是，地月转移轨道具有更显著的偏心率。为此，本文提出一种基于动量矩$\boldsymbol{L}$与拉普拉斯向量$\boldsymbol{A}$的降阶入轨条件，计算公式分别为

$$\boldsymbol{L}(\boldsymbol{r},\boldsymbol{V})=\boldsymbol{r}\times\boldsymbol{V} \tag{6}$$

$$\boldsymbol{A}(\boldsymbol{r},\boldsymbol{V})=\boldsymbol{V}\times\boldsymbol{L}(\boldsymbol{r},\boldsymbol{V})-\mu\boldsymbol{r}/\|\boldsymbol{r}\| \tag{7}$$

式中：μ为地球引力常数。

本文针对地月转移入轨问题的特点，提出以下形式的入轨条件：

$$\boldsymbol{\beta}[\boldsymbol{r}(t_{\mathrm{f}}),\boldsymbol{V}(t_{\mathrm{f}})]=\begin{bmatrix}\boldsymbol{L}[\boldsymbol{r}(t_{\mathrm{f}}) & \boldsymbol{V}(t_{\mathrm{f}})]-\boldsymbol{L}_{\mathrm{ref}}\\ \boldsymbol{n}_1^{\mathrm{T}}\boldsymbol{A}[\boldsymbol{r}(t_{\mathrm{f}}) & \boldsymbol{V}(t_{\mathrm{f}})]-\boldsymbol{n}_1^{\mathrm{T}}\boldsymbol{A}_{\mathrm{ref}}\\ \boldsymbol{n}_2^{\mathrm{T}}\boldsymbol{A}[\boldsymbol{r}(t_{\mathrm{f}}) & \boldsymbol{V}(t_{\mathrm{f}})]-\boldsymbol{n}_2^{\mathrm{T}}\boldsymbol{A}_{\mathrm{ref}}\end{bmatrix}=\boldsymbol{0} \tag{8}$$

式中：$\boldsymbol{\beta}$为入轨条件；$\boldsymbol{n}_1$、$\boldsymbol{n}_2$分别为入轨条件中第1、第2个方向；t_{f}为入轨时刻；$\boldsymbol{L}_{\mathrm{ref}}$和$\boldsymbol{A}_{\mathrm{ref}}$分别为目标地月转移轨道的动量矩和拉普拉斯向量，计算公式为

$$\boldsymbol{L}_{\mathrm{ref}}=\sqrt{\mu a_{\mathrm{ref}}(1-e_{\mathrm{ref}}^2)}\begin{bmatrix}\sin i_{\mathrm{ref}}\sin\Omega_{\mathrm{ref}}\\ -\sin i_{\mathrm{ref}}\cos\Omega_{\mathrm{ref}}\\ \cos i_{\mathrm{ref}}\end{bmatrix} \tag{9}$$

$$\boldsymbol{A}_{\mathrm{ref}}=\mu e_{\mathrm{ref}}\begin{bmatrix}\cos\Omega_{\mathrm{ref}}\cos\omega_{\mathrm{ref}}-\cos i_{\mathrm{ref}}\sin\Omega_{\mathrm{ref}}\sin\omega_{\mathrm{ref}}\\ \sin\Omega_{\mathrm{ref}}\cos\omega_{\mathrm{ref}}+\cos i_{\mathrm{ref}}\cos\Omega_{\mathrm{ref}}\sin\omega_{\mathrm{ref}}\\ \sin i_{\mathrm{ref}}\sin\omega_{\mathrm{ref}}\end{bmatrix} \tag{10}$$

与现有文献[25]中使用的入轨条件相比，式(8)只约束了$\boldsymbol{n}_1$和$\boldsymbol{n}_2$这2个特定方向的拉普拉斯向量$\boldsymbol{A}$。这一方面降低了约束的数量，同时消除了约束间线性相关性，为邻近最优方法的应用提供了必要条件。考虑到拉普拉斯向量的幅值与偏心率相关，且指向近地点，将$\boldsymbol{n}_1$和$\boldsymbol{n}_2$设计为

$$\boldsymbol{n}_1=\boldsymbol{A}_{\mathrm{ref}}/\|\boldsymbol{A}_{\mathrm{ref}}\| \tag{11}$$

$$\boldsymbol{n}_2=(\boldsymbol{L}_{\mathrm{ref}}\times\boldsymbol{A}_{\mathrm{ref}})/\|\boldsymbol{L}_{\mathrm{ref}}\times\boldsymbol{A}_{\mathrm{ref}}\| \tag{12}$$

对地月转移任务(偏心率e_{ref}不为0)，$\boldsymbol{n}_1$和$\boldsymbol{n}_2$具有良好(非奇异)的定义。式(8)中的约束等价于指定轨道五要素，入轨点的真近点角将优化确定。

1.3 地月转移入轨的最优控制建模

地月转移入轨问题可建模为以下最优控制问题：

问题1 地月转移入轨问题。

$$\begin{cases}\text{minimize} & m(t_0)-m(t_{\mathrm{f}})\\ \text{subject to} & \dot{\boldsymbol{r}}(t)=\boldsymbol{V}(t)\\ & \begin{cases}\dot{\boldsymbol{V}}(t)=\boldsymbol{g}[\boldsymbol{r}(t)]+\dfrac{T}{m(t)}\boldsymbol{u}(t)\\ \dot{m}(t)=-R\\ \|\boldsymbol{u}(t)\|=1\end{cases}\\ & \boldsymbol{r}(t_0)=\boldsymbol{r}_0\\ & \boldsymbol{V}(t_0)=\boldsymbol{V}_0\\ & m(t_0)=m_0\\ & \boldsymbol{\beta}[\boldsymbol{r}(t_{\mathrm{f}}),\boldsymbol{V}(t_{\mathrm{f}})]=\boldsymbol{0}\end{cases} \tag{13}$$

式中：t_0为当前时刻；$\boldsymbol{r}_0$、$\boldsymbol{V}_0$和m_0为火箭当前的飞行状态(质心位置、质心速度和质量)。

问题1的性能指标是最小化飞行过程中的推进剂消耗，其解表示一条以推进剂最优方式将火箭从当前状态转移至地月转移轨道的飞行轨迹，是闭环轨迹优化制导方法的基础。

然而，问题1中没有考虑火箭运动中的多源不确定性δ_T、$\boldsymbol{\delta}_{\boldsymbol{u}}$和$\delta_R$。在使用问题1解中的最优控制量(推力方向)时，火箭将在不确定性作用下持续偏离优化出的标称轨迹，难以精确入轨。

针对不确定性带来的影响，需要闭环轨迹规划方法重复求解问题1进行修正。制导性能的关键在于求解问题的计算效率：求解问题1越快，不确定性在运动方程中的累计作用越小。考虑到在制导周期间，当前时刻t_0和当前飞行状态$\boldsymbol{r}_0$、$\boldsymbol{V}_0$和m_0的变化较小，如何利用上一制导周期得到的轨迹初始化当前周期中的轨迹优化是提高计算效率的关键。需要注意的是，在多源不确定性δ_T、$\boldsymbol{\delta}_{\boldsymbol{u}}$和$\delta_R$的影响下，上一制导周期中的最优轨迹不一定是理想的初始参考。

2 基于实时轨迹优化的混合制导

本文提出一种综合邻近最优和二阶锥规划的混合制导方法：在每个制导周期中(设当前制导周期为k)，首先，基于当前的飞行状态$\boldsymbol{r}_0^k$、$\boldsymbol{V}_0^k$和m_0^k，应用邻近最优方法校正上一制导周期的最优飞行时间t_{f}^{k-1}和飞行轨迹$(\boldsymbol{r}^{k-1},\boldsymbol{V}^{k-1},m^{k-1},\boldsymbol{u}^{k-1})$；然后，基于邻近最优方法得到的初始参考，应用二阶锥规划方法实时求解问题1，得到最优飞行轨迹。该混合制导方法的流程如图1所示。

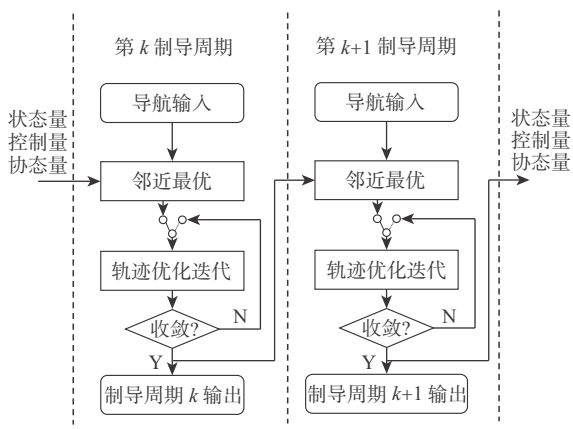


图1 基于实时轨迹规划混合制导方法的流程

Fig. 1 Flowchart of a hybrid guidance method based on real-time path planning

图1所示制导方法的主要特点是：在二阶锥规划开展较为耗时的迭代前，先使用邻近最优方法根据当前飞行状态更新参考轨迹，降低多源不确定性

的影响, 使轨迹规划方法能够快速收敛, 计算得到抵达目标地月转移轨道的高精度轨迹。

本文提出的混合制导策略同时使用了 2 种轨迹规划方法: 邻近最优方法与基于二阶锥规划的数值优化方法。每个制导周期中的计算过程如图 2 所示, 每个制导周期中将先使用计算效率高的邻近最优方法估计多源不确定性作用下新的最优轨迹, 再使用二阶锥规划方法对轨迹进行进一步的校正。

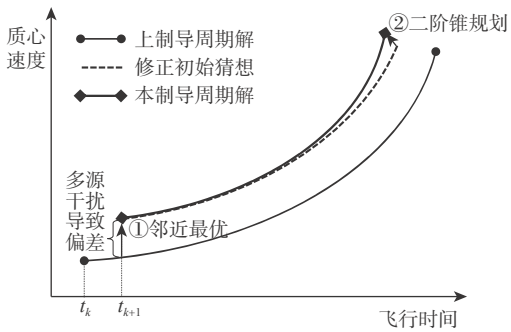


图 2　混合制导方法示意图

Fig. 2　Schematic diagram of a hybrid guidance method

在本文提出的混合制导策略下, 邻近最优方法和二阶锥规划方法实现了互补: 邻近最优方法计算效率高, 但需要准确的协态初始参考才能发挥其优势; 二阶锥规划方法虽然计算效率较低, 但轨迹精度高, 且能够可靠地生成协态参考。混合制导策略充分结合两者优势, 应用二阶锥规划方法为邻近最优方法提供高质量初始参考, 并应用邻近最优方法快速修正多源干扰带来的偏差, 避免二阶锥规划进行过多的耗时迭代, 从而实现兼具高轨迹精度和高计算效率的制导方法。

3　基于邻近最优的轨迹校正方法

在多源不确定性导致火箭偏离预期的飞行轨迹时, 应用邻近最优方法校正上一制导周期中的最优轨迹, 作为本制导周期中的最优参考。具体地, 给定上一制导周期中最优飞行时间 t_{f}^{k-1} 和飞行轨迹 $(\boldsymbol{r}^{k-1},\boldsymbol{V}^{k-1},m^{k-1},\boldsymbol{u}^{k-1})$, 并给定当前时间 t_0^k 和飞行状态 $\boldsymbol{r}_0^k$、$\boldsymbol{V}_0^k$ 和 m_0^k, 将初始状态偏差定义为

$$\delta\boldsymbol{x}_0^k=\begin{bmatrix}\boldsymbol{r}_0^k\\ \boldsymbol{V}_0^k\\ m_0^k\end{bmatrix}-\begin{bmatrix}\boldsymbol{r}^{k-1}(t_0^k)\\ \boldsymbol{V}^{k-1}(t_0^k)\\ m^{k-1}(t_0^k)\end{bmatrix}\tag{14}$$

邻近最优校正的目标是计算初始状态偏差 $\delta\boldsymbol{x}_0^k$ 下近似最优的飞行时间和飞行轨迹, 为二阶锥规划提供高质量的初始参考。

将地月转移入轨问题(问题 1)中的入轨时刻 t_{f} 视作优化参数, 并重定义控制量, 可得到下述等价固定时间、无控制约束的最优控制问题:

问题 2　参数化地月转移入轨问题。

$$\begin{cases}\text{minimize}\quad m(0)-m(1)\\ \text{subject to}\quad \dot{\boldsymbol{r}}(\tau)=(t_{\mathrm{f}}-t_0)\boldsymbol{V}(\tau)\\ \begin{cases}\dot{\boldsymbol{V}}(\tau)=(t_{\mathrm{f}}-t_0)\left(\boldsymbol{g}[\boldsymbol{r}(\tau)]+\dfrac{T}{m(\tau)}\boldsymbol{u}'(\tau)\right)\\ \dot{m}(\tau)=-(t_{\mathrm{f}}-t_0)R\\ \boldsymbol{r}(0)=\boldsymbol{r}_0\\ \boldsymbol{V}(0)=\boldsymbol{V}_0\\ m(0)=m_0\\ \boldsymbol{\beta}[\boldsymbol{r}(1),\boldsymbol{V}(1)]=\boldsymbol{0}\end{cases}\end{cases}\tag{15}$$

问题 2 对运动方程的自变量时间 t 进行了变换, 将问题 1 由终止时间 t_{f} 自由的问题转化为固定时间区间的问题。具体地, 引入新的时间自变量 τ:

$$\tau=\frac{t-t_0}{t_{\mathrm{f}}-t_0}\tag{16}$$

则 $\tau=0$ 对应 $t=t_0$, $\tau=1$ 对应 $t=t_{\mathrm{f}}$。此时, 入轨时刻 t_{f} 成为待优化的参数。

此外, 问题 2 中将控制量 $\boldsymbol{u}$ 重新定义为 $\boldsymbol{u}'=(u_a,u_b)$, 其与原本控制量 $\boldsymbol{u}$ 的关系为

$$\begin{bmatrix}\cos u_a\cos u_b\\ \cos u_a\sin u_b\\ \sin u_a\end{bmatrix}=\boldsymbol{u}\tag{17}$$

通过上述时间变换和控制量重定义, 地月转移入轨问题转化为含参数的固定时间最优控制问题。问题 2 的必要条件包括:

$$\begin{cases}\dot{\boldsymbol{x}}(\tau)=\boldsymbol{f}[\boldsymbol{x}(\tau),\boldsymbol{u}'(\tau),t_{\mathrm{f}}]\\ \dot{\boldsymbol{\lambda}}(\tau)=-\left(\dfrac{\partial\boldsymbol{H}}{\partial\boldsymbol{x}}[\boldsymbol{x}(\tau),\boldsymbol{u}'(\tau),\boldsymbol{\lambda}(\tau),t_{\mathrm{f}}]\right)^{\mathrm{T}}\\ \dfrac{\partial\boldsymbol{H}}{\partial\boldsymbol{u}'}[\boldsymbol{x}(\tau),\boldsymbol{u}'(\tau),\boldsymbol{\lambda}(\tau),t_{\mathrm{f}}]=\boldsymbol{0}\end{cases}\tag{18}$$

式中: $\boldsymbol{x}=(\boldsymbol{r},\boldsymbol{V},m)$ 为状态量; $\boldsymbol{f}$ 为问题 2 的运动方程; $\boldsymbol{\lambda}$ 为协态量; $\boldsymbol{H}=\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{f}$ 为最优控制问题的哈密顿函数。

应用邻近最优理论[21-23], 在最优飞行时间 t_{f}^{k-1} 和飞行轨迹 $(\boldsymbol{r}^{k-1},\boldsymbol{V}^{k-1},m^{k-1},\boldsymbol{u}^{k-1})$ 附近展开问题 2 的必要条件, 可得到火箭状态修正值 $\delta\boldsymbol{x}$、协态修正值 $\delta\boldsymbol{\lambda}$ 和控制修正值 $\delta\boldsymbol{u}'$ 应满足的运动方程。该运动方程可写为紧凑形式:

$$\begin{cases}\delta\dot{\boldsymbol{x}}(\tau)=\boldsymbol{B}_1(\tau)\delta\boldsymbol{x}(\tau)-\boldsymbol{B}_2(\tau)\delta\boldsymbol{\lambda}(\tau)+\boldsymbol{B}_4(\tau)\mathrm{d}t_{\mathrm{f}}\\ \delta\dot{\boldsymbol{\lambda}}(\tau)=-\boldsymbol{B}_3(\tau)\delta\boldsymbol{x}(\tau)-\boldsymbol{B}_1{}^{\mathrm{T}}(\tau)\delta\boldsymbol{\lambda}(\tau)-\boldsymbol{B}_5(\tau)\mathrm{d}t_{\mathrm{f}}\\ \delta\boldsymbol{u}'(\tau)=-\boldsymbol{H}_{\boldsymbol{u}'\boldsymbol{u}'}^{-1}(\tau)\left[\boldsymbol{H}_{\boldsymbol{u}'\boldsymbol{x}}(\tau)\delta\boldsymbol{x}(\tau)+\right.\\ \qquad\left.\boldsymbol{f}_{\boldsymbol{u}'}^{\mathrm{T}}(\tau)\delta\boldsymbol{\lambda}(\tau)+\boldsymbol{H}_{\boldsymbol{u}'t_{\mathrm{f}}}(\tau)\mathrm{d}t_{\mathrm{f}}\right]\end{cases}\tag{19}$$

式中: $\mathrm{d}t_{\mathrm{f}}$ 为飞行时间偏差; 带下标的 $\boldsymbol{f}$ 和 $\boldsymbol{H}$ 表示其对各变量的一阶和二阶偏导数, 例如:

$$\boldsymbol{H}_{\boldsymbol{u}'\boldsymbol{x}}(\tau)=\frac{\partial^2\boldsymbol{H}}{\partial\boldsymbol{u}'\,\partial\boldsymbol{x}}\left[\boldsymbol{x}^{k-1}(\tau),\boldsymbol{u}'^{k-1}(\tau),\boldsymbol{\lambda}^{k-1}(\tau),t_{\mathrm{f}}^{k-1}\right]\tag{20}$$

$\boldsymbol{B}_i(i=1,2,\cdots,5)$表示以下形式的矩阵：

$$\begin{cases}\boldsymbol{B}_1(\tau)=\boldsymbol{f}_{\boldsymbol{x}}(\tau)-\boldsymbol{f}_{\boldsymbol{u}'}(\tau)\boldsymbol{H}_{\boldsymbol{u}'\boldsymbol{u}'}^{-1}(\tau)\boldsymbol{H}_{\boldsymbol{u}'\boldsymbol{x}}(\tau)\\ \boldsymbol{B}_2(\tau)=\boldsymbol{f}_{\boldsymbol{u}'}(\tau)\boldsymbol{H}_{\boldsymbol{u}'\boldsymbol{u}'}^{-1}(\tau)\boldsymbol{f}_{\boldsymbol{u}'}^{\mathrm{T}}(\tau)\\ \boldsymbol{B}_3(\tau)=\boldsymbol{H}_{\boldsymbol{x}\boldsymbol{x}}(\tau)-\boldsymbol{H}_{\boldsymbol{x}\boldsymbol{u}'}(\tau)\boldsymbol{H}_{\boldsymbol{u}'\boldsymbol{u}'}^{-1}(\tau)\boldsymbol{H}_{\boldsymbol{u}'\boldsymbol{x}}(\tau)\\ \boldsymbol{B}_4(\tau)=\boldsymbol{f}_{t_{\mathrm{f}}}(\tau)-\boldsymbol{f}_{\boldsymbol{u}'}(\tau)\boldsymbol{H}_{\boldsymbol{u}'\boldsymbol{u}'}^{-1}(\tau)\boldsymbol{H}_{\boldsymbol{u}t_{\mathrm{f}}}(\tau)\\ \boldsymbol{B}_5(\tau)=\boldsymbol{H}_{\boldsymbol{x}t_{\mathrm{f}}}(\tau)-\boldsymbol{H}_{\boldsymbol{x}\boldsymbol{u}'}(\tau)\boldsymbol{H}_{\boldsymbol{u}'\boldsymbol{u}'}^{-1}(\tau)\boldsymbol{H}_{\boldsymbol{u}'t_{\mathrm{f}}}(\tau)\end{cases}\tag{21}$$

在第k个制导周期中，量测得到初始状态偏差$\delta\boldsymbol{x}_0^k$后，若初始协态修正量$\delta\boldsymbol{\lambda}_0^k$和飞行时间修正量$\mathrm{d}t_{\mathrm{f}}$已知，则可求解式(19)中的微分方程，从而获得飞行轨迹的修正量$\delta\boldsymbol{x}^k(\tau)$和$\delta\boldsymbol{u}'^k(\tau)$。本文应用扫掠法计算初始协态修正量$\delta\boldsymbol{\lambda}_0^k$和飞行时间修正量$\mathrm{d}t_{\mathrm{f}}^k$。问题2对应的扫掠变量$\boldsymbol{S}_i$满足的运动方程为

$$\begin{bmatrix}\dot{\boldsymbol{S}}_1\\ \dot{\boldsymbol{S}}_2\\ \dot{\boldsymbol{S}}_3\\ \dot{\boldsymbol{S}}_4\\ \dot{\boldsymbol{S}}_5\\ \dot{\boldsymbol{S}}_6\end{bmatrix}=\begin{bmatrix}-\boldsymbol{B}_3-\boldsymbol{B}_1^{\mathrm{T}}\boldsymbol{S}_1-\boldsymbol{S}_1\boldsymbol{B}_1+\boldsymbol{S}_1\boldsymbol{B}_2\boldsymbol{S}_1\\ (\boldsymbol{S}_1\boldsymbol{B}_2-\boldsymbol{S}_1^{\mathrm{T}})\boldsymbol{S}_2\\ (\boldsymbol{S}_1\boldsymbol{B}_2-\boldsymbol{S}_1^{\mathrm{T}})\boldsymbol{S}_3-\boldsymbol{S}_1\boldsymbol{B}_4-\boldsymbol{B}_5\\ \boldsymbol{S}_2^{\mathrm{T}}\boldsymbol{B}_2\boldsymbol{S}_2\\ \boldsymbol{S}_2^{\mathrm{T}}(\boldsymbol{B}_2\boldsymbol{S}_3-\boldsymbol{B}_4)\\ \boldsymbol{S}_3^{\mathrm{T}}(\boldsymbol{B}_2\boldsymbol{S}_3-\boldsymbol{B}_4)-\boldsymbol{B}_4^{\mathrm{T}}\boldsymbol{S}_3-\boldsymbol{B}_6\end{bmatrix}\tag{22}$$

式中：

$$\boldsymbol{B}_6=\boldsymbol{H}_{t_{\mathrm{f}}t_{\mathrm{f}}}-\boldsymbol{H}_{t_{\mathrm{f}}\boldsymbol{u}'}\boldsymbol{H}_{\boldsymbol{u}'\boldsymbol{u}'}^{-1}\boldsymbol{H}_{\boldsymbol{u}'t_{\mathrm{f}}}\tag{23}$$

设地月转移入轨条件$\boldsymbol{\beta}=\boldsymbol{0}$对应的拉格朗日乘子为$\boldsymbol{\nu}$，定义函数：

$$G(\boldsymbol{x},\boldsymbol{\nu})=-m(1)+\boldsymbol{\nu}^{\mathrm{T}}\boldsymbol{\beta}[\boldsymbol{r}(1),\boldsymbol{V}(1)]\tag{24}$$

则扫掠变量$\boldsymbol{S}_i$的终止状态满足以下条件：

$$\begin{cases}\boldsymbol{S}_1(1)=\boldsymbol{G}_{\boldsymbol{x}\boldsymbol{x}}\\ \boldsymbol{S}_2(1)=\boldsymbol{\beta}_{\boldsymbol{x}}^{\mathrm{T}}\\ \boldsymbol{S}_3(1)=\boldsymbol{0}\\ \boldsymbol{S}_4(1)=\boldsymbol{0}\\ \boldsymbol{S}_5(1)=\boldsymbol{0}\\ \boldsymbol{S}_6(1)=\boldsymbol{0}\end{cases}\tag{25}$$

应用数值积分方法在时间区间$\tau\in[0,1]$上反向求解式(22)中的状态方程，可获得扫掠变量$\boldsymbol{S}_i$在$\tau=0$处的状态。通过求解线性化后的必要条件，初始协态修正量$\delta\boldsymbol{\lambda}_0^k$和飞行时间偏差$\mathrm{d}t_{\mathrm{f}}$应满足以下必要条件：

$$\delta\boldsymbol{\lambda}_0^k=\left(\boldsymbol{S}_1-\begin{bmatrix}\boldsymbol{S}_2,\boldsymbol{S}_3\end{bmatrix}\begin{bmatrix}\boldsymbol{S}_4 & \boldsymbol{S}_5\\ \boldsymbol{S}_5^{\mathrm{T}} & \boldsymbol{S}_6\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{S}_2^{\mathrm{T}}\\ \boldsymbol{S}_3^{\mathrm{T}}\end{bmatrix}\right)\delta\boldsymbol{x}_0^k\tag{26}$$

$$\begin{bmatrix}\mathrm{d}\boldsymbol{\nu}^k\\ \mathrm{d}t_{\mathrm{f}}^k\end{bmatrix}=-\begin{bmatrix}\boldsymbol{S}_4 & \boldsymbol{S}_5\\ \boldsymbol{S}_5^{\mathrm{T}} & \boldsymbol{S}_6\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{S}_2^{\mathrm{T}}\\ \boldsymbol{S}_3^{\mathrm{T}}\end{bmatrix}\delta\boldsymbol{x}_0^k\tag{27}$$

式中：为简明起见，$\boldsymbol{S}_i$均表示$\boldsymbol{S}_i(\tau)$在0处的取值；$\mathrm{d}\boldsymbol{\nu}^k$为拉格朗日乘子$\boldsymbol{\nu}$对应的修正量。

计算得到$\delta\boldsymbol{\lambda}_0^k$和$\mathrm{d}t_{\mathrm{f}}$后，可求解(19)中的微分方程，获得状态修正值$\delta\boldsymbol{x}^k$和控制修正值$\delta\boldsymbol{u}'^k$，从而将第$k$个制导周期中的初始猜想设置为

$$\begin{bmatrix}\boldsymbol{r}^k(\tau)\\ \boldsymbol{V}^k(\tau)\\ m^k(\tau)\end{bmatrix}=\begin{bmatrix}\boldsymbol{r}^{k-1}(\tau)\\ \boldsymbol{V}^{k-1}(\tau)\\ m^{k-1}(\tau)\end{bmatrix}+\delta\boldsymbol{x}^k(\tau)\tag{28}$$

$$\boldsymbol{u}^k(\tau)=\begin{bmatrix}\cos\left[u_a^{k-1}(\tau)+\delta u_a^k(\tau)\right]\cos\left[u_b^{k-1}(\tau)+\delta u_b^k(\tau)\right]\\ \cos\left[u_a^{k-1}(\tau)+\delta u_a^k(\tau)\right]\sin\left[u_b^{k-1}(\tau)+\delta u_b^k(\tau)\right]\\ \sin\left[u_a^{k-1}(\tau)+\delta u_a^k(\tau)\right]\end{bmatrix}\tag{29}$$

$$t_{\mathrm{f}}^k=t_{\mathrm{f}}^{k-1}+\mathrm{d}t_{\mathrm{f}}^k\tag{30}$$

式(28)～式(30)中的初始猜想具有以下特点：① 在$\tau=0$处，初始状态等于火箭当前飞行状态；②满足问题2的线性化一阶必要条件。因此，邻近最优方法提供的初始猜想一方面修正了多源干扰产生初始状态偏差$\delta\boldsymbol{x}_0^k$的影响，另一方面具有近似最优性，可为二阶锥规划提供高质量的初始参考。

4 基于二阶锥规划的轨迹规划方法

考虑问题1中描述的最优控制问题，本节提出一种基于高斯伪谱法和二阶锥规划的轨迹优化方法及一种基于协态映射定理的协态量估计方法。问题1的哈密顿函数可表示为

$$\boldsymbol{H}=\boldsymbol{\lambda}^{\mathrm{T}}\boldsymbol{f}=\boldsymbol{\lambda}_{\boldsymbol{r}}^{\mathrm{T}}\dot{\boldsymbol{r}}+\boldsymbol{\lambda}_{\boldsymbol{V}}^{\mathrm{T}}\dot{\boldsymbol{V}}+\lambda_m\dot{m}\tag{31}$$

式中：$\boldsymbol{\lambda}=(\boldsymbol{\lambda}_{\boldsymbol{r}},\boldsymbol{\lambda}_{\boldsymbol{V}},\lambda_m)$为协态量；$\boldsymbol{\lambda}_{\boldsymbol{r}}$、$\boldsymbol{\lambda}_{\boldsymbol{V}}$和$\lambda_m$分别为位置协态、速度协态和质量协态。

应用极小值原理，最优控制量$\boldsymbol{u}^*$的选择应使哈密顿函数H最小化，即

$$\boldsymbol{u}^*(t)=\arg\min\boldsymbol{H}=-\frac{\boldsymbol{\lambda}_{\boldsymbol{V}}(t)}{\|\boldsymbol{\lambda}_{\boldsymbol{V}}(t)\|}\qquad\boldsymbol{\lambda}_{\boldsymbol{V}}(t)\neq0\tag{32}$$

注意到上述关系在等式约束$\|\boldsymbol{u}(t)\|=1$或不等式约束$\|\boldsymbol{u}(t)\|\leqslant1$下均成立。因此，可应用无损凸化方法将问题1中的等式约束松弛为不等式约束，且不改变其最优解的特性，相关理论分析与证明见文献[24]。与问题2相同，同样将时间区间从$t\in[t_0,t_{\mathrm{f}}]$变换为$\tau\in[0,1]$，得到如下形式的最优控制问题：

问题3 固定时间区间的地月转移入轨问题。

$$\begin{cases}\text{minimize}\quad m(0)-m(1)\\ \text{subject to}\quad\dot{\boldsymbol{x}}(\tau)=(t_{\mathrm{f}}-t_0)\boldsymbol{f}[\boldsymbol{x}(\tau),\boldsymbol{u}(\tau)]\\ \|\boldsymbol{u}(\tau)\|\leqslant1\\ \boldsymbol{r}(0)=\boldsymbol{r}_0\\ \boldsymbol{V}(0)=\boldsymbol{V}_0\\ m(0)=m_0\\ \boldsymbol{\beta}[\boldsymbol{r}(t_{\mathrm{f}}),\boldsymbol{V}(t_{\mathrm{f}})]=0\end{cases}\tag{33}$$

在问题3中，初始时间t_0为常量，终止时间t_{f}为待优化的问题参数。应用序列线性化方法，可将问题3转化为一系列的二阶锥规划问题，并应用内点法实时求解。

在邻近最优方法提供的初始参考轨迹 $\left[\boldsymbol{x}^k(\tau), \boldsymbol{u}^k(\tau), t_{\mathrm{f}}^k\right]$ 附近将运动方程线性展开, 得到以下形式的线性化运动方程:

$$\dot{\boldsymbol{x}}(\tau)=\boldsymbol{f}_{\mathrm{linear}}^k[\boldsymbol{x}(\tau),\boldsymbol{u}(\tau),t_{\mathrm{f}}] \tag{34}$$

应用高斯伪谱法离散化问题 3, 离散点数量为 N, 则各离散点 τ_i 的位置为

$$\tau_i=\begin{cases}0 & i=1\\ 1 & i=N\\ (P_{N,i-1}+1)/2 & \text{其他}\end{cases} \tag{35}$$

式中: $P_{N,i-1}$ 为 $[-1,1]$ 区间上的勒让德-高斯点。

状态量 $\boldsymbol{x}(\tau)$ 使用前 $N-1$ 个点组成的拉格朗日插值多项式 l_i 近似表示:

$$\boldsymbol{x}(\tau)=\sum_{i=1}^{N-1} l_i(\tau)\boldsymbol{x}_i \tag{36}$$

控制量 $\boldsymbol{u}(\tau)$ 使用中间 $N-2$ 个点组成的拉格朗日插值多项式近似表示:

$$\boldsymbol{u}(\tau)=\sum_{i=2}^{N-1} l_i(\tau)\boldsymbol{u}_i \tag{37}$$

终端状态可使用高斯积分公式获得:

$$\boldsymbol{x}_N=\boldsymbol{x}_0+\frac{1}{2}\sum_{k=2}^{N-1} w_k\boldsymbol{f}_{\mathrm{linear}}^k(\boldsymbol{x}_k,\boldsymbol{u}_k) \tag{38}$$

式中: w_k 为高斯伪谱法的积分系数。

令运动方程在中间 $N-2$ 个点上成立, 建立非线性方程组:

$$\sum_{j=1}^{N-1}\dot{l}_j(\tau_i)\boldsymbol{x}_j=(t_{\mathrm{f}}-t_0)\boldsymbol{f}_{\mathrm{linear}}^k(\boldsymbol{x}_i,\boldsymbol{u}_i)\qquad i=2,3,\cdots,N-1 \tag{39}$$

将式 (20) 和式 (21) 中的非线性方程组写为紧凑形式:

$$\boldsymbol{X}_N=\boldsymbol{D}_N\boldsymbol{F}_N \tag{40}$$

式中: $\boldsymbol{X}_N$ 为状态量组成的矩阵; $\boldsymbol{F}_N$ 为状态方程组成的矩阵; $\boldsymbol{D}_N$ 为高斯伪谱法的微分方程。

至此, 建立一个二阶锥规划问题:

问题 4　离散化后的地月转移入轨问题。

$$\begin{cases}\text{minimize} & m_1-m_N\\ \text{subject to} & \boldsymbol{X}_N=\boldsymbol{D}_N\boldsymbol{F}_N\\ & \|\boldsymbol{u}_k\|\leqslant 1\\ & \boldsymbol{r}_1=\boldsymbol{r}_0\\ & \boldsymbol{V}_1=\boldsymbol{V}_0\\ & m_1=m_0\\ & \boldsymbol{\beta}_{\mathrm{linear}}(\boldsymbol{r}_N,\boldsymbol{V}_N)=0\end{cases} \tag{41}$$

式中: $\boldsymbol{\beta}_{\mathrm{linear}}$ 为线性化地月转移入轨条件。

求解问题 4 可以获得新的轨迹 $\left[\boldsymbol{x}^k(\tau),\boldsymbol{u}^k(\tau),t_{\mathrm{f}}^k\right]$, 通过迭代方式, 其收敛后的解将满足问题 3 的最优必要条件, 可作为制导指令使用。

本文应用基于协态映射定理的协态量估计方法, 能够基于线性化矩阵方程 $\boldsymbol{X}_N=\boldsymbol{D}_N\boldsymbol{F}_N$ 对应的拉格朗日乘子矩阵 $\boldsymbol{\Lambda}$ 估计协态量 $\boldsymbol{\lambda}(\tau)$。具体地, 协态量 $\boldsymbol{\lambda}(\tau)$ 在中间 $N-2$ 的离散点处的估计值为

$$\boldsymbol{\lambda}_k=\boldsymbol{\Lambda}_k/w_k+\boldsymbol{\Lambda}_N\qquad k=2,3,\cdots,N-1 \tag{42}$$

协态量在最后一个离散点处的估计值为

$$\boldsymbol{\lambda}_N=\boldsymbol{\Lambda}_N \tag{43}$$

随后, 可应用拉格朗日插值方法估计协态量的连续形式:

$$\boldsymbol{\lambda}(\tau)=\sum_{k=2}^{N} l_k(\tau)\boldsymbol{\lambda}_k \tag{44}$$

协态映射定理估计出的协态量 $\boldsymbol{\lambda}(\tau)$ 为邻近最优方法的应用提供了必要基础。

迭代终止条件设计为迭代间终止时间 t_{f} 和终止速度 $\boldsymbol{V}_N$ 的相对变化:

$$\delta_{\mathrm{opt}}=\min\left\{\frac{\left|t_{\mathrm{f}}-t_{\mathrm{f}}^{\mathrm{pre}}\right|}{t_{\mathrm{f}}^{\mathrm{pre}}},\frac{\left\|\boldsymbol{V}_N-\boldsymbol{V}_N^{\mathrm{pre}}\right\|}{\boldsymbol{V}_N^{\mathrm{pre}}}\right\} \tag{45}$$

式中: 上标 pre 表示上一次迭代。

当 δ_{opt} 小于预设的阈值 ϵ_{opt} 时, 二阶锥规划迭代停止, 输出制导指令。综上所述, 提出的地月转移入轨混合制导方法的算法流程如下:

输入: 火箭飞行状态 $\boldsymbol{r}$、$\boldsymbol{V}$、m, 目标轨道参数 $\boldsymbol{L}_{\mathrm{ref}}$ 和 $\boldsymbol{V}_{\mathrm{ref}}$。

输出: 推力方向指令 $\boldsymbol{u}$。

初始化:

1. 应用线性插值方法, 生成状态量和控制量的初始参考 $(\boldsymbol{r}^0,\boldsymbol{V}^0,m^0,\boldsymbol{u}^0)$;

2. 将入轨时间 t_{f}^0 的初始参考设为标称入轨时间。

每个制导周期:

3. 从导航系统获取当前的飞行时间 t_0 和状态 $\boldsymbol{r}_0$、$\boldsymbol{V}_0$ 和 m_0;

4. 对第一个制导周期 $k=1$, 转步骤 10。否则, 转步骤 5;

5. 应用式 (14) 计算初始状态偏差 $\delta\boldsymbol{x}_0^k$;

6. 求解式 (22) 和式 (25) 描述的微分方程, 获得 $\boldsymbol{S}_i(0)$;

7. 应用式 (26) 和式 (27) 计算初始协态修正量 $\delta\boldsymbol{\lambda}_0^k$ 和飞行时间修正量 $\mathrm{d}t_{\mathrm{f}}$;

8. 求解式 (19) 描述的微分方程, 获得状态修正值 $\delta\boldsymbol{x}^k$ 和控制修正值 $\delta\boldsymbol{u}'^k$;

9. 应用式 (28)～式 (30) 修正初始参考;

10. 基于初始参考线性化问题 3 获得问题 4;

11. 应用二阶锥规划方法求解问题 4, 并应用

式 (42)～式 (44) 计算协态量;

12. 应用式 (45) 计算判据δ_{opt};

13. 若δ_{opt}小于预设的阈值ϵ_{opt}, 转步骤 14; 否则, 更新初始参考后转步骤 10;

14. 输出制导指令$\boldsymbol{u}$, 转步骤 3。

终止条件:

15. 当待飞时间$t_{\text{f}}-t_0\leqslant t_{\text{stop}}$时, 制导迭代停止;

当半长轴达到目标时, 关闭发动机完成入轨。

5 数值实验与分析

本节应用文献 [7] 中提供的地月转移工况, 验证本文混合制导方法的性能, 共进行了 2 类数值实验: 与现有方法的对比及对动力故障的适应能力。

在 2 类数值实验中, 制导方法的参数保持一致, 具体包括: 制导周期为 2 s, 制导停止时间t_{stop}为 60 s; 邻近最优中求解微分方程的方法为定步数四阶龙格-库塔方法, 步数为 50 步; 二阶锥规划中高斯伪谱法的离散点个数N为 10, 迭代停止阈值ϵ_{opt}为10^{-9}。混合制导方法使用 C 语言实现, 运行环境为主频为 1 GHz 的 Arm 处理器。

5.1 与现有方法的对比

为验证本文混合制导方法对地月转移入轨任务的适应能力, 在推力和秒耗量同时下降 35% 且存在 2% 未知偏差($\delta_T=\delta_R=0.02$)的情况下, 对比了本文混合制导方法与文献 [2,13] 中为 SLS Block 1 开发的改进 PEG 方法, 结果如图 3 和图 4 所示。

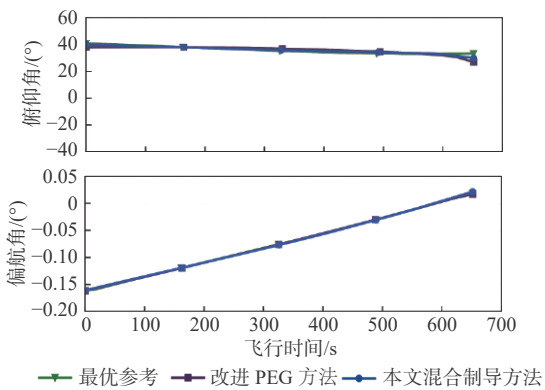


图 3 高度与速度曲线

Fig. 3 Altitude-versus-speed curve

在数值仿真中, 改进 PEG 与本文混合制导方法均完成了入轨任务。混合制导方法的入轨时间为 651.692 s。在 2% 的推力和秒耗量偏差下, 偏心率偏差为 0.000 147, 近地点幅角偏差为−0.004 154° (以半长轴作为入轨条件, 故半长轴无偏差)。与离线轨迹规划工具箱优化得到的最优参考对比, 燃料消耗相对误差为 5.9 kg。与之相对的, 改进 PEG 的燃料消耗相对误差为 30.9 kg, 较本文混合制导方法

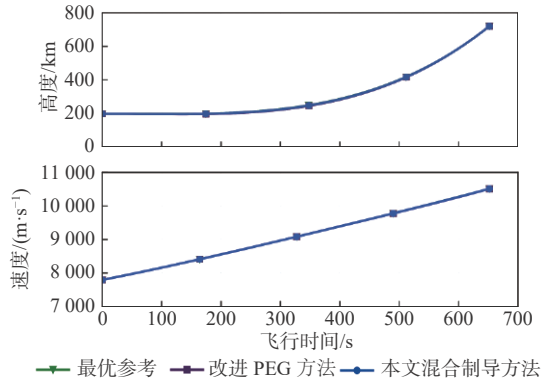


图 4 控制量曲线

Fig. 4 Control variable curve

消耗更多燃料。上述结果表明, 本文混合制导方法具有较好的入轨精度和燃料最优性, 能够满足一些工程应用的需求。

为验证本文混合制导方法综合应用 2 种轨迹规划方法的意义, 与纯二阶锥规划方法(即不使用邻近最优修正初始参考)进行了对比。混合制导方法和纯二阶锥规划方法均能完成上述地月转移入轨任务, 精度在同一数量级。两者的计算时间如图 5 所示。

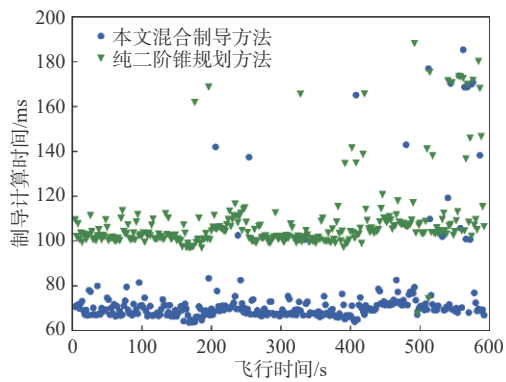


图 5 与纯二阶锥规划方法的计算时间对比

Fig. 5 Comparison of computation times with pure second-order conic programming methods

本文混合制导方法在 296 个制导周期中的平均计算时间为 75.353 ms, 与纯二阶锥规划方法(108.691 ms)相比, 平均计算时间下降 30.67%。表明本文混合制导策略有效, 应用高计算效率的邻近最优方法降低多源干扰影响后, 二阶锥规划方法能够快速迭代收敛。这一方面能够降低箭载计算机的计算负载, 另一方面允许更短的制导周期, 从而进一步提升对多源干扰的适应能力。

5.2 对动力故障的适应能力

为验证本文混合制导方法对动力故障的适应能力, 在推力和秒耗量下降 40%、45%、50% 的 3 种工况且存在未知 2% 偏差的情况下进行了仿真测试, 结果如图 6、图 7 和表 1 所示。

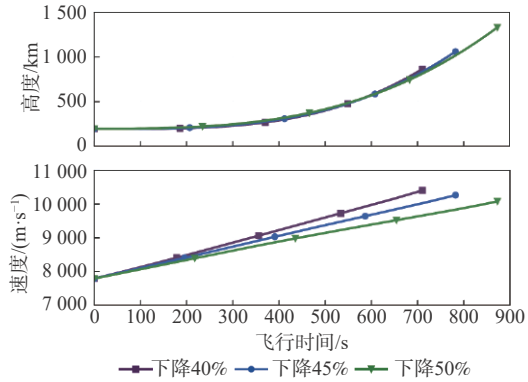


图 6　推力进一步下降时的高度与速度曲线

Fig. 6　Altitude-versus-speed curve as thrust decreases further

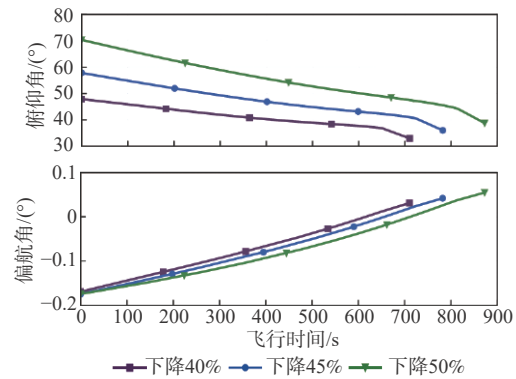


图 7　推力进一步下降时的控制量曲线

Fig. 7　Control variable curve as thrust decreases further

表 1　不同动力故障下的入轨精度

Table 1　Orbital insertion accuracy under different cases

工况	偏心率偏差	轨道倾角偏差/(°)	升交点赤经偏差/(°)	近地点幅角偏差/(°)
下降40%	0.000 200	0.000 006	0.000 015	−0.004 397
下降45%	0.000 270	0.000 005	0.000 016	−0.004 574
下降50%	0.000 373	0.000 004	0.000 017	−0.004 923

本文混合制导方法在 3 种不同动力故障下均能够完成制导任务, 表 1 中的各项轨道要素偏差说明入轨精度无明显差异。与离线轨迹规划得到的最优参考对比, 推进剂消耗误差小于 18.200 kg。改进 PEG 在这 3 种动力故障下无法完成制导任务, 在推力下降 40% 和 45% 的情况下, 改进 PEG 分别于 652 s 和 546 s 处迭代发散, 单制导周期内迭代 20 次仍未收敛, 在推力 50% 的情况下, 初始迭代不能收敛。上述结果表明, 本文混合制导方法对该工况下的动力故障具有较强的适应能力。在现有制导方法无法适应 40% 及以上推力下降的情况下, 本文混合制导方法能够适应 50% 的推力下降。

6　结　论

1) 基于实时轨迹规划的混合轨迹规划方法成功解决了飞行时间在 650 s 左右的小推重比地月转移入轨问题, 且燃料消耗较现有制导方法低。

2) 本文提出的邻近最优校正方法有效降低了轨迹优化的迭代次数。使用邻近最优校正后, 制导计算时间从 108.691 ms 降低至 75.353 ms, 降幅达到 30.67%。

3) 本文混合制导方法的计算时间满足现有制导周期(2 s)的需求, 且存在改善空间。实际应用中可选择更短制导周期, 从而进一步提高对多源干扰的适应能力。

参考文献（References）

[1] 包为民. 航天智能控制技术让运载火箭“会学习”[J]. 航空学报, 2021, 42(11): 1-10.
Bao W M. Space intelligent control technology enables launch vehicle to “self-learning”[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(11): 1-10(in Chinese).

[2] Von Der Porten P, Ahmad N, Hawkins M, et al. Powered explicit guidance modifications and enhancements for space launch system Block-1 and Block-1B vehicles[C]//Proceedings of the 41st American Astronautical Society Guidance, Navigation, and Control Conference. Breckenridge: AAS, 2018: 136.

[3] Chandler D C, Smith I E. Development of the iterative guidance mode with its application to various vehicles and missions[J]. Journal of Spacecraft and Rockets, 1967, 4(7): 898-903.

[4] 吕新广, 宋征宇. 载人运载火箭迭代制导方法应用研究[J]. 载人航天, 2009, 15(1): 9-14.
Lyu X G, Song Z Y. Study of the iterative guidance engineering application to manned launch vehicle[J]. Manned Spaceflight, 2009, 15(1): 9-14(in Chinese).

[5] 茹家欣. 液体运载火箭的一种迭代制导方法[J]. 中国科学 (E 辑), 2009(4): 696-706.
Ru J X. An iterative guidance method for liquid launch vehicle[J]. Science in China (Series E), 2009(4): 696-706(in Chinese).

[6] 施国兴, 吕新广, 巩庆海. 满足多终端约束的二次曲线迭代制导方法研究[J]. 中国空间科学技术, 2018, 38(2): 24-31.
Shi G X, Lyu X G, Gong Q H. Research on quadratic curve IGM for multi-terminal constraints[J]. Chinese Space Science and Technology, 2018, 38(2): 24-31(in Chinese).

[7] 马宗占, 许志, 唐硕, 等. 一种改进的运载火箭迭代制导方法[J]. 航空学报, 2021, 42(2): 217-229.
Ma Z Z, Xu Z, Tang S, et al. Improved iterative guidance method for launch vehicles[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(2): 217-229(in Chinese).

[8] 张珂洋, 王鹏, 于琦, 等. 运载火箭助推段大攻角剖面高精度弹道解析解[J]. 北京航空航天大学学报, 2025, 51(10): 3524-3534.
Zhang K Y, Wang P, Yu Q, et al. Analytical solutions for ascent trajectory under high angle-of-attack conditions[J]. Journal of Beijing University of Aeronautics and Astronautics, 2025, 51(10): 3524-3534(in Chinese).

[9] Ahmad N, Hawkins M, Von Der Porten P, et al. Closed loop guidance trade study for space launch system Block-1B vehicle[C]//Proceedings of the AAS/AIAA Astrodynamics Specialist Conference.

San Diego: Univelt, 2019.

[10] 王大轶, 符方舟, 孟林智, 等. Research of Autonomous Control Technology for Deep Space Probes[J]. 深空探测学报, 2019, 6(4): 317-327.
Wang D Y, Fu F Z, Meng L Z, et al. Research of autonomous control technology for deep space probes[J]. Journal of Deep Space Exploration, 2019, 6(4): 317-327(in Chinese).

[11] 巩庆海, 宋征宇, 吕新广. 迭代制导在月面上升段的应用研究[J]. 载人航天, 2015, 21(3): 231-236.
Gong Q H, Song Z Y, Lyu X G. Study on application of iterative guidance in lunar ascent[J]. Manned Spaceflight, 2015, 21(3): 231-236(in Chinese).

[12] 宋春, 郭延宁, 郭敏文, 等. 火星上升器入轨段自适应迭代制导方案设计[J]. 深空探测学报 (中英文), 2024, 11(1): 31-39.
Song C, Guo Y N, Guo M W, et al. Design of adaptive iterative guidance scheme for Mars ascent vehicle orbiting phase[J]. Journal of Deep Space Exploration, 2024, 11(1): 31-39(in Chinese).

[13] Mchenry R L, Long A L, Cockrell B F, et al. Space shuttle ascent guidance, navigation, and control[J]. Journal of the Astronautical Sciences, 1979, 27: 1-38.

[14] Jaggers R F. Proposed powered explicit guidance thrust integrals derivation/implementation. Mission planning, mission analysis and software formulation[R]. Washington, D. C. : NASA, 1976.

[15] GOODMAN J. Roland jaggers and the development of space shuttle powered explicit guidance (PEG)[C]//Proceedings of the AIAA Scitech 2021 Forum. Reston: AIAA, 2021: 2021.

[16] Hawkins M, Ahmad N. Guidance modifications and enhancements for space launch system Block-1 in support of Artemis I and beyond[C]//Proceedings of the 2020 AAS/AIAA Astrodynamics Specialist Conference. San Diego: Univelt, 2021.

[17] 张洪华, 李骥, 于萍, 等. 嫦娥五号月面起飞上升制导导航与控制技术[J]. 中国科学 (技术科学), 2021, 51(8): 921-937.
Zhang H H, Li J, Yu P, et al. Guidance navigation and control technology for the lunar ascent vehicle of the Chang'e-5 mission[J]. SCIENTIA SINICA Technologica, 2021, 51(8): 921-937(in Chinese).

[18] Brown K R, Harrold E F, Johnson G W. Rapid optimization of multiple-burn rocket flights: NASA CR-1430[R]. Washington, D. C. : NASA, 1969.

[19] Lu P, Griffin B J, Dukeman G A, et al. Rapid optimal multiburn ascent planning and guidance[J]. Journal of Guidance, Control, and Dynamics, 2008, 31(6): 1656-1664.

[20] 陈虹坤, 谭述君, 崔梦真, 等. 大气层外运载火箭上升段轨迹规划的保辛伪谱序列凸优化方法[J]. 北京航空航天大学学报, 2026, 52(4): 1211-1220.
Chen H K, Tan S J, Cui M Z, et al. A symplectic pseudo-spectral successive convex optimization method for trajectory planning of ascent stage of exo-atmosphere launch vehicle[J]. Journal of Beijing University of Aeronautics and Astronautics, 2026, 52(4): 1211-1220(in Chinese).

[21] 张志国, 余梦伦, 耿光有, 等. 应用伪谱法的运载火箭在线制导方法研究[J]. 宇航学报, 2017, 38(3): 262-269.
Zhang Z G, Yu M L, Geng G Y, et al. Research on application of pseudo-spectral method in online guidance method for a launch vehicle[J]. Journal of Astronautics, 2017, 38(3): 262-269(in Chinese).

[22] Cheng X M, Li H F, Zhang R. Efficient ascent trajectory optimization using convex models based on the Newton-Kantorovich/Pseudospectral approach[J]. Aerospace Science and Technology, 2017, 66: 140-151.

[23] 张永华, 王铮, 王时雨, 等. 基于无迹模型预测静态规划的可回收火箭着陆段鲁棒制导方法[EB/OL]. (2024-09-14)[2025-05-28]. https://doi.org/10.13700/j.bh.1001-5965.2024.0491.
Zhang Y H, Wang Z, Wang S Y, et al. Robust guidance method for reusable rocket landing phase based on UMPSP[EB/OL]. (2024-09-14)[2025-05-28]. https://doi.org/10.13700/j.bh.1001-5965.2024.0491 (in Chinese).

[24] Hao Z M, Zhang R. Onboard real-time generation of launch vehicle abort orbits[J]. Journal of Guidance, Control, and Dynamics, 2021, 44(8): 1541-1549.

[25] Wang J W, Zhang R, Li H F. Onboard optimization of multi-arc trajectories with constraints on duration of arcs[J]. Acta Astronautica, 2022, 192: 434-442.

Hybrid guidance method for launch vehicle trans-lunar injection

GUO Kang[1, 2], LI Huifeng[1, 2, *], BAO Weimin[3]

(1. School of Astronautics, Beihang University, Beijing 100191, China;
2. State Key Laboratory of High-Efficiency Reusable Aerospace Transportation Technology, Beihang University, Beijing 100191, China;
3. China Aerospace Science and Technology Corporation, Beijing 100048, China)

Abstract: This paper addresses the critical guidance problem for launch vehicles performing trans-lunar injection missions with low thrust-to-weight ratios, where the extended flight duration and substantial multi-source uncertainties present formidable challenges to conventional closed-loop guidance methods. A novel hybrid guidance approach integrating second-order cone programming with neighboring optimal control is developed to overcome these challenges. The proposed methodology begins with establishing a comprehensive trans-lunar injection problem model that explicitly accounts for multi-source uncertainties. A hybrid guidance strategy based on real-time trajectory planning is formulated to handle these uncertainties effectively. The main innovation is the application of a real-time trajectory planning technique that uses second-order cone programming in conjunction with Gauss pseudospectral discretization. By synergistically applying the costate mapping theorem and exploiting the dual output of second-order cone programming, the method achieves high-precision costate estimation, which provides the essential foundation for the neighboring optimal control component. A trajectory correction method based on neighboring optimal control is developed to address situations where uncertainties cause significant trajectory deviations. This method generates high-quality initial references for the second-order cone programming solver, substantially improving convergence characteristics and computational efficiency. Extensive numerical experiments demonstrate the superiority of the proposed guidance method in trans-lunar injection missions. The computational performance shows remarkable improvement, with the average computation time reduced to 75.353 milliseconds per guidance cycle, representing a 30.67% reduction. In comparison to current methods, this hybrid guidance method's successful demonstration offers a novel and feasible technical pathway for trans-lunar injection missions.

Keywords: launch vehicle guidance; trans-lunar injection; real-time trajectory planning; second-order cone programming; neighboring optimal control

Received: 2025-06-06; **Accepted**: 2025-12-05; **Published Online**: 2025-12-17 10:49
URL: link.cnki.net/urlid/11.2625.V.20251216.1600.002
Foundation item: National Natural Science Foundation of China (92471204)
* **Corresponding author**. E-mail: lihuifeng@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0247

基于控制向量参数化的临近空间飞行器轨迹优化

王远卓，代洪华*，岳晓奎
(西北工业大学 航天学院，西安 710072)

摘 要：针对计算制导场景下临近空间飞行器的在线轨迹优化问题，提出了一种基于实际动力学方程控制向量参数化的轨迹优化方法。通过翻转 Radau 伪谱离散非线性动力学方程，建立了含初始和终端状态约束的轨迹优化问题。通过参数化控制向量，将各离散点处控制量的优化问题转化为几个参数向量的优化问题，简化了计算过程，并提升了计算效率。此外，针对参数幅值约束，融合凸规划技术提出了一种基于控制向量参数化的轨迹优化方法。考虑到中制导任务对终端飞行时间不敏感，将优化问题的自变量从时间转换为航程，进一步缩短了计算时间。仿真结果表明，所提方法能够有效满足计算制导中高精度高效率轨迹优化的任务需求。

关 键 词：临近空间飞行器；最优控制；轨迹优化；凸规划；控制向量参数化设计
中图分类号：V448.1
文献标志码：A **文章编号**：1001-5965(2026)10-3602-16

随着科学技术的迅猛发展，平台体系正加速迈向机械化、信息化与智能化深度融合的崭新阶段。未来，主战平台的战略定位、技术形态以及作战运用模式将迎来革命性的变革[1]。在进攻型空战场景下，有人/无人异构协同战斗机群需要建立区域制空，保护编队后方的战略/战术轰炸机实现载荷(临近空间飞行器)投放并精确打击高价值目标[2]。本文聚焦于临近空间滑翔飞行器的制导律研究，致力于为其飞行性能与攻击精度的提升提供坚实的理论支撑与创新的技术方案。

临近空间滑翔飞行器具有飞行速度快、突防能力强、航程远、气动特性复杂等特点[3]。围绕突防环境下飞行器的制导与轨迹规划，闫循良等[4]系统梳理了现有轨迹规划方法，并深入分析了不同方法的优缺点。围绕多临近空间飞行器“探测-打击-评估”协同任务需求，刘哲等[5]实现了时间协同约束下的轨迹规划任务。在助推及滑翔段分别以俯仰角及攻角为控制量，An 等[6]提出了一种面向临近空间飞行器弹道设计与优化的框架，并讨论了设计与优化结果之间的差异。在协同制导框架下，Yu 等[7]提出了一种融合多飞行约束的高超声速滑翔飞行器协同制导策略。

临近空间飞行器的初、末制导律一般分别采用比较成熟的方案制导与比例制导。中制导律的研究方案如下：基于自适应动态规划的中制导律[8]、基于在线轨迹优化的中制导律[9]、基于微分对策的中制导律[10]、数据驱动的中制导律[11]、基于启发式方法的中制导律[12]。未来的中制导任务将面临更高的需求：不仅需要精确控制飞行器的终端位置，还需控制终端速度的大小和方向。这要求制导系统必须通过法向与侧向过载 2 个控制量实现 6 个状态的精确控制。在机载计算资源受限的约束下，针对该典型欠驱动制导问题[13]，飞行器一般通过在线解算轨迹优化问题实现各终端状态的精确控制[14]。因此，本文主要研究中制导的在线轨迹优化方法。

收稿日期：2025-04-29；**录用日期**：2025-07-25；**网络出版时间**：2025-07-29 10:14
网络出版地址：link.cnki.net/urlid/11.2625.V.20250729.0951.001
基金项目：国家杰出青年科学基金 (52425212)；国家重点研发计划 (2021YFA0717100)；国家自然科学基金 (12072270,U2013206)
*通信作者. E-mail：hhdai@nwpu.edu.cn

引用格式：王远卓，代洪华，岳晓奎. 基于控制向量参数化的临近空间飞行器轨迹优化 [J]. 北京航空航天大学学报，2026，52（10）：3602-3617. Wang Y Z，Dai H H，Yue X K. Near-space vehicle trajectory optimization based on control vector parameterization [J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（10）：3602-3617（in Chinese）.

近年来, 许多学者对各种轨迹优化方法进行了系统的综述, 并深入分析了它们的优缺点[15-16]。轨迹优化方法主要分为直接法与间接法[17]。间接法基于最优控制理论中的哈密尔顿函数和一阶必要条件, 构建两点边值问题, 并通过解算该问题实现轨迹优化的目标[18]。基于间接打靶轨迹优化架构, 刘宇航等[19]融合人工智能机器学习技术实现了深空探测的高精度高效率轨迹优化。Nurre 和 Taheri[20]提出了一种改进的间接优化方法, 用于生成固定时间下燃料最优的日心轨迹。间接轨迹优化方法具有解算精度高、满足一阶必要条件等优点, 但是存在收敛半径小与计算效率低等缺点[21]。与间接法不同的是, 直接法首先将动力学方程离散化, 而后通过非线性规划方法解算[22]。面向大气层边缘高超声速飞行器补能-跳跃滑翔飞行, 雍恩米等[23]基于 Gauss 伪谱实现了分段轨迹设计。赵吉松等[24]融合配点法实现了临近空间飞行器助推-滑翔段的最优轨迹设计, 并分析了典型参数的敏感性。尽管上述直接轨迹优化方法能够获得精确的数值解, 但通常依赖于离线计算, 难以满足轨迹优化问题在线解算的实时性需求。

Lu 将轨迹优化的在线解算问题定义为计算制导问题[14,25]。近 20 年来, 出现了 3 种在线解算方法: 模型预测控制(model predictive control, MPC)、凸规划与模型预测静态规划(model predictive static programming, MPSP)[26]。MPC 广泛应用于航空航天制导控制领域, 是一种基于系统模型的先进控制策略, 通过在每个控制周期内预测系统的未来行为, 优化控制输入保证预测输出满足约束条件并最小化性能指标, 从而实现对复杂动力学系统的精确控制。Patel 和 Subbarao[27]融合可达性分析与 MPC, 保证了干扰和不确定条件下高超声速飞行器再入过程的安全性。面向控制能力更强的带襟翼火星登陆器, Engel 等[28]融合 MPC 提升了终端状态的控制精度。Açıkmeşe 及其团队成员对凸规划理论的发展做出了巨大贡献[29]。该方法借助梯度下降快速收敛到全局最优解, 是一种广泛使用的轨迹优化方法。通过凸规划技术, Hayner 等[30]成功实现了视线约束下六自由度无人机轨迹优化问题的快速精确解算。针对小行星着陆任务, Ma 等[31]将深度神经网络与序列凸规划相结合, 巧妙消除了雅可比矩阵复杂的计算过程, 提出了一种基于机器学习的轨迹优化方法。Benedikter 等[32]基于虚拟控制量与自适应信赖域, 依托融合无损凸化和序列凸化技术, 构建了一种适用于火箭上升段的高效凸规划轨迹优化架构。针对临近空间飞行器在线制导问题, Wang 和 Grant[33]基于现有序列凸规划技术, 通过二次规划方法设计最优反馈制导律, 实现了参考轨迹的精确跟踪。

Padhi 及其团队成员提出了基于偏差动力学方程的 MPSP 方法[26]。基于参数化后的控制向量, Dwivedi 等[34]提出了一种模型预测扩展控制(model predictive spread control, MPSC)方法。基于无迹 MPSP 方法, 张永华等[35]提出了一种风场干扰下用于火箭着陆的鲁棒制导方法。融合谱形式的控制量与小维权重矩阵, 谭一廷等[36]针对防空反导导弹的中制导律设计, 提出了一种广义准谱 MPSP 方法。基于临近空间滑翔飞行器俯仰力矩瞬时平衡假设, Ma 等[37]将这一动力学条件转化为路径约束, 进而提出了一种考虑俯仰运动准静态假设的快速 MPSP 轨迹优化方法。对于临近空间飞行器快速轨迹优化, Wang 和 Dai[9]基于信赖域约束提出了一种融合翻转 Radau 伪谱法(flipped Radau pseudospectral method, FRPM)的改进 MPSP 方法。2019 年, Pan 等[11]证明了上述 MPSP 方法是一种基于实际动力学方程的牛顿类轨迹优化方法。MPSP 方法采用解析递推架构, 无需借助非线性规划方法, 即可实现轨迹优化问题的快速准确解算。然而, 该方法的计算时间受到离散点数量的影响。如果离散点数量不足, 可能导致终端状态误差较大, 从而影响优化结果的准确性。

受上述 MPSP 方法启发, 笔者提出了基于实际动力学方程控制历史直接更新的准谱轨迹优化方法(quasi-spectral trajectory optimization method via direct control history update, QS-DCHU)[38]。受上述 MPSC 与 DCHU 方法启发, 本文进一步提出了一种面向计算制导基于实际动力学方程控制向量参数化的临近空间飞行器轨迹优化方法(direct control history update with spread control, DCHU-SC)。本文的创新点主要归纳为以下 2 点:

1) 通过参数化控制向量, 将各离散点处的控制量优化问题转化为参数优化问题, 降低了问题的复杂度, 并提升了计算效率。此外, 与大多数航空航天轨迹优化问题类似, 临近空间飞行器的中制导任务对终端飞行时间并不敏感。因此, 本文将以时间为自变量的轨迹优化问题转化为以航程为自变量的优化问题, 使状态量的个数从 6 个降低至 5 个, 有效降低了欠驱动程度, 进一步提升了优化效率。该方法具有解析递推架构, 易转化为 C++程序, 符合工程实践需求。因此, 面向计算制导, 本文方法能够同时满足实时性和精确性的需求。

2) 针对现有方法未能充分考虑参数幅值约束的局限, 本文通过融合凸规划技术提出了一种基于

参数化控制向量的轨迹优化方法，进一步改善了方法的适应性。

1 飞行器任务场景与动力学模型

本节首先建立了临近空间飞行器的任务场景，而后建立了以时间为自变量的动力学模型。针对以时间为自变量的轨迹优化问题存在终端飞行时间难以确定、可能无可行解的局限性，建立了以航程为自变量的动力学模型，降低了轨迹优化问题的复杂度，得到更加贴合工程实践的轨迹优化方法。

1.1 临近空间飞行器任务场景

未来高价值目标打击任务中，首先需要有人/无人战斗机群建立进攻型区域制空。在此基础上，编队后方的轰炸机群发射临近空间飞行器实现对目标的精确打击，其任务场景如图1所示。

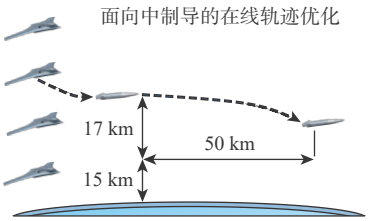


图1 飞行任务场景

Fig. 1 Flight task scenario

1.2 以时间/航程为自变量的动力学模型

飞行器以时间为自变量的动力学模型为[39]

$$\begin{cases}\dot{v}=g\left(n_x-\sin\gamma\right)\\ \dot{\gamma}=\dfrac{g}{v}\left(n_y-\cos\gamma\right)\\ \dot{\psi}=-\dfrac{gn_z}{v\cos\gamma}\\ \dot{r}_x=v\cos\gamma\cos\psi\\ \dot{r}_y=v\sin\gamma\\ \dot{r}_z=-v\cos\gamma\sin\psi\end{cases}\tag{1}$$

式中：6个状态量 v、γ、ψ、r_x、r_y、r_z 分别为速度、航迹角、航向角、航程、高度、侧向位置；g 为重力加速度；状态向量 $\boldsymbol{x}=[v,\gamma,\psi,r_x,r_y,r_z]^{\mathrm{T}}$；控制向量 $\boldsymbol{u}=[n_y,n_z]^{\mathrm{T}}$，$n_x$、$n_y$ 和 n_z 分别为法向过载和侧向过载。关于动力学模型的参数设定，详见笔者前期的研究工作[9]。

控制向量选取法向与侧向过载的原因在于：本文后续提出方法源于现有MPSP方法，通过符号递推解确保了快速准确解算。但是，此类方法已被证明属于一种改进的牛顿类方法。因此，该方法存在收敛半径小的问题，并且采用线性化近似模型时非线性程度不宜过高[11]。为平衡解算效率与现有临近空间飞行器模型中非线性较高的问题，通过提出的轨迹优化方法获取虚拟控制量（法向过载与侧向过载），并反向递推得到实际控制量（攻角与倾侧角）。反向递推公式详见文献[38]，3.1节将验证该思路的可行性及效果。

结合式(1)，航程为自变量的动力学方程为

$$\begin{cases}\dfrac{\mathrm{d}v}{\mathrm{d}r_x}=\dfrac{g\left(n_x-\sin\gamma\right)}{v\cos\gamma\cos\psi}\\ \dfrac{\mathrm{d}\gamma}{\mathrm{d}r_x}=\dfrac{g\left(n_y-\cos\gamma\right)}{v^2\cos\gamma\cos\psi}\\ \dfrac{\mathrm{d}\psi}{\mathrm{d}r_x}=-\dfrac{gn_z}{v^2\cos^2\gamma\cos\psi}\\ \dfrac{\mathrm{d}r_y}{\mathrm{d}r_x}=\dfrac{\tan\gamma}{\cos\psi}\\ \dfrac{\mathrm{d}r_z}{\mathrm{d}r_x}=-\tan\psi\end{cases}\tag{2}$$

鉴于中段飞行轨迹的高度不一定是单调的，使用航程作为自变量。考虑到滑翔飞行器实际的任务需求，中制导段航程可假设单调递增，因此航程可选取为自变量。

以时间为自变量的轨迹优化问题需要通过法向与侧向过载2个控制量实现6个终端状态的控制；与之不同的是，以航程为自变量的优化问题只需控制5个终端状态。因此，以航程为自变量建立动力学方程降低了原优化问题的维度，更有利于在线解算轨迹优化问题[40-41]。

1.3 轨迹优化问题

首先，轨迹优化问题的性能指标将在2.2节中详细给出；然后，该问题需要满足初始状态约束和固定的终端状态约束；最后，系统动力学约束需严格遵循式(1)或式(2)。

2 基于控制向量参数化的轨迹优化

基于实际动力学方程，本节提出了一种控制向量参数化的轨迹优化方法。首先，使用翻转Radau伪谱离散动力学方程；然后，通过控制向量参数化实现了轨迹优化问题的精确高效解算；最后，针对参数幅值约束，融合凸规划技术实现了轨迹优化问题的求解。值得注意的是，本节中的推导过程以时间为自变量。若以航程为自变量，只需将所有公式中的时间 t 换成航程 r_x 即可。

2.1 翻转Radau伪谱离散

动力学方程式(1)简化表达为

$$\dot{\boldsymbol{x}}=\boldsymbol{h}\left(\boldsymbol{x},\boldsymbol{u},t\right)\tag{3}$$

式中：$\boldsymbol{h}$ 为动力学方程向量；$\boldsymbol{x}$ 和 $\boldsymbol{u}$ 分别为状态和控制向量；t 为当前时间。

在参考轨迹附近对式(3)进行一阶泰勒展开[38]：

$$\dot{\boldsymbol{x}}\approx\boldsymbol{H}=\boldsymbol{h}^{\mathrm{ref}}+\boldsymbol{h}_{\boldsymbol{x}}^{\mathrm{ref}}\left(\boldsymbol{x}-\boldsymbol{x}^{\mathrm{ref}}\right)+\boldsymbol{h}_{\boldsymbol{u}}^{\mathrm{ref}}\left(\boldsymbol{u}-\boldsymbol{u}^{\mathrm{ref}}\right)\tag{4}$$

式中：$\boldsymbol{H}$ 为近似的动力学方程；“ref”表示参考轨迹；$\boldsymbol{h}_{\boldsymbol{x}}^{\mathrm{ref}}$和$\boldsymbol{h}_{\boldsymbol{u}}^{\mathrm{ref}}$分别表示动力学方程向量对状态和控

制向量的偏导数。

通过翻转 Radau 伪谱离散式 (4), 该过程与笔者先前研究中的方法类似, 理论详见文献 [9,38]。式 (5)～式 (7) 给出了该方法中的关键公式。

首先, 原时间间隔 $t\in[t_0, t_f]$ 通过归一化被转化至 $\tau\in[-1, 1]$:

$$\tau=\frac{2t}{t_f-t_0}-\frac{t_f+t_0}{t_f-t_0} \tag{5}$$

式中: t_0 和 t_f 分别为给定的初始和终端时间。

然后, 基于加权余量法则动力学约束表达为

$$\boldsymbol{0}=\sum_{i=0}^{N-1}\boldsymbol{D}_{ki}\mathrm{d}\boldsymbol{X}(\tau_i)-\frac{t_f-t_0}{2}\boldsymbol{H}(\mathrm{d}\boldsymbol{X}(\tau_k),\mathrm{d}\boldsymbol{U}(\tau_k),\tau_k;t_0,t_f) \tag{6}$$

式中: $\boldsymbol{X}$ 和 $\boldsymbol{U}$ 分别为离散后的状态和控制向量; τ_i 为节点 (i=0,1,…,K); τ_k 为配点 (k=1,2,…,K); $\boldsymbol{D}_{ki}$ 为微分矩阵中的元素, 由参考轨迹计算得到。

最后, 基于文献 [38] 中符号递推的思想, 通过归纳构建了各状态间的映射关系:

$$\begin{bmatrix}\boldsymbol{X}_0 & \boldsymbol{X}_1 & \cdots & \boldsymbol{X}_k & \cdots & \boldsymbol{X}_K\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}\boldsymbol{I} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \cdots & \boldsymbol{0}\\ \boldsymbol{D}_{10}\boldsymbol{I} & \boldsymbol{D}_{11}\boldsymbol{I}-\dfrac{t_f-t_0}{2}\boldsymbol{h}_{\boldsymbol{x}_1}^{\mathrm{ref}} & \cdots & \boldsymbol{D}_{1k}\boldsymbol{I} & \cdots & \boldsymbol{D}_{1K}\boldsymbol{I}\\ \vdots & \vdots & & \vdots & & \vdots\\ \boldsymbol{D}_{k0}\boldsymbol{I} & \boldsymbol{D}_{k1}\boldsymbol{I} & \cdots & \boldsymbol{D}_{kk}\boldsymbol{I}-\dfrac{t_f-t_0}{2}\boldsymbol{h}_{\boldsymbol{x}_k}^{\mathrm{ref}} & \cdots & \boldsymbol{D}_{kK}\boldsymbol{I}\\ \vdots & \vdots & & \vdots & & \vdots\\ \boldsymbol{D}_{K0}\boldsymbol{I} & \boldsymbol{D}_{K1}\boldsymbol{I} & \cdots & \boldsymbol{D}_{Kk}\boldsymbol{I} & \cdots & \boldsymbol{D}_{KK}\boldsymbol{I}-\dfrac{t_f-t_0}{2}\boldsymbol{h}_{\boldsymbol{x}_K}^{\mathrm{ref}}\end{bmatrix}^{-1}\begin{bmatrix}\boldsymbol{X}_0\\ \dfrac{t_f-t_0}{2}\left[\boldsymbol{h}_1^{\mathrm{ref}}-\boldsymbol{h}_{\boldsymbol{x}_1}^{\mathrm{ref}}\boldsymbol{X}_1^{\mathrm{ref}}+\boldsymbol{h}_{\boldsymbol{u}_1}^{\mathrm{ref}}\left(\boldsymbol{U}_1-\boldsymbol{U}_1^{\mathrm{ref}}\right)\right]\\ \vdots\\ \dfrac{t_f-t_0}{2}\left[\boldsymbol{h}_k^{\mathrm{ref}}-\boldsymbol{h}_{\boldsymbol{x}_k}^{\mathrm{ref}}\boldsymbol{X}_k^{\mathrm{ref}}+\boldsymbol{h}_{\boldsymbol{u}_k}^{\mathrm{ref}}\left(\boldsymbol{U}_k-\boldsymbol{U}_k^{\mathrm{ref}}\right)\right]\\ \vdots\\ \dfrac{t_f-t_0}{2}\left[\boldsymbol{h}_K^{\mathrm{ref}}-\boldsymbol{h}_{\boldsymbol{x}_K}^{\mathrm{ref}}\boldsymbol{X}_K^{\mathrm{ref}}+\boldsymbol{h}_{\boldsymbol{u}_K}^{\mathrm{ref}}\left(\boldsymbol{U}_K-\boldsymbol{U}_K^{\mathrm{ref}}\right)\right]\end{bmatrix} \tag{7}$$

2.2　基于符号递推的在线轨迹优化

本节提出了一种面向实际动力学方程具有符号递推解析架构的准确高效轨迹优化方法。定义式 (7) 动力学方程中的逆矩阵为 $\boldsymbol{P}$。据此, 式 (7) 可进一步整理为

$$\begin{bmatrix}\boldsymbol{X}_0\\ \boldsymbol{X}_1\\ \vdots\\ \boldsymbol{X}_k\\ \vdots\\ \boldsymbol{X}_K\end{bmatrix}=\begin{bmatrix}\boldsymbol{P}_{00} & \boldsymbol{P}_{01} & \cdots & \boldsymbol{P}_{0k} & \cdots & \boldsymbol{P}_{0K}\\ \boldsymbol{P}_{10} & \boldsymbol{P}_{11} & \cdots & \boldsymbol{P}_{1k} & \cdots & \boldsymbol{P}_{1K}\\ \vdots & \vdots & & \vdots & & \vdots\\ \boldsymbol{P}_{k0} & \boldsymbol{P}_{k1} & \cdots & \boldsymbol{P}_{kk} & \cdots & \boldsymbol{P}_{kK}\\ \vdots & \vdots & & \vdots & & \vdots\\ \boldsymbol{P}_{K0} & \boldsymbol{P}_{K1} & \cdots & \boldsymbol{P}_{Kk} & \cdots & \boldsymbol{P}_{KK}\end{bmatrix}\cdot\begin{bmatrix}\boldsymbol{X}_0\\ \dfrac{t_f-t_0}{2}\left(\boldsymbol{h}_1^{\mathrm{ref}}-\boldsymbol{h}_{\boldsymbol{x}_1}^{\mathrm{ref}}\boldsymbol{X}_1^{\mathrm{ref}}+\boldsymbol{h}_{\boldsymbol{u}_1}^{\mathrm{ref}}\left(\boldsymbol{U}_1-\boldsymbol{U}_1^{\mathrm{ref}}\right)\right)\\ \vdots\\ \dfrac{t_f-t_0}{2}\left(\boldsymbol{h}_k^{\mathrm{ref}}-\boldsymbol{h}_{\boldsymbol{x}_k}^{\mathrm{ref}}\boldsymbol{X}_k^{\mathrm{ref}}+\boldsymbol{h}_{\boldsymbol{u}_k}^{\mathrm{ref}}\left(\boldsymbol{U}_k-\boldsymbol{U}_k^{\mathrm{ref}}\right)\right)\\ \vdots\\ \dfrac{t_f-t_0}{2}\left(\boldsymbol{h}_K^{\mathrm{ref}}-\boldsymbol{h}_{\boldsymbol{x}_K}^{\mathrm{ref}}\boldsymbol{X}_K^{\mathrm{ref}}+\boldsymbol{h}_{\boldsymbol{u}_K}^{\mathrm{ref}}\left(\boldsymbol{U}_K-\boldsymbol{U}_K^{\mathrm{ref}}\right)\right)\end{bmatrix} \tag{8}$$

式中: $\boldsymbol{P}_{kk}$ 与式 (7) 中逆矩阵的对应元素保持一致。

基于式 (8), 可进一步归纳递推关系:

$$\boldsymbol{X}_k=\boldsymbol{P}_{k0}\boldsymbol{X}_0+\frac{t_f-t_0}{2}\cdot\sum_{i=1}^{K}\boldsymbol{P}_{ki}\left[\boldsymbol{h}_i^{\mathrm{ref}}-\boldsymbol{h}_{\boldsymbol{x}_i}^{\mathrm{ref}}\boldsymbol{X}_i^{\mathrm{ref}}+\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\left(\boldsymbol{U}_i-\boldsymbol{U}_i^{\mathrm{ref}}\right)\right] \tag{9}$$

通过式 (9), 终端时刻的状态转移方程为

$$\boldsymbol{X}_K=\boldsymbol{P}_{K0}\boldsymbol{X}_0+\frac{t_f-t_0}{2}\cdot\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\left[\boldsymbol{h}_i^{\mathrm{ref}}-\boldsymbol{h}_{\boldsymbol{x}_i}^{\mathrm{ref}}\boldsymbol{X}_i^{\mathrm{ref}}+\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\left(\boldsymbol{U}_i-\boldsymbol{U}_i^{\mathrm{ref}}\right)\right] \tag{10}$$

面向计算制导, 为在线高精度解算轨迹优化问题, 将各离散点处控制向量的优化问题转化为几个参数向量的优化问题, 降低了维度与复杂性, 提升了解算效率。参数化后的控制向量表达为

$$\boldsymbol{U}_i=\boldsymbol{p}_{N+1}\tau_i^N+\boldsymbol{p}_N\tau_i^{N-1}+\cdots+\boldsymbol{p}_{n+1}\tau_i^n+\cdots+\boldsymbol{p}_1 \tag{11}$$

式中: $\boldsymbol{p}$ 为参数矩阵 ($\boldsymbol{p}=[\boldsymbol{p}_{N+1}\quad\boldsymbol{p}_N\quad\cdots\quad\boldsymbol{p}_1]$), 由 N+1 个待优化的参数列向量组成。

控制历史曲线设计为式 (11) 的原因为: 多项式的形式不仅确保了曲线的连续性和光滑性, 还为后续的内环姿态跟踪控制系统提供了良好的基础。此外, 光滑的控制曲线决定了光滑的轨迹曲线, 从而减小了阻力、增加了航程、降低了结构应力, 进一步提升了飞行器的整体性能和可靠性。

为了简化推导, 令 N=2:

$$\boldsymbol{U}_i=\boldsymbol{a}\tau_i^2+\boldsymbol{b}\tau_i+\boldsymbol{c} \tag{12}$$

式中: $\boldsymbol{a}$、$\boldsymbol{b}$、$\boldsymbol{c}$ 分别对应离散轨迹优化问题中 3 个待优化的参数向量。因此, 轨迹优化问题的优化变量由各离散点处的控制向量转化为 3 个参数向量。式 (12) 呈现出典型的抛物线特征, $\boldsymbol{a}$ 决定了抛物线的开口大小及方向; $\boldsymbol{a}$ 与 $\boldsymbol{b}$ 共同决定了抛物线对称轴的位置; $\boldsymbol{c}$ 决定了初始 0 时刻控制量的大小。

为保证控制曲线的光滑性，二次型离散性能指标J设计为

$$J=\frac{t_{\mathrm{f}}-t_0}{4}\left(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{K}_{\boldsymbol{a}}\boldsymbol{a}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{K}_{\boldsymbol{b}}\boldsymbol{b}+\boldsymbol{c}^{\mathrm{T}}\boldsymbol{K}_{\boldsymbol{c}}\boldsymbol{c}\right) \tag{13}$$

式中：$\boldsymbol{K}_{\boldsymbol{a}}$、$\boldsymbol{K}_{\boldsymbol{b}}$与$\boldsymbol{K}_{\boldsymbol{c}}$分别为与3个参数向量对应的惩罚矩阵。

为证明通过式(13)进行迭代的收敛性，可将其改写为

$$\begin{aligned}J=\frac{t_{\mathrm{f}}-t_0}{4}&\left[(\boldsymbol{a}_{\mathrm{ref}}-\mathrm{d}\boldsymbol{a})^{\mathrm{T}}\boldsymbol{K}_{\boldsymbol{a}}(\boldsymbol{a}_{\mathrm{ref}}-\mathrm{d}\boldsymbol{a})+(\boldsymbol{b}_{\mathrm{ref}}-\mathrm{d}\boldsymbol{b})^{\mathrm{T}}\cdot\right.\\&\left.\boldsymbol{K}_{\boldsymbol{b}}(\boldsymbol{b}_{\mathrm{ref}}-\mathrm{d}\boldsymbol{b})+(\boldsymbol{c}_{\mathrm{ref}}-\mathrm{d}\boldsymbol{c})^{\mathrm{T}}\boldsymbol{K}_{\boldsymbol{c}}(\boldsymbol{c}_{\mathrm{ref}}-\mathrm{d}\boldsymbol{c})\right]\end{aligned} \tag{14}$$

式中："ref"表示参考值；设j表示$\boldsymbol{a}$、$\boldsymbol{b}$、$\boldsymbol{c}$向量中的某一个分量，dj表示偏差值，值得注意的是，需要j与dj异号：

$$\frac{\partial J}{\partial j}>0,\frac{\partial J}{\partial \mathrm{d}j}<0\rightarrow\begin{cases}J\searrow,j\searrow,\mathrm{d}j\nearrow\\J\nearrow,j\nearrow,\mathrm{d}j\searrow\end{cases} \tag{15}$$

通过式(15)即可证明轨迹收敛[9]。

结合式(10)中的终端状态转移方程与式(13)中的性能指标，增广性能指标$\overline{J}$表达为

$$\begin{aligned}\overline{J}=&\frac{t_{\mathrm{f}}-t_0}{4}\left(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{K}_{\boldsymbol{a}}\boldsymbol{a}+\boldsymbol{b}^{\mathrm{T}}\boldsymbol{K}_{\boldsymbol{b}}\boldsymbol{b}+\boldsymbol{c}^{\mathrm{T}}\boldsymbol{K}_{\boldsymbol{c}}\boldsymbol{c}\right)+\boldsymbol{\lambda}^{\mathrm{T}}\cdot\\&\left\{\boldsymbol{X}_K-\boldsymbol{P}_{K0}\boldsymbol{X}_0-\frac{t_{\mathrm{f}}-t_0}{2}\cdot\right.\\&\left.\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\left[\boldsymbol{h}_i^{\mathrm{ref}}-\boldsymbol{h}_{\boldsymbol{x}_i}^{\mathrm{ref}}\boldsymbol{X}_i^{\mathrm{ref}}+\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\left(\boldsymbol{a}\tau_i^2+\boldsymbol{b}\tau_i+\boldsymbol{c}-\boldsymbol{U}_i^{\mathrm{ref}}\right)\right]\right\}\end{aligned} \tag{16}$$

式中：$\boldsymbol{\lambda}$为拉格朗日乘子向量。

不考虑参数幅值约束，3个参数向量的一阶最优必要条件为

$$\frac{\partial\overline{J}}{\partial\boldsymbol{a}}=\frac{t_{\mathrm{f}}-t_0}{2}\boldsymbol{K}_{\boldsymbol{a}}\boldsymbol{a}-\frac{t_{\mathrm{f}}-t_0}{2}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\tau_i^2\right)^{\mathrm{T}}\quad\boldsymbol{\lambda}=\boldsymbol{0} \tag{17}$$

$$\frac{\partial\overline{J}}{\partial\boldsymbol{b}}=\frac{t_{\mathrm{f}}-t_0}{2}\boldsymbol{K}_{\boldsymbol{b}}\boldsymbol{b}-\frac{t_{\mathrm{f}}-t_0}{2}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\tau_i\right)^{\mathrm{T}}\quad\boldsymbol{\lambda}=\boldsymbol{0} \tag{18}$$

$$\frac{\partial\overline{J}}{\partial\boldsymbol{c}}=\frac{t_{\mathrm{f}}-t_0}{2}\boldsymbol{K}_{\boldsymbol{c}}\boldsymbol{c}-\frac{t_{\mathrm{f}}-t_0}{2}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\right)^{\mathrm{T}}\quad\boldsymbol{\lambda}=\boldsymbol{0} \tag{19}$$

通过式(17)～式(19)，可以得到

$$\boldsymbol{a}=\boldsymbol{\lambda}\boldsymbol{K}_{\boldsymbol{a}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\tau_i^2\right)^{\mathrm{T}} \tag{20}$$

$$\boldsymbol{b}=\boldsymbol{\lambda}\boldsymbol{K}_{\boldsymbol{b}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\tau_i\right)^{\mathrm{T}} \tag{21}$$

$$\boldsymbol{c}=\boldsymbol{\lambda}\boldsymbol{K}_{\boldsymbol{c}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\right)^{\mathrm{T}} \tag{22}$$

结合式(12)与式(20)～式(22)：

$$\begin{aligned}\boldsymbol{U}_i=&\left[\tau_i^2\boldsymbol{K}_{\boldsymbol{a}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\tau_i^2\right)^{\mathrm{T}}+\tau_i\boldsymbol{K}_{\boldsymbol{b}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\tau_i\right)^{\mathrm{T}}+\right.\\&\left.\boldsymbol{K}_{\boldsymbol{c}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\right)^{\mathrm{T}}\right]\boldsymbol{\lambda}\end{aligned} \tag{23}$$

借助式(23)，式(10)进一步转化为

$$\begin{aligned}\boldsymbol{X}_K=&\boldsymbol{P}_{K0}\boldsymbol{X}_0+\frac{t_{\mathrm{f}}-t_0}{2}\cdot\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\\&\left(\boldsymbol{h}_i^{\mathrm{ref}}-\boldsymbol{h}_{\boldsymbol{x}_i}^{\mathrm{ref}}\boldsymbol{X}_i^{\mathrm{ref}}+\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\left\{\left[\tau_i^2\boldsymbol{K}_{\boldsymbol{a}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\tau_i^2\right)^{\mathrm{T}}+\right.\right.\right.\\&\tau_i\boldsymbol{K}_{\boldsymbol{b}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\tau_i\right)^{\mathrm{T}}+\\&\left.\left.\left.\boldsymbol{K}_{\boldsymbol{c}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\right)^{\mathrm{T}}\right]\left(\boldsymbol{\lambda}-\boldsymbol{U}_i^{\mathrm{ref}}\right)\right\}\right)\end{aligned} \tag{24}$$

而后，通过式(24)解算符号形式的向量$\boldsymbol{\lambda}$：

$$\begin{aligned}\boldsymbol{\lambda}=&\left\{\frac{t_{\mathrm{f}}-t_0}{2}\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\left[\tau_i^2\boldsymbol{K}_{\boldsymbol{a}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\tau_i^2\right)^{\mathrm{T}}+\right.\right.\\&\left.\left.\tau_i\boldsymbol{K}_{\boldsymbol{b}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\tau_i\right)^{\mathrm{T}}+\boldsymbol{K}_{\boldsymbol{c}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\right)^{\mathrm{T}}\right]\right\}^{-1}\cdot\\&\left\{\boldsymbol{X}_K-\boldsymbol{P}_{K0}\boldsymbol{X}_0-\frac{t_{\mathrm{f}}-t_0}{2}\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\left(\boldsymbol{h}_i^{\mathrm{ref}}-\boldsymbol{h}_{\boldsymbol{x}_i}^{\mathrm{ref}}\boldsymbol{X}_i^{\mathrm{ref}}-\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\boldsymbol{U}_i^{\mathrm{ref}}\right)\right\}\end{aligned} \tag{25}$$

根据式(23)与式(25)，各离散点处的控制历史直接更新为

$$\begin{aligned}\boldsymbol{U}_i=&\left[\tau_i^2\boldsymbol{K}_{\boldsymbol{a}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\tau_i^2\right)^{\mathrm{T}}+\tau_i\boldsymbol{K}_{\boldsymbol{b}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\tau_i\right)^{\mathrm{T}}+\right.\\&\left.\boldsymbol{K}_{\boldsymbol{c}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\right)^{\mathrm{T}}\right]\cdot\\&\left\{\frac{t_{\mathrm{f}}-t_0}{2}\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\left[\tau_i^2\boldsymbol{K}_{\boldsymbol{a}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\tau_i^2\right)^{\mathrm{T}}+\right.\right.\\&\left.\left.\tau_i\boldsymbol{K}_{\boldsymbol{b}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\tau_i\right)^{\mathrm{T}}+\boldsymbol{K}_{\boldsymbol{c}}^{-1}\left(\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\right)^{\mathrm{T}}\right]\right\}^{-1}\cdot\\&\left\{\boldsymbol{X}_K-\boldsymbol{P}_{K0}\boldsymbol{X}_0-\frac{t_{\mathrm{f}}-t_0}{2}\cdot\right.\\&\left.\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\left[\boldsymbol{h}_i^{\mathrm{ref}}-\boldsymbol{h}_{\boldsymbol{x}_i}^{\mathrm{ref}}\boldsymbol{X}_i^{\mathrm{ref}}-\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\boldsymbol{U}_i^{\mathrm{ref}}\right]\right\}\end{aligned} \tag{26}$$

根据上述推导, 本文基于实际动力学方程, 通过参数化控制向量, 构建了一种面向计算制导的高效精确轨迹优化方法。值得注意的是, 式 (4) 通过一阶泰勒展开得到线性形式的近似动力学方程与实际非线性动力学方程存在一定的误差。因此, 本文方法借鉴了序列凸规划或 MPSP 中序列迭代的思想, 实现了各终端状态的精确控制。

最后, 上述轨迹优化问题的截止条件为达到最大迭代次数或满足绝对误差容限:

$$s = \mathrm{Norm}\left(\left[\boldsymbol{x}_0 - \boldsymbol{x}_0^{\mathrm{ref}} \quad \cdots \quad \boldsymbol{x}_i - \boldsymbol{x}_i^{\mathrm{ref}} \quad \cdots \quad \boldsymbol{x}_K - \boldsymbol{x}_K^{\mathrm{ref}}\right]\right) < s_{\mathrm{abs}} \tag{27}$$

式中: Norm() 表示 2-范数; s 为相邻 2 次迭代轨迹间的距离; s_{abs} 为绝对误差容限。

2.3　融合凸规划考虑参数幅值约束的轨迹优化

2.2 节方法中 3 个一阶必要条件成立的前提是未考虑参数幅值约束。为有效处理该约束, 给出了下列凸规划(二次规划)问题。

参数幅值约束表达为

$$p_{\min}^{l} \leqslant p_n^{l} \leqslant p_{\max}^{l} \tag{28}$$

式中: l 表示第 n 个参数向量的第 l 个分量; $p_{\min}^{l}$ 和 $p_{\max}^{l}$ 分别为该分量的最小值和最大值。值得注意的是, 参数幅值约束和控制输入约束密切相关。因此, 在轨迹优化问题中引入参数幅值约束, 能够更精准地契合实际的中制导场景。

通过式 (10) 与式 (12), 建立动力学等式约束:

$$\begin{aligned}
&\boldsymbol{X}_K - \boldsymbol{P}_{K0}\boldsymbol{X}_0 - \frac{t_{\mathrm{f}} - t_0}{2}\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\left[\boldsymbol{h}_i^{\mathrm{ref}} - \boldsymbol{h}_{\boldsymbol{x}_i}^{\mathrm{ref}}\boldsymbol{X}_i^{\mathrm{ref}} - \boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\boldsymbol{U}_i^{\mathrm{ref}}\right] = \\
&\quad \frac{t_{\mathrm{f}} - t_0}{2}\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\left(\boldsymbol{a}\tau_i^2 + \boldsymbol{b}\tau_i + \boldsymbol{c}\right) = \\
&\quad \frac{t_{\mathrm{f}} - t_0}{2}\left[\sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\tau_i^2 \quad \sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\tau_i \quad \sum_{i=1}^{K}\boldsymbol{P}_{Ki}\boldsymbol{h}_{\boldsymbol{u}_i}^{\mathrm{ref}}\right]\begin{bmatrix}\boldsymbol{a}\\ \boldsymbol{b}\\ \boldsymbol{c}\end{bmatrix}
\end{aligned} \tag{29}$$

针对式 (13) 中的性能指标、式 (28) 中的参数幅值约束、式 (29) 中的等式约束及初始和终端状态约束, 通过凸规划(二次规划)理论实现优化问题的精确高效解算。

需要注意的是, 2.2 节方法通过符号递推解直接更新控制历史, 因此具备较高的计算效率, 适用于在线解算计算制导下的轨迹优化问题。尽管该方法未能有效处理参数幅值约束(控制输入约束), 但通过设计二次型性能指标, 能够有效保证控制曲线的光滑性。在一般情况下, 只要轨迹优化问题存在可行解, 通常不会触碰输入约束。3.3 节证明了该情况。

与之不同的是, 2.3 节方法因为考虑了额外的参数幅值约束, 不具备符号递推解, 仅能通过凸规划(二次规划)技术解算轨迹优化问题, 牺牲了一定的解算效率[42]。因此, 在实际应用中, 需结合战场态势动态评估并综合考虑, 从 2.2 节或本节方法中选取合适的轨迹优化方法。

3　仿真验证与方法比较

在一致的条件下, 本节首先验证了基于实际动力学方程控制向量参数化的临近空间飞行器轨迹优化方法, 并与现有通用伪谱最优控制软件(general pseudospectral optimal control software, GPOPS)轨迹优化方法、序列凸规划方法、基于偏差动力学方程的 MPSC 方法进行了对比。然后, 针对大部分航空航天轨迹优化问题对终端飞行时间不敏感而对计算效率敏感的特点, 将本文方法的自变量由时间改为航程, 降低了优化问题的复杂度, 并提升了优化效率。最后, 面向参数幅值约束, 对 2.2 节与 2.3 节中提出的 2 种方法进行了对比与分析。需要注意的是, 2.2 节方法未考虑参数幅值约束可构建符号递推解, 因此不需要使用求解器。与之不同, 2.3 节方法因为处理幅值约束而不满足式 (17)～式 (19) 中的一阶必要条件, 因此可采用 MATLAB 环境下的 CVXGEN 求解器解算凸规划问题。

对于临近空间飞行器受限的机载计算资源, 最大迭代次数设为 20; 绝对误差容限设为 10; 翻转 Radau 伪谱中离散点个数设为 84; $\boldsymbol{K}_a$、$\boldsymbol{K}_b$、$\boldsymbol{K}_c$ 这 3 个与参数向量对应的惩罚矩阵均设为单位阵; 初始控制历史猜想(初值)采用比例制导生成; 在仿真结果中, 第 7 次迭代曲线表示最终优化得到的轨迹(见图 2～图 9)。初始和终端状态约束设定如表 1 所示。

表 1　初始和终端状态约束

Table 1　Initial and terminal state constraints

状态约束	速度/(m·s^{-1})	航迹角/(°)	航向角/(°)	航程/km	高度/km	侧向位置/km
初始	1 870	2	−40	0	32	0
终端	1 647	−5	−48	35	31	35

在完成轨迹优化后, 需要通过开环积分验证本文方法的终端控制精度。开环积分采用欧拉积分, 步长设为 0.001 s。

3.1　以时间为自变量控制向量参数化的轨迹优化

不考虑参数幅值约束, 本节验证了以时间为自变量控制向量参数化的轨迹优化方法。首先, 对 2.2 节方法进行了验证与分析。然后, 将本文方法与现有 MPSC 方法[34]、基于 Radau 伪谱的非线性轨迹优化方法(, GPOPS)[43] 及序列凸规划方法进行对

比，说明了本文方法的有效性及意义。仿真中，终端飞行时间设为28.28 s，初始和终端状态约束如表1所示。图2～图4为轨迹曲线。

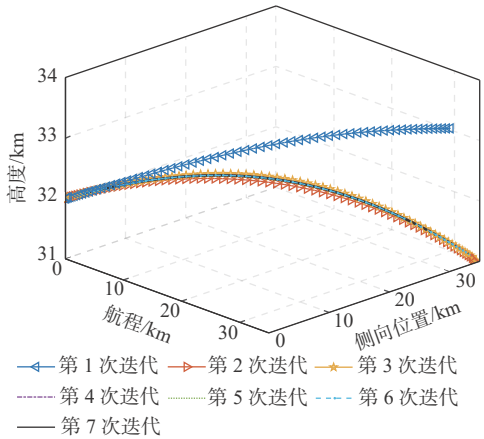


图2 航程-高度-侧向位置曲线

Fig. 2 Curves of downrange, altitude, and lateral position

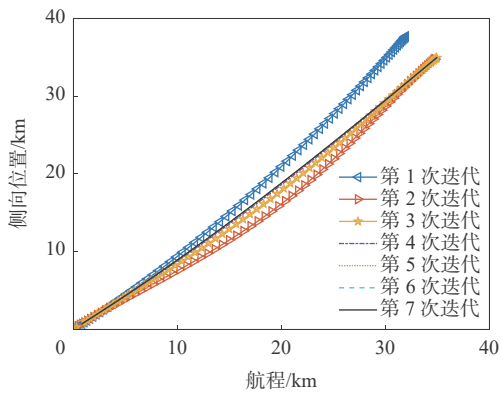


图3 航程-侧向位置曲线

Fig. 3 Curves of lateral position with downrange

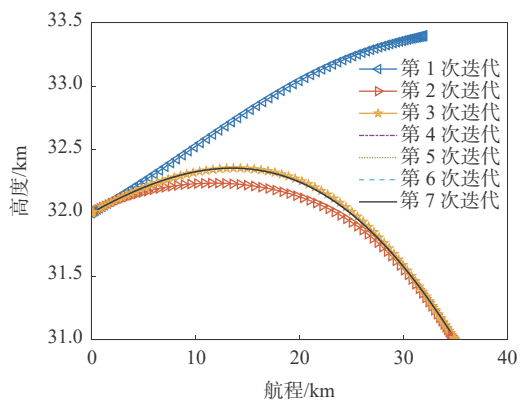


图4 航程-高度曲线

Fig. 4 Curves of altitude with downrange

图2～图4中的第1次迭代曲线表示初始控制历史猜想积分得到的轨迹。可以明显看出，初始猜想与最终得到的最优解存在较大误差。表明即使初值较差，本文方法凭借较强的鲁棒性仍能保证终端状态的收敛。此外，终端速度误差为-2.90×10^{-3} m/s；航迹角与航向角误差分别为-4.90×10^{-5}(°)与-2.60×10^{-4}(°)；三轴位置误差分别为0.52 m、0.14 m与0.46 m。因此，以上数据充分说明了本文方法具有较高的终端状态控制精度。

图5～图7为航迹角、航向角与性能指标随时间的变化曲线。

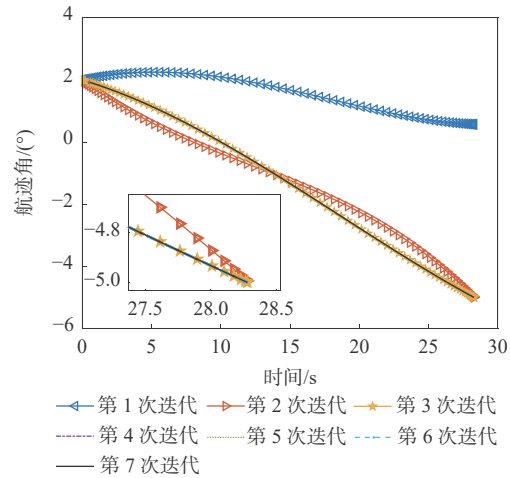


图5 时间-航迹角曲线

Fig. 5 Curves of flight-path angle with time

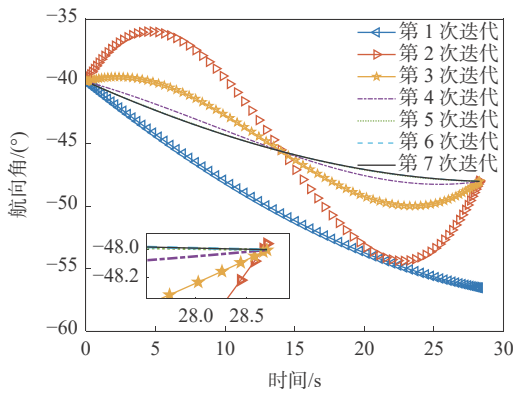


图6 时间-航向角曲线

Fig. 6 Curves of heading angle with time

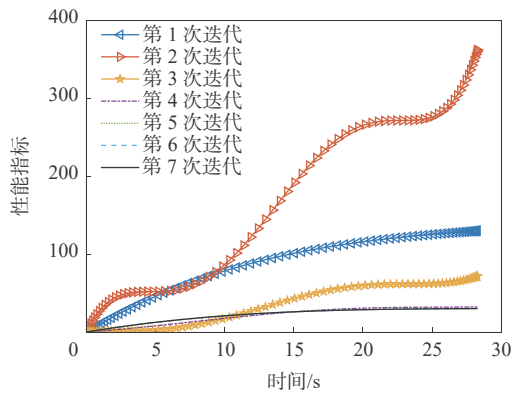


图7 时间-性能指标曲线

Fig. 7 Curves of performance index with time

图5表明，航迹角曲线在第3次迭代时收敛；图6和图7显示，航向角与性能指标曲线的收敛速度稍慢，直至第5次迭代才完成收敛。最终的性能

指标为 30.28。多次迭代的目的是在确保终端状态高精度控制的前提下降低性能指标。虽然该过程增加了计算时间, 但是本文方法在 MATLAB 环境中单次优化的时间约为 200 ms, 若转化成 C++程序, 推测耗时可进一步降低至 50 ms 左右[44]。即使迭代次数增加, 轨迹优化的计算时间也不会过多增加, 满足临近空间飞行器机载计算资源受限下的计算制导问题。

通过 7 次迭代, 3 个参数向量表示为

$$\begin{cases} \boldsymbol{a} = \begin{bmatrix} 0.287\,3 \\ -0.031\,5 \end{bmatrix} \\ \boldsymbol{b} = \begin{bmatrix} -0.134\,9 \\ -0.774\,0 \end{bmatrix} \\ \boldsymbol{c} = \begin{bmatrix} 0.135\,0 \\ 0.907\,5 \end{bmatrix} \end{cases} \tag{30}$$

法向/侧向过载曲线如图 8 和图 9 所示。

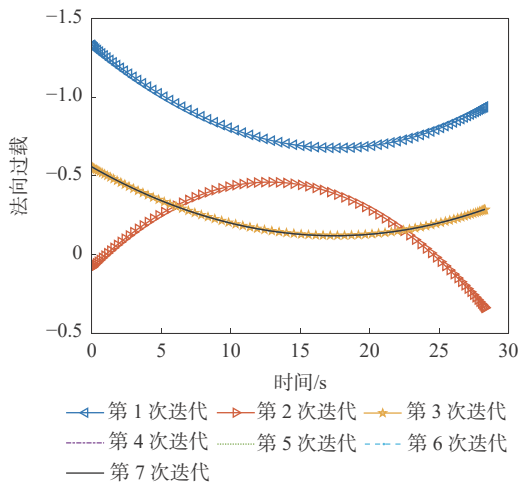


图 8　时间-法向过载曲线

Fig. 8　Curves of normal overload with time

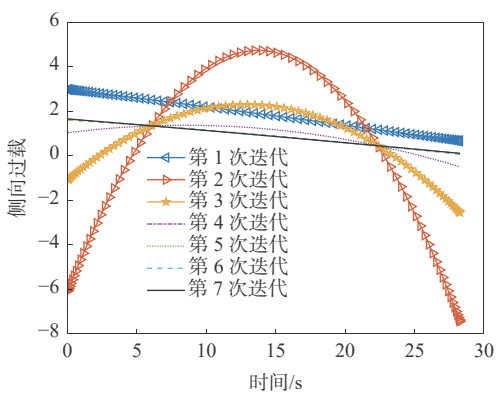


图 9　时间-侧向过载曲线

Fig. 9　Curves of lateral overload with time

通过式 (5) 与式 (12), 可得控制历史表示为

$$n_y(t) = 0.287\,3\left(\frac{t}{14.14}-1\right)^2 - 0.134\,9\left(\frac{t}{14.14}-1\right) + 0.135\,0 \tag{31}$$

$$n_z(t) = -0.031\,5\left(\frac{t}{14.14}-1\right)^2 - 0.774\,0\left(\frac{t}{14.14}-1\right) + 0.907\,5 \tag{32}$$

通过式 (31) 与式 (32) 可以看出, 控制曲线较为光滑, 未出现斜率及幅值过大的情况, 这一特性为后续的内环姿态跟踪控制系统提供了有利条件。

根据参考文献 [38], 实际控制量攻角/倾侧角曲线反向解算结果如图 10 和图 11 所示。

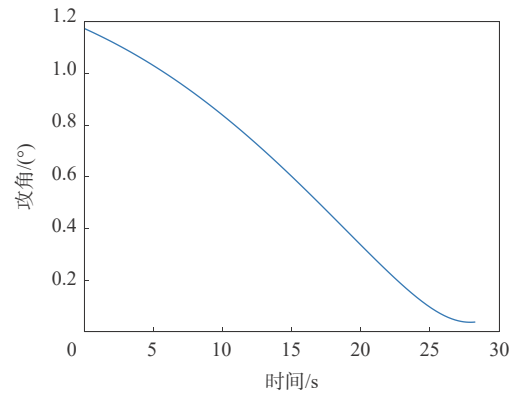


图 10　时间-攻角曲线

Fig. 10　Curves of attack angle with time

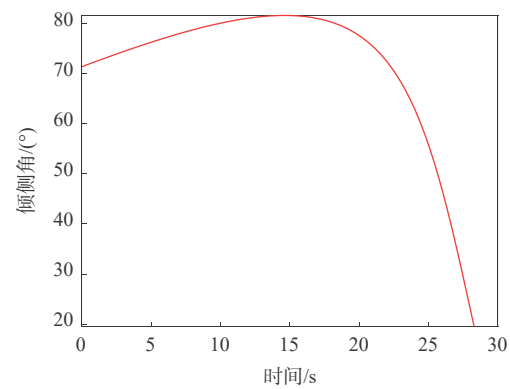


图 11　时间-倾侧角曲线

Fig. 11　Curves of heel angle with time

从图 10 和图 11 可以看出, 攻角曲线的幅值较小, 不会导致失速的现象; 鉴于倾侧角的范围为 [−90°, 90°], 虽然曲线存在细微的抖动, 但仍能满足控制需求。攻角与倾侧角曲线较为光滑的原因在于式 (31) 和式 (32) 中选取的控制历史曲线本身是光滑的。因此, 通过虚拟控制量(法向/侧向过载)反向解算实际控制量(攻角/倾侧角)的方案是可行的。

然后, 基于 MPSC、序列凸规划(sequential convex programming, SCP)与 GPOPS 这 3 种对比方法, 通过仿真说明本文方法(DCHU-SC)能够在计算资源受限的约束下实现终端状态的精确控制。因为协态映射定理的存在, 融合 Radau 伪谱与非线性规划的 GPOPS轨迹优化方法生成的解满足一阶最优必要条件, 可视为最优解。但是, 该方法使用了

大规模的非线性规划，一般离线计算，不适合在线解算轨迹优化问题。

MPSC方法的理论详见文献[28]。该方法中，3个待优化的参数向量更新方程为

$$\begin{bmatrix}\boldsymbol{a} & \boldsymbol{b} & \boldsymbol{c}\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}\boldsymbol{C}_z & \boldsymbol{D}_z & \boldsymbol{E}_z\end{bmatrix}^{-1}\left[\boldsymbol{B}_\lambda-\mathrm{d}\boldsymbol{X}_K\right] \tag{33}$$

式中：$\boldsymbol{C}_z$、$\boldsymbol{D}_z$、$\boldsymbol{E}_z$分别为与3个参数向量相关的矩阵；$\boldsymbol{B}_\lambda$表示通过参考控制历史与敏感度矩阵计算得到的参数矩阵；$\mathrm{d}\boldsymbol{X}_K$表示终端状态向量与参考值的偏差，$\mathrm{d}\boldsymbol{X}_K=\boldsymbol{X}_K-\boldsymbol{X}_K^{\mathrm{ref}}$。

通过$\mathrm{d}\boldsymbol{X}_K$的公式可以看出，该方法基于偏差动力学方程进行推导。与之不同，本文方法基于实际动力学方程进行推导。图12～图14为4种不同方法的轨迹优化结果。其中，括号内的数字表示离散点个数，该参数影响SCP、DCHU-SC及MPSC这3种方法的计算效率与精度。

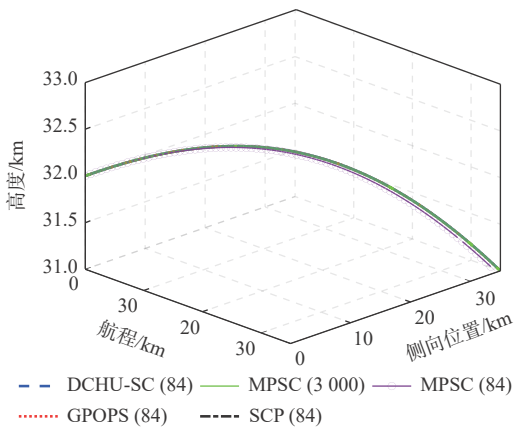


图12 4种方法的航程-高度-侧向位置曲线

Fig. 12 Curves of downrange, altitude, and lateral position for 4 methods

从图12～图14可以明显看出，相比于其他4条轨迹，采用84个离散点的MPSC方法存在较大误差。终端状态误差如表2所示。

从表2可以看出，仿真“MPSC (84)”的终端状态误差较大，位置误差超过了96 m，原因是终端速度未能精确控制。但是，仿真“MPSC (3 000)”不存在该问题，保证了较高的控制精度。产生该现象的原因为：当离散点个数从84个增加至3 000个，优化精度提升。但是，由于离散点个数直接决定轨迹优化所需的计算时间，随着个数的增加，解算时间也随之增加。在MATLAB环境下，仿真“MPSC (3 000)”的单轮轨迹优化时间约为5.5 s，推测在C++环境中约为1.8 s[44]。即使在初始控制历史猜想较好的情况下，考虑到每次轨迹优化大约需要7轮迭代，其在C++环境下的优化时间为12～13 s。临近空间飞行器的平均速度约为1 700 m/s，优化过程中大约飞行21 km，这意味着优化结果在实际飞行中无法直接使用。因此，现有基于偏差动力学方程的MPSC方法并不适用于机载计算资源受限下的飞行器在线轨迹优化任务。此外，与仿真“MPSC (84)”相比，仿真“SCP (84)”虽然降低了终端状态误差，但是位置误差仍大于16 m；同时，通过凸规划技术解算轨迹优化问题需进行非线性规划，牺牲了一定的解算效率[42]。

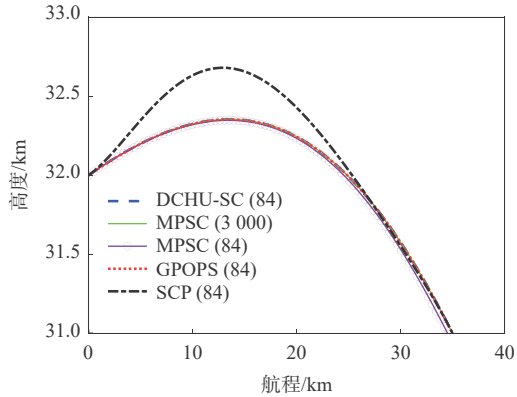


图13 4种方法的航程-高度曲线

Fig. 13 Curves of altitude with downrange for 4 methods

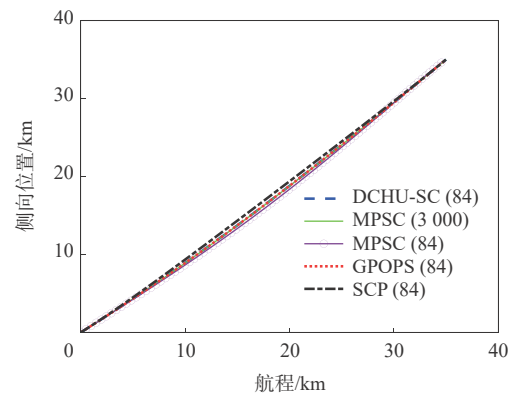


图14 4种方法的航程-侧向位置曲线

Fig. 14 Curves of lateral position with downrange for 4 methods

表2 终端状态误差

Table 2 Terminal state errors

终端状态方法	速度误差/(m·s^{-1})	航迹角误差/(°)	航向角误差/(°)	航程误差/m	高度误差/m	侧向位置误差/m
DCHU-SC (84)	-2.90×10^{-3}	-4.90×10^{-5}	-2.60×10^{-4}	0.52	0.14	0.46
MPSC (3 000)	0.10	-5.43×10^{-4}	1.96×10^{-3}	−0.41	−1.34	−0.10
MPSC (84)	−7.11	0.05	0.24	−76.23	−52.18	26.95
GPOPS(84)	−0.02	-4.32×10^{-6}	-4.18×10^{-6}	0.54	0.10	0.54
SCP (84)	0.02	0.02	0.03	16.25	2.12	3.26

与现有采用 3 000 个离散点的 MPSC 方法相比, 本文方法(DCHU-SC)仅采用 84 个离散点, 缩短了优化时间, 同时进一步提升了控制精度。此外, 本文方法生成的轨迹与 GPOPS 最优解基本重合, 从仿真角度验证了轨迹的最优性, 与 2.2 节中的理论推导高度契合。

上述分析表明, 本文方法仅需优化 3 个参数向量即可获得最优解, 降低了优化问题的复杂性, 并提升了解算效率。因此, 该方法在确保优化效率的同时, 能够实现准确的终端状态控制, 兼具高效性与精确性, 为面向计算制导的轨迹优化领域提供了一种创新且实用的解决方案。

3.2　以航程为自变量控制向量参数化的轨迹优化

由于大多数航空航天轨迹优化问题对终端飞行时间不敏感, 本节将验证以航程为自变量的轨迹优化效果。以航程为自变量而不是以时间为自变量的优点是: 以时间为自变量的轨迹优化问题需通过法向与侧向过载 2 个控制量实现终端速度、航迹角、航向角、航程、高度与侧向位置 6 个状态的精确控制; 与之不同, 以航程为自变量的优化问题需通过 2 个控制量实现终端速度、航迹角、航向角、高度与侧向位置 5 个状态的控制。因此, 相比于以时间为自变量, 以航程为自变量的优化问题复杂度低、收敛半径大。即使初始控制历史猜想较差, 轨迹也更容易收敛。

仿真初始设定如表 1 所示。终端速度误差为 −0.02 m/s; 航迹角与航向角误差分别为-6.50×10^{-4}(°)与-5.45×10^{-4}(°); 三轴位置误差分别为 3.04 m、−0.29 m 与 3.45 m。从终端状态误差的统计结果来看, 各状态均实现了高精度的控制。

然后, 说明以航程为自变量方法的优越性。由于控制量直接作用于式 (2) 中的航迹/航向角微分方程, 以上 2 个状态与控制量之间存在高度关联性。2 个角度的变化曲线如图 15 和图 16 所示。控制曲线如图 17 和图 18 所示。

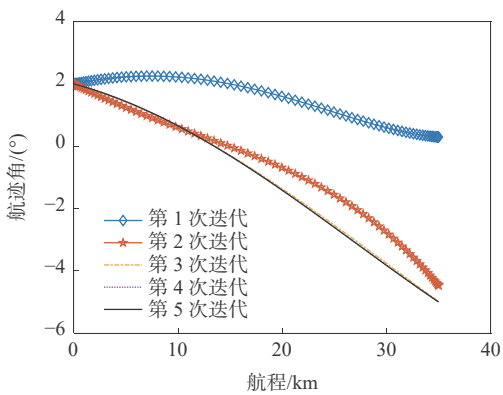


图 15　航程-航迹角曲线

Fig. 15　Curves of flight-path angle with downrange

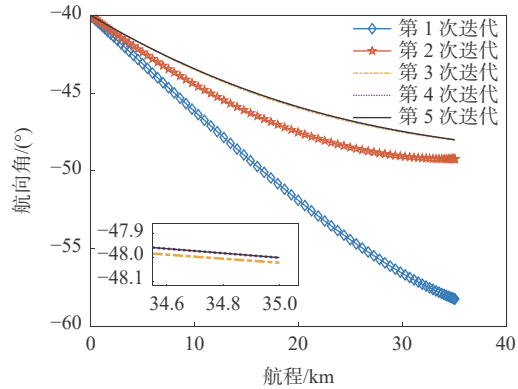


图 16　航程-航向角曲线

Fig. 16　Curves of heading angle with downrange

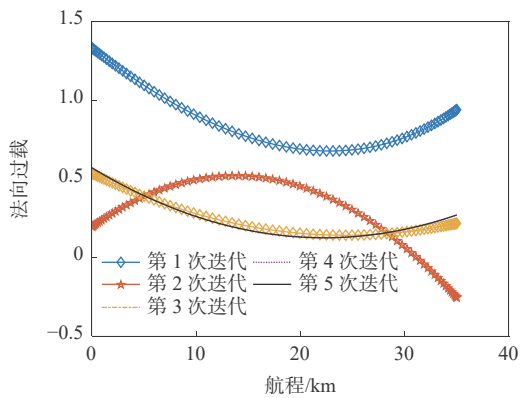


图 17　航程-法向过载曲线

Fig. 17　Curves of normal overload with downrange

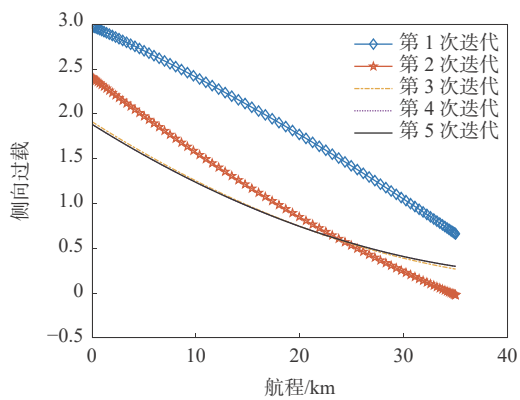


图 18　航程-侧向过载曲线

Fig. 18　Curves of lateral overload with downrange

与图 5、图 6、图 8 与图 9 中以时间为自变量的航迹角、航向角、过载曲线相比, 图 15～图 18 表明以航程为自变量的轨迹优化方法具有更快的收敛速度。此外, 本文方法生成的轨迹幅值较小、更为光滑, 表明该方法不依赖于较大的收敛半径, 进一步凸显了其性能的优越性。

图 19 表示性能指标的变化曲线。图 7 与图 19 的对比结果显示, 误差容限相同时, 以时间为自变量的曲线在第 5 次迭代时趋于收敛, 最终迭代次数为 7 次; 相比之下, 以航程为自变量的曲线在第 3

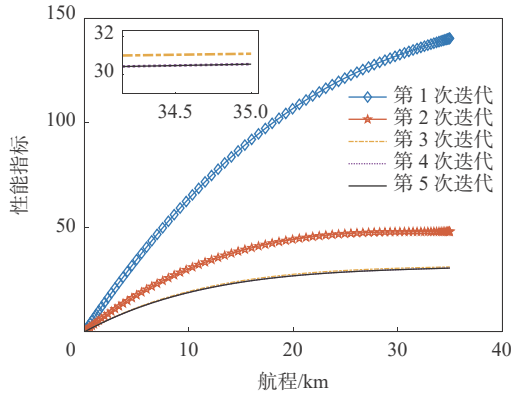


图 19 航程-性能指标曲线

Fig. 19 Curves of performance index with downrange

次迭代时已实现收敛，且最终的迭代次数为 5 次。需要注意，尽管图中显示轨迹已收敛，但后续的几次迭代仍然是必要的。这些额外的迭代能够进一步提升终端状态的控制精度，从而确保最终结果的准确性和可靠性。

为验证本文方法对初始状态摄动的鲁棒性，开展了 2 000 次蒙特卡罗模拟测试。其中，在表 1 初始状态的基础上，初始高度摄动范围为 [−300, 300] m，侧向位置摄动范围为 [−100, 100] m，速度摄动范围为 [−2, 2] m/s，航迹/航向角摄动范围为 [−1°, 1°]，期望的终端状态均保持不变。需要注意的是，未施加初始航程摄动的原因是航程为自变量。蒙特卡罗仿真结果如图 20～图 24 所示。

仿真结果表明，最大终端位置误差不超过 75 m，最大航迹角与航向角误差不超过 0.02°。此外，终端速度误差不超过 0.3 m/s。因此，以上数据表明各终端状态实现了高精度收敛，满足中制导结束时的精度要求。因此，在对终端飞行时间敏感度相对较低的航空航天轨迹优化场景中，采用航程作为自变量取代传统的时间自变量，能够增强迭代过程的鲁棒性、提升收敛速度与优化效率。在面向计算制导的快速轨迹优化问题中，实时性是至关重要的核心指标。本文

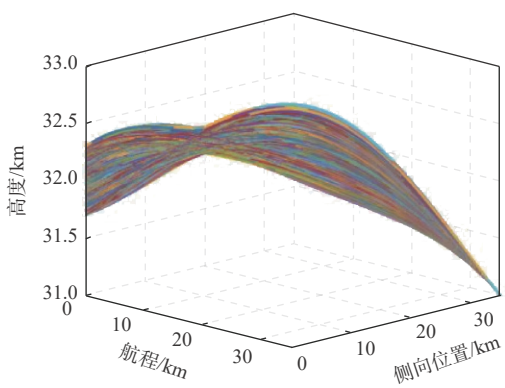


图 20 航程-高度-侧向位置曲线(初始状态摄动)

Fig. 20 Curves of downrange, altitude, and lateral position (initial state perturbations)

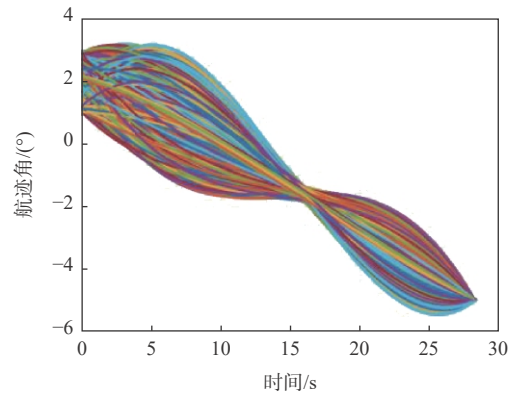


图 21 时间-航迹角曲线(初始状态摄动)

Fig. 21 Curves of flight-angle with time (initial state perturbations)

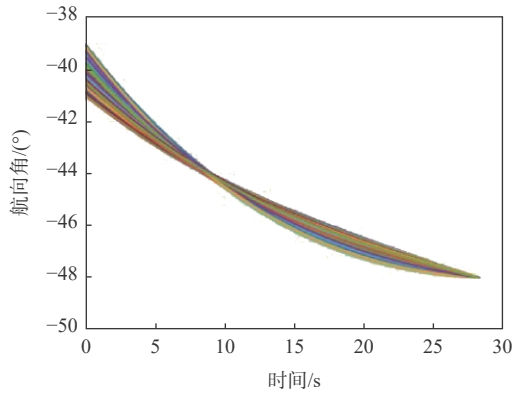


图 22 时间-航向角曲线(初始状态摄动)

Fig. 22 Curves of heading angle with time (initial state perturbations)

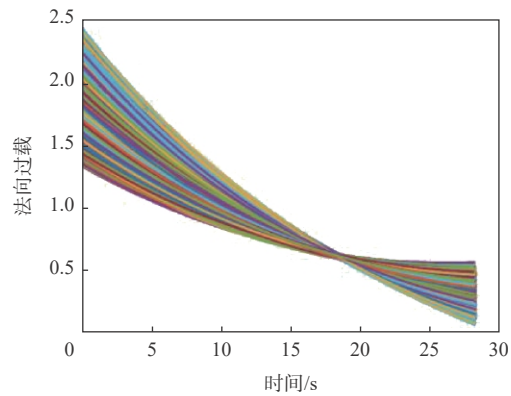


图 23 时间-法向过载曲线(初始状态摄动)

Fig. 23 Curves of normal overload with time (initial state perturbations)

方法能够缩短计算时间，这一改进具有重要意义。

针对不确定建模下的鲁棒性进行蒙特卡罗模拟。部分论文通过非线性无迹变换的思想实现了鲁棒轨迹优化的效果，但导致了问题规模的增大，延长了计算时间[45]。需要注意的是，轨迹优化方法普遍为开环方法，上述仿真已经证明本文方法具有

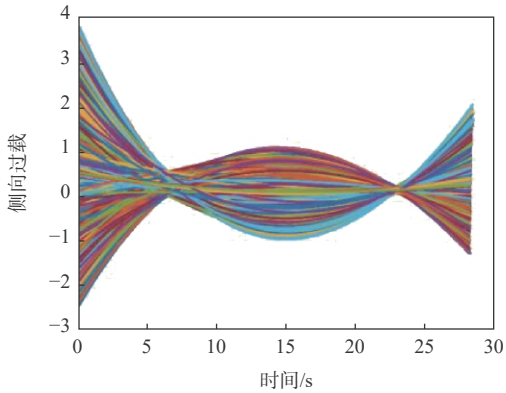


图 24　时间-侧向过载曲线(初始状态摄动)

Fig. 24　Curves of lateral overload with time (initial state perturbations)

快速性, 因此可通过在线多次计算实现终端状态的精确控制[38]。基于标准模型单次轨迹优化的输出, 仅验证不确定模型开环积分的对比结果。首先, 依据表 1 中的初始和终端状态约束构建参考轨迹。然后, 考虑到模型不确定性可能导致的建模误差, 对飞行器的阻力系数 C_D、参考面积 S 及质量 m 分别施加±5% 的摄动。在此基础上, 利用上述随机摄动模型, 通过 2 000 次蒙特卡罗模拟进行开环积分验证, 以评估本文方法对不确定模型的适应性和鲁棒性, 仿真结果如图 25 所示。

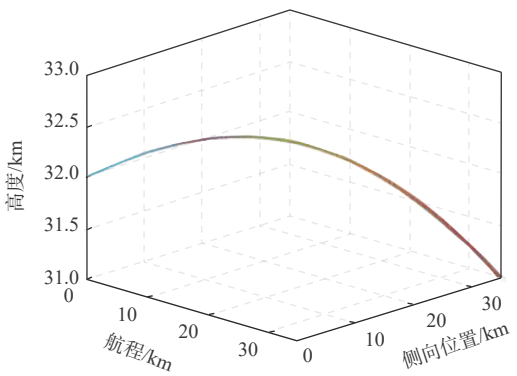


图 25　航程-高度-侧向位置曲线(模型摄动)

Fig. 25　Curves of downrange, altitude, and lateral position (model perturbations)

图 25 表明, 尽管存在模型摄动, 但本文提出的轨迹优化方法终端状态误差仍较小, 表现出良好的适应性及鲁棒性。由于该方法生成的控制曲线是最优且光滑的, 结合临近空间滑翔飞行器中制导段简单、稳定的气动条件, 其受摄动的影响较小, 因此仍能保持轨迹的稳定性。尽管在开环积分条件下, 终端状态误差受到摄动的影响有所增加, 但该方法具备快速计算的优势, 能够通过多次迭代优化, 实现高精度的终端状态控制[38]。

3.3　融合凸规划考虑参数幅值约束的轨迹优化

本节将对比 2.2 节与 2.3 节中 2 种不同轨迹优化方法的结果。其中, 2.2 节提出未考虑参数幅值约束, 但是式 (13) 中设计的二次型性能指标保证了控制曲线的光滑性。在轨迹优化问题存在可行解时, 该方法通常不会违反参数幅值约束。与之不同, 2.3 节方法基于凸规划理论, 成功实现了参数幅值约束的处理。

仿真以航程为自变量, 初始和终端状态约束如表 1 所示。3 个参数的幅值约束均设置为 [−5, 5]。其中, “SC” 代表 2.2 节提出基于扩展控制的轨迹优化方法, 其终端状态误差统计结果同 3.2 节。“CP” 代表 2.3 节中基于凸规划的轨迹优化方法, 其终端状态误差基本与仿真 “SC” 一致。

2 种方法最终生成的控制历史基本接近:

$$n_y(r_x)=0.27\left(\frac{r_x}{17\,500}-1\right)^2-0.15\left(\frac{r_x}{17\,500}-1\right)+0.15 \tag{34}$$

$$n_z(r_x)=0.24\left(\frac{r_x}{17\,500}-1\right)^2-0.79\left(\frac{r_x}{17\,500}-1\right)+0.85 \tag{35}$$

尽管在优化过程中存在参数幅值约束, 但 2.2 节方法最终生成的控制曲线并未触发该约束。原因在于二次型性能指标确保了控制曲线的光滑性。因此, 在优化问题存在可行解时, 该方法通常不会触发参数幅值约束。

在构建航程-法向/侧向过载曲线后, 进一步以时间为自变量, 对式 (1) 中的实际动力学方程数值积分。法向/侧向过载曲线如图 26 和图 27 所示。与前述结论相似, 2 种方法生成的轨迹基本一致。

2.3 节方法主要用于处理参数饱和的问题。然而, 在上述仿真中, 所有参数均未达到其幅值约束, 因此未能充分展示该方法在处理饱和问题时的有效性。为了更全面地验证该方法的实际应用价值, 后续仿真将针对参数饱和的情况展开研究。3 个参数的幅值约束分别设置为: $\boldsymbol{K}_a\in[-3,\,3]$, $\boldsymbol{K}_b\in[-0.7,\,0.7]$, $\boldsymbol{K}_c\in[-2,\,2]$。通过凸规划方法解算含有参数幅值约束的优化问题, 控制历史表达为

$$n_y(r_x)=0.30\left(\frac{r_x}{17\,500}-1\right)^2-0.15\left(\frac{r_x}{17\,500}-1\right)+0.14 \tag{36}$$

$$n_z(r_x)=-1.13\left(\frac{r_x}{17\,500}-1\right)^2-0.70\left(\frac{r_x}{17\,500}-1\right)+1.31 \tag{37}$$

从式 (36) 和式 (37) 可以看出, 参数向量 $\boldsymbol{K}_b$ 触碰了参数幅值约束, 证明了本文方法具备幅值约束的处理能力。数值仿真结果表明, 三轴位置误差分别为−8.81 m、1.27 m 与−9.86 m; 速度误差为 0.06 m/s; 航迹角与航向角误差分别为 1.96×10^{-3}(°) 与−1.41×

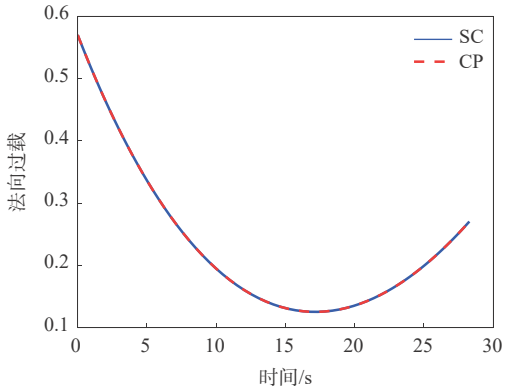


图 26 时间-法向过载曲线

Fig. 26 Curves of normal overload with time

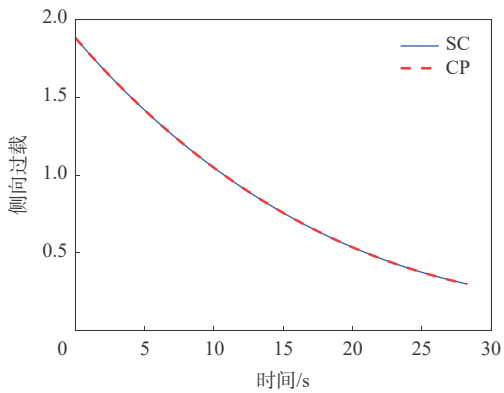


图 27 时间-侧向过载曲线

Fig. 27 Curves of lateral overload with time

10^{-2}(°)。需要特别指出的是，本文方法因为采用符号递推解析架构，从而实现了较高的计算效率。相比之下，凸规划类方法需要将优化问题转化为工具箱要求的标准形式并通过非线性规划方法解算，这一过程在一定程度上牺牲了求解效率[42]。因此，在实际应用中，需综合权衡约束条件与计算精度及效率之间的关系，以选择最合适的轨迹优化策略。

与上述仿真类似，以时间为自变量对动力学方程数值积分，法向/侧向过载曲线如图 28 和图 29 所示。

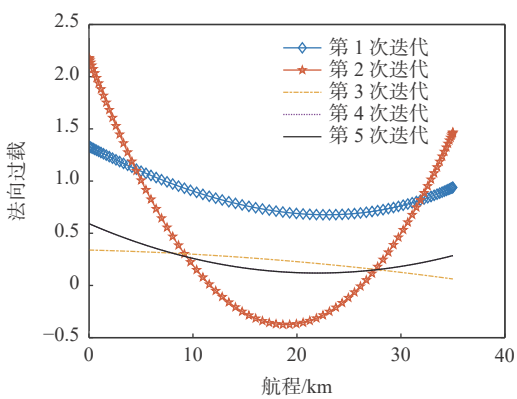


图 28 航程-法向过载曲线

Fig. 28 Curves of normal overload with downrange

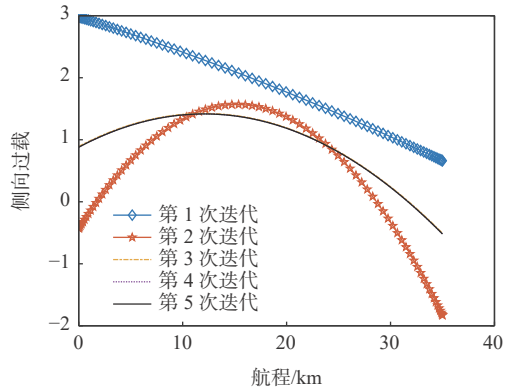


图 29 航程-侧向过载曲线

Fig. 29 Curves of lateral overload with downrange

图 28 和图 29 的对比结果表明，侧向过载曲线在采用凸规划的改进方法后发生了显著变化。这说明该方法通过调整控制曲线，有效处理了参数幅值约束问题，进而实现了快速准确的轨迹优化。

4 结 论

本文提出了一种处理初始和“硬”终端状态约束的在线轨迹优化方法，并得出以下结论：

1）基于实际动力学方程控制向量参数化的临近空间飞行器轨迹优化方法实现了终端状态的精确控制。终端速度误差为-2.90×10^{-3} m/s；航迹角与航向角误差分别为-4.90×10^{-5}(°)与-2.60×10^{-4}(°)；三轴位置误差分别为 0.52 m、0.14 m 与 0.46 m，生成的轨迹与 GPOPS 所得的最优解高度一致，从仿真角度验证了其最优性，并与 2.2 节中的理论相契合。与现有 MPSC 方法相比，本文方法在离散点数量相同的条件下，能够实现终端状态的高精度控制。同时，该方法在 C++环境下完成优化的时间仅为 300～400 ms，满足计算制导下轨迹优化问题在线解算的实时性需求。

2）鉴于临近空间飞行器的中制导任务对终端飞行时间的敏感度较低，本文将以时间为自变量的轨迹优化问题转化为以航程为自变量的优化问题，降低了复杂度。该转换不仅减少了迭代次数，还提升了收敛速度。因此，本文方法适用于面向计算制导的轨迹优化任务。

3）针对参数幅值约束，提出了一种融合凸规划控制向量参数化的轨迹优化方法。由于额外约束的引入，导致问题的复杂度增加，从而产生了一定的终端状态误差。因此在实际应用中，需要平衡约束条件与计算精度之间的关系，以选择合适的轨迹优化策略。

参考文献（References）

[1] 孙智孝，杨晟琦，朴海音，等. 未来智能空战发展综述[J]. 航空学

报, 2021, 42(8): 28-42.
Sun Z X, Yang S Q, Piao H Y, et al. A survey of air combat artificial intelligence[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(8): 28-42(in Chinese).

[2] 傅莉, 谢福怀, 孟光磊, 等. 基于滚动时域的无人机空战决策专家系统[J]. 北京航空航天大学学报, 2015, 41(11): 1994-1999.
Fu L, Xie F H, Meng G L, et al. An UAV air-combat decision expert system based on receding horizon control[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(11): 1994-1999(in Chinese).

[3] Hu Y D, Liu Z C, Jiang C W, et al. Maneuver mode parametric modeling based on trajectory curve evolution laws for hypersonic glide vehicles[J]. Aerospace Science and Technology, 2025, 157: 109856.

[4] 闫循良, 王培臣, 郭杨. 再入滑翔机动突防轨迹规划与制导方法研究综述[J]. 航空学报, 2025, 46(17): 239-266.
Yan X L, Wang P C, Guo Y. Review of research on trajectory planning and guidance methods for entry glide maneuvering penetration[J]. Acta Aeronautica et Astronautica Sinica, 2025, 46(17): 239-266(in Chinese).

[5] 刘哲, 陆浩然, 郑伟, 等. 多滑翔飞行器时间协同轨迹快速规划[J]. 航空学报, 2021, 42(11): 310-324.
Liu Z, Lu H R, Zheng W, et al. Rapid time-coordination trajectory planning method for multi-glide vehicles[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(11): 310-324(in Chinese).

[6] An K, Guo Z Y, Xu X P, et al. A framework of trajectory design and optimization for the hypersonic gliding vehicle[J]. Aerospace Science and Technology, 2020, 106: 106110.

[7] Yu J L, Dong X W, Li Q D, et al. Cooperative guidance strategy for multiple hypersonic gliding vehicles system[J]. Chinese Journal of Aeronautics, 2020, 33(3): 990-1005.

[8] 孙景亮, 刘春生. 基于自适应动态规划的导弹制导律研究综述[J]. 自动化学报, 2017, 43(7): 1101-1113.
Sun J L, Liu C S. An overview on the adaptive dynamic programming based missile guidance law[J]. Acta Automatica Sinica, 2017, 43(7): 1101-1113(in Chinese).

[9] Wang Y Z, Dai H H. Secure model predictive static programming with initial value generator for online computational guidance of near-space vehicles[J]. Aerospace Science and Technology, 2025, 156: 109768.

[10] 王雨琪, 宁国栋, 王晓峰, 等. 基于微分对策的临近空间飞行器机动突防策略[J]. 航空学报, 2020, 41(增刊 2): 69-78.
Wang Y Q, Ning G D, Wang X F, et al. Maneuver penetration strategy of near space vehicle based on differential game[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(Sup 2): 69-78(in Chinese).

[11] Pan B F, Ma Y Y, Yan R. Newton-type methods in computational guidance[J]. Journal of Guidance, Control, and Dynamics, 2019, 42(2): 377-383.

[12] 卢青, 周军, 呼卫军. 基于改进 A*算法的滑翔飞行器轨迹规划[J]. 系统工程与电子技术, 2016, 38(12): 2758-2763.
Lu Q, Zhou J, Hu W J. Trajectory planning for gliding aircraft based on improved A* algorithm[J]. Systems Engineering and Electronics, 2016, 38(12): 2758-2763(in Chinese).

[13] Wang Z, Bao W M, Li H F. Second-order dynamic sliding-mode control for nonminimum phase underactuated hypersonic vehicles[J]. IEEE Transactions on Industrial Electronics, 2017, 64(4): 3105-3112.

[14] Lu P. Introducing computational guidance and control[J]. Journal of Guidance, Control, and Dynamics, 2017, 40(2): 193.

[15] Malyuta D, Yu Y, Elango P, et al. Advances in trajectory optimization for space vehicle control[J]. Annual Reviews in Control, 2021, 52: 282-315.

[16] Shirazi A, Ceberio J, Lozano J A. Spacecraft trajectory optimization: a review of models, objectives, approaches and solutions[J]. Progress in Aerospace Sciences, 2018, 102: 76-98.

[17] Kazim M, Hong J, Kim M G, et al. Recent advances in path integral control for trajectory optimization: an overview in theoretical and algorithmic perspectives[J]. Annual Reviews in Control, 2024, 57: 100931.

[18] Wu D, Cheng L, Jiang F H, et al. Analytical costate estimation by a reference trajectory-based least-squares method[J]. Journal of Guidance, Control, and Dynamics, 2022, 45(8): 1529-1537.

[19] 刘宇航, 杨洪伟, 李爽. 小推力最优轨迹协态估计的高效机器学习方法[J]. 宇航学报, 2022, 43(5): 593-602.
Liu Y H, Yang H W, Li S. Efficient machine learning method for co-state estimation of low-thrust optimal trajectories[J]. Journal of Astronautics, 2022, 43(5): 593-602(in Chinese).

[20] Nurre N P, Taheri E. Comparison of indirect and direct methods for operationally constrained low-thrust trajectories[J]. Journal of Guidance, Control, and Dynamics, 2025, 48(4): 885-902.

[21] 武迪, 郭祥, 蒋方华, 等. 连续小推力轨迹优化间接法研究进展综述[J]. 宇航学报, 2024, 45(7): 995-1007.
Wu D, Guo X, Jiang F H, et al. Progress of indirect method for continuous low-thrust trajectory optimization[J]. Journal of Astronautics, 2024, 45(7): 995-1007(in Chinese).

[22] 任鹏飞, 王洪波, 周国峰. 基于自适应伪谱法的高超声速飞行器再入轨迹优化[J]. 北京航空航天大学学报, 2019, 45(11): 2257-2265.
Ren P F, Wang H B, Zhou G F. Reentry trajectory optimization for hypersonic vehicle based on adaptive pseudospectral method[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(11): 2257-2265(in Chinese).

[23] 雍恩米, 魏桐, 刘深深, 等. 一种基于伪谱法的助推-补能跳跃滑翔轨迹优化方法[J]. 宇航学报, 2024, 45(6): 893-902.
Yong E M, Wei T, Liu S S, et al. A trajectory optimization method for boost-energy supplementing skip gliding motion base on pseudospectral method[J]. Journal of Astronautics, 2024, 45(6): 893-902(in Chinese).

[24] 赵吉松, 李畅, 王晨曦, 等. 吸气式高超声速飞行器助推-巡航轨迹优化[J]. 宇航学报, 2021, 42(7): 820-829.
Zhao J S, Li C, Wang C X, et al. Boost-cruise trajectory optimization of air breathing hypersonic vehicles[J]. Journal of Astronautics, 2021, 42(7): 820-829(in Chinese).

[25] Lu P. What is guidance?[J]. Journal of Guidance, Control, and Dynamics, 2021, 44(7): 1237-1238.

[26] Padhi R, Banerjee A, Mathavaraj S, et al. Computational guidance using model predictive static programming for challenging space missions: an introductory tutorial with example scenarios[J]. IEEE Control Systems Magazine, 2024, 44(2): 55-80.

[27] Patel J, Subbarao K. Reachable-set constrained model predictive control for safe hypersonic reentry[J]. Journal of Guidance, Control, and Dynamics, 2025, 48(6): 1253-1265.

[28] Engel D L, Putnam Z R, Dutta S. Assessment of model predictive control for Mars entry vehicles with flaps[J]. Journal of Spacecraft and Rockets, 2025, 62(5): 1528-1545.

[29] Malyuta D, Reynolds T P, Szmuk M, et al. Convex optimization for trajectory generation: a tutorial on generating dynamically feasible trajectories reliably and efficiently[J]. IEEE Control Systems Magazine, 2022, 42(5): 40-113.

[30] Hayner C R, Carson III J M, Açıkmeşe B, et al. Continuous-time line-of-sight constrained trajectory planning for 6-degree of freedom systems[J]. IEEE Robotics and Automation Letters, 2025, 10(5): 4332-4339.

[31] Ma Y Y, Pan B F, Tan Q D. Neural convex optimization for real-time trajectory generation of asteroid landings[J]. Acta Astronautica, 2025, 229: 606-615.

[32] Benedikter B, Zavoli A, Colasurdo G, et al. Convex approach to three-dimensional launch vehicle ascent trajectory optimization[J]. Journal of Guidance, Control, and Dynamics, 2021, 44(6): 1116-1131.

[33] Wang Z B, Grant M J. Autonomous entry guidance for hypersonic vehicles by convex optimization[J]. Journal of Spacecraft and Rockets, 2018, 55(4): 993-1006.

[34] Dwivedi P N, Bhattacharya A, Padhi R. Suboptimal midcourse guidance of interceptors for high-speed targets with alignment angle constraint[J]. Journal of Guidance, Control, and Dynamics, 2011, 34(3): 860-877.

[35] 张永华, 王铮, 王时雨, 等. 基于无迹模型预测静态规划的可回收火箭着陆段鲁棒制导方法[J/OL]. 北京航空航天大学学报, 2024 (2024-09-14)[2025-04-18]. https://doi.org/10.13700/j.bh.1001-5965.2024.0491.

Zhang Y H, Wang Z, Wang S Y, et al. Robust guidance method for reusable rocket landing phase based on UMPSP[J/OL]. Journal of Beijing University of Aeronautics and Astronautics, 2024 (2024-09-14)[2025-04-18].https://doi.org/10.13700/j.bh.1001-5965.2024.0491 (in Chinese).

[36] 谭一廷, 荆武兴, 高长生. 考虑零控交班视窗角约束的拦截中制导设计[J]. 宇航学报, 2021, 42(10): 1257-1270.

Tan Y T, Jing W X, Gao C S. Design of interception midcourse guidance with zero effort handover visual angle constraint[J]. Journal of Astronautics, 2021, 42(10): 1257-1270(in Chinese).

[37] Ma S D, Yang Y X, Yang H, et al. Trajectory optimization of hypersonic vehicle considering the quasi-static assumption of pitch motion[J]. Aerospace Science and Technology, 2024, 146: 108969.

[38] 王远卓, 代洪华, 岳晓奎. 基于实际动力学方程的飞行器准谱轨迹优化[J]. 航空学报, 2025, 46(22): 184-203.

Wang Y Z, Dai H H, Yue X K. Quasi-spectral trajectory optimization for vehicle based on actual dynamic equations[J]. Acta Aeronautica et Astronautica Sinica, 2025, 46(22): 184-203(in Chinese).

[39] Lu P, Doman D B, Schierman J D. Adaptive terminal guidance for hypervelocity impact in specified direction[J]. Journal of Guidance, Control, and Dynamics, 2006, 29(2): 269-278.

[40] 王远卓. 高超声速飞行器邻近最优制导律研究[D]. 北京: 北京航空航天大学, 2024: 13-19.

Wang Y Z. Research on neighboring optimal guidance law for hypersonic vehicle[D]. Beijing: Beihang University, 2024: 13-19(in Chinese).

[41] 王远卓, 张冉, 李惠峰. 基于协态估计的火箭动力下降邻近最优制导[J]. 宇航学报, 2024, 45(5): 741-752.

Wang Y Z, Zhang R, Li H F. Neighboring optimal guidance for rocket powered descent based on costate estimation[J]. Journal of Astronautics, 2024, 45(5): 741-752(in Chinese).

[42] Liu X, Li S, Xin M. Pseudospectral convex optimization based model predictive static programming for constrained guidance[J]. IEEE Transactions on Aerospace and Electronic Systems, 2023, 59(3): 2232-2244.

[43] Patterson M A, Rao A V. GPOPS-II: a MATLAB software for solving multiple-phase optimal control problems using hp-adaptive gaussian quadrature collocation methods and sparse nonlinear programming[J]. ACM Transactions on Mathematical Software, 2014, 41(1): 1-37.

[44] Agamawi Y M, Rao A V. CGPOPS: a C++ software for solving multiple-phase optimal control problems using adaptive Gaussian quadrature collocation and sparse nonlinear programming[J]. ACM Transactions on Mathematical Software, 2020, 46(3): 1-38.

[45] Mathavaraj S, Padhi R. Quasi-spectral unscented MPSP guidance for robust soft-landing on asteroid[J]. Journal of Optimization Theory and Applications, 2021, 191(2): 823-845.

Near-space vehicle trajectory optimization based on control vector parameterization

WANG Yuanzhuo，DAI Honghua*，YUE Xiaokui

(School of Astronautics，Northwestern Polytechnical University，Xi'an 710072，China)

Abstract：For online trajectory optimization of the near-space vehicle in computational guidance, a trajectory optimization method based on actual dynamic equations is proposed via control vector parameterization. Firstly, a trajectory optimization problem with initial and terminal state constraints is developed by discretizing the nonlinear dynamic equations using the flipped-Radau pseudospectral approach. Secondly, control vector parameterization reduces complexity and boosts computational efficiency by converting the optimization problem for control quantities at each discrete point into an optimization problem for multiple parameter vectors. Thirdly, to deal with parameter amplitude constraints, a trajectory optimization method based on convex programming techniques is proposed via control vector parameterization. Finally, given that the midcourse guidance mission is insensitive to terminal flight time, the optimization problem is transformed from being time-dependent to downrange-dependent, thereby further reducing the computational time. Simulation results demonstrate that the proposed method satisfies the high-efficiency and high-precision requirements for trajectory optimization in computational guidance.

Keywords：near-space vehicle；optimal control；trajectory optimization；convex programming；control vector parametric design

Received：2025-04-29；**Accepted**：2025-07-25；**Published Online**：2025-07-29 10：14

URL：link.cnki.net/urlid/11.2625.V.20250729.0951.001

Foundation items：National Science Fund for Distinguished Young Scholars (52425212); National Key Research and Development Program of China (2021YFA0717100); National Natural Science Foundation of China (12072270,U2013206)

* **Corresponding author**. E-mail: hhdai@nwpu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0444

考虑安全性评估的无人机集群鸽群优化避碰决策

彭雅兰，段海滨*，魏晨
(北京航空航天大学 自动化科学与电气工程学院 飞行器一体化控制全国重点实验室，北京 100191)

摘 要：为提高无人机(UAV)集群在复杂环境中的避碰能力与协同效率，提出一种考虑安全性评估的无人机集群鸽群优化(PIO)避碰决策机制。通过分析无人机集群局部冲突风险与整体网络协调程度，构建多维安全性评估指标体系；利用主成分分析法对集群状态进行综合安全评估，并将避碰决策分为安全、警戒与高危3个阶段。基于综合安全性评估得分动态调节的人工势场函数，实现个体避碰控制参数的自适应调整。为进一步提升高风险状态下集群全局避碰协同效率，引入PIO算法对无人机运动方向与避碰参数进行协同优化。仿真测试结果表明：所提方法在保持集群稳定性的同时有效降低了局部冲突概率，验证了该方法的有效性与实用性。

关 键 词：无人机；鸽群优化；避碰决策；安全性评估；主成分分析法；人工势场法
中图分类号：V221+.3; TB553
文献标志码：A **文章编号**：1001-5965(2026)10-3618-09

随着人工智能、无线通信与导航定位等技术的飞速发展，无人机(unmanned aerial vehicle, UAV)已广泛应用于环境监测、应急救援、空中测绘、物流配送、军事侦察等多种复杂任务场景。由多架无人机组成的集群因其具有高效协同、自主性强、空间覆盖能力广等优势，逐渐成为空中自主系统的重要发展方向[1-3]。然而，在复杂空域中，无人机集群在自主飞行过程中不可避免地面临空间拥堵、路径交叉、局部冲突等问题[4]，因此，如何实现高效、鲁棒的无人机集群避碰与协同控制策略，已成为当前无人系统研究的热点与难点问题。

传统无人机避碰方法可分为基于规则约束、基于轨迹预测、基于优化算法以及基于人工势场法(artificial potential field, APF)等类别[5-7]。其中，APF由于计算简单，适用于分布式控制及机载实时控制，广泛应用于多无人机的避碰控制与路径引导[8-10]。其基本思想是将目标建模为吸引源，障碍物建模为排斥源，通过构造势函数合成控制力，引导个体运动。然而，该方法在处理复杂空间结构或多障碍环境时，易陷入局部极小值、产生路径振荡，且缺乏对集群整体拓扑结构与安全态势的感知与反馈机制，导致避碰效果有限[11]。Hao等[12]在复杂障碍场景中结合风险评估，对人工势场项进行结构化改进，显著提升了路径可行性与平滑性，缓解了传统APF的局部极小值与振荡问题。针对群体编队协同的APF与一致性控制耦合问题，Wang等[13]将改进APF与一致性控制优化用于无人机编队碰撞与通信距离的双约束维护，验证了在通信约束下的收敛与稳定性。Chang等[14]提出基于速度阈值的多一致性APF框架，可在保持连通与避碰的同时实现多子群编队的自适应分合与变换。

近年来，研究者将各种智能优化算法引入无人机路径优化与避碰控制中。鸽群优化(pigeon-inspired optimization, PIO)算法作为一种新型生物启发式群体智能算法，模拟鸽子迁徙中的“地图-指南针”导航与“地标”记忆策略，在路径规划、多目标

收稿日期：2025-07-04；**录用日期**：2025-09-05；**网络出版时间**：2025-09-28 09:08
网络出版地址：link.cnki.net/urlid/11.2625.V.20250926.1757.004
基金项目：国家自然科学基金(U24B20156,62350048,T2121003)
* **通信作者**. E-mail：hbduan@buaa.edu.cn

引用格式：彭雅兰，段海滨，魏晨. 考虑安全性评估的无人机集群鸽群优化避碰决策[J]. 北京航空航天大学学报，2026，52(10)：3618-3626. Peng Y L，Duan H B，Wei C. A safety-evaluation-driven UAV swarm collision avoidance decision-making based on pigeon-inspired optimization[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3618-3626 (in Chinese).

调度与多无人机协同中展现出优越性能[15-17]。Shen 等[18]将仿生社交学习机制应用于 PIO 算法，用于解决具有协同检测功能的多无人机协同路径规划问题。Sun 等[19]研究了基于学习辅助进化 PIO 算法的无人机集群机动决策，通过结合学习进化机制，解决无人机集群空战中的子任务，在全面评估战场态势后，选择最优策略。

当前研究多聚焦于几何避碰或优化路径生成，缺乏对无人机集群飞行安全性与行为一致性的量化评估。影响集群安全性能的因素不仅包括个体间的空间距离，还涉及接近速率、排斥力、拓扑结构等多维信息[20-21]。如何在避碰策略中引入动态安全性评估机制，从而实现对避碰参数的自适应调整，是提高无人机集群智能性和鲁棒性的关键[22]。

从风险感知与智能控制角度出发，本文提出一种考虑安全性评估的无人机集群 PIO 避碰决策，将多维安全性评估机制引入无人机集群避碰控制。基于无人机个体运动状态与集群结构稳定性，构建了完整的“安全评估—策略分级—优化调度”协同框架。利用主成分分析法（principal component analysis, PCA）融合多维评估指标，生成综合安全性评估得分用于调节集群避碰控制策略。引入 PIO 算法，通过搜索加速度方向空间与控制参数组合，提升个体路径选择的全局性与自适应性。该集群避碰策略兼顾快速响应与优化收敛，旨在构建一种面向动态环境、自适应性强、可扩展的无人机集群避碰决策机制。

本文主要贡献如下：

1）提出了“安全评估—策略分级—优化调度”的一体化避碰框架，以综合安全性评估（formal safety assessment, FSA）得分打通风险感知与控制参数之间的闭环，实现基于安全等级的自适应权重 APF 与高危触发 PIO 算法的协同。

2）设计了兼顾局部几何冲突与全局拓扑稳定性的多维安全性评估指标集，并利用主成分分析法生成正交、低冗余的综合安全性评估得分，驱动控制算法参数。

3）在高危态势下引入改进 PIO 算法，在不显著增加计算负担的前提下提升全局协同避碰效率与鲁棒性。

1　无人机集群安全性评估模型

1.1　无人机集群安全性评估指标体系

为全面刻画无人机集群在复杂环境中的风险态势，综合考虑无人机之间的距离、速度、方向与集群结构稳定性等因素，从局部冲突风险与整体网络协调维度出发，构建无人机集群多维度安全性评估指标体系，量化集群在不同态势下的潜在碰撞风险与结构脆弱性[23]。以四旋翼无人机为例，建立安全性评估模型，双无人机间冲突如图 1 所示。

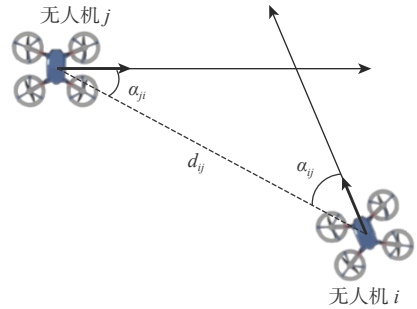


图 1　双无人机冲突

Fig. 1　Two-UAV conflict scenario

1）接近速率 A 反映无人机个体与邻居之间是否存在逼近趋势，是冲突预测的重要依据。无人机 i、无人机 j 与路径冲突交点构成一个三角形。机间接近速率越大，冲突风险越高。假设无人机 i 与其感知半径 R 内所有无人机进行信息交互，组成无人机 i 的邻居集合 N_i。定义无人机 i 与其邻居无人机 $j \in N_i$ 之间的接近速率为

$$A_{ij} = V_i \cos\alpha_{ij} + V_j \cos\alpha_{ji} \tag{1}$$

$$\alpha_{ij} = \begin{cases} |\chi_i - \varphi_{ij}| & |\chi_i - \varphi_{ij}| \leqslant \pi \\ 2 - |\chi_i - \varphi_{ij}| & |\chi_i - \varphi_{ij}| > \pi \end{cases} \tag{2}$$

式中：V_i 和 V_j 分别为无人机 i 和无人机 j 的速度；χ_i 为无人机 i 的航向角；φ_{ij} 为无人机 j 相对于无人机 i 的方位角；α_{ij}、α_{ji} 分别为无人机 i、无人机 j 的航向与 2 架无人机连线之间形成的锐角。

无人机 i 的接近速率指标 I 定义为其与周围无人机之间接近速率的最大值：

$$I_i = \max_{j \in N_i} (A_{ij}) \tag{3}$$

2）局部势能指标 E 衡量无人机 i 与其邻居集合 N_i 内所有无人机之间距离的紧迫程度，定义为

$$E_i = \sum_{j \in N_i} \left(\frac{1}{d_{ij}} - \frac{1}{R} \right)^{\tau} \tag{4}$$

式中：$d_{ij} = \|\boldsymbol{p}_i - \boldsymbol{p}_j\|$；$\tau > 1$ 为指数因子，一般取 $\tau = 2$，其中，$\boldsymbol{p}_i$ 和 $\boldsymbol{p}_j$ 分别为无人机 i 和无人机 j 的位置。局部势能越大，个体处于越拥挤、危险的位置。

3）最大势能指标 P 度量集群中最不安全个体的极端风险程度。当无人机 i 与无人机 j 的局部势能指标相等时，即 E_i=P_j，为区分冲突风险，定义无人机 i 的最大势能为

$$P_i = \max_{i \in \{1,2,\cdots,N\}} (E_i) \tag{5}$$

式中：N为集群中无人机数量。

4）排斥力指标F为无人机i所受人工势场排斥力之和：

$$F_i=\sum_{j\in N_i}\left(k_{\mathrm{r}}\left(\frac{1}{d_{ij}}-\frac{1}{d_0}\right)\frac{1}{d_{ij}^2}\right) \tag{6}$$

式中：k_{r}为排斥系数；d_0为排斥力作用阈值；$\boldsymbol{p}_i$为无人机i的位置。

5）最大排斥力指标$F_i^{\max}$与最大势能指标类似，即当无人机i与无人机j所受排斥力之和相等时，最大排斥力更大的个体，冲突风险更高，定义为

$$F_i^{\max}=\max_{j\in N_i}\left(k_{\mathrm{r}}\left(\frac{1}{d_{ij}}-\frac{1}{d_0}\right)\frac{1}{d_{ij}^2}\right) \tag{7}$$

6）节点度D为与无人机i存在信息交互的无人机个体数。一般而言，机间距离越近，冲突风险越高，将2架无人机之间距离的倒数作为权重，定义无人机的加权节点度W为

$$D_i=\sum_{j=1}^{N}l_{ij} \tag{8}$$

$$l_{ij}=\begin{cases}1 & d_{ij}\leqslant R\\ 0 & d_{ij}>R\end{cases} \tag{9}$$

$$W_i=\sum_{j=1}^{N}l_{ij}\omega_{ij} \tag{10}$$

$$\omega_{ij}=\frac{1}{d_{ij}} \tag{11}$$

式中：l_{ij}为无人机i与无人机j间的邻接系数，$l_{ij}=1$代表机间存在数据交互；ω_{ij}为无人机i与无人机j间节点度权重。

7）平均强度S_i指标衡量无人机i与其他无人机间接近程度，定义为

$$S_i=\frac{1}{N}\sum_{j=1}^{N}l_{ij}\omega_{ij} \tag{12}$$

平均强度值越高，代表无人机个体信息传输密度与结构紧密度越高。

8）平均聚类系数C量化所有无人机的密集程度。考虑复杂网络具有明显的聚类特征，定义聚类系数量化相邻无人机间的互联程度：

$$c_\omega(i)=\frac{1}{k_i^\omega\left(k_i^\omega-1\right)}\sum_{j,k}\frac{\omega_{ij}+\omega_{jk}}{2}l_{ij}l_{jk}l_{ki} \tag{13}$$

$$C_i=\frac{1}{N}\sum_{i=1}^{N}c_\omega(i) \tag{14}$$

式中：$c_\omega(i)$为无人机i的聚类系数；k_i^ω为聚类因子。$C_i\in[0,1]$，当$C_i=0$，集群中各无人机相互隔离，不存在冲突风险；当$C_i=1$，集群中任意2架无人机之间都存在至少一条信息交互链路，冲突风险较高。

通过上述指标的联合评估，可同时检测无人机集群内部冲突与整体网络结构演化趋势，为避碰方法提供有效反馈基础。

1.2 主成分分析法与综合安全性评估模型

为克服多指标评估中数据维度冗余、指标间相关性高等问题[24]，本文引入主成分分析法，通过对标准化后的指标数据矩阵进行特征分解，提取能反映集群避碰风险特征的主成分，并将其映射为综合安全性评估值，作为后续改进人工势场权重调节和避碰优化的关键驱动因子。

1.2.1 主成分分析法

设无人机集群安全性指标矩阵为

$$\boldsymbol{Y}=\left[y_{ij}\right]_{N\times M} \tag{15}$$

式中：y_{ij}为无人机i的第j个安全性评估指标值；N为无人机总数；M为评估指标个数。为消除数据量纲差异，采用零均值单位方差标准化，得到安全性评估指标标准化矩阵$\boldsymbol{Y}'=[y'_{ij}]_{M\times N}$：

$$y'_{ij}=\frac{y_{ij}-\overline{y_j}}{\sigma_j} \tag{16}$$

$$\overline{y_j}=\frac{1}{N}\sum_{i=1}^{N}y_{ij} \tag{17}$$

$$\sigma_j=\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left(y_{ij}-\overline{y_j}\right)^2} \tag{18}$$

式中：$\overline{y_j}$和σ_j分别为第j个安全性评估指标的均值和标准差。

计算标准化评估数据的协方差矩阵：

$$\boldsymbol{Z}=\frac{1}{N-1}\boldsymbol{Y}'^{\mathrm{T}}\boldsymbol{Y}' \tag{19}$$

对协方差矩阵$\boldsymbol{Z}$进行特征值分解，得到其特征值$\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_M$，对应的特征向量为$\boldsymbol{v}_1,\boldsymbol{v}_2,\cdots,\boldsymbol{v}_M$。

主成分第k项的贡献率为

$$\gamma_k=\lambda_k\left(\sum_{j=1}^{M}\lambda_j\right)^{-1} \tag{20}$$

选取前M'个主成分，使得累计贡献率达到设定阈值μ：

$$\sum_{k=1}^{M'}\gamma_k\geqslant\mu \tag{21}$$

式中：设定阈值μ一般取值为0.85、0.9或0.95。

通过上述主成分分析法，将原始存在相关性的评估指标数据映射到相互正交的主成分空间中，确保综合评估结果更具独立性与代表性。

1.2.2　综合安全性评估得分

对于每架无人机，计算其主成分得分 $\boldsymbol{\epsilon}_i=\left[\epsilon_i^1,\epsilon_i^2,\cdots,\epsilon_i^k\right]$：

$$\epsilon_i^k=\sum_{j=1}^{M'}v_{kj}y'_{ij}\qquad k=1,2,\cdots,M' \tag{22}$$

定义每架无人机的综合安全性评估得分为主成分得分的加权和：

$$S_{\mathrm{FSA}i}=\sum_{k=1}^{k'}\gamma_k\epsilon_i^k \tag{23}$$

为便于计算，将综合安全性评估得分进行归一化处理至 [0,1] 区间：

$$S'_{\mathrm{FSA}i}=\frac{S_{\mathrm{FSA}i}-\min(S_{\mathrm{FSA}i})}{\max(S_{\mathrm{FSA}i})-\min(S_{\mathrm{FSA}i})} \tag{24}$$

$S'_{\mathrm{FSA}i}$越小，说明无人机 i 所处环境越安全，发生碰撞概率越低；反之，则表明其面临较高的冲突风险。为简化表述，下文统一使用$S_{\mathrm{FSA}i}$表示无人机 i 的归一化综合安全性评估得分。依据每架无人机的综合安全性评估得分，可将其所处状态划分为 3 个安全等级，如表 1 所示。

表 1　无人机集群安全等级划分

Table 1　UAV swarm security classification

等级	综合安全性评估得分	模式类型	控制策略
安全区	$S_{\mathrm{FSA}i}\leqslant\theta_1$	协同前进	吸引力占主导，维持队形或目标引导
警戒区	$\theta_1<S_{\mathrm{FSA}i}<\theta_2$	正常避碰	平衡吸引力与排斥力，进行协调避碰
高危区	$\theta_2\leqslant S_{\mathrm{FSA}i}$	紧急避碰	排斥权重最大化，抑制吸引力，优化避碰

注：θ_1、θ_2为策略阈值，可根据任务需求设定。

安全区无人机处于低风险区域，与周围邻居保持良好间距和协调状态，此时集群优先保持飞行效率。警戒区无人机安全性中等，存在一定冲突风险，需适度增强排斥控制，保持避碰与协同平衡。高危区无人机个体安全性较差，面临局部拥堵或障碍物碰撞风险，此时需最大限度增强避碰控制，保障个体快速逃离高危区域。

2　基于综合安全性评估的无人机集群避碰决策

2.1　考虑安全性评估的无人机集群避碰决策框架

为提升无人机集群在复杂动态环境中的避碰响应能力与协同控制水平，本文构建了一种考虑安全性评估的优化机制，实现多阶段集群避碰决策。该方法以无人机集群综合安全性评估为基础，综合考虑无人机间距离、速度、方向、局部拓扑等特征，建立量化评估模型。结合主成分分析法压缩指标维度，得到综合安全性评估得分，进而驱动无人机个体避碰策略的自适应调节。

无人机集群安全等级划分与避碰决策对应机制如图 2 所示。整个控制体系由 4 个核心部分组成：①感知与通信层。每个循环周期内，无人机通过本地传感器与组网通信设备获取自身与邻居无人机的位置信息、速度信息以及任务目标位置，同时读取包括感知半径、势场参数、风险阈值、自适应系数和 PIO 算法参数在内的控制配置。②安全性评估层。无人机基于感知信息，实时计算与邻居间的相对状态，通过预定义的评估函数获得安全性指标。③分级策略判定层。依据综合安全性评估得分与两级风险阈值θ_1、θ_2，将状态划分为安全区、警戒区和高危区，安全区与警戒区调整吸引力与排斥力权重以增强 APF 避碰能力，而高危区则触发 PIO 优化模块，对 APF 方向进行全局-局部两阶段搜索优化。④策略执行与状态更新层。对于高危状态下得到的 PIO 优化结果，与 APF 方案进行基于综合安全性评估得分适应度的策略融合，选取更优加速度指令，并经过动力学约束与饱和处理后更新无人机的位置和速度，实现下一周期的安全飞行与协同推进。

APF 可由综合安全性评估得分驱动对权重进行即时调节，并提供稳定、可控的初始加速度，与 PIO 全局方向搜索形成局部与全局互补，在高危态势下形成两阶段优化。该框架通过“评估—分级—自适应控制—优化融合”的闭环方式，实现了风险感知驱动的无人机集群动态避碰决策，有效兼顾了安全性与协同性。

2.2　自适应权重人工势场法

传统 APF 中，经典吸引力势场函数U_{att}为[25]

$$U_{\mathrm{att}}(X)=\frac{1}{2}k_{\mathrm{a}}\left|\boldsymbol{x}-\boldsymbol{x}_g\right|^2 \tag{25}$$

式中：k_{a}为吸引力系数；$\boldsymbol{x}$ 和 $\boldsymbol{x}_g$ 分别为无人机和目标位置矢量。

通过计算目标对无人机产生吸引力势函数的梯度，可以计算得到吸引力$\boldsymbol{F}_{\mathrm{att}}$为

$$\boldsymbol{F}_{\mathrm{att}}(X)=-\nabla U_{\mathrm{att}}(X)=k_{\mathrm{a}}(\boldsymbol{x}-\boldsymbol{x}_g) \tag{26}$$

对于躲避障碍物的无人机，建立排斥力势场U_{rep}：

$$U_{\mathrm{rep}}(X)=\begin{cases}\dfrac{1}{2}k_{\mathrm{r}}\left(\dfrac{1}{d}-\dfrac{1}{d_0}\right)^2 & d\leqslant d_0\\ 0 & d>d_0\end{cases} \tag{27}$$

式中：d_0为排斥力势场有效距离，当障碍物与无人机间距离d大于d_0，障碍物不再对无人机产生排斥力。

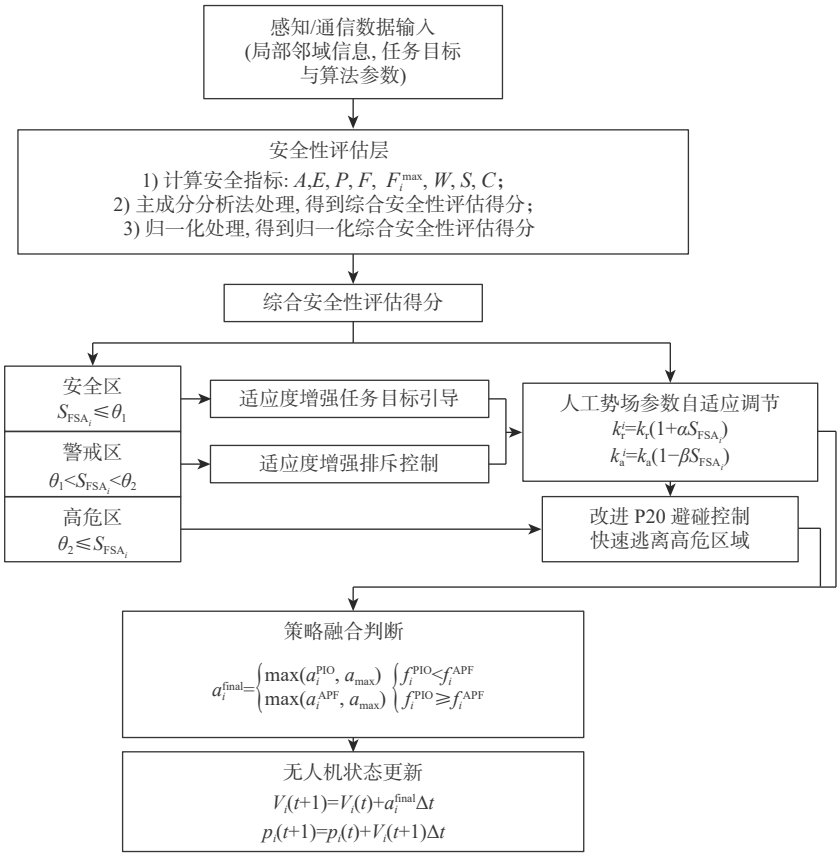


图2 安全等级驱动避碰决策框架

Fig. 2 Framework of security-level-driven collision avoidance decision

排斥力$\boldsymbol{F}_{\text{rep}}$为排斥力势场梯度:

$$\boldsymbol{F}_{\text{rep}}(X)=-\nabla U_{\text{rep}}(X)=\begin{cases}k_{\text{r}}\left(\dfrac{1}{d}-\dfrac{1}{d_0}\right)\dfrac{\boldsymbol{P}_i-\boldsymbol{P}_j}{d^2} & d\leqslant d_0\\ \boldsymbol{0} & d>d_0\end{cases}\tag{28}$$

合力$\boldsymbol{F}_{\text{sum}}$可表示为

$$\boldsymbol{F}_{\text{sum}}=\boldsymbol{F}_{\text{att}}-\boldsymbol{F}_{\text{rep}}\tag{29}$$

传统 APF 在避碰控制中存在参数固定、响应迟钝等问题,难以适应高动态环境下的无人机集群协同避碰需求。为提升避碰控制算法智能化与灵活性,引入基于综合安全性评估得分的权重调节机制,实现势场参数的个体化动态调整,使得风险较高的个体表现出更强烈的趋避行为。排斥力与吸引力权重调节函数分别为

$$k_{\text{r}}^i=k_{\text{r}}(1+\alpha S_{\text{FSA}i})\tag{30}$$

$$k_{\text{a}}^i=k_{\text{a}}(1-\beta S_{\text{FSA}i})\tag{31}$$

式中:$S_{\text{FSA}i}\in[0,1]$为归一化处理后的综合安全性评估得分;α、$\beta\in[0.5,1]$为权重调节灵敏度系数。

该权重调节机制能够使处于高冲突风险的无人机自动增强排斥力并弱化目标吸引力,从而迅速避开冲突区域,且不易陷入局部最小值点;而处境较为安全的无人机个体将继续保持正常航向,维持无人机集群整体飞行任务的推进效率。

2.3 基于改进鸽群优化的无人机集群避碰控制

在多无人机集群自主飞行过程中,由于集群内部个体间频繁交互及动态环境干扰,容易出现局部冲突、路径拥挤、协调失效等问题。APF 具有良好的实时避碰能力,但存在局部极小值、路径振荡和集群行为失控等局限。为弥补其不足,引入 PIO 算法,通过全局协同优化机制对集群状态进行动态调整,从而实现局部避碰与群体协同的统一融合。

若无人机综合安全性评估得分超过θ_2,说明局部风险较大,个体立即启动 PIO 避碰控制策略进行强化避碰。

融合 APF 与 PIO 的集群避碰决策机制,以人工势场计算无人机的加速度作为初始状态,在此基础上,构建基于综合安全性评估指标的适应度函数,通过 PIO 迭代搜索优化加速度方向,从而得到更为安全和协调的控制策略。

该融合策略分为以下 4 个主要步骤:

1)局部人工势场感知避碰。利用人工势场计算无人机个体受目标吸引与障碍物、邻居排斥的加速度。

2)全局方向优化。将该加速度方向作为 PIO 的初始解,构建适应度函数进行群体优化。

① 初始化。假设 N_{p} 只鸽子在 L 维搜索空间中，第 i 只鸽子在第 t 次迭代中位置与速度分别为 $\boldsymbol{x}_i^t$ 与 $\boldsymbol{V}_i^t$。鸽子初始化位置为无人机加速度单位方向[26]：

$$\boldsymbol{x}_i^o=\frac{\boldsymbol{a}_i}{\|\boldsymbol{a}_i\|} \tag{32}$$

式中：a_i 为无人机加速度。

优化方向目标为提高无人机集群安全性与协调性，引入综合安全性评估指标，构造代价函数：

$$f_i(t)=S_{\mathrm{FSA}i}(\boldsymbol{x}_i^t) \tag{33}$$

② 地图-指南针算子阶段，最大迭代次数为 N_{c1}，鸽子速度与位置更新规则为

$$\begin{cases}\boldsymbol{V}_i^t=\boldsymbol{V}_i^{t-1}\mathrm{e}^{-rt}+N_{\mathrm{rand}}\left(\boldsymbol{x}_{\mathrm{best}}^{t-1}-\boldsymbol{x}_i^{t-1}\right)\\ \boldsymbol{x}_i^t=\boldsymbol{x}_i^{t-1}+\boldsymbol{V}_i^t\end{cases} \tag{34}$$

式中：r 为地图和指南针因子；N_{rand} 为 [0,1] 范围内的随机数；x_{best}^{t-1} 为当前鸽群最优位置。

③ 地标算子阶段，最大迭代次数为 N_{c2}。鸽子依靠鸽巢附近的地标导航，远离目的地的鸽子在迭代中被依次抛弃。当迭代次数 $N_{c1}<t\leqslant N_{c2}$ 时，鸽子速度与位置更新规则为

$$\boldsymbol{x}_{\mathrm{center}}^{t-1}=\frac{\sum_{i=1}^{N_{\mathrm{p}}^{t-1}}\boldsymbol{x}_i^{t-1}f_i\left(\boldsymbol{x}_i^{t-1}\right)}{N_{\mathrm{p}}^{t-1}\sum_{i=1}^{N_{\mathrm{p}}^{t-1}}f_i\left(\boldsymbol{x}_i^{t-1}\right)} \tag{35}$$

式中：$\boldsymbol{x}_{\mathrm{center}}^{t-1}$ 为鸽群在 $t-1$ 时刻的中心位置；N_{p}^{t-1} 为 $t-1$ 时刻鸽群中鸽子数量。

$$T\left(\boldsymbol{x}_i^{t-1}\right)=\frac{1}{f_i(t-1)+\varepsilon} \tag{36}$$

式中：T 为第 $t-1$ 代鸽子 i 的适应度值；ε 表示一个极小数，以确保分母不为零。

$$\overline{T}_{\mathrm{global}}^{t-1}=\frac{\sum_{i=1}^{N_{\mathrm{p}}^{t-1}}T\left(\boldsymbol{x}_i^{t-1}\right)}{N_{\mathrm{p}}^{t-1}} \tag{37}$$

式中：$\overline{T}_{\mathrm{global}}^{t-1}$ 表示 $t-1$ 时刻鸽群的平均适应度。将鸽群中每个鸽子个体随机排列为环形拓扑结构，每只鸽子及其左、右两边的鸽子构成小组：

$$\overline{T}_{\mathrm{local}}^{t-1}=\frac{T\left(\boldsymbol{x}_i^{t-1}\right)+T\left(\boldsymbol{x}_{\mathrm{left}}^{t-1}\right)+T\left(\boldsymbol{x}_{\mathrm{right}}^{t-1}\right)}{3} \tag{38}$$

式中：$\overline{T}_{\mathrm{local}}^{t-1}$ 表示小组内在 $t-1$ 时刻的平均适应度值；$T\left(\boldsymbol{x}_{\mathrm{left}}^{t-1}\right)$、$T\left(\boldsymbol{x}_{\mathrm{right}}^{t-1}\right)$ 分别为鸽子 i 在环形拓扑结构中左、右邻居的适应度值。

$$\boldsymbol{x}_{\mathrm{local,center}}^{t-1}=\frac{\boldsymbol{x}_i^{t-1}+\boldsymbol{x}_{\mathrm{left}}^{t-1}+\boldsymbol{x}_{\mathrm{right}}^{t-1}}{3} \tag{39}$$

式中：$\boldsymbol{x}_{\mathrm{local,center}}^{t-1}$ 为 $t-1$ 时刻小组的中心坐标；$\boldsymbol{x}_{\mathrm{left}}^{t-1}$、$\boldsymbol{x}_{\mathrm{right}}^{t-1}$ 分别为鸽子 i 在环形拓扑结构中左、右邻居的位置坐标。

$$\begin{cases}\boldsymbol{x}_i^t=\boldsymbol{x}_i^{t-1}+N_{\mathrm{rand}}\left(\boldsymbol{x}_{\mathrm{center}}^{t-1}-\boldsymbol{x}_i^{t-1}\right) & \overline{T}_{\mathrm{local}}^{t-1}<\overline{T}_{\mathrm{global}}^{t-1}\\ \boldsymbol{x}_i^t=\boldsymbol{x}_i^{t-1}+N_{\mathrm{rand}}\left(\boldsymbol{x}_{\mathrm{local,center}}^{t-1}-\boldsymbol{x}_i^{t-1}\right) & \overline{T}_{\mathrm{local}}^{t-1}\geqslant\overline{T}_{\mathrm{global}}^{t-1}\end{cases} \tag{40}$$

鸽群中的个体在环形拓扑结构中两两对比位置优劣，优胜鸽子向整个鸽群中最优位置学习，落败鸽子不仅需要向整个鸽群学习，还需向优胜鸽子学习。

$$N_{\mathrm{p}}^t=\frac{N_{\mathrm{p}}^{t-1}}{2} \tag{41}$$

按照代价函数值将鸽子排序，在每次迭代中舍弃鸽群中代价函数值低的一半个体，直至 $t>N_{c1}+N_{c2}$，结束循环，得到最优加速度方向 $\boldsymbol{a}_i^{\mathrm{PIO}}$。

3）策略融合判断。比较 APF 和 PIO 输出的综合安全性评估得分，选择最优控制策略。引入控制融合判断函数，对比 2 种控制策略下适应度值：

$$\boldsymbol{a}_i^{\mathrm{final}}=\begin{cases}\max\left(\boldsymbol{a}_i^{\mathrm{PIO}},a_{\max}\right) & f_i^{\mathrm{PIO}}<f_i^{\mathrm{APF}}\\ \max\left(\boldsymbol{a}_i^{\mathrm{APF}},a_{\max}\right) & f_i^{\mathrm{PIO}}\geqslant f_i^{\mathrm{APF}}\end{cases} \tag{42}$$

式中：$a_{\max}$ 为无人机最大加速度限制。

4）无人机状态更新。根据最终加速度更新个体速度与位置。

$$\begin{cases}\boldsymbol{V}_i(t+1)=\boldsymbol{V}_i(t)+\boldsymbol{a}_i^{\mathrm{final}}\Delta t\\ \boldsymbol{p}_i(t+1)=\boldsymbol{p}_i(t)+\boldsymbol{V}_i(t+1)\Delta t\end{cases} \tag{43}$$

通过上述综合安全性评估得分驱动避碰机制，无人机个体可根据所处安全态势实时自适应调整避碰策略，避免单一、固定控制行为所带来的低效率问题，为后续优化算法提供良好初始基础和局部避碰能力保障。

3　仿真实验设计与结果分析

为验证本文所提考虑安全性评估的无人机集群 PIO 避碰决策，设计 5 架无人机穿越设置有多个障碍物的复杂环境。每架无人机初始速度为 10 m/s，初始加速度为 0，无人机最大速度 $V_{\max}$=15 m/s，最大加速度 $a_{\max}$=10 m/s^2，感知半径 R=300 m，基于综合安全性评估得分的无人机集群安全等级划分阈值 $\theta_1=0.3$，$\theta_2=0.7$，在集群前进方向上，设置 6 个圆柱形障碍物。设定最大仿真时间 200 s，步长 0.05 s，算法中各参数设置如表 2 所示。

安全阈值 θ_1 与 θ_2 决定了 APF 与 PIO 控制模式的切换区间，较小的 θ_1 会使集群更早进入“警戒区”，从而提高避碰控制保守性，显著降低冲突次数，但同时也导致路径绕行与飞行时间延长；而较大的 θ_2 会延迟进入“高危区”，延缓 PIO 启动，在短时间内冲突次数上升，但能够减小全局优化计算频

表2 算法参数

Table 2 Parameters of algorithms

算法	参数	数值
APF	吸引力系数k_a	5
	排斥力系数k_r	15
	排斥力作用阈值d_0/m	200
	权重调节灵敏度系数α	0.75
	权重调节灵敏度系数β	0.75
PIO	鸽子种群数量N_p	20
	地图和指南针算子迭代次数N_{c1}	100
	地标算子迭代次数N_{c2}	80
	地图和指南针因子r	0.02

率，从而降低计算负担。APF权重系数k_r与k_a在不同安全等级下动态调节排斥力与吸引力，较大的排斥力系数k_r能在高风险态势下放大排斥力，提升避碰效果，但过大可能导致排斥过冲，引发飞行轨迹振荡；较小的吸引力系数k_a在高风险态势中能够抑制吸引力，强化避障动作，但若设置过低，则可能导致任务执行延迟。

图3、图4分别给出了考虑安全性评估的无人机集群PIO算法避碰三维飞行轨迹与飞行轨迹俯视图，图3中，H为无人机高度。

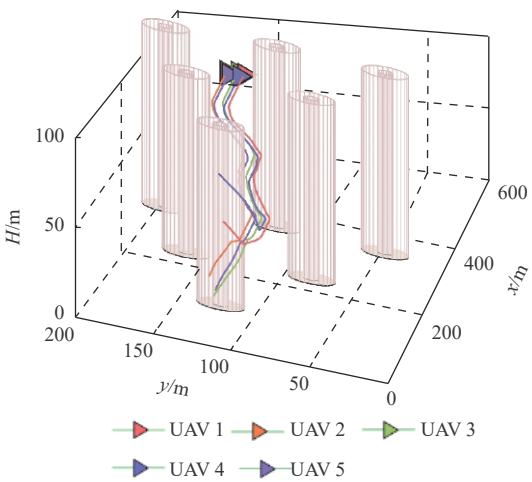


图3 无人机集群三维飞行轨迹

Fig. 3 3D flight path of UAV swarm

由图3和图4可知，在本文所提出避碰控制策略作用下，无人机集群可稳定飞行通过复杂环境，且未发生机间碰撞。

为验证所提避碰控制策略必要性，在相同初始条件和参数设置下，将本文方法分别与单独利用APF、粒子群优化(particle swarm optimization, PSO)算法与PIO算法避碰控制的无人机集群飞行进行仿真验证。4种避碰控制算法下关键安全性评估指标对比结果如表3所示(取10次仿真实验结果平均值)。在任务约束下，PIO通常在20～50次迭代内

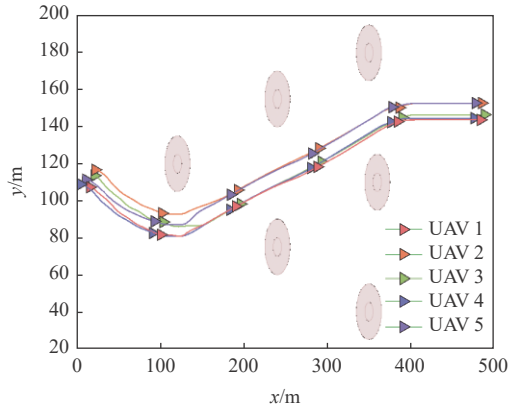


图4 无人机集群飞行轨迹俯视图

Fig. 4 Top view of flight path of UAV swarm

表3 安全性评估指标数据

Table 3 Data on security assessment indicators

指标	A	E	F^{max}	W	S	C	F_i	冲突次数
APF	1.42	117.80	19.30	5.10	0.49	0.53	0.75	5.10
PSO	1.17	96.50	14.70	4.60	0.42	0.45	0.67	2.40
PIO	1.12	91.10	13.50	4.40	0.38	0.41	0.62	1.20
FSA-PIO(本文)	0.95	78.90	10.20	3.80	0.32	0.35	0.51	0.20

收敛至稳定解，且在多次实验中收敛结果波动较小，验证了其在高危状态下的快速响应与稳定性。

由表3可见，本文所提考虑安全性评估的无人机集群PIO避碰决策方法在各个安全性评估指标方面均不同程度优于其他3种基本避碰控制算法。其优越性在于：算法中主成分分析法将各个评估指标压缩为互为正交的主成分并以贡献率加权，形成多元风险统一评分。分级控制将抽象评分转换为具体控制方式，并通过策略融合判断，确保不会因个别采样波动而生成不安全指令。本文算法具有较好的全局路径规划与局部避碰能力，可有效解决复杂环境中的无人机集群避碰问题。

4 结 论

针对无人机集群安全性评估与避碰控制问题，本文提出一种考虑安全性评估的无人机集群鸽群优化算法避碰决策。

1) 构建了一种无人机集群多维安全性评估指标体系，包含接近速率、势能、排斥力、网络拓扑特征等关键指标，从微观运动特性与宏观结构稳定性2方面评估无人机安全状态。

2) 利用主成分分析法，实现了多评估指标融合为单一综合安全性评估得分，为后续控制器调节提供了统一度量依据。

3) 根据无人机综合安全性评估得分调节人工势场吸引力与排斥力权重，构建了一种响应式避碰策略。

4）引入鸽群优化算法对加速度方向与路径进行优化，有效克服了局部最优与集群内部冲突困境，提升了无人机集群整体避碰质量。

未来可进一步考虑复杂对抗环境中无人机集群中的信息不完备与通信滞后问题，扩展无人机集群鲁棒控制的协同安全性评估理论体系。

参考文献（References）

[1] Sharma R R, Ghose D. Collision avoidance between UAV clusters using swarm intelligence techniques[J]. International Journal of Systems Science, 2009, 40(5): 521-538.

[2] Javed S, Hassan A, Ahmad R, et al. State-of-the-art and future research challenges in UAV swarms[J]. IEEE Internet of Things Journal, 2024, 11(11): 19023-19045.

[3] Huang S Y, Zhang H B, Huang Z Y. E2CoPre: energy efficient and cooperative collision avoidance for UAV swarms with trajectory prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(7): 6951-6963.

[4] Rezaee M R, Hamid N A W A, Hussin M, et al. Comprehensive review of drones collision avoidance schemes: challenges and open issues[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(7): 6397-6426.

[5] Debnath D, Vanegas F, Sandino J, et al. A review of UAV path-planning algorithms and obstacle avoidance methods for remote sensing applications[J]. Remote Sensing, 2024, 16(21): 4019.

[6] Roberge V, Tarbouchi M, Labonte G. Comparison of parallel genetic algorithm and particle swarm optimization for real-time UAV path planning[J]. IEEE Transactions on Industrial Informatics, 2013, 9(1): 132-141.

[7] Zhou X, Zhu J C, Zhou H Y, et al. EGO-swarm: a fully autonomous and decentralized quadrotor swarm system in cluttered environments[C]//2021 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE Press, 2021: 4101-4107.

[8] Wu Z Y, Dong S P, Yuan M, et al. Rotate artificial potential field algorithm toward 3D real-time path planning for unmanned aerial vehicle[J]. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 2023, 237(4): 940-955.

[9] Zhang Y X, Chen J B, Chen M, et al. Integrated the artificial potential field with the leader-follower approach for unmanned aerial vehicles cooperative obstacle avoidance[J]. Mathematics, 2024, 12(7): 954.

[10] Cao L J, Wang L, Liu Y, et al. 3D trajectory planning based on the rapidly-exploring random tree-connect and artificial potential fields method for unmanned aerial vehicles[J]. International Journal of Advanced Robotic Systems, 2022, 19(5): 172988062211188.

[11] 朱元军, 李妍, 张学军, 等. 基于风险约束的城市空域无人机安全路径规划[J/OL]. 北京航空航天大学学报, 2025(2025-04-29)[2025-07-01]. https://doi.org/10.13700/j.bh.1001-5965.2024.0843.
Zhu Y J, Li Y, Zhang X J, et al. Risk-constrained safe path planning for unmanned aerial vehicle in urban airspace[J/OL]. Journal of Beijing University of Aeronautics and Astronautics, 2025(2025-04-29)[2025-07-01]. https://doi.org/10.13700/j.bh.1001-5965.2024.0843(in Chinese).

[12] Hao H, Xue Y, Zhai G, et al. UAV path planning based on improved artificial potential field method[J]. Aerospace, 2023, 10(6): 562.

[13] Wang M, Zhong Y L, Zhang Y H, et al. Unmanned aerial vehicle formation control method based on improved artificial potential field and consensus[J]. IET Control Theory & Applications, 2024, 18(18): 2555-2567.

[14] Chang X F, Jiao J Y, Li Y N, et al. Multi-consensus formation control by artificial potential field based on velocity threshold[J]. Frontiers in Neuroscience, 2024, 18: 1367248.

[15] Duan H B, Qiu H X. Advancements in pigeon-inspired optimization and its variants[J]. Science China Information Sciences, 2019, 62(7): 70201.

[16] Chen W K, Hai X S, Hu Y M, et al. Hierarchical decision-making framework for multi-UAV task assignment via enhanced pigeon-inspired optimization[J]. Guidance, Navigation and Control, 2023, 3(4): 2350028.

[17] Hu C H, Xia Y, Zhang J G. Adaptive operator quantum-behaved pigeon-inspired optimization algorithm with application to UAV path planning[J]. Algorithms, 2019, 12(1): 3.

[18] Shen Y K, Liu X N, Ma X, et al. A bionic social learning strategy pigeon-inspired optimization for multi-unmanned aerial vehicle cooperative path planning[J]. Applied Sciences, 2025, 15(2): 910.

[19] Sun Y B, Chen Y, Wei C, et al. Swarm maneuver decision method based on learning-aided evolutionary pigeon-inspired optimization for UAV swarm air combat[J]. Drones, 2025, 9(3): 218.

[20] Tisdale J, Kim Z, Hedrick J K. Autonomous UAV path planning and estimation[J]. IEEE Robotics & Automation Magazine, 2009, 16(2): 35-42.

[21] Lei Y Q, She Z K, Quan Q. Pigeon-inspired UAV swarm control and planning within a virtual tube[J]. Drones, 2025, 9(5): 333.

[22] 陈艺君, 余莎莎, 张学军. 城市低空场景下无人机运行对地风险量化评估[J]. 北京航空航天大学学报, 2025, 51(3): 806-815.
Chen Y J, Yu S S, Zhang X J. Ground risk quantitative assessment for UAV operations in urban low-altitude scenarios[J]. Journal of Beijing University of Aeronautics and Astronautics, 2025, 51(3): 806-815(in Chinese).

[23] Li S M, Cheng X C, Huang X D, et al. Cooperative conflict detection and resolution and safety assessment for 6G enabled unmanned aerial vehicles[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(2): 2183-2198.

[24] Chen E, Chen X, Jing W B, et al. Distributed tensor principal component analysis with data heterogeneity[J]. Journal of the American Statistical Association, 2025, 120(552): 2619-2631.

[25] 梅艺林, 崔立堃, 胡雪岩, 等. 基于人工势场法的复杂环境下多无人车避障与编队控制[J]. 工程科学学报, 2025, 47(2): 364-373.
Mei Y L, Cui L K, Hu X Y, et al. Obstacle avoidance and formation control of multiple unmanned vehicles in complex environments based on artificial potential field method[J]. Chinese Journal of Engineering, 2025, 47(2): 364-373(in Chinese).

[26] Zhang Z, Jiang J, Ling K V, et al. Real-time path planning for autonomous UAVs: an event-triggered multimodal adaptive pigeon-inspired optimization approach[J]. IEEE Transactions on Aerospace and Electronic Systems, 2025, 61(4): 10972-10981.

A safety-evaluation-driven UAV swarm collision avoidance decision-making based on pigeon-inspired optimization

PENG Yalan，DUAN Haibin*，WEI Chen

(National Key Laboratory of Aircraft Integrated Flight Control，School of Automation Science and Electrical Engineering，Beihang University，Beijing 100191，China)

Abstract：This research proposes an adaptive avoidance decision-making mechanism that combines safety assessment with pigeon-inspired optimization (PIO) to enhance the collision avoidance capabilities and coordination efficiency of unmanned aerial vehicle (UAV) swarms operating in complicated environments. A multidimensional safety evaluation index system is constructed by analyzing both local conflict risks and the overall coordination level of the swarm. Principal component analysis is employed to derive a comprehensive safety score, based on which a three-stage avoidance control strategy is formulated. An artificial potential field function with adaptively tuned parameters is designed according to the safety score, enabling responsive and autonomous individual-level control. Furthermore, to improve global cooperative avoidance under high-risk scenarios, the improved PIO algorithm is embedded to jointly optimize the velocity direction and avoidance parameters of swarm members. The effectiveness and resilience of the suggested strategy in dynamic contexts are validated by simulation findings, which show that it dramatically lowers the likelihood of local conflicts while preserving swarm cohesion.

Keywords：unmanned aerial vehicle；pigeon-inspired optimization；collision avoidance decision；safety evaluation；principal component analysis；artificial potential field

Received：2025-07-04；**Accepted**：2025-09-05；**Published Online**：2025-09-28 09：08

URL：link.cnki.net/urlid/11.2625.V.20250926.1757.004

Foundation items：National Natural Science Foundation of China (U24B20156, 62350048,T2121003)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0491

一种高精度的星光折射天文导航观星方法

詹先军[1,2]，王新龙[1,2,*]

(1. 北京航空航天大学 宇航学院，北京 100191；2. 航天器设计优化与动态模拟技术教育部重点实验室，北京 100191)

摘 要：在星光折射间接敏感地平定位方法中，传统固定视轴观星方案导致折射星数量少，且分布不均匀，进而影响星光折射定位的精度和可靠性。为此，提出了一种高精度的星光折射天文导航观星方案。根据平流层大气所形成的折射环带与星敏感器视场(FOV)之间的几何关系，建立了星敏感器视场对折射环带的覆盖面积与视轴高度角之间的关系模型，通过折射环带的覆盖面积最大化，结合飞行器的飞行参数，确定星敏感器视轴的最佳高度角，以提高可观测折射星数量；通过推导构建折射星空间构型与定位精度之间的定量关系模型，基于最小定位精度因子(PDOP)准则，确定折射星最佳空间几何构型，进而得到星敏感器视轴的最佳方位角，从而借助星光折射间接敏感地平定位方法实现对折射星连续、均匀的观测，以提高天文导航的定位精度。仿真验证结果表明，与传统观测方案相比，所提观测方案的位置与速度估计误差分别提升了75.7%和66.7%。

关 键 词：空天飞行器；星光折射；几何构型；星敏感器；天文导航
中图分类号：V448.22+4
文献标志码：A **文章编号**：1001-5965(2026)10-3627-08

空天飞行器具有全球到达、高超声速飞行、高动态等特点，对导航系统的自主性、精确性和可靠性提出了较高要求[1]。惯导系统自主性强、输出信息连续、可全空域全天候工作，但其误差随时间累积；而天文导航系统精度高、自主性强，且误差不随时间发散，可有效校正惯导误差[2]，从而为空天飞行器提供精确、自主、可靠的导航信息[3-4]。

目前，常用的天文导航定位方法主要有2种：直接敏感地平定位和星光折射间接敏感地平定位。直接敏感地平定位利用星敏感器和地平仪分别测量恒星方向和地心矢量方向，从而获得星光角距信息，并结合轨道动力学和滤波方法估计载体位置。然而，由于地球表面不规则，地平仪提供的地平基准的测量精度较低，与星敏感器的测量精度不相匹配，这将极大地影响系统的定位精度[5]。星光折射间接敏感地平定位方法是一种新型高精度天文定位技术，其利用星敏感器敏感经过地球大气折射的星光，结合大气对星光折射的数学模型，精确敏感地平，进而根据载体位置和折射星光之间的几何关系，解算得到载体的高精度三维位置信息[6]。研究表明，该方法无需依赖地平仪提供的地平基准，便可达到较高的定位精度，其不仅适用于地球卫星、载人飞船和空间站等大气层外的航天器[7-8]，而且可推广应用于高空远程长航时飞行器、临近空间飞行器等大气层内飞行器[9]，具有广阔的应用前景。

星光折射间接敏感地平定位方法通常观测经过20～50 km平流层大气折射的星光。因此，飞行器需通过姿态机动，使星敏感器视轴持续指向平流层，以稳定观测折射星并实现连续定位。而折射星观测方案直接决定了可观测折射星的数量及其空间几何构型，进而影响导航精度。另外，空天飞行器动态高、高度变化大，导致星敏感器的有效观星

收稿日期：2025-07-17；录用日期：2025-09-08；网络出版时间：2025-09-30 08:37
网络出版地址：link.cnki.net/urlid/11.2625.V.20250929.1317.004
基金项目：国家自然科学基金(61673040)；航空科学基金(20170151002)；天地一体化信息技术国家重点实验室基金(2015-SGIIT-KFJJ-DH-01)
*通信作者. E-mail：xlwon@163.com

引用格式：詹先军，王新龙. 一种高精度的星光折射天文导航观星方法[J]. 北京航空航天大学学报，2026，52(10)：3627-3634.
Zhan X J，Wang X L. A high-precision starlight refraction celestial navigation observation method[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3627-3634 (in Chinese).

时间窗口有限。因此，如何在有限的观星时间窗口内，快速捕获并稳定跟踪足够数量的折射星，并确保折射星空间构型最优，是实现高精度星光折射定位的关键[10]。

宁晓琳[11]、Qian[12]等根据折射星的空间分布特征，通过将星敏感器视轴指向地球边缘固定高度处，以捕获折射星光。Wang等[13]提出了一种适用于捷联式星敏感器的折射星观测姿态控制律，通过建立载体姿态角和星敏感器视轴高度角之间的约束关系，生成姿态控制指令，使星敏感器视轴指向大气平流层固定高度处。然而，这种固定观测方式使得星敏感器观测的折射星数量少，且空间分布不均匀，从而导致状态量的可观测度严重下降，进而影响星光折射定位的精度和可靠性。魏诗卉等[14]提出了一种通过观测不同方向的折射星，以提高系统可观测度的调姿观星策略，但未能建立具体的观星方向优化准则，从而限制了导航精度的进一步提升。

基于此，本文提出了一种高精度的星光折射天文导航观星方案，通过设计星敏感器视轴的最佳指向，连续、均匀地对折射星进行观测，进而提高天文导航的定位精度。

1 传统折射星光观测方案问题分析

由于地球大气为不均匀介质，当星光穿过大气层时，会因大气折射作用而向地心方向偏折，如图1所示。折射星是指星光通过地球大气层并发生折射的恒星。折射后的星光矢量为$\boldsymbol{u}'_{\mathrm{s}}$，而折射前的星光矢量为$\boldsymbol{u}_{\mathrm{s}}$。从飞行器$F$上看，恒星$s$的视光线相对于地表的视高度为$h_{\mathrm{a}}$。传统折射星光观测方案是将星敏感器固定安装在飞行器F上，并使星敏感器视轴$\boldsymbol{V}_{\mathrm{a}}$与飞行器位置矢量$\boldsymbol{r}$成固定夹角$\gamma$，从而测量得到星光穿过大气后星光矢量方向的变化量，即星光折射角R。

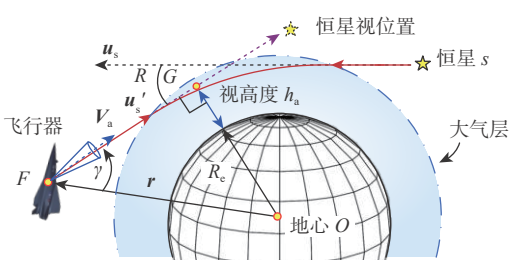


图1 星光折射间接敏感地平定位原理

Fig. 1 Principle of starlight refraction indirect horizon-sensing positioning

星光折射角R与视高度h_{a}之间的关系通常称为星光大气折射模型[15]，可根据折射角R确定折射星光在大气层中的视高度h_{a}，这个视高度观测量h_{a}反映了飞行器与地球之间的几何关系。根据图1中直角三角形OGF的几何关系，可得到星光折射间接敏感地平的量测方程为

$$h_{\mathrm{a}}=\sqrt{r_{\mathrm{s}}^2-\left|\boldsymbol{r}\boldsymbol{u}'_{\mathrm{s}}\right|^2}-R_{\mathrm{e}}+\delta h \tag{1}$$

式中：$\boldsymbol{r}=[x,y,z]^{\mathrm{T}}$为飞行器在地心惯性坐标系中的位置矢量；$r_{\mathrm{s}}=\sqrt{x^2+y^2+z^2}$为飞行器的地心距；$\boldsymbol{u}'_{\mathrm{s}}=[u'_x,u'_y,u'_z]^{\mathrm{T}}$为折射后星光矢量在地心惯性坐标系下的投影；$R_{\mathrm{e}}$为地球半径；$\delta h$为视高度误差，其主要来源于大气模型误差、地气光干扰和星敏感器测量误差，可建模为均值为零、标准差为σ的高斯白噪声[13]。

式(1)可进一步展开表示为

$$h_{\mathrm{a}}=\sqrt{x^2+y^2+z^2-\left(xu'_x+yu'_y+zu'_z\right)^2}-R_{\mathrm{e}}+\delta h \tag{2}$$

从式(2)中可以看出，视高度h_{a}可通过折射角R结合大气对星光折射的数学模型计算得到，折射后星光矢量(u'_x,u'_y,u'_z)可通过星敏感器测得，而地球半径R_{e}已知，式中只有飞行器位置x,y,z为未知量。因此，当观测3颗以上折射星时，便可根据折射星视高度h_{a}计算得到飞行器在地心惯性坐标系中的位置信息x,y,z。

然而，由于传统的折射星光观测方案采用固定观测方式，将星敏感器视轴指向大气平流层固定高度处，使得星敏感器观测的折射星数量少，且空间分布不均匀，进而严重影响星光折射导航系统的定位精度和可靠性。

2 高精度星光折射天文导航观星方案

针对传统固定视轴观星方案的不足，提出了一种高精度星光折射天文导航观星方案。通过设计星敏感器视轴的最佳高度角和最佳方位角，实现连续、均匀地观测折射星，从而提升定位精度。

2.1 星敏感器最佳视轴指向的确定方法

2.1.1 星敏感器视轴最佳高度角的确定

如果折射星视高度h_{a}太高，则折射效应非常微弱；如果视高度h_{a}太低，星光会经过大气不稳定的对流层，难以建立精确的大气折射模型；而在20～50 km平流层范围内，大气比较稳定。因此，星光折射定位方法通常选取视高度h_{a}为20～50 km的折射星光进行观测，如图2所示。从图2所示的几何关系可以看出，视高度h_{a}为20～50 km的折射星光，落在2条临界星光矢量$\boldsymbol{u}_{\mathrm{s1}}$和$\boldsymbol{u}_{\mathrm{s2}}$所夹的范围(即图中阴影区域)内。因此，星敏感器视轴方向矢量$\boldsymbol{V}_{\mathrm{a}}$与飞行器位置矢量$\boldsymbol{r}$的夹角$\gamma$(即星敏感器视轴高度角)应满足：

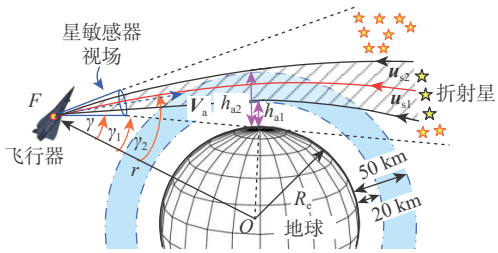


图 2　视高度为 20～50 km 的折射星光

Fig. 2　Refracted starlight with apparent height of 20-50 km

$$\gamma_1 \leqslant \gamma \leqslant \gamma_2 \tag{3}$$

式中: γ、γ_1和γ_2可分别表示为

$$\begin{cases} \gamma = \arccos[-(\boldsymbol{V}_a \boldsymbol{r})/r_s] \\ \gamma_1 = \arcsin[(R_e + h_{a1})/r_s] \\ \gamma_2 = \arcsin[(R_e + h_{a2})/r_s] \end{cases} \tag{4}$$

其中: $h_{a1} = 20$ km; $h_{a2} = 50$ km。

进一步, 将图 2 中的几何关系投影到天球坐标系中, 如图 3 所示。

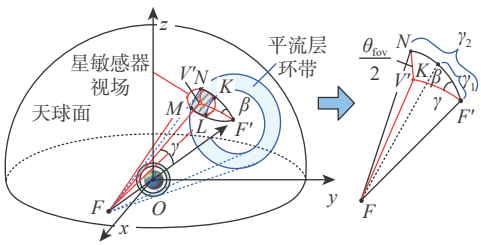


图 3　天球坐标系下星敏感器视轴指向示意图

Fig. 3　Schematic diagram showing the line of sight of a star sensor in celestial coordinate system

在天球面上, F'为飞行器位置F的投影; 20～50 km 平流层大气的投影为一个球面圆环; 星敏感器视场的投影为一个圆, 该圆和平流层圆环重叠的区域$KLMN$, 即为星敏感器观测的折射区域。通过调整星敏感器视轴指向, 该区域的面积可达到最大, 折射星观测概率可达到最高。

由图 3 可知, 当球面角β取最大值时, 折射区域$KLMN$的面积达到最大。在球面三角形$NV'F'$中, 根据球面三角形余弦定理, 可得

$$\cos(\theta_{fov}/2) = \cos\gamma\cos\gamma_2 + \sin\gamma\sin\gamma_2\cos\beta \tag{5}$$

式中: θ_{fov}为星敏感器视场角。

对式 (5) 进行变换可得球面角β的表达式为

$$\beta = \arccos\left[\frac{\cos(\theta_{fov}/2) - \cos\gamma\cos\gamma_2}{\sin\gamma\sin\gamma_2}\right] \tag{6}$$

将式 (6) 关于γ求偏导, 可得

$$\frac{\partial\beta}{\partial\gamma} = \frac{[\cos(\theta_{fov}/2)\cos\gamma - \cos\gamma_2]\sin\gamma_2}{(\sin\gamma\sin\gamma_2)^2\sqrt{1-\cos^2\beta}} \tag{7}$$

令 $\partial\beta/\partial\gamma = 0$, 则可得视轴最佳高度角$\gamma_o$为

$$\gamma_o = \arccos\left[\cos\gamma_2/\cos(\theta_{fov}/2)\right] \tag{8}$$

进一步, 将临界高度角γ_2的表达式 (4) 代入式 (8), 可得最佳高度角γ_o为

$$\gamma_o = \arccos\left[\frac{\sqrt{1-(R_e+h_{a2})^2/r_s^2}}{\cos(\theta_{fov}/2)}\right] \tag{9}$$

从式 (9) 可以看出, 星敏感器视场角θ_{fov}和折射星视高度上界h_{a2}均已知, 根据飞行器的地心距r_s, 可确定视轴最佳高度角γ_o。

2.1.2　星敏感器视轴最佳方位角的确定

1) 折射星构型与定位精度关系模型的建立。通过将n颗折射星的量测方程式 (2) 对飞行器位置x,y,z求偏导数可得

$$\boldsymbol{A}\Delta\boldsymbol{r} = \boldsymbol{V} \tag{10}$$

式中: $\Delta\boldsymbol{r} = [\mathrm{d}x \quad \mathrm{d}y \quad \mathrm{d}z]^{\mathrm{T}}$为飞行器位置误差向量; $\boldsymbol{V} = [\mathrm{d}h_{a1} \quad \mathrm{d}h_{a2} \quad \cdots \quad \mathrm{d}h_{an}]^{\mathrm{T}}$为$n$颗折射星的视高度误差组成的测量误差向量; 矩阵$\boldsymbol{A}$为

$$\boldsymbol{A} = [\boldsymbol{b}_1^{\mathrm{i}} \quad \boldsymbol{b}_2^{\mathrm{i}} \quad \cdots \quad \boldsymbol{b}_n^{\mathrm{i}}]^{\mathrm{T}} \tag{11}$$

其中: $\boldsymbol{b}_j^{\mathrm{i}} = [A_{jx} \quad A_{jy} \quad A_{jz}]^{\mathrm{T}}(j = 1,2,\cdots,n)$表示过地心$O$且与第$j$颗折射星折射后的星光矢量$\boldsymbol{u}'_{sj}$相垂直的单位矢量, 称为折射星地心法向矢量, A_{jx},A_{jy},A_{jz}可表示为

$$\begin{cases} A_{jx} = \left[x - u'_{jx}\left(xu'_{jx} + yu'_{jy} + zu'_{jz}\right)\right]/L_j \\ A_{jy} = \left[y - u'_{jy}\left(xu'_{jx} + yu'_{jy} + zu'_{jz}\right)\right]/L_j \\ A_{jz} = \left[z - u'_{jz}\left(xu'_{jx} + yu'_{jy} + zu'_{jz}\right)\right]/L_j \end{cases} \tag{12}$$

式中: $L_j = \sqrt{x^2 + y^2 + z^2 - \left(xu'_{jx} + yu'_{jy} + zu'_{jz}\right)^2}$。

解方程 (10) 可得, 载体位置误差$\Delta\boldsymbol{r}$可表示为

$$\Delta\boldsymbol{r} = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}\boldsymbol{V} \tag{13}$$

则星光折射定位误差方差阵$\boldsymbol{P}$为

$$\boldsymbol{P} = E\left[(\Delta\boldsymbol{r}_s)(\Delta\boldsymbol{r}_s)^{\mathrm{T}}\right] = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathrm{T}}E\left[\boldsymbol{V}\boldsymbol{V}^{\mathrm{T}}\right]\boldsymbol{A}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1} \tag{14}$$

当n颗折射星独立等精度观测时, 测量误差向量$\boldsymbol{V}$的协方差阵$E\left[\boldsymbol{V}\boldsymbol{V}^{\mathrm{T}}\right]$为

$$E\left[\boldsymbol{V}\boldsymbol{V}^{\mathrm{T}}\right] = \sigma_{dha}^2\boldsymbol{I}_{n\times n} \tag{15}$$

式中: σ_{dha}^2为视高度误差方差; $\boldsymbol{I}_{n\times n}$为$n$阶单位阵。

将式 (15) 代入式 (14), 可得

$$\boldsymbol{P} = \sigma_{dha}^2(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1} \tag{16}$$

则可得位置误差标准差σ_p为

$$\sigma_p = \sqrt{\mathrm{tr}(\boldsymbol{P})} = \sigma_{dha}\sqrt{\mathrm{tr}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}} \tag{17}$$

定义星光折射定位精度因子(position dilution of precision, PDOP)为

$$P_{DOP} = \sqrt{\mathrm{tr}(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})^{-1}} \tag{18}$$

将式(18)代入式(17),可得

$$\sigma_{\mathrm{p}}=\sigma_{\mathrm{dha}}P_{\mathrm{DOP}} \tag{19}$$

从式(19)可以看出,P_{DOP}越小,视高度误差σ_{dha}被放大为位置误差σ_{p}的程度越低,折射星几何构型越好。

2)折射星最优几何构型的确定。从式(18)可以看出,P_{DOP}取决于矩阵$\boldsymbol{A}$,而从式(11)可以看出,矩阵$\boldsymbol{A}$由n颗折射星的地心法向矢量$\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_n$决定。折射星地心法向矢量$\boldsymbol{b}_j$分布在以飞行器$F$为顶点,地心到飞行器的连线$OF$为轴线的圆锥面上,如图4所示。

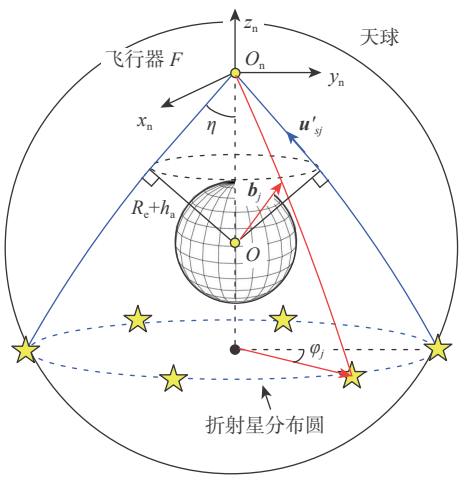


图4 折射星地心法向矢量的几何分布示意图

Fig. 4 Distribution of refracted star geocentric normal vector between the star and the Earth

折射后星光矢量$\boldsymbol{u}'_{sj}$和折射星地心法向矢量$\boldsymbol{b}_j$在地理坐标系$O_{\mathrm{n}}x_{\mathrm{n}}y_{\mathrm{n}}z_{\mathrm{n}}$下的分量可分别表示为

$$\boldsymbol{b}_j^{\mathrm{n}}=\begin{bmatrix}b_{jx}^{\mathrm{n}} & b_{jy}^{\mathrm{n}} & b_{jz}^{\mathrm{n}}\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}\sin\varphi_j\cos\eta & \cos\varphi_j\cos\eta & \sin\eta\end{bmatrix}^{\mathrm{T}} \tag{20}$$

$$\boldsymbol{u}_{sj}'^{\mathrm{n}}=\begin{bmatrix}u_{jx}'^{\mathrm{n}} & u_{jy}'^{\mathrm{n}} & u_{jz}'^{\mathrm{n}}\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}-\sin\varphi_j\sin\eta & -\cos\varphi_j\sin\eta & \cos\eta\end{bmatrix}^{\mathrm{T}} \tag{21}$$

式中:η为天底角,即折射后星光矢量$\boldsymbol{u}_{sj}'^{\mathrm{n}}$与飞行器位置矢量$\boldsymbol{r}$的夹角;$\varphi_j$为方位角,即折射星地心法向矢量$\boldsymbol{b}_j$在水平面内投影与$y_{\mathrm{n}}$轴的夹角。

而折射星地心法向矢量$\boldsymbol{b}_j$在地心惯性坐标系下的方向余弦可表示为

$$\boldsymbol{b}_j^{\mathrm{i}}=\boldsymbol{C}_{\mathrm{n}}^{\mathrm{i}}\boldsymbol{b}_j^{\mathrm{n}} \tag{22}$$

式中:$\boldsymbol{C}_{\mathrm{n}}^{\mathrm{i}}$为地理坐标系到地心惯性坐标系的转换矩阵。

将式(22)代入矩阵$\boldsymbol{A}$的表达式(11),可得

$$\boldsymbol{A}=\boldsymbol{B}\boldsymbol{C}_{\mathrm{i}}^{\mathrm{n}} \tag{23}$$

式中:$\boldsymbol{B}=\begin{bmatrix}\boldsymbol{b}_1^{\mathrm{n}} & \boldsymbol{b}_2^{\mathrm{n}} & \cdots & \boldsymbol{b}_n^{\mathrm{n}}\end{bmatrix}^{\mathrm{T}}$。

这样,将式(23)代入PDOP的定义式(18),可得

$$P_{\mathrm{DOP}}=\sqrt{\mathrm{tr}\left[\boldsymbol{C}_{\mathrm{n}}^{\mathrm{i}}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}\boldsymbol{C}_{\mathrm{i}}^{\mathrm{n}}\right]}=\sqrt{\mathrm{tr}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}} \tag{24}$$

将折射星地心法向矢量$\boldsymbol{b}_j$在地理坐标系下的方向余弦式(20)代入矩阵$\boldsymbol{B}$的表达式,可得

$$\boldsymbol{B}=\begin{bmatrix}\sin\varphi_1\cos\eta & \cos\varphi_1\cos\eta & \sin\eta\\ \sin\varphi_2\cos\eta & \cos\varphi_2\cos\eta & \sin\eta\\ \vdots & \vdots & \vdots\\ \sin\varphi_n\cos\eta & \cos\varphi_n\cos\eta & \sin\eta\end{bmatrix} \tag{25}$$

则有

$$\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}=\begin{bmatrix}\boldsymbol{C} & \boldsymbol{U}\\ \boldsymbol{U}^{\mathrm{T}} & \boldsymbol{D}\end{bmatrix} \tag{26}$$

式中:

$$\boldsymbol{C}=\begin{bmatrix}\cos^2\eta\sum_{j=1}^{n}\sin^2\varphi_j & \cos^2\eta\sum_{j=1}^{n}\cos\varphi_j\sin\varphi_j\\ \cos^2\eta\sum_{j=1}^{n}\cos\varphi_j\sin\varphi_j & \cos^2\eta\sum_{j=1}^{n}\cos^2\varphi_j\end{bmatrix}$$

$$\boldsymbol{U}=\begin{bmatrix}\cos\eta\sin\eta\sum_{j=1}^{n}\sin\varphi_j & \cos\eta\sin\eta\sum_{j=1}^{n}\cos\varphi_j\end{bmatrix}^{\mathrm{T}}$$

$$\boldsymbol{D}=n\sin^2\eta$$

由分块矩阵的求逆法可得,$(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}$为

$$(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}=\begin{bmatrix}(\boldsymbol{C}-\boldsymbol{U}\boldsymbol{D}^{-1}\boldsymbol{U}^{\mathrm{T}})^{-1} & -\boldsymbol{C}^{-1}\boldsymbol{U}(\boldsymbol{D}-\boldsymbol{U}^{\mathrm{T}}\boldsymbol{C}^{-1}\boldsymbol{U})^{-1}\\ -\boldsymbol{D}^{-1}\boldsymbol{U}^{\mathrm{T}}(\boldsymbol{C}-\boldsymbol{U}\boldsymbol{D}^{-1}\boldsymbol{U}^{\mathrm{T}})^{-1} & (\boldsymbol{D}-\boldsymbol{U}^{\mathrm{T}}\boldsymbol{C}^{-1}\boldsymbol{U})^{-1}\end{bmatrix} \tag{27}$$

由于$\boldsymbol{B}$由n个线性无关矢量$\boldsymbol{b}_j^{\mathrm{n}}(j=1,2,\cdots,n)$组成,$\boldsymbol{B}$为列满秩矩阵,$\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}$为正定矩阵。因此,$\boldsymbol{C}$、$\boldsymbol{D}$也为正定矩阵,则$\boldsymbol{U}^{\mathrm{T}}\boldsymbol{C}^{-1}\boldsymbol{U}$、$\boldsymbol{U}\boldsymbol{D}^{-1}\boldsymbol{U}^{\mathrm{T}}$均为半正定矩阵,有

$$\begin{aligned}\mathrm{tr}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}&=\mathrm{tr}(\boldsymbol{C}-\boldsymbol{U}\boldsymbol{D}^{-1}\boldsymbol{U}^{\mathrm{T}})^{-1}+\\&\quad\mathrm{tr}(\boldsymbol{D}-\boldsymbol{U}^{\mathrm{T}}\boldsymbol{C}^{-1}\boldsymbol{U})^{-1}\geqslant\mathrm{tr}(\boldsymbol{C}^{-1})+\mathrm{tr}(\boldsymbol{D}^{-1})\end{aligned} \tag{28}$$

当且仅当$\boldsymbol{U}=\boldsymbol{0}$时等号成立,即

$$\min(P_{\mathrm{DOP}})=\min\left[\sqrt{\mathrm{tr}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})^{-1}}\right]=\sqrt{\mathrm{tr}(\boldsymbol{C}^{-1})+\mathrm{tr}(\boldsymbol{D}^{-1})} \tag{29}$$

根据$\boldsymbol{U}$的定义式可知,$\boldsymbol{U}=\boldsymbol{0}$要求:

$$\sum_{j=1}^{n}\cos\varphi_j=\sum_{j=1}^{n}\sin\varphi_j=0 \tag{30}$$

将式(30)代入式(21)可得,P_{DOP}最小的条件为

$$\sum_{j=1}^{n}u_{jx}'^{\mathrm{n}}=\sum_{j=1}^{n}u_{jy}'^{\mathrm{n}}=0 \tag{31}$$

从式(31)可以看出,当n颗折射星的星光矢量

东向和北向分量之和均为零时, P_{DOP}最小。因此, 当n颗折射星之间的方位角相差$2\pi/n$时, 折射星几何构型最优, 定位精度最高。图 5 为 3、5 颗折射星对应的最优几何构型。

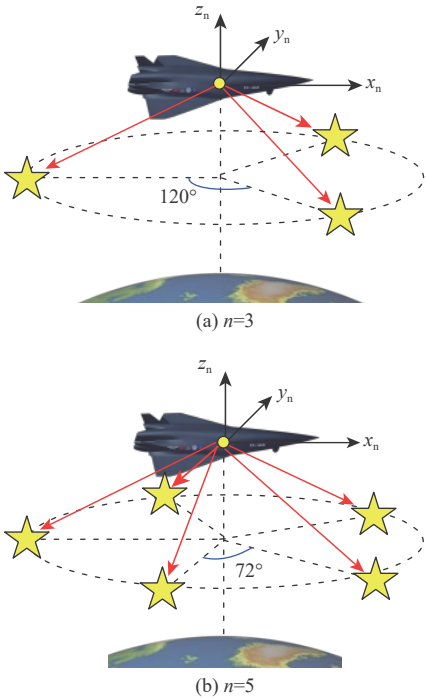


图 5　折射星的最优几何构型

Fig. 5　Optimal geometry of refracted stars

因此, 应用星敏感器进行n次连续观星定位时, 若相邻 2 次观星时星敏感器视轴方位角相差$2\pi/n$, 折射星几何构型最优, 定位精度最高。

2.2　折射星光观测实施方法

根据星敏感器视轴的最佳高度角和最佳方位角, 提出了一种高精度星光折射天文导航观星方案, 其实施流程如图 6 所示。

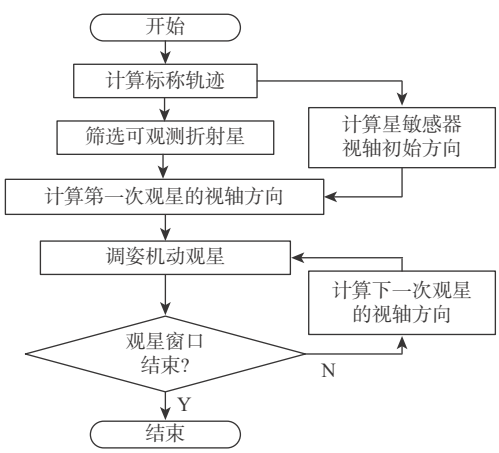


图 6　折射星光观测方案实施流程

Fig. 6　Flowchart of the refracted starlight observation plan

具体实施步骤如下:

1) 可观测折射星的筛选。根据空天飞行器的标称轨迹, 从星表中筛选出可观测折射星$s_1, s_2, \cdots, s_N$构成集合$U_{\mathrm{s}} = \{s_1, s_2, \cdots, s_N\}$。

2) 计算第一次观星的星敏感器视轴方向。根据星敏感器的初始视轴指向, 从可观测折射星集合U_{s}中, 选择距离初始视轴指向最近的折射星s_1进行观测。折射星s_1的方位角φ_1为

$$\varphi_1 = \arctan\left(s_{1y}^{\mathrm{n}} / s_{1x}^{\mathrm{n}}\right) \tag{32}$$

式中: s_{1x}^{n}、s_{1y}^{n}分别为折射星s_1的星光矢量$\boldsymbol{u}_{\mathrm{s}1}$在地理系下的东向、北向分量。

根据式 (9) 计算得到视轴最佳高度角γ_{o}, 结合式 (32) 计算得到折射星s_1的方位角φ_1, 则第一次观星的星敏感器视轴方向矢量$\boldsymbol{V}_1$为

$$\boldsymbol{V}_1 = \left[\sin\gamma_{\mathrm{o}}\cos\varphi_1 \quad \sin\gamma_{\mathrm{o}}\sin\varphi_1 \quad -\cos\gamma_{\mathrm{o}}\right]^{\mathrm{T}} \tag{33}$$

通过飞行器姿态机动将星敏感器视轴调整至式 (33) 所示的方向矢量$\boldsymbol{V}_1$, 即可使星敏感器视轴指向折射星s_1, 进行第一次观星。

3) 计算下一次观星的星敏感器视轴方向。将第一次观星的视轴方向矢量$\boldsymbol{V}_1$, 以飞行器位置矢量$\boldsymbol{r}$为轴线旋转$2\pi/n$, 得到第二次观星的视轴方向矢量$\boldsymbol{V}_2$, 如图 7 所示。

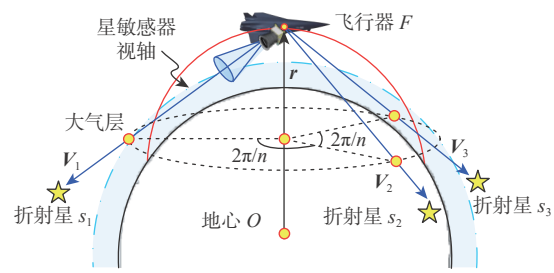


图 7　星敏感器折射星光观测示意图

Fig. 7　Observation of refracted starlight by star sensor

同理, 根据第二次观星的视轴方向矢量$\boldsymbol{V}_2$, 从可观测折射星集合U_{s}中, 选择距离星敏感器视轴最近的折射星s_2进行观测。以此类推, 继续进行下一次观星, 直至观星时间窗口结束。

3　仿真验证与分析

3.1　仿真条件

设定空天飞行器的发射点位置为 (34.5°N, 112.0°E), 海拔高度为 120 m, 发射方位角为 90°; 初始滚转角、偏航角和俯仰角分别为 0°, 0°, 90°; 飞行时长为 1 700 s。飞行器飞出大气层后, 通过姿态机动实现对 3 个方向折射星光的观测, 并通过卡尔曼滤波对惯导位置、速度误差进行校正。

星敏感器视场角为 5°×5°, 视高度测量误差为 120 m(1σ), 极限星等为 6.0 Mv, 输出频率为 10 Hz; 陀

螺常值漂移为 0.01 (°)/h, 随机漂移为 0.01 (°)/h; 加速度计常值误差为 100 μg, 随机误差为 50 μg; 惯导采样时间为 0.01 s, 初始平台失准角为 30″、50″、30″; 滤波器状态误差和误差协方差矩阵的初值为: $\delta \boldsymbol{X}_0 = [500\ \text{m}, 500\ \text{m}, 500\ \text{m}, 0.1\ \text{m/s}, 0.1\ \text{m/s}, 0.1\ \text{m/s}]^{\mathrm{T}}$, $\boldsymbol{P}_0 = \text{diag}\left[(500\ \text{m})^2, (500\ \text{m})^2, (500\ \text{m})^2, (0.1\ \text{m/s})^2, (0.1\ \text{m/s})^2, (0.1\ \text{m/s})^2\right]$。

所有算法均在配备 Core i7-10700H @ 2.90 GHz CPU 和 16 GB 内存的台式计算机上运行, 所使用的仿真平台为 MATLAB R2019a。

3.2 仿真结果与分析

3.2.1 不同折射星观测方案的导航精度

图 8 和图 9 分别给出了采用传统观测方案和本文观测方案时, 系统的位置和速度估计误差。

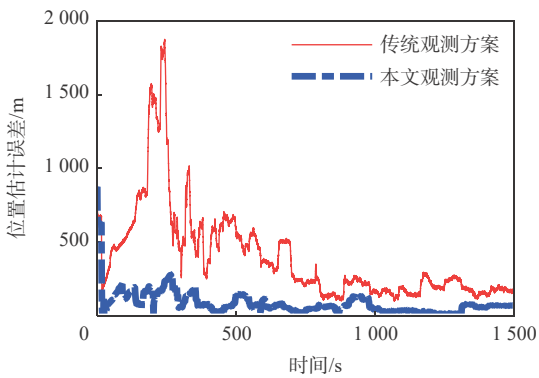


图 8 不同折射星观测方案的位置估计误差

Fig. 8 Position errors of different refracted star observation schemes

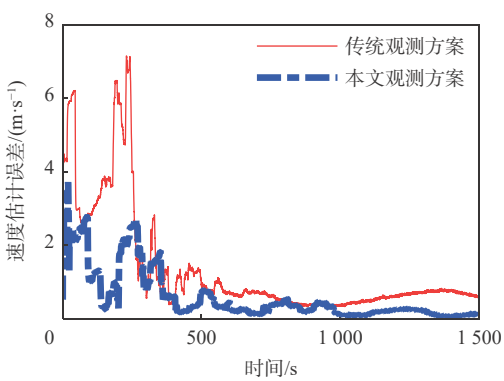


图 9 不同折射星观测方案的速度估计误差

Fig. 9 Velocity errors of different refracted star observation schemes

从图 8 和图 9 可以看出, 本文方案的速度与位置估计精度优于传统方案。这是由于传统方案将星敏感器视轴固定指向地球边缘特定高度处, 导致所观测的折射星几何构型单一, 使得系统可观测度降低, 进而导致星光折射定位精度下降。而本文方案通过建立折射星几何构型与定位精度的定量关系模型, 选择最优空间构型的折射星进行观测, 从而能够实现高精度定位。

为进一步验证本文方案的性能, 统计了不同折射星观测方案下的导航精度, 结果如表 1 所示。从表 1 可以看出, 与传统方案相比, 本文方案的位置、速度误差均方差分别降低了 75.7% 和 66.7%。这表明本文方案通过观测具有最优空间构型的折射星, 能够有效提高定位精度。

表 1 不同折射星观测方案的导航精度对比

Table 1 Comparison of navigation accuracy by different refracted star observation schemes

折射星光观测方案	位置估计误差/m		速度估计误差/(m·s⁻¹)	
	均值	均方差	均值	均方差
传统方案	100.28	143.85	0.34	0.51
本文观测方案	45.48	34.95	0.06	0.17

3.2.2 不同观测方位角对导航精度的影响

观测方位角直接影响折射星几何构型, 从而影响定位精度。图 10 和图 11 给出了观测方位角之差 $\Delta\varphi$ 分别为 30°、60°和 120°时, 系统的位置和速度估计误差。

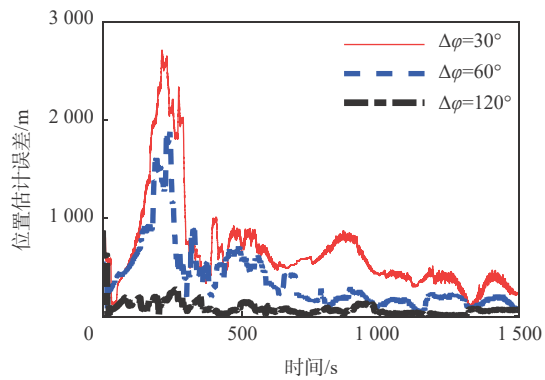


图 10 不同观测方位角下的位置估计误差

Fig. 10 Position errors under different observation azimuths

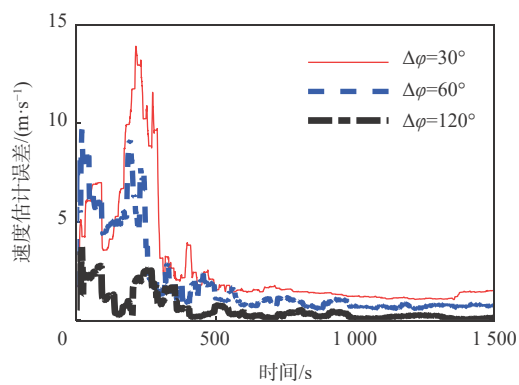


图 11 不同观测方位角下的速度估计误差

Fig. 11 Velocity errors under different observation azimuths

从图 10 和图 11 可以看出, 当观测方位角之差为 120°时, 由于几何精度因子 P_{DOP} 最小, 折射星空间构型最优, 因此导航精度最高。

为进一步验证观测方位角对导航性能的影响, 统计了不同观测方位角的导航精度, 如表 2 所示。从表 2 可以看出, 当观测方位角之差为 120°时, 位置估计误差均方差较 30°和 60°构型分别降低了 90.0% 和 72.2%, 速度估计误差均方差分别降低了 80.2% 和 71.2%。本文方案通过建立折射星方位角与定位精度的关系模型, 选择最优方位角的折射星进行观测, 从而能够提高定位精度。

表 2　不同观测方位角的导航精度对比

Table 2　Comparison of navigation accuracy of different observation azimuths

观测方位角之差/(°)	位置估计误差/m		速度估计误差/(m·s^{-1})	
	均值	均方差	均值	均方差
30	186.22	349.77	0.81	0.86
60	103.72	125.74	0.24	0.59
120	45.48	34.95	0.06	0.17

3.2.3　不同飞行高度对导航精度的影响

图 12 和图 13 给出了载体飞行高度 H 分别为 500 km、1 000 km 和 2 000 km 时, 系统的位置和速度估计误差。不同飞行高度下的导航精度如表 3 所示。

从图 12 和图 13 可以看出, 当载体飞行高度较低时, 由于 20～50 km 平流层大气形成的折射环带

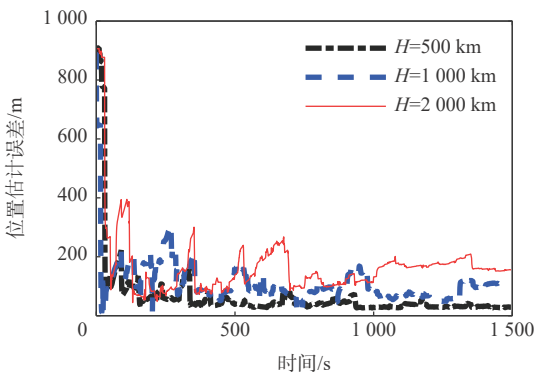


图 12　不同飞行高度下的位置估计误差

Fig. 12　Position errors at different flight heights

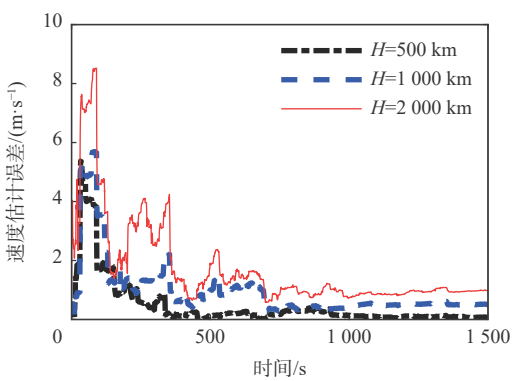


图 13　不同飞行高度下的速度估计误差

Fig. 13　Velocity errors at different flight heights

表 3　不同飞行高度的导航精度对比

Table 3　Comparison of navigation accuracy at different flight heights

飞行高度/km	位置估计误差/m		速度估计误差/(m·s^{-1})	
	均值	均方差	均值	均方差
500	45.48	34.95	0.06	0.17
1 000	86.97	75.74	0.33	0.28
2 000	163.76	112.40	0.56	0.53

在星敏感器视场内的覆盖面积增加, 使得可见折射星数量增多, 定位精度较高; 而当载体的飞行高度增大时, 虽然折射星数量会随高度增加而减少, 但通过所提折射星光观测方案, 仍可维持足够的可观测星数, 并保持较高的导航精度。

从表 3 可以看出, 当飞行高度为 500 km 时, 系统的位置估计误差均方差较 1 000 km 和 2 000 km 时分别降低了 53.9% 和 68.9%, 速度估计误差均方差分别降低了 39.3% 和 67.9%。这表明不同载体飞行高度会对星光折射定位精度产生影响。

4　结　论

针对传统固定视轴观星方案导致可观测折射星少且分布不均, 影响星光折射定位精度和可靠性的问题, 提出一种高精度星光折射天文导航观星方案, 经理论分析与仿真验证得到以下结论:

1) 本文方案通过建立星敏感器视场对折射环带的覆盖面积与视轴高度角的关系模型, 并使折射环带覆盖面积最大化, 能够确定最佳视轴高度角, 进而增加了可观测折射星数。

2) 当n颗折射星的方位角相差$2\pi/n$时, 星光折射定位精度因子 P_{DOP} 最小, 折射星空间构型最优, 定位精度最高。

3) 与传统方案相比, 本文方案通过观测具有最优空间构型的折射星, 能够显著提升系统的可观测度与定位精度, 其位置、速度估计误差分别降低了 75.7% 和 66.7%。

本文方案不仅适用于空天飞行器, 也可推广应用于地球卫星、载人飞船和空间站等大气层外航天器, 以及高空远程长航时飞行器、临近空间飞行器等大气层内飞行器, 具有广阔的应用前景。

参考文献 (References)

[1] 魏昊, 蔡光斌, 凡永华, 等. 高超声速飞行器再入滑翔段在线制导[J]. 北京航空航天大学学报, 2025, 51(1): 183-192.
Wei H, Cai G B, Fan Y H, et al. Online guidance for hypersonic vehicles in glide-reentry segment[J]. Journal of Beijing University of Aeronautics and Astronautics, 2025, 51(1): 183-192(in Chinese).

[2] Song G X, Yang Y Q, Ning X L, et al. A new SINS/RCNS integ-

rated navigation method based on star pixel coordinates[J]. Science China Information Sciences, 2023, 66(9): 199201.

[3] 王鼎杰, 吕汉峰, 吴杰, 等. 星光折射辅助惯性/天文紧组合高精度导航方法[J]. 中国惯性技术学报, 2020, 28(1): 74-81.
Wang D J, Lv H F, Wu J, et al. Stellar refraction-aided SINS/CNS tightly-integrated navigation method[J]. Journal of Chinese Inertial Technology, 2020, 28(1): 74-81(in Chinese).

[4] Ning X L, Sun X H, Fang J C, et al. Satellite stellar refraction navigation using star pixel coordinates[J]. Navigation, 2019, 66(1): 129-138.

[5] Zhu J F, Wang X L, Li H N, et al. A high-accuracy SINS/CNS integrated navigation scheme based on overall optimal correction[J]. Journal of Navigation, 2018, 71(6): 1567-1588.

[6] Zhao Y, Wang D J, Zhang H B, et al. A novel tightly-coupled SINS/RCNS integrated navigation method considering atmospheric density error[J]. IEEE Transactions on Instrumentation and Measurement, 2023, 72: 8503011.

[7] Ning X L, Sun X H, Chao W. Integration of star pixel coordinates and their time differential measurement in satellite stellar refraction navigation[J]. Acta Astronautica, 2019, 159: 286-293.

[8] Yang H, Wang D J, Zhang H B, et al. Autonomous strapdown inertial/refractive celestial integrated navigation using dual-band stellar refraction observations[J]. IEEE Transactions on Aerospace and Electronic Systems, 2025, 62: 2900-2912.

[9] Zhan X J, Wang X L, Hu X D, et al. A new method for high-precision starlight refraction indirect horizon-sensing positioning of atmospheric flight vehicles[J]. Advances in Space Research, 2024, 73(8): 3884-3895.

[10] 宋峰, 张奕群. 星光大气折射观测导航方法可观性研究[J]. 现代防御技术, 2012, 40(2): 65-70.
Song F, Zhang Y Q. Research on observability of navigation approach using starlight refraction measurements[J]. Modern Defence Technology, 2012, 40(2): 65-70(in Chinese).

[11] 宁晓琳, 王龙华, 白鑫贝, 等. 一种星光折射卫星自主导航系统方案设计[J]. 宇航学报, 2012, 33(11): 1601-1610.
Ning X L, Wang L H, Bai X B, et al. A scheme design of satellite autonomous navigation system based on stellar refraction[J]. Journal of Astronautics, 2012, 33(11): 1601-1610(in Chinese).

[12] Qian H M, Sun L, Cai J N, et al. A starlight refraction scheme with single star sensor used in autonomous satellite navigation system[J]. Acta Astronautica, 2014, 96: 45-52.

[13] Wang H Y, Gao Z Q, Wang T F, et al. Study on command attitude law for refracted starlight observation in SINS/RCNS integrated navigation[J]. Advances in Space Research, 2018, 62(3): 721-731.

[14] 魏诗卉, 杨春伟, 刘炳琪, 等. 星光折射连续修正方法及制导规划[J]. 航空学报, 2020, 41(8): 54-63.
Wei S H, Yang C W, Liu B Q, et al. Stellar refraction: continuous correction and guidance planning[J]. Acta Aeronauticaet Astronautica Sinica, 2020, 41(8): 54-63(in Chinese).

[15] Ning X L, Wang F, Fang J C. Implicit UKF and its observability analysis of satellite stellar refraction navigation system[J]. Aerospace Science and Technology, 2016, 54: 49-58.

A high-precision starlight refraction celestial navigation observation method

ZHAN Xianjun[1, 2], WANG Xinlong[1, 2, *]

(1. School of Astronautics, Beihang University, Beijing 100191, China;
2. Key Laboratory of Spacecraft Design Optimization and Dynamic Simulation Technology, Ministry of Education, Beijing 100191, China)

Abstract: The traditional fixed optical axis observation approach reduces positioning accuracy and reliability in the starlight refraction indirect horizon-sensing positioning method by producing a small number of refracted stars with an uneven distribution. Therefore, a high-precision starlight refraction celestial navigation observation method is proposed. Based on the geometric relationship between the stratospheric refraction band and the star sensor's field of view (FOV), a relation model is established between the FOV coverage area of the refraction band and the star sensor's elevation angle. By maximizing the FOV coverage area and combining the flight parameters, the optimal elevation angle is determined to increase the observable refracted stars. Additionally, a quantitative relationship model is developed between positioning precision and the spatial arrangement of refracted stars. Based on the minimum position dilution of precision (PDOP) criterion, the optimal spatial geometric configuration of refracted stars is derived, thereby obtaining the optimal azimuth angle of the optical axis. This enables continuous and uniform observation for refracted starlight, so as to improve the positioning accuracy of celestial navigation. Simulation results demonstrate that compared with the conventional method, the proposed method achieves 75.7% and 66.7% improvements in position and velocity estimation accuracy, respectively.

Keywords: aerospace vehicle; starlight refraction; geometric configuration; star sensor; celestial navigation

Received: 2025-07-17; **Accepted**: 2025-09-08; **Published Online**: 2025-09-30 08:37

URL: link.cnki.net/urlid/11.2625.V.20250929.1317.004

Foundation items: National Natural Science Foundation of China (61673040); Aeronautical Science Foundation of China (20170151002); Open Research Fund of State Key Laboratory of Space-Ground Integrated Information Technology (2015-SGIIT-KFJJ-DH-01)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0468

基于在轨加注的载人登月任务分段式轨道设计

李其玲[1]，师鹏[1,*]，陈勇[2]
(1. 北京航空航天大学 宇航学院，北京 102206；2. 成都流体动力创新中心，成都 610031)

摘 要：针对单颗运载器运载能力有限但大规模目标载荷部署需求庞大的矛盾，采用由在轨加注衔接天地往返与空间转移的“两段式”载人登月任务模式，对使用在轨加注方法的载人登月任务的规模进行分析与计算，建立了发射段、交会段、地月段、落月段等4个主要过程的动力学模型，并完成任务的总体设计仿真分析。运用质量估算方法对比了传统模式与在轨加注模式下的任务总质量与各级规模；构建了包含发射、交会对接、地月转移和月面着陆等4个阶段的任务模型，明确了各阶段约束，并建立了考虑低温推进剂在轨蒸发损耗的质量模型；计算任务窗口并开展全流程轨道仿真，分析各阶段参数与推进剂消耗；采用遗传算法优化落月段参数。质量估算结果表明：相较于传统模式所需数千吨级起飞质量，在轨加注方案起飞总质量为500 t以下，空间转移级质量为数十吨级；全流程仿真结果进一步明确了在轨加注方案各阶段推进剂消耗量，为工程设计提供必要参考。

关 键 词：载人登月；在轨加注；推进剂；航天运输；质量规模
中图分类号：V412
文献标志码：A **文章编号**：1001-5965(2026)10-3635-12

以探月为代表的未来航天任务，具有规模大、距离远的特点，对航天运输系统的运载能力提出了更高的要求。现有载人登月轨道设计方法存在两大局限：一是传统“一站式”运输模式需携带全部推进剂入轨，导致发射规模达数千吨级；二是航天器设计需同时兼顾大气层穿越与深空转移功能，制约运输效率提升[1-2]。

针对上述矛盾，有研究者提出应用在轨推进剂仓库(即“太空加注站”)的太空运输架构，由低成本火箭向加注站输送推进剂，目标航天器抵达在轨加注站后实现推进剂补给。基于在轨加注，航天器获得燃料补给后，可以直接搭载有效载荷从空间轨道出发，实现天地往返与空间转移运输两大过程“解耦”[3-10]。

然而，现有研究存在2点关键不足：①多将轨道设计与加注站部署解耦，缺乏轨道参数(如地月转移窗口)与加注站位置(如近地轨道(low Earth orbit, LEO))的联合分析，导致加注效率与轨道经济性难以协同；②传统轨道设计以“一站式”运输为基础，未充分考虑在轨加注带来的多阶段任务拆分特性，模型约束与实际任务需求匹配度低。

基于上述背景，本文进行了以下工作：①采用“两段式”载人登月任务模式，以LEO加注站为加注节点，实现天地往返与空间转移的解耦；②建立耦合加注站位置的发射、交会、地月、落月4段动力学模型，明确各阶段速度增量、推力约束。

1 任务规模对比分析

考虑到构建在轨加注站的成本，参考汪小卫等[11-13]提出的“接力式”地月空间航天运输模式，采用由在轨加注衔接天地往返与空间转移“两段式”的载人登月任务飞行模式：空间转移运输系统长期停泊于轨道，执行任务前前往加注站进行推进剂在轨加注并等待空间转移窗口；月球着陆器发射至加

收稿日期：2025-07-14；**录用日期**：2025-10-17；**网络出版时间**：2025-11-05 18:19
网络出版地址：link.cnki.net/urlid/11.2625.V.20251105.1107.003
基金项目：国家自然科学基金面上项目(12672454)
*通信作者. E-mail：shipeng@buaa.edu.cn

引用格式：李其玲，师鹏，陈勇. 基于在轨加注的载人登月任务分段式轨道设计[J]. 北京航空航天大学学报，2026，52(10)：3635-3646.
Li Q L，Shi P，Chen Y. Phased orbit design for manned lunar landing mission with on-orbit refueling[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3635-3646 (in Chinese).

注站，与此同时，载人飞船也由运载火箭搭载发射，将航天员送至加注站；航天员转移至月球着陆器后，由空间转移运输系统送至环月轨道（lunar orbit, LO）；航天员完成月表勘测任务后，乘坐月球着陆器返回LO，由空间转移运输系统将月球着陆器送至加注站，随后航天员换乘载人飞船的返回舱返回地球。上述过程如图1所示。

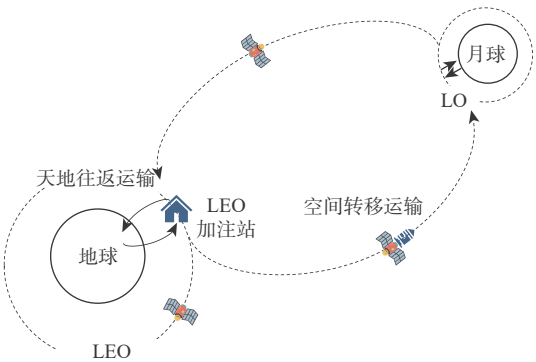


图1 基于在轨加注的地月空间往返运输轨道示意图

Fig. 1 Schematic of cislunar round-trip transportation orbit with on-orbit refueling

为评估在轨加注方案的优势，定义2种任务模式进行对比分析：方案1为在轨加注模式，即基于LEO加注站的“两段式”任务架构；方案2为传统模式，即基于近月轨道一次对接的传统任务模式。

采用模块化质量规模估算方法进行质量规模估算，主要包括建立质量数据库、确定任务过程、求解质量规模三大步骤[14]。具体计算过程如下：

1）建立质量数据库。方案1和方案2中的航天器包括运载火箭、载人飞船、月球着陆器、空间转移运载器。方案1中，运载火箭将载人飞船和月球着陆器分别发射至LEO，地月转移所需速度增量由空间转移运载器提供，月球着陆和起飞由月球着陆器推进系统支持；方案2中，运载火箭将载人飞船和月球着陆器分别送达LO，航天器在月球轨道完成组装。由此可建立结构质量数据库，如表1所示。

表1 结构质量数据库

Table 1 Structural mass database

参数	数值/t
方案1火箭结构质量	30
方案2火箭结构质量	130
载人飞船结构质量	8
载人飞船返回舱结构质量	2
月球着陆器结构质量	4.2
空间转移运载器结构质量	8

2）确定任务过程。根据航天器在轨飞行速度确定速度增量，2种方案下的任务方案及速度增量分别如表2和表3所示（仅展示航天员参与的飞行阶段）。方案1任务阶段包括：地面发射、LEO、LEO→加注站、加注站、加注站→LO（进入转移轨道）、加注站→LO（制动进入LO）、LO→月面、月面→LO、LO、LO→加注站、加注站、再入返回地球12个阶段。方案2任务阶段包括：地面发射、LEO、LEO→LO（进入转移轨道）、LEO→LO（制动进入LO）、LO、LO→月面、月面→LO、LO、LO→LEO、再入返回地球10个阶段。由于加注站的存在，方案1避免了使用重型火箭从LEO直接发射到近月轨道，其所需的速度增量需求明显小于方案2，从而显著减小质量规模。

3）求解质量规模。由各任务方案最后一个任务阶段的质量开始，根据齐奥尔科夫斯基公式和所抛弃的航天器质量可以依次推导上一阶段系统的总质量，进而获得任务开始时系统的质量规模。

表2 方案1任务剖面及速度增量需求

Table 2 Mission profile and velocity increment requirements for Solution 1

序号	任务阶段	航天器	速度增量/(km·s^{-1})	动力来源
1	地面发射	运载火箭1、载人飞船	10.00	运载火箭1
2	LEO	载人飞船	0	无
3	LEO→加注站	载人飞船	0.31	载人飞船推进舱
4	加注站	月球着陆器、空间转移运载器	0	无
5	加注站→LO（进入转移轨道）	月球着陆器、空间转移运载器	3.08	空间转移运载器
6	加注站→LO（制动进入LO）	月球着陆器、空间转移运载器	3.48	空间转移运载器
7	LO→月面	月球着陆器	2.10	月球着陆器
8	月面→LO	月球着陆器	2.00	月球着陆器
9	LO	月球着陆器、空间转移运载器	0	无
10	LO→加注站	月球着陆器、空间转移运载器	0.65	空间转移运载器
11	加注站	载人飞船	0	无
12	再入返回地球	载人飞船返回舱	0.50	返回舱气动减速

表 3　方案 2 任务剖面及速度增量需求

Table 3　Mission profile and velocity increment requirements for Solution 2

序号	任务阶段	包含航天器	速度增量/(km·s^{-1})	动力来源
1	地面发射	运载火箭2、载人飞船	10.00	运载火箭2
2	LEO	运载火箭2、载人飞船	0	无
3	LEO→LO(进入转移轨道)	运载火箭2、载人飞船	3.13	运载火箭2
4	LEO→LO(制动进入LO)	载人飞船	3.48	运载火箭2
5	LO	月球着陆器、载人飞船	0	无
6	LO→月面	月球着陆器	2.10	月球着陆器
7	月面→LO	月球着陆器	2.00	月球着陆器
8	LO	月球着陆器、载人飞船	0	无
9	LO→LEO	载人飞船	0.65	载人飞船推进舱
10	再入返回地球	载人飞船返回舱	0.50	返回舱气动减速

假设整个飞行过程共有 n 个阶段, 在第 i 个阶段中($2\leqslant i\leqslant n$), 速度增量为 Δv_i, 抛弃的质量为 Δm_i, 第 i 阶段结束时质量为 m_i, 则第 i−1 阶段结束时的质量为[15]

$$m_{i-1}=(m_i+\Delta m_i)e^{\frac{\Delta v_i}{I_{sp}g}} \tag{1}$$

式中: I_{sp} 为发动机比冲; g 为重力加速度; e 为轨道偏心率。

由此求得 2 种方案所需的系统质量规模对比如表 4 所示。基于在轨加注的载人登月任务方案系统质量规模显著低于基于近月轨道一次对接的载人登月任务方案, 充分展现了基于在轨加注的载人登月任务方案的优势。

表 4　2 种方案所需质量规模对比

Table 4　Mass scale comparison for both solutions

方案	质量规模/t
基于在轨加注的载人登月任务方案	486.7
基于一次对接的载人登月任务方案	3 906.7

2　LEO 加注轨道飞行任务建模

LEO 加注轨道飞行任务包括发射段、交会段、地月段、落月段, 飞行时序如图 2 所示。发射段主要确定火箭的飞行时序、质量参数及航天器从地面发射进入 LEO 的入轨参数; 交会段需实现航天器与 LEO 加注站的交会对接, 主要确定变轨时刻、速

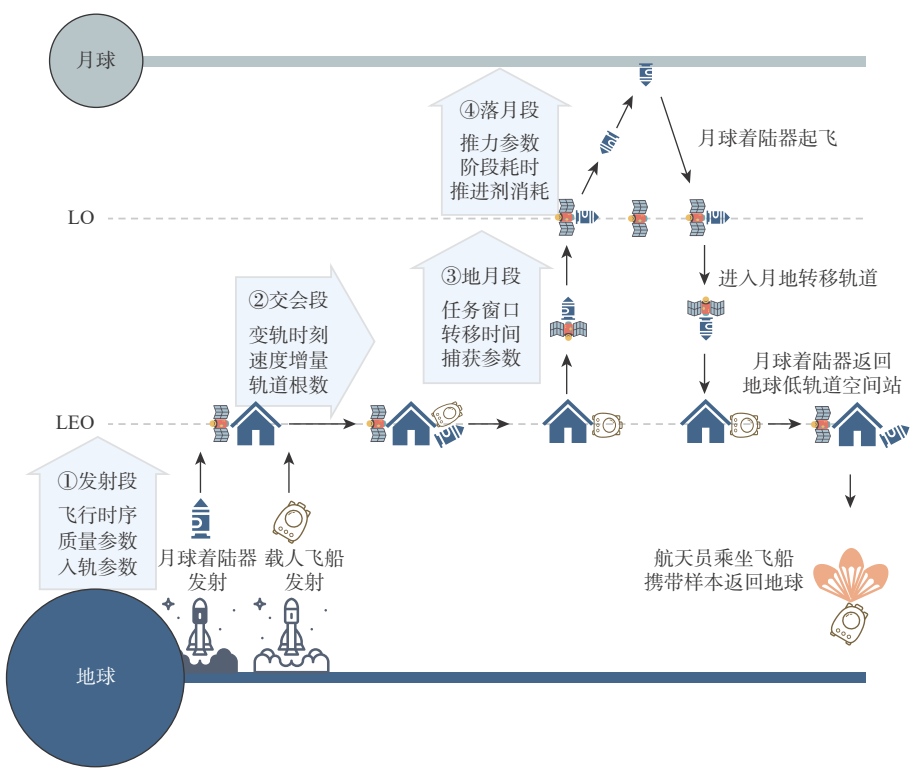


图 2　LEO 加注轨道飞行任务时序图

Fig. 2　LEO orbit injection mission timeline diagram

度增量和轨道根数；地月段为空间转移运输系统与月球着陆器组合体的地月转移轨道，主要确定任务窗口、转移时间和捕获参数；落月段主要确定推力大小、阶段耗时和推进剂消耗量参数。

2.1 发射段建模

发射段考虑两级火箭发动机、助推器的作用过程，将飞行阶段分为5个时序。在发射坐标系下建立火箭飞行的动力学和运动学微分方程，采用四阶龙格库塔法进行数值积分，即可得到火箭发射段的入轨参数。发射段的主要模型如下：

1）飞行时序。

① t_0 时刻：火箭助推器和一级火箭发动机开始工作。

② t_1 时刻：火箭助推器分离，一级火箭发动机继续工作。

③ t_2 时刻：一级火箭发动机分离，二级火箭发动机工作。

④ t_2+t_3：二级火箭主机发动机关机，游机继续工作。

⑤ t_2+t_4：二级火箭游机发动机关机，完成器箭分离，有效载荷入轨。

2）质量模型。

根据火箭的参数有2个子级，助推器质量为

$$\begin{cases} m_{o2}=\dfrac{m_{o1}}{e_1} \\ m_{o3}=\dfrac{m_{o1}}{e_2} \end{cases} \tag{2}$$

式中：m_{o1}为第一级火箭质量；m_{o2}为第二级火箭质量；m_{o3}为助推器质量；e_1为第一级和第二级火箭质量比；e_2为第一级火箭和助推器质量比；e_1、e_2均为无量纲。

由于火箭发动机喷射推进剂，火箭的质量随时间变化，其微分方程为

$$\dot{m}=-\dot{m}_{\mathrm{f}} \tag{3}$$

式中：$\dot{m}_{\mathrm{f}}$为推进剂的质量流量，kg/s。

假设火箭飞行各阶段燃料质量占燃料所在级的总质量的比值为定值，则火箭各级累计消耗燃料质量为

$$\begin{cases} m_{f1}=m_{o1}\dfrac{m_{f1}}{m_{o1}} \\ m_{f2}=\dfrac{m_{o1}}{e_1}\cdot\dfrac{m_{f2}}{m_{o2}} \\ m_{f3}=\dfrac{m_{o1}}{e_2}\cdot\dfrac{m_{f3}}{m_{o3}} \end{cases} \tag{4}$$

式中：m_{f1}为0～t_2时间内消耗的第一级燃料的质量；m_{f2}为t_2～t_3时间内消耗的第二级燃料的质量；m_{f3}为0～t_1时间内消耗的助推器燃料的质量。

同步考虑助推器、一级、整流罩等结构分离，则火箭在飞行过程中的质量为

$$m=\begin{cases} m_{o1}\left(1+\dfrac{1}{e_1}+\dfrac{1}{e_2}\right)-(\dot{m}_{f1}+\dot{m}_{f3})t+m_s+m_p & 0<t<t_1 \\ m_{o1}\left(1+\dfrac{1}{e_1}\right)-\dot{m}_{f1}t+m_s+m_p & t_1<t<t_2 \\ \dfrac{m_{o1}}{e_1}-(\dot{m}_{f21}+\dot{m}_{f22})(t-t_2)+m_p & t_2<t<t_2+t_3 \\ \dfrac{m_{o1}}{e_1}+m_p-\dot{m}_{f21}t_3-\dot{m}_{f22}(t-t_2) & t_2+t_3<t<t_2+t_4 \\ \dfrac{m_{o1}}{e_1}+m_p-m_{f2} & t>t_2+t_4 \end{cases} \tag{5}$$

式中：m_s为整流罩、逃逸塔等箭体结构质量；m_p为有效载荷质量。为方便计算，假设一级结构分离后，整流罩、逃逸塔等同时分离。

对应时刻有

$$\begin{cases} t_1=\dfrac{m_{f3}}{\dot{m}_{f3}} \\ t_2=\dfrac{m_{f1}}{\dot{m}_{f1}} \\ t_3=C \\ t_4=\dfrac{m_{f2}-\dot{m}_{f21}t_3}{\dot{m}_{f22}} \end{cases} \tag{6}$$

3）动力学及运动学模型。在发射坐标系下建立火箭飞行的动力学和运动学微分方程如下。

动力学方程：

$$\begin{cases} \dfrac{\mathrm{d}v_x}{\mathrm{d}t}=\dfrac{P}{m}\cos\varphi\cos\psi+\dfrac{D}{m}\cdot\dfrac{v_x}{v}-\dfrac{x}{r}g+2(\omega_x v_y-\omega_y v_z) \\ \dfrac{\mathrm{d}v_y}{\mathrm{d}t}=\dfrac{P}{m}\sin\varphi\cos\psi+\dfrac{D}{m}\cdot\dfrac{v_x}{v}-\dfrac{R_0+y}{r}g+2(\omega_x v_z-\omega_z v_x) \\ \dfrac{\mathrm{d}v_z}{\mathrm{d}t}=-\dfrac{P}{m}\sin\psi+\dfrac{D}{m}\cdot\dfrac{v_x}{v}-\dfrac{z}{r}g+2(\omega_y v_x-\omega_x v_y) \end{cases} \tag{7}$$

式中：P为火箭发动机推力；R_0为发射点至地心的距离，取地球参考半径；D为气动阻力；r为月球着陆器质心到月心的距离；v_x、v_y、v_z分别为飞行器速度在坐标系x、y、z轴的分量；φ为俯仰角，火箭纵轴与发射坐标系Oxy水平面的夹角；ψ为偏航角，火箭纵轴在发射坐标系Oxy平面的投影与Ox轴的夹角；ω_x、ω_y、ω_z分别为地球自转角速度矢量在发射坐标系x、y、z轴的分量。

运动学方程：

$$\begin{cases} \dfrac{\mathrm{d}x}{\mathrm{d}t}=v_x \\ \dfrac{\mathrm{d}y}{\mathrm{d}t}=v_y \\ \dfrac{\mathrm{d}z}{\mathrm{d}t}=v_z \end{cases} \tag{8}$$

2.2 交会段建模

入轨后，载人飞船与月球着陆器需进行自主变

轨, 与加注站进行交会对接。由于动力学过程类似, 故仅考虑载人飞船的情况。采用近圆轨道的变轨方式进行轨道机动, 变轨过程设计如下。

1) 第 1 次变轨。由于变轨前后角度改变量是小量, 变轨的作用点是载人飞船和加注站轨道的交点, 所需速度增量 Δv_1 与角度改变量 δ(即轨道平面夹角)和当地轨道速度 v_{orb} 成正比, 即 $\Delta v_1 \approx v_{\text{orb}}\delta$, 方向与当地轨道速度垂直, 作用点纬度幅角为 u_1。解球面三角形:

$$\begin{cases} \cos i_2 = \cos i_1 \cos\delta - \sin i_1 \sin\delta \cos u_1 \\ \sin(\Omega_2 - \Omega_1) = \dfrac{\sin\delta \sin u_1}{\sin i_2} \end{cases} \tag{9}$$

式中: i 为轨道倾角; δ 为轨道平面夹角; Ω 为升交点赤经。

代入参数解得 u_1 与 δ, 确定第 1 次变轨点。在该变轨点, Δv_1 的方向沿轨道平面动量矩相反方向。距离初始时刻经过时间:

$$\Delta t_1 = \frac{\Delta M_1}{n_{\text{MS}}} \tag{10}$$

式中: n_{MS} 为载人飞船在当前轨道上的平均角速度; M 为平近点角。

第 1 次变轨后, 载人飞船轨道的近地点幅角发生改变:

$$\Delta\omega = \frac{2\sin\theta}{ev}\Delta v_{1t} - \frac{a(1+e^2)-r}{e^2 av}\Delta v_{1m} - \frac{r\sin u}{\sqrt{\mu p}\tan i}\Delta v_{1h} = -\frac{r\sin u}{\sqrt{\mu p}\tan i}\Delta v_1 \tag{11}$$

式中: θ 为真近点角; μ 为地球引力常数; p 为椭圆轨道半通径; a 为轨道半长轴; 下标 t 表示椭圆轨道切向, m 表示椭圆轨道法向, h 表示垂直于轨道平面方向。

2) 第 2 次变轨。在近地点减速将载人飞船运行轨道变成圆轨道。

近地点的速度:

$$\begin{cases} r_{\text{a}} = \dfrac{p}{1-e} \\ r_{\text{p}} = \dfrac{p}{1+e} \\ v_{\text{p}} = \sqrt{\dfrac{2\mu r_{\text{a}}}{r_{\text{p}}(r_{\text{a}}+r_{\text{p}})}} \end{cases} \tag{12}$$

圆轨道速度:

$$v_2 = \sqrt{\frac{\mu}{r_{\text{p}}}} \tag{13}$$

第 2 次变轨速度增量:

$$\Delta v_2 = v_2 - v_{\text{p}} \tag{14}$$

距离第 1 次变轨经过时间:

$$\Delta t_2 = \frac{\Delta M_2}{n_{\text{MS}}} \tag{15}$$

第 2 次变轨后载人飞船角速度发生改变:

$$\omega_{\text{MS}} = \sqrt{\frac{\mu}{r_{\text{p}}^3}} \tag{16}$$

式中: r_{a}、r_{p} 分别为椭圆轨道远地点半径、近地点半径; v_2 为第 2 次变轨后圆轨道速度; v_{p} 为椭圆轨道近地点速度。

3) 第 3 次变轨。采用霍曼转移轨道, 实现载人飞船和加注站的交会对接。第 3 次变轨是在过渡椭圆轨道的近地点加速, 使载人飞船进入过渡椭圆轨道。

霍曼变轨的提前角 Θ:

$$\Theta = \pi\left[1 - \left(\frac{r_1+r_2}{2r_2}\right)^{\frac{3}{2}}\right] \tag{17}$$

故等待时间为

$$\Delta t_3 = \frac{\Delta\theta - \Theta}{n_{\text{MS}} - n_{\text{CSS}}} \tag{18}$$

式中: $\Delta\theta$ 为初始相位角差; n_{CSS} 为空间中在当前轨道上的平均角速度。

在过渡椭圆近地点处的速度增量:

$$\Delta v_3 = \sqrt{\frac{2\mu r_2}{r_1(r_1+r_2)}} - \sqrt{\frac{\mu}{r_1}} \tag{19}$$

4) 第 4 次变轨。第 4 次变轨是在过渡椭圆轨道的远地点减速使载人飞船进入加注站的圆轨道, 与加注站相遇, 完成对接。

转移时间为过渡椭圆轨道的半周期:

$$\Delta t_4 = \pi\sqrt{\frac{(r_1+r_2)^3}{8\mu}} \tag{20}$$

在过渡椭圆远地点处的速度增量:

$$\Delta v_4 = \sqrt{\frac{\mu}{r_2}} - \sqrt{\frac{2\mu r_1}{r_2(r_1+r_2)}} \tag{21}$$

2.3　地月段建模

地月段基于月球影响球模型建立月球着陆器-空间转移运输系统的运动模型, 主要模型如下[16]。

1) 飞行时序。

① t_0 时刻: 月球着陆器-空间转移运输系统组合体从加注站出发, 进入地月转移轨道的入轨时刻。

② t_1 时刻: 月球着陆器-空间转移运输系统组合体转移至月球影响球的入口时刻。

③ t_2 时刻: 到达地月影响球区域后, 减速抵达近月轨道的时刻。

2) 质量模型。根据地月转移-月球着陆的飞行阶段, 飞行器在运动过程中的质量变化为

$$m_{sc}=\begin{cases}m_{sL}+m_{sT}+m_{p0} & t=t_0\\ m_{sL}+m_{sT}+m_{p0}(1-\eta)^{\frac{t-t_0}{86\,400}} & t_0<t<t_1\\ m_{sL}+m_{sT}+m_{p1} & t=t_1\\ m_{sL}+m_{sT}+m_{p1}(1-\eta)^{\frac{t-t_1}{86\,400}} & t_1<t<t_2\\ m_{sL}+m_{sT}+m_{p2} & t=t_2\\ m_{sL}+m_{pL}(1-\eta)^{\frac{t-t_0}{86\,400}} & t>t_2\end{cases} \tag{22}$$

式中：η为推进剂的日蒸发率；下标L、T分别代表月球着陆器、转移运输系统，s、p分别代表结构、推进剂。

式(22)还涉及时刻t_0、t_1、t_2时的推进剂质量m_{pi}，计算式为

$$\begin{cases}m_{p0}=m_{pL}+m_{pT}-\Delta m_0\\ m_{p1}=(m_{pL}+m_{pT}-\Delta m_0)(1-\eta)^{\frac{t_1-t_0}{86\,400}}-\Delta m_1\\ m_{p2}=\left[(m_{pL}+m_{pT}-\Delta m_0)(1-\eta)^{\frac{t_1-t_0}{86\,400}}-\Delta m_1\right]\cdot\\ \qquad(1-\eta)^{\frac{t_2-t_1}{86\,400}}-\Delta m_2\end{cases}$$

式中：Δm_i为t_i时刻用于加速(制动)消耗的推进剂质量，可根据式(1)计算。

该过程的约束条件如下：器月距离满足$r_{min}\leqslant\rho$，ρ为月球引力范围半径：

$$\rho=r_{moon}\left(\frac{m_{moon}}{m_{earth}}\right)^{\frac{2}{5}} \tag{23}$$

式中：r_{moon}为地月平均距离；m_{moon}为月球质量；m_{earth}为地球质量。

2.4 落月段建模

如图3所示，将月面下降软着陆分为环月降轨、动力下降、定点软着陆、缓速下降4个阶段。建模时忽略引力摄动，预定着陆点在着陆准备轨道平面内，且主减速之后月球着陆器基本位于落月点正上方[17]。

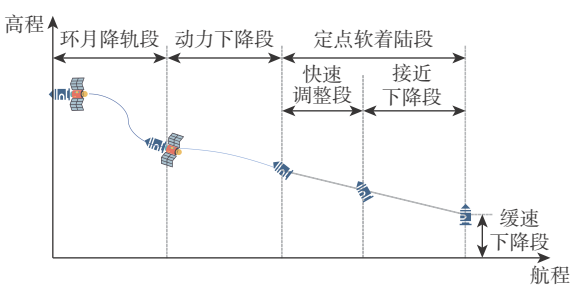


图3 月面下降着陆轨迹方案

Fig. 3 Lunar descent and landing trajectory

1）飞行时序。

① t_0时刻：月球着陆器-空间转移运输系统从近月轨道出发，点火减速进入着陆准备轨道的时刻。

② t_1时刻：月球抵达着陆准备轨道近月点，月球着陆器与空间转移运输系统分离，开启动力下降阶段的时刻。

2）质量模型。根据月球着陆的飞行阶段，飞行器在运动过程中的质量变化为

$$m_{sc}=\begin{cases}m_{sL}+m_{sT}+m_{p0} & t=t_0\\ m_{sL}+m_{sT}+m_{p0}(1-\eta)^{\frac{t-t_0}{86\,400}} & t_0<t<t_1\\ m_{sL}+m_{p1} & t=t_1\\ m_{sL}+m_{p1}(1-\eta)^{\frac{t-t_1}{86\,400}} & t>t_1\end{cases} \tag{24}$$

式(24)还涉及t_0、t_1时刻的推进剂质量m_{pi}，计算式为

$$\begin{cases}m_{p0}=m_{pL}+m_{pT}-\Delta m_0\\ m_{p1}=m_{pL}(1-\eta)^{\frac{t_1-t_0}{86\,400}}\end{cases} \tag{25}$$

3）运动阶段。

① 环月降轨段。动力下降段自距月面一定高度时开始，故环月降轨段的着陆准备轨道近月点高度为动力下降段的月面高度。在本阶段，由空间转移运输系统将月球着陆器自环月停泊圆轨道霍曼变轨送至着陆准备椭圆轨道，如图4所示。该椭圆准备轨道的近月点、远月点半径分别为r_A、r_B。

② 动力下降段。该阶段目的是减小月球着陆器速度，并使月球着陆器接近月面。考虑月球曲率并建立月球极坐标系，如图5所示。

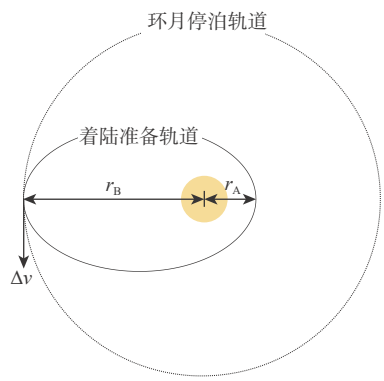


图4 环月降轨段变轨示意图

Fig. 4 Schematic of orbital maneuvering during lunar descent phase

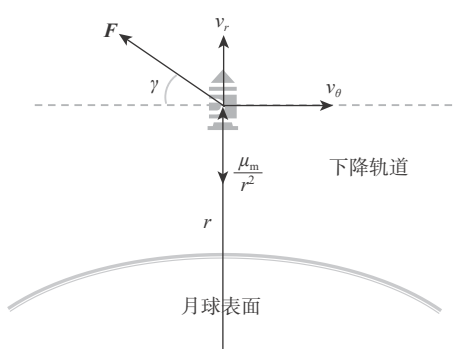


图5 动力下降段运动示意图

Fig. 5 Schematic of powered descent phase dynamics

上述过程对应的动力学方程为

$$\begin{cases}\dot{v}_r=-\dfrac{\mu_{\mathrm{m}}}{r^2}+\dfrac{v_\theta^2}{r}+\dfrac{F}{m}\sin\gamma\\ \dot{v}_\theta=-\dfrac{v_r v_\theta}{r}-\dfrac{F}{m}\cos\gamma\end{cases}\tag{26}$$

运动学方程为

$$\begin{cases}\dot{r}=v_r\\ \dot{\theta}=\dfrac{v_\theta}{r}\end{cases}\tag{27}$$

式中: γ 为推力的方向角, 即推力与当地水平面的夹角; μ_{m} 为月球引力常数; $\boldsymbol{F}$ 为航天器推力的大小; $\dot{\theta}$为在月心极坐标系下的极角; v_{r}、v_θ 分别为在月心极坐标系下, 月球着陆器的径向速度、横向速度。

在本阶段的最优推力控制过程中, 认为推力大小全程保持恒定, 推力方向与速度方向的夹角可以写成关于时间的多项式形式:

$$\gamma=\sum_{i=0}^{n}a_i t^i\tag{28}$$

式中: a_i 为 t^i 的系数。

以燃料消耗最小为优化目标, 即

$$\min\Delta m=\int_0^t\dot{m}\mathrm{d}t\tag{29}$$

③ 快速调整段。由于运动时间较短, 本阶段及之后所有阶段均不考虑月球曲率影响, 可以分解竖直方向与水平方向的运动, 建立平面直角坐标系, 如图 6 所示。

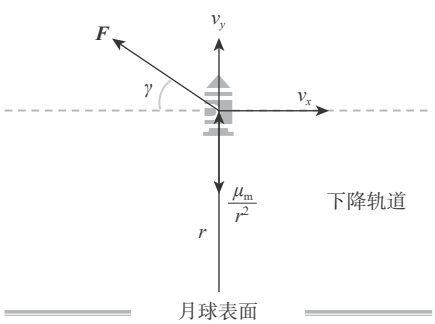


图 6　快速调整段运动示意图

Fig. 6　Schematic of attitude adjustment phase dynamics

上述过程对应的动力学方程为

$$\begin{cases}\dot{v}_x=-\dfrac{F}{m}\cos\gamma\\ \dot{v}_y=-\dfrac{\mu_{\mathrm{m}}}{r^2}+\dfrac{F}{m}\sin\gamma\end{cases}\tag{30}$$

运动学方程为

$$\begin{cases}\dot{x}=v_x\\ \dot{y}=v_y\end{cases}\tag{31}$$

④ 接近下降段。接近下降段的核心目标是避障, 在竖直方向主要有 2 种控制方案: ① 先保持匀速直线运动, 达到预定高度时施加额外速度脉冲达到悬停; ② 在竖直方向上先不施加主动力, 保持一段时间的匀加速运动再开机反推, 实现悬停。建立月球着陆器的受力分析如图 7 所示。

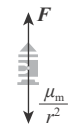


图 7　接近下降段受力分析

Fig. 7　Force analysis for approach and descent phase

上述过程对应的动力学方程为

$$\ddot{y}=-\frac{\mu_{\mathrm{m}}}{r^2}+\frac{F}{m}\tag{32}$$

⑤ 缓速下降段。该阶段与上一阶段的运动微分方程相同, 仅边界条件不同。

3　仿真分析

在第 2 节建立的模型基础上, 本节对载人登月任务的主要阶段进行轨道仿真分析。在天地往返(发射段)阶段, 火箭采用四氧化二氮/偏二甲肼推进剂; 空间转移(交会段、地月段、落月段)阶段, 轨道器使用低温推进剂(液氢液氧), 暂不考虑推进剂在空间环境中蒸发损失的影响。

3.1　任务窗口计算

以 LEO 上空间站某参考时刻的轨道参数为基准, 作为加注站轨道的初始条件, 考虑其在地球 J2 项摄动影响下的轨道运动, 计算其轨道参数。选择以地-月连线为远地点半径、加注站运行半径为近地点半径、与加注站运行平面共面的轨道, 对变轨时刻、轨道倾角进行遍历, 查询合适的窗口。

基于上述条件和约束, 自 2025-05-01T00:00:00 起进行 5 月全月的地月转移窗口遍历搜索, 得出结果如表 5 所示。

表 5　地月转移时间窗口

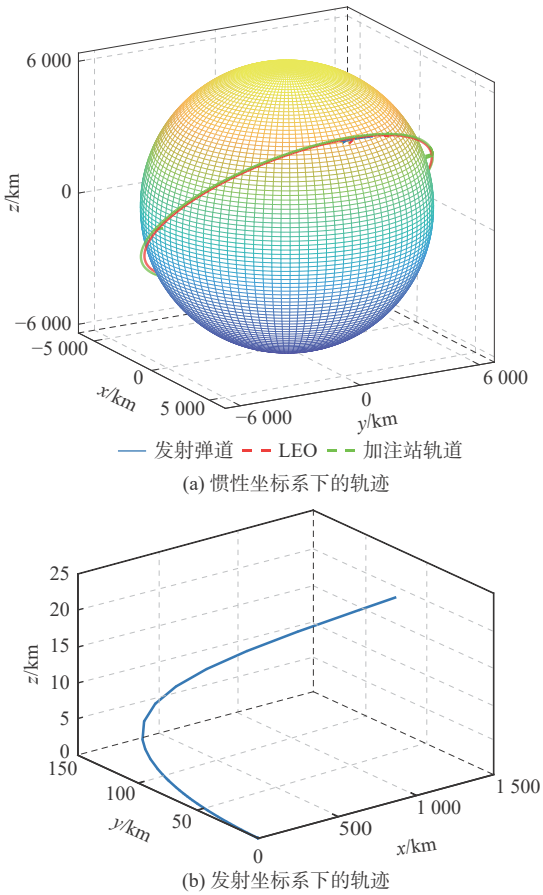
Table 5　Earth-Moon transfer window

变轨时刻	转移时间/s	捕获时相对月球的速度/(km·s^{-1})
2025-05-04T01:00:00	286 300	1.050
2025-05-04T18:00:00	297 700	1.006
2025-05-04T21:00:00	248 600	1.127
2025-05-05T14:00:00	264 300	1.056
2025-05-06T10:00:00	245 600	1.093
2025-05-07T06:00:00	248 600	0.949
2025-05-19T04:00:00	213 900	0.890
2025-05-31T15:00:00	271 600	1.054

3.2　发射段弹道仿真

惯性坐标系下火箭、载人飞船、加注站的飞行轨迹图 8 所示。火箭从地面发射, 到达 LEO 后与载

人飞船分离，随即载人飞船进入LEO。通过调整发射时间，可以保证LEO与加注站轨道的升交点赤经相同，进而保证2个轨道处于同一平面，为载人飞船与加注站的交会对接奠定了基础。最终得到载人飞船入轨参数如表6所示。



(a) 惯性坐标系下的轨迹

(b) 发射坐标系下的轨迹

图8 发射段飞行轨迹

Fig. 8 Ascent flight trajectory during launch phase

表6 载人飞船入轨参数

Table 6 Orbital insertion parameters for manned spacecraft

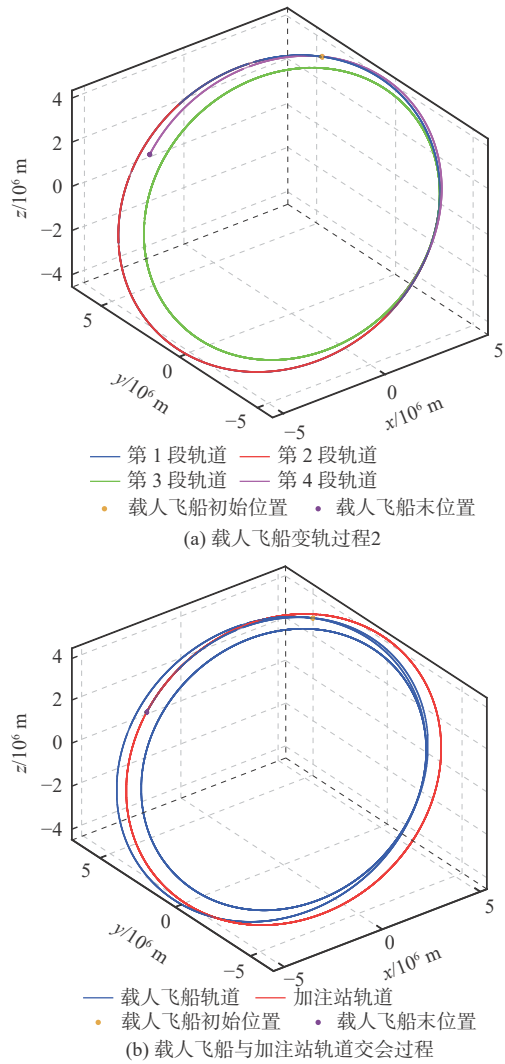
参数	数值
飞行时间/s	469.79
关机点速度/(km·s^{-1})	7.82
关机点高度/km	211.04

3.3 交会段轨道仿真

以2025-05-03T00:00:00作为当前时刻，经3.2节计算得选定的发射时刻为2025-05-03T10:30:39，入轨时刻为2025-05-03T10:31:29，上述过程所需飞行时间为470 s。

计算所得载人飞船交会对接时的4段轨道如图9(a)所示，载人飞船与加注站的交会轨道如图9(b)所示。

各段变轨的时刻和速度增量如表7、表8所示。速度增量为负代表减速。



(a) 载人飞船变轨过程2

(b) 载人飞船与加注站轨道交会过程

图9 交会段运动轨迹

Fig. 9 Rendezvous phase trajectory profile

表7 变轨时刻及速度增量

Table 7 Maneuver epoch and velocity increment

次数	变轨距离上次用时/s	速度增量/(m·s^{-1})
第1次	5 699.44	6.11
第2次	3 896.31	−316.82
第3次	18 576.35	199.51
第4次	2 567.61	194.58

表8 交会对接后航天器的轨道根数

Table 8 Post-docking orbital elements of spacecraft

轨道要素	取值
升交点赤经/(°)	292.41
轨道倾角/(°)	41.37
近地点幅角/(°)	107.50
半长轴/km	6 755.08
偏心率	0
平近点角/(°)	280.56

3.4　地月段轨道仿真

根据 3.1 节中遍历搜索, 获得捕获效果较好的转移轨道, 如图 10 所示。

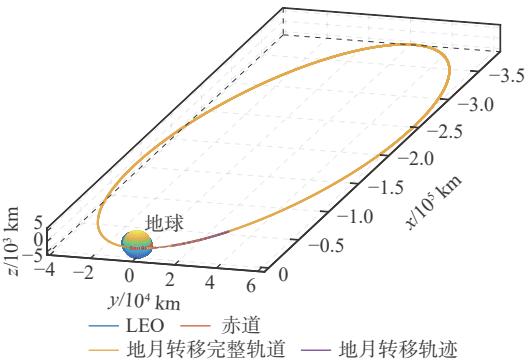


图 10　地月转移轨道

Fig. 10　Earth-Moon transfer trajectory

基于上述轨道, 绘制地月转移过程中的月球运动轨迹与航天器运动轨迹, 如图 11 所示。

由图 12 可见, 在地月转移的过程中, 航天器与月球的距离是单调递减的。

3.5　落月段轨道仿真

月球着陆器载人落月过程的参数选定如表 9 所示。

4 个下降阶段的仿真与优化计算具体如下。

1) 动力下降段。该下降阶段的切向速度较大, 故需要考虑月球形状的影响, 在极坐标系下进行计算。使用搜索算法, 遍历搜索推力; 在同一个推力下, 采用浮点数编码的遗传算法对推力方向角的多项式系数进行优化, 使得燃料消耗最少。该遗传算法的过程如图 13所示[18]。

选择交叉概率为 0.6, 变异概率为 0.1。根据计算结果绘制的动力下降段运动参数如图 14 所示。图 14 (a) 中, 本运动阶段月球着陆器相对月球做环绕运动较径向运动显著。推力方向变化曲线如图 14 (b) 所示, 本阶段推力与当地水平方向的夹角从 0°缓缓增加至 12°。速度变化曲线如图 14(c) 所示, 由月球着陆器指向月心的方向为径向负向, 该阶段主要消除了月球着陆器环绕月球的切向速度。

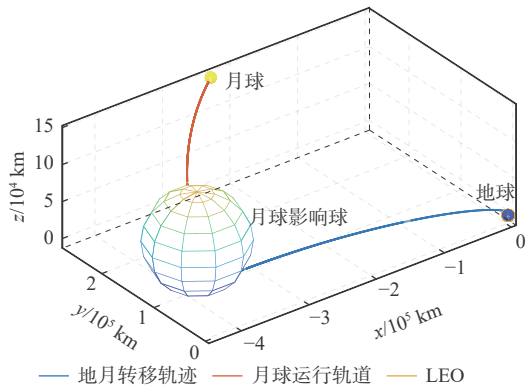


图 11　地月转移过程

Fig. 11　Earth-Moon transfer process

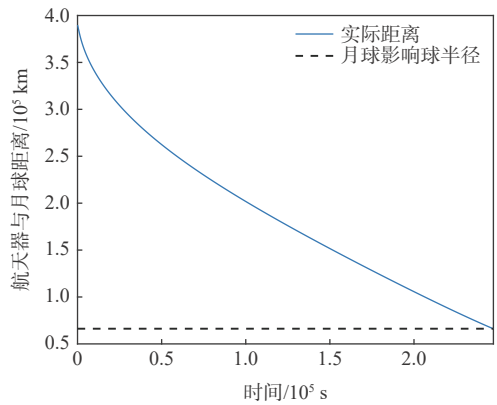


图 12　地月转移过程中航天器与月球距离

Fig. 12　Spacecraft-to-Moon distance during Earth-Moon transfer process

表 9　落月过程各阶段参数

Table 9　Phase-specific parameters during lunar landing sequence

落月段运动阶段	高度/km	推力大小/N	阶段耗时/s	推进剂消耗/kg	剩余质量/kg
环月降轨段	200×15(椭圆轨道)				
动力下降段	15.000～3.000	45 000	342.0	3 469.4	7 172.4
快速调整段	3.000～2.400	32 800	19.3	143.5	7 028.9
接近下降段	2.400～0.050	72 762	57.1	149.6	6 879.2
缓速下降段	0.050～0.002	28 156	10.0	25.3	6 853.9

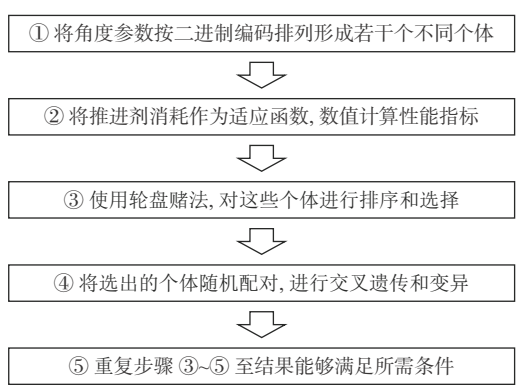


图 13　推力方向优化算法流程[18]

Fig. 13　Thrust vector optimization algorithm flowchart[18]

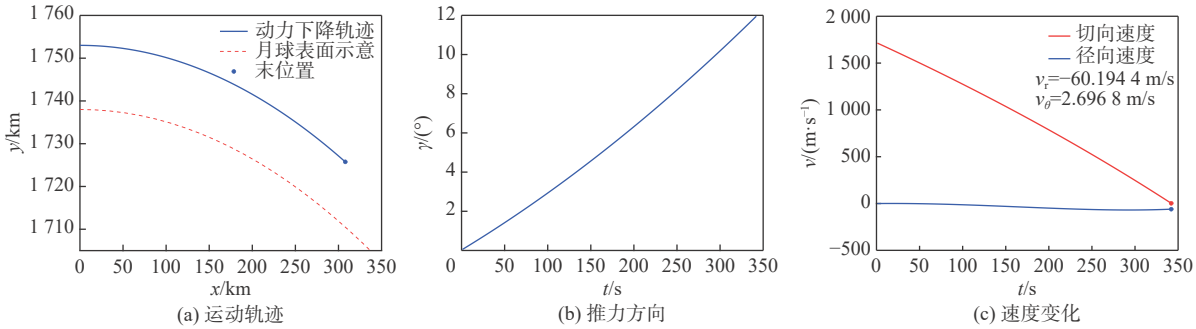


图 14 动力下降段运动参数曲线

Fig. 14 Powered descent phase kinematic profile

2) 快速调整段。该阶段及后续下降阶段的切向速度较小，故不考虑月球形状影响，在直角坐标系下计算。采用与动力下降段一致的优化方法。绘制快速调整段运动参数如图 15 所示。由图 15 (a) 运动轨迹示意图可见，该阶段月球着陆器接近于沿直线运动的状态。推力方向变化曲线如图 15 (b) 所示，该阶段推力方向几乎不变。速度变化曲线如图 15 (c) 所示，该运动阶段推力方向和速度方向夹角保持恒定，合速度的方向也几乎不变，月球着陆器在横向和纵向 2 个方向各自以恒定的加速度减速，水平方向减速至接近 0，竖直方向上月球着陆器仍以一定速度下降。

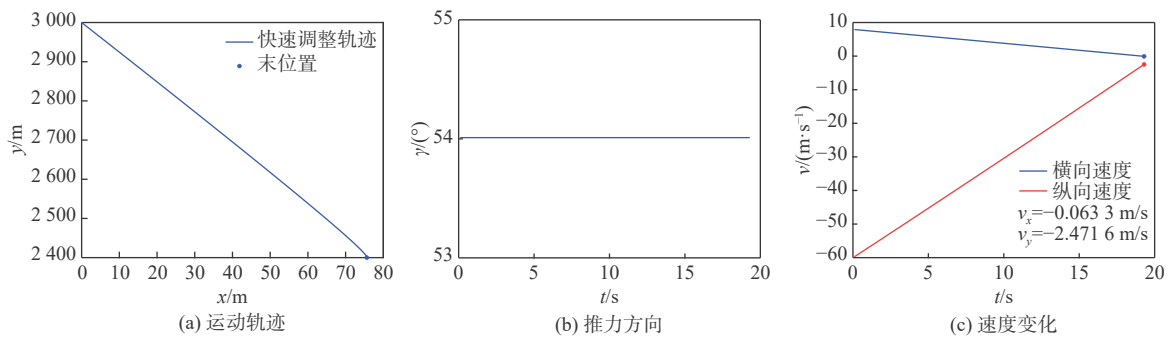


图 15 快速调整段运动参数曲线

Fig. 15 Attitude adjustment phase kinematic profile

3) 接近下降段。该阶段在竖直方向上采用的控制方案为：前期先关闭发动机，在月球引力下进行一段时间的匀加速运动，再启动发动机做一段时间的反推，达到预定高度时恰好实现悬停。

该过程匀加速下降时无需推力，开机反推时推力保持恒定，至距离月面 50 m 时速度为 0，故采用搜索算法遍历匀加速下降段的时间 t 与匀减速段的反推推力 $\boldsymbol{F}$，目标为推进剂消耗最少。算法步骤如图 16 所示[19]。

4) 缓速下降段。本阶段初末速度均接近 0，与上一阶段的微分方程与求解方法完全一致，仅有边界条件存在差异。由于存在缓速要求，将下降过程中最大速度也纳入考虑。至此，完成了 4 段落月过程的优化设计。

将上述工作所得的各阶段的仿真结果进行汇总，可以得到整个载人登月任务的流程，汇总结果如表 10 所示。

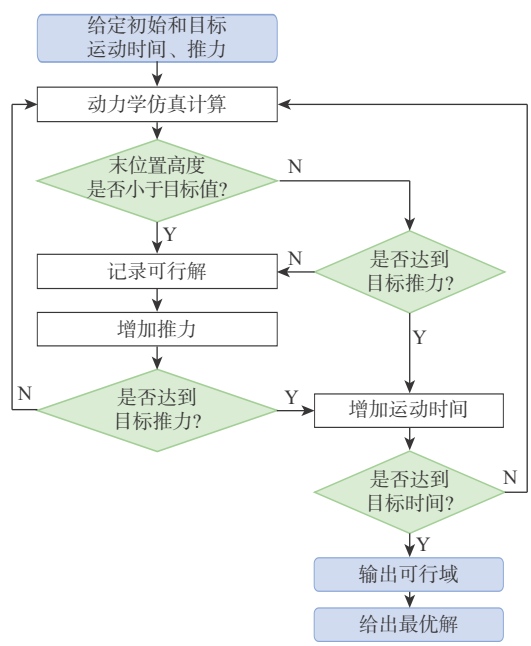


图 16 时间和推力优化算法流程[19]

Fig. 16 Time-thrust co-optimization algorithm flowchart[19]

表 10　任务节点汇总

Table 10　Mission phase node summary

序号	任务阶段	时间节点	推进剂消耗/kg	动力来源
1	地面发射	2025-05-03T10:30:39	430 372.4	运载火箭1
2	LEO→加注站	2025-05-03T10:38:29	1 262.2	载人飞船推进舱
3	与加注站交会对接	2025-05-03T19:10:49		
4	进入地月轨道	2025-05-04T01:00:00	58 309.4	空间转移运载器
5	制动进入LO	2025-05-07T08:31:40	17 611.5	空间转移运载器
6	进入落月轨道	2025-05-07T09:31:37	359.6	空间转移运载器
7	LO→月面	2025-05-07T09:38:45	3 762.6	月球着陆器

4　结　论

针对单次发射运载能力有限与大质量载人登月任务间的矛盾, 本文参考前人的“接力式”地月运输方案, 采用基于在轨加注的“两段式”载人登月任务模式, 完成了任务规模分析、动力学建模与全流程轨道仿真。主要结论如下:

1) 通过模块化质量估算方法对比表明, 基于在轨加注的载人登月任务方案系统质量规模, 显著低于基于近月轨道一次对接的方案, 可大幅减轻发射负担。

2) 建立了从发射至月面着陆的四阶段动力学模型, 完成全流程数值仿真, 明确了各阶段速度增量、推进剂消耗量等关键参数, 为任务工程设计提供了依据。

3) 针对月球软着陆过程, 对推力参数进行优化, 实现燃料最省控制, 为月球着陆器系统设计提供了参考。

开展新型地月空间转移运输系统的任务规划研究, 对于高效、经济、灵活的航天运输任务具有巨大前景。后续考虑进一步研究基于在轨加注的多目标协同任务规划方法, 在任务模型中耦合低温推进剂动力学, 并开展加注时序、多任务调度与在轨存储机制的联合优化。

参考文献（References）

[1] Smith D A. Space launch system (SLS) mission planner’s guide[R]. Washington, D. C.: National Aeronautics and Space Administration, 2018.

[2] Gillespie R W. Impulse propulsion gains resulting from" free" starship[J]. Scientific Reports, 2024, 14(1): 11804.

[3] Kutter B, O'Neil G, Pitchford B, et al. A practical, affordable cryogenic propellant depot based on ULA’s flight experience[C]//AIAA Space 2008 Conference & Exposition. Reston: AIAA, 2008: 7644.

[4] Hartwig J W. Liquid acquisition devices for advanced in-space cryogenic propulsion systems[M]. Cleveland: Case Western Reserve University, 2014.

[5] Fikes M J, Howell M J T, Henley M M. In-space cryogenic propellant depot (ISCPD) architecture definitions and systems studies [C]//57th International Astronautical Congress. Reston: AIAA, 2006.

[6] 陈小前, 袁建平, 姚雯. 航天器在轨服务技术 [M]. 北京: 中国宇航出版社, 2009.
Chen X Q, Yuan J P, Yao W. Spacecraft on-orbit servicing technology[M]. Beijing: China Astronautic Publishing House, 2009(in Chinese).

[7] Jones C A, Merrill R G, Mcvay E. Cis-lunar reusable in-space transportation architecture for the evolvable Mars campaign[C]//AIAA SPACE 2016. Reston: AIAA, 2016: 5493.

[8] National Academies of Sciences, Engineering, and Medicine. Space nuclear propulsion for human Mars exploration[M]. Washington, D. C.: The National Academies Press, 2021.

[9] 高永飞. 以地月空间站为空间港的载人月球探测轨道问题研究[D]. 长沙: 国防科技大学, 2019.
Gao Y F. Study on transfer trajectories of manned lunar exploration with LEO space station and lunar orbital station as spaceport[D]. Changsha: National University of Defense Technology, 2019(in Chinese).

[10] 王昊, 王浩苏, 陈士强, 等. 低温推进剂在轨加注关键技术研究进展与展望[J]. 导弹与航天运载技术 (中英文), 2024(5): 21-27.
Wang H, Wang H S, Chen S Q, et al. Review and development on key techniques of cryogenic propellants on-orbit refueling progress[J]. Missiles and Space Vehicles, 2024(5): 21-27(in Chinese).

[11] 鲁宇, 汪小卫, Saccoccia G, 等. 在轨加注站概念研究[J]. 导弹与航天运载技术, 2015(1): 1-7.
Lu Y, Wang X W, Saccoccia G, et al. Space transportation concept supported by on-orbit propellant depots[J]. Missiles and Space Vehicles, 2015(1): 1-7(in Chinese).

[12] 汪小卫, 鲁宇, 唐琼, 等. 基于在轨加注站的空间运输系统规模分析[J]. 载人航天, 2017, 23(3): 342-347.
Wang X W, Lu Y, Tang Q, et al. Scale analysis of space transportation system supported by space propellant depots[J]. Manned Spaceflight, 2017, 23(3): 342-347(in Chinese).

[13] 陈蓉, 汪小卫, 邓思超, 等. 大规模低成本地月空间航天运输体系研究[J]. 深空探测学报 (中英文), 2023, 10(5): 525-531.
Chen R, Wang X W, Deng S C, et al. Research on large-scale and low-cost cislunar aerospace transportation system[J]. Journal of Deep Space Exploration, 2023, 10(5): 525-531(in Chinese).

[14] 王晓慧, 毛李恒, 赵建文, 等. 一种载人登月质量规模估算的方法与系统: CN201510362335.2[P]. 2018-01-23.
Wang X H, Mao L H, Zhao J W, et al. Method and system for estimating mass scale of manned lunar landing: CN105005643B[P].

2018-01-23(in Chinese).

[15] 龙乐豪. 总体设计[M]. 北京: 宇航出版社, 1991: 42-99.
Long L H. System design[M]. Beijing: Astronautic Publishing House, 1991: 42-99(in Chinese).

[16] 阮开利. 小型月球探测器轨道设计[D]. 西安: 西北工业大学, 2001.
Ruan K L. Orbit design for small lunar probe[D]. Xi'an: Northwestern Polytechnical University, 2001(in Chinese).

[17] 杨孟飞, 张高, 张伍, 等. 月面无人自动采样返回任务技术设计与实现[J]. 中国科学: 技术科学, 2021, 51(7): 738-752.
Yang M F, Zhang G, Zhang W, et al. Technical design and implementation of Chang'e-5 robotic sample return mission on lunar surface[J]. Scientia Sinica Technologica, 2021, 51(7): 738-752(in Chinese).

[18] 王劼, 李俊峰, 崔乃刚, 等. 登月飞行器软着陆轨道的遗传算法优化[J]. 清华大学学报 (自然科学版), 2003, 43(8): 1056-1059.
Wang L, Li J F, Cui N G, et al. Genetic algorithm optimization of lunar probe soft-landing trajectories[J]. Journal of Tsinghua University (Science and Technology), 2003, 43(8): 1056-1059(in Chinese).

[19] 刘炎东, 黄震, 侯振东. 载人月面着陆器中途分离模式下降着陆策略优化研究[J]. 载人航天, 2024, 30(4): 501-509.
Liu Y D, Huang Z, Hou Z D. A strategy optimization study of descent and landing process for manned lunar lander in mid-course separation mode[J]. Manned Spaceflight, 2024, 30(4): 501-509(in Chinese).

Phased orbit design for manned lunar landing mission with on-orbit refueling

LI Qiling[1], SHI Peng[1,*], CHEN Yong[2]

(1. School of Astronautics, Beihang University, Beijing 102206, China;
2. Chengdu Fluid Dynamics Innovation Center, Chengdu 610031, China)

Abstract: To address the conflict between the limited payload capacity of a single launch vehicle and the substantial deployment demands of large-scale target payloads, this paper employs a two-phase manned lunar landing mission architecture, where on-orbit refueling connects Earth ascent and space transfer. Mission scale analysis and calculations were performed for manned lunar landing utilizing on-orbit refueling. Dynamic models for the four main phases—launch, rendezvous, Earth-Moon transfer, and lunar landing—were established, and overall mission design simulation analysis was completed. A mass estimation method was employed to compare the total mission mass and the scale of each stage between the traditional mode and the on-orbit refueling mode. In order to define phase-specific limitations and develop a mass model that takes into account in-orbit evaporation losses of cryogenic propellant, a mission model covering the four phases of launch, rendezvous and docking, Earth-Moon transfer, and lunar landing was built. Subsequently, mission windows were calculated, and full-process orbital simulations were conducted to analyze trajectory parameters and propellant consumption for each phase. Genetic algorithms were used to optimize parameters for the lunar landing phase. Mass estimation results show that, compared to the thousands of tonnes launch mass required by the traditional mode, the total launch mass for the on-orbit refueling scheme is below 500 tonnes, and the space transfer stage mass is on the order of tens of tonnes. The full-process simulation findings provide crucial references for engineering design, further elucidating the propellant consumption for each stage of the on-orbit refueling strategy.

Keywords: manned lunar landing; on-orbit refueling; propellant; space transportation; mass scale

Received: 2025-07-14; **Accepted**: 2025-10-17; **Published Online**: 2025-11-05 18:19

URL: link.cnki.net/urlid/11.2625.V.20251105.1107.003

Foundation item: General Program of National Natural Science Foundation of China (12672454)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2026.0062

多模态生理信号驾驶舒适性机器学习模型构建与比较

姜维胜[1]，周前祥[1,*]，柳忠起[1]，刘宇双[2]，王家涛[1]，靳萌萌[1]

(1. 北京航空航天大学 生物与医学工程学院，北京 100191；2. 中国康复研究中心（北京博爱医院）超声科，北京 100068)

摘 要：为实现驾驶舒适性的客观评估，基于多模态生理信号构建并系统比较了多种机器学习模型。采集64名受试者在模拟驾驶中的心电(ECG)、表面肌电(SEMG)、肌肉氧含量(SmO_2)及下肢动脉血流信号，以舒适性评分为标签，建立支持向量机(SVM)、随机森林(RF)、自适应提升(AdaBoost)等6类模型，采用主成分分析、合成少数类过采样技术(SMOTE)及网格搜索进行优化。结果表明：多信号融合能够显著提升分类性能；支持向量机在包含动脉血流参数的特征组合中表现最优，平均曲线下面积(AUC)达0.97，准确率为0.90；随机森林在多信号融合中平均AUC超过0.95。研究明确了面向便携监测的有效信号组合："表面肌电+心电"双信号组合(平均AUC为0.90)及"肌肉氧含量+心电+表面肌电"三信号组合(平均AUC为0.92)。研究成果可为驾驶舒适性的客观评估提供模型优选与信号组合的实证依据，也可为车载健康监测系统开发提供参考。

关 键 词：驾驶舒适性；生理信号；机器学习；分类模型；信号融合；健康监测

中图分类号：U461.4

文献标志码：A **文章编号**：1001-5965(2026)10-3647-11

随着出租车、长途客货运、工程车辆等长时间或长距离驾驶场景日益普遍，驾乘人员持续暴露于车辆振动与固定坐姿构成的复合负荷环境中[1]。这种"振动-静态负荷-时间"的多因素耦合作用，不仅易引发驾驶员肌肉疲劳、血液循环受阻及注意力下降，也会直接影响乘客的舒适体验与身心健康，已成为关乎驾乘安全、职业健康与出行品质的突出人因工程问题[2]。汽车座椅作为连接人体与车辆的关键界面，其动态特性与人机工学设计直接影响振动传递与体压分布，是改善驾乘体验的核心部件[3]。然而，当前座椅舒适性评价高度依赖主观问卷，存在反馈滞后、个体差异大、难以量化内在生理机制等局限，制约了座椅设计的精准优化与个性化适配。

为弥补主观评价的不足，基于客观信号的驾驶状态监测与舒适性评估研究日益受到关注。主要技术路径包括基于视觉的行为分析、基于车辆动态的间接推断及基于生理信号的直接监测[4]。基于视觉的行为分析方法易于实施，但受环境条件干扰和驾驶行为的影响。基于心电(electrocardiogram, ECG)、肌电、脑电等生理信号的直接监测方法，因其能直接反映自主神经与肌肉骨骼系统的状态变化，被认为在客观性和前瞻性上更具潜力[5]。在利用客观信号构建舒适性预测模型方面，机器学习技术已成为主流技术[6]。已有研究成功应用线性回归、支持向量机(support vector machine, SVM)、神经网络等模型，建立了座椅物理参数(如压力分布、振动加速度)与主观舒适度评分之间的非线性映射关系[7]。支持向量机与神经网络模型在相关预测任

收稿日期：2026-01-26；**录用日期**：2026-03-30；**网络出版时间**：2026-04-08 12:16
网络出版地址：link.cnki.net/urlid/11.2625.V.20260407.1830.001
基金项目：国家重点研发计划 (2023YFF0615904)
* **通信作者**. E-mail：zqxg@buaa.edu.cn

引用格式：*姜维胜，周前祥，柳忠起，等. 多模态生理信号驾驶舒适性机器学习模型构建与比较[J]. 北京航空航天大学学报，2026，52(10)：3647-3657. Jiang W S，Zhou Q X，Liu Z Q，et al. Construction and comparison of machine learning models for driving comfort assessment using multimodal physiological signals[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3647-3657 (in Chinese).*

务中均展现出优于传统线性模型的性能，证实了数据驱动方法在此领域具有适用性[8]。

尽管现有研究已取得显著进展，但面向实际驾驶场景的舒适性客观评估仍面临挑战。一方面，多数研究侧重于单一类型信号或少数物理参数，未能充分利用多模态生理信号所蕴含的互补信息，以全面捕捉复杂驾驶负荷下人体多系统的协同响应[9]；另一方面，在模型构建中，缺乏针对多模态生理特征的系统性筛选与不同机器学习算法的比较研究，对于何种特征组合与何种分类模型在驾驶舒适性判别任务中具有最优性能，尚缺乏明确的实证结论。

为此，本文聚焦于普通车辆驾驶场景，旨在系统构建并比较基于多模态生理信号的驾驶舒适性机器学习模型。核心研究内容为：①探究多模态生理信号特征融合对提升驾驶舒适性状态分类性能的作用；②系统比较不同机器学习算法在该任务中的性能，确定最优模型；③评估在面向便携应用的精简信号组合下，模型的性能表现，以寻求精度与实用性的平衡。本文期望为建立一套客观、精准且易于部署的驾驶舒适性评估方法提供实证依据与算法基础，并为未来智能座舱的实时健康监测系统集成提供技术参考。

1 实验设计与数据采集

1.1 受试者与实验平台

为实现对长时间驾驶负荷下人体生理响应的系统性评估，并确保所采集数据的可靠性，本文在受试者筛选、实验设计与平台搭建方面均严格遵循科学规范。

1.1.1 受试者与样本量

研究共招募64名健康青年男性受试者，平均年龄为(24.98±2.83)岁。所有受试者均持有C1或C2类驾驶证，具备常规驾驶经验(平均(3.39±0.92)年)，且无心血管、肌肉骨骼及代谢性系统病史，以排除潜在疾病对生理信号的干扰。实验前一周，受试者被要求停用任何药物或补充剂，并在实验前24 h内避免摄入咖啡因、酒精及剧烈运动，以确保生理状态稳定。本文实验方案已获得伦理委员会批准(批准号：BM20240195)，所有受试者在参加实验前均已签署知情同意书。

样本量通过G*Power 3.1软件，基于重复测量方差分析(repeated measures analysis of variance, ANOVA)进行估算[10]。以预实验数据为基础，设定效应量$f=0.25$(中等效应)，显著性水平$\alpha=0.05$，统计功效$1-\beta=0.8$，并考虑了实验设计的组数与重复测量次数。计算得出总样本量为64人，可满足统计推断的可靠性要求。

1.1.2 实验平台与驾驶任务设计

实验在高保真动态模拟驾驶平台上进行，该平台由六自由度Stewart运动平台、实车驾驶操控模块、视景仿真系统及中央控制系统构成[11]。该平台依据ISO 2631—1997等标准中定义的全身振动测试方法进行标定与验证，其产生的振动频谱和加速度幅值能够有效模拟真实车辆在典型路况下的动态输入[12]。

平台能够依据预设的车辆动力学模型与路面激励谱，精确复现车辆在不同路况(如平整铺装路面与颠簸非铺装路面)及车速(低速(50±5) km/h与高速(100±10)km/h)条件下产生的多轴向振动与姿态变化，其输出的物理激励与相关国际标准中定义的全身振动暴露条件具有高度一致性。

为系统探究驾驶环境因素的影响，实验采用道路平顺性(高平顺性：城市道路或高速公路；低平顺性：山区道路或坑洼土路)和车速(低速(50±5) km/h与高速(100±10) km/h)的两因素两水平完全随机设计。64名受试者被随机并均匀分配至4个实验条件组合(高平顺性-高速、高平顺性-低速、低平顺性-高速、低平顺性-低速)。每位受试者需完成连续2 h的模拟驾驶任务。要求受试者在晴朗日间驾驶自动挡的轻型越野车，将车速维持在规定范围内，进行长时间巡航驾驶。实验在温湿度受控的实验室环境中进行(温度(23±0.5) ℃；湿度(50±5)%)，以排除环境干扰。

1.1.3 实验流程与数据采集

实验准备阶段，让受试者充分了解实验目的与流程，知晓实验条件及量表填写方法，签署知情同意书和数据协议。对受试者测量信号点位的皮肤进行清理、标记，并粘贴传感器。实验开始后，受试者首先在静息仰卧位下完成5 min的基线生理数据采集。随后，受试者进入模拟驾驶器并调整至最舒适的驾驶坐姿。驾驶任务分为2个1 h阶段。在驾驶前(0 min)、驾驶中(60 min)及驾驶结束后(120 min) 3个时间节点，受试者在静态坐姿下同步完成多通道生理信号的采集与主观量表的填写。信号监测过程中平台暂停运行，避免平台振动对信号产生干扰，确保数据的准确性。实验在受试者完成驾驶任务或在心理、生理上无法继续实验时结束。实验结束后，保存实验数据、关闭实验设备并移除电极贴和传感器等。

1.2 多模态生理数据采集

为实现对长时间驾驶负荷下人体多系统生理响应的综合量化评估，本文同步采集了多模态生

理信号: 心电、表面肌电(surface electromyogram, sEMG)、肌肉氧含量(muscle oxygen saturation, SmO_2)、下肢动脉血流和浅层血流灌注参数。所有信号均严格按照标准化流程采集, 并通过软件时标对齐方案保障多模态数据在时间轴上的一致性, 为后续融合分析奠定基础。

1) 心电信号: 心电信号反映了心脏的电活动节律, 是评估自主神经系统功能状态的核心无创指标。本文采用美国 BIOPAC MP150 多通道生理记录系统及其 ECG100C 放大器模块进行采集, 采样频率设置为 1 000 Hz。电极粘贴严格遵循标准肢体三导联位置[13]。基于连续心搏间期序列, 可进一步计算心率变异性(heart rate variability, HRV)的各类时域、频域特征, 用以分析驾驶任务中心血管系统的调节与适应机制。

2) 表面肌电信号: 表面肌电信号直接表征了骨骼肌的神经肌肉活动, 是量化肌肉激活强度、时序模式及疲劳状态的重要工具[14]。研究采用 Delsys Trigno 无线表面肌电系统, 以 2 148 Hz 的采样频率同步采集受试者双侧下肢主要肌群(包括右侧股直肌、胫骨前肌、腓肠肌、比目鱼肌和双侧的股外侧肌、股内侧肌)的电信号。电极放置位置均依据相关肌群的解剖学定位及国际表面肌电学会的推荐标准确定[15]。

3) 肌肉氧含量参数: 局部肌肉组织的氧含量水平是评估其代谢状态与能量供需平衡的关键。本文使用中科搏锐 WORTH 型无线近红外光谱仪, 以 33 Hz 的采样频率连续监测双侧腓肠肌肌腹处的氧合血红蛋白、脱氧血红蛋白及总血红蛋白浓度的相对变化, 进而衍生出组织氧饱和度等关键参数, 用于表征肌肉的氧代谢动力学[16]。

4) 下肢动脉血流参数: 采用深圳华声 Clover 系列便携式彩色多普勒超声诊断仪进行测量, 配备成人腹部探头 C5-1(B)与下肢动脉专用探头 L15-4(A)。在实验设定的静息与驾驶后间歇期, 对受试者左侧髂外动脉及腘动脉进行测量, 获取收缩期峰值流速、舒张末期流速、血管内径等关键参数, 并计算阻力指数, 用以量化下肢动脉的血流力学状态。

5) 浅层血流灌注参数: 采用 PeriFlux 6000 激光多普勒血流仪, 以 100 Hz 的采样频率连续监测左脚背体表区域微血管的浅层血流灌注, 其读数以灌注单位表示。该参数可灵敏反映皮肤及皮下组织的微循环灌注变化[17]。

为确保所有采集自不同设备的生理信号在时间上的同步性, 本文对于所有可连续记录并具备软件标记功能的设备, 在每一个实验阶段(如静息开始、驾驶开始、驾驶结束)开始时, 操作员均在对应的设备控制软件中手动插入精确的时间标记。对于需要在特定时刻进行测量的非连续采集设备(如超声多普勒), 精确记录其测量操作的开始与结束时刻。在后续数据处理阶段, 所有原始数据均以时间标记为基础, 将各信号统一对齐至同一绝对时间轴, 从而实现多模态数据在时间维度上的精确同步。

1.3　主观舒适性评价标签获取

为构建数据驱动的驾驶舒适性分类模型, 需要获取可靠的主观状态标签作为监督学习的真值。本文采用主观评价量表对驾驶过程中的整体舒适性感受进行量化, 并以此为基础生成二分类标签。

舒适性评价采用 10 分制舒适度评分量表, 评分范围从 1 分(极不舒适)到 10 分(极其舒适)[18]。在每次生理信号采集的同步时间点(即驾驶前、驾驶中(60 min)、驾驶结束后(120 min)), 受试者在静态坐姿下, 依据自身即刻的整体感受, 独立完成一次评分。所有受试者在实验开始前均接受了统一的量表使用指导, 以确保其对评分标准理解一致。

为将连续的舒适度评分转化为适用于二分类模型的离散标签, 参照美国机动车工程师协会相关研究中使用 10 分制量表对可接受舒适性界限为 6 分的定义, 本文设定明确的阈值, 对舒适性状态进行明确划分。将评分大于 6 分的主观感受定义为“舒适”状态, 并赋以标签“1”; 将评分小于或等于 6 分的感受定义为“不舒适”状态, 并赋以标签“0”[19]。该阈值划分能够区分积极与消极的舒适性体验, 从而为后续机器学习模型提供清晰的分类目标。最终, 每位受试者在 3 个时间点获得的评分均被转换为对应的二分类标签, 与同期采集的多模态生理信号数据严格对应, 构成模型训练与验证的数据样本。

2　机器学习模型构建

2.1　信号预处理与特征提取

为确保从多模态生理信号中提取有效、可比的特征, 所有原始信号均需经过系统预处理, 以消除噪声、统一尺度。首先进行时间对齐: 对连续信号(心电、表面肌电、肌肉氧含量及浅层血流灌注), 依据设备软件的时间标记进行截取; 对离散的下肢动脉血流信号, 则根据手动时间标记进行获取。该策略实现了秒级时间窗内的数据同步, 能够满足评估持续性舒适状态的分析需求。数据处理均在 MATLAB R2021b 平台上完成, 具体流程如下:

1) 心电信号: 对心电信号进行预处理, 获取高保真的 R 波位置, 以进行精确的心率变异性分析[20]。

首先，采用0.5～40 Hz的带通滤波器，有效抑制基线漂移和高频肌电噪声，并保留P波、QRS波群与T波等关键成分；同时，应用50 Hz陷波滤波器消除工频干扰。然后，将所有心电信号统一截取为300 s（5 min）的标准时长，以保证时间维度的一致性。最后，对信号幅度进行Z-score标准化，消除个体间的固有幅值差异。基于预处理后的信号，提取时域特征（如RR间期标准差、相邻RR间期差值的均方根）和频域特征（如低频功率、高频功率及其比值）。

2）表面肌电信号：对表面肌电信号进行处理，以识别肌肉活动成分与各种噪声[21]。采用20～500 Hz的四阶巴特沃斯带通滤波器，在保留表面肌电信号主频成分的同时抑制低频运动伪迹和高频噪声，并辅以50 Hz陷波滤波器。处理后的信号同样被统一为300 s时长并进行Z-score标准化。从标准化信号中提取如下特征：①时域特征：计算积分肌电值（反映肌肉激活总量）和均方根值（反映平均收缩强度）；②频域特征：通过快速傅里叶变换计算功率谱，提取平均功率频率，其下降趋势是肌肉疲劳的典型频域表征。

3）肌肉氧含量信号：肌肉氧含量数据来源于近红外光谱设备直接输出的浓度时间序列。处理流程包括：依据实验时间标记进行数据分段；采用线性插值法填补可能的信号缺失；使用滑动平均滤波器平滑数据以降低随机波动。特征提取主要计算各关键生理参数在对应时间窗内的统计特征，包括氧饱和度、脱氧血红蛋白浓度、氧合血红蛋白浓度、总血红蛋白浓度及组织血红蛋白指数的均值与标准差。这些指标综合反映了局部肌肉的氧含量水平、代谢状态及血流容积变化[22]。

4）下肢血流参数和动脉血流参数：下肢血流参数包括来自超声多普勒的动脉循环系统指标和来自激光多普勒的微循环指标[23]。动脉血流参数由超声多普勒设备在特定时刻（静息期、驾驶间歇）测得，包括收缩期峰值流速、舒张末期流速、平均流速、血管直径及计算所得的阻力指数等。这些参数作为瞬时测量值直接用作特征，不进行额外的连续滤波处理。

5）浅层血流灌注参数：数据由设备直接输出为CSV格式，其中，已包含均值、标准差等基本统计量。处理时，将所有受试者的数据统一进行格式标准化与字段对齐处理。

由于从上述多源异构信号中提取的特征具有不同的量纲和数值范围（如时间、频率、浓度、流速、无量纲比值等），在输入机器学习模型前，对所有特征进行Z-score标准化处理。此步骤将每个特征转换为均值为0、标准差为1的分布。标准化处理至关重要，其消除了特征尺度差异对模型（特别是支持向量机、线性回归等基于距离或梯度优化的模型）训练的负面影响，确保所有特征在模型学习过程中具有可比的贡献度，从而提升模型的收敛速度、稳定性和最终分类性能。

最终，每位受试者在每个实验阶段（驾驶前、中、后）的所有生理信号，经过上述流程，生成一个统一的高维特征向量。该特征向量与对应时刻的主观舒适性二分类标签（舒适性为1/不舒适性为0）相关联，共同构成用于机器学习模型训练与验证的完备数据集。

2.2 特征优选与降维

在完成多模态生理信号的特征提取后，所获得的高维特征集内可能存在信息冗余与多重共线性问题。为提升后续机器学习模型的训练效率、稳定性与泛化能力，并增强特征集的判别力，本文采用了两阶段的特征优化策略：首先，针对信号内部进行特征筛选；随后，应用主成分分析对全局特征进行降维。

2.2.1 表面肌电信号内部特征优选

对表面肌电信号提取的积分肌电值、均方根值和平均功率频率3个特征进行相关性分析表明，均方根值与另外2个特征之间均具有高度相关性（相关系数均大于0.85），表明三者所承载的肌电活动信息存在显著重叠。在统计建模中，高度相关的特征会引入多重共线性，可能导致模型参数估计不稳定、方差增大及可解释性下降。为在保留核心信息的同时避免冗余，本文选择均方根值作为表面肌电信号的核心表征指标。其优势在于：计算过程包含平方、均值与开方运算，能有效平滑信号随机波动，抗噪性强；在物理意义上表征信号的有效振幅，能更直接地反映肌肉收缩的能量输出水平，且具有良好的时变响应特性，适用于动态功能评估[24]。

2.2.2 基于主成分分析的特征降维

本文从5种信号中提取了37个特征。鉴于样本量相对有限，为应对高维特征可能引发的过拟合风险并消除多重共线性，采用主成分分析进行降维。具体步骤包括：对所有特征进行Z-score标准化，随后计算标准化数据的协方差矩阵并进行特征分解，依据特征值贡献率选取主成分，最终，保留累计方差贡献率达85%的前15个主成分作为模型输入。此处理在有效保留原始信息的同时，显著降低了特征维度与模型复杂度，旨在从源头缓解小样本下的过拟合风险，从而提升模型训练的稳定性与泛化能力[25]。

2.3　机器学习模型

为系统比较不同算法范式在多模态生理信号分类任务中的性能, 本文选取了 6 种具有代表性的机器学习模型[26]。这些模型涵盖了从线性基准到复杂非线性集成的不同思想, 旨在全面评估其对高维、非线性生理数据的适应能力。所选模型包括: 线性回归、支持向量机、决策树、随机森林(ramdom forest, RF)、自适应提升(adaptive Boosting, AdaBoost)和梯度提升。

线性回归作为一种广义线性模型, 被选作性能比较的基线, 其模型简单、计算高效, 且系数具有良好可解释性[27]。然而, 其核心的线性假设可能难以描述舒适性感知与复杂生理信号之间的非线性关系。

支持向量机通过核函数将数据映射到高维空间以寻找最优分类超平面, 擅长处理非线性问题且泛化能力较强[28], 但其性能对特征质量、核函数及惩罚参数的选择极为敏感, 这使其成为检验不同特征组合判别力的理想模型。

决策树模型结构直观, 无需复杂数据预处理这一步骤, 但单棵决策树容易过拟合, 泛化能力弱, 常作为集成学习中弱学习器的基础[29]。

随机森林是一种基于 Bagging 的集成算法, 通过构建大量决策树来汇总其结果, 其引入的双重随机性能有效降低方差, 提升模型的稳定性和抗过拟合能力, 适合处理可能存在噪声和相关性的高维特征[30]。

自适应提升是一种基于 Boosting 的序列集成算法, 其核心机制是迭代调整样本权重, 使后续的弱学习器更关注先前被误分类的样本。凭借这一特性, 其特别适用于描述复杂的分类边界, 在数据类别不平衡时也具有潜在优势[31]。

梯度提升是另一种 Boosting 算法, 其以迭代方式拟合当前模型的残差(负梯度), 依托加法模型实现模型的不断优化。该算法具有较高的预测精度, 但对参数设置和过拟合控制要求也较高。

综上, 本节构建的 6 种模型在假设、复杂度与稳健性上形成梯度对比。通过系统比较其性能, 旨在为驾驶舒适性分类这一具体任务提供模型选型的实证依据, 并深入揭示不同算法在处理此类生物医学信号时的内在特性与局限。

2.4　模型训练与优化策略

为确保模型训练的稳健性与评估的可靠性, 本文在训练流程中系统集成了 3 项关键策略。

1) 采用基于受试者的十折交叉验证。将全部 64 名受试者随机分为 10 折, 在每一轮迭代中, 使用其中 9 折的全部时间点数据进行训练, 剩余 1 折受试者的全部数据进行测试。此过程重复 10 次, 确保每位受试者仅被测试一次。该策略能够严格保证在受试者层面, 训练集与测试集之间完全独立, 从而实现对模型泛化能力的稳健估计[32]。

2) 为缓解数据中存在的类别不平衡(舒适与不舒适之比为 2.69∶1), 在每一折交叉验证的训练集内部应用合成少数类过采样技术(synthetic minority oversampling technique, SMOTE)。该技术可在特征空间内对少数类样本进行插值合成新样本, 以平衡类别分布, 从而提升模型对少数类(不舒适状态)的识别能力, 避免模型因数据分布偏斜而产生分类偏差, 从而提升模型在不舒适状态识别上的灵敏度[33]。

3) 为充分挖掘各模型的潜力并规避过拟合或欠拟合风险, 本文采用网格搜索法进行超参数优化。针对每个模型设定关键超参数的搜索空间, 通过遍历网格内所有参数组合并在验证集上评估性能, 最终选取最优参数配置用于测试。具体包括支持向量机的正则化参数 C(0.1～100)与径向基核系数 γ(0.1～10)、随机森林的决策树数量(100～300)与最大深度(10～100)、自适应提升的弱学习器数量(100～300)与学习率(0.1～0.3)等。通过系统性的遍历, 能够找到在给定搜索范围内性能最佳的超参数, 从而充分挖掘每个模型的潜力。

2.5　性能评估指标

本文利用分类问题常用的混淆矩阵对不同机器学习模型在驾驶舒适性分类任务上的性能进行全面、客观的评估和比较, 包含对模型整体判别能力、对特定类别的识别能力以及在不同决策阈值下稳健性的多角度度量[34]。

本文采用的评估指标有准确率、精确度、召回率、F_1 分数和曲线下面积(area unden curve, AVC)。通过综合分析以上指标, 可从模型的整体正确率、对“不舒适”状态判别的精准性、全面性以及模型的综合判别力等多个维度, 对 6 种机器学习分类器的性能进行系统、深入的评估与比较, 从而为模型优选提供坚实的数据支持。

3　结果与分析

本文系统评估了 6 种机器学习模型在 5 种不同生理信号组合场景下的性能。通过统计学方法对模型精度及信号贡献度进行了量化分析, 主要结果如下。

3.1　机器学习模型整体性能对比

基于融合全部生理信号构建的 6 种机器学习模型, 其综合性能评估结果如图 1 所示。整体分析表

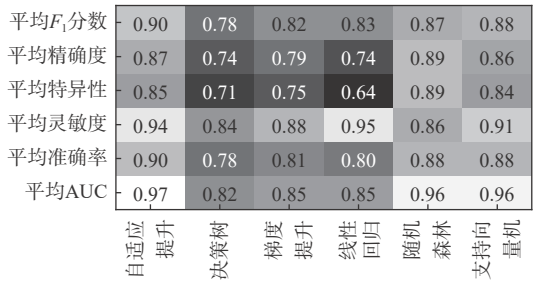


图1 全部生理信号机器学习模型评价指标热图

Fig. 1 Heatmap of evaluation metrics for all physiological signal machine learning models

明,在处理此类高维、非线性的多模态生理数据分类任务时,集成学习方法展现出显著且稳定的优势。

具体而言,自适应提升算法在绝大多数评价指标上取得了最优或接近最优的表现,其平均AUC值达到0.97,平均准确率与平均F_1分数均为0.90。该算法的优异性能主要归功于其自适应权重调整机制,能有效应对类别不平衡,并精细捕捉复杂的非线性映射关系。随机森林同样表现卓越,其平均AUC值为0.96,平均特异性达0.89,在所有模型中最高,表明其在判断“舒适”状态时误报率低、可靠性强。支持向量机取得了具有竞争力的结果(平均AUC值为0.96),其平均灵敏度高达0.91,在所有模型中较优,意味着其在识别“不舒适”状态时漏报率低,适用于安全预警场景。

相比之下,线性回归模型呈现出高灵敏度(0.95)但低特异性(0.64)的明显类别判别偏差,揭示了其线性假设难以有效解决此类复杂非线性生理关联的根本局限。决策树作为单一模型表现最不理想(平均AUC值为0.82),直观暴露了其易过拟合、泛化能力弱的缺点。梯度提升算法在实验参数设置下的表现相对较弱(平均AUC值为0.85)。

为评估不同模型间性能差异的统计显著性,对十折交叉验证得到的平均AUC值进行了Friedman检验。结果如图2所示,不同模型的平均AUC值

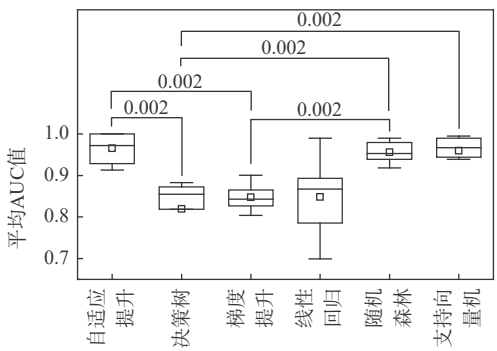


图2 全部生理信号机器学习模型平均AUC值箱线图

Fig. 2 Box plot of average AUC values for all physiological signal machine learning models

存在显著差异($\chi^2(5)=32.45$, $P<0.001$)。为进一步明确具体差异,采用Bonferroni校正的Wilcoxon符号秩检验进行两两比较(校正后显著性水平$\alpha=0.003\,3$)。在全部15对比较中,有5对模型之间的差异达到了统计学显著水平($P<0.003\,3$),具体包括:自适应提升与决策树、自适应提升与梯度提升、决策树与随机森林、决策树与支持向量机及梯度提升与随机森林。性能同属最优梯队模型(如自适应提升、随机森林与支持向量机)之间的两两比较P值均大于校正后α水平,表明其之间的性能差异在统计上不显著。这一结果量化证实了集成学习模型(自适应提升、随机森林)及支持向量机的性能显著优于决策树等单一弱学习器,在达到高性能水平后,不同集成策略或算法之间可能不存在统计学上的本质性能差异。

综上,从实际车辆人因工程系统应用的角度,可根据不同需求进行模型选型:自适应提升适用于追求高精度、稳定综合评估的场景;随机森林更适用于对“舒适”状态判别的特异性(即降低误报警)有更高要求的场景;而支持向量机则更适合部署于以“不舒适”状态早期、灵敏检测为核心目标的安全预警模块中。统计检验表明,这3类模型构成了本次任务中高性能且彼此等效的优选集合。

3.2 生理信号的判别效能分析

为深入探究不同生理系统对驾驶舒适性状态的独立表征能力,评估各生理信号作为单一信息来源的判别潜力,本文系统分析了5种单一生理信号在多种机器学习模型上的分类性能,结果如图3所示。分析表明,不同生理信号所蕴含的舒适性判别信息存在显著差异,且同一信号在不同算法模型上的性能表现也各具特点,这为理解舒适性的多生理维度基础及后续的特征筛选提供了重要依据。

在5种信号中,下肢动脉血流信号展现出最强的独立判别能力,其最优模型(随机森林)的平均AUC达到0.93,平均准确率与平均F_1分数分别为0.83和0.84,综合性能显著领先。这从数据层面直接印证了下肢血流动力学状态是评估长时间坐姿负荷下舒适性的一个核心生理维度。该信号在支持向量机模型上的平均特异性与平均精确度均达1.00,但平均灵敏度相对较低(0.61),表明其判别策略极为保守;而自适应提升算法则取得了更高的平均灵敏度(0.89),在识别不舒适状态上更为积极。这种差异体现了不同算法对同一生理特征的模式提取各有侧重。

表面肌电信号的独立判别效能次之,其最优模型(支持向量机)的平均AUC为0.89,平均精确度

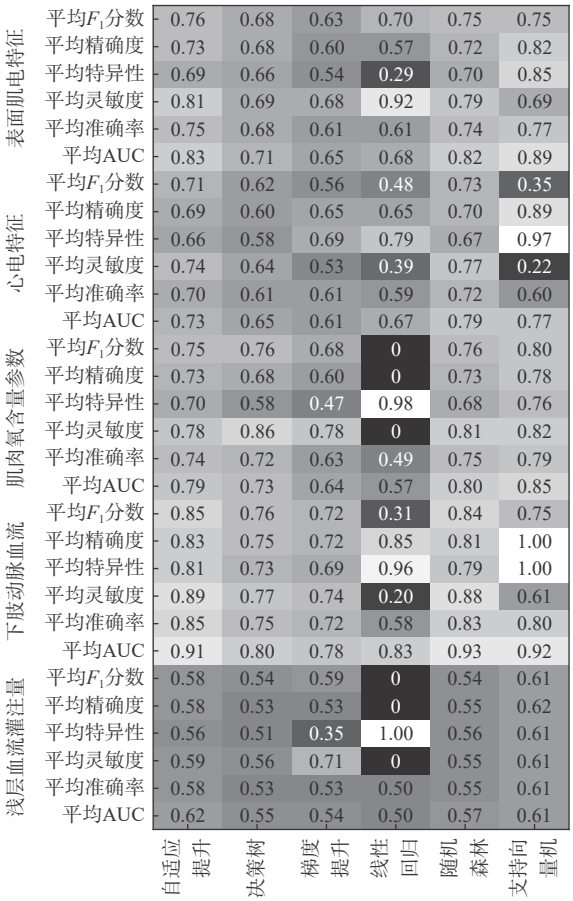


图 3　单一生理信号机器学习模型评价指标热图

Fig. 3　Heatmap of evaluation metrics for single physiological signal machine learning models

为 0.82。该信号主要反映骨骼肌的神经肌肉活动与疲劳状态。自适应提升算法在该信号上获得了最高的平均灵敏度(0.81), 再次显示了其在捕捉正类样本上的优势。肌肉氧含量信号的判别力处于中等水平(支持向量机的平均 AUC 为 0.85 最优), 随机森林与自适应提升性能接近, 表明该信号所反映的局部肌肉氧代谢与微循环信息对舒适性有一定贡献, 但作为单一信息来源强度有限。

心电信号(随机森林的平均 AUC 为 0.79 最优)的独立判别能力相对较弱, 可能是因为心电所表征的全身性自主神经调节是一个更为宏观、系统的指标, 其对座椅界面局部负荷的特异性响应可能不如动脉血流等局部信号直接和敏感。

浅层血流灌注量信号的独立判别效能最弱, 各模型平均 AUC 普遍偏低(最优自适应提升仅为 0.62)。梯度提升模型虽获得了 0.71 的平均灵敏度, 但平均特异性低至 0.35, 已失去了判别意义。这可能由于该信号表征的是皮肤及浅表组织的微循环, 其波动易受环境温度、局部压力点变化等多种混杂因素影响, 信噪比较低, 导致其与整体舒适性状态间的稳定关联较弱。

因此, 从独立判别效能的视角进行评估, 下肢动脉血流信号是驾驶舒适性最敏感、最具鉴别力的单一生理指标, 表面肌电信号次之, 二者构成了舒适性客观评估的核心信息源。心电信号与肌肉氧含量信号可作为有效的辅助信息来源, 而浅层血流灌注信号的独立应用价值有限。此结果量化了不同生理系统对驾驶不舒适的贡献度差异, 明确了特征优选中的优先顺序。

3.3　信号融合的增益效应与核心信号识别

为探究多模态生理信号融合对驾驶舒适性评估的协同增强作用, 并识别其中的关键判别特征, 本文系统比较了从双信号到五信号多种组合下的模型性能演变, 结果如表 1 所示。分析表明, 多源信号融合能产生显著的性能增益, 且下肢动脉血流信号在各类组合中均被证实为不可或缺的核心判别特征。

信号融合的增益轨迹清晰可见: 相较于单信号输入, 双信号组合普遍带来性能的第一次跃升。其中, “下肢动脉血流+肌肉氧含量”的组合表现突出, 随机森林模型在该组合上平均 AUC 达 0.94, 平均准确率达 0.89, 平均 F_1 分数达 0.90, 显示出 2 类信号在表征循环与代谢状态上的良好互补性。同样, “下肢动脉血流+表面肌电”的组合在自适应提升模型上也取得了平均 AUC 为 0.94 的优异表现,

表 1　多生理信号组合机器学习模型评价指标

Table 1　Evaluation metrics for multi-physiological signal combination machine learning models

信号组合	机器学习模型	平均AUC	平均F_1分数	平均精确度	平均特异性	平均灵敏度	平均准确率
下肢动脉血流+肌肉氧含量	随机森林	0.94	0.90	0.88	0.87	0.91	0.89
下肢动脉血流+表面肌电	自适应提升	0.94	0.87	0.83	0.80	0.93	0.86
表面肌电+心电	随机森林	0.90	0.81	0.76	0.71	0.88	0.80
下肢动脉血流+肌肉氧含量+心电	随机森林	0.95	0.88	0.86	0.84	0.90	0.87
肌肉氧含量+心电+表面肌电	支持向量机	0.92	0.64	1.00	1.00	0.48	0.74
肌肉氧含量+心电+表面肌电	随机森林	0.92	0.81	0.80	0.78	0.83	0.80
下肢动脉血流+肌肉氧含量+心电+表面肌电	支持向量机	0.97	0.90	0.88	0.87	0.93	0.90
下肢动脉血流+肌肉氧含量+心电+表面肌电	随机森林	0.97	0.89	0.88	0.87	0.90	0.89
下肢动脉血流+肌肉氧含量+心电+表面肌电	自适应提升	0.96	0.89	0.85	0.84	0.94	0.89

印证了循环与肌骨系统负荷关联的有效性。当信号融合增至三路时，模型性能得到进一步巩固和提升。例如，“下肢动脉血流+肌肉氧含量+心电”的三信号组合在随机森林模型上实现了平均AUC为0.95，较优的双信号组合又有提高，说明心电信号所贡献的自主神经调节信息提供了额外的判别维度。

在“肌肉氧含量+心电+表面肌电”组合中，支持向量机表现出平均AUC (0.92)与平均灵敏度(0.48)、平均准确率(0.74)的显著分离。高AUC值表明，该模型能将大多数“不舒适”样本的预测概率排在“舒适”样本之前。但是，极低的灵敏度说明，当应用固定阈值进行分类时，模型倾向于做出保守预测，从而漏检。这可能是由于缺少动脉血流这一强判别特征后，特征空间的分离度下降，支持向量机为最大化间隔而找到了一个严重偏向某一方的分类超平面。相比之下，随机森林在同一特征组合上取得了更均衡的性能(平均AUC为0.92，平均F_1分数为0.81)，体现了集成方法在不均衡或重叠特征空间中的稳健优势。此现象进一步说明，在评估模型时需综合考量多项指标，单一的高AUC值并不等同于可直接应用的良好分类器。

最终，在四信号全融合状态下，模型性能达到峰值。支持向量机在“下肢动脉血流+肌肉氧含量+心电+表面肌电”组合上取得了本次研究的最优性能，平均AUC高达0.97，平均准确率与平均F_1分数均为0.90。随机森林与自适应提升在该组合下同样表现卓越，平均AUC分别达到0.97和0.96。这标志着融合“循环-代谢-神经-肌骨”多系统信息的完备特征集，能够最全面、最稳定地表述驾驶负荷下的复杂生理状态。核心信号的不可替代性在对比中得到反复验证。在所有高性能组合中，下肢动脉血流信号均为共有元素。相反，未包含下肢动脉血流的信号组合，其模型性能仅停留在中等水平，提升有限。这从反面对比中确证了下肢动脉血流动力学参数是驾驶舒适性判别体系中具有基石性作用的特征。而浅层血流灌注量信号，尽管在融合中能提供微弱的信息补充，但其单独或作为核心的判别贡献始终有限。

3.4 面向应用场景的方案选择

基于前述系统的性能评估与对比，本文可进一步为不同工程应用场景提供明确、量化的模型与特征优化配置建议。对于追求极限判别精度、主要用于深入机理研究或作为标准验证的高精度实验室评估场景，推荐采用包含核心判别信息的完备特征集，即“下肢动脉血流+肌肉氧含量+心电+表面肌电”四信号组合，并配合高性能分类算法。其中，支持向量机能够利用该特征集达到最高的分类精度(平均AUC为0.97)；若对模型的稳定性、抗过拟合能力及特征重要性分析有额外需求，随机森林是表现同样卓越(平均AUC为0.97)的理想备选。此方案能实现对驾驶舒适性状态最为精细化、高可信度的判别。

然而，考虑到在实际车辆环境中部署超声多普勒等设备监测下肢动脉血流的操作复杂性、成本与用户侵扰性，面向驾驶员实时健康管理或疲劳预警的便携式可穿戴监测系统，需寻求在性能与实用性之间取得平衡的精简方案。为此，本文验证了2种可行的技术路径：①双信号精简路径，即采用“表面肌电+心电”这一兼具良好判别力与易测性的组合，在随机森林模型上可保持平均AUC为0.90、平均F_1分数为0.81的可靠性能，且传感器易于集成到穿戴设备中；②三信号优化路径，采用“肌肉氧含量+心电+表面肌电”组合，该组合在支持向量机或随机森林等模型上可实现平均AUC为0.92的性能，同时能够避免复杂的血管参数测量。这2种方案为开发高集成度的实用化车载监测系统提供了技术依据。

4 讨 论

4.1 多模态生理信号融合的机制阐释

本文证实，多模态生理信号融合是实现高精度驾驶舒适性评估的有效途径。下肢动脉血流信号被验证为核心判别特征，这一发现与生物力学和生理学经典理论高度契合。早期研究已明确指出，长时间坐姿会因血管的几何弯曲与受压，导致下肢动脉血流动力学环境恶化，表现为壁面剪应力下降、血流速度减缓，影响血管内皮功能，并易引发局部不适乃至血管疾病风险升高[35-37]。本文证实了下肢动脉血流信号具有最强的独立判别力，从而将主观的“不舒适”与具体的生理指标建立关联，为循环系统引起主观不适提供了实证。

同时，本文融合表面肌电、心电信号的做法，顺应了舒适性评估从单一维度走向系统整合的思路。早期的相关研究，如Andersson等[38]关于坐姿下腰部负荷与肌电活动的工作，以及Lin等[39]对脑电、心电与驾驶表现关系的研究，均已揭示单一生理信号在表征特定负荷方面的价值与局限。De Looze等[40]的综述指出，体压分布、肌电等多源信号与主观感受间存在复杂关联，需综合考量。本文在此基础上，通过系统性融合循环系统和肌骨系统相关信号，构建了跨系统的生理响应模式。结果表明，多模态融合策略能够捕捉单一信号无法揭示的

协同与代偿机制, 如“下肢动脉血流+表面肌电”组合同时反映了循环障碍与肌肉的代偿性紧张, 从而实现了对驾驶舒适性状态更全面的表述。这验证了集成多维度生理信息是深入理解并精准评估复杂人-机-环境系统中人体状态变化的必然方向。

4.2　模型性能差异的算法根源分析

不同模型的性能差异源于其算法原理。随机森林通过构建大量差异化的决策树并集成其结果, 有效降低了方差, 赋予了模型出色的稳定性和抗过拟合能力, 使其在多信号融合时表现稳健。支持向量机则通过核函数将数据映射到高维空间寻找最优分类面, 当输入如下肢动脉血流等判别力强的特征时, 能实现极高的分类精度, 这也导致其性能对特征选择异常敏感。自适应提升通过聚焦于被前序模型分错的样本, 逐步提升整体性能, 尤其擅长处理复杂的决策边界。而线性回归的线性假设与复杂生理现象间的非线性本质严重不符, 导致其性能受限。

4.3　研究的技术贡献与工程启示

本文通过系统的实证分析, 在驾驶舒适性客观评估领域取得了以下 3 方面启示:

1) 确立了下肢动脉血流信号的核心判别地位, 深化了生理机制认知。研究通过数据驱动对比证实, 下肢动脉血流动力学是表征驾驶舒适性最敏感的单一生理指标。

2) 针对高精度评估与便携监测的不同需求, 本文提供了经过性能量化的具体方案, 为工程开发中的传感器配置与算法选型提供了清晰的决策依据。

3) 实证了线性模型的根本局限, 强化了非线性方法的必要性。 线性回归模型在本文复杂数据上表现出的严重性能失衡, 实证了驾驶舒适性与其生理信号之间存在高度非线性的映射关系。在类似涉及多系统耦合的生物医学信号分析中, 应优先采用非线性模型, 以避免因方法选择不当导致结论失效。

本文在识别了核心生理特征的同时, 得出了能够落地的具体技术方案, 为构建实用性的驾驶舒适性智能监测系统奠定了基础。

4.4　与现有研究的对比

本文构建的模型, 特别是基于下肢动脉血流信号与支持向量机或随机森林的组合, 其性能指标(如平均 AUC > 0.95, 平均 F_1 分数 > 0.85)超过了文献中广泛报告的线性回归模型($R^2 = 0.76$)[6]; 同时超过了部分支持向量机模型(平均准确率约为 0.75)[41]; 与当前先进的神经网络模型($R^2 = 0.92$)达到了可比水平[8]。本文主要创新在于通过模型和信号组合的网格化比较, 量化揭示了下肢动脉血流信号在驾驶舒适性判别中的核心地位, 明确了集成学习方法在处理此类多模态生理数据上的稳定优势。同时, 提出了便携式信号组合方案, 为研究成果向可穿戴设备转化提供了关键的指导。

4.5　局限性与未来工作

本文的局限性一方面是数据源于模拟驾驶环境, 另一方面是受试者均为健康青年男性。这虽然有利于控制混杂因素, 但其限制了结论的外部效度与泛化能力。未来工作可围绕以下方向实施: 开展实车实验进行模型验证; 纳入更多样化的受试者样本, 增加受试者数量, 以降低模型的过拟合风险, 建立更具普适性和稳健的评估模型; 通过更精细的舒适性评分分段或采用回归模型来避免对单一阈值的依赖; 以及探索更精确的多源数据的时间窗对齐(如硬件级同步)和更优的特征提取方法, 尝试融合更多维度的生理或行为信号, 以构建更优的舒适性状态识别算法; 推动监测模型产品化, 向智能座舱健康监测系统迈进。

5　结　论

本文通过构建与系统比较基于多模态生理信号的机器学习模型, 得出以下结论:

1) 模型优选方面: 集成学习方法(随机森林、自适应提升)在处理多模态生理信号时综合性能更优、更稳定; 支持向量机能达到最高分类精度, 但对特征输入敏感。

2) 信号贡献方面: 下肢动脉血流信号是驾驶舒适性判别的核心生理指标, 其与表面肌电、心电信号的融合能显著提升模型性能, 印证了多系统生理响应耦合机制。

3) 应用方案方面: 研究验证了便携式精简信号组合(如“表面肌电+心电+肌肉氧含量”)的可行性, 为开发兼顾精度与实用性的车载实时监测系统提供了明确的特征选择与模型配置依据。

本文通过详实的实验数据与系统分析, 为驾驶舒适性的客观智能评估提供了从理论机制、算法选型到工程化路径的完整方案支持。

参考文献(References)

[1] Casey G J, Miles-Johnson T, Stevens G J. Lack of consistency in truck driver fatigue management: Australian law enforcement officer's and transport regulation agent's perceptions of enforcement[J]. Transportation Research Part F: Traffic Psychology and Behaviour, 2025, 112: 322-341.

[2] Nader M. Experimental investigations of special vehicles vibration with regard to their influence on human organism[J]. Engineering Transactions, 2000, 48(2): 119-134.

[3] Voicu D, Stoica R M, Vilău R, et al. Frequency analysis of vibrations in terms of human exposure while driving military armoured personnel carriers and logistic transportation vehicles[J]. Electronics, 2023, 12(14): 3152.

[4] Sikander G, Anwar S. Driver fatigue detection systems: a review [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(6): 2339-2352.

[5] Sanjaya K H, Lee S, Katsuura T. Review on the application of physiological and biomechanical measurement methods in driving fatigue detection[J]. Journal of Mechatronics, Electrical Power, and Vehicular Technology, 2016, 7(1): 35-48.

[6] Kolich M, Seal N, Taboun S. Automobile seat comfort prediction: statistical model vs. artificial neural network[J]. Applied Ergonomics, 2004, 35(3): 275-284.

[7] Huang S R, Zhang Z F, Xu Z M, et al. Modeling of human model for static pressure distribution prediction[J]. International Journal of Industrial Ergonomics, 2015, 50: 186-195.

[8] 孙守迁, 吴群, 吴剑锋, 等. 一种基于支持向量回归的驾驶座椅舒适度评价方法[J]. 中国机械工程, 2008, 19(11): 1326-1330.
Sun S Q, Wu Q, Wu J F, et al. An evolution method of driving seat comfort based on support vector regression[J]. China Mechanical Engineering, 2008, 19(11): 1326-1330(in Chinese).

[9] Elgendi M, Menon C. Machine learning ranks ECG as an optimal wearable biosignal for assessing driving stress[J]. IEEE Access, 2020, 8: 34362-34374.

[10] Yenipinar A, Koc S, Canga D, et al. Determining sample size in logistic regression with G-Power[J]. Black Sea Journal of Engineering and Science, 2019, 2(1): 16-22.

[11] Davliakos I, Papadopoulos E. Model-based control of a 6-dof electrohydraulic Stewart-Gough platform[J]. Mechanism and Machine Theory, 2008, 43(11): 1385-1400.

[12] International Organization for Standardization. Mechanical vibration and shock-evaluation of human exposure to whole-body vibration-part 1: general requirements: ISO 2631—1997[S]. Geneva: International Organization for Standardization, 1997: 1-17.

[13] Jiang W S, Yin C, Zhou Q X, et al. Evaluation model of running fatigue of young students based on characteristic parameters of ECG signal[C]//HCI International 2023 Posters. Berlin: Springer, 2023: 336-342.

[14] Matsuda T, Miyamori T, Fujino Y, et al. Reliability and validity of muscle activity analysis using wearable electromyographs[J]. Journal of Physical Therapy Science, 2024, 36(3): 123-127.

[15] Jastania R, Wang P, Alqahtani B, et al. The Development of an interface instrument for collecting electromyography data and controlling a continuous passive motion machine[J]. Applied Sciences, 2023, 13(22): 12221.

[16] Boushel R, Langberg H, Olesen J, et al. Monitoring tissue oxygen availability with near infrared spectroscopy (NIRS) in health and disease[J]. Scandinavian Journal of Medicine & Science in Sports, 2001, 11(4): 213-222.

[17] Fredriksson I, Larsson M, Strömberg T. Measurement depth and volume in laser doppler flowmetry[J]. Microvascular Research, 2009, 78(1): 4-13.

[18] El Falou W, Duchêne J, Grabisch M, et al. Evaluation of driver discomfort during long-duration car driving[J]. Applied Ergonomics, 2003, 34(3): 249-255.

[19] SAE International. Subjective rating scale for evaluation of noise and ride comfort characteristics related to motor vehicle tires: SAE J1060_201405[S]. Warrendale: SAE International, 2014.

[20] Teich M C, Lowen S B, Jost B M, et al. Heart rate variability: measures and models[M]//Akay M. Nonlinear biomedical signal processing, dynamic analysis and modeling. Piscataway: IEEE Press, 2001, 2: 159-213.

[21] Chowdhury R H, Reaz M B I, Mohd Alauddin B M A, et al. Surface electromyography signal processing and classification techniques[J]. Sensors, 2013, 13(9): 12431-12466.

[22] Koirala B, Concas A, Sun Y, et al. Blood volume versus deoxygenated NIRS signal: computational analysis of the effects muscle O_2 delivery and blood volume on the NIRS signals[J]. Journal of Applied Physiology, 2021, 131(5): 1418-1431.

[23] Hwang J Y. Doppler ultrasonography of the lower extremity arteries: anatomy and scanning guidelines[J]. Ultrasonography, 2017, 36(2): 111-119.

[24] De Luca C J. The use of surface electromyography in biomechanics[J]. Journal of Applied Biomechanics, 1997, 13(2): 135-163.

[25] Bharadiya J. A tutorial on principal component analysis for dimensionality reduction in machine learning[J]. International Journal of Innovative Science and Research Technology, 2023, 8(5): 2028-2032.

[26] Alalawi H H, Alsuwat M S. Detection of cardiovascular disease using machine learning classification models[J]. International Journal of Engineering Research and Technology, 2021, 10(7): 151-157.

[27] Liu Q, Liu Z, Cui B Y, et al. Driving safety and comfort enhancement in urban underground interchanges via driving simulation and machine learning[J]. Sustainability, 2024, 16(21): 9601.

[28] Abdullah D M, Abdulazeez A M. Machine learning applications based on SVM classification a review[J]. Qubahan Academic Journal, 2021, 1(2): 81-90.

[29] Charbuty B, Abdulazeez A M. Classification based on decision tree algorithm for machine learning[J]. Journal of Applied Science and Technology Trends, 2021, 2(1): 20-28.

[30] Speiser J L, Miller M E, Tooze J, et al. A comparison of random forest variable selection methods for classification prediction modeling[J]. Expert Systems with Applications, 2019, 134: 93-101.

[31] Hatwell J, Gaber M M, Atif Azad R M. Ada-WHIPS: explaining AdaBoost classification with applications in the health sciences[J]. BMC Medical Informatics and Decision Making, 2020, 20(1): 250.

[32] Nti I K, Nyarko-Boateng O, Aning J. Performance of machine learning algorithms with different K values in K-fold CrossValidation[J]. International Journal of Information Technology and Computer Science, 2021, 13(6): 61-71.

[33] Widodo A O, Setiawan B, Indraswari R. Machine learning-based intrusion detection on multi-class imbalanced dataset using SMOTE[J]. Procedia Computer Science, 2024, 234: 578-583.

[34] Novakovic J D, Veljovic A, Ilic S S, et al. Evaluation of classification models in machine learning[J]. Theory and Applications of Mathematics and Computer Science, 2017, 7(1): 39-46.

[35] Hou J X, Li X Y, Li Z J, et al. An in vivo data-based computational study on sitting-induced hemodynamic changes in the external iliac artery[J]. Journal of Biomechanical Engineering, 2022, 144(2): 021007.

[36] Restaino R M, Holwerda S W, Credeur D P, et al. Impact of prolonged sitting on lower and upper limb micro- and macrovascular dilator function[J]. Experimental Physiology, 2015, 100(7): 829-838.

[37] Thosar S S, Bielko S L, Wiggins C C, et al. Differences in brachial and femoral artery responses to prolonged sitting[J]. Cardiovascular Ultrasound, 2014, 12(1): 50.

[38] Andersson B J, Ortengren R, Nachemson A, et al. Lumbar disc pressure and myoelectric back muscle activity during sitting. IV. studies on a car driver's seat[J]. Scandinavian Journal of Rehabilitation Medicine, 2013, 6(3): 128-133.

[39] Lin C T, Wu R C, Jung T P, et al. Estimating driving performance based on EEG spectrum analysis[J]. EURASIP Journal on Advances in Signal Processing, 2005(19): 3165-3174.

[40] De Looze M P, Kuijt-Evers L F M, Van Dieën J. Sitting comfort and discomfort and the relationships with objective measures[J]. Ergonomics, 2003, 46(10): 985-997.

[41] Luo S W. An automobile seat comfort evaluation method based on fuzzy support vector machine[C]//2008 9th International Conference on Computer-Aided Industrial Design and Conceptual Design. Piscataway: IEEE Press, 2008: 206-208.

Construction and comparison of machine learning models for driving comfort assessment using multimodal physiological signals

JIANG Weisheng[1], ZHOU Qianxiang[1,*], LIU Zhongqi[1], LIU Yushuang[2], WANG Jiatao[1], JIN Mengmeng[1]

(1. School of Biological Science and Medical Engineering, Beihang University, Beijing 100191, China;
2. Department of Ultrasound, China Rehabilitation Research Center (Beijing Boai Hospital), Beijing 100068, China)

Abstract: To achieve an objective assessment of driving comfort, this study developed and systematically compared multiple machine learning models based on multimodal physiological signals. 64 individuals had their electrocardiograms (ECGs), surface electromyograms (sEMGs), muscle oxygen saturation (SmO_2), and arterial blood flow signals recorded while they drove. Using comfort scores as labels, six types of models, including support vector machine (SVM), random forest (RF), and adaptive Boosting (AdaBoost), were constructed and optimized with principal component analysis, synthetic minority oversampling technique (SMOTE), and grid search. Results showed that multimodal signal fusion significantly improved classification performance. The SVM model achieved optimal performance with the feature set containing arterial blood flow parameters, reaching an average area under the curve (AUC) of 0.97 and an accuracy of 0.90. RF models consistently exceeded 0.95 AUC with fused signals. Practical signal combinations for portable monitoring were identified: "sEMG + ECG" dual-signal combination (average AUC is 0.90) and "SmO_2 + ECG + sEMG" triple-signal combination (average AUC is 0.92). This study offers empirical support for model selection and signal setup in objective driving comfort evaluation, providing insightful information for the development of in-car health monitoring systems.

Keywords: driving comfort; physiological signal; machine learning; classification models; signal fusion; health monitoring

Received: 2026-01-26; **Accepted**: 2026-03-30; **Published Online**: 2026-04-08 12:16

URL: link.cnki.net/urlid/11.2625.V.20260407.1830.001

Foundation item: National Key Research and Development Program of China (2023YFF0615904)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0695

2020—2024年火箭发射排放特性研究

宋佳，杨宇，韩旺*，杨立军
(北京航空航天大学 宇航学院，北京 100191)

摘 要：航天事业的快速发展在推动科技进步的同时，也带来了不容忽视的环境影响。为评估全球航天活动对大气环境的影响，基于2020—2024年的全球火箭发射数据，对不同类型火箭产生的CO_2、黑碳(BC)、H_2O、Al_2O_3和含氯化合物(Cl_y)等产物在水平和垂直空间的排放分布特性进行了研究和分析。结果表明：2020—2024年期间，全球火箭发射排放的气相排放物以CO_2和H_2O为主，两者排放量分别高达2.71×10^8 kg和1.44×10^8 kg，固相排放物主要包括BC和Al_2O_3，排放量分别为6.70×10^6 kg和7.80×10^6 kg；在水平空间上，火箭排放物呈现出显著的地域集中性，美国和中国主导贡献了全球BC、CO_2和H_2O的排放，Al_2O_3和Cl_y的排放则主要集中于中国、法国、美国和印度等国家；在大气层的垂直空间上，CO_2的排放集中于对流层，BC排放核心区位于平流层和中间层，受限于高海拔大气层的垂直对流强度低，BC易出现积聚效应，导致平流层及以上大气层的综合辐射强迫提升近1倍，表明相比处于对流层的航空运输业及传统工业，航天发射活动排放的BC对环境气候的危害更为严重。

关 键 词：二氧化碳；黑碳；液氧煤油；平流层排放；温室效应
中图分类号：V11
文献标志码：A **文章编号**：1001-5965(2026)10-3658-09

航天事业的发展有助于人类探索宇宙、拓展生存空间，推动了科学技术的进步，也为人类社会带来了深远的影响。航天技术不仅在通信、导航、气象监测等传统领域取得广泛的应用，同时，推动了太空旅游、低轨卫星星座建设和空间资源开发等新兴商业模式的兴起与演进[1-3]。

火箭是有效载荷安全可靠地进入太空的运载工具。随着各国航天活动的加速推进，火箭发射需求剧烈增长。2024年，中国执行了68次航天发射任务，美国火箭发射次数达到154次，创下历史新高。此外，欧洲、印度、俄罗斯等国家和地区的航天机构及私营企业也在加大发射频次。这一趋势表明，全球航天产业正进入高速发展阶段，但随之而来的环境问题也亟待关注。

运载火箭多采用碳氢燃料和固体燃料作为推进剂，燃烧过程中不仅会释放大量CO_2、H_2O、黑碳(black carbon, BC)等气候强迫因子，同时，还涉及Al_2O_3和含氯化合物(Cl_y)等污染物。尤其是当前商业航天正加速向低成本、高运载能力的可复用运载火箭方向发展，其单次发射可消耗数百吨推进剂，以美国的猎鹰9号运载火箭为例，其单次发射可消耗约5.23×10^5 kg的推进剂，排放出4.77×10^5 kg CO_2和1.15×10^4 kg BC[4-5]，CO_2是导致全球变暖的主要温室气体，而Ross等[6-7]研究表明，BC在平流层的单位质量气候影响效应是CO_2的10^5倍以上，对大气辐射强迫和气候系统的破坏作用更为严重。

与排放到对流层的航空运输业或生产工业不同，火箭尾气排放涉及对流层、平流层、中间层及

收稿日期：2025-09-29；**录用日期**：2025-10-17；**网络出版时间**：2025-11-12 14:57
网络出版地址：link.cnki.net/urlid/11.2625.V.20251112.1201.001
基金项目：北京市自然科学基金(L248007,3254053)；国家重点研发计划青年科学家项目(2022YFF0504500)；博士后创新人才支持计划(BX20240464)；国家自然科学基金青年基金项目（C类）(52506122)
* **通信作者**. E-mail：drwanghan@buaa.edu.cn

引用格式：宋佳，杨宇，韩旺，等. 2020—2024年火箭发射排放特性研究[J]. 北京航空航天大学学报，2026，52（10）：3658-3666.
Song J，Yang Y，Han W，et al. A study on emission characteristics of rocket launches (2020—2024)[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（10）：3658-3666（in Chinese）.

更高海拔的大气层区域, 由于平流层及以上高度区域垂直对流强度弱, 会产生积聚效应, 据研究估算, 平流层排放物的平均清除时间约为 3～5 年[8]。Kokkinakis 等[9] 以 Falcon 9 火箭执行 Thaicom-8 任务为原型, 通过高精度计算流体动力学模拟研究发现, 火箭发射会在高层大气中排放大量 CO_2 和氮氧化物等污染物, 在 70 km 高度, 单次排放的 CO_2 质量相当于 26 km^3 当地大气中所含的 CO_2。除各种会导致温室效应和臭氧消耗的废气外, 火箭燃烧和再入过程中产生的 BC 等固相颗粒可能对气候影响更大[10]。Samset 等的研究表明[11], 积聚效应导致大气层顶部 BC 的直接辐射强迫显著强于平流层, 两者相差 10 倍以上。Ryan 等[12] 基于 2019 年的火箭发射排放数据, 评估了航天排放对环境气候的影响, 结果表明, 火箭排放的单位质量 BC 的气候影响程度远高于航空运输业。Tsigaridis 等[13] 通过地球系统模型模拟发现, 在年发射 1 万次的场景下, 火箭排放的 BC 可使平流层负荷增加 3 倍, 导致对流层顶升温超 1.5 K, 并间接使平流层水汽含量增加约 50%, 进而对全球气候产生显著影响。

在全球航天活动迅猛增长的背景下, 亟需系统性评估火箭发射活动对气候和大气环境的影响。特别地, 火箭发射排放不仅具有显著的地区排放分布不均特性, 其发射过程连续穿越对流层、平流层、中间层、热层等多种大气层结构, 在不同高度上的分布存在巨大差异[10], 这种分布不均对环境的影响特性仍不清晰。因此, 本文从已有的航天发射频次数据出发, 定量分析了多种火箭发射排放物在水平、垂直空间上的排放特征, 进一步确立了垂直空间排放不均对环境的影响规律。本文不仅有助于完善航天活动的环境评估体系, 还能为未来绿色推进技术的发展提供数据基础和科学依据。需要指出的是, 现有研究表明, 排放物对环境的长周期影响需考虑大气流动[14-15], 而本文重点聚焦于火箭发射这一短时过程中的源头排放特性, 旨在揭示日益频繁的航天发射活动对环境的直接危害性, 因此, 未将大气流动的影响纳入考虑。

1　计算模型

1.1　排放估算模型

火箭发动机燃烧过程中会排放出多种产物, 其中, CO_2、BC、H_2O、Al_2O_3 等具有强烈的红外、可见光辐射吸收能力, 增强大气中的辐射强迫, 是不可忽视的温室因子[16-17], 而 Cl_y 在大气中受热易分解成氯自由基, 催化臭氧发生裂解, 对臭氧层造成破坏。因此, 本文基于 2020—2024 年间火箭发射频次及其发动机相关型号, 计算了 CO_2、BC、H_2O、Al_2O_3 和 Cl_y 的排放量, 并生成了全球火箭发射的排放物水平和垂直空间排放分布图。火箭发射频次信息及推进剂数据来源于公开网站(DISCOSweb、Gunter's Space Page、SatWiki 等)、学术论文[4, 18-19] 和官方文件[20-22]。一些型号的火箭缺少相关文件的支持, 本文综合考虑火箭长度、级数、起飞质量、推进剂类型和运载能力, 采用型号相近的火箭数据对其估计, 如表 1 所示。在被替代的 13 个型号中, 除谷神星 1 号、长征 11 号等少数型号外, 其余型号在 2020—2024 年间发射频次均低于 5 次, 其对总体排放估算的累积影响有限。缺失火箭主要集中于使用液氧煤油和固体推进剂的型号, 可能对这 2 类推进剂的消耗总量估算带来轻微偏差, 但替换型号的推进剂消耗量仅占全部型号推进剂总量的 2.3%, 因此, 该替代方法可能造成的误差很小。

表 1　数据缺失的火箭型号及其参考型号

Table 1　Rocket types with incomplete data and associated reference models

缺失型号	替代/参考型号	依据
火箭3号 (Astra Roket 3)	电子号 (Elecrton)	推进剂类型
千里马1号 (Chollima 1)	长征2号丙	推进剂类型
韩国固体燃料运载火箭 (GYUB-TV2)	米诺陶1号 (Minotaur 1)	推进剂类型、级数
凯洛斯(Kairos)	快舟1号	推进剂类型、级数
谷神星1号	沙维特 (Shavit)	推进剂类型、级数
捷龙3号	艾普斯龙CLPS (Epsilon 2 CLPS)	运载能力、级数
快舟11号	长征6号/快舟1号	起飞规模、运载能力、推进剂类型
长征11号	米诺陶1号 (Minotaur 1)	推进剂类型、级数
长征12号	长征7号甲	运载能力、推进剂类型
长征6号甲	长征7号甲	起飞规模
朱雀2号	安塔瑞斯230 (Antares 230)	起飞规模
朱雀2号改	安塔瑞斯230 (Antares 230)	起飞规模

排放物的排放因子(每千克推进剂的排放质量)因推进剂类型而异, 2020—2024 年间, 主流推进剂组合包括液氧-煤油、液氧-液氢、固体推进剂(高氯酸铵/铝)、自燃推进剂(肼类/四氧化二氮)和液氧-甲烷, 单位推进剂的排放如表 2 所示[4, 23]。

基于单次火箭发射的推进剂类型和消耗量, 可以计算出不同排放物释放量[23]。需特别指出, 本文未将发射任务失败、可回收火箭残留燃料等情形所引起的排放量减少纳入考虑。

1.2　排放物垂直空间排放分布计算模型

CO_2 和 BC 被视为温室效应贡献最强的 2 种物

表2 每千克推进剂燃烧产生的排放物质量

Table 2 Mass of pollutants emitted per kilogram of emissions burned

推进剂	排放物质量/kg				
	CO_2	BC	H_2O	Al_2O_3	Cl_y
液氧-煤油	0.91	0.022	0.36		
自燃推进剂	0.46	0.016	0.38		
固体推进剂	0.44	0.016	0.39	0.34	0.217
液氧-甲烷	0.55	0.005	0.45		
液氧-液氢			1.27		

质，本文重点针对这2种物质计算垂直空间上的排放特性。火箭的垂直排放取决于各级推进剂类型和质量及各级发动机的点火高度。Ross等[7]在研究中给出了0～100 km内不同高度下火箭推进剂质量消耗百分比。基于该数据和火箭各级点火高度，可以获得不同高度区间的推进剂消耗情况。本文编译的发射事件高度数据涵盖了目标高度范围（0～100 km），包括各型号火箭助推器发动机点火、助推器发动机截止、一级发动机点火、一级发动机截止和二级发动机点火的高度，第二发动机点火以外的指令通常发生在100 km以上。具体计算过程如下：

1）若事件结束高度大于100 km，即火箭相应级携带的推进剂并未完全在100 km内消耗，假定推进剂消耗量与高度成正比，以此得到本级实际消耗量为

$$m_{\mathrm{eff}} = m\frac{100 - h_{\mathrm{s}}}{h_{\mathrm{e}} - h_{\mathrm{s}}} \tag{1}$$

式中：m为推进剂质量；m_{eff}为本级推进剂实际消耗质量；h_{s}为阶段起始高度；h_{e}为阶段结束高度。

2）Ross等[7]给出的不同高度下推进剂消耗占比覆盖范围为0～100 km。然而，实际火箭发射的各阶段高度跨度并非0～100 km。因此，需对Ross等给出的占比进行缩放，以得到各发射阶段对应的推进剂消耗占比：

$$P_{\mathrm{norm}}(h) = P(h) \Big/ \sum_{h=h_{\mathrm{s}}}^{h_{\mathrm{e}}} P(h) \tag{2}$$

式中：$P(h)$为每单位高度下推进剂消耗量占比；$P_{\mathrm{norm}}(h)$为不同火箭发射阶段归一化后的占比。

3）由高度区间内的推进剂消耗质量和相应的排放因子计算排放量：

$$E(i,h) = P_{\mathrm{norm}}(h) m_{\mathrm{eff}} E_{\mathrm{If}}(i,p) \tag{3}$$

式中：$E(i,h)$为高度h下物质i的排放质量；$E_{\mathrm{If}}(i,p)$为推进剂p的物质i最终排放因子。

本文将排放的垂直空间限定在0～100 km的范围内（以卡门线为界），垂直空间分辨率为5 km。

对于BC，其排放后温度较高，会和大气中的空气发生氧化反应。由于不同高度下氧含量不同，氧化速率亦不相同，因此，最终BC排放量随海拔变化差异较大。本文采用Wayson等[24]基于Atlas II火箭排放数据构建的指数拟合公式，对不同高度下因氧化速率不同导致的BC排放量差异进行计算：

$$E_{\mathrm{If,BC}} = E_{\mathrm{Ip,BC}} \max\left\{0.04, \min\left[1, 0.04\mathrm{e}^{\frac{0.12}{h-15}}\right]\right\} \tag{4}$$

式中：$E_{\mathrm{Ip,BC}}$为初等排放因子，使用表2中的取值；$E_{\mathrm{If,BC}}$为考虑不同海拔下氧化速率后的BC排放垂直分布特性。

2 结果与讨论

2.1 总体排放情况

图1展示了火箭发射次数随年份的变化关系。可以看到，全球航天发射活动呈现出明显的阶段性特征：1957—1965年间，年发射次数从2次激增至1965年的112次，这一爆发式增长主要源于1957年斯普特尼克1号发射成功后引发的美苏太空竞赛；1970—1990年，进入相对稳定的平台期，年发射量维持在100～130次之间，期间航天发展从军事对抗转向实用化卫星应用，包括通信、气象等民用系统的部署[25]；1990年至21世纪初，发射次数出现了显著下降趋势，2004年降至53次，主要归因于苏联解体后航天预算缩减及美国航天飞机计划的结束[26]；而2013年后，随着商业发射市场逐渐成熟，发射量重现指数级增长趋势，2024年全球火箭发射达254次，较2023年增长约17%，若维持当前增速，预计到2050年年发射量可能突破3 000次。可以预见的是，指数型增长的发射频次势必在未来带来更为严峻的环境问题。本文基于Kaya恒等式[27]的火箭BC排放预测模型，假定未来火箭市场以常规增速场景发展（发射次数年增长率为5.6%），预计到2050年，火箭发射产生的BC排放将达到5.741 5×10^7 kg/年，是2024年的26倍。

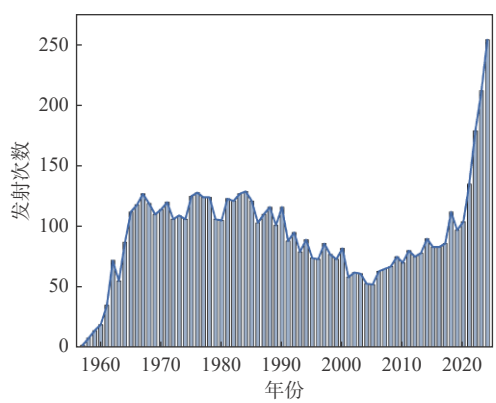


图1 1957—2024火箭发射次数

Fig. 1 Number of rocket launch from 1957 to 2024

基于 2020—2024 年火箭发射频次, 结合相关推进剂类型计算了不同物质的排放量, 如图 2 所示。5 年内, 排放最多的分别为 CO_2 和 H_2O, 分别为 2.71×10^8 kg 和 1.44×10^8 kg。以 BC 和 Al_2O_3 为代表的固相污染物排放量低于 CO_2 和 H_2O 排放量 2 个量级, 分别为 6.70×10^6 kg 和 7.80×10^6 kg。排放量最低的为 Cl_y 污染物, 仅为 5.10×10^6 kg。为验证本文采用的排放模型有效性, 将 2020 年计算结果与 Ryan 等[12] 计算的 2019 年结果进行比较(考虑到 2019 年与 2020 年火箭发射次数十分接近, 分别为 97 次和 104 次)。本文估算的 2020 年火箭排放量为 H_2O: 1.6×10^7 kg、BC: 6.9×10^5 kg、Cl_y: 8.2×10^5 kg、Al_2O_3: 1.2×10^6 kg, 其与 Ryan 等[12] 计算的 2019 年结果基本吻合, 分别为 H_2O: 1.1×10^7 kg、BC: 5.0×10^5 kg、Cl_y: 9.0×10^5 kg、Al_2O_3: 1.6×10^6 kg。

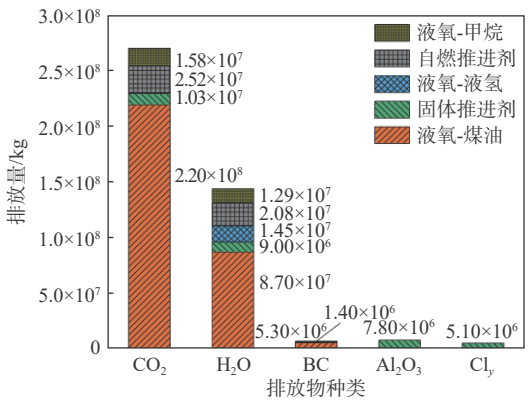


图 2　2020—2024 年间排放物释放量

Fig. 2　Emissions during 2020 to 2024

从推进剂种类消耗量来看(如图 3 所示), 5 年内火箭发射消耗量最多的燃料为液氧-煤油推进剂, 消耗量达到 2.2×10^8 kg, 因其具备高推重比、低成本的优势, 兼顾可靠性和经济性, 仍是当前火箭发射任务的首要选择, 从表 2 可以看出, 液氧-煤油的 CO_2 和 BC 排放因子显著高于其余组合推进剂, 因此, 液氧-煤油对 CO_2 和 BC 的排放贡献最高, 如图 2 所示。自燃推进剂消耗量仅次于液氧-煤油推进剂, 为 4.8×10^7 kg, 其 CO_2 和 BC 排放因子低于液氧-煤油推进剂, CO_2 和 BC 排放贡献也远低于液氧-煤油推进剂, 需指出的是, 受限于自燃推进剂高毒性特性, 许多商业航天已弃用此类推进剂。固体推进剂和液氧/液氢 2 种燃料的消耗量相当, 显著低于传统的液氧-煤油和自燃推进剂, 两者消耗量均为 1.0×10^7 kg 左右, 由于固体推进剂含有铝、氯等元素, 是 Al_2O_3 和 Cl_y 污染物排放的唯一来源; 液氧/液氢燃烧产物仅包含 H_2O, 其 H_2O 的排放因子显著高于其余推进剂。液氧-甲烷推进剂主要用在超重-星舰、朱雀 2 号和火神-半人马等新一代运载火箭, 消

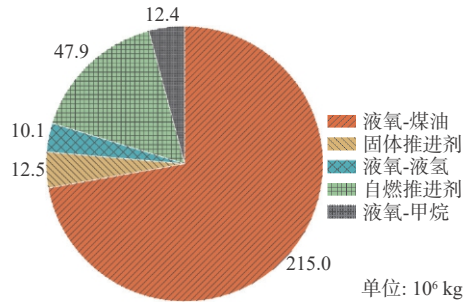


图 3　2020—2024 年间推进剂消耗量

Fig. 3　Propellant consumption during 2020—2024

耗量与固体推进剂相当, 为 1.24×10^7 kg, 从表 2 可以看出, 该推进剂的 BC 排放因子显著低于液氧-煤油推进剂, 污染低, 同时, 甲烷不易结焦, 复用性能强, 与未来商业化重复使用需求高度匹配, 是未来航天最具潜力的推进剂。2024 年液氧煤油的消耗量达到 8.38×10^7 kg, 占推进剂总量的 70%; 若将其中 50% 的液氧-煤油任务替换为液氧-甲烷, 经计算, BC 排放量将由 2.17×10^6 kg 降至 1.46×10^6 kg, 减少了 32.7%。若在 2050 年, 液氧甲烷推进剂的占比达到 50%, BC 排放量将从 $5.741\ 5\times10^7$ kg 降至 $4.154\ 8\times10^7$ kg 左右, 减少了 27.6%。

2.2　水平空间排放分布特性

图 4 展示了不同物质排放量之和的水平分布特性, 同时, 给出了各国的不同排放物的占比。各

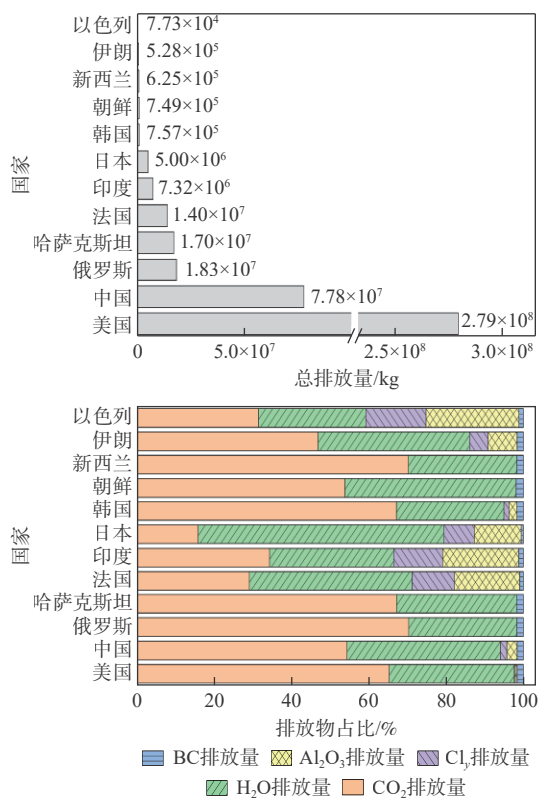


图 4　2020—2024 年间总排放物的空间分布及占比

Fig. 4　Spatial distribution and proportion of total emissions from 2020 to 2024

国火箭发射的主要排放物为 CO_2 和 H_2O，两者合计占各国总排放量的 60%～97%，反映了当前火箭推进剂以碳氢燃料和液氧为主的技术特点。中国、美国、俄罗斯等航天大国 CO_2 占比高达 54%～70%，显著高于其他排放物。法国的欧洲航天发射场和日本则相反，H_2O 占比高于 CO_2，且 Cl_y 和 Al_2O_3 也具有较高的占比。

图 5～图 9 展示了 2020—2024 年火箭发射产生的 CO_2、BC、H_2O、Al_2O_3 和 Cl_y 排放量的全球水平空间排放分布特性。

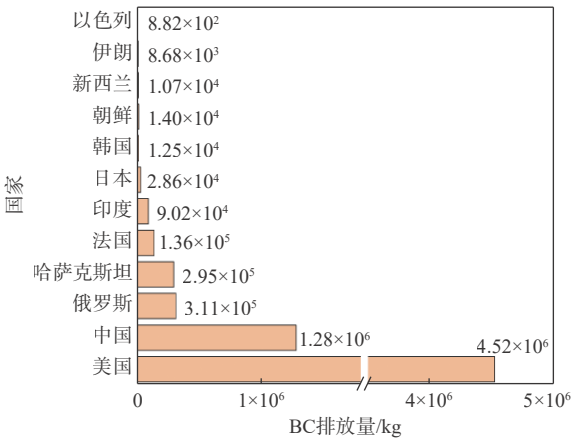


图 5 2020—2024 年间 BC 排放的空间分布

Fig. 5 Spatial distribution of total BC emissions during from 2020 to 2024

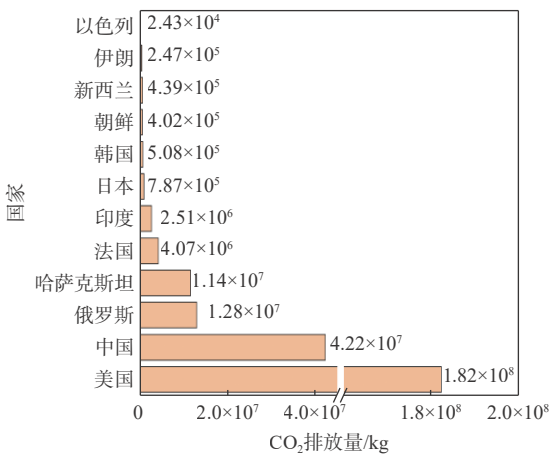


图 6 2020—2024 年间 CO_2 排放的空间分布

Fig. 6 Spatial distribution of total CO_2 emissions from 2020 to 2024

从图 5 可以看出，全球 BC 排放呈现出明显的水平空间分布不均特征，BC 排放主要集中于北半球，其中，美国、中国和俄罗斯为主导排放国家，美国排放量最为显著，达到 4.52×10^6 kg，主要是由液氧-煤油为推进剂的猎鹰 9 号高频次发射带来的。中国的 BC 排放量仅次于美国，为 1.28×10^6 kg，约为美国的 1/3。除中国、美国和俄罗斯外，印度、伊朗、法国和日本等国家也开展了一定的火箭发射任务，BC 排放总量约为 9.1×10^5 kg。

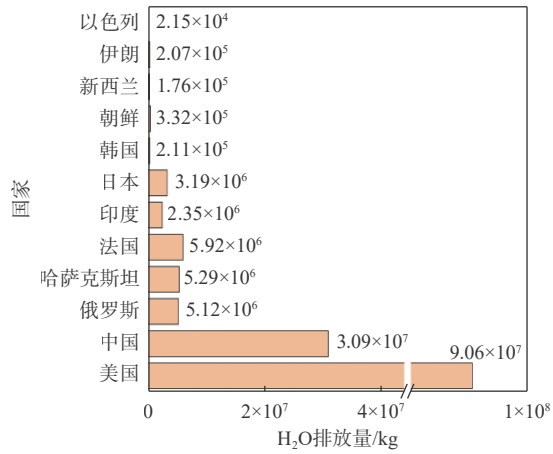


图 7 2020—2024 年间 H_2O 排放的空间分布

Fig. 7 Spatial distribution of total H_2O emissions from 2020 to 2024

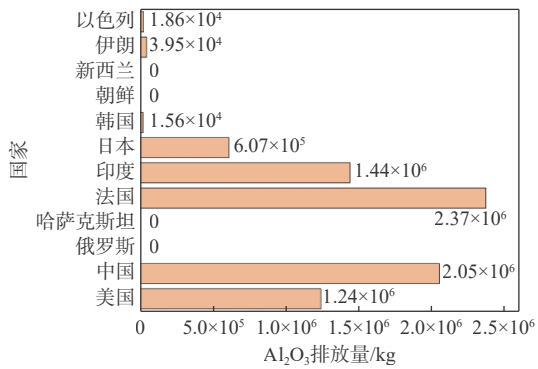


图 8 2020—2024 年间 Al_2O_3 排放的空间分布

Fig. 8 Spatial distribution of total Al_2O_3 emissions from 2020 to 2024

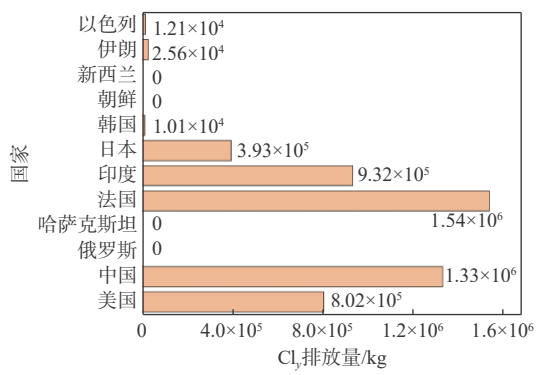


图 9 2020—2024 年间 Cl_y 空间分布

Fig. 9 Spatial distribution of total Cl_y emissions from 2020 to 2024

从图 6 可以看出，CO_2 与 BC 在水平空间上的排放分布特性一致，主要集中于美国、中国和俄罗斯 3 个国家，其中，美国排放量仍位居第一，排放量达 1.82×10^8 kg，表明 BC 和 CO_2 两者排放量具有一定的关联特性，这主要是因为液氧-煤油是中国、美

国和俄罗斯火箭发射的首要推进剂选择, 且该推进剂的 BC 和 CO_2 排放因子均显著高于其余推进剂。

图 7 展示了 H_2O 的水平空间排放分布特性。可以看出, H_2O 与 CO_2 及 BC 的水平空间排放分布一致, 集中于美国、中国和俄罗斯 3 个国家。尽管液氧-液氢推进剂的 H_2O 排放因子为液氧-煤油的 3 倍, 但其在 2020—2024 年间消耗量不足液氧-煤油的 1/10, 因此, 主导 H_2O 排放的仍为液氧-煤油推进剂。

图 8 和图 9 分别展示了 Al_2O_3 和 Cl_y 的水平空间排放分布, 由于 Al_2O_3 和 Cl_y 这 2 种污染物均来源于固体推进剂, 因此, 两者的排放特性呈现出高度的一致性。图 8 和图 9 表明, Al_2O_3 和 Cl_y 这 2 种污染物排放呈现出强区域集中性, 且主要集中于北半球国家, 这与上述 CO_2、BC 及 H_2O 的排放特性相同。然而, Al_2O_3 和 Cl_y 排放量较多的国家由美国转变为中国和法国, 其在 2020—2024 年的 Al_2O_3 排放量均超过 2.0×10^6 kg, Cl_y 的排放量则超过 1.2×10^6 kg。这一数据反映出当前全球主要航天国家在推进剂选择上的差异: 中国在发射任务中呈现液氧-煤油与固体推进剂并重的格局, 法国则以固体推进系统为主导, 尤其体现在阿丽亚娜系列火箭中, 而美国则更倾向于采用液氧-煤油, 固体推进剂使用比例明显偏低。对于美国而言, 其 Al_2O_3 和 Cl_y 排放量与印度相当, 两者污染物排放量分别约为 1.24×10^6 kg 和 8.02×10^5 kg, 但与中法两国排放量差距并不显著。在 CO_2、BC 和 H_2O 排放量较为靠前的俄罗斯, 由于极少采用固体推进技术, 其 Al_2O_3 和 Cl_y 排放量十分有限。

总体上, 在 2020—2024 年间, 各物质的排放量呈现出极强的区域集中性, 液氧-煤油推进剂是 CO_2、BC 和 H_2O 等物质的主要来源, 排放集中于美国、中国和俄罗斯; 而 Al_2O_3 和 Cl_y 这 2 种污染物则主要集中于法国、中国、印度和美国。

需指出的是, 本文聚焦于航天发射污染物初排阶段的分布特性, 并未考虑大气扩散与输送对后续污染物分布特性的影响, 因此, 对长周期下气候的影响评估存在一定偏差。航天发射污染物排放对平流层增温、大气环流和臭氧变化的长周期影响评估需基于污染物的化学演化机理、气候效应及大气输运模型, 进一步结合本文数据进行模拟计算。

2.3　不同级别运载火箭的排放特性

本节以 CO_2、BC 和 H_2O 排放量较高的中国和美国为对象, 分析不同近地轨道运载能力的火箭排放特征, 确立高排放量的火箭类型, 以精准评估航天领域的碳足迹, 优化发射任务结构。图 10 展示了中国和美国的小型(运载能力$<2\times10^3$ kg)、中型(运载能力为 2×10^3～1.5×10^4 kg)、大型(运载能力为 1.5×10^4～5×10^4 kg)和重型(运载能力$>5\times10^4$ kg)火箭 CO_2、H_2O 和 BC 排放的占比。图 10 表明, 中国航天任务以长征系列中型运载火箭为主, 其 CO_2 排放量达到 3.1×10^7 kg, 占总排放量的 73.8%; H_2O 和 BC 排放也呈现出类似的分布模式, 占比分别为 76.4% 和 78.0%。对于美国而言, 大型火箭是 CO_2、H_2O 和 BC 排放的主要来源, 其中, CO_2 和 BC 占比分别为 82.7% 和 86.4%, 而 H_2O 占比略低, 为 75.3%。大型火箭是中国的 CO_2、H_2O 和 BC 的第二大来源, 占比在 17%～22% 之间。相比之下, 美国第二大排放来源为重型火箭。由于小型火箭载荷较小, 单位发射成本较高, 发射次数不多且单次规模较小, 因此, 在 2020—2024 年间, 小型火箭的排放量占比低, 尤其是转向超重型火箭发展的美国, 其小型火箭排放量占比均低于 1%。

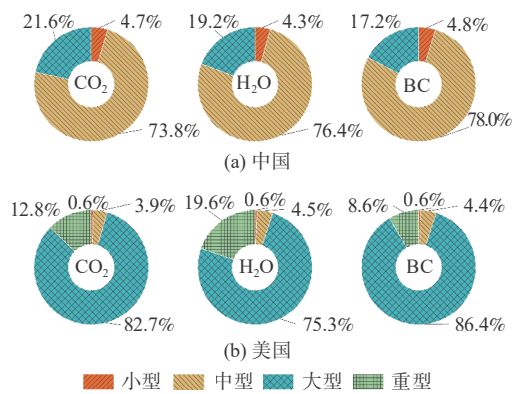


图 10　中国和美国不同级别运载火箭的排放占比

Fig. 10　Pollutant emission proportions from different launch vehicle classes in China and the United States

综上所述, 中美两国的排放来源上存在显著差异, 中国排放主要源于中型火箭, 次要排放源均为大型火箭; 而美国的主要排放源为大型火箭, 第二大来源为重型火箭。这一差异反映了两国在火箭运载体系发展方向上的分化, 美国更倾向于向重型平台演进, 而中国目前仍以中大型火箭为主力。

2.4　垂直空间排放分布特性

图 11 展示了 2020—2024 年 BC 和 CO_2 排放量随大气层高度的变化。可见, 2020—2024 年全球火箭发射所产生的 CO_2 排放量呈现出明显的底层集中型分布, 45.9% 的 CO_2 排放于对流层区域, 随着大气层高度增加, CO_2 排放量逐步下降, 在海拔高于 80 km 的热层, CO_2 排放量不及对流层的 1/10。在对流层区域, CO_2 排放量随高度的上升迅速下降, 从 5 km 的 4.8×10^7 kg 下降至 15 km 的 1.4×10^7 kg, 降幅达 70.8%。在平流层、中间层及热层区域, CO_2 排

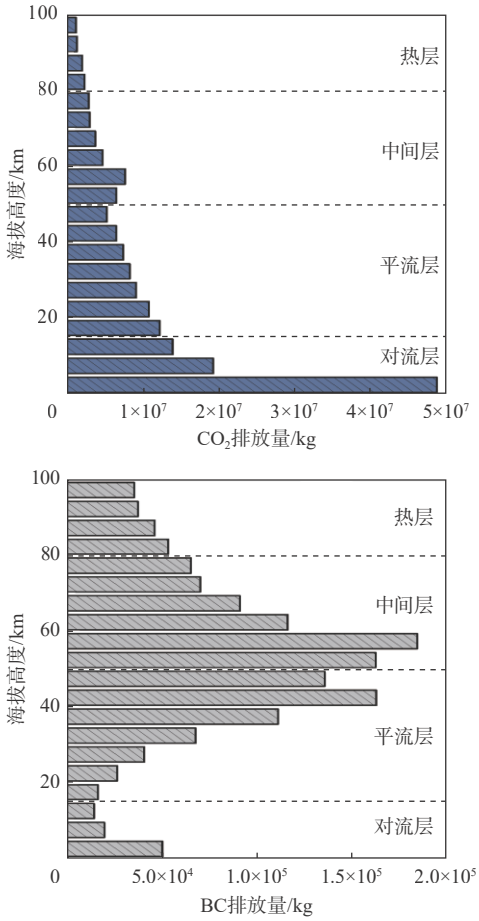


图 11 2020—2024 年间 0～100 km CO_2 和 BC 的垂直空间排放分布

Fig. 11 Vertical distribution (0-100 km altitude) of CO_2 and BC emissions during 2020—2024

放量随高度降幅较为平缓。CO_2 的底层集中型排放特性主要源于一级发动机作为火箭的主要推进段，燃料消耗量占总量的 60%～80%，同时，一级助推阶段主要在对流层完成。

相较于 CO_2，BC 的排放在垂直方向上呈现出显著不同的分布特性。图 11 表明，随高度变化的 BC 排放量呈现出近似正态分布，74% 的 BC 排放于平流层和中间层。峰值出现于 60 km 高度的中间层，这与 Barker[4] 和 Ryan[12] 等预测的分布规律一致。在对流层，BC 排放量随高度上升而下降，这一趋势与 CO_2 一致。随着高度上升至对流层，BC 排放量呈现出显著的上升趋势，当高度从 20 km 上升至 55 km 时，BC 排放量增加 9 倍。在中间层和热层区域，BC 排放量开始随高度的增加逐步下降。由于高温废气羽流中的后燃，喷嘴出口平面处的 BC 被部分氧化为 CO 和 CO_2。由于氧化需要周围空气中的 O_2，因此，在较高海拔地区氧化量会减少，BC 排放量增加。

平流层及以上高度的大气层由于垂直对流作用减弱，BC 易发生积聚效应，对环境更严重的破坏。前人通常采用辐射强迫衡量 BC 对地球大气环境中输入输出能量平衡的影响程度。为进一步定量评估不同高度下 BC 对总辐射强迫的相对贡献，本文计算了随高度变化的单位 BC 质量辐射强迫（normalized direct radiative forcing, NDRF）[11] 与各高度区间上 BC 排放量的乘积，获得了综合辐射强迫，更直观地反映 BC 辐射效应的高度相关性，图 12 绘制了未考虑积聚效应的综合辐射强迫估算结果（NDRF 数值参考 2.5 km 海拔的数据，约为 2 000 W/g）。可以看出，在对流层内，是否考虑积聚效应对综合辐射强迫影响较小，主要原因在于该区域 BC 排放量较低，且 NDRF 在该区域处于上升趋势，整体处于较低水平。当海拔来到平流层和中间层后，NDRF 数值升高至 3 750 W/g，较 2.5 km 提升近 1 倍，使该区域的综合辐射强迫提升 98.9%。在海拔更高的热层区域，NDRF 几乎稳定在 4 000 W/g 左右，积聚效应带来的辐射强迫增幅进一步上升至 101.0%。综上，在对流层，火箭排放的 BC 引发的积聚效应对环境影响较为有限，主要受限于 BC 排放量较低；而在海拔更高的平流层、中间层及热层，BC 的积聚效应显著增强其气候影响，导致综合辐射强迫提升近 1 倍。因此，相比处于对流层的航空运输业及传统工业，航天发射活动排放的单位质量 BC 对气候环境的危害可能更为严重，火箭发动机在专注燃烧性能的同时，也不可忽视 BC 的生成与控制。为有效降低火箭发动机的碳烟排放，最根本的技术路径在于从源头抑制其生成，可从燃料组分、喷注雾化、燃烧组织和结构设计等多个维度协同优化，如控制关键燃烧区域的混合比[28]、在可行范围内适度降低燃烧室压、优化喷注器构型以提升

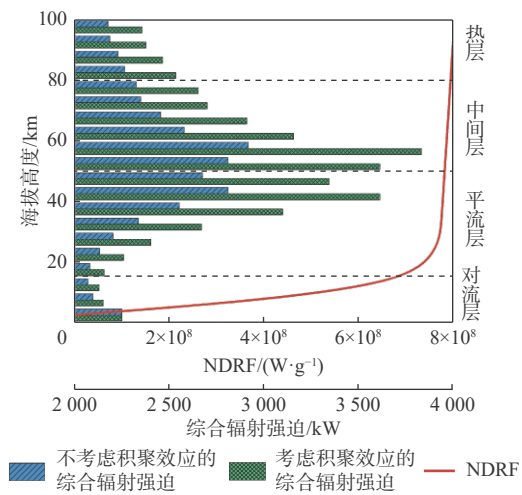


图 12 考虑和不考虑积聚效应时 BC 综合辐射强迫的垂直剖面

Fig. 12 Vertical profiles of BC comprehensive radiative forcing with and without considering accumulation effects

掺混均匀性, 从而消除局部富燃区[29], 减少因不完全燃烧产生的 BC 排放。

3　结　论

本文从已有的航天发射频次数据出发, 计算了 2020—2024 年全球火箭发射产生的 CO_2、H_2O、BC、Al_2O_3 及 Cl_y 等多种物质排放量, 评估了不同燃料、火箭运载量对不同排放物的贡献, 进一步分析了航天发射活动产生的排放物在水平、垂直空间上的排放分布特性, 以此确立了 BC 的积聚效应对环境影响程度。研究主要结论如下:

1) 从排放物类型分析, CO_2 和 H_2O 是火箭发射排放的主要气相产物, 排放量分别为 2.71×10^8 kg 和 1.44×10^8 kg, 以 BC 和 Al_2O_3 为代表性的固相污染物的排放量低于 CO_2 和 H_2O 排放量 2 个量级, 分别为 6.70×10^6 kg 和 7.80×10^6 kg, 液氧-煤油是 CO_2、H_2O、BC 主要来源, 固体推进剂则是 Al_2O_3 唯一来源。因此, 应大力推行液氧-甲烷替代传统液氧–煤油推进剂, 从源头上降低 BC 排放。

2) 在水平空间上, 火箭发射排放呈现出显著的地域集中性, 美国和中国贡献了全球大部分的 BC、CO_2 和 H_2O 排放, 这与两国大量采用液氧-煤油推进剂开展发射活动相关。相比之下, Al_2O_3 和 Cl_y 则主要集中于中国、法国、美国和印度等国家。对于中美两国, 除应减少液氧-煤油推进剂的使用外, 亟需开发传统固体推进剂的绿色替代燃料。

3) 从火箭运载能力分析, 中国的排放主要源于中型火箭, 其 CO_2、H_2O 和 BC 的次要排放源为大型火箭; 而美国主要源于大型火箭, 次要排放源为重型火箭, 表明美国更倾向于向超重型平台发展, 而中国目前仍以中大型火箭为主力。

4) 在垂直空间上, CO_2 的排放集中于对流层, 其排放强度随高度迅速降低, 而 BC 排放核心区位于平流层中上部和中间层中下部。由于这些高海拔大气层的垂直对流强度低, BC 易出现积聚效应, 导致高海拔区域的综合辐射强迫提升近 1 倍。鉴于此, 应优先建立针对平流层排放的长期监测与评估体系, 并将高空积聚效应纳入航天活动的环境影响评价中。同时, 优先在高海拔区域采用低 BC 排放的燃料替代传统液氧-煤油推进剂。

参考文献（References）

[1] 严一粟, 郭继峰. 月球原位资源开发体系研究进展[J]. 宇航学报, 2024, 45(8): 1155-1171.
Yan Y S, Guo J F. Progress of lunar in-situ resource utilization system[J]. Journal of Astronautics, 2024, 45(8): 1155-1171(in Chinese).

[2] 王巍, 姚伟. 近地小天体采矿体系构想及效益初步评估[J]. 空间科学与试验学报, 2024, 24(1): 13-22.
Wang W, Yao W. Concept and preliminary evaluation of near earth asteroids mining architecture[J]. Journal of Space Science and Experiment, 2024, 24(1): 13-22(in Chinese).

[3] 邵文凯, 刘尘旻, 张克非, 等. 近地小行星矿物分离与资源提取的挑战与展望[J]. 宇航学报, 2025, 46(1): 17-36.
Shao W K, Liu C M, Zhang K F, et al. Challenges and prospects for mineral separation and resource extraction from near-earth asteroids[J]. Journal of Astronautics, 2025, 46(1): 17-36(in Chinese).

[4] Barker C R, Marais E A, Mcdowell J C. Global 3D rocket launch and re-entry air pollutant and CO_2 emissions at the onset of the megaconstellation era[J]. Scientific Data, 2024, 11: 1079.

[5] SpaceX. Falcon 9[EB/OL]. (2024-11-20)[2025-07-18]. https://www.spacex.com/vehicles/falcon-9/.

[6] Ross M N, Mills M, Toohey D. Potential climate impact of black carbon emitted by rockets[J]. Geophysical Research Letters, 2010, 37(24): 2010GL044548.

[7] Ross M N, Sheaffer P M. Radiative forcing caused by rocket engine emissions[J]. Earth's Future, 2014, 2(4): 177-196.

[8] Dallas J A, Raval S, Alvarez Gaitan J P, et al. The environmental impact of emissions from space launches: a comprehensive review[J]. Journal of Cleaner Production, 2020, 255: 120209.

[9] Kokkinakis I W, Drikakis D. Atmospheric pollution from rockets[J]. Physics of Fluids, 2022, 34(5): 056107.

[10] Ross M N, Jones K L. Implications of a growing spaceflight industry: climate change[J]. Journal of Space Safety Engineering, 2022, 9(3): 469-477.

[11] Samset B H, Myhre G. Vertical dependence of black carbon, sulphate and biomass burning aerosol radiative forcing[J]. Geophysical Research Letters, 2011, 38(24): L24802.

[12] Ryan R G, Marais E A, Balhatchet C J, et al. Impact of rocket launch and space debris air pollutant emissions on stratospheric ozone and global climate[J]. Earth's Future, 2022, 10(6): e2021EF002612.

[13] Tsigaridis K, Field R, Bauer S E, et al. Composition and climate impacts of increasing launches to low earth orbit[C]//Proceedings of the AIAA SCITECH 2024 Forum. Reston: AIAA, 2024: 2168.

[14] Butchart N. The brewer-Dobson circulation[J]. Reviews of Geophysics, 2014, 52(2): 157-184.

[15] Maloney C M, Portmann R W, Ross M N, et al. The climate and ozone impacts of black carbon emissions from global rocket launches[J]. Journal of Geophysical Research: Atmospheres, 2022, 127(12): e2021JD036373.

[16] Solomon S, Rosenlof K H, Portmann R W, et al. Contributions of stratospheric water vapor to decadal changes in the rate of global warming[J]. Science, 2010, 327(5970): 1219-1223.

[17] Danilin M Y, Shia R L, Ko M K W, et al. Global stratospheric effects of the alumina emissions by solid-fueled rocket motors[J]. Journal of Geophysical Research: Atmospheres, 2001, 106(D12): 12727-12738.

[18] Pradon C V M, Eastham S D, Chossière G, et al. Global three-dimensional emission inventory for launch vehicles from 2009 to 2018[J]. Journal of Spacecraft and Rockets, 2023, 60(3): 716-727.

[19] Brown T F M, Bannister M T, Revell L E, et al. Worldwide rocket

launch emissions 2019: an inventory for use in global models[J]. Earth and Space Science, 2024, 11(10): e2024EA003668.

[20] ArianeGroup. Ariane 6, a versatile launcher for all missions [EB/OL]. (2024-08-20)[2025-07-18]. https://www.ariane.group/en/space-transportation/ariane-6/.

[21] JAXA. H3 launch vehicle[EB/OL]. (2025-06-10)[2025-07-18]. https://global.jaxa.jp/projects/rockets/h3/.

[22] United Launch Alliance. Vulcan centaur[EB/OL]. (2025-06-08)[2025-07-18]. https://www.ulalaunch.com/rockets/vulcan-centaur.

[23] Desain J D, Brady B B. Potential atmospheric impact generated by space launches worldwide-update for emission estimates from 1985 to 2013: TOR-2014-02140[R]. El Segundo: The Aerospace Corporation, 2014.

[24] Consulting L, Wayson A R L. Commercial space vehicle emissions modeling[M]. Washington, D. C. : Transportation Research Board, 2021.

[25] Launius R D. The smithsonian history of space exploration: from the ancient world to the extraterrestrial future[M]. Washington, DC: Smithsonian Books, 2018.

[26] Bainbridge W S, Roco M C. Handbook of science and technology convergence[M]. Berlin: Springer International Publishing, 2014.

[27] Bergero C, Gosnell G, Gielen D, et al. Pathways to net-zero emissions from aviation[J]. Nature Sustainability, 2023, 6(4): 404-414.

[28] 张锋, 严宇, 杨伟东, 等. 气氧/煤油富燃燃气发生器积碳特性试验研究[J]. 火箭推进, 2017, 43(6): 76-81.
Zhang F, Yan Y, Yang W D, et al. Experimental investigation on soot deposition in a fuel-rich GOX/kerosene gas generator[J]. Journal of Rocket Propulsion, 2017, 43(6): 76-81(in Chinese).

[29] Turushov V, Orlov V, Efremov Y, et al. Development status of a soot-free fuel-rich kerosene tripropellant preburner for reusable liquid rocket engine applications[C]//Proceedings of the 31st Joint Propulsion Conference and Exhibit. Reston: AIAA, 1995: 3002.

A study on emission characteristics of rocket launches (2020—2024)

SONG Jia，YANG Yu，HAN Wang*，YANG Lijun

(School of Astronautics，Beihang University，Beijing 100191，China)

Abstract：The rapid development of spaceflight has significantly advanced science and technology, but it also poses non-negligible environmental challenges. This study focuses on major emissions from rocket launches, including CO_2, black carbon (BC), H_2O, Al_2O_3, and chlorine species (Cl_y). To assess the atmospheric influence of space activities, their horizontal and vertical spatial distributions are examined using worldwide launch data from 2020 to 2024. The results indicate that gaseous emissions are dominated by CO_2 and H_2O, with total quantities of approximately 2.71×10^8 kg and 1.44×10^8 kg, respectively. In contrast, particulate emissions consist mainly of BC and Al_2O_3, with estimated releases of 6.70×10^6 kg and 7.80×10^6 kg, respectively. Horizontally, the emissions exhibit strong regional concentration, with the United States and China accounting for the majority of CO_2, H_2O, and BC emissions. Al_2O_3 and Cl_y emissions are mainly concentrated in China, France, the United States, and India. Vertically, CO_2 emissions are concentrated in the troposphere, while BC emissions peak in the stratosphere and mesosphere. Due to the persistence of BC in the upper atmosphere, a near twofold increase in net radiative forcing can be observed in these regions. This suggests that because of their high-altitude BC emissions, space launch operations are more dangerous to the climate than aircraft and conventional industrial sources that are limited to the troposphere.

Keywords：carbon dioxide; black carbon; kerosene; stratospheric emission; greenhouse effect

Received：2025-09-29；**Accepted**：2025-10-17；**Published Online**：2025-11-12 14：57

URL：link.cnki.net/urlid/11.2625.V.20251112.1201.001

Foundation items：Beijing Municipal Natural Science Foundation (L248007,3254053); Young Scientist Program of the National Key Research and Development Program of China (2022YFF0504500); Postdoctoral Innovative Talent Support Program (BX20240464); Young Scientists Fund of National Natural Science Foundation of China (Category C) (52506122)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0594

基于触发式批处理辨识的固定翼无人机自适应跟踪控制

邸尚明[1]，申佳军[1]，刘烨晖[1]，王薇[1,2,*]

(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 杭州创新研究院，杭州 310051)

摘 要：针对欠驱动与参数不确定性耦合导致的固定翼无人机控制设计与稳定性分析难题，提出一种基于触发式批处理辨识的自适应分层控制方法。为处理欠驱动问题，基于独轮车模型将原固定翼无人机模型转换为虚拟积分器与线性参数化模型；结合反步法与固定时间控制技术，建立等效飞行跟踪控制器。通过分析参数不确定性对无人机运动状态的作用机理，采用谱分解技术构造线性回归方程以定量收集激励信息，确保收集激励信息的丰富性和线性回归方程的有界性。为消除参数不确定性的影响，建立触发式批处理参数估计器，在触发函数满足特定条件时充分利用已收集的激励信息更新参数估计。与固定翼无人机的现有分层控制策略相比，所提框架实现了误差收敛性、激励丰富性与参数辨识精度的统一。

关 键 词：数据驱动控制；自适应控制；固定时间控制；固定翼无人机；欠驱动特性

中图分类号：TP273; V249

文献标志码：A **文章编号**：1001-5965(2026)10-3667-10

固定翼无人机作为航空作业体系内一类具备典型优势的飞行器，凭借其续航时间长、飞行速度稳定、航程覆盖范围广等特性，在战略决策[1]、环境监测[2]、测绘建模[3]等领域具有广泛应用。然而，固定翼无人机模型存在欠驱动特性，系统鲁棒性较差，对外部扰动敏感且难以实现高精度轨迹跟踪。此外，无人机动力学模型的固有参数不确定性会影响飞行控制性能，甚至带来安全隐患。针对固定翼无人机存在的欠驱动特性与参数不确定性因素耦合问题，如何设计飞行控制器以保证闭环系统的稳定性并实现精确的轨迹跟踪，是飞行控制领域亟待解决的科学问题。

在飞控领域研究中，围绕欠驱动问题，已取得诸多令人满意的研究成果。固定翼无人机的运动通常具有六自由度，却受限于4个控制输入。文献[4]采用内外环控制策略，内环以跟踪误差为代价函数，通过求解博弈方程获得最优控制输入，实现对无人机高度和姿态角的精确跟踪，外环基于模型预测控制(model predictive control, MPC)，根据x-y方向误差生成俯仰角和滚转角的参考值，并反馈至内环以实现整体的轨迹跟踪，但该方法仅能保证内环跟踪误差收敛到零。文献[5]将无人机运动方程分解为高度和偏航角的全驱动子系统及其余状态的欠驱动子系统，全驱动部分通过比例-积分-微分控制(proportional integral derivative, PID)与滑模控制器实现快速收敛，欠驱动部分经坐标变换为级联形式后，再设计滑模控制器镇定其余4个状态量，但该方法仅能有效跟踪高度和偏航角，其余状态仅能稳定在常值，限制了其在动态轨迹跟踪等复杂场景中的应用。文献[6]针对欠驱动特性，为偏航角ϕ引入反正弦横截函数，使得偏航角误差ϕ_e与纵向误差y_e能够同时镇定，通过设计速度指令实现横向误差

收稿日期：2025-08-26；**录用日期**：2026-02-24；**网络出版时间**：2026-03-09 16:40
网络出版地址：link.cnki.net/urlid/11.2625.V.20260309.1012.001
基金项目：国家自然科学基金(62373019)；中央高校基本科研业务费专项资金(501XYGG2025103009,501XSKC2025103001)
* **通信作者**. E-mail：w.wang@buaa.edu.cn

引用格式：邸尚明，申佳军，刘烨晖，等. 基于触发式批处理辨识的固定翼无人机自适应跟踪控制[J]. 北京航空航天大学学报，2026，52(10)：3667-3676. Di S M，Shen J J，Liu Y H，et al. Batch-triggered identification based adaptive tracking control for fixed-wing UAV[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(10)：3667-3676 (in Chinese).

x_e的收敛，并以角速度指令同步达成y_e与ϕ_e的收敛，进而解决欠驱动问题。

自适应控制作为现代控制理论的核心分支之一，在处理参数不确定性方面具有广泛应用。经典的自适应控制方法大致可以分为基于Lyapunov和基于估计的两大类。前者设计自适应律，使得Lyapunov函数满足某些期望性质，如$\dot{V}(\boldsymbol{q},\hat{\boldsymbol{\theta}}) \leqslant -\kappa(\|\boldsymbol{q}\|)$，从而保证状态$\boldsymbol{q}$的收敛性，但难以实现对未知参数的准确估计，其中，$\hat{\boldsymbol{\theta}}$为未知常参数的估计值，$k>0$，表示正常数。后者根据系统输入和状态数据中蕴含的激励信息估计未知参数，可在持续激励（persistent excitation, PE）条件下保证系统状态和估计误差的收敛性，但此类激励条件与系统未来状态相关，在控制设计阶段难以确保和验证。

近年来，许多数据增强型自适应算法因能放宽激励条件而逐渐受到重视，如并行学习[7-8]、复合学习[9-11]和触发式批处理估计器[12-14]。并行学习采用激励检测算法，在特定时间间隔内收集激励信息，但由于采样机制的限制，容易出现激励遗漏的情况。复合学习和触发式批处理估计器相关成果利用积分器连续地收集激励信息，保证了激励收集的丰富性，却可能导致线性回归方程变得无界，且未考虑不同方向激励丰富度的差异。因此，如何在充分收集激励信息并兼顾其方向丰富度差异的前提下，保证线性回归方程的有界性，是数据驱动自适应控制领域面临的核心挑战。

在飞控领域中，已有许多研究采用自适应控制方法处理参数不确定性。文献[15]对无人机动力学中的不确定性进行参数化拟合，设计相应的自适应律动态调整拟合表达式中的系数，进而有效补偿模型中的不确定性。文献[16]针对无人机姿态角动力学中的不确定性，基于复合学习算法，通过积分器构造线性回归方程并引入参数更新律更新参数估计，与传统基于估计的算法相比放宽了激励条件，仅在区间激励（interval excitation, IE）下即可实现跟踪误差和参数估计误差的指数收敛。但自适应控制方法无法保证回归方程的有界性。

针对线性回归方程无界的问题，文献[17-18]根据激励矩阵的最大特征值设计遗忘因子，并将其直接引入线性回归方程两端，有效解决了回归方程无界的问题，但未能考虑不同方向激励丰富度的差异。文献[19-21]则采用定向遗忘策略，仅在激励新增方向，即信息子空间中进行遗忘，而其正交方向上的激励保持不变，确保回归方程有界的同时，保证参数估计误差的收敛性。

当固定翼无人机同时存在欠驱动特性与参数不确定性时，控制器设计需应对控制输入维度不足与不确定性动态耦合的复合约束，进而面临系统稳定性与轨迹跟踪精度难以协同保障的挑战。对此，文献[22]通过坐标变换将无人机状态分解为全驱动与欠驱动2部分，定义辅助误差将欠驱动状态误差通过耦合项引入全驱动通道，并设计低维自适应律对动力学参数不确定性进行参数化拟合。文献[23]采用分层控制策略，将原系统分解为姿态角-空速全驱动子系统与高度欠驱动子系统，通过俯仰通道补偿欠驱动特性实现高度控制，为各子系统设计反步滑模控制器，结合梯度法自适应律估计姿态角参数不确定性，同步保证系统稳定性与跟踪性能。

综上所述，固定翼无人机飞行控制器的相关成果在处理欠驱动特性与参数不确定性时，通常难以兼顾控制性能与参数估计准确性。现有的自适应控制方法往往未能充分利用激励信息处理不确定性，且参数估计的收敛性依赖于充分激励或区间激励等条件，在控制设计阶段难以先验。针对上述问题，本文提出一种复合控制设计方法，基于独轮车模型处理欠驱动问题，采用谱分解技术构造线性回归方程，并通过触发式批处理参数估计器应对参数不确定性。本文主要贡献如下：

1）与现有处理欠驱动特性的成果[4-6,22-23]相比，基于独轮车模型处理欠驱动问题，采用反步法-固定时间控制策略，使跟踪误差收敛到零，进一步提升跟踪性能。

2）与现有激励收集方案[7-21]相比，采用谱分解技术构造线性回归方程，基于激励矩阵特征值设计遗忘因子，不仅能保证回归方程的有界性，还充分考虑了不同方向激励丰富度的差异。

3）与现有自适应控制成果[15-16,22-25]相比，基于触发式批处理辨识技术处理参数不确定性，将欠驱动特性与参数不确定性解耦，在非持续部分激励条件下，通过有限次参数估计消除参数不确定性的影响。

1 系统建模与问题阐述

1.1 系统建模

本文重点研究固定翼无人机的平面运动，其运动学模型可描述为

$$\begin{cases} \dot{x} = v\cos\phi \\ \dot{y} = v\sin\phi \\ \dot{\phi} = \omega_{\mathrm{y}} \\ \dot{v} = a \end{cases} \tag{1}$$

式中：向量$\boldsymbol{p} = [x,y]^{\mathrm{T}}$表示无人机的位置坐标；$\phi$为无人机的偏航角；$v$和$\omega_{\mathrm{y}}$分别为无人机的线速度和偏

航角速度; a为无人机的线加速度。

固定翼无人机的示意图如图 1 所示, 其中, D 为空气阻力。

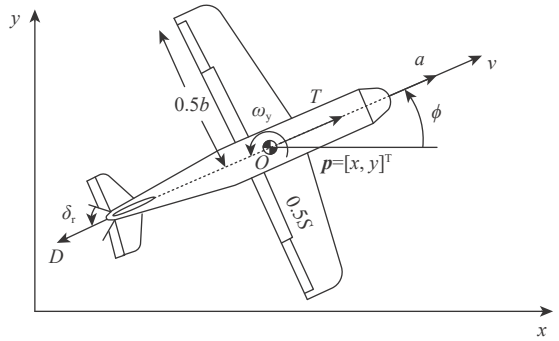


图 1　固定翼无人机示意图

Fig. 1　Fixed-wing UAV schematic diagram

固定翼无人机的动力学模型通常可分解为线运动方程和角运动方程, 如式 (2) 所示:

$$\begin{cases} m\dot{v} = T - \dfrac{1}{2}\rho_a v^2 S C_D \\ I_z\dot{\omega}_y = \dfrac{1}{2}\rho_a v^2 S b \delta_r C_n \end{cases} \tag{2}$$

式中: m为无人机的质量, kg; ρ_a为空气密度, kg/m^3; S 为机翼面积, m^2; C_D为阻力系数; I_z为无人机绕z轴的转动惯量; b为机翼展长, m; C_n为方向舵力矩系数; T和δ_r分别为发动机推力和方向舵偏角, 其与无人机控制指令$\boldsymbol{u} = [T_c, \delta_{rc}]^T$满足如下线性转换关系: $T = k_1 T_c, \delta_r = k_2 \delta_{rc}$, T_c为无人机发动机推力, δ_{rc}为方向舵偏角, k_1和k_2分别为推力系数和舵机传动系数。

本文假定k_1、k_2、C_D、I_z、C_n为未知常参数, 为处理不确定性, 对式 (2) 进行线性参数化, 可得下述动力学模型:

$$\dot{\boldsymbol{s}} = \boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})^T \boldsymbol{\theta} \tag{3}$$

式中: $\boldsymbol{s} = [v, \omega_y]^T$为状态向量; $\boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})$为非线性回归项; $\boldsymbol{\theta}$为未知参数向量。

$$\begin{cases} \boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u}) = \begin{bmatrix} -\dfrac{\rho_a v^2 S}{2m} & \dfrac{T_c}{m} & 0 \\ 0 & 0 & \dfrac{\rho_a v^2 b S \delta_{rc}}{2} \end{bmatrix}^T \\ \boldsymbol{\theta} = [C_D, k_1, k_2 C_n / I_z]^T \triangleq [\theta_1, \theta_2, \theta_3]^T \end{cases} \tag{4}$$

由于未知参数k_2、I_z、C_n以乘积的形式出现, 故将其视为一个整体, 记为θ_3。

1.2　问题阐述

固定翼无人机模型 (1)～(4) 中有 x、y、ϕ这 3 个输出, 却仅有T_c和δ_{rc}这 2 个控制输入, 属于典型的欠驱动非线性系统。此外, 系统动力学模型的固有参数不确定性也会显著影响跟踪性能和稳定性。

本文旨在为固定翼无人机设计一个自适应飞行控制器, 实现以下控制目标。

1) 无人机的位置坐标能够精确跟踪参考轨迹, 跟踪误差最终收敛到零。

2) 回归项$\boldsymbol{\psi}$的所有谱均被收集到线性回归方程 $\boldsymbol{Z}(t) = \boldsymbol{W}(t)\boldsymbol{\theta}$中, 同时保证方程的有界性, 其中, $\boldsymbol{W}(t)$表示激励矩阵。

3) 动力学模型中参数不确定性的影响将在有限次参数估计后被消除。

为消除参数不确定性的影响并解决欠驱动问题, 本文将基于以下假设和引理展开研究。

假设 1　参考轨迹 $\boldsymbol{p}_r(t)$幅值有界, 且二阶连续可导, 即存在正常数 M, 使得对于任意 $t \geqslant 0$满足: $\|\boldsymbol{p}_r(t)\| \leqslant M$, 同时 $\dot{\boldsymbol{p}}_r(t)$、$\ddot{\boldsymbol{p}}_r(t)$存在且连续。

假设 2[26]　存在一个已知集合 $B \subset \mathbf{R}$, 满足 $k_1, k_2 \in B$且$0 \notin B$。

引理 1[27]　对于一个正定变量ζ, 如果存在常数 $\varepsilon_1, \varepsilon_2 > 0, 0 < r_1 < 1, r_2 > 1$, 使其导数满足:

$$\dot{\zeta} = -\varepsilon_1 \zeta^{r_1} - \varepsilon_2 \zeta^{r_2} \tag{5}$$

则ζ将在固定时间内收敛到零, 其收敛时间的上界估计如下:

$$T_{\max} = \frac{1}{\varepsilon_1(1-r_1)} + \frac{1}{\varepsilon_2(r_2-1)} \tag{6}$$

引理 2[28]　考虑一个n维正规矩阵$\boldsymbol{W}$, 设其具有h个互异的特征值$\lambda_1, \lambda_2, \cdots, \lambda_h$, 设$\boldsymbol{e}_{i,1}, \boldsymbol{e}_{i,2}, \cdots, \boldsymbol{e}_{i,k}$为特征值$\lambda_i$的特征空间的一组单位正交基, 则谱阵$\boldsymbol{E}_i$定义为 $\boldsymbol{E}_i = \sum_{j=1}^{k} \boldsymbol{e}_{i,j}\boldsymbol{e}_{i,j}^T$, 矩阵 $\boldsymbol{W}$的谱分解形式为 $\boldsymbol{W} = \sum_{i=1}^{h} \lambda_i \boldsymbol{E}_i$, 其中, 谱阵集合$\{\boldsymbol{E}_1, \boldsymbol{E}_2, \cdots, \boldsymbol{E}_h\}$唯一且满足以下性质:

$$\sum_{i=1}^{h} \boldsymbol{E}_i = \boldsymbol{I} \tag{7}$$

$$\boldsymbol{E}_i = \boldsymbol{E}_i^T = (\boldsymbol{E}_i)^2 \qquad \forall i = 1, 2, \cdots, h \tag{8}$$

$$\boldsymbol{E}_i \boldsymbol{E}_j = 0 \qquad \forall i, j = 1, 2, \cdots, h, i \neq j \tag{9}$$

注 1　假设 1 旨在确保参考轨迹具备足够的平滑性, 这一条件在实际工程场景中易于满足。假设 2 是为了通过自适应控制设计, 规避参数估值不准确导致的控制输入奇异性问题, 相关内容将在下文展开分析。

2　控制器设计

控制器设计分为模型转换、自适应反步-固定时间控制、线性回归方程和触发式批处理参数估计器 4 个部分。控制器设计思路如图 2 所示。通过模型转换将原欠驱动系统重构为标准形式下的控制问题。采用反步法-固定时间控制分层设计运动

学虚拟控制输入与无人机实际控制指令。基于谱分解技术构造线性回归方程，引入触发式批处理参数估计算法，处理动力学模型的参数不确定性。

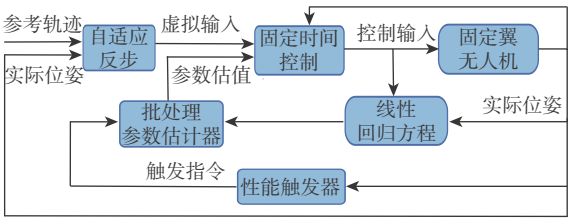


图2 固定翼无人机自适应控制器框架

Fig. 2 Fixed-wing UAV adaptive controller framework

2.1 模型转换

为处理欠驱动问题，基于独轮车模型[29]变换运动学方程。由式(1)可得：$v_x=v\cos\phi$，$v_y=v\sin\phi$，则v_x和v_y的导数计算如下：

$$\begin{cases}\dot{v}_x=\dot{v}\cos\phi-v\dot{\phi}\sin\phi\\ \dot{v}_y=\dot{v}\sin\phi+v\dot{\phi}\cos\phi\end{cases}\tag{10}$$

将$u_x=\dot{v}_x$和$u_y=\dot{v}_y$作为虚拟控制输入，则其与ω_y和a之间的转换关系如下：

$$\begin{cases}\omega_y=\dfrac{u_y\cos\phi-u_x\sin\phi}{v}\\ a=u_x\cos\phi+u_y\sin\phi\end{cases}\tag{11}$$

令无人机的速度向量为$\boldsymbol{v}=[v_x,v_y]^{\mathrm{T}}$，虚拟控制输入向量为$\boldsymbol{u}_{\mathrm{v}}=[u_x,u_y]^{\mathrm{T}}$，则变换后的运动学模型为

$$\begin{cases}\dot{\boldsymbol{p}}=\boldsymbol{v}\\ \dot{\boldsymbol{v}}=\boldsymbol{u}_{\mathrm{v}}\end{cases}\tag{12}$$

式中：虚拟控制输入$\boldsymbol{u}_{\mathrm{v}}$与$\omega_y$、$a$之间的变换关系如式(11)所示。

至此，2个控制输入T_{c}、δ_{rc}便可控制2个输出x、y，在无人机偏航角ϕ可测量的前提下，欠驱动问题得到解决。

2.2 自适应反步-固定时间控制设计

本节首先基于反步法设计无人机虚拟控制输入的期望值$\boldsymbol{u}_{\mathrm{vr}}$，进而得到角速度期望值$\omega_{\mathrm{yr}}$与加速度期望值$a_{\mathrm{r}}$，再以二者为辅助变量进行固定时间控制设计，最终求解出无人机实际控制输入$\boldsymbol{u}$。

根据反步法：

步骤1 定义跟踪误差：

$$\begin{cases}z_1=\boldsymbol{p}-\boldsymbol{p}_{\mathrm{r}}\\ z_2=\boldsymbol{v}-\boldsymbol{\alpha}\end{cases}\tag{13}$$

式中：$\boldsymbol{p}_{\mathrm{r}}$为$\boldsymbol{p}$的参考值；$\boldsymbol{\alpha}$为镇定函数。

对z_1求导：

$$\dot{z}_1=\boldsymbol{v}-\dot{\boldsymbol{p}}_{\mathrm{r}}\tag{14}$$

定义Lyapunov函数$V_1=\dfrac{1}{2}z_1^{\mathrm{T}}z_1$，对$V_1$求导数：

$$\dot{V}_1=z_1^{\mathrm{T}}(z_2+\boldsymbol{\alpha}-\dot{\boldsymbol{p}}_{\mathrm{r}})\tag{15}$$

为镇定z_1，令

$$\boldsymbol{\alpha}=-c_1z_1+\dot{\boldsymbol{p}}_{\mathrm{r}}\tag{16}$$

式中：c_1为正常数。

由式(15)可知，此时$\dot{V}_1$满足：

$$\dot{V}_1=z_1^{\mathrm{T}}z_2-c_1z_1^{\mathrm{T}}z_1\tag{17}$$

步骤2 对z_2求导，由式(13)可得

$$\dot{z}_2=\boldsymbol{u}_{\mathrm{v}}+c_1\dot{\boldsymbol{p}}-c_1\dot{\boldsymbol{p}}_{\mathrm{r}}-\ddot{\boldsymbol{p}}_{\mathrm{r}}\tag{18}$$

定义Lyapunov函数$V_2=V_1+\dfrac{1}{2}z_2^{\mathrm{T}}z_2$，对$V_2$求导：

$$\dot{V}_2=z_1^{\mathrm{T}}z_2-c_1z_1^{\mathrm{T}}z_1+z_2^{\mathrm{T}}(\boldsymbol{u}_{\mathrm{v}}+c_1\dot{\boldsymbol{p}}-c_1\dot{\boldsymbol{p}}_{\mathrm{r}}-\ddot{\boldsymbol{p}}_{\mathrm{r}})\tag{19}$$

设计虚拟控制输入$\boldsymbol{u}_{\mathrm{v}}$的期望值$\boldsymbol{u}_{\mathrm{vr}}$为

$$\boldsymbol{u}_{\mathrm{vr}}=-c_2z_2-z_1-c_1\dot{\boldsymbol{p}}+c_1\dot{\boldsymbol{p}}_{\mathrm{r}}+\ddot{\boldsymbol{p}}_{\mathrm{r}}\tag{20}$$

式中：c_2为正常数。

当$\boldsymbol{u}_{\mathrm{v}}$收敛于其期望值$\boldsymbol{u}_{\mathrm{vr}}$时，$\dot{V}_2$满足：

$$\dot{V}_2=-c_1z_1^{\mathrm{T}}z_1-c_2z_2^{\mathrm{T}}z_2\tag{21}$$

将$\boldsymbol{u}_{\mathrm{vr}}=[u_{\mathrm{xr}},u_{\mathrm{yr}}]^{\mathrm{T}}$代入式(11)可得偏航角速度和线加速度的期望值，记为ω_{yr}和a_{r}，表示为

$$\begin{cases}\omega_{\mathrm{yr}}=\dfrac{u_{\mathrm{yr}}\cos\phi-u_{\mathrm{xr}}\sin\phi}{v}\\ a_{\mathrm{r}}=u_{\mathrm{xr}}\cos\phi+u_{\mathrm{yr}}\sin\phi\end{cases}\tag{22}$$

由于固定翼无人机在飞行过程中，速度始终不为零，因此式(22)中角速度的期望值不会出现奇异。接下来将以ω_{yr}和a_{r}为辅助变量进行固定时间控制设计。

为使实际偏航角速度跟踪期望值，定义误差：

$$z_3=\omega_{\mathrm{y}}-\omega_{\mathrm{yr}}\tag{23}$$

定义Lyapunov函数$V_3=\dfrac{1}{2}z_3^2$，对V_3求导：

$$\dot{V}_3=z_3\dot{z}_3=z_3(\alpha_{\mathrm{y}}-\dot{\omega}_{\mathrm{yr}})\tag{24}$$

式中：α_{y}为偏航角加速度。

根据引理1，将α_{y}的期望值α_{yr}设计为

$$\alpha_{\mathrm{yr}}=-c_3\mathrm{sgn}(z_3)|z_3|^{\beta_1}-c_4\mathrm{sgn}(z_3)|z_3|^{\beta_2}+\dot{\omega}_{\mathrm{yr}}\tag{25}$$

式中：$c_3,c_4>0,0<\beta_1<1,\beta_2>1$，均为正常数。

无人机的实际控制输入$\boldsymbol{u}=[T_{\mathrm{c}},\delta_{\mathrm{rc}}]^{\mathrm{T}}$可通过对式(3)～式(4)进行变换并结合式(22)与式(25)获得：

$$\begin{bmatrix}T_{\mathrm{c}}\\ \delta_{\mathrm{rc}}\end{bmatrix}=\begin{bmatrix}\dfrac{2ma_{\mathrm{r}}+\rho_{\mathrm{a}}v^2S\hat{\theta}_1}{2\hat{\theta}_2}\\ \dfrac{2\alpha_{\mathrm{yr}}}{\rho_{\mathrm{a}}v^2bS\hat{\theta}_3}\end{bmatrix}\tag{26}$$

式中：未知参数估值$\hat{\boldsymbol{\theta}}=[\hat{\theta}_1,\hat{\theta}_2,\hat{\theta}_3]^{\mathrm{T}}$将由触发式批处理参数估计器得到。

在参数估计准确的前提下, 若实际偏航角加速度收敛于其期望值, 由式 (24) 可知$\dot{V}_3$满足:

$$\dot{V}_3 = -c_3|z_3|^{\beta_1+1} - c_4|z_3|^{\beta_2+1} = -2c_3 V_3^{\frac{\beta_1+1}{2}} - 2c_4 V_3^{\frac{\beta_2+1}{2}} \tag{27}$$

即式 (25) 可使V_3在固定时间内收敛到零。

2.3　线性回归方程

根据系统辨识与最小二乘理论, 式 (3) 中的回归项$\boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})$可以看作激励信息的载体。本文通过激励矩阵$\boldsymbol{W}$收集$\boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})$中的谱信息[11]。为便于分析, 暂定$\boldsymbol{W}$为半正定矩阵, 后文将证明这一性质。

为完成参数估计, 首先应为方程 (3) 构造线性回归方程:

$$\boldsymbol{Z}(t) = \boldsymbol{W}(t)\boldsymbol{\theta} \tag{28}$$

为确保线性回归方程的有界性, 同时考虑不同方向激励丰富度的差异, 本文将采用谱分解法实现上述目标。根据引理 2, 激励矩阵$\boldsymbol{W}$的谱分解形式可以改写为

$$\boldsymbol{W}(t) = \sum_{i=1}^{h(t)} \lambda_i(t)\boldsymbol{E}_i(t) \tag{29}$$

根据式 (9) 可知谱阵$\boldsymbol{E}_i(i=1,2,\cdots,h(t))$彼此正交, 因此, 谱阵可以表示先前收集激励信息的$h(t)$个不同激励方向。特征值$\lambda_i(t)(i=1,2,\cdots,h(t))$表示相应激励方向上收集的激励丰富度。

根据$\boldsymbol{W}$的分解式 (29) 以及式 (7), t时刻的线性回归方程可以改写为

$$\boldsymbol{Z}(t) = \sum_{i=1}^{h(t)} \boldsymbol{E}_i(t)\boldsymbol{Z}(t) = \sum_{i=1}^{h(t)} \lambda_i(t)\boldsymbol{E}_i(t)\boldsymbol{\theta} \tag{30}$$

式 (9) 表明谱阵$\boldsymbol{E}_i(t)(i=1,2,\cdots,h(t))$彼此正交, 因此, 式 (30) 可以分解为$h(t)$个不同的激励方向, 即

$$\boldsymbol{E}_i(t)\boldsymbol{Z}(t) = \lambda_i(t)\boldsymbol{E}_i(t)\boldsymbol{\theta} \qquad i=1,2,\cdots,h(t) \tag{31}$$

对式 (3) 两边同时左乘$\boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})$:

$$\boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})\dot{\boldsymbol{s}} = \boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})\boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})^{\mathrm{T}}\boldsymbol{\theta} \tag{32}$$

结合式 (32) 与分解后的线性回归方程 (31), 可以构造以下方程:

$$\boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})\dot{\boldsymbol{s}} - \sum_{i=1}^{h(t)} \rho(\lambda_i,\boldsymbol{s})\boldsymbol{E}_i(t)\boldsymbol{Z}(t) = \left[\boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})\boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})^{\mathrm{T}} - \sum_{i=1}^{h(t)} \rho(\lambda_i,\boldsymbol{s})\lambda_i(t)\boldsymbol{E}_i(t)\right]\boldsymbol{\theta} \tag{33}$$

式中: ρ为激励收集的遗忘因子, 定义如下:

$$\rho(\lambda_i,\boldsymbol{s}) = \begin{cases} 0 & \lambda_i \leqslant \sigma_{\min} \\ \dfrac{\lambda_{\max}(\boldsymbol{\psi}\boldsymbol{\psi}^{\mathrm{T}})}{2\lambda_i}\left\{\mathrm{sat}\left(\dfrac{\lambda_i-\mu}{\eta}\right)+1\right\} & \text{其他} \end{cases} \tag{34}$$

其中: $\lambda_{\max}(\boldsymbol{\psi}\boldsymbol{\psi}^{\mathrm{T}})$为矩阵$\boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})\boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})^{\mathrm{T}}$的最大特征值; $\sigma_{\max}$和$\sigma_{\min}$分别对应$\rho(\lambda_i,\boldsymbol{s})=0$的最大特征值和$\rho(\lambda_i,\boldsymbol{s})=\lambda_{\max}(\boldsymbol{\psi}\boldsymbol{\psi}^{\mathrm{T}})/\lambda_i$的最小特征值; 常数$\mu=(\sigma_{\max}+\sigma_{\min})/2$和$\eta=(\sigma_{\max}-\sigma_{\min})/2$分别为饱和函数sat的中心和宽度。饱和函数定义为

$$\mathrm{sat}(x) = \begin{cases} -1 & x \leqslant -1 \\ x & -1 < x < 1 \\ 1 & x \geqslant 1 \end{cases} \tag{35}$$

在时间区间$[0,t]$内对式 (33) 进行积分, 可得线性回归方程$\boldsymbol{Z}(t)=\boldsymbol{W}(t)\boldsymbol{\theta}$, 其中, 向量$\boldsymbol{Z}(t)$和矩阵$\boldsymbol{W}(t)$可通过下述常微分方程获得:

$$\dot{\boldsymbol{Z}} = \boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})\dot{\boldsymbol{s}} - \sum_{i=1}^{h(t)} \rho(\lambda_i,\boldsymbol{s})\boldsymbol{E}_i(t)\boldsymbol{Z}(t) \tag{36}$$

$$\dot{\boldsymbol{W}} = \boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})\boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})^{\mathrm{T}} - \sum_{i=1}^{h(t)} \rho(\lambda_i,\boldsymbol{s})\lambda_i(t)\boldsymbol{E}_i(t) \tag{37}$$

式中: 方程的初始条件为$\boldsymbol{Z}(0)=\boldsymbol{0}, \boldsymbol{W}(0)=\boldsymbol{0}$。

微分方程中的第 1 项收集了当前的激励信息, 添加第 2 项是为了保证矩阵$\boldsymbol{Z}(t)$和$\boldsymbol{W}(t)$的有界性。根据遗忘因子 (34) 和矩阵$\boldsymbol{W}$的定义 (37), 属于$\lambda_i \leqslant \sigma_{\min}$方向的信息不会被遗忘, 而对应于$\lambda_i > \sigma_{\min}$方向的信息将以$\rho(\lambda_i,\boldsymbol{s})$的速率被遗忘。

线性回归方程 (28) 的性质总结为定理 1, 因篇幅所限, 其详细证明可参见文献 [11]。

定理 1[11]　考虑具有参数不确定性的动力学模型 (3)～(4) 以及构造的线性回归方程 (28), 以下结论成立:

1) 激励矩阵$\boldsymbol{W}(t)$在$t \geqslant 0$时半正定。

2) 如果回归项$\boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})$有界, 那么矩阵$\boldsymbol{W}$特征值有界, 其上界为$\sigma_{\max}$。

3) 对于任意时刻$\tau \in [0,t]$, $\boldsymbol{\psi}[\boldsymbol{s}(\tau),\boldsymbol{u}(\tau)]$的列空间是矩阵$\boldsymbol{W}(t)$列空间的子集。

4) 给定三维常向量$\boldsymbol{\varphi}$, $\boldsymbol{W}(t)\boldsymbol{\varphi}=\boldsymbol{0}$等价于所有$\tau \in [0,t]$, $\boldsymbol{\psi}[\boldsymbol{s}(\tau),\boldsymbol{u}(\tau)]^{\mathrm{T}}\boldsymbol{\varphi}=\boldsymbol{0}$。

根据定理 1, 先前出现的所有激励信息均被收集到矩阵$\boldsymbol{W}$中, 由于矩阵$\boldsymbol{W}$与回归项$\boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})$具有相同的零空间, 这表明估值误差$\tilde{\boldsymbol{\theta}}=\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}$在零空间$N[\boldsymbol{W}(t)]$上的投影在时间区间$[0,t]$内从未受到激励, 因此, 零空间$N[\boldsymbol{W}(t)]$可以看作参数估计误差的非受激励空间。由于矩阵$\boldsymbol{W}(t)$对称, 且列空间$R[\boldsymbol{W}(t)]$是零空间$N[\boldsymbol{W}(t)]$的正交补, 故列空间$R[\boldsymbol{W}(t)]$可以看作参数估计误差的受激励空间。

同时, 矩阵$\boldsymbol{W}$的秩$\mathrm{rank}[\boldsymbol{W}(t)]$增加的一些新的激励信息被收集到矩阵$\boldsymbol{W}$中。相应地, 受激励空间维数增加, 非受激励空间维数减少。

注 2　与现有激励收集方案[7-21]相比, 本文基

于谱分解定理构建线性回归方程，定量收集激励信息，引入遗忘因子，考虑了不同方向激励丰富度的差异，使得当强激励方向的信息被遗忘时，弱激励方向不受影响，同时保证了线性回归方程的有界性。

2.4 触发式批处理参数估计器

由式(26)可知，当未知参数估值$\hat{\theta}_2$、$\hat{\theta}_3$接近0时，控制输入将产生奇异，为解决这一问题，在假设2的前提下，根据触发式批处理辨识相关理论[14]，可得下述参数更新律：

$$\begin{cases}\hat{\boldsymbol{\theta}}(\tau_i)=\underset{\boldsymbol{\vartheta}\in\mathbf{R}\times\mathbf{B}^2}{\arg\min}\left\{\left\|\boldsymbol{\vartheta}-\hat{\boldsymbol{\theta}}(\tau_{i-1})\right\|^2\right\}\\ \text{s.t.}\quad \boldsymbol{Z}(\tau_i)=\boldsymbol{W}(\tau_i)\boldsymbol{\vartheta}\end{cases}\tag{38}$$

参数估值$\hat{\boldsymbol{\theta}}$在时刻$\tau_i$利用区间$(\tau_0,\tau_i)$内收集的激励更新，并在任意2个连续更新时刻之间保持不变：

$$\hat{\boldsymbol{\theta}}(t)=\hat{\boldsymbol{\theta}}(\tau_i)\qquad\forall t\in(\tau_i,\tau_{i+1})\tag{39}$$

根据假设2中集合B的约束以及参数更新律(38)，$\hat{\theta}_2$、$\hat{\theta}_3$在任意触发时刻均不会等于0，从而避免了控制输入奇异。此时，参数估计$\hat{\boldsymbol{\theta}}(\tau_i)$不再是$\hat{\boldsymbol{\theta}}(\tau_{i-1})$在线性方程(28)定义的超平面上的投影。更新后的参数估计$\hat{\boldsymbol{\theta}}(\tau_i)$应该是在集合$B$与式(28)定义的超平面交集中，与先前参数估计$\hat{\boldsymbol{\theta}}(\tau_{i-1})$距离最近的点。

由于Lyapunov函数能够有效反映控制性能，因此可用于设计调节触发条件，当其达到预设的触发阈值时，应激活批处理参数估计器。此外，当收集到的激励信息明显比上一次触发更丰富时，也应更新参数估计。故调节触发条件如下：该条件决定激活参数更新律(38)的时刻：

$$\tau_i=\min_{t>\tau_{i-1}}\left\{t=\tau_{i-1}+\Delta\tau,\text{or}V_3(t)=\chi_i,\text{or}K_{\mathrm{W}}^{l}(t)>K_{\mathrm{W}}^{l}(\tau_{i-1})\right.\tag{40}$$

式中：$\Delta\tau>0$为最大触发间隔；V_3为固定时间控制Lyapunov函数；$K_{\mathrm{W}}^{l}(t)$为激励矩阵特征值中大于l的数量；触发阈值χ_i设计如下：

$$\chi_i=\gamma_1 V_3(\tau_{i-1})+\gamma_2\qquad\forall t\in(\tau_{i-1},\tau_i)\tag{41}$$

式中：$\gamma_1\geqslant 1$，$\gamma_2>0$为待设计的常参数，这样能严格保证$\chi_i>V_3(\tau_{i-1})$。初始触发时刻设计为$\tau_0=0\,\mathrm{s}$，对于任意正整数i，触发阈值χ_i在区间(τ_{i-1},τ_i)内保持不变，并将在下一个触发时刻τ_i发生跳变。至此，完成了参数更新律的设计。

3 稳定性分析

定理2 考虑由式(3)～式(4)、式(11)～式(12)给定的系统模型，当具有控制输入(26)与参数更新律(38)～(40)，并满足假设1和假设2时，以下结论成立：

1）参数估计$\hat{\boldsymbol{\theta}}$的更新次数不超过3次。

2）设τ_k为参数估值$\hat{\boldsymbol{\theta}}$的最后更新时刻，则对$\forall t\geqslant\tau_k$，未知参数$\boldsymbol{\theta}$的影响将被完全消除，即$\boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})^{\mathrm{T}}\tilde{\boldsymbol{\theta}}\equiv\boldsymbol{0},\forall t\geqslant\tau_k$成立。

3）闭环系统稳定，跟踪误差最终收敛到零。

4）系统不会产生芝诺行为(有限时间内触发无限次参数更新)。

证明 1）设$R[\boldsymbol{W}(\tau_i)]$和$N[\boldsymbol{W}(\tau_i)]$分别表示矩阵$\boldsymbol{W}$在$\tau_i$时刻的列空间和零空间，当条件(38)触发并且$\tilde{\boldsymbol{\theta}}(\tau_i)\notin N[\boldsymbol{W}(\tau_{i+1})]$时，应该更新参数估计来确保$\tilde{\boldsymbol{\theta}}(\tau_{i+1})\in N[\boldsymbol{W}(\tau_{i+1})]$，由于在相邻触发时刻$(\tau_i,\tau_{i+1})$内存在新的激励信息，即矩阵$\boldsymbol{W}(t)$的秩增加，故$\dim R[\boldsymbol{W}(\tau_{i+1})]\geqslant\dim R[\boldsymbol{W}(\tau_i)]+1$，其中，dim表示线性空间的维数。相应地，受激励空间维数增加，非受激励空间维数减少，因此可进一步得到$\dim N[\boldsymbol{W}(\tau_{i+1})]\leqslant\dim N[\boldsymbol{W}(\tau_i)]-1$。根据单调有界定理，$\dim N[\boldsymbol{W}(\tau_i)]$最终将减小到一个非负整数，由$\dim N[\boldsymbol{W}(0)]=3$，可知参数估计$\hat{\boldsymbol{\theta}}$的更新次数不会超过3次。

2）此外，$\boldsymbol{\psi}(\boldsymbol{s},\boldsymbol{u})^{\mathrm{T}}\tilde{\boldsymbol{\theta}}\equiv\boldsymbol{0},\forall t\geqslant\tau_k$显然成立，否则$\hat{\boldsymbol{\theta}}$必须根据新的激励信息继续更新，这与$\tau_k$为最后更新时刻矛盾，矛盾表明上式成立，未知参数的影响将被完全消除。

3）采用控制器(26)，由于参数不确定性的影响将在至多3次更新后被消除，因此，根据式(27)可知V_3将在固定时间内收敛到零，即误差z_3收敛到零，进而虚拟控制输入实际值u_x、u_y将收敛到其期望值$u_{x\mathrm{r}}$、$u_{y\mathrm{r}}$，故式(21)成立，即闭环系统稳定，跟踪误差z_1、z_2指数收敛到零。

4）根据触发条件(40)，参数更新律(38)的触发情况可以分为以下3种类型：① 控制器性能触发，即$V_3(\tau_i)=\chi_i$；② 周期性触发，即$\tau_i=\tau_{i-1}+\Delta\tau$；③ 激励丰富性触发，即$K_{\mathrm{W}}^{l}(\tau_i)>K_{\mathrm{W}}^{l}(\tau_{i-1})$。上述3种类型中，类型①和类型③引起的触发，表明在区间(τ_{i-1},τ_i)内存在新激励信息，此时$\boldsymbol{\psi}(\boldsymbol{s}(\tau_{i-1}),\boldsymbol{u}(\tau_{i-1}))^{\mathrm{T}}\tilde{\boldsymbol{\theta}}(\tau_{i-1})\neq\boldsymbol{0}$，参数估计$\hat{\boldsymbol{\theta}}$必定被更新，以确保$\boldsymbol{\psi}(\boldsymbol{s}(\tau_i),\boldsymbol{u}(\tau_i))^{\mathrm{T}}\tilde{\boldsymbol{\theta}}(\tau_i)=\boldsymbol{0}$。当在区间$(\tau_{i-1},\tau_i)$内不存在新激励信息时，将由类型②引起触发，此时参数估计$\hat{\boldsymbol{\theta}}$保持不变。

根据上述分析，类型①和类型③引起的触发必定会使参数估计得到更新，根据定理2的第1条结论可知，参数估计的更新次数不会超过3次，故二者所引起的触发次数也不会超过3次；而类型②引起的估计器触发最小时间间隔为$\Delta\tau$。因此，以上

3 种类型均不会导致有限时间内触发无限次参数更新, 也即排除了触发条件 (40) 的芝诺行为。　证毕

注 3　与传统自适应控制方法相比, 本文采用触发式批处理辨识技术来处理参数不确定性, 在非持续部分激励条件下消除了未知参数的影响, 实现了对未知参数受激励分量的准确估计。该方法可拓展至智能汽车自适应巡航、多无人系统协同定位等控制场景。

注 4　本文方法在激励收集方式与自适应控制方法方面, 与现有研究成果的对比分别如表 1 和表 2 所示。

表 1　激励收集方式对比

Table 1　Comparison of excitation collection methods

激励收集方式	激励信息是否丰富	回归方程是否有界
积分器[7-16]	是	否
积分器+一般遗忘[17-18]	否	是
积分器+定向遗忘[19-21]	否	是
谱分解+定向遗忘(本文)	是	是

表 2　自适应控制方法对比

Table 2　Comparison of adaptive control methods

自适应控制方法	所需激励条件	参数更新律连续性
并行学习[7-8]	弱持续激励	连续
复合学习[9-11]	弱持续激励、充分激励 非持续部分激励	连续
触发式批处理[12-14](本文)	非持续部分激励	离散

4　仿真结果

为证明本文控制方案的有效性, 将提供数值仿真的结果。

仿真过程的参数选择如下: 控制器参数 $c_1=0.8, c_2=0.8, c_3=1, c_4=1, \beta_1=0.6, \beta_2=1.2$, 触发条件缩放系数 $\gamma_1=5, \gamma_2=1.5$, 遗忘因子预设上下界 $\sigma_{\min}=3, \sigma_{\max}=8$, 饱和函数上下界为 ±1, 未知参数真值 $\boldsymbol{\theta}=(0.015, 4.0, -0.3)^{\mathrm{T}}$, 未知参数估值初值为 $\hat{\boldsymbol{\theta}}(0)=(0.03, 3.0, -0.5)^{\mathrm{T}}$, 触发参数更新律的最小时间间隔 $\Delta\tau=3$ s, 无人机质量 $m=0.55$ kg, 空气密度 $\rho_{\mathrm{a}}=1.225\ \mathrm{kg/m^3}$, 机翼面积 $S=0.21\ \mathrm{m^2}$, 机翼展长 $b=1.37$ m。

给定期望轨迹 $x_{\mathrm{r}}=5-10\cos(0.7t)$, $y_{\mathrm{r}}=-10\sin(0.7t)$, 实验效果如图 3 所示。可以看到, 初始时刻二者存在偏差, 随着无人机的运动, 二者逐渐重合。从图 4 的误差曲线中同样可以看出, 跟踪误差呈现逐渐收敛的趋势, 其中跟踪误差定义为: $x_{\mathrm{error}}=x-x_{\mathrm{r}}, y_{\mathrm{error}}=y-y_{\mathrm{r}}$。

控制输入(油门推力 T_{c} 和方向舵偏角 δ_{rc})经限幅以后的变化曲线如图 5 所示。二者均为归一化后的值, 其中, $T_{\mathrm{c}}\in[0,1]$, $\delta_{\mathrm{rc}}\in[-1,1]$, 可知稳态后二者均处于合理范围, 且限幅并未影响系统的稳定性与跟踪精度。

未知参数估值 $\hat{\boldsymbol{\theta}}$ 与真值 $\boldsymbol{\theta}$ 曲线如图 6 所示, 如第 3 节所述, 3 个参数估值均在 3 次更新以后等于其真值。

无人机速度曲线如图 7 所示, 其稳态值约为

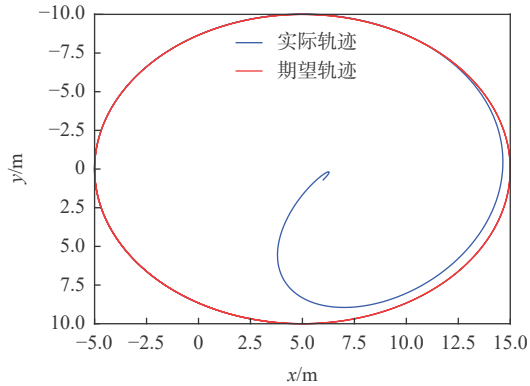


图 3　实际轨迹与期望轨迹

Fig. 3　Real trajectory and desired trajectory

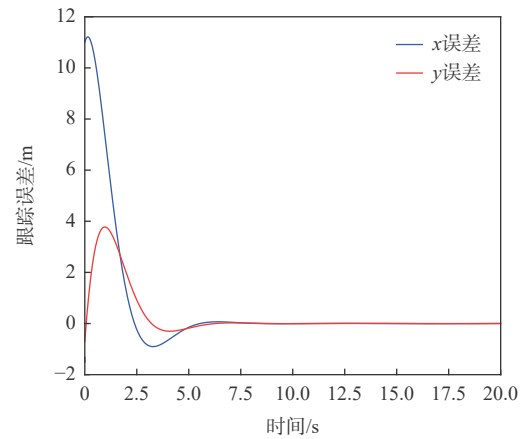


图 4　x、y方向的跟踪误差

Fig. 4　Tracking errors in the x and y directions

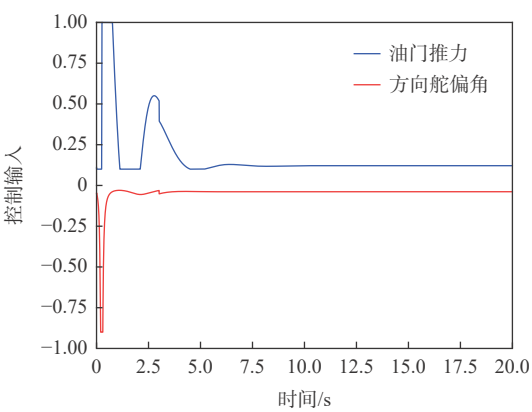


图 5　控制输入曲线

Fig. 5　Control input curves

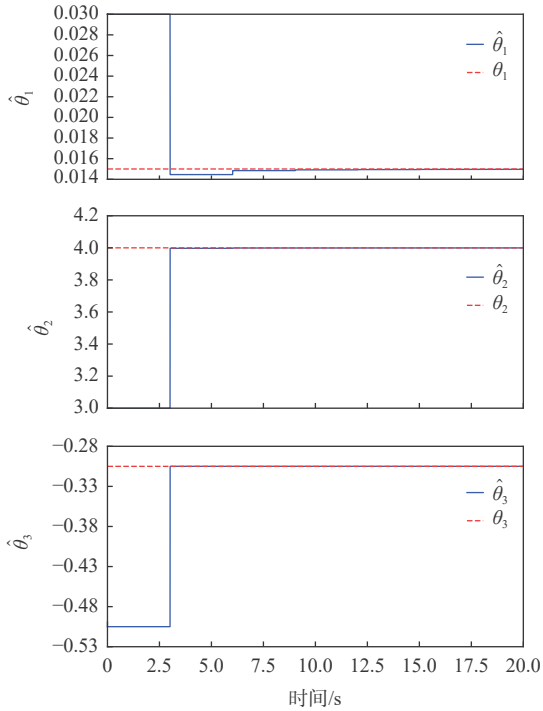


图 6 参数估计$\hat{\boldsymbol{\theta}}$

Fig. 6 Parameter estimate $\hat{\boldsymbol{\theta}}$

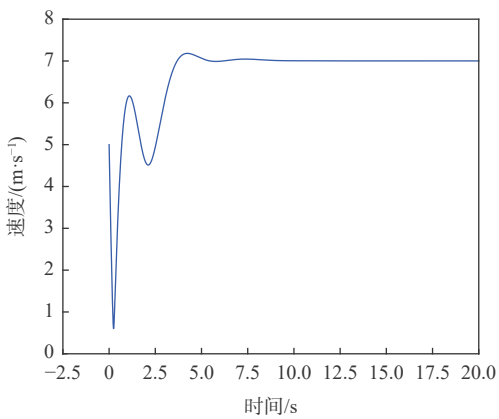


图 7 无人机速度曲线

Fig. 7 UAV velocity curve

7 m/s, 同样符合轨迹要求。

仿真结果表明, 本文控制方法较好地完成了轨迹跟踪任务, 同时消除了参数不确定性的影响, 证明了方法的有效性。

5 结 论

1) 本文针对固定翼无人机欠驱动特性与模型参数不确定性, 通过模型转换结合反步法与固定时间控制, 分层设计运动学虚拟输入与实际指令。

2) 利用谱分解技术构建线性回归方程, 引入触发式批处理参数估计器处理参数不确定性, 在保障方程有界性的同时, 通过有限次更新消除估计误差的受激励分量, 确保系统稳定并避免芝诺行为。

未来研究将聚焦于外部干扰下, 无人机的跟踪与镇定控制, 并尝试将所提框架拓展至多智能体系统。

参考文献 (References)

[1] Zhuang X, Li D G, Wang Y, et al. Optimization of high-speed fixed-wing UAV penetration strategy based on deep reinforcement learning[J]. Aerospace Science and Technology, 2024, 148: 109089.

[2] Nugroho G, Taha Z, Nugraha T S, et al. Development of a fixed wing unmanned aerial vehicle (UAV) for disaster area monitoring and mapping[J]. Journal of Mechatronics, Electrical Power, and Vehicular Technology, 2015, 6(2): 83-88.

[3] Hinzmann T, Schneider T, Dymczyk M, et al. Robust map generation for fixed-wing UAVs with low-cost highly-oblique monocular cameras[C]//2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Piscataway: IEEE Press, 2016: 3261-3268.

[4] Raffo G V, Ortega M G, Rubio F R. Path tracking of a UAV via an underactuated control strategy[J]. European Journal of Control, 2011, 17(2): 194-213.

[5] Wang Z, Bao W M, Li H F. Second-order dynamic sliding-mode control for nonminimum phase underactuated hypersonic vehicles [J]. IEEE Transactions on Industrial Electronics, 2017, 64(4): 3105-3112.

[6] Huang J S, Wen C Y, Wang W, et al. Adaptive output feedback tracking control of a nonholonomic mobile robot[J]. Automatica (Journal of IFAC), 2014, 50(3): 821-831.

[7] Long J, Guo Y M, Liu Z, et al. Filtering-based concurrent learning adaptive attitude tracking control of rigid spacecraft with inertia parameter identification[J]. International Journal of Robust and Nonlinear Control, 2023, 33(8): 4562-4576.

[8] Parikh A, Kamalapurkar R, Dixon W E. Integral concurrent learning: adaptive control with parameter convergence using finite excitation[J]. International Journal of Adaptive Control and Signal Processing, 2019, 33(12): 1775-1787.

[9] Guo K, Pan Y P, Zheng D D, et al. Composite learning control of robotic systems: a least squares modulated approach[J]. Automatica (Journal of IFAC), 2020, 111: 108612.

[10] Shen J J, Liu Y H, Wang W, et al. Composite learning adaptive safety critical control with application to adaptive cruise of Intelligent vehicles[J]. IEEE Transactions on Industrial Electronics, 2025, 72(10): 10793-10803.

[11] Shen J J, Wang W, Wen C Y, et al. Composite learning adaptive control under non-persistent partial excitation[EB/OL]. (2024-08-11)[2025-08-26]. https://arxiv.org/abs/2408.01731.

[12] Karafyllis I, Krstic M. Adaptive certainty-equivalence control with regulation-triggered finite-time least-squares identification[J]. IEEE Transactions on Automatic Control, 2018, 63(10): 3261-3275.

[13] Shen J J, Wang W, Zhou J, et al. Adaptive safety with control barrier functions and triggered batch least-squares identifier[J]. Automat-

ica (Journal of IFAC), 2025, 173: 112059.

[14] Karafyllis I, Kontorinaki M, Krstic M. Adaptive control by regulation-triggered batch least squares[J]. IEEE Transactions on Automatic Control, 2020, 65(7): 2842-2855.

[15] Roshanian J, Rahimzade E. Model reference adaptive control using a new adaptation law based on power-series quadratic lyapunov functions for the single degree of freedom wing rock phenomenon[J]. Journal of Aerospace Science and Technology, 2021, 14(1): 11-29.

[16] Xu B. Composite learning finite-time control with application to quadrotors[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2018, 48(10): 1806-1815.

[17] Cho N, Shin H S, Kim Y, et al. Composite model reference adaptive control with parameter convergence under finite excitation[J]. IEEE Transactions on Automatic Control, 2018, 63(3): 811-818.

[18] Pan Y P, Aranovskiy S, Bobtsov A, et al. Efficient learning from adaptive control under sufficient excitation[J]. International Journal of Robust and Nonlinear Control, 2019, 29(10): 3111-3124.

[19] Lee H I, Shin H S, Tsourdos A. Concurrent learning adaptive control with directional forgetting[J]. IEEE Transactions on Automatic Control, 2019, 64(12): 5164-5170.

[20] Goel A, Bruce A L, Bernstein D S. Recursive least squares with variable-direction forgetting: compensating for the loss of persistency [lecture notes][J]. IEEE Control Systems Magazine, 2020, 40(4): 80-102.

[21] Lai B, Bernstein D S. SIFt-RLS: subspace of information forgetting recursive least squares[EB/OL]. (2024-04-16)[2025-08-26]. https://arxiv.org/abs/2404.10844.

[22] Baldi S, Roy S, Yang K, et al. An underactuated control system design for adaptive autopilot of fixed-wing drones[J]. IEEE/ASME Transactions on Mechatronics, 2022, 27(5): 4045-4056.

[23] Bao C C, Guo Y F, Luo L R, et al. Design of a fixed-wing UAV controller based on adaptive backstepping sliding mode control method[J]. IEEE Access, 2021, 9: 157825-157841.

[24] Schlanbusch S M, Zhou J, Schlanbusch R. Adaptive attitude control of a rigid body with input and output quantization[J]. IEEE Transactions on Industrial Electronics, 2022, 69(8): 8296-8305.

[25] Fjetland A K, Zhou J, Abeyrathna D, et al. Kick detection and influx size estimation during offshore drilling operations using deep learning[C]//2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA). Piscataway: IEEE Press, 2019: 2321-2326.

[26] Shen J J, Wang W, Wu X K, et al. Regulation-triggered batch identifier-based adaptive control for nonlinear systems with unknown control coefficients[J]. Journal of Control and Decision, 2026, 13(5): 1324-1339.

[27] Bhat S P, Bernstein D S. Finite-time stability of continuous autonomous systems[J]. SIAM Journal on Control and Optimization, 2000, 38(3): 751-766.

[28] Lancaster P, Tismenetsky M. The theory of matrices[M]. 2nd ed. San Diego: Academic Press, 1985.

[29] De Luca A, Oriolo G, Vendittelli M. Stabilization of the unicycle via dynamic feedback linearization[J]. IFAC Proceedings Volumes, 2000, 33(27): 687-692.

Batch-triggered identification based adaptive tracking control for fixed-wing UAV

DI Shangming[1]，SHEN Jiajun[1]，LIU Yehui[1]，WANG Wei[1, 2, *]

(1. School of Automation Science and Electrical Engineering，Beihang University，Beijing 100191，China；
2. Hangzhou Innovation Institute，Beihang University，Hangzhou 310051，China)

Abstract：In this paper, an adaptive hierarchical control method based on a regulation-triggered batch identifier is proposed. This approach addresses the challenges in stability analysis and control design of fixed-wing UAVs, which are complicated by the coupling of underactuated characteristics and parametric uncertainty. In order to deal with the underactuated problem, the original fixed-wing UAV model is transformed into a virtual integrator and linear parameterized model based on the unicycle model. Backstepping is then combined with fixed-time control methods to create an equivalent flight tracking controller. By analyzing the mechanism of parameter uncertainty on the motion state of the UAV, a linear regression equation is constructed by spectral decomposition technology to quantitatively collect excitation information, which ensures the richness of the collected excitation information and the boundedness of the linear regression equation. A regulation-triggered batch identifier is created to update the parameter estimation using the gathered excitation data when the trigger function satisfies specific requirements in order to remove the impact of parametric uncertainty. Compared with the existing hierarchical control strategies for fixed-wing UAVs, the proposed framework achieves the unity of error convergence, excitement richness and parameter identification accuracy.

Keywords：data-driven control；adaptive control；fixed-time control；fixed-wing UAV；underactuated characteristics

Received：2025-08-26；**Accepted**：2026-02-24；**Published Online**：2026-03-09 16：40

URL：link.cnki.net/urlid/11.2625.V.20260309.1012.001

Foundation items：National Natural Science Foundation of China (62373019); The Fundamental Research Funds for the Central Universities (501XYGG2025103009, 501XSKC2025103001)

*** Corresponding author**. E-mail：w.wang@buaa.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西配楼412~416房间
电　　话:(010)82316698,82316981,82315499,82317448
E-mail: jbuaa@ buaa. edu. cn
http://bhxb. buaa. edu. cn
http://www. buaa. edu. cn

IISTJA 工信科技期刊集群 集群成员期刊

《北京航空航天大学学报》
第五届编辑委员会

主　任（主　编）：赵沁平

（以下按姓氏笔画为序）

副主任（副主编）：丁希仑　王少萍　孙志梅　李秋实　李焕喜　杨嘉陵
苗俊刚　相　艳　徐立军　钱德沛　曹晋滨

编　委：马殿富　王　琪　王　聪　邓小燕　王青云　王荣明　刘　宇
刘　红　江　洁　刘　强　闫　鹏　朱天乐　刘铁钢　齐铂金
陈万春　邹正平　苏东林　杨世春　沈成平　邱志平　宋知人
杨树斌　张晓林　杨晓奕　杨继萍　李惠峰　吴新开　张瑞丰
杨照华　宋凝芳　周　锐　林宇震　林贵平　战　强　姚仰平
胡庆雷　赵秋红　段海滨　赵巍胜　席　平　郭　宏　徐　洁
徐世杰　郭洪波　康　锐　翟　锦　熊华钢

编辑部主任：张　嵘

编　　辑：孙　芳　万玉敏　王　茜　卞欢欢　李亚泰

北京航空航天大学学报
Beijing Hangkong Hangtian Daxue Xuebao
（原《北京航空学院学报》）
（月刊　1956年创刊）
第52卷 第10期 2026年10月

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS (JBUAA)
(Monthly, Started in 1956)
Vol.52 No.10 October 2026

主管单位　中华人民共和国工业和信息化部
主办单位　北京航空航天大学
主　　编　赵沁平
编辑出版　《北京航空航天大学学报》编辑部
邮　　编　100191
地　　址　北京市海淀区学院路37号
印　　刷　北京科信印刷有限公司
发　　行　《北京航空航天大学学报》编辑部
发行范围　国内外发行
联系电话　(010) 82316698 82316981
82315499 82317448
电子信箱　jbuaa@ buaa.edu.cn

Administrated by Ministry of Industry and Information Technology of the People's Republic of China
Sponsored by Beijing University of Aeronautics and Astronautics (BUAA)
(Beijing 100191, P. R. China)
Chief Editor ZHAO Qinping
Edited and Published by Editorial Board of JBUAA
Printed by Beijing Kexin Printing Co., Ltd.
Distributed by Editorial Board of JBUAA
Telephone (010) 82316698 82316981
82315499 82317448
E-mail jbuaa@buaa.edu.cn
http://bhxb.buaa.edu.cn

中国标准连续出版物号：ISSN 1001-5965
CN 11-2625/V

ISSN 1001-5965
9 771001 596267 10>

微信公众号　超星电子书

国内定价：50.00元／期