ISSN 1001-5965
CODEN BHHDE8

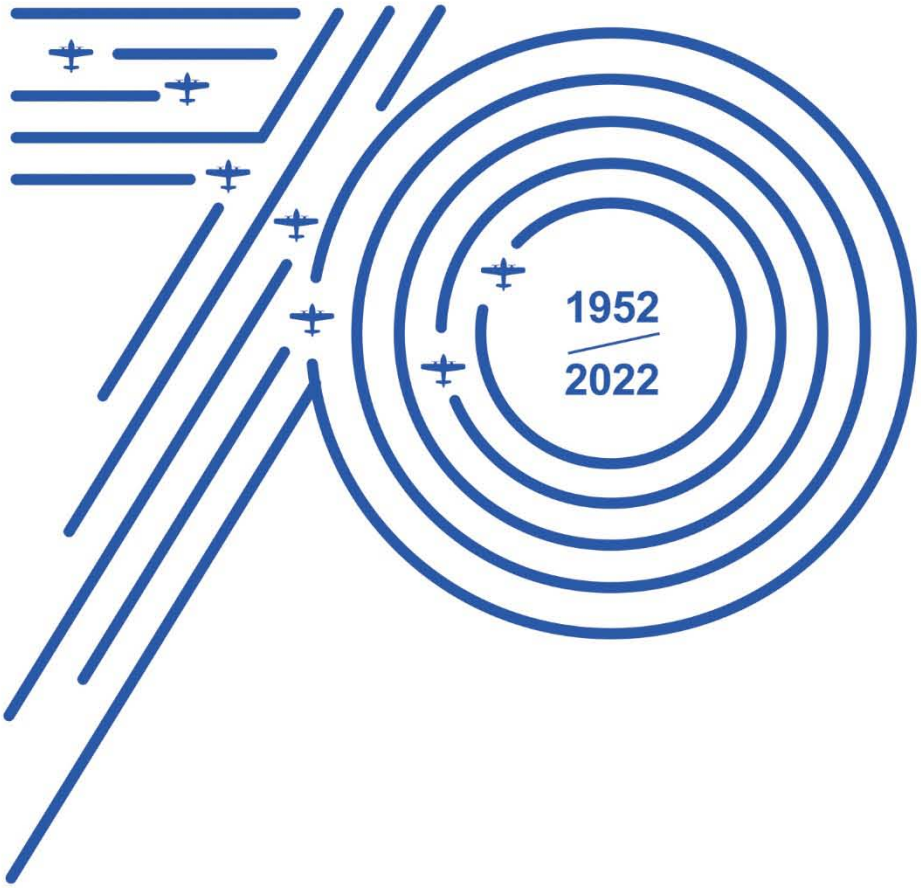
北京航空航天大學学报

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS



70周年校庆专刊

2022-09
Vol.48 No.9



七十载空天报国
新时代逐梦一流

卷首语

柏彦故地，巍巍学府。春秋代序，峥嵘航路。2022年10月25日，北京航空航天大学将迎来建校70周年华诞。

为国而生，与国同行。1952年，抗美援朝的烽火催生了“急需办一所航空大学”的国家需求，汇聚全国八校航空精英，新中国第一所航空航天高等学府——北京航空学院成立。作为新中国红色工程师的摇篮，北航建校7年即成为全国首批重点高校，在新中国各个历史时期都是国家重点建设的大学，首批进入“211工程”，2001年进入“985工程”，2013年入选首批“2011计划”国家协同创新中心，2017年入选国家“双一流”建设高校名单，始终奋进在中国高等教育的第一方阵。七十载创业兴学，服务国家是北航人永恒不变的精神追求，空天报国是北航人最为鲜亮的精神底色。

矢志空天，立德树人。70年来，北航传承听党话、跟党走的红色基因，始终与国家发展和民族振兴同频共振，为新中国国民经济建设和国防事业，特别是航空航天事业作出重要贡献。70年来，北航坚守为党育人、为国育才的初心使命，培养出一大批学术精英、专业人才和治国栋梁。70年来，北航践行爱国奉献、敢为人先的价值追求，服务国家战略，打造国之重器。70年来，北航人心系国家，追寻和坚守三代人的“中国心”“长鹰魂”“陀螺梦”，淬炼了“艰苦朴素、勤奋好学、全面发展、勇于创新”的校风，“德才兼备，知行合一”的校训和“空天报国”的北航精神。

校庆盛典，学术盛会。适逢北京航空航天大学70周年校庆之际，作为北京航空航天大学主办的高水平学术期刊，在学校期刊中心的统一部署下，在各位专家学者的大力支持下，《北京航空航天大学学报》特推出70周年校庆专刊。



70周年校庆专刊刊发了北航院士、知名专家的学术成果以及学科前沿的科研信息，共计25篇，涵盖了航空航天科学、信息科学、能源科学、材料科学、制造与工程科学、空间环境科学、综合交叉科学等国家重点研究领域，旨在体现近年来北航学术研究的技术前沿新视角、科技发展新方向、理论研究新思考、创新攻关新探索，给广大读者一场高水准的学术盛宴。

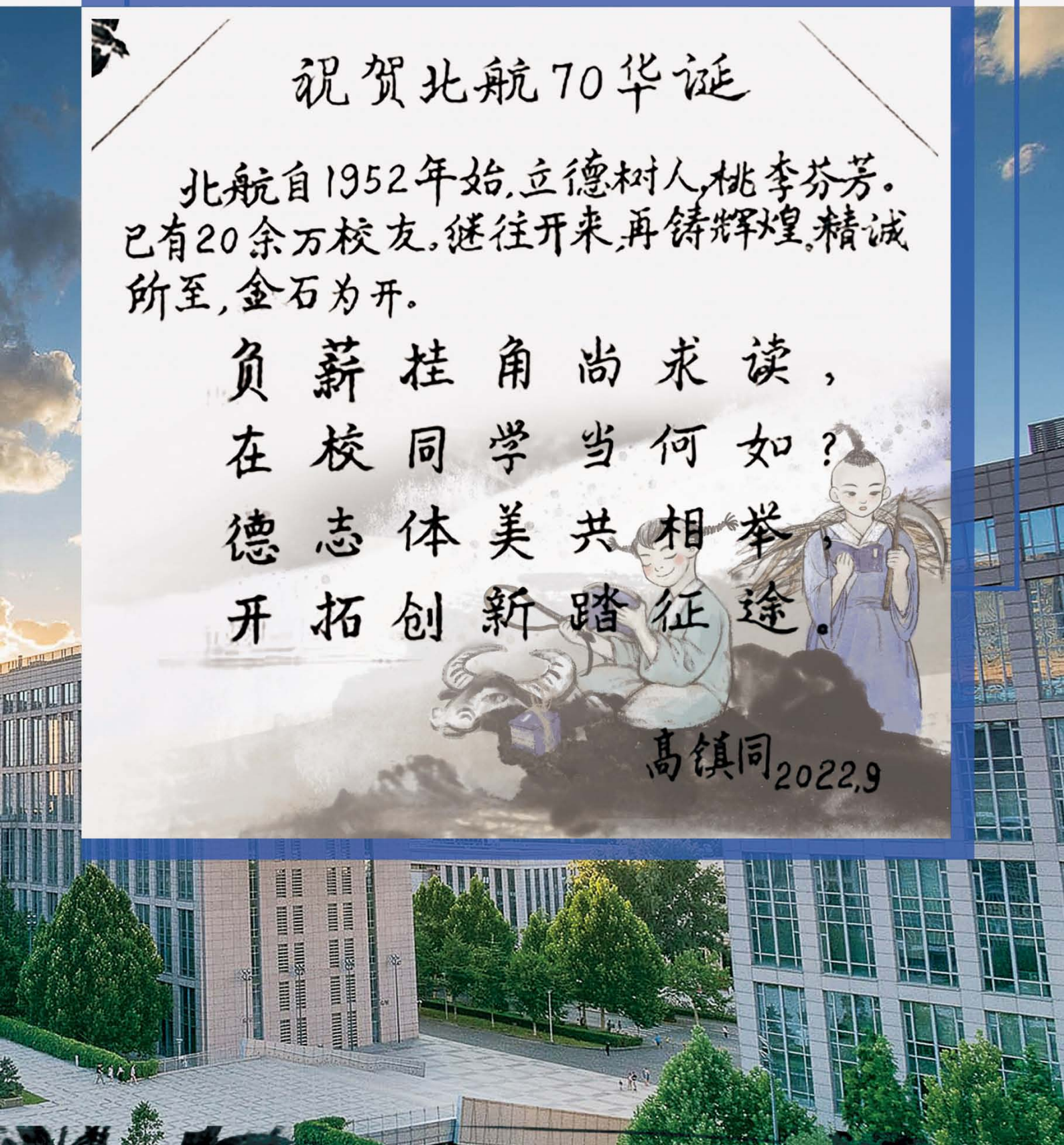
春华秋实，辛勤耕耘，作为航空航天及相关学科优秀学术论文的重要载体、国内核心期刊、中国精品科技期刊，《北京航空航天大学学报》自1956年创刊以来伴随着航空航天事业的迅猛发展，始终屹立学术潮头，不断展现和引领航空航天领域的前沿成果，勇于开拓，不断创新，为航空航天学术的发展做出了应有的贡献。

逐梦一流，再创辉煌。新时代、新征程、新使命，立足新发展阶段，北航期刊将坚持以习近平新时代中国特色社会主义思想为指导，紧紧围绕建设创新型国家的总体目标和学校“双一流”建设目标，以质量建设为核心，依托优势学科，凝聚前沿创新成果，打造高质量期刊，支撑人才培养和学科发展，不断提升学校的学术影响力、文化影响力和国际影响力！

赵沁平院士亲笔题词



高镇同院士亲笔赋诗



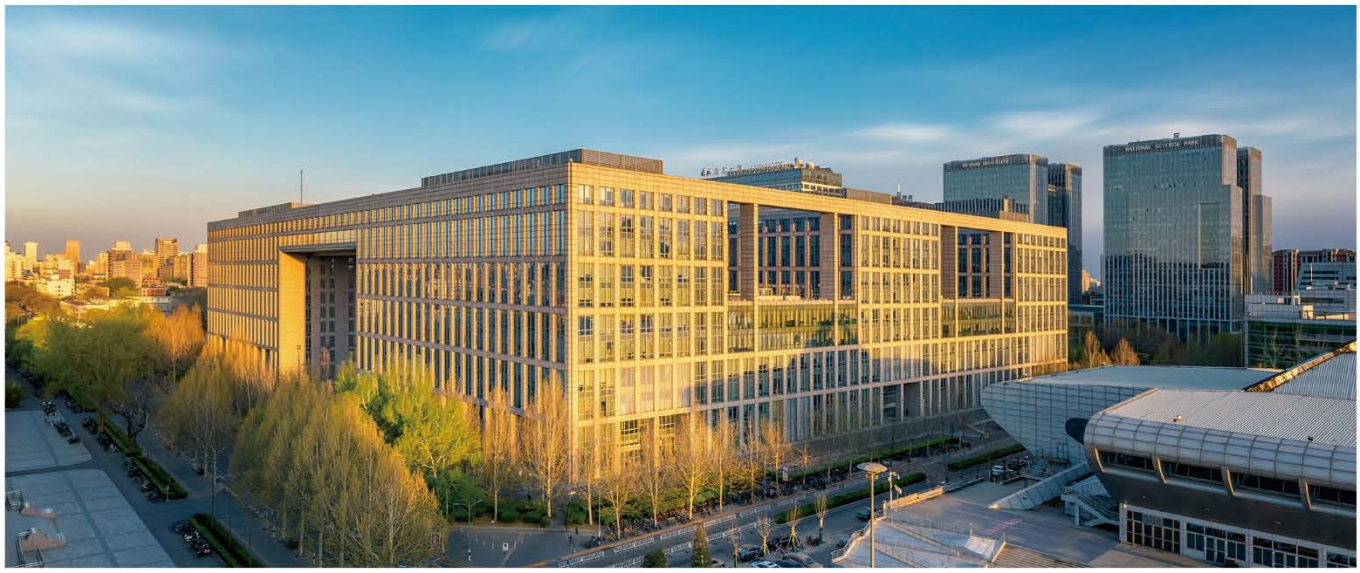
祝贺北航70华诞

北航自1952年始，立德树人，桃李芬芳。已有20余万校友，继往开来，再铸辉煌。精诚所至，金石为开。

负薪挂角尚求读，
在校同学当何如？
德志体美共相举，
开拓创新踏征途。

高镇同 2022.9

校园风采



新主楼





北京航空航天博物馆



图书馆

沙河校区主楼







绿园



沙河校区校门



庆祝沈元学院（原高等工程学院）建院20周年

沈元学院是现高等理工学院、未来空天技术学院、国家卓越工程师学院的校内统称，其前身高等工程学院成立于 2002 年，是北航最早的拔尖创新人才培养基地和教育教学改革试验基地。作为北航的荣誉学院，沈元学院目前承担基础学科拔尖学生培养计划、未来技术学院、工程硕博士培养改革专项试点等国家级创新人才培养改革任务。

砥砺奋进二十载，矢志一流育英才。自成立以来，学院明确并始终坚持“强化基础、突出实践、重在素质、面向创新”的人才培养方针，汇聚国内外一流师资和优势资源，将“导师制、小班化、个性化、国际化”培养模式贯穿人才培养全过程，全面实施“书院制”育人，在本科生和博士生培养方面推行了一系列卓有成效的改革，探索了一条符合教育教学规律、人才成长规律、适应和促进创新人才成长的拔尖创新人才培养之路，培养了一大批具有远大理想、高尚情怀、国际化视野和一流竞争力的学术领军人才和卓越工程师，打造了北航高素质领军领导人才培养的“第一品牌”，赢得了教育界的广泛关注和良好的社会声誉。

立足新时代，学院将牢记为党育人、为国育才的初心使命，落实立德树人根本任务，踔厉奋发，守正创新，扎根基础，面向未来，追求卓越，努力培养堪当大用的时代新人。



高等工程学院成立



更名高等理工学院并冠名沈元荣誉学院



未来空天技术学院成立



国家卓越工程师学院成立

庆祝建院十周年大会



本科生开展军事训练



“重逢少年时·共话高工情”主题活动







“躬逢新时代·筑梦新高工”校友年会



参观中国共产党历史展览



北京航空航天大学 沈元学院 | 学院风采

“永远跟党走·唱支歌儿给党听”快闪活动



“空天杯”篮球赛



航空航天器创新设计与实践专项竞赛



“师生合力育春芽，花草满园生光辉”主题党日活动



本研新生两校区报到

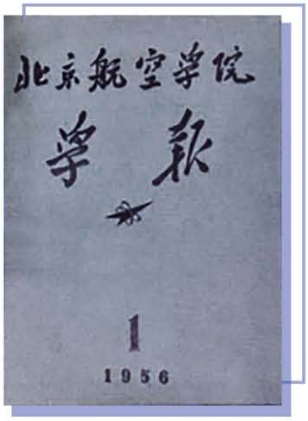




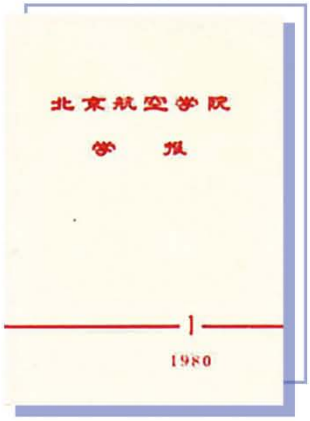


校盛刊强

1956



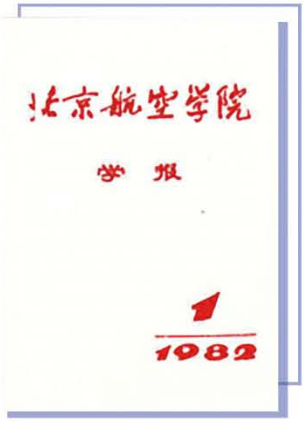
1980



1981



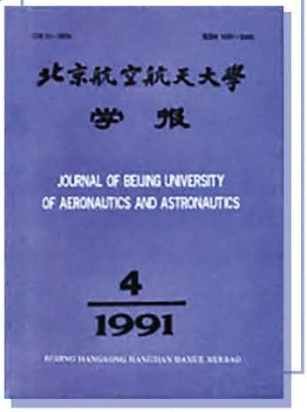
1982



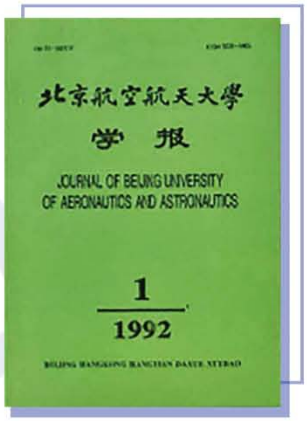
1988



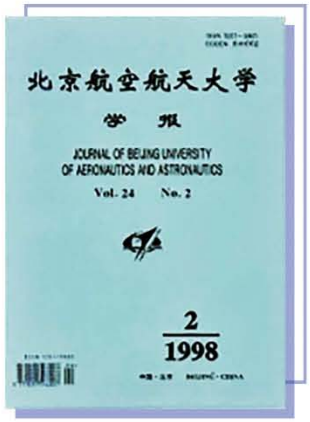
1991



1992



1998



2002



2022



北京航空航天大学学报

第 48 卷　第 9 期　（总第 355 期）　2022 年 9 月

目　　次

复杂电磁环境辐射-散射耦合场快速预估方法…… 苏东林，崔朔，白江飞，李尧尧（1553）
从网格到“东数西算”：构建国家算力基础设施 …… 钱德沛，栾钟治，刘轶（1561）
材料基因工程加速新材料设计与研发 …… 孙志梅，王冠杰，张烜广，周健（1575）
星基 ADS-B 系统及关键技术发展综述…… 张学军，谭元晧，李雪缘，简鑫慧（1589）
基于“时间-特征”协同注意力的机场快轨客流预测 …… 杜文博，石婉君，廖盛时，朱熙（1605）
飞机液压泵二维性能退化的可靠性评估 …… 王少萍，陈仁同，张超（1613）
基于指数平均动量鸽群优化的多无人机协同目标防御 …… 段海滨，仝秉达，刘冀川（1624）
重油航空活塞发动机燃油喷射技术 …… 丁水汀，邵龙涛，赵帅，朱锟，杜发荣，周煜（1630）
超临界二氧化碳闭式布莱顿循环系统研究进展 ……
…… 邹正平，王一帆，姚李超，刘火星，许鹏程，李辉（1643）
基于边界元法的三维结构体滑跳运动数值仿真 …… 杨超，姜宇，吴志刚（1678）
时间积分方法的研究进展与挑战 …… 邢誉峰，季奕，张慧敏（1692）
旋转部件 SGBEM-FEM 耦合热弹性断裂分析…… 董雷霆，贺双新（1702）
运动信息引导的目标检测算法 …… 胡海苗，沈柳青，高立崑，李明竹（1710）
自动铺丝成型构件缺陷在线检测技术进展 ……
…… 丁希仑，罗伟恒，刘斐，王国栋，陈维强，张武翔（1721）
面向智能化管理的数字孪生电池构建方法 … 杨世春，李强伟，周思达，张正杰，马源，陈飞（1734）
新型扭簧电连接器的确信可靠性正向设计优化 …… 康锐，刘浩然，张清源，余丽，翟国富（1745）
流动拓扑对液体射流失稳与雾化的影响 …… 杨立军，黄东骐，韩旺，李敬轩，富庆飞（1757）
基于电磁超声横波的管道剩余厚度检测 …… 徐立军，刘福禄，丁一清，李正勇，谢跃东（1767）
基于电润湿液体透镜的显微成像技术 …… 王琼华，袁荣英，刘超（1774）
仿生减阻表面的进展与挑战 …… 刘明杰，吴青山，严昊，于存明，蔺存国，张金伟，赵天艺（1782）
磷酸掺杂型高温质子交换膜燃料电池关键材料研究进展 ……
…… 相艳，李文，郭志斌，张劲，卢善富（1791）
无人机高机动抗扰轨迹跟踪控制方法 …… 王英勋，宋欣屿，赵江，蔡志浩（1806）
飞行过程中头盔对飞行员颈部损伤的影响 ……
…… 王丽珍，刘景龙，赵彦鹏，卜伟平，柳松杨，樊瑜波（1818）
基于层次分析的信息物理网络多维攻击损毁评估 …… 洪晟，岳天羽，李新建（1827）
多飞行器攻击时间一致性协同制导进展综述与展望 ……
…… 董希旺，于江龙，化永朝，李清东，任章（1836）

期刊基本参数： CN 11-2625/V * 1956 * m * A4 * 292 * zh * P * ￥50.00 * 1400 * 25 * 2022-09

（**编　辑**　张　嵘　孙　芳　卞欢欢　王　茜　苏　磊　黄紫月）

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Vol. 48 No. 9 (Sum 355) September 2022

CONTENTS

Fast prediction method for radiated and scattered coupled fields in complex electromagnetic environment
…… SU Donglin, CUI Shuo, BAI Jiangfei, LI Yaoyao (1553)
From grid to "East-west Computing Transfer": Constructing national computing infrastructure
…… QIAN Depei, LUAN Zhongzhi, LIU Yi (1561)
Novel material design and development accelerated by materials genome engineering
…… SUN Zhimei, WANG Guanjie, ZHANG Xuanguang, ZHOU Jian (1575)
A review of development of space-based ADS-B system and its key technologies
…… ZHANG Xuejun, TAN Yuanhao, LI Xueyuan, JIAN Xinhui (1589)
Passenger flow forecasting of airport express based on time and feature cooperative attention
…… DU Wenbo, SHI Wanjun, LIAO Shengshi, ZHU Xi (1605)
Reliability estimation for aircraft hydraulic pump based on bivariate performance degradation analysis
…… WANG Shaoping, CHEN Rentong, ZHANG Chao (1613)
Coordinated target defense for multi-UAVs based on exponentially averaged momentum pigeon-inspired optimization
…… DUAN Haibin, TONG Bingda, LIU Jichuan (1624)
Fuel injection technology of heavy fuel aircraft piston engine
…… DING Shuiting, SHAO Longtao, ZHAO Shuai, ZHU Kun, DU Farong, ZHOU Yu (1630)
Progress in research of closed supercritical carbon dioxide Brayton cycle system
…… ZOU Zhengping, WANG Yifan, YAO Lichao, LIU Huoxing, XU Pengcheng, LI Hui (1643)
Numerical simulation of skipping motion of three-dimensional structure based on boundary element method
…… YANG Chao, JIANG Yu, WU Zhigang (1678)
Advances and challenges in time integration methods
…… XING Yufeng, JI Yi, ZHANG Huimin (1692)
SGBEM-FEM coupling for thermoelastic fracture mechanics analysis of rotational components
…… DONG Leiting, HE Shuangxin (1702)
Object detection algorithm guided by motion information
…… HU Haimiao, SHEN Liuqing, GAO Likun, LI Mingzhu (1710)
Review on automated fiber placement induced defects and their online monitoring technology
…… DING Xilun, LUO Weiheng, LIU Fei, WANG Guodong, CHEN Weiqiang, ZHANG Wuxiang (1721)
Construction of digital twin model of lithium-ion battery for intelligent management
…… YANG Shichun, LI Qiangwei, ZHOU Sida, ZHANG Zhengjie, MA Yuan, CHEN Fei (1734)
Function oriented belief reliability design and optimization of new torsion spring electrical connectors
…… KANG Rui, LIU Haoran, ZHANG Qingyuan, YU Li, ZHAI Guofu (1745)
Influence of flow topology on instability and atomization of liquid jets
…… YANG Lijun, HUANG Dongqi, HAN Wang, LI Jingxuan, FU Qingfei (1757)
Residual thickness detection of pipeline based on electromagnetic ultrasonic shear wave
…… XU Lijun, LIU Fulu, DING Yiqing, LI Zhengyong, XIE Yuedong (1767)
Microscopic imaging technology with electrowetting liquid lens
…… WANG Qionghua, YUAN Rongying, LIU Chao (1774)
Progress and challenges of bionic drag reduction surfaces
…… LIU Mingjie, WU Qingshan, YAN Hao, YU Cunming, LIN Cunguo, ZHANG Jinwei, ZHAO Tianyi (1782)
Research progress on key materials of phosphoric acid doped high-temperature proton exchange membrane fuel cells
…… XIANG Yan, LI Wen, GUO Zhibin, ZHANG Jin, LU Shanfu (1791)
Anti-disturbance trajectory tracking control method for aggressive quadrotors
…… WANG Yingxun, SONG Xinyu, ZHAO Jiang, CAI Zhihao (1806)
Effect of helmet on neck injury of pilots in flight
…… WANG Lizhen, LIU Jinglong, ZHAO Yanpeng, BU Weiping, LIU Songyang, FAN Yubo (1818)
Multidimensional attack damage assessment of cyber-physical systems based on hierarchical analysis
…… HONG Sheng, YUE Tianyu, LI Xinjian (1827)
Review and prospect of cooperative guidance with attack time consensus for multiple aerial vehicles
…… DONG Xiwang, YU Jianglong, HUA Yongzhao, LI Qingdong, REN Zhang (1836)

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0705

复杂电磁环境辐射-散射耦合场快速预估方法

苏东林[1,2], 崔朔[1,3], 白江飞[1], 李尧尧[2,*]

(1. 北京航空航天大学 电子信息工程学院, 北京 100083; 2. 北京航空航天大学 前沿科学技术创新研究院, 北京 100083;
3. 北京航空航天大学 沈元学院, 北京 100083)

摘 要: 针对复杂电磁环境中目标散射体和干扰辐射源同时存在,影响目标回波散射场的问题,提出了一种基于辐射源方向图和极化散射矩阵数据的耦合场快速预估方法。利用可提前独立获取和加载的辐射源方向图及目标各个方向的散射数据,实现了辐射-散射耦合场景下空间总电磁场的近实时快速预估。仿真计算了辐射源与散射体不同距离和相对强度下场景的总电场变化情况,验证了辐射-散射耦合效应对目标回波场存在显著影响。无需使用电磁计算方法对场景进行重新计算,满足复杂电磁环境内场仿真试验对实时提供数据的需求,具有工程应用价值。

关 键 词: 雷达散射截面; 有源干扰; 快速预估; 极化散射矩阵; 辐射方向图
中图分类号: TN974; V221+.92
文献标志码: A **文章编号**: 1001-5965(2022)09-1553-08

随着电子信息技术与电子对抗技术的发展,在当前及未来的作战场景中,干扰辐射源和目标散射体同时存在,战场态势瞬息万变,电磁环境严酷复杂[1]。为了分析、评估电子信息系统的复杂电磁环境适应性能力,掌握系统性能边界,对复杂电磁环境的数值模拟与仿真必不可少[2-3]。

目标电磁散射特性是影响目标回波幅度的关键因素,蕴含着目标的几何结构、材料等信息[4]。在目标电磁散射特性研究中,研究者提出了雷达散射截面(radar cross section, RCS)[5]、散射振幅系数[6]、极化散射矩阵[7-8]等描述参数,为散射特性问题的分析提供了可量化的概念。在辐射研究中,天线远场辐射方向图[9-10]是表征辐射特性的典型特征。

研究者引入多种方法进行电磁场的仿真计算。低频数值方法包括矩量法(method of moments, MOM)[11]、有限元法(finite element method, FEM)[12]和时域有限差分法(finite difference time domain, FDTD)[13]。近似求解方法包括几何光学法(geometric optics, GO)[14]、物理光学法(physical optics, PO)[15-16]、弹跳射线法(shooting and bouncing ray, SBR)[17]、一致性几何绕射理论(uniform geometrical theory of diffraction, UTD)[18-20]等。为了进一步提高计算效率、减少资源占用,研究者提出了多种加速技术,如快速多级子方法(fast multipole method, FMM)、多层快速多极子方法(multilevel fast multipole method, MLFMM)、矩阵预处理技术、递归方法、快速傅里叶变换方法等,来快速求解线性方程组[21-22],基于 CPU/GPU 的硬件加速、树形数据结构快速检索技术等来快速追迹射线管[23],基于高斯回归预测、神经网络建模等策略来建立数学预测模型[24-25]。

就研究对象而言,上述描述方法或关注金属/介质结构的散射特性求解,或关注天线的辐射特性求解。对于目标散射体和辐射源同时存在的复杂电磁环境,空间中既有辐射场,又有散射场,辐射场对目标的散射场、探测雷达的接收场均存在

收稿日期: 2022-08-01; **录用日期**: 2022-08-18; **网络出版时间**: 2022-08-23 10:59
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220823.0954.001.html
* **通信作者**. E-mail: liyaoyao@buaa.edu.cn

引用格式: *苏东林, 崔朔, 白江飞, 等. 复杂电磁环境辐射-散射耦合场快速预估方法[J]. 北京航空航天大学学报, 2022, 48(9): 1553-1560. SU D L, CUI S, BAI J F, et al. Fast prediction method for radiated and scattered coupled fields in complex electromagnetic environment[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1553-1560(in Chinese).*

影响。此时,雷达接收到的空间总场不能用目标单独存在时的散射场和干扰单独存在的辐射场直接叠加的方式来描述,还要考虑辐射源对目标的耦合散射场。综上所述,面向辐射源和目标散射体共同存在的实际场景仿真的应用需求,仍存在一定研究空间:

1)对抗场景中,雷达、目标散射体、干扰辐射源的相对态势瞬息万变,场景电磁环境的仿真预测需要具备近实时计算效率,常规的电磁计算方法难以满足要求。

2)对于不掌握精确几何结构的非合作目标,传统的依赖目标和天线的几何模型的电磁计算方法存在应用困难。

3)对于散射场与辐射场相互交织的复杂场景,传统的单独针对散射或辐射的计算不能描述辐射场与散射场间的物理作用,存在模型误差。

4)动态场景中,孤立目标或干扰辐射源的位置和姿态变化,但散射和辐射特性不变,若能在场景电磁环境计算中利用这一特点,有望进一步节约资源、提升效率。

因此,本文提出一种辐射-散射耦合场快速预估方法。首先,载入天线装机方向图和平台雷达散射截面数据;然后,通过射线定向和引爆计算辐射场对散射场的激励;最后,通过场的极化矢量叠加,得到辐射、散射耦合条件下的场景远场结果。本文方法将复杂电磁环境辐射-散射耦合场的计算与环境中目标的散射特性数据库和干扰的辐射特性数据库建立联系,在先验数据的支撑下,通过建立和调用对应的散射特性、辐射特性数据开展场景的电磁计算,实现了耦合散射场的实时预测。仿真结果验证了辐射源存在情况下,雷达接收的电磁场较目标原始散射场存在明显差异,从而将影响雷达正确识别目标。

1 辐射-散射耦合场模型

基于雷达视角,对包含干扰辐射源和目标散射体的复杂电磁环境进行分析,如图 1 所示。

空间存在的总电场 $\boldsymbol{E}$ 可分为 4 类:雷达入射场 $\boldsymbol{E}_{\mathrm{i}}$、目标在雷达照射下产生的原始散射场 $\boldsymbol{E}_{\mathrm{s}}$、散射体在辐射源照射下产生的次生散射场 $\boldsymbol{E}'_{\mathrm{s}}$及干扰辐射源产生的辐射场 $\boldsymbol{E}_{\mathrm{r}}$,即

$$\boldsymbol{E} = \boldsymbol{E}_{\mathrm{i}} + \boldsymbol{E}_{\mathrm{s}} + \boldsymbol{E}_{\mathrm{r}} + \boldsymbol{E}'_{\mathrm{s}} \tag{1}$$

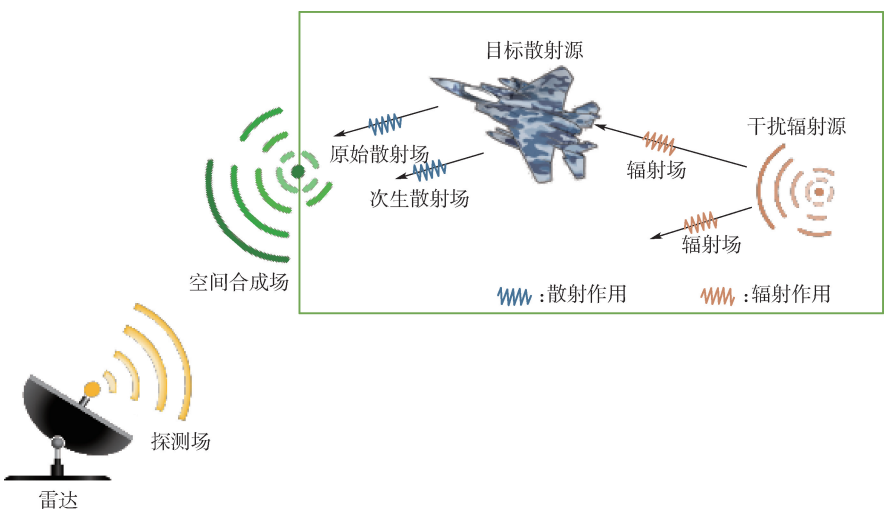


图 1 辐射-散射耦合场景示意图

Fig. 1 Radiated and scattered coupled fields

1.1 目标原始散射场

目标原始散射场 $\boldsymbol{E}_{\mathrm{s}}$ 是目标散射体对雷达入射场 $\boldsymbol{E}_{\mathrm{i}}$ 的响应。从感应电流的观点分析,入射场在散射体表面激发感应电流 $\boldsymbol{J}_{\mathrm{s}}$,进而在空间形成散射场。

$$\boldsymbol{E}_{\mathrm{s}}(\boldsymbol{r}) = -\mathrm{j}\omega\mu\int_{V_{\mathrm{s}}} G(\boldsymbol{r},\boldsymbol{r}')\left[1+\frac{1}{k^2}\nabla'\nabla'\cdot\right]\boldsymbol{J}_{\mathrm{s}}(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}' \tag{2}$$

式中:ω 为角频率;μ 为真空磁导率;V_{s} 为散射体;$\boldsymbol{r}$ 和 $\boldsymbol{r}'$分别为场点和源点;$G(\boldsymbol{r},\boldsymbol{r}')$ 为自由空间点源格林函数;k 为波数。

应注意的是,多散射体同时存在时,各散射体上的感应电流不等于它们孤立地被同样的入射场激发的感应电流。原因在于:各散射体的散射场形成了新的激励源,影响其他散射体上的感应电流。

1.2 干扰源辐射场

实际中的干扰源辐射场通常来自于飞行平台上雷达和干扰机辐射的电磁场。辐射源电流密度为 $\boldsymbol{J}_{\mathrm{r}}$,则引发的辐射电场 $\boldsymbol{E}_{\mathrm{r}}$ 为

$$\boldsymbol{E}_{\mathrm{r}}(\boldsymbol{r}) = -\mathrm{j}\omega\mu\int_{V_{\mathrm{r}}} G(\boldsymbol{r},\boldsymbol{r}')\left[1+\frac{1}{k^2}\nabla'\nabla'\cdot\right]\boldsymbol{J}_{\mathrm{r}}(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}' \tag{3}$$

式中:V_{r} 为干扰辐射源。

1.3　目标次生散射场

对于 1.2 节所述的辐射场,当其照射到目标表面时,会在目标上激发感应电流,并产生散射场。空间中的辐射场也会照射到目标表面,激发感应电流并产生散射场。为与入射场激发的原始散射场相区分,辐射场激发的感应电流成为次生感应电流,辐射场激发的散射场称为目标的次生散射场。

照射到目标上的辐射场 $\boldsymbol{E}'_{\mathrm{i}}(\boldsymbol{r})$ 为

$$\boldsymbol{E}'_{\mathrm{i}}(\boldsymbol{r}) = \boldsymbol{E}_{\mathrm{r}}(\boldsymbol{r}) \qquad \boldsymbol{r}\in V_{\mathrm{s}} \tag{4}$$

$\boldsymbol{E}'_{\mathrm{i}}(\boldsymbol{r})$ 在目标表面激发次生感应电流 $\boldsymbol{J}'_{\mathrm{s}}(\boldsymbol{r})$,引发的目标次生散射场 $\boldsymbol{E}'_{\mathrm{s}}(\boldsymbol{r})$ 为

$$\boldsymbol{E}'_{\mathrm{s}}(\boldsymbol{r}) = -\mathrm{j}\omega\mu\int_{V_{\mathrm{s}}} G(\boldsymbol{r},\boldsymbol{r}')\left[1+\frac{1}{k^2}\nabla'\nabla'\cdot\right]\boldsymbol{J}'_{\mathrm{s}}(\boldsymbol{r}')\,\mathrm{d}\boldsymbol{r}' \tag{5}$$

2　辐射-散射耦合场快速预估方法

实际工程中,单目标、单辐射源的特性数据往往都是具备的,这些数据中隐含了目标散射体和辐射源的几何外形、材料参数等特征信息。利用这些已知先验信息,可以等效替代复杂环境求解中耗时的积分或路径寻迹步骤,直接实现辐射和散射电磁场的快速求解。

2.1　辐射场快速计算方法

2.1.1　远场方向图

天线远场方向图通常描述为

$$F(\theta_{\mathrm{r}},\varphi_{\mathrm{r}}) = \frac{|\boldsymbol{E}_{\mathrm{r}}(\theta_{\mathrm{r}},\varphi_{\mathrm{r}})|}{E_{\mathrm{r,max}}} \tag{6}$$

式中:$\boldsymbol{E}_{\mathrm{r}}$ 为远区球面上任意 $(\theta_{\mathrm{r}},\varphi_{\mathrm{r}})$ 方向某点处的场强;$E_{\mathrm{r,max}}$ 为球面上 $\boldsymbol{E}_{\mathrm{r}}$ 的最大振幅。

2.1.2　辐射场的计算和变换

在干扰源天线处建立本地坐标系。对于与干扰源天线相对方向相同的矢量,其在场景中干扰源天线在本地坐标系下的角度坐标与在干扰源天线远场方向图中的角度坐标一致。干扰源天线本地角度坐标与全局角度坐标的转换关系详见 2.2.2节。

借助天线远场方向图,局部坐标系下的干扰源辐射场 $\boldsymbol{E}_{\mathrm{r,l}}$ 的计算公式为

$$\boldsymbol{E}_{\mathrm{r,l}}(\boldsymbol{r}) = \alpha\hat{\boldsymbol{e}}_{\mathrm{r,l}}F(\theta_{\mathrm{r,l}},\varphi_{\mathrm{r,l}})\,\mathrm{e}^{-\mathrm{j}[\boldsymbol{k}_{\mathrm{r}}\cdot(\boldsymbol{r}-\boldsymbol{r}_{\mathrm{a}})+\beta]} \tag{7}$$

式中:α 为电场幅度系数;$\hat{\boldsymbol{e}}_{\mathrm{r,l}}$ 为干扰源天线的电场单位矢量;$\boldsymbol{r}$ 为空间场点,在干扰源天线局部坐标系中的角度坐标为 $(\theta_{\mathrm{r,l}},\varphi_{\mathrm{r,l}})$;$\boldsymbol{k}_{\mathrm{r}}$ 为干扰源辐射场波矢;$\boldsymbol{r}_{\mathrm{a}}$ 为干扰源天线的等效相位中心;β 为干扰源辐射场初始相位。

将该干扰源辐射场转换到全局坐标系下,表示为

$$\begin{aligned}\boldsymbol{E}_{\mathrm{r}} = {} & (\boldsymbol{E}_{\mathrm{r,l}}\cdot\hat{\boldsymbol{\theta}}_{\mathrm{r,l}} + \boldsymbol{E}_{\mathrm{r,l}}\cdot\hat{\boldsymbol{\varphi}}_{\mathrm{r,l}})\cdot\hat{\boldsymbol{\theta}}_{\mathrm{r}} + \\ & (\boldsymbol{E}_{\mathrm{r,l}}\cdot\hat{\boldsymbol{\theta}}_{\mathrm{r,l}} + \boldsymbol{E}_{\mathrm{r,l}}\cdot\hat{\boldsymbol{\varphi}}_{\mathrm{r,l}})\cdot\hat{\boldsymbol{\varphi}}_{\mathrm{r}}\end{aligned} \tag{8}$$

式中:$\hat{\boldsymbol{\theta}}_{\mathrm{r}}$ 和 $\hat{\boldsymbol{\varphi}}_{\mathrm{r}}$ 分别为全局坐标系中干扰源辐射场的垂直极化方向基矢和水平极化方向基矢;$\hat{\boldsymbol{\theta}}_{\mathrm{r,l}}$ 和 $\hat{\boldsymbol{\varphi}}_{\mathrm{r,l}}$ 分别为本地坐标系中干扰源辐射场的垂直极化方向基矢和水平极化方向基矢。

2.2　散射场快速计算方法

2.2.1　极化散射矩阵

目标的极化散射矩阵完整地描述了目标的极化特性。通过极化散射矩阵,可以快速得到目标的极化散射场。极化散射矩阵 $\overline{\overline{\boldsymbol{S}}}$ 通常表示为

$$\overline{\overline{\boldsymbol{S}}}(\theta_{\mathrm{s}},\varphi_{\mathrm{s}}\mid\theta_{\mathrm{i}},\varphi_{\mathrm{i}}) = \begin{bmatrix} S_{\theta\theta} & S_{\theta\varphi} \\ S_{\varphi\theta} & S_{\varphi\varphi}\end{bmatrix} \tag{9}$$

式中:$S_{ij}(i,j=\theta,\varphi)$ 为目标以 i 极化接收和 j 极化发射时的复标量散射系数,i、j 代表 2 个正交的极化基方向;θ_{s} 和 φ_{s} 分别为散射方向的俯仰角和方位角;θ_{i} 和 φ_{i} 分别为入射方向的俯仰角和方位角。

2.2.2　散射场的计算和变换

由于目标在场景中的姿态与其在极化散射矩阵数据库中的姿态不尽相同,为将目标极化散射矩阵中的角度与场景中的角度做对应,采用坐标系转换的方法。

建立场景的全局坐标系,记为 $\boldsymbol{O}(\hat{\boldsymbol{x}},\hat{\boldsymbol{y}},\hat{\boldsymbol{z}})$。在目标处建立本地坐标系,记为 $\boldsymbol{r}_0(\hat{\boldsymbol{u}},\hat{\boldsymbol{v}},\hat{\boldsymbol{w}})$。对于与目标相对方向相同的矢量,其在场景中目标在本地坐标系下的角度坐标与在目标极化散射矩阵数据库中的角度坐标一致,如图 2 所示。本地坐标系基矢与场景全局坐标系基矢间的关系为

$$\begin{bmatrix}\hat{\boldsymbol{u}}\\ \hat{\boldsymbol{v}}\\ \hat{\boldsymbol{w}}\end{bmatrix} = \begin{bmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z\end{bmatrix}\begin{bmatrix}\hat{\boldsymbol{x}}\\ \hat{\boldsymbol{y}}\\ \hat{\boldsymbol{z}}\end{bmatrix} \tag{10}$$

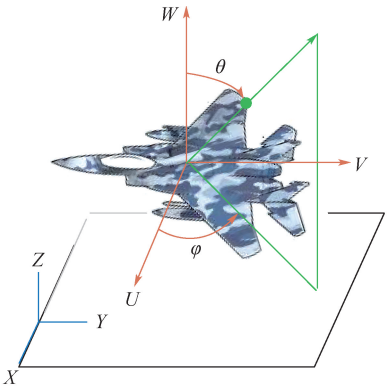
对于矢量 $\boldsymbol{A} = A_x\hat{\boldsymbol{x}} + A_y\hat{\boldsymbol{y}} + A_z\hat{\boldsymbol{z}}$,其本地角度坐标 $(\theta_{\mathrm{l}},\varphi_{\mathrm{l}})$ 与全局角度坐标 (θ,φ) 的转换关系为

$$\theta_{\mathrm{l}} = \arccos(-A_w) \tag{11}$$

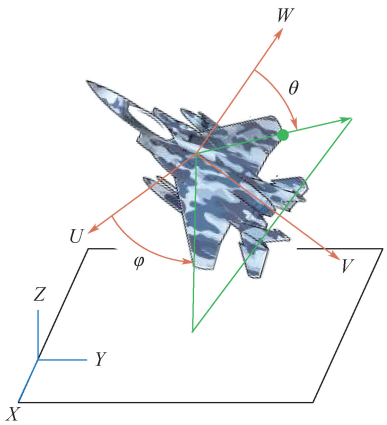
$$\varphi_{\mathrm{l}} = \begin{cases}\arctan(A_v/A_u) & A_u\neq 0\\ 0 & A_u = 0\end{cases} \tag{12}$$

式中:

$$\begin{bmatrix}A_u\\ A_v\\ A_w\end{bmatrix} = \begin{bmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z\end{bmatrix}\begin{bmatrix}\sin\theta\cos\varphi\\ \sin\theta\sin\varphi\\ \cos\varphi\end{bmatrix} \tag{13}$$



(a) 极化散射矩阵数据库中目标的坐标系定义



(b) 场景中目标的坐标系定义

图 2 极化散射矩阵数据库与场景中目标的坐标系定义

Fig. 2 Coordinate system definitions of polarization scattering matrix database and target in the scene

入射场场强在局部坐标系的表达式为

$$\boldsymbol{E}_{\mathrm{i},l} = \mathrm{e}^{-\mathrm{j}\boldsymbol{k}_{\mathrm{i}}\cdot\boldsymbol{r}_0}\left[(\boldsymbol{E}_{\mathrm{i}}\cdot\hat{\boldsymbol{\theta}}_{\mathrm{i}} + \boldsymbol{E}_{\mathrm{i}}\cdot\hat{\boldsymbol{\varphi}}_{\mathrm{i}})\cdot\hat{\boldsymbol{\theta}}_{\mathrm{i},l} + (\boldsymbol{E}_{\mathrm{i}}\cdot\hat{\boldsymbol{\theta}}_{\mathrm{i}} + \boldsymbol{E}_{\mathrm{i}}\cdot\hat{\boldsymbol{\varphi}}_{\mathrm{i}})\cdot\hat{\boldsymbol{\varphi}}_{\mathrm{i},l}\right] \tag{14}$$

式中：$\boldsymbol{k}_{\mathrm{i}}$ 为入射场波矢；$\boldsymbol{r}_0$ 为散射体位置矢量；$\boldsymbol{E}_{\mathrm{i}}$ 为场景的入射场场强；$\hat{\boldsymbol{\theta}}_{\mathrm{i}}$ 和 $\hat{\boldsymbol{\varphi}}_{\mathrm{i}}$ 分别为全局坐标系中入射场的垂直极化方向基矢和水平极化方向基矢；$\hat{\boldsymbol{\theta}}_{\mathrm{i},l}$和 $\hat{\boldsymbol{\varphi}}_{\mathrm{i},l}$分别为本地坐标系中入射场的垂直极化方向基矢和水平极化方向基矢。

假定入射场的源和目标的距离、接收天线和目标的距离均满足远场条件，借助极化散射矩阵，局部坐标系的远场散射场 $\boldsymbol{E}_{\mathrm{s},l}$的计算公式为

$$\boldsymbol{E}_{\mathrm{s},l} = \overline{\overline{\boldsymbol{S}}}(\theta_{\mathrm{s},l},\varphi_{\mathrm{s},l}\mid\theta_{\mathrm{i},l},\varphi_{\mathrm{i},l})\cdot\boldsymbol{E}_{\mathrm{i},l}\mathrm{e}^{-\mathrm{j}\boldsymbol{k}_{\mathrm{s}}\cdot(\boldsymbol{r}-\boldsymbol{r}_0)} \tag{15}$$

式中：$\theta_{\mathrm{s},l}$和 $\varphi_{\mathrm{s},l}$分别为散射场波矢方向在目标本地坐标系中的俯仰角和方位角；$\theta_{\mathrm{i},l}$和 $\varphi_{\mathrm{i},l}$分别为入射场波矢方向在目标本地坐标系中的俯仰角和方位角；$\boldsymbol{k}_{\mathrm{s}}$ 为散射场波矢；$\boldsymbol{r}$ 为远场场点。

将该散射场转换到全局坐标系下，表示为

$$\boldsymbol{E}_{\mathrm{s}} = (\boldsymbol{E}_{\mathrm{s},l}\cdot\hat{\boldsymbol{\theta}}_{\mathrm{s},l} + \boldsymbol{E}_{\mathrm{s},l}\cdot\hat{\boldsymbol{\varphi}}_{\mathrm{s},l})\cdot\hat{\boldsymbol{\theta}}_{\mathrm{s}} + (\boldsymbol{E}_{\mathrm{s},l}\cdot\hat{\boldsymbol{\theta}}_{\mathrm{s},l} + \boldsymbol{E}_{\mathrm{s},l}\cdot\hat{\boldsymbol{\varphi}}_{\mathrm{s},l})\cdot\hat{\boldsymbol{\varphi}}_{\mathrm{s}} \tag{16}$$

式中：$\hat{\boldsymbol{\theta}}_{\mathrm{s}}$ 和 $\hat{\boldsymbol{\varphi}}_{\mathrm{s}}$ 分别为全局坐标系中散射场的垂直极化方向基矢和水平极化方向基矢；$\hat{\boldsymbol{\theta}}_{\mathrm{s},l}$和 $\hat{\boldsymbol{\varphi}}_{\mathrm{s},l}$分别为本地坐标系中散射场的垂直极化方向基矢和水平极化方向基矢。

2.3 次生散射场快速计算方法

对于次生散射场，由于目标的极化散射特性不因入射场的变化而变化，仍可使用极化散射矩阵进行快速求解。

当散射体和干扰辐射源位置满足远场条件、散射体可近似为一点时，干扰源辐射场在散射体处可近似为沿连线方向传播的平面波。此时，目标的次生散射场可表示为

$$\boldsymbol{E}'_{\mathrm{s}}(\theta_{\mathrm{s},l},\varphi_{\mathrm{s},l}) = \overline{\overline{\boldsymbol{S}}}(\theta_{\mathrm{s},l},\varphi_{\mathrm{s},l}\mid\theta_{\mathrm{r},l},\varphi_{\mathrm{r},l})\cdot\boldsymbol{E}_{\mathrm{r},l}(\boldsymbol{r}_0)\cdot\mathrm{e}^{-\mathrm{j}\boldsymbol{k}_{\mathrm{s}}\cdot(\boldsymbol{r}-\boldsymbol{r}_0)} \tag{17}$$

式中：$\theta_{\mathrm{r},l}$和 $\varphi_{\mathrm{r},l}$为干扰辐射源相对于目标的局部角度坐标；$\boldsymbol{r}$ 为远场场点；$\boldsymbol{E}_{\mathrm{r},l}(\boldsymbol{r}_0)$为 $\boldsymbol{r}_0$ 处的远场干扰辐射场。

2.4 场合成方法

为了计算辐射-散射矢量耦合场总场，本文方法对耦合场的各成分分量向全局坐标系下的极化基做投影，再按极化类型分别进行求和。

以 2 个场成分为例。将场 1 整理为全局坐标系下的表达式为

$$\boldsymbol{E}_1 = E_1^{\theta}\hat{\boldsymbol{\theta}} + E_1^{\varphi}\hat{\boldsymbol{\varphi}} \tag{18}$$

式中：$\hat{\boldsymbol{\theta}}$ 和 $\hat{\boldsymbol{\varphi}}$ 分别为全局坐标系中耦合场的垂直极化方向基矢和水平极化方向基矢。

场 2 整理为全局坐标系下的表达式为

$$\boldsymbol{E}_2 = E_2^{\theta}\hat{\boldsymbol{\theta}} + E_2^{\varphi}\hat{\boldsymbol{\varphi}} \tag{19}$$

则场 1 与场 2 的合成场的计算公式为

$$\boldsymbol{E}_{\mathrm{sum}} = (E_1^{\theta} + E_2^{\theta})\hat{\boldsymbol{\theta}} + (E_1^{\varphi} + E_2^{\varphi})\hat{\boldsymbol{\varphi}} \tag{20}$$

3 仿真实验

建立理想导体球和电偶极子辐射源、飞机和喇叭天线辐射源 2 个算例，分别就本文方法原理的正确性和方法的适用性开展验证。所有仿真均在 Intel(R) Core(TM) i7-7700 CPU@3.60 GHz，16.0 GB RAM 的台式计算机上完成。

3.1 理想导体球和电偶极子辐射源

仿真场景设置如图 3 所示。位于 *XOY* 平面上的理想导体球半径为 0.5 m，一个辐射源位于从球心沿 *Y* 方向相距 *D* 处，辐射源为偶极矩 0.005 Am的理想电偶极子。平面波从 $\theta_{\mathrm{i}} = 0°$，$\varphi_{\mathrm{i}} = 0°$处入射。仿真频率为 300 MHz。

图 4 给出了不同距离 *D* 和辐射源强度倍数因子 *M* 时，整个场景的远场电场强度计算结果。

以图 4(a)为例,将商业软件 feko 的 MOM 求解器的求解结果作为参考值,本文方法的均方根误差为 0.002 7,结果一致性良好。每个三维全空间场景平均用时 0.008 9 s。可以分析得到以下规律:

1)在相同距离 D 下,辐射场的倍数因子 M 越大,空间总远场电场分布越接近电偶极子孤立存在时的辐射电场;辐射场的倍数因子 M 越小,空间总远场电场分布越接近理想导体球孤立存在时的散射电场。

2)在相同辐射场倍数因子 M 下,距离 D 越大,空间总远场电场分布越接近此时场强相对较大的场成分。

这些规律符合物理规律,验证了本文方法的正确性。

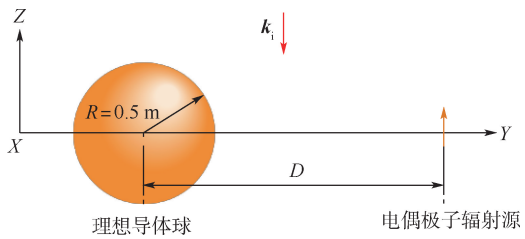
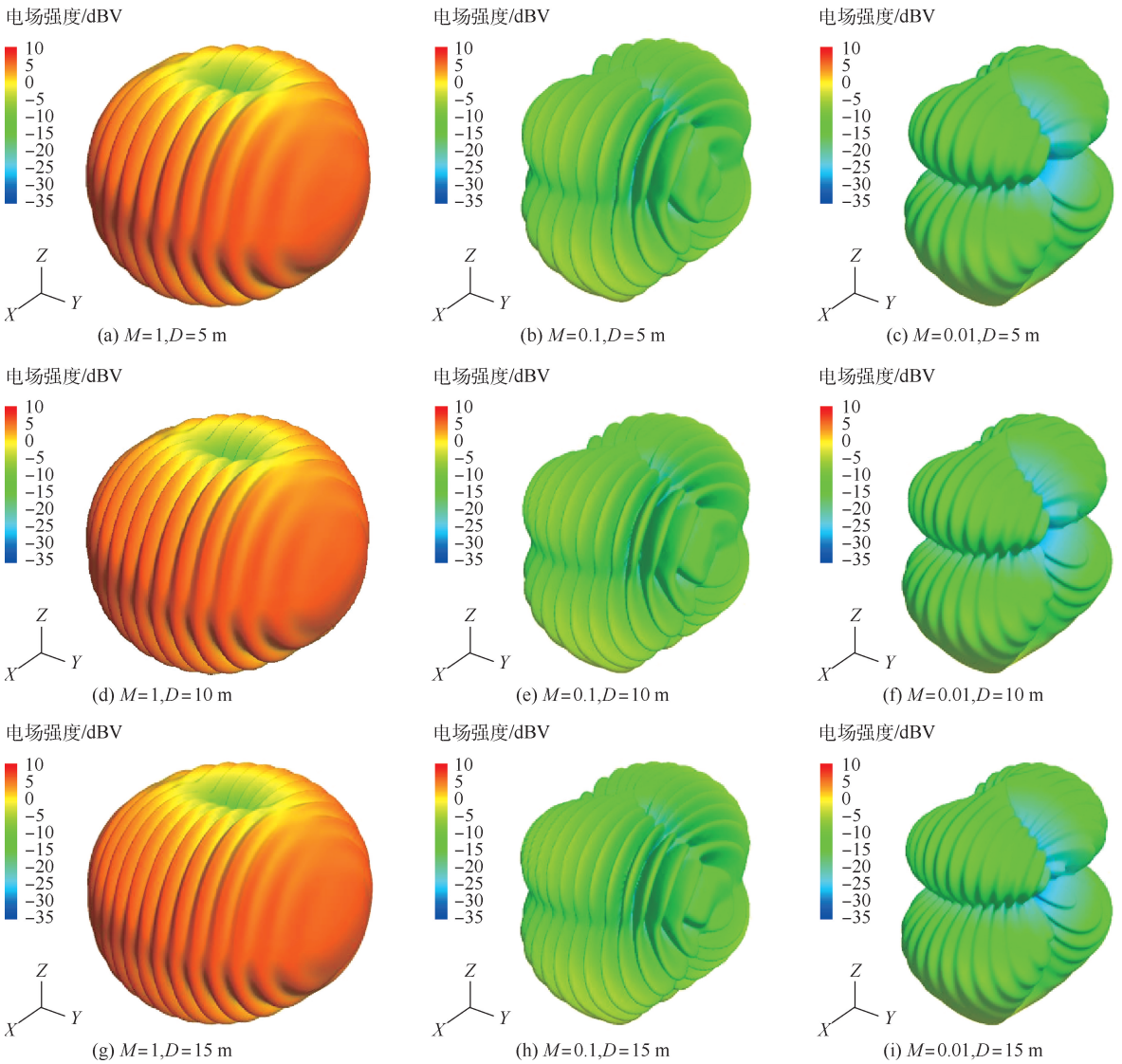


图 3　理想导体球与电偶极子仿真场景示意图

Fig. 3　Simulation scene of a perfect conductor sphere and an electric dipole



(a) M=1,D=5 m　(b) M=0.1,D=5 m　(c) M=0.01,D=5 m

(d) M=1,D=10 m　(e) M=0.1,D=10 m　(f) M=0.01,D=10 m

(g) M=1,D=15 m　(h) M=0.1,D=15 m　(i) M=0.01,D=15 m

图 4　不同距离和辐射源强度倍数因子的理想导体球与电偶极子远场电场强度对比

Fig. 4　Comparison of far-field electric field strength of a perfect conductor sphere and an electric dipole at different distances and radiation source intensity multiplier factors

3.2　飞机和喇叭天线辐射源

为了验证本文方法的实用性,对飞机和喇叭天线辐射源的场景进行计算。仿真场景如图 5 所示。飞机等效相位中心位于 $\boldsymbol{r}_a=(x_a,y_a,z_a)$;辐射源为喇叭天线的辐射场,等效相位中心位于(0,0,0) m。平面波从 $\theta_i=90°$,$\varphi_i=30°$处入射。仿

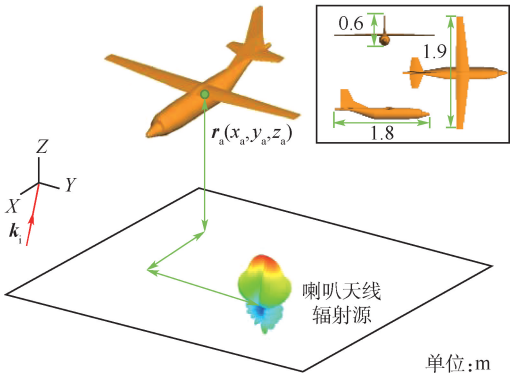
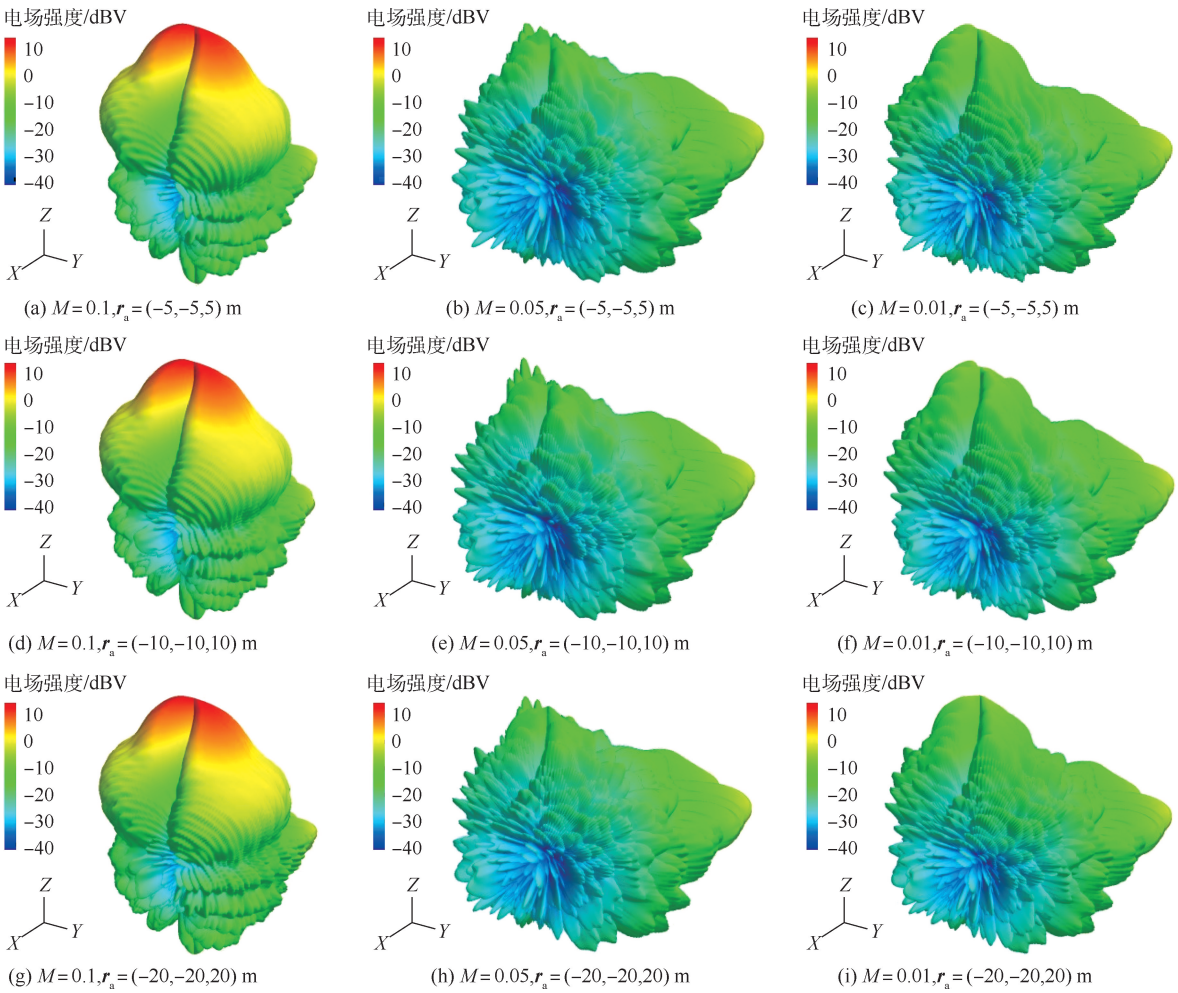


图 5 飞机与喇叭天线辐射源仿真场景示意图

Fig. 5 Simulation scene of an aircraft and a conical horn antenna's radiated field

真频率为 3 GHz。

图 6 给出了不同距离下,整个场景的远场电场强度计算结果。以图 6(i)为例,将商业软件 feko的 MOM 求解器的求解结果作为参考值,本文方法的均方根误差为 0.000 09,结果一致性良好。每个三维全空间场景平均用时 0.009 5 s,与 3.1 节场景算例的计算效率持平。通过对结果分析可以发现,辐射场强度和位置均会对空间中的场分布产生显著影响。其中,辐射场强主要影响目标回波散射场强度,辐射场位置主要影响目标回波散射场分布。



(a) $M=0.1, \boldsymbol{r}_a=(-5,-5,5)$ m (b) $M=0.05, \boldsymbol{r}_a=(-5,-5,5)$ m (c) $M=0.01, \boldsymbol{r}_a=(-5,-5,5)$ m

(d) $M=0.1, \boldsymbol{r}_a=(-10,-10,10)$ m (e) $M=0.05, \boldsymbol{r}_a=(-10,-10,10)$ m (f) $M=0.01, \boldsymbol{r}_a=(-10,-10,10)$ m

(g) $M=0.1, \boldsymbol{r}_a=(-20,-20,20)$ m (h) $M=0.05, \boldsymbol{r}_a=(-20,-20,20)$ m (i) $M=0.01, \boldsymbol{r}_a=(-20,-20,20)$ m

图 6 不同距离和辐射源强度倍数因子的飞机与喇叭天线辐射源远场电场强度对比

Fig. 6 Comparison of far-field electric field strength of an aircraft and a conical horn antenna's radiated field at different distances and radiation source intensity multiplier factors

4 结 论

区别于经典的计算电磁学计算方法,本文提出了一种辐射-散射耦合场快速预估方法,适于工程应用。

1)利用提前获取并加载的天线辐射方向图和目标极化散射特性数据,能够实现对复杂电磁环境近实时预测,适用于动态电磁环境的内场仿真。

2)不需要掌握目标和天线的几何模型信息,即可开展计算。

3)方法的计算效率对求解频率、场景电尺寸不敏感,在同样的仿真计算机配置下,相较其他方法具备更好的电大复杂场景的电磁环境仿真预测

能力。

未来将开展进一步的研究,使本文方法可以处理近场辐射源与散射体耦合的情况。

参考文献(References)

[1] 隋起胜,袁建全. 反舰导弹战场电磁环境仿真及试验鉴定技术[M]. 北京:国防工业出版社,2015:1-8.
SUI Q S, YUAN J Q. Anti-ship missile battlefield electromagnetic environment simulation and test identification technology [M]. Beijing: National Defense Industry Press, 2015: 1-8 (in Chinese).

[2] 苏东林,谢树果,戴飞,等. 系统级电磁兼容性量化设计理论与方法[M]. 北京:国防工业出版社,2015:49-52.
SU D L, XIE S G, DAI F, et al. Theory and method for quantitative design of system-level electromagnetic compatibility[M]. Beijing: National Defense Industry Press, 2015: 49-52 (in Chinese).

[3] 王磊,苏东林,谢树果,等. 飞机进近着陆电磁环境建模与辐射分布分析[J]. 北京航空航天大学学报,2012,38(10):1369-1374.
WANG L, SU D L, XIE S G, et al. Modeling of electromagnetic environment and radiation distribution analysis for aircraft approaching and landing [J]. Journal of Beijing University of Aeronautics and Astronautics, 2012, 38 (10): 1369-1374 (in Chinese).

[4] 黄培康,殷红成,许小剑. 雷达目标特性[M]. 北京:电子工业出版社,2005:4-22.
HUANG P K, YIN H C, XU X J. Radar target characteristics [M]. Beijing: Publishing House of Electronics Industry, 2005: 4-22 (in Chinese).

[5] 何国瑜,卢才成,洪家才,等. 电磁散射的计算和测量[M]. 北京:北京航空航天大学出版社,2004:19-28.
HE G Y, LU C C, HONG J C, et al. Calculation and measurement of electromagnetic scattering [M]. Beijing: Beihang University Press, 2004: 19-28 (in Chinese).

[6] ROY J E. The concept of the scattered field and the finite-difference time-domain method of the scattered-field formulation [J]. IEEE Transactions on Antennas and Propagation, 2016, 64 (9): 4102-4107.

[7] BICKEL S H. Some invariant properties of the polarization scattering matrix[J]. Proceedings of the IEEE, 1965, 53(8): 1070-1072.

[8] LI C, LI Y Z, YANG Y, et al. Moving target's scattering matrix estimation with a polarimetric radar[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(8): 5540-5551.

[9] 吴龙刚,苏东林,陈佳佳,等. 半空间中机载甚低频双拖曳天线电磁辐射特性[J]. 北京航空航天大学学报,2011,37(11):1471-1474.
WU L G, SU D L, CHEN J J, et al. Analysis of electromagnetic radiation for airborne very-low-frequency (VLF) dual trailing antenna in half-space[J]. Journal of Beijing University of Aeronautics and Astronautics, 2011, 37 (11): 1471-1474 (in Chinese).

[10] LIU N W, ZHU L, LIU Z X, et al. Radiation pattern reshaping of a narrow slot antenna for bandwidth enhancement and stable pattern using characteristic modes analysis[J]. IEEE Transactions on Antennas and Propagation, 2022, 70(1): 726-731.

[11] CAO X Y, CHEN M S, QI Q, et al. An improved GMRES method for solving electromagnetic scattering problems by MoM[J/OL]. IEEE Transactions on Antennas and Propagation, 2022 (2022-07-07) [2022-08-10]. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9818956.

[12] BOTELHO D P, MARECHAL Y, RAMDANE B. Higher order NEM and FEM accuracy comparison[J]. IEEE Transactions on Magnetics, 2015, 51(3): 1-4.

[13] CHEN G Z, YANG S C, SU D L. An accurate three-dimensional FDTD(2,4) method on face-centered cubic grids with low numerical dispersion[J]. IEEE Antennas and Wireless Propagation Letters, 2019, 18(9): 1711-1715.

[14] LEAO T F C, CHOPIN V M, TRUEMAN C W. Electromagnetic characterization of a gyproc slab by measurement and 3-D geometrical optics simulation [J]. IEEE Antennas and Wireless Propagation Letters, 2013, 12: 1570-1573.

[15] 苏东林,宗国明,吕善伟. 武装直升飞机雷达散射截面的估算方法[J]. 北京航空航天大学学报,1994,20(3):248-252.
SU D L, ZONG G M, LV S W. The method of calculating radar cross section of fighting helicopters[J]. Journal of Beijing University of Aeronautics and Astronautics, 1994, 20(3): 248-252 (in Chinese).

[16] 苏东林,曾国奇,刘焱,等. 运动目标 RCS 特性分析[J]. 北京航空航天大学学报,2006,32(12):1413-1417.
SU D L, ZENG G Q, LIU Y, et al. RCS study of moving radar targets[J]. Journal of Beijing University of Aeronautics and Astronautics, 2006, 32(12): 1413-1417 (in Chinese).

[17] WANG J J, GUO L X, WEI Y W, et al. Application of the improved SBR-TSM based on MPI to EM scattering from multiple targets above a 3-D rough sea surface[J]. IEEE Antennas and Wireless Propagation Letters, 2022, 21(2): 411-415.

[18] 苏东林,王宝发,张采来. 目标外形拟合及雷达散射截面计算[J]. 北京航空航天大学学报,1990(1):101-108.
SU D L, WANG B F, ZHANG C L. Contour mimicry of targets & radat cross section's calculation[J]. Journal of Beijing University of Aeronautics and Astronautics, 1990 (1): 101-108 (in Chinese).

[19] 王冰切,苏东林,张晓雷. 飞机表面绕射射线的寻迹方法[J]. 北京航空航天大学学报,2007,33(7):785-788.
WANG B Q, SU D L, ZHANG X L. Discrete ray path tracing on aircraft[J]. Journal of Beijing University of Aeronautics and Astronautics, 2007, 33(7): 785-788 (in Chinese).

[20] 李尧尧,苏东林,刘焱,等. 高精度法矢下切割面自适应的凸曲面射线寻迹[J]. 北京航空航天大学学报,2016,42(12):2632-2639.
LI Y Y, SU D L, LIU Y, et al. Convex surface ray tracing based on adaptive cutting surface adjustment under exact normal vector[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(12): 2632-2639 (in Chinese).

[21] TZOULIS A, EIBERT T F. A hybrid FEBI-MLFMM-UTD method for numerical solutions of electromagnetic problems including arbitrarily shaped and electrically large objects [J]. IEEE

Transactions on Antennas and Propagation, 2005, 53(10): 3358-3366.

[22] KARAGOUNIS G, ZUTTER D D, GINSTE D V. A hybrid MLFMM-UTD method for the solution of very large 2-D electromagnetic problems[J]. IEEE Transactions on Antennas and Propagation, 2016, 64(1): 224-234.

[23] HUO J C, XU L, SHI X W, et al. An accelerated shooting and bouncing ray method based on GPU and virtual ray tube for fast RCS prediction[J]. IEEE Antennas and Wireless Propagation Letters, 2021, 20(9): 1839-1843.

[24] XIAO D H, GUO L X, LIU W, et al. Improved Gaussian process regression inspired by physical optics for the conducting target's RCS prediction[J]. IEEE Antennas and Wireless Propagation Letters, 2020, 19(12): 2403-2407.

[25] LIU Z, WANG L F, WEN Z D, et al. Multilevel scattering center and deep feature fusion learning framework for SAR target recognition[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-14.

Fast prediction method for radiated and scattered coupled fields in complex electromagnetic environment

SU Donglin[1,2], CUI Shuo[1,3], BAI Jiangfei[1], LI Yaoyao[2,*]

(1. School of Electronic and Information Engineering, Beihang University, Beijing 100083, China;
2. Research Institute for Frontier Science, Beihang University, Beijing 100083, China;
3. Shenyuan Honors College, Beihang University, Beijing 100083, China)

Abstract: A coupled scattered field fast prediction method based on the radiation source pattern and polarization scattering matrix is proposed for the complex electromagnetic environment in which the target and interfering radiation sources exist simultaneously and affect the scattered field of the target. The near real-time fast prediction of the overall electromagnetic field in the radiation-scattering coupled scenario is made possible by utilizing the radiation source pattern and the scattering data of all directions of the target, which can be independently obtained and loaded in advance. The effects of the radiation-scattering coupling on the scene's echo field are determined through simulations by calculating the changes in the total electric field of the radiation source and scatter at different distances and relative strengths. Calculation of the whole field of the scene doesn't involve electromagnetic calculation methods, which meets the demand for real-time data provision in the complex electromagnetic environment infield simulation test and is engineering applicable.

Keywords: radar cross section; active interference; fast prediction; polarization scattering matrix; radiation pattern

Received: 2022-08-01; **Accepted**: 2022-08-18; **Published online**: 2022-08-23 10:59

URL: kns.cnki.net/kcms/detail/11.2625.V.20220823.0954.001.html

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0715

从网格到“东数西算”:构建国家算力基础设施

钱德沛*, 栾钟治, 刘轶
(北京航空航天大学 计算机学院, 北京 100083)

摘 要: 简要回顾了几十年来计算机使用方式的变迁,介绍了基于网络计算技术的国家高性能计算基础设施 CNGrid 的设计与实现。讨论了在“东数西算”战略工程背景下中国算力发展的新趋势,以及国家算力基础设施发展面临的新的技术挑战,并对中国未来超算应用生态和算力基础设施建设提出了展望。

关 键 词: 高性能计算; 基础设施; 网格计算; CNGrid; 东数西算
中图分类号: TP311; TP316; TP391
文献标志码: A **文章编号**: 1001-5965(2022)09-1561-14

1 计算机使用模式的演变

计算机是20世纪人类社会最伟大的发明之一,它的出现,彻底改变了人类生活、工作的面貌。计算与模拟和理论与分析、实验与观测一起,成为人类认识客观世界、开展科学研究的重要手段。对现代科学研究而言,计算的重要性不言而喻。利用计算,可以更清晰地揭示客观世界的发展规律,探索和预测未知的事物。例如,大数据处理分析、人工智能大模型训练和推理、新能源及其利用、新材料设计、工业产品创新设计、创新药物研发、精确天气预报、全球气候变化预测、社会治理和决策支持等,都依赖计算机的强大算力。因此,算力已经成为一个国家创新能力和综合国力的体现。

伴随算力的提高,如何能更容易地使用计算机,便捷地获得所需的算力,也是人们一直追求的目标。计算机诞生70多年来,其使用方式一直在不断变化。早期,人围着计算机转,用户要跑到专门的机房去上机,计算机被一个用户单道程序所独占。随着操作系统的进步,计算机的使用方式逐渐从单道程序的主机模式向多道程序的批作业模式[1]和分时交互模式[2]发展。在分时计算系统[3]中,众多用户可以通过终端,同时使用一台计算机,每个用户在分配给他的时间片内使用计算机,但感觉上似乎是在独占这台计算机。20世纪80年代,个人计算机的出现使计算机进入千家万户,但是个人计算机的性能有限,孤立的个人计算机难以胜任大的计算任务。直到20世纪另一个伟大发明计算机网络的出现和普及,带来网络计算的新变革,计算机的使用方式才发生了影响更深远的变化。所谓网络计算,就是通过网络连接网上分散的计算机,汇聚网络连接的各类硬件和软件资源,形成能力更为强大的计算系统。用户可以通过网络随时随地访问计算机,使用计算资源,完成自身任务,而无需关心计算资源的物理所在。

美国的超算中心联网是网络计算系统的早期范例。20世纪80年代中期,美国国防用途的Arpanet进入商用,Internet诞生。在美国国家科学基金会(NSF)支持下,建设了主干速率56 Kbps、运行TCP/IP协议的NSFNET网络。NSFNET将美国加利福尼亚大学的圣地亚哥超级计算机中心(SDSC)、伊利诺伊大学的国家超级计算应用中心

收稿日期: 2022-08-02; **录用日期**: 2022-08-18; **网络出版时间**: 2022-08-23 09:07
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220822.1844.002.html
* **通信作者**. E-mail: depeiq@buaa.edu.cn

引用格式: 钱德沛, 栾钟治, 刘轶. 从网格到“东数西算”:构建国家算力基础设施[J]. 北京航空航天大学学报, 2022, 48(9): 1561-1574. QIAN D P, LUAN Z Z, LIU Y. From grid to “East-west Computing Transfer”: Constructing national computing infrastructure[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1561-1574(in Chinese).

(NCSA)、康奈尔大学的康奈尔国家超级计算机研究室(CNSF)、匹兹堡超级计算机中心(PBC)、冯·诺依曼国家超级计算机中心(JVNNSC)和美国国家大气研究中心(NCAR)的科学计算分部连接起来,对大学和科研用户提供可远程使用的计算资源[4]。

20 世纪 90 年代中期,网格计算(grid computing)的概念在美国兴起[5]。网格(grid)一词最初是指电力网(power grid),网格计算借用电力网的概念,提出要利用高速互联网把分布于不同地理位置的计算、数据、存储和软件等资源连为一体,通过调度、管理和安全保障机制,建立一个像电网一样的计算网格,把算力像电力那样输送给终端用户,支持共享使用和协同工作[6]。在美国 NSF 的支持下,分别由 NCSA 和 SDSC 牵头,实施了 2 个网格计算项目,初步建立了计算网格的雏形[7]。网格计算研究在 20 世纪末到 21 世纪最初 10 年达到高潮。在美国的倡导下,成立了全球网格论坛 GGF,与此对应,国际 IT 大公司联合成立了企业网格论坛 EGF。2006 年,GGF 和 EGF 合并,成为开放网格论坛 OGF。GGF 提出了开放网格服务基础设施 OGSI 和开放网格服务体系架构 OGSA 等标准[8-9],协调全球网格计算的研究和开发力量,研究资源管理,安全、信息服务及数据管理等网格计算基本理论和关键技术。在 Globus 项目[10]支持下研发了 Globus Toolkit 3.0 (GT3) 软件,GT3 作为 OGSI 的一个完整的参考实现,成为网格计算的事实标准。

在网格计算热潮中,美国、欧盟、日本、中国都实施了一批网格计算研究计划或项目。部分代表性项目如表 1 所示。

表 1 世界部分网格相关研究计划

Table 1 Part of grid-related programs in the world

国别	网格计算相关研究计划/项目
美国	TeraGrid, XSEDE
欧盟	EGEE, EGI
英国	UK e-Science
日本	NAREGI, HPCI
韩国	K* Grid
中国	CNGrid, ChinaGrid

美国是网格计算的发源地,该方向的研究计划持续时间最长,实施的项目数量最多。美国 NSF 专门设立 Cyberinfrastructure 部门,持续稳定支持网格计算方向的研究。美国的网格项目主要有2 类。第 1 类由美国 NSF 支持,在先进计算伙伴计划 PACI 之后,从 20 世纪末开始实施 TeraGrid 项目[11],其主要目标是用网格计算技术推动国家科技进步,保持美国的科技领先地位。2011 年,TeraGrid 的后继项目 XSEDE 项目(https://www.xsede.org/)正式启动,该项目旨在连接全球的计算机、数据和研究人员,建立可供科学家共享的计算环境。美国 NSF 资助的开放科学网格(OSG)在其基础软件 HTCondor(https://htcondor.org/)支持下,实现了众多大学与国家实验室的计算资源共享,为科学家提供了科学计算的环境。第 2 类网格项目由美国国防部、能源部等支持,其主要目标是更好地完成本部门的任务。2 类研究的应用目标有所不同,但共同点是要发展先进的基于网络的应用基础设施,实现应用层面的互联互通、资源共享、协同工作。

欧盟于 2000 年和 2001 年分别启动了欧洲网格计划(EuroGrid)(https://www.eurogrid.org/)和欧洲数据网格计划(European DataGrid)[12]。在欧洲数据网格计划的基础上,2004 年 3 月,欧盟框架研究计划启动了 EGEE 项目[13],其目标是基于网格技术开发欧洲的服务网格基础设施,供科学家全天候使用。2011 年,欧盟框架计划又启动了 EGEE 的后继项目 EGI(https://www.egi.eu/)。在这些项目支持下,研发了欧盟的网格中间件 gLite (http://glite.cern.ch/),建立了可持续运维的泛欧计算基础设施。

英国的网格研究计划是 UK e-Science[14],其目标是用网格技术改变科学研究的模式,推动科学技术的进步,长远目标是影响未来的信息技术基础设施。在 UK e-Science 计划支持下,英国在大学和研究机构建立了一批国家 e-Science 中心,依托 OMII-UK 项目研发了英国的开放网格中间件,开发了一批面向 e-Science 的网格应用系统。

日本文部科学省(MEXT)在 2003 年启动了“国家研究网格基础设施”项目 NAREGI[15]。NAREGI 构建在日本教育科研网 SuperSINET 之上,旨在研制并部署面向科学研究的网格基础设施,并参与全球开放网格组织 OGF 的工作,为网格的标准化活动提供支持。在 NAREGI 之后,日本政府又结合 E 级超级计算机的研制,启动了日本高性能计算基础设施项目 HPCI。HPCI 通过 SuperSINET 连接日本大学和研究机构中的 10 个大超算中心和 2 个大数据中心,形成日本的国家级计算基础设施。

中国的网格计算研究起步于 20 世纪 90 年代末,科学技术部(以下简称科技部)是支持网格计算研究的主要政府部门,从 1999 年起,中国在高

性能计算和网格方向连续实施了多个国家 863 重大项目和国家重点研发专项,表 2 列出了科技部在该方向支持的主要项目。在这些项目的持续支持下,研发了国家高性能计算环境系统软件 CNGrid GOS 和 CNGrid Suite,使用环境系统软件,聚合了分布在全国各地近 20 个超算中心和高性能计算中心的计算资源,实现了资源的互联互通与统一共享、作业的提交与全局调度、数据的全局管理和环境的安全管控,在此基础上,成功构建了基于网格/网络计算技术的国家高性能计算环境——中国国家网格服务环境 CNGrid(参见 http://www.cngrid.org)。CNGrid 历经 20 余年的发展,正从“可用”迈向“好用”,目前的聚合计算能力超过 50 亿亿次,存储容量近 500 PB,部署了 600 多个应用软件和工具软件,支撑了数千项国家科技计划项目和重要工程项目的研究工作,用户覆盖基础研究、工业设计、能源环境和信息服务等众多领域,极大促进了中国科技创新能力的提高,已经成为科学研究、技术创新、工程设计中不可或缺的新型信息基础设施。

表 2　中国科技部的网格和高性能计算项目

Table 2　Grid and high performance computing projects under the Ministry of Science and Technology of China

项目来源	项目名称	执行周期	项目成果
国家 863 重大课题	国家高性能计算环境	1999—2000 年	4 000 亿次曙光 3000;包含 5 个高性能计算中心的国家高性能计算环境原型
国家 863 重大专项	高性能计算机及核心软件	2002—2005 年	11.2 万亿次曙光 4000,5.36 万亿次的联想深腾 6800;国家高性能计算环境实验床“中国国家网格 CNGrid”,8 个结点,18 万亿次计算能力;一批网格应用
国家 863 重大项目	高效能计算机及网格服务环境	2006—2010 年	4 700 万亿次的天河 1A,3 000 万亿次的曙光 6000,1 071 万亿次的神威蓝光;具有服务特征的国家网格服务环境 CNGrid,11 个结点,8 000万亿次计算能力;一批网格和高性能计算应用
国家 863 重大项目	高效能计算机及应用服务环境	2011—2015 年	12.5 亿亿次的神威·太湖之光,10 亿亿次的天河 2A;以服务支持应用的国家高性能计算环境 CNGrid,14 个结点,20 亿亿次计算能力;一批高性能计算应用
国家重点研发专项	高性能计算	2016—2021 年	E 级计算机;初步具备基础设施形态的国家高性能计算环境 CNGrid,19 个结点,52 亿亿次计算能力;一批高性能计算应用

2006 年兴起的云计算是网络计算技术与应用模式的一次大变革。与以往由学术界主导的技术热潮不同,云计算从开始就是由 IT 公司提出并引领的。2006 年 3 月,亚马逊公司推出弹性计算云 EC2(http://aws.amazon.com/ec2),2006 年 8 月,谷歌公司首席执行官埃里克·施密特在搜索引擎大会首次提出“云计算”的概念。此后,微软、戴尔、IBM 等国际 IT 巨头和百度、阿里等中国互联网公司都纷纷跟进。在学术界,美国加利福尼亚大学伯克利分校的 Armbrust 等也专门发文,阐述云计算的学术问题[16]。几年之内,云计算已从新兴技术发展成为全球的热点技术。云的资源被虚拟化,可以动态升级,资源被所有云计算用户通过网络方便地使用。云计算的出现改变了 IT 应用系统部署运行的方式。在传统 IT 应用模式下,应用部门需要自行采购计算机硬件和软件,在私有的计算系统上安装部署自己的应用软件,运行和维护应用系统。在云计算模式下,用户无须自行采购维护计算机,而是从云服务商那里租赁所需的计算资源,在云中安装特定的应用软件,存放应用的数据,完成应用系统的部署,应用系统就能够运行在云端。应用部门本身不需要采购和维护私有的计算机,当应用需求变化时,可以根据需要增加或减少租赁的云计算资源。服务和按用付费是云计算的商业模式,是计算向基础设施形态迈出的一大步。根据所提供的服务内容,云计算可分为 IaaS(提供基础资源)、PaaS(提供平台服务)和 SaaS(提供应用软件)[17]。根据服务的范围和应用的性质,云计算又可分为公有云、私有云和混合云[18]。按照服务封装部署方式又可分为虚拟机、容器、裸金属服务器等[19]。今天,几乎所有大数据中心都在某种程度上使用云计算技术,提供云服务。云计算技术也被引入传统的高性能计算领域,出现了以云方式运行超级计算中心的“云超算”和提供高性能计算能力的“超算云”[20]。

中国的云计算和国际同步发展。国家 863 计划在 2010 年就启动了“中国云”重大项目,支持阿里、百度等互联网公司研发云计算系统。“十三五”期间实施了“云计算与大数据”重点专项,更加系统全面地推进云计算关键技术和系统的研发与应用。今天,阿里云、华为云、百度云、浪潮云等已经在国内市场举足轻重。

物联网(IoT)、基于移动互联网应用的蓬勃发展催生了边缘计算[21]。边缘计算的目的是使应用程序、数据和计算能力(服务)更加靠近端用户,而不是更靠近集中的云,这样就能减少数据的移动,降低数据传输的延迟,降低端系统和数据中心之间的传输带宽需求,达到更低的成本和更好的用户体验的效果。随着边缘计算技术的进步,云-边-端融合的 IT 应用模式也日趋流行,成为渗透更宽广应用领域的网络计算的新形态。

2 国家高性能计算基础设施 CNGrid

建设国家级高性能计算基础设施是创新型国家建设的战略需求。基于网络计算的计算基础设施具有如下特征:①动态性。系统的状态和行为动态变化,资源动态接入和退出、设备随时会出故障、网络可能拥塞甚至断开、用户的数量会不断变化等。②自治性。地理分散的资源在支持广泛共享的同时,仍能保持原有的隶属和管理属性。③开放性。硬件、软件和服务来自不同的厂商,由不同的团队开发,遵循不同的技术规范,兼容并蓄,形成自然生长演化的计算生态环境。这种动态、自治、开放的基础设施不同于资源集中拥有和控制的云计算环境。

在开放、动态的互联网环境下,聚合网上异构、自治的分散资源,构建在全国范围共享使用的国家高性能计算基础设施,面临重大技术挑战:①在动态环境下如何应对系统资源的不确定性,对用户提供稳定的高质量服务;②在不改变原有资源隶属关系和管理模式的条件下,如何实现受控共享;③在开放异构的环境下,如何高效开发和运行大规模分布并行应用,建立高性能计算应用的生态环境。国家高性能计算基础设施 CNGrid 通过体系结构、系统软件、应用模式、应用开发与优化技术等创新应对上述挑战,为在中国形成高性能计算资源提供、应用开发和运行服务的完整产业链奠定了技术基础。

2.1 非集中层次虚拟化体系结构及系统软件

针对开放动态环境下分布异构资源的统一管理与受控共享、系统安全及服务质量保障等重大技术难题,设计并实现了国家高性能计算基础设施“三横两纵”的非集中层次虚拟化体系结构,如图 1 所示。“三横”是指自底向上的内核系统层、系统服务层和应用层。内核系统层通过资源实体和虚拟组织等抽象,将地理分布、自治的高性能计算物理资源抽象和聚合为可动态划分、申请和调度的虚拟资源,通过运行时虚拟地址空间和自治安全策略,解决资源视图、资源发现及定位、异构资源统一访问等基础性问题。系统服务层通过访问虚拟资源,以服务化形式向上层应用提供作业管理、数据访问与传输、应用编程、用户映射等功能。应用层使用系统服务层提供的功能,实现应用的业务逻辑,服务最终用户。“两纵”是贯穿内核系统层、系统服务层、应用层 3 个层次的环境监控管理和安全机制,保障环境的可管理性和安全性。

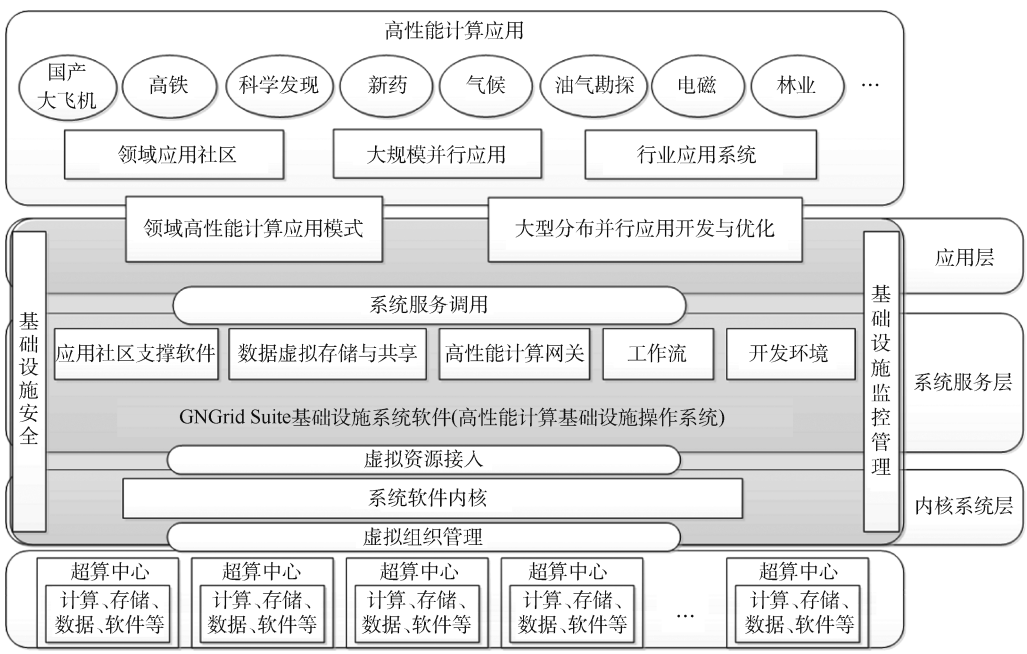


图 1 国家高性能计算基础设施的非集中层次虚拟化体系结构

Fig. 1 Decentralized hierarchical virtualization architecture for national high performance computing infrastructure

体系结构的非集中是指 CNGrid 的管控采用地理分布模式,即在每个 CNGrid 结点部署一台运行系统软件的服务器,通过覆盖网络将各个 CNGrid 结点动态组织成星型、网状或混合结构,以适应国家高性能计算基础设施对资源的分层分域管理的需求。

基于非集中层次虚拟化体系结构,研发了基础设施系统软件 CNGrid Suite,其系统架构如图 2 所示。CNGrid Suite 提出并实现了“资源实体”、“虚拟组织”和“运行时虚拟地址空间”等 3 个系统核心抽象,来表达系统中的各种资源要素、要素间的静态关系和运行时的动态关系,通过这些抽象,将分散、异构、无序的计算机硬件资源、软件资源和用户组织成逻辑有序、可受控共享的虚拟资源,支持资源的动态聚合、调度和安全访问。

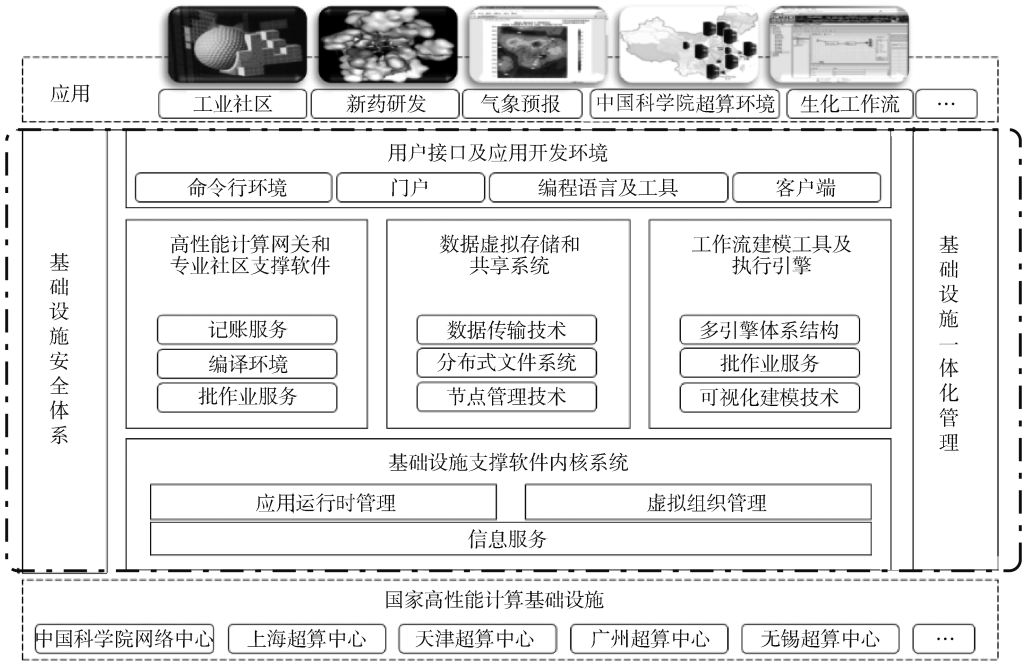


图 2　CNGrid Suite 系统架构

Fig. 2　CNGrid suite system architecture

2.2　资源组织与作业调度

针对国家高性能计算基础设施特点和应用需求,提出了“资源实体”、“虚拟组织”和“运行时虚拟地址空间”等系统软件创新概念,应对资源描述、组织和访问的挑战。

在 CNGrid 中,用户、资源和社区都被统一抽象为资源实体。每个资源实体由一个全局 id 来标识,代表一个可以访问其他实体,也可被其他实体所访问的全局资源要素。系统软件的全局命名管理模块对资源实体实施统一管控,完成资源实体创立与消除、资源定位、资源解耦等功能。

多个相关资源实体可构成一个虚拟组织。虚拟组织描述资源实体之间的静态关系,实现资源实体注册、资源实体接入与剔除、资源实体元信息管理、资源实体权限管理、资源实体访问控制等功能。通过虚拟组织把资源实体组织成可有效管控、相互协同的资源集合。

运行时虚拟地址空间描述了资源实体间的动态访问和调用关系,结合动态绑定的安全策略,解决了资源命名、资源视图、资源发现及定位、资源统一安全访问等基础性问题。与传统操作系统的进程概念相对应,CNGrid 提出了网程(Grip)概念。网程在运行时虚拟地址空间中代表资源实体动态访问其他资源,实施访问控制,分配、管理和回收资源,实现应用的可控启动和终止。

CNGrid 聚合了分布在不同地域、不同组织机构中的各种各样的高性能计算资源,面向用户提供统一的系统映像和透明的作业调度是基本需求。CNGrid 的作业调度由服务端、驱动器和客户端 3 部分组成[22]。面向用户提供统一的访问入口和使用方式。系统软件根据用户作业请求性质为其自动匹配适当的高性能计算资源。另外,也提供开放接口,为作业调度模型和调度策略的优化提供了可能。作业调度核心模块由资源收集器、资源匹配器和资源调度器构成,通过引入多种优先级作业队列,细化作业的系统状态,改善了作业调度策略的可配置性[23]。

2.3 监控管理和安全机制

针对资源的分布性、异构性和动态性,服务和应用的多样性及管理需求的各异性等特征,CNGrid提出了统一实体监控管理、管理功能动态构造、管理功能跨域动态部署及协同工作等创新概念与机制,设计了基于统一实体的监控管理体系架构,研发了一体化的监控管理系统(见图3),实现了CNGrid的资源监控和运行管理,为多层次资源的按需共享和自主协同提供了支撑。

与CNGrid系统软件中的资源实体抽象相对应,设计了基于实体的资源管理信息描述方法,采用统一的"被管对象"抽象建立全局信息模型,对各类资源信息进行有效的建模与表示,形成对不同层次、不同类别资源精确监控和管理的基础。

针对资源的动态性和多样性,提供了监控管理功能的动态生成、部署、运行的能力,支持管理功能的动态扩展和更新,实现监控管理系统的动态构造与演化。针对资源的跨域特点,提出了监控管理功能跨域动态部署的概念。基础设施的监控管理按分布层次式组织,监控管理功能分布在各个监控管理域中,各管理域既局部自治又相互协作。管理域设立自身的监控管理中心,形成多级监控管理中心的协同机制。同时,实现了基于复杂事件处理的监控信息高效获取、传输和控制的机制。这些措施有效减少了与监控管理有关的数据流量,降低了监控管理对基础设施正常应用业务的影响。

针对CNGrid环境下资源种类繁多、数目巨大的现状,设计实现了单维度、多维度及基于日志等多种故障扫描、识别和应对方法,能够准确定位故障,分析故障根因,及时通告故障事件并推荐应对的策略,为提高CNGrid的可用性、可靠性和可管理性提供了保障。

CNGrid的安全机制采用基于证书的身份认证和访问权限控制,系统软件基于代理证书实现用户认证和权限代理。首先定义访问控制的策略构建资源共享操作上下文,操作上下文包含用户在CNGrid中的身份信息(用户的代理证书)、用户所属虚拟组织及其所在组别,以及虚拟组织签发的资源访问令牌。资源提供者能够在虚拟组织中注册资源并对其进行持续的管理,通过向用户分配相应的权限,控制应用占用的资源并有效支持多个应用间的协同。在运行时,由网程维护用户身份并实施访问控制。当用户需要访问资源时,把自己的操作上下文从用户端传送到资源端,基于证书权限验证的结果控制资源的访问[24]。CNGrid中部署了证书的认证中心(CA),用户可通过CA的Web界面申请用户证书。CNGrid安全机制基于公钥基础设施(PKI),使用标准的X.509证书,提供用户和资源的双向认证。

CNGrid的安全机制在权限控制的前提下,尽可能地支持基础设施资源的共享。受控共享是CNGrid提出的一个重要概念。在受控共享机制下,只要访问控制权限允许,非属主用户和属主用户均可完成对资源的操作,此过程称为属主用户和非属主用户对资源的受控共享。

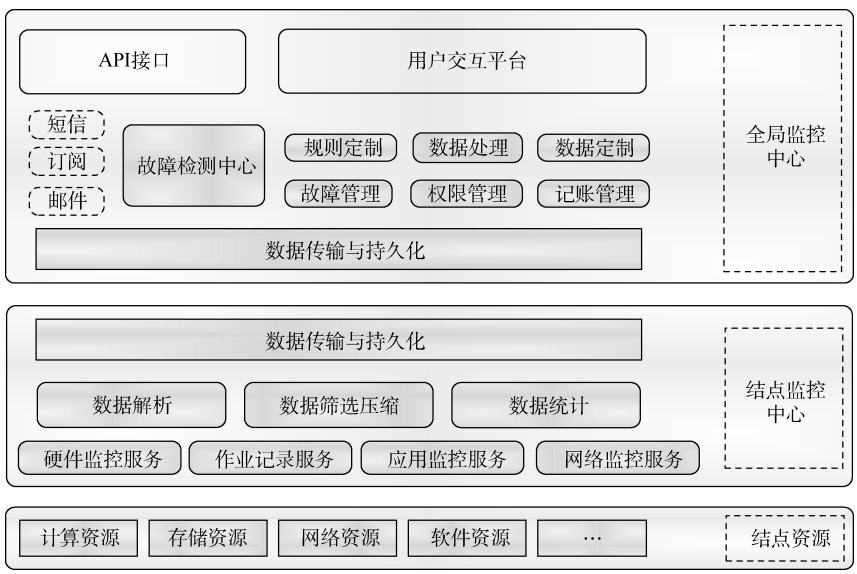


图3 一体化基础设施监控管理体系结构框架

Fig. 3 Architectural framework of integrated infrastructure monitoring and management

2.4　数据管理和高效传输

高效存储和访问分布、异构、自治的数据资源是 CNGrid 要解决的另一个关键问题。针对科学研究和行业应用的实际需求,设计并实现了基于虚拟数据空间的数据管理体系,有效集成了环境中的数据资源,构建了统一的数据管理空间,为用户提供了透明统一的数据存储、访问和管理能力。CNGrid 的数据资源主要包括文件系统和数据库系统。虚拟数据空间为文件系统的集成共享提供虚拟文件系统,为数据库系统的集成共享提供虚拟数据库系统。在这两者之上,虚拟数据空间提供数据基础服务,简化了存储和数据的使用逻辑,为用户或应用提供便利。

针对 CNGrid 中数据分布存储和自治管理等特点,虚拟数据空间采用面向服务的分布式层次结构进行构建。设计了基于分布域的联邦数据存储管理机制,在各数据域的自治管理基础上实现全局统一管理,保障数据管理的可扩展性。通过数据域之间的协作来满足应用的分布式存储需求,系统根据用户的访问位置等信息实现数据资源的就近存储和管理,以便提高用户对数据的访问效率。设计实现了异构数据库的整合机制,以统一的接口实现对不同数据库管理系统的数据访问,并且通过并行机制保障在大规模分布式环境下的访问效率。

影响 CNGrid 中数据传输效率的主要因素包括单次传输的数据量、网络带宽的利用率和传输引入的额外开销。因此,提高效率的关键在于减小单次传输的数据量,充分利用网络带宽,降低传输额外开销。数据传输不可靠的主要原因是网络链路及主机的不稳定,增强可靠性的关键在于克服不稳定因素,减少数据传输错误造成的损失。CNGrid 通过多个副本并行传输来提高带宽利用率,通过文件的分块传输来减少每次传输的数据量,通过就近传输来提高传输速度和可靠性,提供断点续传和三方传输来提高数据传输效率,减少额外开销。

CNGrid 环境的动态变化特性使数据存储资源难以保证持续的服务。CNGrid 引入数据副本管理机制来保证数据服务的可靠性。数据副本的引入带来数据一致性维护问题。为此设计实现了并行化的一致性有限状态机,有效降低了数据一致性维护的代价。数据按照其使用频度被定义为冷热数据,系统根据数据温度动态调整其副本数量,在提高访问效率的同时减少了不必要的开销。此外,还设计实现了位置及网络实时状态感知的数据副本放置策略,在保证一致性的同时提高了数据访问的效率。

CNGrid 的数据管理服务通过灵活和自适应的数据访问授权控制,解决了数据安全性与环境复杂动态性之间的矛盾。其采用细粒度的访问控制策略,为不同的资源拥有者和使用者对不同粒度的数据资源的访问,提供个性化的访问控制策略,满足了自治性和个性化的要求。

2.5　基于应用社区的应用新模式

针对以公共计算平台支撑个性化领域应用的需求,CNGrid 提出了体现领域应用特点的个性化领域应用社区概念。应用社区具有“批零”结合的资源管控与按需服务机制,既有网格聚合分散资源的能力,又有云计算集中管控、按用付费的特点,成为国家高性能计算基础设施的应用新模式。为了支撑应用社区的构建和运行,发展了领域应用中间件 Xfinity,实现了多层次的用户管理机制、按域划分的资源管理模式、基于模板的应用零开发热部署技术及资源动态绑定的工作流技术等体系架构和关键技术创新。

按需定制的服务模式体现在服务方式和内容的定制,可为不同用户定制满足其特定需求的专用社区。通过社区动态配置、资源动态绑定与复用、应用按需集成与动态部署等技术,实现了服务的按需定制。按需付费的交易模式贯穿服务交易全过程。资源拥有者在社区发布资源与价格信息,用户通过社区选择能满足其需求且价格合适的服务。社区监督服务交易过程和服务完成情况,保证交易各方的利益。

为实现按需调配的资源管理模式,提出了权属策略灵活配置的社区资源管理技术,将特定资源组织成资源子域授权给不同用户使用。不同资源子域的用户相互隔离,互不干扰。用户对其资源子域拥有完全的支配权,可做更精细的分级授权管理,实现社区内资源的有效调配和充分共享。

社区通过基于角色的权限访问控制、组管理和双层映射等技术实现了多层次、分角色的用户与资源的精细管理。实现了与企业业务系统相容的低开销安全机制,允许独立制定和修改国家高性能计算基础设施、社区、企业这 3 个管理域的安全机制,在管理域之间建立信任关系和映射机制,消解各管理域不同安全策略间的矛盾。

工业创新设计社区是该新型应用模式的一个实例,其系统框架如图 4 所示。工业社区将国家高性能计算基础设施的计算服务推送到汽车制造、

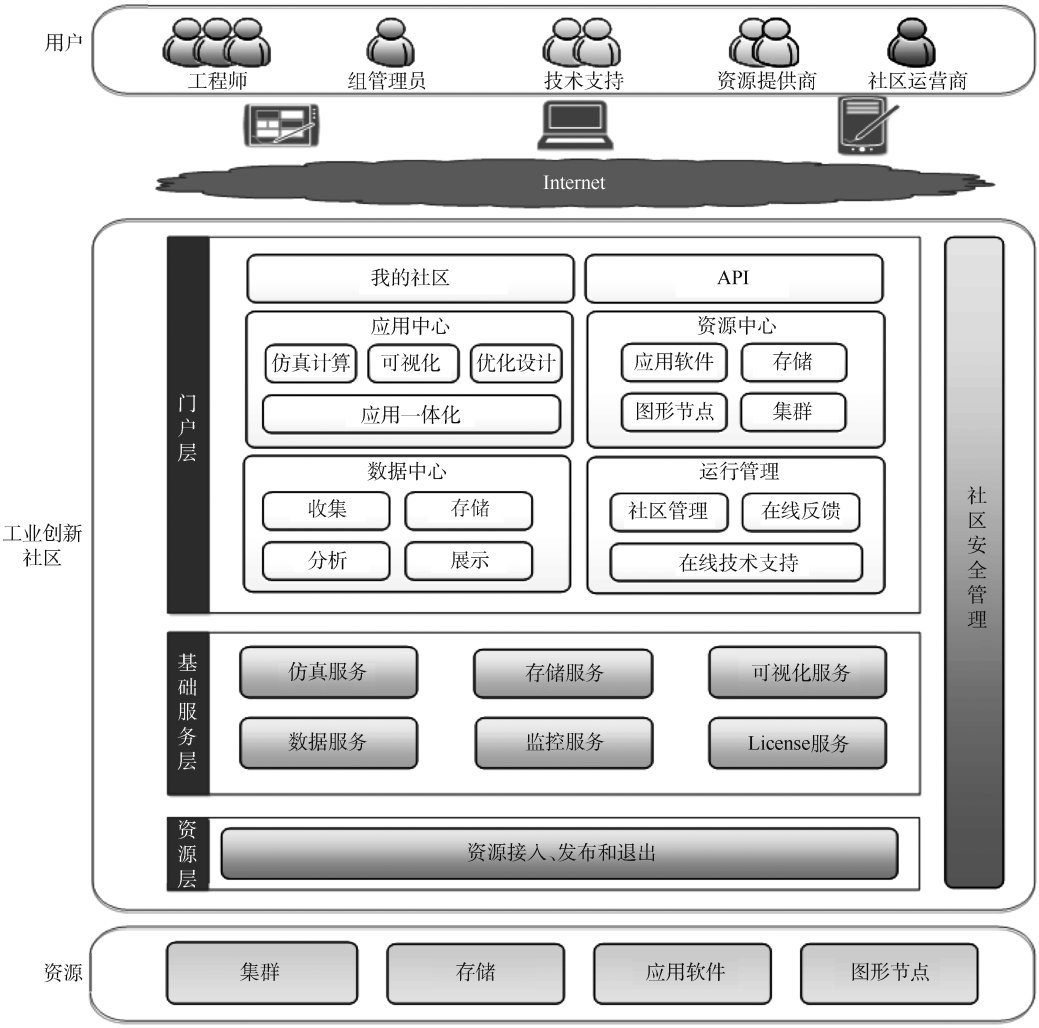


图4 工业社区系统框架

Fig. 4 Industrial community system framework

核电、飞机制造等行业的企业内部,加快了产品设计,降低了研发成本,提高了企业竞争力,取得很好的经济效益。

2.6 分布异构环境下应用软件的开发与优化

CNGrid的地理分布、资源异构的特征给大型应用的开发带来新的挑战。针对大规模应用的流程化与跨结点分布特征,CNGrid突破了构件与流程相结合的工作流编排、部署和运行技术,实现了流程在线组装、即时协作、即插即用的服务适配机制,支持分布资源的动态绑定、应用的快速开发和灵活部署执行,允许领域专业人员以低代码方式开发应用。

针对国产超级计算机多级并行和多种异构的特点,提出并实现了一系列并行程序优化方法和技术。例如,提出了节点间MPI、节点内OpenMP、处理器内多核并行的多级混合并行模式。提出了适用于不同异构平台的区域分解和动态负载平衡方法,通过动态可调的区域划分,实现加速器和通用处理器之间的负载均衡,隐藏加速器和通用处理器间的通信开销。提出了定制缓存及计算/访存重叠技术,充分发挥数据在片内核间的最佳重用。提出以定制DMA传输等方式实现计算和访存的最优化重叠,缓解内存带宽对应用整体性能的限制,大幅提升系统效率。研发了应用级断点保护技术,保证了大规模长时间作业的正确执行。发展了屏蔽硬件细节的并行算法库和编程接口,使不熟悉并行计算的应用领域专家能编写高效的并行应用软件。

在上述关键技术突破的基础上,研发了面向国家高性能计算基础设施的应用集成开发环境。集成开发环境包含基础算法库、应用模块库、程序模板库、优化工具库、拖拽式的工作流编排器、适配多种国产处理器的跨异构结点编译环境等,其系统架构如图5所示。开发人员可以使用集成开发环境中基于模板库的开发向导,自动生成程序代码框架,重用基本算法和模块库中的代码,快速构建应用程序,并在国家高性能计算基础设施中交互式地部署、优化和运行。

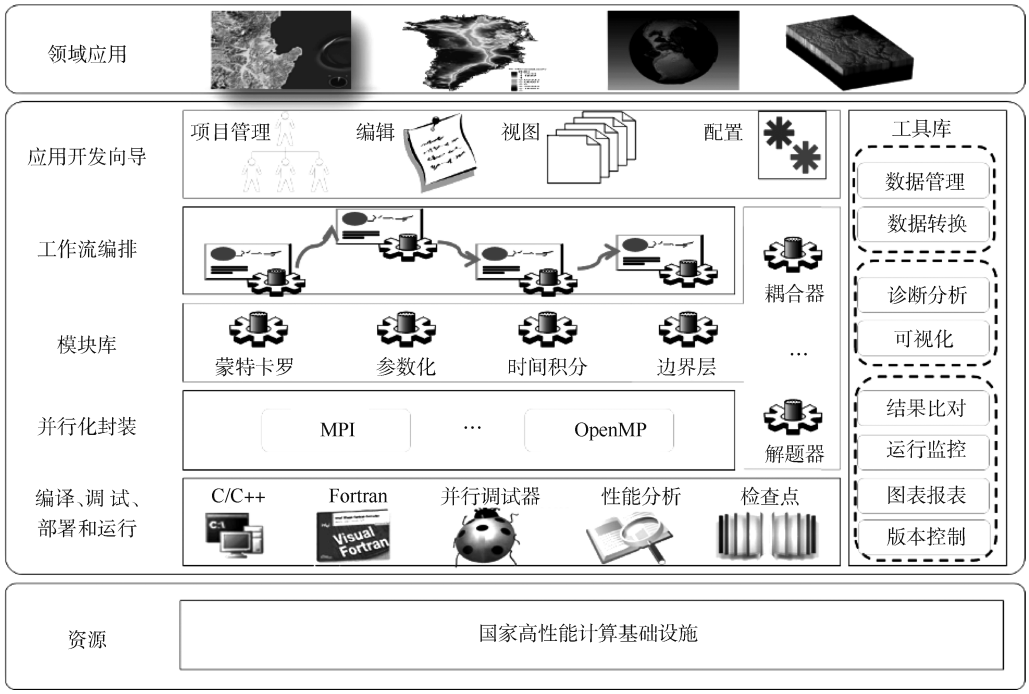


图 5　高性能计算应用集成开发环境系统架构

Fig. 5　System architecture of integrated development environment for high performance computing applications

3　算力基础设施发展趋势与展望

3.1　新兴应用和技术趋势

3.1.1　新兴应用及算力需求

近年来,一系列新兴技术与应用的快速发展对算力基础设施提出了更高的要求。其中,最具代表性的有人工智能、大数据和云计算等。

新一代人工智能的核心驱动力来自深度学习技术。通过对多层大规模人工神经网络进行训练并用于推理,促进了计算机视觉、自然语言处理等领域的突破性进展。虽然人工神经网络概念的出现和应用已有数十年,但之所以近年来才取得快速发展,离不开算力的支持。深度学习是一种计算和数据驱动的技术,由于深度神经网络规模庞大,且通常需要使用大量的数据进行训练,这带来了巨大的计算量。表 3[25] 给出了几种典型神经网络模型的训练计算量。庞大的计算量使得神经网络训练通常需要借助加速部件进行,即使这样,一次模型训练也需要花费数小时到数天时间。

表 3　典型深度神经网络模型的训练计算量[25]

Table 3　Training computations of typical deep neural networks[25]

神经网络模型	应用领域	训练计算量/Flops
AlexNet	图像分类	4.7×10^{17}
VGG16	图像分类	8.5×10^{18}
YOLOv3	图像目标检测	5.1×10^{19}
Transformer	自然语言处理	7.4×10^{18}
GPT-3	自然语言处理	3.1×10^{23}

注:数据来源于 https://docs.google.com/spreadsheets/d/1AAIebjNsnJj_uKALHbXNfn3_YsT6sHXtCU0q7OIPuc4。

大数据和云计算是另一种推动算力设施发展的新兴技术。在虚拟机和容器技术的支持下,人们可以在硬件平台上实现计算资源的灵活划分和隔离,以及软件环境的快速部署,这使得用户可以从云平台获得按需分配、可动态伸缩、易于部署且稳定可靠的硬软件平台和算力服务,这一特性吸引越来越多的用户将网站、业务平台和信息系统等迁移托管到云数据中心。由于企事业单位数量众多,此类应用的需求聚合到一起形成了对算力资源的庞大需求。

3.1.2　体系结构及算力设施的发展趋势

随着集成电路延续数十年的“摩尔定律”减缓并走向停滞,计算机体系结构进入了变革期,多样化的体系结构不断涌现。为了在集成电路规模和性能增长减缓的背景下持续提升应用性能,定制化(customization)成为近年来体系结构发展的一大特点,即通过设计面向不同应用的加速部件/处理器,持续提升应用性能,典型代表有 GPU 和深度学习处理器/加速器。在 GPU 方面,除了 Nvidia GPU 外,传统处理器厂商 AMD 和 Intel 也陆续推出自己的 GPU,国内也已研发出多款自主 GPU 芯片,为人工智能、科学计算、图形/图像处理等应用提供了高性能计算平台;在深度学习处理器/加速器方面,比较有代表性的有寒武纪、

Google TPU、华为昇腾等,这些处理器/加速器专为神经网络计算而设计,其性价比和能效均优于通用 CPU 和 GPU。

在多种新兴应用的推动下,各种处理器/加速器被应用于算力设施中,这带来了以下 2 方面的变化:

1) 算力中心内的异构化。在算力中心内部,异构已成为主流架构。表 4[26] 给出了 TOP500 超级计算机排行榜中前十位高性能计算机的体系结构。可以看出,10 台机器中仅有 1 台(富岳)采用同构架构,其他 9 台均为异构架构,除 CPU + GPU 结构外,中国的神威·太湖之光采用片内异构众核处理器,天河 2A 采用 CPU + 加速器结构。算力中心异构化的另一个体现是面向人工智能应用的异构体系结构,除了 CPU + GPU 结构外,CPU + 深度学习处理器/加速器结构也被智算中心广泛采用,如 CPU + Google TPU、CPU + 寒武纪、CPU + 华为昇腾等。

2) 算力中心的多样化(算力中心间的异构化)。传统的超算中心主要面向科学/工程计算,应用类型以并行数值模拟为主,主要特征是以双精度浮点运算为核心的计算密集型应用。与之相比,人工智能应用的计算类型主要是单精度/半精度浮点和定点运算,而大数据和云计算则以数据密集型应用为主。为了适应这些新兴应用的需求,算力中心的硬件配置也开始出现变化,出现了配置深度学习处理器/加速器、主要面向人工智能应用的智算中心,以及配置大容量内存和网络虚拟化设备、主要面向大数据和云计算应用的云算中心,同时,超算中心也开始支持人工智能和大数据应用。

表 4 TOP500 排名前十的高性能计算机(2022 年 6 月)[26]

Table 4 TOP10 in TOP500 high performance computing systems (June 2022)[26]

排名	系统	处理器/加速器	Linpack 性能/PFlops
1	Frontier	AMD 64C + AMD MI250X	1 102
2	Fugaku(富岳)	A64FX 48C	442.01
3	LUMI	AMD 64C + AMD MI250X	151.90
4	Summit(顶点)	IBM Power + Nvidia V100	148.60
5	Sierra(山脊)	IBM Power + Nvidia V100	94.64
6	Sunway TaihuLight(神威·太湖之光)	Sunway SW26010	93.01
7	Perlmutter	AMD 64C + Nvidia A100	70.87
8	Selene	AMD 64C + Nvidia A100	63.46
9	Tianhe-2A(天河 2A)	Intel Xeon + Matrix2000	61.44
10	Adastra	AMD 64C + AMD MI250X	46.10

注:数据来源于 http://www.top500.org。

3.2 算力基础设施面临的技术挑战

随着“东数西算”国家战略的实施,西部多个算力枢纽将建设算力中心,并面向东部经济发达地区提供算力服务。在这一背景下,如果由各个算力中心单打独斗,分散运营,则算力中心需投入人力物力自行发展用户,容易出现算力中心间的负载不均衡,导致算力碎片化和算力资源浪费;而在用户侧,由于各算力平台的硬件配置、软件资源、服务接口存在差异,也将给用户的软件开发和资源使用带来诸多不便。因此,通过将多个算力中心互联,向用户提供一站式、集成化的算力服务,形成覆盖全国的算力基础设施,对于提升算力资源利用效率和服务水平,促进国产软件和应用生态发展,支撑“东数西算”国家战略具有重要意义。

为了构建算力基础设施,需要解决算力中心异构化和多样化带来的诸多技术挑战,主要体现在以下几方面:

1) 计算任务在异构算力中心间的透明调度。算力基础设施必须具备的一项功能是:用户通过算力基础设施的服务平台提交一个计算任务后,可以直接得到计算结果,无需关心该计算任务在哪个算力中心上运行。这需要按照计算任务的类型和需求确定其所需资源,并根据各个算力中心的硬软件配置及可用资源数量进行任务分配和调度。例如,如果用户提交的是一个使用 CUDA 编写的可执行程序,就要在各算力中心寻找配置 Nvidia GPU 的节点并获取其当前使用状态,在此基础上进行调度;而如果用户只是希望在指定的数据集上完成深度神经网络模型训练,或者只是希望对所设计零件进行结构强度分析,那么计算任务的分派调度就不但要考虑硬件资源,还要考虑是否具备所需的软件资源。在以上基本调度功能的基础上,如果进一步考虑用户的服务质量和计费需求,如限定任务完成时间、限定资费水平等,任务调度要考虑的因素就更多。

2) 如何提供多层次、多样化的算力服务。算力服务的层次与云计算服务相似,也可分为基础设施(IaaS)、平台(PaaS)、应用软件(SaaS)3 个层次。在基础设施层,算力服务以处理器/加速器/计算节点的形式提供;在平台层,算力服务在基础设施之上还提供系统软件和支撑软件,如高性能计算用户需要的 MPI 环境和基础算法库,人工智能用户需要的深度学习框架,大数据用户需要的分布式处理框架等;在应用软件层次,用户则可以直接使用不同种类的应用软件和服务。在算力基础设施中,算力中心的异构化和多样化使得算力服务变得更加复杂。以平台层为例,即使是高性能计算所需的 MPI 环境和基础算法库,在不同的算力中心中也会有不小的差异,用户在不同算力中心中使用这些服务时往往需要进行适配和修改(如修改作业脚本等)。由于算力基础设施需要向用户提供透明、一致的算力服务,如何屏蔽这些差异,就是需要着力研究和解决的问题。

3) 多算力中心分布式协同计算与虚拟超级计算机。在过去,由于算力中心间的网络带宽有限,计算任务通常分配在单个算力中心上完成,即使进行跨中心计算,一般也仅限于工作流中的不同计算阶段。近年来,随着网络基础设施和算网融合技术的发展,算力中心间的网络互联带宽不断提升,传输延迟显著下降。国内已有超算中心间网络互联达到传输带宽10 Gbps、延迟接近 1 ms 的水平,这给多个算力中心进行分布式协同计算提供了可能性。这一方面需要算力中心基础支撑软件互联互通构成分布式计算环境,另一方面,还需要研究在这种环境中的任务划分、调度和迁移等技术。更进一步,在互联带宽和延迟满足要求的情况下,是否可能将多个超算中心互联成为统一管理、统一调度的单一计算系统,形成可完成数倍于原有规模并行计算的“虚拟超级计算机”,也是值得研究和探讨的问题。

4) 多样化算力中心和异构化体系结构的编程问题。异构体系结构显著增加了并行编程的复杂性,该问题在多样化算力中心场景下更加突出。为了支持异构处理器/加速器编程,厂商推出了相应的编程语言/接口,如用于 Nvidia GPU 的 CUDA、用于申威众核处理器的 Athread、用于 AMD GPU 的 ROCm/HIP 等,并基于这些编程接口开发/移植了各种基础算法库、求解器、深度学习框架等,但异构平台编程的复杂度仍然远高于传统的 CPU 平台。在多样化算力中心中,异构硬件平台的种类更多,为使程序具有更好的平台适应性,在软件编程模型和语言方面还需要进行更多的工作。虽然近年来已经出现了一些独立于厂商的加速器编程接口,如 OpenCL、OpenAcc、SYCL 等,但这些编程接口在不同硬件平台上的实现仍然有差异,为一种平台编写的程序通常难以不加修改地在另一种平台上编译和运行。有鉴于此,仍然有必要提出独立于硬件平台且可屏蔽硬件细节的编程模型/语言,与此同时,研究开发异构程序转换工具,实现并行程序在不同硬件平台间的透明转换和自动编译,进而支持并行程序在多样化算力中心的透明调度和运行,也将是一项很有价值的工作。

5) 数据在分布式算力中心间的放置问题。无论是科学工程计算,还是人工智能或大数据应用,其数据规模都较为庞大。由于算力中心间的数据传输和访问开销较大,在算力中心间进行计算任务调度和迁移时,数据放置就成为必须考虑的一个重要因素。

6) 公共算力中心的数据安全和隐私问题。在人工智能领域,为了满足数据隐私和安全需求,提出了联邦学习技术(federated learning)[27],通过多个数据拥有者协同完成训练,避免了数据向其他实体公开。这种数据隐私和安全性需求同样存在于大数据分析、科学与工程计算等领域。为了使用公共算力中心的计算服务,用户往往需要将数据上传至算力中心,虽然通过 VPN 等技术可以保证数据在网络中传输时的安全性,但在多用户共享的算力中心中,数据在外存中的存放,以及计算过程中数据在内存的存放,仍然存在数据外泄的可能性。为满足对数据安全性要求较高用户的需求,如何在大数据分析和科学工程计算领域提供类似于联邦学习的机制,或实现“可计算但不可读写”,是值得深入研究的问题。

3.3　中国超算应用生态存在问题和算力基础设施未来展望

3.3.1　中国超算应用生态存在的问题

近年来,中国高性能计算技术水平取得了长足进步,超级计算机研制水平已处于国际前列,拥有性能排名前列的超级计算机,生产和部署的高性能计算系统数量也世界领先。在高性能计算应用软件方面,面向国产超算系统研发了一批重点行业/领域应用软件,取得了众多的应用成果,大规模并行算法及应用也 2 次获得代表国际超算应用最高水平的 Gordon Bell 奖,但总体上,高性能计算软件与应用的发展相对不足,应用生态也不够丰富。产生这种现象的原因有多个方面:

1）软件和应用研发投入不足。中国科研领域和产业界长期存在着“重硬轻软”的现象,在高性能计算领域,国家投入经费的大部分都用于高性能计算机系统研制,软件和应用处于配合和支撑地位,与发达国家超算研究计划中硬软件投入接近 1:1 相比,中国的软件和应用研发投入明显不足。

2）软件和应用种类多、研发持续时间长。软件和应用种类繁多,应用生态的建立需要长期持续的努力。一种新型处理器/加速器推出后,初期往往只具备操作系统和编译等核心软件,而单靠硬件研发单位完成多种编程语言编译器、调试及性能分析工具、基础算法库、求解器、各种领域应用软件的研发非常困难,需要多方参与,经过若干年的持续努力,逐步研发和完善。

3）用户使用习惯不利于国产软件的推广。中国超算应用软件研发起步相对较晚,西方国家在很多行业和领域已推出了商业化软件,用户已形成了商业软件使用习惯和对商业软件的认知度,某些行业甚至只认可某种软件的仿真结果,这给国内应用软件的自主研发和推广应用带来了很大困难。一种软件研发完成后,需要通过推广应用来支撑驱动软件维护和持续升级,通过软件持续升级,不断增强功能并改善用户体验,进而吸引更多用户使用,然而,目前国产超算软件研发还未能形成这种良性循环。

软件和应用生态存在的不足使得中国高性能计算领域存在着“大而不强”的现象,与此同时,高性能基础和应用软件大量依赖国外软件,也存在“卡脖子”的风险。

3.3.2 中国算力基础设施的发展展望

“东数西算”战略的实施将形成算力中心建设的新高潮,为了构建国家算力基础设施,需要在研发突破关键技术的基础上,补足中国超算软件与应用的短板,并通过运营模式和机制创新,建立起丰富且自我发展的国产软件应用生态。为此,需要在以下方面开展重点工作:

1）研究突破关键技术,支撑算力基础设施发展。围绕算力中心异构化和多样化带来的技术挑战,解决算力基础设施面临的技术难题,研发核心软件和服务平台,实现多样化算力中心的互联互通、资源共享和服务提供,为国家算力基础设施的构建和发展提供技术支撑。

2）强化计算软件研发,补足国产软件与应用短板。以国产处理器/加速器的兴起为契机,加强基于国产硬件的工具链、算法库、求解器、领域应用等基础和应用软件研发,通过若干年的持续努力,建立较为完备的国产硬件支撑和应用软件栈,形成可自我发展的国产软件与应用生态。

3）改变单一机时服务方式,推动算力中心能力建设。国内已建成的超算中心为了弥补运行经费的不足,普遍以提供机时服务为主,即俗称的“卖机时”。这种低层次的算力服务消耗了较多的人力和精力,制约了算力中心向更高水平发展。通过建立国家算力基础设施,可以推动算力中心从机时提供者向应用研发者和解决方案提供者转变。同时,各超算中心通过研发建立领域应用平台,可以突出自己的技术特色,进而形成算力中心各有所长的态势,也可以避免算力中心发展同质化。

4）创新算力运营模式和机制,打造多方共赢的应用生态。算力中心涉及地方政府、投资方、设备提供方和运营方,在构建国家算力基础设施的过程中,需要以多方共赢为目标,通过运营模式和机制创新,鼓励多方参与,通过竞争促进技术进步和服务水平提升,与此同时,以应用商店(App store)等模式打造研发、服务、运营等多方共赢的软件和应用生态,进而推动国家算力基础设施做大做强。

4 结束语

CNGird 在国家科技计划支持下历经 20 余年发展,已经成为不可或缺的国家高性能计算基础设施,并为“东数西算”背景下国家算力基础设施的建设积累了宝贵经验,奠定了技术基础。温故而知新。面对“东数西算”的新任务,要认真总结 CNGrid 建设的历史经验,分析应用和技术发展的新趋势,定位亟待解决的瓶颈技术问题,探索新的应用模式和机制,更高效地构建新一代国家算力基础设施,实现“东数西算”的国家战略。这是国家创新发展赋予的历史使命。

参考文献(References)

[1] DENNIS J. Segmentation and the design of multiprogrammed computer systems[J] Journal of the ACM, 1965, 12(4): 589-602.

[2] SACKMAN H. Time-sharing versus batch processing: The experimental evidence[C]// Proceedings of the American Federation of Information Processing Societies. New York: ACM, 1968:1-10.

[3] SCHWARTZ J, COFFMAN E, WEISSMAN C. A general-purpose time-sharing system[C]// Proceedings of the American Federation of Information Processing Societies. New York:

ACM,1964:397-411.

[4] MILLS D L,BRAUN H. The NSFNET backbone network[C]// Proceedings of the ACM Workshop on Frontiers in Computer Communications Technology. New York:ACM,1987:191-196.

[5] FOSTER I T,KESSELMAN C. The grid:Blueprint for a new computing infrastructure [M]. San Francisco:Morgan Kaufman Publishers,1998.

[6] STEVENS R,WOODWARD P,DEFANTI T,et al. From the I-WAY to the national technology grid[J]. Communications of the ACM,1997,40(11):50-60.

[7] THOMAS M,BOISSEAU J,DAHAN M,et al. Development of NPACI grid application portals and portal Web services[J]. Cluster Computing,2003,6(3):177-188.

[8] FOSTER I,CZAJKOWSKI K,FERGUSON D,et al. Modeling and managing state in distributed systems:The role of OGSI and WSRF[J]. Proceedings of the IEEE,2005,93(3):604-612.

[9] TALIA D. The open grid services architecture:Where the grid meets the Web[J]. IEEE Internet Computing,2002,6(6):67-71.

[10] FOSTER I,KESSELMAN C. Globus:A metacomputing infrastructure toolkit[J]. International Journal of Supercomputer Application,1998,11(2):115-129.

[11] REED D. A. Grids,the TeraGrid,and beyond[J]. IEEE Computer,2003,36(1):62-68.

[12] KUNSZT P. European DataGrid project:Status and plans[J]. Nuclear Instruments and Methods in Physics Research Section A:Accelerators,Spectrometers,Detectors and Associated Equipment,2003,502(2-3):376-381.

[13] GAGLIARDI F,JONES B,GREY F,et al. Building an infrastructure for scientific grid computing:Status and goals of the EGEE project[J]. Philosophical Transactions of the Royal Society A:Mathematical,Physical and Engineering Sciences,2005,363(1833):1729-1742.

[14] HEY T,TREFETHEN A E. The UK e-Science core programme and the grid[J]. Future Generation Computer Systems,2002,18(8):1017-1031.

[15] MATSUOKA S,SHINJO S,AOYAGI M,et al. Japanese computational grid research project:NAREGI[J]. Proceedings of the IEEE,2005,93(3):522-533.

[16] ARMBRUST M,FOX A,GRIFFITH R,et al. Above the clouds:A Berkeley view of cloud computing:UCB/EECS-2009-28[R]. Berkeley:EECS Department University of California,Berkeley Technical Report,2009.

[17] SARASWAT M,TRIPATHI R C. Cloud computing:Analysis of top 5 CSPs in SaaS,PaaS and IaaS platforms[C]// 2020 9th International Conference on System Modeling and Advancement in Research Trends,2020:20421390.

[18] SOTOMAYOR B,MONTERO R,LLORENTE I,et al. Virtual infrastructure management in private and hybrid clouds[J]. IEEE Internet Computing,2009,13(5):14-22.

[19] BARIK R,LENKA R,RAO K,et al. Performance analysis of virtual machines and containers in cloud computing[C]// 2016 International Conference on Computing,Communication and Automation. Piscataway:IEEE Press,2016:16585534.

[20] SIMONS J. HPC cloud bad;HPC in the cloud good[C]// 2013 IEEE 27th International Symposium on Parallel and Distributed Processing. Piscataway:IEEE Press,2013:13683523.

[21] MOR N. Edge computing:Scaling resources within multiple administrative domains[J]. Queue,2018,16(6):106-116.

[22] 乔健,查礼. 中国国家网格作业管理设计与实现[J]. 计算机应用,2008,28(8):2003-2009.
QIAO J,ZHA L. Design and implementation of grid job management for China national grid[J]. Computer Applications,2008,28(8):2003-2009(in Chinese).

[23] 王小宁,肖海力,曹荣强. 面向高性能计算环境的作业优化调度模型的设计与实现[J]. 计算机工程与科学,2017,39(4):619-626.
WANG X N,XIAO H L,CAO R Q. Design and implementation of an optimal job scheduling model for the high performance computing environment[J]. Computer Engineering & Science,2017,39(4):619-626(in Chinese).

[24] 喻林,邹永强,查礼. CNGrid GOS 安全:设计与实现[J]. 华中科技大学学报(自然科学版),2010,38(S1):6-10.
YU L,ZOU Y Q,ZHA L. CNGrid GOS security:Design and implementation[J]. Journal of Huazhong University of Science & Technology (Natural Science Edition),2010,38(S1):6-10(in Chinese).

[25] SEVILLA J,VILLALOBOS P,C ERON J,et al. Parameter,compute and data trends in machine learning[EB/OL]. [2022-05-30]. https://docs.google.com/spreadsheets/d/1AAIebjNsnJj_uKALHbXNfn3_YsT6sHXtCU0q7OIPuc4.

[26] TOP500 list[EB/OL]. [2022-06-20]. https://top500.org/lists/top500/2022/06/.

[27] BONAWITZ K,EICHNER H,GRIESKAMP W,et al. Towards federated learning at scale:System design[C]// Proceedings of the Conference on Machine Learning and Systems. Piscataway:IEEE Press,2019:1-15.

From grid to "East-west Computing Transfer": Constructing national computing infrastructure

QIAN Depei*, LUAN Zhongzhi, LIU Yi

(School of Computer Science and Engineering, Beihang University, Beijing 100083, China)

Abstract: This article gives a review of the evolution of computer use-mode over the time since the invention of modern computers and presents the challenges and tasks in building the national computing infrastructure. The first section of this article provides a brief overview of how computer use-mode has evolved over the past several decades. Then the design and implementation of China's national high performance computing infrastructure CNGrid are introduced. Following that, the trends and new technical challenges in developing the computing infrastructure under the circumstance of the national strategic project of "East-west Computing Transfer" are discussed. Finally, the perspectives of building the supercomputing eco-system and constructing the new type of computing infrastructure in China are presented.

Keywords: high performance computing; infrastructure; grid computing; CNGrid; East-west Computing Transfer

Received: 2022-08-02; **Accepted**: 2022-08-18; **Published online**: 2022-08-23 09:07

URL: kns.cnki.net/kcms/detail/11.2625.V.20220822.1844.002.html

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0318

材料基因工程加速新材料设计与研发

孙志梅[1,*], 王冠杰[1,2], 张烜广[1], 周健[1]

(1. 北京航空航天大学 材料科学与工程学院, 北京 100083;
2. 北京航空航天大学 集成电路科学与工程学院, 北京 100083)

摘 要: 在未知材料化学成分和性能关系的情况下,通过传统的"试错-纠错"方法研发具有特定功能的新材料成本高且经常失败。随着人工智能和数据驱动的第四科学范式的发展,材料基因工程(MGE)已经成为材料设计与研发的新模式。综述了材料基因工程中高通量计算、材料数据库和人工智能方法的研究进展。介绍了材料高通量计算常用的框架和方法;阐述了材料数据库在材料数据类型和数据标准两方面的发展现状和有待解决的难题;总结了人工智能方法在材料关键基础问题中的应用。从高通量可视化计算方法、材料多类型数据库和可视化机器学习框架三方面重点证述了自主开发的多尺度集成可视化的高通量自动计算和数据管理智能平台 ALKEMIE。展望了材料基因工程未来的发展趋势。

关 键 词: 材料基因工程(MGE); 高通量计算; 材料数据库; 机器学习; 智算平台
中图分类号: TB39
文献标志码: A **文章编号**: 1001-5965(2022)09-1575-14

先进材料是国家工业的支柱,在传统"试错-纠错"研发模式下,材料学家基于自身知识储备和认知能力,通过反复迭代的试错-纠错改进材料性能,实现新材料的设计与研发。随着新一轮工业革命和互联网时代的到来,新材料的研发速度严重滞后于材料性能需求速度,按需逆向设计和精准控制性能已成为新材料设计的必然趋势。20世纪末,美国兴起组合材料学(combinatorial materials science, CMS)[1],通过并行合成和高通量表征技术,实现了新材料的快速制备和筛选;21世纪初,美国和欧洲部分国家提出的集成计算材料工程(integrated computational materials engineering, ICME)[2]将不同时间尺度和空间尺度的多种材料模拟方法相结合,在新材料设计领域取得了突破。如今,随着计算机和人工智能技术的飞速发展,材料基因工程(materials genome engineering, MGE)被视为实现材料科学技术飞跃和新材料高效研发与设计的基础,是新材料研发的加速器。

材料基因工程是受人类基因组计划(human genome project, HGP)的启发而建立的。在生物学中,基因是一组编码信息,被视为生物体生长和发育的蓝图,而在材料领域,基因可被看作决定其宏观性能的微观特征单元。作为基于数据驱动的科学发展第四范式,材料基因工程将高通量计算和设计、高通量制备、高通量表征、材料数据库和人工智能相结合,大大缩短了材料研发周期、降低了研发成本,从而快速研发出满足日益增长的性能需求的新材料。

2002年,美国宾夕法尼亚州立大学 Liu[3-4] 首次提出材料基因的概念,并于2005年设立材料基

收稿日期: 2022-05-06; **录用日期**: 2022-05-18; **网络出版时间**: 2022-06-01 15:07
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220531.1837.003.html
基金项目: 国家自然科学基金(51872017); 中国博士后科学基金(2022TQ0019); 北航高性能计算平台
* **通信作者**. E-mail: zmsun@buaa.edu.cn

引用格式: 孙志梅, 王冠杰, 张烜广, 等. 材料基因工程加速新材料设计与研发[J]. 北京航空航天大学学报, 2022, 48(9): 1575-1588. SUN Z M, WANG G J, ZHANG X G, et al. Novel material design and development accelerated by materials genome engineering[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1575-1588(in Chinese).

因基金会。2011 年,美国正式发布提升美国全球竞争力的材料基因组计划(materials genome initiative, MGI)[5],确立了面向未来的集成计算、实验和数据库的材料研发新模式。MGI 构建了包含上百万条先进能源材料的计算模拟结果材料数据库,并提出了计算机辅助材料研发、模块化的材料模拟体系、开放式的材料高性能数据库以及多尺度计算融合等研究方向。欧盟也相继提出了“新材料发现 NOMAD”计划、德国推出了工业 4.0 战略、俄罗斯推出“2030 年前材料与技术发展战略”、中国提出了“材料基因工程”等一系列政策将新材料探索和材料创新设计与研发作为首要发展目标。

本文首先介绍了国内外材料基因工程领域常用的高通量计算模拟软件和框架。其次,从材料数据来源、多类型数据库和数据标准方面介绍了目前常用的材料数据库。然后,总结了机器学习方法在材料学中的热点应用,重点介绍了笔者团队自主开发的多尺度集成可视化的高通量自动计算和数据管理智能平台 ALKEMIE 研究进展。最后,总结提出了材料基因工程未来的重点发展方向。

1 材料高通量计算方法

科学发展经历了如图 1 所示的 4 个过程:从文艺复兴时期实验主导的经验范式、以经典力学和热力学为主导的理论模型范式、基于分子动力学和密度泛函理论的计算科学范式,到如今的大数据驱动的科学研究范式[6]。近年来,美国、欧洲、日本和中国等国家科研人员在数据驱动研发模式的推动下,开发了一系列用于材料计算基础设施的高通量计算框架和“即插即用”功能完善的高通量计算软件,如表 1 所示。2011 年,在美国国家科学技术委员会、能源部和教育部支持下,加利福尼亚大学伯克利分校劳伦斯伯克利国家实验室 Jain 等[7]主导开展了材料基因组项目(materials project, MP),该项目开发了 4 款分别用于材料建模、材料计算模拟、自动纠错和服务器部署的高通量计算分析软件。其中,FireWorks 用于构建材料高通量计算模拟框架,解决高通量材料计算过程中多任务间的依赖关系和任务间的参数及数据传递;Custodian 用于高通量计算过程中的自动纠错;Pymatgen 通过抽象的 Python 对象解析材料结构对称性,自动分析不同尺度计算模拟软件的输入参数和结果,工作原理如图 2 所示;Atomate 则实现了完整的高通量流程及数据存储和服务器的配置,如图 3 中能带计算工作流所示。该项目以其独特的命令行操作方式为高通量计算奠定了软件基础。

2012 年,同属美国材料基因组项目的杜克大学 Curtarolo 等[8]基于 Python2 开发了适用于第一性原理计算的高通量计算软件 AFLOW-π。该软件针对高通量第一性原理计算,集成了数据实时反馈、错误控制、数据管理和归档等功能,可用于实现能带结构、态密度、声子色散、弹性特性、复介电常数、扩散系数等高通量计算,并针对性地优化了紧束缚的哈密顿量(tight-binding Hamiltonians, TBH)计算和数据分析流程。

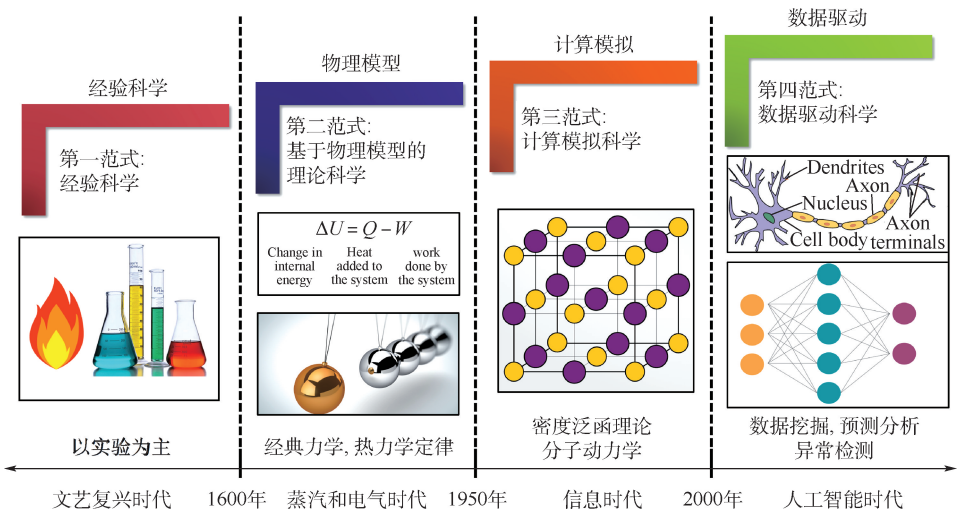


图 1 科学发展的四个范式[6]

Fig. 1 Four paradigms of science[6]

表 1　材料高通计算软件和框架发展现状

Table 1　Software and frameworks of material high-throughput calculation

名称	说明	网址
ALKEMIE	高通量计算框架和软件	https://alkemine.org/
Pymatgen	高通量计算框架	http://pymatgen.org/
FireWorks	高通量计算框架	https://materialsproject.github.io/fireworks/
Atomate	高通量计算框架和软件	https://atomate.org/
AFLOW-π	高通量计算软件	http://aflowlib.org/scr/aflowpi/
ASE	高通量计算框架和软件	http://wiki.fysik.dtu.dk/ase/
AiiDA	高通量计算框架和软件	http://aiida.net/
Abipy	高通量计算软件	https://github.com/abinit/abipy
MPInterfaces	高通量计算框架	http://henniggroup.github.io/MPInterfaces/
Imeall	高通量计算软件	https://github.com/Montmorency/imeall
MedeA	高通量计算软件	https://www.materialsdesign.com/
IprPy	高通量计算框架	https://www.ctcms.nist.gov/potentials/iprPy/
Pylada	高通量计算框架	http://pylada.github.io/pylada
MIP	高通量计算软件	http://www.mip3d.org/
MatCloud	高通量计算软件	http://matcloud.cnic.cn/
JAMIP	高通量计算框架	http://jamip-code.com
MatAi	高通量计算框架	https://www.mat.ai/
中国材料基因工程高通量计算平台	高通量计算框架	http://mathtc.nscc-tj.cn/

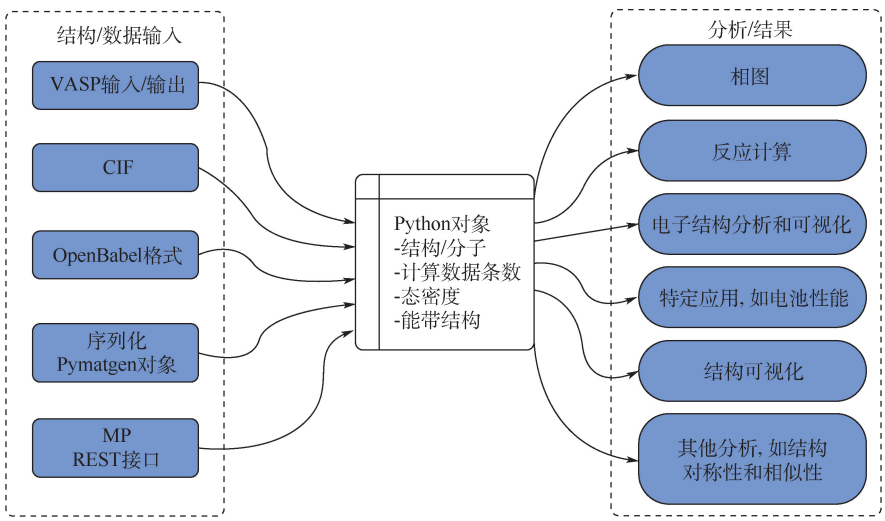


图 2　Pymatgen 高通量软件工作原理[7]

Fig. 2　Principle of Pymatgen high-throughput software[7]

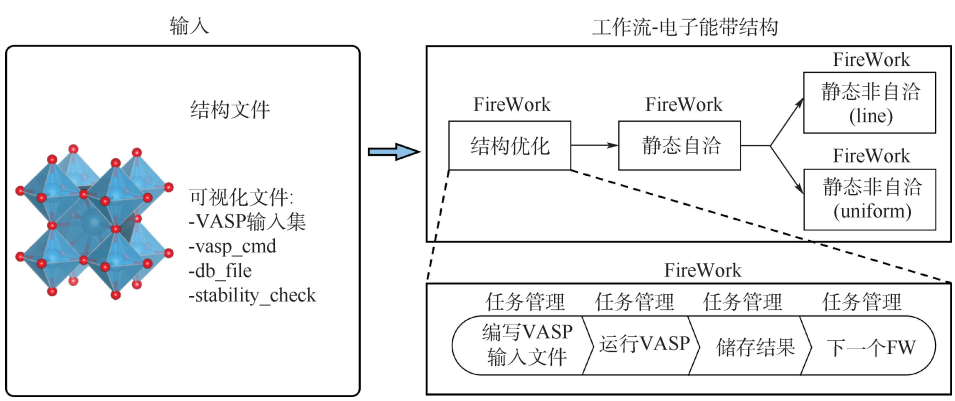


图 3　Atomate 中高通量能带计算流程[7]

Fig. 3　High-throughput computational workflow of band structure in Atomate[7]

丹麦科技大学 Larsen 团队[9]开发了材料批量化计算平台原子模拟环境(atom simulation environment, ASE),该软件由于缺少工作流程、计算参数和结果的自动纠错功能,并非完整意义上的高通量计算。而随着版本的迭代,科研人员为其进一步添加了可视化的用户界面、多个材料软件计算器、多种算法的分子动力学计算软件和多种晶体结构优化算法及边界条件,可以满足不同用户不同功能的计算需求。目前,该平台包含了 ABINIT、CASTEP、CP2K、VASP 和 LAMMPS 等常用的材料计算模拟软件。

2016 年,瑞士洛桑联邦理工大学 Pizzi 等[10]开发了高通量计算引擎(automated interactive infrastructure and database for computational science, AiiDA),该软件系统基于自动化、数据库和开源共享理念,开发了支持数万个材料计算任务并发运行的高通量算法。材料科学家不仅关注计算模拟的输入和输出,更关注计算模拟过程中的精度及构型的变化是否准确,因此,该软件保存了材料高通量计算中的子任务依赖关系,并自动跟踪记录所有计算和工作流程的输入、输出和中间元数据,以便在其开放式数据库 Materials Cloud 中查询数据。

基于 MP 发展的材料高通量计算基础框架,比利时天主教鲁汶大学 Gonze 等[11]开发了基于多体微扰论的第一性原理高通量计算软件 Abipy;美国佛罗里达大学 Mathew 等[12]发展了针对二维材料表面和异质结的高通量计算流程 MPInterfaces(见图 4);英国伦敦国王学院 Lambert 等[13]发展了用于原子晶界的高通量计算框架 Imeall 等。

相比于国外材料基因工程的研究成果,中国高通量计算起步较晚,但在 2016 年《中华人民共和国国民经济和社会发展第十三个五年规划纲要》中材料基因国家重点专项支持下,中国也涌现了多个成熟的材料高通量计算框架和软件。

多尺度集成可视化的高通量自动计算和数据管理智能平台(artificial learning and knowledge enhanced materials informatics engineering, ALKEMIE)是由笔者团队 Wang 等[14-15]基于 Python 开源框架自主开发的中国第一个高通量自动流程可视化计算和数据管理智能平台。该平台从设计出发吸取国外材料基因相关软件的先进理念,克服了计算过程中可能遇到的兼容性差、接口不统一和功能拓展困难等问题,开发了包含材料高通量自动计算模拟、材料数据库及数据管理、基于人工智能和机器学习的材料数据挖掘 3 个核心理念的智能平台。ALKEMIE 平台适用于数据驱动的材料研发,详细内容见第 4 节。

MIP(materials information platform)是由上海大学 Yang 等[16]开发的适用于热电材料高通量筛选的高通量计算软件;MatCloud 是由北京迈高材云科技有限公司 Yang 等[17]开发的第一性原理计算引擎,目前支持 VASP 和 ABINIT 等第一性原理高通量计算软件;JAMIP 是由吉林大学 Zhao 等[18]开发的开源高通量集成软件,该软件利用人工智能算法在高通量计算的海量材料数据中智能寻求新材料和新原理。中国材料基因工程高通量计算平台(CNMGE)则是由国家超级计算天津中心开发的网页版高通量集成计算平台,该平台可以集成多种不同的高通量计算软件,包括 ALKEMIE 高通量智算平台、含能材料分子专用高通量筛选系统 EM-Studio 和用于无机骨架材料的晶体结构解析与预测软件(framework generator, FraGen)。

上述高通量计算软件和框架,一方面为科研人员提供了快捷方便的自动计算工作流,可以在高性能超算中实现高效并行计算,显著提升计算效率;另一方面避免了传统试错-纠错法中可能出现的人为误差,使得研究人员有更多的精力关注材料科学问题本身,而非高通量所涉及的技术难题。近年来,高通量计算模拟已经在材料构型预测、材料结构稳定性和相稳定性预测、能源材料能量转化效率、最优掺杂元素分级筛选等方面获得了广泛应用。举例来说,Curtarolo 等[19-20]通过高

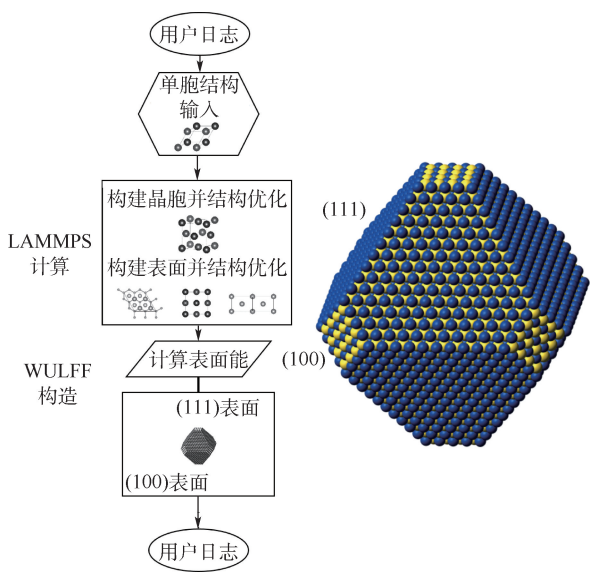


图 4 二维材料表面和异质结的高通量计算软件 MPInterfaces 工作流程[12]

Fig. 4 Workflow for 2D material surfaces and heterojunctions in MPInterfaces software[12]

通量方法从 435 个含 d 电子的二元金属间化合物中筛选出了 283 个能量稳定化合物,其中 273 个(96.5%)化合物最终获得了实验验证;Hu 等[21]基于第一原理计算的层级筛选,从 29 个过渡金属元素中筛选出可以提高相变材料 Sb_2Te_3 性能的最佳掺杂元素 Y、Sc 和 Hg,除了有毒的 Hg 元素,Y 和 Sc 均获得了实验验证;Gan 等[22]基于第一原理计算的高通量分级筛选,通过结构能量、声子谱、力学稳定性和转换效率等 4 个筛选标准从 21 060个候选光电材料中筛选出了 78 个稳定化合物,且其中 22 个化合物的性能优于当前太阳能电池材料 GaAs。

2　材料数据库研究进展

高通量计算大幅提高了计算模拟的效率,并产生了海量的数据,这些数据中既包含了有用的材料性质数据,又包含了大量重复的无效数据。而材料学中由于所研究的材料体系、成分、结构等的不同,材料测试、制备工艺和流程也不尽相同。对于不同用途材料,所关注的材料性能和关键指标也有差异,因此,在数据库的构建过程中面临数据存储类型、数据库的兼容性和泛化能力等一系列问题。本文根据材料数据来源,将数据大致分为材料计算数据、材料测试数据和已发表的文本数据三大类。本节将介绍目前国内外常用的材料数据库和数据共享标准。

2.1　材料多类型数据库

目前常用的材料数据库如表 2 所示。ICSD 是由莱布尼茨信息基础设施研究所 Belsky 等[23]构建的无机材料结构数据库,收录了自 1913 年以来在 1 600 个学术期刊发表的超过 8 万篇论文中的共计 21 万条晶体结构数据,覆盖金属、有机物、同素异形体等各种形态的材料体系;COD (crystallography open database)是由英国剑桥大学 Quirós 等[24]开发的包含超过 700 万个有机、无机、金属有机化合物和矿物的晶体结构数据库;MP Database 是由美国加利福尼亚大学伯克利分校劳伦斯伯克利国家实验室 Jain 等[7]建立的材料计算模拟专用数据平台(见图 5),不仅收录了材料结构数据,也收录了元素性质、电子结构、弹性张量和能源转换电极性能等数据;AFLOWLIB (automatic flow lib)是由美国杜克大学 Curtarolo 等[19]基于 AFLOW-π 高通量软件开发的材料计算数据库,数据库中收录了 6 400 余条热力学相

表 2　材料多类型数据库的发展

Table 2　Development of multi-type databases of materials

名称	说明	网址
ALKEMIE-DB	通用材料数据库	https://alkemine.net/
ICSD	晶体结构数据库	http://www2.fiz-karlsruhe.de/icsd_home.html
COD	晶体结构数据库	http://crystallography.net
AFLOWLIB	材料计算数据库	http://aflowlib.org/
MP Database	材料计算数据库	http://materialsproject.org/
NOMAD	材料计算数据库	https://repository.nomad-coe.eu/
OQDM	材料计算数据库	http://oqmd.org/
CMR	材料计算数据库	https://cmr.fysik.dtu.dk/
OMDB	电子结构数据库	https://omdb.mathub.io/
JARVIS-DFT	材料赝势数据库	https://ctcms.nist.gov/-knc6/JVASP.html
C2DB	二维材料数据库	https://cmr.fysik.dtu.dk/c2db/c2db.html
AMSD	矿物材料数据库	http://rruff.geo.arizona.edu/AMS/amcsd.php
ASM	合金材料数据库	https://mio.asminternational.org/ac
CatApp	催化材料数据库	https://suncat.stanford.edu/theory/it-facilities
HSMD	储氢材料数据库	https://hydrogenmaterialssearch.govtools.us
Thermocalc	热力学数据库	https://thermocalc.com/products/databases/
Citrination	通用材料数据库	https://citrination.com
MatNavi	通用材料数据库	https://mits.nims.go.jp/en
NIST	通用材料数据库	https://www.nist.gov/programs-projects
Materials Web	通用材料数据库	https://materialsweb.org/twodmaterials
Materials Cloud	通用材料数据库	https://materialscloud.org/discover
MSDSN	通用材料数据库	http://www.materdata.cn/
Atomly	通用材料数据库	https://atomly.net/
MCDC	腐蚀防护数据库	https://www.corrdata.org.cn/
MGED	材料计算数据库	https://www.mgedata.cn/

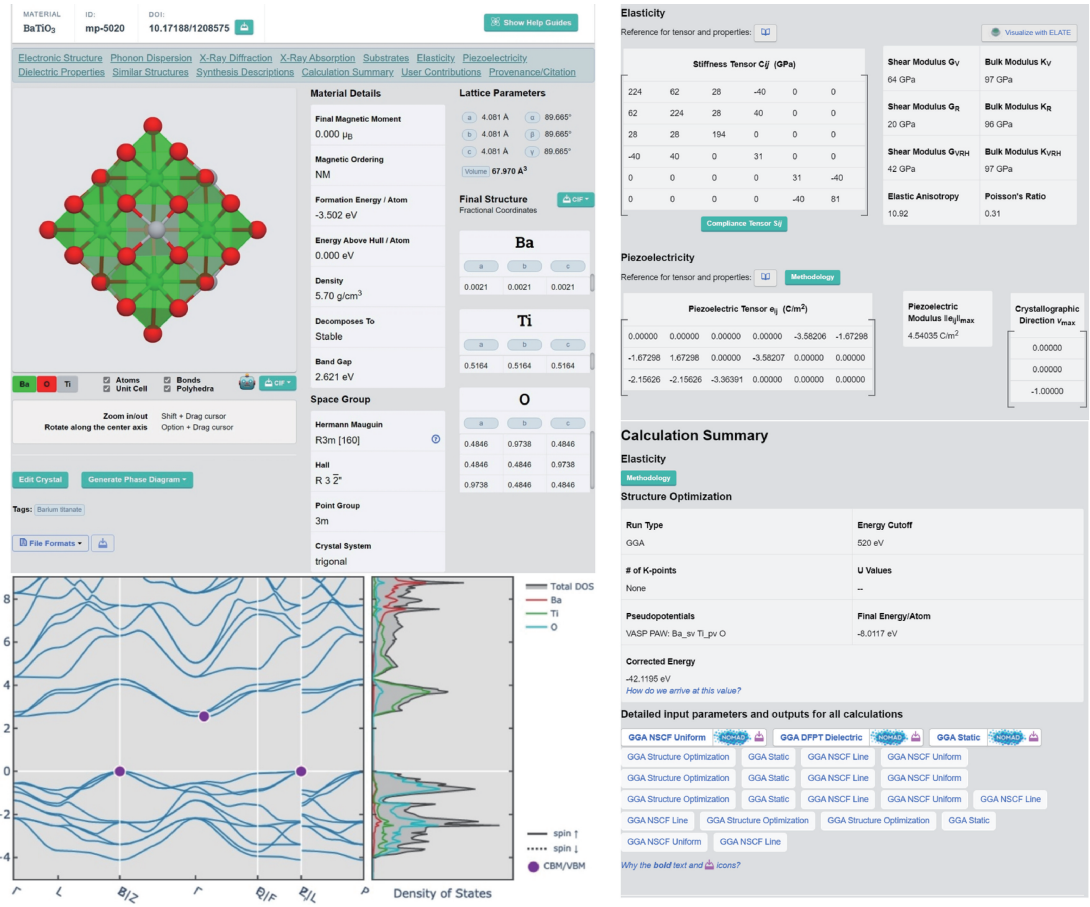


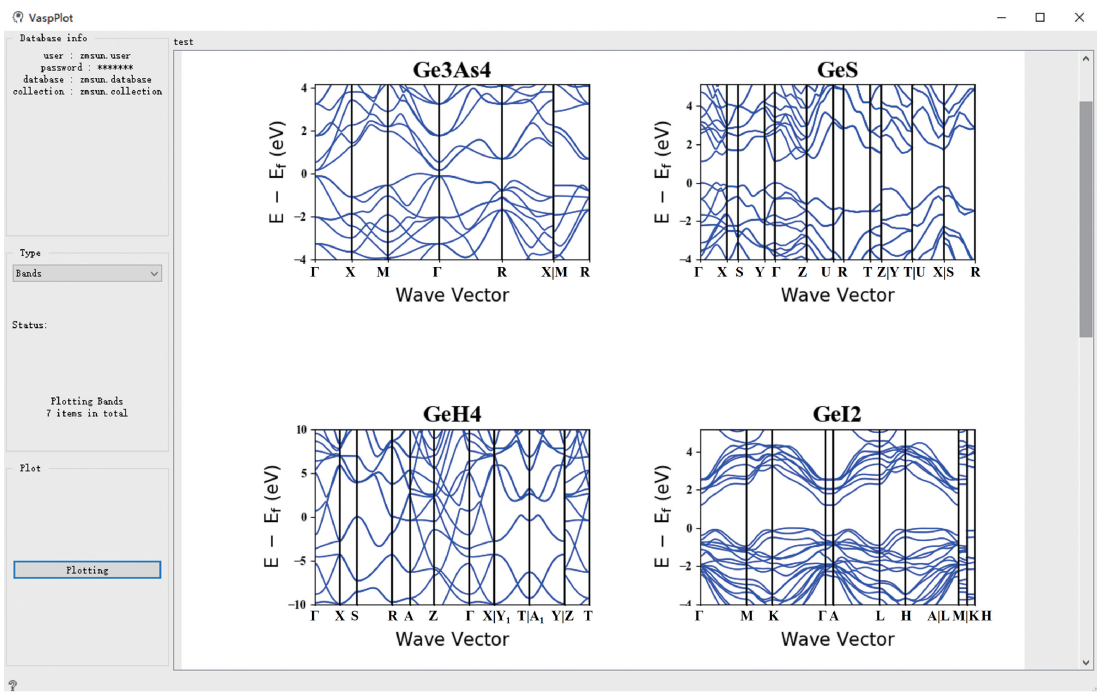
图 5 Materials Project Database 数据库概况[7]

Fig. 5 Snapshot of Materials Project Database[7]

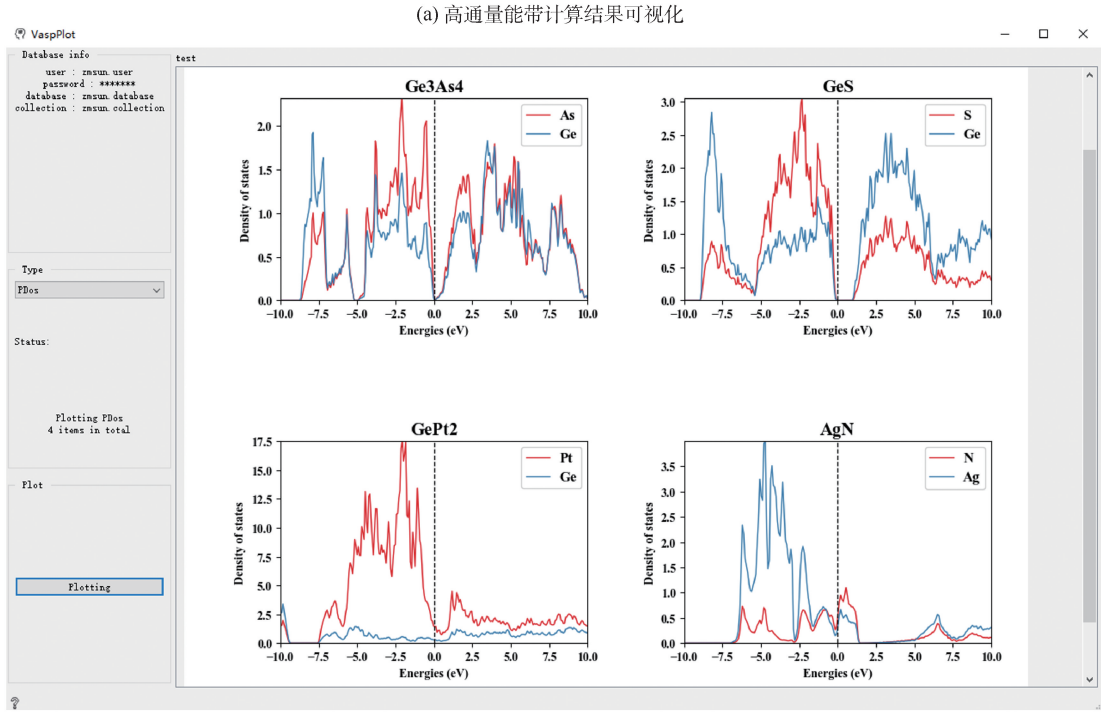
图数据和超过 45 万个四元混合物的材料性质数据;Materials Cloud 是由瑞士洛桑联邦理工大学 Pizzi 等[10]开发的第一性原理计算元数据的数据库,包括超过 752 万条第一性原理结构弛豫流程和纳米多孔材料吸收和扩散相关的材料性质数据;OQDM(open quantum database for materials)是由美国西北大学 Saal 等[25]开发的第一性原理计算热力学数据库,包含了数万个二元、三元和四元相图;NOMAD(novel materials discovery)是由欧洲马克斯·普朗克学会 Draxl 等[26]开发的欧洲最大的新材料共享数据库,该数据库包含了 49 TB 的各类材料数据;MatNavi 是由日本国家材料科学研究所 Ogata 等[27]开发的多种材料数据的集合,包含聚合物数据库(化学结构、加工、物理性质、NMR 光谱数据)、无机材料数据库(晶体结构、相图、物理性质)、金属材料数据库(密度、弹性模量、蠕变性质、疲劳特性)、电子结构计算数据库等。

除了上述通用的多类型材料数据库,还有众多针对材料某个特定领域的特色数据库。例如,美国佛罗里达大学开发的材料表面界面数据库 Materials Web;美国国家标准技术研究所 Choudhary 等[28]开发的赝势数据库 JARVIS-DFT;丹麦科技大学 Larsen 等[9]开发的二维材料、硒化物和硫化物数据库(computational materials repository, CMR);北欧理论物理研究所 Borysov 等[29]开发的三维有机晶体材料电子结构和带隙数据库(organic materials database, OMDB);美国斯坦福大学 Hummelshøj 等[30]开发的催化材料活化能数据库 CatApp;此外,还有 ASM 出版社开发的热处理数据表及应力应变和蠕变曲线材料数据库(ASM alloy center database, ASMDB)、矿物材料数据库(American mineralogist crystal structure database, AMSD)和储氢材料数据库(hydrogen storage materials database, HSMD)等。

中国目前也发展了多个大型材料数据库共享平台。ALKEMIE-DB 是由笔者团队 Wang 等[14-15]基于高通量智能计算平台 ALKEMIE 开发的多类型材料数据库(见图 6),该数据库分为隐私数据库和共享数据库两大类,根据数据类型进一步细分为含 60 余万组数据的晶体结构数据库、含 1 万余条声子能带的声子谱数据库、含 20 余万组数据的深度学习赝势数据库、高通量计算工作流数据库等。MSDSN 是由国家统筹建设、北京科技大学



(a) 高通量能带计算结果可视化



(b) 高通量态密度计算结果可视化

图 6　ALKEMIE-DB 数据库中高通量能带和态密度计算结果可视化[14]

Fig. 6　Visualization of high-throughput band structures and density of states calculations in ALKEMIE-DB databases[14]

实施完成的材料科学数据共享平台,主要分为包含有色金属材料和特种合金和微观组织模拟的实验数据库、热力学和动力学相关的计算模拟数据库等。Atomly 是由中国科学院物理研究所松山湖材料实验室和怀柔材料基因平台共建的材料计算数据库,目前包含 20 万条材料结构数据和 5 万条相图数据等。MCDC(national materials corrosion and protection data center)是由国家科技部门建设的腐蚀防护数据平台,包含环境数据、腐蚀数据、腐蚀检测和腐蚀预测等。

2.2　材料数据标准

随着高通量计算、高通量实验和超级计算机计算能力的发展,材料数据形成的数据海(data ocean)面临着 5 个重要的挑战("5V"特性):

1) 速率(velocity)。新数据产生速率和旧数据更新速率越来越快,对数据的格式化存储和快速读写提出了更高的要求。

2) 数据量(volume)。材料数据以 TB 量级不

断增加,需要可靠的数据存储、高效的数据检索和开放共享的可重复利用。

3) 多样性(variety)。材料数据的存储形式、材料体系、测试和计算方法,以及数据蕴含的物理意义更加多样化。

4) 真实性(veracity)。数据的不确定性和可靠性决定了数据是否真实有效,进一步决定了数据挖掘和机器学习模型的精度和泛化能力。

5) 低价值密度(value)。数据价值密度高低与数据总量大小成反比,数据量越大,数据价值密度越低。

如何在海量数据中分析预测数据隐藏的真实意义和价值是材料基因工程方法努力探索的主要方向。

为了解决上述问题,欧洲马克斯·普朗克学会 Draxl 等[26]提出了材料数据库建设的 FAIR Data 准则,即可发现(findable)、可获取(accessible)、可互操作(interoperable)和可再利用(reusable),来提升材料数据的开源共享性。欧洲用于材料设计的开放式数据库集成团队 Andersen 等[31]提出了材料数据共享标准 OPTIMADE(见图 7),该标准通过 JSON 格式定义了材料数据共享的统一标准,目前大多数材料数据库 MP Database、AFLOWLIB、NOMAD、Materials Cloud 和 ALKEMIE-DB 等均提供了 OPTIMADE 通用接口支持。

中国中关村材料试验技术联盟于 2019 年也提出了材料基因工程数据通则,将材料科学数据分为样品信息、源数据(未经处理的数据)与衍生数据(经分析处理得到的数据)三大类,并从宏观上定义了材料数据的通用性与专用性;北京科技大学材料基因工程北京市重点实验室在 2020 年 MSDSN 数据库中定义了材料科研数据 DOI 编码规则:“10. 11961/classification. project. date. sequence”,编码中包含了材料数据分类号、项目支撑信息、注册日期和流水号;北京航空航天大学 ALKEMIE-DB 数据库也发展了用于材料数据共享的唯一标识符:“alkemie. date. classification/user_defined_label. number.”,其中 alkemie 为数据库社区唯一标识,date 代表数据创建日期(精确到 μs),classification 表示数据类别,user_defined_label 为用户自定义字段,number 为数据唯一索引序号。

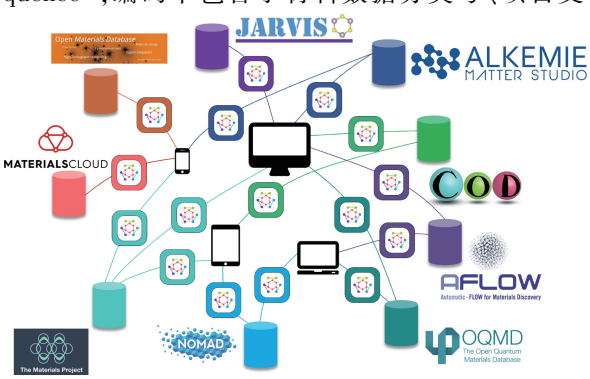


图 7 材料数据共享标准 OPTIMADE 概况[31]

Fig. 7 Overview of materials data sharing standard of OPTIMADE[31]

3 机器学习在材料设计中的应用

机器学习传统上分为监督学习和非监督学习两大类。监督学习是指给算法一个数据集,对于数据集中的每一个样本,都给出对应的映射(即标签),算法的目的是给出更多的映射,得到更多的答案;而非监督学习常常被用于在大量无标签数据中探索并发现规律。进一步,监督学习分为处理连续值的回归问题和处理离散值的分类问题。目前,常用的机器学习算法已经有很多相关综述,本文不再赘述。随着计算机科学和人工智能的发展,机器学习在材料结构设计、材料性能预测和材料分析图像识别等领域扮演着越来越重要的角色。本节主要介绍机器学习在材料性能预测、材料数据文本挖掘和机器学习原子间作用势。

3.1 材料性能预测

机器学习在材料学中的应用主要包含格式化材料数据、机器学习模型训练和高效材料性能预测 3 个步骤,并在二维、光伏、催化、合金和热电等材料中均获得了显著的研究成果。近年来,瑞士洛桑联邦理工大学 Lin 等[32]通过高通量筛选和机器学习算法,从 ICSD 和 COD 数据库的 108 423 个三维晶体构型中,通过对称性和几何算法筛选了 1 036 个容易合成和 789 个具有潜在可能性的层状二维材料混合物;东南大学 Lu 等[33]通过高通量筛选和机器学习描述符,从 5 158 个候选材料中筛选了能量转化效率高且带隙在 0.9 ~ 1.6 eV 之间有机无机结合的钙钛矿光伏材料(HOIPS),并构建了可以高效预测带隙值的结构描述符和机器学习模型,进一步将模型拓展到双层 HOIPS 中,成功从 11 370 个混合物中预测了 204 个无毒且稳定的光伏材料[34];北京科技大学 Zhang 等[35]在铜合金体系中通过贝叶斯优化迭代方法,分别构建了误差小于 7% 的硬度模型和误差小于 9% 的电导率模型,并通过迭代优化设计了兼具优异力学和电学性能的 Cu-Ni-Co-Si-Mg 合金。

3.2 材料数据文本挖掘

2018 年,瑞士联邦理工学院 Villars 等[36]通

过计算机视觉和自然语言方法从已经正式发表的论文中自动识别有效材料结构和数据,并通过数据挖掘探索数据背后隐藏的物理模型和机理,构建了 MPDS(materials platform for data science),近年来,该团队首次从 35 000 篇论文中解析了 15 500个化学成分,并基于其中的 2 330 个二元体系构建了机器学习分类模型;进一步从超过 8 000 篇发表的论文中构建了 2 800 个二元混合物的原子配位环境多面体分析算法(atomic environment types,AETs);随后,从超过 50 000 篇已发表论文中分析了 290 000 个原子配位环境数据,将该算法模型拓展到了多元无机混合物中。2019 年,Tshitoyan 等[37]发展了 Word2vec 非监督机器学习模型,成功从 330 万材料文本中筛选了超过 7 000 个候选的热电材料。除了材料计算模拟数据,在实验合成方式上,Kononova 等[38]通过文本挖掘和自然语言处理方式从 53 538 篇科学文献数据中筛选了 19 488 个无机金属合成方式,包括材料成分、制备条件、化学平衡方程和反应过程,该数据库为实验中无机材料的制备过程提供了有力的数据支持。

3.3　机器学习原子间作用势

材料计算模拟根据模拟时长和体系大小分为原子尺度、分子尺度、介观尺度和宏观尺度模拟,尺度越小模拟精度越高,尺度越大越接近真实体系,但是不同的模拟尺度采用的物理模型和近似原理不同,数据耦合非常困难,而数据驱动的机器学习方法被视为材料多尺度模拟的耦合剂。经典大规模分子动力学常被用来模拟近似真实材料体系的服役性能,但可靠、精确的原子间势函数的匮乏限制了其广泛应用。基于密度泛函原理(DFT)的第一性原理模拟具有精确的赝势库,但求解本征值所需的巨大计算量限制了该方法在大的原子尺度和时间尺度上的模拟,常用的 VASP 仅限数百原子的体系。因此,简单方便地获得适用于经典分子动力学的可靠势函数非常重要。随着计算机技术、计算机视觉和材料基因理念的快速发展,通过机器学习结合大数据、高通量计算的方法拟合可靠的适用于经典分子动力学模拟的势函数成为了研究热点。

机器学习势函数的发展主要经历了原子个数受限的低维度势函数和泛化能力强的高维度神经网络势函数 2 个发展过程。1995 年,Blank 等[39]开发了第一个基于统计学的势函数模型,用于研究氢原子的分子动力学模拟;2009 年,Malshe 等[40]进一步提出了通过神经网络预测经典多体势方程参数的模型。但是,上述模型均不能改变输入的原子个数,因此限制了机器学习势函数的应用。

2011 年,Behler[41]提出了原子中心对称函数,通过数学方程解析原子局域环境,构建了输入原子个数不受限的高维度神经网络模型。2018 年,Gastegger 等[42]发展了权重相关的对称函数(wACSF),通过卷积神经网络提升了模型的精度和实用性,但是由于局域近似,无法包含超过截断半径的原子长程相互作用。2018 年,Yao 等[43]提出了包含长程静电作用和散射作用的 HDNNP 方法,但是该方法并未获得广泛应用,一方面由于物理学中超过 6 ~ 10 Å 的静电作用通常对体系的整体影响较小,另一方面添加长程作用会急剧增加模型训练成本,与其对精度的微小提升相比有待进一步优化。笔者团队 Wang 等[14]开发了适用于相变材料 Sb 单质的跨尺度机器学习势函数PotentialMind,该势函数模型与 DFT 比较,对能量预测的精度达到 98%,平均到每个原子上的能量误差值小于 0.045 eV/atom,对力的预测精度达到 89%,该算法具有很强的扩展性和通用性,易于扩展到多元材料体系中。

机器学习势函数方法一方面实现了具有第一原理精度且更大原子数体系和更长时间尺度的大规模分子动力学模拟,另一方面通过替代求解复杂多体薛定谔方程本征值,使得模拟速度提升 2 ~ 3个数量级,目前,该方法已经在多个材料体系中获得了应用。例如,Sosso 等[44]发展了适用于二元相变存储材料 GeTe 的人工神经网络势函数,并实现了具有第一原理精度的 4 096 个原子体系的大规模分子动力学模拟,通过模拟相变材料的多个淬火过程(100 ~ 300 ps),探究了淬火速度和模拟体系大小对 GeTe 非晶结构的影响;Artrith和 Urban[45]基于 Fortran 语言开发了适用于钙钛矿 TiO_2 的神经网络势函数软件(atomic energy network, AENET),并面向科研人员开源使用,加速了机器学习势函数方法在能源材料中的应用;Mocanu 等[46]通过高斯近似方法构建了三元相变材料 $Ge_2Sb_2Te_5$ 的势函数模型,实现了含7 200个原子的非晶体系的大规模分子动力学模拟,揭示了模型大小和原子个数对非晶局域结构的影响,并阐明了非晶构型中化学键和晶化过程中的微观结构演化;Zhang 等[47]开发了适用于高性能并行计算的深度神经网络势函数方法 DeePMD,实现了模拟体系超过 1 亿原子、模拟时长超过 1 ns 的大规模分子动力学模拟,显著加速了新材料设计与研发。

4 多尺度集成可视化的高通量自动计算和数据管理智能平台 ALKEMIE

2016 年,笔者团队在国家重点研发计划材料基因工程专项的支持下,基于 Python 开源框架自主开发了一套多尺度集成可视化的高通量自动计算和数据管理智能平台 ALKEMIE,主要包含高通量自动工作流 ALKEMIE Matter Studio(MS)、数据管理及材料数据库 ALKEMIE Database(DB)、基于机器学习的材料数据挖掘 ALKEMIE Potential Mind(PM)三部分[14-15]。

4.1 高通量智能计算 ALKEMIE-MS

ALKEMIE 基于 AMDIV 设计理念,解决了材料高通量智能模拟过程中 5 个核心问题:自动化计算(automation)、模块化拓展(modular)、材料数据库(database)、人工智能方法(intelligence)和可视化界面(visualization),可实现从建模、运行到数据分析,全程自动无人工干预。

ALKEMIE 中,多尺度集成的高通量自动计算可以通过不同模块间以搭积木的方式实现自动耦合并完成计算。不同模块的连接方式如图 8 所示。首先,由建模模块控制输入,通过多种建模方式将材料构型导入高通量预处理器,科学计算模块控制任务的计算顺序和纠错(可进行电子尺度、原子尺度、分子尺度和介观尺度的高通量自动流程)。然后,服务器用来协调计算资源,配置远程节点,实现本地与远程服务器通信并提交任务,数据存储系统负责保存整个流程中所有的元数据,将计算结果保存在不同类型的数据库中。最后,通过数据分析和人工智能进行数据挖掘。

目前,该软件可以实现超过 10^4 量级的高通量并发计算,包含第一性原理计算 VASP、QE 和 ABINIT,分子动力学模拟 LAMMPS 和 ASE,热力学计算软件 Gibbs,动态蒙特卡罗模拟(KMC),相场相图模拟 OpenPhase 和 OpenCalphd 等多尺度模拟软件,可通过参数传递的方式实现跨尺度计算。目前,该软件系统已部署在 9 家超算中心,包括 4 家国家超算、4 家高校超算和 1 家企业超算。软件可移植性和可拓展性强,适用于对材料模拟掌握程度从初级到专业的所有材料研究人员,可提供材料建模、高通量智能计算、数据挖掘和人工智能一体化的新材料设计方案。

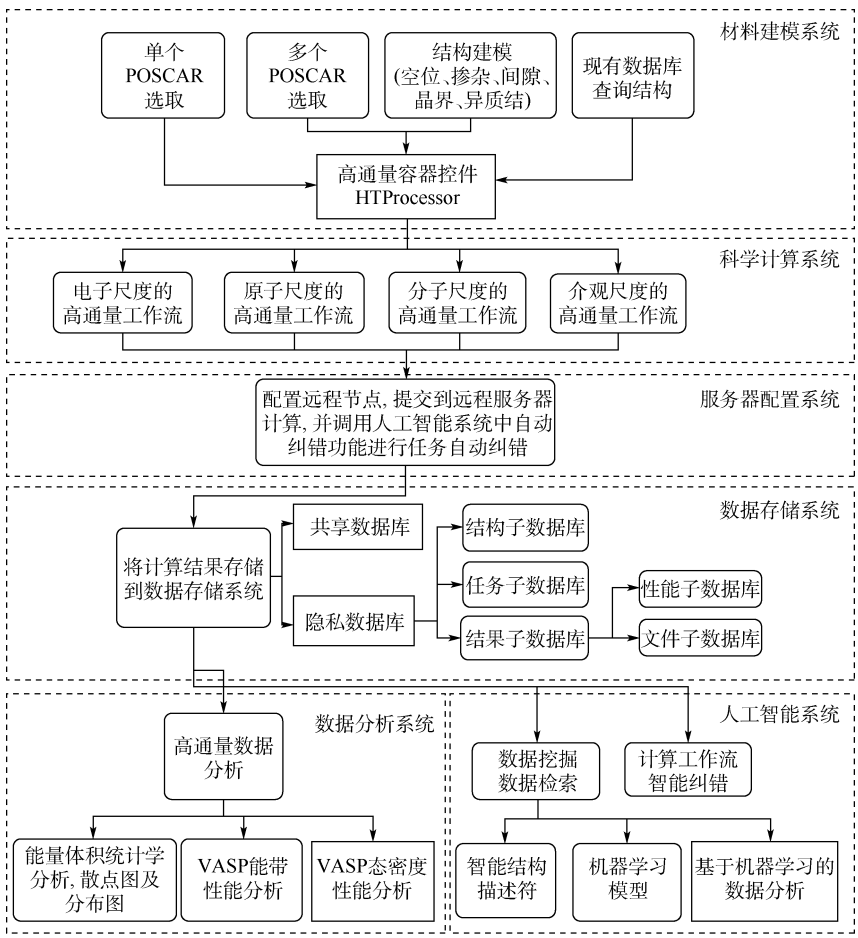


图 8 多尺度集成可视化的高通量自动计算和数据管理智能平台 ALKEMIE 计算模块概况

Fig. 8 Overview of platform with multi-scale integration of visualized automatic high-throughput calculation and intelligent data management ALKEMIE

4.2　材料多类型数据库 ALKEMIE-DB

ALKEMIE-DB 材料多类型数据库分为五大类:材料结构数据库、工作流源数据库、材料性能数据库、机器学习描述符数据库和论文数据库。目前,收录了超过 64 万条材料结构数据、296 条相图数据、1 万条声子能带数据、1 418 条赝势数据和 20 余万条机器学习描述符数据。通过 ALKEMIE 中 JSON 格式的 API 和国际通用的 OPTIMADE 格式的 API 实现数据高效查询检索,并通过 Finder 控件实现结构的数据可视化,如图 9 所示[14]。

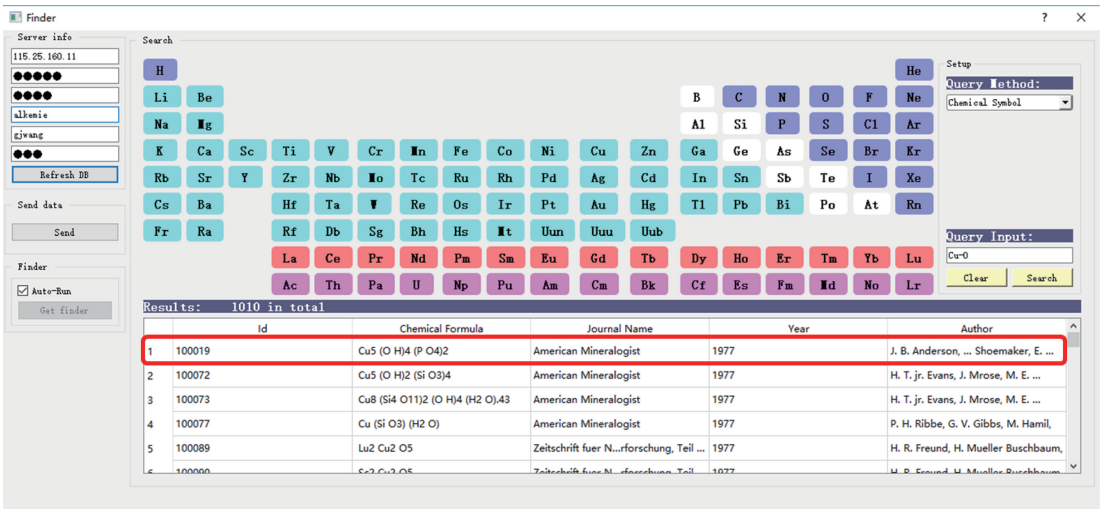


图 9　ALKEMIE-DB 材料结构数据库[14]

Fig. 9　Material structure database of ALKEMIE-DB[14]

4.3　材料学机器学习应用框架 ALKEMIE-PM

机器学习在材料学中通常被视为未知的黑盒模型,而材料研究人员相比于机器学习方法更关注材料性能、成分和工艺问题,因此,如何构建简单通用的可视化机器学习框架至关重要。ALKEMIE 通过抽象凝练高级 API 及规范化和格式化机器学习的每个具体步骤开发了一套通用的可视化机器学习的流程。Datasets 模块给定了数据集及特征的输入格式,Model 模块定义了不同的机器学习算法,Evaluate 模块给出了机器学习训练过程的收敛情况及模型在测试集或模型在部署过程中的应用情况,Plotter 模块中多种分析方法提供了将训练过程及其结果可视化的功能,如图 10 所示。

目前,笔者团队基于 ALKEMIE 可视化机器学习框架发展了一系列高效机器学习模型。Chen 等[48]通过团簇展开法和高通量第一性原理计算搜索了二维过渡金属硫化物的单层和双层无序掺杂结构,得到了稳定掺杂结构和能量之间的对应关系,通过机器学习挖掘了影响半导体-金属转变(SMT)的 2 个关键特征,即范德华间隙内掺杂氧原子的浓度差和 Mo-S/O 键角正切的平均值($\overline{\tan\theta}$);Peng 等[49]通过线性回归方法构建了适用于 MXene 材料单原子催化剂结构与氧还原反应(ORR)性能构效关系的组分描述符,基于简单的元素性质可以高效快速预测材料的催化活性;

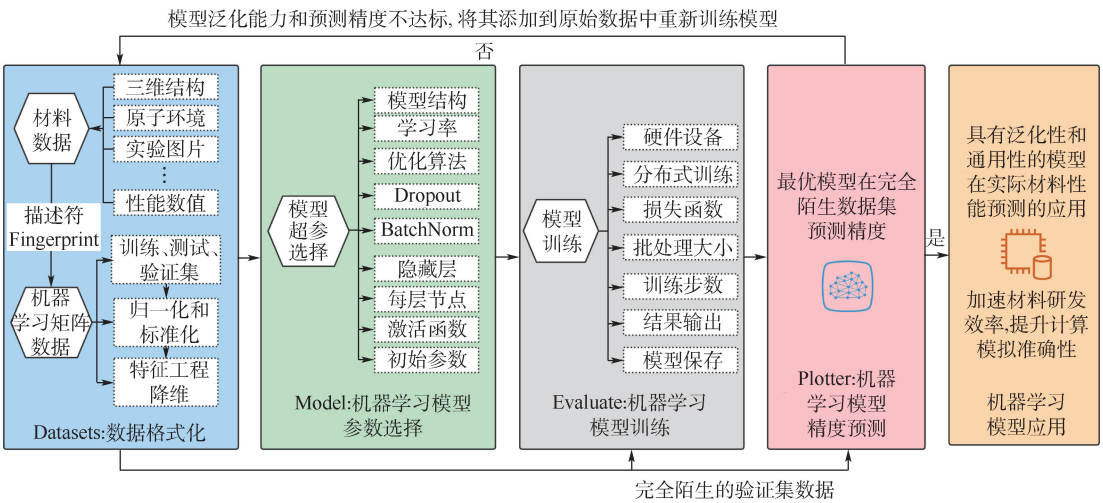


图 10　ALKEMIE-PM 机器学习框架

Fig. 10　Framework of ALKEMIE-PM machine learning

Gan 等[50]通过高通量筛选和可视化机器学习框架,开发了2个精度分别为90.90%和91.67%的机器学习模型预测层状IV-V-VI族半导体不同温度下的最大热电优值(ZT_{max})和实现ZT_{max}所对应的最佳掺杂类型,并成功从840个候选成分中筛选出数种具有优良潜力的热电材料。

4.4 未来的发展方向

综上所述,ALKEMIE已经研发了集可视化高通量自动计算流程、材料多类型数据库和人工智能方法于一体的新材料设计研发智能平台,但是未来仍有亟须发展的新方向和新方法。

在高通量计算方面,开发从原子、分子、介观到器件的跨尺度模拟方法是目前极具挑战且具有广阔应用前景的热点问题。Martin Karplus、Michael Levitt和Arieh Warshel三位科学家在分子领域凭借量子力学和分子动力学跨尺度模拟方法(QM/MM)获得了2013年诺贝尔化学奖。而在周期性材料的研究方面,由于体系周期性边界条件和原子局域环境的复杂性,使得跨尺度模拟的精度非常难以控制,发展高通量跨尺度高并发、自动纠错及数据耦合方法,通过机器学习数据挖掘等算法进一步提升跨尺度模拟精度是未来的研究热点之一。

材料数据库方面,应该保持开源和共享的发展理念,基于FAIR原则,构建包含材料计算和实验元数据及中间数据的高效数据库,发展数据规模更大、种类更丰富且具有航空特色的共享数据平台,完善更加通用兼容的数据标准和共享标识均是未来重要的研究方向。

在机器学习领域,材料学中数据集的构建非常困难,因此,研发基于小数据集的高效机器学习模型训练算法至关重要;由于机器学习模型的黑盒特性,探索可解释的机器学习模型,阐明模型背后隐藏的物理意义,实现逆向材料成分和结构设计也是未来的热门研究领域。

5 结束语

材料基因工程颠覆了传统的"试错-纠错"材料研发模式,通过数据驱动的高通量方法和人工智能模型加速新材料的研发与设计。本文系统总结了国内外知名的材料高通量计算框架、常用的针对多种材料体系的大型材料数据库和机器学习方法,并概述了多尺度集成可视化的高通量自动计算和数据管理智能平台ALKEMIE的研究进展,提出了未来发展的研究方向,为实现按需逆向设计新材料提供参考。

参考文献(References)

[1] XIANG X, SUN X, BRICENO G, et al. A combinatorial approach to materials discovery[J]. Science, 1995, 268(5218): 1738-1740.

[2] ALLISON J. Integrated computational materials engineering: A perspective on progress and future steps[J]. JOM, 2011, 63(4): 15-18.

[3] LIU Z. High throughput CALPHAD modeling and the materials genome[J]. Technology, 2010, 617: 622.

[4] LIU Z. View and comments on the dataecosystem: Ocean of data[J]. Engineering, 2020, 6(6): 604-608.

[5] WHITE A. The materials genome initiative: One year on[J]. MRS Bulletin, 2012, 37(8): 715-716.

[6] AGRAWAL A, CHOUDHARY A. Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science[J]. APL Materials, 2016, 4(5): 053208.

[7] JAIN A, ONG S, HAUTIER G, et al. Commentary: The materials project: A materials genome approach to accelerating materials innovation[J]. APL Materials, 2013, 1(1): 011002.

[8] CURTAROLO S, SETYAWAN W, HART G, et al. AFLOW: An automatic framework for high-throughput materials discovery[J]. Computational Materials Science, 2012, 58: 218-226.

[9] LARSEN A, MORTENSEN J, BLOMQVIST J, et al. The atomic simulation environment-A Python library for working with atoms[J]. Journal of Physics: Condensed Matter, 2017, 29(27): 273002.

[10] PIZZI G, CEPELLOTTI A, SABATINI R, et al. AiiDA: Automated interactive infrastructure and database for computational science[J]. Computational Materials Science, 2016, 111: 218-230.

[11] GONZE X, AMADON B, ANGLADE P M, et al. ABINIT: First-principles approach to material and nanosystem properties[J]. Computer Physics Communications, 2009, 180(12): 2582-2615.

[12] MATHEW K, SINGH A K, GABRIEL J J, et al. MPInterfaces: A materials project based Python tool for high-throughput computational screening of interfacial systems[J]. Computational Materials Science, 2016, 122: 183-190.

[13] LAMBERT H, FEKETE A, KERMODE J R, et al. Imeall: A computational framework for the calculation of the atomistic properties of grain boundaries[J]. Computer Physics Communications, 2018, 232: 256-263.

[14] WANG G J, PENG L, LI K, et al. ALKEMIE: An intelligent computational platform for accelerating materials discovery and design[J]. Computational Materials Science, 2021, 186: 110064.

[15] WANG G J, LI K, PENG L, et al. High-throughput automatic integrated material calculations and data management intelligent platform and the application in novel alloys[J]. Acta Metallurgica Sinica, 2021, 58(1): 75-88.

[16] YANG J, LI H, WU T, et al. Evaluation of half-Heusler compounds as thermoelectric materials based on the calculated electrical transport properties[J]. Advanced Functional Materials,

2008,18(19):2880-2888.

[17] YANG X,WANG Z,ZHAO X,et al. MatCloud:A high-throughput computational infrastructure for integrated management of materials simulation,data and resources[J]. Computational Materials Science,2018,146:319-333.

[18] ZHAO X G,ZHOU K,XING B,et al. JAMIP:An artificial-intelligence aided data-driven infrastructure for computational materials informatics[J]. Science Bulletin,2021,66(19):1973-1985.

[19] CURTAROLO S,SETYAWAN W,WANG S,et al. AFLOWLIB.ORG:A distributed materials properties repository from high-throughput ab initio calculations[J]. Computational Materials Science,2012,58:227-235.

[20] CURTAROLO S,HART G L,NARDELLI M B,et al. The high-throughput highway to computational materials design[J]. Nature Materials,2013,12(3):191-201.

[21] HU S,LIU B,LI Z,et al. Identifying optimal dopants for Sb_2Te_3 phase-change material by high-throughput ab initio calculations with experiments[J]. Computational Materials Science,2019,165:51-58.

[22] GAN Y,MIAO N,LAN P,et al. Robust design of high-performance optoelectronic chalcogenide crystals from high-throughput computation[J]. Journal of the American Chemical Society,2022,144(13):5878-5886.

[23] BELSKY A,HELLENBRANDT M,KAREN V L,et al. New developments in the inorganic crystal structure database (ICSD):Accessibility in support of materials research and design[J]. Acta Crystallographica,2002,58(3):364-369.

[24] QUIRÓS M,GRAŽULIS S,GIRDZIJAUSKAITÉ S,et al. Using SMILES strings for the description of chemical connectivity in the crystallography open database[J]. Journal of Cheminformatics,2018,10(1):1-17.

[25] SAAL J E,KIRKLIN S,AYKOL M,et al. Materials design and discovery with high-throughput density functional theory:The open quantum materials database (OQMD)[J]. JOM,2013,65(11):1501-1509.

[26] DRAXL C,SCHEFFLER M. The NOMAD laboratory:From data sharing to artificial intelligence[J]. Journal of Physics:Materials,2019,2(3):036001.

[27] OGATA T,YAMAZAKI M. New stage of MatNavi,materials database at NIMS[C]// ASME Pressure Vessels and Piping Conference. New York:ASME,2009:1561-1568.

[28] CHOUDHARY K,DECOST B,TAVAZZA F. Machine learning with force-field-inspired descriptors for materials:Fast screening and mapping energy landscape[J]. Physical Review Materials,2018,2(8):083801.

[29] BORYSOV S S,GEILHUFE R M,BALATSKY A V. Organic materials database:An open-access online database for data mining[J]. PLOS One,2017,12(2):171501.

[30] HUMMELSHØJ J S,ABILD-PEDERSEN F,STUDT F,et al. CatApp:A Web application for surface chemistry and heterogeneous catalysis[J]. Angewandte Chemie International Edition,2012,51(1):272-274.

[31] ANDERSEN C W,ARMIENTO R,BLOKHIN E,et al. OPTIMADE,an API for exchanging materials data[J]. Scientific Data,2021,8(1):217.

[32] LIN X,YANG W,WANG K L,et al. Two-dimensional spintronics for low-power electronics[J]. Nature Electronics,2019,2(7):274-283.

[33] LU S,ZHOU Q,OUYANG Y,et al. Accelerated discovery of stable lead-free hybrid organic-inorganic perovskites via machine learning[J]. Nature Communications,2018,9(1):3405.

[34] WU Y,LU S,JU M G,et al. Accelerated design of promising mixed lead-free double halide organic-inorganic perovskites for photovoltaics using machine learning[J]. Nanoscale,2021,13(28):12250-12259.

[35] ZHANG H,FU H,ZHU S,et al. Machine learning assisted composition effective design for precipitation strengthened copper alloys[J]. Acta Materialia,2021,215:117118.

[36] VILLARS P,CENZUAL K,GLADYSHEVSKII R,et al. PAULING FILE-towards a holistic view[J]. Chemistry of Metals and Alloys,2018,11(3):43-76.

[37] TSHITOYAN V,DAGDELEN J,WESTON L,et al. Unsupervised word embeddings capture latent knowledge from materials science literature[J]. Nature,2019,571(7763):95-98.

[38] KONONOVA O,HUO H,HE T,et al. Text-mined dataset of inorganic materials synthesis recipes[J]. Scientific Data,2019,6(1):203.

[39] BLANK T B,BROWN S D,CALHOUN A W,et al. Neural network models of potential energy surfaces[J]. The Journal of Chemical Physics,1995,103(10):4129-4137.

[40] MALSHE M,NARULKAR R,RAFF L M,et al. Development of generalized potential-energy surfaces using many-body expansions,neural networks,and moiety energy approximations[J]. The Journal of Chemical Physics,2009,130(18):184102.

[41] BEHLER J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials[J]. The Journal of Chemical Physics,2011,134(7):074106.

[42] GASTEGGER M,SCHWIEDRZIK L,BITTERMANN M,et al. wACSF:Weighted atom-centered symmetry functions as descriptors in machine learning potentials[J]. The Journal of Chemical Physics,2018,148(24):241709.

[43] YAO K,HERR J E,TOTH D W,et al. The TensorMol-0.1 model chemistry:A neural network augmented with long-range physics[J]. Chemical Science,2018,9(8):2261-2269.

[44] SOSSO G C,MICELI G,CARAVATI S,et al. Neural network interatomic potential for the phase change material GeTe[J]. Physical Review B,2012,85(17):174103.

[45] ARTRITH N,URBAN A. An implementation of artificial neural-network potentials for atomistic materials simulations:Performance for TiO_2[J]. Computational Materials Science,2016,114:135-150.

[46] MOCANU F,KONSTANTINOU K,LEE T,et al. Modeling the phase-change memory material,$Ge_2Sb_2Te_5$,with a machine-learned interatomic potential[J]. The Journal of Physical Chemistry B,2018,122(38):8998-9006.

[47] ZHANG L,HAN J,WANG H,et al. Deep potential molecular dynamics:A scalable model with the accuracy of quantum mechanics[J]. Physical Review Letters,2018,120(14):143001.

[48] CHEN Q,CHEN M,ZHU L,et al. Composition gradient media-

ted semiconductor metal transition in ternary transition metal dichalcogenide bilayers[J]. ACS Applied Materials & Interfaces,2020,12(40):45184-45191.

[49] PENG Q,ZHOU J,CHEN J,et al. Cu single atoms on Ti_2CO_2 as a highly efficient oxygen reduction catalyst in a proton exchange membrane fuel cell[J]. Journal of Materials Chemistry A,2019,7(45):26062-26070.

[50] GAN Y,WANG G,ZHOU J,et al. Prediction of thermoelectric performance for layered IV-V-VI semiconductors by high-throughput ab initio calculations and machine learning[J]. NPJ Computational Materials,2021,7(1):176.

Novel material design and development accelerated by materials genome engineering

SUN Zhimei[1,*], WANG Guanjie[1,2], ZHANG Xuanguang[1], ZHOU Jian[1]

(1. School of Materials Science and Engineering, Beihang University, Beijing 100083, China;
2. School of Integrated Circuit Science and Engineering, Beihang University, Beijing 100083, China)

Abstract: Without knowing the relationship between the chemical composition and properties of the material, the development of new materials with desired properties based on conventional trial-and-error methods is cost inefficient and sometimes ends fruitlessly. With the development of artificial intelligence and the fourth science paradigm, data-driven science, materials genetic engineering (MGE) has become a new approach for designing novel materials. In this paper, recent advances in high-throughput computation, materials database, and artificial intelligence methods in MGE are reviewed. First, the framework and software for high-throughput computations of materials are introduced. Then, the progress and critical problems of materials databases are presented in terms of types and standard interfaces of materials data. Next, we summarize the applications of artificial intelligence methods in critical issues of materials science. In particular, from the aspects of visualized high-throughput calculation methods, multi-type materials database and visualized machine learning framework, we emphasize an in-house developed platform ALKMIE, featured by multi-scale integration of automatic high-throughput calculation, visualization, and intelligent data management. Finally, we highlight the future directions of MGE.

Keywords: materials genome engineering (MGE); high-throughput calculation; materials database; machine learning; artificial intelligence calculation platform

Received: 2022-05-06; **Accepted**: 2022-05-18; **Published online**: 2022-06-01 15:07

URL: kns.cnki.net/kcms/detail/11.2625.V.20220531.1837.003.html

Foundation items: National Natural Science Foundation of China (51872017); China Postdoctoral Science Foundation (2022TQ0019); The High Performance Computing (HPC) Resources at Beihang University

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0309

星基 ADS-B 系统及关键技术发展综述

张学军*, 谭元皓, 李雪缘, 简鑫慧
(北京航空航天大学 电子信息工程学院, 北京 100083)

摘 要: 广播式自动相关监视(ADS-B)系统是国际民航组织在空管全球化的背景下提出的新一代监视技术。相较陆基系统,星基 ADS-B 系统能够实现全球空域覆盖,可增强现有空管系统能力,促进国家低空空域开放与通航产业发展。概述了星基 ADS-B 的由来、概念及运行原理;介绍了星基 ADS-B 系统国内外发展概况与发展历程;针对微弱信号解调、多波束接收、解交织、防欺骗、星座设计、路由转发算法及监视性能评估等星基 ADS-B 关键技术梳理了国内外相关研究现状,介绍了北京航空航天大学联合团队在星基 ADS-B 上的研究工作与发射试验星情况;结合未来星基空管技术发展与应用需求对星基 ADS-B 系统发展趋势进行总结与展望。

关 键 词: 星基广播式自动相关监视(ADS-B); 微弱信号解调; 多波束接收; 解交织; 防欺骗; 星座设计; 路由算法; 监视性能评估
中图分类号: V19
文献标志码: A **文章编号**: 1001-5965(2022)09-1589-16

21 世纪以来,随着全球经济和航空业的飞速发展,航空器规模不断扩张,航线航班迅速增加,全球飞行日益密集,跨洋区、荒漠、偏远地区飞行任务迅速增长,民用运输航空成为国家交流和跨境货运不可或缺的重要方式,军事航空全球战略飞行和远程奔袭打击成为国家军事战略威慑重要样式,全球各国对航空实时监视服务和监管基础设施的需求不断增长。

广播式自动相关监视(automatic dependent surveillance broadcast, ADS-B)是国际民航组织(International Civil Aviation Organization, ICAO)确定的空中交通管理航空监视新技术,航空器从全球卫星定位系统和机载设备获取标识号、三维位置、速度、航向等信息,通过空-空数据链及空-地数据链周期性自动向其他航空器或地面站进行广播式发送,并接收其他航空器的广播信息。美国、欧洲、澳大利亚、中国等全球各国均在大力推广 ADS-B 技术[1-2]。

陆基 ADS-B 系统作为目前推行的航空监视系统,存在一定程度的应用局限性,尚不能满足未来航空监视全球化、全程无缝监视覆盖的要求。多起航空事故都暴露了全球民航基础设施建设不足,尤其是海洋及南北两极区域监视缺乏的问题;另外,低空空域各种非法飞行活动对民航飞行造成了巨大威胁。传统陆基 ADS-B 系统受地面部署条件的限制,要实现海洋全覆盖需要在大量浮标上部署 ADS-B 地面站,要实现地面全覆盖需要在大量无人区部署 ADS-B 地面站,由于建设难度目前只能覆盖全球 30% 左右的空域[3],海洋、极地等偏远地区与地形复杂区域存在大量盲区,这些盲区使得各类航空系统的运行效率无法发挥,空域使用效率受到制约,飞行安全也存在隐患。同时,要完全依靠陆基 ADS-B 实现空域 100% 全覆盖,建站规模和后期运营维护成本过高。尤其中国地域辽阔,东西部地理环境差异大,发展程度不均衡,难以实现全国范围内大面积航空监视和

收稿日期: 2022-05-01; **录用日期**: 2022-06-09; **网络出版时间**: 2022-07-04 16:03
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220701.1950.002.html
* **通信作者**. E-mail: zhxj@buaa.edu.cn

引用格式: 张学军, 谭元皓, 李雪缘, 等. 星基 ADS-B 系统及关键技术发展综述[J]. 北京航空航天大学学报, 2022, 48(9): 1589-1604. ZHANG X J, TAN Y H, LI X Y, et al. A review of development of space-based ADS-B system and its key technologies[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1589-1604(in Chinese).

空管服务。

针对陆基 ADS-B 系统的不足,国际民航组织提出了星基监视的概念,通过在低轨卫星上搭载 ADS-B 接收设备(即星基 ADS-B 系统),利用卫星星座全球覆盖、无地形遮挡等特点,实现对全球航班的实时连续无缝监视[4]。

星基 ADS-B 是实现远洋远域航空器监视的主要技术手段,相较传统的陆基监视系统,其具有全球覆盖的优势[5],可有效克服陆基星基 ADS-B 系统不足,在陆基 ADS-B 地面站难以覆盖区域引入星基 ADS-B 服务,提供一个可靠有效的监视方案,以提升航空的安全性、效率和容量,并可开拓全新的军事应用。

1) 提升航空安全性。星基 ADS-B 系统可以提供接近实时的持续飞行监视服务,增强态势感知、冲突检测与解脱能力;同时,星基 ADS-B 系统可以提供独立于全球航空通信网络的监视模式,更好地管理安全风险,有助于重建事故的相关因素与情况,提高对全球航空器搜索和救援的安全响应能力,增强航空安全保障。

2) 提升飞行效率和容量。星基 ADS-B 系统可以有效管理航空器在空域内飞行的航迹,促进和支持更准确、更高效的航迹规划,改善航空器反方向和交叉飞行的交通状况,有效提升飞行效率并降低航空公司运行成本;同时,星基 ADS-B 系统支持航空监视信息共享和运行协同,减小飞行间隔,提升空域容量。

3) 开拓军事新应用。星基 ADS-B 系统可以为军事航空全球飞行提供一种监视模式,可获取全球军事航空飞行态势数据和活动规律,实现对周边及域外大国军事航空活动早期预警,为空防安全提供保障。

中国民用航空局制定了《中国民航航空器追踪监控体系建设实施路线图》,其核心内容就是提升航空器追踪监控水平,实现全球范围内对航空器的持续监视与安全管理[6],其中包括自主星基 ADS-B 系统的建设。因此,星基 ADS-B 系统对于未来航空业发展具有重要作用和前瞻性意义。

本文概述了星基 ADS-B 概念及其运行原理,介绍了星基 ADS-B 系统国内外发展历程与相关卫星、载荷概况,总结了包括微弱信号解调、多波束接收、解交织、防欺骗、星座设计、路由转发算法、监视性能评估在内的星基 ADS-B 关键技术研究现状;进一步介绍了北京航空航天大学(以下简称北航)联合团队在星基 ADS-B 上的研究工作与试验星情况。结合星基技术发展与空管应用需求展望了星基 ADS-B 系统未来趋势。

1 星基 ADS-B 系统与应用

星基 ADS-B 系统是一种基于全球卫星定位系统、航空数据链及卫星通信的新一代广域航空器监视系统,其由低轨星座及卫星 ADS-B 接收机、卫星空间组网子系统、机载 ADS-B 设备和地面空管综合应用子系统等共同组成,低轨星座中每一颗卫星都加装了高灵敏度的 1090ES ADS-B 接收机,具备 ADS-B In 功能,接收民用和军用航空器的机载 ADS-B Out 设备发送的报文。其总体架构如图 1 所示。

星基 ADS-B 系统运行原理如图 2 所示。航空器从全球卫星定位系统和机载设备获取标识号、三维位置、速度、航向等信息,通过机载 ADS-B 发射机以 1 090 MHz 频率周期性向外进行广播式发送。卫星上的 ADS-B 接收机在全球范围内接收航空器广播的 ADS-B 信息,解调为报文并按照卫星星间通信协议组包,经星间链路转发至国内卫星地面网关站上方的卫星后,通过星地链路下传落地,并统一传回中国星基 ADS-B 地面数据中心。数据中心处理后,按照不同用户的应用需求,分发给管制部门、航空公司和通航企业等航空用户,实现对全球各国加装 1090ES ADS-B Out 设备的军用和民用航空器连续实时监视。

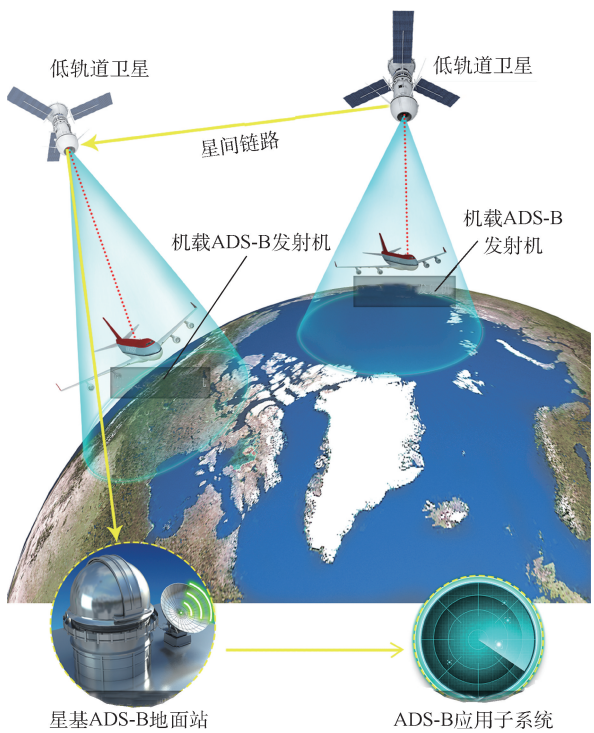


图 1 星基 ADS-B 系统总体架构

Fig. 1 Overall architecture of space-based ADS-B system

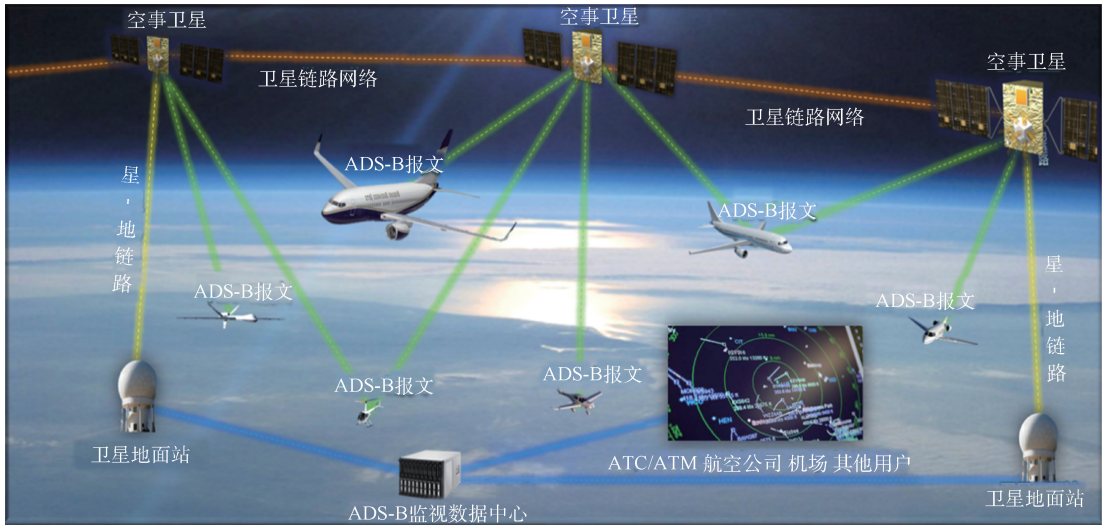


图 2　星基 ADS-B 系统运行原理

Fig. 2　Space-based ADS-B system operation principle

星基 ADS-B 系统兼容原先陆基 ADS-B 系统,只是将 ADS-B 接收机安装于低轨卫星星座上,从而满足洋区、荒漠、偏远地区及地球南北两极等区域的监视需求,有效弥补陆基 ADS-B 系统布站的局限性,实现空域无缝监视,在全时段、全球范围对军用和民用航空器飞行全过程进行连续无间断位置跟踪与目标识别,并能大幅提升空管监视系统的精确性和可靠性,可广泛应用于军民航领域。

在民航方面,星基 ADS-B 可应用于空中交通管制、航空公司运行控制、全球航班跟踪、通用航空监视等。例如,可利用星基 ADS-B 获取的包括空域信息、机场信息等在内的海量数据进行分析,辅助预测航迹,在全球层面上进行战术和战略交通管理规划,以减少任何特定空域的延误和中断,帮助航空公司节省燃料成本、高效运行;可以为全球航空器提供安全运行状态监测及救援信息数据支持,缩小航空安全间隔,优化空域管理,增加空域容量,提升航空安全。

在军航方面,星基 ADS-B 可应用于军用航空器远海远域飞行任务保障、外军情报收集分析、南海和中印边界等敏感地区巡航监视、战事指挥等。一方面适应现代化强国战略需求,为军事航空承担的反恐维稳、抢险救灾、空中安保、国际救援等任务提供一种监视手段和保障,实现全球战略巡航飞行及域外作战部署;另一方面建立有效的覆盖沿边沿海空域的监视和识别手段,实现领海领空入侵检测,同时可获取并掌握全球军事航空飞行态势数据和活动规律,如重点军用机场起降情况、军事部署力量调度情况、周边及域外大国军航行动等,实现全球范围内军情收集,保障国家空域安全。

2　星基 ADS-B 系统发展趋势

“MH370”事件发生后,通过星基 ADS-B 系统实现全球航空器持续监视引起全球高度重视。目前,世界主要航空大国均针对星基 ADS-B 开展对应研究与关键技术攻关,积极推进星基 ADS-B 系统建设发展,瞄准全球航空监视体系构建,依托空间互联网、低轨卫星星座系统开展有关的关键技术测试验证,并且主要技术架构、设备功能和应用形态初见成型,尤其 Aireon 公司基于铱星二代组网卫星,构建的星基 ADS-B 系统已开始试点提供全球服务。国内主要是国防科技大学、上海微小卫星工程中心、航天五院、航天八院及北京航空航天大学等单位,先后启动 ADS-B 星基接收机等关键技术攻关并开展了相应研究。

2.1　国外发展

国外在政策上对星基 ADS-B 建设大力支持,2015 年世界无线电通信大会将 1 090 MHz 分配给航空移动卫星服务,用于接收按照国际民航组织标准运行的航空器发射的 ADS-B 信息[7]。国际民航组织也在持续关注并积极推动星基 ADS-B 的测试与发展,根据星基 ADS-B 技术的建设进展制定相关标准[8]。

美国、加拿大、欧洲航天局(ESA)、德国、丹麦等均启动了各自的星基航空监视计划,开展了在轨试验,其中铱星二代系统已经能够满足全球非繁忙空域的空管监视应用需求。

1) Proba-V

德国宇航中心(DLR)于 2008 年开始研究低轨卫星是否能接收到飞机广播的 1090ES ADS-B

信号,为星基 ADS-B 提供了概念证明[9]。2013 年 5 月欧洲航天局发射了 Proba-V 卫星,携带 DLR 研制的 ADS-B 接收机,在开机运行的 2 h 内接收到了超过 12 000 条 ADS-B 信息[10-11],其观测到的飞机航迹如图 3 所示。Proba-V 首先验证了利用低轨卫星接收 ADS-B 信号的可行性,为该领域的未来发展奠定了基础。

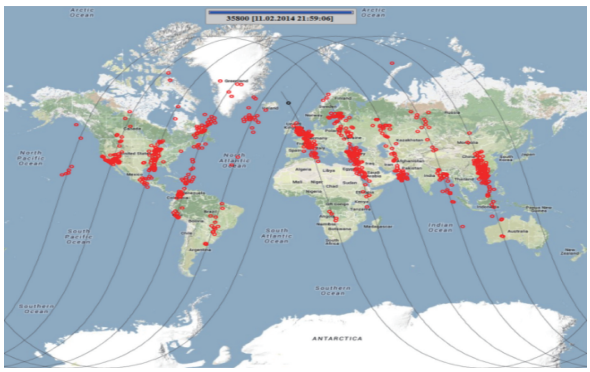


图 3 Proba-V 观测到的飞机航迹[10]

Fig. 3 Aircraft tracks observed by Proba-V[10]

2) GOMX 系列

GomSpace 在 2013 年底发射了搭载 ADS-B 有效载荷的 GOMX-1 立方星,根据在轨运行期间接收的数据分析,载荷具备跟踪飞机的能力,验证了基于立方体卫星的 ADS-B 载荷可行性[12-13]。2015 年,GomSpace 发射了 GOMX-3 卫星,更新了 ADS-B 接收机,图 4 为 GOMX-3 采集的全球 ADS-B 数据,在全球除极地地区外的区域收集 ADS-B 数据任务中表现良好[14]。2019 年又发射了 GOMX-4 卫星,使用与 GOMX-3 相同的 ADS-B 接收器,并测试不同距离的卫星间通信,最高可达 4 500 km[15-16]。

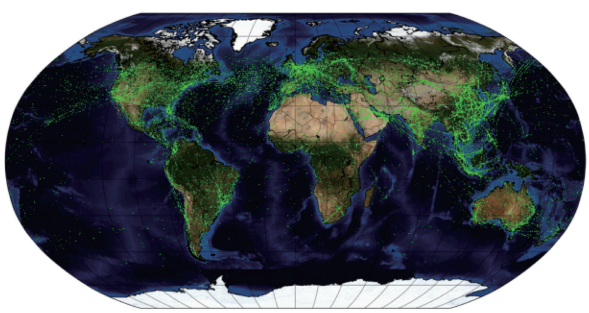


图 4 GOMX-3 采集的全球 ADS-B 数据[14]

Fig. 4 Global ADS-B data collected by GOMX-3[14]

3) CanX-7

CanX-7 卫星于 2016 年 9 月发射,其上搭载由加拿大皇家军事学院(RMCC)研制的 ADS-B 载荷。其主要关注地点是北大西洋,自发射以来的 6 个月内共接收到超过 430 万条 ADS-B 信息,在轨采集数据分布如图 5 所示,测试结果表明搭载在低轨卫星上的接收机能够探测和跟踪飞机[17-18]。

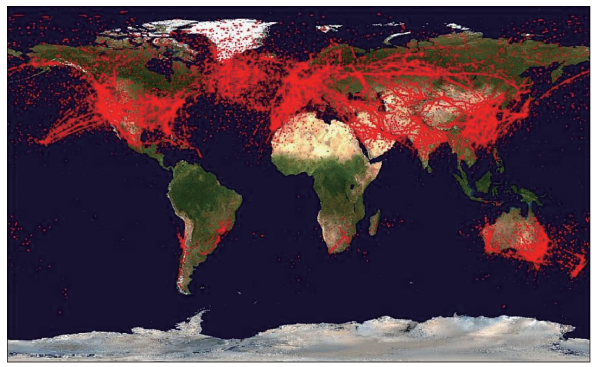


图 5 CanX-7 在轨采集数据分布[18]

Fig. 5 Distribution of CanX-7 in-orbit acquisition data[18]

4) ALAS 系统

2008 年美国 GlobalStar 和 ADS-B Technologies 公司开始联合研发可用的星基 ADS-B 系统,借助 GlobalStar 卫星星座提供全球 ADS-B 服务,称为 ADS-B 链路增强系统(ADS-B link augmentation system, ALAS),并在美国进行了飞行测试,如图 6 所示[19]。目前由于全球星座建设情况基本停滞,ALAS 系统前景很不明朗;并且由于 ALAS 系统需要对飞机设备进行改装,其在民用航空领域的推广应用难度也十分大[20-21]。



图 6 GlobalStar ADS-B(ALAS)飞行测试[19]

Fig. 6 GlobalStar ADS-B (ALAS) flight test[19]

5) Spire

Spire Global 于 2016 年 12 月宣布推出一项名为 Spire AirSafe 的星基全球航空器追踪服务。2018 年年中 Lemur 2 卫星搭载 ADS-B 有效载荷发射,目前 Spire 公司已经发射了超过 100 颗卫星,并在其官网提供了 ADS-B 数据服务[22-24],Spire 的 ADS-B 数据如图 7 所示。

6) 铱星二代星基 ADS-B 系统

Aireon 公司的星基 ADS-B 系统以铱星二代网络为基础,将 ADS-B 1090ES 接收机安装在 75 颗低轨卫星上(其中 66 颗主用卫星,9 颗备用卫星),主用卫星分布在 6 个轨道上,每个轨道 11 颗

卫星,轨道高度 780 km,采用多波束多通道体制,如图 8 所示[25]。表 1 为 Aireon 星基 ADS-B 系统设计监视性能指标。



图 7　Spire 的 AIS 与 ADS-B 数据[23]

Fig. 7　Spire dataset of AIS and ADS-B[23]



图 8　铱星二代卫星星座轨道示意图[25]

Fig. 8　Iridium-Next satellites orbit diagram[25]

表 1　Aireon 星基 ADS-B 系统设计监视性能指标

Table 1　Surveillance performance metrics for Aireon's space-based ADS-B system

监视性能指标	取值
数据链	1090ES
监视数据格式	ASTERIX CAT021,CAT023,CAT025
监视范围	≥99%(全球)
可用性	≥99.9%
传输延迟	≤1.5 s
报文更新间隔	≤8 s(95%概率)

该系统 2015 年开始发射,2019 年 1 月完成星座部署,并于同年投入全球航空监视试运行,促进了国际民航组织洋区飞行间隔标准的更新,缩小了洋区的飞行间隔,并且在冰岛、加拿大地区及大西洋空域开展测试,表 2 为铱星二代星基 ADS-B 系统测试指标。

Aireon 星基 ADS-B 系统设计之初就是以商业运行为目的,是目前唯一系统架构与指标体系支撑管制运行、采用商业化运作、具备可持续运行能力的星基 ADS-B 系统。目前,铱星二代系统经过多次测试验证,能够满足全球非繁忙空域的空管监视应用需求[26-29],图 9 显示了"铱星二代"星基 ADS-B 系统获取的全球空中交通数据。

表 2　铱星二代星基 ADS-B 系统测试指标

Table 2　Iridium-Next space-based ADS-B system test metrics

技术指标	测试场景	取值
报文更新间隔	冰岛凯夫拉维克国际机场	≤8 s(99%)
	冰岛、加拿大实测	≤2 s(绝大部分)
传输延迟	星到地面实测	226 ms(均值)
		312 ms(95%)
		345 ms(99%)
	加拿大实测	≤50 ms(95%)
数据更新频率	北大西洋 Gander FIR	2.87 s(95%)
	北大西洋 Shanwick FIR	3.36 s(95%)
	Nav Canada 非繁忙空域	<2.8 s(95%)



图 9　"铱星二代"星基 ADS-B 系统获取的全球空中交通数据[29]

Fig. 9　Air traffic data captured globally by Iridium-Next space-based ADS-B system[29]

7)其他

西班牙的 Aistech 公司设计了一个 120 颗卫星的星座,将包括地球观测、ADS-B 等应用,2019 年该星座的前 2 颗卫星成功被送入轨道,预计将在未来几年内全面运行。澳大利亚航空服务公司与 Skykraft 公司合作,开发由 210 颗卫星组成的星基通信与监视空管系统,计划于 2022 年 6 月进行首次发射。西班牙 Indra 和 ENAIRE 公司创建了名为 Startical 的项目,计划部署一个由 200 多颗小型卫星组成的星座以改善空中交通管理并提供全球服务。葡萄牙 Tekever 公司计划发射一个 12 颗卫星的星座,将携带包括合成孔径雷达、AIS、ADS-B 在内的多个有效载荷,目前仍处于研究阶段[30-33]。

2.2　国内发展

国内政策方面,在《中国民航航空器追踪监控体系建设实施路线图》和民航科技发展"十三五"规划中均强调了星基 ADS-B 技术是未来发展

关键。国内多家单位针对天地一体化、空间互联网提出计划并发射验证星,包括“鸿雁”星座、“虹云”工程等,均将星基 ADS-B 作为了其可搭载载荷之一。

目前,国内已有多家科研单位紧跟国际星基 ADS-B 系统研究步伐,针对星基 ADS-B 关键技术进行论证与研制,并发射了试验星,具有了一定的技术基础与测试条件。

1) 国防科技大学

国防科技大学设计与研制的“天拓三号”卫星(TianTuo-3,TT-3)于 2015 年 9 月发射,搭载了 ADS-B 接收机进行在轨验证,采用单通道单波束体制,灵敏度约 -93 dBm。其接收系统平均每天接收约 40 万条 ADS-B 报文数据,幅宽超过 2 000 km[34],验证了低轨卫星接收 ADS-B 信号的可行性。图 10显示了“天拓三号”ADS-B 数据获得的 2018 年 5 月全球飞机静态分布图。

2020 年 8 月发射的“天拓五号”卫星(TianTuo-5,TT-5)在其基础上提升了星基 ADS-B 载荷接收性能,增加了天线增益,每天双机能接收星基 ADS-B 报文超过 300 万条,幅宽达到了 4 500 km 的范围[35-36]。

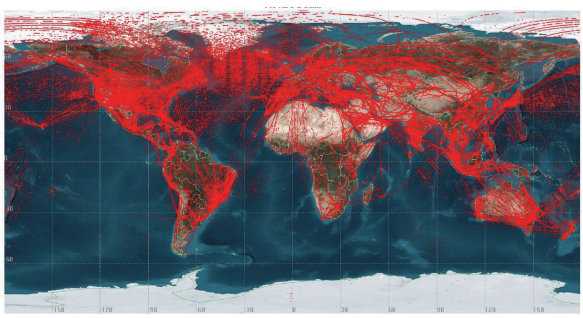


图 10 “天拓三号”ADS-B 数据获得的 2018 年 5 月全球飞机静态分布图[34]

Fig. 10 Static distribution of global aircraft obtained from TianTuo-3 ADS-B dataset in May 2018[34]

2) 上海微小卫星工程中心

上海微小卫星工程中心于 2015 年 9 月发射了“上科大二号”卫星(STU-2),采用 2U 立方体卫星架构,星上搭载有星基 ADS-B 载荷,灵敏度约 -94 dBm[37]。STU-2 一个月获取的飞机数据如图 11 所示,其直接采用丹麦 GomSpace 公司的立方体卫星方案,接收机每天可以收到超过 10 000 条信息,飞机与卫星间的距离可以达到 1 500 km,验证了基于立方体卫星的 ADS-B 载荷技术方案的可行性。

3) 北京和德宇航技术有限公司

北京和德宇航技术有限公司规划建设“天行者”星座,主要搭载 AIS、VDES、ADS-B 等多种有效载荷,将由 48 颗卫星组网,均匀分布在 6 个近地轨道面上。目前星座处于建设中[38-39]。

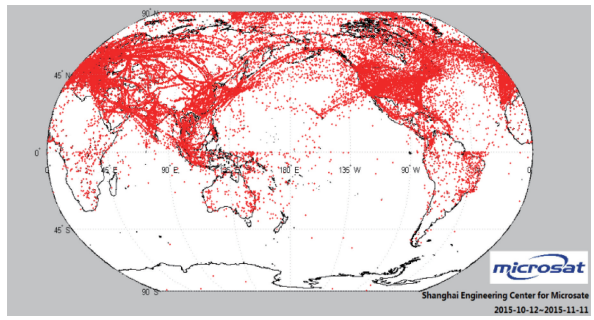


图 11 STU-2 一个月获取的飞机数据[37]

Fig. 11 STU-2 captured aircraft data over one month[37]

3 星基 ADS-B 关键技术研究现状

相对于传统的陆基 ADS-B 系统,星基 ADS-B 系统通过卫星组网建立星座,具有覆盖范围广的特点,但系统复杂度高、部署周期长,同时也存在星基端接收 ADS-B 信号时信噪比低、信号交织严重等问题,严重制约了系统的接收性能。针对星基 ADS-B 各项关键技术,国内外学者已经开展了深入研究。

3.1 微弱信号解调技术

陆基 ADS-B 系统的信号接收范围不超过 450 km,而星基 ADS-B 系统由于轨道高度和覆盖范围的增加,信号接收距离远大于陆基 ADS-B 系统,导致星基 ADS-B 接收机接收到的信号功率和信噪比远远小于陆基 ADS-B 接收设备。欧美在轨技术验证表明,基于低轨卫星的星基 ADS-B 系统接收到的 ADS-B 信号功率将低至 -102 dBm,而陆基 ADS-B 系统接收机设计时未考虑处理功率在 -90 dBm 以下的信号,在星基条件下性能急剧下降[40]。因此,必须突破星基 ADS-B 高灵敏度解调技术,才能推动星基 ADS-B 系统在空管领域应用的进程。

Delovski 等提出了一种改进的 S 模式相关机制,借助 ADS-B 报文 8 位帧头和 5 位 DF17 标志位序列提高信号帧头的同步率[41]。余孙全等在帧头检测时将匹配滤波器的低电平设置为负值,并采用基于置信度的纠错方法,实验验证在 95% 的译码率下得到了 -95 dBm 的灵敏度[42]。Ren 等根据 ADS-B 信号帧格式的特点,提出了一种基于恒虚警率的多准则前导头检测算法,提高低信噪比下 DF17 格式的 ADS-B 信号的检测性能[43]。Qin 和 Yang 研究 ADS-B S 模式消息前导码检测

和数据位提取算法的设计与实现,缓解了传统算法反复验证计算量大以及占用资源过多的问题[44]。Zhang 和 Li 设计了一种 RS 译码器,采用改进的 BM 迭代算法对传输过程中 UAT 模式 ADS-B 信号产生的误码进行纠错,避免了有限域逆运算的出现[45]。Zhang 等研究增强型 ADS-B 接收技术,使用多个幅度样本以及能力更强的错误检测/纠正算法在高干扰环境下提高了正确解调概率[46]。Ren 等提出了一种具有最优阈值的 N 置信度纠错算法和具有恒定复杂度的降序纠错算法,改进低信噪比情况下 ADS-B 的接收性能[47]。

3.2　解交织技术

陆基 ADS-B 系统的信号接收范围相对有限,接收到的信号数量少,而星基 ADS-B 系统覆盖范围广,单个载荷需满足上千架以上航空器同时监视,大量的报文数量导致信号重叠交织情况十分严重。同时和陆基系统相比,星基 ADS-B 系统中信号到达卫星时的信噪比更低、信号之间的功率差更小,进一步加剧了交织信号分离的难度,易造成接收机误解码率高、监视丢点不连续等问题,导致无法获取重要信息而影响航空器安全,降低空域监视性能。为了避免这种情况,将重叠信号分离出来的解交织技术十分关键。

Petrochilos 等研究出了经典的投影算法(projection algorithm, PA)和扩展投影算法(extended projection algorithm, EPA)来在真实环境中分离多个二次监视雷达信号[48-50],其信号结构与 ADS-B 的信号结构相同。国际民航组织制定的标准文件中通过采用对强信息或第 1 个到达的消息进行解码,而放弃其他消息的方式处理重叠报文[40]。陈为桢提出了一种基于干扰抵消的 ADS-B 交叠信号处理算法,对成功检测的大功率信号的参数进行估计来重构信号,从交叠信号中抵消掉重构信号可以成功恢复小功率信号[51]。吴杰等从脉冲幅度特征出发,研究了一种在时域中分离 2 个重叠 ADS-B 信号的方法[52]。吴仁彪等对信号的采样点进行累加,并采用 K 均值聚类方法分类处理 ADS-B 交织信号[53]。Wang 等使用未校准天线阵列提出了一种基于盲自适应波束成形的 ADS-B 交织信号分离方法[54]。Yu 等提出了 2 种 ADS-B 重叠信号分离方法,一种采用基于功率差的自适应阈值调整技术进行分离,另一种通过重构和相消来分离混合信号,并分析了 2 种方法的应用场景[55]。刘慧等研究了松弛改进的快速独立成分分析算法(fast independent component analysis, FastICA)在解交织中的使用,在经典算法中考虑松弛因子,避免了收敛性能取决于初始权值的缺陷[56]。Zhang 等提出了一种分离算法,根据测量的特征值分布确定重叠信号的数量后,利用 MUSIC 算法中的峰值搜索估计信号到达方向来重构混合矩阵,并通过计算重构混合矩阵的 Moore-Penrose 逆来估计分离矩阵[57]。Li 等分析了两个重叠信号的时域特性,提出了一种基于时域对消的 ADS-B 分离算法[58]。

3.3　多波束接收技术

由于星基 ADS-B 系统覆盖范围远大于陆基 ADS-B 系统的监视范围,信号冲突概率大大增加,而星基 ADS-B 系统多采用微小卫星,通过软件算法解决交织信号分离问题的计算复杂度和资源要求对小卫星能力是一项巨大挑战。国内外另一种解决星基 ADS-B 信号冲突问题的有效方法是采用多波束接收的方式,将卫星可视范围分为几个波束覆盖区,以降低单个波束覆盖范围内的航空器数量,有效降低信号冲突概率,提升星基 ADS-B 监视性能。

Bettray 等研究了 ADS-B 多波束天线的工程模型,提出一种星基 ADS-B 信号的 7 波束接收天线来提升接收性能[59]。Budroweit 等讨论了多通道 ADS-B 接收机的设计与实现,并给出了初步的测试结果[60]。Yu 等提出了一种自适应多波束形成方法,通过动态调整波束方向图来有效减小 ADS-B 信号冲突概率,提升 ADS-B 信息的更新概率[61];其于 2020 年又介绍了一种具有自校准功能的数字波束形成系统,采用一个4 ×4 微带阵列天线和一个 16 通道接收机,可以同时产生 19 个以上波束大小和方向可灵活控制的独立波束,减少信号碰撞,提高星基 ADS-B 系统的可靠性[62]。

3.4　防欺骗技术

ADS-B 消息以公开形式广播,由于缺乏加密和身份验证等基本安全措施使其在应用具有被窃听、干扰、消息注入和修改等风险。恶意用户可以很容易通过设备实时进行窃听,在 1 090 MHz 频率上发送干扰信号妨碍消息正常接收,或通过发送人工 ADS-B 信号向系统注入虚假航空器信息,以伪装不存在的航空器或受信任的航空器。以上攻击威胁可能导致系统可用性和可靠性中断,对 ADS-B 系统造成严重破坏。因此,随着 ADS-B 系统广泛应用,为了消除攻击行为的影响,需要进行欺骗干扰检测,以避免攻击数据注入决策流。

Amin[63] 和 Kacem[64] 等研究了基于密码学思想保护 ADS-B 信号的方法,但添加加密或认证机

制需要修改当前的协议体系结构。Ghose 和 Lazos 提出了一种基于 ADS-B 传输射频特性的物理层验证方法来验证航空器的速度和位置,利用短相干时间和多普勒扩展现象检测恶意广播的 ADS-B 信息[65]。Kim 等添加时间戳数据,根据接收到 GPS 信号时间与到达接收机时间的时间差对 ADS-B 信息进行安全性验证,并分析了时间戳的范围与精度[66]。Naganawa 等利用扇形天线的振幅单脉冲技术测量到达角(angle of arrival, AoA),同时根据 ADS-B 报文中航空器的位置计算预期 AoA,通过比较 2 种 AoA 检测欺骗[67]。Habler 和 Shabtai 基于 LSTM 编解码算法检测异常 ADS-B 消息,通过分析合法 ADS-B 消息序列对飞行路线建模,航空器可以自主评估接收到的 ADS-B 消息与合法飞行路径的偏差[68]。Ying 等对 ADS-B 数据的特征进行挖掘,建立分类模型,提出一种使用深度神经网络监测 ADS-B 攻击行为的方法[69]。吕宗平等利用全球导航卫星系统(GNSS)完好性信息对虚假 ADS-B 报文进行识别,通过水平保护级(HPL)与导航不确定类别(NUC)的一致性比对防止恶意欺骗[70]。Li 等基于黏性层次 Dirichlet 过程的隐马尔可夫模型建立 ADS-B 数据序列隐状态预测生成模型,通过分析隐藏状态序列的上下文偏差信息检测攻击数据[71]。吴庆基于多普勒频偏变化与位置变化的对应关系,提出一种深度学习的 ADS-B 防欺骗方法,利用改进型 AlexNet 识别虚假信号[72]。

3.5 星座与星间链路技术

虽然卫星上的 ADS-B 接收机覆盖范围比陆基系统大,但单颗卫星无法完全满足全球连续覆盖的空管监视要求,需要建立和部署卫星星座来集成并提高单颗卫星的性能。此外,具有良好可靠星间链路的星座可以完成监视数据的即时、连续传输与下载。低轨卫星星间链路多具有较大的动态性,可能会受到多普勒频移的显著影响,因此如何设计一个合适的星座构型并优化以及建立稳定可靠的星间链路,是需要研究的重要内容之一。

Nguyen 等对链路可能性和卫星星座跟踪能力的性能进行分析,定义了包括卫星数量、轨道高度、最大间隔时间等在内的星座参数,提出一种低成本的星基 ADS-B 系统[73]。Nag 等基于模拟空中交通数据描述了一种适用于区域覆盖任务的立方体卫星星座设计方法,以提高监视区的安全性和态势感知能力[74]。Chen 等分析星基系统的设计约束,设计了由 81 颗微纳卫星组成的 Walker 卫星星座系统,通过数值仿真验证了卫星系统设计的合理性[75]。Leyva-Mayorga 等提出了一个在低轨星座中动态建立星间链路的框架和相应的算法,以确保所有卫星在任何时候都至少具备一个完整的星间链路连接[76]。Guo 等提出了一个卫星星座性能评估模型,综合考虑覆盖范围、星间链路、网络传输等方面给出了一个 55 颗卫星的星座设计方案[77]。考虑到低轨卫星星座规模,部署策略对发射次数和待发射卫星有重大影响,de Weck 等论证了分阶段部署在低轨通信卫星星座中的价值,并针对全球星应用进行了研究[78]。Lee 等结合铱星和全球星的经验,提出一种使用遗传算法优化的分阶段部署策略,使星座大小适应市场需求[79]。

3.6 路由转发技术

虽然陆基 ADS-B 系统已经较为成熟,但卫星网络传输与地面网络传输存在较大的差别。中国目前没有能力在世界各地部署星基 ADS-B 地面站,因此所有 ADS-B 数据将通过卫星网络汇聚到中国地面站上空的卫星,距离汇聚节点较近的星间链路更有可能出现拥塞,在这种情况下大多数路由算法和流量平衡方法会导致大量不必要的消耗。同时由于航路航线的设置,ADS-B 数据在全球的分布很不均衡,在繁忙空域的数据量可能是非繁忙空域的数十倍,这种不均衡分布在进行汇聚时会导致路径流量的不均衡,加剧了某些链路的负担,容易超过链路的最大传输能力而造成丢包现象或进一步导致卫星网络的拥塞。因此,针对星基 ADS-B 系统的网络路由转发技术进行研究是必要且关键的。

目前,卫星网络路由技术根据实现方式可以大致分为虚拟拓扑法、虚拟节点法和动态拓扑法 3 类,流量均衡技术可以按照全局方案与局部方案分为 2 类。Mohorčič等将流量同时分布在最短路径和次短路径上,提出了 ALR 多径流量均衡路由算法[80]。Bai 等提出了 CEMR 多径流量均衡算法,将流量分布到 K 条路径上解决拥塞问题[81]。Hussein 等针对全球交通流量分布采用流量预测来进行流量均衡[82]。Taleb 等提出了 ELB 局部流量均衡算法,在一跳范围内进行均衡[83]。Dong 等提出针对不同业务需求,组合使用基于人工智能的 QoS 算法、多层卫星网络架构、多路径路由算法以提升路由性能[84]。叶红军等针对星基航空监视系统设计了星地一体化综合监视网络,并进行了理论分析与网络拓扑验证[85]。郑晓冬等考虑低轨星间链路网络中航空安全信息传输的路由规划需求进行了算法设计,分别提出了多站就近

落地、高可靠备份传输以及分区负载均衡 3 种算法,并基于星间链路网络仿真平台进行了试验验证[86]。

3.7　监视性能评估技术

星基 ADS-B 是空天地一体化的大系统,其监视数据量大、动态变化快,监视数据质量如位置误差、报文更新周期、数据传输延迟等受 GNSS 导航系统、星间链路传输等多种因素影响。为验证星基 ADS-B 系统的可用性及其监视性能能否达到空中交通管理系统的最低接入标准,通过试验卫星接收的数据,研究适用于星基 ADS-B 系统的监视性能评价指标,开展在轨运行数据分析与性能评估并确立相应评估方法,及时发现和告警达不到空中交通管制所需监视性能等级的情况将具有十分重要的意义。

Delovski 等依据陆基 ADS-B 监视要求给出了包括目标获取概率、目标检测概率、目标识别概率和消息识别概率在内的 4 个星基 ADS-B 监视性能参数计算方法,并对 Proba-V 卫星 ADS-B 载荷在北大西洋、太平洋越洋航线上的监视性能进行了评估[11]。Schäfer 等开发和部署了面向中欧地区的 ADS-B 系统 OpenSky,并分析了其获取的数据[87]。van der Pryt 等建立了一个确定到达卫星的功率水平的模型,为 CanX-7 性能提供评估[88]。Garcia 等提出了系统和接收机模型,评估是否能够实现提供服务所需的更新间隔[89];并于 2017 年进行了铱星二代卫星的在轨试验情况分析,采用了陆基参考发射机,用于铱星系统性能模型的深入校准[26];其在 2018 年铱星二代系统较为完整之后,又针对不同的 ADS-B 发射机单独进行实验,定义了可用性指标,并分析了铱星二代系统的关键指标[27]。钟建华等开发了一个 ADS-B 数据质量分析软件,通过解析报文计算数据的完好率、漏点率和跳点率[90]。于克非和赵嶷飞利用“天拓三号”卫星数据对更新时间间隔、数据完好性和数据准确性 3 项指标进行统计分析[91-92]。王运帷提出一套 ADS-B 数据质量评估体系,从接收能力、刷新时间间隔、精确性、连续性 4 个角度进行评估分析[93]。

4　北航联合团队星基 ADS-B 研究情况

北航四川九洲空管科技有限责任公司(简称九洲空管公司)、中国民航大学、西华大学等单位经过多年合作,形成“产学研用”联合科研团队。张军院士带领科研团队,在国家重大科研项目支持下,依托国家空管监视与通信系统工程技术研究中心等多个国家级和省部级科研平台,共同开展星基 ADS-B 研究与技术攻关,并发射了试验验证星。

4.1　星基 ADS-B 关键技术研究

北航张学军等设计了以相干解调为基础的星基 ADS-B 联合解调算法,重点解决了信号帧头的同步和信号的频偏问题,在之前的研究基础上有效提高了解调灵敏度[94]。九洲空管公司冯涛和梁俊提出了一种帧头检测方法,并使用遗传算法对检测到的微弱信号进行高精度频率和相位估计,实现直接相干解调[95];通过反演修正的方法在低信噪比下实现了 ADS-B 信号译码,提升星基 ADS-B 接收信号解调性能[96]。张学军等利用最大似然比法则建立数学模型提取低置信度矩阵,对低置信度位在错误图表中对应的错误图样进行异或纠错,解决了在任意位置的 5 个及以下的错误纠错,提高了译码率[97];并充分考虑星上硬件资源,结合 FPGA 技术完成星基 ADS-B 接收机信号处理模块的实现,减少了FPGA资源消耗与处理延时[98]。

北航刘凯等研究了基于泊松过程的星基 ADS-B 信号高阶碰撞模型,并据此提出一种迭代投影算法用于星基 ADS-B 信号分离,在仿真中比较了算法与 MDA 和 EPA 的性能差异[99]。北航张涛等通过对混叠信号进行白化处理判断重叠信号数量,依此在每个时间段内进行矩阵变化,利用波束赋形矩阵得到分离信号,降低了多信号碰撞情况下分离算法的复杂度[100]。

针对防欺骗问题,九洲空管的李家蓬等分别针对单通道和四通道 ADS-B 地面站进行防欺骗方法研究,对单通道采用时序模式特征验证、几何相符、意图相符等多种验证手段组合验证;对四通道则结合四通道脉冲幅度测向法和信号幅度-距离测量法实现目标定位,判断虚假目标,并将相应研究成果应用于星基四通道接收机的研究中[101-102]。安强等则利用信号 TDOA 解算目标空间位置,同 ADS-B 报告位置比对判断目标的真实性与有效性[103]。

针对星座设计,北航张学军等结合当前空管发展与监视需求,提出了满足全球覆盖的 60 星、96 星的 2 种星座构型方案,同时考虑卫星重量与尺寸包络,探讨了一箭多星的星座部署方案;并采用“瓦式”架构射频前端设计,将辐射阵元通过功分网络合成后与后端射频模块按照分层结构层层堆砌在基板上,对星间链路技术展开研究。

北航张涛等建立路径中间节点到目的节点之间的 M 条待选转发路径,在目的节点周围建立多径,共同负担全球 ADS-B 业务流量业务,缓解目的节点周围的拥塞和延时,提高网络传输的可靠性[104]。针对 ADS-B 业务在低轨卫星网络中产生汇聚的特点,提出一种可以缓解网络拥塞的动态路由协议,引入了链路代价机制和链路状态反馈机制,结合最短路径算法,缓解空间信息网络的链路拥塞,降低丢包率[105-106]。张学军等提出了一种基于目标节点泛洪的全局-局部路由算法 DF-GL,在全局范围内使用多目的节点的泛洪算法,在汇聚节点的 N 跳范围内进行多径路由优化,随后通过动态代价估计和概率路径选择对算法进一步改进,并利用空间网络仿真(SNS)系统上进行仿真验证,提高了路由算法在实际场景中的可用性[107]。

中国民航大学刘海涛等理论推导了在仅有 ADS-B 信号干扰情况下星基接收机的监视容量[108],之后又基于共信道干扰模型理论分析 ADS-B 消息被星基接收机正确接收的概率以及位置报告更新间隔,并结合航空器的空-时分布特性分析了星基系统的监视容量,对单通道接收机能否满足当前监视需要作评估[109]。北航张学军等提出通过水平位置更新率 PUR、气压高度更新率 PUH、航空器身份更新率 PUI 计算监视系统所需监视性能可用性[110];此外提出一种监视性能评估指标形成方法,利用航空器的航行信息建立参考航迹,依据预设有效数据标准,通过比对参考航迹中数据项与航行信息数据项筛选并统计监视性能指标[111]。同时,张学军等搭建了国内首套空事卫星系统半实物仿真演示验证平台,该平台可将实物载荷与真实卫星映射为虚拟卫星网络中的卫星节点,通过真实数据和模拟数据共同模拟星座运行状态,评估系统监视性能。

4.2 北航星基 ADS-B 验证星

北航于 2019 年发射了星基 ADS-B 平台技术验证星,验证了所需卫星平台的功能及性能指标,并在 2020 年发射了北航星基 ADS-B 技术验证星,如图 12 所示,其上搭载张学军等针对星基 ADS-B 信号传输衰减大、空域内信号易产生冲突等重点问题研制的星基 ADS-B 航空监视载荷开展在轨技术验证。该载荷是中国首个面向全球空管需求,针对全球航班广域监视应用研制的载荷,突破了星基 ADS-B 高灵敏度解调算法、数字多波束接收等关键技术。

自成功发射以来,卫星工作正常,获得了全球空域的大量飞行数据,经过数据分析可知,繁忙空域内,平均每小时接收 ADS-B 消息数超过 100 万条;最远探测距离超过 2 500 km;半径 800 km 的设计覆盖范围内,95% 位置消息更新间隔低于 8 s,报文更新率初步满足空管运行需求,性能指标达到国内领先、国际先进水平[112]。北航星基 ADS-B 系统技术验证星获取的全球航空器分布如图 13 所示。

根据北航星基 ADS-B 技术验证星在轨试验期间接收的数据分析,充分考虑运输航空起飞/降落、离场/进近、巡航各个阶段运行场景与通用航空监视跟踪运行场景,其具备在全球范围内对运输航空器各飞行阶段进行监视跟踪的能力,具备对在全球范围内对通用航空飞行器监视跟踪的能力。

国外对星基 ADS-B 系统的研究起步较早,2013 年发射的 Proba-V 卫星就首先验证了卫星接收 ADS-B 信号的可行性,而 GOMX-1 卫星则验证了基于立方星的 ADS-B 载荷的可行性,随后 Aireon 公司建设的铱星二代系统采用相控阵天线、多波束接收的星基 ADS-B 载荷,具备了提供空管监

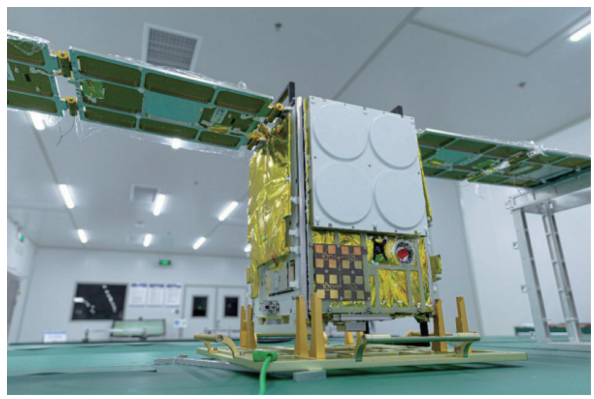


图 12 北航星基 ADS-B 技术验证星

Fig. 12 Beihang space-based ADS-B technology verification satellite

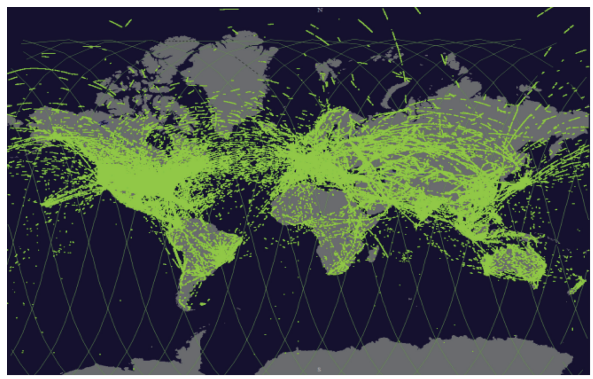


图 13 北航星基 ADS-B 系统技术验证星获取的全球航空器分布

Fig. 13 Global aircraft distribution obtained by Beihang space-based ADS-B technology verification satellite

视服务的能力,也是目前国际上唯一建成并投入商业试运行的星基 ADS-B 系统。

国内 2015 年后有一些不成系统的低轨 ADS-B 监视试验卫星,国防科技大学、上海微小卫星工程中心相继发射了试验星“天拓三号”和“上科大二号”,均实现了星基 ADS-B 信号的接收,这些试验星参考国外前期研究,基于单颗卫星对 ADS-B 信号接收的可行性做了验证,无法满足全球空管监视应用需求。2020 年国防科技大学沿用“天拓三号”技术体制,发射了改进的“天拓五号”,在星基 ADS-B 接收性能上有所提升,但仍旧无法满足空管运行需求。

根据国内已发射的低轨 ADS-B 卫星测试情况,北航星基 ADS-B 技术验证星面向空管监视实际运行设计与研制,在高灵敏度接收算法、数字多波束合成等关键技术上有所突破,初步满足了空管运行需求。其作为单星的技术验证星,与目前已经成功组网、唯一商业化运作的铱星二代系统还存在一定差距,在性能指标与星间链路等方面还有改进与提升的空间。

但北航星基 ADS-B 技术验证星在轨验证的成功仍标志着中国在星基航空监视方面达到国际先进水平,未来将有能力给航空用户提供基础的全球航空监视和跟踪服务,极大推动中国空事卫星系统全面组网建设的进程,对中国军民航和通航的发展具有极为重要的意义。

5　结 束 语

星基 ADS-B 系统利用低轨卫星空间覆盖优势,搭载 ADS-B 接收机,实现对全球高中低空各类空域的全覆盖,可为所有安装了 ADS-B 设备的航空器提供空中交通管制监视服务,确保其不受陆基系统地形限制的影响,解决陆基空管系统在洋区、荒漠、偏远地区的盲区问题,提升全球空管综合监控能力,实现全球范围内无缝空域监视和空中交通管理。北航、九洲空管公司、中国民航大学等单位的联合团队针对星基 ADS-B 进行了广泛而深入的研究,成功发射的北航星基 ADS-B 技术验证星为中国空管系统全球化建设奠定了基础。

为了推动中国星基 ADS-B 系统的全面建设,未来还需要进一步发展并突破关键技术,提升其解调灵敏度,改进其相控阵天线和多通道处理设计等,以提高有效覆盖范围,适应不同轨道高度和不同航空用户的需求;此外,要想为全球范围内的航空用户提供监视服务必须完成卫星全球组网,其中包括星座与卫星平台设计、低时延星上路由等多项关键技术,可以结合航空网络流量预测,主动分散卫星网络流量以防止拥塞;另外,还需结合现有标准进一步完善监视性能评估体系,支撑星基 ADS-B 系统运行。

同时随着中国军事现代化进程不断推进,军民航空域使用需求多元化、多样化趋势明显加快,空域使用结构需求更加复杂,空域使用矛盾将继续加剧。星基 ADS-B 系统的发展时刻都有军方身影,其不仅是服务全球空中交通的一种高效基础设施,更是军航作战全球飞行的一种支撑体系。未来需要综合考虑军民航空管在国家战略和全球化发展中的作用,建立统一的兼容技术架构,将航天航空技术要求和军民航用户需求整合在一起,以发展需求为牵引、拓展军民航应用,实现与航空企业、地区经济协同发展,实现资源共用、信息共享为重点的军民航协同发展。

参考文献(References)

[1] 中国民用航空局. 中国民用航空 ADS-B 实施规划[S]. 北京:中国民用航空局,2012:1-45.
Civil Aviation Administration of China. ADS-B implementation plan for civil aviation in China[S]. Beijing:Civil Aviation Administration of China,2012:1-45(in Chinese).

[2] International Civil Aviation Organization Asia and Pacific Office. Guidance material on comparison of surveillance technologies (GMST)[R]. Montreal:ICAO,2007:1-44.

[3] BLOMENHOFER H, PAWLITZKI A, ROSENTHAL P, et al. Space-based automatic dependent surveillance broadcast (ADS-B) payload for in-orbit demonstration[C]// 2012 6th Advanced Satellite Multimedia Systems Conference (ASMS) and 12th Signal Processing for Space Communications Workshop (SPSC). Piscataway:IEEE Press,2012:160-165.

[4] Canada. The concept of space-based reception of automatic dependent surveillance-broadcast (ADS-B): A38-WP/132[R]. Montreal:ICAO,2013:2-3.

[5] Flight Safety Foundation. Benefits analysis of space-based ADS-B[R]. Alexandria:Flight Safety Foundation,2016:18-35.

[6] 中国民用航空局. 中国民航航空器追踪监控体系建设实施路线图[S]. 北京:中国民用航空局,2017:1-7.
Civil Aviation Administration of China. Implementation roadmap for the construction of aircraft tracking and monitoring system for civil aviation in China[S]. Beijing:Civil Aviation Administration of China,2017:1-7(in Chinese).

[7] 马斌. 卫星系统将为全球航班跟踪保驾护航[J]. 中国无线电,2015(11):20.
MA B. Satellite system will escort global flight tracking[J]. China Radio,2015(11):20(in Chinese).

[8] ICAO Western and Central Africa Office. ASEPS using ADS-B trials in parts of the nat airspace:SAT/24-WP/17[R]. Montreal:ICAO,2019:2-4.

[9] BRODSKY Y, RIEBER R, NORDHEIM T. Balloon-borne air traffic management (ATM) as a precursor to space-based ATM [J]. Acta Astronautica, 2012, 70: 112-121.

[10] DELOVSKI T, WERNER K, RAWLIK T, et al. ADS-B over satellite the world's first ADS-B receiver in space [C]// Small Satellite Systems and Services Symposium, 2014.

[11] WERNER K, BREDEMEYER J, DELOVSKI T. ADS-B over satellite: Global air traffic surveillance from space [C]// 2014 Tyrrhenian International Workshop on Digital Communications-Enhanced Surveillance of Aircraft and Vehicles (TIWDC/ESAV). Piscataway: IEEE Press, 2014: 47-52.

[12] ALMINDE L K, CHRISTIANSEN J, LAURSEN K K, et al. GOMX-1: A nano-satellite mission to demonstrate improved situational awareness for air traffic control [C]// 26th Annual AIAA/USU Conference on Small Satellites. Reston: AIAA, 2012.

[13] ALMINDE L, KAAS K, BISGAARD M, et al. GOMX-1 flight experience and air traffic monitoring results [C]// 28th Annual AIAA/USU Conference on Small Satellites. Reston: AIAA, 2014.

[14] GomSpace. GOMX-3 (GomSpace Express-3) [EB/OL]. (2015-10-02) [2022-03-25]. https://directory.eoportal.org/web/eoportal/satellite-missions/g/gomx-3.

[15] LÉON L, KOCH P, WALKER R. GOMX-4: The twin European mission for IOD purposes [C]// 32nd Annual AIAA/USU Conference on Small Satellites. Reston: AIAA, 2018.

[16] GomSpace. GomX-4 (GomSpace Express-4) mission [EB/OL]. (2017-09-06) [2022-03-25]. https://directory.eoportal.org/web/eoportal/satellite-missions/g/gomx-4.

[17] VINCENT R, VAN DER PRYT R. The CanX-7 nanosatellite ADS-B mission: A preliminary assessment [J]. Positioning, 2017, 8(1): 1-11.

[18] BENNETT I, COTTEN B, ZEE R E. On-orbit results from the CanX-7 ADS-B payload [C]// 11th IAA Symposium on Small Satellites for Earth Observation, 2017.

[19] HENRY C, VAN WAGENEN J. Global, ADS-B technologies complete space-based aircraft tracking demonstration [EB/OL]. (2014-09-18) [2022-04-12]. https://www.aviationtoday.com/2014/09/18/global-adsb-technologies-complete-space-based-aircraft-tracking-demonstration/.

[20] RAY J. Space-based ADS-B [EB/OL]. (2013-09-01) [2022-04-12]. https://www.aviationtoday.com/2013/09/01/space-based-ads-b/.

[21] BAKER K. Space-based ADS-B: Performance, architecture and market options [C]// 2019 Integrated Communications, Navigation and Surveillance Conference (ICNS). Piscataway: IEEE Press, 2019: 1-18.

[22] TRAUTVETTER C. Spire to provide lower-cost ADS-B aircraft tracking [EB/OL]. (2016-12-05) [2022-03-26]. https://www.ainonline.com/aviation-news/aerospace/2016-12-05/spire-provide-lower-cost-ads-b-aircraft-tracking.

[23] CAPPAERT J. The spire small satellite network [M]// PELTON J N, MADRY S. Handbook of small satellites: Technology, design, manufacture, applications, economics and regulation. Berlin: Springer, 2020: 1101-1121.

[24] Spire (Lemur / Minas) [EB/OL]. (2021-12-26) [2022-04-08]. https://www.newspace.im/constellations/spire.

[25] AIREON. Global ATS Surveillance [EB/OL]. [2022-04-06]. https://aireon.com/products/global-ats-surveillance/.

[26] GARCIA M A, DOLAN J, HOAG A. Aireon's initial on-orbit performance analysis of space-based ADS-B [C]// 2017 Integrated Communications, Navigation and Surveillance Conference (ICNS). Piscataway: IEEE Press, 2017: 4A1-1.

[27] GARCIA M A, DOLAN J, HABER B, et al. A compilation of measured ADS-B performance characteristics from Aieron's on-orbit test program [C]// international symposium Enhanced Solutions for Aircraft and Vehicle Surveillance Applications, 2018.

[28] CANSO. Space-based ADS-B lifts off [EB/OL]. (2019-05-14) [2022-04-06]. https://canso.org/airspace-q2-2019-space-based-ads-b-lifts-off/.

[29] BELLAMY W. EASA issues first space-based ADS-B surveillance as a service certification [EB/OL]. (2019-06-30) [2022-04-06]. https://www.aviationtoday.com/2019/06/05/easa-issues-first-space-based-ads-b-surveillance-as-a-service-certification/.

[30] PICOLA BARÉ A. Study of the benefits and applications of LEO (low Earth orbit) for communications and definition of space new business plan: Surveillance & broadcast for aircraft [D]. Barcelona: Universitat Politècnica de Catalunya, 2021: 14-15.

[31] HARRISON R. Using satellites to improve air traffic control coverage [EB/OL]. (2022-02-09) [2022-04-10]. https://spaceaustralia.com/news/using-satellites-improve-air-traffic-control-coverage.

[32] CANSO. ENAIRE and Indra will launch a constellation of satellites into orbit to improve air traffic management [EB/OL]. (2021-05-18) [2022-04-09]. https://canso.org/enaire-and-indra-will-launch-a-constellation-of-satellites-into-orbit-to-improve-air-traffic-management/.

[33] HENRY C. Portuguese company embarks on first domestic satellite project [EB/OL]. (2017-11-10) [2022-04-09]. https://spacenews.com/portuguese-company-embarks-on-first-domestic-satellite-project/.

[34] CHEN L H, YU S Q, CHEN Q, et al. Data reception analysis of ADS-B on board the TianTuo-3 satellite [J]. Journal of Physics: Conference Series, 2020, 1438(1): 012030.

[35] NI J S, CHEN L H, YU S Q, et al. Analysis and application of spaceborne mode S and ADS-B data fusion [C]// 2021 International Conference on Big Data Engineering and Education (BDEE). Piscataway: IEEE Press, 2021: 51-55.

[36] 倪久顺,陈利虎,余孙全,等. 星载 ADS-B 相关研究进展及展望[J]. 中国空间科学技术,2022,42(1):30-37.
NI J S, CHEN L H, YU S Q, et al. A review for space-based ADS-B [J]. Chinese Space Science and Technology, 2022, 42(1): 30-37 (in Chinese).

[37] WU S, CHEN W, CHAO C. The STU-2 CubeSat mission and in-orbit test results [C]// 30th AIAA/USU Conference on Small Satellites, 2016.

[38] 北京和德宇航技术有限公司. "天行者"星座和应用服务 [EB/OL]. [2022-04-09]. http://www.head-aerospace.com/#/Mainwork/Skywalker.

China HEAD Aerospace Technology Co. "Skywalker" constellation and application services[EB/OL]. [2022-04-09]. http://www. head-aerospace. com/#/Mainwork/Skywalker (in Chinese).

[39] 李国利,朱霄雄. 高分九号 03 星发射成功搭载发射皮星三号 A 星、和德五号卫星[J]. 科技传播,2020,12(13):1.
LI G L, ZHU X X. Gaofen IX 03 successfully launched with the launch of Pixin III A and He De V satellites [J]. Public Communication of Science & Technology, 2020, 12(13): 1(in Chinese).

[40] RTCA. Minimum operational performance standards (MOPS) for 1 090 MHz extended squitter automatic dependent surveillance-broadcast (ADS-B) and traffic information services-broadcast (TIS-B): DO-260B [S]. Washington, D. C.: RTCA, 2011.

[41] DELOVSKI T, BREDEMEYER J, WERNER K. ADS-B over satellite coherent detection of weak Mode-S signals from low earth orbit[C]// Small Satellites Systems and Services, 2016.

[42] 余孙全,陈利虎,李松亭,等. 高灵敏度星载 ADS-B 信号解调算法[J]. 太赫兹科学与电子信息学报,2018,16(5):886-891.
YU S Q, CHEN L H, LI S T, et al. High sensitivity detection algorithm for space-based ADS-B[J]. Journal of Terahertz Science and Electronic Information Technology, 2018, 16(5): 886-891(in Chinese).

[43] REN P, WANG J X, SONG H W, et al. A novel multi-criteria preamble detection algorithm for ADS-B signals[J]. IEEE Access, 2019, 7: 97319-97332.

[44] QIN J X, YANG J. Design and implementation of spaceborne ADS-B message detection algorithm[J]. DEStech Transactions on Engineering and Technology Research, 2018, 298(8): 295-302.

[45] ZHANG C Z, LI N R. The design and implementation of UAT mode ADS-B signal's RS decoder[C]// 2014 23rd International Conference on Computer Communication and Networks (ICCCN). Piscataway: IEEE Press, 2014: 1-6.

[46] ZHANG C Z, ZHANG C L, WANG Y. Enhanced ADS-B reception techniques research[C]// 10th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM 2014). Piscataway: IEEE Press, 2014: 657-661.

[47] REN P, WANG J X, ZHANG P X. Novel error correction algorithms for ADS-B signals with matched filter based decoding [J]. Physical Communication, 2019, 36: 100788.

[48] PETROCHILOS N, VAN DER VEEN A J. Algorithms to separate overlapping secondary surveillance radar replies[C]// 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing. Piscataway: IEEE Press, 2004: 2-49.

[49] PETROCHILOS N, GALATI G, MENE L, et al. Separation of multiple secondary surveillance radar sources in a real environment by a novel projection algorithm[C]// Proceedings of the 5th IEEE International Symposium on Signal Processing and Information Technology. Piscataway: IEEE Press, 2005: 125-130.

[50] PETROCHILOS N, GALATI G, PIRACCI E. Projection techniques for separation of multiple secondary surveillance radar sources in a real environment[C]// 4th IEEE Workshop on Sensor Array and Multichannel Processing. Piscataway: IEEE Press, 2006: 344-348.

[51] 陈为桢. 星载 ADS-B 中频接收机分离算法的研究与实现[D]. 成都:电子科技大学,2017:1-59.
CHEN W Z. The research and implementation of spaceborne ADS-B IF receiver separation algorithm[D]. Chengdu: University of Electronic Science and Technology of China, 2017: 1-59 (in Chinese).

[52] 吴杰,郭建华,蒋凯,等. ADS-B 二重交织信号时域分离算法[J]. 通信技术,2017,50(10):2184-2189.
WU J, GUO J H, JIANG K, et al. ADS-B double intertwined signal separation algorithm in time domain[J]. Communications Technology, 2017, 50(10): 2184-2189(in Chinese).

[53] 吴仁彪,吴琛琛,王文益. 基于累加分类的 ADS-B 交织信号处理方法[J]. 信号处理,2017,33(4):572-576.
WU R B, WU C C, WANG W Y. A method of overlapped ADS-B signal processing based on accumulation and classification [J]. Journal of Signal Processing, 2017, 33(4): 572-576 (in Chinese).

[54] WANG W Y, WU R B, LIANG J L. ADS-B signal separation based on blind adaptive beamforming[J]. IEEE Transactions on Vehicular Technology, 2019, 68(7): 6547-6556.

[55] YU S Q, CHEN L H, LI S T, et al. Separation of space-based ADS-B signals with single channel for small satellite[C]// 2018 IEEE 3rd International Conference on Signal and Image Processing. Piscataway: IEEE Press, 2018: 315-321.

[56] 刘慧,倪育德,刘鹏. 基于松弛改进 FastICA 算法的星基 ADS-B 信号分离[J]. 电讯技术,2020,60(2):203-209.
LIU H, NI Y D, LIU P. Space-based ADS-B signal separation based on loose modified FastICA[J]. Telecommunication Engineering, 2020, 60(2): 203-209(in Chinese).

[57] ZHANG C S, ZHANG T, ZHANG H. Overlapping ADS-B signals separation algorithm based on MUSIC[C]// 2019 6th International Conference on Information Science and Control Engineering (ICISCE). Piscataway: IEEE Press, 2019: 1094-1098.

[58] LI K D, KANG J, REN H, et al. A reliable separation algorithm of ADS-B signal based on time domain [J]. IEEE Access, 2021, 9: 88019-88026.

[59] BETTRAY A, LITSCHKE O, BAGGEN L. Multi-beam antenna for space-based ADS-B[C]// 2013 IEEE International Symposium on Phased Array Systems and Technology. Piscataway: IEEE Press, 2013: 227-231.

[60] BUDROWEIT J, JAKSCH M P, DELOVSKI T. Design of a multi-channel ADS-B receiver for small satellite-based aircraft surveillance[C]// 2019 IEEE Radio and Wireless Symposium. Piscataway: IEEE Press, 2019: 1-4.

[61] YU S Q, CHEN L H, LI S T, et al. Adaptive multi-beamforming for space-based ADS-B[J]. Journal of Navigation, 2019, 72(2): 359-374.

[62] YU S Q, CHEN L H, FAN C G, et al. Integrated antenna and receiver system with self-calibrating digital beamforming for space-based ADS-B [J]. Acta Astronautica, 2020, 170: 480-486.

[63] AMIN S, CLARK T, OFFUTT R, et al. Design of a cyber security framework for ADS-B based surveillance systems[C]// 2014

Systems and Information Engineering Design Symposium (SIEDS). Piscataway: IEEE Press, 2014: 304-309.

[64] KACEM T, WIJESEKERA D, COSTA P, et al. Key distribution mechanism in secure ADS-B networks[C]// 2015 Integrated Communication, Navigation and Surveillance Conference (ICNS). Piscataway: IEEE Press, 2015: 1-3.

[65] GHOSE N, LAZOS L. Verifying ADS-B navigation information through Doppler shift measurements[C]// 2015 IEEE/AIAA 34th Digital Avionics Systems Conference (DASC). Piscataway: IEEE Press, 2015: 2A-4A.

[66] KIM Y, JO J Y, LEE S. A secure location verification method for ADS-B[C]// 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC). Piscataway: IEEE Press, 2016: 1-10.

[67] NAGANAWA J, TAJIMA H, MIYAZAKI H, et al. ADS-B anti-spoofing performance of monopulse technique with sector antennas[C]// 2017 IEEE Conference on Antenna Measurements & Applications. Piscataway: IEEE Press, 2017: 87-90.

[68] HABLER E, SHABTAI A. Using LSTM encoder-decoder algorithm for detecting anomalous ADS-B messages[J]. Computers & Security, 2018, 78: 155-173.

[69] YING X H, MAZER J, BERNIERI G, et al. Detecting ADS-B spoofing attacks using deep neural networks[C]// 2019 IEEE Conference on Communications and Network Security. Piscataway: IEEE Press, 2019: 187-195.

[70] 吕宗平, 倪育德, 陈君, 等. 基于GNSS完好性的ADS-B防欺骗[J]. 雷达科学与技术, 2018, 16(4): 359-365.
LYU Z P, NI Y D, CHEN J, et al. Anti-spoofing for ADS-B based on GNSS integrity[J]. Radar Science and Technology, 2018, 16(4): 359-365(in Chinese).

[71] LI T Y, WANG B H, SHANG F T, et al. Dynamic temporal ADS-B data attack detection based on sHDP-HMM[J]. Computers & Security, 2020, 93: 101789.

[72] 吴庆. 基于深度学习的ADS-B欺骗式干扰检测[D]. 天津: 中国民航大学, 2020: 1-48.
WU Q. ADS-B spoofing detection based on deep learning[D]. Tianjin: Civil Aviation University of China, 2020: 1-48(in Chinese).

[73] NGUYEN T H, TSAFNAT N, CETIN E, et al. Low-Earth orbit satellite constellation for ADS-B based in-flight aircraft tracking[J]. Advances in Aircraft and Spacecraft Science, 2015, 2(1): 95-108.

[74] NAG S, RIOS J L, GERHARDT D, et al. CubeSat constellation design for air traffic monitoring[J]. Acta Astronautica, 2016, 128: 180-193.

[75] CHEN C C, LIU Z Q, FAN W, et al. Design and application analysis of the global coverage satellite system for space aeronautics ATM information collection[C]// Space Information Networks, 2017, 688: 266-273.

[76] LEYVA-MAYORGA I, SORET B, POPOVSKI P. Inter-plane inter-satellite connectivity in dense LEO constellations[J]. IEEE Transactions on Wireless Communications, 2021, 20(6): 3430-3443.

[77] GUO J M, YANG L, CHEN Q, et al. Design of a low earth orbit satellite constellation network for air traffic surveillance[J]. Journal of Navigation, 2020, 73(6): 1263-1283.

[78] DE WECK O L, DE NEUFVILLE R, CHAIZE M. Staged deployment of communications satellite constellations in low earth orbit[J]. Journal of Aerospace Computing, Information, and Communication, 2004, 1(3): 119-136.

[79] LEE H W, JAKOB P C, HO K, et al. Optimization of satellite constellation deployment strategy considering uncertain areas of interest[J]. Acta Astronautica, 2018, 153: 213-228.

[80] MOHORČIČ M, WERNER M, ŠVIGELJ A, et al. Alternate link routing for traffic engineering in packet-oriented ISL networks[J]. International Journal of Satellite Communications, 2001, 19(5): 463-480.

[81] BAI J J, LU X C, LU Z X, et al. Compact explicit multi-path routing for LEO satellite networks[C]// 2005 Workshop on High Performance Switching and Routing. Piscataway: IEEE Press, 2005: 386-390.

[82] HUSSEIN M, ABU-ISSA A, ELAYYAN I. Location-aware load balancing routing protocol for LEO satellite networks[C]// 2018 International Conference on Advanced Communication Technologies and Networking (CommNet). Piscataway: IEEE Press, 2018: 1-7.

[83] TALEB T, MASHIMO D, JAMALIPOUR A, et al. SAT04-3: ELB: An explicit load balancing routing protocol for multi-hop NGEO satellite constellations[C]// IEEE Globecom 2006. Piscataway: IEEE Press, 2006: 1-5.

[84] DONG C Y, XU X, LIU A J, et al. Research on new methods of LEO satellite networks routing[C]// 2019 6th International Conference on Information Science and Control Engineering (ICISCE). Piscataway: IEEE Press, 2019: 495-499.

[85] 叶红军, 刘亮, 贾诗雨. 星空地一体化航空安全监控体制设计[J]. 无线电工程, 2019, 49(9): 801-806.
YE H J, LIU L, JIA S Y. Design on satellite-air-ground integrated aviation safety monitoring system[J]. Radio Engineering, 2019, 49(9): 801-806(in Chinese).

[86] 郑晓冬, 顾青涛, 鲍亚川, 等. 低轨航空安全监视星座路由规划算法设计与仿真[J]. 无线电通信技术, 2019, 45(3): 253-257.
ZHENG X D, GU Q T, BAO Y C, et al. Design and simulation of low-orbit aviation surveillance constellation route algorithm[J]. Radio Communications Technology, 2019, 45(3): 253-257(in Chinese).

[87] SCHÄFER M, STROHMEIER M, LENDERS V, et al. Bringing up OpenSky: A large-scale ADS-B sensor network for research[C]// Proceedings of the 13th International Symposium on Information Processing in Sensor Networks. Piscataway: IEEE Press, 2014: 83-94.

[88] VAN DER PRYT R, VINCENT R. A simulation of signal collisions over the North Atlantic for a spaceborne ADS-B receiver using aloha protocol[J]. Positioning, 2015, 6(3): 23-31.

[89] GARCIA M A, STAFFORD J, MINNIX J, et al. Aireon space based ADS-B performance model[C]// 2015 Integrated Communication, Navigation and Surveillance Conference (ICNS). Piscataway: IEEE Press, 2015: C1-C2.

[90] 钟建华, 刘卫东, 王冬冬, 等. ADS-B监视数据质量分析研究[J]. 西安航空学院学报, 2013, 31(3): 72-75.
ZHONG J H, LIU W D, WANG D D, et al. Analysis on ADS-B

surveillance data quality[J]. Journal of Xi'an Aeronautical University,2013,31(3):72-75(in Chinese).

[91] 于克非. 星基 ADS-B 系统监视性能可用性评估[D]. 天津:中国民航大学,2018:1-54.
YU K F. Availability evaluation of surveillance performance on space-based ADS-B[D]. Tianjin:Civil Aviation University of China,2018:1-54(in Chinese).

[92] 赵嶷飞,于克非. 星基广播式自动相关监视系统监视数据空中位置信息质量分析[J]. 科学技术与工程,2018,18(14):279-284.
ZHAO Y F,YU K F. Quality analysis of air position on space-based automatic dependent surveillance broadcast surveillance data[J]. Science Technology and Engineering,2018,18(14):279-284(in Chinese).

[93] 王运帷. 陆基与星基 ADS-B 系统数据质量研究[D]. 天津:中国民航大学,2018:1-92.
WANG Y W. Research of land-based and space-based ADS-B system data quaility[D]. Tianjin:Civil Aviation University of China,2018:1-92(in Chinese).

[94] 张学军,李雪缘,王子润. 一种适用于高灵敏度星载 ADS-B 接收机的信号处理方法:CN111884981B[P]. 2021-05-18.
ZHANG X J,LI X Y,WANG Z R. Signal processing method suitable for high-sensitivity satellite-borne ADS-B receiver:CN111884981B[P]. 2021-05-18(in Chinese).

[95] FENG T,LIANG J. Parameter estimation of weak space-based ADS-B signals using genetic algorithm[J]. ETRI Journal,2021,43(2):324-331.

[96] 冯涛,梁俊. 基于反演策略的星基 ADS-B 信号译码方法[J]. 电子科技大学学报,2020,49(1):64-70.
FENG T,LIANG J. Space-based ADS-B signal decoding by inversion method[J]. Journal of University of Electronic Science and Technology of China,2020,49(1):64-70(in Chinese).

[97] 张学军,简鑫慧,黄如. 适用于星基 ADS-B 的低置信度矩阵可靠提取方法和检错纠错高性能方法:CN111865332A[P]. 2021-07-27.
ZHANG X J,JIAN X H,HUANG R. Reliable low confidence matrix extraction method and high performance method of error detection and correction for space-based ADS-B:CN111865332A[P]. 2021-07-27(in Chinese).

[98] 张学军,王子润,简鑫慧. 基于 FPGA 的星载 ADS-B 接收机的信号处理模块:CN112039578B[P]. 2021-05-04.
ZHANG X J,WANG Z R,JIAN X H. Signal processing module of satellite ADS-B receiver based on FPGA:CN112039578B[P]. 2021-05-04(in Chinese).

[99] LIU K,ZHANG T,DING Y. Blind signal separation algorithm for space-based ADS-B[C]// Proceedings of the 2016 International Conference on Mechatronics Engineering and Information Technology. Paris:Atlantis Press,2016:214-220.

[100] 张涛,丁洋,刘雪飞. 信号碰撞分离方法和装置:CN107276937B[P]. 2020-02-18.
ZHANG T,DING Y,LIU X F. Signal collision separation method and apparatus:CN107276937B[P]. 2020-02-18(in Chinese).

[101] 李家蓬,刘志刚,安强,等. 一种基于单通道 ADS-B 地面站的防欺骗综合解决方法:CN111142126B[P]. 2022-03-01.
LI J P,LIU Z G,AN Q,et al. Anti-spoofing comprehensive solution based on single-channel ADS-B ground station:CN111142126B[P]. 2022-03-01(in Chinese).

[102] 李家蓬,安强,付磊,等. 一种基于四通道 ADS-B 地面站的防欺骗解决方法:CN110988865B[P]. 2021-08-10.
LI J P,AN Q,FU L,et al. Anti-spoofing solution based on four-channel ADS-B ground station:CN110988865B[P]. 2021-08-10(in Chinese).

[103] 安强,李家蓬,黄枭,等. 基于信号 TDOA 计算的星基 ADS-B 目标验证[C]// 中国航空学会航空电子与空中交通管理学术会议(CCATM2021)暨航电与空管分会 2021 年学术年会,2021.
AN Q,LI J P,HUANG X,et al. Satellite-based ADS-B target verification based on signal TDOA calculation[C]// Proceedings of the China Aviation Society Avionics and Air Traffic Management Conference (CCATM2021) and the 2021 Annual Academic Conference of the Avionics and ATC Branch,2021(in Chinese).

[104] 张涛,张丽鑫,范伟强. 基于星基 ADS-B 报文卫星网络的局部多径路由方法和装置:CN107241268A[P]. 2020-05-12.
ZHANG T,ZHANG L X,FAN W Q. Method and apparatus for local multipath routing based on space-based ADS-B messaging satellite network:CN107241268A[P]. 2020-05-12(in Chinese).

[105] 张涛,曹思源,龚思龙. 一种低轨道卫星网络拥塞控制方法及装置:CN110958640B[P]. 2021-07-20.
ZHANG T,CAO S Y,GONG S L. Low-orbit satellite network congestion control method and device:CN110958640B[P]. 2021-07-20(in Chinese).

[106] 范伟强. 空事卫星星上路由技术[D]. 北京:北京航空航天大学,2019:1-57.
FAN W Q. Research on low orbit satellite routing technology[D]. Beijing:Beihang University,2019:1-57(in Chinese).

[107] WANG Y C,ZHANG X J,ZHANG T. A flooding-based routing algorithm for ADS-B packets transmission in LEO satellite network[C]// 2019 Integrated Communications,Navigation and Surveillance Conference (ICNS). Piscataway:IEEE Press,2019:1-9.

[108] 刘海涛,王松林,秦定本,等. 星基 ADS-B 接收机监视容量分析[J]. 航空学报,2018,39(5):182-189.
LIU H T,WANG S L,QIN D B,et al. Performance analysis of surveillance capacity of satellite-based ADS-B receiver[J]. Acta Aeronautica et Astronautica Sinica,2018,39(5):182-189(in Chinese).

[109] 李少洋. 星基 ADS-B 监视性能仿真研究[D]. 天津:中国民航大学,2020:1-61.
LI S Y. Simulation study of satellite-based ADS-B surveillance performance[D]. Tianjin:Civil Aviation University of China,2020:1-61(in Chinese).

[110] 张学军,李雪缘. 所需监视性能可用性评估方法:CN111754817B[P]. 2021-06-01.
ZHANG X J,LI X Y. Required monitoring performance availability evaluation method:CN111754817B[P]. 2021-06-01(in Chinese).

[111] 张学军,简鑫慧. 监视性能评估指标形成方法:CN111785095B[P]. 2021-06-01.
ZHANG X J, JIAN X H. Monitoring performance index evaluation method: CN111785095B[P]. 2021-06-01 (in Chinese).

[112] 李雪缘,张学军,杨宁. 北航空事卫星一号监视应用分析[C]//中国航空学会航空电子与空中交通管理学术会议(CCATM2021)暨航电与空管分会2021年学术年会,2021.
LI X Y, ZHANG X J, YANG N. Analysis of Beihang aviation satellite-1 surveillance applications[C]// Proceedings of the China Aviation Society Avionics and Air Traffic Management Conference (CCATM2021) and the 2021 Annual Academic Conference of the Avionics and ATC Branch, 2021 (in Chinese).

A review of development of space-based ADS-B system and its key technologies

ZHANG Xuejun*, TAN Yuanhao, LI Xueyuan, JIAN Xinhui

(School of Electronic Information Engineering, Beihang University, Beijing 100083, China)

Abstract: In the context of the globalization of air traffic control, the ICAO has proposed a new generation of surveillance technology called the automatic dependent broadcast surveillance (ADS-B) system. Compared with land-based systems, the space-based ADS-B system can achieve global airspace coverage, which can enhance the capabilities of the existing air traffic control system and promote the opening of national low-altitude airspace as well as the development of the general aviation industry. Firstly, the origin, concept, and operation principle of space-based ADS-B are introduced. Secondly, a brief overview and history of the domestic and international space-based ADS-B systems are provided. Then, the research status on the space-based ADS-B key technologies are sorted out by both domestic and international scholars, including weak signal demodulation, multi-beam reception, de-interleaving, anti-spoofing, and constellation design, routing algorithm, and surveillance performance evaluation. The research work and Beihang space-based ADS-B technology verification satellite of the Beihang University, is introduced. Finally, the development trend and outlooks of the space-based ADS-B system are summarized by combining the future development of space-based air traffic control technology and application needs.

Keywords: space-based automatic dependent broadcast surveillance (ADS-B); weak signal demodulation; multi-beam reception; de-interleaving; anti-spoofing; constellation design; routing algorithm; surveillance performance evaluation

Received: 2022-05-01; **Accepted**: 2022-06-09; **Published online**: 2022-07-04 16:03
URL: kns.cnki.net/kcms/detail/11.2625.V.20220701.1950.002.html
* **Corresponding author**. E-mail: zhxj@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0321

基于“时间-特征”协同注意力的机场快轨客流预测

杜文博[1], 石婉君[1], 廖盛时[1], 朱熙[2,*]

(1. 北京航空航天大学 电子信息工程学院, 北京 100083; 2. 北京航空航天大学 前沿科学技术创新研究院, 北京 100083)

摘 要: 机场快轨客流的准确预测是实现机场轨道交通系统智能化、精细化、高效化管控的基础,对提升机场服务水平和运行效率有着重要意义。由于影响因素众多、相互耦合,且因素对客流时序影响机理复杂,机场快轨客流的准确预测极具挑战。提出了一种基于“时间-特征”协同注意力机制的机场快轨客流预测模型,实现了精准捕捉多维因素在不同时序上对机场快轨客流的影响。基于北京首都国际机场快轨实际客流数据进行实验,结果表明了所提方法的有效性。

关 键 词: 机场快轨; 时间序列; 客流预测; “时间-特征”协同注意力; 长短时记忆网络

中图分类号: U121

文献标志码: A **文章编号**: 1001-5965(2022)09-1605-08

相比于城市道路交通,机场快轨以准点率高、速度快、运能大、舒适性好等优势,成为衔接机场与市区的重要交通方式。然而,目前中国各大机场的快轨系统时常出现繁忙时段运力不足、空闲时段运力浪费的“客流-运力”不匹配现象,成为进一步提升民航旅客出行效率和机场运转效能的瓶颈之一。因此,有必要依据机场客流动态变化情况对快轨运力进行高效调配。为达到该目的,首要一步便是设计机场快轨客流的准确预测模型。

机场快轨客流受众多非线性因素影响,如进离港航班旅客量、天气、时段、是否为工作日等,这些因素对快轨客流的影响机理及相互之间的影响方式和大小各不相同;同时,这些影响在时间维度上也有着复杂的先后关系,难以定性和定量描述,如从机场到市区方向快轨的客流形成过程中,恶劣天气会使到港航班靠桥时间延后,加上到港旅客从下机到抵达机场快轨站需要一定时间,造成了机场快轨客流变化在时序上相对于天气与到港航班旅客数量等因素变化的复杂耦合性。可见,机场快轨客流的形成是一个复杂多变的过程,准确预测机场快轨客流极具挑战。

目前专门针对机场快轨客流预测的研究较少,而对于地铁、铁路客流预测问题,国内外已有不少研究,方法大体可分为 2 类:①基于数学建模的方法;②基于机器学习的方法。第 1 类方法认为客流在历史上的统计规律也适用于未来,基于对历史客流数据的挖掘分析构建描述客流变化规律的统计学模型,从而实现对未来客流的预测。具体来说,Li[1]使用交通四阶段预测模型,即基于交通产生量、吸引量、交通分布和交通方式划分因素建立数学模型,对哈尔滨地铁客流数据进行了预测分析;Ding 等[2]在自回归移动平均(autoregressive integrated moving average, ARIMA)模型中增加自回归条件异方差模块,从地铁历史客流数据中捕捉轨道客流在一天中不同时段的平均值和

收稿日期: 2022-05-06; **录用日期**: 2022-06-02; **网络出版时间**: 2022-06-22 10:26
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220621.1152.003.html
基金项目: 国家重点研发计划 (2019YFF0301400)
* **通信作者**. E-mail: zhuxi@buaa.edu.cn

引用格式: *杜文博, 石婉君, 廖盛时, 等. 基于“时间-特征”协同注意力的机场快轨客流预测[J]. 北京航空航天大学学报, 2022, 48(9): 1605-1612. DU W B, SHI W J, LIAO S S, et al. Passenger flow forecasting of airport express based on time and feature cooperative attention[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1605-1612(in Chinese).*

方差;Zhang 和 Liang[3]综合考虑了客流数据的日相似性,计算 2 周中同周次的客流量比值得到归一化客流数据,再基于该数据建立改进后的卡尔曼滤波预测模型,实现了对北京地铁站的客流预测。总体上,基于数学建模的方法通常计算复杂度低、易于实现,但就本文所研究的机场快轨在未来较短一段时间内客流的预测问题来说,该类方法难以准确建模与短时客流密切相关的气象、航班载客数等高动态因素对客流的影响,因此对机场快轨短时客流预测的精度较低。

第 2 类方法是基于机器学习的轨道客流预测方法,通过使用历史轨道客流样本训练机器学习模型,从样本中挖掘客流受多种因素影响而动态变化的知识,从而实现对未来客流量的预测。具体来说,Roos 等[4]从所研究站点客流与其自身历史客流及临近站点客流具有因果关系这一事实出发,采用贝叶斯网络对客流进行预测,且提出了一种缺失数据拟合策略,在客流数据不完整时采用最大期望算法学习网络结构和参数,提升了该方法在数据缺失条件下的鲁棒性;惠阳等[5]通过反向传播神经网络(back propagation neural network, BPNN)构建节假日、是否为工作日、时段、天气等多种因素与地铁客流间的映射,实现了对西安地铁各站客流的准确预测;Sun 等[6]针对地铁客流数据变化兼备规律性及随机性,将客流变化曲线视作信号,提出了先使用小波分析对客流变化曲线的高频和低频成分进行拆解,再使用多个支持向量机对各成分分别学习预测,最后综合预测结果,避免了客流曲线不同频段成分在预测过程中的相互影响。上述轨道客流预测方法所采用的机器学习模型较为经典,较基于数学建模的方法对多维因素影响客流的非线性机理挖掘得更深,但在捕捉高动态因素与快轨客流间的复杂时序关联上仍较为欠缺,导致预测性能并未达到最佳。为此,一些学者使用时序深度学习网络来解决该问题。Yang 等[7]使用长短时记忆(long short-term memory, LSTM)网络对地铁客流进行预测,且考虑到每天相同时段的客流变化趋势较为类似,因此,借助时间注意力机制重点关注所预测时段在历史数据上的客流变化情况,提升了预测准确性;Ma 等[8]将客流数据分布转化为静态图像,采用卷积神经网络(convolutional neural network, CNN)提取其中的空间特征,进而用 Bi-LSTM 网络对客流时间序列分别从前向和后向学习时间特征,实现了地铁客流数据空间特征和周期性的有效提取;Liu 等[9]使用 3 个 LSTM 网络分别提取地铁客流的时、日、周特性,结合天气、是否为工作日、降水、地铁运营时刻、站点间的交通时长等影响客流的因素对地铁换乘站及常规站点客流进行了准确预测。

综上所述,基于时序深度学习的轨道客流预测方法因对深层次时空特征挖掘的能力强已成为轨道客流预测的主流方法。但已有相关工作大多针对市区内轨道交通,而本文所研究的机场快轨客流形成机理较之更为复杂。机场快轨客流由乘机、下机、接机、送机等多类人员组成,受气象、航班延误、旅客通关等高动态因素影响较大,同时这些因素变化与机场快轨客流形成之间存在显著但难以量化的时序关联,这些特点导致传统时序深度学习网络难以高效学习机场快轨客流的变化规律。

为此,本文提出了一种基于“时间-特征”协同注意力机制的机场快轨客流预测方法。该方法在 LSTM 时序预测模型的基础上嵌入了“时间-特征”注意力机制,用于精细化捕捉各维特征在不同时序上对机场快轨客流的影响规律。将所提方法的性能与支持向量机(support vector machine, SVM)、BPNN、ARIMA、标准 LSTM 等现有轨道客流预测代表性方法进行了对比,验证了本文方法的优越性。

1 预测方法

1.1 长短时记忆网络

如图 1 所示,LSTM 网络[10]由循环训练的 LSTM 单元构成,展开后呈现链状,链中每个 LSTM 单元对应着一个时间片。LSTM 单元内部具体结构如图 2 所示。

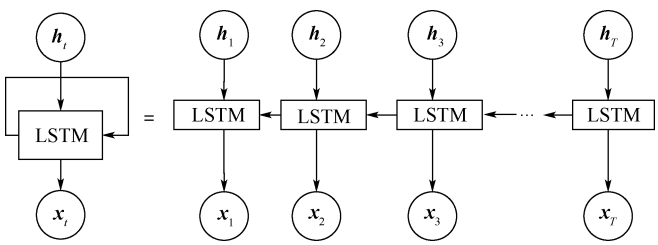


图 1 LSTM 网络

Fig. 1 LSTM network

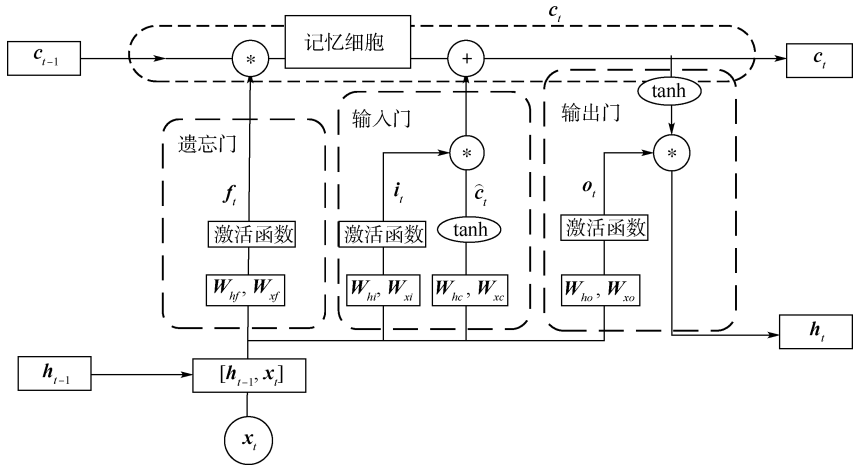


图 2　LSTM 单元结构

Fig. 2　LSTM unit structure

LSTM 单元包含输入门、遗忘门、输出门和记忆细胞。输入门用来确定当前输入有多少新的信息可以加入到记忆细胞中;遗忘门用来确定当前细胞需要遗忘的信息;输入门和遗忘门共同更新记忆细胞在每个时间片的状态;输出门用来控制记忆细胞要输出的信息。LSTM 单元涉及的计算公式如下:

$$\boldsymbol{f}_t = \sigma(\boldsymbol{W}_{hf}\boldsymbol{h}_{t-1} + \boldsymbol{W}_{xf}\boldsymbol{x}_t + \boldsymbol{b}_f) \tag{1}$$

$$\boldsymbol{i}_t = \sigma(\boldsymbol{W}_{hi}\boldsymbol{h}_{t-1} + \boldsymbol{W}_{xi}\boldsymbol{x}_t + \boldsymbol{b}_i) \tag{2}$$

$$\boldsymbol{o}_t = \sigma(\boldsymbol{W}_{ho}\boldsymbol{h}_{t-1} + \boldsymbol{W}_{xo}\boldsymbol{x}_t + \boldsymbol{b}_o) \tag{3}$$

$$\hat{\boldsymbol{c}}_t = \tanh(\boldsymbol{W}_{hc}\boldsymbol{h}_{t-1} + \boldsymbol{W}_{xc}\boldsymbol{x}_t + \boldsymbol{b}_c) \tag{4}$$

$$\boldsymbol{c}_t = \boldsymbol{f}_t \odot \boldsymbol{c}_{t-1} + \boldsymbol{i}_t \odot \hat{\boldsymbol{c}}_t \tag{5}$$

$$\boldsymbol{h}_t = \boldsymbol{o}_t \odot \tanh(\boldsymbol{c}_t) \tag{6}$$

式中:$\boldsymbol{f}_t$、$\boldsymbol{i}_t$、$\boldsymbol{o}_t \in \mathbf{R}^m$分别为遗忘门、输入门、输出门的计算结果;$\hat{\boldsymbol{c}}_t$、$\boldsymbol{c}_t$、$\boldsymbol{h}_t \in \mathbf{R}^m$分别为新生成的细胞状态、当前细胞状态、隐藏状态;$\boldsymbol{W}_{hf}, \boldsymbol{W}_{hi}, \boldsymbol{W}_{ho}, \boldsymbol{W}_{hc} \in \mathbf{R}^{m\times m}$和 $\boldsymbol{W}_{xf}, \boldsymbol{W}_{xi}, \boldsymbol{W}_{xo}, \boldsymbol{W}_{xc} \in \mathbf{R}^{m\times n}$为遗忘门、输入门、输出门和单元状态分别涉及的 2 个权重矩阵;$\boldsymbol{b}_f, \boldsymbol{b}_i, \boldsymbol{b}_o, \boldsymbol{b}_c \in \mathbf{R}^m$ 分别为遗忘门、输入门、输出门和单元状态的偏置向量;$\boldsymbol{x}_t$ 为 t 时刻的输入时间序列值;$\sigma(\cdot)$为 Sigmoid 激活函数;$\tanh(\cdot)$为双曲正切函数激活函数;$\odot$为哈达玛积。

1.2　“时间-特征”协同注意力

影响机场快轨客流的各个因素对快轨客流预测的贡献程度、作用方式存在差异,为获得准确的预测效果,需要对不同特征受关注程度,以及同一特征在不同时间片上的受关注程度进行学习和优化。为此,本文提出了一种“时间-特征”协同注意力机制,使得预测模型能够对不同时间片上不同特征的关注焦点进行动态调配。

近年来,注意力机制被广泛应用于基于深度学习的各种分类、预测、聚类等任务中[11-12],促使模型在学习样本过程中重点关注相对重要的特征或历史时序(对于时序样本),以获得更好的样本泛化能力。自注意力机制在捕获全局长时序依赖关系上较传统注意力机制优势明显,然而自注意力机制对不同时间片使用的是同一组权重矩阵,不同特征之间权重不同,因此自注意力机制只能得到不同特征之间的相对注意力大小,无法得到一组时间序列中同一特征在不同时间片的注意力分布。因此,本文通过对每一个时间片设置独立的自注意力模块,构建“时间-特征”协同注意力机制,实现为不同时间片上的多维特征进行独立的权重分配。

本文所提出的基于“时间-特征”协同注意力(time and feature cooperative attention, TFATT)的快轨客流预测模型如图 3 所示。首先,对客流及相关影响因素的原始数据进行清洗,包括对数据缺失部分进行推测补全,以及对异常数据进行剔除;然后,为了消除不同客流影响因素(或称输入特征)间的量纲差异,对数据进行 0-1 归一化处理[13]。

归一化后的输入特征时序数据 $\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_i, \cdots, \boldsymbol{x}_T) \in \mathbf{R}^{n\times T}$,其中 n 为输入特征的数量,T 为时间序列长度,隐藏层输出 $\boldsymbol{H} = (\boldsymbol{h}_1, \boldsymbol{h}_2, \cdots, \boldsymbol{h}_T) \in \mathbf{R}^{m\times T}$,其中 m 为 LSTM 单元数量;再将 $\boldsymbol{H}$ 经过全连接层得到新的时间序列 $\boldsymbol{A} = \boldsymbol{W}_F \cdot \boldsymbol{H} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_i, \cdots, \boldsymbol{a}_T) \in \mathbf{R}^{n\times T}$,其中 $\boldsymbol{W}_F \in \mathbf{R}^{n\times m}$ 为全连接矩阵。

为构建“时间-特征”协同注意力机制,本文方法针对每一个时间片设置了独立的特征注意力模块,即对于第 i 个时间片分别计算查询(Query)矩阵 $\boldsymbol{Q}_i$、关键字(Key)矩阵 $\boldsymbol{K}_i$、值(Value)矩阵 $\boldsymbol{V}_i$,$1 \leqslant i \leqslant T$。

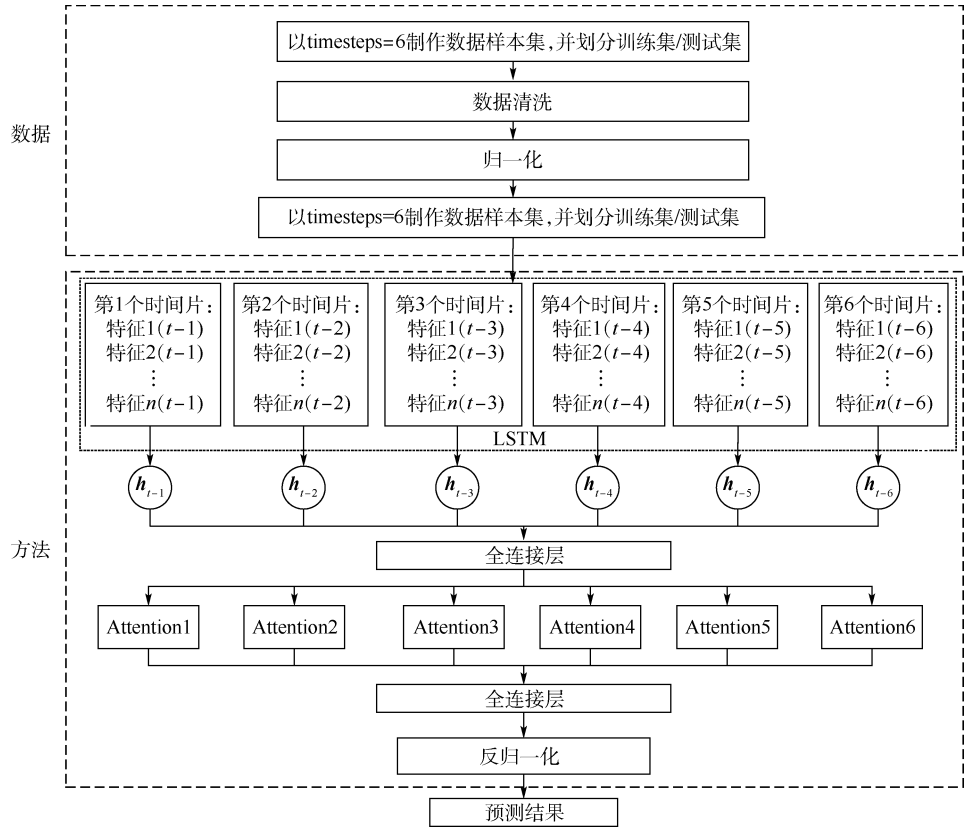


图 3 机场快轨客流预测模型

Fig. 3 Passenger flow prediction model of airport express rail

$$\boldsymbol{Q}_i = \boldsymbol{A}\boldsymbol{W}^{\boldsymbol{Q}_i} \tag{7}$$

$$\boldsymbol{K}_i = \boldsymbol{A}\boldsymbol{W}^{\boldsymbol{K}_i} \tag{8}$$

$$\boldsymbol{V}_i = \boldsymbol{A}\boldsymbol{W}^{\boldsymbol{V}_i} \tag{9}$$

式中:$\boldsymbol{W}^{\boldsymbol{Q}_i}, \boldsymbol{W}^{\boldsymbol{K}_i}, \boldsymbol{W}^{\boldsymbol{V}_i} \in \mathbf{R}^{n\times n}$为权重矩阵,通过样本训练得到。

进而通过式(10)计算每一个时间片加入注意力之后的值:

$$\text{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}) = \text{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{e}}\right) \tag{10}$$

式中:e 为 $\boldsymbol{K}$ 的维度大小。

对于每一个计算结果 $\text{Attention}_i(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}) \in \mathbf{R}^{T\times n}$,取其中第 i 个向量,记为$\mathbf{att}_i, \mathbf{att}_i \in \mathbf{R}^n$。将所有注意力模块计算出的向量组合得到新的时间序列:$\mathbf{ATT} = (\mathbf{att}_1, \mathbf{att}_2, \cdots, \mathbf{att}_i, \cdots, \mathbf{att}_T) \in \mathbf{R}^{n\times T}$,再经过一个全连接层得到 $\hat{y}_{\text{st}} = \boldsymbol{w}_{F2} \cdot \mathbf{ATT} \cdot \boldsymbol{w}_{F3}$($\boldsymbol{w}_{F2} \in \mathbf{R}^{1\times n}, \boldsymbol{w}_{F3} \in \mathbf{R}^{T\times 1}$ 为全连接向量),最后对 $\hat{y}_{\text{st}}$进行反归一化处理,得到机场快轨客流预测值 $\hat{y}$。

2 实验验证

2.1 实验数据

本文选取 2019 年 1 月 5 日至 7 日、5 月 9 日至 11 日、8 月 4 日至 15 日、9 月 27 日至 30 日、12 月 14 日至 17 日共 26 天的北京首都国际机场快轨站的进出站刷卡数据、机场天气数据(含晴、雾、雨3 类),以及航班起降计划数据和实际起降数据。目标预测时段为 6:30—23:00,选取 10 min 作为一个时间片长度(共计 100 个时间片/天 × 26 天 =2 600个时间片),统计每个时间片对应的市区到机场方向客流、机场到市区方向客流、天气、机场实际/计划到离港航班载客数、时段信息(以 1 h 为单元,如 19 ~20 时段)、工作日/非工作日信息。

2.2 模型参数

本文将所提 TFATT 模型与近年来最常见的轨道客流预测模型 SVR[6]、BPNN[5]、ARIMA[2]和标准 LSTM[7]模型进行对比,通过咨询行业领域专家与调研文献,得到影响机场快轨客流的主要因素包括快轨客流历史数据、进离港航班旅客量、天气、时段、是否工作日等。为了明确机场到市区、市区到机场 2 个方向上机场快轨客流的关键影响因素,本文基于 2 600 条样本,使用 Pearson 相关系数、SNR 信噪比对各个客流影响因素与 2 个方向客流的关联程度作初步检验,并挑选其中关联度较大的特征作为模型输入(ARIMA 模型除外,根据文献[1]可知其输入特征仅为客流)。具体为:对于机场到市区方向客流,特征包括机场

到市区方向历史客流、机场实际到港航班载客数(机场实际到港航班载客数与快轨客流间 Pearson 系数为 -0.106)、实际与计划离港航班载客数(实际离港航班载客数与快轨客流间 Pearson 系数为 0.302,计划离港航班载客数与快轨客流间 Pearson 系数为 0.244)、天气(晴、雾、雨天气与快轨客流间信噪比为 0.114)、时段(时段与快轨客流间信噪比为 1.426)、是否工作日(是否工作日与快轨客流间信噪比为 0.108);对于市区到机场方向客流,特征包括市区到机场方向历史客流、机场计划到港航班载客数(机场计划到港航班载客数与快轨客流 Pearson 系数为 -0.078)、机场实际与计划离港航班载客数(机场实际离港航班数与快轨客流 Pearson 系数为 0.136,机场计划离港航班数与快轨客流 Pearson 系数为 0.156)、天气(晴、雾、雨天气与快轨客流间信噪比为 0.265)、时段(时段与快轨客流间信噪比为 0.671)、是否工作日(是否工作日与快轨客流间信噪比为 0.311)。各模型采用目标预测时间片的前 6 个时间片的特征向量,对目标预测时间片客流进行预测。

考虑到模型的参数对模型性能有很大的影响,目前还没有成熟的模型参数选取理论方法,因此,本文进行了参数遍历实验来挑选参数。参数设置如下:①SVM。核函数选择径向基函数(radio basic function,RBF),核系数 gamma 设为 auto,误差项的惩罚参数 C 设为 10。②BPNN。BPNN 隐含层数设为 2,每层大小设为 36。③ARIMA。自回归系数 p 设为 4,滑动平均系数设为 1,差分阶数 d 设为 0。④LSTM。batch_size = 128,timesteps = 6,epochs = 300,LSTM 单元数设为 70,LSTM 隐藏层数设为 2,损失函数设为 MAE,激活函数设为 Sigmoid,优化函数设为 Adam。⑤TFATT。batch_size = 128,epochs = 300,隐藏层数设为 2,损失函数设为 MAE,激活函数为 Sigmoid,优化函数设为 Adam。

本文实验在 Windows10 系统中完成,CPU 为 i7-8550U@ 1.8 GHz,内存 8.0 GB,IDE 为 Pycharm2018.2,Python 版本为 3.6,实验框架为 Tensorflow2.6.2。

2.3　实验设计

将 26 天的数据中某一天的数据作为测试集(含 100 条样本),剩下 25 天作为训练集(含 2 500 条样本),设置 26 组实验,每组实验独立开展 10 次,共开展 260 次实验。

本文使用均方误差(mean square error,MSE)和平均绝对误差(mean absolute error, MAE)来定量评估 2 种模型的性能[14]。计算公式如下:

$$\mathrm{MSE} = \frac{1}{m}\sum_{i=1}^{m}(\hat{y} - y_i)^2 \tag{11}$$

$$\mathrm{MAE} = \frac{1}{m}\sum_{i=1}^{m}|\hat{y} - y_i| \tag{12}$$

式中:y_i 为实际值;$\hat{y}$ 为预测值;m 为样本总数。

2.4　实验结果对比分析

基于 2.3 节所述实验设计,表 1 为 260 次实验的平均结果。

由表 1 可知,BPNN 和 SVR 预测效果不如 ARIMA、LSTM 及本文所提 TFATT 模型,这是由于其虽然能够在一定程度上挖掘出多维特征对客流的非线性影响,但不能捕捉多维输入特征在前后时序上的关联;ARIMA 仅使用历史客流数据作为输入实现对未来客流的预测,但其只能挖掘内生变量的时序影响而未考量其他多因素的综合影响,因此预测效果不如本文所提 TFATT 模型;LSTM 可学习客流受到的多元特征时序影响,但面对机场快轨运行中复杂的多维因素与客流耦合机理,对于特征时序关联的泛化准确度仍不足;本文所提 TFATT 模型通过“时间-特征”协同注意力机制充分捕捉不同特征影响客流的时序滞后性,由表 1 可知,本文所提模型在 2 个方向上的客流预测效果均优于其他 4 个对比模型。

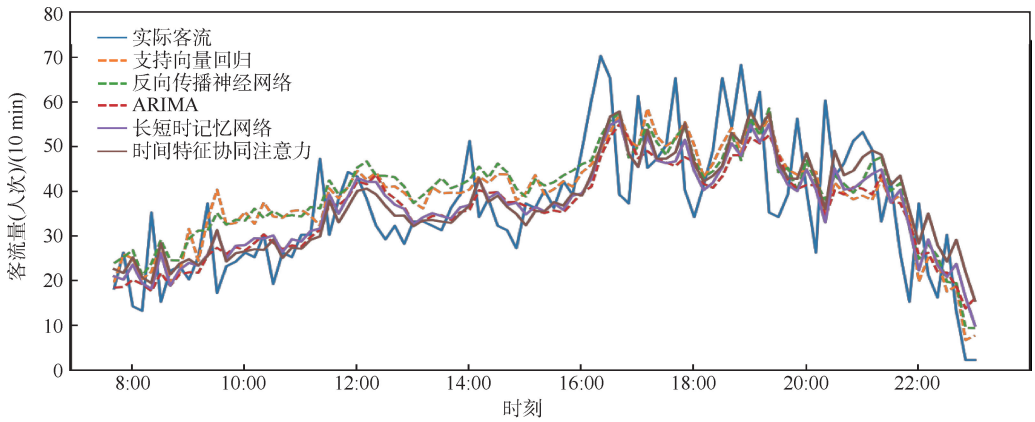
此外,由表 1 还可看出,各模型对市区到机场方向客流的预测精度普遍低于机场到市区方向(后文将对该现象原因作进一步分析),但本文 TFATT 模型相比于其他 4 个模型在 2 个出行方向上均有更好的预测效果,说明了“时间-特征”协同注意力机制对于提升模型预测性能的有效性。

为更直观地查看各模型的预测效果,图 4 展示了将前 25 天的样本作为训练集、最后 1 天样本作为测试集的预测曲线。

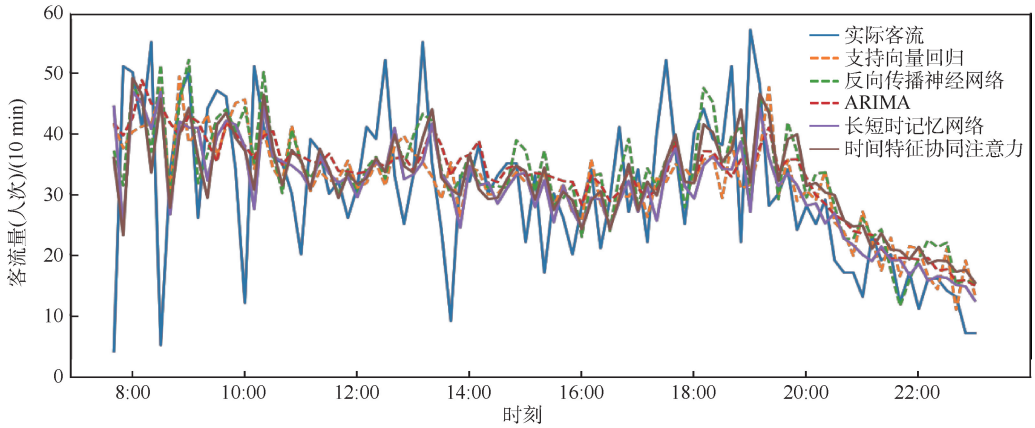
表 1　机场快轨客流预测性能

Table 1　Forecasting performance of airport express rail passenger flow

客流方向	预测模型	MSE	MAE
机场→市区	BPNN	112.245	8.81
	SVR	107.72	8.90
	ARIMA	101.89	7.89
	LSTM	105.35	8.11
	TFATT	96.94	7.67
市区→机场	BPNN	129.95	8.16
	SVR	128.06	8.22
	ARIMA	127.89	8.49
	LSTM	132.22	8.70
	TFATT	123.52	8.12



(a) 2019年12月17日机场到市区方向客流预测曲线



(b) 2019年12月17日市区到机场方向客流预测结果

图4 客流预测曲线

Fig. 4 Forecast results of passenger flow

由图4所示结果可知,5种模型均能够预测出机场快轨实际客流变化的总体趋势,但仔细对比可发现,本文TFATT模型较其他模型在对实际客流曲线的拟合度上更优。机场快轨实际客流曲线均出现了在短时间内波动较大的现象,特别是在市区到机场方向,这是由一些偶然因素所致。例如,图4(b)中市区到机场方向列车到达机场站时刻在各预测时间片上并非均匀分布,一旦有车到站则客流猛增,无车到站则客流稀少,造成实际客流曲线跳跃明显。由于偶然因素尚未被纳入到本文实验模型的输入特征中,造成各模型的预测客流曲线均不能严密拟合于实际客流曲线,但实验结果仍能说明在现有数据基础上,本文方法相比于其他方法具有优越性。

3 结 论

1) 本文面向机场快轨运力配置不能适应客流动态变化导致运力不足或浪费的问题,提出了一种基于"时间-特征"协同注意力机制的机场快轨客流预测模型TFATT,通过对不同因素在不同时序下对于机场快轨客流影响的显著性进行精细化建模,提升了机场快轨客流预测的精度,为机场运控部门开展机场快轨运力调控提供支撑。

2) 基于北京首都国际机场及其快轨实际运行数据开展快轨客流预测实验,将本文所提客流预测模型TFATT与常用的基于BPNN、SVR、ARIMA、LSTM模型的客流预测方法进行了实验对比,结果验证了TFATT模型的有效性。

目前仅考虑了机场路侧的一种交通方式机场快轨进行客流预测,而实际上路侧具有多种交通方式,彼此之间互有影响,后续可以开展机场快轨[15]、巴士[16]、出租车[17]等多种交通方式客流协同预测研究,以提升预测精度。

参考文献(References)

[1] LI Y. Passenger flow forecast of Harbin the Line 1 subway[C]// 2010 International Conference On Computer Design and Applications. Piscataway: IEEE Press, 2010: 11524216.

[2] DING C, DUAN J, ZHANG Y, et al. Using an ARIMA-GARCH modeling approach to improve subway short-term ridership fore-

casting accounting for dynamic volatility[J]. IEEE Transactions on Intelligent Transportation Systems,2018,19(4):1054-1064.

[3] ZHANG Z,LIANG T. Short-term forecasting of passenger flow on the metro platform using an improved Kalman filtering method[C]//19th COTA International Conference of Transportation Professionals,2019.

[4] ROOS J,BONNEVAY S,GAVIN G. Short-term urban rail passenger flow forecasting:A dynamic bayesian network approach [C]//2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA). Piscataway:IEEE Press,2016:16640539.

[5] 惠阳,王永岗,彭辉,等. 基于优化 PSO-BP 算法的耦合时空特征下地铁客流预测[J]. 交通运输工程学报,2021,21(4):210-222.
HUI Y,WANG Y G,PENG H,et al. Prediction of subway passenger flow based on optimized PSO-BP algorithm[J]. Journal of Traffic and Transportation Engineering,2021,21(4):210-222(in Chinese).

[6] SUN Y,LENG B,GUAN W. A novel wavelet-SVM short-time passenger flow prediction in Beijing subway system[J]. Neurocomputing,2015,166:109-121.

[7] YANG B,SUN S,LI J,et al. Traffic flow prediction using LSTM with feature enhancement[J]. Neurocomputing,2019,332:320-327.

[8] MA X,ZHANG J,DU B,et al. Parallel architecture of convolutional bi-directional LSTM neural networks for network-wide metro ridership prediction[J]. IEEE Transactions on Intelligent Transportation Systems,2019,20(6):2278-2288.

[9] LIU Y,LIU Z,JIA R. DeepPF:A deep learning based architecture for metro passenger flow prediction[J]. Transportation Research Part C:Emerging Technologies,2019,101:18-34.

[10] GERS F A,SCHMIDHUBER J,CUMMINS F. Learning to forget:Continual prediction with LSTM[J]. Neural Computation,2000,12(10):2451-2471.

[11] TIAN D,LIN C,ZHOU J,et al. SA-YOLOv3:An efficient and accurate object detector using self-attention mechanism for autonomous driving[J]. IEEE Transactions on Intelligent Transportation Systems,2022,23(5):4099-4110.

[12] SUN J,KIM J. Joint prediction of next location and travel time from urban vehicle trajectories using long short-term memory neural networks[J]. Transportation Research Part C:Emerging Technologies,2021,128:103114.

[13] MUNKHDALAI L,MUNKHDALAI T,PARK K H,et al. Mixture of activation functions with extended min-max normalization for forex market prediction[J]. IEEE Access,2019,7:183680-183691.

[14] BOGAERTS T,MASEGOSA A D,ANGARITA-ZAPATA J S,et al. A graph CNN-LSTM neural network for short and long-term traffic forecasting based on trajectory data[J]. Transportation Research Part C:Emerging Technologies,2020,112:62-77.

[15] ZHANG J,CHE H,CHEN F,et al. Short-term origin-destination demand prediction in urban rail transit systems:A channel-wise attentive split-convolutional neural network method[J]. Transportation Research Part C:Emerging Technologies,2020,124:102928.

[16] MA Q,LI S,ZHANG H,et al. Robust optimal predictive control for real-time bus regulation strategy with passenger demand uncertainties in urban rapid transit[J]. Transportation Research Part C:Emerging Technologies,2021,127:103086.

[17] DU B,PENG H,WANG S,et al. Deep irregular convolutional residual LSTM for urban traffic passenger flows prediction[J]. IEEE Transactions on Intelligent Transportation Systems,2020,21(3):972-985.

Passenger flow forecasting of airport express based on time and feature cooperative attention

DU Wenbo[1], SHI Wanjun[1], LIAO Shengshi[1], ZHU Xi[2,*]

(1. School of Electronic and Information Engineering, Beihang University, Beijing 100083, China;
2. Research Institute for Frontier Science, Beihang University, Beijing 100083, China)

Abstract: Accurate prediction of passenger flow of airport express rail is the basis for realizing intelligent, refined and efficient control of airport rail transit system. And it is of great significance to improve airport service levels and operational efficiency. The correct prediction of airport express rail passenger flow is particularly difficult due to the complex interplay of multiple factors that influence passenger flow in complicated ways. In this paper, an airport express passenger flow prediction model based on time and feature cooperative attention mechanism is proposed to accurately capture the influence degree of multi-dimensional factors on express passenger flow in different time series. Based on the actual passenger flow data of Beijing Capital International Airport, the experiment results show that the proposed method is effective.

Keywords: airport express rail; time series; passenger flow forecast; time and feature cooperative attention; long short-term memory

Received: 2022-05-06; **Accepted**: 2022-06-02; **Published online**: 2022-06-22 10:26

URL: kns.cnki.net/kcms/detail/11.2625.V.20220621.1152.003.html

Foundation item: National Key R & D Program of China (2019YFF0301400)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0297

飞机液压泵二维性能退化的可靠性评估

王少萍[1,2,*], 陈仁同[1], 张超[2,3]

(1. 北京航空航天大学 自动化科学与电气工程学院, 北京 100083;
2. 北京航空航天大学宁波创新研究院, 宁波 315800; 3. 北京航空航天大学 前沿科学技术创新研究院, 北京 100083)

摘 要: 飞机液压泵为飞机操纵、起落架收放和刹车提供高压能源,其安全性和可靠性至关重要。针对飞机液压泵二维退化可靠性评估问题,构建了基于液压泵微观表面混合润滑的转子-配流盘摩擦副和轴尾密封副的摩擦磨损退化模型,基于统一随机过程模型描述2个性能指标的退化规律,利用 Copula 函数表征二维性能退化的相关关系,使用贝叶斯马尔可夫链蒙特卡罗法得到其考虑二元相关退化模型的未知参数估计,进而对飞机液压泵进行了可靠性评估。通过液压泵回油流量及轴尾密封摩擦力矩2个性能退化指标的实际退化数据,验证了所提出的理论和方法。结果表明,考虑2个性能退化指标相关的退化模型有效提高了可靠性评估及寿命预测精度。

关 键 词: 二维性能退化; Copula 函数; 参数估计; 飞机液压泵; 可靠性评估
中图分类号: V215.7; TP206+.3
文献标志码: A **文章编号**: 1001-5965(2022)09-1613-11

飞机液压系统是为飞机的起飞、操纵、起落架收放和刹车等提供能源的关键系统,其结构复杂、使用环境恶劣,故障发生率高,维修成本高,可靠性、安全性和维修保障性要求极高[1]。液压泵是液压系统的心脏,其通过发动机机匣直接与发动机连接,因此承受高压、宽温和强振动等恶劣工况,极易出现摩擦副摩擦磨损、承力件疲劳和橡胶件老化等故障现象[2]。据统计,飞机液压泵故障在液压系统故障中占比最高,因此飞机液压系统故障诊断大多聚焦于飞机液压泵[3]。

飞机液压泵最常见的故障模式是其3对摩擦副(即转子-配流盘、柱塞-柱塞腔和滑靴-斜盘)的摩擦磨损故障和其轴尾密封故障[4]。20世纪80年代,英国 Bath 大学 Hunt 在泵出口安装了压力传感器监测液压泵摩擦副磨损故障,通过压力信号的对数倒频谱发现磨损征兆[5]。德国 Langen 通过在外部加装加速度传感器,进行液压泵的摩擦副摩擦磨损的振动故障诊断[6]。为了揭示液压泵内在的故障机理,美国普渡大学 Ivantysynova 和 Lasaar 深入研究了柱塞泵柱塞副,通过耦合求解柱塞副动力学方程、Reynolds 方程和能量方程等,分析了柱塞副间油膜厚度、压力、温度和摩擦力等润滑特性,为柱塞泵柱塞副故障这段提供了理论依据[7-9]。Kassem 和 Bahr 建立了液压泵的流量脉动的故障诊断模型,深入分析液压泵故障映射的压力变化,给出基于动态压力信号的液压泵故障诊断方法[10]。为了进一步研究液压泵摩擦副表面材料对性能的影响,Hong 等对比研究了表面镀膜材料对液压泵性能的影响,试验验证了镀膜可以使配油盘耐摩擦性能明显改善[11]。考虑到液压泵使用过程中关键摩擦副间存在流体润滑,因此其故障会引起混合润滑情况,

收稿日期: 2022-04-29; **录用日期**: 2022-05-29; **网络出版时间**: 2022-06-21 13:26
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220620.1546.001.html
基金项目: 国家自然科学基金国际合作重点项目(51620105010); 国家科技重大专项(J2019-V-0016-0111)
* **通信作者**. E-mail: shaopingwang@buaa.edu.cn

引用格式: *王少萍, 陈仁同, 张超. 飞机液压泵二维性能退化的可靠性评估[J]. 北京航空航天大学学报, 2022, 48(9): 1613-1623.*
WANG S P, CHEN R T, ZHANG C. Reliability estimation for aircraft hydraulic pump based on bivariate performance degradation analysis[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1613-1623(in Chinese).

Fang 和 Shirakashi 定义了接触系数描述摩擦副的润滑特性[12]，Manring 采用理论分析和试验测量的方法给出了柱塞-柱塞腔之间摩擦力的建模[13]。日本著名学者 Yamaguchi 对液压泵滑靴与斜盘摩擦副的混合润滑建立了基于 Greenwood-Williamson 和 Patir-Cheng 模型的综合混合润滑模型，并通过实验证明了模型的正确性[14-15]。

在国内，浙江大学诸葛起等最先开展了液压能源系统的故障诊断与建模分析，通过液压泵壳体的振动信号监测进行故障诊断[16]。浙江大学徐兵教授团队对轴向柱塞泵滑靴副和配流副等关键摩擦副的润滑特性，搭建试验台实现了对油膜厚度、滑靴自旋速度等关键参数的测试[17]。北京理工大学刘洪等考虑了滑靴轮廓磨损和弹性变形等因素对滑靴副动态润滑油膜特性的影响，分析了导致滑靴偏磨磨损的因素[18]。燕山大学刘思远等对轴向柱塞泵滑靴副剧烈磨损过程中滑靴润滑油膜特性进行了分析，并将轮廓磨损高度纳入了滑靴副油膜厚度修正量[19]。1994 年开始，北京航空航天大学王少萍团队持续开展飞机液压泵基于信号和基于模型的故障诊断研究[20-21]，通过在配流盘周向设置了压力传感器和油膜厚度传感器，构建了液压泵转子-配流盘混合润滑摩擦磨损模型，研究了转子-配流盘摩擦副的动态压力和油膜变化关系[22-23]。随后将以上微观和宏观建模思想推广应用到柱塞-柱塞腔、滑靴-斜盘及密封副，从而得到液压泵关键摩擦副和密封副的性能退化关系，为飞机液压泵基于性能退化的健康状态评估奠定了基础[24-26]。基于随机过程的建模方法由于具有良好的统计特性，被广泛应用于液压元件可靠性建模[27-29]。常用的随机过程模型有维纳过程模型、伽马过程模型以及逆高斯过程模型等。近年来，随着液压泵性能退化模型研究的不断深入，多故障竞争退化成为研究热点，如何准确表征液压泵多故障综合退化关系进而实现健康状态评估，成为飞机液压系统故障的关键问题。

为了准确描述飞机液压系统多故障引发的性能退化关系，本文采用多场耦合失效物理建模方法构建飞机液压泵转子-配流盘和轴尾密封性能退化模型，采用 Copula 函数[30]描述多维退化和可靠性，采用贝叶斯马尔可夫链蒙特卡罗法进行参数估计，并进行试验验证，实现飞机液压泵的准确退化刻画和可靠性评估。

1 转子-配流盘多场耦合退化建模

飞机液压泵广泛采用柱塞式液压泵，其主要摩擦副包括转子-配流盘、柱塞-转子副和滑靴-斜盘副。由于柱塞式液压泵通过柱塞在柱塞孔内往复运动实现吸油和排油，转子-配流盘间存在高低压非对称分布油液，从而在载荷作用下造成液压泵转子-配流盘摩擦磨损失效，如图 1 所示。

通常，液压泵转子-配流盘摩擦副表面并不是完全光滑的，存在粗糙峰分布，图 2 为使用电子显微轮廓仪得到的配流盘表面形貌。为了准确描述液压泵转子-配流盘摩擦磨损关系，本文用楔形油膜表征液压泵转子-配流盘高低压造成的油膜形貌(见图 3(a))。为了方便数学描述，本文假设液压泵缸体表面为光滑表面，仅考虑配流盘为粗糙表面(见图 3(b))。

对于转子-配流盘上任意一点(x,y)的微元面积为 $\mathrm{d}x\mathrm{d}y$，缸体与配流盘的真实接触面积只占名义接触面积很小的一部分，其受力关系为

$$p(x,y) = p_{\mathrm{l}}(x,y)(1 - a_{\mathrm{w}}) + p_{\mathrm{s}}(x,y) \tag{1}$$

式中：$p(x,y)$为 $\mathrm{d}x\mathrm{d}y$ 处微元所受压力；$p_{\mathrm{l}}(x,y)$为(x,y)处无粗糙峰的油膜压力；$p_{\mathrm{s}}(x,y)$为粗糙峰接触压力；a_{w} 为真实接触面积占名义接触面积的比例数，表示为

$$a_{\mathrm{w}} = \int_{h(x,y)}^{h_{\max}} g(z)\,\mathrm{d}z \tag{2}$$

式中：$h(x,y)$为小微元处的油膜厚度；$h_{\max}$为最大粗糙峰高度；$g(z)$为综合表面高度分布。

在转子-配流盘油膜区域，其油膜厚度、压力

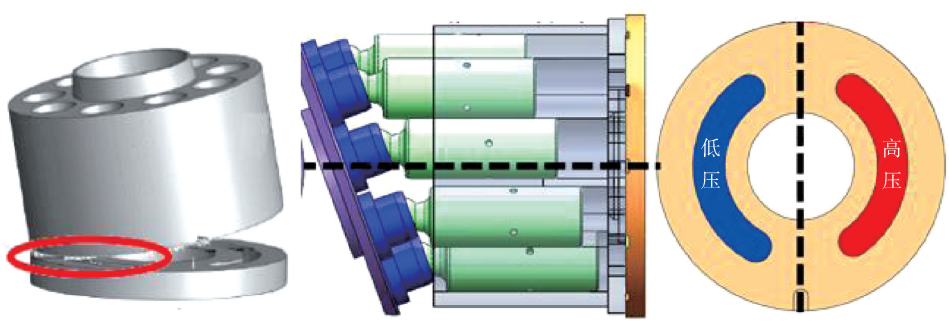


图 1 柱塞式液压泵的转子-配流盘摩擦副

Fig. 1 Valve plate-cylinder block friction of piston pump

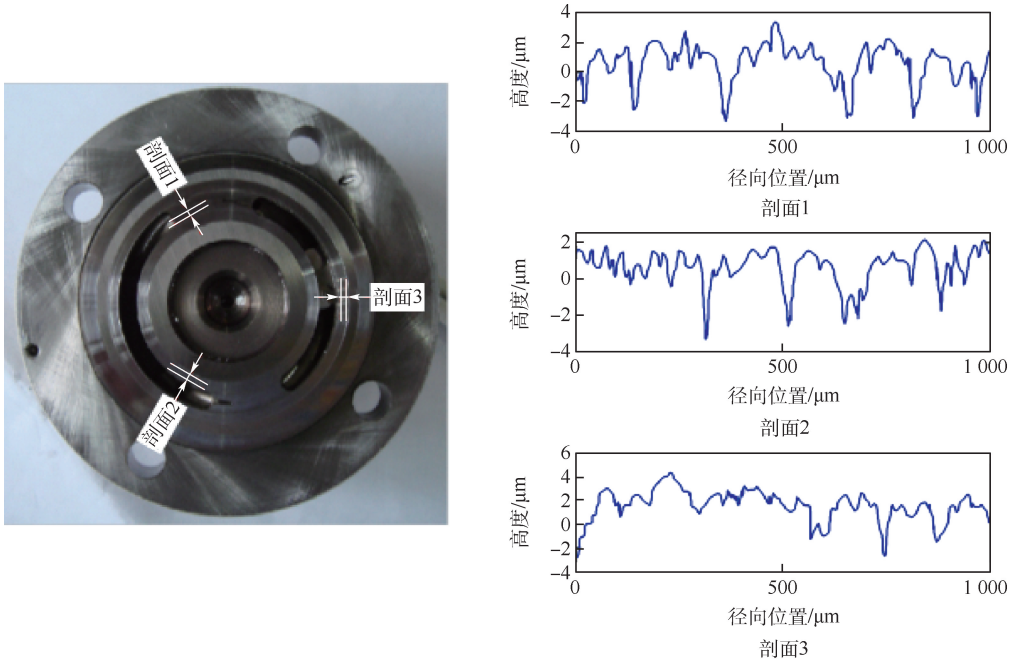
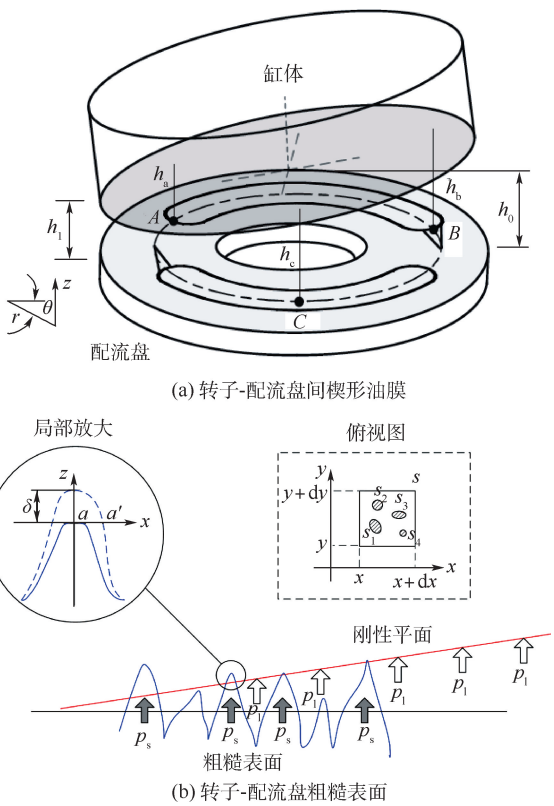


图 2　柱塞式液压泵的配流盘表面形貌

Fig. 2　Surfare topography of valve plate for piston pump



(a) 转子-配流盘间楔形油膜

(b) 转子-配流盘粗糙表面

图 3　柱塞式液压泵转子-配流盘微观关系

Fig. 3　Detailed relationship of valve plate-cylinder block for piston pump

与转速等的关系满足 Reynolds 方程:

$$\frac{\partial}{\partial x}\left(\frac{\rho h^3}{\mu_{\mathrm{P}}}\cdot\frac{\partial p_{\mathrm{l}}}{\partial x}\right)+\frac{\partial}{\partial y}\left(\frac{\rho h^3}{\mu_{\mathrm{P}}}\cdot\frac{\partial p_{\mathrm{l}}}{\partial y}\right)=6\left[\frac{\partial}{\partial x}(v_x\rho h)+\frac{\partial}{\partial y}(v_y\rho h)+2\rho\frac{\partial h}{\partial t}\right] \tag{3}$$

式中:p_{l} 为油膜压力;h 为油膜厚度;μ_{P} 为润滑介质动力黏度;ρ 为油液密度;v_x 和 v_y 分别为 x 方向和 y 方向上的边界流速;t 为时间。

考虑黏温效应和黏压效应,动力黏度 μ 受到温度和压力的影响:

$$\mu=\mu_0\exp[\alpha_{\mathrm{p}}p_{\mathrm{l}}-\alpha_{\mathrm{T}}(T-T_0)] \tag{4}$$

式中:μ_0 为 $p_{\mathrm{l}}=0$ 并且 $T=T_0$ 时油液的动力黏度;α_{p} 和 α_{T} 为常数。

油膜温度分布可以由能量方程解出:

$$c_{\mathrm{p}}\rho\left(\frac{\partial T}{\partial t}+\boldsymbol{v}\cdot\mathrm{grad}\ T\right)=\lambda\cdot\mathrm{div}(\mathrm{grad}\ T)+\mu\Phi_{\mathrm{D}}(\boldsymbol{v}) \tag{5}$$

式中:c_{p} 为流体比热容;$\boldsymbol{v}$ 为流速向量;λ 为流体热导率;Φ_{D} 为能量耗散系数。

为了精确估计油膜厚度,需要考虑金属表面的热变形和压力引起的弹性变形。热变形的计算公式为

$$\varepsilon_{\mathrm{T}}(x,y)=[\boldsymbol{H}]^{-1}\int_V[\boldsymbol{D}]^{\mathrm{T}}[\boldsymbol{E}]\{[\boldsymbol{a}_{\mathrm{T}}]\cdot T(x,y)\}\mathrm{d}V \tag{6}$$

式中:$\boldsymbol{H}$ 为刚度矩阵;$\boldsymbol{D}$ 为节点形状参数矩阵;$\boldsymbol{E}$ 为弹性模量矩阵;$\boldsymbol{a}_{\mathrm{T}}$ 为热膨胀系数矩阵。

弹性变形的计算公式为

$$\varepsilon_{\mathrm{p}}(x,y)=\frac{1-\nu^2}{\pi E}\iint_{\Omega}\frac{p(\xi,\zeta)}{\sqrt{(x-\xi)^2+(y-\zeta)^2}}\mathrm{d}\xi\mathrm{d}\zeta \tag{7}$$

式中:E 为弹性模量;ν 为泊松比;(ξ,ζ) 为转子-配流盘上积分区域 Ω 的坐标。

由于 Reynolds 方程是偏微分方程,难以求出

其解析解,采用数值方法来求封油带上各点油膜压力,又因为转子在不停转动,数值解算的边界条件需要动态更新,图 4 所示为液压泵油膜压力分布和油膜厚度分布,φ 为转子转过的角度。

当配流盘粗糙峰与缸体接触并发生磨粒磨损(见图 5),则单个磨粒嵌入转子-配流盘的深度为

$$\begin{cases}\Delta_{\mathrm{A}}(x,y,D_{\mathrm{d}}) = \dfrac{H}{1+H}\left[\dfrac{D_{\mathrm{d}}}{\tan\beta_{\mathrm{d}}} - h(x,y)\right] \\ \Delta_{\mathrm{B}}(x,y,D_{\mathrm{d}}) = \dfrac{1}{1+H}\left[\dfrac{D_{\mathrm{d}}}{\tan\beta_{\mathrm{d}}} - h(x,y)\right]\end{cases} \tag{8}$$

式中:Δ_{A} 为磨粒嵌入转子表面深度;Δ_{B} 为磨粒嵌入配流盘表面深度;H 为配流盘和转子表面硬度比;D_{d} 为磨粒最大直径;β_{d} 为磨粒形状角。

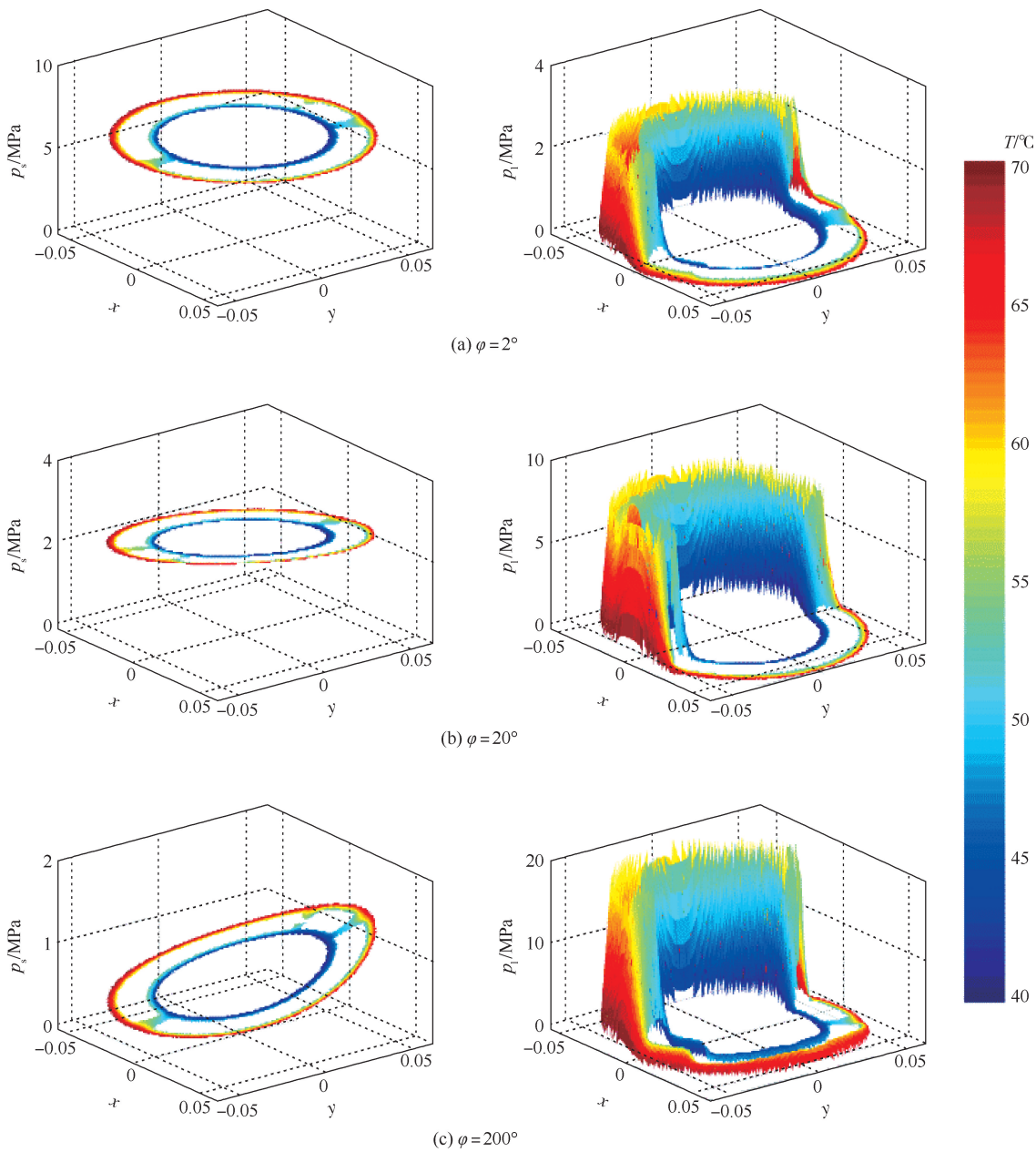
一定时间内单个磨粒对转子和配流盘造成的磨损体积分别为

$$\begin{cases}\Delta V_{\mathrm{A}}(x,y,D_{\mathrm{d}}) = f\Delta_{\mathrm{A}}^2(x,y,D_{\mathrm{d}})\tan\beta_{\mathrm{d}}\omega_{\mathrm{p}}r(x,y) \\ \Delta V_{\mathrm{B}}(x,y,D_{\mathrm{d}}) = f\Delta_{\mathrm{B}}^2(x,y,D_{\mathrm{d}})\tan\beta_{\mathrm{d}}\omega_{\mathrm{p}}r(x,y)\end{cases} \tag{9}$$

式中:f 为磨损系数;ω_{p} 为液压泵转速;$r(x,y)$ 为点(x,y)对应的半径。

一定时间内(x,y)区域内所有磨粒对转子和配流盘造成的磨损体积分别为

$$\begin{cases}V_{\mathrm{A}}(x,y) = \displaystyle\sum_{D_{\mathrm{d}}}\Delta V_{\mathrm{A}}(x,y,D_{\mathrm{d}})k(D_{\mathrm{d}})h(x,y)\,\mathrm{d}x\mathrm{d}y \\ V_{\mathrm{B}}(x,y) = \displaystyle\sum_{D_{\mathrm{d}}}\Delta V_{\mathrm{B}}(x,y,D_{\mathrm{d}})k(D_{\mathrm{d}})h(x,y)\,\mathrm{d}x\mathrm{d}y\end{cases} \tag{10}$$



(a) $\varphi=2°$

(b) $\varphi=20°$

(c) $\varphi=200°$

图 4 柱塞式液压泵转子-配流盘间的油膜厚度、压力和温度变化

Fig. 4 Oil film thickness, pressure and temperature distribution between valve plate-cylinder block for piston pump

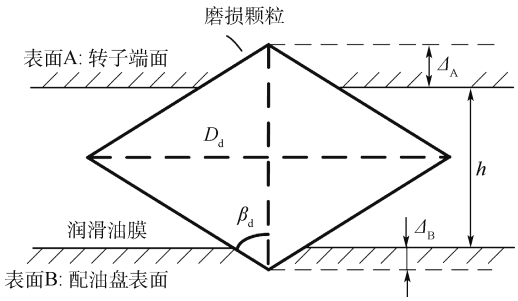


图 5　柱塞式液压泵转子-配流盘润滑磨粒磨损

Fig. 5　Abrasive wear of valve plate-cylinder block for piston pump

式中:$k(D_d)$为尺寸 D_d 的磨粒对应的浓度。则一定时间内转子和配流盘磨损总体积 V_w 为

$$V_w = \iint_{(x,y)\in\Omega} [V_A(x,y) + V_B(x,y)]\,\mathrm{d}x\mathrm{d}y \tag{11}$$

2　液压泵轴尾密封退化建模

飞机液压泵轴尾密封是典型的机械密封,由动环、静环及其之间的密封油液组成,如图 6 所示。其中坐标系定义为:α 为圆周方向,z 为轴向方向,r 为径向方向。r_2、r_1 和 r_b 分别为密封环配合面的外圈半径、内圈半径和平衡半径。

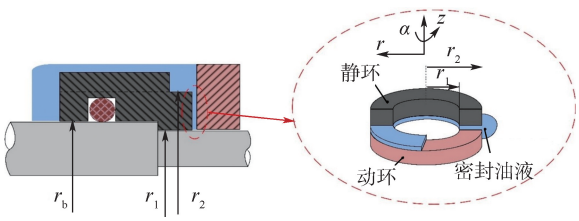


图 6　轴尾机械密封动环与静环间的润滑区域

Fig. 6　Lubrication area between moving ring and stationary ring of axial pump mechanical seal

假设机械密封油膜处于层流工作状态,动环和静环表面的微凸体形变不影响旋转机械密封系统的宏观几何形状,静环和动环的粗糙表面的高度分布是随机的。则动环与静环之间的润滑油膜的压力分布是通过 Reynolds 方程来表征:

$$F\frac{\partial}{\partial r}\left(\frac{rh_n^3}{12\mu}\cdot\frac{\partial D}{\partial r}\right)+F\frac{1}{r}\cdot\frac{\partial}{\partial\alpha}\left(\frac{h_n^3}{12\mu}\cdot\frac{\partial D}{\partial\alpha}\right)=\frac{r\omega}{2}\left[\frac{\partial h_n}{\partial\alpha}+(1-F)\frac{\partial(h_nD)}{\partial\alpha}\right]+r\frac{\partial h_n}{\partial t} \tag{12}$$

式中:r 为径向方向;α 为角度坐标;μ 为油液的动力黏度;$\omega=\partial\alpha/\partial t$ 为动环和静环之间的相对旋转角速度;h_n 为油膜厚度,与机械密封的轴向载荷相关;D 为通用变量,可表示油膜压力 p 或密度 ρ,取决于油膜是否发生空化效应;F 为开关函数,其决定所描述的空化是否发生。

由于机械密封通常在混合润滑状态下工作,而且动环和静环之间微凸体会发生相互作用,可以得到沿密封环油膜承载力为

$$F_f = \int_{r_i}^{r_o} 2\pi r p(r)\,\mathrm{d}r \tag{13}$$

式中:$p(r)$为沿密封环径向油膜压力;r_i、r_o 分别为密封件内、外环半径。

鉴于边界润滑和混合润滑条件下,机械密封表面会产生微凸体接触,采用 Greenwood-Williamson[26] 模型来计算微凸体接触压力。假设密封副表面一个是光滑刚性表面,另一个是服从为概率密度 $\phi(z)$的粗糙表面(见图 7),则接触是有限个独立的 Hertz 接触,微凸体接触压力为

$$p_{ac}(z,t) = \frac{W_a}{A_n} = \frac{4}{3}\eta_e R_e^{\frac{1}{2}} E_e \int_{\bar{h}(z,t)/\sigma_e}^{\infty} \phi(z)(z-\bar{h}(z,t)/\sigma_e)^{\frac{3}{2}}\,\mathrm{d}z \tag{14}$$

式中:微凸体具有球形的顶点,半径 R_e 为常数;$\bar{h}(z,t)$为油膜的平均厚度;η_e 为等效微凸体密度;W_a 为总接触载荷;A_n 为名义接触面积;E_e 为等价弹性模量;等效粗糙度 σ_e 定义为密封环表面粗糙度的均方根值。

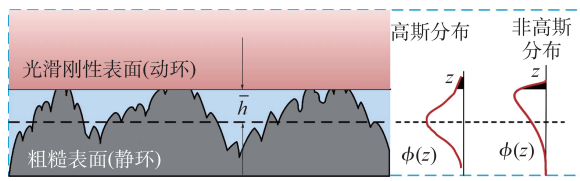


图 7　微凸体接触示意图

Fig. 7　Contact relation of rough surface

如果机械密封的微凸体接触部分为 b_m:

$$b_m(r) = \int_h^{\infty} \phi(z)\,\mathrm{d}z \tag{15}$$

则端面接触压力为

$$p_c(r) = b_m(r)S_C \tag{16}$$

式中:S_C 为屈服应力。

机械密封微凸体的接触压力为

$$F_c = \int_{r_i}^{r_o} 2\pi r p_c(r)\,\mathrm{d}r \tag{17}$$

如果密封副的微凸体接触生热,则机械密封的热传导方程为

$$\frac{\partial T}{\partial t} = \frac{\alpha_{temp}}{r}\cdot\frac{\partial}{\partial r}\left(r\frac{\partial T}{\partial r}\right)+\alpha\frac{\partial^2 T}{\partial z^2}+\frac{\alpha_{temp}}{k_{temp}}q_r \tag{18}$$

式中:T 为密封油液的温度;t 为时间;α_{temp}为热扩散系数;k_{temp} 为导热系数;q_r 为径向 r 处的热流。考虑到油膜轴向 z 方向的厚度很薄,一般重点分析径向 r 方向的温度场分布。

微凸体接触就会产生摩擦磨损。根据 Archard磨损公式,得到密封副的磨损深度为

$$h_w = k_w \frac{F_W}{\pi(r_2^2 - r_1^2)} L \tag{19}$$

式中：k_w 为磨损系数；L 为滑动距离；F_W 为法向接触载荷。继而可以得到时间磨损率 k_t 及距离磨损率 k_s 为

$$\begin{cases} k_t = \dfrac{\mathrm{d}h_w}{\mathrm{d}t} = k_m \dfrac{F_W}{S_a} V_a \\ k_s = \dfrac{\mathrm{d}h_w}{\mathrm{d}L} = k_m \dfrac{F_W}{S_a} \end{cases} \tag{20}$$

式中：$V_a = \omega(r_1 + r_2)/2$ 为密封环旋转平均速度；k_m 为磨损系数；$S_a = \pi(r_2^2 - r_1^2)$ 为磨损接触面积。

3 液压泵二维性能退化可靠性模型

鉴于飞机液压系统的性能退化与飞行剖面有关，而不同的载荷谱作用下性能退化的规律不同，常用随机过程描述。为了统一常用的随机过程，如维纳过程、高斯过程和伽马过程，本文构建统一随机过程模型。

3.1 统一随机过程模型

定义具有统计独立增量的双参数统一随机过程模型为 $X(t) = \mathrm{USP}[\alpha(t), \beta(t)]$，其均值和方差为

$$\begin{cases} E[X(t)] = \mu\Lambda(t) \\ \mathrm{Var}[X(t)] = \sigma^2\Lambda(t) \end{cases} \tag{21}$$

式中：μ 和 $\sigma^2 > 0$ 为常数；退化过程 $\{X(t), t > 0\}$ 中退化量 $X(t)$ 依赖于 2 个参数 $\alpha(t)$ 和 $\beta(t)$ 的取值；$\Lambda(t)$ 为时间尺度函数。不同的随机过程，其统一随机过程模型的参数不同，如表 1 所示。根据表 1可以得到不同退化量的 $X(t)$ 和退化增量 $\Delta X(t)$ 的概率密度函数和累积分布函数的表达式。

表 1 统一随机过程模型定义

Table 1 Definition of unified stochastic process model

模型	$\alpha(t)$	$\beta(t)$
维纳过程	$\mu\Lambda(t)$	$\sigma^2\Lambda(t)$
逆高斯过程	$\mu\Lambda(t)$	$\dfrac{\mu^3}{\sigma^2}\Lambda(t)$
伽马过程	$\dfrac{\mu^2}{\sigma^2}\Lambda(t)$	$\dfrac{\sigma^2}{\mu}$

维纳过程 $\Delta X(t) \sim N(\mu\Delta\Lambda(t), \sigma^2\Delta\Lambda(t))$ 的概率密度函数可以表示为

$$f(\Delta X(t)) = \frac{1}{\sqrt{2\pi\sigma^2\Delta\Lambda(t)}} \exp\left[-\frac{(\Delta X(t) - \mu\Delta\Lambda(t))^2}{2\sigma^2\Delta\Lambda(t)} \right] \tag{22}$$

逆高斯过程 $\Delta X(t) \sim \mathrm{IG}\left(\mu\Delta\Lambda(t), \dfrac{\mu^3}{\sigma^2}\Delta\Lambda^2(t)\right)$ 的概率密度函数可以表示为

$$f(\Delta X(t)) = \sqrt{\frac{\mu^3\Delta\Lambda^2(t)}{2\pi\sigma^2\Delta X(t)^3}} \exp\left[-\frac{\mu(\Delta X(t) - \mu\Delta\Lambda(t))^2}{2\sigma^2\Delta X(t)} \right] \tag{23}$$

伽马过程 $\Delta X(t) \sim \mathrm{Ga}\left(\dfrac{\mu^2}{\sigma^2}\Delta\Lambda(t), \dfrac{\sigma^2}{\mu}\right)$ 的概率密度函数可以表示为

$$f(\Delta X(t)) = \frac{\left(\dfrac{\sigma^2}{\mu}\right)^{-\frac{\mu^2}{\sigma^2}\Delta\Lambda(t)}}{\Gamma\left(\dfrac{\mu^2}{\sigma^2}\Delta\Lambda(t)\right)} \Delta X(t)^{\frac{\mu^2}{\sigma^2}\Delta\Lambda(t)-1} \exp\left(-\frac{\mu}{\sigma^2}\Delta X(t) \right) \tag{24}$$

3.2 二维 Copula 性能退化模型

假设元件退化共监测 k 个性能指标，其对应的性能退化过程用 $\{X_k(t), t > 0\}$ 表示。要描述多个性能指标的相关关系，可以采用 Copula 函数表征其相关关系。k 维 Copula 函数可以定义为

$$C(x_1, x_2, \cdots, x_k) = \Pr(X_1 \leqslant x_1, X_2 \leqslant x_2, \cdots, X_k \leqslant x_k) \tag{25}$$

式中：$x_i (i = 1, 2, \cdots, k)$ 为 Copula 函数的变量；X_i 为服从均匀分布的随机变量且其分布区间为 $[0, 1]$。由式(25)可知，$C(x_1, x_2, \cdots, x_k)$ 实际上是一个 k 维随机变量的联合累积分布函数。如果 k 维随机变量 $X_1, X_2, \cdots, X_k$ 的边缘分布为 $F_i(X_i \leqslant x_i) = F_i(x_i)$ 时，可以将各随机变量的边缘分布函数 $F_i(x_i)$ 作为 Copula 函数的输入，则这 k 个随机变量之间的相关关系描述为

$$F(x_1, x_2, \cdots, x_k) = \Pr(X_1 \leqslant x_1, X_2 \leqslant x_2, \cdots, X_k \leqslant x_k) = C(F_1(x_1), F_2(x_2), \cdots, F_k(x_k)) \tag{26}$$

若只考虑 2 个性能指标，采用二维 Copula 函数得到 2 个性能指标的联合累积概率密度函数为

$$F(X_1, X_2) = C[F_1(x_1), F_2(x_2); \theta] \tag{27}$$

式中：$F(X_1, X_2)$ 为 2 个性能指标的联合累积概率密度函数；$F_1(x_1)$ 和 $F_2(x_2)$ 分别为 2 个性能指标的边缘分布函数；θ 为 Copula 函数中的相关系数，描述了 2 个性能指标的相关程度。

令 $u = F_1(x_1), v = F_2(x_2)$，则 2 个随机变量的联合概率密度函数可以表示为

$$f(u, v) = c(u, v; \theta) f(u) f(v) = \frac{\partial C(u, v; \theta)}{\partial u \partial v} f(u) f(v) \tag{28}$$

式中：$f(u)$、$f(v)$ 分别为 2 个性能指标的边缘概率密度函数；$c(u, v; \theta)$ 为对应的 Copula 函数 $C(u, v; \theta)$ 的密度函数，且有

$$c(u,v;\theta)=\frac{\partial C(u,v;\theta)}{\partial u\partial v} \tag{29}$$

常用的 Copula 函数主要有 3 类,即椭圆 Copula 函数(如 Gaussian Copula 函数,Student-t Copula 函数)、阿基米德 Copula 函数(如 Frank Copula 函数,Clayton Copula 函数)和 FGM Copula 函数等[29]。

3.3　二维 Copula 性能退化的可靠性模型

假设 d_1 和 d_2 为 2 个性能退化指标的失效阈值,当任意一个性能指标达到失效阈值,则该元件或系统失效,则二维性能退化可靠度函数为

$$R(t)=\Pr\{X_1(t)<d_1,X_2(t)<d_2\} \tag{30}$$

当 2 个性能退化指标相关,且相关关系使用 Copula 函数描述时,则二维性能退化可靠度为

$$R(t)=C[R_1(t),R_2(t)] \tag{31}$$

式中:R_1、R_2 分别为 2 个性能退化指标的可靠度函数。

其故障概率密度函数可以表示为

$$f_s(t)=-\frac{\partial R(t)}{\partial t}=\sum_{k=1}^{2}\frac{\partial C[R_1(t),R_2(t)]}{\partial R_k(t)}f_k(t) \tag{32}$$

3.4　未知参数估计

假设对第 $i(i=1,2,\cdots,N)$ 个液压泵进行退化试验,第 $k(k=1,2)$ 个性能指标的第 $j(j=1,2,\cdots,M)$ 个观测量为 $X_k(t_{ij})$。退化增量 $\Delta X_k(t_{ij})=X_k(t_{ij})-X(t_{i,j-1})$。考虑每个性能指标的退化机理不同,可单调递增也可非严格单调退化。采用统一随机过程模型对 2 个性能退化指标的边缘分布进行建模,时间尺度函数定义为 $\Lambda(t)=t^q$;2 个性能退化指标的相关关系采用 Copula 函数进行建模。由此,可以建立起考虑 2 个性能指标相关的可靠性评估模型:

$$\begin{aligned}X_k(t_{ij})&=f[\alpha_k(t_{ij}),\beta_k(t_{ij})]\cdot\\&f(\Delta X_1(t_{ij}),\Delta X_2(t_{ij})\mid\boldsymbol{\Theta}_k,\theta)=\\&\prod_{k=1}^{2}\prod_{i=1}^{N}\prod_{j=1}^{M}f_k(\Delta X_k(t_{ij});\boldsymbol{\Theta}_k)\cdot\\&\prod_{i=1}^{N}\prod_{j=1}^{M}c(F_1(\Delta X_k(t_{ij})),F_2(\Delta X_k(t_{ij}));\theta)\end{aligned} \tag{33}$$

式中:$\boldsymbol{\Theta}_k(k=1,2)$ 为第 $k(k=1,2)$ 个性能指标随机过程模型中的未知参数,即 $\boldsymbol{\Theta}_k=(\mu_k,\sigma_k,q_k)$。

若已获得性能退化数据,式(33)的对数似然函数为

$$\begin{aligned}\ln L(\boldsymbol{\Theta}_k,\theta)=&\sum_{k=1}^{2}\sum_{i=1}^{N}\sum_{j=1}^{M}\ln[f_k(\Delta X_k(t_j);\boldsymbol{\Theta}_k)]+\\&\sum_{i=1}^{N}\sum_{j=1}^{M}\ln[c(F_1(\Delta X_k(t_j)),F_2(\Delta X_k(t_j));\theta)]\end{aligned} \tag{34}$$

利用贝叶斯马尔可夫链蒙特卡罗法可以得到未知参数 $\boldsymbol{\Theta}_k(k=1,2)$ 及 θ 的估计值。

4　液压泵二维性能退化可靠性分析

4.1　试验数据

对飞机液压泵施加载荷谱,获得表征液压泵转子-配流盘摩擦磨损退化的间接泄漏流量参数关系(见图 8)及表征轴尾密封退化的间接摩擦力矩参数变化(见图 9)。图 8 中不同颜色的曲线表示 1 ~ 5 号液压泵回油流量随时间的变化关系,可见随着时间的增加,液压泵由于转子-配流盘磨损回油流量不断增加。图 9 中不同颜色的曲线表示 1 ~ 5 号液压泵轴尾密封间接参数—摩擦力矩随时间的变化关系,可见随着运行时间的增加轴尾密封的摩擦力矩不断减小。

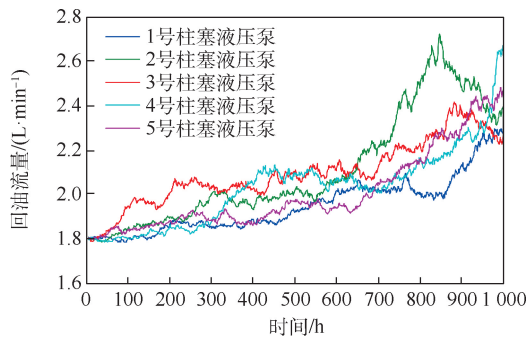


图 8　液压泵回油流量退化曲线

Fig. 8　Degradation paths for return oil flow of hydraulic pump

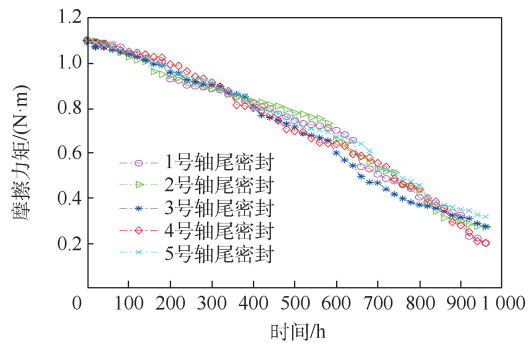


图 9　液压泵轴尾密封磨损摩擦力矩退化曲线

Fig. 9　Degradation paths for friction torque of axial pump mechanical seal

4.2　二维 Copula 性能退化模型参数估计

令图 8 的性能退化量(液压泵回油流量)为 $X_1(t_{ij})$,采用维纳过程描述其性能退化过程;图 9 的性能退化量(液压泵轴尾密封磨损摩擦力矩)为 $X_2(t_{ij})$,采用逆高斯过程描述其性能退化过程;2 个性能指标之间的相关性由 Gaussian Copula 来描述。选取 2 个性能指标对应的时间间隔为时刻点。根据前述的参数估计方法,用贝叶斯马尔可夫链蒙特卡罗法对参数模型中未知参数进行估

计,共进行了 50 000 次的迭代,选取最后 30 000 次迭代的结果来生成未知参数的后验分布的均值。未知参数的后验分布如图 10 所示,图中纵坐标为每个未知参数的后验概率分布,横坐标表示每个未知参数的估计值。估计结果如表 2 所示。

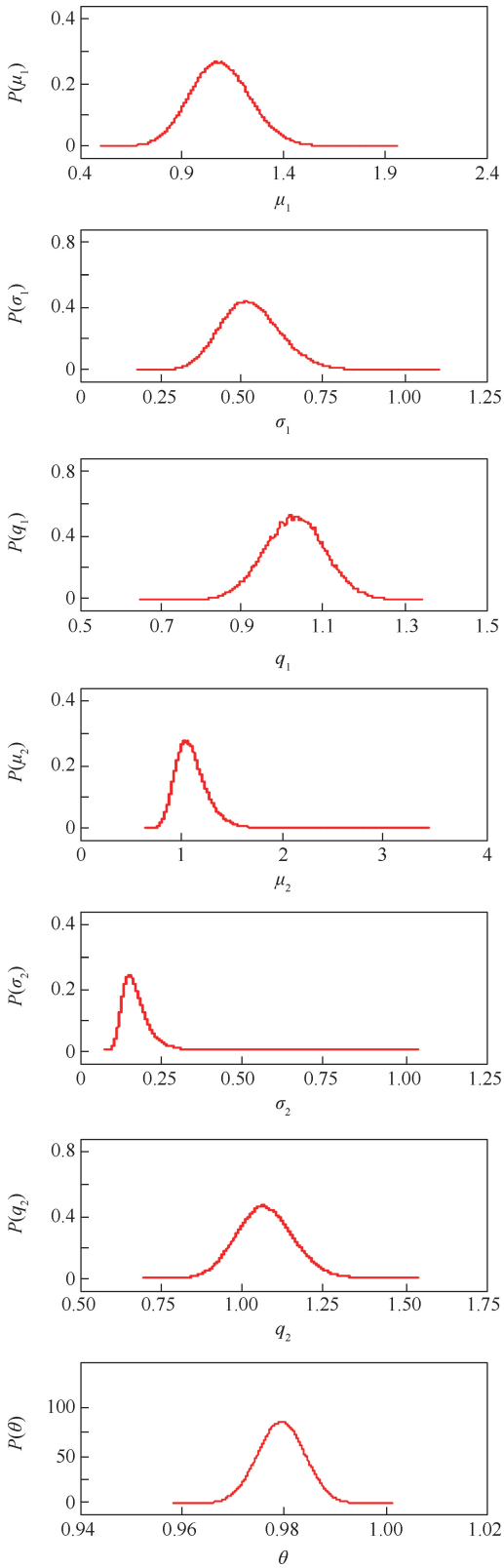


图 10 模型中未知参数的后验分布

Fig. 10 Posterior distribution for unknown parameters in degradation model

通过参数估计的结果可以得到只考虑回油流量、只考虑轴尾密封摩擦力矩以及考虑 2 个退化指标相关的液压泵可靠度随时间变化曲线,如图 11所示。

表 2 模型中未知参数估计结果

Table 2 Estimation results for unknown parameters in degradation model

性能指标	μ	σ	q
X_1	1.189	0.528	1.027
X_2	1.098	0.171 9	1.069

注:相关系数 $\theta=0.874\ 0$。

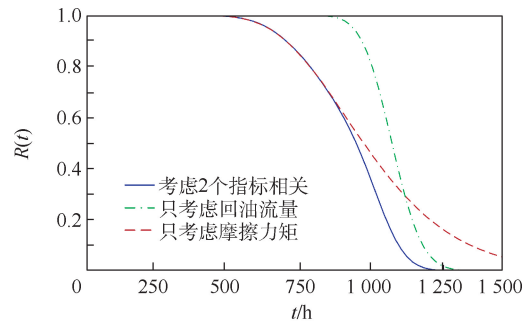


图 11 回油流量、轴尾密封摩擦力矩独立与考虑 2 个性能指标相关性液压泵可靠度随时间曲线

Fig. 11 Reliability curves with time for only considering return oil flow, only considering friction torque and both two dependent performance indicators

从图 11 中可以看出,若只考虑液压泵回油流量这一性能退化指标,可以得到图 11 的绿色点划线可靠度随时间变化曲线。在这种情况下,液压泵在 1 000 h 内可靠度可以处于 0.9 很高的可靠度水平,说明液压泵转子-配流盘设计的润滑油膜是非常有效的,可以保证相当长时间的高可靠度水平。但随着液压泵使用时间的不断增加,楔形油膜最小油膜处会出现短暂的摩擦磨损,进而不断污染润滑油膜,快速使转子-配流盘磨损加剧,影响液压泵回油流量的变化,液压泵可靠度快速下降。

若只考虑轴尾密封磨损这一性能退化指标,可以得到图 11 的红色虚线可靠度随时间变化曲线。在这种情况下,液压泵轴尾密封可以在750 h 内具有 0.9 较高的可靠度水平。但如果其机械密封处轴向载荷变化或油膜出现空化效应,由于其密封油液很薄,其摩擦磨损性能退化会较快,但可靠度达到 0.5 后其下降速度变缓,这时如果其他性能退化变快(如转子-配流盘摩擦磨损),其他性能退化会与轴尾密封产生竞争失效。

如果同时考虑转子-配流盘磨损的性能退化和轴尾密封磨损的性能退化,利用本文提出的考

虑 2 个性能退化指标相关的可靠度评估模型可以得到图 11 的蓝色曲线。可见,综合考虑二维性能退化其可靠度曲线会较单独考虑一维性能退化要更为保守。再看其寿命指标,考虑 2 个相关性能退化指标的可靠度曲线计算得到液压泵寿命为 1 187 h,而实际寿命为 1 130 h,精度为 95.20%。若只考虑单一性能退化指标,从图 11 中可以发现其可靠度曲线都比考虑两相关的可靠度曲线高,寿命预测精度显然不如考虑两相关的性能退化指标。因此,考虑二维相关退化指标的可靠寿命评估比单独的可靠寿命评估与实际更加吻合。

5　结　论

本文针对飞机液压泵多失效机理的现状,提出了二维性能退化的可靠性评估方法。通过试验验证,得到如下结论:

1) 构建的飞机液压泵混合润滑多场耦合摩擦磨损性能退化模型和轴尾密封摩擦磨损退化模型可以准确表征液压泵摩擦副、密封副的多场作用下的性能退化关系。

2) 将液压泵各退化过程用统计随机过程表征,基于 Copula 函数将二维性能退化进行融合,可以实现 2 个性能退化过程的故障概率与可靠度综合表征。

3) 采用贝叶斯马尔可夫链蒙特卡罗法得到的二维性能退化的后验分布及参数估计,进而给出二维性能退化下的可靠性评估,试验验证了其有效性。

参考文献 (References)

[1] 焦宗夏,孔祥东,王少萍,等. 大型飞机电液动力控制与作动系统新体系基础研究[J]. 中国基础科学,2018,20(2):41-47.
JIAO Z X,KONG X D,WANG S P,et al. Advancements of basic researches on large aircraft of electro-hydraulic power and actuation system new architecture[J]. China Basic Science,2018,20(2):41-47(in Chinese).

[2] SHI C,WANG S P,WANG X J,et al. Variable load failure mechanism for high-speed load sensing electro-hydrostatic actuator pump of aircraft[J]. Chinese Journal of Aeronautics,2018,31(5):949-964.

[3] 王少萍. 液压泵计算机辅助测试及故障诊断研究[D]. 北京:北京航空航天大学,1994.
WANG S P. Computer aided test and fault diagnosis to hydraulic pump[D]. Beijing:Beihang University,1994(in Chinese).

[4] GUO S R,CHEN J H,LU Y L,et al. Hydraulic piston pump in civil aircraft:Current status,future directions and critical technologies[J]. Chinese Journal of Aeronautics,2020,33(1):16-30.

[5] HUNT T M. Condition monitoring of ship borne hydraulic pump [D]. Bath:University of Bath,1981.

[6] LANGEN H J. Early detection of damage in hydraulic components by measurements of structure-borne sound[J]. Oelhydraul. Pneum,1981,25(7):568-573.

[7] IVANTYSYNOVA M,LASAAR R. An investigation into micro- and macrogeometric design of piston/cylinder assembly of swash plate machines[J]. International Journal of Fluid Power,2004,5(1):23-36.

[8] PELOSI M,IVANTYSYNOVA M. A novel thermal model for the piston/cylinder interface of piston machines[C]// Proceedings of ASME 2009 Dynamic Systems and Control Conference. New York:ASME,2010:37-44.

[9] PPELOSI M,IVANTYSYNOVA M. Heat transfer and thermal elastic deformation analysis on the piston/cylinder interface of axial piston machines[J]. Journal of Tribology,2012,134(4):041101.

[10] KASSEM S A,BAHR M K. Effect of port plate silencing grooves on performance of swash plate axial piston pumps[C]// Current Advances in Mechanical Design and Production VII. Amsterdam:Elsevier,2000:139-148.

[11] HONG Y S,LEE S Y,KIM S H,et al. Improvement of the low-speed friction characteristics of a hydraulic piston pump by PVD-coating of TiN[J]. Journal of Mechanical Science and Technology,2006,20(3):358-365.

[12] FANG Y,SHIRAKASHI M. Mixed lubrication characteristics between the piston and cylinder in hydraulic piston pump-motor [J]. Journal of Tribology,1995,117(1):80-85.

[13] MANRING N D. Friction forces within the cylinder bores of swash-plate type axial-piston pumps and motors[J]. Journal of Dynamic Systems,Measurement,and Control,1999,121(3):531-537.

[14] YAMAGUCHI A,MATSUOKA H. A mixed lubrication model applicable to bearing/seal parts of hydraulic equipment[J]. Journal of Tribology,1992,114(1):116-121.

[15] KAZAMA T,YAMAGUCHI A. Application of a mixed lubrication model for hydrostatic thrust bearings of hydraulic equipment[J]. Journal of Tribology,1993,115(4):686-691.

[16] 诸葛起,路甬祥. 用振动信号检测柱塞泵松靴故障[J]. 振动、测试与诊断,1989,9(3):11-15.
ZHUGE Q,LU Y X. Using vibration signal to detect failure of piston pump's slipper pair[J]. Journal of Vibration,Measurement & Diagnosis,1989,9(3):11-15(in Chinese).

[17] 徐兵,李迎兵,张斌,等. 轴向柱塞泵滑靴副倾覆现象数值分析[J]. 机械工程学报,2010,46(20):161-168.
XU B,LI Y B,ZHANG B,et al. Numerical simulation of overturning phenomenon of axial piston pump slipper pair[J]. Journal of Mechanical Engineering,2010,46(20):161-168(in Chinese).

[18] 刘洪,苑士华,彭增雄. 轴向柱塞泵滑靴油膜动态仿真[J]. 北京理工大学学报,2011,31(11):1282-1286.
LIU H,YUAN S H,PENG Z X. Dynamic simulation of slipper bearings in axial piston pumps[J]. Transactions of Beijing Institute of Technology,2011,31(11):1282-1286(in Chinese).

[19] 刘思远,王闯,杨梦雪,等.斜盘泵滑靴副剧烈磨损过程的动态特性分析[J].液压与气动,2017(1):1-5.
LIU S Y, WANG C, YANG M X, et al. Analysis of dynamic characteristics for severe wear process of swash plate axial piston pump slipper pair[J]. Chinese Hydraulics & Pneumatics, 2017(1):1-5(in Chinese).

[20] 杨光琴.多传感器航空液压系统信息融合故障诊断研究[D].北京:北京航空航天大学,2003.
YANG G Q. Multi-sensors fusion based fault diagnosis of aerial hydraulic system[D]. Beijing: Beihang University, 2003 (in Chinese).

[21] 张文超.机载液压泵源健康管理[D].北京:北京航空航天大学,2008.
ZHANG W C. Health management of airborne hydraulic pump [D]. Beijing: Beihang University, 2008(in Chinese).

[22] 赵四军.飞机液压泵源故障预测与健康管理系统关键技术研究[D].北京:北京航空航天大学,2010.
ZHAO S J. Fault prognostics and health management of aircraft hydraulic power supply[D]. Beijing: Beihang University, 2010 (in Chinese).

[23] 杜隽.航空液压能源系统动态失效机理与健康管理关键技术研究[D].北京:北京航空航天大学,2012.
DU J. Fault mechanism and health management of aviation hydraulic system[D]. Beijing: Beihang University, 2012(in Chinese).

[24] 何兆民.航空液压泵健康状态识别与故障预测关键技术研究[D].北京:北京航空航天大学,2014.
HE Z M. Health state identification and fault prediction of aircraft hydraulic pump[D]. Beijing: Beihang University, 2014(in Chinese).

[25] 卢传奇.基于分层递进的航空液压系统故障诊断与寿命预测研究[D].北京:北京航空航天大学,2017.
LU C Q. Fault diagnosis and life prediction of aircraft hydraulic system based on hierarchical progression[D]. Beijing: Beihang University, 2017(in Chinese).

[26] 贺文涛.旋转机械密封失效机理与寿命预测研究[D].北京:北京航空航天大学,2020.
HE W T. Failure mechanism and life prediction of rotating machinery seal[D]. Beijing: Beihang University, 2020 (in Chinese).

[27] MA Z H, WANG S P, LIAO H T, et al. Engineering-driven performance degradation analysis of hydraulic piston pump based on the inverse Gaussian process[J]. Quality and Reliability Engineering International, 2019, 35(7):2278-2296.

[28] CHEN R T, ZHANG C, WANG S P, et al. Reliability estimation of mechanical seals based on bivariate dependence analysis and considering model uncertainty[J]. Chinese Journal of Aeronautics, 2021, 34(5):554-572.

[29] ZHANG C, CHEN R T, BAI G H, et al. Reliability estimation of rotary lip seal in aircraft utility system based on time-varying dependence degradation model and its experimental validation [J]. Chinese Journal of Aeronautics, 2020, 33(8):2230-2241.

[30] 张超.星载传动机构加速寿命试验理论及评估方法研究[D].北京:北京航空航天大学,2012.
ZHANG C. Accelerated life test and evaluation of spaceborne transmission mechanism[D]. Beijing: Beihang University, 2012 (in Chinese).

Reliability estimation for aircraft hydraulic pump based on bivariate performance degradation analysis

WANG Shaoping[1,2,*], CHEN Rentong[1], ZHANG Chao[2,3]

(1. School of Automation Science and Electrical Engineering, Beihang University, Beijing 100083, China;
2. Ningbo Institute of Technology, Beihang University, Ningbo 315800, China;
3. Research Institute for Frontier Science, Beihang University, Beijing 100083, China)

Abstract: An Aircraft hydraulic pump provides high-pressure energy for aircraft manipulation, landing gear retraction and braking. Maintaining its reliability and safety is critical. In this paper, we construct a mixed lubrication wear degradation model based on the micro topography for rotor-valve plate friction pairs and axial pump mechanical seals. The unified stochastic process model is used to describe the degradation process of these two degradation performance indicators, Copula functions are used to describe the dependent relationships between two performance indicators. Bayesian Markov chain Monte Carlo method is used to estimate unknown parameters in the degradation model considering two dependent performance indicators. The results of the estimation are then utilized to conduct reliability evaluations for aircraft hydraulic pumps. The proposed theory and method were verified through real degradation test data of return oil flow for hydraulic pumps and friction torque for mechanical seals. The results show that the degradation model considering two dependent performance indicators can increase the accuracy of reliability estimation and lifetime prediction.

Keywords: bivariate performance degradation; Copula function; parameter estimation; aircraft hydraulic pump; reliability estimation

Received: 2022-04-29; **Accepted**: 2022-05-29; **Published online**: 2022-06-21 13:26

URL: kns.cnki.net/kcms/detail/11.2625.V.20220620.1546.001.html

Foundation items: NSFC Projects of International Cooperation and Exchanges (51620105010); National Science and Technology Major Project (J2019-V-0016-0111)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0308

基于指数平均动量鸽群优化的多无人机协同目标防御

段海滨[1,*], 仝秉达[1], 刘冀川[2,3]

(1. 北京航空航天大学 自动化科学与电气工程学院, 北京 100083;
2. 中国电子科技集团公司第五十四研究所, 石家庄 050081; 3. 西安电子科技大学 电子工程学院, 西安 710071)

摘 要: 针对多无人机(UAV)协同目标防御问题,提出了一种基于指数平均动量鸽群优化(EM-PIO)算法。针对三维空间中的多无人机协同目标防御系统进行建模,得到了无人机支配区域的曲面约束方程,并获得了双方无人机的最优控制输入量。采用多级罚函数法构造了优化算法的目标函数,并通过所提出的 EM-PIO 算法来求解最优目标点。将所提 EM-PIO 算法与遗传算法(GA)和粒子群优化(PSO)算法进行仿真对比实验,验证了所提 EM-PIO 算法更加有效解决多无人机协同目标防御问题。

关 键 词: 无人机(UAV)对抗; 鸽群优化(PIO); 目标防御; 微分博弈; 鞍点策略
中图分类号: V221$^+$.3; TB553
文献标志码: A **文章编号**: 1001-5965(2022)09-1624-06

无人机(unmanned aerial vehicle, UAV)自主控制技术及低成本传感器技术的快速发展,使得无人机系统越来越广泛地应用于民用和军事领域[1-2]。由于任务环境复杂而多变,单架无人机往往不能有效完成任务,而使用多架无人机协作能够有效提高成功率[3],顺利完成各种定位、搜索、攻击、安全、监视、评估等复杂任务[4]。本文主要针对与安全和防御应用相关场景的无人机协同目标防御问题进行了探索研究,在目标防御问题中,入侵无人机在收集到防御无人机的状态信息后,试图抵达目标区域而不被防御无人机拦截,而防御方若干架无人机的任务是尽快拦截入侵无人机,防止对方抵达目标区域。

von Moll 等[5]将多无人机协同目标防御场景描述为一个微分博弈问题。场景中的无人机具有简单的运动学方程,防御无人机扮演微分博弈中的追捕者角色,入侵无人机扮演逃跑者角色。Garcia 等[6]研究了多参与者的边界防御问题,并给出了团队合作最优解,场景中的智能体能够利用对手的非最优策略使得己方的收益最大化。然而,上述研究只考虑了当防御者与入侵者的位置重合时,入侵者才视为被捕获这种情况,且防御无人机要保护的目标边界是无限大的。事实上,无人机可以在一定距离时使用自身携带的武器,干扰或者摧毁对方,且目标区域可能有界。Shishika 和 Kumar[7]研究了一类具有任意凸形状的边界防御问题,入侵者团队试图突破防御者团队对目标区域的保护,而防御者团队试图通过拦截入侵者智能体。Sinha 等[8]研究了 3 个智能体之间的追逃场景,并分别为入侵者和防御者设计了控制策

收稿日期: 2022-04-30; **录用日期**: 2022-05-18; **网络出版时间**: 2022-06-06 14:40
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220602.1834.001.html
基金项目: 科技创新 2030-"新一代人工智能"重大项目(2018AAA0102303); 国家自然科学基金(U20B2071, 91948204, T2121003, U1913602)
* **通信作者**. E-mail: hbduan@buaa.edu.cn

引用格式: *段海滨, 仝秉达, 刘冀川. 基于指数平均动量鸽群优化的多无人机协同目标防御[J]. 北京航空航天大学学报, 2022, 48(9): 1624-1629. DUAN H B, TONG B D, LIU J C. Coordinated target defense for multi-UAVs based on exponentially averaged momentum pigeon-inspired optimization[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1624-1629(in Chinese).*

略。Wang 等[9]考虑了具有通信约束的追逃问题,并分别求解了追逃双方应当使用的策略。然而,上述研究只考虑了防御者被限制在二维目标区域的边界中运动情况。实际上,入侵无人机或防御无人机均可以在三维空间中自由运动。

基于上述研究,考虑一个复杂环境下面向多无人机协同目标区域防御问题。其中,防御无人机的数量 $M>1$,入侵无人机的数量为 1。防御无人机的速度可能各不相同,但均大于入侵无人机。防御团队要保护的目标区域为一个有限大的三维空间,且防御无人机在与入侵无人机一定距离时即可拦截。基于上述假设,本文将无人机目标区域防御问题建模为约束最优化问题,解决问题的关键是求解出双方无人机的最优拦截(目标)点。因此,本文设计了一种新型的改进鸽群优化算法解决此类问题。

本文设计了一种无人机协同目标防御系统策略,系统中的防御无人机根据系统实时状态进行合作,对进入捕获半径的入侵无人机进行拦截。另外,针对无人机协同目标防御问题需要求解的约束最优化问题定义了多级非稳态罚函数,便于优化算法找出可行的最优解。对基本鸽群优化(pigeon-inspired optimization, PIO)算法进行了改进,有效解决了原始算法在收敛性和准确性方面的不足,并将改进后的 PIO 算法应用于解决多无人机协同目标防御问题。

1　无人机协同目标防御系统建模

考虑一个由 M 架防御无人机 $P_1,P_2,\cdots,P_M$ 和 1 架入侵无人机 E 构成的无人机协同目标防御系统,系统中所有无人机均在欧氏三维空间中运动。受文献[5]的启发,系统中的各无人机具有如下运动学方程:

$$\begin{cases}\dot{x}_E=\cos\theta_E\cos\psi_E\\ \dot{y}_E=\cos\theta_E\sin\psi_E\\ \dot{z}_E=\sin\theta_E\\ \dot{x}_{P_i}=\beta_i\cos\theta_{P_i}\cos\psi_{P_i}\\ \dot{y}_{P_i}=\beta_i\cos\theta_{P_i}\sin\psi_{P_i}\\ \dot{z}_{P_i}=\beta_i\sin\theta_{P_i}\end{cases}\quad i=1,2,\cdots,M \tag{1}$$

式中:$\beta_i>1$ 为防御无人机 P_i 与入侵无人机 E 的速度比;$\theta_E\in[-\pi,\pi)$ 和 $\psi_E\in[0,2\pi)$ 分别为入侵无人机的俯仰角和航向角;$\theta_{P_i}\in[-\pi,\pi)$ 和 $\psi_{P_i}\in[0,2\pi)$ $(i=1,2,\cdots,M)$ 分别为防御无人机的俯仰角和航向角。双方无人机的控制量分别为 $\boldsymbol{u}_E=\{\theta_E,\psi_E\}$ 和 $\boldsymbol{u}_P=\{\theta_{P_1},\psi_{P_1},\cdots,\theta_{P_M},\psi_{P_M}\}$。该无人机协同目标防御系统的状态由系统中每架无人机的位置确定:

$$\boldsymbol{x}=(\boldsymbol{x}_E,\boldsymbol{x}_{P_i})\in\mathbf{R}^{3M+3} \tag{2}$$

$$\begin{cases}\boldsymbol{x}_E=(x_E,y_E,z_E)\in\mathbf{R}^3\\ \boldsymbol{x}_{P_i}=(x_{P_i},y_{P_i},z_{P_i})\in\mathbf{R}^{3M}\end{cases}\quad i=1,2,\cdots,M \tag{3}$$

设系统初始时刻为 t_0,初始状态为 $\boldsymbol{x}_0=\boldsymbol{x}(t_0)=(\boldsymbol{x}_{E0},\boldsymbol{x}_{P_i0})$,其中 $\boldsymbol{x}_{E0}=\boldsymbol{x}_E(t_0)$ 和 $\boldsymbol{x}_{P_i0}=\boldsymbol{x}_{P_i}(t_0)$ $(i=1,2,\cdots,M)$。系统终端时刻为 t_f,终端状态为 $\boldsymbol{x}_{f0}=\boldsymbol{x}_f(t_f)=(\boldsymbol{x}_{Ef},\boldsymbol{x}_{P_if})$,其中 $\boldsymbol{x}_E(t_f)=\boldsymbol{x}_{Ef}$ 和 $\boldsymbol{x}_{P_i}(t_f)=\boldsymbol{x}_{P_if}$。

防御无人机要防御的目标区域为球体,球心为 $\boldsymbol{x}_T=(x_T,y_T,z_T)$,半径为 r_T。防御无人机的目标是拦截入侵无人机,使入侵无人机与所要防御目标区域距离最远,其捕获半径为 r_c。入侵无人机的目标为在终端时刻 t_f 时尽量缩短自身与目标区域之间的距离。假设入侵无人机不可能到达目标平面,即考虑防御无人机能够拦截成功的定量博弈问题。因此博弈对抗的终止条件为

$$T=\left\{\prod_{i=1}^{M}\|\boldsymbol{x}_E-\boldsymbol{x}_{P_i}\|\leqslant r_c\right\}\quad i=1,2,\cdots,M \tag{4}$$

性能指标泛函为

$$J(\boldsymbol{u}_E(\boldsymbol{x}),\boldsymbol{u}_P(\boldsymbol{x});\boldsymbol{x}_0)=\|\boldsymbol{x}_{Ef}-\boldsymbol{x}_T\| \tag{5}$$

值函数定义为

$$V(\boldsymbol{x})=\min_{\boldsymbol{u}_E(\cdot)}\max_{\boldsymbol{u}_P(\cdot)}J(\boldsymbol{u}_E(\boldsymbol{x}),\boldsymbol{u}_P(\boldsymbol{x});\boldsymbol{x}_0)=\min_{\boldsymbol{u}_E(\cdot)}\max_{\boldsymbol{u}_P(\cdot)}\|\boldsymbol{x}_{Ef}-\boldsymbol{x}_T\| \tag{6}$$

另外设双方无人机均使用开环策略,可以获得系统当前状态,无法获得其他无人机当前或未来的控制量。同时,假设此目标防御系统的最优鞍点策略是存在的,即存在最优控制输入 $\boldsymbol{u}_E^*$ 和 $\boldsymbol{u}_P^*$,使得以下不等式组成立:

$$J(\boldsymbol{u}_E,\boldsymbol{u}_P^*;\boldsymbol{x}_0)\leqslant J(\boldsymbol{u}_E^*,\boldsymbol{u}_P^*;\boldsymbol{x}_0)\leqslant J(\boldsymbol{u}_E^*,\boldsymbol{u}_P;\boldsymbol{x}_0) \tag{7}$$

防御无人机与入侵无人机均以恒定的速度运动,故双方无人机的最优路径均为直线,双方的支配区域由以下等式确定的曲面分隔:

$$\frac{\|\boldsymbol{x}-\boldsymbol{x}_{P_i0}\|-r_c}{\|\boldsymbol{x}-\boldsymbol{x}_{E0}\|}=\beta_i\quad i=1,2,\cdots,M \tag{8}$$

式中:$\boldsymbol{x}=(x,y,z)\in\mathbf{R}^3$。式(8)给出了入侵无人机的可行解区域。当式(8)中等号成立时,入侵无人机 E 可以在中途不被防御无人机 P_i 拦截的情况下到达点曲面上的任意一点。另外设入侵无

人机 E 的最优目标点为 $\boldsymbol{x}_{\mathrm{I}}=(x_{\mathrm{I}},y_{\mathrm{I}},z_{\mathrm{I}})$，除了应当满足式(8)之外，还应当有如下等式成立：

$$\boldsymbol{x}_{\mathrm{I}}=\min_{\boldsymbol{x}}(\|\boldsymbol{x}-\boldsymbol{x}_{\mathrm{T}}\|) \tag{9}$$

此时双方无人机的最优控制输入量为

$$\begin{cases}\theta_E^*=\arctan\left(\dfrac{z_{\mathrm{I}}-z_{E0}}{\sqrt{(x_{\mathrm{I}}-x_{E0})^2+(y_{\mathrm{I}}-y_{E0})^2}}\right)\\ \psi_E^*=\arctan\left(\dfrac{y_{\mathrm{I}}-y_{E0}}{x_{\mathrm{I}}-x_{E0}}\right)\\ \theta_{P_i}^*=\arctan\left(\dfrac{z_{\mathrm{I}}-z_{P_i0}}{\sqrt{(x_{\mathrm{I}}-x_{P_i})^2+(y_{\mathrm{I}}-y_{P_i})^2}}\right)\\ \psi_{P_i}^*=\arctan\left(\dfrac{y_{\mathrm{I}}-y_{P_i0}}{x_{\mathrm{I}}-x_{P_i0}}\right)\end{cases} \tag{10}$$

2 罚函数法构造优化目标函数

由第1节的系统建模可知，确定防御无人机和入侵无人机的最优目标点实际是求解一个由式(1)确定的约束最优化问题：

$$\min_{\boldsymbol{x}}f(\boldsymbol{x})=\min_{\boldsymbol{x}}(\|\boldsymbol{x}-\boldsymbol{x}_{\mathrm{T}}\|) \tag{11}$$

满足以下非线性约束：

$$g_i(\boldsymbol{x})=\frac{\|\boldsymbol{x}-\boldsymbol{x}_{P_i0}\|-r_{\mathrm{c}}}{\|\boldsymbol{x}-\boldsymbol{x}_{E0}\|}-\beta_i=0$$
$$i=1,2,\cdots,M \tag{12}$$

解决约束最优化问题的常用方法是使用罚函数法构建目标函数 $F(\boldsymbol{x})$，转化为无约束最优化问题然后使用优化算法求解。约束最优化问题由可行解和不可行解组成，其中可行解满足所有约束条件，而不可行解至少违反其中一个约束条件。目前为止，除了试错法(trial-and-error)之外，还没有其他方式定义罚函数的方法。然而，罚函数的定义仍具有挑战性，如果惩罚值过高，最优化算法通常会陷入局部最优解；如果惩罚值过低，优化算法可能很难得到可行的最优解。

罚函数通常分为稳态罚函数和非稳态罚函数两类。稳态罚函数在整个最优化的过程中使用固定的惩罚值；非稳态罚函数中，惩罚值是动态变化的。参考文献[10-11]中的结果显示，使用非稳态罚函数得到的结果几乎总是优于通过稳态罚函数的结果。

本文采用的罚函数可定义如下：

$$F(\boldsymbol{x})=f(\boldsymbol{x})+h(k)H(\boldsymbol{x}) \tag{13}$$

式中：$f(\boldsymbol{x})$为式(11)中约束最优化问题的原始目标函数；$h(k)$为一个动态调整的惩罚值；k 为优化算法当前迭代次数；$H(\boldsymbol{x})$为惩罚因子，定义为

$$H(\boldsymbol{x})=\sum_{i=1}^{M}\sigma(|g_i(\boldsymbol{x})|)g_i(\boldsymbol{x})^{2\gamma(|g_i(\boldsymbol{x})|)} \tag{14}$$

式中：$\sigma(\cdot)$为一个多级函数；$\gamma(\cdot)$为罚函数的指数函数；$g_i(\boldsymbol{x})$为式(12)中的约束项。

为了优化算法能够尽快收敛，本文使用的罚函数设置了容差 $|g_i(\boldsymbol{x})|\leqslant 10^{-5}$，当低于该阈值时 $\sigma(|g_i(\boldsymbol{x})|)=0$。函数 $h(\cdot)$、$\sigma(\cdot)$和 $\gamma(\cdot)$的定义如下：

$$h(k)=k\sqrt{k} \tag{15}$$

$$\sigma(|g_i(\boldsymbol{x})|)=\begin{cases}10 & 10^{-5}<|g_i(\boldsymbol{x})|\leqslant 10^{-3}\\ 20 & 10^{-3}<|g_i(\boldsymbol{x})|\leqslant 10^{-1}\\ 100 & 10^{-1}<|g_i(\boldsymbol{x})|\leqslant 1\\ 300 & 1<|g_i(\boldsymbol{x})|\end{cases} \tag{16}$$

$$\gamma(|g_i(\boldsymbol{x})|)=\begin{cases}1 & |g_i(\boldsymbol{x})|<1\\ 2 & |g_i(\boldsymbol{x})|\geqslant 1\end{cases} \tag{17}$$

3 指数平均动量鸽群优化算法

本文所要解决的约束最优化问题可采用确定性或者随机性方法求解。确定性方法，如可行方向法或者广义梯度下降法，对目标函数 $f(\boldsymbol{x})$的连续性和可微性具有一定要求。因此，使用随机性方法解决约束最优化问题是近年来的热门发展方向。虽然进化算法(evolutionary algorithms, EA)主要是解决无约束最优化问题而发展起来的，但其也是解决约束最优化问题的一种可行的替代方法。典型的进化算法有遗传算法(genetic algorithm, GA)[12]和粒子群优化(particle swarm optimization, PSO)算法[13]，均已经被用于解决约束最优化问题中。

针对无人机航路规划问题，Duan 和 Qiao[14]提出了一种新的生物启发式群体优化算法——鸽群优化算法。该算法基于鸽子的归巢行为，设计了地图和指南针算子、地标算子，以求解最优化问题。假设搜索空间的维度为 D，鸽群中的第 i 只鸽子由 D 维向量 $\boldsymbol{X}_i=(x_{i1},x_{i2}\cdots,x_{iD})$表示，鸽群中取得全局最优值的鸽子用向量 $\boldsymbol{X}_g=(x_{g1},x_{g2},\cdots,x_{gD})$表示。第 i 只鸽子的速度由向量 $\boldsymbol{V}_i=(v_{i1},v_{i2},\cdots,v_{iD})$表示。

在地图和指南针算子中，鸽群中的鸽子位置根据式(18)和式(19)进行更新：

$$\boldsymbol{V}_i^{k+1}=\boldsymbol{V}_i^k\mathrm{e}^{-Rt}+r(\boldsymbol{X}_g^k-\boldsymbol{X}_i^k) \tag{18}$$

$$\boldsymbol{X}_i^{k+1}=\boldsymbol{X}_i^k+\boldsymbol{V}_i^{k+1} \tag{19}$$

式中：$i=1,2,\cdots,N$ 为鸽群中的鸽子序号；R 为地图和指南针因子；r 为在[0,1]范围内均匀分布的随机数。式(18)用于确定鸽群中第 i 只鸽子第

$k+1$ 次迭代的速度,式(19)用于确定鸽群中第 i 只鸽子第$k+1$ 迭代的位置,即将第 k 次迭代的位置与第$k+1$ 次迭代的速度相加。

在地标算子中,每次迭代之后鸽子的数量会减少一半,目标函数值较低的一半鸽子将被舍弃,即

$$N^{k+1} = \frac{N^k}{2} \tag{20}$$

式中:N^{k+1}和 N^k 分别为第 $k+1$ 次和 k 次迭代中鸽群中鸽子数量。此时鸽群中的鸽子位置通过如下更新:

$$\boldsymbol{X}_c^{k+1} = \frac{\sum_{i=1}^{N^{k+1}} \boldsymbol{X}_i^k F(\boldsymbol{X}_i^k)}{\sum_{i=1}^{N^{k+1}} F(\boldsymbol{X}_i^k)} \tag{21}$$

$$\boldsymbol{X}_i^{k+1} = \boldsymbol{X}_i^k + r(\boldsymbol{X}_c^k - \boldsymbol{X}_i^k) \tag{22}$$

式中:$F(\cdot)$为第 2 节定义的罚函数;$\boldsymbol{X}_c^{k+1}$ 为第$k+1$次迭代鸽群中 N^{k+1}只鸽子的中心。

尽管基本鸽群优化算法能够求解许多函数最优化问题,但仍然存在收敛性和准确性不足、效率不高等问题。基于此,本文提出了一种新的改进鸽群优化算法-指数平均动量鸽群优化(exponentially averaged momentum PIO,EM-PIO)算法,以解决多无人机协同目标防御问题。

在机器学习中,反向传播(back propagation,BP)算法是用于训练多层前馈神经网络的最常用算法之一。BP 算法使用梯度下降法来最小化实际输出和期望输出之间的误差,但这种算法常常取得局部最优或者在附近振荡,无法收敛到全局最优值。因此可以引入一个动量项来解决此问题,该动量项可作为一个低通滤波器来平滑输出[15]。受此启发,本文在基本鸽群优化算法中的地图和指南针算子速度更新方程(式(18))中,对方程中的探索部分赋予更多的权重。新的地图和指南针算子中速度和位置更新方程表示如下:

$$\boldsymbol{M}_i^{k+1} = \alpha \boldsymbol{M}_i^k + (1-\alpha)\boldsymbol{V}_i^k \tag{23}$$

$$\boldsymbol{V}_i^{k+1} = \boldsymbol{M}_i^{k+1} + r(\boldsymbol{X}_g^k - \boldsymbol{X}_i^k) \tag{24}$$

$$\boldsymbol{X}_i^{k+1} = \boldsymbol{X}_i^k + \boldsymbol{V}_i^{k+1} \tag{25}$$

式中:$\alpha<1$ 为动量因子;$\boldsymbol{M}_i^{k+1}$ 为第 $k+1$ 次迭代中的动量效应。递归展开式(23),有

$$\begin{cases} \boldsymbol{M}_i^{N+1} = \alpha \boldsymbol{M}_i^N + (1-\alpha)\boldsymbol{V}_i^N \\ \boldsymbol{M}_i^{N+1} = \alpha^2 \boldsymbol{M}_i^{N-1} + \alpha(1-\alpha)\boldsymbol{V}_i^{N-1} + (1-\alpha)\boldsymbol{V}_i^N \\ \quad \vdots \\ \boldsymbol{M}_i^{N+1} = \sum_{k=0}^{N} \alpha^k (1-\alpha)\boldsymbol{V}_i^{N-k} \end{cases} \tag{26}$$

式中:N 为算法总迭代次数;α 为式(23)中的动量因子;$\boldsymbol{V}$ 为式(24)中鸽群中某只鸽子的速度。由于动量因子 $\alpha<1$,动量的分布方式更多地在当前速度上。随着迭代次数的增加,旧速度的系数与动量因子 α 共同累积,旧的速度值对动量 $\boldsymbol{M}$ 的贡献将降低,这会有效增强鸽群优化算法的搜索能力,同时防止鸽子被其历史速度加权相加而陷入局部最优值。另外,由于速度值 $\boldsymbol{V}$ 是迭代累积求和得到的,不需要额外的空间来存储速度的历史值。

本文提出的 EM-PIO 算法解决多无人机协同目标防御问题的具体实现流程如图 1 所示。

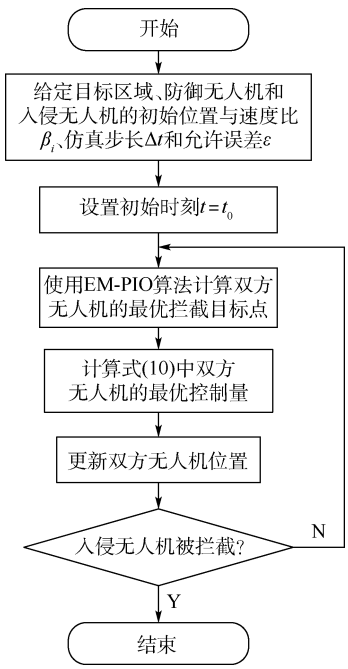


图 1　EM-PIO 算法解决多无人机协同防御问题实现流程

Fig. 1　Procedure of coordinated target defense with multi-UAVs cooperative using EM-PIO algorithm

4　仿真对比实验

本文设防御无人机的数量 $M=2$。入侵无人机和防御无人机的初始位置分别为 $\boldsymbol{x}_{E0}=(6,6,3)$,$\boldsymbol{x}_{P_{10}}=(5,4,2)$和 $\boldsymbol{x}_{P_{20}}=(3,5,3)$,防御无人机与入侵无人机之间的速度比 $\beta_1=1.1$ 和 $\beta_2=1.2$。防御无人机要防御的目标区域为球体,球心坐标 $\boldsymbol{x}_T=(3,3,2)$,半径 $r_c=1$ m。双方无人机位置、目标区域和约束曲面 $g_1(\boldsymbol{x})=0$ 和 $g_2(\boldsymbol{x})=0$ 的图像如图 2 所示。可以看出,$g_1(\boldsymbol{x})=0$ 与 $g_2(\boldsymbol{x})=0$ 相交形成一条曲线,双方无人机的最优目标点一定在曲线上。

同时,本节设计了仿真实验对比 EM-PIO 算法与 PSO、GA 算法。3 种算法的参数值如表 1 所示。

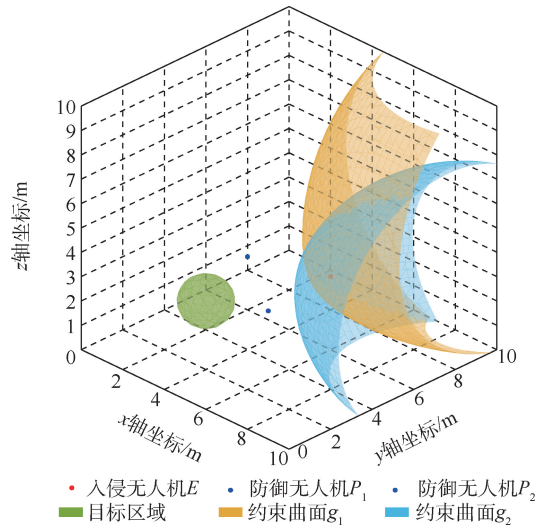


图 2 约束曲面示意图

Fig. 2 Schematic of constraint surface

表 1 EM-PIO、PSO 和 GA 算法参数

Table 1 Parameters of EM-PIO, PSO and GA algorithms

算法	参数	数值	描述
EM-PIO	N	20	鸽群中鸽子数目
	N_{c1}	100	地图和指南针算子迭代次数
	N_{c2}	120	地标算子迭代次数
	α	0.5	动量因子
PSO	N	20	种群粒子数目
	N_c	120	迭代次数
	ω	0.5	惯性权重
	c_1	0.6	社会学习因子
	c_2	0.4	个体学习因子
GA	N	20	种群个体数目
	N_c	120	迭代次数
	P_c	0.8	交叉概率
	P_m	0.05	变异概率

3 种算法的进化曲线如图 3 所示。由图 3 进化曲线可见，本文提出的 EM-PIO 算法具有更好的全局优化性能和收敛速度，可有效解决复杂态势下的多无人机约束最优化问题。

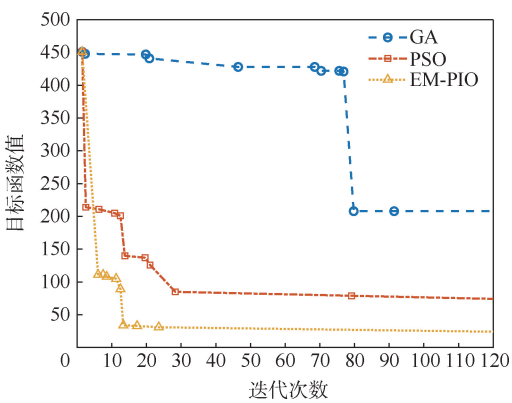


图 3 EM-PIO、PSO 和 GA 算法进化曲线对比

Fig. 3 Comparison of evolution curves of EM-PIO, PSO and GA algorithms

5 结 论

1）本文构造的多级非稳态罚函数可以通过逐渐增加惩罚值有效地控制遗传算法收敛速度，确保获得可行解。

2）提出的 EM-PIO 算法通过在地图和指南针算子处引入动量因子，增强了鸽群优化算法的搜索能力，同时防止被历史速度加权相加而陷入局部最优值。

3）提出的 EM-PIO 算法具有较为优异的搜索性能，在仿真实验中相比于 PSO 和 GA 算法具有较快的收敛速度且具有更好的全局优化性能。

本文在分析无人机协同目标防御问题时，只考虑了入侵无人机数量为 1 架时的情况，后续将进一步研究入侵无人机数量大于 1 的情况。

参考文献（References）

[1] 段海滨，邱华鑫. 基于群体智能的无人机集群自主控制[M]. 北京：科学出版社，2018：1-10.
DUAN H B，QIU H X. Unmanned aerial vehicle swarm autonomous control based on swarm intelligence[M]. Beijing：Science Press，2018：1-10(in Chineses).

[2] 段海滨，何杭轩，赵彦杰，等. 2021 年无人机热点回眸[J]. 科技导报，2022，40(1)：215-227.
DUAN H B，HE H X，ZHAO Y J，et al. Review of technological hotspots of unmanned aerial vehicle in 2021[J]. Science & Technology Review，2022，40(1)：215-227(in Chineses).

[3] ZHU K，HAN B，ZHANG T. Multi-UAV distributed collaborative coverage for target search using heuristic strategy[J]. Guidance，Navigation and Control，2021，1(1)：2150002.

[4] HUO M Z，DUAN H B，FAN Y M. Pigeon-inspired circular formation control for multi-UAV system with limited target information[J]. Guidance，Navigation and Control，2021，1(1)：2150004.

[5] VON MOLL A，GARCIA E，CASBEER D，et al. Multiple-pursuer，single-evader border defense differential game[J]. Journal of Aerospace Information Systems，2020，17(8)：407-416.

[6] GARCIA E，CASBEER D W，VON MOLL A，et al. Multiple pursuer multiple evader differential games[J]. IEEE Transactions on Automatic Control，2020，66(5)：2345-2350.

[7] SHISHIKA D，KUMAR V. Perimeter-defense game on arbitrary convex shapes[EB/OL]. (2021-05-01)[2022-04-28]. https：//arxiv. org/abs/1909. 03989.

[8] SINHA A，KUMAR S R，MUKHERJEE D. Three-agent time-constrained cooperative pursuit-evasion[J]. Journal of Intelligent & Robotic Systems，2022，104(2)：1-27.

[9] WANG J，LI G，LIANG L，et al. Pursuit-evasion games of multiple cooperative pursuers and an evader：A biological-inspired perspective[J]. Communications in Nonlinear Science and Numerical Simulation，2022，110：106386.

[10] GANJEFAR S，TOFIGHI M. Dynamic economic dispatch solu-

tion using an improved genetic algorithm with non-stationary penalty functions[J]. European Transactions on Electrical Power,2011,21(3):1480-1492.

[11] KONNOV I V. Application of penalty methods to non-stationary variational inequalities[J]. Nonlinear Analysis: Theory, Methods & Applications,2013,92:177-182.

[12] KATOCH S,CHAUHAN S S,KUMAR V. A review on genetic algorithm:Past, present, and future[J]. Multimedia Tools and Applications,2021,80(5):8091-8126.

[13] PARSOPOULOS K E,VRAHATIS M N. Particle swarm optimization method for constrained optimization problems[J]. Intelligent Technologies-Theory and Application:New Trends in Intelligent Technologies,2002,76(1):214-220.

[14] DUAN H B, QIAO P. Pigeon-inspired optimization: A new swarm intelligence optimizer for air robot path planning[J]. International Journal of Intelligent Computing and Cybernetics, 2014,7(1):24-37.

[15] JIVANI U N, VAIDYA O B, BHATTACHARYA A, et al. A swarm variant for the schrodinger solver[C]// Proceedings of 2021 International Joint Conference on Neural Networks (IJCNN). Piscataway:IEEE Press,2021:1-8.

Coordinated target defense for multi-UAVs based on exponentially averaged momentum pigeon-inspired optimization

DUAN Haibin[1,*], TONG Bingda[1], LIU Jichuan[2,3]

(1. School of Automation Science and Electrical Engineering, Beihang University, Beijing 100083, China;
2. The 54th Research Institute of China Electronics Technology Group Corporation, Shijiazhuang 050081, China; 3. School of Electronic Engineering, Xidian University, Xi'an 710071, China)

Abstract: This paper proposes a multi-unmanned aerial vehicle (UAV) cooperative target defense method based on exponentially averaged momentum pigeon-inspired optimization (EM-PIO). Firstly, the multi-UAV cooperative target protection system in three-dimensional space is modeled. The surface constraint equation of the UAV-dominated area and the optimal control input of UAVs are obtained. Secondly, in order to address the constrained optimization problem, the multi-level penalty function method is used to generate the objective function of the optimization algorithm. In addition, an EM-PIO algorithm is proposed to solve the optimal point. Comparative experiments with the genetic algorithm (GA) and particle swarm optimization (PSO) are conducted. The simulation results show that the EM-PIO method can solve the multi-UAV cooperative target defense problem more effectively.

Keywords: unmanned aerial vehicle (UAV) confrontation; pigeon-inspired optimization (PIO); target defense; differential games; saddle point strategy

Received: 2022-04-30; **Accepted**: 2022-05-18; **Published online**: 2022-06-06 14:40

URL: kns.cnki.net/kcms/detail/11.2625.V.20220602.1834.001.html

Foundation items: Science and Technology Innovation 2030-Key Project of "New Generation Artificial Intelligence" (2018AAA0102303); National Natural Science Foundation of China (U20B2071,91948204,T2121003,U1913602)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0012

重油航空活塞发动机燃油喷射技术

丁水汀[1], 邵龙涛[2], 赵帅[1], 朱锟[3], 杜发荣[1], 周煜[1,*]

(1. 北京航空航天大学 航空发动机研究院, 北京 100083;
2. 北京航空航天大学 能源与动力工程学院, 北京 100083; 3. 中国航空发动机研究院, 北京 101300)

摘 要: 燃油喷射技术作为航空活塞发动机的关键技术之一,其结构形式及特性直接影响航空活塞发动机的整机性能,随着通用航空对发动机性能要求的不断提升,燃油喷射技术的发展也在不断延续。综述了当前各种重油航空活塞发动机采用的燃油供给系统,对其适用的经典机型进行了梳理,总结了各种燃油供给系统的优缺点。阐述了研究重油喷射技术采用的前沿理论、仿真计算及试验方法,分析了研究重油雾化过程存在的难点,给出了研究重油喷射技术的相关建议。提出了未来重油航空活塞发动机发展思路和关键技术,重点希望对重油航空活塞发动机的正向研制提供技术指导,推动中国通用航空的可持续发展。

关 键 词: 航空活塞发动机; 重油; 喷射技术; 二冲程; 负碳生物燃料; 控制策略
中图分类号: V233
文献标志码: A **文章编号**: 1001-5965(2022)09-1630-13

燃油喷射系统(fuel injection system, FIS)是确保重油航空活塞发动机(heavy fuel aircraft piston engine, HF-APE)健康、稳定、最大限度发挥动力性能,保证飞机正常工作的关键系统。随着重油航空活塞发动机技术水平的不断提升,燃油喷射系统也由简单到复杂,并且由机械控制向电子控制过渡。

最近几十年来,航空汽油在燃烧特性、雾化特性及低温流动性方面的优势使其一直作为航空活塞发动机的主要燃料[1],但航空汽油饱和蒸汽压高、闪点低、挥发性强的特性又使其常温下遇明火容易发生爆炸,给燃料储运和使用方面带来了很大的安全隐患[2],燃料管理难度加大,尤其在军用领域的使用受到限制,欧美国家对于含铅燃料的禁用时间已经有了明确的规定,并且在军舰等军事设备上严禁配备航空汽油[3]。

相比于航空汽油,航空重油(煤油、柴油)具有黏度大、挥发性差、闪点高及安全性高的特性,全世界范围都有广泛供应,除此之外,压燃重油采用更高的压缩比,得到更高的燃烧效率及更低的油耗[4],具有缓解航空汽油等轻质燃油紧缺、提高安全系数等优点,在通用航空领域具有广阔的应用前景[5]。除此之外,重油在军方后勤保障系统中有深厚的根基,因此使用航空煤油或者柴油的航空器更受军方青睐。但是由于重油燃料黏度较大、蒸发性差、燃油雾化效果不理想的限制[6],导致部分重油航空活塞发动机在起动阶段必须采取预热、引燃、高能点火等辅助手段,且起动过程转速不稳定,发动机易熄火[7];发动机高速运转时,燃油蒸发时间不足,排气逸出损失大[8];随着飞行高度的增加,航空活塞发动机的进气温度下降,压缩终了缸内温度低,燃油蒸发速率降低,混合气质量下降,有效热效率下降,有效燃油消耗率上升,碳烟排放显著增加[9]。

收稿日期: 2022-01-12; **录用日期**: 2022-02-16; **网络出版时间**: 2022-04-29 08:57
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220428.1555.003.html
基金项目: 国家重点基础研究发展计划 (2018YFB010400); 国家自然科学基金 (51775025,51775013)
* **通信作者**. E-mail: zybuaa@hotmail.com

引用格式: 丁水汀, 邵龙涛, 赵帅, 等. 重油航空活塞发动机燃油喷射技术[J]. 北京航空航天大学学报, 2022, 48(9): 1630-1642.
DING S T, SHAO L T, ZHAO S, et al. Fuel injection technology of heavy fuel aircraft piston engine[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1630-1642(in Chinese).

本文对当前重油航空活塞发动机采用的燃油喷射系统进行梳理;对研究燃油喷射技术的相关理论、仿真模拟及试验方法进行总结;对二冲程缸内直喷燃油喷射技术、先进燃油喷射的控制策略、燃油喷射与燃烧室的匹配、负碳燃料燃油喷射技术等进行前瞻性探索;对重油航空活塞发动机的发展起到参考与借鉴作用。

1　重油航空活塞发动机燃油喷射系统的研究进展

虽然目前航空活塞发动机大多为点燃式航空汽油活塞发动机,但以柴油和航空煤油等重油为燃料的压燃式发动机价格更便宜、热效率更高、系统结构简单、安全性好,逐渐成为航空活塞发动机发展的新趋势[10]。

重油航空活塞发动机燃油喷射系统历经百年的发展,形成了以进气道燃油喷射(port fuel injection,PFI)系统、机械燃油直喷(mechanical fuel direct injection,MFDI)系统、高压共轨燃油喷射(high pressure common rail,HP-CR)系统及空气辅助喷射(air assisted direct injection,AADI)系统为代表的重油航空活塞发动机燃油喷射系统[11]。

1.1　进气道重油喷射系统

进气道重油喷射系统通过预热发动机进气道与曲轴箱、改进化油器、提高燃油温度的方式来加快重油的蒸发速度,重油以油气混合气的形式进入气缸。采用高能火花塞提高点火能量,使重油能被点燃。由于进气道重油喷射需要电辅助预热及更多的点火能量,因此多用于小型重油航空活塞发动机。进气道重油喷射的原理如图 1 所示。

国外率先开展进气道重油喷射系统研究的单位有:德国 3W 公司,其开发的 3W-157xiB2HFEFi 发动机采用进气道喷射和电辅助加热措施燃烧 JP8 航空煤油,该系统使得发动机结构异常复杂[12];美国罗切斯特理工学院的 Sonex 燃烧系统采用电加热气缸盖,改进化油器的方式将二冲程汽油机改烧 JP8 航空煤油[13]。国内,北京交通大学的宁智教授团队对进气道喷射的二冲程涡轮增压发动机进行了仿真分析,分别从排气能量、排气背压、扫气系数、逃逸率及匹配工作点等方面,对进气道喷射二冲程发动机采用废气涡轮增压器的影响因素进行了量化分析[14]。南京航空航天大学对二冲程航空汽油机 3203E 进行了改进,其采用 PTC 元件进行电辅助加热的方式来提高 RP-3 航空煤油的蒸发效果[8]。

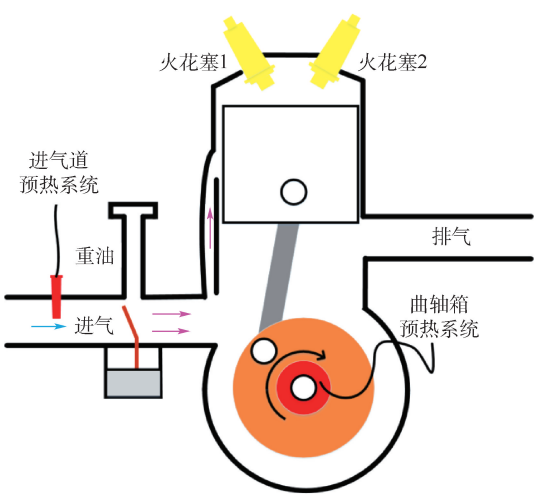


图 1　进气道重油喷射原理

Fig. 1　Schematic of port fuel injection

虽然众多科研人员对进气道重油喷射技术进行了研究,但是从原理上分析,进气道重油喷射使进入气缸燃烧室的部分混合气未经燃烧直接排出气缸,造成发动机耗油率上升的缺点不可规避;重油航空辛烷值较低,点燃会造成发动机的明显爆震[15],容易造成发动机性能迅速恶化,引起发动机机体、零部件金属软化等一系列严重问题。

1.2　机械燃油直喷系统

机械燃油直喷系统直接由凸轮轴驱动柱塞对燃油进行加压,由发动机附属机械机构完成燃油喷射和调节,采用机械喷油器,喷油泵和机械喷油器之间采用一小段高压油管完成连接。图 2 为机械燃油直喷技术的原理。

美国 XRDI 公司研制的小型二冲程煤油航空活塞发动机采用机械燃油喷射系统 MCDI(mechanical compression direct injection),功率达到了 12.5 kW[16]。法国 SMA 航空活塞发动机采用博世的分布式单体泵管嘴供油系统,转速为 2 200 r/min,塞斯纳和乌克兰飞机制造商都装配了 SR305-230E 发动机[17]。美国 Deltahawk 公司开发的系列发动机 DH-160A4、DH-180A4、DH-200A4 均采用机械燃油供给系统,转速为 2 000 r/min,采

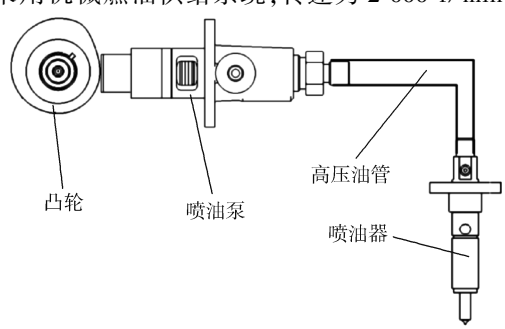


图 2　机械燃油直喷技术原理

Fig. 2　Schematic of mechanical fuel direct injection

用机械燃油供给系统的好处是:飞机失电之后,发动机依然可以保持部分功率运行,直至安全降落[18]。英国 WAM 系列的发动机采用机械燃油喷射系统,其特点是使用了预燃室,并采用多通道喷射进主燃室进行间接燃烧,发动机转速为 2 750 r/min[19]。北京航空航天大学航空微小型动力团队自主研发的系列二冲程重油航空活塞发动机(见图 3)均采用机械燃油直喷系统,发动机转速为 2 400 r/min。

机械泵管嘴系统虽然装备了多型号重油航空活塞发动机,但其多需要配置机械调速器,调节精度较差,调速器结构相对笨重,对重油航空活塞发动机的轻量化配置不利,其采用的机械式喷油器喷油压力相对较低,大多约为 20 MPa,不利于重油充分地雾化燃烧。

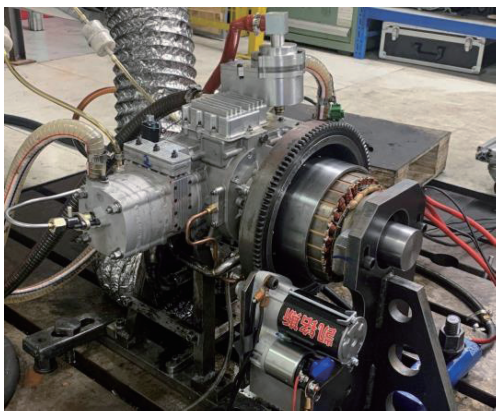


图 3 二冲程重油航空活塞发动机

Fig. 3 Two-stroke heavy fuel aircraft piston engine

1.3 高压共轨燃油喷射系统

航空活塞发动机高压共轨燃油喷射系统是由电控单元(ECU)、传感器和执行器 3 部分组成,附件包括油轨、压力传感器和压力控制阀,高压油泵不断将高压燃油送入共轨管,并维持轨压;共轨管起到储存燃油并保持油压、消除燃油压力波动的作用[20],各种传感器将测定的实时运行参数与计算机中设定 MAP 图进行比较,计算出最佳喷油定时和喷油脉宽,精确地控制电控喷油器将燃油喷入燃烧室,使发动机在各种工况下都能获得最佳浓度的混合气,其原理如图 4 所示。作为电控高压共轨燃油喷射系统的执行器,电控喷油器主要由喷油嘴、控制活塞、控制油量孔和控制电磁阀组成,采用压力-时间的计量方式,用高速泄油电磁阀或压电晶体执行器控制其喷射过程。

最早应用高压共轨燃油喷射系统的航空活塞发动机是 Thielert Aircraft Engines(TAE)公司的 Centurion 系列发动机,并大量装备在钻石飞机公司的飞机上[21]。Austro 航空活塞发动机公司的 AE300 同样采用高压共轨燃油喷射系统[22],并在 2009 年通过 EASA 认证,现在役数量大约 3 000 台[4]。国内目前对于高压共轨燃油喷射系统也逐渐从理论研究向工程应用迈进。陆军军事交通学院研究了 RP-3 航空煤油在高压共轨柴油机中的应用[23]。江苏大学对高压共轨柴油机过渡工况采用的转矩控制策略进行了研究,并进行了相关控制软件的开发[24]。北京理工大学的孙柏刚等[25]对高压共轨燃油供给系统的压力波动特性进行了试验研究,得到了压力波传播速度、波动过程幅值、喷油压降下降幅值、周期及衰减时间的变化规律。产业化方面,以成都威特电喷有限责任公司、龙口龙泵燃油喷射有限公司、重油高科电控燃油喷射系统(重庆)有限公司与南岳电控(衡阳)工业技术股份有限公司为代表的公司也在高压共轨项目上投入巨资,目前均有批量生产的产品面世。

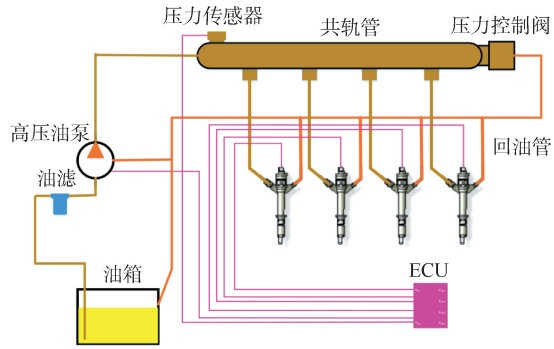


图 4 高压共轨燃油喷射系统原理图

Fig. 4 Schematic of high pressure common rail system

从原理上讲,电控高压共轨燃油供给系统属于恒压式供油,喷油规律为矩形,相对三角形喷油规律,其产生的爆发压力高,燃烧相对粗暴,不利于降低 NO_x 排放。市场对高压共轨技术的价格难以接受,推广高压共轨技术在重油航空活塞发动机上应用压力较大[26]。除此之外,受到结构的限制,高压共轨系统布置在结构紧凑的二冲程发动机中较为困难[27],并且需要频繁地检查发动机与螺旋桨之间的变速箱,降低了航空活塞发动机的大修间隔[28]。在航空低温、电磁干扰等极端环境下,电控高压共轨燃油供给系统喷油器、溢流阀等执行元件的可靠性同样受到挑战。

1.4 空气辅助喷射系统

空气辅助喷射系统利用高压压缩空气从喷孔喷出时的气动力克服燃油表面张力来雾化重油,使重油在相对较低的喷射压力下获得较小的喷雾粒径[29]。空气辅助喷射系统主要由高压气源、气

压调节阀、油箱、油泵、油压表、油气调节阀、油嘴、气嘴和 ECU 等组成,其工作原理如图 5 所示。

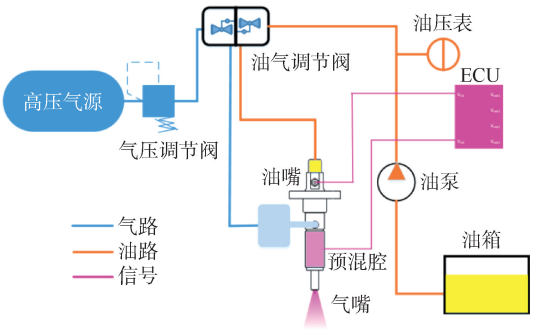


图 5　空气辅助喷射系统示意图

Fig. 5　Schematic of air assisted direct injection

空气辅助喷射系统对燃料种类不敏感,适用于多燃料压燃航空活塞发动机;喷雾特性受缸内压力变化影响较大,贯穿距离随气缸内压力升高而减小,有利于实现混合气分层;对燃油适应性较强,对于黏度较大的重油燃料仍可以保证良好的雾化效果,燃油计量由燃油喷嘴实现,喷油正时由混合气喷嘴执行,二者相互独立[30];系统的功耗远小于高压共轨燃油喷射系统。

20 世纪 90 年代,以澳大利亚 Orbital 公司为代表的多家研究机构先后提出了低压空气辅助喷射系统,为二冲程缸内直喷技术带来了革命性的变革[31]。国内北京理工大学的赵振峰等[32]在一台二冲程四缸直喷重油发动机上研究了空气辅助喷射系统的混合气形成。高宏力、张付军教授等针对空气辅助喷射技术开展了关于喷雾特性、喷射时刻和喷射持续期等多项研究[33]。南京航空航天大学将传统的二冲程进气道喷射发动机改造为空气辅助缸内直喷煤油发动机,并对其控制策略进行了研究[34]。

但是,由于空气辅助喷射系统喷油压力低,对实际发动机缸内环境比较敏感,过高的缸内压力将影响喷雾的雾化质量。燃油供给系统需要配备压缩空气,以及其他相关附件,结构相对复杂,当前多应用于小型二冲程航空发动机[35-36]。

2　重油航空活塞发动机燃油喷射技术的研究进展

2.1　燃油喷射雾化理论研究进展

燃油雾化特性(液滴破碎、蒸发、碰壁及混合气的形成)对重油发动机的燃烧、排放均有重要影响,重油航空活塞发动机研究过程的难点之一便是缸内雾化过程。

1878 年,Rayleigh 最早提出关于液体射流破碎的理论,之后 Bergwerk 认为空化作用才是雾化产生的主要因素。1995 年,Li[37]提出了针对具有三维扰动的无黏性气体介质中黏性液体射流的线性稳定性分析。随后,大批学者从不同的角度考虑建立了多种燃油雾化的模型。1999 年,Arine 等[38]基于两相流的欧拉法建立了液体雾化模型。Senecal 等[39]从黏性液体高速运动的角度建立了雾化模型。2003 年,Koo[40]从超剪切喷嘴的液体雾化角度出发,建立了燃油雾化的多维模型。Iyer和 Abraham[41]从缸内雾化的环境考虑,建立了欧拉液体和欧拉气体同时存在的双流体模型。2007 年,法国的 Demoulin 等[42]在 Borgh 提出的模型基础上,建立了用来描述液滴初始破碎过程的模型。目前,被认为较为合理的是由 Castleman[43]提出的空气动力干扰理论,也被认为是最有发展前途的雾化机理的解释之一。但是随着喷射压力的提高,当前国际主流观点认为[44-46],当喷油压力达到 300 MPa 时,即为超高压状态,该状态下的燃油雾化和燃烧性能的研究非常匮乏。从目前的研究来看,燃油喷雾超过声速之后,喷雾前端会产生激波,激波会对燃油的雾化及燃烧过程产生一定的影响[47]。

上述有关燃油喷射与雾化机理的研究都有一定的理论基础,同时又是在一定的假设前提下提出的,具有一定的局限性,到目前为止还没有一套完整的理论能够对燃油雾化做出完美的解释,对于实际情况下的燃油喷射,将不同的理论有机结合是具有实际工程意义的。

2.2　燃油喷射仿真模拟研究进展

重油航空活塞发动机燃油喷射系统中实际的喷雾、蒸发、扩散、燃烧与传热等过程十分复杂,加上航空活塞发动机运行时的循环变动,要在实际运行中进行参数控制并获得准确可靠的实验数据并不现实[48]。因此,集成了电磁学、热力学和静力学等多个学科的现代燃油喷射仿真软件对燃油喷射特性的研究起到了关键作用[49]。计算机辅助仿真用于开发和优化匹配新型燃油喷射系统,可以大量减少试验次数,节省资源和时间,其在系统变参数分析上具有的低成本、短开发周期的优越性是试验研究无可比拟的,同时仿真分析还可以发现一些在试验中无法观察到的新现象和新规律。

用于发动机燃油喷射及缸内流动的仿真模拟软件有 KIVA-3V 程序、ANSYS 公司的 FLUENT、里卡多的 VECTIS、奥地利 AVL 公司的 Fire、英国 Adapco 公司的 STAR-CD、英国 Gamma 公司的

GT-Power、Ricardo 公司的 Wave、美国西南研究院的 VIPRE 等商业软件[50]。

意大利摩德纳大学的 Mattarelli 等[51]对一台额定功率高达 110 kW 的二冲程直喷柴油发动机进行了数值模拟研究,仿真结果为发动机燃油供给系统的优化提供了理论依据。英国斯塔福德郡大学的 Hooper 等[52]使用 Wave 软件,对多燃料的活塞式发动机进行了 CFD 仿真研究,并用台架试验对仿真结果进行了验证。西班牙瓦伦西亚理工大学的 Salvador 等[53]采用仿真手段研究了生物柴油对共轨系统电磁阀动态特性的影响。La 等[54]在一台空气辅助燃油喷射的二冲程发动机上进行了标准化台架试验,该标准化台架试验用于校准该发动机的一维 CFD 模型。Sener 等[55]利用试验和三维仿真的方式分析了不同喷射压力和喷雾锥角下的燃烧过程,并与 DSCS 燃烧室和 MSCS 燃烧室进行了匹配标定。

国内各单位同样对燃油喷射技术进行了仿真研究。天津大学的胡春明等[56]利用发动机建模软件 AMESim 建立了发动机模型,用 VPSO-Elman 空燃比预测模型在 MATLAB/SimuLink 中建立了 VPSO-Elman 空燃比预测模型控制系统,对航空活塞发动机瞬态空燃比控制进行了研究。北京航空航天大学的王振宇等[57]建立了某重油航空发动机燃油供给系统的数学模型,用于仿真燃油供给系统内的非定常流动,通过仿真计算得到了该系统的压力波动特性和喷油规律。南京航空航天大学的陈林林[58]利用 GT-Power 软件建立了煤油发动机工作循环数值模型,通过分析发动机主要结构参数,调整参数(进气压力、空燃比、点火提前角)对煤油发动机性能的影响,为煤油发动机的参数优化匹配提供理论基础。北京交通大学的李长胜[59]对 Dr. Schrick 公司研发的 TKDI600 发动机的混合气形成及燃烧过程进行了仿真分析,应用 AVL Fire 软件建立了 TKDI600 发动机的仿真计算模型,研究并分析了发动机缸内流场、燃油喷射、混合气形成和燃烧过程的特性,并探讨了发动机的燃油碰壁现象和油膜形成机理。此外,石允[60]应用 Fire 软件建立了 HS-700 发动机的三维 CFD 模型,研究了对重油混合气形成有明显影响的因素,并分析了各个影响因素及燃油喷射方式对缸内流场和混合气形成质量的影响规律。北京理工大学的刘波澜等[61]建立了空气辅助燃油喷射系统的仿真模型,研究了结构、运行及环境参数对该系统的影响,为空气辅助燃油喷射系统的研究提供了理论依据。

重油航空活塞发动机缸内燃油喷射是极其复杂的三维湍流运动,具有强瞬变、强压缩、强涡流和各向异性的特点,加之燃烧室的形状复杂、运动边界及循环变动,航空活塞发动机缸内湍流变得异常复杂。因此,建立可信的发动机缸内喷射优化仿真模型对燃油喷射系统研发具有十分重要的理论支撑与实践指导意义。

2.3 燃油喷射试验研究进展

试验研究是理论研究的前提和基础,对重油航空活塞发动机的技术进步起着巨大的推动作用。由于燃油喷射过程是一个瞬变过程,其试验研究对于测量设备与试验手段都有很高的要求[62]。重油航空活塞发动机燃油喷射系统的试验主要从宏观和微观 2 个角度进行。宏观的试验方法主要包括缸压采集方法、油压信号采集法[63];微观的试验方法主要是对燃油喷射后的喷雾特性[64-65],如贯穿距离、喷雾锥角、液滴速度等进行研究,微观研究采用的技术主要有高速摄影、纹影法、阴影法、相位多普勒技术(phase Doppler anemometer, PDA)和平面激光粒径测试技术(laster sheet drop size, LSD)等,试验装置有光学发动机、快速压缩机和定容弹等[66]。

缸压采集系统采用缸压传感器和电荷放大器,把缸压信号与曲轴相位信号同步接入燃烧分析仪,通过燃烧分析仪对发动机工作过程的数据进行实时采集、分析及存储。通过分析缸压,得到滞燃期与燃烧持续期之间的相关性,由发动机性能表现推导出燃油喷射与燃烧之间的对应关系。北京理工大学使用 Dewetron-5000 燃烧分析仪采集高压油管泵端压力、驱动电路和喷油器针阀升程信号,利用 Kistler4067BB2000 型油压传感器和电荷放大器采集高压油管中的燃油压力,使用 EFS8246 型单次喷射仪采集喷油器循环喷油量[67],对单体泵系统的供油特性进行了试验研究[68],其实验原理如图 6 所示。天津大学的台架试验系统主要由试验发动机、电力测功机、燃油供给系统、冷却系统、数据采集系统及发动机控制系统等组成。在一台自主研发的单缸航空活塞发动机上,针对起喷转速、燃油温度、点火能量对航空煤油发动机冷起动性能的影响开展了试验研究,并对燃烧特性进行了分析[69]。

燃油供给系统的循环供油量主要通过瞬时油量测量仪进行,其试验原理如图 7 所示。

对单一喷油器瞬时油量的测量与高压油泵的凸轮轴转速同步,喷出的燃油完全进入到可变容积的盛油腔内。上海交通大学的韦雄等[70]采用

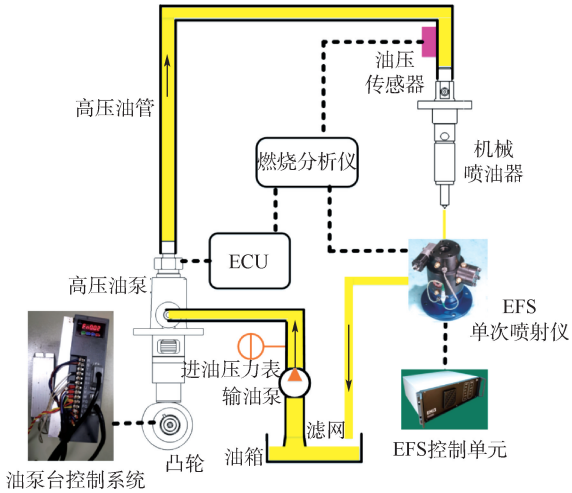


图 6　油泵试验台系统示意图

Fig. 6　Schematic of diesel pump test bench system

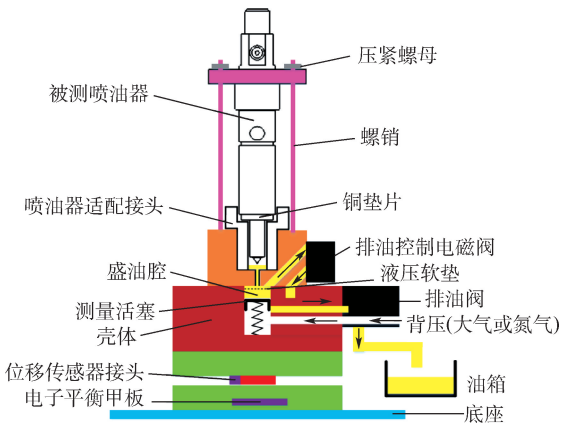


图 7　EFS-EMI2 机械结构示意图

Fig. 7　Schematic of EFS-EMI2 mechanical structure

该测量设备对电控单体泵供油量的精确测量与一致性进行了分析。洛阳拖拉机研究所的王睿等[71]采用 EFS 试验台对共轨喷油器的性能进行了测试试验。北京理工大学的严明等[72]采用 EFS 试验台研究了起喷压力对电控单体泵喷油量的影响机理。

可视化试验系统主要由燃油供给、燃油喷射、图像采集及数据处理系统组成。试验中,通过输油泵给燃油供给系统输送燃油,采用涡轮流量计和压力传感器来监测进油压力和流量,喷射后的燃油用输油泵送回油箱。燃油喷射过程中,使用 LED 光源照亮喷油器的透明油嘴,高速相机在定容弹的另一侧进行拍摄。高速相机可以拍摄到喷嘴内部的空穴流动及喷孔喷出的油雾,随后将数字图像传输到计算机中进行后续的分析和处理,其试验系统示意图如 8 所示。

上海交通大学采用由定容燃烧弹、预混容器、高压共轨燃油喷射系统、控制系统、数据采集系统和 CMOS 高速摄影系统组成的蒸发喷雾液相贯穿

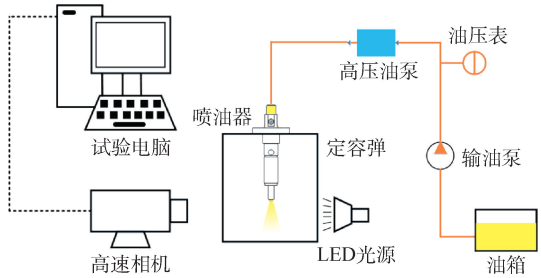


图 8　喷雾可视化试验系统示意图

Fig. 8　Schematic diagram of spray visual test system

距光学测试系统,研究了喷射压力、环境温度及环境密度对柴油蒸发喷雾液相贯穿距发展的影响,为柴油机喷雾燃烧系统的设计优化提供了参考[73]。中南大学能源科学与工程学院的动力试验室利用 GS-1000 型高压共轨燃油喷射试验台和可视化耐高压容器,通过 Motion Pro X-3 高速摄影,采用频闪照相的方法对图像进行定量分析,提高了图像分析的效率和精度,直观地对燃油雾化特性的影响参数进行了分析[74]。西安交通大学的魏衍举等[75]采用 Phantom Miro eX4 型高速摄像机利用背光法拍摄缸内柴油油束的混合过程,对强涡流场中柴油喷雾扩散特性的影响参数进行了研究。北京理工大学的吴晗等[76]采用背光成像和阴影技术,研究了空气辅助喷射系统煤油喷雾特性。南京航空航天大学利用 Hotron 公司的 FASTCAMSA1.1 高速摄影机对不同工况喷雾过程进行高速拍摄,研究了不同操作工况燃油喷雾的发展形态[77]。

3　未来重油喷射关键技术及发展方向

当前,重油航空活塞发动机的研发大多还停留在基本性能的实现阶段,对于控制策略的合理性、电磁系统的可靠性、燃油喷射技术的适应性等暂未做深入的研究,并且随着“碳中和”“碳达峰”概念在资本市场的爆发,负碳生物燃料等可持续燃料的应用也成为通用航空发展的大方向[78]。

3.1　二冲程缸内直喷燃油喷射技术

当前,压燃式重油航空活塞发动机的功重比多分布于 1 kW/kg 左右,并且产品多以较成熟的四冲程发动机改制为主。二冲程点燃式发动机虽能满足部分通航飞机的巡航动力要求,但是其工作原理存在扫气损失及过后排气问题,未燃碳氢化合物的排放量高,点燃重油存在不同程度的爆震问题[16]。

二冲程压燃式航空活塞发动机没有点火系统,对标四冲程,其零件数量少、功重比大、超载能

力强、过量空气系数大、对油气混合比例的变化不敏感,发生空中停车的概率极小。因此,二冲程压燃式重油航空活塞发动机是未来主要发展方向。二冲程压燃式航空活塞发动机性能提升的关键是更短的时间内在缸内形成均匀的混合气,得到适用于二冲程压燃式航空活塞发动机的喷油规律,减少未燃碳氢化合物的排放,这给燃油喷射系统提出了极大的挑战[79]。对于直喷系统,缸内直喷喷油器的设计主要考虑其孔数、孔径、喷孔锥角及分布等方面对缸内混合气形成的影响[80]。喷油器作为燃油喷射系统的核心部件,对发动机的性能有至关重要的影响。发动机转速高、功率大、爆压高的特点,以及航空活塞发动机紧凑的空间,都给喷油器的设计增加了很大的难度[81]。高速电磁阀技术的发展给电控缸内直喷技术的实现提供了保证,虽然当前的电控高压共轨燃油供给系统已经在部分重油航空发动机上应用,但是其大部分是车用发动机燃油喷射系统改进而成,更多的是为四冲程发动机定制,其响应速度、喷射频率、重量、布局与可靠性并不适用于二冲程航空活塞发动机的需求。未来依托结构优化、新型材料、智能诊断、故障保护和冗余设计等技术的进步,正向设计出布局紧凑、质量轻、能持续稳定工作在高负荷区间、适用于低温环境的缸内直喷脉冲式电控燃油喷射系统是提升重油航空活塞发动机性能的一个关键[82]。

3.2 燃油喷射与燃烧室匹配技术

航空活塞发动机燃烧室内燃油喷射、缸内气流组织及燃烧室形状的匹配需要考虑到航空活塞发动机的性能需求及外界环境的变化[83]。航空活塞发动机转速较高,采用直喷技术后,要在极短的时间内使燃油与气流在燃烧室内混合形成可燃混合气,在进行直喷设计时,必须考虑影响缸内混合气形成的参数,重视燃油喷射参数与燃烧室形状的合理匹配,以及缸内气体流动对雾化效果的影响;还要考虑到海拔高度变化所引起的外界环境参数的变化,随着海拔升高,进气压力下降,环境温度下降,导致缸内进气量下降,进而气缸中的流场强度和扫气效率均下降,此时航空活塞发动机燃油喷射与燃烧室壁面碰撞的现象比普通内燃机强烈得多,撞壁重油的蒸发及壁面堆积重油的蒸发在蒸发总油量中比例变大,对缸内燃烧质量有重要的影响[84]。由于结构限制无法进一步提高压缩压力和压缩比,如何保证二冲程重油航空发动机的油气混合、燃烧效率,成为航空活塞发动机提升动力性能的关键。研究适用于重油航空活塞发动机的喷射参数与燃烧室匹配技术成为研究重油航空活塞发动机的一个重要切入点。

3.3 先进的燃油喷射控制策略

电控燃油喷射系统对发动机控制的优劣主要取决于对 ECU 的控制策略[85]。因此,对二冲程重油活塞发动机喷油策略进行研究,具有重要的现实意义。重油航空黏度大,蒸发性差,在冷机状态下雾化效果差,混合气形成困难,冷起动性能差,这与重油航空发动机高安全性、高稳定性的适航准则背道而驰;航空活塞发动机飞行时高度变化较快,进入发动机气缸内的氧含量变化也较快,并且随着海拔高度的上升,大气压力不断下降,单位体积内的氧含量也不断下降,因此采用与车用发动机相同的控制策略显然不能满足航空活塞发动机的性能需求。

当前,关于航空活塞发动机控制策略的研究主要集中在航空汽油活塞发动机的 3 种控制策略,即基于 MAP 图的开环、怠速工况中的闭环及基于模型的空燃比控制策略[86]。压燃式重油航空活塞发动机与点燃式不同,发动机不能直接调节进气量,只能根据实时转速、转矩调节循环供油量,而进气量和循环供油量之间是随动变化的,当前针对压燃式航空活塞发动机的控制策略研究很少。采用先进的供油策略使缸内空燃比或过量空气系数在一个合理的阈值内波动,保证良好的燃烧过程,从而更好地控制发动机,这是未来重油航空活塞发动机研究的热点。

3.4 适用于负碳燃料的燃油喷射系统

在燃料的原料来源、制备、输运和使用等过程中使用碳捕获、碳封存和低碳高效利用等技术,使其全生命周期总碳排放小于零的燃料称之为负碳燃料[87]。

近年来,通用航空保持着持续增长的态势,对航空燃料的需求与日俱增,燃用传统航油带来的碳排放逐年增加,为了保证通用航空的可持续发展,推进低碳革命,通航产业对碳排放的追求也达到了前所未有的高度[88]。为了降低碳排放,负碳生物燃料成为一种完美的重油替代品,其作为一种可再生的能源,来源广泛,尤其是可以用微藻作为原料进行负碳生物燃料的制备,具有生长快、廉价易得和含油量高的优点。负碳生物燃料替代部分传统燃料,从环境的角度来说,对节约化石能源、实现可持续发展有着重要意义。从航空业的需求上来说,是一项有价值有前景的课题。除此之外,如果对航空活塞发动机的燃油喷射过程实现智能化可调参数控制,就可以对燃烧过程、热效

率及燃油消耗率产生积极的影响。探索智能可调参数在柴油机燃烧过程和喷油之间的关系,建立燃烧模型并采用仿真或者试验手段进行验证,对于节能减排具有重大意义[89]。

现有的重油航空活塞发动机燃油喷射系统燃用负碳生物燃料后,由于负碳生物燃料的黏度、雾化效率、燃烧充分程度等与传统重油不同,容易引起发动机缸内的积碳、结焦,对发动机工作性能产生重要影响。开发适用于负碳燃料的航空活塞发动机燃油喷射系统对于推进通用航空的低碳革命而言非常关键[90]。

4　结　论

通用航空正快速向高端发展,将涡轴发动机、活塞发动机、纯电动系统及混合动力系统进行对比,从整个动力系统来看,涡轴发动机由于转速高需要配备减速器,虽然发动机质量较轻,但系统功重比优势不明显,且燃油消耗率较高。纯电动系统充电时间长、续航时间短、过放电会导致电池性能下降,高空低温时电池性能下降明显。混合动力系统结构相对复杂,成本高。因此,活塞发动机在短时间内仍占据主导地位[4,91]。重油航空活塞发动机又因其低油耗、高可靠性成为航空活塞发动机一个重要的分支,重油喷射技术作为重油航空活塞发动机的核心技术之一,对发动机的缸内燃烧过程具有重要影响。

本文对当前重油航空活塞发动机各类燃油供给系统的研究进展进行了总结,分析各类燃油供给系统的优缺点,对其适用的经典机型进行了梳理;将重油航空活塞发动机燃油喷射技术的研究进展从理论、仿真及试验的角度进行了分析,提出了研究重油航空活塞发动机燃油喷射技术的难点,对具有实际工程意义的研究方案提出了建议;综合通用航空对发动机动力性能、可靠性和环境友好性等方面的需求,对未来重油航空活塞发动机燃油喷射技术的发展方向提出了建议。本文的研究结论如下:

1) 重油航空活塞发动机的燃油供给系统具有多样性,匹配发动机时需要综合考虑功重比、可靠性、安全性及动力性能。进气道重油喷射系统和空气辅助喷射系统目前多适用于小型航空活塞发动机;基于车用发动机改制的航空活塞发动机多用高压共轨燃油喷射系统,正向研发的重油航空活塞发动机多用机械脉冲式燃油供给系统。

2) 重油喷射技术是重油在毫秒级的时间里经过一系列雾化发展成小液滴并与缸内气流混合的过程,这一过程的复杂程度很难用单一的方法进行研究,因此重油喷射技术的发展需要依托理论分析、仿真计算及大量的试验研究相互配合,这样能够缩短研发周期,节约成本,具有实际的工程意义。

3) 高功重比、高燃烧效率、低油耗及低污染是未来重油航空活塞发动机发展的必然趋势,因此,二冲程压燃的结构形式、缸内直喷与燃烧室匹配的技术方案、适用于高空特性的先进燃油喷射控制策略及适用于负碳燃料的燃烧系统是未来重油航空活塞发动机的发展方向。

未来应该抓住通用航空井喷式发展的局势,充分借鉴国内外先进的技术,以二冲程压燃式航空活塞发动机为研究关键,针对各类航空器的动力需求,对重油航空活塞发动机进行正向自主设计,全面推进通用航空动力系统的发展。

参考文献(References)

[1] CHEN L F, DING S R, LIU H Y, et al. Comparative study of combustion and emissions of kerosene (RP-3), kerosene-pentanol blends and diesel in a compression ignition engine[J]. Applied Energy, 2017, 203: 91-100.

[2] PICKETT B M, BURNETTE P F, WELLS S P, et al. Fire safety tests comparing synthetic jet and diesel fuels with JP-8[J]. Fire Safety Journal, 2011, 46(3): 89-95.

[3] 丁水汀,宋越,杜发荣,等. 航空重油活塞发动机发展趋势及关键技术分析[J]. 航空动力学报, 2021, 36(6): 1121-1136.
DING S T, SONG Y, DU F R, et al. Analysis on development trend and key technology of aircraft heavy fuel piston engine [J]. Journal of Aerospace Power, 2021, 36(6): 1121-1136 (in Chinese).

[4] 郑君. 通用航空活塞发动机现状及发展趋势探讨[J]. 内燃机与配件, 2020(19): 196-198.
ZHENG J. Discussion on current situation and development trend of general aviation piston engine[J]. Internal Combustion Engine & Parts, 2020(19): 196-198 (in Chinese).

[5] 张奇,杜发荣. 小功率航空活塞发动机重油技术进展[J]. 小型内燃机与车辆技术, 2014, 43(4): 81-85.
ZHANG Q, DU F R. Small aero-piston engine heavy fuel technology development summary [J]. Small Internal Combustion Engine and Vehicle Technique, 2014, 43(4): 81-85 (in Chinese).

[6] 杜春媛,田梦园,胡春明. 二冲程重油航空发动机技术发展分析[J]. 小型内燃机与车辆技术, 2020, 49(3): 92-96.
DU C Y, TIAN M Y, HU C M. Technical development analysis of two-stroke heavy fuel aero engine[J]. Small Internal Combustion Engine and Vehicle Technique, 2020, 49(3): 92-96 (in Chinese).

[7] HOOPER P. Experimental experience of cold starting a spark ignition UAV engine using low volatility fuel[J]. Aircraft Engi-

neering and Aerospace Technology,2017,89(1):106-111.

[8] 王在良.15 kW 级二冲程点燃式 PFI 煤油活塞发动机性能试验研究[D].南京:南京航空航天大学,2018:5.
WANG Z L. Experimental study on performance of 15 kW two-stroke ignited PFI kerosene piston engine[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2018: 5 (in Chinese).

[9] 沈颖刚,聂珂,徐劲松,等.不同海拔下压燃式航空活塞发动机的燃烧与排放特性[J].推进技术,2022,43(4):248-256.
SHEN Y G, NIE K, XU J S, et al. Combustion and emission characteristics of compression-ignition aero piston engine at different altitudes[J]. Journal of Propulsion Technology, 2022, 43(4): 248-256 (in Chinese).

[10] TSACH S, CHEMLA J, PENN D. UAV systems development in IAI-Past, present & future[C]//2nd AIAA "Unmanned Unlimited" Conference and Workshop & Exhibit. Reston: AIAA, 2003: 1-19.

[11] 王金华,邹季灿,吴小任,等.通过改进出油阀结构完善喷油泵性能[J].内燃机,2000(6):3-7.
WANG J H, ZOU J C, WU X C, et al. Improving the structure of fuel delivery valve to perfect the performance of fuel injection pump[J]. Internal Combustion Engines, 2000(6): 3-7 (in Chinese).

[12] 3W International GmbH. Heavy fuel system[EB/OL]. [2022-01-02]. https://3w-internatio nal. com/.

[13] DUDDY B J, LEE J, WALLUK M, et al. Conversion of a spark-ignited aircraft engine to JP-8 heavy fuel for use in unmanned aerial vehicles[J]. SAE International Journal of Engines, 2011, 4(1): 82-93.

[14] 尹君.小型航空进气道喷射二冲程涡轮增压发动机的仿真分析及改进研究[D].北京:北京交通大学,2014.
YIN J. Simulation analysis and improvement research for small aviation intake port injection two-stroke engine[D]. Beijing: Beijing Jiaotong University, 2014 (in Chinese).

[15] FALKOWSKI D T, ABATA D L, CHO P. The performance of a spark-ignited stratified-charge two stroke engine operating on a kerosine based aviation fuel[C]//SAE International off-Highway and Powerplant Congress and Exposition, 1997: 1-11.

[16] LA C, MURPHY P, CAKEBREAD S. Benchmarking a 2-stroke spark ignition heavy fuel engine[C]//SAE World Congress & Exhibition, 2012.

[17] 冯光烁,周明.重油航空活塞发动机技术路线分析[J].清华大学学报(自然科学版),2016,56(10):1114-1121.
FENG G S, ZHOU M. Assessment of heavy fuel aircraft piston engine types[J]. Journal of Tsinghua University (Science and Technology), 2016, 56(10): 1114-1121 (in Chinese).

[18] Deltahawk Engines Inc. Two-cycle diesel engine configured for operation with high temperature combustion chamber surfaces: US, EP16907512.4[P]. 2019-05-01.

[19] MATTARELLI E, PALTRINIERI F, PERINI F, et al. 2-stroke diesel engine for light aircraft: IDI vs. DI combustion systems[C]//SAE Powertrains Fuels & Lubricants Meeting, 2010.

[20] 于顺立,韩建立.柴油机的四种供油系统[J].工程机械与维修,2008(9):166.
YU S L, HAN J L. Four fuel supply systems for diesel engines[J]. Construction Machinery & Maintenance, 2008(9): 166 (in Chinese).

[21] UNDERWOOD S C. Performance and emission characteristics of an aircraft turbo diesel engine using JET-A fuel[D]. Atlanta: Georgia Institute of Technology, 2008.

[22] 陈霄阳.AE300 发动机燃油系统分析[J].新型工业化,2019,9(11):105-107.
CHEN X Y. Fuel system analysis of AE300 engine[J]. The Journal of New Industrialization, 2019, 9(11): 105-107 (in Chinese).

[23] 刘伍权,张光超,杨春浩,等.RP-3 航空煤油在高压共轨柴油机中的应用[J].军事交通学院学报,2017,19(11):40-45.
LIU W Q, ZHANG G C, YANG C H, et al. Application of RP-3 aviation kerosene in high pressure common rail diesel engine[J]. Journal of Military Transportation University, 2017, 19(11): 40-45 (in Chinese).

[24] 韩业扬.高压共轨柴油机过渡工况控制策略研究[D].镇江:江苏大学,2019.
HAN Y Y. Research on control strategy of high pressure common rail diesel engine under transient condition[D]. Zhenjiang: Jiangsu University, 2019 (in Chinese).

[25] 孙柏刚,刘志超,王沛,等.恒压式燃油系统压力波动特性研究[J].车用发动机,2016(6):36-40.
SUN B G, LIU Z C, WANG P, et al. Research on pressure fluctuation characteristics of constant pressure fuel system[J]. Vehicle Engine, 2016(6): 36-40 (in Chinese).

[26] 郝磊,HELLER M,金江善,等.技术创新,迎接挑战:燃油系统专家圆桌访谈[J].柴油机,2019,41(6):1-6.
HAO L, HELLER M, JIN J S, et al. Technological innovation, meeting challenges—Roundtable interview with fuel system experts[J]. Diesel Engine, 2019, 41(6): 1-6 (in Chinese).

[27] 杨白凡.二冲程点燃式缸内直喷重油发动机性能试验研究[D].南京:南京航空航天大学,2016.
YANG B F. Research on a 2-stroke DISI heavy fuel engine's performance[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2016 (in Chinese).

[28] NARDELLA F. Torsional vibration reduction for geared aviation compression ignition engines with power transmission through the camshaft or dedicated internal driveshaft a sweep through 2 and 4-stroke engines with differing numbers of cylinders and two comparison power train configurations indicate the 4-stroke, 6-cylinder engine is ideally suited for this application[C]// Noise and Vibration Conference and Exhibition, 2017: 1-8.

[29] 杨海青,陈茂杰,黄丽萍,等.夹气喷嘴瞬态喷雾的 CFD 仿真及试验[J].航空动力学报,2015,30(12):2897-2903.
YANG H Q, CHEN M J, HUANG L P, et al. CFD simulation and experiment of transient spray for an air-assisted injector[J]. Journal of Aerospace Power, 2015, 30(12): 2897-2903 (in Chinese).

[30] JIN S H, BREAR M, WATSON H, et al. An experimental study of the spray from an air-assisted direct fuel injector[J]. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 2008, 222(10): 1883-1894.

[31] CATHCART G, DICKSON G, AHERN S. The application of air-assist direct injection for spark-ignited heavy fuel 2-stroke and 4-stroke engines[R]. Warrendale: SAE Technical Papers, 2005.

[32] 赵振峰,俞春存,董雪飞,等. 二冲程重油直喷发动机混合气形成研究[J]. 航空动力学报,2021,36(8):1569-1577.
ZHAO Z F, YU C C, DONG X F, et al. Research on mixture formation of two? stroke heavy fuel direct injection engine[J]. Journal of Aerospace Power, 2021, 36(8): 1569-1577(in Chinese).

[33] 高宏力,张付军,刘波澜,等. 空气辅助喷射闪急沸腾喷雾特性试验[J]. 航空动力学报,2019,34(1):63-72.
GAO H L, ZHANG F J, LIU B L, et al. Experiment on spray characteristics of air-assisted injection under flash boiling conditions[J]. Journal of Aerospace Power, 2019, 34(1): 63-72(in Chinese).

[34] 刘锐,魏民祥,杨海青,等. 二冲程点燃式直喷煤油发动机冷起动控制策略[J]. 航空动力学报,2017,32(1):213-220.
LIU R, WEI M X, YANG H Q, et al. Cold start control strategy of a two-stroke direct injection spark-ignited kerosene engine[J]. Journal of Aerospace Power, 2017, 32(1): 213-220(in Chinese).

[35] 高宏力,张付军,王苏飞,等. 航空煤油活塞发动机空气辅助喷射系统喷雾特性试验研究[J]. 兵工学报,2019,40(5):927-937.
GAO H L, ZHANG F J, WANG S F, et al. Experimental study of air-assisted spray characteristics of aviation kerosene piston engine[J]. Acta Armamentarii, 2019, 40(5): 927-937(in Chinese).

[36] 王家雄,徐岩. 供油系统中的液体可压缩性[C]//第十六届全国大功率柴油机学术年会,2009:235-237.
WANG J X, XU Y. Liquid compressibility in the oil supply system[C]// The 16th National High-power Diesel Engine Academic Annual Conference, 2009: 235-237(in Chinese).

[37] LI X. Mechanism of atomization of a liquid jet[J]. Atomization & Sprays, 1995, 5: 89-105.

[38] ARIANE V, BURLUKA A A, BORGHI R, et al. Development of a eulerian model for the "atomization" of a liquid jet[J]. Atomization & Sprays, 2001, 11(6): 619-642.

[39] SENECAL P K, SCHMIDT D P, NOUAR I, et al. Modeling high-speed viscous liquid sheet atomization[J]. International Journal of Multiphase Flow, 1999, 25(6-7): 1073-1097.

[40] KOO J Y. An overview of liquid spray modeling formed by high-shear nozzle/swirler assembly[J]. KSME International Journal, 2003, 17(5): 726-739.

[41] IYER V, ABRAHAM J. An evaluation of a two-fluid eulerian-liquid eulerian-gas model for diesel sprays[J]. Journal of Fluids Engineering, 2003, 125(4): 660-669.

[42] DEMOULIN F X, BEAU P A, BLOKKEEL G, et al. A new model for turbulent flows with large density fluctuations: Application to liquid atomization[J]. Atomization and Sprays, 2007, 17(4): 315-345.

[43] CASTLEMAN R A. The mechanism of the atomization of liquids[J]. Bureau of Standards Journal of Research, 1930, 6: 360-376.

[44] WANG X G, HUANG Z H, KUTI O A, et al. Experimental and analytical study on biodiesel and diesel spray characteristics under ultra-high injection pressure[J]. International Journal of Heat and Fluid Flow, 2010, 31(4): 659-666.

[45] LIN X D, HAN X M, LI D G. Design and evaluation for target indicated torque based engine starting control strategy in a high pressure common rail diesel engine[J]. Mathematical Problems in Engineering, 2016, 2016: 1-8.

[46] GHASEMI A, BARRON R M, BALACHANDAR R. Spray-induced air motion in single and twin ultra-high injection diesel sprays[J]. Fuel, 2014, 121: 284-297.

[47] 林伟迪. 超高压柴油雾化及燃烧特性的三维数值模拟研究[D]. 北京:北京交通大学,2019.
LIN W D. Three-dimensional numerical simulation of atomization and combustion characteristics of diesel under ultra-high injection pressures[D]. Beijing: Beijing Jiaotong University, 2019(in Chinese).

[48] 李宝刚. 柴油机喷雾燃烧特性实验研究[D]. 重庆:重庆交通大学,2015.
LI B G. Experiment research on the spray combustion characteristics of diesel engine[D]. Chongqing: Chongqing Jiaotong University, 2015(in Chinese).

[49] 周凯. 压燃式航空活塞发动机关键技术分析[J]. 中国科技信息,2017(8):26-27.
ZHOU K. Analysis of key technology of compression ignition aero piston engine[J]. China Science and Technology Information, 2017(8): 26-27(in Chinese).

[50] 倪计民. 汽车内燃机试验技术[M]. 上海:同济大学出版社,1998.
NI J M. Test technology in vehicle-engine[M]. Shanghai: Tongji University Press, 1998(in Chinese).

[51] MATTARELLI E, RINALDINI C A, WILKSCH M. 2-stroke high speed diesel engines for light aircraft[J]. SAE International Journal of Engines, 2011, 4(2): 2338-2360.

[52] HOOPER P R, AL-SHEMMERI T, GOODWIN M J. An experimental and analytical investigation of a multi-fuel stepped piston engine[J]. Applied Thermal Engineering, 2012, 48: 32-40.

[53] SALVADOR F J, GIMENO J, DE LA MORENA J, et al. Using one-dimensional modeling to analyze the influence of the use of biodiesels on the dynamic behavior of solenoid-operated injectors in common rail systems: results of the simulations and discussion[J]. Energy Conversion and Management, 2012, 54(1): 122-132.

[54] LA C, MURPHY P, CAKEBREAD S. Benchmarking a 2-stroke spark lgnition heavy fuel engine[C]// SAE 2012 World Congress & Exhibition, 2012.

[55] SENER R, YANGAZ M U, GUL M Z. Effects of injection strategy and combustion chamber modification on a single-cylinder diesel engine[J]. Fuel, 2020, 266: 117122.

[56] 胡春明,毕延飞,王齐英,等. 航空活塞式发动机瞬态空燃比控制仿真研究[J]. 航空动力学报,2018,33(5):1236-1244.
HU C M, BI Y F, WANG Q Y, et al. Simulation research on transient air-fuel ratio control of aero piston engine[J]. Journal of Aerospace Power, 2018, 33(5): 1236-1244(in Chinese).

[57] 王振宇,丁水汀,杜发荣. 航空重油发动机燃油供给系统动

力学模型分析[J]. 航空动力学报,2012,27(4):846-853.
WANG Z Y,DING S T,DU F R. Analysis of dynamics modeling of aviation heavy oil piston engine fuel delivery system[J]. Journal of Aerospace Power,2012,27(4):846-853(in Chinese).

[58] 陈林林. 二冲程煤油发动机性能数值模拟与喷油控制研究[D]. 南京:南京航空航天大学,2009.
CHEN L L. Research on numerical simulation of performance and fuel injection control for two-stroke kerosene engine[D]. Nanjing:Nanjing University of Aeronautics and Astronautics,2009(in Chinese).

[59] 李长胜. 小型高速航空重油活塞发动机混合气形成及燃烧特性的研究[D]. 北京:北京交通大学,2014:27.
LI C S. Study on the mixture formation and combustion of light-weight,high-speed,heavy oil aeroengine[D]. Beijing:Beijing Jiaotong University,2014:27(in Chinese).

[60] 石允. 小型点燃式二冲程活塞重油发动机混合气形成的模拟研究[D]. 北京:北京交通大学,2012:49-92.
SHI Y. Study on heavy-fuel mixture formation of light-weight,spark-ignition,two-stroke engine[D]. Beijing:Beijing Jiaotong University,2012:49-92(in Chinese).

[61] 刘波澜,颜超,于飞,等. 重油活塞发动机夹气燃油喷射系统仿真研究[J]. 北京理工大学学报,2021,41(6):597-602.
LIU B L,YAN C,YU F,et al. Simulation research on air-assisted fuel injection system for keorsene piston engine[J]. Transactions of Beijing Institute of Technology,2021,41(6):597-602(in Chinese).

[62] 陈舟. 高压共轨喷油器原尺寸透明喷嘴内线空化特性及其对喷雾影响的试验研究[D]. 镇江:江苏大学,2019.
CHEN Z. Experimental study of string cavitation in real-size optical high-pressure common rail injector nozzles and its effects on spray[D]. Zhenjiang:Jiangsu University,2019(in Chinese).

[63] 石智成,吴晗,路佩,等. 冷起动工况柴油/煤油燃料喷雾燃烧特性研究[J]. 工程热物理学报,2020,41(5):1279-1285.
SHI Z C,WU H,LU P,et al. Spray and combustion characteristics of diesel/kerosene blends under cold-start conditions[J]. Journal of Engineering Thermophysics,2020,41(5):1279-1285(in Chinese).

[64] EAGLE W E,MORRIS S B,WOOLDRIDGE M S. High-speed imaging of transient diesel spray behavior during high pressure injection of a multi-hole fuel injector[J]. Fuel,2014,116:299-309.

[65] PANDEY R K,REHMAN A,SARVIYA R M. Impact of alternative fuel properties on fuel spray behavior and atomization[J]. Renewable and Sustainable Energy Reviews,2012,16(3):1762-1778.

[66] 刘正平. 高压定容喷雾模拟实验系统研制及应用研究[D]. 北京:北京交通大学,2013.
LIU Z P. Development and application of constant volume spray experiment system under high ambient pressure[D]. Beijing:Beijing Jiaotong University,2013(in Chinese).

[67] 胡若,刘福水,王沛,等. 电控单体泵燃油系统低速供油特性试验研究[J]. 内燃机工程,2017,38(1):70-75.
HU R,LIU F S,WANG P,et al. Experimental study on low speed fuel delivery characteristics of electronic unit pump fuel system[J]. Chinese Internal Combustion Engine Engineering,2017,38(1):70-75(in Chinese).

[68] 樊志强. 电控单体泵系统供油特性及其凸轮型线参数化设计[D]. 北京:北京理工大学,2014.
FAN Z Q. Research on fuel supply characteristics of electronic unit pump and parametric design of its cam profile[D]. Beijing:Beijing Institute of Technology,2014(in Chinese).

[69] 苏思源,胡春明,刘娜,等. 活塞式航空煤油发动机冷启动性能试验[J]. 航空动力学报,2020,35(1):97-105.
SU S Y,HU C M,LIU N,et al. Experiment on cold start of piston aviation kerosene engine[J]. Journal of Aerospace Power,2020,35(1):97-105(in Chinese).

[70] 韦雄,冒晓建,肖文雍,等. 国产电控单体泵供油量精确测量与一致性规律分析[J]. 内燃机工程,2011,32(6):54-57.
WEI X,MAO X J,XIAO W Y,et al. Accurate measurement and consistency analysis of fuel delivery for domestic electronic unit pump[J]. Chinese Internal Combustion Engine Engineering,2011,32(6):54-57(in Chinese).

[71] 王睿,杨卫平,雷军,等. 某10支共轨喷油器性能测试试验[J]. 拖拉机与农用运输车,2021,48(1):23-26.
WANG R,YANG W P,LEI J,et al. Performance test of 10 common-rail injectors[J]. Tractor & Farm Transporter,2021,48(1):23-26(in Chinese).

[72] 严明,杨青,王沛,等. 启喷压力对电控单体泵供油系统喷油量的影响机理[J]. 车用发动机,2015(1):59-63.
YAN M,YANG Q,WANG P,et al. Influencing mechanism of opening pressure on fuel injection quantity of electronic unit pump injection system[J]. Vehicle Engine,2015(1):59-63(in Chinese).

[73] 周昕毅,李铁,王宁. 柴油蒸发喷雾瞬态液相贯穿距模型及试验验证[J]. 内燃机工程,2020,41(2):8-14.
ZHOU X Y,LI T,WANG N. Modeling and experimental verification of transient liquid penetration of diesel evaporating spray[J]. Chinese Internal Combustion Engine Engineering,2020,41(2):8-14(in Chinese).

[74] 岂斌. 高压共轨燃油喷射雾化特性试验研究[D]. 长沙:中南大学,2010.
QI B. Experiments research on spray characteristics of diesel fuel injection system for high-pressure common rail[D]. Changsha:Central South University,2010(in Chinese).

[75] 魏衍举,章旭东,邓胜才,等. 强涡流场中柴油喷雾扩散特性研究[J]. 农业机械学报,2019,50(8):394-399.
WEI Y J,ZHANG X D,DENG S C,et al. Diffusion characteristics of diesel spray in swirl flows field[J]. Transactions of the Chinese Society for Agricultural Machinery,2019,50(8):394-399(in Chinese).

[76] WU H,WANG L L,WU Y,et al. Spray performance of air-assisted kerosene injection in a constant volume chamber under various in-cylinder GDI engine conditions[J]. Applied Thermal Engineering,2019,150:762-769.

[77] 高岩飞. 点燃式煤油发动机燃油雾化技术研究[D]. 南京:南京航空航天大学,2010.
GAO Y F. Research on atomization technology of kerosene in spark ignition engine[D]. Nanjing:Nanjing University of Aero-

nautics and Astronautics,2010(in Chinese).

[78] 韩玉琪,袁善虎,王飒."碳中和"目标牵引下的航空动力发展分析[J].航空动力,2021(6):28-30.
HAN Y Q,YUAN S H,WANG S. Analysis to the development of aero engine to achieve carbon neutrality[J]. Aerospace Power,2021(6):28-30(in Chinese).

[79] 郭世龙.基于柴油机电控喷油器结构参数与喷油特性相关关系的试验研究[D].石河子:石河子大学,2019.
GUO S L. Experimental research on correlation between structural parameters and fuel injection characteristics of diesel engine controlled fuel injector[D]. Shihezi:Shihezi University,2019(in Chinese).

[80] 杨耿.轻量化单缸直喷式柴油机喷油器与燃烧系统匹配研究[D].南宁:广西大学,2016.
YANG G. Characteristics study on light-weight single cylinder direct injection diesel engine in cylinder flow[D]. Nanning:Guangxi University,2016(in Chinese).

[81] 高广新,袁竹林,王仁辉,等.高速大功率柴油机共轨喷油器关键结构设计与参数匹配[C]// 2019 中国汽车工程学会年会,2019:77-81.
GAO G X,YUAN G L,WANG R H, et al. Key structural design and parameter matching of common rail injector for high speed and high power diesel engine[C]//2019SAECCE-IEE,2019:77-81(in Chinese).

[82] 周兴利.电控柴油机故障智能诊断研究[D].上海:上海交通大学,2009.
ZHOU X L. Study of intelligence diagnosis for electronic controlled diesel engine faults[D]. Shanghai:Shanghai Jiao Tong University,2009(in Chinese).

[83] PILOTO-RODRÍGUEZ R,DÍAZ Y,MELO-ESPINOSA E A,et al. Conversion of fatty acid distillates into biodiesel:Engine performance and environmental effects[J]. Energy Sources,Part A: Recovery, Utilization, and Environmental Effects, 2020,42(4):387-398.

[84] 崔宇航,卫海桥,王祥庭,等.高海拔模拟环境下柴油机燃烧粗暴可视化试验研究[J].天津大学学报(自然科学与工程技术版),2022,55(4):383-390.
CUI Y H,WEI H Q,WANG X T,et al. Optical experiments on diesel knock under simulated high-altitude conditions[J]. Journal of Tianjin University (Science and Technology),2022,55(4):383-390(in Chinese).

[85] 王正江,申立中,万明定,等.柴油机起动过程瞬态喷油量的控制策略[J].内燃机学报,2021,39(4):320-325.
WANG Z J,SHEN L Z,WAN M D,et al. Transient fuel control strategy for diesel engine under starting process[J]. Transactions of CSICE,2021,39(4):320-325(in Chinese).

[86] 徐劲松,聂珂,黄国勇,等.压燃式航空活塞发动机转矩与空燃比的控制[J].航空动力学报,2021,36(5):1083-1093.
XU J S,NIE K,HUANG G Y,et al. Control of torque and air-fuel ratio of compression-ignition aero piston engine[J]. Journal of Aerospace Power,2021,36(5):1083-1093(in Chinese).

[87] MATHEWS J A. Carbon-negative biofuels[J]. Energy Policy,2008,36(3):940-945.

[88] RAVI V,GAO A H,MARTINKUS N B,et al. Air quality and health impacts of an aviation biofuel supply chain using forest residue in the northwestern United States[J]. Environmental Science & Technology,2018,52(7):4154-4162.

[89] 李鹏宇,王新校,栾军山,等.基于韦伯模型的柴油机燃烧控制模型与一维仿真[J].内燃机与配件,2021(21):5-8.
LI P Y,WANG X X,LUAN J S,et al. Diesel engine combustion control model based on Wiebe expression and one-dimensional simulation[J]. Internal Combustion Engine & Parts,2021(21):5-8(in Chinese).

[90] 孙超.柴油机燃用小球藻生物柴油 Soot 及 NOX 生成历程的数值分析[D].镇江:江苏大学,2020:91.
SUN C. Numerical analysis of soot and NOX generation process of diesel engine with chlorella biodiesel fuel[D]. Zhenjiang:Jiangsu University,2020:91(in Chinese).

[91] 孔祥恩,刘海峰.无人机用航空活塞发动机关键技术的研究进展[J].小型内燃机与车辆技术,2021,50(3):79-87.
KONG X E,LIU H F. Research progress of key technologies of aviation piston engine for UAV[J]. Small Internal Combustion Engine and Vehicle Technique,2021,50(3):79-87(in Chinese).

Fuel injection technology of heavy fuel aircraft piston engine

DING Shuiting[1], SHAO Longtao[2], ZHAO Shuai[1], ZHU Kun[3], DU Farong[1], ZHOU Yu[1,*]

(1. Research Institute of Aero-Engine, Beihang University, Beijing 100083, China;
2. School of Energy and Power Engineering, Beihang University, Beijing 100083, China;
3. Areo Engine Academy of China, Beijing 101300, China)

Abstract: Fuel injection technology is one of the key technologies of aviation piston engines. Its structure and characteristics directly affect the integral performance of aircraft piston engines. With the requirements of general aviation for better engine performance, the development of fuel injection technology continues. This paper summarizes the current fuel supply systems used in various heavy fuel aircraft piston engines, sorts out the classic models that are applicable to them, and reviews the advantages and disadvantages of the various fuel supply systems. It expounds some cutting-edge theories, simulation calculations and experimental methods used in the study of heavy oil injection technology, and analyzes some difficulties in heavy fuel atomization, followed by relevant suggestions for the study of heavy fuel injection technology. Finally, this paper puts forward some ideas and key technologies for the future development of heavy fuel aircraft piston engine, with a focus on providing technical guidance for the forward development of heavy fuel aircraft piston engines, thus contributing to the sustainable development of China's general aviation.

Keywords: aircraft piston engine; heavy fuel; inject technology; two stroke; carbon-negative biofuel; control strategy

Received: 2022-01-12; **Accepted**: 2022-02-16; **Published online**: 2022-04-29 08:57

URL: kns.cnki.net/kcms/detail/11.2625.V.20220428.1555.003.html

Foundation items: National Basic Research Program of China (2018YFB010400); National Natural Science Foundation of China (51775025, 51775013)

* **Corresponding author**. E-mail: zybuaa@hotmail.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0196

超临界二氧化碳闭式布莱顿循环系统研究进展

邹正平[1,2], 王一帆[2,3], 姚李超[1,2,*], 刘火星[1,2], 许鹏程[2,3], 李辉[2,3]

(1. 北京航空航天大学 航空发动机研究院, 北京 102206; 2. 航空发动机气动热力国防科技重点实验室, 北京 100083;
3. 北京航空航天大学 能源与动力工程学院, 北京 100083)

摘 要: 超临界二氧化碳(SCO_2)闭式布莱顿循环凭借高热效率、高紧凑性和高经济-环保性等优势,已成为能源与动力领域的热点技术之一。针对超临界二氧化碳闭式布莱顿循环,详细介绍了工作原理、优势及国内外相关研究进展,总结了循环总体热力、超临界工质叶轮机、紧凑高效换热器、控制及储热等相关关键技术的研究现状,并对当前工程应用面临的问题和未来技术发展方向进行了分析和展望。分析表明,循环总体热力设计阶段应涵盖部件低维性能分析,以评估部件性能指标的可实现性,并综合考虑全寿命周期性能、紧凑性、经济性等指标。工质的剧烈物性变化导致叶轮机与换热器内部特殊流动与换热机理,需发展充分考虑工质特殊物性影响的叶轮机和紧凑换热器设计方法;通过理论分析和机器深度学习相结合构建不同工质叶轮机相似方法,可为超临界二氧化碳叶轮机气动性能试验验证提供理论基础。此外,鲁棒高效的控制策略可实现超临界二氧化碳闭式布莱顿循环有效可靠调控,而集成新型介质储热技术的超临界二氧化碳闭式布莱顿循环系统将为高温光热发电提供关键技术支撑。

关 键 词: 超临界二氧化碳(SCO_2); 闭式布莱顿循环; 热力建模与分析; 叶轮机; 紧凑高效换热器

中图分类号: TK123

文献标志码: A **文章编号**: 1001-5965(2022)09-1643-35

当前,绝大多数能源均以热能动力利用的形式被转化为电能或机械能,如何提高动力循环的性能成为能源深度开发的主要技术途径。在现有成熟循环构型和部件设计技术基础上,通过改变循环工质以改善循环性能是热能动力利用领域的热点研究方向,诸如超临界氦[1-2]、氦氙混合物[3-4]、超临界二氧化碳(SCO_2)[5]等特殊工质被逐步挖掘,其中SCO_2闭式布莱顿循环具备诸多潜在优势,被视作有望突破现有动力循环技术发展瓶颈的革新技术[6]。

早在20世纪60年代,Feher等[7-8]便率先提出将SCO_2用作动力循环工质的构想,指出SCO_2循环凭借高热效率和高紧凑性等优势在地面/空天发电及动力推进等领域具有广阔应用前景。之后,Angelino[9-10]通过对比不同SCO_2循环和蒸汽循环的性能,证实无再热SCO_2循环的热效率便可超过再热蒸汽循环效率,且具备更高的结构紧凑性和经济性,其工作掀起了关于SCO_2循环的研究热潮,所提出的再压缩循环构型为后续众多的设计研究奠定了坚实基础[11-13]。然而,受当时叶轮机和紧凑换热器的设计/制造水平等因素的限制,研究热潮在20世纪70年代末期便戛然而止,相关研究成果也并未得到工程实际应用。随着工业领域对动力循环效率和经济性要求日益严苛,

收稿日期: 2022-03-29; **录用日期**: 2022-04-22; **网络出版时间**: 2022-05-17 17:06
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220517.1414.002.html
基金项目: 航空发动机气动热力国防科技重点实验室基金 (2021-JCJQ-LB-062-0205)
* **通信作者**. E-mail: ylcpersonal@buaa.edu.cn

引用格式: 邹正平, 王一帆, 姚李超, 等. 超临界二氧化碳闭式布莱顿循环系统研究进展[J]. 北京航空航天大学学报, 2022, 48(9): 1643-1677. ZOU Z P, WANG Y F, YAO L C, et al. Progress in research of closed supercritical carbon dioxide Brayton cycle system[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1643-1677(in Chinese).

SCO_2循环于21世纪初重新受到关注。Dostal等[14]率先探究了SCO_2循环在新一代核电领域的应用前景,从热力性能、经济性、叶轮机/换热器部件可行性等多方面系统地论述了该循环的优势。虽然其研究愿景为新一代核能发电,但研究人员敏锐地洞察了SCO_2闭式布莱顿循环在热动力利用领域的广阔应用前景,很快便在全球范围内掀起了新的研究热潮。

本文综述了国内外在SCO_2闭式布莱顿循环技术领域的相关研究现状,并结合北京航空航天大学高超声速强预冷团队(简称北航团队)的研究,对循环总体热力、叶轮机、紧凑换热器、控制策略、储热等关键技术进行了介绍。

1 SCO_2闭式布莱顿循环国内外研究现状

1.1 循环原理及优势

较传统蒸汽/燃气循环而言,SCO_2闭式布莱顿循环具备高热效率和高紧凑性2个突出的性能优势,这主要得益于工质的特殊物性。当处于超临界状态时(临界温度304 K,临界压力7.38 MPa),二氧化碳兼具气体的低黏性、强流动性、易压缩/膨胀的性质,以及液体的高密度、高传热效率、强做功能力。更重要的是,其在临界点附近物性变化非常剧烈(见图1),对位于近临界点的压缩过程而言,较小的压缩功耗便足以带来工质物性的显著变化[15];此外,二氧化碳在近临界点区域偏离理想气体的程度很大(压缩因子0.2~0.5),低可压缩性使得其压缩过程非常类似于液体压缩过程,相应的压缩功耗也较低[16]。SCO_2闭式布莱顿循环的工作原理正是将压缩过程起点置于临界点附近,通过降低压气机功耗进而有效提升循环热效率。

研究表明,在最高循环温度t_{max}为500~800℃时,间接式SCO_2闭式布莱顿循环热效率可达40%~55%,且可在更低温度下(300~500℃)保持20%以上;当t_{max}超过1 000℃,直接式SCO_2闭式布莱顿循环的热效率可达到60%~65%,远高于燃气循环,更高温度下甚至优于蒸汽-燃气轮机联合循环[16],如图2(a)所示。另外,SCO_2的密度非常大(接近空气密度的200倍),使得SCO_2闭式布莱顿循环中的叶轮机、换热器、管路附件等部

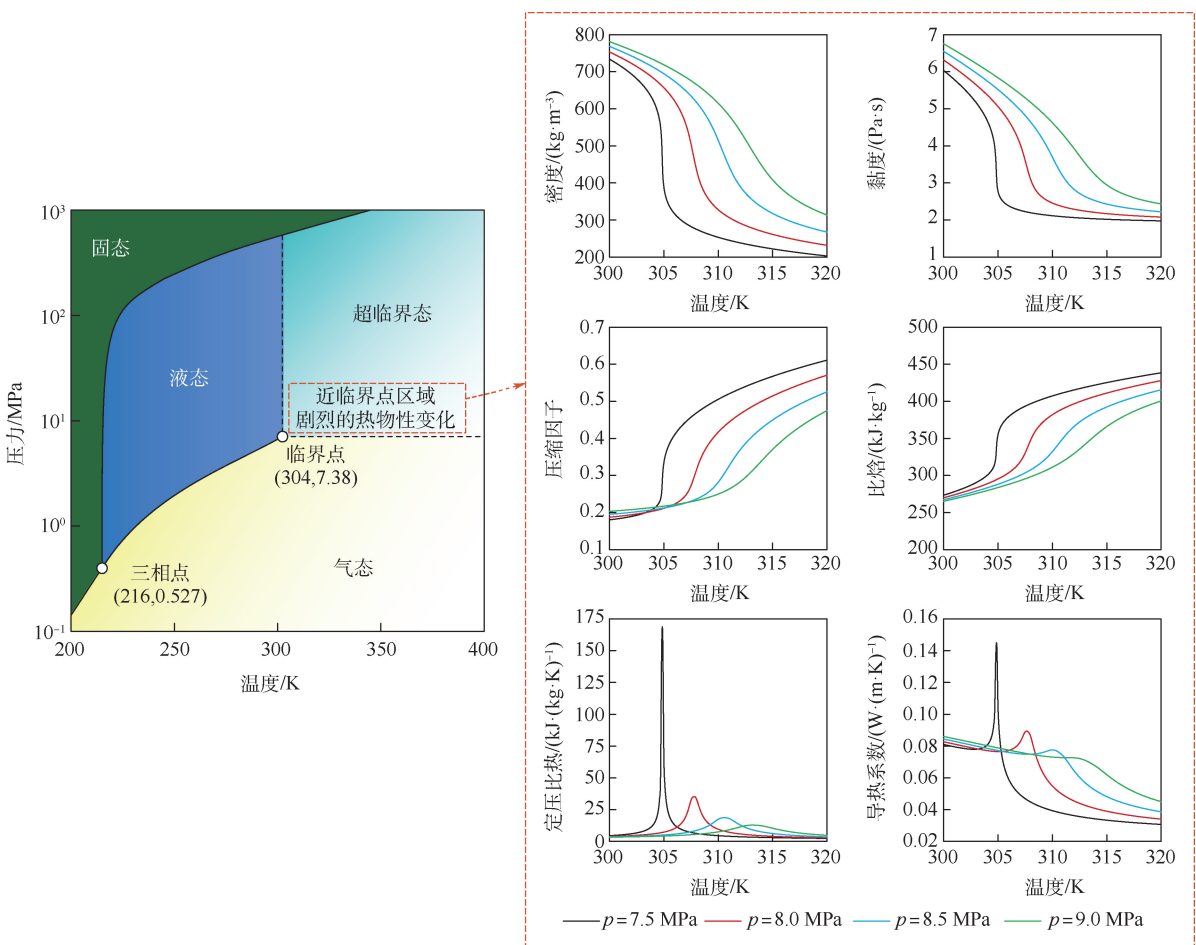
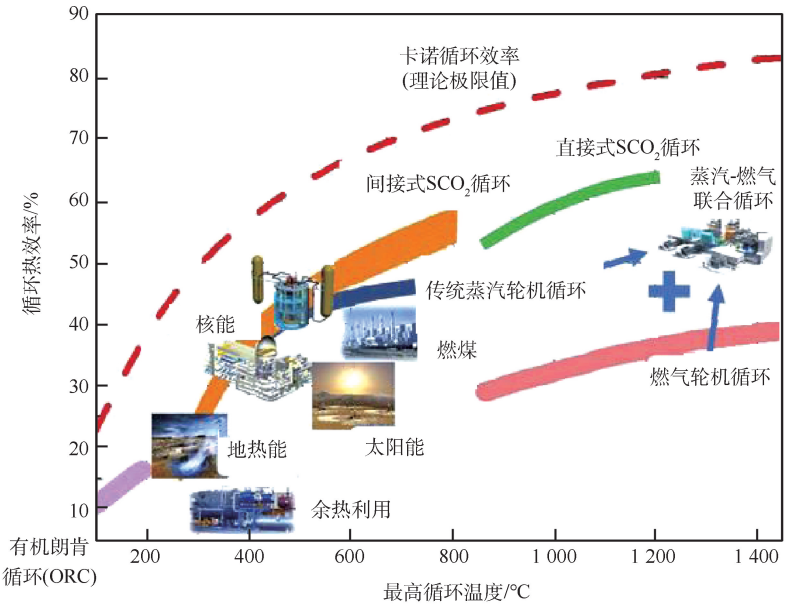
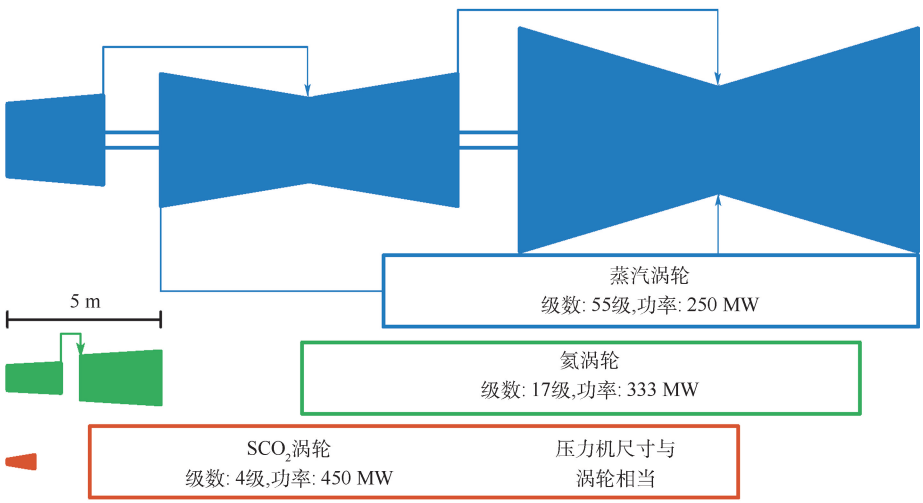


图1 SCO_2在临界点附近的剧烈物性变化

Fig. 1 Dramatic variation of properties of SCO_2 near critical point



(a) 高循环热效率[16]



(b) 低部件尺寸和高紧凑性[14]

图 2　SCO_2闭式布莱顿循环的性能优势

Fig. 2　Performance advantages of closed SCO_2 Brayton cycle

件的尺寸较小;同时,叶轮机级数也较少(10 MW 级系统仅需采用单级径流式叶轮机设计),相应地,系统紧凑性大幅提升,如图 2(b)所示。相同功率量级下,SCO_2闭式布莱顿循环中叶轮机部件的体积可降至 0.1% ~5%;10 MW 级循环中叶轮机部件的直径仅接近 0.15 m,而 300 MW 级循环中叶轮机部件的直径也仅 1.2 m 左右[17]。以蒸汽循环和 SCO_2闭式布莱顿循环对比为例,在 480℃热源温度下,蒸汽循环的热效率接近 40%,而 SCO_2闭式布莱顿循环的热效率则可达 42%,同时系统总体积不到蒸汽循环系统的 50%[18]。此外,SCO_2工质本身也具备以下天然优势[19]:①安全性高。热物性稳定,无毒且不可燃。②环境友好。对臭氧层无破坏,温室效应低。③经济性高。自然界中的二氧化碳资源充足,且 SCO_2制备工艺成熟。

由于上述诸多优势,SCO_2闭式布莱顿循环被视作当前蒸汽/燃气循环的有力潜在替代者,在核能发电、太阳能光热发电、燃煤发电、燃料电池发电、余热利用、舰船动力系统甚至高超声速飞行器动力系统等领域均具备广阔的应用前景,如图 3 所示。在新一代核电领域,SCO_2闭式布莱顿循环有望在当前钠冷反应堆出口温度(510 ~525℃)下达到 43%的热效率,在未来第四代反应堆更高出口温度下热效率甚至有望突破 50%[20]。在太阳能光热发电领域,协同熔盐储热和高效风冷的 SCO_2闭式布莱顿循环被视为未来大幅降低光热发电度电成本的首选技术途径,主要原因在于[21-22]:①相同材料下,工质低腐蚀性允许更高的

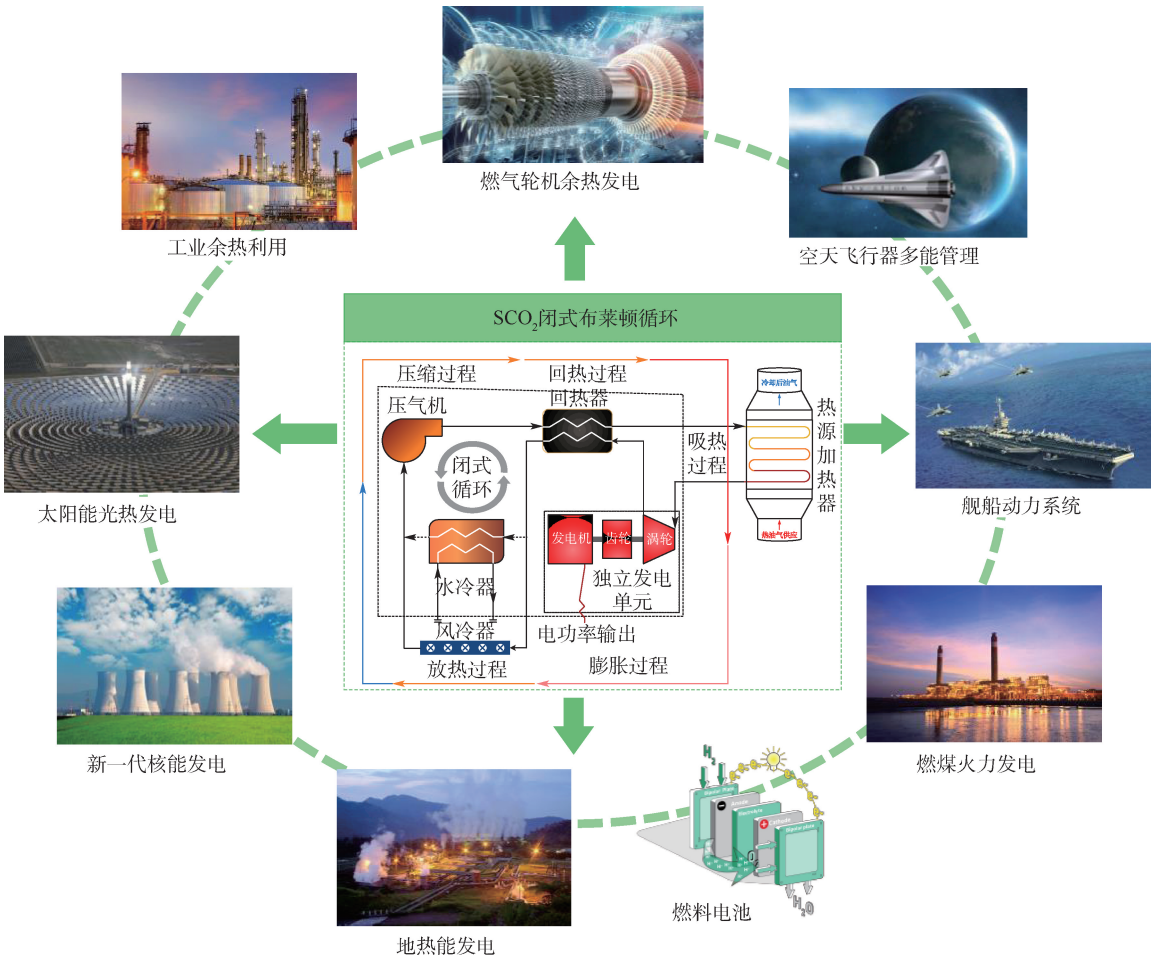


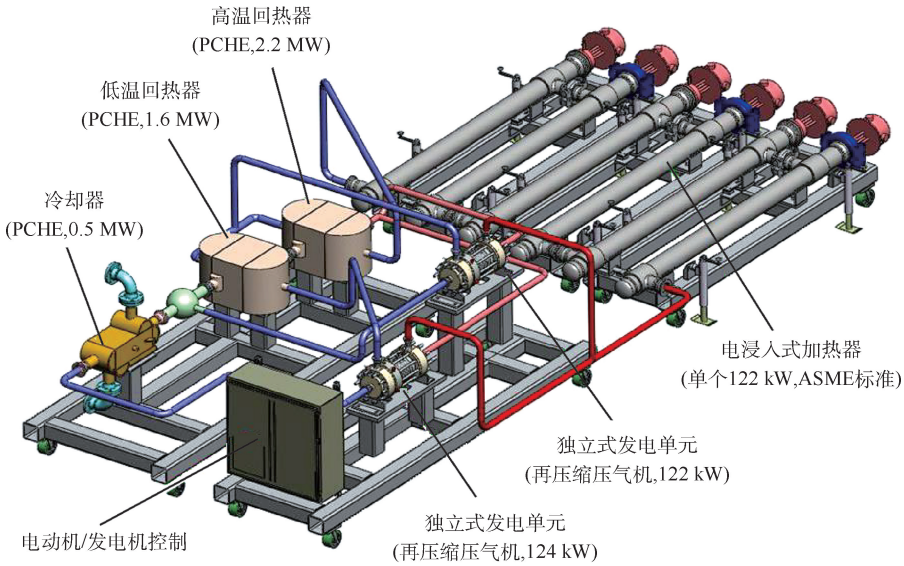
图 3 SCO_2闭式布莱顿循环潜在应用领域

Fig. 3 Potential applications of closed SCO_2 Brayton cycle

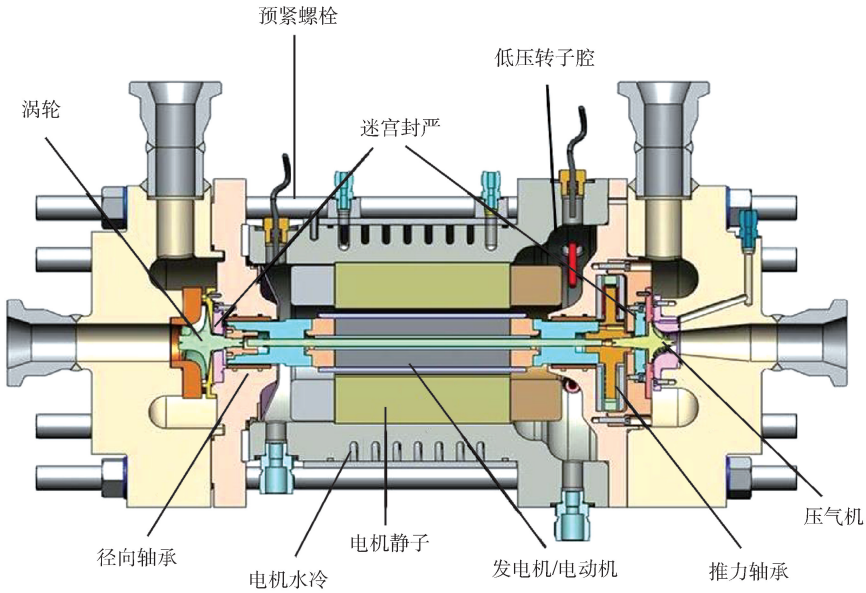
循环温度;②热效率更高(在 600 ~ 700℃甚至更高温度下,热效率可达到甚至超过 50%),系统体积更小,质量更轻,可大幅降低制造、安装和维护成本(包括定日镜场-集热-储热系统);③高温风冷下热效率也相对更高,适用于干旱地区。在低温余热利用领域,当前有机朗肯循环存在易燃、有毒、化学稳定性较差等诸多劣势,而 SCO_2闭式布莱顿循环则更加安全可靠,并能在低温热源下稳定高效运行。在燃煤发电领域,目前基于蒸汽朗肯循环并采用 MEA 燃烧后脱碳技术的发电系统在最大温度 700℃时的热效率接近 42.5%[23],而采用 SCO_2闭式布莱顿循环则有望将现有热效率提高 6.2% ~ 7.4%,同时获得度电成本下降 7.8% ~ 13.6% 的经济性能收益[24]。在舰船/空天动力系统领域,SCO_2闭式布莱顿循环既能保证高热效率,同时高紧凑性又能满足在狭小受限空间中布置的要求,非常符合未来“综合电站”和“能量综合管理系统”对高效-紧凑-轻质动力系统的严苛需求[25-27]。研究表明,基于 SCO_2闭式布莱顿循环的涡扇发动机联合循环可降低 1.9% 油耗[28],而采用 SCO_2闭式布莱顿循环的超燃冲压发动机冷却燃油流量有望降低 21.5% ~ 28.8%,同时壁面温度可控制在 995 K 安全温度以内[29]。在燃料电池发电领域,采用 SCO_2闭式布莱顿循环不仅可带来热效率的提高和成本的降低,更能减少发电系统有害物质的排放,同时几乎对环境无污染[28-31]。

1.2 国内外相关机构研究进展

国外起步较早,现已进入试验验证、系统集成和示范电站应用阶段。美国能源部(DOE)早在 2003 年便率先启动了第四代核电先进动力循环项目,其下属桑迪亚国家实验室(SNL)联合麻省理工学院(MIT)等多家科研机构,建成了 250 kW 级再压缩循环试验平台(见图 4),完成了包括叶轮机特性试验、二氧化碳混合工质试验、系统动态特性等多项关键技术验证[32-34]。2011 年,DOE 又启动了 Sunshot 计划,旨在采用 SCO_2闭式布莱顿循环将太阳能光热发电成本降至 6 美分/(kW · h),该项目在 10 MW 级轴流涡轮、50 MW 涡轮和 450 MW 压气机、微通道换热器等方面取得众多研究成果[35-37]。2012 年,DOE 启动在燃煤发电领域的研究项目,旨在以 SCO_2闭式布莱顿循环替代蒸汽朗肯循环和燃气布莱顿-朗肯联合循环,研发叶轮机、



(a) 循环系统集成



(b) 集成发电单元体

图 4　SNL SCO_2闭式布莱顿循环系统试验验证平台[34]

Fig. 4　Experimental platform of closed SCO_2 Brayton cycle at SNL[34]

回热器、燃烧器及关键材料,并将技术成熟度提升至 7 级。之后,DOE 又相继启动了 STEP 计划和 APOLLO 计划,设计、建造和运行 10 MW 级示范电站,实现 700℃涡轮入口温度,并验证 50% 以上的循环热效率。目前,美国西南研究院(SwRI)已在 720℃、27.5 MPa 全温全压下完成了 10 MW 循环涡轮全尺寸性能试验和耐久性测试。此外,美国海军核实验室也针对舰船推进系统推出海军堆项目,联合 SNL、NASA、MIT 等机构开展包括循环及部件设计、流动与传热试验和材料相容性试验等研究,搭建了 100 kW 级系统试验平台,并完成了叶轮机低转速性能试验,以及启动过程、加热、停机等过程动态性能研究[38-41]。美国 Echogen 公司则针对商用余热利用研发了 2 款型号产品:EPS100(8 MW)和 EPS5(300 kW)。其中,EPS100 主要用于 500 ~ 550℃、65 ~ 70 kg/s 流量的气体燃烧掺入(如燃机排气)的余热回收,其已完成调试的最大功率达到 3.1 MW[42]。

韩国和日本也在 SCO_2闭式布莱顿循环领域开展了大量研究。韩国原子能研究院(KAERI)早在 2009 年便探究了 SCO_2闭式布莱顿循环在第四代核能发电系统的应用前景,系统开展了循环设计与热力分析、叶轮机和紧凑换热器设计和性能分析、变工况性能分析,以及动态特性分析及控制策略等研究[43]。此后,KAERI 联合韩国科学技术院(KAIST)规划并设计了兼具部件和循环试验能力的平台,计划分 4 个阶段逐步完成再压缩循环的原理验证[44]。第 1 阶段完成了高速压气机

设计和性能分析[45];第2阶段搭建了100 kW级试验平台,实现了电功率输出并获得了压气机低转速性能曲线[46]。韩国能源研究所(KIER)也搭建了不同功率量级的循环试验平台,正在开展相关试验研究[47-48]。此外,日本东京工业大学(TIT)针对中低温工业余热利用设计了2.3 MWe级循环和10 kW级的缩尺试验台,并计划开展压气机压缩功耗下降验证和换热器性能测试等工作[49]。

此外,在Horizon 2020计划的支持下,欧盟也在积极推进SCO_2闭式布莱顿循环技术的理论与应用研究,并启动了2个特别研究计划:sCO_2-Flex和sCO_2-HeRo计划,旨在开发并验证用以燃煤发电和核燃料余热利用的系统。目前,包括英国、德国、捷克等国多个研究机构参与该计划,并在SCO_2高效换热机理、SCO_2压气机、50 kW级循环试验验证等方面取得一定成果[50-52]。

国内虽在SCO_2闭式布莱顿循环技术领域起步较晚,但已在各关键技术研究、试验验证和示范机组建设等方面取得显著成果。北京航空航天大学针对工业余热利用完成3 MW级循环和叶轮机、紧凑换热器等部件设计;中国核动力研究设计院初步建立了用于研究循环流阻、换热和流动不稳定性等问题的试验平台[53];西安热工研究院完成了5 MW级循环试验平台设计建设及试验机组72 h满负荷试运行[54];中国科学院工程热物理研究所建成MW级SCO_2压气机试验平台并完成全载测试运行,建成示范机组并实现锅炉点火和稳定运行;首航高科能源技术股份有限公司正与法国企业合作,推进利用SCO_2闭式布莱顿循环改造敦煌10 MW太阳能光热发电站;中国船舶集团有限公司第七一一研究所也在积极探索SCO_2闭式布莱顿循环在舰船动力领域的应用可行性。

2 SCO_2闭式布莱顿循环构型及总体技术

2.1 循环构型

优化循环组构、构建高性能循环构型一直是SCO_2闭式布莱顿循环的研究热点之一。出于不同的应用场景和性能提升目的,研究人员发展了众多循环构型。Crespi等[6]对超过60种循环构型进行了全面的总结,并将其分为2类:独立循环(stand-alone cycle)和联合循环(combined-cycle)。下面对几种典型的独立循环构型进行介绍,如图5所示。

2.1.1 典型独立循环构型演化及性能对比

最基本的SCO_2闭式布莱顿循环为简单布莱顿循环,其包含一个压气机、一个涡轮、一个冷却器和一个热源换热器。由于工质经涡轮膨胀直接进入冷却器放热,其所携带的大量余热被浪费,导致该循环热效率较低。为提高循环热效率,引入回热器来回收涡轮膨胀后的余热以预热压气机出口的低温流体,达到降低循环吸热量的目的,即为简单回热循环。在简单回热循环基础上,采用间冷或再热等手段可进一步提高循环热效率。间冷循环中,放热和压缩过程分2次进行,如此通过降低再次压缩过程的起始温度可有效减少循环压缩耗功;再热循环中,吸热和膨胀过程分2次进行,如此通过提高再次膨胀过程起始温度可有效增加循环膨胀做功。

受工质物性剧烈变化的影响,简单回热循环中回热器易出现温差夹点问题,导致其热效率进一步提升受限[14]。为此,可将回热器拆分为高温回热器和低温回热器,通过改变低温回热器冷侧(或热侧)的流动参数来控制其换热特征,进而规避温差夹点问题。典型的采用高/低温回热器布局的循环包括再压缩循环和预压缩循环。预压缩循环中,高温回热器热侧出口的低压流体经过预先压缩进入低温回热器,预压缩的增温升压作用减小了低温回热器冷热两侧的热容差异,进而有效地增强其换热能力。再压缩循环中,低温回热器热侧出口的流体被分流成两支路:一股流体经过放热-压缩后回到低温回热器冷侧,另一股流体进入并行的压气机,之后2股流体汇合后进入高温回热器冷侧。对低温回热器而言,分流使得其冷侧流量降低,不仅使得循环放热量降低,同时通过减小冷热两侧的热容差异有效削弱了温差夹点问题。在再压缩循环的基础上演化而来的局部冷却循环则在分流前引入新的预冷-压缩过程,降低放热支路的温度并提高其压力,循环放热量可进一步降低,进而带来热效率收益。此外,局部冷却循环比功率更高,同时其热效率对最优循环增压比变化的敏感性更低。

为分析不同循环构型性能优劣,北航团队自主开发了SCO_2闭式布莱顿循环模块化建模/仿真程序,并对比了上述几种独立循环构型的热力性能[55],如图6所示。图中:循环增压比为3.2,最低压力为7.5 MPa,最高循环温度为550℃。可以看到,较简单回热循环而言,再压缩、预压缩和局部冷却循环具有明显的效率优势,其热效率增幅可达5%以上;而传统的再热和间冷增效手段所带

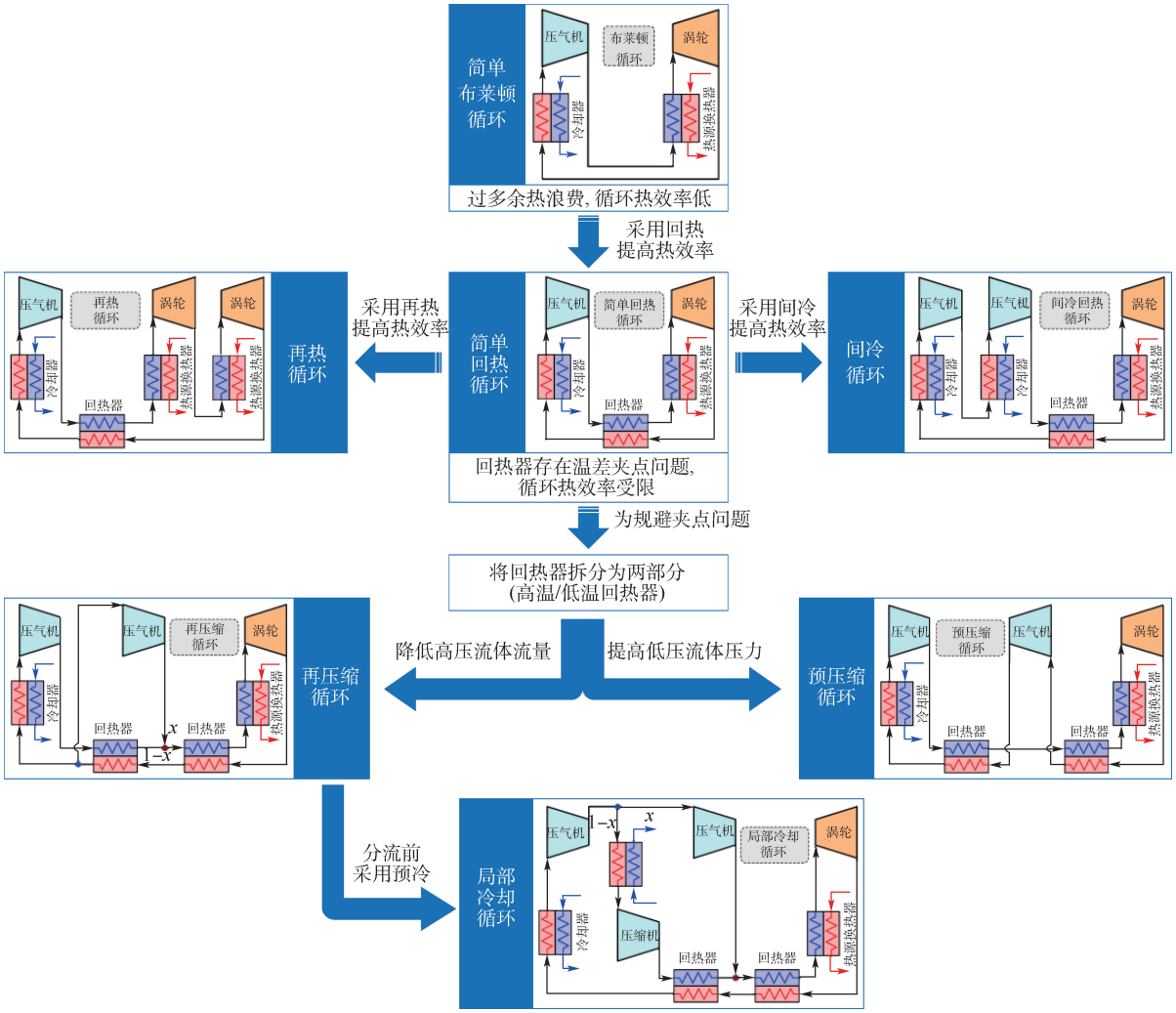


图 5　典型的 SCO_2 闭式布莱顿循环构型及其演化关系

Fig. 5　Layouts and derivative relationships of typical closed SCO_2 Brayton cycles

来的循环热效率增幅有限；在这 3 种高效率循环构型中，再压缩循环和预压缩循环的热效率基本持平，而局部冷却循环的热效率相对更高。此外，再压缩循环具备高流量-低比功率的特点，而局部冷却循环则表现出低流量-高比功率的特点。

2.1.2　多再压缩循环

在再压缩循环基础上，北航团队提出了一种多再压缩循环构型，可进一步提高循环热效率[55]。如图 7 所示，该循环包含 n 个再压缩单元（图中绿色虚线所示），其按顺序串行排列于高温回热器和主压缩压气机之间，每个单元包含串行低温回热器、分流三通、并行再压缩压气机和合流三通。当 $n=1$ 时，该循环退化为再压缩循环；当 $n=0$ 时，则进一步退化为简单回热循环。多再压缩循环热效率的增加归功于再压缩单元中的分流作用（流向再压缩压气机的分流比 x），其分流特征以分流向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 表征。图 8 对比了不同再压缩循环的热效率。从图 8（a）中可以看到，在简单回路的基础上引入一个再压缩单元构建再压缩循环，能够大幅提高循环热效率，最大热效率增幅超过 7.5%。同样地，从图 8（b）和图 8（c）中可以看到，继续引入再压缩单元后，循环的热效率可进一步提高，其中双再压缩循环较单再压缩循环的最大热效率增幅约为 1.5%。然而，随着再压缩回路数的增加，循环热效率提升幅度逐渐降低，同时分流向量可行域持续变小。因此，双再压缩循环构型是一种较优的选择，其在不大幅增加构型复杂度的条件下可获得明显的热效率收益，同时分流向量可行域较宽。

2.2　循环建模及分析方法

循环建模在于构建各部件的热力模型方程组，而循环分析则通过求解部件热力方程组以确定各状态点参数和热力-经济性能。准确地预测 SCO_2 的物性（尤其是临界点附近）是进行循环热力建模与分析的基础。目前，开源代码 NIST REFPROP[56] 被广泛采用，其提供了专门针对二氧化碳而建立的 SW EOS 物性模型[57]，具备非常高的物性预测精度，尤其是临界点附近[58-59]。下

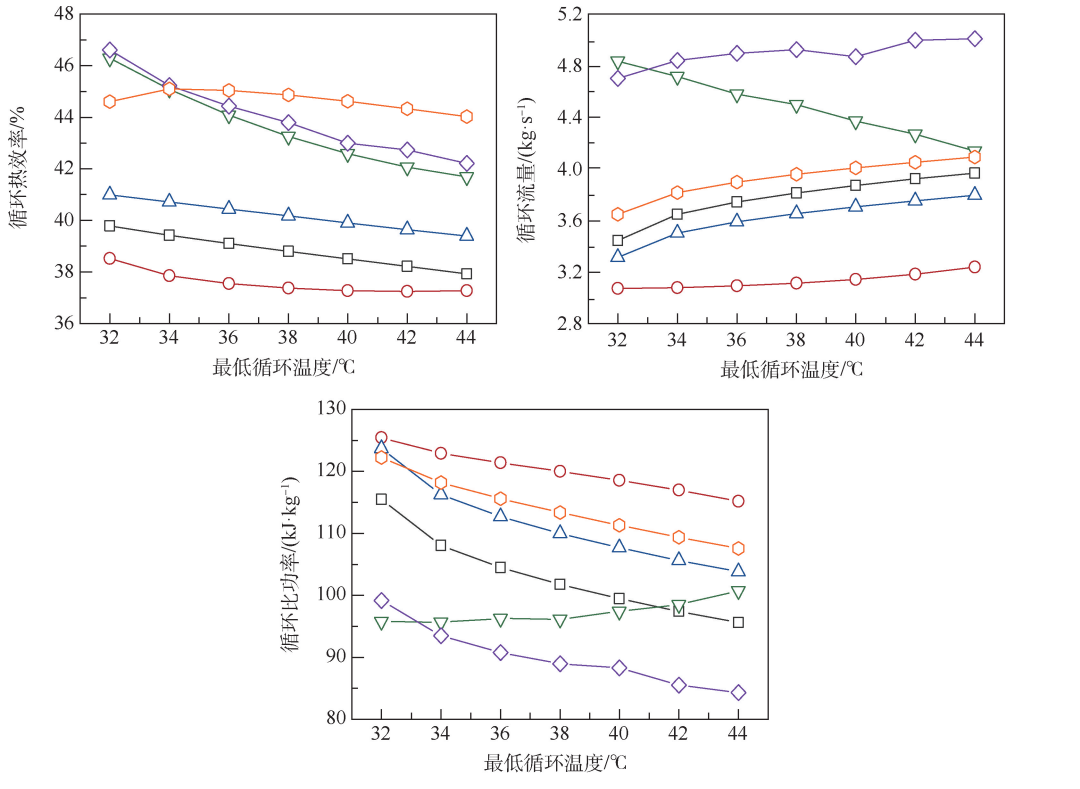


图6 典型的 SCO_2 闭式布莱顿循环性能对比[55]

Fig. 6 Performance comparison of typical close SCO_2 Brayton cycles[55]

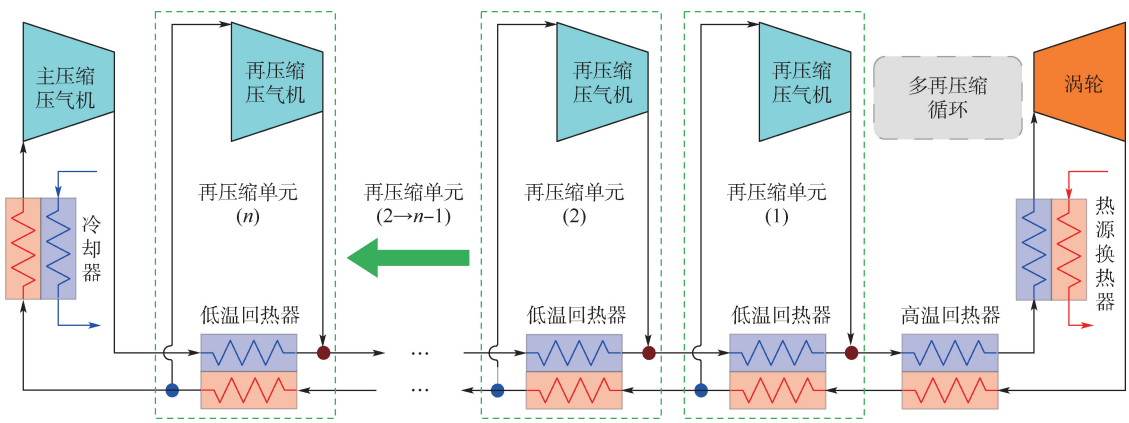


图7 多再压缩循环构型[55]

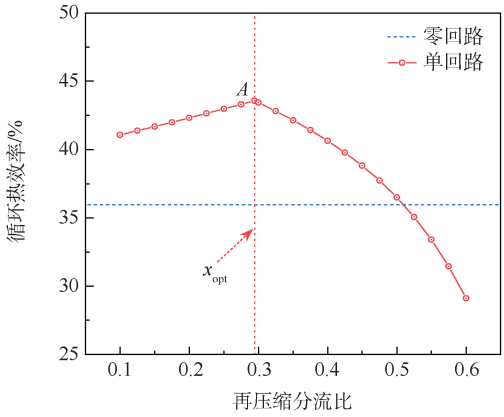
Fig. 7 Layout of multi-recompression cycle[55]

面对部件建模和循环性能分析方法进行简要总结。

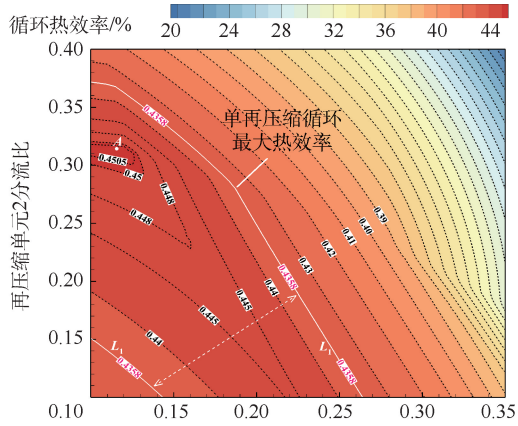
2.2.1 “零维”建模方法

当前,“零维”建模方法被广泛用于 SCO_2 闭式布莱顿循环热力建模与分析。该方法将各部件视为“黑匣子”,建模时仅关注部件总性能和进出口的热力参数,并不关注内部参数分布及尺寸特征。建模过程实际上通过求解流量方程和能量方程,在给定的进口边界条件下确定部件出口边界条件。按照工作原理的不同,各部件可分为三大类:叶轮机(压气机、涡轮)、换热器(回热器、热源换热器、冷却器)和管路系统(管道、三通)。表1总结了不同部件的“零维”热力学模型。表中,$H=\dot{m}h$ 和 $\dot{E}=\dot{m}\dot{e}$ 分别代表总焓和总㶲,W 为叶轮机功率,$\Delta\dot{Q}$ 为总换热量,$\dot{m}$ 为质量流量,h 为比焓,$\dot{e}$ 为比㶲,下标 inl 和 out 分别表示进口和出口,C 表示压气机,T 表示涡轮,c 和 h 分别表示换热器冷侧和热侧。

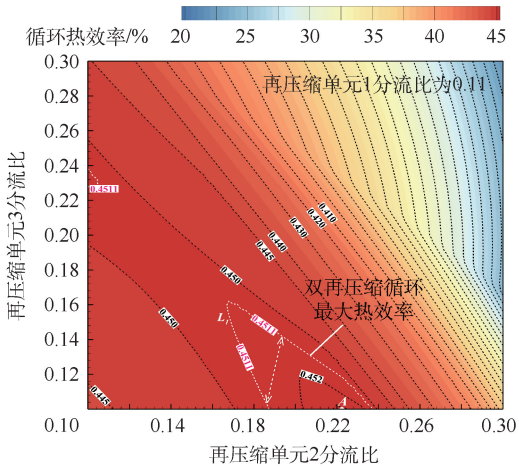
对叶轮机而言,在等熵压缩/膨胀假设下,其出口参数按式(1)和式(2)确定,式中 η_{is} 为等熵效率。Dyreby 等指出[60-61],等熵效率对叶轮机总压比存在依赖性,因此多变效率 η_{po} 更适合用于叶轮机建模。此时,将整个压缩/膨胀过程分为 n 个子过程,每个子过程中的总压比按式(3)求得,而



(a) 零回路与单回路对比



(b) 单回路与双回路对比



(c) 双回路与三回路对比

图 8　不同再压缩单元数时多再压缩循环性能对比[55]

Fig. 8　Performance comparison of multi-recompression cycle with different recompression units[55]

表 1　循环部件“零维”热力学模型

Table 1　0D thermodynamic model for cycle components

部件	模型
压气机	$\dot{m}_{inl}=\dot{m}_{out}$ $H_{inl}=H_{out}+W_C$
涡轮	$\dot{m}_{inl}=\dot{m}_{out}$ $H_{inl}=H_{out}+W_T$
换热器	$\dot{m}_{inl,c}=\dot{m}_{out,c}$ $\dot{m}_{inl,h}=\dot{m}_{out,h}$ $\lvert\Delta\dot{Q}\rvert=H_{out,c}-H_{inl,c}$ $\lvert\Delta\dot{Q}\rvert=H_{inl,h}-H_{out,h}$
n_1个进口，n_2个出口 $n_1=n_2=1$:管道 $n_1=1,n_2=2$:分流三通 $n_1=2,n_2=1$:合流三通	$\sum_{i=1}^{n_1}\dot{m}_{inl}\big\rvert^{i}=\sum_{i=1}^{n_2}\dot{m}_{out}\big\rvert^{j}$ $\sum_{i=1}^{n_1}H_{inl}\big\rvert^{i}=\sum_{i=1}^{n_2}H_{out}\big\rvert^{j}$

其等熵效率则等于多变效率。

$$\begin{cases}p_{out,C}=p_{inl,C}\pi\\ p_{out,T}=p_{inl,T}/\pi\end{cases}\tag{1}$$

$$\begin{cases}h_{out,C}=h_{inl,C}-(h_{inl,C}-h_{out,is,C})/\eta_{is}\\ h_{out,T}=h_{inl,T}-(h_{inl,T}-h_{out,is,T})\eta_{is}\end{cases}\tag{2}$$

$$\pi_i=\sqrt[n]{\pi}\qquad i=1,2,\cdots,n\tag{3}$$

式中：p 为压力；π 为叶轮机增压比。

对换热器而言，一般指定冷/热侧进口参数、压力损失系数 c_p、效能 ε 或导热率 UA。效能的一种通用表述为焓差效能 $\varepsilon_{\Delta h}$，其定义为实际换热量 $|\Delta\dot{Q}|$ 与理论最大换热量 $\Delta\dot{Q}_{max}$ 之比，如式(4)所示。此外，部分研究[62]也采用温差效能 $\varepsilon_{\Delta t}$，其定义为热侧进出口温差与冷热两侧进口温差之比，如式(5)所示。然而，由于 SCO_2 工质在临界点附近物性变化剧烈，采用温差效能 $\varepsilon_{\Delta t}$ 进行换热器模拟可能导致不符合物理定律的结果，如冷/热侧换热量绝对值不相等。

$$\varepsilon_{\Delta h}=\frac{|\Delta\dot{Q}|}{\Delta\dot{Q}_{max}}=\frac{|\Delta\dot{Q}|}{\min(|\Delta\dot{Q}_{max,h}|,|\Delta\dot{Q}_{max,c}|)}\tag{4}$$

$$\varepsilon_{\Delta t} = (t_{\text{inl,h}} - t_{\text{out,h}})/(t_{\text{inl,h}} - t_{\text{inl,c}}) \tag{5}$$

Dyreby 等[60-61]指出,给定效能进行换热器模拟有可能导致回热器在不同设计工况下的导热率相差显著,进而造成显著的尺寸和成本差异;而导热率直接关系到回热器的尺寸和成本,其更适用于不同循环构型的性能(尤其是经济性能)对比分析。此时,将换热器沿管程离散为多个子换热单元,认为每个子单元内的换热量和压力损失相等,通过迭代求解每个子单元内的温度和压力,最终可得到换热器沿程的所有参数分布,如此可更为准确地考虑到二氧化碳物性变化的影响(尤其是低温回热器存在夹点问题时)。

2.2.2 循环性能分析

能量分析以热力学第一定律为基础,立足于能量的数量,以热效率 η_{cyc} 为指标,其定义为净输出功率占从热源吸热量的比例,即

$$\eta_{\text{cyc}} = \frac{\sum W_{\text{T}}\big|^{i} + \sum W_{\text{C}}\big|^{i}}{\Delta\dot{Q}_{\text{inj}}} = \frac{|\Delta\dot{Q}_{\text{inj}}| - |\Delta\dot{Q}_{\text{rej}}|}{|\Delta\dot{Q}_{\text{inj}}|} \tag{6}$$

式中:$|\Delta\dot{Q}_{\text{inj}}|$ 和 $|\Delta\dot{Q}_{\text{rej}}|$ 分别为吸热量和放热量,二者之差即为循环净输出功率。

㶲分析以热力学第二定律为基础,立足于能量的数量和质量,以㶲效率 ε_{cyc} 为指标,其定义为净输出功率占进入循环的热量㶲的比例[63]:

$$\varepsilon_{\text{cyc}} = \frac{\sum W_{\text{T}}\big|^{i} + \sum W_{\text{C}}\big|^{i}}{\dot{E}_{Q}} \tag{7}$$

式中:$\dot{E}_{Q}$ 为热量㶲,其决定于吸热量、热源温度 t_{r} 和环境温度 t_0:

$$\dot{E}_{Q} = |\Delta\dot{Q}_{\text{inj}}|(1 - t_0/t_{\text{r}}) \tag{8}$$

经济性分析基于经济性评估模型量化单位收益的成本,常以平准化度电成本 LCOE 为评价指标,其定义为

$$\text{LCOE} = \frac{\text{TAC}}{\tau\dot{W}_{\text{net}}} \tag{9}$$

式中:τ 为年平均运行时间,h;$\dot{W}_{\text{net}}$ 为净输出电功率,kWe;TAC 为总年费用。总年费用由投资成本 C_{inv} 和运行维护成本 C_{mt} 组成,而运行维护成本与投资成本正相关,即[64]

$$\text{TAC} = C_{\text{inv}} + C_{\text{mt}} = (1 + 0.05)C_{\text{inv}} \tag{10}$$

式中:投资成本 C_{inv} 主要包含部件研发成本 C_{eq} 和系统选址建造成本 C_{site},即

$$C_{\text{inv}} = \text{CRF}(C_{\text{eq}} + C_{\text{site}}) = \text{CRF}(1 + 0.2)C_{\text{eq}} \tag{11}$$

式中:系统选址建造成本与研发成本正相关,资金回收系数 CRF 则与回报率 i_{r} 和系统生命周期 l 有关。

$$\text{CRF} = \frac{i_{\text{r}}(1 + i_{\text{r}})^{l}}{(1 + i_{\text{r}})^{l} - 1} \tag{12}$$

一般地,回报率 i_{r} 可取 12%,而生命周期 l 可取 20 年[65-66]。部件研发成本直接决定于各部件的热力学性能参数,具体如表 2 所示。

表 2 循环部件经济性模型

Table 2 Economic model for cycle components

部件	成本公式
回热器	$2\,500\text{UA}$[64]
冷却器	$2\,300\text{UA}$[67-68]
热源换热器	$3\,500\text{UA}$[67-68] 或 $283\Delta\dot{Q}$[66,69]
压气机	$6\,898W^{0.7865}$[68] 或 $71.1\dot{m}(0.92-\eta)^{-1}\pi\ln\pi$[66,69]
涡轮	$7\,790W^{0.6842}$[68] 或 $479.34\dot{m}(0.93-\eta)^{-1}\ln\pi(1+\mathrm{e}^{0.036t_{\text{inl}}-54.4})$[66,69]

注:模型中成本单位为美元,导热率 UA 单位为 kW/K,换热量 $\Delta\dot{Q}$ 和功率 W 单位为 kW。

2.3 循环总体热力设计方法

目前,大多数循环总体热力设计均基于“零维”部件模型的“零维”设计方法,该方法可结合优化算法快速实现初步热力方案设计与优化。此外,为考虑循环与部件的耦合效应,部分研究将部件低维设计集成至循环总体设计中,发展了循环总体“一维”耦合设计方法。下面对这 2 种方法进行简要总结。

2.3.1 循环总体“零维”设计/优化

一般而言,循环总体热力设计是复杂的多参数优化问题,确定较优的设计参数组合是总体设计的首要任务。由于“零维”方法简单高效、鲁棒性强,借助其可快速进行参数影响规律分析和大量方案对比,进而总结参数选取准则并确定较优的设计参数组合。基于“零维”方法,Sarkar 和 Bhattacharyya[62]对再压缩再热循环进行了优化设计,指出再热可带来 3.5% 的热效率增益;Cheng 等[70]对再压缩循环进行了全局优化,基于大量数据统计建立了优化设计准则;Akbari 和 Mahmoudi[71]对再压缩-有机朗肯循环进行了热力-经济性优化设计。此外,“零维”方法的简单和高鲁棒性特点也可与优化方法相结合实现循环总体方案快速优化设计[72]。

北航团队提出了一种基于粒子群算法的再压缩循环“零维”优化设计方法[73]。该方法以循环热效率为目标,以合流三通两进口温差为限制条件,将再压缩循环总体设计转化为约束优化问题:

$$\max \quad f(\boldsymbol{a}) = \eta_{\mathrm{cyc}}(\boldsymbol{a})$$
$$\text{s.t.} \quad g(\boldsymbol{a}) = |\Delta t(\boldsymbol{a})| \leqslant \Delta t_{\max} \tag{13}$$

式中:优化参数向量 $\boldsymbol{a}$ 包括循环最低压力 $p_{\min}$、增压比 π_{cyc} 和分流比 x;$g(\boldsymbol{a})$ 为温差不等式约束条件,其中 $\Delta t_{\max}$ 为合流三通两进口最大允许温差。合流三通温差限制的引入可有效考虑大温差条件下三通热条带现象可能导致的热疲劳失效,对部件寿命和循环运行安全性十分必要。

图 9 给出了不同温差限制条件下的再压缩循环优化设计结果,图中 A 点和 B 点分别为粒子群算法寻优得到的最大温差限制 $\Delta t_{\max}=5$℃ 和 $\Delta t_{\max}=15$℃ 时循环的最优设计点。可以看到,A 点和 B 点与合流三通进口温差等高线(黑色虚线)所反映出的 $\Delta t_{\max}=5$℃ 和 $\Delta t_{\max}=15$℃ 温差限制下循环能达到的最高热效率点十分吻合,说明该优化设计方法能有效地寻找到不同温差限制下的循环最优设计点。

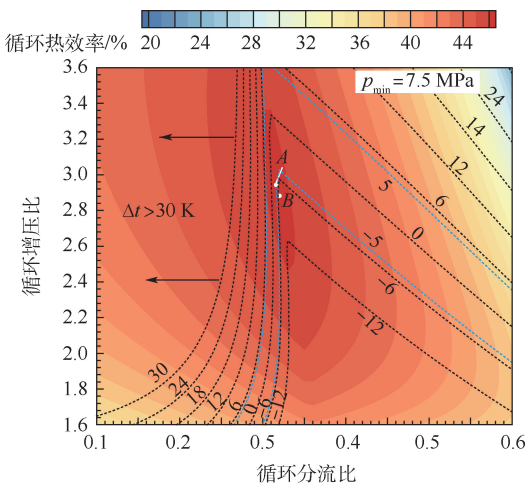


图 9　再压缩循环“零维”优化设计结果[73]

Fig. 9　0D optimal design result of recompression cycle[73]

2.3.2　循环总体“一维”耦合设计

SCO_2 闭式布莱顿循环总体与部件间存在强烈的耦合作用:一方面体现为循环性能高度依赖并直接决定于部件性能[5,70],另一方面部件性能又受循环参数(尤其是功率量级)的显著影响[74-75]。在“零维”设计方法中,总体设计阶段无法考虑此种耦合作用,为此发展出总体“一维”耦合设计方法,其核心思想是将部件初步设计(或“一维”模型)涵盖至总体设计。Dostal 等[14]将回热器初步设计引入循环方案设计与分析中,通过换热与压降模型来评估回热器性能;类似的回热器初步设计方法也被用于再压缩循环设计和温度夹点的影响研究[76-77];Saeed 等[78-79]则将涡轮初始设计集成进循环设计中,并指出采用“零维”涡轮模型与“一维”涡轮设计模型所预测的最大循环热效率偏差高达 4.14%。

北航团队发展了一种适用于工程应用的 SCO_2 闭式布莱顿循环“一维”设计方法[80]。该方法将叶轮机和换热器部件初步设计耦合至总体设计,一方面可在总体设计阶段准确定量评估部件性能,另一方面又能反馈部件的尺寸和内部参数分布等重要信息。以图 10 所示的再压缩循环为例,压气机和涡轮部件采用径流式构型,初步设计时采用损失模型预测其热效率;回热器采用逆流式直通道印刷电路板式换热器,基于离散换热微元法利用换热和压降模型求得沿程参数分布。利用该“一维”耦合设计方法分别对 500 kW(CYC-A)和 5 MW(CYC-B)再压缩循环进行了设计,结果如表 3 所示。可以看到,循环流量的增加使得大功率循环 CYC-B 中叶轮机部件效率均得到不同幅度的提高,并最终使得循环热效率提高 1.0%。这表明该“一维”方法能有效考虑循环总体与部件间的强耦合作用。

2.4　循环非设计状态仿真方法

SCO_2 闭式布莱顿循环非设计状态性能仿真依赖于对部件非设计性能的预测。对叶轮机部件而言,一般基于特性曲线图预测其非设计状态性能。一种简单的做法是通过特性曲线拟合出性能与特征参数的准经验关系式[60,81],但其准确度和普适性非常低;另一种通用做法是通过折合特性图换算得到叶轮机的非设计性能,然而目前多数研究均采用传统空气叶轮机的折合准则[82],其并不能有效考虑到 SCO_2 工质的强真实气体效应的影响。研究表明,SCO_2 的强真实气体效应会对压气机的非设计状态性能产生显著的影响[83]。Carstens[84]认为该影响主要反映在密度的不同上,通过引入密度修正来计算叶轮机在非设计状态下的物理流量和增压比,然而仅考虑密度这一单一因素并不够。

为此,北航团队发展了一种能够较为全面地考虑工质真实气体影响的叶轮机非设计性能预测方法[55]。该方法首先引入压缩因子 Z_t 和等熵指数 n_s 建立起如下气体工质叶轮机通用折合准则:

$$\begin{cases} \dot{m}_{\mathrm{cor}} = \dfrac{\dot{m}\sqrt{Z_t t_t}}{p_t\sqrt{n_s}} \\ N_{\mathrm{cor}} = \dfrac{N}{\sqrt{n_s Z_t t_t}} \end{cases} \tag{14}$$

式中:$\dot{m}_{\mathrm{cor}}$ 为折合流量;N_{cor} 为折合转速;t_t 和 p_t 分别为进口总温和总压;N 为物理转速。

压缩因子 Z_t 和等熵指数 n_s 为表征真实气体

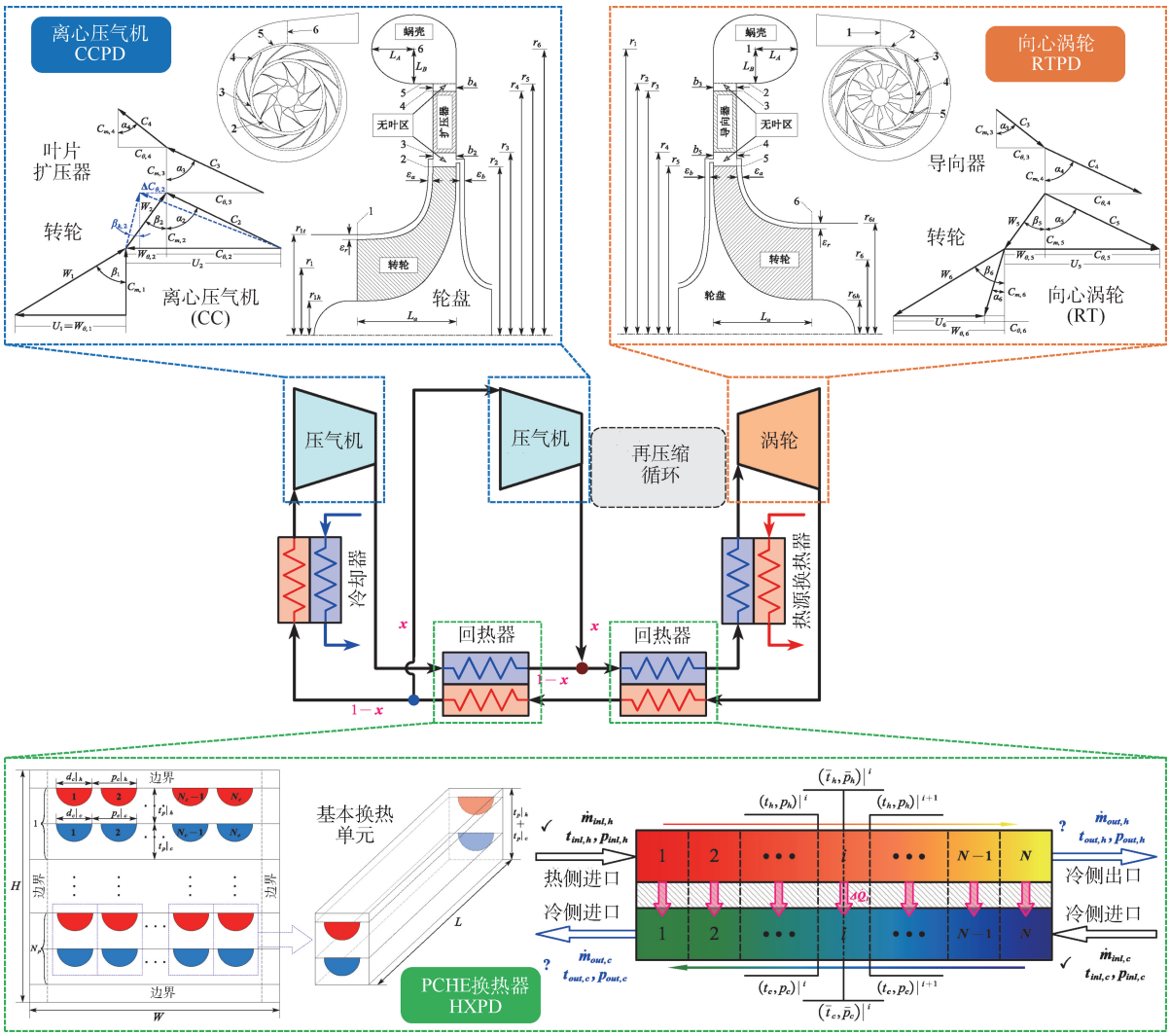


图 10 再压缩循环“一维”耦合设计方法

Fig. 10 1D coupled design method of recompression cycle

表 3 再压缩循环“一维”耦合设计结果

Table 3 1D coupled design results of recompression cycle

循环	循环流量/(kg·s^{-1})	循环热效率/%	主压缩压气机等熵效率/%	再压缩压气机等熵效率/%	涡轮等熵效率/%
CYC-A	7.85	37.00	78.67	78.58	82.44
CYC-B	75.49	38.00	81.00	81.76	83.37
相对增量/%	~860	1.00	2.33	3.18	0.93

效应的关键参数，如图 11 所示。对理想气体而言，压缩因子 $Z_t=1$，等熵指数 $n_s=\gamma$（γ 为比热比，常数），式(14)退化为传统空气工质叶轮机折合准则：

$$\begin{cases}\dot{m}_{\mathrm{cor}}=\dfrac{\dot{m}\sqrt{t_t}}{p_t\sqrt{\gamma}}\\ N_{\mathrm{cor}}=\dfrac{N}{\sqrt{\gamma}\sqrt{t_t}}\end{cases}\tag{15}$$

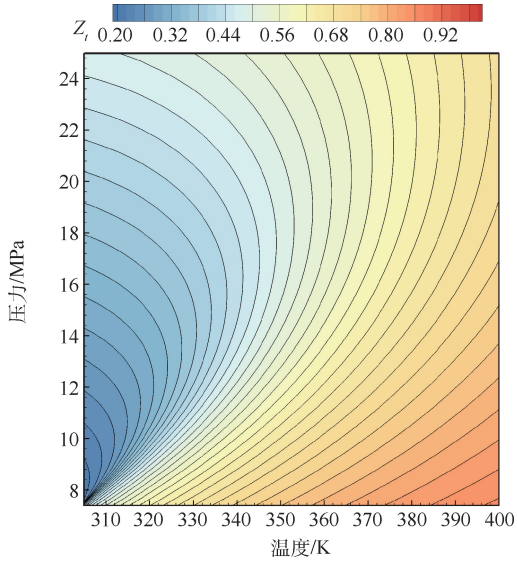
基于上述折合准则，采用特性图插值的方法便可快速计算出 SCO_2 叶轮机非设计性能，具体如下：

1）根据实验或仿真结果绘制特性曲线图，根据设计点进口总参数采用折合准则绘制折合特性图。

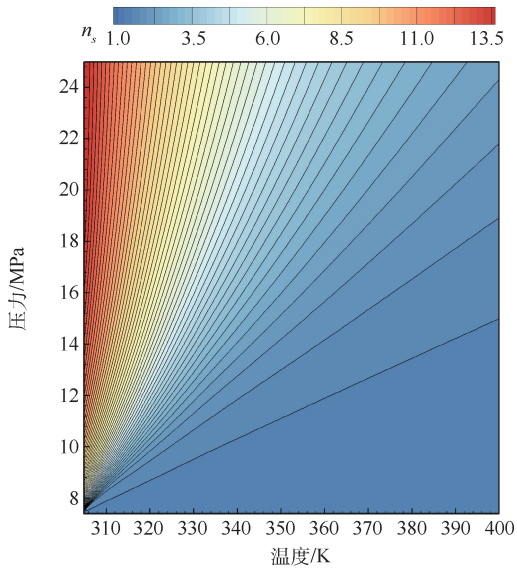
2）按照非设计状态进口总温/总压和物理转速计算此状态的折合转速，在折合特性图上插值得到非设计状态的折合流量、增压比和效率。

3）根据折合流量反算出非设计状态下的物理流量。

对换热器而言，若其设计方案（即形式和几何尺寸）已知，可利用正问题方法计算其非设计状态性能。若在总体分析阶段仅知道其设计状态性能而未知其形式和几何尺寸，则可采用基于性



(a) 压缩因子Z_t



(b) 等熵指数n_s

图 11　SCO_2工质真实气体效应表征参数

Fig. 11　Parameters for real gas effect of SCO_2

能缩放的方法[85]进行换热器非设计性能预测。非设计状态下,换热器冷/热侧局部换热系数 α 与面积 A 之积根据设计值按式(16)进行缩放:

$$(\alpha A)_{\text{off}} = (\alpha A)_{\text{on}} \left(\frac{\lambda_{\text{off}}}{\lambda_{\text{on}}}\right)\left(\frac{\dot{m}_{\text{off}}/\mu_{\text{off}}}{\dot{m}_{\text{on}}/\mu_{\text{on}}}\right)^{a_1}\left(\frac{Pr_{\text{off}}}{Pr_{\text{on}}}\right)^{a_2} \tag{16}$$

式(16)基于换热模型 $\alpha \propto \lambda Re^{a_1} Pr^{a_2}$ 得到。式中:λ 为导热系数;μ 为动力黏度;Pr 为普朗特数;下标 on 和 off 分别表示设计状态和非设计状态;指数 a_1 和 a_2 根据不同换热模型指定。

换热器子单元的导热率 UA 表示为

$$\frac{1}{(\text{UA})_{\text{off}}} = \frac{1}{(\alpha A)_{\text{off,c}}} + \frac{1}{(\alpha A)_{\text{off,h}}} \tag{17}$$

类似地,基于摩擦压力损失 $\Delta p \propto (\Delta l/d_{\text{h}})\rho u^2$,非设计状态下换热器冷/热侧压力损失可按式(18)进行缩放:

$$\Delta p_{\text{off}} = \Delta p_{\text{on}} \frac{\dot{m}^2_{\text{off}}/\rho_{\text{off}}}{\dot{m}^2_{\text{on}}/\rho_{\text{on}}} \tag{18}$$

式中:ρ 为密度。

3　SCO_2叶轮机技术

临界点附近的低压缩功耗是 SCO_2 闭式布莱顿循环巨大性能优势的根本原因,这一优势的实现与 SCO_2 叶轮机密不可分。SCO_2 工质的物性迥异于常规工质,特别是其在临界点附近所表现出的强真实气体效应,使得叶轮机内部会出现一些特殊的流动特征,同时也给叶轮机的气动设计带来了挑战。本节对 SCO_2 叶轮机气动设计技术、相似技术及现阶段国内外相关试验研究进展进行了介绍。

3.1　SCO_2工质物性对叶轮机的影响

如图 1 所示,临界点附近二氧化碳定压比热陡然增大,而压缩因子 Z_t 急剧减小,这意味着压气机能以更低的温升获得更高的压比。同时,二氧化碳的比热和焓同时受温度和压力的影响,压缩因子亦如此,且其值远小于 1,这意味着常规设计方法中所采用的理想气体假设与定比热假设均不再适用,在进行叶轮机气动设计时需另行考虑。此外,二氧化碳的黏性低,使得叶轮机具有极高的雷诺数(可达 10^7),对于工作在临界点附近的压气机其雷诺数甚至可以达到 10^8。更高的雷诺数意味着更小的摩擦损失,对于粗糙度低的叶轮机而言这一优势是较为明显的。

对于工作在临界点附近的压气机而言,还存在一个特殊的流动现象——进口冷凝相变,如图 12所示。受叶片前缘曲率的影响,SCO_2 工质在压气机进口叶尖吸力面区域存在局部加速,加速导致其静温静压降低,并极可能降至饱和线以下

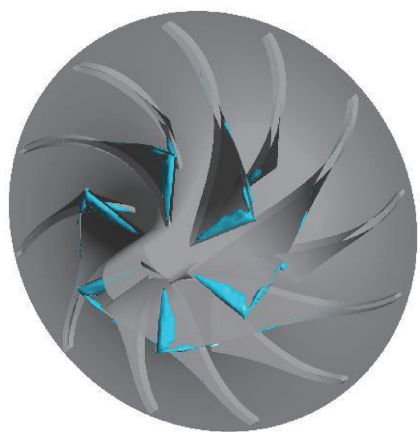


图 12　SCO_2压气机进口局部冷凝相变

Fig. 12　Local condensation phrase change at SCO_2 compressor inlet

而落入两相区。部分试验表明,压气机进口流动大面积进入冷凝区会降低其效率[86-87],也有试验发现冷凝对压气机效率没有影响[88],关于冷凝对压气机性能的影响还有待展开更加深入和广泛的试验研究。至于冷凝对压气机稳定性的影响,就现有试验结果来看,进口工作在亚临界区的压缩基本都能稳定运行,未出现明显的振动[88-89]。

值得说明的是,工质物性的影响不仅局限于气动方面,SCO_2的高压、高密度工作环境还对转子的支撑方案、封严方式等结构设计带来了挑战,其特殊的化学性质也对加工制造的选材提出了一些要求,这些都需要在具体应用中进行相关经验的积累。

3.2 SCO_2叶轮机气动设计技术

SCO_2叶轮机的气动设计需要结合设计手段与分析手段进行反复迭代调试优化,其基本步骤如图13所示。首先,在一维层面借助各种经验关系进行设计参数的选择,并利用一维性能预测方

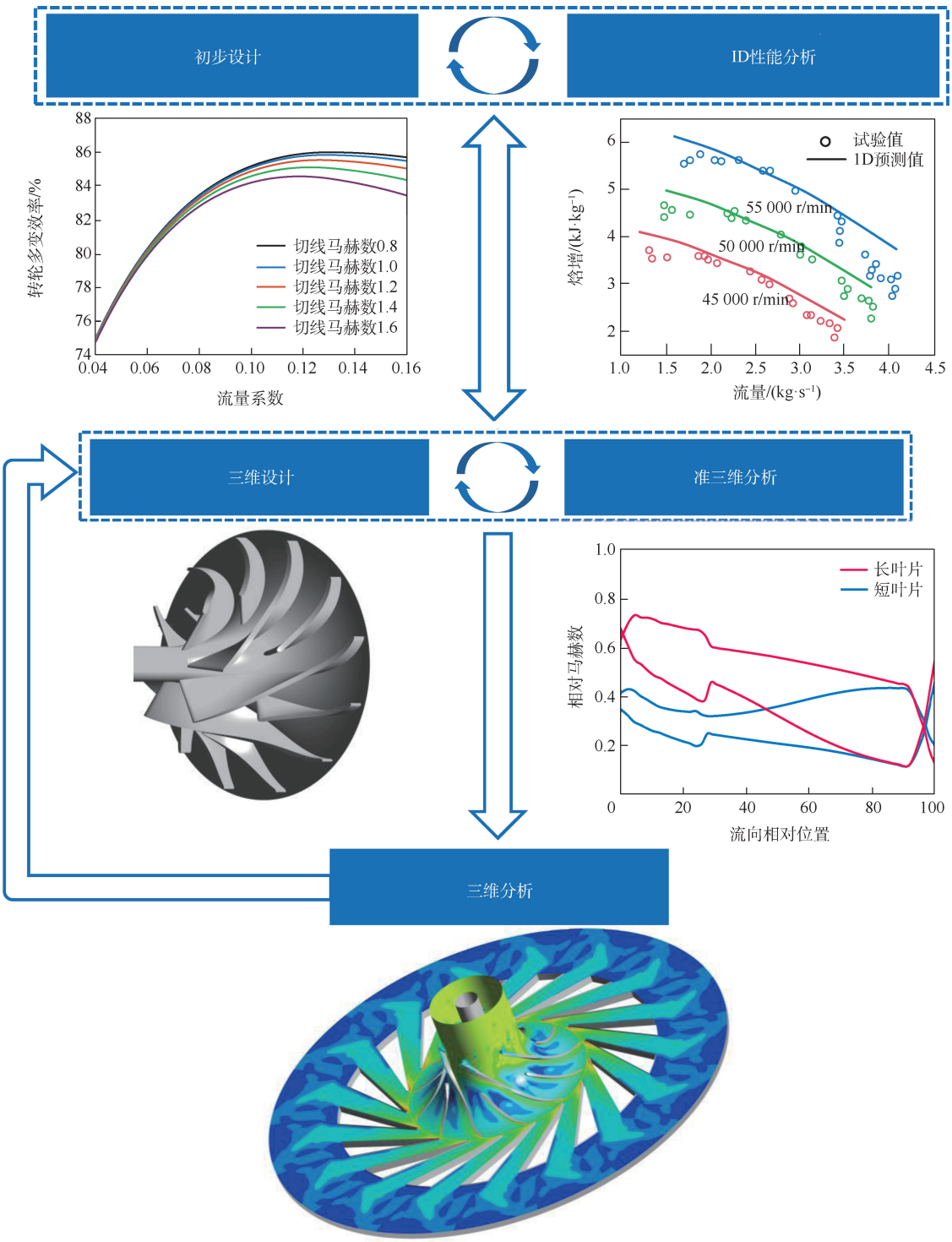


图13 SCO_2叶轮机气动设计流程

Fig. 13 Aerodynamic design process of SCO_2 turbomachinery

法进行性能评估,据此进行参数调整,以确定一组最优的速度三角形,为后续设计提供一个良好的起点;然后,在一维几何参数的基础上进行细节补充,借助准三维分析工具进行三维几何的优化选取;最后,使用三维数值模拟手段进行更精细的性能分析和细节调整。

3.2.1　一维性能预测方法

在 SCO_2 叶轮机气动设计流程中,一维层面的速度三角形设计直接决定着设计方案的优劣,而速度三角形的优化则依赖于一维性能预测方法。为此,发展有效可靠的一维性能预测方法(关键是低维流动损失模型)是 SCO_2 叶轮机气动设计的关键。

北航团队分别针对 SCO_2 径流式叶轮机建立了低维损失模型,并在此基础上发展了一维性能预测方法。对离心压气机而言,所发展的混合模型[90]结合了传统一区模型[91-92]和二区模型[93-94]的优点,其将压气机内部流动划分为射流区和尾迹区,对于射流区采用一区模型的分析思路和部分损失模型经验关联式;对尾迹区则引入 2 个状态参数以间接地考虑损失机制尚未明确的端壁损失的影响。该混合模型对射流区和尾迹区的损失预测均基于明确的物理意义,因此具有更高的普适性。对向心涡轮而言,所发展的一区损失模型[95]建立于对现有不同损失经验关联式组合的对比评估,涵盖攻角损失、转子通道损失、叶尖泄漏损失、尾缘损失、鼓风损失等损失项。表 4 总结了离心压气机和向心涡轮低维损失模型的基本方程,图 14 给出了某离心压气机和某向心涡轮的一维特性预测曲线,分别与实验值和三维 CFD 值进行了对比。可以看到,所发展的一维性能预测方法能较为准确地预测 SCO_2 径流式叶轮机性能。

3.2.2　三维几何设计影响

三维几何设计通过几何细节优化,合理地组织叶轮机内部流动,达到控制流动损失的目的。北航团队系统地研究了 SCO_2 径流式叶轮机关键三维几何设计参数的影响规律,给出了关键设计参数的选取准则。

对离心压气机而言,叶轮分流叶片位置、后弯角、叶尖间隙高度等参数对其性能影响最为显著[96]。研究表明,通过合理调整分流叶片流向的起始位置,可有效改善分流叶片表面前缘附近负荷分布,获得叶轮效率的提升,如图 15 所示;后弯角的增大会导致叶轮增压比的明显降低,但其可改善尾缘附近的负荷;由于 SCO_2 离心压气机叶轮出口叶高较小,叶尖间隙高度的影响尤为显著,当出口叶尖间隙从 8.3% 增大至 12.9% 时,叶尖泄漏流量的增加导致间隙泄漏损失增大,尾迹区明显增大,转轮效率降低约 0.7%;同时尾缘附近负荷明显降低,叶轮增压比也有所减小,如图 16 所示。在静子部件方面,采用叶片扩压器可有效降低进入蜗壳的流体速度进而减小蜗壳损失,此外,

表 4　SCO_2 叶轮机一维性能预测模型基本方程

Table 4　Basic equations for 1D performance prediction of SCO_2 turbomachinery

叶轮机	损失项	基本方程
离心压气机	射流区	$P_{02\mathrm{rel}} = P_{02\mathrm{rel,is}} - (P_{01\mathrm{rel}} - P_{01}) \sum_i \omega_i$
	尾迹区	$\kappa, \chi = k_0 + k_1 \cdot Mu + k_2 \cos \beta_{2b} + k_3 (cl/b_2)$
	掺混面	$\dot{m} = \rho_2 A_2 c_{2m}$ $\dot{m}_{\mathrm{jet}} c_{2m,\mathrm{jet}} + \dot{m}_{\mathrm{wake}} c_{2m,\mathrm{wake}} - \dot{m} c_{2m} = P_2 A_2 - P_{2\mathrm{jet}} A_{2b}$ $\dot{m}_{\mathrm{jet}} c_{2u,\mathrm{jet}} + \dot{m}_{\mathrm{wake}} c_{2u,\mathrm{wake}} = \dot{m} c_{2u}$ $\dot{m}_{\mathrm{jet}} H_{02\mathrm{jet}} + \dot{m}_{\mathrm{wake}} H_{02\mathrm{wake}} + \dot{m} (\Delta H_{\mathrm{disc}} + \Delta H_{\mathrm{rec}}) = \dot{m} H_{02}$
向心涡轮	攻角损失	$\begin{cases} 0.5 W_3^2 \sin^2 (\beta_3 - \beta_{3\mathrm{opt}}) & i < 0 \\ 0.5 W_3^2 \sin^3 (\beta_3 - \beta_{3\mathrm{opt}}) & i > 0 \end{cases}$
	转子通道损失	$\frac{1}{2} M_J K_P \left\{ \left(\frac{L_H}{D_H} \right) + 0.68 \left[1 - \left(\frac{R_{4m}}{R_3} \right)^2 \right] \frac{\cos \beta_4}{B_4 / c} \right\} \frac{1}{2} (W_3^2 + W_4^2)$
	叶尖泄漏损失	$0.6 \frac{\varepsilon_r}{B_3} C_{3t} \sqrt{2 \frac{\pi}{Z} C_{3t} C_{4m} \frac{D_{4s}^2 - D_{4h}^2}{(D_3 - D_{4s}) \left(1 + \frac{\rho_3}{\rho_4} \right)}}$
	尾缘损失	$\frac{\rho_5 W_5^2}{2} \left[\frac{N_R t_4}{\pi (R_{4H} + R_{4S}) \cos \beta_4} \right]$
	鼓风损失	$f_{\mathrm{disk}} \frac{\rho_{\mathrm{ave}} R_3^2 U_3^3}{2 \dot{m}}$

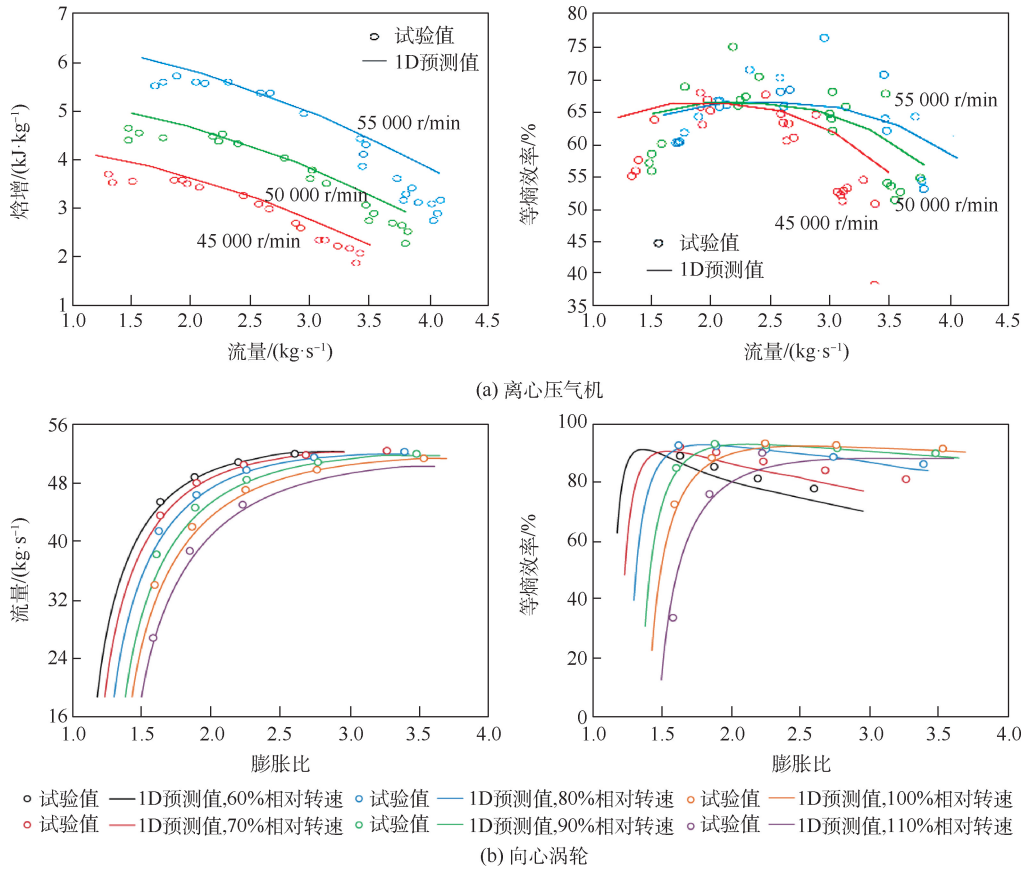


图14 SCO_2叶轮机一维性能预测方法校验

Fig. 14 Validation of 1D performance prediction method for SCO_2 turbomachinery

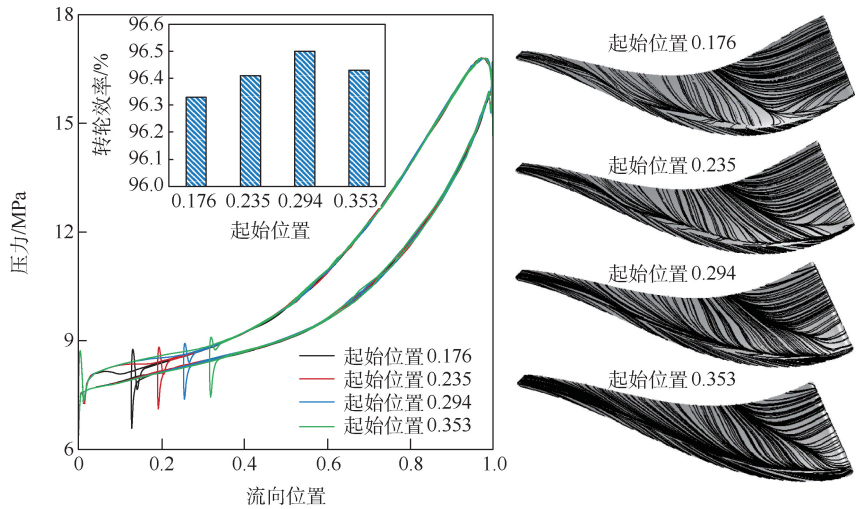


图15 离心压气机分流叶片流向起始位置的影响

Fig. 15 Effect of streamwise location of centrifugal compressor splitter leading edge

通过合理地选择蜗壳截面形式、调整蜗壳进口扩张角及其流通面积与扩压器的匹配可有效降低蜗壳内部流动损失。

对向心涡轮而言,研究发现[95],叶轮子午流道形式(尤其是曲率)对其性能影响显著,合理的流道曲率和流向面积变化可获得较高的转轮效率,如分别采用最大圆弧和多项式构造轮毂和机匣型线并保证中间流向截面面积为进出口面积的平均值;叶轮进口余高可有效抑制泄漏流动,降低泄漏流量和损失,增加15%的余高可提高近0.2%的转轮效率(见图17);存在最佳转轮叶片数使得涡轮效率较优,较小的进口绝对气流角对应较大的最佳叶片数(见图18)。对静子部件而言,采用导向器可改善转轮进口气流条件,选择合适的导向器进口气流能有效降低静子部件损失,然而导向器的引入可能会限制涡轮在高膨胀比下的流量;导向器和蜗壳间无叶段长度会同时影响转轮进口参数展向分布的不均匀性和蜗壳内的流动

速度,其长度的选择需综合考虑转轮效率和静子部件的损失。

此外,清华大学张扬军教授团队针对 SCO_2 轴流涡轮叶片高度尺寸小、泄漏流动强的特征,基于旋涡流动控制方法提出了一种复合弯掠前加载叶片构型,可有效减小动叶泄漏流动损失,提高小流量 SCO_2 涡轮的气动效率,并得到试验的成功验证,如图 19 所示[97]。

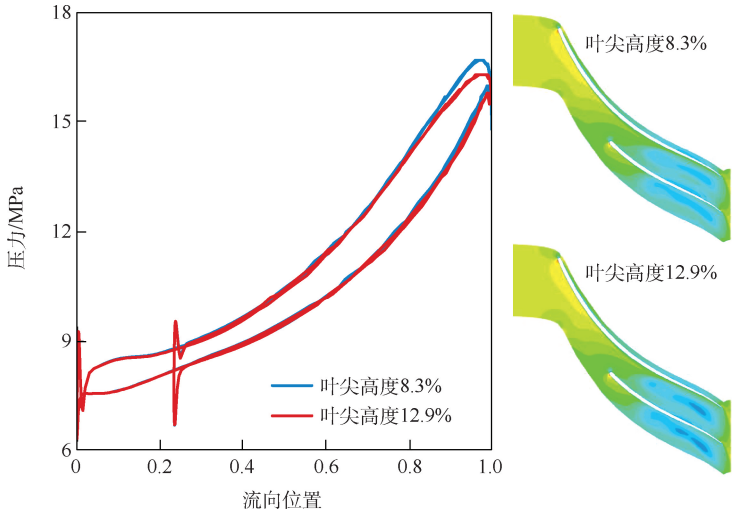


图 16　离心压气机叶轮出口叶尖间隙的影响

Fig. 16　Effect of tip clearance at centrifugal compressor runner outlet

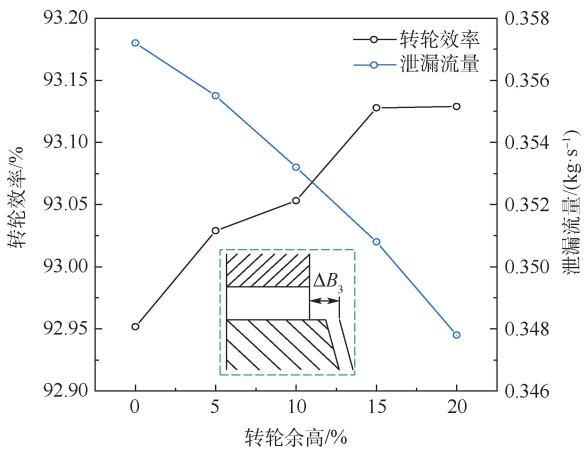


图 17　向心涡轮转轮余高的影响

Fig. 17　Effect of extra height of radial turbine runner

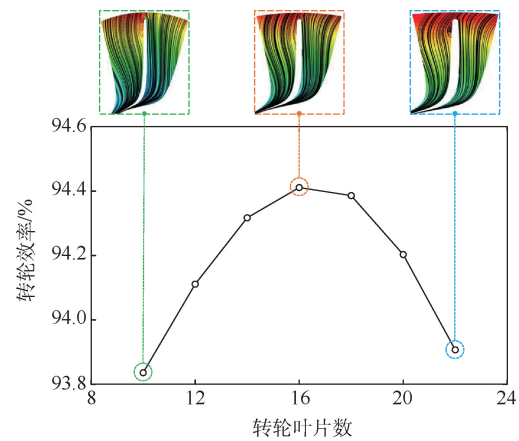


图 18　向心涡轮转轮叶片数的影响

Fig. 18　Effect of blade number of radial turbine runner

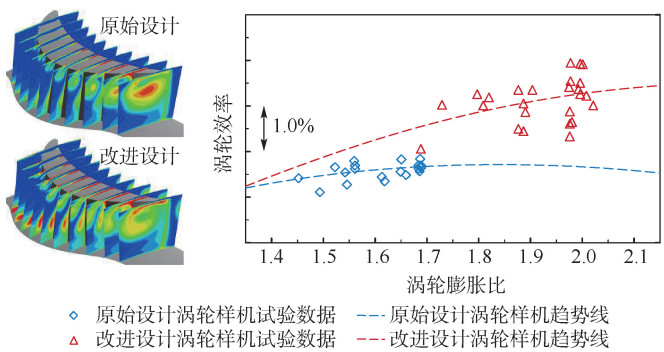


图 19　基于复合弯掠前加载叶片的 SCO_2 涡轮性能改进[97]

Fig. 19　Performance improvement of SCO_2 turbine based on composite curved and swept front-loading blade[97]

3.3　SCO_2 叶轮机性能相似试验技术

开展真实工质下试验测量是获取 SCO_2 叶轮机性能的最可靠方法,然而其却对试验条件与能力要求非常苛刻。①高压工作环境构建与长时运行依赖于闭式循环系统,这会导致试验成本高昂、系统启动复杂和高压气体密封困难等问题;②保证压气机来流温度/压力稳定于临界点附近要求试验系统具备极高的控制精度。为此,如何规避上述问题对 SCO_2 叶轮机性能试验尤为重要。一种简单高效的思路是:基于不同工质叶轮机相似

变换思想[98-99]，通过相似方法建立常规空气叶轮机与 SCO_2 叶轮机间的特性变换准则，如此便可先基于开式循环系统试验得到空气叶轮机特性，而后间接变换得到 SCO_2 叶轮机的特性。

北航团队发展了一种空气叶轮机与 SCO_2 叶轮机间的多变相似变换方法[100]（Poly 方法），该方法基于 PI 定理得到真实气体工质叶轮机的流量相似参数 Π_m 和转速相似参数 Π_U，基于真实气体压缩过程的多变分析推导得到性能转换参数 $\Pi_{\Delta H}$ 和 Π_η，如表 5 所示。值得说明的是，理论推导证明等熵效率 η_{is} 不再适合作为真实气体工质叶轮机与空气叶轮机之间的相似参数，而多变效率 η_{po} 由于不受增压比、焓增及物性的影响，更适合作为一种通用的相似参数。

表 5 真实气体工质叶轮机特性相似参数

Table 5 Similarity parameters for real gas turbomachinery

相似参数	表达式
Π_m	$\dfrac{\dot{m}\sqrt{R_g T_{01}}}{P_{01} D^2} \Big/ \sqrt{\dfrac{n_s}{Z_0}\left(\dfrac{2}{n_s+1}\right)^{\frac{n_s+1}{n_s-1}}}$
Π_U	$ND/\sqrt{n_s Z_0 R_g T_0}$
$\Pi_{\Delta H}$	$\Delta H_0/(n_s Z_{01} R_g T_{01})$
Π_η	η_{po}

由于 SCO_2 涡轮工作于远离临界点工况，上述理论推导结果可直接使用。然而，对 SCO_2 压气机而言，工质近临界点剧烈的物性变化导致上述由理论推导得到的相似参数无法保证 SCO_2 压气机与空气压气机间完全相似，且越靠近临界点不完全相似程度越大。为此，引入深度神经网络（DNN）建立不完全相似度与真实气体效应间的联系，用以对相似变换结果进行修正。具体思路为：首先引入 2 个修正系数用以对相似转换得到的性能参数进行修正，通过理论分析得到 2 个修正系数的影响参数；然后借助一维性能预测方法得到一系列相似的 SCO_2 压气机和空气压气机的数据库，搭建 DNN（见图 20）进行训练；最后得到一种采用 DNN 改进后的相似方法（Poly + DNN）。以一台工作于临界点附近的 SCO_2 压气机为例对该方法进行校验，结果如图 21 所示。图中：圆点表示通过数值模拟直接得到的 SCO_2 压气机特性点，虚线表示由数值模拟得到的空气压气机特性通过 Poly 方法相似变换间接得到的 SCO_2 压气机特性，实线则表示结合深度学习进行相似变换（Poly + DNN）得到的结果。可以看出，利用该方法由空气压气机特性通过相似变换间接得到的特性曲线与 SCO_2 压气机实际特性曲线之间的误差基本不超过 2%，满足工程研制需求。

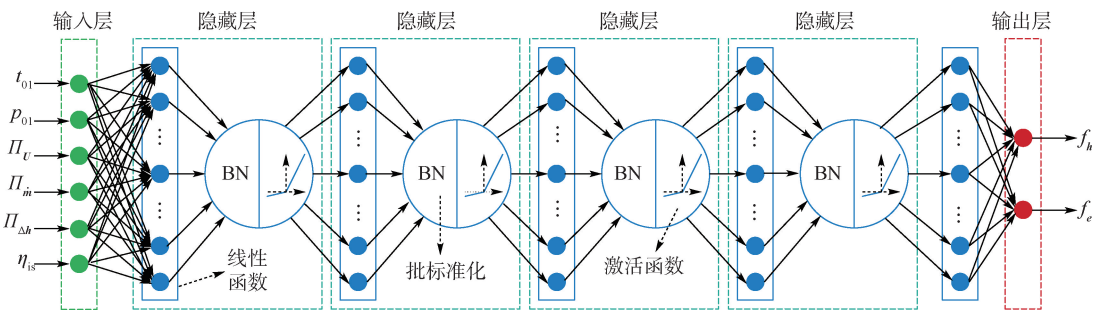


图 20 用以训练数据库的 DNN 结构[100]

Fig. 20 DNN structure for data training[100]

3.4 SCO_2 叶轮机试验研究

在 SCO_2 叶轮机的试验研究方面，国外开展了相对深入和全面的工作。早在 2010 年 SNL 就建成了 250 kW 级再压缩循环试验平台，完成了压气机/涡轮原理验证和特性测试等一系列试验[101-102]，如图 22 所示。由于主压缩压气机工作在临界点附近，试验主要关注主压缩压气机的性能和稳定性。试验中测试的最高转速达到 64 000 r/min，为设计转速的 85.33%，压气机效率基本达到设计预估效率。

KAERI 以一台增压泵为研究对象，在其搭建的循环-部件综合试验平台上对进口液相来流下的泵性能进行了试验测试，并对 SCO_2 增压泵数值模拟问题进行了探讨[45]，如图 23 所示。此后，其与 KAIST、浦项科技大学共同设计并搭建了 100 kW 级 SCO_2 闭式布莱顿循环综合试验测试平台（SCIEL），获得了循环电功率输出，同时完成了压气机低转速性能试验测量[103]。此外，日本 TIT 针对高温气冷堆提出了一种 SCO_2 部分预冷直接循环模式[104-105]，开发了 10 kW 实验室级发电循环，并开展了叶轮机相关试验测量。

国内关于 SCO_2 叶轮机部件试验研究相对较少，公开报道中，中国科学院工程热物理研究所于 2018 年搭建了 1 MW 级压气机试验测试平台，目

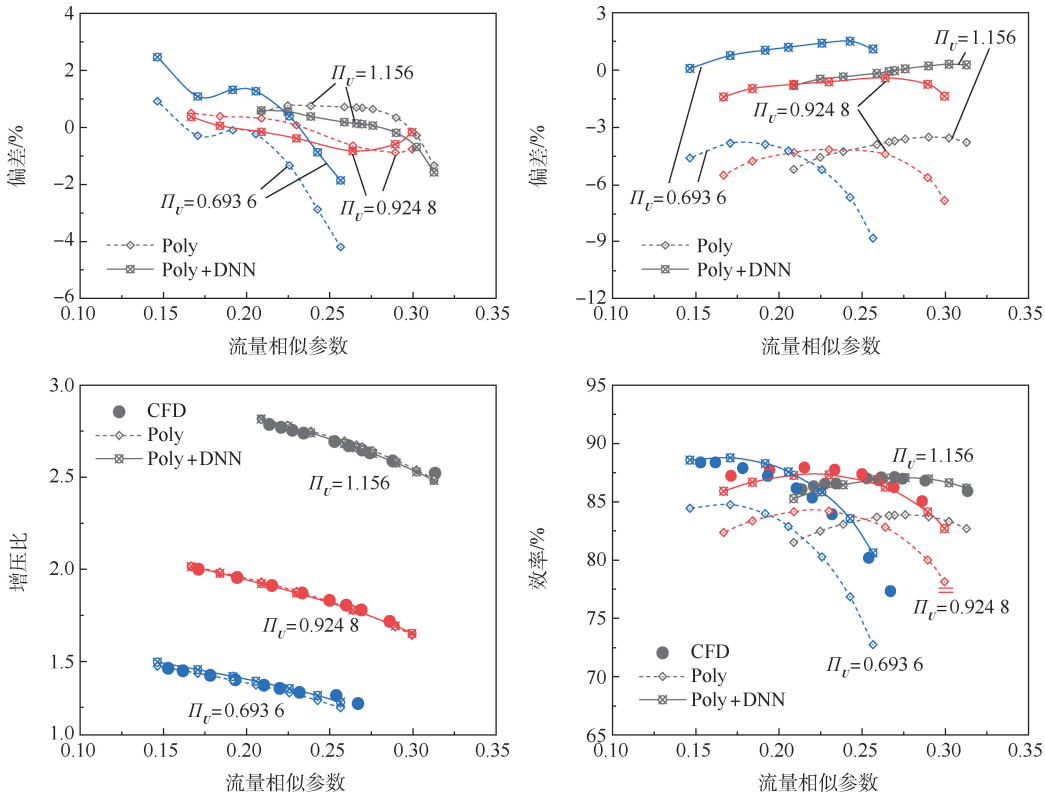
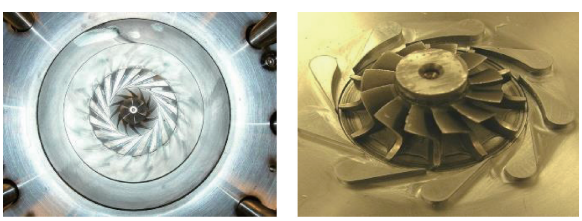


图 21　SCO_2压气机相似方法校验[100]

Fig. 21　Validation of similarity method for SCO_2 compressor[100]



(a) 离心压气机[101]　　(b) 向心涡轮[102]

图 22　SNL 离心压气机和向心涡轮[101-102]

Fig. 22　Centrifugal compressor and radial turbine at SNL[101-102]

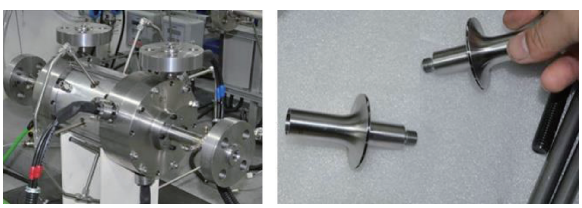


图 23　KAERI 压气机试验[103]

Fig. 23　Compressor experiment at KAERI[103]

前完成了包括饱和态、临界态、超临界态等多进口工况下的全载实验。

4　SCO_2紧凑高效换热器技术

换热器是 SCO_2闭式布莱顿循环实现高热效率、高紧凑度和高可靠性的核心部件。一般地,SCO_2闭式布莱顿循环包含热源换热器、冷却器和回热器。热源换热器是循环工作温度最高的部件,其从高温热源吸热将工质加热至指定的涡轮入口温度;冷却器将工质热量传递给水或空气等冷却工质,实现指定的压气机入口温度;回热器的冷热两侧流体均为 SCO_2工质,通过热量回收提升循环系统的热效率。上述 3 类换热器应具有超紧凑、高效能、低流阻及强换热等特征,并能在高温/超高压(>500℃,20 MPa)等极端环境中长期服役[106]。在众多换热器形式中,紧凑式换热器(单位体积换热面积大于 700 m^2/m^3)以其比表面积大、换热能力强、结构完整性好等独特优势,逐渐成为 SCO_2闭式布莱顿循环系统中诸多换热器的最佳选择[107]。图 24 给出了几种具有较大应用前景的紧凑式换热器,包括印刷电路板式换热器(PCHE)[108]、微细管式换热器(MSTE)[109]、板翅式换热器(PFHE)[110]和增材制造式换热器(AM-HX)[111]。可以发现,微小尺度换热通道是上述几类换热器的基本特征,其不仅有利于满足高温/超高压条件的工作要求,还能通过增加换热面积、提升换热系数的方式增强换热性能,如图 24 所示。

极端的工作环境和严苛的性能要求无疑对换热器设计、制造和实验技术等方面提出严峻挑战。SCO_2工质在换热器内存在的强物性变化特征会加剧浮升力和热加速的影响,使得传统的常物性

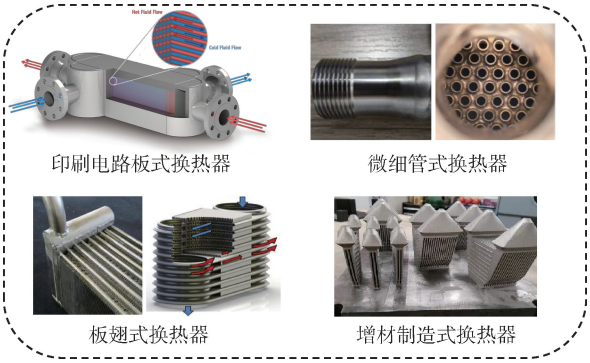


图 24 应用于 SCO_2 闭式布莱顿循环系统的紧凑式换热器

Fig. 24 Compact heat exchangers used for closed SCO_2 Brayton cycles system

换热器设计方法(如对数温差法和效能-换热单元数法)和流动换热经验关联式不再适用[112]。超高的热负荷需求也迫使研究人员不断开发更加新颖高效的通道类型,但如何保证新型换热器在极端条件下可靠工作是此类换热器流道成型和焊接需要解决的首要问题[113]。此外,高温超高压环境模拟也是 SCO_2 换热器试验面临的技术难题,工质密度和定压比热的剧烈变化同样会对流量和温度的精确控制产生影响。因此,国内外诸多研究机构针对 SCO_2 换热器内部的流动传热机理、设计及加工制造方法和试验验证技术开展了大量研究。

4.1 SCO_2 流动换热机理

探究 SCO_2 在微小尺度通道内的流动换热机理可以深入理解强物性变化下的热量输运机制,有利于构建适用于超临界流体传热的经验关联式,指导 SCO_2 换热器设计。

受热物性参数剧烈变化的影响(见图 1),SCO_2 的流动与换热过程与常规工质存在显著区别,可能会导致强化换热、恶化换热或正常传热 3 种结果[114]。总体而言,SCO_2 流动和换热特性与热流密度、质量流速、系统压力及流动方向等边界条件密切相关[115]。高热流密度结合低质量流速易造成换热恶化,而相反条件则会产生换热强化现象;高压下 SCO_2 热物性变化幅度较小,系统压力远离临界压力时,换热恶化/强化效果会得到抑制;流动方向的影响则较为复杂,竖直向上流动相较于竖直向下流动更易产生换热恶化现象,而水平流动则会在流道截面方向产生非均匀传热,流道内产生的“肾形涡”通常会强化通道底部的换热能力[116-118]。

异常传热现象产生的根本原因在于强物性变化及其导致的特殊流动结构。北航团队的研究表明[119],流体流动换热过程中,工质的热物性变化迅速,浮升力对流动换热的影响逐渐增大,导致流场在不同空间位置处呈现出非均匀分布,如图 25 所示。受浮升力作用,圆周方向上近壁面处的高温低密度流体随主流向下游运动的过程中,逐渐向圆管顶部聚集,低温高密度流体则受重力的影响向底部壁面处聚集,导致局部换热出现差异。近壁面处湍动能与湍流换热也随之发生改变,增强了换热特性的复杂性。同时,上述过程的发生将伴随近壁面处湍流流动结构的改变,在不同空间位置处的湍流猝发行为出现差异。如图 26 所示,在浮升力影响显著的水平管流中,发现在底部壁面处低速上抛和高速下扫现象得到强化,尤其是低速上抛现象。然而,在顶部壁面处相同现象的发生却在一定程度上受到了抑制,进而导致局部流动换热能力显著不同。

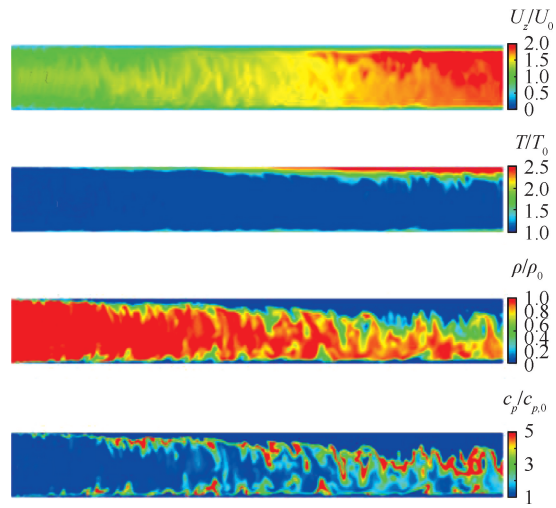


图 25 水平加热圆管内超临界流体瞬时参量分布(混合对流工况)[119]

Fig. 25 Instantaneous parameter distribution of supercritical fluid in a horizontal heated tube (mixed convection)[119]

此外,流体密度的剧烈变化会增强浮升力的作用,从而改变壁面切应力分布及边界层内湍流脉动的生成量,增强换热特性的复杂性[120]。加热条件下密度的迅速降低同样会导致流速增加,产生热加速效应。当热流密度较高时,流体沿流向的加速行为会降低边界层对数区的湍动能,从而导致湍流层流化[121]。此外,贴近壁面处的流体也会受到热加速影响,从而导致特殊的流动现象(主流速度低于近壁区),如图 27 所示[122]。一般地,浮升力和热加速在 SCO_2 流动换热过程中产生耦合影响,大尺寸通道中浮升力效果较强,而小尺寸通道中热加速效果则更为明显[123]。Kurganov 和 Kaptil' ny[124] 在竖直加热管流的实验研究中发现,受浮升力和热加速的耦合作用,SCO_2 的速度型会经历由“U”型变化至平坦型,再演化

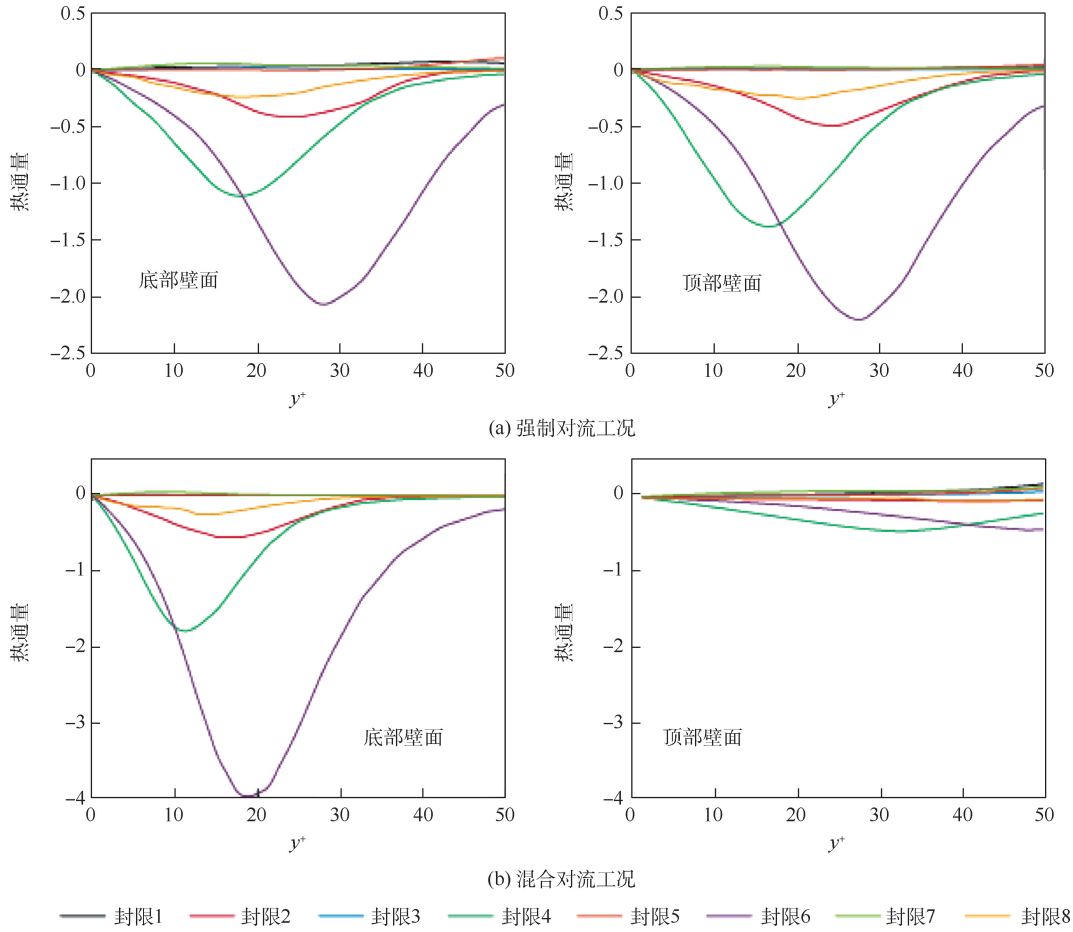


图 26　水平加热圆管内超临界流体湍流行为与湍流换热量关系[119]

Fig. 26　Relationship between turbulence behavior and turbulent heat transfer of supercritical fluid in a horizontal heated tube[119]

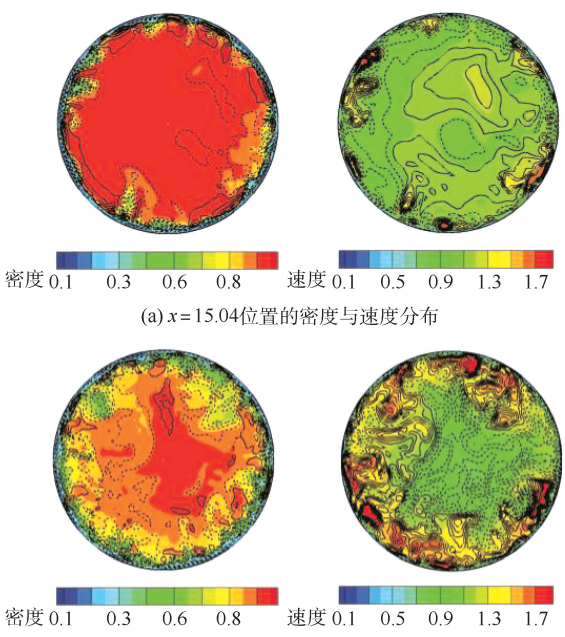


图 27　竖直加热圆管内 SCO_2 密度与速度分布[122]

Fig. 27　Density and velocity distributions of SCO_2 in a heated vertical tube[122]

为“M”型的复杂过程,而这 2 个变化过程又对应着切应力的减小和恢复,分别造成换热能力的削弱与增强。

4.2　紧凑式换热器设计及制造技术

4.2.1　热力设计方法

由于 SCO_2 在流动换热过程中存在强物性变化,使得传统基于常物性假设的设计方法不再适用[76]。为了捕捉物性变化对换热系数、流动阻力和平均温差等关键参数的影响,研究人员基于“离散换热微元”思想提出了紧凑式换热器的离散设计方法,如图 28 所示[125-127]。该方法根据换热量(设计)或流动长度(校核)将换热器划分为若干个子换热器,使得各子换热器内部的参数可视为常数。在此基础上,针对每个子换热器利用传统方法进行设计或校核,即可获得考虑关键参数变化影响的换热器结果。当需要考虑输入条件变化和时间参数影响时,基于该设计方法也可构建紧凑式换热器的动态分析模型[128]。

针对目前最常用于 SCO_2 闭式布莱顿循环系统的逆流式 PCHE,Guo[129] 和 Ke 等[130] 分别基于效能-传热单元数方法和对数温差法建立了相应的离散分析模型,并证明了方法的有效性。在此基础上,北航团队提出了考虑肋效率影响的基本换热单元并联热阻模型(见图 29)[131],该模型以翅片 $\theta=(\pi-2)/4$ 位置为临界截面,将传统基本

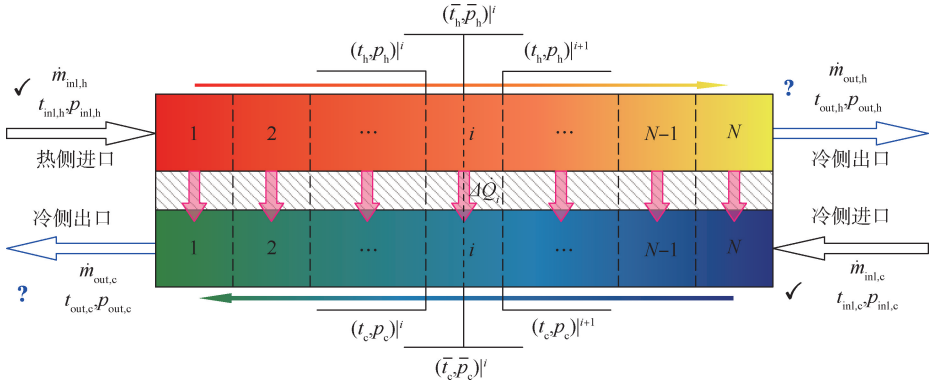


图 28 紧凑式换热器离散设计方法

Fig. 28 Segmented design method for compact heat exchanger

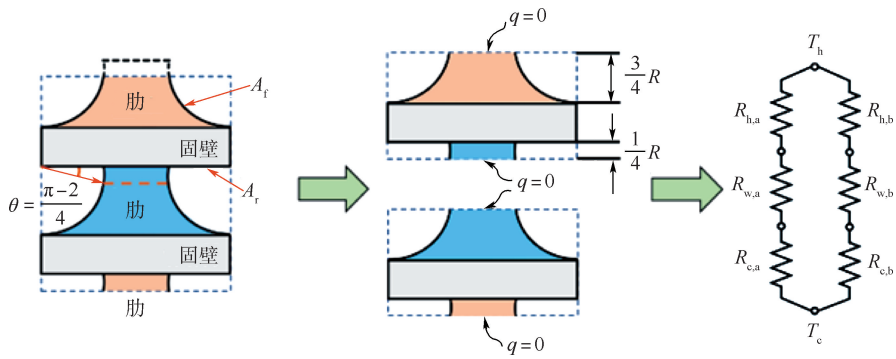


图 29 基本换热单元并联热阻模型[131]

Fig. 29 Parallel thermal resistance model for heat transfer unit[131]

换热单元进一步分割为 2 个换热子单元,则换热子单元的设计过程均可视为一个带第三类边界条件的一维导热问题,从而实现了翅片影响的精细化考虑。对于错流式 PCHE,Zhao 等[132]发展了一种基于热网络的二维离散设计方法,并通过实验验证了该方法对强物性变化的捕捉能力。此外,由于封头的存在,换热器内部可能会出现逆流、顺流和错流同时存在的情况。因此,Son 等[133]基于能量微分方程提出了考虑真实路径变化的二维离散设计方法,并指出采用纯逆流假设模型会对低效能换热器产生较大预测误差。对于有望用于热源换热器和回热器的微细管式换热器,Li[134]和 Jiang[135]等也提出了相应的离散设计模型。

4.2.2 换热器结构设计

综合发展现状来看,PCHE 是 SCO_2 闭式布莱顿循环系统中应用最多的换热器,因为其具有耐压高、单位体积换热面积大、轻质高效等优势。图 30展示了 PCHE 的典型通道类型,可以分为连续型通道(直线型、Z 字型和正弦型)和间断型通道(翼形翅片、S 形翅片和矩形翅片)2 类[136]。目前,直线型和 Z 字型是最为成熟的通道类型,其加工简单、结构完整性好,且用于换热器设计的经验关系式较为可靠[137]。相较于直线型通道,由于换热面积增加和湍流增强,Z 字型通道具有更好的换热效果,但相应的流动损失也会增加。正弦型通道及由其发展而来的 S 形翅片通道可有效降低 Z 字型通道中流体分离所造成的压力损失[138]。Tsuzuki 等[139]发现,在相同换热能力下,S 形翅片通道的压力损失仅为 Z 字型通道的 1/5。类似地,Kim 等[140]提出了 NACA 翼形翅片通道,指出在相同换热能力下该型通道的压降仅为 Z 字型通道的 1/20。此外,矩形翅片通道能有效增强换热能力,但相同雷诺数下其摩擦因子却高于 S 形翅片通道和翼形翅片通道[141]。尽管基于新型通道的 PCHE 展现出了优异的热工水力性能,但其较高的加工制造难度和较低的承压能力在一定程度上限制了自身的发展。

除了换热器芯体通道结构设计外,换热器的封头设计同样重要,较差的封头设计会造成流体流量分配不均,从而降低换热器性能[142-143]。Chu 等[144]针对 PCHE 设计了 4 种不同型线的封头形式(见图 31),发现基于双曲线形的封头可以有效降低气流不均匀性,提高换热器的综合性能。Siddiqui 和 Zubair[145]发现封头歧管面积和换热器通道长度越大,流量分配越均匀。Koo 等[146]则通过调整封头角度和半径对封头进行形状优化,以获得质量流量和压力损失的均匀分布。此外,Ma 等[147]提出了流量分布的物理模型来预测流量不均匀性,

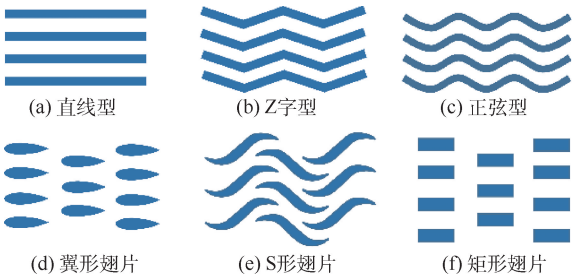


图 30　PCHE 典型通道类型

Fig. 30　Typical flow passage type of PCHE

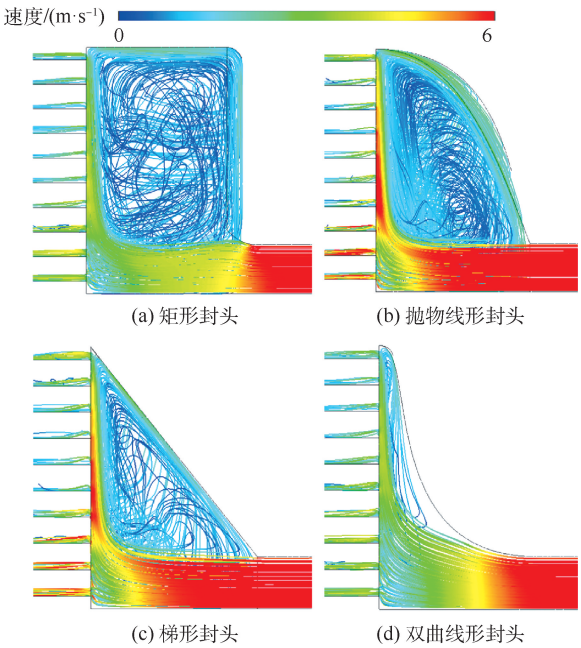


图 31　PCHE 换热器不同入口封头的影响[144]

Fig. 31　Effects of different inlet headers of PCHE[144]

并分析了不同因素对流量不均匀性的影响。

4.2.3　换热器制造技术

为满足高温超高压条件长期服役要求,PCHE 一般采用光化学蚀刻及扩散焊制造技术,以使得芯体接合面达到母材的同等强度[148]。图 32 展示了 PCHE 的基本制造过程。具体为:①通过光刻和蚀刻技术在金属薄板上加工出特定的通道图案,常规通道尺寸约为 0.5 ~ 2.0 mm[149]。②将加工过的薄板按照规律“堆叠”并放入真空炉中,通过高温高压实现扩散焊接。受到焊接设备的限制,PCHE 芯体的外形尺寸一般小于 600 mm × 600 mm × 1 200 mm[150]。③将芯体焊接上封头和歧管等结构形成最终换热器。由于特殊的结构设计和扩散焊技术,PCHE 耐压和耐温最高可达 90 MPa和 900℃。此外,与传统管壳式换热器相比,PCHE 基本换热单元的特征尺度相对较小,紧凑度可高达 2 500 m^2/m^3,故体积最高可减小 85%[151]。对于通道直径小于 2 mm 的微管换热器,其一般采用钎焊技术将换热管焊接于管板上,并被证实可承受 15 MPa 的超高压力[109]。

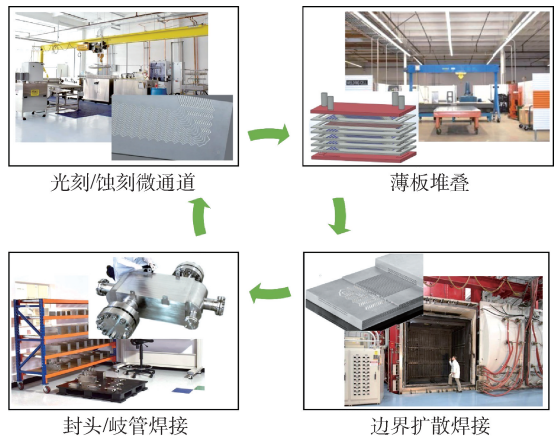


图 32　PCHE 加工过程[131]

Fig. 32　Processing route of PCHE[131]

近年来,增材制造技术(3D 打印)为 SCO_2 换热器加工开辟了全新的技术途径[152]。Rasouli 等[111]利用增材制造技术完成了加热器制造,试验证明该加热器可承受约 20 MPa 高压和约 550℃高温。同时,增材制造技术摆脱了传统加工方式的约束,可用于加工十分复杂、新颖的通道形式。Ding 等[153]利用该技术完成了包含自适应通道的换热器加工,试验表明,该换热器在效能增加 2% 的同时压力损失可降低 69%,具有非常优异的工作性能。

4.3　紧凑式换热器试验研究

目前,针对 SCO_2 PCHE,包括美国 SNL[101,154] 和阿贡国家实验室[155],以及日本 TIT[156]、韩国 KAIST[157]、中国科学院工程热物理研究所[158] 等国内外研究机构均开展了丰富的试验研究。Liu 等[137]对特征长度为 1 140 μm 的直线型通道 PCHE 进行了压降特性测试,实验雷诺数范围为 64 ~ 14 160,在较为宽广的区间内验证了常规尺度流阻经验模型的精度,最大偏差不超过 4.6%。Chu 等[159]以水和 SCO_2 为工质研究了直线型通道 PCHE 的流阻和换热性能,发现相同流量下 SCO_2 的换热功率约为水的1.2 ~ 1.5 倍,并且在跨临界状态换热性能会增强。Kim 等[127,160-161]分别以水、氦气及氦气-二氧化碳混合物为工质,对特征长度 922 μm、弯角 15°的 Z 字型通道 PCHE 进行了详尽的试验研究,提出了全新的流阻和换热经验关联式以指导换热器设计。此外,Nikitin 等[156]开展了 3 kW 级 PCHE 的性能测试,结果表明,该换热器的总换热系数可以达到 300 ~ 650 W/(m^2K)。Zhou 等[158]设计制造了 100 kW 量级的 PCHE,并通过试验获得了 98.4% 的换热器效能,如图 33 所示。鉴于连续型通道 PCHE 具有良好的承压和换热能力,北航团队也试验探究了微

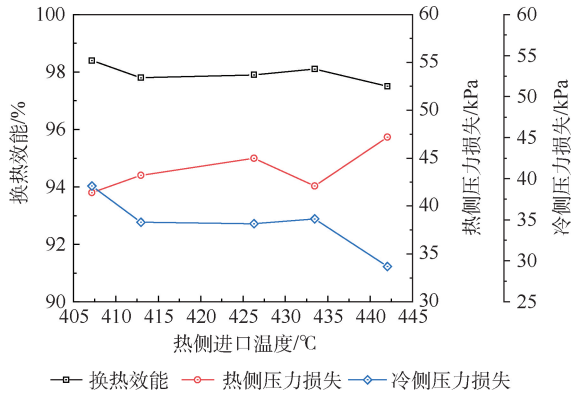


图 33 100 kW 级 PCHE 试验结果[158]

Fig. 33 Experimental results of 100 kW class PCHE[158]

通道 PCHE(直线型通道和 Z 字型通道)应用于高超声速强预冷发动机循环的可行性[125,131],试验结果表明,所研制的微通道 PCHE 具有 105 kW/kg 的超高功重比性能,且总传热系数高达 8.1 kW/(m^2K),在航空航天领域极具应用前景。

针对间断型通道,Pidaparti 等[143]通过试验研究了翼形翅片及矩形翅片通道的换热和流阻特性,结果显示,与 52° Z 字型通道相比,这 2 种间断型通道均能显著降低流阻,且矩形翅片通道具有最好的换热性能,可减小 25% 的制造成本。Ngo 等[162]对比了 S 形翅片通道和 52° Z 字型通道的流动与传热性能,试验结果表明,Z 字型通道的努塞尔数虽比 S 形翅片通道高 24% ~34%,但其流阻却是后者的 4 ~5 倍。对于其他类型的紧凑式换热器,目前仅有部分研究对微管式换热器[109]和增材制造式换热器[111]的热工水力性能进行探究。在此基础上,诸多研究人员根据试验结果提出了适用于 SCO_2 换热器设计的流动与换热低维预测模型,以考虑强物性变化对换热器热工水力特性的影响。表 6 给出了 PCHE 典型通道类型的部分经验关联式。可以发现,目前针对 SCO_2 在复杂通道内的试验研究多集中于较低雷诺数范围,仍需开展更多高雷诺区间的试验研究以扩宽经验关联式的适用范围。

表 6 SCO_2 PCHE 流动与换热经验关联式

Table 6 Flow and heat transfer correlations for SCO_2 PCHE

通道类型	研究人员	经验关联式	适用范围
直线型	Chu 等[159]	$Nu_b/Nu_{Fc} = 0.58 - 53(Gr/Re^{2.7})^{0.36}\ (T_w \approx T_{pc})$ $Nu_b/Nu_{Fc} = 0.36 - 22(Gr/Re^{2.7})^{0.42}\ (T_w > T_{pc})$	$3\,000 < Re < 60\,000$ $3\,000 < Re < 70\,000$
	Liu 等[163]	$Nu_b = 0.122\,9\,Re_b^{0.602\,1}\,Pr_b^{0.3}(c_{p,w}/c_{p,b})^{0.131\,0}$ $f_b = 0.057\,76\,Re_b^{-0.219\,2}$	$3\,600 < Re < 36\,500$
Z 字型	Ngo 等[162]	$Nu = (0.169\,6 \pm 0.014\,4)Re^{0.629\pm0.009}Pr^{0.317\pm0.014}$ $f = (0.192\,4 \pm 0.029\,9)Re^{-0.091\pm0.016}$	$3\,500 < Re < 22\,000$
	Nikitin 等[156]	$\begin{cases} h_{hot} = 2.52Re^{0.681} \\ h_{cold} = 5.49Re^{0.625} \end{cases}$ $\begin{cases} f_{hot} = (-1.402\times10^{-6} \pm 0.087\times10^{-6})Re + (0.044\,95 \pm 0.000\,38) \\ f_{cold} = (-1.545\times10^{-6} \pm 0.099\times10^{-6})Re + (0.093\,18 \pm 0.000\,9) \end{cases}$	$\begin{cases} 2\,800 < Re < 5\,800 \\ 6\,200 < Re < 12\,100 \end{cases}$ $\begin{cases} 2\,800 < Re < 5\,800 \\ 6\,200 < Re < 12\,100 \end{cases}$
	Ngo 等[162]	$Nu = (0.174\,0 \pm 0.011\,8)Re^{0.593\pm0.007}Pr^{0.430\pm0.014}$ $f = (0.454\,5 \pm 0.040\,5)Re^{-0.340\pm0.009}$	$3\,500 < Re < 23\,000$
S 形翅片	Zhao 等[164]	$\begin{cases} Nu_{hot} = 0.177\,2Re^{0.605\,1}Pr^{0.212\,7} \\ Nu_{cold} = 0.100\,1Re^{0.663\,7}Pr^{0.353\,6} \end{cases}$ $\begin{cases} f_{hot} = 0.437\,4Re^{-0.237\,2} \\ f_{cold} = 0.207\,2Re^{-0.181\,4} \end{cases}$	$\begin{cases} 2\,700 < Re < 7\,000 \\ 1\,000 < Re < 2\,700 \end{cases}$ $\begin{cases} 2\,700 < Re < 7\,000 \\ 1\,000 < Re < 2\,700 \end{cases}$
翼形翅片	Pidaparti 等[143]	$Nu = 0.060\,1\,Re^{0.732\,6}Pr^{0.345\,3}(\rho_b/\rho_w)^{0.432\,9}(C_{pb}/C_p)^{-0.355\,6}$ $f = 0.025\,6$	$4\,000 < Re < 38\,000$ $3\,800 < Re < 38\,000$

5 控制及储热系统

5.1 循环控制策略

为保证 SCO_2 闭式布莱顿循环安全稳定运行,有效可靠的控制策略十分重要。一般地,SCO_2 闭式布莱顿循环的控制参数包括主压缩压气机进口温度和压力、涡轮进口温度、压气机喘振等,常用的控制策略包括各类阀门、叶轮机转速、冷/热源换热量等[165]。

图 34 给出了典型的再压缩循环控制策略。主压缩压气机进口参数(尤其是温度)对循环系统性能影响显著,保证其进口参数稳定维持在临界点附近(超临界区)是循环控制策略的重点。通常,主压缩压气机进口温度通过调整冷却器冷

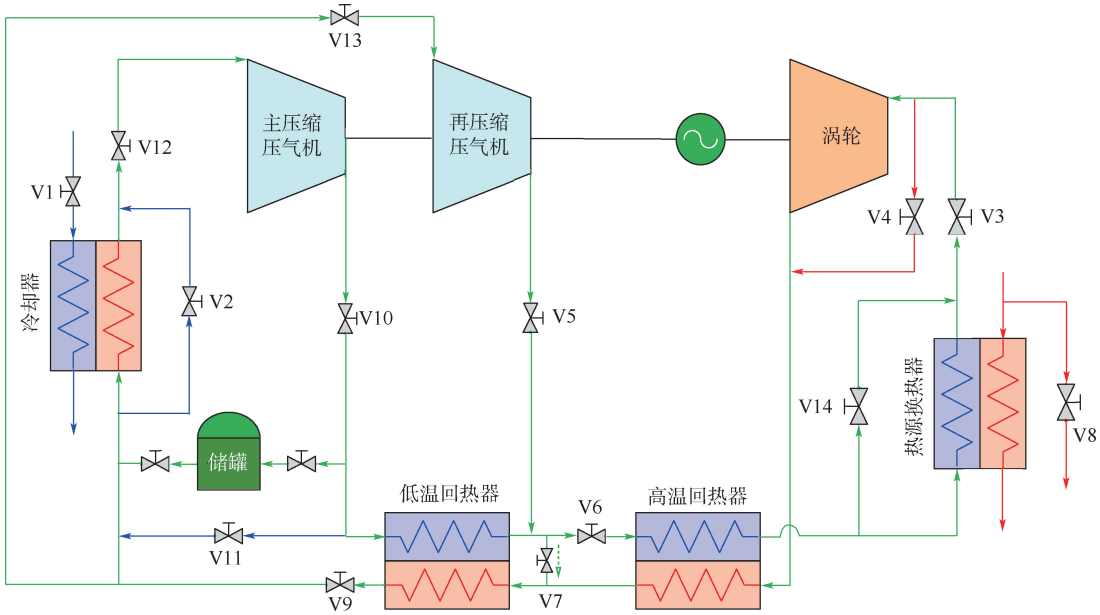


图 34　典型的 SCO_2 再压缩循环控制策略

Fig. 34　Typical control strategy for SCO_2 recompression cycle

侧流量进行控制(调整阀门 V1 开度或者冷却泵转速)[166],必要时也可调整旁通阀 V2 进行辅助控制[167];而进口压力则通常采用库存管理策略通过向循环系统充/放工质进行控制,工质储罐一般配置于主压缩压气机出口和冷却器进口间[168-169]。此外,引入低温回热器节流阀(V9)并配合冷却器流量调节也可实现主压缩压气机进口压力和温度的有效控制[170]。涡轮进口温度,直接影响循环系统功率和运行安全性,一般可采用调整热源工质流量、阀门开度、压气机转速等策略实现有效控制。Luu 等[171]采用调整热源工质流量(调整阀门 V8)的策略,通过改变循环吸热功率以控制涡轮进口温度;而通过旁通流路改变循环吸热量也可实现涡轮进口温度控制,如涡轮旁通(阀门 V4)和热源换热器旁通(阀门 V14)。此外,Matteo 等[172]指出涡轮进口温度对循环流量的敏感性更高,采用库存管理也可控制涡轮进口温度,且更为合适。由于循环流量受压气机转速的影响显著,调整压气机转速是一种控制涡轮进口温度的策略[173]。为规避压气机喘振,循环运行过程中需采用特殊的控制策略。通常,采用调整叶轮机转速[174]或阀门开度[166]的策略实现压气机喘振保护。由于喘振发生在小流量工况下,通过提高转速以增大流量可规避压气机进入喘振状态。此外,引入压气机循环阀(V11)和节流阀(V10 和 V12)也可有效规避喘振。Clementoni 等[41]采用循环阀在较宽的工况范围内将压气机维持在喘振边界内运行;Gao 等[175]则采用节流阀进行压气机喘振保护,并指出节流阀的引入可有效提高控制系统的鲁棒性。

除上述控制参数外,换热器进口温度、分流比、压气机转速等也时常被作为控制参数。Moiseytsev 和 Sienicki[176]引入旁通阀(V7)来控制高温换热器进口温度,并间接地实现对热源的有效控制。循环分流比可通过调整阀门开度和压气机转速等方法进行控制,Carstens[84]指出可通过调整主压缩压气机和再压缩压气机转速,或分别在其出口引入阀门 V10 和 V5,来实现对分流比的有效控制。由于压气机通常由电动机或涡轮驱动,也可通过调整电动机或涡轮的转速以控制分流比[40-41]。

5.2　储热系统

当前,储热技术被广泛应用于太阳能光热发电中,以在无日照条件下维持系统电功率输出。在集热-储热端,多采用直接集成高/低温熔盐储罐的直接式储热布局,熔盐被同时用作换热和储热介质;在发电端,可直接采用 SCO_2 闭式布莱顿循环系统替代现有蒸汽朗肯循环系统,如图 35 所示。当集热温度超过 565℃时,现有的硝酸熔盐难以适用作集热-储热介质,包括液态金属(如 Na、Na-K、Pb-Bi)、氯盐、碳酸盐等新型集/储热介质有望得以利用[177-178]。此外,利用固体颗粒进行集热-储热的技术也有望得到应用,但接收器和熔盐储罐需特殊设计[179]。Kelly[180]给出了一种采用硅岩或铝珠等惰性介质填充层的储热系统设计,采用内隔热设计可有效避免高温端加装压力安全壳;在此基础上,Bindra 等[181]通过动态改变流动进出口的设计有效提高了储热系统的㶲效率。

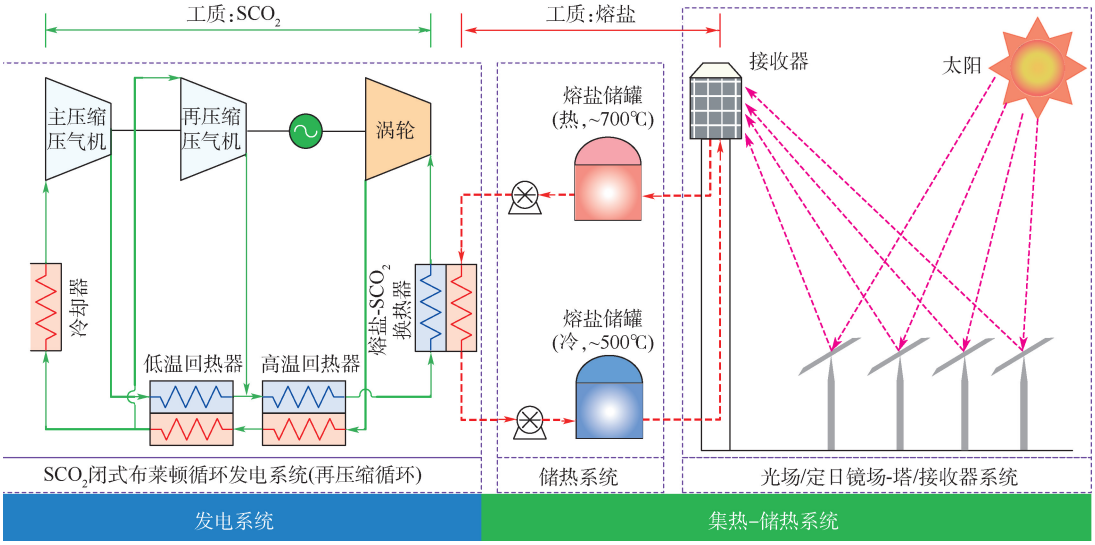


图35 聚焦式太阳能发电系统(集成SCO_2再压缩循环)

Fig. 35 Concentrating solar power system (integrated with SCO_2 recompression cycle)

此外,相变储热系统是另一种技术途径,其通常采用化学盐或金属相变材料。相较而言,高熔化热的铝或铝合金是此类储热系统较为理想的可选材料,而化学盐相变材料则存在导热率限制问题。为改善化学盐相变材料性能,可将其注入至石墨等多孔导热介质中[182],亦或将其封装至热稳定壳体并置入填充层中[183]。相变储热系统有望更好地与构型简单的SCO_2闭式布莱顿循环发电系统集成,然而目前尚未在太阳能光热发电领域得到商业化应用。

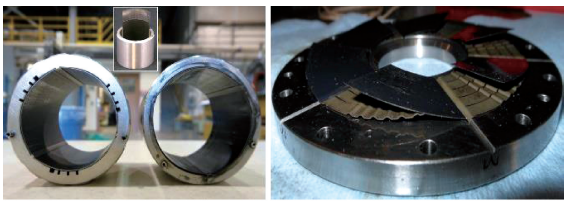
6 工程应用面临的问题与挑战

现阶段,SCO_2闭式布莱顿循环技术虽已取得长足发展,但离大范围工程应用尚存在一定距离。未来,需围绕以下方面开展更为全面深入的研究。

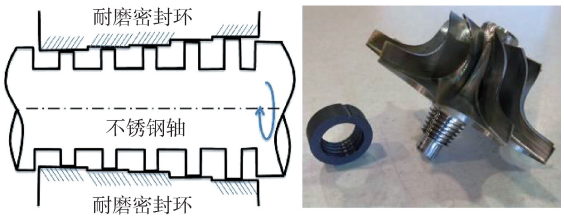
1)计及不确定性效应的全寿命周期变工况热力-经济综合性能的总体技术。SCO_2闭式布莱顿循环系统的运行状态受外部工况条件影响显著,工作环境变化幅度大,不确定性效应明显,其中最为典型的如太阳能光热发电存在多变的光照时间和强度等,有必要发展计及多变工况综合性能及不确定性影响的SCO_2总体技术。此外,在工程应用中考虑全寿命周期的效费比(性能/成本)最佳的气动热力和结构布局、最佳运行控制规律等仍需进一步探讨。有必要构建适用于总体热力分析的、考虑长时间运行的性能衰减的模型,以真正建立考虑全工况范围内综合性能的评价模型和方法。由于闭式循环多子系统的紧密耦合,使得系统的控制和调节更为复杂和困难,尚需突破多变工况下系统动态响应机制并掌握高鲁棒动态调控策略,以及精确的气动热力边界条件控制技术。完整的SCO_2闭式布莱顿循环系统运行时中涉及储能等相关子系统,发展更稳定的高温储热材料和储热系统或其他形式的储能系统,有助于提高整个系统对多变工况应用场景下的适应性。

2)冷凝相变对压气机和系统的影响机制及其在设计中的考虑。当SCO_2闭式布莱顿循环系统中的压气机工作在临界点附近时,压气机进口存在的局部加速可能导致冷凝相变,从而导致局部的气动热力状态和流动发生较为明显的变化,进而影响压气机的性能、稳定工作范围,甚至导致压气机的激振等。液滴与叶片的相互作用还可能会导致叶片的腐蚀和磨损等,进一步引起压气机乃至系统性能的衰减,并降低其稳定工作范围。目前,缺乏有效且可准确模拟冷凝的非稳态相变过程的数值手段;对冷凝的动力学特征及影响机制尚不明晰,包括液相形成过程、尺寸/形态/分布范围、演化规律等,液滴撞击等因素导致的叶片磨损和激振机理、过程及力学特征等。此外,冷凝相变的有效调控方法及其在气动设计层面(如低维速度三角形设计参数选取、叶片造型等)的应用途径等还需进一步研究。

3)高速叶轮机高温高压动封严技术。高速/高压/高温工作环境导致超临界叶轮机泄漏量大、鼓风损失、动静碰磨、气流激振诱发转子振动、高温腐蚀等问题,严重限制运行转速和性能的提高,现有常规迷宫、蜂窝等密封方法难以满足封严要求。采用气浮或磁悬浮轴承支撑转子并做好与空气系统等的匹配是一种可行思路,如SNL采用气浮轴承支撑配合耐磨迷宫封严结构[184],如图36所示。但目前高速气浮轴承面临刚度不足导致的



(a) 气浮轴承(左: 径向轴承 右: 推力轴承)



(b) SCO_2涡轮轴迷宫封严结构

图 36　SNL 叶轮机气浮轴承及迷宫封严结构[184]

Fig. 36　Turbomachinery gas foil bearing and labyrinth seal structure at SNL[184]

转子动力学问题、高温环境下耐磨涂层失效等问题,磁悬浮轴承则存在体积大、质量高、结构复杂、控制精度要求苛刻等不足,相关技术仍需进一步提升。此外,空气系统的设计及转子系统大轴向力的平衡仍是限制气浮或磁悬浮轴承得到工程实用的难题。特别是在闭式循环系统中,无论是气浮轴承还是磁悬浮轴承,能否有效工作与二次流系统紧密相关,因此实现轴承与二次流系统在全工况范围的匹配以提供轴承安全可靠的工作环境,也是需要进一步研究的技术。

4) 轻质耐腐蚀高温材料及加工技术。SCO_2闭式布莱顿循环系统包含涡轮、换热器、高温调节阀和高温高压管路等多个高温/高压环境工作的热端零部件,这些热端零部件是限制系统寿命和安全性的关键,其对材料有着极为苛刻的要求:耐高温和高压、耐腐蚀、低成本并具备良好加工性能等,使得其成本是整个系统生产成本中最为昂贵的部分。为进一步提高系统寿命、降低生产成本,以满足工程使用要求,需进一步开发适用于SCO_2闭式布莱顿循环系统的耐高温、耐腐蚀、轻质、低成本和加工性能好的新材料,并突破高温换热器、高温调节阀等高温零部件的加工制备技术以降低热端零部件的生产成本。

5) 系统全面的试验验证和数据库的建立。目前的研究在试验验证方面远不足以支撑形成成熟的产品,现有报道的试验大多数是零散、不系统的部件级试验,整机的试验很少,且已有的试验中大多数仍然偏向于性能,并未覆盖到所有设计流程,如耐久性考核试验、性能衰减试验、环境试验、部件失效试验等亦未见报道,即没有完整的覆盖全设计流程的试验验证相关工作,导致实际工程应用中缺乏诸多数据和信息支撑,同时设计技术、准则和全工况的控制策略、系统的安全性和稳定性、环境的适用性等未得到验证或充分验证。因此,建立适用于SCO_2闭式布莱顿循环系统的试验体系,制定相应的试验流程和规范,开展系统全流程全要素试验验证工作,并建立相应的数据库,对于形成成熟的SCO_2闭式布莱顿循环发电设备等产品具有十分重要的意义。

6) 高速电机技术、高温高压调节阀技术、辅助系统与主机系统全工况的匹配和调控技术、考虑不确定性的部件设计方法等也需在未来开展更全面深入的研究。

7　结 束 语

SCO_2闭式布莱顿循环是一项具有巨大性能优势和广阔应用前景的热动力利用革新技术。总结和分析国内外相关研究表明:

1) 热力循环是SCO_2闭式布莱顿循环系统的关键,总体热力设计和分析时有必要涵盖部件低维特性分析,以充分考虑所提出的部件性能指标的可实现性。此外,循环设计目标并非单纯追求设计点的高热效率,而应综合考虑全寿命周期性能、紧凑性、经济性等指标。

2) SCO_2工质的强真实气体效应使得叶轮机设计不能简单借鉴成熟的常规工质叶轮机的设计方法,需针对性地发展有效的设计方法。此外,通过理论分析和机器深度学习相结合得到的不同工质叶轮机相似方法为利用常规工质叶轮机试验台开展SCO_2等特殊工质叶轮机气动性能的试验验证奠定了理论基础。

3) 剧烈的热物性变化导致SCO_2特殊的微尺度流动换热机理,采用离散模型进行紧凑换热器设计非常必要。此外,通过机理性试验构建适用于SCO_2工质的流动换热模型可为紧凑高效换热器设计提供重要支撑。

4) 鲁棒的控制策略是SCO_2闭式布莱顿循环系统稳定运行的关键,旁通流路、库存管理等策略可实现循环的可靠高效调控。此外,储能对太阳能光热发电或多能综合管理/利用具有重要意义,集成新型介质储热技术的SCO_2闭式布莱顿循环系统是高温光热发电的关键技术储备。

5) 为促进SCO_2闭式布莱顿循环技术大范围工程应用,未来仍需在计及不确定性效应的全寿命周期热力-经济综合性能的总体技术、冷凝相变的影响机制和调控方法、高速叶轮机高温高压动

封严技术、轻质耐腐蚀高温材料及加工技术、更全面的系统试验验证和数据库建立、高速电机技术、考虑不确定性的部件设计方法等方面开展更全面深入的研究。

参考文献（References）

[1] TAKIZUKA T, TAKADA S, YAN X, et al. R&D on the power conversion system for gas turbine high temperature reactors[J]. Nuclear Engineering and Design, 2004, 233(1-3): 329-346.

[2] HERRANZ L E, LINARES J I, MORATILLA B Y. Power cycle assessment of nuclear high temperature gas-cooled reactors[J]. Applied Thermal Engineering, 2009, 29(8): 1759-1765.

[3] EL-GENK M S, TOURNIER J M. Noble-gas binary mixtures for closed-Brayton-cycle space reactor power systems[J]. Journal of Propulsion and Power, 2007, 23(4): 863-873.

[4] EL-GENK M S, TOURNIER J M. Selection of noble gas binary mixtures for Brayton space nuclear power systems[C]// Proceedings of the International Energy Conversion Engineering Conference and Exhibit, 2006: 4168-4176.

[5] KLAUS B, PETER F, RICHARD D. Fundamentals and applications of supercritical carbon dioxide (sCO_2) based power cycles[M]. Cambridge: Woodhead Publishing, 2017.

[6] CRESPI F, GAVAGNIN G, SANCHEZ D, et al. Supercritical carbon dioxide cycles for power generation: A review[J]. Applied Energy, 2017, 195: 152-183.

[7] FEHER E G. The supercritical thermodynamic power cycle[J]. Energy Conversion, 1968, 8(2): 85-90.

[8] HOFFMANN J R, FEHER E G. 150 kwe supercritical closed cycle system[J]. Journal of Engineering for Power, 1971, 93(1): 70-80.

[9] ANGELINO G. Perspectives for the liquid phase compression gas turbine[J]. Journal of Engineering for Power, 1967, 89(2): 229-236.

[10] ANGELINO G. Carbon dioxide condensation cycles for power production[J]. Journal of Engineering for Power, 1968, 90(3): 287-295.

[11] DIEVOT J P. The sodium-CO_2 fast breeder reactor concept[C]// International Conference on Sodium Technology and Large Fast Reactor Design, 1968: 7-9.

[12] GOKHSTEIN D P. Use of carbon dioxide as a heat carrier and working substance in atomic power station[J]. Soviet Atomic Energy, 1969, 26(4): 378-380.

[13] GOKHSTEIN D P, VERKHIVKER G P. Future design of thermal power stations operating on carbon dioxide[J]. Thermal Engineering, 1971, 18(4): 36-38.

[14] DOSTAL V, DRISCOLL M J, HEJZLAR P. A supercritical carbon dioxide cycle for next generation nuclear reactors: MIT-ANP-TR-100[R]. Cambridge: MIT, 2004.

[15] DOSTAL V, HEJZLAR P, DRISCOLL M J. High-performance supercritical carbon dioxide cycle for next-generation nuclear reactors[J]. Nuclear Technology, 2006, 154(3): 264-282.

[16] AHN Y, BAE S J, KIM M, et al. Review of supercritical CO_2 power cycle technology and current status of research and development[J]. Nuclear Engineering and Technology, 2015, 47(6): 647-661.

[17] FLEMING D D, HOLSCHUH T V, CONBOY T M, et al. Scaling considerations for a multi-megawatt class supercritical CO_2 Brayton cycle and commercialization[C]// Proceedings of the ASME Turbo Expo 2012: Turbine Technical Conference and Exposition. New York: ASME, 2012, 5: 953-960.

[18] ISHIYAMA S, MUTO Y, KATO Y, et al. Study of steam, helium and supercritical CO_2 turbine power generations in prototype fusion power reactor[J]. Progress in Nuclear Energy, 2008, 50(2-6): 325-332.

[19] LIAO G L, LIU L J, E J Q, et al. Effects of technical progress on performance and application of supercritical carbon dioxide power cycle: A review[J]. Energy Conversion and Management, 2019, 199: 11986.

[20] CLEMENTONI E M, COX T L. Practical aspects of supercritical carbon dioxide Brayton system testing[C]// Proceedings of the 4th International Symposium on Supercritical CO_2 Power Cycles, 2014: 1-7.

[21] WANG K, HE Y L, ZHU H H. Integration between supercritical CO_2 Brayton cycles and molten salt solar power towers: A review and a comprehensive comparison of different cycle layouts[J]. Applied Energy, 2017, 195: 819-836.

[22] HO C K, CARLSON M, GARG P, et al. Technoeconomic analysis of alternative solarized s-CO_2 Brayton cycle configurations[J]. Journal of Solar Energy Engineering, 2016, 138(5): 051008.

[23] VISWANATHAN R, COLEMAN K, RAO U. Materials for ultra-supercritical coal-fired power plant boilers[J]. International Journal of Pressure Vessels and Piping, 2006, 83(11-12): 778-783.

[24] PARK S H, KIM J Y, YOON M K, et al. Thermodynamic and economic investigation of coal-fired power plant combined with various supercritical CO_2 Brayton power cycle[J]. Applied Thermal Engineering, 2018, 130: 611-623.

[25] MANJUNATH K, SHARMA O P, TYAGI S K, et al. Thermodynamic analysis of a supercritical/transcritical CO_2 based waste heat recovery cycle for shipboard power and cooling applications[J]. Energy Conversion and Management, 2018, 155: 262-275.

[26] SHARMA O P, KAUSHIK S C, MANJUNATH K. Thermodynamic analysis and optimization of a supercritical CO_2 regenerative recompression Brayton cycle coupled with a marine gas turbine for shipboard waste heat recovery[J]. Thermal Science and Engineering Progress, 2017, 3: 62-74.

[27] CHENG K L, QIN J, SUN H C, et al. Power optimization and comparison between simple recuperated and recompression supercritical carbon dioxide closed-Brayton-cycle with finite cold source on hypersonic vehicles[J]. Energy, 2019, 181: 1189-1201.

[28] JACOB F, ROLT A M, SEBASTIAMPILLAI J M, et al. Performance of a supercritical CO_2 bottoming cycle for aero applications[J]. Applied Sciences, 2017, 7(2): 255.

[29] MIAO H Y, WANG Z W, NIU Y B. Performance analysis of cooling system based on improved supercritical CO_2 Brayton cycles for scramjet[J]. Applied Thermal Engineering, 2020,

167:114774.

[30] MEHRPOOYA M, BAHRAMIAN P, POURFAYAZ F, et al. Introducing and analysis of a hybrid molten carbonate fuel cell-supercritical carbon dioxide Brayton cycle system[J]. Sustainable Energy Technologies and Assessments, 2016, 18: 100-106.

[31] MAHMOUDI S M S, GHAVIMI A R. Thermoeconomic analysis and multi objective optimization of a molten carbonate fuel cell-supercritical carbon dioxide-organic Rankine cycle integrated power system using liquefied natural gas as heat sink[J]. Applied Thermal Engineering, 2016, 107: 1219-1232.

[32] WRIGHT S A, RADEL R F, VERNON M E, et al. Operation and analysis of a supercritical CO_2 Brayton cycle: SAND2010-0171[R]. Albuquerque: Sandia National Laboroatories, 2010.

[33] WRIGHT S A, CONBOY T M, PARMA E J. Summary of the Sandia supercritical CO_2 development program[C]// Proceedings of the 3rd International Symposium on Supercritical CO_2 Power Cycles, 2011: 1-16.

[34] PASCH J, CONBOY T, FLEMING D, et al. Supercritical CO_2 recompression Brayton cycle: Completed assembly description: SAND2012-9546[R]. Albuquerque: Sandia National Laboroatories, 2012.

[35] CHIRANJEEV K, DOUGLAS S H, EDIP S. Development of high efficiency hot gas turbo-expander for optimized CSP supercritical CO_2 power block operation[C]// Proceedings of the 4th International Symposium on Supercritical CO_2 Power Cycles, 2014: 1-11.

[36] BIDKAR R A, MANN A, SINGH R, et al. Conceptual designs of 50 MW and 450 MW supercritical CO_2 turbomachinery trains for power generation from coal. Part 1: Cycle and turbine[C]// Proceedings of the 5th International Symposium on Supercritical CO_2 Power Cycles, 2016: 1-18.

[37] BIDKAR R A, MANN A, SINGH R, et al. Conceptual designs of 50 MW and 450 MW supercritical CO_2 turbomachinery trains for power generation from coal. Part 2: Compressors[C]// Proceedings of the 5th International Symposium on Supercritical CO_2 Power Cycles, 2016: 1-18.

[38] CLEMENTONI E M, COX T L, KING M A. Off-nominal component performance in a supercritical carbon dioxide Brayton cycle[J]. Journal of Engineering for Gas Turbine and Power, 2016, 138(1): 011703-2.

[39] CLEMENTONI E M, COX T L, SPRAGUE C P. Startup and operation of a supercritical carbon fioxide Brayton cycle[J]. Journal of Engineering for Gas Turbines and Power, 2014, 136(7): V008T34A006.

[40] CLEMENTONI E M, COX T L, KING M A, et al. Initial transient power operation of a supercritical carbon dioxide Brayton cycle with thermal-hydraulic control[C]// Proceedings of the 5th International Symposium on Supercritical CO_2 Power Cycle, 2016: 1-10.

[41] CLEMENTONI E M, COX T L, KING M A, et al. Transient power operation of a supercritical carbon dioxide Brayton cycle[C]// Proceedings of the ASME Turbo Expo 2017: Turbomachinery Technical Conference & Exposition. New York: ASME, 2017: V009T38A001.

[42] HELD T J. Initial test results of a megawatt-class supercritical CO_2 heat engine[C]// Proceedings of the 4th International Symposium on Supercritical CO_2 Power Cycles, 2014: 1-12.

[43] CHA J E, LEE T H, EOH J H, et al. Development of a supercritical CO_2 Brayton energy conversion system coupled with a sodium cooled fast reactor[J]. Nuclear Engineering and Technology, 2009, 41(8): 1025-1044.

[44] AHN Y, LEE J, KIM S G, et al. The design study of supercritical carbon dioxide integral experiment loop[C]// Proceedings of the ASME Turbo Expo 2013: Turbine Technical Conference and Exposition. New York: ASME, 2013: V008T34A003.

[45] KIM S G, AHN Y, LEE J, et al. Numerical investigation of a centrifugal compressor for supercritical CO_2 as a working fluid[C]// Proceedings of the ASME Turbo Expo 2014: Turbine Technical Conference and Exposition. New York: ASME, 2014: V03BT36A005.

[46] LEE J, KIM S G, CHA J E, et al. Uncertainty on performance measurement of S-CO_2 compressor operating near the critical point[C]// Proceedings of the 4th International Symposium on Supercritical CO_2 Power Cycles, 2014: 1-9.

[47] CHO J, SHIN H, CHO J, et al. Development of the supercritical carbon dioxide power cycle experimental loop with a turbo-generator[C]// Proceedings of the ASME Turbo Expo 2017: Turbomachinery Technical Conference & Exposition. New York: ASME, 2017: V009T38A018.

[48] CHO J, SHIN H, RA H S, et al. Research on the development of a small-scale supercritical carbon dioxide power cycle experimental test loop[C]// Proceedings of the 5th International Symposium on Supercritical CO_2 Power Cycles, 2016: 1-12.

[49] UTAMURA M, HASUIKE H, YAMAMOTO T. Demonstration test plan of closed cycle gas turbine with supercritical CO_2 as working fluid[J]. Strojarstvo, 2010, 52(4): 459-465.

[50] MIOL M D, BIANCHI G, HENRY G, et al. Design of a single-shaft compressor, generator, turbine for small-scale supercritical CO_2 systems for waste heat to power conversion applications[C]// Proceedings of the 2nd European Supercritical CO_2 Conference, 2018: 1-8.

[51] STRAETZ M, MERTZ R, STRAFLINGER J. Experimental investigation on the heat transfer between condensing steam and SCO_2 in a compact heat exchanger[C]// Proceedings of the 2nd European Supercritical CO_2 Conference, 2018: 1-10.

[52] HACKS A, VOJACEK A, DOHMEN H J, et al. Experimental investigation of the SCO_2-HeRo compressor[C]// Proceedings of the 2nd European Supercritical CO_2 Conference, 2018: 1-10.

[53] WANG J, HUANG Y, ZANG J, et al. Research activities on supercritical carbon dioxide power conversion technology in China[C]// Proceedings of the ASME Turbo Expo 2014: Turbine Technical Conference and Exposition. New York: ASME, 2014: V03BT36A009.

[54] LI H Z, ZHANG Y F, GAO W, et al. Preliminary design study on a multi-megawatts fossil-based supercritical CO_2 recompression and reheat integral test facility[C]// Proceedings of the 5th International Symposium on Supercritical CO_2 Power Cycles, 2016: 1-17.

[55] 姚李超. 超临界二氧化碳布雷顿循环气动热力分析方法及建模技术[D]. 北京:北京航空航天大学, 2020.

YAO L C. Methodology for thermaodynamic analysis and modeling of supercritical carbon dioxide Brayton cycles[D]. Beijing: Beihang University, 2020(in Chinese).

[56] LEMMON E, HUBER M, MCLINDEN M. NIST standard reference database 23: Reference fluid thermodynamic and transport properties-REFPROP, Version 9. 1[S]. Gaithersburg: NIST, 2013.

[57] SPAN W, WAGNER W. A new equation of state for carbon dioxide covering the fluid region from the triple-point temperature to 1 100 K at pressures up to 800 MPa[J]. Journal of Physical and Chemical Reference Data, 1996, 25(6): 1509-1558.

[58] ZHAO Q, MECHERI M, NEVEUX T, et al. Selection of a proper equation of state for the modeling of a supercritical CO_2 Brayton cycle: Consequences on the process design[J]. Industrial & Engineering Chemistry Research, 2017, 56(23): 6841-6853.

[59] BALTADJIEV N D. An investigation of real gas effects in supercritical CO_2 compressors[D]. Cambridge: MIT, 2012.

[60] DYREBY J. Modeling the supercritical carbon dioxide Brayton cycle with recompression[D]. Madison: University of Wisconsin-Madison, 2014.

[61] DYREBY J, KLEIN S, NELLIS G, et al. Design considerations for supercritical carbon dioxide Brayton cycles with recompression[J]. Journal of Engineering for Gas Turbines and Power, 2014, 136(10): 101701.

[62] SARKAR J, BHATTACHARYYA S. Optimization of recompression S-CO_2 power cycle with reheating[J]. Energy Conversion and Management, 2009, 50(8): 1939-1945.

[63] BEJAN A, TSATSARONIS G, MORAN M. Thermal design and optimization[M]. New York: John Wiley & Sons, 1996.

[64] CHEN Y, WANG M, LISO V, et al. Parametric analysis and optimization for exergoeconomic performance of a combined system based on solid oxide fuel cell-gas turbine and supercritical carbon dioxide Brayton cycle[J]. Energy Conversion and Management, 2019, 186: 66-81.

[65] THANGANADAR D, ASFAND F, PATCHIGOLLA K. Thermal performance and economic analysis of supercritical carbon dioxide cycles in combined cycle power plant[J]. Applied Energy, 2019, 255: 113836.

[66] WANG X R, DAI Y P. Exergoeconomic analysis of utilizing the transcritical CO_2 cycle and the ORC for a recompression supercritical CO_2 cycle waste heat recovery: A comparative study[J]. Applied Energy, 2016, 170: 193-207.

[67] RAO Z H, XUE T C, HUANG K X, et al. Multi-objective optimization of supercritical carbon dioxide recompression Brayton cycle considering printed circuit recuperator design[J]. Energy Conversion and Management, 2019, 201: 112094.

[68] CARLSON M D, MIDDLETON B M, HO C K. Techno-economic comparison of solar-driven SCO_2 Brayton cycles using component cost models baselined with vendor data and estimates[C]// Proceedings of the ASME 2017 11th International Conference on Energy Sustainability. New York: ASME, 2017: 1-7.

[69] LI H, FAN G, CAO L Y, et al. A comprehensive investigation on the design and off-design performance of supercritical carbon dioxide power system based on the small-scale lead-cooled fast reactor[J]. Journal of Cleaner Production, 2020, 256: 120720.

[70] CHENG W L, HUANG W X, NIAN Y L. Global parameter optimization and criterion formula of supercritical carbon dioxide Brayton cycle with recompression[J]. Energy Conversion and Management, 2017, 150: 669-677.

[71] AKBARI A D, MAHMOUDI S. Thermoeconomic analysis & optimization of the combined supercritical CO_2 (carbon dioxide) recompression Brayton/organic Rankine cycle[J]. Energy, 2014, 78: 501-512.

[72] WANG K, HE Y L. Thermodynamic analysis and optimization of a molten salt solar power tower integrated with a recompression supercritical CO_2 Brayton cycle based on integrated modeling[J]. Energy Conversion and Management, 2017, 135: 336-350.

[73] 姚李超, 邹正平, 付超, 等. 超临界二氧化碳再压缩布雷顿循环性能分析及优化设计方法研究[J]. 推进技术, 2022, 43(3): 200452.
YAO L C, ZOU Z P, FU C, et al. Performance analysis and optimization design method of supercritical carbon dioxide recompression Brayton cycle[J]. Journal of Propulsion Technology, 2022, 43(3): 200452(in Chinese).

[74] ZHANG H Z, ZHAO H, DENG Q H, et al. Aerothermodynamic design and numerical investigation of supercritical carbon dioxide turbine[C]// Proceedings of the ASME Turbo Expo 2015: Turbine Technical Conference and Exposition. New York: ASME, 2015: V009T36A007.

[75] LIU Z Y, LUO W W, ZHAO Q J, et al. Preliminary design and model assessment of a supercritical CO_2 compressor[J]. Applied Sciences, 2018, 8(4): 595.

[76] GKOUNTAS A A, STAMATELOS A M, KALFAS A. Recuperators investigation for high temperature supercritical carbon dioxide power generation cycles[J]. Applied Thermal Engineering, 2017, 125: 1094-1102.

[77] KIM S, CHO Y, KIM M S, et al. Characteristics and optimization of supercritical CO_2 recompression power cycle and the influence of pinch point temperature difference of recuperators[J]. Energy, 2019, 147: 1216-1226.

[78] SAEED M, KIM M H. Analysis of a recompression supercritical carbon dioxide power cycle with an integrated turbine design/optimization algorithm[J]. Energy, 2018, 165: 93-111.

[79] SAEED M, KHATOON S, KIM M H. Design optimization and performance analysis of a supercritical carbon dioxide recompression Brayton cycle based on the detailed models of the cycle components[J]. Energy Conversion and Management, 2019, 196: 242-260.

[80] YAO L C, ZOU Z P. A one-dimensional design methodology for supercritical carbon dioxide Brayton cycles: Integration of cycle conceptual design and components preliminary design[J]. Applied Energy, 2020, 276: 115354.

[81] DYREBY J J, KLEIN S A, NELLIS G F, et al. Modeling off-design and part-load performance of supercritical carbon dioxide power cycles[C]// Proceedings of the ASME Turbo Expo 2013: Turbine Technical Conference and Exposition. New York: ASME, 2013: V008T34A014.

[82] JAHN I H, KEEP J A. On the off-design performance of supercritical carbon dioxide power cycles[C]// Proceedings of

Shanghai 2017 Global Power and Propulsion Forum, 2017: 1-10.

[83] BALTADJIEV N D, LETTIERI C, SPAKOVSZKY Z S. An investigation of real gas effects in supercritical CO_2 centrifugal compressors[J]. Journal of Turbomachinery, 2015, 137(9): 091003.

[84] CARSTENS N A. Control strategies for supercritical carbon dioxide power conversion systems[D]. Cambridge: MIT, 2007.

[85] HOOPES K, SANCHEZ 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 International Symposium on Supercritical CO_2 Power Cycles, 2016: 1-14.

[86] 朱玉铭,姜玉雁,梁世强,等. 超临界二氧化碳布雷顿发电循环压气机实验研究进展[J]. 热力发电, 2020, 49(10): 11-20.
ZHU Y M, JIANG Y Y, LIANG S Q, et al. Experimental research progress of supercritical carbon dioxide Brayton cycle compressor[J]. Thermal Power Generation, 2020, 49(10): 11-20(in Chinese).

[87] NOALL J, NICHOLS B, PASCH J. Achievable efficiency and stability of supercritical CO_2 compression systems[C]// Proceedings of the 4th International Symposium on Supercritical CO_2 Power Cycles, 2014: 1-10.

[88] UTAMURA M, HASUIKE H, OGAWA K, et al. Demonstration of supercritical CO_2 closed regenerative Brayton cycle in a bench scale experiment[C]// Proceedings of the ASME Turbo Expo 2012: Turbine Technical Conference and Exposition. New York: ASME, 2012: 155.

[89] ARITOMI M, ISHIZUKA T, MUTO Y, et al. Performance test results of the supercritical CO_2 compressor for a new gas turbine generating system[C]// Proceedings of the 18th International Conference on Nuclear Engineering. New York: ASME, 2010: 43-52.

[90] XU P C, ZOU Z P, XUAN L M. A hybrid performance prediction method for centrifugal compressors based on single-zone and two-zone models[J]. Aerospace Science and Technology, 2021, 108: 106358.

[91] AUNGIER R H. Mean streamline aerodynamic performance analysis of centrifugal compressors[J]. Journal of Turbomachinery, 1995, 117(3): 360.

[92] OH H W, YOON E S, CHUNG M K. An optimum set of loss models for performance prediction of centrifugal compressors[J]. Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy, 1997, 211(4): 331-338.

[93] JAPIKSE D. Assessment of single-and two-zone modeling of centrifugal compressors, studies in component performance: Part 3[C]// Proceedings of the ASME Turbo Expo 1985: Power for Land, Sea, and Air. New York: ASME, 1985: V001T03A023.

[94] JAPIKSE D. Centrifugal compressor design and performance[M]. White River Junction: Concepts NREC, 1996.

[95] 吴宇博. 向心涡轮非设计点低维性能预测和三维造型方法研究[D]. 北京:北京航空航天大学, 2021.
WU Z B. Research on performance prediction of radial turbine and aerodynamic design of turbine blades[D]. Beijing: Beihang University, 2021(in Chinese).

[96] 王志强. SCO_2再压缩循环发电系统再压缩压气机设计技术研究[D]. 北京:北京航空航天大学, 2020.
WANG Z Q. Research on recompression compressor design technology for supercritical carbon dioxide recompression cycle power generation system[D]. Beijing: Beihang University, 2020(in Chinese).

[97] 应祺煜. 超临界二氧化碳轴流涡轮旋涡流动机理与控制方法研究[D]. 北京:清华大学, 2022.
YING Q Y. Vortex flow mechanism investigations and flow control method study in a supercritical carbon dioxide axial turbine[D]. Beijing: Tsinghua University, 2022(in Chinese).

[98] ZOU Z P, DING C. A new similarity method for turbomachinery with different working media[J]. Applied Thermal Engineering, 2018, 133: 170-178.

[99] 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.

[100] XU P C, ZOU Z P, YAO L C. A unified performance conversion method for similar compressors working with different gases based on polytropic analysis and deep-learning improvement[J]. Energy Conversion and Management, 2021, 247: 114747.

[101] WRIGHT S, PICKARD P, VERNON M, et al. Description and test results from a supercritical CO_2 Brayton cycle development program: AIAA 2009-4607[R]. Reston: AIAA, 2009.

[102] CONBOY T, WRIGHT S, PASCH J, et al. Performance characteristics of an operating supercritical CO_2 Brayton cycle[J]. Journal of Engineering for Gas Turbines and Power, 2012, 134(11): 111703.

[103] CHA J E, BAE S W, LEE J, et al. Operation results of a closed supercritical CO_2 simple Brayton cycle[C]// Proceedings of the 5th International Symposium on Supercritical CO_2 Power Cycles, 2016: 1-10.

[104] YASUYOSHI K, TAKESHI N, YOSHIO Y. A carbon dioxide partial condensation direct cycle for advanced gas cooled fast and thermal reactors[C]// Proceedings of the Global 2001 International Conference on: "Back-End of the Fuel Cycle: From Research to Solutions", 2001: 1-8.

[105] KATO Y, NITAWAKI T, MUTO Y. Medium temperature carbon dioxide gas turbine reactor[J]. Nuclear Engineering and Design, 2004, 230(1-3): 195-207.

[106] MECHERI M, MOULLEC Y L. Supercritical CO_2 Brayton cycles for coal-fired power plants[J]. Energy, 2016, 103: 758-771.

[107] LI Q, FLAMANT G, YUAN X, et al. Compact heat exchangers: A review and future applications for a new generation of high temperature solar receivers[J]. Renewable and Sustainable Energy Reviews, 2011, 15(9): 4855-4875.

[108] ZHANG H Z, CHENG K Y, HUAI X L, et al. Experimental and numerical study of an 80-kW zigzag printed circuit heat exchanger for supercritical CO_2 Brayton cycle[J]. Journal of Thermal Science, 2021, 30(4): 1289-1301.

[109] CAI H F, JIANG Y Y, WANG T, et al. Experimental investigation on convective heat transfer and pressure drop of supercrit-

ical CO_2 and water in microtube heat exchangers[J]. International Journal of Heat and Mass Transfer,2020,163:120443.

[110] SULLIVAN S D,FARIAS J,KESSELI J,et al. Mechanical design and validation testing for a high-performance supercritical carbon dioxide heat exchanger[C]// Proceedings of the ASME Turbo Expo 2017:Turbomachinery Technical Conference and Exposition. New York:ASME,2017:V009T38A012.

[111] RASOULI E,DAVIS C A,SUBEDI S,et al. Design and performance characterization of an additively manufactured primary heat exchanger for sCO_2 waste heat recovery cycles[C]// Proceedings of the 6th International Symposium on Supercritical CO_2 Power Cycles,2018:1-12.

[112] HAO J H,GE Z H,CHEN Q,et al. The coupled process-component modeling and optimization for heat exchanger of supercritical CO_2 with property variation based on heat current method[J]. Applied Thermal Engineering,2020,168:114833.

[113] KWON J S,SON S,HEO J Y,et al. Compact heat exchangers for supercritical CO_2 power cycle application[J]. Energy Conversion and Management,2020,209:112666.

[114] CABEZA L F,GRACIA A D,FERNÁNDEZ A I,et al. Supercritical CO_2 as heat transfer fluid:A review[J]. Applied Thermal Engineering,2017,125:799-810.

[115] EHSAN M M,GUAN Z Q,KLIMENKO A Y. A comprehensive review on heat transfer and pressure drop characteristics and correlations with supercritical CO_2 under heating and cooling applications[J]. Renewable and Sustainable Energy Reviews,2018,92:658-675.

[116] 李辉,汝卓霖,邹正平,等. 微小尺度通道内超临界甲烷传热特征研究[J]. 南京航空航天大学学报,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,2021,53(4):513-520(in Chinese).

[117] DUFFEY R B,PIORO I L. Experimental heat transfer of supercritical carbon dioxide flowing inside channels (survey)[J]. Nuclear Engineering and Design,2005,235(8):913-924.

[118] JIANG P X,SHI R S,XU Y J,et al. Experimental investigation of flow resistance and convection heat transfer of CO_2 at supercritical pressures in a vertical porous tube[J]. The Journal of Supercritical Fluids,2006,38(3):339-346.

[119] ZHANG J Q,ZOU Z P,WANG Y F. Large-eddy simulation on the aerodynamic and thermal characteristics in a micropipe of the hypersonic engine precooler[J]. Micromachines,2022,13(4):637.

[120] JIANG P X,ZHANG Y,XU Y J,et al. Experimental and numerical investigation of convection heat transfer of CO_2 at supercritical pressures in a vertical tube at low Reynolds numbers[J]. International Journal of Thermal Sciences,2008,47(8):998-1011.

[121] NARASIMHA R,SREENIVASAN K R. Relaminarization of fluid flows[J]. Advances in Applied Mechanics,1979,19:221-309.

[122] BAE J H,YOO J Y. Direct numerical simulation of turbulent supercritical flows with heat transfer[J]. Physics of Fluids,2005,17(10):105104.

[123] ZHANG S J,XU X X,LIU C,et al. A review on application and heat transfer enhancement of supercritical CO_2 in low-grade heat conversion[J]. Applied Energy,2020,269:114962.

[124] 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.

[125] 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.

[126] 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.

[127] 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.

[128] JIANG Y,LIESE E,ZITNEY S E,et al. Design and dynamic modeling of printed circuit heat exchangers for supercritical carbon dioxide Brayton power cycles[J]. Applied Energy,2018,231:1019-1032.

[129] GUO J F. Design analysis of supercritical carbon dioxide recuperator[J]. Applied Energy,2016,164:21-27.

[130] KE H B,XIAO Q,CAO Y P,et al. Simulation of the printed circuit heat exchanger for S-CO_2 by segmented methods[J]. Energy Procedia,2017,142:4098-4103.

[131] 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).

[132] 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.

[133] SON S,LEE Y,LEE J I. Development of an advanced printed circuit heat exchanger analysis code for realistic flow path configurations near header regions[J]. International Journal of Heat and Mass Transfer,2015,89:242-250.

[134] 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.

[135] JIANG Y,LIESE E,ZITNEY S E,et al. Optimal design of microtube recuperators for an indirect supercritical carbon dioxide recompression closed Brayton cycle[J]. Applied Energy,2018,216:634-648.

[136] CHAI L,TASSOU S A. A review of printed circuit heat exchangers for helium and supercritical CO_2 Brayton cycles[J].

Thermal Science and Engineering Progress,2020,18:100543.

[137] LIU S H,HUANG Y P,WANG J F,et al. Experimental study on transitional flow in straight channels of printed circuit heat exchanger[J]. Applied Thermal Engineering, 2020, 181: 115950.

[138] ANEESH A,SHARMA A,SRIVASTAVA A,et al. Effects of wavy channel configurations on thermal-hydraulic characteristics of printed circuit heat exchanger (PCHE)[J]. International Journal of Heat and Mass Transfer,2018,118:304-315.

[139] TSUZUKI N,KATO Y,ISHIDUKA T. High performance printed circuit heat exchanger[J]. Applied Thermal Engineering, 2007,27(10):1702-1707.

[140] 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.

[141] ZHANG K,LI M J,LIU H,et al. A general and rapid method to evaluate the effect of flow maldistribution on the performance of heat exchangers[J]. International Journal of Thermal Sciences,2021,170:107152.

[142] HAIDER P,FREKO P,ACHER T,et al. Influence of inlet configuration and distributor geometry on the performance of cryogenic plate-fin heat exchangers[J]. Applied Thermal Engineering,2021,195:117197.

[143] PIDAPARTI S R,ANDERSON M H,RANJAN D. Experimental investigation of thermal-hydraulic performance of discontinuous fin printed circuit heat exchangers for supercritical CO_2 power cycles[J]. Experimental Thermal and Fluid Science, 2019,106:119-129.

[144] CHU W X,BENNETT K,CHENG J,et al. Numerical study on a novel hyperbolic inlet header in straight-channel printed circuit heat exchanger[J]. Applied Thermal Engineering,2019, 146:805-814.

[145] SIDDIQUI O K,ZUBAIR S M. Efficient energy utilization through proper design of microchannel heat exchanger manifolds:A comprehensive review[J]. Renewable and Sustainable Energy Reviews,2017,74:969-1002.

[146] KOO G W,LEE S M,KIM K Y. Shape optimization of inlet part of a printed circuit heat exchanger using surrogate modeling[J]. Applied Thermal Engineering,2014,72(1):90-96.

[147] MA T,ZHANG P,SHI H N,et al. Prediction of flow maldistribution in printed circuit heat exchanger[J]. International Journal of Heat and Mass Transfer,2020,152:119560.

[148] MENG B,WAN M,ZHAO R,et al. Micromanufacturing technologies of compact heat exchangers for hypersonic precooled airbreathingpropulsion:A review[J]. Chinese Journal of Aeronautics,2021,34(2):79-103.

[149] MYLAVARAPU S K,SUN X D,CHRISTENSEN R N,et al. Fabrication and design aspects of high-temperature compact diffusion bonded heat exchangers[J]. Nuclear Engineering and Design,2012,249:49-56.

[150] MA Y,XIE G N,HOOMAN K. Review of printed circuit heat exchangers and its applications in solar thermal energy[J]. Renewable and Sustainable Energy Reviews,2022,155:111933.

[151] OH C H,KIM E S,PATTERSON M. Design option of heat exchanger for the next generation nuclear plant[J]. Journal of Engineering for Gas Turbines and Power, 2010, 132(3): 032903.

[152] KAUR K,SINGH P. State-of-the-art in heat exchanger additive manufacturing[J]. International Journal of Heat and Mass Transfer,2021,178:121600.

[153] DING M,LIU J,CHENG W L,et al. An adaptive flow path regenerator used in supercritical carbon dioxide Brayton cycle[J]. Applied Thermal Engineering,2018,138:513-522.

[154] CARLSON M D,KRUIZENGA A K,SCHALANSKY C,et al. Sandia progress on advanced heat exchangers for SCO_2 Brayton cycles[C]// Proceedings of the 4th International Symposium on Supercritical CO_2 Power Cycles,2014:1-10.

[155] MOISSEYTSEV A,SIENICKI J J,CHO D H,et al. Comparison of heat exchanger modeling with data from CO_2-to-CO_2 printed circuit heat exchanger performance tests[C]// Proceedings of the International Congress on Advances in Nuclear Power Plants,2010:459-468.

[156] 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.

[157] BAIK S,KIM S G,LEE J,et al. Study on CO_2-water printed circuit heat exchanger performance operating under various CO_2 phases for S-CO_2 power cycle application[J]. Applied Thermal Engineering,2017,113:1536-1546.

[158] ZHOU J Z,CHENG K Y,ZHANG H Z,et al. Test platform and experimental test on 100 kW class printed circuit heat exchanger for supercritical CO_2 Brayton cycle[J]. International Journal of Heat and Mass Transfer,2020,148:118540.

[159] CHU W X,LI X H,MA T,et al. Experimental investigation on SCO_2-water heat transfer characteristics in a printed circuit heat exchanger with straight channel[J]. International Journal of Heat and Mass Transfer,2017,113:184-194.

[160] KIM I H,NO H C. Thermal hydraulic performance analysis of a printed circuit heat exchanger using a helium-water test loop and numerical simulations[J]. Applied Thermal Engineering, 2011,31(17-18):4064-4073.

[161] KIM I H,NO H C. Thermal-hydraulic physical models for a printed circuit heat exchanger covering He,He-CO_2 mixture, and water fluids using experimental data and CFD[J]. Experimental Thermal and Fluid Science,2013,48:213-221.

[162] 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.

[163] 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.

[164] ZHAO J W,ZHAO R,NIAN Y L,et al. Experimental study of supercritical CO_2 in a vertical adaptive flow path heat exchanger[J]. Applied Thermal Engineering,2021,188:116597.

[165] WANG X,WANG R,BIAN X Y,et al. Review of dynamic performance and control strategy of supercritical CO_2 Brayton cy-

cle[J]. Energy and AI,2021,5:100078.

[166] LIESE E,ALBRIGHT J,ZITNSY S A. Start up,shutdown and load-following simulations of a 10 MWe supercritical CO_2 recompression closed Brayton cycle[J]. Applied Energy,2020,277:115628.

[167] MOISSEYTSEV A,SIENICKI J J. Investigation of plant control strategies for the supercritical CO_2 Brayton cycle for a sodium-cooled fast reactor using the plant dynamics code:ANL-GenIV-147[R]. Argonne:Argonne National Laboratory,2010.

[168] LUU M T,MILANI D,MCNAYUGHTON R,et al. Dynamic modeling and start-up operation of a solar-assisted recompression supercritical CO_2 Brayton power cycle[J]. Applied Energy,2017,199:247-263.

[169] MILANI D,LUU M T,MCNAYUGHTON R,et al. A comparative study of solar heliostat assisted supercritical CO_2 recompression Brayton cycles:Dynamic modelling and control strategies[J]. The Journal of Supercritical Fluids,2017,120:113-124.

[170] OLUMAYEGUN O,WANG M. Dynamic modelling and control of supercritical CO_2 power cycle using waste heat from industrial processes[J]. Fuel,2019,249:89-102.

[171] LUU M T,MILANI D,MCNAYUGHTON R,et al. Advanced control strategies for dynamic operation of a solar-assisted recompression supercritical CO_2 Brayton cycle[J]. Applied Thermal Engineering,2018,136:682-700.

[172] MATTEO M,GIUSEPPE B,SAVVAS A T. Transient analysis and control of a heat t power conversion unit based on a simple regenerative supercritical CO_2 Joule-Brayton cycle[J]. Applied Thermal Engineering,2021,183:116214.

[173] DU X,HU J S,XIA G L. Operation characteristic of supercritical carbon dioxide-cooled reactor system under coordination control scheme[J]. International Journal of Advanced Robotic Systems,2020,17(3):172988142093383.

[174] CONBOY T,PASCH J,FLEMING D. Control of a supercritical CO_2 recompression Brayton cycle demonstration loop[J]. Journal of Engineering for Gas Turbines and Power,2013,135(11):111701.

[175] GAO C T,WU P,LIU W H,et al. Development of a bypass control strategy for supercritical CO_2 Brayton cycle cooled reactor system under load-following operation[J]. Annals of Nuclear Energy,2021,151:107917.

[176] MOISSEYTSEV A,SIENICKI J. Recent development in S-CO_2 cycle dynamic modeling and analysis at ANL[C]// Proceedings of the 4th International Symposium on Supercritical CO_2 Power Cycles,2014:1-21.

[177] VIGNAROOBAN K,XU X H,WANG K,et al. Vapor pressure and corrosivity of ternary metal-chloride molten-salt based heat transfer fluids for use in concentrating solar power systems[J]. Applied Energy,2015,159:206-213.

[178] WU Y T,REN N,WANG T,et al. Experimental study on optimized composition of mixed carbonate salt for sensible heat storage thermal power plant[J]. Solar Energy,2011,85(9):1957-1966.

[179] HO C K,IVERSON B D. Review of high-temperature central receiver designs for concentrating solar power[J]. Renewable and Sustainable Energy Reviews,2014,29:835-846.

[180] KELLY B. Advanced thermal storage for central receivers with supercritical coolants:DE-FG36-08GO18149[R]. Lakewood:Abengoa Solar Inc,2010.

[181] BINDRA H,BUENO P,MORRIS J F. Sliding flow method for exergetically efficient packed bed thermal storage[J]. Applied Thermal Engineering,2014,64(1-2):201-208.

[182] ZHAO W H,FRANCE D M,YU W H,et al. Phase change material with graphite foam for applications in high-temperature latent heat storage systems of concentrated solar plant[J]. Renewable Energy,2014,69:134-146.

[183] STEKLI J,IRWIN L,PITCHUMANI R. Technical challenges and opportunities for concentrating solar power with thermal energy storage[J]. Journal of Thermal Science and Engineering Applications,2013,5:021011.

[184] CONBOY T. An approach to turbomachinery for supercritical Brayton space power cycles[C]// Proceedings of Nuclear and Emerging Technologies for Space 2013,2013:1-8.

Progress in research of closed supercritical carbon dioxide Brayton cycle system

ZOU Zhengping[1,2], WANG Yifan[2,3], YAO Lichao[1,2,*], LIU Huoxing[1,2], XU Pengcheng[2,3], LI Hui[2,3]

(1. Research Institute of Aero-Engine, Beihang University, Beijing 102206, China;
2. National Key Laboratory of Science and Technology on Aero-Engine Aero-Thermodynamics, Beijing 100083, China;
3. School of Energy and Power Engineering, Beihang University, Beijing 100083, China)

Abstract: The supercritical carbon dioxide (SCO_2) closed Brayton cycle has received considerable attention in the field of energy and power due to its advantages of high thermal efficiency and compactness, economy, and environment friendliness. This study reviews research on the SCO_2 closed Brayton cycle in terms of its operating principle, advantages, and domestic and overseas research progress. Key techniques such as cycle thermodynamics, turbomachinery working with supercritical medium, high-efficiency compact heat exchangers, control strategies, and thermal storage are analyzed. Moreover, difficulties in and challenges of engineering applications are discussed, and future development directions are presented. It is indicated that the low-dimensional performance analysis of components should be involved in the conceptual design of cycle thermodynamics to estimate the attainable performance of components, considering the lifecycle performance, compactness, and economy of the cycle. The dramatic variations of working medium properties would lead to special flow and heat transfer mechanisms in turbomachinery and heat exchanger, respectively, motivating the design methodologies for turbomachinery and compact heat exchanger that fully consider the effect of special properties of working medium. The similarity method constructed by theoretical analysis and deep machine learning could provide theoretical foundations for experimental validation of the aerodynamic performance of SCO_2 turbomachinery. Furthermore, robust and efficient control strategies could control the cycle effectively, and SCO_2 closed Brayton cycles integrated with thermal storage techniques adopting novel thermal medium would provide key technical supports for concentrating solar power at high operation temperature.

Keywords: supercritical carbon dioxide (SCO_2); closed Brayton cycle; thermodynamic modeling and analysis; turbomachinery; compact heat exchanger

Received: 2022-03-29; **Accepted**: 2022-04-22; **Published online**: 2022-05-17 17:06

URL: kns.cnki.net/kcms/detail/11.2625.V.20220517.1414.002.html

Foundation item: Fundation of National Key Laboratory of Science and Technology on Aero-Engine Aero-Thermodynamics (2021-JCJQ-LB-062-0205)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0141

基于边界元法的三维结构体滑跳运动数值仿真

杨超, 姜宇, 吴志刚*

(北京航空航天大学 航空科学与工程学院, 北京 100083)

摘 要: 利用滑跳运动实现近水面高速机动飞行的击水式飞行器是近年来的热点问题,三维结构体高速斜入水冲击载荷求解是滑跳运动数值仿真的关键步骤。针对滑跳运动开展研究,建立了一种基于边界元法的可以考虑结构弹性效应的三维结构体近水面滑跳运动时域数值仿真方法。高速斜入水冲击载荷由边界元法求得,利用状态方程直接积分法求解得到结构节点位移动力响应,并通过滑跳运动动力学模型更新结构体的质心位置及运动速度。针对钢球高速斜入水冲击现象进行数值仿真并与试验结果对比,验证了所提方法的准确性。开展三维刚性球冠体的近水面滑跳运动时域数值仿真和变参分析,得到了结构质量、初始高度、水平抛出速度和半径大小4种参数对球冠体滑跳运动的影响规律。考虑结构弹性效应后,高速斜入水冲击载荷减小,对球冠体滑跳运动也有一定影响。

关 键 词: 滑跳运动; 边界元法; 入水冲击; 流固耦合; 球冠体

中图分类号: O353.4

文献标志码: A **文章编号**: 1001-5965(2022)09-1678-14

水面滑跳是结构体近水面小角度斜撞水过程中发生的一种特殊运动,日常生活中在河边进行的“打水漂游戏”就是一种典型的滑跳现象[1]。结构体在滑跳运动过程中,不断地以较大的水平速度和较小的垂向速度斜向撞击水面,损失一部分动能的同时,还会被水面弹起。近年来,不断有学者提出利用这种滑跳运动来实现近水面高速机动飞行的新概念跨介质飞行器[2-5],也称为击水式飞行器。因此,水面滑跳问题作为跨介质飞行器的基础问题也逐渐受到关注。

结构体近水面滑跳问题涉及复杂多相流动与结构体大幅运动的相互作用,尤其是结构体高速斜入水冲击过程中带来的受力剧烈变化。目前,国内外对结构物高速斜入水冲击载荷问题的研究,主要集中在结构体垂向砰击水面带来的载荷,在工程应用领域具有一定的局限性。例如,Zhao和Faltinsen[6]在1993年针对二维楔形体的垂直入水进行了计算;樊征和刘莹莹[7]采用解析法分析了二维小升角楔形体的垂向入水问题;2008年,许国冬等[8]讨论了V形楔多种入水形式的相似解;Faltinsen和Chezhian[9]以非轴对称的船体模型入水为研究对象,基于Wagner理论,研究了三维垂向入水砰击的数值方法。也有部分研究考虑了结构体的弹性效应。例如,1991年,顾懋祥等[10]采用迭代法计算了平头旋转壳体垂直入水时的水弹性效应;Takagi[11]在1994年运用Wagner的平板近似理论计算了弹性结构的入水冲击问题;2017—2020年,张岳青等[12-13]对楔形体和弧形体的垂直入水进行了试验和商用软件数值仿真;2018年,Mohammad和Maurizio[14]对柔性细长体的垂直入水冲击载荷进行了理论分析和试验验证;2004—2006年,Wu等[15-16]提出了二维楔形体的迭代求解方法,并运用二维边界元法,使用柯西积分计算了楔形体入水的相似解。边界元法是基

收稿日期: 2022-03-11; **录用日期**: 2022-05-27; **网络出版时间**: 2022-06-02 11:26
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220601.1111.001.html
* **通信作者**. E-mail: wuzhigang@buaa.edu.cn

引用格式: *杨超, 姜宇, 吴志刚. 基于边界元法的三维结构体滑跳运动数值仿真[J]. 北京航空航天大学学报, 2022, 48(9): 1678-1691. YANG C, JIANG Y, WU Z G. Numerical simulation of skipping motion of three-dimensional structure based on boundary element method[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1678-1691 (in Chinese).*

于边界积分方程,采用与有限元法[17]类似的分元、离散思想建立起来的。由于边界积分方程是定义在边界上的,边界元法只要在边界上划分单元,相比于区域解法降低了维度,具备较高的计算效率和计算精度[18]。相对而言,结构体高速斜入水冲击问题的研究较少。2008 年,王永虎、石秀华等[19]以入水冲击的理论动力学和入水弹道学理论为基础,完成了鱼雷头部尖拱的斜入水高速冲击载荷的分析;2009 年,陈占晖和卢永锦[20]利用流固耦合有限元法研究了立方体斜向冲击水面的运动特性;2012 年,郑金伟和宗智[21]利用 LS-DYNA 软件计算了三维刚性椭圆头结构的高速斜入水冲击载荷;2014 年,Scolan[22]完成了三维不规则外形的刚体结构低速斜入水冲击载荷试验;2021 年,蔡维轩等[23]完成了无旋圆盘的滑跳运动试验;同年,王聪等[24]对旋转圆盘的滑跳特性进行了理论仿真。可以看出,考虑结构弹性效应影响的三维结构体高速斜入水冲击载荷数值求解算法的研究较少。

为了高效求解三维结构体滑跳运动问题,本文基于边界元法,建立了一种可以考虑弹性效应的三维结构体近水面滑跳运动的时域数值仿真方法。在三维结构体的滑跳运动过程中,高速斜入水冲击载荷由边界元法求得,利用状态方程直接积分法(direct integration of state equation method)求解得到结构节点位移动力响应,并通过滑跳运动动力学模型更新结构体的质心位置及运动速度。这使得本文方法既可以应用于三维弹性结构体,也拥有较高的计算效率。针对刚球高速斜入水冲击现象进行了数值仿真,并与试验值进行了对比,验证了该方法流体载荷求解部分的准确性。针对三维球冠体的近水面滑跳运动进行了时域数值仿真,得到了对应的运动特性。研究了不同的初始参数(结构质量、初始高度、水平抛出速度和半径大小)和结构弹性效应对球冠体近水面滑跳运动的影响,总结了相应的规律。

1　问题描述

本文分析的三维结构体近水面滑跳运动问题可以由图 1 表示,半径为 r_0 的球冠体结构在质心距离水面高度 H_0 的位置上以速度 V_0 水平抛出。图中采用了如下坐标系描述:流场坐标系 O-xyz,原点固定在流场正中心的 O 点,x 轴和 y 轴固定在与 O 点同一高度的水平面上,z 轴垂直向上。结构坐标系 O'-$x'y'z'$,原点固定在球冠体的质心 O'点,x'轴和 y'轴固定在与 O'点同一高度的水平面上,z'轴垂直向上。图 1 中,Ω 为流体域,S 为流体域边界。

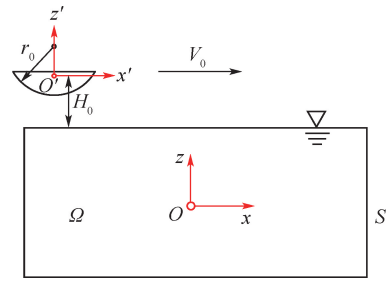


图 1　球冠体近水面滑跳运动问题示意图

Fig. 1　Schematic diagram of near-water-surface skipping motion of hemispherical structure

2　理论方法

三维结构体高速斜入水冲击是流体和结构相互作用的强非线性过程。在很多情况下,流体运动的速度比声速小得多,因此,除平头物体垂直入水外,一般可以忽略流体的可压缩性[25]。入水冲击载荷具有瞬态极大值的特点,与之相比,入水冲击时流体的重力因素所造成的影响较小,因此作为简化条件,忽略流体重力的影响。

本文建立的基于边界元法的三维弹性结构体近水面滑跳运动时域迭代仿真方法的基本原理如图 2 所示。三维弹性结构体在近水面滑跳过程中受到斜入水冲击载荷、摩擦阻力、重力和浮力的共同作用,每个时间步内流场作用在三维结构体上的斜入水冲击载荷由边界元法求解得到,利用状态方程直接积分法求解得到结构节点位移动力响应。流场载荷连同重力和浮力代入到滑跳运动动力学模型中进行分析,得到结构体新的质心位置及运动速度。重新构造流场的形状和边界条件,并进入下一步迭代计算。

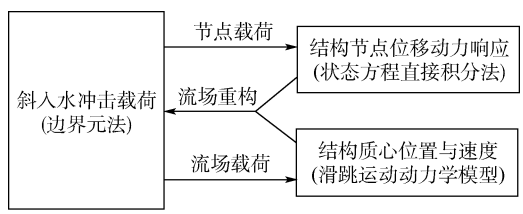


图 2　时域迭代求解方法基本原理

Fig. 2　Basic principle of time domain iterative solution method

2.1　边界元法

根据流体区域的简化和假设,流体为无旋、无黏、不可压流体,则三维流体运动的速度势 ϕ 满足如下 Laplace 方程:

$$\nabla^2\phi = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\phi = 0 \tag{1}$$

为了确定方程(1)在空间某个内部域或外部域的解,还必须附加定解条件,即边界条件。对于本文分析的问题,满足 Laplace 方程的第三边界条件,即混合条件。流体区域自由表面上的边界条件为

$$\phi = 0 \tag{2}$$

流体区域底面和侧面边界的法向速度为 0,则边界条件为

$$\frac{\partial \phi}{\partial \boldsymbol{n}} = 0 \tag{3}$$

三维结构体浸湿表面的法向速度为结构体运动速度的法向分量,则对应的边界条件为

$$\frac{\partial \phi}{\partial \boldsymbol{n}} = \boldsymbol{V} \cdot \boldsymbol{n} \tag{4}$$

式中:$\boldsymbol{n}$ 为流体域的法向单位矢量,指向流体域外为正;$\frac{\partial}{\partial \boldsymbol{n}}$为沿该法向的导数;$\boldsymbol{V}$ 为结构体相对全局坐标系的运动速度矢量。

对于三维 Laplace 方程,其基本解 ϕ^* 应满足条件[15]:

$$\nabla^2 \phi^*(P;Q) = -\delta(P,Q) \tag{5}$$

式中:$P(x_P, y_P, z_P)$ 点代表域内源点,x_P、y_P、z_P 代表域内源点的坐标;$Q(x_Q, y_Q, z_Q)$ 点代表域内场点,x_Q、y_Q、z_Q 代表域内场点的坐标;$\delta(P,Q)$ 为以 P 点为奇异点 Dirac Delta 函数。

求解之后得到三维 Laplace 方程的基本解:

$$\phi^*(P;Q) = \frac{1}{4\pi r(P,Q)} \tag{6}$$

式中:r 为任意场点 Q 和 P 源点之间的距离,表示为

$$r(P,Q) = \sqrt{(x_Q - x_P)^2 + (y_Q - y_P)^2 + (z_Q - z_P)^2} \tag{7}$$

若以 ϕ 为所求三维 Laplace 方程边值问题的解,而 ϕ^* 作为辅助解 ψ 代入 Green 第二公式(由 Gauss 公式推得):

$$\int_{\Omega} (\psi \nabla^2 \phi - \phi \nabla^2 \psi) \mathrm{d}\Omega = \int_S (\psi \nabla \phi - \phi \nabla \psi) \cdot \boldsymbol{n} \mathrm{d}S \tag{8}$$

可以将三维 Laplace 方程表示为边界积分的形式,得到流体域 Ω 上任一点三维 Laplace 方程的解用边界变量的积分来表示的表达式:

$$\phi(P) = \int_S \left[\phi^*(P;q) \frac{\partial \phi}{\partial \boldsymbol{n}}(q) - \frac{\partial \phi^*(P;q)}{\partial \boldsymbol{n}(q)} \phi(q) \right] \mathrm{d}S(q) \quad \forall P \in \Omega \tag{9}$$

式(9)中的积分是对 $\mathrm{d}S(q)$ 积分,即在边界面对场点进行积分,q 代表边界面上的场点。

为了建立边界积分方程,只需将基本解的域内源点 P 从流体域内趋向于边界面源点 p,于是得到

$$\alpha \phi(p) = \int_S \left[\phi^*(p;q) \frac{\partial \phi}{\partial \boldsymbol{n}}(q) - \frac{\partial \phi^*(p;q)}{\partial \boldsymbol{n}(q)} \phi(q) \right] \mathrm{d}S(q) \quad \forall p \in S \tag{10}$$

式中:α 为边界面上源点 p 处的立体角系数。

对于本文分析的三维结构体,进行常值三角形单元的边界划分,如图 3 所示。对于常值三角形单元来说,相邻单元间未知量本身都不连续。在每个边界元上 ϕ 和 $\partial\phi/\partial \boldsymbol{n}$ 都是常量,并用该边界元中点的值来代替。

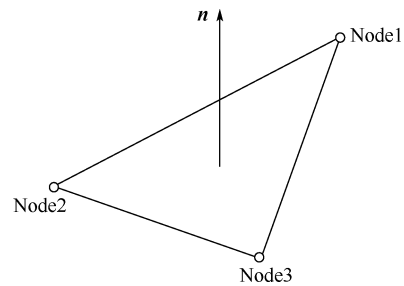


图 3 常值三角形单元

Fig. 3 Constant triangular element

p 点取遍所有边界单元的中点,就可以得到如下矩阵:

$$\boldsymbol{H} \cdot \boldsymbol{\phi} = \boldsymbol{G} \cdot \frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{n}} \tag{11}$$

式中:$\boldsymbol{H}$ 和 $\boldsymbol{G}$ 分别为 $\boldsymbol{\phi}$ 和 $\partial\boldsymbol{\phi}/\partial \boldsymbol{n}$ 的积分系数矩阵,仅仅与求解域的几何特性有关。

通过求解式(11),就可以得到三维结构体的入水过程中任意边界元中点的速度势。在求得边界单元中点处的势函数值后,可以用一种简便的数值微分方法[26-27]计算浸湿表面上速度分布。设浸湿表面方程为 $y = y(x,z)$,势函数在浸湿表面上的增量可以写成

$$\Delta\phi = u\Delta x + v\Delta y + w\Delta z = u\Delta x + v\left(\frac{\partial y}{\partial x}\Delta x + \frac{\partial y}{\partial z}\Delta z\right) + w\Delta z = u^*\Delta x + w^*\Delta z \tag{12}$$

式中:$u^* = u + v\frac{\partial y}{\partial x}$;$w^* = w + \frac{\partial y}{\partial z}$。

在该边界元中点附近任意找到 2 个点,2 次使用式(12),便可求解出等效速度 u^* 和 w^*。再根据图 4,用式(13)求解速度矢量:

$$\begin{cases} u = k(ku^* - lw^*) + m^2 u^* \\ v = -lmu^* - kmw^* \\ w = m^2 w^* - l(ku^* - lw^*) \end{cases} \tag{13}$$

式中:u、v、w 分别为边界元中点速度的 3 个分量;l、m、k 分别为边界元中点单位法向量的 3 个分量。

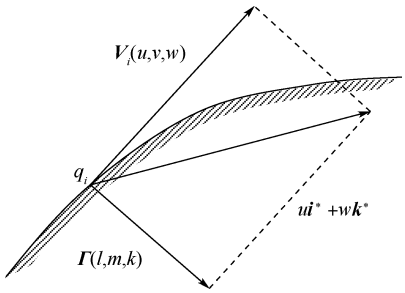


图 4　速度分量求解示意图

Fig. 4　Schematic diagram of velocity component solution

图 4 中,q_i 为控制点,$\boldsymbol{V}_i$ 为控制点速度向量,$\boldsymbol{\Gamma}$ 为控制点单位法向量,$\boldsymbol{i}$ 和 $\boldsymbol{j}$ 分别为 x 坐标方向和 z 坐标方向的单位向量。

在计算得到每个边界单元中心点的速度和速度势后,就可以利用下述改进的非线性伯努利方程求得每个边界单元对应的水动力压强[28]:

$$p_i = -\rho\left(\frac{\mathrm{d}\phi}{\mathrm{d}t} - \boldsymbol{V}\cdot\nabla\phi + \frac{1}{2}\nabla\phi\cdot\nabla\phi\right) \tag{14}$$

式中:p_i 为边界单元的压强;ρ 为流体密度。

最后,每个边界单元的压强乘以每个边界单元的面积,即可得到每个边界单元对结构体的作用力。

2.2　高速斜入水冲击载荷模型

由于浸没在水中的体积排挤流体产生液面隆起现象[29],又称 Wagner 型入水冲击,如图 5 所示。取决于入水结构体的外形和入水角等初始状态,下面给出确定液面隆起等效位置 H_e 的表达式,即

$$H_e = \eta(x,y,z,\theta,t) \tag{15}$$

式中:θ 为入水倾角;t 为时间;$\eta(x,y,z,\theta,t)$ 为函数表达式。参考文献[30],圆球体倾斜入水时,H_e 的取值固定为 1.35 倍的初始液面入水深度。考虑自由液面的隆起现象后,会导致结构体浸湿表面扩大,高速斜入水冲击载荷增大。

结构体高速斜入水冲击载荷可以分为较大的水平速度带来的划水力和较小的垂向速度带来的垂向入水砰击载荷 2 部分。当结构体高速斜入水时,流体被排开,因此球冠体被浸湿的表面只有与速度矢量 $\boldsymbol{V}$ 夹角小于 90°且低于液面隆起等效位置 H_e 的部分[31-32],示意图如图 5 所示。同样,只有这部分浸湿表面会受到高速斜入水冲击载荷。

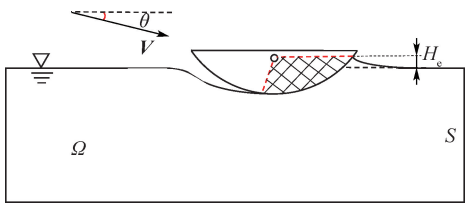


图 5　浸湿表面示意图

Fig. 5　Schematic diagram of water soaked surface

对于结构体浸湿表面的三角形常值单元,计算高速斜入水冲击载荷时的边界条件为结构体水平运动速度 V_x 的法向分量、垂向运动速度 V_z 的法向分量与结构体旋转角速度 ω 带来的速度 V_ω 的法向分量之和,示意图如图 6 所示。

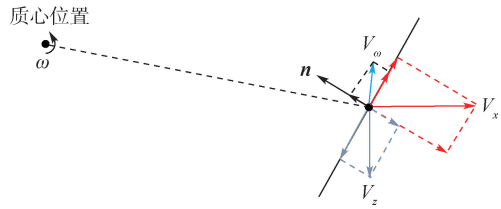


图 6　斜入水冲击载荷模型单元边界条件

Fig. 6　Element boundary conditions of oblique water impact load model

当使用边界元法计算斜入水冲击载荷时,假设流体为理想流体。而实际流体是有黏性的,因此在斜入水冲击载荷模型中,结构体与流体相接触的表面受到摩擦阻力的作用。摩擦阻力系数 c_f 可以直接选用帕兰特-许立汀紊流摩阻计算公式[3,33],即

$$c_f = \frac{0.075}{(\lg Re - 2)^2} \tag{16}$$

式中:Re 为雷诺数。

对于球冠体来说,结构底部较为平顺,因此流体黏性带来的摩擦阻力主要作用于水平方向,后续的分析中忽略垂直方向的摩擦阻力。

2.3　结构动力响应分析

本文采用状态方程直接积分法计算每个时间步内,弹性结构在流场载荷下的弹性动力响应[34]。结构体在模态向量坐标下的动力平衡方程[35-36]可以表示为

$$\overline{\boldsymbol{M}}\ddot{\boldsymbol{q}} + \overline{\boldsymbol{C}}\dot{\boldsymbol{q}} + \overline{\boldsymbol{K}}\boldsymbol{q} = \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{F}_{\mathrm{str}} \tag{17}$$

式中:$\boldsymbol{q}$ 为结构体模态向量坐标;$\overline{\boldsymbol{M}}$ 为结构体广义质量矩阵;$\overline{\boldsymbol{C}}$ 为结构体广义阻尼矩阵;$\overline{\boldsymbol{K}}$ 为结构体广义刚度矩阵;$\boldsymbol{F}_{\mathrm{str}}$ 为有限元模型节点处所受的外力向量;$\boldsymbol{\xi}$ 为结构体模态矩阵。

令

$$\boldsymbol{x} = \begin{bmatrix}\boldsymbol{q}\\ \dot{\boldsymbol{q}}\end{bmatrix} \tag{18}$$

$$\boldsymbol{y} = \boldsymbol{q} \tag{19}$$

$$\boldsymbol{u} = \boldsymbol{\xi}^{\mathrm{T}}\boldsymbol{F}_{\mathrm{str}} \tag{20}$$

式中:$\boldsymbol{x}$ 为状态向量;$\boldsymbol{y}$ 为输出向量;$\boldsymbol{u}$ 为输入向量。

可以得到结构动力平衡方程的状态空间模型为

$$\begin{cases}\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u}\\ \boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} + \boldsymbol{D}\boldsymbol{u}\end{cases} \tag{21}$$

式中：

$$\begin{cases} \boldsymbol{A} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{I} \\ -\overline{\boldsymbol{M}}^{-1}\overline{\boldsymbol{K}} & -\overline{\boldsymbol{M}}^{-1}\overline{\boldsymbol{C}} \end{bmatrix} \\ \boldsymbol{B} = \begin{bmatrix} \boldsymbol{0} \\ \overline{\boldsymbol{M}}^{-1} \end{bmatrix} \\ \boldsymbol{C} = [\boldsymbol{I} \quad \boldsymbol{0}] \\ \boldsymbol{D} = [\boldsymbol{0}] \end{cases}$$

当有限元模型节点处所受的外力矩阵确定时，就可以通过状态方程直接积分法计算得到三维弹性结构体的位移响应[37]。

2.4 滑跳运动动力学模型

进行滑跳运动的结构体一般为对称结构，因此只需要考虑结构体的纵向平面内运动：水平运动、垂直运动和俯仰运动。通过动量定理和动量矩定理[38]，可以建立滑跳运动的全空间运动方程组。

当结构体在空中飞行时，没有流体载荷的作用，水平方向的运动方程为

$$m_0 a_x = 0 \tag{22}$$

式中：m_0 为结构体总质量；a_x 为水平方向加速度。

垂直方向的运动方程为

$$m_0 a_z = F_{\text{gravity}} + F_{\text{buoyancy}} \tag{23}$$

式中：a_z 为垂直方向加速度；F_{gravity} 为重力；F_{buoyancy} 为浮力。

俯仰方向上的运动方程为

$$J\dot{\omega} = 0 \tag{24}$$

式中：J 为结构体绕质心俯仰旋转的转动惯量；$\dot{\omega}$ 为俯仰角加速度。

当结构体接触水面时，受到高速斜入水冲击载荷的作用，水平方向的运动方程为

$$m_0 a_x = \sum_{i=1}^{b} F_{lxi} + F_{fx} \tag{25}$$

式中：F_{lxi} 为浸湿表面单元高速斜入水冲击载荷的阻力分量；F_{fx} 为摩擦阻力的水平方向分量；b 为浸湿表面单元总数。

垂直方向的运动方程为

$$m_0 a_z = \sum_{i=1}^{b} F_{lzi} + F_{\text{gravity}} + F_{\text{buoyancy}} + F_{fz} \tag{26}$$

式中：F_{lzi} 为浸湿表面单元高速斜入水冲击载荷的升力分量；F_{fz} 为摩擦阻力的垂直方向分量。

俯仰方向上的运动方程为

$$J\dot{\omega} = \sum_{i=1}^{b} \left(F_{lxi} \cdot r_z(P_c, P_i) + F_{lzi} \cdot r_x(P_c, P_i) \right) \tag{27}$$

式中：P_c 为结构体质心位置；P_i 为浸湿表面单元中点位置；$r_x(P_c, P_i)$ 为结构体质心与浸湿表面单元中点 x 坐标的差值；$r_z(P_c, P_i)$ 为结构体质心与浸湿表面单元中点 z 坐标的差值。

2.5 时域迭代求解方法

本文建立的基于边界元法的三维结构体近水面滑跳运动时域迭代仿真方法流程如图7所示。首先，进行分析模型的初始化和分析参数的设置，开始时域迭代计算。在每一个时间步内，对结构体是否触水进行判断。如果触水，由该时刻下的结构体表面和初始流场进行组合构造，确定该时间步的流场形状及流场边界条件，并提交边界元法程序进行求解，得到该时间步内流场对结构体的高速斜入水冲击载荷。然后，将流场载荷插值到结构浸湿表面节点，并利用状态方程直接积分法得到结构节点的弹性位移动力响应。如果结构体没有触水，则该时间步内流场对结构体的高速斜入水冲击载荷为0。最后，根据滑跳运动动力学模型，更新结构体的位置与速度信息，并进入下一步迭代计算。

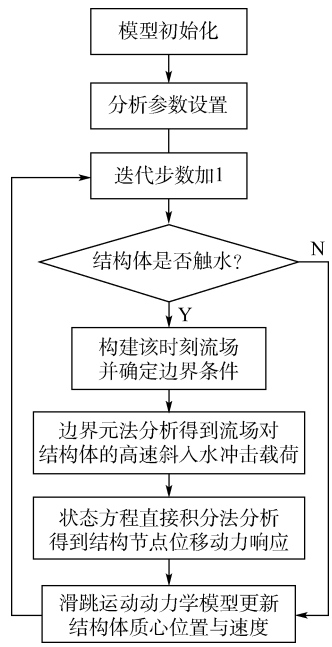


图7 时域迭代求解方法流程

Fig. 7 Flow chart of time domain iterative solution method

3 仿真计算及分析

3.1 数值方法验证

1976年，为了研究高速斜入水冲击载荷的数值，Soliman等[39]完成了钢球高速斜入水冲击试验，得到了一系列相关数据。

在Soliman的试验中，高速斜入水冲击结构为钢制实心球体，直径为25.4 mm，入水速度大小为64.62 m/s。试验使用安装在铸铁平台上的弹药枪，将炮弹射入装满水的有机玻璃水箱(6 ft长、1 ft宽、1 ft深，1 ft(英尺)=0.304 8 m)。相

关参数采用 35 mm 摄像机及频闪仪进行测量。试验的示意图如图 8 所示。图中的钢球以入水速度 $\boldsymbol{V}$、入水倾角 θ 高速冲击水面,c 代表钢球质心。

为了验证本文提出的三维结构体高速斜入水冲击载荷时域迭代求解方法的可行性和准确性,针对 Soliman 的试验进行了数值仿真。钢球具有较大的水平运动速度及较小的垂向运动速度。数值计算结果与试验值的对比如图 9 和图 10所示。其中,钢球离开水面时运动速度的数值计算结果与试验值对比如图 9 所示;钢球离开水面时出水倾角的数值计算结果与试验值对比如图 10 所示。

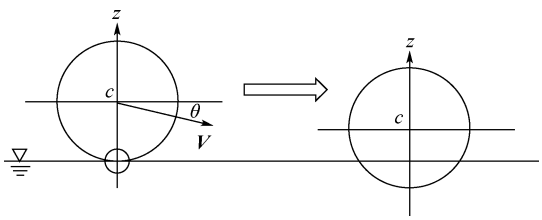


图 8　钢球高速斜入水试验示意图

Fig. 8　Schematic diagram of high speed oblique water entry test of steel sphere

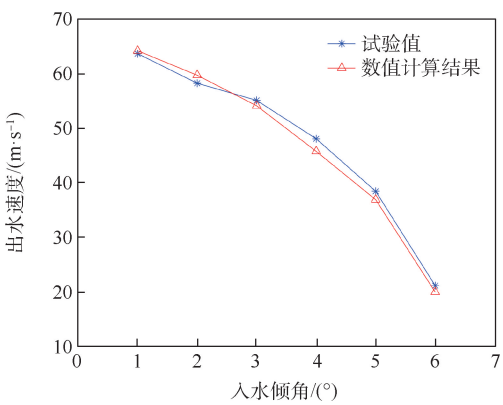


图 9　离开水面速度数值计算结果与试验值对比

Fig. 9　Comparison between calculated and experimental values of velocity leaving water surface

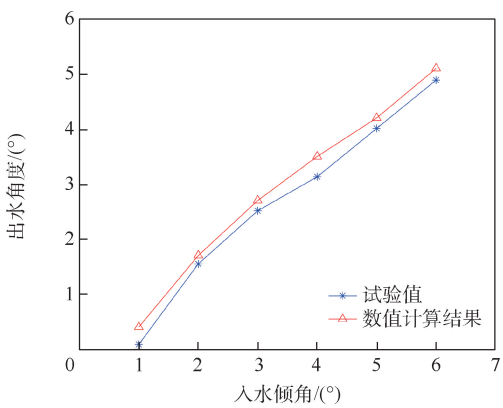


图 10　离开水面倾角数值计算结果与试验值对比

Fig. 10　Comparison between calculated and experimental values of dip angle leaving water surface

通过图 9 和图 10 可知,当钢球入水倾角 θ 增大时,钢球在触水滑跳过程中的动能衰减增大,其离开水面时的速度倾角增大。此外还可以看出,本文方法的数值计算结果与试验数值吻合良好,验证了本文提出的三维结构体高速斜入水冲击载荷时域迭代求解方法的可行性和准确性。

3.2　仿真建模

在本文方法中,每一个时间步都需要用到结构体的浸湿表面和未受扰动的流场来对下一个时间步的流场进行重构,结构动力响应分析也需要将结构体的模态分析结果作为原始数据。因此,需要提前对结构体进行有限元建模和模态分析,建立未受扰动的流场,设置初始参数。

为了方便构成流场边界,需要预先对球冠体进行网格划分,即有限元建模。球冠体半径为0.152 4 m,材料密度为 1 800 kg/m^3,质量为 1.554 kg。在结构有限元建模时,球冠体表面均采用三角形单元进行网格划分,构建完成的结构有限元模型如图 11 所示。材料的弹性模量为70 GPa,剪切模量为 26 GPa,泊松比为 0.33。通过实心球冠体的模态分析可以得到,一阶固有频率高达 3 800 Hz,球冠体的弹性效应很小。因此,实心球冠体可近似地看作一个刚性物体。

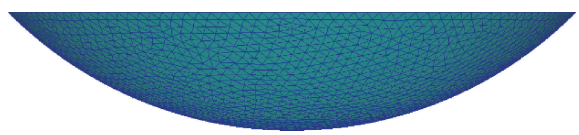


图 11　球冠体有限元模型

Fig. 11　Finite element model of spherical crown

针对直径为 0.304 8 m 的球冠体结构,建立一个边长为 1.0 m 的立方体流场,初始未受扰动的流场如图 12 所示。由于采用边界元法进行高速斜入水冲击载荷分析,流场内部没有网格划分,只是在边界表面上划分了共 10 682 个三角形单元。在滑跳运动计算过程中,流场随结构体质心的水平运动进行平移,流场的坐标原点始终处于质心的正下方。

对于标准算例,球冠体在距离水平面 1.0 m 的高度处,以 25.0 m/s 的速度被水平抛出。时域仿真总时长为 1.4 s,迭代步长为 0.4 ms,液体的密度设置为 1 000 kg/m^3,重力加速度 g 为9.8 m/s^2。从第一个时间步开始,每个时间步的流场域上表面都是由流场的自由表面和结构浸湿表面重构而成。

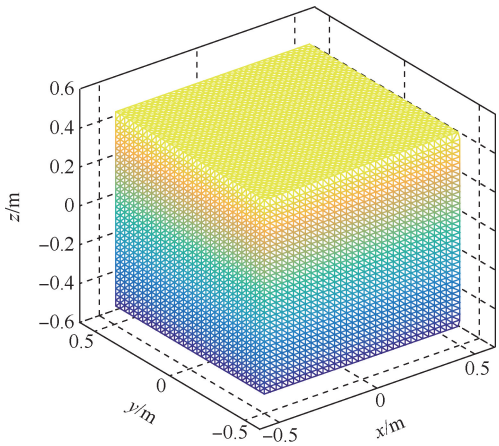


图 12 初始未受扰动的流场

Fig. 12 Initial undisturbed flow field

3.3 刚性球冠体滑跳运动仿真

根据 3.2 节的标准算例参数设置,应用本文建立的数值仿真方法对刚性球冠体进行近水面滑跳运动时域仿真,仿真总时长为 1.4 s,仿真结果如图 13 ~ 图 19 所示。其中,图 13 为球冠体近水面滑跳运动的轨迹仿真结果,图 14 为球冠体近水面滑跳运动的垂向速度时历曲线,图 16 为球冠体近水面滑跳运动的水平速度时历曲线。

根据图 13 可知,本文建立的基于边界元法的三维结构体近水面滑跳运动时域数值仿真方法适用于球冠体近水面滑跳运动的数值仿真。球冠体在 1.4 s 内经历了 3 次触水滑跳运动,每次触水弹起的质心最高位置逐渐降低。此外还可以看出,第 1 次触水滑跳过程时长约 24 ms,入水深度为 0.025 m;第 2 次触水滑跳过程时长约为 28 ms,入水深度为 0.018 m;第 3 次触水滑跳过程时长约 32 ms,入水深度为 0.014 m。可以看出,随着触水滑跳次数的增加,触水滑跳过程的时长增加,但入水深度减小。通过图 14 和图 16 可

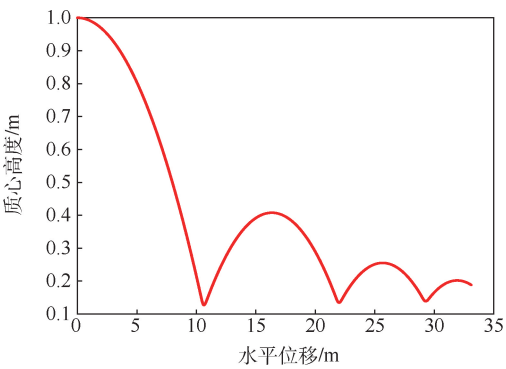


图 13 滑跳运动轨迹仿真结果

Fig. 13 Simulation results of skipping motion trajectory

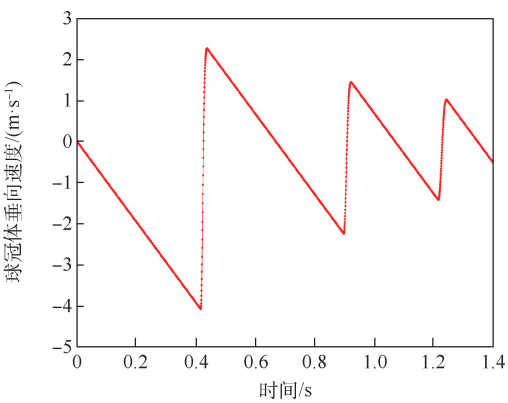


图 14 滑跳运动垂向速度时历曲线

Fig. 14 Vertical velocity time history curve of skipping motion

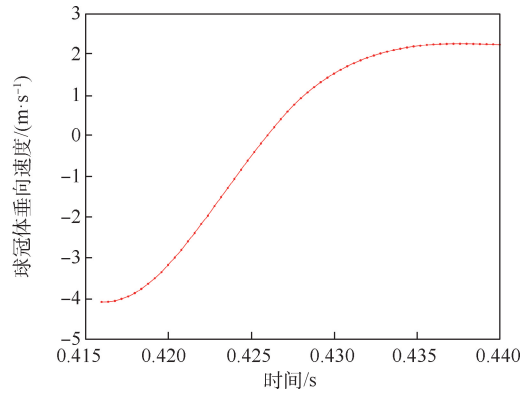


图 15 第 1 次触水垂向速度时历曲线

Fig. 15 Vertical velocity time history curve in the first contact with water

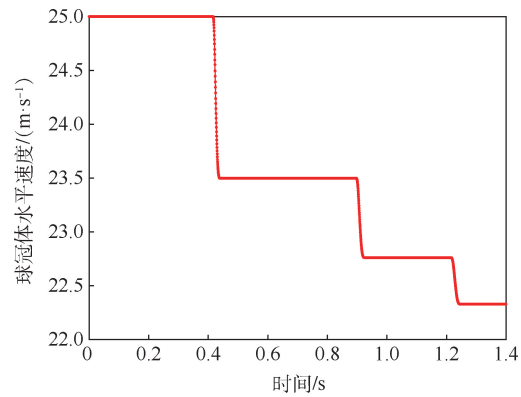


图 16 滑跳运动水平速度时历曲线

Fig. 16 Horizontal velocity time history curve of skipping motion

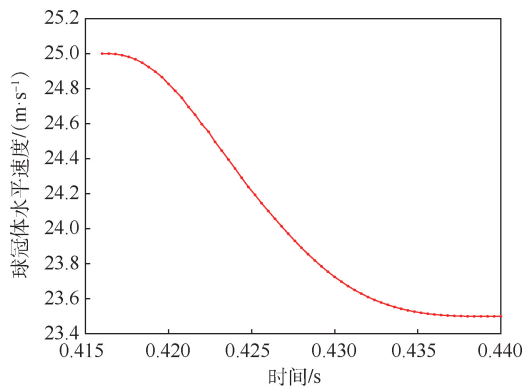


图 17 第 1 次触水水平速度时历曲线

Fig. 17 Horizontal velocity time history curve in the first contact with water

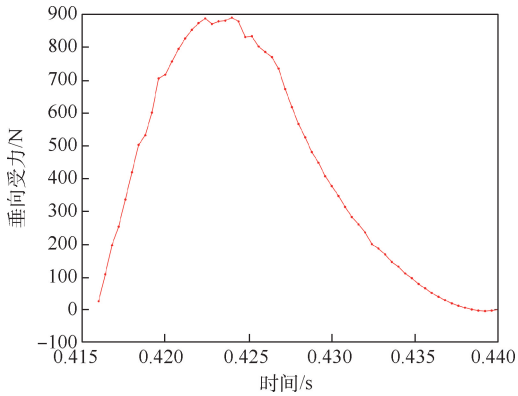


图 18　第 1 次触水垂向受力时历曲线

Fig. 18　Vertical stress time history curve in the first contact with water

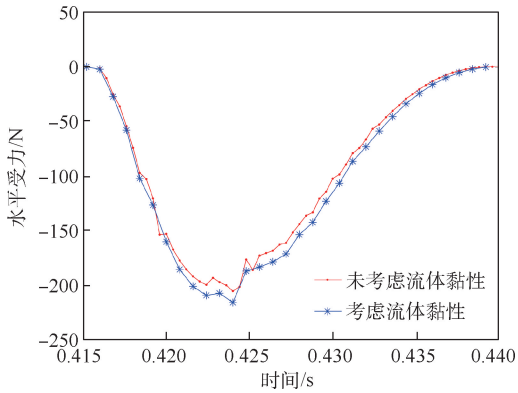


图 19　第 1 次触水水平受力时历曲线

Fig. 19　Horizontal stress time history curve in the first contact with water

知,球冠体在结构未触水的时间段内只受到重力作用,下降速度匀速增大,水平速度不变;当球冠体触水时,受到流场力的作用,在极短的时间内垂向速度由负转正,水平速度减小。此外还可以看出,球冠体每次触水弹起后,垂向速度的最大值都会发生衰减。垂向速度最大值和水平速度的减小,代表了球冠体动能在不断的触水滑跳过程中被消耗。滑跳运动过程中结构体的水平速度不断下降,当触水弹起高度不足以使结构体脱离水面之后,滑跳运动结束。

由于球冠体每次触水滑跳时间过短,体现在图 14 和图 16 中,触水滑跳的垂向速度变化和水平速度变化近似折线形式。而事实上,触水弹跳的过程中速度变化是连续的。球冠体第 1 次触水滑跳过程中的垂向速度变化曲线和水平速度变化曲线分别如图 15 和图 17 所示。通过图 15 可知,第 1 次触水滑跳时间历程为 0.416 ~ 0.44 s,时长为 24 ms。当球冠体刚接触水面时,垂向速度约为 -4.0 m/s,受到流场力的作用,垂向速度迅速变化,约为 0.426 s 时由负转正,并继续增加至 2.2 m/s离开水面。通过图 17 可知,水平速度由刚接触水面时的 25 m/s 持续减小至 23.5 m/s 离开水面,衰减率为 6%(衰减率 $\lambda = (V_{in} - V_{out})/V_{in}$,其中 V_{in}为球冠体第 1 次触水之前的水平速度,V_{out}为球冠体第 1 次触水之后的水平速度,下同)。

球冠体第 1 次触水过程中的高速斜入水冲击载荷仿真结果如图 18 和图 19 所示。其中,图 18 为球冠体在第 1 次触水过程中的高速斜入水冲击载荷垂直方向分量时历图,图 19 为球冠体在第 1 次触水过程中的高速斜入水冲击载荷水平方向分量时历图。

通过图 18 和图 19 可以看出,在球冠体的第 1 次触水过程中,高速斜入水冲击载荷提供了阻力和升力,都是很快达到峰值并逐渐减小。分开来看,垂直方向受到的载荷方向向上,起到升力的作用,数值相对较大,在 0.422 s 时达到峰值 887.2 N,在其作用下,球冠体下降速度迅速减为 0 并加速上升,直至离开水面;水平方向持续受到阻力的作用,数值相对较小,0.422 s 时达到峰值 -216.0 N,在其作用下水平运动速度不断减小。

对于球冠体而言,直径为 0.304 8 m,第 1 次触水滑跳时运动速度为 25 m/s,流体密度为 1 000 kg/m^3。经过计算,此时的雷诺数高达 7.5×10^6。由图 19 可知,在球冠体第 1 次触水滑跳的过程中,考虑流体黏性的高速斜入水冲击载荷水平方向分量峰值为 -216.0 N,未考虑流体黏性的高速斜入水冲击载荷水平方向分量峰值为 -205.5 N,影响约为 5.1%。由此可见,在高速斜入水冲击仿真过程中,流体黏性带来的影响较小,但是也不可忽略。本文后续的仿真中,都已考虑流体黏性带来的摩擦阻力影响。

3.4　变参分析

3.4.1　结构质量对滑跳运动的影响

为了得到结构质量对球冠体滑跳运动的影响规律,在不改变其他分析参数的情况下,改变球冠体的密度,使得结构质量 m_0 分别为 1.0,1.554,3.0 kg,仿真总时长为 1.3 s。不同结构质量的球冠体滑跳轨迹对比如图 20 所示,不同结构质量的滑跳运动垂向速度时历对比如图 21 所示,不同结构质量的滑跳运动水平速度时历对比如图 22 所示。

由图 20 和图 21 可知,当结构质量增大时,结构体每次触水滑跳入水深度增加,弹起速度和高度也会增大,导致每次滑跳落点之间的水平距离加大。在相同的时间和距离内,随着结构质量的增加,触水滑跳的次数减少。由图 22 可知,质量为 1.0 kg 时,第 1 次触水滑跳后水平速度降为

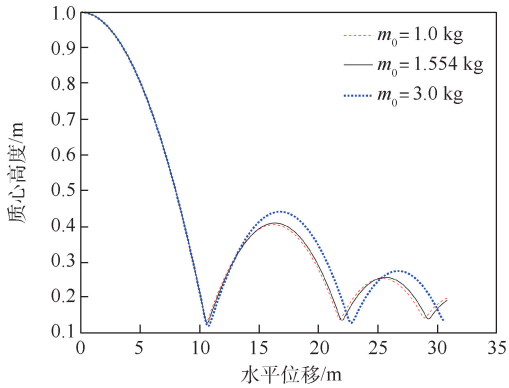


图20 不同结构质量的滑跳轨迹对比

Fig. 20 Comparison of skipping motion trajectory with different structural quality

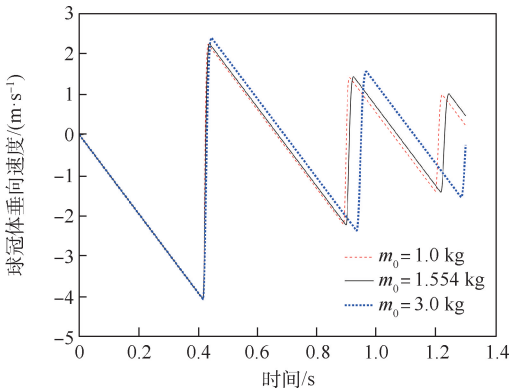


图21 不同结构质量的滑跳垂向速度时历对比

Fig. 21 Time history comparison of skipping motion vertical velocity with different structural quality

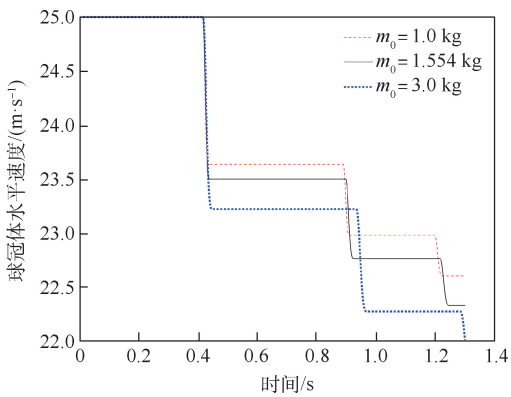


图22 不同结构质量的滑跳水平速度时历对比

Fig. 22 Time history comparison of skipping motion horizontal velocity with different structural quality

23.64 m/s,衰减率为5.44%;质量为1.554 kg时,第1次触水滑跳后水平速度降为23.5 m/s,衰减率为6%;质量为3.0 kg时,第1次触水滑跳后水平速度降为23.22 m/s,衰减率为7.12%。当结构质量增大时,每次触水弹跳水平速度的下降量增大。

3.4.2 初始高度对滑跳运动的影响

为了得到初始高度对球冠体滑跳运动的影响规律,在不改变其他分析参数的情况下,改变初始高度H_0分别为0.5,1.0,1.5 m,仿真总时长为1.3 s。不同初始高度的滑跳轨迹对比如图23所示,不同初始高度的滑跳垂向速度时历对比如图24所示,不同初始高度的滑跳水平速度时历对比如图25所示。

通过图23和图24可知,初始抛出高度增加,结构体每次触水滑跳入水深度增加,触水弹起的高度和弹起速度也随之增大,导致每次滑跳落点

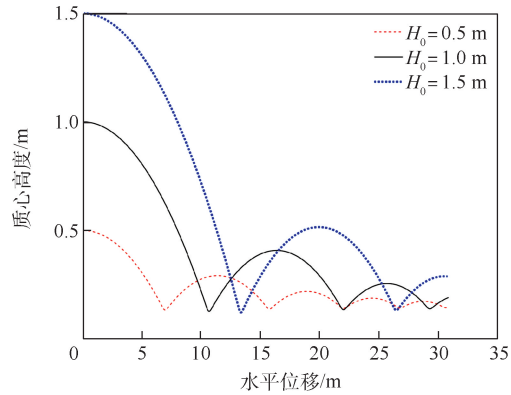


图23 不同初始高度的滑跳轨迹对比

Fig. 23 Comparison of skipping motion trajectories with different initial heights

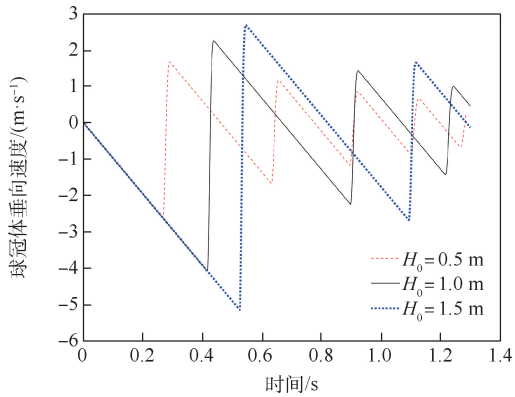


图24 不同初始高度的滑跳垂向速度时历对比

Fig. 24 Time history comparison of skipping motion vertical velocity with different initial heights

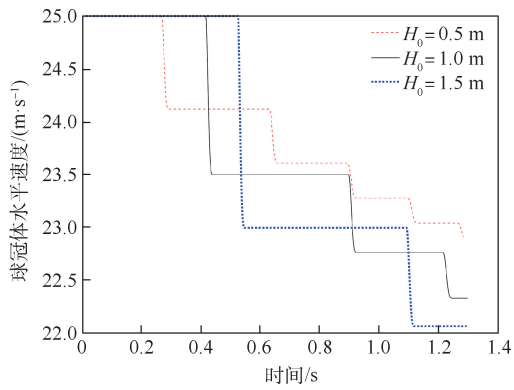


图25 不同初始高度的滑跳水平速度时历对比

Fig. 25 Time history comparison of skipping motion horizontal velocity with different initial heights

之间的水平距离加大,在同样的时间和距离内触水滑跳的次数变少。通过图 25 可知,初始高度为 0.5 m 时,第 1 次触水滑跳后水平速度降为 24.12 m/s,衰减率为 3.52%;初始高度为 1.0 m 时,第 1 次触水滑跳后水平速度降为 23.5 m/s,衰减率为 6%;初始高度为 1.5 m 时,第 1 次触水滑跳后水平速度降为 23.0 m/s,衰减率为 8%。初始抛出高度的增大,会导致每次触水弹跳水平速度的下降量增加,动能损耗增多。

3.4.3　水平抛出速度对滑跳运动的影响

为了得到水平抛出速度对球冠体滑跳运动的影响规律,在不改变其他分析参数的情况下,改变球冠体的水平抛出速度 V_0 分别为 20,25,30 m/s,仿真总时长为 1.3 s。不同水平抛出速度的滑跳轨迹对比如图 26 所示,不同水平抛出速度的滑跳运动垂向速度时历对比如图 27 所示。

通过图 26 和图 27 可以看出,随着水平抛出速度的增大,相同时间内球冠体的水平位移增加,每次触水滑跳的弹起高度和弹起速度也会增加,相同时间和距离内触水滑跳的次数减少。

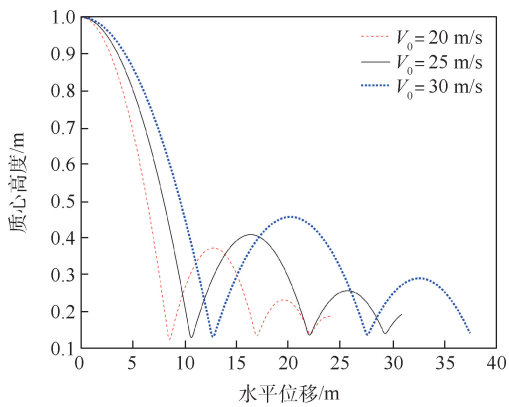


图 26　不同水平抛出速度的滑跳轨迹对比

Fig. 26　Comparison of skipping motion trajectories with different horizontal throw speeds

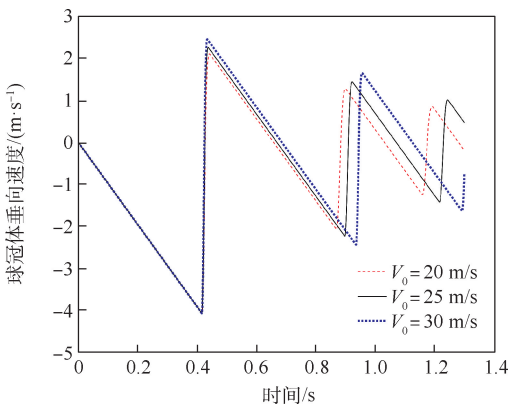


图 27　不同水平抛出速度的滑跳垂向速度时历对比

Fig. 27　Time history comparison of skipping motion vertical velocity with different horizontal throw speeds

3.4.4　半径大小对滑跳运动的影响

为了得到半径大小对球冠体滑跳运动的影响规律,在不改变其他分析参数的情况下,改变球冠体的半径大小 r_0 分别为 0.1,0.152 4,0.2 m,仿真总时长为 1.3 s。不同半径大小的滑跳轨迹对比如图 28 所示,不同半径大小的滑跳运动垂向速度时历对比如图 29 所示,不同半径大小的滑跳运动水平速度时历对比如图 30 所示。

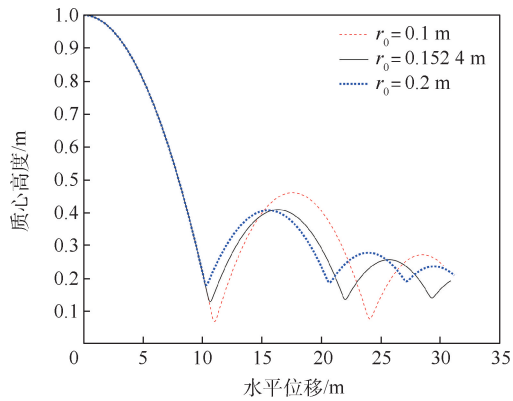


图 28　不同半径大小的滑跳轨迹对比

Fig. 28　Comparison of skipping motion trajectories with different radius sizes

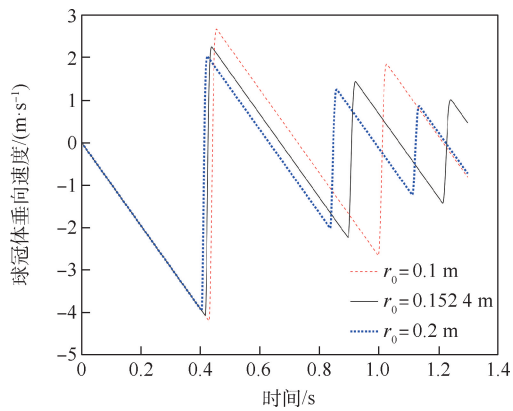


图 29　不同半径大小的滑跳垂向速度时历对比

Fig. 29　Time history comparison of skipping motion vertical velocity with different radius sizes

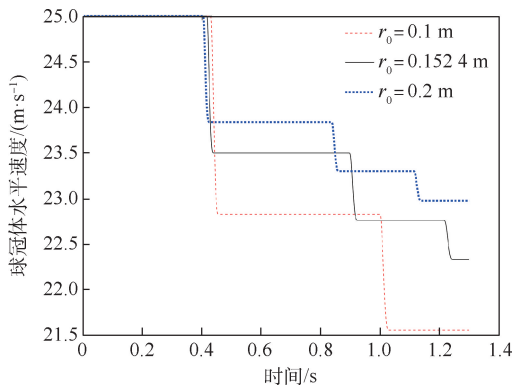


图 30　不同半径大小的滑跳水平速度时历对比

Fig. 30　Time history comparison of skipping motion horizontal velocity with different radius sizes

通过图 28 和图 29 所示,随着球冠体的半径减小,结构每次触水弹跳的入水深度增加,每次触水滑跳的高度和弹起速度也随之增加,导致每次触水滑跳运动之间飞行的距离增大,在同样的时间或者距离内触水弹起的次数减少。此外,通过图 30 可知,球冠体半径为 0.1 m 时,第 1 次触水滑跳后水平速度降为 22.83 m/s,衰减率为 8.68%;球冠体半径为 0.152 4 m 时,第 1 次触水滑跳后水平速度降为 23.5 m/s,衰减率为 6%;球冠体半径为 0.2 m 时,第 1 次触水滑跳后水平速度降为 23.84 m/s,衰减率为 4.64%,随着球冠体半径的减小,结构每次触水弹跳水平速度的下降量增大,动能损耗增多。

3.5 弹性效应的影响

为了得到结构体弹性效应对球冠体滑跳运动的影响规律,在不改变其他分析参数的情况下,考虑结构体的弹性效应并进行时域仿真分析。在考虑结构弹性效应的影响时,发现实心高刚度材料的球冠体几乎没有弹性效应。为了放大结构的弹性效应,将实心球冠体改为空心,壳体的厚度为 1.0 mm,材料刚度同时降低,弹性模量改为 0.7 MPa,泊松比为 0.33。进行模态分析后得到,空心球冠体的一阶弹性模态频率为 12.32 Hz,二阶弹性模态频率为 14.91 Hz,仿真总时长为 1.1 s。有无弹性效应时结构体第 1 次触水所受垂向流场载荷时历对比如图 31 所示,有无弹性效应时结构体第 1 次触水所受水平方向流场载荷时历对比如图 32所示,有无弹性效应的滑跳轨迹对比如图 33 所示,有无弹性效应的滑跳运动垂向速度时历对比如图 34 所示,有无弹性效应的滑跳运动水平速度时历对比如图 35 所示。考虑弹性效应时,球冠体最低点的弹性变形时历图如图 36 所示。

通过图 31 和图 32 可知,考虑结构体的弹性效应之后,触水滑跳过程中流体作用在结构上的垂向载荷和水平方向载荷都有所减小,这也影响了结构体的滑跳运动。垂向受力峰值由 887.2 N 降为 792.7 N;水平方向受力峰值由 −205.5 N 变为 −165.3 N。通过图 33 和图 34 可知,考虑弹性效应之后,球冠体每次触水滑跳的弹起高度降低,弹起速度也降低。同时,每次触水弹起飞行的距离也会降低,在同样的时间或者距离内触水弹起的次数会增多。通过图 35 可以知道,考虑弹性效

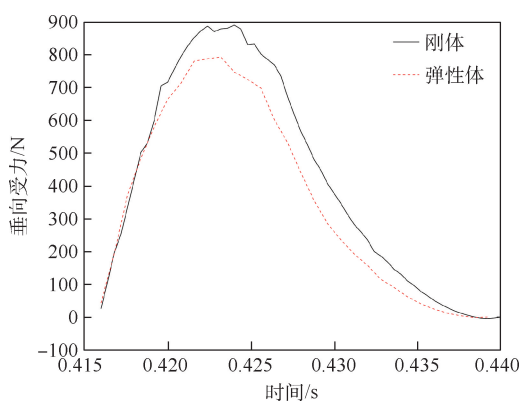


图 31 有无弹性效应的第 1 次触水垂向受力时历对比

Fig. 31 Time history comparison of vertical stress in the first contact with water with or without elastic effect

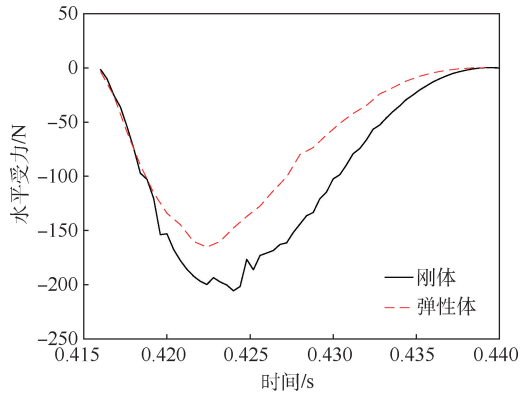


图 32 有无弹性效应的第 1 次触水水平受力时历对比

Fig. 32 Time history comparison of horizontal stress in the first contact with water with or without elastic effect

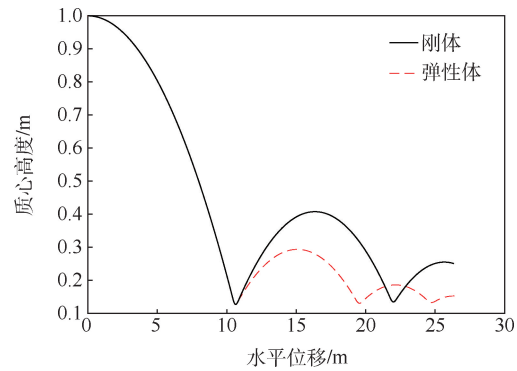


图 33 有无弹性效应的滑跳轨迹对比

Fig. 33 Comparison of skipping motion trajectories with or without elastic effect

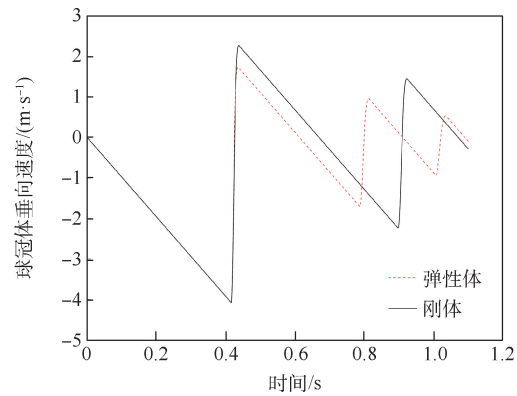


图 34 有无弹性效应的滑跳垂向速度时历对比

Fig. 34 Time history comparison of skipping motion vertical velocity with or without elastic effect

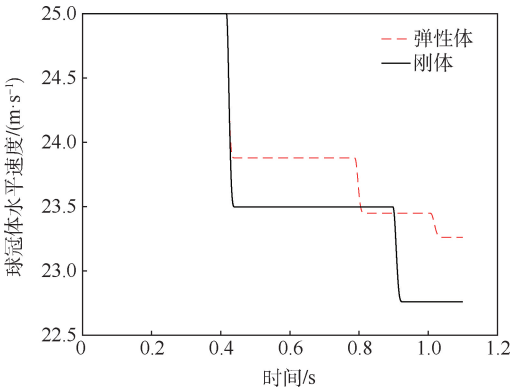


图 35　有无弹性效应的滑跳水平速度时历对比

Fig. 35　Time history comparison of skipping motion horizontal velocity with or without elastic effect

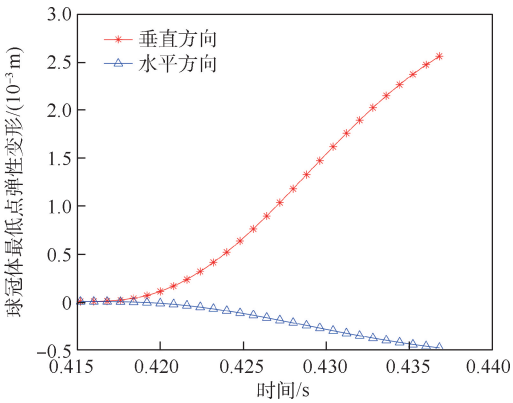


图 36　球冠体最低点的弹性变形时历曲线

Fig. 36　Time history curves of elastic deformation at the lowest point of spherical crown

应之后,第 1 次触水滑跳后水平速度降为 23.88 m/s,衰减率为 4.48%。这说明考虑弹性效应后,球冠体每次触水滑跳水平速度衰减率降低。

通过图 36 可知,球冠体最低点的垂向弹性形变为一直向上压缩,水平方向弹性形变为向后变形。这与高速斜入水冲击载荷的作用方向一致。此外,垂直方向的弹性形变大于水平方向的弹性形变。

4　结　论

1) 本文基于边界元法,建立了一种三维结构体近水面滑跳运动的时域数值仿真方法。该方法既可以应用于任意外形的三维结构体,也拥有较高的计算效率。针对三维刚性球冠体的近水面滑跳运动进行了时域数值仿真,得到了对应的运动特性,验证了该方法的可行性和准确性。在此基础上,可以进一步应用于击水式飞行器的滑跳弹道仿真分析。

2) 在球冠体的每次触水过程中,高速斜入水冲击载荷提供的阻力和升力,均快速达到峰值并逐渐减小,且垂向受力相较于水平方向受力会先一步达到峰值。垂向受力的数值相对较大,水平方向的受力数值相对较小。

3) 初始参数对球冠体滑跳运动有较大的影响。随着结构质量的增加,每次触水滑跳的弹起速度和高度增大,每次滑跳落点之间的水平距离加大,在相同的时间和距离内触水滑跳的次数减少,每次触水弹跳水平速度的衰减量增大。初始高度的增加,会导致每次触水滑跳的弹起高度和弹起速度增大,每次滑跳落点之间的水平飞行距离加大,在同样的时间和距离内触水滑跳的次数变少,每次触水弹跳水平速度的衰减量增大。随着水平抛出速度的增大,每次触水滑跳的弹起速度和高度增大,每次滑跳落点之间的水平距离加大,在相同的时间和距离内触水滑跳的次数减少。球冠体半径的减小,会导致每次触水滑跳的弹起高度和弹起速度增加,每次触水弹跳运动之间飞行的距离增大,在同样的时间或者距离内触水弹起的次数减少,每次触水弹跳水平速度的衰减量增大。

4) 弹性效应对球冠体滑跳运动有一定的影响,考虑弹性效应后,会减小高速斜入水冲击载荷,导致球冠体弹起速度和高度降低,且滑跳过程中水平速度衰减率减小。因此,对弹性效应较明显的结构体进行滑跳运动数值仿真时,应当考虑弹性效应的影响。

参考文献 (References)

[1] 孙士明,陈玮琪,王宝寿,等. 航行体近水面滑跳运动试验研究[J]. 数字海洋与水下攻防,2019,2(2):79-87.
SUN S M, CHEN W Q, WANG B S, et al. Experimental research on near-water-surface skipping motion of vehicle[J]. Digital Ocean & Underwater Warfare, 2019, 2(2): 79-87 (in Chinese).

[2] 孙士明,陈玮琪,王宝寿. 跨介质航行体研究现状及关键技术[J]. 水中兵器,2016 (4):1-7.
SUN S M, CHEN W Q, WANG B S. Research status and key technologies of cross-media navigation vehicles[J]. Underwater Weapon, 2016(4): 1-7 (in Chinese).

[3] 裴譞,张宇文,李闻白,等. 跨介质飞行器气/水两相弹道仿真研究[J]. 工程力学,2010,27(8):223-228.
PEI X, ZHANG Y W, LI W B, et al. Simulation and analysis on the gas/water two-phase ballistics of trans-media aircraft[J]. Engineering Mechanics, 2010, 27(8): 223-228 (in Chinese).

[4] 王伟,张宇文,朱灼. 跨介质飞行器弹道仿真分析[J]. 计算机仿真,2011,28(12):1-4.
WANG W, ZHANG Y W, ZHU Z. Simulation and analysis on ballistic trajectory of trans-media aircraft[J]. Computer Simula-

tion,2011,28(12):1-4(in Chinese).

[5] 李金洪,杨安强,粟凌云.跨介质UAV水面滑跳转向特性建模与仿真[J].鱼雷技术,2012,20(6):401-406.
LI J H,YANG A Q,SU L Y. Modeling and simulation of dynamic characteristics of slide jump steering trajectory for trans-media UAV[J]. Torpedo Technology,2012,20(6):401-406(in Chinese).

[6] ZHAO R,FALTINSEN O. Water entry of two-dimensional bodies[J]. Journal of Fluid Mechanics,1993,246(1):593-612.

[7] 樊征,刘莹莹.掠海飞行器主动触水撞水力分析[J].固体火箭技术,2012,35(6):722-725.
FAN Z,LIU Y Y. Analysis on slamming force of active water impact for sea skimming vehicle[J]. Journal of Solid Rocket Technology,2012,35(6):722-725(in Chinese).

[8] XU G D,DUAN W Y,WU G X. Numerical simulation of oblique water entry of an asymmetrical wedge[J]. Ocean Engineering,2008,35(16):1597-1603.

[9] FALTINSEN O M,CHEZHIAN M. A generalized Wagner method for three-dimensional slamming[J]. Journal of Ship Research,2005,49(4):279-287.

[10] 顾懋祥,程贯一,张效慈.平头旋转壳撞水水弹性效应的研究[J].水动力学研究与进展:A辑,1991,6(1):42-51.
GU M X,CHENG G Y,ZHANG X C. Study on hydroelastic effect of a blunt rotatory shell impacting with water[J]. Journal of Hydrodynamics:Ser. A,1991,6(1):42-51(in Chinese).

[11] TAKAGI K. Influence of elasticity on hydrodynamic impact problem[J]. Journal of the Kansai Society of Naval Architects,Japan,1994,222:97-106.

[12] 张岳青,徐绯,李建辰,等.结构物入水冲击表面压力的模型研究及应用[J].机械强度,2017,39(1):33-39.
ZHANG Y Q,XU F,LI J C,et al. Model study and application on surface pressure measurement of structure water entry[J]. Journal of Mechanical Strength,2017,39(1):33-39(in Chinese).

[13] 张岳青,白治宁,曾小凡,等.楔形和弧形结构入水冲击响应研究[J].船舶力学,2020,24(3):400-408.
ZHANG Y Q,BAI Z N,ZENG X F,et al. Study of water impact response of wedge-and-arc-shaped structures[J]. Journal of Ship Mechanics,2020,24(3):400-408(in Chinese).

[14] MOHAMMAD J,MAURIZIO P. Water entry of compliant slender bodies:Theory and experiments[J]. International Journal of Mechanical Sciences,2018,149:514-529.

[15] WU G X,SUN H,HE Y S. Numerical simulation and experimental study of water entry of a wedge in free fall motion[J]. Journal of Fluids and Structures,2004,19(3):277-289.

[16] WU G X. Numerical simulation of water entry of twin wedges[J]. Journal of Fluids and Structures,2006,22(1):99-108.

[17] BATHE K J. 有限元法[M]. 2版. 轩建平,译. 北京:高等教育出版社,2016.
BATHE K J. Finite element procedures[M]. 2nd ed. XUAN J P,translated. Beijing:Higher Education Press,2016(in Chinese).

[18] 姚振汉,王海涛. 边界元法[M]. 北京:高等教育出版社,2010.
YAO Z H,WANG H T. Boundary element methods[M]. Beijing:Higher Education Press,2010(in Chinese).

[19] 王永虎,石秀华,李文哲,等.斜入水高速冲击的理论建模及缓冲分析[J].机械科学与技术,2008,27(6):766-769.
WANG Y H,SHI X H,LI W Z,et al. Modeling and cushioning analysis of oblique water entry with high velocity[J]. Mechanical Science and Technology for Aerospace Engineering,2008,27(6):766-769(in Chinese).

[20] 陈占晖,卢永锦.高速运动物体砰击水面后的动力学特性研究[J].船舶工程,2009,31(3):62-65.
CHEN Z H,LU Y J. Study on the dynamic characteristics of high-speed object slamming water[J]. Ship Engineering,2009,31(3):62-65(in Chinese).

[21] 郑金伟,宗智.三维刚体椭圆头结构高速倾斜入水冲击模拟[J].船海工程,2012,41(3):7-9.
ZHENG J W,ZONG Z. 3-dimensional numerical simulation of rigid elliptic structure inclined water-entry at high speed[J]. Ship & Ocean Engineering,2012,41(3):7-9(in Chinese).

[22] SCOLAN Y. Oblique water entry of a three dimensional body[J]. International Journal of Naval Architecture and Ocean Engineering,2014,6(4):1197-1208.

[23] TSAI H W,WU W F,LAI C L. An experimental study of non-spinning stone-skipping process[J]. Experimental Thermal and Fluid Science,2021,123:1-10.

[24] LI C H,WANG C,WEI Y J,et al. Numerical investigation on the cavity dynamics and deviation characteristics of skipping stones[J]. Journal of Fluids and Structures,2021,104:1-16.

[25] 卢炽华,何友声.二维弹性结构入水冲击过程中的流固耦合效应[J].力学学报,2000,32(2):129-140.
LU C H,HE Y S. Coupled analysis of nonlinear interaction between fluid and structure during impact[J]. Acta Mechanica Sinica,2000,32(2):129-140(in Chinese).

[26] 李井煜,卢晓平,赵鹏伟.直接边界元法解势流速度场问题[J].中国舰船研究,2015,10(1):68-75.
LI J Y,LU X P,ZHAO P W. Direct boundary element method for the problem of potential flow velocity field[J]. Chinese Journal of Ship Research,2015,10(1):68-75(in Chinese).

[27] 王献孚.计算船舶流体力学[M].上海:上海交通大学出版社,1992.
WANG X F. Computational ship hydrodynamics[M]. Shanghai:Shanghai Jiaotong University Press,1992(in Chinese).

[28] 闫发锁,于龙超,赵九龙,等.轴对称体砰击载荷的三维动网格数值计算[J].中国海洋大学学报,2017,47(5):113-120.
YAN F S,YU L C,ZHAO J L,et al. Numerical calculation of the slamming load for axisymmetric geometries using 3D dynamic mesh[J]. Periodical of Ocean University of China,2017,47(5):113-120(in Chinese).

[29] WAGNER V H. Phenomena associated with impacts and sliding on liquid surfaces[J]. Journal of Applied Mathematics and Mechanics,1932,12:193-215.

[30] WARDLAW A B,MORRISON A M,BALDWIN J L. Prediction of impact pressures,forces,and moments during vertical and oblique water entry[R]. [S. l.]:Naval Surface Weapons Center White Oak Lab Silver Spring Md,1977.

[31] JOHNSON W,REID S R. Ricochet of spheres off water[J].

Journal of Mechanical Engineering Science, 1975, 17(2): 71-81.

[32] HUTCHINGS I M. The ricochet of spheres and cylinders from the surface of water[J]. International Journal of Mechanical Sciences, 1976, 18(5): 243-247.

[33] 是长春. 平板湍流摩阻系数的显式表达式[J]. 航空学报, 1986, 9(6): B299-B302.
SHI C C. An explicit expression of the turbulent frictional drag coefficient for a flat plate[J]. Acta Aeronautica et Astronautica Sinica, 1986, 9(6): B299-B302(in Chinese).

[34] 张森文, 曹开彬. 计算结构动力响应的状态方程直接积分法[J]. 计算力学学报, 2000, 17(1): 94-118.
ZHANG S W, CAO K B. Direct integration of state equation method for dynamic response of structure[J]. Chinese Journal of Computational Mechanics, 2000, 17(1): 94-118(in Chinese).

[35] 方同, 薛璞. 振动理论及应用[M]. 西安: 西北工业大学出版社, 1998.
FANG T, XUE P. Vibration theory and application[M]. Xi'an: Northwest University of Technology Press, 1998(in Chinese).

[36] 诸德超, 邢誉峰, 程伟, 等. 工程振动基础[M]. 北京: 北京航空航天大学出版社, 2004.
ZHU D C, XING Y F, CHENG W, et al. Engineering vibration foundation[M]. Beijing: Beihang University Press, 2004(in Chinese).

[37] 杨超, 吴志刚, 万志强, 等. 飞行器气动弹性原理[M]. 北京: 北京航空航天大学出版社, 2011.
YANG C, WU Z G, WAN Z Q, et al. Principle of aircraft aeroelasticity[M]. Beijing: Beihang University Press, 2011(in Chinese).

[38] 钱杏芳, 林瑞雄, 赵亚男. 导弹飞行力学[M]. 北京: 北京理工大学出版社, 1987.
QIAN X F, LIN R X, ZHAO Y N. Missile flight mechanics[M]. Beijing: Beijing Institute of Technology Press, 1987(in Chinese).

[39] SOLIMAN A S, REID S R, JOHNSON W. The effect of spherical projectile speed in ricochet off water and sand[J]. International Journal of Mechanical Sciences, 1976, 18(6): 279-284.

Numerical simulation of skipping motion of three-dimensional structure based on boundary element method

YANG Chao, JIANG Yu, WU Zhigang*

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

Abstract: In recent years, the water striking aircraft which realizes high-speed maneuvering flight near the water surface by using skipping motion has become one of the hot issues in research, and the solution of high-speed oblique water impact load of three-dimensional structure is the key step of skipping motion numerical simulation. Based on the boundary element method, a time-domain numerical simulation method of water surface skipping of three-dimensional structure considering structural elastic effect is established. The high-speed oblique water impact load is obtained by the boundary element method, and the displacement dynamic response of the structural node is obtained by the direct integration method of the state equation. Moreover, the centroid position and motion speed of the structure are updated by the skipping motion dynamic model. To verify the accuracy of the fluid load solution algorithm of this method, the numerical simulation of high-speed oblique water impact of steel sphere is carried out and compared with the experimental value. The time-domain numerical simulation and variable parameter analysis of skipping motion of three-dimensional rigid spherical crown are carried out. Then, the effects of structural mass, initial height, horizontal throw speed and radius on the skipping motion of the spherical crown are obtained. After considering the elastic effect, the impact load of high-speed oblique water entry decreases, which also has a certain influence on the sliding and jumping motion of the spherical crown.

Keywords: skipping motion; boundary element method; water inflow impact; fluid solid coupling; spherical crown

Received: 2022-03-11; **Accepted**: 2022-05-27; **Published online**: 2022-06-02 11:26

URL: kns.cnki.net/kcms/detail/11.2625.V.20220601.1111.001.html

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0288

时间积分方法的研究进展与挑战

邢誉峰[1,*], 季奕[2], 张慧敏[3]

(1. 北京航空航天大学 固体力学研究所, 北京 100083;
2. 北京理工大学 宇航学院, 北京 100081; 3. 北京宇航系统工程研究所, 北京 100076)

摘 要: 时间积分方法是动力学常微分方程的一种有效数值求解工具,广泛应用于航空航天、土木工程、机械制造等工程领域的动力学分析。介绍了近几十年来时间积分方法的研究进展。回顾了该领域的经典工作,包括级数展开法、Runge-Kutta 法和 Newmark 方法;介绍了为解决经典时间积分方法在精度、效率、耗散和稳定性方面的不足而发展出的先进时间积分方法,主要包括参数方法、高阶无条件稳定方法、保能量方法、线性多步法、复合方法和 BN 稳定型方法;分析比较了已有方法的性能特点和适用范围,指出了时间积分方法发展中值得关注的若干问题。

关 键 词: 动力学系统; 时间积分方法; 精度; 效率; 耗散; 稳定性
中图分类号: O302
文献标志码: A **文章编号**: 1001-5965(2022)09-1692-10

时间积分方法[1](time integration method, TIM)是求解动力学常微分方程最有力的数值方法之一,也是目前计算机辅助工程(computer aided engineering, CAE)软件中的主流时域求解器。时间积分方法的思想[1]是:将待求时间域离散成为一系列时间区间或时间单元,并建立每个时间区间内位移、速度和加速度随时间的变化规律,进而由前一时刻的已知状态变量递推计算当前时刻的未知状态变量。由此可见,时间积分方法有 2 个特点[1]:①动力学平衡方程仅在离散时间点或广义时间点上得到满足,而不是在任意时刻都成立;②需要构造状态变量在时间区间或时间步内的变化规律。这些变化规律决定一个方法的精度、耗散、效率和稳定性。这与空间有限元方法类似,也需要设计单元的容许位移函数。也就是位移随着空间坐标的变化规律,并且代数平衡方程也仅在结点上得到满足。有限单元的容许位移函数决定了精度、效率和适用性。

级数展开法、Runge-Kutta 法和 Newmark 方法是经典的时间积分方法。基于级数展开思想直接构造的格式皆属于条件稳定的显式时间积分方法,代表性工作有 Taylor 级数法和 Lie 级数法。Taylor 级数法直接对时间坐标求导,而 Lie 级数法则对状态变量求导。求导方式的不同使得 Lie 级数法可以解决 Taylor 级数法中高阶求导困难和计算量大的问题。Runge-Kutta 法是基于 Taylor 级数展开思想于 19 世纪末、20 世纪初提出的高精度方法。Runge-Kutta 法[2]的求解对象是形如 $\mathrm{d}\boldsymbol{z}/\mathrm{d}t=\boldsymbol{f}(\boldsymbol{z},t)$ 的一阶常微分方程(ordinary differential equation, ODE),其利用函数 $\boldsymbol{f}$ 在若干时间离散点上信息的线性组合来表示 $\boldsymbol{f}$ 的导数,再利用 Taylor 级数法确定其中的系数。与级数展开法和 Runge-Kutta 法不同,虽然 Newmark 方法设计思想也是基于级数展开,但其求解对象为二阶 ODE。

收稿日期: 2022-04-28; **录用日期**: 2022-05-18; **网络出版时间**: 2022-06-07 10:36
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220606.1607.001.html
基金项目: 国家自然科学基金 (12172023,11872090)
* **通信作者**. E-mail: xingyf@buaa.edu.cn

引用格式: 邢誉峰, 季奕, 张慧敏. 时间积分方法的研究进展与挑战[J]. 北京航空航天大学学报, 2022, 48(9): 1692-1701.
XING Y F, JI Y, ZHANG H M. Advances and challenges in time integration methods[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1692-1701(in Chinese).

Newmark 方法包括许多著名的辛格式,如梯形法则(trapezoidal rule, TR)、中心差分方法(central difference method, CDM)、Fox-Goodwin 方法等。

随着科学技术的发展,人们对时间积分方法的精度、效率、耗散和稳定性的要求越来越高,经典方法已经难以满足需求,这推动了时间积分方法在过去几十年中的持续发展。激发人们设计高性能时间积分方法的主要问题可以概括如下:

1) 具有数值耗散的 Newmark 方法[1]只有一阶精度。

2) 收敛阶数大于 2 时,线性多步法只能是条件稳定或不稳定[3]。

3) 对线性系统具有无条件稳定性的二阶 Newmark 辛几何方法,如 TR,对于简单的非线性问题可能发散[4-5]。

4) 如何同时提高时间积分方法的精度、效率、耗散性能和稳定性。

5) 对非线性动力系统而言,时间积分方法的失稳机理,二阶非线性系统的无条件稳定性方法的设计原则。

针对问题 1,通过对动力学平衡方程进行加权处理[6-10],或在差分格式中引入辅助状态变量[11-12],学者们发展了参数类耗散方法,该类耗散方法具有二阶精度。为了解决问题 2,学者们利用微分求积法[13-15]和加权残量法[16-18]等构造出了具有无条件稳定性的高阶时间积分方法。针对问题 3,借助能量有界准则[19],学者们发展了可稳定计算非线性动力响应的保能量方法[20-30]。为了全面提高时间积分方法的数值性能,通过使用更多时刻的已知状态变量,或不同分步采用不同类型的时间积分方法,学者们提出了线性多步法[31-38]和复合方法[39-50],这 2 类方法可有效解决问题 4。问题 5 更具有挑战性。为了提高稳定性,学者们提出了结构相关时间积分方法[51-54],在该类方法中,算法参数与系统质量、阻尼和刚度矩阵相关,数值结果表明,该类方法能够有效提高数值积分方法处理软刚度系统时的稳定性。此外,学者们还发现可利用幅值有界准则[55]或时变参数[31]解释在求解非线性问题时已有时间积分方法的失稳机理,并据此幅值有界准则设计了对一阶和二阶 ODE 都是无条件稳定的 BN 稳定型方法[56]。

本文着重介绍了针对上述问题提出的先进时间积分方法的发展状况,并对时间积分方法的未来发展趋势进行展望。

1　参数方法

为了提高 Newmark 方法耗散格式的收敛阶次,学者们基于 Newmark 方法发展了具有数值阻尼的二阶参数方法。根据构造方式不同,参数方法可以分成 2 类:一类是加权动力学平衡方程,另一类是在差分格式中引入辅助变量。

第一类参数方法的代表性工作有广义-α 方法[6-8]、HHT-α 方法[9]和 WBZ-α 方法[10]。这类方法沿用了 Newmark 方法的差分格式,即

$$\begin{cases} \boldsymbol{x}_{t+\Delta t} = \boldsymbol{x}_t + \Delta t\dot{\boldsymbol{x}}_t + \Delta t^2[(0.5-\beta)\ddot{\boldsymbol{x}}_t + \beta\ddot{\boldsymbol{x}}_{t+\Delta t}] \\ \dot{\boldsymbol{x}}_{t+\Delta t} = \dot{\boldsymbol{x}}_t + \Delta t[(1-\gamma)\ddot{\boldsymbol{x}}_t + \gamma\ddot{\boldsymbol{x}}_{t+\Delta t}] \end{cases} \tag{1}$$

式中:Δt 为时间步长;β 和 γ 为算法参数;$\boldsymbol{x}_t$、$\dot{\boldsymbol{x}}_t$ 和 $\ddot{\boldsymbol{x}}_t$ 分别为 t 时刻的位移、速度和加速度,以此类推。

为了解决 Newmark 方法耗散格式仅具有一阶精度问题,该类参数方法对动力学平衡方程进行了加权。以广义-α 方法为例,其在动力学平衡方程中引入 2 个参数 α 和 δ,即

$$\boldsymbol{M}[(1-\alpha)\ddot{\boldsymbol{x}}_{t+\Delta t} + \alpha\ddot{\boldsymbol{x}}_t] + \boldsymbol{C}[(1-\delta)\dot{\boldsymbol{x}}_{t+\Delta t} + \delta\dot{\boldsymbol{x}}_t] + \boldsymbol{K}[(1-\delta)\boldsymbol{x}_{t+\Delta t} + \delta\boldsymbol{x}_t] = (1-\delta)\boldsymbol{R}_{t+\Delta t} + \delta\boldsymbol{R}_t \tag{2}$$

式中:$\boldsymbol{M}$、$\boldsymbol{C}$、$\boldsymbol{K}$ 和 $\boldsymbol{R}$ 分别为质量矩阵、阻尼矩阵、刚度矩阵和外载荷列向量。

由式(2)可以看出,第一类参数方法不满足离散时间节点的平衡方程。广义-α 方法的参数可统一写成高频极限处谱半径 ρ_∞ 的函数,即

$$\begin{cases} \alpha = \dfrac{2\rho_\infty - 1}{\rho_\infty + 1} \\ \delta = \dfrac{\rho_\infty}{\rho_\infty + 1} \\ \gamma = \dfrac{3 - \rho_\infty}{2(\rho_\infty + 1)} \\ \beta = \dfrac{1}{(\rho_\infty + 1)^2} \end{cases} \tag{3}$$

上述参数可保证广义-α 方法具有二阶精度,并且高频段的数值耗散程度可由 ρ_∞ 精确调控。当 $\alpha=\delta=0$ 时,广义-α 方法退化为 Newmark 方法;当 $\alpha=0$ 时,广义-α 方法退化为 HHT-α 方法;当 $\delta=0$ 时,广义-α 方法退化为 WBZ-α 方法。以广义-α 方法为代表的第一类参数方法采用了加权思想,使得动力学平衡方程不能被严格满足,这导致由该类参数方法计算的加速度仅具有一阶精度。为了改善该性能,张慧敏和邢誉峰[11]提出了第二类参数方法,其在时间节点处严格满足动力

学平衡方程,即

$$\boldsymbol{M}\ddot{\boldsymbol{x}}_{t+\Delta t}+\boldsymbol{C}\dot{\boldsymbol{x}}_{t+\Delta t}+\boldsymbol{K}\boldsymbol{x}_{t+\Delta t}=\boldsymbol{R}_{t+\Delta t} \tag{4}$$

在差分格式方面,该类参数方法引入1个额外变量$\boldsymbol{\theta}$(可类比为位移对时间的三阶导数)为

$$\begin{cases}\boldsymbol{x}_{t+\Delta t}=\boldsymbol{x}_t+\Delta t\dot{\boldsymbol{x}}_t+\dfrac{1}{2}\Delta t^2\ddot{\boldsymbol{x}}_t+\\ \qquad\dfrac{1}{6}\Delta t^3[(1-\alpha)\boldsymbol{\theta}_t+\alpha\boldsymbol{\theta}_{t+\Delta t}]\\ \dot{\boldsymbol{x}}_{t+\Delta t}=\dot{\boldsymbol{x}}_t+\Delta t\ddot{\boldsymbol{x}}_t+\dfrac{1}{2}\Delta t^2[(1-\delta)\boldsymbol{\theta}_t+\delta\boldsymbol{\theta}_{t+\Delta t}]\\ \ddot{\boldsymbol{x}}_{t+\Delta t}=\ddot{\boldsymbol{x}}_t+\Delta t[(1-\gamma)\boldsymbol{\theta}_t+\gamma\boldsymbol{\theta}_{t+\Delta t}]\end{cases} \tag{5}$$

因此,该类方法被称为三参数单步法(three parameters single-step method, TPSM)。引入的辅助变量$\boldsymbol{\theta}$是一个中间变量,其本身不需要满足任何状态方程。因此,在初始时刻,$\boldsymbol{\theta}$可以设置为0以避免额外的启动程序。采用相同的构造方式,邢誉峰和张慧敏[12]还发展了四参数保辛单步法(four parameters single-step method, FPTM)。

2 高阶无条件稳定方法

通过减小时间步长来提高精度最容易实现,这相当于有限元方法中的h收敛性,但不可避免会导致计算量增加,而且会引入舍入误差。升高时间积分方法的收敛阶次是另一种提高计算精度的有效方式,相当于有限元方法中的p收敛性。虽然线性多步法的稳定性受Dahlquist定理[3]的约束,即当收敛阶次大于2时,算法是条件稳定或不稳定,不过学者们利用多级升阶思想构造出了具有无条件稳定性的高阶时间积分方法。

多级方法的定义是:在每个时间步内,在多个时刻或多次在同一时刻满足动力学平衡方程的时间积分方法。代表性工作有Runge-Kutta法[2]、微分求积法[13-15]和加权残量法[16-18]等。多级方法的稳定性不受Dahlquist定理约束,可以同时实现高阶精度和无条件稳定性,随着算法级数的增加,方法的收敛阶次随之提高,但伴随着成倍增加的计算量。

基于多级思想建立的Runge-Kutta法[2]因其精度高而被广泛用于求解一阶ODE。若用其求解二阶ODE,则需将二阶ODE转化为一阶ODE。Runge-Kutta法按构造格式可以分为显式、对角隐式和完全隐式3类。利用高斯积分,s级完全隐式Runge-Kutta法可以达到$2s$阶精度。但是,如此改善精度导致计算成本大幅度增加,因此其实用价值不高。综合精度、效率和稳定性,对角隐式Runge-Kutta法更受欢迎。Runge-Kutta法的求解格式为

$$\boldsymbol{y}_{t+\Delta t}=\boldsymbol{y}_t+\Delta t\sum_{j=1}^{s}b_j\boldsymbol{k}_j \tag{6}$$

$$\boldsymbol{k}_j=\boldsymbol{f}\left(t+c_j\Delta t,\boldsymbol{y}_t+\Delta t\sum_{i=1}^{s}a_{ij}\boldsymbol{k}_i\right)\quad j=1,2,\cdots,s \tag{7}$$

对角隐式Runge-Kutta法的Butcher表如下:

$$\begin{array}{c|c}\boldsymbol{c}&\boldsymbol{A}\\\hline&\boldsymbol{b}^{\mathrm{T}}\end{array}\rightarrow$$

$$\begin{array}{c|cccccccc}0&0&0&0&\cdots&0&0&0\\4c_2&2c_2&2c_2&0&\cdots&0&0&0\\c_3&a_{31}&a_{32}&2c_2&\cdots&0&0&0\\\vdots&\vdots&\vdots&\vdots&&\vdots&0&0\\c_{s-2}&a_{s-2,1}&a_{s-2,2}&a_{s-2,3}&\cdots&2c_2&0&0\\c_{s-1}&a_{s-1,1}&a_{s-1,2}&a_{s-1,3}&\cdots&a_{s-2,s-2}&2c_2&0\\1&b_1&b_2&b_3&\cdots&b_{s-2}&b_{s-1}&2c_2\\\hline&b_1&b_2&b_3&\cdots&b_{s-2}&b_{s-1}&2c_2\end{array} \tag{8}$$

由式(8)可以看出,对角隐式Runge-Kutta法的系数矩阵$\boldsymbol{A}$是一个对角元素不全为0的下三角矩阵。通过参数设计和选择足够的级数,隐式Runge-Kutta法可以具有任意高阶精度和无条件稳定性。

基于微分求积思想、加权残量思想、最小二乘法等,Fung[13-14,16,18]构造出一些高阶无条件稳定方法。在Fung的工作基础上,一些性能更加优秀的高阶方法被提出,如Huang方法[57]和Kim方法[58]。Huang和Fu[57]将不同格式的Fung方法混合使用在一个时间步长内,从而有效提高了低频精度。在不损失高阶精度和无条件稳定性的前提下,Kim方法[58]优化了一个时间步内时间离散点的位置,使得最后一个时间配点与时间步的终点重合,从而简化了Fung方法的执行流程。下面以Kim方法为例,简述该类高阶方法的构造流程。在Kim方法中,位移、速度和加速度的表达式如下:

$$\begin{cases}\boldsymbol{u}(\bar{t})\cong\bar{\boldsymbol{u}}(\bar{t})=\sum_{j=0}^{n}\psi_j(\bar{t})\boldsymbol{u}_j\\ \boldsymbol{v}(\bar{t})\cong\bar{\boldsymbol{v}}(\bar{t})=\sum_{j=0}^{n}\psi_j(\bar{t})\boldsymbol{v}_j\\ \boldsymbol{a}(\bar{t})\cong\bar{\boldsymbol{a}}(\bar{t})=\sum_{j=0}^{n}\psi_j(\bar{t})\boldsymbol{a}_j\end{cases} \tag{9}$$

式中:ψ_j为第j个时间点的Lagrange插值函数;$\boldsymbol{u}_j$、$\boldsymbol{v}_j$和$\boldsymbol{a}_j$分别为第j个时间点的位移、速度和加速度。

Lagrange 插值函数的定义为

$$\psi_j(\bar{t}) = \prod_{\substack{s=0 \\ (s\neq j)}}^{n} \frac{\tilde{t}-\tau_s}{\tau_j-\tau_s} \qquad \tilde{t} = \frac{\bar{t}-t_k}{\Delta t}, \bar{t} \in [t_k, t_{k+1}] \tag{10}$$

式中:t_k 为第 k 个时间步的初始时刻;τ_s 和 Δt 为第 k 个时间步内的配点。

在式(9)中,位移、速度和加速度是彼此独立的。Kim 方法采用加权残量法建立速度-位移和加速度-速度之间的关系,即定义如下残量:

$$\begin{cases} \boldsymbol{r}_1(\bar{t}) = \bar{\boldsymbol{v}}(\bar{t}) - \dot{\bar{\boldsymbol{u}}}(\bar{t}) \\ \boldsymbol{r}_2(\bar{t}) = \bar{\boldsymbol{a}}(\bar{t}) - \dot{\bar{\boldsymbol{v}}}(\bar{t}) \end{cases} \tag{11}$$

通过要求上述 2 个残量的积分为零,即可得到速度-位移和加速度-速度之间的关系,即

$$\begin{bmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \\ \vdots \\ \boldsymbol{v}_n \end{bmatrix} = \begin{bmatrix} \alpha_{11}\boldsymbol{I} & \cdots & \alpha_{1n}\boldsymbol{I} \\ \alpha_{21}\boldsymbol{I} & \cdots & \alpha_{2n}\boldsymbol{I} \\ \vdots & & \vdots \\ \alpha_{n1}\boldsymbol{I} & \cdots & \alpha_{nn}\boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_n \end{bmatrix} + \begin{bmatrix} \beta_1\boldsymbol{I} \\ \beta_2\boldsymbol{I} \\ \vdots \\ \beta_n\boldsymbol{I} \end{bmatrix} \boldsymbol{u}_0 + \begin{bmatrix} \gamma_1\boldsymbol{I} \\ \gamma_2\boldsymbol{I} \\ \vdots \\ \gamma_n\boldsymbol{I} \end{bmatrix} \boldsymbol{v}_0 \tag{12}$$

$$\begin{bmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \\ \vdots \\ \boldsymbol{a}_n \end{bmatrix} = \begin{bmatrix} \alpha_{11}\boldsymbol{I} & \cdots & \alpha_{1n}\boldsymbol{I} \\ \alpha_{21}\boldsymbol{I} & \cdots & \alpha_{2n}\boldsymbol{I} \\ \vdots & & \vdots \\ \alpha_{n1}\boldsymbol{I} & \cdots & \alpha_{nn}\boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_1 \\ \boldsymbol{v}_2 \\ \vdots \\ \boldsymbol{v}_n \end{bmatrix} + \begin{bmatrix} \beta_1\boldsymbol{I} \\ \beta_2\boldsymbol{I} \\ \vdots \\ \beta_n\boldsymbol{I} \end{bmatrix} \boldsymbol{v}_0 + \begin{bmatrix} \gamma_1\boldsymbol{I} \\ \gamma_2\boldsymbol{I} \\ \vdots \\ \gamma_n\boldsymbol{I} \end{bmatrix} \boldsymbol{a}_0 \tag{13}$$

结合 $t+\tau_i\Delta t(i=1,2,\cdots,n)$ 时刻的动力学平衡方程,即可求解出所有未知的状态变量。基于微分求积思想、加权残量思想、最小二乘法等建立的高阶方法在有数值耗散时具有 $(2n-1)$ 阶精度,在无数值耗散时是 $2n$ 阶精确的。

3　保能量方法

著名的 TR(即欧拉中点辛差分格式)谱半径特性表明其是一种无条件稳定方法,但将其应用于求解一些简单的非线性问题时,如刚性单摆[4]和软弹簧方程[5],发现若时间步长选取的不合适,TR 会发散。虽然数值阻尼可以增强稳定性,但当把耗散方法用于非线性问题求解时,其结果仍可能不收敛。

为了稳定地计算非线性动力学响应,学者们基于能量有界准则[19]建立了保能量方法。在该类方法中,当前时间步的机械能小于或等于上一时间步的机械能与外力功之和。据此准则,Hughes 等[22]首先提出了约束能量型方法(constraint energy method, CEM)。CEM 以 TR 为基础,通过 Lagrange 乘子法施加能量准则约束。然而,该方法需要额外计算 Lagrange 乘子和离散的能量值,在非线性迭代中还会遇到收敛困难问题。不同于 CEM,Simo 和 Wong[30]提出了另一种保能量方法,即能量-动量法(energy momentum method, EMM)。EMM 是基于修正的中点法则建立的,在使用 Green 应变的非线性系统中可以保守能量,但其难以推广到一般非线性系统。

针对一般非线性系统,已有 Gonazlez 方法[26]、Krenk 方法[25]等。这些方法均基于中点法则,修正后的中点内力能够满足能量守恒的要求,但也都需要计算离散点的能量值。为了解决这个问题,笔者直接从能量准则出发,提出了一种更为通用的保能量方法[21](energy-conserving method, ECM)。该方法利用 Gauss-Legendre 求积法则计算平均内力,数值结果表明,只要使用足够的节点数,就可以精确地保守系统能量。更重要的是,该方法仅需已知内力形式即可,不需要计算能量,并且适用于同时包含阻尼和刚度非线性的动力学系统。

下面以 CEM 为例,简述保能量方法的实现过程。考虑如下非线性保守系统:

$$\boldsymbol{M}\ddot{\boldsymbol{x}} + \boldsymbol{N}(\boldsymbol{x}) = \boldsymbol{0} \tag{14}$$

式中:$\boldsymbol{N}$ 为弹性内力。

对于保守系统,动能的离散形式为

$$T_t = \frac{1}{2}\dot{\boldsymbol{x}}_t^{\mathrm{T}}\boldsymbol{M}\dot{\boldsymbol{x}}_t \tag{15}$$

由于势能的相对性,其在一个时间步内的增量可以写为

$$\Delta V = V_t - V_{t+\Delta t} = \int_{\boldsymbol{x}_t}^{\boldsymbol{x}_{t+\Delta t}} N(\boldsymbol{x})^{\mathrm{T}}\mathrm{d}\boldsymbol{x} \tag{16}$$

根据能量守恒定律,保守系统的动能和势能之和应为常数,即机械能增量为零。

$$\Delta E = T_{t+\Delta t} + V_{t+\Delta t} - T_t - V_t = \Delta T + \Delta V = 0 \tag{17}$$

式(17)是所有保能量方法的设计准则。

CEM 是基于 TR 设计的,对于方程(14),其递推格式如下:

$$\frac{4}{\Delta t^2}\boldsymbol{M}\boldsymbol{x}_{t+\Delta t} + \boldsymbol{N}_{t+\Delta t} = \boldsymbol{M}\left[\ddot{\boldsymbol{x}}_t + \frac{4}{\Delta t^2}(\boldsymbol{x}_t + \Delta t\dot{\boldsymbol{x}}_t)\right] \tag{18}$$

$$\begin{cases} \dot{\boldsymbol{x}}_{t+\Delta t} = \dfrac{2}{\Delta t}(\boldsymbol{x}_{t+\Delta t} - \boldsymbol{x}_t) - \dot{\boldsymbol{x}}_t \\ \ddot{\boldsymbol{x}}_{t+\Delta t} = \dfrac{2}{\Delta t}(\dot{\boldsymbol{x}}_{t+\Delta t} - \dot{\boldsymbol{x}}_t) - \ddot{\boldsymbol{x}}_t \end{cases} \tag{19}$$

在 CEM 中,定义一个泛函 $\Gamma(\boldsymbol{x}_{t+\Delta t})$,其形式如下:

$$\Gamma(\boldsymbol{x}_{t+\Delta t}) = \frac{2}{\Delta t^2}\boldsymbol{x}_{t+\Delta t}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{x}_{t+\Delta t} + V(\boldsymbol{x}_{t+\Delta t}) - \boldsymbol{x}_{t+\Delta t}^{\mathrm{T}}\boldsymbol{M}\left(\ddot{\boldsymbol{x}}_t + \frac{4}{\Delta t^2}(\boldsymbol{x}_t + \Delta t\dot{\boldsymbol{x}}_t)\right) \tag{20}$$

其变分为

$$\delta\boldsymbol{x}_{t+\Delta t}^{\mathrm{T}}\frac{\partial\Gamma(\boldsymbol{x}_{t+\Delta t})}{\partial\boldsymbol{x}_{t+\Delta t}} = \delta\boldsymbol{x}_{t+\Delta t}^{\mathrm{T}}\left[\frac{4}{\Delta t^2}\boldsymbol{M}\boldsymbol{x}_{t+\Delta t} + \boldsymbol{N}(\boldsymbol{x}_{t+\Delta t}) - \boldsymbol{M}\left(\ddot{\boldsymbol{x}}_t + \frac{4}{\Delta t^2}(\boldsymbol{x}_t + \Delta t\dot{\boldsymbol{x}}_t)\right)\right] \tag{21}$$

由式(21)可见,式(18)等价于泛函 $\Gamma(\boldsymbol{x}_{t+\Delta t})$ 的一阶变分等于零。为了满足能量守恒方程(17),CEM 将 $\Gamma(\boldsymbol{x}_{t+\Delta t})$ 修正为

$$\tilde{\Gamma}(\boldsymbol{x}_{t+\Delta t},\lambda) = \Gamma(\boldsymbol{x}_{t+\Delta t}) + \lambda\Delta E(\boldsymbol{x}_{t+\Delta t}) \tag{22}$$

式中:λ 为 Lagrange 乘子。

而式(22)对应的 Euler-Lagrange 方程为

$$\frac{\partial\tilde{\Gamma}(\boldsymbol{x}_{t+\Delta t},\lambda)}{\partial\boldsymbol{x}_{t+\Delta t}} = \frac{\partial\Gamma(\boldsymbol{x}_{t+\Delta t})}{\partial\boldsymbol{x}_{t+\Delta t}} + \lambda\frac{\partial\Delta E(\boldsymbol{x}_{t+\Delta t})}{\partial\boldsymbol{x}_{t+\Delta t}} = (1+\lambda)\left(\frac{4}{\Delta t^2}\boldsymbol{M}\boldsymbol{x}_{t+\Delta t} + \boldsymbol{N}(\boldsymbol{x}_{t+\Delta t})\right) - \boldsymbol{M}\left[\ddot{\boldsymbol{x}}_t + \frac{4}{\Delta t^2}\left((1+\lambda)\boldsymbol{x}_t + \Delta t\left(1+\frac{\lambda}{2}\right)\dot{\boldsymbol{x}}_t\right)\right] = \boldsymbol{0} \tag{23}$$

和

$$\frac{\partial\tilde{\Gamma}(\boldsymbol{x}_{t+\Delta t},\lambda)}{\partial\lambda} = \Delta E(\boldsymbol{x}_{t+\Delta t}) = 0 \tag{24}$$

通过式(23)和式(24)可联立求解出 $\boldsymbol{x}_{t+\Delta t}$ 和 λ,进而利用式(19)计算 $t+\Delta t$ 时刻的速度和加速度。从方程(24)可以看出,每一步 CEM 都需要计算能量,与 TR 相比计算量略有增加。

4 线性多步法和复合方法

为了综合提高时间积分方法的数值性能,包括精度、效率、耗散和稳定性,可以采用线性多步法和复合方法。线性多步法属于单步时间积分方法中的一种,但为了计算当前时刻的响应 $\boldsymbol{x}_{t+\Delta t}$,需要利用其他自启动方法计算前几个时刻的响应,以获得足够的启动信息。与线性多步法不同,复合方法是将一个时间步划分成若干个分步,计算 $\boldsymbol{x}_{t+\Delta t}$ 需要每个分步的信息。

4.1 线性多步法

工程界目前使用较多的线性多步法[2]有 Adams 类方法、显式 Adams-Bashforth 方法、隐式 Adams-Moulton 方法及预测校正类格式等。这些方法虽然具有高阶精度,但稳定性难以提高。Dahlquist[3]曾证明,所有显式方法和二阶以上的线性多步法均无法实现无条件稳定。因此,在实际应用中,显式和隐式线性多步方法常作为预测和校正方法成对使用,并结合自适应步长技术,控制局部截断误差,以避免结果发散。此外,向后差分格式,如 Houbolt 方法[59]和 Park 方法[1],也属于经典的线性多步法。它们一般具备二阶精度、无条件稳定性及强烈的高频耗散能力,可以快速滤掉高频响应成分,因此更适用于刚性动力学系统的分析。

为进一步提高线性多步法的计算精度,学者们构造出了更加精确的线性三步法[33]和线性四步法[33]。另外,为了解决线性多步法无法自启动的问题,通过引入一些辅助状态变量,学者们还提出了与一些线性多步法具有相同数值性能的单步法[32-33]。

下面以线性两步法[33](linear two-step method, LTM)为例,介绍该类方法的求解流程。LTM 的递推格式为

$$\begin{cases}\boldsymbol{M}\ddot{\boldsymbol{x}}_{t+\Delta t} + \boldsymbol{C}\dot{\boldsymbol{x}}_{t+\Delta t} + \boldsymbol{K}\boldsymbol{x}_{t+\Delta t} = \boldsymbol{R}(t+\Delta t)\\ \boldsymbol{x}_{t+\Delta t} = \alpha_1\boldsymbol{x}_t + \alpha_2\boldsymbol{x}_{t-\Delta t} + \Delta t(\beta_0\dot{\boldsymbol{x}}_{t+\Delta t} + \beta_1\dot{\boldsymbol{x}}_t + \beta_2\dot{\boldsymbol{x}}_{t-\Delta t})\\ \dot{\boldsymbol{x}}_{t+\Delta t} = \alpha_1\dot{\boldsymbol{x}}_t + \alpha_2\dot{\boldsymbol{x}}_{t-\Delta t} + \Delta t(\beta_0\ddot{\boldsymbol{x}}_{t+\Delta t} + \beta_1\ddot{\boldsymbol{x}}_t + \beta_2\ddot{\boldsymbol{x}}_{t-\Delta t})\end{cases} \tag{25}$$

或

$$\widehat{\boldsymbol{K}}\boldsymbol{x}_{t+\Delta t} = \widehat{\boldsymbol{R}} \tag{26}$$

式中:

$$\widehat{\boldsymbol{K}} = \frac{1}{\beta_0^2\Delta t^2}\boldsymbol{M} + \frac{1}{\beta_0\Delta t}\boldsymbol{C} + \boldsymbol{K}$$

$$\widehat{\boldsymbol{R}} = \boldsymbol{R}(t+\Delta t) + \boldsymbol{M}\left[\frac{1}{\beta_0^2\Delta t^2}(\alpha_1\boldsymbol{x}_t + \alpha_2\boldsymbol{x}_{t-\Delta t}) + \frac{\beta_1+\alpha_1\beta_0}{\beta_0^2\Delta t}\dot{\boldsymbol{x}}_t + \frac{\beta_2+\alpha_1\beta_0}{\beta_0^2\Delta t}\dot{x}_{t-\Delta t} + \frac{\beta_1}{\beta_0}\ddot{\boldsymbol{x}}_t + \frac{\beta_2}{\beta_0}\ddot{\boldsymbol{x}}_{t-\Delta t}\right] + \boldsymbol{C}\left[\frac{1}{\beta_0\Delta t}(\alpha_1\boldsymbol{x}_t + \alpha_2\boldsymbol{x}_{t-\Delta t}) + \frac{\beta_1}{\beta_0}\dot{\boldsymbol{x}}_t + \frac{\beta_2}{\beta_0}\dot{\boldsymbol{x}}_{t-\Delta t}\right]$$

由式(26)可以看出,LTM 需利用其他自启动方法计算 $t-\Delta t$ 时刻的状态变量,之后才能执行式(26)。

4.2 复合方法

复合方法在一个时间步的不同分步内使用不同的时间积分方法,以发挥不同方法各自的优势,从而获得优于各分步方法的数值性能。

Bank 等[60]在 1985 年把 TR 和后向差分公式(backward difference formula, BDF)组合使用,并将其用于求解一阶 ODE。Bathe 和 Baig[46]在

2005 年将这种复合方法 TR-BDF 用于求解二阶 ODE,并发展了复合方法概念。随后,Bathe 等进一步研究了 TR-BDF 方法在非线性系统[61]、波传播[62]等问题中的性能。TR-BDF 方法优越的低频精度和高频耗散能力受到了学术界和工程界的广泛关注。在 Bathe 方法基础上,学者们提出了一系列性能更加优越的复合方法,如三分步 TR-TR-BDF[50]、三分步 TR-BDF-Houbolt[48]、四分步 Optimal-TR-TR-TR-BDF[42] 等。与两分步 Bathe 方法[46]相比,这些方法的低频精度更高。

由 TR 和 BDF(或 Houbolt 方法)组合的复合方法具有 L 型稳定性。为了能够精确地调控高频段的数值耗散量,基于标准 Bathe 方法[46],Rezaiee-Pajand 和 Sarafrazi[37] 提出了 β_1/β_2-Bathe 方法,该方法的第二分步采用两点后向差分公式(backward interpolation formula, BIF),通过调整 β_1 和 β_2 的大小和比例,可以控制高频段的数值耗散量。之后,Noh 和 Bathe[44] 进一步优化了 β_1/β_2-Bathe 方法,提出了 ρ_∞-Bathe 方法,仅使用一个参数 ρ_∞ 来光滑和准确地改变数值耗散的大小。在具有可控数值耗散的前提下,为了进一步提高该类方法的低频精度,季奕和邢誉峰[43] 提出了一种优化三分步方法。

下面以 Bathe 方法[46] 为例介绍复合方法的求解思路。在 Bathe 方法中,把一个时间步 $[t,\ t+\Delta t]$ 划分为 2 个分步 $[t,\ t+\gamma\Delta t]$ 和 $[t+\gamma\Delta t,\ t+\Delta t]$,其中 $\gamma\Delta t$ 表示第 1 个分步的大小。TR 被用在第 1 个分步中,即

$$\begin{cases} \boldsymbol{x}_{t+\gamma\Delta t} = \boldsymbol{x}_t + \dfrac{\gamma\Delta t}{2}(\dot{\boldsymbol{x}}_t + \dot{\boldsymbol{x}}_{t+\gamma\Delta t}) \\ \dot{\boldsymbol{x}}_{t+\gamma\Delta t} = \dot{\boldsymbol{x}}_t + \dfrac{\gamma\Delta t}{2}(\ddot{\boldsymbol{x}}_t + \ddot{\boldsymbol{x}}_{t+\gamma\Delta t}) \end{cases} \tag{27}$$

第 2 个分步采用三点 BDF,其格式如下:

$$\begin{cases} \Delta t\dot{\boldsymbol{x}}_{t+\gamma\Delta t} = \theta_2\boldsymbol{x}_{t+\Delta t} + \theta_1\boldsymbol{x}_{t+\gamma\Delta t} + \theta_0\boldsymbol{x}_t \\ \Delta t\ddot{\boldsymbol{x}}_{t+\gamma\Delta t} = \theta_2\dot{\boldsymbol{x}}_{t+\Delta t} + \theta_1\dot{\boldsymbol{x}}_{t+\gamma\Delta t} + \theta_0\dot{\boldsymbol{x}}_t \end{cases} \tag{28}$$

式中:θ_0、θ_1 和 θ_2 为算法参数。

联立求解 $t+\gamma\Delta t$ 和 $t+\Delta t$ 时刻的动力学平衡方程,即可得到一个时间步内所有状态变量。

另外,针对由 TR 和 BDF(或 BIF)复合得到的时间积分方法,邢誉峰和季奕等[42] 研究了如何选择分步数、分步步长、差分点数、各分步采用的方法,并得到了一些重要的结论,为复合方法的构造提供指导,这些结论包括:

1) 当 BDF(或 BIF)采用的差分点数不变时,随着分步数的增加,TR 在一个时间步中的占比增加,从而提高了低频精度。但是,这种方法不能明显提高低频精度。

2) 当分步数固定时,提高 BDF 的差分点数(或 BIF 中的插值点数),可以有效提高低频精度。实际上,随着差分点数目增加,一个时间步内采用 TR 的分步的占比略有降低。因此,可以认为增加差分点数比增加分步数更有助于提高低频精度。

3) 当分步数为常数时,只要采用相同的参数优化方式,就可以构造出性能相近的复合方法。如三分步优化 Optimal-TR-BDF-BDF 方法中采用 TR 的分步的占比大约是三分步优化 Optimal-TR-TR-BDF 方法中的一半,但二者的计算精度几乎相同。

5　BN 稳定型方法

引入数值阻尼是改善时间积分方法在非线性系统中稳定性的一种直接且有效的方式。与非耗散方法相比,一般情况下耗散方法在非线性系统中具有稳定性优势,但其精度变低。为了保证在计算非线性动力学响应时的稳定性,保能量方法是一种选择。但保能量方法具有额外的计算量,并且适用范围有限。

与耗散方法和保能量方法的思想不同,季奕、邢誉峰和 Wiercigroch[56] 基于 BN 稳定性理论[2] 设计出一种对任意非线性初值问题均具有无条件稳定性的两分步方法,记为 ρ_∞-TSSBN (ρ_∞-two sub step BN-stable) 方法。BN 稳定性理论是由 Butcher 于 1975 年提出的,其目的是为了分析 Runge-Kutta 法在求解一阶初值问题时的稳定性。该理论认为:若 Runge-Kutta 法的解始终满足 $\|\boldsymbol{y}_{t+\Delta t}\| \leqslant \|\boldsymbol{y}_t\|$,则认为其对一阶线性和非线性 ODE 均是无条件稳定的。

下面简单介绍 ρ_∞-TSSBN 方法的求解格式和算法参数。该方法包含 2 个分步 $[t,\ t+c_1\Delta t]$ 和 $[t+c_1\Delta t,\ t+c_2\Delta t]$,其中,$0<c_1<c_2<1$。第 1 个分步的求解格式为

$$\begin{cases} \boldsymbol{M}\ddot{\boldsymbol{x}}_{t+c_1\Delta t} + \boldsymbol{F}(\boldsymbol{x}_{t+c_1\Delta t}, \dot{\boldsymbol{x}}_{t+c_1\Delta t}, t+c_1\Delta t) = \boldsymbol{0} \\ \boldsymbol{x}_{t+c_1\Delta t} = \boldsymbol{x}_t + c_1\Delta t\dot{\boldsymbol{x}}_{t+c_1\Delta t} \\ \dot{\boldsymbol{x}}_{t+c_1\Delta t} = \dot{\boldsymbol{x}}_t + c_1\Delta t\ddot{\boldsymbol{x}}_{t+c_1\Delta t} \end{cases} \tag{29}$$

第 2 个分步格式为

$$\begin{cases} \boldsymbol{M}\ddot{\boldsymbol{x}}_{t+c_2\Delta t} + \boldsymbol{F}(\boldsymbol{x}_{t+c_2\Delta t}, \dot{\boldsymbol{x}}_{t+c_2\Delta t}, t+c_2\Delta t) = \boldsymbol{0} \\ \boldsymbol{x}_{t+c_2\Delta t} = \boldsymbol{x}_t + c_2\Delta t[(1-\alpha)\dot{\boldsymbol{x}}_{t+c_1\Delta t} + \alpha\dot{\boldsymbol{x}}_{t+c_2\Delta t}] \\ \dot{\boldsymbol{x}}_{t+c_2\Delta t} = \dot{\boldsymbol{x}}_t + c_2\Delta t[(1-\alpha)\ddot{\boldsymbol{x}}_{t+c_1\Delta t} + \alpha\ddot{\boldsymbol{x}}_{t+c_2\Delta t}] \end{cases} \tag{30}$$

式中:$\boldsymbol{F}$ 为包含内力和外力的向量。

由式(29)和式(30)可求解出 $t+c_1\Delta t$ 和 $t+c_2\Delta t$ 两个时刻的状态变量。时间步终点 $t+\Delta t$ 的信息是由这 2 个时刻的状态变量加权得到,即

$$\begin{cases}\boldsymbol{x}_{t+\Delta t}=\boldsymbol{x}_t+\Delta t[b_1\dot{\boldsymbol{x}}_{t+c_1\Delta t}+(1-b_1)\dot{\boldsymbol{x}}_{t+c_2\Delta t}]\\ \dot{\boldsymbol{x}}_{t+\Delta t}=\dot{\boldsymbol{x}}_t+\Delta t[b_1\ddot{\boldsymbol{x}}_{t+c_1\Delta t}+(1-b_1)\ddot{\boldsymbol{x}}_{t+c_2\Delta t}]\\ \boldsymbol{M}\ddot{\boldsymbol{x}}_{t+\Delta t}+\boldsymbol{F}(\boldsymbol{x}_{t+\Delta t},\dot{\boldsymbol{x}}_{t+\Delta t},t+\Delta t)=\boldsymbol{0}\end{cases}\tag{31}$$

从式(29)可以看出,$t_0+c_1\Delta t$ 时刻状态变量的求解不需要初始时刻 t_0 的加速度信息,因此,该方法是一种真正的自启动方法。

从式(31)可以看出,$t+\Delta t$ 时刻不需要有效刚度矩阵分解和 Newton 迭代运算。因次,从计算量的角度来说,该方法属于两分步格式。

为了确保 ρ_∞-TSSBN 方法能够稳定地处理非线性系统,即具有 BN 稳定性,利用 BN 稳定性理论的充分条件(即代数稳定性)设计了该算法参数。当 $\rho_\infty=1$ 时,算法参数为

$$c_1=\frac{1}{4},c_2=\frac{3}{4},b_1=\frac{1}{2},\alpha=\frac{1}{3}\tag{32}$$

当 $0\leqslant\rho_\infty<1$ 时,算法参数为

$$\begin{cases}c_1=\dfrac{2-\sqrt{2(\rho_\infty+1)}}{2(1-\rho_\infty)}\\ c_2=\dfrac{\sqrt{2}(\rho_\infty+1)-2\rho_\infty\sqrt{\rho_\infty+1}}{2(1-\rho_\infty)\sqrt{\rho_\infty+1}}\\ b_1=\dfrac{1}{2}\\ \alpha=\dfrac{\sqrt{2}-\sqrt{\rho_\infty+1}}{\sqrt{\rho_\infty+1}-\sqrt{2}\rho_\infty}\end{cases}\tag{33}$$

数值算例表明,在 TR 和 ρ_∞-Bathe 方法失效的非线性问题中,ρ_∞-TSSBN 方法可以给出稳定且精确的结果。

6 结束语

本文介绍了用于求解动力学常微分方程的时间积分方法,并着重介绍了近年来发展出的一些先进时间积分方法,包括参数方法、高阶无条件稳定方法、保能量方法、线性多步法、复合方法和 BN 稳定型方法。更详细的资料读者可以阅读相关的综述文章[63]和专著[64]。

关于时间积分方法的未来发展问题,可以从其求解的是线性问题和非线性问题角度进行讨论。

对于线性动力系统,时间积分方法的设计已有成熟的理论基础。虽然已经存在高精度、快速计算方法[65-67],但如何更快、更精确地计算仍然是人们追求的目标。对于线性单自由度动力系统,已经可以利用时间积分方法得到精确的结点位移[68-69],其具有精确的幅值和相位。对于线性多自由度系统,虽然辛几何算法[70]具有保幅值或保能量特性,或者说从根本上解决了幅值误差累计问题,但相位误差累计问题仍然存在。幅值有界准则和能量准则都是针对幅值衰减和发散问题提出的,但建立一个针对相位误差的约束准则是值得探索的。

对于复杂非线性动力系统,时间积分方法的精度和稳定性应该是更加本质的问题。季奕等[56]基于 BN 稳定性理论已经设计出两分步二阶无条件稳定时间积分方法,但发展更精确高效的 BN 稳定型方法是一项有意义和挑战的工作。此外,虽然季奕和邢誉峰[31]已经建立了时变参数谱分析理论,并利用此理论设计了对非线性刚度系统是无条件稳定的方法,但尚未将这种方法推广用于非线性阻尼系统等。非线性系统的高阶时间积分方法的设计及多尺度快速计算问题等也都是值得关注的问题。

现代科学发展的特征是多学科交叉、多物理场耦合和多尺度分析和非线性特征,对于结构动力学系统时间积分方法的研究工作也要与现代科学发展的步伐协调一致。

参考文献(References)

[1] HUGHES T J R. The finite element method: Linear static and dynamic finite element analysis[M]. Upper Saddle River: Prentice-Hall Press, 1987.

[2] BUTCHER J C. Numerical methods for ordinary differential equations[M]. New York: John Wiley & Sons, Ltd, 2016.

[3] DAHLQUIST G G. A special stability problem for linear multistep methods[J]. BIT Numerical Mathematics, 1963, 3(1): 27-43.

[4] KUHL D, CRISFIELD M A. Energy-conserving and decaying algorithms in nonlinear structural dynamics[J]. International Journal for Numerical Methods in Engineering, 1999, 45(5): 569-599.

[5] PARK K C. An improved stiffly stable method for direct integration of nonlinear structural dynamics[J]. Journal of Applied Mechanics-Transactions of the ASME, 1975, 42(2): 464-470.

[6] 邵慧萍,蔡承文. 结构动力学方程数值积分的三参数算法[J]. 应用力学学报,1988,5(4):77-81.
SHAO H P, CAI C W. A three parameters algorithm for numerical integration of structural dynamic equations[J]. Chinese Journal of Applied Mechanics, 1988, 5(4): 77-81(in Chinese).

[7] CHUNG J, HULBERT G. A time integration algorithm for structural dynamics with improved numerical dissipation: The generalized-α method[J]. Journal of Applied Mechanics-Transactions of the ASME, 1993, 60(2): 371-375.

[8] CHUNG J,LEE J M. A new family of explicit time integration methods for linear and nonlinear structural dynamics[J]. International Journal for Numerical Methods in Engineering,1994,37(23):3961-3976.

[9] HILBER H M,HUGHES T J R,TAYLOR R L. Algorithms in structural dynamics[J]. Earthquake Engineering and Structural Dynamics,1977,5:283-292.

[10] WOOD W L,BOSSAK M,ZIENKIEWICZ O C. An alpha modification of Newmark's method[J]. International Journal for Numerical Methods in Engineering,1980,15(10):1562-1566.

[11] ZHANG H M,XING Y F. A three-parameter single-step time integration method for structural dynamic analysis[J]. Acta Mechanica Sinica,2019,35(1):112-128.

[12] XING Y F,ZHANG H M. An efficient non-dissipative higher-order single-step integration method for long-term dynamics simulation[J]. International Journal of Structural Stability and Dynamics,2018,18(9):1850113.

[13] FUNG T C. Solving initial value problems by differential quadrature method: Part 1-First-order equations[J]. International Journal for Numerical Methods in Engineering,2001,50(6):1411-1427.

[14] FUNG T C. Solving initial value problems by differential quadrature method: Part 2-Second-and higher-order equations[J]. International Journal for Numerical Methods in Engineering,2001,50(6):1429-1454.

[15] KIM W,LEE J H. A comparative study of two families of higher-order accurate time integration algorithms[J]. International Journal of Computational Methods,2020,17(8):1950048.

[16] FUNG T C. Weighting parameters for unconditionally stable higher-order accurate time step integration algorithms: Part 1: First-order equations[J]. International Journal for Numerical Methods in Engineering,1999,45(8):941-971.

[17] JI Y,XING Y F. An improved higher-order time integration algorithm for structural dynamics[J]. CMES-Computer Modelling in Engineering & Sciences,2021,126(2):549-575.

[18] FUNG T C. Weighting parameters for unconditionally stable higher-order accurate time step integration algorithms: Part 2: Second-order equations[J]. International Journal for Numerical Methods in Engineering,1999,45(8):971-1006.

[19] BELYTSCHKO T,SCHOEBERLE D F. On the unconditional stability of an implicit algorithm for nonlinear structural dynamics[J]. Journal of Applied Mechanics-Transactions of the ASME,1975,42(4):865-869.

[20] LAVRENČIČ M,BRANK B. Comparison of numerically dissipative schemes for structural dynamics: Generalized-alpha versus energy-decaying methods[J]. Thin-Walled Structures,2020,157:107075.

[21] ZHANG H M,XING Y F,JI Y. An energy-conserving and decaying time integration method for general nonlinear dynamics[J]. International Journal for Numerical Methods in Engineering,2020,121(5):925-944.

[22] HUGHES T J R,CAUGHEYV T K,LIU W K. Finite-element methods for nonlinear elastodynamics which conserve energy[J]. Journal of Applied Mechanics-Transactions of the ASME,1978,45(2):366-370.

[23] WU B,PAN T L,YANG H W,et al. Energy-consistent integration method and its application to hybrid testing[J]. Earthquake Engineering and Structural Dynamics,2020,49(5):415-433.

[24] ZHANG R,ZHONG H Z. A quadrature element formulation of an energy-momentum conserving algorithm for dynamic analysis of geometrically exact beams[J]. Computers & Structures,2016,165:96-106.

[25] KRENK S. Global format for energy-momentum based time integration in nonlinear dynamics[J]. International Journal for Numerical Methods in Engineering,2014,100(6):458-476.

[26] GONZALEZ O. Exact energy and momentum conserving algorithms for general models in nonlinear elasticity[J]. Computer Methods in Applied Mechanics and Engineering,2000,190(13-14):1763-1783.

[27] MAMOURI S, HAMMADI F, IBRAHIMBEGOVIC A. Decaying/conserving implicit scheme and non-linear instability analysis of 2D frame structures[J]. International of Journal Non-Linear Mechanics,2014,67:144-152.

[28] ORDEN J C. Energy and symmetry-preserving formulation of nonlinear constraints and potential forces in multibody dynamics[J]. Nonlinear Dynamics,2019,95:823-837.

[29] KUHL D, EKKELARD R. Constraint energy momentum algorithm and its application to non-linear dynamics of shells[J]. Computer Methods in Applied Mechanics and Engineering,1996,136(3-4):293-315.

[30] SIMO J C, WONG K K. Unconditionally stable algorithms for rigid body dynamics that exactly preserve energy and momentum[J]. International Journal for Numerical Methods in Engineering,1991,31(1):19-52.

[31] JI Y,XING Y F. A two-step time integration method with desirable stability for nonlinear structural dynamics[J]. European Journal of Mechanics-A/Solids,2022,94:104582.

[32] ZHANG J. A-stable linear two-step time integration methods with consistent starting and their equivalent single-step methods in structural dynamics analysis[J]. International Journal for Numerical Methods in Engineering,2021,122(9):2312-2359.

[33] ZHANG H M,ZHANG R S,MASARATI P. Improved second-order unconditionally stable schemes of linear multi-step and equivalent single-step integration methods[J] Computational Mechanics,2021,67:289-313.

[34] DONG S. BDF-like methods for nonlinear dynamic analysis[J]. Journal of Computational Physics,2010,229(8):3019-3045.

[35] ZHANG J. A-stable two-step time integration methods with controllable numerical dissipation for structural dynamics[J]. International Journal for Numerical Methods in Engineering,2020,121(1):54-92.

[36] YANG C,WANG X,LI Q,et al. An improved explicit integration algorithm with controllable numerical dissipation for structural dynamics[J]. Archive of Applied Mechanics,2020,90:2413-2431.

[37] REZAIEE-PAJAND M,SARAFRAZI S R. A mixed and multi-step higher-order implicit time integration family[J]. Proceedings of the Institution of Mechanical Engineers,Part C:Journal of Mechanical Engineering Science, 2010, 224 (10): 2097-

2108.

[38] ZHOU X, TAMMA K K. Design, analysis, and synthesis of generalized single step single solve and optimal algorithms for structural dynamics[J]. International Journal for Numerical Methods in Engineering, 2004, 59(5): 597-668.

[39] LI J Z, YU K P. Development of composite sub-step explicit dissipative algorithms with truly self-starting property[J]. Nonlinear Dynamics, 2021, 103: 1-26.

[40] JI Y, XING Y F. Optimization of a class of n-sub-step time integration methods for structural dynamics[J] International Journal of Applied Mechanics, 2021, 13(6): 2150064.

[41] LIU T H, HUANG F L, WEN W B, et al. Further insights of a composite implicit time integration scheme and its performance on linear seismic response analysis[J]. Engineering Structures, 2021, 241: 112490.

[42] XING Y F, JI Y, ZHANG H M. On the construction of a type of composite time integration methods[J] Computers & Structures, 2019, 221: 157-178.

[43] JI Y, XING Y F. An optimized three-sub-step composite time integration method with controllable numerical dissipation[J] Computers & Structures, 2020, 231: 106210.

[44] NOH G, BATHE K J. The Bathe time integration method with controllable spectral radius: The ρ_∞ -Bathe method[J] Computers & Structures, 2019, 212: 299-310.

[45] KIM W. An improved implicit method with dissipation control capability: The simple generalized composite time integration algorithm[J]. Applied Mathematical Modelling, 2020, 81: 910-930.

[46] BATHE K J, BAIG M M I. On a composite implicit time integration procedure for nonlinear dynamics[J]. Computers & Structures, 2005, 83(31-32): 2513-2524.

[47] LI J Z, ZHAO R, YU K P, et al. Directly self-starting higher-order implicit integration algorithms with flexible dissipation control for structural dynamics[J]. Computer Methods in Applied Mechanics and Engineering, 2022, 389: 114274.

[48] WEN W B, WEI K, LEI H S, et al. A novel sub-step composite implicit time integration scheme for structural dynamics[J]. Computers & Structures, 2017, 182: 176-186.

[49] NOH G, HAM S, BATHE K J. Performance of an implicit time integration scheme in the analysis of wave propagations[J]. Computers & Structures, 2013, 123: 93-105.

[50] CHANDRA Y, ZHOU Y, STANCIULESCU I, et al. A robust composite time integration scheme for snap-through problems[J]. Computational Mechanics, 2015, 55: 1041-1056.

[51] KOLAY C, RICLES J M. Development of a family of unconditionally stable explicit direct integration algorithms with controllable numerical energy dissipation[J]. Earthquake Engineering and Structural Dynamics, 2014, 43(9): 1361-1380.

[52] KOLAY C, RICLES J M. Improved explicit integration algorithms for structural dynamic analysis with unconditional stability and controllable numerical dissipation[J]. Journal of Earthquake Engineering, 2019, 23(5): 771-792.

[53] DU X Q, YANG D X, ZHOU J L, et al. New explicit integration algorithms with controllable numerical dissipation for structural dynamics[J]. International Journal of Structural Stability and Dynamics, 2018, 18(3): 1850044.

[54] LI J Z, YU K P. Noniterative integration algorithms with controllable numerical dissipations for structural dynamics[J]. International Journal of Computer Mathematics, 2018, 15(3): 1850111.

[55] BURRAGE K, BUTCHER J C. Stability criteria for implicit Runge-Kutta methods[J]. SIAM Journal on Numerical Analysis, 1979, 16(1): 46-57.

[56] JI Y, XING Y F, WIERCIGROCH M. An unconditionally stable time integration method with controllable dissipation for second-order nonlinear dynamics[J]. Nonlinear Dynamics, 2021, 105: 3341-3358.

[57] HUANG C, FU M H. A composite collocation method with low-period elongation for structural dynamics problems[J]. Computers & Structures, 2018, 195: 74-84.

[58] KIM W, REDDY J N. A new family of higher-order time integration algorithms for the analysis of structural dynamics[J]. International Journal for Numerical Methods in Engineering, 2017, 84: 071008.

[59] HOUBOLT J C. A recurrence matrix solution for the dynamic response of elastic aircraft[J]. Journal of the Aeronautical Sciences, 1950, 17(9): 540-550.

[60] BANK R E, COUGHRAN W M, FICHTNER W, et al. Transient simulations of silicon devices and circuits[J]. IEEE Transactions on Electro Devices, 1985, 32(10): 1992-2007.

[61] BATHE K J. Conserving energy and momentum in nonlinear dynamics: A simple implicit time integration scheme[J]. Computers & Structures, 2007, 85(7-8): 437-445.

[62] ZAKIAN P, BATHE K J. Transient wave propagations with the propagations with the Noh-Bathe scheme and the spectral element method[J]. Computers & Structures, 2021, 254: 106531.

[63] TAMMA K K, HAR J, ZHOU X M, et al. An overview and recent advances in vector and scalar formalisms: Space/time discretizations in computational dynamics-A unified approach[J]. Archives of Computational Methods in Engineering, 2011, 18: 119-283.

[64] 邢誉峰,张慧敏,季奕. 动力学常微分方程的时间积分方法[M]. 北京:科学出版社,2022.
XING Y F, ZHANG H M, JI Y. Time integration methods in dynamic ordinary differential equations[M]. Beijing: Science Press, 2022(in Chinese).

[65] ZHONG W X, WILLIAMS F W. A precise time step integration method[J]. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 1994, 208(6): 427-430.

[66] XING Y F, ZHANG H M, WANG Z K. Highly precise time integration method for linear structural dynamic analysis[J]. International Journal for Numerical Methods in Engineering, 2018, 116(8): 505-529.

[67] JI Y, XING Y F. Highly precise and efficient solution strategy for linear heat conduction and structural dynamics[J]. International Journal for Numerical Methods in Engineering, 2022, 123(2): 366-395.

[68] 邢誉峰,郭静. 与结构动特性协同的自适应 Newmark 方法[J]. 力学学报,2012,44(5): 904-911.

XING Y F, GUO J. A self-adaptive Newmark method with parameters dependent upon structural dynamic characteristics[J] Chinese Journal of Theoretical and Applied Mechanics, 2012, 44(5): 904-911(in Chinese).

[69] 邢誉峰,杨蓉. 单步辛算法的相位误差分析及修正[J]. 力学学报,2007,39(5):668-671.
XING Y F, YANG R. Phase errors and their correction in symplectic implicit single-step algorithm[J]. Chinese Journal of Theoretical and Applied Mechanics 2007, 39(5): 668-671(in Chinese).

[70] FENG K, QIN M Z. Hamiltonian algorithm for Hamiltonian dynamical systems[J]. Progress in Natural Science, 1991, 1(2): 105-116.

Advances and challenges in time integration methods

XING Yufeng[1,*], JI Yi[2], ZHANG Huimin[3]

(1. Institute of Solid Mechanics, Beihang University, Beijing 100083, China;
2. School of Aerospace Engineering, Beijing Institute of Technology, Beijing 100081, China;
3. Beijing Institute of Astronautical Systems Engineering, Beijing 100076, China)

Abstract: Time integration methods are a powerful tool for solving transient responses, which have been widely used to solve dynamic problems in aerospace, civil engineering, machinery manufacturing, and other fields. This paper reviews the advances in time integration methods in the past decades. Firstly, some classical methods, such as the series expansion method, the Runge-Kutta method and the Newmark method, are introduced. To solve the drawbacks involved in the classical methods, several time integration methods with more desirable numerical properties, including accuracy, efficiency, dissipation and stability, have been developed. Moreover, the advanced methods, including parameters methods, higher-order unconditionally stable methods, energy-conserving methods, linear multistep methods, composite methods and BN-stable methods, are introduced in this paper. Finally, the numerical properties and application scope of the existing time integration methods are compared, and some issues worthy of attention are given.

Keywords: dynamic systems; time integration method; accuracy; efficiency; dissipation; stability

Received: 2022-04-28; **Accepted**: 2022-05-18; **Published online**: 2022-06-07 10:36
URL: kns.cnki.net/kcms/detail/11.2625.V.20220606.1607.001.html
Foundation items: National Natural Science Foundation of China (12172023, 11872090)
* **Corresponding author**. E-mail: xingyf@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0140

旋转部件 SGBEM-FEM 耦合热弹性断裂分析

董雷霆*, 贺双新
(北京航空航天大学 航空科学与工程学院, 北京 100083)

摘 要: 旋转部件载荷复杂,易萌生裂纹,引发疲劳断裂失效。针对含裂纹受热旋转部件,在全局无裂纹区域使用有限元法(FEM),发挥其在大型结构仿真分析方面的优势;在裂纹附近局部子域使用对称伽辽金边界单元法(SGBEM),发挥其在结构断裂分析方面的优势。构造考虑旋转惯性载荷和热载荷影响的 SGBEM 超单元,该超单元刚度矩阵对称半正定,且施加的旋转惯性载荷和热载荷仅影响等效力向量,可实现与有限元刚度矩阵和力向量的直接装配,从而耦合求解旋转部件的热弹性断裂问题。采用构造的 SGBEM 超单元 FEM 耦合法计算了含裂纹受热旋转圆盘的应力强度因子,验证了该方法的有效性。

关 键 词: 对称伽辽金边界单元法(SGBEM); 超单元; 热弹性; 旋转惯性载荷; 应力强度因子
中图分类号: V214.1; TB125
文献标志码: A **文章编号**: 1001-5965(2022)09-1702-08

涡轮盘和转子叶片等旋转部件的疲劳断裂是航空发动机的常见多发故障[1-3],严重影响航空发动机和飞机的安全性与可靠性。旋转部件的高效仿真分析,特别是应力强度因子等断裂参数的精确快速计算,对飞机和航空发动机的结构仿真分析和结构完整性评估具有重要意义[4-5]。

有限元法(finite element method, FEM)[6]广泛应用于大型复杂结构的仿真分析,已成为当今工程技术领域应用最为广泛、成效最为显著的数值分析方法。FEM 采用连续函数作为形函数,然而在裂纹几何面处的位移场具有不连续性,并且在裂纹尖端处的应力场具有奇异性[7],因此 FEM 在处理裂纹这种强不连续问题方面具有难度,需要在裂纹尖端的高应力区划分致密的网格[8],这增加了问题求解规模,且限制了断裂力学参数的求解精度。

对称伽辽金边界单元法(symmetric Galerkin boundary element method, SGBEM)[9-10]在结构断裂仿真分析方面具有优势。对于三维断裂力学问题,由于 SGBEM 无需内部网格,仅在裂纹面划分网格,相比有限元,大大降低了裂纹尖端应力集中区域网格划分难度;其面力边界积分方程中的检验函数选取为位移场[11],可以精确计算裂纹张开位移,配以裂尖 1/4 结点单元后[12],SGBEM 可得到很高的应力强度因子计算精度[13];并且 SGBEM 的位移和面力边界积分方程是弱奇异的[14],这降低了数值积分计算难度。但由于 SGBEM 的系数矩阵是满阵[15],随着自由度数目的增加,会导致计算规模快速增加,因此难以用于求解复杂结构。

SGBEM 超单元 FEM 耦合法[16]结合了 FEM 求解大型复杂结构的优势和 SGBEM 求解断裂问题的优势。该方法在裂纹附近局部子域使用 SGBEM 求解,构造了 SGBEM 超单元;在全局无裂纹区域使用 FEM 求解;由于SGBEM超单元对称且半正定,可直接装配至全局有限元的刚度矩阵中,从

收稿日期: 2022-03-11; **录用日期**: 2022-04-08; **网络出版时间**: 2022-04-20 15:47
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220419.1306.002.html
基金项目: 航空科学基金 (201909051001)
* **通信作者**. E-mail: ltdong@buaa.edu.cn

引用格式: 董雷霆, 贺双新. 旋转部件 SGBEM-FEM 耦合热弹性断裂分析[J]. 北京航空航天大学学报, 2022, 48(9): 1702-1709.
DONG L T, HE S X. SGBEM-FEM coupling for thermoelastic fracture mechanics analysis of rotational components[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1702-1709(in Chinese).

而耦合求解含裂纹的复杂结构。该方法对飞机二维结构如加筋板裂纹黏结修复、耳片接头裂纹疲劳失效等实际问题可进行高效的仿真分析。但该方法没有考虑旋转惯性载荷和热载荷对含裂纹局部子域产生的影响,不能直接用于三维旋转部件的热弹性断裂分析。

本文将 SGBEM 超单元 FEM 耦合法推广至三维旋转部件的热弹性断裂问题。推导了三维情况下考虑热载荷和旋转惯性载荷的 SGBEM 超单元,将 SGBEM 超单元应用于含裂纹局部子域,并给出 SGBEM 超单元 FEM 耦合的过程,计算算例验证了该方法。

1　旋转部件热弹性 SGBEM 超单元

1.1　位移和面力边界积分方程

如图 1 所示,设弹性体 Ω 内存在体力 $\boldsymbol{f}$ 和温度场 T,源点 $\boldsymbol{\xi}$ 和场点 $\boldsymbol{x}$ 均位于弹性体 Ω 的边界 $\partial\Omega$上。伽辽金式位移边界积分方程[17-18]为

$$\begin{aligned}&\int_{\partial\Omega}\frac{1}{2}\delta t_p(\boldsymbol{\xi})u_p(\boldsymbol{\xi})\mathrm{d}S_{\xi}+\\&\int_{\partial\Omega}\delta t_p(\boldsymbol{\xi})\mathrm{d}S_{\xi}\int_{\partial\Omega}F_p(\boldsymbol{x}-\boldsymbol{\xi})\mathrm{d}\Gamma_x=\\&\int_{\partial\Omega}\delta t_p(\boldsymbol{\xi})\mathrm{d}S_{\xi}\int_{\partial\Omega}t_j(\boldsymbol{x})u_j^{*p}(\boldsymbol{x}-\boldsymbol{\xi})\mathrm{d}\Gamma_x+\\&\int_{\partial\Omega}\delta t_p(\boldsymbol{\xi})\mathrm{d}S_{\xi}\int_{\partial\Omega}D_iu_j(\boldsymbol{x})G_{ij}^{*p}(\boldsymbol{x}-\boldsymbol{\xi})\mathrm{d}\Gamma_x+\\&\int_{\partial\Omega}\delta t_p(\boldsymbol{\xi})\mathrm{d}S_{\xi}\int_{\partial\Omega}n_i(\boldsymbol{x})u_j(\boldsymbol{x})\phi_{ij}^{*p}(\boldsymbol{x}-\boldsymbol{\xi})\mathrm{d}\Gamma_x\end{aligned}\tag{1}$$

式中:重复的下标表示求和;$u_p(\boldsymbol{\xi})$ 为源点 $\boldsymbol{\xi}$ 处沿坐标 p 方向位移;$t_p(\boldsymbol{\xi})$ 为源点 $\boldsymbol{\xi}$ 处沿坐标 p 方向面力;$\delta t_p(\boldsymbol{\xi})$ 为边界上与面力具有相同插值空间的任意函数(即虚面力);函数 $F_p(\boldsymbol{x}-\boldsymbol{\xi})$ 为旋转惯性载荷和热载荷在边界积分中的表达式,其具体形式见附录 B;函数 $u_j^{*p}(\boldsymbol{x}-\boldsymbol{\xi})$ 为开尔文(Kelvin)问题的位移基本解,具体形式见附录 A 中式(A1),本文中所有与开尔文位移基本解相关的函数均在上标加“ * ”;函数 $G_{ij}^{*p}(\boldsymbol{x}-\boldsymbol{\xi})$ 见附录 A 中式(A5);函数 $\phi_{ij}^{*p}(\boldsymbol{x}-\boldsymbol{\xi})$ 见附录 A 中式(A6);$n_i(\boldsymbol{x})$ 为场点 $\boldsymbol{x}$ 处单位外法线方向向量分量。

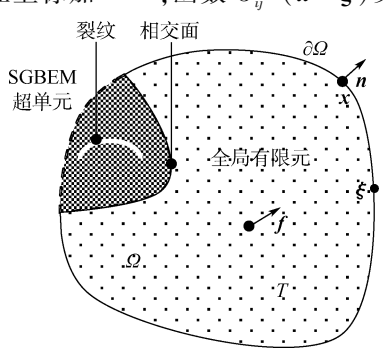


图 1　热弹性体

Fig. 1　Thermoelastic body

伽辽金式面力边界积分方程[17-18]为

$$\begin{aligned}&-\int_{\partial\Omega}\frac{1}{2}\delta u_b(\boldsymbol{\xi})t_b(\boldsymbol{\xi})\mathrm{d}S_{\xi}+\int_{\partial\Omega}A(\boldsymbol{\xi})\mathrm{d}S_{\xi}+\\&\int_{\partial\Omega}n_a(\boldsymbol{\xi})\delta u_b(\boldsymbol{\xi})\mathrm{d}S_{\xi}\int_{\partial\Omega}B_{ab}(\boldsymbol{x}-\boldsymbol{\xi})\mathrm{d}\Gamma_x+\\&\int_{\partial\Omega}D_t\delta u_b(\boldsymbol{\xi})\mathrm{d}S_{\xi}\int_{\partial\Omega}C_{tb}(\boldsymbol{x}-\boldsymbol{\xi})\mathrm{d}\Gamma_x=\\&\int_{\partial\Omega}D_a\delta u_b(\boldsymbol{\xi})\mathrm{d}S_{\xi}\int_{\partial\Omega}t_j(\boldsymbol{x})G_{ab}^{*j}(\boldsymbol{x}-\boldsymbol{\xi})\mathrm{d}\Gamma_x-\\&\int_{\partial\Omega}n_a\delta u_b(\boldsymbol{\xi})\mathrm{d}S_{\xi}\int_{\partial\Omega}t_j(\boldsymbol{x})\phi_{ab}^{*j}(\boldsymbol{x}-\boldsymbol{\xi})\mathrm{d}\Gamma_x+\\&\int_{\partial\Omega}D_a\delta u_b(\boldsymbol{\xi})\mathrm{d}S_{\xi}\int_{\partial\Omega}D_pu_q(\boldsymbol{x})H_{abpq}^{*}(\boldsymbol{x}-\boldsymbol{\xi})\mathrm{d}\Gamma_x\end{aligned}\tag{2}$$

式中:$\delta u_b(\boldsymbol{\xi})$ 为边界上与位移具有相同插值空间的任意函数(即虚位移);函数 $A(\boldsymbol{x})$,$B_{ab}(\boldsymbol{x}-\boldsymbol{\xi})$,$C_{tb}(\boldsymbol{x}-\boldsymbol{\xi})$ 为域积分中的旋转惯性载荷和热载荷在边界积分中的表达式,其具体形式见附录 B;函数 $H_{abpq}^{*}(\boldsymbol{x}-\boldsymbol{\xi})$ 见附录 A 中式(A7);D_a 为切向微分算符,见附录 A 中式(A8)。

伽辽金式位移边界积分方程式(1)是将传统边界元法中的位移边界积分方程写为弱形式,并对核函数使用亥姆霍兹(Helmholtz)分解进行弱奇异化后得到的;伽辽金式面力边界积分方程式(2)是将强奇异的应力边界积分方程写为弱形式,并对核函数使用亥姆霍兹分解进行弱奇异化后得到的。将伽辽金式位移和面力边界积分方程在边界网格上离散,再求解相应的线性方程组即可得到边界上的位移和面力,从而可以求解裂纹的张开位移,并使用位移外推法获得裂纹应力强度因子。

1.2　SGBEM 超单元构造格式

将伽辽金式位移边界积分方程式(1)和伽辽金式面力边界积分方程式(2)离散。为简化推导过程,本文不列出一个边界单元系数矩阵组装到对称伽辽金边界单元法中系数矩阵的过程。不失一般性,假设边界上位移和面力均为未知的。

边界上位移可离散为

$$u_p=N^mu_p^m\tag{3}$$

式中:$N^m=N^m(\boldsymbol{x})$ 为边界单元上第 m 个结点处的形函数 $N(\boldsymbol{x})$;$u_p^m=u_p^m(\boldsymbol{x})$ 为边界单元上第 m 个结点处沿 p 方向位移,本文也用 u_p^m 表示该位移在 SGBEM 中的自由度编号。

边界上面力可离散为

$$t_p=M^nt_p^n\tag{4}$$

式中:$M^n = M^n(\boldsymbol{x})$为边界单元上第 n 个结点形函数;$t_p^n = t_p^n(\boldsymbol{x})$为边界单元上第 n 个结点处沿 p 方向位移,本文也用 t_p^n 表示该面力在 SGBEM 中的自由度编号。

伽辽金式位移边界积分方程式(1)可离散为

$$\frac{1}{2}\delta \boldsymbol{t}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{u} + \delta \boldsymbol{t}^{\mathrm{T}}\boldsymbol{G}\boldsymbol{f} = \delta \boldsymbol{t}^{\mathrm{T}}\boldsymbol{U}\boldsymbol{t} + \delta \boldsymbol{t}^{\mathrm{T}}\boldsymbol{V}\boldsymbol{u} \tag{5}$$

式中:向量 $\boldsymbol{t}$ 为面力自由度 t_p^n 组成的列向量;向量 $\boldsymbol{u}$ 为位移自由度 u_p^m 组成的列向量;向量 $\boldsymbol{f}$ 为边界单元上结点处的旋转惯性载荷或温度场组成的列向量,为已知量;矩阵 $\boldsymbol{M}$ 的元素为式(1)第 1 项离散后形函数在单元上的积分;矩阵 $\boldsymbol{G}$ 的元素为式(1)第 2 项离散后函数 $F_p(\boldsymbol{x}-\boldsymbol{\xi})$和形函数之积在单元上的积分;矩阵 $\boldsymbol{U}$ 的元素为式(2)第 3 项离散后函数 $u_j^{*p}(\boldsymbol{x}-\boldsymbol{\xi})$和形函数之积在单元上的积分;矩阵 $\boldsymbol{V}$ 的元素为式(1)第 4、5 项离散后 2 个函数 $G_{ij}^{*p}(\boldsymbol{x}-\boldsymbol{\xi})$和 $\phi_{ij}^{*p}(\boldsymbol{x}-\boldsymbol{\xi})$分别与形函数求积后结果的和在单元上的积分。

伽辽金式面力边界积分方程式(2)可离散为

$$-\frac{1}{2}\delta \boldsymbol{u}^{\mathrm{T}}\boldsymbol{N}\boldsymbol{t} + \delta \boldsymbol{u}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{f} = \delta \boldsymbol{u}^{\mathrm{T}}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{t} + \delta \boldsymbol{u}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{u} \tag{6}$$

式中:矩阵 $\boldsymbol{N}$ 的元素为式(2)第 1 项离散后形函数在单元上的积分;矩阵 $\boldsymbol{H}$ 的元素为式(2)第 2、3、4 项离散后函数 $A(\boldsymbol{x})$、$B_{ab}(\boldsymbol{x}-\boldsymbol{\xi})$、$C_{tb}(\boldsymbol{x}-\boldsymbol{\xi})$分别与形函数求积后结果的和在单元上的积分;矩阵 $\boldsymbol{W}$ 的元素为式(2)最后 1 项离散后函数 $H_{abpq}^{*}(\boldsymbol{x}-\boldsymbol{\xi})$和形函数之积在单元上的积分。观察式(1)、式(2)可知,矩阵 $\boldsymbol{M}$ 和 $\boldsymbol{N}$ 互为转置矩阵,即 $\boldsymbol{M}=\boldsymbol{N}^{\mathrm{T}}$;矩阵 $\boldsymbol{U}$ 和 $\boldsymbol{W}$ 为对称矩阵,即 $\boldsymbol{U}=\boldsymbol{U}^{\mathrm{T}}$, $\boldsymbol{W}=\boldsymbol{W}^{\mathrm{T}}$。

离散的位移边界积分方程式(5)和离散的面力边界积分方程式(6)均含有位移自由度和面力自由度。但位移有限元中仅出现位移自由度,且在断裂力学仿真中更关心裂纹面张开位移。因此,可以将式(5)和式(6)中的面力自由度视为内部自由度缩聚掉,只保留位移自由度。

离散的位移边界积分方程式(5)可改写为

$$\boldsymbol{t} = \boldsymbol{U}^{-1}\left(\frac{1}{2}\boldsymbol{M} - \boldsymbol{V}\right)\boldsymbol{u} + \boldsymbol{U}^{-1}\boldsymbol{G}\boldsymbol{f} \tag{7}$$

离散的面力边界积分方程式(6)可改写为

$$-\delta \boldsymbol{u}^{\mathrm{T}}\boldsymbol{N}\boldsymbol{t} + \delta \boldsymbol{u}^{\mathrm{T}}\boldsymbol{H}\boldsymbol{f} = \delta \boldsymbol{u}^{\mathrm{T}}\left(\boldsymbol{V}^{\mathrm{T}} - \frac{1}{2}\boldsymbol{N}\right)\boldsymbol{t} + \delta \boldsymbol{u}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{u} \tag{8}$$

综合式(7)和式(8),可得 SGBEM 超单元的构造格式为

$$\delta \boldsymbol{u}^{\mathrm{T}}\bar{\boldsymbol{t}} + \delta \boldsymbol{u}^{\mathrm{T}}\bar{\boldsymbol{f}} = \delta \boldsymbol{u}^{\mathrm{T}}\bar{\boldsymbol{K}}\boldsymbol{u} \tag{9}$$

式中:

$$\bar{\boldsymbol{t}} = \boldsymbol{N}\boldsymbol{t} \tag{10}$$

$$\bar{\boldsymbol{f}} = \left[\left(\boldsymbol{V}^{\mathrm{T}} - \frac{1}{2}\boldsymbol{N}\right)\boldsymbol{U}^{-1}\boldsymbol{G} - \boldsymbol{H}\right]\boldsymbol{f} \tag{11}$$

$$\bar{\boldsymbol{K}} = \left(\boldsymbol{V}^{\mathrm{T}} - \frac{1}{2}\boldsymbol{N}\right)\boldsymbol{U}^{-1}\left(\boldsymbol{V} - \frac{1}{2}\boldsymbol{M}\right) - \boldsymbol{W} \tag{12}$$

注意到,由于 $\boldsymbol{M}=\boldsymbol{N}^{\mathrm{T}}$, $\boldsymbol{U}=\boldsymbol{U}^{\mathrm{T}}$, $\boldsymbol{W}=\boldsymbol{W}^{\mathrm{T}}$,则系数矩阵$\bar{\boldsymbol{K}}$为对称矩阵,即 $\bar{\boldsymbol{K}}=\bar{\boldsymbol{K}}^{\mathrm{T}}$。并且经数值算例表明,SGBEM 超单元系数矩阵 $\bar{\boldsymbol{K}}$ 是半正定的,且仅有刚体运动为其零特征向量。

1.3 含裂纹 SGBEM 超单元

当弹性体存在裂纹时,如图 1 所示,假设一个裂纹的 2 个几何曲面 S_c^+ 和 S_c^- 形状相同,在裂纹面 S_c^+ 和 S_c^- 上同一点 $\boldsymbol{x}$ 处的张开位移为

$$\Delta u_i(\boldsymbol{x}) = u_i^+(\boldsymbol{x}) - u_i^-(\boldsymbol{x}) \tag{13}$$

添加此约束条件后,无需将 2 个裂纹面分别划分网格,仅需在一个裂纹面划分网格即可,可直接计算裂纹面张开位移。

式(13)代入 SGBEM 超单元的构造格式(9),即得含裂纹弹性体的 SGBEM 超单元构造格式。

2 SGBEM 超单元 FEM 耦合法

2.1 SGBEM 超单元 FEM 耦合构造格式

如图 1 所示,局部子域和全局有限元区域相交面上有重合的面单元和重合的结点。

位移有限元的表达格式可写为

$$\begin{bmatrix}\delta \boldsymbol{u}_{\mathrm{G}}\\ \delta \boldsymbol{u}_{\mathrm{I}}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{t}_{\mathrm{G}}\\ \boldsymbol{t}_{\mathrm{I}}\end{bmatrix} + \begin{bmatrix}\delta \boldsymbol{u}_{\mathrm{G}}\\ \delta \boldsymbol{u}_{\mathrm{I}}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{f}_{\mathrm{G}}\\ \boldsymbol{f}_{\mathrm{I}}\end{bmatrix} = \begin{bmatrix}\delta \boldsymbol{u}_{\mathrm{G}}\\ \delta \boldsymbol{u}_{\mathrm{I}}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{K}_{\mathrm{GG}} & \boldsymbol{K}_{\mathrm{GI}}\\ \boldsymbol{K}_{\mathrm{IG}} & \boldsymbol{K}_{\mathrm{II}}\end{bmatrix}\begin{bmatrix}\boldsymbol{u}_{\mathrm{G}}\\ \boldsymbol{u}_{\mathrm{I}}\end{bmatrix} \tag{14}$$

式中:仅在全局(global)有限元中出现的结点,其位移 $\boldsymbol{u}$、面力 $\boldsymbol{t}$ 和相关有限元刚度矩阵 $\boldsymbol{K}$ 标注下标 G;仅在局部(local)子域中出现的结点,其位移 $\boldsymbol{u}$、面力 $\bar{\boldsymbol{t}}$ 和相关 SGBEM 超单元刚度矩阵$\bar{\boldsymbol{K}}$标注下标 L;在全局与局部相交面(interface)上的结点,其位移、面力和相关的有限元刚度矩阵以及 SGBEM 超单元刚度矩阵标注下标 I。

全局有限元中内部结点不需要考虑边界条件。这里为了简化起见,不再详细区分有限元的内部结点和边界结点。

SGBEM 超单元式(9)可改写为

$$\begin{bmatrix}\delta \boldsymbol{u}_{\mathrm{I}}\\ \delta \boldsymbol{u}_{\mathrm{L}}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\bar{\boldsymbol{t}}_{\mathrm{I}}\\ \bar{\boldsymbol{t}}_{\mathrm{L}}\end{bmatrix} + \begin{bmatrix}\delta \boldsymbol{u}_{\mathrm{I}}\\ \delta \boldsymbol{u}_{\mathrm{L}}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\bar{\boldsymbol{f}}_{\mathrm{I}}\\ \bar{\boldsymbol{f}}_{\mathrm{L}}\end{bmatrix} = \begin{bmatrix}\delta \boldsymbol{u}_{\mathrm{I}}\\ \delta \boldsymbol{u}_{\mathrm{L}}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\bar{\boldsymbol{K}}_{\mathrm{II}} & \bar{\boldsymbol{K}}_{\mathrm{IL}}\\ \bar{\boldsymbol{K}}_{\mathrm{LI}} & \bar{\boldsymbol{K}}_{\mathrm{LL}}\end{bmatrix}\begin{bmatrix}\boldsymbol{u}_{\mathrm{I}}\\ \boldsymbol{u}_{\mathrm{L}}\end{bmatrix} \tag{15}$$

分别装配式(14)、式(15)的刚度矩阵和等效结点力向量，可得 SGBEM 超单元 FEM 耦合构造格式：

$$\begin{bmatrix}\delta\boldsymbol{u}_{\mathrm{G}}\\ \delta\boldsymbol{u}_{\mathrm{L}}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{t}_{\mathrm{G}}\\ \bar{\boldsymbol{t}}_{\mathrm{L}}\end{bmatrix}+\begin{bmatrix}\delta\boldsymbol{u}_{\mathrm{G}}\\ \delta\boldsymbol{u}_{\mathrm{I}}\\ \delta\boldsymbol{u}_{\mathrm{L}}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{f}_{\mathrm{G}}\\ \boldsymbol{f}_{\mathrm{I}}+\bar{\boldsymbol{f}}_{\mathrm{I}}\\ \bar{\boldsymbol{f}}_{\mathrm{L}}\end{bmatrix}=$$

$$\begin{bmatrix}\delta\boldsymbol{u}_{\mathrm{G}}\\ \delta\boldsymbol{u}_{\mathrm{I}}\\ \delta\boldsymbol{u}_{\mathrm{L}}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{K}_{\mathrm{GG}} & \boldsymbol{K}_{\mathrm{GI}} & \boldsymbol{O}\\ \boldsymbol{K}_{\mathrm{IG}} & \boldsymbol{K}_{\mathrm{II}}+\bar{\boldsymbol{K}}_{\mathrm{II}} & \bar{\boldsymbol{K}}_{\mathrm{IL}}\\ \boldsymbol{O} & \bar{\boldsymbol{K}}_{\mathrm{LI}} & \bar{\boldsymbol{K}}_{\mathrm{LL}}\end{bmatrix}\begin{bmatrix}\boldsymbol{u}_{\mathrm{G}}\\ \boldsymbol{u}_{\mathrm{I}}\\ \boldsymbol{u}_{\mathrm{L}}\end{bmatrix}\tag{16}$$

注意到，有限元区域与局部子域相交面上存在着相交结点，有限元在相交结点上的面力 $\boldsymbol{t}_{\mathrm{I}}$ 与 SGBEM 超单元在相交结点上的面力 $\bar{\boldsymbol{t}}_{\mathrm{I}}$ 大小相等，方向相反，即 $\boldsymbol{t}_{\mathrm{I}}=-\bar{\boldsymbol{t}}_{\mathrm{I}}$，故式(16)不显示相交结点上的面力。

2.2　SGBEM 超单元 FEM 耦合求解流程

SGBEM 超单元 FEM 耦合求解流程采用自动化的软件完成，用户仅需要独立划分有限元网格与裂纹面网格。如图 2 所示，详细耦合求解过程如下。

1）搜索局部子域。当旋转部件划分好有限元网格后，根据旋转部件中裂纹位置，找到裂纹附近有限元单元，这些有限元单元构成了局部子域有限元网格。

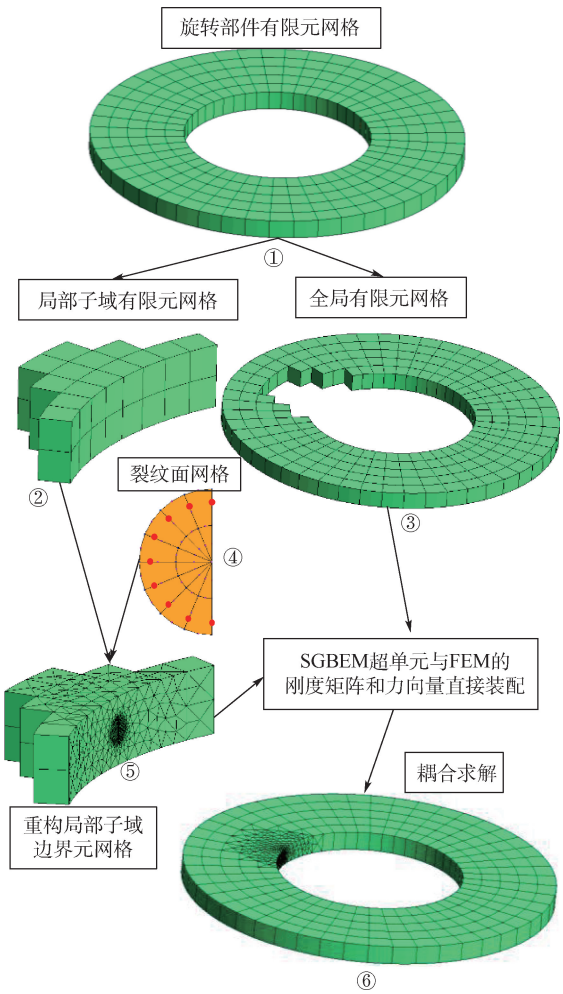


图 2　SGBEM 超单元 FEM 耦合求解流程

Fig. 2　Process of SGBEM super element and FEM coupling solution

2）消除内部结点。注意到，裂纹面网格是另外插入的，裂纹面网格结点与局部子域有限元网格结点没有任何关系。对于内部裂纹，需要将局部子域有限元网格中内部结点消去，仅保留边界面单元，这些边界面单元和裂纹面单元共同构成了局部子域边界网格，如图 3 所示，为清楚显示内部，该图隐藏了部分边界单元。

3）重构局部子域。对于面裂纹，以图 2 中④所示半圆形裂纹为例，其在局部子域中的位置如图 3所示。裂纹前沿(即半圆曲线)上的结点在局部子域内部，无需处理；但该裂纹的开口线(即半圆直径线)上的结点既在裂纹面上，也在局部子域边界面上。因此，需要重新划分局部子域边界面(见图 2 中⑤部分)。注意，局部子域和全局有限元相交面上的边界面单元不可以重构，因为相交面上的结点是局部子域和全局有限元共有的。

4）装配刚度矩阵。在获得局部子域边界元网格后，分别求解局部子域 SGBEM 超单元和全局有限元，获取 SGBEM 超单元的刚度矩阵和有限元的刚度矩阵，再通过结点编号直接装配。

5）施加载荷。根据式(9)等号左端第 2 项求解局部子域旋转惯性载荷和热载荷产生的等效载荷列向量，并求解全局有限元中的载荷列向量。再按照式(16)等号左端第 3 项将局部子域等效载荷列向量装配至全局有限元中的等效载荷列向量。

6）耦合求解。步骤 4)和步骤 5)给出局部子域 SGBEM 超单元和全局有限元的刚度矩阵和等效载荷列向量的整体装配形式。求解该线性方程组，可获得局部子域 SGBEM 超单元和全局有限元所有结点的位移解，包括裂纹张开位移，从而可以使用位移外推法求解裂纹应力强度因子。

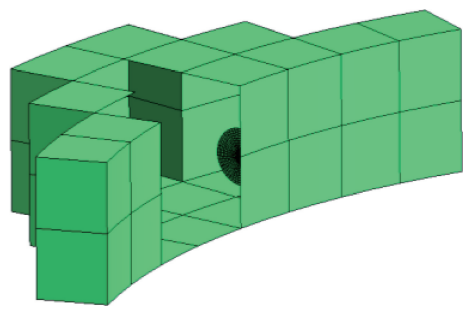


图 3　局部子域边界网格和裂纹网格

Fig. 3　Subdomain boundary mesh and crack mesh

3 含裂纹受热旋转圆盘算例

本节使用 SGBEM 超单元 FEM 耦合法求解图 2中的受热旋转圆盘,以展示其优点。

如图 4 所示,圆盘内径 $R_i = 0.1$ m,外径 $R_o = 0.2$ m,厚 $t = 0.02$ m,弹性模量 $E = 1\ 000$ MPa,泊松比 $\nu = 0.3$,线膨胀系数 $\alpha = 2.17 \times 10^{-5}/℃$。绕圆盘中心轴的转速为 ω,转速 ω 的取值如表 1 所示。沿径向的温度分布为

$$T_r(\boldsymbol{x}) = k(x^2 + y^2) \tag{17}$$

式中:系数 k 的取值如表 1。在圆盘内环面中间存在一半圆形面裂纹,半径为 $r = 0.004$ m(见图 3);在圆盘内环面中心处施加刚体位移约束。

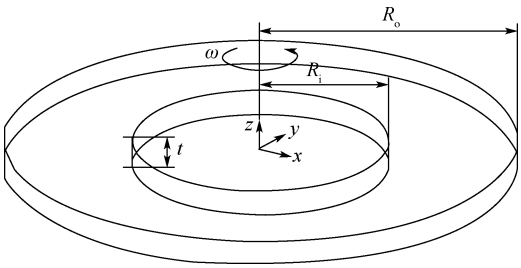


图 4 受热旋转圆盘

Fig. 4 Rotational disk undergoing thermal loading

该圆盘的 SGBEM 超单元 FEM 耦合法求解流程见 2.2 节。分别使用 SGBEM 超单元 FEM 耦合法和商业有限元软件 ANSYS Workbench 计算半圆形裂纹圆弧中点处的 I 型裂纹应力强度因子。ANSYS Workbench 所使用的有限元网格见(见图 5(a)),该网格比图 2 中⑥部分的耦合法所用的网格更加密集;且 ANSYS Workbench 在裂纹附近使用了多种体单元,有五面体棱柱单元、六面体单元、五面体金字塔单元以及四面体单元(见图 5(b))。可见,耦合方法使用了更加稀疏的网格,在裂纹尖端处仅使用了四边形单元,网格也更加简单(见图 2 中④部分)。

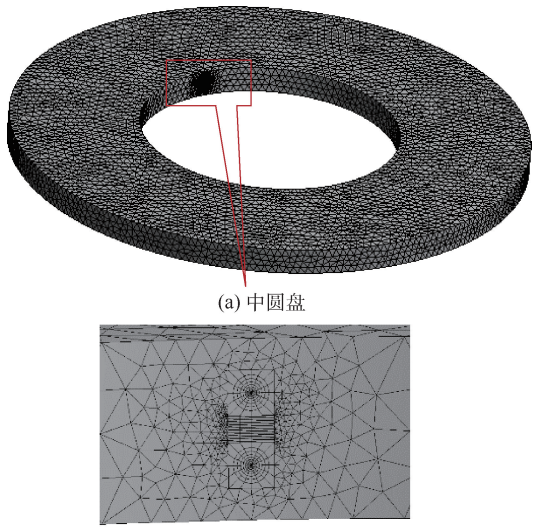
受热旋转圆盘 I 型裂纹应力强度因子计算结果如表 1 所示,不同转速和温度下,SGBEM 超单元 FEM 耦合法与 ANSYS Workbench 的计算结果相比,误差均小于 2%,可见 SGBEM 超单元 FEM 耦合法可有效用于旋转部件的热弹性断裂分析。

该算例中本文方法在不含裂纹的全局区域进行稀疏的有限元剖分;在裂纹附近仅使用二维单元进行边界元的计算,仅需要对裂纹面和包络裂纹的局部子域表面进行简单的二维网格划分,避免了裂纹附近细致的三维有限元网格剖分,可以显著减少单元数量。关于 SGBEM-FEM 法模拟断裂与疲劳问题的在网格剖分和计算效率方面的优势,请见文献[19-20]中给出的详细论述。

表 1 不同转速和温度下圆盘裂纹应力强度因子

Table 1 Stress intensity factors of crack in disk for different rotational speeds and temperature

转速/(10^3 r·min^{-1})	系数 $k/10^3$	应力强度因子/(MPa·m$^{\frac{1}{2}}$) SGBEM 超单元 FEM 耦合法	应力强度因子/(MPa·m$^{\frac{1}{2}}$) ANSYS Workbench	相对误差/%
6	6	3.652 81	3.719 6	-1.80
7	7	4.792 74	4.877 8	-1.74
8	8	6.084 43	6.189 6	-1.70
9	9	7.527 88	7.655 3	-1.66
10	10	9.120 92	9.276 6	-1.68



(a) 中圆盘

(b) 裂纹

图 5 ANSYS Workbench 中圆盘和裂纹的有限元网格

Fig. 5 Disk and crack FEM mesh used by ANSYS Workbench

4 结 论

本文给出了三维旋转部件 SGBEM 超单元的具体构造格式。本文将 SGBEM 超单元与有限元耦合,计算了旋转圆盘在热载荷作用下半圆形裂纹的应力强度因子,并与商业有限元软件 ANSYS Workbench 结果进行对比,对比结果表明了该 SGBEM 超单元 FEM 耦合法可有效用于旋转部件的热弹性断裂分析。通过 SGBEM 超单元 FEM 耦合求解流程可知:

1) 有限元网格独立划分,不受裂纹干扰。由于裂纹面不使用有限元单元,有限元网格可自由按需划分,无需考虑裂纹尖端应力集中区有限元网格加密的困难。

2) 裂纹面网格独立划分,难度降低。裂纹仅使用边界单元,可自由按需划分,且内部裂纹无需重构网格。虽然面裂纹存在网格重构问题,但仅需在局部子域边界曲面上划分二维网格,相比传

统有限元需在裂纹附近区域划分致密的三维网格,难度大大降低。

3) 耦合过程简单。由于 SGBEM 超单元仅存在位移自由度,不存在面力自由度,从 SGBEM 超单元 FEM 耦合构造格式可知,其与有限元的装配耦合过程与有限元中一个单元的刚度矩阵装配至整体刚度矩阵的过程完全一致。

4) 裂纹尖端网格简单。SGBEM 超单元 FEM,耦合法在裂纹前沿使用 1/4 结点面单元。而有限元需要使用多种形状的体单元划分不规则形状的裂纹前沿。

5) 有助于快速仿真裂纹扩展。当进行裂纹扩展仿真时,每当裂纹扩展一段距离,对于内部裂纹,耦合方法仅需在裂纹前沿增加一层边界面单元即可继续仿真;对于面裂纹,仅需再次重构局部子域裂纹附近处的边界面网格。但对有限元而言,需要重新划分裂纹附近区域体单元网格。

参考文献 (References)

[1] 陈予恕,张华彪. 航空发动机整机动力学研究进展与展望[J]. 航空学报,2011,32(8):1371-1391.
CHEN Y S,ZHANG H B. Review and prospect on the research of dynamics of complete aero-engine systems[J]. Acta Aeronautica et Astronautica Sinica,2011,32(8):1371-1391(in Chinese).

[2] 陈博,朱剑寒,鲁辉军. 航空发动机涡轮叶片高周疲劳裂纹故障分析与思考[J]. 燃气涡轮试验与研究,2020,33(3):33-36.
CHEN B,ZHU J H,LU H J. Analysis and consideration for turbine blade high-cycle fatigue crack failure of aero-engine[J]. Gas Turbine Experiment and Research,2020,33(3):33-36(in Chinese).

[3] 马利丽,何立强,任伟峰. 航空发动机自由涡轮叶片裂纹故障分析[J]. 航空发动机,2018,44(6):54-58.
MA L L,HE L Q,REN W F. Blade crack fault analysis of aero-engine free turbine[J]. Aeroengine,2018,44(6):54-58(in Chinese).

[4] 董雷霆,周轩,赵福斌,等. 飞机结构数字孪生关键建模仿真技术[J]. 航空学报,2021,42(3):023981.
DONG L T,ZHOU X,ZHAO F B,et al. Key technologies for modeling and simulation of airframe digital twin[J]. Acta Aeronautica et Astronautica Sinica,2021,42(3):023981(in Chinese).

[5] 李其汉. 航空发动机结构完整性研究进展[J]. 航空发动机,2014,40(5):1-6.
LI Q H. Investigation progress on aeroengine structural integrity[J]. Aeroengine,2014,40(5):1-6(in Chinese).

[6] 王勖成. 有限单元法[M]. 北京:清华大学出版社,2003:1-8.
WANG X C. Finite element method[M]. Beijing: Tsinghua University Press,2003:1-8(in Chinese).

[7] 段静波,李道奎,雷勇军. 断裂力学分析的奇异有限元法研究综述[J]. 机械强度,2012,34(2):262-269.
DUAN J B,LI D K,LEI Y J. Advances in singular finite element method for fracture mechanics analysis[J]. Journal of Mechanical Strength,2012,34(2):262-269(in Chinese).

[8] 徐杰,周迅,陈文华,等. 基于有限元法的表面疲劳裂纹扩展模拟[J]. 浙江理工大学学报,2012,29(1):66-69.
XU J,ZHOU X,CHEN W H,et al. Simulation of fatigue crack growth of surface cracked plates by finite element method[J]. Journal of Zhejiang Sci-Tech University,2012,29(1):66-69(in Chinese).

[9] BONNET M,MAIER G,POLIZZOTTO C. Symmetric Galerkin boundary element methods[J]. Applied Mechanics Reviews,1998,51(11):669-704.

[10] SUTRADHAR A,PAULINO G H,GRAY L J. Symmetric Galerkin boundary element method[M]. Berlin: Springer-Verlag,2008:171-196.

[11] LI S,MEAR M E,XIAO L. Symmetric weak-form integral equation method for three-dimensional fracture analysis[J]. Computer Methods in Applied Mechanics and Engineering,1998,151(3-4):435-459.

[12] NIKISHKOV G P. Accuracy of quarter-point element in modeling crack-tip fields[J]. Computer Modeling in Engineering and Sciences,2013,93(5):335-361.

[13] FRANGI A. Fracture propagation in 3D by the symmetric Galerkin boundary element method[J]. International Journal of Fracture,2002,116(4):313-330.

[14] FRANGI A. Regularization of boundary element formulations by the derivative transfer method[M]// BREBBIA C A,ALIABADI M H. Singular integrals in boundary element methods[M]. Southampton: Computational Mechanics Publications, 1998: 125-164.

[15] HAN Z,ATLURI S. SGBEM (for cracked local subdomain): FEM (for uncracked global structure) alternating method for analyzing 3D surface cracks and their fatigue-growth[J]. Computer Modeling in Engineering & Sciences,2002,3(6):699-716.

[16] DONG L T,ATLURI S N. SGBEM (using non-hyper-singular traction BIE), and super elements, for non-collinear fatigue-growth analyses of cracks in stiffened panels with composite-patch repairs[J]. Computer Modeling in Engineering & Sciences,2012,89(5):41-458.

[17] HAN Z,ATLURI S. On simple formulations of weakly-singular traction & displacement BIE,and their solutions through petrov-Galerkin approaches[J]. Computer Modeling in Engineering & Sciences,2003,4(1):5-20.

[18] HE S X,WANG C Y,ZHOU X,et al. Weakly singular symmetric Galerkin boundary element method for fracture analysis of three-dimensional structures considering rotational inertia and gravitational forces[J]. Computer Modeling in Engineering & Sciences,2022,131(3):1857-1882.

[19] DONG L T,ATLURI S N. Fracture & fatigue analyses:SGBEM-FEM or XFEM? Part 1:2D structures[J]. Computer Modeling in Engineering & Sciences,2013,90(2):91-146.

[20] DONG L T,ATLURI S N. Fracture & fatigue analyses:SGBEM-

FEM or XFEM? Part 2:3D solids[J]. Computer Modeling in Engineering & Sciences,2013,90(5):379-413.

附录A:

$$u_j^{*p}(\boldsymbol{x}-\boldsymbol{\xi}) = \frac{1}{16\pi G(1-\nu)}\frac{1}{r}[(3-4\nu)\delta_{pi}+r_{,p}r_{,i}] \tag{A1}$$

式中:r为源点$\boldsymbol{\xi}$到场点$\boldsymbol{x}$的距离,下标中的逗号表示求偏导,即$r_{,p}=\partial/\partial x_p$;$G$为剪切模量;$\nu$为泊松比;$\delta_{pi}$为Kronecker delta符号,即

$$\delta_{pi} = \begin{cases} 0 & p \neq i \\ 1 & p = i \end{cases} \tag{A2}$$

函数$\sum_{abqm}^{*}(\boldsymbol{x}-\boldsymbol{\xi})$为

$$\sum_{abqm}^{*}(\boldsymbol{x}-\boldsymbol{\xi}) = E_{abpk}e_{qnk}\sigma_{mn}^{*p}(\boldsymbol{x}-\boldsymbol{\xi}) \tag{A3}$$

式中:e_{qnk}为排列符号,表示为

$$\begin{cases} e_{123}=e_{231}=e_{312}=1 \\ e_{132}=e_{321}=e_{213}=-1 \\ e_{qnk}=0 \quad q,n,k\text{取其他值} \end{cases} \tag{A4}$$

$$G_{ij}^{*p}(\boldsymbol{x}-\boldsymbol{\xi}) = \frac{1}{8\pi(1-\nu)}\frac{1}{r}[(1-2\nu)e_{ipj}+e_{ikj}r_{,k}r_{,p}] \tag{A5}$$

$$\phi_{ij}^{*p}(\boldsymbol{x}-\boldsymbol{\xi}) = \frac{1}{4\pi}\cdot\frac{1}{r^2}\delta_{pj}r_{,i} \tag{A6}$$

$$H_{abpq}^{*}(\boldsymbol{x}-\boldsymbol{\xi}) = \frac{G}{8\pi(1-\nu)}\cdot\frac{1}{r}[4\nu\delta_{aq}\delta_{bp}-\delta_{ap}\delta_{bq}-2\nu\delta_{ab}\delta_{pq}+\delta_{ab}r_{,p}r_{,q}+\delta_{pq}r_{,a}r_{,b}-2\delta_{ap}r_{,b}r_{,q}-\delta_{bq}r_{,a}r_{,p}] \tag{A7}$$

切向微分算符为

$$D_a = n_i e_{ija}\frac{\partial}{\partial x_j} \tag{A8}$$

附录B:

当弹性体作用有旋转惯性载荷时,如果使用散度定理将旋转惯性载荷引起的域积分转换为边界积分,则伽辽金式位移和面力边界积分方程式(1)和式(2)中关于旋转惯性载荷的函数为

$$F_p(\boldsymbol{x}-\boldsymbol{\xi}) = -\frac{1}{8\pi G}h_{jm}\{\delta_{pj}n_k(\boldsymbol{x})(x_m r_{,k}-r\delta_{mk})-\frac{1}{2(1-\nu)}[n_p(\boldsymbol{x})x_m r_{,j}-r\delta_{mp}n_j(\boldsymbol{x})]\} \tag{B1}$$

$$A(\boldsymbol{\xi}) = 0 \tag{B2}$$

$$B_{ab}(\boldsymbol{x}-\boldsymbol{\xi}) = -\frac{1}{8\pi}h_{jm}\left\{\frac{\nu}{1-\nu}\delta_{ab}[n_j(\boldsymbol{x})x_m r_{,kk}-\delta_{mj}n_k(\boldsymbol{x})r_{,k}]+\delta_{ja}[n_b(\boldsymbol{x})x_m r_{,kk}-\delta_{mb}n_k(\boldsymbol{x})r_{,k}]+\delta_{jb}[n_a(\boldsymbol{x})x_m r_{,kk}-\delta_{ma}n_k(\boldsymbol{x})r_{,k}]-\frac{1}{1-\nu}[n_j(\boldsymbol{x})x_m r_{,ab}-\delta_{mj}n_b(\boldsymbol{x})r_{,a}]\right\} \tag{B3}$$

式中:h_{jm}的分量写为矩阵的形式,即

$$h_{jm} = \rho\begin{bmatrix} \omega_2^2+\omega_3^2 & -\omega_1\omega_2 & -\omega_3\omega_1 \\ -\omega_1\omega_2 & \omega_3^2+\omega_1^2 & -\omega_2\omega_3 \\ -\omega_3\omega_1 & -\omega_2\omega_3 & \omega_1^2+\omega_2^2 \end{bmatrix} \tag{B4}$$

其中:ρ为弹性体的密度;$\boldsymbol{\omega}=(\omega_1,\omega_2,\omega_3)$为弹性体绕定轴转动的转速。

$$C_{tb}(\boldsymbol{x}-\boldsymbol{\xi}) = 0 \tag{B5}$$

当弹性中存在温度场时,如果使用散度定理和径向积分法将温度场引起的域积分转换为边界积分,则伽辽金式位移和面力边界积分方程式(1)和式(2)中关于温度场的函数为

$$F_p(\boldsymbol{x}-\boldsymbol{\xi}) = -\frac{1-2\nu}{8\pi G(1-\nu)}\cdot\frac{1}{r^2}\cdot\frac{\partial r}{\partial n}r_{,p}\int_0^r\beta T(\boldsymbol{x})\mathrm{d}R \tag{B6}$$

$$A(\boldsymbol{\xi}) = \frac{1-2\nu}{2(1-\nu)}n_b(\boldsymbol{\xi})\delta u_b(\boldsymbol{\xi})\beta T(\boldsymbol{\xi}) \tag{B7}$$

$$B_{ab}(\boldsymbol{x}-\boldsymbol{\xi}) = \frac{1-2\nu}{4\pi(1-\nu)}\cdot\frac{1}{r^2}\cdot\frac{\partial r}{\partial n}(3r_{,a}r_{,b}-\delta_{ab})\cdot\int_0^r\beta[T(\boldsymbol{x})-T(\boldsymbol{\xi})]\frac{1}{R}\mathrm{d}R+\frac{1-2\nu}{8\pi(1-\nu)}\beta T(\boldsymbol{x})\cdot\frac{1}{r^2}(3r_{,a}r_{,b}r_{,m}-\delta_{mb}r_{,a}-2\delta_{ab}r_{,m})n_m(\boldsymbol{x}) \tag{B8}$$

$$C_{tb}(\boldsymbol{x}-\boldsymbol{\xi}) = \frac{1-2\nu}{8\pi(1-\nu)}\cdot\frac{1}{r}e_{mbt}n_m(\boldsymbol{x})\beta T(\boldsymbol{x}) \tag{B9}$$

SGBEM-FEM coupling for thermoelastic fracture mechanics analysis of rotational components

DONG Leiting*, HE Shuangxin

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

Abstract: Rotational components undergo complex loads which can easily initiate cracks, resulting in fatigue and fracture failures. For thermal-loaded rotational components with cracks, the finite element method (FEM) was applied to the global structure without cracks, thus exerting the advantage of FEM in numerical analysis of large-scale structure; meanwhile, the symmetric Galerkin boundary element method (SGBEM) was employed to the subdomain around the crack, thereby developing the advantage of SGBEM in fracture analysis. The SGBEM super element, which took the influence of the rotational inertia loading and the thermal loading into consideration, was developed in this paper. The stiffness matrix of the obtained SGBEM super element was symmetric and positive semidefinite, where the rotational inertia loading and the thermal loading only influence the equivalent force vector. Thus, the SGBEM super element can be directly assembled with the stiffness matrix and the equivalent force vector of the FEM, and be coupled with the FEM to solve the thermoelastic fracture problem of rotational components. Stress intensity factors of the cracked rotational disk undergoing the thermal loading were computed, which validates the effectiveness of SGBEM super element and FEM coupling method.

Keywords: symmetric Galerkin boundary element method (SGBEM); super element; thermoelasticity; rotational inertia load; stress intensity factor

Received: 2022-03-11; **Accepted**: 2022-04-08; **Published online**: 2022-04-20 15:47

URL: kns.cnki.net/kcms/detail/11.2625.V.20220419.1306.002.html

Foundation item: Aeronautical Science Foundation of China (201909051001)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0291

运动信息引导的目标检测算法

胡海苗*, 沈柳青, 高立崑, 李明竹
(北京航空航天大学 计算机学院, 北京 100083)

摘 要: 在室外监控视频的场景下,由于场景的复杂性及目标的多样性,监控视频中的目标存在难以检测的情况,如目标被遮挡、目标尺寸变化等,目标检测任务仍然存在挑战。基于此,提出了一种利用运动信息引导基于卷积神经网络的目标检测算法来提高目标检测的准确率。对运动目标检测算法进行一定的改进,使得到的运动前景图中能够保持静止目标前景的存在;利用运动前景图中的前景可以指示目标空间位置的特点,在特征层面将网络提取的特征图与获取的以运动前景图为主的运动信息相融合,提高特征图可能存在目标区域的响应值;在目标检测算法的检测器中,引入一个定位分支,利用视频帧的运动前景图,学习候选目标的定位置信度,并与目标的分类置信度加权求和,作为目标最终的置信度,再通过非极大值抑制方法得到检测结果。实验证明,在固定摄像机下采集的数据集中,所提算法能够提升目标检测的准确率。

关 键 词: 运动信息; 前景区域; 特征融合; 定位分支; 目标检测
中图分类号: TP391
文献标志码: A **文章编号**: 1001-5965(2022)09-1710-11

近年来,随着监控设备的不断更新换代,越来越多的便捷实惠的视频监控设备被大范围地应用于日常生活之中[1]。这些随处可见的监控设备不仅记录着人们的生活,而且保障着人们的安全。但现实监控场景是比较复杂的,场景中目标的尺寸、姿态也是多种多样,会出现被遮挡目标、小目标等容易漏检的困难目标,降低目标检测算法的准确率。因此,针对固定摄像机拍摄出来的监控视频进行目标检测依然是一项充满挑战的任务。

在目标检测算法的发展历程中,算法种类可以分为传统的目标检测算法和在卷积神经网络基础上发展起来的目标检测算法。

传统的目标检测算法又可以分为 2 类:一类是利用手工设计特征的检测算法;另一类是利用运动变换信息的检测算法。

对于利用手工设计特征的目标检测算法,都遵循着同样的流程,分别是区域选择、特征提取和分类器分类。首先是区域选择,也被称为候选区域提取,是为了从图像中生成大量的候选区域,用于下一步的特征提取。在特征提取阶段,一般使用手工设计的特征,也就是纹理、颜色等底层特征,如梯度方向直方图特征(histogram of oriented gridients, HOG)[2]、局部二值模式(local binary pattern, LBP)[3]、尺度不变特征变换(scale invariant feature transform, SIFT)[4],这些底层特征经过编码学习后,能够得到更深一层的特征,包括主成分分析(principal component analysis, PCA)特征、线性判别分析(linear discriminant analysis, LDA)特征等。总体来说,手工设计的特征在光照变化、遮挡等复杂场景下的鲁棒性较差,因此会导致目标检测的准确率不高。进入到分类阶段,使用训练过的分类器对提取的特征进行分类,得到目标

收稿日期: 2022-04-28; **录用日期**: 2022-05-13; **网络出版时间**: 2022-05-17 14:33
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220517.1059.001.html
基金项目: 国家自然科学基金(62122011,U21A20514); 浙江省重点研发项目(2022C01082)
* **通信作者**. E-mail: frank0139@163.com

引用格式: *胡海苗, 沈柳青, 高立崑, 等. 运动信息引导的目标检测算法[J]. 北京航空航天大学学报, 2022, 48(9): 1710-1720.*
HU H M, SHEN L Q, GAO L K, et al. Object detection algorithm guided by motion information[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1710-1720(in Chinese).

检测的结果。另一种传统的目标检测算法是利用运动变换信息对目标进行检测。具体来说,是利用运动目标检测算法,如帧间差分法[5]、背景差分法[6-8]等,生成运动前景图,从而得到运动目标的位置。

基于卷积神经网络的目标检测算法可以分为单阶段算法和两阶段算法,不同在于是否有明确的感兴趣区域(region of interest,ROI)提取步骤。在单阶段算法中,会将整幅图像划分为小格子,在小格子上按照预先设定的锚框直接对图像进行分类和回归,典型算法有 SSD 系列算法[9]、YOLO 系列算法[10-13]、RetinaNet 算法[14]。两阶段算法通常会使用区域建议网络(region proposal network,RPN),通过预先设定的锚框,提取出 ROI,再通过全连接网络针对提取出的 ROI 进行分类和回归操作,典型算法有 R-CNN[15]、Fast R-CNN[16]、Faster R-CNN[17]、Cascade R-CNN[18]。相比较而言,单阶段检测器时间效率更高,但两阶段检测器的准确率更高。

基于卷积神经网络的目标检测算法的性能很大程度上依赖于训练时所使用的数据集,而在大部分数据集中,困难目标的数量较少,导致模型分布不均衡,无法对该类目标有充分的学习;同时,困难目标由于特征不显著,很容易被检测器忽略,造成漏检;不仅如此,基于卷积神经网络的目标检测算法还要求对数据要有较高质量的标注,如果要对困难目标进行标注,需要花费较多的人力,对数据集中的目标进行仔细的辨认和精准的标注,代价比较昂贵,甚至在标注时很可能会引入干扰,反而降低了检测器的性能。除此之外,现有的基于卷积神经网络的目标检测算法对候选目标框置信度的预测考虑不够全面,只考虑到了分类置信度,即目标属于该类别的概率。但是,分类置信度高并不能说明该候选目标框的位置也是十分精准的,只考虑分类置信度,可能会导致定位精准的候选目标框被筛除,影响目标检测算法的准确率。

传统目标检测算法中的利用运动变换信息的目标检测算法[5-8],可以快速发现困难目标,并且运动前景图中的前景位置能够体现目标的空间位置信息。而在室外监控场景下,大部分的目标都是能提取到运动信息的,同时,处于运动状态的目标是更应该关注的目标。因此,可以考虑将运动信息与目标检测算法相结合,提升目标检测的准确性。但是运动目标检测算法的缺点也是显而易见的,该类算法主要针对的是运动目标的检测,因此,在目标静止或运动幅度不大时,很难或不能完整提取出目标区域。

本文提出了一种运动信息引导的目标检测算法主要创新点如下:①提出一种改进的运动目标检测算法,在 ViBe 背景建模算法[7]的基础上进行了改进,改进后的算法能够减少静止目标融入背景的情况;②提出一种基于运动信息的多尺度特征融合模型,将以运动前景图为主的运动信息作为空间权重与特征金字塔网络提取的特征图相融合;③提出一种运动前景图指导的定位分支模块,根据前景的位置信息学习预测候选目标的定位置信度,作为目标分类置信度的补充。

1　相关工作

目前来说,运动信息在目标检测算法中的应用分为 2 种:①传统目标检测算法中利用运动变换信息的检测算法;②将运动信息(主要是光流信息)引入到卷积神经网络中,简化卷积神经网络特征的提取或进行多帧特征的融合。

深度特征流(deep feature flow, DFF)[19]方法首次将光流信息引入到了目标检测网络中,Zhu 等发现,在基于卷积神经网络的目标检测算法中,最耗时的步骤是特征提取,并且对于一个连续的视频序列来说,相邻帧提取的特征差别并不大,因此,在卷积神经网络中引入了光流信息来预测特征图,减少神经网络在提取特征方面的时间开销。DFF 网络中,将视频帧分为关键帧和非关键帧。卷积神经网络对关键帧提取特征;而对非关键帧,会与关键帧一起输入到光流网络中,得到光流信息,再根据光流信息对关键帧的特征进行修正,作为非关键帧的特征进行后续的分类回归。

光流引导的特征融合(flow guild feature aggregation, FGFA)[20]网络在 DFF 网络的基础上,利用光流信息聚集相邻帧上的特征,提高检测的精度。对于当前帧,FGFA 网络会将其前后各 k 帧的特征按照 DFF 网络的方法预测出对应的当前帧特征,再将这 $2k$ 帧特征按照权值平均的方法与当前帧特征聚合,将聚合后的特征进行后续的分类回归。

光流信息可以通过深度学习方法或传统方法获取。深度学习方法需要进行额外的标注和训练,传统方法在背景发生剧烈变化时会失效,并且光流信息对光照的变化比较敏感,也不能像运动前景图那样明确地反映出来前景的位置。而本文的应用场景是针对固定摄像机拍摄的监控视频,不用考虑到周围环境的运动,因此,可以使用帧间差分法或背景差分法来获取运动信息,

运动信息的获取更加快速,能更加准确地反映前景位置。同时,本文引入运动信息,目的是采用一种便捷的方式为卷积神经网络提供目标的位置信息,而不用进行代价昂贵的标注,光流信息的深度学习方法需要引入额外的标注,反而违背了这一初衷。

为了提高检测器的定位精度,目前使用的改进方法大致有 2 类:一类是选择更加能够反映定位准确性的定位损失函数;另一类是在检测器中加入定位置信度的预测。一般来说,判断预测目标定位是否精准,是通过计算预测目标的检测框和标注信息中真值框的 IoU 值。但是 IoU 存在一个问题,对于没有重叠的 2 个框,因为 IoU 值为 0,所以无法进行学习训练。在这种情况下,有学者提出了 GIoU (generalized intersection over union)[21]的概念。GIoU 同时关注了重叠区域和不重叠区域,即使两框没有重叠,也能计算出距离值,取值范围为[-1,1],越接近 1 表示两框的重叠程度越大,越接近 -1 表示两框的距离越远。而对于在检测器中加入定位置信度预测的方法,IoU-Net[22]中为检测器设计了一个独立的定位分支,在 Faster R-CNN 网络的基础上,使用真值框和检测框 IoU 作为定位置信度进行学习;IoU-aware RetinaNet[23]是在 RetinaNet 网络的基础上,在回归分支中增加了一个模块预测定位置信度。

以上方法对于定位精度的衡量都是根据检测框和真值框来计算的,需要精确的标注信息,但是对于数据集中的困难目标,在很多时候是缺少标注的,而在运动前景图中可以显示出这类目标的位置,因此,可以通过计算运动前景图中前景的外接矩形框和预测目标检测框的匹配程度,作为候选目标的定位置信度指标。同时,相较于采用独立的定位分支,本文在检测器中直接增加一个定位分支,使用与回归分支和分类分支相同的特征,学习候选目标的定位置信度,保证特征的一致性。

2 本文方法

在室外视频监控的环境下,针对基于卷积神经网络的目标检测算法对困难目标标注代价高和置信度预测时考虑不全面这 2 个问题,本文提出了一种运动信息引导的目标检测算法。从基于卷积神经网络的目标检测算法的特征提取和置信度预测这 2 个方面展开研究,分为 3 个模块:运动前景图获取模块、多尺度特征融合模块和定位分支模块。网络结构如图 1 所示。

本文所提目标检测算法以原始视频帧和视频帧对应的运动前景图作为输入。运动前景图获取模块的作用是获取视频帧对应的运动前景图。本文针对的是摄像机固定下的监控场景,因此选取了快速的、无监督的 ViBe 背景建模算法来获取运动前景图,并对 ViBe 算法进行了改进。改进后的算法能够减少静止目标融入背景或提取前景有残缺的情况。整个网络结构可以看作 2 个阶段:①特征提取阶段。通过 ResNet 的骨架网络输出一系列的特征图,经过特征金字塔网络(feature pyramid networks, FPN)输出金字塔网络的各层尺寸不同的特征,通过本文提出的多尺度特征融合模块,将运动信息与金字塔网络的各层输出进行融合,形成各层最终的特征图。②检测阶段。将特征提取阶段获取的特征图输入到检测器中进

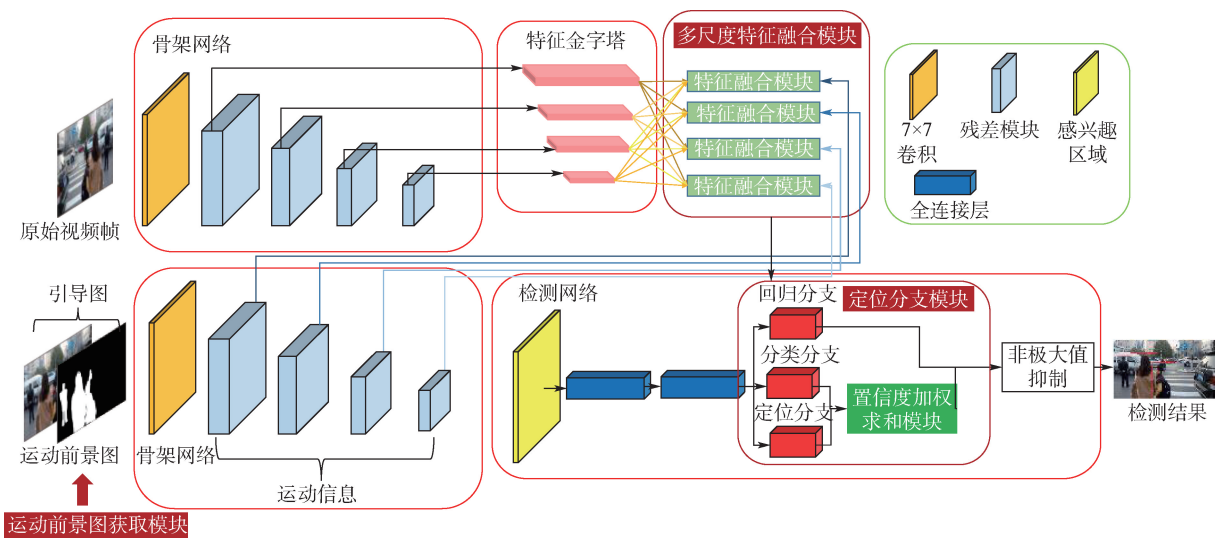


图 1 本文方法网络结构

Fig. 1 Network structure of the proposed method

行检测,经过分类分支、回归分支及本文提出的定位分支,得到检测框、分类置信度和定位置信度,再经过非极大值抑制(non maximum suppression, NMS)[24]得到最终的检测结果。

2.1　运动前景图获取模块

ViBe[7]算法是一种像素级的背景建模算法。在初始化时会对像素点建立一个邻域集合,对于之后新的图像帧,为每个像素点计算其与每个邻域像素点的欧氏距离。如果该像素与足够多的邻域像素点(一般设置为 20)欧氏距离小于设置的阈值(一般为 16),则认为该像素点属于背景像素;否则认为该像素点属于前景像素,由此生成运动前景图。

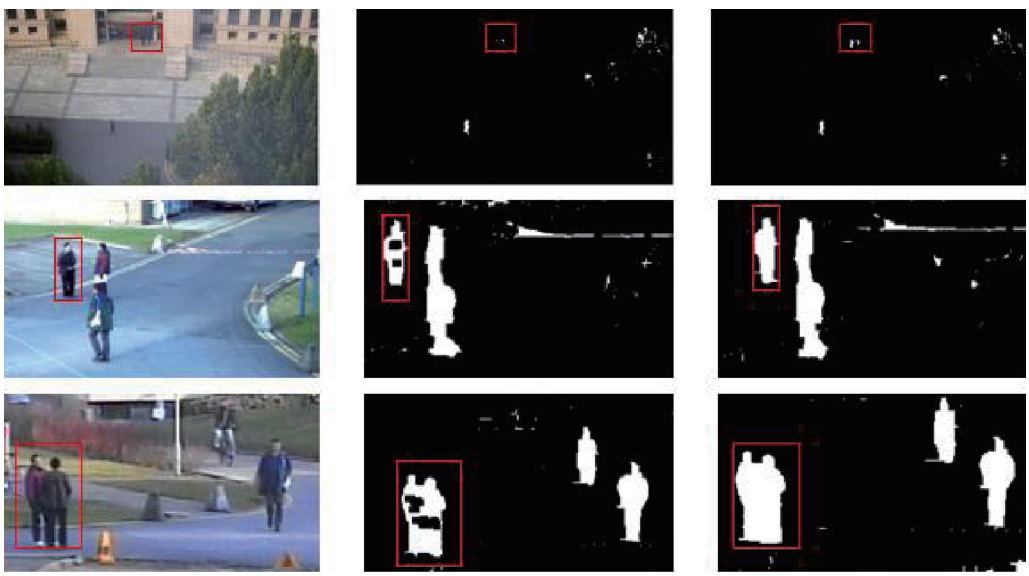
运动目标检测算法存在一个问题,对于视频中运动一段时间后长时间静止的目标,会将其判断为静止目标,或在检测的前景上出现空洞。为了减少这种情况的发生,本文对 ViBe 算法进行了改进,利用历史帧与当前帧的相似度对静止目标作二次判断,避免静止目标融入到背景之中。

改进的 ViBe 背景建模算法会为每个像素点设置一个前景计数器 fgCount 和一个前景标志 fgFlag。初始时前景计数器 fgCount 设置为 0,表示该像素点被判断为前景的次数为 0,前景标志 fgFlag 也设置为 0,表示该像素为背景。

对于输入的图像,将图像中的各个像素与初始背景模型进行匹配度计算。如果像素点与背景模型匹配,则认为该像素点为背景点,将该像素点的 fgCount 和 fgFlag 都置为 0,对该像素点对应的邻域背景模型进行更新;如果不匹配,则认为像素点为前景点,fgCount 值加 1,fgFlag 置为 1,表示该像素点为前景。

当一个像素的前景计数器 fgCount 达到阈值(设置为 10)时,认为该前景已经持续存在了一段时间,需要判断该前景对应的是否为静止目标。判断方法是:计算该像素与所保存的历史背景帧的对应像素的误差平方,确定相似度,如果相似,认为该点为静止目标,fgCount 置为 0,fgFlag 保持为 1;否则 fgCount 和 fgFlag 都置为 0,对该像素点对应的邻域背景模型进行更新,并用当前帧的像素更新历史背景帧的对应像素。

图 2(a)为原始视频帧,图 2(b)为 ViBe 算法得出的前景图,图 2(c)为改进的 ViBe 算法得出的前景图,第 1 行中目标长时间静止,消失在前景图中,改进的 ViBe 算法能够得到静止目标的前景,第 2、3 行中目标静止导致前景图中有空洞,改进的 ViBe 算法能够填补前景中的空洞。



(a) 原始视频帧　(b) ViBe算法　(c) 改进的ViBe算法

图 2　运动前景图结果比较

Fig. 2　Comparison of foreground map results

2.2　基于运动信息的多尺度特征融合模块

基于运动信息的多尺度特征融合方法网络结构如图 3 所示。

为了获取更多目标的信息,本文在运动前景图的基础上,将原始视频帧的灰度图与运动前景图进行通道拼接,得到引导图。这样,不仅可以对运动前景图补充信息,而且可以将运动前景图也扩充为三通道的图像,方便后续操作。同时,基于运动信息的多尺度特征融合模块是在特征层面进行运动信息与特征图的融合操作,因此,为了保证最终获得的运动信息和卷积神经网络产生的特征图具有相同的尺寸和通道数,会将获得的引导图也输入到

相同的骨架网络中,得到最终的运动信息。

输入的原始视频帧经过骨架网络和特征金字塔网络之后得到了自上而下 X_1、X_2、X_3、X_4四层特征,同样,运动前景图与原始图像灰度图通道拼接后的引导图经过骨架网络后得到了 G_1、G_2、G_3、G_4四层的运动信息,这些信息是特征融合模块的输入。下文就以第 4 层的特征融合模块为例进行介绍,特征融合方式如图 4 所示。

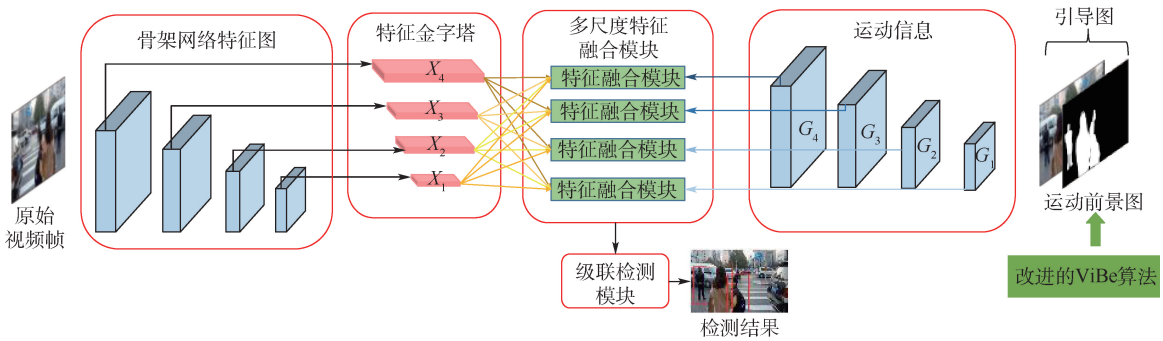


图 3 基于运动信息的多尺度特征融合模块网络结构

Fig. 3 Network structure of multi-scale feature fusion module based on motion information

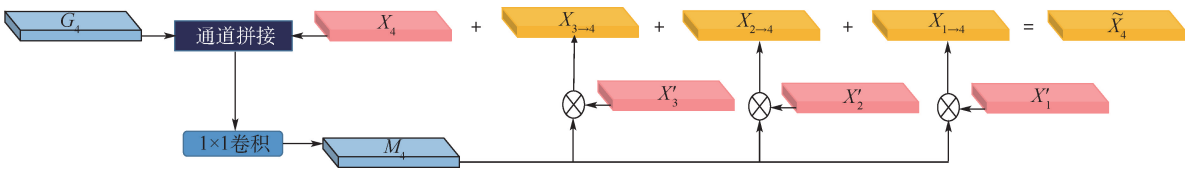


图 4 多尺度特征融合模块示意图

Fig. 4 Schematic diagram of multi-scale feature fusion module

$$X_{1\to4} = X'_1 \otimes M_4 \tag{1}$$

$$X_{2\to4} = X'_2 \otimes M_4 \tag{2}$$

$$X_{3\to4} = X'_3 \otimes M_4 \tag{3}$$

$$\tilde{X}_4 = X_4 + X_{1\to4} + X_{2\to4} + X_{3\to4} \tag{4}$$

式中:X'_1、X'_2、X'_3分别为将 X_1、X_2、X_3 经过上采样操作处理成与 X_4相同大小的特征;M_4 为将 X_4 与获取的引导图拼接在一起,经过一个简单的卷积网络后得到的注意力分布图;$X_{1\to4}$、$X_{2\to4}$、$X_{3\to4}$分别为特征金字塔网络前 3 层特征图与注意力分布图按元素相乘后的特征;$\tilde{X}_4$ 为第 4 层融合后的特征。

其他层的特征融合模块的操作与第 4 层类似,通过这种方式将运动信息与特征金字塔网络各层特征融合,使得特征图重点关注可能存在目标的区域。

2.3 运动前景图指导的定位分支模块

本文根据计算的候选目标检测框与前景区域外接矩形框的匹配程度作为衡量候选目标定位置信度的指标。受到 IoU 计算公式的启发,考虑可以用两者区域交集的面积比上并集的面积来表示匹配程度。但是通过对运动目标前景区域的观察发现,在目标比较密集的区域,其前景区域会粘连在一起,提取的外接矩形框也会包含多个目标,使用两者区域的并集作为分母就会得到一个比较小的值,无法正确地反映两者的匹配程度。因此,在计算检测网络提取到的预测目标检测框与前景图中的外接矩形框匹配程度时,对 IoU 的计算方法进行了改进,如下:

$$\text{Match} = \frac{(\hat{y}_2 - y_1)(x_2 - \hat{x}_1)}{(\hat{y}_2 - \hat{y}_1)(\hat{x}_2 - \hat{x}_1)} \tag{5}$$

式中:Match 表示预测目标检测框与前景图中的外接矩形框的匹配程度;$(\hat{x}_1, \hat{y}_1, \hat{x}_2, \hat{y}_2)$为预测目标检测框的左上、右下坐标;$(x_1, y_1, x_2, y_2)$为运动前景图中外接矩形框的左上、右下坐标;分子表示候选目标与外接矩形框相交区域的面积,分母表示候选目标框的面积。

本文在目标检测算法的检测器中增加一个分支作为定位分支,该分支与分类分支和回归分支共享 2 个全连接层,3 个分支保持特征的一致性。增加定位分支后的检测器结构如图 5所示,图中各组成结构上的数字代表该结构输出的大小,C 表示使用的数据集中目标种类的数目。

定位分支和分类分支、回归分支一样,会共同参与到对于全连接层的训练中,提升全连接层在提取特征时的定位准确率,使候选目标能够得到一个定位置信度的同时,也能够提升回归分支中检测框的定位精度。

3 个分支使用的训练损失函数各不相同,分

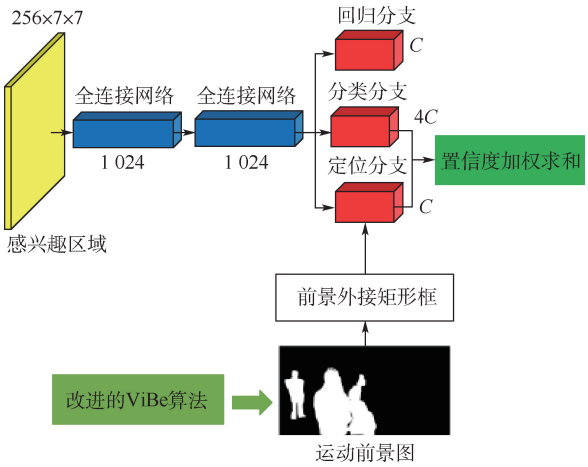


图 5　引入定位分支后的检测器结构

Fig. 5　Detection head structure after introducing localization branch

类分支使用了交叉熵损失函数:

$$L_{\text{cls}} = -\sum_{i=1}^{C} y_i \log_2 p_i \tag{6}$$

式中:L_{cls}为分类损失;y_i 为示性函数,当目标预测的类别和真实类别相同时取 1,其余取 0;p_i 表示目标类别为当前类别的概率;C 为类别数。

回归分支使用了 Smooth L1 损失函数,如下:

$$L_{\text{bbox}} = \frac{1}{n}\sum_{i=1}^{n} \text{Smooth}_{\text{L1}}(\text{bbox}_{\text{pred}_i} - \text{bbox}_{\text{target}_i}) \tag{7}$$

$$\text{Smooth}_{\text{L1}} = \begin{cases} 0.5x^2 & |x| < 0.5 \\ |x| - 0.5 & \text{其他} \end{cases} \tag{8}$$

式中:L_{bbox}为回归损失;$\text{bbox}_{\text{pred}_i}$为回归分支得到的检测框;$\text{bbox}_{\text{target}_i}$为真值框;$n$ 为正样本的数量;$\text{Smooth}_{\text{L1}}$为计算 Smooth L1 函数。

定位分支使用均方误差损失函数(MSE),如下:

$$L_{\text{loc}} = \frac{1}{n}\sum_{i=1}^{n} (\text{loc}_{\text{pred}_i} - \text{loc}_{\text{true}_i})^2 \tag{9}$$

式中:L_{loc}为定位损失;$\text{loc}_{\text{pred}_i}$为全连接网络中定位分支得到的候选目标框的定位置信度;$\text{loc}_{\text{true}_i}$为根据式(5)计算的候选目标框和前景外接框的匹配程度的真值,只对属于正样本的候选框计算匹配程度,因此 n 代表正样本数量。

分类分支、回归分支和定位分支 3 个分支在训练时采用联合训练,将 3 个分支的损失求和作为网络的总损失:

$$L_{\text{total}} = L_{\text{cls}} + L_{\text{bbox}} + L_{\text{loc}} \tag{10}$$

式中:L_{total}为整个网络的损失。

在测试阶段,通过定位分支获得每个预测目标的定位置信度,再与分类分支获得的分类置信度进行加权求和,获得目标最终的置信度:

$$\text{confidence}_{\text{score}} = \partial \cdot \text{cls}_{\text{score}} + (1 - \partial)\text{loc}_{\text{score}} \tag{11}$$

式中:$\text{confidence}_{\text{score}}$为目标最终的置信度;$\text{cls}_{\text{score}}$从分类分支获得,为候选目标的分类置信度;$\text{loc}_{\text{score}}$从定位分支得到,为候选目标的定位置信度;$\partial$为权重,本文取值 0.6。得到预测目标的最终置信度后,用这一置信度指导 NMS 方法进行重复候选框的筛除,获得检测结果。

3　实验结果与分析

3.1　实验设置

本文实验在 Ubuntu 系统下进行,整体神经网络框架基于 Pytorch 构建而成。程序编写中使用的主要编程语言为 Python3,同时为了方便程序的管理和使用,通过 anaconda 搭建了一个虚拟环境,全部实验都在虚拟环境下进行。除此之外,在硬件方面,使用了 GPU 加速训练过程。

本文使用的基线方法是 Cascade R-CNN[18] 网络,包括 ResNet-50 骨架网络、特征金字塔网络、RPN 网络和级联检测网络。在训练策略上采用联合训练方法,使用随机梯度下降法(stochastic gradient descent, SGD)训练策略,初始学习率设置为 0.001,使用 step 方法进行学习率的调整,动量设置为 0.9,权值衰减系数为 0.000 1。

本文在进行实验时选用了 3 个数据集:①笔者所在实验室自建的且已经公开发表的数据集 DML_det[25];②DukeMTMC[26] 数据集中的一部分数据;③PETS09[27] 数据集中的部分数据。其中,DML_det 数据集来自于笔者所在实验室,拍摄于北京航空航天大学,时间跨度较长,数据集中的目标较小。训练阶段使用了 DML_det 中的 18 段视频段,测试阶段使用了其中 10 个视频段,并且对视频段进行了采样间距为 100 帧的采样。DukeMTMC 数据集拍摄于杜克大学校园。本文选择了 Cam1 和 Cam9 两个场景进行实验,同样使用了采样方法,每隔 100 帧选取一张图像,共获得了 2 556帧图像,其中一半图像作为训练集,另一半图像作为测试集。PETS09 数据集拍摄于英国雷丁大学,选择了 S0 作为训练集,S2 中的 L2 难度的测试集作测试。S0 中共包含 7 163 张训练图像,S2 的 L2 中共包含 1 131 张测试图像。针对使用的 3 个数据集,在训练测试时只对其中的行人目标进行检测,骑自行车或三轮车的行人也归属于行人类别。

评价标准方面,本文采用了与 COCO 数据集相同的评价方式,即通过平均精确率(average precision, AP)和召回率(Recall)来评价模型[28]。

对于网络给出的检测框,只有当其与真值框

计算出的 IoU 大于设定的阈值时,才认为检测结果是正确的。这个目标被成功地检测出来,则属于正确的正样本(true positive, TP),而如果有一个被标注的目标,没有检测框与其对应,就认为这是一个错误的负样本(false negtive, FN),当一个没有被标注过的区域出现了检测框,就认为这是一个错误的正样本(false positive, FP)。基于上述对不同检测结果的定义,精确率(Precision)和召回率(Recall)的计算公式如下:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{12}$$

式中:Precision 表示正确检测出来的目标占所有检测目标的比率。

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{13}$$

式中:Recall 表示正确检测出来的目标占所有目标的比率。

本文采用了多个 IoU 阈值对提出的算法进行评估,其中 AP@[0.5:0.95]代表选取了从 0.5、0.55 到 0.95 的 IoU 阈值,计算这些阈值下的平均精确率,再求这些平均精确率的平均值,Recall@[0.5:0.95]代表从 0.5、0.55 到 0.95 的 IoU 阈值下的平均召回率。

3.2 对比实验

本文提出了一种基于运动信息的多尺度特征融合方法。利用运动目标检测算法可以获得目标位置信息,尤其是困难目标信息的特点,并对运动目标检测算法进行改进,将运动前景图作为空间权重融合到卷积神经网络提取的特征图中。同时,提出了一种运动前景图指导的定位分支模块,根据运动前景图中前景的空间位置,衡量卷积神经网络候选目标的定位置信度。

为了说明本文算法对目标检测算法准确率的提升,选取了几种近年提出的目标检测算法,同样在 DML_det 数据集、DukeMTMC 数据集和 PETS09 数据集上进行了训练和测试,与本文算法做对比,实验结果如表 1 ~ 表 3 所示。

从表 1 ~ 表 3 的实验结果可以看出,YOLOX、YOLOF、Sparse R-CNN、Deformable DETR 等算法对于本文所使用的数据集表现不好。一方面,本文选用的数据集数据量远小于 COCO 数据集;另一方面,如 Sparse R-CNN 算法需要多个轮次的训练才可以达到最好的效果。整体的实验结果表明,对于固定摄像机拍摄的监控视频场景下的行人检测,本文算法表现较好。

表 1 DML_det 数据集上本文算法与其他目标检测算法对比

Table 1 Comparison of the proposed algorithm with other object detection algorithms on DML_det dataset %

方法	AP@[0.5:0.95]	Recall@[0.5:0.95]
TOOD[29]	36.3	49.5
Dynamic R-CNN[30]	36.1	46.5
ATSS[31]	36.8	48.4
VarifocalNet[32]	37.2	48.2
YOLOF[33]	19.0	33.2
Sparse R-CNN[34]	24.0	44.7
YOLOX[35]	25.1	36.7
FoveaBox[36]	25.8	38.5
Double-Head R-CNN[37]	33.7	41.3
NAS-FCOS[38]	35.5	46.1
本文方法	43.6	57.7

表 2 DukeMTMC 数据集上本文算法与其他目标检测算法对比

Table 2 Comparison of the proposed algorithm with other object detection algorithms on DukeMTMC dataset %

方法	AP@[0.5:0.95]	Recall@[0.5:0.95]
TOOD[29]	59.0	64.9
Dynamic R-CNN[30]	58.1	63.5
ATSS[31]	56.3	63.1
VarifocalNet[32]	56.5	63.0
YOLOF[33]	40.0	52.4
Sparse R-CNN[34]	45.0	63.4
YOLOX[35]	44.6	53.2
FoveaBox[36]	54.6	60.8
Double-Head R-CNN[37]	55.7	60.5
NAS-FCOS[38]	56.8	63.6
本文方法	62.0	71.0

表 3 PETS09 数据集上本文算法与其他目标检测算法对比

Table 3 Comparison of the proposed algorithm with other object detection algorithms on PETS09 dataset %

方法	AP@[0.5:0.95]	Recall@[0.5:0.95]
TOOD[29]	38.9	53.8
Dynamic R-CNN[30]	40.4	52.8
ATSS[31]	38.9	54.6
VarifocalNet[32]	39.6	55.3
YOLOF[33]	36.2	50.4
Deformable DETR[39]	29.5	45.3
YOLOX[35]	36.0	47.8
FoveaBox[36]	36.3	52.3
Double-Head R-CNN[37]	38.5	50.0
NAS-FCOS[38]	40.7	56.0
本文方法	42.3	57.9

3.3　消融实验

将本文提出的多尺度特征融合方法和定位分支模块,分别应用于基线方法 Cascade R-CNN 中,在 DML_det 数据集、DukeMTMC 数据集和 PETS09 数据集上对算法进行了消融实验,得到的结果如表 4 ~ 表 6 所示。

表 4　DML_det 数据集性能评价指标

Table 4　Performance evaluation indexes of DML_det dataset

方法				评价指标/%	
Cascade R-CNN	改进 ViBe 算法	多尺度特征融合	定位分支	AP@[0.5:0.95]	Recall@[0.5:0.95]
√				35.5	52.5
√	√	√		40.1	56.4
√	√		√	40.7	57.8
√		√	√	41.9	58.5
√	√	√	√	43.6	57.7

表 5　DukeMTMC 数据集性能评价指标

Table 5　Performance evaluation indexes of DukeMTMC dataset

方法				评价指标/%	
Cascade R-CNN	改进 ViBe 算法	多尺度特征融合	定位分支	AP@[0.5:0.95]	Recall@[0.5:0.95]
√				58.1	68.2
√	√	√		60.4	69.3
√	√		√	60.2	70.3
√		√	√	61.2	70.7
√	√	√	√	62.0	71.0

表 6　PETS09 数据集性能评价指标

Table 6　Performance evaluation indexes of PETS09 dataset

方法				评价指标/%	
Cascade R-CNN	改进 ViBe 算法	多尺度特征融合	定位分支	AP@[0.5:0.95]	Recall@[0.5:0.95]
√				38.2	54.8
√	√	√		40.0	57.2
√	√		√	39.9	57.3
√		√	√	41.2	56.7
√	√	√	√	42.3	57.9

从表 4 ~ 表 6 实验结果可以看出,引入多尺度特征融合模块和定位分支都可以在基线方法上带来精确率和召回率的提升,并且将 2 种方法整合之后,在精确率和召回率方面有更近一步的提升,可以说明本文算法的有效性。同时,对改进的 ViBe 算法进行了实验验证,从表 4 ~ 表 6 中的结果可以看出,使用改进后的算法获取运动前景图,能够提升目标检测的准确率。

图 6 ~ 图 11 分别展示了 3 个数据集上基线方法和引入了本文算法后的部分结果示意图,绿框代表真值框,红框代表检测框。图 6 中的第 1 幅图和第 4 幅图中,环境中存在与行人目标相似的物体,因此产生了误检。在引入本文算法后,增加了运动前景图的信息,误检的物体在前景图中没有对应的前景,因此特征响应值降低,从而去除误检。图 8 中的第 1、3、4 幅图和图 10,因为测试数据中的目标被遮挡或目标太小,而没有被检测器检测出来,引入本文算法后,这些困难目标在前景图中可以反映出来,提升了特征图的响应值,因



图 6　DML_det 数据集基线方法结果

Fig. 6　Baseline method results of DML_det dataset

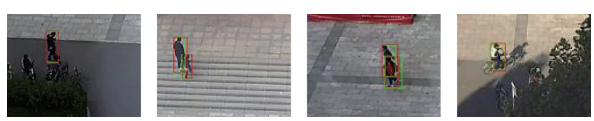


图 7　DML_det 数据集本文算法结果

Fig. 7　The proposed algorithm results of DML_det dataset



图 8　DukeMTMC 数据集基线方法结果

Fig. 8　Baseline method results of DukeMTMC dataset



图 9 DukeMTMC 数据集本文算法结果

Fig. 9 The proposed algorithm results of DukeMTMC dataset



图 10 PETS09 数据集基线方法结果

Fig. 10 Baseline method results of PETS09 dataset



图 11 PETS09 数据集本文算法结果

Fig. 11 The proposed algorithm results of PETS09 dataset

此能够被检测到。图 6 中的第 2、3 幅图和图 8 中的第 2 幅图,使用基线方法检测框的位置偏离目标,而引入本文算法后,能够使检测框位置更加准确。

4 结 论

本文提出了一种运动信息引导的目标检测算法,适用于固定摄像头下的监控视频场景,包括改进的运动前景图获取方法、基于运动信息的多尺度特征融合方法和运动前景图指导的定位分支模块。所提算法在设计时没有引入复杂的先验信息或需要进行精细的标注,并且通过在3 个固定监控视频行人检测数据集上所作的实验,验证了本文算法可以提升目标检测算法的准确率。

参考文献(References)

[1] DOLLAR P, WOJEK C, SCHIELE B, et al. Pedestrian detection: An evaluation of the state of the art[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 34(4):743-761.

[2] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2005, 1:886-893.

[3] AHONEN T, HADID A, PIETIKÄINEN M. Face recognition with local binary patterns[C]// European Conference on Computer Vision. Berlin: Springer, 2004:469-481.

[4] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2):91-110.

[5] SINGLA N. Motion detection based on frame difference method[J]. International Journal of Information & Computation Technology, 2014, 4(15):1559-1565.

[6] ZIVKOVIC Z. Improved adaptive Gaussian mixture model for background subtraction[C]// Proceedings of the 17th International Conference on Pattern Recognition. Piscataway: IEEE Press, 2004, 2:28-31.

[7] BARNICH O, VAN DROOGENBROECK M. ViBe: A universal background subtraction algorithm for video sequences[J]. IEEE Transactions on Image Processing, 2010, 20(6):1709-1724.

[8] PICCARDI M. Background subtraction techniques: A review[C]// 2004 IEEE International Conference on Systems, Man and Cybernetics. Piscataway: IEEE Press, 2004, 4:3099-3104.

[9] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]// European Conference on Computer Vision. Berlin: Springer, 2016:21-37.

[10] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016:779-788.

[11] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017:7263-7271.

[12] REDMON J, FARHADI A. YOLOv3: An incremental improvement[EB/OL]. (2018-04-08)[2022-04-08]. https://arxiv.org/abs/1804.02767v1.

[13] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[J]. Artificial Intelligence, 2020, 4:1-17.

[14] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]// Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017:2980-2988.

[15] GIRSHICK R, DONAHUE J, DARRELL T, et al. Region-based convolutional networks for accurate object detection and segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 38(1):142-158.

[16] GIRSHICK R. Fast R-CNN[C]// Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2015:1440-1448.

[17] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. Advances in Neural Information Processing Systems, 2017, 39(6):1137-1149.

[18] CAI Z, VASCONCELOS N. Cascade R-CNN: Delving into high quality object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018:6154-6162.

[19] ZHU X, XIONG Y, DAI J, et al. Deep feature flow for video recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017:2349-2358.

[20] ZHU X, WANG Y, DAI J, et al. Flow-guided feature aggregation for video object detection[C]// Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017:408-417.

[21] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized in-

tersection over union: A metric and a loss for bounding box regression[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2019: 658-666.

[22] JIANG B, LUO R, MAO J, et al. Acquisition of localization confidence for accurate object detection[C]// European Conference on Computer Vision. Berlin: Springer, 2018: 784-799.

[23] WU S, LI X, WANG X. IoU-aware single-stage object detector for accurate localization[J]. Image and Vision Computing, 2020, 97: 103911.

[24] NEUBECK A, VAN GOOL L. Efficient non-maximum suppression[C]// Proceedings of the 18th International Conference on Pattern Recognition. Piscataway: IEEE Press, 2006, 3: 850-855.

[25] WANG X, HU H M, ZHANG Y. Pedestrian detection based on spatial attention module for outdoor video surveillance[C]// 2019 IEEE 15th International Conference on Multimedia Big Data (BigMM). Piscataway: IEEE Press, 2019: 247-251.

[26] ZHANG Z, WU J, ZHANG X, et al. Multi-target, multi-camera tracking by hierarchical clustering: Recent progress on dukemtmc project[EB/OL]. (2017-11-27)[2022-04-08]. https://arxiv.org/abs/1712.09531.

[27] FERRYMAN J, SHAHROKNI A. PETS2009: Dataset and challenge[C]// 2009 12th IEEE International Workshop on Performance Evaluation of Tracking and Surveillance. Piscataway: IEEE Press, 2009: 1-6.

[28] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: Common objects in context[C]// European Conference on Computer Vision. Berlin: Springer, 2014: 740-755.

[29] FENG C, ZHONG Y, GAO Y, et al. TOOD: Task-aligned one-stage object detection[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 3490-3499.

[30] ZHANG H, CHANG H, MA B, et al. Dynamic R-CNN: Towards high quality object detection via dynamic training[C]// European Conference on Computer Vision. Berlin: Springer, 2020: 260-275.

[31] ZHANG S, CHI C, YAO Y, et al. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 9759-9768.

[32] ZHANG H, WANG Y, DAYOUB F, et al. VarifocalNet: An iou-aware dense object detector[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 8514-8523.

[33] CHEN Q, WANG Y, YANG T, et al. You only look one-level feature[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 13039-13048.

[34] SUN P, ZHANG R, JIANG Y, et al. Sparse R-CNN: End-to-end object detection with learnable proposals[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 14454-14463.

[35] GE Z, LIU S, WANG F, et al. YOLOX: Exceeding YOLO series in 2021[EB/OL]. (2021-08-06)[2022-04-08]. https://arxiv.org/abs/2107.08430.

[36] KONG T, SUN F, LIU H, et al. FoveaBox: Beyound anchor-based object detection[J]. IEEE Transactions on Image Processing, 2020, 29: 7389-7398.

[37] WU Y, CHEN Y, YUAN L, et al. Rethinking classification and localization for object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 10186-10195.

[38] WANG N, GAO Y, CHEN H, et al. NAS-FCOS: Fast neural architecture search for object detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 11943-11951.

[39] ZHU X, SU W, LU L, et al. Deformable DETR: Deformable transformers for end-to-end object detection[EB/OL]. (2021-03-18)[2022-04-08]. https://arxiv.org/abs/2010.04159v1.

Object detection algorithm guided by motion information

HU Haimiao*, SHEN Liuqing, GAO Likun, LI Mingzhu

(School of Computer Sience and Engineering, Beihang University, Beijing 100083, China)

Abstract: Due to the complexity of the scene and the diversity of objects, the objects in the scene of outdoor surveillance video are difficult to detect, which involves such problems like the object is blocked, or the size of object changes. Therefore, the object detection task is still challenging. To improve the accuracy of the object detection algorithm, this paper proposed a method of using motion information to guide the object detection algorithm based on convolutional neural network. Firstly, the motion object detection algorithm is improved to keep the foreground of stationary target in the motion foreground map; secondly, using the feature that the foreground in the motion foreground map can indicate the spatial position of the object, the feature map extracted by the network is fused with the motion information to improve the response value of the possible object area in the feature map; finally, in the detector of the object detection algorithm, a localization branch is introduced. Using the motion foreground map of the video frame, the location reliability of the candidate object is learned, and weighted sum with the classification confidence of the object is used as the final confidence of the object. The detection result is obtained through the non maximum suppression method. Experiments show that the proposed method can improve the accuracy of object detection in the data set collected under the fixed camera.

Keywords: motion information; foreground area; feature fusion; localization branch; object detection

Received: 2022-04-28; **Accepted**: 2022-05-13; **Published online**: 2022-05-17 14:33

URL: kns.cnki.net/kcms/detail/11.2625.V.20220517.1059.001.html

Foundation items: National Natural Science Foundation of China (62122011, U21A20514); Key Research and Development Program of Zhejiang Province (2022C01082)

* **Corresponding author**. E-mail: frank0139@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0307

自动铺丝成型构件缺陷在线检测技术进展

丁希仑[1,2], 罗伟恒[1], 刘斐[1], 王国栋[3], 陈维强[3], 张武翔[1,2,*]

(1. 北京航空航天大学 机械工程及自动化学院, 北京 100083;
2. 北京航空航天大学宁波创新研究院, 宁波 315800; 3. 北京卫星制造厂有限公司, 北京 100190)

摘 要: 自动铺丝(AFP)工艺为连续纤维增强复合材料结构提供了高效和稳定的成型方法,但铺放工艺参数和成型装备性能的不稳定极易造成制件出现间隙、重叠、纤维波纹等各类缺陷,从而严重影响整体结构的机械性能及使用寿命,因此,利用在线检测系统对成型过程中铺层表面或内部缺陷进行实时识别与评估尤为重要。归纳了复合材料铺放过程中常见缺陷的产生原因、表现形式及其对结构产品综合性能的影响。基于不同检测原理综述了复合材料制件自动铺丝成型过程在线检测系统的发展及应用特点,涵盖激光技术、可见光图像技术、热成像技术及布拉格光纤光栅技术等。总结了当前在线检测系统存在的问题,并展望了其未来发展趋势。

关 键 词: 连续纤维增强复合材料; 自动铺丝(AFP); 缺陷; 在线检测技术; 识别方法
中图分类号: TB332
文献标志码: A **文章编号**: 1001-5965(2022)09-1721-13

复合材料是由 2 种或 2 种以上的组分材料按一定方式复合而成的多相材料,其中连续纤维增强复合材料由增强纤维和均质基体组成,相较于钢、铝合金等传统结构材料,其具有质量轻、比强度高、比刚度高、耐高温、性能可设计和多功能兼容等诸多优点[1]。随着研究的不断深入,复合材料在航空及航天等重要领域的应用日益广泛,使用比例不断攀升,已成为影响科技发展的关键战略材料之一,如在波音 787、空客 A350 等先进民用大型飞机中,复合材料的用料达到了 50% 以上[2]。因此,面对结构轻量化与高性能的发展需求,复合材料结构的成型制造水平已成为衡量一个国家科技竞争力的重要标志[3]。

在复合材料结构制造过程中,铺叠作为制造关键工序之一,其质量直接影响最终成型构件的性能。手工铺叠方法作为目前中国航空领域主要采用的铺放方式,可以满足大多数复合材料制件的铺叠要求,且能够实现型面变化大、结构复杂制件的成型。然而,手工铺叠存在精度差、效率低、成本高、材料利用率不足等问题,无法满足高质量复合材料结构产品的制造要求。因此,出现了如热压罐成型、树脂传递模塑成型、纤维缠绕成型、自动铺带成型、自动铺丝成型、拉挤成型等复合材料结构自动成型工艺技术,有效提高了复合材料构件生产效率及成型质量稳定性。

自动铺丝(automated fiber placement, AFP)技术结合了纤维缠绕技术与自动铺带技术的优点,通过铺放头将多根预浸料丝束或干纤维集束成一条宽度可变的窄带,利用压辊及加热装置将其铺放到模具表面并加热压实定型。由于基于 AFP 成型制件质量稳定、生产效率高、生产成本低,且 AFP 铺放自由度高,可用于大曲率制件的铺放成

收稿日期: 2022-04-30; **录用日期**: 2022-05-18; **网络出版时间**: 2022-06-16 17:29
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220616.1430.001.html
基金项目: 浙江省自然科学基金 (LD22E050011); 宁波 2025 重大科技专项 (2022Z070)
* **通信作者**. E-mail: zhangwuxiang@buaa.edu.cn

引用格式: *丁希仑, 罗伟恒, 刘斐, 等. 自动铺丝成型构件缺陷在线检测技术进展[J]. 北京航空航天大学学报, 2022, 48(9): 1721-1733. DING X L, LUO W H, LIU F, et al. Review on automated fiber placement induced defects and their online monitoring technology[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1721-1733(in Chinese).*

型,该技术彻底改变了航空、航天工业复合结构的生产,且在其他市场领域中也有重要应用,如风能涡轮叶片中的复合材料结构制造[4-5]。

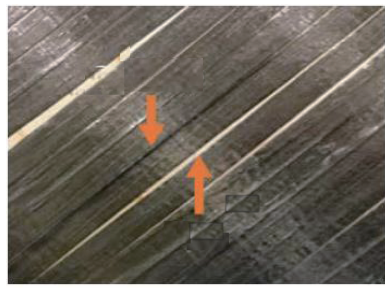
尽管AFP技术带来了一种更高效、稳定的复合材料成型方法,然而作为一种复杂、高维、非线性的生产技术,其工艺过程涉及多种工艺参数,无缺陷复合材料构件的生产制造仍是瓶颈问题。制件自动铺放成型过程中的缺陷形成与材料质量、工艺参数、轨迹规划及装备精度息息相关,严重影响复合材料结构的尺寸一致性及综合性能。因此,近年来围绕AFP过程实时监控、诊断与评估的研究一直是相关领域专家学者的研究热点。

本文针对基于AFP成型的复合材料构件在铺放过程中易产生的缺陷类型,简单介绍了其外观、成因及对制件性能产生的影响,基于不同检测原理,梳理了目前针对制件AFP工艺成型过程中铺层缺陷的在线检测技术,并指出了其未来发展方向。

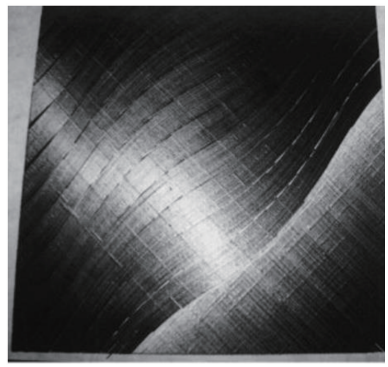
1 AFP成型构件中的缺陷

复合材料作为多相材料,材料本身对成型质量的影响与单相材料相比要更为复杂。此外,作为一个材料与结构同时成型的过程,丝束质量及工艺过程都会影响复合材料制件的宏观及微观结构。研究表明,AFP制件中的缺陷形成受到多种因素的影响,如AFP过程参数(进给速度、固化温度、压实力及铺设速度等)选择不当、丝束在制造过程中留存的缺陷、设备精度不足、轨迹规划不合理都会使得铺层表面在AFP铺放过程中出现多种缺陷,从而影响成型制件的综合性能[6]。

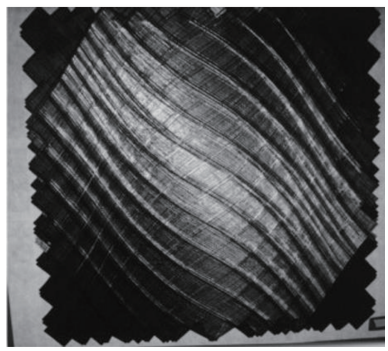
制造过程中产生的铺层表面缺陷为AFP成型技术应用的主要限制因素,对其进行分类有助于对缺陷形成机制的理解、缺陷对制件影响规律的分析及铺层表面缺陷的检测。Harik等[7]从设计、制造、分析和检测的角度出发,对AFP铺放过程中出现的多种表面缺陷进行了系统性的分类,其中,铺层表面的间隙与重叠作为传统层合板及变刚度层合板中最常见的缺陷,受到了广泛关注,如图1所示[7-8]。这2种缺陷在不同结构层合板中的形成原因也存在差异。在传统层合板的铺放过程中,铺放设备精度不足会导致实际铺放轨迹与预设轨迹不一致,从而导致间隙或重叠区域;而在铺放变刚度层合板时,除设备精度外,零件几何形状、复杂的丝束铺放路径、不同铺覆策略的选取,也会导致丝束在铺放过程中沿着牵引路径不可避免地出现呈三角形的间隙或重叠区域[7-8]。



(a) 传统层合板中的间隙及重叠



(b) 变刚度层合板中的间隙



(c) 变刚度层合板中的重叠

图1 传统层合板和变刚度层合板中的间隙与重叠缺陷[7-8]

Fig. 1 Gaps and overlaps in conventional laminates and variable-stiffness laminates[7-8]

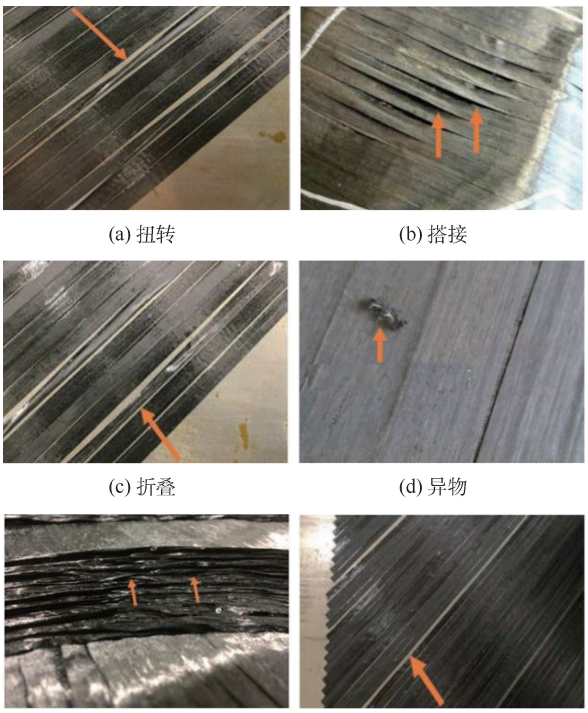
除间隙与重叠外,铺层在铺放过程中受操作手段、设备精度、工艺参数、零件结构等因素的影响,会出现多种如扭转、搭接、折叠、异物、褶皱、缺丝等表面缺陷,如图2所示[7]。其中,丝束扭转主要由进料系统操作不当、导向系统堵塞、丝束张力设置不当、铺放设备振动等原因造成;而丝束张力过大、压辊压实力不足、压辊尺寸选择不当都会影响丝束与铺放表面的贴合质量,导致丝束从铺放表面脱落,形成缺丝缺陷,或在凹面形成搭接区域。此外,当丝束沿曲线牵引路径铺放时,张紧轮张力问题同样会导致丝束在横向方向上折叠,从而产生间隙与重叠区域;异物指在铺设过程中铺层表面上堆积的不规则碳纤维球、树脂球或来自其他生产区域的碎片。

以上缺陷随着铺放过程不断累积,不仅会影响尺寸一致性,还会影响铺层之间的贴合质量,在制件内部形成多个贫树脂区或富树脂区,在厚度

方向上发生多处纤维错位,从而导致局部的纤维波浪,这一缺陷称为面外纤维波纹,如图 3 所示[9]。面外纤维波纹作为工业复合材料部件设计和制造中最常见的问题之一,通常发生于厚截面复合材料结构中,常作为层合板铺设过程中铺层多个表面缺陷(间隙、重叠、褶皱、搭接)的综合宏观表征。

在机械载荷的作用下,制件内部的缺陷会影响结构的可靠性及使用寿命,从而导致结构整体失效,因此在设计 AFP 成型大型航空构件时,准确评估缺陷对复合材料性能的影响对结构损伤容限设计至关重要[10]。研究人员通过实验或仿真等方式,探究了制件拉伸、压缩、剪切等力学性能对缺陷的敏感性,发现制件性能受铺设环境、缺陷类型、缺陷尺寸、缺陷位置、制件宽度、铺层规划等多种因素的影响[11-19]。近年来,复合材料逐渐应用在船舶中的水翼、螺旋桨及涡轮叶片等结构,这些结构往往比航空材料更厚,内部的缺陷使结构更易出现冲击损伤的问题,也对制件层间力学性能及损伤容限提出了更高要求。对此,许多研究人员定性或定量分析了缺陷对制件层间性能、断裂韧性、损伤容限、疲劳寿命的影响[20-30]。然而,大部分缺陷难以在铺设过程中通过人工方式引入,因此,目前大部分针对缺陷对制件综合性能影响的研究主要集中在间隙、重叠与面外波纹等 3 种缺陷。

随着现代工业的不断发展,复合材料越来越多地被应用于不同领域,如航空、航天及交通运输等领域,也对复合材料结构质量性能提出了进一步的要求。由以上研究可以看出,复合材料 AFP 成型工艺过程会给制件带来各类缺陷,影响制件的尺寸一致性、机械性能及疲劳寿命。在外部扰动及制件内部缺陷的耦合作用下,复合材料结构在其运行过程中可能会出现机械故障或结构失效。



(a) 扭转　(b) 搭接　(c) 折叠　(d) 异物　(e) 褶皱　(f) 缺丝

图 2　AFP 铺放过程中常见的制造缺陷[7]

Fig. 2　Common manufacturing defects induced by AFP process[7]

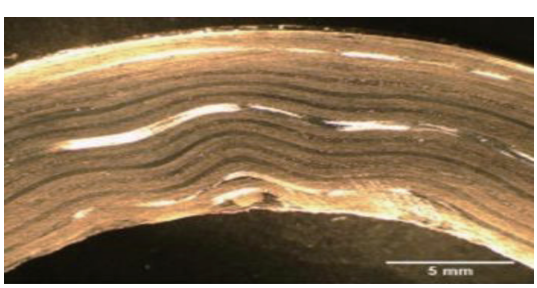


图 3　层合板中的纤维波纹缺陷[9]

Fig. 3　Out-of-plane waviness in composite laminates[9]

2　AFP 工艺在线检测技术研究进展

随着复合材料制造工艺自动化水平的不断上升,对高质量、低损伤部件的需求不断增加,然而 AFP 成型过程受工艺参数、材料质量、设备精度、轨迹规划等因素的影响,造成超出容许程度的各类缺陷,严重影响制件的力学性能,制约了其在诸多场景中的应用。传统的无损检测方法(如 X 光成像技术、剪切成像技术)受设备尺寸、质量、环境要求及响应时间等因素的限制,难以在 AFP 工艺过程中实时识别缺陷,因此,稳定可靠的在线检测系统对于 AFP 成型过程中结构的质量控制尤为重要[31-33]。

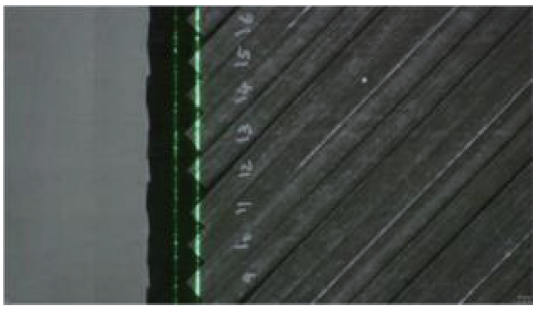
在 AFP 工艺发展的早期,工作人员通过目测评估铺层铺放质量。然而,铺层表面缺陷一般只有毫米级,且与预浸料之间的外观辨识度不高,使得表面缺陷难以判别,存在易漏检、过于依赖经验等问题,为缺陷的人工定位及评估带来了巨大的挑战[34]。据统计,利用目测法对复合材料进行缺陷检测与修复所需时间在自动铺丝工艺过程中的占比达到了 63%,是铺设过程耗时的 2.5 倍,严重制约了成形效率[35]。因此,对铺放过程进行实时监控,采用各类自动化缺陷检测方法极为重要。随着研究的不断进展,出现了一系列针对 AFP 过程的在线检测系统,如基于激光技术、基于可见光图像识别技术、基于热成像技术、基于布拉格光纤光栅技术的在线检测系统,以及基于其他技术原理的各类探索性在线检测系统,本文将分别总结

各技术的原理、优缺点及发展现状。

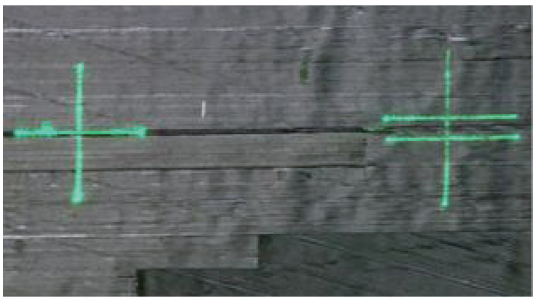
2.1 基于激光技术的复合材料制件 AFP 成型过程在线检测系统

激光检测技术作为一种非接触式测量方式，具有实时性好、精度高、分辨率高等优势，因此，许多研究人员针对复合材料制件 AFP 成型过程，利用激光投影仪、激光位移传感器、激光轮廓仪等设备，搭建了基于激光检测技术的在线检测系统，除辅助丝束定位外，还能够实现 AFP 铺放过程中铺层表面间隙、重叠、异物、扭转、搭接、褶皱、脱黏等多种缺陷的识别。

针对丝束铺放轨迹与预设轨迹不一致所带来的一系列缺陷（如间隙、重叠等），工作人员需要对缺陷的来源丝束进行定位以实现进一步的修复，而缺陷及丝束外观特征使得丝束定位及缺陷评估耗费的时间成本急剧增加。针对这一问题，激光投影仪由于其投影亮度高、设备体积小、准确率高等优点，常被用于复合材料手糊成型工艺过程中辅助铺层质量评估[36]。美国 Electroimpact 公司[37-38]利用激光投影仪在 AFP 过程中标记铺层边界、定位缺陷来源丝束，辅助检测人员区分预设间隙及生产过程带来的间隙缺陷，引导修复人员对缺陷进行修复。此外，Shadmehri 等[39-40]搭建的 AFP 过程在线检测系统利用激光投影仪获得制件三维空间信息，并根据缺陷种类设计的不同“缺陷特征”投影到制件表面，通过相机捕获铺层图像以辅助决策，如图 4 所示。除了标记铺层边界辅助铺层边缘质量评估外，该系统还能够实现丝束偏移检测及间隙大小测量。



(a) 铺层边缘质量评估



(b) 偏移检测及间隙测量

图 4 基于激光投影仪的 AFP 过程在线检测系统[39-40]

Fig. 4 AFP online monitoring systems based on laser projector[39-40]

与人工目测法相比，利用激光投影仪搭建的 AFP 过程在线检测系统提供了具有更高成本效益的解决方案，但是该技术目前多为人工检测辅助手段，未能实现 AFP 工艺过程中的在线自动化监测，主要有以下 2 个原因：

1）实时性不足。要在复杂表面上实现精准投影，需要确认激光投影仪与工作表面的空间关系。当在旋转芯轴上制造大型复合零件时，空间关系计算需要耗费大量时间，从而使得投影滞后于芯轴旋转速度，无法在铺放过程中定位丝束及缺陷。

2）适用场景受限。在复杂曲面上投影易产生误差，且误差随着入射角的增大而增大，难以精准定位缺陷。

激光三角位移传感器则为 AFP 工艺过程的自动检测带来了可能性，其作为数据采集方式，除了可以测量被测物体的空间位置信息外，还能获取铺层表面的深度信息，其主要由激光器、激光检测器和测量电路组成，具有高分辨率、高响应、适应恶劣环境等优点[41-44]。早在 1996 年，Kitson[41]将激光三角位移传感器集成在铺丝头以在铺放过程中提取丝束边缘空间位置信息，并通过计算机成像技术及特征提取算法处理获得的数据，进一步评估间隙、重叠的尺寸及识别铺放表面异物；然而有研究人员指出，通过丝束边缘位置信息计算铺面表面异常区域准确率不高，且大量的数据存储使得检测速度大幅下降[42-43]。针对这一问题，Zambal 等[44]基于激光三角位移传感器采集到的铺层表面深度信息生成铺层表面灰度图像，利用神经网络对其进行语义分割处理，从而实现铺层表面的间隙、重叠及异物的识别、分类与评估，如图 5 所示。图中，绿色、蓝色、红色及黄色区域分别代表无缺陷铺层表面、间隙、重叠及异物区域。

轮廓扫描仪通过激光线扫描的方式测量目标物表面位置、形状的变化，获取被测表面轮廓点云



图 5 基于激光三角位移传感器的表面制造缺陷识别与分类[44]

Fig. 5 Identification and classification of manufacturing defects on surface based on laser triangulation sensor[44]

信息,通过计算机成像技术进一步实现表面特征的识别与评估。针对激光投影仪难以实现高度自动化监测的问题,美国 FlightAware 公司[45]为 NASA 设计了一套 AFP 在线检测系统,利用安装在铺丝头或机械臂上的线激光扫描仪获取铺层表面的离散点数据,基于特征检测算法检测丝束边缘、重叠、间隙及搭接等特征,从而实现辅助决策。Cemenska 等[46-47]进一步结合激光投影仪、轮廓扫描仪、工业相机及特征识别软件,成功实现了在 AFP 过程中铺层边缘质量评估、多种表面缺陷识别、重叠和间隙宽度测量的自动化,并通过人机交互界面将超出容许程度的缺陷以图像与实时摄像的形式呈现给工作人员。Meister 等[48-50]搭建了基于激光轮廓仪的 AFP 过程在线检测系统(见图 6),利用 CLAHE 算法预处理其得到的铺层表面轮廓信息,结合 DCGAN 算法与传统数据增强算法扩充数据量,利用不同的卷积神经网络对铺层表面缺陷进行识别与分类,其对于褶皱、扭转、异物的分类准确率达到 100%,而间隙分类准确率达到了 96.81%。

尽管轮廓扫描仪具有实时性好、稳定性高等优点,且能够实现多种缺陷的自动化在线识别,但是目前系统铺层边界检测精度仍有待提高,要求工作人员对未检测到的铺层边界进行手工标记,且需要采集大量生产数据进一步优化缺陷识别及边缘检测算法,从而增大时间成本。

神经网络作为机器学习技术中的一种,常用于预测、控制、故障诊断,有不少学者结合神经网络及激光检测技术以实现自动铺放过程中铺层表面的缺陷检测,如 Zambal、Meister 等学者结合神经网络及激光传感器实现铺层表面间隙、重叠、褶皱、扭转、异物的评估[44,48-50]。此外,Sacco 等[51-53]基于美国 NCDMM 公司研发的复合材料结构检测系统(ACSIS),利用残差卷积神经网络和遗传算法对轮廓扫描仪获取的制件表面图像进行语义分割,从而实现制件表面缺丝、扭转、间隙、重叠、褶皱、搭接等多种缺陷的识别与分类。

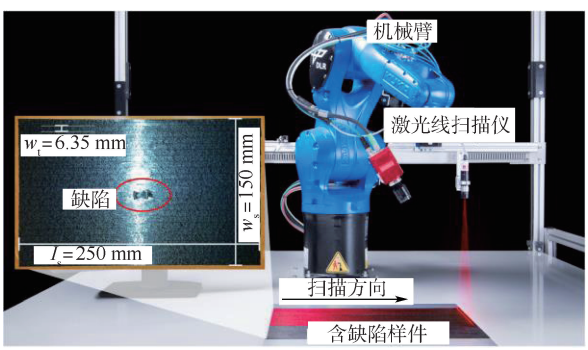


图 6　基于激光轮廓仪的 AFP 过程在线检测系统[48-50]

Fig. 6　In-situ AFP monitoring system based on profilometer[48-50]

综上所述,目前已有大量研究人员利用激光技术对 AFP 过程进行在线监测,且随着机器学习技术的发展,神经网络也被用于处理激光位移传感器及轮廓扫描仪获得的铺层表面数据,识别铺层表面的各类缺陷,进一步提升了监测系统的可靠性及响应速度。然而,激光检测效果仍然受到环境、数据、设备的限制,只能检测到铺设过程中的表面缺陷,无法识别制件在固化过程中内部结构的变化。

2.2　基于可见光图像识别技术的复合材料制件 AFP 成型过程在线检测系统

基于激光技术的 AFP 在线检测系统主要是利用激光投影仪辅助定位丝束和标记铺层边界,或是利用轮廓扫描仪实现铺放表面缺陷的检测。然而环境光易对检测过程产生干扰,从而降低缺陷识别精度。除了激光传感器外,工业摄像头作为常见的图像采集方法,在复合材料 AFP 成型过程中也有着许多应用。

Soucy[54]通过在铺丝头上安装摄像头系统以辅助工人进行铺设过程的在线检测,但是实验证明该方法同样受到环境光影响,且预浸料丝束与铺设表面之间辨识度不高。针对这一问题,Engelbart[55]利用工业相机实时监测铺放表面以评估表面缺陷,并通过照明系统为制件上的缺陷提供更高的对比度,减轻环境光源和制件材料反射率对缺陷检测的影响,然而该方法要求操作人员对捕获图像进行实时解读,自动化程度不高,缺陷识别主观依赖性强。与之相似,Shadmehri 等[39]搭建了激光视觉系统,但该系统作为辅助决策方法,同样依赖工作人员实现铺放质量评估,并未实现自动化检测。

随着计算机视觉技术的不断发展,工业相机不再仅仅作为提供远程监测方法的工具,通过结合图像处理算法为利用工业相机进行 AFP 过程铺层表面状态监测提供了更多的可能性。最初针对丝束铺放角度的实时监测,陈文等[56]提出了一种基于质心的二维图像分割法辅助铺层表面纹理提取,根据预先设定的纤维角度信息设计滤波器用于判断纹理方向,为铺放角度检测和表面缺陷检测提供了良好的基础。倪金辉等[57]指出采用标准 Hough 变换算法无法满足 AFP 过程中直线丝束边缘检测的实时性要求,并针对该问题提出了一种基于直线拟合的预浸料边缘直线在线检测

算法,与标准 Hough 变换算法相比,该算法能够保持相同精度及抗噪性能,大幅降低了检测耗时。

针对制件 AFP 铺放过程中铺层表面缺陷的检测,刘辉[58]搭建了纤维带在线检测系统,通过对图像进行预处理实现纤维带提取、边缘检测、图像细化,基于简化动态 Hough 变换算法、投影重心法实现铺放角度、间隙与重叠的评估与测量。Tao 等[59]设计了集成多光源照明系统及 CCD 相机的检测系统并置于铺丝头设备中,提出了基于直方图拉伸的图像增强算法,结合区域提取、边缘检测、Hough 变换等一系列图像处理算法,在铺设过程中定位丝束,测量间隙宽度,如图 7 所示。美国 Boeing 公司[60-61]设计了机器视觉系统用于监测 AFP 过程中出现的间隙、重叠、缺丝、异物等缺陷,其中铺丝头集成了数码相机及照明系统,以记录丝束铺放全过程,实现丝束铺放过程的三维可视化,但随着该系统的不断更新,其复杂程度也在不断提高。

与传统人工检测方式相比,视觉检测系统能够实现更时效的铺丝工艺在线检测和更集成的监测过程,并且为在线自动化检测提供了具有成本效益的解决方案。然而,基于可见光图像识别的铺层表面在线检测系统主要依赖于边缘检测技术,易受环境光及设备精度影响,存在检测精度不足、缺陷检测种类受限等问题,尽管研究人员利用照明系统减轻了环境光源的影响,该系统仍然受到丝束与铺带表面的光线反射特性的限制;此外,工业相机响应时间会对丝束最大铺放速度带来限制。

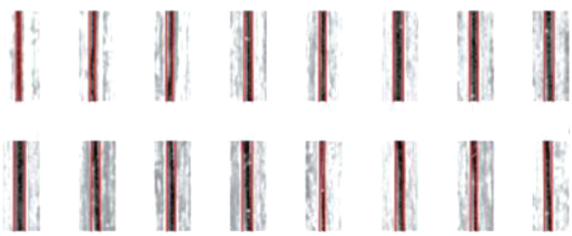


图 7 基于图像识别的不同间隙区域检测及测量[59]

Fig. 7 Identification and measurement of different gap regions during AFP process using image recognition technology[59]

2.3 基于热成像技术的复合材料制件 AFP 成型过程在线检测系统

在 AFP 工艺过程中,存放在铺放头中的预浸料丝束在铺放前需要保持低温以防止粘连,且铺丝头上集成的加热装置(热风枪、石英灯等)会对铺放表面进行预热,以保证层间贴合效果。由于预浸料在铺放过程中会比铺放平面的表面温度低,基于热成像技术的 AFP 在线检测系统可以利用温度差信息分析铺放质量。在此基础上,研究人员通过结合不同热源、不同算法,搭建了各类基于热成像技术的 AFP 工艺过程在线检测系统。

美国 Boeing 公司[62-66]在早期提出利用倾斜照明系统及红外摄像仪辅助监测,以实现间隙、重叠、拖拽、起皱、扭转、异物等多种铺层表面缺陷的在线识别,减少铺层表面检测所耗费的停机时间,同时利用多角度照明系统引导工作人员对缺陷进行修复。但随着该检测系统的迭代更新,系统复杂度也在不断提升,且利用短波红外摄像头只能实现对特定种类的缺陷检测[36,67]。

如图 8 所示,为了实现铺层表面缺陷的自动在线识别,Schmidt 和 Denkena 等[68-73]搭建了基于热成像技术的 AFP 在线检测系统,并通过一系列实验验证了其可行性及检测精度。在最初的设计中,研究人员利用红外摄像头监控不同的感兴趣区域(region of interest, ROI),结合边缘检测算法提取铺设丝束位置信息,通过与规划轨迹比较实现了间隙与重叠的实时测量;同时,针对丝束扭转、搭接、异物等多种缺陷,利用铺放表面温度场信息进行在线评估,并将该系统集成在 IFW 铺丝头上以提高监测可靠性及响应速度[68-70]。此外,Schmidt 和 Denkena 等[71]系统研究了路径规划、制件几何、工艺参数对铺放表面温度传导行为的影响,据此建立了热传导模型以预测在特定工艺参数下的铺放表面及铺设丝束的温度分布和变化,进一步提高了监测精度。为使监测系统更加独立于材料种类,Schmidt 等[72-73]利用卷积神经网络识别铺层表面缺陷(间隙、重叠、不同异物)并对其进行分类,通过比较不同模型分类效果选取最佳网络结构,其分类准确率达到了 92% 以上。上述监测系统主要通过与预设轨迹作对比,利用计算机视觉技术识别铺层表面缺陷,与之不同的是,NASA 的 Juarez 和 Gregory 等[74-78]开发了用于 AFP 过程监测的热成像系统,主要利用铺层表面热流信息,根据不同缺陷对热传导行为的影响不同这一机理实现铺层表面缺陷的检测。如图 9所示,该系统通过石英灯加热铺层表面,利用铺层表面作为透射热源,通过铺丝头上的红外摄像头与工业相机,记录铺放过程中铺层表面的温度场变化,发现即使没有经过任何图像处理,铺层表面缺陷也清晰可见(见图 10)。此外,研究人员利用数据规约算法、CLAHE 算法、有监督机器学习算法处理热成像仪采集到的铺层表面原始数据,减轻环境温度与镜头畸变带来的影响,检测间



图 8　基于热成像技术的 AFP 过程在线检测系统[68-73]

Fig. 8　Thermal image technology based online monitoring system for AFP process[68-73]

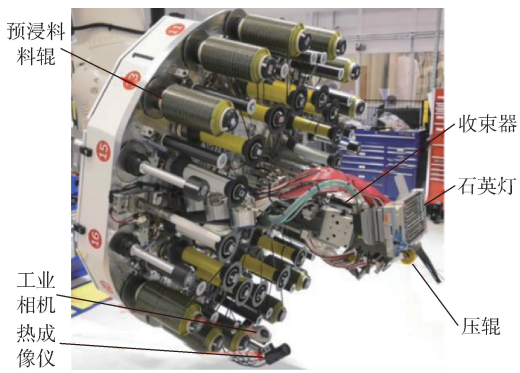


图 9　NASA 开发的基于热成像技术的 AFP 过程在线监测系统[74-78]

Fig. 9　Thermal image technology based online monitoring system for AFP process developed by NASA[74-78]

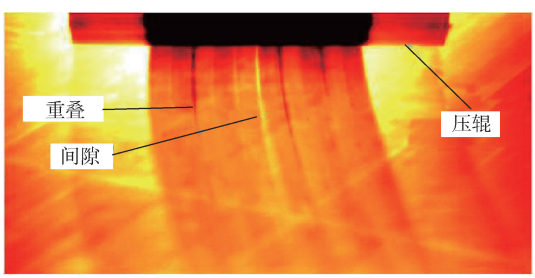


图 10　铺放过程中铺层表面热成像图[74-78]

Fig. 10　Thermal images of surface of laying ply[74-78]

隙、重叠、缺丝、起皱、脱黏等表面缺陷，在线评估铺放质量及压实质量，实验结果表明，该系统成功检测到复合材料制件 AFP 成型过程中出现的所有小于 0.762 mm 的缺陷，此外，该系统通过集成硬件及软件系统，实现了铺层表面的三维空间重构。Chen 等[67] 同样通过集成数据采集、缺陷识别、缺陷测量及缺陷可视化 4 个模块，利用卷积神经网络实现了多种缺陷（如间隙、重叠、缺丝、扭转、异物、搭接等）在 AFP 过程中的在线识别，同时基于图像提取与分割算法定量评估了缺陷尺寸，并通过三维重构对其进行可视化处理。

与可见光图像识别技术相比，热成像监测技术能够很好地集成到自动铺丝系统中以实现自动铺丝工艺过程监测的自动化，并且能够实现对多种缺陷的在线检测及贴合质量的评估。然而，热成像技术仍然存在一定的问题：

1）热源加热稳定性为该技术的应用带来了挑战，热源输出热量应该根据铺设速度进行自动校正，否则将会造成误检，降低检测精度。

2）铺设速度受到限制。热成像摄像头及光学摄像头的参数限制了最大铺设速度，当铺设速度过快时，容易造成漏检。

3）温度数据受到环境影响。由于热成像技术是利用铺放表面及铺设丝束的温度信息进行分析，环境温度会对铺设过程的温度信息带来噪声，使得原始温度数据信噪比过高。尽管研究人员通过数据处理的方法改善了这一现象，但是环境仍然会对缺陷识别精度带来一定影响。

2.4　基于布拉格光纤光栅技术的复合材料制件 AFP 成型过程在线检测系统

自动铺丝工艺过程中的过程监测系统大多采用激光成像系统、可见光成像系统及热成像系统，然而这些监测方法只能识别制件 AFP 成型过程中铺放表面的缺陷，无法监测到制件固化或服役过程中内部的变化。早期，研究人员尝试通过植入应变片监测结构内部应变的变化，然而该传感器容易失效，且会严重影响制件性能，难以应用在实际工程环境中[79-80]。布拉格光纤光栅（FBG）传感器作为光纤传感器中的一种，可使光在光纤传播过程中产生折射及反射，并通过监控该反射光波长变化计算光纤周围环境的应变及温度，具有体积小、质量轻、抗电磁干扰能力强、可远程操作的优势，在近年的无损检测应用中备受青睐[81-87]。

Oromiehie 等[88-95] 基于 FBG 传感器搭建了 AFP 在线检测系统，通过在某一铺层铺放结束后埋入 FBG 传感器，监测制件后续铺设过程中内部压力及温度变化，识别铺放过程中的缺陷，其系统结构如图 11 所示。研究人员通过实验发现，FBG 传感器能够监测到制件在 AFP 过程中残余应力的形成，并通过带有多种缺陷（间隙、重叠、异物）的层合板与无缺陷样本，发现缺陷种类与尺寸会对 FBG 波长响应曲线的轮廓及斜率造成不同的影响，证明 FBG 传感器可用于在线识别 AFP 过程制件内部缺陷的形成[88-92]。

FBG 反射光波长受应变及温度的共同影响，而在上述实验中，系统利用单根 FBG 传感器监测制件 AFP 成型过程，没有实现制件内部温度及应

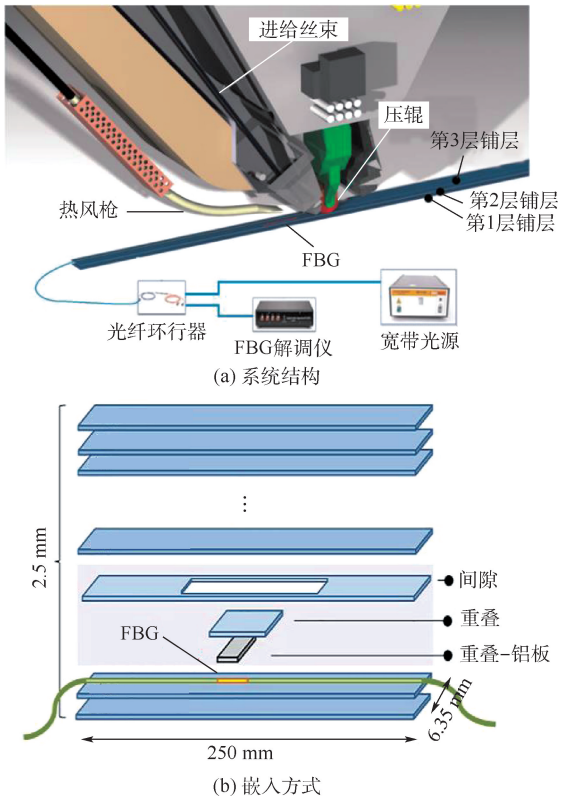


图 11 基于 FBG 传感器的 AFP 过程在线检测系统[88-95]

Fig. 11 AFP online monitoring system based on FBG sensors[88-95]

变的单独定量测量。针对这一问题,研究人员将 2 个应变敏感度不同的 FBG 传感器交错布置,通过与有限元分析结果对比验证其应变测量准确率在容许范围内。此外,研究人员指出可利用嵌入的 FBG 传感器监测制件服役过程中的声信号,并利用快速傅里叶变换算法对数据进行处理,实现制件在服役过程中的结构健康监测[93-95]。

与其他在线检测技术相比,光纤监测系统能够实现制件内部全生命周期的在线监测,且 FBG 波长响应信息作为一维信号,对数据存储要求较小,具有低成本、小尺寸、轻便、实时性高、能够实现远程操控等优点。然而,目前基于光纤传感器的 AFP 过程在线检测系统仍处在实验室阶段,其面临的主要挑战如下:

1) 测量精度不足。上述实验中,研究人员通过与有限元分析结果作对比,虽然应变测量准确率在容许范围内,但最大误差达到将近 20%。

2) 存活率不足。与应变片相比,FBG 传感器的存活率大幅提高,然而该传感器的可靠性仍受 AFP 成型过程中温度及压力的影响。

3) 只能实现局部测量。FBG 传感器只能实现局部点的测量,不能提供反映 AFP 过程中成型构件全区域的应变信息,尽管一根光纤可以集成多个 FBG 传感器,但是目前并没有相关研究及应用。

4) 影响制件性能。尽管与应变片相比,嵌入式 FBG 传感器对复合材料层压板结构完整性影响不大,但当嵌入方向垂直于纤维方向时,会导致纤维周围的应力集中,从而影响结构性能。

2.5 基于其他检测技术的复合材料制件 AFP 成型过程在线检测系统

除了上述几种针对复合材料成型过程的在线检测系统,也有许多研究人员针对以上监测系统存在的问题,探究了其他传感器用于 AFP 过程监测的可能性,如应变计、涡流传感器、分布式光纤传感器、触觉传感器等[96-99]。

研究表明,缺陷的形成与演化是由制造过程中多个应力状态及其动态波的耦合效应引起的,且缺陷的存在会进一步影响应力波的传播特性[100-101]。利用这一特性,Han 等[96]将压辊作为激励源,通过应变计实时监测 AFP 过程中的应力波传播信号,结合傅里叶变换及小波变换分析应力波特征指标(如平均应力、应力峰值),建立了孔隙密度与超声衰减系数之间的关系,用于评估制件固化前的孔隙密度。

针对基于可见光图像识别的 AFP 成型过程在线检测系统易受环境光源及材料外观影响的问题,Schmidt 等[97]搭建了基于涡流检测的 AFP 过程在线检测系统,结果表明,除了纤维方向外,该系统能够识别丝束间隙、重叠、异物及搭接等缺陷。然而,由于涡流传感器的形状及扫描速度的限制,该传感器仍未集成到自动铺丝头中。

与 FBG 传感器不同,分布式光纤传感器可以实现沿光纤方向的环境温度及应变的高精度多点检测。Shamsuddoha 等[98]结合该传感器与数字图像相关法(DIC)监控带切口热塑性复合材料管在 AFP 工艺成型过程中及在双轴加载情况下的位移及机械应变数据,通过与数值仿真结果比较验证了其准确率,但该系统仍处在实验室阶段,且与 FBG 传感器相同,存在存活率不足与影响制件性能等问题。

在过去的三十多年,触觉传感器作为机器人领域中重要的组成部分,主要用于检测仅通过视觉难以检测到的特性,如接触力或接触表面 3D 形状。欧洲空客公司[102-103]提出了在压辊内集成压电传感器以实时反馈 AFP 或 ATL 过程中的局部压力分布,但是这些设计仍然停留在概念阶段,缺乏实际的应用验证。Elkington 等[99]制作了集成触觉传感器的压辊,在压实过程中通

过集成的 LED 灯及工业相机捕获柔性压辊内部光学标记的位置信息,利用最近邻方法对数据进行处理,成功实现了 AFP 过程中的铺层表面异常检测,如褶皱、搭接、异物、间隙、重叠、缺丝的在线识别。

本文总结了部分复合材料制件 AFP 成型过程在线检测原理和系统的适用缺陷类型、测量参数、测量精度和数据采集速率,结果如表 1 所示。

表 1　各类 AFP 检测系统技术细节汇总

Table 1　Technical review on online monitoring systems for AFP process

测试技术	适用缺陷种类	测量参数	测试精度	数据采集速率
激光技术[37-53]	间隙、重叠、扭转、搭接、异物、折叠、缺丝	物体表面轮廓	0.06 ~ 0.25 mm	<1 kHz
可见光图像识别[54-62]	间隙、重叠	表面图像	0.04 ~ 0.92 mm	
热成像技术[68-79]	间隙、重叠、扭转、搭接、异物、褶皱、缺丝	表面温度场	0.18 ~ 0.44 mm(空间分辨率)	30 ~ 60 Hz
布拉格光纤光栅技术[89-94]	间隙、重叠、异物	应变及温度	4.29% ~ 19.46%(应变误差)	1 kHz
涡流传感技术[98]	间隙、重叠、搭接、异物	电阻抗	0.3 mm(空间分辨率)	
触觉传感技术[100]	间隙、重叠、搭接、异物、褶皱、缺丝	物体表面轮廓		

3　结 束 语

铺放过程中出现的各类缺陷不断累积,严重影响了复合材料结构的机械性能及疲劳寿命,限制了 AFP 技术的应用发展。为了有效提升制件成型过程中的质量,国外大量研究人员基于不同技术对复材制件自动铺放成型过程的在线检测方法进行了深入的研究,其中激光、可见光图像、热成像等技术已成功应用。然而自动铺丝工艺在中国发展起步较晚,对铺放过程监控技术的研究仍处于起步阶段,主要依赖操作者目视进行铺放成型过程的缺陷检测。

研究表明,目前各类监测系统具有不同的优势,可结合计算机成像、边缘检测或神经网络等方法实现多类缺陷的识别及铺层质量的评估。然而,这些检测系统在实际的工程应用中易受到设备、环境、数据及材料等因素的影响,存在检测精度不足、易受环境影响、应用场景受限、实时性差等问题。针对以上问题,未来可从以下 3 个方面开展研究:

1) 通过不同检测技术的结合实现制件在 AFP 成型过程中的在线检测。基于激光技术或热成像技术的 AFP 过程在线检测系统能够有效检测到制件铺放过程中的表面缺陷;而利用布拉格光纤光栅、分布式光纤传感器及应变计可以实现对 AFP 过程中制件内部缺陷的演变的监测。因此,结合不同传感器的不同优势,开发基于多传感器的在线缺陷检测系统,将成为未来在线检测和控制缺陷的研究重点。

2) 结合机器学习算法实现制件在 AFP 成型过程中的在线检测。近年神经网络已经在各种工程应用中得到了普及,包括预测、控制和故障诊断,而针对 AFP 过程的在线检测,已有不少学者利用神经网络模型实现铺层表面缺陷的在线识别,然而这些研究工作仍处于起步阶段。因此,应密切关注人工智能技术的发展,结合新型神经网络模型,进一步提高铺层表面缺陷的在线识别精度。

3) 深入研究不同因素对缺陷形成及演变的影响。目前已有的在线检测技术可实现铺层表面缺陷的有效识别,但是无法根据设备、材料、工艺参数等因素的变化实现未来铺层质量的预测。因此,深入分析不同因素对缺陷形成及演变的影响机理,有助于实现铺放过程中结构缺陷形成及演变的预测,进一步提高在线检测系统识别与评估的准确率,为后续的工艺参数调控、成型质量控制打下基础。

参考文献 (References)

[1] 陈祥宝. 先进复合材料技术导论[M]. 北京:航空工业出版社, 2017.
CHEN X B. Introduction of advanced composite technology [M]. Beijing: Aviation Industry Press, 2017 (in Chinese).

[2] 邢丽英,包建文,礼嵩明,等. 先进树脂基复合材料发展现状和面临的挑战[J]. 复合材料学报, 2016, 33(7): 1327-1338.
XING L Y, BAO J W, LI S M, et al. Development status and facing challenge of advanced polymer matrix composites[J]. Acta Materiae Compositae Sinica, 2016, 33(7): 1327-1338 (in Chinese).

[3] 单忠德,战丽,缪云良,等. 复合材料构件数字化精确成形技术与装备[J]. 科技导报, 2020, 38(14): 63-67.
SHAN Z D, ZHAN L, MIAO Y L, et al. Digital precision forming technology and equipment for composite materials[J]. Science and Technology Review, 2020, 38(14): 63-67 (in Chinese).

[4] BRASINGTON A, SACCO C, HALBRITTER J, et al. Automated fiber placement: A review of history, current technologies, and future paths forward [J]. Composites Part C: Open Access, 2021, 6: 100182.

[5] BARILE M, LECCE L, ESPOSITO M, et al. Advanced AFP applications on multifunctional composites[C]// The Second International Symposium on Automated Composites Manufacturing, 2015: 23-24.

[6] QURESHI Z, SWAIT T, SCAIFE R, et al. In situ consolidation of thermoplastic prepreg tape using automated tape placement technology: Potential and possibilities[J]. Composites Part B: Engineering, 2014, 66: 255-267.

[7] HARIK R, SAIDY C, WILLIAMS S J, et al. Automated fiber placement defect identity cards: Cause, anticipation, existence, significance, and progression[C]// SAMPE 2018, 2018.

[8] LOPES C S, CAMANHO P P, GÜRDAL Z, et al. Progressive failure analysis of tow-placed, variable-stiffness composite panels[J]. International Journal of Solids and Structures, 2007, 44(25-26): 8493-8516.

[9] SHUART M J. Failure of compression-loaded multidirectional composite laminates[J]. AIAA Journal, 1989, 27(9): 1274-1279.

[10] NOEVERE A T, COLLIER C S. Design for manufacturing tool for automated fiber placement structures-verification and validation: AIAA 2020-1477[R]. Reston: AIAA, 2020.

[11] SAWICKI A J, MINGUET P J. The effect of intraply overlaps and gaps upon the compression strength of composite laminates: AIAA-98-1786[R]. Reston: AIAA, 1998.

[12] CROFT K, LESSARD L, PASINI D, et al. Experimental study of the effect of automated fiber placement induced defects on performance of composite laminates[J]. Composites Part A: Applied Science and Manufacturing, 2011, 42(5): 484-491.

[13] LAN M, CARTIÉ D, DAVIES P, et al. Influence of embedded gap and overlap fiber placement defects on the microstructure and shear and compression properties of carbon-epoxy laminates [J]. Composites Part A: Applied Science and Manufacturing, 2016, 82: 198-207.

[14] LAN M, CARTIÉ D, DAVIES P, et al. Microstructure and tensile properties of carbon-epoxy laminates produced by automated fibre placement: Influence of a caul plate on the effects of gap and overlap embedded defects[J]. Composites Part A: Applied Science and Manufacturing, 2015, 78: 124-134.

[15] NGUYEN M H, VIJAYACHANDRAN A A, DAVIDSON P, et al. Effect of automated fiber placement (AFP) manufacturing signature on mechanical performance of composite structures [J]. Composite Structures, 2019, 228: 111335.

[16] MAROUENE A, LEGAY P, BOUKHILI R. Experimental and numerical investigation on the open-hole compressive strength of AFP composites containing gaps and overlaps[J]. Journal of Composite Materials, 2017, 51(26): 3631-3646.

[17] MUKHOPADHYAY S, JONES M I, HALLETT S R. Tensile failure of laminates containing an embedded wrinkle: Numerical and experimental study[J]. Composites Part A: Applied Science and Manufacturing, 2015, 77: 219-228.

[18] MUKHOPADHYAY S, JONES M I, HALLETT S R. Compressive failure of laminates containing an embedded wrinkle: Experimental and numerical study[J]. Composites Part A: Applied Science and Manufacturing, 2015, 73: 132-142.

[19] NARTEY M, ZHANG T, GONG B, et al. Understanding the impact of fibre wrinkle architectures on composite laminates through tailored gaps and overlaps[J]. Composites Part B: Engineering, 2020, 196: 108097.

[20] MAUNG P, PRUSTY B G, WHITE J M, et al. Structural performance of a shape-adaptive composite hydrofoil using automated fibre placement[J]. Engineering Structures, 2019, 183: 351-365.

[21] MAUNG P T, PRUSTY B G, PHILLIPS A W, et al. Curved fibre path optimisation for improved shape adaptive composite propeller blade design[J]. Composite Structures, 2021, 255: 112961.

[22] ROBERT C, PECUR T, MAGUIRE J M, et al. A novel powder-epoxy towpregging line for wind and tidal turbine blades[J]. Composites Part B: Engineering, 2020, 203: 108443.

[23] RHEAD A T, DODWELL T J, BUTLER R. The effect of tow gaps on compression after impact strength of robotically laminated structures[J]. Computers, Materials and Continua, 2013, 35(1): 1-16.

[24] ELSHERBINI Y M, HOA S V. Experimental and numerical investigation of the effect of gaps on fatigue behavior of unidirectional carbon/epoxy automated fiber placement laminates[J]. Journal of Composite Materials, 2017, 51(6): 759-772.

[25] ELSHERBINI Y M, HOA S V. Fatigue threshold-stress determination in AFP laminates containing gaps using IR thermography [J]. Composites Science and Technology, 2017, 146: 49-58.

[26] CARTIÉ D, LAN M, DAVIES P, et al. Influence of embedded gap and overlap fiber placement defects on interlaminar properties of high performance composites[J]. Materials, 2021, 14(18): 5332.

[27] QIAN S, LIU X, YE Y, et al. Effect of gap and overlap fiber placement defects on the delamination behavior of L-shaped composite laminates[J]. Composite Structures, 2021, 268: 113963.

[28] DAVIDSON P, WAAS A M. The effects of defects on the compressive response of thick carbon composites: An experimental and computational study[J]. Composite Structures, 2017, 176: 582-596.

[29] MUKHOPADHYAY S, NIXON-PEARSON O J, HALLETT S R. An experimental and numerical study on fatigue damage development in laminates containing embedded wrinkle defects [J]. International Journal of Fatigue, 2018, 107: 1-12.

[30] XU X, JONES M I, ALI H, et al. Effect of out-of-plane wrinkles in curved multi-directional carbon/epoxy laminates[J]. Composites Science and Technology, 2020, 197: 108282.

[31] LUKASZEWICZ D H J A, POTTER K D, EALES J. A concept for the in situ consolidation of thermoset matrix prepreg during automated lay-up[J]. Composites Part B: Engineering, 2013, 45(1): 538-543.

[32] OROMIEHIE E, PRUSTY B G, COMPSTON P, et al. Automated fibre placement based composite structures: Review on the defects, impacts and inspections techniques[J]. Composite Structures, 2019, 224: 110987.

[33] CAI J, QIU L, YUAN S, et al. Structural health monitoring for composite materials[M/OL]// HU N. Composites and their applications. (2012-08-22)[2022-04-01]. https://www.intechopen.com/chapters/38395.

[34] GUIN W E,JACKSON J R,BOSLEY C M. Effects of tow-to-tow gaps in composite laminates fabricated via automated fiber placement[J]. Composites Part A:Applied Science and Manufacturing,2018,115:66-75.

[35] HALBRITTER A,HARPER R. Big parts demand big changes to the fiber placement status quo[R]. [S. l.]:SME Composites Manufacturing,2012.

[36] 马少博,文立伟,王若舟,等. 复合材料自动铺放表面缺陷检测技术研究进展[J]. 复合材料科学与工程,2020(12):121-128.
MA S B,WEN L W,WANG R Z,et al. A review of surface defect inspection technology of automated fiber placement manufacturing[J]. Science and Engineering of Composite Materials,2020(12):121-128(in Chinese).

[37] RUDBERG T,NIELSON J,HENSCHEID M,et al. Improving AFP cell performance[J]. SAE International Journal of Aerospace,2014,7(2):317-321.

[38] RUDBERG T,CEMENSKA J. Incorporation of laser projectors in machine cell controller reduces ply boundary inspection time,on-part course identification and part probing[J]. SAE International Journal of Aerospace,2012,5(1):74-78.

[39] SHADMEHRI F,IOACHIM O,PAHUD O,et al. Laser-vision inspection system for automated fiber placement (AFP) process[C]//20th International Conference on Composite Materials,2015.

[40] SHADMEHRI F. Laser vision inspection system and method:US 20170248406A1[P]. 2017-08-31.

[41] KITSON L E. Composite material laser flaw detection: US 05562788A[P]. 1996-10-08.

[42] OLDANI T. Visual fiber placement inspection:US07835567B2[P]. 2010-11-16.

[43] OLDANI T. Visual fiber placement inspection:US20110017381A1[P]. 2011-01-27.

[44] ZAMBAL S,HEINDL C,EITZINGER C,et al. End-to-end defect detection in automated fiber placement based on artificially generated data[C]//14th International Conference on Quality Control by Artificial Vision,2019:11172.

[45] MAASS D. Progress in automated ply inspection of AFP layups[J]. Reinforced Plastics,2015,59(5):242-245.

[46] CEMENSKA J,RUDBERG T,HENSCHEID M,et al. AFP automated inspection system performance and expectations[R]. [S. l.]:SAE,2017:1-8.

[47] CEMENSKA J,RUDBERG T,HENSCHEID M. Automated in-process inspection system for AFP machines[J]. SAE International Journal of Aerospace,2015,8(2):303-309.

[48] MEISTER S. Automated defect analysis using optical sensing and explainable artificial intelligence for fibre layup process in composite manufacturing[D]. Delft:Delft University of Technology,2022.

[49] MEISTER S,WERMES M,STÜVE J,et al. Investigations on explainable artificial intelligence methods for the deep learning classification of fibre layup defect in the automated composite manufacturing[J]. Composites Part B:Engineering,2021,224:109160.

[50] MEISTER S,MÖLLER N,STÜVE J,et al. Synthetic image data augmentation for fibre layup inspection processes:Techniques to enhance the data set[J]. Journal of Intelligent Manufacturing,2021,32(6):1767-1789.

[51] SACCO C,RADWAN A B,HARIK R,et al. Automated fiber placement defects:Automated inspection and characterization[C]//SAMPE 2018,2018.

[52] SACCO C. Machine learning methods for rapid inspection of automated fiber placement manufactured composite structures[D]. Columbia:University of South Carolina,2019.

[53] SACCO C,RADWAN A B,BEATTY T,et al. Machine learning based AFP inspection:A tool for characterization and integration[C]//SAMPE 2019,2019.

[54] SOUCY K A. In-process monitoring for quality assurance of automated composite fabrication[M]//THOMPSON D O,CHIMENTI D E. Review of progress in quantitative nondestructive evaluation. Berlin:Springer,1996:2225-2231.

[55] ENGELBART R. System and method for identifying defects in a composite structure:US07171033B2[P]. 2007-01-30.

[56] 陈文,曹力,肖军,等. 基于视觉的复合材料自动铺放表面纹理实时检测方法[J]. 科学技术与工程,2011,11(18):4143-4149.
CHEN W,CAO L,XIAO J,et al. A vision-based real-time detection method of texture in automatic composite fiber placement[J]. Science Technology and Engineering,2011,11(18):4143-4149(in Chinese).

[57] 倪金辉,肖军,文立伟. 一种基于直线拟合的预浸料边缘直线在线视觉检测方法[J]. 计算机科学,2015,42(Z6):16-19.
NI J H,XIAO J,WEN L W. Real-time visual inspection method to detect prepreg edge straight based on line fitting[J]. Computer Science,2015,42(Z6):16-19(in Chinese).

[58] 刘辉. 基于机器视觉的纤维铺放角度及间距检测系统研究[D]. 武汉:武汉理工大学,2008.
LIU H. The research of image detection system of fiber tape laying based on machine vision[D]. Wuhan:Wuhan University of Technology,2008(in Chinese).

[59] TAO Y,JIA S,DUAN Y,et al. An online detection method for composite fibre tow placement accuracy[J]. Proceedings of the Institution of Mechanical Engineers,Part B:Journal of Engineering Manufacture,2016,230(9):1614-1621.

[60] ENGELBART R W. Verification of tow cut for automatic fiber placement:EP2072224B1[P]. 2012-02-08.

[61] ENGELBART R W. Mapping tow splices in composite structures:US20120216957A1[P]. 2012-08-30.

[62] RITTER J A. Real-time infrared thermography inspection and control for automated composite marterial layup: US 07513964B2[P]. 2009-04-07.

[63] ENGELBART R W. Systems and methods enabling automated return to and/or repair of defects with a material placement machine:US07039485B2[P]. 2006-05-02.

[64] ENGELBART R W. Systems and methods for using light to indicate defect locations on a composite structure:EP05735387A[P]. 2005-04-11.

[65] ENGELBART R W. Systems and method for identifying foreign objects and debris (FOD) and defects during fabrication of a

composite structure: US20050025350A1[P]. 2005-02-03.

[66] ENGELBART R W. In-process vision detection of flaws and FOD by back field illumination: US07678214B2[P]. 2010-03-16.

[67] CHEN M, JIANG M, LIU X, et al. Intelligent inspection system based on infrared vision for automated fiber placement[C]// Proceedings of 2018 IEEE International Conference on Mechatronics and Automation. Piscataway: IEEE Press, 2018: 918-923.

[68] DENKENA B, SCHMIDT C, WEBER P. Automated fiber placement head for manufacturing of innovative aerospace stiffening structures[J]. Procedia Manufacturing, 2016, 6: 96-104.

[69] DENKENA B, SCHMIDT C, VÖLTZER K, et al. Thermographic online monitoring system for automated fiber placement processes[J]. Composites Part B: Engineering, 2016, 97: 239-243.

[70] SCHMIDT C, DENKENA B, VÖLTZER K, et al. Thermal image-based monitoring for the automated fiber placement process[J]. Procedia CIRP, 2017, 62: 27-32.

[71] SCHMIDT C, DENKENA B, HOCKE T, et al. Influence of AFP process parameters on the temperature distribution used for thermal in-process monitoring[J]. Procedia CIRP, 2017, 66: 68-73.

[72] SCHMIDT C, HOCKE T, DENKENA B. Deep learning-based classification of production defects in automated-fiber-placement processes[J]. Production Engineering, 2019, 13(3-4): 501-509.

[73] SCHMIDT C, HOCKE T, DENKENA B. Artificial intelligence for non-destructive testing of CFRP prepreg materials[J]. Production Engineering, 2019, 13(5): 617-626.

[74] JUAREZ P D, GREGORY E D. In situ thermal inspection of automated fiber placement for manufacturing induced defects[J]. Composites Part B: Engineering, 2021, 220: 109002.

[75] JUAREZ P D, GREGORY E D. In situ thermal inspection of automated fiber placement manufacturing[C]// AIP Conference, 2019: 120005.

[76] JUAREZ P D, CRAMER K E, SEEBO J P. Advances in in-situ inspection of automated fiber placement systems[C]// Thermosense: Thermal Infrared Applications, 2016: 986109.

[77] GREGORY E D, JUAREZ P D. In-situ thermography of automated fiber placement parts[C]// AIP Conference Proceedings, 2018: 060005.

[78] GREGORY E D, JUAREZ P D. In situ thermal nondestructive evaluation for assessing part quality during composite automated fiber placement[J]. Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems, 2018, 1(4): 1-6.

[79] SHOKRIEH M M, MOHAMMADI A R G. Non-destructive testing (NDT) techniques in the measurement of residual stresses in composite materials: An overview[M]// SHOKRIEH M M. Residual stresses in composite materials. Amsterdam: Elsevier, 2014: 58-75.

[80] PARLEVLIET P P, BERSEE H E N, BEUKERS A. Residual stresses in thermoplastic composites-A study of the literature. Part I: Formation of residual stresses[J]. Composites Part A: Applied Science and Manufacturing, 2006, 37(11): 1847-1857.

[81] BAEJEWSKI W, GSIOR P, KALET J. Application of optical fibre sensors to measuring the mechanical properties of composite materials and structures[M/OL]// ATTAF B. Advances in composite materials-Ecodesign and analysis. (2011-03-16)[2022-04-01]. https://www.intechopen.com/chapters/14299.

[82] GROS X E. Current and future trends in non-destructive testing of composite materials[J]. Annales De Chimie Science Des Matériaux, 2000, 25(7): 539-544.

[83] MAJUMDER M, GANGOPADHYAY T K, CHAKRABORTY A K, et al. Fibre Bragg gratings in structural health monitoring-Present status and applications[J]. Sensors and Actuators, A: Physical, 2008, 147(1): 150-164.

[84] MURUKESHAN V M, CHAN P Y, ONG L S, et al. Cure monitoring of smart composites using fiber Bragg grating based embedded sensors[J]. Sensors and Actuators, A: Physical, 2000, 79(2): 153-161.

[85] LENG J, ASUNDI A. Structural health monitoring of smart composite materials by using EFPI and FBG sensors[J]. Sensors and Actuators, A: Physical, 2003, 103(3): 330-340.

[86] RUZEK R, KUDRNA P, KADLEC M, et al. Strain and damage monitoring in CFRP fuselage panels using fiber Bragg grating sensors. Part II: Mechanical testing and validation[J]. Composite Structures, 2014, 107: 737-744.

[87] OROMIEHIE E, RAJAN G, PRUSTY G B. Thermal sensitivity and relaxation of carbon fibre-foam sandwich composites with fibre optic sensors[J]. Journal of Sandwich Structures and Materials, 2016, 18(5): 652-664.

[88] OROMIEHIE E, PRUSTY G B, COMPSTON P, et al. Structural health monitoring and processing of composites using photonic sensing technology[C]// Proceedings of the 20th International Conference on Composite Materials, 2015: 19-24.

[89] OROMIEHIE E. In-situ process monitoring in automated fibre placement-based manufacturing of advanced composites[D]. Sydney: University of New South Wales, 2017.

[90] OROMIEHIE E, PRUSTY B G, RAJAN G, et al. Optical fiber Bragg grating sensors for process monitoring in advanced composites[C]// 2016 IEEE Sensors Applications Symposium (SAS). Piscataway: IEEE Press, 2016: 1-5.

[91] OROMIEHIE E, PRUSTY B G, COMPSTON P, et al. In situ process monitoring for automated fibre placement using fibre Bragg grating sensors[J]. Structural Health Monitoring, 2016, 15(6): 706-714.

[92] OROMIEHIE E, PRUSTY B G, COMPSTON P, et al. Characterization of process-induced defects in automated fiber placement manufacturing of composites using fiber Bragg grating sensors[J]. Structural Health Monitoring, 2018, 17(1): 108-117.

[93] OROMIEHIE E, PRUSTY G B, COMPSTON P, et al. In-situ simultaneous measurement of strain and temperature in automated fiber placement (AFP) using optical fiber Bragg grating (FBG) sensors[J]. Advanced Manufacturing: Polymer and Composites Science, 2017, 3(2): 52-61.

[94] OROMIEHIE E, PRUSTY B G, RAJAN G, et al. Machine learning based process monitoring and characterisation of automated composites[C]// SAMPE 2017, 2017: 398-410.

[95] OROMIEHIE E, DAS CHAKLADAR N, RAJAN G, et al. Online monitoring and prediction of thermo-mechanics of AFP based thermoplastic composites[J]. Sensors (Basel), 2019, 19(6):1310.

[96] HAN Z, SUN S, LI W, et al. Experimental study of the effect of internal defects on stresswaves during automated fiber placement[J]. Polymers, 2018, 10(4):1-14.

[97] SCHMIDT C, SCHULTZ C, WEBER P, et al. Evaluation of eddy current testing for quality assurance and process monitoring of automated fiber placement[J]. Composites Part B: Engineering, 2014, 56:109-116.

[98] SHAMSUDDOHA M, DAVID M, OROMIEHIE E, et al. Distributed optical fibre sensor based monitoring of thermoplastic carbon composite cylinders under biaxial loading: Experimental and numerical investigations[J]. Composite Structures, 2021, 261:113277.

[99] ELKINGTON M, ALMAS E, WARD-CHERRIER B, et al. Real time defect detection during composite layup via tactile shape sensing[J]. Science and Engineering of Composite Materials, 2021, 28(1):1-10.

[100] ZHANG H, HUA J, GAO F, et al. Efficient Lamb-wave based damage imaging using multiple sparse Bayesian learning in composite laminates[J]. NDT and E International, 2020, 116: 102277.

[101] CAIRNS D S, NELSON J W, WOO K, et al. Progressive damage analysis and testing of composite laminates with fiber waves[J]. Composites Part A: Applied Science and Manufacturing, 2016, 90:51-61.

[102] SCHLEGEL K, FRIEDBERGER A, LOPEZ-BRAVO D. Fiber composite laying device and fiber composite laying method for producing a fiber composite scrim for forming a fiber composite component: US161033300[P]. 2019-01-17.

[103] SCHLEGEL K, ALOIS F, DAVID L B. Fiber composite laying device and fiber composite laying method for producing a fiber composite scrim for forming a fiber composite component: US202210032561A1[P]. 2022-02-03.

Review on automated fiber placement induced defects and their online monitoring technology

DING Xilun[1,2], LUO Weiheng[1], LIU Fei[1], WANG Guodong[3], CHEN Weiqiang[3], ZHANG Wuxiang[1,2,*]

(1. School of Mechanical Engineering and Automation, Beihang Unviersity, Beijing 100083, China;
2. Ningbo Innovation Research Institute of Beihang University, Ningbo 315800, China;
3. Beijing Satellite Manufacturing Factory Co., Ltd., Beijing 100190, China)

Abstract: The automated fiber placement (AFP) process has become a promising solution to the fabrication of composite structures, due to its efficiency and stability. However, the instability of the lay-up process parameters and the performance of the equipment can easily cause such defects as gaps, overlaps, and waviness, which severely affect the overall performance and service life of composite structures. Therefore, automatic identification and evaluation of the defects on the surface or inside the laminates during the AFP process through in-situ monitoring systems are of particular importance. In this paper, the causes and appearance of common defects on the ply during the AFP process and their effects on the overall performance of products are first concluded. Then, the advantages, development, and applications of online monitoring systems based on different detection techniques are systematically reviewed, mainly including laser technology, image recognition technology, thermal imaging technology, and fiber Bragg grating technology. Finally, the limitations in the current online monitoring systems are summarized, and outlooks for their future trends are given.

Keywords: fiber reinforced plastics; automated fiber placement (AFP); defects; online monitoring technology; identification methods

Received: 2022-04-30; **Accepted**: 2022-05-18; **Published online**: 2022-06-16 17:29

URL: kns.cnki.net/kcms/detail/11.2625.V.20220616.1430.001.html

Foundation items: Zhejiang Provincial Natural Science Foundation of China (LD22E050011); Ningbo Key Projects of Science and Technology Innovation 2025 Plan (2022Z070)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0593

面向智能化管理的数字孪生电池构建方法

杨世春, 李强伟, 周思达, 张正杰, 马源, 陈飞*

(北京航空航天大学 交通科学与工程学院, 北京 100083)

摘 要: 动力电池在长时宽温域运行时,其使用性能、寿命和安全性随时间动态演变,存在单体性能不一致、系统容量快速衰减或内部缺陷诱发的电池热失控等问题,需要全气候、全生命周期电池精准管理技术。突破智能化管理的数字孪生技术、构建数字孪生电池为提升电池管理能力带来了新的解决方案,已逐步成为行业发展趋势之一。围绕动力电池精细化管理技术发展趋势,针对数字孪生动力电池构建需求,从系统建模与管控需求等方面分析了数字孪生电池建模的基本准则,系统性阐述多维度、多尺度、多物理场融合的数字孪生电池的构建方法,并结合团队前期研究分析了某电池数字孪生的实践案例,探索了数字孪生电池在生产设计、全生命周期管理等场景下的应用可能性,为电池管理技术发展提供思路与参考。

关 键 词: 新能源汽车; 动力电池; 数字孪生; 电池模型; 电池管理
中图分类号: U469
文献标志码: A **文章编号**: 1001-5965(2022)09-1734-11

发展新能源汽车是中国从汽车大国迈向汽车强国的必经之路。在“十四五”规划建议中,新能源汽车已被列入十大战略性新兴产业之一。中国的新能源汽车产销量已连续 7 年位居全球首位;国务院印发的《新能源汽车产业发展规划(2021—2035 年)》中指出,到 2025 年,中国新能源汽车新车销售量达到汽车新车销售总量的 20% 左右。动力电池是新能源汽车能源动力系统的核心,其性能影响了整车安全性、寿命、经济性。然而,动力电池在全气候环境下的使用性能、寿命和安全性随时间动态演变,产生了单体性能不一致诱发的系统容量快速衰减、短板电池或内部缺陷诱发的电池热失控等问题[1-2]。

动力电池管理技术通过对电池的使用性能进行评估、合理的规划电池充放电过程等,改善动力电池寿命和安全性[3]。传统动力电池管理方法受限于车载硬件单元的有限计算/存储能力,只能实现简单的性能监管、状态估算等功能,准确度差、可靠性低[4]。随着数字孪生、虚拟现实、云端管控等数字化管理技术的发展,为动力电池管理技术发展提供了新契机和新思路[5-6]。数字孪生是一种全新的建模思想和方法,基于通信、仿真、数据分析等前沿技术,通过构建物理实体的逼真仿真模型及数字化实时状态映射,实现对物理实体全生命周期的管理决策和进化学习[7-9]。数字孪生在卫星、海洋、林业、交通等领域已开展了大量探索式研究和示范应用[10-12];在中国“十四五”规划中的新能源汽车重点专项等也阐述了数字孪生在汽车、动力电池行业的应用。发展数字孪生技术、构建数字孪生电池为提升电池管理能力带来了新的解决方案,已逐步成为行业发展趋势之一,宁德时代、比亚迪、国轩高科等电池企业均已开展前瞻性技术研究[13-14]。但目前对于动力电池的数字孪生仍缺乏系统的建模理论、评价方法、

收稿日期: 2022-07-07; **录用日期**: 2022-08-05; **网络出版时间**: 2022-08-12 13:56
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220812.1323.002.html
基金项目: 国家自然科学基金 (U1864213)
* **通信作者**. E-mail: cf2020@buaa.edu.cn

引用格式: 杨世春, 李强伟, 周思达, 等. 面向智能化管理的数字孪生电池构建方法[J]. 北京航空航天大学学报, 2022, 48(9): 1734-1744. YANG S C, LI Q W, ZHOU S D, et al. Construction of digital twin model of lithium-ion battery for intelligent management[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1734-1744(in Chinese).

应用范围等。本文围绕动力电池精细化管理技术发展趋势，针对数字孪生动力电池构建需求，系统性阐述动力电池数字孪生模型的构建理论，探索了其在电池设计、全生命周期管理、梯次利用等场景下的应用可能性。希望通过本文对电池数字孪生模型的详细介绍和解释，为电池行业发展提供思路和理论支撑。

1　动力电池数字孪生模型概念

模型是对物理实体的数学描述，可以实现物理实体的属性、特征、功能的复现和重构。模型始终是动力电池管理研究的重点，现有动力电池模型主要分为等效电路模型、电化学模型和数据驱动模型[15]。3 类模型均可实现对电池静动态特性的模拟或时空尺度上的性能仿真，但这 3 类模型均缺乏对实体电池的差异化适配和全生命周期演化，如电化学模型对初值和参数具有极大的依赖性；等效电路模型不仅依赖于大量标定实验，也无法满足全生命周期的参数自适应，精度较差；数据驱动模型对训练集具有极大的依赖性，对于超出训练集的场景泛化性能有限[16-17]。

数字孪生模型吸取上述建模的优势和缺点，通过刻画电池内部的反应过程或映射关系，对物理实体的实际运行环境、工况等进行抽象或提取，使模型在符合实际场景的环境下进行虚拟仿真，使模型无限逼近物理实体并与实体“同时空”演化，从而实现对实体电池的数字化仿真、监控、预测、优化等。如图 1 所示，数字孪生模型具有以下 3 大特征：

1）全生命周期建模。数字孪生电池的核心思想是伴随着物理实体共同演化，从而在数字空间内构建出与物理实体相映射的电池模型。该模型需要反映出电池内部的各类微观参数和宏观性能在全生命周期内的演化趋势、规律等信息，并且具备对未来演化趋势的预测和分析。例如以电化学模型为基础的数字孪生电池，应时刻保持与物理实体相一致的循环次数、状态、环境等，从而预测电池在该状态下的剩余寿命、安全状态，如在现有建模思路与框架下，假定负极电位达到 0 V 后发生析锂反应，从而在充电结束时可分析该次充电条件下的电池内部析锂量，并可用于反向推断充电策略和充电边界等。全生命周期建模思想是保证电池全寿命管理的基础。

2）对物理实体的逼近。数字孪生电池具备在全生命周期向物理实体逼近的功能，这一功能体现在实体电池在运行过程中应实时上传并反馈电池的各类信息，而后在云计算平台中对信息进行整合、归类、深入挖掘，解构电池信息与模型之间的映射关系，从而通过提取和辨识实体电池的信息来更新数字孪生电池。这使得数字孪生模型并不完全依赖初始条件，可以先对电池应用并在云计算平台内部署数字孪生模型，而后随着电池

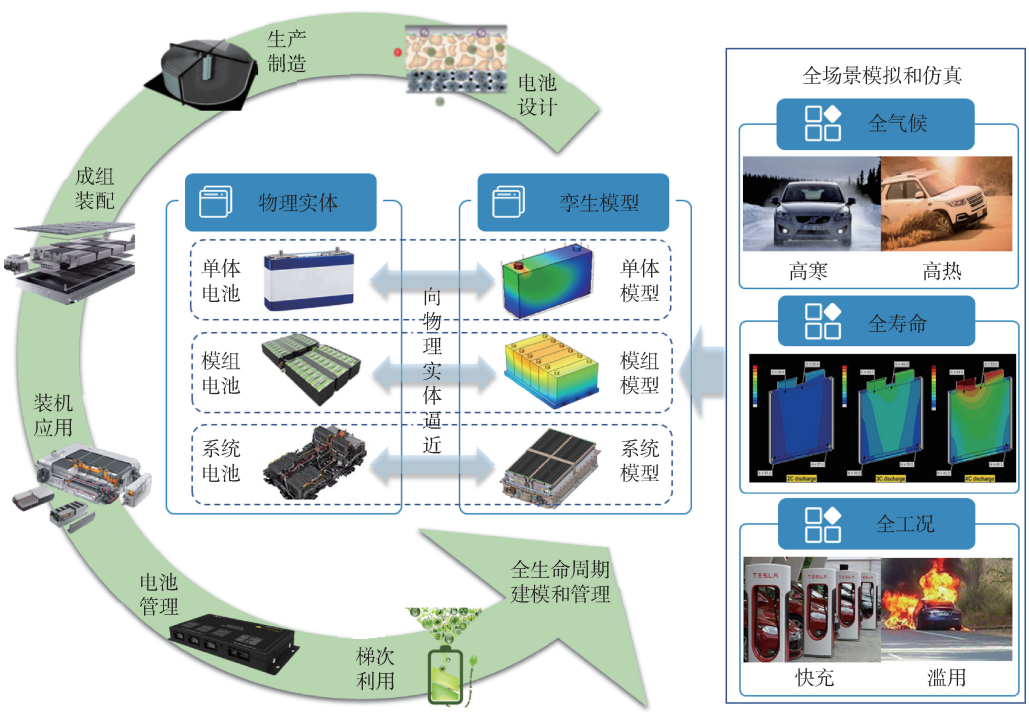


图 1　数字孪生电池三大特征

Fig. 1　Three main features of digital twin cell

的使用过程再对数字孪生电池进行优化和更新，有益于解决目前电池研发速度快和性能测试周期长之间的矛盾。

3）全场景模拟和仿真。数字孪生电池在数字空间内运行，因此其可以模拟物理实体的各类场景、工况等，并对物理实体的性能进行仿真和分析。例如，通过分析实体电池的使用工况和环境可以分析其是否存在过充、过放等可能，随后可以在云计算平台内通过数字孪生电池进行虚拟仿真，分析极端工况等对电池内部造成的损伤及危害严重性，分析其安全边界演化等。此外，数字孪生电池也可以模拟实体电池在高原、高寒、高温等环境下的仿真性能，对新能源汽车长途、跨省运行时管理策略的适配和更新提供依据。

2 数字孪生电池的发展和需求

传统车载电池管理系统以车载嵌入式系统硬件平台为核心，通过基础的逻辑功能和简单算法，能够实现较为全面的动力电池管理要求，涵盖数据监测、均衡、通信、故障诊断、充放电控制等。车规级的嵌入式系统存储空间有限（小于1 GB，通常为MB量级）、计算速度慢（主频约为500 MB/s），全生命周期数据无法存储，高精度算法、精度复杂模型等无法部署，导致现有的车端电池管理系统内部模型和算法依赖于出厂标定，无法实现全生命周期自适应演化。同时，基于电池机理或全生命周期数据的动力电池寿命预测算法与热失控预警算法也依赖于高精度的电池模型，因此车载电池管理系统无法支撑寿命、安全管理系统开发。由于存在算力和存储的限制，车载电池管理系统已无法支持更精细和全生命周期的电池管理；随着新能源汽车的大量普及，批量化的产品仍存在小概率偶发的失效问题，统计数据表明新能源汽车的起火概率达到了万分之二，如何预测预警热失控等小概率事件是电池管理亟待解决的难题[18]；此外，动力电池制造缺陷等导致电池加速老化、寿命异常衰退甚至存在容量跳水等现象，也是目前车载电池管理系统难以解决的问题之一[19-20]。

数字孪生为上述问题提供了新的解决方案，结合车云融合的云端动力电池管控平台，在平台中部署数字孪生模型及自适应算法，通过车端上传的全生命周期数据，可实现数字孪生模型的全生命周期演化与逼近物理实体，同时也为实现模型实时运行、单车单控策略设计提供了实现契机。数字孪生建模的核心是在全生命周期向物理实体逼近并随着物理实体共同演化。由于动力电池本身电化学反应原理的复杂性、车载工况的时变迟滞性等，数字孪生模型的构建难度较高，主要存在以下问题和需求：

1）动力电池内部电化学反应机理复杂，如图2所示，除充放电时在正负极发生的粒子迁移过程外，还存在多种副反应，包括负极表面锂的析出、固体电解质膜的生长/破裂、电解液分解产气、隔膜老化收缩等，导致动力电池性能存在微观模糊性、演化复杂性、实际多变性。准确描述动力电池内部各类老化反应、建立电池副反应的量化方程使精细化电池模型构建存在难题，需要进一步探索动力电池等效机理模型或数据驱动模型，揭示数字孪生构建的基础理论。

2）动力电池的微观参数和宏观性能直接映射关系模糊，如固态电解质隔膜厚度、析锂层厚度等影响电池寿命和内阻等，其量化映射关系不明晰；电池的微观参数在全生命周期的演化过程中存在随机性、与使用工况成相关性，因此揭示其在全生命周期的演化过程、构建“宏观-微观”耦合参数辨识优化算法存在难题，需要解决全生命周期下模型自学习优化机制等，揭示数字孪生构建

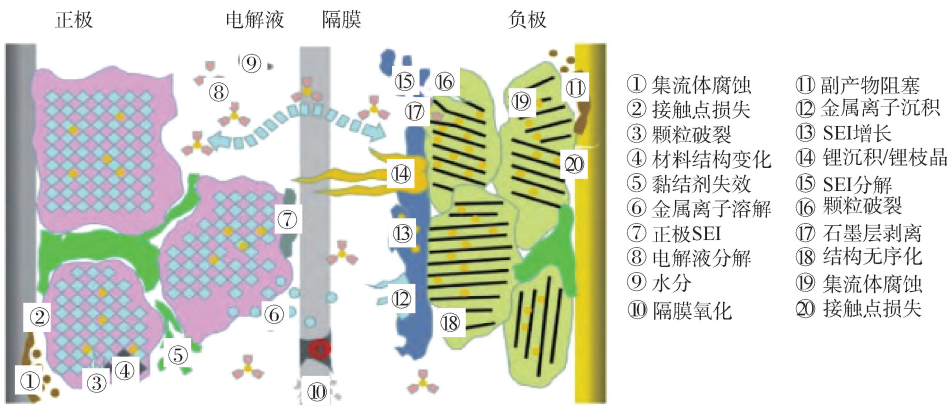


图2 动力电池内部复杂副反应过程

Fig. 2 Complex side reaction processes within power cell

的全生命周期演化理论。

3)数字孪生模型可以解决电池全生命周期的性能精准预测难题,但数字孪生模型难以直接应用到电池管理过程中,仍需要结合逻辑、控制算法等,对电池在特定工况、特定场景下的性能及演化趋势进行分析。需要与电池的控制策略相结合,构建动力电池全生命周期动态安全/健康边界、性能评估、长时安全预警等算法,揭示数字孪生的场景化应用理论。

3　数字孪生电池构建理论

通常情况下,数字孪生电池与物理实体之间是相互映射的,因此数字孪生电池主要包括物理实体、数字模型及其之间的数据和交互关系等。物理实体层在电池装机后既已确定,是数字孪生研究的主要对象,而数据和交互关系则依托于云计算、云服务等技术实现[21]。本文重点围绕数字孪生电池的模型构建方法,探究其构建准则和关键组件。

3.1　数字孪生模型构建准则

数字孪生模型通过多维度、多尺度、多物理场建模方法,全方位动态模拟电池在实际环境、工况条件下的性能演化过程。现有动力电池种类较多(磷酸铁锂电池、三元锂电池、钴酸锂电池、固态电池等)、尺寸差异大(方形电池、软包电池、圆柱形电池等)、应用场景多(商用车、乘用车、重型车辆、船用电池等),数字孪生模型应具备较高的普适性和面向用户的特异性,这对数字孪生模型的构建提出了较高要求,需要一系列建模准则对模型进行规范。如图 3 所示,基于用户使用场景、便捷性、通用性等考虑,数字孪生模型应满足以下 6 项构建基本准则。

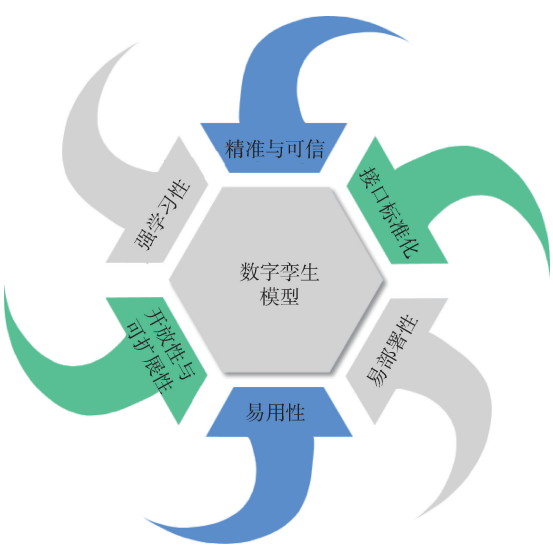


图 3　数字孪生模型构建 6 项基本准则

Fig. 3　Six basic guidelines for digital twin model construction

1)精准与可信。对物理实体的精准化描述是数字孪生模型应用的基础。精准化有多层次的含义,首先,数字孪生模型不仅应准确描述电池的静动态特性,还应准确描述出其随时间变化而产生的参数异变、演化等性能,在时间尺度上具有精准性;其次,数字孪生模型效仿电池的成组方式,采用材料-单体-模组-系统-整车层级的建模方法,在每一尺度上的建模方法、特征等可能存在差异,但均应满足高精度、高可靠的要求,从根本上减少模型误差的传递和积累,在空间尺度上也具有精准性。此外,数字孪生模型在云平台上运行时还应具有高可信的计算环境,除了要保障计算本身的准确性,还应对计算产生的数据、资源等进行保护。

2)接口标准化。由于电池种类的复杂性,数字孪生模型通常需要快速、准确的适配各类电池,因此需要对电池的接口进行标准化,从而提高适配速度和准确度,为不同领域、不同行业的使用人员提供便捷的使用方式。需要标准化的接口主要包括:模型定义、参数输入、特征曲线输入、数据接口、通讯协议、求解方式、输出内容、曲线和其他需要封装的内容等。此外,除接口标准化之外,对于接口的形式、格式、内容等也需要相应的定义,使其容易获取、解析、读取,并在不同电池间相互兼容,以满足通用性和差异性的要求。

3)易部署性。数字孪生模型需要部署在云计算平台中才可以运行,因此数字孪生模型应该对不同种类的云平台具备部署能力,这主要涵盖 2 方面的内容,强兼容性和轻量化。强兼容性是指数字孪生模型应采用标准、规范的建模软件或语言实现,同时这一软件或语言应能在大部分云计算平台中部署,例如常用的多物理场仿真软件 Comsol、科学计算软件 MATLAB 或基础语言 Python、C + + 等。轻量化是指数字孪生模型在满足主要功能的基础上,在几何描述、求解方法、计算复杂度、逻辑处理等方面归约或者简化,从而减少模型的复杂度,提高在云计算平台的运行速度、减少计算和存储开销,提高云计算平台的用户接入数量。

4)强学习性。在全生命周期内向物理实体无限逼近和学习是数字孪生模型的本质特征之一,其功能实现要求数字孪生模型具备可学习性,这是指模型应能够随物理实体或者系统自发的进行更新、演化,并在全生命周期内进行主动的持续

的性能优化。可学习性的实现要依赖于模型内部的自学习、自修正算法,通过获取的原始测试数据、实际运行数据、模型仿真数据等,通过智能优化算法或数据驱动算法等使模型的参数动态更新,以满足模型的准确性、智能性。

5)开放性与可扩展性。随着电池管理功能性不断拓展、管理复杂度不断提升,数字孪生模型的功能也应该随之更新,这要求数字孪生模型具有开放性和扩展性,在部署后仍具备扩展的能力。数字孪生模型的开放性可以表现为模型的增广或融合,如对动力电池机理的深入研究可能探寻到更多影响电池寿命、安全等性能的因素,在对新发现的因素进行建模描述后,应可以有机地与孪生模型进行整合,通过模型融合、合并、关联、协同等方式,使模型重新成为一个整体。可扩展性则要求模型应在不同的应用场景、工况条件下可以灵活的更改自身结构、参数配置等条件,从而快速满足用户需求和场景匹配等。

6)易用性。数字孪生模型的使用者可能来自不同领域、不同专业,因此其对动力电池领域的专业知识理解不深入、专业性不强,所以数字孪生模型的使用应具备易用性,包括可视化的前端界面和可交互的用户界面等。前端界面应具有直观可视的形式,包括电池所处的环境信息、历史运行工况、电池状态、历史故障信息等,由于各类信息的重复性高、交叉性强,因此前端界面应具备强大的数据分析、统计能力,并从多要素、多维度、多尺度、多物理场等形式展现电池信息。用户界面则是指数字孪生模型为用户提供的一个动态可配置的交互界面,在该界面中,用户应可以用图形化或流程化描述对模型的内置逻辑、组件等进行配置,并且可以修改、兼容不同的接口信息等,从而使用户在数字孪生模型内部庞大的模型系统和参数空间内了解应配置、修改的信息。

3.2 数字孪生模型关键组件

在分析了数字孪生模型构建的 6 项准则后,还应具体的分析数字孪生电池所应涵盖的基本内容,形成其构建理论。数字孪生模型仍属于模型驱动方法,因此经典的建模理论仍适用于数字孪生模型的构建;此外,由于数字孪生模型具有全生命周期性、自学习性和场景应用性,因此,还需要采用特殊的多维度、多尺度、多物理场耦合建模方法,从而实现对动力电池的复杂系统进行全方位的抽象和建模[22]。因此,对于数字孪生模型的构建,应包含以下 5 个步骤:基本模型构建,模型融合,模型矫正,模型验证,模型管理,各步骤关系如图 4 所示。基本模型将为数字孪生电池提供基础,通过各类建模方式,实现从材料、单体层级到模组、系统层级的模型构建和组装,随后采用模型融合方法,将不同层级的模型通过数据交互、参数传递、仿真对接等方式进行融合,保障系统是有机的整体,各子模型之间参数可以互联互通并且一致;模型矫正是在基本模型构建完成后,采用智能学习的算法对模型的参数、结果等进行优化和更新,这也将伴随着实体对象的工作而一并演化;模型的验证和管理是在数字孪生模型基础之前对其性能进行测试和优化,由于云计算平台中可能会运行多个模型,模型之间的交叉、冲突等需要合理的方法进行评价和管理。上述 5 个步骤是数字孪生模型构建的基本方法,除此之外还可额外的扩展出模型扩展、模型迭代等其他步骤。

1)基本模型构建。基本模型描述了动力电池的工作原理和使用特性,是数字孪生电池的基石。根据 6 项基本规则,对基本模型的构建应采用多层级、多尺度、多物理场耦合的方式进行建模。在层级上,应满足从材料到单体、模组、系统等多个层级的建模方法,对于材料层级,可选用分子动力学等建模方法描述原子、分子尺度的性能差异对电池最终性能的影响,而对于单体、模组、系统等层级而言,可以依据材料尺度的仿真结果结合实际的电池成组方式、电气连接方式、热管理系统等进行全系统的数字化建模;在尺度上,应采用微观-宏观相耦合的建模方法,通过分析微观层级的电池特性演化,分析宏观性能的衰退、劣化规律等;在多物理场耦合方面,基本模型应涵盖电化学、热、力、流体、应力等各类物理场,描述电池内部的电化学反应基本原理、老化原理、传质传热机



图 4 数字孪生模型五大步骤

Fig. 4 Five steps of digital twin model

理、机械应力、膨胀力、热管理冷媒流动机理等各类基本原理。在理想情况下,数字孪生电池应涵盖上述的多层级、多尺度、多物理场耦合,但是从实际建模和应用角度出发,可以适度的根据需求进行降维或简化。

2)模型融合。动力电池机理复杂、层级众多,采用基本原理构建的电池基本模型需要进一步的交叉和融合,使数字孪生电池从简单模型堆砌变为系统级的复杂模型。模型的融合需要涵盖以下方面。首先,模型的参数应该互联互通,如使用分子动力学仿真的电池材料参数应传递至单体层级,基于单体模型仿真的电、热等参数传递至模组或系统层级,而系统层级的工况、环境等也将反应单体电池工况的细微差别。对于参数的传递应包括尺寸、场参数、经验参数、物性参数等。其次,模型融合还应包括物理场的融合和交叉,这体现在不同物理场之间的参数传递和相互作用影响,例如电池的生热需依赖于电池内部的电化学反应和副反应过程,而电池的散热过程还依赖于热管理系统的热对流、热交换等过程。

3)模型矫正。模型矫正是数字孪生电池的最核心机制之一,通过模型矫正可以实现电池模型的自学习和自优化,从而实现模型全生命周期的更新和迭代。模型矫正部分应该包含 3 个方面。①应详细分析模型内的参数敏感度,探究微观/宏观参数与电池性能之间的相关性,如电极厚度、固/液相比例等对电极膨胀等影响较大,而对析锂层等影响较弱。通过敏感度分析可以对参数的重要程度进行划分,提取关键参数和所需要的参数精度。②应构建参数与性能之间的相互表征关系,目前车端上传的云端数据比较有限,一般包括电压、电流、温度等,因此需要建立参数与可获取的电池信息的方程或关系,从而为参数的解析和辨识提供基础。③应建立基于智能学习的参数解析和算法,通过集群式数据统计分析和并行加速计算等方法,从电池组的群体特征出发研究参数的提取和辨识,并反馈至个体中优化电池的参数,提高电池参数辨识的特异性和普适性。

4)模型验证。模型验证是保障数字孪生模型高效、可靠的有力证明。模型的验证主要包括 3 方面:软件验证、实体验证和线下测试验证。软件验证是指通过电池软件仿真形成闭环,通过多个模型之间的相互校正或基于历史数据的软件模型校正等,采用大量的仿真工况、场景等探究模型的精度和输出,分析数字孪生电池在各类场景、工况下的性能。实体验证通过模型仿真的结果与实际获取的信息进行对比,不仅可以验证数字孪生模型的准确度,更可以为模型的矫正提供依据,以模型误差为反馈矫正电池的参数、公式等。线下测试验证则是通过在电池装机后同步开展的一些电池工况、循环测试实验等,为数字模型提供精确、可靠的数据依据,可以有效提高模型的准确度,并为模型的自学习、参数自校正提供有效参考。

5)模型管理。模型管理是在模型的构建、融合和自学习基础上,在云平台上进一步对模型进行分类存储和智能化管理、调度。由于数字孪生模型功能复杂、架构庞大,因此对于每一项功能单独调度模型极易造成计算和存储资源的浪费,因此需要采用智能化的方法对模型进行管理、调度和维护。对于单模型运行的云计算平台而言,在模型调用时可以将计算信息、结果等分类存储,在其他服务和功能调用时首先查找已有数据结果,从而减少模型调用次数和计算开销,提高功能的调度速率。而对于多模型并行计算的云计算平台而言,则可以通过竞争机制对模型进行归类和优选,对于用户的调度首先提供历史最优的模型,而后在后台空闲时采用指定案例、服务需求等对多模型进行调用、计算,并根据计算分析结果确定当前最优解的模型,在下一次调度时采用该模型为用户提供服务。这将有助于模型在各类场景、工况下的应用并指导后续模型的构建。

基本模型构建,模型融合,模型矫正,模型验证,模型管理是数字孪生模型构建的关键步骤,但是在实际的应用过程中,受限于产品开发速度、需求、目标、场景等,可以依据企业自身的特点进行组合和选择,特别是基本模型构建方法众多,可针对所应用的场景进行适应的模型选择,优选模型复杂度、存储量、计算开销较为均衡的模型组件。

4　动力电池数字孪生实践

针对新能源汽车动力电池精细化设计、全生命周期精准管理的实际应用需求,本文基于数字孪生模型构建理论,在模型构建、性能仿真和预测等方面进行了探索和研究,如图 5 所示。

在数字孪生模型构建方面,以车用电池系统为研究对象,按材料-单体-模组-系统 4 个层级建立与物理实体相映射的电池模型,在材料层级包括正负极材料及缺陷构造、电解液浓度及离子浓度分布、隔膜孔隙率缩小或破坏、SEI 膜增厚或分解等子模型;在单体层级包括电池正负极离子扩

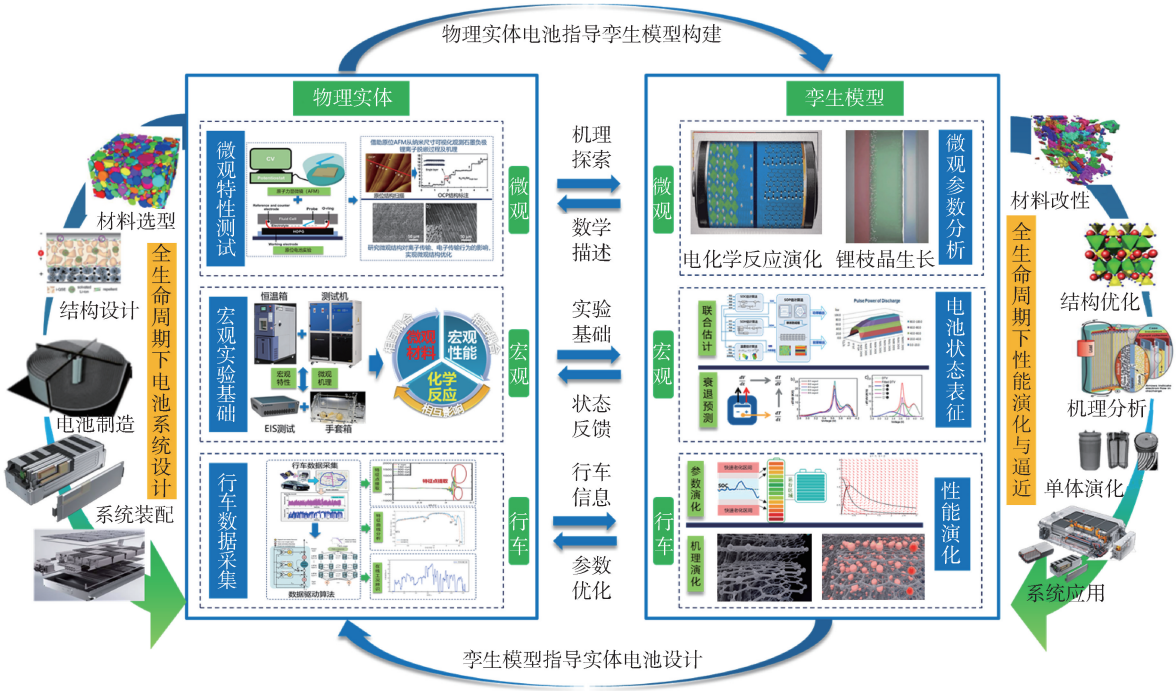


图 5 以电化学模型为基础的数字孪生电池

Fig. 5 Digital twin cell based on electrochemical model

散、电势产生、电化学-热产生等子模型;在模组层级,耦合电池产热传热机理,包括散热系统及其他电附属元件、均衡管理系统等;在系统层级,包括电池系统内的箱体结构设计、线束布置、焊接情况、铜排布置等。通过将材料-单体-模组-系统的不同类型、不同尺度、不同物理场的模型进行耦合和加权、参数传递、物理场融合,实现了从材料到系统层级的模型深度融合。

在电池性能和预测方面,数字孪生模型通过实体电池上传的电压、电流、温度、故障数、车速、位置等信息,在对实体电池组状态、参数等进行迭代优化后,实现电池膨胀/铜排老化等几何构型更新、电池内部微观参数的更新(包括电池正负极材料的膨胀、电解液分解、气体析出、隔膜损伤或收缩、孔隙率下降、正负极材料固相损失、SEI 膜增厚或分解、电极的扩散诱导应力、负极表面锂的析出和电解液内活性锂离子的损失、集流体腐蚀等)、宏观性能的衰退(包括内阻增长、可用容量损失、电池充电边界下降、可用容量区间缩短、极化电压加剧等)、全生命周期性能模拟(包括剩余寿命、最大可用容量、电池组安全性、可靠性衰退图谱等)。

此外,在基于数字孪生模型的管理方面,通过数字孪生模型在全生命周期的性能演化和模拟仿真,可以分析得到与电池使用工况呈强相关的电池内部参数的劣化特性,通过与神经网络算法相结合,采用模型 + 数据融合驱动的方法建立电池管理算法,从而实现电池系统的高安全动态充电边界管控、基于负极无析锂控制的充电策略设计、基于电池极端工况下失效边界预判的长时热失控预警等功能,为传统动力电池管理提供扩展性功能,有助于提高电池全生命周期全气候全工况的健康和安全管理能力。

5 数字孪生电池技术难点与挑战

数字孪生电池为新能源汽车能源系统智能化管控提供了新思路和方向,但目前阶段,构建完整的数字孪生电池还存在技术难点与挑战,主要包括电池精细化建模、全生命周期跨平台跨尺度建模等方面。

1)电池建模还需更精细化、更系统化。目前动力电池建模主要包括分子动力学模型、等效电路模型、电化学机理模型、数据驱动模型等。分子动力学模型依赖于对电池微观性机理的研究,从材料角度预测电池的扩散系数、离子电导率等参数,虽然具备不依赖于电池实体进行建模,但计算量大、模拟能力有限。等效电路模型是目前车端电池管理较为常用的模型之一,仅通过电阻、电容等元件模拟电池静、动态充放电特性,无法反应电池微观结构的变化和全生命周期的衰退。电化学机理模型是从电池化学反应原理出发等效的模型,主要包括单粒子模型、伪二维模型、异构模型等多种形式,通过模拟锂离子在正负极之间的迁

移及副反应的发生,可以预测全生命周期的性能衰退规律,但受限于模型的假设条件、有限的测量参数,电化学模型的实用性还不强。数据驱动模型是近年来电池建模的热点方向之一,通过不可解释的复杂神经网络模拟电池输入与输出间的映射关系,建模方式简单,但依赖于特征选取、训练集等。综上,目前动力电池建模理论还不完善,缺乏具有普适性的、能够切实反应真实条件下微观变化及宏观特效演化的模型,未来可基于电化学模型与数据驱动模型进一步研究探索,通过融合驱动建模方法等提升电池建模能力与理论水平。

2) 电池跨平台、跨领域建模难度大。现有动力电池建模方法依赖于有限元建模软件或科学计算软件平台,为实现从生产制造到车载动力到梯级利用到拆解回收的全链条建模与仿真,需要各平台间具备互融互通性并可实现跨平台、跨领域的建模仿真。目前,动力电池领域的常用建模软件如 COMSOL Multiphysics、ANSYS FLUENT、MATLAB 等,均已完成或已具备跨平台仿真的功能,但受限于模型接口复杂、各平台建模语言存在差异,跨平台的建模难度仍较大。此外,由于动力电池系统不仅包含电芯建模,也包括连接件、汇流排、热管理系统、均衡系统、热防护系统等各方面组件,组件建模并融合的难度也较大,为系统性的仿真设计带来了挑战。需要对电池系统内各组件的特性等进行独立建模仿真,形成模块化仿真组件和平台,才能够提升数字孪生电池的建模速度和质量,并提高组件可复用性、实用性等。

6　数字孪生电池应用与展望

信息技术的发展也为数字孪生带来了新思路和新需求。随着新能源汽车向电动化、智能化、网联化方向变革,基于数字孪生模型的动力电池管理与云计算、人工智能、虚拟现实等技术深度融合,对动力电池的设计、管理等产生了深远影响,诞生了基于数字孪生电池的新服务和新理念。

1) 基于模型的电池设计理念。传统动力电池的设计方法大多依靠经验选择,通过研发人员对材料的理解和实验经验,对电池正负极材料的配比、关键物性等进行选择和控制,导致了电池在设计完成前的性能属于未知状态,必须要依靠组装好的成品电池进行实验标定测试。基于模型的设计理念通过在电池组装前建立电池的等效模型,通过给出正负极材料、正负极厚度、离子浓度、活性颗粒尺寸、固相/液相比例等关键特征、物性参数等,通过电化学原理或数据驱动方法对电池组装后的性能进行预测,从而形成设计-仿真的闭环过程,减少设计阶段的测试和验证时间,提高电池开发效率、减少成本。数字孪生电池与基于模型的电池设计理念相吻合,数字孪生电池不仅能通过电池的材料参数预测组装后的静动态特性,更可以预测其全生命周期的性能,从而为更好的电池设计和优化提供了强有力的支撑。

2) 基于模型的模型测试、验证方法。目前动力电池的测试和验证几乎完全依赖于线下检测设备,如电池的容量、内阻、老化实验等要依赖于充放电测试仪、电化学阻抗谱测试等依赖于电化学工作站,设备依赖性强、测试费用高;针对动力电池管理的关键算法和模型的验证,也均为采用标准化的流程测试,如电池荷电状态、健康状态、功率状态等,测试周期长、测试范围窄、测试精度有限。基于数字孪生电池的测试和验证方法为研究电池性能、算法性能等提供了新的解决方案,如基于电化学原理建立的数字孪生电池,可以在仿真过程中实时获取电池的电化学阻抗信息,并且随着电池全生命周期的老化过程而观测到电池阻抗谱的演化过程;将数字孪生电池封装为黑盒模型后,可与电池管理的算法、模型等形成软件在环测试或全场景虚拟测试,通过数字孪生电池模拟在高原、高寒、高温等全气候全工况场景和过充、过放、过温等滥用场景,实现对动力电池管理算法和模型的大范围、高精度测试,同时由于仿真软件的计算能力强、仿真加速方法多,基于数字孪生模型的测试周期也大大缩短。

3) New IT 技术与数字孪生深度融合。云计算、车联网、大数据、人工智能等 New IT 技术快速发展,与数字孪生相结合已成为发展趋势。数字孪生电池需要可供部署和运行的大规模计算平台,并与车端的实体电池建立良好的沟通反馈渠道,云计算和车联网将打通数字孪生电池与实体电池间的屏障,使得数字孪生电池具备在全生命周期逼近实体并随实体共同演化的可能性。大数据和人工智能算法为数字孪生赋予了新的建模方法和自学习思路,对于传统的电化学模型或者等效电路模型而言,在模型建立后其性能便已经固定,与实体电池的相关性低,难以模拟不同用户、不同电池的差异性。大数据和人工智能算法可以为模型的参数提供自学习方法,通过上传的云端数据等分析电池在全寿命内的性能差异、参数匹配等,进而数字孪生在全生命周期的演化提供了 New IT 的解决方案。

4）智能应用/服务新需求。数字孪生电池为动力电池管理提供了的新思路，其全生命周期演化的优势将极大地丰富电池管理内容和策略，例如动力电池微观层级的性能演化分析（SEI 膜增长、充电析锂预测、电解液分解、电极膨胀、电池产气等），电池宏观层析的性能预测（充放电边界、快充策略、一致性均衡管理策略、热仿真分析和热薄弱环节管理）及长时安全预警、剩余寿命预测、综合性能评价、残值评估等功能，如图 6 所示。数字孪生电池可以为深入的分析电池性能提供依据和参考，基于数字孪生电池的智能化电池管理功能和服务也将推动行业技术进步、快速健康发展。

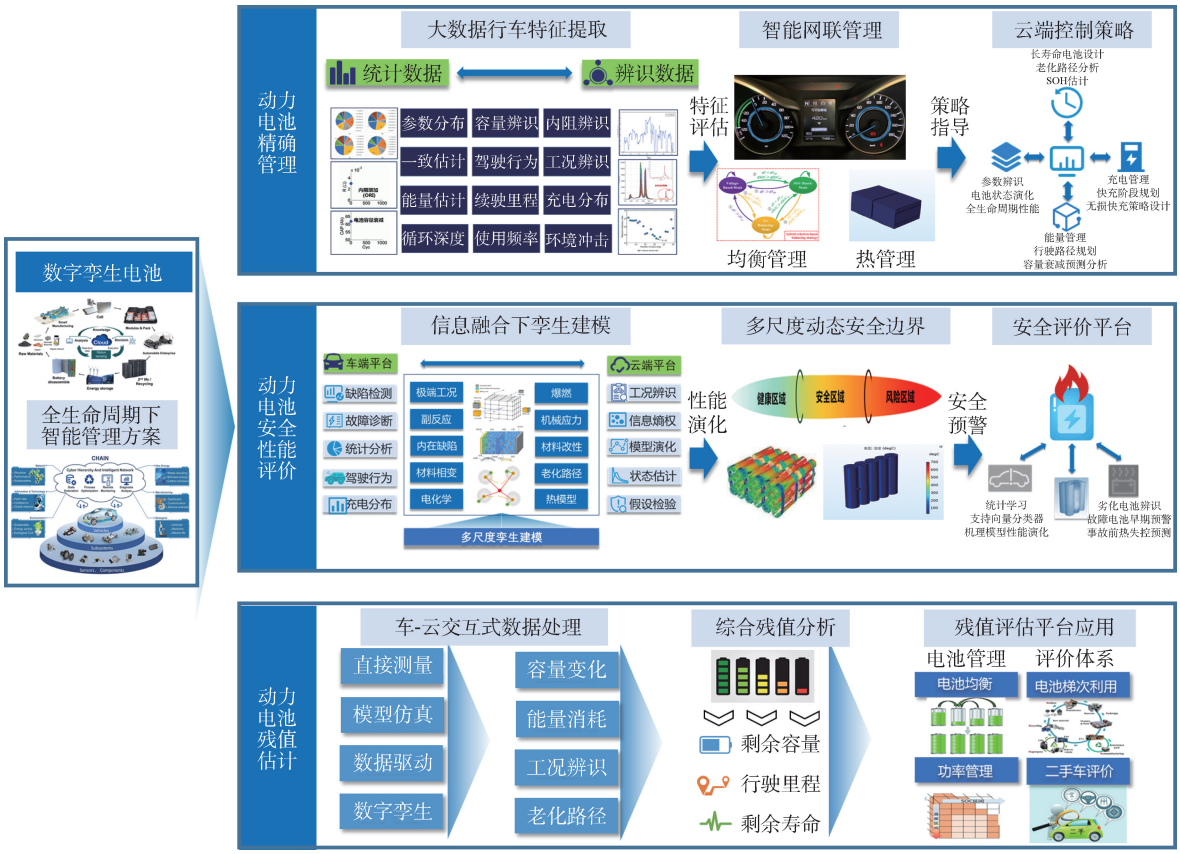


图 6　数字孪生与智能化电池管理应用

Fig. 6　Digital twin and intelligent battery management applications

7　结　论

随着新能源汽车技术的快速发展与普及，电池寿命、安全性已成为制约新能源汽车性能提升的关键。电池管理技术是保障电池健康、安全使用的核心，合理管控电池可实现延寿、性能充分释放、安全预警等功能。但目前车端电池管理系统在算力、存储等方面仍存在瓶颈，导致大量电池管控理论、方法、模型等不能应用到车端电池管理中。数字孪生技术的发展为电池管理提供了新契机，通过建立端-云融合式的智能化电池管理方法，将传统电池管控拓展至超算力、大存储的云平台中，通过数字孪生模型将各类电池建模方法、管控方法等有机融合，从而实现精准、健康的电池管理。

本文围绕目前数字孪生电池的发展需求和技术挑战，针对智能化管理需求下的数字孪生电池构建方法开展了研究：

1）从系统建模与管控需求等方面分析了数字孪生电池建模的基本准则，系统性阐述多维度、多尺度、多物理场融合的数字孪生电池的构建准则，基于用户使用场景、便捷性、通用性等考虑，数字孪生模型应具备精准可信、标准化、易部署、易用、可扩展、强学习等基本特性。

2）围绕数字孪生在全生命周期性、自学习性和场景应用性的需求，其建模方法可分解为基本模型构建、模型融合、模型矫正、模型验证、模型管理 5 个步骤，并针对所应用的场景进行在模型复杂度、存储量、计算开销等方面进行优选。

3）结合团队前期研究，分析了某电池数字孪生的实践案例，按材料-单体-模组-系统 4 个层级建立与物理实体相映射的电池模型，并分析了现有电池建模方法的不足与挑战。

随着数字能源、智慧能源需求的发展，数字孪

生电池在生产设计、全生命周期管理等场景均具备应用能力,通过指导电芯设计、测试、残值评估、智能管控等,拓展现有电池管理方法与平台,为电池管理技术发展提供思路与参考。

参考文献(References)

[1] 王亚楠,韩雪冰,卢兰光,等. 电动汽车动力电池研究展望:智能电池、智能管理与智慧能源[J]. 汽车工程,2022,44(4):617-637.
WANG Y N,HAN X B,LU L G,et al. Prospects of research on traction batteries for electric vehicles:Intelligent battery,wise management,and smart energy[J]. Automotive Engineering,2022,44(4):617-637(in Chinese).

[2] ZALOSH R,GANDHI P,BAROWY A. Lithium-ion energy storage battery explosion incidents[J]. Journal of Loss Prevention in the Process Industries,2021,72:104560.

[3] LI J,QIAO Z,BAO J,et al. An activity theory-based analysis approach for end-of-life management of electric vehicle batteries[J]. Resources,Conservation and Recycling,2020,162:105040.

[4] ZHANG J,JIANG Q,PAN A Q,et al. An optimal dispatching strategy for charging and discharging of electric vehicles based on cloud-edge collaboration[C]// 2021 3rd Asia Energy and Electrical Engineering Symposium (AEEES). Piscataway:IEEE Press,2021:827-832.

[5] YANG S C,HE R,ZHANG Z J,et al. CHAIN:Cyber hierarchy and interactional network enabling digital solution for battery full-lifespan management[J]. Matter,2020,3(1):27-41.

[6] 陶飞,张辰源,张贺,等. 未来装备探索:数字孪生装备[J]. 计算机集成制造系统,2022,28(1):1-16.
TAO F,ZHANG C Y,ZHANG H,et al. Future equipment exploration:Digital twin equipment[J]. Computer Integrated Manufacturing Systems,2022,28(1):1-16(in Chinese).

[7] 张辰源,陶飞. 数字孪生模型评价指标体系[J]. 计算机集成制造系统,2021,27(8):2171-2186.
ZHANG C Y,TAO F. Evaluation index system for digital twin model[J]. Computer Integrated Manufacturing Systems,2021,27(8):2171-2186(in Chinese).

[8] 朱凯,陈健,吕桃林,等. 空间电源数字孪生系统[J]. 上海航天(中英文),2021,38(3):197-206.
ZHU K,CHEN J,LYU T L,et al. Digital twin system for space power-sources[J]. Aerospace Shanghai (Chinese & English),2021,38(3):197-206(in Chinese).

[9] 刘蔚然,陶飞,程江峰,等. 数字孪生卫星:概念、关键技术及应用[J]. 计算机集成制造系统,2020,26(3):565-588.
LIU W R,TAO F,CHENG J F,et al. Digital twin satellite:Concept,key technologies and applications[J]. Computer Integrated Manufacturing Systems,2020,26(3):565-588(in Chinese).

[10] 陶飞,张贺,戚庆林,等. 数字孪生十问:分析与思考[J]. 计算机集成制造系统,2020,26(1):1-17.
TAO F,ZHANG H,QI Q L,et al. Ten questions towards digital twin:Analysis and thinking[J]. Computer Integrated Manufacturing Systems,2020,26(1):1-17(in Chinese).

[11] 陶飞,马昕,胡天亮,等. 数字孪生标准体系[J]. 计算机集成制造系统,2019,25(10):2405-2418.
TAO F,MA X,HU T L,et al. Research on digital twin standard system[J]. Computer Integrated Manufacturing Systems,2019,25(10):2405-2418(in Chinese).

[12] 陶飞,刘蔚然,张萌,等. 数字孪生五维模型及十大领域应用[J]. 计算机集成制造系统,2019,25(1):1-18.
TAO F,LIU W R,ZHANG M,et al. Five-dimension digital twin model and its ten applications[J]. Computer Integrated Manufacturing Systems,2019,25(1):1-18(in Chinese).

[13] KHAYYAM H,ABAWAJY J,JAVADI B,et al. Intelligent battery energy management and control for vehicle-to-grid via cloud computing network[J]. Applied Energy,2013,111:971-981.

[14] LI W H,CUI H,NEMETH T,et al. Cloud-based health-conscious energy management of hybrid battery systems in electric vehicles with deep reinforcement learning[J]. Applied Energy,2021,293:116977.

[15] SASAKI T,UKYO Y,NOVÁK P. Memory effect in a lithium-ion battery[J]. Nature Materials,2013,12(6):569-575.

[16] LIPU M S H,HANNAN M A,KARIM T F,et al. Intelligent algorithms and control strategies for battery management system in electric vehicles:Progress,challenges and future outlook[J]. Journal of Cleaner Production,2021,292:126044.

[17] ZHANG Y,LIU H P,ZHANG Z G,et al. Cloud computing-based real-time global optimization of battery aging and energy consumption for plug-in hybrid electric vehicles[J]. Journal of Power Sources,2020,479:229069.

[18] DENG J,BAE C,MARCICKI J,et al. Safety modelling and testing of lithium-ion batteries in electrified vehicles[J]. Nature Energy,2018,3(4):261-266.

[19] HARPER G,SOMMERVILLE R,KENDRICK E,et al. Recycling lithium-ion batteries from electric vehicles[J]. Nature,2019,575(7781):75-86.

[20] SEVERSON K A,ATTIA P M,JIN N,et al. Data-driven prediction of battery cycle life before capacity degradation[J]. Nature Energy,2019,4(5):383-391.

[21] 陶飞,张贺,戚庆林,等. 数字孪生模型构建理论及应用[J]. 计算机集成制造系统,2021,27(1):1-15.
TAO F,ZHANG H,QI Q L,et al. Theory of digital twin modeling and its application[J]. Computer Integrated Manufacturing Systems,2021,27(1):1-15(in Chinese).

[22] 卢锦玲,颜禄涵,腊志源,等. 基于数字孪生与动态能效模型的综合能源系统实时优化调度策略[J/OL]. 电网技术,2022(2022-05-06)[2022-07-06]. https://doi.org/10.13335/j.1000-3673.pst.2022.0044.
LU J L,YAN L H,LA Z Y,et al. Real-time optimal scheduling strategy for integrated energy system based on digital twins and dynamic energy efficiency model[J]. Power System Technology,2022(2022-05-06)[2022-07-06]. https://doi.org/10.13335/j.1000-3673.pst.2022.0044(in Chinese).

Construction of digital twin model of lithium-ion battery for intelligent management

YANG Shichun, LI Qiangwei, ZHOU Sida, ZHANG Zhengjie, MA Yuan, CHEN Fei*

(School of Transportation Science and Engineering, Beihang University, Beijing 100083, China)

Abstract: To achieve peak carbon and carbon neutral goals, the development of electric cars has become strategically important. It is necessary to have precise battery management technology because the lifespan and safety of power batteries change dynamically as they are used. This leads to rapid capacity degradation brought on by the inconsistent performance of single cells and thermal runaway brought on by short board batteries or internal defects. New battery management capabilities have been made possible by the development of the digital twin model, which is now one of the technical trends in the industry. Based on the development trend of battery management technology, the article concentrates on the analysis of basic principles of battery digital twin modeling from the aspects of system modeling to management and control requirements. The article systematically introduces the construction method of multi-dimensional, multi-scale and multi-physical field fusion of digital twin battery. Combined with the previous research of the team, the practical case of the digital twin battery was analyzed. Finally, the application perspective of digital twin battery is discussed in production design, life cycle management and other scenarios, which provided ideas and references for the development of battery management technology.

Keywords: new energy vehicles; power battery; digital twinning; battery model; battery management system

Received: 2022-07-07; **Accepted**: 2022-08-05; **Published online**: 2022-08-12 13:56

URL: kns.cnki.net/kcms/detail/11.2625.V.20220812.1323.002.html

Foundation item: National Natural Science Foundation of China (U1864213)

* **Corresponding author**. E-mail: cf2020@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0323

新型扭簧电连接器的确信可靠性正向设计优化

康锐[1,2], 刘浩然[1,2], 张清源[2,3,*], 余丽[1,2], 翟国富[4]

(1. 北京航空航天大学 可靠性与系统工程学院, 北京 100083; 2. 可靠性与环境工程技术重点实验室, 北京 100083;
3. 北京航空航天大学 航空科学与工程学院, 北京 100083; 4. 哈尔滨工业大学 电气工程及自动化学院, 哈尔滨 150001)

摘 要: 新产品正向设计需综合考虑新结构工艺、多设计变量、复杂工作环境的影响,并合理量化不确定性影响,评估可靠性水平,从而在源头上确保高可靠性。基于确信可靠性理论,研究了基于性能裕量的确信可靠性正向设计优化方法。提出了面向新型扭簧电连接器研发的确信可靠性正向设计优化基本流程,包括设计变量初值优选、性能裕量建模、不确定性量化和确信可靠性分析与优化4部分。以某新型扭簧电连接器的正向设计为例,基于正交实验、响应面建模和模拟退火启发式优化等方法实现了多维离散-连续变量及其不确定性作用下的优化设计。

关 键 词: 确信可靠性; 性能裕量; 正向设计优化; 实验设计; 多维离散-连续变量
中图分类号: TB114.3
文献标志码: A **文章编号**: 1001-5965(2022)09-1745-12

电连接器是用于传输信号、接通电流的关键器件,其性能降级或失效对整个系统都有重要影响。近年来,扭簧电连接器因载流改善能力强、接触力稳定性高,在追求小型化、大电流可靠传输的市场中广受关注[1]。由于存在较高的技术壁垒,要在国产材料、工艺条件下实现扭簧结构,同时满足对竞品低成本的竞争优势,必须要开展多次迭代的概念、结构和详细方案的正向设计,自主研发经济可靠的新型扭簧电连接器。

当前,电连接器的设计主要是在少数理论分析的基础上,以优化性能为目标。一种常见的设计思路是先参考传统型号的经验结构给出设计图稿[2],再利用经验模型、手册指南等进行理论分析和性能建模[3],之后基于经验取值和区间范围进行实验设计(DOE)和仿真[4-5],最后通过试验测试分析变量的不确定性影响,从而改进设计和控制工艺[6]。例如,不同学者基于DOE方法针对材料组合、结构尺寸和工艺公差中少数设计变量进行参数化建模与仿真[7-8]。文献[9]基于接触电阻模型磨损方程等先后开展了某型片簧电连接器材料、结构的设计优化。这些设计方法虽能在一定程度上实现电连接器的功能性能设计,但通常无法在设计中保证批量产品在复杂工作环境下的可靠性,且设计过程存在许多经验性的、主观的判断,仍然可能带来失效风险。

为了在新型扭簧电连接器正向设计过程中充分考虑可靠性因素,必须对其进行可靠性正向设计优化。然而,由于新型扭簧电连接器工艺过程的创新性与结构原理的复杂性,其可靠性正向设计主要面临2方面挑战:

1) 新型扭簧电连接器是复杂弹性体结构,设计变量众多,交互作用复杂,对可靠性影响形式多样,方案优选的难度大。正向设计需要同时考虑

收稿日期: 2022-05-06; **录用日期**: 2022-05-27; **网络出版时间**: 2022-06-14 09:16
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220613.1329.001.html
基金项目: 国家自然科学基金(62073009); 可靠性与环境工程技术重点实验室基金(6142004210401)
* **通信作者**. E-mail: zhangqingyuan@buaa.edu.cn

引用格式: 康锐, 刘浩然, 张清源, 等. 新型扭簧电连接器的确信可靠性正向设计优化[J]. 北京航空航天大学学报, 2022, 48(9): 1745-1756. KANG R, LIU H R, ZHANG Q Y, et al. Function oriented belief reliability design and optimization of new torsion spring electrical connectors[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1745-1756(in Chinese).

从原料采买到加工装配整个过程,往往存在多维离散-连续变量[10],而现有方法往往是针对某一种或某几个设计变量的设计,并非系统化和全局化的设计,效率低、成本高及优化空间小。

2)新型扭簧电连接器采用新工艺过程和新加工设备实现小尺寸加工,且面临复杂工作环境,可靠性受选材、工艺、手册标准、外界环境及工作条件等多源不确定性影响[11],而现有研究通常只考虑其中部分因素的不确定性,模型不完整。这在缺乏认知、工艺不成熟及需求复杂多变的正向设计中无法基于设计方案得到产品在不同工作场景下的可靠性指标,易导致设计偏差,造成潜在失效风险。

因此,针对多设计变量和多源不确定性的特征,本文以确信可靠性理论为基础,提出直接面向市场需求的确信可靠性正向设计优化方法,以解决系统性设计优化与不确定性量化分析的问题。确信可靠性理论是一种新的可靠性理论[12],其基于裕量可靠、退化永恒和不确定性原理,从性能裕量的角度构建起产品性能与可靠性之间的相关关系[13],并通过综合度量工程中的固有不确定和认知不确定因素[14],实现产品定量建模与定量设计[15],更适合面向用户需求的产品可靠性正向设计优化过程。

综上,本文基于确信可靠性理论,以性能裕量为核心,研究新型扭簧电连接器的确信可靠性正向设计优化方法,以解决新型电连接器设计中的多维设计变量和多源不确定性问题。

1 基于确信可靠性的正向设计优化流程

1.1 确信可靠性理论

确信可靠性理论以裕量可靠、退化永恒及不确定性3方面可靠性科学原理为指导,构建了基于学科方程、裕量方程、退化方程和度量方程的理论体系[12],具体表示为[12,16]

$$\boldsymbol{P} = f_0(\boldsymbol{X},\boldsymbol{Y},t) \tag{1}$$

$$\boldsymbol{M} = d(\boldsymbol{P},\boldsymbol{P}_{\text{th}}) \tag{2}$$

$$\boldsymbol{P} = f_t(\boldsymbol{X},\boldsymbol{Y},t) \tag{3}$$

$$R = \mu(\tilde{\boldsymbol{M}} > 0) \tag{4}$$

在这些方程中,核心是产品性能 $\boldsymbol{P}$、裕量 $\boldsymbol{M}$、可靠度 R,相关影响因素包括内因变量 $\boldsymbol{X}$、外因变量 $\boldsymbol{Y}$、可逆时间 t、不可逆时间 $\vec{t}$[17] 及性能阈值 $\boldsymbol{P}_{\text{th}}$,它们之间的关系由距离函数 $d(\cdot)$、退化函数 $f_t(\cdot)$ 和测度函数 $\mu(\cdot)$ 量化。另外,学科方程中 f_0 表示未发生退化时产品性能与内外因变量及可逆时间的函数,$\tilde{\boldsymbol{M}}$ 表示受不确定性影响的性能裕量。

在确信可靠性理论框架下,产品性能裕量 $\boldsymbol{M}$ 决定了产品的可靠程度,可靠度就是对性能裕量大于0的度量。要在产品正向设计过程中达到高可靠度,核心即围绕性能裕量建立4个方程。

1.2 基本流程

在实际正向设计中,最首要的任务是保证初始状态下产品裕量可靠,否则考虑产品的退化也没有意义。因此,本文确信可靠性正向设计优化的根本目标是在功能需求 $\boldsymbol{P}_{\text{th}}$ 要求下,以关键性能 $\boldsymbol{P}$ 的裕量 $\boldsymbol{M}$ 为核心,考虑外因变量 $\boldsymbol{Y}$ 的影响,通过优化内因变量 $\boldsymbol{X}$ 的设计值使产品更可靠。为此,本文基于初始状态提出了确信可靠性正向设计优化的基本流程,如图1所示。

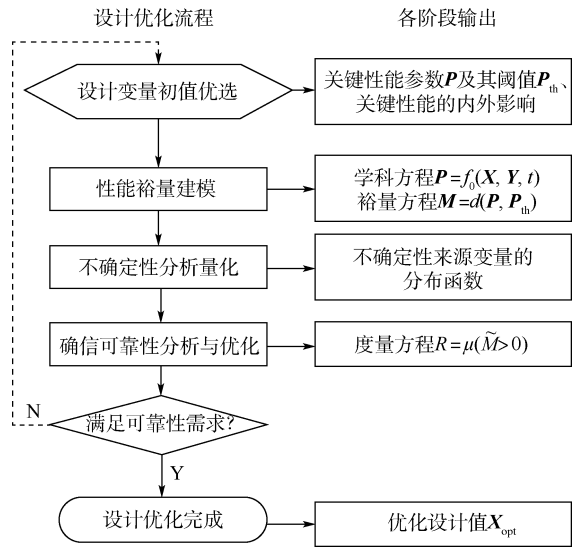


图1 确信可靠性正向设计优化流程

Fig. 1 Process of function oriented belief reliability design and optimization

1)设计变量初值优选。分析产品功能原理和确定关键性能,识别系统内外影响因素,并通过设计正交实验初步确定使得性能最优的设计变量组合作为设计初值。

2)性能裕量建模,即构建关键性能的学科方程和裕量方程。主要有2类方法:①基于产品功能性能关系进行理论推导,构建解析模型;②通过实验构建代理模型,通常可使用响应面方法。

3)不确定性分析量化。一方面,明确内外因素可能受到的不确定性影响,分析不确定性的来源;另一方面,确定不确定性因素的分布类别与分布函数,实现不确定性量化。

4)确信可靠性分析与优化。建立初始确信

可靠性度量方程,并基于产品可靠性要求建立优化模型,求得使可靠度最大的设计值。

由于对产品认知和实验条件的限制,上述流程通常需要随着研发过程推进和实验能力提升不断迭代进行。另外,除本文关注的初始确信可靠度要求外,还要在进一步的设计中关注产品的性能退化过程,从而实现产品的迭代优化与可靠性增长。

2　确信可靠性正向设计优化方法

本节将基于图 1 中给出的确信可靠性正向设计优化流程,详细阐述针对新型扭簧电连接器的确信可靠性正向设计优化方法。

2.1　设计变量初值优选方法

在新产品正向设计之初往往要考虑大量影响因素,众多的设计变量与水平数让组合方案数量庞大。为提高设计优化效率和节约设计成本,从众多的设计方案中选出性能最优的设计变量组合,本文将设计变量初值优选的过程分为 3 个递进的关键步骤,即确定关键性能和阈值、内外影响因素分析和基于正交实验的设计变量优选。

2.1.1　确定关键性能和阈值

关键性能 $\boldsymbol{P}$ 是指产品使用者和设计者最关心的功能评价指标或指标体系,可基于 FPMA 方法展开分析[13]。新型扭簧电连接器的功能是在大功率传输和信息传递中保证电连接器的稳定导通和安全使用,其性能包括温升、接触电阻、插拔力、插拔寿命、电压降等。在正向设计中,通常将温升 ΔT 确定为新型扭簧电连接器的关键性能,主要有 2 方面原因:①新型扭簧电连接器往往需要在几十安培充电电流下可靠工作,通电后温升过高可能导致焊点或结构熔化,造成极大安全隐患[18];②温升是最直观、最易测量的性能指标,且可在一定程度上从侧面反映接触电阻、插拔力等其他性能的水平,具有更好的综合性。

阈值 $\boldsymbol{P}_{\mathrm{th}}$ 是性能可满足要求的边界值,根据不同的要求分为规定阈值 $\boldsymbol{P}_{\mathrm{th,c}}$、破坏阈值 $\boldsymbol{P}_{\mathrm{th,d}}$ 等。规定阈值是由行业标准手册和用户需求规定的性能边界;破坏阈值是产品发生不可逆破坏或产生安全隐患的失效边界,由材料属性决定[15]。对于扭簧电连接器而言,规定阈值通常不能覆盖全部的工作场景。例如,电动车行业标准 SAE/USCar2 中规定,连接端子通电后任意时刻的最大温升不得超过 55 K[19],即 $\Delta T_{\mathrm{th,c}}=55$ K,但却未考虑使用环境温度。因为焊锡在熔点(183℃)90% 的温度下(即 $T_{\max}=164.7$℃)就将发生松脱,所以不同环境温度 T_{e} 下均对应着不同的温升破坏阈值 $\Delta T_{\mathrm{th,d}}$。由于温升越小越好,取各类阈值中的最小值作为温升阈值:

$$\Delta T_{\mathrm{th}}=\min(\Delta T_{\mathrm{th,c}},\Delta T_{\mathrm{th,d}}) \tag{5}$$

2.1.2　内外影响因素分析

内外影响因素是指影响关键性能的产品内部因素(结构、材料、工艺等)和外部因素(工作应力、使用环境等)。为了基于确信可靠性进行新型扭簧电连接器正向设计,必须明确内外影响因素,为裕量建模奠定基础。本文通过分析温升产生的机理来确定内外影响因素。

温升的机理可从产热和散热两大方面考虑。产热方面,由焦耳定律 $Q=I^2R_{\mathrm{eq}}t$ 可知,工作电流大小和接触电阻大小都对温升有决定性影响。工作电流作为环境载荷通常由实际使用情况而定,而体电阻、收缩电阻及膜层电阻 3 部分组成的接触电阻由产品设计决定。其中,体电阻与端子材料、结构有关,收缩电阻、膜层电阻则与插拔过程的过盈量与端子磨损程度相关。散热方面,材料与环境的热量交互会影响温升。当扭簧电连接器温度高于环境温度时,就会通过热传导、热辐射和热对流的方式向环境散热,因此由材料热特性决定[20]。

本文使用向量 $\boldsymbol{X}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 和 $\boldsymbol{Y}=(y_1,y_2,\cdots,y_r)^{\mathrm{T}}$ 分别表示内外因变量,其中 n 和 r 分别表示内外因变量个数。具体而言,内因向量 $\boldsymbol{X}$ 需要进行设计以使产品获得最优性能,因此也称为设计变量,其既包括离散变化的离散型变量 $\boldsymbol{X}_{\mathrm{D}}$,也包括在一定区间范围内连续变化的连续型变量 $\boldsymbol{X}_{\mathrm{C}}$。例如,板材厚度只能从供应商提供的某几个类型中选择,是离散型变量;扭簧的扭转角度可以在(0°,90°)间选取,是连续型变量。新型扭簧电连接器端子的外因向量 $\boldsymbol{Y}$ 主要涉及电流 I_i 和环境温度 T_{e}。其中,I_i,$i\in V$ 表示不同工作场景 i 下的电流,V 表示工作场景集合,T_{e} 表示扭簧电连接器外环境温度。

2.1.3　基于正交实验的设计变量优选

根据上述分析,新型扭簧电连接器端子的设计变量初值优选需要实现 2 方面目标:针对离散型设计量 $\boldsymbol{X}_{\mathrm{D}}$,从有限的取值可能中确定使得温升最低的取值组合 $\boldsymbol{X}_{\mathrm{D,0}}$;针对连续型设计变量 $\boldsymbol{X}_{\mathrm{C}}$,探索其对温升性能的影响,初步选出最优水平组合 $\boldsymbol{X}_{\mathrm{C,0}}$ 作为设计优化的初值。

为了实现上述目标,本文使用正交实验设计方法[21]实现新型扭簧电连接器端子的多维离散-

连续设计变量初值优选,并初步确定各设计变量及其相互作用对温升的影响和程度[22]。

首先,要确定所有内因变量的水平。若第 k 个变量有 J_k 个水平,则记 x_{kj} 为其第 j 个水平对应的变量取值,其中 $k=1,2,\cdots,n$,$j=1,2,\cdots,J_k$。随后,根据变量个数及水平数选择合适的正交表进行实验。最后,通过主效应和交互作用分析每个内因变量及其各个水平 x_{kj} 对温升的影响并选出在现有水平下内因变量初值的最优组合 $\boldsymbol{X}_0=(\boldsymbol{X}_{\mathrm{D},0},\boldsymbol{X}_{\mathrm{C},0})^{\mathrm{T}}=(x_{1,0},x_{2,0},\cdots,x_{n,0})^{\mathrm{T}}$。

2.2 性能裕量建模方法

新型扭簧电连接器温升的学科方程难以求得解析表达式,在构建其性能裕量模型时往往需采用响应面方法求解多变量-响应问题,建立温升性能与多种变量间的关系,即学科方程的代理模型。

2.2.1 学科方程的构建

响应面法的核心原理是将变量与响应间的关系视为响应曲面[23],通过选取采样点设计有限的实验得到响应面上采样点的响应,加权回归后用多项式代理模型拟合响应目标和变量之间的变化关系[24]。因此,对于新型扭簧电连接器端子而言,其学科方程的构建主要包括如下步骤:

步骤1 确定所有变量取值范围,由于离散型内因变量选型已确定,设计响应面所考虑的变量 $\boldsymbol{\Theta}$ 应包括连续型内因变量 $\boldsymbol{X}_{\mathrm{C}}$ 和外因变量 $\boldsymbol{Y}$,即

$$\boldsymbol{\Theta}=(\Theta_1,\Theta_2,\cdots,\Theta_{m+r})^{\mathrm{T}}=(\boldsymbol{X}_{\mathrm{C}}^{\mathrm{T}},\boldsymbol{Y}^{\mathrm{T}})^{\mathrm{T}}=(x_1,x_2,\cdots,x_m,y_1,y_2,\cdots,y_r)^{\mathrm{T}} \tag{6}$$

在使用条件分析和正交实验水平优选的基础上,确定相关变量的取值范围:

$$\{\Theta_c\mid\Theta_c^{\mathrm{L}}\leqslant\Theta_c\leqslant\Theta_c^{\mathrm{U}},\ c=1,2,\cdots,m+r\}\in S \tag{7}$$

式中:Θ_c^{U}、Θ_c^{L} 分别为变量 Θ_c 的上、下边界;S 为响应曲面。为尽可能消除量纲影响,取每个变量的区间边界和中心,基于 Min-Max 归一化方法建立单位空间 $\Omega=\{\zeta^*\mid\zeta_i^*\in[-1,1]\}$,即

$$\begin{cases}\Theta_c^*=\dfrac{2\Theta_c-(\Theta_c^{\mathrm{U}}+\Theta_c^{\mathrm{L}})}{\Theta_c^{\mathrm{U}}-\Theta_c^{\mathrm{L}}}\\ S^*=\{\Theta_c^*\mid-1\leqslant\Theta_c^*\leqslant1,c=1,2,\cdots,m+r\}\end{cases} \tag{8}$$

需要指出,由于变量对性能有确定性的响应关系,直接将变量归一化会削弱响应效果,应根据需要对变量进行处理。例如,电流 I_i 的平方项 I_i^2 对温升有决定性影响,因此用电流平方归一化数值 I_i^{2*} 拟合响应面,从而突出平方响应效果,即

$$I_i^{2*}=\frac{2I_i^2-((I_i^{\mathrm{U}})^2+(I_i^{\mathrm{L}})^2)}{(I_i^{\mathrm{U}})^2-(I_i^{\mathrm{L}})^2} \tag{9}$$

步骤2 建立二阶代理模型。由于二阶模型能综合考虑变量的独立作用和交互作用,本文用常见的二阶模型建模,即

$$\Delta T=f(\boldsymbol{\Theta}^*)=a_0+\boldsymbol{A}\boldsymbol{\Theta}^*+\boldsymbol{\Theta}^{*\mathrm{T}}\boldsymbol{B}\boldsymbol{\Theta}^*+\varepsilon \tag{10}$$

式中:a_0 为常数项;矩阵 $\boldsymbol{A}=\mathbf{R}^{1\times(m+r)}$ 为一次项系数;矩阵 $\boldsymbol{B}=\mathbf{R}^{(m+r)\times(m+r)}$ 为二次项系数;ε 为随机扰动项。随着项数增加,拟合的方程通常更精确,但计算过程更加复杂,且并非所有项都会对结果产生决定性影响,容易因项数过多产生过拟合问题,因此,本文用方差齐性检验(F-test)检验式(11)中各项的显著性,过滤出显著项[22-23],并得到简化的学科方程:

$$\Delta T=\varepsilon+a_0+\sum_{i=1}^{m+r}w_i\alpha_i\Theta_i^*+\sum_{i=1}^{m+r-1}\sum_{j=i+1}^{m+r}[\omega_{ii}\beta_{ii}(\Theta_i^*)^2+\omega_{ij}\beta_{ij}\Theta_i^*\Theta_j^*] \tag{11}$$

式中:α_i、β_i 和 β_{ii} 为各项的拟合系数,w_i 和 ω_{ij} 为各项的过滤系数,本文依据变量的显著性水平将其定义为0/1变量,即

$$\begin{cases}w_i=\begin{cases}0 & \Theta_i\text{ 不显著}\\ 1 & \Theta_i\text{ 显著}\end{cases}\\ \omega_{ij}=\begin{cases}0 & \Theta_i\Theta_j\text{ 不显著}\\ 1 & \Theta_i\Theta_j\text{ 显著}\end{cases}\\ \forall\, i,j=1,2,\cdots,m+r\end{cases} \tag{12}$$

2.2.2 裕量方程的建立

根据简化方程中的非零项可得到影响关键性能的核心变量 $\boldsymbol{\Theta}^\star$:

$$\boldsymbol{\Theta}^\star=(\boldsymbol{X}_{\mathrm{C}}^\star,\boldsymbol{Y}^\star)=(x_1,\cdots,x_{m_1},y_1,\cdots,y_{r_1})\quad m_1\leqslant m,r_1\leqslant r \tag{13}$$

由于温升是望小性能,进一步可以用欧氏距离建立关键性能的裕量方程:

$$\boldsymbol{M}=\Delta T_{\mathrm{th}}-\Delta T=\Delta T_{\mathrm{th}}-f_0(\boldsymbol{X}^\star,\boldsymbol{Y}^\star) \tag{14}$$

2.3 不确定性分析与量化方法

为了分析新型扭簧电连接器端子在不同内外因变量作用下的确信可靠性水平,还需针对其裕量方程开展不确定性分析与量化。

裕量方程的不确定性主要来源于3个方面:模型形式及系数的不确定性、内外因变量的不确定性以及性能阈值的不确定性。

模型的不确定性是代理模型构建过程中引入的,这使得模型预测值与实际值存在一定偏差,因

此本文在学科方程中加入随机扰动项 ε,其分布为 $\boldsymbol{\Phi}_{\varepsilon}$。

内外因变量的不确定性源于生产制造的尺寸偏差和外部环境的变化与波动。对裕量方程中的内因变量 $\boldsymbol{X}^{\star}$,要综合考虑制造公差、设备精度和测量精度的影响,往往可以通过收集生产制造的相关数据给出其分布,表示为 $\boldsymbol{\Phi}_{x_1},\boldsymbol{\Phi}_{x_2},\cdots,\boldsymbol{\Phi}_{x_{m_1}}$。对于外因变量 $\boldsymbol{Y}^{\star}$,通常可假定电流和环境温度服从正态分布 $\boldsymbol{\Phi}_I$、$\boldsymbol{\Phi}_{T_e}$,即 $I_i \sim N(\mu_{I_i},\sigma_{I_i})$,$T_e \sim N(\mu_{T_e},\sigma_{T_e})$,其中,$i\in V$ 表示工作场景,电流均值是对应场景下的额定电流值,环境温度均值为规定的工作环境温度。根据 3σ 原则,分别取其波动范围的 1/3 为标准差,即 $\sigma_I=\varepsilon_I/3$,$\sigma_{T_e}=\varepsilon_{T_e}/3$。

性能阈值的不确定性由工作场景的变化和波动导致。因为温升的阈值本质上是工作环境温度的函数,可以据此计算阈值的分布 $\boldsymbol{\Phi}_{\Delta T_{\text{th}}}$。

2.4　确信可靠性分析与优化方法

基于上述确定性模型与不确定性量化结果,新型扭簧电连接器端子的确信可靠度即为性能裕量大于 0 的概率,即

$$\begin{aligned}R_i &= \Pr\{\tilde{\boldsymbol{M}} > 0\} = \\ &\Pr\{d(g(\boldsymbol{X}^{\star},I_i,T_e),\Delta T_{\text{th}}) > 0\} = \\ &1-\boldsymbol{\Phi}(\nu_M)\end{aligned} \tag{15}$$

式中:R_i 为工作场景 i 下扭簧电连接器端子的确信可靠度;$\nu_M=-\mu_M/\sigma_M$,μ_M 和 σ_M 分别为裕量的均值和标准差。当裕量分布形式简单时,可直接用式(15)求解,否则可用蒙特卡罗仿真。

为得到最优设计值,本文以最大化任意规定工作环境下确信可靠度为目标,以核心内因变量 $\boldsymbol{X}^{\star}$ 的均值 $\boldsymbol{\mu}_{X^{\star}}=(\mu_{x_1},\mu_{x_2},\cdots,\mu_{x_{m_1}})$ 为决策变量,以成本和设计变量取值范围为约束建立优化模型(16)。其中成本由加工精度、工艺、仪器设备等多项因素共同决定,可认为是设计变量目标值与标准差的函数。

$$\begin{aligned}&\text{Maximize}\ \bigwedge_{i\in V} R_i \\ &\text{s.t.}\quad \text{Cost}(\boldsymbol{\mu}_{X^{\star}},\boldsymbol{\sigma}_{X^{\star}}) \leqslant \text{Cost}(\boldsymbol{\mu}_{X_0},\boldsymbol{\sigma}_{X_0}) \\ &\mu_{x_k}^{\text{L}} \leqslant \mu_{x_k} \leqslant \mu_{x_k}^{\text{U}} \\ &k=1,2,\cdots,m_1,x_k\in\boldsymbol{X}^{\star}\end{aligned} \tag{16}$$

3　案例研究

本文以某新型 3.6 mm 扭簧电连接器的正向设计过程为例,说明本文方法的有效性。

扭簧电连接器的研发目的是设计高可靠的电连接器通用标准件,以满足电动车电连接器的使用需求。在现阶段受到生产加工、实验能力等限制,主要优化目标是在新结构、小尺寸、多设计变量和复杂工作环境下,提高初始确信可靠度。

3.1　设计变量初值优选

首先,确定关键性能和阈值。由于选择了温升作为关键性能,通过调研用户需求和电动车行业标准 IEC 60320-1[24]、SAE/USCAR-2[19]、GB/T 20234.1—2015[25] 等,可以获取扭簧电连接器的工作场景集 V 和不同工作场景 $i\in V$ 下的规定阈值 $\Delta T_{\text{th,c}}$,如下:

$$\begin{aligned}V = \{I,T_e \mid &\mu_I=\{40,60,70\},\sigma_I=0.05\mu_I/3, \\ &\mu_{T_e}=[-40,125],\sigma_{T_e}=5/3\}\end{aligned} \tag{17}$$

$$\Delta T_{\text{th,c}}^{i}=\begin{cases}30\ \text{K} & i=\{\mu_I=40,\ -40\leqslant\mu_{T_e}\leqslant 125\}\\ 50\ \text{K} & i=\{\mu_I=60,\ -40\leqslant\mu_{T_e}\leqslant 125\}\\ 55\ \text{K} & i=\{\mu_I=70,\ -40\leqslant\mu_{T_e}\leqslant 125\}\end{cases} \tag{18}$$

式中:$i=\{I=70\ \text{A},T_e\in[-40,125]\,℃\}$ 的工作场景最常用。之后可由式(5)得到温升阈值 ΔT_{th}。

其次,内外影响因素分析。由于外部影响因素只需考虑电流和环境温度,在此主要对内部影响因素进行分析,得到各变量的类型、备选水平数和用以开展实验的各水平具体取值。

对于离散型变量,按照实际生产制造等条件将可选水平一一列出。例如,扭簧材料共有 C17410、C17510、QBe2.0 共 3 种可供选择,因此扭簧材料的水平数为 3。对于连续型变量,则根据设计需求在变量取值范围内确定水平。例如,扭簧的扭转角度可以在(0°,90°)间选取,而根据产品设计初期的经验和工艺条件,本文给出了 45°、50°及 55°这 3 个水平。为能更清晰地了解新型扭簧电连接器的正向设计中有哪些设计变量,水平如何选取,本文将 13 个影响因素及其水平选择均展示在表 1 中。

表 1 中,内因变量 $\boldsymbol{X}$ 共 13 个,7 个是离散型变量,6 个是连续型变量,包括加工工艺、粗糙度、材料、扭转角等。有 2 个变量是 2 水平,11 个变量是 3 水平。表中,D 表示离散型变量,C 表示连续型变量;孔套内径与扭簧材料厚度有关,孔套与扭簧直接配合,但二者需分开加工,将二者设计目标值分开考虑。

在此基础上,初步确定了扭簧电连接器的结构模型,如图 2 所示。

然后,基于正交实验的设计变量优选。根据表 1 分析的各变量水平取值,本文用有限元方法

表 1 内因变量分析

Table 1 Analysis of endogenous variables

序号	影响因素	变量类型	水平数	水平
1	加工工艺	D	2	工艺 A、工艺 B
2	粗糙度 Ra	D	3	0.8、1.6、3.2
3	镀层材料	D	3	金、银、镍
4	插针材料	D	3	T2 紫铜、磷铜、H62
5	孔套材料	D	3	T2 紫铜、磷铜、H62
6	扭簧材料	D	3	C17410、C17510、QBe2.0
7	扭簧材料厚度 δ	D	2	0.2 mm、0.25 mm
8	扭转角 θ	C	3	45°、50°、55°
9	格栅间距 s	C	3	0.6 mm、0.7 mm、0.8 mm
10	插针外径 d_p	C	3	3.57 mm、3.60 mm、3.63 mm
11	扭簧基圆直径 d_s	C	3	3.75 mm、3.80 mm、3.85 mm
12	孔套内径 d_h	C	3	$2\delta+\{3.75,3.80,3.85\}$ mm
13	格栅长度 l	C	3	7.25 mm、9.25 mm、11.25 mm

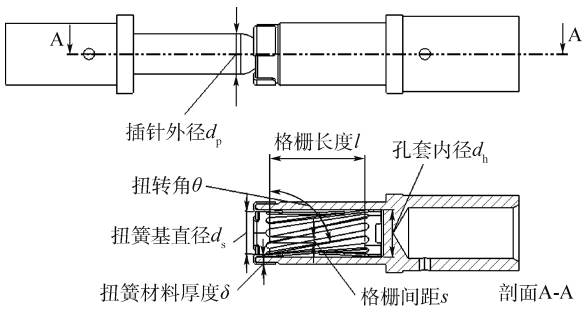


图 2 结构参数

Fig. 2 Structural parameters

开展正交实验。在温升仿真过程中,由于镀层和粗糙度水平的变化影响小,可暂时视为常量,用摩擦因数、热辐射率等材料参数等效。选择混合正交表 $L_{36}(3^9\times2^2)$ 对其余 11 个变量的取值组合进行仿真。基于效应分析评估各内因变量的主效应和交互作用,图 3 为正交实验数据结果的可视化展示。

由图 3(a)主效应图可以看出单变量对温升的影响。点越靠近下方代表在相应水平作用下的温升越低。以扭转角为例,可以看出,温升与角度成明显的负相关。而图 3(b)交互作用图反映变量间的耦合影响,折线的交叉反映变量间有交互作用,平行则代表变量间独立。例如,格栅长度和插针材料间交互作用明显。

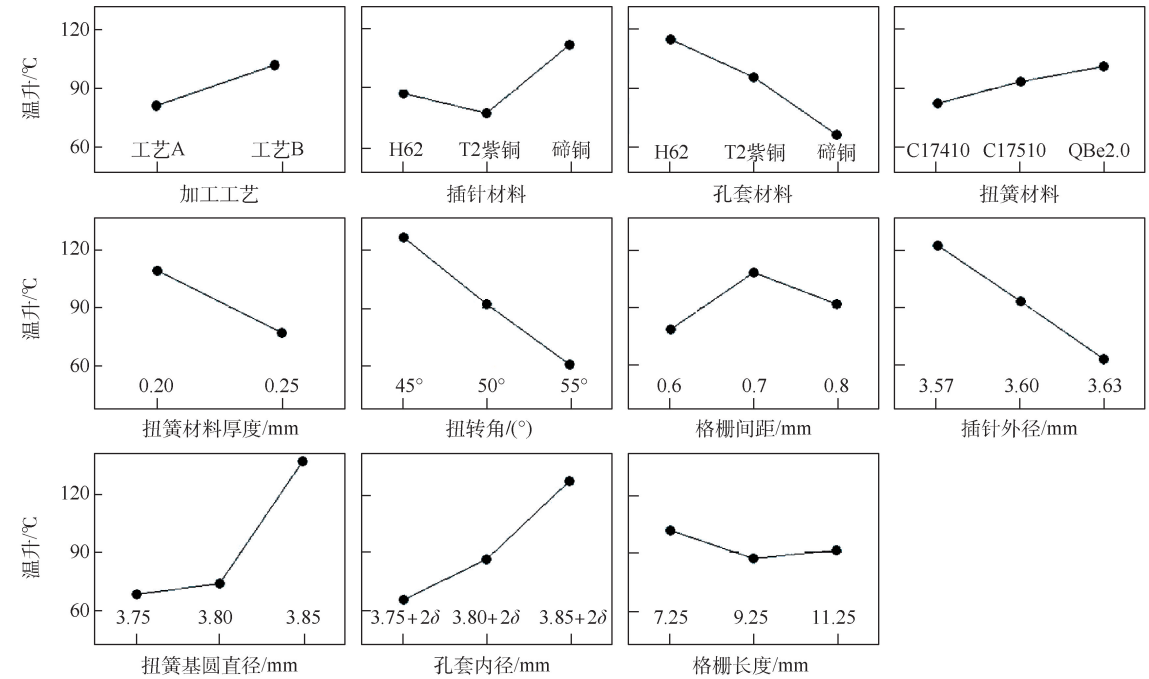
要开展初值优选,还需考虑 3 方面问题。①由于交互作用的存在,最优水平组合可能与主效应最优水平集不一致。例如,扭簧材料选择 C17410,扭簧材料厚度选择 0.25 mm 时各自的主效应最优,但当扭簧材料选择 C17410 时,扭簧材料厚度选择 0.2 mm 时的组合效果最佳。②用户需求发生变更,或者生产加工条件受限,所选水平无法作为设计值。例如,在原有加工工艺水平下工艺 A 更成熟,但用户要求采用新工艺 B。③水平选取的范围过于保守或过于粗放,实际最优值落在水平范围外或某两水平之间。例如,图 3(a)主效应图显示扭转角与温升基本呈线性关系,这说明 55°的设计值较为保守。再如,孔套内径与扭簧材料厚度和扭簧基圆直径的选择相关,当二者取值确定后,孔套内径取值范围落在 4.15 ~ 4.25 mm之间,这说明之前的水平确定过于粗放。因此,本文结合工艺难度、成本、产品定位及用户需求等多项因素更新离散型变量选型和连续型变量设计范围。例如,采用工艺 B 的设计方案后尽量通过改进工艺提高加工精度,通过增大扭转角设计范围进一步寻找最优设计值。

分别对扭转角、插针外径、孔套内径的水平取值作出变更后,在更新的最优水平组合下,进行粗糙度和镀层的 2 因素 3 水平全数实验,发现采用 Ra 为 0.8 的设计粗糙度和金作为镀层材料组合性能最优。基于图 3 的结果给出的设计变量初步优选水平,并在用户需求等因素的限制下进行了额外的判断和相应的调整后,表 2 给出了设计初值优选的输出结果。

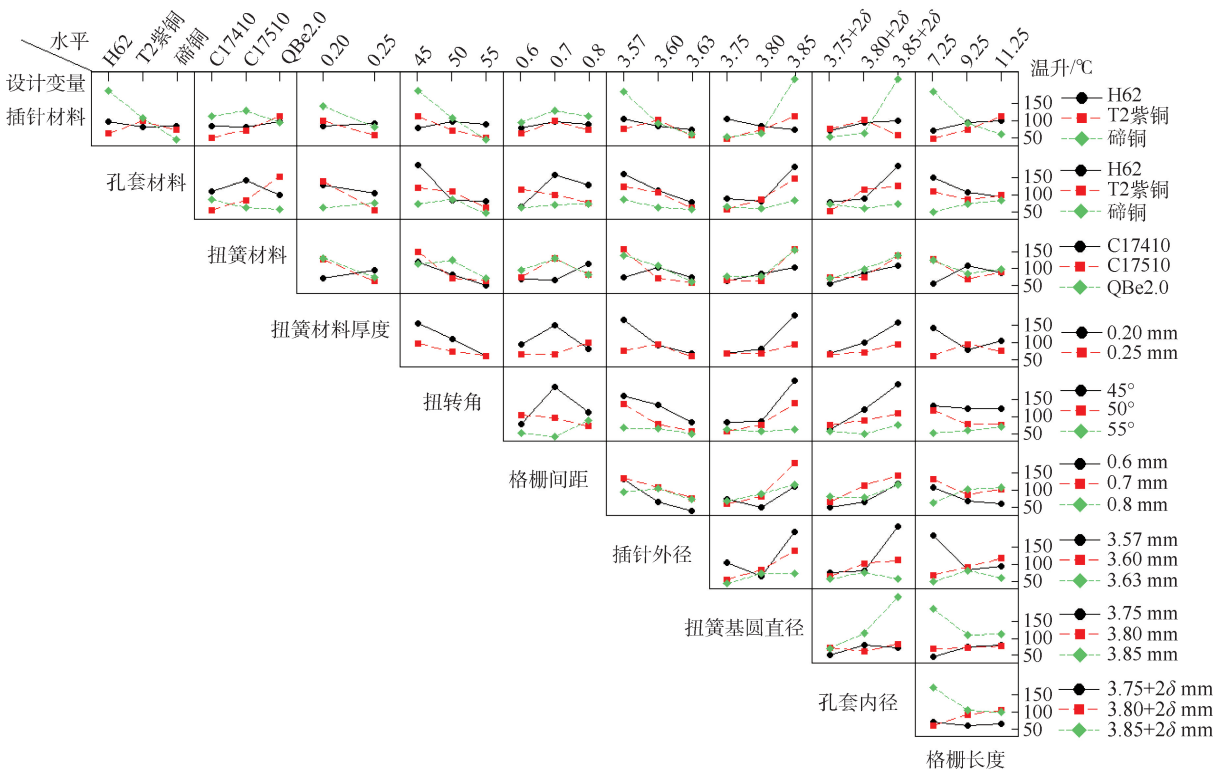
3.2 性能裕量建模

基于表 2 中的设计选型结果,离散型内因变量的选型可以直接确定,而各连续型内因变量和外因变量的最优设计值则需通过响应面法进一步确定,本文按表 3 设计响应面采样点。

需要注意的是,在定义电流的采样边界时,以额定电流上下界的平方作为电流响应的采样边界,从而放大电流平方的响应。定义环境温度采样边界时,虽然用户需求规定扭簧电连接器的工作环境温度大致在[-40,125]℃之间,但扭簧电连接器在相对密闭的环境中温度会高于连接器表



(a) 主效应图



(b) 交互作用图

图 3　效应分析结果

Fig. 3　Results of effect analysis

面温度,因此为使模型能更好覆盖实际工作剖面的环境温度,取响应面采样环境温度上边界为 150℃。

用二阶模型拟合后,经方差齐性检验(F-test)在 95% 的置信水平下得到了仅含显著项的归一化代理模型,该模型 F 值为 111.4,R^2 为0.854 1,有效信噪比为 39.06,说明拟合效果较好。

$$\Delta T = 29.17 - 3.79\theta^* + 2.49d_h^* + 2.07s^* + 14.91I^{2*} - 1.37T_e^* - 3.37d_s^* d_h^* - 2.14\theta^* I^{2*} + 3.23d_s^{*2} + 3.53\theta^{*2} + 2.22d_h^{*2} \quad (19)$$

按各项的显著性排序,影响温升的核心变量依次为电流、扭转角、孔套内径、格栅间距、环境温度、扭簧基圆直径,即外因变量与核心设计变量为

表 2 确定最优设计变量初值

Table 2 Determination of optimal initial values

序号	内因变量	数值特征	初步优选水平	是否变更	水平变更
1	加工工艺	D	工艺 A	√	工艺 B
2	粗糙度 Ra	D	0.8		
3	镀层材料	D	金		
4	插针材料	D	T2		
5	孔套材料	D	磷铜		
6	扭簧材料	D	C17410		
7	扭簧材料厚度 δ	D	0.2 mm		
8	扭转角 θ	C	55°	√	65 mm
9	格栅间距 s	C	0.6 mm		
10	插针外径 d_{p}	C	3.63 mm	√	3.57 mm
11	扭簧基圆直径 d_{s}	C	3.75 mm		
12	孔套内径 d_{h}	C	$2\delta+3.75$ mm	√	4.2 mm
13	格栅长度 l	C	7.25 mm		

表 3 响应面采样点

Table 3 Sampling points of response surface

变量	采样边界	
	-1	1
扭转角 θ/(°)	55	75
格栅间距 s/mm	0.5	0.9
插针外径 d_{p}/mm	3.57	3.60
扭簧基圆直径 d_{s}/mm	3.75	3.80
孔套内径 d_{h}/mm	4.2	4.25
格栅长度 l/mm	5.25	9.25
电流 I_i^2/A^2	1 600	4 900
环境温度 T_{e}/℃	-40	150

$$\begin{cases}\boldsymbol{\Theta}^{\star}=(\theta,d_{\mathrm{h}},s,d_{\mathrm{s}},I_i,T_{\mathrm{e}})\\ \boldsymbol{X}_C^{\star}=(\theta,d_{\mathrm{h}},s,d_{\mathrm{s}})\\ \boldsymbol{Y}^{\star}=\boldsymbol{Y}=(I_i,T_{\mathrm{e}})\end{cases}\tag{20}$$

由式(14)可得到扭簧电连接器温升的裕量方程:

$$\begin{aligned}\boldsymbol{M}=\Delta T_{\mathrm{th}}-(&29.17-3.79\theta^{*}+2.49d_{\mathrm{h}}^{*}+2.07s^{*}+\\&14.91I^{2*}-1.37T_{\mathrm{e}}^{*}-3.37d_{\mathrm{s}}^{*}d_{\mathrm{h}}^{*}-2.14\theta^{*}I^{2*}+\\&3.23d_{\mathrm{s}}^{*2}+3.53\theta^{*2}+2.22d_{\mathrm{h}}^{*2})\end{aligned}\tag{21}$$

3.3 不确定性分析与量化

由裕量方程可知,模型形式及系数不确定性、内外因变量的不确定性、性能阈值的不确定性同时存在,且有复杂的关联影响。

首先,模型形式及系数的不确定性分析量化。假设随机扰动 ε 服从标准正态分布,本文基于响应面实验结果,用式(22)算出了 σ 的无偏估计。其中,N 为采样点总数,p 为模型的显著项项数,e 为响应面残差。

$$\hat{\sigma}^2=\sum e_i^2/(N-p-1)\tag{22}$$

之后,内外因变量本身的不确定性分析量化。对于扭簧基圆直径、孔套内径、格栅间距这些加工精度较高的内因变量,认为它们服从以设计值为均值,加工/测量仪器精度为标准差的正态分布。而扭转角由于采用新工艺加工,且有一定的测量难度,本文采用均值已知标准差未知的样本假设,通过收集实测数据,求出了扭转角的极大似然估计,经皮尔逊卡方检验以及偏峰度检验,正态分布假设成立。电流和环境温度这 2 个外因变量则服从以额定工作条件为均值,波动范围为 3σ 的正态分布。

最后,性能阈值的不确定性分析量化。温升阈值的期望和方差可根据阈值与电流、环境温度、工作场景的函数间接求解。

在上述分析与假设下,表 4 给出了所有随机变量的正态分布参数。

3.4 确信可靠性分析与优化

基于上述分析,本文构建如下确信可靠性优化模型。由于扭簧电连接器尺寸小,设计值变化范围小,用料对成本影响小,现阶段的成本主要由加工精度和工艺决定,优化模型也就转化为标准差不变的随机期望优化问题。

$$\begin{aligned}&\text{Maximize}\quad \bigwedge_{i\in V}R_i\\&\text{s.t.}\begin{cases}\sigma_X=\sigma_{X_0}\\ 3.75\leqslant\mu_{d_{\mathrm{s}}}\leqslant 3.80\\ 55\leqslant\mu_{\theta}\leqslant 75\\ 4.2\leqslant\mu_{d_{\mathrm{h}}}\leqslant 4.25\\ 0.5\leqslant\mu_{s}\leqslant 0.9\end{cases}\end{aligned}\tag{23}$$

其中,期望的变化范围即设计变量的取值范围。为了求解该模型,本文使用蒙特卡罗方法求解确信可靠度,以 3.1 节中优选得到的设计方案为初值 $\boldsymbol{X}_0$,用模拟退火的启发式方法求解设计变量最优取值 $\boldsymbol{X}_{\mathrm{opt}}$。

表 5 展示了使得初始确信可靠度最高的设计方案。可见,扭簧基圆直径和扭转角的设计值对可靠度的影响存在拐点,最优值分别为 3.773 mm

表 4　随机变量的分布

Table 4　Distribution of random variables

变量	分布	均值 μ	标准差 σ
随机扰动 ε	Φ_{ε}	0	5.35℃
扭簧基圆直径 d_s	Φ_{d_s}	μ_{d_s}	0.001 mm
扭转角 θ	Φ_{θ}	μ_{θ}	4°
孔套内径 d_h	Φ_{d_p}	μ_{d_p}	0.01 mm
格栅间距 s	Φ_{s}	μ_{s}	0.01 mm
电流 I_i	Φ_{I}	μ_{I}	$0.05\mu_I/3$ A
环境温度 T_e	Φ_{T_e}	μ_{T_e}	5/3℃
温升阈值 ΔT_{th}	$\Phi_{\Delta T_{th}}$	$E[\Delta T_{th}(I,T_e)]$	$\sqrt{D[\Delta T_{th}(I,T_e)]}$℃

表 5　确信可靠度优化结果

Table 5　Results of belief reliability optimization

变量	取值范围	正向设计优化	
		初值 $\boldsymbol{X}_0$	优化值 $\boldsymbol{X}_{opt}$
扭簧基圆直径 d_s/mm	[3.75,3.80]	3.75	3.773
扭转角 θ/(°)	[55,75]	65	73.87
孔套内径 d_h/mm	[4.20,4.25]	4.2	4.202
格栅间距 s/mm	[0.5,0.9]	0.6	0.5
确信可靠度理论值 R		0.978 5	0.994 7

和 73.87°。孔套内径和格栅间距的最优设计值则是各自设计范围的下界,即 4.202 mm 和 0.5 mm。由此可以进一步更新设计值,从而定量提高扭簧电连接器的可靠度。

进一步地,本文用蒙特卡罗方法分别计算了在 3 种额定电流下,初值和优化值对应的温升、温升阈值。图 4 展示了优化前后温升-阈值对比结果。其中,$f_{\Delta T}(x)$ 为温升的概率密度函数,$\Delta T_{X=X_0}$ 和 $\Delta T_{X=X_{opt}}$ 分别表示初值和优化值下的温升,ΔT_{th} 为温升阈值。

在 3 种工作场景中,温升随着额定电流的增大明显升高,$I=70$ A 是最严酷的工作条件。

可看出由于不确定性的作用,扭簧电连接器虽然在设计初值下能保持较低的温升,但仍存在较大的温升超过阈值的风险。而基于上述优化模型得到的优化值,本质上是降低了温升的均值,从而增大了裕量,使产品更可靠。

为了更直观地展示裕量的优化效果,图 5 对优化前后的裕量分布进行对比。其中,$f_m(x)$ 为性能裕量的概率密度函数,$F_m(x)$ 为性能裕量的累积分布函数。

可见同初值相比,优化值下扭簧电连接器的裕量显著提高,尤其是最恶劣环境 $I=70$ A 下优化效果最为明显,裕量分布的均值增大,标准差减小,仅有少数仿真结果中裕量小于 0。由此实现了在多种工作场景下,产品初始确信可靠度的定量提高。

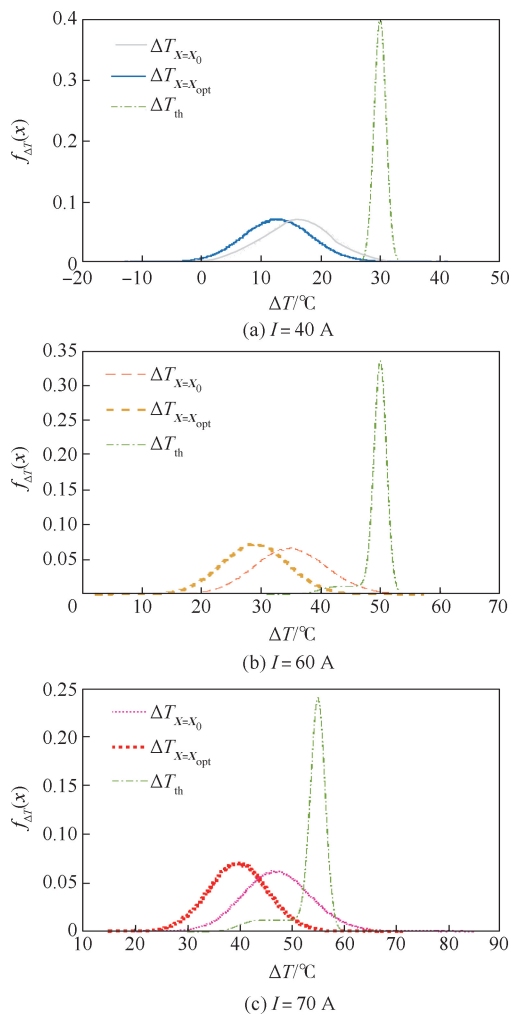
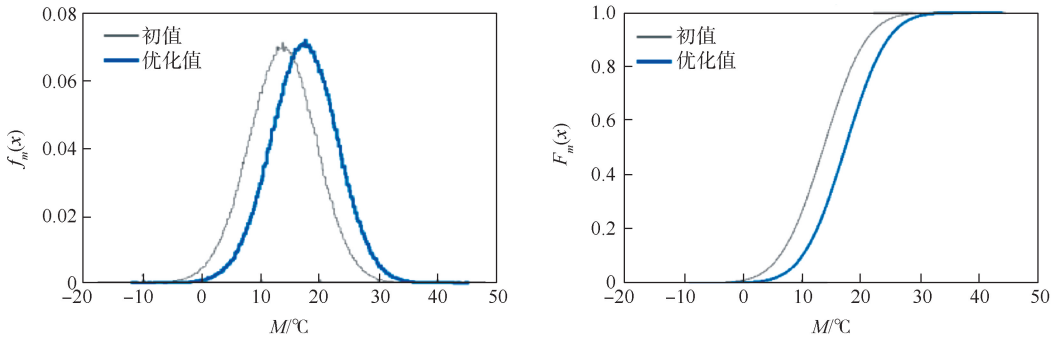
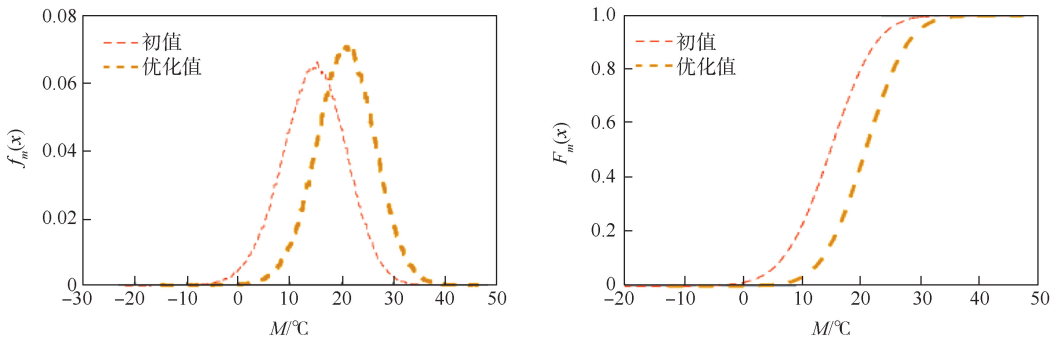


图 4　优化前后温升分布对比

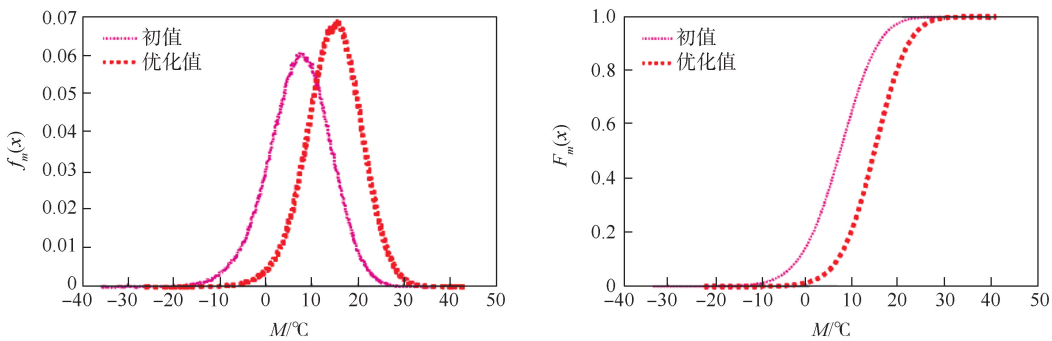
Fig. 4　Comparison of distributions of temperature rise before and after optimization



(a) $I=40$ A



(b) $I=60$ A



(c) $I=70$ A

图 5 优化前后裕量分布对比

Fig. 5 Comparison of margin distribution before and after optimization

4 结 论

本文依托工程案例系统性总结了新型扭簧电连接器的确信可靠性正向设计优化方法,将正向设计优化划分为周期迭代过程,每周期按照设计变量初值优选、性能裕量建模、不确定性分析量化、确信可靠性分析与优化 4 个阶段依次展开,从而在确信可靠性理论框架下考虑多维离散-连续变量及其不确定性作用,实现新产品的正向优化设计。主要贡献如下:

1)面向新型扭簧电连接器的可靠性正向设计需求,提出了基于确信可靠性的系统化可靠性正向设计方法;

2)综合考虑和量化了选材、工艺、手册标准、外界环境、工作条件等多源不确定性影响,构建了确信可靠性设计优化模型,实现了多维离散-连续设计变量下的设计方案优化;

3)通过开展正向设计的实际案例研究,验证了设计变量优选、性能裕量建模、确信可靠性优化建模等方法的有效性,可为其他新产品的可靠性正向设计提供参考。

需要指出的是,在案例分析阶段,本文基于实际案例只进行了初始确信可靠度的正向设计优化。因此,在未来研究中,还应根据设计需求开展考虑产品退化的可靠寿命优化设计方法研究。

参考文献(References)

[1] 李志萍. 电动汽车传导式充电接口关键技术研究[D]. 成都:电子科技大学, 2014:18-21.
LI Z P. Research on key techniques of conductive charge coupler of electric vehicle[D]. Chengdu: University of Electronic Science and Technology of China, 2014:18-21(in Chinese).

[2] 吴珊珊. 电动汽车充电连接器的热性能研究[D]. 北京:北

京邮电大学,2020:4-7.
WU S S. Study on thermal performance of electric vehicle connectors[D]. Beijing: Beijing University of Posts and Telecommunications,2020:4-7(in Chinese).

[3] 王世娇,陈文华,钱萍,等. 航天电连接器的可靠性设计建模[J]. 机械工程学报,2017,53(10):180-186.
WANG S J,CHEN W H,QIAN P,et al. Reliability design model of aerospace electrical connector[J]. Journal of Mechanical Engineering,2017,53(10):180-186(in Chinese).

[4] 汪日超,张伟,徐佳济,等. 电连接器接触件结构分析与结构参数优化设计[J]. 机械制造与自动化,2021,50(3):136-139.
WANG R C,ZHANG W,XU J J,et al. Structural ansysis of electrical connector contacts and optimization design of structural parameters[J]. Machine Building & Automation,2021,50(3):136-139(in Chinese).

[5] 关凌宇. 线簧连接器的特性研究[D]. 北京:北京邮电大学,2018:25-28.
GUAN L Y. Research on the characteristics of wire spring connector[D]. Beijing: Beijing University of Posts and Telecommunications,2018:25-28(in Chinese).

[6] KYEONG S,PECHT M. Electrical connectors[M]. New York: Wiley,2020:173-196.

[7] 田楠. 连接器不同触点材料滑动接触特性及可靠性分析[D]. 北京:北京邮电大学,2019:11-15.
TIAN N. Sliding contact characteristics and reliability analysis of different contact materials of connector[D]. Beijing: Beijing University of Posts and Telecommunications,2019:11-15(in Chinese).

[8] 凌三强,徐乐,付饶,等. 基于稳健性设计原理的接触件插拔力质量一致性优化方法[J]. 机械工程学报,2017,53(4):190-197.
LING S Q,XU L,FU R,et al. Optimization method of quality consistency for insertion force of electrical contact based on robust design principle[J]. Journal of Mechanical Engineering,2017,53(4):190-197(in Chinese).

[9] WANG R Y,XU L J,ZHOU Y L. A systematic approach for the reliability evaluation of electric connector[J]. Journal of Electrical and Computer Engineering,2021,2021:5514674.

[10] JENSEN H A,SEPULVEDA J G. Structural optimization of uncertain dynamical systems considering mixed-design variables[J]. Probabilistic Engineering Mechanics,2011,26(2):269-280.

[11] ZHU N,ZHAI G F,GUO Z Y. Research on contact reliability design of electrical connector contact[C]//2019 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE). Piscataway: IEEE Press,2019:155-160.

[12] 康锐. 确信可靠性理论与方法[M]. 北京:国防工业出版社,2020:26-29.
KANG R. Belief reliability theory and methodology[M]. Beijing: National Defense Industry Press,2020:26-29(in Chinese).

[13] 张清源,文美林,康锐,等. 基于确信可靠性的功能、性能与裕量分析方法[J]. 系统工程与电子技术,2021,43(5):1413-1419.
ZHANG Q Y,WEN M L,KANG R,et al. Function,performance and margin analysis method based on belief reliability[J]. Systems Engineering and Electronics,2021,43(5):1413-1419(in Chinese).

[14] 范梦飞,曾志国,康锐. 基于确信可靠度的可靠性评价方法[J]. 系统工程与电子技术,2015,37(11):2648-2653.
FAN M F,ZENG Z G,KANG R. An approach to measure reliability based on belief reliability[J]. Systems Engineering and Electronics,2015,37(11):2648-2653(in Chinese).

[15] 李泊远,康锐,余丽. 基于概率测度的确信可靠性建模与分析方法[J]. 系统工程与电子技术,2021,43(7):1995-2004.
LI B Y,KANG R,YU L. Belief reliability modelling and analysis method based on the probability measure[J]. Systems Engineering and Electronics,2021,43(7):1995-2004(in Chinese).

[16] 陈文彬,李晓阳,童邦安,等. 谐波减速器的传动效率确信可靠性建模与分析[J]. 振动工程学报,2022,35(1):237-245.
CHEN W B,LI X Y,TONG B G,et al. Belief reliability modeling and analysis for transmission efficiency of harmonic gear[J]. Journal of Vibration Engineering,2022,35(1):237-245(in Chinese).

[17] 李晓阳. 加速退化试验——不确定性量化与控制[M]. 北京:国防工业出版社,2022:25-26.
LI X Y. Accelerated degradation testing: Quantification and control of uncertainties[M]. Beijing: National Defense Industry Press,2022:25-26(in Chinese).

[18] 汪天宇. 大功率连接器的可靠性研究与设计[D]. 北京:北京邮电大学,2021:1-7.
WANG T Y. Research and reliability design of high power connector[D]. Beijing: Beijing University of Posts and Telecommunications,2021:1-7(in Chinese).

[19] USCAR. Performance specification for automotive electrical connector systems: SAE/USCAR-2 Rev. 7[S]. Warrendale: SAE International,2020:22-26.

[20] 陈孟,黄庆,翁羽,等. 自然对流换热系数计算方法研究[C]//第十七届全国反应堆结构力学会议,2012:217-221.
CHEN M,HUANG Q,WENG Y,et. al. The computational method for heat transfer film coefficient of nature convection[C]//Proceedings of the 17th National Conference on Reactor Structural Mechanics,2012:217-221(in Chinese).

[21] 王静. 高强度正交表的构造及 D-最优增广设计[D]. 上海:上海师范大学,2021:1-5.
WANG J. Construction of orthogonal arrays with high strength and D-optimal augmented designs[D]. Shanghai: Shanghai Normal University,2021:1-5(in Chinese).

[22] KAYMAZ I,MCMAHON C A. A response surface method based on weighted regression for structural reliability analysis[J]. Probabilistic Engineering Mechanics,2005,20(1):11-17.

[23] HUANG X Z,LIU Y,ZHANG Y M,et al. Reliability analysis of structures using stochastic response surface method and saddlepoint approximation[J]. Structural and Multidisciplinary Optimization,2017,55(6):2003-2012.

[24] IEC. Appliance couplers for household and similar general purposes-Part 1: General requirements: IEC 60320-1 ED4[S]. Ge-

neva:IEC,2021:36-38,68-71.

[25] 国家质量监督检验检疫总局,中国国家标准化管理委员会.电动汽车传导充电用连接装置第1部分:通用要求:GB/T 20234.1—2015[S].北京:中国标准出版社,2016:3-10.

General administration of quality supervision, inspection and quarantine, China national standardization administration committee. Connection set of conductive charging for electric vehicles-Part 1: General requirements: GB/T 20234.1—2015[S]. Bejing: China Standard Press, 2016: 3-10(in Chinese).

Function oriented belief reliability design and optimization of new torsion spring electrical connectors

KANG Rui[1,2], LIU Haoran[1,2], ZHANG Qingyuan[2,3,*], YU Li[1,2], ZHAI Guofu[4]

(1. School of Reliability and Systems Engineering, Beihang University, Beijing 100083, China;
2. Science & Technology on Reliability & Environmental Engineering Laboratory, Beijing 100083, China;
3. School of Aeronautic Science and Engineering, Beihang University, Beijing 100083, China;
4. School of Electrical Engineering and Automation, Harbin Institute of Technology, Harbin 150001, China)

Abstract: To evaluate reliability of new products, function-oriented design should comprehensively consider the influence of new structure and technology, multiple design variables, and complex working environments, and reasonably quantify uncertainty effects. This will ensure high reliability from the original design. Based on the belief reliability theory, this study investigates a performance margin-based and function-oriented method for belief reliability design and optimization. The basic procedure of the function-oriented reliability design is proposed for the development of new torsion spring electrical connectors, which includes four steps, i.e., initial value selection of design variables, performance margin modeling, uncertainty quantification, and reliability analysis and optimization. This procedure is exemplified in the design and development of a new type of torsion spring electrical connector. The optimal design of multidimensional discrete-continuous variables under the influence of their uncertainties is realized based on orthogonal experiments, response surface modeling, and simulated annealing heuristic optimization.

Keywords: belief reliability; performance margin; function oriented design and optimization; design of experiment; multidimensional discrete-continuous variables

Received: 2022-05-06; **Accepted**: 2022-05-27; **Published online**: 2022-06-14 09:16

URL: kns.cnki.net/kcms/detail/11.2625.V.20220613.1329.001.html

Foundation items: National Natural Science Foundation of China (62073009); Fund for Science & Technology on Reliability & Environmental Engineering Laboratory (6142004210401)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0608

流动拓扑对液体射流失稳与雾化的影响

杨立军[1,2], 黄东骐[1], 韩旺[1,2], 李敬轩[1,2], 富庆飞[1,2,*]

(1. 北京航空航天大学 宇航学院, 北京 100083; 2. 北京航空航天大学宁波创新研究院, 宁波 315800)

摘 要: 在大推力液体火箭发动机燃烧过程中,推进剂射流失稳与雾化是起始环节,会对后续蒸发与燃烧等过程产生显著影响。尽管前人做过很多研究,但对湍流射流雾化机理的认知还存在盲区。基于此通过流动拓扑理论来揭示湍流液体平面射流的雾化机理。采用直接数值模拟方法对静止空气环境下的液体平面射流雾化过程进行了高分辨率数值模拟,分析了流场中不同拓扑结构与气液界面曲率的相互影响,阐明了流动拓扑对液体平面射流雾化的影响机制。研究发现,所有流动拓扑结构都有助于产生压缩应变率和拉伸应变率,其中不稳定焦点结构(UFC)拓扑结构对流场应变率的影响最大;在流动拓扑结构影响下,液体体积分数等值面的曲率与应变呈现负相关关系。另外,UFC 主要产生拉伸应变率,而其余流动拓扑结构主要产生压缩应变率。研究结果表明:射流雾化过程主要受到 UFC 拓扑结构的影响,UFC 会促进气液界面产生较大的拉伸应变率,进而促进片状或管状结构液体结构生成,从而引起液体射流破碎。

关 键 词: 射流雾化; 平面射流; 流动拓扑; 界面曲率; 应变率
中图分类号: V221+.3; TB553
文献标志码: A **文章编号**: 1001-5965(2022)09-1757-10

推进剂雾化包括液体射流失稳、断裂、破碎、液滴生成等一系列过程。推进剂的雾化性能会显著影响发动机性能;研究射流雾化机理有助于实现对雾化过程的精准调控,从而优化后续的蒸发燃烧过程,提升发动机的推力性能。尽管前人对液体射流雾化开展了大量的研究,但对其雾化机理还存在认识不清的地方[1]。

射流雾化的研究方法主要分为 2 类:实验测量和数值模拟。实验测量方法的优势在于容易开展参数可调的大规模研究,但也面临着测量精度有限,测量物理量较少等挑战。数值模拟方法虽然计算量较大,但能够提供更高的精度,同时也能提供更多实验中不易获得的数据,帮助开展更加细致的研究,因此受到广泛的关注。

在雾化的直接数值模拟(DNS)研究中,通常采用水平集[2]、流体体积[3]等方法来捕捉气液界面。例如,Leboissetier 和 Zaleski[4] 最早对柴油燃料射流进行了 DNS,该研究表明喷嘴内部的液体射流在层流条件下不会发生破裂雾化。这一结果与实验对比显示出了较为良好的一致性。对于喷嘴外部的射流,Klein[5] 指出,当平面液体射流的韦伯数小于 270 时,没有发现射流破裂及雾化的行为。Sander 和 Weigand[6] 发现,除了喷嘴出口处的湍流特性会对射流液面的稳定性造成影响,速度型也会显著影响射流雾化的结果。Ménard 等[7] 首次在两相环境下研究了柴油燃料射流 DNS 的网格分辨率问题;Desjardins 和 Pitsch[8] 则使用 DNS 研究了中等韦伯数下(We = 500 ~ 2 000)湍流

收稿日期: 2022-07-11; **录用日期**: 2022-08-14; **网络出版时间**: 2022-08-19 16:16
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220819.1337.002.html
基金项目: 国家自然科学基金 (11922201,11872091,11927802,U1837211)
* **通信作者**. E-mail: fuqingfei@buaa.edu.cn

引用格式: *杨立军, 黄东骐, 韩旺, 等. 流动拓扑对液体射流失稳与雾化的影响[J]. 北京航空航天大学学报, 2022, 48(9): 1757-1766. YANG L J, HUANG D Q, HAN W, et al. Influence of flow topology on instability and atomization of liquid jets[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1757-1766(in Chinese).*

射流液面的破裂行为。Shinjo 和 Umemura[9] 在层流环境及静止空气中，在大韦伯数下（We = 14 000）使用高分辨率网格对圆柱柴油燃料射流进行了数值模拟。

以上数值模拟研究展现了较为清晰的射流雾化过程，发现韦伯数和流场环境是影响射流雾化的重要因素，但是有关射流雾化的详细机制仍然不清。基于此，本文尝试通过流动拓扑理论分析液体射流雾化机理。流动拓扑理论由 Chong 等[10-11] 提出。流动拓扑理论主要是基于速度梯度张量的 3 个不变量来定义的不同流场拓扑结构，这种方法被广泛运用于解决流体问题及燃烧问题。Chong 等[12] 运用该方法研究了壁面剪切流动，Elsinga 和 Marusic[13] 研究了均匀且各向同性的湍流流动。Hasslberger 等[14] 首次在气泡空气和水的两相流流动中研究了流动拓扑的分布，并发现代表旋转运动的拓扑结构主要存在于气泡内部。Watanabe 等[15] 研究了在湍流与非湍流界面附近的拓扑结构。Josef 等[16] 研究了柴油燃料射流雾化过程中的流动拓扑结构影响；Han 等[17] 研究了湍流流动拓扑结构对湍流非预混火焰切向扩散的影响，从流动拓扑结构角度给出了经典火焰面模型假设不成立的条件。

基于上述讨论，本文采用水平集方法，对湍流液体平面射流直接进行三维数值模拟。水平集方法有 2 个优点：①相比于传统的欧拉法，水平集方法可以直接在直角网格中对曲线和曲面进行数值计算，而不用将曲线和曲面参数化；②水平集方法易于追踪物体结构的改变。因此，使用水平集方法进行 DNS 可以有效提高计算精度，并减少计算成本。本文通过使用水平集方法得到高分辨率数值模拟数据，从流动拓扑角度揭示液体平面射流雾化的机理。

1 模型与算法

1.1 控制方程与算例设置

考虑在静止气体环境中的液体平面射流，其液体射流湍流初速度场在一个槽道中产生（见图 1，右侧液体初始湍流场由左侧充分发展的槽道流速度场提供）。射流的质量方程和动量方程由经典的不可压缩 Navier-Stokes 方程给出：

$$\frac{\partial \boldsymbol{u}}{\partial t}+\boldsymbol{u}\cdot\nabla\boldsymbol{u}=-\frac{1}{\rho}\nabla p+\frac{1}{\rho}\nabla\cdot(\mu[\nabla\boldsymbol{u}+\nabla\boldsymbol{u}^{\mathrm{T}}]) \tag{1}$$

$$\frac{\partial \rho}{\partial t}+\nabla\cdot(\rho\boldsymbol{u})=\frac{\partial \rho}{\partial t}+\boldsymbol{u}\cdot\nabla\rho=0 \tag{2}$$

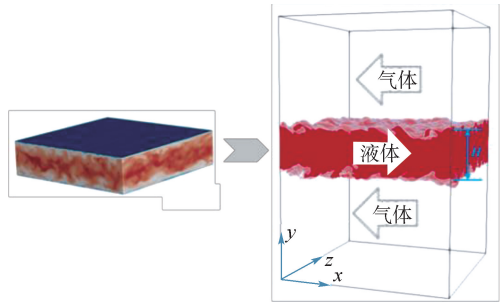


图 1 湍流液体平面射流算例设置

Fig. 1 Turbulent liquid plane jet arithmetic setup

式中：$\boldsymbol{u}$ 为速度场；ρ 为密度；p 为压强；μ 为黏度。

界面 Γ 将液相和气相分隔开。值得注意的是，在气液多相流中，物理量在两相界面处存在突变。令液相密度 $\rho=\rho_{\mathrm{l}}$，气相密度 $\rho=\rho_{\mathrm{g}}$，液相动力黏度 $\mu=\mu_{\mathrm{l}}$，气相动力黏度 $\mu=\mu_{\mathrm{g}}$。因此，在界面 Γ 处满足：$[\rho]_{\Gamma}=\rho_{\mathrm{l}}-\rho_{\mathrm{g}}$，$[\mu]_{\Gamma}=\mu_{\mathrm{l}}-\mu_{\mathrm{g}}$。而由于速度连续，在界面处 $[\boldsymbol{u}]_{\Gamma}=0$。影响气液界面传输计算精度的主要是如下的压力跳跃：

$$[P]_{\Gamma}=\sigma_{\kappa}+2[\mu]_{\Gamma}\boldsymbol{n}^{\mathrm{T}}\cdot(\nabla\boldsymbol{u})\cdot\boldsymbol{n} \tag{3}$$

式中：σ 为表面张力系数；κ 为界面曲率；$\boldsymbol{n}$ 为界面法向量。

本文使用低马赫多相湍流燃烧求解器 NGA[18] 对式（1）~式（3）方程进行数值求解。在进行数值模拟时，为解决界面处的压力跳跃和密度跳跃，本文使用了精确水平集守恒方法 ACLS，在处理表面张力时应用了虚拟流动方法 GFM[19]。

图 1 展示了流动计算域。射流初始高度为 H，入口处平均速度为 U_0，整个计算域在流向和展向上（即 x，z 方向）是周期性边界条件。在流动高度方向（即 y 方向）是出口边界条件。整个计算域在 x-y-z 方向上尺寸是 $4H\times6H\times4H$，对应每个方向的网格数目是 256 × 384 × 256，网格大小为 $H/64$[8]。而提供湍流初速度流场的槽道流计算域如图 1 所示，其计算尺寸是 $4H\times H\times4H$，对应每个方向的网格数目是 256 × 64 × 256，槽道流的高度为 H，与后续射流初始高度对应。射流初始高度处于 $-H/2<y<H/2$ 的范围内，液体射流的雷诺数 Re = 5 000，韦伯数 We = 1 000，其中雷诺数 $Re=\rho U_0H/\mu$，韦伯数 $We=\rho U_0^2H/\sigma$。气液密度比被设置为 $\rho_{\mathrm{l}}/\rho_{\mathrm{g}}=40$，为使气液运动黏度相等，设置 $\mu_{\mathrm{l}}/\mu_{\mathrm{g}}=40$。

图 2 展示了不同时刻下气液界面的演化。本文气液界面通过液体体积百分数为 0.5 的等值面表征，图 2（a）~（d）分别对应的时间是 0.01，0.03，0.06，0.10 ms 时刻，对应界面失稳，表面波增长，界面出现破碎，射流初次雾化的时间点。

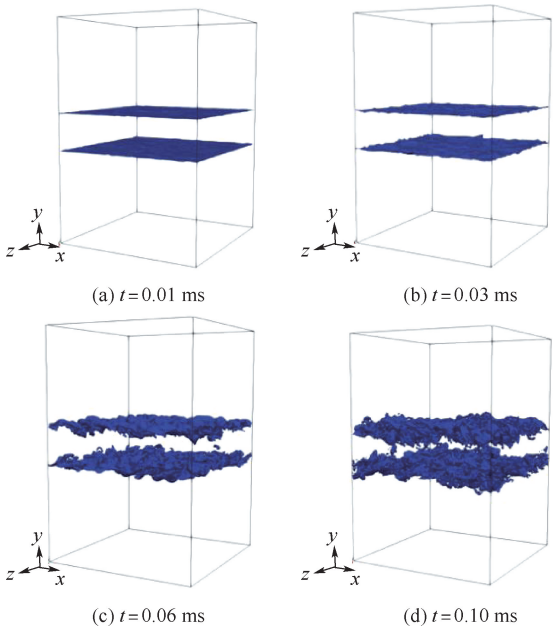


图 2　射流气液界面的演化

Fig. 2　Evolution of jet gas-liquid interface

1.2　流动拓扑理论

本节介绍流动拓扑分析方法。根据 Perry 和 Chong 等[10-12]的研究,速度梯度张量 $\boldsymbol{A}_{ij}$ 由式(4)给出:

$$\boldsymbol{A}_{ij} \equiv \frac{\partial \boldsymbol{u}_i}{\partial \boldsymbol{x}_j} = \boldsymbol{S}_{ij} + \boldsymbol{W}_{ij} \tag{4}$$

同时,可将速度梯度张量分解为 $\boldsymbol{S}_{ij}$ 和 $\boldsymbol{W}_{ij}$ 两部分之和,其中 $\boldsymbol{S}_{ij}=0.5(\boldsymbol{A}_{ij}+\boldsymbol{A}_{ji})$,$\boldsymbol{W}_{ij}=0.5(\boldsymbol{A}_{ij}-\boldsymbol{A}_{ji})$,分别称为对称张量和反对称张量。速度梯度张量 $\boldsymbol{A}_{ij}$ 的 3 个不变量由式(5)给出:

$$\lambda^3 + P\lambda^2 + Q\lambda + R = 0 \tag{5}$$

式中:P、Q、R 分别为张量的第 1、第 2、第 3 不变量。P、Q 和 R 分别由下式给出:

$$P = -\boldsymbol{S}_{ii} = 0 \tag{6}$$

$$Q = Q_S + Q_W = -\boldsymbol{S}_{ij}\boldsymbol{S}_{ji}/2 + \boldsymbol{W}_{ij}\boldsymbol{W}_{ji}/2 \tag{7}$$

$$R = (-\boldsymbol{S}_{ij}\boldsymbol{S}_{jk}\boldsymbol{S}_{ki} - 3\boldsymbol{W}_{ij}\boldsymbol{W}_{jk}\boldsymbol{W}_{ki})/3 \tag{8}$$

显然,第 1 不变量 P 代表流体体积的变化,由质量守恒知 $P=0$,因此本文将主要在 Q-R 平面内分析流动拓扑结构。式(5)是一元三次方程,其判别式为

$$\Delta = (27R^2 + 4Q^3)/108 \tag{9}$$

判别式将 Q-R 平面分为 2 个区域,当 $\Delta>0$,方程有 1 个实数根和 2 个复数根,当 $\Delta<0$,方程有 3 个实数根,当 $\Delta=0$ 时,对应 2 条直线,这 2 条直线分离不同的拓扑结构区域(见图 3):

$$r_{1a} = -2(-3Q)^{3/2}/27 \tag{10}$$

$$r_{1b} = +2(-3Q)^{3/2}/27 \tag{11}$$

当 2 个复数根为纯虚数时,对应 $r_2=0$。这 3 条直线 r_{1a}、r_{1b}、r_2 将整个 Q-R 平面分为 4 个区域,分别命名为 S1 ~ S4,如图 3 所示。利用速度梯度张量的第 2 不变量 Q 可以解释每个部分所表示的物理意义,Q 可以被分成 2 个部分:

$$Q = Q_S + Q_W = -\boldsymbol{S}_{ij}\boldsymbol{S}_{ji}/2 + \boldsymbol{W}_{ij}\boldsymbol{W}_{ji}/2 \tag{12}$$

式中:Q_S 和 Q_W 分别为 $\boldsymbol{S}_{ij}$ 和 $\boldsymbol{W}_{ij}$ 的第 2 张量不变量。Q_S 代表每单位质量的动能耗散为热能,代表变形;Q_W 与涡量相关,代表旋转,因此 $Q>0$,代表系统以旋转为主,$Q<0$,则代表系统以变形为主。同时 $R>0$,代表压缩行为,$R<0$,代表拉伸行为。通过流动拓扑分析,可以看出在不同的拓扑区域内,流动受到不同的机制控制。总而言之,S1 拓扑区域代表不稳定焦点结构(UFC),主要受旋转压缩机制影响。S2 拓扑区域代表不稳定节点/鞍点结构(UN/S/S),主要受到压缩变形机制影响。S3 拓扑区域代表稳定节点/鞍点结构(SN/S/S),主要受到拉伸变形机制影响。S4 拓扑区域代表稳定焦点结构(SFS),主要受到旋转拉伸机制影响。

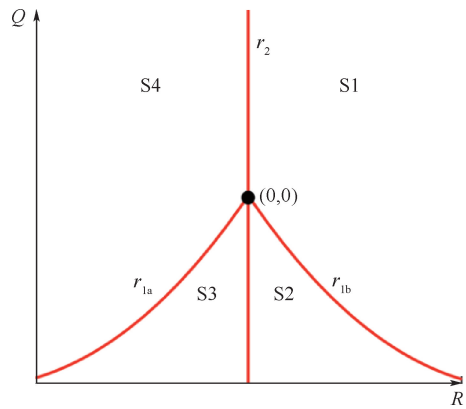


图 3　不同拓扑区域分布示意图

Fig. 3　Different topology area distribution diagram

1.3　气液界面曲率

为了研究流动拓扑与气液界面的耦合关系,还需研究气液界面曲率。下面对界面曲率给出定义。本文中所考虑的界面是液体体积分数等值面,以 α 表示。$\alpha=1$ 时表示全为液相,$\alpha=0$ 时表示全为气相。该等值面的法向量 $\boldsymbol{n}$ 定义为

$$\boldsymbol{n} = \frac{\nabla\alpha}{|\nabla\alpha|} \tag{13}$$

类比流动速度张量的 3 个不变量,定义界面的曲率张量如下:

$$\boldsymbol{n}_{i,j} = \partial \boldsymbol{n}_i/\partial \boldsymbol{x}_j \tag{14}$$

同理,也可以找出 3 个不变量,此处以 I_1、I_2、I_3 表示,分别表示为

$$I_1 = -\nabla\cdot\boldsymbol{n} \tag{15}$$

$$I_2 = (\boldsymbol{n}_{i,i}\boldsymbol{n}_{j,j} - \boldsymbol{n}_{i,j}\boldsymbol{n}_{j,i})/2 \tag{16}$$

$$I_3 = -\det(\boldsymbol{n}_{i,j}) \tag{17}$$

由于向量大小恒定,满足 $I_3 = 0$。因此,只剩下第 1、第 2 张量不变量,对应的曲率张量的特征值方程为

$$\kappa^2 + I_1\kappa + I_2 = 0 \tag{18}$$

方程(18)的 2 个解分别对应着 1 个点附近的最大曲率 κ_1 和最小曲率 κ_2,因此,引入高斯曲率 κ_g 和平均曲率 κ_m:

$$\kappa_g = \kappa_1\kappa_2 \tag{19}$$

$$\kappa_m = (\kappa_1 + \kappa_2)/2 \tag{20}$$

图 4 展示了液体体积百分数等值面表面的几何形状[11]。可以看出,所有结构都落在 $\kappa_g = \kappa_m^2$ 曲线的下方,曲线上方的情况没有物理意义。若平均曲率大于 0,则界面呈现出凸界面的形式;反之若平均曲率小于 0,则为凹界面。若高斯曲率大于 0,主要呈现出抛物面结构;若高斯曲率小于 0,主要呈现双曲面结构。而在高斯曲率和平均曲率都接近于 0 的情况下,界面呈现出片状的平面结构;而在平均曲率等于 0、高斯曲率小于 0 的部分,则主要呈现出管状结构。

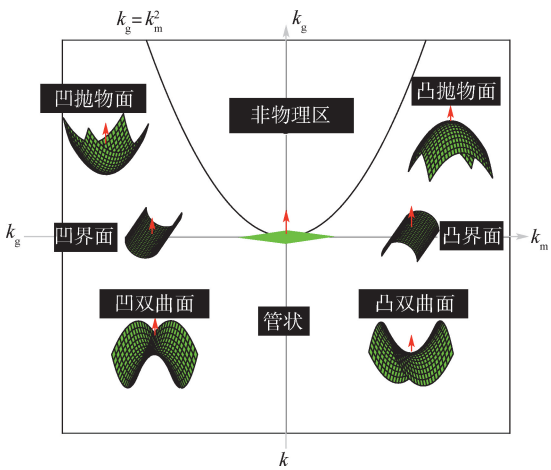
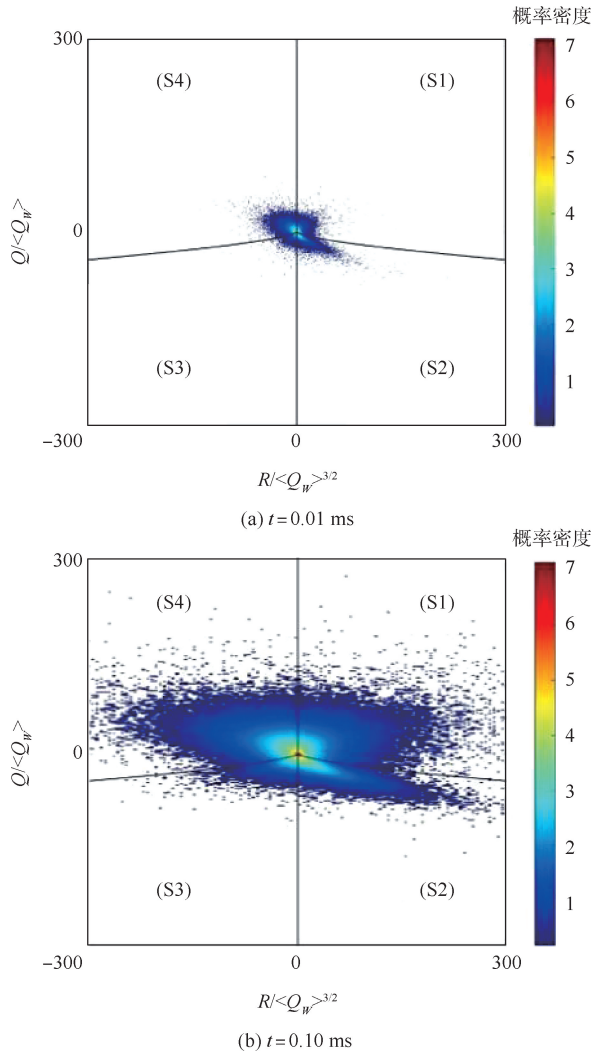


图 4 使用平均曲率和高斯曲率对等值面结构进行分类[11]

Fig. 4 Classification of structure of equivalence surface using mean curvature and Gaussian curvature[11]

2 结果与讨论

本节讨论液体射流雾化过程中流动拓扑结构演化以及流动拓扑结构对气液界面曲率的影响。首先,图 5 展示了第 2 不变量 Q 和第 3 不变量 R 的联合概率密度分布图(JPDF)。为方便对比,Q 使用 Q_W 进行无量纲化($Q/<Q_W>$),同理 R 使用 Q_W 的 3/2 次方进行无量纲化($R/<Q_W>^{3/2}$)。选取 $t = 0.01$ ms 和 $t = 0.10$ ms 的结果进行展示。如图 5所示,2 个时刻最大的概率密度分布均出现



(a) $t = 0.01$ ms

(b) $t = 0.10$ ms

图 5 第 2、第 3 张量不变量的联合概率密度分布

Fig. 5 Joint probability density function of the second and the third tensor invariance

在原点附近,因此将图 5 中的结果原点处在图 6 中进行放大展示。可以看出,在液体射流初始时刻,S1 拓扑结构出现的概率略微不同,S1 概率稍大,而随着时间推移,到了初次雾化的时刻,S1 拓扑结构出现的概率较大,表明流动主要受 UFC 拓扑结构的影响,即在雾化过程中,流体微团主要受旋转压缩作用机制的影响;而且随着雾化过程的进行,这种机制的重要性越发明显。值得注意的是,不同时刻的图像与之前研究中观察到的水滴形状比较相似,且在雾化后,区域面积变化较小,呈现出一定的自相似性质。

为了进一步研究流动拓扑对不同液体体积分数等值面包括气液界面的影响,本文引入条件应变率来对结果进行分析,定义如下:

$$a = \boldsymbol{n}\cdot(\nabla\boldsymbol{u})\cdot\boldsymbol{n} \tag{21}$$

式中:a 为等值面切向应变率,该值的正与负分别对应于拉伸应变率和压缩应变率。

图 7 展示了射流雾化过程中应变率在 Q-R

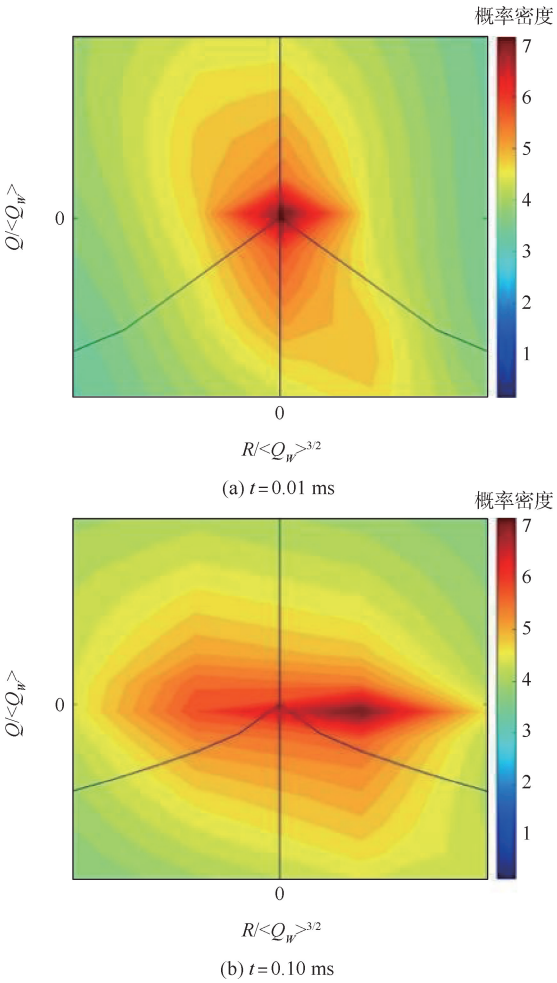


图 6　图 5 在原点附近的放大图

Fig. 6　Fig. 5 enlarged view near origin point

空间的条件平均分布。如图 7 所示,在 UFC(S1)拓扑分区,主要引起拉伸应变率(应变大于 0)。而在 SFS(S4)拓扑分区,主要引起压缩应变率(应变小于 0)。而在 UN/S/S(S2)、SN/S/S(S3)区域,则由应变主导雾化。另外,图 7 表明,初始时刻,应变能主要是拉伸应变率主导。而随着雾化进行,压缩应变率与拉伸应变率效应共存。其中,UFC(S1)流动拓扑区域产生拉伸应变率的概率最高。

图 8 展示了同一时刻 t = 0.06 ms 下,不同液体体积分数(α = 0.2,0.5,0.8)等值面上应变率的差异。如图 8 所示,对于 UFS 拓扑结构,当气相占比较大时(α = 0.2),具有较高的拉伸应变率,而液相占比较大(α = 0.8),拉伸应变率较小。对于其他流动拓扑结构,均为产生压缩应变率的概率较高。其中 SFS 结构,压缩应变率较小,而在变形主导的区域(S2,S3),压缩应变率较大。

选取同一时刻 t = 0.06 ms 下不同的液体体积分数等值面(α = 0.2,0.5,0.8)曲率进行分析,通过 κ_g-κ_m 联合概率密度分布图像探讨不同界面

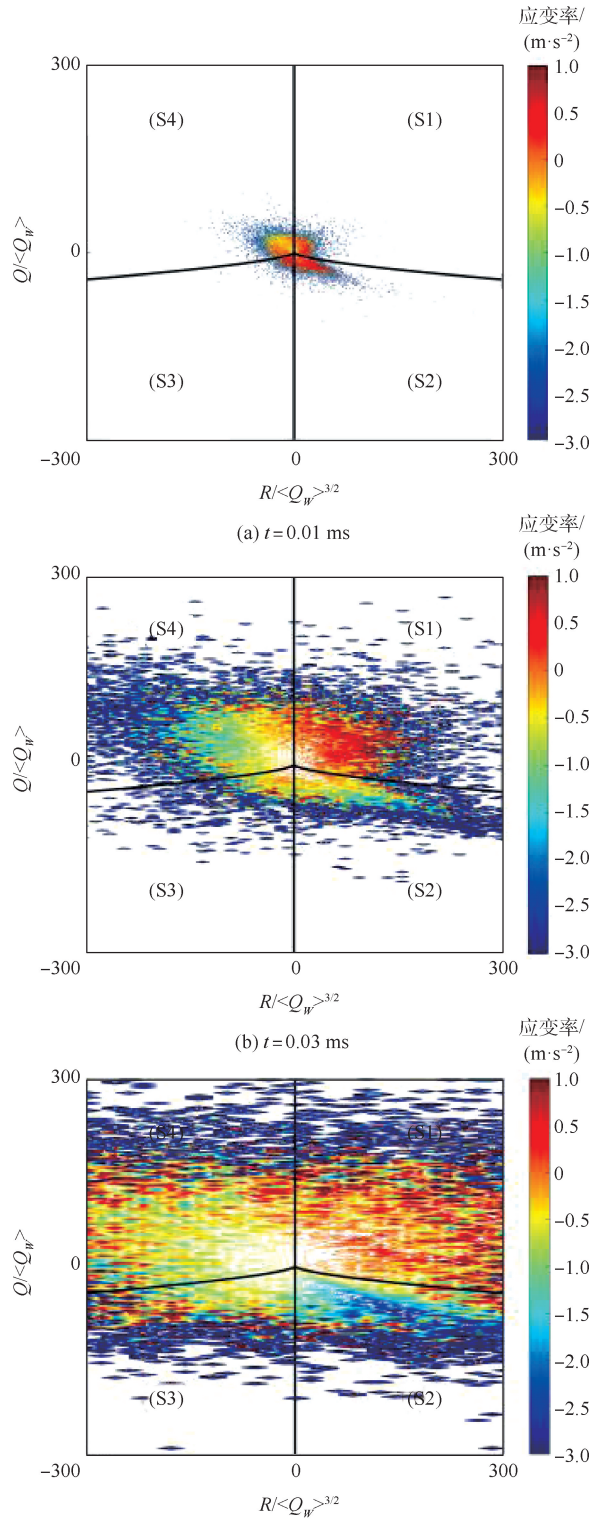


图 7　不同时刻下应变率在 Q-R 平面内的条件平均分布

Fig. 7　Conditional average distribution of strain rate in Q-R plane in different moments

的几何结构。如图 9 所示,等值面呈现出不同的结构。在气液界面处(α = 0.5 处),概率密度分布较为对称,最高概率密度出现在原点附近,说明主要呈现出片状的结构。而在 α = 0.2 时,气相占比较大,平均曲率大于 0 概率较大,这种情况下主要呈现出凸形表面结构。在 α = 0.8 时,液相占

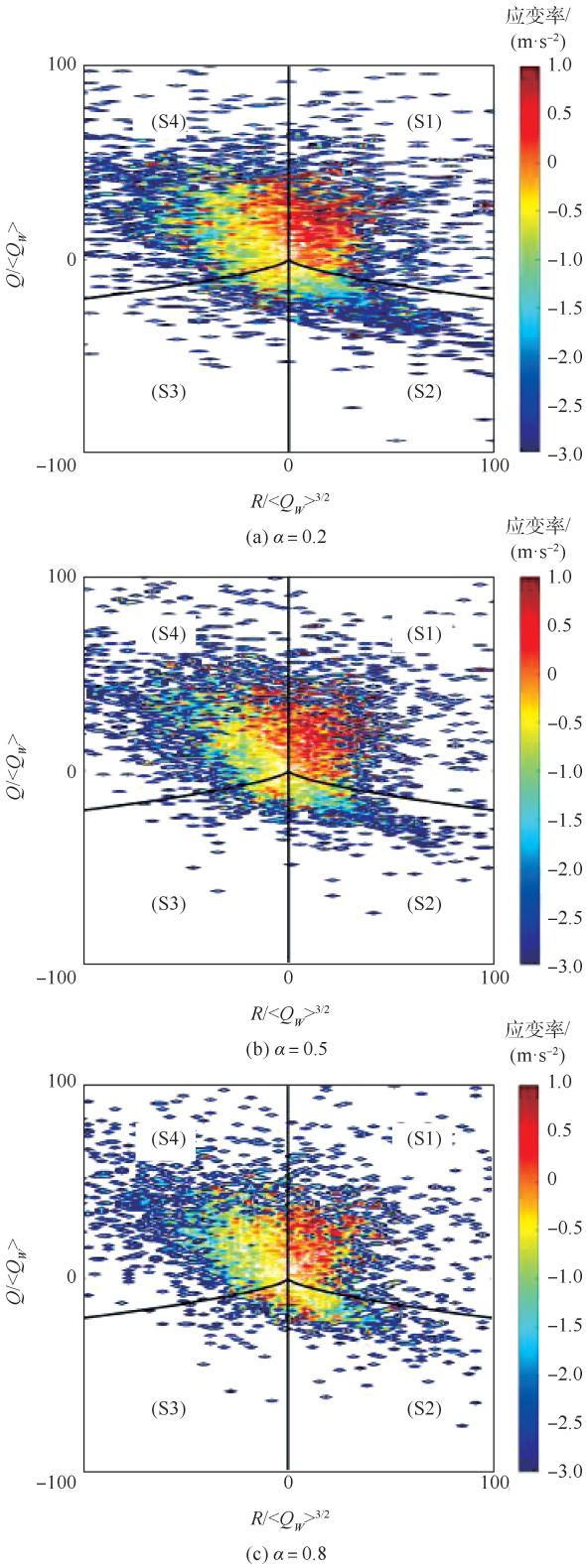


图8 不同等值面应变率在 Q-R 平面内的条件平均分布

Fig. 8 Conditional average distribution of strain rate in Q-R plane on different iso-surfaces

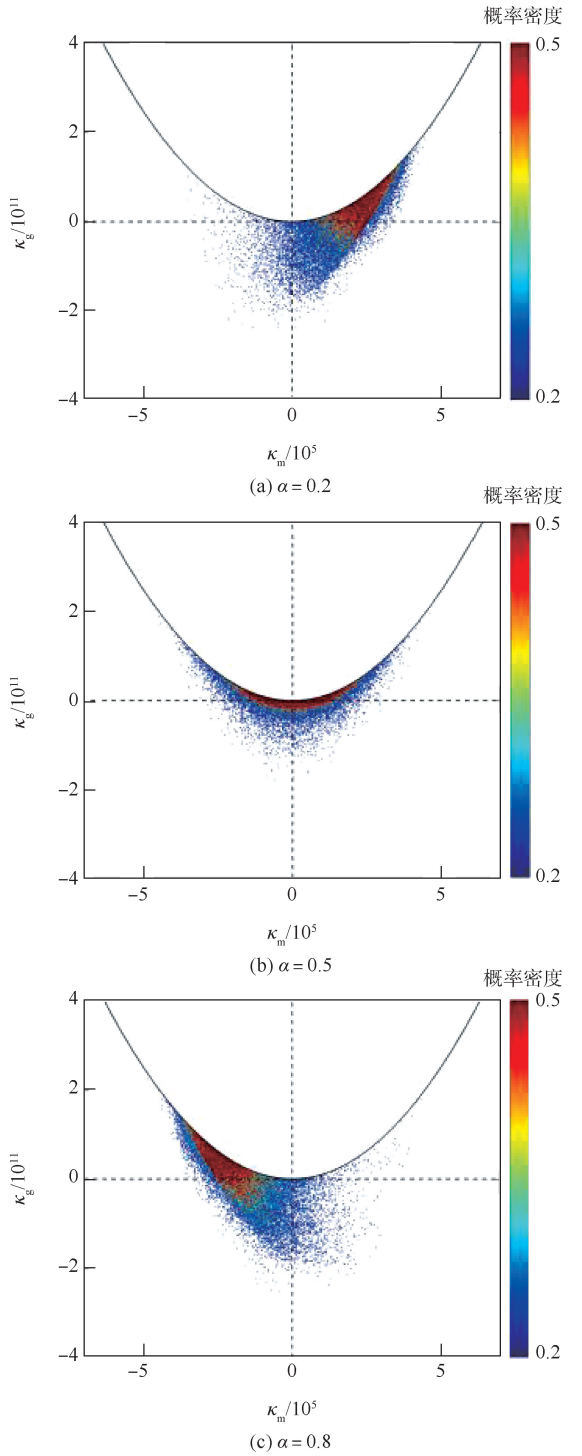


图9 不同等值面上的 κ_g-κ_m 联合概率密度分布

Fig. 9 Joint probability density distribution of κ_g-κ_m on different equivalence sufales

比较大,平均曲率小于0概率较大,主要呈现出凹形表面结构,这些表明,局部界面几何结构会受到局部拓扑结构影响很大。

取不同时刻研究所有液体体积百分数等值面上的 κ_g-κ_m 联合概率密度分布,以此探讨雾化过程的不同时刻整个射流流场的曲率分布,从而获得整个流场界面的大致形状。除了选取初次雾化之前的 $t=0.01$ ms,$t=0.03$ ms 和 $t=0.06$ ms,还选取了雾化发生后的时间 $t=0.10$ ms,$t=0.15$ ms 和 $t=0.20$ ms。如图10(a)~(f)所示,雾化初始时刻主要存在的是凸面、凹面、片状结构,其中片状结构概率最高。随着雾化进行,出现了管状的等值面结构,同时平均曲率及高斯曲率的绝对值

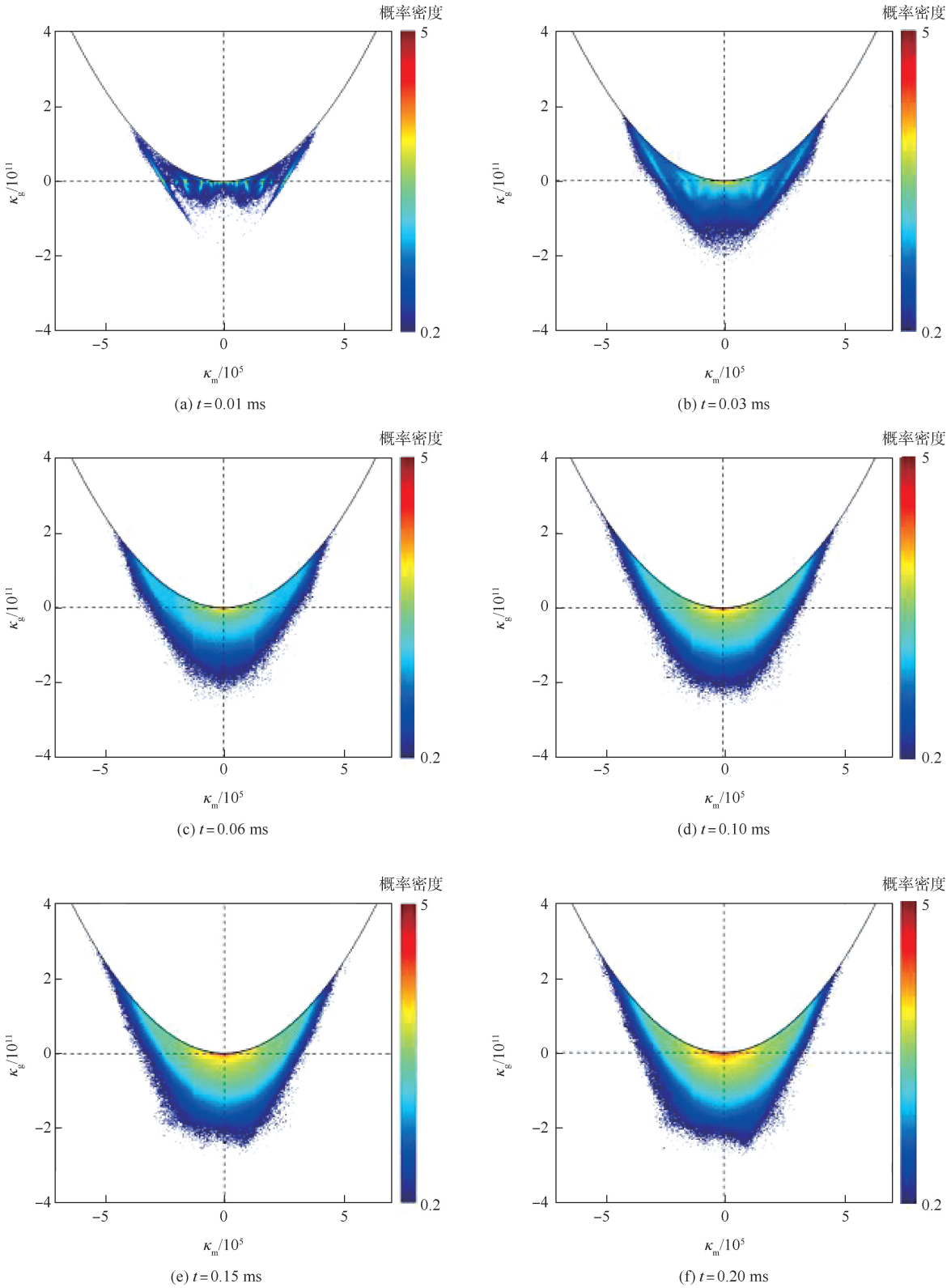
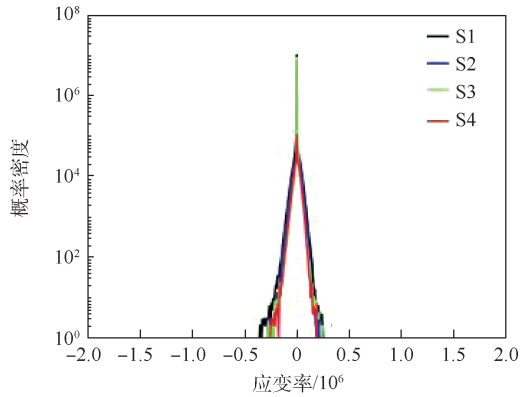


图 10　不同时刻下整个流场的 κ_g-κ_m 联合概率密度分布

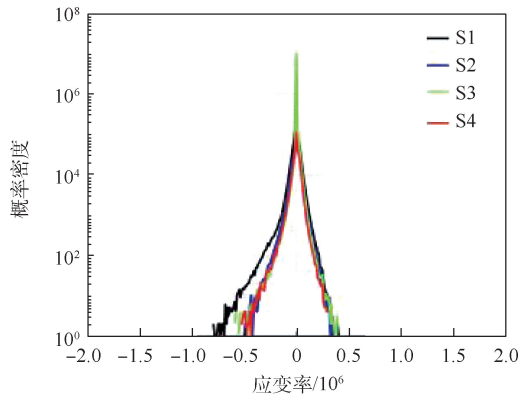
Fig. 10　Joint probability density distribution of κ_g-κ_m for whole flow field at different moments

不断增加,这意味着随着雾化进程,也会出现凹双曲面、凹抛物面、凸双曲面、凸抛物面结构,但在每个时刻,仍然是出现片状结构的概率最高。而在初次雾化后,逐渐呈现出图像不关于 $\kappa_m=0$ 轴对称的现象,产生凹双曲面的概率略微大于产生凸双曲面的概率。由图 9 可知,液相占比较大时产生凹面的概率较大,可能是由于雾化后期液滴弥散至整个空间所致。

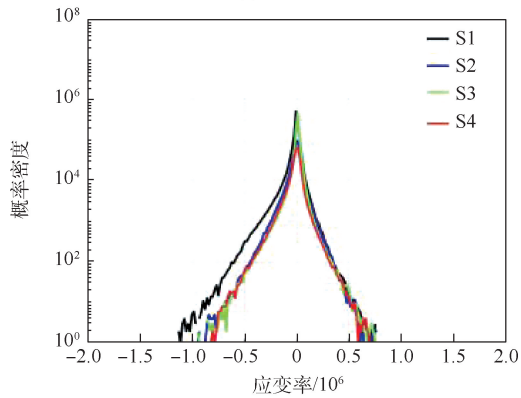
图 11 展示了应变率在不同拓扑区域内的概率分布图(PDF)。可以看出,所有流动拓扑结构都有助于在不同区域内形成压缩应变率和拉伸应变率。但在每个区域内,产生零应变的概率均为



(a) $t=0.01$ ms



(b) $t=0.03$ ms



(c) $t=0.06$ ms

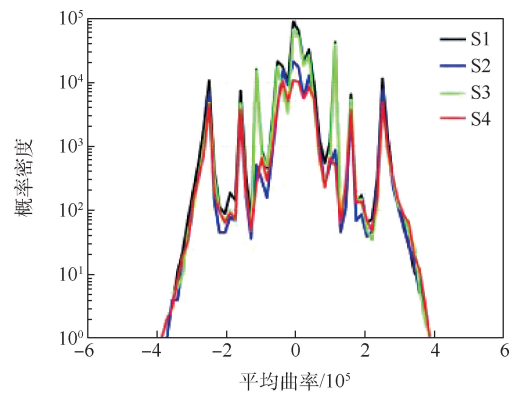
图 11 不同拓扑区域内的应变率概率密度分布

Fig. 11 Probability density distribution of strain rate in different topological regions

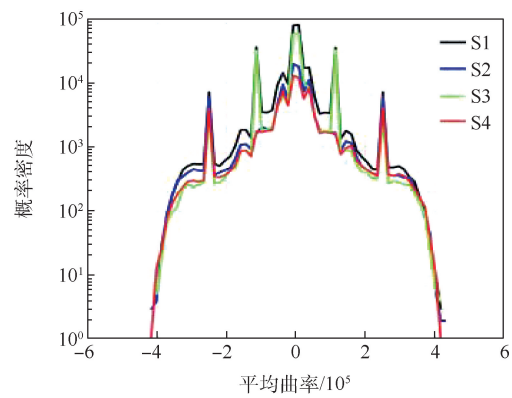
最高。而随着雾化进行,每个拓扑区域产生的压缩应变率都会大于其产生的拉伸应变率,即偏向于负的应变率。由图 11 可以看出,S1 和 S3 的概率密度显著大于 S2 和 S4 的概率密度。随着雾化进行,S1 拓扑会产生更大的压缩应变率。另外,S1 和 S3 对拉伸与压缩应变率的贡献最高,由于 S1 拓扑比 S3 拓扑概率更大,S1 概率增大,应变率变大,曲率变小。

图 12 展示了平均曲率在不同拓扑区域内的概率分布。可以看出,所有流动拓扑结构都会产生正平均曲率和负平均曲率,每个区域内,产生零曲率的概率最大,表明产生片状结构的概率最高,这与之前的结论一致。同时,曲率的概率分布大致关于 $\kappa_m=0$ 对称,表明在雾化过程中,产生凸界面结构与凹界面结构的概率大致相等。

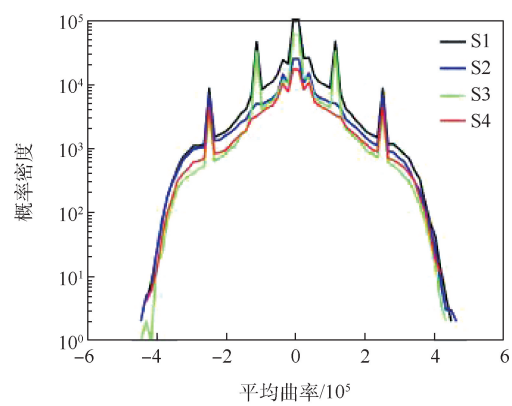
图 13 展示了 $t=0.10$ ms 时刻应变率和平均曲率的联合概率密度函数分布。图 13 表明,随着应变增加,等值面界面趋向于片状,管状结构,在压缩应变率区域(应变小于 0),曲率下降的速度比拉伸区域(应变大于 0)快,在压缩应变率区域,应变和曲率的负相关关系更为明显。这里负相关关系指一个变量随着另一个变量的减少而增加的关系。



(a) $t=0.01$ ms



(b) $t=0.03$ ms



(c) $t=0.06$ ms

图 12 不同拓扑区域内的平均曲率概率密度分布

Fig. 12 Mean curvature probability density distribution in different topological regions

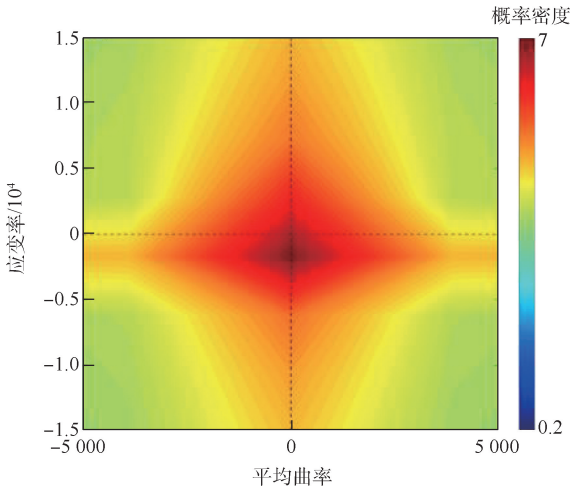


图 13　整个流场中应变率与平均曲率的联合概率密度函数分布

Fig. 13　Joint probability density function distribution of strain and mean curvature over the entire flow field

基于上述结论,可以总结出如下因果链,其中,上下箭头代表该量绝对值的增加或者减少,水平箭头代表因果链:$PDF_{(UFC)}$ ↑→应变率↑→曲率↓→雾化性能↓。

3　结　论

本文对湍流液体平面射流进行了三维直接数值模拟。通过流动拓扑理论,分析了射流雾化过程中流动拓扑的影响机制。通过分析气液界面的曲率,揭示了流动拓扑与界面结构之间的关系。从流动拓扑的角度阐述了射流雾化气液界面不断演化的机制。研究结果表明:

1)所有流动拓扑结构都有助于产生压缩应变率和拉伸应变率。射流雾化过程主要受到 UFC 拓扑结构的影响,从而产生拉伸应变率,且 UFC 的增多会导致更多的等值面界面压缩应变率,进而导致气液界面曲率下降,降低雾化性能。

2)随着雾化进行,气相区域主要产生凸界面,液相区域主要产生凹界面,气液界面处界面主要为片状结构。除 UFC 外,其余拓扑结构主要产生压缩应变率,在 SFS 区域产生的压缩应变率较小,而在 UN/S/S 区域和 SN/S/S 区域,产生的压缩应变率较大。

3)随着雾化进行,产生压缩应变率的概率会略微大于拉伸应变率概率。

根据上述结论,总结出如下的因果链:$PDF_{(UFC)}$ ↑→应变率↑→曲率↓→雾化性能↓。该因果链说明,随着 UFC 拓扑结构的增加,会使得拉伸应变率增大,界面受到拉伸,曲率则会下降,不利于雾化过程的进行,因此雾化性能下降率。为了验证以上因果链的普适性,本文作者团队正在进行多参数变化的直接数值模拟研究,旨在全面地揭示不同雷诺数与韦伯数下流动拓扑对液体射流雾化的影响机制。

参考文献(References)

[1] DESJARDINS O, MCCASLIN J, OWKES M, et al. Direct numerical and large-eddy simulation of primary atomization in complex geometries [J]. Atomization and Sprays, 2013, 23(11):1001-1048.

[2] SUSSMAN M, SMEREKA P, OSHER S. A level set approach for computing solutions to incompressible two-phase flow [J]. Journal of Computational Physics, 1994, 114 (1):146-159.

[3] HIRT C W, NICHOLS B D. Volume of fluid (VOF) method for the dynamics of free boundaries [J]. Journal of Computational Physics, 1981, 39 (1):201-225.

[4] LEBOISSETIER A, ZALESKI S. Direct numerical simulation of the atomization of liquid jet [C] // Proceedings of the ILASS-Europe, 2001:2-6.

[5] KLEIN M. Directed numerical simulation of a spatially developing water sheet at moderate Reynolds number [J]. International Journal of Heat and Fluid Flow, 2005, 26(5):722-731.

[6] SANDER W, WEIGAND B. Direct numerical simulation and analysis of instability enhancing parameters in liquid sheets at moderate Reynolds numbers [J]. Physics of Fluids, 2008, 20(5):053301.

[7] MÉNARD T, TANGUY S, BERLEMONT A. Coupling level set/VOF/ghost fluid methods: Validation and application to 3D simulation of the primary break-up of a liquid jet [J]. International Journal of Multiphase Flow, 2007, 33(5):510-524.

[8] DESJARDINS O, PITSCH H. Detailed numerical investigation of turbulent atomization of liquid jets [J]. Atomization and Sprays, 2010, 20(4):311-336.

[9] SHINJO J, UMEMURA A. Simulation of liquid jet primary breakup: Dynamics of ligament and droplet formation [J]. International Journal of Multiphase Flow, 2010, 36(7):513-532.

[10] PERRY A E, CHONG M S. A description of eddying motions and flow patterns using critical-point concepts [J]. Annual Review of Fluid Mechanics, 1987, 19(1):125-155.

[11] CHONG M S, PERRY A E, CANTWELL B J. A general classification of three-dimensional flow fields [J]. Physics of Fluids, 1990, 2(5):765-777.

[12] CHONG M S, SORIA J, PERRY A E, et al. Turbulence structures of wall-bounded shear flows found using DNS data [J]. Journal of Fluid Mechanics, 1998, 357:225-247.

[13] ELSINGA G E, MARUSIC I. Universal aspects of small-scale motions in turbulence [J]. Journal of Fluid Mechanics, 2010, 662:514-539.

[14] HASSLBERGER J, KLEIN M, CHAKRABORTY N. Local flow topology analysis applied to bubble-induced turbulence [C] // 12th International ERCOFTAC Symposium on Engineering Turbulence Modelling and Measurements. Montpellier, 2018:270-290.

[15] WATANABE T, SAKAI Y, NAGATA K, et al. Vortex stretching and compression near the turbulent/non-turbulent interface in a planar jet[J]. Journal of Fluid Mechanics, 2014, 758: 754-785.

[16] JOSEF H, SEBASTIAN K, MARKUS K, et al. Flow topologies in primary atomization of liquid jets: A direct numerical simulation analysis [J]. Journal of Fluid Mechanics, 2019, 859: 819-838.

[17] HAN W, ARNE S, FELIX D, et al. Influence of flow topology and scalar structure on flame-tangential diffusion in turbulent non-premixed combustion [J]. Combustion and Flame, 2019, 206: 21-36.

[18] DESJARDINS O, BLANQUART G, BALARAC G, et al. High order conservative finite difference scheme for variable density low Mach number turbulent flows[J]. Journal of Computational Physics, 2008, 227(15): 7125-7159

[19] DESJARDINS O, MOUREAU V, PITSCH H. An accurate conservative level set/ghost fluid method for simulating turbulent atomization[J]. Journal of Computational Physics, 2008, 227(18): 8395-8416.

Influence of flow topology on instability and atomization of liquid jets

YANG Lijun[1,2], HUANG Dongqi[1], HAN Wang[1,2], LI Jingxuan[1,2], FU Qingfei[1,2,*]

(1. School of Astronautics, Beihang University, Beijing 100083, China;
2. Ningbo Institute of Technology, Beihang University, Ningbo 315800, China)

Abstract: In the combustion process of liquid fuels, jet instabilities and atomization are the starting points, which can have significant effects on following processes such as evaporation and combustion. There remain gaps in our under standing of the turbulent jet atomization mechanism despite the extensive prior research. This work aims to reveal the atomization mechanism of turbulent liquid planar jets through the use of flow topology. A high-resolution direct numerical simulation method is used to solve the atomization process of a liquid plane jet in a still air environment. In order to clarify the process by which flow topology affects the atomization of liquid plane jets, the interaction between various topologies in the flow field and the curvature of the gas-liquid interface is analyzed. It is found that all flow topologies contribute to the generation of compressive and extensive strain rate, among which the UFC topology has the greatest effect on the flow field strain rate; the curvature of the liquid volume fraction iso-surface shows a negative correlation with the strain rate under the influence of the flow topology. In addition, the UFC structure mainly generates extensive strain, while the remaining flow topologies mainly generate compressive strain. These results suggest that the jet atomization process is mainly influenced by the UFC topology, which facilitates a large extensive strain at the gas-liquid interface, which in turn promotes the generation of sheet or saddle structures, thus causing liquid jet fragmentation.

Keywords: jet atomization; plane jet; flow topology; curvature; strain rate

Received: 2022-07-11; **Accepted**: 2022-08-14; **Published online**: 2022-08-19 16:16

URL: kns.cnki.net/kcms/detail/11.2625.V.20220819.1337.002.html

Foundation items: National Natural Science Foundation of China (11922201, 11872091, 11927802, U1837211)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0301

基于电磁超声横波的管道剩余厚度检测

徐立军[1], 刘福禄[1], 丁一清[2], 李正勇[3], 谢跃东[1,*]

(1. 北京航空航天大学 仪器科学与光电工程学院, 北京 100083; 2. 北京航空航天大学 沈元学院, 北京 100083;
3. 上海航天动力技术研究所, 上海 200000)

摘 要: 为实现管道剩余厚度的精确检测,设计了一种基于电磁超声横波的非接触式管道厚度检测系统。采用自研的电磁超声大功率激励源与换能器产生测厚横波,并由接收器对回波电压信号进行实时滤波和处理,得到铝制管道的精确剩余厚度。针对电磁超声回波的小信号和低信噪比对激励线圈参数进行优化,在此基础上对横波声束在圆形管道界面辐射特性进行分析。根据换能器线圈匝数、线圈宽度分别为回波信号峰峰值和信噪比的最大影响因子,设计优化后检测换能器并实现了误差小于0.2%的剩余厚度检测。

关 键 词: 管道剩余厚度; 电磁超声换能器; 换能器; 声束辐射; 横波
中图分类号: TB51+7
文献标志码: A **文章编号**: 1001-5965(2022)09-1767-07

管道作为石油、天然气、燃气等液体和气体的主要输送方式,在人们的日常生活和工业生产中承载着关键性工作,其服役状态直接影响着能源运输和生产生活的安全、稳定。随着服役时间的增加,在役管道不可避免地会出现裂纹、腐蚀等缺陷,严重的缺陷会直接导致管道泄露等事故,进而造成重大的经济损失。管道剩余壁厚的及时检测可以避免因腐蚀造成管道泄漏事故,同时管道剩余壁厚是管道可靠性的重要指标之一。

在实际的管道检测过程中,管道的腐蚀多发生在外表面,且具有腐蚀残留、工作环境恶劣等特点。电磁超声换能器(electromagnetic acoustic transducer,EMAT)具有非接触、无需耦合剂、对被检测表面光洁度无严格要求并能够产生多种声波模态等特性,适合用于管道剩余壁厚的在线检测。电磁超声相较于压电超声测厚具有更好的环境适用性,无需对被测管道进行表面清洁处理,避免因高低温等恶劣情况造成耦合剂失效导致压电超声无法使用,同时电磁超声移动更加灵活。EMAT测厚是利用洛伦兹力机制、磁致伸缩机制和磁化力机制在被检测对象内部激发并接收超声波信号,根据声波波速和传播时间计算出被检测部件的厚度。电磁超声系统中换能器线圈和磁铁的形状、大小与位置设置灵活,可以激发多种类型的超声波,但由于在金属管道中横波的波速约为纵波的一半,且波模态稳定,成为最适合作为回波法测量剩余厚度的声波类型。利用电磁超声测厚,国内外研究人员开展了很多研究工作。Asano 和 Yoshida[1]应用电磁超声共振技术,设计了一种测量钢板表面氧化膜厚度的电磁超声换能器;Hobbis和 Aruleswaran[2]使用英国华威大学开发的宽带 EMAT 系统对厚度范围在 0.28 ~ 2.8 mm 之间的铝合金样品进行了静态测量,实验的标准偏差小于等于 0.08 μm;Parra-Raad 等[3]设计了一款新型 EMAT 结构,通过 2 个彼此正交的线圈在金属材料中激发出 2 个正交极化的剪切波,这使

收稿日期: 2022-04-29; **录用日期**: 2022-05-18; **网络出版时间**: 2022-06-07 13:09
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220606.1801.004.html
基金项目: 国家自然科学基金(61901022); 上海航天科技创新基金(118SAST2021117001); 中央高校基本科研业务费(YWF-19-BJ-J-358)
* **通信作者**. E-mail: yuedongxie@buaa.edu.cn

引用格式: *徐立军, 刘福禄, 丁一清, 等. 基于电磁超声横波的管道剩余厚度检测[J]. 北京航空航天大学学报, 2022, 48(9): 1767-1773. XU L J, LIU F L, DING Y Q, et al. Residual thickness detection of pipeline based on electromagnetic ultrasonic shear wave [J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1767-1773(in Chinese).*

EMAT 同时进行厚度测量和裂纹检测成为了可能。国内许多高校均开展了电磁超声测厚方面的研究,哈尔滨工程大学的王相豪[4]研制了快速精确测量的 EMAT 测厚系统,实现了对标准钢块的厚度测量,测量精度为 0.5 mm;浙江大学的唐志峰等[5]设计了一款具有定点超声测厚功能的复合式电磁超声设备,实现了钢管测厚误差小于 0.1 mm;哈尔滨工业大学的孙峥等[6]研制了一种用于管道内检测的 EMAT 在线测厚装置,实现了 8 ~ 35 mm 范围内的在役管道厚度测量。目前,针对管道电磁超声测厚装置和声波在管道壁内声束辐射的研究较少。

本文主要研究基于电磁超声横波模态的管道壁厚测量方法及其换能器激励线圈优化设计,并利用自研的电磁超声系统实现了对铝制管道剩余厚度的精确测量。利用有限元模拟仿真的方法分析了螺旋线圈参数对横波在铝制管道传播的影响,以回波信号信噪比为主要参考因素制作电磁超声换能器。根据管道壁内横波声束辐射指向性验证了系统可行性,实现了铝制管道剩余壁厚误差小于 0.2% 的精确测量。

1 电磁超声横波测厚原理

EMAT 一般由永磁体和激励线圈组成。永磁体提供偏置磁场,激励线圈搭载高频电流从而在试件表面激发高频涡流场。通过不同方向的偏置磁场和不同形状的线圈组合,能够产生体波、表面波、导波等不同类型的声波[7]。在试件厚度测量过程中,横波在结构件边界处不存在模式转换且波速较慢,因此成为广泛采用的声波类型。典型的横波电磁超声换能器如图 1 所示,横波电磁超声换能器由永磁体和螺旋形激励线圈组成。永磁体产生方向垂直于试件表面的偏置磁场,

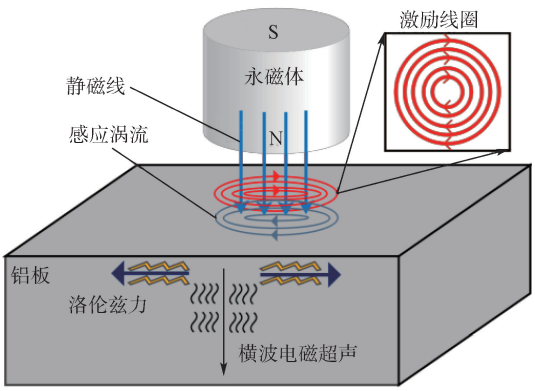


图 1 典型的横波电磁超声换能器示意图

Fig. 1 Schematic diagram of a typical shear wave electromagnetic acoustic transducer

流过螺旋形激励线圈的高频电流在试件表面感应出涡流场,在垂直偏置磁场作用下产生洛伦兹力,使得试件表面晶粒形变带动相邻质点振动成为超声波波源。电磁超声的接收过程是电磁超声激励的逆过程。对于非铁磁性的铝制材料,在垂直偏置磁场作用下的横波电磁超声主要是由洛伦兹力产生的,磁致伸缩力和磁性力可以忽略不计[8]。

电磁超声的换能过程涉及到静态磁场、脉冲涡流场和固体力学场等多种物理场的耦合。永磁体产生的偏置磁场为静态磁场,没有电流的产生,即 $\boldsymbol{J}=\boldsymbol{0}$,为

$$\nabla\times\boldsymbol{H}=\boldsymbol{0} \tag{1}$$

式中:$\boldsymbol{H}$ 为磁场强度,可以用标量函数的梯度进行表示,即

$$\boldsymbol{H}=-\nabla\varphi_{\mathrm{m}} \tag{2}$$

其中:φ_{m} 为标量磁位。

永磁体在试件中产生的偏置磁感应强度 $\boldsymbol{B}$ 为

$$\boldsymbol{B}=\mu\boldsymbol{H}+\boldsymbol{B}_{\mathrm{r}} \tag{3}$$

式中:μ 为被测试件磁导率;$\boldsymbol{B}_{\mathrm{r}}$ 为永磁体剩余磁感应强度。

忽略位移电流的影响,激励线圈中交变电流在试件中产生的磁感应强度 $\boldsymbol{B}_{\mathrm{j,m}}$ 为

$$\nabla\times\boldsymbol{E}_{\mathrm{E}}=-\frac{\partial\boldsymbol{B}_{\mathrm{j,m}}}{\partial t} \tag{4}$$

式中:$\boldsymbol{E}_{\mathrm{E}}$ 为涡流的电场强度。

同时,激励线圈中高频电流在试件表面感应出涡流密度 $\boldsymbol{J}_{\mathrm{E}}$ 为

$$\boldsymbol{J}_{\mathrm{E}}=\sigma_{\mathrm{m}}\boldsymbol{E}_{\mathrm{E}} \tag{5}$$

式中:σ_{m} 为被测试件的电导率。

在偏置磁场和激励线圈作用下,被测试件中产生的洛伦兹力 $\boldsymbol{F}_{\mathrm{L}}$ 为

$$\boldsymbol{F}_{\mathrm{L}}=\boldsymbol{J}_{\mathrm{E}}\times(\boldsymbol{B}+\boldsymbol{B}_{\mathrm{j,m}}) \tag{6}$$

在非铁磁性的铝材料中,洛伦兹力 $\boldsymbol{F}_{\mathrm{L}}$ 是由静态磁场 $\boldsymbol{B}$ 和被测试件内部涡流感生的交变磁场 $\boldsymbol{B}_{\mathrm{j,m}}$ 共同作用产生的。电磁超声接收过程是其产生的逆过程:声波在传播过程中遇到声阻抗不同的地方会发生反射,当声波经过反射再次回到试件上表面时,试件中带电粒子的运动在外部静磁场作用下会形成动态电流[9]。回波涡流密度为

$$\boldsymbol{J}'_{\mathrm{E}}=\sigma_{\mathrm{m}}(\boldsymbol{\nu}\times\boldsymbol{B}) \tag{7}$$

式中:$\boldsymbol{J}'_{\mathrm{E}}$ 为试件中涡流密度;$\boldsymbol{\nu}$ 为与磁场 $\boldsymbol{B}$ 相互作用的粒子速度。

动态电流在导体周围产生交变磁场,该磁场被换能器中螺旋线圈接收,产生电压回波信号。横波电磁超声通过测量超声始波和回波或 2 次回波之间的时间差确定试件的厚度。回波传播路程

为试件厚度的 2 倍,因此试件厚度 h 为

$$h = v_m \frac{\Delta t}{2} \tag{8}$$

式中:v_m 为试件中声速;Δt 为回波时间间隔。

2　换能器优化与测试

2.1　电磁超声换能器的仿真与优化

横波电磁超声测量管道剩余壁厚虽然具有精度高、非接触、不易受测试环境影响等优点,但是其能量交换效率低下、接收信号微弱[10-11]等缺点不能忽视。通过对超声换能器参数优化可以在一定程度上改善电磁超声测量的不足,使之满足剩余厚度测量系统的检测要求。利用有限元仿真软件对测量系统进行建模并利用正交化试验思想优化仿真参数设计。对换能器激励线圈参数进行优化可以提高系统性能,提高横波回波电压信号的峰峰值和信噪比。超声换能器参数优化中,国内外的许多学者已经进行了广泛的研究得到改变线圈提离距离 d 对回波信号幅值的影响,即在系统可制作范围内,应尽量减小线圈提离距离[12-13];同时,换能器中永磁体的尺寸与线圈的尺寸不存在严格固定比例关系,可以根据线圈做调整[9,14]。因此,本文系统换能器优化过程中充分考虑工艺情况和应用场景,将线圈参数中的提离距离、永磁体尺寸和永磁体与线圈的间距设为固定值,不予讨论研究。对于线圈中的其他参数,如线圈匝数(c_n)、线圈宽度(ω)和线圈间距(l)作为主要的优化参数。

根据现实参照和使用环境,将 EMAT 优化参数设置如表 1 所示。其他参数根据实际制作情况分别设置:线圈提离高度为 0.112 5 mm(PCB 印刷电路中阻焊层厚度为 0.012 5 mm,绝缘胶带厚度为 0.1 mm);线圈厚度采用 1 盎司铜箔的 PCB 印刷电路标准为 0.035 mm(1 盎司 =28.350 g);线圈与永磁体的距离为 0.9 mm(PCB 印刷电路板材厚度为 0.8 mm,绝缘胶带厚度为 0.1 mm);永磁体采用半径为 20 mm、高为 20 mm 的圆柱形磁铁,材料为 NdFeB(N54)。

有限元仿真的几何模型如图 2 所示。为节省

表 1　电磁超声换能器参数优化范围

Table 1　Electromagnetic acoustic tranolucer parameter optimization range

线圈匝数	线圈宽度/mm	线圈间距/mm
4	0.2	0.8
6	0.4	1.0
8	0.6	1.2

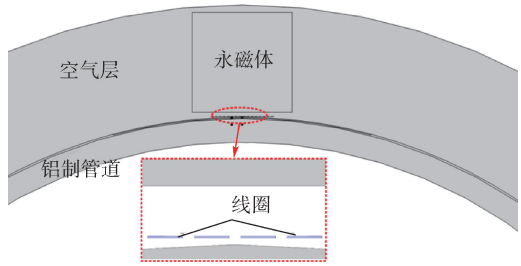


图 2　横波电磁超声有限元仿真模型

Fig. 2　Shear wave electromagnetic acoustic finite element simulation model

计算资源,将完整铝制管道截取 90°,即周向四分之一管道作为检测剩余厚度的模型,管道外直径为 200 mm,厚度为 5 mm。

根据需要优化的换能器参数设计了一个正交阵列($L_9(3^4)$),如表 2 所示,利用有限元仿真软件对正交测试阵列进行仿真,并记录仿真中回波电压信号中的峰峰值和信噪比。为了表征离散化的程度并评估各个因子对测试结果的影响程度,需要对正交试验的结果进行极差分析,其计算的过程为[15]

$$K_{xN} = \frac{1}{m}\sum_{i=1}^{n} y_{xi} \tag{9}$$

$$R_x = \max(K_{xn}) - \min(K_{xn}) \tag{10}$$

式中:y_{xi}为含有参数因子 x 的第 i 次测试结果,x 因子为需要优化的参数;m 为参数因子 x 在水平 N 的测试次数,$N=1,2,3$;n 为测试总数;K_{xN}为该水平下参数因子 x 试验结果的平均水平;R_x 为参数因子 x 对实验结果的影响程度。

根据有限元法仿真结果得到平均值 K_{xN}和影响程度 R_x 的结果,如表 3 所示。可得到,对于回波电压信号的峰峰值有 $R_{cn} > R_l > R_\omega$,而对于信噪比存在 $R_\omega > R_{cn} \approx R_l$。随着线圈匝数的增加回波信号峰峰值逐渐减小,信噪比下降;线圈宽度对峰峰值的影响可以忽略,但信噪比随其增加而增加;

表 2　横波电磁超声换能器参数正交测试阵列

Table 2　Shear wave electromagnetic acoustic transducer parameter orthogonal test array

试验号	线圈匝数	线圈宽度/mm	线圈间距/mm	峰峰值/mV	信噪比/dB
1	4	0.2	1.0	0.285 8	14.218
2	6	0.2	0.8	0.243 9	15.095
3	8	0.2	1.2	0.144 4	9.768
4	4	0.4	0.8	0.316 7	15.559
5	6	0.4	1.2	0.193 1	18.129
6	8	0.4	1.0	0.172 0	15.657
7	4	0.6	1.2	0.243 9	15.458
8	6	0.6	1.0	0.213 5	16.682
9	8	0.6	0.8	0.197 7	18.603

表 3 正交试验结果分析

Table 3 Analysis of orthogonal experiment results

分析因子	对峰峰值影响	对信噪比影响
K_{cn1}	2.8×10^{-4}	15.08
K_{cn2}	2.2×10^{-4}	16.64
K_{cn3}	1.7×10^{-4}	14.68
R_{cn}	1.1×10^{-4}	1.96
$K_{\omega1}$	2.2×10^{-4}	13.03
$K_{\omega2}$	2.3×10^{-4}	16.45
$K_{\omega3}$	2.2×10^{-4}	16.91
R_{ω}	8.9×10^{-6}	3.89
K_{l1}	2.5×10^{-4}	16.42
K_{l2}	2.2×10^{-4}	15.52
K_{l3}	1.9×10^{-4}	14.45
R_{l}	5.9×10^{-5}	1.97

线圈间距与线圈匝数对回波峰峰值及信噪比的影响情况相同。

回波式横波电磁超声测厚系统中，螺旋形线圈接收到的回波电压仿真结果如图 3 所示，电压信号微弱，通常为几十至几百微伏量级，对于检测系统的接收是困难的。同时，在回波电压信号中通常夹杂着反射纵波和其他波模态的干扰及电路噪声的影响。

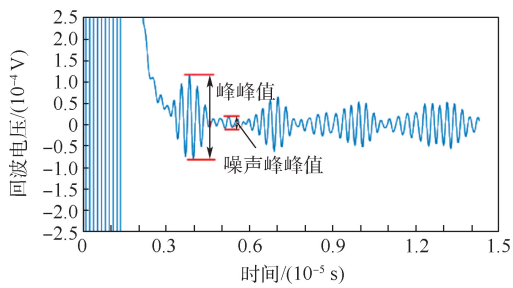
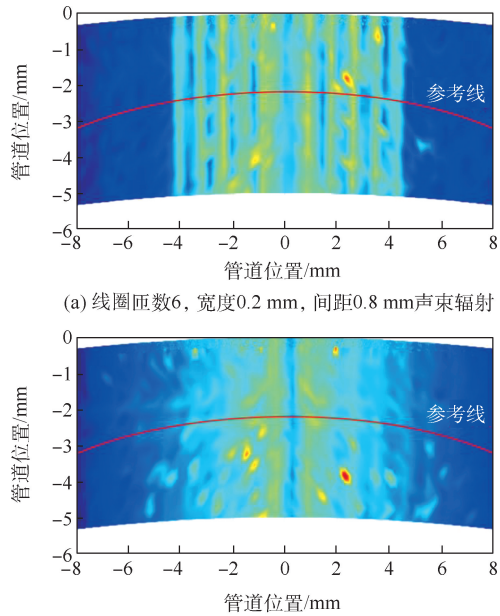


图 3 横波电磁超声回波电压信号

Fig. 3 Shear wave electromagnetic acoustic echo voltage signal

测量系统要求回波电压信号不仅具有较高的幅值还应有足够的信噪比来提高接收系统中反射横波电压信号的时间准确性。在仿真中可以看到回波电压信号幅值较小，信噪比较低，为保证接收处理系统对回波信号的时间处理精度，优先保证信噪比参数。同时考虑到较少的线圈匝数会激励出较少的涡流，造成电磁超声信号能量微弱影响检测性能。在此基础上考虑到被测管道为圆形与 PCB 印刷电路板上的线圈不能完全贴合，针对此情况对超声传播过程中的声束指向性进行仿真，结果如图 4 所示。

以管道壁厚截面的中心弧线为参考，将横波传播中参考线处声波能量做曲线，如图 5 所示。当线宽较宽、线间距较大时虽然声波能量较小，但声波指向性更好、发散角小，能量更多集中在激励线圈的中心路径上，得到的回波电压信号信噪比更高。经过以上仿真分析，最终决定使用线圈匝数为 8，宽度为 0.6 mm，间距为 1.2 mm 的 PCB 印刷电路板线圈，并对其进行了相应制作，如图 6 所示。通过有限元仿真看到使用的激励线圈所产生的横波在管道内传播路径主要集中在中心位置，在其反射的过程中能量较为集中，与平板中传播情况基本相同，可忽略管道弧形对测量产生的影响，证实方案的可行性。由管壁超声的传播路径和指向性可以得到，该传感器的管壁剩余厚度测量区域为传感器中心区域覆盖处的管道厚度。

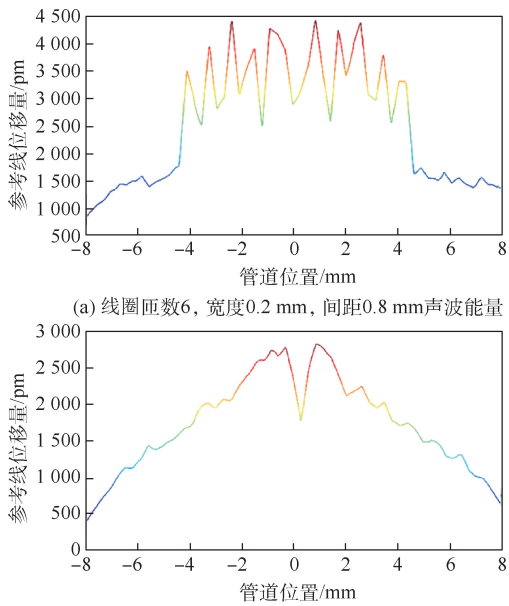


(a) 线圈匝数6，宽度0.2 mm，间距0.8 mm声束辐射

(b) 线圈匝数8，宽度0.6 mm，间距1.0 mm声束辐射

图 4 横波在管道壁中声束辐射图

Fig. 4 Schematic of beam radiation of shear wave in pipe wall



(a) 线圈匝数6，宽度0.2 mm，间距0.8 mm声波能量

(b) 线圈匝数8，宽度0.6 mm，间距1.0 mm声波能量

图 5 横波在管道壁中声波指向性图

Fig. 5 Energy distribution of shear wave in acoustic directivity of pipe wall

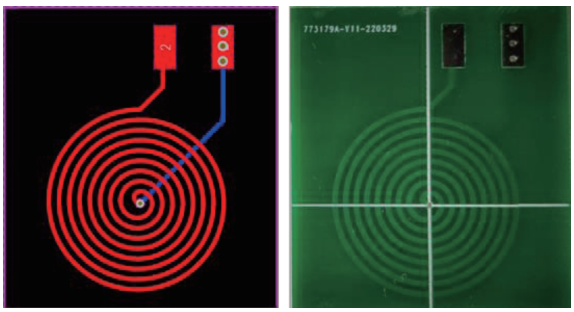


图 6　电磁超声激励用 PCB 螺旋线线圈

Fig. 6　Spiral PCB coil for electromagnetic acoustic excitation

2.2　横波电磁超声测厚系统测试

横波电磁超声激励信号一般采用具有一定占空比的正弦脉冲信号,从信号处理角度来说,其实质是加了零次幂的矩形窗的正弦调制信号。添加不同窗函数调制下的激励信号会对回波电压信号的峰峰值和信噪比产生显著影响。因为不同的调制窗函数会有不同的带宽,所以应尽量选择窄带脉冲激励以实现最大化信噪比[16]。测试系统采用的是与有限元仿真一样的带有海明窗调制的脉冲激励正弦电流信号为

$$f(t) = 20\sin(2\pi ft)\sin\left(\frac{2\pi ft}{2n}\right) \qquad t < 2T \qquad (11)$$

式中:信号频率 f 为 3.5 MHz;脉冲周期数 n 为 5。换能器中线圈激励信号如图 7 所示。

试验测试系统采用自研的高功率脉冲激励源和接收器,如图 8 所示,大功率激励源与换能器中激励线圈经过阻抗匹配之后相连接,由于匹配不完美问题,实际测量中激励线圈端激励电压峰峰值为 726 V,流经线圈的电流峰峰值为 18.15 A,峰峰功率为 13 kW。将换能器放置在外径为 200 mm、壁厚为 5 mm 的铝制管道外表面,将永磁体放置在激励和接收线圈之上,永磁体尺寸、材料及牌号与仿真设置完全一致。换能器中,将与激励线圈相同规格的接收线圈放置在激励线圈上方,经过阻抗匹配之后与接收端电路相连接,经过限幅、滤波、60 dB 的放大和数据处理之后得到回波信号如图 9 所示。

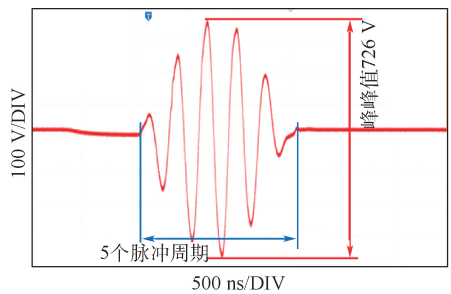


图 7　电磁超声脉冲激励信号

Fig. 7　Electromagnetic acoustic pulse excitation signal

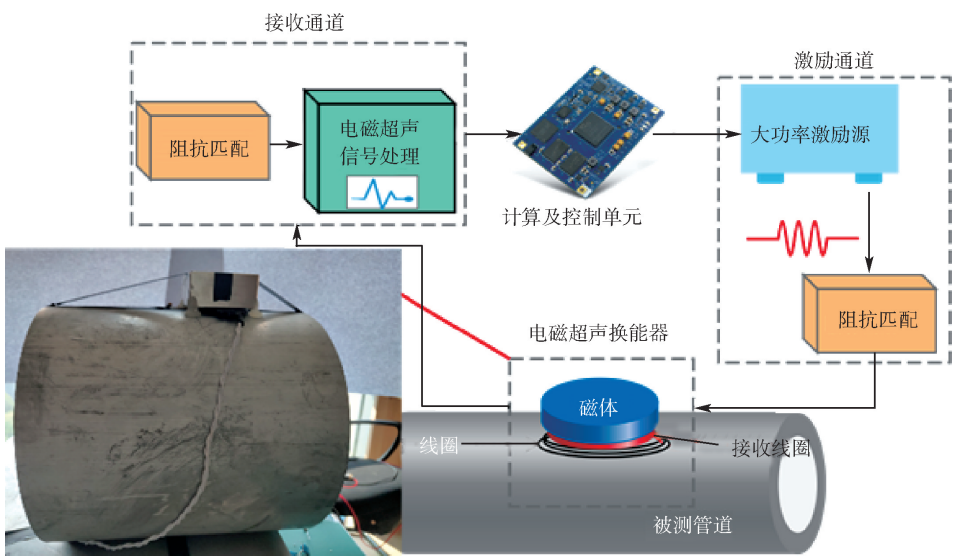


图 8　管道剩余厚度测试系统

Fig. 8　Pipeline residual thickness test system

在处理完成后的回波电压信号(见图 9)可以明显看到管道剩余厚度的 2 次回波电压信号,根据 2 次回波电压的时间间隔可以计算所测量的管道剩余厚度。铝材质中的横波传播速度[17]为

$$v_s = \sqrt{\frac{G}{\rho}} = \sqrt{\frac{E}{\rho}\cdot\frac{2}{2(1+\nu)}} \qquad (12)$$

式中:G 为铝材质的剪切模量;ρ 为铝材质密度;E 为铝材质弹性模量;ν 为铝材质泊松比。对于铝材质而言,$E = 7.0\times10^{10}$ Pa,$\rho = 2\ 700$ kg/m^3,$\mu = 0.33$,故有 $v_s = 3.122\times10^3$ m/s。

根据多次测量数据平均可得到 2 次回波之间的时间间隔为 3.24 μs,根据计算得出的横波声速可得到管道剩余厚度为 5.057 mm,这与用电子游标卡尺多次测量得到的厚度 5.05 mm 误差为

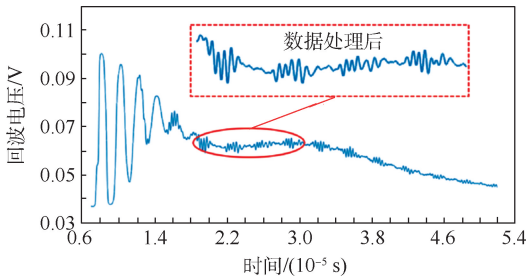


图9 横波电磁超声实际回波电压信号

Fig. 9 Actual echo voltage signal of shear wave electromagnetic acoustic

0.007 mm。同时对标称外直径为200 mm,壁厚为10 mm的铝制管道进行同样操作的壁厚检测,得到2次回波之间的平均时间间隔为6.49 μs,计算管道壁厚为10.131 mm,与用电子游标卡尺多次测量的厚度10.12 mm误差为0.011 mm。

2.3 结果分析

在测试系统中,由于线圈复阻抗测量误差、阻抗匹配计算的数据截断和匹配感容器件的制造误差造成了阻抗匹配后的阻抗还存在感抗,进而引起大功率激励源输出信号的拖尾现象。同时,超声激励时接收电路会进入饱和状态,受限于系统中运算放大器的过载恢复时间限制,会造成激励过后存在约10 μs恢复时间。在此时间内的回波信号淹没在饱和输出中,难以通过数据处理的手段分离和计算。

由于电子游标卡尺的测量分辨率限制,因此在多次测量后只能精确到0.01 mm,在与横波电磁超声测量的厚度进行误差比较时,存在数据的不准确性。

3 结 论

本文根据管道剩余厚度检测的应用需求,设计了基于电磁超声横波的管道剩余厚度检测系统,主要结论如下:

1)针对测量对象将系统中换能器线圈参数进行优化,并仿真横波在管壁中声束辐射指向性验证了换能器受弧形管壁影响较小,满足厚度测量的可行性。

2)针对换能器线圈匝数和线圈宽度分别为回波信号峰峰值和信噪比的最大影响因子,设计了满足系统检测需求的换能器并利用自研电磁超声系统实现了误差小于0.2%的高精度管道剩余壁厚测量。

参考文献(References)

[1] YOSHIDA M, ASANO T. A new method to measure the oxide layer thickness on steels using electromagnetic acoustic resonance[J]. Journal of Nondestructive Evaluation, 2003, 22: 11-19.

[2] HOBBIS A, ARULESWARAN A. Non-contact thickness and profile measurements of rolled aluminium strip using EMAT[C]// AIP Conference, 2006, 820: 1772-1779.

[3] PARRA-RAAD J, KHALILI P, CEGLA F. Shear waves with orthogonal polarisations for thickness measurement and crack detection using EMATs[J]. NDT & E International, 2020, 111: 1-7.

[4] 王相豪. 电磁超声测厚系统的设计与实现[D]. 哈尔滨:哈尔滨工程大学, 2018: 63-69.
WANG X H. Design and realization of electromagnetic ultrasonic thickness measurement system[D]. Harbin: Harbin Engineering University, 2018: 63-69(in Chinese).

[5] 唐志峰,孙兴涛,张鹏飞,等. 集测厚与导波检测于一体的复合式电磁超声换能器研究[J]. 仪器仪表学报, 2020, 41(9): 98-109.
TANG Z F, SUN X T, ZHANG P F, et al. Research on composite electromagnetic ultrasonic transducer integrating thickness measurement and guided wave detection[J]. Chinese Journal of Scientific Instrument, 2020, 41(9): 98-109(in Chinese).

[6] 孙峥,李永虔,杨金旭,等. 管道内检测电磁超声在线测厚装置[J]. 中国测试, 2017, 43(2): 69-72.
SUN Z, LI Y W, YANG J X, et al. Thickness gauging equipment for ILI of pipelines using EMATs[J]. China Measurement & Test, 2017, 43(2): 69-72(in Chinese).

[7] 黄松岭,王坤,赵伟. 电磁超声导波理论与应用[M]. 北京:清华大学出版社, 2013: 2-29.
HUANG S L, WANG K, ZHAO W. Theory and application electromagnetic ultrasonic guided wave[M]. Beijing: Tsinghua University Press, 2013: 2-29(in Chinese).

[8] RIBICHINI R, NAGY P, OGI H. The impact of magnetostriction on the transduction of normal bias field emats[J]. NDT & E International, 2012, 51: 8-15.

[9] MIRKHANI K, CHAGGARES C, MASTERSON C, et al. Optimal design of EMAT transmitters[J]. NDT & E International, 2004, 37(3): 181-193.

[10] JIA X J, QI O Y. Optimal design of point-focusing shear vertical wave electromagnetic ultrasonic transducers based on orthogonal test method[J]. IEEE Sensors Journal, 2018, 18(19): 8064-8073.

[11] YUAN W, LIU Z, LI Y, et al. Comparative analysis of emat receiving process between ferromagnetic and nonferromagnetic materials[C]// 2019 IEEE International Ultrasonics Symposium (IUS). Piscataway: IEEE Press, 2019: 611-614.

[12] KANG L, DIXON S, WANG K C, et al. Enhancement of signal amplitude of surface wave EMATs based on 3D simulation analysis and orthogonal test method[J]. Ndt & E International, 2013, 59: 11-17.

[13] HUANG S L, ZHAO W, ZHANG Y S, et al. Study on the lift-off effect of EMAT[J]. Sensors and Actuators A: Physical, 2009, 153(2): 218-221.

[14] 唐旭明,阳能军,李平. 电磁超声横波换能器中永磁体的建模与优化[J]. 无损探伤, 2017, 41(6): 17-22.
TANG X M, YANG N J, LI P. Modeling method and optimum

design of permanent magnet in elctromagnetic acoustic shear wave transduceers[J]. NDT,2017,41(6):17-22(in Chinese).

[15] PENG J,DONG F,XU Q,et al. Orthogonal test design for optimization of supercritical fluid extraction of daphnoretin,7-methoxy-daphnoretin and 1,5-diphenyl-1-pentanone from Stellera chamaejasme L. and subsequent isolation by high-speed counter-current chromatography[J]. Journal of Chromatography A,2006,1135(2):151-157.

[16] AUGUSTYNIAK M,USAREK Z. Finite element method applied in electromagnetic NDTE:A review[J]. Journal of Nondestructive Evaluation,2016,35(3):1-15.

[17] DAVID J,CHEEKE N. Fundamentals and applications of ultrasonic waves[M]. 2nd ed. Boca Raton:CRC Press,2010:76-90.

Residual thickness detection of pipeline based on electromagnetic ultrasonic shear wave

XU Lijun[1], LIU Fulu[1], DING Yiqing[2], LI Zhengyong[3], XIE Yuedong[1,*]

(1. School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing 100083, China;
2. Shenyuan Honors College, Beihang University, Beijing 100083, China;
3. Shanghai Institute of Aerospace Power Technology, Shanghai 200000, China)

Abstract: To accurately detect the residual thickness of pipeline, a non-contact pipeline thickness detection system based on the electromagnetic ultrasonic shear wave was designed. The system adopts a self-developed electromagnetic ultrasonic high-power excitation source and a transducer to generate shear wave, and the receiver is used to filter and process the echo voltage signal in real-time to obtain the accurate residual thickness of the aluminum pipe. The excitation coil parameters were optimized to improve the small-signal and low signal-to-noise ratio of the receiving signal. On this basis, the beam radiation of shear waves propagating within the pipe was analyzed. Based on the fact that the number of turns of the coil and the width of the coil are the majority influence factors to the peak-to-peak value of the echo signal and the SNR respectively, the transducer is designed and the residual thickness detection accuracy with the error less than 0.2% is achieved.

Keywords: residual pipeline thickness; electromagnetic acoustic transducer; transducer; beam radiation; shear wave

Received: 2022-04-29; **Accepted**: 2022-05-18; **Published online**: 2022-06-07 13:09

URL: kns.cnki.net/kcms/detail/11.2625.V.20220606.1801.004.html

Foundation items: National Natural Science Foundation of China (61901022); Shanghai Academy of Spaceflight Technology Funds (118SAST2021117001); The Fundamental Research Funds for the Central Universities (YWF-19-BJ-J-358)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0302

基于电润湿液体透镜的显微成像技术

王琼华*, 袁荣英, 刘超
(北京航空航天大学 仪器科学与光电工程学院, 北京 100083)

摘 要: 电润湿液体透镜因其焦距可调、响应速度快、操作方便和功耗低等特点越来越受到研究者的青睐,可以解决显微成像系统中诸如机械抖动、离散变焦等技术问题,推进显微成像系统向智能化、轻量化发展。阐述了电润湿液体透镜的基本概念和国内外研究现状;重点综述了基于电润湿液体透镜的显微成像技术的最新研究,详细介绍了连续光学变焦和轴向扫描2类显微成像技术;总结和展望了基于电润湿液体透镜的显微成像技术的机遇和挑战。

关 键 词: 液体透镜; 电润湿液体透镜; 显微成像; 连续光学变焦; 轴向扫描
中图分类号: O435
文献标志码: A **文章编号**: 1001-5965(2022)09-1774-08

显微镜是当今社会不可或缺的重要精密仪器,在医学诊疗、科学研究、地质考古、微纳制造等方面发挥着重要作用。

光学倍率一直是显微镜的重要指标,决定了实际的观察视场和光学分辨率。传统的显微物镜具有固定的离散倍率,通过更换镜头来实现倍率的切换,这带来的是一系列专业且复杂的操作,并且更换镜头后需要重新对焦,不利于对动态样本的实时观察。同时,传统显微物镜也无法实现将细节连续地放大或缩小,仅能进行有限的数码变倍,牺牲了光学分辨率。显微镜连续光学变焦的解决方案之一是改变系统透镜之间的距离[1-5],通过外部装置驱动的机械或光学补偿来达到像面的稳定。然而,这种方式的补偿系统体积庞大,并且由于机械运动,样品振动仍然是一个问题。另外,变焦速度慢也影响了样本的实时观察。

另外,基于显微成像的特点,大多数显微镜的景深很小,只能观测切片式样本。但切片式样本在制作时会破坏样本的形貌,且不能观测样本的深度信息,从而不能还原样本真实形态。传统的显微镜可以通过轴向扫描的方式来实现扩展景深[6-11]。一种方式是通过折射率差的透镜来实现,另一种方式是通过机械移动镜头或载物台来轴向扫描,这会存在机械抖动和响应时间慢等缺点。

液体透镜是指由液体填充所形成的具有透镜功能的光学器件[12],可以解决以上2个传统显微镜面临的技术难题。按驱动原理可将液体透镜笼统地分为电润湿液体透镜、介电泳液体透镜和弹力膜液体透镜。其中,电润湿液体透镜具有较大的有效孔径、光滑的表面形状和更低的重力效应,因此被认为更适用于显微成像系统。基于电润湿液体透镜的显微成像系统不仅可以实现连续的光学变焦,还可以实现无机械辅助的轴向扫描,大大推动了显微成像技术的发展。

本文先介绍了电润湿液体透镜,再详细论述了电润湿液体透镜在显微成像系统中连续光学变焦和轴向扫描的应用,最后讨论了基于电润湿液体透镜的显微成像技术的机遇和尚待解决的难题。

收稿日期: 2022-04-29; **录用日期**: 2022-05-18; **网络出版时间**: 2022-06-01 11:40
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220531.1422.001.html
基金项目: 国家自然科学基金 (61927809)
* **通信作者**. E-mail: qionghua@buaa.edu.cn

引用格式: 王琼华, 袁荣英, 刘超. 基于电润湿液体透镜的显微成像技术[J]. 北京航空航天大学学报, 2022, 48(9): 1774-1781.
WANG Q H, YUAN R Y, LIU C. Microscopic imaging technology with electrowetting liquid lens[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1774-1781(in Chinese).

1　电润湿液体透镜

电润湿液体透镜是利用电润湿效应原理,通过外加电压改变液体界面的接触角,使得液面曲率发生变化,从而实现透镜焦距的改变。电润湿液体透镜的组成包含 2 种或 2 种以上的互不相溶、密度匹配和具有折射率差的透明液体。其中一种是导电液体,一般为有机盐溶液;另一种为非导电液体,一般为硅油等有机高分子溶液。

电润湿液体透镜最早是在 2000 年由法国物理光谱实验室的 Berge 和 Peseux 提出[13],他们利用电毛细作用改变透明液滴的接触角,实现了电润湿液体透镜,该透镜的光焦度变化是人眼的 5 ~ 10倍,响应时间为 0.03 s,功耗为毫瓦量级。随后,荷兰的 Kuiper 和 Hendriks[14]在此基础上提出了最典型的圆柱形电润湿液体透镜,并在照相系统中使用,其结构和原理如图 1(a)所示。该电润湿液体透镜主体部分包含 2 种液体,其中导电液体采用高浓度的氯化锂盐溶液,密度为 1.12 g/cm^3,折射率为 1.38;另一种非导电液体采用溶解了少量四溴化碳的甲基苯基环三硅氧烷溶液,折射率为 1.55。为了增加电润湿液体透镜的稳定性和初始接触角,圆柱结构内壁镀有介电层和疏水层。在初始状态时,由于液体与侧壁界面张力的作用,液-液界面为凸型,如图 1(a)所示,透镜表现为负透镜,光线通过透镜后发散。外加电压后,接触角随电压而改变,透镜焦距随之变化,最终可以表现为正透镜,光线通过透镜后会聚,如图 1(b)所示。图 1(c)展示了圆柱形电润湿液体透镜的实物及工作状态。圆柱形的结构不仅增大了电润湿液体透镜的有效孔径,还实现了正负焦距的变化,增大了变焦范围。

2017 年,德国的 Zappe 等[15]提出了一种双层液-液界面的电润湿液体透镜,通过在不同位置设计不同高度的电极结构,实现多层导电液体的曲率操控,进一步扩大了电润湿液体透镜的焦距变化范围。2021 年,该团队又通过优化液体特性、利用液体黏度和表面张力之间的相互作用提升了电润湿液体透镜的响应速度[16]。

电润湿液体透镜已经实现了商业化,并可以应用到多种光学系统中[17]。2018 年,美国康宁公司开发了一系列经验模型来预测电润湿液体透镜的液-液两相的多个物理和化学性质[18],研制了高折射率差的电润湿液透镜的液体配方,获得了用于电润湿透镜的 2 种液体的最佳匹配范围,进一步指导了电润湿液体透镜的商业发展。

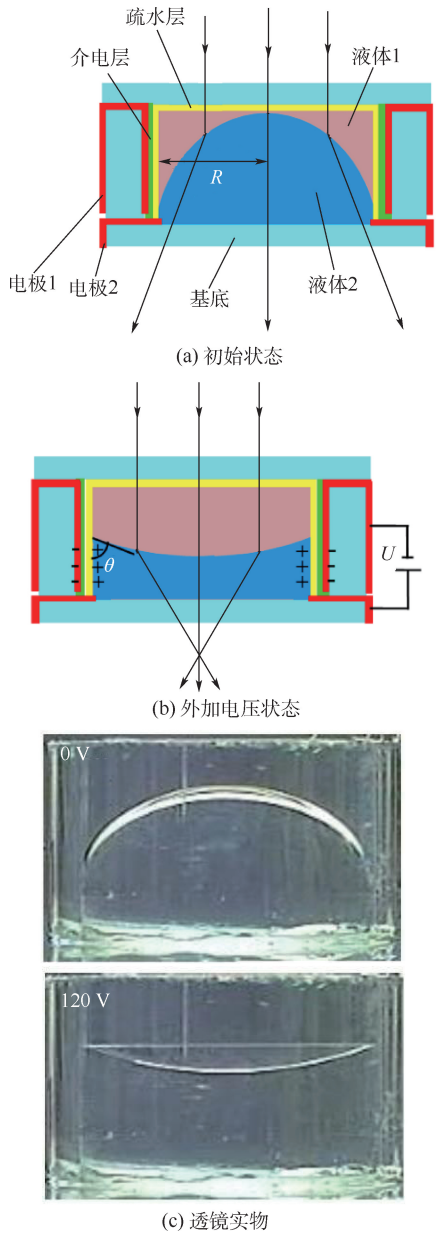


图 1　圆柱形电润湿液体透镜[14]

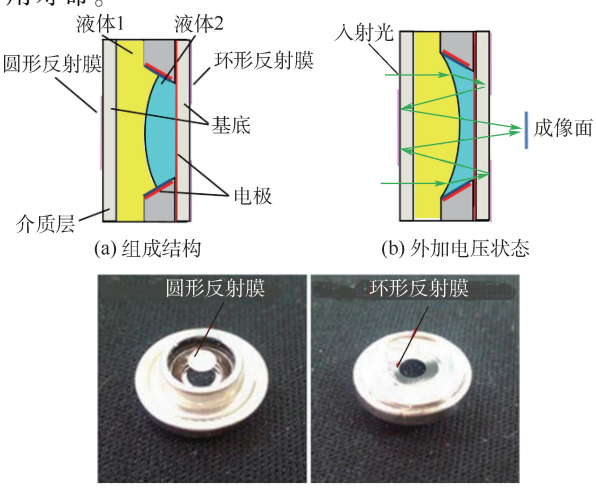
Fig. 1　Cylindrical electrowetting liquid lens[14]

国内,清华大学岳瑞峰教授团队[19]在 2012 年提出了一种由涂有导电氧化铟锡(ITO)膜、疏水电介质膜的载玻片、中空锥形金属环和液体组成的电润湿液体透镜,该透镜通过改变施加在金属环和 ITO 控制电极上的电压,可逆地调整液体弯月面的位置和曲率,从而调整透镜的焦距,该电润湿液体透镜的最短焦距可以达到 2.5 cm。2020 年,天津大学的张红霞等[20]计算了不同孔径的电润湿液体透镜的液-液界面面型,分析了不同电压下电润湿液体透镜的焦距范围、波前像差和孔径之间的关系,提出了一种三层液体的电润湿液体透镜,提高了电润湿液体透镜的数值孔径,其均方根波前像差误差小于四分之一波长。

笔者团队一直致力于电润湿液体透镜的研

究。2015 年,笔者团队提出了具有多重环形反射面的电润湿液体透镜[21],其结构如图 2(a)所示。在外加电压后,液-液界面曲率的调节结合内部多重环形反射结构可有效提升透镜的光焦度变化量,如图 2(b)所示。图 2(c)展示了多重环形反射面的电润湿液体透镜的实物,该液体透镜具有结构紧凑、变焦范围大的优点。2019 年,笔者团队又结合电润湿液体透镜的可调焦性质并辅助非球面透镜,提出了一种集成非球面的电润湿液体透镜[22],如图 3(a)所示。随着外加电压的变化,液体界面曲率发生改变,电润湿液体透镜的焦距随之改变,同时,当光通过非球面时,由于界面折射率不同,光可以被精确地聚集以消除球差和畸变。图 3(b)、(c)展示了该集成非球面的电润湿液体透镜的实物及施加不同电压时该透镜的成像结果,该透镜在一定调焦范围内同时消除部分像差,提高了电润湿液体透镜成像质量。

笔者团队对电润湿液体透镜的结构进行了优化设计,并对其液体材料进行了研究,提出了一种基于无水有机溶液的球形电润湿液体透镜[23],其结构如图 4(a)所示。透镜的腔体为球形,可以扩大接触角范围,从而进一步扩大透镜的焦距变化范围。透镜实物和成像结果如图 4(b)、(c)所示,其中导电液体为有机盐醇溶液,密度为 1.048 g/cm^3,折射率为 1.438 7,非导电液体为芳香卤代烃和汽油分段烷烃的混合溶液,密度为 1.048 g/cm^3,折射率为 1.490 1。无水有机溶液可以避免加电压后的水解现象,即使电润湿液体透镜的介电层失效,该透镜仍能继续使用,这将大大增加电润湿液体透镜的稳定性,延长使用寿命。

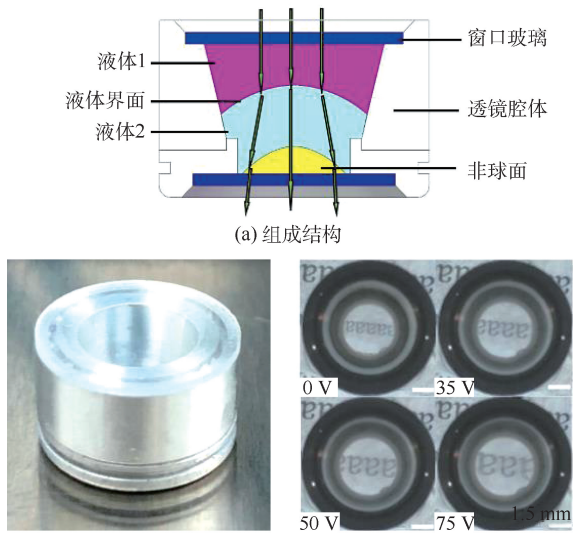


(a) 组成结构 (b) 外加电压状态

(c) 透镜实物

图 2 多重环形反射面的电润湿液体透镜[21]

Fig. 2 Electrowetting liquid lens with multiple annular reflective surfaces[21]

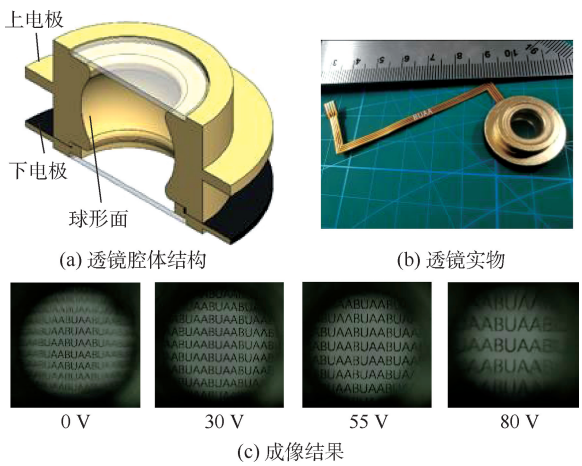


(a) 组成结构

(b) 透镜实物 (c) 成像结果

图 3 集成非球面的电润湿液体透镜[22]

Fig. 3 Electrowetting liquid lens integrated with aspherical surface[22]



(a) 透镜腔体结构 (b) 透镜实物

(c) 成像结果

图 4 基于无水有机溶液的球形电润湿液体透镜[23]

Fig. 4 Spherical electrowetting liquid lens with non-aqueous organic solution[23]

2 电润湿液体透镜在显微成像中的应用

研究表明,由于电润湿液体透镜的机理限制,电润湿液体透镜的口径较小,在小口径的成像系统中更具有优势,显微镜属于典型的小口径成像系统,电润湿液体透镜已经广泛应用于各种显微成像系统中。根据电润湿液体透镜在显微成像系统中的功能,本文将基于电润湿液体透镜的显微成像技术大致分为 2 类:连续光学变焦和轴向扫描。

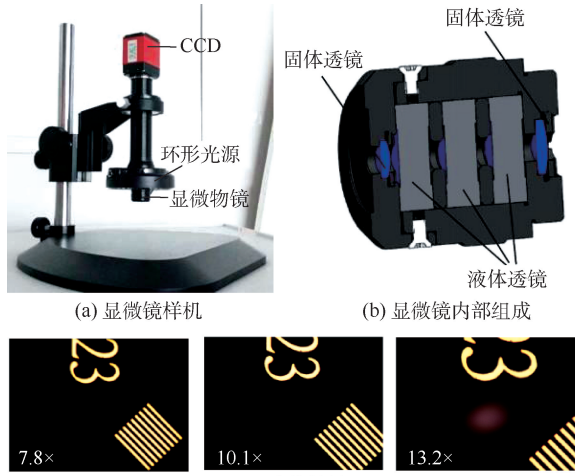
2.1 连续光学变焦

电润湿液体透镜最大的优势在于具有自适应变焦功能,因此,基于电润湿液体透镜的显微镜可以实现无机械移动的连续光学变焦,打破了传统显微镜离散变焦的局限。

2016 年，笔者团队首次研制出了基于电润湿液体透镜的连续光学变焦显微镜样机[24]，如图 5(a)所示。该样机主要由显微物镜、环形光源、镜筒和 CCD 组成。其中，显微物镜由 3 个电润湿液体透镜和 2 个固体透镜组成，如图 5(b)所示。通过外加电压控制电润湿液体透镜的焦距，在无任何机械移动的情况之下，该显微镜可以实现 7.8 ~ 13.2 倍的连续光学变焦，其倍率切换时间约为 50 ms。此外，在变焦的过程中，通过 3 个电润湿液体透镜的配合还可以实时校正成像的像差，提高显微镜的成像分辨率，成像结果如图 5(c)所示。

近期，笔者团队通过优化设计进一步提升了基于电润湿液体透镜的连续光学变焦显微成像技术[25]，提高了成像的分辨率，增大了变焦范围。其中，显微镜的物镜由 4 个电润湿液体透镜和6 个固体透镜组成，如图 6(a)所示。固体透镜承担了系统大部分的光焦度，再通过协调施加在4 个液体透镜上的驱动电压，可以实现从 9.6 倍到 22.2 倍连续光学变焦。计算并验证了电润湿液体透镜的曲率半径和系统放大倍率的关系，结果如图 6(b)所示。本文研制了该显微物镜样机并测试了其成像结果，如图 6(c)、(d)所示，结果表明，该显微物镜在变焦过程中可以保持 200 lp/mm 的高分辨率(lp/mm 指线对每毫米，是镜头分辨率计算单位)。

为了进一步扩大连续光学变焦显微镜的变焦范围，笔者团队研制了一种由变焦物镜和变焦目镜组成的大变焦范围的连续光学变焦显微镜[26]，将显微镜的变焦范围提升到了 59.1 ~ 159.2 倍。该显微镜的系统组成和原理如图 7(a)所示，其中，变焦物镜由 3 个固体透镜和 4 个电润湿液体透镜组成，变焦目镜由消色差目镜和 1 个电润湿

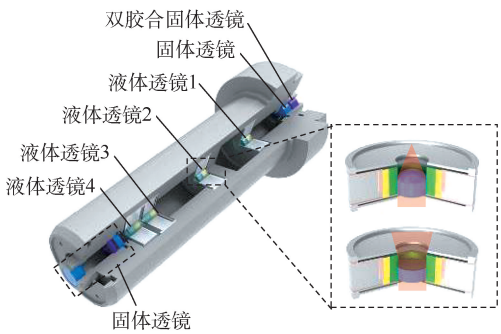


(a) 显微镜样机　(b) 显微镜内部组成

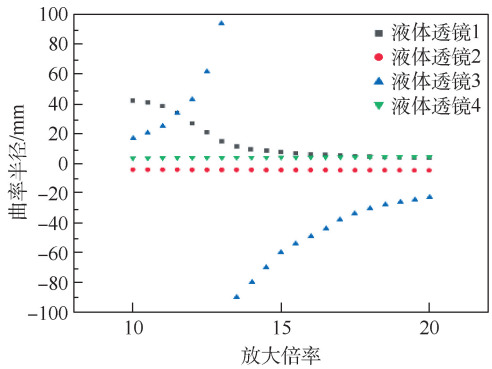
(c) 成像结果

图 5　基于电润湿液体透镜的连续光学变焦显微镜[24]

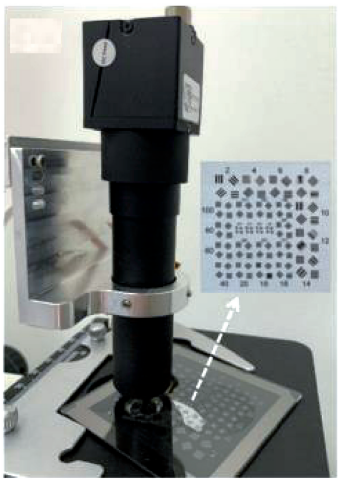
Fig. 5　Continuous optical zoom microscope with electrowetting liquid lens[24]

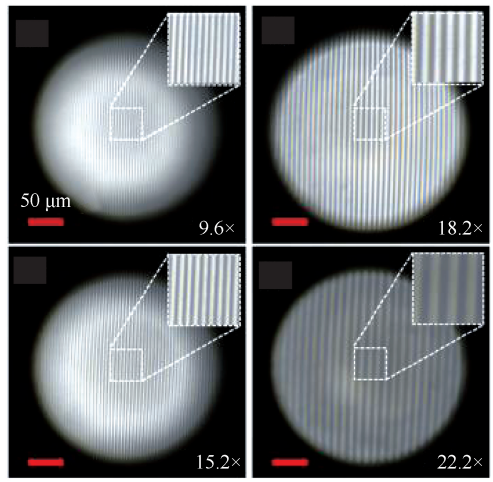


(a) 显微物镜的结构组成



(b) 电润湿液体透镜的曲率半径与系统的放大倍率关系

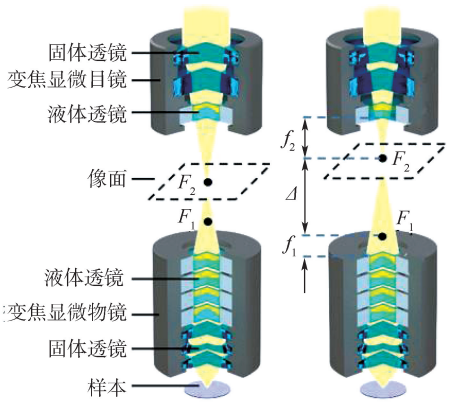

(c) 显微物镜样机



(d) 成像结果

图 6　高分辨率的连续光学变焦显微物镜[25]

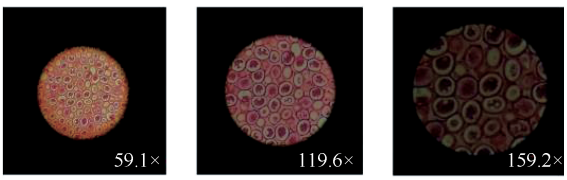
Fig. 6　Microscopic objective of high resolution continuous optical zoom[25]



(a) 显微镜组成及原理



(b) 显微镜样机



(c) 成像结果

图 7 大变焦范围的连续光学变焦显微镜[26]

Fig. 7 Continuous optical zoom microscope with large zoom range[26]

液体透镜组成。外加电压改变目镜和物镜中电润湿液体透镜的焦距时,物镜和目镜的焦距及两者之间的像面随之变化,这在很大程度上增大了焦距的调节范围。本文研制的显微镜如图 7(b)所示,其成像结果如图 7(c)所示,该显微镜的最大数值孔径为 0.212,在变焦过程中具有良好的成像质量。

2.2 轴向扫描

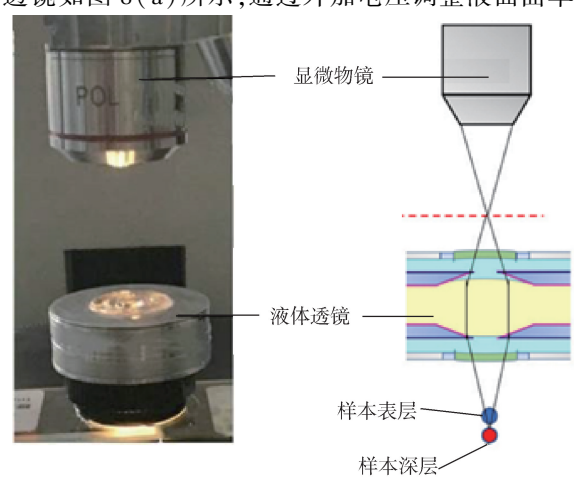
显微镜轴向扫描是实现大景深显微的关键技术,该技术不仅能够提升成像面的深度范围,还可以通过层析扫描获取样本的深度信息和 3D 重建,大大推动了显微成像技术及相关行业的研究进展。电润湿液体透镜可以自适应调焦,是实现轴向扫描最佳的方式之一。

美国中佛罗里达大学的 Murali 等[27]基于电润湿液体透镜设计了一种高分辨率三维扫描显微系统,该显微系统具有自适应调焦功能,不仅可以无机械运动实现快速聚焦,还可以对重新聚焦产生的光学像差进行补偿。该显微镜的分辨率达到 250 lp/mm,成像深度为 2 mm,可应用于生物医学成像、干涉测量及 3D 成像等。随后,该团队又提出了基于电润湿液体透镜的具有像差补偿动态重聚焦的光学相干显微镜[28],通过实验对显微镜的聚焦深度进行了量化表征,应用该显微镜获得了横向 2 mm、纵向 0.8 mm、8 个不同聚焦位置的一系列体内清晰的横断面图像。

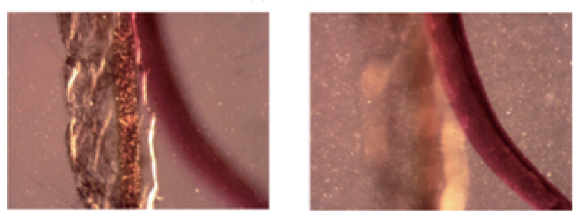
美国乔治亚大学的 Tehrani 等[29]使用电润湿液体透镜在多光子显微镜中进行远程聚焦扫描,该方法可以应用于多个活体标本的成像和测量,实现了活体动态活动的五维双光子体积显微成像。

华南师范大学的 Yang 等[30]基于电润湿液体透镜提出了一个快速可控的共聚焦光声显微系统,以实现不同深度表面不规则或多层结构标本的整体切片成像。该共聚焦光声显微系统的物镜主要由电润湿液体透镜和聚偏二氟乙烯传感器组成,通过电润湿液体透镜的自适应变焦,物镜可以调整共聚焦光声显微系统的焦距,实现了约 6~43 mm的共聚焦扫描范围,且具有较高的横向分辨率和较短的聚焦时间。

笔者团队提出了一种双液-液曲面的电润湿液体透镜[31],并将其用于显微成像系统中,实现了在无机械移动的情况下任意倍率的显微物镜都可以进行轴向扫描,且保持放大倍率不变。该显微系统的组成结构及双液-液曲面的电润湿液体透镜如图 8(a)所示,通过外加电压调整液面曲率,



(a) 系统组成



(b) 聚焦在样本表层 (c) 聚焦在样本深层

图 8 基于双液-液曲面电润湿液体透镜的轴向扫描显微系统

Fig. 8 Axial scanning microscopic system using electrowetting liquid lens with two liquid-liquid curved interfaces

可以实现对样本的轴向扫描。显微系统的扫描成像结果如图 8(b)、(c)所示。实验表明,该系统对深度 1 mm 的样本每一个层面均能清晰成像。

3　结 束 语

电润湿液体透镜是近年来光学成像领域冉冉升起的新星,具有灵活的透镜界面,被认为是下一代微型光学透镜。基于电润湿液体透镜的显微成像系统可以快速连续光学变焦来实时观测样本,且能实现活体样本的 3D 观测,解决了部分显微成像系统存在的技术瓶颈,一定程度上推动了显微成像技术的发展。本文介绍了电润湿液体透镜的发展,综述了基于电润湿液体透镜的显微成像技术的国内外研究现状,并讨论了基于电润湿液体透镜的显微成像技术的机遇和挑战。

显微系统的轴向扫描仅仅需要一片电润湿液体透镜就能实现,可以与任意的现有显微镜结合,是电润湿液体透镜在显微成像系统应用中相对比较简单和成熟的技术手段。基于电润湿液体透镜实现显微镜的连续光学变焦被认为是电润湿液体透镜在显微成像系统应用中最有前景同时又是最有挑战的研究,这将打破显微镜离散变焦的传统局限,同时大大减小显微镜的尺寸,给显微镜的发展带来新的发展契机。

然而,电润湿液体透镜依然存在一些科学问题亟待解决,基于电润湿液体透镜的显微成像技术仍存在一些挑战。

从电润湿液体透镜方面来看,其在显微系统中的使用目前还存在以下 3 个技术难题需要去解决:

1) 电润湿液体透镜达到最大/最小焦距所需的驱动电压约为 60 V,需要单独的驱动器来控制变焦,不利于显微系统的集成化。

2) 电润湿液体透镜由 2 种液体组成,液体的折射率、阿贝数会影响透镜的光焦度、成像质量。目前的电润湿液体透镜光焦度和成像质量还有待提高。

3) 电润湿液体透镜的孔径依然有限,目前商业化的最大口径仅为 5.8 mm,这也限制了显微镜获得更大的分辨率。

从显微成像系统方面来看,基于电润湿液体透镜的显微技术还存在以下 2 个挑战:

1) 电润湿液体透镜的变焦范围较小,导致显微系统的变焦比有限。为了获得更大的变焦范围,可以使用更多的电润湿液体透镜,而目前电润湿液体透镜孔径较小,增加透镜数量会导致显微成像系统的分辨率和亮度下降,这与显微镜的高分辨要求相矛盾,如何在保持较高的成像质量的前提下增大变焦范围是基于电润湿液体透镜的显微成像技术面临的最大挑战。

2) 尽管电润湿液体透镜的尺寸较小,但其驱动尺寸相对较大。如何将电润湿液体透镜的驱动与显微成像系统集成,使基于电润湿液体透镜的显微镜微型化、便携化,这是研究者们面临的另一个挑战。

参考文献 (References)

[1] HUANG Y, CHEN X, ZHANG H, et al. Design of a high-performance digital slit-lamp microscope with five-switched zoom [J]. Applied Sciences, 2020, 10(8): 2757.

[2] 张润泽. 长工作距变焦显微系统物镜设计[D]. 长春:长春理工大学, 2019: 12-16.
ZHANG R Z. Design of zoom microscope objective with long working distance [D]. Changchun: Changchun University of Science and Technology, 2019: 12-16 (in Chinese).

[3] 阎帅, 向阳, 李琦, 等. 用于细胞发酵罐的长工作距变焦显微系统设计[J]. 长春理工大学学报(自然科学版), 2020, 43(2): 8-13.
YAN S, XIANG Y, LI Q, et al. Design oflong working distance zoom microscopy system for cell fermenter [J]. Journal of Changchun University of Science and Technology (Natural Science Edition), 2020, 43(2): 8-13 (in Chinese).

[4] 许利峰, 张新, 曲贺盟, 等. 高变倍比连续变焦体视显微镜物镜设计[J]. 应用光学, 2013, 34(1): 38-44.
XU L F, ZHANG X, QU H M, et al. Optical design of high-magnification zoom stereo microscope objective [J]. Journal of Applied Optics, 2013, 34(1): 38-44 (in Chinese).

[5] 王佳楠, 向阳. 变焦电子腹腔镜物镜光学系统的设计[J]. 长春理工大学学报(自然科学版), 2018, 41(1): 67-70.
WANG J N, XIANG Y. Design of optical system for zoom electronic laparoscope objective lens [J]. Journal of Changchun University of Science and Technology (Natural Science Edition), 2018, 41(1): 67-70 (in Chinese).

[6] ZONG W, ZHAO J, CHEN X, et al. Large-field high-resolution two-photon digital scanned light-sheet microscopy [J]. Cell Research, 2015, 25(2): 254-257.

[7] 宋贤林. 大景深轴向高分辨光声显微成像方法研究[D]. 武汉:华中科技大学, 2019: 14-29.
SONG X L. Photoacoustic microscopy with high axial resolution in a large depth of field [D]. Wuhan: Huazhong University of Science and Technology, 2019: 14-29 (in Chinese).

[8] BENSCHOP J, VAN ROSMALEN G. Confocal compact scanning optical microscope based on compact disc technology [J]. Applied Optics, 1991, 30(10): 1179-1184.

[9] DIXON A E, DAMASKINOS SATKINSON M R. Transmission and double-reflection scanning stage confocal microscope [J]. Scanning, 1991, 13(4): 299-306.

[10] GREWE B F, LANGER D, KASPER H, et al. High-speed in vi-

vo calcium imaging reveals neuronal network activity with near-millisecond precision[J]. Nature Methods, 2010, 7(5): 399-405.

[11] SHI R, JIN C, XIE H, et al. Multi-plane, wide-field fluorescent microscopy for biodynamic imaging in vivo[J]. Biomedical Optics Express, 2019, 10(12): 6625-6635.

[12] 王琼华,刘超,王迪,等. 液体光子器件[M]. 北京:科学出版社,2021:76.
WANG Q H, LIU C, WANG D, et al. Liquid photonic device [M]. Beijing: Science Press, 2021: 76(in Chinese).

[13] BERGE B, PESEUX J. Variable focal lens controlled by an external voltage: An application of electrowetting[J]. The European Physical Journal E, 2000, 3(2): 159-163.

[14] KUIPER S, HENDRIKS B H W. Variable-focus liquid lens for miniature cameras[J]. Applied Physics Letters, 2004, 85(7): 1128-1130.

[15] KOPP D, BRENDER T, ZAPPE H. All-liquid dual-lens optofluidic zoom system[J]. Applied Optics, 2017, 56(13): 3758-3763.

[16] ZHAO P, LI Y, ZAPPE H. Accelerated electrowetting-based tunable fluidic lenses[J]. Optics Express, 2021, 29(10): 15733-15746.

[17] ERIC S, BRUNO B, FRANCK F, et al. Optical design rules of a camera module with a liquid lens and principle of command for AF and OIS functions[C]// Proceedings of SPIE, 2010, 7849: 784903-1-784903-14.

[18] OBER M S, DERMODY D, MAILLARD M, et al. Development of biphasic formulations for use in electrowetting-based liquid lenses with a high refractive index difference[J]. ACS Combinatorial Science, 2018, 20(9): 554-566.

[19] KANG M, YUE R F. Variable-focus liquid lens based on EWOD[J]. Journal of Adhesion Science and Technology, 2012, 26(12-17): 1941-1946.

[20] SONG X, ZHANG H X, LI D, et al. Electrowetting lens with large aperture and focal length tunability[J]. Scientific Reports, 2020, 10(1): 16318.

[21] LI L, LIU C, REN H, et al. Annular folded electrowetting liquid lens[J]. Optics Letters, 2015, 40(9): 1968-1971.

[22] WANG J H, ZHOU X, LUO L, et al. Tunable liquid lens integrated with aspheric surface[J]. Optics Communications, 2019, 445: 56-63.

[23] XU J, ZHAO Y, LIU C, et al. Non-aqueous organic solution based on a large-aperture spherical electrowetting liquid lens with a wide tunable focal length range[J]. Journal of Materials Chemistry C, 2022, 10(17): 6778-6793.

[24] LI L, WANG D, LIU C, et al. Zoom microscope objective using electrowetting lenses[J]. Optics Express, 2016, 24(3): 2931-2940.

[25] JIANG Z, WANG D, ZHENG Y, et al. Continuous optical zoom microscopy imaging system based on liquid lenses[J]. Optics Express, 2021, 29(13): 20322-20335.

[26] KUANG F L, YUAN R Y, WANG Q H, et al. Large zooming range adaptive microscope employing tunable objective and eyepiece[J]. Scientific Reports, 2020, 10(1): 14644.

[27] MURALI S, THOMPSON K, ROLLAND J. Three-dimensional adaptive microscopy using embedded liquid lens[J]. Optics Letters, 2009, 34(2): 145-147.

[28] MURALI S, MEEMON P, LEE K S, et al. Assessment of a liquid lens enabled in vivooptical coherence microscope[J]. Applied Optics, 2010, 49(16): 145-156.

[29] TEHRANI K F, LATCHOUMANE C V, SOUTHERN W M, et al. Five-dimensional two-photon volumetric microscopy of in-vivo dynamic activities using liquid lens remote focusing[J]. Biomedical Optics Express, 2019, 10(7): 3591-3604.

[30] MA H, CHENG Z, WANG Z, et al. Fast controllable confocal focus photoacoustic microscopy using a synchronous zoom optosono objective[J]. Optics Letters, 2019, 44(7): 1880-1883.

[31] LI L, XIAO L, WANG J H, et al. Movable electrowetting optofluidic lens for optical axial scanning in microscopy[J]. Opto-Electronic Advances, 2019, 2(2): 10-14.

Microscopic imaging technology with electrowetting liquid lens

WANG Qionghua*, YUAN Rongying, LIU Chao

(School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing 100083, China)

Abstract: The electrowetting liquid lens is increasingly favored by researchers due to its tunable focal length, fast response, convenient operation, and low power consumption. It can solve some technical problems such as mechanical jitter and discrete zoom in microscopic imaging systems, thus promoting the intelligence and lightweight of these systems. Firstly, the basic concept of and the research on electrowetting liquid lenses are described. Then, the latest research on the microscopic imaging technology with the electrowetting liquid lens is reviewed. Two microscopic imaging technologies, continuous optical zoom and axial scanning, are introduced in detail. Finally, the prosects and challenges of the microscopic imaging technology with the electrowetting liquid lens are presented.

Keywords: liquid lens; electrowetting liquid lens; microscopic imaging; continuous optical zoom; axial scanning

Received: 2022-04-29; **Accepted**: 2022-05-18; **Published online**: 2022-06-01 11:40

URL: kns.cnki.net/kcms/detail/11.2625.V.20220531.1422.001.html

Foundation item: National Natural Science Foundation of China (61927809)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0295

仿生减阻表面的进展与挑战

刘明杰[1,*], 吴青山[1], 严昊[1], 于存明[1], 蔺存国[2], 张金伟[2], 赵天艺[1]

(1. 北京航空航天大学 化学学院, 北京 100083;
2. 中国船舶重工集团公司第七二五研究所 海洋腐蚀与防护重点实验室, 青岛 266237)

摘 要: 减阻表面由于其在海洋船舶、管道运输、飞机和国防军事等各个领域展现出的巨大潜力和广泛的应用前景而受到越来越多的关注。自然界中的动植物进化出了许多独特的形态或表面结构,均展现出了卓越的减阻性能。对近年来受海豚表面启发的顺服表面减阻、鲨鱼皮表面启发的微结构减阻及受荷叶表面启发的超疏水减阻等进行了总结,并介绍了这些技术的研究进展、减阻机理和挑战,对目前仿生减阻表面的研究进行了总结与展望。

关 键 词: 仿生; 减阻; 顺服表面; 微结构; 超疏水
中图分类号: Q692; O357.4
文献标志码: A **文章编号**: 1001-5965(2022)09-1782-09

自 20 世纪 70 年代开始,全球能源危机日益加剧,导致各种新型节能减排技术的发展变得日益紧迫[1]。幸运的是,自然界中的生物总是可以提供灵感,经过数十亿年的自然选择和进化,它们在如何更高效地利用能量方面展现出了卓越的本领。例如,高效飞行的鸟类为优化各种飞行器提供了灵感,长期快速游动的鱼类则为优化各种水下舰艇提供了许多指导等[2]。近几十年来,仿生学蓬勃发展,尤其是关于流体控制领域。流体控制领域的目标主要包括减阻、提高升力、减少流体诱导的噪声和促进传热等[3-4],其中减阻技术对于飞机、潜水器、水面舰船和长距离管道等系统都有着巨大的应用需求,因此显得尤其重要[1,5-6]。

当流体在真实环境中流过一个物体时,根据不同的情况可能会产生许多类型的阻力,如摩擦阻力、形状阻力(压差阻力)、诱导阻力、波动阻力和干涉阻力等[7],其中摩擦阻力和形状阻力是最常见也是影响最大的 2 类阻力,目前的减阻技术主要研究如何降低这 2 类阻力的负面影响。对于形状阻力,自然界中的生物一般是尽可能地使自身形态成流线型,这样可以大大减少压差阻力。对于摩擦阻力,不同的生物采用了不同的策略,其中最典型 3 类分别为:①海豚,可以利用其光滑柔性表面的黏弹特性形成顺服表面,从而延缓表面流体的转捩,实现高效的游动[8];②鲨鱼,表面具有精细的三维齿状结构,可以有效地调控近壁面的湍流结构,从而实现高速的游动[9];③荷叶,可以在表面形成一层气垫,从而导致水滴在其表面的摩擦阻力减少[10]。经过近几十年的研究,这 3 种策略已经逐渐发展成为减阻技术中最重要的三大分支,即行波减阻、微结构减阻和超疏水减阻。

本文聚焦于目前仿生减阻技术方面的研究进

收稿日期: 2022-04-29; **录用日期**: 2022-06-02; **网络出版时间**: 2022-06-22 09:34
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220621.1021.001.html
基金项目: 国家自然科学基金-"叶企孙"科学基金(U2141251); 国家自然科学基金(22175010); 国家杰出青年自然科学基金(21725401)
* **通信作者**. E-mail: liumj@buaa.edu.cn

引用格式: 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).

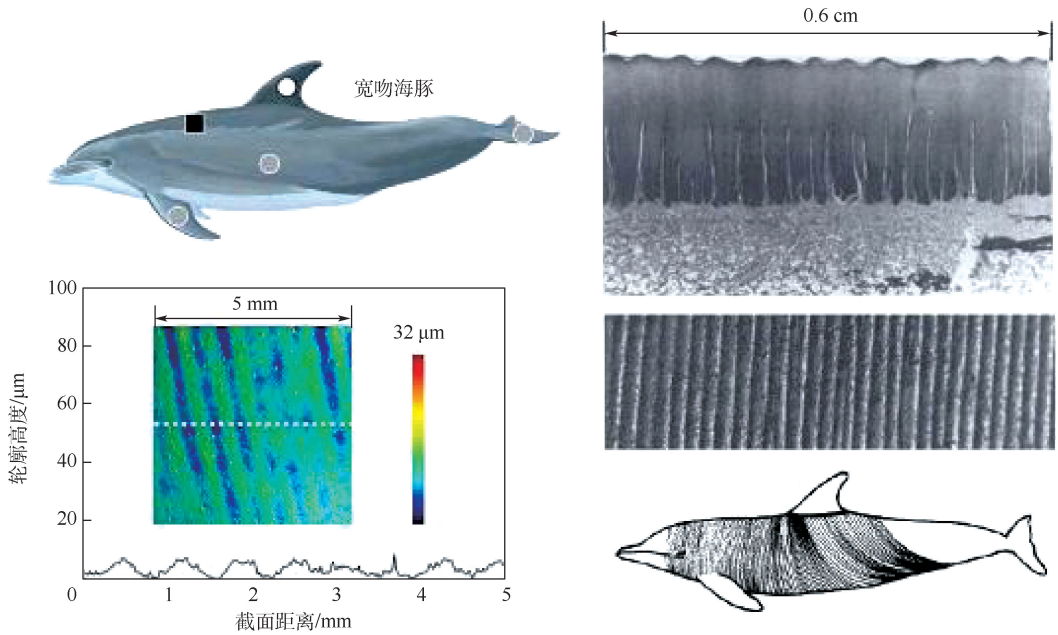
展,主要讨论以海豚、鲨鱼和荷叶表面为仿生原型的顺服表面行波减阻、微结构减阻和超疏水减阻技术,并总结其最新进展、减阻机理和应用前景,旨在对目前仿生流体控制领域中的减阻研究进行全面的梳理。

1　海豚表面启发的顺服表面减阻

1.1　顺服表面减阻

顺服表面是指在水流的冲击下会被动发生形变的柔性表面。顺服表面可以更好地适应水流流过,使水流中引起转捩发生的不稳定波变弱,进而推迟层流在表面的转捩,导致层流区域在表面变长,从而实现减阻[11]。顺服表面是最典型的也是最早启发学者们开始研究其减阻性能的,即海豚表面。海豚表面非常光滑,平均粗糙度仅 5.3 μm 左右[12],如图 1(a)所示。通过进一步切片研究发现,海豚表皮下方有许多可以感受水压的乳头结构。这些被包裹在皮下组织的液体会随着压力的变化流出或流入细管,细管嵌入皮下组织导致皮肤上下收缩或肿胀,进而产生振动[8],如图 1(b)所示。这种受到湍流的压力变化而被动地振动并导致行波在表皮上传播,从而推迟表面流体转捩进而减少摩擦阻力的情况被称为顺服表面减阻。

顺服表面的减阻实验研究最早是由 Kramer[13]在 1960 年进行的,在一个圆柱表面覆盖了一层模仿海豚表皮的柔性涂层,测试了其减阻性能,相同条件下,覆盖柔性层的圆柱比纯圆柱的速度提高近一倍,不过该实验结果一直存在争议。直到 20 世纪 80 年代,Carpenter 和 Garrad[14]详细讨论了“Kramer 争议”,并在理论上证实了 Kramer 的柔性表面确实可以推迟层流到湍流的转捩。随着进一步的研究,发现在优化顺服表面减阻性能时,尽管顺服表面可以有效地减少 TS(Tollmien-Schlichting)波的增长,但是顺服表面的阻尼会造成不稳定波的增长。1984 年,Gad-el-Hak 等[15]发现了行波减阻中的色散会导致非线性效应。1990 年,Carpenter[11]研究了行波颤振对行波减阻的影响,并提出了弹簧-平板-杆模型来解释减阻机理。至此,关于顺服表面的行波减阻的优化思路主要是:抑制 TS 波增长的同时,减少顺服表面的行波颤振和由色散所导致的非线性效应。之后的研究开始以理论研究为主。2017 年,Tsigklifis 和Lucey[16]对零压力梯度边界层中的顺服表面进行了局部和全局稳定性分析,发现 TS 波和行波颤振失稳都可以形成全局失稳。2018 年,Nagy 和 Paál[17]对只能沿顺流方向运动的顺服表面的转捩延迟性能进行了研究,发现顺服表面的展向运动不影响 TS 波不稳定性的增长速率,且不会产生对流不稳定性。从上述研究进展可以发现,2000 年之前的研究主要讨论影响行波减阻的因素,探索并证明其减阻机理。进入 21 世纪之后,学者们将目光投向了理论模拟,希望利用理论模拟来为实验提供指导。



(a) 海豚表面的形貌表征[12]　　(b) 海豚表面内部的乳头结构与遍布全身的褶皱[8]

图 1　海豚表面结构

Fig. 1　Dolphin surface structure

1.2 顺服表面的减阻机理

顺服表面与水流相互作用的关系对于研究减阻机理是非常重要的。实验中,顺服表面对水流产生的非定常力的响应性区别较大,因此很难用统一的模型去解释。目前的理论模型有 2 类:一类是以表面为主体的模型,即考虑表面的运动方程时不考虑垂直于表面的变量影响;第二类则是以体积为主体的模型,该模型同时考虑表面的运动和垂直于表面的运动。对于表面型减阻模型,理论认为产生一个负的雷诺剪应力是减少湍流不稳定发展的方法。于是 Grosskreutz[18] 根据该思路设计了一种各向异性的顺服表面,如图 2 所示[11]。Carpenter[11] 根据该表面提出了各向异性的弹簧-平板-杆模型,用来分析和解释减阻的原因。当杆的偏转角度等于 0 时,就变成了经典的 Kramer 顺服表面,即弹簧-平板模型。对于体积型减阻模型,则是从理论上最简单的顺服表面出发,认为表面是一层理想均匀的黏弹性层。Duncan 等[19] 根据 Navier 方程,适当扩展黏弹性相,从而提出了体积模型。Yeo[20] 随后发展出了单层和双层的纤维复合的各向异性顺服表面模型。

非定常流体在顺服表面的流动涉及 2 种含波介质的相互作用,与刚性平面的区别在于顺服表面涉及失稳模式的扩散。对于刚性平面和顺服表面,TS 波是导致边界层不稳定从而转捩的原因。除此之外,对于顺服表面目前还有 2 类不稳定,即色散和行波颤振。通过对边界层的雷诺剪切力的研究表明,壁面的柔性对扰动速度既有局部影响

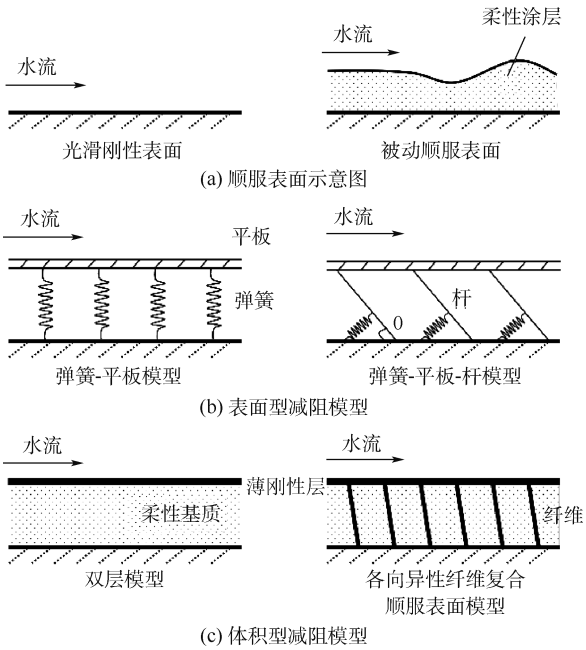


图 2 顺服表面的减阻机理[11]

Fig. 2 Drag reduction mechanism of compliant surface[11]

也有较长范围的影响。尽管局部增加雷诺应力不利于稳定,但是长程来看总体的雷诺剪切力会大大下降。由于雷诺应力产生的能量减少和向壁面传递的能量增加,柔度的增加对 TS 波具有稳定作用。对于行波颤振,其产生的主要原因是壁面压力脉动所产生的不可逆功。阻尼会起到不利的结果,Gad-el-Hak[21] 发现,当阻尼增大时,壁面的行波颤振会被色散作用取代。行波颤振的临界速度小于色散的临界速度,因此阻尼的增加更倾向于稳定行波颤振,从而使得色散成为主要的不稳定性。总之,TS 波的抑制有利于减阻,行波颤振和色散则对减阻不利。

1.3 挑 战

目前的实验表明,在使用针对局部雷诺数定制的较小的顺服表面时效果更好,因此在不影响抑制 TS 波发展的情况下,顺服表面可以做多小是值得进一步研究的问题[11]。另外,这种顺服表面对于气流来说,需要很精细的壁面,似乎很难在实际应用中采用[22]。总之,柔性壁面为边界层流体的控制开辟了新的思路,在实际应用中也有着巨大的潜力。

2 鲨鱼皮表面启发的微结构减阻

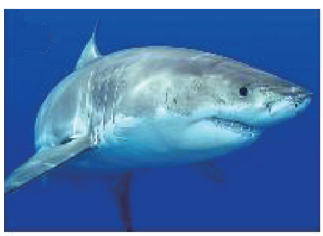
2.1 微结构减阻

海洋中高速游动的生物为减少船舶和其他水下航行器(如潜艇、鱼雷等)的阻力提供了大量的灵感,其中鲨鱼可以称为海洋中的“速度之王”,世界上速度最快的鲨鱼(Isurus oxyrinchus)速度可以超过 56 km/h。之所以可以游动的如此之快,除了完美的流线型身体将压差阻力降至最低,还与表面的三维齿状结构有关。这与海豚光滑的表面截然相反。鲨鱼皮表面含有精细的表面微结构,可以有效调控近壁面的湍流结构,从而大大降低表面的摩擦阻力[23]。这打破了人们曾经认为只有越光滑的表面才更能有效减阻的常规认知,并开启了微结构减阻的新思路。

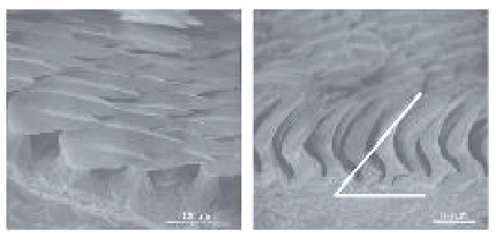
鲨鱼表面的齿状结构大小约为 0.2 ~ 0.5 mm,分为外层和内层,如图 3 所示。外层由牙釉质组成,内层则是坚硬的骨骼结构。齿状结构表面还存在平行于水流方向的沟槽,沟槽高度约为 8 μm,宽度约为 60 μm[24]。值得注意的是,这种齿状结构不会随着鲨鱼的增长而变大,其大小主要取决于生长的位置和鲨鱼的种类[25]。通过直接用新鲜鲨鱼皮或复制鲨鱼皮牙齿结构进行减阻分析的结果表明,鲨鱼皮确实可以减少阻力,

提高游速和防污[26]。由于鲨鱼皮肤齿状体结构复杂,难以大规模生产,为了方便研究,研究人员往往将鲨鱼皮表面精细的三维结构简化成不同形状的肋条结构。一般将含有沟槽的三维齿状结构简化成截面形状为三角形(或 V 型)、梯形、圆形、矩形和波浪形等的二维肋条结构[27],再根据不同肋条结构的形状和尺寸,在一定条件下(如流体介质和流体速度)获得最佳的减阻能力。最早开始研究肋条结构减阻的是 NASA 兰利研究中心,20 世纪 80 年代,Walsh[28]研究了含有肋条结构的平板的湍流减阻性能,发现当 V 型肋条的无量纲高度 $h^+ \leqslant 25$,无量纲宽度 $s^+ \leqslant 30$ 时会出现减阻效果,当 $h^+ = 13$,$s^+ = 15$,能实现 7% ~8% 的最佳减阻效果,该结果也为之后的研究提供了指导。之后肋条结构减阻的研究采用的无量纲尺度基本都在 10 ~30 之间。1985 年,Bacher 和 Smith[29]研究了三角肋条的减阻性能,提出了二次涡理论。1989 年,Bechert 和 Bartenwerfer[30]研究了含有肋条结构的湍流边界层中黏性底层的性质,提出了突出高度理论,为之后的沟槽结构减阻提供了理论基础。1997 年,Bechert 等[31]结合理论与实验优化了肋条结构的条件,提出当肋条的高度为相邻肋条间距的一半时具有最佳的减阻效果(能达到 8.7%)。

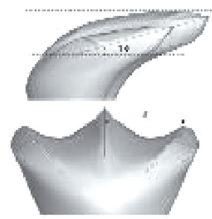
尽管早年关于肋条结构表面的减阻效果进行了许多研究与优化,但是显然人造肋条表面与真实的鲨鱼皮还有很大的差距,这主要是因为真实的鲨鱼皮表面的齿状结构是三维的,除了肋条之外,还可以在水流的作用下改变攻角,使得鲨鱼皮可以更加适应水流。随着 3D 打印技术的兴起,鲨鱼皮复杂的三维结构也可以实现完美的复制。2014 年,Lauder 课题组[32]利用 micro-CT 技术扫描鲨鱼皮的结构,并且利用 3D 打印技术进行了复制,制备了一个含有人造三维的鲨鱼皮肤的游动装置,发现其游泳速度可以提高 6.6%,认为这是因为真实的三维结构可以产生一个更大的前缘漩涡。2018 年,Miyazaki 等[33]提出了三维鲨鱼皮齿状结构的形态不均匀性在被动控制局部湍流中起着关键作用。总之,目前关于鲨鱼皮简化的二维肋条结构的减阻研究已经比较全面,基本上可以实现 8% 左右的减阻效果,不过对于三维的齿状结构的减阻机理研究还不够,甚至出现很多矛盾的地方。例如,2016 年 Boomsma 和 Sotiropoulos[34]利用数值模拟的三维鲨鱼皮结构不仅没有减阻,还增加了阻力。因此,关于三维齿状结构的减阻机理还需要进一步的研究。



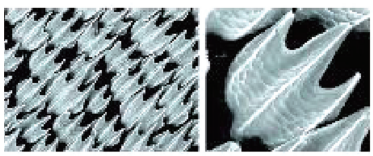
(a) 鲨鱼表面的齿状结构[24]



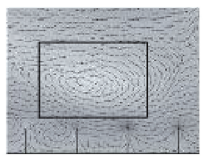
(b) 齿状结构的侧视图[35]



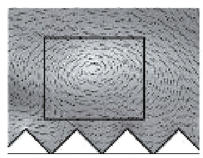
(c) 齿状结构的三维模型[34]



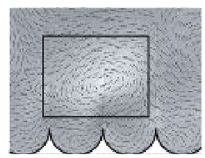
齿状结构



刀片



锯齿



扇形

(d) 二维肋条结构[27]

图 3　鲨鱼皮的结构

Fig. 3　Structure of shark skin

2.2　二维肋条结构的减阻机理

关于鲨鱼皮减阻的机理主要是针对二维肋条结构,有 2 种理论[36],如图 4 所示。第一种理论是突出高度理论,突出高度是指高于黏性底层的高度。突出高度分又为 2 种情况:①当水流方向与肋条结构平行时,突出高度如图 4(c)上部所示,指从肋条顶端到沟槽底部黏性底层处的高度;②当水流方向与肋条结构垂直时,由于沟槽内部会产生低流速的漩涡,相当于提高了黏性底层的高度,这种情况的突出高度如图 4(c)下部所示,这时纵向的突出高度与横向的突出高度会有一个高度差,该高度差说明了水流更容易沿着肋条方向运动,很少发生横向的运动,说明肋条结构可以将水流限制在肋条之间,减少了横向运动的消耗,从而减少了阻力。另外,当流向涡的尺度比肋条的间距要大时,漩涡与肋条尖端的接触面积很小,

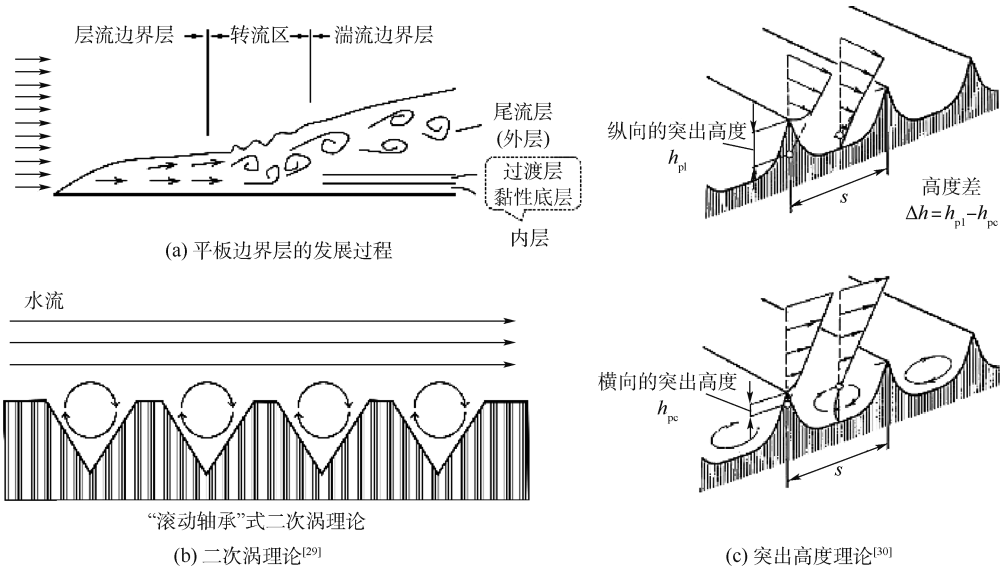


图4 鲨鱼皮的减阻机理[36]

Fig. 4 Drag reduction mechanism of shark skin[36]

这也减少了与壁面的摩擦阻力。

第二种理论是二次涡理论。尺度较大的流向涡不会进入槽内,但是会有一些尺度较小的二次涡进入槽内,而这些停留在槽内的二次涡会对阻力产生影响,当肋条的宽度和深度适当时,沟槽底部会出现涡流,这些旋涡像轴承一样不断滚动,使其与壁面的摩擦阻力变成了动力,从而实现减阻,如图4(b)所示[29]。二次涡可以有效地削弱湍流动量的交换,减少流向涡射流和扫流事件的发生。同时,二次涡的存在会导致流体在流经微结构表面时发生部分滑移。总之,肋条结构的存在可以阻碍流向涡的展向运动,并且肋条结构还会在流动方向上产生涡流,将高湍动能区域推离肋条表面,导致湍流动量传递和表面摩擦减小。

2.3 挑 战

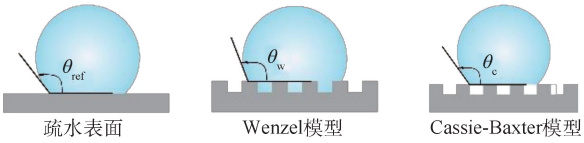
虽然关于鲨鱼皮仿生微结构减阻的研究已经取得了巨大的进展,但是要想在实际场合中应用,仍有许多问题需要解决。首先,大面积制备的工艺仍需要改进,目前关于3D打印的制备技术还不能实现大面积制备,而滚动热压方式的精度还需要提高[24,34]。其次,真实的鲨鱼皮三维齿状结构是可以运动的[35],会根据不同场合改变其攻角,目前人造的鲨鱼皮结构还不具备实时运动的能力。最后,关于鲨鱼皮三维齿状结构的减阻机理目前还缺乏统一的解释,需要进一步的研究[29]。当然,受鲨鱼皮启发的微结构减阻打破了之前认为只有光滑表面才减阻的认知壁垒,为之后各种新型的基于非光滑表面的减阻技术提供了不同的研究途径。

3 荷叶表面启发的超疏水减阻

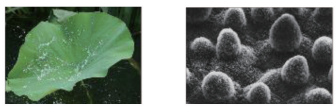
3.1 超疏水减阻

新型的非光滑表面减阻技术中最典型的就是超疏水减阻。超疏水表面通常由经过化学处理的疏水的微/纳米尺度结构组成,其最普遍的区分方式是根据接触角来定义,即当水与表面的接触角大于150°时[37],称该表面为超疏水表面。值得注意的是,对于一般光滑的疏水表面,其与水的接触角一般不会超过120°,因此要想实现超疏水,即进一步增加疏水表面的接触角,需要在疏水的基础上引入微/纳结构[10]。之所以如此,是因为对于粗糙表面的浸润性而言(见图5(a)),分为2种情况:对于已经浸润的粗糙表面,根据Wenzel模型[38],其实际的表观接触角 $\cos\theta_w = r\cos\theta_{ref}$;对于非浸润的粗糙表面,根据Cassie-Baxter模型[39],其实际表观接触角 $\cos\theta_c = \varphi_s\cos\theta_{ref} - \varphi_g$,$\varphi_s$ 表示表面固体组分的占比,φ_g 表示表面气体组分的占比。这2种模型都表明随着表面粗糙度的增大,表观接触角会比固有的光滑表面的接触角要大。即引入表面粗糙结构可以使亲水表面更亲水,疏水表面更疏水。这也解释了为什么目前发现的大自然中的各种超疏水表面往往都具有微/纳结构。

超疏水现象在许多自然界中的生物表面都能发现,其中,最典型的例子就是荷叶表面。早在1997年,Barthlott和Neinhuis[10]就发现荷叶表面的结构由许多微纳米的突起结构和表层蜡晶体组成(见图5(b)),并且提出正是由于表面的这些微结构和疏水表层蜡导致了超疏水现象,在水滴



(a) 不同表面的浸润性[37]



(b) 荷叶表面的微观结构[10]

图 5　超疏水表面

Fig. 5　Superhydrophobic surface

的滚动作用下可以带走表面的固体污染物从而实现自清洁。学者们将这种出淤泥而不染的超疏水自清洁现象称为荷叶效应。2002 年,Kim 等[40]制备了具有微纳米尺度微结构的超疏水表面,发现液滴在这种表面上具有减阻效果,并提出了有效滑移的概念。至此开启了学者们对于超疏水表面减阻功能的研究。2003 年,Lauga 和 Stone[41]研究了纵向条纹(沿水流方向)和横向条纹(垂直水流方向)这 2 种超疏水微纳结构在压力驱动的斯托克斯管道流中的减阻情况,发现当滑移区的间距较小时,纵向条纹的有效滑移长度是横向条纹的 2 倍。2006 年,Henoch 等[42]研究了“纳米草”和“纳米砖”结构在湍流中的减阻效果,发现与光滑的 PVC 平板相比,层流中的阻力下降了 50%,湍流中的阻力下降了约 10%。2009 年,Daniello[43]和 Woodlford[44]等分别通过压力降及 PIV 的方式测试了纵向条纹在湍流中的减阻效果,发现纵向条纹约有 18% 的减阻效果,但是横向条纹减阻效果不明显甚至会增加阻力。2013 年,Aljallis 等[45]研究了大面积疏水纳米粒子喷涂涂层的减阻效果,发现在高雷诺数下可以实现约 30% 的减阻效果。2015 年,Cohen 课题组[46]研究了泰勒-库埃特流中超疏水转子表面的减阻能力,提出在高雷诺数下,有效滑移长度与雷诺数的平方根相关。2020 年,Kim 课题组[47]首次报道了在真实户外环境中超疏水表面减阻的成功案例。

通过上述关于超疏水减阻的研究进展可以发现,从开始关于有效滑移长度的理论研究,再到实验室条件下层流湍流实验的验证,以及大面积的超疏水表面减阻测试和实际开放水域超疏水表面减阻的成功案例报道,超疏水减阻似乎越来越接近实际应用,这对于海洋运输、远距离液体运输管道系统等来说是一个令人振奋的消息。

3.2　超疏水减阻的机理

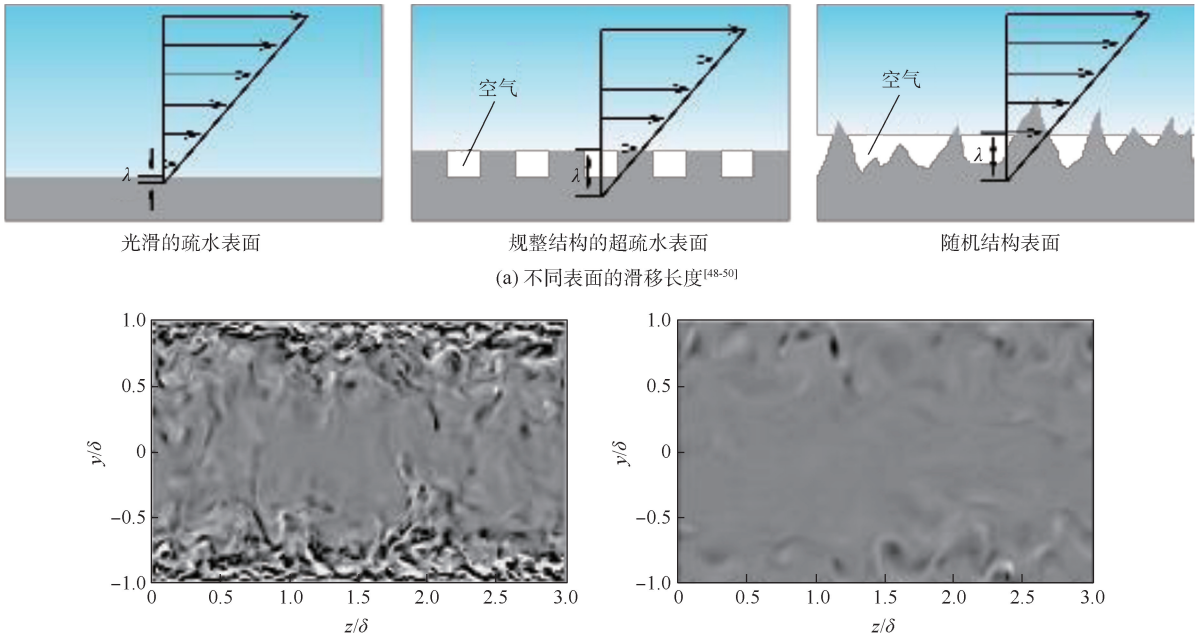
关于超疏水减阻机理,目前普遍用有效滑移长度给予解释。当超疏水表面处于水下时,在粗糙结构之间会捕获空气。与传统的固-液界面无滑移的边界条件不同,超疏水表面上的部分气-液界面将导致有效滑移。当滑移长度与流体结构的特征尺寸相当时,可以出现减阻效果。如图 6(a)所示,根据 Navier 润滑模型[48],润滑长度 λ 表示润滑速度线性外推至 0 的虚拟表面与真实表面($y=0$)之间的距离。

在层流中[49],超疏水表面减阻主要与滑移长度和流动几何形状的特征长度有关,当滑移长度与流体几何形状的特征长度相当时,可以显著减少表面摩擦阻力。但是在湍流中[50],滑移长度不再是决定湍流减阻的唯一决定性因素,不仅涉及固体壁面的有效滑移,还涉及近壁面湍流结构的抑制[51],如图 6(b)所示。

要想实现有效的超疏水减阻,有 3 个基本原则:①表面微结构之间要有足够的宽度,即产生尽可能大的滑移长度;②微结构的高度要低,要小于流体的黏性尺度;③最好有多级结构来帮助保持气体和抵抗液体的浸润。

3.3　挑　战

超疏水减阻目前的主要挑战在于空气层的寿命问题,大多数实验室流动系统(如水洞)中,超疏水表面的被困空气很少受到挑战,甚至大部分是设计成对表面空气层有利的,而在真实的开放水域中,大部分情况下的水是欠饱和水(如海水和拖曳水池),并且随着深度的增加,被困的气泡层还会受到越来越大的静压挑战,根据亨利定律可知,静水压的增大会增加气体在水中的扩散速率,这进一步减少了空气层的寿命,对超疏水减阻构成了根本性挑战[49-50]。此外,如果开放水域处于典型的野外自然环境条件下,则水中可能还会含有大量的化学物质(如表面活性剂)和颗粒污染物,并存在大量的不确定性(如温度和盐度),这些不确定性大多都不利于维持长效性的减阻。目前,关于延长空气层寿命学者们也做出了许多努力:第一种思路是增强表面束缚气体的能力,如采用多级次的结构来增强固定空气的能力[52];第二种则是直接给表面补充气体,如人工地去给表面气体增加压力[53],或利用表面电解等方式产生气体来自我补充[54]。总的来说,目前关于超疏水表面的研究都展现了其在减阻方面的巨大潜力,不过超疏水表面还可能在自清洁、防雾、防冰、防生物黏附和防腐蚀等其他领域发挥出令人期待的作用。



(a) 不同表面的滑移长度[48-50]

(b) 抑制近壁面湍流的结构[51]

图6 滑移长度

Fig. 6 Slip length

4 结束语

减阻技术在很多应用场合都有着巨大的潜力,本文从仿生流体控制领域出发,总结了3类受生物启发的减阻技术的研究进展、减阻的机理和存在的挑战与问题。海豚表面的柔性顺服表面具有更强的水流适应性,能有效抑制表面TS波发生推迟转捩,从而实现了有效的减阻效果。而鲨鱼皮表面进化出的精细的三维齿状结构,打破了只有光滑表面能减阻的常规认知,开启了微结构减阻。研究发现,鲨鱼皮的微结构可以有效调控表面湍流的结构,并且抑制表面水流的横向位移,减少了壁面消耗,从而实现减阻。新型微结构减阻中最典型的超疏水减阻,其受到荷叶等具有疏水微/纳结构的启发,发现可以通过表面疏水的微结构束缚气泡,从而有效增加壁面的滑移长度,实现减阻。总之,大自然中的生物可以为新型减阻技术提供源源不断的灵感,并且这些新型减阻技术目前都已经得到了蓬勃的发展。当然,现今最先进的技术往往需要多学科的交叉融合,这也是新型减阻技术需要进一步探索和研究的地方。

参考文献(References)

[1] TIAN G, FAN D, FENG X, et al. Thriving artificial underwater drag-reduction materials inspired from aquatic animals: Progresses and challenges[J]. RSC Advances, 2021, 11(6): 3399-3428.

[2] CHOI H, PARK H, SAGONG W, et al. Biomimetic flow control based on morphological features of living creatures[J]. Physics of Fluids, 2012, 24(12): 121302.

[3] GAD-EL-HAK M. Flow control: The future[J]. Journal of Aircraft, 2001, 38(3): 402-402.

[4] KROO I. Drag due to lift: Concepts for prediction and reduction[J]. Annual Review of Fluid Mechanics, 2001, 33(1): 587-617.

[5] SINDAGI S, VIJAYAKUMAR R. Succinct review of MBDR/BDR technique in reducing ship's drag[J]. Ships and Offshore Structures, 2021, 16(9): 968-979.

[6] KORNILOV V I, BOIKO A V. Advances and challenges in periodic forcing of the turbulent boundary layer on a body of revolution[J]. Progress in Aerospace Sciences, 2018, 98: 57-73.

[7] 侯辉昌. 减阻力学[M]. 北京: 科学出版社, 1987.
HOU H C. Drag reduction mechanics[M]. Beijing: Science Press, 1987(in Chinese).

[8] LANG A W, JONES E M, AFROZ F. Separation control over a grooved surface inspired by dolphin skin[J]. Bioinspiration & Biomimetics, 2017, 12(2): 1-35.

[9] AFROZ F, LANG A, HABEGGER M L, et al. Experimental study of laminar and turbulent boundary layer separation control of shark skin[J]. Bioinspiration & Biomimetics, 2016, 12(1): 016009.

[10] NEINHUIS C, BARTHLOTT W. Purity of the sacred lotus, or escape from contamination in biological surfaces[J]. Planta, 1997, 202(1): 1-8.

[11] CARPENTER P W. Status of transition delay using compliant walls[M]// BUSHNELL D M, HEFNER J N. Viscous drag reduction in boundary layers. Reston: AIAA, 1990: 79-113.

[12] WAINWRIGHT D K, FISH F E, INGERSOLL S, et al. How smooth is a dolphin? The ridged skin of odontocetes[J]. Biolo-

gy Letters,2019,15(7):20190103.

[13] KRAMER M O. Boundary layer stabilization by distributed damping[J]. Journal of the American Society for Naval Engineers,1960,72(1):25-34.

[14] CARPENTER P W,GARRAD A D. The hydrodynamic stability of flow over Kramer-type compliant surfaces. Part 1. Tollmien-Schlichting instabilities[J]. Journal of Fluid Mechanics,1985,155:465-510.

[15] GAD-EL-HAK M,BLACKWELDER R F,RILEY J J. On the interaction of compliant coatings with boundary-layer flows[J]. Journal of Fluid Mechanics,1984,140:257-280.

[16] TSIGKLIFIS K,LUCEY A D. The interaction of Blasius boundary-layer flow with a compliant panel:Global,local and transient analyses[J]. Journal of Fluid Mechanics,2017,827:155-193.

[17] NAGY P T,PAÁL G. The effect of spanwise and streamwise flexible coating on the boundary layer transition[J]. Journal of Fluids and Structures,2018,110:103521.

[18] GROSSKREUTZ R. An attempt to control boundary-layer turbulence with nonisotropic compliant walls[J]. University Science Journal (Dar es Salaam),1975,1:67-73.

[19] DUNCAN J H,WAXMAN A M,TULIN M P. The dynamics of waves at the interface between a viscoelastic coating and a fluid flow[J]. Journal of Fluid Mechanics,1985,158:177-197.

[20] YEO K S. The stability of boundary-layer flow over single- and multi-layer viscoelastic walls[J]. Journal of Fluid Mechanics,1988,196:359-408.

[21] GAD-EL-HAK M. The response of elastic and viscoelastic surfaces to a turbulent boundary layer[J]. Journal of Applied Mechanics,1986,53(1):206-212.

[22] CARPENTER P W. The optimization of compliant surfaces for transition delay[M]// LIEPMANN H W,NARASIMHA R. Turbulence management and relaminarisation. Berlin:Springer,1988:305-313.

[23] PU X,LI G,LIU Y. Progress and perspective of studies on biomimetic shark skin drag reduction[J]. ChemBioEng Reviews,2016,3(1):26-40.

[24] LI W,WEAVER J C,LAUDER G V. Biomimetic shark skin:Design,fabrication and hydrodynamic function[J]. Journal of Experimental Biology,2014,217(10):1656-1666.

[25] BECHERT D W,BRUSE M,HAGE W. Experiments with three-dimensional riblets as an idealized model of shark skin[J]. Experiments in Fluids,2000,28(5):403-412.

[26] YU C,LIU M,LEI J,et al. Bio-inspired drag reduction:From nature organisms to artificial functional surfaces[J]. Giant,2020,2:100017.

[27] SAMUEL M,BHARAT B. Modeling and optimization of shark-inspired riblet geometries for low drag applications[J]. Journal of Colloid and Interface Science,2016,474:206-215.

[28] WALSH M J. Riblets as a viscous drag reduction technique[J]. AIAA Journal,1983,21(4):485-486.

[29] BACHER E V,SMITH C R. A combined visualization-anemometry study of the turbulent drag reducing mechanisms of triangular micro-groove surface modifications:AIAA 1985-548[R]. Reston:AIAA,1985.

[30] BECHERT D W,BARTENWERFER M. The viscous flow on surfaces with longitudinal ribs[J]. Journal of Fluid Mechanics,1989,206:105-129.

[31] BECHERT D W,BRUSE M,HAGE W,et al. Experiments on drag-reducing surfaces and their optimization with an adjustable geometry[J]. Journal of Fluid Mechanics,1997,338:59-87.

[32] 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.

[33] MIYAZAKI M,HIRAI Y,MORIYA H,et al. Biomimetic riblets inspired by sharkskin denticles:Digitizing,modeling and flow simulation[J]. Journal of Bionic Engineering,2018,15(6):999-1011.

[34] BOOMSMA A,SOTIROPOULOS F. Direct numerical simulation of sharkskin denticles in turbulent channel flow[J]. Physics of Fluids,2016,28(3):59-87.

[35] LANG A W,BRADSHAW M T,SMITH J A,et al. Movable shark scales act as a passive dynamic micro-roughness to control flow separation[J]. Bioinspiration & Biomimetics,2014,9(3):036017.

[36] TIAN G Z,FENG X M. Focus on bioinspired textured surfaces toward fluid drag reduction:Recent progresses and challenges[J]. Advanced Engineering Materials,2022,24:2100696.

[37] QUÉRÉ D. Non-sticking drops[J]. Reports on Progress in Physics,2005,68(11):2495-2532.

[38] WENZEL R Z. Resistance of solid surfaces to wetting by water[J]. Industrial & Engineering Chemistry,1936,28(8):988-994.

[39] CASSIE A B D,BAXTER S. Wettablity of porous surfaces[J]. Transactions of the Faraday Society,1944,40:546-551.

[40] KIM J,KIM C J. Nanostructured surfaces for dramatic reduction of flow resistance in droplet-based microfluidics[C]// Proceedings of the 15th IEEE International Conference on Micro Electro Mechanical System. Piscataway:IEEE Press,2002:479-482.

[41] LAUGA E,STONE H A. Effective slip in pressure-driven Stokes flow[J]. Journal of Fluid Mechanics,2003,489:55-77.

[42] HENOCH C,KRUPENKIN T N,KOLODNER P,et al. Turbulent drag reduction using superhydrophobic surfaces:AIAA 2006-3192[R]. Reston:AIAA,2006.

[43] DANIELLO R J,WATERHOUSE N E,ROTHSTEIN J P. Drag reduction in turbulent flows over superhydrophobic surfaces[J]. Physics of Fluids,2009,21(8):085103.

[44] WOOLFORD B,PRINCE J,MAYNES D,et al. Particle image velocimetry characterization of turbulent channel flow with rib patterned superhydrophobic walls[J]. Physics of Fluids,2009,21(8):3395-3478.

[45] ALJALLIS E,SARSHAR M A,DATLA R,et al. Experimental study of skin friction drag reduction on superhydrophobic flat plates in high Reynolds number boundary layer flow[J]. Physics of Fluids,2013,25(2):351-412.

[46] SRINIVASAN S,KLEINGARTNER J A,GILBERT J B,et al. Sustainable drag reduction in turbulent taylor-couette flows by depositing sprayable superhydrophobic surfaces[J]. Physical Review Letters,2015,114(1):014501.

[47] XU M,GRABOWSKI A,YU N,et al. Superhydrophobic drag reduction for turbulent flows in open water[J]. Physical Review

Applied,2020,13:034056.

[48] NETO C,EVANS D R,BONACCURSO E,et al. Boundary slip in Newtonian liquids:A review of experimental studies[J]. Reports on Progress in Physics,2005,68(12):2859.

[49] LEE C,CHOI C H,KIM C J. Superhydrophobic drag reduction in laminar flows:A critical review[J]. Experiments in Fluids,2016,57(12):1-20.

[50] HYUNGMIN P,CHANG H C,KIM C J. Superhydrophobic drag reduction in turbulent flows:A critical review[J]. Experiments in Fluids,2021,62(11):229.

[51] PARK H,PARK H,KIM J. A numerical study of the effects of superhydrophobic surface on skin-friction drag in turbulent channel flow[J]. Physics of Fluids,2013,25(11):110815.

[52] BARTHLOTT W,SCHIMMEL T,WIERSCH S,et al. The salvinia paradox:Superhydrophobic surfaces with hydrophilic pins for air retention under water[J]. Advanced Materials,2010,22(21):2325-2328.

[53] CARLBORG C F,DO-QUANG M,STEMME G,et al. Continuous flow switching by pneumatic actuation of the air lubrication layer on superhydrophobic microchannel walls[C]// Proceedings of the 21st IEEE International Conference on Micro Electro Mechanical systems. Piscataway:IEEE Press,2008:599-602.

[54] LEE C,KIM C J. Underwater restoration and retention of gases on superhydrophobic surfaces for drag reduction[J]. Physical Review Letters,2011,106(1):014502.

Progress and challenges of bionic drag reduction surfaces

LIU Mingjie[1,*], WU Qingshan[1], YAN Hao[1], YU Cunming[1], LIN Cunguo[2], ZHANG Jinwei[2], ZHAO Tianyi[1]

(1. School of Chemistry, Beihang University, Beijing 100083, China;
2. State Key Laboratory for Marine Corrosion and Protection, Luoyang Ship Material Research Institute (LSMRI), China Shipbuilding Industry Corporation (CSIC), Qingdao 266237, China)

Abstract: Drag reduction surfaces have attracted more and more attention due to their great potentiality and wide applications in marine ships, pipeline transportation, aircraft, and national defense and military. Interestingly, animals and plants in nature have evolved many unique shapes or surface structures which show excellent drag reduction performance. In this review, we introduce the recent drag reduction technologies of compliant, microstructural, and superhydrophobic surfaces inspired by dolphin surfaces, shark skins, and lotus leaves, respectively. We summarize the research progress, mechanism and challenges of these technologies, and provide an overview of the current research on bionic drag reduction surfaces, and its future perspectives.

Keywords: bionics; drag reduction; compliant surface; microstructure; superhydrophobicity

Received: 2022-04-29; **Accepted**: 2022-06-02; **Published online**: 2022-06-22 09:34

URL: kns.cnki.net/kcms/detail/11.2625.V.20220621.1021.001.html

Foundation items: Ye Qisun Science Foundation of National Natural Science Foundation of China (U2141251); National Natural Science Foundation of China (22175010); National Natural Science Foundation for Distinguished Young Scholars (21725401)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0575

磷酸掺杂型高温质子交换膜燃料电池关键材料研究进展

相艳[1], 李文[1], 郭志斌[2,*], 张劲[1], 卢善富[1]

(1. 北京航空航天大学 空间与环境学院, 北京 100083; 2. 北京海得利兹新技术有限公司, 北京 100162)

摘 要: 高温质子交换膜燃料电池(HT-PEMFC)由于较高的工作温度(130~220℃),具有较快的电极反应动力学、较强的抗燃料/空气中杂质毒化能力、广泛的燃料来源(甲醇重整气、工业副产氢等)及简单的水/热管理系统等优点。因此,HT-PEMFC 将成为聚合物膜燃料电池的重要前沿发展方向之一。重点介绍了北京航空航天大学团队近十年来在 HT-PEMFC 关键材料-高温膜、催化层和膜电极等方面的研究进展,针对磷酸(PA)掺杂型高温膜的质子传导率和机械性能之间的最佳平衡点、催化层中 PA 分布和迁移对电池性能的影响机制,以及大尺寸膜电极一致性对电堆性能影响与衰减机制等科学问题,从聚电解质膜材料的分子设计、有序催化层结构调控和大尺寸膜电极电堆优化等工作进行了梳理,对 HT-PEMFC 技术所面临的技术挑战问题与未来发展趋势做出了评述和展望。

关 键 词: 燃料电池; 高温质子交换膜; 催化层; 膜电极; 高温燃料电池堆
中图分类号: TM911.47
文献标志码: A **文章编号**: 1001-5965(2022)09-1791-15

燃料电池是一种将化学能直接转换为电能的电化学能源转换装置,具有功率密度大、能量转换效率高、环境污染小等众多优点[1-3],可广泛应用于便携式电源、电动汽车等领域。因此,发展燃料电池技术对解决能源与环境问题具有重要的社会意义。其中,高温质子交换膜燃料电池(HT-PEMFC)由于较高的工作温度(130~220℃),使其具有较快的电极反应动力学、较高的杂质气体(CO, SO_2 等)耐受性、简单的水/热管理系统及燃料来源广泛等优点,受到广泛关注与应用[4-8]。

1 高温质子交换膜燃料电池技术

HT-PEMFC 的核心部件是基于磷酸(PA)掺杂的质子交换膜(HT-PEM)和催化层组装的膜电极(MEA),这对电池的输出性能和长期稳定性有直接影响[9-10]。在 PA 掺杂型 HT-PEM 中,通常聚合物高分子链中主链骨架提供膜材料的力学性能,碱性/弱碱性基团提供 PA 吸附位点。PA 分子以结合和游离 2 种形式存在于聚合物高分子链的网络结构中,其中结合形式的 PA 以强氢键作用"锚定"在聚合物高分子链周围,而游离形式的 PA(也称"自由 PA")与聚合物分子相互作用较弱[11],主要在 PA 分子间形成氢键网络,氢键网络是高温质子传输的主要来源。质子通过氢键网络以 Grotthuss 机制进行传输,即质子跳跃机制。膜内的质子供体和质子受体之间可形成氢键,并互相连接构成氢键网络,在该网络中,质子沿着氢键在载体分子间移动,在相邻分子间每完成一次跳跃,随后都发生载体分子的重排,即氢键的形成和

收稿日期: 2022-07-05; **录用日期**: 2022-07-28; **网络出版时间**: 2022-08-16 15:12
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220815.1633.001.html
* **通信作者**. E-mail: guozhibin@heracles-tec.com

引用格式: *相艳, 李文, 郭志斌, 等. 磷酸掺杂型高温质子交换膜燃料电池关键材料研究进展[J]. 北京航空航天大学学报, 2022, 48(9): 1791-1805. XIANG Y, LI W, GUO Z B, et al. Research progress on key materials of phosphoric acid doped high-temperature proton exchange membrane fuel cells[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1791-1805(in Chinese).*

断裂,这种不断重排过程形成质子的传输通道。因此,较高的游离 PA 掺杂能够获得较高的质子电导率[12]。然而较高 PA 掺杂水平会导致膜中大量的游离 PA 分子填充在聚合物高分子链的周围,减弱了分子链之间的范德华力,导致膜的尺寸稳定性大幅降低[13-14],在实际应用过程中面临质子传输性能与机械强度难以协同兼顾的严峻问题。对于催化层,在构建高温膜电极(HT-MEA)时,在组装压力和催化层微结构毛细力作用下,PA 从电解质膜内迁移进入催化层中,并在电流驱动下在催化层中进行再分布,形成电化学反应三相界面。然而,PA 在 MEA 内的动态迁移过程使得 PA 填充在催化层微孔结构中,造成反应气体与催化剂的接触受阻,增大了物质传输电阻[15];并且由于 PA 及 PA 解离的阴离子吸附在铂基催化剂表面,降低了催化剂反应活性[16-17]。

近年来,人们对 HT-PEMFC 存在的上述科学问题取得一定进展,但实现 HT-PEMFC 保持优异的电化学性能、技术国产化及产品商业化发展仍然面临着巨大挑战。人们希望通过廉价膜材料的选择以降低使用成本,聚合物高分子结构设计以优化膜材料性能,大尺寸 MEA 的一致性工艺与性能优化,HT-PEMFC 与甲醇重整器等耦合联用以拓宽燃料气体选择范围等途径,有效解决高性能 HT-PEMFC 的技术开发与实用化。

本文在梳理 HT-PEMFC 关键科学问题的基础上,着重针对北京航空航天大学(简称北航)相艳教授团队近十年来在 HT-PEMFC 关键材料的相关研究进展方面进行了详细阐述。如图 1 所示,主要包括:发展了 3 代不同体系的 HT-PEM 材料;调控了催化层微观结构,实现其稳定化运行;制备了大尺寸 MEA 并对其一致性工艺进行了深入研究等工作。

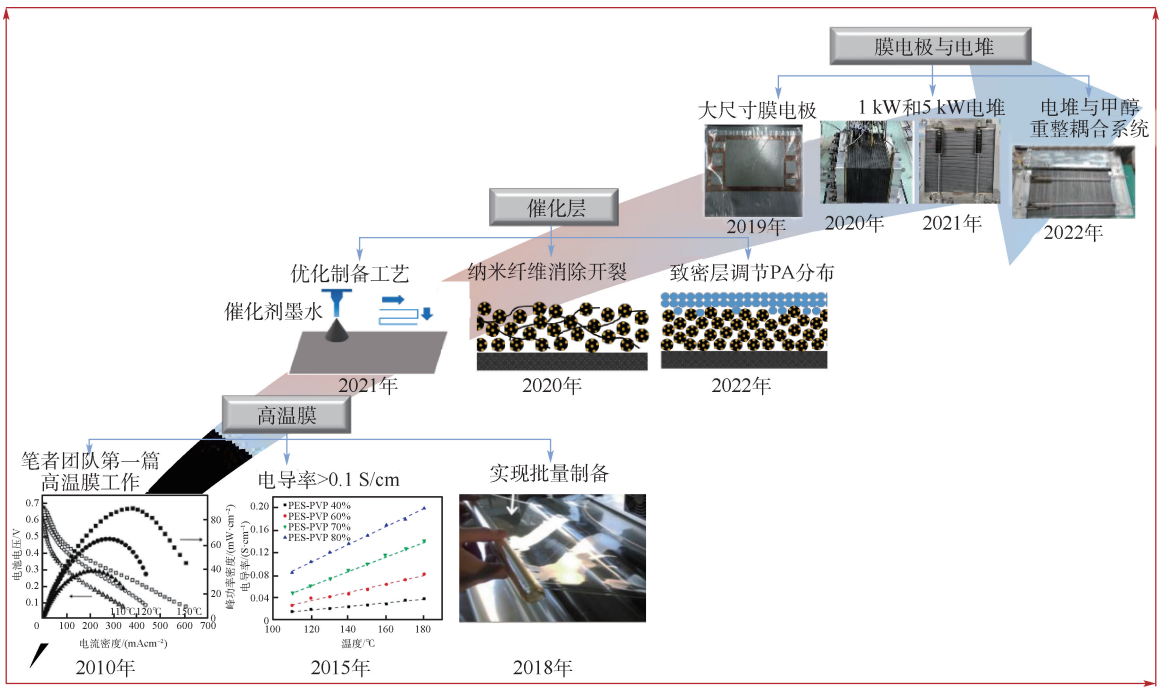


图 1 北航相艳教授团队在高温膜、催化层和膜电极与电堆的研究发展历程

Fig. 1 Research and development history of Professor Xiang Yan's team of Beihang University in high-temperature membrane, catalytic layer, membrane electrode and stack

2 高温质子交换膜

在众多高温膜材料体系中,PA 掺杂型 PEM 由于在高温低湿环境下具有较高的质子传导率而被认为是最具发展应用前景的高温聚合物膜材料[18-22]。以早期研究中典型的 PA 掺杂聚苯并咪唑(PA/PBI)为例,其骨架上的碱性咪唑基团在掺杂 PA 后,咪唑基团与 PA 之间形成连续氢键,增强了质子传导[23]。然而,由于 PBI 材料合成工艺复杂,成本较高,限制了其大规模应用。因此,开发更高效且低成本的膜材料对于 HT-PEMFC 的大规模应用与发展尤为重要。

笔者研究团队自 2010 年前后开展高性能低成本的聚合物膜材料的研究工作,从膜基材的选择、微观结构的调控及质子导体的筛选 3 个方面入手,发展了 3 代新型 HT-PEMs 材料(见图 2)。第 1 代:聚乙烯吡咯烷酮(PVP)基 HT-PEM。PVP 具有优异的溶解性及热稳定性能,且具有独特的

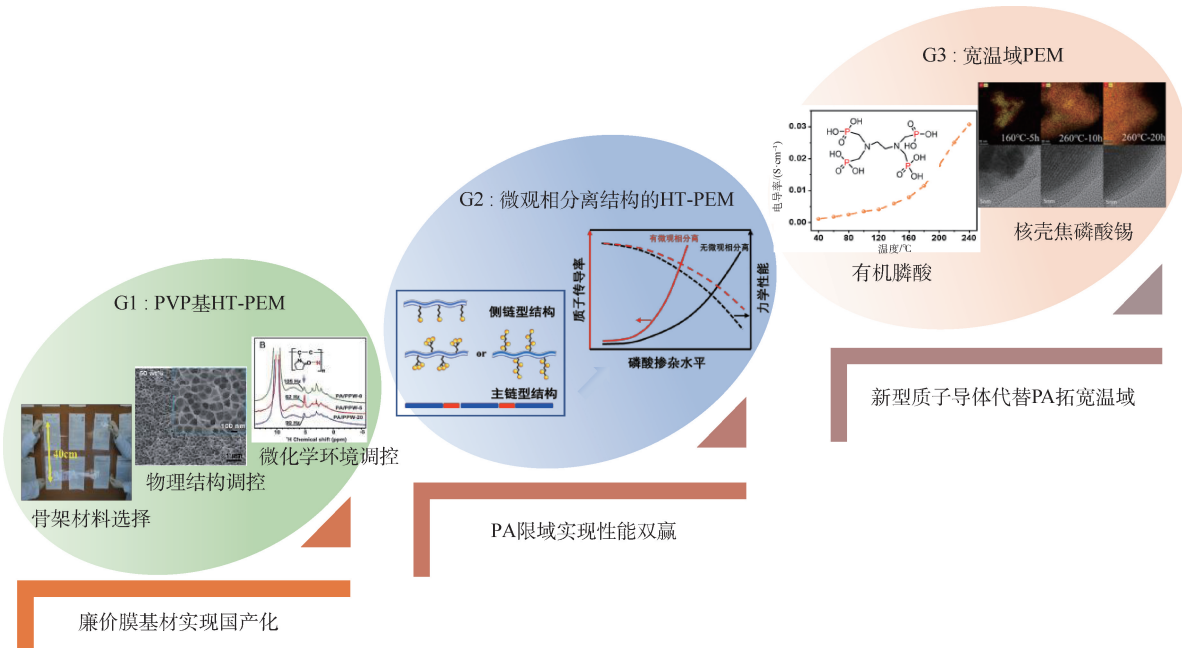


图 2　3 代不同 HT-PEM 的设计理念与研究进展

Fig. 2　Design concepts and research progress of three generations of different HT-PEM

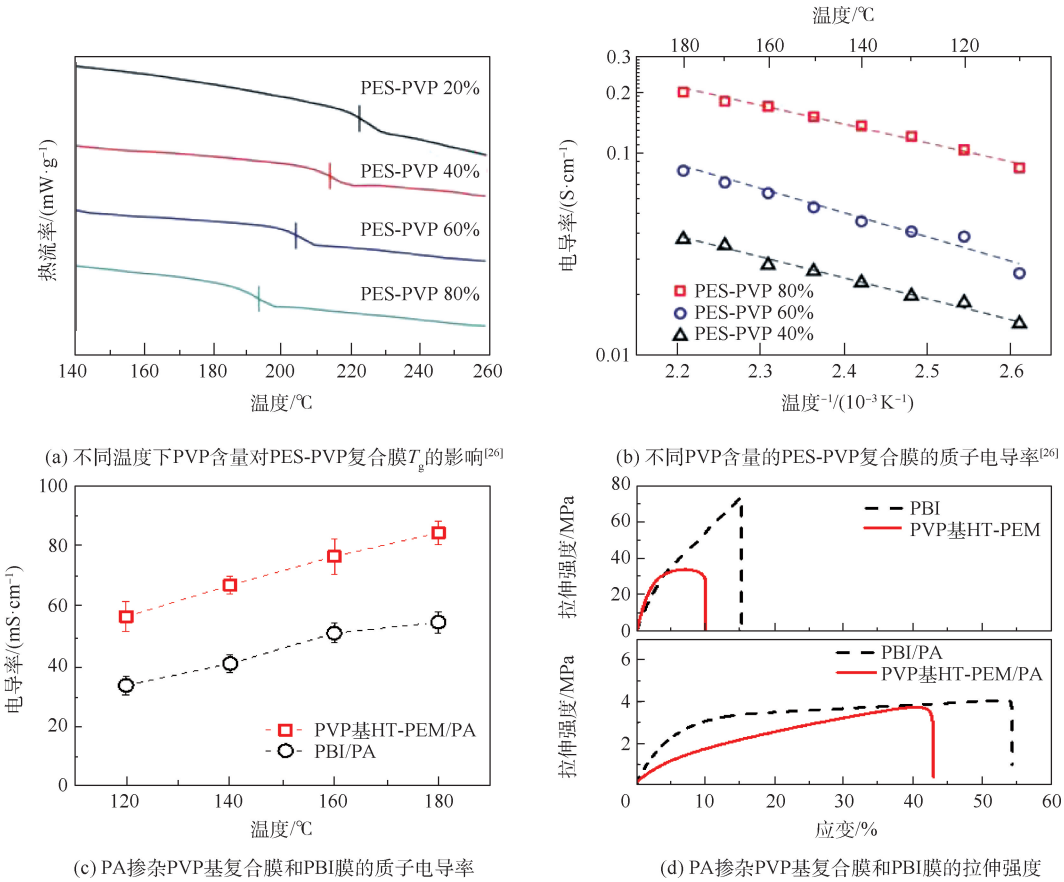
分子单元结构,PVP 含有的吡咯烷酮环上的碱性 N 和 O 原子均可在酸性环境中被质子化,类似于 PBI 上咪唑基团可以吸收 PA,这为 PVP 制备高性能 PEM 提供了理论前提条件[24];第 2 代:具有微观相分离结构的 HT-PEM。无论是 PBI 基或 PVP 基,其聚合物的传导功能基团均在主链或者相连主链上,因此 PA 掺杂膜面临质子传导率和拉伸强度之间的最佳平衡点的问题。受商业化 Nafion 聚合物独特的高分子链结构启发:侧链亲水的离子团簇提供质子传输通道,憎水的主链形成膜的骨架。从分子工程角度对高分子链进行源头设计,改变高分子链上吸附 PA 功能基团的位置及个数,实现 PA 限域聚集分布,形成类似于 Nafion 膜的微相分离结构,在不损失质子传导率的同时降低 PA 对主链的塑化作用,兼顾优异的质子传导率与拉伸强度。第 3 代:宽温域 PEM。在前两代 PEM 的研究中,均使用 PA 作为质子导体,传质性能非常依赖 PA 掺杂量,而 PA 与聚合物基体的相互作用较弱,在电池运行过程中易流失,在室温条件下难以启动,造成性能的衰减。如能开发非 PA 型固态质子导体,可彻底解决 PA 流失问题,并可将电池操作温度拓宽至低温实现低温启动,拓宽工作温度范围。因此,笔者研究团队基于质子导体开发了第 3 代新型 PEM,以期使用新型固态质子导体(如有机膦酸、核壳焦磷酸锡)代替 PA 进行质子传输,解决 PA 流失对电池带来的负面影响,实现高性能 HT-PEMFC 的开发与应用。本节将逐一详细介绍3 代HT-PEM 的研究进展。

2.1　PVP 基 HT-PEM

在 PVP 基 HT-PEM 中,PVP 的碱性 N 杂环可用作 PA 吸附位点,质子在 PA 与 N 杂环之间传递或 PA 分子间的传递以实现质子传输[25-26]。但由于 PVP 脆性较大,一般不能单独形成有用的膜材料。因此,笔者研究团队通过将 PVP 与其他具有良好机械性能的骨架材料进行复合,并通过调控物理结构及微化学环境,制备出了高质子传导率与优异机械性能的 PVP 基 HT-PEM 复合材料。

2.1.1　PVP 基 PEM 骨架材料的选择

聚偏氟乙烯(PVDF)作为一种化学稳定性优异,价格低廉的热塑性材料,笔者研究团队将其与 PVP 混合制备了 PVDF-PVP 复合膜[27],PVP 质量分数为 80 % 时,复合膜质子电导率在 180℃ 可高达 9×10^{-2} S/cm。然而由于 PVDF 较低的玻璃化转变温度(T_g),当 PVDF 含量为 40% 时复合膜的 T_g 仅为 120℃,严重限制了高温燃料电池的优势。针对这一问题,本文采用 T_g 为 230℃ 且化学稳定性良好的聚醚砜(PES)替代 PVDF 与 PVP 共混,制备了 PES-PVP 高温膜[26]。当 PVP 质量分数为 80% 时即 PES-PVP 80%,复合膜的 T_g 仍可高达 192℃(见图 3(a)),极大提高了复合膜应用温度范围。此外,PVDF 与 PVP 之间的氢键作用占据了部分 PA 吸附位点,而 PES 与 PVP 之间没有相互作用。因此,PVP 的吡咯环中除 N 原子外,羰基氧也被释放出来吸附 PA,使得 PES-PVP 膜的 PA 吸附量(PA 掺杂水平为 9.1)远高于同样 80% PVP



(a) 不同温度下PVP含量对PES-PVP复合膜T_g的影响[26]

(b) 不同PVP含量的PES-PVP复合膜的质子电导率[26]

(c) PA掺杂PVP基复合膜和PBI膜的质子电导率

(d) PA掺杂PVP基复合膜和PBI膜的拉伸强度

图3 PES-PVP膜的T_g和质子传导率及PVP基膜和PBI膜的性能对比

Fig. 3 T_g and proton conductivity of PES-PVP membrane and performance comparison between PVP-based membrane and PBI membrane

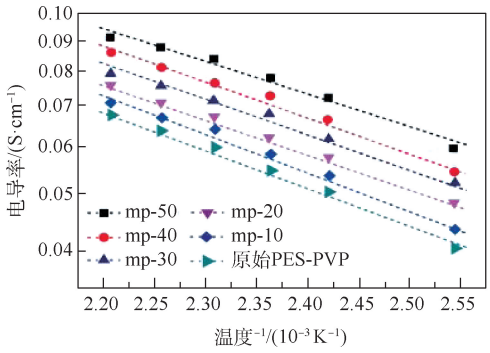
含量下的PVDF-PVP复合膜(PA掺杂水平为2.7),在180℃质子电导率达0.21 S/cm(见图3(b)),并获得850 mW/cm^{2}峰功率密度。以上研究表明不同复合材料对PVP基膜的物理化学性质有重要影响。Ren等[28]证实了与掺有脂肪族聚合物相比,芳族聚合物(如PES和PSU等)与PVP制备的复合膜表现出更小的体积溶胀及更高的酸掺杂量,获得高电导率的同时具有良好的机械性能,更适合应用于PEM。

此外,笔者研究团队还以HT-PEM商业化应用为目标,成功使用流延法放大制备了小批量幅宽为40 cm的PVP基复合膜,实现了国产化应用。该膜中每个PVP结构单元的吸PA水平可达4.9,质子传导率在180℃达85 mS/cm(见图3(c)),优于同等条件下商业化PBI/PA膜,掺杂PA后的机械性能与PBI/PA膜相当(见图3(d))。

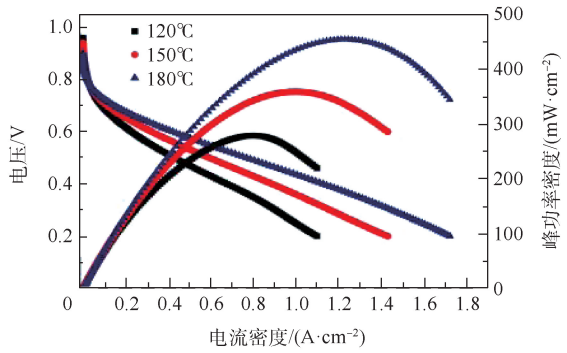
增加PVP含量可有效提高膜的PA掺杂水平,提高电导率,但也会严重影响膜的机械性能[26]。有文献表明,多孔聚四氟乙烯(PTFE)作为增强材料可显著提高拉伸强度[29]。笔者研究团队在PES-PVP膜的基础上使用同时具备亲疏水结构的Triton表面活性剂作为PES-PVP膜与PTFE之间的界面,提高了二者之间的相容性。掺杂PA后,优化后的复合膜PES-PVP/PTFE-5(PTFE的掺杂量为5%)(6.50 MPa)相较PES-PVP膜(1.54 MPa)拉伸强度得到明显提升,质子电导率在180℃达0.26 S/cm,并获得607 mW/cm^{2}的峰功率密度,且运行过程中开路电压无明显下降[30]。

2.1.2 物理结构与微化学环境调控

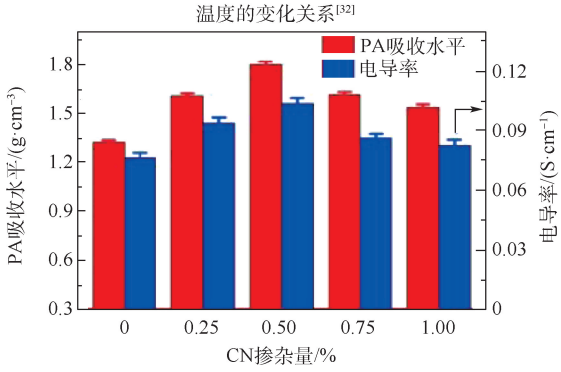
为了进一步提高质子传导率,在不破坏聚合物主链结构的情况下对PEM结构进行调控,提高局部自由PA含量和自由PA之间的连通性,同时也减弱PA对主链的塑化作用,加强质子传递的同时保证膜具有良好的机械强度。研究表明,在PEM中引入多孔结构有利于负载大量自由PA[31]。笔者研究团队基于PES-PVP复合膜,以单分散的SiO_2固体球为硬模板制备了含亚微孔结构的PES-PVP/SiO_2膜[32]。获得的多孔结构显著提高了局部区域的自由PA含量和自由PA之间的连通性,在优化的孔结构下SiO_2含量为50%的多孔膜(mp-50)质子传导率在180 ℃可高达0.09 S/cm(见图4(a)),峰功率密度为454 mW/cm^2(见图4(b))。



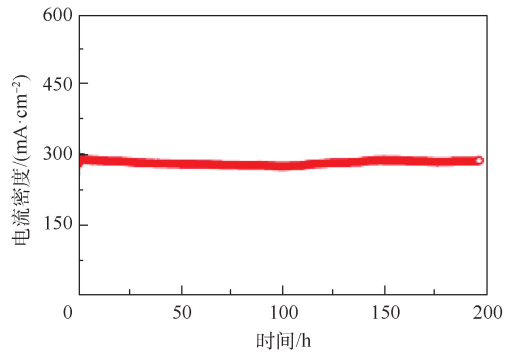
(a) 不同SiO_2含量制备的多孔膜质子传导率随温度的变化关系[32]



(b) 不同温度下mp-50多孔膜的单电池极化曲线[32]



(c) 不同CN含量复合膜的PA掺杂水平和质子传导率[35]



(d) 基于P/CN-0.5/PA复合膜的单电池耐久性测试(温度：160℃，电压：0.6 V)[35]

图 4　以 SiO_2 为模板制备的 PES-PVP 多孔膜及 g-C_3N_4 掺杂 PES-PVP 复合膜的质子电导率和燃料电池性能

Fig. 4　Proton conductivity and fuel cell performance of PES-PVP porous membrane prepared with SiO_2 as template and PES-PVP composite membrane doped with g-C_3N_4

有机-无机共混作为提高 PEM 尺寸稳定性及机械性能的方式之一，因制备方式简单受到广泛应用[12,33-34]。笔者研究团队将具有二维片层结构的石墨相碳氮化合物(g-C_3N_4)纳米片(CN)引入 PVP 基 PEM 中[35]，利用 CN 中丰富的含氮功能基团(—NH 和 C—N—C)提高了 PA 掺杂量(见图 4(c))。此外，^{1}H MAS NMR 结果表明，CN 能够促进 PA 以及 PVP 上的质子解离，这归因于 CN 的碱性(解离常数 pK_a = 8.95)强于吡咯烷酮环碱性(pK_a = 7.6)，更易与 PA 形成氢键，优化了质子传递环境。优化的复合膜 P/CN-0.5/PA(P 代表 PVP 基膜，CN 代表 g-C_3N_4，0.5 为 CN 的含量为 0.5%)在 160℃ 的质子传导率达 104 mS/cm，拉伸强度为 6.0 MPa，分别比 PES-PVP 膜提升了 36% 和 60%，并获得优异的恒压放电稳定性(见图 4(d))。除了物理共混，还通过化学交联制备了 PVP 与氯甲基聚砜的交联膜[36]，利用交联生成的季铵基团增强了膜的 PA 吸附能力，在 160℃ 获得了 120 mS/cm 高质子传导率。

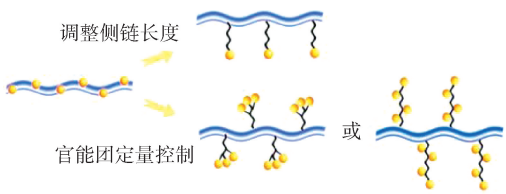
在众多无机材料中，磷钨酸(HPW)作为最具吸引力的无机固体质子导体具有优异的稳定性及质子传导性能，在室温和 100 % 湿度(RH)下获得高达 0.18 S/cm 电导率，被广泛引入到聚合物基质中进行膜修饰或制备新型 PEM[37-42]。笔者研究团队在前期工作中将 HPW 负载到介孔二氧化硅 MCM-41 中，并成功将其用作燃料电池的无机固体电解质，高效的微孔限域结构使得 HPW 纳米团簇构筑了有效的质子传输途径，最终获得优异的 H_2-O_2 燃料电池性能及甲醇燃料电池性能[43]。基于此，进一步探究了 HPW 与 PVP 膜基质的作用机制，PVP 吡咯环的 N 原子可与 HPW 上的质子作用生成了质子化 NR_3H^+，从而将 HPW 锚定在聚合物网络结构中，构建质子传输通道。在 60℃ 及全湿状态下，质子电导率达到 0.066 S/cm，与同样测试条件下的 Nafion 115 质子电导率相当[39]。此后，本文将 HPW 用于 HT-PEM-FC[44]，^{1}H NMR 结果表明掺有 HPW 的复合膜比无 HPW 掺杂膜表现出更高的可移动质子含量，显示出更窄的线宽信号，表明该复合膜中聚合物链更易移动。在 160℃ 获得 1.44×10^{-1} S/cm 高质子传导率，单电池的峰值功率密度为416 mW/cm^2。PA 作为一种中强酸在膜中难以完全解离，膜内质子浓度远低于 PA 浓度，使得 PA 利用率较低。综上研究可知，可以通过复合其他无机材料优化膜

内质子传输环境,提高 PA 解离度,获得更多可移动自由质子,进而提高质子传导率。

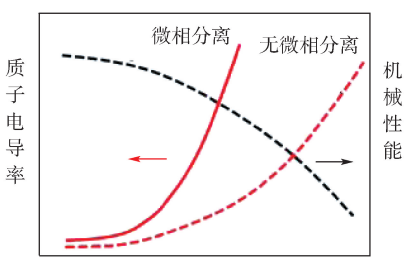
2.2 微相分离结构 HT-PEM

通过 2.1 节研究中的物理增强、化学交联等方式虽然很大程度可以改善膜材料的质子传导性能与机械性能,但同时也会带来一系列新的问题:纤维增强的复合膜吸收 PA 后易出现增强纤维与膜基体剥离的问题;多孔结构的膜存在 PA 流失速度快、稳定性差的问题。因此,解决 HT-PEM 目前所面临的困境,必须从膜材料的化学结构出发,对聚合物高分子链进行源头设计。笔者研究团队针对高分子链进行了精准设计,改变可吸附 PA 的功能基团的位置和个数(见图 5(a)),制备了一系列具有微相分离结构的新型膜材料,实现 PEM 质子传导率和机械性能双赢(见图 5(b))。

通过调控聚苯醚聚合物膜(PPO-PIm-Cx)的末端咪唑官能化的侧链长度调控膜的 PA 吸收水平,在侧链亚甲基个数为 6 的 PPO-PIm-C6 膜中,膜内形成明显的微观相分离结构,团簇尺寸约为 5.56 nm,在 160℃ 获得质子传导率 0.042 S/cm,以及良好的拉伸强度 12.1 MPa。随后,在另一项工作中探究了侧链碱性基团数量对质子传导率和机械性能的影响,将具有 3 个叔胺基的 2,4,6-三(二甲胺甲基)-苯酚(TDAP)接枝到 PSU(TDAP-PSU)上,与单叔胺接枝的 PSU 膜(DMA-PSU)相比,1 个接枝位点上的胺基束明显减少了体积膨胀。接枝度 75% 的 TDAP-PSU 膜与接枝度 99% 的 DMA-PSU 膜的 PA 吸收量相当,但前者的机械强度却是后者的 2.2 倍。基于 TDAP-PSU 膜的单电池峰功率密度达到 453 mW/cm^2,且获得优异的稳定



(a) 聚合物高分子结构设计示意图[9]



(b) 质子传导率和拉伸强度随PA掺杂水平示意图[9]

图 5 微相分离结构质子交换膜的设计与性能示意图

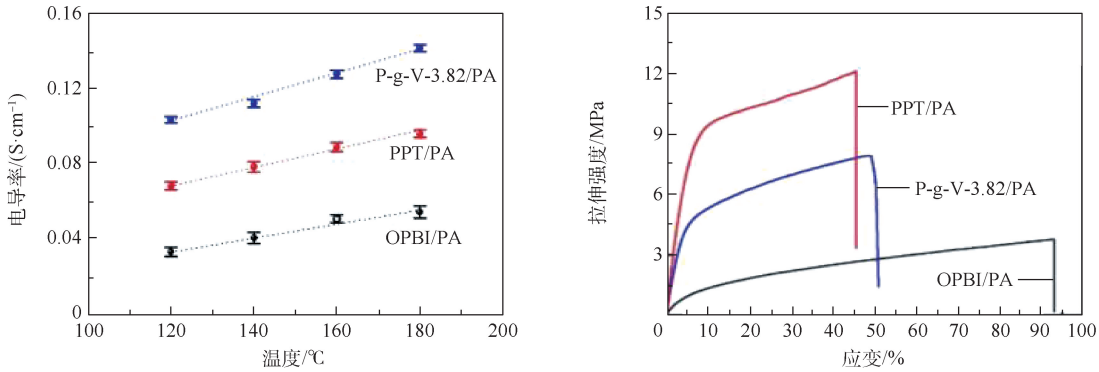
Fig. 5 Schematic diagram of design and performance of PEM with microphase separation structure

性[45]。此外,本文通过原子转移自由基聚合将聚 1-乙烯基咪唑(P-g-V)作为 PA 掺杂位点接枝到聚砜骨架上,获得系列具有不同侧链长度和功能基团数量的聚合物材料[46],其中 P-g-V-3.82/PA(每个接枝位点平均接枝 3.82 个基团)复合膜在 160℃ 的质子电导率为127 mS/cm,膜的拉伸强度保持为 7.94 MPa(见图 6(a)),远高于相同测试条件下 PBI/PA 膜。优异的性能归因于 P-g-V-3.82膜的功能性侧链自组装形成微相分离结构,如图 6(d)所示的原子力显微镜(AFM)图像显示膜内出现侧链的亲水微相(暗区)和聚合物主链的疏水微相(明亮区),进而实现了质子快速传输并维持了良好的机械性能。

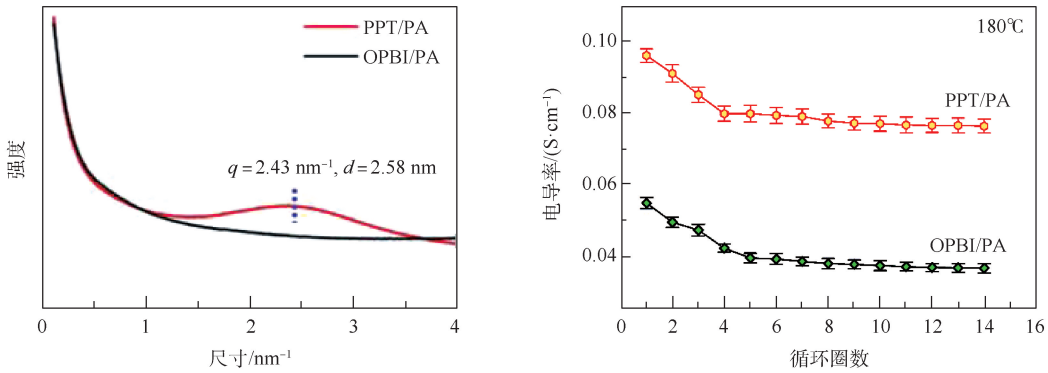
然而,侧链型 HT-PEMs 的热稳定性和化学稳定性仍有待提高,一般而言,类似 PBI 主链型聚合物的热稳定性要优于侧链结构的聚合物。若结合主链型聚合物优异的热稳定性及侧链型微相分离结构,设计一类掺杂 PA 后可以形成微相分离结构的主链型聚合物,则可以更好地兼顾质子传导性、拉伸强度以及热稳定性的问题。因此,本文研究团队制备了基于刚性疏水芳基骨架和哌啶基团的聚亚芳基哌啶(PPT)[47]。通过 AFM(见图 6(d))和小角 X 射线散射(SAXS)(见图 6(b))测试发现,OPBI 膜在掺杂 PA 前后均未形成微相分离结构。而 PPT/PA 膜在 AFM 图像中出现明显的微相分离结构区域(见图 6(d)),其中暗区和亮区分别代表来自哌啶环基的亲水微相和来自聚合物骨架的疏水微相,且 SAXS 数据显示在 PPT 膜内成功形成了 PA 聚集的离子团簇(见图 6(b))。良好的微相分离结构使得 PA 更易聚集在哌啶环周围,与憎水的苯环分离,降低了 PA 对分子链的塑化作用,获得了 12 MPa 的优异拉伸强度,在 180℃ 无外部加湿条件下显示出 96.0 mS/cm 的高质子电导率(见图 6(a)),且在质子传导率循环稳定性测试中 PPT/PA 膜始终优于 OPBI/PA 膜(见图 6(b))。在 180 ℃ 基于 PPT/PA 膜的 H_2-O_2 燃料电池的最高输出功率高达 1 220.2 mW/cm^2,是同等测试条件下 OPBI/PA 膜的单电池输出性能的 1.85 倍(660.8 mW/cm^2),且在 1 600 h 的耐久性测试中保持了优异的输出稳定性(见图 6(c))。

2.3 宽温域 PEM

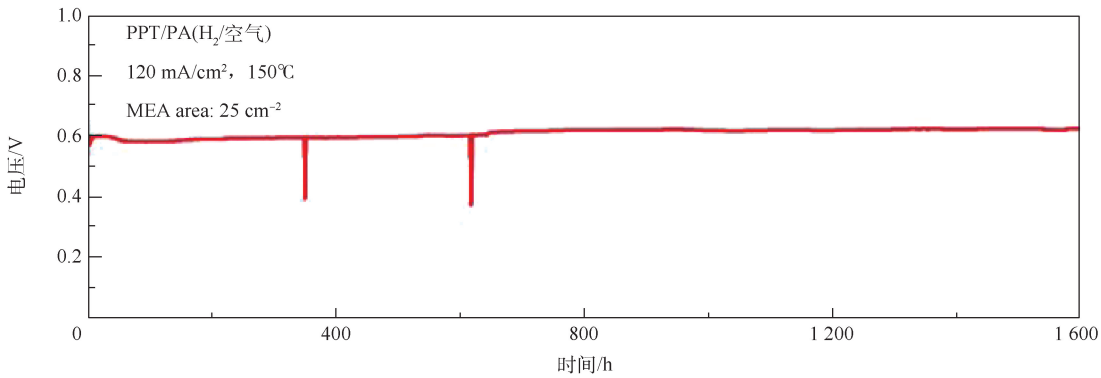
PA 掺杂 PEM 的质子传递性与膜内 PA 掺杂量有直接关系,然而小分子 PA 易受到各种因素影响而流失[48-49],导致无法室温启动等问题[18]。尽管前人已做了很多努力,笔者研究团队也已发展了 2 代新型的 HT-PEM,但对于 PA 流失问题仍



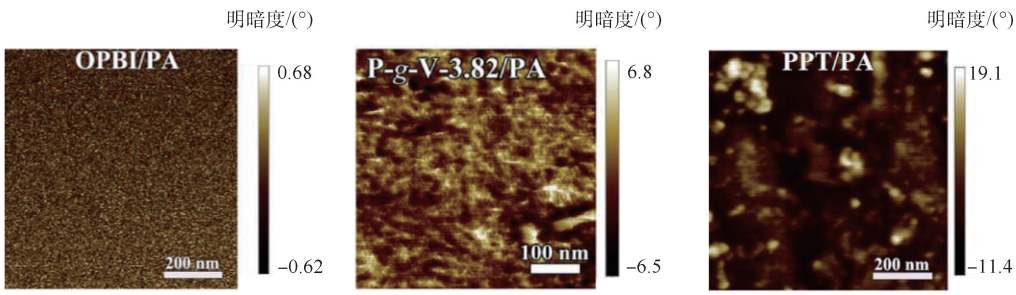
(a) P-g-V-3.82/PA膜、PPT/PA膜及OPBI/PA膜的质子传导率及拉伸强度



(b) OPBI/PA和PPT/PA的SAXS曲线及电导率随温度循环变化情况[9,47]



(c) 基于PPT/PA膜的单电池的输出稳定性[47]



(d) 不同膜的AFM图像[9,47]

图 6　P-g-V-3.82/PA 膜、PPT/PA 膜及 OPBI/PA 膜的微相分离结构及性能对比

Fig. 6　Comparison of microphase separation structure and performance of P-g-V-3.82/PA membrane, PPT/PA membrane and OPBI/PA membrane

没有彻底解决。因此，笔者研究团队从质子导体角度出发，拟发展新型质子导体代替 PA 进行质子传输，进一步拓宽 PEMFC 的操作温度范围：拓宽低温区实现电池的室温启动，拓宽高温区可将电池与甲醇重整器等进行耦合联用。

研究证明有机膦酸（OPA）是一类良好的有机质子导体[50]，较大的分子尺寸使其分子动态迁移能力低于 PA，理论上不易流失。已有研究表明，烷基或苯基取代 OPA 在高温无水条件下可获得 10^{-2} ~ 10^{-1} S/cm 的高电导率[51-52]。因此，将

OPA 作为质子导体用于制备新型 HT-PEM，有望解决 PA 掺杂型 HT-PEMFC 面临的困境。为此，笔者研究团队探索了不同 OPA 在 HT-PEM 中的质子传输行为与机制。将乙二胺四亚甲基膦酸（EDTMPA）引入 PVP 膜基质中构建了新型 PEM，EDTMPA 和 PA 双质子导体协同改善了单位体积膜内的有效质子浓度及复合膜质子扩散系数，使得该膜在 180℃ 无水条件下质子电导率可达 0.065 S/cm，约为仅含有单一 PA 质子导体膜的 1.8 倍，并获得 716 mW/cm^2 高峰功率密度及优异的输出稳定性。此外，EDTMPA 对 PA 的锚定作用使得复合膜在 80℃ 同样获得了优异的电导率（0.025 S/cm）及 326 mW/cm^2 峰功率密度，这有望实现电池的低温启动。此外，还发现氨基三亚甲基膦酸质子导体（ATMP）与 OPBI 膜有明显交联作用，这区别于传统仅有交联作用的交联剂。因此，该现象启发本文将兼具质子传导功能和交联作用的 ATMP 引入 OPBI 膜中制备交联膜，在较低 PA 掺杂量下（156%）在 80℃ 获得了 0.033 S/cm和在 160℃ 0.112 S/cm 的高质子传导率，单电池分别在 80℃达到了 0.268 W/cm^2 和在 160℃ 0.98 W/cm^2 的高峰功率密度。除 OPA 质子导体之外，有研究发现，核壳焦磷酸锡（SnP_2O_7）在 200℃ 以上也具有良好的质子电导率[53-55]。笔者研究团队对核壳 SnP_2O_7 在 200℃ 及以上更高温度的质子传导机制进行探究发现，核壳 SnP_2O_7 优异的质子电导率主要来源于最外层的凝胶层，该凝胶层主要成分是焦磷酸，260℃ 无水条件下可获得 0.084 S/cm 的高质子传导率，高温电导率可稳定 170 h 左右，在 200℃ 以上的高温条件下极具质子导体应用潜力。

综上，笔者研究团队共开发了 3 代 PEM，首先开发了低成本 PVP 膜基材，探究了与其他骨架材料形成复合膜的性能，并调控了 PVP 基 PEM 的物理结构和微化学环境以改善膜的质子传导率与机械性能的平衡问题，并实现了 PES-PVP 复合膜国产化批量制备；接着进一步通过高分子链设计制备了具有微相分离结构的 PEM，将吸附 PA 的侧链功能基团与提供机械性能的主链分离开来，实现质子传导率与拉伸强度的双赢；以上 2 种策略均是基于开发新型膜材料且均使用 PA 作为质子导体，而 PA 带来的一系列问题仍不可避免，因此，非 PA 型质子导体的开发对于 PEMFC 的实际应用至关重要，我们探究了新型质子导体用于 PEM，并对其在宽温域的传导机制进行研究，进一步实现 PEMFC 的实用化发展。

3 催化层

在 PA 掺杂 PEM 中，PA 受各种因素在 MEA 内进行分布与迁移[9,56-59]。如图 7 所示，在电池运行过程中，膜内 PA 受到电流驱动，PA 水解的 $H_2PO_4^-$ 从膜内向阳极迁移，与阳极侧的氢离子反应形成 PA，提高了阳极 PA 含量。随后，由于浓度梯度差使得 PA 扩散到阴极，与阴极产物水结合形成 $H_2PO_4^-$，此循环过程促进了 PA 在催化层中的再分布，导致催化层出现“酸淹”现象，大量 PA 以及 PA 解离的负离子占据铂催化剂的活性位点，使得催化层内的物质传输阻力增大。因此，笔者研究团队拟通过从催化层体相和界面等角度对 PA 在 MEA 中的分布进行适当调控，以期缓解 PA 对 MEA 造成的不利影响，提高燃料电池性能。

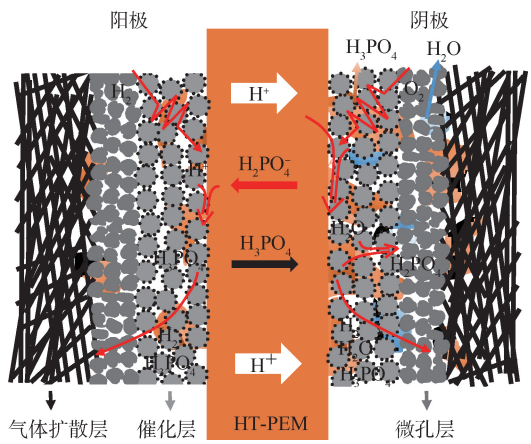
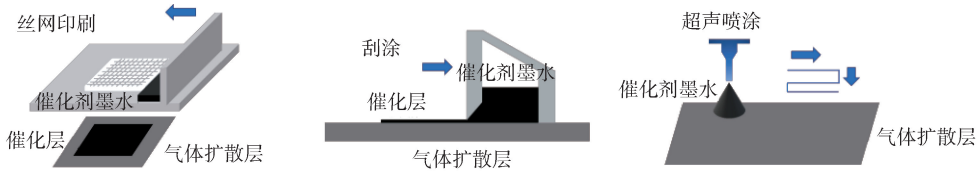


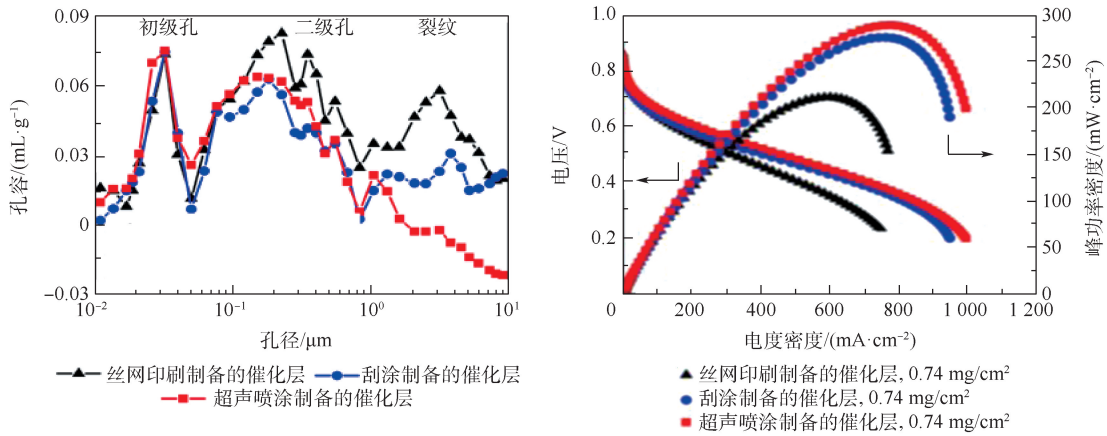
图 7 磷酸在膜电极中的分布与迁移示意图[9]

Fig. 7 Distribution and migration of phosphoric acid in membrane electrode[9]

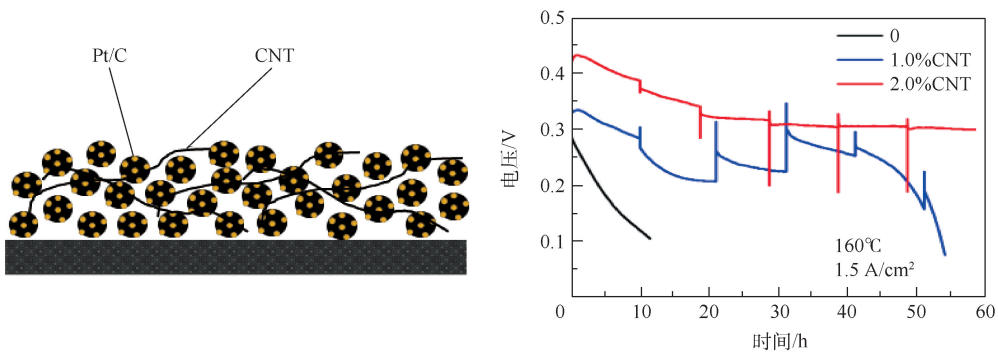
在催化层制备过程中，催化层的结构受到催化剂墨水中使用的溶剂沸点，黏结剂含量及制备方法等影响，导致在催化层中形成不同尺寸的裂纹[60-62]。较大的裂纹会降低催化层的毛细力作用，使 PA 快速迁移进入催化层内部，形成“酸淹”[63]。Halter 等[64]发现 PA 充填裂纹的毛细压力随着裂纹宽度的减小而增大，裂纹的连通性和可及性对 PA 的保留能力起主要作用。而催化层内的 PA 压力等物理特性易受到制备方法的影响。因此，笔者研究团队研究了超声喷涂、丝网印刷、刮涂 3 种方式构建的具有不同微观结构的催化层（见图 8(a)），探究了制备方法对催化层微观结构的影响[15]。结果表明，超声喷涂制备的催化层表面平坦，裂纹少且互相不连通，具有良好的二级孔分布，促进了催化层内 PA 的均匀分布，构建了更多电化学反应三相界面；而丝网印刷和刮涂制备的催化层表面分别显示出 20 ~ 30 μm 和



(a) 催化层的3种制备方法示意图[15]



(b) 不同裂纹程度催化层的孔径分布以及基于0.74 mg/cm² Pt载量的膜电极性能[15]



(c) 添加CNT减少催化层开裂的示意图以及不同CNT含量催化层组装的HT-MEA每10 h启停的稳定性测试[67]

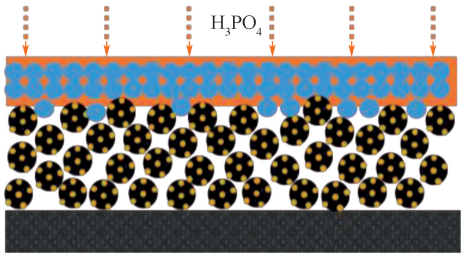
图 8　改善催化层裂纹对膜电极性能的影响

Fig. 8　Effect of improving catalytic layer cracks on membrane electrode performance

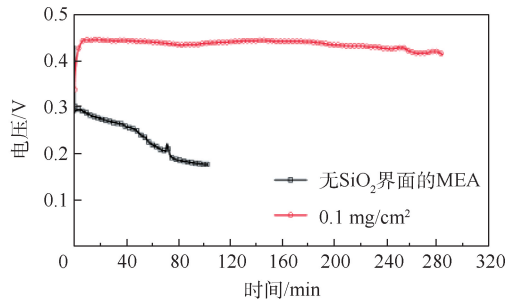
10 ~ 15 μm 的相互连通的明显裂纹，表现出较发达的二级孔结构（见图 8(b)）。基于超声喷涂制备的催化层 HT-PEMFC 的峰功率密度高于其他 2 种方式制备的催化层（见图 8(b)），且保持了优异的稳定性。因此，减少催化层裂纹的生成，降低裂纹程度，能够有效避免催化层内 PA 的不均匀分布，降低极化过程中的物质传输损失。

除了制备方法的优化，本文利用具有强韧性和高长径比的一维柔性碳纳米管[65-66]（CNT）进一步消除溶剂蒸发过程中催化剂团聚体间的拉伸应力（见图 8(c)），达到降低催化层裂纹数量和宽度的目的[67]。同时，CNT 高比表面积也提高了催化层内的二级孔含量，促进 PA 在催化层内的均匀分布，致使电化学反应动力学得到提升，物质传输阻力降低。当催化层中加入质量分数为 2.0% 的 CNT 时，电池峰值功率密度为 673 mW/cm²，比不含 CNT 的电池高 1.5 倍，且在多次启停过程中表现出良好的稳定性。

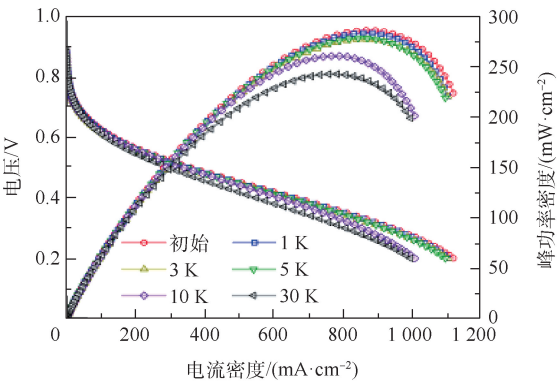
尽管加入一维 CNT 材料可以抑制催化层内裂纹的生成及增加催化层二级孔含量，达到降低物质传输阻力和提高 Pt 利用率的目的。然而催化层的二级孔孔径范围跨度较大且大小不一，影响了催化层内 PA 含量和分布，并且催化层内相互贯通的孔结构增加了 PA 流失的风险。因此，为了降低靠近膜侧催化层的 PA 含量，笔者研究团队构建了具有更致密孔径结构的界面层，在催化层与 PEM 之间构建 SiO_2 界面（见图 9(a)），利用纳米 SiO_2 较小粒径（10 ~ 20 nm）、与 PA 作用力强以及能与 PA 在高温反应的特性，有效地降低了催化层中 PA 的初始含量，减缓了 PA 向催化层的迁移速率，提高了 HT-PEMFC 的输出功率和稳定性。基于 SiO_2 界面层组装的 HT-MEA 即使在 30℃，0.3 A/cm² 下也能够稳定运行超过 280 min（见图 9(b)），在 30 K 圈电势循环的加速老化测试过程中，其峰功率密度仅衰减 15.4%（见图 9(c)），远低于没有 SiO_2 界面层组装的高



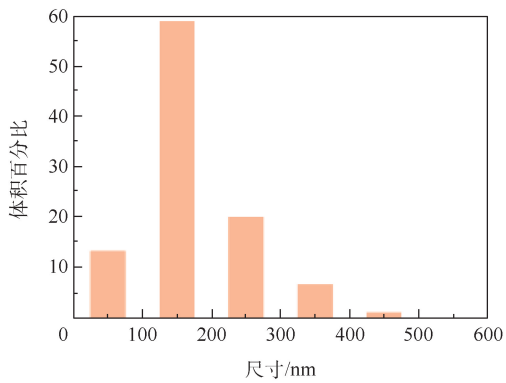
(a) 催化层与PEM之间构建界面示意图



(b) 有/无SiO_2界面层的MEA的稳定性



(c) 0.1 mg/cm² SiO_2界面层的MEA在加速老化后的极化曲线和功率密度曲线



(d) 加速老化后催化剂团聚体尺寸分布

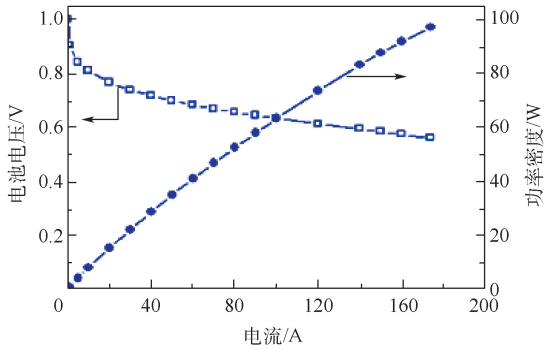
图9 SiO_2 界面层对MEA性能的影响

Fig. 9 Influence of SiO_2 interface layer on MEA performance

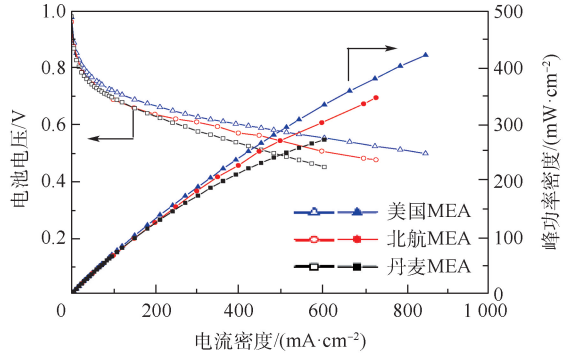
温MEA(56.6%)。通过对老化后催化层3D结构重建分析,由于PA向催化层迁移量降低,催化剂团聚体尺寸仅发生轻微的变化(见图9(d)),合理的致密界面层有效改善了MEA的输出性能及稳定性。

4 高温膜电极和燃料电池堆

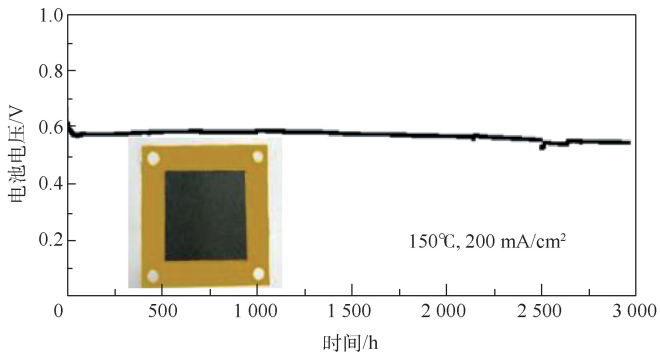
随着PVP基复合膜的批量制备工艺逐渐成熟,本文将其与Pt/C催化剂电极组装成不同尺寸的MEA,进一步考察了其在HT-PEMFC中的实际应用。分别制备了25,200,400 cm^2 的大尺寸MEA。图10(a)展示了基于MEA尺寸为200 cm^2 的PVP基PEM/PA燃料电池在H_2-O_2环境下的极化曲线,电池表现出1.01 V高开路电压,表明膜极低的气体渗透率,且展示出优异的输出性能,在0.6 V下的输出电流密度为0.62 A/cm^2。随后,在H_2-空气环境下,由于空气中的氧气分压降低导致电池的开路电压略微降低至0.98 V,0.6 V下的输出电流密度为0.32 A/cm^2。为了评价该MEA的商业化水平,在相同测试条件下对美国Advent公司和丹麦电力系统公司制备的基于PBI/PA的MEA性能进行测试,2种基于PBI/PA膜的MEA电池的开路电压与本研究的PVP基PEM/PA的电池基本一致(0.98 V),0.6 V下的输出电流密度分别为0.39 A/cm^2 及0.25 A/cm^2,而基于PVP基PEM/PA膜的MEA性能处于上述2种商业化MEA之间,表明本研究的大尺寸MEA具有极大的商业化潜力(见图10(b))。MEA的运行稳定性对于燃料电池的实际应用至关重要。在150℃,0.2 A/cm^2 电流密度下对PVP基PEM/PA膜燃料电池进行稳定性测试,如图10(c)所示,电池在长达3 000 h的运行时间内表现出优异的稳定性,电压衰减速率仅为9.7 μV/h,与商业化PBI/PA膜在同等放电条件下性能相当。基于单电池良好的输出性能及稳定性,将其组装成燃料电池电堆(MEA面积为200 cm^2)进行测试。采用多蛇形流场,并对其进行优化,获得更高输出性能以及更均匀的电流密度分布(见图10(e))。如图10(d)所示,具有3片膜电极的电堆具有高的开路电压,证明其具有良好的密封性能。在稳定性测试中,分别在电堆组装完成的第1~17天和搁置100余天后,对电堆进行从室温升至工作温度的启停、输出性能测试及停机氮气吹扫等工序来测试其稳定性。发现在整个测试时长内,电堆的开路电压基本保持在3 V,且输出性能几乎维持在稳定水平。在第121天和第162天的测试中,输出电流密度较初始性能略有降低,这可能是催化层吸潮所致。此外,本文还继续组装了20片



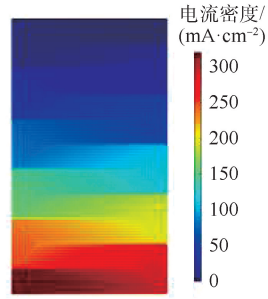
(a) 基于PVP基PEM/PA的燃料电池放电曲线，150℃无加湿/背压，计量比H_2:O_2=3:9



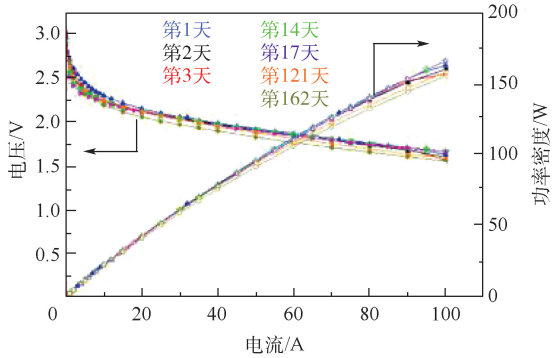
(b) 不同MEA输出性能比较，150℃无加湿/背压，计量比H_2:O_2=2:5，电极面积为165 cm²



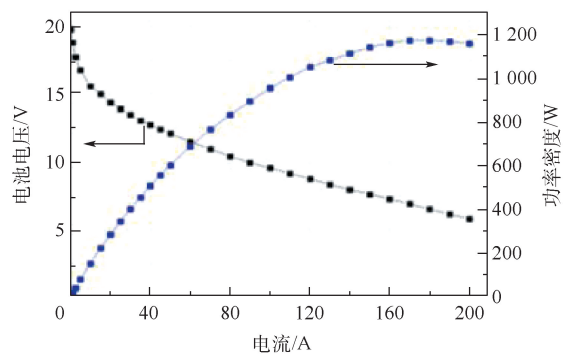
(c) 恒流稳定性测试，150℃，无加湿/背压，计量比H_2:O_2=3:6，电极面积为25 cm²



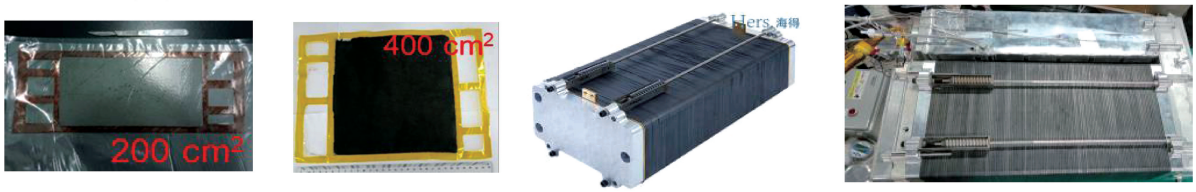
(d) 多蛇形流场电流密度分布云图



(e) 基于3片MEA的电池稳定性测试



(f) 20个MEA电池的极化曲线



(g) 200 cm²和400 cm²的MEA照片；长堆的实物照片；电堆与甲醇重整器耦合的系统照片

图 10　基于 PVP 基 PEM 组装的不同尺寸的膜电极性能及部分器件实物照片

Fig. 10　Performance investigation of membrane electrodes with different sizes assembled based on PVP-based PEM and photos of some devices

大尺寸 MEA 的长电堆，开路电压达到 20 V，150℃下 12 V 下的输出电流达到 45 A，且获得了高达 1.15 kW 的高峰功率密度（见图 10(f)），初步实现了 PVP 基 PEM/PA 燃料电池千瓦级输出性能的目标，并且到目前为止，本文尝试将其与甲醇重整器系统耦合（见图 10(g)），其部分成果已经在北京海得利兹新技术有限公司实现产业化转化。以上研究为 HT-PEMFC 关键材料以及燃料电池的国产化提供了重要的研究基础。

5　结　论

HT-PEMFC 具有较强的杂质气体耐受性、广泛的燃料选择性及可实现高效热电联动等诸多优势，且其发展迅速，目前成为燃料电池未来的主要发展方向之一。本文基于 HT-PEMFC 的关键部件进行了广泛深入的探究。

1）对聚合物膜材料进行开发和探究，分别发展了 PVP 基 PEM、微相分离结构 PEM 及宽温域 PEM。PVP 基 PEM 得益于 PVP 吡咯烷酮环上的官能基团对 PA 的吸附作用，并通过物理化学调控显著提升了膜的酸吸收量及质子传导率，并且初步实现了 PVP 基复合膜的批量制备及国产化应用；进一步通过对聚合物高分子链进行精细设计，构筑了具有微观相分离结构的新型高温膜材料，使得 PA 限域聚集分布，保证质子传导率的同时降低 PA 对主链的塑化作用，实现质子传导率与拉伸强度双赢；此外，PA 掺杂 PEMFC 的 PA 流失问题严重影响其实际应用，通过发展基于新型非 PA 质子导体的宽温域 PEM，利用有机膦酸或核壳焦磷酸锡质子导体代替 PA 进行传质，有望缓解 PA 流失，降低电池启动温度，缩短启动时间，另一方面新型质子导体使得电池工作温度提高，使其与甲醇重整器耦合，提升热利用率和能量转化效率。因此，兼顾质子传导率与拉伸强度的低成本且具有宽工作温域的电解质膜材料是后续研究的重点方向之一。

2）对于催化层，基于 PA 在 MEA 内的迁移和分布，对催化层的微观结构进行系列调控，包括优化催化层的制备方法以及引入纳米材料 CNT 进一步提升性能，还通过在三相界面构筑更致密的 SiO_2 层，调控 PA 分布，缓解 PA 流失。然而，PA 在催化层内的迁移是一个动态过程，目前缺乏更多的原位测试手段，后续研究可以此为方向，进一步明晰其动态迁移过程，为催化层设计提供依据。

3）为了实现商业化应用，对基于 PVP 复合膜的大尺寸 MEA 的制备及一致性工艺进行探究，并成功组装了千瓦级 HT-PEMFC 电堆，展示出优异的电池稳定性。综上系列工作为 HT-PEMFC 关键组成部件的发展提供了切实的参考依据，更高性能膜材料的开发，以及 MEA 的商业化发展仍是 PA 掺杂 PEMFC 未来的重点研究方向之一。

此外，HT-PEMFC 普遍使用的质子导体 PA 对催化剂的毒化作用也是限制 HT-PEMFC 商业化和大规模应用的重要影响因素。因此，开发低成本、高活性的抗 PA 催化剂也是提高 HT-PEMFC 稳定运行的重要方向之一。目前研究表明，通过电子结构调控或物理阻隔等方式可以有效改善催化剂单电极活性，但实现其在 HT-PEMFC 中的高效应用依旧面临挑战。在后续研究中进一步发展可以表征 PA 对催化剂毒化机制的原位表征手段，对于深入指导催化剂的理性设计及开发更有效的抗 PA 毒化策略具有重要的科学意义。

参考文献（References）

[1] HAIDER R, WEN Y C, MA Z F, et al. High temperature proton exchange membrane fuel cells: Progress in advanced materials and key technologies[J]. Chemical Society Reviews, 2021, 50(2): 1138-1187.

[2] CANO Z P, BANHAM D, YE S Y, et al. Batteries and fuel cells for emerging electric vehicle markets[J]. Nature Energy, 2018, 3(4): 279-289.

[3] ZHANG H W, SHEN P K. Advances in the high performance polymer electrolyte membranes for fuel cells[J]. Chemical Society Reviews, 2012, 41(6): 2382-2394.

[4] RIBEIRINHA P, SCHULLER G, BOAVENTURA M, et al. Synergetic integration of a methanol steam reforming cell with a high temperature polymer electrolyte fuel cell[J]. International Journal of Hydrogen Energy, 2017, 42(19): 13902-13912.

[5] CHENG Y, ZHANG J, LU S F, et al. Significantly enhanced performance of direct methanol fuel cells at elevated temperatures[J]. Journal of Power Sources, 2020, 450: 227620.

[6] YAN W R, XIANG Y, ZHANG J, et al. Substantially enhanced power output and durability of direct formic acid fuel cells at elevated temperatures[J]. Advanced Sustainable Systems, 2020, 4(7): 2000065.

[7] CHENG Y, ZHANG J, LU S F, et al. High CO tolerance of new SiO_2 doped phosphoric acid/polybenzimidazole polymer electrolyte membrane fuel cells at high temperatures of 200-250 degrees C[J]. International Journal of Hydrogen Energy, 2018, 43(49): 22487-22499.

[8] ZHANG J, LU S F, ZHU H J, et al. Amino-functionalized mesoporous silica based polyethersulfone-polyvinylpyrrolidone composite membranes for elevated temperature proton exchange membrane fuel cells[J]. RSC Advances, 2016, 6(89): 86575-86585.

[9] ZHANG J J, ZHANG J, WANG H N, et al. Advancement in distribution and control strategy of phosphoric acid in membrane electrode assembly of high-temperature polymer electrolyte membrane fuel cells[J]. Acta Physico-Chimica Sinica, 2021, 37(9): 2010071.

[10] BARRON O, SU H N, LINKOV V, et al. Enhanced performance and stability of high temperature proton exchange membrane fuel cell by incorporating zirconium hydrogen phosphate in catalyst layer[J]. Journal of Power Sources, 2015, 278: 718-724.

[11] LEE H S, ROY A, LANE O, et al. Synthesis and characterization of poly(arylene ether sulfone)-b-polybenzimidazole copolymers for high temperature low humidity proton exchange membrane fuel cells[J]. Polymer, 2008, 49(25): 5387-5396.

[12] AILI D, ZHANG J, JAKOBSEN M T D, et al. Exceptional durability enhancement of PA/PBI based polymer electrolyte membrane fuel cells for high temperature operation at 200 ℃[J]. Journal of Materials Chemistry A, 2016, 4(11): 4019-4024.

[13] LI X B, WANG P, LIU Z C, et al. Arylether-type polybenzimidazoles bearing benzimidazolyl pendants for high-temperature proton exchange membrane fuel cells[J]. Journal of Power Sources, 2018, 393: 99-107.

[14] CHEN J C, CHEN P Y, LEE S W, et al. Synthesis of soluble polybenzimidazoles for high-temperature proton exchange membrane fuel cell (PEMFC) applications[J]. Reactive and Functional Polymers, 2016, 108: 122-129.

[15] ZHANG J J, WANG H N, LI W, et al. Effect of catalyst layer microstructures on performance and stability for high temperature polymer electrolyte membrane fuel cells[J]. Journal of Power Sources, 2021, 505: 230059.

[16] KASERER S, CALDWELL K M, RAMAKER D E, et al. Analyzing the influence of H_3PO_4 as catalyst poison in high temperature PEM fuel cells using in-operando X-ray absorption spectroscopy[J]. The Journal of Physical Chemistry C, 2013, 117(12): 6210-6217.

[17] PIMPERL N, BEVILACQUA N, SCHMID M A, et al. Nitrogen-functionalized carbon-supported Pt catalysts implemented in high-temperature polymer electrolyte membrane fuel cell[J]. Journal of Power Sources, 2021, 507: 229971.

[18] LEE K S, SPENDELOW J S, CHOE Y K, et al. An operationally flexible fuel cell based on quaternary ammonium-biphosphate ion pairs[J]. Nature Energy, 2016, 1: 16120.

[19] TANG H Y, GENG K, WU L, et al. Fuel cells with an operational range of −20℃ to 200℃ enabled by phosphoric acid-doped intrinsically ultramicroporous membranes[J]. Nature Energy, 2022, 7: 153-162.

[20] ATANASOV V, LEE A S, PARK E J, et al. Synergistically integrated phosphonated poly(pentafluorostyrene) for fuel cells[J]. Nature Materials, 2021, 20: 370-377.

[21] LIM K H, LEE A S, ATANASOV V, et al. Protonated phosphonic acid electrodes for high power heavy-duty vehicle fuel cells[J]. Nature Energy, 2022, 7: 248-259.

[22] ZHANG J, AILI D, LU S F, et al. Advancement toward polymer electrolyte membrane fuel cells at elevated temperatures[J]. Research, 2020, 2020: 9089405.

[23] LI Q F, JENSEN J O, SAVINELL R F, et al. High temperature proton exchange membranes based on polybenzimidazoles for fuel cells[J]. Progress in Polymer Science, 2009, 34(5): 449-477.

[24] LI W, WANG H N, ZHANG J, et al. Advancements of polyvinylpyrrolidone-based polymer electrolyte membranes for electrochemical energy conversion and storage devices[J]. ChemSusChem, 2022, 15(10): e202200071.

[25] LU S F, XU X, ZHANG J, et al. Progress of phosphoric acid doped high temperature proton exchange membrane for fuel cells[J]. Scientia Sinica Chimica, 2017, 47(5): 565-572.

[26] XU X, WANG H N, LU S F, et al. A novel phosphoric acid doped poly(ethersulphone)-poly(vinyl pyrrolidone) blend membrane for high-temperature proton exchange membrane fuel cells[J]. Journal of Power Sources, 2015, 286: 458-463.

[27] GUO Z B, XU X, XIANG Y, et al. New anhydrous proton exchange membranes for high-temperature fuel cells based on PVDF-PVP blended polymers[J]. Journal of Materials Chemistry A, 2015, 3(1): 148-155.

[28] REN X R, LI H H, LIU K, et al. Preparation and investigation of reinforced PVP blend membranes for high temperature polymer electrolyte membranes[J]. Fibers and Polymers, 2019, 19(12): 2449-2457.

[29] LI M Q, SCOTT K. Polymer electrolyte membrane for high temperature fuel cells to fit vehicle applications[J]. Electrochimica Acta, 2010, 55(6): 2123-2128.

[30] LU S F, XIU R J, XU X, et al. Polytetrafluoroethylene (PTFE) reinforced poly(ethersulphone) – poly(vinyl pyrrolidone) composite membrane for high temperature proton exchange membrane fuel cells[J]. Journal of Membrane Science, 2014, 464: 1-7.

[31] MONG A L, YANG S, KIM D. Pore-filling polymer electrolyte membrane based on poly(arylene ether ketone) for enhanced dimensional stability and reduced methanol permeability[J]. Journal of Membrane Science, 2017, 543: 133-142.

[32] GUO Z B, XIU R J, LU S F, et al. Submicro-pore containing poly(ether sulfones)/polyvinylpyrrolidone membranes for high-temperature fuel cell applications[J]. Journal of Materials Chemistry A, 2015, 3(16): 8847-8854.

[33] GUO Z M, CHEN J N, BYUM J J, et al. Insights into the performance and degradation of polybenzimidazole/muscovite composite membranes in high-temperature proton exchange membrane fuel cells[J]. Journal of Membrance Science, 2022, 641: 119868.

[34] KUMAR S R, WANG J J, WU Y S, et al. Synergistic role of graphene oxide-magnetite nanofillers contribution on ionic conductivity and permeability for polybenzimidazole membrane electrolytes[J]. Journal of Power Sources, 2020, 445: 227293.

[35] BAI H J, WANG H N, ZHANG J, et al. Simultaneously enhancing ionic conduction and mechanical strength of poly(ether sulfones) – poly(vinyl pyrrolidone) membrane by introducing graphitic carbon nitride nanosheets for high temperature proton exchange membrane fuel cell application[J]. Journal of Membrane Science, 2018, 558: 26-33.

[36] BAI H J, ZHANG J, WANG H N, et al. Highly conductive quaternary ammonium-containing cross-linked poly(vinyl pyrrolidone) for high-temperature PEM fuel cells with high-performance[J]. Journal of Membrance Science, 2022, 645: 120194-120203.

[37] KOURASI M, WILLS R G A, SHAH A A, et al. Heteropolyacids for fuel cell applications[J]. Electrochimica Acta, 2014, 127(5): 454-466.

[38] STAILI P, MINUTOLI M, HOCEVAR S. Membranes based on phosphotungstic acid and polybenzimidazole for fuel cell application[J]. Journal of Power Sources, 2000, 90(2): 231-235.

[39] LU S F, XU X, ZHANG J, et al. A self-anchored phosphotungstic acid hybrid proton exchange membrane achieved via one-step synthesis[J]. Advanced Energy Materials, 2014, 4(17): 1400842.

[40] XU X, WANG H N, LU S F, et al. A phosphotungstic acid self-anchored hybrid proton exchange membrane for direct methanol fuel cells[J]. RSC Advances, 2016, 6(49): 43049-43055.

[41] YUAN Z Z, DAI Q, QIAO L, et al. Highly stable aromatic poly(ether sulfone) composite ion exchange membrane for vanadium flow battery[J]. Journal of Membrane Science, 2017, 541: 465-473.

[42] ZHANG J, LIU J, LU S F, et al. Ion-exchange-induced selective

etching for the synthesis of amino-functionalized hollow mesoporous silica for elevated-high-temperature fuel cells[J]. ACS Applied Materials & Interfaces,2017,9(37):31922-31930.

[43] LU S F,WANG D L,JIANG S P P,et al. HPW/MCM-41 phosphotungstic acid/mesoporous silica composites as novel proton-exchange membranes for elevated-temperature fuel cells[J]. Advanced Materials,2010,22(9):971-976.

[44] ZHANG J,CHEN S A,BAI H J,et al. Effects of phosphotungstic acid on performance of phosphoric acid doped polyethersulfone-polyvinylpyrrolidone membranes for high temperature fuel cells[J]. International Journal of Hydrogen Energy,2021,46(19):11104-11114.

[45] ZHANG J J,ZHANG J,BAI H J,et al. A new high temperature polymer electrolyte membrane based on trifunctional group grafted polysulfone for fuel cell application[J]. Journal of Membrane Science,2019,572:496-503.

[46] BAI H J,WANG H N,ZHANG J,et al. High temperature polymer electrolyte membrane achieved by grafting poly(1-vinylimidazole) on polysulfone for fuel cells application[J]. Journal of Membrane Science,2019,592:117395.

[47] BAI H J,PENG H G,XIANG Y,et al. Poly(arylene piperidine)s with phosphoric acid doping as high temperature polymer electrolyte membrane for durable, high-performance fuel cells[J]. Journal of Power Sources,2019,443:227219.

[48] DANG J C,ZHAO L P,ZHANG J,et al. Imidazole microcapsules toward enhanced phosphoric acid loading of polymer electrolyte membrane for anhydrous proton conduction[J]. Journal of Membrane Science,2018,545:88-98.

[49] TANG H Y,GENG K,HAO J K,et al. Properties and stability of quaternary ammonium-biphosphate ion-pair poly(sulfone)s high temperature proton exchange membranes for H_2/O_2 fuel cells[J]. Journal of Power Sources,2020,475:228521.

[50] JIANG Z Y,ZHENG X H,WU H,et al. Proton conducting membranes prepared by incorporation of organophosphorus acids into alcohol barrier polymers for direct methanol fuel cells[J]. Journal of Power Sources,2008,185(1):85-94.

[51] SCHUSTER M, RAGER T, NODA A,et al. About the choice of the protogenic group in PEM separator materials for intermediate temperature, low humidity operation: A critical comparison of sulfonic acid,phosphonic acid and imidazole functionalized model compounds[J]. Fuel Cells,2005,5(3):355-365.

[52] RAGER T,SCHUSTER M,STEININGER H,et al. Poly(1,3-phenylene-5-phosphonic acid), a fully aromatic polyelectrolyte with high ion exchange capacity[J]. Advanced Materials,2007,19(20):3317-3321.

[53] NAGAO M, TAKEUCHI A, HEO P,et al. A proton-conducting In^{3+}-doped SnP_2O_7 electrolyte for intermediate-temperature fuel cells[J]. Electrochemical and Solid-State Letters,2006,9(3):105-109.

[54] WU X,VERMA A,SCOTT K. A Sb-doped SnP_2O_7 solid proton conductor for intermediate temperature fuel cells[J]. Fuel Cells,2008,8(6):453-458.

[55] FORAN G Y,GOWARD G R. Site-specific proton dynamics in indium-doped tin pyrophosphate[J]. The Journal of Physical Chemistry C,2020,124(52):28407-28416.

[56] KANNAN A,LI Q F,CLEEMANN L N,et al. Acid distribution and durability of HT-PEM fuel cells with different electrode supports[J]. Fuel Cells,2018,18(2):103-112.

[57] BECKER H,REIMER U,AILI D,et al. Determination of anion transference number and phosphoric acid diffusion coefficient in high temperature polymer electrolyte membranes[J]. Journal of the Electrochemical Society,2018,165(10):863-869.

[58] BECKER H,CLEEMANN L N,AILI D,et al. Probing phosphoric acid redistribution and anion migration in polybenzimidazole membranes[J]. Electrochemistry Communications,2017,82:21-24.

[59] MARTIN S,LI Q,STEENBERG T,et al. Binderless electrodes for high-temperature polymer electrolyte membrane fuel cells[J]. Journal of Power Sources,2014,272:559-566.

[60] LEE E,KIM D H,PAK C. Effects of cathode catalyst layer fabrication parameters on the performance of high-temperature polymer electrolyte membrane fuel cells[J]. Applied Surface Science,2020,510:145461.

[61] MARTIN S,GARCIA YBRRA P L,CASTILLO J L. Ten-fold reduction from the state-of-the-art platinum loading of electrodes prepared by electrospraying for high temperature proton exchange membrane fuel cells[J]. Electrochemistry Communications,2018,93:57-61.

[62] MACK F,KLAGES M,SCHOLTA J,et al. Morphology studies on high-temperature polymer electrolyte membrane fuel cell electrodes[J]. Journal of Power Sources,2014,255:431-438.

[63] HALTER J,BEVILACQUA N,ZEIS R,et al. The impact of the catalyst layer structure on phosphoric acid migration in HT-PEMFC-An operando X-ray tomographic microscopy study[J]. Journal of Electroanalytical Chemistry,2020,859:113832.

[64] HALTER J,GLOOR T,AMOROSO B,et al. Wetting properties of porous high temperature polymer electrolyte fuel cells materials with phosphoric acid[J]. Physical Chemistry Chemical Physics,2019,21(24):13126-13134.

[65] JIN H,NAYEEM M O G,LEE S,et al. Highly durable nanofiber-reinforced elastic conductors for skin-tight electronic textiles[J]. ACS Nano,2019,13(7):7905-7912.

[66] XIA L,LI X,WU Y,et al. Electrodes derived from carbon fiber-reinforced cellulose nanofiber/multiwalled carbon nanotube hybrid aerogels for high-energy flexible asymmetric supercapacitors[J]. Chemical Engineering Journal,2020,379:122325.

[67] ZHANG J J,BAI H J,YAN W R,et al. Enhancing cell performance and durability of high temperature polymer electrolyte membrane fuel cells by inhibiting the formation of cracks in catalyst layers[J]. Journal of the Electrochemical Society,2020,167(11):114501.

Research progress on key materials of phosphoric acid doped high-temperature proton exchange membrane fuel cells

XIANG Yan[1], LI Wen[1], GUO Zhibin[2,*], ZHANG Jin[1], LU Shanfu[1]

(1. School of Space and Environment, Beihang University, Beijing 100083, China;
2. Beijing Heracles Novel Technology Co., Ltd., Beijing 100162, China)

Abstract: High-temperature proton exchange membrane fuel cells (HT-PEMFC) has fast electrode reaction kinetics, strong resistance to fuel/air impurity poisoning, a wide range of fuel sources (pure hydrogen, methanol-reforming gas, formic acid, etc.), and simple water/thermal management systems due to their high operating temperature (130℃ – 200℃). They have become one of the important development directions of polymer membrane fuel cells. This paper mainly introduces the research progress of Beihang University in HT-PEMFC key materials-high-temperature membrane, catalytic layer, and membrane electrode assemblies in recent ten years. Aiming at the best balance between proton conductivity and mechanical properties of phosphoric acid (PA) doped high-temperature membrane, the influence mechanism of PA distribution and migration in the catalytic layer on cell performance, and the influence and attenuation mechanism of large-size membrane electrode consistency on stack performance. The molecular design of polyelectrolyte membrane materials, the regulation of ordered catalytic layer structure, and the optimization of large-size membrane electrode stack were reviewed, and the technical challenges faced by HT-PEMFC technology as well as the future development trend are reviewed and prospected.

Keywords: fuel cell; high-temperature proton exchange membrane; catalytic layer; membrane electrode; high-temperature fuel cell stack

Received: 2022-07-05; **Accepted**: 2022-07-28; **Published online**: 2022-08-16 15:12

URL: kns.cnki.net/kcms/detail/11.2625.V.20220815.1633.001.html

* **Corresponding author**. E-mail: guozhibin@heracles-tec.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0216

无人机高机动抗扰轨迹跟踪控制方法

王英勋, 宋欣屿, 赵江, 蔡志浩*

(北京航空航天大学 自动化科学与电气工程学院, 北京 100083)

摘 要: 针对传统的四旋翼控制方法在高机动飞行时控制效果不佳,难以跟踪机动性较强的轨迹且控制精度较低的问题,设计了基于增量非线性动态逆控制(INDI)方法和微分平坦前馈的轨迹跟踪控制器,不仅提高了高机动轨迹的跟踪精度,也增强了抗扰能力。由于角加速度无法直接获得,该方法对其非常敏感,设计了多种角加速度估计方法进行对比,通过飞行试验选择了效果最佳的互补滤波方法。试验结果证明:设计的基于互补滤波的前馈 INDI 方法可以控制飞行器快速、准确地跟踪高机动轨迹,且具有较强的抗扰能力。

关 键 词: 四旋翼; 轨迹跟踪; 增量非线性动态逆控制(INDI); 抗扰控制; 角加速度估计

中图分类号: V249.1

文献标志码: A **文章编号**: 1001-5965(2022)09-1806-12

近年来,四旋翼无人机以其独特的外形、结构与飞行方式逐渐成为国内外关注的热点。与常规布局的直升机相比,四旋翼无人机机械结构简单,成本较低,易于维护,4 个螺旋桨对称分布,使得四旋翼无人机的机动能力更强,静态盘旋的稳定性更好,也更容易实现机型的微小型化。四旋翼无人机特别适合在近地面环境(如室内、城市和丛林等)中执行监视、侦察等任务,具有广阔的应用前景[1]。在进行这些任务时,往往需要无人机进行高速度、高加速度、高角速度、高角加速度的运动,然而,四旋翼在高机动飞行时,系统呈现严重的非线性和非定常特性,气动阻力难以建模,给控制器设计带来了很大的难度。此外,精准的跟踪高机动轨迹还需要参考其高阶时间导数,即加加速度和加加加速度。传统的控制方法难以解决建模不准和扰动的问题,也无法考虑轨迹的高阶时间导数,在高机动飞行时控制效果不佳,难以跟踪机动性较强的轨迹[2]。

非线性动态逆控制(nonlinear dynamic inversion, NDI)通过对动态系统求逆来实现非线性控制系统反馈线性化,是非线性控制中的一种较为有效的方法[3],在飞行器控制领域被广泛应用[4],对非线性控制系统具有很好的跟踪性能。但是 NDI 对建模误差十分敏感,控制器的鲁棒性较低[5]。针对这一问题,Sieberling 等[6]在 NDI 基础上提出了增量非线性动态逆控制(incremental nonlinear dynamic inversion, INDI),通过角加速度反馈解决了 NDI 鲁棒性不足的问题,实现了对固定翼无人机姿态的控制,但只进行了仿真,未进行试飞测试。Simplício 等[7]将 INDI 用于直升机悬停控制,验证了 INDI 对模型的依赖性较低,具有较强的鲁棒性,但不适用于轨迹跟踪控制。Lu 等[8]用 INDI 实现了固定翼有人机作动器故障时的轨迹跟踪控制。近年来,INDI 逐渐被用于四旋翼无人机的控制领域[9-14],但在估计角加速度时,均采用了延迟较大的方法,对控制精度有着较大的影响。

角加速度不易测量[11,15],角加速度计是测量

收稿日期: 2022-04-02; **录用日期**: 2022-05-26; **网络出版时间**: 2022-06-14 11:49
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220613.1819.005.html
* **通信作者**. E-mail: czh@buaa.edu.cn

引用格式: 王英勋, 宋欣屿, 赵江, 等. 无人机高机动抗扰轨迹跟踪控制方法[J]. 北京航空航天大学学报, 2022, 48(9): 1806-1817. WANG Y X, SONG X Y, ZHAO J, et al. Anti-disturbance trajectory tracking control method for aggressive quadrotors[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1806-1817(in Chinese).

飞行器角加速度最直接的装置,但由于成本和飞行器尺寸的限制,还未应用于小型无人机[16-18]。虽然也可以通过线加速度计间接测量角加速度值,但需要在机身上合理布置至少 3 组加速度计,难以实现[19-21]。除却测量角加速度外,国内外也提出了多种数值估计方法用以计算角加速度的估计值。最常见的方法是对角速度进行微分获取角加速度的估计值,但其会放大噪声的影响,为此引入低通滤波器,但又会带来一定的延迟。为了在不引入延迟的同时抑制噪声,Valiviita 和 Ovaska[22]提出了一种预测低通滤波器,虽然取得了很好的效果,但其假设加速度可以通过分段低次多项式准确近似,不适用于机动性较强的无人机。Smeur 等[9]使用低通滤波同时处理角加速度和控制输入,实现了滞后同步,虽然看起来很简单,但滞后时间是不确定的。此外,卡尔曼滤波器基于以高斯白噪声为输入的恒定加速度模型[23-25],但实际上无人机的飞行并不一定表现出类似的稳定随机过程[26-27]。

本文主要创新点如下:①设计了基于 INDI 方法的轨迹跟踪控制器,为减小轨迹跟踪的延迟,引入了微分平坦前馈。②由于角加速度无法直接获得,INDI 方法对其又非常敏感,本文设计了多种角加速度估计方法,并通过飞行试验选择了效果最佳的估计方法。③通过飞行试验验证了所提控制器具有高精度的高机动轨迹跟踪效果,且具有较强的抗扰能力。

1　预备知识

1.1　无人机模型

本文研究对象为小型四旋翼无人机,惯性参考坐标系选择北东地坐标系,其正交基为($\boldsymbol{i}_x$,$\boldsymbol{i}_y$,$\boldsymbol{i}_z$),机体坐标系固定在四旋翼飞行器上,坐标原点与其质心重合,如图 1 所示。

图 1 中所示的单位矢量是机体坐标系下的基向量,形成从机体坐标系到惯性参考坐标系的转换旋转矩阵 $\boldsymbol{R}=(\boldsymbol{b}_x \quad \boldsymbol{b}_y \quad \boldsymbol{b}_z)\in \mathrm{SO}(3)$,即

$$\boldsymbol{R}=\begin{bmatrix} \cos\theta\cos\psi & \sin\theta\cos\psi\sin\phi-\sin\psi\cos\phi & \sin\theta\cos\psi\cos\phi+\sin\psi\sin\phi \\ \cos\theta\sin\psi & \sin\theta\sin\psi\sin\phi+\cos\psi\cos\phi & \sin\theta\sin\psi\cos\phi-\cos\psi\sin\phi \\ -\sin\theta & \cos\theta\sin\phi & \cos\theta\cos\phi \end{bmatrix} \tag{1}$$

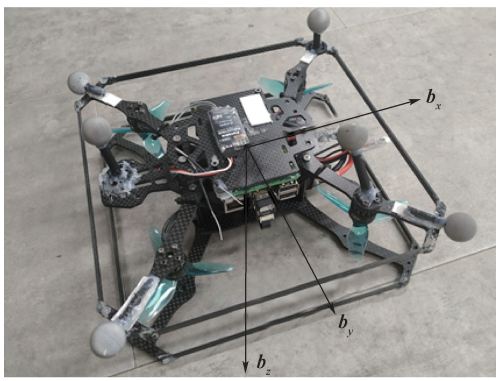


图 1　四旋翼及机体坐标系定义

Fig. 1　Quadrotor with body-fixed reference system

四旋翼无人机的质心运动模型为

$$\dot{\boldsymbol{x}}=\boldsymbol{v} \tag{2}$$

$$\dot{\boldsymbol{v}}=g\boldsymbol{i}_z+\tau\boldsymbol{b}_z+m^{-1}\boldsymbol{f}_{\mathrm{ext}} \tag{3}$$

式中:$\boldsymbol{x}$ 和 $\boldsymbol{v}$ 分别为惯性参考坐标系下的位置和速度。加速度主要考虑 3 项:①向下的重力加速度 g;②推重比 τ 为总推力 T 和机体质量 m 之比,由于推力矢量沿机体坐标系的 $\boldsymbol{b}_z$ 的反方向,四旋翼向前、向后或侧向加速时必须进行俯仰或滚转运动;③外部干扰力矢量 $\boldsymbol{f}_{\mathrm{ext}}$,其包含所有作用于机体的其他力(如空气阻力)。

四旋翼无人机的姿态运动模型为

$$\dot{\boldsymbol{\xi}}=\boldsymbol{S}\boldsymbol{\Omega} \tag{4}$$

$$\dot{\boldsymbol{\Omega}}=\boldsymbol{J}^{-1}(\boldsymbol{M}+\boldsymbol{M}_{\mathrm{ext}}-\boldsymbol{\Omega}\times\boldsymbol{J}\boldsymbol{\Omega}) \tag{5}$$

式中:$\boldsymbol{\Omega}=(p \quad q \quad r)^{\mathrm{T}}$ 为机体坐标系下的角速度向量,p 为滚转角速度,q 为俯仰角速度,r 为偏航角速度;$\boldsymbol{\xi}=(\phi \quad \theta \quad \psi)^{\mathrm{T}}$ 为飞行器的姿态角向量,ϕ 为滚转角,θ 为俯仰角,ψ 为偏航角;$\boldsymbol{J}$ 为机体的惯性张量矩阵;$\boldsymbol{M}$ 为控制力矩;$\boldsymbol{M}_{\mathrm{ext}}$为外部总干扰力矩;$\boldsymbol{S}$ 为机体角速度到姿态角速率的坐标转换矩阵,即

$$\boldsymbol{S}=\begin{bmatrix} 1 & \sin\phi\tan\theta & \cos\phi\tan\theta \\ 0 & \cos\phi & -\sin\phi \\ 0 & \dfrac{\sin\phi}{\cos\theta} & \dfrac{\cos\phi}{\cos\theta} \end{bmatrix} \tag{6}$$

假定每个螺旋桨轴均与 $\boldsymbol{b}_z$ 轴完美平行,用 T_i 和 Q_i 来描述每个电机产生的推力和扭矩,则有

$$\begin{cases} T_i=k_\tau\omega_i^2 \\ Q_i=k_q\omega_i^2 \end{cases} \qquad i=1,2,3,4 \tag{7}$$

式中:$\omega_i>0\,(i=1,2,3,4)$ 为每个电机的旋转速度;k_τ 为电机推力系数;k_q 为电机扭矩系数。

电机的序号和旋转方向如图 2 所示。

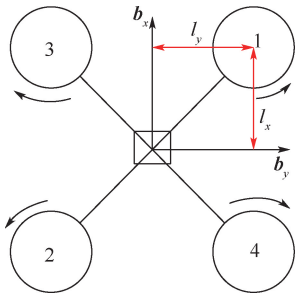


图 2 旋翼序号及旋转方向

Fig. 2 Rotor serial number and rotation direction

可以得到机体坐标系下的控制力矩 $\boldsymbol{M}$ 为

$$\boldsymbol{M} = \begin{bmatrix} l_y(-T_1+T_2+T_3-T_4) \\ l_x(T_1-T_2+T_3-T_4) \\ Q_1+Q_2-Q_3-Q_4+J_r(\dot{\omega}_1+\dot{\omega}_2-\dot{\omega}_3-\dot{\omega}_4) \end{bmatrix} =$$

$$\begin{bmatrix} -l_yk_\tau & l_yk_\tau & l_yk_\tau & -l_yk_\tau \\ l_xk_\tau & -l_xk_\tau & l_xk_\tau & -l_xk_\tau \\ k_q & k_q & -k_q & -k_q \end{bmatrix} \begin{bmatrix} \omega_1^2 \\ \omega_2^2 \\ \omega_3^2 \\ \omega_4^2 \end{bmatrix} +$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ J_r & J_r & -J_r & -J_r \end{bmatrix} \begin{bmatrix} \dot{\omega}_1 \\ \dot{\omega}_2 \\ \dot{\omega}_3 \\ \dot{\omega}_4 \end{bmatrix} = \boldsymbol{G}_1\boldsymbol{\omega}\circ\boldsymbol{\omega} + \boldsymbol{G}_2\dot{\boldsymbol{\omega}} \tag{8}$$

式中:“∘”表示 Hadamard 积;l_x 和 l_y 表示如图 2 所示的力臂长度;J_r 为电机的惯性矩。

$$\boldsymbol{G}_1 = \begin{bmatrix} -l_yk_\tau & l_yk_\tau & l_yk_\tau & -l_yk_\tau \\ l_xk_\tau & -l_xk_\tau & l_xk_\tau & -l_xk_\tau \\ -k_q & -k_q & k_q & k_q \end{bmatrix} \tag{9}$$

$$\boldsymbol{G}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -J_r & -J_r & J_r & J_r \end{bmatrix} \tag{10}$$

忽略电机绕 x 轴和 y 轴的附加力矩,只考虑电机力矩的 z 轴分量。

$$T = T_1+T_2+T_3+T_4 = k_\tau(\omega_1^2+\omega_2^2+\omega_3^2+\omega_4^2) \tag{11}$$

1.2 微分平坦前馈

控制器的目标是精准地跟踪由式(12)定义的参考轨迹:

$$\boldsymbol{\sigma}_{\mathrm{ref}}(t) = (\boldsymbol{x}_{\mathrm{ref}}(t)^{\mathrm{T}} \quad \psi_{\mathrm{ref}}(t))^{\mathrm{T}} \tag{12}$$

$\boldsymbol{\sigma}_{\mathrm{ref}}(t)$ 由 4 个微分平坦输出组成,即惯性参考坐标系下的四旋翼位置 $\boldsymbol{x}_{\mathrm{ref}}(t) \in \mathbf{R}^3$ 和机身偏航角 $\psi_{\mathrm{ref}}(t) \in \mathcal{T}$,$\mathcal{T}$ 表示圆组。

假设 $\boldsymbol{x}_{\mathrm{ref}}$ 的前四阶导数存在且连续,ψ_{ref} 前二阶导数存在且连续。在惯性参考坐标系中,$\boldsymbol{x}_{\mathrm{ref}}$ 对时间求导依次产生参考速度 $\boldsymbol{v}_{\mathrm{ref}}$、参考加速度 $\boldsymbol{a}_{\mathrm{ref}}$、参考加加速度 $\boldsymbol{j}_{\mathrm{ref}}$ 和参考加加加速度 $\boldsymbol{s}_{\mathrm{ref}}$,即

$$\begin{cases} \boldsymbol{v}_{\mathrm{ref}} = \dot{\boldsymbol{x}}_{\mathrm{ref}} \\ \boldsymbol{a}_{\mathrm{ref}} = \dot{\boldsymbol{v}}_{\mathrm{ref}} \\ \boldsymbol{j}_{\mathrm{ref}} = \dot{\boldsymbol{a}}_{\mathrm{ref}} \\ \boldsymbol{s}_{\mathrm{ref}} = \dot{\boldsymbol{j}}_{\mathrm{ref}} \end{cases} \tag{13}$$

同理,ψ_{ref} 对时间求导给出参考偏航角速度 $\dot{\psi}_{\mathrm{ref}}$ 和参考偏航角加速度 $\ddot{\psi}_{\mathrm{ref}}$。

对式(3)求导,可以得到加加速度 $\boldsymbol{j}_{\mathrm{ref}}$ 的表达式为

$$\begin{aligned} \boldsymbol{j}_{\mathrm{ref}} &= \frac{\mathrm{d}}{\mathrm{d}t}(g\boldsymbol{i}_z + \tau\boldsymbol{b}_z + m^{-1}\boldsymbol{f}_{\mathrm{ext}}) = \\ &\dot{\tau}\boldsymbol{R}_{\mathrm{ref}}\boldsymbol{i}_z + \dot{\tau}_{\mathrm{ref}}\boldsymbol{b}_z = \tau\boldsymbol{R}[\boldsymbol{\Omega}_{\mathrm{ref}}]_\times \boldsymbol{i}_z + \dot{\tau}_{\mathrm{ref}}\boldsymbol{b}_z = \\ &\tau\boldsymbol{R}[\boldsymbol{i}_z]_\times^{\mathrm{T}}\boldsymbol{\Omega}_{\mathrm{ref}} + \dot{\tau}_{\mathrm{ref}}\boldsymbol{b}_z = \\ &\tau(-\boldsymbol{b}_y \quad \boldsymbol{b}_x \quad \boldsymbol{0})\boldsymbol{\Omega}_{\mathrm{ref}} + \dot{\tau}_{\mathrm{ref}}\boldsymbol{b}_z \end{aligned} \tag{14}$$

式中:$[\ ]_\times$ 表示斜对称矩阵,$\boldsymbol{a}\times\boldsymbol{b} = [\boldsymbol{a}]_\times\boldsymbol{b}$;$\dot{\boldsymbol{R}}_{\mathrm{ref}} = \boldsymbol{R}[\boldsymbol{\Omega}_{\mathrm{ref}}]_\times$,$[\boldsymbol{\Omega}_{\mathrm{ref}}]_\times\boldsymbol{i}_n = [\boldsymbol{i}_n]_\times^{\mathrm{T}}\boldsymbol{\Omega}_{\mathrm{ref}}$,$n = x,y,z$。

又因为

$$\boldsymbol{\Omega}_{\mathrm{ref}} = \boldsymbol{S}^{-1}\dot{\boldsymbol{\xi}}_{\mathrm{ref}} \tag{15}$$

式中:$\boldsymbol{S}^{-1}$ 为姿态角速率到机体角速度的坐标旋转矩阵。

$$\boldsymbol{S}^{-1} = \begin{bmatrix} 1 & 0 & -\sin\theta \\ 0 & \cos\phi & \sin\phi\cos\theta \\ 0 & -\sin\phi & \cos\phi\cos\theta \end{bmatrix} \tag{16}$$

所以有

$$\begin{aligned} \boldsymbol{j}_{\mathrm{ref}} &= \tau(-\boldsymbol{b}_y \quad \boldsymbol{b}_x \quad \boldsymbol{0})\boldsymbol{S}^{-1}\dot{\boldsymbol{\xi}}_{\mathrm{ref}} + \dot{\tau}_{\mathrm{ref}}\boldsymbol{b}_z = \\ &\tau(-\boldsymbol{b}_y \quad \cos\phi\boldsymbol{b}_x \quad \sin\theta\boldsymbol{b}_y + \sin\phi\cos\theta\boldsymbol{b}_x)\dot{\boldsymbol{\xi}}_{\mathrm{ref}} + \\ &\dot{\tau}_{\mathrm{ref}}\boldsymbol{b}_z = [-\tau\boldsymbol{b}_y \quad \tau\cos\phi\boldsymbol{b}_x \quad \boldsymbol{b}_z]\begin{bmatrix} \dot{\phi}_{\mathrm{ref}} \\ \dot{\theta}_{\mathrm{ref}} \\ \dot{\tau}_{\mathrm{ref}} \end{bmatrix} + \\ &\tau(\sin\theta\boldsymbol{b}_y + \sin\phi\cos\theta\boldsymbol{b}_x)\dot{\psi}_{\mathrm{ref}} \end{aligned} \tag{17}$$

求逆可得

$$\begin{bmatrix} \dot{\phi}_{\mathrm{ref}} \\ \dot{\theta}_{\mathrm{ref}} \\ \dot{\tau}_{\mathrm{ref}} \end{bmatrix} = (-\tau\boldsymbol{b}_y \quad \tau\cos\phi\boldsymbol{b}_x \quad \boldsymbol{b}_z)^{-1} \cdot$$

$$(\boldsymbol{j}_{\mathrm{ref}} - \tau(\sin\theta\boldsymbol{b}_y + \sin\phi\cos\theta\boldsymbol{b}_x)\dot{\psi}_{\mathrm{ref}}) =$$

$$\begin{bmatrix} -\dfrac{1}{\tau}\boldsymbol{b}_y^{\mathrm{T}} \\ \dfrac{1}{\tau\cos\phi}\boldsymbol{b}_x^{\mathrm{T}} \\ \boldsymbol{b}_z^{\mathrm{T}} \end{bmatrix}\boldsymbol{j}_{\mathrm{ref}} + \begin{bmatrix} \sin\theta \\ -\cos\theta\tan\phi \\ 0 \end{bmatrix}\dot{\psi}_{\mathrm{ref}} \tag{18}$$

即

$$\dot{\boldsymbol{\xi}}_{\mathrm{ref}} = \frac{1}{\tau}\begin{bmatrix} -\boldsymbol{b}_y^{\mathrm{T}} \\ \dfrac{\boldsymbol{b}_x^{\mathrm{T}}}{\cos\phi} \\ 0 \end{bmatrix}\boldsymbol{j}_{\mathrm{ref}} + \begin{bmatrix} \sin\theta \\ -\cos\theta\tan\phi \\ 1 \end{bmatrix}\psi_{\mathrm{ref}} \tag{19}$$

$$\dot{\tau}_{\mathrm{ref}} = \boldsymbol{b}_z^{\mathrm{T}}\boldsymbol{j}_{\mathrm{ref}} \tag{20}$$

由式(15)可得

$$\boldsymbol{\Omega}_{\mathrm{ref}} = \boldsymbol{S}^{-1}\dot{\boldsymbol{\xi}}_{\mathrm{ref}} = \frac{1}{\tau}\begin{bmatrix} -\boldsymbol{b}_y^{\mathrm{T}} \\ \boldsymbol{b}_x^{\mathrm{T}} \\ -\boldsymbol{b}_x^{\mathrm{T}}\tan\phi \end{bmatrix}\boldsymbol{j}_{\mathrm{ref}} + \begin{bmatrix} 0 \\ 0 \\ \dfrac{\cos\theta}{\cos\phi} \end{bmatrix}\dot{\psi}_{\mathrm{ref}} \tag{21}$$

对式(21)求导可得

$$\dot{p}_{\mathrm{ref}} = \frac{-\boldsymbol{b}_z^{\mathrm{T}}p_{\mathrm{ref}} + \boldsymbol{b}_x^{\mathrm{T}}r_{\mathrm{ref}}}{\tau}\boldsymbol{j}_{\mathrm{ref}} + \frac{\dot{\tau}_{\mathrm{ref}}\boldsymbol{b}_y^{\mathrm{T}}}{\tau^2}\boldsymbol{j}_{\mathrm{ref}} - \frac{\boldsymbol{b}_y^{\mathrm{T}}}{\tau}\boldsymbol{s}_{\mathrm{ref}} \tag{22}$$

将式(20)、式(21)代入可得

$$\dot{p}_{\mathrm{ref}} = \frac{-\dot{\tau}_{\mathrm{ref}}}{\tau}p_{\mathrm{ref}} + q_{\mathrm{ref}}r_{\mathrm{ref}} - \frac{\dot{\tau}_{\mathrm{ref}}}{\tau}p_{\mathrm{ref}} - \frac{\boldsymbol{b}_y^{\mathrm{T}}}{\tau}\boldsymbol{s}_{\mathrm{ref}} = -\frac{\boldsymbol{b}_y^{\mathrm{T}}}{\tau}\boldsymbol{s}_{\mathrm{ref}} - \frac{2\dot{\tau}_{\mathrm{ref}}}{\tau}p_{\mathrm{ref}} + q_{\mathrm{ref}}r_{\mathrm{ref}} \tag{23}$$

同理可得

$$\dot{q}_{\mathrm{ref}} = \frac{-\boldsymbol{b}_z^{\mathrm{T}}q_{\mathrm{ref}} + \boldsymbol{b}_y^{\mathrm{T}}r_{\mathrm{ref}}}{\tau}\boldsymbol{j}_{\mathrm{ref}} - \frac{\dot{\tau}_{\mathrm{ref}}\boldsymbol{b}_x^{\mathrm{T}}}{\tau^2}\boldsymbol{j}_{\mathrm{ref}} + \frac{\boldsymbol{b}_x^{\mathrm{T}}}{\tau}\boldsymbol{s}_{\mathrm{ref}} = \frac{-\dot{\tau}_{\mathrm{ref}}}{\tau}q_{\mathrm{ref}} - p_{\mathrm{ref}}r_{\mathrm{ref}} - \frac{\dot{\tau}_{\mathrm{ref}}}{\tau}q_{\mathrm{ref}} + \frac{\boldsymbol{b}_x^{\mathrm{T}}}{\tau}\boldsymbol{s}_{\mathrm{ref}} = \frac{\boldsymbol{b}_x^{\mathrm{T}}}{\tau}\boldsymbol{s}_{\mathrm{ref}} - \frac{2\dot{\tau}_{\mathrm{ref}}}{\tau}q_{\mathrm{ref}} - p_{\mathrm{ref}}r_{\mathrm{ref}} \tag{24}$$

$$\dot{r}_{\mathrm{ref}} = -\frac{1}{\cos^2\phi}q_{\mathrm{ref}}\dot{\phi}_{\mathrm{ref}} - \tan\phi\dot{q}_{\mathrm{ref}} - \frac{\sin\theta}{\cos\phi}\dot{\theta}_{\mathrm{ref}}\dot{\psi}_{\mathrm{ref}} + \frac{\sin\phi\cos\theta}{\cos^2\phi}\dot{\phi}_{\mathrm{ref}}\dot{\psi}_{\mathrm{ref}} + \frac{\cos\theta}{\cos\phi}\ddot{\psi}_{\mathrm{ref}} = -\frac{1}{\cos\phi}\dot{\theta}_{\mathrm{ref}}\dot{\phi}_{\mathrm{ref}} - \tan\phi\dot{q}_{\mathrm{ref}} - \frac{\sin\theta}{\cos\phi}\dot{\theta}_{\mathrm{ref}}\dot{\psi}_{\mathrm{ref}} + \frac{\cos\theta}{\cos\phi}\ddot{\psi}_{\mathrm{ref}} \tag{25}$$

通过上面的推导,可以将轨迹跟踪问题重新构造为状态跟踪问题,将推导出的参考角速度 $\boldsymbol{\Omega}_{\mathrm{ref}}$和参考角加速度 $\dot{\boldsymbol{\Omega}}_{\mathrm{ref}}$关于参考轨迹加加加速度 $\boldsymbol{s}_{\mathrm{ref}}$、加加速度 $\boldsymbol{j}_{\mathrm{ref}}$、偏航角速度 $\dot{\psi}_{\mathrm{ref}}$和偏航角加速度 $\ddot{\psi}_{\mathrm{ref}}$的表达式,作为轨迹跟踪控制中的前馈输入。

1.3　增量非线性动态逆控制原理

常规非线性系统的状态空间表达式可以写为

$$\begin{cases} \dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t)) + g(\boldsymbol{x}(t),\boldsymbol{u}(t)) \\ \boldsymbol{y}(t) = h(\boldsymbol{x}(t)) \end{cases} \tag{26}$$

式中:$\boldsymbol{x}(t)$为状态向量;$\boldsymbol{u}(t)$为控制输入向量;$\boldsymbol{y}(t)$为输出向量。不同于 NDI 方法将整个系统的动态特性求逆,INDI 方法在当前时刻的状态 $\boldsymbol{x}_0$和控制输入 $\boldsymbol{u}_0$ 下对系统的动态特性进行线性化并求逆。

使用一阶泰勒级数展开式重写式(26),忽略其他高阶项可得

$$\begin{cases} \dot{\boldsymbol{x}} = \dot{\boldsymbol{x}}_0 + \left.\dfrac{\partial f(\boldsymbol{x})}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_0,\boldsymbol{u}_0}\cdot(\boldsymbol{x}-\boldsymbol{x}_0) + \left.\dfrac{\partial g(\boldsymbol{x},\boldsymbol{u})}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}_0,\boldsymbol{u}_0}\cdot(\boldsymbol{x}-\boldsymbol{x}_0) + \left.\dfrac{\partial g(\boldsymbol{x},\boldsymbol{u})}{\partial \boldsymbol{u}}\right|_{\boldsymbol{x}_0,\boldsymbol{u}_0}\cdot(\boldsymbol{u}-\boldsymbol{u}_0) + \Delta \\ \dot{\boldsymbol{y}} = \dfrac{\mathrm{d}h}{\mathrm{d}\boldsymbol{x}}\cdot\dot{\boldsymbol{x}} \end{cases} \tag{27}$$

由于飞控计算机的采样时间很短,作动器的响应速度很快,与 $\Delta\boldsymbol{u}$ 相比,$\Delta\boldsymbol{x}$ 的作用可以忽略不计。因此,可以将式(27)简化为

$$\dot{\boldsymbol{y}} \approx \frac{\mathrm{d}h}{\mathrm{d}\boldsymbol{x}}\cdot\left[\dot{\boldsymbol{x}}_0 + \left.\frac{\partial g(\boldsymbol{x},\boldsymbol{u})}{\partial \boldsymbol{u}}\right|_{\boldsymbol{x}_0,\boldsymbol{u}_0}\cdot(\boldsymbol{u}-\boldsymbol{u}_0)\right] = \dot{\boldsymbol{y}}_0 + \boldsymbol{B}(\boldsymbol{x}_0,\boldsymbol{u}_0)\cdot\Delta\boldsymbol{u} \tag{28}$$

由此推导出 INDI 的控制律为

$$\begin{cases} \Delta\boldsymbol{u} = \boldsymbol{B}^{-1}(\boldsymbol{x}_0,\boldsymbol{u}_0)(\dot{\boldsymbol{y}}_{\mathrm{des}} - \dot{\boldsymbol{y}}_0) \\ \boldsymbol{u} = \boldsymbol{u}_0 + \Delta\boldsymbol{u} \end{cases} \tag{29}$$

式中:$\dot{\boldsymbol{y}}_{\mathrm{des}}$为期望输出的导数;$\dot{\boldsymbol{y}}_0$ 为当前输出的导数;$\boldsymbol{B}^{-1}(\boldsymbol{x}_0,\boldsymbol{u}_0)$为当前的操纵导数矩阵 $\boldsymbol{B}(\boldsymbol{x}_0,\boldsymbol{u}_0)$的广义逆。

2　基于前馈 INDI 的抗扰控制器设计

总体控制框架如图 3 所示。其中,无人机的加速度由加速度计测得,用 $\boldsymbol{a}^{\mathrm{b}}$ 表示在机体坐标系下未经过滤波的加速度测量值,用 $\boldsymbol{a}$ 表示惯性参考坐标系下经过重力加速度校正的加速度,则有

$$\boldsymbol{a} = \boldsymbol{R}\boldsymbol{a}^{\mathrm{b}} + g\boldsymbol{i}_z \tag{30}$$

无人机的位置信息由 OptiTrack 提供,为了融合加速度计和 OptiTrack 的信息,采用了扩展卡尔曼滤波(EKF)方法,获得了经过 EKF 处理后的位置 $\boldsymbol{x}_{\mathrm{f}}$、速度 $\boldsymbol{v}_{\mathrm{f}}$、加速度 $\boldsymbol{a}_{\mathrm{f}}$。角速度由陀螺仪测量得到,根据角速度信息和角加速度估计方法得到角加速度的估计值 $\dot{\boldsymbol{\Omega}}_{\mathrm{f}}$。

2.1　位置和速度控制

位置和速度控制基于 2 个串联的 PD 控制器,所得的控制器表达式为

$$\boldsymbol{a}_{\mathrm{c}} = \boldsymbol{K}_v(\boldsymbol{K}_x(\boldsymbol{x}_{\mathrm{ref}} - \boldsymbol{x}) + \boldsymbol{v}_{\mathrm{ref}} - \boldsymbol{v}) + \boldsymbol{a}_{\mathrm{ref}} = \boldsymbol{K}_v\boldsymbol{K}_x(\boldsymbol{x}_{\mathrm{ref}} - \boldsymbol{x}) + \boldsymbol{K}_v(\boldsymbol{v}_{\mathrm{ref}} - \boldsymbol{v}) + \boldsymbol{a}_{\mathrm{ref}} \tag{31}$$

式中:$\boldsymbol{K}_x$ 为位置环对角增益矩阵;$\boldsymbol{K}_v$ 为速度环对角增益矩阵;下标 ref 表示从参考轨迹获得的参考值,下标 c 表示控制器计算的命令值。式中的前 2 项确保跟踪参考位置和参考速度,最后 1 项作为前馈输入以确保跟踪参考加速度,得到的加速度命令用于计算推力和姿态命令。

位置和速度控制原理框图如图 4 所示。

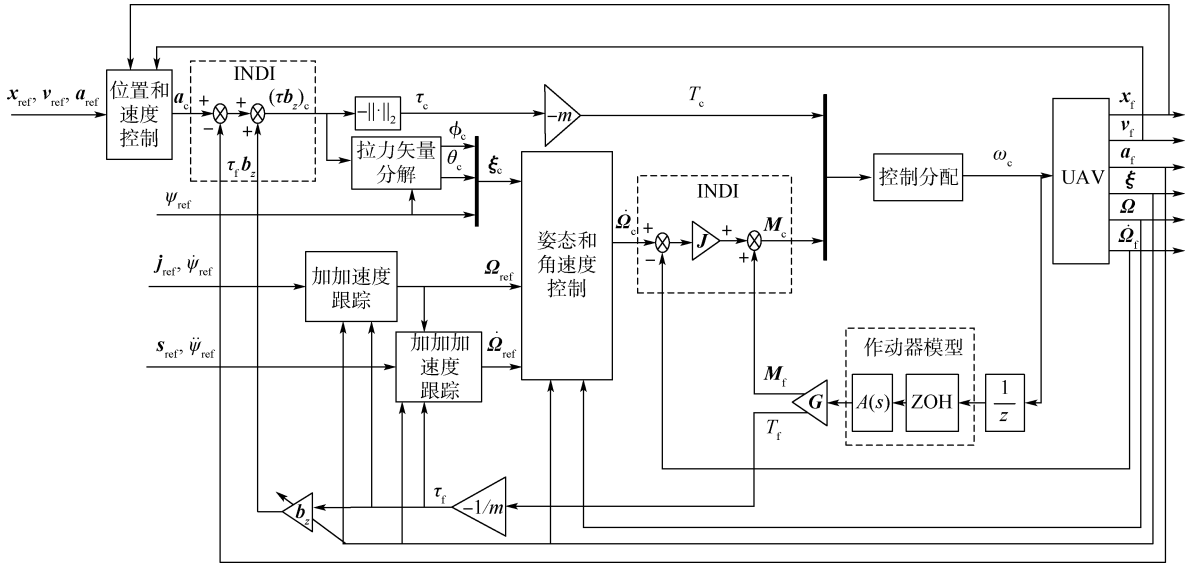


图 3 总体控制框图

Fig. 3 Overall control diagram

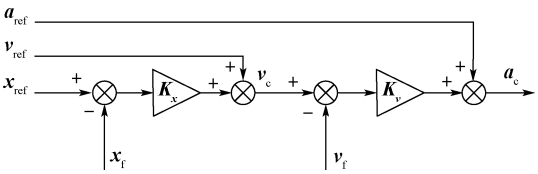


图 4 位置和速度控制

Fig. 4 Position and velocity control

2.2 INDI 加速度控制

根据式(3)获得基于测得的加速度和推重比的干扰力表达式:

$$\boldsymbol{f}_{\text{ext}} = m(\boldsymbol{a}_{\text{f}} - (\tau\boldsymbol{b}_z)_{\text{f}} - g\boldsymbol{i}_z) \tag{32}$$

将式(32)代入式(3)可得

$$\begin{aligned}\boldsymbol{a} &= \tau\boldsymbol{b}_z + g\boldsymbol{i}_z + m^{-1}\boldsymbol{f}_{\text{ext}} = \\ &\tau\boldsymbol{b}_z + g\boldsymbol{i}_z + m^{-1}(m(\boldsymbol{a}_{\text{f}} - (\tau\boldsymbol{b}_z)_{\text{f}} - g\boldsymbol{i}_z)) = \\ &\tau\boldsymbol{b}_z - (\tau\boldsymbol{b}_z)_{\text{f}} + \boldsymbol{a}_{\text{f}}\end{aligned} \tag{33}$$

利用式(33)的增量关系来计算式(31)给出的加速度命令所需的推重比命令:

$$(\tau\boldsymbol{b}_z)_{\text{c}} = \tau_{\text{f}}\boldsymbol{b}_z + \boldsymbol{a}_{\text{c}} - \boldsymbol{a}_{\text{f}} \tag{34}$$

式中:τ_{f} 为根据电机转速估计值和推力模型估算的当前推重比。

式(34)增量性质使控制器可以在存在干扰或建模误差的情况下实现对加速度的控制,所需的推重比为

$$\tau_{\text{c}} = -\left\|(\tau\boldsymbol{b}_z)_{\text{c}}\right\|_2 \tag{35}$$

推力矢量$(\tau\boldsymbol{b}_z)_{\text{c}}$的方向与机体坐标系 $\boldsymbol{b}_z$ 轴的方向一致,假设不存在 $\theta = \pm\pi/2$ 的奇异情况,可以根据几何关系解算滚转角和俯仰角:

$$\begin{cases}\phi_{\text{c}} = \arcsin\left(\dfrac{(-\sin\psi \quad \cos\psi \quad 0)(\tau\boldsymbol{b}_z)_{\text{c}}}{\left\|(\tau\boldsymbol{b}_z)_{\text{c}}\right\|_2}\right) \\ \theta_{\text{c}} = \arctan\left(\dfrac{(\cos\psi \quad \sin\psi \quad 0)(\tau\boldsymbol{b}_z)_{\text{c}}}{\boldsymbol{i}_z^{\text{T}}(\tau\boldsymbol{b}_z)_{\text{c}}}\right)\end{cases} \tag{36}$$

加速度控制原理框图见图 3 左侧虚线部分。

2.3 姿态和角速度控制

式(36)和已知的偏航角参考值 ψ_{ref} 共同构成姿态角指令:

$$\boldsymbol{\xi}_{\text{c}} = \begin{bmatrix}\phi_{\text{c}} \\ \theta_{\text{c}} \\ \psi_{\text{ref}}\end{bmatrix} \tag{37}$$

姿态和角速度控制基于 2 个串联的 PD 控制器,所得的控制器等效于:

$$\begin{aligned}\dot{\boldsymbol{\Omega}}_{\text{c}} &= \boldsymbol{K}_{\boldsymbol{\Omega}}(\boldsymbol{S}^{-1}\boldsymbol{K}_{\boldsymbol{\xi}}(\boldsymbol{\xi}_{\text{c}} - \boldsymbol{\xi}) + \boldsymbol{\Omega}_{\text{ref}} - \boldsymbol{\Omega}_{\text{f}}) + \dot{\boldsymbol{\Omega}}_{\text{ref}} = \\ &\boldsymbol{S}^{-1}\boldsymbol{K}_{\boldsymbol{\Omega}}\boldsymbol{K}_{\boldsymbol{\xi}}(\boldsymbol{\xi}_{\text{c}} - \boldsymbol{\xi}) + \boldsymbol{K}_{\boldsymbol{\Omega}}(\boldsymbol{\Omega}_{\text{ref}} - \boldsymbol{\Omega}_{\text{f}}) + \dot{\boldsymbol{\Omega}}_{\text{ref}}\end{aligned} \tag{38}$$

式中:$\boldsymbol{K}_{\boldsymbol{\xi}}$ 为姿态环对角增益矩阵;$\boldsymbol{K}_{\boldsymbol{\Omega}}$ 为角速度环对角增益矩阵;$\boldsymbol{\Omega}_{\text{ref}}$ 由式(21)计算得到,$\dot{\boldsymbol{\Omega}}_{\text{ref}}$ 由式(24)、式(25)计算得到。式中的前 2 项确保跟踪姿态指令和参考角速度,最后 1 项作为前馈输入以确保跟踪参考角加速度,实现对高阶加速度的跟踪。

姿态和角速度控制的原理框图如图 5 所示。

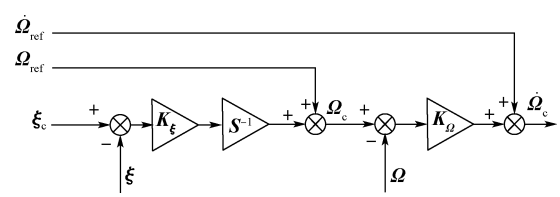


图 5 姿态和角速度控制

Fig. 5 Attitude and angular rate control

2.4 INDI 角加速度控制

由式(5)可知

$$\boldsymbol{M}_{\text{ext}} = \boldsymbol{J}\dot{\boldsymbol{\Omega}}_{\text{f}} - \boldsymbol{M}_{\text{f}} + \boldsymbol{\Omega}_{\text{f}} \times \boldsymbol{J}\boldsymbol{\Omega}_{\text{f}} \tag{39}$$

式中:$\boldsymbol{M}_{\mathrm{f}}$ 为机体坐标系下根据转速指令计算的控制力矩。

将式(39)代入式(5)可得

$$\begin{aligned}\dot{\boldsymbol{\Omega}}_{\mathrm{c}} &= \boldsymbol{J}^{-1}(\boldsymbol{M}_{\mathrm{c}} + \boldsymbol{M}_{\mathrm{ext}} - \boldsymbol{\Omega}_{\mathrm{f}} \times \boldsymbol{J}\boldsymbol{\Omega}_{\mathrm{f}}) = \\ &\boldsymbol{J}^{-1}(\boldsymbol{M}_{\mathrm{c}} + (\boldsymbol{J}\dot{\boldsymbol{\Omega}}_{\mathrm{f}} - \boldsymbol{M}_{\mathrm{f}} + \boldsymbol{\Omega}_{\mathrm{f}} \times \boldsymbol{J}\boldsymbol{\Omega}_{\mathrm{f}}) - \\ &\boldsymbol{\Omega}_{\mathrm{f}} \times \boldsymbol{J}\boldsymbol{\Omega}_{\mathrm{f}}) = \boldsymbol{J}^{-1}(\boldsymbol{M}_{\mathrm{c}} + \boldsymbol{J}\dot{\boldsymbol{\Omega}}_{\mathrm{f}} - \boldsymbol{M}_{\mathrm{f}})\end{aligned} \tag{40}$$

于是有

$$\boldsymbol{M}_{\mathrm{c}} = \boldsymbol{J}\dot{\boldsymbol{\Omega}}_{\mathrm{c}} - \boldsymbol{J}\dot{\boldsymbol{\Omega}}_{\mathrm{f}} + \boldsymbol{M}_{\mathrm{f}} \tag{41}$$

由于角加速度无法直接获得,需要通过角加速度估计方法进行估算,$\dot{\boldsymbol{\Omega}}_{\mathrm{f}}$ 有一定时间延迟,式(41)中所有变量应为同一时刻,否则该式是错误的,造成系统振荡发散。因此,在利用当前时刻电机转速估算控制力矩 $\boldsymbol{M}_{\mathrm{f}}$ 时,应使其延迟与角加速度估计延迟保持一致。角加速度控制原理框图见图 3 中间虚线部分。

2.5　电机转速估计

由于缺少电机转速的传感器和电调的转速反馈,引入作动器模型以获得当前时刻的电机转速的估计值。电机转速指令与其响应之间的动态关系可以表示为连续的一阶传递函数。

$$\begin{cases}\omega_i = A(s)\omega_{ci} \\ A(s) = \dfrac{1}{\tau_{\mathrm{a}}s + 1}\end{cases} \qquad i = 1,2,3,4 \tag{42}$$

式中:$\omega_{ci}(i=1,2,3,4)$ 为每个电机的转速指令;τ_{a} 为可以通过参数辨识获得的作动器时间常数。

根据估计的转速和 1.1 节中的无人机模型,即可估算无人机当前时刻的推重比 τ_{f} 和控制力矩 $\boldsymbol{M}_{\mathrm{f}}$。

2.6　基于互补滤波的低延迟角加速度估计

INDI 基于加速度反馈,降低了对模型气动参数的敏感性,提高了控制的鲁棒性,因此,加速度值的准确性和实时性是实现上述方法的基础。在实际应用中,小型无人机由于体积和成本的原因,主要装备的是惯性传感器(inertial measurement unit, IMU),无法直接获得角加速度,因此需要设计延迟低、精度高的方法对角加速度进行估计。

2.6.1　基于模型预测法的角加速度估计

根据飞行器的动力学方程(5),可以通过理论估计方法计算出角加速度信号,如下:

$$\hat{\dot{\boldsymbol{\Omega}}}_{\mathrm{MP}} = \boldsymbol{J}^{-1}(\hat{\boldsymbol{M}}_{\mathrm{c}} + \hat{\boldsymbol{M}}_{\mathrm{ext}} - \boldsymbol{\Omega} \times \boldsymbol{J}\boldsymbol{\Omega}) \tag{43}$$

式中:$\hat{\boldsymbol{M}}_{\mathrm{c}}$ 和 $\hat{\boldsymbol{M}}_{\mathrm{ext}}$ 分别为控制力矩和干扰力矩的估计值;$\hat{\dot{\boldsymbol{\Omega}}}_{\mathrm{MP}}$ 为模型预测(model prediction, MP)方法估计的角加速度。在计算时忽略了难以估测的干扰力矩 $\hat{\boldsymbol{M}}_{\mathrm{ext}}$,控制力矩 $\hat{\boldsymbol{M}}_{\mathrm{c}}$ 采用 2.5 节中根据转速估计的 $\boldsymbol{M}_{\mathrm{f}}$,因此有

$$\hat{\dot{\boldsymbol{\Omega}}}_{\mathrm{MP}} = \boldsymbol{J}^{-1}(\boldsymbol{M}_{\mathrm{f}} - \boldsymbol{\Omega} \times \boldsymbol{J}\boldsymbol{\Omega}) \tag{44}$$

由于实际中无法对飞行器的参数和空气动力学准确建模,控制力矩会受到模型不确定性的影响,无法准确估计,此外系统扰动无法准确建模,模型预测方法无法考虑实际飞行中外界干扰的影响,外界干扰的影响会导致估计结果偏离真值。虽然该方法的鲁棒性较差,在实际中不能很好地估计角加速度,但其不使用微分运算,也不需要低通滤波器,因此噪声很小,相位滞后几乎为零。由于模型预测方法是一种基于系统非线性动力学方程的理论估计方法,其动态响应非常快。

2.6.2　基于微分法和低通滤波的角加速度估计

微分法是现实中估计角加速度信号最直接和最常用的方法。由于微分运算对系统噪声有严重的放大作用,会导致信号性能变差,通常使用二阶巴特沃斯低通滤波器(low-pass filter, LPF)对 IMU 信号进行滤波,以减轻机身振动和其他噪声的影响,再进行微分运算获取角加速度。

$$H(s) = \frac{\omega_{\mathrm{n}}^2}{s^2 + 1.41\omega_{\mathrm{n}}s + \omega_{\mathrm{n}}^2} \tag{45}$$

式中:ω_{n} 为低通滤波器的截止频率。

很难找到一种低通滤波器,既能表现出小的相位滞后,又能抑制微分操作引起的放大噪声。但该方法不依赖于系统模型参数,其估计结果可以有效地反映外界干扰和系统不确定性的影响,从而能够真实地估计出实际角加速度信号的变化。

2.6.3　基于互补滤波的角加速度估计

模型预测方法具有噪声少、相位滞后小的优点,可以弥补微分法的不足。微分法可以估计实际角加速度信号的真值,弥补模型预测估计结果的误差。因此,基于互补滤波(complementary filter, CF)理论的思想对估计结果进行数据融合,可以得到较好的估计结果。

互补滤波器结构如图 6 所示。其中,$G(s)$ 为积分器,$\hat{\dot{\boldsymbol{\Omega}}}_{\mathrm{MP}}$ 为模型预测方法提供的角加速度估计值,$\boldsymbol{\Omega}$ 为由 IMU 测量的角速度信息,因此 $\hat{\dot{\boldsymbol{\Omega}}}_{\mathrm{D}}$ 等价于通过微分法获得的角加速度估计值,$\hat{\dot{\boldsymbol{\Omega}}}_{\mathrm{CF}}$ 为互补滤波方法提供的角加速度估计值。

通过推导,角加速度的最终估计值如下:

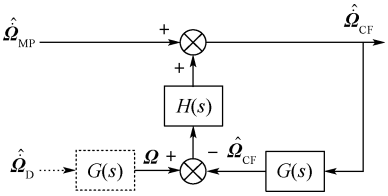


图 6 互补滤波

Fig. 6 Complementary filter

$$\hat{\boldsymbol{\Omega}}_{\mathrm{CF}}(s) = \frac{1}{1+G(s)H(s)}\hat{\boldsymbol{\Omega}}_{\mathrm{MP}}(s) + \frac{G(s)H(s)}{1+G(s)H(s)}\hat{\boldsymbol{\Omega}}_{\mathrm{D}}(s) = T(s)\hat{\boldsymbol{\Omega}}_{\mathrm{MP}}(s) + S(s)\hat{\boldsymbol{\Omega}}_{\mathrm{D}}(s) = \frac{1}{1+G(s)H(s)}\hat{\boldsymbol{\Omega}}_{\mathrm{MP}}(s) + \frac{H(s)}{1+G(s)H(s)}\boldsymbol{\Omega}(s) \tag{46}$$

将 $H(s)$ 设计为如下 PI 控制结构：

$$H(s) = K_{\mathrm{P}} + \frac{K_{\mathrm{I}}}{s} \tag{47}$$

式中：K_{P} 为比例系数；K_{I} 为积分系数。

则最终估计值的表达式为

$$\hat{\dot{\boldsymbol{\Omega}}}_{\mathrm{CF}}(s) = \frac{s^2}{s^2+K_{\mathrm{P}}s+K_{\mathrm{I}}}\hat{\dot{\boldsymbol{\Omega}}}_{\mathrm{MP}}(s) + \frac{K_{\mathrm{P}}s+K_{\mathrm{I}}}{s^2+K_{\mathrm{P}}s+K_{\mathrm{I}}}\hat{\dot{\boldsymbol{\Omega}}}_{\mathrm{D}}(s) = \frac{s^2}{s^2+K_{\mathrm{P}}s+K_{\mathrm{I}}}\cdot\hat{\dot{\boldsymbol{\Omega}}}_{\mathrm{MP}}(s) + \frac{K_{\mathrm{P}}s+K_{\mathrm{I}}}{s^2+K_{\mathrm{P}}s+K_{\mathrm{I}}}[s\boldsymbol{\Omega}(s)] \tag{48}$$

式(48)是一个阻尼比 $\zeta = \dfrac{K_{\mathrm{P}}}{2\sqrt{K_{\mathrm{I}}}}$、自然频率 $\omega_{\mathrm{n}} = \sqrt{K_{\mathrm{I}}}$ 的二阶系统。

从方程(48)可以看出，$T(s)$ 为一个高通滤波器，其保留了模型预测方法的有益的高通特性，并抑制影响低频不确定性和干扰；$S(s)$ 为一个低通滤波器，可滤除 $\hat{\boldsymbol{\Omega}}_{\mathrm{D}}(s)$ 的高频噪声并保留信号的低通特性。由于 $G(s)$ 是一个积分器，$G(s)\hat{\boldsymbol{\Omega}}_{\mathrm{D}}(s)$ 是可以通过 IMU 测量的角速度信号。因此，IMU 的测量值可以直接用于数据融合方法的实际实现中，巧妙地避免了整个估计过程中的微分运算。因此，该方法不会遇到由差分操作引起的噪声放大问题。

$T(s)+S(s)=1$ 是互补性的体现。它将每种方法的有利部分保留在最终估计结果中，并确保截止频率彼此相同。因此，微分法的相位滞后问题通过补偿模型预测方法的高通特性来解决。通过角速度信号的反馈，可以消除外界干扰和模型不确定性的影响。综上所述，该估计方法的最大优点是在估计过程中巧妙避免了微分运算，有效降低噪声的放大效应，提高信号性能。

采集实飞数据，分别用较为常用的直接微分法、卡尔曼滤波方法和互补滤波方法进行角加速度估计。直接微分法由于噪声被放大，根据多次试验，选择了延迟相对较小且对噪声抑制效果较好的、截止频率为 70 rad/s 的二阶巴特沃斯低通滤波器滤波处理。卡尔曼滤波器中由于控制器的控制周期为 250 Hz，因此 $\Delta T = 0.004$ s。互补滤波器的 $H(s)$ 根据试验反复调整到最佳效果。三者对角加速度估计结果如图 7 所示，其中虚线为真实值，红色实线为互补滤波方法估计结果，蓝色虚线为卡尔曼滤波方法估计结果，绿色点线为对角速度直接微分后低通滤波的估计结果。

低通滤波方法虽然抑制噪声的效果很好，但其时间延迟较大，卡尔曼滤波方法虽然可以实时计算最优卡尔曼增益，但由于其计算角加速度预测值时采用了线性模型，在姿态变化剧烈时测量值与预测值误差较大，也导致了时间延迟较大。与卡尔曼滤波和微分后低通滤波方法相比，互补滤波方法虽然计算量更大，但其抑制噪声的能力更强，时间延迟也更小。因此，在实现高机动轨迹跟踪控制时，采用互补滤波方法获得角加速度反馈以提高 INDI 方法的精度。

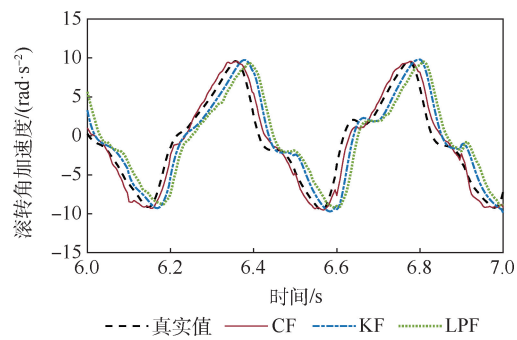


图 7 角加速度估计结果

Fig. 7 Angular acceleration estimation results

3 飞行试验

为了验证本文所设计控制器的精度和抗扰性能，采用图 1 中所示四旋翼无人机进行飞行试验。该四旋翼重 667 g，推进系统由 T-MOTOR F1507 3 800 kV 电机、T-MOTOR T3140 桨叶和 T-MOTOR V45A V2 电调组成，相邻电机相距 15 cm，由一个 4S LiPo 电池供电。飞控选用 Pixraptor，控制频率为 250 Hz，飞机的姿态、角速度、加速度等由其内部 IMU 测量得到，位置信息由 120 Hz 的

OptiTrack动作捕捉系统提供。利用 EKF 方法对 OptiTrack 的位置信息和 IMU 的加速度信息进行处理融合,得到融合后的位置、速度和加速度信息,根据试验选择延迟最小、精度最高的互补滤波方法对角加速度进行估计。

首先,使跟踪速度和加速度都很大的高机动轨迹验证其对高机动轨迹的跟踪能力;其次,通过附加超过飞行器正面面积 4 倍以上的纸板增大飞行过程中的阻力,增大其模型误差,验证其对内部干扰的鲁棒性;然后,利用风扇测试其在 5 m/s 的风干扰中的悬停能力,进一步展示了对外部干扰的鲁棒性;最后,将设计的控制器性能与常规 PID 控制器进行对比分析。

3.1　轨迹跟踪试验

为了保证试验的安全,轨迹跟踪的起始和终止状态飞机应保持悬停,根据参考文献[27]中的 Minimum snap 方法,设计了不同加速度、速度的 8 字形参考轨迹,分别利用前馈 PID 方法和本文所设计的前馈 INDI 方法对其进行跟踪。当参考轨迹的最大速度为 4.481 m/s,最大加速度为 5.117 m/s^2 时,前馈 PID 方法和前馈 INDI 方法均能很好地跟踪参考轨迹,但前馈 INDI 方法跟踪的均方根误差远小于前馈 PID 方法。试验结果如图 8和表 1 所示。

在跟踪最大速度为 6.098 m/s、最大加速度为 12.648 m/s^2 的 8 字形轨迹时,前馈 PID 方法的跟踪误差为 17.238 cm,而前馈 INDI 方法仅为 7.175 cm。此时,前馈 PID 方法控制飞行器进行轨迹跟踪时,姿态较为振荡,稳定性下降。试验结果如图 9和表 2 所示。

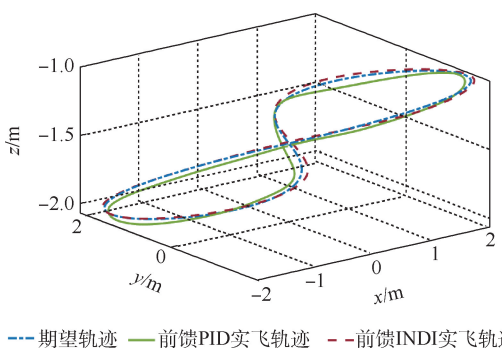


图 8　低速轨迹跟踪曲线

Fig. 8　Low speed trajectory tracking curves

表 1　低速轨迹跟踪效果

Table 1　Low speed trajectory tracking performance

参数	前馈 PID 方法	前馈 INDI 方法
位置均方根误差/cm	8.091	4.232
偏航角均方根误差/(°)	2.735	0.433
最大速度/(m·s^{-1})	3.769	4.116
平均速度/(m·s^{-1})	1.944	2.086
最大加速度/(m·s^{-2})	8.855	9.192
平均加速度/(m·s^{-2})	3.864	4.122

当参考轨迹的速度和加速度进一步提高时,前馈 PID 方法无法跟踪,而前馈 INDI 方法仍然能保持很高的精度。采用不同的角加速度估计方法跟踪最大速度 10.18 m/s、最大加速度 15.36 m/s^2 的参考轨迹时,跟踪结果如图 10 所示,飞行过程中的位置误差、偏航角误差、速度和加速度曲线如图 11所示,试验数据如表 3 所示。

跟踪过程中,基于互补滤波的前馈 INDI 控制方法的跟踪误差仅为 10.310 cm,偏航角的均方根误差为 2.645°。飞行器的最大速度为 10.416 m/s,平均速度为 4.654 m/s,最大加速度为 17.119 m/s^2,平均加速度为 7.434 m/s^2。可见设

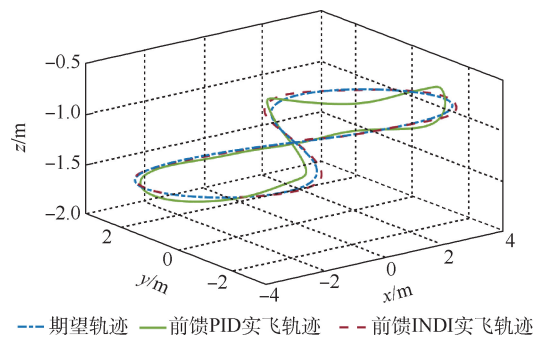


图 9　中速轨迹跟踪曲线

Fig. 9　Medium speed trajectory tracking curves

表 2　中速轨迹跟踪效果

Table 2　Medium speed trajectory tracking performance

参数	前馈 PID 方法	前馈 INDI 方法
位置均方根误差/cm	17.24	6.021
偏航角均方根误差/(°)	2.042	1.972
最大速度/(m·s^{-1})	6.233 1	6.425
平均速度/(m·s^{-1})	2.414 8	2.601
最大加速度/(m·s^{-2})	16.097 5	13.839
平均加速度/(m·s^{-2})	5.299 5	4.919

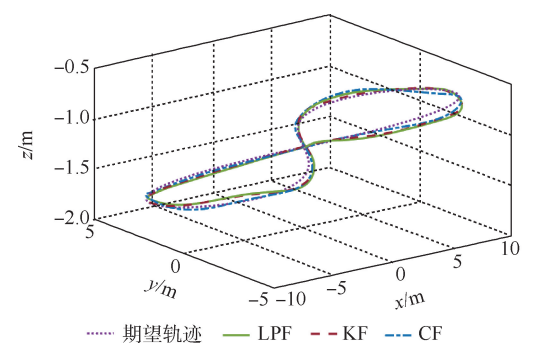


图 10　高机动轨迹跟踪曲线

Fig. 10　Aggressive trajectory tracking curves

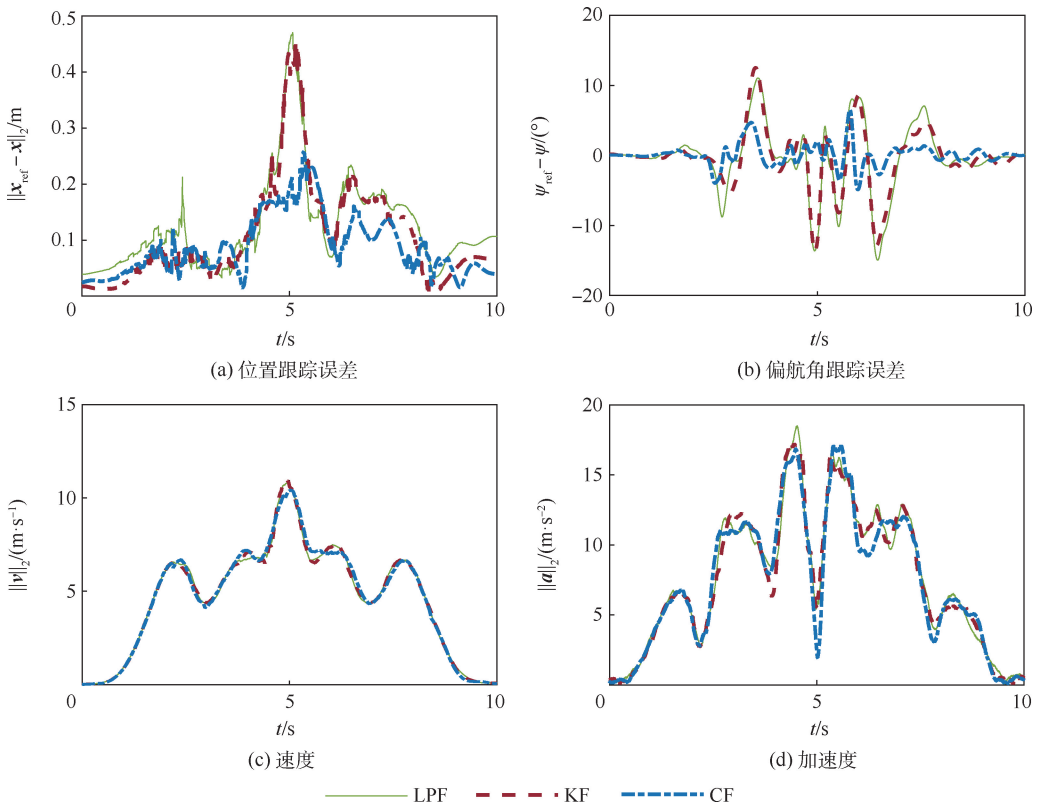


图 11 高机动轨迹跟踪曲线

Fig. 11 Aggressive trajectory tracking curves

表 3 高机动轨迹跟踪效果

Table 3 Aggressive trajectory tracking performance

参数	LPF	KF	CF
位置均方根误差/cm	15.072	14.084	10.310
偏航角均方根误差/(°)	4.751	4.298	1.552
平均速度/(m·s^{-1})	4.654	4.643	4.654
最大速度/(m·s^{-1})	10.815	10.922	10.416
平均加速度/(m·s^{-2})	7.604	7.555	7.434
最大加速度/(m·s^{-2})	18.510	17.239	17.119

计的基于互补滤波的前馈 INDI 控制器可以跟踪高机动轨迹,有着较高的精度。

3.2 抗扰试验

3.2.1 内部干扰

在飞机下方悬挂如图 12 所示面积大于其 4 倍的硬纸板,增大其运动时的空气阻力,增大其与仿真模型的不匹配度,以测试该控制器对内部模型误差干扰的鲁棒性。

分别用微分平坦前馈 PID 控制器和本文所设计的控制器控制四旋翼跟踪最大速度 4.07 m/s、最大加速度 8.45 m/s^2 的轨迹,效果如图 13 所示,飞行过程中的位置误差、偏航角误差、速度和加速度曲线如图 14 所示。

前馈 PID 控制器跟踪该轨迹时,均方根误差为 8.091 cm,但加入纸板后,均方根误差变为 10.00 cm,精度有所下降。前馈 INDI 控制器跟踪

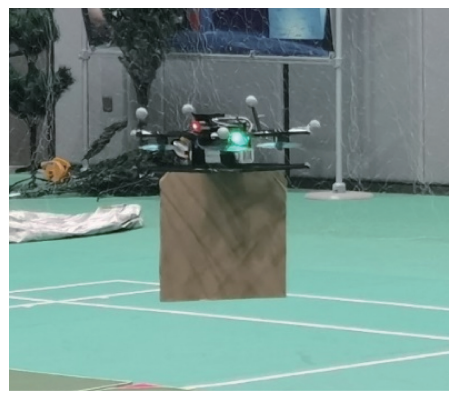


图 12 悬挂纸板的四旋翼无人机

Fig. 12 Quadrotor with cardboard drag plate

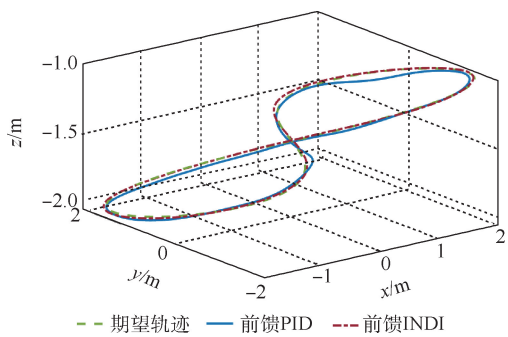


图 13 带纸板轨迹跟踪曲线

Fig. 13 Tajectory tracking curves with cardboard drag plate

轨迹时,均方根误差为 4.232 cm,加入纸板后为 4.494 cm,变化不明显。其他参数如表 4 所示。可

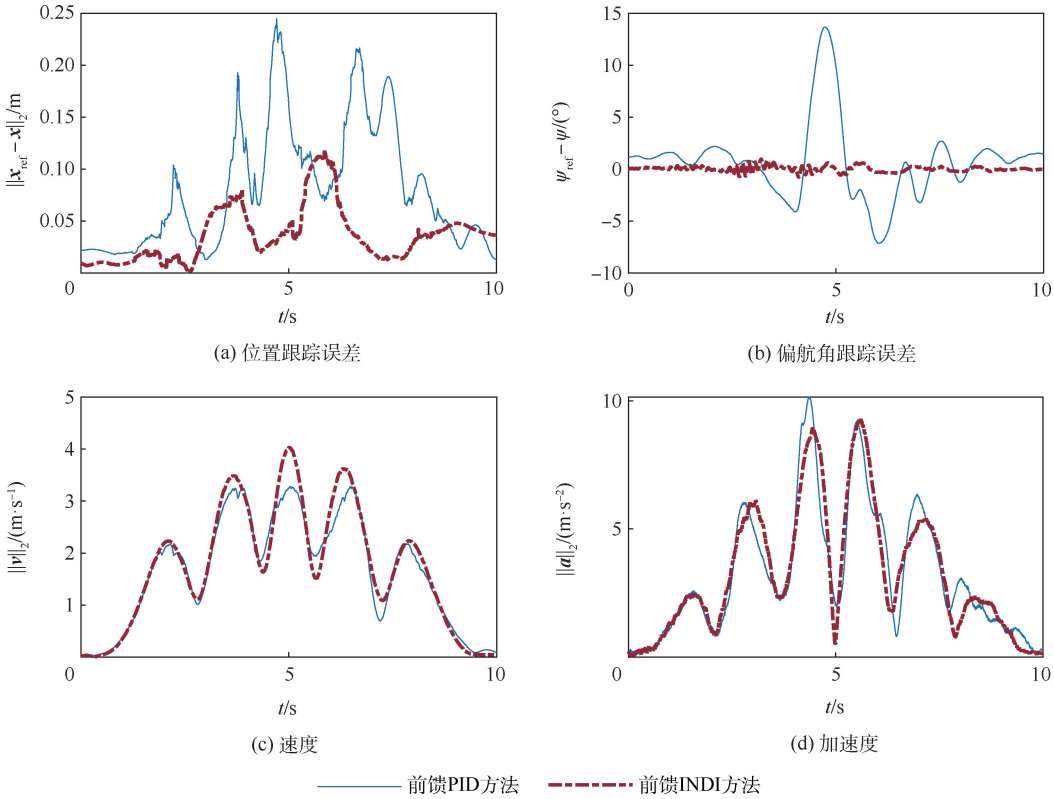


图 14　带纸板轨迹跟踪曲线

Fig. 14　Trajectory tracking curves with cardboard drag plate

表 4　轨迹跟踪效果

Table 4 Trajectory tracking performance

参数	前馈 INDI 方法(不带纸板)	前馈 INDI 方法(带纸板)	前馈 PID 方法(不带纸板)	前馈 PID 方法(带纸板)
位置均方根误差/cm	4.232	4.494	8.091	10.00
偏航角均方根误差/(°)	0.433	0.460	2.735	6.592
最大速度/(m·s^{-1})	4.116	4.034	3.769	3.283
平均速度/(m·s^{-1})	2.086	2.081	1.944	1.950
最大加速度/(m·s^{-2})	9.192	9.299	8.855	10.147
平均加速度/(m·s^{-2})	4.122	4.096	3.864	4.078

见,前馈 INDI 控制器不仅精度更高,对模型误差的容忍性也更强,对内部干扰有着优秀的抗扰能力。

3.2.2　外界干扰

为了验证前馈 INDI 控制器对外界干扰的抗扰能力,将功率为 280 W 的华生 FL750 工业风扇置于飞机悬停点前方 2 m 处(见图 15),用风速仪测得此处风速约为 5 m/s,进行悬停抗扰试验,得到的悬停效果如图 16 所示。

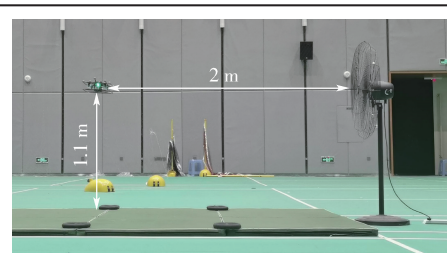
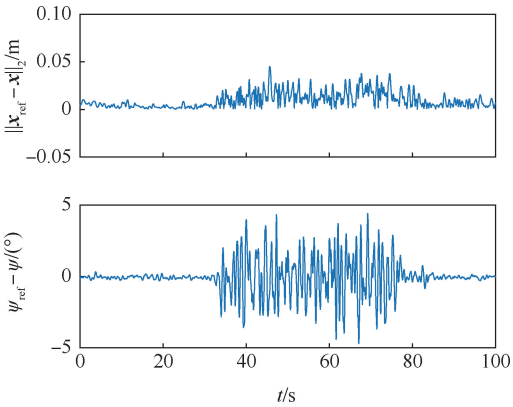


图 15　悬停抗扰试验

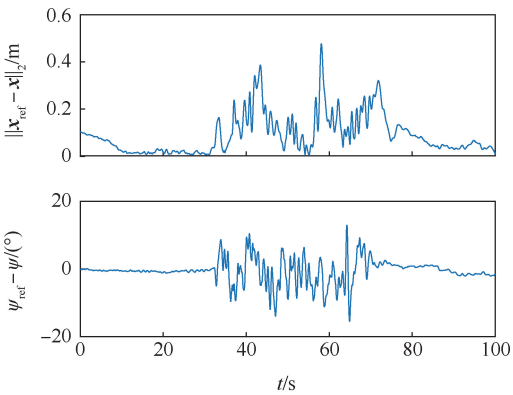
Fig. 15　Hover test with disturbance

前馈 PID 控制器控制四旋翼进行悬停时,风扇启动后位置和偏航角误差都显著增大,位置最大误差达 12.273 cm。前馈 INDI 控制悬停加入风干扰后,位置误差虽有所增大,但位置误差最大时也仅为 4.502 cm,风干扰对其悬停的影响不大。可见对于外界干扰,前馈 INDI 方法的鲁棒性明显优于前馈 PID 方法。其他参数如表 5 所示。

综上所述,不论是对于建模不准确引起的内部干扰,还是外界阵风干扰,前馈 INDI 方法都表现出了优秀的抗扰能力。



(a) 前馈INDI方法



(b) 前馈PID方法

图 16 悬停抗扰曲线

Fig. 16 Curves for hover with disturbance

表 5 悬停抗扰效果

Table 5 Performance for hover with disturbance

参数	前馈 INDI 方法	前馈 PID 方法
位置均方根误差/cm	1.199	12.273
位置最大误差/cm	4.502	47.670
偏航角均方根误差/(°)	0.018 9	0.059 6
偏航角最大误差/(°)	0.081 9	0.269

4 结 论

1）设计的前馈 INDI 控制器在跟踪高机动轨迹时具有较高的控制精度，在跟踪最大速度为 10.416 m/s，最大加速度为 17.119 m/s^2，平均速度为 4.654 m/s，平均加速度为 7.434 m/s^2 的高机动轨迹时，均方根误差仅为 10.310 cm。

2）该控制器对内部扰动具有较强的抗扰能力，加入纸板增大模型误差后，跟踪最大速度为 4.034 m/s，最大加速度为 9.299 m/s^2 的轨迹时，均方根误差由 4.232 cm 变为 4.494 cm，无明显增大。

3）该控制器对外界风干扰具有较强的抗干扰能力，令控制器控制飞行器保持悬停，在其悬停正前方 2 m 处引入风速为 5 m/s 的风干扰，位置均方根误差仍能保持在 1.199 cm。

飞行试验表明，本文设计的基于互补滤波的前馈 INDI 轨迹跟踪控制器可以控制飞行器精准、快速跟踪参考轨迹，在高机动飞行时仍保持极高的控制精度，对外界风干扰和内部模型误差有着较强的抗扰能力。

参考文献（References）

[1] ROOT P. Fast lightweight autonomy (FLA) (Archived)[EB/OL]. (2020-04-08)[2022-04-01]. https://www.darpa.mil/program/fast-lightweight-autonomy.

[2] 吕震华，高亢. 美国无人集群城市作战应用发展综述[J]. 中国电子科学研究院学报，2020，15(8)：738-745.
LV Z H，GAO K. Review of the development of drone swarm urban combat applications in the USA[J]. Journal of CAEIT，2020，15(8)：738-745(in Chinese).

[3] BYRNES C I，ISIDORI A. Nonlinear control systems[J]. Lecture Notes in Control and Information Sciences，2003，242(65)：408.

[4] SNELL S A，NNS D F，ARRARD W L. Nonlinear inversion flight control for a supermaneuverable aircraft[J]. Guidance，2015，15(4)：976-984.

[5] LEE D，KIM H J，SASTRY S. Feedback linearization vs. adaptive sliding mode control for a quadrotor helicopter[J]. International Journal of Control Automation and Systems，2009，7(3)：419-428.

[6] SIEBERLING S，CHU Q P，MULDER J A. Robust flight control using incremental nonlinear dynamic inversion and angular acceleration prediction[J]. Journal of Guidance，Control，and Dynamics，2010，33(6)：1732-1742.

[7] SIMPLÍCIO P，PAVEL M D，VAN KAMPEN E，et al. An acceleration measurements-based approach for helicopter nonlinear flight control using incremental nonlinear dynamic inversion[J]. Control Engineering Practice，2013，21(8)：1065-1077.

[8] LU P，VAN KAMPEN E J，DE VISSER C，et al. Aircraft fault-tolerant trajectory control using incremental nonlinear dynamic inversion aircraft fault-tolerant trajectory control using incremental nonlinear dynamic inversion[J]. Control Engineering Practice，2016，57：126-141.

[9] SMEUR E J，CHU Q P，CROON G. Adaptive incremental nonlinear dynamic inversion for attitude control of micro aerial vehicles：AIAA 2016-1390[R]. Reston：AIAA，2016.

[10] SMEUR E J J，DE CROON G C H E，CHU Q. Cascaded incremental nonlinear dynamic inversion for MAV disturbance rejection[J]. Control Engineering Practice，2018，73：79-90.

[11] TAL E，KARAMAN S. Accurate tracking of aggressive quadrotor trajectories using incremental nonlinear dynamic inversion and differential flatness[J]. IEEE Transactions on Control Systems Technology，2021，29(3)：1203-1218.

[12] ZHAO K，ZHANG J，MA D，et al. Composite disturbance rejection attitude control for quadrotor with unknown disturbance[J]. IEEE Transactions on Industrial Electronics，2020，67(8)：6894-6903.

[13] WANG Z，ZHAO J，CAI Z，et al. Onboard actuator model-based

incremental nonlinear dynamic inversion for quadrotor attitude control:Method and application[J]. Chinese Journal of Aeronautics,2021,34(11):216-227.

[14] JI C H,KIM C S,KIM B S. A hybrid incremental nonlinear dynamic inversion control for improving flying qualities of asymmetric store configuration aircraft[J]. Aerospace,2021,8(5):23.

[15] CHEN G,LIU A,HU J,et al. Attitude and altitude control of unmanned aerial-underwater vehicle based on incremental nonlinear dynamic inversion[J]. IEEE Access,2020,8:156129-156138.

[16] OVASKA S J,VALIVIITA S. Angular acceleration measurement:A review[J]. IEEE Transactions on Instrumentation and Measurement,1998,47(5):1211-1217.

[17] BUBNOV A V,EMASHOV V A,CHUDINOV A N,et al. Measurement methods for angular acceleration and errors for angular velocity of synchrophase electric drive[J]. Measurement Techniques,2014,57(8):860-865.

[18] CHENG S,FU M,WANG M,et al. Modeling for fluid transients in liquid-circular angular accelerometer[J]. IEEE Sensors Journal,2017,17(2):267-273.

[19] MOKHTARI E,ELHABIBY M,SIDERIS M G. Wavelet spectral techniques for error mitigation in the superconductive angular accelerometer output of a gravity gradiometer system[J]. IEEE Sensors Journal,2017,17(12):3782-3793.

[20] ALROWAIS H,GETZ P,KIM M G,et al. Bio-inspired fluidic thermal angular accelerometer[C]// IEEE International Conference on Micro Electro Mechanical Systems. Piscataway:IEEE Press,2016:15821613.

[21] COSTELLO M,JITPRAPHAI T. Determining angular velocity and angular acceleration of projectiles using triaxial acceleration measurements[J]. Journal of Spacecraft and Rockets,2002,39(1):73-80.

[22] VALIVIITA S,OVASKA S J. Delayless acceleration measurement method for elevator control[J]. IEEE Transactions on Industrial Electronics,1998,45(2):364-366.

[23] CHOUKROUN D,BAR-ITZHACK I Y,OSHMAN Y. Novel quaternion Kalman filter[J]. IEEE Transactions on Aerospace and Electronic Systems,2006,42(1):174-190.

[24] VALENTI R G,DRYANOVSKI I,XIAO J. A linear Kalman filter for MARG orientation estimation using the algebraic quaternion algorithm[J]. IEEE Transactions on Instrumentation and Measurement,2016,65(2):467-481.

[25] BELANGER P R,DOBROVOLNY P,HELMY A,et al. Estimation of angular velocity and acceleration from shaft encoder measurements[J]. International Journal of Robotics Research,2002,17(11):1225-1233.

[26] HAN J D,HE Y Q,XU W L. Angular acceleration estimation and feedback control: An experimental investigation[J]. Mechatronics,2007,17(9):524-532.

[27] MELLINGER D,KUMAR V. Minimum snap trajectory generation and control for quadrotors[C]// IEEE International Conference on Robotics and Automation. Piscataway:IEEE Press,2011:2520-2525.

Anti-disturbance trajectory tracking control method for aggressive quadrotors

WANG Yingxun, SONG Xinyu, ZHAO Jiang, CAI Zhihao*

(School of Automation Science and Electrical Engineering, Beihang University, Beijing 100083, China)

Abstract: It is difficult for the traditional quadrotor control method with poor control effect in aggressive flight and low control accuracy to track trajectory with high speed and high acceleration. To solve this problem, a control method is proposed based on incremental nonlinear dynamic inversion (INDI) and differential flatness, as well as the complementary filter. The proposed control method not only improves the tracking accuracy of aggressive trajectories, but also enhances the anti-disturbance ability. Since the angular acceleration, which the proposed method is very sensitive to, cannot be directly obtained, a variety of angular acceleration estimation methods are designed for comparison, and the complementary filtering method with the best performance is selected through the flight test. The experimental results show that the proposed control method using INDI and differential flatness based on complementary filter can control the quadrotor to track aggressive trajectories quickly and accurately, and has strong anti-interference ability.

Keywords: quadrotor; trajectory tracking; incremental nonlinear dynamic inversion (INDI); disturbance rejection control; angular acceleration estimation

Received: 2022-04-02; **Accepted**: 2022-05-26; **Published online**: 2022-06-14 11:49

URL: kns.cnki.net/kcms/detail/11.2625.V.20220613.1819.005.html

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0609

飞行过程中头盔对飞行员颈部损伤的影响

王丽珍[1,2,3], 刘景龙[1,2,3], 赵彦鹏[4], 卜伟平[4], 柳松杨[4], 樊瑜波[1,2,3,5,*]

(1. 北京航空航天大学 生物与医学工程学院, 北京 100083;
2. 北京航空航天大学 北京市生物医学工程高精尖创新中心, 北京 100083;
3. 北京航空航天大学 生物力学与力生物学教育部重点实验室, 北京 100083;
4. 空军特色医学中心, 北京 100142; 5. 北京航空航天大学 医学科学与工程学院, 北京 100083)

摘 要: 为研究飞行过程中不同头盔对飞行员颈部损伤的影响,基于头颈部多刚体动力学模型,对急转弯及稳定盘旋飞行工况下4种头盔情况开展了动力学分析。得到了不同头盔作用下颈部肌肉力、椎间力、力矩及颈部损伤指数 N_{ij} 的变化规律,对不同头盔作用下的颈部损伤风险进行了评估。结果表明:所建立的头颈部模型与实验结果较为吻合,可以用来进行仿真模拟分析;相同飞行工况下,B头盔的佩戴使得头部质心前移,增加了颈部前屈时斜方肌的肌肉力;对于同一颈椎节段,B头盔的轴向压缩力最大,而C头盔导致的后伸力矩最大;在所有的计算工况中,上下颈椎的颈部损伤指数 N_{ij} 均小于0.5,满足航空领域的安全要求;B头盔和C头盔的质量相近,而B头盔质心位置靠前,C头盔靠后,说明头盔质心的前后位置会对颈部的损伤产生重要影响。研究结果可为飞行员头盔的设计优化及其对颈部损伤的影响提供研究支撑。

关 键 词: 飞行员; 头盔; 头颈部; 多刚体动力学模型; 损伤分析
中图分类号: R318.01; TP319.9
文献标志码: A **文章编号**: 1001-5965(2022)09-1818-09

颈部疼痛是困扰飞行员的一种常见疾病[1]。在一项针对超过8 000名飞行员的统计分析中,飞行员颈部疼痛的发生率高达51%[2]。各国的研究显示,丹麦空军在飞行中及飞行后经历颈部疼痛的飞行员比例甚至高达97%[3],挪威空军为72%[4],中国战斗机飞行员颈部疼痛的比例也达33.7%[5]。颈部肌肉的急性拉伤[6]、韧带撕裂[7]、椎间盘突出[8]甚至椎体的骨折[9]都会引发颈部疼痛,造成颈部损伤的原因有很多,包括头盔的佩戴[10]、头颈部姿势[11]、飞行时长[12]、飞行工况[13]等因素,尤其在实际飞行过程中,多种因素的共同作用加剧了颈部损伤风险。

头盔的佩戴能够保护飞行员头部免受损伤,为了增强飞行员的作战效能,往往会在头盔上集成显示器等装置,这导致了不同头盔的质量及质心位置不同。研究显示,头盔质量的增加和质心位置的改变,增加了飞行过程中的颈部载荷,加大了颈部的损伤风险[10]。自20世纪80年代瑞典空军开展头盔对颈部损伤影响的研究以来,国内外开展了大量的相关研究,主要包括流行病学调查、实验研究和数值仿真研究。Sovelius等[14]利用蹦床实验,测量了飞行员佩戴头盔在不同加载

收稿日期: 2022-07-12; **录用日期**: 2022-08-02; **网络出版时间**: 2022-08-11 10:30
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220809.1527.001.html
基金项目: 国家自然科学基金(11822201); 空军军医大学飞行人员效能提升计划(2019ZTA04)
* **通信作者**. E-mail: yubofan@buaa.edu.cn

引用格式: 王丽珍, 刘景龙, 赵彦鹏, 等. 飞行过程中头盔对飞行员颈部损伤的影响[J]. 北京航空航天大学学报, 2022, 48(9): 1818-1826. WANG L Z, LIU J L, ZHAO Y P, et al. Effect of helmet on neck injury of pilots in flight[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1818-1826(in Chinese).

下颈部的肌电信号,发现头盔质量的增加,增大了颈部肌肉的拉力。吴明磊等[15]利用志愿者和 HYBRID Ⅲ假人实验,分析了不同弹射条件下头盔对人体头颈响应的影响,发现头盔对颈部载荷的影响随着加速度的增加而增加。贾晓红等[16]利用 AnyBody 软件平台建立头颈部肌骨模型,分析了不同头盔质量、质心位置及加速度对颈部肌肉力的影响,发现头盔质量的增加导致肌肉力的增加,而质心的后移会增加前伸肌肉的作用力。Mathys 和 Ferguson[17]也利用头颈部模型分析了不同头盔在特定载荷下对颈部损伤的影响,同一工况下下颈椎的载荷较大。包佳仪等[18]也建立了飞行员头颈部有限元模型,分析了阻拦着舰过程中飞行员颈部的损伤。然而以上关于飞行员颈部损伤的相关研究中,主要集中在较短时间内的恒定载荷作用下颈部的损伤分析,对不同飞行工况下不同头盔对飞行员颈部损伤影响的研究还较为缺乏。

因此,本文采用头颈部多刚体动力学模型,对急转弯及稳定盘旋飞行工况下,不同头盔对飞行员颈部损伤的影响进行了仿真研究,通过肌肉力、椎间力及颈部损伤准则研究颈部损伤情况,分析飞行过程中头盔对飞行员颈部损伤的影响,研究结果可为飞行员头盔的设计优化及其颈部损伤的影响提供仿真数据支撑。

1　模型的建立与验证

1.1　模型的建立

本文中的人体头颈部多刚体动力学模型是基于笔者课题组的研究建立的[19]。基于 1 名中国成年男性头颈部的 CT 图像,通过逆向建模得到头颈部的几何模型,导入多刚体动力学建模软件 ADAMS 中建立多刚体模型。模型包括头骨、7 个颈椎、1 个胸椎及胸部和肩部的骨骼结构,这些结构均设为刚体。胸骨和肩骨用来给肌肉提供附着点,不考虑胸椎和胸骨与肩骨之间的相对运动,将三者连接成一个整体。

模型中的全局坐标系位于胸椎 T1 的中心,x 轴为人体前后向,y 轴为人体侧向,根据右手定则 z 轴为竖直方向,如图 1 所示。

参考 de Jager[20]设置模型中头骨、椎体的质量特性,建立包含韧带、椎间盘和关节共同作用的集总参数关节,通过力和位移及速度之间的关系

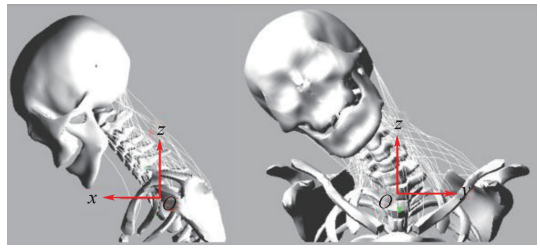


图 1　多刚体动力学模型及坐标系

Fig. 1　Multi-body dynamic model and definition of coordinates

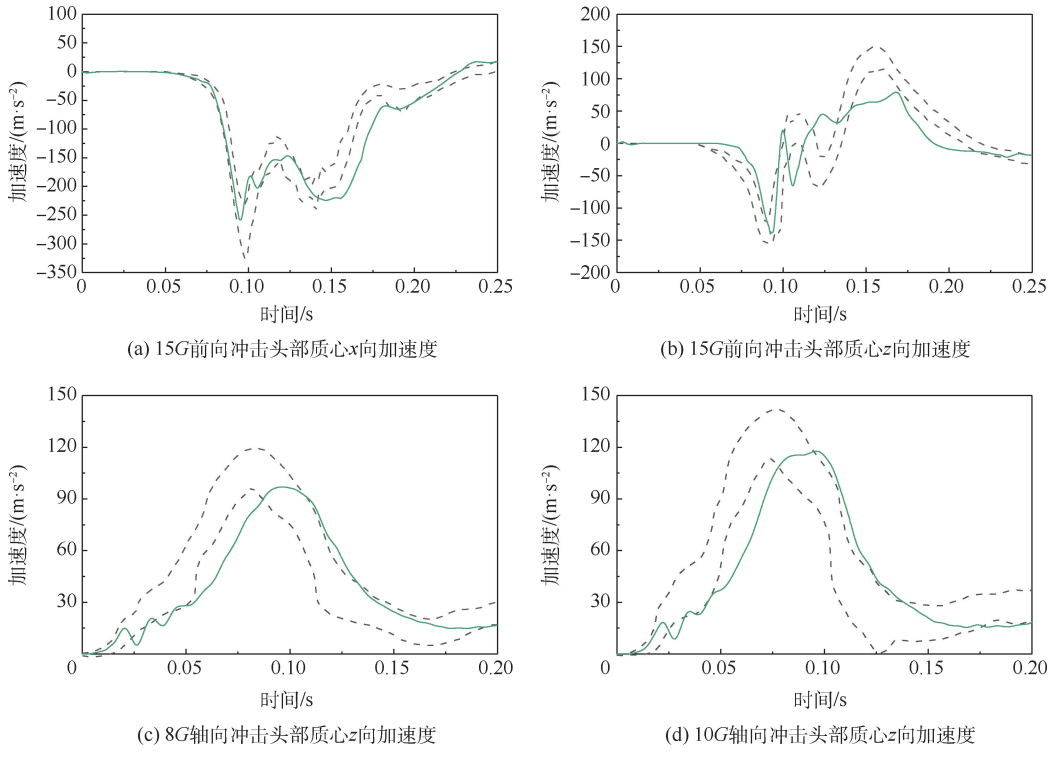
对集总关节进行定义,忽略了椎间盘、韧带和肌肉的质量[19-22]。基于解剖位置,采用“绳索”单元建立了 15 组颈部肌肉,在肌肉起止点之间建立滑动分割点,表示肌肉的弯曲效应,如图 1 所示,使用 Hill 三元素模型表示肌肉的收缩作用[23],所建多刚体模型如图 1 所示。

1.2　模型的验证

使用关于志愿者的加速度峰值为 15G 的前向冲击数据[24]及针对志愿者的 8G 和 10G 沿脊柱方向的冲击数据[25]对模型进行了验证。将验证输入的运动学数据施加在模型的第一胸椎 T1 处,开展动力学计算。提取模型头部质心处的 x 向和 z 向加速度与志愿者的实验数据进行对比,结果如图 2 所示。

图 2(a)、(b)展示了模型及实验的前向冲击结果,可以看出,前向验证下头部质心 x 向和 z 向加速度能够较好地复现实验结果。图 2(c)、(d)为沿脊柱方向的冲击结果,可以看出,沿脊柱方向 8G 和 10G 冲击的 z 向冲击结果与实验数据有着相似的趋势。

使用相关性分析方法对模型仿真度进行定量评估,采用区间评估和截面评估来评价曲线相似度,表示模型的仿真度。区间评估用于评价仿真曲线在实验区间内的相似度,截面评估通过幅值、相位等评价曲线的相似度,可用来评估仿真假人或模型与实验数据的相似度[26]。通过总得分定量评估模型的仿真度,对 4 种验证结果分配了相同的权重计算总得分,得分在 0 ~ 0.26 表明模型是不合格无法接受的,0.26 ~ 0.44 表明模型仿真度尚可,0.44 ~ 0.65 表明模型仿真度合适,在 0.65 ~ 0.86 则被认为仿真度良好,本文模型的综合得分为 0.694,如表 1 所示,在 0.65 ~ 0.86 区间内,表明本文中的头颈部模型可以用来进行仿真计算。



(a) 15G前向冲击头部质心x向加速度　(b) 15G前向冲击头部质心z向加速度

(c) 8G轴向冲击头部质心z向加速度　(d) 10G轴向冲击头部质心z向加速度

– – – – – 实验数据上下限　——— 仿真结果

图2　模型仿真结果与实验数据对比

Fig. 2　Comparison of simulation results and experimental data

表1　头颈部模型的生物仿真度评估

Table 1　Biofidelity assessment of head and neck model

载荷工况	验证变量	区间评估	截面评估	整体得分
15G 前向冲击	x 向加速度	0.779	0.833	0.806
	z 向加速度	0.429	0.626	0.528
8G 轴向冲击	z 向加速度	0.651	0.840	0.726
10G 轴向冲击	z 向加速度	0.618	0.864	0.716
整个模型				0.694

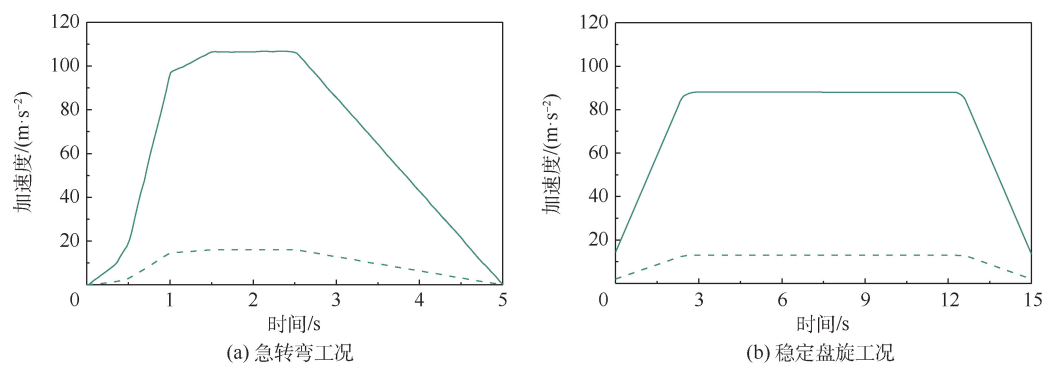
2　仿真计算及结果分析

2.1　计算工况

本文采用2种输入工况,分别为急转弯工况和稳定盘旋工况。提取飞行实验中2种工况下飞行员胸椎处的加速度载荷曲线,如图3所示。本文评价3种头盔在2种飞行工况下对颈部损伤的影响,采用无头盔情况作为对照组,共计4种头盔情况。添加头盔后,按照重心分布及平行移轴定理,得头部的整体质心及转动惯量,如表2所示。将2种飞行工况的载荷曲线施加在4种头盔模型的第一胸椎处开展动力学计算,共计8种计算工况。

2.2　肌肉力分析

分别提取2种飞行工况下4种头盔情况对应的头夹肌、斜方肌及集总舌肌的肌肉力,如图4和图5所示。在2种飞行工况下,无头盔和B头盔情况的头夹肌和斜方肌肌肉力较大,对应的集总舌肌肌肉力很小,而A头盔和C头盔情况的集总



(a) 急转弯工况　(b) 稳定盘旋工况

– – – – – x向加速度　——— z向加速度

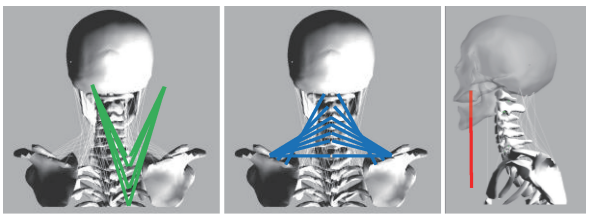
图3　飞行中加速度时间曲线

Fig. 3　Acceleration-time curves in-flight

表 2　头部质量特性

Table 2　Mass properties of head

头盔工况	质量/kg	重心位置/m		转动惯量/(kg·m²)	
		x 向	z 向	x 向	z 向
无头盔	4.16	0.027	0.021	0.018 1	0.023 6
A 头盔	5.16	0.022	0.025	0.031 5	0.035 7
B 头盔	5.75	0.035	0.036	0.056 1	0.053 0
C 头盔	5.80	0.021	0.038	0.043 1	0.050 9



(a) 头夹肌　(b) 斜方肌　(c) 集总舌肌

图 4　三组肌肉的示意图

Fig. 4　Schematic diagram of three muscle groups

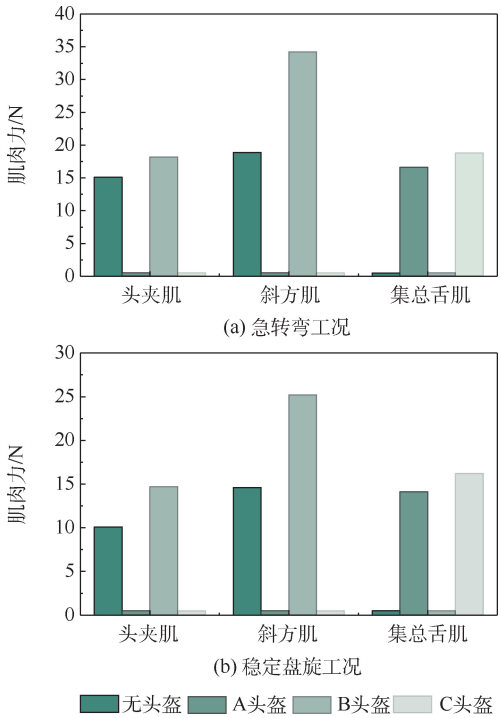


图 5　不同头盔工况下的最大肌肉力

Fig. 5　Maximum muscle force for different helmet conditions

舌肌肌肉力较大,对应的头夹肌和斜方肌的肌肉力较小。头夹肌和斜方肌为屈肌,位于颈部后侧,在头部前倾时被拉长,而集总舌肌为伸肌,在颈部前侧,在头部后仰时被拉长。表明在 2 种飞行工况下,无头盔和 B 头盔情况下头颈部均发生前倾,而 A 头盔和 C 头盔情况头颈部均发生后仰。此外,在 2 种飞行工况下,佩戴 B 头盔时,斜方肌肌肉力均最大,在急转弯工况下为 34.2 N,在稳定盘旋工况下为 25.2 N,均明显高于其他头盔情况,表明 B 头盔对颈部斜方肌具有较大影响,斜方肌更容易受到损伤。

2.3　椎间受力分析

通常,使用上下颈椎的最大作用力和弯矩对颈部损伤进行评估[27]。根据输入曲线特点,头颈部在人体矢状面内运动,椎体主要受到压力作用及前屈后伸的力矩作用。因此,提取各计算工况下头部枕骨大孔与第一颈椎(OC-C1)之间和第七颈椎与第一胸椎(C7-T1)之间的轴向压力和前屈后伸力矩的时间曲线,如图 6 所示。此处取轴向压缩及前屈力矩为正,轴向拉伸及后伸力矩为负。

图 6(a)、(c)显示,急转弯工况下 OC-T1 与 C7-T1 之间的轴向压力具有相似的变化趋势,与输入曲线中的 z 向载荷趋势相似。稳定盘旋工况中,两颈椎节段的轴向压力也具有相似的变化趋势。对于急转弯工况下的同一颈椎节段,B 头盔情况下的轴向压力高于其他头盔情况。图 6(b)、(d)显示,急转弯工况下无头盔和 B 头盔情况的力矩变化趋势相似,且均为正值,表明颈部发生前屈;而 A 头盔和 C 头盔情况的力矩变化趋势相似,均为负值,表明颈部发生后仰。图 6(e)、(g)显示,针对稳定盘旋工况,OC-C1 与 C7-T1 之间的轴向压力也具有相似的变化趋势,对于同一节段,B 头盔情况的轴向压力最大。图 6(f)、(h)显示,稳定盘旋下,两节段在无头盔和 B 头盔均为前屈力矩,而 A 头盔和 C 头盔为后伸力矩。由表 2 可得,无头盔和 B 头盔情况下,头部质心的整体位置靠前,可能导致头部更容易发生前倾。

表 3 展示了 OC-C1 及 C7-T1 节段的轴向压力和前屈后伸力矩极值及相应的损伤评估参考值[28],超过损伤评估值,则认为相应椎体发生严重损伤。取轴向压力及前屈力矩为正,拉伸力及后伸力矩为负。在2 种飞行工况佩戴不同头盔的情况下,轴向发生压缩,压缩力均小于临界值。其中,急转弯工况佩戴 B 头盔的情况下,C7-T1 节段的轴向压力最大,为 1 141.6 N,小于临界值 4 000 N;前屈力矩也最大,为 72.2 N·m,小于临界值 190 N·m。同样,急转弯工况中佩戴 C 头盔的情况下,C7-T1 节段的后伸力矩最大为 -80.1 N·m,小于临界值 -192 N·m。表 3 中的结果表明,2 种飞行工况中,佩戴 4 种头盔情况下均不会造成颈部的严重损伤。表 2 显示,B 头盔和 C 头盔质量相近,B 头盔质心位置更靠前,结合轴向压力结果显示,质心靠前增加了轴向压缩作用;C 头盔质心较为靠后,结合前屈后伸力矩结果,质心靠后可能会造成颈部后仰,增加后仰力矩。由表 3 可知,同一颈椎节段,后仰的损伤临界值小于前屈的临界值,意味着 C 头盔容易造成更大的损伤风险。

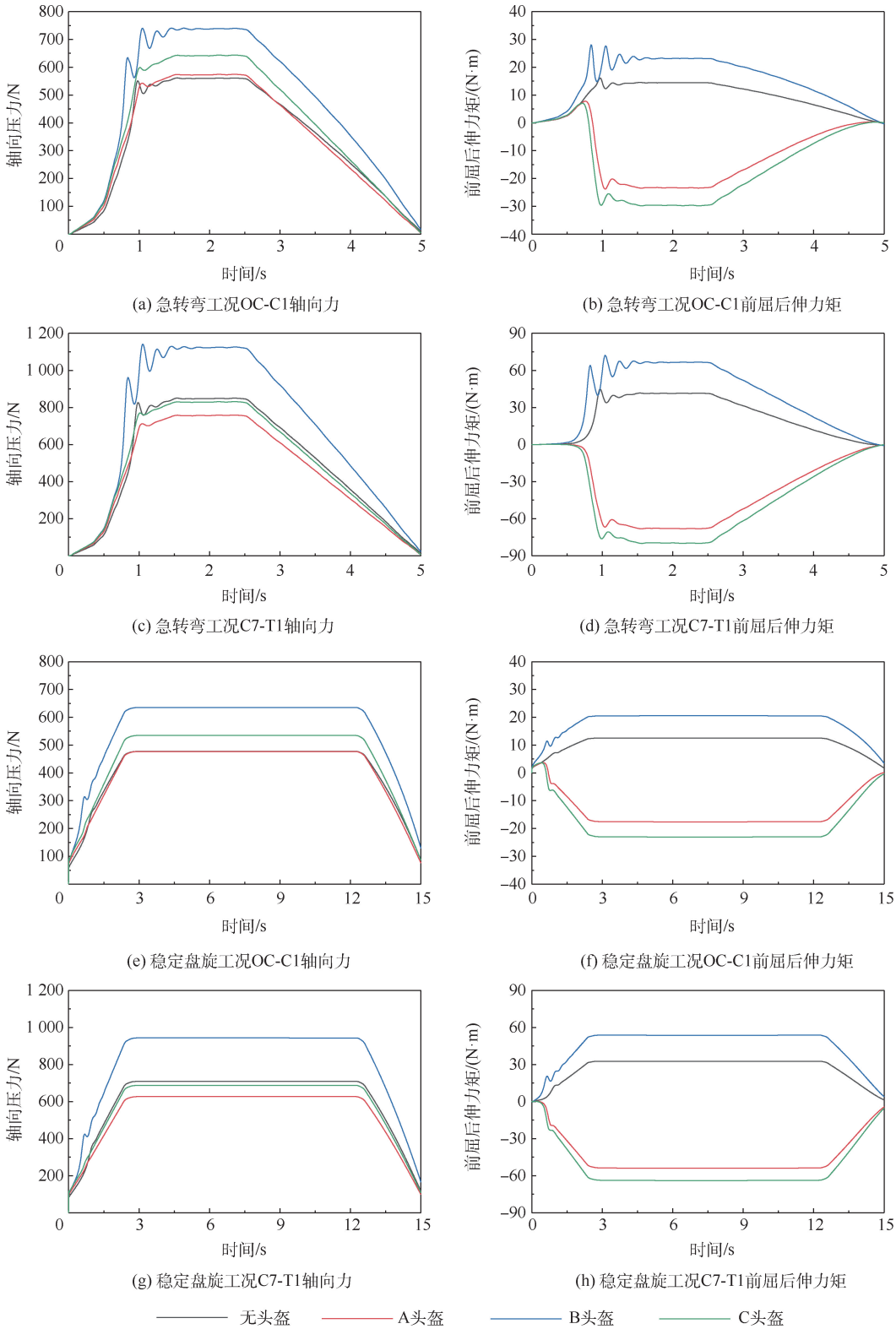


图6 相邻椎体间轴向压力和前屈后伸力矩随时间的变化曲线

Fig. 6 Axial force and flexion/extension moment time curves of adjacent segment

2.4 N_{ij}损伤风险分析

为了综合分析颈部轴向压力和前屈后伸力矩对颈部损伤的影响,使用颈部损伤N_{ij}准则进行定量评估。颈部损伤的N_{ij}准则是由美国高速公路安全管理局(NHTSA)提出的用于评价汽车前向碰撞中乘员颈部发生严重损伤的评估标准[29],计算公式为

$$N_{ij} = \left|\frac{F_z}{F_{\text{int}}}\right| + \left|\frac{M_y}{M_{\text{int}}}\right| \tag{1}$$

式中:F_z为轴向的压缩力或拉伸力;M_y为矢状面内的前屈后伸力矩;F_{int}为压缩/拉伸力的临界值,M_{int}为前屈/后伸力矩的临界值。

针对 OC-C1 及 C7-T1 颈椎节段,临界值取值如表 4 所示[28,30]。N_{ij}越大,表明颈部损伤风险越大,在航空领域,为了保证飞行员的安全不发生严重损伤,一般认为 N_{ij}不超过 0.5[31]。

表 3　OC-C1 及 C7-T1 节段的最大轴向力与前屈后伸力矩

Table 3　Maximum axial force and flexion/extension moment for OC-C1 and C7-T1 segment

载荷形式	临界值		计算工况	急转弯工况		稳定盘旋工况	
	OC-C1	C7-T1		OC-C1	C7-T1	OC-C1	C7-T1
轴向力(压缩力/拉伸力)/N	4 000/－4 170	4 000/－4 170	无头盔	561.65/－0.3	850.4/－0.3	447.4/0	709.4/0
			A 头盔	574.7/－0.4	758.6/－0.4	476.9/0	626.8/0
			B 头盔	740.7/－0.4	1 141.6/－0.5	635.3/0	944.0/0
			C 头盔	643.8/－0.4	831.4/－0.5	534.7/0	687.4/0
前屈后伸力矩(前屈力矩/后伸力矩)/(N·m)	190/－96	380/－192	无头盔	16.2/－0.3	45.0/－0.8	12.5/0	32.7/－0.2
			A 头盔	7.8/－23.8	0.1/－68.2	3.7/－17.6	0.3/－53.9
			B 头盔	28.1/－0.4	72.2/－1.0	20.6/0	53.8/－0.4
			C 头盔	7.0/－29.9	0.1/－80.1	3.7/－23.1	0.3/－64.0

表 4　OC-C1 和 C7-T1 节段 N_{ij}临界值的选取

Table 4　N_{ij} critical intercept values for OC-C1 and C7-T1 segment

N_{ij}临界值	OC-C1	C7-T1
轴向拉伸力/N	6 780	6 780
轴向压缩力/N	6 200	6 200
前屈力矩/(N·m)	305	610
后伸力矩/(N·m)	133	266

图 7 和图 8 分别显示了急转弯工况和稳定盘旋工况下 N_{ij}随时间的变化曲线,图 9 为各工况下 N_{ij}的最大值。可以看出,各工况 N_{ij}值均小于0.5,表明在 2 种飞行工况下,佩戴 4 种头盔均不会对颈部造成严重损伤。对于同一飞行工况,两节段的 N_{ij}随时间变化的曲线具有相似的趋势。

对于急转弯工况,同一头盔佩戴情况下 C7-T1 节段的 N_{ij}高于 OC-C1 节段,表明下颈椎损伤

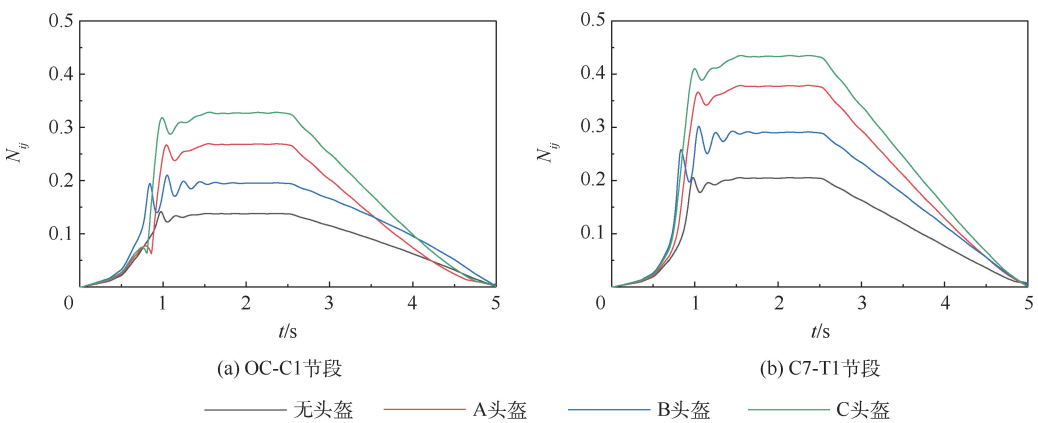


图 7　急转弯工况 N_{ij}随时间的变化曲线

Fig. 7　N_{ij} time curves under sharp turn condition

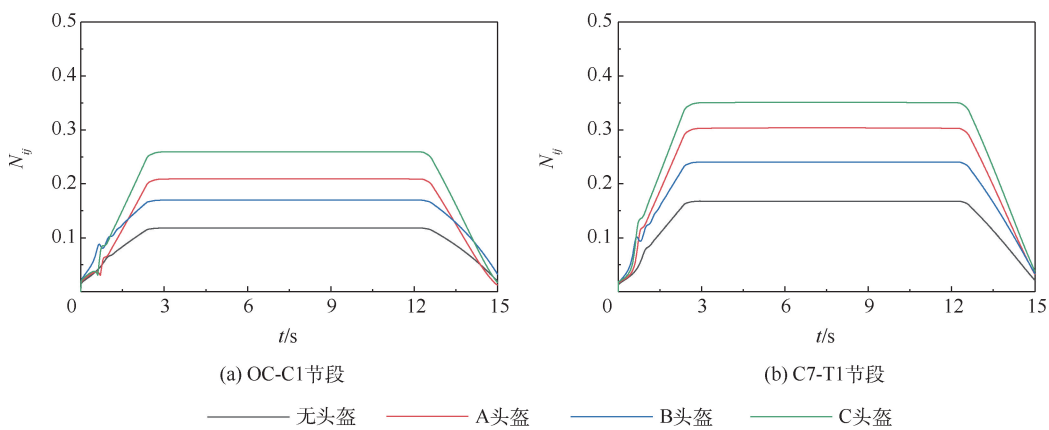


图 8　稳定盘旋工况 N_{ij}随时间的变化曲线

Fig. 8　N_{ij} time curves under steady hover condition

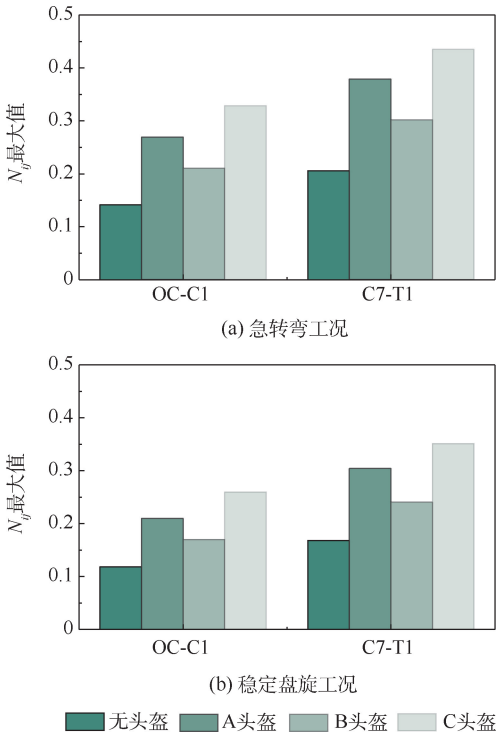


图9 N_{ij}的最大值

Fig. 9 Maximum N_{ij} values

风险较高;针对同一节段下,C头盔的N_{ij}最大,A头盔次之,B头盔最小。而在最大轴向压力及最大前屈后伸力矩的分析中,对于同一颈椎节段,B头盔的轴向压力最大,远高于其他头盔情况,而C头盔情况下的后伸力矩大于B头盔的前屈力矩。由表4可看出,前屈力矩的临界值约为后伸力矩的2倍,说明颈部在后伸情况下更易损伤,因此导致了C头盔情况下颈部具有最大的N_{ij}值。由表2可知,B头盔和C头盔质量相近,而C头盔质心靠后,显示质心的过度后移会增加颈部的损伤风险。

3 结 论

本文对急转弯及稳定盘旋飞行工况下,不同头盔对飞行员颈部损伤的影响进行了仿真模拟研究,得到了不同工况下颈部肌肉力、椎间力与力矩及颈部损伤N_{ij}的变化规律,分析了不同飞行工况下头盔对颈部损伤的影响,得到以下结论:

1) 在2种飞行工况下,B头盔的佩戴使得头部质心前移,均增加了颈部前屈时斜方肌的肌肉力,容易造成斜方肌的拉伤;A头盔和C头盔的佩戴使得质心后移,导致集总舌肌受到拉伸。

2) 对于特定头盔情况,下颈椎的轴向压缩力高于上颈椎;对于同一颈椎节段,B头盔的情况下轴向压缩力最大,C头盔导致的后伸力矩最大。

3) 在所有的飞行工况下,上下颈椎的N_{ij}均小于0.5,表明4种头盔情况均不会造成颈部严重损伤,满足航空领域的安全要求。

4) 对于同一飞行工况相同的颈椎节段,C头盔导致的N_{ij}最大,其次是A头盔情况,而B头盔导致的轴向压缩力大,N_{ij}却较小,这与颈部在前屈后伸过程中的损伤容限有关。B头盔和C头盔质量相近,而B头盔靠前,C头盔靠后,说明头盔质心的前后位置对颈部的损伤产生严重影响。

参考文献 (References)

[1] HONKANEN T, SOVELIUS R, MÄNTYSAARI M, et al. + Gz exposure and spinal injury-induced flight duty limitations[J]. Aerospace Medicine and Human Performance, 2018, 89(6): 522-556.

[2] RICHES A, SPATFORD W, WICHALLS J, et al. A systematic review and meta-analysis about the prevalence of neck pain in fast jet pilots[J]. Aerospace Medicine and Human Performance, 2019, 90(10): 882-890.

[3] LANGE B, TORP-SVENDSEN J, TOFT P. Neck pain among fighter pilots after the introduction of the JHMCS helmet and NVG in their environment[J]. Aviation Space and Environmental Medicine, 2011, 82(5): 559-563.

[4] WAGSTAFF A S, JAHR K I, RODSKIER S. + Gz-induced spinal symptoms in fighter pilots: Operational and individual associated factors[J]. Aviation Space and Environmental Medicine, 2012, 83(11): 1092-1096.

[5] 柳松杨,丛红,王鹤,等. 军机飞行员的颈部损伤研究[J]. 医用生物力学, 2010, 25(4): 262-265.
LIU S Y, CONG H, WANG H, et al. Study on neck injuries in military pilots[J]. Journal of Medical Biomechanics, 2010, 25(4): 262-265(in Chinese).

[6] POUSETTE M W, MARTIRE R L, LINDER J, et al. Neck muscle strain in air force pilots wearing night vision goggles[J]. Aerospace Medicine and Human Performance, 2016, 87(11): 928-932.

[7] ANDERSEN H T. Neck injury sustained during exposure to high-*G* forces in the F16b[J]. Aviation Space and Environmental Medicine, 1988, 59(4): 356-358.

[8] KESKIMOLO T, PERNU J, KARPPINEN J, et al. Degenerative cervical spine changes among early career fighter pilots: A 5-year follow-up[J/OL]. BMJ Military Health, 2021 (2021-06-15) [2022-07-01]. https: // pubmed. ncbi. nlm. nih. gov/34131064/.

[9] MANEN O, CLÉMENT J, BISCONTE S, et al. Spine injuries related to high-performance aircraft ejections: A 9-year retrospective study[J]. Aviation Space and Environmental Medicine, 2014, 85(1): 66-70.

[10] CHUMBLEY E M, STOLFI A, MCEACHEN J C. Risk factors for cervical pain in F-15c pilots[J]. Aerospace Medicine and Human Performance, 2017, 88(11): 1000-1007.

[11] HARRISON M F, FORDE K A, ALBERT W J, et al. Posture and helmet load influences on neck muscle activation[J]. Aerospace Medicine and Human Performance, 2016, 87(1): 48-53.

[12] TUCKER B, NETTO K, HAMPSON G, et al. Predicting neck pain in royal Australian air force fighter pilot[J]. Military Medicine, 2012, 177(4): 444-450.

[13] SHIRI R, FRILANDER H, SAINIO M, et al. Cervical and lumbar pain and radiological degeneration among fighter pilots: A systematic review and meta-analysis[J]. Occupational and Environmental Medicine, 2015, 72(2): 145-150.

[14] SOVELIUS R, OKSA J, RINTALA H, et al. Neck muscle strain when wearing helmet and NVG during acceleration on a trampoline[J]. Aviation Space and Environmental Medicine, 2008, 79(2): 112-116.

[15] 吴明磊,马春生,刘威,等. 弹射时装显示器头盔对人体的生物力学效应[J]. 中华航空航天医学杂志, 2005, 16(4): 267-271.
WU M L, MA C S, LIU W, et al. Biodynamic effect of helmet with mounted display system on human during ejection[J]. Chinese Journal of Aerospace Medicine, 2005, 16(4): 267-271 (in Chinese).

[16] 贾晓红,茆军兵,王人成,等. 头盔质量和质心对军机飞行员颈部肌力的影响[J]. 医用生物力学, 2012, 27(4): 416-420.
JIA X H, MAO J B, WANG R C, et al. Effect of helmet mass and mass center on neck muscle strength in military pilots[J]. Journal of Medical Biomechanics, 2012, 27(4): 416-420 (in Chinese).

[17] MATHYS R, FERGUSON S J. Simulation of the effects of different pilot helmets on neck loading during air combat[J]. Journal of Biomechanics, 2012, 45(14): 2362-2367.

[18] 包佳仪,王兴伟,周前祥,等. 阻拦着舰过程中飞行员颈部的损伤分析与预测[J]. 北京航空航天大学学报, 2019, 45(3): 499-507.
BAO J Y, WANG X W, ZHOU Q X, et al. Analysis and prediction of neck injury of pilots during carrier aircraft arrest deck-landing[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(3): 499-507 (in Chinese).

[19] WANG Y W, WANG L Z, LIU S Y, et al. A two-step procedure for coupling development and usage of a pair of human neck models[J]. Computer Methods in Biomechanics and Biomedical Engineering, 2018, 21(5): 413-426.

[20] DE JAGER M K J. Mathematical head-neck models for acceleration impacts[D]. Eindhoven: Eindhoven University of Technology, 1996.

[21] DENG Y C, GOLDSMITH W. Response of a human head/neck/upper-torso replica to dynamic loading-II. Analytical/numerical model[J]. Journal of Biomechanics, 1987, 20(5): 487-497.

[22] HORST M. Human head neck response in frontal, lateral and rear end impact loading[D]. Eindhoven: Eindhoven University of Technology, 2002.

[23] PANZER M B, FICE J B, CRONIN D S. Cervical spine response in frontal crash[J]. Medical Engineering and Physics, 2011, 33(9): 1147-1159.

[24] THUNNISSEN J, WISMANS J, EWING C, et al. Human volunteer head-neck response in frontal flexion: A new analysis[J]. Stapp Car Crash Journal, 1995, 39: 3065-3086.

[25] JACOB B P, JEFFREY T S, JESSICA A W, et al. Development and evaluation of a finite element model of the THOR for occupant protection of spaceflight crewmembers[J]. Accident Analysis and Prevention, 2015, 82: 244-256.

[26] GEHRE C, GADES H, WERNICKE P. Objective rating of signals using test and simulation responses[C]// Proceeding of 21st International Technical Conference on the Enhanced Safety of Vehicles, 2009.

[27] 柯鹏,诸斌,柳松杨,等. 拦阻着舰过程中不同百分位人体颈部损伤风险的差异对比[J]. 医用生物力学, 2018, 33(5): 383-389.
KE P, ZHU B, LIU S Y, et al. Differences of neck injury risks for different dummies during arrested landing[J]. Journal of Medical Biomechanics, 2018, 33(5): 383-389 (in Chinese).

[28] MERTZ H J, IRWIN A L, PRASAD P. Biomechanical and scaling bases for frontal and side impact injury assessment reference values[J]. Stapp Car Crash Journal, 2003, 47: 155-188.

[29] EPPINGER R, SUN E, BANDAK F, et al. Development of improved injury criteria for the assessment of advanced automotive restraint systems-II: 1999-6407-5[R]. Washington, D. C.: NHTSA, 1999.

[30] WHITE N A, MORENO D P, GAYZIK F S, et al. Cross-sectional neck response of a total human body Fe model during simulated frontal and side automobile impacts[J]. Computer Methods in Biomechanics and Biomedical Engineering, 2015, 18(3): 293-315.

[31] PARR J C, MILLER M E, PELLETTIERE J A, et al. Neck injury criteria formulation and injury risk curves for the ejection environment: A pilot study[J]. Aviation Space and Environmental Medicine, 2013, 84(12): 1240-1248.

Effect of helmet on neck injury of pilots in flight

WANG Lizhen[1,2,3], LIU Jinglong[1,2,3], ZHAO Yanpeng[4], BU Weiping[4], LIU Songyang[4], FAN Yubo[1,2,3,5,*]

(1. School of Biological Science and Medical Engineering, Beihang University, Beijing 100083, China;
2. Beijing Advanced Innovation Center for Biomedical Engineering, Beihang University, Beijing 100083, China;
3. Key Laboratory for Biomechanics and Mechanobiology of Ministry of Education, Beihang University, Beijing 100083, China;
4. Air Force Medical Center, Beijing 100142, China; 5. School of Engineering Medicine, Beihang University, Beijing 100083, China)

Abstract: To explore the effect of helmet on neck injury of pilots in-flight, numerical simulations were carried out on four different helmets during sharp turn and stable hover condition based on the established head-neck multi-body dynamic model. Combining the muscle force, intervertebral force and moment, and the neck injury index N_{ij}, the neck injury of pilots was evaluated and predicted. The results show that the dynamic responses of the model are in agreement with the experimental results, with fair bio-fidelity. In the same flight condition, the wearing of helmet B shifts the center of head forward and increases the muscle force of the trapezius accordingly. For the same cervical segment, the axial compression force for helmet B and the extension moment for helmet C is maximum respectively. The value of N_{ij} for all simulated conditions is less than 0.5, which conforms to the safety requirements in aviation. The mass of helmet B and helmet C are close, while the mass center of helmet B is forward compared to helmet C, which indicates the anteroposterior position of helmet has an important effect on neck injury. This study provides insights for the optimization of pilot's helmet and its effect on pilot's neck injury.

Keywords: pilots; helmet; head and neck; multi-body dynamic model; injury analysis

Received: 2022-07-12; **Accepted**: 2022-08-02; **Published online**: 2022-08-11 10:30

URL: kns.cnki.net/kcms/detail/11.2625.V.20220809.1527.001.html

Foundation items: National Natural Science Foundation of China (11822201); Air Force Medical University Effectiveness Improvement Plan for Flight Personnel (2019ZTA04)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0137

基于层次分析的信息物理网络多维攻击损毁评估

洪晟[1,*], 岳天羽[1], 李新建[2]

(1. 北京航空航天大学 网络空间安全学院, 北京 100083; 2. 湖北中烟工业有限责任公司, 武汉 430040)

摘 要: 由于信息物理网络交叉互联、多重连接的特点,易成为网络攻击的对象目标。通过对信息物理网络的网络体系架构和网络应用特点进行相关研究,提出了一种信息物理网络多维攻击损毁评估模型。将信息物理网络分为物理层、规则层、服务层3个维度,基于每个维度分别给出网络攻击损毁评估属性计算方法。通过对网络状态监测,可以得到拓扑稳定性、连通可用性、失效可控性,传输稳定性、丢包可用性、误码可控性,响应稳定性、业务可用性、风险可控性3个维度的攻击损毁评估属性。采用层次分析法得到信息物理网络3个不同维度及网络整体的攻击损毁评估结果,为信息物理网络的安全防护提供支撑与依据。模拟信息物理网络遭受分布式拒绝服务攻击,并采用所提方法进行攻击损毁评估,验证了所提模型的实用性和有效性。

关 键 词: 信息物理网络; 网络攻击; 多维攻击损毁评估; 攻击损毁评估属性; 层次分析法

中图分类号: TN915

文献标志码: A **文章编号**: 1001-5965(2022)09-1827-09

信息物理网络是一个由各种不同类型和规模、独立运行和管理的计算机网络连接组成的互联物理网络,其顺应了信息集成管理和控制一体化的发展趋势与需求,被广泛应用于能源、交通、市政、食品和航空航天等关系国计民生的重要行业和领域,密切关联着人们的生产生活,同时也关系国家发展的命脉。除此之外,信息物理网络依托各类应用程序携带了大量的数据、资源和服务,如互链接的超文本文档、电子邮件、微博与视频等应用均通过信息物理网络实现信息共享,随着全球范围网络攻击事件的不断增多,信息物理网络也成为国家级黑客攻击的重点目标,网络战已成为军事斗争的新形态,各国围绕网络战纷纷展开激烈较量和争夺。

由于信息物理网络表现出的多重互联、交叉连接及约束较少的特点,当网络中一个或少数几个实体节点或连接路径遭受攻击时,攻击会通过实体节点的连接关系向其他实体节点扩散传播,形成规模性的攻击失效连锁反应,如瞬时爆发及持续性的网络震荡、计算中断及服务拒止等现象,甚至产生灾难性的影响。

近年来,由于信息物理网络系统遭到病毒、攻击等导致网络失效连锁反应并造成巨大影响的案例数不胜数。2010 年“震网事件”爆发,黑客利用 Windows 系统中的多个漏洞来突破网络物理限制,释放“震网”病毒攻击伊朗铀浓缩基地信息控制网络,使至少五分之一的离心机被迫关闭[1]。2015 年 12 月,黑客利用“BlackEnergy”恶意软件

收稿日期: 2022-03-10; **录用日期**: 2022-03-25; **网络出版时间**: 2022-06-07 11:24
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220606.1706.003.html
基金项目: 国家重点研发计划(2019YFB1706001); 工信部高质量发展专项(TC210804R-2)
* **通信作者**. E-mail: shenghong@buaa.edu.cn

引用格式: 洪晟, 岳天羽, 李新建. 基于层次分析的信息物理网络多维攻击损毁评估[J]. 北京航空航天大学学报, 2022, 48(9): 1827-1835. HONG S, YUE T Y, LI X J. Multidimensional attack damage assessment of cyber-physical systems based on hierarchical analysis[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1827-1835(in Chinese).

在乌克兰电力信息网络系统中植入病毒,导致乌克兰 70 万户家庭断电,给人民的生产生活带来巨大影响[2]。2017 年 5 月,“永恒之蓝”勒索病毒爆发,“永恒之蓝”利用 Windows 系统的远程代码执行漏洞获取系统最高权限,乌克兰、俄罗斯、西班牙、法国与英国等多国的政府、银行、电力系统、通信系统、能源企业与机场等重要基础设施信息网络均遭遇到袭击[3]。2018 年 8 月,台积电公司生产园区控制信息网络遭大规模病毒攻击,使其大部分生产车间全线停工,且对全球的高性能计算芯片交付造成了很大影响,造成高达17.4 亿元的损失。2020 年 9 月,智利国家银行由于被恶意邮件在银行信息网络中安插后门,遭到勒索软件攻击,加密了银行内网中大部分服务工作站。

由此可见,大量黑客借助邮件、微博、视频、网页等网络应用攻击信息物理网络致使网络大范围连锁失效已经成为危害信息物理网络的主要问题,为了能更好地分析度量网络攻击对信息物理网络所造成的危害,判断信息物理网络遭受攻击后的网络损毁程度,需要对信息物理网络进行攻击损毁评估。

本文分析了信息物理网络攻击损毁评估的发展现状;从物理层、规则层、服务层 3 个不同层次维度构建了信息物理网络多维攻击损毁评估架构;将“稳定性”、“可用性”与“可控性”网络属性融入信息物理网络攻击损毁评估中;建立了信息物理网络多维攻击损毁评估模型并进行了案例实验分析。

1 信息物理网络攻击损毁评估发展

目前,对于信息物理网络损毁评估主要分为 2 方面:①对网络攻击损毁评估指标的研究;②对网络攻击损毁评估方法的研究。

在网络攻击损毁评估指标研究方面, Shi 等[4]分析了网络空间的影响,提出了基于模糊综合评价法的网络攻击损毁评价指标体系。Lin 等[5]通过提出了一个高级持续威胁(APT)攻击下的评价指标体系,通过该指标体系量化 APT 攻击损毁效果。鲜明等[6]提出了建立“目标-属性-指标”层次结构的信息网络损毁评估指标体系,通过分层分析、综合评估的方法,增强了信息网络损毁效果评估的系统性和实用性。汪生和孙乐昌[7]从攻击者的角度出发,对信息物理网络损毁评估指标体系进行了改进,采用状态迁移分析技术,提升了评估过程中的可控性和规范性。赵博夫和殷肖川[8]从破坏效果、信息获取及控制能力 3 个方面,建立了攻击损毁评估指标体系,利用层次分析法(AHP)灰色理论对网络攻击损毁进行评估。胡影等[9]提出了原子指标的概念,通过对攻击库的研究,根据原子功能提炼信息物理网络损毁评估指标,增强了评估方法的适用性。Niu 等[10]对无线信息网络攻击干扰损毁评价指标进行了研究,采用网络划分率作为评价指标。

在网络攻击损毁评估方法研究方面,刘进等[11]采用层次分析法,通过专家层层打分进行信息物理网络损毁评估计算,提高了评估的准确性。Maciel 等[12]通过层次建模方法对物联网系统在 DDoS 攻击下的影响进行了评估。李雄伟等[13]将模糊理论引入信息网络攻击效果评估领域,提出基于模糊综合评价法的评估模型,简化了计算的复杂度。Chai 等[14]采用间接变权法计算权值,采用模糊变权综合评价方法,进行网络攻击损毁评估。钟远和郝建国[15]提出基于网络熵的信息物理网络攻击效果评估方法,从一定程度上缓解了网络损毁评估的主观性,但存在误差较大的情况。戴方芳等[16]在网络熵的基础上提出基于二维熵分量的 K 均值信息网络攻击效果评估方法,可以更高效地处理评估数据,得出评估结果。曾伟渊[17]通过分析各种熵差的计算方法,提出了一种基于网络“熵差”的攻击损毁评估方法。魏兴[18]将模糊层次分析法与熵权法权值的接近度作为权重准确性依据,以模糊层次权重趋势为基准确定赋权方法,建立了基于接近度赋权的网络攻击损毁评估模型。Jajodia 等[19]将灰色理论用于评估计算信息网络损毁效果,增强了损毁评估各指标之间的关联性。Lim 等[20]提出了一种二阶段数学优化模型,通过二阶段随机整数规划建立网络损伤评估优化模型。郭琳茜[21]综合考虑攻击复杂性与攻击损毁之间的关联,将网络分析法与层次分析法相结合,提出了基于变权逼近理想解排序方法的网络攻击损毁评估模型。贾薇等[22]针对无线信息网络,提出了一种基于攻击树的网络攻击损毁评估方法。

上述对信息物理网络攻击损毁评估的方法均从信息物理网络整体网络安全的角度出发,即利用完整性、可靠性及可用性作为评估属性,没有充分考虑信息物理网络应用(如邮件、微博、视频、网页等)遭受攻击时的损毁特性,无法从信息物理网络本质架构特性上进行攻击损毁评估。

本文从信息物理网络架构特性出发,面对信息物理网络遭受攻击及攻击扩散传播过程,提出

对信息物理网络进行分层分析,从物理层、规则层、服务层 3 个不同层次维度进行网络攻击损毁评估,并将信息物理网络的“稳定性”、“可用性”与“可控性”等网络属性融入其 3 个不同层次维度的攻击损毁评估模型中,构建一种信息物理网络多维攻击损毁评估模型。

2　信息物理网络多维攻击损毁评估架构

鉴于信息物理网络邮件、微博、视频等应用的特点,并参考黄宁和伍志韬[23]对网络可靠性评估提出的 3 层评估模型,将信息物理网络分为物理层、规则层、服务层 3 个维度,从信息物理网络的 3 个维度进行信息物理网络的多维度攻击损毁评估,信息物理网络多维攻击损毁评估架构如图 1 所示。

这种维度划分能够很好地对邮件、微博、视频等信息物理网络应用的功能进行表述。物理层构建了信息物理网络的基础物理层面部分,包括了网络中节点硬件管理和网络拓扑规划等内容;规则层是链接物理层面和服务层面的关键,包括了网络传输设置、网络协议配置和设备使用配置等内容,决定了网络数据的传输、分发与处理;而服务层直接面向网络用户,包括了网络软件应用的运行和网络数据的计算转化等内容。

在进行了信息物理网络物理层、规则层、服务层 3 个维度的划分后,基于每一维度的攻击损毁评估指标,形成每一维度相应的攻击损毁评估的网络属性,通过层次分析法对各网络属性及相关加权融合,建立信息物理网络物理层、规则层、服务层 3 个维度的多维攻击损毁评估模型。

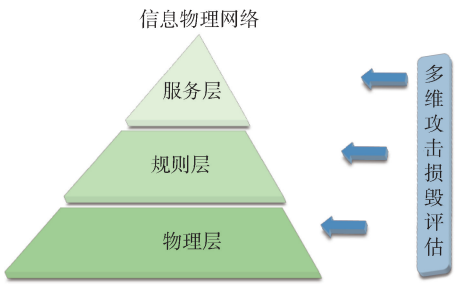


图 1　信息物理网络多维攻击损毁评估架构

Fig. 1　Framework for multidimensional attack damage assessment of cyber-physical systems

3　信息物理网络多维攻击损毁评估属性

3.1　物理层维度

对于信息物理网络物理层面,在遭受网络攻击后会出现网络中被攻击节点失效,并且失效将会向其他节点传播扩散,进而导致网络失效节点数量增加,网络节点间连接减少,网络整体连通性下降。因此,对信息物理网络物理层面的评估主要评估网络的拓扑稳定性、连通可用性及失效可控性属性。

3.1.1　拓扑稳定性

信息物理网络遭受攻击后会产生节点失效,而失效节点也会进一步扩散传播,进而破坏整个网络的拓扑结构,通过寻找网络中最大连通子图,计算信息物理网络在遭受攻击下的拓扑连通性,可表示为

$$g = \frac{N_G}{N} \tag{1}$$

式中:g 为网络拓扑连通性;N 为网络中总节点数;N_G 为网络中最大连通子图节点数。

随着攻击的扩散传播,网络的拓扑结构不断发生变化,最大连通子图也随之改变,拓扑连通性也不断变化。因此,本文基于网络的拓扑连通性,采用拓扑稳定性来表示攻击对网络的拓扑结构损害的程度,拓扑稳定性可表示为

$$S_g = 1 - \frac{1}{n\bar{g}} \sum_{i=1}^{n} |g - \bar{g}| \tag{2}$$

式中:S_g 为网络的拓扑稳定性;$\bar{g}$ 为网络中拓扑连通性的平均值;n 为对网络拓扑的观测记录次数。

3.1.2　连通可用性

在信息物理网络的物理层维度中,其可用性评估的重点在于网络中节点与网络连边可用与否。因此,运用连通可用性来评估网络攻击对信息物理网络物理层可用性的影响,连通可用性可表示为

$$A_c = P(g > g_0) \tag{3}$$

式中:A_c 为网络的连通可用性;g_0 为网络拓扑连通性阈值,认为只有当网络当前连通性大于 g_0 时,网络才处于可用状态;P 为概率。

连通可用性表示网络拓扑连通性大于阈值的可用状态的概率,但对单个时间的网络连通可用性计算存在较大的随机。因此,采用长时间 T 内的求和平均值来表示网络连通可用性:

$$A_c = \frac{1}{T} \sum_{T} P(g > g_0) \tag{4}$$

基于概率计算对式(4)进一步变形,得到

$$A_c = \frac{\mathrm{MTBF}}{\mathrm{MTBF} + \mathrm{MTTR}} = \frac{T_{g>g_0}}{T_{g>g_0} + T_{g\leqslant g_0}} \tag{5}$$

式中:MTBF(即 $T_{g>g_0}$)为网络连通可用的时间;MTTR(即 $T_{g\leqslant g_0}$)为网络连通不可用的时间。

3.1.3 失效可控性

信息物理网络在受到攻击后,相较于未受攻击时的网络失效率 λ 大大增加,运用网络熵的原理和计算方法,失效率 λ 增加,网络中熵值增加,网络的不可控性增大。通过熵计算,计算信息物理网络失效可控性来评估网络攻击损毁,其计算式为

$$H_\lambda = 1 + \log_2(1 - \lambda) \tag{6}$$

式中:H_λ 为网络失效可控性,当失效率 λ 越大时,网络失效可控性 H_λ 越小。

3.2 规则层维度

信息物理网络的规则层主要负责网络协议配置、网络设备配置和网络运行规则的制定。在网络受到攻击后,主要影响网络传输时延、网络节点转发速率、网络传输丢包、网络传输误码与网络传输流量等指标。因此,在对信息物理网络规则层的攻击损毁评估中主要考虑网络的传输稳定性、丢包可用性及误码可控性属性。

3.2.1 传输稳定性

在规则层中,网络遭受攻击会影响网络传输时延,传输时延指数据通过网络连接,从一个节点发送到另一个节点所需的时间,网络传输时延计算式为

$$D = \frac{s}{v_t} + \frac{\mathrm{DB}}{\mathrm{sp}_{\mathrm{sr}}} \tag{7}$$

式中:s 为传输距离;v_t 为网络传输速率;DB 为传输网络数据大小;$\mathrm{sp}_{\mathrm{sr}}$ 表示网络节点收发数据速率。

同物理层维度一样,本文基于网络传输时延,采用传输稳定性来表示攻击对网络规则层维度损毁程度,网络传输稳定性为

$$S_D = 1 - \frac{1}{n\overline{D}} \sum_{i=1}^{n} |D - \overline{D}| \tag{8}$$

式中:S_D 为网络传输稳定性;$\overline{D}$ 为网络传输时延的平均值。

3.2.2 丢包可用性

在规则层维度,攻击还会影响到信息物理网络的丢包率,因此,采用丢包可用性来评估网络攻击对信息物理网络规则层可用性的影响。设置丢包率阈值 L_0,当实时丢包率 L 大于或等于阈值时,认为网络损毁严重,无法正常使用,反之则认为网络可以正常使用运行。丢包可用性 A_{loss} 为

$$A_{\mathrm{loss}} = P(L < L_0) \tag{9}$$

同物理层维度一样,计算在长时间 T 中丢包可用性:

$$A_{\mathrm{loss}} = \frac{1}{T} \sum_{T} P(L < L_0) \tag{10}$$

基于概率计算对式(10)进一步变形,得到公式如下:

$$A_{\mathrm{loss}} = \frac{\mathrm{MUT}_{\mathrm{loss}}}{\mathrm{MUT}_{\mathrm{loss}} + \mathrm{MDT}_{\mathrm{loss}}} \tag{11}$$

式中:$\mathrm{MUT}_{\mathrm{loss}}$ 为网络丢包可用时间;$\mathrm{MDT}_{\mathrm{loss}}$ 为网络丢包不可用时间。

3.2.3 误码可控性

在受到攻击后,信息物理网络在传输数据时误码会大大增加,基于误码率来表示信息物理网络的可控性,误码率计算式为

$$\mathrm{SER} = \frac{B_e}{B} \times 100\% \tag{12}$$

式中:B 为单位时间传输的总数据量;B_e 为单位时间传输的误码数据量。

运用网络熵的原理和计算方法,随着网络中误码率 SER 的增加,网络熵值增加,从而网络的不可控性增大。通过熵计算,计算信息物理网络误码可控性来评估网络攻击损毁,其计算式为

$$H_{\mathrm{SER}} = 1 + \log_2(1 - \mathrm{SER}) \tag{13}$$

式中:H_{SER} 为网络误码可控性,当误码率 SER 越大时,网络误码可控性 H_{SER} 越低。

3.3 服务层维度

在信息物理网络的服务层面,主要功能是对外界用户操作请求给与响应,当信息物理网络遭受网络攻击后,网络对用户请求的响应将会受到影响,并且网络应用服务所面临的风险会大大增加,网络提供的应用业务也会受到影响,可用性降低甚至无法使用。因此,在信息物理网络服务层面的攻击损毁评估中主要分析网络的响应稳定性、业务可用性及风险可控性属性。

3.3.1 响应稳定性

在服务层中,攻击会影响网络对用户请求业务的响应,以响应时间为例,服务响应时间计算如下:

$$\mathrm{RES} = 2\left(\frac{s}{v_t} + \mathrm{delay}\right) \tag{14}$$

式中:delay 为服务处理时延。

同物理层、规则层维度一样,本文基于网络响应时间,采用响应稳定性来表示攻击对网络服务层维度损毁程度,网络响应稳定性 S_{RES} 为

$$S_{\mathrm{RES}} = 1 - \frac{1}{n \cdot \overline{\mathrm{RES}}} \sum_{i=1}^{n} |\mathrm{RES} - \overline{\mathrm{RES}}| \tag{15}$$

式中:$\overline{\mathrm{RES}}$ 为网络响应时间的平均值。

3.3.2 业务可用性

对于信息物理网络服务层维度,其可用性主

要在于网络对用户所请求业务响应的可用与否,因此,采用业务响应时间来计算业务可用性。网络业务可用性计算式为

$$A_{\mathrm{RES}} = P(\mathrm{RES} < (1+\rho)\cdot \mathrm{RES}_0) \tag{16}$$

式中:RES_0 为网络业务响应时间阈值;ρ 为响应抖动比例。

通常采用可用性原则取标准正态分布,认为当网络当前响应时间小于抖动的响应时间阈值 RES_0 时,网络处于可用状态。

同物理层、规则层维度一样,对单个时间的网络业务可用性计算存在较大的随机性,因此,采用长时间 T 内的求和平均值来表示网络的业务可用性:

$$A_{\mathrm{RES}} = \frac{1}{T}\sum_{T} P(\mathrm{RES} < (1+\rho)\cdot \mathrm{RES}_0) \tag{17}$$

基于概率计算对式(17)进一步变形,得到

$$A_{\mathrm{RES}} = \frac{\mathrm{MUT}_{\mathrm{RES}}}{\mathrm{MUT}_{\mathrm{RES}} + \mathrm{MDT}_{\mathrm{RES}}} \tag{18}$$

式中:$\mathrm{MUT}_{\mathrm{RES}}$ 为网络业务可用时间;$\mathrm{MDT}_{\mathrm{RES}}$ 为网络业务不可用时间。

3.3.3　风险可控性

在信息物理网络服务层维度,用户请求的网络业务服务本身就存在一定的风险值,而在遭受网络攻击后业务服务风险值会增大,本文采用风险值来计算信息物理网络服务层的可控性,风险值 R 计算式为

$$R = ph \tag{19}$$

式中:p 为风险频率;h 为风险损失。

运用网络熵的原理和计算方法,网络遭受攻击,网络风险值 R 增加,网络的熵值增加,从而网络不可控性增大。通过熵计算,计算信息物理网络风险可控性来评估网络攻击损毁,其表示为

$$H_R = 1 + \log_2(1-R) \tag{20}$$

式中:H_R 为网络风险可控性,当风险值 R 越大时,网络风险可控性 H_R 越小。

4　信息物理网络多维攻击损毁评估模型

信息物理网络多维攻击损毁评估模型如图 2 所示。

通过建立对邮件、微博、视频、网页等信息物理网络及应用网络攻击行为全面跟踪、记录、描述与分析的机制,分析受到攻击后信息物理网络连通子图、误码率、丢包率、时延、响应时间等网络性能指标的变化,根据第 3 节所提出的计算公式计算出信息物理网络物理层、规则层、服务层 3 个维度的网络稳定性、可用性、可控性属性。如图 3 所示,采用层次分析法按照“目标-维度-属性”的原则对攻击毁损评估问题逐层进行分析。目标层是信息物理网络攻击毁损评估结果;维度层是信息物理网络物理层、规则层、服务层 3 个不同评估维度;属性层为信息物理网络不同层次维度的攻击毁损评估属性,包括稳定性、可用性与可控性。所使用层次分析法中不出现指标层,网络相关性能指标依据第 3 节所提出的计算公式分别计算信息物理网络物理层、规则层、服务层 3 个维度的网络稳定性、可用性、可控性属性。

首先,通过层次分析法得到物理层、规则层、服务层各维度的相对权重。接着,通过权重系数的复合计算各个网络属性相对于最终攻击损毁评估的绝对权重。随后,采用加权融合计算信息物理网络物理层、规则层、服务层 3 个维度的攻击损毁评估结果及信息物理网络整体的攻击损毁评估结果。最后,依据信息物理网络在不同攻击场景下的毁损评估结果,将信息物理网络损毁情况划分为轻微损毁、轻度损毁、较大损毁、严重损毁 4 个级别。

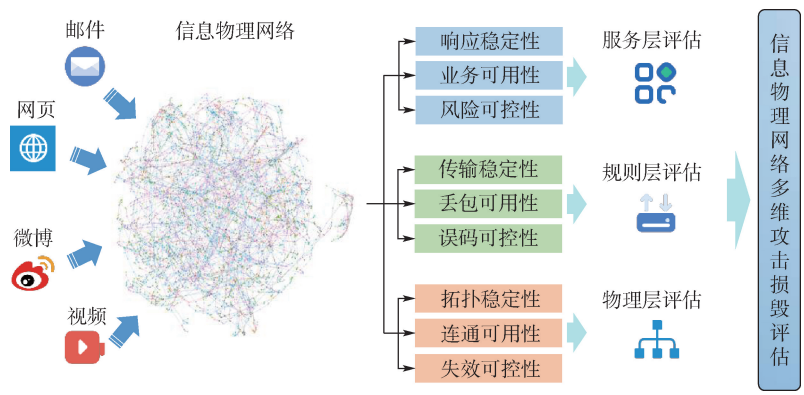


图 2　信息物理网络多维攻击损毁评估模型

Fig. 2　Model for multidimensional attack damage assessment of cyber-physical systems

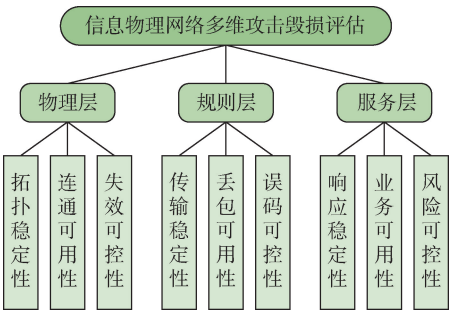


图 3 基于层次分析法的网络攻击损毁评估

Fig. 3 Network attack damage assessment based on analytic hierarchy process

5 案例分析

案例所研究的信息物理网络架构如图 4 所示。

通过模拟信息物理网络遭受分布式拒绝服务(DDoS)攻击后的网络状态数据变化,监测记录并计算出网络中网络拓扑、失效节点、失效率、丢包率、传输时延、误码率、风险值等指标值,随后依据第 3 节所给出的信息物理网络物理层、规则层、服务层 3 个维度的属性计算公式计算相应的网络攻击损毁评估属性,计算出的各属性如表 1 和图 5 所示。

从表 1 和图 5 中可以看出,DDoS 攻击在服务层中业务可用性的损毁程度最大,这是由于 DDoS 攻击不断进行服务访问,使得正常业务使用受到极大影响。而在规则层中,同样由于 DDoS 攻击导致的大量设备不断访问网络系统,使得网络中存在大量用户不断上下线并发送数据,数据互相干扰,从而在网络传输中形成大量乱码,使得误码可控性损毁程度加大。

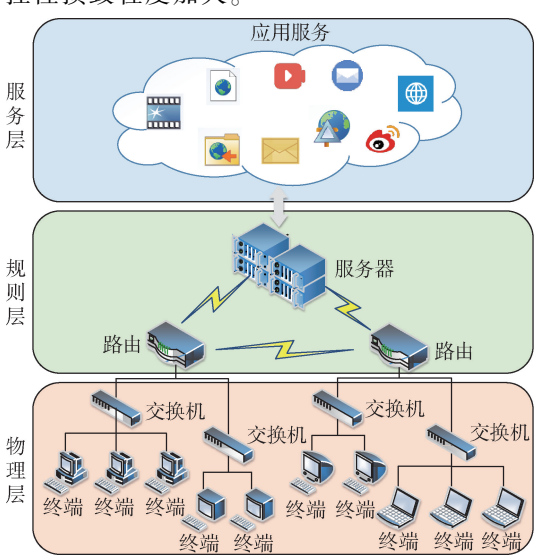


图 4 信息物理网络架构

Fig. 4 Architecture of cyber-physical systems

表 1 网络攻击损毁评估属性

Table 1 Attributes for network attack damage assessment %

维度	攻击损毁评估属性	数值
物理层	拓扑稳定性	74.35
	连通可用性	96.65
	失效可控性	92.60
规则层	传输稳定性	97.41
	丢包可用性	88.22
	误码可控性	67.81
服务层	响应稳定性	81.14
	业务可用性	72.22
	风险可控性	84.80

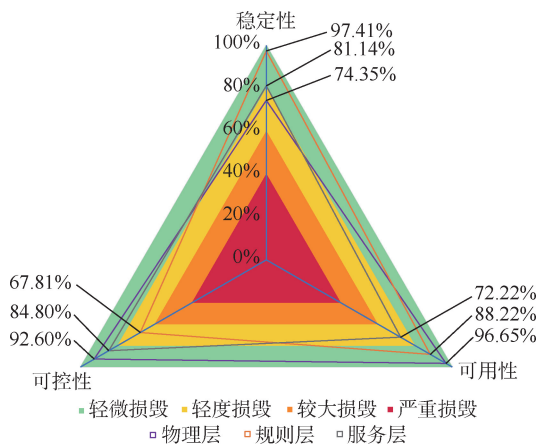


图 5 网络多维攻击损毁评估展现

Fig. 5 Display of network multidimensional attack damage assessment

采用层次分析法计算信息物理网络多维攻击损毁评估模型中属性与维度层的相对权值。利用加权融合得到信息物理网络物理层、规则层、服务层 3 个维度的攻击损毁评估结果如图 6 所示。

从图 6 中可以看出,信息物理网络遭受网络攻击后,物理层、规则层、服务层的损毁评估结果分别为 89.32%、87.67%、77.08%,该值表示网络完好程度。可以看出,服务层受攻击损毁最为

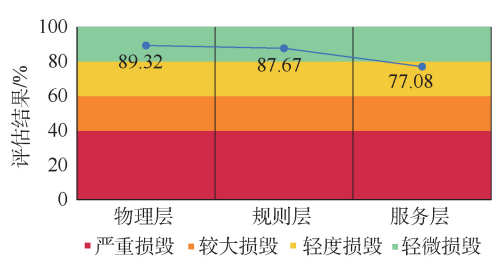


图 6 信息物理网络多维攻击损毁评估结果

Fig. 6 Results of multidimensional attack damage assessment of cyber-physical systems

严重,物理层损毁最轻,这是因为服务层需要基于物理层与规则层来为用户提供服务业务,因此,其损毁程度会受到下两层的影响,所以其损毁程度最高。此外 DDoS 攻击是通过大量设备对网络协同进行不断的服务访问,从而消耗网络服务资源,致使网络瘫痪,无法响应正常应用服务,由于 DDoS 攻击的主要是服务层面,评估结果中服务层的损毁程度最高,与攻击方式相符合。最后,可以计算得到信息物理网络整体的攻击损毁评估结果,如图 7 所示。

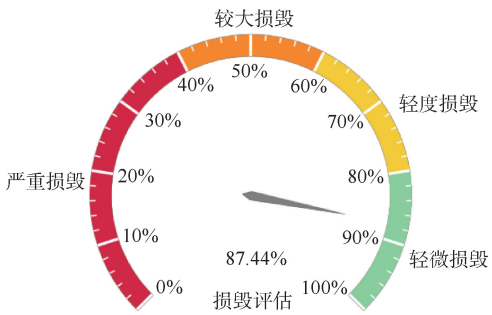


图 7　信息物理网络整体攻击损毁评估结果

Fig. 7　Results of overall attack damage assessment of cyber-physical systems

6　结　论

信息物理网络在国家能源、交通、通信、社交与航空航天等领域都有广泛应用,并且由于信息物理网络存在多重互联、交叉连接与约束较少的特点,已成为各国网络战攻击的重要目标,因此,对信息物理网络遭受攻击后的攻击损毁评估至关重要。本文基于现有的网络攻击评估研究,提出信息物理网络多维攻击损毁评估模型,具有如下优点:

1) 本文方法构建了信息物理网络多维攻击损毁评估架构,基于信息物理网络协议架构,从物理层、规则层、服务层 3 个维度进行网络攻击损毁评估分析。

2) 本文方法建立了信息物理网络多维攻击损毁评估模型,在物理层、规则层、服务层 3 个不同维度中分别采用稳定性、可用性、可控性属性进行评估,运用层次分析法得出 3 个维度及信息物理网络整体的攻击损毁评估结果。

3) 本文方法对提出的评估属性,拓扑稳定性、连通可用性、失效可控性、传输稳定性、丢包可用性、误码可控性、响应稳定性、业务可用性与风险可控性均给出了具体计算公式,能更好地对信息物理网络攻击损毁评估进行定量分析。

参考文献(References)

[1] ALLADI T, CHAMOLA V, ZEADALLY S. Industrial control systems: Cyberattack trends and counter-measures[J]. Computer Communications, 2020, 155: 1-8.

[2] 郭庆来,辛蜀骏,王剑辉,等. 由乌克兰停电事件看信息能源系统综合安全评估[J]. 电力系统自动化,2016,40(5):145-147.
GUO Q L, XIN S J, WANG J H, et al. Comprehensive security assessment for a cyber physical energy system: A lesson from Ukraine's blackout[J]. Automation of Electric Power Systems, 2016, 40(5): 145-147(in Chinese).

[3] 王冬. 解析勒索软件 Cerber 的攻击方法[J]. 计算机与网络,2016,15:52-53.
WANG D. Analyze the attack method of ransomware Cerber [J]. Computer and Network, 2016, 15: 52-53(in Chinese).

[4] SHI Z Y, ZHAO G L, LIU J H. The effect evaluation of the network attack based on the fuzzy comprehensive evaluation method[C]// 2016 3rd International Conference on Systems and Informatics (ICSAI). Piscataway: IEEE Press, 2016: 367-372.

[5] LIN J, JIANG R, QI Y L, et al. Construction of index system based on advanced persistent threat[C]// ITM Web of Conferences, 2017, 12: 03009.

[6] 鲜明,包卫东,王永杰. 网络攻击效果评估导论[M]. 长沙:国防科技大学出版社,2007.
XIAN M, BAO W D, WANG Y J. Introduction to effect evaluation of network attack[M]. Changsha: National University of Defense Technology Press, 2007(in Chinese).

[7] 汪生,孙乐昌. 网络攻击效果评估系统的研究与实现:基于指标体系[J]. 计算机工程与应用,2005,41(34):149-153.
WANG S, SUN L C. Research and implementation of evaluation system for network attack effect based on indicator system[J]. Computer Engineering and Applications, 2005, 41(34): 149-153(in Chinese).

[8] 赵博夫,殷肖川. 多维网络攻击效果评估方法研究[J]. 计算机工程与设计,2011,32(8):2596-2599.
ZHAO B F, YIN X C. Research on multidimensional attack effect evaluation method[J]. Computer Engineering and Design, 2011, 32(8): 2596-2599(in Chinese).

[9] 胡影,郑康锋,杨义先. 一种基于原子功能的网络攻击效果评估指标体系[J]. 计算机工程与科学,2008,30(10):1-4.
HU Y, ZHENG K F, YANG Y X. A metric system to evaluate network attack effects based on atomic functions[J]. Computer Engineering & Science, 2008, 30(10): 1-4(in Chinese).

[10] NIU Z, MA T, MA C L, et al. Analysis of the evaluation index for jamming partition battlefield wireless network[C]// 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference. Piscataway: IEEE Press, 2019: 2438-2442.

[11] 刘进,王永杰,张义荣,等. 层次分析法在网络攻击效果评估中的应用[J]. 计算机应用研究,2005,22(3):113-115.
LIU J, WANG Y J, ZHANG Y R, et al. Application of analytic hierarchy process to network attack effect evaluation[J]. Appli-

cation Research of Computers, 2005, 22(3): 113-115(in Chinese).

[12] MACIEL R, ARAUJO J, MELO C, et al. Impact evaluation of DDoS attacks using IoT devices[C]//2021 IEEE International Systems Conference (SysCon). Piscataway: IEEE Press, 2021: 1-8.

[13] 李雄伟,于明,杨义先,等. Fuzzy-AHP法在网络攻击效果评估中的应用[J]. 北京邮电大学学报,2006,29(1):124-127.
LI X W, YU M, YANG Y X, et al. Application of Fuzzy-AHP method in the effect evaluation of network attack[J]. Journal of Beijing University of Posts and Telecommunications, 2006, 29(1): 124-127(in Chinese).

[14] CHAI Z L. An indirect variable weights method to compute fuzzy comprehensive evaluation values[J]. Soft Computing, 2019, 23(23): 12511-12519.

[15] 钟远,郝建国. 基于系统熵的网络攻击信息支援效能评估方法[J]. 解放军理工大学学报(自然科学版),2014,15(2):127-132.
ZHONG Y, HAO J G. Estimation method for information support efficiency of network attacks based on system entropy[J]. Journal of PLA University of Science and Technology (Natural Science Edition), 2014, 15(2): 127-132(in Chinese).

[16] 戴方芳,郑康锋,胡影,等. 基于二维熵分量的K均值攻击效果评估[J]. 北京邮电大学学报,2014,37(1):71-75.
DAI F F, ZHENG K F, HU Y, et al. A K-means cluster evaluation of attack effect based on Bi-dimensional entropy components[J]. Journal of Beijing University of Posts and Telecommunications, 2014, 37(1): 71-75(in Chinese).

[17] 曾伟渊. 一种基于网络熵的计算机网络攻击效果定量评估方法[J]. 长春师范大学学报,2016,35(8):24-29.
ZENG W Y. A quantitative evaluation method of computer network attack effects based on network entropy[J]. Journal of Changchun Normal University, 2016, 35(8): 24-29(in Chinese).

[18] 魏兴. 面向目标对象的网络攻击效果评估技术研究与实现[D]. 北京:北京邮电大学,2018.
WEI X. Research and implementation of targeted object attack evaluation[D]. Beijing: Beijing University of Posts and Telecommunications, 2018(in Chinese).

[19] JAJODIA S, NOEL S, O'BERRY B. Topological analysis of network attack vulnerability[M]. Berlin: Springer, 2007: 247-266.

[20] LIM G J, KIM S, CHO J, et al. Multi-UAV pre-positioning and routing for power network damage assessment[J]. IEEE Transactions on Smart Grid, 2018, 9(4): 3643-3651.

[21] 郭琳茜. 面向Ad Hoc网络的攻击效果评估技术研究与实现[D]. 北京:北京邮电大学,2019.
GUO L X. Research and implementation of attack effect evaluation technology for Ad Hoc networks[D]. Beijing: Beijing University of Posts and Telecommunications, 2019(in Chinese).

[22] 贾薇,胡影,戴方芳. 基于攻击树的无线局域网攻击效果评估[C]//第十九届全国青年通信学术年会. 北京:国防工业出版社,2014:127-131.
JIA W, HU Y, DAI F F. WLAN attack effect assessment based on attack tree[C]//Proceedings of the 19th National Youth Communication Academic Conference. Beijing: National Defense Industry Press, 2014: 127-131(in Chinese).

[23] 黄宁,伍志韬. 网络可靠性评估模型与算法综述[J]. 系统工程与电子技术,2013,35(12):2651-2660.
HUANG N, WU Z T. Survey of network reliability evaluation models and algorithms[J]. Systems Engineering and Electronics, 2013, 35(12): 2651-2660(in Chinese).

Multidimensional attack damage assessment of cyber-physical systems based on hierarchical analysis

HONG Sheng[1,*], YUE Tianyu[1], LI Xinjian[2]

(1. School of Cyber Science and Technology, Beihang University, Beijing 100083, China;
2. China Tobacco Hubei Industrial LLC, Wuhan 430040, China)

Abstract: With characteristics, such as cross interconnection and multiple connections, cyber-physical systems is susceptible to network attack. Through relevant research on its network architecture and on the characteristics of its network application, a model for multidimensional attack damage assessment is established. Firstly, the cyber-physical systems was divided in three dimensions: the physical layer, the rule layer and the service layer. For each layer, an attribute-based method for assessing the network attack damage was given. Then, through monitoring the network status, the attributes for attack damage assessment were obtained, which include topological stability, connectivity availability, failure controllability, transmission stability, packet loss availability, bit error controllability, response stability, service availability and risk controllability. Finally, the analytic hierarchy process was used to gain the evaluation results of the attack damage in each of the three different layers and the whole network, thus providing support and basis for the security protection of the network. Simulation of the DDoS attack on the cyber-physical systems, along with the application of the proposed method to evaluate the attack damage, verifies the practicability and effectiveness of the model.

Keywords: cyber-physical systems; network attacks; multidimensional attack damage assessment; attack damage assessment attributes; analytic hierarchy process

Received: 2022-03-10; **Accepted**: 2022-03-25; **Published online**: 2022-06-07 11:24

URL: kns.cnki.net/kcms/detail/11.2625.V.20220606.1706.003.html

Foundation items: National Key R & D Program of China (2019YFB1706001); High-quality Development Project of Ministry of Industry and Information Technology (TC210804R-2)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0368

多飞行器攻击时间一致性协同制导进展综述与展望

董希旺[1,2], 于江龙[2,*], 化永朝[1], 李清东[2], 任章[1,2]

(1. 北京航空航天大学 人工智能研究院, 北京 100083; 2. 北京航空航天大学 自动化科学与电气工程学院, 北京 100083)

摘 要: 攻击时间一致性协同制导是协同制导的一个重要分支。综述了多飞行器攻击时间一致性协同制导的主要研究成果和方法分类。对攻击时间一致性协同制导的战术意义进行了分析;按照目标的运动形式,分为对固定目标的攻击实现一致性协同制导和对运动/机动目标的协同制导2类问题;归纳了国内外对这2类协同制导问题的主要研究成果,从研究方法层面综述各种协同制导方法的优缺点及其适应的问题场景;对攻击时间一致性协同制导方向的相关技术挑战进行了总结,并对未来发展方向进行展望。

关 键 词: 协同制导; 攻击时间一致性; 攻击时间; 一致性跟踪; 通信

中图分类号: V249.1; TJ765.3

文献标志码: A **文章编号**: 1001-5965(2022)09-1836-09

世界各军事强国逐渐发展并完善本国的防空反导系统,单枚精确制导类飞行器的打击/拦截作战效能急剧下降。然而,通过高速、机动、隐身、抗干扰等方面来提升单体飞行器作战效能,将会花费极大代价。多飞行器协同制导技术将会是提升作战效能的可行途径[1-2]。

协同制导指的是多枚相同种类(或不同种类)的飞行器,在一定的分布式协同策略基础上,通过个体之间的信息传输,实现制导系统状态变量的协调一致(如攻击角度、攻击时间等)。随着弹载计算机技术、弹载通信技术的发展,多弹协同制导具备初步的工程可行性。多飞行器协同制导可以极大提高低成本飞行器编队的对机动目标的拦截能力、对高价值固定/慢速移动目标的毁伤能力,受到了各军事强国的广泛关注。

针对协同制导问题,可以通过制导任务表现形式来进行划分,一般可以分为攻击时间一致性协同制导[3]、攻击角度一致性协同制导[4]、博弈对抗协同制导[5]、任务约束协同制导[6]等。上述不同协同制导方法均有不同优点,不同协同制导方法的使用应该结合具体场景选择,甚至可以组合[7]及分段[8]使用。

本文主要针对攻击时间一致性协同制导技术的发展进行综述。攻击时间一致性协同制导具有十分重要的战术任务意义,笔者认为主要有以下4方面原因:①在对高价值固定/慢速移动目标的协同毁伤场景下,攻击时间一致性能够实现“饱和攻击”。在此场景下,多低成本的自杀式固定翼无人机、巡航类飞行器通过高密度饱和打击能够极大提升低空突防概率和目标毁伤效能。②在对高速高机动无人机/导弹的协同拦截场景下,多拦截飞行器同时到达遭遇点实现协同杀伤,可以

收稿日期: 2022-05-17; **录用日期**: 2022-06-09; **网络出版时间**: 2022-07-05 13:46

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20220704.1510.001.html

基金项目: 科技创新2030-“新一代人工智能”重大项目(2020AAA0108200); 国家自然科学基金(62103016,61922008,61873011,61973013,61803014); 博士后创新人才支持计划(BX20200034); 中国博士后科学基金(2020M680297)

* **通信作者**. E-mail: sdjxyjl@buaa.edu.cn

引用格式: *董希旺, 于江龙, 化永朝, 等. 多飞行器攻击时间一致性协同制导进展综述与展望[J]. 北京航空航天大学学报, 2022, 48(9): 1836-1844. DONG X W, YU J L, HUA Y Z, et al. Review and prospect of cooperative guidance with attack time consensus for multiple aerial vehicles[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1836-1844(in Chinese).*

使得多个破片类战斗部协同工作,提升拦截效能。③针对有人战斗机的拦截问题,多地空/空空导弹同时实现对目标锁定与末制导,则飞行员的心理压力将会提升,导致其突防逃逸效能下降。④针对协同末制导阵位的形成,必须通过多约束协同中制导来实现,协同中制导的本质也是多飞行器能够同时到达预测交班点。综上,攻击时间一致性协同制导技术在大多数制导问题中有非常大的应用价值。

目前,已经有许多学者发表了多篇较为完善的协同制导技术的综述论文[9-12]。这些综述类论文主要站在整体的角度对协同制导架构[9]、协同制导模式[10]和协同制导方法[11]进行全面的综述,让读者能够对协同制导理论和方法有一个全局的认识和把握。而本文则是对协同制导问题的一个重要分支进行方法和技术层面的综述。攻击时间协同制导尽管是研究较早、成果较完善的一种理论方法(文献[12]较为完善地综述了单飞行器攻击时间约束下的制导方法),然而尚缺乏对其优势和研究途径的总结。本文用“打击目标”来刻画协同制导问题,用“攻击时间一致性”来刻画多飞行器同时命中目标的协同特性。首先,本文将协同制导问题分为协同打击固定目标和协同打击运动/机动目标的协同制导问题。从控制系统角度来看,这 2 类协同制导系统分别为自治系统和非自治系统,具有方法层面的差异。其次,通过归纳总结国内外对多飞行器攻击时间一致性的协同制导技术的主要研究成果,从研究方法层面来综述各种协同制导方法的优缺点及其适应的问题场景。最后,在现有的研究基础上,对本领域的相关技术挑战与难点进行了总结,并且尝试对未来发展方向进行展望。

1　打击固定目标的攻击时间一致性协同制导

针对高价值固定目标的时间一致性协同制导问题,由于其映射“饱和攻击”概念具有十分鲜明的工程意义,研究成果相对较多。从方法技术角度上,主要分为以下 5 类方法:基于误差反馈的时间一致性协同制导方法、基于优化技术的时间一致性协同制导方法、基于轨迹跟踪的时间一致性协同制导方法、基于推力控制的时间一致性协同制导方法和基于智能学习技术的时间一致性协同制导方法。下面对这 5 类方法进行详细介绍。

1.1　基于误差反馈的时间一致性协同制导方法

在基于误差反馈的时间一致性协同制导问题中,先建立剩余飞行时间动力学模型,然后以设定的期望剩余飞行时间或者通过局部信息交互得到的平均剩余飞行时间作为跟踪指令,构建剩余飞行时间跟踪误差动力学模型。文献[13]较早地提出了一种基于期望剩余飞行时间误差反馈的偏置比例导引律形式(称为攻击时间控制制导律,ITCG),实现了对静止目标带有期望攻击时间的打击,并且扩展成为经典的时间一致协同制导。基于上述架构,文献[14]较早地构建了双层协同制导架构,顶层为一致性协调变量层,通过一致性协调算法得到期望剩余飞行时间指令,底层则采用攻击时间控制制导律实现协同制导。上述方法中,攻击时间一致性误差通过与比例导引律进行“加性”负反馈,同样,如果进行“乘性”负反馈,则从直观上也能实现攻击时间一致性误差的收敛。文献[15-16]则应用此思想,将剩余飞行时间跟踪误差乘性作用于比例导引律的导航比中,也有效地实现了期望的攻击时间一致性。由于近年来一致性协同控制理论的发展与应用,多飞行器剩余飞行时间一致性收敛恰好可以作为一致性协同控制的典型应用,二者在原理和思想方面存在天然相同点。为此,文献[17-21]则构建了剩余飞行时间动力学模型,将其转化为低阶非线性模型,进而利用一致性控制[17-19]和一致性跟踪控制[20-21]研究了无领导者/有领导者模式下的攻击时间一致性协同制导方法。

双层架构的优点是直观、物理意义明显,在未来协同决策-协同制导一体化作战模式下,对制导时间要求严格的场景中将会有较大应用价值,然而现有的方法对于期望攻击飞行时间指令大多为固定值,选择区间相对有限。“乘性”负反馈类协同制导方法,可能会由于初始误差加大导致初始过载较大。基于一致性算法的协同制导方法尽管期望攻击时间依赖于初始条件或领弹,但是具备一定的自组织性,实现起来较为方便,在近距高动态格斗、期望攻击时间要求相对宽松的场景下具有应用价值。

1.2　基于优化技术的时间一致性协同制导方法

由于针对固定目标的制导系统本质是自治系统,因此可利用最优控制/优化方法,将攻击时间作为一个约束,进行优化设计。文献[22]将带有攻击时间约束的制导问题转化为带有终端约束的连续系统最优控制问题,并通过数值方法得到攻击时间控制制导律,并可以直接扩展为攻击时间

一致性协同制导律。文献[23]则是将协同制导系统进行离散化,构建多约束离散的最优控制问题,通过模型预测控制方法获得攻击时间一致性协同制导律指令。文献[24]则构建了圆弧式协同制导律,构建了圆弧制导律参数与剩余飞行时间的参数映射关系,通过优化制导参数可以实现多飞行器对目标的同时打击。在弹目速度较大时,为了提高上述优化问题求解速度,文献[25-26]则将带有终端攻击时间约束的最优协同制导问题进一步线性化与离散化,引入序列凸优化技术实现了协同制导律的快速求解。

基于优化技术的时间一致性协同制导方法的优点较多,且能够同时处理复杂多约束条件。当飞行器规模增大的时候,基于优化技术的协同制导算法实时性、制导精度、稳定性将会受到一定局限,但是随着弹载计算机技术的发展,这一类方法在未来还是具有较大的应用前景。

1.3 基于轨迹跟踪的时间一致性协同制导方法

基于轨迹跟踪技术的时间一致性协同制导方法,一般根据1.1节介绍的双层协同制导架构进行设计,在设定期望剩余飞行时间或通过局部信息交互得到平均剩余飞行时间作为指令以后,构建带有期望飞行时间指令构型的参数化轨迹模型,通过优化轨迹参数可以获得带有期望攻击时间指令的轨迹解析表达式,在制导层面可以设计轨迹跟踪协同制导律实现多飞行器同时攻击。由于参数化轨迹模型的自变量不同,协同制导模型也会不同。目前较多的学者采用前置角作为自变量,文献[27-30]构建了前置角参数化轨迹模型,同时引入攻击时间指令作为优化代价函数或约束条件,通过参数优化方法可以得到带有期望攻击时间指令的轨迹解析表达式。文献[31]构建了视线角作为自变量的协同制导模型,并且得到带有攻击时间约束的视线角轨迹曲线。文献[32]则是直接将过载曲线进行参数化设计,通过闭环系统动力学分析求解剩余飞行时间,进而优化过载参数模型。

上述策略中,前置角参数化模型能够处理视场约束,这是因为前置角在一定条件下可以和视场角关联,通过对前置角轨迹的规划,能够同时保障目标在视场范围之内。视线角、过载参数化模型的优点在于参考轨迹的跟踪控制更加方便。

1.4 基于推力控制的时间一致性协同制导方法

当飞行器具备速度调节能力(或推力控制能力)时,则可以通过轴向过载和法向过载同时调节剩余飞行时间,实现攻击时间一致性协同制导。文献[33-34]将攻击时间一致性协同制导转化为分布式一致性控制问题,并且基于一致性理论设计了攻击时间一致性协同制导律。文献[35]基于领导者-跟随者架构,将攻击时间一致性协同制导转化为分布式一致性跟踪控制问题,构建了攻击时间可调的协同制导律。在此基础上,有学者研究了复杂多约束条件下的协同制导问题。文献[36]研究了飞行速度带有上下界约束的攻击时间一致性协同制导问题。文献[37]研究了无视线角速率测量情况下的攻击时间一致性协同制导问题。文献[38]则研究了无弹目相对速度测量情况下的攻击时间一致性协同制导问题。

上述策略引入了速度调节能力,这在现实中也是可以获得解决方案的。例如,在固定翼/旋翼无人机的协同制导场景、巡飞弹/巡航弹的协同制导场景下均是可以应用的。

1.5 基于智能学习技术的时间一致性协同制导方法

在上述4类方法中,为了实现攻击时间的协调一致,不可避免地需要使用剩余飞行时间这一状态变量。然而,剩余飞行时间与轨迹(或过载)是耦合的,这也就使得剩余飞行时间只能通过近似估计得到。剩余飞行时间在轨迹相对平直的情况下近似误差较小,然而在轨迹较为弯曲或目标存在运动/机动的情况下,其近似误差则可能会变大。对于这种具有非随机不确定性因素的协同制导模型,人工智能技术有一定的解决可行性。文献[39]指出可以采用离线学习和动态调整相结合的方法,实现智能自主制导。目前也有学者开展相关的研究。文献[40]采用神经网络来估计剩余飞行时间,并且构建了相应的攻击时间一致的智能协同制导律。文献[41-42]则利用强化学习方法,将攻击时间协同制导问题用数值优化的方式进行智能求解,通过离线大规模训练及在线使用的方式,实现了多飞行器对目标的同时打击。

随着人工智能技术的发展,智能导弹是未来导弹的发展方向,智能协同制导也将会成为未来智能导弹的重要技术基础。

2 打击运动/机动目标的攻击时间一致性协同制导

针对运动/机动目标的时间一致性协同制导问题,由于目标的未知运动/机动信息给协同制导系统带来不确定性,而且协同制导可行域将会进一步被压缩,存在一定的研究难度和技术挑战。然而,在工程应用中,由于这种方法能够实现对空

中运动/机动目标的有效打击,十分有现实意义。从方法技术角度上,本节的问题解决方法与固定目标协同制导问题解决方法较为类似,主要分为以下 5 类方法:基于剩余飞行时间误差反馈的时间一致性协同制导方法、基于预测命中点的时间一致性协同制导方法、基于期望轨迹跟踪的时间一致性协同制导方法、基于速度调节的时间一致性协同制导方法和基于动态规划的时间一致性协同制导方法。下面对这 5 类方法进行详细介绍。

2.1　基于剩余飞行时间误差反馈的时间一致性协同制导方法

针对运动/机动目标,在基于误差反馈的时间一致性协同制导问题中,剩余飞行时间跟踪误差动力学模型将会受到目标的未知运动/机动信息影响,这一部分将会作为未知外部干扰而影响系统的稳定性与制导精度。文献[43]在静止目标时间一致性协同制导律基础上,将目标机动进行补偿。文献[44-46]针对带有目标未知运动/机动信息的剩余飞行时间跟踪误差动力学模型,设计期望攻击时间指令,并且利用鲁棒控制或变结构控制方法,实现了对未知扰动的补偿,保障攻击时间一致性误差或期望剩余飞行时间跟踪误差能够有限时间收敛。此外,文献[47]利用反馈线性化方法,建立了期望相对距离跟踪误差系统动力学模型,并且利用变结构控制与干扰估计方法实现了多飞行器弹目相对距离的一致,进而实现了攻击时间的一致性。文献[48]构建了目标未知运动/机动信息滤波器,并且将协同制导系统线性化,引入期望攻击时间作为终端约束,利用最优控制方法得到了攻击时间协同制导律。

针对目标的运动/机动给协同制导系统带来的未知干扰信息,上述方法具有一定的鲁棒性。然而,期望攻击时间指令的选择将会影响协同制导系统的收敛性。

2.2　基于预测命中点的时间一致性协同制导方法

针对单飞行器对运动/机动目标的打击/拦截问题,将其转化为对预测命中点的打击/拦截是一个较为有效的技术。在协同制导问题中,也有一些学者将此技术进行应用。文献[49-51]利用预测命中点技术,将对机动目标的协同制导问题转化为对固定目标的协同制导问题。文献[52]则在预测命中点思想上,将目标的机动信息也引入到目标轨迹预测,修正了仅依赖目标当前位置和速度外推的预测命中点计算方法。文献[53]则是构建了一个虚拟目标点,并且结合分段制导思想,构建了带有攻击时间和角度约束的制导律,进而在指定期望攻击实现后,可以通过双层协同制导架构实现攻击时间一致性协同制导。

2.3　基于期望轨迹跟踪的时间一致性协同制导方法

由于存在目标运动/机动,期望协同制导轨迹规划会受到不确定因素的影响,因此,协同制导轨迹的形式及参数选择将会影响协同制导效果。文献[54]构建了前置角参数化轨迹模型,同时引入攻击时间指令作为期望终端条件,通过参数优化方法可以得到带有期望攻击时间约束的轨迹解析表达式,以匀速直线运动目标的协同制导场景进行了仿真验证。文献[55]则建立了带有期望攻击时间误差收敛特性的动力学模型,并且参数优化得到了带有期望收敛性能的协同制导律,以匀速直线运动目标的协同制导场景进行了仿真验证。文献[56]提出了一种新的虚拟目标轨迹跟踪控制方法,能够实现带有攻击角度和时间约束的协同制导。文献[57]设计了带有攻击时间约束的视线角/角速率曲线,并且给出了视线角跟踪制导律,实现了带有期望攻击时间控制的制导,同样,通过双层协同制导架构,可以实现多飞行器对机动目标的攻击时间一致性协同制导。

可以看出,由于目标未知运动/机动的影响,这一类协同制导方法对于期望轨迹的设计要求较高,相关参数也有一定的敏感性。

2.4　基于速度调节的时间一致性协同制导方法

与 1.4 节类似,当飞行器具备速度调节能力时,则可通过轴向过载实现攻击时间一致性协同制导。文献[58-59]针对无领导者架构,将攻击时间一致性协同制导转化为分布式一致性控制问题,并且设计了攻击时间一致性协同制导律。同样,文献[60-62]基于领导者-跟随者架构,将攻击时间一致性协同制导问题转化为分布式一致性跟踪控制问题,利用有限时间收敛方法,构建了攻击时间误差有限时间收敛的协同制导律。在此基础上,有学者研究了复杂多约束条件下的协同制导问题,如文献[63]则研究了无相对速度测量情况下的攻击时间一致性协同制导问题。

同样,上述策略引入了速度调节能力,在实际固定翼/旋翼无人机的协同制导打击地面车辆场景、巡飞弹/巡航弹的协同制打击水面/地面移动目标场景中均有应用潜力。

2.5　基于动态规划的时间一致性协同制导方法

由于目标存在机动能力时,从控制论的角度,制导问题也可以看作一个微分对策问题。因此,多飞行器协同制导就成为了一个协同微分对策问

题[64]。协同微分对策刻画了目标逃逸和协同打击这2个方面的博弈,体现了二者之间的智能特性。对于带有攻击时间一致性的协同博弈制导问题,在运动学回路,可以通过求解协同微分对策问题得到解析解或者数值解。然而,当考虑复杂非线性运动特性(或制导系统动力学特性)时,往往得不到解析解。自适应动态规划方法则是相对有效的解决方案。文献[65-67]将攻击时间一致性约束下的机动目标协同制导问题转化为非线性最优控制问题,利用自适应动态规划方法得到了协同制导律的近似解,实现了攻击时间一致性。

针对运动/机动目标的协同打击问题,动态规划制导方法或微分对策制导方法本质也是最优制导方法的推广,并且能够发展成为带有智能特性的协同博弈制导方法,可以有效处理攻击时间一致性等多约束条件下的协同制导问题,并且具备最优特性,同时此类方法需要的信息量相对较多,问题求解时间将会随着问题规模增大而增大。随着人工智能技术的发展,这一类方法在智能导弹中也会有一定的应用前景。

2.6 关于目标运动/机动特性分析

2.1节~2.5节综述了目标存在运动/机动特性下的攻击时间一致性协同制导问题。然而,针对合作型目标(运动/机动特性已知)和非合作型目标(运动/机动特性未知),上述方法的适用性仍需要进一步分析。

针对合作型目标,可以通过轨迹预测、卫星/载机信息探测等手段获取其运动/机动信息,因此,目标的运动/机动影响可以在协同制导律中进行补偿。针对基于预测命中点的时间一致性协同制导方法[49-53],大多将目标的未知动态进行估计与预测,转化为对合作类目标的协同制导问题。针对基于轨迹跟踪的时间一致性协同制导方法[54-57]、基于微分对策的协同制导方法[64],其协同制导律中也大多用到了目标的运动信息。

针对非合作目标,带有攻击时间一致性的协同制导问题本身就是一个比较有挑战性的问题,而且如果不依赖于目标的运动/机动信息,处理起来会更加困难。因此,仅有部分基于误差反馈的时间一致性协同制导方法[43-47]、基于自适应动态规划的时间一致性协同制导方法[65-67]能够实现期望的同时攻击,然而攻击时间一致性误差也有一定的残差、协同制导可行域也相对较窄。另外,基于推力控制的时间一致性协同制导方法则能够同时针对2类目标实现攻击时间一致性。因此,针对非合作目标带有攻击时间一致性的协同制导问题,仍是本领域的研究热点和难点。

2.7 关于制导信息获取途径分析

从制导信息获取途径的角度,协同制导方法可以分为分布式协同制导和集中式协同制导2类。可以看出,基于误差反馈的时间一致性协同制导方法、基于推力控制的时间一致性协同制导方法,以及部分基于动态规划的时间一致性协同制导方法可以实现分布式信息获取;而基于预测命中点的时间一致性协同制导方法、基于轨迹跟踪的时间一致性协同制导方法,以及第1.2节中的基于优化技术/智能学习的时间一致性协同制导方法,大多数是集中式的协同制导架构。由于分布式协同制导架构自组织性、自适应性更好,因此这也是时间一致性协同制导领域的重点问题。

3 攻击时间一致性协同制导律研究方向和展望

随着飞行器技术的发展,集群作战将会成为未来重要的作战方式。从世界各国的集群武器发展现状来看,飞行器集群武器已初现端倪,美国“小精灵”、“郊狼”等飞行器集群已经逐渐形成作战能力,初步具备对高价值目标的协同毁伤能力。针对防御方,多空空/地空导弹拦截机动目标的作战样式也在逐渐完善。攻击时间一致性协同制导技术是上述协同作战任务执行层面的关键环节之一。随着任务场景的复杂化、目标隐身/突防/机动能力的提升,攻击时间一致性协同制导应该向着智能性、实时性、鲁棒性、宽域性和多约束性等方面发展。主要方向介绍如下。

3.1 智能性

在未来作战场景中,目标运动/机动的能力将会提升,攻击时间一致性协同制导问题必然将会遇到目标不确定机动干扰带来的影响,形成协同博弈对抗态势。随着弹载计算机技术的发展,将基于强化学习、深度学习的智能控制、智能估计等方法引入到协同制导问题中,实现对机动目标作用下剩余时间的精准估计及攻击时间一致性协同制导律的设计,将会非常有意义。

3.2 实时性

基于1.2节和2.5节的分析,基于优化技术的攻击时间一致性协同制导方法能够处理更多约束、更复杂的模型。然而随着飞行器规模的增大,无论是集中式还是分布式的算法,实时性将会受到一定的挑战。因此,在大规模飞行器协同作战的场景下,需要建立更加精准的协同制导模型,结

合一些智能算法,简化方法,来提升最优协同制导算法的实时性。

3.3 鲁棒性

在实际应用中,鲁棒性体现在协同制导系统对不确定噪声、不确定环节的不敏感性。在现有的研究中,有部分结果并未考虑实际导引信息获取途径是否可行,如红外导引头较难获取相对速度,因此剩余飞行时间较难直接测量。导引信息中的噪声也会对协同制导效果产生较大影响。此外,目标的未知运动/机动信息也会对协同制导精度产生较大影响。因此在实际应用中,需要进一步探讨攻击时间一致性鲁棒协同制导问题。

3.4 宽域性

在现有研究中,部分攻击时间一致性协同制导算法收敛性对初始条件非常敏感,同时期望的攻击时间可行域相对狭窄。在高动态环境下,为了形成可行的攻击时间协同制导,多飞行器初始阵位必须处于可行域,这就要求协同制导可行域具备宽域性,即对初始条件、攻击时间指令具有大范围适应能力。

3.5 多约束性

攻击时间协同制导律是协同制导的一种典型制导律。正如引言中提到的,未来协同制导问题将会收到多约束条件限制,如攻击时间/攻击角度/视场角等多约束组合。尽管现在已经有较多成果,然而大多数都是基于有推力的情况,对于无速度调节能力的飞行器,能够实现多约束条件下的协同制导也是一个重要的发展方向。

4 结　论

攻击时间一致性协同制导是协同制导的基本问题,具有很重要的战术意义。尽管现有成果较为丰富,通过对现有方法的总结来看,仍存在一些局限性。从解析解角度来看,需要进一步探索攻击时间一致性协同制导律的新形式、新设计方法,来提升对初始条件和目标机动的适应性。从数值解角度来看,随着人工智能技术和弹载计算机技术的发展,一些先进的、智能的攻击时间一致性协同制导方法将会在未来得到应用。此外,作为基本的协同制导模式,尽管攻击时间、攻击角度、视场角 3 类协同制导约束均有分离或组合模式下的研究,然而相关方法适应性和智能性仍需进一步提高,未来一方面需要继续探索多种协同条件约束下的协同制导律设计,另一方面需要进一步面向实战场景,探索作战任务牵引的新型协同制导模式及其协同制导律设计方法。

参考文献(References)

[1] 肖增博,雷虎民,滕江川,等. 多导弹协同制导规律研究现状及展望[J]. 航空兵器,2011(6):18-22.
XIAO Z B,LEI H M,TENG J C,et al. The present status and cooperative prospect of multi-missile guidance law[J]. Aero Weaponry,2011(6):18-22(in Chinese).

[2] 任章,于江龙. 多临近空间拦截器编队拦截自主协同制导控制技术研究[J]. 导航定位与授时,2018,5(2):1-6.
REN Z,YU J L. Research on the autonomous cooperative guidance control for the formation interception of multiple near space interceptors[J]. Navigation Positioning and Timing,2018,5(2):1-6(in Chinese).

[3] ZHAO S Y,ZHOU R. Cooperative guidance for multimissile salvo attack[J]. Chinese Journal of Aeronautics,2008,21(6):533-539.

[4] 于江龙,董希旺,李清东,等. 拦截机动目标的分布式协同围捕制导方法[J/OL]. 航空学报,2021(2021-08-17)[2022-05-10]. http://kns.cnki.net/kcms/detail/11.1929.V.20210817.1013.010.html.
YU J L,DONG X W,LI Q D,et al. Distributed cooperative encirclement hunting guidance method for intercepting the maneuvering target[J/OL]. Acta Aeronautica et Astronautica Sinica,2021(2021-08-17)[2022-05-10]. http://kns.cnki.net/kcms/detail/11.1929.V.20210817.1013.010.html(in Chinese).

[5] KUMAR S R,SHIMA T. Cooperative nonlinear guidance strategies for aircraft defense[J]. Journal of Guidance,Control,and Dynamics,2017,40(1):124-138.

[6] 张帅,郭杨,王仕成. 带有引诱角色的有限时间协同制导方法[J]. 宇航学报,2018,39(3):308-317.
ZHANG S,GUO Y,WANG S C. Finite time cooperative guidance method with a lure role[J]. Journal of Astronautics,2018,39(3):308-317(in Chinese).

[7] 王晓芳,王紫扬,林海. 一种同时具有攻击时间和攻击角度约束的协同制导律[J]. 弹道学报,2017,29(4):1-8.
WANG X F,WANG Z Y,LIN H. A cooperative guidance law with constraints of impact time and impact angle[J]. Journal of Ballistics,2017,29(4):1-8(in Chinese).

[8] 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.

[9] 赵建博,杨树兴. 多导弹协同制导研究综述[J]. 航空学报,2017,38(1):22-34.
YANG J B,YANG S X. Review of multi-missile cooperative guidance[J]. Acta Aeronautica et Astronautica Sinica,2017,38(1):22-34(in Chinese).

[10] 王建青,李帆,赵建辉,等. 多导弹协同制导律综述[J]. 飞行力学,2011,29(4):6-10.
WANG J Q,LI F,ZHAO J H,et al. Summary of guidance law based on cooperative attack of multi-missile method[J]. Flight Dynamics,2011,29(4):6-10(in Chinese).

[11] 姚禹正,余文斌,杨立军,等. 多导弹协同制导技术综述[J]. 飞航导弹,2021(6):112-121.

YAO Y Z, YU W B, YANG L J, et al. Overview of multi missile cooperative guidance technology[J]. Aerodynamic Missile Journal, 2021(6): 112-121(in Chinese).

[12] 张友安,王星亮,吴华丽,等.带攻击时间约束的导引律综述[J].海军航空工程学院学报,2015,30(4):301-309.
ZHANG Y A, WANG X L, WU H L, et al. Survey of guidance law with attack time constraint[J]. Journal of Naval Aeronautical and Astronautical University, 2015, 30(4): 301-309(in Chinese).

[13] JEON I S, LEE J I, TAHK M J. Impact-time-control guidance law for anti-ship missiles[J]. IEEE Transactions on Control Systems Technology, 2006, 14(2): 260-266.

[14] 赵世钰,周锐.基于协调变量的多导弹协同制导[J].航空学报,2008,29(6):1605-1611.
ZHAO S Y, ZHOU R. Multi-missile cooperative guidance using coordination variables[J]. Acta Aeronautica et Astronautica Sinica, 2008, 29(6): 1605-1611(in Chinese).

[15] 张春妍,宋建梅,侯博,等.带落角和时间约束的网络化导弹协同制导律[J].兵工学报,2016,37(3):431-438.
ZHANG C Y, SONG J M, HOU B, et al. Cooperative guidance law with impact angle and impact time constraints for networked missiles[J]. Acta Armamentarii, 2016, 37(3): 431-438(in Chinese).

[16] JEON I S, LEE J I, TAHK M J. Homing guidance law for cooperative attack of multiple missiles[J]. Journal of Guidance, Control, and Dynamics, 2010, 33(1): 275-280.

[17] SINHA A, KUMAR S R, MUKHERJEE D. Cooperative integrated guidance and control design for simultaneous interception[J]. Aerospace Science and Technology, 2022, 120: 107262.

[18] LI Z H, DING Z T. Robust cooperative guidance law for simultaneous arrival[J]. IEEE Transactions on Control Systems Technology, 2019, 27(3): 1360-1367.

[19] LI K, WANG J N, LEE C H, et al. Distributed cooperative guidance for multivehicle simultaneous arrival without numerical singularities[J]. Journal of Guidance, Control, and Dynamics, 2020, 43(7): 1365-1373.

[20] 张友安,马国欣,王兴平.多导弹时间协同制导:一种领弹-被领弹策略[J].航空学报,2009,30(6):1109-1118.
ZHANG Y A, MA G X, WANG X P. Time-cooperative guidance for multi-missiles: A leader-follower strategy[J]. Acta Aeronautica et Astronautica Sinica, 2009, 30(6): 1109-1118(in Chinese).

[21] 林德福,何绍溟,王江,等.基于虚拟领弹-从弹的集群分布式协同制导技术研究[J].中国科学:技术科学,2020,50(5):506-515.
LIN D F, HE S M, WANG J, et al. Research on cluster distributed collaborative guidance technology based on virtual leader-follower[J]. Scientia Sinica: Technologica, 2020, 50(5): 506-515(in Chinese).

[22] KORAY S E, RAZIYE T. Impact time and angle control based on constrained optimal solutions[J]. Journal of Guidance, Control, and Dynamics, 2016, 39(10): 2445-2451.

[23] KANG S, WANG J N, LI G, et al. Optimal cooperative guidance law for salvo attack: An MPC-based consensus perspective[J]. IEEE Transactions on Aerospace and Electronic Systems, 2018, 54(5): 2397-2410.

[24] 张友根,张友安,施建洪,等.基于双圆弧原理的协同制导律研究[J].海军航空工程学院学报,2009,24(5):537-542.
ZHANG Y G, ZHANG Y A, SHI J H, et al. Research on cooperative guidance law based on double arc principle[J]. Journal of Naval Aeronautical and Astronautical University, 2009, 24(5): 537-542(in Chinese).

[25] 姜欢,赵建博,熊芬芬,等.多约束下基于凸优化的协同制导研究[C]//第36届中国控制会议,2017:1302-1307.
JIANG H, ZHAO J B, XIONG F F, et al. Cooperative guidance with constrained impact using convex optimization[C]// Proceedings of the 36th Chinese Control Conference, 2017: 1302-1307(in Chinese).

[26] JIANG H, AN Z, YU Y N, et al. Cooperative guidance with multiple constraints using convex optimization[J]. Aerospace Science and Technology, 2018, 79: 426-440.

[27] CHEN X T, WANG J Z. Nonsingular sliding-mode control for field-of-view constrained impact time guidance[J]. Journal of Guidance, Control, and Dynamics, 2018, 41(5): 1214-1222.

[28] TEKIN R, ERER K S, HOLZAPFEL F. Adaptive impact time control via look-angle shaping under varying velocity[J]. Journal of Guidance, Control, and Dynamics, 2017, 40(12): 3247-3255.

[29] HOU Z W, YANG Y, LIU L, et al. Terminal sliding mode control based impact time and angle constrained guidance[J]. Aerospace Science and Technology, 2019, 93: 105142.

[30] WANG P Y, GUO Y N, MA G F, et al. New look-angle tracking guidance strategy for impact time and angle control[J]. Journal of Guidance, Control, and Dynamics, 2022, 45(3): 545-557.

[31] 李浩,毛昱天,吴丹,等.基于理想视线制导的多导弹协同攻击策略[J].导航定位与授时,2015,2(4):1-6.
LI H, MAO Y T, WU D, et al. Cooperative attack of multiple missiles with ideal-line-of-sight-guidance[J]. Navigation Positioning and Timing, 2015, 2(4): 1-6(in Chinese).

[32] 李斌,林德福,何绍溟,等.基于最优误差动力学的时间角度控制制导律[J].航空学报,2018,39(11):157-167.
LI B, LIN D F, HE S M, et al. Time and angle control guidance law based on optimal error dynamics[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(11): 157-167(in Chinese).

[33] WANG C, DONG W, WANG J, et al. Guidance law design with fixed-time convergent error dynamics[J]. Journal of Guidance, Control, and Dynamics, 2021, 44(7): 1389-1398.

[34] ZHOU J L, YANG J Y. Distributed guidance law design for cooperative simultaneous attacks with multiple missiles[J]. Journal of Guidance, Control, and Dynamics, 2016, 39(10): 2439-2447.

[35] 李国飞,朱国梁,吕金虎,等.主-从多飞行器三维分布式协同制导方法[J].航空学报,2021,42(11):236-245.
LI G F, ZHU G L, LU J H, et al. Three-dimensional distributive cooperative guidance law for multiple leader-follower flight vehicles[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(11): 236-245(in Chinese).

[36] CHEN Y D, WANG J N, SHAN J Y, et al. Cooperative guidance for multiple powered missiles with constrained impact and bounded speed[J]. Journal of Guidance, Control, and Dynam-

ics,2021,44(4):825-841.

[37] WANG X H,TAN C P. 3-D impact angle constrained distributed cooperative guidance for maneuvering targets without angular-rate measurements[J]. Control Engineering Practice,2018,78:142-159.

[38] LYU T,LI C J,GUO Y N,et al. Three-dimensional finite-time cooperative guidance for multiple missiles without radial velocity measurements[J]. Chinese Journal of Aeronautics,2019,32(5):1294-1304.

[39] 方洋旺,邓天博,符文星. 智能制导律研究综述[J]. 无人系统技术,2020,3(6):36-42.
FANG Y W,DENG T B,FU W X. An overview on the intelligent guidance law[J]. Unmanned Systems Technology,2020,3(6):36-42(in Chinese).

[40] 金泽宇,刘凯,尹中杰,等. 基于神经网络剩余时间模型的协同制导律设计[J]. 战术导弹技术,2021(4):103-109.
JIN Z Y,LIU K,YIN Z J,et al. Impact time cooperation guidance law design based on time-to-go estimating model using neural network[J]. Tactical Missile Technology,2021(4):103-109(in Chinese).

[41] 陈中原,韦文书,陈万春. 基于强化学习的多发导弹协同攻击智能制导律[J]. 兵工学报,2021,42(8):1638-1647.
CHEN Z Y,WEI W S,CHEN W C. Reinforcement learning-based intelligent guidance law for cooperative attack of multiple missiles[J]. Acta Armamentarii,2021,42(8):1638-1647(in Chinese).

[42] 蔡远利,闫明明,刘佳琪. 基于深度强化学习的时间协同制导方法及仿真[C]//第 22 届中国系统仿真技术及其应用学术年会,2021:371-375.
CAI Y L,YAN M M,LIU J Q. Research on deep reinforcement learning based temporal cooperative guidance method and simulation[C]//22nd CCSSTA,2021:371-375(in Chinese).

[43] 周华,刘红军,王泽,等. 一种基于目标机动补偿的协同制导律[J]. 导弹与航天运载技术,2015(1):92-95.
ZHOU H,LIU H J,WANG Z,et al. Synergistic guidance law based on target maneuvering compensation[J]. Missiles And Space Vehicles,2015(1):92-95(in Chinese).

[44] 马萌晨,宋申民. 拦截机动目标的多导弹协同制导律[J]. 航空兵器,2021,28(6):19-27.
MA M C,SONG S M. Multi-missile cooperative guidance law for intercepting maneuvering target[J]. Aero Weaponry,2021,28(6):19-27(in Chinese).

[45] KUMAR S R,MUKHERJEE D. Terminal time constrained nonlinear interception strategies against maneuvering targets[J]. Journal Guidance,Control,and Dynamics,2021,44(1):200-209.

[46] BISHWASH H U S,SINHA A,KUMAR S R. Deviated pursuit based nonlinear impact-time guidance with finite-time convergence[J]. International Federation of Automatic Control,2020,53(1):93-98.

[47] 张保峰. 带攻击角度约束的多导弹协同导引律设计[D]. 哈尔滨:哈尔滨工业大学,2013:35-47.
ZHANG B F. Design of multi-missile cooperative guidance law with attack angle constraint[D]. Harbin:Harbin Institute of Technology,2013:35-47(in Chinese).

[48] FONOD R,SHIMA T. Blinding guidance against missiles sharing bearings-only measurements[J]. IEEE Transactions on Aerospace and Electronic Systems,2018,54(1):205-216.

[49] 杨雪冬,朱蕊蕾,李君龙. 交会时间可控的多子弹协同制导方法[J]. 现代防御技术,2013,41(1):59-63.
YNAG X D,ZHU R P,LI Y L. Submunitions cooperative guidance method of controllable rendezvous time[J]. Modern Defence Technology,2013,41(1):59-63(in Chinese).

[50] TSALIK R,SHIMA T. Circular impact-time guidance[J]. Journal of Guidance,Control,and Dymaics,2019,42(8):1836-1847.

[51] CHO D,KIM H J. Nonsingular sliding mode guidance for impact time control[J]. Journal of Guidance,Control,and Dymaics,2016,39(1):61-68.

[52] CHO N,LEE Y I. Guidance synthesis with separate impact point prediction considering speed variations[J]. Journal of Guidance,Control,and Dymaics,2021,44(2):428-440.

[53] HU Q L,HAN T. New impact time and angle guidance strategy via virtual target approach[J]. Journal of Guidance,Control,and Dymaics,2018,41(8):1755-1765.

[54] TEKIN R,ERER K S. Impact time and angle control against moving targets with look angle shaping[J]. Journal of Guidance,Control,and Dynamics,2020,43(5):1020-1025.

[55] SINHA A,KUMAR S R,MUKHERJEE D. Nonsingular impact time guidance and control using deviated pursuit[J]. Aerospace Science and Technology,2021,115:106776.

[56] 张友安,梁勇,刘京茂,等. 基于轨迹成型的攻击角度与时间控制[J]. 航空学报,2018,39(9):138-146.
ZHANG Y A,LIANG Y,LIU J M,et al. Trajectory reshaping based impact angle and time control[J]. Acta Aeronautica et Astronautica Sinica,2018,39(9):138-146(in Chinese).

[57] HU Q L,HAN T,XIN M. Sliding-mode impact time guidance law design for various target motions[J]. Journal of Guidance,Control,and Dynamics,2019,42(1):136-148.

[58] 吕腾,吕跃勇,李传江,等. 带空间协同的多导弹时间协同制导律[J]. 航空学报,2018,39(10):185-196.
LYU T,LYU Y Y,LI C J,et al. Time cooperative guidance law for multiple missiles with spatial cooperation[J]. Acta Aeronautica et Astronautica Sinica,2018,39(10):185-196(in Chinese).

[59] DONG W,WANG C,WANG J,et al. Fixed-time terminal angle constrained cooperative guidance law against maneuvering target[J]. IEEE Transactions on Aerospace and Electronic Systems,2022,58(2):1352-1366.

[60] 赵恩娇. 多飞行器编队控制及协同制导方法[D]. 哈尔滨:哈尔滨工业大学,2018:70-95.
ZHAO E J. Formation control and cooperative guidance methods for multiple flight vehicles[D]. Harbin:Harbin Institute of Technology,2018:70-95(in Chinese).

[61] 宋俊红. 拦截机动目标的有限时间制导律及多弹协同制导律研究[D]. 哈尔滨:哈尔滨工业大学,2018:124-146.
SONG J H. Research on finite-time guidance law and cooperative guidance law of multi-missiles for intercepting maneuvering target[D]. Harbin:Harbin Institute of Technology,2018:124-146(in Chinese).

[62] ZHANG S,GUO Y,LIU Z,et al. Finite-time cooperative guid-

ance strategy for impact angle and time control[J]. IEEE Transactions on Aerospace and Electronic Systems,2021,57(2):806-819.

[63] ZHANG P,ZHANG X. Multiple missiles fixed-time cooperative guidance without measuring radial velocity for maneuvering targets interception[J]. ISA Transactions,2022,126:388-397.

[64] TAN Z W,FONOD R,SHIMA T. Cooperative guidance law for target pair to lure two pursuers into collision[J]. Journal of Guidance,Control,and Dynamics,2018,41(8):1687-1699.

[65] 袁斐然,刘春生,陈必露. 基于自适应动态规划的多对一追逃博弈策略[J]. 电光与控制,2022,29(1):1-6.
YUAN F R,LIU C S,CHEN B L. Many-to-one pursuit-evasion game strategy based on adaptive dynamic programming[J]. Electronics Optics & Control,2022,29(1):1-6(in Chinese).

[66] 孙景亮. 自适应动态规划研究及其在导弹拦截制导中的应用[D]. 南京:南京航空航天大学,2019:77-89.
SUN J L. Research on adaptive dynamic programming and its applications to the missile interception guidance[D]. Nanjing:Nanjing University of Aeronautics and Astronautics,2019:77-89(in Chinese).

[67] LONG T,CAO Y,SUN J L,et al. Adaptive event-triggered distributed optimal guidance design via adaptive dynamic programming[J]. Chinese Journal of Aeronautics,2022,35(7):113-127.

Review and prospect of cooperative guidance with attack time consensus for multiple aerial vehicles

DONG Xiwang[1,2], YU Jianglong[2,*], HUA Yongzhao[1], LI Qingdong[2], REN Zhang[1,2]

(1. Institute of Artificial Intelligence, Beihang University, Beijing 100083, China;
2. School of Automation Science and Electrical Engineering, Beihang University, Beijing 100083, China)

Abstract: Cooperative guidance issue with attack time consensus is an important branch of cooperative guidance issues. The main research results and classification methods of cooperative guidance with attack time consensus for multiple aerial vehicles are summarized. Firstly, the tactical significance of cooperative guidance with attack time consensus is analyzed. Secondly, two categories of issues are researched depending on how the target moves: cooperative guidance for fixed targets and cooperative guidance for moving or maneuvering targets. Furthermore, a review of the main research results of these two types of cooperative guidance problems is provided. The advantages, disadvantages and scenarios of the above cooperative guidance methods are summarized from the level of research methods. Finally, a summary of the related technical challenges of cooperative guidance with attack time consensus is given, along with a prospective direction for future development.

Keywords: cooperative guidance; attack time consensus; attack time; consensus tracking; communication

Received: 2022-05-17; **Accepted**: 2022-06-09; **Published online**: 2022-07-05 13:46

URL: kns.cnki.net/kcms/detail/11.2625.V.20220704.1510.001.html

Foundation items: Science and Technology Innovation 2030-Key Project of "New Generation Artificial Intelligence" (2020AAA0108200); National Natural Science Foundation of China (62103016,61922008,61873011,61973013,61803014); China National Postdoctoral Program for Innovative Talents (BX20200034); China Postdoctoral Science Foundation (2020M680297)

* **Corresponding author**. E-mail: sdjxyjl@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:100083 北京市海淀区学院路 37 号 《北京航空航天大学学报》编辑部

办公地点:北京航空航天大学办公楼 405,407 房间

电　　话:(010)82316698,82316981,82315499

E-mail: jbuaa@ buaa. edu. cn

http://bhxb. buaa. edu. cn

http://www. buaa. edu. cn

《北京航空航天大学学报》
第五届编辑委员会

主　任（主　编）：赵沁平

（以下按姓氏笔画为序）

副主任（副主编）：丁希仑　王少萍　孙志梅　李秋实　李焕喜　杨嘉陵
苗俊刚　相　艳　徐立军　钱德沛　曹晋滨

编　委：马殿富　王　琪　王　聪　邓小燕　王青云　王荣明　刘　宇
刘　红　江　洁　刘　强　闫　鹏　朱天乐　刘铁钢　齐铂金
陈万春　邹正平　苏东林　杨世春　沈成平　邱志平　宋知人
杨树斌　张晓林　杨晓奕　杨继萍　李惠峰　吴新开　张瑞丰
杨照华　宋凝芳　周　锐　林宇震　林贵平　战　强　姚仰平
胡庆雷　赵秋红　段海滨　赵巍胜　席　平　郭　宏　徐　洁
徐世杰　郭洪波　康　锐　翟　锦　熊华钢

北京航空航天大学学报

Beijing Hangkong Hangtian Daxue Xuebao

（原《北京航空学院学报》）

（月刊　1956年创刊）

第48卷第9期　2022年9月

主管单位　中华人民共和国工业和信息化部
主办单位　北京航空航天大学
主　　编　赵沁平
编辑出版　《北京航空航天大学学报》编辑部
邮　　编　100083
地　　址　北京市海淀区学院路37号
印　　刷　北京科信印刷有限公司
发　　行　《北京航空航天大学学报》编辑部
发行范围　国内外发行
联系电话　(010) 82316698 82316981 82315499
电子信箱　jbuaa@buaa.edu.cn

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS (JBUAA)

(Monthly, Started in 1956)

Vol.48 No.9 September 2022

Administrated by Ministry of Industry and Information Technology of the People's Republic of China
Sponsored by Beijing University of Aeronautics and Astronautics (BUAA)
(Beijing 100083, P. R. China)
Chief Editor ZHAO Qinping
Edited and Published by Editorial Board of JBUAA
Printed by Beijing Kexin Printing Co., Ltd.
Distributed by Editorial Board of JBUAA
Telephone (010) 82316698 82316981 82315499
E-mail jbuaa@buaa.edu.cn
http://bhxb.buaa.edu.cn

中国标准连续出版物号：ISSN 1001-5965 / CN 11-2625/V　　国内定价：50.00元／期

ISSN 1001-5965