主管单位：中华人民共和国工业和信息化部
主办单位：北京航空航天大学

ISSN 1001-5965
CN 11-2625/V

北京航空航天大學学报

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS



非生物环境因子干扰对微生物群落演替影响的研究进展

胡大伟　李帅帅　刘光辉

9 2024
Vol.50

北京航空航天大学学报

第 50 卷　第 9 期　（总第 379 期）　2024 年 9 月

目　　次

非生物环境因子干扰对微生物群落演替影响的研究进展 …… 胡大伟，李帅帅，刘光辉（2677）
多源知识融合的方面级情感分析模型 …… 韩虎，郝俊，张千锟，赵启涛（2688）
基于到达时间间隔差的 ADS-B 位置消息验证方法 …… 刘海涛，刘家祥，李冬霞，王磊（2696）
新型 ADETCS 下工业信息物理系统综合安全控制 …… 李炜，程雪，李亚洁（2704）
基于暗通道先验引导的图像去雾网络 …… 黄淑英，夏钰锟，杨勇，万伟国，邱根莹（2717）
舵机电动负载模拟器迭代学习控制 …… 刘晓琳，李晋恺（2727）
深度神经网络模型任务切分及并行优化方法 …… 巨涛，刘帅，王志强，李林娟（2739）
面部三维生理特征点在睡眠剥夺中的疲劳检测 …… 李敬强，房秋，樊天辰，马龙（2753）
基于图结构和序列特征融合的关系抽取 …… 武同心，纪鑫，王宏刚，杨智伟，陈屹婷，赵加奎（2763）
基于 BP 神经网络的甲板运动预报与补偿技术 …… 赵所，林立，李震，侯中喜（2772）
基于联合仿真的曲面共形多孔结构拓扑优化方法 …… 付君健，徐勇，周祥曼，舒正涛，田启华（2781）
基于多通道图神经网络的属性聚合式实体对齐 ……
…… 纪鑫，武同心，王宏刚，杨智伟，何禹德，赵晓龙（2791）
类双锥型升力式返回器横航向不稳定性分析及改进 ……
…… 陈刚，孙雪，李广兴，赵丹，马建颖，何程（2800）
基于中国鸟情的鸟撞适航条款充分性和适宜性 …… 陆晓华，蔡景，张柱国，张迎春（2810）
考虑视觉搜索特性的指挥控制系统 KLSM …… 郭仲，赵刚，付斌，刘佳烨（2819）
基于机器学习的高超声速飞行器双曲率前缘气动热预测方法 …… 杨帆，林明月，胡宗民，罗长童（2826）
复合坐标系跨临界液滴表面 Kelvin-Helmholtz 不稳定增长高效预测方法 ……
…… 崔海港，闫啸宇，高岩飞，陈飞，杨世春（2835）
扫频雷达双频连续波伪信号干扰效应 …… 赵宏泽，魏光辉，潘晓东，杜雪，卢新福（2843）
大行程磁阻驱动微定位平台及其轨迹跟踪控制 …… 张旭，陆晓伟，赖磊捷（2852）
多策略融合改进的飞蛾火焰优化算法 …… 何加文，许贤泽，高波（2862）
动态场景中基于特征点几何约束的 3D SLAM 算法 …… 刘铭坚，罗景文，秦世引（2872）
大规模卫星星座组网的码分多址干扰分析 … 徐可笛，徐兆斌，郭晓旭，杨嘉，成恒飞，金仲和（2885）
基于近似计算技术的多级逻辑电路面积优化 …… 赵维凯，王伦耀，岑旭梦，夏银水，储著飞（2893）
基于竞争情景的双枢纽航线网络优化模型 …… 何湘妮，李艳华，阳杰（2902）
双弯曲旋转超声椭圆振动加工系统设计 …… 邵壮壮，郑侃，董松，孙振文，孙连军，严嘉文（2912）
柱塞泵配流盘卸荷槽开度对瞬态流场特征影响 …… 刘珊珊，雷欣瑞，宋宇宇，杨凯，黄家海（2919）
空间站舱内悬浮实验台制导控制方法 …… 卢力铮，刘伟，张皓（2930）
面向区域覆盖的多无人机动态通信资源分配方法 …… 卢毛毛，刘春辉，董赞亮（2939）
空间飞网系统动力学建模与仿真 …… 陈鹏旭，吴晨晨，倪智宇（2951）
多网络控制器 FC-AE-1553 动态带宽调度机制 …… 李心茹，张馨月，李铮，何锋（2963）

期刊基本参数：CN 11-2625/V * 1956 * m * A4 * 298 * zh * P * ￥50.00 * 400 * 30 * 2024-09

（**编　辑**　张　嵘　孙　芳　卞欢欢　王　茜　李亚泰）

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Vol. 50 No. 9 (Sum 379) September 2024

CONTENTS

Research progress on influence of disturbance of abiotic environmental factors on microbial community succession
…… HU Dawei, LI Shuaishuai, LIU Guanghui (2677)
Multi-source knowledge fusion model for aspect-based sentiment analysis
…… HAN Hu, HAO Jun, ZHANG Qiankun, ZHAO Qitao (2688)
Verification scheme of position message for automatic dependent surveillance-broadcast system based on interval difference of arrival
…… LIU Haitao, LIU Jiaxiang, LI Dongxia, WANG Lei (2696)
Integrated security control of industrial cyber-physical systems based on new type ADETCS
…… LI Wei, CHENG Xue, LI Yajie (2704)
Image dehazing network based on dark channel prior guidance
…… HUANG Shuying, XIA Yukun, YANG Yong, WAN Weiguo, QIU Genying (2717)
Iterative learning control of electric load simulator of aircraft steering gear
…… LIU Xiaolin, LI Jinkai (2727)
Task segmentation and parallel optimization of DNN model
…… JU Tao, LIU Shuai, WANG Zhiqiang, LI Linjuan (2739)
Fatigue detection of facial 3D physiological feature points in sleep deprivation
…… LI Jingqiang, FANG Qiu, FAN Tianchen, MA Long (2753)
Relation extraction based on fusion of graph structure and sequence features
…… WU Tongxin, JI Xin, WANG Honggang, YANG Zhiwei, CHEN Yiting, ZHAO Jiakui (2763)
Deck motion prediction and compensation technology based on BP neural network
…… ZHAO Suo, LIN Li, LI Zhen, HOU Zhongxi (2772)
Topological optimization method for conformal cellular structures on surfaces based on co-simulation
…… FU Junjian, XU Yong, ZHOU Xiangman, SHU Zhengtao, TIAN Qihua (2781)
Attribute aggregation entity alignment based on multi-channel graph neural network
…… JI Xin, WU Tongxin, WANG Honggang, YANG Zhiwei, HE Yude, ZHAO Xiaolong (2791)
Analysis and improvement of lateral instability of quasi-biconical lifting reentry spacecraft
…… CHEN Gang, SUN Xue, LI Guangxing, ZHAO Dan, MA Jianying, HE Cheng (2800)
Adequacy and suitability of airworthiness clause of bird strike based on bird situation in China
…… LU Xiaohua, CAI Jing, ZHANG Zhuguo, ZHANG Yingchun (2810)
Considering visual search characteristics on command and control system KLSM
…… GUO Zhong, ZHAO Gang, FU Bin, LIU Jiaye (2819)
Prediction method of aero-heating of hypersonic vehicle bi-curvature leading edge based on machine learning
…… YANG Fan, LIN Mingyue, HU Zongmin, LUO Changtong (2826)
Efficient prediction method for Kelvin-Helmholtz instability growth on transcritical droplet surface in composite coordinate system
…… CUI Haigang, YAN Xiaoyu, GAO Yanfei, CHEN Fei, YANG Shichun (2835)
Dual-frequency continuous wave pseudo-signal interference effect in swept-frequency radar
…… ZHAO Hongze, WEI Guanghui, PAN Xiaodong, DU Xue, LU Xinfu (2843)
Large-stroke microposition stage driven by reluctance actuator and its trajectory tracking control
…… ZHANG Xu, LU Xiaowei, LAI Leijie (2852)
Improved moth-flame optimization algorithm with multi-strategy integration
…… HE Jiawen, XU Xianze, GAO Bo (2862)
3D SLAM algorithm based on geometric constraints of feature points in dynamic scenarios
…… LIU Mingjian, LUO Jingwen, QIN Shiyin (2872)
Analysis of CDMA interference in large-scale satellite constellation
…… XU Kedi, XU Zhaobin, GUO Xiaoxu, YANG Jia, CHENG Hengfei, JIN Zhonghe (2885)
Area optimization of multilevel logic circuits using approximate computing
…… ZHAO Weikai, WANG Lunyao, CEN Xumeng, XIA Yinshui, CHU Zhufei (2893)
Optimization model for dual hub airline networks based on competition scenario
…… HE Xiangni, LI Yanhua, YANG Jie (2902)
Design of double bending rotary ultrasonic elliptical vibration machining system
…… SHAO Zhuangzhuang, ZHENG Kan, DONG Song, SUN Zhenwen, SUN Lianjun, Yan Jiawen (2912)
Influence of unloading groove opening of port plate of plunger pumps on transient flow field characteristics
…… LIU Shanshan, LEI Xinrui, SONG Ziyu, YANG Kai, HUANG Jiahai (2919)
Guidance and control method of levitation experiment facility inside China's Space Station
…… LU Lizheng, LIU Wei, ZHANG Hao (2930)
Dynamic communication resource allocation for multi-UAV area coverage
…… LU Maomao, LIU Chunhui, DONG Zanliang (2939)
Dynamic modelling and simulation of a tethered-net in space
…… CHEN Pengxu, WU Chenchen, NI Zhiyu (2951)
FC-AE-1553 dynamic bandwidth scheduling mechanism for multi network controllers
…… LI Xinru, ZHANG Xinyue, LI Zheng, HE Feng (2963)

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0736

非生物环境因子干扰对微生物群落演替影响的研究进展

胡大伟[1]，李帅帅[1,2]，刘光辉[3,4,*]

(1. 北京航空航天大学 生物与医学工程学院，北京 100191；2. 麦吉尔大学 生物系，蒙特利尔 H3A1B1；
3. 北京航空航天大学 航空科学与工程学院，北京 100191；4. 中国科学院空间应用工程与技术中心，北京 100094)

摘 要：了解驱动群落演替的因素是微生物生态学的中心目标之一，非生物环境因子影响着单个种群的生长及种群之间的相互作用。微生物群落非常敏感，很容易对作为干扰的非生物环境因子的变化产生响应。研究不同类型的干扰如何产生不同的影响，将增强对微生物群落多样性、结构和功能之间关系的认识。以作用对象为分类标准，梳理了各种非生物环境因子干扰对微生物群落演替的影响研究，包括非特异性、特异性及航天等特殊环境下的非生物环境因子。从干扰程度的量化、多因素干扰影响及干扰后演替的时间过程等方面，得出研究干扰下群落多样性演替机制的关键问题。展望了未来需要探索的方向，并提出相关研究建议。

关 键 词：非生物环境因子；干扰；群落演替；微生物；特殊环境
中图分类号：Q938.1+5; V7
文献标志码：A **文章编号**：1001-5965(2024)09-2677-11

生态学的机理方法寻求在功能上将个体特征与更高层次的过程联系起来[1]，如种群动态和多种群相互作用、物种分布的决定因素、群落结构的发展和维持[2-3]。生态系统的各个组成部分在不同时间和空间尺度上以复杂的方式相互作用，并推动系统的组建和功能维持。群落是具有多种相互作用种群的集合[4]，微生物群落是相互作用的微生物的组合；几乎所有的微生物都生活在群落中。在土壤和人体等多种生境中，微生物群落是关键生态过程的实际驱动者[5]，但目前对于微生物群落的形成和演替知之甚少[4]。由于群落的多样性和组成方式调节了生态系统的功能[6-7]，了解驱动群落演替的因素仍然是生态学[8-9]，尤其是微生物生态学的中心目标[10]。

群落的基本特征包含物种组成、多样性和相互作用[11]。非生物环境因子影响着单个种群的生长及种群之间的相互作用[12]，这种影响在生态学上被定义为干扰，其在物理意义上抑制、伤害或杀死群落中的某些个体，产生绝对数量或相对多度的变化[13]，为其他个体的生长或繁殖创造机会[14]，被认为是影响物种多样性[15]和生态系统结构变化的主要因素[16-19]。大多数微生物群落对干扰很敏感[16,20]，如抗生素的使用对肠道微生物的影响[21]，土壤细菌群落的组成在野火后由基于生态位转变为随机组合[22]。在这些过程中，干扰的不同梯度驱动着多物种群落结构组成以复杂的方式变化[1,19]。由于梯度揭示了物种特征和物种集合在不断变化的生态环境下的一致性模式[23]，对非生物环境因子的梯度进行研究，可以更深入理解低层级过程如何影响高层级过程[1]。

虽然非生物环境因子对群落的作用力会产生干扰，造成破坏并可导致生物的死亡或生物量的损

收稿日期：2022-08-22；**录用日期**：2022-09-09；**网络出版时间**：2022-11-02 16:30
网络出版地址：link.cnki.net/urlid/11.2625.V.20221102.1028.001
基金项目：国家重点研发计划 (2021YFA0716100)；国家自然科学基金国际 (地区) 合作研究项目 (32261133528)
* **通信作者**. E-mail：liugh1991@126.com

引用格式：胡大伟，李帅帅，刘光辉. 非生物环境因子干扰对微生物群落演替影响的研究进展 [J]. 北京航空航天大学学报，2024，50（9）：2677-2687. HU D W，LI S S，LIU G H. Research progress on influence of disturbance of abiotic environmental factors on microbial community succession [J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50（9）：2677-2687（in Chinese）.

失,其本身虽不能使得生物多样性增加,但作为群落的一种选择力量是显而易见的。通常的观点认为,干扰通过降低物种密度来减缓竞争排斥作用,防止少数最具有竞争优势的物种占有主导地位,从而有利于生物多样性的维持[24]。但也有观点认为,当选择过程对不同的物种有偏好时就会产生竞争排斥,干扰只是减缓了竞争排斥过程,并未改变物种多样性减少的结果,只有负频率依赖选择同时存在时,干扰才会促进物种共存和多样性的维持[25];波动较大的群落有着较低的“有效群落大小”,干扰实际上减少了物种多样性[26];对微生物群落来说,其多样性通常会因为环境压力扰乱了群落内种群相互作用的生态平衡而减少[27],但尽管如此,在不同的条件下,被破坏的微生物群落表现出了生存能力提高的现象,表明了对干扰的普遍适应[28]。总的来说,群落结构会随着非生物环境因子而改变[12]。

干扰有不同维度的分类方式,研究不同类型的干扰如何产生不同的影响,将增强对多样性、结构和功能之间关系的认识[29]。从因果逻辑上,干扰或改变环境,并可能对群落产生间接影响;或直接改变群落[13]。从持续时间上,干扰可以被归类为脉冲事件或压力事件[30]。一般来说,脉冲干扰(disturbance)是相对离散、短期的事件,系统功能最终可能恢复到原来的状态[13],并且对功能表现基本没有影响[29];而压力干扰(stressor,也称压力源,由于行文连贯性需要,本文统称为干扰)是长期、连续的,将导致一个新的状态[13],功能表现变化较慢,并且对额外的干扰更敏感[29]。当然,根据所研究时间尺度的不同,对长短的定义也会有相应的变化[16]。从作用形式上,非特异性的干扰会广泛影响各种生物,而特异性的干扰只改变少数生物[29]。从作用程度上,频率依赖选择,如人为干扰,有选择性地消除某些物种并偏向于保留当地更常见的物种,β多样性可能比未受干扰的原始环境低;反之,负频率依赖选择可能会产生更高的β多样性。以上2种情况中,群落的成分更替与人为干扰的环境梯度密切相关;如果干扰是完全随机的,则β多样性与原始环境相似[19]。

目前,对于干扰在生态学领域的研究主要集中于植物群落等宏观尺度,而对于微观尺度(如微生物群落的演替)的研究尚处于起步阶段。从理论生态学模型建立的角度,以作用形式的分类更能清晰地在数学公式所含项中区分。因此,本文以作用形式为分类标准,同时考虑特殊环境中的独有情况,梳理了各种非生物环境因子干扰对微生物群落演替的影响研究。

1 非特异性非生物环境因子

非特异性非生物环境因子是在生态系统中对生物体产生无差别影响的因子,如稀释、环境连通性、物理位置变化等。

1.1 稀释对微生物群落演替的影响

非生物环境因子干扰的常见效应包括死亡率的改变[12],以及在没有进一步干扰的情况下,群落可以维持多久。这2个效应可在实验室实验中得到实现[31],即培养和共培养细菌的标准方法——周期性地将微生物群落(或种群)稀释到新鲜培养基中。虽然这种连续稀释法很少能反映自然条件,但可以估计种群密度和演替中的特征,从而检验各种假设并理解生态学原理。在稀释的过程中,稀释率(转移体积)和稀释频率(转移间隔)2个关键参数对于多物种群落的组成影响的研究还处在比较初步的阶段。

1.1.1 稀释率对微生物群落演替的影响

理论模型预测,由于稀释导致的群落中所有物种的死亡率同等增加,可对群落组成产生巨大影响,尤其有利于成对共培养中生长较快的物种,从而使占优势地位的物种由生长较慢者转变为生长较快者[11]。而一些在恒化器中进行的不同生长速率的微生物实验也支持理论预测[32-33]。

Abreu等[12]对稀释率的研究,通过实验证实了理论模型的预测,使用简单的两种群Lotka-Volterra(LV)竞争模型,并且增加了由稀释引发的死亡率一项,从而为稀释条件下群落的变化提供了合理的模型;发现在实验所构建的体系中,生长速率和物种竞争力之间存在普遍的权衡,在高密度、低稀释率环境中生长较慢的物种具有优势地位,而随着稀释率的增加,结果会发生显著变化,稀释造成的死亡率在很大程度上决定了物种生长速度对竞争表现的重要性。这种权衡在自然界中普遍存在,如在微生物中,抗生素的抗药性虽然明显具有竞争优势,但会使得细菌付出代价;在植物中,种子较大的植物必然产生较少的种子,但这种种子在育苗时具有更强的竞争力。

而在自然界中,环境条件会随着时间的推移而发生变化,季节变化导致植物群落、湖泊微生物群落和海洋微生物群落具有可重复性的演替模式,不同的环境条件甚至可能导致生态系统突然和不可逆转的变化。但到目前为止,很少有研究关注干扰除增加多样性外的其他结果。因此,Abreu等[34]探究了时间波动的干扰对微生物群落的影响,通过交

替稀释率来实现液体培养基中环境的波动, 在双种群共同培养中, 观察到波动的稀释率导致了双稳态, 即依赖于每个种群初始相对多度的 2 种可互相替代的稳定状态, 在 2 个极端稀释率之间的恒定中等稀释率也具有双稳态, 表明简单的群落可以由平均增长率对竞争系数的线性依赖模拟而来。而由添加了稀释导致的死亡率的 Lotka-Volterra 竞争模型模拟预测结果来看, 作为线性模型, 其预测的与稀释率时间波动无关的结果与实际也相符[35]; 更广泛的推测, 稀释率波动的环境, 与时间上具有平均稀释率的环境, 具有相同的结果。Abreu 等[34] 使用一个更复杂的三物种群落检验了预测, 结果表明, 用稀释率的波动预测对群落结构的影响是合理的。

在 Scheuerl 和 Kaitala 对稀释影响的研究中, 稀释率控制了每次稀释间隔的生长期中的初始条件, 作为主要因素用于研究具有捕食者-猎物关系的纤毛虫-细菌的双种群群落的变化[31]。使用添加了允许相互作用共同进化因子的半连续 Lotka-Volterra 竞争模型[36], 模拟了不同稀释率下捕食者-猎物系统的生态动力学和进化动力学现象; 发现提高稀释率会使纤毛虫和细菌的共存关系受到威胁, 进化上的变化受限, 而降低稀释率则会导致进化; 认为可能是由于稀释率过高时, 没有足够的猎物, 捕食者无法获取足够的食物, 因此, 纤毛虫最先灭绝。这项研究中使用的模型表明, 从高稀释率到中等稀释率有明显的进化上的变化, 但从中等稀释率到低稀释率进化上的变化并不明显, 推测是因为在未受进化保护的情况下, 作为猎物的细菌即使发生了很小的防御上的进化也会受益匪浅, 而在后续的阶段中这种影响变得不再明显。

1.1.2　稀释频率对微生物群落演替的影响

在 Scheuerl 和 Kaitala[31] 对稀释影响的研究中, 稀释频率也是主要因素, 其主要决定下一个生长期开始时捕食者纤毛虫和猎物细菌的比例。发现稀释频率的上升会加速种群的灭绝, 并且使得进化率降低; 而稀释频率的下降则有相反的影响。随着稀释率的增加和稀释频率的降低, 细菌的防御及纤毛虫的捕食都有所增强, 显示出了军备竞赛的动态效应 (arms-race dynamics)。通过对稀释率和稀释频率的比较, 认为稀释频率对所构建的双物种群落生态动力学和进化动力学的影响更大, 并且生态动力学对稀释频率的变化比对稀释率的变化更敏感, 特别是在实验的初始阶段。

生长速率、竞争能力及资源利用效率之间的权衡可能是微生物对干扰作出反应的能力基础[37]。高生长速率的微生物往往有着较高的核糖体组分合成速率, 从而减少了对有利生长条件的反应时间, 当有利的生长条件出现在脉冲式干扰之后时, 具有最快反应的微生物将最先增长并由此改变群落组成[38]。生长速率对微生物群落在脉冲式干扰下的适应能力具有重要意义, 特别是在干扰产生新资源的情况下, 少数存活的个体能够在突然的脉冲式干扰后迅速生长, 恢复到干扰前的种群规模[16]。

1.2　空间因素对微生物群落演替的影响

微生物群落通常是在结构化的环境中进行演替的[39], 交互通常以局部的方式发生[40]。Harcombe[41] 的实验直接证明了结构化的空间环境是导致个体倾向于合作行为和局部相互作用的决定性因素, 这有助于交叉互惠的稳定。Raynaud 和 Nunan[42] 的研究表明, 尽管土壤中的微生物个体数量和物种数量非常多, 但其只与临近的一些个体相互作用。空间结构在引起物种多样性的变化[4] 及群落演替[43] 的过程中起着重要的作用, 当需要完整地描述生态过程时, 需要同时考虑时间和空间因素, 尤其是空间因素[44]。

1.2.1　环境物理性质对微生物群落演替的影响

环境干扰对陆地水生生态系统的维持构成了巨大的威胁, 尤其是在水体流向和分布改变的情况下[45], 对液态水主要来自融化冰雪的冰冻圈的影响尤为突出[46]。南极洲极地沙漠的淡水生态系统是南极洲内陆生物多样性的重要组成部分, 且其中绝大部分是微生物[47]。南极淡水生态系统的几个特定条件影响着栖息地之间的连接程度[48]。Ramoneda 等[46] 对南极淡水生态系统细菌群落的研究认为, 这与细菌繁殖体在生境之间的传播有关, 可能在确定局部微生物群落的组成上发挥着重要的作用。尽管该系统浅部的细菌群落之间有较高的物理连接性和扩散机会, 但非生物环境条件在驱动群落结构演替上主导了扩散。功能代谢途径预测表明, 浅水和深水微生物群落之间的功能基因潜力存在差异。因此, 未来气候变化导致的温度升高和融水将影响南极淡水生态系统中细菌的多样性和功能。

河流生态系统中的细颗粒泥沙沉降与水的流速对微生物群落也具有不可忽视的影响。Juvigny-Khenafou 等[49] 的研究评估了溪流底栖细菌群落在与城市化和农业集约化相关的多重干扰(如降低流速和添加细颗粒沉积物)下的变化。细颗粒沉积物通过覆盖河床表面, 形成了一道物理屏障, 阻碍了水柱和水源带之间的氧气和营养物质的交换[50]; 使得河床表面的基底颗粒尺寸和表面粗糙度减小, 这

些变化都有利于能够适应新环境条件的细菌群落[51]，可能影响微生物的定殖模式[49]。流速的减小增加了河床上方边界层的厚度，从而降低了与水柱的资源交换率[52]；降低的流速有利于细菌从流动的水中沉降，这与之前的研究[53-54]相反。

1.2.2 位置变换对微生物群落演替的影响

人为造成的微生物所处位置的变换可以增加或减少群落之间的成分周转[19]，会对其群落结构造成影响，在现实生活中，最常见的例子是对土壤的耕作[55-56]。越来越多的证据表明，土壤微生物在多种生态系统功能中起着至关重要的作用，如支持植物生产力，维持地上和地下群落之间养分循环的过程等[57-58]。维持土壤微生物群落的稳定性和丰富程度有利于提升土壤资源循环效率和土壤健康程度，并且有助于植物的生产力和产量的提高[55]。有研究证明，土壤微生物群落组成在时间上的变化受到土地耕作的影响[59]；强烈的土地耕作干扰会导致土壤微生物多样性和多度降低，并且诱导成分上的变化，从而改变群落组成，影响生态系统中资源的循环和每日、每季度和每年度的维持过程[60]。有研究观察到，土壤耕作导致微生物多度的短期变化，这种变化与增加养分的淋溶和土壤反硝化作用对养分循环的破坏是一致的[61]。

Wagg 等[55]通过长期的土地管理实验，对比了有无对土壤进行耕作的条件下，土壤微生物群落的多样性、同步性和稳定性的差异，发现对土壤生态系统的耕作，作为干扰，会降低土壤微生物群落的丰富度 (richness)，扰乱时间动态，破坏土壤微生物的多度 (abundance)。真菌和细菌净多度时间变化机制存在差异，真菌和细菌群落之间不同的时间反应可能与土壤结构物理变化引起的土壤性质的变化有关。由于真菌在生长过程中产生大量的菌丝，在土壤结构和菌丝网络被物理破坏后，真菌的丰富度和多度会大大减少，需要更长的时间来恢复[55]。与真菌群落相比，耕作干扰对细菌群落的丰富度和多度的负面影响不显著，但也有研究认为这可能反映了土壤细菌群落在环境干扰后快速恢复和适应的能力[20]。另外，Wagg 等[55]的研究认为，土地耕作降低了细菌群落组成的时间变异性，使得细菌群落变得不稳定。这与 Lauber 等[59]的研究结果也互相支持。因为具有变异性的种群间时间的不同步会导致更稳定的细菌多度，即补偿动态(compensatory dynamics)[62-64]，其机理除土壤中淤泥含量的减少、pH 值的增加以外，还有可能与耕作引起的一般土壤均质化有关[55]。Habig 等[56]研究了在 4 年的耕作周期中，不同耕作顺序(麦-苜蓿轮作和麦单作)和不同土壤干扰程度(常规耕作和免耕)对土壤微生物多样性指数的影响，并结合其他指标综合评价土壤生态系统的功能，结果表明，免耕和轮作可以提高土壤微生物的丰富度和均匀度。Silva[65]利用 16S-rDNA-DGGE 图谱发现，随着耕作的加强，细菌的多样性会减少。

微生物群落在空间上的规模化和快速扩散的能力在干扰后的群落恢复中发挥了重要作用[66]。分散是元群落统一理论的一个关键特征，该理论认为群落是通过个体在不同环境中的移动而联系在一起的[67]，干扰通过杀死局部的某些从属同一分类群的个体而释放资源、创造空的小生境，从而启动作为迭代事件的群落重构。在重构过程中，这些被释放出来的生态位可以由来自本群落其他位置的微生物填补，从而在受到干扰后促进群落的恢复能力[16]。例如，由于来自未受干扰群落的微生物资源的不断扩散，被脉冲干扰而盐度增加的淡水细菌群落在组成和功能上都具有较强的恢复力[68]。但从另一个角度来讲，如果干扰广泛蔓延，影响到整个区域，那么扩散可促进群落之间耐干扰物种的传播，从而改变群落的优势种，降低局部的整体群落恢复力。此外，干扰事件所打开的生态位也受到随机定殖的、分散的微生物的影响，由于“优先效应”(priority effects)，早期干扰后的定殖微生物可能会长期存在，与原有的群落成员竞争[69]，影响随后分散的微生物定殖的可能性[70]。这些研究表明，与局部群落的生物特性、干扰特性产生相互作用的微生物的扩散，可以对群落的重构产生影响。

2 特异性非生物环境因子

特异性非生物环境因子是在生态系统中对生物体产生有差别影响的因子，如对不同物种各有最优条件的温度、氧气、pH 值等自然环境参数，还包括人类活动排出的重金属和有机物等。

2.1 环境参数对微生物群落演替的影响

在微生物群落演替的过程中，虽然局部的环境参数基本一致，但由于群落内各个物种具有不同的最适温度、氧气、pH 值等参数条件，会造成群落的演替。

温度和氧气是水生细菌群落演替的关键驱动力[71]。Shade 等[17]通过 16S 核糖体 RNA 基因标签焦磷酸测序和自动核糖体基因间隔区分析(automated ribosomalintergenic spacer analysis, ARISA)来评估浮游细菌群落，研究了其对温带湖泊中温度和氧气干扰的反应，结果表明，在干扰之后的 7 d (湖泊表层微生物群落)和 11 d(湖泊深层微生物群

落)内, 群落的组成和多样性均恢复到干扰前的水平, 显示出微生物群落在遭受重大干扰后有恢复的能力。

土壤的盐分和 pH 值属于影响土壤微生物群落的最重要因素, 其调节了群落内部的相互作用[72]。Zeng 等[73] 为了解土壤微生物对多重干扰的梯度响应, 分析了土壤微生物群落的结构和功能, 结果表明, 土壤电导率与 OTU 水平上的细菌群落密切相关, pH 值与细菌群落呈显著正相关, 高电导率和低 pH 值降低了大多数优势菌门的多度, 但提高了厚壁菌门(*Firmicutes*)、异常球菌-栖热菌门(*Deinococcus-Thermus*)和硝化螺旋菌门(*Nitrospirae*)的多度, 显著增加了几个双组分系统相关、电子转移相关或甲烷合成相关模块的多度。在 1.2.2 节所述的长期土地耕作中, 对于土壤的耕作, 除了更改土壤结构和破坏真菌菌丝网络外, 还会增加土壤的 pH 值, 减少土壤的淤泥含量[55]。这种 pH 值和淤泥特性的变化被认为与细菌和真菌的多度和群落组成变化有关[74], 有研究观察到真菌的多度和丰富度与较高的黏土和淤泥含量及较低的 pH 值呈正相关[75], 这也可能导致了真菌和细菌在多度和群落组成上的不同反应[55]。然而, 针对不同的土壤类型, 受干扰产生的结果也各不相同。一项基于长期施用石灰试验的典型研究表明, pH 值对土壤细菌群落的影响比对土壤真菌群落的影响更显著[74]; 另一项研究表明, 增施氮肥可以显著提高土壤的电导率、降低 pH 值, 土壤电导率与微生物群落结构呈正相关, 而 pH 值与微生物群落结构和其功能结构呈负相关[76]。

2.2　重金属和有机物对微生物群落演替的影响

在人类活动中, 向自然环境释放的物质会对其中的微生物群落产生影响, 如意外事故排出的化石燃料, 由于环境保护手段不够完全而在工业化过程中排出的重金属离子, 为控制生态系统的状态而主动释出的杀虫剂、杀菌剂等。在这些过程中, 微生物群落受到化学污染的干扰, 分类学意义和遗传学意义上的多样性均低于未受干扰的相同群落。但受干扰群落中的优势种群具有较强的生理耐受性和基质利用能力, 表明广义生理多样性是受干扰群落中竞争成功种群的适应性特征, 具有对生态干扰的普遍适应能力[28]。

工业化生产和农业灌溉导致了具有毒性的土壤重金属含量的增加[77], 长期暴露于吸收高浓度的重金属离子对包括微生物群落在内的生态系统的各个成分都是有害的。重金属离子污染物影响了微生物群系和功能基因, 减少了微生物的多样性, 进一步改变了土壤微生物群落的结构[78-80]。Müller 等[29] 研究了长期暴露于重金属汞对土壤细菌群落的影响, 发现其多样性减少, 而基质利用率与对照组相比没有明显差异, 汞污染的土壤对新干扰的抵抗力降低。在 Zeng 等[73] 对中国西北干旱和轻度盐渍化黄土中微生物对重金属干扰的反应研究中, 土壤中的高电导率和重金属元素的高含量明显降低了土壤细菌和真菌的多样性、丰富度和 OTU 数。以往的研究在高度铅污染的土壤中发现了丰富的变形菌门(*Proteobacteria*)细菌[81-82], 但令人惊讶的是, 在 Zeng 等[73] 的研究中, 虽然土壤中的铅含量很高, 但在高 pH 值下, 铅与细菌群落的关系并不显著, 与低 pH 值下铅含量较高的亚热带湿润土壤中观察到的结果[83] 并不一致, 表明在这一情景下, 高 pH 值可能降低了干旱黄土的铅生物的有效性和毒性。

Atlas 等[28] 通过分析测量的分类、遗传、生理和营养多样性参数数据来确定暴露在原油中、暴露在含铅汽油中及暴露在除草剂 2,4,5-三氯苯氧乙酸(2,4,5-T)中的群落受什么特性的微生物主导, 发现各干扰下的分类多样性均有显著降低, 其中, 影响最大的是含铅汽油。在传统的认知中, 受干扰群落中, 存活的稀有种群具有某些特定属性, 以使得其自身得以继续存活, 因此, 这些种群被认为是高度“专门化”的, 这是传统浓缩培养技术中用于分离具有特定性状的微生物的原理; 但事实并非如此, 数据表明, 受干扰的群落中演替出更高的生理耐受性和代谢通量, 受原油干扰的群落表现出更强的利用烃类和羧酸的能力, 受含铅汽油干扰的群落对碳氢化合物和羧酸的营养适应性增强。

包括杀虫剂、除草剂和杀真菌剂在内的除害剂是具有潜在毒性的物质, 虽然大部分都是针对特定物种而设计的, 但往往会对非目标物种产生影响, 能够改变生境中微生物群落的结构和功能, 可能对其造成直接的毒性伤害或有益影响[84]。有些微生物(尤其是细菌)可能利用除害剂作为促进其本身生长和存活的营养来源, 而其他对除害剂敏感的物种可能会受损或大量死亡[85]。这些生态变化可能会引发一连串的群落演替。除害剂消除或减少某些微生物种群可能会使耐除害剂的微生物在争夺共有资源的过程中得到释放, 从而促进了这些物种的生长[86]; 有些原生动物以细菌为食但却由于除害剂被抑制, 这有助于细菌的存活。除害剂也可能会改变水生生境的物理和化学特性, 抑制某些微生物的同时刺激其他微生物的生长和存活, 间接影响到微生物群落[87]。Muturi 等[84] 对于一系列常用的除害剂如何单独或共同影响水生栖息地微生物群落

进行了研究，结果表明，低浓度的单独和混合除害剂可以减少微生物的多样性，并且诱导微生物群落结构的转变，除草剂对微生物多样性和群落结构没有影响，3种杀虫剂混合或全部5种杀虫剂混合可以降低微生物的多样性，导致微生物群落结构发生重大变化。Atlas等[28]研究发现，受2,4,5-三氯苯氧乙酸干扰的群落表现出对羧酸营养的多样性，另外，受2,4,5-三氯苯氧乙酸干扰的群落遗传多样性也有显著降低。

3 特殊环境中的干扰

在特定的环境中，限制生长和生存的干扰因素会对群落产生重大影响[88-89]。许多微生物群落是有营养限制的，意味着利用现有碳源的能力是微生物受到自然选择的最主要因素[90]。然而，特殊环境中的其他因素，如水的可利用性、氮、温度、pH值或盐度可能更为关键[91-92]，使得这些系统成为研究微生物群落演替过程的有力模型[93]。

3.1 沙漠生态系统

在沙漠生态系统中，微生物群落演替过程中核心决定性因素是水。其已经进化到依赖于空气相对湿度的变化，是研究局部气候变化对微生物群落组成和功能影响的模型[94]。Uritskiy等[95]探究了从区域性的沙漠尺度到单个岩盐尺度上主导微生物群落组成的非生物因素，发现岩盐外围的环境对需要可持续用水的物种不太有利，虽然它们都能够在长期干燥的环境下生存。

3.2 极地生态系统

在极地环境中，微生物群落对环境连通性干扰的响应也非常明显。如2.2.1节所述，Ramoneda等[46]对南极淡水生态系统细菌群落的研究注意到了环境之间的连通性和各栖息地条件在控制微生物群落组成中的作用。

3.3 外太空环境中的空间站

宇宙辐射由微粒辐射和电磁辐射2部分组成，主要发生在太空和地球的高海拔处，且在各种不同的条件下辐射源都会有一定的生态效应[24]。高剂量电离辐射可以直接杀灭微生物，常用于不耐热物品的消毒，如橡胶、塑料、高分子聚合物、食品等。空间站中滋生的微生物受来自宇宙的低剂量电离辐射的影响。越来越多的证据证明，在低剂量电离辐射的贫营养环境中，微生物可以快速增殖，并且物种多样性较高[96]。而在温度、湿度和培养基质等因子相同的情况下，空间站微生物群落的物种多样性大大高于地面生长的微生物群落[97]，如国际空间站环境下Shannon多样性指数能达到4.2～5.7，其平均值高于人体及家居环境的平均水平[98]。Yang等[99-100]认为低剂量电离辐射驱动了空间站贫营养环境下微生物群落的多样性的原理是底物的反馈调节导致了不同的生长延迟，从而发生了时间生态位的分化，使得更多的物种存活下来。

空间站与地面环境的另一个显著不同的环境因子就是重力，即空间站处于微重力的环境中。Rosado等[101]的研究发现，空间微重力可以影响微生物的生长。但Bouma和Pierson[102]通过在地面旋转模拟微重力效应下微生物培养实验表明，微重力不会导致微生物群落物种多样性的形成。因此，空间站中影响微生物群落演替的非生物环境因子主要是辐射。

季铵盐为阳离子表面活性剂，可改变细菌胞浆膜的通透性，使菌体物质外渗，阻碍其代谢而使细菌死亡，是一种有效的杀菌剂[103]。空间站舱室作为密闭环境，外部环境的微生物对群落的组成影响较小，封闭系统内的人类活动和运输的设备材料是微生物群落的主要来源。其中，最主要的来源是宇航员本身，因为他们的身体携带大量微生物，包括细菌、真菌、原生动物和病毒[97]。而空间站中宇航员的因素又无法被消除，微生物的滋生不可避免，因此，需要对空间站中存在的微生物进行清除。国际空间站进行清洁工作，采用吸尘器吸尘，用含0.1%季铵盐（0.039 9%辛基癸基二甲基氯化铵、0.019 95%二辛基二甲基氯化铵、0.019 95%二癸基二甲基氯化铵、0.053 2%烷基二甲基苄基氯化铵（50%^{14}C 40%^{12}C 10%^{16}C））的专用湿巾对舱内表面进行擦拭，但对微生物的清除虽然当下有效但并不能彻底，过一段时间后微生物又会长出来[104]。季铵盐类是发现耐药性最早的杀菌剂[105]，早在1961年，Lowbury[106]就观察到了铜绿假单胞菌（*Pseudomonas aeruginosa*）对季铵盐类化合物有耐受现象，其对革兰氏阴性菌的效能不是最理想，只有高浓度才能杀灭。可Adair等[107]还是发现了铜绿假单胞菌对高浓度季铵盐类化合物有耐受。因此，使用含季铵盐的杀菌剂却不能完全杀灭微生物群落的事件可能会造成进一步的演替，产生“生态意外”（ecological surprise）的现象[108]。

4 结束语

虽然针对各种不同形式微生物群落受到干扰产生的演替都有所研究，但相关的探索仍有很大发展空间，在干扰下观察到的群落多样性模式背后的机制仍有待阐明。

1）对于干扰本身，量化的报道较少，更多的研

究关注于定性的描述。当然, 有一部分原因在于具有频率、强度、范围和周期性等属性的干扰很难定义, 干扰的特征和生态影响差异巨大, 阻碍了理论和预测模型的进步。在 1.1.1 节所述的最近, 在实验室条件下的稀释可以作为量化干扰的典型例子, 有助于以后开展相关研究提供思路。

2) 大部分研究仍然停留在单因素阶段, 只有少数研究关注多个干扰因素共同作用的影响, 而在自然环境中仍然存在着大量情景是多因素干扰共同作用的。

3) 很少有研究明确地关注干扰后微生物群落组成的时间过程; 相反, 大多数研究仅仅关注组成在某一时刻的敏感性, 这一方面源于群落生态学的思考惯性, 另一方面源于微生物领域研究手段的思考惯性, 而在 1.1.1 节所述的实验室条件下的稀释同样给出了构造人工简单群落来进行群落演替基础性研究的思路。但其研究也存在着不足之处, 当干扰按照脉冲式干扰和持续式干扰区分时, 稀释作为脉冲式干扰的典型代表, 不适合直接整合进 Lotka-Volterra 方程中; 应用类似手段对于特异性非生物环境因子, 以及多种因子共同作用的研究还未见报道。

因此, 今后的研究可以更关注于将干扰划分梯度定量, 描述在一段时间内微生物群落演替变化的情景, 并且将实际环境中的多因素干扰抽离出来, 以探究它们共同作用的影响。

致谢　感谢北京航空航天大学计算机学院虚拟现实技术与系统国家重点实验室为本文提供的支持。

参考文献 (References)

[1] WELLBORN G A, SKELLY D K, WERNER E E. Mechanisms creating community structure across a freshwater habitat gradient [J]. Annual Review of Ecology, Evolution, and Systematics, 1996, 27: 337-363.

[2] WERNER E E. Species packing and niche complementarity in three sunfishes[J]. The American Naturalist, 1977, 111(979): 553-578.

[3] TILMAN D. The importance of the mechanisms of interspecific competition[J]. The American Naturalist, 1987, 129(5): 769-774.

[4] GORTER F A, MANHART M, ACKERMANN M. Understanding the evolution of interspecies interactions in microbial communities[J]. Philosophical Transactions of the Royal Society B: Biological Sciences, 2020, 375(1798): 20190256.

[5] KONOPKA A. What is microbial community ecology?[J]. The ISME Journal, 2009, 3(11): 1223-1230.

[6] MOUILLOT D, GRAHAM N A J, VILLÉGER S, et al. A functional approach reveals community responses to disturbances[J]. Trends in Ecology & Evolution, 2013, 28(3): 167-177.

[7] BRIONES A, RASKIN L. Diversity and dynamics of microbial communities in engineered environments and their implications for process stability[J]. Current Opinion in Biotechnology, 2003, 14(3): 270-276.

[8] STEGEN J C, LIN X J, FREDRICKSON J K, et al. Quantifying community assembly processes and identifying features that impose them[J]. The ISME Journal, 2013, 7(11): 2069-2079.

[9] POWELL J R, KARUNARATNE S, CAMPBELL C D, et al. Deterministic processes vary during community assembly for ecologically dissimilar taxa[J]. Nature Communications, 2015, 6: 8444.

[10] ZHOU J Z, NING D L. Stochastic community assembly: Does it matter in microbial ecology?[J]. Microbiology and Molecular Biology Reviews, 2017, 81(4): e00002-e00017.

[11] TILMAN D. Resource competition and community structure[M]. Princeton: Princeton University Press, 1982.

[12] ABREU C I, FRIEDMAN J, ANDERSEN WOLTZ V L, et al. Mortality causes universal changes in microbial community composition[J]. Nature Communications, 2019, 10: 2120.

[13] RYKIEL E J. Towards a definition of ecological disturbance[J]. Australian Journal of Ecology, 1985, 10(3): 361-365.

[14] SANTILLAN E, SESHAN H, CONSTANCIAS F, et al. Frequency of disturbance alters diversity, function, and underlying assembly mechanisms of complex bacterial communities[J]. NPJ Biofilms and Microbiomes, 2019, 5: 8.

[15] MACKEY R L, CURRIE D J. The diversity-disturbance relationship: Is it generally strong and peaked?[J]. Ecology, 2001, 82(12): 3479.

[16] SHADE A, PETER H, ALLISON S D, et al. Fundamentals of microbial community resistance and resilience[J]. Frontiers in Microbiology, 2012, 3: 417.

[17] SHADE A, READ J S, YOUNGBLUT N D, et al. Lake microbial communities are resilient after a whole-ecosystem disturbance[J]. The ISME Journal, 2012, 6(12): 2153-2167.

[18] SOUSA W. The role of disturbance in natural communities[J]. Annual Review of Ecology, Evolution, and Systematics, 1984, 15: 353-391.

[19] MYKRÄ H, TOLKKINEN M, HEINO J. Environmental degradation results in contrasting changes in the assembly processes of stream bacterial and fungal communities[J]. Oikos, 2017, 126(9): 1291-1298.

[20] ALLISON S D, MARTINY J B H. Resistance, resilience, and redundancy in microbial communities[J]. Proceedings of the National Academy of Sciences of the United States of America, 2008, 105(Suppl 1): 11512-11519.

[21] DETHLEFSEN L, RELMAN D A. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation[J]. Proceedings of the National Academy of Sciences of the United States of America, 2011, 108(Suppl 1): 4554-4561.

[22] FERRENBERG S, O'NEILL S P, KNELMAN J E, et al. Changes in assembly processes in soil bacterial communities following a wildfire disturbance[J]. The ISME Journal, 2013, 7(6): 1102-1111.

[23] CONNELL J H. The influence of interspecific competition and other factors on the distribution of the barnacle chthamalusstellatus [J]. Ecology, 1961, 42(4): 710-723.

[24] 尚玉昌. 普通生态学[M]. 3 版. 北京: 北京大学出版社, 2010.

SHANG Y C. General ecology[M]. 3rd ed. Beijing: Peking University Press, 2010(in Chinese).

[25] FOX J W. The intermediate disturbance hypothesis should be abandoned[J]. Trends in Ecology & Evolution, 2013, 28(2): 86-92.

[26] ORROCK J L, WATLING J I. Local community size mediates ecological drift and competition in metacommunities[J]. Proceedings Biological Sciences, 2010, 277(1691): 2185-2191.

[27] ATLAS R M. Diversity of microbial communities[M]//MARSHALL K C. Advances in microbial ecology. Berlin: Springer, 1984: 1-47.

[28] ATLAS R M, HOROWITZ A, KRICHEVSKY M, et al. Response of microbial populations to environmental disturbance[J]. Microbial Ecology, 1991, 22(3): 249-256.

[29] MÜLLER A K, WESTERGAARD K, CHRISTENSEN S, et al. The diversity and function of soil microbial communities exposed to different disturbances[J]. Microbial Ecology, 2002, 44(1): 49-58.

[30] BENDER E A, CASE T J, GILPIN M E. Perturbation experiments in community ecology: Theory and practice[J]. Ecology, 1984, 65(1): 1-13.

[31] SCHEUERL T, KAITALA V. The effect of dilution on eco-evolutionary dynamics of experimental microbial communities[J]. Ecology and Evolution, 2021, 11(19): 13430-13444.

[32] SOMMER U. Phytoplankton competition along a gradient of dilution rates[J]. Oecologia, 1986, 68(4): 503-506.

[33] SPIJKERMAN E, COESEL P F M. Competition for phosphorus among planktonic desmid species in continuous-flow culture[J]. Journal of Phycology, 1996, 32(6): 939-948.

[34] ABREU C I, ANDERSEN WOLTZ V L, FRIEDMAN J, et al. Microbial communities display alternative stable states in a fluctuating environment[J]. PLoS Computational Biology, 2020, 16(5): e1007934.

[35] CHESSON P, HUNTLY N. The roles of harsh and fluctuating conditions in the dynamics of ecological communities[J]. The American Naturalist, 1997, 150(5): 519-553.

[36] KAITALA V, HILTUNEN T, BECKS L, et al. Co-evolution as an important component explaining microbial predator-prey interaction[J]. Journal of Theoretical Biology, 2020, 486: 110095.

[37] STEVENSON B S, SCHMIDT T M. Life history implications of rRNA gene copy number in Escherichia coli[J]. Applied and Environmental Microbiology, 2004, 70(11): 6670-6677.

[38] KLAPPENBACH J A, DUNBAR J M, SCHMIDT T M. rRNA operon copy number reflects ecological strategies of bacteria[J]. Applied and Environmental Microbiology, 2000, 66(4): 1328-1333.

[39] WANG M X, LIU X N, NIE Y, et al. Selfishness driving reductive evolution shapes interdependent patterns in spatially structured microbial communities[J]. The ISME Journal, 2021, 15(5): 1387-1401.

[40] ESCALANTE A E, REBOLLEDA-GÓMEZ M, BENÍTEZ M, et al. Ecological perspectives on synthetic biology: Insights from microbial population biology[J]. Frontiers in Microbiology, 2015, 6: 143.

[41] HARCOMBE W. Novel cooperation experimentally evolved between species[J]. Evolution, 2010, 64(7): 2166-2172.

[42] RAYNAUD X, NUNAN N. Spatial ecology of bacteria at the microscale in soil[J]. PLoS One, 2014, 9(1): e87217.

[43] AMOR D R, DAL BELLO M. Bottom-up approaches to synthetic cooperation in microbial communities[J]. Life, 2019, 9(1): 22.

[44] LI S S, LIU H, LIU G H, et al. Simulated artificial interventions maintain species diversity of spatially structured microbial communities in closed ecosystem[J]. Acta Astronautica, 2022, 201: 39-47.

[45] WOODWARD G, PERKINS D M, BROWN L E. Climate change and freshwater ecosystems: Impacts across multiple levels of organization[J]. Philosophical Transactions of the Royal Society B: Biological Sciences, 2010, 365(1549): 2093-2106.

[46] RAMONEDA J, HAWES I, PASCUAL-GARCÍA A, et al. Importance of environmental factors over habitat connectivity in shaping bacterial communities in microbial mats and bacterioplankton in an Antarctic freshwater system[J]. FEMS Microbiology Ecology, 2021, 97(4): fiab044.

[47] CHOWN S L, CLARKE A, FRASER C I, et al. The changing form of Antarctic biodiversity[J]. Nature, 2015, 522: 431-438.

[48] HAWES I, SMITH R, HOWARD-WILLIAMS C, et al. Environmental conditions during freezing, and response of microbial mats in ponds of the McMurdo Ice Shelf, Antarctica[J]. Antarctic Science, 1999, 11(2): 198-208.

[49] JUVIGNY-KHENAFOU N P D, PIGGOTT J J, ATKINSON D, et al. Fine sediment and flow velocity impact bacterial community and functional profile more than nutrient enrichment[J]. Ecological Applications, 2021, 31(1): e02212.

[50] HARTWIG M, BORCHARDT D. Alteration of key hyporheic functions through biological and physical clogging along a nutrient and fine-sediment gradient[J]. Ecohydrology, 2015, 8(5): 961-975.

[51] BESEMER K. Biodiversity, community structure and function of biofilms in stream ecosystems[J]. Research in Microbiology, 2015, 166(10): 774-781.

[52] STEVENS C L, HURD C L. Boundary-layers around bladed aquatic macrophytes[J]. Hydrobiologia, 1997, 346(1): 119-128.

[53] SALIS R K, BRUDER A, PIGGOTT J J, et al. High-throughput amplicon sequencing and stream benthic bacteria: Identifying the best taxonomic level for multiple-stressor research[J]. Scientific Reports, 2017, 7: 44657.

[54] NUY J K, LANGE A, BEERMANN A J, et al. Responses of stream microbes to multiple anthropogenic stressors in a mesocosm study[J]. The Science of the Total Environment, 2018, 633: 1287-1301.

[55] WAGG C, DUDENHÖFFER J H, WIDMER F, et al. Linking diversity, synchrony and stability in soil microbial communities[J]. Functional Ecology, 2018, 32(5): 1280-1292.

[56] HABIG J, LABUSCHAGNE J, MARAIS M, et al. The effect of a medic-wheat rotational system and contrasting degrees of soil disturbance on nematode functional groups and soil microbial communities[J]. Agriculture, Ecosystems & Environment, 2018, 268: 103-114.

[57] PELLKOFER S, VAN DER HEIJDEN M G A, SCHMID B, et al. Soil communities promote temporal stability and species asynchrony in experimental grassland communities[J]. PLoS One, 2016, 11(2): e0148015.

[58] WAGG C, BENDER S F, WIDMER F, et al. Soil biodiversity and

soil community composition determine ecosystem multifunctionality[J]. Proceedings of the National Academy of Science of the United States of America, 2014, 111(14): 5266-5270.

[59] LAUBER C L, RAMIREZ K S, AANDERUD Z, et al. Temporal variability in soil microbial communities across land-use types[J]. The ISME Journal, 2013, 7(8): 1641-1650.

[60] BRADFORD M A, WOOD S A, BARDGETT R D, et al. Discontinuity in the responses of ecosystem processes and multifunctionality to altered soil community composition[J]. Proceedings of the National Academy of Sciences of the United States of America, 2014, 111(40): 14478-14483.

[61] GRIFFITHS B S, KUAN H L, RITZ K, et al. The relationship between microbial community structure and functional stability, tested experimentally in an upland pasture soil[J]. Microbial Ecology, 2004, 47(1): 104-113.

[62] LOREAU M. From populations to ecosystems[M]. Princeton: Princeton University Press, 2010.

[63] LOREAU M, DE MAZANCOURT C. Species synchrony and its drivers: Neutral and nonneutral community dynamics in fluctuating environments[J]. The American Naturalist, 2008, 172(2): 48-66.

[64] GONZALEZ A, LOREAU M. The causes and consequences of compensatory dynamics in ecological communities[J]. Annual Review of Ecology, Evolution, and Systematics, 2009, 40: 393-414.

[65] SILVA A P. Bacterial diversity under different tillage and crop rotation systems in an oxisol of southern Brazil[J]. The Open Agriculture Journal, 2013, 7(1): 40-47.

[66] FINLAY B J. Global dispersal of free-living microbial eukaryote species[J]. Science, 2002, 296(5570): 1061-1063.

[67] LOGUE J B, MOUQUET N, PETER H, et al. Empirical approaches to metacommunities: A review and comparison with theory[J]. Trends in Ecology & Evolution, 2011, 26(9): 482-491.

[68] BAHO D L, PETER H, TRANVIK L J. Resistance and resilience of microbial communities: Temporal and spatial insurance against perturbations[J]. Environmental Microbiology, 2012, 14(9): 2283-2292.

[69] URBAN M C, DE MEESTER L. Community monopolization: Local adaptation enhances priority effects in an evolving metacommunity[J]. Proceedings Biological Sciences, 2009, 276(1676): 4129-4138.

[70] SHULMAN M J, OGDEN J C, EBERSOLE J P, et al. Priority effects in the recruitment of juvenile coral reef fishes[J]. Ecology, 1983, 64(6): 1508-1513.

[71] SHADE A, READ J S, WELKIE D G, et al. Resistance, resilience and recovery: Aquatic bacterial dynamics after water column disturbance[J]. Environmental Microbiology, 2011, 13(10): 2752-2767.

[72] WANG M, CHEN S B, CHEN L, et al. Responses of soil microbial communities and their network interactions to saline-alkaline stress in Cd-contaminated soils[J]. Environmental Pollution, 2019, 252: 1609-1621.

[73] ZENG X Y, LI S W, LENG Y, et al. Structural and functional responses of bacterial and fungal communities to multiple heavy metal exposure in arid loess[J]. The Science of the Total Environment, 2020, 723: 138081.

[74] ROUSK J, BÅÅTH E, BROOKES P C, et al. Soil bacterial and fungal communities across a pH gradient in an arable soil[J]. The ISME Journal, 2010, 4(10): 1340-1351.

[75] DE VRIES F T, MANNING P, TALLOWIN J R B, et al. Abiotic drivers and plant traits explain landscape-scale patterns in soil microbial communities[J]. Ecology Letters, 2012, 15(11): 1230-1239.

[76] WEI W, YANG M, LIU Y X, et al. Fertilizer N application rate impacts plant-soil feedback in a Sanqi production system[J]. The Science of the Total Environment, 2018, 633: 796-807.

[77] MEENA R, DATTA S P, GOLUI D, et al. Long-term impact of sewage irrigation on soil properties and assessing risk in relation to transfer of metals to human food chain[J]. Environmental Science and Pollution Research, 2016, 23(14): 14269-14283.

[78] FENG G, XIE T, WANG X, et al. Metagenomic analysis of microbial community and function involved in Cd-contaminated soil[J]. BMC Microbiology, 2018, 18(1): 11.

[79] CUI H, LIU L L, DAI J R, et al. Bacterial community shaped by heavy metals and contributing to health risks in cornfields[J]. Ecotoxicology and Environmental Safety, 2018, 166: 259-269.

[80] ZHANG H, WAN Z W, DING M J, et al. Inherent bacterial community response to multiple heavy metals in sediment from river-lake systems in the Poyang Lake, China[J]. Ecotoxicology and Environmental Safety, 2018, 165: 314-324.

[81] KUPPUSAMY S, THAVAMANI P, MEGHARAJ M, et al. Pyrosequencing analysis of bacterial diversity in soils contaminated long-term with PAHs and heavy metals: Implications to bioremediation[J]. Journal of Hazardous Materials, 2016, 317: 169-179.

[82] SCHNEIDER G K. Dürrenmatts Besuch der alten dame und hebbels Die nibelungen. urbilder und jugendlektüre[J]. The German Quarterly, 2017, 90(1): 36-54.

[83] PAN X M, ZHANG S R, ZHONG Q M, et al. Effects of soil chemical properties and fractions of Pb, Cd, and Zn on bacterial and fungal communities[J]. The Science of the Total Environment, 2020, 715: 136904.

[84] MUTURI E J, DONTHU R K, FIELDS C J, et al. Effect of pesticides on microbial communities in container aquatic habitats[J]. Scientific Reports, 2017, 7: 44565.

[85] RUSSELL R J, SCOTT C, JACKSON C J, et al. The evolution of new enzyme function: lessons from xenobiotic metabolizing bacteria versus insecticide-resistant insects[J]. Evolutionary Applications, 2011, 4(2): 225-248.

[86] WIDENFALK A, BERTILSSON S, SUNDH I, et al. Effects of pesticides on community composition and activity of sediment microbes: Responses at various levels of microbial community organization[J]. Environmental Pollution, 2008, 152(3): 576-584.

[87] SCHÄFER R B, BUNDSCHUH M, ROUCH D A, et al. Effects of pesticide toxicity, salinity and other environmental variables on selected ecosystem functions in streams and the relevance for ecosystem services[J]. The Science of the Total Environment, 2012, 415: 69-78.

[88] GOLDFORD J E, LU N X, BAJIĆ D, et al. Emergent simplicity in microbial community assembly[J]. Science, 2018, 361(6401): 469-474.

[89] ZHAO P Y, BAO J B, WANG X, et al. Deterministic processes dominate soil microbial community assembly in subalpine coniferous forests on the Loess Plateau[J]. PeerJ, 2019, 7: e6746.

[90] MELLO B L, ALESSI A M, MCQUEEN-MASON S, et al. Nutrient availability shapes the microbial community structure in sugarcane bagasse compost-derived consortia[J]. Scientific Reports, 2016, 6: 38781.

[91] ARMSTRONG A, VALVERDE A, RAMOND J B, et al. Temporal dynamics of hot desert microbial communities reveal structural and functional responses to water input[J]. Scientific Reports, 2016, 6: 34434.

[92] MERINO N, ARONSON H S, BOJANOVA D P, et al. Living at the extremes: Extremophiles and the limits of life in a planetary context[J]. Frontiers in Microbiology, 2019, 10: 780.

[93] SCHMID A K, ALLERS T, DIRUGGIERO J. SnapShot: Microbial extremophiles[J]. Cell, 2020, 180(4): 818.

[94] URITSKIY G, GETSIN S, MUNN A, et al. Halophilic microbial community compositional shift after a rare rainfall in the Atacama Desert[J]. The ISME Journal, 2019, 13(11): 2737-2749.

[95] URITSKIY G, MUNN A, DAILEY M, et al. Environmental factors driving spatial heterogeneity in desert halophile microbial communities[J]. Frontiers in Microbiology, 2020, 11: 578669.

[96] YANG X B, XU X M, HU D W. Succession mechanism of microbial community with high species diversity in nutrient-deficient environments with low-dose ionizing radiation[J]. Ecological Modelling, 2020, 435: 109270.

[97] NOVIKOVA N, DE BOEVER P, PODDUBKO S, et al. Survey of environmental biocontamination on board the International Space Station[J]. Research in Microbiology, 2006, 157(1): 5-12.

[98] LANG J M, COIL D A, NECHES R Y, et al. A microbial survey of the International Space Station (ISS)[J]. PeerJ, 2017, 5: e4029.

[99] YANG X B, LI S S, SONG G Y, et al. Microbial diversity formed and maintained through substrate feedback regulation and delayed responses induced by low-dose ionizing radiation[J]. Acta Astronautica, 2021, 188: 239-251.

[100] YANG X B, SONG G Y, LIU H, et al. Microbial diversity formation and maintenance due to temporal niche differentiation caused by low-dose ionizing radiation in oligotrophic environments[J]. Life Sciences in Space Research, 2021, 31: 92-100.

[101] ROSADO H, DOYLE M, HINDS J, et al. Low-shear modelled microgravity alters expression of virulence determinants of Staphylococcus aureus[J]. Acta Astronautica, 2010, 66(3-4): 408-413.

[102] BOUMA J E, PIERSON D L. Combined effects of simulated microgravity and multi-strain interactions on population dynamics of a constructed microbial community[R]. Warrendale: SAE International, 1998.

[103] 李亚红, 陈冲, 武杰, 等. 紫外分光光度法测定海水中季铵盐的含量[J]. 理化检验-化学分册, 2010, 46(2): 136-137.
LI Y H, CHEN C, WU J, et al. UV-spectrophotometric determination of quaternary ammonium salt in seawater[J]. Physical Testing and Chemical Analysis (Part B: Chemical Analysis), 2010, 46(2): 136-137(in Chinese).

[104] 杨宏, 侯永青, 张兰涛. 微生物控制: 我国空间站面临的新挑战[J]. 载人航天, 2013, 19(2): 38-46.
YANG H, HOU Y Q, ZHANG L T. Microbe control: A new challenge to faced by Chinese space station[J]. Manned Spaceflight, 2013, 19(2): 38-46(in Chinese).

[105] 顾春英, 薛广波. 细菌对消毒剂的抗性研究进展[J]. 中华医院感染学杂志, 1997, 7(4): 62-66.
GU C Y, XUE G B. Research progress of bacterial resistance to disinfectants[J]. Chinese Journal of Nosocomiology, 1997, 7(4): 62-66(in Chinese).

[106] LOWBURY E J L. Skin disinfection[J]. Journal of Clinical Pathology, 1961, 14(1): 85.

[107] ADAIR F W, GEFTIC S G, GELZER J. Resistance of pseudomonas to quaternary ammonium compounds. I. Growth in benzalkonium chloride solution[J]. Applied Microbiology, 1969, 18(3): 299-302.

[108] LI S, LIU X, CHEN L, et al. Cure or curse? Simulation indicates that microbes proliferate under disinfection measures in the space station[J/OL]. BiorXiv, 2024(2024-05-29)[2024-06-23]. https://doi.org/10.1101/2024.04.16.589799.

Research progress on influence of disturbance of abiotic environmental factors on microbial community succession

HU Dawei[1]，LI Shuaishuai[1, 2]，LIU Guanghui[3, 4, *]

(1. School of Biological Science and Medical Engineering，Beihang University，Beijing 100191，China；
2. Department of Biology，McGill University，Montreal H3A1B1，Canada；
3. School of Aeronautic Science and Engineering，Beihang University，Beijing 100191，China；
4. Technology and Engineering Center for Space Utilization，Chinese Academy of Sciences，Beijing 100094，China)

Abstract: It is one of the main goals of microbial ecology to understand factors that drive the succession of communities. Abiotic environmental factors affect the growth of populations and the interactions between them. Microbial communities are very sensitive and can easily respond to changes in abiotic environmental factors as the disturbance. Studying the influence of different types of disturbance will enhance the understanding of the relationship among diversity, structure, and function of the microbial community. With the action subject as the classification criteria, the influence of various abiotic environmental factors on the succession of microbial communities was sorted out, including abiotic environmental factors in non-specific, specific, and special (e.g. space) environments. Although studies have been conducted on the succession of microbial communities in various forms due to disturbance, there is still a long way to go for related exploration. The quantification, multi-factor effect of disturbance, and the time course of succession have yet to be studied enough.

Keywords: abiotic environmental factors; disturbance; community succession; microbe; special environment

Received：2022-08-22；**Accepted**：2022-09-09；**Published Online**：2022-11-02 16：30
URL：link.cnki.net/urlid/11.2625.V.20221102.1028.001
Foundation items：National Key Research and Development Program of China (2021YFA0716100); Fund for International Cooperation and Exchange of the Natural Science Foundation of China (32261133528)

* **Corresponding author**. E-mail：liugh1991@126.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0728

多源知识融合的方面级情感分析模型

韩虎*，郝俊，张千锟，赵启涛
(兰州交通大学 电子与信息工程学院，兰州 730070)

摘 要：方面级情感分析(ABSA)是一项细粒度情感分析任务，其目的是针对评论语句中出现的特定方面给出对应的情感极性。现有的基于深度学习的ABSA方法大多侧重于评论语句语义和句法的挖掘，往往忽略了评论语句可能涉及的概念知识和情感程度信息。针对此问题，提出一种融合多源知识的神经网络模型，通过句法依赖揭示句子的结构框架、词共现捕捉单词之间的语义联系、情感网络和概念图谱的嵌入为模型提供情感和背景知识，共同实现评论语句上下文与评价方面的增强表示，并通过双交互注意力模式实现评论语句上下文与评价方面的协调优化。通过在4个公开数据集上的实验验证，该模型在ABSA任务中，准确率分别达到了75.00%、77.90%、81.55%、90.10%，与基准模型相比均有所提高。研究成果不仅验证了多源知识融合在ABSA任务中的有效性，也为未来的研究提供了新的思路和方法。

关 键 词：方面级情感分析；图卷积网络；多源融合；知识图谱；交互注意力机制
中图分类号：TP391
文献标志码：A **文章编号**：1001-5965(2024)09-2688-08

方面级情感分析(aspect-based sentiment analysis, ABSA)指针对句子中不同的方面给出对应的情感标签。区别于给定整个文本情感极性，方面级情感分析是一项粒度较细的任务[1-2]。例如，“This tool is simple to use, but it costs too much！”一句提到了2个方面：“use”和“cost”，方面“use”情感是积极的，而“cost”是消极的。

近年来，深度学习在情感分析领域取得了令人瞩目的成功，其基本思想可以总结为基于语义角度和语法角度的建模。基于语义角度的模型利用卷积神经网络(convolutional neural network, CNN)、长短期记忆(long short-term memory, LSTM)网络等提取方面和上下文的特征，再利用注意力机制捕捉特定方面的上下文表征[3-4]。然而，句子中存在多个方面的情况，基于语义的模型往往容易混淆方面与对应情感之间的联系。而句法信息蕴含着更深层次的文本信息，能够反映词与词的依赖关系，这种依赖关系是联系方面与情感的有效媒介。

因此，基于语法角度的模型利用图卷积网络(graph convolutional network, GCN)处理句法依赖树这类数据[5-7]。此外，语料库层面的词共现信息可以反映语言学中的固定搭配，加强这类语义特征能够更好地把握全局词汇信息[8]。上述模型虽能很好地集成语言知识，但由于评论文本存在自身较短的特点，容易在实际分类过程中造成歧义。因此，尝试利用常识和情感等先验知识为模型提供监督信号，提高模型语义解析能力[9-11]。

受上述工作的启发，本文提出一种多源知识融合的方面级情感分析(multi-source knowledge fusion for aspect-based sentiment analysis, MSKFSA)模型，从以下3个方面增强了模型的分类能力：

1) 提出一种同时融入语言知识、情感知识、概念知识的交互注意力网络MSKFSA，将上下文语义特征与多源化知识进行融合表示。

收稿日期：2022-08-17；**录用日期**：2023-01-14；**网络出版时间**：2023-03-23 09:01
网络出版地址：link.cnki.net/urlid/11.2625.V.20230322.0955.001
基金项目：国家自然科学基金(62166024)
*通信作者. E-mail：hanhu_lzjtu@mail.lzjtu.cn

引用格式：韩虎，郝俊，张千锟，等. 多源知识融合的方面级情感分析模型[J]. 北京航空航天大学学报，2024，50(9)：2688-2695.
HAN H，HAO J，ZHANG Q K，et al. Multi-source knowledge fusion model for aspect-based sentiment analysis[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2688-2695 (in Chinese).

2) 通过句法依赖、词共现、情感网络和概念图谱的嵌入对评论语句上下文与评价方面增强表示。

3) 使用双交互注意力机制建模方面和上下文, 完成方面与上下文之间的协调优化, 提高情感分类的准确率。

1　相关工作

方面级情感分析是一种细粒度的情感分析任务, 能够识别句子给定方面的情感倾向, 经典模型主要使用神经网络从上下文中提取特定于方面的特征表示。Tang 等[4] 提出目标依赖的长短期记忆(target-denpendent long short-term mermery, TD-LSTM)网络, 通过简单的拼接完成目标信息的增强。Wang 等[12] 对 TD-LSTM 进行了改进, 引入注意力机制挖掘方面和观点词之间的潜在相关性。Ma 等[13] 考虑到句子上下文与方面词之间互有影响, 借助 LSTM 网络分别编码上下文和方面, 通过交互注意力机制达成上下文与方面的增强表示。

对于句子中具有不同极性的多个方面的情况, 这类基于语义的模型往往混淆了方面与对应情感之间的联系。句法依存关系通过分析给定的语法体系获取句法结构, 建立依存关系树识别词语之间的依存关系。Zhang 等[5] 利用依存关系树表示句子(见图 1), 提出利用 GCN 挖掘依存关系树中的句法信息。Wang 等[6] 提出了一种面向方面的依赖树结构, 并利用关系图注意力网络对树结构进行编码。此外, 语义的关联也同样重要。Zhang 和 Qian[8] 研究了语言学中词语共现的特点, 设计了层次句法图和词汇图(见图 2), 更好地利用了语料库级别的单词共现信息及不同类型的句法依存关系。图 2 中: S_i 表示依存关系类型, d_i 表示词可在语料库中的共现频率。Pang 等[14] 提出了一种多通道 GCN, 用于有效融合句法和语义信息。Li 等[15] 使用了由 SemGCN 和 SynGCN 模块组成的双重 GCN 模型, 以利用句法结构和语义关联的互补性。

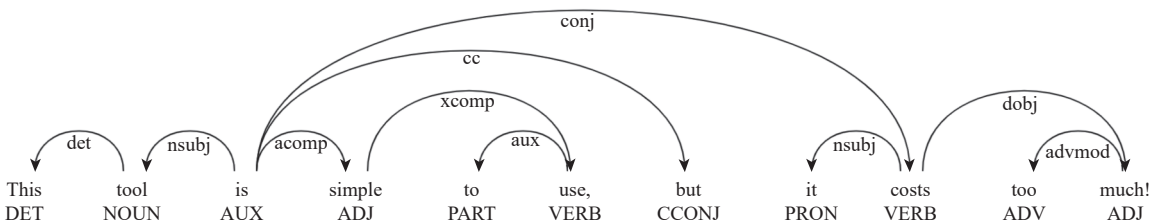
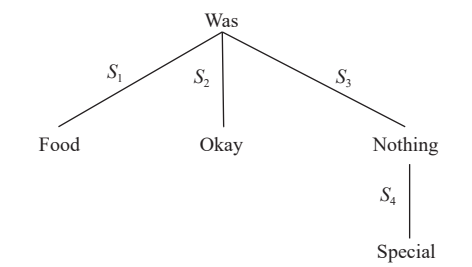
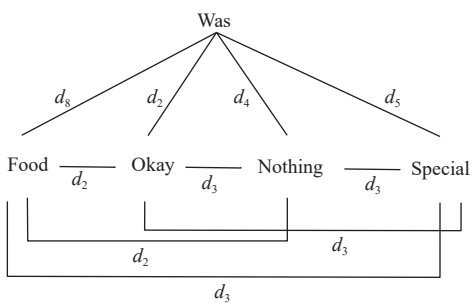


图 1　依存句法树[5]

Fig. 1　Dependency syntax tree[5]



(a) 层次句法图



(b) 层次词汇图

图 2　层次图[8]

Fig. 2　Hierarchical graph[8]

同时, 评论语句在内容表达上非常灵活, 仅从句子本身出发会造成偏差, 外部知识在许多自然语言处理 (natural language processing, NLP) 任务中的重要作用已经得到了验证。Zhou 等[10] 提出将句法信息和外部知识联合编码, 知识通过单个节点进行采样。Zhong 等[16] 提出了一种知识图增强网络, 并设计分层融合模块实现知识间的互补性。Chen 等[17] 通过引入知识图谱来解决不同语境下的“一词多义”问题。Bian 等[18] 使用多头注意力机制来表示目标和上下文, 并从外部知识库中提取概念知识融合来增强上下文和目标的语义表示。

在情感分析任务中, 外部情感常识知识通常被用作增强情感特征表征的来源。例如, 利用 SenticNet 情感词典获得额外的情感分析资源[19-20], 作为一种多功能的情感知识库, SenticNet 在增强情感表示任务上表现出色。Ma 等[9] 利用 SenticNet 增强 LSTM 模型, 提取方面和上下文的情感特征。Xing 等[21] 的研究结果表明, SenticNet 的表现优于其他情感词典。Liang 等[22] 利用 SenticNet 作为常识知识库, 对图进行润色, 进而增强图模型提取的情

感表示。

受上述工作的启发，语言知识、外部知识、句子本身不应只是简单的结合，而是将这些知识与上下文进行交叉融合，以获得更丰富的特征表示，指导模型做出正确的情感极性判别。本文提出了一种融合多种知识的方面级情感分析模型MSKFSA，有效融合方面的上下文信息和依赖信息，并在实验中验证了模型的效果。

2 模 型

本文提出的MSKFSA模型如图3所示，由如下4个部分组成。

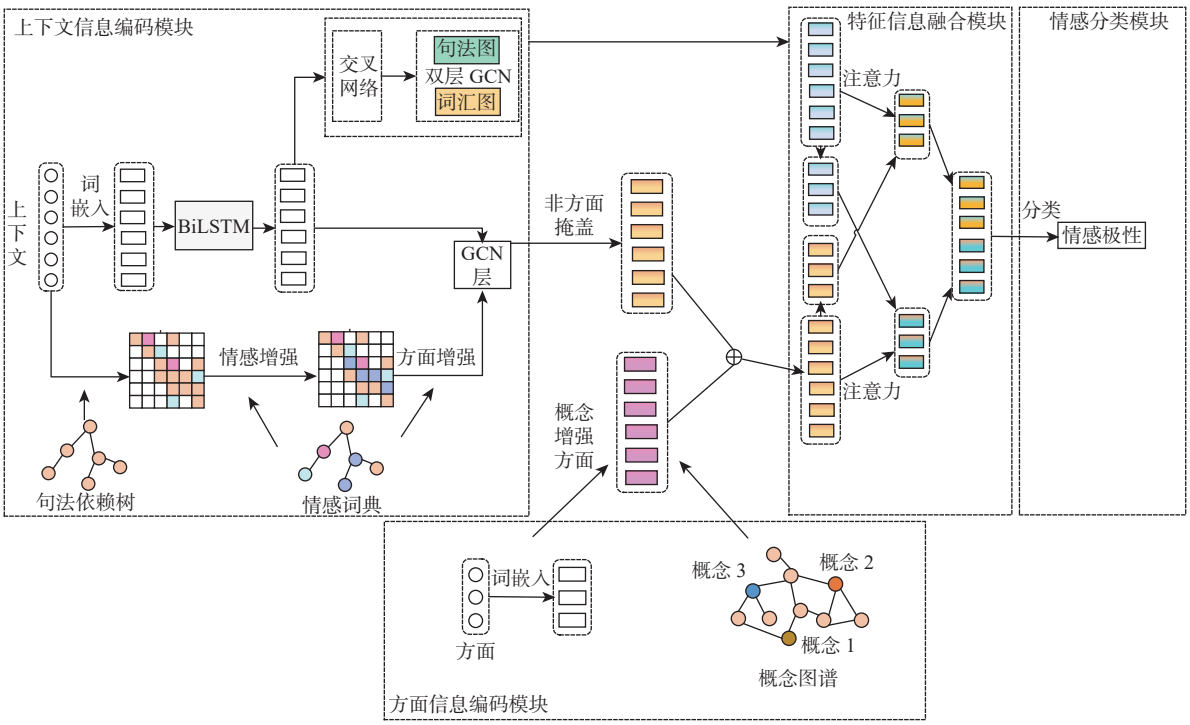


图3 MSKFSA模型网络整体框架

Fig. 3 Overall framework of MSKFSA model

2.1 上下文信息编码模块

给定单词序列$\boldsymbol{s}=[\boldsymbol{w}_1,\boldsymbol{w}_2,\cdots,\boldsymbol{w}_{a_1},\cdots,\boldsymbol{w}_{a_m},\cdots,\boldsymbol{w}_n]$，其中，$\boldsymbol{a}=[\boldsymbol{w}_{a_1},\cdots,\boldsymbol{w}_{a_m}]$为包含$m$个词的方面(方面词或方面短语)。使用GLoVe[23]将文本单词嵌入d_w维的特征空间中，得到对应的连续词向量表示$\boldsymbol{e}=[\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_{a_1},\cdots,\boldsymbol{e}_{a_m},\cdots,\boldsymbol{e}_n]$，$\boldsymbol{e}_i\in\mathbf{R}^{d_w}$，将词向量输入双向长短期记忆(bidirectional long short-term memory, BiLSTM)网络中，得到上下文信息的句子表示$\boldsymbol{h}^{\mathrm{c}}=[\boldsymbol{h}_1^{\mathrm{c}},\boldsymbol{h}_2^{\mathrm{c}},\cdots,\boldsymbol{h}_n^{\mathrm{c}}]$。

将得到的向量与词汇图、句法图作为输入，借助局部层次词汇图和层次句法图[8]获得更优表示，基本思想是使2个图与经过BiLSTM处理的上下文进一步交互。局部层次句法图和词汇图含有层次结构，因此，利用可以聚合关系的双层GCN进行操作。

1) 将相同关系的节点聚合到虚拟节点，在GCN中使用相同的归一化隐藏特征总和作为聚合函数来获取虚拟节点嵌入，每种关系类型r对应的表示为$\tilde{\boldsymbol{h}}_i^{l,r}$，$l$为层号，$i$为聚合的目标节点。

2) 聚合所有虚拟节点及其关系，使用平均聚合函数，针对不同关系类型(虚拟节点)更新目标节点i的表示形式：

$$\boldsymbol{h}_i^l=\mathrm{ReLU}\left(\boldsymbol{W}^l\cdot\left(\oplus_r\tilde{\boldsymbol{h}}_i^{l,r}\right)\right)\tag{1}$$

式中："$\oplus_r$"表示不同关系类型对应表示的连接；$\boldsymbol{W}^l$为第l层中的权重矩阵；$\boldsymbol{h}_i^l$为聚合操作后的输出。

2.2 方面信息编码模块

2.2.1 情感增强部分

在句法依赖树上获得邻接矩阵$\boldsymbol{D}\in\mathbf{R}^{n\times n}$。其中，邻接矩阵$\boldsymbol{D}$的构建规则为：若节点$i$、$j$间存在依存关系，则$D_{ij}=1$，否则$D_{ij}=0$。另外，本文在构建邻接矩阵时，均考虑单词自身的依存关系，即$D_{ii}=D_{jj}=1$。

在此基础上，为充分利用上下文词与方面词之间的情感信息，借助SenticNet情感词典，其中，$\mathrm{SenticNet}(\boldsymbol{w}_i)\in[-1,1]$表示词$\boldsymbol{w}_i$在SenticNet中的情感分数。当$\mathrm{SenticNet}(\boldsymbol{w}_i)=0$，表示词$\boldsymbol{w}_i$是中性词或在SenticNet中不存在。在句子的依存树中提取相邻节点的情感依存关系强度，表示为S_{ij}：

$$S_{ij}=\text{SenticNet}(\boldsymbol{w}_i)+\text{SenticNet}(\boldsymbol{w}_j) \tag{2}$$

此外, 考虑到$\boldsymbol{w}_i$、$\boldsymbol{w}_j$是否为方面词的情况, 构建方面增强矩阵$\boldsymbol{Z}$, 当其中含有方面词时, $Z_{ij}=1$, 否则$Z_{ij}=0$, 得到情感增强后的矩阵$\boldsymbol{A}$:

$$A_{ij}=D_{ij}(S_{ij}+Z_{ij}+1) \tag{3}$$

将处理后的情感邻接矩阵$\boldsymbol{A}$作为输入, 根据其邻域的隐藏表示来更新第l个 GCN 层中的每个节点:

$$\boldsymbol{h}_i^l=\text{ReLU}\left(\tilde{\boldsymbol{A}}_i\boldsymbol{g}_i^{l-1}\boldsymbol{W}^l+\boldsymbol{b}^l\right) \tag{4}$$

$$\boldsymbol{g}_i^{l-1}=F\left(\boldsymbol{h}_i^{l-1}\right) \tag{5}$$

式中: $\boldsymbol{g}_i^{l-1}$为节点i在 GCN 中第l层的输入隐藏状态; $\tilde{\boldsymbol{A}}$为标准化对称情感邻接矩阵, $\tilde{\boldsymbol{A}}_i=\boldsymbol{A}_i/(\boldsymbol{E}_i+\boldsymbol{O})$, $\boldsymbol{O}$为元素全为 1 的矩阵, $\boldsymbol{E}_i=\left[\sum_{j=1}^{n}A_{ij}\right]$; $\boldsymbol{W}^l$为线性变换权重矩阵; $\boldsymbol{b}^l$为偏置项; $F()$为位置感知转换函数, 能够加强靠近该方面的上下文的重要性, 减少过程中产生的噪声[5]。

l层 GCN 的输出结果为: $\boldsymbol{H}^l=\left[\boldsymbol{h}_1^l,\boldsymbol{h}_2^l,\cdots,\boldsymbol{h}_{a_1}^l,\cdots,\boldsymbol{h}_{a_m}^l,\cdots,\boldsymbol{h}_n^l\right]$。

为突出方面词的重要特征, 利用特定于方面的掩蔽, 来掩蔽由最终 GCN 层学习的输出向量的非方面词, 并保持方面词表示不变:

$$\boldsymbol{h}_t^l=\boldsymbol{0}\qquad 1\leqslant t<a_1,a_m<t\leqslant n \tag{6}$$

式中: $\boldsymbol{h}_t^l$为 GCN 层最终输出学习到的第t个单词的表示; a_1为方面词的开始索引; a_m为方面词的结束索引。

因此, 得到情感方面表示:

$$\boldsymbol{S}=\left[\boldsymbol{0},\cdots,\boldsymbol{h}_{a_1},\cdots,\boldsymbol{h}_{a_m},\cdots,\boldsymbol{0}\right] \tag{7}$$

2.2.2　概念增强部分

使用 Concept Graph 知识图谱[24]对方面进行概念化, 获取方面相关概念集合$C=\{\boldsymbol{c}_1,\boldsymbol{c}_2,\cdots,\boldsymbol{c}_m\}$, m为所取概念的个数, $\boldsymbol{c}_i$为概念集中的第i个概念向量。另外, 对概念进行自注意力计算, 筛选出恰当的概念, 注意力计算公式如下:

$$\boldsymbol{\delta}_i=\text{softmax}\left(\boldsymbol{v}^{\mathrm{T}}\tanh(\boldsymbol{W}\boldsymbol{c}_i)+\boldsymbol{b}\right) \tag{8}$$

式中: $\boldsymbol{\delta}_i$为概念集中的第i个概念向量的注意力权重; $\boldsymbol{W}\in\mathbf{R}^{d_a\times d}$为权重矩阵, d为维度, d_a为超参数; $\boldsymbol{v}\in\mathbf{R}^{d_a}$为权重向量; $\boldsymbol{b}$为偏置。注意力机制根据重要程度给予概念权重。

对每个概念向量进行加权计算, 获取最终的概念表示:

$$\boldsymbol{p}=\sum_{i=1}^{m}\delta_i\boldsymbol{c}_i \tag{9}$$

式中: m为概念个数。

组合方面向量$\boldsymbol{a}$与其对应的概念向量$\boldsymbol{p}$, 获取概念增强方面部分$\boldsymbol{K}=(\boldsymbol{k}_1,\boldsymbol{k}_2,\cdots,\boldsymbol{k}_m)$。将概念增强方面$\boldsymbol{K}$与情感增强方面$\boldsymbol{S}$进行拼接, 得到最终的方面表示$\tilde{\boldsymbol{a}}$。

2.3　特征信息融合模块

使用交互注意力机制对 2 部分表示进行交互建模。首先, 对隐藏状态向量取平均, 得到上下文的初始表示:

$$\boldsymbol{c}_{\text{avg}}=\sum_{i=1}^{n}\boldsymbol{h}_i^l/n \tag{10}$$

$$\boldsymbol{a}_{\text{avg}}=\sum_{i=1}^{m}\tilde{\boldsymbol{a}}_i/m \tag{11}$$

1) 方面-上下文注意力计算。对于上下文表示, 利用γ函数计算上下文的隐藏状态向量$\boldsymbol{h}_i^l\in\boldsymbol{H}^l$与方面向量$\boldsymbol{a}_{\text{avg}}$的语义关联, 为每个上下文词$\boldsymbol{h}_i^l$分配相关的注意力权重$\alpha_i$:

$$\alpha_i=\frac{\exp\left(\gamma\left(\boldsymbol{h}_i^l,\boldsymbol{a}_{\text{avg}}\right)\right)}{\sum_{j=1}^{n}\exp\left(\gamma\left(\boldsymbol{h}_j^l,\boldsymbol{a}_{\text{avg}}\right)\right)} \tag{12}$$

式中: γ函数的计算形式为

$$\gamma\left(\boldsymbol{h}_i^l,\boldsymbol{a}_{\text{avg}}\right)=\tanh\left(\boldsymbol{h}_i^l\boldsymbol{W}_a\boldsymbol{a}_{\text{avg}}^{\mathrm{T}}+\boldsymbol{b}_a\right) \tag{13}$$

其中: $\boldsymbol{W}_a$和$\boldsymbol{b}_a$分别为权重矩阵和偏置。

2) 上下文-方面注意力计算。同理, 使用上下文表示$\boldsymbol{c}_{\text{avg}}$来计算方面表示的注意力向量:

$$\beta_i=\frac{\exp\left(\gamma\left(\tilde{\boldsymbol{a}}_i,\boldsymbol{c}_{\text{avg}}\right)\right)}{\sum_{j=1}^{m}\exp\left(\gamma\left(\tilde{\boldsymbol{a}}_j,\boldsymbol{c}_{\text{avg}}\right)\right)} \tag{14}$$

经过上述计算后, 得到上下文和方面表示:

$$\boldsymbol{c}_r=\sum_{i=1}^{n}\alpha_i\boldsymbol{h}_i^l \tag{15}$$

$$\boldsymbol{a}_r=\sum_{i=1}^{n}\beta_i\tilde{\boldsymbol{a}}_i \tag{16}$$

连接方面表示$\boldsymbol{a}_r$和上下文表示$\boldsymbol{c}_r$, 作为最终表示:

$$\boldsymbol{z}=\boldsymbol{c}_r\oplus\boldsymbol{a}_r \tag{17}$$

2.4　情感分类模块

利用交互注意力层的输出$\boldsymbol{z}$作为全连接层输入, 通过softmax函数输出最终情感极性, 即

$$\boldsymbol{y}=\text{softmax}(\boldsymbol{W}\boldsymbol{z}+\boldsymbol{b}) \tag{18}$$

式中: $\boldsymbol{W}$为全连接层权重矩阵; $\boldsymbol{b}$为偏置。

使用交叉熵损失函数进行模型优化:

$$\zeta=-\sum_{i=1}^{J}y_i\ln\hat{y}_i+\lambda\|\boldsymbol{\theta}\| \tag{19}$$

式中：J为所有的训练数据；y_i和$\hat{y}_i$分别为训练集的真实值和预测值；$\boldsymbol{\theta}$为所有可训练参数；λ为L_2正则化系数。

3 实 验

3.1 数据集

使用Twitter[25]、SemEval-2014[26]、SemEval-2015[27]、SemEval-2016[28]4个数据集进行实验，样本如表1所示。

表1 ABSA任务数据集

Table 1 ABSA task datasets

数据集	训练集个数			测试集个数		
	积极	中性	消极	积极	中性	消极
Twitter	1 561	3 127	1 560	173	346	173
Lap14	994	464	870	341	169	128
Rest15	912	36	256	326	34	182
Rest16	1 240	69	439	469	30	117

3.2 参数设置

实验使用GloVe初始词嵌入向量，BiLSTM的隐藏层节点数为300，概念向量维度为300，模型中的权重采用均匀分布进行初始化。使用Adam作为优化器，学习率设置为0.002，L_2正则项系数设置为0.000 02，Twitter、Lap14、Rest15、Rest16批量训练样本分别设置为16、32、16、64。

3.3 对比实验

本文选取10组模型进行对比实验。

1）LSTM[4]。利用LSTM编码上下文信息，将最后一层隐藏向量作为文本特征表示输入分类器中。

2）RAM[29]。将多重注意力捕获的语义特征和LSTM捕获的特征整合，得到最终表示。

3）IAN[13]。一种利用注意力机制对上下文和目标交互学习的模型。

4）MGAN[30]。一种多粒度注意力网络，可以捕获方面与上下文之间的词级交互，降低信息损失。

5）AOA[31]。联合建模方面与上下文，并自动关注句子中的重要部分。

6）ASGCN[5]。依赖树上的GCN来挖掘句法信息和单词依赖，并结合注意力机制进行情感分类。

7）BiGCN[8]。一种结合层次句法和词汇图的双级GCN，更好地利用语料库级别的单词共现信息及不同类型的句法依存关系。

8）GL-GCN[32]。一种引入全局和局部信息的GCN模型，有效提高了文本分类的性能。

9）KEATGCN[33]。一种借助知识图谱和GCN的交互注意力模型，为评论文本注入背景信息和语言知识。

10）MIGCN[34]。一种多交互GCN模型，可以同时考虑到位置距离特征和文本词间的语法信息，并设计2种交互用来捕获其中的关系。

3.4 实验结果

本文按照基准分类进行了3组实验，对比实验结果如表2所示。本文提出的MSKFSA模型的分类准确率和Marco-F_1与3类基准模型相比均展示出优越性。其中，准确率在Twitter、Lap14、Rest15、Rest16数据集上均取得最优，分别达到了75.00%、77.90%、81.55%、90.10%，相较于基准模型的最优结果分别提高了0.43%、0.99%、0.39%、0.6%。在Twitter、Lap14、Rest15数据集上，相较于对比模型最优结果的Marco-F_1，分别提高了0.69%、0.93%、1.57%。另外，在Twitter、Lap14、Rest15数据集上，准确率和Marco-F_1均达到了最优结果，验证了本文模型相较于其他模型的优越性。

根据结果可知，相比仅使用LSTM的模型，基于GCN的模型在4个数据集上表现较优。综上所

表2 不同模型结果对比

Table 2 Results of different models %

类别	模型	准确率				Marco-F_1			
		Twitter	Lap14	Rest15	Rest16	Twitter	Lap14	Rest15	Rest16
LSTM模型	LSTM[4]	69.56	69.28	77.37	86.80	67.70	63.09	55.17	63.88
	RAM[29]	69.36	74.49			67.30	71.35		
交互模型	IAN[13]	72.50	72.05	78.54	84.74	70.81	67.38	52.65	55.21
	MGAN[30]	72.54	75.27			70.81	70.81		
	AOA[31]	72.30	76.62	78.17	87.50	70.20	67.52	57.02	66.21
GCN模型	ASGCN[5]	72.15	75.55	79.89	88.99	70.40	71.05	61.89	67.48
	BiGCN[8]	74.16	74.59	81.16	88.96	73.35	71.84	64.79	70.84
	GL-GCN[32]	73.26	76.91	80.81	88.47	71.26	72.76	64.99	69.64
	MIGCN[34]	73.31	76.59	80.81	89.50	72.12	72.44	64.21	71.97
	KEATGCN[33]	74.57	76.65	80.07	89.45	73.34	73.21	65.90	74.81
本文模型	MSKFSA	75.00	77.90	81.55	90.10	74.03	74.14	67.47	71.99

述, LSTM 能够有效捕捉上下文信息, 但句子长度越长, 与方面情感判别无关的噪声信息也随之增加, 而 GCN 可以较好地解决该问题。通过把握句法依存关系, 进一步学习单词与单词之间的依赖关系, 得到方面的最终表示。

相较于 3 类模型, 本文模型的性能进一步提高, 这是因为对于长度较短的文本, 其本身包含的信息有限, 而模型可以有效融入先验知识来丰富表示。另外, 针对特定情感分析任务, 融合词与词之间的情感得分能够有效指引模型对特定方面情感进行正确的判定。

3.5　消融实验

为进一步检查模型的每个组件对性能的影响, 本文对模型进行消融研究, 结果如表 3 所示。

表 3　消融实验结果

Table 3　Results of ablation experiments　　%

模型	准确率				Marco-F_1			
	Twitter	Lap14	Rest15	Rest16	Twitter	Lap14	Rest15	Rest16
MSKFSA	75.00	77.90	81.55	90.10	74.03	74.14	67.47	71.99
w/o lan	73.99	72.26	79.89	87.18	72.37	67.28	62.81	68.95
w/o em	74.13	72.10	78.78	87.01	72.29	67.16	64.83	64.78
w/o co	74.42	70.85	78.60	86.69	72.91	64.71	60.06	67.65
w/o in	74.57	76.65	80.63	88.96	73.03	72.95	66.05	67.51

1) w/o lan。去掉模型中的上下文信息编码模块中的词汇、句法编码部分。

2) w/o em。去掉模型中的情感增强部分。

3) w/o co。去掉模型中的概念增强部分。

4) w/o in。去掉模型中的特征信息融合模块, 将 2 部分表示拼接做最后的分类。

可以看出, 去除交互注意力指导的 w/o in 模型整体表现优于其他消融模型, 且与完整模型之间差距较小。表明语言知识和外部知识深刻的影响整个情感分析任务, 如果能够对这些知识进行充分的挖掘利用, 有相当重要的作用。比较分别去除情感知识和概念知识的 w/o em 和 w/o co 模型, 在 4 个数据集上的表现基本相同, w/o co 模型在 Lap14、Rest15 数据集上的 Marco-F_1 弱于 w/o em 模型, 但在 Twitter、Rest16 上又优于 w/o em 模型。表明情感知识与概念知识同等重要, 均不可忽略。同时, 去除语言知识的 w/o lan 模型与 w/o in 模型相比, 在 Lap14、Rest15、Rest16 上的准确率有较大下降, 说明虽然有情感、概念知识的武装, 但上下文中具有更重要的语义信息, 语言知识不可或缺。本文所设计的 MSKFSA 模型结合多源知识, 将方面的 3 部分特征较好地融合, 增强了模型的分类能力。

综上所述, MSKFSA 模型可以通过外部知识和 GCN 分别引入概念、情感知识及上下文语义知识, 通过双交互注意力机制学习上下文与方面之间的交互关系, 进一步提高模型的性能。

3.6　GCN 层数的影响

为研究 GCN 层数对 MSKFSA 模型的影响, 在上述数据集中将 GCN 层数设置为 1, 2, …, 8, 以评估 MSKFSA 模型。如图 4 和图 5 所示, 当 MSKFSA 模型层数为 2 时, 综合性能最佳, 但随着网络深度的增加, 模型的 Marco-F_1 波动剧烈。一方面, 当层数较少时, 节点表示不能传播很远; 另一方面, 如果层

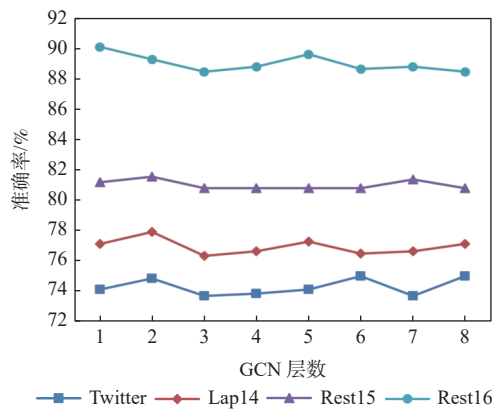


图 4　GCN 层数与准确率的关系

Fig. 4　Relationship between number of GCN layers and accuracy

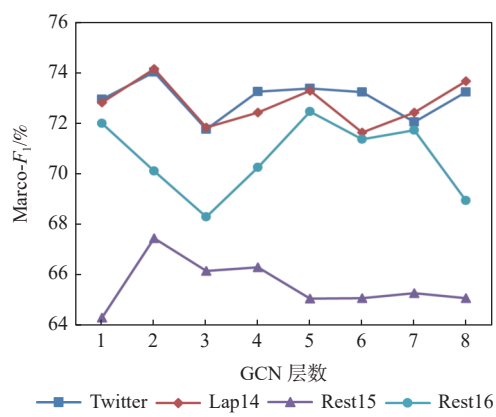


图 5　GCN 层数与 Marco-F_1 的关系

Fig. 5　Relationship between number of GCN layers and Marco-F_1

数过多,由于梯度消失和信息冗余,模型会变得不稳定。

4 结 论

本文提出了一种利用多源知识的交互注意力网络模型。

1)通过融合句法依赖、词共现信息、情感网络和概念图谱等多源知识实现了对评论语句更全面的理解和分析。

2)通过设计双交互注意力机制增强评论语句上下文与评价方面之间的关联,用于实现两部分的协调优化。

3)实验结果显示,语言知识及外部知识同等重要,均不可忽略,且对上下文与方面进行交互是有利的,所提模型在公开数据集上的准确率与Macro-F_1大部分性能指标优于基准模型。

随着预训练语言模型的出现,未来将进一步探索如何融合更多元的知识源,以提高情感分析的效率和准确性。

参考文献(References)

[1] LIU B, ZHANG L. A survey of opinion mining and sentiment analysis[M]//AGGARWAL C C, ZHAI C X. Mining text data. Berlin: Springer, 2012: 415-463.

[2] 唐晓波, 刘广超. 细粒度情感分析研究综述[J]. 图书情报工作, 2017, 61(5): 132-140.
TANG X B, LIU G C. Research review on fine-grained sentiment analysis[J]. Library and Information Service, 2017, 61(5): 132-140 (in Chinese).

[3] KIM Y. Convolutional neural networks for sentence classification [C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2014: 1746-1751.

[4] TANG D Y, QIN B, FENG X C. Effective LSTMs for target-dependent sentiment classification[C]//Proceedings of the 26th International Conference on Computational Linguistics. Stroudsburg: ACL, 2016: 3298-3307.

[5] ZHANG C, LI Q C, SONG D W. Aspect-based sentiment classification with aspect-specific graph convolutional networks[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing. Stroudsburg: ACL, 2019: 4567-4577.

[6] WANG K, SHEN W Z, YANG Y Y, et al. Relational graph attention network for aspect-based sentiment analysis[C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2020: 3229-3238.

[7] 谢珺, 王雨竹, 陈波, 等. 基于双指导注意力网络的属性情感分析模型[J]. 计算机研究与发展, 2022, 59(12): 2831-2843.
XIE J, WANG Y Z, CHEN B, et al. Aspect-based sentiment analysis model with bi-guide attention network[J]. Journal of Computer Research and Development, 2022, 59(12): 2831-2843(in Chinese).

[8] ZHANG M, QIAN T Y. Convolution over hierarchical syntactic and lexical graphs for aspect level sentiment analysis[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2020: 3540-3549.

[9] MA Y K, PENG H Y, CAMBRIA E. Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive LSTM[C]//Proceedings of the 32nd Conference on Artificial Intelligence. Palo Alto: AAAI, 2018: 5876-5883.

[10] ZHOU J, HUANG J X, HU Q V, et al. SK-GCN: Modeling syntax and knowledge via graph convolutional network for aspect-level sentiment classification[J]. Knowledge-Based Systems, 2020, 205: 106292.

[11] ZHAO A P, YU Y. Knowledge-enabled BERT for aspect-based sentiment analysis[J]. Knowledge-Based Systems, 2021, 227: 107220.

[12] WANG Y Q, HUANG M L, ZHU X Y, et al. Attention-based LSTM for aspect-level sentiment classification[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2016: 606-615.

[13] MA D H, LI S J, ZHANG X D, et al. Interactive attention networks for aspect-level sentiment classification[C]//Proceedings of the 26th International Joint Conference on Artificial Intelligence. New York: ACM, 2017: 4068-4074.

[14] PANG S G, XUE Y, YAN Z H, et al. Dynamic and multi-channel graph convolutional networks for aspect-based sentiment analysis [C]//Proceedings of the Findings of the Association for Computational Linguistics. Stroudsburg: ACL, 2021: 2627-2636.

[15] LI R F, CHEN H, FENG F X, et al. Dual graph convolutional networks for aspect-based sentiment analysis[C]//Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing. Stroudsburg: ACL, 2021: 6319-6329.

[16] ZHONG Q H, DING L, LIU J H, et al. Knowledge graph augmented network towards multiview representation learning for aspect-based sentiment analysis[EB/OL]. (2022-01-13)[2022-08-01]. http://arxiv.org/abs/2201.04831.

[17] CHEN J D, HU Y Z, LIU J P, et al. Deep short text classification with knowledge powered attention[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 6252-6259.

[18] BIAN X M, FENG C, AHMAD A, et al. Targeted sentiment classification with knowledge powered attention network[C]//Proceedings of the IEEE 31st International Conference on Tools with Artificial Intelligence. Piscataway: IEEE Press, 2019: 1073-1080.

[19] CAMBRIA E, SPEER R, HAVASI C, et al. SenticNet: A publicly available semantic resource for opinion mining[C]//Proceedings of the AAAI Fall Symposium. Palo Alto: AAAI, 2010: 14-18.

[20] CAMBRIA E, LI Y, XING F Z, et al. SenticNet 6: Ensemble application of symbolic and subsymbolic AI for sentiment analysis [C]//Proceedings of the 29th ACM International Conference on Information & Knowledge Management. New York: ACM, 2020: 105-114.

[21] XING F Z, PALLUCCHINI F, CAMBRIA E. Cognitive-inspired domain adaptation of sentiment lexicons[J]. Information Processing and Management, 2019, 56(3): 554-564.

[22] LIANG B, SU H, GUI L, et al. Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks[J]. Knowledge-Based Systems, 2022, 235: 107643.

[23] PENNINGTON J, SOCHER R, MANNING C. GloVe: Global vec-

tors for word representation[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2014: 1532-1543.

[24] JI L, WANG Y J, SHI B T, et al. Microsoft concept graph: Mining semantic concepts for short text understanding[J]. Data Intelligence, 2019, 1(3): 238-270.

[25] DONG L, WEI F R, TAN C Q, et al. Adaptive recursive neural network for target-dependent twitter sentiment classification[C]//Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2014: 49-54.

[26] PONTIKI M, GALANIS D, PAVLOPOULOS J, et al. SemEval-2014 task 4: Aspect based sentiment analysis[C]//Proceedings of the 8th International Workshop on Semantic Evaluation. Stroudsburg: ACL, 2014: 27-35.

[27] PONTIKI M, GALANIS D, PAPAGEORGIOU H, et al. SemEval-2015 task 12: Aspect based sentiment analysis[C]//Proceedings of the 9th International Workshop on Semantic Evaluation. Stroudsburg: ACL, 2015: 486-495.

[28] PONTIKI M, GALANIS D, PAPAGEORGIOU H, et al. SemEval-2016 task 5: Aspect based sentiment analysis[C]//Proceedings of the 10th International Workshop on Semantic Evaluation. Stroudsburg: ACL, 2016: 19-30.

[29] CHEN P, SUN Z Q, BING L D, et al. Recurrent attention network on memory for aspect sentiment analysis[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2017: 452-461.

[30] FAN F F, FENG Y S, ZHAO D Y. Multi-grained attention network for aspect-level sentiment classification[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg: ACL, 2018: 3433-3442.

[31] HUANG B X, OU Y L, CARLEY K M. Aspect level sentiment classification with attention-over-attention neural networks[M]// THOMSON R, DANCY C, HYDER A, et al. Social, cultural, and behavioral modeling. Berlin: Springer, 2018: 197-206.

[32] ZHU X F, ZHU L, GUO J F, et al. GL-GCN: Global and local dependency guided graph convolutional networks for aspect-based sentiment classification[J]. Expert Systems with Applications, 2021, 186: 115712.

[33] 韩虎, 郝俊, 张千锟, 等. 知识增强的交互注意力方面级情感分析模型[J]. 计算机科学与探索, 2023, 17(3): 709-718.
HAN H, HAO J, ZHANG Q K, et al. Knowledge-enhanced interactive attention model for aspect-based sentiment analysis[J]. Journal of Frontiers of Computer Science and Technology, 2023, 17(3): 709-718(in Chinese).

[34] 王汝言, 陶中原, 赵容剑, 等. 多交互图卷积网络用于方面情感分析[J]. 电子与信息学报, 2022, 44(3): 1111-1118.
WANG R Y, TAO Z Y, ZHAO R J, et al. Multi-interaction graph convolutional networks for aspect-level sentiment analysis[J]. Journal of Electronics & Information Technology, 2022, 44(3): 1111-1118(in Chinese).

Multi-source knowledge fusion model for aspect-based sentiment analysis

HAN Hu*, HAO Jun, ZHANG Qiankun, ZHAO Qitao

(School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: Aspect-based sentiment analysis (ABSA) is a fine-grained sentiment analysis task that aims to give the corresponding sentiment polarity for specific aspects that appear in review statements. Most existing aspect-based sentiment analysis methods relying on deep learning focus on mining the semantics and syntax of review statements, often ignoring the conceptual knowledge and sentiment degree information that may be involved in the review statements. To address this problem, a neural network model incorporating multi-source knowledge. It reveals the structural framework of sentences through syntactic dependencies, captures semantic connections between words through word co-occurrence, and embeds emotional networks and concept graphs to provide emotional and background knowledge for the model, and coordinated optimization of the contextual and evaluative aspects of review statements was realized through a dual-interaction attention model. Experimental results on four public datasets show that the model achieves better performance than existing models, with accuracy reaching 75.00%, 77.90%, 81.55%, and 90.10%, respectively, all of which were improved compared to the benchmark model. This achievement not only verifies the effectiveness of multi-source knowledge fusion in ABSA tasks, but also provides new ideas and methods for future research.

Keywords: aspect-based sentiment analysis; graph convolutional network; multi-source fusion; knowledge graph; interactive attention mechanisms

Received: 2022-08-17; **Accepted**: 2023-01-14; **Published Online**: 2023-03-23 09:01

URL: link.cnki.net/urlid/11.2625.V.20230322.0955.001

Foundation item: National Natural Science Foundation of China (62166024)

* **Corresponding author**. E-mail: hanhu_lzjtu@mail.lzjtu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0729

基于到达时间间隔差的 ADS-B 位置消息验证方法

刘海涛*，刘家祥，李冬霞，王磊
(中国民航大学 电子信息与自动化学院，天津 300300)

摘 要：广播式自动相关监视 (ADS-B) 是一种新型的航空器监视技术，具有定位精度高、更新速度快及覆盖范围广等优势，在民用航空监视领域获得广泛应用。由于 ADS-B 消息以广播方式传输，未采用加密与认证机制，ADS-B 系统易受外部欺骗源的干扰。为解决该问题，提出了一种基于到达时间间隔差 (IDOA) 的 ADS-B 位置消息验证方法。建立基于 IDOA 的 ADS-B 位置消息验证系统模型，理论分析给出检验统计量的表达式及其统计特性，进一步分析给出检测门限的确定方法，通过仿真验证所提方法的正确性与有效性。结果表明：所提方法与基于到达时间差 (TDOA) 的验证方法具有相同的检测性能，且对时间测量误差和 ADS-B 位置误差不敏感，但所提方法可克服基于 TDOA 的验证方法存在的对地面站时间同步误差敏感的缺点。

关 键 词：广播式自动相关监视；欺骗干扰；到达时间间隔差；检测门限；地面站时间同步误差

中图分类号：TN925

文献标志码：A **文章编号**：1001-5965(2024)09-2696-08

广播式自动相关监视 (automatic dependent surveillance-broadcast, ADS-B) 是一种新型的航空器监视技术[1-3]。机载 ADS-B 发射机通过全球卫星导航系统(global navigation satellite system, GNSS)及机载设备获取航空器位置、速度、身份等信息，并生成 ADS-B 消息。ADS-B 消息通过脉冲位置调制转换为 1 090 MHz ADS-B 信号，并随机广播发射，ADS-B 地面站接收 ADS-B 信号，并完成信号的解调与解码，最终多个 ADS-B 消息封装为 ASTERIX CAT 021 报告，并通过地面网络分发给空中交通管制终端，管制员依据管制终端显示的航空器监视信息提供空中交通管制服务。与传统的航空器监视手段相比，ADS-B 系统具有定位精度高、更新速率快、覆盖范围广及建设运行成本低等多方面优势，因此，其在民用航空监视领域获得广泛应用[4]。由于 ADS-B 系统以广播方式工作，且 ADS-B 消息格式公开，ADS-B 系统没有采取任何加密与认证机制，容易受到外部欺骗源的干扰[5-7]，严重威胁民用航空的飞行安全。因此，针对 ADS-B 系统存在的欺骗式干扰问题，开展 ADS-B 位置消息验证方法的研究对保障民航飞行安全具有重要意义。

近年来，围绕着 ADS-B 系统存在的欺骗式干扰问题，国内外开展了大量研究，可分为如下 5 类：基于加密认证的验证方法、基于波达方向的验证方法、基于信号特征的验证方法、基于机器学习和深度学习的验证方法及基于到达时间差(time difference of arrival, TDOA)的验证方法。

基于加密认证的验证方法的核心思想为[8-12]：通过密钥和密码算法对 ADS-B 链路进行加密，此类方法要求修改 ADS-B 系统协议架构[13]，所有机载及地面站设备均需要更换。基于波达方向的验证方法的核心思想[14-15]为：地面站通过接收航空器播发 ADS-B 位置消息，获得航空器的空间位置坐标，同时根据地面站的位置坐标，得到航空器的电波来

收稿日期：2022-08-17；**录用日期**：2022-09-16；**网络出版时间**：2022-11-14 14:52
网络出版地址：link.cnki.net/urlid/11.2625.V.20221111.1711.002
基金项目：国家重点研发计划 (2016YFB0502402)；国家自然科学基金 (U1733120)；天津市多元投入基金项目青年项目 (21JCQNJC00770)
* **通信作者**. E-mail：htliucauc@qq.com

引用格式：刘海涛，刘家祥，李冬霞，等. 基于到达时间间隔差的 ADS-B 位置消息验证方法 [J]. 北京航空航天大学学报，2024，50(9)：2696-2703. LIU H T，LIU J X，LI D X，et al. Verification scheme of position message for automatic dependent surveillance-broadcast system based on interval difference of arrival [J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2696-2703 (in Chinese).

向, 此外, 通过波达方向估计方法获得接收信号的电波来向, 对 2 个来向信息进行比较, 判定航空器播发 ADS-B 位置消息的真实性, 此类方法要求地面站具备信号测向的能力, 而空间测向设备较复杂。基于信号特征的验证方法的核心思想[16-18]为: 地面站利用接收信号的强度、多普勒频偏等信号特征来验证航空器位置消息的真实性, 此类方法存在的问题是检测性能较差[19]。基于机器学习和深度学习的验证方法的核心思想[20-21]为: 通过对 ADS-B 数据的特征进行挖掘, 建立不同的检测模型来验证航空器位置消息的真实性, 此类方法存在的问题是未考虑航空器机动状态改变所带来的 ADS-B 数据变化, 适应性不高。

基于 TDOA 的验证方法的研究现状如下: 文献 [22-23] 提出通过多点定位对航空器位置消息进行验证的方法, 要求地面站数量至少 4 个, 且各地面站之间保持精确的时间同步; 文献 [24] 提出利用信号传输时间的航空器位置消息验证方法, 该方法存在的主要问题是需要精确知晓航空器 ADS-B 位置消息的播发时刻, 在实际应用中, 地面站无法精确获得航空器 ADS-B 消息的播发时刻, 难以应用于实际系统; 文献 [25] 提出利用 TDOA 的位置消息验证方法, 该方法要求地面站数量为 2 个, 但存在的问题仍是要求各地面站之间保持精确的时间同步。

为克服基于 TDOA 的验证方法存在的问题, 本文提出了一种基于到达时间间隔差(interval difference of arrival, IDOA)的 ADS-B 位置消息验证方法。在不增加地面站数量的情况下, 可获得与基于 TDOA 的验证方法相同的检测性能, 且对时间测量误差和 ADS-B 位置误差不敏感, 克服了基于 TDOA 的验证方法对地面站定时同步误差敏感的缺点。本文创新点如下: ①提出了基于 IDOA 的 ADS-B 位置消息验证方法; ②理论分析给出了检测门限的设定方法。

1　系统模型

图 1 给出了基于 IDOA 的 ADS-B 位置消息验证系统模型。系统由 4 个节点组成: 地面站 A(GSA)、地面站 B(GSB)、航空器(AV)和位置验证终端(LVT)。在飞行过程中, 机载 ADS-B 发射机通过 GNSS 及机载设备获取航空器位置、速度、身份等信息, 并生成 ADS-B 消息, ADS-B 消息经过脉冲位置调制后, 随机广播发射。GSA 节点与 GSB 节点均安装 ADS-B 接收机, 其接收航空器播发的 ADS-B 消息, 并记录 ADS-B 位置消息的接收时刻, 2 个节点将接收到的 ADS-B 位置消息及接收时刻信息

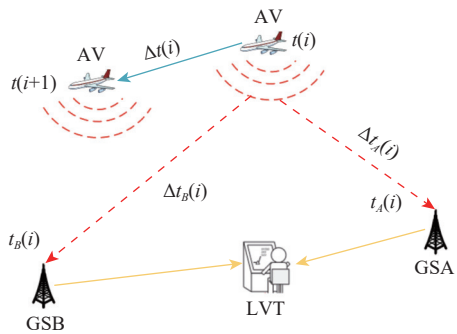


图 1　基于 IDOA 的 ADS-B 位置消息验证系统模型

Fig. 1　ADS-B position verification system model based on IDOA

传输至 LVT 节点。LVT 节点接收 2 个节点传输的 ADS-B 位置消息及接收时刻, 计算相同航空器 ADS-B 位置消息的 IDOA 值, 并与预先设定的检测门限进行比较, 以确定航空器播发 ADS-B 位置消息的真实性。

假设 GSA 节点和 GSB 节点的三维位置坐标分别为 (x_A,y_A,z_A)和(x_B,y_B,z_B)。航空器播发 ADS-B 位置消息的序号记为$\{i=1,\cdots,I\}$, 第i条 ADS-B 位置消息的播发时刻为$t(i)$, 该时刻航空器位置坐标为$\left[x(i),y(i),z(i)\right]$, 而航空器通过 GNSS 获得的航空器位置坐标为$\left[\tilde{x}(i),\tilde{y}(i),\tilde{z}(i)\right]$, 此外航空器播发第$i+1$条位置消息与第$i$条位置消息的时间间隔记为$\Delta t(i)$, $t(i)$时刻航空器播发的第i条位置消息经时延$\Delta t_A(i)$传输至 GSA 节点, 假设该消息被 GSA 节点正确接收, 且 GSA 节点记录该消息的接收时刻为$t_A(i)$。同理, 航空器播发的第i条位置消息经时延$\Delta t_B(i)$传输至 GSB 节点, 假设该消息同时也被 GSB 节点正确接收, 且 GSB 节点记录该消息的接收时刻为$t_B(i)$。

2　本文方法

2.1　基于消息接收时刻的 IDOA

GSA 节点接收第i条和第$i+1$条位置消息的时间间隔为

$$T_A(i)=t_A(i+1)-t_A(i) \tag{1}$$

同理, GSB 节点接收第i条和第$i+1$条位置消息的时间间隔为

$$T_B(i)=t_B(i+1)-t_B(i) \tag{2}$$

根据式 (1) 和式 (2), 引入参量$\Delta T_{AB}(i)$:

$$\Delta T_{AB}(i)=T_B(i)-T_A(i) \tag{3}$$

式中: $\Delta T_{AB}(i)$ 为 GSA 节点接收第i条和第$i+1$条位置消息的时间间隔与 GSB 节点接收第i条和第$i+1$条位置消息的时间间隔的差值。

根据附录 A 的推导, $\Delta T_{AB}(i)$ 进一步表示为

$$\Delta T_{AB}(i)=\Delta t_B(i+1)-\Delta t_B(i)-\Delta t_A(i+1)+\Delta t_A(i) \tag{4}$$

2.2 基于 ADS-B 位置消息的 IDOA

$t(i)$时刻，航空器位于$[x(i),y(i),z(i)]$点，其播发的第i条位置消息于$t(i)+\Delta t_A(i)$和$t(i)+\Delta t_B(i)$时刻分别被 GSA 节点和 GSB 节点接收，GSA 节点与 GSB 节点解析位置消息得到航空器的报告点坐标为$[\tilde{x}(i),\tilde{y}(i),\tilde{z}(i)]$，则航空器报告点与 GSA 节点的距离为

$$d_A(\tilde{x}(i),\tilde{y}(i),\tilde{z}(i))=\sqrt{(x_A-\tilde{x}(i))^2+(y_A-\tilde{y}(i))^2+(z_A-\tilde{z}(i))^2} \tag{5}$$

同理，航空器报告点与 GSB 节点距离为

$$d_B(\tilde{x}(i),\tilde{y}(i),\tilde{z}(i))=\sqrt{(x_B-\tilde{x}(i))^2+(y_B-\tilde{y}(i))^2+(z_B-\tilde{z}(i))^2} \tag{6}$$

根据式 (4)，以及第i条和第$i+1$条位置消息，计算得到的 IDOA 表示为

$$\begin{aligned}\Delta\tilde{T}_{AB}(i)=&d_B(\tilde{x}(i+1),\tilde{y}(i+1),\tilde{z}(i+1))/c-\\&d_B(\tilde{x}(i),\tilde{y}(i),\tilde{z}(i))/c-d_A(\tilde{x}(i+1),\tilde{y}(i+1),\tilde{z}(i+1))/c+\\&d_A(\tilde{x}(i),\tilde{y}(i),\tilde{z}(i))/c\end{aligned} \tag{7}$$

式中：c为电磁波传播速度。

为比较 2 种方法获得的航空器 IDOA 值，设置以下仿真场景：地面站 A 的三维坐标为 (10,10,0) km，地面站 B 的三维坐标为 (150,150,0) km。仿真航空器以 1 000 km/h 的飞行速度从 (0,40,10) km 直线飞行至 (180,100,0) km，飞行过程中，航空器播发 ADS-B 位置消息速率为 2 条/s。静止欺骗源位于 (60,160,0) km，其欺骗产生一条与仿真航空器航迹相同的虚假航迹，虚假航迹的 ADS-B 位置消息播发速率也设置为 2 条/s。ADS-B 水平位置误差标准差设置为 4×10^{-3} km[26]，ADS-B 垂直位置误差标准差为 6.6×10^{-3} km[26]，地面站时间测量误差标准差设置为 6.4×10^{-9} s[25]，仿真时长为 1 366 s。

图 2 给出仿真场景下$\Delta T_{AB}\cdot c$与仿真时间的关系曲线，$\Delta T_{AB}\cdot c$代表播发相邻 2 条位置消息时航空器至 GSA 节点的距离差与航空器至 GSB 节点距离差的差值。比较表明：①对于真实航空器，2 种方法获得的 ADS-B 位置消息的$\Delta T_{AB}\cdot c$非常接近；②对于欺骗源，2 种方法获得的 ADS-B 位置消息的$\Delta T_{AB}\cdot c$差异非常大。

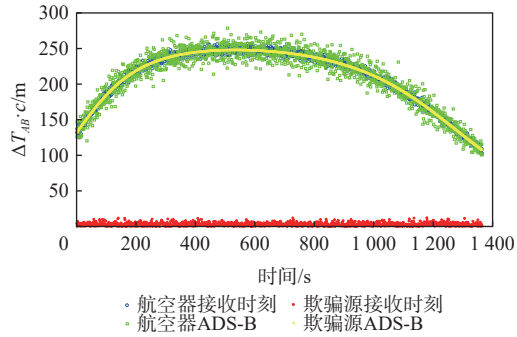


图 2 仿真场景下$\Delta T_{AB}\cdot c$与仿真时间的关系曲线

Fig. 2 Relationship curve between $\Delta T_{AB}\cdot c$ and simulation time in simulation scenario

2.3 检验统计量与检测门限

为检验 ADS-B 位置消息的异常，引入检验统计量J：

$$J=\Delta T_{AB}(i)-\Delta\tilde{T}_{AB}(i) \tag{8}$$

假设检测门限设置为γ，则判定 ADS-B 位置消息异常的准则为

$$\begin{cases}H_0:|J|\leqslant\gamma\\H_1:|J|>\gamma\end{cases} \tag{9}$$

式中：H_0表示航空器 ADS-B 位置消息正常的事件；H_1表示航空器 ADS-B 位置消息异常的事件。

由于地面站接收机记录 ADS-B 信号达到时刻时存在测量误差[25]，式 (8) 给出的检验统计量J修正为

$$\begin{aligned}J=&\Delta t_B(i+1)+\varepsilon_B(i+1)-\Delta t_B(i)-\varepsilon_B(i)-\\&\Delta t_A(i+1)-\varepsilon_A(i+1)+\Delta t_A(i)+\varepsilon_A(i)-\\&d_B(\tilde{x}(i+1),\tilde{y}(i+1),\tilde{z}(i+1))/c+d_B(\tilde{x}(i),\tilde{y}(i),\tilde{z}(i))/c+\\&d_A(\tilde{x}(i+1),\tilde{y}(i+1),\tilde{z}(i+1))/c-d_A(\tilde{x}(i),\tilde{y}(i),\tilde{z}(i))/c\end{aligned} \tag{10}$$

式中：$\varepsilon_A(i)$为 GSA 节点接收第i条 ADS-B 位置消息的时间测量误差；$\varepsilon_B(i)$为 GSB 节点接收第i条 ADS-B 位置消息的时间测量误差。

航空器播发 ADS-B 位置消息时，考虑到航空器通过 GNSS 获得的航空器位置与真实位置存在一定的误差，为引入 GNSS 系统的定位误差[25]，假设$t(i)$时刻，航空器位于$[x(i),y(i),z(i)]$，其播发的第i条 ADS-B 消息传输至地面站 A 的时间延迟$\Delta t_A(i)$表示为

$$\Delta t_A(i)=d_A(x(i),y(i),z(i))/c \tag{11}$$

同理，该消息传输至地面站 B 的时间延迟$\Delta t_B(i)$可表示为

$$\Delta t_B(i)=d_B(x(i),y(i),z(i))/c \tag{12}$$

将式 (11) 和式 (12) 代入式 (10) 得

$$\begin{aligned}J=&d_B(x(i+1),y(i+1),z(i+1))/c-d_B(x(i),y(i),z(i))/c-\\&d_A(x(i+1),y(i+1),z(i+1))/c+d_A(x(i),y(i),z(i))/c-\\&d_B(\tilde{x}(i+1),\tilde{y}(i+1),\tilde{z}(i+1))/c+d_B(\tilde{x}(i),\tilde{y}(i),\tilde{z}(i))/c+\\&d_A(\tilde{x}(i+1),\tilde{y}(i+1),\tilde{z}(i+1))/c-d_A(\tilde{x}(i),\tilde{y}(i),\tilde{z}(i))/c+\\&\varepsilon_B(i+1)-\varepsilon_B(i)-\varepsilon_A(i+1)+\varepsilon_A(i)\end{aligned} \tag{13}$$

假设$[\Delta x(i),\Delta y(i),\Delta z(i)]$代表 GNSS 系统的定位误差，将其代入到式 (13)，J进一步修正为

$$J = d_B(\tilde{x}(i+1)+\Delta x(i+1),\tilde{y}(i+1)+\Delta y(i+1),\tilde{z}(i+1)+\Delta z(i+1))/c - d_B(\tilde{x}(i)+\Delta x(i+1),\tilde{y}(i)+\Delta y(i+1),\tilde{z}(i)+\Delta z(i+1))/c - d_A(\tilde{x}(i+1)+\Delta x(i+1),\tilde{y}(i+1)+\Delta y(i+1),\tilde{z}(i+1)+\Delta z(i+1))/c + d_A(\tilde{x}(i)+\Delta x(i+1),\tilde{y}(i)+\Delta y(i+1),\tilde{z}(i)+\Delta z(i+1))/c - d_B(\tilde{x}(i+1),\tilde{y}(i+1),\tilde{z}(i+1))/c + d_B(\tilde{x}(i),\tilde{y}(i),\tilde{z}(i))/c + d_A(\tilde{x}(i+1),\tilde{y}(i+1),\tilde{z}(i+1))/c - d_A(\tilde{x}(i),\tilde{y}(i),\tilde{z}(i))/c + \varepsilon_B(i+1)-\varepsilon_B(i)-\varepsilon_A(i+1)+\varepsilon_A(i) \quad (14)$$

为计算检验统计量 J 的概率密度函数分布形式, 根据附录 B 的推导, 将检验统计量 J 表示为

$$J \cong (b_{x,i+1}\Delta x(i+1)+b_{y,i+1}\Delta y(i+1)+b_{z,i+1}\Delta z(i+1))/c - (b_{x,i}\Delta x(i+1)+b_{y,i}\Delta y(i+1)+b_{z,i}\Delta z(i+1))/c - (a_{x,i+1}\Delta x(i+1)+a_{y,i+1}\Delta y(i+1)+a_{z,i+1}\Delta z(i+1))/c - (a_{x,i}\Delta x(i+1)+a_{y,i}\Delta y(i+1)+a_{z,i}\Delta z(i+1))/c + \varepsilon_B(i+1)-\varepsilon_B(i)-\varepsilon_A(i+1)+\varepsilon_A(i) \quad (15)$$

进一步根据附录 C 的推导, 得到检验统计量 J 的概率密度函数服从高斯分布。

根据 J 的分布特性, 得到航空器 ADS-B 位置消息正常判决为航空器 ADS-B 位置消息异常的概率(虚警概率) P_{FA} 表示为

$$P_{\mathrm{FA}} = P(H_1/H_0) = P(|J|>\gamma/H_0) = 2Q\left(\frac{\gamma}{\sigma_J}\right) \quad (16)$$

式中: $Q(x)$ 为高斯 Q 函数; σ_J 为 J 的标准差。

假设虚警概率 P_{FA} 取值为 0.01, 根据式 (16) 得到检测门限为

$$\gamma = \sigma_J Q^{-1}\left(\frac{P_{\mathrm{FA}}}{2}\right) \quad (17)$$

3　仿真实验

3.1　仿真参数

为验证本文方法的正确性, 构建基于 IDOA 的 ADS-B 位置消息验证仿真系统。仿真系统由地面站 A、地面站 B、静止欺骗源(SC)、航空器(AV)等 4 个节点组成。表 1 给出仿真系统的主要技术参数。

表 1　仿真参数设置

Table 1　Simulation parameter setting

参数	数值
地面站A的位置坐标/km	(140,140,0)
地面站B的位置坐标/km	(40,40,0)
静止欺骗源SC的位置坐标/km	(157.68,4.67,0)
航空器AV的航迹	随机生成900条航迹
时间测量误差的标准差/s	6.4×10^{-9}
地面站时间同步误差/s	2×10^{-8}
ADS-B水平位置误差的标准差/km	4×10^{-3}
ADS-B垂直位置误差的标准差/km	6.6×10^{-3}
虚警概率	0.01

3.2　仿真结果

图 3 给出了 2 个地面站位置关系及航迹示意图。为保证实验的可靠性, 900 条航空器 AV 航迹均为随机生成, 每条航迹包含 1 099 个连续的航空器空中位置, 航空器飞行高度恒定为 10 km。

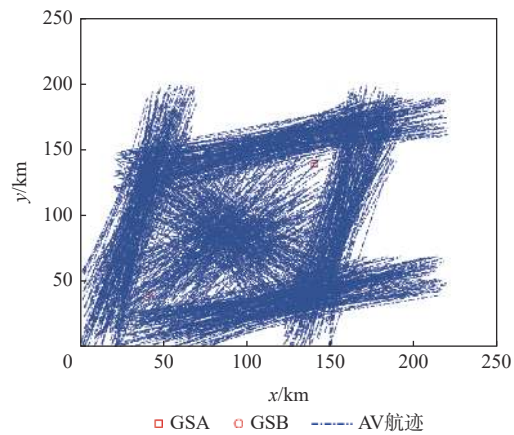


图 3　2 个地面站位置关系及航迹示意图

Fig. 3　Position relationship between two ground stations and flight path diagram

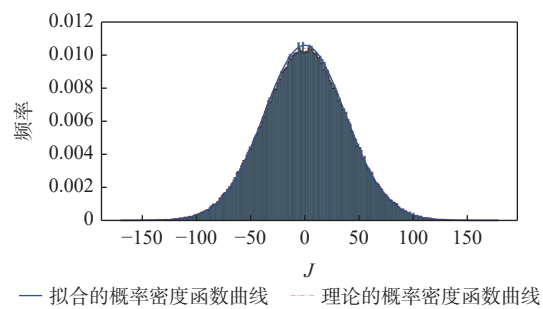


图 4　检验统计量 J 的分布直方图

Fig. 4　Histogram of distribution of test statistic J

图 4 给出了检验统计量 J 的分布直方图(根据图 3)。2 条曲线比较表明, 检验统计量 J 的理论概率密度曲线与拟合得到的概率密度曲线完全一致, 验证了推导出的检验统计量 J 分布特性的正确性。

图 5 给出了 2 个地面站与欺骗源位置关系及航迹示意图。为保证实验的可靠性, 900 条航空器 AV 航迹和 900 条虚假航迹均为随机生成, 每条航迹包含 1 099 个连续的航空器空中位置, 航空器飞行高度恒定为 10 km。

表 2 给出了 TDOA 与 IDOA 检测性能比较(地面站时间同步误差为 2×10^{-8} s[26])。2 组数据比较表明: ①TDOA 将虚假航迹判决为虚假航迹的概率(检测概率)为 0.999 8, 将正常航空器航迹判决为虚

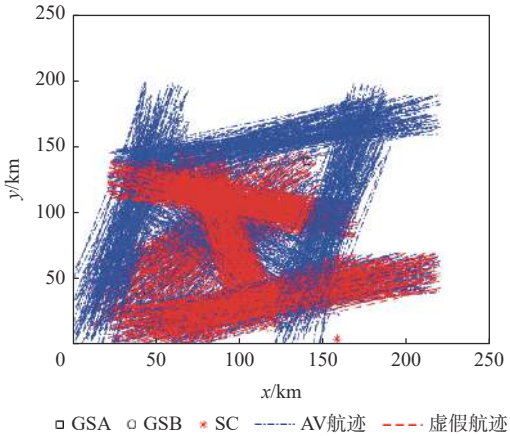


图5 2个地面站与欺骗源位置关系及航迹示意图

Fig. 5 Position relationship between two ground stations and spoofing source and flight path diagram

表2 TDOA与IDOA的检测性能比较

Table 2 Comparison of detection performance between TDOA and IDOA

方法	虚警概率	检测概率
TDOA	0.032 5	0.999 8
IDOA	0.015 8	0.998 6

假航迹的概率(虚警概率)为0.032 5; IDOA的检测概率为0.998 6, 虚警概率为0.015 8。表明IDOA与TDOA具有相同的检测性能。②IDOA虚警概率统计值为0.015 8, 与确定检测门限时所设置的虚警概率0.01非常接近, 进一步验证了检验统计量J概率密度函数的分析是正确的。

图6给出了地面站时间同步误差对TDOA和IDOA检测性能的影响。仿真结果表明: ①随着地面站时间同步误差由1.5×10^{-8} s逐渐增大至3.5×10^{-8} s, IDOA的虚警概率始终维持在0.01左右, TDOA的虚警概率由0.026 3逐渐增大至0.116 7; ②随着地面站时间同步误差由1.5×10^{-8} s逐渐增大至3.5×10^{-8} s, IDOA和TDOA的检测概率均维持在0.990 0附近; ③TDOA对地面站时间同步误差非常敏感, 而IDOA对地面站时间同步误差不敏感。导致该结果的原因为: 根据式(3), 基于消息接收时刻的IDOA值的计算公式不涉及地面站时间同步误差, 故IDOA对地面站时间同步误差不敏感; 基于消息接收时刻的TDOA值的计算公式为$\Delta T_{AB}(i)=t_B(i)-t_A(i)$, 故TDOA对地面站时间同步误差敏感。

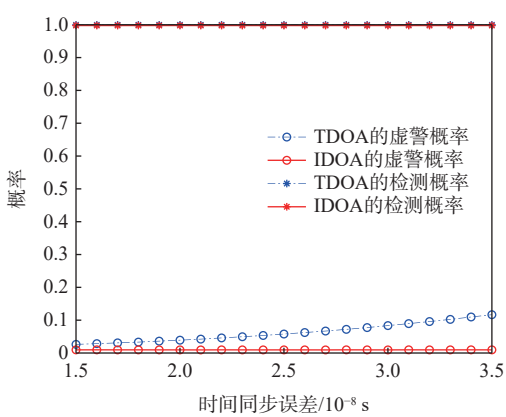


图6 地面站时间同步误差对TDOA与IDOA检测性能的影响

Fig. 6 Influence of time synchronization error of ground station on detection performance of TDOA and IDOA

图7给出了时间测量误差对TDOA和IDOA检测性能的影响。仿真结果表明: ①随着时间测量误差由0.6×10^{-8} s逐渐增大至1.5×10^{-8} s, IDOA的虚警概率由0.032 0逐渐减小至0附近, TDOA的虚警概率由0.062 6逐渐减小至0附近; ②随着时间测量误差由0.6×10^{-8} s逐渐增大至1.5×10^{-8} s, IDOA的检测概率始终维持在0.990 0附近, TDOA的检测概率始终维持在0.990 0附近; ③IDOA和TDOA都对时间测量误差不敏感。

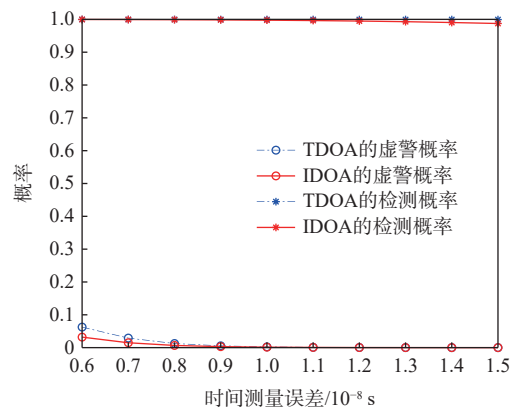


图7 时间测量误差对TDOA与IDOA检测性能的影响

Fig. 7 Influence of time measurement error on detection performance of TDOA and IDOA

图8给出了ADS-B水平位置误差对TDOA和IDOA检测性能的影响。仿真结果表明: ①随着ADS-B水平位置误差由1×10^{-3} km逐渐增大至7.5×10^{-3} km, IDOA的虚警概率始终维持在0.01左右, TDOA方法的虚警概率始终维持在0.01左右; ②随着ADS-B水平位置误差由1×10^{-3} km逐渐增大至7.5×10^{-3} km, IDOA的检测概率始终维持在0.990 0附近, TDOA的检测概率始终维持在0.990 0附近; ③IDOA和TDOA对ADS-B水平位置误差不敏感。

图9给出了ADS-B垂直位置误差对TDOA和IDOA检测性能的影响。仿真结果表明: ①随着ADS-B垂直位置误差由6.5×10^{-3} km逐渐增大至17.5×10^{-3} km, IDOA的虚警概率始终维持在0.01左右, TDOA的虚警概率始终维持在0.02左右; ②随着ADS-B垂直位置误差由6.5×10^{-3} km逐渐增大至17.5×10^{-3} km, IDOA的检测概率始终维持在0.990 0

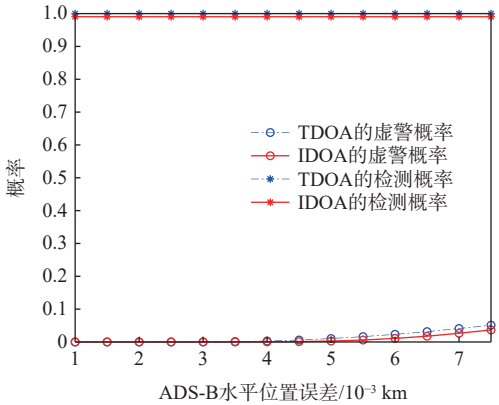


图 8　ADS-B 水平位置误差对 TDOA 与 IDOA 检测性能的影响

Fig. 8　Influence of horizontal position error of ADS-B on detection performance of TDOA and IDOA

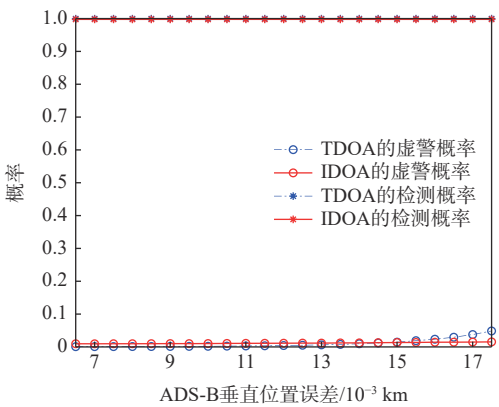


图 9　ADS-B 垂直位置误差对 TDOA 与 IDOA 检测性能的影响

Fig. 9　Influence of vertical position error of ADS-B on detection performance of TDOA and IDOA

附近, TDOA 的检测概率始终维持在 0.990 0 附近; ③IDOA 和 TDOA 对 ADS-B 垂直位置误差不敏感。

4　结　论

本文提出了一种基于 IDOA 的 ADS-B 位置消息验证方法, 包括给出检验统计量表达式、统计特性及检测门限。通过仿真实验, 得出以下结论:

1) 相同条件下, IDOA 与 TDOA 具有相同的检测性能, 且对时间测量误差和 ADS-B 位置误差不敏感。

2) 随着地面站时间同步误差的不断增大, TDOA 的检测能力逐渐下降, 而 IDOA 的检测能力则不受时间同步误差的影响。

值得说明的是, 本文研究的是基于 IDOA 的陆基 ADS-B 位置消息验证方法, 针对陆基 ADS-B 系统尚无法满足航空监视全球化的缺陷, 未来可以进一步将 IDOA 应用于星基 ADS-B 系统中。

参考文献 (References)

[1] 张军. 空域监视技术的新进展及应用[J]. 航空学报, 2011, 32(1): 1-14.
ZHANG J. New development and application of airspace surveillance technology[J]. Acta Aeronautica et Astronautica Sinica, 2011, 32(1): 1-14(in Chinese).

[2] ICAO. Technical provisions for mode S services and extended squitter: ICAO 9871[S]. Montreal: ICAO, 2012.

[3] RTCA DO-260B. Minimum operational performance standards for 1 090 MHz extended squitter automatic dependent surveillance-broadcast (ADS-B) and traffic information services-broadcast (TIS-B): ED-102A[S]. Washington, D. C. : RTCA, 2009.

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

[5] STROHMEIER M, SCHAFER M, LENDERS V, et al. Realities and challenges of nextgen air traffic management: The case of ADS-B[J]. IEEE Communications Magazine, 2014, 52(5): 111-118.

[6] STROHMEIER M, LENDERS V, MARTINOVIC I. On the security of the automatic dependent surveillance-broadcast protocol[J]. IEEE Communications Surveys & Tutorials, 2015, 17(2): 1066-1087.

[7] SCHÄFER M, LENDERS V, MARTINOVIC I. Experimental analysis of attacks on nextgeneration air traffic communication [M]//ABDALLA M, POINTCHEVAL D, FOUQUE P A, et al. Applied cryptography and network security. Berlin: Springer, 2013: 253-271.

[8] FINKE C, BUTTS J, MILLS R, et al. Enhancing the security of aircraft surveillance in the next generation air traffic control system[J]. International Journal of Critical Infrastructure Protection, 2013, 6(1): 3-11.

[9] HABLEEL E, BAEK J, BYON Y J, et al. How to protect ADS-B: Confidentiality framework for future air traffic communication [C]//Proceedings of the IEEE Conference on Computer Communications Workshops. Piscataway: IEEE Press, 2015: 155-160.

[10] AMIN S, CLARK T, OFFUTT R, et al. Design of a cyber security framework for ADS-B based surveillance systems[C]//Proceedings of the Systems and Information Engineering Design Symposium. Piscataway: IEEE Press, 2014: 304-309.

[11] LI W, KAMAL P. Integrated aviation security for defense-in-depth of next generation air transportation system[C]//Proceedings of the IEEE International Conference on Technologies for Homeland Security. Piscataway: IEEE Press, 2011: 136-142.

[12] BAEK J, BYON Y J, HABLEEL E, et al. Making air traffic surveillance more reliable: A new authentication framework for automatic dependent surveillance-broadcast (ADS-B) based on online/offline identity-based signature[J]. Security and Communication Networks, 2015, 8(5): 740-750.

[13] 丁建立, 邹云开, 王静, 等. 基于深度学习的 ADS-B 异常数据检测模型[J]. 航空学报, 2019, 40(12): 323220.
DING J L, ZOU Y K, WANG J, et al. ADS-B anomaly data detection model based on deep learning[J]. Acta Aeronautica et Astro-

nautica Sinica, 2019, 40(12): 323220(in Chinese).

[14] WANG W Y, CHEN G, WU R B, et al. A low-complexity spoofing detection and suppression approach for ADS-B[C]//Proceedings of the Integrated Communication, Navigation and Surveillance Conference. Piscataway: IEEE Press, 2015: 1-8.

[15] 王文益, 陈庚, 吴仁彪, 等. 基于十字阵列的ADS-B欺骗式干扰抑制方法: CN104360323A[P]. 2015-02-18.
WANG W Y, CHEN G, WU R B, et al. ADS-B deception jamming restraining method based on cross array: CN104360323A[P]. 2015-02-18(in Chinese).

[16] STROHMEIER M, LENDERS V, MARTINOVIC I. Intrusion detection for airborne communication using PHY-layer information [M]//JULISCH K, KRUEGEL C. Detection of intrusions and malware, and vulnerability assessment. Berlin: Springer, 2015: 67-77.

[17] SCHÄFER M, LEU P, LENDERS V, et al. Secure motion verification using the Doppler effect[C]//Proceedings of the 9th ACM Conference on Security & Privacy in Wireless and Mobile Networks. New York: ACM, 2016: 135-145.

[18] GHOSE N, LAZOS L. Verifying ADS-B navigation information through Doppler shift measurements[C]//Proceedings of the IEEE/AIAA 34th Digital Avionics Systems Conference. Piscataway: IEEE Press, 2015: 1-27.

[19] 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.

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

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

[22] NUSEIBEH B, HALEY C B, FOSTER C. Securing the skies: Inrequirements we trust[J]. Computer, 2009, 42(9): 64-72.

[23] STROHMEIER M, MARTINOVIC I, LENDERS V. A k-NN-based localization approach for crowdsourced air traffic communication networks[J]. IEEE Transactions on Aerospace Electronic Systems, 2018, 54(3): 1519-1529.

[24] SCHAFER M, LENDERS V, SCHMITT J. Secure track verification[C]//Proceedings of the IEEE Symposium on Security and Privacy. Piscataway: IEEE Press, 2015: 199-213.

[25] NAGANAWA J, MIYAZAKI H. A theory of aircraft position verification using TDOA[C]//Proceedings of the Asia-Pacific Microwave Conference. Piscataway: IEEE Press, 2018: 833-835.

[26] US Department of Defense. Global positioning system standard positioning service performance standard[S]. Washington, D. C. : US Department of Defense, 2008.

附录 A

图 A1 给出航空器播发第i条和第$i+1$条位置消息时各节点时间示意图。其中，$t(i)$和$t(i+1)$的关系为：$t(i+1)=t(i)+\Delta t(i)$，其他时刻之间的关系类似。

结合图 A1 和式 (1)，$T_A(i)$可以进一步表示为

$$T_A(i)=\Delta t(i)+\Delta t_A(i+1)-\Delta t_A(i) \tag{A1}$$

结合图 A1 和式 (2)，$T_B(i)$可以进一步表示为

$$T_B(i)=\Delta t(i)+\Delta t_B(i+1)-\Delta t_B(i) \tag{A2}$$

将式 (A1) 和式 (A2) 代入式 (3)，$\Delta T_{AB}(i)$可以进一步表示为

$$\Delta T_{AB}(i)=\Delta t_B(i+1)-\Delta t_B(i)-\Delta t_A(i+1)+\Delta t_A(i)$$

即得到式 (4)。

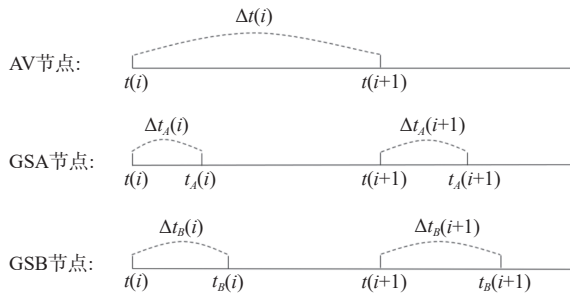


图 A1 航空器播发第 i 条和第 i+1 条位置消息时各节点时间示意图

Fig. A1 Time diagram of each node when aircraft transmits position message of Article i and Article i +1

附录 B

根据函数全微分的定理可知：

$$f(x+\Delta x,y+\Delta y)-f(x,y)=\frac{\partial f}{\partial x}\Delta x+\frac{\partial f}{\partial y}\Delta y+o(\rho)(\rho\to 0) \tag{B1}$$

式中：$f(x,y)$为关于变量x、y的函数；Δx和Δy分别为变量x和y的增量；$\rho=\sqrt{(\Delta x)^2+(\Delta y)^2}$。

根据式 (B1) 可得

$$\begin{cases}
d_A(\tilde{x}(i+1)+\Delta x(i+1),\tilde{y}(i+1)+\Delta y(i+1),\tilde{z}(i+1)+\Delta z(i+1))-d_A(\tilde{x}(i+1),\tilde{y}(i+1),\tilde{z}(i+1))\cong \\
\quad \dfrac{\partial d_A}{\partial x}\Big|_{\substack{x=x'_{i+1}\\y=y'_{i+1}\\z=z'_{i+1}}}\Delta x_{i+1}+\dfrac{\partial d_A}{\partial y}\Big|_{\substack{x=x'_{i+1}\\y=y'_{i+1}\\z=z'_{i+1}}}\Delta y_{i+1}+\dfrac{\partial d_A}{\partial z}\Big|_{\substack{x=x'_{i+1}\\y=y'_{i+1}\\z=z'_{i+1}}}\Delta z_{i+1} \\
d_A(\tilde{x}(i)+\Delta x(i+1),\tilde{y}(i)+\Delta y(i+1),\tilde{z}(i)+\Delta z(i+1))-d_A(\tilde{x}(i),\tilde{y}(i),\tilde{z}(i))\cong \\
\quad \dfrac{\partial d_A}{\partial x}\Big|_{\substack{x=x'_i\\y=y'_i\\z=z'_i}}\Delta x_i+\dfrac{\partial d_A}{\partial y}\Big|_{\substack{x=x'_i\\y=y'_i\\z=z'_i}}\Delta y_i+\dfrac{\partial d_A}{\partial z}\Big|_{\substack{x=x'_i\\y=y'_i\\z=z'_i}}\Delta z_i \\
d_B(\tilde{x}(i+1)+\Delta x(i+1),\tilde{y}(i+1)+\Delta y(i+1),\tilde{z}(i+1)+\Delta z(i+1))-d_B(\tilde{x}(i+1),\tilde{y}(i+1),\tilde{z}(i+1))\cong \\
\quad \dfrac{\partial d_B}{\partial x}\Big|_{\substack{x=x'_{i+1}\\y=y'_{i+1}\\z=z'_{i+1}}}\Delta x_{i+1}+\dfrac{\partial d_B}{\partial y}\Big|_{\substack{x=x'_{i+1}\\y=y'_{i+1}\\z=z'_{i+1}}}\Delta y_{i+1}+\dfrac{\partial d_B}{\partial z}\Big|_{\substack{x=x'_{i+1}\\y=y'_{i+1}\\z=z'_{i+1}}}\Delta z_{i+1} \\
d_B(\tilde{x}(i)+\Delta x(i+1),\tilde{y}(i)+\Delta y(i+1),\tilde{z}(i)+\Delta z(i+1))-d_B(\tilde{x}(i),\tilde{y}(i),\tilde{z}(i))\cong \\
\quad \dfrac{\partial d_B}{\partial x}\Big|_{\substack{x=x'_i\\y=y'_i\\z=z'_i}}\Delta x_i+\dfrac{\partial d_B}{\partial y}\Big|_{\substack{x=x'_i\\y=y'_i\\z=z'_i}}\Delta y_i+\dfrac{\partial d_B}{\partial z}\Big|_{\substack{x=x'_i\\y=y'_i\\z=z'_i}}\Delta z_i
\end{cases} \tag{B2}$$

将式 (B2) 代入式 (14), J表示为

$$\begin{aligned}J \cong & (b_{x,i+1}\Delta x(i+1)+b_{y,i+1}\Delta y(i+1)+b_{z,i+1}\Delta z(i+1))/c- \\ & (b_{x,i}\Delta x(i+1)+b_{y,i}\Delta y(i+1)+b_{z,i}\Delta z(i+1))/c- \\ & (a_{x,i+1}\Delta x(i+1)+a_{y,i+1}\Delta y(i+1)+a_{z,i+1}\Delta z(i+1))/c- \\ & (a_{x,i}\Delta x(i+1)+a_{y,i}\Delta y(i+1)+a_{z,i}\Delta z(i+1))/c+ \\ & \varepsilon_B(i+1)-\varepsilon_B(i)-\varepsilon_A(i+1)+\varepsilon_A(i)\end{aligned}$$

式中: $a_{m,n}$为$d_A(x,y,z)$在点$[x(n),y(n),z(n)]$处对m的偏导数; $b_{m,n}$为$d_B(x,y,z)$在点$[x(n),y(n),z(n)]$处对m的偏导数。

即得到式 (15)。

附录 C

由式 (15) 可知, J的分布特性由时间测量误差和 ADS-B 位置误差决定。

当地面站采用 GPS 授时系统对接收时刻进行标定时, 时间测量误差$\varepsilon_A(i)$和$\varepsilon_B(i)$均服从均值为 0、标准差为σ_{t}的高斯分布[25]。当航空器生成 ADS-B 位置报文时, 位置导航精度类别给出了估计的位置不确定性, 将估计的位置不确定性转化为 ADS-B 位置误差, 则$\Delta x(i)$、$\Delta y(i)$均服从均值为 0、标准差为σ_{h}的高斯分布[26], $\Delta z(i)$服从均值为 0、标准差为σ_{v}的高斯分布[26]。

由于$\varepsilon_A(i)$、$\varepsilon_B(i)$、$\Delta x(i)$、$\Delta y(i)$和$\Delta z(i)$均服从高斯分布, J最终服从高斯分布:

$$J \sim N(\mu_J,\sigma_J^2) \tag{C1}$$

式中:

$$\mu_J = 0$$

$$\begin{aligned}\sigma_J^2=&\frac{\sigma_{\mathrm{h}}^2}{c^2}(b_{x,2}^2+b_{y,2}^2-b_{x,1}^2-b_{y,1}^2-a_{x,2}^2-a_{y,2}^2+a_{x,1}^2+a_{y,1}^2)+ \\ & \frac{\sigma_{\mathrm{v}}^2}{c^2}(b_{z,2}^2-b_{z,1}^2-a_{z,2}^2+a_{z,1}^2)+4\sigma_{\mathrm{t}}^2\end{aligned}$$

其中: σ_{h}、σ_{v}、σ_{t}分别为 ADS-B 水平位置误差、ADS-B 垂直位置误差、时间测量误差的标准差。

Verification scheme of position message for automatic dependent surveillance-broadcast system based on interval difference of arrival

LIU Haitao*, LIU Jiaxiang, LI Dongxia, WANG Lei

(School of Electronic Information and Automation, Civil Aviation University of China, Tianjin 300300, China)

Abstract: Because of its precise aircraft position, quick update interval, and broad surveillance coverage, automatic dependent surveillance-broadcast (ADS-B) is a rapidly developing aeronautical surveillance system that has been extensively utilized in civil aviation surveillance. Nevertheless, since ADS-B communications are transmitted without cryptographic security protections, ADS-B is susceptible to spoofing sources. To detect spoofing, a verification scheme of ADS-B position messagesbased on interval difference of arrival (IDOA) is proposed. Firstly, the system model of ADS-B position message verification based on IDOA is presented.Then, the expression of the test statistic for ADS-B position message verification is derived. In addition, the detection threshold is given by the derived statistical property of the test statistic. Finally, computer simulation is used to confirm the accuracy and efficacy of the suggested strategy.The simulation results show that the probability of detection and probability of false of the IDOA scheme is equivalent to that of the time difference of arrival (TDOA) scheme,and IDOA is not sensitive to time measurement error and ADS-B position error. However, the benefit of IDOA is that it is not affected by time synchronization errors.

Keywords: automatic dependent surveillance-broadcast; spoofing interference; interval difference of arrival; detection threshold; time synchronization error of ground station

Received: 2022-08-17; **Accepted**: 2022-09-16; **Published Online**: 2022-11-14 14:52

URL: link.cnki.net/urlid/11.2625.V.20221111.1711.002

Foundation items: National Key Research and Development Program of China (2016YFB0502402); National Natural Science Foundation of China (U1733120); Multiple Input Foundation Youth Project of Tianjin in China (21JCQNJC00770)

* **Corresponding author**. E-mail: htliucauc@qq.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0734

新型 ADETCS 下工业信息物理系统综合安全控制

李炜[1, 2, 3, *]，程雪[1, 2, 3]，李亚洁[1, 2, 3]

(1. 兰州理工大学 电气工程与信息工程学院，兰州 730050；2. 甘肃省工业过程先进控制重点实验室，兰州 730050；
3. 兰州理工大学 电气与控制工程国家级实验教学示范中心，兰州 730050)

摘 要：针对一类虚假数据注入 (FDI) 攻击与故障共存的工业信息物理系统 (ICPS)，在新型自适应离散事件触发通信机制 (ADETCS) 下，研究了综合安全控制与通信的协同设计问题。引入 ADETCS 取代传统离散事件触发通信机制，构建安全与通信可自适应协同优化运行的 ICPS 架构；通过对 ADETCS 中阈值函数的优化设计，获得一种可随系统动态行为双向连续变化的新型 ADETCS；在该机制下，基于主动容侵和容错的思想，通过构造 Lyapunov 泛函，借助新型 Bessel-Legebdre 不等式、锥补线性化 (CCL) 等少保守性技术，在统一的自适应变采样框架下，推证出鲁棒估计器与综合安全控制器的求解方法，使 ICPS 可同时防御 FDI 攻击与故障，并具有更优的自适应通信调节能力。仿真结果表明：新型 ADETCS 中阈值函数的巧妙设计及 CCL 的应用，相较现存通信机制和设计方法，可使 ICPS 的综合安全控制性能与资源节约获得更优的折中平衡。

关 键 词：新型自适应离散事件触发通信机制；工业信息物理系统；综合安全控制；锥补线性化；动态优化平衡

中图分类号：TP277

文献标志码：A **文章编号**：1001-5965(2024)09-2704-13

信息物理系统 (cyber-physical system,CPS) 是计算单元和物理对象在网络环境中高度集成交互而成的智能系统[1]。随着控制、通信、计算机学科的快速发展，在工业 4.0 和智能制造的背景下，工业信息物理系统 (industry cyber-physical system,ICPS) 以其独特的优势蓬勃兴起。然而，ICPS 在快速推进的同时，有限网络带宽与海量数据传输之间的冲突，始终是 ICPS 优质发展的一大桎梏。针对该问题，学者们在周期时间触发通信机制的基础上，根据系统动态行为与性能需求，通过引入各种触发条件[2-4]，形成了不同的通信机制。其中，离散事件触发通信机制 (discrete event-triggered communication scheme, DETCS)，因能通过判断“事件”发生与否来实现对数据的筛选，在节约网络通信资源的同时，使系统控制性能与网络服务质量的协同设计成为可能，而备受关注。

尽管在 DETCS 下，ICPS 的控制性能与通信资源节约已得到较好的平衡折中，并取得了不少成果，但 DETCS 中静态不变的触发参数仍无法达到人们对通信节约的更高期许。因而，近年来学者们在 DETCS 的基础上，通过将静态不变的触发参数更换为动态可变的触发参数，继而又提出了自适应离散事件触发通信机制 (adaptive discrete event-triggered communication scheme, ADETCS)，以期提升 ICPS 的性能与资源占用间动态平衡的水平。虽然现有的 ADETCS[5-11] 各有特点，但触发参数能依据系统行为动态变化的本质是相同的，此类通信机制的确缓解了 DETCS 适应性不足的问题，但仍有可改进空间。因此，如何通过进一步改善触发参数对系统行为的适应度，给出一种新型的 ADETCS，

收稿日期：2022-08-20；**录用日期**：2022-12-30；**网络出版时间**：2023-01-30 15:16
网络出版地址：link.cnki.net/urlid/11.2625.V.20230130.1249.002
基金项目：国家自然科学基金 (62163022)
* **通信作者**. E-mail：liwei@lut.edu.cn

引用格式：李炜，程雪，李亚洁. 新型 ADETCS 下工业信息物理系统综合安全控制 [J]. 北京航空航天大学学报，2024，50（9）：2704-2716.
LI W，CHENG X，LI Y J. Integrated security control of industrial cyber-physical systems based on new type ADETCS[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50（9）：2704-2716（in Chinese）.

提高 ICPS 控制性能与通信资源节约之间的自适应折中平衡水平, 是本文的研究动机之一。

ICPS 是信息系统与物理系统高度融合的产物, 因此, 深受网络安全与物理安全的双层困扰。在网络安全层面, ICPS 主要面临虚假数据注入(false data injection, FDI)、拒绝服务 (denial of service, DOS)、数据重放等攻击的威胁, 其中, FDI 由于高时空隐蔽性特点, 更加难以识别防范, 也更受业界关注。仅从控制视角而言, 解决该问题的主要方法是对攻击进行主动或被动容侵[12-14]。在物理安全层面, ICPS 主要面临传感器、执行器等器件故障问题, 应对策略是对故障的被动、主动及主被动混合容错等[15-17]。综合已有的研究成果不难看出, 可共同防御 FDI 攻击和物理故障的成果还非常有限, 尽管文献 [18-20] 在 DETCS 下实现了攻击与故障的双重安全控制, 但在 ADETCS 下还少有涉足。因此, 如何在新型 ADETCS 下, 实现攻击主动容侵、故障主动容错与通信的协同设计, 是本文的又一研究动机。

此外, 无论是基于 DETCS 还是 ADETCS, 就数据的传输时序间隔和对象的属性而言, 可将 ICPS 等价为非均匀数据采样系统。而研究该类问题的优选方法即是将其转换为时延系统, 借助时滞系统理论[21] 开展研究。因此, 如何设计出少保守性的估计器与综合安全控制器, 进一步提高解的优性, 也是本文研究动机所在。

综上所述, 本文针对 FDI 攻击与故障共存的 ICPS, 在 ADETCS 下研究综合安全控制与通信的协同设计问题。首先, 从通信机制的改善着眼, 以 ADETCS 替代传统 DETCS, 构建集主动容侵和容错一体的双重安全架构; 其次, 针对现存 ADETCS 中触发参数 $\sigma(t_k h)$ 只可单向连续增大或以固定值离散增减等不足, 将正态型函数引入阈值函数 λ 的改良设计, 得到一种事件触发参数可随系统行为自适应连续双向变化的新型 ADETCS; 然后, 在该通信机制对应下的统一估计与控制的非均匀自适应数据传输周期, 并借助时延系统理论、新型 Bessel-Legebdre 不等式等技术, 推证出鲁棒估计器与综合安全控制与通信的协同设计方法; 最后, 采用锥补线性化 (cone complementarity linearization,CCL) 算法取代变量替换法, 减少非线性矩阵不等式(nonlinear matrix inequality, NLMI)求解的保守性, 以期使 ICPS 可自适应地兼顾系统安全与通信性能, 实现二者更优协同运行的目标。

1　系统建模

1.1　新型 ADETCS 下 ICPS 综合控制架构

考虑实际中一类 FDI 攻击与执行器故障共存的 ICPS, 在确保系统安全运行的前提下, 为进一步改善系统性能、节约网络资源, 通过引入自适应离散事件发生器, 可建立 ADETCS 下 ICPS 综合安全控制架构, 如图 1 所示。该系统组件主要包括被控对象、智能传感单元 (传感器、采样器、自适应离散事件发生器)、控制单元 (估计器、综合安全控制器)、零阶保持器、执行器及双端网络。

注 1　与文献 [20] 不同的是, 图 1 架构中将原有的离散事件发生器更换为一种自适应离散事件发生器, 使得原本触发参数不变的静态 DETCS, 被事件触发参数可随系统行为自适应变化的动态 ADETCS 所取代。由于 ADETCS 在平衡系统控制性能与节约通信资源方面的天然优势, 在该通信机制下进行 ICPS 综合安全估计、综合安全控制与通

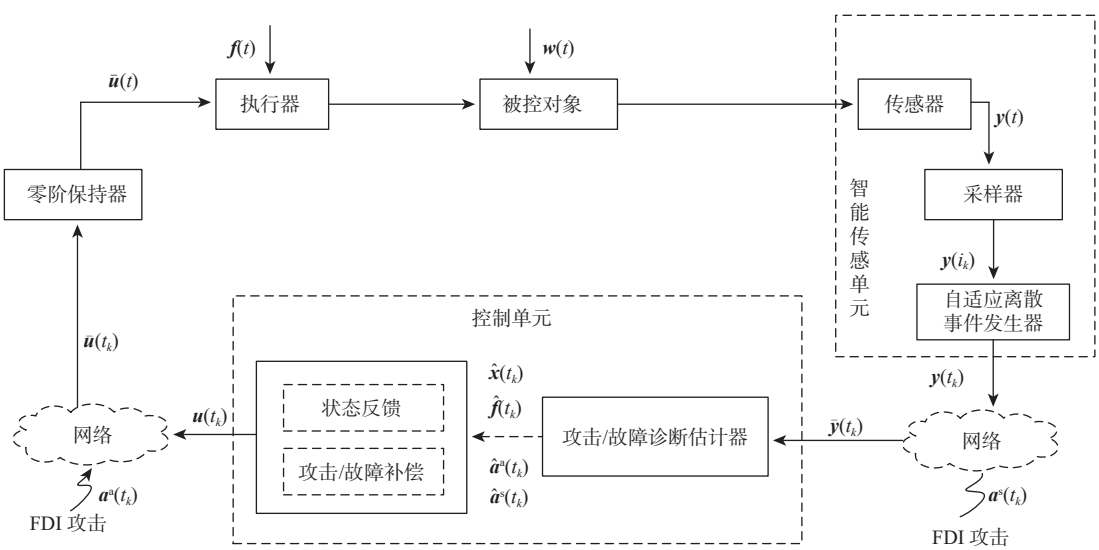


图 1　新型 ADETCS 下 ICPS 综合安全控制架构

Fig. 1　Integrated security control framework of ICPS in new type ADETCS

信协同设计，有望达到更优的折中平衡目标。

1.2 新型 ADETCS 的提出

受文献 [4]DETCS 设计的启发，提出如下自适应事件触发条件，用于判断采样数据是否传输：

$$\boldsymbol{e}(i_kh)^{\mathrm{T}}\boldsymbol{\phi}\boldsymbol{e}(i_kh)\geqslant\sigma(t_kh)\boldsymbol{y}(i_kh)^{\mathrm{T}}\boldsymbol{\phi}\boldsymbol{y}(i_kh) \tag{1}$$

式中：$\boldsymbol{e}(i_kh)=\boldsymbol{y}(i_kh)-\boldsymbol{y}(t_kh)$；$\boldsymbol{\phi}$为需设计的事件触发权矩阵；$\boldsymbol{y}(i_kh)$为当前时刻系统输出的量测采样值；$\boldsymbol{y}(t_kh)$为当前时刻满足 ADETCS 的系统输出采样值；$\sigma(t_kh)$为所需设计的自适应事件触发参数。

假设传感器由时钟驱动，采样周期为h，传输周期$T_k=t_{k-1}h-t_kh$。分析$\boldsymbol{y}(i_kh)$、$\boldsymbol{y}(t_kh)$、$\boldsymbol{y}(t_{k+1}h)$三者对应的时序序列，则i_kh与t_kh满足如下关系：

$$i_kh=t_kh+lh\qquad l\in 1,2,\cdots,j \tag{2}$$

式中：$j=\min\{j|t_kh+(j+1)h\geqslant t_{k+1}h\}$；$i_kh$为采样器采样时刻；$t_kh$为数据成功触发时刻。

因此，在 ADETCS 下的数据将按照如下方式传输：如果触发条件 (1) 满足，则将$\boldsymbol{y}(i_kh)$传输到控制单元，反之，则丢弃。

可以看出，ADETCS 中，触发条件的设定与 DETCS 类似，均是通过判定系统当前采样数据与上一个发送数据的差异是否超过预先设定的阈值，从而决定当前采样数据是否需要被传输，其自适应性的体现是将触发参数由原来的固定值，更替为随系统行为变化的阈值函数$\sigma(t_kh)$。从现存的阈值函数的变化来看，主要有 2 种方式：以某种非线性特性连续变化，或以某值离散增减。其中，较为典型的为文献 [5,7] 中的 ADETCS。下文以此为例，介绍二者阈值函数的设计机理，并分析问题所在，进而进行有针对性的改进，提出本文的新型 ADETCS。

在自适应阈值函数的连续选取方式中，文献 [5] 用反正切函数设计的自适应触发参数是代表之一：

$$\sigma(t_kh)=\max\{\sigma_{\min},\lambda\sigma(t_{k-1}h)\} \tag{3}$$

$$\lambda=\begin{cases}0 & \|\boldsymbol{y}(t_kh)\|\geqslant\|\boldsymbol{y}(t_{k-1}h)\|\\ 1-\dfrac{2\alpha}{\pi}\arctan\left(\dfrac{\|\boldsymbol{y}(t_kh)\|-\|\boldsymbol{y}(t_{k-1}h)\|}{\|\boldsymbol{y}(t_kh)\|}\right) & \text{其他}\end{cases} \tag{4}$$

式中：$\sigma(t_kh)$的下界为$\sigma_{\min}$，文献 [5] 未对其上界进行界定，即系统最大时延不确定，当相邻 2 次触发时刻发送数据的$\|\boldsymbol{y}(t_kh)\|$和$\|\boldsymbol{y}(t_{k-1}h)\|$差值大于或等于零时，$\lambda=0$，此时，$\sigma(t_kh)=\sigma_{\min}$，退化为 DETCS，只有当$\|\boldsymbol{y}(t_kh)\|$和$\|\boldsymbol{y}(t_{k-1}h)\|$差值小于零，且$\alpha\neq 1$时，系统才可在 ADETCS 下运行。可以看出，尽管阈值函数$\sigma(t_kh)$是随系统行为动态变化的，但当系统输出误差较大时，为保证系统性能，需要较小的$\sigma(t_kh)$，然而根据式 (4)，$\sigma(t_kh)$只可随λ单向连续增大，难以满足大误差、小触发参数的需求。

在自适应阈值函数的离散选取方式中，文献 [7] 是较有代表性的一种，为便于与本文中的新型 ADETCS 进行比较，将其时序与变量作实质等价调整，文献 [7] 的 ADETCS 可表示为

$$\sigma(t_kh)=\sigma(t_{k-1}h)+\lambda \tag{5}$$

$$\lambda=\begin{cases}-0.1 & \boldsymbol{\eta}^{\mathrm{T}}(t_kh)\boldsymbol{\eta}(t_kh)>\rho\\ 0 & \boldsymbol{\eta}^{\mathrm{T}}(t_kh)\boldsymbol{\eta}(t_kh)=\rho\\ 0.1 & \boldsymbol{\eta}^{\mathrm{T}}(t_kh)\boldsymbol{\eta}(t_kh)<\rho\end{cases} \tag{6}$$

式中：$\sigma(t_kh)\in[\sigma_{\min},\sigma_{\max}]$，$\sigma_{\max}$、$\sigma_{\min}$分别为自适应触发参数$\sigma(t_kh)$变化的上、下界；$\rho$为正数。将文献 [7] 中$\boldsymbol{\eta}(t_kh)$的时序与本文进行对等转化，可得到：$\boldsymbol{\eta}(t_kh)=\boldsymbol{y}(t_kh)-\boldsymbol{y}(t_{k-1}h)$，当$\boldsymbol{\eta}^{\mathrm{T}}(t_kh)\boldsymbol{\eta}(t_kh)$分别小于、大于、等于$\rho$时，其阈值函数$\lambda$随之双向改变为 0.1、−0.1、0，即$\sigma(t_kh)$递减或递增或不变。

可以看出，尽管触发函数$\sigma(t_kh)$是随系统行为动态变化的，且计算简便，但自适应触发参数随λ是以固定增减值离散变化的，其对系统行为依赖的精细度、平滑性均显不足。

综合文献 [5,7] 的优点，本文对阈值函数进行改进，使得$\sigma(t_kh)$可在上下界内随系统行为变化，具有双向且连续自适应性。$\sigma(t_kh)$满足如下条件：

$$\sigma(t_kh)=\max\{\sigma_{\min},\lambda\sigma(t_{k-1}h)\}$$
$$\sigma(t_kh)\in[\sigma_{\min},\sigma_{\max}] \tag{7}$$

$$\lambda=\alpha+\mathrm{e}^{-\beta z^2}\qquad 0<\alpha<1,\beta>0 \tag{8}$$

$$z(t_kh)=\frac{\|\boldsymbol{y}(t_kh)\|-\|\boldsymbol{y}(t_{k-1}h)\|}{\|\boldsymbol{y}(t_kh)\|} \tag{9}$$

式中：$\sigma_{\max}$、$\sigma_{\min}$分别为$\sigma(t_kh)$的上、下界，取值范围为$0\leqslant\sigma_{\min}<\sigma_{\max}<1$，$\sigma_{\max}$对应确立系统最大时延上界$h_1$。

式 (8) 中，当$\alpha\in(0,1)$时，$\lambda\in(0,2)$。当$\|\boldsymbol{y}(t_kh)\|$与$\|\boldsymbol{y}(t_{k-1}h)\|$之间差异较大时，$0<\lambda<1$，使得$\sigma(t_kh)<\sigma(t_{k-1}h)$，系统传输数据增加；当$\|\boldsymbol{y}(t_kh)\|$与$\|\boldsymbol{y}(t_{k-1}h)\|$之间差异较小时，$1\leqslant\lambda<2$，使得$\sigma(t_kh)>\sigma(t_{k-1}h)$，系统传输数据减少；当$\|\boldsymbol{y}(t_kh)\|=\|\boldsymbol{y}(t_{k-1}h)\|$，即系统处于平稳状态时，$\lambda=\alpha$，由于$0<\alpha<1$，$0<\lambda<1$，此时$\sigma(t_kh)=\sigma_{\min}$保持不变。由此可见，由于$\lambda$的巧妙设计，使阈值函数$\sigma(t_kh)$的选取由离散、单向转化为连续、双向变化，同时还可在平衡点处收敛于某一常值。

注 2 本文设计的新型 ADETCS，不仅考虑了自适应触发参数的上下界，由于λ的特殊设计，触发参数可随系统行为自适应双向且连续变化，从而使数据的传输也可适应系统行为，更有针对性地调整事件触发参数，因此，有望达到节约更多网络资源和更优性能的目的。

1.3　新型 ADETCS 下 ICPS 的描述

针对具有双侧 FDI 攻击、执行器未知连续时变故障[22-23]及外部扰动的线性 ICPS, 可表示如下:

$$\begin{cases}\dot{\boldsymbol{x}}(t)=\boldsymbol{A}\boldsymbol{x}(t)+\boldsymbol{B}\overline{\boldsymbol{u}}(t)+\boldsymbol{E}_f\boldsymbol{f}(t)+\boldsymbol{E}_w\boldsymbol{w}(t)\\ \boldsymbol{y}(i_kh)=\boldsymbol{C}\boldsymbol{x}(i_kh)+\boldsymbol{E}_v\boldsymbol{v}(i_kh)\end{cases}\tag{10}$$

式中: $\boldsymbol{x}(t)\in\mathbf{R}^n$为状态向量; $\overline{\boldsymbol{u}}(t)\in\mathbf{R}^{n_u}$为执行器侧受到 FDI 攻击的输入向量; $\boldsymbol{f}(t)\in\mathbf{R}^{n_f}$、$\boldsymbol{w}(t)\in\mathbf{R}^{n_w}$分别为执行器故障、外部扰动; $\boldsymbol{y}(i_kh)\in\mathbf{R}^{n_f}$、$\boldsymbol{v}(i_kh)\in\mathbf{R}^{n_v}$分别为采样输出向量、传感器量测噪声; $\boldsymbol{A}$、$\boldsymbol{B}$、$\boldsymbol{E}_f$、$\boldsymbol{E}_w$、$\boldsymbol{C}$、$\boldsymbol{E}_v$为已知适维矩阵。

ICPS 传感器侧、执行器侧网络攻击建模为

$$\begin{cases}\overline{\boldsymbol{u}}(t_kh)=\boldsymbol{u}(t_kh)+\boldsymbol{E}_{\mathrm{a}}\boldsymbol{a}^{\mathrm{a}}(t_kh)\\ \overline{\boldsymbol{y}}(t_kh)=\boldsymbol{y}(t_kh)+\boldsymbol{E}_{\mathrm{s}}\boldsymbol{a}^{\mathrm{s}}(t_kh)\end{cases}\tag{11}$$

式中: $\boldsymbol{u}(t_kh)$为执行侧控制向量; $\overline{\boldsymbol{y}}(t_kh)$为传感侧输出向量; $\boldsymbol{a}^{\mathrm{a}}(t_kh)\in\mathbf{R}^{n_{\mathrm{a}}}$、$\boldsymbol{a}^{\mathrm{s}}(t_kh)\in\mathbf{R}^{n_{\mathrm{s}}}$分别为执行侧、传感侧受到的 FDI 攻击, $\boldsymbol{a}^{\mathrm{a}}(t_kh)=\boldsymbol{a}^{\mathrm{s}}(t_kh)$; $\boldsymbol{E}_{\mathrm{a}}$、$\boldsymbol{E}_{\mathrm{s}}$为已知适维矩阵 。

ADETCS 的引入, 经过自适应离散事件发生器筛选后的数据, 仍以非均匀周期方式传输, 而被控对象是连续的, 因此, 本文仍可基于时滞系统理论进行研究。另外, 图 1 给出的 ICPS 框架中, 由于估计器和综合安全控制器均处于双端网络之间的控制单元, 若不考虑现代硬性体系下的网络传输和计算时延, 估计器和综合安全控制器在 ADETCS 下有相同的传输间隔, 因而可构造如下统一的时延函数:

$$\tau(t)=t-t_kh\qquad t\in[t_kh,t_{k+1}h)\tag{12}$$

式中: 时延函数 $0\leqslant\tau(t)<h_1$, h_1为由$\sigma_{\max}$确定的最大时延上界。

2　新型 ADETCS 下鲁棒估计器的设计

为同时估计传感侧攻击 $\boldsymbol{a}^{\mathrm{s}}(t_kh)$、执行侧攻击 $\boldsymbol{a}^{\mathrm{a}}(t_kh)$及执行器故障$\boldsymbol{f}(t)$, 将三者增广为一个广义故障向量$\overline{\boldsymbol{f}}(t)$。由式 (10)～式 (12) 可得增广后的系统状态方程为

$$\begin{cases}\dot{\boldsymbol{x}}(t)=\boldsymbol{A}\boldsymbol{x}(t)+\boldsymbol{B}\overline{\boldsymbol{u}}(t)+\overline{\boldsymbol{E}}_1\overline{\boldsymbol{f}}(t)+\boldsymbol{E}_w\boldsymbol{w}(t)\\ \overline{\boldsymbol{y}}(t)=\boldsymbol{C}\boldsymbol{x}(t-\tau(t))+\boldsymbol{E}_v\boldsymbol{v}(t-\tau(t))+\overline{\boldsymbol{E}}_2\overline{\boldsymbol{f}}(t)\end{cases}\tag{13}$$

式中: $\overline{\boldsymbol{f}}(t)=\left[\boldsymbol{f}^{\mathrm{T}}(t)\quad(\boldsymbol{a}^{\mathrm{a}}(t-\tau(t)))^{\mathrm{T}}\quad(\boldsymbol{a}^{\mathrm{s}}(t-\tau(t)))^{\mathrm{T}}\right]^{\mathrm{T}}$; $\overline{\boldsymbol{E}}_1=\left[\boldsymbol{E}_f\quad\boldsymbol{B}\boldsymbol{E}_{\mathrm{a}}\quad\boldsymbol{0}\right]$; $\overline{\boldsymbol{E}}_2=\left[\boldsymbol{0}\quad\boldsymbol{0}\quad\boldsymbol{E}_{\mathrm{s}}\right]$。

攻击/故障诊断估计器建模为

$$\begin{cases}\dot{\hat{\boldsymbol{x}}}(t)=\boldsymbol{A}\hat{\boldsymbol{x}}(t)+\boldsymbol{B}\boldsymbol{u}(t)+\overline{\boldsymbol{E}}_1\hat{\overline{\boldsymbol{f}}}(t)-\boldsymbol{L}[\hat{\overline{\boldsymbol{y}}}(t)-\overline{\boldsymbol{y}}(t)]\\ \dot{\hat{\overline{\boldsymbol{f}}}}(t)=-\boldsymbol{F}\left[\hat{\overline{\boldsymbol{y}}}(t)-\overline{\boldsymbol{y}}(t)\right]\\ \hat{\overline{\boldsymbol{y}}}(t)=\boldsymbol{C}\hat{\boldsymbol{x}}(t)+\overline{\boldsymbol{E}}_2\hat{\overline{\boldsymbol{f}}}(t)\end{cases}\tag{14}$$

式中: $\hat{\boldsymbol{x}}(t)$为状态向量的估计量; $\boldsymbol{L}$、$\boldsymbol{F}$分别为待设计状态估计器增益矩阵、攻击/故障增益矩阵。

定义 1　$\boldsymbol{e}_x(t)=\hat{\boldsymbol{x}}(t)-\boldsymbol{x}(t)$, $\boldsymbol{e}_{\overline{f}}(t)=\hat{\overline{\boldsymbol{f}}}(t)-\boldsymbol{f}(t)$, $\boldsymbol{e}_{\overline{y}}(t)=\hat{\overline{\boldsymbol{y}}}(t)-\overline{\boldsymbol{y}}(t)$, 由式 (13)、式 (14) 可得误差系统为

$$\begin{cases}\dot{\boldsymbol{e}}_x(t)=\boldsymbol{A}\boldsymbol{e}_x(t)+\left(\overline{\boldsymbol{E}}_1-\boldsymbol{L}\overline{\boldsymbol{E}}_2\right)\boldsymbol{e}_{\overline{f}}(t)-\boldsymbol{E}_w\boldsymbol{w}(t)-\\ \qquad\boldsymbol{L}\boldsymbol{C}\boldsymbol{e}_x(t-\tau(t))+\boldsymbol{L}\boldsymbol{E}_v\boldsymbol{v}(t-\tau(t))\\ \dot{\boldsymbol{e}}_{\overline{f}}(t)=-\boldsymbol{F}\boldsymbol{C}\boldsymbol{e}_x(t-\tau(t))+\boldsymbol{F}\boldsymbol{E}_v\boldsymbol{v}(t-\tau(t))-\\ \qquad\boldsymbol{F}\boldsymbol{E}_2\boldsymbol{e}_{\overline{f}}(t)-\dot{\overline{\boldsymbol{f}}}(t)\end{cases}\tag{15}$$

为便于计算, 令$\overline{\boldsymbol{e}}(t)=\left[\boldsymbol{e}_x^{\mathrm{T}}(t)\quad\boldsymbol{e}_{\overline{f}}^{\mathrm{T}}(t)\right]^{\mathrm{T}}$, 结合式 (15) 可得如下增广误差系统:

$$\dot{\overline{\boldsymbol{e}}}(t)=\overline{\boldsymbol{A}}\overline{\boldsymbol{e}}(t)-\overline{\boldsymbol{E}}_w\boldsymbol{w}(t)-\overline{\boldsymbol{B}}\boldsymbol{e}_x(t-\tau(t))+\overline{\boldsymbol{L}}\boldsymbol{E}_v\boldsymbol{v}(t-\tau(t))\tag{16}$$

式中: $\overline{\boldsymbol{A}}=\begin{bmatrix}\boldsymbol{A}&\boldsymbol{E}_f\\ \boldsymbol{0}&\boldsymbol{0}\end{bmatrix}$; $\overline{\boldsymbol{B}}=\overline{\boldsymbol{L}}\,\overline{\boldsymbol{C}}$, $\overline{\boldsymbol{L}}=\begin{bmatrix}\boldsymbol{L}\\ \boldsymbol{F}\end{bmatrix}$, $\overline{\boldsymbol{C}}=\begin{bmatrix}\boldsymbol{C}&\overline{\boldsymbol{E}}_2\end{bmatrix}$; $\overline{\boldsymbol{E}}_w=\begin{bmatrix}\boldsymbol{E}_w&\boldsymbol{0}\\ \boldsymbol{0}&\boldsymbol{I}\end{bmatrix}$。

定理 1　在 ADETCS 下, 对于给定的正数α、β、h_1、γ_1, 若存在正定对称矩阵 $\boldsymbol{P}$与适维矩阵 $\boldsymbol{X}$、$\boldsymbol{Y}$, 满足线性矩阵不等式 (17) 和 (18), 则存在状态增益矩阵 $\boldsymbol{L}$与增广故障增益矩阵 $\boldsymbol{F}$, 使得增广误差系统 (16) 在无扰动时渐近稳定; 在有扰动时, 满足$\|\overline{\boldsymbol{e}}(t)\|_2^2\leqslant\gamma_1^2\left[\|\overline{\boldsymbol{w}}(t)\|_2^2+\sum\limits_{k=0}^{+\infty}(t_{k+1}h-t_kh)\|\boldsymbol{v}(t_kh)\|_2^2\right]$的性能指标, $\overline{\boldsymbol{w}}(t)=\begin{bmatrix}\boldsymbol{w}(t)\\ \dot{\overline{\boldsymbol{f}}}(t)\end{bmatrix}$, $\boldsymbol{L}$、$\boldsymbol{F}$可由$\overline{\boldsymbol{L}}=\begin{bmatrix}\boldsymbol{L}\\ \boldsymbol{F}\end{bmatrix}=\boldsymbol{P}^{-1}\boldsymbol{Y}$求解。

$$\begin{bmatrix}\boldsymbol{I}_{11}&\boldsymbol{I}_{12}&\boldsymbol{I}_{13}&\boldsymbol{I}_{14}&\boldsymbol{I}_{15}&\boldsymbol{N}&\boldsymbol{I}_{17}&\boldsymbol{I}_{18}\\ &\boldsymbol{I}_{22}&\boldsymbol{I}_{23}&\boldsymbol{I}_{24}&\boldsymbol{0}&\boldsymbol{N}&\boldsymbol{0}&\boldsymbol{0}\\ &&\boldsymbol{I}_{33}&\boldsymbol{I}_{34}&\boldsymbol{I}_{35}&\boldsymbol{N}&\boldsymbol{0}&\boldsymbol{0}\\ &&&\boldsymbol{I}_{44}&\boldsymbol{I}_{45}&\boldsymbol{N}&\boldsymbol{0}&\boldsymbol{0}\\ &&&&\boldsymbol{I}_{55}&\boldsymbol{0}&\boldsymbol{I}_{57}&\boldsymbol{I}_{58}\\ &&&&&\boldsymbol{I}_{66}&\boldsymbol{0}&\boldsymbol{0}\\ &&&&&&\boldsymbol{I}_{77}&\boldsymbol{0}\\ &&&&&&&\boldsymbol{I}_{88}\end{bmatrix}<\boldsymbol{0}\tag{17}$$

$$\begin{cases}\begin{bmatrix}\boldsymbol{J}_1&\boldsymbol{J}_2\\ &\boldsymbol{J}_4\end{bmatrix}<\boldsymbol{0}\\ \boldsymbol{J}_1=\begin{bmatrix}\boldsymbol{J}_{11}&\boldsymbol{J}_{12}&\boldsymbol{J}_{13}&\boldsymbol{J}_{14}&\boldsymbol{J}_{15}\\ &\boldsymbol{J}_{22}&\boldsymbol{J}_{23}&\boldsymbol{J}_{24}&\boldsymbol{0}\\ &&\boldsymbol{J}_{33}&\boldsymbol{J}_{34}&\boldsymbol{0}\\ &&&\boldsymbol{J}_{44}&\boldsymbol{0}\\ &&&&\boldsymbol{J}_{55}\end{bmatrix}\\ \boldsymbol{J}_2=\begin{bmatrix}\boldsymbol{J}_{16}&\boldsymbol{J}_{17}&\boldsymbol{N}&\boldsymbol{J}_{19}&\boldsymbol{J}_{20}\\ \boldsymbol{0}&\boldsymbol{0}&\boldsymbol{N}&\boldsymbol{0}&\boldsymbol{0}\\ \boldsymbol{0}&\boldsymbol{0}&\boldsymbol{N}&\boldsymbol{0}&\boldsymbol{0}\\ \boldsymbol{0}&\boldsymbol{0}&\boldsymbol{N}&\boldsymbol{0}&\boldsymbol{0}\\ \boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{J}_{49}&\boldsymbol{J}_{50}\end{bmatrix}\\ \boldsymbol{J}_4=\begin{bmatrix}-\gamma_1^2\boldsymbol{I}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{J}_{69}&\boldsymbol{J}_{60}\\ &-\gamma_1^2\boldsymbol{I}&\boldsymbol{0}&\boldsymbol{J}_{79}&\boldsymbol{J}_{70}\\ &&\boldsymbol{J}_{88}&\boldsymbol{0}&\boldsymbol{0}\\ &&&\boldsymbol{J}_{99}&\boldsymbol{0}\\ &&&&\boldsymbol{J}_{100}\end{bmatrix}\end{cases}\tag{18}$$

式中: $\boldsymbol{N}$为单位向量; $\boldsymbol{I}_{11}=\boldsymbol{P}\overline{\boldsymbol{A}}+\overline{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{P}+\boldsymbol{Q}-\dfrac{\pi^2}{4}\boldsymbol{R}+3\boldsymbol{N}-3\boldsymbol{N}^{\mathrm{T}}$, $\boldsymbol{Q}$、$\boldsymbol{R}$为正定对称矩阵, $\boldsymbol{I}_{12}=3\boldsymbol{N}-\boldsymbol{N}^{\mathrm{T}}$, $\boldsymbol{I}_{13}=3\boldsymbol{N}-8\boldsymbol{N}^{\mathrm{T}}$, $\boldsymbol{I}_{14}=3\boldsymbol{N}+12\boldsymbol{N}^{\mathrm{T}}$, $\boldsymbol{I}_{15}=\dfrac{\pi^2}{4}\boldsymbol{R}$, $\boldsymbol{I}_{17}=h_1\overline{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{P}$, $\boldsymbol{I}_{18}=h_1^{\frac{1}{2}}\overline{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{P}$, $\boldsymbol{I}_{22}=-\boldsymbol{Q}-\boldsymbol{N}-\boldsymbol{N}^{\mathrm{T}}$, $\boldsymbol{I}_{23}=-\boldsymbol{N}-8\boldsymbol{N}^{\mathrm{T}}$, $\boldsymbol{I}_{24}=$

$-\boldsymbol{N}+12\boldsymbol{N}^{\mathrm{T}}$, $\boldsymbol{I}_{33}=-8\boldsymbol{N}-\boldsymbol{N}^{\mathrm{T}}$, $\boldsymbol{I}_{34}=-8\boldsymbol{N}+12\boldsymbol{N}^{\mathrm{T}}$, $\boldsymbol{I}_{35}=\boldsymbol{I}_{45}=\boldsymbol{0}$, $\boldsymbol{I}_{44}=12\boldsymbol{N}+12\boldsymbol{N}^{\mathrm{T}}$, $\boldsymbol{I}_{55}=-\dfrac{\pi^2}{4}\boldsymbol{R}$, $\boldsymbol{I}_{57}=-h_1\left(\boldsymbol{Y}\overline{\boldsymbol{A}}\right)^{\mathrm{T}}$, $\boldsymbol{I}_{58}=-h_1^{\frac{1}{2}}\left(\boldsymbol{Y}\overline{\boldsymbol{C}}\right)^{\mathrm{T}}$, $\boldsymbol{I}_{66}=-\dfrac{1}{9}h_1\boldsymbol{P}$, $\boldsymbol{I}_{77}=-\boldsymbol{P}\boldsymbol{R}^{-1}\boldsymbol{P}$, $\boldsymbol{I}_{88}=-\boldsymbol{P}\boldsymbol{Z}^{-1}\boldsymbol{P}$, $\boldsymbol{Z}$为正定对称矩阵, $\boldsymbol{J}_{11}=\boldsymbol{I}_{11}+\boldsymbol{I}$, $\boldsymbol{J}_{12}=\boldsymbol{I}_{12}$, $\boldsymbol{J}_{13}=\boldsymbol{I}_{13}$, $\boldsymbol{J}_{14}=\boldsymbol{I}_{14}$, $\boldsymbol{J}_{15}=\dfrac{\pi^2}{4}\boldsymbol{R}-\boldsymbol{Y}\overline{\boldsymbol{C}}$, $\boldsymbol{J}_{16}=\boldsymbol{Y}\boldsymbol{E}_v$, $\boldsymbol{J}_{17}=-\boldsymbol{P}\overline{\boldsymbol{E}}_w$, $\boldsymbol{J}_{19}=h_1^{\frac{1}{2}}\overline{\boldsymbol{A}}^{\mathrm{T}}\boldsymbol{P}$, $\boldsymbol{J}_{20}=-\boldsymbol{J}_{19}$, $\boldsymbol{J}_{22}=\boldsymbol{I}_{22}$, $\boldsymbol{J}_{23}=\boldsymbol{I}_{23}$, $\boldsymbol{J}_{24}=\boldsymbol{I}_{24}$, $\boldsymbol{J}_{33}=\boldsymbol{I}_{33}$, $\boldsymbol{J}_{34}=\boldsymbol{I}_{34}$, $\boldsymbol{J}_{44}=\boldsymbol{I}_{44}$, $\boldsymbol{J}_{55}=\boldsymbol{I}_{55}$, $\boldsymbol{J}_{49}=-h_1\left(\boldsymbol{Y}\overline{\boldsymbol{C}}\right)^{\mathrm{T}}$, $\boldsymbol{J}_{50}=-h_1^{\frac{1}{2}}\left(\boldsymbol{Y}\overline{\boldsymbol{C}}\right)^{\mathrm{T}}$, $\boldsymbol{J}_{69}=-h_1(\boldsymbol{Y}\boldsymbol{E}_v)^{\mathrm{T}}$, $\boldsymbol{J}_{60}=-h_1^{\frac{1}{2}}(\boldsymbol{Y}\boldsymbol{E}_v)^{\mathrm{T}}$, $\boldsymbol{J}_{79}=h_1\boldsymbol{Y}\boldsymbol{E}_v$, $\boldsymbol{J}_{70}=-h_1^{\frac{1}{2}}\overline{\boldsymbol{E}}_w^{\mathrm{T}}\boldsymbol{P}$, $\boldsymbol{J}_{88}=-\dfrac{1}{9}h_1\boldsymbol{P}$, $\boldsymbol{J}_{99}=-\boldsymbol{P}\boldsymbol{R}^{-1}\boldsymbol{P}$, $\boldsymbol{J}_{100}=-\boldsymbol{P}\boldsymbol{Z}^{-1}\boldsymbol{P}$。

证明 构造如下 Lyapunov-Krasovskii 泛函[5]:

$$V(t)=V_1(t)+V_2(t)+V_3(t)+V_4(t)+V_5(t) \tag{19}$$

式中:

$$\begin{cases} V_1(t)=\bar{\boldsymbol{e}}^{\mathrm{T}}(t)\boldsymbol{P}\bar{\boldsymbol{e}}(t) \\ V_2(t)=\displaystyle\int_{t-h_1}^{t}\bar{\boldsymbol{e}}^{\mathrm{T}}(s)\boldsymbol{Q}\bar{\boldsymbol{e}}(s)\mathrm{d}s \\ V_3(t)=h_1^2\displaystyle\int_{t-\tau(t)}^{t}\dot{\bar{\boldsymbol{e}}}^{\mathrm{T}}(s)\boldsymbol{R}\dot{\bar{\boldsymbol{e}}}(s)\mathrm{d}s \\ V_4(t)=-\dfrac{\pi^2}{4}\displaystyle\int_{t-\tau(t)}^{t}[\bar{\boldsymbol{e}}(s)-\bar{\boldsymbol{e}}(s-\tau(t))]^{\mathrm{T}}\cdot \\ \qquad \boldsymbol{R}[\bar{\boldsymbol{e}}(s)-\bar{\boldsymbol{e}}(s-\tau(t))]\mathrm{d}s \\ V_5(t)=\displaystyle\int_{-h_1}^{0}\int_{t+\theta}^{t}\dot{\bar{\boldsymbol{e}}}^{\mathrm{T}}(s)\boldsymbol{Z}\dot{\bar{\boldsymbol{e}}}(s)\mathrm{d}s\mathrm{d}\theta \end{cases}$$

其中: $\boldsymbol{P}$、$\boldsymbol{Q}$、$\boldsymbol{R}$、$\boldsymbol{Z}$为正定对称矩阵。

在初始状态为零且无外界扰动的情况下, 沿系统 (15) 对$V(t)$求导, 有

$$\dot{V}(t)=\dot{V}_1(t)+\dot{V}_2(t)+\dot{V}_3(t)+\dot{V}_4(t)+\dot{V}_5(t) \tag{20}$$

式中:

$$\begin{cases} \dot{V}_1(t)=2\bar{\boldsymbol{e}}^{\mathrm{T}}(t)\boldsymbol{P}\dot{\bar{\boldsymbol{e}}}(t) \\ \dot{V}_2(t)=\bar{\boldsymbol{e}}^{\mathrm{T}}(t)\boldsymbol{Q}\bar{\boldsymbol{e}}(t)-\bar{\boldsymbol{e}}^{\mathrm{T}}(t-h_1)\boldsymbol{Q}\bar{\boldsymbol{e}}(t-h_1) \\ \dot{V}_3(t)=h_1^2\dot{\bar{\boldsymbol{e}}}^{\mathrm{T}}(t)\boldsymbol{R}\dot{\bar{\boldsymbol{e}}}(t) \\ \dot{V}_4(t)=-\dfrac{\pi^2}{4}[\bar{\boldsymbol{e}}(t)-\bar{\boldsymbol{e}}(t-\tau(t)]^{\mathrm{T}}\boldsymbol{R}[\bar{\boldsymbol{e}}(t)-\bar{\boldsymbol{e}}(t-\tau(t)] \\ \dot{V}_5(t)=h_1\dot{\bar{\boldsymbol{e}}}^{\mathrm{T}}(t)\boldsymbol{Z}\dot{\bar{\boldsymbol{e}}}(t)-\displaystyle\int_{t+h_1}^{t}\dot{\bar{\boldsymbol{e}}}^{\mathrm{T}}(s)\boldsymbol{Z}\dot{\bar{\boldsymbol{e}}}(s)\mathrm{d}s \end{cases}$$

针对式 (20) 中的积分项$-\displaystyle\int_{t+h_1}^{t}\dot{\bar{\boldsymbol{e}}}^{\mathrm{T}}(s)\boldsymbol{Z}\dot{\bar{\boldsymbol{e}}}(s)\mathrm{d}s$, 采用新型 Bessel-Legebdre 不等式[24] 处理, 即

$$-\int_{t+h_1}^{t}\dot{\bar{\boldsymbol{e}}}^{\mathrm{T}}(s)\boldsymbol{Z}\dot{\bar{\boldsymbol{e}}}(s)\mathrm{d}s\leqslant\boldsymbol{\xi}^{\mathrm{T}}\left[\mathrm{Sym}\{\boldsymbol{\Pi}^{\mathrm{T}}\boldsymbol{N}\}+h_1\boldsymbol{N}\overline{\boldsymbol{Z}}\boldsymbol{N}\right]\boldsymbol{\xi} \tag{21}$$

式中: Sym( ) 表示 Bessel-Legebdre 不等式。

$$\boldsymbol{\xi}=[\bar{\boldsymbol{e}}(t)\quad\bar{\boldsymbol{e}}(t-h_1)\quad\gamma_1\quad\gamma_2\quad\bar{\boldsymbol{e}}(t-\tau(t))]^{\mathrm{T}}$$

式中: γ_1、γ_2为常数。

$$\overline{\boldsymbol{Z}}=\begin{bmatrix}\boldsymbol{Z}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{0}&3\boldsymbol{Z}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{0}&5\boldsymbol{Z}\end{bmatrix}$$

$$\boldsymbol{\Pi}=\begin{bmatrix}\boldsymbol{I}&-\boldsymbol{I}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{I}&\boldsymbol{I}&-2\boldsymbol{I}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{I}&-\boldsymbol{I}&-6\boldsymbol{I}&12\boldsymbol{I}&\boldsymbol{0}\end{bmatrix}$$

将式 (20) 和式 (21) 联立, 并记: $\boldsymbol{N}_{11}=[\boldsymbol{I}\ \boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{0}]$, $\boldsymbol{N}_{12}=[\overline{\boldsymbol{A}}\ \boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{0}\ -\overline{\boldsymbol{L}}\,\overline{\boldsymbol{C}}]$, $\boldsymbol{N}_{13}=[\boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{I}]$, $\boldsymbol{N}_{14}=[\boldsymbol{0}\ \boldsymbol{I}\ \boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{0}]$, 有: $\bar{\boldsymbol{e}}(t)=\boldsymbol{N}_{11}\boldsymbol{\xi}$, $\dot{\bar{\boldsymbol{e}}}(t)=\boldsymbol{N}_{12}\boldsymbol{\xi}$, $\bar{\boldsymbol{e}}(t-\tau(t))=\boldsymbol{N}_{13}\boldsymbol{\xi}$, $\bar{\boldsymbol{e}}(t-h_1)=\boldsymbol{N}_{14}\boldsymbol{\xi}$, 最终化简整理得

$$\dot{V}(t)\leqslant\boldsymbol{\xi}^{\mathrm{T}}\theta\boldsymbol{\xi} \tag{22}$$

式中:

$$\begin{aligned}\theta=&\boldsymbol{N}_{11}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{N}_{12}+\boldsymbol{N}_{12}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{N}_{11}+\boldsymbol{N}_{11}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{N}_{11}-\boldsymbol{N}_{14}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{N}_{14}-\\&\frac{\pi^2}{4}\left[\boldsymbol{N}_{11}^{\mathrm{T}}-\boldsymbol{N}_{13}^{\mathrm{T}}\right]\boldsymbol{R}[\boldsymbol{N}_{11}-\boldsymbol{N}_{13}]+\boldsymbol{\Pi}^{\mathrm{T}}\boldsymbol{N}+\boldsymbol{N}^{\mathrm{T}}\boldsymbol{\Pi}+\\&h_1{}^2\boldsymbol{N}_{12}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{N}_{12}+h_1(\boldsymbol{N}_{12}^{\mathrm{T}}\boldsymbol{Z}\boldsymbol{N}_{12}+\boldsymbol{N}^{\mathrm{T}}\overline{\boldsymbol{Z}}\boldsymbol{N})\end{aligned}$$

根据 Lyapunov 稳定性理论, 显然只要$\theta<0$, 则$\dot{V}(t)<0$, 增广动态误差系统 (16) 渐近稳定。

在零初始条件下, 当系统所受外界扰动$\overline{\boldsymbol{w}}(t)\neq\boldsymbol{0}$, 且存在估计误差$\boldsymbol{e}_y(t)\neq\boldsymbol{0}$时, 考虑如下性能指标:

$$\begin{aligned}J=&\dot{V}(t)+\bar{\boldsymbol{e}}_y^{\mathrm{T}}(t)\bar{\boldsymbol{e}}_y(t)-\\&\gamma_1^2\left[\overline{\boldsymbol{w}}^{\mathrm{T}}(t)\overline{\boldsymbol{w}}(t)+\boldsymbol{v}^{\mathrm{T}}(t-\tau(t))\boldsymbol{v}(t-\tau(t))\right]<0\end{aligned} \tag{23}$$

将式 (20) 和式 (21) 联立, 并令:

$$\boldsymbol{\xi}_1=\left[\bar{\boldsymbol{e}}(t)\ \ \bar{\boldsymbol{e}}(t-h_1)\ \ \gamma_1\ \ \gamma_2\ \ \bar{\boldsymbol{e}}(t-\tau(t))\ \ \boldsymbol{v}(t-\tau(t))\ \ \overline{\boldsymbol{w}}(t)\right]^{\mathrm{T}}$$

$$\boldsymbol{M}_{21}=[\boldsymbol{I}\ \boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{0}]$$

$$\boldsymbol{M}_{22}=[\overline{\boldsymbol{A}}\ \boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{0}\ -\overline{\boldsymbol{L}}\,\overline{\boldsymbol{C}}\ \ \overline{\boldsymbol{L}}\boldsymbol{E}_w\ \ -\overline{\boldsymbol{E}}_w]$$

$$\boldsymbol{M}_{23}=[\boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{I}\ \boldsymbol{0}\ \boldsymbol{0}]$$

$$\boldsymbol{M}_{24}=[\boldsymbol{0}\ \boldsymbol{I}\ \boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{0}\ \boldsymbol{0}]$$

$$\boldsymbol{M}_{25}=\begin{bmatrix}\boldsymbol{I}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{I}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{I}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{I}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{I}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\end{bmatrix}$$

$$\boldsymbol{M}_{26}=\begin{bmatrix}\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{I}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{I}&\boldsymbol{0}\end{bmatrix}$$

式中: $\bar{\boldsymbol{e}}(t)=\boldsymbol{M}_{21}\boldsymbol{\xi}_1$, $\dot{\bar{\boldsymbol{e}}}(t)=\boldsymbol{M}_{22}\boldsymbol{\xi}_1$, $\bar{\boldsymbol{e}}(t-\tau(t))=\boldsymbol{M}_{23}\boldsymbol{\xi}_1$, $\bar{\boldsymbol{e}}(t-h_1)=\boldsymbol{M}_{24}\boldsymbol{\xi}_1$, $\boldsymbol{\xi}=\boldsymbol{M}_{25}\boldsymbol{\xi}_1$, $\left[\boldsymbol{v}(t-\tau(t))\quad\overline{\boldsymbol{w}}(t)\right]^{\mathrm{T}}=\boldsymbol{M}_{26}\boldsymbol{\xi}_1$。

$$J\leqslant\boldsymbol{\xi}_1^{\mathrm{T}}\theta_1\boldsymbol{\xi}_1 \tag{24}$$

式中:

$$\begin{aligned}\theta_1=&\boldsymbol{M}_{21}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{M}_{22}+\boldsymbol{M}_{22}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{M}_{21}+\boldsymbol{M}_{21}^{\mathrm{T}}\boldsymbol{M}_{21}+\boldsymbol{M}_{21}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{M}_{21}-\\&\boldsymbol{M}_{24}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{M}_{24}+\boldsymbol{M}_{22}^{\mathrm{T}}\boldsymbol{Z}\boldsymbol{M}_{22}+\boldsymbol{M}_{25}^{\mathrm{T}}\left[\boldsymbol{\Pi}^{\mathrm{T}}\boldsymbol{N}+\boldsymbol{N}^{\mathrm{T}}\boldsymbol{\Pi}+\right.\\&\left.h_1\boldsymbol{N}^{\mathrm{T}}\overline{\boldsymbol{Z}}\boldsymbol{N}\right]\boldsymbol{M}_{25}-\frac{\pi^2}{4}\left[\boldsymbol{M}_{21}^{\mathrm{T}}-\boldsymbol{N}_{23}^{\mathrm{T}}\right]\boldsymbol{R}[\boldsymbol{M}_{21}-\boldsymbol{N}_{23}]+\\&\boldsymbol{\Pi}^{\mathrm{T}}\boldsymbol{N}+\boldsymbol{N}^{\mathrm{T}}\boldsymbol{\Pi}+h_1^2\boldsymbol{M}_{22}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{M}_{22}-\gamma_1^2\boldsymbol{M}_{26}^{\mathrm{T}}\boldsymbol{M}_{26}\end{aligned}$$

根据 Lyapunov 稳定性理论, 只要 $\theta_1<0$, 则 $J<0$, 增广动态误差系统满足给定的性能指标。

综上所述, 若增广误差系统 (16) 渐近稳定且满足相应的性能指标, 需要满足以下条件:

$$\begin{cases}\theta<0\\ \theta_1<0\end{cases} \tag{25}$$

将式 (25) 展开并用 Schur 引理, 令 $\boldsymbol{Y}=\boldsymbol{P}\bar{\boldsymbol{L}}$, 即可得到矩阵不等式 (17) 和 (18)。

对式 (23) 从 0 到+∞积分, 有

$$V(+\infty)-V(0)\leqslant-\int_0^{+\infty}\bar{\boldsymbol{e}}^{\mathrm{T}}(t)\bar{\boldsymbol{e}}(t)\mathrm{d}t+\gamma_1^2\left[\int_0^{+\infty}\bar{\boldsymbol{w}}^{\mathrm{T}}(t)\bar{\boldsymbol{w}}(t)\mathrm{d}t+\sum_{k=0}^{+\infty}(t_{k+1}h-t_kh)\boldsymbol{v}^{\mathrm{T}}(t_kh)\boldsymbol{v}(t_kh)\right] \tag{26}$$

最终可得 $\|\bar{\boldsymbol{e}}(t)\|_2^2\leqslant\gamma_1^2\left[\|\bar{\boldsymbol{w}}(t)\|_2^2+\sum_{k=0}^{+\infty}(t_{k+1}h-t_kh)\|\boldsymbol{v}(t_kh)\|_2^2\right]$。

证毕

由于矩阵不等式 (17) 和 (18) 中含有非线性项, 无法直接利用 LMI 工具箱求取 $\boldsymbol{L}$、$\boldsymbol{F}$。为解决该问题, 常用方法是通过变量替换[23]将非线性项转换为线性项, 再利用 LMI 求取 $\boldsymbol{L}$、$\boldsymbol{F}$, 但会增加结果的保守性, 而采用 CCL[25]算法, 则可直接从 NLMI 中得到少保守性的解。本文基于 CCL 得到求取 $\boldsymbol{L}$、$\boldsymbol{F}$ 的方法, 具体表述如下。

定理 2　存在适当维数矩阵 $\boldsymbol{W}_1$、$\boldsymbol{W}_2$, 使得 $\boldsymbol{W}_1<-\boldsymbol{P}\boldsymbol{R}^{-1}\boldsymbol{P}$, $\boldsymbol{W}_2<-\boldsymbol{P}\boldsymbol{Z}^{-1}\boldsymbol{P}$ 成立, 矩阵不等式 (17) 和 (18) 就可转化为

$$\begin{bmatrix}\boldsymbol{I}_{11}&\boldsymbol{I}_{12}&\boldsymbol{I}_{13}&\boldsymbol{I}_{14}&\boldsymbol{I}_{15}&\boldsymbol{N}&\boldsymbol{I}_{17}&\boldsymbol{I}_{18}\\ &\boldsymbol{I}_{22}&\boldsymbol{I}_{23}&\boldsymbol{I}_{24}&\boldsymbol{0}&\boldsymbol{N}&\boldsymbol{0}&\boldsymbol{0}\\ &&\boldsymbol{I}_{33}&\boldsymbol{I}_{34}&\boldsymbol{I}_{35}&\boldsymbol{N}&\boldsymbol{0}&\boldsymbol{0}\\ &&&\boldsymbol{I}_{44}&\boldsymbol{I}_{45}&\boldsymbol{N}&\boldsymbol{0}&\boldsymbol{0}\\ &&&&\boldsymbol{I}_{55}&\boldsymbol{0}&\boldsymbol{I}_{57}&\boldsymbol{I}_{58}\\ &&&&&\boldsymbol{I}_{66}&\boldsymbol{0}&\boldsymbol{0}\\ &&&&&&-\boldsymbol{W}_1&\boldsymbol{0}\\ &&&&&&&-\boldsymbol{W}_2\end{bmatrix}<\boldsymbol{0} \tag{27}$$

$$\begin{bmatrix}\boldsymbol{J}_1&\boldsymbol{J}_2\\ &\boldsymbol{J}_4'\end{bmatrix}<\boldsymbol{0} \tag{28}$$

$$\boldsymbol{J}_1=\begin{bmatrix}\boldsymbol{J}_{11}&\boldsymbol{J}_{12}&\boldsymbol{J}_{13}&\boldsymbol{J}_{14}&\boldsymbol{J}_{15}\\ &\boldsymbol{J}_{22}&\boldsymbol{J}_{23}&\boldsymbol{J}_{24}&\boldsymbol{0}\\ &&\boldsymbol{J}_{33}&\boldsymbol{J}_{34}&\boldsymbol{0}\\ &&&\boldsymbol{J}_{44}&\boldsymbol{0}\\ &&&&\boldsymbol{J}_{55}\end{bmatrix}$$

$$\boldsymbol{J}_2=\begin{bmatrix}\boldsymbol{J}_{16}&\boldsymbol{J}_{17}&\boldsymbol{N}&\boldsymbol{J}_{19}&\boldsymbol{J}_{10}\\ \boldsymbol{0}&\boldsymbol{0}&\boldsymbol{N}&\boldsymbol{0}&\boldsymbol{0}\\ \boldsymbol{0}&\boldsymbol{0}&\boldsymbol{N}&\boldsymbol{0}&\boldsymbol{0}\\ \boldsymbol{0}&\boldsymbol{0}&\boldsymbol{N}&\boldsymbol{0}&\boldsymbol{0}\\ \boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{J}_{49}&\boldsymbol{J}_{50}\end{bmatrix}$$

$$\boldsymbol{J}_4'=\begin{bmatrix}-\gamma_1^2\boldsymbol{I}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{J}_{69}&\boldsymbol{J}_{60}\\ &-\gamma_1^2\boldsymbol{I}&\boldsymbol{0}&\boldsymbol{J}_{79}&\boldsymbol{J}_{70}\\ &&\boldsymbol{J}_{88}&\boldsymbol{0}&\boldsymbol{0}\\ &&&\boldsymbol{W}_1&\boldsymbol{0}\\ &&&&\boldsymbol{W}_2\end{bmatrix}$$

对于 $\boldsymbol{W}_1<-\boldsymbol{P}\boldsymbol{R}^{-1}\boldsymbol{P}$, $\boldsymbol{W}_2<-\boldsymbol{P}\boldsymbol{Z}^{-1}\boldsymbol{P}$, 由 Schur 引理可得如下不等式组约束:

$$\begin{cases}\begin{bmatrix}\boldsymbol{W}_1^{-1}&\boldsymbol{P}^{-1}\\ \boldsymbol{P}^{-1}&\boldsymbol{R}^{-1}\end{bmatrix}\geqslant\boldsymbol{0}\\ \begin{bmatrix}\boldsymbol{W}_1^{-1}&\boldsymbol{I}\\ \boldsymbol{I}&\boldsymbol{W}_1^{-1}\end{bmatrix}\geqslant\boldsymbol{0}\\ \begin{bmatrix}\boldsymbol{P}&\boldsymbol{I}\\ \boldsymbol{I}&\boldsymbol{P}^{-1}\end{bmatrix}\geqslant\boldsymbol{0}\\ \begin{bmatrix}\boldsymbol{R}&\boldsymbol{I}\\ \boldsymbol{I}&\boldsymbol{R}^{-1}\end{bmatrix}\geqslant\boldsymbol{0}\\ \begin{bmatrix}\boldsymbol{W}_2^{-1}&\boldsymbol{P}^{-1}\\ \boldsymbol{P}^{-1}&\boldsymbol{Z}^{-1}\end{bmatrix}\geqslant\boldsymbol{0}\\ \begin{bmatrix}\boldsymbol{W}_2^{-1}&\boldsymbol{I}\\ \boldsymbol{I}&\boldsymbol{W}_2^{-1}\end{bmatrix}\geqslant\boldsymbol{0}\\ \begin{bmatrix}\boldsymbol{Z}&\boldsymbol{I}\\ \boldsymbol{I}&\boldsymbol{Z}^{-1}\end{bmatrix}\geqslant\boldsymbol{0}\end{cases} \tag{29}$$

则利用 CCL 可将 NLMI 的求解转化为如下最小化问题:

$$\min\operatorname{tr}\left(\boldsymbol{W}_1\boldsymbol{W}_1^{-1}+\boldsymbol{P}\boldsymbol{P}^{-1}+\boldsymbol{R}\boldsymbol{R}^{-1}+\boldsymbol{W}_1\boldsymbol{W}_1^{-1}+\boldsymbol{Z}\boldsymbol{Z}^{-1}\right) \tag{30}$$

约束条件为式(27)～式(29)。

式(30)的可行解即为 ADETCS 下系统状态、FDI 攻击/执行器故障的广义鲁棒估计器增益 $\boldsymbol{L}$、$\boldsymbol{F}$。

注 3　在 ADETCS 下, 由于触发参数依赖系统行为可动态自适应变化, 进而使时延函数 $\tau(t)$ 随之变化, 时滞依赖的鲁棒估计器由原 DETCS 下的静态事件驱动变成了动态事件驱动。这种自适应动态可变的通信机制, 在确保对状态、FDI 攻击/执行器故障估计性能的同时, 也会更大程度地节约网络通信和计算资源。

3　综合安全控制器的设计

在鲁棒估计器的基础上, 本节对双端网络 FDI 攻击、执行器故障的主动容侵、主动容错与通信协同设计方法进行研究。

基于前期的研究工作[20], 对于双端网络 FDI 攻击和执行器故障, 假设 FDI 攻击与执行器故障是可分离的, 首先, 将传感侧 FDI 攻击 $\boldsymbol{a}^{\mathrm{s}}(t_kh)$ 从广义故障估计 $\hat{\bar{\boldsymbol{f}}}(t_kh)$ 中分离出来, 不再纳入后续控制量的计

算；然后，对于执行侧FDI攻击$\boldsymbol{a}^{\mathrm{a}}(t_kh)$与执行器故障$\boldsymbol{f}(t)$，基于故障调节的思想，对二者的估计量$\hat{\boldsymbol{a}}^{\mathrm{a}}(t_kh)$、$\hat{\boldsymbol{f}}(t)$进行补偿。显然，无论是前者的分离，还是后者的补偿，均是以主动方式实施的安全防御策略，由于其兼顾了网络攻击的容侵和物理器件的容错，具有综合安全控制效能。

进一步，基于状态估计量$\hat{\boldsymbol{x}}(t_kh)$，对系统设计输出动态反馈控制，则有如下综合安全控制策略：

$$\boldsymbol{u}(t_kh)=-\boldsymbol{K}\hat{\boldsymbol{x}}(t_kh)-\boldsymbol{B}^{*}\boldsymbol{E}_f\hat{\boldsymbol{f}}(t_kh)-\boldsymbol{B}^{*}\boldsymbol{E}_{\mathrm{a}}\hat{\boldsymbol{a}}^{\mathrm{a}}(t_kh) \tag{31}$$

式中：$\boldsymbol{K}$为需要设计的控制器增益；$\boldsymbol{B}^{*}$为$\boldsymbol{B}$的伪逆矩阵，满足$(\boldsymbol{I}-\boldsymbol{B}\boldsymbol{B}^{*})\boldsymbol{E}_f=\boldsymbol{0}$；$\hat{\boldsymbol{x}}(t_kh)$、$\hat{\boldsymbol{a}}^{\mathrm{a}}(t_kh)$、$\hat{\boldsymbol{f}}(t_kh)$分别为状态估计向量、广义故障估计向量中分离出的执行侧FDI、执行器故障的估计量。

具体分离是：基于前述设计鲁棒估器得到的广义故障向量$\hat{\bar{\boldsymbol{f}}}(t_kh)=\left[\hat{\boldsymbol{f}}(t_kh)\quad\hat{\boldsymbol{a}}^{\mathrm{a}}(t_kh)\quad\hat{\boldsymbol{a}}^{\mathrm{s}}(t_kh)\right]^{\mathrm{T}}$，经如下分离处理，即有

$$\hat{\boldsymbol{f}}(t_kh)=\left[\boldsymbol{I}\quad\boldsymbol{0}\quad\boldsymbol{0}\right]*\hat{\bar{\boldsymbol{f}}}(t_kh)$$

$$\hat{\boldsymbol{a}}^{\mathrm{a}}(t_kh)=\left[\boldsymbol{0}\quad\boldsymbol{I}\quad\boldsymbol{0}\right]*\hat{\bar{\boldsymbol{f}}}(t_kh)$$

$$\hat{\boldsymbol{a}}^{\mathrm{s}}(t_kh)=\left[\boldsymbol{0}\quad\boldsymbol{0}\quad\boldsymbol{I}\right]*\hat{\bar{\boldsymbol{f}}}(t_kh)$$

在时滞理论下，将式(10)与式(31)联立，可得ICPS闭环模型为

$$\begin{aligned}\dot{\boldsymbol{x}}(t)=&\boldsymbol{A}\boldsymbol{x}(t)-\boldsymbol{B}\boldsymbol{K}\boldsymbol{x}(t-\tau(t))-\boldsymbol{B}\boldsymbol{K}\boldsymbol{e}_{\boldsymbol{x}}(t-\tau(t))-\\&\boldsymbol{E}_f\boldsymbol{e}_f(t-\tau(t))+\tau(t)\boldsymbol{E}_f\dot{\boldsymbol{f}}(t)-\\&\boldsymbol{B}\boldsymbol{E}_{\mathrm{a}}\boldsymbol{e}_{\mathrm{a}}(t-\tau(t))+\boldsymbol{E}_w\boldsymbol{w}(t)+\tau(t)\boldsymbol{B}\boldsymbol{E}_{\mathrm{a}}\dot{\boldsymbol{a}}^{\mathrm{a}}(t)\end{aligned} \tag{32}$$

式中：$\boldsymbol{e}_{\boldsymbol{x}}(t-\tau(t))$、$\boldsymbol{e}_{\mathrm{a}}(t-\tau(t))$、$\boldsymbol{e}_f(t-\tau(t))$分别为状态估计误差、执行器侧所受FDI攻击误差、执行器故障估计误差。

定理3 在ADETCS下，给定正数α、β、h_1、γ_2、$\bar{\sigma}$，若存在正定对称矩阵$\overline{\boldsymbol{P}}$、$\overline{\boldsymbol{Q}}$、$\overline{\boldsymbol{R}}$、$\hat{\boldsymbol{Z}}$与适维矩阵$\overline{\boldsymbol{K}}$、$\boldsymbol{\phi}$、$\boldsymbol{X}_1$、$\boldsymbol{X}_2$、$\boldsymbol{X}_3$、$\boldsymbol{X}_4$、$\boldsymbol{X}_5$、$\boldsymbol{X}_6$、$\boldsymbol{X}_7$、$\boldsymbol{X}_8$、$\boldsymbol{X}_9$、$\boldsymbol{X}_{10}$、$\boldsymbol{X}_{11}$、$\boldsymbol{X}_{12}$，使得矩阵不等式(33)～(35)成立，则存在控制器增益矩阵$\boldsymbol{K}$与事件触发权矩阵$\boldsymbol{\phi}$，使得闭环ICPS(32)在无扰动时渐近稳定，在有扰动时，满足性能指标$\|\boldsymbol{x}\|_2^2\leqslant\gamma_2^2\left(\|\boldsymbol{w}(t)\|_2^2+\sum_{k=0}^{+\infty}(t_{k+1}h-t_kh)\cdot(\|\boldsymbol{e}_{\boldsymbol{x}}(t_kh)\|_2^2+\left\|\boldsymbol{e}_f(t_kh)\right\|_2^2+\|\boldsymbol{e}_{\mathrm{a}}(t_kh)\|_2^2)\right)$，同时$\boldsymbol{K}$、$\boldsymbol{B}^{*}$可通过$\boldsymbol{K}=\left(\overline{\boldsymbol{P}}\boldsymbol{B}\right)^{-1}\overline{\boldsymbol{K}}$、$(\boldsymbol{I}-\boldsymbol{B}\boldsymbol{B}^{*})\boldsymbol{E}_f=\boldsymbol{0}$分别求取，其中，$\bar{\sigma}=(\sigma_{\min}+\sigma_{\max})/2$。

$$\begin{bmatrix}\boldsymbol{\pi}_{11}&\boldsymbol{\pi}_{12}\\&\boldsymbol{\pi}_{22}\end{bmatrix}<\boldsymbol{0} \tag{33}$$

$$\begin{bmatrix}\boldsymbol{\rho}_{11}&\boldsymbol{\rho}_{12}\\&\boldsymbol{\rho}_{22}\end{bmatrix}<\boldsymbol{0} \tag{34}$$

$$\begin{cases}\begin{bmatrix}\boldsymbol{X}_2&\boldsymbol{E}_f^{\mathrm{T}}\overline{\boldsymbol{P}}\\&\boldsymbol{X}_1\end{bmatrix}>\boldsymbol{0}\\\begin{bmatrix}\boldsymbol{X}_4&\boldsymbol{E}_{\mathrm{a}}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\overline{\boldsymbol{P}}\\&\boldsymbol{X}_3\end{bmatrix}>\boldsymbol{0}\\\begin{bmatrix}\boldsymbol{X}_6&\boldsymbol{E}_f^{\mathrm{T}}\hat{\boldsymbol{Z}}\\&\boldsymbol{X}_5\end{bmatrix}>\boldsymbol{0}\\\begin{bmatrix}\boldsymbol{X}_8&\boldsymbol{E}_{\mathrm{a}}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\hat{\boldsymbol{Z}}\\&\boldsymbol{X}_7\end{bmatrix}>\boldsymbol{0}\\\begin{bmatrix}\boldsymbol{X}_{10}&\boldsymbol{E}_f^{\mathrm{T}}\overline{\boldsymbol{R}}\\&\boldsymbol{X}_9\end{bmatrix}>\boldsymbol{0}\\\begin{bmatrix}\boldsymbol{X}_{12}&\boldsymbol{E}_{\mathrm{a}}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\overline{\boldsymbol{R}}\\&\boldsymbol{X}_{11}\end{bmatrix}>\boldsymbol{0}\end{cases} \tag{35}$$

式中：

$$\boldsymbol{\pi}_{11}=\begin{bmatrix}\boldsymbol{\vartheta}_{11}&\boldsymbol{\vartheta}_{12}&\boldsymbol{\vartheta}_{13}&\boldsymbol{\vartheta}_{14}&\boldsymbol{\vartheta}_{15}&\overline{\boldsymbol{N}}\\&\boldsymbol{\vartheta}_{22}&\boldsymbol{\vartheta}_{23}&\boldsymbol{\vartheta}_{24}&\boldsymbol{0}&\overline{\boldsymbol{N}}\\&&\boldsymbol{\vartheta}_{33}&\boldsymbol{\vartheta}_{34}&\boldsymbol{0}&\overline{\boldsymbol{N}}\\&&&\boldsymbol{\vartheta}_{44}&\boldsymbol{0}&\overline{\boldsymbol{N}}\\&&&&\boldsymbol{\vartheta}_{55}&\boldsymbol{0}\\&&&&&\boldsymbol{\vartheta}_{66}\end{bmatrix}$$

$$\boldsymbol{\pi}_{12}=\begin{bmatrix}\boldsymbol{\vartheta}_{17}&\boldsymbol{\vartheta}_{18}&\boldsymbol{\vartheta}_{19}&\boldsymbol{0}&\boldsymbol{\vartheta}_{101}&\boldsymbol{\vartheta}_{102}\\\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\\boldsymbol{\vartheta}_{57}&\boldsymbol{\vartheta}_{58}&\boldsymbol{\vartheta}_{59}&\boldsymbol{\vartheta}_{500}&\boldsymbol{\vartheta}_{501}&\boldsymbol{\vartheta}_{502}\\\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\end{bmatrix}$$

$$\boldsymbol{\pi}_{22}=\begin{bmatrix}\boldsymbol{\vartheta}_{77}&&&&&\\&\boldsymbol{\vartheta}_{88}&&&&\\&&\boldsymbol{\vartheta}_{99}&&&\\&&&\boldsymbol{\vartheta}_{1010}&&\\&&&&\boldsymbol{\vartheta}_{1111}&\\&&&&&\boldsymbol{\vartheta}_{1212}\end{bmatrix}$$

其中：$\boldsymbol{\vartheta}_{11}=\overline{\boldsymbol{P}}\,\overline{\boldsymbol{A}}+\overline{\boldsymbol{A}}^{\mathrm{T}}\,\overline{\boldsymbol{P}}+\overline{\boldsymbol{Q}}-\dfrac{\pi^2}{4}\boldsymbol{R}+3\overline{\boldsymbol{N}}-3\overline{\boldsymbol{N}}^{\mathrm{T}}+h_1(\boldsymbol{X}_1+\boldsymbol{X}_3)$，$\boldsymbol{\vartheta}_{12}=3\overline{\boldsymbol{N}}-\overline{\boldsymbol{N}}^{\mathrm{T}}$，$\boldsymbol{\vartheta}_{13}=3\overline{\boldsymbol{N}}-8\overline{\boldsymbol{N}}^{\mathrm{T}}$，$\boldsymbol{\vartheta}_{14}=3\overline{\boldsymbol{N}}+12\overline{\boldsymbol{N}}^{\mathrm{T}}$，$\boldsymbol{\vartheta}_{15}=\dfrac{\pi^2}{4}\overline{\boldsymbol{R}}$，$\boldsymbol{\vartheta}_{17}=h_1\overline{\boldsymbol{A}}^{\mathrm{T}}\overline{\boldsymbol{P}}$，$\boldsymbol{\vartheta}_{18}=h_1^{\frac{1}{2}}\overline{\boldsymbol{A}}^{\mathrm{T}}\overline{\boldsymbol{P}}$，$\boldsymbol{\vartheta}_{19}=h_1\overline{\boldsymbol{A}}^{\mathrm{T}}\overline{\boldsymbol{P}}$，$\boldsymbol{\vartheta}_{101}=h_1^{\frac{3}{2}}\overline{\boldsymbol{A}}^{\mathrm{T}}\overline{\boldsymbol{P}}$，$\boldsymbol{\vartheta}_{102}=h_1^{\frac{3}{2}}\overline{\boldsymbol{A}}^{\mathrm{T}}\overline{\boldsymbol{P}}$，$\boldsymbol{\vartheta}_{22}=-\overline{\boldsymbol{Q}}-\overline{\boldsymbol{N}}-\overline{\boldsymbol{N}}^{\mathrm{T}}$，$\boldsymbol{\vartheta}_{23}=-\overline{\boldsymbol{N}}-8\overline{\boldsymbol{N}}^{\mathrm{T}}$，$\boldsymbol{\vartheta}_{24}=-\overline{\boldsymbol{N}}+12\overline{\boldsymbol{N}}^{\mathrm{T}}$，$\boldsymbol{\vartheta}_{33}=-8\overline{\boldsymbol{N}}-\overline{\boldsymbol{N}}^{\mathrm{T}}$，$\boldsymbol{\vartheta}_{34}=-8\overline{\boldsymbol{N}}+12\overline{\boldsymbol{N}}^{\mathrm{T}}$，$\boldsymbol{\vartheta}_{44}=12\overline{\boldsymbol{N}}+12\overline{\boldsymbol{N}}^{\mathrm{T}}$，$\boldsymbol{\vartheta}_{55}=-\dfrac{\pi^2}{4}\overline{\boldsymbol{R}}$，$\boldsymbol{\vartheta}_{57}=-h_1\overline{\boldsymbol{K}}^{\mathrm{T}}$，$\boldsymbol{\vartheta}_{58}=-h_1^{\frac{1}{2}}\overline{\boldsymbol{K}}^{\mathrm{T}}$，$\boldsymbol{\vartheta}_{59}=-h_1\overline{\boldsymbol{K}}^{\mathrm{T}}$，$\boldsymbol{\vartheta}_{500}=-h_1\overline{\boldsymbol{K}}^{\mathrm{T}}$，$\boldsymbol{\vartheta}_{501}=-h_1^{\frac{3}{2}}\overline{\boldsymbol{K}}^{\mathrm{T}}$，$\boldsymbol{\vartheta}_{502}=-h_1^{\frac{3}{2}}\overline{\boldsymbol{K}}^{\mathrm{T}}$，$\boldsymbol{\vartheta}_{66}=-\dfrac{1}{9h_1}\overline{\boldsymbol{Z}}$，$\boldsymbol{\vartheta}_{77}=-\overline{\boldsymbol{P}}\boldsymbol{X}_3^{-1}\overline{\boldsymbol{P}}$，$\boldsymbol{\vartheta}_{88}=-\overline{\boldsymbol{P}}\,\overline{\boldsymbol{Z}}^{-1}\overline{\boldsymbol{P}}$，$\boldsymbol{\vartheta}_{99}=-\overline{\boldsymbol{P}}\boldsymbol{X}_5^{-1}\overline{\boldsymbol{P}}$，$\boldsymbol{\vartheta}_{1010}=-\overline{\boldsymbol{P}}\boldsymbol{X}_7^{-1}\overline{\boldsymbol{P}}$，$\boldsymbol{\vartheta}_{1111}=-\overline{\boldsymbol{P}}\boldsymbol{X}_9^{-1}\overline{\boldsymbol{P}}$，$\boldsymbol{\vartheta}_{1212}=-\overline{\boldsymbol{P}}\boldsymbol{X}_{11}^{-1}\overline{\boldsymbol{P}}$。

$$\boldsymbol{\rho}_{11}=\begin{bmatrix}\boldsymbol{\delta}_{11} & \boldsymbol{\delta}_{12} & \boldsymbol{\delta}_{13} & \boldsymbol{\delta}_{14} & \boldsymbol{\delta}_{15} & \boldsymbol{\delta}_{16} & \boldsymbol{\delta}_{17} & \boldsymbol{\delta}_{18} & \boldsymbol{\delta}_{19}\\ & \boldsymbol{\delta}_{22} & \boldsymbol{\delta}_{23} & \boldsymbol{\delta}_{24} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ & & \boldsymbol{\delta}_{33} & \boldsymbol{\delta}_{34} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ & & & \boldsymbol{\delta}_{44} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ & & & & \boldsymbol{\delta}_{55} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ & & & & & \boldsymbol{\delta}_{66} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ & & & & & & \boldsymbol{\delta}_{77} & \boldsymbol{0} & \boldsymbol{0}\\ & & & & & & & \boldsymbol{\delta}_{88} & \boldsymbol{0}\\ & & & & & & & & \boldsymbol{\delta}_{99}\end{bmatrix}$$

$$\boldsymbol{\rho}_{12}=\begin{bmatrix}\boldsymbol{\rho}_1 & \boldsymbol{\rho}_2\\ \boldsymbol{\rho}_3 & \boldsymbol{\rho}_4\end{bmatrix}$$

$$\boldsymbol{\rho}_1=\begin{bmatrix}\boldsymbol{0} & \boldsymbol{0} & \overline{\boldsymbol{N}}\\ \boldsymbol{0} & \boldsymbol{0} & \overline{\boldsymbol{N}}\\ \boldsymbol{0} & \boldsymbol{0} & \overline{\boldsymbol{N}}\\ \boldsymbol{0} & \boldsymbol{0} & \overline{\boldsymbol{N}}\end{bmatrix}$$

$$\boldsymbol{\rho}_3=\begin{bmatrix}\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}$$

$$\boldsymbol{\rho}_2=\begin{bmatrix}\boldsymbol{\delta}_{113} & \boldsymbol{\delta}_{114} & \boldsymbol{\delta}_{115} & \boldsymbol{\delta}_{116} & \boldsymbol{\delta}_{117} & \boldsymbol{\delta}_{118}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}$$

$$\boldsymbol{\rho}_4=\begin{bmatrix}\boldsymbol{\delta}_{513} & \boldsymbol{\delta}_{514} & \boldsymbol{\delta}_{515} & \boldsymbol{\delta}_{516} & \boldsymbol{\delta}_{517} & \boldsymbol{\delta}_{518}\\ \boldsymbol{\delta}_{613} & \boldsymbol{\delta}_{614} & \boldsymbol{\delta}_{615} & \boldsymbol{\delta}_{616} & \boldsymbol{\delta}_{617} & \boldsymbol{\delta}_{618}\\ \boldsymbol{\delta}_{713} & \boldsymbol{\delta}_{714} & \boldsymbol{\delta}_{715} & \boldsymbol{\delta}_{716} & \boldsymbol{\delta}_{717} & \boldsymbol{\delta}_{718}\\ \boldsymbol{\delta}_{813} & \boldsymbol{\delta}_{814} & \boldsymbol{\delta}_{815} & \boldsymbol{\delta}_{816} & \boldsymbol{\delta}_{817} & \boldsymbol{\delta}_{818}\\ \boldsymbol{\delta}_{913} & \boldsymbol{\delta}_{914} & \boldsymbol{\delta}_{915} & \boldsymbol{\delta}_{916} & \boldsymbol{\delta}_{917} & \boldsymbol{\delta}_{918}\end{bmatrix}$$

$$\boldsymbol{\rho}_{22}=\begin{bmatrix}\boldsymbol{\rho}_1' & \boldsymbol{\rho}_2'\\ & \boldsymbol{\rho}_4'\end{bmatrix}$$

$$\boldsymbol{\rho}_1'=\begin{bmatrix}\boldsymbol{\delta}_{1010} & & & \\ & \boldsymbol{\delta}_{1111} & & \\ & & \boldsymbol{\delta}_{1212} & \\ & & & \boldsymbol{\delta}_{1313}\end{bmatrix}$$

$$\boldsymbol{\rho}_2'=\begin{bmatrix}\boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}$$

$$\boldsymbol{\rho}_4'=\begin{bmatrix}\boldsymbol{\delta}_{1414} & & & & \\ & \boldsymbol{\delta}_{1515} & & & \\ & & \boldsymbol{\delta}_{1616} & & \\ & & & \boldsymbol{\delta}_{1717} & \\ & & & & \boldsymbol{\delta}_{1818}\end{bmatrix}$$

其中: $\boldsymbol{\delta}_{11}=\boldsymbol{\vartheta}_{11}+\boldsymbol{I}$, $\boldsymbol{\delta}_{12}=\boldsymbol{\vartheta}_{12}$, $\boldsymbol{\delta}_{13}=\boldsymbol{\vartheta}_{13}$, $\boldsymbol{\delta}_{14}=\boldsymbol{\vartheta}_{14}$, $\boldsymbol{\delta}_{15}=\dfrac{\pi^2}{4}\overline{\boldsymbol{R}}-\overline{\boldsymbol{K}}$, $\boldsymbol{\delta}_{16}=\dfrac{\pi^2}{4}\overline{\boldsymbol{R}}+\overline{\boldsymbol{K}}$, $\boldsymbol{\delta}_{17}=-\overline{\boldsymbol{P}}\boldsymbol{E}_f$, $\boldsymbol{\delta}_{18}=-\overline{\boldsymbol{P}}\,\overline{\boldsymbol{B}}\boldsymbol{E}_{\mathrm{a}}$, $\boldsymbol{\delta}_{19}=\overline{\boldsymbol{P}}\boldsymbol{E}_w$, $\boldsymbol{\delta}_{113}=\boldsymbol{\vartheta}_{17}$, $\boldsymbol{\delta}_{114}=\boldsymbol{\vartheta}_{18}$, $\boldsymbol{\delta}_{115}=\boldsymbol{\vartheta}_{19}$, $\boldsymbol{\delta}_{116}=\boldsymbol{\vartheta}_{100}$, $\boldsymbol{\delta}_{117}=\boldsymbol{\vartheta}_{101}$, $\boldsymbol{\delta}_{118}=\boldsymbol{\vartheta}_{102}$, $\boldsymbol{\delta}_{22}=\boldsymbol{\vartheta}_{22}$, $\boldsymbol{\delta}_{23}=\boldsymbol{\vartheta}_{23}$, $\boldsymbol{\delta}_{24}=\boldsymbol{\vartheta}_{24}$, $\boldsymbol{\delta}_{33}=\boldsymbol{\vartheta}_{33}$, $\boldsymbol{\delta}_{34}=\boldsymbol{\vartheta}_{34}$, $\boldsymbol{\delta}_{44}=\boldsymbol{\vartheta}_{44}$, $\boldsymbol{\delta}_{55}=\boldsymbol{\vartheta}_{55}$, $\boldsymbol{\delta}_{513}=\boldsymbol{\vartheta}_{57}$, $\boldsymbol{\delta}_{514}=\boldsymbol{\vartheta}_{58}$, $\boldsymbol{\delta}_{515}=\boldsymbol{\vartheta}_{59}$, $\boldsymbol{\delta}_{516}=\boldsymbol{\vartheta}_{500}$, $\boldsymbol{\delta}_{517}=\boldsymbol{\vartheta}_{501}$, $\boldsymbol{\delta}_{518}=\boldsymbol{\vartheta}_{502}$, $\boldsymbol{\delta}_{66}=-\gamma_2^2\boldsymbol{I}$, $\boldsymbol{\delta}_{613}=\boldsymbol{\delta}_{615}=\boldsymbol{\delta}_{616}=-h_1\overline{\boldsymbol{K}}^{\mathrm{T}}$, $\boldsymbol{\delta}_{614}=-h_1^{\frac{1}{2}}\overline{\boldsymbol{K}}^{\mathrm{T}}$, $\boldsymbol{\delta}_{617}=\boldsymbol{\delta}_{618}=-h_1^{\frac{3}{2}}\overline{\boldsymbol{K}}^{\mathrm{T}}$, $\boldsymbol{\delta}_{77}=-\gamma_2^2\boldsymbol{I}$, $\boldsymbol{\delta}_{713}=-h_1\boldsymbol{E}_f^{\mathrm{T}}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{714}=-h_1^{\frac{1}{2}}\cdot\boldsymbol{E}_f^{\mathrm{T}}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{715}=\boldsymbol{\delta}_{716}=-h_1\boldsymbol{E}_f^{\mathrm{T}}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{717}=\boldsymbol{\delta}_{718}=-h_1^{\frac{3}{2}}\boldsymbol{E}_f^{\mathrm{T}}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{88}=-\gamma_2^2\boldsymbol{I}$, $\boldsymbol{\delta}_{813}=-h_1\boldsymbol{E}_{\mathrm{a}}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{814}=-h_1^{\frac{1}{2}}\boldsymbol{E}_{\mathrm{a}}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{815}=\boldsymbol{\delta}_{816}=-h_1\boldsymbol{E}_{\mathrm{a}}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{817}=\boldsymbol{\delta}_{818}=-h_1^{\frac{3}{2}}\boldsymbol{E}_{\mathrm{a}}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{99}=-\gamma_2^2\boldsymbol{I}$, $\boldsymbol{\delta}_{913}=-h_1\boldsymbol{E}_w^{\mathrm{T}}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{914}=-h_1^{\frac{1}{2}}\boldsymbol{E}_w^{\mathrm{T}}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{915}=\boldsymbol{\delta}_{916}=-h_1\boldsymbol{E}_w^{\mathrm{T}}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{917}=\boldsymbol{\delta}_{918}=-h_1^{\frac{3}{2}}\boldsymbol{E}_w^{\mathrm{T}}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{1010}=\overline{\sigma}\boldsymbol{\phi}$, $\boldsymbol{\delta}_{1111}=-\boldsymbol{\phi}$, $\boldsymbol{\delta}_{1212}=-\dfrac{1}{9h_1}\hat{\boldsymbol{Z}}$, $\boldsymbol{\delta}_{1313}=-\overline{\boldsymbol{P}}\,\overline{\boldsymbol{R}}^{-1}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{1414}=-\overline{\boldsymbol{P}}\hat{\boldsymbol{Z}}^{-1}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{1515}=-\overline{\boldsymbol{P}}\boldsymbol{X}_5^{-1}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{1616}=-\overline{\boldsymbol{P}}\boldsymbol{X}_7^{-1}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{1717}=-\overline{\boldsymbol{P}}\boldsymbol{X}_9^{-1}\overline{\boldsymbol{P}}$, $\boldsymbol{\delta}_{1818}=-\overline{\boldsymbol{P}}\boldsymbol{X}_{11}^{-1}\overline{\boldsymbol{P}}$。

定理 1 与定理 3 的证明类似, 此处不再赘述。

注 4　由于本文设计的新型 ADETCS 触发参数可随系统行为双向连续变化, 且介于 $[\sigma_{\min},\sigma_{\max}]$ 之间。若 $\overline{\sigma}$ 以上界 $\sigma_{\max}$ 设计控制器, 虽有利于系统通信资源节约, 但控制器会因此更加保守, 系统性能难达优性; 若 $\overline{\sigma}$ 以下界 $\sigma_{\min}$ 设计控制器, 在对应的小时延下更易使系统性能优性, 但对大时延有一定风险且不利于通信资源节约。显然, $\overline{\sigma}$ 的选取会使系统性能与通信资源节约之间出现“此消彼长”的关系, 因而本文中初步选取了 $\sigma(t_kh)$ 上下界的均值 $\overline{\sigma}=(\sigma_{\min}+\sigma_{\max})/2$, 以兼顾系统性能与通信资源节约之间的折中, 但 $\overline{\sigma}$ 的取值尚需进一步优化。

由于鲁棒估计器与综合安全控制器均处于控制单元, 在 ADETCS 下, 无论是估计还是控制, 均统一于同一非均匀传输框架下, 尽管观测器的设计并未直接关联通信, 但因与控制器的工作周期协同工作, 能够依据 ICPS 的行为自适应地动态变更数据传输周期, 从而使系统安全性能与资源节约得到更有效的折中平衡。

4　仿真实验与结果分析

为验证本文提出的新型 ADETCS, 以及在该通信机制下给出的鲁棒估计器和综合安全控制器设计的有效性和优势, 应用工业四容水箱系统进行仿真实验研究。

为进一步验证本文方法的优势, 用同一工业中常见的四容水箱系统, 分别与文献 [5,7,13,20] 进行对比性实验研究。

4.1　工业实例描述

工业四容水箱系统选自文献 [26], 具体参数为: $\boldsymbol{w}(t)$、$\boldsymbol{v}(i_kh)$ 为高斯白噪声过程, 服从高斯分布 $N(0,0.01)$, 系统初始状态为: $\boldsymbol{x}_0=[4,4,2,2]^{\mathrm{T}}$。

$$\boldsymbol{A}=\begin{bmatrix}-0.016 & 0 & 0.042 & 0\\ 0 & -0.011 & 0 & 0.033\\ 0 & 0 & -0.042 & 0\\ 0 & 0 & 0 & -0.033\end{bmatrix},\ \boldsymbol{B}=\begin{bmatrix}0.083 & 0\\ 0 & 0.063\\ 0 & 0.048\\ 0.031 & 0\end{bmatrix},$$

$$\boldsymbol{E}_w=\begin{bmatrix}0.01\\ 0.01\\ 0\\ 0.01\end{bmatrix},\boldsymbol{E}_v=\begin{bmatrix}0.01\\ 0\\ 0.01\\ 0.01\end{bmatrix},\boldsymbol{E}_f=-\begin{bmatrix}0.083\\ 0\\ 0\\ 0.031\end{bmatrix},\boldsymbol{C}=\frac{\boldsymbol{I}_4}{2},$$

$$\boldsymbol{E}_{\mathrm{s}}=\begin{bmatrix}1\\ 1\\ 1\\ 1\end{bmatrix},\boldsymbol{E}_{\mathrm{a}}=\begin{bmatrix}1\\ 1\end{bmatrix}。$$

假设连续时变故障为

$$f(t)=\begin{cases}0 & t<250\\ 2+2\sin(0.1\pi(t-250)) & 250\leqslant t\leqslant 800\end{cases}$$

传感侧、执行器侧FDI攻击$a^{\mathrm{s}}(t)$、$a^{\mathrm{a}}(t)$均为

$$a^{\mathrm{s}}(t)=a^{\mathrm{a}}(t)=\begin{cases}0 & t<150\\ 1+\sin(0.1\pi(t-150)) & 150\leqslant t\leqslant 800\end{cases}$$

4.2 新型ADETCS下综合安全控制的有效性与少保守性分析

选取$h_1=2.2$、$\gamma_1=3.9$、$\gamma_2=1.5$、$\alpha=0.5$、$\beta=50$、$\sigma_{\min}=0.001$、$\sigma_{\max}=0.005\,2$，则设计控制器时$\bar{\sigma}=0.003\,1$，利用定理1～定理3可求得四容水箱系统的如下设计参数。

状态与增广故障估计增益为

$$\boldsymbol{L}=\begin{bmatrix}0.498\,9 & -0.237\,4 & 0.478\,0 & -0.737\,5\\ 0.250\,7 & -0.072\,3 & 0.280\,2 & -0.457\,2\\ 0.113\,7 & -0.055\,8 & 0.152\,6 & -0.209\,9\\ 0.136\,2 & -0.073\,9 & 0.126\,1 & -0.187\,7\end{bmatrix}$$

$$\boldsymbol{F}=\begin{bmatrix}-0.116\,5 & 0.056\,5 & -0.079\,2 & 0.138\,8\\ 0.172\,8 & -0.076\,8 & 0.189\,2 & -0.284\,5\\ -0.073\,0 & 0.071\,1 & -0.086\,4 & 0.206\,6\end{bmatrix}$$

伪逆矩阵$\boldsymbol{B}^*$为

$$\boldsymbol{B}^*=\begin{bmatrix}2.163\,0 & -3.883\,4 & 5.097\,0 & 0\\ -12.625\,9 & 10.043\,0 & 7.651\,8 & 0\end{bmatrix}$$

综合安全控制器增益为

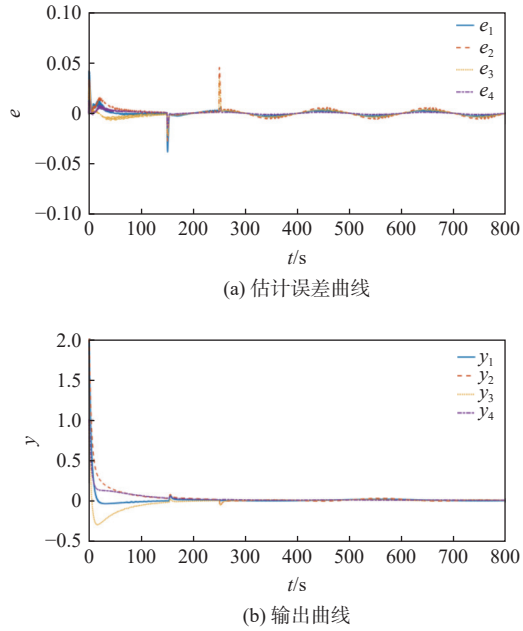
$$\boldsymbol{K}=\begin{bmatrix}1.352\,1 & -0.128\,0 & 0.697\,5 & 2.091\,6\\ 0.286\,5 & 0.273\,9 & 2.022\,3 & 1.027\,0\end{bmatrix}$$

事件触发权矩阵为

$$\boldsymbol{\phi}=10^{-8}\begin{bmatrix}0.223\,2 & & & \\ & 0.223\,2 & & \\ & & 0.223\,2 & \\ & & & 0.223\,2\end{bmatrix}$$

图2给出了四容水箱系统估计误差及系统输出曲线。为体现本文采用新型ADETCS与CCL后的效果，图3给出了采用文献[20]的系统相应结果。

观察图2(a)可知，初始只有外部扰动，系统估计误差约为0.03，之后趋于0；当150 s或250 s，双侧遭遇攻击或同时存在FDI攻击与执行器故障时，系统估计误差除瞬间约为−0.05与0.06，之后均迅速趋于0；鲁棒观测器的最小干扰抑制率$\gamma_{1\,\min}^{*}=1.829\,5$。观察图3(a)可知，相同时间段内，文献[20]中估计误差约为0.08，之后迅速趋于0.02，并维持在大小为0.02的误差带内，其最小抑制率$\gamma_{1\,\min}=2.147\,9$。观察图2(b)与图3(b)，基于鲁棒估计器估计值的综合安全控制器，亦可快速地对攻击与故障进行抵御与补偿，最小干扰抑制率分别

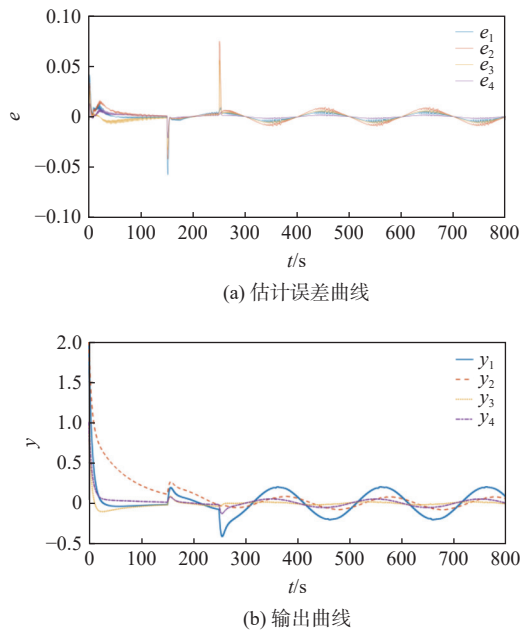


(a) 估计误差曲线

(b) 输出曲线

图2 本文方法系统估计误差与输出曲线

Fig. 2 System estimation error and output curves of the proposed method



(a) 估计误差曲线

(b) 输出曲线

图3 文献[20]系统估计误差与输出曲线

Fig. 3 System estimation error and output curves in Ref. [20]

为 $\gamma_{2\min}^{*}=2.082\,1$, $\gamma_{2\min}^{*}=3.263\,7$。

此外, 在 800 s 的仿真时段内, 文献 [20] 需要传输 987 次, 本文仅需传输 518 次即可获得更优的性能, 表明本文用 ADETCS 置换 DETCS 后, 可更有效地节约通信资源。

由此可见, 利用本文方法, 无论是估计还是防御, 性能均优于文献 [20], 即对 ICPS 的 FDI 攻击与故障起到了更有效的防范作用, 同时也显示出本文在求解时由于引入了 CCL 算法, 使结果具有了更少的保守性。

进一步, 为体现主动容侵策略对攻击防御的优势, 在 FDI 攻击与执行器故障不变的情形下, 图 4 给出了采用文献 [13] 容侵策略的四容水箱系统输出响应曲线。可以看出, 在 150 s 加入 FDI 攻击后, 相较于图 2(b), 图 4 所示曲线的波动程度明显增大。这充分说明本文所设计的主动容侵策略, 可更有效地防御双端 FDI 攻击。究其原因, 文献 [13] 采用弹性控制方式应对 FDI 攻击, 未对 FDI 攻击进行在线估计, 本质上属于被动容侵; 而本文采用对攻击分离、补偿的主动应对策略, 可增强系统对 FDI 攻击的主动防御能力, 因而确保了系统在攻击下更优的性能。

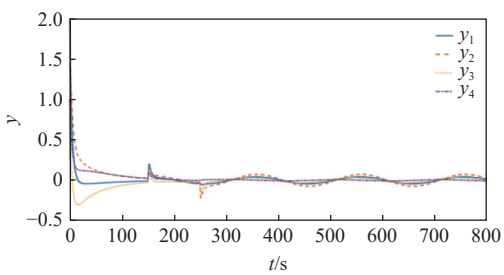


图 4　文献 [13] 弹性控制下系统输出曲线

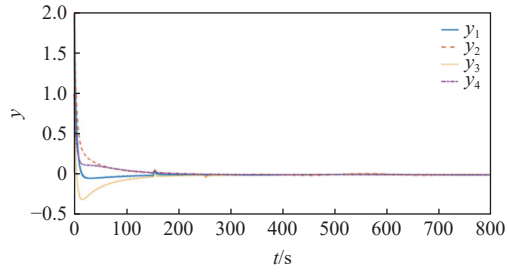
Fig. 4　System output curves of resiliently control in Ref. [13]

4.3　新型 ADETCS 通信资源节约与系统性能的对比分析

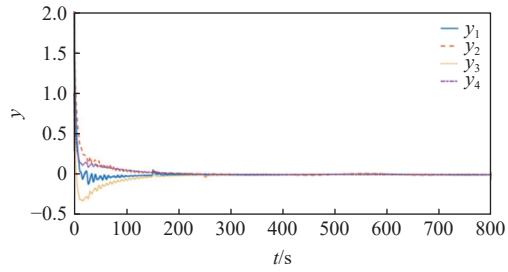
为进一步验证本文设计的新型 ADETCS 在通信资源节约和系统性能方面的优势, 分别将文献 [5,7] 的自适应触发方式引入前述设计结果中, 即在相同性能指标约束和控制参数下, 进行对比性实验。

实验中, 文献 [5] 中仅有最小值约束, 取为 $\sigma_{\min}=0.001$, 可调参数 $\alpha=0.01$; 由于文献 [7] 与本文中自适应触发参数均有上下界约束, 即 $\sigma(t_kh)\in[\sigma_{\min},\sigma_{\max}]$, 仍取 $\sigma_{\min}=0.001$、$\sigma_{\max}=0.005\,2$、$\rho=0.001$。

图 2(b) 与图 5 为不同 ADETCS 下的系统输出曲线, 图 6 为不同 ADETCS 中 $\sigma(t_kh)$ 的变化曲线, 表 1 为同约束、不同通信机制下系统传输数据的次数。



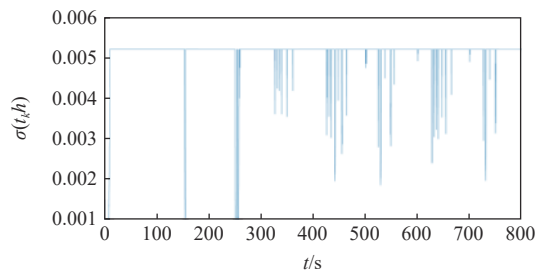
(a) 文献 [5] 系统输出曲线



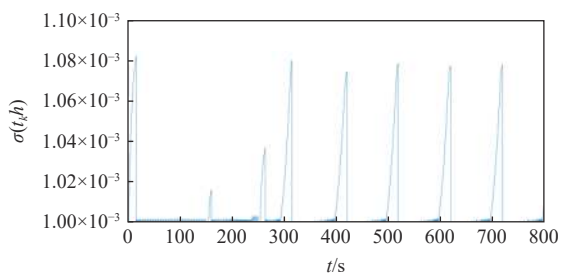
(b) 文献 [7] 系统输出曲线

图 5　不同 ADETCS 下系统输出曲线

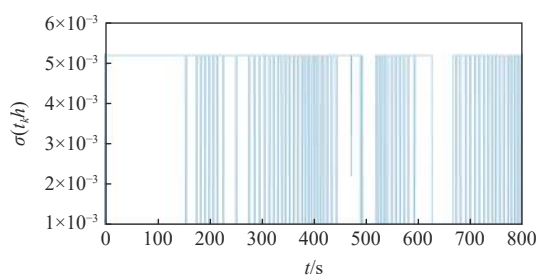
Fig. 5　System output curves in different ADETCSs



(a) 新型 ADETCS



(b) 文献 [5]



(c) 文献 [7]

图 6　不同 ADETCS 中 $\sigma(t_kh)$ 的变化曲线

Fig. 6　$\sigma(t_kh)$ variation curves in different ADETCSs

就通信资源的节约而言, 由表 1 可知, 在 800 s 的仿真时段内, ADETCS 下的数据传输量较 DETCS

表1 同约束、不同通信机制的数据传输次数

Table 1 Data transmission times of different communication mechanisms under the same constraints

ADETCS	传输次数
文献[20] DETCS	987
文献[5] ADETCS	643
文献[7] ADETCS	681
本文新型ADETCS	528

均有明显减少，即均可有效节约通信资源。文献[5]中自适应触发参数上限未人为界定，系统运行中最大达到了0.002 3，数据需传输643个；文献[7]和本文均限定了0.005 2的上界，前者共计传输681个数据，而后者仅需传输528个数据，显然新型ADETCS对网络资源的节约更有效。

就系统性能而言，比较图2(b)与图5的系统输出曲线发现，除故障和攻击发生的瞬间外，新型ADETCS下的系统输出均很平稳，文献[5]的输出出现了轻微波动，文献[7]的输出则出现了小幅高频振荡。由此可得，在相同性能指标约束和控制参数下，尽管不同的ADETCS均较DETCS节约了通信资源，但对系统性能也产生了程度不一的影响，本文中新型ADETCS下系统性能得到了良好保障。

追根溯源，不同的ADETCS，其本质是阈值函数依赖的λ不同。进一步观察图6可发现，λ的选取直接影响了阈值函数$\sigma(t_kh)$的变化。文献[5]中λ的实质是反正切函数，虽然连续但因判决条件只关注$\|\boldsymbol{y}(t_kh)\|$小于$\|\boldsymbol{y}(t_{k-1}h)\|$的误差情形，忽略了$\|\boldsymbol{y}(t_kh)\|$大于$\|\boldsymbol{y}(t_{k-1}h)\|$一侧，$\lambda$仅随单向误差连续变化；文献[7]中$\lambda$虽是双向变化的，但却以离散方式增减，其值的跃变使数据传输率也随之跃变，导致系统发生了小幅高频振荡；新型ADETCS通过选取正态型λ，使$\sigma(t_kh)$可随系统双向误差的大小连续变化，从而确保了系统的性能。

综上所述，新型ADETCS在维持系统性能的同时，还能兼顾资源的有效节约，最终实现了系统性能与资源节约之间更优的动态平衡。

5 结 论

本文针对一类同时具有FDI攻击与故障的ICPS，设计了一种新型ADETCS，并在该机制下进行了攻击与故障的估计、综合安全控制与通信的协同设计等方法的研究，得到如下结论：

1）针对DETCS中静态不变的触发参数，不利于ICPS控制性能与通信资源节约的更优折中平衡，通过引入ADETCS，并对其触发参数依赖的阈值函数λ巧妙设计，提出了一种新型ADETCS，实验结果表明，相较于现存阈值函数单向连续或双向离散变化的ADETCS，新型ADETCS使$\sigma(t_kh)$可随系统双向误差的大小连续自适应变化，从而在确保系统性能的同时，节约了更多的资源。

2）针对攻击与故障共存的ICPS，在新型ADETCS下构建了具有主动容侵和主动容错能力的ICPS综合安全架构，由于随系统行为自适应变化的触发参数$\sigma(t_kh)$取代了DETCS中静态不变的触发参数，使得鲁棒估计器与综合安全控制器，可在自适应动态事件驱动下，有效地估计和防御攻击与故障对ICPS的影响，并进一步节约通信与计算资源。

3）在新型ADETCS下，通过对自适应非均匀传输周期的时延转换，在统一的自适应变采样框架下，推证出了估计器及综合安全控制器的求取方法。尤其是以CCL替换传统的线性法，可直接从NLMI中得到更少保守性的解，提升了解的优性，进而在新型ADETCS下，使ICPS的综合安全控制达到了系统性能与资源节约之间更优的动态平衡。

在ADETCS下，如何使控制器与观测器的设计更具自适应性，将是极有意义和挑战性的工作。

参考文献（References）

[1] 刘烃, 田决, 王稼舟, 等. 信息物理融合系统综合安全威胁与防御研究[J]. 自动化学报, 2019, 45(1): 5-24.
LIU T, TIAN J, WANG J Z, et al. Integrated security threats and defense of cyber-physical systems[J]. Acta Automatica Sinica, 2019, 45(1): 5-24(in Chinese).

[2] ANTA A, TABUADA P. To sample or not to sample: Self-triggered control for nonlinear systems[J]. IEEE Transactions on Automatic Control, 2010, 55(9): 2030-2042.

[3] WANG X F, LEMMON M D. Event-triggering in distributed networked control systems[J]. IEEE Transactions on Automatic Control, 2011, 56(3): 586-601.

[4] PENG C, HAN Q L, YUE D. To transmit or not to transmit: A discrete event-triggered communication scheme for networked Takagi-Sugeno fuzzy systems[J]. IEEE Transactions on Fuzzy Systems, 2013, 21(1): 164-170.

[5] PENG C, ZHANG J, YAN H. Adaptive event-triggering H_∞ load frequency control for network-based power system[J]. IEEE Transactions on Industrial Electronics, 2016, 63(8): 1685-1694.

[6] YUAN S, YU C P, SUN J. Adaptive event-triggered consensus control of linear multi-agent systems with cyber attacks[J]. Neurocomputing, 2021, 442: 1-9.

[7] 乔伟豪, 朱凤增, 彭力. 基于自适应事件触发的时滞系统分布式l_2-l_∞滤波[J]. 控制与决策, 2022, 37(4): 1074-1080.
QIAO W H, ZHU F Z, PENG L. Distributed l_2-l_∞ filtering based on adaptive event triggering for time-delay systems[J]. Control and Decision, 2022, 37(4): 1074-1080(in Chinese).

[8] ZHANG J, PENG C, DU D J, et al. Adaptive event-triggered communication scheme for networked control systems with randomly occurring nonlinearities and uncertainties[J]. Neurocomputing, 2016,

174: 475-482.

[9] PAN Y N, YANG G H. Novel event-triggered filter design for nonlinear networked control systems[J]. Journal of the Franklin Institute, 2018, 355(3): 1259-1277.

[10] GU Z, SHI P, YUE D. An adaptive event-triggering scheme for networked interconnected control system with stochastic uncertainty[J]. International Journal of Robust and Nonlinear Control, 2017, 27(2): 236-251.

[11] GU Z, TIAN E G, LIU J L. Adaptive event-triggered control of a class of nonlinear networked systems[J]. Journal of the Franklin Institute, 2017, 354(9): 3854-3871.

[12] HU L, WANG Z D, HAN Q L, et al. State estimation under false data injection attacks: Security analysis and system protection[J]. Automatica, 2018, 87: 176-183.

[13] SUN Z W, XUE W F, LIU J L, et al. Adaptive event-triggered resilient control of industrial cyber physical systems under asynchronous data injection attack[J]. Journal of the Franklin Institute, 2022, 359(7): 3000-3023.

[14] GU Y P, YU X, GUO K X, et al. Detection, estimation, and compensation of false data injection attack for UAVs[J]. Information Sciences, 2021, 546: 723-741.

[15] MAO Z H, JIANG B, SHI P. Observer-based fault-tolerant control for a class of networked control systems with transfer delays[J]. Journal of the Franklin Institute, 2011, 348(4): 763-776.

[16] HUO Z H, ZHANG Z X, FANG H J. Research on fault-tolerant control of networked control systems based on information scheduling[J]. Journal of Systems Engineering and Electronics, 2008, 19(5): 1024-1028.

[17] HAN S Y, CHEN Y H, TANG G Y. Fault diagnosis and fault-tolerant tracking control for discrete-time systems with faults and delays in actuator and measurement[J]. Journal of the Franklin Institute, 2017, 354(12): 4719-4738.

[18] YE D, LUO S P. A co-design methodology for cyber-physical systems under actuator fault and cyber attack[J]. Journal of the Franklin Institute, 2019, 356(4): 1856-1879.

[19] LI W, SHI Y, LI Y. Research on secure control and communication for cyber-physical systems under cyber attacks[J]. Transactions of the Institute of Measurement & Control, 2019, 41(12): 3421-3437.

[20] 李炜, 张建军. 攻击与故障共存的 ICPS 综合安全控制方法[J]. 浙江大学学报 (工学版), 2021, 55(6): 1185-1198.
LI W, ZHANG J J. Integrated security control method for industrial cyber-physical system with attack and fault[J]. Journal of Zhejiang University (Engineering Science), 2021, 55(6): 1185-1198(in Chinese).

[21] FRIDMAN E. A refined input delay approach to sampled-data control[J]. Automatica, 2010, 46(2): 421-427.

[22] ZHANG K, JIANG B, STAROSWIECKI M. Dynamic output feedback fault tolerant controller design for Takagi-Sugeno fuzzy systems with actuator faults[J]. IEEE Transactions on Fuzzy Systems, 2010, 18(1): 194-201.

[23] 邱爱兵, 吉虹钢, 顾菊平. 非均匀采样数据系统时变故障估计与调节最优集成设计[J]. 自动化学报, 2014, 40(7): 1493-1504.
QIU A B, JI H G, GU J P. Optimal integrated design of time-varying fault estimation and accommodation for nonuniformly sampled data systems[J]. Acta Automatica Sinica, 2014, 40(7): 1493-1504(in Chinese).

[24] ZHANG X M, HAN Q L, SEURET A, et al. Overview of recent advances in stability of linear systems with time-varying delays[J]. IET Control Theory & Applications, 2019, 13(1): 1-16.

[25] EL GHAOUI L, OUSTRY F, AITRAMI M. A cone complementarity linearization algorithm for static output-feedback and related problems[J]. IEEE Transactions on Automatic Control, 1997, 42(8): 1171-1176.

[26] JOHANSSON K H. The quadruple-tank process: A multivariable laboratory process with an adjustable zero[J]. IEEE Transactions on Control Systems Technology, 2000, 8(3): 456-465.

Integrated security control of industrial cyber-physical systems based on new type ADETCS

LI Wei[1, 2, 3, *], CHENG Xue[1, 2, 3], LI Yajie[1, 2, 3]

(1. College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China;
2. Gansu Provincial Key Laboratory of Advanced Control for Industrial Processes, Lanzhou 730050, China;
3. National Experimental Teaching Demonstration Center of Electrical and Control Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: For industrial cyber-physical system (ICPS) with false data injection (FDI) attacks and faults, the collaborative design of integrated security control and communication in the new type adaptive discrete event-triggered communication scheme (ADETCS) was studied. Firstly, the ADETCS was introduced to replace the traditional discrete event-triggered communication scheme, and an ICPS architecture with adaptive collaborative optimization of security control and communication was constructed. Secondly, through the optimization design of the threshold function in ADETCS, a new type ADETCS was obtained, which could change continuously in both directions with the system's dynamic behaviors. Then, based on the active attack-tolerant and active fault-tolerant ideas, the Lyapunov function was constructed, and less conservative techniques such as new Bessel-Legebdre inequality and cone complementarity linearization (CCL) were used. Under the unified adaptive variable sampling framework, the solution methods of the robust estimator and integrated security controller were derived so that ICPS could defend against FDI attacks and faults at the same time and have better adaptive communication regulation ability. The simulation results show that the ingenious design of threshold function in the new type ADETCS and the application of CCL can achieve a better compromise between the integrated security control performance and resource saving of ICPS.

Keywords: new type adaptive discrete event-triggered communication scheme; industrial cyber-physical system; integrated security control; cone complementarity linearization; dynamic optimal balance

Received: 2022-08-20; **Accepted**: 2022-12-30; **Published Online**: 2023-01-30 15:16
URL: link.cnki.net/urlid/11.2625.V.20230130.1249.002
Foundation item: National Natural Science Foundation of China (62163022)
* **Corresponding author**. E-mail: liwei@lut.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0758

基于暗通道先验引导的图像去雾网络

黄淑英[1]，夏钰锟[2]，杨勇[3,*]，万伟国[4]，邱根莹[2]

(1. 天津工业大学 软件学院，天津 300387；2. 江西财经大学 信息管理学院，南昌 330032；
3. 天津工业大学 计算机科学与技术学院，天津 300387；4. 江西财经大学 软件与物联网工程学院，南昌 330032)

摘 要：基于深度学习的去雾方法多数直接学习有雾图像和无雾图像之间的映射关系，未结合有雾图像自身特点，存在雾信息检测不精确、去雾不彻底的问题。针对该问题，提出一种基于暗通道先验引导的图像去雾网络(DCPDNet)。通过卷积层提取有雾图像的浅层特征；构建2个特征增强模块(FEB)来增强图像的空间特征，该模块在2个尺度上对图像特征进行增强，即利用深层特征图实现语义特征的增强，浅层特征图对实现图像细节特征的增强；为使提取的特征更关注雾的区域，基于有雾图像中雾的成像特点设计基于引导图的特征校正模块(FCB)，利用暗通道先验理论构建引导图将网络学习的注意力引导到有雾区域，对提取的深层特征图做进一步的细化和校正；利用残差结构的跳转连接，将增强的浅层特征补充网络丢失的细节特征，并经过卷积操作重建去雾后图像。实验结果证明：DCPDNet可以在保持模型轻量型及运行速度较快的情况下实现良好的去雾效果。与近年先进的去雾方法进行比较，DCPDNet不仅在效率上占有优势，其去雾效果在主观视觉感受和客观评价结果上都获得了更好的效果。

关 键 词：图像去雾；暗通道先验；引导图；残差学习；特征增强
中图分类号：TP391.41
文献标志码：A **文章编号**：1001-5965(2024)09-2717-10

雾霾天气下拍摄到的图像通常存在对比度低、细节丢失等问题，会影响后续高级视觉任务，如目标识别、物体分类等[1]。对雾天图像进行去雾可以有效恢复被雾气遮挡的图像细节，提高目标识别等任务的正确率。现有的图像去雾方法主要分为2类，即传统方法的图像去雾和基于深度学习的图像去雾。

传统方法的图像去雾大多依赖大气散射模型，通过估计模型中的参数(大气光值、透射率)反推物理模型，求解得到去雾图像。但由于缺少图像的深度信息，单幅有雾图像的透射率很难被准确估计，为此，一些学者提出采用先验知识来约束模型的求解。He等[2]提出暗通道先验算法，利用图像暗像素对透射率进行估计，但该算法不适用于处理天空区域的去雾。Gao等[3]提出一种基于暗通道先验和形态学重构的单幅图像去雾方法，来解决现有方法结果存在的过于平滑、缺失细节的问题。Wang等[4]假设有雾图像及对应的清晰图像在最小通道内存在线性关系，提出一种基于线性变换的单幅图像快速去雾方法来实现运行时间和处理效果之间的平衡。Zhu等[5]提出利用能量函数最小化来估计透射率的去雾方法，解决了暗通道先验方法中的缺陷，避免了对图像的后续细化工作。Liu等[6]提出一种透射率自适应正则化图像去雾方法来实现高质量的单图像去雾，能有效抑制复原图像的伪影。传统的图像去雾方法虽然具有较强的可解释性，但由于

收稿日期：2022-09-04；**录用日期**：2022-12-16；**网络出版时间**：2023-01-05 09:36
网络出版地址：link.cnki.net/urlid/11.2625.V.20230104.1422.001
基金项目：国家自然科学基金(61862030,62072218,62261025)；江西省自然科学基金(20192ACB20002,20192ACBL21008)；江西省博士后科研项目(2020KY44)
* **通信作者**. E-mail：greatyangy@126.com

引用格式：黄淑英，夏钰锟，杨勇，等. 基于暗通道先验引导的图像去雾网络[J]. 北京航空航天大学学报，2024，50(9)：2717-2726.
HUANG S Y，XIA Y K，YANG Y，et al. Image dehazing network based on dark channel prior guidance[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2717-2726 (in Chinese).

图像退化是一个复杂问题，大气散射模型很难精确描述图像的退化过程，该类方法的泛化性较差[7]。

近年来，由于深度学习方法具有强大的特征学习能力，能较好地表征网络输入与输出之间的映射，基于深度学习的去雾方法相继被提出。Yeh等[8]将图像去雾定义为低频和高频无监督学习分支的网络结构，提出一种基于多尺度残差学习和图像分解的单幅图像去雾网络（multi-scale residual learning-based single image dehazing network via image decomposition, MSRL-DehazeNet）架构。Chen等[9]提出一种自适应的图像块选择网络（patch map selection network, PMS-Net）来实现有效去雾，解决了传统雾霾去除方法可能导致的过饱和及颜色失真问题。Li等[10]利用卷积神经网络（convolutional neural network, CNN）构建端到端的一体化图像去雾网络（all-in-one dehazing network, AOD-Net），该网络作为轻量级网络，虽然可以有效降低网络的运行时间，但去雾能力并不理想。Mei等[11]提出一种采用编-解码结构的渐进式特征融合网络（progressive feature fusion network, PFFNet）来实现端到端的图像去雾，PFFNet虽然可以实现有效的去雾，但网络中特征维度较大，导致模型整体偏大。Wang等[12]认为雾天的大气光主要影响YCrCb空间的亮度分量，并据此提出基于大气光照先验的网络（network based on atmospheric illumination prior, AIPNet）来识别有雾区域。Song等[13]提出一种新的排序CNN，构造新的排序层来扩展CNN的结构，从而同时捕获有雾图像的统计属性和结构属性。Zheng等[14]通过使用3个CNN网络分别提取特征、学习模型及融合结果，达到在单个GPU上实现4K图像的实时去雾的效果。为解决图像中存在的非均匀雾问题，Chen等[15]提出一种包含超分辨率方法和知识迁移方法的模型，有效解决了不均匀雾的去除问题。Li等[16]认为现有依赖物理模型的方法大多无法获得主观真实的去雾结果，因此，利用神经网络之间的对抗博弈来获得主观效果更加真实的复原图像。Kim等[17]基于像素级Wasserstein自编码器提出一种可以生成不同风格去雾图像的高生成去雾方法，可应用于低光增强、夜间去雾及水下图像增强等领域。

基于深度学习的方法虽然在图像去雾领域大多表现良好，但有些方法为了追求更好的复原效果，增大了网络的尺寸，导致网络消耗过多计算及内存资源，同时在去雾的过程中没有考虑有雾图像自身的特点，存在去雾不彻底的问题。为此，本文结合有雾图像自身的特点和深度网络的学习能力，提出一种基于暗通道先验引导的图像去雾网络（dark channel prior-guided image dehazing network, DCPDNet），在保证网络轻量型及运行速度较快的同时实现较好的图像复原效果。该网络主要采用端到端的残差结构，首先，针对浅层特征，利用构建的2个特征增强模块（feature enhancement block, FEB）来增加图像空间特征的相关性，并实现特征的增强；然后，利用暗通道先验理论构建引导图，通过构建的特征校正模块（feature correction block, FCB），将网络学习的注意力引导至图像有雾区域，精细化所学习到的雾信息；最后，重建具有良好视觉效果的去雾图像。为验证本文模型的有效性，做了大量的实验，并与近年来的先进方法在主观和客观评价2个方面进行了对比。实验结果表明，本文模型在保持模型轻量型及运行速度较快的情况下，实现了更好的去雾效果。

1 本文方法

为实现深度框架中更关注有雾区域特征的学习，本文结合暗通道先验理论构建了DCPDNet模型来实现图像的有效去雾。该网络结构是一个残差网络，所学习的残差部分相当于有雾图像成分，有雾图像特征学习越精确，去雾就越彻底。图1展示了本文提出的DCPDNet整体框架。首先，根据暗通道先验，从有雾图像中提取粗略的有雾图像暗通道 $\boldsymbol{I}_{\text{dark}}$，作为后期网络学习的引导图；其次，通过2个卷积层提取有雾图像的浅层特征，并利用构建的Dou-conv块将特征图的值进行归一化，归一化后的初始特征图 $\boldsymbol{F}_0$ 被送入2个FEB模块中进行图像特征的增强；然后，将增强的特征送入FCB模块中，并以粗略的有雾图像暗通道 $\boldsymbol{I}_{\text{dark}}$ 作为引导图，来构建透射率模块，引导网络更关注有雾区域的特征学习，进一步精细化有雾图像特征；最后，利用包含像素级注意力（pixel-level attention, PA）模块[18]的残差操作获得去雾后的特征图，并利用3个卷积层进行特征整合，输出最终的去雾结果。

1.1 特征增强模块

在深度学习方法中，通常会采用通道注意力和空间注意力来选择性地增强图像的特征。通道注意力对每个特征图采用一个加权系数，空间注意力对所有特征图采用相同的加权矩阵，这种加权方式并没有考虑每个特征图内部空间信息的相关性。而自然语言处理中的自注意力机制则更擅长捕捉数据或特征的内部相关性。因此，为增加图像特征空间信息的相关性，基于自注意力机制思想，本文构建了一种FEB模块，在2个尺度上对图像特征进行增强，模块整体结构如图2所示。

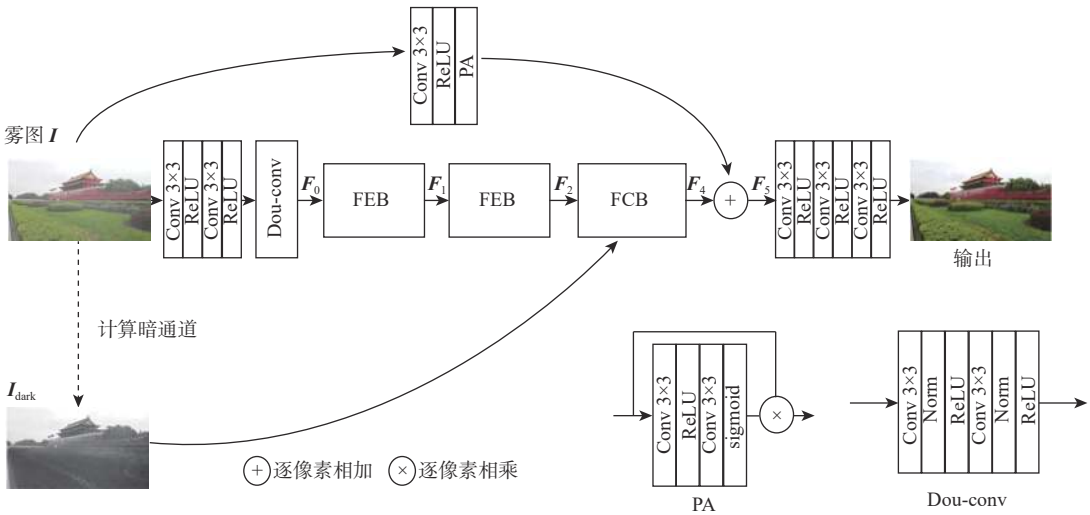


图 1　DCPDNet 整体框架

Fig. 1　Overall framework of DCPDNet

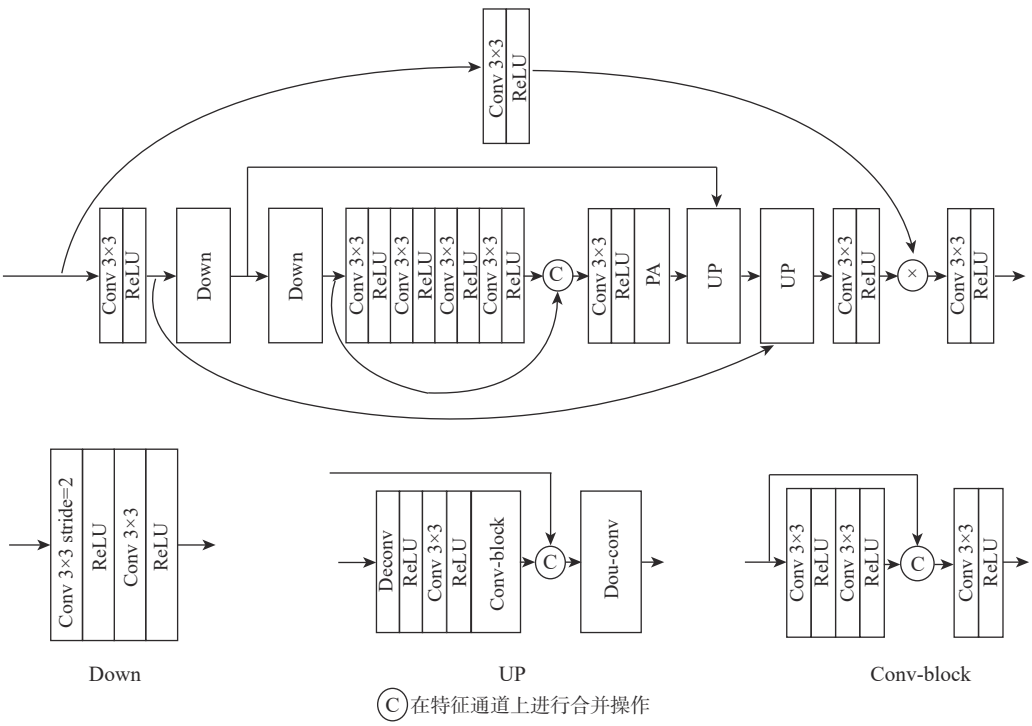


图 2　FEB 结构

Fig. 2　Structure of FEB

首先, 初始的特征图 $\boldsymbol{F}_0$ 通过 1 个卷积层和 2 个下采样的 Down 模块, 获得原始输入特征尺寸 1/4 大小的特征图, 目的在于增大卷积核的感受野。Down 模块中, 使用卷积核为 3×3、步长 stride 为 2 的卷积进行下采样。其次, 使用 4 个卷积层进行特征提取, 为减少信息丢失, 将第 2 次下采样后的特征图与卷积输出的特征图进行合并, 同时使用卷积操作对特征图进行降维以减少网络的复杂度。降维后的特征图被送入空间注意力模块(PA 模块), 使网络更加集中对感兴趣区域的学习。然后, 2 个上采样模块(UP 模块)对特征进行 2 次上采样, 进一步细化特征和恢复特征图的尺寸。UP 模块主要由卷积、反卷积、卷积块 Conv-block 及 Dou-conv 组成, 其中, Conv-block 模块主要由 2 个卷积层和 1 个跳转连接组成。Dou-conv 模块如图 1 右下角所示, 其由 2 个卷积层组成, 每个卷积层后增加了一个归一化操作(instance normalization, Norm), 将特征图的值约束在 0～1 区间。为实现浅层信息的复用, 在 UP 模块中, 将下采样后的特征与 Conv-block 模块的输出进行合并, 合并后的特征再经过一个 Dou-conv 模块进行特征整合和归一化。最后, 由于卷积过程中特征图的值被归一化, 将经过 2 个 UP 模块的特征图与经过对 $\boldsymbol{F}_0$ 进行一次卷积细化的特征图相乘, 来对提取的特征进行

增强,并增加特征图空间上的相关性。前后2个FEB模块输出被记为$\boldsymbol{F}_1$和$\boldsymbol{F}_2$,模块中卷积核大小均为3×3,每个卷积层使用32个卷积核,激活函数均使用ReLU函数。

1.2 特征校正模块

为使网络学习的特征更关注图像的有雾区域,本文提出FCB模块,利用暗通道先验构建引导图,进一步校正和精细化所学习到的特征。图3展示了FCB模块的结构。FCB模块的前半部分为特征提取,其结构与FEB模块相似。这2部分结构的区别在于:FCB模块的前半部分在浅层特征传递时采用了PA模块,主干结构特征提取最后也采用一个PA模块,其目的是使网络学习的特征更集中在感兴趣区域。FCB模块的后半部分则针对FCB模块前半部分提取的特征,暗通道先验被利用来将网络学习注意力引导至图像的有雾区域,达到特征的进一步精细化及对特征进行校正的目的。

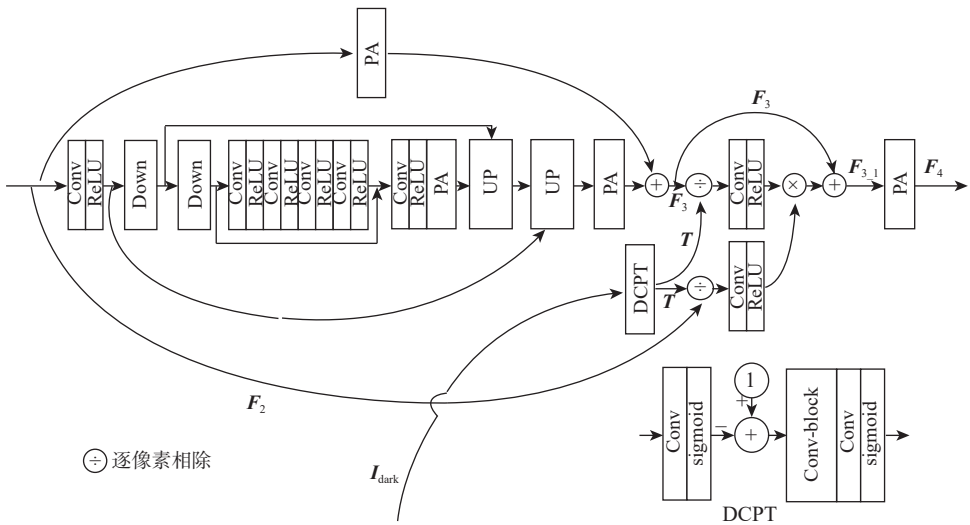


图3 FCB结构

Fig. 3 Structure of FCB

FCB模块的后半部分先利用暗通道先验透射(dark channel prior transmission, DCPT)模块和引导图$\boldsymbol{I}_{\text{dark}}$来学习计算透射率$\boldsymbol{T}$,再与主干网络提取的特征进行除法操作,来放大有雾区域的特征值,使网络更加关注目标区域,起到对特征进一步细化及校正的作用。DCPT模块的构建是基于暗通道先验中透射率计算的思想,其公式如下:

$$T(X)=1-\omega\min_{c\in\{R,G,B\}}\left(\min_{y\in\Omega(x)}\frac{\boldsymbol{I}^c(X)}{\boldsymbol{A}^c}\right)\approx 1-\frac{\omega}{A_{\min}}\min_{y\in\Omega(x)}(\boldsymbol{I}_{\text{dark}}(y)) \tag{1}$$

式中:X为像素的空间坐标;$\boldsymbol{I}$为有雾图像;$\boldsymbol{A}$为大气光值;ω为超参数,取值为0.95;$A_{\min}$为$\boldsymbol{A}$的最小值;$\Omega(x)$表示以像素点x为中心的窗口。

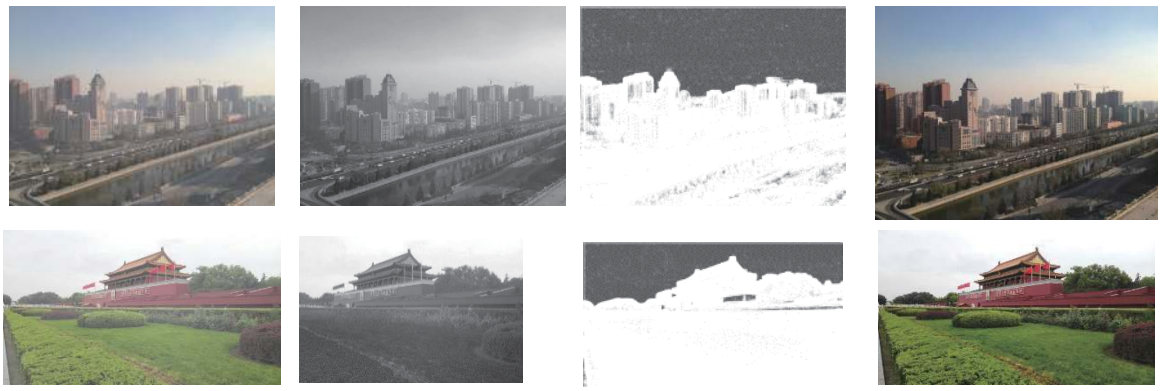
式(1)中第2项可看成对有雾图像RGB三通道采用最小值滤波求取最小值的结果,并乘以常数ω后再进行取反。其中,对RGB三通道求最小值得到粗略的有雾图像暗通道$\boldsymbol{I}_{\text{dark}}$,也即网络引导图。但是,直接对$\boldsymbol{I}_{\text{dark}}$使用最小值滤波容易导致伪边缘现象[19],因此,本文选择使用网络卷积替代最小值滤波。普通卷积操作是对特征进行滤波加偏置的过程,同时滤波器中的参数可训练,该滤波器不仅可以表现得比简单的最小值滤波更好,还可以对特征做进一步细化。因此,本文使用式(2)代替式(1)计算粗略透射率$\boldsymbol{T}_{\text{rough}}$:

$$\boldsymbol{T}_{\text{rough}}=\boldsymbol{O}-\text{Conv}_{3\times3}(\boldsymbol{I}_{\text{dark}}) \tag{2}$$

式中:$\boldsymbol{O}$为元素全为1的矩阵;$\text{Conv}_{3\times3}()$表示卷积层,包含3×3的卷积和ReLU操作。

根据式(2),DCPT模块可以被构建为图3右下方的结构,$\boldsymbol{I}_{\text{dark}}$先执行式(2)的操作,即通过一个卷积层,并对特征图的值取负数,再与1做加法,获得初始的透射率图$\boldsymbol{T}_{\text{rough}}$。再使用卷积块Conv-block、卷积层和sigmoid激活函数对初始的透射率图$\boldsymbol{T}_{\text{rough}}$进行细化,得到最终的$\boldsymbol{T}$。为方便后续计算,获得更多的特征细节,本文设置DCPT模块输出的$\boldsymbol{T}$为32通道。图4为2张有雾图像得到的透射率,选择其中第2个通道的特征图进行展示。可以观察到,DCPT模块计算得到的透射率可以有效区分出天空区域。

为避免网络特征提取过程中特征信息的丢失,本文将FCB模块的输入特征图$\boldsymbol{F}_2$和前半部分学习的特征图$\boldsymbol{F}_3$都除以$\boldsymbol{T}$来放大有雾区域的特征值,再分别通过一个卷积层进行特征细化,并将2个引



(a) 有雾图像　(b) 暗通道 $\boldsymbol{I}_{\text{dark}}$ (引导图)　(c) 透射率 $\boldsymbol{T}$　(d) 去雾结果

图 4　图像去雾结果及中间结果

Fig. 4　Image dehazing results and intermediate results

导后的特征图进行相乘, 进一步实现有雾区域特征的增强, 最终通过加上 $\boldsymbol{F}_3$ 来完整地保留已学习到的特征。上述操作利用式(3)计算得到特征 $\boldsymbol{F}_{3_1}$:

$$\boldsymbol{F}_{3_1}=\text{Conv}_{3\times3}\left(\frac{\boldsymbol{F}_2}{\boldsymbol{T}}\right)\cdot\text{Conv}_{3\times3}\left(\frac{\boldsymbol{F}_3}{\boldsymbol{T}}\right)+\boldsymbol{F}_3 \tag{3}$$

图 4 展示了有雾图像、引导图、求取的一幅透射率特征图和去雾结果。从图 4 的透射率特征图可以观察到, 有雾处的透射率值明显小于清晰处, 由于使用了 sigmoid 激活函数, $\boldsymbol{T}$ 的值均在(0,1)之间。因此, 特征图除以 $\boldsymbol{T}$ 可以将浓雾区域的特征值放大, 使主干网络的学习更加集中在图像有雾的区域, 实现对雾的有效学习。

由于本文网络是一个基于残差学习的端到端网络, 主干网络主要进行有雾图像的学习, 在从有雾图像转变到特征 $\boldsymbol{F}_{3_1}$ 的过程中, 经历多次的上下采样, 抛去了部分细节, 专注于图像中雾的学习。最终细节的补充则使用残差学习, 但 $\boldsymbol{F}_{3_1}$ 中并不是所有学习到的特征都是有用的, 因此, 最终将 $\boldsymbol{F}_{3_1}$ 送入 PA 模块中做最后的筛选, 得到 FCB 模块的输出 $\boldsymbol{F}_4$。为得到更好的复原效果, 本文采用残差结构构建网络, 选择直接从原始有雾图像中补充需要的特征细节, 利用一个跳转链接, 将有雾图像送入一个单独的卷积层和 PA 模块来实现有选择的抽取特征, 与学习的特征图 $\boldsymbol{F}_4$ 相加得到 $\boldsymbol{F}_5$, 再通过 3 个卷积层进行特征整合得到清晰的去雾图像。

1.3　损失函数

对于损失函数的选取, 本文选择平滑 L_1 函数及感知损失函数作为损失函数对网络进行训练。平滑 L_1 函数可以更直观地计算出网络输出与输入之间的差距, 而感知损失函数则更擅长于衡量潜在特征之间的不同。平滑 L_1 函数可以计算网络输出结果与 GroundTruth 之间的像素差值, 同时能够避免由于零点处不可导而引发的无法产生稳定解的问题。平滑 L_1 函数可表示为

$$L_1=\frac{1}{N}\sum_{i=1}^{N}G_{\text{s}}(I_{\text{O}i}-J_i) \tag{4}$$

$$G_{\text{s}}(x)=\begin{cases}0.5x^2 & |x|<1\\ |x|-0.5 & \text{其他}\end{cases} \tag{5}$$

式中: I_{O} 为网络输出; J 为 GroundTruth; N 为图像总像素数量; i 为第 i 个像素; x 为网络输出与 GroundTruth 之间的差值。

感知损失函数通过 VGG16 网络提取特征, 本文使用 ImageNet[20] 预训练的模型分别提取网络输出及 GroundTruth 在 VGG16 网络中 3 个阶段的输出特征图, 使用均方差计算两者之间的差值。感知损失函数 L_{p} 可表示为

$$L_{\text{p}}=\sum_{j=1}^{3}\frac{1}{C_jH_jW_j}\left\|\boldsymbol{V}_j(I_{\text{O}})-\boldsymbol{V}_j(J)\right\|_2^2 \tag{6}$$

式中: $\boldsymbol{V}$ 为从 VGG16 网络中提取的特征; j 为所提取的第 j 个特征层; C、H、W 分别为特征图尺寸中的通道、长、宽。

因此, 整体损失函数表示为

$$L=L_1+\lambda L_{\text{p}} \tag{7}$$

式中: λ 为超参数, 调整感知损失所占比重, 本文在实验中将其设置为 0.04。

2　实验与结果分析

2.1　实验设置

本文 DCPDNet 参数设置如下: 网络训练中, 初始学习率设置为 0.001, 每 20 个 epoch 将学习率衰减至原来的一半, 共训练 100 个 epoch。优化器使用 Adam 优化器, 参数均使用默认参数值。为降低训练计算量, BatchSize 设置为 5, 同时训练数据投入网络前均随机裁剪至 224 像素×224 像素大小。

本文方法在 RESIDE 数据集[21] 上进行训练,

室外的训练集包含310 000张合成有雾图像，对应8 970张GroundTruth图像。为保持足够信息量的同时减少训练量，本文对每张GroundTruth图像都在对应的合成有雾图像中随机选择一张来组成训练集，最终选择10 000张合成有雾图像对网络进行训练。对于室内训练集，则使用RESIDE-Standard数据集[21]中的训练数据。测试图像选择SOTS数据集[21]中的500张室内测试图像、500张室外测试图像及HSTS测试数据集对本文方法进行测试。

为验证本文方法的去雾效果，将其与2个传统方法进行对比，即Meng[22]和Berman[23]等的方法，并与5个深度学习的方法进行对比，即AOD-Net[10]、PFFNet[11]、EPDN[24]、Light-DehazeNet[25]、RefineD-Net[26]。

本文方法在Windows10系统下进行训练和测试，使用Python编程语言及PyTorch深度学习框架来搭建网络框架和实现算法，硬件环境为InterI CoreI i7-9700K CPU @ 3.60 GHz, 32.0 GB RAM, GeForce RTX 2 080 Ti。

2.2 主观效果

从主观效果上比较本文方法与其他方法的性能。图5展示了2组室外有雾图像及对比方法的去雾结果。可以观察到，Meng等[22]方法的结果存在比较严重的偏色和噪声，如有雾图像1的结果中桌上的面饼在去雾之后出现了偏色问题，有雾图像2的结果中天空区域包含很多噪声和伪影，也呈现出了严重的偏色；Berman等[23]方法的去雾结果中还存有残留的雾，如有雾图像1中右侧框中还存在薄雾，有雾图像2中的立柱之间同样能观察到薄雾；AOD-Net[10]明显去雾能力不强，从2张有雾图像的结果中都可以观察到还存在大量残余的雾；PFFNet[11]的去雾结果则存在偏暗的问题，有雾图像1和有雾图像2的结果中右侧区域较暗，导致细节纹理不清晰；EPDN[24]的去雾结果同样存在细节处偏暗的问题，2张有雾图像的右侧对比GroundTruth都存在过暗的情况，导致细节模糊不清；Light-DehazeNet[25]的去雾结果则存在残留的雾和偏色问题，如有雾图像的结果还存在大量的薄雾，2组图像中的地板及人物的服装都能观察到不同程度的偏



图5 室外有雾图像去雾结果对比

Fig. 5 Comparison of dehazing results of outdoor hazy images

色现象; RefineDNet[26] 的结果则存在伪影的情况, 在 2 组有雾图像的结果中, 在地面区域都可观察到不同程度的伪影; 本文方法的结果和 Ground Truth 最接近, 没有出现伪影、色偏等现象, 更好地复原了图像的细节。

图 6 展示了 2 组室内有雾图像及对比方法的去雾结果。从有雾图像 3 和有雾图像 4 可以观察到, Meng[22] 及 Berman[23] 等方法获得的结果都存在偏色情况; AOD-Net[10] 及 Light-DehazeNet[25] 则存在去雾不彻底的问题, 从有雾图像 3 可以明显观察到, 图像中间区域还残留大量的雾; PFFNet[11]、EPDN[24] 及 RefineDNet[26] 则都存在不同程度的过增强、去雾不彻底的现象, 如有雾图像 3 中的桌面区域的颜色比 GroundTruth 更亮, 有雾图像 4 中凳子区域还存在明显的残留雾; 本文方法在 2 幅有雾图像上的复原效果都最好, 没有出现偏色和去雾不彻底等现象。综合主观实验结果可以证明本文方法在去雾方面的有效性。



图 6　室内有雾图像去雾结果对比

Fig. 6　Comparison of dehazing results of indoor hazy images

2.3　客观评价

为更客观地评价本文方法, 在 3 个测试数据集上分别使用峰值信噪比 PSNR、结构相似性 SSIM 及均方误差 MSE 指标来进行评价。PSNR 是用来衡量图像噪声水平和图像质量的客观评价指标, 其值越大, 表示图像质失真越小。SSIM 是用于衡量 2 个图像在亮度、对比度和结构方面相似程度的指标, 其值越大, 表示复原图像的质量越好。MSE 是一种常用的衡量复原图像与实际观测图像之间差异的指标, 其值越小, 表示复原图像的质量越好。表 1 展示了本文方法及对比方法在 SOTS 室外数据

表 1　SOTS 室外数据集测试结果

Table 1　Test results of SOTS outdoor dataset

方法	PSNR	SSIM	MSE
Meng[22]	15.78	0.821 5	0.032 7
Berman[23]	18.15	0.846 5	0.021 5
AOD-Net[10]	19.59	0.816 3	0.012 2
PFFNet[11]	21.05	0.796 8	0.012 7
EPDN[24]	20.57	0.872 7	0.013 1
Light-DehazeNet[25]	18.81	0.871 8	0.015 2
RefineDNet[26]	21.10	0.908 1	0.010 3
本文	31.06	0.974 9	0.001 0

集上的测试结果。可以观察到，本文方法在3个指标上都是最优，其次是RefineDNet[26]，传统方法中Meng等[22]方法在SOTS室外数据集上的表现较差，PSNR和MSE指标均为最低，SSIM指标倒数第三，可能与传统方法局限性较大有关。表2展示了本文方法及对比方法在HSTS数据集上的评价结果。同样也是本文方法最优，EPDN[24]获得了次优的PSNR值，RefineDNet[26]获得的PSNR值排序第三，而SSIM和MSE值排序第二。传统方法Meng等[22]方法获得的3种指标都是最差的。表3展示了SOTS室内数据集的结果。可观察到，本文方法在3个指标上同样还是最优，其次是EPDN[24]，评价结果较差的仍然是传统的Meng等[22]方法。从对比结果可以看出，与其他方法相比，本文方法的去雾评价指标始终能达到最优，进一步验证了方法的有效性和优越性。

表2 HSTS数据集测试结果

Table 2 Test results of HSTS dataset

方法	PSNR	SSIM	MSE
Meng[22]	15.32	0.742 3	0.033 2
Berman[23]	17.60	0.779 0	0.024 1
AOD-Net[10]	18.97	0.789 9	0.014 5
PFFNet[11]	20.06	0.769 4	0.020 4
EPDN[24]	21.74	0.855 7	0.009 9
Light-DehazeNet[25]	19.75	0.886 1	0.011 1
RefineDNet[26]	21.54	0.903 4	0.008 6
本文	32.79	0.979 0	0.000 5

表3 SOTS室内数据集测试结果

Table 3 Test results of SOTS indoor dataset

方法	PSNR	SSIM	MSE
Meng[22]	17.26	0.787 5	0.022 1
Berman[23]	17.41	0.778 1	0.021 9
AOD-Net[10]	17.63	0.793 0	0.022 7
PFFNet[11]	24.76	0.887 7	0.003 9
EPDN[24]	25.13	0.913 7	0.004 0
Light-DehazeNet[25]	19.69	0.837 7	0.013 8
RefineDNet[26]	20.56	0.835 9	0.011 0
本文	31.21	0.972 3	0.000 9

2.4 方法时间对比分析

为更全面地对比本文方法及其他方法，分别计算了各方法的运行时间。表4展示了各方法在SOTS室外数据集上的平均运行时间。可以观察到，运行时间最短的为AOD-Net[10]，平均用时12 ms，其次为PFFNet[11]，用时29 ms，本文方法在平均运行时间上与前4种方法的运行时间差距相对较小。传统方法中，运行时间最长的为Berman等[23]方法，用时4 483 ms，同时也是对比方法中用时最长的方法。对比方法中，深度学习方法用时最长的为RefineDNet[26]，用时757 ms。然而，结合去雾实验结果来看，耗时排名前三的深度学习方法中，PFFNet[11]容易导致复原图像颜色偏暗、过增强的问题，AOD-Net[10]和Light-DehazeNet[25]都存在去雾不彻底的问题。因此，综合来看，本文方法在图像复原质量和计算效率之间达到了一个较好的平衡。

表4 平均运行时间对比

Table 4 Average runtime comparison

方法	时间/ms
Meng[22]	1 304
Berman[23]	4 483
AOD-Net[10]	12
PFFNet[11]	29
EPDN[24]	84
Light-DehazeNet[25]	71
RefineDNet[26]	757
本文	211

2.5 方法模型大小对比分析

表5展示了本文方法及对比方法的模型大小结果。由于Meng等[22]方法和Berman等[23]方法属于传统方法，并未用到学习模型，未对这2个方法进行对比。从表5中可以观察到，模型尺寸最小的为AOD-Net[10]，大小为0.016 6 MB，其次为Light-DehazeNet[25]，大小为0.120 1 MB，第3为本文方法，大小为3.985 MB。对比方法中，模型最大的为RefineDNet[26]，大小为261.6 MB。模型较小的AOD-Net[10]和Light-DehazeNet[25]都存在去雾不彻底的问题，原因是过小的模型可能存在学习能力不足的问题，无法实现高质量的图像复原。而本文方法的模型虽然不是最小，但在对比方法中仍然属于模型较小的行列，同时运行速度也较快，复原结果的主观效果及客观评价指标都明显优于其他对比方法，因此，本文方法可在保证模型运行速度和占用内存不大的情况下获得更优的复原结果。

表5 模型大小对比

Table 5 Comparison of model size

方法	模型大小/MB
AOD-Net[10]	0.016 6
PFFNet[11]	57.05
EPDN[24]	66.31
Light-DehazeNet[25]	0.120 1
RefineDNet[26]	261.6
本文	3.985

2.6 消融实验

本节对DCPDNet中的FEB模块和FCB模块进

行消融实验，来验证 2 个模块的有效性及对整个网络的贡献。表 6 展示了 SOTS 室外数据集上 4 个方法的平均实验结果，方法 1 为 Baseline，是在 DCPDNet 中同时去掉了 FEB 模块和 FCB 模块；方法 2 为在 DCPDNet 中去除 FCB 模块，只使用 FEB 模块；方法 3 为去除 FEB 模块，只使用 FCB 模块。可以看出，无论是去除 FEB 模块还是 FCB 模块，都会导致评价结果下降，完整的 DCPDNet 的结果是最优的。由此可以说明，本文提出的 FEB 模块和 FCB 模块都能有效提升网络性能，对图像复原起到实际作用。

表 6　消融实验

Table 6　Ablation experiments

Baseline	FEB	FCB	SSIM	PSNR
√	×	×	0.957 9	27.17
√	√	×	0.967 3	29.24
√	×	√	0.973 1	29.63
√	√	√	0.974 9	31.06

3　结　论

针对目前深度学习的去雾方法存在的去雾不彻底的问题，本文提出一种基于暗通道先验引导的图像去雾网络 DCPDNet。

1）构建特征增强模块，利用具有丰富语义特征的深层特征图作为权重图，来增加图像空间特征的相关性，实现图像浅层特征的增强。

2）设计基于引导图的特征校正模块，利用暗通道先验构建引导图，引导网络学习包含雾信息的精细化特征图。

3）将本文模型与先进的去雾模型进行对比，并对模型中构建的模块做了消融实验比较，来验证各模块的有效性。实验结果表明，本文方法获得了更好的视觉效果及更优的客观评价结果，同时，具有轻量型及运行速度较快的优势。

未来计划考虑真实有雾图像除了雾霾因素，通常存在噪声、夜间曝光不足等因素影响产生的降质问题，将提出的去雾方法应用于实际的夜间图像去雾，以生成更高视觉质量的结果。

参考文献（References）

[1] SALAZAR-COLORES S, CABAL-YEPEZ E, RAMOS-ARREGUIN J M, et al. A fast image dehazing algorithm using morphological reconstruction[J]. IEEE Transactions on Image Processing, 2019, 28(5): 2357-2366.

[2] HE K M, SUN J, TANG X O. Single image haze removal using dark channel prior[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(12): 2341-2353.

[3] GAO Y Y, HU H M, LI B, et al. Detail preserved single image dehazing algorithm based on airlight refinement[J]. IEEE Transactions on Multimedia, 2019, 21(2): 351-362.

[4] WANG W C, YUAN X H, WU X J, et al. Fast image dehazing method based on linear transformation[J]. IEEE Transactions on Multimedia, 2017, 19(6): 1142-1155.

[5] ZHU M Z, HE B W, WU Q. Single image dehazing based on dark channel prior and energy minimization[J]. IEEE Signal Processing Letters, 2018, 25(2): 174-178.

[6] LIU Q, GAO X B, HE L H, et al. Single image dehazing with depth-aware non-local total variation regularization[J]. IEEE Transactions on Image Processing, 2018, 27(10): 5178-5191.

[7] YANG Y, WANG Z W. Haze removal: Push DCP at the edge[J]. IEEE Signal Processing Letters, 2020, 27: 1405-1409.

[8] YEH C H, HUANG C H, KANG L W. Multi-scale deep residual learning-based single image haze removal via image decomposition [J]. IEEE Transactions on Image Processing, 2020, 29: 3153-3167.

[9] CHEN W T, DING J J, KUO S Y. PMS-Net: Robust haze removal based on patch map for single images[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2019: 11681-11689.

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

[11] MEI K F, JIANG A W, LI J C, et al. Progressive feature fusion network for realistic image dehazing[C]//Proceedings of the Asian Conference on Computer Vision. Berlin: Springer, 2018: 203-215.

[12] WANG A N, WANG W H, LIU J L, et al. AIPNet: Image-to-image single image dehazing with atmospheric illumination prior[J]. IEEE Transactions on Image Processing, 2018, 28(1): 381-393.

[13] SONG Y F, LI J, WANG X G, et al. Single image dehazing using ranking convolutional neural network[J]. IEEE Transactions on Multimedia, 2018, 20(6): 1548-1560.

[14] ZHENG Z R, REN W Q, CAO X C, et al. Ultra-high-definition image dehazing via multi-guided bilateral learning[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 16180-16189.

[15] CHEN T Y, FU J H, JIANG W T, et al. SRKTDN: Applying super resolution method to dehazing task[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2021: 487-496.

[16] LI Y N, LIU Y H, YAN Q X, et al. Deep dehazing network with latent ensembling architecture and adversarial learning[J]. IEEE Transactions on Image Processing, 2021, 30: 1354-1368.

[17] KIM G, PARK S W, KWON J. Pixel-wise Wasserstein autoencoder for highly generative dehazing[J]. IEEE Transactions on Image Processing, 2021, 30: 5452-5462.

[18] QIN X, WANG Z L, BAI Y C, et al. FFA-Net: Feature fusion attention network for single image dehazing[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI, 2020, 34(7): 11908-11915.

[19] LU Z W, LONG B Y, YANG S Q. Saturation based iterative approach for single image dehazing[J]. IEEE Signal Processing Letters, 2020, 27: 665-669.

[20] JOHNSON J, ALAHI A, LIF F. Perceptual losses for real-time style

transfer and super-resolution[C]//Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2016: 694-711.

[21] LI B Y, REN W Q, FU D P, et al. Benchmarking single-image dehazing and beyond[J]. IEEE Transactions on Image Processing, 2018, 28(1): 492-505.

[22] MENG G F, WANG Y, DUAN J Y, et al. Efficient image dehazing with boundary constraint and contextual regularization[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2013: 617-624.

[23] BERMAN D, TREIBITZ T, AVIDAN S. Non-local image dehazing[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 1674-1682.

[24] QU Y Y, CHEN Y Z, HUANG J Y, et al. Enhanced pix2pix dehazing network[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2019: 8160-8168.

[25] ULLAH H, MUHAMMAD K, IRFAN M, et al. Light-DehazeNet: Anovel lightweight CNN architecture for single image dehazing[J]. IEEE Transactions on Image Processing, 2021, 30: 8968-8982.

[26] ZHAO S Y, ZHANG L, SHEN Y, et al. RefineDNet: A weakly supervised refinement framework for single image dehazing[J]. IEEE Transactions on Image Processing, 2021, 30: 3391-3404.

Image dehazing network based on dark channel prior guidance

HUANG Shuying[1], XIA Yukun[2], YANG Yong[3,*], WAN Weiguo[4], QIU Genying[2]

(1. School of Software, Tiangong University, Tianjin 300387, China;
2. School of Information Management, Jiangxi University of Finance and Economics, Nanchang 330032, China;
3. School of Computer Science and Technology, Tiangong University, Tianjin 300387, China;
4. School of Software and Internet of Things Engineering, Jiangxi University of Finance and Economics, Nanchang 330032, China)

Abstract: Currently, the majority of dehazing techniques that utilize deep learning focus on directly acquiring the mapping relationship between a foggy image and a non-fog image.Because of the lack of combination with the characteristics of fog images, there are problems such as inaccurate detection of fog information and incomplete dehazing. In order to address the aforementioned issues, this study introduces a novel approach called the dark channel prior-guided image dehazing network (DCPDNet), which operates in an end-to-end manner. First, the shallow features of the input foggy image are extracted by several convolution layers. Secondly, two feature enhancement blocks (FEB) are constructed to enhance the spatial features of the image, which can enhance the image features on two scales, that is, the deep feature map is used to enhance the semantic features, and the shallow feature map is used to enhance the image details. Thirdly, in order to make the extracted features pay more attention to the fog area, a feature correction block (FCB) based on the guidance map is designed by considering the imaging characteristics of fog in the fog image. The FCB uses the dark channel theory to build a guidance map to guide the attention of network learning to the fog area, and further refine and correct the extracted deep feature map. Finally, by using the skip connection of the residual structure, the enhanced shallow features are used to supplement the details lost in the network, and the image after dehazing is reconstructed through several convolution operations. Multiple trials have demonstrated that DCPDNet is capable of achieving a satisfactory dehazing effect while maintaining a lightweight model and quick execution. Compared with some advanced dehazing methods proposed in recent years, DCPDNet has better performance in terms of efficiency, subjective visual perception and objective evaluation results.

Keywords: image dehazing; dark channel prior; guidance map; residual learning; feature enhancement

Received: 2022-09-04; **Accepted**: 2022-12-16; **Published Online**: 2023-01-05 09:36

URL: link.cnki.net/urlid/11.2625.V.20230104.1422.001

Foundation items: National Natural Science Foundation of China (61862030,62072218,62261025); Jiangxi Provincial Natural Science Foundation (20192ACB20002,20192ACBL21008); Postdoctoral Research Projects of Jiangxi Province, China (2020KY44)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0711

舵机电动负载模拟器迭代学习控制

刘晓琳*，李晋恺
(中国民航大学 电子信息与自动化学院，天津 300300)

摘 要：针对舵机电动负载模拟器 (ELSSG) 受到多余力矩干扰的问题，提出了一种结合迭代学习控制与误差符号鲁棒积分控制 (RISE) 的复合控制器。建立了 ELSSG 的数学模型，分析了多余力矩的产生机理。在硬件结构上，引入了金属-橡胶缓冲弹簧以提高系统稳定性及加载精度；在控制策略上，基于误差符号鲁棒积分控制与自适应遗忘因子改进的引导信号迭代学习控制设计了复合控制器，并证明了该迭代学习控制的收敛条件。通过仿真实验证明了复合控制器与传统迭代学习控制方法、PID 控制方法相比，能够更加有效地消除多余力矩干扰，提高 ELSSG 加载力矩输出精度。

关 键 词：舵机电动负载模拟器；多余力矩；迭代学习控制；金属-橡胶缓冲弹簧；遗忘因子
中图分类号：V249.1; TP273
文献标志码：A **文章编号**：1001-5965(2024)09-2727-12

舵机电动负载模拟器 (electric load simulator of aircraft steering gear, ELSSG) 是一种用来测试飞机舵机性能指标的半实物仿真设备[1-2]，其使用加载电机输出的力矩来模拟飞机飞行中舵机受到的空气铰链力矩，从而测试在该力矩影响下舵机的控制性能。与传统的破坏性全实物实验方法相比，使用 ELSSG 在地面实验室条件下进行的半实物实验破坏性小、成本低，更易进行重复实验，从而获得更全面的实验数据[3-5]。

ELSSG 的加载电机与受测舵机直接相连，当加载电机向舵机输出模拟力矩时，舵机的主动运动会产生干扰加载电机输出的力矩，该力矩被称为多余力矩[6-8]。多余力矩引起的加载力矩误差会严重影响舵机测试实验数据的准确性和可靠性。因此，如何补偿多余力矩引起的输出力矩误差是该领域亟待解决的研究课题，对于降低舵机研究成本、缩短研发周期具有重要意义。

多余力矩的补偿方法主要有基于结构不变性的前馈控制方法[9]、改进 PID 控制方法[10-13]、滑模控制法[14-15] 及迭代学习控制 (iterative learning control, ILC) 法[16-18] 等。由于多余力矩由舵机主动运动产生，研究者们提出了根据舵机角速度信号对其进行前馈补偿的控制方法，虽然结构简单，但补偿效果较差。更进一步地，文献 [9] 提出结合舵机位置信号与角速度信号进行前馈控制，降低多余力矩干扰。与前馈控制方法相比，PID 控制方法的补偿效果更好。文献 [10] 在改进 PID 控制结构的基础上，采用积分先行的 PID 控制设计了复合控制器，配合小脑模型神经网络前馈控制降低了跟踪误差，但仍需手动整定 PID 控制参数，同时需要重新训练前馈控制网络。文献 [12] 提出采用群智能算法对 PID 控制参数进行寻优的控制策略，无须手动调参。对于 PID 控制存在抗扰性差的缺点，文献 [15] 设计了自抗扰反演滑模复合控制器，利用扩张状态观测器得到系统总和扰动，结合滑模控制器提高跟踪精度，实验表明，相比于传统手段，该方法能够有效降低输出力矩误差。迭代学习控制在迭代时通过不断调整控制量来提高控制精度。文献 [16] 根据 ELSSG 输出呈周期性的特点，提出了按照正弦周期补偿的迭代学习控制，即每经过一个正弦周期，便

收稿日期：2022-08-13；**录用日期**：2023-05-15；**网络出版时间**：2023-06-06 14:21
网络出版地址：link.cnki.net/urlid/11.2625.V.20230605.1445.002
基金项目：中国民航大学第十一期波音基金 (20210715113)
* **通信作者**. E-mail：caucyanjiusheng@163.com

引用格式：*刘晓琳，李晋恺. 舵机电动负载模拟器迭代学习控制 [J]. 北京航空航天大学学报，2024，50（9）：2727-2738.*
LIU X L，LI J K. Iterative learning control of electric load simulator of aircraft steering gear[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50（9）：2727-2738（in Chinese）.

调节预设输入，实验证明，该方法能够有效抑制多余力矩干扰。文献 [17] 提出了一种引导信号迭代学习控制策略，但收敛速度慢且跟踪误差大，需要结合其他控制策略才能实现对负载模拟器的精确控制。文献 [18] 提出结合 P 型迭代学习与 PID 控制的复合控制器，对存在控制时滞的电动加载系统进行了实验分析。与上述其他控制方法相比，迭代学习控制只需要系统运行中的误差信息，不会因为系统参数不准确而降低控制精确度，且其可以通过迭代对控制量进行调整，从而不断降低误差，更适用于 ELSSG 多余力矩补偿问题。

针对 ELSSG 存在的多余力矩干扰问题，本文提出一种结合迭代学习控制与误差符号鲁棒积分控制 (robust integral control of sign of error, RISE) 的复合控制方式，该控制方式不需要 ELSSG 精确参数，通过迭代更新控制量从而逐渐降低跟踪误差，实现 ELSSG 加载力矩的高精度输出。在变频率正弦信号的舵机工作信号扰动下，通过仿真实验验证了该方法的可行性和有效性。

1 舵机电动负载模拟器结构设计

飞机飞行过程中，舵面始终受到空气铰链力矩的作用，舵机需要克服该力矩才能操纵舵面调整飞机姿态。因此，在进行测试时，需要对工作中的舵机施加模拟铰链力矩来还原舵机实际工作状态。基于此，本文设计的 ELSSG 工作原理如图 1 所示。

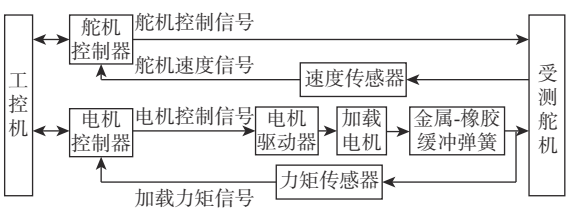


图 1 ELSSG 工作原理

Fig. 1 Working principle of ELSSG

按照力矩传递关系，ELSSG 分为工控机、加载装置、测量装置和受测舵机 4 个部分。ELSSG 工作时，2 条控制回路同时运行。工控机、加载装置与力矩传感器构成第 1 条控制回路。工控机输出预设力矩信号至电机控制器，驱动加载电机输出模拟铰链力矩，经由金属-橡胶缓冲弹簧加载至受测舵机，再通过力矩传感器测量输出力矩，并返回至工控机。工控机、受测舵机与速度传感器构成第 2 条控制回路。工控机输出舵机工作信号，利用速度传感器测量舵机工作状态。

受测舵机与加载电机直接相连。一方面，在加载电机输出加载力矩时，受测舵机主动运动会影响加载电机输出的加载力矩。另一方面，加载电机输出力矩也会影响到受测舵机主动运动。为降低两者之间的机械结构耦合，使用金属-橡胶缓冲弹簧作为受测舵机与加载电机的连接装置，以增加系统稳定性。

2 舵机电动负载模拟器数学建模

ELSSG 存在强位置干扰与复杂的机械耦合关系，因此，建立符合实际工作状况的数学模型较为困难。考虑到系统工作原理与实际应用场景，本文在现有数学模型的基础上提出以下 2 点改进之处：

1）引入金属-橡胶缓冲弹簧作为电机与舵机之间的连接装置。金属-橡胶缓冲弹簧能抑制刚性连接无法处理的突变扰动现象，减弱系统耦合，增加稳定性；与传统弹簧杆相比，其能够实现连续变刚度加载，不仅保证了实验的连续性，而且节省了实验成本。现有数学模型未考虑缓冲弹簧，本文根据力矩传递方式，重新建立了 ELSSG 的数学模型。

2）舵机工作信号使用变频率正弦信号替代固定频率正弦信号。现有研究中，通常使用固定频率的正弦信号作为舵机工作信号。由于多余力矩干扰的存在会导致舵机工作信号发生变化，使用频率周期变化的正弦信号更加符合实际工作情况。同时，变频率舵机信号引起的多余力矩更为复杂，对控制方法的稳定性要求更高，更适于验证 ELSSG 控制方法的有效性。

2.1 永磁直流力矩电机数学模型

加载电机是 ELSSG 力矩输出元件。根据加载力矩的特点，选择体积小、转矩大、能长期在堵转状态下稳定运行的永磁直流力矩电机。电机电枢回路电压平衡方程为

$$\begin{cases} U_{\mathrm{m}} = i_{\mathrm{m}} R_{\mathrm{m}} + L_{\mathrm{m}} \dot{i}_{\mathrm{m}} + E \\ E = C_{\mathrm{e}} \dot{\theta}_{\mathrm{m}} \end{cases} \tag{1}$$

式中：U_{m} 为电机电枢电压；i_{m} 为电机电枢电流；R_{m} 为电机电枢电阻；L_{m} 为电机电枢电感；E 为电机反电动势；C_{e} 为电机反电动势系数；θ_{m} 为电机旋转角度；$\dot{\theta}_{\mathrm{m}}$ 为电机旋转角速度。

在电机运行过程中，考虑到 ELSSG 所受摩擦及外部扰动，电机电磁转矩方程可以表示为

$$C_{\mathrm{m}} i_{\mathrm{m}} - T_{\mathrm{l}} - F_{\mathrm{f}} + D = J_{\mathrm{m}} \ddot{\theta}_{\mathrm{m}} + B_{\mathrm{m}} \dot{\theta}_{\mathrm{m}} \tag{2}$$

式中：C_{m} 为电机转矩常数；T_{l} 为输出力矩；F_{f} 为摩擦力矩；D 为外部扰动力矩；J_{m} 为电机转动惯量；B_{m} 为电机阻尼系数。

2.2 电机驱动器数学模型

选用脉宽调制 (pulse width modulation, PWM) 驱动器接收电机控制器传输的控制信号，输出驱动

电机所需电枢电压。由于驱动器工作频率远大于系统加载频率, 在建模时可视为理想比例环节, 即

$$K_{\mathrm{PWM}} = U_{\mathrm{m}}/u_{\mathrm{in}} \tag{3}$$

式中: u_{in} 为 ELSSG 控制信号; K_{PWM} 为驱动器比例放大系数。

2.3　金属-橡胶缓冲弹簧数学模型

理论证明, 选择合适的缓冲弹簧可以提高 ELSSG 的稳定性, 而且在实现变刚度加载的同时, 能够降低多余力矩对加载力矩的部分影响[12]。忽略弹簧质量等因素, 可以分析得出加载力矩大小与舵机及加载电机之间的转动角度差呈正比例关系。对金属-橡胶缓冲弹簧建模为

$$T_{\mathrm{l}} = K_{\mathrm{l}}(\theta_{\mathrm{m}} - \theta_{\mathrm{r}}) \tag{4}$$

式中: θ_{r} 为舵机转角; K_{l} 为金属-橡胶缓冲弹簧的弹性系数。

联立式 (1)～式 (4), 可以建立 ELSSG 数学模型, 如图 2 所示。

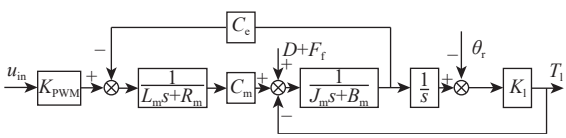


图 2　ELSSG 数学模型

Fig. 2　Mathematical model of ELSSG

为分析金属-橡胶缓冲弹簧弹性系数大小对系统工作稳定性的影响, 本文利用劳斯稳定判据对闭环系统稳定性进行判别, 并给出保证系统稳定性条件。

结合式 (1)～式 (4), 可以得到系统特征方程为

$$1+\frac{C_{\mathrm{e}}C_{\mathrm{m}}}{(L_{\mathrm{m}}s+R_{\mathrm{m}})(J_{\mathrm{m}}s+B_{\mathrm{m}})}+\frac{K_{\mathrm{PWM}}C_{\mathrm{m}}K_{\mathrm{l}}}{s(L_{\mathrm{m}}s+R_{\mathrm{m}})(J_{\mathrm{m}}s+B_{\mathrm{m}})}+\frac{K_{\mathrm{l}}}{s(J_{\mathrm{m}}s+B_{\mathrm{m}})}=0 \tag{5}$$

即

$$s^3J_{\mathrm{m}}L_{\mathrm{m}}+s^2(B_{\mathrm{m}}L_{\mathrm{m}}+R_{\mathrm{m}}J_{\mathrm{m}})+s(R_{\mathrm{m}}B_{\mathrm{m}}+C_{\mathrm{e}}C_{\mathrm{m}}+K_{\mathrm{l}}L_{\mathrm{m}})+K_{\mathrm{PWM}}C_{\mathrm{m}}K_{\mathrm{l}}+R_{\mathrm{m}}K_{\mathrm{l}}=0 \tag{6}$$

代入 ELSSG 结构参数, R_{m}=4.32 Ω, L_{m}=0.005 H, C_{e}=3.6 (V·s)/rad, J_{m}=0.12 kg·m^2, B_{m}=0.232 (Nm·s)/rad, C_{m}=7.835 Nm/A, K_{PWM}=15。

根据劳斯稳定判据, 为使闭环系统稳定, 则有

$$\begin{cases} K_{\mathrm{l}} > 0 \\ 15.175 - 0.07K_{\mathrm{l}} > 0 \end{cases} \tag{7}$$

由式 (7) 可得, 当弹性系数 K_{l} 范围为 0～216.7 Nm/rad 时, ELSSG 能够保持稳定。

金属-橡胶缓冲弹簧作为弹性连接装置, 在力矩加载时会因形变而引起输出迟滞及输出死区, 而且迟滞时间与死区长度和弹簧的弹性系数有关。若弹性系数太小, 迟滞时间与死区长度都会增大, 导致系统加载精度和动态响应速度下降, 无法满足高频加载的要求; 若弹性系数过大, 迟滞时间与死区长度都会减小, 虽然可以增加系统加载频宽, 但高频扰动分量也会随之增加, 导致系统动态性能变差。总之, 设置不合理的弹簧弹性系数将会使得输出力矩误差增大。

在 MATLAB/Simulink 中建立 ELSSG 仿真模型, 更改弹性系数并绘制频率特性曲线。当 K_{l} 为 1, 100, 200 Nm/rad 时, 其频率特性曲线如图 3 中所示。分析可知, 当弹性系数 K_{l} 增大时, 系统带宽也随之提高。K_{l} 为 1 Nm/rad 时, 系统带宽为 0.489 Hz; K_{l} 为 100 Nm/rad 时, 系统带宽为 10.4 Hz; K_{l} 为 200 Nm/rad 时, 系统带宽为 15.1 Hz。随着弹性系数的提高, 相同频率时, 系统的相位延迟也逐渐减少, 在信号频率为 5 Hz 的条件下, K_{l} 为 1 Nm/rad 时, 相位延迟为−184°, K_{l} 为 100 Nm/rad 时, 相位延迟为−128°, K_{l} 为 200 Nm/rad 时, 相位延迟为−55.1°。但弹性系数越高, 金属-橡胶缓冲弹簧特性越趋近刚性杆, 对系统高频扰动分量的抑制能力越低。因此, 弹性系数并不是越高越好。

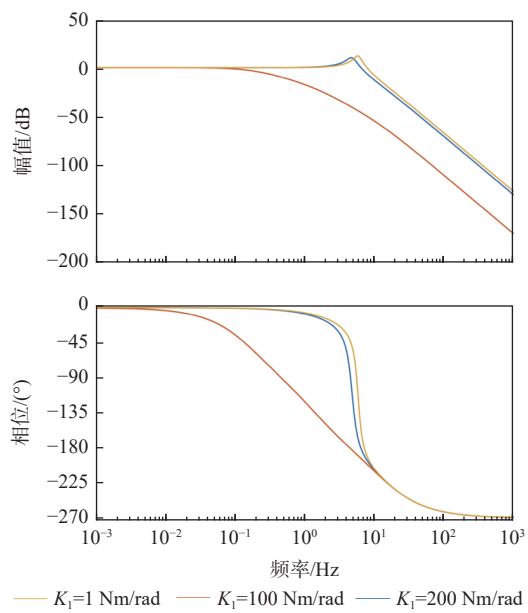


图 3　不同弹性系数的系统频率特性曲线

Fig. 3　Frequency characteristic curves of systems with different elastic coefficients

因此, 目前研究学者普遍采用的方法是: 在保证系统稳定的条件下, 首先通过劳斯稳定判据确定弹性系数的取值范围, 然后根据实验需要更换不同弹性系数的弹簧以获得更好的多余力抑制效果, 或者使用相同弹性系数的弹簧进行不同控制方法的对比实验。本文设定弹性系数为 100 Nm/rad, 既能

保证系统稳定和足够的带宽，又可以基于该弹性系数对不同的控制方法进行实验对比。

3 舵机电动负载模拟器控制器设计

在ELSSG数学模型的基础上，针对多余力矩干扰问题，采用抗扰性强的RISE控制与迭代学习控制设计了复合控制器，结构如图4所示。相比于传统的控制方法，本文提出以下3点创新之处：

1）使用RISE控制方法控制系统力矩输出。相比于传统PID控制，RISE控制不仅控制精度更高、抗扰性更强，而且可以根据系统传感器的实际情况来调整控制器阶数。更进一步地，针对RISE控制中因符号积分项导致控制初期出现振荡的问题，提出基于误差的自适应参数调整方式，即根据当前误差自动调整符号积分项的系数。

2）在RISE控制的基础上引入迭代学习控制。通过迭代调整输出量，使跟踪误差不断降低。同时，针对迭代学习控制跟踪误差收敛慢的问题，采用引导信号机制实现对系统的误差控制，并结合自适应遗忘因子加速迭代学习误差收敛过程。

3）由于舵机的主动运动是引起多余力矩的原因，研究者们基于不变性原理，引入舵机当前运动信息或舵机预设信号设计前馈控制器[9-10,12]。实验证明，该方法可提高控制精度并抑制多余力矩。而本文基于迭代学习控制，可以不需要舵机运动信息，也不需要系统精准参数设计控制器，仅根据误差进行控制量的更新，就可以实现对系统多余力矩的控制。

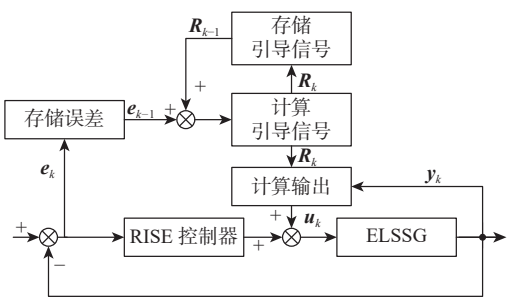


图4 复合控制器结构示意图

Fig. 4 Structure diagram of compound controller

3.1 RISE控制器设计

RISE控制是一种数据驱动的控制方法，不仅结构简单，而且能够在不需要系统详细参数与建模的情况下实现对系统的控制，已广泛应用于水下机器人、自动飞行系统、多机器人系统轨迹跟踪等领域[19-23]。同样作为数据驱动的控制方法，PID控制只利用到了系统的误差、误差的微分及误差积分这3项系统信息。相比之下，RISE控制可以根据传感器及系统实际情况，通过改变控制器结构的方式，利用误差的二阶微分及其他系统信息。另外，RISE控制器特有的误差符号积分能够在误差较小时实现控制量的快速调节，提高控制精度与抗扰性。综合上述优点，RISE控制更适合解决有强干扰的ELSSG控制问题。

设ELSSG的加载力矩预设值为x_r，实际加载力矩输出值为x，定义力矩误差为

$$o_1 = x_r - x \tag{8}$$

式中：o_1为力矩跟踪误差。

定义辅助误差为

$$\begin{cases} o_2 = \dot{o}_1 + o_1 \\ \vdots \\ o_n = \dot{o}_{n-1} + o_{n-1} + o_{n-2} \end{cases} \tag{9}$$

根据误差与辅助误差定义RISE控制器为

$$u(t) = qo_n + \int_0^t (k_s \mu o_n + \gamma \mathrm{sgn}(o_n))\mathrm{d}\tau \tag{10}$$

式中：q、μ、γ为正数；k_s为常数；符号函数$\mathrm{sgn}(x)$表达式为

$$\mathrm{sgn}(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases} \tag{11}$$

根据不同的系统结构，通过改变n的大小，可以改变RISE控制器的结构，从而利用不同的系统误差信息。根据上述定义及ELSSG内部结构，取n为2，设计控制器为

$$u(t) = qo_2 + \int_0^t (k_s \mu o_2 + \gamma \mathrm{sgn}(o_2))\mathrm{d}\tau \tag{12}$$

符号函数积分作为RISE控制的独创之处，理论上可以在误差较小时输出较大控制量，增加系统抗扰性，并调高系统控制精度。但实际应用中，符号积分项系数γ值的选取无法兼顾控制精度与稳定性。若γ过小，控制器输出中符号函数积分项所占比重小，控制精度便会降低；若γ过大，控制初期符号积分会引起加载力矩输出振荡，不利于系统稳定。

针对该问题，本文提出自适应符号函数项改进方案，对符号函数项引入基于误差的自适应策略，具体改进为

$$u(t) = qo_2 + \int_0^t (k_s \mu o_2 + \gamma \mathrm{sgn}(o_2)\mathrm{e}^{-\lambda|o_2|})\mathrm{d}\tau \tag{13}$$

式中：λ为正常数。在误差较大及控制初期时，符号积分项输出较小，不会引起系统振荡；当误差较小时，符号积分项在控制量中占比增大，控制ELSSG精准输出加载力矩，减小跟踪误差。

3.2 迭代学习控制器设计

ELSSG中，加载力矩与多余力矩均具有周期

性, 而迭代学习控制正适用于具有周期性运行性质的系统。传统迭代学习控制跟踪系统预设输入信号, 通过迭代减小输出与预设之间的误差, 迭代学习控制律为

$$\begin{cases} \boldsymbol{e}_k = \boldsymbol{y}_{\mathrm{d}} - \boldsymbol{y}_k \\ \boldsymbol{u}_k = \boldsymbol{u}_{k-1} + f(\boldsymbol{e}_k) \end{cases} \tag{14}$$

式中: k 为迭代次数; $\boldsymbol{y}_{\mathrm{d}}$ 为系统预设输入; $\boldsymbol{y}_k$、$\boldsymbol{e}_k$、$\boldsymbol{u}_k$ 分别为系统在第 k 次迭代时的输出、跟踪误差、输入。

根据上一次迭代的误差, 迭代学习控制可以逐渐减小系统跟踪误差。但是, 由于系统中存在时滞及输入输出间相位延迟, 传统迭代学习控制会在 ELSSG 运行中出现失控问题。为此, 文献 [10] 提出的引导信号迭代学习控制可以表示为

$$\begin{cases} \boldsymbol{e}_k(t) = \boldsymbol{y}_{\mathrm{d}}(t) - \boldsymbol{y}_k(t) \\ \boldsymbol{l}_k(t) = \boldsymbol{l}_{k-1}(t) + \beta \boldsymbol{e}_{k-1}(t) \\ \boldsymbol{u}_k(t) = \varepsilon(\boldsymbol{l}_k(t) - \boldsymbol{y}_k(t)) \end{cases} \tag{15}$$

式中: $\boldsymbol{l}_k(t)$ 为第 k 次迭代时的引导信号; ε 为系统学习参数; β 为迭代学习系数。

相比于传统迭代学习方法, 该方法不再跟踪预设信号, 而是结合 PID 控制与迭代学习控制, 对引导信号进行跟踪。该方法实现了迭代过程中对跟踪误差的控制, 避免因控制时滞与初始状态引起系统失控。但由于根据引导信号进行迭代, 而非直接根据误差进行迭代, 使得该方法存在收敛速度慢的缺点。

针对该问题, 本文提出带自适应遗忘因子的迭代学习控制, 其控制律为

$$\begin{cases} \alpha = \mathrm{e}^{1-k_{\mathrm{s}}\delta} \quad \delta > 0 \\ \boldsymbol{e}_k(t) = \boldsymbol{y}_{\mathrm{d}}(t) - \boldsymbol{y}_k(t) \\ \boldsymbol{R}_k(t) = \boldsymbol{R}_{k-1}(t) + \boldsymbol{L}_k(t)\boldsymbol{e}_{k-1}(t) \\ \boldsymbol{u}_k(t) = (1-\alpha)p(\boldsymbol{R}_k(t) - \boldsymbol{y}_k(t)) + \alpha(\boldsymbol{e}_k(t) - \boldsymbol{e}_{k-1}(t)) \end{cases} \tag{16}$$

式中: α 为自适应遗忘因子; k_{s}、δ、p 为常数; $\boldsymbol{R}_k(t)$ 为改进后的第 k 次迭代时的引导信号; $\boldsymbol{L}_k(t)$ 为系数矩阵, 其维度根据误差的维度进行调整。

相比于传统迭代学习控制, 本文迭代学习控制方法有以下 3 点创新之处:

1) 采用引导信号迭代学习控制。引导信号迭代学习在迭代过程中不再跟踪预设信号, 实现了迭代过程中对误差的控制, 从而防止系统失控。引导信号机制的引入提高了迭代学习控制的抗扰性。

2) 引入遗忘因子, 结合引导信号迭代学习控制与传统迭代学习控制设计复合迭代学习控制器。在迭代前期, 通过前 2 次跟踪误差的差值, 加速迭代学习过程。随着迭代次数的增加, 通过遗忘因子增加引导信号迭代学习控制在输出中的比例, 控制系统收敛。

3) 对遗忘因子引入自适应机制。固定遗忘因子不能同时提高迭代速度与误差精度, 而且会导致系统在迭代中误差发散。自适应遗忘因子可以同时兼顾两者, 根据当前误差决定加快迭代速度, 减少跟踪误差精度。

该方法收敛性证明如下:

对于一个存在 n 个状态变量的系统:

$$\begin{cases} \boldsymbol{x}_k(t+1) = \boldsymbol{A}\boldsymbol{x}_k(t) + \boldsymbol{B}\boldsymbol{u}_k(t) \\ \boldsymbol{y}_k(t) = \boldsymbol{C}\boldsymbol{x}_k(t) \end{cases} \tag{17}$$

式中: $k = 1,2,\cdots,n$; $\boldsymbol{A}$、$\boldsymbol{B}$、$\boldsymbol{C}$ 为系统系数矩阵; $\boldsymbol{x}_k(t)$、$\boldsymbol{y}_k(t)$、$\boldsymbol{u}_k(t)$ 分别为系统在第 k 次迭代时的状态向量、输出向量、输入向量。

$$\boldsymbol{x}_k(t) = [\boldsymbol{x}_{k1}(t), \boldsymbol{x}_{k2}(t), \cdots, \boldsymbol{x}_{kn}(t)]^{\mathrm{T}} \tag{18}$$

$$\boldsymbol{y}_k(t) = [\boldsymbol{y}_k(1), \boldsymbol{y}_k(2), \cdots, \boldsymbol{y}_k(n)]^{\mathrm{T}} \tag{19}$$

$$\boldsymbol{u}_k(t) = [\boldsymbol{u}_k(0), \boldsymbol{u}_k(1), \cdots, \boldsymbol{u}_k(n-1)]^{\mathrm{T}} \tag{20}$$

系统第 k 次的输出为

$$\begin{aligned} \boldsymbol{y}_k = {} & \boldsymbol{C}\boldsymbol{A}^{N-1}\boldsymbol{B}\boldsymbol{u}_k(0) + \boldsymbol{C}\boldsymbol{A}^{N-2}\boldsymbol{B}\boldsymbol{u}_k(1) + \cdots + \\ & \boldsymbol{C}\boldsymbol{B}\boldsymbol{u}_k(n-1) + \boldsymbol{C}\boldsymbol{A}^{N}\boldsymbol{x}_k(0) \end{aligned} \tag{21}$$

简化为

$$\boldsymbol{y}_k = \boldsymbol{G}\boldsymbol{u}_k + \boldsymbol{D}\boldsymbol{x}_k(0) \tag{22}$$

式中:

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{C}\boldsymbol{B} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{C}\boldsymbol{A}\boldsymbol{B} & \boldsymbol{C}\boldsymbol{B} & \cdots & \boldsymbol{0} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{C}\boldsymbol{A}^{n-1}\boldsymbol{B} & \boldsymbol{C}\boldsymbol{A}^{n-2}\boldsymbol{B} & \cdots & \boldsymbol{C}\boldsymbol{B} \end{bmatrix}$$

$$\boldsymbol{D} = [\boldsymbol{C}\boldsymbol{A}, \boldsymbol{C}\boldsymbol{A}^2, \cdots \boldsymbol{C}\boldsymbol{A}^n]^{\mathrm{T}}$$

系统在 k+1 次迭代后的误差 $\boldsymbol{e}_{k+1}$ 为

$$\begin{aligned} \boldsymbol{e}_{k+1} = {} & \boldsymbol{y}_{\mathrm{d}} - \boldsymbol{y}_{k+1} = \boldsymbol{y}_{\mathrm{d}} - \boldsymbol{G}\boldsymbol{u}_{k+1} - \boldsymbol{D}\boldsymbol{x}_{k+1}(0) = \boldsymbol{y}_{\mathrm{d}} - \\ & \boldsymbol{G}[(1-\alpha)p(\boldsymbol{R}_{k+1} - \boldsymbol{y}_{k+1}) + \alpha(\boldsymbol{e}_k - \boldsymbol{e}_{k-1})] - \boldsymbol{D}\boldsymbol{x}_{k+1}(0) = \\ & \boldsymbol{y}_{\mathrm{d}} - \boldsymbol{G}\boldsymbol{u}_k - \boldsymbol{D}\boldsymbol{x}_{k+1}(0) + \boldsymbol{D}\boldsymbol{x}_k(0) + \boldsymbol{G}[(1-\alpha)p\boldsymbol{y}_k - \\ & \alpha(\boldsymbol{e}_{k-1} - \boldsymbol{e}_{k-2}) + (1-\alpha)p(\boldsymbol{L}_k\boldsymbol{e}_k - \boldsymbol{y}_{k+1}) + \alpha(\boldsymbol{e}_{k-1} - \boldsymbol{e}_{k-2})] \end{aligned} \tag{23}$$

根据式 (22), 式 (23) 可转化为

$$\begin{aligned} \boldsymbol{e}_{k+1} = {} & \boldsymbol{e}_k - \boldsymbol{G}[(1-\alpha)p(\boldsymbol{e}_{k+1} - \boldsymbol{e}_k + \boldsymbol{L}_k\boldsymbol{e}_k) + \\ & \alpha(\boldsymbol{e}_k - 2\boldsymbol{e}_{k-1} + \boldsymbol{e}_{k-2})] - \boldsymbol{D}\boldsymbol{x}_{k+1}(0) + \boldsymbol{D}\boldsymbol{x}_k(0) \end{aligned} \tag{24}$$

根据式 (15) 与式 (16) 可得

$$\boldsymbol{x}_{k+1}(0) = \boldsymbol{x}_k(N) = \boldsymbol{A}^N\boldsymbol{x}_k(0) + \boldsymbol{Q}\boldsymbol{u}_k \tag{25}$$

式中: $\boldsymbol{Q} = [\boldsymbol{A}^{n-1}\boldsymbol{B}, \boldsymbol{A}^{n-2}\boldsymbol{B}, \cdots, \boldsymbol{A}\boldsymbol{B}, \boldsymbol{B}]$。

ELSSG 引导信号与输入输出信号均具有周期

性，即当$\boldsymbol{L}_k(t)=\boldsymbol{L}_{k-1}(t)$时可得

$$\begin{cases}\boldsymbol{y}_k(t)=\boldsymbol{y}_{k-1}(t)\\ \boldsymbol{u}_k(t)=\boldsymbol{u}_{k-1}(t)\end{cases} \tag{26}$$

同时，对于迭代学习时间上的连续性，有$\boldsymbol{x}_{k+1}(0)=\boldsymbol{x}_k(N)$，以及系统初始状态的一致性$\boldsymbol{x}_{k+1}(0)=\boldsymbol{x}_k(0)$。结合式(25)可得

$$\boldsymbol{x}_{k+1}(0)=\boldsymbol{A}^n\boldsymbol{x}_k(0)+\boldsymbol{Q}\boldsymbol{u}_{k-1} \tag{27}$$

根据式(24)可得

$$\begin{aligned}\boldsymbol{x}_{k+1}(0)-\boldsymbol{x}_k(0)=&(1-\alpha)p\boldsymbol{Q}(\boldsymbol{L}_k\boldsymbol{e}_{k-1}+\boldsymbol{e}_k-\boldsymbol{e}_{k-1})+\\&\alpha\boldsymbol{Q}(\boldsymbol{e}_{k-1}-2\boldsymbol{e}_{k-2}+\boldsymbol{e}_{k-3})\end{aligned} \tag{28}$$

结合式(19)与式(22)可得

$$\begin{aligned}\boldsymbol{e}_{k+1}=&\boldsymbol{e}_k-\boldsymbol{G}[(1-\alpha)p(\boldsymbol{e}_{k+1}-\boldsymbol{e}_k+\boldsymbol{L}_k\boldsymbol{e}_k)+\alpha(\boldsymbol{e}_k-2\boldsymbol{e}_{k-1}+\\&\boldsymbol{e}_{k-2})]-\boldsymbol{D}[(1-\alpha)p\boldsymbol{Q}(\boldsymbol{L}_k\boldsymbol{e}_{k+1}+\boldsymbol{e}_k-\boldsymbol{e}_{k-1})+\\&\alpha\boldsymbol{Q}(\boldsymbol{e}_{k-1}-2\boldsymbol{e}_{k-2}+\boldsymbol{e}_{k-3})]\end{aligned} \tag{29}$$

由此可得

$$\begin{aligned}\boldsymbol{e}_{k+1}=&[\boldsymbol{I}+(1-\alpha)p(\boldsymbol{G}+\boldsymbol{D}\boldsymbol{Q}\boldsymbol{L}_k)]^{-1}\cdot\\&\{[\boldsymbol{I}+(1-\alpha)p(\boldsymbol{G}-\boldsymbol{L}_k\boldsymbol{G}-\boldsymbol{D}\boldsymbol{Q})-\alpha\boldsymbol{G}]\boldsymbol{e}_k+\\&[2\alpha\boldsymbol{G}+(1-\alpha)p\boldsymbol{D}\boldsymbol{Q}-\alpha\boldsymbol{D}\boldsymbol{Q}]\boldsymbol{e}_{k-1}+\\&(2\alpha\boldsymbol{D}\boldsymbol{Q}-\alpha\boldsymbol{G})\boldsymbol{e}_{k-2}-\alpha\boldsymbol{D}\boldsymbol{Q}\boldsymbol{e}_{k-3}\}\end{aligned} \tag{30}$$

式中：$\boldsymbol{I}$为单位矩阵。

对式(30)两边取范数，根据范数的相容性，可得

$$\begin{aligned}\|\boldsymbol{e}_{k+1}\|<&\left\|[\boldsymbol{I}+(1-\alpha)p(\boldsymbol{G}+\boldsymbol{D}\boldsymbol{Q}\boldsymbol{L}_k)]^{-1}\cdot\right.\\&\left.[\boldsymbol{I}+(1-\alpha)p(\boldsymbol{G}-\boldsymbol{L}_k\boldsymbol{G}-\boldsymbol{D}\boldsymbol{Q})-\alpha\boldsymbol{G}]\right\|\|\boldsymbol{e}_k\|+\\&\left\|[2\alpha\boldsymbol{G}+(1-\alpha)p\boldsymbol{D}\boldsymbol{Q}-\alpha\boldsymbol{D}\boldsymbol{Q}]\right\|\|\boldsymbol{e}_{k-1}\|+\\&\left\|(2\alpha\boldsymbol{D}\boldsymbol{Q}-\alpha\boldsymbol{G})\right\|\|\boldsymbol{e}_{k-2}\|+\left\|\alpha\boldsymbol{D}\boldsymbol{Q}\right\|\|\boldsymbol{e}_{k-3}\|\end{aligned} \tag{31}$$

引理1[12] 当$p_1+p_2+\cdots+p_n<1, p_i\geqslant 0\ (i=0,1,\cdots,n)$时，若$a_k\leqslant p_1a_{k-1}+p_2a_{k-2}+\cdots+p_na_{k-n}\ (k=n+1,n+2,\cdots)$，则$\lim\limits_{k\to\infty}a_k=0$。

综上，若满足如下条件：

$$\begin{aligned}&\left\|[\boldsymbol{I}+(1-\alpha)p(\boldsymbol{G}+\boldsymbol{D}\boldsymbol{Q}\boldsymbol{L}_k)]^{-1}\cdot\right.\\&\left.[\boldsymbol{I}+(1-\alpha)p(\boldsymbol{G}-\boldsymbol{L}_k\boldsymbol{G}-\boldsymbol{D}\boldsymbol{Q})-\alpha\boldsymbol{G}]\right\|+\\&\left\|[2\alpha\boldsymbol{G}+(1-\alpha)p\boldsymbol{D}\boldsymbol{Q}-\alpha\boldsymbol{D}\boldsymbol{Q}]\right\|+\\&\left\|(2\alpha\boldsymbol{D}\boldsymbol{Q}-\alpha\boldsymbol{G})\right\|+\|\alpha\boldsymbol{D}\boldsymbol{Q}\|<1\end{aligned} \tag{32}$$

根据该引理，当$k\to\infty$时，系统跟踪误差$\boldsymbol{e}_k\to\boldsymbol{0}$，即自适应遗忘因子迭代学习控制在满足式(32)的条件下误差收敛于$\boldsymbol{0}$，该控制方法可以使得系统收敛。

4 实验验证

为验证本文方法的有效性，实验共分为3个步骤来验证各部分内容：第1步验证改进引导信号迭代学习方法的有效性，第2步验证使用非固定频率正弦信号作为舵机工作信号对系统力矩跟踪误差的影响，第3步验证复合控制方法在ELSSG控制中的有效性。

4.1 改进引导参数迭代学习控制性能实验

为验证改进引导信号迭代学习的有效性，选择如下系统进行测试：

$$y(k+1)=\begin{cases}\dfrac{y(k)u(k)}{1+y(k)^2}+(u(k)+a(k)\sin(y(k)))^3 & k<250\\ \dfrac{y(k)u(k)^3}{1+y(k)^2}+u(k)^3 & 250\leqslant k\leqslant 500\end{cases} \tag{33}$$

式中：$a(k)=0.1\mathrm{round}(k/500)$，round( )函数用于舍入到最近的整数。

系统期望信号为

$$y_{\mathrm{d}}(k+1)=\begin{cases}0.5\times(-1)^{\mathrm{round}(k/100)} & k<250\\ 0.5\sin(k\pi/100)+0.3\cos(k\pi/50) & 250\leqslant k\leqslant 500\end{cases} \tag{34}$$

式(34)是参数时变且结构时变的系统。分别采用本文提出的自适应遗忘因子的引导信号迭代学习控制与传统引导信号迭代学习控制，在不同迭代次数下获得系统的跟踪曲线及跟踪误差，如图5、图6和表1所示。

由图5分析可知，在10次迭代后，改进迭代学习控制相比传统迭代学习控制更接近预设信号。结合表1分析可知，自适应遗忘因子引导信号迭代

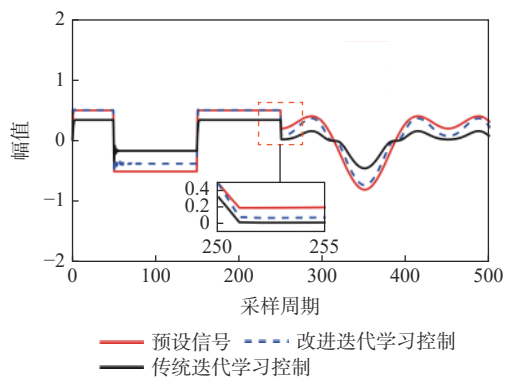


图5 10次迭代下的系统输出

Fig. 5 System output at 10 iterations

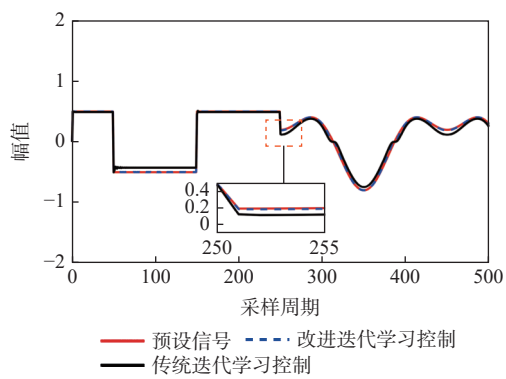


图6 20次迭代下的系统输出

Fig. 6 System output at 20 iterations

表 1　不同迭代学习控制下系统的跟踪误差

Table 1　Tracking error of system under different iterative learning control

方法	迭代次数	最大跟踪误差/(N·m)	误差率/%
传统迭代学习控制	10	0.41	51.2
	20	0.051 7	6.46
改进迭代学习控制	10	0.199 5	24.9
	20	0.023 4	2.925

学习的最大跟踪误差率为 24.9%, 相比传统迭代学习控制跟踪误差降低了 51.4%。说明与传统迭代学习控制相比, 本文提出的改进迭代学习控制误差收敛更快。由图 6 分析可知, 20 次迭代后, 传统迭代学习控制与预设信号仍有明显差距, 此时本文改进迭代学习控制方法的跟踪误差率仅为 2.925%。可以看出, 自适应遗忘因子引导信号迭代学习控制能够应用于参数时变与结构时变的控制系统, 且误差收敛更快, 控制精度相比传统迭代学习控制更高。

4.2　舵机电动负载模拟器力矩加载实验

在 MATLAB/Simulink 平台建立 ELSSG 模型结构, 如图 7 所示, 包括 MATLAB 程序部分和 Simulink 模型部分。在迭代过程中, 不随迭代次数变化的 ELSSG 数学模型、PID 控制器及 RISE 控制器由 Simulink 模型实现。随迭代次数变化的迭代学习控制器由 MATLAB 代码实现。MATLAB 主程序输出预设力矩、迭代学习控制器迭代所需要的参数与跟踪力矩误差值, 并接收当前迭代周期的跟踪力矩误差与迭代学习控制器的输出值, 调用 Simulink 模型实现迭代。系统具体参数如下: R_m=4.32 Ω, L_m=0.005 H, C_e=3.6 (V·s)/rad, J_m=0.12 kg·m^2。预设加载力矩信号为正弦信号, 幅值为 50 N·m, 频率为 1 Hz。在仿真实验中, 设置固定步长为 0.01 s, 此时系统输出曲线足够平滑, 不需再额外增加滤波装置。

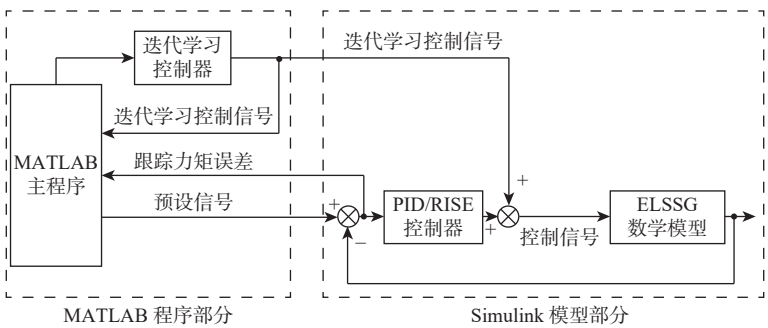


图 7　TLAB/Simulink 模型结构示意图

Fig. 7　Schematic of MATLAB/Simulink model structure

4.2.1　舵机工作信号对加载力矩影响实验

实验中, 舵机运动采用 2 种不同信号。第 1 种是频率从 0.01 Hz 逐渐变化至 1 Hz 的正弦波信号(扫频信号), 其数学表达式为 $y(t)=5\mathrm{e}^{\mathrm{j}2\pi(0.1t+0.5\times0.999t^2)}$, 信号图像如图 8 所示。第 2 种是常用的固定频率正弦波信号(正弦信号), 其数学表达式为 $y(t)=5\sin t$。

实验分为 2 组。第 1 组实验中, ELSSG 模型预设力矩信号为幅值 50 N·m、频率 1 Hz 的正弦信号, 其数学表达式为 $y(t)=50\sin t$。第 2 组实验中, 预设力矩信号幅值为 0, 其数学表达式为 $y(t)=0$。2 组实验均使用传统 PID 控制做为控制器, 实验结果如图 9 所示。

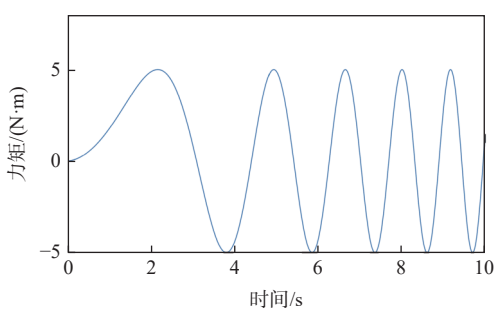


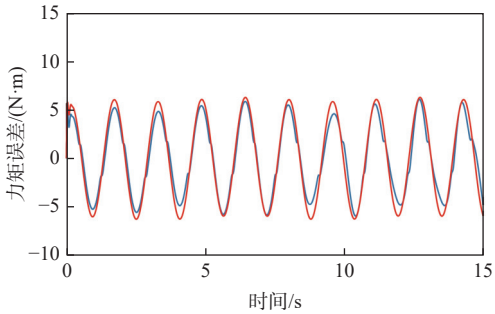
图 8　变频率正弦信号(扫频信号)图像

Fig. 8　Sinusoidal signal (sweep frequency signal) image with variable frequency

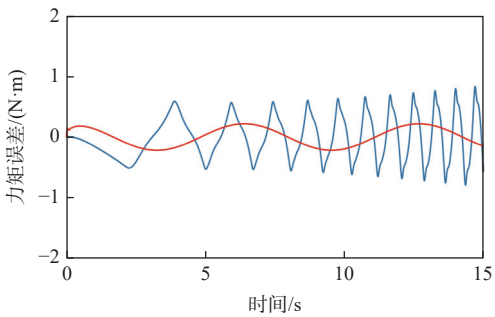
由图 8 和图 9 分析可知, 当舵机信号为变频率的正弦信号时, 由于信号频率发生了变化, 传统 PID 控制无法对多余力矩进行有效补偿, 导致系统跟踪误差发生变化。图 9 中, 当系统预设力矩为 0 时, 单一参数 PID 控制已经无法补偿多余力矩, 系统跟踪误差随着时间逐渐发散。由此可得以下结论:

1) 当舵机信号为固定频率正弦信号时, PID 控制可以实现多余力矩的补偿, 但频率变化时, 单一参数 PID 控制器无法单独实现多余力矩的补偿, 不能保证系统长期稳定运行。

2) 由于舵机在 ELSSG 工作时与加载电机相互影响, 频率变化的正弦信号作为舵机工作信号引起的跟踪力矩误差情况更复杂, 即多余力矩干扰情况更复杂, 更适于验证 ELSSG 控制方法的有效性与泛用性。



(a) 预设力矩为 50sin t 时系统跟踪误差曲线



(b) 预设力矩为 0 时系统跟踪误差曲线

图 9 不同预设力矩下系统跟踪误差曲线

Fig. 9 System tracking error curves under different preset torques

3）根据文献 [16,18]，迭代学习控制适用于解决 ELSSG 控制问题。虽然舵机信号频率发生了变化，导致该问题更加复杂，但由于迭代学习控制只与不同迭代周期中的误差有关，与系统结构和信号变化形式无关，舵机信号频率变化时，不会影响迭代学习控制通过迭代改变控制量。因此，理论上迭代学习控制仍然适用于该问题。

4.2.2 多余力矩补偿实验

ELSSG 模型参数与 4.2.1 节相同，将实验分为 4 组，控制方法如下：①第 1 组。由本文提出的自适应遗忘因子引导信号迭代学习控制与自适应 RISE 控制构成的复合控制方法。②第 2 组。由迭代学习控制与 PID 控制构成的复合控制方法。③第 3 组。由引导信号迭代学习控制与 PID 控制构成的复合控制方法。④第 4 组。由迭代学习控制与 RISE 控制构成的复合控制方法。PID 控制参数为：P=10, I=0, D=5。RISE 控制器控制参数为：k_s =7, μ =7, γ =15。

第 25 次迭代后力矩跟踪效果、第 25 次迭代后加载力矩跟踪误差、系统迭代误差如图 10～图 12 及表 2 所示。

由上述数据分析可知：

1）对比 25 次迭代后各实验组的跟踪误差。由图 10 及表 2 分析可知，在 25 次迭代后，所有实验组的加载力矩均能跟随预设力矩，且误差率都在

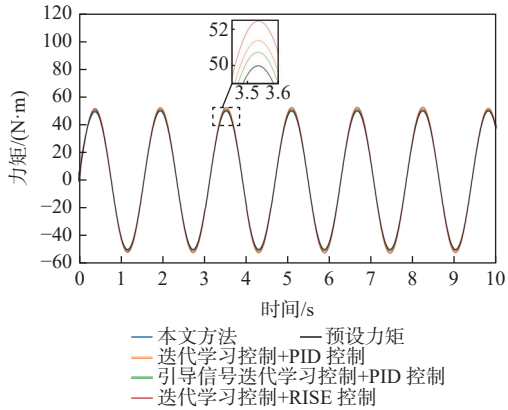


图 10 预设信号频率为 1 Hz 时 ELSSG 输出力矩

Fig. 10 ELSSG output torque when preset signal frequency is 1 Hz

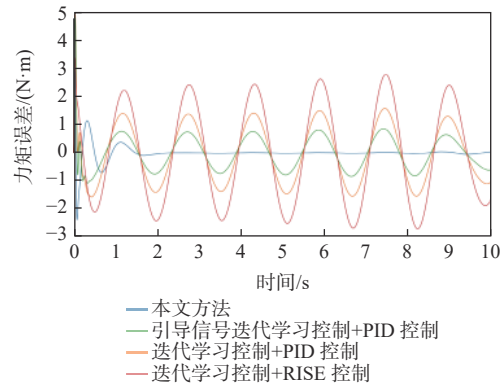


图 11 预设信号频率为 1 Hz 时 ELSSG 输出力矩误差

Fig. 11 ELSSG output torque error when preset signal frequency is 1 Hz

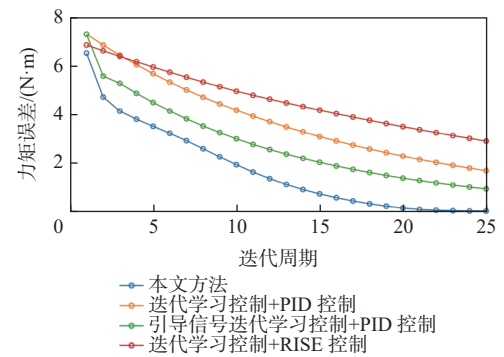


图 12 预设信号频率为 1 Hz 时力矩误差随迭代周期变化

Fig. 12 Variation of torque error with iterative periods when preset signal frequency is 1 Hz

表 2 预设信号为 1 Hz 时力矩跟踪误差

Table 2 Torque tracking error when preset signal is 1 Hz

方法	最大跟踪误差/(N·m)	误差率/%
本文方法	0.06	0.12
迭代学习控制+PID控制	1.67	3.34
引导信号迭代学习控制+PID控制	0.92	1.84
迭代学习控制+RISE控制	2.59	5.18

6% 以内, 证明传统 PID 控制器无法补偿的多余力矩干扰可以通过迭代学习复合控制器进行补偿。由图 12 分析可知, 使用迭代学习控制器配合其他方法均能进一步降低跟踪误差, 下降了 2%～5%。其中, 使用迭代学习控制+RISE 控制, 误差率最高为 5.18%, 而使用本文方法的实验组最终误差率为 0.12%, 远低于其他控制方法。

2) 对比各实验组第 1 次迭代后的跟踪误差与第 25 次迭代后的跟踪误差。由图 12 分析可知, 迭代学习控制+PID 控制的跟踪误差与引导信号迭代学习控制 + PID 控制的初始误差相同。但随着迭代, 引导信号迭代学习控制 + PID 控制误差收敛更快, 最终误差率仅为 1.84%, 远低于迭代学习控制 + PID 控制的 3.34%, 说明引导信号迭代学习控制在控制精度与收敛速度比迭代学习控制更优。由图 12 分析可知, 虽然在第 1 次迭代时 RISE 控制的误差低于 PID 控制, 但配合相同参数的迭代学习控制器进行迭代后, 迭代学习控制+RISE 控制的误差下降反而比迭代学习控制+PID 控制缓慢, 最终使用迭代学习控制+RISE 控制的 ELSSG 力矩跟踪误差比迭代学习控制+PID 控制的力矩跟踪误差率高 1.84%。这是由于 RISE 控制器中符号积分项参数为固定值, 每次迭代后系统误差本应下降, 但误差减小时, 符号积分项由于符号函数的特性仍为 1 或 −1, 输出较大的控制量, 导致实际 RISE 控制配合迭代学习控制后系统误差下降速度受到影响。因此, 虽然 RISE 控制在首次迭代时跟踪误差小于相同条件下的 PID 控制, 但 RISE 控制结合迭代学习控制构成的复合控制器降低输出力矩跟踪误差的能力却更差。

3) 纵向对比每次迭代后力矩跟踪误差。由图 11 及图 12 分析可知, 采用本文方法的 ELSSG 在每次迭代后输出误差均为最低。由表 2 可知, 在最后一次迭代后, 相比传统的迭代学习控制+PID 控制, 应用本文方法的实验组力矩跟踪误差降低了 96.4%。由图 12 分析可知, 本文方法在迭代中误差下降最快, 这是由于自适应 RISE 控制能够调整符号积分项的系数, 从而配合引导信号迭代学习控制迅速降低系统误差。

为测试本文提出的控制器对于高频率信号的控制性能, 在 ELSSG 模型参数和 RISE 控制器控制参数不变的情况下, 预设力矩信号改为幅值 50 N·m、频率 10 Hz 的正弦信号, 其数学表达式为 $y(t)=50\sin(10t)$。同样分为 4 组进行对比实验, 得到第 20 次迭代后力矩跟踪效果、第 20 次迭代后加载力矩跟踪误差、系统迭代误差如图 13～图 15 及表 3 所示。

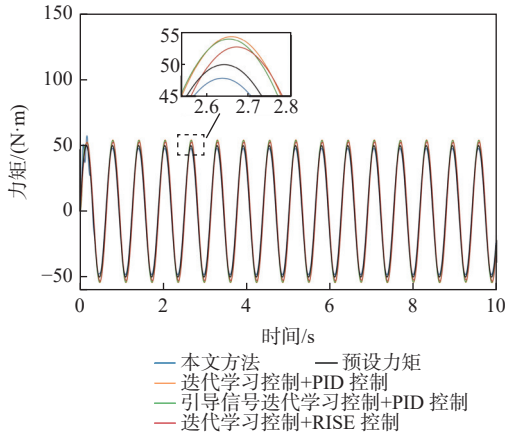


图 13　预设信号频率为 10 Hz 时 ELSSG 输出力矩

Fig. 13　ELSSG output torque when preset signal frequency is 10 Hz

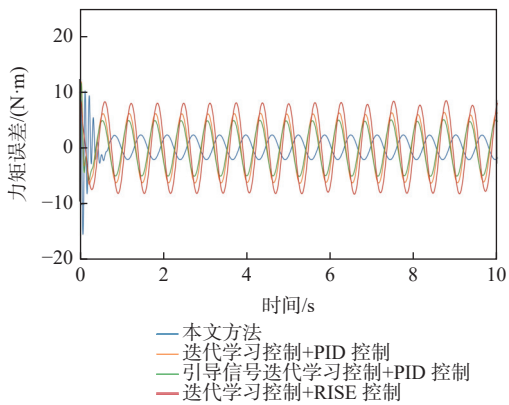


图 14　预设信号频率为 10 Hz 时 ELSSG 输出力矩误差

Fig. 14　ELSSG output torque error when preset signal frequency is 10 Hz

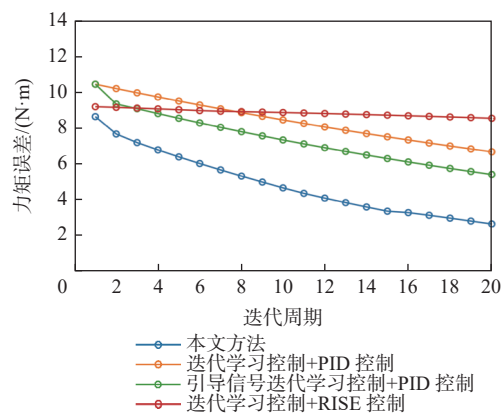


图 15　预设信号频率为 10 Hz 时力矩误差随迭代周期变化

Fig. 15　Variation of torque error with iterative periods when preset signal frequency is 10 Hz

表 3　预设信号为 10 Hz 时力矩跟踪误差

Table 3　Torque tracking error when preset signal is 10 Hz

方法	最大跟踪误差/(N·m)	误差率/%
本文方法	2.60	5.20
迭代学习控制+PID控制	6.70	13.40
引导信号迭代学习控制+PID控制	5.44	10.88
迭代学习控制+RISE控制	8.57	17.14

由上述数据分析可知：

1）对比20次迭代后各实验组的跟踪误差。由图13分析可知，当预设信号频率从1 Hz提高至10 Hz时，在20次迭代后，所有实验组的输出力矩均能跟随预设力矩输出正弦波形。由图14及表3分析可知，在20次迭代后，使用本文方法的实验组跟踪误差率最低，仅为5.20%，而迭代学习控制+RISE控制的跟踪误差率最高，达到了17.14%，说明本文提出的控制器在预设信号频率为10 Hz时仍能有效抑制多余力矩。

2）对比各实验组第1次迭代后的跟踪误差与第20次迭代后的跟踪误差。由图15分析可知，当预设信号频率从1 Hz提高至10 Hz后，迭代学习控制+PID控制、引导信号迭代学习控制+PID控制及本文方法在每次迭代后均能降低跟踪误差，最大跟踪误差分别降低了35.9%、48%、68.9%，而迭代学习控制+RISE控制跟踪误差在迭代20次后仅降低了7%。一方面，当预设信号频率提高时，迭代学习控制+RISE控制自身迭代误差下降慢的问题仍然存在，且在高频率时更加严重。另一方面，应用本文方法的控制器能够克服该问题，可以通过迭代快速降低力矩跟踪误差，说明在高频率条件下，本文提出的自适应RISE与改进引导信号迭代学习方法仍然有效。同时，上述实验表明，本文方法能够实现ELSSG高精度力矩输出，在不使用舵机前馈控制的条件下抑制多余力矩。

4.3 舵机信号前馈控制实验

根据相关文献[9-10,12,24]，添加舵机信号前馈控制能够提高复合控制器的抗扰性与控制精度。为测试舵机信号前馈控制是否能够在舵机信号变频率条件下发挥作用，并提高本文方法的控制精度，基于不变性原理[24-25]与系统数学模型，设计了舵机前馈控制器，系统控制结构示意图如图16所示。其中，舵机前馈控制器传递函数$G_V(s)$与加载电机传递函数$G_m(s)$满足如下关系：

$$G_V(s)G_m(s)=1$$

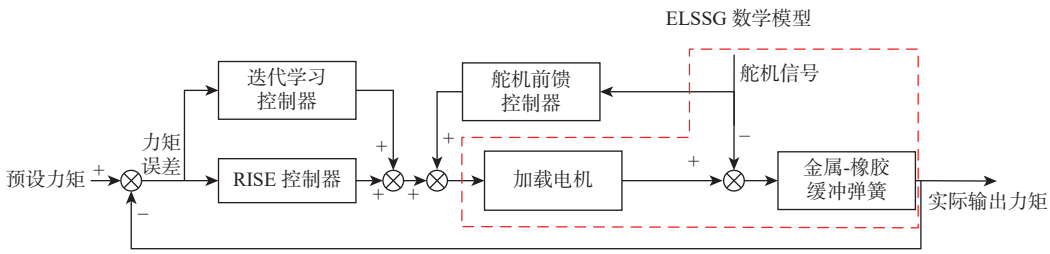


图16 添加舵机前馈控制后控制结构示意图

Fig. 16 Control structure after adding feed forward control of aircraft steering gear

在ELSSG模型参数和RISE控制器控制参数不变的情况下，设定预设力矩信号为幅值50 N·m、频率10 Hz的正弦信号，其数学表达式为$y(t)=50\sin(10t)$。分为8组进行实验，前4组不做改变，后4组在仿真系统中根据不变性原理添加相同的舵机前馈控制器。共迭代20次，系统的跟踪力矩误差随迭代周期数变化如图17所示，第20次迭代后的跟踪误差如表4所示。

从上述数据可知：

1）图17中，虚线均在相同颜色的实线之下，说明对于每种方法，添加舵机前馈控制器后，都能够有效提高对于多余力矩的抑制能力，从而降低力矩跟踪误差。

2）对于每种控制方法，每个迭代周期中添加舵机前馈控制器的实验组跟踪误差都更低。由表4分析可知，第20次迭代后，本文提出的控制策略误差率下降了0.96%，其他3种控制方法误差率分别下降了0.44%、0.32%、0.84%。

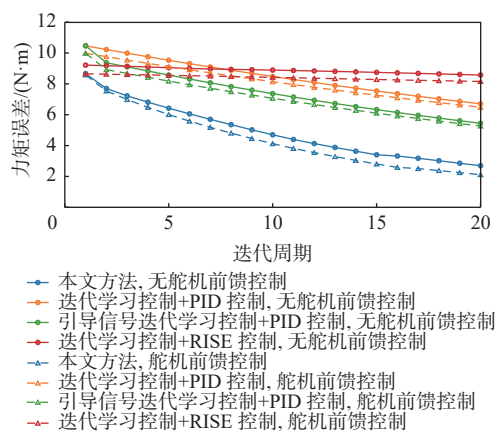


图17 添加不同前馈控制方法后不同迭代周期下的力矩跟踪误差

Fig. 17 Torque tracking error of different iteration periods after adding different feedforward control methods

上述实验数据及分析说明，在舵机信号频率变化的条件下，前馈控制器仍然有效。如果无法获取准确的舵机信号，只使用迭代学习控制或本文提出

表 4　前馈控制实验力矩跟踪误差

Table 4　Torque tracking error of feedforward experiment

是否添加舵机前馈控制器	方法	最大跟踪误差/(N·m)	误差率/%
否	本文方法	2.60	5.20
	迭代学习控制+PID控制	6.70	13.40
	引导信号迭代学习控制+PID控制	5.44	10.88
	迭代学习控制+RISE控制	8.57	17.14
是	本文方法	2.12	4.24
	迭代学习控制+PID控制	6.48	12.96
	引导信号迭代学习控制+PID控制	5.26	10.56
	迭代学习控制+RISE控制	8.15	16.30

的复合控制策略, 仍然可以有效抑制多余力矩。当条件允许时, 应在系统中引入舵机运动信息设计的前馈控制器, 进一步提高系统的抗扰性与控制性能, 降低跟踪误差, 加快迭代速度。

5　结　论

1) 针对 ELSSG 输出力矩受到多余力矩干扰问题, 提出了一种自适应遗忘因子引导信号迭代学习控制结合 RISE 控制的复合控制器, 通过数学证明得到该方法的收敛条件, 并通过实验证明其有效性。

2) 研究了舵机工作信号频率变化对 ELSSG 输出力矩误差的影响。首先, 通过实验证明了舵机工作信号频率变化时, 传统 PID 控制器的控制效果下降。其次, 与传统的固定频率舵机工作信号相比, 变化频率的信号引起的多余力矩更为复杂, 更适于验证 ELSSG 控制方法的稳定性与控制精度。

3) 与常规迭代学习控制配合 PID 控制及引导信号迭代学习控制相比, 本文提出的改进迭代学习控制配合自适应参数 RISE 控制的误差收敛速度更快, 跟踪误差更低。由自适应 RISE 控制提供迭代学习所需的初态, 并由改进引导信号迭代学习控制提高复合控制器的控制精度。通过实验证明了本文提出的复合控制器能够实现对 ELSSG 的高精度控制。

4) 当舵机信号频率变化时, 在迭代学习控制中引入基于不变性原理设计的舵机前馈控制器, 能够提高系统的控制精度。如果无法获得舵机运动信息, 仅使用本文方法也能实现对多余力矩的抑制。

参考文献 (References)

[1] 李成成, 王广林, 潘旭东, 等. 电动负载模拟器摩擦模型参数辨识方法[J]. 振动测试与诊断, 2020, 40(3): 519-525.
LI C C, WANG G L, PAN X D, et al. Research on parameter identification method of friction model for electric load simulator[J]. Journal of Vibration, Measurement & Diagnosis, 2020, 40(3): 519-525(in Chinese).

[2] LIU H T, LIU H R, SHAN X L. Linear active disturbance rejection control with torque compensation for electric load simulator[J]. Journal of Power Electronics, 2021, 21(1): 195-203.

[3] 王超, 侯远龙, 王力, 等. 炮控系统电动负载模拟器性能影响因素分析[J]. 电机与控制学报, 2016, 20(12): 74-81.
WANG C, HOU Y L, WANG L, et al. Influence analysis on electric load simulator for the gun control system[J]. Electric Machines and Control, 2016, 20(12): 74-81(in Chinese).

[4] 税洋, 尉建利, 闫杰. 基于模型参考自适应控制的舵机加载系统研究[J]. 西北工业大学学报, 2018, 36(2): 246-251.
SHUI Y, WEI J L, YAN J. Research on motor drive load simulator based on model reference adaptive control[J]. Journal of Northwestern Polytechnical University, 2018, 36(2): 246-251(in Chinese).

[5] FU Z, WANG S P, WANG X J. ESO-based adaptive robust force control of linear electric load simulator[C]//Proceedings of the CSAA/IET International Conference on Aircraft Utility Systems. London: IET, 2018: 1-6.

[6] 林辉, 吕帅帅, 陈晓雷, 等. 导弹尾翼电动负载模拟器快速终端滑模控制[J]. 哈尔滨工业大学学报, 2017, 49(3): 22-28.
LIN H, LU S S, CHEN X L, et al. Fast terminal sliding mode control for missile rudder electric dynamic load simulator systems[J]. Journal of Harbin Institute of Technology, 2017, 49(3): 22-28(in Chinese).

[7] LUO C Y, YAO J Y, CHEN F H, et al. Adaptive repetitive control of hydraulic load simulator with RISE feedback[J]. IEEE Access, 2017, 5: 23901-23911.

[8] 刘晓琳, 姜梦馨. 基于 WOA 的飞机舵机电动加载系统双环复合控制研究[J]. 振动与冲击, 2021, 40(12): 246-253.
LIU X L, JIANG M X. A study on double loop composite control based on WOA for an aircraft rudder electric loading system[J]. Journal of Vibration and Shock, 2021, 40(12): 246-253(in Chinese).

[9] 郭行, 陈康, 孙力, 等. 一种引入位置比例控制提高电动负载模拟器性能的控制系统设计方法[J]. 西北工业大学学报, 2014, 32(2): 235-239.
GUO H, CHEN K, SUN L, et al. An effective method for designing control system of motor-driven torque control simulator by introducing position proportional control[J]. Journal of Northwestern Polytechnical University, 2014, 32(2): 235-239(in Chinese).

[10] 刘晓琳, 李卓. 飞机舵机电动加载系统多余力矩抑制方法[J]. 系统工程与电子技术, 2019, 41(6): 1366-1373.
LIU X L, LI Z. Method to restrain extra torque of aircraft rudder-electric loading system[J]. Systems Engineering and Electronics, 2019, 41(6): 1366-1373(in Chinese).

[11] HAN H T, YANG B. A CMAC based self-tuning intelligent PID controller for electric load simulator[C]//Proceedings of the 3rd International Conference on Instrumentation, Measurement, Computer, Communication and Control. Piscataway: IEEE Press, 2013: 1443-1448.

[12] 潘卫东, 范元勋, 雷建杰, 等. 摩擦对电动直线负载模拟器的影响及其抑制研究[J]. 兵工学报, 2019, 40(10): 2050-2059.
PAN W D, FAN Y X, LEI J J, et al. Effect of friction on electric linear load simulator and research on friction suppression[J]. Acta Armamentarii, 2019, 40(10): 2050-2059(in Chinese).

[13] ZHANG M, YANG B. A naive method of applying fuzzy logic to

CMAC in electric load simulator[J]. Transactions of the Institute of Measurement and Control, 2017, 39(10): 1590-1599.

[14] ULLAH N, KHAN W, WANG S P. High performance direct torque control of electrical aerodynamics load simulator using fractional calculus[J]. Acta Polytechnica Hungarica, 2014, 11(10): 59-78.

[15] 代明光, 齐蓉. 基于扩展状态观测器的电动负载模拟器反演滑模控制[J]. 航空学报, 2020, 41(5): 323683.
DAI M G, QI R. Backstepping sliding mode control of electric dynamic load simulator based on extended state observer[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(5): 323683(in Chinese).

[16] 牛国臣, 王巍, 宗光华. 基于迭代学习的电动负载模拟器复合控制[J]. 控制理论与应用, 2014, 31(12): 1740-1747.
NIU G C, WANG W, ZONG G H. Composite control for electric load simulator based on iterative learning[J]. Control Theory & Applications, 2014, 31(12): 1740-1747(in Chinese).

[17] 黄静, 郑华义, 李宏, 等. 带参数学习的引导信号迭代学习控制方法[J]. 兵工学报, 2019, 40(11): 2363-2369.
HUANG J, ZHENG H Y, LI H, et al. Guiding signal iterative learning control method with parameter learning[J]. Acta Armamentarii, 2019, 40(11): 2363-2369(in Chinese).

[18] 代明光, 齐蓉. 具有控制时滞的电动加载系统迭代学习复合控制[J]. 北京航空航天大学学报, 2020, 46(2): 340-349.
DAI M G, QI R. Composite iterative learning control for electric dynamic loading system with control time delay[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(2): 340-349(in Chinese).

[19] LI Z, MA X, LI Y B. Robust tracking control strategy for a quadrotor using RPD-SMC and RISE[J]. Neurocomputing, 2019, 331: 312-322.

[20] SU Z K, XIE M Y, LI C T. RISE based active vibration control for the flexible refueling hose[J]. Aerospace Science and Technology, 2019, 92: 387-404.

[21] YAO Z K, YAO J Y, SUN W C. Adaptive RISE control of hydraulic systems with multilayer neural-networks[J]. IEEE Transactions on Industrial Electronics, 2019, 66(11): 8638-8647.

[22] HU C, GAO H B, GUO J H, et al. RISE-based integrated motion control of autonomous ground vehicles with asymptotic prescribed performance[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2021, 51(9): 5336-5348.

[23] YAO J Y, DENG W X, JIAO Z X. RISE-based adaptive control of hydraulic systems with asymptotic tracking[J]. IEEE Transactions on Automation Science and Engineering, 2017, 14(3): 1524-1531.

[24] 韩松杉, 焦宗夏, 尚耀星, 等. 基于舵机指令前馈的电液负载模拟器同步控制[J]. 北京航空航天大学学报, 2015, 41(1): 124-132.
HAN S S, JIAO Z X, SHANG Y X, et al. Synchronizing compensation control of electro-hydraulic load simulator using command signal of actuator[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(1): 124-132(in Chinese).

[25] 汪成文, 焦宗夏, 罗才瑾. 基于改进的速度同步控制的电液负载模拟器[J]. 航空学报, 2012, 33(9): 1717-1725.
WANG C W, JIAO Z X, LUO C J. An improved velocity synchronization control on electro-hydraulic load simulator[J]. Acta Aeronautica et Astronautica Sinica, 2012, 33(9): 1717-1725(in Chinese).

Iterative learning control of electric load simulator of aircraft steering gear

LIU Xiaolin*, LI Jinkai

(School of Electronic Information and Automation, Civil Aviation University of China, Tianjin 300300, China)

Abstract: The electric load simulator of aircraft steering gear (ELSSG) is disturbed by redundant torque. To address this issue,a compound controller combining iterative learning control and robust integral control of sign of error (RISE) was proposed. The mathematical model of the load simulator was established, and the generation mechanism of excess torque was analyzed.In terms of the hardware structure, the metal-rubber buffer spring was introduced to improve the stability and loading accuracy of the system. In terms of the control strategy, the compound controller was designed based on the robust integral control of error signs and the iterative learning control of the guidance signal improved by adaptive forgetting factor, and the convergence condition of the iterative learning control was proven. The simulation experiment proves that the compound controller can eliminate the interference from redundant torque more effectively than the traditional iterative learning control method and PID control method and improve the load torque output accuracy of the load simulator.

Keywords: electric load simulator of aircraft steering gear; redundant torque; iterative learning control; metal-rubber buffer spring; forgetting factor

Received: 2022-08-13; **Accepted**: 2023-05-15; **Published Online**: 2023-06-06 14:21

URL: link.cnki.net/urlid/11.2625.V.20230605.1445.002

Foundation item: The 11th Boeing Fund of Civil Aviation University of China (20210715113)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0731

深度神经网络模型任务切分及并行优化方法

巨涛*，刘帅，王志强，李林娟
(兰州交通大学 电子与信息工程学院，兰州 730070)

摘 要：为解决传统手工切分神经网络模型计算任务并行化方法面临的并行化难度大、训练耗时长、设备利用率低等问题，提出了一种基于深度神经网络(DNN)模型特性感知的任务切分及并行优化方法。结合硬件计算环境，对模型计算特性进行动态分析，获取模型内部相关性和各类参数属性，构建原始计算任务有向无环图(DAG)；利用增强反链，构建DAG节点间可分区聚类的拓扑关系，将原始DAG转换为易于切分的反链DAG；通过拓扑排序生成反链DAG状态序列，并使用动态规划将状态序列切分为不同执行阶段，分析最佳分割点进行模型切分，实现模型分区与各GPU间动态匹配；对批量进行微处理，通过引入流水线并行实现多迭代密集训练，提高GPU利用率，减少训练耗时。实验结果表明：与已有模型切分方法相比，在CIFAR-10数据集上，所提模型切分及并行优化方法可实现各GPU间训练任务负载均衡，在保证模型训练精度的同时，4 GPU加速比达到3.4，8 GPU加速比为3.76。

关 键 词：深度神经网络模型并行；模型切分；流水线并行；反链；并行优化
中图分类号：TP311
文献标志码：A **文章编号**：1001-5965(2024)09-2739-14

近年来，深度神经网络(deep neural network, DNN)的训练和推理在模型规模和计算量要求上都有了迅速的增长[1]。较大规模的神经网络模型在多个应用领域带来显著的求解效果，但不断增大的模型规模和随之对计算能力更高的要求，给深度神经网络模型应用带来了实际的挑战[2]。由于硬件计算资源的限制，如单计算节点计算能力、内存容量的不足，大规模神经网络模型很难在单个计算节点或单个加速器中直接处理[3]。有效的处理方法是将较大规模的模型进行切分，跨多个计算节点或加速节点对模型进行并行处理，从而通过模型并行获得较好的处理效果[4]。相关的研究工作表明，模型并行可以有效加快深度学习模型的训练速度，进一步缩短模型的训练时间[5-6]。

设计和实现高效的模型并行方法，通常要同时考虑神经网络模型的可并行性、硬件计算资源的硬件架构特性及可扩展性，不同类型的神经网络模型在不同的并行化策略下会表现出截然不同的计算性能特征[7]，并行优化性能高度依赖于神经网络模型特点和计算资源的硬件架构特性，只有将神经网络模型的计算任务特性和计算资源硬件架构特征较好匹配，才能充分发挥硬件计算资源的计算能力，提升神经网络模型的训练速度和计算性能。

为应对以上问题，本文提出一种深度神经网络模型任务切分及并行优化方法。首先，结合硬件资源特性，对大规模神经网络卷积过程进行分析，构建有向无环图(directed acyclic graph, DAG)。然后，通过增强反链模型分析任务间计算特性，生成拓扑排序后的状态序列。利用动态规划进行任务切分，实现任务与资源的合理匹配，并结合流水线并行技术加速训练，达到细粒度任务并行切分和并行加速的目的。

收稿日期：2022-08-17；**录用日期**：2022-11-07；**网络出版时间**：2022-11-22 11:41
网络出版地址：link.cnki.net/urlid/11.2625.V.20221118.1909.004
基金项目：国家自然科学基金(61862037)；兰州交通大学天佑创新团队项目(TY202002)；甘肃省科技计划项目(23CXGA0028)
*通信作者. E-mail：jutao@mail.lzjtu.edu.cn

引用格式：巨涛，刘帅，王志强，等. 深度神经网络模型任务切分及并行优化方法[J]. 北京航空航天大学学报，2024，50(9)：2739-2752. JU T，LIU S，WANG Z Q，et al. Task segmentation and parallel optimization of DNN model[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2739-2752 (in Chinese).

1 相关工作

对神经网络模型计算任务进行切分并进行并行加速已有大量的相关研究工作。文献[8]通过手工切分模型后，将模型不同的计算任务分配在各个加速器执行，但未充分考虑加速设备不同的计算性能，存在计算负载不均衡、训练加速效果不明显、任务切分耗时较长的问题。文献[9]提出使用数学方法对模型运行时间进行预估后切分，但整个过程没有实际训练数据的支撑，缺乏准确性。文献[10-11]使用强化学习来自动确定模型并行性，以及硬件计算资源配置，但这种在线决策技术需要大量额外的计算时间和计算资源，不能高效实现模型并行。文献[12]提出的FlexFlow使用模型并行和数据并行性混合方式进行模型任务切分，但未考虑计算资源的计算特性，存在计算资源利用率低的问题。文献[13]提出了一种通过统计模型各层计算量和CUDA内存占用情况的模型切分方法，在固定加速器数量的前提下，通过对输入的分区数进行预设，实现静态切分的模型并行，在切分过程中忽略了模型内部各层之间的执行顺序、值传递和依赖关系，不能较好地对大规模神经网络模型进行动态分析和切分，在使用场景及具体切分策略中存在一定的局限性。近期，很多工作使用Transformer[14]、NASNet[15]、BERT[16]、GPT-3[17]等大规模模型进行训练。文献[18-19]提出通过在层内进行任务切分实现Tensor级模型并行，并引入流水线进一步提升计算性能。这种层内模型切分方式可以避免层间切分引起的硬件算力分配不均的问题，但只能在Transformer模型中实现，缺乏泛化性，同时在进行细粒度切分时会引入冗余的计算和通信量。

针对以上深度学习模型并行化过程中存在的问题，本文提出一种基于动态计算特性感知的深度学习模型计算任务动态切分及并行优化方法，整体处理框架如图1所示。首先，结合硬件计算资源特性，对大规模深度学习神经网络模型卷积过程进行动态分析，构建原始计算任务DAG。然后，通过增强反链模型分析原始DAG中任务之间的内部计算特性，构建符合计算迭代规范的反链DAG，再进行深度优先搜索生成拓扑排序后的反链DAG状态序列。在此基础上，结合硬件计算资源特性，通过动态规划进行任务切分，实现计算任务和计算资源的合理匹配。同时，结合流水线并行进一步加速模型的训练速度，实现细粒度深度学习模型任务并行切分和并行加速。

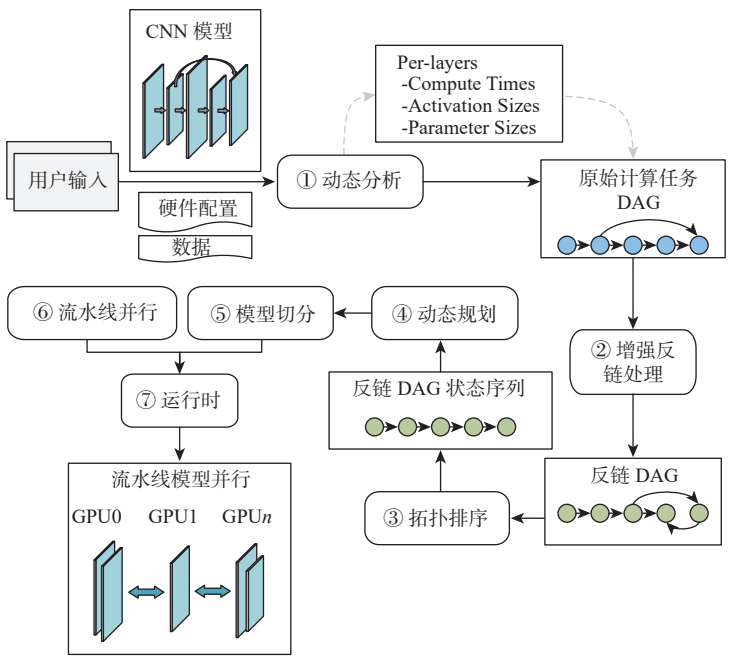


图1 深度神经网络模型切分及并行优化计算框架

Fig. 1 Deep neural network model segmentation and parallel optimization computation framework

2 基于增强反链的计算任务图构建

2.1 卷积神经网络模型计算特性动态分析

分析卷积神经网络(convolutional neural network, CNN)模型计算特性，主要通过在具体的硬件计算平台上，对输入的多个小批量数据进行训练，探测模型不同层计算时间和内存占用信息，以及层间执行顺序和依赖关系。通过稀疏化卷积神经网络层间的密集连接性来聚焦关键拓扑连接，将复杂的模型结构转换为易于分区聚类的原始计算任务

DAG, 进一步提高模型的表示能力, 作为后续任务切分的基础。

2.1.1　动态分析方法

动态分析方法是将模型结构转换为具体的 DAG 节点信息和边信息。首先, 在具体的计算平台上输入训练脚本, 对卷积神经网络模型进行训练, 识别层及层间各类信息。然后, 将层及层间各类信息封装构成 DAG 节点, 跟踪层间执行顺序得到 DAG 的边。最后, 组合节点和边, 将卷积神经网络执行过程转换为原始计算任务 DAG, 完成模型动态分析。

2.1.2　动态分析过程

1) 相关变量及数据结构定义。在整个动态分析过程中, 将多个小批量的数据输入模型进行训练。使用字典键值存储模型层信息, 将各层执行次序Layer$i$作为键, 层及层间操作类型Layer_type(卷积、激活、池化和完全连接等层类型, 以及 Concat、View、Add、Size、BatchNorm 等层间操作类型)和具体操作信息Layer_desc、计算时间(前向传播F_i和反向传播B_i)、内存占用信息(每层输出激活大小、反向传播中的梯度大小A_i、每层参数量P_i)作为值进行存储; 将层间执行顺序和依赖关系使用各层键Layer($i \rightarrow j$)存储; 模型在前向传播过程中每层的输出形状使用通道数、高、宽组成的列表{channels, h, w}记录, 用来构成各层和层间操作的输入或输出值传递。

定义原始计算任务 DAG 类, 有节点和边属性, 其中, 节点对象包含序号node$i$、类型、描述和各类参数信息, 边为节点间执行次序, 使用字典存储各节点对应的出边和入边。

2) 状态信息收集。根据定义的数据结构及记录的相关信息, 在模型训练过程中完成状态信息的收集。收集时, 主要将层执行次序键与节点序号对应, 层信息的值与节点剩余属性对应, 封装各节点信息构成节点对象, 形式为: [node$i$-Layer_type (Layer_ desc) -F_i, B_i, A_i, P_i]。将层及层间操作的执行顺序和依赖关系Layer($i \rightarrow j$)封装为 DAG 边, 边信息为: (node$i \rightarrow$ node$j$)。原始计算任务 DAG 中, 节点表示卷积神经网络模型执行过程中特征的聚合和转换, 边表示信息流[20]。同时, 在 DAG 构建过程中记录节点关系, 将node$i$的出边节点定义为Outnode$i$, 前序节点定义为Prenode$i$, 后序节点定义为Sucnode$i$。后续算法以节点信息和边信息作为输入。

2.1.3　动态分析结果表示

图 2 左侧是一个卷积神经网络模型局部结构示例, 对其进行动态分析。首先, 使用字典键值存储模型层信息, 将各层执行次序作为键, 层间操作类型和具体操作信息、计算时间、内存占用信息作为值。然后, 将各字典键值封装为节点对象(见图 2 中虚线上方信息), 映射到图 2 右侧所对应的节点, 层间的连接Layer($i \rightarrow j$)(表示执行顺序和依赖关系)映射为指示信息流(node$i \rightarrow$ node$j$)。最后, 将模型结构转换为原始计算任务 DAG, 如图 2 右侧所示。

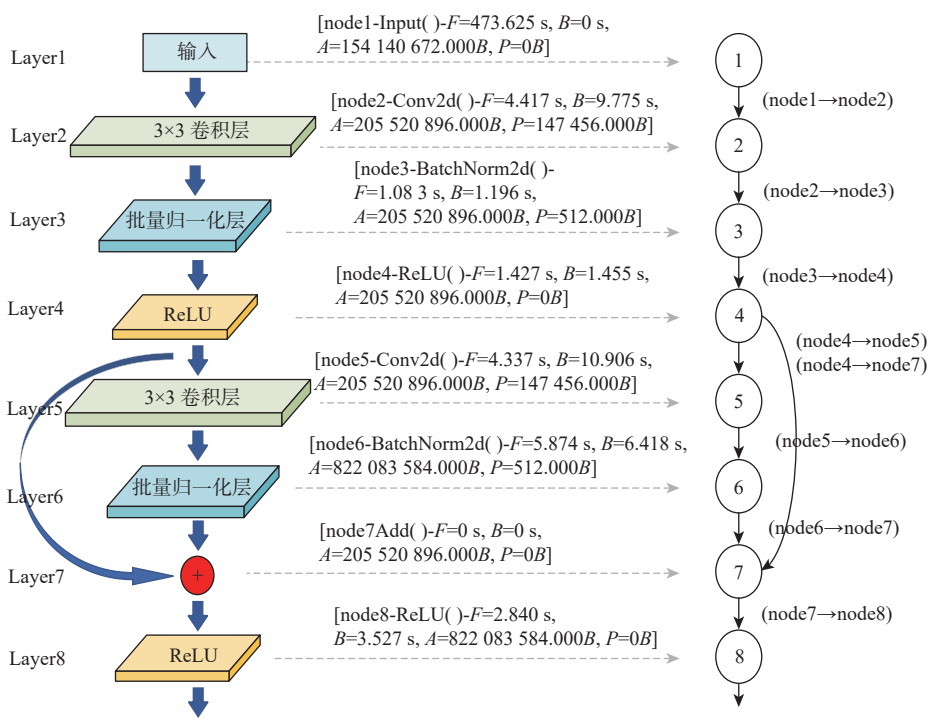


图 2　模型结构转换 DAG 示例

Fig. 2　DAG example of model structure transformation

通过以上卷积神经网络模型计算特性动态分析，重新构建了神经网络层之间的连通关系，聚焦关键拓扑连接，以便于检索同一空间变量的拓扑结构，从而为实现将复杂的模型层间计算任务的切分问题转化为高效的DAG节点分区聚类做好准备。

2.2 增强反链处理

在DAG分区聚类过程中，如果某些节点之间存在连接或重叠依赖，通过将DAG图中节点切分为边缘独立的子集，使得子集边缘指向同一方向，进而将子集转换为独立节点，组织各独立节点生成分区聚类后的DAG[21]。DAG分区聚类后增强了对节点间关系的粗粒度描述，从而可作为卷积神经网络模型计算任务切分的依据。具体是：将DAG切分为反链的形式，即将DAG中同一层次级别的相关节点构成一个集合。任意反链间关系独立，可作为分区的切入点。组合各反链可以达到DAG分区聚类的目的，同时也保持了DAG的拓扑结构和执行顺序[22]。

根据DAG中节点任意反链关系独立的特性，为解决由于原始DAG节点间连接紧密、逻辑关系复杂、存在连接或重叠依赖，导致任务切分难度大的问题，本文提出利用增强反链来组合各节点和部分前序节点，将原始DAG节点切分为边缘独立的集合，作为DAG分区的切入点，从而构成新的节点拓扑连接，弱化原始节点间的依赖关系，以减少模型切分难度。

2.2.1 增强反链理论

反链是DAG中的节点子集，通常由小于或等于2的节点数构成。具体是：将层次结构中同一级别内没有任何联系、执行工作相同的节点进行组合，其中，任何节点间都不存在路径连接。即反链是一种其内部节点作用相似、关系不可比较、存在重叠依赖的拓扑关系组合。通过反链可将DAG切分为多个节点子集，可实现DAG分区聚合的目的。任意反链中节点 u 与 v 间关系可表示为 $d(u,v)=\infty$，即 u 与 v 不可达、不可比较，反链的形式化表示为[23]

$$a=\{u,v\in r|\forall u,v\in a:d(u,v)=\infty,d(v,u)=\infty\} \tag{1}$$

式中：r 为DAG中所有节点的集合。

结合卷积神经网络模型层次间复杂的拓扑关系，以及转换后的计算任务DAG，通常存在节点数大于2的满足反链概念的节点子集。本文提出增强反链Aa，通过扩大DAG节点拓扑关系的搜索范围，组合同一层次级别的节点node$i$与其部分前序节点Prenode$i$构成边缘独立的节点子集{node$i$,Prenode$i$}，在较大邻域内遍历各节点与相似节点的重叠关系，主要遍历节点间的重叠依赖关系，即某一节点同一层级别的前序节点间的并列执行的关系，探索由多个节点构成的子集，进而对转换后的原始计算任务DAG进行更合理的切分。

2.2.2 增强反链处理过程

增强反链处理是通过分析逻辑关系紧密的原始DAG节点与其前序节点间的连接或重叠依赖关系，将原始DAG节点切分为边缘独立的节点集合，进而将集合转换为独立的反链DAG节点，弱化了原始DAG节点的依赖关系，将模型切分问题转化为高效的原始DAG分区聚类问题。

具体增强反链处理过程如下：

1）假设node$i$为原始DAG节点，若前序节点集合Prenode$i$中存在分支节点node$k$且绕开node$i$指向其他（非本身）节点。此时，由于节点内部运行机制和执行优先级模糊，且node$k$到后续节点在较大邻域内存在依赖关系重叠。

2）将node$k$与node$i$构成边缘独立的集合。原始DAG节点的增强反链为本身节点及其分支节点集合，即node$i$的增强反链Aai为{node$i$,node$k$}。

3）通过上述操作，得到原始DAG节点的各个增强反链，并转换为边缘独立的节点，作为构建反链DAG的基础。

4）分支节点的选取。在DAG中对关系重叠或相似的节点进行组合时，由于分支节点不同分支间关系并列且在同一层级，在前序节点的遍历过程中，分支节点是重点选取判断的目标。

具体分支节点选取算法如下。

算法1 分支节点node$k$选取算法。

输入： node$i$，出边节点集合{Outnode$i$}。

输出： node$i$增强反链Aai。

1. 初始化nodes_K = set()
2. 获取当前节点的所有前序节点{Prenode$i$}
3. for node$k$ in {Prenode$i$}:
4. if Outnode$k$ not in {Prenode$i$} and Outnode$k$!= node$i$ /*如果出边不在前序节点集合之中，且出边节点不是本身节点*/
5. 将node$k$添加到nodes_K
6. end for
7. 返回Aai = nodes_K + node$i$

以下为增强反链处理实例。对图3上部得到的原始DAG各节点进行增强反链处理，具体过程如下：对于node5，得到其前序节点集合(node1,node2,node3,node4)，其中，node4满足分支节点的条件。由于node4与其出边节点node5和node7存在依赖关系重叠，且不清楚到出边节点的内部运行机制和执行

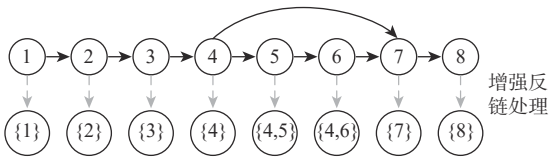


图 3　原始 DAG 和各节点对应的增强反链

Fig. 3　Original DAG and corresponding augmented antichain of each node

优先级, 组合 node4和 node5构成边缘独立的集合, 即 node5的增强反链为{node4, node5}。依据上述操作, 遍历原始 DAG 节点, 分析本身节点和其部分前序节点, 将重叠依赖的同一层次节点组合生成对应的增强反链, 进而将增强反链转换为如图 3 下部所示的边缘独立的节点。

2.3　反链 DAG 构建

通过增强反链处理, 获得原始 DAG 中各节点的增强反链后, 原始 DAG 被切分为各边缘独立的节点集合, 实现了将连接紧密、重叠依赖的节点关系拆分成松散的粗粒度描述, 将卷积神经网络模型切分问题转化成高效的原始 DAG 分区问题。反链 DAG 是为了组织各增强反链的执行次序, 在不破坏原始 DAG 的逻辑关系上, 通过遍历原始 DAG 节点间的执行顺序, 构建增强反链为节点与边的依赖关系图。

1) 反链 DAG 构建过程。具体构建反链 DAG 时, 先在增强反链的基础上计算后继反链, 从而作为构建反链 DAG 的依据, 后继反链是在本身节点已得到增强反链的基础上, 通过遍历本节点增强反链的后序节点, 确定后序节点的增强反链与本节点增强反链的执行顺序, 保证各增强反链有序执行, 从而生成反链 DAG。

后继反链获取过程如下。

算法 2　node*i*后继反链 nextnode*i*获取算法。

输入: node*i*, {Outnode*i*}, Sucnode*i*。

输出: node*i*后继反链 nextnode*i*。

1. 初始化 nextnode*i*s = []
2. 获取节点的增强反链 Aa*i*
3. for node*i* in Aa*i*
4. for Outnode*i* in {Outnode*i*}
5. if SucOut_node*i* not in Aa*i*

/*出边节点的后序节点如果不在增强反链中*/

6. 构建 nextnode*i* = AaSuc_Out_node*i*
7. 将 nextnode*i*添加到 nextnode*i*s
8. end for
9. end for
10. 返回 nextnode*i*s

假设 node*i*为原始 DAG 节点, Aa*i*为节点的增强反链。首先, 获取 Aa*i*中节点的出边节点(顺序执行的下一个节点)集合{Outnode*i*}。其次, 对于每个出边节点 Outnode*i*, 若其后序节点 SucOut_node*i*(顺序执行的后续所有节点)不在本身增强反链 Aa*i*中, 则确定此后序节点的增强反链 AaSuc_Out_node*i*是 node*i*的后继反链之一。然后, 对 Outnode*i*的后序节点依次遍历, 获得满足条件所有节点的增强反链, 完成对 node*i*后继反链的计算过程。最后, 将 node*i*增强反链 Aa*i*与后序节点的增强反链 AaSuc_Out_node*i*转换为节点, 并连接构成 DAG 的边。

2) 反链 DAG 构建实例。通过图 4 中的实例进一步说明反链 DAG 构建过程。为获取图 4 上部中 node5的后继反链, 首先, 遍历其增强反链中节点 node4和node5, 得到出边节点集合(node5, node6, node7), 其中, node5已在本身增强反链中不予考虑。然后, 得到node6的后序节点集合(node7, node8), 且后序节点都不在增强反链{node4, node5}中, 因此, node6的增强反链{node4, node6}为 node5的后继反链之一; 同理, node7的增强反链{node7}也是 node5后继反链之一。因此, 得到 node5的后继反链为{node4, node6}和{node7}。最后, 连接 node5的增强反链和其后继反链构成反链 DAG 的边。重复以上步骤, 计算各节点后继反链信息, 连接各节点的增强反链和其后继反链构成反链 DAG 的边, 进一步完善图结构的边, 从而将原始 DAG 转换为反链 DAG, 如图 4 下部所示。

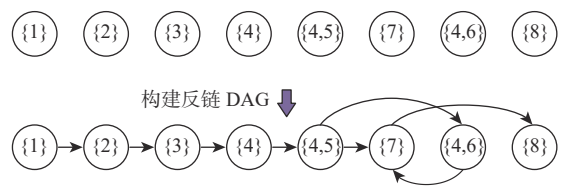


图 4　原始 DAG 各节点增强反链及构建反链 DAG

Fig. 4　Augmented antichain of each node of original DAG and construction of antichain DAG

3　反链 DAG 状态序列生成

由于生成的反链 DAG 执行次序不符合原始卷积神经网络模型执行顺序, 而原始卷积神经网络模型计算任务的执行逻辑和原始 DAG 图是一致的。为在不破坏原始卷积神经网络模型计算任务执行逻辑关系的基础上组织各增强反链的执行次序, 通过深度优先遍历, 对反链 DAG 中增强反链执行次序进行拓扑排序, 形成反链 DAG 节点的一个线性序列, 即反链 DAG 状态序列, 以确定任务的正确执行顺序, 确保在任意节点任务到达之前, 其前序节点任务都已完成, 避免冲突现象发生, 从而保证原始卷积神经网络模型的训练逻辑。最终, 将得到的

反链DAG状态序列作为后续模型切分及并行优化算法的输入。

实例说明，对图5上部中反链DAG各节点进行深度优先搜索的拓扑排序，将原本执行次序混乱的增强反链进行排列，并连接各节点边缘，形成符合模型迭代规范的增强反链线性序列，构成反链DAG状态序列，如图5下部所示。

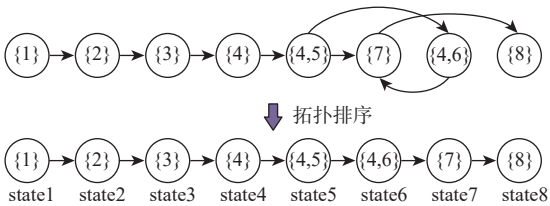


图5 反链DAG和反链DAG状态序列

Fig. 5 Antichain DAG and antichain DAG state sequence

4 流水线并行优化

流水线并行是训练大规模深度神经网络模型的有效方法。流水线并行训练的研究通常分为2类，即同步训练的流水线并行和异步训练的流水线并行。同步训练的流水线并行方法需要在相邻训练迭代之间进行必要的梯度同步，以确保收敛。在运行时，尽可能多地调度并发流水阶段，以最大限度地提高设备利用率[24-25]。异步训练的流水线并行的训练方式为：将微批量连续插入流水阶段，并丢弃原始同步操作以实现最大吞吐量[26-28]。由于设备间通信量少、设备利用率高、训练耗时更短，流水线并行已成为加速大规模神经网络模型并行训练的有效手段。

流水线并行将要训练的模型层切分为多个执行阶段，每个执行阶段包含一组连续处理层，每个执行阶段分别映射到一个单独的GPU上进行处理。每个GPU对其执行阶段中的所有层执行前向传播和反向传播[29-31]。图6中对比了传统模型并行和流水线并行方法。其中，神经网络模型跨4个GPU进行切分，F_k为第k个单元的前向传播函数，B_k为反向传播函数，$F_{i,k}$为第i个GPU上运行第k个微批量的前向传播函数。可以看到，在相同训练时间下，传统模型并行方法中，每个小批量在任何时刻只有一个阶段处于活动状态，导致其他GPU处于空闲状态，产生很多气泡开销。流水线并行将输入的小批量切分为更小的微批量，使不同的GPU能够同时处理不同的微批量，提高了GPU的利用率，并且减少了气泡开销，同步应用梯度。

本文并行优化的目标是最小化模型训练的总体时间。通过将反链DAG状态序列切分为流水线并行中的不同阶段进行训练，使用动态规划逐次优化流水线中最慢阶段的训练耗时，最终得到具体神经网络模型任务切分和计算任务到计算资源的动态配置策略。

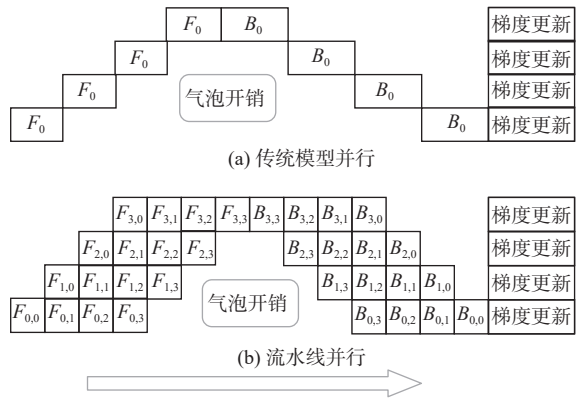


图6 传统模型并行和流水线并行

Fig. 6 Traditional model parallelization and pipeline parallelization

4.1 动态规划

动态规划将拓扑排序后的反链DAG状态序列作为输入，结合具体硬件计算资源，将状态序列切分为多个执行阶段，最小化使用不同GPU数量时流水线并行最慢阶段所花费的时间。通过逐次优化各阶段计算时间，找到流水线并行最优解。分析最优解，得到反链DAG状态序列最佳分割点数组，并作为模型切分的输入。

动态规划具体执行过程主要包括3个步骤：准备阶段、计算阶段和分析阶段。准备阶段计算状态序列相关参数，计算阶段依据硬件环境得到流水线并行最优化结果，分析阶段通过分析最优化结果计算反链DAG状态序列得到最佳分割点数组。

4.1.1 准备阶段

首先，将反链DAG状态序列(state1, state$N$)作为输入。计算每个执行状态state$i$的计算时间$T_{\text{state}i}$、激活大小$O_{\text{state}i}$和参数大小$P_{\text{state}i}$，具体是得到state$i$的增强反链，遍历其中每个反链节点的前序节点，累加求和。然后，通过逐次相减计算每个状态到后续状态间(state($i \to j$))的计算时间、激活大小和参数大小，并将state$i$增强反链节点的激活值进行累加，得到state$i$的输出激活值。最后，将结果保存输入计算阶段。

4.1.2 计算阶段

定义$T'(1 \to N, m)$表示使用m个GPU在state1和state$N$之间最佳流水线并行中最慢阶段所用的时间，公式为

$$T'(1 \to N, m) = \min_{1 \leqslant k \leqslant N} \min_{1 \leqslant m' < m} \max\{T'(1 \to k, m-m'), 2C_k, T(k+1 \to N, m')\} \tag{2}$$

式中：$T'(1 \to k, m-m')$为state1和state$k$之间使用$m-m'$

个 GPU 流水线并行最慢阶段所用时间, 是优化迭代的主要部分; $2C_k$为state$k$和state$(k+1)$之间传递激活和梯度通信所用的时间; $T(k+1\to N,m')$为已切分阶段所用的最短时间。

根据准备阶段保存的各状态间相关参数信息, 结合当前硬件环境下 GPU 数及 GPU 间通信带宽, 使用式(2)对状态序列(state1, state$N$)进行计算, 具体步骤如下:

步骤 1　遍历各state$(i\to j)$使用不同 GPU 数时当前流水线并行的最小运行时间(计算时间和通信时间之和), 并使用三元数组 Ta[i][j][m]存储各个state$(i\to j)$最小运行时间和当前 GPU 数量m。

步骤 2　从后往前依次将各 state$(i\to j)$切分为 state$(i\to k)$和 state$(k+1\to j)$前后 2 部分, 通过对state$(k+1\to j)$的激活和梯度内存占用情况进行判断, 如果超出内存则跳过此切分情况, 不予分割考虑。

步骤 3　分别获取前部分 state$(i\to k)$运行时间、输入传输时间和后部分 state$(k+1\to j)$运行时间、输出传输时间, 将四者中最大值与步骤 1 的最小运行时间进行比较, 得到当前 GPU 数量下state$(i\to j)$分区中的最小流水线运行时间。

步骤 4　将各 GPU 数量情况下 state$(i\to j)$的最小流水线并行运行时间、最优分区间隔、所用 GPU 数量m保存到三元数组 Ta[i][j][m]中。

通过计算阶段得到的结果为使用m个 GPU 切分反链 DAG 状态序列(state1, state$N$)的最优化三元数组 Ta[1][N][m], 包括流水线并行最优阶段所用的最短时间$T(k+1,N,m')$、各阶段间最优分割点间隔(state$k$, state$(k+1)$)和使用的最优机器数, 并将最优化三元数组输入分析阶段进行反链 DAG 状态序列的最优分割点分析。

4.1.3　分析阶段

通过计算阶段得到的最优化三元数组, 遍历在不同 GPU 数量下, 得到反链 DAG 状态序列所有最优分割点间隔, 构成最优分割点数组作为模型切分的输入。

根据流水线并行训练中 GPU 从高层到底层依次分配原则, 通过分析最优化三元数组 Ta[1][N][m], 将反链 DAG 状态序列(state1, state$N$)在m个 GPU 上切分为各个阶段, 并得到最优分割点数组。

步骤 1　定义 start和 end分别为反链 DAG 状态序列的起始点和终止点, 得到反链 DAG 状态序列在m个 GPU 上流水线并行最优化三元数组 Ta[start][end][m]。

步骤 2　定义数组splits记录分割点位置。获取在使用m个 GPU 下 (state1, state$N$)的最优分割点。输入起始点start = 1和终止点end = N到Ta[start][end][m]中, 遍历最优化三元数组 Ta[1][N][m]得到最优分割点间隔为 (state$k$, state$(k+1)$), 即分割点为k, 将k添加到splits中。

步骤 3　定义 nextsplit为首个分割点, prevsplit为末尾分割点。此时, nextsplit = k, prevsplit = N。将状态序列切分为一个最优阶段stage k(state$(k+1\to N)$)和剩余最慢阶段状态序列分区remainstates(state$(1\to k)$)。同时, 将末尾分割点 prevsplit更新为当前最优分割点位置nextsplit = k。

步骤 4　对剩余最慢阶段状态序列进行分析, 获取在$m-1$个 GPU 下(state0, state$k$)的最优分割间隔。即输入 start = 1, end = prevsplit到 Ta[start][end][m]中, 遍历剩余状态序列最优化三元数组 Ta[1][k][$m-1$], 得到下一个最优分割点。

步骤 5　nextsplit不为空是继续分割的条件。从后往前通过依次遍历反链 DAG 状态序列的最优化三元数组 Ta, 重复上述步骤直到nextsplit为空, 即分割结束, 返回最优分割点数组。

图 7 为反链 DAG 状态序列切分示意图。

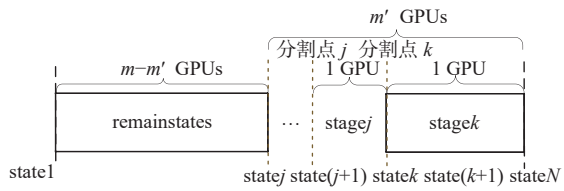


图 7　反链 DAG 状态序列切分示意图

Fig. 7　Schematic of antichain DAG state sequence segmentation

4.2　模型切分

模型切分是实现模型并行的重要步骤, 通过将卷积神经网络模型切分为各个子模型, 并部署到多个 GPU 上并行执行, 可实现大规模计算任务高效执行, 提高训练效率[32]。

本文在模型切分时利用动态规划得到的最优分割点数组, 将反链 DAG 状态序列切分为各个执行阶段, 再回溯各阶段内部增强反链信息到原始 DAG, 生成各个子图, 最终将子图映射回卷积神经网络模型, 得到各子模型结构, 完成模型切分过程。模型切分示意如图 8 所示, 具体分为以下步骤:

1) 根据动态规划得到反链 DAG 状态序列最优分割点数组, 使用各分割点对反链 DAG 状态序列进行切分, 生成反链 DAG 状态序列分区, 如图 8 中反链 DAG 状态序列分区部分所示, 处在不同状态区间内的增强反链构成各个执行阶段状态信息。

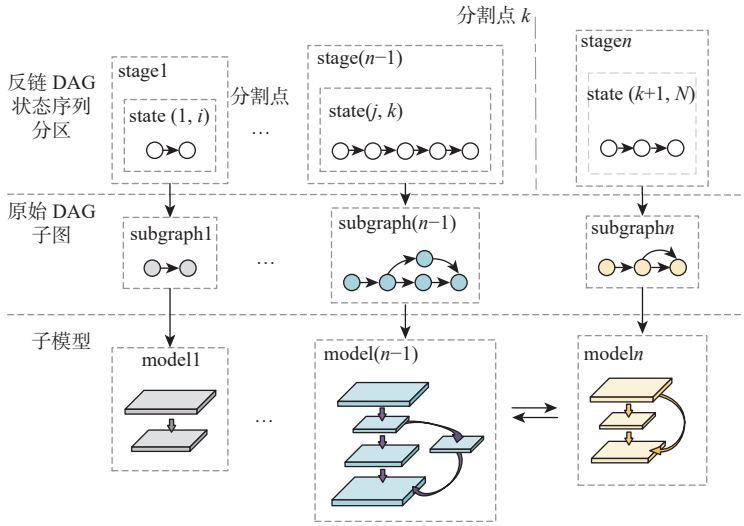


图 8 模型切分示意图

Fig. 8 Model segmentation process

2) 将各个阶段状态序列的增强反链节点回溯到原始 DAG 节点，为其对应的原始 DAG 节点设置分区编号，生成一系列子图，如图 8 中原始 DAG 子图部分所示。

3) 将各原始 DAG 子图节点及信息流，映射回原始神经网络模型，应用动态分析时记录的模型内部层输入、输出形状和相关参数信息，整合各分区内层间值传递和依赖关系，应用模板文件生成各个卷积神经网络子模型，如图 8 中子模型部分所示。

4) 遍历各子模型每层的输入，构建子模型输入张量；调用对应的 forward 函数，构建子模型输出张量；遍历每层的输出，设置子模型的张量类型和形状。连接各子模型间输入张量、输出张量、张量类型和形状，生成总体模型迭代规范配置文件，确保各子模型在执行前向传播和反向传播过程时符合原模型的逻辑顺序，从而完成模型计算任务切分，图 8 中子模型部分双向箭头所示为各子模型间前向传播和反向传播过程。

5) 将生成的总体模型任务切分迭代规范配置文件作为后续流水线并行优化算法的输入。

4.3 流水线并行优化算法

模型并行由于采用单纯的多加速器顺序执行，不同加速器训练时空闲状态产生的气泡开销过大，导致训练时间过长。本文流水线并行方法将批量切分为更小的微批量，通过 GPU 间同时处理多个微批量达到密集迭代的训练效果，以提高 GPU 的利用率，减少气泡开销，大幅降低模型并行训练时间开销。

流水线并行优化整体算法主要包括如下过程：首先，得到模型切分后的子模型 model$i$、各子模型间的输入和输出连接；其次，将子模型依次放在 GPUi上，引入总体模型任务切分迭代规范配置文件，连接各子模型，保证子模型间数据传输符合原模型的逻辑顺序；然后，将不同大小的小批量切分为微批量，在各 GPU 上依次密集迭代执行，并使用非默认 CUDA 流设置每个 GPU，对每个微批量使用不同的流，确保计算任务之间的明确依赖关系，使得 GPU 计算不必等待微批量数据的输入，实现 GPU 通信与后续微批量计算的显著重叠，减少了训练时间；最后，在每次小批量迭代结束后同步更新模型参数信息。

具体优化执行过程如下。

算法 3 流水线并行优化执行算法。

输入： 数据集D，配置文件{model$i$}、{inputmodel$i$}、{outputmodel$i$}，微批量大小M。

输出： 更新后的参数。

1. 初始化使用D构建 batch size 迭代数组
2. for(i,model,gpu) in {model$i$,GPUi}
3. 将各子模型依次放在GPU上
4. end for
5. 重写 forward 函数
6. splits=batch size.split(M)

/*将切分 batch size 为多个M大小*/

7. set inputmodel$i$.to('CUDA : i')
8. set outputmodel$i$.to('CUDA : $i+1$')
9. 定义累积梯度列表 ret=[]
10. for split in splits:
11. set inputmodel$i$.to('CUDA : i')
12. set outputmodel$i$.to('CUDA : $i+1$')
13. end for
14. 将每次迭代累积的梯度值添加到 ret

15. Return ret/*同步参数信息*/

5　实验验证

为验证本文提出的基于特性感知的卷积神经网络模型任务切分及模型并行优化方法的有效性, 首先, 将本文切分方法(antichain_mp、antichain_pp)和传统手工切分方法(handsplit_mp)、基于每层 CUDA 内存占用情况切分方法(cudasize_mp)[13]、基于每层算力切分方法(hashrate_mp)[13]进行切分模型耗时对比。然后, 选取 CIFAR-10 数据集, 将各切分方法使用 ResNet-50 模型进行并行训练, 对比各方法并行训练耗时和训练精度。最后, 分析微批量大小(micro-batch size)对本文流水线并行训练计算效率的影响。

5.1　实验设置

本文实验是在单机多卡硬件平台上进行测试, 分为 2 种情况, 包括 4 个 GPU 和 8 个 GPU 进行模型并行训练, 采用 NVIDIA Tesla V100-SXM2 GPU, 显存为 32 GB, 各 GPU 间通信带宽为 100 GB/s。选取 CIFAR-10 数据集, 包含训练数据集 50 000 张图像, 测试数据集 10 000 张图像, 并进行图像增强操作, 设置图像尺寸为 224 像素×224 像素。使用深度学习框架 PyTorch 1.10.0、Python 3.8、CUDA 10.2、cuDNN 8.1.1 搭建实验运行环境。

5.2　实验结果分析

5.2.1　不同模型切分方法耗时对比

将本文方法和传统手工切分方法 (handsplit_mp)、基于每层 CUDA 内存占用情况切分方法 (cudasize_mp)、基于每层算力切分方法 (hashrate_mp), 针对不同类型卷积神经网络模型[33-38]在 4 GPU 和 8 GPU 中进行切分, 对不同切分方法所需时间进行对比, 实验结果如表 1 所示。

实验结果表明, 针对不同类型卷积神经网络模型, 分别在 4 GPU 和 8 GPU 切分过程中所花费的时间, 本文提出的基于特性感知的模型任务切分策略对比其他方法整体耗时最少。相比传统手工切分方法 (handsplit_mp), 本文方法大幅度提高了模型切分的效率, 在节省手工切分成本的同时, 降低了人为因素带来的切分误差。

由于 Inception-v3[35]和 DenseNet-121[37]模型参数量大, 层之间关系紧密, 相比基于每层 CUDA 内存占用情况切分方法 (cudasize_mp) 不需考虑模型内部机制, 本文方法基于增强反链对模型任务计算图进行切分, 深度分析层间执行顺序和依赖关系, 因而所花费的时间稍多。总体来说, 本文方法通过感知硬件计算资源特性, 在切分过程中根据不同类型卷积神经网络模型特点进行动态切分, 使得各 GPU 间达到负载均衡, 可在保证原始模型执行

表 1　不同模型切分方法在 4 GPU 和 8 GPU 上的耗时对比

Table 1　Time consumption comparison of different model segmentation methods on 4 GPU and 8 GPU

模型	GPU数	handsplit_mp 耗时/min	cudasize_mp 耗时/s	hashrate_mp 耗时/s	本文方法耗时/s
VGG-16[33]	4	30	6.76	8.72	0.01
	8	>60	7.84	9.34	0.03
ResNet-18[34]	4	35	1.98	2.14	0.07
	8	>60	3.02	4.11	0.21
ResNet-50	4	45[36]	2.76	3.56	0.38
	8	>60	3.80	6.31	1.17
ResNet-101	4	>60	4.12	5.78	1.20
	8	>60	5.89	9.24	3.69
AlexNet	4	>60	7.12	8.45	0.02
	8	>60	8.94	10.26	0.01
ResNext-50	4	>60	2.79	3.72	0.41
	8	>60	4.97	8.11	1.17
ResNext-101	4	>60	4.53	6.01	1.21
	8	>60	6.12	8.37	3.58
Inception-v3[35]	4	>60	2.64	12.25	3.07
	8	>60	3.76	6.23	4.12
DenseNet-121[37]	4	>60	1.28	1.59	1.29
	8	>60	3.86	4.11	3.73
SqueezeNet1_0[38]	4	>60	2.73	3.59	0.06
	8	>60	8.46	4.27	0.20

逻辑的基础上，提高模型切分的效率、准确性和泛化性。

5.2.2 并行训练耗时对比

为验证本文方法在提升并行训练速度方面的有效性，将 ResNet-50 卷积神经网络模型分别使用手工切分方法（handsplit_mp）、基于每层算力切分方法（hashrate_mp）、基于每层 CUDA 内存占用情况切分方法（cudasize_mp）和本文切分方法（antichain_mp、antichain_pp），在 4 GPU 和 8 GPU 上完成模型计算任务切分后，进行模型并行训练对比。主要对 ResNet-50 模型在不同批量大小（batch size）下，训练 70 个 Epoch 达到稳定准确率时所花费的时间进行对比，并计算各切分方法并行加速比，实验结果如图 9～图 12 所示。

实验结果表明，本文提出的模型切分流水线并行方法相比其他方法具有更好的训练效果。从实验结果可以看出，当 batch size 较小时，由于任一时刻各方法模型并行训练过程中只有一个 GPU 在工作，其他 GPU 均空闲，气泡开销大，导致各个方法训练时间较长。随着 batch size 逐渐增大，虽然仍存在 GPU 空闲状态，但逐渐变大的 batch size 增加了 GPU 的内存占用率，相对提高了 GPU 的利用率，使得整体的训练耗时相对减少，当 batch size 达到 2^9 时，整体训练时间趋于平稳。

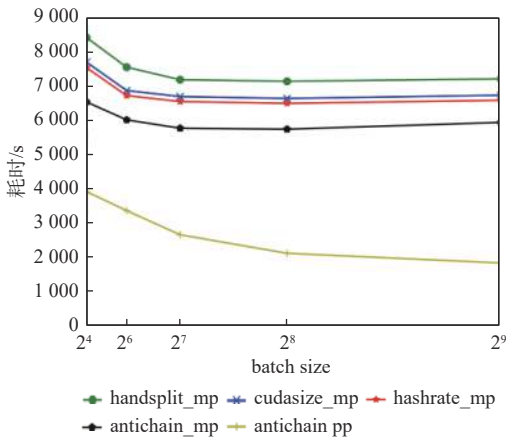


图 9 4 GPU 模型并行执行时间

Fig. 9 Parallel execution time of 4 GPU model

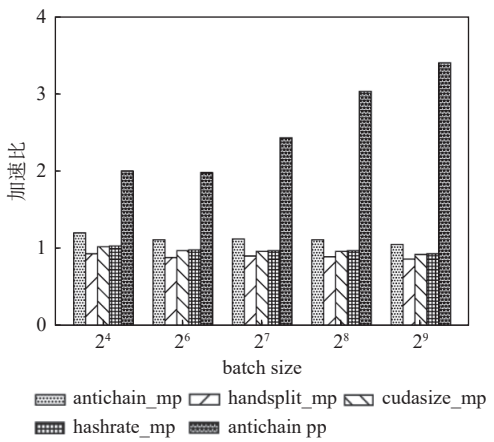


图 10 4 GPU 模型并行加速比

Fig. 10 Parallel speedup of 4 GPU model

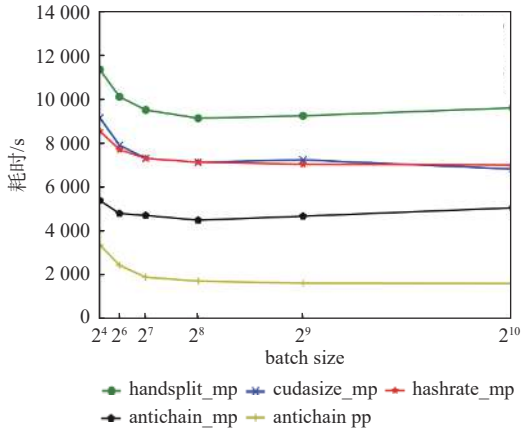


图 11 8 GPU 模型并行执行时间

Fig. 11 Parallel execution time of 8 GPU model

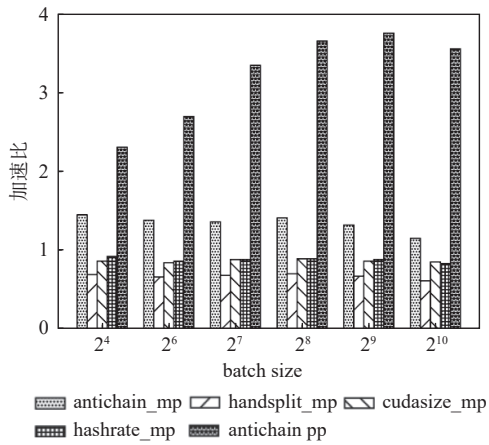


图 12 8 GPU 模型并行加速比

Fig. 12 Parallel speedup of 8 GPU model

本文方法在模型并行基础上，通过加入流水线并行（antichain_pp）后，由于不同微批量密集迭代，可以启动更多微小的 CUDA 内核，增大了训练吞吐量，进一步提高了 GPU 的利用率。随着 batch size 不断增大，实现了 GPU 间计算与通信重叠，可更大程度上减少模型并行中的气泡开销，显著降低训练时间。从实验结果可以看出，当 batch size 达到 2^9 时，本文模型切分并行优化方法在 4 GPU 上与已有方法相比，最大训练时间可以降低近 90 min，平均可以降低 84 min；在 8 GPU 情况下，最大训练时差近 133 min，平均可以降低 107 min；4 GPU 情况下，最大加速比达到 3.4，8 GPU 情况下最大加速比为 3.76。

5.2.3 并行训练精度分析

为验证本文方法在提升模型训练精度方面的有效性，分别使用本文方法（antichain_pp）、手工切

分方法(handsplit_mp)、基于每层算力切分方法(hashrate_mp)和基于每层 CUDA 内存占用情况切分方法(cudasize_mp), 对 ResNet-50 模型在 4 GPU 和 8 GPU 情况下, 并行训练 70 个 Epoch 的精度变化进行分析, 具体实验结果如图 13 和图 14 所示。

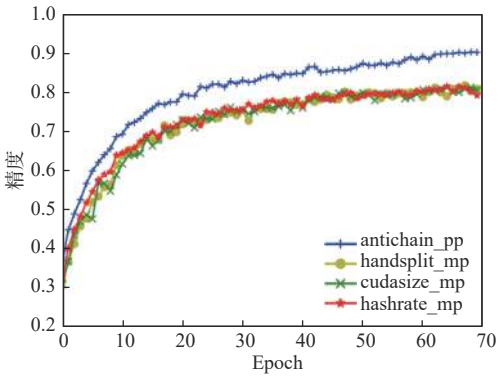


图 13　4 GPU 不同切分方法并行训练精度

Fig. 13　Parallel training accuracy of 4 GPU with different segmentation methods

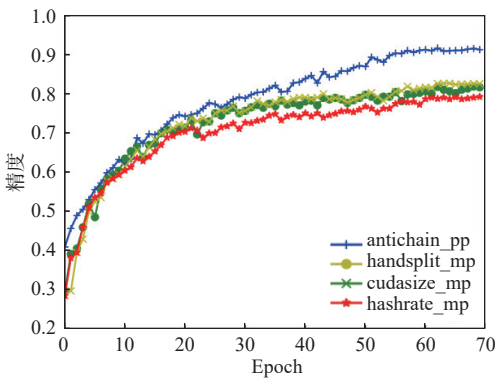


图 14　8 GPU 不同切分方法并行训练精度

Fig. 14　Parallel training accuracy of 8 GPU with different segmentation methods

实验结果表明, 当使用 4 GPU 模型并行训练, 采用 batch size 为 2^9 时, 从图 13 中可以看出, 本文方法和手工切分方法 (handsplit_mp)、基于每层算力切分方法 (hashrate_mp)、基于每层 CUDA 内存占用切分方法 (cudasize_mp) 相比, 在训练开始时训练精度存在小的差异, 随着训练批次的增加, 本文方法训练精度提升效果更加明显。由于其他方法模型并行中数据不能在 GPU 间重叠训练, 各 GPU 需要等待其他 GPU 梯度传递, 延迟了模型参数更新, 训练精度提升速度缓慢。本文方法在保证子模型间依赖关系的同时, 将批量切分为更小的微批量, 在各个 GPU 间密集迭代训练同步更新参数, 实现了计算与通信同步, 加速了模型参数信息学习, 从而进一步提升了训练精度。最终, 在各方法模型训练精度趋于平稳时, 最大训练精度为: 手工切分方法 (handsplit_mp) 为 0.814, 基于每层算力切分方法 (hashrate_mp) 为 0.815, 基于每层 CUDA 内存占用情况切分方法 (cudasize_mp) 为 0.831, 本文方法达到 0.902。

当使用 8 GPU 模型并行训练时, 各 GPU 分摊模型参数量较 4 GPU 少, 可增大 batch size 为 2^{10} 进行训练。从图 14 中可以看出, 本文方法在训练开始时模型精度提升快, 在之后与手工切分方法 (handsplit_mp)、基于每层算力切分方法 (hashrate_mp) 和基于每层 CUDA 内存占用情况切分方法 (cudasize_mp), 在训练中模型精度增长速度相当。当 Epoch 达到 30 以后, 子模型在 GPU 间传输梯度时, 由于来自各 GPU 的延迟对整个梯度更新的影响, 在前向传播和反向传播时会损失一部分精度, 使得各方法训练精度逐渐趋于平稳。本文流水线并行时, 由于使用更多微批量数据进行密集迭代执行, 一个 GPU 在结束一个微批量训练后, 可以立即开始下一个微批量训练, 可提高 GPU 利用率, 可学习更多参数信息来拟合模型, 降低了模型训练损失率。最终, 在训练精度达到平稳后, 最大训练精度为: 手工切分方法 (handsplit_mp) 为 0.818, 基于每层算力切分方法 (hashrate_mp) 为 0.792, 基于每层 CUDA 内存占用情况切分方法 (cudasize_mp) 为 0.826, 本文方法最高精度达到 0.916。

5.2.4　微批量大小对并行优化方法性能影响分析

实验中发现, 本文流水线并行优化方法在进行模型训练时, 批量切分的微批量大小会直接影响并行训练耗时, 需要结合具体 batch size 大小和 GPU 算力调整微批量大小, 以获得较好的并行训练加速效果。本节将 ResNet-50 模型在 4 GPU 和 8 GPU 分别基于 batch size 为 2^6、2^7、2^8 和 2^7、2^8、2^9, 通过切分批量为不同大小微批量训练 70 个 Epoch, 对比本文流水线并行训练所花费的时间, 不同微批量大小情况下, 训练耗时如图 15 和图 16 所示。

从实验结果可以看出, 当将 ResNet-50 使用 4 GPU, batch size 分别为 2^8、2^7 和 2^6 进行流水线并

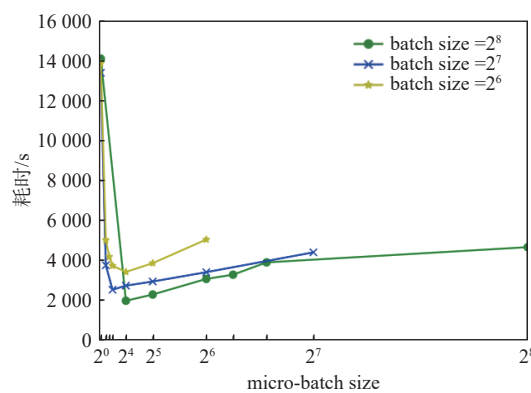


图 15　4 GPU 不同微批量大小流水线并行训练耗时

Fig. 15　Parallel training time of 4 GPU pipeline with different micro-batch sizes

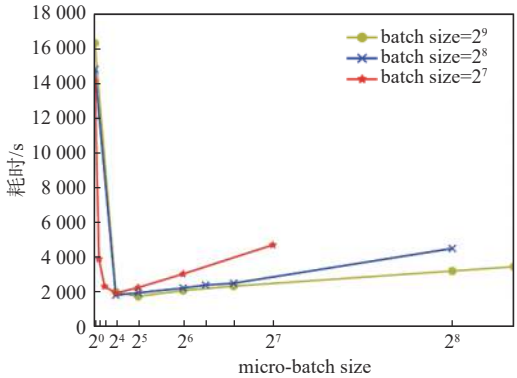


图 16 8 GPU 不同微批量大小流水线并行训练耗时

Fig. 16 Parallel training timeof 8 GPU pipeline with different micro-batch sizes

行训练时，微批量分别为 2^4、2^3 和 2^4 时模型训练耗时最小；使用 8 GPU，batch size 分别为 2^9、2^8 和 2^7 时，微批量大小分别为 2^5、2^4 和 2^4 时流水线并行训练耗时最小。可以看出，使用不同数量 GPU 进行流水线并行训练时，增大微批量的大小并不能直接减少训练耗时。如图 15 和图 16 所示，随着微批量增大，当超过最优微批量时，流水线并行训练所花费的时间反而增长。主要因为在使用不同大小的 batch size 和不同数量的 GPU 训练过程中，虽然较大的微批量可以实现在 GPU 间较快的信息传递，但较大的微批量使得 GPU 间参数同步操作耗时增长，延迟的梯度更新会增加额外的气泡开销，从而导致流水线并行整体执行时间增加。通过分析以上不同微批量大小情况下的并行训练耗时，选择训练耗时最小的 micro-batch size 作为最优微批量大小，将最优微批量应用到流水线并行优化训练中，以达到降低训练耗时的目的。

6 结 论

1）本文方法通过增强反链处理，合理匹配计算任务与资源。在 4 GPU 和 8 GPU 切分过程中整体耗时最少，显著提高了模型切分效率，减少了人为误差。

2）本文方法结合流水线并行加速模型训练。在 4 GPU 上，最大训练时间减少近 90 min，平均减少 84 min；在 8 GPU 上，最大训练时间减少 133 min，平均减少 107 min，提升了训练效率。

3）本文方法通过密集迭代训练同步更新参数，提升了模型训练精度。在 4 GPU 和 batch size 为 2^9 时，训练精度达 0.902；在 8 GPU 和 batch size 为 2^{10} 时，训练精度达 0.916。

后续将研究分布式环境下的任务切分及并行优化策略，以满足更大规模数据和计算需求，并探索多模态深度学习模型的并行优化方法，以适应不同应用场景下大模型的高效训练需求。

参考文献（References）

[1] KIM Y, CHOI H, LEE J, et al. Efficient large-scale deep learning framework for heterogeneous multi-GPU cluster[C]//Proceedings of the IEEE 4th International Workshops on Foundations and Applications of Self Systems. Piscataway: IEEE Press, 2019: 176-181.

[2] NARAYANAN D, PHANISHAYEE A, SHI K, et al. Memory-efficient pipeline-parallel dnn training[C]//Proceedings of the International Conference on Machine Learning. New York: PMLR Press, 2021: 7937-7947.

[3] 朱泓睿, 元国军, 姚成吉, 等. 分布式深度学习训练网络综述[J]. 计算机研究与发展, 2021, 58(1): 98-115.
ZHU H R, YUAN G J, YAO C J, et al. Survey on network of distributed deep learning training[J]. Journal of Computer Research and Development, 2021, 58(1): 98-115(in Chinese).

[4] TAN M, LE Q. EfficientNet: Rethinking model scaling for convolutional neural networks[C]//Proceedings of the International Conference on Machine Learning. New York: PMLR Press, 2019: 6105-6114.

[5] LEE S, KIM J K, ZHENG X, et al. On model parallelization and scheduling strategies for distributed machine learning[C]//Proceedings of the 27th International Conference on Neural Information Processing Systems. New York: ACM, 2014: 2834-2842.

[6] KIM J K, HO Q, LEE S, et al. STRADS: A distributed framework for scheduled model parallel machine learning[C]//Proceedings of the 11th European Conference on Computer Systems. New York: ACM, 2016: 1-16.

[7] 朱虎明, 李佩, 焦李成, 等. 深度神经网络并行化研究综述[J]. 计算机学报, 2018, 41(8): 1861-1881.
ZHU H M, LI P, JIAO L C, et al. Review of parallel deep neural network[J]. Chinese Journal of Computers, 2018, 41(8): 1861-1881 (in Chinese).

[8] KRIZHEVSKY A. One weird trick for parallelizing convolutional neural networks[EB/OL]. (2014-04-26)[2022-08-15]. https://arxiv.org/abs/1404.5997.

[9] BÁRÁNY I, GRINBERG V S. Block partitions of sequences[J]. Israel Journal of Mathematics, 2015, 206(1): 155-164.

[10] MIRHOSEINI A, PHAM H, LE Q V, et al. Device placement optimization with reinforcement learning[C]//Proceedings of the International Conference on Machine Learning. New York: PMLR Press, 2017: 2430-2439.

[11] DING Z X, CHEN Y R, LI N N, et al. Device placement optimization for deep neural networks via one-shot model and reinforcement learning[C]//Proceedings of the IEEE Symposium Series on Computational Intelligence. Piscataway: IEEE Press, 2020: 1478-1484.

[12] JIA Z, ZAHARIA M, AIKEN A. Beyond data and model parallelism for deep neural networks[EB/OL]. (2018-07-14)[2022-08-15]. https://arxiv.org/abs/1807.05358.

[13] KIM C, LEE H, JEONG M, et al. Torchgpipe: On-the-fly pipeline parallelism for training giant models[EB/OL]. (2020-04-21)[2022-08-15]. https://arxiv.org/abs/2004.09910.

[14] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM, 2017:

6000-6010.

[15] ZOPH B, VASUDEVAN V, SHLENS J, et al. Learning transferable architectures for scalable image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 8697-8710.

[16] DEVLIN J, CHANG M W, LEE K, et al. BERT: Pre-training of deep bidirectional transformers for language understanding[EB/OL]. (2019-05-24)[2022-08-15]. https://arxiv.org/abs/1810.04805.

[17] BROWN T B, MANN B, RYDER N, et al. Language models are few-shot learners[C]//Proceedings of the 34th International Conference on Neural Information Processing Systems. New York: ACM, 2020: 1877-1901.

[18] SHOEYBI M, PATWARY M, PURI R, et al. Megatron-LM: Training multi-billion parameter language models using model parallelism[EB/OL]. (2020-03-13)[2022-08-15]. https://arxiv.org/abs/1909.08053.

[19] NARAYANAN D, SHOEYBI M, CASPER J, et al. Efficient large-scale language model training on GPU clusters using megatron-LM[C]//Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. New York: ACM, 2021: 1-15.

[20] YUAN K, LI Q Q, SHAO J, et al. Learning connectivity of neural networks from a topological perspective[C]//Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2020: 737-753.

[21] FORTUNATO S. Community detection in graphs[J]. Physics Reports, 2010, 486(3-5): 75-174.

[22] VASILIAUSKAITE V, EVANS T S. Making communities show respect for order[J]. Applied Network Science, 2020, 5(1): 15.

[23] VASILIAUSKAITE V, EVANS T S, EXPERT P. Cycle analysis of directed acyclic graphs[J]. Physica A: Statistical Mechanics and its Applications, 2022, 596: 127097.

[24] HUANG Y, CHENG Y, BAPNA A, et al. GPipe: Efficient training of giant neural networks using pipeline parallelism[C]//Proceedings of the 33rd Internatioanl Conference on Neural Information Processing Systems. New York: ACM, 2019: 103-112.

[25] GENG J K, LI D, WANG S. ElasticPipe: An efficient and dynamic model-parallel solution to DNN training[C]//Proceedings of the 10th Workshop on Scientific Cloud Computing. New York: ACM, 2019: 5-9.

[26] HARLAP A, NARAYANAN D, PHANISHAYEE A, et al. PipeDream: Fast and efficient pipeline parallel DNN training[EB/OL]. (2018-06-08)[2022-08-15]. https://arxiv.org/abs/1806.03377.

[27] RASLEY J, RAJBHANDARI S, RUWASE O, et al. DeepSpeed: System optimizations enable training deep learning models with over 100 billion parameters[C]//Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2020: 3505-3506.

[28] YANG B, ZHANG J, LI J, et al. PipeMare: Asynchronous pipeline parallel DNN training[EB/OL]. (2020-02-09) [2022-08-15]. https://arxiv.org/abs/1910.05124v2.

[29] NARAYANAN D, HARLAP A, PHANISHAYEE A, et al. PipeDream: Generalized pipeline parallelism for DNN training[C]//Proceedings of the 27th ACM Symposium on Operating Systems Principles. New York: ACM, 2019: 1-15.

[30] GUAN L, YIN W, LI D, et al. XPipe: Efficient pipeline model parallelism for multi-GPU DNN training[EB/OL]. (2020-11-09)[2022-08-15]. https://arxiv.org/abs/1911.04610.

[31] FAN S Q, RONG Y, MENG C, et al. DAPPLE: A pipelined data parallel approach for training large models[C]//Proceedings of the 26th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming. New York: ACM, 2021: 431-445.

[32] RAJBHANDARI S, RASLEY J, RUWASE O, et al. ZeRO: Memory optimizations toward training trillion parameter models[C]//Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. Piscataway: IEEE Press, 2020: 1-16.

[33] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2015-04-10)[2022-08-15]. https://arxiv.org/abs/1409.1556.

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

[35] ZAREMBA W, SUTSKEVER I, VINYALS O. Recurrent neural network regularization[EB/OL]. (2015-02-19)[2022-08-15]. https://arxiv.org/abs/1409.2329.

[36] 王丽, 郭振华, 曹芳, 等. 面向模型并行训练的模型拆分策略自动生成方法[J]. 计算机工程与科学, 2020, 42(9): 1529-1537.
WANG L, GUO Z H, CAO F, et al. An automatic model splitting strategy generation method for model parallel training[J]. Computer Engineering & Science, 2020, 42(9): 1529-1537(in Chinese).

[37] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 4700-4708.

[38] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB modelsize[EB/OL]. (2016-11-04)[2022-08-15]. https://arxiv.org/abs/1602.07360.

Task segmentation and parallel optimization of DNN model

JU Tao*，LIU Shuai，WANG Zhiqiang，LI Linjuan

(School of Electronic and Information Engineering，Lanzhou Jiaotong University，Lanzhou 730070，China)

Abstract：In order to solve the problems of difficult parallelization, long training time, and low equipment utilization in the traditional parallelization method of manually partitioning computing tasks of the neural network model, a task segmentation and parallel optimization method based on feature perception of deep neural network (DNN) model was proposed.Firstly, in the hardware computing environment, the computing characteristics of the model were dynamically analyzed to obtain the internal correlation and various parameter attributes of the model, and the directed acyclic graph (DAG) of original computing tasks was constructed. Secondly, the topological relationship of DAG nodes that could be partitioned into clusters was constructed by using augmented antichain, and the original DAG was transformed into an antichain DAG for easy partition. Thirdly, the antichain DAG state sequence was generated by topological sorting, and the state sequence was divided into different execution stages by dynamic programming.The optimal segmentation points were analyzed to divide the model and achieve dynamic matching between the model partition and each GPU. Finally, micro-processing in batches was carried out, and multi-iteration intensive training was realized by introducing pipeline parallelization, which improved GPU utilization and reduced training time. The experimental results show that on the CIFAR-10 dataset, compared with the existing model segmentation methods, the proposed model segmentation and parallel optimization method can balance the training task load among GPUs. The 4 GPU speedup reaches 3.4, and the 8 GPU speedup reaches 3.76 while ensuring the model training accuracy.

Keywords：deep neural network model parallelization；model segmentation；pipeline parallelization；antichain；parallel optimization

Received：2022-08-17；**Accepted**：2022-11-07；**Published Online**：2022-11-22 11：41

URL：link.cnki.net/urlid/11.2625.V.20221118.1909.004

Foundation items：National Natural Science Foundation of China (61862037); Tianyou Innovation Team Project of Lanzhou Jiaotong University (TY202002);Science and Technology Project of Gansu Province (23CXGA0028)

* **Corresponding author**. E-mail：jutao@mail.lzjtu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0733

面部三维生理特征点在睡眠剥夺中的疲劳检测

李敬强[1,2,*]，房秋[1,2]，樊天辰[1,2]，马龙[3]

(1. 中国民航大学 安全科学与工程学院，天津 300300；2. 中国民航大学 民航安全科学研究所，天津 300300；
3. 中国民航大学 中欧航空工程师学院，天津 300300)

摘　　要：睡眠不足、生理节律紊乱是导致交通运输岗位人员疲劳的重要因素，随着值勤“刷脸”技术的兴起，及时识别岗位人员疲劳状态是安全风险防范关口前移的必要举措。为深入探索面部疲劳的三维变化，通过30 h睡眠剥夺实验，基于主成分分析(PCA)法融合面部双侧曲率特征，提出皱纹严重程度指数(WSI)，结合主观疲劳评估、精神运动警觉度测试和客观心率(HR)监测方法，综合判断受试者的面部疲劳程度。结果表明：在睡眠剥夺条件下，WSI变化和疲劳的波动特征明显，整体呈逐渐升高趋势；30 h实验数据显示，WSI变化和主观疲劳指数、反应时的波动趋势有较强相关性(P<0.01)和HR的波动趋势有一定的关联性(P<0.05)。基于WSI指数将疲劳划分为4个等级，从三维生理特征的角度验证面部疲劳静态检测的可行性，研究结果为疲劳快速检测技术提供了新思路。

关　键　词：面部特征；疲劳检测；睡眠剥夺；三维建模；昼夜节律
中图分类号：X951
文献标志码：A　　**文章编号**：1001-5965(2024)09-2753-10

随着道路货物运输、民用航空、航空航天等领域全天候服务的属性越来越明显，岗位人员疲劳现象，已经成为重要的安全隐患。疲劳会降低作业人员的警觉性和认知能力，也会对任务决策、判断能力产生较大影响[1-2]。疲劳后个体的面部特征会发生变化，例如眼睑闭合、嘴角张开等局部行为的频繁发生，还包含静态特征诸如“脸色难看”“黑眼圈”“脸色疲惫”[3-4]等。当前，人脸识别技术已经被广泛用于交通运输环境的实时监测之中。靳慧斌等[5]通过受试者工作特征曲线(receiver operating characteristic curve, ROC)分析瞳孔直径推导出管制员的疲劳状态；Lim等[6]提出了一种结合Viola-Jones人脸检测技术、霍夫变换法、直方图均衡法、Canny边缘检测算法和眼睑闭合程度百分比(percentage of eyelid closure over the pupil over time, PERCLOS)的算法，能够准确检测出驾驶员的眼睛状况。上述技术通常需要连续不断地监测个体眼动运动轨迹或捕捉面部关键部位动作，对个体头部位置、角度和运动幅度有较高要求，该局限性制约了提取疲劳特征点的精确度。随着面部识别在签到、签出环节的普及，如果能够快速而准确地识别岗位人员的疲劳并同时采取相应的预警措施将会是控制疲劳风险的重要辅助技术手段。然而，目前尚未有利用个体静态面部疲劳特征变化来识别疲劳状态的系统性研究。

皱纹作为面部疲劳状态的表现是可以被检测到的[7-9]，其中，泪沟和鼻唇沟是中面部皱纹的突出表现。泪沟是位于内侧眶下缘的自然凹陷，从内眦向内测延伸至2 cm的位置[10]；鼻唇沟则起于鼻翼两侧并延伸到嘴角位置[11]。泪沟和鼻唇沟在面部疲

收稿日期：2022-08-20；**录用日期**：2023-02-03；**网络出版时间**：2023-03-09 15:10
网络出版地址：link.cnki.net/urlid/11.2625.V.20230309.1121.003
基金项目：国家自然科学基金(U1933122)；中央高校基本科研业务费专项资金(3122019186)
*通信作者. E-mail：jqli@cauc.edu.cn

引用格式：李敬强，房秋，樊天辰，等. 面部三维生理特征点在睡眠剥夺中的疲劳检测[J]. 北京航空航天大学学报，2024，50(9)：2753-2762. LI J Q，FANG Q，FAN T C，et al. Fatigue detection of facial 3D physiological feature points in sleep deprivation[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2753-2762 (in Chinese).

劳的过程中逐渐加深，但是其在短时间内由于脂肪和皮肤的张性会因为疲劳而有不同程度的加深和恢复。See 等[12]提出了一种三维彩色散斑立体摄影测量的新方法及其在鼻唇沟体积 (nasolabial fold volume, NLFV) 评估中的应用。Sadick 等[13]定义了泪沟的4个解剖学特征，并描述了一种新的泪沟畸形分类量表 (tear trough rating scale, TTRS)。Arthur 和 Hassan [14]使用 TTRS 确定了泪沟深度和其他特征，评估了泪沟深度、色素沉着、鼻脂肪垫脱垂和下眼睑皱纹。采用传统的二维建模方法较难提取上述区域皱纹的特征点，而且因缺乏对皱纹的纵深维度量化无法对疲劳程度进行准确评估。然而，对于三维视角的疲劳特征点提取及评估还鲜有研究。

基于此，本文通过睡眠剥夺实验诱发受试者产生疲劳，并利用摄像头采集面部疲劳照片，通过建模分析双侧面部特征点曲率和相关指标，与主观疲劳量表评分和精神运动警觉度测试进行比对研究，以验证面部三维生理特征检测疲劳的科学性和有效性。本文构建的面部三维生理特征点能够对睡眠剥夺个体基于静态面部特征的疲劳风险预警技术提供借鉴。

1 实验设计

1.1 被试者的选取

招募中国民航大学学生40名，通过实验遵守程度标准筛选被试者31名，其中飞行学员8名，管制学员9名，男性，年龄为20±1.23岁，身体健康，视力、听力正常，无睡眠不良、失眠等疾病，睡眠类型为中间型[15]。受试者进行实验前24 h内不得饮用刺激性饮品，不得服用影响生理节律的药物。所有的受试者都会在实验开始前详细地了解实验目的、实验工具和具体流程，并在公平公开的原则下签署知情同意书。本实验已经得到中国民航大学安全科学与工程学院伦理委员会的审批同意。

1.2 实验工具

1）主观疲劳量表。主观评估疲劳采用卡罗林斯卡嗜睡量表 (Karolinska sleepiness scale, KSS)[16]，采用1～9级评分来获得人的疲劳等级。1、5、9级分别代表警觉度极高、既不警觉也不困倦和极度困倦无法保持清醒。

2）精神运动警觉度测试仪。采用 E-prime3.0 编制的精神运动警觉度测试 (psychomotor vigilance task, PVT)[17]程序，以白色为底，呈现在屏幕正中心，时长为5 min。受试者会在屏幕面前进行观察，其间，会随机出现一个红色圆点，从其出现到消失共会持续500 ms，红色圆点的直径约为4 mm，当受试者看到红色圆点后，需迅速触碰屏幕以做出反应，此时终端将自动记录受试者从看到红色圆点到反应所用的时间 (reaction time, RT)。

3）心率监测手环。使用 Polar H10 胸带，质量为60 g，厚度为10 mm，心率（heart rate, HR) 采样频率为1 kHz。佩戴注意事项：HR 监测带应与皮肤接触并紧贴，松紧程度为一指左右，佩戴至受试者的胸部正下方部位，佩戴时应标志朝上；胸带的背带需要进行湿润，可用清水或酒精进行均匀擦拭；将胸带缠绕受试者的胸部一周后连接至扣环。

4）面部数据记录仪器。图像采集使用以太网相机、图像采集卡和数码光处理投影仪，该组合可实现对面部图像的精准采集。

1.3 实验流程

实验采取固定作息法[18]，受试者保持从第1天早晨8:00到第2天中午11:00的30 h清醒状态。每位受试者在早晨8:00清醒，8:30到达实验室，前半小时专门用于训练和设备佩戴，正式实验从9:00开始。

实验期间保持恒定的环境温度 (24±1℃) 和光照强度 (100 lux 以下)，为保持固定的热量消耗和摄取，受试者除了去卫生间之外，其余时间需以固定姿势和角度坐在椅子上，每小时进食含有个体每日所必须摄入热量1/24的食物[19-20]。实验过程中禁止任何形式的睡眠和剧烈运动。

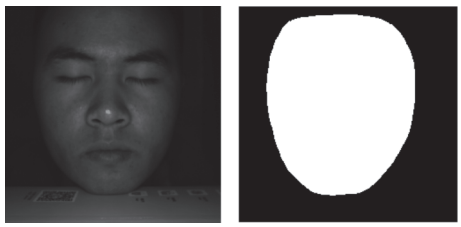
实验过程中持续监测 HR 波动，并且每小时结束后，受试者需填写 KSS 量表进行疲劳自我评价、使用 PVT 的行为能力测试程序检测反应时 RT，并且用面部数据记录仪器进行面部拍照采样。

2 面部特征点提取和处理

2.1 面部三维建模

采用一种基于条纹投影级次识别的三维测量方法[21]。通过主动双目视觉法，实现投影条纹级次的判别，对获取的信息进行旋转轴标定，通过迭代最近点算法进行精度高于0.08 mm的三维测量，将其运用到三维面部建模中，如图1和图2所示。

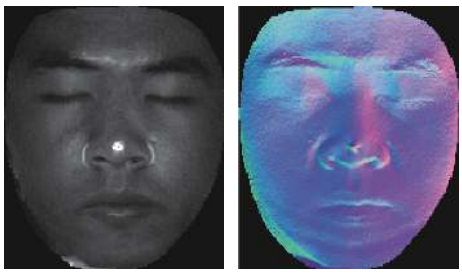
本实验采用 Lab VIEW 进行3D面部建模，摄像机的内外参数组成4个坐标系：(x, y)、(x_c, y_c, z_c)、(x_w, y_w, z_w)、(u, v) 分别为像平面、相机、世界和图像坐标系。实验采用棋盘标定法，点 (u, v) 与点 (x_w, y_w, z_w) 的关系式为



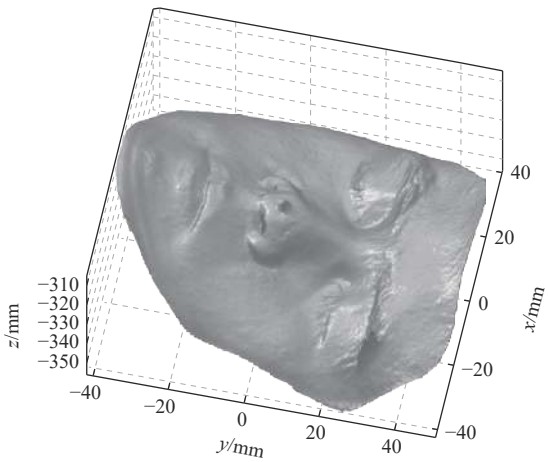
(a) 相机图像　(b) 面部提取区域

图 1　面部轮廓识别

Fig. 1　Facial contour recognition



(a) 灰度图像　(b) 法向量分布图



(c) 面部建模

图 2　面部三维建模

Fig. 2　3D face modeling

$$z_c\begin{bmatrix}u\\v\\1\end{bmatrix}=\boldsymbol{M}_1\boldsymbol{M}_2\begin{bmatrix}x_w\\y_w\\z_w\end{bmatrix}=\begin{bmatrix}\dfrac{f}{\lambda_u} & -\dfrac{f}{\lambda_u}\cot\alpha & u_0\\ 0 & \dfrac{f}{\lambda_v}\cdot\dfrac{1}{\sin\alpha} & v_0\\ 0 & 0 & 1\end{bmatrix}[\boldsymbol{R}\quad\boldsymbol{T}] \tag{1}$$

式中: $\boldsymbol{M}_1$ 为内参数矩阵; $\boldsymbol{M}_2$ 为外参数矩阵; $\boldsymbol{R}$ 为旋转矩阵; $\boldsymbol{T}$ 为平移向量; λ_u 和 λ_v 分别为 u 轴和 v 轴上的尺寸; α 为 u 轴和 v 轴的夹角。

由式(1)可知, 通过采集面部疲劳实时图像, 可以得到内参数矩阵 $\boldsymbol{M}_1$ 及畸变系数矩阵 $\boldsymbol{C}$, 分别为

$$\boldsymbol{M}_1=\begin{bmatrix}2\,213 & 0 & 622.35\\ 0 & 2\,219.64 & 488.691\\ 0 & 0 & 1\end{bmatrix}$$

$$\boldsymbol{C}=\begin{bmatrix}-0.069\,3 & 0.017\,8 & 0.001\,8 & -0.002 & 0\end{bmatrix}$$

首先, 采用相移法进行条纹投影, 相位值为

$$\theta(x,y)=\arctan\left(\frac{\boldsymbol{L}_4(x,y)-\boldsymbol{L}_2(x,y)}{\boldsymbol{L}_1(x,y)-\boldsymbol{L}_3(x,y)}\right) \tag{2}$$

式中: $\boldsymbol{L}_n(x,y)$ 为光强分布, n=1, 2, 3, 4。

其次, 通过基于主动双目识别法, 可以获得投影坐标 (u_x, u_y), 从而获得真实相位数:

$$\vartheta=\theta-2\left[\mathrm{INT}\left(\frac{\theta+\pi}{2\pi}\right)-\mathrm{INT}\left(\frac{u_x-1}{16}\right)\right]\pi \tag{3}$$

式中: INT 为取整函数。

通过迭代最近点算法与旋转轴相结合对双面点云数据进行拼接。设 2 组具有 m 个点的点云数据为点集 A 和 B, 点 $A_i=(x_i, y_i, z_i)$ 与点 $B_i=(X_i, Y_i, Z_i)$ 坐标一一对应, 得到欧氏距离 $d(A_i, B_i)$ 为

$$d(A_i,B_i)=\sqrt{(x_i-X_i)^2+(y_i-Y_i)^2+(z_i-Z_i)^2} \tag{4}$$

最后, 通过对旋转矩阵 $\boldsymbol{R}$ 和平移向量 $\boldsymbol{T}$ 多次迭代, 使最优解 ξ 取值最小:

$$\xi=\sum_{i=1}^{m}\left|\left(\boldsymbol{R}\overrightarrow{OA_i}+\boldsymbol{T}\right)-\overrightarrow{OB_i}\right|^2 \tag{5}$$

通过点云拼接连通区域对双面轮廓线进行链接, 从而完成面部建模。

2.2　面部特征点定位

通过观察睡眠剥夺前后的面部生理特征发现, 面部双侧皱纹存在一定程度的变化, 如图 3 和图 4 所示。为准确提取面部三维生理特征点的变化指标, 拟定每个皱纹区域采用 10 个定位点, 如图 5 所示。计算三维平均曲率 (mean curvature, MC)[22-23], 通过 MC 反映的面部皱纹弯曲度可以很好地比较皱纹的加深程度。

2.3　面部特征点曲率计算

设面部特征点的坐标为 $\delta(x_i, y_i, z_i)$, i=0, 1, ⋯, n, 按 x 坐标从小到大排列。

引入相邻点连线向量:

$$\boldsymbol{\mu}_i=\overrightarrow{\delta_{i-1}\delta_i}=\left[x_i-x_{i-1},y_i-y_{i-1},z_i-z_{i-1}\right] \tag{6}$$

其中: i=0, 1, ⋯, n。

3 个相邻观测点之间曲线的平均弯曲度为

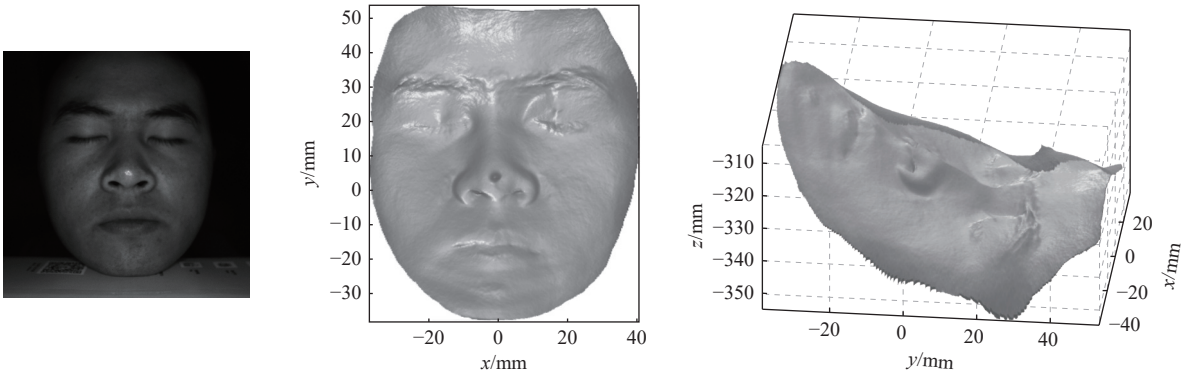
$$\omega_i=\arccos\frac{\boldsymbol{\mu}_i\cdot\boldsymbol{\mu}_{i+1}}{|\boldsymbol{\mu}_i|\cdot|\boldsymbol{\mu}_{i+1}|}(|\boldsymbol{\mu}_i|+|\boldsymbol{\mu}_{i+1}|)^{-1} \tag{7}$$

其中: 等式右侧分母为折线的长度, 分子为折线的转角。

从点 δ_0 到点 δ_n 整条曲线的全线平均弯曲度为

$$\omega_p=\left(\sum_{i=1}^{n-1}\arccos\frac{\boldsymbol{\mu}_i\cdot\boldsymbol{\mu}_{i+1}}{|\boldsymbol{\mu}_i|\cdot|\boldsymbol{\mu}_{i+1}|}\right)\left(\sum_{i=1}^{n}|\boldsymbol{\mu}_i|\right)^{-1} \tag{8}$$

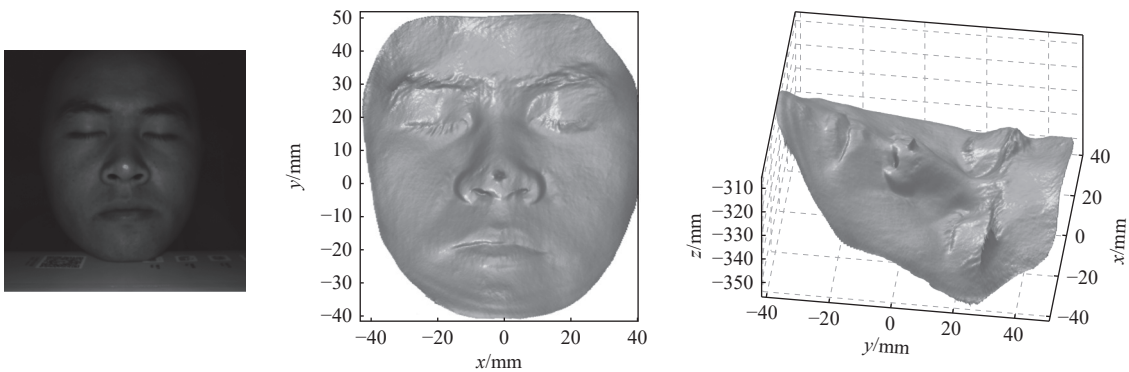
其中: 等式右侧分母为折线的总长度, 分子为折线转角的累计值。



(a) 相机图像 (b) 面部建模正面 (c) 面部建模侧面

图 3 n=1 时睡眠剥夺前的面部建模

Fig. 3 Facial modeling before sleep deprivation with n=1



(a) 相机图像 (b) 面部建模正面 (c) 面部建模侧面

图 4 n=30 时睡眠剥夺后的面部建模

Fig. 4 Facial modeling after sleep deprivation with n=30

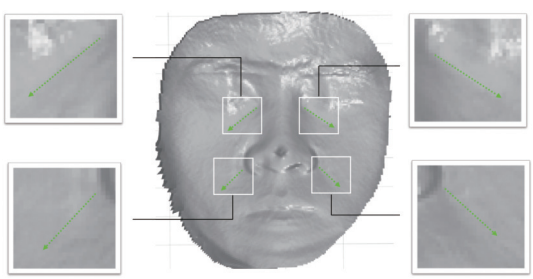


图 5 面部特征点定位

Fig. 5 Facial feature point localization

将式(8)应用于表 1 数据，得出皱纹的平均弯曲度为 ω_p=0.103。

表 1 面部特征点坐标

Table 1 Facial feature point coordinates mm

序号	x	y	z	序号	x	y	z
1	−19.97	28.43	−338.2	6	−22.77	25.61	−338
2	−20.53	27.87	−338.2	7	−23.33	25.04	−338
3	−21.09	27.3	−338.2	8	−23.9	24.49	−338.1
4	−21.65	26.73	−338.1	9	−24.46	23.93	−338.1
5	−22.21	26.17	−338	10	−25.03	23.37	−338.2

2.4 面部特征点数据融合

实验前通过拉伊达准则对数据剔除异常值，得出左右泪沟、鼻唇沟双侧面部曲率的平均值，如图 6 所示。

由于不同个体有不同的面部生理特征，为解决因个体差异而导致的疲劳检测不精确的缺陷，本文采用了主成分分析 (principal component analysis, PCA) 法[24]。PCA 法是一种统计学方法，其能将数据进行简化处理后，通过线性变换，提取出能概括所有信息的综合指标。每个维度都在不同程度上反映了整体指标的某些信息。本文将提取到的左泪沟、右泪沟、左鼻唇沟及右鼻唇沟 4 个特征值，采用 PCA 法进行简化处理。

1）建立相关矩阵 $\boldsymbol{E}=\left[e_{ij}\right]_{m\times m}$，$e_{ij}$的表达式为

$$e_{ij}=\frac{\sum_{k=1}^{m}\left(x_{ki}-\overline{x_i}\right)\left(x_{kj}-\overline{x_j}\right)}{\sqrt{\sum_{k=1}^{m}\left(x_{ki}-\overline{x_i}\right)^2\sum_{k-1}^{m}\left(x_{kj}-\overline{x_j}\right)^2}} \tag{9}$$

式中：e_{ij} 为初始变量 x_i、x_j 的相关系数，i,j=1, 2, ⋯, m。

2）将相关矩阵 $\boldsymbol{E}$ 的特征向量 $\boldsymbol{\varphi}_i$ 按照特征值 ρ_i 从大到小依次排序，组成特征向量矩阵 $\boldsymbol{\varphi}_j$=[φ_{1j}, φ_{2j}, ⋯, φ_{nj}]$^{\mathrm{T}}$，j=1, 2, ⋯, m，得出 m 个主成分指标变量 ρ_j：

$$\rho_j=\varphi_{1j}x_1+\varphi_{2j}x_2+\cdots+\varphi_{nj}x_n \tag{10}$$

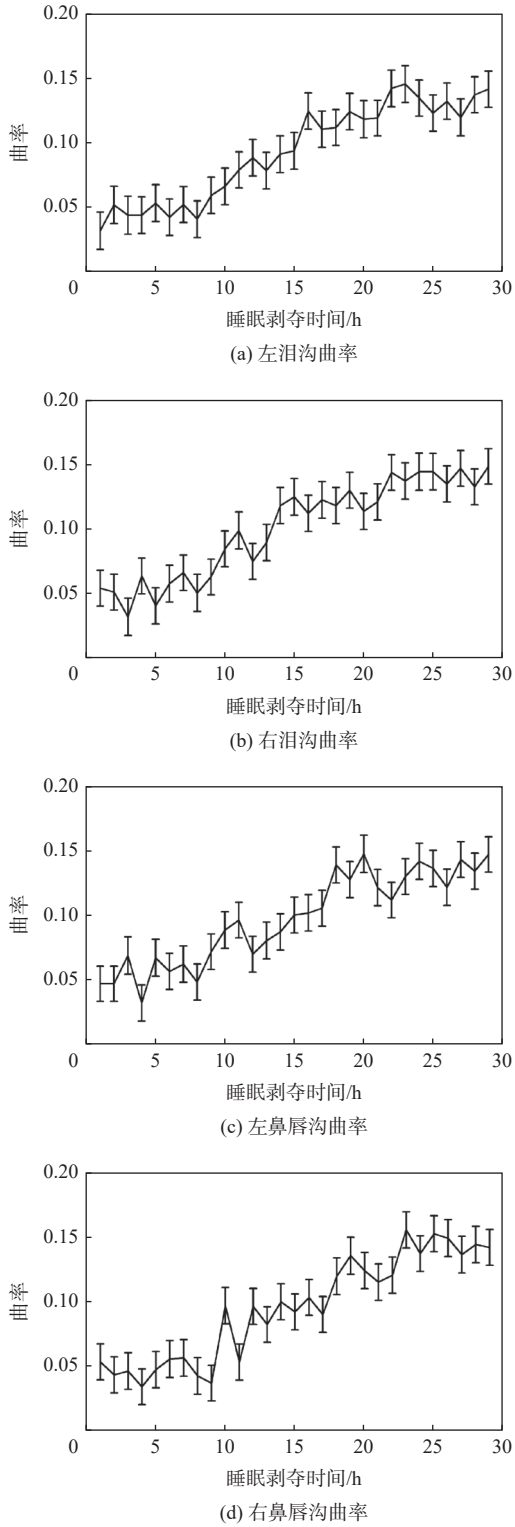


图 6　双侧面部曲率统计

Fig. 6　Curvature in bilateral facial statistics

3) 计算综合评价指标, 主成分贡献率 σ_j 及累计方差贡献率 η:

$$\begin{cases}\sigma_j=\rho_j\Big/\sum\limits_{k=1}^{m}\rho_k\\ \eta=\sum\limits_{k=1}^{p}\rho_k\Big/\sum\limits_{k=1}^{m}\rho_k\end{cases}\tag{11}$$

4) 取 η>85%,前 p 个指标 $x_1, x_2, \cdots, x_p$ 组成矩阵 $\boldsymbol{F}$ 计算主成分矩阵 $\boldsymbol{\kappa}$:

$$\boldsymbol{\kappa}=\sum_{j=1}^{p}\boldsymbol{\sigma}_j\boldsymbol{F}\tag{12}$$

通过 PCA 降维方法, 计算出各指标的公因子形成新特征, 各新特征值均大于 90%, 这减少了降维过程中的信息缺失, 保证了后续分析结果的可靠性。成分 1 的贡献率大于 90%, 故提取成分 1 进行后续降维分析, 如表 2 所示。

表 2　PCA 的总方差解释

Table 2　Total variance interpretation of PCA

新特征	成分	总计		方差百分比/%	
		初始特征值	提取载荷平方和	初始特征值	提取载荷平方和
0.979	1	3.739	3.739	93.485	93.485
0.965	2	0.118		2.95	
0.956	3	0.092		2.297	
0.967	4	0.051		1.269	

将通过 PCA 降维方法从左泪沟曲率 L_{TTC}、右泪沟曲率 R_{TTC}、左鼻唇沟曲率 L_{NLFC}、右鼻唇沟曲率 R_{NLFC} 提取出的综合指标定义为皱纹严重程度 (wrinkle severity index, WSI), 其表达式为

$$F_{\mathrm{WSI}}=0.506\,3L_{\mathrm{TTC}}+0.499\,1R_{\mathrm{TTC}}+0.494\,4L_{\mathrm{NLFC}}+0.500\,1R_{\mathrm{NLFC}}\tag{13}$$

3　疲劳指标分析

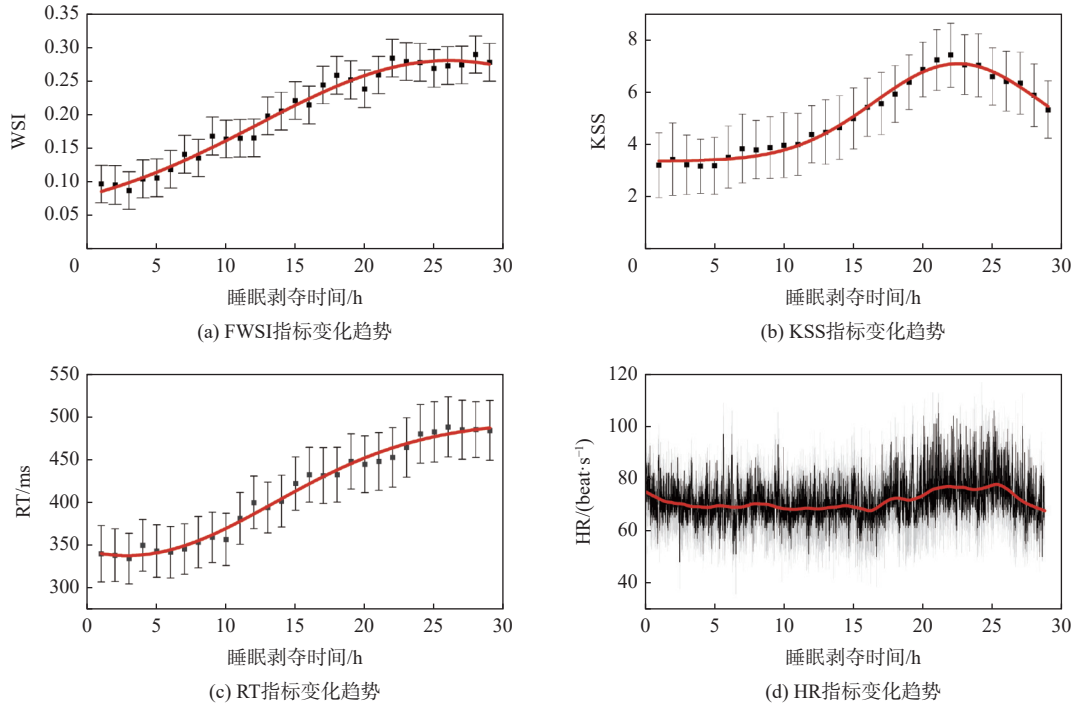
本文对 WSI、KSS 得分、RT 进行高斯拟合, 使用局部加权回归平滑 (locally weighted scatter-plot smoothing, LOWESS) 法对 HR 进行平滑处理, 如图 7 所示。通过线性回归分析, 得到了 WSI、KSS、RT 和 HR 的线性回归系数 β; 使用斯皮尔曼秩相关, 分析 WSI、KSS、RT 之间的相关性, 采用交叉相关函数 (cross-correlation function, CCF), 计算出指标与 HR 之间的交叉相关性; 将统计检验的显著性水平均设定为 P<0.05。

3.1　面部皱纹指数

面部皱纹指数 WSI 在早晨处于较低水平, 5 h 后中午 1:00 左右开始缓慢波动上升, 直至晚上 8:00 左右开始加速升高, 直至第 2 天凌晨 6:00 达到最大值后趋于稳定。WSI 整体处于 (0.08, 0.29) 之间, 平均皱纹指数 $\overline{X}_{\mathrm{WSI}}$ 为 0.2, 变化范围 σ_{WSI} 为 0.07。对 WSI 进行高斯拟合, 如图 7 所示, 拟合程度 R^2=0.98, P<0.05。

3.2　主观嗜睡得分

主观 KSS 得分在早晨同样处于较低水平, 11 h 后晚上 7:00 左右开始迅速增加, 直至次日凌晨



(a) FWSI指标变化趋势 (b) KSS指标变化趋势

(c) RT指标变化趋势 (d) HR指标变化趋势

图 7 4种类型疲劳指标随睡眠剥夺变化趋势

Fig. 7 Variation trend of four types of fatigue indexes with sleep deprivation

6:00达到最大值后开始缓慢恢复。KSS值在(3.16, 7.43)之间波动，平均KSS得分$\overline{X}_{KSS}$=5.08，变化范围σ_{KSS}为1.44。对KSS进行高斯拟合，如图7所示，拟合程度R^2=0.985，P<0.05。

3.3 精神运动警觉度测试反应时

PVT测试在早晨反应迅速，10 h后傍晚6:00左右反应时RT开始快速波动增加，早晨3:00～6:00维持在一个稳定的状态，随后上午6:00～10:00反应时又开始缓慢增加直至达到最高点后趋于稳定。RT整体范围为(333.52, 487.31) ms之间，平均反应时$\overline{X}_{RT}$为410.13 ms，变化幅度σ_{RT}为55.38 ms。对RT进行高斯拟合，如图7所示，拟合程度R^2=0.98，P<0.05。

3.4 心率节律

如图7所示，对HR进行LOWESS，可以看出睡眠剥夺开始时HR处在一个相对较高的水平，之后前3 h左右缓慢下降并在前20 h，即第2天凌晨4:00前维持在一个较为稳定的水平，之后开始波动上升，直至上午10:00左右升至最高值，最后4 h回落至与睡眠剥夺中期相当的水平。HR整体处于(67.63, 77.69) beat/s之间，平均心率$\overline{X}_{RT}$=71.33 beat/s，变化幅度σ_{HR}为3.03 beat/s。

随着睡眠剥夺时间的延长，WSI、KSS、RT和HR随睡眠剥夺实验过程的变化如表3所示。WSI与KSS得分值、PVT的RT、HR的显著性均小于0.01，离散程度小，且显著相关，具有统计学意义。计算分析表明，4种疲劳指标的线性回归均大于0，且相关分析P<0.01，说明WSI、KSS、RT、HR均随睡眠剥夺时间的延长而呈上升趋势。WSI与主观疲劳KSS分值、PVT的相关性为均大于0.9，与HR的相关性大于0.4，这表明面部疲劳综合指标与各主观、客观疲劳指标相关性良好。

表 3 疲劳指标的统计分析

Table 3 Statistical analysis of fatigue index

指标	β	相关性			
		WSI	KSS	RT	HR
WSI	0.971	1	0.904**	0.961**	0.407*
KSS	0.887	0.904**	1	0.875**	0.527**
RT	0.982	0.961**	0.875**	1	0.404*
HR	0.489	0.407*	0.527**	0.404*	1

注：**表示P<0.01，*表示P<0.05。

PVT是快速检测持续性认知注意力的最直接指标之一[25]，本文实验中，30 h不间断的睡眠剥夺和PVT反应时测试能在以往的研究程度上更进一步造成了个体的生理疲劳并影响了其行为和生理特征。研究显示，主观疲劳程度KSS和RT指标与WSI的变化趋势呈现出明显的相关性，这表明嗜睡和警觉性都受到睡眠压力和昼夜节律的影响[26]，且发现RT和HR与其他指标之间有1～2 h的相位延迟，造成上述现象的原因可能是由于睡眠剥夺过程中的警觉性缺失，涉及感觉运动网络内部的连接、皮层下网络与皮层网络之间连接的多个脑网络的交互作用[27]；KSS在后6 h显现出微弱的恢复过程，

与主观嗜睡性的功能连接仅与皮层下丘脑和体感觉皮层之间的通信有关, 这与现有的研究结果一致[28]。

中面部皱纹是由于皮肤、脂肪室和肌肉层的多重作用产生的[29], 在面部特征融合指标 WSI 的变化趋势中发现, 前 5 h 变化较缓慢, 这可能是由于颧骨内侧深层脂肪室和肌肉[30-31]的阶段稳定性使泪沟和鼻唇沟的体积保持一个相对稳定的状态; 晚上 8:00 到凌晨 6:00 时 WSI 进入一个较白天更高的水平且快速上升, 主观疲劳水平和警觉水平也都有显著增加, 有研究表明, 褪黑激素的分泌与人的睡眠意愿有关, 其促使个体的主观嗜睡感加深, 褪黑素通常在傍晚开始分泌, 在凌晨 3:00～4:00 达到较高水平, 在调节睡眠-觉醒周期中发挥着重要作用[32-33]。警觉性水平和面部生理特征的改变可能是褪黑激素的分泌和睡眠的缺失导致。

同时实验发现, 第 2 天凌晨 2:00～6:00 时昼夜节律处于低谷期, WSI 与主观疲劳值、警觉度和 HR 都有较大程度的增加, 与褪黑素的分泌周期相吻合, 这可能是大脑前额受到睡眠剥夺的累积性损伤导致的, 这种损伤被定义为“睡眠债”[34-35]。睡眠剥夺最后 7 h 的 WSI 值和警觉度趋于稳定, 主观疲劳水平有缓慢恢复的趋势, 这可能是脑功能的生理补偿和昼夜节律的恢复导致的[36-37]。研究发现, 睡眠压力和昼夜节律对疲劳程度与面部生理特征的影响均存在显著的交互作用, 这证实了通过面部生理特征对疲劳检测的合理性, 并深入解析了睡眠压力和昼夜节律对疲劳指标的调节作用。

4　面部疲劳指标分级

近年来, 对于面部疲劳现象的检测主要集中在捕捉面部动作轨迹的动态变化上, 尽管面部皱纹的发生率广且观察较为直观, 但由于其需要精确地识别面部三维特征的细微变化, 故少有学者关注个体疲劳程度对此方面的影响, 为此, 本文建立面部疲劳融合指标 WSI, 与 4 种主客观疲劳指标进行比较分析, 探究由睡眠剥夺引起的疲劳程度对面部三维生理指标的影响。

为进一步细化疲劳对 WSI 的影响, 对 WSI 进行了高斯多峰拟合, 拟合得到峰值在 9, 18, 22 h, 将其作为分隔 WSI 变化的临界睡眠剥夺时间节点, 如图 8 所示。WSI 的高斯拟合函数 $F(t)$ 为

$$F(t)=0.04+\frac{8.248\,8}{32.261\,4\sqrt{\pi/(4\ln 2)}}\mathrm{e}^{\frac{-4\ln 2(t-26.040\,6)^2}{32.261\,4^2}} \quad (14)$$

式中: t 为睡眠剥夺时间。拟合程度 R^2=0.980 2, 拟合的方程回归性良好。

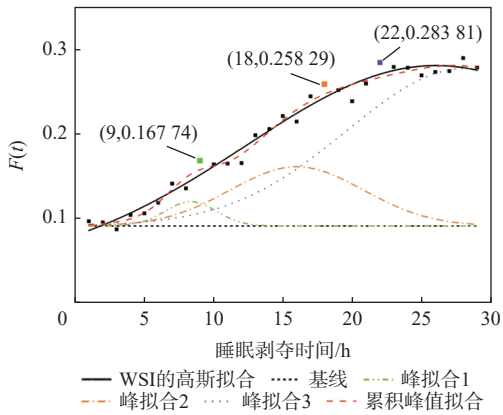


图 8　WSI 的高斯多峰拟合

Fig. 8　Gaussian multi-peak fitting of the WSI

WSI 在睡眠剥夺期间的总体变化可以分为 4 个阶段: 前 9 h 处于缓慢上升阶段, 之后的 9 h 快速上升, 中间的 4 h 上升趋势有所减弱, 最后 7 h 稳定在一个较高的水平, 总体指数随着睡眠剥夺时间的延长波动上升, 如图 9 所示, **表示 P<0.01, *表示 P<0.05。各阶段的 WSI 值均与 KSS 有良好的相关性, 第 1、第 2 和第 4 阶段 WSI 与 RT 相关性较好, 以上数据的对比表明, 面部融合指标 WSI 在主观和客观疲劳检测上具有一定的有效性和科学性。故为更准确地划分面部生理特征疲劳等级, 将随 WSI 的变化将疲劳过程划分为 4 个等级, 通过 4 个面部疲劳等级的划分能够对面部疲劳程度有更精准的判断, 如表 4 所示。

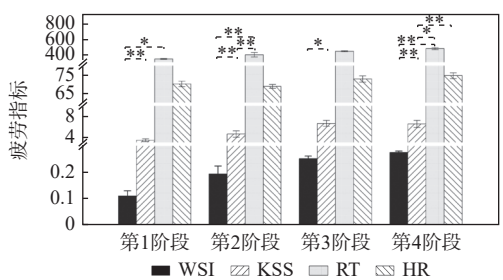


图 9　4 阶段各疲劳指标的变化

Fig. 9　Changes of fatigue indexes in four stages

表 4　基于 WSI 的疲劳等级划分

Table 4　Fatigue grade classification based on WSI

疲劳等级	WSI	等级描述	具体表现
1	0～0.15	精力充沛	注意程度高, 精神饱满, 面部皱纹肉眼不可见
2	0.15～0.25	轻度疲劳	注意程度一般, 稍觉疲倦, 面部皱纹较浅
3	0.25～0.28	中度疲劳	注意程度较低, 非常疲倦, 面部皱纹加深
4	0.28以上	重度疲劳	注意程度极低, 非常疲倦想睡觉, 面部皱纹很深

5 结 论

1）在睡眠剥夺30 h过程中，通过基于条纹投影级次识别的三维测量方法进行面部三维建模，定位面部三维特征点计算平均曲率分析面部双侧皱纹加深程度，利用PCA法降维融合提出面部生理三维特征点的疲劳融合指标WSI，并结合KSS量表、PVT反应时和HR变化，通过相关性分析发现WSI与3个指标均显著相关，具有很好的一致性。

2）将WSI代表的面部三维生理疲劳程度分成4个等级，能够更直观地区分个体的面部疲劳状态，为睡眠剥夺后面部生理疲劳特征的研究提供了新思路，也为全天候交通服务类岗位人员疲劳监测技术开发提供了参考。

本文实验受试者群体特征年龄差异较小，该局限性使得本实验结论尚不能推广至其他年龄群体。今后的研究会进一步丰富其他人口学变量，进一步揭示面部疲劳的三维变化特征并验证检测技术的准确性。

参考文献（References）

[1] SKALA F, ZEMKOVÁ E. Effects of acute fatigue on cognitive performance in team sport players: Does it change the way they perform? A scoping review[J]. Applied Sciences, 2022, 12(3): 1736.

[2] VAN DEN BERG M J, WU L J, GANDER P H. Subjective measurements of In-flight sleep, circadian variation, and their relationship with fatigue[J]. Aerospace Medicine and Human Performance, 2016, 87(10): 869-875.

[3] SUNDELIN T, LEKANDER M, KECKLUND G, et al. Cues of fatigue: Effects of sleep deprivation on facial appearance[J]. Sleep, 2013, 36(9): 1355-1360.

[4] HOLDING B C, SUNDELIN T, CAIRNS P, et al. The effect of sleep deprivation on objective and subjective measures of facial appearance[J]. Journal of Sleep Research, 2019, 28(6): e12860.

[5] 靳慧斌, 于桂花, 刘海波. 瞳孔直径检测管制疲劳的有效性分析[J]. 北京航空航天大学学报, 2018, 44(7): 1402-1407.
JIN H B, YU G H, LIU H B. Effectiveness analysis of pupil diameter detection for air traffic controller's fatigue[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(7): 1402-1407 (in Chinese).

[6] LIM T J, LEONG H Y, JIA Y P, et al. Eye fatigue algorithm for driver drowsiness detection system[M]. Berlin: Springer, 2021: 638-652.

[7] MUHIMMAH I, ABRIYANI F, KURNIAWARDHANI A. Automatic wrinkles detection on face image[J]. IOP Conference Series: Materials Science and Engineering, 2019, 482: 012026.

[8] ELBASHIR R M, HOON YAP M. Evaluation of automatic facial wrinkle detection algorithms[J]. Journal of Imaging, 2020, 6(4): 17.

[9] SABINA U, WHANGBO T K. Edge-based effective active appearance model for real-time wrinkle detection[J]. Skin Research and Technology: Official Journal of International Society for Bioengineering and the Skin and International Society for Digital Imaging of Skin and International Society for Skin Imaging, 2021, 27(3): 444-452.

[10] LIEW S, DE SILVA D. Tear trough and lid–cheek junction anatomy: Current understanding and nonsurgical treatment[J]. Facial Plastic Surgery, 2022, 38(2): 143-151.

[11] MAGNUSSON M R. Commentary on: Three-dimensional description of the angular artery in the nasolabial fold[J]. Aesthetic Surgery Journal, 2021, 41(6): 705-706.

[12] SEE M S, FOXTON M R, MIEDZIANOWSKI-SINCLAIR N A, et al. Stereophotogrammetric measurement of the nasolabial fold in repose: A study of age and posture-related changes[J]. European Journal of Plastic Surgery, 2007, 29(8): 387-393.

[13] SADICK N S, BOSNIAK S L, CANTISANO-ZILKHA M, et al. Definition of the tear trough and the tear trough rating scale[J]. Journal of Cosmetic Dermatology, 2007, 6(4): 218-222.

[14] ARTHUR C A, HASSAN A S. Tear trough depth and quality across age groups: A cross-sectional study[J]. European Journal of Plastic Surgery, 2022, 45(1): 65-72.

[15] 李敬强, 王庆福, 张璐, 等. 不同睡眠类型选择性注意在睡眠剥夺中内稳态与昼夜节律研究[J]. 生物医学工程学杂志, 2022, 39(2): 248-256.
LI J Q, WANG Q F, ZHANG L, et al. Study on homeostasis and circadian rhythm of attention performance of different chronotypes in sleep deprivation[J]. Journal of Biomedical Engineering, 2022, 39(2): 248-256 (in Chinese).

[16] SHAHID A, WILKINSON K, MARCU S, et al. Stanford sleepiness scale (SSS)[M]. New York: Springer New York, 2011: 369-370.

[17] DINGES D F, POWELL J W. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations[J]. Behavior Research Methods, Instruments, & Computers, 1985, 17(6): 652-655.

[18] RODHEIM K, JUNG C, WRIGHT K. 016 associations between circadian melatonin and temperature amplitudes during constant routine[J]. Sleep, 2021, 44(S2): A8.

[19] 汪锡, 王远, 许杨, 等. 非肥胖健康成人静息代谢率预测公式准确性研究[J]. 北京生物医学工程, 2017, 36(6): 607-613.
WANG X, WANG Y, XU Y, et al. Accuracy of predictive equations for resting metabolic rate in non-obese healthy Chinese adults[J]. Beijing Biomedical Engineering, 2017, 36(6): 607-613 (in Chinese).

[20] PUTZ P, KOGLER B, BERSENKOWITSCH I. Reliability and validity of assessing energy and nutrient intake with the Vienna food record: A cross-over randomised study[J]. Nutrition Journal, 2019, 18(1): 1-10.

[21] 马龙, 徐泓悦, 胡艳敏, 等. 基于投影条纹级次识别的发动机叶片三维测量方法[J]. 激光与光电子学进展, 2018, 55(4): 041201.
MA L, XU H Y, HU Y M, et al. Three-dimensional measurement method of aero-engine blades based on projected fringe order identification[J]. Laser & Optoelectronics Progress, 2018, 55(4): 041201 (in Chinese).

[22] CHARYTANOWICZ K, CIEŚLAK W, MOZGAWA W. A generalization of the total mean curvature[J]. Rendiconti Del Seminario

Matematico Della Università Di Padova, 2021, 146: 223-231.

[23] FAIRAG F, CHEN K, BRITO-LOEZA C, et al. A two-level method for image denoising and image deblurring models using mean curvature regularization[J]. International Journal of Computer Mathematics, 2022, 99(4): 693-713.

[24] ZHENG J W, RAKOVSKI C. On the application of principal component analysis to classification problems[J]. Data Science Journal, 2021, 20(1): 26.

[25] ANTLER C A, YAMAZAKI E M, CASALE C E, et al. The 3-minute psychomotor vigilance test demonstrates inadequate convergent validity relative to the 10-minute psychomotor vigilance test across sleep loss and recovery[J]. Frontiers in Neuroscience, 2022, 16(2): 815697.

[26] 王海玮, 温惠英, 刘敏. 夜间环境驾驶员精神负荷的生理特性评估与实验[J]. 吉林大学学报 (工学版), 2017, 47(2): 420-428.
WANG H W, WEN H Y, LIU M. Experimental evaluation of nighttime driver's physiological characteristics in driving simulator[J]. Journal of Jilin University (Engineering and Technology Edition), 2017, 47(2): 420-428 (in Chinese).

[27] PAN L Y, MAI Z F, WANG J H, et al. Altered vigilant maintenance and reorganization of rich-clubs in functional brain networks after total sleep deprivation[J]. Cerebral Cortex, 2023, 33(4): 1140-1154.

[28] THOMAS S J. Sleep deprivation: Time to intervene[J]. Sleep Medicine, 2021, 86(10): 116-117.

[29] ROHRICH R J, AVASHIA Y J, SAVETSKY I L. Prediction of facial aging using the facial fat compartments[J]. Plastic and Reconstructive Surgery, 2021, 147(1S-2): 38-42.

[30] WONG C H, HSIEH M K H, MENDELSON B. The tear trough ligament: Anatomical basis for the tear trough deformity[J]. Plastic and Reconstructive Surgery, 2012, 129(6): 1392-1402.

[31] HUR M S, O J, YANG H M, et al. Heights and spatial relationships of the facial muscles acting on the nasolabial fold by dissection and three-dimensional microcomputed tomography[J]. PLoS One, 2020, 15(8): e0237043.

[32] BANI ISSA W, ABDUL RAHMAN H, ALBLUWI N, et al. Morning and evening salivary melatonin, sleepiness and chronotype: A comparative study of nurses on fixed day and rotating night shifts[J]. Journal of Advanced Nursing, 2020, 76(12): 3372-3384.

[33] KUSZTOR A, RAUD L, JUEL B E, et al. Sleep deprivation differentially affects subcomponents of cognitive control[J]. Sleep, 2019, 42(4): zsz016.

[34] VASEY C, MCBRIDE J, PENTA K. Circadian rhythm dysregulation and restoration: The role of melatonin[J]. Nutrients, 2021, 13(10): 3480.

[35] MAIRE M, REICHERT C F, GABEL V, et al. Human brain patterns underlying vigilant attention: Impact of sleep debt, circadian phase and attentional engagement[J]. Scientific Reports, 2018, 8: 970.

[36] CRAIG A, TRAN Y, WIJESURIYA N, et al. Regional brain wave activity changes associated with fatigue[J]. Psychophysiology, 2012, 49(4): 574-582.

[37] AHMED T, WILSON D. Exploiting circadian memory to hasten recovery from circadian misalignment[J]. Chaos, 2021, 31(7): 073130.

Fatigue detection of facial 3D physiological feature points in sleep deprivation

LI Jingqiang[1, 2, *], FANG Qiu[1, 2], FAN Tianchen[1, 2], MA Long[3]

(1. College of Safety Science and Engineering, Civil Aviation University of China, Tianjin 300300, China;
2. Research Institute of Civil Aviation Safety, Civil Aviation University of China, Tianjin 300300, China;
3. Sino-European Institute of Aviation Engineering, Civil Aviation University of China, Tianjin 300300, China)

Abstract: Sleep deprivation and circadian rhythm disorder are important factors affecting the fatigue of transportation staff. In order to advance the safety risk prevention threshold, it is now vital to identify staff fatigue concurrently with the advent of "face brushing" technology while on duty. This study uses the 30 h sleep deprivation experiment to further explore the three-dimensional changes of facial fatigue. Based on the fusion of facial bilateral curvature characteristics and principal component analysis (PCA), it proposes the wrinkles severity index (WSI), which when paired with the objective heart rate (HR) monitoring method, psychomotor vigilance test, and subjective fatigue assessment, integrated subjects' facial fatigue judgment. The results showed that under the condition of sleep deprivation, the WSI changes and fatigue fluctuated significantly, and the overall trend was gradually increasing. The results of the 30 h trial revealed a high correlation between the change in WSI and the trends in subjective tiredness index, response time (P<0.01), and HR (P<0.05). Based on the WSI index, the fatigue was divided into four grades, and the feasibility of static detection of facial fatigue was verified from the perspective of three-dimensional physiological characteristics. The results provided a new idea for rapid fatigue detection technology.

Keywords: facial features; fatigue detection; sleep deprivation; three-dimensional modeling; circadian rhythm

Received: 2022-08-20; **Accepted**: 2023-02-03; **Published Online**: 2023-03-09 15:10

URL: link.cnki.net/urlid/11.2625.V.20230309.1121.003

Foundation items: National Natural Science Foundation of China (U1933122); The Fundamental Research Funds for the Central Universities (3122019186)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0706

基于图结构和序列特征融合的关系抽取

武同心[1,*]，纪鑫[1,2]，王宏刚[1]，杨智伟[1]，陈屹婷[1]，赵加奎[1]
(1. 国家电网有限公司大数据中心，北京 100052；2. 北京航空航天大学 计算机学院，北京 100191)

摘 要：关系抽取是自然语言处理应用中的一项重要任务。现有的关系抽取方法主要基于语言序列特征或句子结构信息来预测关系，并不能有效地反映实体间关系的内在结构和特征。为此，提出一种融合句子中图结构和序列特征信息的关系抽取模型。该模型利用基于注意力的图卷积网络（GCN）学习语句中的结构信息，利用双向长短期记忆（BiLSTM）网络学习语句的序列语义特征，通过注意力机制结合句子结构特征和序列语义对关系进行分类。在公共数据集和手工构建的数据集上进行了大量实验，验证了所提模型的优越性。

关 键 词：信息抽取；关系抽取；图神经网络；序列模型；特征融合
中图分类号：TP391.1
文献标志码：A **文章编号**：1001-5965(2024)09-2763-09

关系抽取是识别文本中给定实体之间的关系。作为信息提取的子任务，其在知识图构建[1]和问答系统[2]等各种自然语言处理应用中发挥着重要作用。随着大数据在生物医学工程[3]和电力系统[4]等领域的发展，关系抽取在工程领域也得到了极大的关注。现有的关系抽取模型主要分为2类：基于序列和基于结构的关系抽取模型[5]。基于序列的关系抽取模型主要利用单词序列信息。例如，利用卷积神经网络（convolutional neural networks, CNN）[6]来捕获词语之间的相关性进行关系抽取[7]，利用长短期记忆（long short-term memory, LSTM）网络对词语的上下文信息进行编码以预测关系[8]。然而，这些模型无法捕获句子的非局部句法关系，对关系抽取造成影响。另外，基于结构的关系抽取模型利用句子的结构信息来预测关系。例如，利用树型LSTM网络[9]或图神经网络（graph neural network, GNN）[10]编码依存树。这些模型能够学习词语之间的长距离非线性相关性。然而，其不能有效地利用词语的序列上下文语义。由于许多关系是由多个句子表示的，因此，该模型很难确定从某一词语出发应追溯的路径长度。

为综合利用这2种模型的优点。本文提出了一种融合图卷积网络（graph convolution network, GCN）[11]和双向LSTM（BiLSTM）网络[12]的关系抽取模型（简称CaStr模型），并结合序列语义和图结构信息进行关系抽取。首先，为有效捕获词语之间的长距离非线性相关性，使用“软邻接”策略将依存树的原始结构转换为完全连通的边加权图。这些权重作为词语之间相关性的强度，可以通过自注意力方法学习[13]。为对词语之间的相关性进行编码，利用GCN学习词语表示。由于浅层的GCN模型可能无法有效捕获大型图的非局部交互信息，为训练GCN以捕捉更深的相关性，本文采用级联连接机制来组合GCN的多层进行表示。其次，为学习序列语义信息，本文采用BiLSTM学习另一种类型的词语表示，该种词语表示可以捕获句子序列的双向序列语义信息。最后，利用注意机制将这2种类型的特征表示进行融合以用于关系预测。为评

收稿日期：2022-08-11；**录用日期**：2022-10-04；**网络出版时间**：2022-11-07 09:19
网络出版地址：link.cnki.net/urlid/11.2625.V.20221104.1415.002
基金项目：国家电网有限公司大数据中心科技项目 (52999021N005)
*通信作者. E-mail：tongxin-wu@sgcc.com.cn

引用格式：武同心，纪鑫，王宏刚，等. 基于图结构和序列特征融合的关系抽取[J]. 北京航空航天大学学报，2024，50(9)：2763-2771.
WU T X，JI X，WANG H G，et al. Relation extraction based on fusion of graph structure and sequence features[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2763-2771 (in Chinese).

估本文模型的性能，本文在2个公共数据集和一个手动构建的数据集上进行了大量的实验。实验结果均证明了本文模型的优越性。

1 相关工作

本文工作涉及基于序列和基于结构的2种关系抽取模型。

1.1 基于序列的关系抽取模型

基于序列的关系抽取模型有2种常见的神经网络结构，即CNN和LSTM网络。基于CNN的模型采用CNN[6]来挖掘词语之间的相关性进行嵌入学习[7]。在最原始CNN结构的基础上，产生了一系列改进模型[7,14-17]。然而，基于CNN的模型不能有效地处理远程关系，这是因为CNN主要针对局部信息进行编码，所以这些模型的性能受到很大的限制。另外，基于LSTM网络的模型采用LSTM来捕获词语的上下文信息进行嵌入学习[8]。为提高LSTM网络的性能，将一些方法引入注意力机制来学习重要信息[18-19]。由于LSTM网络仅能利用序列中位于目标词语之前的信息，因此，可利用BiLSTM网络用来扩展LSTM网络并捕获双向信息[20]。当前已经有一些方法采用BiLSTM网络对文本中的关系进行分类，但基于LSTM网络的方法不能有效利用词之间的长距离相关性，因而影响了关系抽取的性能。

1.2 基于结构的关系抽取模型

基于结构的关系抽取模型主要是基于GNN[21]，如使用GCN[7]来识别实体之间的关系。GNN是一种能够捕捉数据图型拓扑特征的神经网络。基于GNN的模型[10,22-23]通常使用文本依赖树作为先验图结构，以获得比CNN和LSTM网络更高效的表示。另外，也有一些模型是在树形LSTM网络[9]的基础上进行扩展来利用结构信息，有的模型将句子的依存关系建模为树形LSTM网络结构来进行关系抽取[24]，还有一些模型则采用以路径为中心的依存树修剪策略[25]。

通常，有些关系由多个句子表达[26-27]，因此，需要研究跨句子的关系抽取。这些模型也主要基于图结构特征[28]。其用节点来表示单词，并用边建模内部和句子间依存关系（依存结构、共指等）。这种跨句子的图结构可有效提取出实体对的特征。有一些模型通过添加更多类型的边[29-30]或利用神经网络[31-32]在基于图结构特征模型的基础上进行扩展。例如，文献[30]先构建了不同粒度（句子、提及到的内容、实体）的节点，然后，将其与启发式生成的边连接，最终，通过面向边的模型推断关系[33]。文献[29]将跨句图视为一个由结构化注意模型[34]诱导的潜在变量，并提出了一种细化机制来挖掘整个文档中的多跳信息，但是这种模型依赖关系解析器，可能会存在流水线式累积误差。

2 本文模型

在实践中，句子结构和词语序列在关系抽取中都起着重要作用。因此，本文提出了一种融合这2种特征的关系抽取模型——CaStr模型。

图1为CaStr模型结构，CaStr模型包括3个组件，即结构信息提取器、上下文语义提取器和特征融合模块。结构信息提取器利用GCN对句子的依

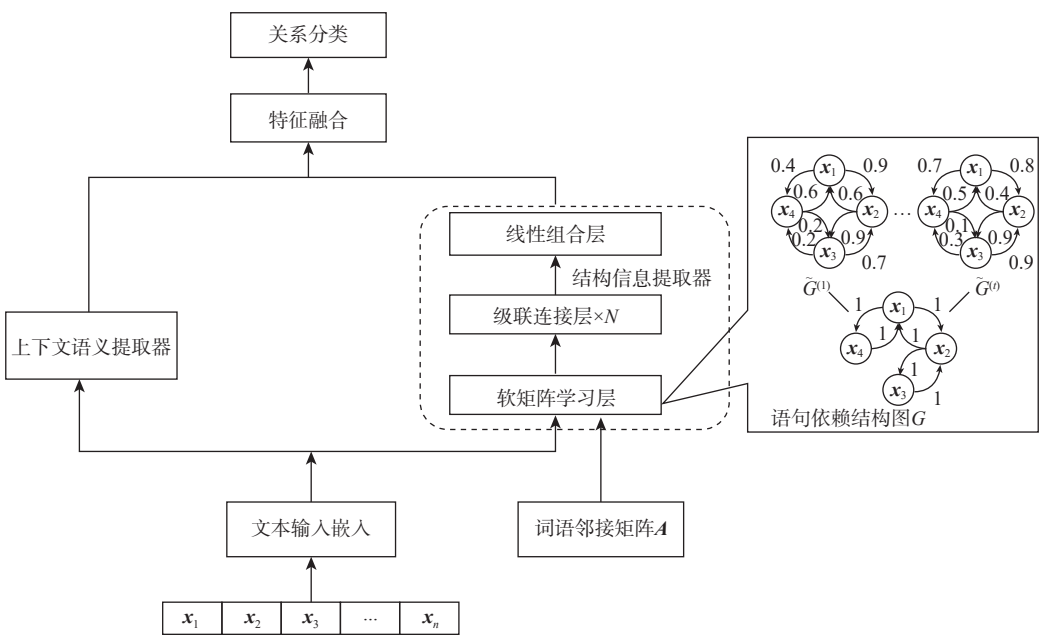


图1 CaStr模型结构

Fig. 1 Structure of CaStr model

存树进行建模, 从句子中学习结构特征。在上下文语义提取器中, BiLSTM 网络用于学习单词的上下文特征。特征融合模块利用注意机制融合这 2 种类型的特征, 融合后的特征可用于关系分类。该模型的优点是可以通过 GCN 提取长距离同步图特征, 也可以通过 BiLSTM 网络提取上下文特征。图 1 中, $\boldsymbol{x}_i$ 为词语向量。

2.1　文本输入嵌入

首先, 输入的文本被嵌入向量表示中。本文使用 GloVe[35] 初始化单词嵌入 $\boldsymbol{e}_{w_i}$。其次, 词性信息 (part-of-speech, POS) 对关系预测也很重要。例如, 在关系预测中, 动词和名词往往比形容词更重要。为利用 POS 信息进行关系抽取, 本文将其映射到随机初始化的向量 $\boldsymbol{e}_{\mathrm{pos}_i}$。

因此, 得到一个词的融合嵌入表示 $\boldsymbol{e}_i$, 其中有 $\boldsymbol{e}_i=[\boldsymbol{e}_{w_i},\boldsymbol{e}_{\mathrm{pos}_i}]$, w_i 为词语索引, $\boldsymbol{e}_{w_i}$ 为词语向量。因此, 给定一个句子 $X=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n\}$, 其最终嵌入表示为 $E=\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$。

2.2　上下文语义提取器

由于每个句子都是一个单词序列, 因此, 单词的语义也依赖于近邻的其他单词, 因此, 还需要结合上下文信息来进行关系抽取。在上下文语义提取器组件中, 本文在输入嵌入上使用 BiLSTM 网络来进一步学习句子中的上下文信息。给定输入嵌入表示 $E=\{\boldsymbol{e}_1,\boldsymbol{e}_2,\cdots,\boldsymbol{e}_n\}$, 使用 BiLSTM 网络提取句子的特征 $H_{\mathrm{b}}=\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_n\}$, 其中, 每个单词向量的隐藏层维数为 d。在 BiLSTM 中, 每个 $\boldsymbol{h}_t$ 由一个 LSTM 单元生成, 其中, $t=1,2,\cdots,n$。生成 $\boldsymbol{h}_t$ 的过程如下:

$$\boldsymbol{g}_t=\tan(\boldsymbol{W}_{gx}\boldsymbol{x}_t+\boldsymbol{W}_{gh}\boldsymbol{h}_{t-1}+\boldsymbol{W}_{gc}\boldsymbol{c}_{t-1}+\boldsymbol{b}_g) \tag{1}$$

$$\boldsymbol{c}_t=\boldsymbol{f}_t\boldsymbol{c}_{t-1}+\boldsymbol{g}_t\boldsymbol{i}_t \tag{2}$$

$$\boldsymbol{h}_t=\boldsymbol{o}_t\tan(\boldsymbol{c}_t) \tag{3}$$

式中: $\boldsymbol{W}_{gx}$、$\boldsymbol{W}_{gh}$、$\boldsymbol{W}_{gc}$ 为需要学习的参数矩阵, $\boldsymbol{b}_g$ 为需要学习的参数向量。

每个 LSTM 单元有 3 个门: 输入门 $\boldsymbol{i}_t$、忘记门 $\boldsymbol{f}_t$ 和输出门 $\boldsymbol{o}_t$。其计算式如下:

$$\boldsymbol{i}_t=\sigma(\boldsymbol{W}_{ix}\boldsymbol{x}_t+\boldsymbol{W}_{ih}\boldsymbol{h}_{t-1}+\boldsymbol{W}_{ic}\boldsymbol{c}_{t-1}+\boldsymbol{b}_i) \tag{4}$$

$$\boldsymbol{f}_t=\sigma(\boldsymbol{W}_{fx}\boldsymbol{x}_t+\boldsymbol{W}_{fh}\boldsymbol{h}_{t-1}+\boldsymbol{W}_{fc}\boldsymbol{c}_{t-1}+\boldsymbol{b}_f) \tag{5}$$

$$\boldsymbol{o}_t=\sigma(\boldsymbol{W}_{ox}\boldsymbol{x}_t+\boldsymbol{W}_{oh}\boldsymbol{h}_{t-1}+\boldsymbol{W}_{oc}\boldsymbol{c}_t+\boldsymbol{b}_o) \tag{6}$$

式中: σ 为激活函数。

由于 BiLSTM 网络可以更好地捕获双向语义依赖的信息, 因此, 本文分别从 BiLSTM 网络中的 LSTM 单元获得 $\boldsymbol{h}_{t_1}$ 和 $\boldsymbol{h}_{t_r}$ 并将其结合生成 $\boldsymbol{h}_t=[\boldsymbol{h}_{t_1},\boldsymbol{h}_{t_r}]$。至此, 通过 BiLSTM 网络层, 生成了每个单词的向量表示 $\boldsymbol{h}_t$。对于整个句子, 则得到 H_{b} 作为句子的表示。

2.3　结构信息提取器

通常依存树表示句子结构的信息, 能够捕捉难以从句子表面得出的非局部和非线性的句法关系。因此, 除了序列信息之外, 本文还学习了用于关系抽取的结构特征。

本文为每个句子构造了一个原始的依存树。然后, 将原始依存树转换为原始邻接矩阵 $\boldsymbol{A}(n\times n)$。在 $\boldsymbol{A}$ 中, 每个元素表示 2 个对应的词语节点之间是否存在连接, 但这种硬表示无法反映 2 个节点之间关系的强度。如图 2 所示, 首先, 本文使用多头注意机制学习多个软矩阵[13,36], 从不同角度反映节点之间的关系; 其次, 通过 GCN 将结构信息编码到表示学习中, 进一步学习单词表示; 最后, 将从不同角度学习到的单词表示组合起来, 以获得统一的向量表示, 这就是结构信息提取器的输出。总的来说, 该组件主要包含 3 层, 即软矩阵学习层、级联连接层和线性组合层。

2.3.1　软矩阵学习层

在这一层中, 本文将原始依赖树转换为完全连

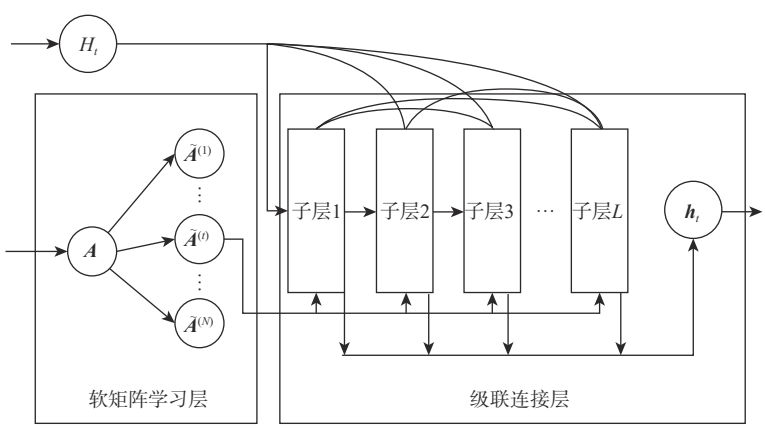


图 2　基于密集连接的级联连接

Fig. 2　Cascaded connections based on dense connections

通的边加权图，其中，软邻接矩阵$\tilde{\boldsymbol{A}}$是基于注意力机制学习的。如图1所示，$\tilde{\boldsymbol{A}}^{(1)}$表示一个全连通图$\tilde{G}^{(t)}$，其是原始图$G$的全连通形式。$\tilde{\boldsymbol{A}}$中的每个元素$\tilde{A}_{ij}$表示从节点$i$到节点$j$的边权重，其大小与原始邻接矩阵$\boldsymbol{A}(n\times n)$元素的大小相同。通过学习软矩阵，可以更有效地建模节点之间的关系，尤其是通过间接多跳路径连接的节点，而这些软关系可以通过模型中的可微函数来捕捉。本文$\tilde{\boldsymbol{A}}$是基于多头注意力方法计算的，该方法允许模型将来自不同角度的信息进行整合。该过程的计算包括一个查询值和一组键值对，输出的是值的加权和，其中，权重由具有相应键的查询函数来计算，$\tilde{\boldsymbol{A}}$的第t个注意力头计算方式如下：

$$\tilde{\boldsymbol{A}}^{(t)}=\mathrm{softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{W}_i^{\boldsymbol{Q}}\times(\boldsymbol{K}\boldsymbol{W}_i^{\boldsymbol{K}})^{\mathrm{T}}}{\sqrt{d}}\right) \tag{7}$$

式中：$\boldsymbol{W}_i^{\boldsymbol{Q}}\in\mathbf{R}^{d\times d}$，$\boldsymbol{W}_i^{\boldsymbol{K}}\in\mathbf{R}^{d\times d}$，$\tilde{\boldsymbol{A}}^{(t)}\in\mathbf{R}^{d\times d}$，$\boldsymbol{Q}$和$\boldsymbol{K}$均等于$\boldsymbol{h}^{l-1}$。$\tilde{\boldsymbol{A}}^{(t)}$为由第$t$个注意力头生成的软邻接矩阵。该层总共构造了$N$个不同的软矩阵$\tilde{\boldsymbol{A}}^{(t)}$，其中，$N$为超参数。图1展示了原始邻接图被转换为对应于软邻接矩阵的多个图的示例。

2.3.2 级联连接层

为更有效地学习到丰富的局部和非局部信息，类似于图2所示的密集连接[37]，引入级联连接层来更好地学习图表示。在密集连接中，节点表示是前几层中节点表示的串联，其中，每个前一层的输出都对当前表示具有同等重要的贡献。然而，在这里的级联连接组件中，较近的层被认为对当前表示起到更重要的作用。

每个级联连接层具有L个子层，其中，L为超参数。这些子层的维度为$d_{\mathrm{hidden}}=L/d$，其中，d为输入特征维度。经过一系列计算，每个子层的结果最终都会被组合起来，生成级联连接层的输出，其维数也为d。为更好地解释后续操作，将此层的输出定义如下：

$$\boldsymbol{g}_j^{(l)}=[\boldsymbol{x}_j,\gamma^{(l-1)}\boldsymbol{h}_j^{(1)},\cdots,\gamma\boldsymbol{h}_j^{(l-1)}] \tag{8}$$

式中：$0<\gamma<1$为罚因子；l为第l个子层；j为第j节点；$\boldsymbol{x}_j$为该级联连接层的原始输入。

由于存在N个不同的邻接矩阵$\tilde{\boldsymbol{A}}$，因此，可以使用N个级联连接层来分别处理这些矩阵。使用GCN在第t个矩阵$\tilde{\boldsymbol{A}}^{(t)}$上的主要运算式如下：

$$\boldsymbol{h}_{t_i}^{(l)}=\rho\left(\sum_{j=1}^{n}\tilde{\boldsymbol{A}}_{ij}^{(t)}\boldsymbol{W}_t^{(l)}\boldsymbol{g}_j^{(l)}+\boldsymbol{b}_t^{(l)}\right) \tag{9}$$

式中：$t=1,2,\cdots,N$，权重矩阵$\boldsymbol{W}$为不同的$\tilde{\boldsymbol{A}}^{(t)}$选择不同的权重矩阵和偏差向量$\boldsymbol{b}$；$i$为第$i$个节点；$l$为第$l$个子层的输出；$\rho$为激活函数(如RELU)。由于每个子层的列维数都会随着子层深度的增加而增加，因此，权重矩阵的列维数也会相应变化，即，$\boldsymbol{W}_t^{(l)}\in\mathbf{R}^{d_{\mathrm{hidden}}\times d^{(l)}}$，其中，$d^{(l)}=d+d_{\mathrm{hidden}}\times(l-1)$。依次计算出每个子层的结果$\boldsymbol{h}_{t_i}^{(l)}$，有$\boldsymbol{h}_{t_i}^{(l)}\in\mathbf{R}^{d_{\mathrm{hidden}}}$。然后，将每个子层输出合并，得到维数为$d$的统一输出$\boldsymbol{h}_{t_i}$，有$\boldsymbol{h}_{t_i}=[\boldsymbol{h}_{t_i}^{(1)},\boldsymbol{h}_{t_i}^{(2)},\cdots,\boldsymbol{h}_{t_i}^{(l)}]$。

2.3.3 线性组合层

该层用于组合从N个不同级联连接层学习到的表示，其中原始输入用多头注意力机制来学习。对应于第i个词的输出定义如下：

$$\boldsymbol{h}_{\mathrm{out}i}=[\boldsymbol{h}_{1i},\boldsymbol{h}_{2i},\cdots,\boldsymbol{h}_{Ni}] \tag{10}$$

$$\boldsymbol{h}_{\mathrm{comb}i}=[\boldsymbol{W}_{\mathrm{comb}i}\boldsymbol{h}_{\mathrm{out}i}+\boldsymbol{b}_{\mathrm{comb}i}] \tag{11}$$

式中：$\boldsymbol{W}_{\mathrm{comb}i}\in\mathbf{R}^{(d\times N)\times d}$，$\boldsymbol{h}_{\mathrm{comb}i}\in\mathbf{R}^{d}$。$\boldsymbol{h}_{\mathrm{comb}i}$为结构信息提取器的输出，该提取器将结构信息编码为单词表示。

2.4 特征融合

由于已经学习到了反映结构信息的特征$H_{\mathrm{g}}=\{\boldsymbol{h}_{\mathrm{comb}1},\boldsymbol{h}_{\mathrm{comb}2},\cdots,\boldsymbol{h}_{\mathrm{comb}n}\}$和反应上下文信息的特征$H_{\mathrm{b}}=\{\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_n\}$，因此，在本节将其融合进行关系预测。一种最直接的方法是直接将其进行线性组合，然而，这种方法不能捕捉到这2类特征之间的潜在相关性。因此，本文采用了一种基于注意力的方法进行来有效的融合，如图3所示。首先，将H_{b}和H_{g}结合起来生成$\boldsymbol{H}=[H_{\mathrm{b}},H_{\mathrm{g}}]$，其中，$\boldsymbol{H}\in\mathbf{R}^{n\times 2d}$。其次，使用$\boldsymbol{H}$作为输入，通过卷积运算生成3个新的节点特征矩阵$\boldsymbol{C}$、$\boldsymbol{D}$和$\boldsymbol{E}$，这3个特征矩阵大小也为$\mathbf{R}^{n\times 2d}$。用于生成位置注意矩阵$\boldsymbol{Q}\in\mathbf{R}^{2d\times 2d}$的元素计算式为

$$q_{ij}=\frac{\exp(D_jC_i)}{\sum_{i=1}^{2d}\exp(D_jC_i)} \tag{12}$$

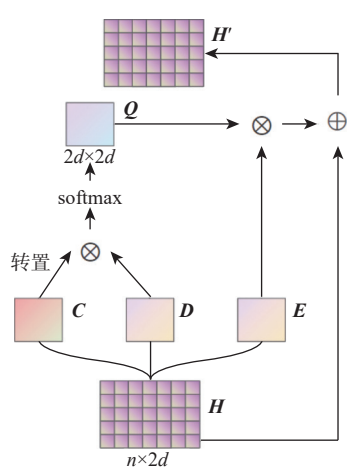


图3 基于注意力机制的融合

Fig. 3 Attention mechanism-based fusion

式中: $q_{ij} \in \boldsymbol{Q}$, 其表示第 j 个位置对第 i 个位置的影响, $i=1,2,\cdots,2d$, $j=1,2,\cdots,2d$。通过注意力学习, 更重要的特征对结果的影响可以增强。最后, 融合后的特征向量计算如下:

$$\boldsymbol{B}'_j = \alpha \sum_{i=1}^{2d} (q_{ij}\boldsymbol{E}_i) + \boldsymbol{B}_j \tag{13}$$

式中: 学习因子 α 为超参数。

2.5　关系分类

与现有工作类似, 本文将关系抽取视为一项分类任务。如 2.4 节所述, 已经学习到了输入文本的融合特征表示 $B=\{\boldsymbol{b}_1,\boldsymbol{b}_2,\cdots,\boldsymbol{b}_n\}$。为对关系进行分类, 首先, 通过池化运算从 B 获得句子表示和 2 个目标实体的表示, 并将这 3 种表示连接起来。其次, 将其送入前馈神经网络(feed forward neural network, FFNN), 以获得最终的特征表示 $\boldsymbol{h}_{\text{final}}$, 该过程计算式如下:

$$\boldsymbol{h}_{\text{sen}} = f(\boldsymbol{h}_{\text{mask0}}) \tag{14}$$

$$\boldsymbol{h}_{\text{obj1}} = f(\boldsymbol{h}_{\text{mask1}}) \tag{15}$$

$$\boldsymbol{h}_{\text{obj2}} = f(\boldsymbol{h}_{\text{mask2}}) \tag{16}$$

$$\boldsymbol{h}_{\text{final}} = \text{FFNN}(\boldsymbol{h}_{\text{sen}}, \boldsymbol{h}_{\text{obj1}}, \boldsymbol{h}_{\text{obj2}}) \tag{17}$$

式中: $\boldsymbol{h}_{\text{mask}}$ 为经掩码处理的综合隐藏层表示。$\boldsymbol{h}_{\text{mask0}}$ 中, 只有句子中不在实体内的词语的表示会被选择。类似地, 对于 $\boldsymbol{h}_{\text{mask1}}$ 和 $\boldsymbol{h}_{\text{mask2}}$, 仅选择组成实体的词语的表示。$f:\mathbf{R}^{2d\times n}\rightarrow\mathbf{R}^{2d\times 1}$ 是将 n 个向量映射到一个句子向量的最大池化函数。最后使用 softmax 将特征 $\boldsymbol{h}_{\text{final}}$ 分类为不同的关系。

3　实验与分析

3.1　数据集

为评估 CaStr 模型的性能, 本文在 3 个数据集上与基线模型进行比较。第 1 个数据集是 Semeval-10[38], 其包含 10 717 个实例, 具有 9 个关系和 1 个特殊的“其他”类。将该数据集用于评估句子级关系抽取的性能, 即从单个句子中抽取关系。为评估跨语句关系抽取的性能, 使用数据集 PubMed[39], 其包含 6 987 个三元关系样本和 6 087 个二元关系样本。大多数样本都包含多个句子, 每个样本都被标记上 5 个关系标签中的 1 个, 即“抵抗或无反应”“敏感”“有反应”“抵抗”和“无”。在该数据集上进行了多类别 n 元关系抽取实验。

为进一步评估本文模型在实践中的性能。本文从工业领域收集数据构建另一个数据集, 即在收集中国国家电网公司 2021 年技术监督大数据分析报告语料的基础上, 构建了一个电力系统关系抽取数据集 CSGC-3。手工选择了 3 480 段文本进行注释, 注释了 3 种在应用中非常有用的关系。在这个数据集中, 35% 的关系是跨句子的。最后, 将数据集进行分割, 以进行训练、验证和测试。该数据集关系的数量如表 1 所示。

表 1　CSGC-3 数据子集中不同关系的数量

Table 1　Number of relations in CSGC-3 subdatasets

子集	从属/个	缺陷/个	原因/个
训练集	1 796	860	930
验证集	143	64	83
测试集	151	62	105

3.2　模型实现细节

在输入文本嵌入层, 使用 GloVe 初始化本文所涉及语料库的单词的嵌入矩阵, 其维数为 300。为获得更多信息作为后续层的输入, 使用 Stanford CoreNLP[40] 处理句子, 以获得分词结果、词性信息和依存树。使用 Adam[41] 优化器, 设置其初始学习率为 10^{-4}, 然后每 20 个迭代后以 0.9 的速率减小。批量大小设置为 50。通过调整参数, 最终选择验证集中 F_1 值最高的模型。通过初步实验, 本文将多头注意机制中的头数设置为 3。与相关工作类似, 使用关系抽取的精确率、召回率和 F_1 值作为评估指标。

3.3　基线模型

为验证本文 CaStr 模型的有效性, 本文与基于序列特征和基于结构的模型进行比较, 包括:

1) CNN[7] 模型。其使用基于句子级注意力机制的 CNN, 可以将其视为基于序列特征的模型。

2) BLSTM[18] 网络模型。其使用 BiLSTM 网络来提取关系。与 LSTM 网络模型相比, 其可以捕获双向语义依赖, 这也是一种基于序列特征的模型。

3) Tree-LSTM[25] 网络模型。其使用了常用于关系抽取的长短期记忆网络, 并且结合了句子和图结构信息。

4) AGGCN(attention guided graph convolutional networks)[10] 模型。其提出了一种用于关系抽取的注意力引导 GCN, 以端到端的方式学习软剪枝策略, 是一种基于结构的模型。

3.4　句子级关系抽取

本文在 CSGC-3 和 Semeval-10[38] 数据集上评估句子级关系抽取的性能。实验结果分别如表 2 和表 3 所示。如表 2 所示, 本文 CaStr 模型获得了最高的精确率、召回率和 F_1 分数, 这表明 CaStr 模型中可以有效地从复杂文本中提取关系。如表 3 所示, CaStr 模型在基准数据集上的性能也优于其他

表2 在Semeval-10[38]数据集上的句级关系抽取性能表现

Table 2 Sentence-level relation extraction performance on Semeval-10[38] dataset %

模型	精确率	召回率	F_1
CNN[7]	63.63	57.76	60.55
BiLSTM[18]	66.01	59.21	62.40
Tree-LSTM[25]	67.73	60.03	63.65
AGGCN[10]	69.92	60.91	65.11
CaStr	71.03	61.76	66.07

表3 在CSGC-3数据集上的句级关系抽取性能表现

Table 3 Sentence-level relation extraction performance on CSGC-3 dataset %

模型	精确率	召回率	F_1
CNN[7]	94.13	87.62	90.75
BiLSTM[18]	94.48	88.17	91.21
Tree-LSTM[25]	96.74	89.64	93.05
AGGCN[10]	97.14	89.84	93.35
CaStr	98.22	90.27	94.08

模型。从结果中可以得出以下3个结论：

1）基于序列特征的模型倾向于将关系标签预测为高频率标签，使其具有相对较高的精确率。因为这些模型主要利用的是上下文信息，而这对于学习词语之间的长距离相关性是无效的。

2）基于结构的模型（如AGGCN[10]）比基于序列特征的模型具有更好的性能。这主要是因为基于结构的模型使用图形神经网络来捕获词语之间更复杂的相关性，从而提高特征表示在文本中反映关系的能力。

3）CaStr模型优于基于序列特征和基于结构的模型。由于CaStr模型使用序列模块提取上下文语义特征，使用GNN模块提取结构特征，因此，CaStr模型是综合利用词语之间的局部和全局相关性进行的关系抽取。

3.5 跨句关系抽取

在本文CSGC-3和PubMed[39]数据集上评估了跨句关系抽取的性能，实验结果分别如表4和表5所示。从这些表中可以看出，CaStr模型在精确率、召回率和F_1得分最高，这表明CaStr模型对跨句关系也是有效的。CaStr模型优于基线模型的几个原因如下：首先，基于序列特征的模型很难有效地扩展成能学习跨句子特征的模型。其次，基于结构的模型可以学习基于图神经网络的句子间相关性，从而提升跨句子关系抽取的性能。最后，通过结合序列特征中的上下文信息，基于结构的模型能更有效地学习不同句子中实体之间的语义相关性。因此，CaStr模型进一步提升了跨句关系抽取的性能。

表4 在PubMed[39]数据集上的跨句关系抽取

Table 4 Cross-sentence relation extraction on PubMed[39] dataset %

模型	精确率	召回率	F_1
CNN[7]	81.85	76.79	79.24
BiLSTM[18]	84.78	78.74	81.65
Tree-LSTM[25]	86.87	80.34	83.48
AGGCN[10]	88.56	81.67	84.98
CaStr	90.23	83.17	86.56

表5 在CSGC-3数据集上的跨句关系抽取性能

Table 5 Cross-sentence relation extraction performance on CSGC-3 dataset %

模型	精确率	召回率	F_1
CNN[7]	68.78	63.56	66.06
BLSTM[18]	70.08	65.27	67.59
Tree-LSTM[25]	73.75	68.34	70.94
AGGCN[10]	77.91	72.11	74.90
CaStr	79.83	73.22	76.38

3.6 消融实验

截止目前，CaStr模型已被证明具有良好的表现。由于这一部分的实验仅用于确认每个模块对整个模型的贡献，因此，还需要再建立2种模型配置。第1种配置是舍弃BiLSTM网络层，将此模型标记为CaStr-NO-BiLSTM；第2种配置是舍弃位置注意层，这使得模型直接结合上下文语义特征和结构特征，将此模型标记为CaStr-NO-ATT。本文测试了这些配置的句子级关系抽取的性能，表6和表7在2个数据集上展示了上述2种配置的实验结果。从表中可以看出这2种模型在2个数据集上的表现都比Caster模型差。该消融实验表明，BiLSTM网络和位置注意力机制都有助于该模型效果的提升。该实验还表明，局部上下文信息对提高表示能力以反映文本中的关系也很重要。在

表6 在Semeval-10[38]数据集上的消融实验

Table 6 Ablation experiments on Semeval-10[38] dataset %

模型	精确率	召回率	F_1
CaStr-NO-BiLSTM	65.35	56.78	60.76
CaStr-NO-ATT	68.41	59.01	63.36
CaStr	71.03	61.76	66.07

表7 在CSGC-3数据集上的消融实验

Table 7 Ablation experiments on CSGC-3 dataset %

模型	精确率	召回率	F_1
CaStr-NO-BiLSTM	93.12	84.91	88.82
CaStr-NO-ATT	96.03	86.85	91.21
CaStr	98.22	90.27	94.08

CaStr 模型中, 直接组合方法不能有效捕获由 2 个提取器提取的特征之间的相关性。

3.7　案例研究

表 8 选取了 2 个关系三元组示例, 分别展示了 BiLSTM[18] 网络模型、AGGCN[10] 模型及 CaStr 模型在 CSGC-3 数据集上抽取三元组的例子, 其中三元组中的第 3 项表示前两项之间的关系。对于左边的第 1 个例子, CaStr 模型和 BiLSTM[18] 网络模型都可以准确识别“主变压器”和“控制箱”的从属关系, 但 BiLSTM[18] 网络模型却错误地将“温差”的主语判断为“主变压器”。这可能是因为“主变压器”和“温差”之间的序列距离比“控制箱”和“温差”之间的距离更近。对于右边的第 2 个例子, AGGCN[10] 模型错误地颠倒了原因关系的主语和宾语, 这意味着 AGGCN[10] 模型无法学习到“松动”和“开关的多个滑动齿轮”的语义。然而, 基于从上下文和多特征融合中学习到的语义信息, CaStr 模型却可以准确地识别。

表 8　不同模型从 CSGC-3 数据集中抽取出的关系实例

Table 8　Examples of relation extracted by different models from CSGC-3 dataset

文本	由CaStr模型输出的三元组	由BiLSTM[18]网络模型输出的三元组	由AGGCN[10]模型输出的三元组
1号主变压器控制箱内液晶显示器与变压器的温差较大, 怀疑显示不准确。	{主变压器, 控制箱, 从属}, {控制箱, 温差, 缺陷}, {温差, 显示不准确, 原因}。	{主变压器, 控制箱, 从属}, {主变压器, 温差, 缺陷}, {温差, 显示不准确, 原因}。	{主变压器, 机械限制, 从属}, {机械限制, 松动, 缺陷}, {松动, 开关的多个齿轮滑动, 原因}。
变压器的机械限制松动, 导致开关的多个齿轮滑动。	{主变压器, 机械限制, 从属}, {机械限制, 开关的多个齿轮, 缺陷}, {开关的多个齿轮, 松动, 原因}。		

4　结　论

1) 本文提出了融合 GCN 和 BiLSTM 网络的关系抽取模型, 将上下文语义和结构信息结合起来进行关系抽取。

2) 本文既利用了级联连接的 GCN 用于编码词语之间的长距离相关性, 又利用了 BiLSTM 网络用于编码语言序列信息以进行表示学习, 还基于注意力机制的融合方法用于组合特征表示以便后续进行关系抽取。

3) 在公共数据集和手工构建的数据集上进行的大量实验证明了本文模型的优越性。

本文模型兼备基于序列的关系抽取模型和基于结构的关系抽取模型的优点, 下一步将结合预训练模型提取更有效的上下文特征, 以进一步提高模型的关系特征描述能力。

参考文献 (References)

[1] BOLLACKER K, EVANS C, PARITOSH P, et al. Freebase: A collaboratively created graph database for structuring human knowledge[C]//Proceedings of the ACM SIGMOD International Conference on Management of Data. New York: ACM, 2008: 1247-1250.

[2] YU M, YIN W P, HASAN K S, et al. Improved neural relation detection for knowledge base question answering[EB/OL]. (2017-04-20)[2022-04-10]. http://arxiv.org/abs/1704.06194.

[3] 宁康, 陈挺. 生物医学大数据的现状与展望[J]. 科学通报, 2015, 60(5): 534-546.
NING K, CHEN T. Big Data for biomedical research: Current status and prospects[J]. Chin Sci Bull, 2015, 60(5): 534-546(in Chinese).

[4] 王群弼. 电力领域实体关系抽取及知识图谱构建研究[D]. 北京: 中国地质大学 (北京), 2020: 23-70.
WANG Q B. Research on relation extraction and knowledge graph construction in the filed of electric power[D]. Beijing: China University of Geosciences (Beijing), 2020: 23-70(in Chinese).

[5] BEKOULIS G, DELEU J, DEMEESTER T, et al. Adversarial training for multi-context joint entity and relation extraction[EB/OL]. (2018-08-21)[2022-04-12]. http://arxiv.org/abs/1808.06876.

[6] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[7] LIU C Y, SUN W B, CHAO W H, et al. Convolution neural network for relation extraction[C]//Proceedings of the International Conference on Advanced Data Mining and Applications. Berlin: Springer, 2013: 231-242.

[8] CHENG J P, DONG L, LAPATA M. Long short-term memory-networks for machine reading[EB/OL]. (2016-01-25)[2022-04-12]. http://arxiv.org/abs/1601.06733.

[9] TAI K S, SOCHER R, MANNING C D. Improved semantic representations from tree-structured long short-term memory networks [EB/OL]. (2015-02-28)[2022-04-12]. http://arxiv.org/abs/1503.00075.

[10] GUO Z J, ZHANG Y, LU W. Attention guided graph convolutional networks for relation extraction[EB/OL]. (2019-06-18)[2022-04-13]. http://arxiv.org/abs/1906.07510.

[11] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks[EB/OL]. (2016-09-09)[2022-04-13]. http://arxiv.org/abs/1609.02907.

[12] SHI M Y, HUANG J Y, LI C F. Entity relationship extraction based on BLSTM model[C]//Proceedings of the IEEE/ACIS International Conference on Computer and Information Science. Piscataway: IEEE Press, 2019: 266-269.

[13] LI J, WANG X, TU Z P, et al. On the diversity of multi-head attention[J]. Neurocomputing, 2021, 454: 14-24.

[14] ZENG D J, LIU K, LAI S W, et al. Relation classification via convolutional deep neural network[C]//Proceedings of the COLING In-

ternational Conference on Computational Linguistics: Technical Papers. Stroudsburg: Association for Computational Linguistics, 2014: 2335-2344.

[15] JIANG X T, WANG Q, LI P, et al. Relation extraction with multi-instance multi-label convolutional neural networks[C]//Proceedings of the COLING International Conference on Computational Linguistics: Technical Papers. Stroudsburg: Association for Computational Linguistics, 2016: 1471-1480.

[16] LIN Y K, SHEN S Q, LIU Z Y, et al. Neural relation extraction with selective attention over instances[C]//Proceedings of the Annual Meeting of the Association for Computational Linguistics. Stroudsburg: Association for Computational Linguistics, 2016: 2124-2133.

[17] NGUYEN T H, GRISHMAN R. Relation extraction: Perspective from convolutional neural networks[C]//Proceedings of the Workshop on Vector Space Modeling for Natural Language Processing. Stroudsburg: Association for Computational Linguistics, 2015: 39-48.

[18] ZHOU P, SHI W, TIAN J, et al. Attention-based bidirectional long short-term memory networks for relation classification[C]//Proceedings of the Annual Meeting of the Association for Computational Linguistics. Stroudsburg: Association for Computational Linguistics, 2016: 207-212.

[19] ZHANG Y H, ZHONG V, CHEN D Q, et al. Position-aware attention and supervised data improve slot filling[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg: Association for Computational Linguistics, 2017: 35-45.

[20] HUANG Z H, XU W, YU K. Bidirectional LSTM-CRF models for sequence tagging[EB/OL]. (2015-08-09)[2022-04-15] http://arxiv.org/abs/1508.01991.

[21] SCARSELLI F, GORI M, TSOI A C, et al. The graph neural network model[J]. IEEE Transactions on Neural Networks, 2009, 20(1): 61-80.

[22] FU T J, LI P H, MA W Y. GraphRel: Modeling text as relational graphs for joint entity and relation extraction[C]//Proceedings of the Annual Meeting of the Association for Computational Linguistics. Stroudsburg: Association for Computational Linguistics, 2019: 1409-1418.

[23] ZHANG Y H, QI P, MANNING C D. Graph convolution over pruned dependency trees improves relation extraction[EB/OL]. (2018-09-26)[2022-04-16]. http://arxiv.org/abs/1809.10185.

[24] GENG Z Q, CHEN G F, HAN Y M, et al. Semantic relation extraction using sequential and tree-structured LSTM with attention[J]. Information Sciences, 2020, 509: 183-192.

[25] MIWA M, BANSAL M. End-to-end relation extraction using LSTMs on sequences and tree structures[EB/OL]. (2016-01-05)[2022-04-16]. http://arxiv.org/abs/1601.00770.

[26] VERGA P, STRUBELL E, MCCALLUM A. Simultaneously self-attending to all mentions for full-abstract biological relation extraction[EB/OL]. (2018-02-28)[2022-04-17]. http://arxiv.org/abs/1802.10569.

[27] YAO Y, YE D M, LI P, et al. DocRED: A large-scale document-level relation extraction dataset[EB/OL]. (2019-06-14)[2022-04-17]. http://arxiv.org/abs/1906.06127.

[28] QUIRK C, POON H. Distant supervision for relation extraction beyond the sentence boundary[EB/OL]. (2019-09-15)[2022-04-17]. http://arxiv.org/abs/1609.04873.

[29] NAN G S, GUO Z J, SEKULIĆ I, et al. Reasoning with latent structure refinement for document-level relation extraction[EB/OL]. (2020-05-13)[2022-04-18]. http://arxiv.org/abs/2005.06312.

[30] CHRISTOPOULOU F, MIWA M, ANANIADOU S. Connecting the dots: Document-level neural relation extraction with edge-oriented graphs[EB/OL]. (2019-08-31)[2022-04-18]. http://arxiv.org/abs/1909.00228.

[31] JIA R, WONG C, POON H. Document-level N-ary relation extraction with multiscale representation learning[EB/OL]. (2019-04-04)[2022-04-18]. http://arxiv.org/abs/1904.02347.

[32] GUPTA P, RAJARAM S, SCHÜTZE H, et al. Neural relation extraction within and across sentence boundaries[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Washton, D. C. : AAAI, 2019, 33(1): 6513-6520.

[33] CHRISTOPOULOU F, MIWA M, ANANIADOU S. A walk-based model on entity graphs for relation extraction[C]//Proceedings of the Annual Meeting of the Association for Computational Linguistics. Stroudsburg: Association for Computional Linguistics, 2018: 81-88.

[34] LIU Y, LAPATA M. Learning structured text representations[J]. Transactions of the Association for Computational Linguistics, 2018, 6: 63-75.

[35] PENNINGTON J, SOCHER R, MANNING C. Glove: Global vectors for word representation[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg: Association for Computational Linguistics, 2014: 1532-1543.

[36] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the International Conference on Neural Information Processing Systems. New York: ACM, 2017: 6000–6010.

[37] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 2261-2269.

[38] HENDRICKX I, KIM S N, KOZAREVA Z, et al. SemEval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals[EB/OL]. (2019-11-23)[2022-04-20]. http://arxiv.org/abs/1911.10422.

[39] PENG N Y, POON H, QUIRK C, et al. Cross-sentence N-ary relation extraction with graph LSTMs[J]. Transactions of the Association for Computational Linguistics, 2017, 5: 101-115.

[40] MANNING C, SURDEANU M, BAUER J, et al. The stanford CoreNLP natural language processing toolkit[C]//Proceedings of the Annual Meeting of the Association for Computational Linguistics: System Demonstrations. Stroudsburg: Association for Computational Linguistics, 2014: 55-60.

[41] KINGMA D P, BA J. Adam: A method for stochastic optimization[EB/OL]. (2014-12-22)[2022-04-20]. https://arxiv.org/abs/1412.6980.

Relation extraction based on fusion of graph structure and sequence features

WU Tongxin[1, *], JI Xin[1, 2], WANG Honggang[1], YANG Zhiwei[1], CHEN Yiting[1], ZHAO Jiakui[1]

(1. Big Data Center of State Grid Corporation of China, Beijing 100052, China;
2. School of Computer Science and Engineering, Beihang University, Beijing 100191, China)

Abstract: Relation extraction is an important task for natural language processing applications. Most of the existing relation extraction methods mainly predict the relation based on language sequence features or structure information of sentences, which fails to effectively reflect the internal structure and features of the relation between entities. In this paper, a relation extraction model fusing graph structure and sequence feature information in sentences was proposed. The model used an attention-based graph convolutional neural network (GCN) to learn the structure information of sentences and utilized bi-directional long short-term memory (BiLSTM) to learn the sequence semantics. The relation was classified by considering the two features through the attention mechanism. Extensive experiments were conducted on a public dataset and a manually constructed dataset, which demonstrated the priority of the proposed model.

Keywords: information extraction; relation extraction; graph neural network; sequence model; feature fusion

Received: 2022-08-11; **Accepted**: 2022-10-04; **Published Online**: 2022-11-07 09:19

URL: link.cnki.net/urlid/11.2625.V.20221104.1415.002

Foundation item: Big Data Center of State Grid Corporation of China Technology Project (52999021N005)

* **Corresponding author**. E-mail: tongxin-wu@sgcc.com.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0743

基于BP神经网络的甲板运动预报与补偿技术

赵所[1,2,*]，林立[2]，李震[2]，侯中喜[1]
(1. 国防科技大学 空天科学学院，长沙 410073；2. 沈阳飞机设计研究所，沈阳 110135)

摘 要：在全自动着舰过程中，考虑甲板运动对全自动着舰成功率的影响。针对舰载机响应滞后带来的甲板和舰载机纵向高度不同步的问题，采用甲板运动预报技术对舰船未来有效时间内的运动参数进行预报。建立基于正弦波组合的六自由度甲板运动模型，基于反向传播（BP）人工神经网络建立甲板运动预报模型，提出采用反向传播优化算法减小预报模型训练更新时间，提高预报模型速率和精度。基于超前网络建立甲板运动补偿模型，结合着舰导引系统确保舰载机与甲板运动同相位。建立飞机运动模型，并通过仿真实验，验证甲板运动预报和补偿技术的可行性。

关 键 词：全自动着舰；甲板运动预报；甲板运动补偿；BP神经网络；超前网络
中图分类号：V249
文献标志码：A **文章编号**：1001-5965(2024)09-2772-09

与飞机着陆相比，舰载机在着舰过程中，需要保持与甲板运动在高度上的同步[1]。但航空母舰由于受到海风和浪涌等多种因素的影响，将会出现偏航、横摇、纵摇、沉浮等多种形式的甲板运动，运动的不确定性及环境干扰，极大地增加了着舰难度，这往往会导致舰载机无法与阻拦索啮合，发生逃逸、撞击等事故[2-3]。由于甲板运动严重影响了舰载机着舰过程的安全性和成功率。因此，需要通过甲板运动预报来预估得到未来一段时间的甲板运动信号，并将其引入着舰导引系统中，以抵消舰载机响应滞后带来的甲板和舰载机纵向高度、横纵角度不同步的问题。既有的预估方法包括卡尔曼滤波、粒子滤波、神经网络及自回归AR模型等，但由于导引系统与实际甲板运动不可避免地会存在一定的相位滞后，就需要与甲板运动补偿相结合[4-5]，保持着舰前飞机与甲板运动同步，以保证舰载机可以在理想着舰点附近安全着舰。目前研究的甲板预报与补偿技术已经有了长足的发展，预报可以达到12 s内相对精确的预估，但在实时速率和精度上还有很大提升空间。

本文首先对甲板运动进行基于正弦波的建模；其次，基于反向传播(back-propagation, BP)人工神经网络对甲板运动进行预报，提出甲板运动预报模型使用优化算法和在线训练来减少模型参数训练更新所需时间，提高模型精度，并采用超前网络进行补偿，确保舰载机与甲板运动同相位；最后，建立飞机纵向运动模型，对本文预报模型和补偿模型进行了仿真验证。

1 甲板运动模型建立

甲板运动模型主要分为基于正弦波组合的甲板运动模型和基于功率谱的甲板运动模型[6-7]，本文采用基于正弦波组合的甲板运动模型。以舰船质心为坐标，建立六自由度的甲板运动坐标系，如图1所示，其中，x_V为甲板运动方向。绕x轴、y轴和z轴的3种运动分别为横摇、首摇和纵摇，沿着x轴、y轴和z轴的运动为纵荡、升沉和横荡。其中，横摇、纵摇和升沉能够返回初始位置，因此，将其看作窄频带的平稳随机过程。

实验过程中将运动简化为

收稿日期：2022-08-29；录用日期：2023-02-07；网络出版时间：2023-03-30 14:48
网络出版地址：link.cnki.net/urlid/11.2625.V.20230330.1044.001
*通信作者. E-mail：250772665@qq.com

引用格式：赵所，林立，李震，等. 基于BP神经网络的甲板运动预报与补偿技术[J]. 北京航空航天大学学报，2024，50(9)：2772-2780.
ZHAO S，LIN L，LI Z，et al. Deck motion prediction and compensation technology based on BP neural network[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2772-2780 (in Chinese).

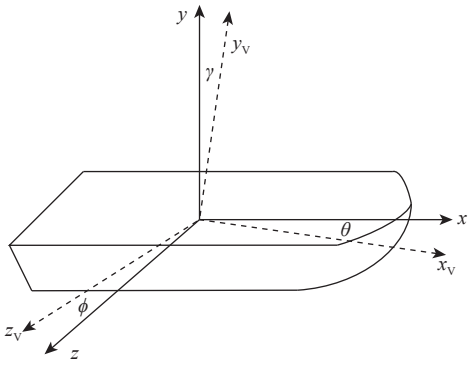


图 1　六自由度甲板运动模型

Fig. 1　Deck motion model in six degree of freedom

$$\begin{cases}\gamma=\sum\limits_{i=0}^{N}\gamma_{mi}\sin\left(\dfrac{2\pi}{T_{yi}}t+\gamma_{0i}\right)\\ \theta=\sum\limits_{i=0}^{N}\theta_{mi}\sin\left(\dfrac{2\pi}{T_{yi}}t+\theta_{0i}\right)\\ h=\sum\limits_{i=0}^{N}h_{mi}\sin\left(\dfrac{2\pi}{T_{yi}}t+h_{0i}\right)\end{cases}\tag{1}$$

式中：γ和θ分别为甲板的横摇和纵摇角度；h为升沉位移；γ_m指横摇运动的幅值；t为甲板运动周期。

考虑到实际着舰过程中，舰载机需要的是着舰点处的甲板运动参数，构建船舰质心到着舰点的运动模型：

$$h_z=h+l\sin\theta+w\tag{2}$$

式中：h_z为着舰点的升沉位移；l为船舰质心与着舰点之间的距离；w为白噪声。

根据式(2)可以得到甲板着舰点的横摇、纵摇及升沉运动模型，将其作为实验数据，在运动模型数据上叠加确定方差的白噪声w。航母在 5 级海况，以 15.4 m/s(30 kn) 的航速状态行驶时，海浪及风作用于舰体引起的甲板运动响应曲线如图 2 所示。

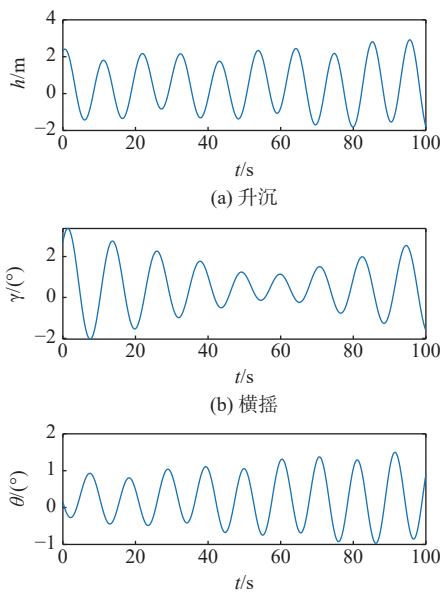


图 2　甲板运动响应曲线

Fig. 2　Deck motion response curve

2　基于 BP 神经网络的甲板运动预报

针对甲板运动预报采取经典的精确数学模型，通常会面临计算复杂、不确定、非线性等问题，而人工神经网络作为一种机器学习模拟生物神经反馈过程的计算模型，能够达到并行、分布式处理数据信号，通过调节模型内部计算单元之间的相互连接关系的智能学习训练。采取神经网络方法进行甲板运动预报，能够极大地降低计算的复杂度。

人工神经网络领域中，BP 神经网络因其具有很强的非线性映射能力和柔性的网络结构而被广泛应用，但其也存在学习速率慢的问题，即使是简单问题也需要几百次甚至上千次的学习才能收敛。本文基于 BP 神经网络提出的优化反向传输算法和在线训练，能够在继承 BP 神经网络优势的同时进一步提高精度及速率。

2.1　基于经典 BP 的甲板运动预报

1986 年，Rumelhart 等提出了 BP 神经网络的概念，因其具有很强的非线性映射能力，所以一经提出就引发了广泛关注[8]。

BP 神经网络结构图如图 3 所示，其基本处理单元是神经元，$x_1,x_2,\cdots,x_m$为输入信号，$c_1,c_2,\cdots,c_n$为中间层处理结果，$y_1,y_2,\cdots,y_l$为测试结果，e 为无量纲化相对误差，v_{ij}和ω_{jk}分别为各神经元间对应权值[9-10]。

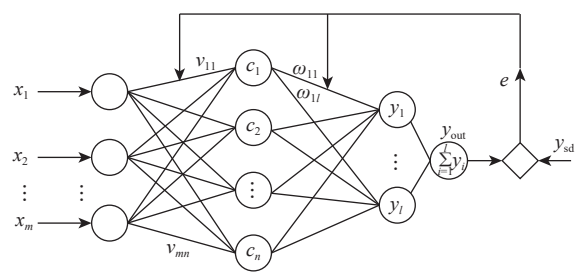


图 3　BP 神经网络

Fig. 3　BP neural network

数据信号在各层间传输处理过程中如果直接求解，那么无论网络结构多么复杂，输出和输入信号都是线性组合关系。因此，在网络模型各层级的计算单元间加入非线性的激活函数，能够使神经网络面向非线性问题，从而提升网络对模型的表征能力。信号在各层间的计算如下：

$$c_j=f\left(\sum_{i=0}^{m}v_{ij}x_i\right)\qquad j=1,2,\cdots,n\tag{3}$$

$$y_k=g\left(\sum_{j=0}^{n}\omega_{jk}c_j\right)\qquad k=1,2,\cdots,l\tag{4}$$

式中：函数$f(x)$和$g(x)$为各层间的激活函数。在实

验中本文采取双曲正切S型激活函数：

$$\mathrm{tansig}(n)=\frac{2}{1+\mathrm{e}^{-2n}}-1 \tag{5}$$

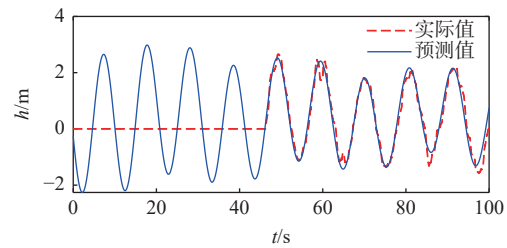
当神经网络预估得到的输出值不满足误差精度要求时，模型就会进入反向传播。要使平方和误差减小，反向传输过程各层权值的调整量需要与误差的梯度下降成正比：

$$\Delta\omega=-\eta\frac{\partial E}{\partial\omega}=-\eta\frac{\partial E}{\partial y}\cdot\frac{\partial y}{\partial\omega}=\eta\delta_k c_j \tag{6}$$

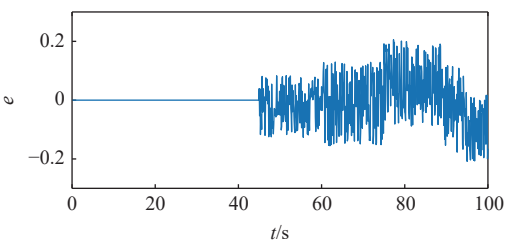
式中：η为更新过程中的更新步长；E为平方和误差函数；δ为误差函数对输出层的各神经元的偏导数。

BP神经网络在使用训练样本进行训练时，主要通过调整权值来训练更新模型，调整后的权值再次通过数据信号进行预估、误差比对，循环往复直到满足精度需求。

使用经典BP神经网络对第1节利用着舰点处的甲板运动模型产生的升沉运动状态进行甲板运动预报，使用前40 s甲板运动数据作为训练样本，设输入层神经元为80，中间层神经元为10，输出层神经元数目设为50，提前5 s对5级海况下的甲板运动的预报情况及相对误差如图4所示。



(a) 基于BP神经网络的升沉运动预测值与实际值对比



(b) 误差

图4 基于BP神经网络的升沉运动预估

Fig. 4 Prediction of heaving motion based on BP neural network

实验结果均方误差(mean-square error, MSE)为0.131 83 h/m。从预测结果可以看出，基于经典BP神经网络的预报运动曲线与实际曲线趋势基本吻合，能够比较准确地对甲板运动进行预报，但预报误差较大，且采用离线模型训练无法在实际应用中实时跟进，在模型精度和实时性能上仍有提升空间。

2.2 反向传输算法

BP神经网络的反向传输过程要求各层权值的调整更满足精度需求，因此，调整量需要与误差的梯度下降成正比，即梯度下降法，但这个方法在迭代中寻找局部最优的方向时，会选择与之前十分相近的方向，即网络在该方向上的参数更新还未满足误差需求，这就导致在更新过程中的轨迹呈现锯齿状。

为减小更新时调整方向所需时间，选择优化方向，以确保步长能够使该方向的参数更新满足优化需求。

对于反向传输时模型的优化问题：

$$\min_{\boldsymbol{x}} f(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x} \tag{7}$$

式中：$\boldsymbol{x}$为模型参数矩阵；$\boldsymbol{A}$为正定矩阵；$\boldsymbol{b}$为已知常量矩阵。为确保优化方向$\boldsymbol{p}_k$间不相互干扰必须满足共轭正交$\boldsymbol{p}_i^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_j=0(i\neq j)$。若令

$$\boldsymbol{p}_k=-\boldsymbol{g}_k+\beta_{k-1}\boldsymbol{p}_{k-1}+\sum_{i=0}^{k-1}\beta_k^{(i)}\boldsymbol{p}_i \qquad k\geqslant 1 \tag{8}$$

式中：$\boldsymbol{g}_k=\dfrac{\partial f(\boldsymbol{x}_k)}{\partial\boldsymbol{x}_k}=\boldsymbol{A}\boldsymbol{x}_k+\boldsymbol{b}$为梯度。那么系数$\beta_k$必须满足优化方向共轭条件，计算可得

$$\beta_k^{(i)}=\frac{\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_i}{\boldsymbol{p}_i^{\mathrm{T}}\boldsymbol{A}\boldsymbol{p}_i}=\frac{\boldsymbol{g}_k^{\mathrm{T}}(\boldsymbol{g}_{i+1}-\boldsymbol{g}_i)}{\boldsymbol{p}_i^{\mathrm{T}}(\boldsymbol{g}_{i+1}-\boldsymbol{g}_i)}=0 \tag{9}$$

因此，第k步的优化方向为

$$\boldsymbol{p}_k=-\boldsymbol{g}_k+\beta_{k-1}\boldsymbol{p}_{k-1} \tag{10}$$

确定优化方向后，需要求得优化目标进行计算更新以调整参数，但由于一般真实模型比较复杂，因此，需要在求解过程中采用在当前优化点的邻域对函数进行泰勒展开，逼近$\boldsymbol{x}_k$附近点$\boldsymbol{x}_k+\boldsymbol{\Delta}_k$的估计值：

$$f(\boldsymbol{x}_k+\boldsymbol{\Delta}_k)=f(\boldsymbol{x}_k)+\boldsymbol{J}_{\boldsymbol{x}}^k\boldsymbol{\Delta}_k \tag{11}$$

式中：$\boldsymbol{J}_{\boldsymbol{x}}^k$为$\boldsymbol{x}_k$处的雅可比矩阵。若设$\boldsymbol{\Delta}_k$为$\boldsymbol{x}_k$处的模型参数增量，$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\boldsymbol{\Delta}_k$，则模型：

$$\|\boldsymbol{y}-f(\boldsymbol{x}_{k+1})\|=\min_{\boldsymbol{\Delta}_k}\left\|\boldsymbol{e}_k-\boldsymbol{J}_{\boldsymbol{x}}^k\boldsymbol{\Delta}_k\right\| \tag{12}$$

式中：$\boldsymbol{e}_k=\boldsymbol{y}-f(\boldsymbol{x}_k)$。设$\left\|\boldsymbol{e}_k-\boldsymbol{J}_{\boldsymbol{x}}^k\boldsymbol{\Delta}_k\right\|=0$，则可以得到优化点的迭代：

$$\boldsymbol{x}_{k+1}=\boldsymbol{x}_k+\boldsymbol{J}_{\boldsymbol{x}}^{k-1}(\boldsymbol{y}-f(\boldsymbol{x}_k)) \tag{13}$$

使用BP神经网络模型在预估甲板运动时，在反向传输过程中通过优化算法大大减少了调整方向时搜索网络的时间，这确保精确度的同时也能够保证速率。

2.3 甲板运动预报仿真验证

经典的BP神经网络模型采取离线训练的方式进行预估，通过历史样本数据进行模型初始化训练后，直接输入实验样本进行模型计算得到结果，但实际舰载机着舰过程需要对舰船的运动状态进行实时采样以更新甲板预报模型的输入样本，因此，

将预报器的训练方式更改为在线训练。

在线训练模型的甲板运动预报流程如图 5 所示, 即在初始化训练后, 测试数据将以预设时间为周期进行采样, 在模型计算预估结果的同时, 将新采样数据与历史数据相结合作为训练样本对模型进行训练及参数更新, 以保障预报模型的实时性能。

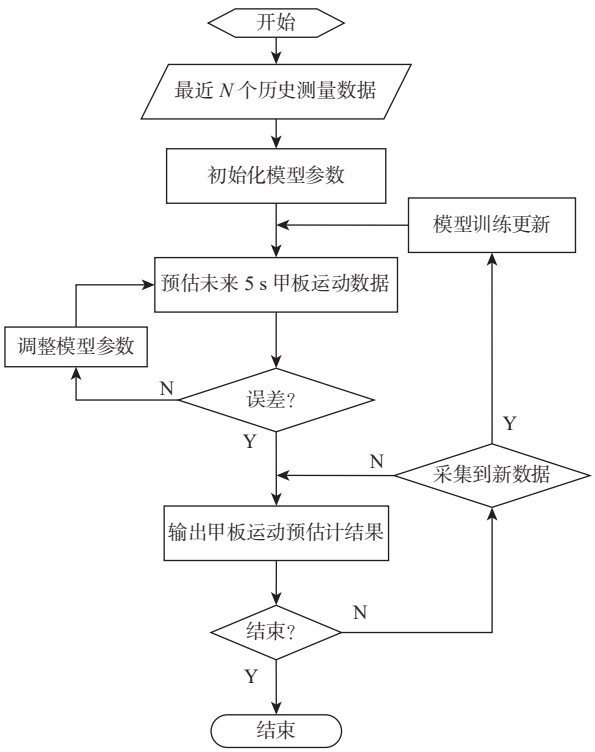


图 5　实时甲板运动预报流程

Fig. 5　Deck motion real-time prediction process

以六自由度甲板运动模型作为测量数据源, 验证基于 BP 神经网络优化模型的甲板运动预报技术的可行性。在第 1 节中确定的多正弦波组合运动模型中加入方差为 0.01 的随机白噪声作为甲板横摇、纵摇和升沉运动的样本数据, 并设系统采样时间为 0.1 s。

实验采用在线训练的方式进行模型更新, 传输周期设为 1.0 s, 将实验所用甲板运动数据集分为训练数据及实验预估数据, 先对输入数据进行归一化处理, 确定神经网路各层级参数, 再设输入层神经元数为 80, 即输入数据样本的维数为 80, 根据预测时间 5 s, 输出层神经元数目设为 50, 中间层经验计算式:

$$l=\sqrt{n+m+a} \tag{14}$$

式中: n为输入层神经元数目; m为输出层神经元个数; a为 1～10 的任意数, 由此确定中间层层数l为 12。设模型训练迭代次数为 1 000, 学习速率为 0.05, 期望目标最小误差为 0.000 01。

本文预估模型对甲板升沉运动的期望值与预测值对比及预估相对误差如图 6 所示。

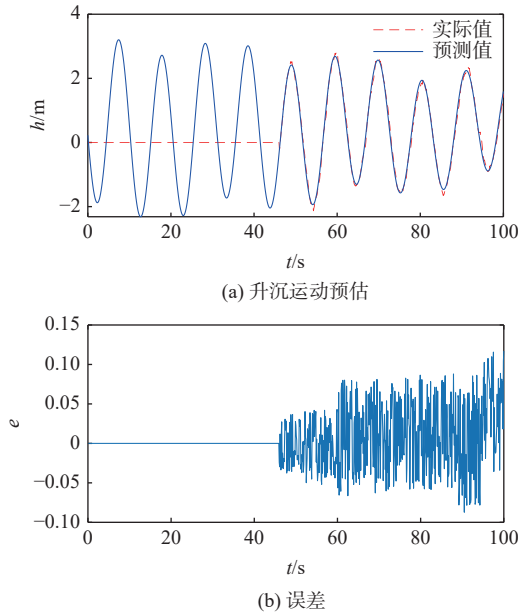


图 6　升沉运动预估

Fig. 6　Prediction of heaving motion

由图 6 可知, 预估曲线与实际样本曲线十分吻合, 基于本文预估模型的甲板运动预报模型能够通过历史数据很好地预测甲板运动未来有效时间内的运动变化趋势。相较于基于经典 BP 的甲板运动预报方法实验结果, 本文提出的预报模型在在线训练过程中, 预测误差有一定程度的减小, 即能够跟上实时训练及预报过程, 并可以更精准快速地预估甲板运动。由此可以说明, 基于优化模型的甲板运动预报技术对于随机甲板运动模型的预估是非常有效的, 且提高学习速率后能够满足实时状态下的甲板运动预报精度要求。

针对预报模型的性能研究, 测试其在 4 级海况下提前 4, 7, 10, 13, 16, 19 s 对甲板的升沉运动进行预报, 实验结果如表 1 所示, 可以得出, 基于优化算法的预报模型在 16 s 内都能比较准确地预估甲板未来运动趋势, 且对升沉、纵摇和横摇 3 种运动的实验结果误差都能满足精度要求。

在预报提前 10 s 的情况下分别对不同海况下的升沉、横摇和纵摇运动进行预估, 如图 7 所示, 可以看到在 3～5 级海况中, 本文预报模型都能够精

表 1　甲板运动预报性能

Table 1　Deck motion prediction performances

时间/s	升沉MSE/m	纵摇MSE/(°)	横摇MSE/(°)
4	0.024 875	0.025 589	0.029 882
7	0.023 695	0.036 056	0.027 329
10	0.029 782	0.039 349	0.043 956
13	0.068 56	0.076 037	0.077 719
16	0.089 719	0.097 24	0.095 743
19	0.125 33	0.146 59	0.131 588

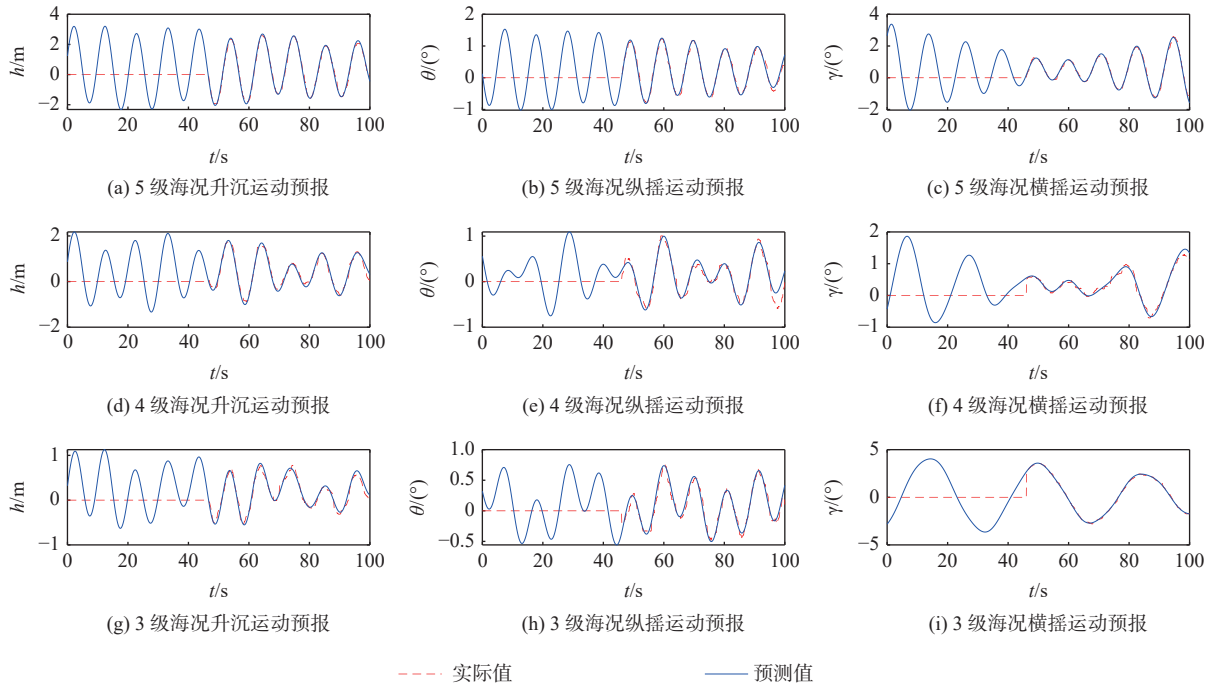


图7 本文预报模型在不同海况下的仿真结果

Fig. 7 Simulation results of the proposed predicting model under different sea conditions

确地预测甲板未来运动趋势。可以说明，本文预报模型能够适应复杂环境及不同程度的剧烈运动挑战。

3 基于超前网络的甲板运动补偿

舰载机着舰过程中，为确保能够在目标着舰点与舰船甲板啮合，舰载机导引需要考虑到甲板的纵向运动问题，使得舰载机在着舰过程中避免因为舰船甲板下降导致舰载机着舰失败逃逸复飞，或舰船突然升高使得飞机撞击甲板或舰船尾部的情况。将理想的甲板着舰点的纵向运动即高度变化信息加入纵向自动着舰导引系统 (longitudinal automatic carrier landing system, $\mathrm{ASCL}|_{\mathrm{Long}}$) 中，以确保舰载机在着舰引导航行的同时能够跟随甲板扰动同相位运动[2,11]。

但由于纵向的导引系统在舰船运动的有效工作特征频段内会有不可避免的相位滞后，导致系统难以实时准确地跟踪甲板纵向运动，因此，针对相位产生的实时性能和误差问题，预先对舰船的高度变化进行补偿，以确保飞机着舰精度。图8为舰载机的纵向导引系统。

图中，H_{DM}为理想的甲板高度变化信号，H_0为舰载机的初始高度，H_{com}为理想下滑轨迹，H_{error}为高度偏差信号，DMP为甲板运动预报模型，DMC为甲板运动补偿模型。着舰过程中，结合甲板运动预报模型和甲板运动补偿模型，先根据甲板历史运动信号对甲板进行实时预报，再利用获得的预估甲板运动信息进行网络补偿，处理后引入到导引系统 $\mathrm{ASCL}|_{\mathrm{Long}}$ 中。

对于甲板的运动补偿采取相位超前网络 $G_{\mathrm{DMC}}(s)$[12-13]，补偿模型需要对导引系统 $\mathrm{ASCL}|_{\mathrm{Long}}$ 的相位滞后进行补偿，则针对有效频段 $\omega_s = 0.2 \sim 1.0$ rad/s，理想的补偿情况应当满足：

$$G_{\mathrm{DMC}}(s)G_{\mathrm{ACLS}}(s)|_{s=j\omega,\omega=0.2\sim1.0} = 1 \tag{15}$$

式中：$G_{\mathrm{ACLS}}(s)$为导引系统 $\mathrm{ASCL}|_{\mathrm{Long}}$ 的闭环传递函数。

为保证着舰系统在有效频段中满足频域响应特性要求，采取如下结构：

$$G_{\mathrm{DMC}}(s) = K_{\mathrm{DMC}}G_{\mathrm{L}}(s)G_{\mathrm{C}}(s)G_{\mathrm{F}}(s) \tag{16}$$

式中：K_{DMC}为补偿滤波器增益；$G_{\mathrm{L}}(s)$为低通滤波器，用于抑制高频噪声；$G_{\mathrm{C}}(s)$为针对低通滤波的补

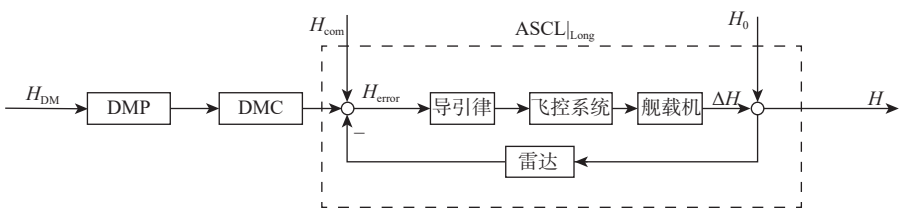


图8 舰载机纵向导引系统

Fig. 8 Carrier aircraft longitudinal guidance system

偿环节; $G_F(s)$为相位超前网络, 对系统相位进行整体调整以确保舰载机在经过导引系统后与甲板运动同步。其中:

$$G_L(s)=c1\Big/\left(\frac{1}{\tau_{DMC}}s+1\right) \tag{17}$$

$$G_C(s)=\left(\frac{s^2}{\omega_{DMC}{}^2}+\frac{2\xi_{DMC}}{\omega_{DMC}}s+1\right)\Big/\left(\frac{1}{\omega_n}s+1\right)^2 \tag{18}$$

$$G_F(s)=(\alpha T s+1)/(Ts+1) \tag{19}$$

取滤波器时间常数τ_{DMC}为 2 s, 自然频率ω_{DMC}为 0.65 rad/s, 阻尼ξ_{DMC}为 0.45 作为仿真的超前补偿网络。

采用确定的甲板运动模型进行补偿模型仿真验证如图 9 所示。

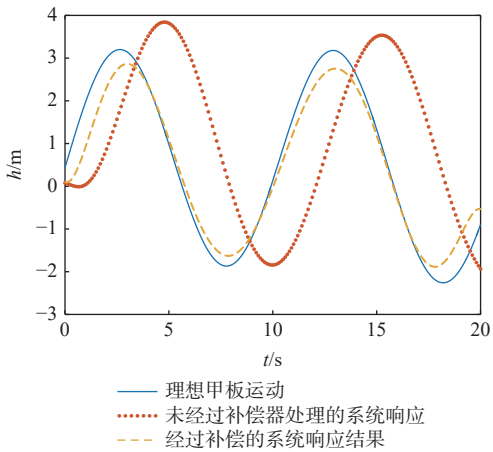


图 9　甲板运动补偿模型实验效果

Fig. 9　Experimental effect of deck motion compensation model

4　甲板运动扰动着舰系统仿真验证

4.1　舰载机纵向运动模型

舰载机纵向控制系统如图 10 所示, 信号通过控制律将指令传送给舵机, 并计算出运动模型获取运动参数。

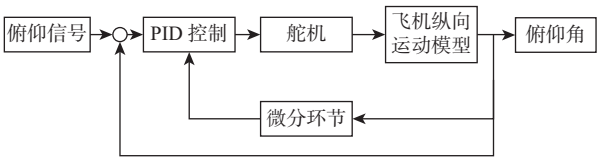


图 10　舰载机纵向控制系统

Fig. 10　Carrier aircraft longitudinal control system

根据飞机运动状态的动力学关系可以得到舰载机的运动学方程组如下:

$$\begin{cases}\dot{\theta}_f=q\cos\phi-r\sin\phi\\ \dot{\psi}=\dfrac{1}{\cos\theta_f}(q\sin\phi+r\cos\phi)\\ \dot{\phi}=p+q\sin\phi\tan\theta_f+r\cos\phi\tan\theta_f\end{cases} \tag{20}$$

式中: θ_f为俯仰角; ψ为偏航角; ϕ为滚转角; p、q和r分别为绕机体坐标系 x 轴、y 轴和 z 轴的角速度。

舰载机着舰的线性化状态空间方程[14-15]可以表示为

$$\dot{\boldsymbol{x}}_h=\boldsymbol{A}_s\boldsymbol{x}_h+\boldsymbol{B}\boldsymbol{u} \tag{21}$$

式中: $\boldsymbol{x}_h$表示飞机的各个状态变量; $\boldsymbol{u}$为飞机的控制量; $\boldsymbol{A}_s$和$\boldsymbol{B}$为$\boldsymbol{x}_h$和$\boldsymbol{u}$线性关系的对应状态空间矩阵。其中:

$$\begin{cases}\boldsymbol{x}_h=\left[\Delta V,\Delta\alpha,\Delta\beta,\Delta p,\Delta q,\Delta r,\Delta\theta_f,\Delta\phi,\Delta\psi\right]^T\\ \boldsymbol{u}=[\Delta\delta_e,\Delta\delta_T,\Delta\delta_a,\Delta\delta_r]^T\end{cases} \tag{22}$$

式中: V为速度; α为迎角; β为侧滑角; δ_e为升降舵偏角; δ_T为油门开合度; δ_a为副翼偏角, δ_r为方向舵偏角。

若围绕舰载机纵向运动来讨论, 建立飞机纵向运动模型, 则可以将状态参数简化为

$$\begin{cases}\boldsymbol{x}_h=\left[\Delta V,\Delta\alpha,\Delta q,\Delta\theta_f\right]^T\\ \boldsymbol{u}=[\Delta\delta_e,\Delta\delta_T]^T\end{cases} \tag{23}$$

设定舰载机的瞬时状态参数如表 2 所示。

表 2　舰载机瞬时状态

Table 2　Transient state parameters of carrier aircraft

参数	数值	参数	数值
高度h/m	120	偏航角速度r/((°)·s^{-1})	0
速度V/(m·s^{-1})	65.0	滚转角速度p/((°)·s^{-1})	0
迎角α/(°)	8	俯仰角速度q/((°)·s^{-1})	0
侧滑角β/(°)	0	升降舵偏角δ_e/(°)	−3.17
俯仰角θ_f/(°)	5	副翼偏角δ_a/(°)	0
偏航角ψ/(°)	0	油门开合度δ_T/%	0.43
滚转角ϕ/(°)	0	方向舵偏角δ_r/(°)	0

得到舰载机的状态空间参数矩阵为

$$\begin{cases}\boldsymbol{A}_s=\begin{bmatrix}-0.087\,3 & 3.405\,3 & -0.024\,5 & -32.735\,1\\ -0.001\,5 & -0.557\,9 & 0.989\,7 & 0.008\,5\\ 0 & -0.300\,1 & -0.185\,4 & 0\\ 0 & 0 & 1 & 0\end{bmatrix}\\ \boldsymbol{B}=\begin{bmatrix}-0.004\,5 & 18.780\,3\\ -0.001\,4 & -0.011\,8\\ -0.028\,5 & 0\\ 0 & 0\end{bmatrix}\end{cases} \tag{24}$$

4.2　仿真验证

为验证本文甲板运动预报模型和补偿模型在实际工程应用的可行性, 使用建立的飞机运动模型和甲板运动模型进行仿真验证。

舰载机着舰过程中纵向运动的示意图如图 11 所示。

设定飞机飞行高度为h_0为 120 m, 空速为 65 m/s, 舰船航速为 15 m/s, 舰船甲板高度为 30 m, 吃水深

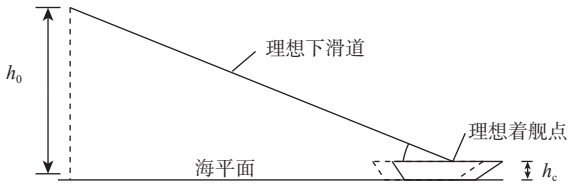
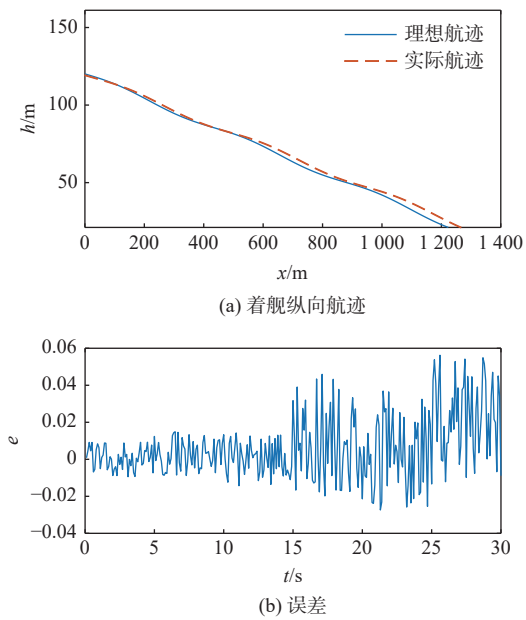


图 11 舰载机着舰下滑道示意图

Fig. 11 Schematic diagram of carrier aircraft landing glide path

度为 10 m, 即甲板距海平面距离h_c为 20 m, 舰载机距离舰船甲板着舰点的水平距离为 1 256.98 m, 进行仿真验证。使用经典 BP 的甲板运动预报如图 12 所示; 将本文预报技术及补偿技术相结合导入舰载机导引系统, 本文预报模型的舰载机纵向航迹图如图 13 所示。



(a) 着舰纵向航迹

(b) 误差

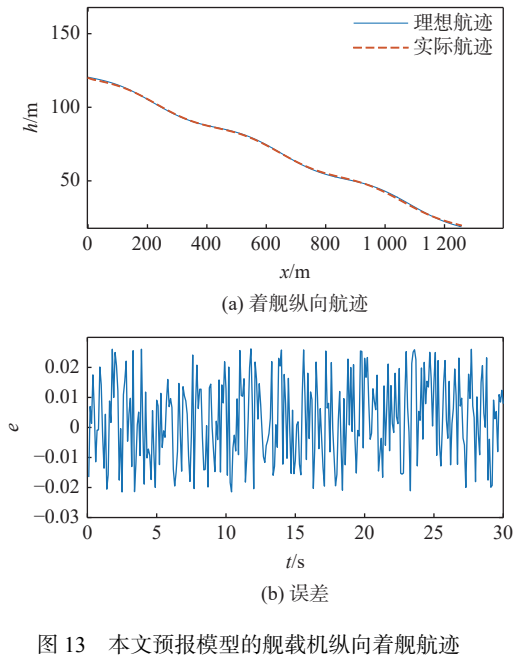
图 12 基于经典 BP 预报技术的舰载机纵向着舰航迹

Fig. 12 Longitudinal landing track of carrier aircraft based on classical BP prediction technology

由图 12 和图 13 可知, 经典的 BP 预报技术会使舰载机着舰的纵向实际航迹与理想航迹的相对误差随时间的变化而越来越大, 且有一定的相位差; 而使用本文预报模型与补偿技术相结合, 则使舰载机着舰的纵向实际航迹与理想航迹基本吻合, 且相对误差保持在较小的范围内。

上述仿真结果可以说明, 本文预报模型可以保证舰载机在着舰情况下能够提前预估到较为精确的甲板运动状态, 同时加入的甲板补偿技术能够消除导引系统自身的问题带来的相位差, 以保证舰载机与甲板同相位运动。

针对舰载机由于甲板运动的扰动产生的纵向位移进行验证分析, 图 14 为舰载机的甲板运动扰动响应与实际着舰点高度变化对比图, 由图可知, 飞机经过预报模型与补偿模型的相互作用, 着舰系

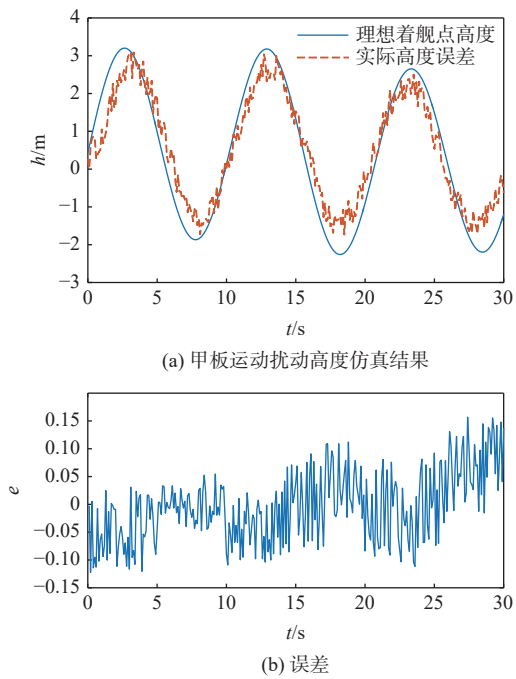


(a) 着舰纵向航迹

(b) 误差

图 13 本文预报模型的舰载机纵向着舰航迹

Fig. 13 Longitudinal landing track of carrier aircraft based on the proposed predicting model



(a) 甲板运动扰动高度仿真结果

(b) 误差

图 14 甲板运动扰动高度仿真结果

Fig. 14 Height simulation results of deck motion disturbance

统能够使机体运动与甲板运动保持同相位, 且误差较小。

图 15 和图 16 为仿真验证中状态参数的动态响应, 由图可知, 舰载机在甲板运动扰动下的迎角和速度响应基本控制在小范围内。

在不同海况下进行验证, 得到舰载机甲板运动扰动的误差情况如表 3 所示, 由表可知, 在不同的海况下, 预报模型和补偿模型都能够配合系统准确地进行着舰, 从而提高着舰成功率。

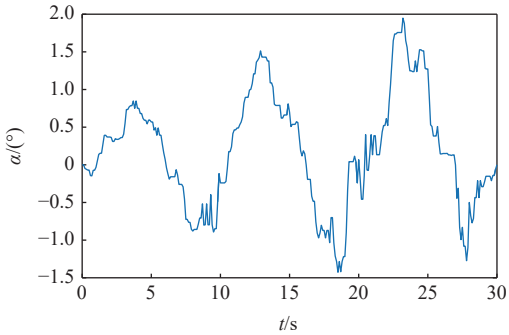


图 15　甲板运动扰动迎角响应

Fig. 15　Angle of attack response of deck motion perturbation

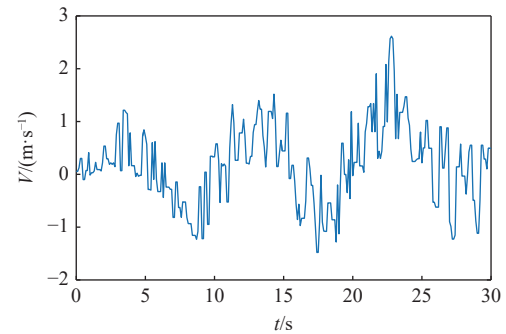


图 16　甲板运动扰动速度响应

Fig. 16　Velocity response of deck motion disturbance

表 3　不同海况下纵向甲板运动扰动实验结果

Table 3　Experimental results of longitudinal deck motion disturbance under different sea conditions

海况	高度平均误差/m	高度误差标准差/m
5级	0.460 34	3.613 54
4级	0.376 16	4.510 79
3级	0.393 52	4.014 54

5　结　论

1) 本文介绍了基于 BP 人工神经网络的甲板运动预报方法和采用超前网络的补偿方法, 提出采用反向传播优化算法和在线训练以确保神经网络模型应用能够高精度、高速率并行。

2) 在基于正弦波组合的甲板运动模型和舰载机纵向运动模型上, 验证本文甲板运动预报模型和补偿模型的性能, 仿真结果表明, 本文预报模型对于随机甲板运动模型的预估在 3～16 s 内是非常有效的, 并且能够适应多种海况各方向运动的挑战, 补偿模型能够确保飞机准确追踪甲板高度信息并同相位运动。针对舰载机着舰过程的甲板运动扰动, 预报和补偿技术相结合能够有效地处理扰动问题, 提高着舰成功率。

参考文献（References）

[1] 武恒州, 罗福平, 石星辰, 等. 全自动着舰技术现状与发展趋势分析[J]. 飞机设计, 2020, 40(6): 1-5.
WU H Z, LUO F P, SHI X C, et al. Analysis on the status quo and development trend of automatic carrier landing technology[J]. Aircraft Design, 2020, 40(6): 1-5(in Chinese).

[2] 杨一栋. 直升机飞行控制[M]. 北京: 国防工业出版社, 2007: 37-70.
YANG Y D. Helicopter flight control[M]. Beijing: National Defense Industry Press, 2007: 37-70(in Chinese).

[3] 许东松, 刘星宇, 王立新. 航母运动对舰载飞机着舰安全性的影响[J]. 北京航空航天大学学报, 2011, 37(3): 289-294.
XU D S, LIU X Y, WANG L X. Influence of carrier motion on landing safety for carrier-based airplanes[J]. Journal of Beijing University of Aeronautics and Astronautics, 2011, 37(3): 289-294(in Chinese).

[4] 王可, 徐明亮, 李亚飞, 等. 一种面向航空母舰甲板运动状态预估的鲁棒学习模型[J]. 自动化学报, 2021, 48: 1-9.
WANG K, XU M L, LI Y F, et al. A robust learning model for deck motion prediction of aircraft carrier[J]. Acta Automatica Sinica, 2021, 48: 1-9(in Chinese).

[5] 杨柳, 徐东昊. 基于极短期运动预报的舰载机着舰过程仿真分析[J]. 中国舰船研究, 2018, 13(4): 99-103.
YANG L, XU D H. Aircraft carrier landing process simulation based on extremely short-term prediction of ship motion[J]. Chinese Journal of Ship Research, 2018, 13(4): 99-103(in Chinese).

[6] LI X L, LV X H, YU J D, et al. Neural network application on ship motion prediction[C]//Proceedings of the International Conference on Intelligent Human-Machine Systems and Cybernetics. Piscataway: IEEE Press, 2017: 414-417.

[7] 王月春, 张素珍, 孙宁. 六自由度舰船运动轨迹跟踪算法分析[J]. 舰船科学技术, 2022, 44(8): 56-59.
WANG Y C, ZHANG S Z, SUN N. Analysis on tracking algorithm of 6-DOF ship motion trajectory[J]. Ship Science and Technology, 2022, 44(8): 56-59(in Chinese).

[8] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J]. Nature, 1986, 323(6088): 533-536.

[9] 石广军, 张旸, 王青林. 舰船甲板运动预估技术研究[J]. 直升机技术, 2020(4): 20-24.
SHI G J, ZHANG Y, WANG Q L. Prediction technology for ship deck motion[J]. Helicopter Technique, 2020(4): 20-24(in Chinese).

[10] 吴鹏飞, 石章松, 吴中红, 等. 无人直升机着舰甲板运动预估与补偿方法[J]. 电光与控制, 2019, 26(12): 22-27.
WU P F, SHI Z S, WU Z H, et al. An estimation and compensation method of deck motion for unmanned helicopter landing on ship[J]. Electronics Optics & Control, 2019, 26(12): 22-27(in Chinese).

[11] 张永花. 舰载机着舰过程甲板运动建模及补偿技术研究[D]. 南京: 南京航空航天大学, 2012: 32-50.
ZHANG Y H. Research on deck motion modeling and deck motion compensation for carrier landing[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2012: 32-50(in Chinese).

[12] 张明廉, 徐军. 舰载飞机自动着舰系统的研究[J]. 北京航空航天大学学报, 1994, 20(4): 386-391.
ZHANG M L, XU J. Research on automatic landing system of carrier-based aircraft[J]. Journal of Beijing University of Aeronautics

and Astronautics, 1994, 20(4): 386-391(in Chinese).

[13] 潘海飞, 王欣, 沙峰. 舰载机着舰过程中甲板运动补偿技术研究[J]. 信息技术, 2013, 37(4): 116-120.
PAN H F, WANG X, SHA F. Study on the deck motion compensation technique during carrier aircraft landing[J]. Information Technology, 2013, 37(4): 116-120(in Chinese).

[14] URNES J M, HESS R K. Development of the F/A-18A automatic carrier landing system[J]. Journal of Guidance, Control, and Dynamics, 1985, 8(3): 289-295.

[15] ANDERSON M R. Inter and outer loop manual control of carrier aircraft landing[C]//Proceedings of the AIAA, Guidance, Navigation, and Control Conference. San Diego: AIAA, 1996.

Deck motion prediction and compensation technology based on BP neural network

ZHAO Suo[1, 2, *], LIN Li[2], LI Zhen[2], HOU Zhongxi[1]

(1. College of Aerospace Sciences, National University of Defense Technology, Changsha 410073, China;
2. Shenyang Aircraft Design & Research Institute, Shenyang 110135, China)

Abstract: The impact of deck motion on the complete automatic landing success rate is taken into account during the landing procedure. Aiming at the asynchrony of the deck and the longitudinal height of the carrier-based aircraft caused by the delay in the response of the carrier-based aircraft, the deck motion prediction technology is used to forecast the motion parameters of the ship in the effective time in the future. A six-degree-of-freedom deck motion model based on sine wave combination is established, and a deck motion prediction model is established based on a back-propagation (BP) artificial neural network. A deck motion compensation model is developed in accordance with the advance network, and the landing guidance system is integrated to guarantee that the vector of the deck and the carrier aircraft are in synchronization. The aircraft motion model is established, and the feasibility of the deck motion prediction and compensation technology is verified through the simulation experiments.

Keywords: automatic carrier landing; deck motion prediction; deck motion compensation; BP neural network; advance network

Received: 2022-08-29; **Accepted**: 2023-02-07; **Published Online**: 2023-03-30 14:48
URL: link.cnki.net/urlid/11.2625.V.20230330.1044.001
* **Corresponding author**. E-mail: 250772665@qq.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0751

基于联合仿真的曲面共形多孔结构拓扑优化方法

付君健[1,2,3]，徐勇[2]，周祥曼[1,2,*]，舒正涛[2]，田启华[1,2]

(1. 水电机械设备设计与维护湖北省重点实验室，宜昌 443002；2. 三峡大学 机械与动力学院，宜昌 443002；
3. 智能制造装备与技术全国重点实验室，武汉 430074)

摘　　要：针对宏观曲面结构上多孔结构单胞构型与设计域外形不匹配，导致曲面结构优化困难的问题，提出一种基于联合仿真的曲面共形多孔结构拓扑优化方法。基于三周期极小曲面隐式水平集函数，实现多孔结构的参数化建模。采用等参单元法理论，建立自然坐标系到笛卡儿坐标系的映射关系，实现曲面多孔结构的共形建模。构造线性插值函数，保证多孔结构单胞之间的 C^0 连续。引入可变切割水平集函数，建立 MATLAB 与 ANSYS 联合仿真的多孔结构拓扑优化框架。数值算例表明：所提方法可有效实现曲面多孔结构的拓扑优化设计，确保多孔结构单胞与设计域外形的匹配，提升多孔结构的力学性能。

关 键 词：拓扑优化；多孔结构；共形建模；可变切割水平集函数；联合仿真
中图分类号：TH112
文献标志码：A　　**文章编号**：1001-5965(2024)09-2781-10

由多孔结构单胞填充而成的宏观结构是一种几何构形复杂、含有大量孔洞的结构，其具有轻量化、高承载、高比刚度等优点[1-2]，广泛应用于航空航天、船舶、医疗器械等领域[3]。然而，实际工程中的机械结构具有复杂的曲面几何特征，若将形状规则的多孔结构单胞填充到曲面上，易导致结构单胞与设计域外形不匹配，从而影响多孔结构的力学性能。因此，为保证曲面多孔结构外形的完整性，提升曲面多孔结构的力学性能，有必要研究曲面多孔结构的拓扑优化设计方法。

针对曲面多孔结构的设计问题，大部分研究都是基于三周期极小曲面和空间单元映射法[4-5]，实现对曲面多孔结构的建模设计，并采用有限元中六面体单元对实体模型进行网格划分，基于等参单元法理论，利用等参变换将三周期极小曲面映射到每个六面体单元当中，从而构建出基于三周期极小曲面的曲面多孔结构。此外，基于距离场和三周期极小曲面的多孔结构设计方法[6-7]也是目前常用的一种方法，该方法通过计算边界点的距离场，并与三周期极小曲面的距离场进行布尔运算，可以设计出内部结构复杂、外表面与实体结构相匹配的多孔结构，同时保证了结构单胞之间的良好连续性。上述2类方法实现了曲面多孔结构的建模设计，但未考虑到多孔结构的优化设计，无法实现多孔结构单胞在设计域空间中的最优分布，难以使结构的力学性能达到最优。因此，需进一步研究曲面多孔结构优化设计方法，提升多孔结构的力学性能，以满足实际工程的需求。

拓扑优化是指在给定的设计空间内，在已知的载荷和边界条件下，对结构材料进行最优分布，使结构的某种力学性能达到最优[8]。采用拓扑优化方法对多孔结构进行优化设计，相较于均匀多孔结构，优化设计后的多孔结构在空间中呈梯度分布，在结构抗承载能力[9-10]、抗屈曲能力[11]和吸能特性[12]

收稿日期：2022-08-31；**录用日期**：2023-01-14；**网络出版时间**：2023-02-28 10:54
网络出版地址：link.cnki.net/urlid/11.2625.V.20230227.1919.006
基金项目：国家自然科学基金 (51775308)；湖北省自然科学基金 (2021CFB236)；智能制造装备与技术全国重点实验室开放基金 (IMETKF2023016)
* **通信作者**. E-mail：zhouxman@ctgu.edu.cn

引用格式：*付君健，徐勇，周祥曼，等. 基于联合仿真的曲面共形多孔结构拓扑优化方法 [J]. 北京航空航天大学学报，2024，50（9）：2781-2790. FU J J，XU Y，ZHOU X M，et al. Topological optimization method for conformal cellular structures on surfaces based on co-simulation [J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50（9）：2781-2790（in Chinese）.*

等方面有着明显的优势。目前，对于多孔结构优化的研究大多都是面向二维平面/三维欧几里德空间中形状规则的设计域[13-14]，鲜少研究曲面上多孔结构的优化问题。针对曲面上多孔结构的优化设计研究，有学者基于水平集方法[15]与共形映射实现对自由曲面的拓扑优化[16-18]，并利用保角映射理论，将三维自由曲面映射到规则的二维矩形域上，通过降低问题的维数，将三维空间中自由曲面上的拓扑优化问题重新表达为欧几里德空间中的二维拓扑优化问题，从而实现了自由曲面结构的优化设计。此外，还有学者提出基于移动可变形组件方法和计算保角映射的复杂曲面拓扑优化方法[19]，该方法基于二维移动可变形组件框架，利用保角映射技术，将三角形网格表示的复杂曲面在数值上映射到规则的二维参数域，该方法可以在较少的设计变量和有限元分析自由度的情况下实现对复杂曲面的拓扑优化设计。上述方法实现了对曲面多孔结构的优化设计，但因其优化设计与共形映射的理论门槛较高，与现有工程软件结合有一定的困难，因此，研究一种简单易用的曲面多孔结构优化方法对于实际工程具有重要意义。

为提升曲面上多孔结构的力学性能，本文提出基于联合仿真的曲面共形多孔结构拓扑优化方法，利用共形建模技术，建立共形多孔结构几何模型，引入可变切割水平集函数[20]，建立MATLAB与ANSYS软件联合的拓扑优化框架，实现多孔结构在曲面上的最优分布，并通过数值案例验证了本文方法的可行性与有效性。

1 共形多孔结构参数化建模

1.1 多孔结构参数化建模

使用水平集函数对多孔结构进行隐式建模是常用的一种建模方法，其基本原理是将曲面隐式地用高一维函数 $\Phi(\boldsymbol{X})$ 的等值面去表达，通过水平集函数的零等值面将设计域分为实体、边界和空洞3部分。水平集函数隐式地描述结构几何形状 Ω 的定义如下[21]：

$$\begin{cases}\Phi(\boldsymbol{X})>0 & \boldsymbol{X}\in\Omega\\ \Phi(\boldsymbol{X})=0 & \boldsymbol{X}\in\partial_\Omega\\ \Phi(\boldsymbol{X})<0 & \boldsymbol{X}\in D/\Omega\end{cases} \tag{1}$$

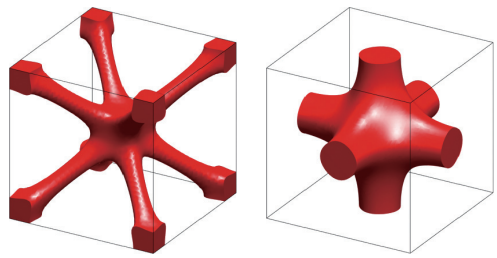
式中：$\boldsymbol{X}$ 为空间物理坐标；∂_Ω 为结构几何边界；D 为包含实体区域和几何边界的空间。本文仅考虑三维多孔结构，对于三维的结构边界，采用四维的水平集函数描述。

三周期极小曲面是一种由隐式水平集函数定义的结构[22]，具有良好的可设计性。本文采用三周期极小曲面作为多孔结构代表体积单胞，在其隐式函数中引入水平控制参数 c，控制单胞的几何形状和体积分数。I-Wrapped Package(I-WP)和Primitive (P)型单胞的四维隐式水平集函数分别为

$$\begin{aligned}\Phi_{\text{I-WP}}=&2(\cos(2\pi x)\cos(2\pi y)+\\&\cos(2\pi y)\cos(2\pi z)+\cos(2\pi z)\cos(2\pi x))-\\&(\cos(4\pi x)+\cos(4\pi y)+\cos(4\pi z))-c\end{aligned} \tag{2}$$

$$\Phi_{\text{P}}=\cos(2\pi x)+\cos(2\pi y)+\cos(2\pi z)-c \tag{3}$$

式中：x、y、z 为三维空间物理坐标矩阵。根据水平集函数定义，得到三周期极小曲面的实体结构如图1所示。



(a) I-WP型 (c=13.5)　　(b) P型 (c=13.5)

图1　三周期极小曲面多孔结构

Fig. 1　Triply periodic minimal surfaces based cellular structures

1.2 多孔结构共形建模

形状规则的单胞无法准确填充到复杂的曲面结构中，为实现曲面多孔结构的单胞域形状随形变化，本文提出共形建模技术。先基于等参单元法，利用六面体单元形函数对形状规则的单胞进行等参变换，将自然坐标系下形状规则的单胞转化成笛卡儿坐标系下形状不规则的单胞，采用的六面体单元形函数矩阵[23]如下：

$$\boldsymbol{N}_i=\frac{1}{8}(\boldsymbol{I}+\boldsymbol{\gamma}\gamma_i)(\boldsymbol{I}+\boldsymbol{\eta}\eta_i)(\boldsymbol{I}+\boldsymbol{\tau}\tau_i)\qquad i=1,2,\cdots,8 \tag{4}$$

式中：$\boldsymbol{I}$ 为单位矩阵；i 为单元节点序号；γ_i、η_i、τ_i 为自然坐标系下单元节点的坐标；$\boldsymbol{\gamma}$、$\boldsymbol{\eta}$、$\boldsymbol{\tau}$ 为自然坐标系下的单元内部点的坐标矩阵。

图2为P型单胞的共形建模，其几何的映射关系如下：

$$\begin{bmatrix}\boldsymbol{x}_{\text{d}}\\ \boldsymbol{y}_{\text{d}}\\ \boldsymbol{z}_{\text{d}}\end{bmatrix}=\sum_i^8\boldsymbol{N}_i\begin{bmatrix}x_i\\ y_i\\ z_i\end{bmatrix}\qquad i=1,2,\cdots,8 \tag{5}$$

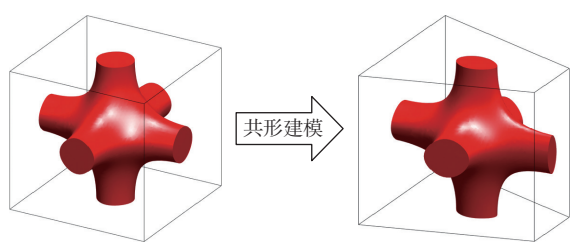


图2　P型单胞共形建模

Fig. 2　Conformal modeling of P unit cell

式中: $\boldsymbol{x}_d$、$\boldsymbol{y}_d$、$\boldsymbol{z}_d$ 为笛卡儿坐标系下的单元内部点坐标矩阵; x_i、y_i、z_i 为笛卡尔坐标系下的单元节点坐标。

复杂曲面共形多孔结构的隐式建模, 关键在于计算多孔结构有限元模型的六面体单元节点坐标。利用 CAD 软件建立复杂曲面结构几何模型。基于有限元软件采用六面体单元对曲面结构进行网格划分, 利用有限元软件相关工具和功能模块, 提取有限元模型上六面体单元节点坐标及单元节点编号。以标准的三周期极小曲面单胞作为结构代表性单胞, 利用共形技术将单胞映射到结构有限元模型的六面体单元上, 以建立复杂曲面多孔结构隐式模型, 图 3 为利用共形技术实现复杂曲面多孔结构建模过程。

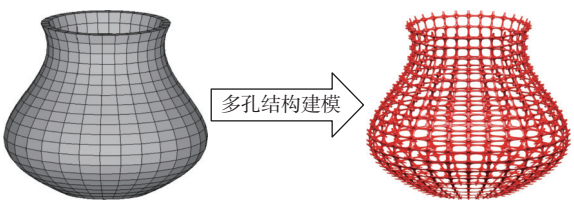


图 3　多孔结构建模过程

Fig. 3　Modeling of cellular structures

1.3　共形多孔结构 C^0 连续建模

在多孔结构设计过程中, 单胞几何边界的连续性对于结构的性能有着至关重要的作用, 可有效避免结构受力时的应力集中现象, 同时也保证了结构的可制造性[24-25]。为保证多孔结构单胞之间的几何连续性, 本文建立 C^0 连续的局部插值模型, 将结构单胞的水平控制参数矩阵 $\boldsymbol{c}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ 定义为三维空间坐标 $\boldsymbol{c}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ [26], 通过在相邻单胞的公共节点上定义一对相同的设计变量 $h(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$, 利用形函数对设计变量进行插值, 从而保证了相邻多孔结构单胞接触面处 $\boldsymbol{c}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ 相等, 实现了在不同体积分数下的多孔结构 C^0 连续。

建立 C^0 连续的局部插值模型, 其核心思想是保证相邻多孔结构单胞接触面处 $\boldsymbol{c}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ 相等。采用六面体单元将宏观设计域划分成 M 个单胞, 共有 m 个节点。每个单胞进一步划分为若干八节点六面体单元, 图 4 为局部插值模型的基本原理。

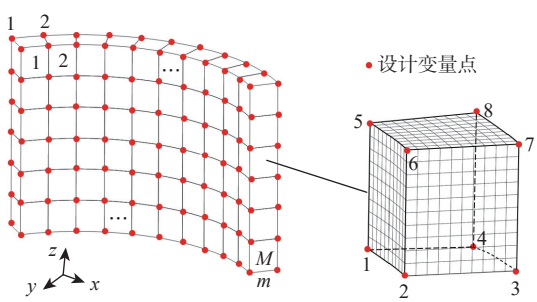


图 4　局部插值模型

Fig. 4　Local interpolation model

采用形函数式 (4) 对任意单胞 C 节点上的设计变量进行线性插值, 得到单胞 C 上所有节点的设计变量矩阵 $\boldsymbol{c}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ 为

$$\begin{aligned}\boldsymbol{c}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) = \frac{\boldsymbol{I}}{8}[&(\boldsymbol{I}+\boldsymbol{\gamma}\gamma_1)(\boldsymbol{I}+\boldsymbol{\eta}\eta_1)(\boldsymbol{I}+\boldsymbol{\tau}\tau_1)h_1+\\ &(\boldsymbol{I}+\boldsymbol{\gamma}\gamma_2)(\boldsymbol{I}+\boldsymbol{\eta}\eta_2)(\boldsymbol{I}+\boldsymbol{\tau}\tau_2)h_2+\\ &(\boldsymbol{I}+\boldsymbol{\gamma}\gamma_3)(\boldsymbol{I}+\boldsymbol{\eta}\eta_3)(\boldsymbol{I}+\boldsymbol{\tau}\tau_3)h_3+\\ &(\boldsymbol{I}+\boldsymbol{\gamma}\gamma_4)(\boldsymbol{I}+\boldsymbol{\eta}\eta_4)(\boldsymbol{I}+\boldsymbol{\tau}\tau_4)h_4+\\ &(\boldsymbol{I}+\boldsymbol{\gamma}\gamma_5)(\boldsymbol{I}+\boldsymbol{\eta}\eta_5)(\boldsymbol{I}+\boldsymbol{\tau}\tau_5)h_5+\\ &(\boldsymbol{I}+\boldsymbol{\gamma}\gamma_6)(\boldsymbol{I}+\boldsymbol{\eta}\eta_6)(\boldsymbol{I}+\boldsymbol{\tau}\tau_6)h_6+\\ &(\boldsymbol{I}+\boldsymbol{\gamma}\gamma_7)(\boldsymbol{I}+\boldsymbol{\eta}\eta_7)(\boldsymbol{I}+\boldsymbol{\tau}\tau_7)h_7+\\ &(\boldsymbol{I}+\boldsymbol{\gamma}\gamma_8)(\boldsymbol{I}+\boldsymbol{\eta}\eta_8)(\boldsymbol{I}+\boldsymbol{\tau}\tau_8)h_8]\end{aligned} \tag{6}$$

式中: h_1, h_2, …, h_8 为单元 8 个节点的设计变量。将 $\boldsymbol{c}(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ 代入结构单胞的隐式水平集函数中, 得到 C^0 连续多孔结构。

图 5 为未经插值的梯度多孔结构, 其相邻单胞之间存在几何突变。图 6 为插值后的梯度多孔结构, 其相邻单胞之间具有良好的几何连续性。



图 5　几何突变的梯度多孔结构

Fig. 5　Graded cellular structures with abrupt geometric features

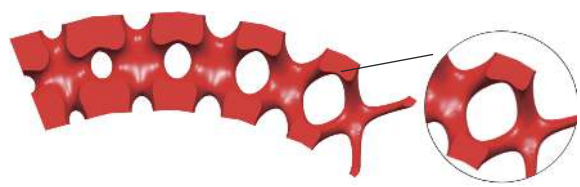


图 6　几何连续的梯度多孔结构

Fig. 6　Graded cellular structures with geometric continuity

2　可变切割水平集函数构造

为实现对复杂曲面上多孔结构的优化设计, 基于结构几何模型, 本文引入可变切割水平集函数[20,27]。如图 4 所示, 先在整个宏观结构的单胞节点上定义一组与时间相关的全局设计变量:

$$\boldsymbol{H}(t) = [h_1(t) \quad h_2(t) \quad \cdots \quad h_m(t)]^{\mathrm{T}} \tag{7}$$

式中: $\boldsymbol{H}(t)$ 为每个单胞节点上的全局设计变量; m 为宏观结构的单胞节点数, 任意一个编号为 C 的结构单胞, 均包含 8 个节点, 则其单胞节点的设计变量 $\boldsymbol{h}_C(t)$ 表示为

$$\boldsymbol{h}_C(t) = \left[h_C^1(t) \quad h_C^2(t) \quad \cdots \quad h_C^8(t)\right] = \boldsymbol{S}_C\boldsymbol{H}(t) \tag{8}$$

式中: $\boldsymbol{S}_C$ 为选择矩阵, 实现从全局设计变量中提取

单胞 C 的全局设计变量。

使用形函数对单胞 C 的全局设计变量 $\boldsymbol{h}_C(t)$ 进行插值，形函数向量 $\boldsymbol{N}_C(\boldsymbol{x})$ 定义为

$$\boldsymbol{N}_C(\boldsymbol{x})=\begin{bmatrix}N_1^C(\boldsymbol{x}) & N_2^C(\boldsymbol{x}) & \cdots & N_8^C(\boldsymbol{x})\end{bmatrix} \tag{9}$$

单胞 C 的切割函数 $C_C(\boldsymbol{x},t)$ 可表示为

$$C_C(\boldsymbol{x},t)=\boldsymbol{N}_C^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{h}_C(t)=\boldsymbol{N}_C^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{S}_C\boldsymbol{H}(t) \tag{10}$$

通过将所有单胞的切割函数 $C_C(\boldsymbol{x},t)$ 组合在一起，形成具有 C^0 连续性的全局切割函数 $C_{\mathrm{g}}(\boldsymbol{x},t)$，结合预定义的全局水平集函数 $\boldsymbol{\Phi}_{\mathrm{g}}(\boldsymbol{x})$，可得多孔结构隐式水平集函数：

$$\begin{cases}\boldsymbol{\Phi}(\boldsymbol{x},t)=\boldsymbol{\Phi}_{\mathrm{g}}(\boldsymbol{x})+C_{\mathrm{g}}(\boldsymbol{x},t)>0 & \forall\boldsymbol{x}\in\Omega/\partial_\Omega\\ \boldsymbol{\Phi}(\boldsymbol{x},t)=\boldsymbol{\Phi}_{\mathrm{g}}(\boldsymbol{x})+C_{\mathrm{g}}(\boldsymbol{x},t)=0 & \forall\boldsymbol{x}\in\partial_\Omega\\ \boldsymbol{\Phi}(\boldsymbol{x},t)=\boldsymbol{\Phi}_{\mathrm{g}}(\boldsymbol{x})+C_{\mathrm{g}}(\boldsymbol{x},t)<0 & \forall\boldsymbol{x}\in D/\Omega\end{cases} \tag{11}$$

式中：$C_{\mathrm{g}}(\boldsymbol{x},t)$ 通过 $\boldsymbol{H}(t)$ 来调整，$\boldsymbol{H}(t)$ 为关于宏观时间 t 的变量，调整 $\boldsymbol{H}(t)$ 可以灵活地控制结构单胞的几何特征，同时改变单胞的体积分数。

3 拓扑优化数学建模与灵敏度分析

3.1 拓扑优化数学模型

本文仅研究复杂曲面上共形多孔结构刚度优化问题，基于可变切割水平集方法，建立以结构柔度最小化为目标函数，体积分数为约束条件的多孔结构拓扑优化数学模型：

$$\begin{cases}\text{find:}\quad \boldsymbol{H}(t)=[h_1(t)\quad h_2(t)\quad\cdots\quad h_m(t)]^{\mathrm{T}}\\ \min:\quad J(\boldsymbol{\Phi})=\dfrac{1}{2}\displaystyle\int_\Omega\boldsymbol{\varepsilon}^{\mathrm{T}}(\boldsymbol{u})\boldsymbol{E}\boldsymbol{\varepsilon}(\boldsymbol{u})\mathrm{d}\Omega\\ \text{s.t.}\begin{cases}a_{\boldsymbol{\Phi}}(\boldsymbol{u},\boldsymbol{v})=l_{\boldsymbol{\Phi}}(\boldsymbol{v}) & \forall\boldsymbol{v}\in U\\ g(\boldsymbol{\Phi})=\int_\Omega\mathrm{d}\Omega-\mu\overline{V}\leqslant 0\\ h_{\min}\leqslant h_k(t)\leqslant h_{\max}\end{cases}\end{cases} \tag{12}$$

式中：m 为全局设计变量的数量；$J(\boldsymbol{\Phi})$ 为结构柔度函数；$\boldsymbol{\varepsilon}$ 为结构应变场；$\boldsymbol{u}$ 为结构位移场；$\boldsymbol{E}$ 为材料弹性矩阵；$a_{\boldsymbol{\Phi}}(\boldsymbol{u},\boldsymbol{v})=l_{\boldsymbol{\Phi}}(\boldsymbol{v})$ 为弹性平衡方程的弱形式；$\boldsymbol{v}$ 和 U 分别为虚拟位移域和动力学上允许的位移域空间；$g(\boldsymbol{\Phi})$ 为结构的体积约束函数；μ 为结构给定的体积分数大小；$\overline{V}$ 为设计域的体积大小；$h_{\min}$ 和 $h_{\max}$ 分别为设计变量的上、下界限；h_k 为设计域中第 k 个节点的全局设计变量。

3.2 灵敏度分析

本文通过求解哈密顿-雅可比偏微分方程来实现结构拓扑形状的演化。根据形状导数[28]的定义及其引理，推导目标函数和约束条件关于设计变量的灵敏度，其目标函数、能量双线性形式、载荷线性形式关于时间变量 t 的微分形式分别为

$$\frac{\partial J(\boldsymbol{\Phi})}{\partial t}=\int_\Omega\boldsymbol{\varepsilon}^{\mathrm{T}}(\dot{\boldsymbol{u}})\boldsymbol{E}\boldsymbol{\varepsilon}(\boldsymbol{u})\mathrm{d}\Omega+\frac{1}{2}\int_\Omega\boldsymbol{\varepsilon}^{\mathrm{T}}(\boldsymbol{u})\boldsymbol{E}\boldsymbol{\varepsilon}(\boldsymbol{u})\delta(\boldsymbol{\Phi})\frac{\partial\boldsymbol{\Phi}}{\partial t}\mathrm{d}\Omega \tag{13}$$

$$\frac{\partial a_{\boldsymbol{\Phi}}(\boldsymbol{u},\boldsymbol{v})}{\partial t}=\int_\Omega\boldsymbol{\varepsilon}^{\mathrm{T}}(\dot{\boldsymbol{u}})\boldsymbol{E}\boldsymbol{\varepsilon}(\boldsymbol{v})\mathrm{d}\Omega+\int_\Omega\boldsymbol{\varepsilon}^{\mathrm{T}}(\boldsymbol{u})\boldsymbol{E}\boldsymbol{\varepsilon}(\dot{\boldsymbol{v}})\mathrm{d}\Omega+\int_\Omega\boldsymbol{\varepsilon}^{\mathrm{T}}(\boldsymbol{u})\boldsymbol{E}\boldsymbol{\varepsilon}(\boldsymbol{v})\delta(\boldsymbol{\Phi})\frac{\partial\boldsymbol{\Phi}}{\partial t}\mathrm{d}\Omega \tag{14}$$

$$\frac{\partial l_{\boldsymbol{\Phi}}(\boldsymbol{v})}{\partial t}=\int_\Omega(f\dot{\boldsymbol{v}}+\mathrm{div}(p\dot{\boldsymbol{v}}\boldsymbol{n}))\mathrm{d}\Omega+(f\boldsymbol{v}+\mathrm{div}(p\boldsymbol{v}\boldsymbol{n}))\delta(\boldsymbol{\Phi})\frac{\partial\boldsymbol{\Phi}}{\partial t}\mathrm{d}\Omega \tag{15}$$

式中：f 为结构体积力；p 为边界上的牵引力；$\boldsymbol{n}$ 为法向矢量；div 为散度算子，可表示为

$$\mathrm{div}(p\boldsymbol{v}\boldsymbol{n})=\nabla(p\boldsymbol{v})\frac{\nabla\boldsymbol{\Phi}}{|\nabla\boldsymbol{\Phi}|}+\left(\nabla\frac{\nabla\boldsymbol{\Phi}}{|\nabla\boldsymbol{\Phi}|}\right)p\boldsymbol{v} \tag{16}$$

由于 $\dot{\boldsymbol{v}}\in U$，可得到共轭方程：

$$\int_\Omega\boldsymbol{\varepsilon}^{\mathrm{T}}(\boldsymbol{u})\boldsymbol{E}\boldsymbol{\varepsilon}(\dot{\boldsymbol{v}})\mathrm{d}\Omega=\int_\Omega(f\dot{\boldsymbol{v}}+\mathrm{div}(p\dot{\boldsymbol{v}}\boldsymbol{n}))\mathrm{d}\Omega \tag{17}$$

弹性平衡方程两边对时间 t 求偏导：

$$\frac{\partial a_{\boldsymbol{\Phi}}(\boldsymbol{u},\boldsymbol{v})}{\partial t}=\frac{\partial l_{\boldsymbol{\Phi}}(\boldsymbol{v})}{\partial t} \tag{18}$$

将式 (14)、式 (15) 和式 (17) 代入式 (18) 得

$$\int_\Omega\boldsymbol{\varepsilon}^{\mathrm{T}}(\dot{\boldsymbol{u}})\boldsymbol{E}\boldsymbol{\varepsilon}(\boldsymbol{v})\mathrm{d}\Omega=\int_\Omega\big(f\boldsymbol{v}+\mathrm{div}(p\boldsymbol{v}\boldsymbol{n})\mathrm{d}\Omega-\boldsymbol{\varepsilon}^{\mathrm{T}}(\boldsymbol{u})\boldsymbol{E}\boldsymbol{\varepsilon}(\boldsymbol{v})\big)\delta(\boldsymbol{\Phi})|\nabla\boldsymbol{\Phi}|v_{\boldsymbol{n}}\mathrm{d}\Omega \tag{19}$$

式中：$v_{\boldsymbol{n}}$ 为法向速度。

因为结构刚度最大化为自伴随问题，可得

$$\int_\Omega\boldsymbol{\varepsilon}^{\mathrm{T}}(\dot{\boldsymbol{u}})\boldsymbol{E}\boldsymbol{\varepsilon}(\boldsymbol{u})\mathrm{d}\Omega=\int_\Omega\big(f\boldsymbol{u}+\mathrm{div}(p\boldsymbol{u}\boldsymbol{n})-\boldsymbol{\varepsilon}^{\mathrm{T}}(\boldsymbol{u})\boldsymbol{E}\boldsymbol{\varepsilon}(\boldsymbol{u})\big)\delta(\boldsymbol{\Phi})|\nabla\boldsymbol{\Phi}|v_{\boldsymbol{n}}\mathrm{d}\Omega \tag{20}$$

$$\mathrm{div}(p\boldsymbol{u}\boldsymbol{n})=\left(\nabla(p\boldsymbol{u})\cdot\frac{\nabla\boldsymbol{\Phi}}{|\nabla\boldsymbol{\Phi}|}+\left(\nabla\cdot\frac{\nabla\boldsymbol{\Phi}}{|\nabla\boldsymbol{\Phi}|}\right)p\boldsymbol{u}\right) \tag{21}$$

将式 (20) 代入式 (13) 得到目标函数 $J(\Omega)$ 的形状导数：

$$\frac{\partial J(\boldsymbol{\Phi})}{\partial t}=\int_\Omega G(\boldsymbol{\Phi})\delta(\boldsymbol{\Phi})|\nabla\boldsymbol{\Phi}|v_{\boldsymbol{n}}\mathrm{d}\Omega \tag{22}$$

$$G(\boldsymbol{\Phi})=f\boldsymbol{u}+\mathrm{div}(p\boldsymbol{u}\boldsymbol{n})-\frac{1}{2}\boldsymbol{\varepsilon}^{\mathrm{T}}(\boldsymbol{u})\boldsymbol{E}\boldsymbol{\varepsilon}(\boldsymbol{u}) \tag{23}$$

同理，可求得体积约束的形状导数：

$$\frac{\partial g(\boldsymbol{\Phi})}{\partial t}=\int_\Omega\delta(\boldsymbol{\Phi})|\nabla\boldsymbol{\Phi}|v_{\boldsymbol{n}}\mathrm{d}\Omega \tag{24}$$

本文以 $\boldsymbol{H}(t)$ 为设计变量，构造可变切割水平集函数对哈密顿-雅可比偏微分方程进行时空解耦，从而将水平集函数参数化，得到关于时间 t 的常微分方程，其数学表达式为

$$|\nabla\boldsymbol{\Phi}|v_{\boldsymbol{n}}+\frac{\partial\boldsymbol{H}(t)}{\partial t}\boldsymbol{N}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{S}=0 \tag{25}$$

得到水平集函数边界演化的法向速度：

$$v_{\boldsymbol{n}}=-\frac{1}{|\nabla\boldsymbol{\Phi}|}\cdot\frac{\partial\boldsymbol{H}(t)}{\partial t}\boldsymbol{N}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{S} \tag{26}$$

将式 (26) 代入式 (22) 得

$$\frac{\partial J(\boldsymbol{\Phi})}{\partial t}=-\sum_{k=1}^{m}\int_{\Omega}G(\boldsymbol{\Phi})\frac{\partial H_k(t)}{\partial t}\boldsymbol{N}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{S}\delta(\boldsymbol{\Phi})\mathrm{d}\Omega \quad (27)$$

利用链式法则求解目标函数 $J(\boldsymbol{\Phi})$ 关于时间变量 t 的导数可得

$$\frac{\partial J(\boldsymbol{\Phi})}{\partial t}=\sum_{k=1}^{m}\frac{\partial J(\boldsymbol{\Phi})}{\partial H_k(t)}\cdot\frac{\partial H_k(t)}{\partial t} \quad (28)$$

比较式 (27) 和式 (28) 可以得到目标函数 $J(\boldsymbol{\Phi})$ 关于设计变量 $\boldsymbol{H}(t)$ 的灵敏度:

$$\frac{\partial J(\boldsymbol{\Phi})}{\partial H_k(t)}=-\int_{\Omega}G(\boldsymbol{\Phi})\boldsymbol{N}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{S}\delta(\boldsymbol{\Phi})\mathrm{d}\Omega \quad (29)$$

同理, 将式 (26) 代入式 (24) 得

$$\frac{\partial g(\boldsymbol{\Phi})}{\partial t}=\sum_{k=1}^{m}\int_{\Omega}\frac{\partial H_k(t)}{\partial(t)}\boldsymbol{N}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{S}\delta(\boldsymbol{\Phi})\mathrm{d}\Omega \quad (30)$$

利用链式法则求解体积约束 $g(\boldsymbol{\Phi})$ 关于时间变量 t 的导数可得

$$\frac{\partial g(\boldsymbol{\Phi})}{\partial t}=\sum_{k=1}^{m}\frac{\partial g(\boldsymbol{\Phi})}{\partial H_k(t)}\cdot\frac{\partial H_k(t)}{\partial t} \quad (31)$$

比较式 (30) 和式 (31) 可以得到体积约束 $g(\boldsymbol{\Phi})$ 关于设计变量 $\boldsymbol{H}(t)$ 的灵敏度:

$$\frac{\partial g(\boldsymbol{\Phi})}{\partial H_k(t)}=\int_{\Omega}\boldsymbol{N}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{S}\delta(\boldsymbol{\Phi})\mathrm{d}\Omega \quad (32)$$

4　基于联合仿真的拓扑优化

针对复杂曲面结构在 MATLAB 中存在的几何模型建立困难、边界条件添加困难和单元网格划分困难等问题, 本文建立 MATLAB 和 ANSYS 联合仿真, 采用 MATLAB 语言编写拓扑优化算法[29], 利用 ANSYS 软件提供的参数化设计语言和相关工具[30], 重复调用有限元分析及其结果数据, 从而实现对曲面多孔结构的优化设计。

MATLAB 和 ANSYS 联合仿真程序流程如图 7 所示, 主要包括有限元分析和拓扑优化 2 部分。其中, 有限元分析在 ANSYS 软件中进行, 包括建立几何模型、定义单元类型、定义材料属性、施加边界条件和求解有限元模型; 拓扑优化部分基于 MATLAB 软件编写优化算法, 主要包括定义参数化水平集函数、调用有限元分析数据、计算灵敏度信息和更新设计变量。拓扑优化过程中, MATLAB 与 ANSYS 需要进行数据交互, 并重复调用 ANSYS 中有限元模型的单元节点编号与坐标、单元应变能等数据。最后, 在 MATLAB 中编写结果显示函数, 显示优化结果。图 7 中, OC 算法表示 Optimality Criteria 算法。

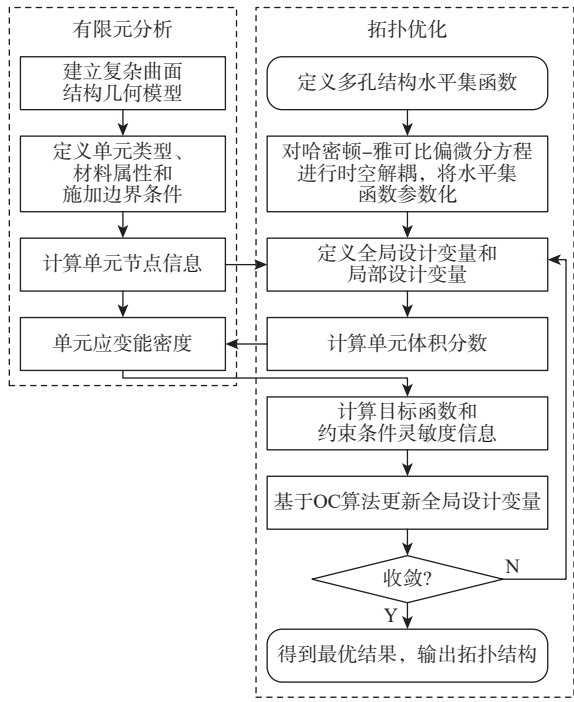


图 7　联合仿真程序流程

Fig. 7　Flow of co-simulation program

5　数值案例及结果分析

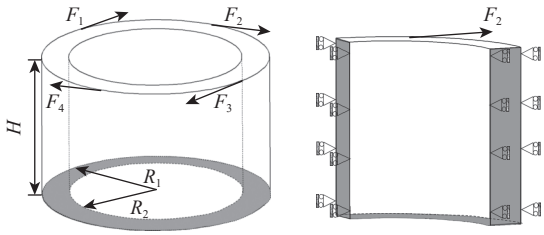
为验证本文方法的有效性, 以带有曲面特征的宏观结构为例进行优化分析。假定构成实体材料的弹性模量为 1、泊松比为 0.3, 以结构柔度最小化为目标建立优化数学模型, 采用 OC 算法作为优化求解算法[31]。

多孔结构的性能和单胞尺度相关, 在相同的设计域和边界条件下, 改变设计域中单胞的数量和离散单元的数量, 优化后的结构柔度是不同的, 因此, 优化后结构的性能也会不同。对于多孔结构, 其结构的优化性能存在着网格依赖性, 当单胞的数量增多时, 其结构最终优化的柔度也会增大; 持续增加单胞的数量, 多孔结构的性能将保持不变[32], 本文对多孔结构的优化都是在给定单胞和单元数量的情况下对结构进行优化。

5.1　圆筒结构优化设计

本节算例选取 I-WP 型单胞作为结构代表性体积单胞, 如图 8 所示。图 8(a) 为圆筒结构设计域, 设计域内、外径和高分别为 R_1=50 mm、R_2=40 mm 和 H=64 mm。设计域的下端面施加固定约束, 上端面沿圆筒面 1/4 位置切线方向施加集中载荷 F_1=F_2=F_3=F_4=5 N。图 9 为圆筒结构有限元网格划分, 图 9(a) 为采用八节点六面体单元对圆筒结构进行网格划分, 将设计域划分为 40×2×8 个单胞, 再将每个单胞进一步划分为 20×20×20 个单元。为简化计算过程, 仅对总设计域的 1/4 部分开展拓扑优化设计, 结

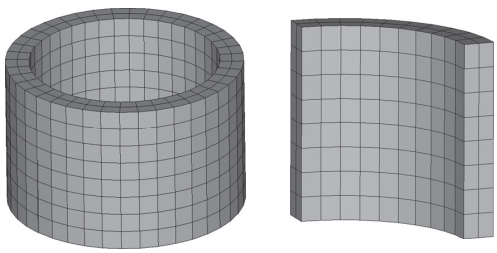
构设计域两端施加对称约束，如图 8(b) 所示为 1/4 圆筒结构设计域，其网格划分如图 9(b) 所示。优化过程中，设计域的体积分数约束为 0.35。



(a) 完整设计域 (b) 1/4圆筒设计域

图 8 圆筒结构设计域

Fig. 8 Design domain of cylinder structure



(a) 完整结构 (b) 1/4圆筒结构

图 9 圆筒结构有限元网格划分

Fig. 9 Finite element meshing of cylinder structure

优化迭代曲线如图 10 所示，结构初始柔度为 3 177.1，随着优化的进行，结构柔度下降速变快，结构轮廓变化明显，之后趋于稳定。经过 60 次迭代后结构柔度收敛至 670.3，体积分数收敛至 0.35，结构刚度提升了约 78.9%。图 11 为圆筒结构最优拓扑构型，图 11(a) 为 1/4 圆筒结构的最优拓扑构型，优化后的结构轮廓清晰，结构在几何上呈现明显的梯度分布。在结构的主要传力路径上结构单胞的体积大，受力小的地方结构单胞的体积小。优化过程中，通过设置每个单胞的体积分数都大于 0，使结构设计域内每个单胞都分布材料，从而提高了多孔结构的稳定性。此外，在优化过程中，通过在相邻的单胞的公共节点上定义相同的设计变量，利用线性插值函数，从而保证了结构单胞之间几何边界的连续性，使得优化后的梯度多孔结构具有良好的力学性能。图 11(b) 为完整结构的最优构型。

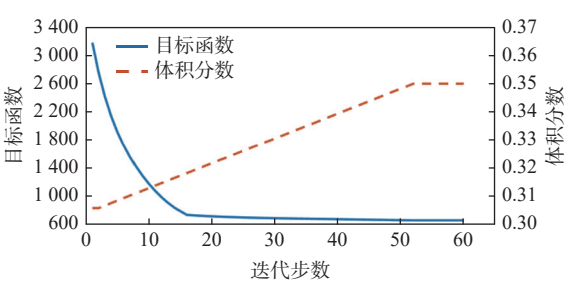
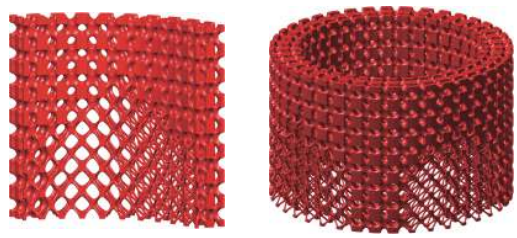


图 10 圆筒结构优化迭代曲线

Fig. 10 Optimize iterative curve of cylinder structure



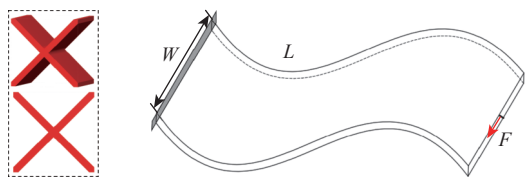
(a) 1/4圆筒结构 (b) 完整结构

图 11 圆筒结构最优拓扑构型

Fig. 11 The optimal configuration of cylinder structure

5.2 自由曲面结构优化设计

本节算例选取如图 12 所示的 X 型结构单胞作为结构代表性体积单胞。图 12(b) 为自由曲面结构设计域，设计域宽和弧长分别为 W=48 mm 和 L=96 mm，结构厚度为 3 mm。结构设计域左端面为固定约束，结构右端面中间位置沿宽度方向施加集中载荷 F=10 N。图 13 为采用八节点六面体单元对自由曲面结构进行网格划分，将设计域划分为 16×8×1 个单胞，再将每个单胞进一步划分为 20×20×20 个单元。优化过程中，设计域的体积分数约束为 0.5。



(a) X型单胞 (b) 结构设计域

图 12 预定义单胞及结构设计域

Fig. 12 Predefined unit cell and structural design domains

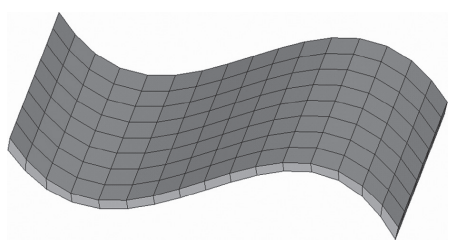


图 13 自由曲面结构有限元网格划分

Fig. 13 Finite element meshing of free-form surface structure

图 14 为自由曲面结构的优化迭代曲线，随着优化的进行，结构柔度下降速度变快，结构的材料分布变化明显。经过 60 次迭代后目标函数和体积分数曲面收敛，体积分数收敛至 0.50，柔度从最开始的 5 471.4 收敛至 2 099.9。相较于初始结构，优化后的结构刚度提升了约 61.6%。优化后的结构最优拓扑构型如图 15 所示，优化后的结构呈明显的梯度分布，结构材料主要分布在结构的传力路径上，受力小的结构材料分布较少。此外，优化后的结构不同体积分数的单胞之间具有良好的连续性，

可有效避免结构出现应力集中的现象, 提升了结构的力学性能。

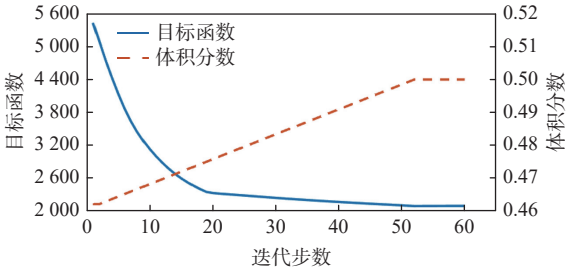
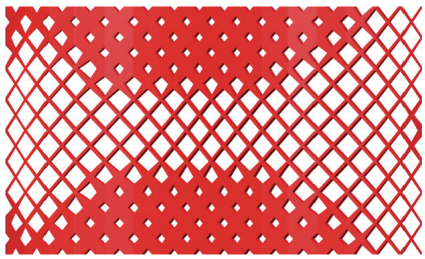
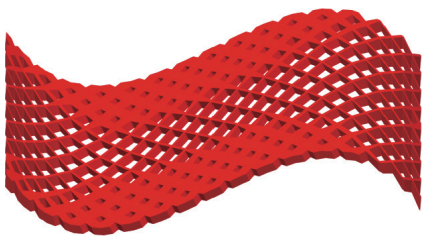


图 14　自由曲面结构优化迭代曲线

Fig. 14　Optimize iterative curve of free-form surface structure



(a) 正视图



(b) 轴测图

图 15　自由曲面结构最优拓扑构型

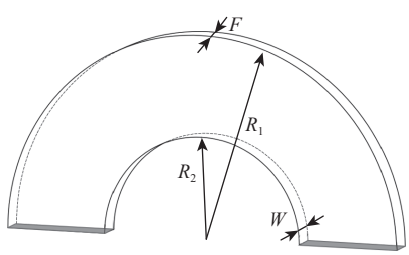
Fig. 15　Optimal configuration of free-form surface structure

5.3　曲面夹层结构优化设计

本节算例选取如图 16 所示, 基于 I-WP 型极小曲面构建的新型夹层结构单胞作为结构代表性体积单胞。图 16(b) 为曲面夹层结构设计域, 设计域内外径和宽分别为 R_1=50 mm、R_2=25 mm 和 W=8 mm。结构设计域底部左右两端面为固定约束, 结构上曲面中间处施加垂直曲面向下的集中载荷 F=20 N。图 17 为采用八节点六面体单元对结构进行网格划分, 将设计域划分为 20×4×1 个单胞, 每个单胞进一步划分为 20×20×20 个单元。优化过程中, 设计域



(a) 夹层单胞



(b) 结构设计域

图 16　预定义单胞及结构设计域

Fig. 16　Predefined unit cell and structural design domains

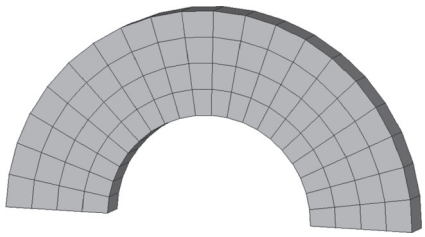


图 17　曲面夹层结构有限元网格划分

Fig. 17　Finite element meshing of curving sandwich structure

的体积分数约束为 0.45。

优化迭代曲线如图 18 所示, 初始结构其夹芯为均匀的多孔结构, 随着优化的进行, 结构柔度的逐渐减小、体积分数的逐渐增加, 结构逐渐呈梯度变化。迭代 60 次后, 结构体积分数收敛至 0.45, 结构的柔度从开始的 5 779.3 收敛至 2 699.3, 优化的结构刚度提升了约 53.3%。其优化后的最优拓扑构型如图 19 所示, 优化后的结构在几何上呈明显的梯度分布, 不同夹层的单胞形状和体积分数都不相同, 在结构受力的地方结构单胞体积分数较大, 其单胞夹芯也更加粗壮。优化后的结构材料主要分布在结构传力路径上, 这使得夹层结构具有了更加良好的力学性能。

为证明优化后的共形曲面夹层结构具有更加

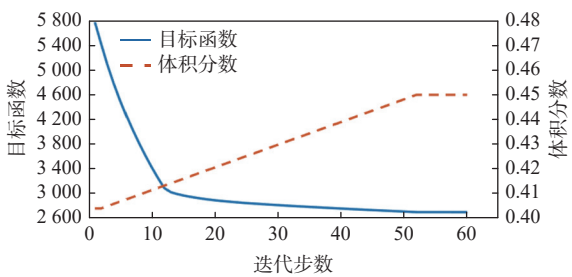
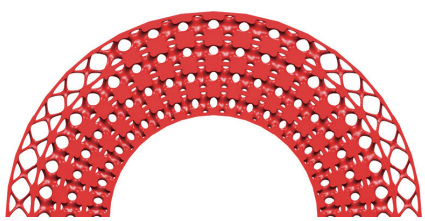


图 18　曲面夹层结构优化迭代曲线

Fig. 18　Optimize iterative curve of curving sandwich structure



(a) 正视图



(b) 轴测图

图 19　曲面夹层结构最优拓扑构型

Fig. 19　Optimal configuration of curving sandwich structure

良好的承载性能，在体积分数均为 0.45 的条件下，分别对优化后的共形曲面夹层结构、均匀的共形曲面夹层结构和未共形曲面夹层结构进行有限元静承载分析对比。图 20 为采用布尔运算方法得到的未共形曲面夹层结构。结构的材料均设置为 PA2200，其材料密度为 1 020 kg/m^3。结构的负载和支撑约束如图 16(b) 所示。图 21 为施加不同负载条件下 3 种结构的最大变形位移曲线，通过有限元分析结果可得，在相同负载情况下，共形曲面夹层结构相比于未共形的曲面夹层结构最大变形位移更小，其在承载性能上表现更加优异。此外，优化后的共形曲面夹层结构最大变形位移明显小于均匀的共形曲面夹层结构，其在承载性能上具有显著的优异性。根据负载-位移曲线分别计算 3 种结构的刚度，可知优化后的共形曲面夹层结构的刚度为 357.1 N/mm，均匀的共形曲面夹层结构的刚度为 200 N/mm，未共形的曲面夹层结构的刚度为 147.3 N/mm。相较于均匀的共形曲面夹层结构，优化后的共形曲面夹层结构的刚度提高了约 78.55%。此外，均匀的共形曲面夹层结构的刚度比未共形曲面夹层结构的刚度提高了约 35.78%。图 22 为在 60 N 负载条件下 3 种曲面夹层结构的最大变形位移云图对比。

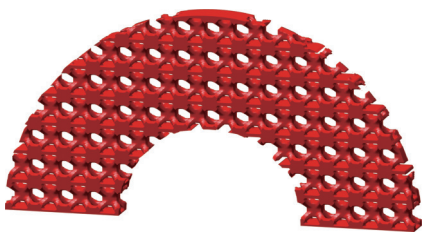


图 20 未共形的曲面夹层结构

Fig. 20 Non-conformal curving sandwich structure

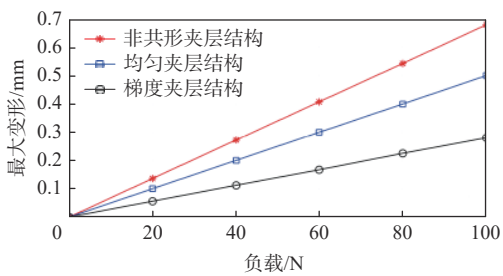


图 21 载荷-最大变形曲线

Fig. 21 Load-maximum deformation curve

有限元分析结果表明，利用共形技术不仅可以实现单胞与设计域外形匹配，避免单胞被割裂而导致夹层结构外形不完整，而且一定程度上提高了曲面夹层结构的力学性能；同时，优化后的曲面夹层结构相比均匀的曲面夹层结构在性能上也更加优异。

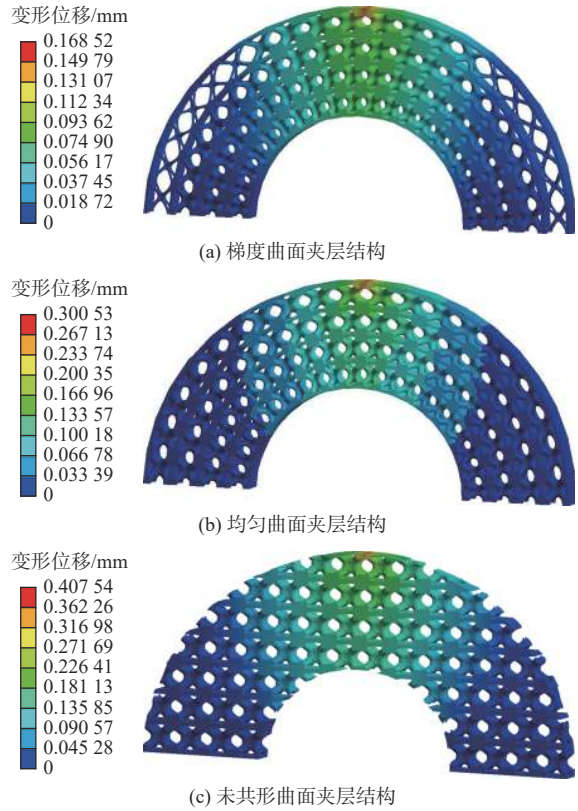


图 22 3 种结构最大变形位移云图

Fig. 22 Maximum deformation nephogram of three structures

6 结 论

本文提出一种基于联合仿真的曲面共形多孔结构拓扑优化方法。通过理论推导与数值算例分析，得出以下结论：

1）三周期极小曲面可实现多孔结构的隐式参数化建模，共形建模技术可实现多孔结构单胞与设计域外形的匹配，构造的局部插值模型，实现了共形多孔结构几何边界的 C^0 连续；

2）建立了基于 MATLAB 与 ANSYS 的联合仿真框架，通过引入可变切割水平集函数，实现了曲面多孔结构的优化设计；

3) 优化后的曲面多孔结构刚度性能比均布曲面多孔结构的性能更优，共形曲面多孔结构的力学性能优于未共形曲面多孔结构的力学性能。

参考文献（References）

[1] 王向明, 苏亚东, 吴斌, 等. 微桁架点阵结构在飞机结构/功能一体化中的应用[J]. 航空制造技术, 2018, 61(10): 16-25.
WANG X M, SU Y D, WU B, et al. Application for additive manufacturing of lattice materials on integrated aircraft structures and functions[J]. Aeronautical Manufacturing Technology, 2018, 61(10): 16-25(in Chinese).

[2] 邓昊宇, 王春洁. 三维点阵结构等效热分析与优化方法[J]. 北京航空航天大学学报, 2019, 45(6): 1122-1128.

DENG H Y, WANG C J. Equivalent thermal analysis and optimization method for three-dimensional lattice structure[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(6): 1122-1128(in Chinese).

[3] WANG B, YANG M S, ZHANG D Y, et al. Alternative approach for imperfection-tolerant design optimization of stiffened cylindrical shells via energy barrier method[J]. Thin-Walled Structures, 2022, 172: 108838.

[4] 张壮雅, 赵珂, 杨晓峰, 等. 基于三周期极小曲面和等参单元法的骨支架建模方法研究[J]. 机械设计与制造, 2017(11): 234-237.
ZHANG Z Y, ZHAO K, YANG X F, et al. Reseachon bone scaffold modeling method basedon triply periodic minimaland isoperimetric element method[J]. Machinery Design & Manufacture, 2017(11): 234-237(in Chinese).

[5] 王清辉, 夏刚, 徐志佳, 等. 面向组织工程的松质骨微观结构 TPMS 建模方法[J]. 计算机辅助设计与图形学学报, 2016, 28(11): 1949-1956.
WANG Q H, XIA G, XU Z J, et al. Modelling the microstructures of cancellous bone based on triply periodic minimal surface for tissue engineering[J]. Journal of Computer-Aided Design & Computer Graphics, 2016, 28(11): 1949-1956(in Chinese).

[6] YOO D J. Porous scaffold design using the distance field and triply periodic minimal surface models[J]. Biomaterials, 2011, 32(31): 7741-7754.

[7] 余学伟, 李琳慧, 张涛, 等. 基于三周期极小曲面和体素距离场的复杂轮廓点阵结构建模[J]. 重庆大学学报, 2022, 45(6): 75-84.
YU X W, LI L H, ZHANG T, et al. Modeling of complex contour lattice structure based on triply periodic minimal surface and voxel distance field[J]. Journal of Chongqing University, 2022, 45(6): 75-84(in Chinese).

[8] 任毅如, 向剑辉, 何杰, 等. 基于导重法的自重载荷下悬臂梁结构拓扑优化[J]. 北京航空航天大学学报, 2021, 47(7): 1338-1344.
REN Y R, XIANG J H, HE J, et al. Topology optimization of cantilever structure with self-weight load based on guide-weight method[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(7): 1338-1344(in Chinese).

[9] 付君健, 舒正涛, 田启华, 等. 功能梯度多孔结构拓扑优化的混合水平集方法[J]. 机械工程学报, 2022, 58(17): 249-260.
FU J J, SHU Z T, TIAN Q H, et al. A hybrid level set method for topology optimization of functionally graded cellular structures[J]. Journal of Mechanical Engineering, 2022, 58(17): 249-260(in Chinese).

[10] 孙鹏飞, 张跃, 尹鹏, 等. 隐式曲面梯度多孔结构拓扑优化设计方法[J]. 西安交通大学学报, 2022, 56(1): 85-95.
SUN P F, ZHANG Y, YIN P, et al. A topology optimization design method of graded cellular structures with implicit surfaces[J]. Journal of Xi'an Jiaotong University, 2022, 56(1): 85-95(in Chinese).

[11] NIKNAM H, AKBARZADEH A H, RODRIGUE D, et al. Architected multi-directional functionally graded cellular plates[J]. Materials & Design, 2018, 148: 188-202.

[12] QI D X, LU Q Y, HE C W, et al. Impact energy absorption of functionally graded chiral honeycomb structures[J]. Extreme Mechanics Letters, 2019, 32: 100568.

[13] 赵芳垒, 敬石开, 刘晨燕. 基于局部相对密度映射的变密度多孔结构设计方法[J]. 机械工程学报, 2018, 54(19): 121-128.
ZHAO F L, JING S K, LIU C Y. Variable density cellular structure design method base on local relative density mapping[J]. Journal of Mechanical Engineering, 2018, 54(19): 121-128(in Chinese).

[14] 蔡金虎, 王春洁. 基于映射的梯度点阵结构设计方法[J]. 振动与冲击, 2020, 39(20): 74-81.
CAI J H, WANG C J. A graded lattice structure design method based on mapping process[J]. Journal of Vibration and Shock, 2020, 39(20): 74-81(in Chinese).

[15] TOWNSEND S, KIM H A. A level set topology optimization method for the buckling of shell structures[J]. Structural and Multidisciplinary Optimization, 2019, 60(5): 1783-1800.

[16] VOGIATZIS P, MA M, CHEN S K, et al. Computational design and additive manufacturing of periodic conformal metasurfaces by synthesizing topology optimization with conformal mapping[J]. Computer Methods in Applied Mechanics and Engineering, 2018, 328: 477-497.

[17] YE Q, GUO Y, CHEN S K, et al. Topology optimization of conformal structures on manifolds using extended level set methods (X-LSM) and conformal geometry theory[J]. Computer Methods in Applied Mechanics and Engineering, 2019, 344: 164-185.

[18] ZHOU Y, GAO L, LI H. Graded infill design within free-form surfaces by conformal mapping[J]. International Journal of Mechanical Sciences, 2022, 224: 107307.

[19] HUO W D, LIU C, DU Z L, et al. Topology optimization on complex surfaces based on the moving morphable component method and computational conformal mapping[J]. Journal of Applied Mechanics, 2022, 89(5): 051008.

[20] ZONG H M, LIU H, MA Q P, et al. VCUT level set method for topology optimization of functionally graded cellular structures[J]. Computer Methods in Applied Mechanics and Engineering, 2019, 354: 487-505.

[21] WANG L, LI Z S, NI B W, et al. A robust topology optimization method considering bounded field parameters with uncertainties based on the variable time step parametric level-set method[J]. Applied Mathematical Modelling, 2022, 107: 441-463.

[22] LI D W, DAI N, TANG Y L, et al. Design and optimization of graded cellular structures with triply periodic level surface-based topological shapes[J]. Journal of Mechanical Design, 2019, 141(7): 071402.

[23] 蔡升勇, 习俊通. 基于形函数控制的组织工程骨架孔隙结构实体建模方法[J]. 机械工程学报, 2009, 45(7): 297-304.
CAI S Y, XI J T. Computer modeling approach for a porous scaffold structure in tissue engineering based on the shape function control[J]. Journal of Mechanical Engineering, 2009, 45(7): 297-304(in Chinese).

[24] WANG Y Q, CHEN F F, WANG M Y. Concurrent design with connectable graded microstructures[J]. Computer Methods in Applied Mechanics and Engineering, 2017, 317: 84-101.

[25] CRAMER A D, CHALLIS V J, ROBERTS A P. Microstructure interpolation for macroscopic design[J]. Structural and Multidisciplinary Optimization, 2016, 53(3): 489-500.

[26] 杜义贤, 张跃, 付君健, 等. 承载散热一体化的功能梯度多孔结构拓扑优化设计[J]. 计算机辅助设计与图形学学报, 2021, 33(7): 1141-1150.

DU Y X, ZHANG Y, FU J J, et al. Topology optimization design of functionally graded cellular structure with integrated load bearing and heat dissipation[J]. Journal of Computer-Aided Design & Computer Graphics, 2021, 33(7): 1141-1150(in Chinese).

[27] LIU H, ZONG H M, SHI T L, et al. M-VCUT level set method for optimizing cellular structures[J]. Computer Methods in Applied Mechanics and Engineering, 2020, 367: 113154.

[28] ALLAIRE G, JOUVE F, TOADER A M. Structural optimization using sensitivity analysis and a level-set method[J]. Journal of Computational Physics, 2004, 194(1): 363-393.

[29] 龙凯, 左正兴. 基于拓扑优化和形状优化方法的主轴承盖结构设计[J]. 农业机械学报, 2008, 39(4): 152-156.
LONG K, ZUO Z X. Structural design of main bearing cap based on topological optimization and shape optimization[J]. Transactions of the Chinese Society for Agricultural Machinery, 2008, 39(4): 152-156(in Chinese).

[30] 范文杰, 徐进永, 张子达. 基于双向渐进结构优化法的装载机动臂拓扑优化[J]. 农业机械学报, 2006, 37(11): 24-27.
FAN W J, XU J Y, ZHANG Z D. Topology optimization on arm of loader based on BESO[J]. Transactions of the Chinese Society for Agricultural Machinery, 2006, 37(11): 24-27(in Chinese).

[31] ZHOU M, ROZVANY G. The COC algorithm, Part II: Topological, geometrical and generalized shape optimization[J]. Computer Methods in Applied Mechanics and Engineering, 1991, 89(1-3): 309-336.

[32] XIE Y M, ZUO Z H, HUANG X D, et al. Convergence of topological patterns of optimal periodic structures under multiple scales[J]. Structural and Multidisciplinary Optimization, 2012, 46(1): 41-50.

Topological optimization method for conformal cellular structures on surfaces based on co-simulation

FU Junjian[1, 2, 3], XU Yong[2], ZHOU Xiangman[1, 2, *], SHU Zhengtao[2], TIAN Qihua[1, 2]

(1. Hubei Key Laboratory of Hydroelectric Machinery Design & Maintenance, Yichang 443002, China;
2. College of Mechanical & Power Engineering, China Three Gorges University, Yichang 443002, China;
3. Stale Key Laboratory of Intelligent Manufacturing Equipment and Technology, Wuhan 430074, China)

Abstract: The optimization design for structures with surface is challenging due to the mismatch between cellular unit cell configuration and the design domain shape. Thus, a topological optimization design method for conformal cellular structures on surfaces based on co-simulation is proposed. The parametric modelling of cellular structures is achieved by utilising the implicit level set functions derived from triply periodic minimal surfaces. By using the isoparametric element method, the mapping relationship between the natural coordinate system and the cartesian coordinate is established to realize the conformal modeling of the cellular structures on surfaces. A linear interpolation function is constructed to ensure the C^0 continuity of the cellular unit cell. An optimization framework for cellular structures is built by combining MATLAB and ANSYS with the use of variable cutting level set functions.Numerical examples show that this method can effectively realize the topological optimization design of cellular structures on surfaces. The method also ensures the matching of the cellular structures unit cell with the shape of the design domain, and improves the mechanical properties of the cellular structures.

Keywords: topology optimization; cellular structures; conformal modeling; variable cutting level set function; co-simulation

Received: 2022-08-31; **Accepted**: 2023-01-14; **Published Online**: 2023-02-28 10:54

URL: link.cnki.net/urlid/11.2625.V.20230227.1919.006

Foundation items: National Natural Science Foundation of China (51775308); Natural Science Foundation of Hubei Province (2021CFB236); Open Fund of Stale Key Laboratory of Intelligent Manufacturing Equipment and Technology (IMETKF2023016)

*** Corresponding author**. E-mail: zhouxman@ctgu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0703

基于多通道图神经网络的属性聚合式实体对齐

纪鑫，武同心*，王宏刚，杨智伟，何禹德，赵晓龙
(国家电网有限公司大数据中心，北京 100052)

摘　　要：随着电力大数据时代的到来，电网企业在多年的技术监督工作中积累了大量的多模态数据。多模态数据的结构化存储和融合是电力系统有效组织和管理的关键。为融合构建大规模的电力缺陷多模态知识图谱，提出基于多通道图神经网络的多模态实体对齐方法，以有效融合多源电力异构数据。收集电力领域缺陷日志记录，构建电力缺陷多模态知识图谱实体对齐数据集(EKG)，将文本、图像等多模态信息整合到知识图谱中，为实体对齐任务提供丰富的语义信息。多模态数据增加了实体和关系表示的难度，通过挖掘电力领域多模态知识图谱的特征信息，设计属性聚合式对齐方法，利用知识图谱中的多模态属性和结构信息从图像、文本、名称和结构4个维度学习节点表示，解决电力缺陷多模态知识图谱无法有效集成的问题。实验结果表明：所提方法在EKG上取得了最好的效果。

关　键　词：电力知识图谱；图神经网络；实体对齐；设备缺陷；知识图谱
中图分类号：TP391.1
文献标志码：A　　**文章编号**：1001-5965(2024)09-2791-09

随着电网规模的扩大，电力系统知识量呈爆炸式增长，电力市场深化改革正处于关键阶段[1]。为有效地组织、管理和利用海量知识，可将知识图谱技术引入电力系统领域。知识图谱以实体-关系-实体或实体-属性-属性值2种三元组的形式存储事实。其支持各种应用程序，如语义搜索、问题回答和推荐系统。多模态知识图谱可以从大量的文献和图像数据中获取结构化知识，并以结构化的方式描述电力系统中的概念、实体、事件和关系，为电力产业链提供更有效的跨媒体大数据组织、管理和认知能力。

随着实体种类和数量的增加，研究多模态实体对齐技术变得越来越重要。多模态实体对齐决定着现实世界中具有不同信息源的2个或多个实体是否指向相同的对象[2-5]。如果多个实体表示相同的对象，则在这些实体之间建立对齐关系，并合并和聚合实体中包含的信息。

电力领域中实体的属性具有多模态特性[6-9]。一些实体可能具有图像属性。例如，实体“变压器001”有描述其形状和颜色的图像，而实体“铁芯”则没有对应的图像。在具有多模态属性的知识图谱中，以往的方法难以挖掘足够的信息并实现实体的对齐。此外，知识图谱在电力领域中的结构是不平衡的[10-11]。例如，一些不太常用的电力设备，没有足够的属性信息和邻居信息。已有的方法大多只聚合邻居实体或多跳子图[12]，而对一些邻居较少的实体的对齐效果较差。

针对上述问题，本文提出基于多通道神经网络的多模态实体对齐方法，用于融合构建大规模电力缺陷多模态知识图谱。首先，构建电力缺陷多模态知识图谱实体对齐数据集(简称EKG)。该数据集包含7种实体和关系类型，描述了设备的基本信

收稿日期：2022-08-11；**录用日期**：2022-09-15；**网络出版时间**：2023-04-18 16:32
网络出版地址：link.cnki.net/urlid/11.2625.V.20230418.1119.001
基金项目：国家电网有限公司大数据中心科技项目(SGSJ0000FXJS2100099)
*通信作者. E-mail：tongxin-wu@sgcc.com.cn

引用格式：纪鑫，武同心，王宏刚，等. 基于多通道图神经网络的属性聚合式实体对齐[J]. 北京航空航天大学学报，2024，50(9)：2791-2799.
JI X，WU T X，WANG H G，et al. Attribute aggregation entity alignment based on multi-channel graph neural network[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2791-2799 (in Chinese).

息、设备之间的关联关系、设备缺陷产生的原因和需要采取的措施，为多模态实体对齐任务提供丰富的属性和结构信息。其次，多模态数据增加了实体和关系表示的难度，本文提出属性聚合式对齐方式，采用 VGG16 初始化图像属性，BERT (bidirectional encoder representations from transformers) 初始化文本属性表示和名称属性表示，利用多模态知识图谱中的属性和结构信息，从图像、文本、名称和结构 4 个维度学习节点表示，经过归一化后对其嵌入进行集成。最后，进行评估模型充分利用知识图谱中的节点和属性信息，解决多模态数据不能被很好地利用和属性不平衡的问题。实验结果表明，本文方法对构建的 EKG 具有较好的性能。

1 多模态数据预处理

将多模态知识图谱技术引入电力领域已迫在眉睫。电力设备结构复杂，运行维护涉及设备信息、缺陷信息、实时监控信息、维护信息等大量数据[13-15]。本文建立多模态知识图谱，方便电力人员对电力信息进行查询和分析[16-18]，提高了电网的信息利用率，为电力设备运维提供了宝贵的知识和工程经验，为电厂设备和故障管理提供决策支持。结合缺陷模式及其影响分析，构建电力缺陷多模态知识图谱，实现其知识的有效管理和应用。本文建立了 2 个电力缺陷多模态知识图谱：电力缺陷多模态知识图谱 1(简称知识图谱 1)和电力缺陷多模态知识图谱 2(简称知识图谱 2)。其包括实体、图像属性、文本属性和名称属性，如图 1 所示，比对关系建立在与比对种子属于同一实体的实体之间。训练集与测试集的比对种子比例为 8:2。由于知识图谱 1 和知识图谱 2 之间难以直接进行实体对齐，因此，需要对这 2 个知识图谱进行如下预处理。

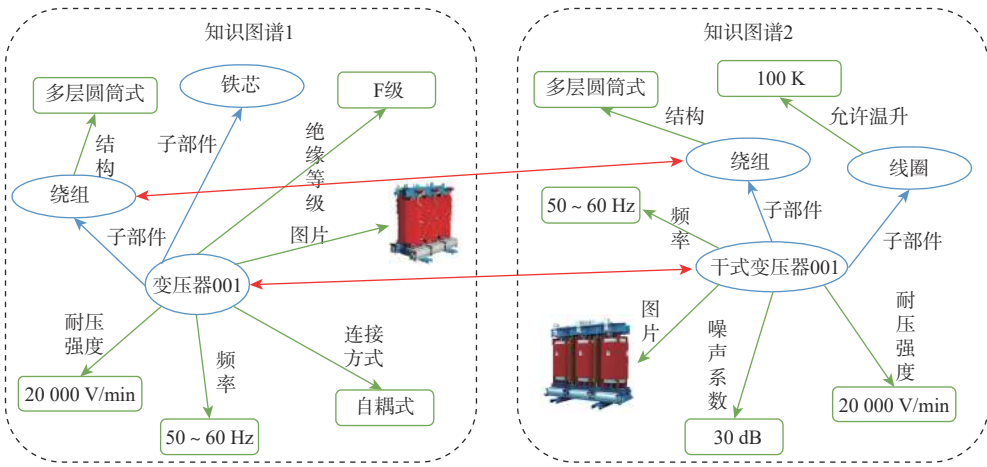


图 1 多模态实体对齐示例

Fig. 1 Multi-modal entity alignment examples

1) 关系对齐预处理。同一关系在不同的知识图谱中可能有 2 种或 2 种以上的不同表达[19-21]，而在不同的知识图谱中，表达方式也是不同的。因此，本文对知识图谱 1 和知识图谱 2 的关系进行了预对齐。通过计算编辑距离，以 0.95 为相似度阈值，实现关系的对齐。当知识图谱 1 和知识图谱 2 之间关系的编辑距离大于 0.95 时，知识图谱 2 中的关系就会被重命名为知识图谱 1 中的关系。

2) 图像统一处理。在多模态知识图谱中，图像的大小和分辨率常常是不一致的[22-26]。图像中所代表的对象有可能只占据了图像的一部分，这将导致模型学习不必要的信息。因此，需将所有图像输入目标检测方法 FastRCNN (fast region-based convolutional neural networks) 中，并将包含必要信息的最小图像作为新图像。

2 多模态实体对齐方法

与传统的知识图谱不同，文本构建的多模态知识图谱包含丰富的属性信息，如图像和节点描述信息 (如特定设备的具体缺陷描述)，这对学习实体表示非常重要。因此，本文方法设计了多个通道来学习具有多个属性信息的节点表示。图 2 为本文方法示意图，该方法利用知识图谱中的多模态属性和结构信息从图像、文本、名称和结构 4 个方面学习节点表示。VGG16 初始化图像，BERT 来初始化文本属性表示和名称属性表示。4 个方面共同学习节点表示，实现多模态知识图谱对齐。图 2 中，GCN 为图卷积网络 (graph convolutional network, GCN)，GAT 为图注意力网络 (graph attention network, GAT)，Mean Aggregate 为均值聚合，h 为节点表征，n 为聚合层数。本文方法包括 2 个阶段。

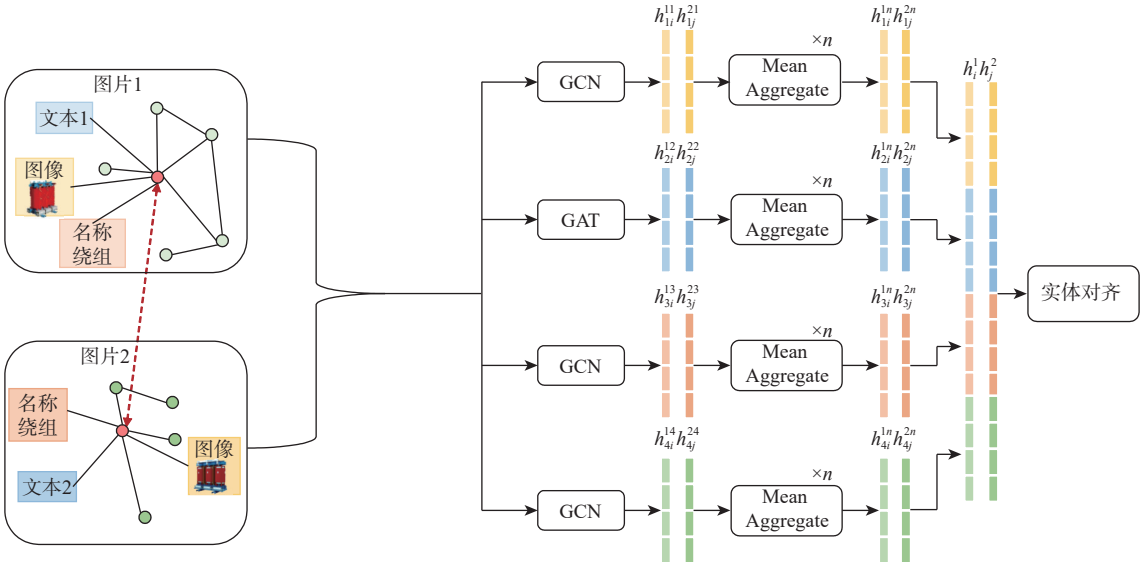


图 2　本文方法示意图

Fig. 2　Schematic diagram of the proposed method

1) 实体表示。实体表示阶段使用图神经网络(graph neural networks, GNN)综合学习带有多模态属性(图像、名称、文本属性)的节点表示。

2) 实体对齐。这个阶段利用目标实体及其邻居的特征来生成嵌入的实体。然后, 利用最终的实体表征来推断实体的相似性。

2.1　实体表示

本文对 4 种不同的数据形式执行不同的特征向量初始化。

1) 图像嵌入G_{p}。保存结构信息和图像属性信息。对于一个实体, 假设有一个图像, 该图像由经过预处理的 VGG16 表示, 得到该实体的初始表示向量$\boldsymbol{v}_{1i}$。G_{p}中实体向量$\boldsymbol{e}_i$的第 1 层表示向量为

$$\boldsymbol{e}_{1i}^{1}=\mathrm{ReLU}(W_{11}\boldsymbol{v}_{1i}) \tag{1}$$

式中: ReLU 为激活函数; W_{11}为第 1 层G_{p}的可学习参数。在G_{p}中, GCN 用于聚合实体的邻居实体$N(\boldsymbol{e}_i)$和其初始表示, 以获得其最终表示为

$$\boldsymbol{E}_{\mathrm{p}}^{l}=\mathrm{GCN}\left(\boldsymbol{R}_{\mathrm{p}}^{l-1},\boldsymbol{E}_{\mathrm{p}}^{l-1}\right)=\mathrm{ReLU}\left(\mathrm{Pool}\left(\boldsymbol{E}_{\mathrm{p}1}^{l},\boldsymbol{E}_{\mathrm{p}2}^{l},\cdots,\boldsymbol{E}_{\mathrm{p}t_{\mathrm{p}}}^{l}\right)\right) \tag{2}$$

式中: $\boldsymbol{R}_{\mathrm{p}}^{l-1}$为关系在 l-1 层的关系表示; Pool操作用于压缩特性。对实体子图的信息进行聚合。G_{p}中$N-1$层的实体表示为$\boldsymbol{h}_{\mathrm{p}i}^{N-1}$, N_i是$\boldsymbol{e}_i$邻居集。第n层中的实体$\boldsymbol{e}_i$表示为

$$\boldsymbol{h}_{\mathrm{p}i}^{n}=\mathrm{ReLU}\left(W_{\mathrm{p}n}\mathrm{MEAN}\left(\left\{\boldsymbol{h}_{\mathrm{p}i}^{n-1}\right\}\cup\left\{\boldsymbol{h}_{\mathrm{p}j}^{n-1},\forall\boldsymbol{e}_j\in N(\boldsymbol{e}_i)\right\}\right)\right) \tag{3}$$

式中: $W_{\mathrm{p}n}$为G_{p}的可学习参数; MEAN 为计算所有向量的平均向量值。为使子图的输出矢量差更小, 需采用 L2 归一化操作:

$$\mathrm{norm}\left(\boldsymbol{h}_{\mathrm{p}i}^{n}\right)=\frac{\boldsymbol{h}_{\mathrm{p}i}^{n}}{\left\|\boldsymbol{h}_{\mathrm{p}i}^{n}\right\|_2} \tag{4}$$

2) 文本嵌入G_{t}。对于实体, 假设其具有n个文本属性。BERT 用于初始化每个实体的文本属性。属性向量为$[\boldsymbol{a}_{21},\boldsymbol{a}_{22},\cdots,\boldsymbol{a}_{2n}]$, 属性值向量为$[\boldsymbol{v}_{21},\boldsymbol{v}_{22},\cdots,\boldsymbol{v}_{2n}]$。第 1 层中每个实体$\boldsymbol{e}_i$的表示使用一组属性和值。第1层实体在文本维度的表示为

$$\boldsymbol{h}_{\mathrm{t}i}^{1}=\mathrm{ReLU}\left(\sum_{j=1}^{n}\alpha_j W_{\mathrm{t}1}\left[\boldsymbol{a}_j,\boldsymbol{v}_j\right]\right) \tag{5}$$

式中: 属性向量表示为

$$\boldsymbol{\alpha}_j=\mathrm{softmax}\left(\mathrm{ReLU}\left(\boldsymbol{u}^{\mathrm{T}}\left[\boldsymbol{h}_j^{0};\boldsymbol{a}_j\right]\right)\right) \tag{6}$$

其中: $\boldsymbol{W}_{\mathrm{t}1}\in\mathbf{R}^{n\times n}$和$\boldsymbol{u}\in\mathbf{R}^{n\times n}$为可学习的矩阵; [;] 表示向量拼接; $\boldsymbol{h}_j^0$为由 node2vec 初始化实体的初始向量。由于一个实体可能有多个文本属性, 因此, 图形注意网络可用于聚合其所有其他属性。最终从文本属性维度看, 实体表示为

$$\boldsymbol{E}_{\mathrm{t}}^{l}=\mathrm{GCN}\left(\boldsymbol{R}_{\mathrm{t}}^{l-1},\boldsymbol{E}_{\mathrm{t}}^{l-1}\right)=\mathrm{ReLU}\left(\mathrm{Pool}\left(\boldsymbol{E}_{\mathrm{t}1}^{l},\boldsymbol{E}_{\mathrm{t}2}^{l},\cdots,\boldsymbol{E}_{\mathrm{t}t_{\mathrm{p}}}^{l}\right)\right) \tag{7}$$

在G_{t}中第n实体表示与G_{p}相同。采用 ReLU 激活函数获取实体表示:

$$\boldsymbol{h}_{ti}^{n}=\mathrm{ReLU}\left(\boldsymbol{W}_{tn}\mathrm{MEAN}\left(\left\{\boldsymbol{h}_{ti}^{n-1}\right\}\cup\left\{\boldsymbol{h}_{tj}^{n-1},\forall\boldsymbol{e}_j\in N(\boldsymbol{e}_i)\right\}\right)\right) \tag{8}$$

3) 图节点的名字嵌入G_{a}。对于实体, BERT 用于获取其实体名称的初始表示向量$\boldsymbol{v}_{3i}$。G_{a}中实体$\boldsymbol{e}_i$的第 1 层表示向量表示为

$$\boldsymbol{h}_{\mathrm{a}i}^{1}=\mathrm{ReLU}(\boldsymbol{W}_{\mathrm{a}1}\boldsymbol{v}_{\mathrm{a}i}) \tag{9}$$

式中: $\boldsymbol{W}_{\mathrm{a}1}$为第 1 层$G_{\mathrm{a}}$的可学习参数矩阵。然后, 本

文使用GCN将其邻居实体$N(\boldsymbol{e}_i)$的初始表示和其初始表示聚合起来，以获得实体的最终表示。最终从名字维度看，实体表示为

$$\boldsymbol{E}_{\mathrm{a}}^{l}=\mathrm{GCN}\left(\boldsymbol{R}_{\mathrm{a}}^{l-1},\boldsymbol{E}_{\mathrm{a}}^{l-1}\right)=\mathrm{ReLU}\left(\mathrm{Pool}\left(\boldsymbol{E}_{\mathrm{a}1}^{l},\boldsymbol{E}_{\mathrm{a}2}^{l},\cdots,\boldsymbol{E}_{\mathrm{a}t_{\mathrm{p}}}^{l}\right)\right) \tag{10}$$

第n层的实体表示$\boldsymbol{h}_{ai}^{n}$与第n-1层的子图表示$\boldsymbol{h}_{ai}^{n-1}$相同。

4）结构嵌入G_{s}。由于该结构嵌入中不存在属性信息，因此，可直接将node2vec初始化得到的初始向量作为初始结构特征向量。对于实体，将随机生成初始表示向量$\boldsymbol{v}_{si}$。由于G_{s}比其他子图获得初始向量的过程少，因此，可直接通过聚合运算得到G_{s}中实体$\boldsymbol{e}_i$的第1层表示向量。然后，文本使用GCN将其邻居实体$N(\boldsymbol{e}_i)$的初始表示和其初始表示聚合起来，以获得实体的最终表示。最终从结构维度看，实体表示为

$$\boldsymbol{E}_{\mathrm{s}}^{l}=\mathrm{GCN}\left(\mathbf{R}_{\mathrm{s}}^{l-1},\boldsymbol{E}_{\mathrm{s}}^{l-1}\right)=\mathrm{ReLU}\left(\mathrm{Pool}\left(\boldsymbol{E}_{\mathrm{s}1}^{l},\boldsymbol{E}_{\mathrm{s}2}^{l},\cdots,\boldsymbol{E}_{\mathrm{s}t_{\mathrm{p}}}^{l}\right)\right) \tag{11}$$

第n层的实体表示$\boldsymbol{h}_{si}^{n}$与第n-1层的子图表示$\boldsymbol{h}_{si}^{n-1}$相同。

2.2 实体对齐

实体对齐通过减少种子等效实体之间的距离，将具有不同数据形式的2个知识图谱表示统一到一个向量空间中。本文训练了4种数据形式的节点表征，经过归一化后对其嵌入进行集成，最后进行评估。

本文采用拼接的方法对4种数据形式的节点表征进行合并，最终进行推理。对于拼接法，保留了所有数据形式的特征，直接拼接4种数据形式的归一化向量节点表征，得到编码器的最终输出向量：

$$\boldsymbol{h}_i^{l}=\left[\mathrm{norm}\left(\boldsymbol{h}_{1i}^{l}\right);\mathrm{norm}\left(\boldsymbol{h}_{2i}^{l}\right);\mathrm{norm}\left(\boldsymbol{h}_{3i}^{l}\right);\mathrm{norm}\left(\boldsymbol{h}_{4i}^{l}\right)\right] \tag{12}$$

3 数据集与实验设置

3.1 数据集统计

本文构建了EKG，表1为电力知识图谱中包含的所有节点类型和关系类型，总共包含7种节点类型和关系类型。为实现实体对齐，本文构建了2个电力缺陷多模态知识图谱，分别为知识图谱1和知识图谱2，如表2所示。表2给出2个知识图谱中的节点类型、关系类型和属性类型信息。电力缺陷多模态知识图谱1中实体类型数为3 229，关系类型数为7，属性类型数为24；电力缺陷多模态知识图谱2中，实体类型、关系类型和属性类型数分别为4 821、7和18。

表3为2个电力缺陷多模态知识图谱的节点和关系数量统计，知识图谱1由3 498个关系、2 363个节点和3 077个属性组成，其中，包括1 032幅图像和2 045个其他类型的属性。知识图谱2由3 531个关系、2 363个节点和3 132个属性组成，包括1 044幅图像和2 088个属性文字类型。2个图谱间的共同实体包括2 117个实体种子。将EKG以8∶1∶1的比例划分为训练集、验证集和测试集。

表1 电力缺陷多模态知识图谱节点和关系类型

Table 1 Node and relationship types of power defects of multi-modal knowledge graph

节点类型	节点关系
设备	子部件关系
缺陷	缺陷关系
缺陷图片	图片
危害	缺陷危害
成因	引起
措施	具体措施
设备型号	型号

表2 数据集统计信息

Table 2 Dataset statistics

类型	实体类型数	关系类型数	属性类型数	图像数
知识图谱1	3 229	7	24	3 821
知识图谱2	4 821	7	18	5 438

表3 电力缺陷多模态知识图谱统计

Table 3 Statistics of power defects of multi-modal knowledge graph

类型	关系数	节点数	属性数
知识图谱1	3 498	2 363	3 077
知识图谱2	3 531	2 363	3 132

3.2 实验设置

本文方法是基于PyTorch实现的，这是一个开源的深度学习框架[22-29]。本文方法是基于GCN的GNN，每批次训练样本的大小为32，节点嵌入维数为128，随机失活比率为0.6，GCN和GAT模型层数为2，训练轮数为100，L2正则化值为0.000 1。对于4个通道的学习率的评价，本文采用网格搜索的方法，步长为0.001，则模型学习速率为0.001。评价指标包括：MR为平均排名（mean rank，MR）、MRR为平均倒数排名（mean reciprocal rank，MRR）、Hits@1为正确答案出现在所有可能答案中的第1个位置的比例（hyperlink-induced topic search@1，Hits@1）、

Hits@10 为正确答案出现在所有可能答案的前十位中的比例 (hyperlink-induced topic search@10, Hits@10)。为防止模型过拟合, 便选择了早期停止策略。具体来说, 如果模型在连续 100 次迭代内没有降低验证集的损失函数, 则操作将停止。为确保公平性所有基线模型都使用相同的数据集分区。并使用相同的节点表示维度, 该维度被设置为 128。所有的实验都是在一台只有一个 GPU (NVIDIA GTX-1080 Ti) 的服务器上进行的。本文方法主要参数如表 4 所示。

表 4　本文方法主要参数

Table 4　Main parameters of the proposed method

参数	本文方法最优值
每批次训练样本的大小	32
节点嵌入维数	128
随机失活比率	0.6
GCN和GAT模型层数	2
训练轮次	100
L2正则化	0.000 1
Adam优化器中的初始学习率	0.001

3.3　基线模型

将本文方法的模型与其他基线模型在本文构建的数据集上进行比较: ①GCC-align (graph convolutional networks for entity alignment)[22] 模型, 其通过 GCN 编码实体和属性。②KDCoE (knowledge distillation via collaborative ensemble)[23] 模型, 其使用节点属性和迭代策略, 并运用多个模型从不同的角度学习。③MuGNN (multi-relational graph neural network)[24] 模型是一个多通道 GNN 模型, 其通过不同的关系加权方案编码一个知识图谱通道。④SEA (self-ensemble alignment)[25] 模型, 其使用对抗训练, 并利用标记和未标记实体实现实体对齐。⑤MultiKE (multi-view knowledge embedding) [26] 模型, 其使用几种组合策略来基于实体名称、关系和属性的视图嵌入实体。⑥RSN4EA (recurrent skipping network for entity alignment)[27] 模型, 其使用偏置随机漫步采样方法生成跨越 KGs 的长路径, 并使用循环跳绳网络对路径进行建模。⑦AliNet (alignment network)[28] 模型, 其以端到端的方式缓解了邻居结构的非同构性。⑧AttrGNN (attribute-aware graph neural network)[29] 模型, 其使用属性值编码器, 将知识图谱划分为子图, 有效地建模各种类型的属性三元组。

4　实　验

4.1　主实验

表 5 为不同模型在 4 个评估指标上的性能。本文模型利用文本、图像和结构多模态信息从多个维度学习节点表示, 从而获得最佳的实体对齐效果, 如表中加粗部分所示。与现有的基于属性的实体对齐模型 (如 GCN-align[22] 和 KDCoE[23]) 相比, 本文模型有了显著的改进。这可能是因为本文模型采用基于 GNN 的方法构造具有不同属性信息的子图, 并分别学习表示。与 MuGNN[24]、MultiKE[26] 和 AttrGNN[29] 等多视图实体对齐模型相比, 本文模型利用图像信息获得语义更丰富的节点表示。与 AttrGNN[29] 模型相比, 本文模型分别学习文本和图像的节点表示, 而不是将文本和图像映射到同一个空间中。与 AttrGNN[29] 模型相比, 本文模型 MR 降低了 2.04%, MRR 值增加了 3.50%, Hits@1 值增加了 4.77%, Hits@1 值增加了 6.31%。这表明, 以基于 GNN 的方式学习文本、图像和结构语义可以使实体包含更丰富的语义信息。

表 5　实体对齐性能

Table 5　Entity alignment performances　%

模型	MR	MRR	Hits@1	Hits@10
GCN-align[22]	51.37	60.73	50.92	77.38
KDCoE[23]	33.83	53.35	36.83	60.24
MuGNN[24]	42.82	47.23	47.85	68.76
SEA[25]	28.23	64.27	51.21	70.72
MultiKE[26]	41.78	54.93	51.19	75.25
RSN4EA[27]	32.72	49.02	52.72	73.82
AliNet[28]	38.83	61.38	53.48	75.63
AttrGNN[29]	28.28	62.82	53.05	75.42
本文模型	**26.24**	**66.32**	**57.82**	**81.73**

4.2　消融实验

在本文的预训练和事件检测任务中所涉及到的主要参数如表 6 所示。在预训练任务中, 评估预训练效果的指标为模型预测被替换单词的准确率; 在事件检测任务中, 评估任务效果的指标为 MR、MRR、Hits@1、Hits@10。其中, 精确率是指返回的真正例占所有返回的正例的比例, 召回率是指返回的真正例占所有正例的比例, F_1 值则是精确率和准确率的调和均值。本文在表 6 中评估了本文的 4 种通道在各种组合中的性能。当本文只使用一个通道时, 命名通道在 MR、MRR 和 Hits@1 指标上表现最好, 而结构化通道在 Hits@10 指标上表现最好。这说明在属性信息稀缺的情况下, 使用名称的语义信息可以取得较好的效果。当本文使用 2 个通道的任意组合时, 2 个通道都比原来的单一通道的性能有所提高, 并且 2 个通道的任意组合都比最佳单一通道的性能更高, 这表明基于 GNN 的多维度数据表征对本文的数据集是有效的。在 MRR、

表6 消融实验

Table 6 Ablation experiments %

数据格式	信息	MR	MRR	Hits@1	Hits@10
使用1种数据格式	图像信息(G1)	40.23	42.73	31.32	62.89
	文本信息(G2)	43.83	39.51	34.27	65.73
	名字信息(G3)	38.29	40.36	38.92	67.28
	结构信息(G4)	39.27	41.64	42.03	69.20
使用2种数据格式	G1+G2	33.02	50.72	49.24	73.54
	G1+G3	36.71	52.20	51.53	77.09
	G1+G4	32.46	53.01	51.01	74.37
	G2+G3	30.06	56.24	50.23	75.40
	G2+G4	34.12	58.91	52.17	77.74
	G3+G4	31.58	57.54	53.92	78.37
使用3种数据格式	G1+G2+G3	29.70	62.87	53.27	79.73
	G1+G2+G4	28.06	65.29	56.84	80.28
	G1+G3+G4	28.25	62.86	54.24	81.07
	G3+G3+G4	26.94	59.72	55.39	80.94
使用所有数据格式	G1+G2+G3+G4	**26.24**	**66.32**	**57.82**	**81.73**

Hits@1 和 Hits@10 中，名称信道和结构信道的结合效果最好。当本文使用任意三通道组合时，图像通道、名称通道和结构通道的组合在 MRR、Hits@1 和 Hits@10 指标上都取得了最好的性能。当本文使用 4 个通道时，本文模型在 4 个指标上都达到了最优的性能，如表中加粗部分所示。结果表明，图像通道、文本通道、名称通道和结构通道的结合使实体表示语义更好。

4.3 模型参数

为进一步测试本文模型对参数的敏感性，本节在 EKG 上进行参数实验。图 3 和图 4 显示了对齐种子率、GCN 层和节点表示的维度的影响。

1）对齐种子率。如图 3 所示，本文在数据集上使用 2 个层将本文模型与 RSN4EA[27]、AliNet[28] 和 AttrGNN[29] 模型进行比较。将不同比例的比对种子作为训练数据，本文模型明显优于基于图学习的方法。结果表明，本文模型利用了多种类型属性，且无论比对种子的数量多少，权重可学习模块都是有效的。当使用完整的数据集时，本文模型表现最好。

2）模型层。图 4 (a) 显示了本文模型具有 1～5 层 GCN 的性能。可以看出，随着模型层数的增加，本文模型的性能明显下降 (语义混淆现象)。本文模型 2 层可以学习到具有可区分度的节点表示，而本文模型 5 层学习到的节点表示已经变得不可区分。本文模型的层数相对较浅 (通常 1～2 层比较好)。其可以捕获异构结构，并在图上保留丰富的语义。因为本文模型在 2 个 GCN 层下的性能最好，所以本文模型使用了 2 层 GCN。

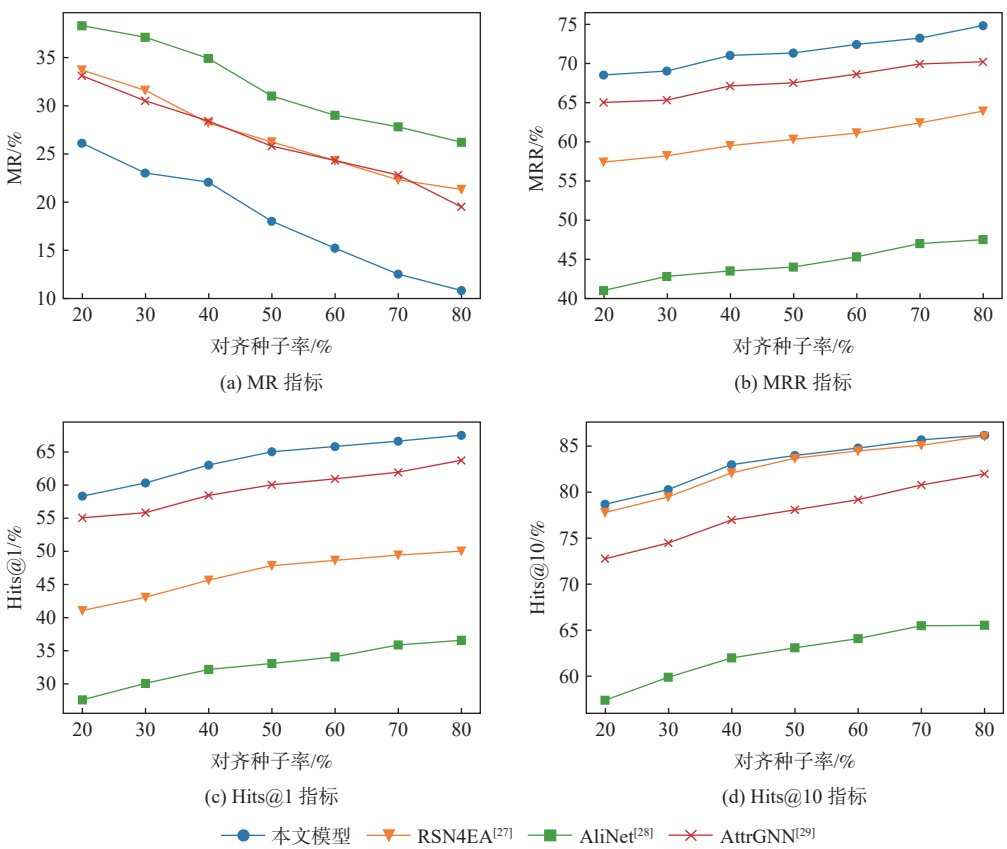


图 3 不同对齐种子率下的各模型性能

Fig. 3 Performance of each model with different alignment seed rates

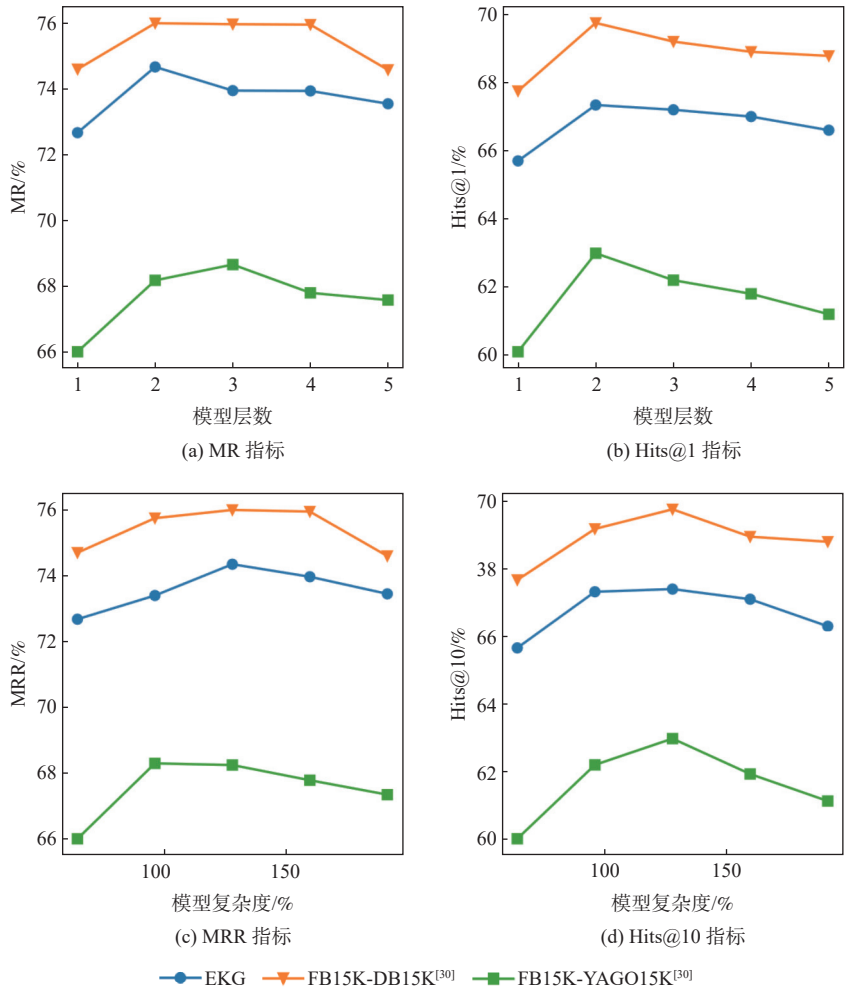


图 4　不同模型层数和维度下的本文模型性能

Fig. 4　Performance of the proposed model with different layers and dimensions

3) 模型维度。节点的表示维度将直接影响模型的性能。本文模型分别在 EKG、FB15K-DB15K[30] 和 FB15K-YAG015K[30] 多模态实体对齐数据集上进行实验。如图 4(b) 所示, 随着节点表示维度的增加, 本文模型的性能先是缓慢上升, 然后保持不变, 最后缓慢下降。这是因为本文模型需要足够的维度来编码语义和结构信息, 但维度太大可能会引入一些冗余, 导致过拟合现象。由于本文模型在节点表示维数为 128 时性能最好, 因此, 本文模型最终将节点表示维度设为 128。

5　结　论

1) 本文利用多模态数据增加实体和关系表示的难度, 设计了基于 GNN 的属性聚合式实体对齐方法, 挖掘电力领域多模态知识图谱的特征信息, 利用知识图谱中的图像、文本、名称等多模态属性和结构信息从学习节点表示, 有效地解决了电力缺陷领域多模态知识图谱融合难题。

2) 本文构建了 EKG, 将文本、图像等多模态信息整合到知识图谱中, 为实体对齐任务提供了丰富的语义信息。

参考文献 (References)

[1] 高海翔, 苗璐, 刘嘉宁, 等. 知识图谱及其在电力系统中的应用研究综述[J]. 广东电力, 2020, 33(9): 66-76.
GAO H X, MIAO L, LIU J N, et al. Review on knowledge graph and its application in power systems[J]. Guangdong Electric Power, 2020, 33(9): 66-76(in Chinese).

[2] 王小鹏. 基于知识图谱的择优分段迭代式实体对齐方法研究[J]. 信息与电脑 (理论版), 2021, 33(18): 48-52.
WANG X P. Research on the method of optimal segmentation iterative entity alignment based on knowledge graph[J]. China Computer & Communication, 2021, 33(18): 48-52(in Chinese).

[3] 魏忠诚, 张洁滢, 连彬, 等. 基于双向 GCN 和 CVm 的实体对齐模型研究[J]. 计算机应用研究, 2021, 38(9): 2716-2720.
WEI Z C, ZHANG J Y, LIAN B, et al. Entity alignment model based on two-way GCN and CVm[J]. Application Research of Computers, 2021, 38(9): 2716-2720(in Chinese).

[4] 马建红, 刘双耀, 杨珺. 多信息加权融合实体对齐算法[J]. 计算机应用与软件, 2021, 38(7): 295-301.
MA J H, LIU S Y, YANG J. Multi-information weighted fusion en-

tity alignment algorithm[J]. Computer Applications and Software, 2021, 38(7): 295-301(in Chinese).

[5] 王悦悦. 面向高铁运维领域的嵌入实体对齐技术研究与实现[D]. 北京: 北京交通大学, 2021.
WANG Y Y. Research and implementation of embedded entity alignment technology for high speed rail operation and maintenance domain[D]. Beijing: Beijing Jiaotong University, 2021(in Chinese).

[6] 田玲, 张谨川, 张晋豪, 等. 知识图谱综述: 表示、构建、推理与知识超图理论[J]. 计算机应用, 2021, 41(8): 2161-2186.
TIAN L, ZHANG J C, ZHANG J H, et al. Knowledge graph survey: Representation, construction, reasoning and knowledge hypergraph theory[J]. Journal of Computer Applications, 2021, 41(8): 2161-2186(in Chinese).

[7] 付雷杰, 曹岩, 白瑀, 等. 国内垂直领域知识图谱发展现状与展望[J]. 计算机应用研究, 2021, 38(11): 3201-3214.
FU L J, CAO Y, BAI Y, et al. Development status and prospect of vertical domain knowledge graph in China[J]. Application Research of Computers, 2021, 38(11): 3201-3214(in Chinese).

[8] 陈烨, 周刚, 卢记仓. 多模态知识图谱构建与应用研究综述[J]. 计算机应用研究, 2021, 38(12): 3535-3543.
CHEN Y, ZHOU G, LU J C. Survey on construction and application research for multi-modal knowledge graphs[J]. Application Research of Computers, 2021, 38(12): 3535-3543(in Chinese).

[9] 张吉祥, 张祥森, 武长旭, 等. 知识图谱构建技术综述[J]. 计算机工程, 2022, 48(3): 23-37.
ZHANG J X, ZHANG X S, WU C X, et al. Survey of knowledge graph construction techniques[J]. Computer Engineering, 2022, 48(3): 23-37(in Chinese).

[10] 宋玮琼, 韩柳, 羡慧竹, 等. 电能计量装置的知识图谱构建与应用[J]. 电测与仪表, 2024, 61(4): 218-224.
SONG W Q, HAN L, XIAN H Z, et al. Construction and application of knowledge graph of electric energy metering device[J]. Electrical Measurement & Instrumentation, 2024, 61(4): 218-224(in Chinese).

[11] 李宏波, 朱永利, 王京保. 基于多层特征融合 CNN 的变压器 PRPD 图谱识别[J]. 电测与仪表, 2020, 57(18): 63-68.
LI H B, ZHU Y L, WANG J B. Transformer PRPD pattern recognition based on multi-layer feature fusion CNN[J]. Electrical Measurement & Instrumentation, 2020, 57(18): 63-68(in Chinese).

[12] 李新鹏, 徐建航, 郭子明, 等. 调度自动化系统知识图谱的构建与应用[J]. 中国电力, 2019, 52(2): 70-77.
LI X P, XU J H, GUO Z M, et al. Construction and application of knowledge graph of power dispatching automation system[J]. Electric Power, 2019, 52(2): 70-77(in Chinese).

[13] 李晓露, 左璇, 刘日亮, 等. 基于形状约束语言的电网模型知识图谱验证方法[J]. 中国电力, 2022, 55(1): 119-125.
LI X L, ZUO X, LIU R L, et al. SHACL-based validation method of knowledge graph for power system model[J]. Electric Power, 2022, 55(1): 119-125(in Chinese).

[14] 宋次剑. 知识图谱构建若干关键技术及公共安全领域应用研究[D]. 合肥: 中国科学技术大学, 2021.
SONG C J. Research on several key technologies of knowledge graph construction and their application in public security field[D]. Hefei: University of Science and Technology of China, 2021(in Chinese).

[15] 刘学壮. 基于主动学习的知识融合系统的设计与实现[D]. 大连: 大连理工大学, 2021.
LIU X Z. Design and implementation of knowledge fusion system based on active learning[D]. Dalian: Dalian University of Technology, 2021(in Chinese).

[16] 阎志刚, 李成城, 林民. 嵌入知识图谱信息的命名实体识别方法[J]. 内蒙古师范大学学报 (自然科学版), 2021, 50(3): 275-282.
YAN Z G, LI C C, LIN M. Named entity recognition method for embedding knowledge graph information[J]. Journal of Inner Mongolia Normal University (Natural Science Edition), 2021, 50(3): 275-282(in Chinese).

[17] 李文娜, 张智雄. 基于联合语义表示的不同知识库中的实体对齐方法研究[J]. 数据分析与知识发现, 2021, 5(7): 1-9.
LI W N, ZHANG Z X. Entity alignment method for different knowledge repositories with joint semantic representation[J]. Data Analysis and Knowledge Discovery, 2021, 5(7): 1-9(in Chinese).

[18] 季一木, 刘艳兰, 刘尚东, 等. 基于 BERT 的多源知识库索引对齐算法[J]. 南京邮电大学学报 (自然科学版), 2021, 41(2): 49-61.
JI Y M, LIU Y L, LIU S D, et al. Multi-source knowledge base index alignment based on BERT[J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2021, 41(2): 49-61(in Chinese).

[19] 田江伟, 李俊锋, 柳青. 结合属性结构的图卷积实体对齐算法[J]. 计算机应用研究, 2021, 38(7): 1979-1982.
TIAN J W, LI J F, LIU Q. GCN-based entity alignment algorithm combined with attribute structure[J]. Application Research of Computers, 2021, 38(7): 1979-1982(in Chinese).

[20] 杭婷婷, 冯钧, 陆佳民. 知识图谱构建技术: 分类、调查和未来方向[J]. 计算机科学, 2021, 48(2): 175-189.
HANG T T, FENG J, LU J M. Knowledge graph construction techniques: Taxonomy, survey and future directions[J]. Computer Science, 2021, 48(2): 175-189(in Chinese).

[21] 鲁佩佩. 基于知识库对齐的命名实体识别方法[J]. 电脑知识与技术, 2021, 17(4): 184-186.
LU P P. Named entity recognition method based on knowledge base alignment[J]. Computer Knowledge and Technology, 2021, 17(4): 184-186(in Chinese).

[22] 曾维新, 赵翔, 唐九阳, 等. 基于重排序的迭代式实体对齐[J]. 计算机研究与发展, 2020, 57(7): 1460-1471.
ZENG W X, ZHAO X, TANG J Y, et al. Iterative entity alignment via re-ranking[J]. Journal of Computer Research and Development, 2020, 57(7): 1460-1471(in Chinese).

[23] WANG Z C, LV Q S, LAN X H, et al. Cross-lingual knowledge graph alignment via graph convolutional networks[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg : Association for Computational Linguistics , 2018: 349-357.

[24] CHEN M H, TIAN Y T, CHANG K W, et al. Co-training embeddings of knowledge graphs and entity descriptions for cross-lingual entity alignment[EB/OL]. (2018-06-18)[2022-05-07]. http://arxiv.org/abs/1806.06478.

[25] CAO Y X, LIU Z Y, LI C J, et al. Multi-channel graph neural network for entity alignment[EB/OL]. (2019-08-26)[2022-05-07]. http://

arxiv.org/abs/1908.09898.

[26] ZHANG Q H, SUN Z Q, HU W, et al. Multi-view knowledge graph embedding for entity alignment[EB/OL]. (2019-06-06)[2022-05-09]. http://arxiv.org/abs/1906.02390.

[27] GUO L B, SUN Z Q, HU W. Learning to exploit long-term relational dependencies in knowledge graphs[EB/OL]. (2019-05-13)[2022-05-09]. http://arxiv.org/abs/1905.04914.

[28] SUN Z Q, WANG C M, HU W, et al. Knowledge graph alignment network with gated multi-hop neighborhood aggregation[EB/OL]. (2019-11-20)[2022-05-10]. http://arxiv.org/abs/1911.08936.

[29] LIU Z Y, CAO Y X, PAN L M, et al. Exploring and evaluating attributes, values, and structures for entity alignment[EB/OL]. (2020-10-07)[2022-05-11]. http://arxiv.org/abs/2010.03249.

[30] LI Q, JI C, GUO S, et al. Multi-modal knowledge graph transformer framework for multi-modal entity alignment[C]//Proceedings of the Findings of the Association for Computational Linguistics. Stroudsburg: Association for Computational Linguistics, 2023: 987-999.

Attribute aggregation entity alignment based on multi-channel graph neural network

JI Xin，WU Tongxin*，WANG Honggang，YANG Zhiwei，HE Yude，ZHAO Xiaolong

(Big Data Center of State Grid Corporation of China，Beijing 100052，China)

Abstract：With the advent of the era of big data in power, power grid enterprises have accumulated a large number of multi-modal data in years of technical supervision work. The structured storage and fusion of multi-modal data are the keys to the effective organization and management of power systems. In order to fuse and construct a large-scale multi-modal knowledge graph of power defects, a multi-modal entity alignment method based on a multi-channel graph neural network was proposed to effectively fuse heterogeneous data of multi-source power. A multi-modal knowledge graph entity alignment dataset (EKG) for power defects was constructed by collecting logs related to many defects in the power field. Multi-modal information such as text and images was integrated into the knowledge graph, which provided rich semantic information for entity alignment tasks. The multi-modal data increased the representation difficulty of the entities and relationships. By mining the characteristics of the multi-modal knowledge graph in the power field, an attribute aggregation alignment method was designed. The node representation was learned from the four dimensions of image, text, name, and structure by using the multi-modal attributes and structure information in the knowledge graph, solving the problem that the power defects of a multi-modal knowledge graph cannot be integrated effectively. Experimental results show that the proposed method achieves the best performance on EKG.

Keywords：power knowledge graph；graph neural network；entity alignment；equipment defect；knowledge graph

Received：2022-08-11；**Accepted**：2022-09-15；**Published Online**：2023-04-18 16：32

URL：link.cnki.net/urlid/11.2625.V.20230418.1119.001

Foundation item：Big Data Center of State Grid Corporation of China Technology Project (SGSJ0000FXJS2100099)

* **Corresponding author**. E-mail：tongxin-wu@sgcc.com.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0708

类双锥型升力式返回器横航向不稳定性分析及改进

陈刚[1,*]，孙雪[1]，李广兴[1]，赵丹[1]，马建颖[1]，何程[2]

(1. 上海航天技术研究院 上海宇航系统工程研究所，上海 201109；2. 南京航空航天大学 无人机研究院，南京 210016)

摘 要：类双锥型升力式返回器的横航向稳定性对其理想再入策略的实施构成严苛的约束，且在设计阶段需要进行专门分析。针对类双锥型升力式返回器的标称再入过程，使用数值仿真方法计算该飞行器在全速域、宽迎角范围内的横航向稳定性，结合流场特征分析亚/跨/超声速速域内的横航向稳定性突变规律；重点分析涡流发生器、航向增强型面、重心偏置对横航向稳定性的影响及其作用机理，最终确定了航向舵偏转与重心偏置相结合的方案就是最佳横航向增稳方案。

关 键 词：升力式返回器；高超声速；横航向稳定性；流动机理；增稳措施

中图分类号：V221$^+$.3; V412.4$^+$.4

文献标志码：A **文章编号**：1001-5965(2024)09-2800-10

类双锥型升力式返回器结合了传统返回舱和航天飞机的特点，能够进行大范围的机动滑翔，并可有效降低过载和热载峰值，是天地往返运输系统的一个重要发展方向[1]。由于升力式返回器再入过程中飞行高度和速度变化巨大，气动特性变化复杂，因此为保障飞行安全，在设计过程中必须保证返回器本体的稳定性。

早在20世纪60年代，研究人员就注意到返回舱的稳定性问题。1968年Moseley等[2]针对阿波罗返回舱的飞行稳定性问题进行了研究，指出返回器表面流动不稳定，之后国内外大量的研究人员对此进行数值模拟。Abe等[3]通过分析返回器的动态俯仰力矩，指出飞船纵向不稳定是返回器后部的涡流造成的；张涵信等[4-5]从非线性动力学理论出发对飞船返回舱的俯仰运动进行了定性理论分析；Teramoto等[6-7]则针对钝头返回舱跨声速产生失稳现象的流动机理进行了研究。然而受限于计算流体动力学(computational fluid dynamics, CFD)当时的发展水平，计算误差较大且对不稳定性产生机理的研究较为粗略。朱德华等[8]采用大涡模拟法分析了钝头返回舱再入过程的分离和转捩现象，并分析了其稳定性特征。部分研究者则采用风洞试验的方法来研究返回器表面的流动。Mitcheltree等[9]基于风洞试验研究了钝头返回器外形半锥角在亚声速飞行时对稳定性的影响；何开锋与和争春[10]在全局稳定性分析和数值模拟的基础上，使用风洞试验验证了返回舱在亚、跨声速飞行时的动稳定性。宋玉辉等[11-12]采用风洞自由振动试验方法研究了钝头返回器的动稳定特性，结果表明在亚、跨声速，甚至超声速范围均出现了动不稳定现象。宋威等[13]也采用风洞自由飞试验技术，研究了钝头类联盟号返回舱在跨声速时的气动特性与运动特性。同时，他们还指出，由于风洞试验中的支撑干扰、动壁干扰等，不可避免会的产生误差。这种误差和数值计算中存在的模型误差、网格误差、离散误差一样，会对试验或计算结果产生影响，因此，需采取撑杆优化、模型尺度控制等措施进行误差控制及修正。

随着高超声速飞行器的发展，研究人员发现此类横航向与纵向特性制约、横航向气动耦合严重等问题也普遍存在于高超声速飞行器中，并对飞行器稳定性的影响因素、失稳后的运动形态和增稳措施开展了大量的研究，而这些研究主要集中在纵向稳

收稿日期：2022-08-11；录用日期：2022-10-04；网络出版时间：2022-10-10 08:44
网络出版地址：link.cnki.net/urlid/11.2625.V.20221009.0850.001
*通信作者. E-mail：agoniage@163.com

引用格式：陈刚，孙雪，李广兴，等. 类双锥型升力式返回器横航向不稳定性分析及改进[J]. 北京航空航天大学学报，2024，50(9)：2800-2809.
CHEN G，SUN X，LI G X，et al. Analysis and improvement of lateral instability of quasi-biconical lifting reentry spacecraft[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2800-2809 (in Chinese).

定性方面。鲍文和姚照辉[14]针对吸气式升力体高超声速飞行器建立了面向巡航控制的纵向运动模型。贾子安等[15-16]结合牛顿理论和流线简化模型，推导了高超飞行器外形流线特征对纵向静稳定性的影响；在此基础上，王柯穆[17]研究了乘波体流线特征对纵向压心的影响。Mangin 等[18]则将飞行器上表面尾部设计为膨胀面，使得纵向压心后移以改善纵向静稳定性。针对横航向稳定性的研究则相对较少，高清等[19-20]基于风洞试验和模态分析对类高超声速技术验证飞行器（hypersonic technology test vehicle, HTV-2 ）外形的横侧向稳定性进行了研究；祝立国等[21]基于横航向小扰动方程推导了开环动态偏航稳定判据，并对高速飞行器的横航向稳定性进行了探讨；闵昌万[22]则针对副翼控制下的高超声速飞行器横侧向稳定性判据开展了研究。

当前，针对类双锥型升力式返回器横航向稳定性变化机理的分析及增稳控制措施方面鲜有研究。为弥补现有研究的不足，本文使用数值仿真方法对类双锥型升力式返回器在全速域、宽迎角范围内的横航向稳定性进行分析，结合流场特征分析了亚/跨/超声速速域内的横航向稳定性突变规律。重点针对该类返回器的横航向稳定性问题研究了涡流发生器、航向增强型面、重心偏置 3 种增稳方案，从流场特征角度分析了 3 种方案的作用机理和效果，最终确定了最佳横航向增强方案。

1　返回器外形及标称再入过程

类双锥型升力式返回飞行器是一种新构型返回器，与传统构型返回器相比，该类返回器具有升阻比高、过载小、落点精度高等特点。典型的类双锥型升力式返回器外形如图 1 所示，该返回器主要包含本体和升降舵 2 个部分。在总体设计过程中，本体外形可基于多约束条件优化设计生成，升降舵几何参数则根据典型状态下的配平要求确定。

再入返回过程中，由返回器纵向特性所决定的理想再入策略如图 2 所示，其中，α 为迎角，Ma 为马赫数，Ma_{cr} 为临界马赫数，H 为飞行高度，K 和 K_{max} 分别为升阻比和最大升阻比。其中：在高马赫数区间（$Ma>Ma_{cr}$），返回器以 45°以上大迎角执行再入，以降低热流峰值；在低马赫数区间（$Ma<Ma_{cr}$），返回器以 20°～30°的小迎角接续飞行，以发挥升力体构型的大升阻比优势，直至速度降至 0.8 马赫附近时执行开伞程序。大小迎角切换的临界马赫数 Ma_{cr} 取 10～12，由热流、动压共同决定。

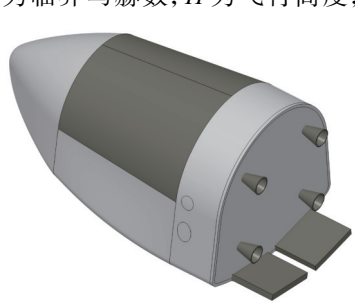


图 1　升力式返回器外形

Fig. 1　Configuration of lifting reentry spacecraft

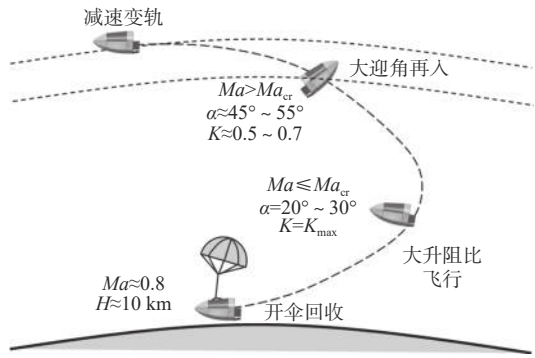


图 2　升力式返回器理想再入策略

Fig. 2　Ideal reentry strategy of lifting reentry spacecraft

这种标称再入策略以返回器纵向特性为基础，理论上可发挥返回器在大迎角和小迎角范围内的不同性能优势，但在实施过程中需进一步考虑横航向特性与纵向特性之间的匹配关系，即在再入过程中的每一个状态点均需保证返回器具有良好的横航向稳定性，这种横纵向匹配是保证飞行安全可控的关键，在设计阶段需要对其开展专门分析。

2　返回器横航向稳定性分析

2.1　数值计算方法

针对图 1 所示的类双锥型升力式返回器外形，使用数值计算方法来获取其在再入过程中的横航向稳定性。计算过程采用如图 3 所示的全模结构化网格逼近返回器外形，同时对头锥、肩部、迎风面进行网格加密以提高流动细节捕捉精度。计算所采用的主控方程为守恒形式的非定常可压缩 Navier-Stokes 方程，无黏通量采用二阶迎风 Roe 通量差分分裂格式离散，网格界面上采用 MUSCL 格式插值以获得高阶精度，黏性项则采用中心差分格式离散[23]。湍流模型采用 SST 两方程湍流模型，湍流控制方程空间离散采用二阶迎风格式。时间推进基于隐式 LU-SGS 格式完成。为验证该数值方法的准确性，将该方法用于文献 [23] 中类神舟返回舱的典型工况（Ma=6, α=20°）气动特性模拟，并将试验结果进行比较，结果如表 1 所示，气动力系数预测值与实验值间的误差在 3% 左右（表中 C_A、C_N、C_L、C_D、C_m 依次为轴向力系数、法向力系数、升力系数、阻力系数、俯仰力矩系数），由此表明数值仿真模型具有较高的可信度。

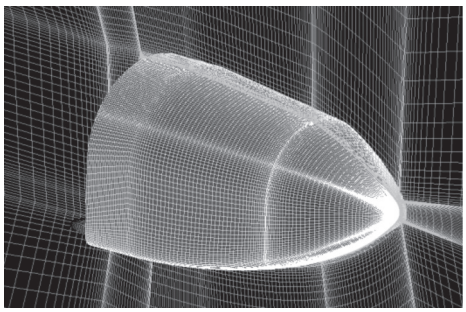


图3 升力式返回器计算网格

Fig. 3 Computational mesh of lifting reentry spacecraft

表1 类神舟返回器计算值与实验值对比

Table 1 Comparison of calculated and experimental values of Shenzhou-like reentry spacecraft

参数项	C_A	C_N	C_L	C_D	C_m
预测值	1.315	0.158 7	−0.300 6	1.290	0.099 9
实验值	1.335	0.159 0	−0.307 2	1.309	0.102 5

2.2 横航向计算及结果讨论

横航向稳定性分析计算工况的设置覆盖了再入返回的整个过程，其中，马赫数为0.5～20，迎角为30°～60°(10°间隔)，侧滑角为2°～4°。使用2.1节所述方法开展数值计算，可得到返回器横航向特性随马赫数的变化，如图4所示，其中，力矩参考点距离头部顶点间距 X_{cg}/L = 60%(L为本体长度)，$Y_{cg}=Z_{cg}=0$。由于小侧滑角范围内偏航力矩系数 C_n 和滚转力矩系数 C_l 随侧滑角变化呈近似线性，因此，在图4(a)和图4(b)中直接将 C_n 和 C_l 表示为侧滑角 β 的偏导数形式，即将 $C_{n\beta}$、$C_{l\beta}$，作为传统横航向静稳定性的主要判据。

在 $C_{n\beta}$ 和 $C_{l\beta}$ 的基础上可进一步导出开环动态偏航稳定性判据 $C_{n\beta DYN}$，如图4(c)所示。$C_{n\beta DYN}=C_{n\beta}\cos\alpha-I_z/I_xC_{l\beta}\sin\alpha$，其中，$I_z/I_x$ 为惯量系数，图1中的返回器取0.498 5。开环动态偏航稳定性判据在传统静稳定性判据的基础上考虑了横航向的相互耦合效应，能够更全面地反映飞行器的真实横航向稳定性能。

传统静稳定性判据要求 $C_{n\beta}>0$，$C_{l\beta}<0$，开环动态偏航稳定要求 $C_{n\beta DYN}>0$。分析图4中的 $C_{n\beta}$、$C_{l\beta}$ 及 $C_{n\beta DYN}$，可得到以下结论。

1) 迎角越小，横航向稳定性越差：再入初期迎角处于30°～50°时返回器呈航向不稳定状态(见图4(a))，在该迎角范围内实施再入无法保障飞行安全；再入迎角达60°时可保证再入初期的横航向稳定性，但在马赫数降至3以下时，亦会进入航向不稳定状态。

2) 横航向耦合对横航向稳定性有改善作用：开环动态偏航稳定性($C_{n\beta DYN}>0$)计算结果表明，考虑横航向耦合的实际横航向静稳定性略优于传统静稳定性判据下的分析结果。以 $C_{n\beta DYN}>0$ 来看，返回器以40°、50°、60°迎角执行再入时均可维持开环动态偏航稳定，直至马赫数降低至与各迎角相对应的稳定速域下限。迎角越大，稳定速域下限马赫数越小(α=40°、50°、60°时下限马赫数分别为2、1.75、

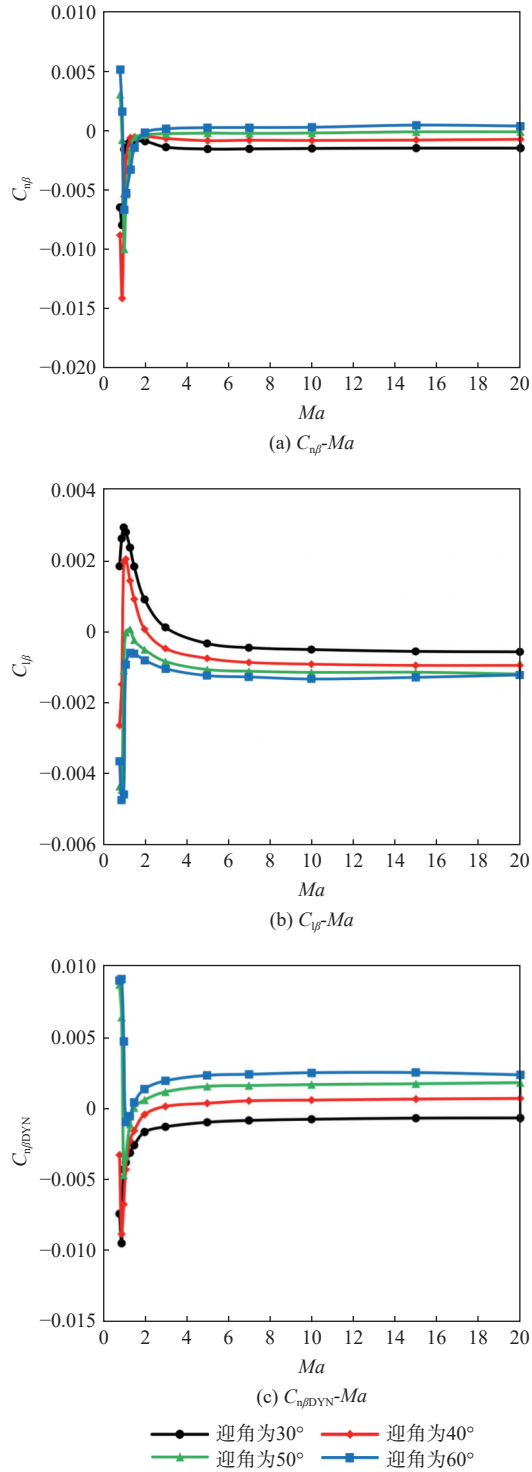


图4 返回器横航向特性随马赫数变化

Fig. 4 Changes in the lateral features of reentry spacecraft with Mach number

1.5), 迎角为 30°以下时在全速域内均无法维持开环动态偏航稳定。

3) 跨声速区间内横航向稳定性发生剧烈变化, 导致横纵向不匹配: 速度降至跨声速区间时, $C_{n\beta}$、$C_{l\beta}$、$C_{n\beta DYN}$ 均出现剧烈变化, 导致返回器横航向特性恶化。随着马赫数的降低, 迎角 40°、50°、60°依次进入开环动态偏航不稳定状态, 如图 4(c)所示, 各马赫数对应的满足开环动态偏航稳定的最小迎角值也开始出现大幅波动。

图 5 为再入迎角走廊, 图中展示了跨声速附近的开环动态偏航稳定迎角下限, 以及升降舵极限偏转(+20°/−10°)所能提供的配平迎角范围。由图可知, 当前横航向稳定最小迎角在部分马赫数下已经超出了升降舵面的配平极限, 造成了横纵向不匹配。由此引发的再入迎角走廊闭塞, 会导致该返回器在任意迎角策略下, 均无法稳定执行再入并减速至 0.8*Ma* 附近的开伞点。

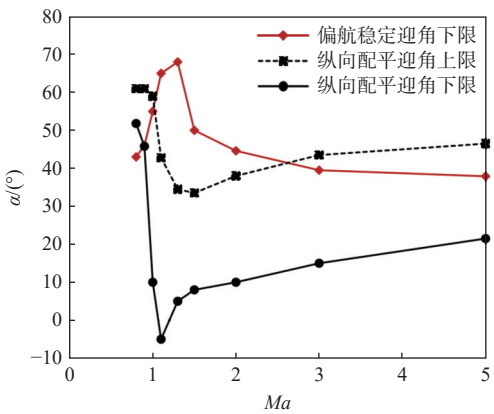


图 5　再入迎角走廊

Fig. 5　Reentry corridor of angle of attack

3　横航向稳定性变化机理分析

根据数值仿真结果可对返回器产生如第 2 节所述的横航向特性变化的机理进行分析。其中, 返回器在马赫数为 5 以上的高超声速流域内的横航向稳定性主要受迎角影响, 且具有良好的规律性: 迎角增加, 返回器压心向后上方移动, 横航向稳定性改善, 如图 6 所示。马赫数为 5 以下的亚跨超声速速域内的横航向稳定性变化则具有更高的复杂性, 其中以跨声速区域的表现最为明显。以 30°迎角为例, 马赫数在 0.7～5 的偏航力矩系数变化如图 7 所示, 该曲线可分为 *AB*、*BC*、*CD* 这 3 个阶段, *AB* 段压心前移, 偏航稳定性恶化; *BC* 段压心迅速后移, 偏航稳定性改善; *CD* 段压心缓慢前移, 压心位置随马赫数的增加而逐渐收敛。

通过流场分析可以获知马赫数在 0.5～5 的压心位置移动机理。图 8 为 30°迎角、4°侧滑角下马

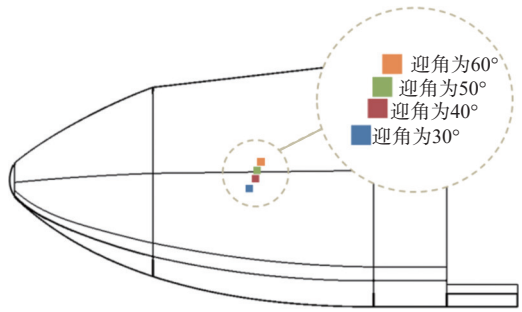


图 6　横航向压心位置随迎角变化(*Ma* = 5)

Fig. 6　Variation of lateral pressure center with angle of attack (*Ma* = 5)

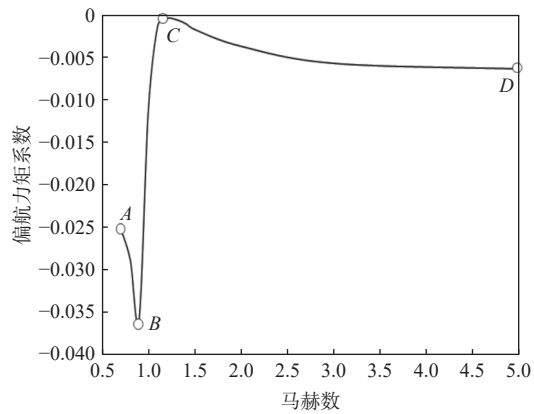


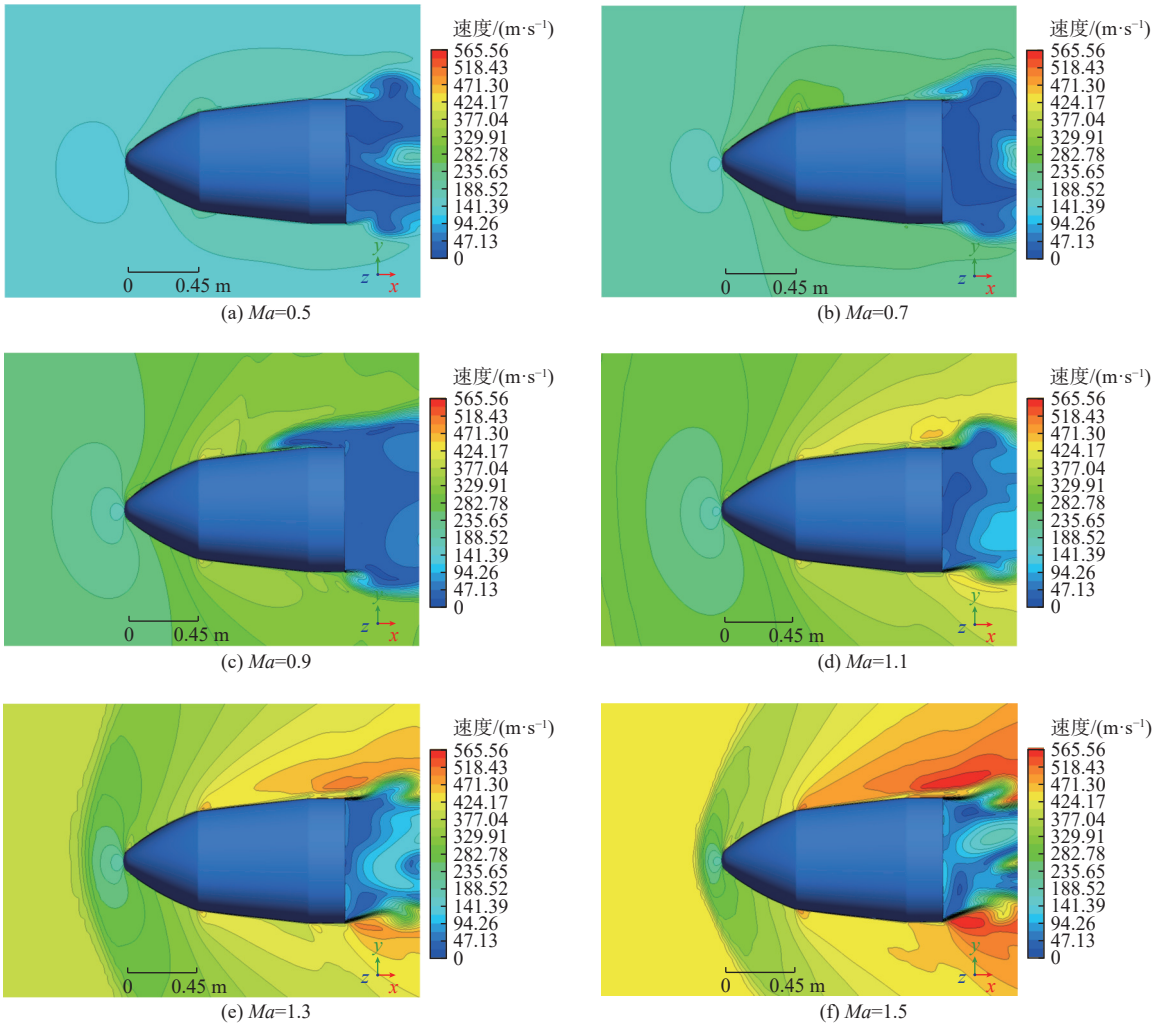
图 7　横航向稳定性变化($\alpha = 30°$)

Fig. 7　Lateral stability variation ($\alpha = 30°$)

赫数分别为 0.5、0.7、0.9、1.1、1.3、1.5 时的流场速度分布。在该迎角下, 当马赫数从亚声速(见图 8(a))增加至高亚声速(见图 8(c))的过程中, 返回器表面逐渐出现局部激波, 激波后流速降低, 在背风一侧诱导出边界层分离(见图 8(c)), 分离区内压强增大、吸力降低, 导致压心前移, 在图 7 中表现为 *AB* 段偏航力矩系数的快速下降; 随着马赫数的继续增加, 背风侧激波位置逐渐后移, 并在 *Ma*=1.1 时到达返回器后缘位置(见图 8(d)), 背风侧分离区在该过程中逐渐消除, 使得压心快速后移, 在图 7 中表现为 *BC* 段偏航力矩系数的明显恢复。该现象与翼型跨声速流动俯仰稳定性变化规律[24] 具有一致性。

马赫数增加至 1.1 以上时, 返回器前方开始出现弓形激波(见图 8(d)和图 8(e)), 且激波距离物面的距离随马赫数的增加而不断缩减, 返回器头部吸力峰逐步增强, 压心向前侧移动, 在图 7 中表现为 *CD* 段偏航力矩系数的缓慢降低。

图 9 为马赫数为 0.5、0.9、1.1 时, 返回器与 *xy* 平面交线上的压力系数分布及马赫数为 1.1、1.3、1.5 时对应的侧向力沿轴向分布(为便于对比进行归一化处理), 从图中可观察到局部激波诱发分离区的产生和消除, 以及超声速区压心随马赫数增加的前移规律。



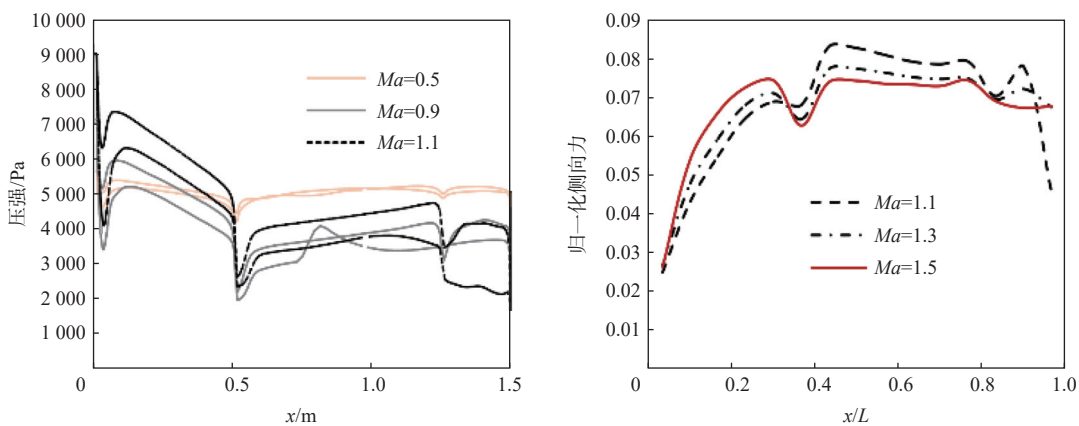
(a) *Ma*=0.5 (b) *Ma*=0.7

(c) *Ma*=0.9 (d) *Ma*=1.1

(e) *Ma*=1.3 (f) *Ma*=1.5

图 8 流场速度分布

Fig. 8 Flow field velocity distribution



(a) 返回器与 *xy* 平面交线压强分布 (b) 侧向力沿轴向分布 (归一化)

图 9 压强及侧向力分布

Fig. 9 Pressure and lateral force distribution

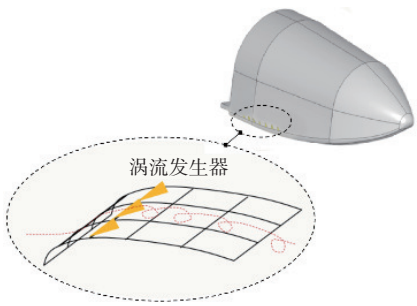
综合来看，由于局部激波、激波诱导边界层分离和激波推移的共同作用，跨声速区域内的横航向稳定性对飞行安全带来了最严峻的挑战，这也是导致返回器无法以第 1 节所述的理想再入策略实施再入的主要原因，因此，需采取增稳措施对该区域内的横航向稳定性进行改善。

4 横航向稳定性控制措施

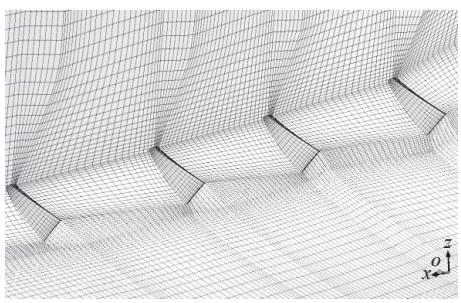
改善跨声速区间内的横航向稳定性可考虑以下 3 种方法：①利用涡流发生器消除因局部激波诱导产生的边界层分离；②使用航向增强型面来提供额外的偏航恢复力矩；③调整重心位置。

4.1　涡流发生器

在分离区域前端加装微型涡流发生器来增加边界层动能是消除或抑制机翼边界层分离的常用方法。本文中的返回器流动分离位置型面与传统机翼存在显著差异, 表现为较大的几何扩张角, 这种大几何扩张角所产生的逆压梯度可能对涡流发生器的应用效果产生影响, 因此, 需使用数值计算方法进行详细分析。所使用的涡流发生器为一种常用的楔形结构, 以串列形式安装在返回器侧部(方式 1)或肩部(方式 2), 如图 10(a)所示, 计算所采用的数值分析方法与 2.1 节相同, 计算网格如图 10(b)所示。



(a) 涡流发生器设置



(b) 数值计算网格

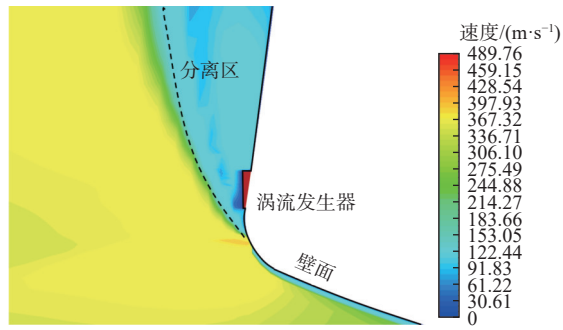
图 10　涡流发生器设置及计算网格

Fig. 10　Vortex generator set and computation mesh

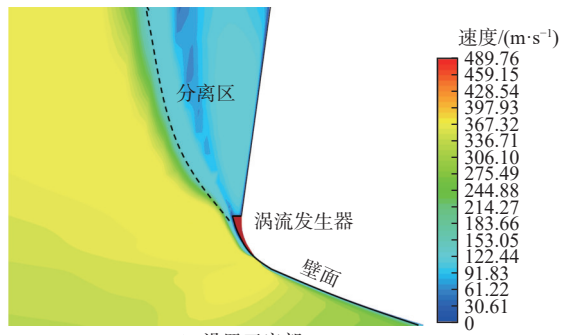
图 11 为数值计算得到的 2 种不同涡流发生器配置下的 yz 平面流场速度分布(对应迎角为 30°、侧滑角为 4°、马赫数为 0.9)。结果显示, 背风侧较大的几何扩张角抑制了涡流向分离区域的有效注入, 使得背风侧边界层分离未受到有效控制或消除。增大涡流发生器的尺寸来改善涡流强度可提升涡流发生器的实际作用, 但也降低了该结构在再入严苛热环境下的生存能力。据此可判定, 利用涡流发生器来消除或抑制边界层分离进而改善跨声速速域内的横航向稳定性, 对当前返回器而言并不具有良好的适应性。

4.2　航向增强型面

在返回器尾部加装航向增强型面形成附加偏航恢复力矩, 是改善跨声速横航向稳定性更直接的方法。可使用的航向增强型面包括航向增强裙边(见图 12(a))、航向舵(见图 12(b))等。其中, 裙边



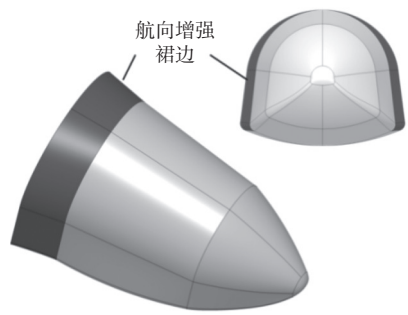
(a) 设置于背风侧



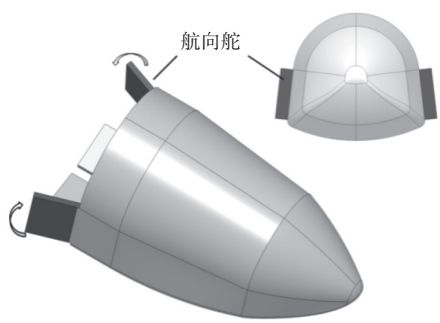
(b) 设置于肩部

图 11　涡流发生器不同安装位置及对应流场速度分布

Fig. 11　Vortex generator locations and corresponding field velocity distribution



(a) 航向增强裙边



(b) 航向舵

图 12　横航向稳定性增强措施

Fig. 12　Lateral stability enhancement measures

构型在原构型基础上增加了侧边几何倾角; 航向舵构型则是将航向舵安装于返回器尾部侧面, 其角度偏转为 0°～20°(外偏为正)。

使用第 3 节所述数值仿真方法, 对 2 种新构型在跨声速区间内的流动特性进行模拟, 可得到流场速度分布如图 13 所示。将图 13 与图 8(c)和图 8(e)

所示的原始构型流场分布进行对比可知，新构型中航向增强型面的使用对激波边界层分离现象的抑制效果不明显，然而图 14 所示的压力分布显示，2 种航向增强型面均不同程度地增加了返回器尾部的吸力峰，产生了额外的偏航恢复力矩，因而对返回器的横航向稳定性产生了改善作用。

图 15 为裙边构型、航向舵 0°及 20°偏转构型在全速域、30°迎角下的的横航向特性计算结果，且与

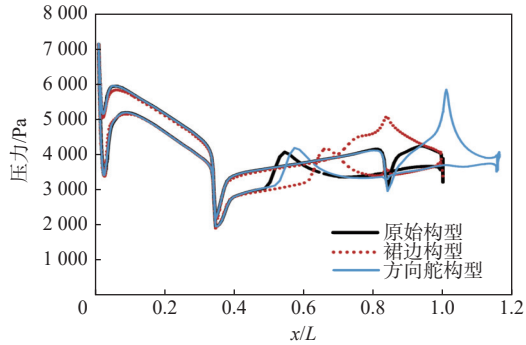
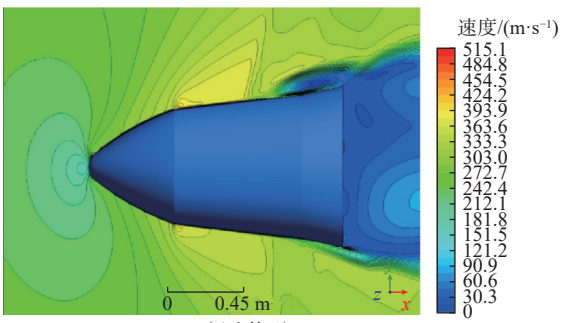
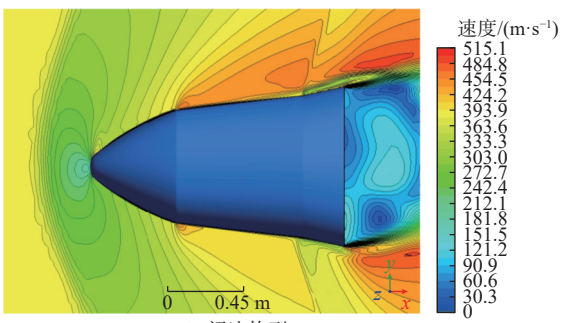


图 14 不同构型返回器与 *xy* 平面交线压力分布对比

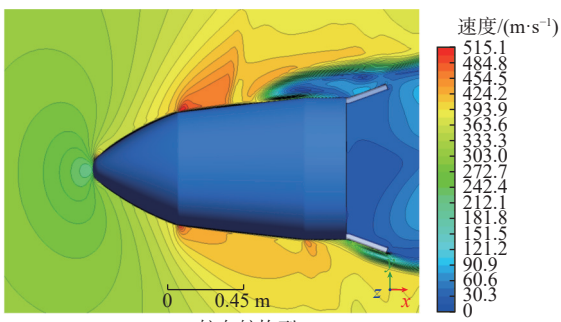
Fig. 14 Pressure distribution on intersection lines of *xy* plane and reentry spacecraft with different configurations



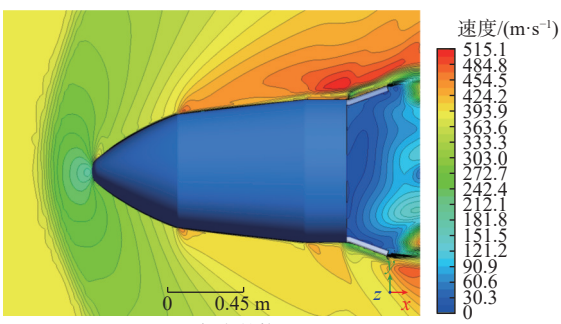
(a) 裙边构型: *Ma*=0.9



(b) 裙边构型: *Ma*=1.3



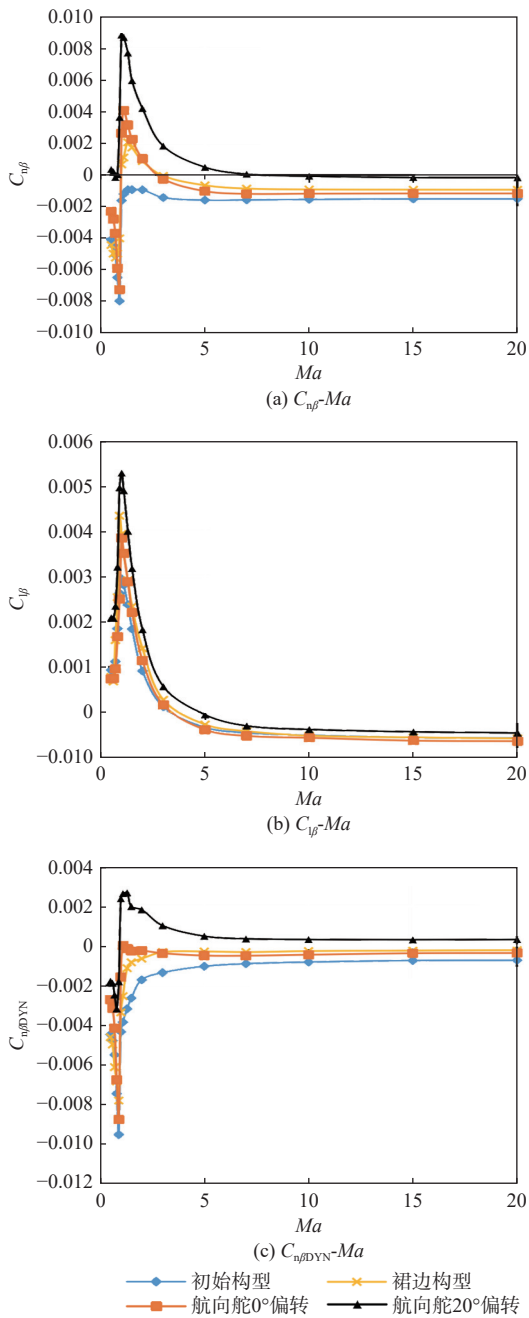
(c) 航向舵构型: *Ma*=0.9



(d) 航向舵构型: *Ma*=1.3

图 13 使用航向增强型面流场速度分布

Fig. 13 Flow field velocity distribution with lateral stability enhancement surfaces



(a) $C_{n\beta}$-*Ma*

(b) $C_{l\beta}$-*Ma*

(c) $C_{n\beta DYN}$-*Ma*

图 15 航向增强返回器横航向特性计算结果分析

Fig. 15 Calculated lateral feature of reentry spacecraft lateral stability enhancement analysis

原始构型进行了比较。数据显示, 裙边方案的横航向增稳效果与 0°航向舵接近, 而航向舵 20°偏转对横航向稳定性的改善效果最为明显。

在航向舵 20°偏转状态下, 返回器在 30° 迎角下的偏航稳定性从全速域不稳定变为在马赫数 0.8 以上稳定($C_{n\beta}$>0)或临界稳定($C_{n\beta}$≈0), 但由于航向舵位置靠下, 因此其横向稳定性出现小幅恶化(横向稳定的最小临界马赫数由 3 提升到 5)。从开环动态偏航稳定性的综合判据($C_{n\beta DYN}$>0)分析, 添加航向舵并偏转 20°可将返回器开环动态偏航稳定的最低速度降至 0.95*Ma*, 此时返回器稳定性和初始外形相比已明显提升, 但依然无法保证返回器稳定降速至开伞点的状态(约 0.8*Ma*)。

4.3　调整重心位置

改变横航向特性的第 3 种方法是调整重心位置。对返回器来说, 重心前后偏移主要影响偏航特性和俯仰特性, 上下偏移主要改变滚转特性。在给定的迎角下, 将重心向前配置既增加了偏航恢复力矩, 又增加了俯仰低头力矩, 偏航恢复力矩的增加有利于改善航向稳定性, 但俯仰低头力矩的增加会降低纵向配平迎角, 这会对偏航恢复力矩产生不利影响(见第 2.2 节)。两者的共同作用限制了重心前后移动对横航向稳定性的改善作用, 因此, 在利用重心偏移来改善横航向稳定性时, 主要考虑重心的上下移动。

图 16 以 30°迎角为例, 给出了重心下偏 6%(Z_{cg}/L=6%)时, 各构型状态返回器横航向特性随马赫数的变化情况。结果显示, 重心向下偏移 6%, 可在不影响航向稳定性(见图 16(a))的前提下明显改善横向稳定性(见图 16(b)), 从而将返回器的开环动态偏航的稳定性(见图 16(c)和图 16(d))显著改善: 在 30°迎角下, 初始构型、裙边构型、航向舵 0°偏转、航向舵 20°偏转构型所对应的横航向稳定速域下限马赫数分别为 1.5、1.0、0.99、0.82。其中, 航向舵外偏 20°对应的不稳定临界马赫数已达到亚声速开伞点附近, 这意味着在该构型状态下, 返回器可以第 2 节理想再入策略执行再入飞行。

尽管航向舵 20°偏转方案为跨声速区间内的横航向稳定性改善提供了最有效的途径, 但该方案同时也会对返回器气动特性、防热性能产生不利影响

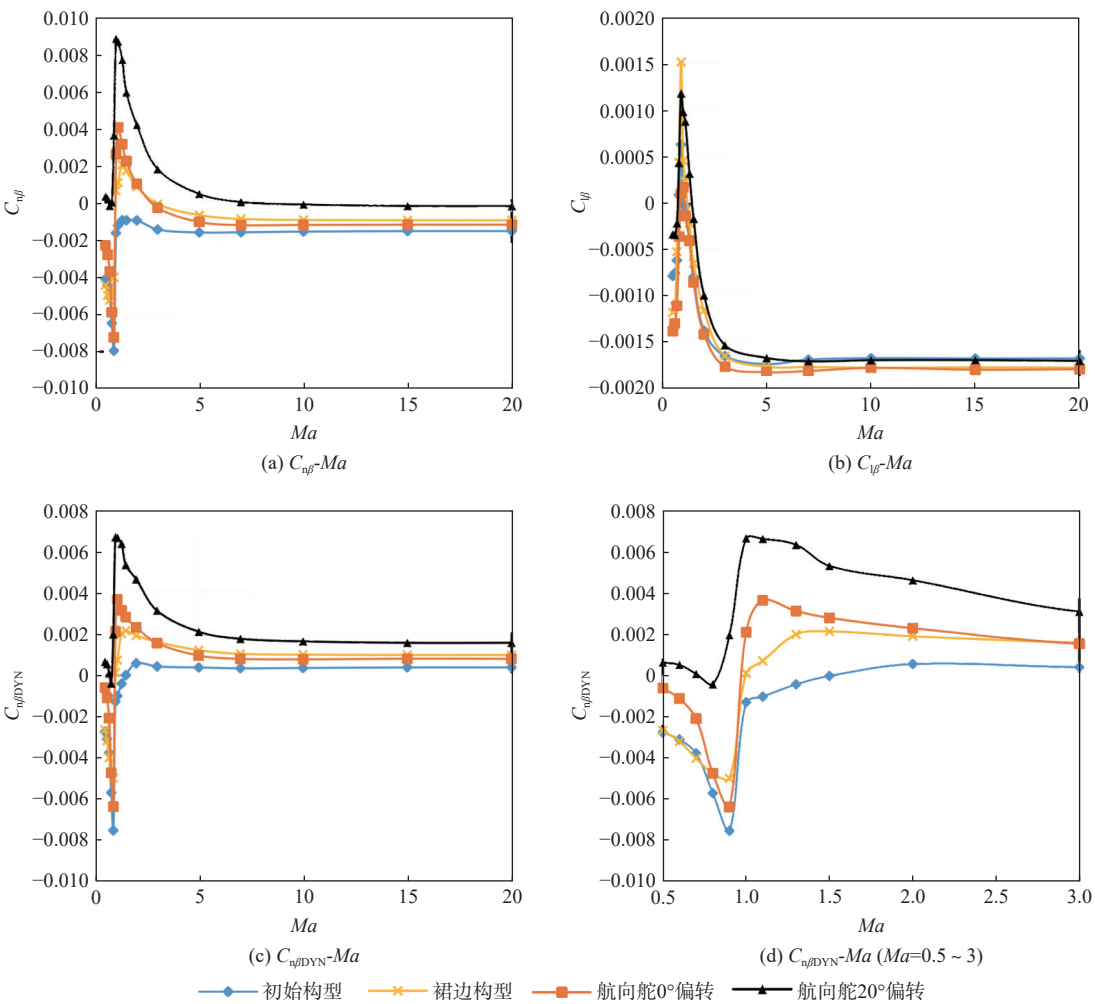


图 16　重心下偏 6% 返回器横航向特性计算结果

Fig. 16　Calculated lateral features of reentry spacecraft center-of-gravity offset 6%

(不同迎角下的升阻比降低3%～5%)。与此同时,通过进一步分析图16可知,在相同重心偏置下,航向舵0°偏转偏在超声速、高超声速阶段亦可保持足够的横航向稳定性,因此,在制定再入策略时,可考虑从航向舵0°(超声速、高超声速阶段)向航向舵外偏20°(跨声速阶段)的切换过程,以实现系统整体效能的最大化。

5 结 论

1) 类双锥型升力式返回器的横航向稳定性在跨声速阶段存在显著恶化,其许用迎角下限超出了升降舵面的配平极限,如无增稳措施将无法实现安全开伞。

2) 该类返回器横航向稳定性波动机理在不同速域内存在显著差异,其中,亚声速阶段为背风侧局部激波的生成和边界层的诱导分离;跨声速阶段为激波推移和分离区的消除;超声速阶段为头部弓形激波的生成和移动。

3) 可加装航向增强型面(裙边、航向舵等)或实行重心偏置来改善该类返回器的横航向稳定性,其中,舵向舵20°偏转、重心下偏6%为将不稳定临界马赫数降至亚声速开伞点附近。

4) 航向舵偏转对返回器气动特性、防热性能存在不利影响,在制定再入策略时可考虑从0°偏转(超声速/高超声速)向20°偏转(亚/跨声速)的控制阶段切换,以优化系统整体效能。

参考文献(References)

[1] VASHCHENKOV P, KASHKOVSKY A, DYADKIN A, et al. Numerical study of high-altitude aerodynamics of the clipper reentry vehicle[C]//Proceedings of the 46th AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2008: 1188.

[2] MOSELEY W C, GRAHAM R E, HUGHES J E. Aerodynamic stability characteristics of the apollo command module: NASA TN D-4688[R]. Houston: Manned Spacecraft Center, 1968: 171-175.

[3] ABE T, SATO S I, MATSUKAWA Y, et al. Study for dynamically unstable motion of reentry capsule[C]//Proceedings of the 34th Thermophysics Conference. Reston: AIAA, 2000: 2589.

[4] 张涵信, 袁先旭, 叶友达, 等. 飞船返回舱俯仰振荡的动态稳定性研究[J]. 空气动力学学报, 2002, 20(3): 247-259.
ZHANG H X, YUAN X X, YE Y D, et al. Research on the dynamic stability of an orbital reentry vehicle in pitching[J]. Acta Aerodynamica Sinica, 2002, 20(3): 247-259(in Chinese).

[5] 袁先旭, 张涵信, 谢昱飞. 飞船返回舱再入俯仰动稳定吸引子数值仿真[J]. 空气动力学学报, 2007, 25(4): 431-436.
YUAN X X, ZHANG H X, XIE Y F. Numerical simulation for dynamic stability in pitching of unfinned reentry capsule and bifurcation with Mach number prediction[J]. Acta Aerodynamica Sinica, 2007, 25(4): 431-436(in Chinese).

[6] TERAMOTO S, FUJII K. Study on the mechanism of the instability of a re-entry capsule at transonic speeds[C]//Proceedings of the Fluids Conference and Exhibit. Reston: AIAA, 2000: 2603.

[7] TERAMOTO S, HIRAKI K, FUJII K. Numerical analysis of dynamic stability of a reentry capsule at transonic speeds[J]. AIAA Journal, 2001, 39(4): 646-653.

[8] 朱德华, 沈清, 杨武兵. 返回舱高雷诺数再入过程底部流动稳定性[J]. 力学学报, 2021, 53(3): 752-760.
ZHU D H, SHEN Q, YANG W B. Base flow stability of return capsule during reentry process with high Reynolds number[J]. Chinese Journal of Theoretical and Applied Mechanics, 2021, 53(3): 752-760(in Chinese).

[9] MITCHELTREE R, FREMAUX C, YATES L. Subsonic static and dynamic aerodynamics of blunt entry vehicles[C]//Proceedings of the 37th Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 1999: 1020.

[10] 何开锋, 和争春. 飞船返回舱跨声速全局稳定性研究[J]. 飞行力学, 1999, 17(3): 34-38.
HE K F, HE Z C. The global stability analysis of reentry capsule transonic flight[J]. Flight Dynamics, 1999, 17(3): 34-38(in Chinese).

[11] 宋玉辉, 陈农, 秦永明. 亚跨超声速返回舱动稳定特性[J]. 航天返回与遥感, 2014, 35(2): 31-38.
SONG Y H, CHEN N, QIN Y M. Sub-, trans- and super-sonic dynamic stability characteristics for reentry capsule[J]. Spacecraft Recovery & Remote Sensing, 2014, 35(2): 31-38(in Chinese).

[12] 胡静, 宋玉辉, 陈农, 等. 返回舱动稳定特性风洞试验的影响参数[J]. 航天返回与遥感, 2013, 34(5): 14-19.
HU J, SONG Y H, CHEN N, et al. Study of key parameters for dynamic stability of reentry capsule[J]. Spacecraft Recovery & Remote Sensing, 2013, 34(5): 14-19(in Chinese).

[13] 宋威, 艾邦成, 蒋增辉, 等. 返回舱跨声速自由飞行的静动稳定性[J]. 实验流体力学, 2019, 33(4): 89-94.
SONG W, AI B C, JIANG Z H, et al. Free flight static and dynamic aerodynamic characteristics for re-entry capsule at transonic speed[J]. Journal of Experiments in Fluid Mechanics, 2019, 33(4): 89-94(in Chinese).

[14] 鲍文, 姚照辉. 综合离心力/气动力的升力体高超声速飞行器纵向运动建模研究[J]. 宇航学报, 2009, 30(1): 128-133.
BAO W, YAO Z H. Study on longitudinal modeling for integrated centrifugal/aero force lifting-body hypersonic vehicles[J]. Journal of Astronautics, 2009, 30(1): 128-133(in Chinese).

[15] 贾子安, 张陈安, 王柯穆, 等. 乘波布局高超声速飞行器纵向静稳定特性分析[J]. 中国科学: 技术科学, 2014, 44(10): 1114-1122.
JIA Z A, ZHANG C A, WANG K M, et al. Longitudinal static stability analysis of hypersonic waveriders[J]. Scientia Sinica (Technologica), 2014, 44(10): 1114-1122(in Chinese).

[16] 贾子安. 乘波飞行器气动布局初步设计研究[D]. 北京: 中国科学院大学, 2014: 56-75.
JIA Z A. Preliminary design of aerodynamic layout of waverider vehicle[D]. Beijing: University of Chinese Academy of Sciences, 2014: 56-75(in Chinese).

[17] 王柯穆. 幂次体乘波飞行器纵向静稳定及升阻比优化设计[D]. 北京: 中国科学院大学, 2013.
WANG K M. Longitudinal static stability and optimal design of lift-drag ratio of power-law waverider[D]. Beijing: University of

Chinese Academy of Sciences, 2013(in Chinese).

[18] MANGIN B, BENAY R, CHANETZ B, et al. Optimization of viscous waveriders derived from axisymmetric power-law blunt body flows[J]. Journal of Spacecraft and Rockets, 2006, 43(5): 990-998.

[19] 高清, 赵俊波, 李潜. 类 HTV-2 横侧向稳定性研究[J]. 宇航学报, 2014, 35(6): 657-662.
GAO Q, ZHAO J B, LI Q. Study on lateral-directional stability of HTV-2 like configuration[J]. Journal of Astronautics, 2014, 35(6): 657-662(in Chinese).

[20] 高清, 李潜. 美国高超声速飞行器横侧向稳定性研究[J]. 飞航导弹, 2012(12): 14-18.
GAO Q, LI Q. Study on lateral stability of American hypersonic vehicle[J]. Aerodynamic Missile Journal, 2012(12): 14-18(in Chinese).

[21] 祝立国, 王永丰, 庄逢甘, 等. 高速高机动飞行器的横航向偏离预测判据分析[J]. 宇航学报, 2007, 28(6): 1550-1553.
ZHU L G, WANG Y F, ZHUANG F G, et al. The lateral-directional departure criteria analysis of high-speed and high maneuverability aircraft[J]. Journal of Astronautics, 2007, 28(6): 1550-1553(in Chinese).

[22] 闵昌万. 高超声速飞行器横侧向气动布局准则研究[J]. 宇航总体技术, 2018, 2(3): 1-10.
MIN C W. The aerodynamic configuration criteria for the lateral-directional stability of hypersonic vehicle[J]. Astronautical Systems Engineering Technology, 2018, 2(3): 1-10(in Chinese).

[23] 王亮, 周玲. 基于改进的 k-ω-γ 转捩模式预测高超声速飞行器气动特性[J]. 空气动力学学报, 2021, 39(3): 51-61.
WANG L, ZHOU L. Prediction of aerodynamic characteristics of hypersonic vehicle by improved k-ω-γ transition model[J]. Acta Aerodynamica Sinica, 2021, 39(3): 51-61(in Chinese).

[24] 钱翼稷. 空气动力学[M]. 北京: 北京航空航天大学出版社, 2004: 287-289.
QIAN Y J. Aerodynamics[M]. Beijing: Beihang University Press, 2004: 287-289(in Chinese).

Analysis and improvement of lateral instability of quasi-biconical lifting reentry spacecraft

CHEN Gang[1,*], SUN Xue[1], LI Guangxing[1], ZHAO Dan[1], MA Jianying[1], HE Cheng[2]

(1. Aerospace System Engineering Shanghai, Shanghai Academy of Spaceflight Technology, Shanghai 201109, China;
2. Unmanned Aerial Vehicles Research Institute, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: The lateral stability of the quasi-biconical lifting reentry spacecraft is a strict constraint on the implementation of its ideal reentry strategy, which requires specialized analysis in the design stage. In this paper, for the nominal reentry process of the quasi-biconical lifting reentry spacecraft, the lateral stability of the spacecraft in the range of full speed and wide angle of attack was calculated by using the numerical simulation method. According to the flow field characteristics, the mutation regularity of the lateral stability in the sub-/trans-/supersonic domain was analyzed. Then, the effect of the eddy current generator, lateral stability enhancement surface, and center-of-gravity offset on the lateral stability was analyzed, as well as the corresponding mechanism. The results reveal that the combination of rudder deflection and center-of-gravity offset is the optimal solution to enhance lateral stability.

Keywords: lifting reentry spacecraft; supersonic; lateral stability; flow mechanism; stability enhancement measure

Received: 2022-08-11; **Accepted**: 2022-10-04; **Published Online**: 2022-10-10 08:44
URL: link.cnki.net/urlid/11.2625.V.20221009.0850.001

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0726

基于中国鸟情的鸟撞适航条款充分性和适宜性

陆晓华[1,*]，蔡景[1]，张柱国[2]，张迎春[2]

(1. 南京航空航天大学 民航学院，南京 211106；2. 中国民用航空上海航空器适航审定中心，上海 200335)

摘 要：以中国民航近十余年运输类飞机鸟撞信息为数据源，利用混合威布尔函数拟合鸟撞能量分布，并应用列文伯格-马夸尔特（L-M）寻优算法得到更为精确的拟合分布参数估计值。对照安全性指标，通过鸟撞能量分布函数评估出现有鸟撞适航条款对机翼而言是具有充分性和适宜性的，对尾翼而言具有充分性但过于保守。以发生灾难性事故的安全性指标为依据，中国鸟情环境下机翼结构抗鸟撞鸟体质量最低适航要求为1.218 kg；同时验证了美国联邦航空局（FAA）在美国鸟情环境下提高机翼结构抗鸟撞鸟体质量最低适航要求的愿景（约3.6 kg），证明了基于两重两参数混合威布尔分布的鸟撞冲击能量统计分析方法的有效性及结果的可信度，为中国民航自主修订相关适航条款修订提供了理论依据和实践参考。

关 键 词：鸟情环境；适航条款；鸟撞能量分布；评估；充分性；适宜性；安全性指标

中图分类号：V221^{+}.1; X915.1

文献标志码：A **文章编号**：1001-5965(2024)09-2810-09

鸟的种类、数量及其地域分布尤其是机场周边的鸟情环境对飞机的安全运行有较大的影响，飞机结构抗鸟撞设计是保障飞机安全运行的重要研究内容。美国联邦航空局(Federal Aviation Administration, FAA)颁布的联邦航空规章(Federal Aviation Regulations, FAR)25.631条款规定，撞击飞机尾翼的鸟体质量上限为8 lb(1 lb≈0.45 kg)[1]；FAA颁布的FAR 25.571(e)(1)条款中将鸟撞作为一种离散源损伤，规定飞机一般结构与风挡的抗鸟撞要求要保持一致(即4 lb)[1]。欧洲航空安全局(European Aviation Safety Agency, EASA)颁布的25部审定规范(Certification Specification, CS)中对飞机机翼和尾翼抗鸟撞鸟体质量的要求都是4 lb[2]。随着欧美各国飞机服役经验和鸟撞案例的不断积累，FAA认为现代飞机使用了很多新技术，包括一些新的关键控制系统和复合材料结构，这些新技术的部位对鸟撞冲击载荷更为敏感，而在最初设计抗鸟撞标准时未考虑这类新技术；此外，最新的美国鸟撞数据显示，所有的鸟撞事件中有7.2%的鸟体质量超过4 lb，有3.6%的鸟体质量超过8 lb，同时鸟的数量也在持续增加，因此鸟撞事件的发生概率在持续增大[3]。因此，FAA从1987年开始就有意将飞机所有结构的抗鸟撞标准提高至8 lb。但欧洲联合航空局(Joint Aviation Agency, JAA, 2002年欧盟成立了具有法律权限的EASA机构，并全面接替了原JAA的职能)认为提高飞机抗鸟撞适航标准无法从技术角度得到证实，且欧洲的鸟撞事件统计表明，只有1.2%的鸟体质量超过4 lb、0.3%的鸟体质量超过8 lb[3]。因此，JAA决定将对飞机尾翼抗鸟撞适航要求与FAR 25.631[1]条款保持差异。尽管FAA从1992年起就联合JAA成立特别工作组来协调鸟撞相关适航条款的差异性问题，但双方始终没有达成一致。美国和欧洲的飞机制造商本着协调一致的精神也分别满足了FAA和EASA的现行要求。中国民航规章(China Civil Aviation Regulation, CCAR)-25[4]部规章和FAR-25[1]部规章除了测量单位换算

收稿日期：2022-08-17；**录用日期**：2022-11-04；**网络出版时间**：2022-12-02 11:12
网络出版地址：link.cnki.net/urlid/11.2625.V.20221201.1457.002
基金项目：国家自然科学基金(U1933202)；中国民航局安全能力建设基金(2021-198)
* **通信作者**. E-mail：luxiaohua@nuaa.edu.cn

引用格式：陆晓华，蔡景，张柱国，等. 基于中国鸟情的鸟撞适航条款充分性和适宜性[J]. 北京航空航天大学学报，2024，50(9)：2810-2818.
LU X H，CAI J，ZHANG Z G，et al. Adequacy and suitability of airworthiness clause of bird strike based on bird situation in China[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2810-2818 (in Chinese).

差异外, 在鸟撞要求上无实质性差异。CAAR/FAR 25.571(e)[1,4] 和 CS 25.571[2] 条款中对离散源损伤项中的鸟撞要求相同, 只是 CS 25.571[2] 条款参照 CS 25.631[2] 条款中的规定执行。按照 FAA 颁布的鸟撞适航条款进行研制的飞机在中国特定的鸟情环境下能否达到所需的安全性水平或当前的适航条款规定是否又过于保守, 这需要对鸟撞适航条款的充分性和适宜性进行研究, 特别是随着中国民用飞机自主研制的推进和商业运营的投入, 制定/修订适宜的鸟撞适航条款是降低鸟撞安全风险和提高研运维制经济性的重要任务之一[5]。

国内外关于鸟撞安全性分析的研究成果主要包括基于鸟撞事件统计的风险评估和关联性分析、根据鸟撞适航条款进行的试验方法、仿真技术及相关优化改进设计研究、鸟撞适航规章的解读等内容。在鸟撞风险评估方面, 文献 [6-8] 统计分析了鸟撞事件的发生率、损伤率及鸟击发生的时间/时段、飞行高度/阶段、鸟体质量、鸟群种类之间的关联; 文献 [9-12] 分别应用层次分析法、模糊数学法、马尔可夫链蒙特卡罗法、矩阵法对机场周边和飞行航路上的鸟撞事件进行了统计分析和风险评估。在鸟撞试验和仿真计算方面, 文献 [13-14] 应用复合材料、人工鸟弹等新技术进行了鸟撞试验和仿真研究; 文献 [15-16] 对飞机的关键系统和重要结构进行了抗鸟撞分析。在鸟撞适航规章的解读方面, 文献 [17] 对鸟撞适航条款的修订背景进行了简要分析; 文献 [18] 针对鸟撞适航要求提出了通用的符合性设计方法及流程; 文献 [19] 对运输类飞机抗鸟撞要求的试验验证和符合性分析方法的认可审定作了进一步的解释说明; 美国材料试验学会将根据鸟撞适航条款颁布的航空透明件鸟撞标准试验方法中规定的鸟体质量、鸟撞速度、试验设备、试验环境等要求都用于指导飞机其他结构的鸟撞试验[20]。

从实际鸟情统计讨论规章条款的充分性和适宜性在国内比较少见, 在国外也未有详细的研究成果。尽管目前中国运输类飞机鸟撞案例信息存在不完备等情况, 但现有的鸟撞案例信息对于研究鸟环境对运输类飞机安全性的影响具有重要的参考价值。本文在收集整理近十余年中国民航鸟击航空器信息的基础上, 通过走访航司航线机械员和飞行员、咨询机场防鸟驱鸟人员及查阅相关飞行手册资料, 选取鸟击数据中各机型在典型航段、航路上不同飞行阶段的平均值作为该飞行阶段的代表速度值, 统计分析鸟体的撞击能量分布规律, 并研究现有鸟撞适航条款在中国鸟情环境下的充分性和适宜性。

1　鸟撞冲击能量的统计分析

1.1　鸟撞冲击能量值确定

鸟撞冲击能量用鸟体撞击飞机前的瞬时鸟体动能表示, 即 $E = mv^2/2$, 其中, m 为鸟体质量, v 为两者相对撞击速度。为确保飞机在遭受鸟撞后仍能完成安全飞行任务, CCAR 对鸟撞发生时的鸟体质量和速度的要求为[4]: ①CCAR 25.571(一般结构), 鸟体质量为 1.8 kg, 飞机与鸟沿飞机飞行航迹的相对速度取海平面 V_c 或海拔 2 450 m 的 $0.85V_c$, 选两者中较严重者; ②CCAR 25.631(尾翼), 鸟体质量为 3.6 kg, 飞机的速度(相对于鸟沿飞机飞行航迹)等于按 CCAR 25.335(a) 选定的海平面 V_c。

发生鸟撞时的速度主要取决于飞机当时的飞行速度, 虽然通常情况下鸟撞发生时的飞行速度常难以精确获取, 但多数情况下可以基本确认鸟撞发生的飞行阶段, 则可取该机型在某飞行阶段的平均估计速度作为鸟撞发生时的速度。

1.2　鸟撞冲击能量分布

1.2.1　混合威布尔分布特征确定

两重混合威布尔分布定义: 如果一个总集由 2 个子集组成, 则各子集均服从两参数威布尔分布, 但参数值不同。假设各子集的概率密度函数分别为 $f_1(t)$ 和 $f_2(t)$, 各子集的混合权重分别为 p_1 和 p_2, 且满足 $p_1 + p_2 = 1$, 则总集的概率密度函数为

$$f(t) = p_1 f_1(t) + p_2 f_2(t) \tag{1}$$

式中: t 为概率密度函数的随机变量。

根据上述定义, 每个威布尔分布子集的概率密度函数 $f_i(t)$ 表达式为

$$f_i(t) = \frac{\beta_i}{\eta_i} t^{\beta_i - 1} \exp\left[-(t/\eta_i)^{\beta_i}\right] \qquad i = 1,2 \tag{2}$$

式中: β_i 和 η_i 分别为第 i 个威布尔分布子集的形状参数和比例参数。

累积分布函数 $F_i(t)$ 为

$$F_i(t) = 1 - \exp\left[-(t/\eta_i)^{\beta_i}\right] \qquad i = 1,2 \tag{3}$$

两重两参数混合威布尔分布可靠度函数可表示为

$$R(t) = p_1 R_1(t) + p_2 R_2(t) = p_1 \exp\left[-(t/\eta_1)^{\beta_1}\right] + p_2 \exp\left[-(t/\eta_2)^{\beta_2}\right] \tag{4}$$

将式(3)变换为

$$\frac{1}{1 - F(t)} = \exp(t/\eta)^{\beta} \tag{5}$$

对式(5)两边取 2 次自然对数, 即

$$\ln\ln\frac{1}{1 - F(t)} = \beta\ln(t) - \beta\ln\eta \tag{6}$$

令

$$\begin{cases} y=\ln\ln[1/(1-F(t))] \\ x=\ln(t) \\ b=-\beta\ln\eta \\ w=\beta \end{cases} \tag{7}$$

将式(7)代入式(6),得到

$$y=wx+b \tag{8}$$

式(8)即转化为二维直角坐标系中的直线方程,其中,w为斜率,b为截距。按照从小到大的顺序排列鸟撞冲击能量,并通过下列中位秩式(9)计算鸟击航空器累积失效概率的估计值$\tilde{F}(t_i)$。取鸟撞冲击能量的自然对数值为x轴,以无量纲值$\ln\ln\left[1/\left(1-\tilde{F}(t)\right)\right]$为$y$轴,如图1所示。

$$\tilde{F}(t_i)=\frac{i-0.3}{n+0.4}\qquad i=1,2,\cdots,n \tag{9}$$

式中:n为鸟撞冲击能量数据的数量。

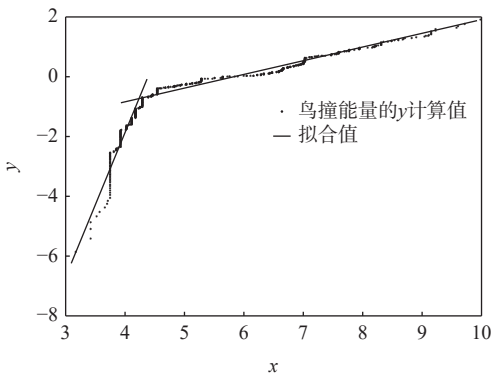


图1 鸟撞冲击能量拟合直线

Fig. 1 Bird strike energy fitting line

图1中的鸟撞冲击能量拟合结果显示,2个子集分别对应上图中的2条逼近直线,且交汇于原始数据附近的一个交点。由此初步判定,鸟撞冲击能量数据基本符合两重两参数混合威布尔分布。

1.2.2 混合威布尔分布参数估计

由式(4)可知,两重两参数混合威布尔分布的可靠度概率函数可表示为

$$P(t)=p_1R_1(t)+p_2R_2(t)=p_1\exp\left[-(t/\eta_1)^{\beta_1}\right]+p_2\exp\left[-(t/\eta_2)^{\beta_2}\right] \tag{10}$$

以$\boldsymbol{x}=[x_1,x_2,x_3,x_4,x_5]^{\mathrm{T}}$表征两重两参数混合威布尔分布函数的参数向量,其中,各分量分别代表参数p_1、η_1、β_1、η_2和β_2。则式(10)可改写为

$$P(t,\boldsymbol{x})=x_1\exp\left[-\left(\frac{t}{x_2}\right)^{x_3}\right]+(1-x_1)\exp\left[-\left(\frac{t}{x_4}\right)^{x_5}\right] \tag{11}$$

由式(9)可得鸟撞冲击能量的概率$\tilde{P}(t)$观测值为

$$\tilde{P}(t_i)=1-\tilde{F}(t_i)=1-\frac{i-0.3}{n+0.4}\qquad i=1,2,\cdots,n \tag{12}$$

建立最优参数$\boldsymbol{x}$的优化目标函数为

$$\min\lambda(\boldsymbol{x})=\frac{1}{2}\sum_{i=1}^{n}r_i^2(\boldsymbol{x}) \tag{13}$$

式中:$r_i(\boldsymbol{x})$为残差量,即$r_i(\boldsymbol{x})=\tilde{P}(t_i)-P(t_i,\boldsymbol{x})$,$i=1,2,\cdots,n$。

根据混合威布尔分布函数的性质,各参数值应满足下列约束条件:$0<p_1<1$,$\eta_1>0$,$\beta_1>0$,$\eta_2>0$,$\beta_2>0$。

列文伯格-马夸尔特法寻优算法[21]是一种无约束优化算法,其核心为:在Gauss-Newton方向和最速下降方向之间采用二阶微分和方向矢量不断调整函数计算的收敛方向,从而获得更快的收敛速度。应用列文伯格-马夸尔特(Levenberg-Marquardt,L-M)算法进行无约束优化求解目标函数的解,拟通过式(14)和式(15)进行参数变换去除约束条件。令

$$x_1=\frac{\exp(\theta_1)}{\exp(\theta_1)+\exp(-\theta_1)} \tag{14}$$

$$x_j=\exp\left(\theta_j\right)\qquad j=2,3,4,5 \tag{15}$$

令$\boldsymbol{\theta}=[\theta_1,\theta_2,\theta_3,\theta_4,\theta_5]^{\mathrm{T}}$,将式(14)和式(15)代入式(11),即

$$P(t,\boldsymbol{\theta})=\frac{\exp(\theta_1)}{\exp(\theta_1)+\exp(-\theta_1)}\exp\left[-\left(\frac{t}{\exp(\theta_2)}\right)^{\exp(\theta_3)}\right]+\left(1-\frac{\exp(\theta_1)}{\exp(\theta_1)+\exp(-\theta_1)}\right)\exp\left[-\left(\frac{t}{\exp(\theta_4)}\right)^{\exp(\theta_5)}\right] \tag{16}$$

经过变换,式(13)转变为无约束优化问题:

$$\min\lambda(\boldsymbol{\theta})=\frac{1}{2}\sum_{i=1}^{n}r_i^2(\boldsymbol{\theta})=\frac{1}{2}\sum_{i=1}^{n}\left(\tilde{P}(t_i)-P(t_i,\boldsymbol{\theta})\right)^2 \tag{17}$$

求得$\boldsymbol{\theta}$的最优解后,可通过式(14)和式(15)求出鸟撞冲击能量数据两重两参数混合威布尔分布的5个未知参数。

L-M算法中非线性最小二乘迭代法为

$$\boldsymbol{\theta}^{(k+1)}=\boldsymbol{\theta}^{(k)}-\left[\boldsymbol{J}\left(\boldsymbol{\theta}^{(k)}\right)^{\mathrm{T}}\boldsymbol{J}\left(\boldsymbol{\theta}^{(k)}\right)+l^{(k)}\times\mathrm{diag}\left(\boldsymbol{J}\left(\boldsymbol{\theta}^{(k)}\right)^{\mathrm{T}}\boldsymbol{J}\left(\boldsymbol{\theta}^{(k)}\right)\right)\right]^{-1}\boldsymbol{J}\left(\boldsymbol{\theta}^{(k)}\right)^{\mathrm{T}}\boldsymbol{r}\left(\boldsymbol{\theta}^{(k)}\right) \tag{18}$$

式中:k为迭代增量变量;$l^{(k)}$为第k次迭代下的步长;$\boldsymbol{\theta}^{(k)}$为第$k$次迭代下的参数估计值向量;$\boldsymbol{r}(\boldsymbol{\theta})$为残差矩阵,即$\boldsymbol{r}(\boldsymbol{\theta})=[r_1(\boldsymbol{\theta}),r_1(\boldsymbol{\theta}),\cdots,r_n(\boldsymbol{\theta})]$;$\boldsymbol{J}\left(\boldsymbol{\theta}^{(k)}\right)$为目标函数$\lambda(\boldsymbol{\theta})$在$\boldsymbol{\theta}^{(k)}$处的Jacobi矩阵,即

$$\boldsymbol{J}(\boldsymbol{\theta})^{\mathrm{T}}=\begin{bmatrix} \dfrac{\partial r_1(\boldsymbol{\theta})}{\partial\theta_1} & \dfrac{\partial r_2(\boldsymbol{\theta})}{\partial\theta_1} & \cdots & \dfrac{\partial r_n(\boldsymbol{\theta})}{\partial\theta_1} \\ \dfrac{\partial r_1(\boldsymbol{\theta})}{\partial\theta_2} & \dfrac{\partial r_2(\boldsymbol{\theta})}{\partial\theta_2} & \cdots & \dfrac{\partial r_n(\boldsymbol{\theta})}{\partial\theta_2} \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial r_1(\boldsymbol{\theta})}{\partial\theta_m} & \dfrac{\partial r_2(\boldsymbol{\theta})}{\partial\theta_m} & \cdots & \dfrac{\partial r_n(\boldsymbol{\theta})}{\partial\theta_m} \end{bmatrix} \tag{19}$$

其中: m 为待估计参数数量, 本文中 m 为 5。

L-M 寻优算法的具体步骤如下:

步骤 1　令 $k=0$, 给定初值 $\boldsymbol{\theta}^{(k)}$, 迭代精度 ε, 初始步长 $l^{(k)}$。

步骤 2　计算 $\boldsymbol{r}\left(\boldsymbol{\theta}^{(k)}\right)$ 和 $\boldsymbol{J}\left(\boldsymbol{\theta}^{(k)}\right)$, 令 $\boldsymbol{A}=\boldsymbol{J}\left(\boldsymbol{\theta}^{(k)}\right)^{\mathrm{T}}\cdot\boldsymbol{J}\left(\boldsymbol{\theta}^{(k)}\right)$, $\boldsymbol{B}=\boldsymbol{J}\left(\boldsymbol{\theta}^{(k)}\right)^{\mathrm{T}}\boldsymbol{r}\left(\boldsymbol{\theta}^{(k)}\right)$。

步骤 3　计算 $\boldsymbol{\theta}^{(k+1)}$ 及目标函数 $\lambda(\boldsymbol{\theta}^{(k+1)})$ 的值。

步骤 4　如果 $\lambda(\boldsymbol{\theta}^{(k+1)})\leqslant\varepsilon$, 则最优解为 $\boldsymbol{\theta}^{(k+1)}$, 迭代结束, 否则跳转到步骤 5。

步骤 5　如果 $\lambda(\boldsymbol{\theta}^{(k+1)})\geqslant\lambda(\boldsymbol{\theta}^{(k)})$, 则改变步长, 令 $l^{(k)}=10l^{(k)}$, 跳转到步骤 6; 如果 $\lambda(\boldsymbol{\theta}^{(k+1)})<\lambda(\boldsymbol{\theta}^{(k)})$, 则改变步长, 令 $l^{(k)}=0.1l^{(k)}$, 跳转到步骤 6。

步骤 6　令 $k=k+1$, 重复步骤 2。

根据近十余年中国民航鸟撞事件统计数据及上述参数估计方法, 鸟撞冲击能量数据的分布拟合结果如图 2 所示, 并与一般两参数威布尔分布拟合结果进行比较, 如图 2、表 1 和表 2 所示。

SSE 为拟合数据和原始数据的误差平方和, RMSE 为 SSE/n 的平方根, R^2 为拟合数据和原始数据的线性相关度指标。SSE 和 RMSE 的值越接近 0, 表明模型选择越合理, 拟合效果越佳; R^2 的取值在 0～1, 该值越接近 1, 表明该模型对数据的拟合效果越好, 数据预测也越精确。

综合所述, 两重两参数混合威布尔分布函数更适合中国的鸟撞冲击能量分布特征。以 e 表征鸟撞冲击能量, 则其累计概率分布函数形式为

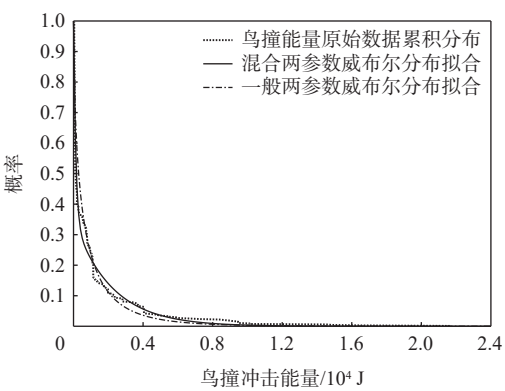


图 2　鸟撞冲击能量威布尔分布拟合

Fig. 2　Fitting of Weibull distribution of bird strike energy

表 1　2 种拟合分布函数的参数估计结果

Table 1　Parameter estimation values of two fitting distribution functions

估计参数	p_1	η	β
两重两参数混合威布尔分布	0.346 4	2 204.685 2 (η_1) 127.346 3 (η_2)	1 (β_1) 1 (β_2)
一般两参数威布尔分布		542.098 7	0.601 1

注: η_1, η_2 为 η 的两重参数; β_1, β_2 为 β 的两重参数。

表 2　2 种拟合分布函数的拟合优度对比

Table 2　Comparison of goodness of fit of two fitting functions

拟合优度评估指标	残差平方和SSE	均方差RMSE	判定系数R^2
两重两参数混合威布尔分布	1.613 3	0.041 3	0.957 9
一般两参数威布尔分布	6.074 5	0.080 2	0.906 8

$$P(e)=0.346\,4\exp\left[-\left(\frac{e}{2\,204.685\,2}\right)^{1}\right]+0.653\,6\exp\left[-\left(\frac{e}{127.346\,3}\right)^{1}\right]\tag{20}$$

其物理意义是冲击能量大于 e 的概率为 $P(e)$。因此, 假定冲击能量处于 $[e_1, e_2]$ 区间内, 则其对应的概率值为

$$P(e_1\leqslant e<e_2)=P(e_1)-P(e_2)\tag{21}$$

2　基于安全性目标的条款充分性评估

2.1　鸟击能量概率估算

CCAR 25.571[4] 和 CCAR 25.631[4] 规定, 参考国内外主流运输类飞机的速度 V_c, 选取 V_c=180 m/s 作为鸟撞速度, 以确定不同适航规章要求对应的鸟撞能量要求。因此, 1.8 kg(4 lb)和 3.6 kg(8 lb)抗鸟撞要求对应的可承受的冲击能量值分别为 $e_4=29\,393$ J 和 $e_8=58\,786$ J, 则鸟撞冲击能量大于 29 393 J 和大于 58 786 J 的概率风险分别为

$$P_4=P(e>e_4)=5.617\,4\times10^{-7}$$

$$P_8=P(e>e_8)=9.109\,5\times10^{-13}$$

飞机抗 1.8 kg 和 3.6 kg 鸟撞适航条款对应的冲击能量及其概率值如图 3 所示。

以 2015 年为例, 中国民航鸟击事件万架次率为 4.455, 假设运输类飞机每架次的平均飞行时间为 3 h, 则每架次飞机每飞行小时发生鸟击的概率为 $P_b=1.485\times10^{-4}$。当年发生鸟击部位的统计数

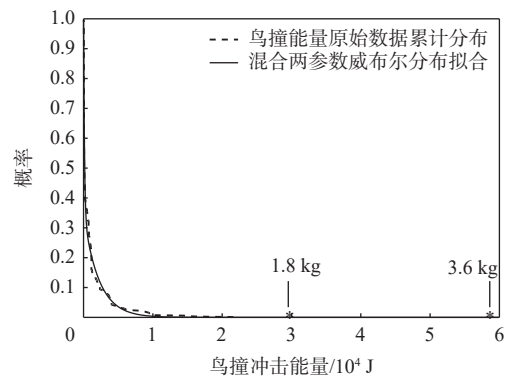


图 3　飞机抗鸟撞适航条款对应的冲击能量及其概率值

Fig. 3　Strike energy and its probability corresponding to airworthiness clause of aircraft against bird strike

据[22]显示，机翼发生鸟击的次数占总体的比例为14.94%，尾翼发生鸟击的次数占总体的比例为2.33%，则每飞行小时机翼和尾翼发生鸟击的概率分别为

$$\begin{cases} P_{\mathrm{b}}^{\mathrm{w}} = 1.485\times10^{-4}\times14.94\% = 2.218\,6\times10^{-5} \\ P_{\mathrm{b}}^{\mathrm{t}} = 1.485\times10^{-4}\times2.33\% = 3.460\,1\times10^{-6} \end{cases}$$

因此，超过1.8 kg和3.6 kg质量的鸟体分别对机翼和尾翼的撞击能量风险为

$$\begin{cases} P_4^{\mathrm{w}} = P(e>e_4)\times P_{\mathrm{b}}^{\mathrm{w}} = \left(5.617\,4\times10^{-7}\right)\times \\ \qquad \left(2.218\,6\times10^{-5}\right) = 1.246\,3\times10^{-11} \\ P_8^{\mathrm{t}} = P(e>e_8)\times P_{\mathrm{b}}^{\mathrm{t}} = \left(9.109\,5\times10^{-13}\right)\times \\ \qquad \left(3.460\,1\times10^{-6}\right) = 3.152\,0\times10^{-18} \end{cases}$$

根据参考文献[23]中具有共识性的发生灾难性事故的目标概率为10^{-9}可知，在中国鸟情环境下，运输类飞机机翼遭遇4 lb鸟体撞击的概率满足安全风险目标，飞机尾翼遭受8 lb鸟体撞击的概率满足安全风险目标，即中国现行的飞机机翼和尾翼抗鸟撞适航条款是充分的，但尾翼抗鸟撞适航条款却过于保守。鉴于尾翼抗鸟撞适航条款过于保守的现状，对尾翼每飞行小时遭受4 lb鸟撞击的概率计算为

$$P_4^{\mathrm{t}} = P(e>e_4)\times P_{\mathrm{b}}^{\mathrm{t}} = \left(5.617\,4\times10^{-7}\right)\times \left(3.460\,1\times10^{-6}\right) = 1.943\,5\times10^{-12}$$

上述计算结果表明，中国鸟情环境下飞机尾翼抗3.6 kg鸟体质量的适航条款也是充分且保守的。

鉴于参考文献[3]中对整架飞机每次飞行时遭遇超过1.8 kg鸟/海平面巡航速度撞击的能量和超过3.6 kg鸟/海平面巡航速度撞击的能量的概率分别为10^{-7}和10^{-8}的安全性目标，则每次飞行发生与1.8 kg和3.6 kg以上质量的鸟撞击概率分别为

$$\begin{cases} P_4^{\mathrm{a}} = P(e>e_4)\times(P_{\mathrm{b}}^{\mathrm{w}}\times3) = (5.617\,4\times10^{-7})\times \\ \qquad \left(4.455\times10^{-4}\right) = 2.502\,6\times10^{-10} << 10^{-7} \\ P_8^{\mathrm{a}} = P(e>e_8)\times(P_{\mathrm{b}}^{\mathrm{t}}\times3) = \left(9.109\,5\times10^{-13}\right)\times \\ \qquad \left(4.455\times10^{-4}\right) = 4.058\,3\times10^{-16} << 10^{-8} \end{cases}$$

上述计算结果表明，中国鸟情环境下整架飞机每次飞行时抗1.8 kg鸟和3.6 kg鸟撞击的适航条款也是充分的，且抗3.6 kg鸟撞击的适航条款也是过于保守的。

2.2 FAA的愿景验证

为验证2.1节研究方法和研究结果的可信性，以FAA网站上公布的国家野生动物撞击数据库（national wildlife strike database, NWSD）中2006—2019年的鸟击事件信息为基础，挑选具有撞击速度记载值及明确鸟种的1万余条记录数据进行统计分析，得出鸟击能量概率分布大致服从对数正态分布（位置参数为5.843，比例参数为1.367），如图4所示。

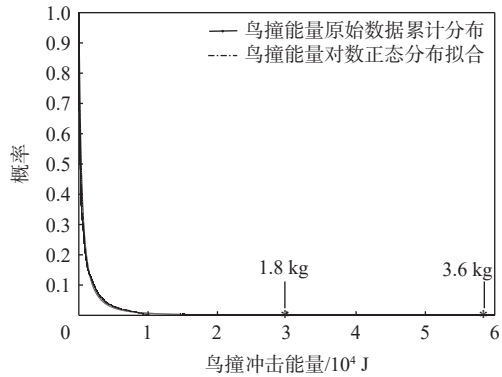


图4 基于FAA鸟击能量分布的飞机适航条款相关的冲击能量及其概率值

Fig. 4 Strike energy and its probability related to aircraft airworthiness clause based on bird strike energy distribution of FAA

图4显示，1.8 kg和3.6 kg的鸟以180 m/s的速度撞击后，大于相应撞击能量的概率分别为$5.847\,8\times10^{-4}$和$8.727\,6\times10^{-5}$。根据文献[24]，以美国2015年为例（对应于中国2015年的运行数据），美国民航鸟击事件万架次率为3.192，同样假设飞机每架次的平均飞行时间为3 h，则每架次飞机每飞行小时发生鸟击的概率为$P_{\mathrm{b}}' = 1.064\times10^{-4}$；同时，机翼和尾翼发生鸟击的次数占鸟击总数的比例分别为14%和1%。因此，对机翼而言，1.8 kg及以上质量的鸟体每飞行小时发生撞击机翼的概率为$5.847\,8\times10^{-4}\times1.064\times10^{-4}\times0.14=8.710\,9\times10^{-9}$，大于灾难性事故安全目标$10^{-9}$；对尾翼而言，3.6 kg及以上质量的鸟体每飞行小时发生撞击平尾或垂尾的概率为$8.727\,6\times10^{-5}\times1.064\times10^{-4}\times0.01=9.286\,2\times10^{-11}\approx10^{-10}$，小于灾难性事故安全目标$10^{-9}$。上述分析说明：机翼抗1.8 kg鸟体质量的适航条款要求已经无法满足北美地区发生灾难性鸟击风险的安全目标，存在较大的发生灾难性鸟击飞机事件的风险，这个结论与FAA提出的将飞机机翼抗鸟撞鸟体质量标准提升至3.6 kg具有趋同性；尾翼发生3.6 kg及以上质量的鸟体撞击的概率也体现了FAA继续提高尾翼抗鸟撞适航标准的意向。这从侧面反映了基于鸟撞冲击能量统计分布的研究方法和中国鸟环境下现行的飞机机翼抗鸟撞适航条款的充分性结论具有一定的可信性。同时，基于美国鸟情的飞机机翼抗1.8 kg鸟撞击的安全不充分性，对机翼每飞行小时遭受3.6 kg鸟撞击的概率计算为$8.727\,6\times10^{-5}\times1.064\times10^{-4}\times14\%=1.300\,1\times10^{-9}$。该结果表明，美国鸟情环境下飞机机翼抗3.6 kg鸟体质量的适航条款接近但仍超出发生灾难性事故的安全性目标，安全

风险裕度为负。

鉴于文献 [3] 中对整架飞机每次飞行时遭遇超过 1.8 kg 鸟/海平面巡航速度撞击的能量和超过 3.6 kg 鸟/海平面巡航速度撞击的能量的概率分别为 10^{-7} 和 10^{-8} 的安全性目标, 则每次飞行发生与 1.8 kg 和 3.6 kg 以上质量的鸟撞击概率分别为

$$\begin{cases} P_4^{\mathrm{a}} = (5.847\,8\times10^{-4})\times(1.064\times10^{-4}\times3) = \\ \qquad 1.866\,7\times10^{-7} > 10^{-7} \\ P_8^{\mathrm{a}} = (8.727\,6\times10^{-5})\times(1.064\times10^{-4}\times3) = \\ \qquad 2.785\,8\times10^{-8} > 10^{-8} \end{cases}$$

上述计算结果表明, 美国鸟情环境下整架飞机每次飞行时抗 1.8 kg 鸟和 3.6 kg 鸟撞击的适航条款是不充分的, 这也是美国局方近年来一直致力于推动和协调提高运输类飞机抗鸟撞适航认证标准的愿景。

3　基于安全性目标的条款适宜性评估

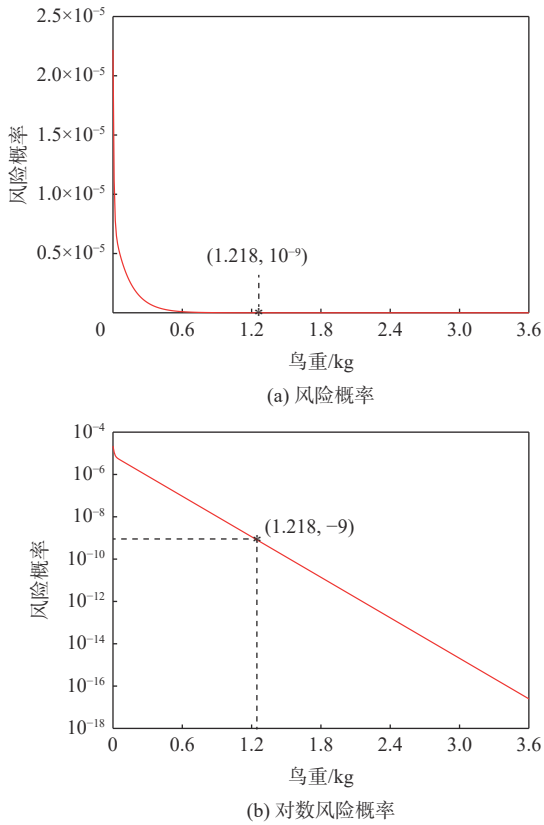
3.1　不同安全性目标下的鸟体质量确定

以 1×10^{-9} 为鸟撞安全性目标, 根据中国鸟击航空器能量概率分布函数、鸟击事件万架次率、机翼部位发生鸟击的频次比例等数据进行鸟体质量反推, 得到机翼结构对应的可承受的鸟体质量为 1.218 kg, 即从安全风险目标的角度考虑, 处于中国鸟情环境下的运输类飞机机翼的抗鸟撞要求中的鸟体质量需达到 1.218 kg, 这小于目前中国对民机机翼抗鸟撞适航要求中对鸟体的质量要求 (1.8 kg)。飞机机翼不同抗鸟撞结构设计所能承受的鸟体质量与风险值的对应关系如图 5 所示。

同理, 如以 10^{-9} 为尾翼抗鸟撞安全性目标(不按平尾和垂尾分别计算)对应的鸟体质量为 0.965 kg (约 2.13 lb), 即如果仅从上述安全风险目标考虑, 对处于中国鸟群环境下的运输类飞机尾翼的抗鸟撞鸟体质量要求只需达到 0.965 kg 即可。因此, 在中国鸟群环境下 CCAR 25.631[4] 条款中对尾翼结构的抗鸟撞鸟体质量要求(8 lb)是非常保守的。

如以 10^{-10} 为尾翼抗鸟撞安全性目标, 反推得到的鸟体质量为 1.278 kg; 如以 10^{-11} 为尾翼抗鸟撞安全性目标, 反推得到的鸟体质量为 1.591 kg。飞机尾翼不同抗鸟撞结构设计所能承受的鸟体质量与风险值的对应关系如图 6 所示。

由图 5 和图 6 可知, 中国鸟情环境下飞机机翼和尾翼结构不同抗鸟撞鸟体质量要求对应的安全风险水平; 根据可接受安全风险阈值, 可对应得到运输类飞机机翼和尾翼结构设计需满足可承受的最大鸟体质量。这对飞机机翼和尾翼结构抗鸟撞设计及优化改进具有重要的参考价值。

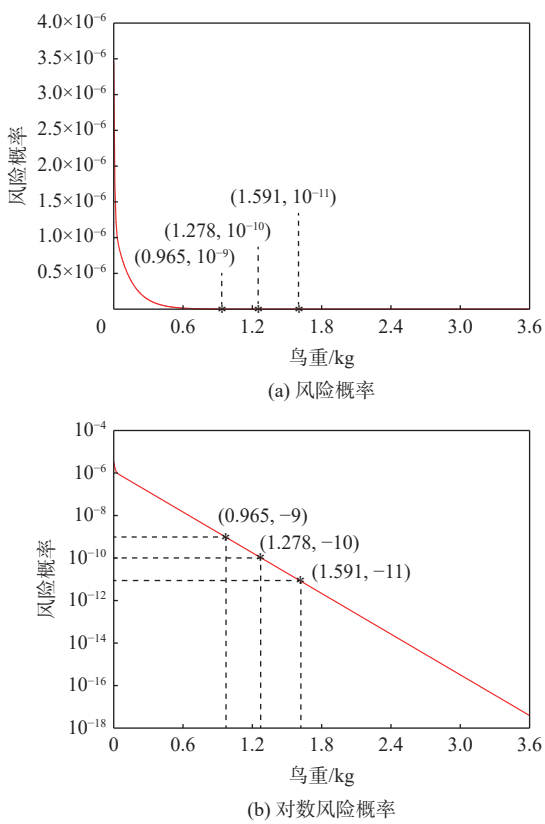


(a) 风险概率

(b) 对数风险概率

图 5　飞机机翼结构遭遇的鸟体质量与对应的风险概率

Fig. 5　Weight of bird affecting aircraft wing and corresponding risk



(a) 风险概率

(b) 对数风险概率

图 6　飞机尾翼结构遭遇的鸟体质量与对应的风险概率

Fig. 6　Weight of bird affecting aircraft tail and corresponding risk

3.2 FAA愿景的进一步验证

以FAA网站上公布的2006—2019年的鸟击事件数据及其鸟击能量分布函数为基础，以10^{-9}为安全性目标，反推机翼应承受的抗鸟撞鸟体质量为3.9 kg，大于目前规章要求的4 lb(约1.8 kg)鸟体质量，这也许正是从一个重要维度佐证了FAA近20年来一直在讨论将机翼抗鸟撞鸟体质量标准提高至8 lb(约3.6 kg)的证据之一[3]；同样，按此安全性目标，反推尾翼应承受的抗鸟撞鸟体质量应为1.5 kg，小于目前规章要求的8 lb(约3.6 kg)。但由于在美国曾经仅出现过1次鸟击尾翼后机毁人亡的事故，故FAA暂时没有调整对尾翼的抗鸟撞鸟体质量标准。

4 结 论

本文提出一种基于两重两参数混合威布尔分布的鸟撞冲击能量统计分析方法，并通过对美国商用飞机鸟击数据的统计分析和FAA修订相关条款的部分愿景验证，证实本文方法的有效性，以及基于中国鸟情环境下的运输类飞机抗鸟撞适航条款的充分性和适宜性结果的一定可信性。

1）应用L-M寻优算法对鸟撞冲击能量进行混和威布尔分布参数寻优估计，得到比一般两参数威布尔分布更加精确和有效的中国鸟撞冲击能量分布。

2）对机翼而言，相关抗鸟撞适航条款具有充分性和适宜性；以10^{-9}为安全性目标，机翼结构对应的可承受鸟体质量为1.218 kg。

3）对尾翼而言，相关适航条款具有充分性但过于保守，可适当降低尾翼结构的抗鸟撞鸟体质量要求(如1.8 kg，与EASA规定的飞机尾翼抗鸟撞鸟体质量要求一致)。

值得注意的是：对鸟情数据的统计分析和对鸟撞航空器冲击能量的风险评估，是保守确定鸟撞适航条款中鸟体质量要求的初始依据，也是持续更新鸟撞适航条款的客观事实。相较于美国鸟撞航空器信息收集的长期性、完善性和规范性，尽管中国民航鸟撞航空器信息还存在缺项较多、分类颗粒度粗犷等问题，也导致本文计算结果与实际情况还存在一定差异，但随着中国鸟情数据收集的不断丰富完善和对鸟撞案例统计分析的持续改进，相信可以得到更加精确的鸟情分布特征，为制定/修订适合中国鸟情环境的运输类飞机抗鸟撞适航要求提供可靠的决策依据。

致谢 感谢中国民航科学技术研究院提供的相关鸟撞航空器信息源，感谢中国民航上海航空器适航审定中心提供的相关规章解读支持。

参考文献（References）

[1] Federal Aviation Administration, Department of Transportation of U. S. . Part 25—Airworthiness standards: Transport category airplanes [S/OL] . Washington, D. C. : Federal Aviation Administration, (2024-05-01)[2024-05-23]. https://www.govinfo.gov/content/pkg/CFR-2023-title14-vol1/pdf/CFR-2023-title14-vol1-part25.pdf.

[2] European Aviation Safety Agency. Certification specifications and acceptable means of compliance for large aeroplanes CS-25: Amendment 19[S]. Brussels: European Aviation Safety Agency, 2017: 1-C-26-1-C-28+1-D-6.

[3] ANDREW H K. Submittal of results of harmonization effort on FAR/JAR §25.631, birdstrike: L350-03-114, version 2[R]. Washington, D. C. : Aviation Rulemaking Advisory Committee of FAA, 2003: 11-32.

[4] 中国民用航空局政策法规司. 运输类飞机适航标准: CCAR-25-R4 [S]. 北京: 中国民用航空局, 2016: 69-71.
Policy and Regulation Department, Civil Aviation Administration of China. Airworthiness standards of transport category airplanes CCAR-25-R4 [S]. Beijing: Civil Aviation Administration of China, 2016: 69-71(in Chinese).

[5] Federal Aviation Administration, Department of Transportation of U. S. . Bird strike requirements for transport category airplanes: FAA–2015–2490[S]. Washington D. C. : Federal Aviation Administration, 2015: 63877-63878.

[6] COLÓN M R, LONG A M. Strike hazard posed by columbids to military aircraft[J]. Human-Wildlife Interactions, 2018, 12(2): 198-211.

[7] PFEIFFER M B, BLACKWELL B F, DEVAULT T L, et al. Quantification of avian hazards to military aircraft and implications for wildlife management[J]. Public Library of Science One, 2018, 13(11): 1-16.

[8] 熊明兰, 王华伟, 徐怡, 等. 基于鸟击事故征候预测的通用航空安全研究[J]. 系统工程与电子技术, 2020, 42(9): 2033-2040.
XIONG M L, WANG H W, XU Y, et al. General aviation safety research based on prediction of bird strike symptom[J]. Systems Engineering and Electronics, 2020, 42(9): 2033-2040(in Chinese).

[9] 陈唯实, 万健, 李敬. 基于机场探鸟雷达数据的鸟击风险评估[J]. 北京航空航天大学学报, 2013, 39(11): 1431-1436.
CHEN W S, WAN J, LI J. Bird strike risk assessment with airport avian radar data[J]. Journal of Beijing University of Aeronautics and Astronautics, 2013, 39(11): 1431-1436(in Chinese).

[10] 郑江萍. 鸟击风险分布及风险评估研究[D]. 天津: 中国民航大学, 2014: 38-43.
ZHENG J P. Study on risk distribution and assessment of bird strike[D]. Tianjin: Civil Aviation University of China, 2014: 38-43 (in Chinese).

[11] 朱贝蓓, 蔡景. 基于MCMC方法的运输类飞机鸟撞冲击能量研究[J]. 航空计算技术, 2017, 47(1): 94-96.
ZHU B B, CAI J. Study on impact energy of bird strike of transport aircraft based on MCMC method[J]. Aeronautical Computing Technique, 2017, 47(1): 94-96(in Chinese).

[12] SARKHEIL H, TALAEIAN ERAGHI M, VATAN KHAH S. Haz-

ard identification and risk modeling on runway bird strikes at Sardar-e-Jangal International Airport of Iran[J]. Modeling Earth Systems and Environment, 2021, 7(4): 2589-2598.

[13] 陈琨, 解江, 裴惠, 等. 明胶鸟弹撞击复合材料蜂窝夹芯板试验[J]. 复合材料学报, 2020, 37(2): 328-335.
CHEN K, XIE J, PEI H, et al. Experiment of composite honeycomb sandwich panels subjected to gelatin bird impact[J]. Acta Materiae Compositae Sinica, 2020, 37(2): 328-335(in Chinese).

[14] LONG S C, MU X L, LIU Y H, et al. Failure modeling of composite wing leading edge under bird strike[J]. Composite Structures, 2021, 255: 113005.

[15] 朱海龙. 民用飞机液压系统防鸟撞适航符合性评估[J]. 液压气动与密封, 2021, 41(10): 67-70.
ZHU H L. CivilAircraft hydraulic system airworthiness compliance assessment for bird bird impact[J]. Hydraulics Pneumatics & Seals, 2021, 41(10): 67-70(in Chinese).

[16] KIM D H, KIM S W. Evaluation of bird strike-induced damages of helicopter composite fuel tank assembly based on fluid-structure interaction analysis[J]. Composite Structures, 2019, 210: 676-686.

[17] 顾晨轩, 苏艳, 王辉. 国内外运输类飞机鸟撞适航条款及其修订背景研究分析[J]. 科技创新导报, 2017, 14(7): 247-249.
GU C X, SU Y, WANG H. Research and analysis on airworthiness clause of bird impact of transport aircraft at home and abroad and its revision background[J]. Science and Technology Innovation Herald, 2017, 14(7): 247-249(in Chinese).

[18] 罗刚, 张海洋, 吴春波, 等. 航空发动机吞鸟与鸟撞飞机适航通用分析方法[J]. 航空发动机, 2019, 45(6): 90-96.
LUO G, ZHANG H Y, WU C B, et al. General airworthiness analysis method of aeroengine bird ingestion and aircraft bird impact[J]. Aeroengine, 2019, 45(6): 90-96(in Chinese).

[19] SUZANNE M. Bird strike requirements for transport category airplane- compliance by analysis[R]. Washington, D. C. : FAA, 2016: 1-15.

[20] ASTM Subcommittee on Transparent Enclosures. Standard test method for bird impact testing of aerospace transparent enclosure: F330-21[S]. West Conshohochen: ASTM International, 2021: 1-5.

[21] 王立, 孙秀清, 张春明, 等. 一种全天时星跟踪器相对惯导的安装阵在线快速估计方法[J]. 航空学报, 2020, 41(8): 624117.
WANG L, SUN X Q, ZHANG C M, et al. Fast online estimation method for installing matrix between all-time star tracker and inertial navigation system[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(8): 624117(in Chinese).

[22] 中国民用航空局机场司和中国民航科学技术研究院. 2015 年度中国民航鸟击信息分析报告[R]. 北京: 中国民用航空局, 2015: 1-24.
Airport Department of Civil Aviation Administration of China and China Academy of Civil Aviation Science and Technology. Civil aviation bird strike aircraft information analysis report of civil aviation of China in 2015[R]. Beijing: Civil Aviation Administration of China, 2015: 1-24(in Chinese).

[23] Society of Automotive Engineers. Guidelines and methods for conducting the safety assessment process on civil airborne system and equipment: ARP 4761[S]. Detroit: SAE International, 1996: 102-103.

[24] RICHARD A D, MICHAEL J B, PHYLLIS R M, et al. Wildlife strikes to civil aircraft in the United States, 1990–2019[R]. Washington D. C. : Federal Aviation AdministrationNational Wildlife Strike Database, 2021: 2-31 .

Adequacy and suitability of airworthiness clause of bird strike based on bird situation in China

LU Xiaohua[1,*], CAI Jing[1], ZHANG Zhuguo[2], ZHANG Yingchun[2]

(1. College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China;
2. Shanghai Aircraft Airworthiness Certification Center of CAAC, Shanghai 200335, China)

Abstract: Based on the data source of bird strike information of airplanes in transportation category from civil aviation of China in recent ten years, the mixed Weibull function was used to fit the bird strike energy distribution, and a Levenberg-Marquardt (L-M) optimization algorithm was applied to obtain more accurate estimations of fitted distribution parameters. Based on the safety index, it was sufficient and appropriate for the wing to evaluate the existing airworthiness clauses of bird strike by the bird strike energy distribution function. However, it was sufficient but too conservative for the tail. According to the safety index of catastrophic accidents, the minimum bird weight meeting the airworthiness requirement of the wing against bird strike under the bird situation in China was 1.218 kg. At the same time, the vision was verified that Federal Aviation Administration (FAA) would increase the minimum airworthiness requirement of the wing against bird strike to 3.6 kg under the bird situation in America, which partially demonstrated the validity of the methods and reliability of the results and provided a theoretical basis and practical reference for the independent revision of relevant airworthiness clauses of civil aviation of China.

Keywords: bird situation; airworthiness clause; bird strike energy distribution; evaluation; adequacy; suitability; safety index

Received: 2022-08-17; **Accepted**: 2022-11-04; **Published Online**: 2022-12-02 11:12

URL: link.cnki.net/urlid/11.2625.V.20221201.1457.002

Foundation items: National Natural Science Foundation of China (U1933202); Safety Capacity Building Fund of CAAC (2021-198)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0120

考虑视觉搜索特性的指挥控制系统KLSM

郭仲[1,2,*]，赵刚[1]，付斌[2]，刘佳烨[3]

(1. 中国人民解放军军事科学院 系统工程研究院，北京 100101；2. 中国人民解放军93216部队，北京 100085；
3. 中国人民解放军93131部队，北京 100038)

摘 要：传统的击键水平模型（KLM）专注于键盘和鼠标的操作分析，适用于简单软件界面任务完成时间的快速定量评估，但将其迁移到包含复杂信息特征的指控系统软件界面评估时，却会产生较大误差。为此，基于KLM，结合用户视觉特性，面向指挥控制系统典型任务，提出基于用户视觉搜索处理时间的物理操作时间单元指标，并通过实验获取用户视觉搜索处理时间典型参数，建立更精确的指控系统任务操作分析模型——击键和视觉搜索模型（KLSM），即考虑视觉搜索特性因素的影响。选取5类典型指控任务，进行对比实验验证，KLSM对任务完成时间的预测评估更为准确。

关 键 词：视觉搜索特性；指控系统；软件界面；KLSM；预测评估
中图分类号：V221^{+}.3; TB472
文献标志码：A **文章编号**：1001-5965(2024)09-2819-07

指挥控制系统(下文称指控系统)是实现战场上各个变量、各模块互相联通，保证战场指挥高效畅通、战场指令精准下达的关键因素[1]。其中，软件界面是指挥员与系统交互的主要通道，是系统中各类战法、算法、数据乃至作战概念的直接呈现，起着至关重要的作用[2]。

随着战争形态的变化，敌我双方作战循环(observe, orient, decide, act, OODA)不断加快，软件界面承载了更加繁杂的战场态势信息和关键作战功能，这给指挥员作战指挥带来巨大的压力和挑战[3]。愈加复杂的界面设计和人机交互操作不可避免地导致界面可用性问题，因此软件界面评估工作越发重要[4]。众多学者提出界面交互系统可用性评价方法，致力于为交互设计优化指明方向[5]。

其中，目标、操作、方法、选择方法的原则(goals, operators, methods, and selection rules, GOMS)是人机交互领域最经典的任务分析方法之一。Card[6]提出GOMS，通过操作时间实现对操作行为的量化分析，从用户操作的角度实现对产品性能和可用性的分析，以改进技术、优化产品。Schrepp[7]利用该方法分析产品中的用户性能问题。Jung和Jang [8]利用GOMS验证了其提出的2步触摸法的有效性，从而提升了智能手机的可用性。Huang和Chiu [9]运用GOMS分析来评估用户能否实现预先设定的性能目标，包括界面分析和接口设计等。Ramkumar等[10]利用GOMS分析用户特征，用来预测用户人机交互的潜在风险。

本文面向指挥控制系统典型任务，针对具有复杂信息的指控系统软件界面，结合用户视觉搜索特性，提出更精确的指控系统任务操作分析模型——击键和视觉搜索模型(keystroke level and visual search model , KLSM)，并结合典型任务对模型效果进行了对比验证分析。

1 击键水平模型

1.1 GOMS

GOMS是一种用于评估和分析人机交互系统的认知模型，其旨在帮助设计师理解和预测用户在

收稿日期：2024-03-04；录用日期：2024-03-15；网络出版时间：2024-03-25 12:17
网络出版地址：link.cnki.net/urlid/11.2625.V.20240325.0956.001
基金项目：国家社会科学基金 (2022-SKJJ-B-043)
*通信作者. E-mail：18803806@qq.com

引用格式：郭仲，赵刚，付斌，等. 考虑视觉搜索特性的指挥控制系统KLSM[J]. 北京航空航天大学学报，2024，50(9)：2819-2825.
GUO Z，ZHAO G，FU B，et al. Considering visual search characteristics on command and control system KLSM[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2819-2825 (in Chinese).

使用计算机界面或系统时的认知和行为过程。GOMS可以定性分析用户交互的特征(如动作序列),并定量预测每个操作员的执行时间[11]。该模型将人与系统软件的交互过程分解为4个部分:目标(goals)、操作(operations)、方法(methods)、选择方法的原则(selection rules)。其核心理念是将用户与界面的交互行为拆解为独立的行为元动作,无须真实用户介入就能预测交互行为的序列,以评估交互的质量。

在实际应用中,GOMS有不同的变种和扩展,如关键路径法(Critical-PathMrthod, CPM-GOMS)、自然语言模型(natural GOMS language, NGOMSL)等,这些模型可用于更精细地描述用户行为和认知过程,但是无法实现对并行任务的评估。Christou等[12]提出基于GOMS改进的Codein模型,可以对并行操作的任务完成时间进行建模和评估。这些模型和方法在人机交互设计中有助于提升用户体验,减少用户错误,以及提高系统效率。但评估过程繁琐,不易执行。为提高模型的实用性,Card等[13]提出击键水平模型(keystroke level model, KLM),该模型专注于键盘和鼠标操作的分析,其使用一个固定的操作集合,包括按键、鼠标点击、鼠标移动等,以估算完成任务所需时间。KLM的简单性使其在快速评估界面设计时非常有效。学者们依据实际场景将KLM进行改进拓展,Findlater等[14]提出基于触屏游戏的KLM,赵欣等[15]提出更适用于手持触摸显示设备的评估模型。目前的研究大多针对简单PC界面或手持设备,对适用于具有复杂信息的指控系统的界面评估模型的研究较匮乏。

1.2 KLM基本原理

KLM采用预定义的操作符来预测用户的无差错任务执行时间,主要涉及3个操作符:点击、指向和心理操作,如图1所示。通过对用户行为的研究,给出了各基本操作行为所需典型时间,如表1所示。

任务执行总时间等于任务中出现的各个操作符关联时间的总和,利用K、P、M及M的放置规则

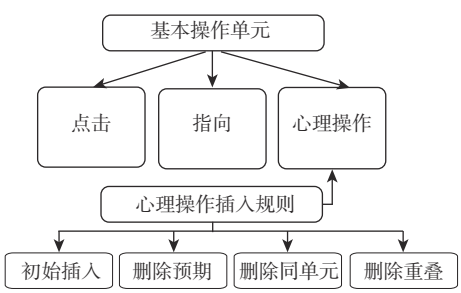


图1 KLM基本原理

Fig. 1 The basic principle of KLM

表1 KLM基本操作时间

Table 1 Basic operation time of KLM

操作	典型值/s	含义
K	0.2	敲击键盘中一个键所用的时间
P	1.1	用户指向显示器某个位置所用的时间
M	1.35	用户进入下一步所需心理准备时间

即可定量分析用户使用人机界面的操作时间,进而分析出界面设计的优劣:

$$T_{\mathrm{KLM}} = T_K + T_P + T_M + \cdots \tag{1}$$

式中:T_{KLM}为任务执行总时间;T_K为点击操作时间的总和;T_P为指向操作时间的总和;T_M为心理操作时间的总和。

2 KLSM构建

2.1 KLM局限性

前期研究发现,采用KLM预测的指控系统软件界面任务执行时间与实际工作时间存在较大误差。分析得出其原因为,除了KLM本身所固有的序列操作、无差错等假设带来的误差外,没有考虑对于界面复杂、信息量繁多的指控系统,用户需要花费大量时间查看、搜索界面信息,而KLM缺少对用户界面搜索加工步骤时间的评估,因而导致其用于指控系统任务评估时的准确性不高。

因此,本文基于KLM,结合用户视觉特性,提出基于用户视觉处理时间的物理操作时间单元指标并量化,建立了更精确的指控系统任务操作分析模型KLSM,以提升模型预测评估的准确性。

2.2 视觉搜索处理时间

视觉搜索处理是指眼睛移动搜索并处理相关线索的过程[16]。指控系统界面信息复杂,用户需要花费较长时间搜索加工信息,而传统的KLM缺乏对用户视觉搜索处理耗费时间的评估。脑电、眼动等无损伤探测技术的快速发展,为研究用户此类特征提供了更多支持。其中,眼动指标中的扫描持续时间表示用户在视觉扫描过程中,相邻的两注视点间所用时间[17]。注视时间是构成注视点的首个采样点与最后一个采样点间的持续时间,其可反映用户信息加工的效率,与用户的视觉搜索效率呈负相关[18-19],并且能够有效地量化表征用户的视觉搜索处理时间。

2.3 实验与模型构建

2.3.1 实验目的

实验运用眼动追踪技术,通过对用户使用指控

系统执行任务的实验数据进行统计分析, 得出视觉搜索处理时间的典型值, 并将其作为可量化的物理操作时间单元, 从而在 KLM 的基础上, 增加物理操作时间单元参数, 构建 KLSM。

2.3.2　实验设计

实验针对某指控系统, 选取日常使用频率较高的、最具代表性的规划管理等 5 类典型任务设计用例。任务开始前, 主试者向被试者介绍实验流程及注意事项, 被试者确认理解后, 佩戴眼动仪, 每完成一项任务后休息 10 min。眼动仪记录被试者在任务过程中的眼动数据, 提取时间特征, 并计算被试者在执行典型任务情况下的平均注视时长。通过对平均注视时长的总体描述的统计结果, 结合统计特征, 确定作为新增用户视觉搜索处理时间参数以优化 KLM。为保证实验数据的客观性, 采取多轮实验取平均值的方式来采集和处理数据, 即每类任务每个被试者至少开展 3 轮实验。

2.3.3　被试者

邀请某部队当前在位的 6 名值班参谋参与实验。被试者共 6 人, 年龄为 22～35 岁, 所有被试者视力或矫正视力正常, 色觉正常。所有人员均能够熟练操作指控系统。被试者均为右利手, 在实验前均已休息充足。

2.3.4　实验设备

台式计算机 6 台, 用于被试者操作指控系统; 系统服务器 1 套, 运行指控系统后台服务, 其 CPU 为鲲鹏 920; 瑞典 Tobii 公司研制的 Tobii Glasses 3 眼动仪 6 台, 如图 2 所示; Tobii Pro Lab 眼动分析软件 1 套。Tobii Glasses 3 眼动仪能够记录和分析各项眼动数据, 并且还能在拍摄真实场景的同时不限制被试者的行动, 确保还原最真实的人类行为, 为科学研究提供深入有效的观察窗口。



图 2　Tobii Glasses 3 眼动仪

Fig. 2　Tobii Glasses 3 eye tracker

2.3.5　数据采集与整理

3 轮实验后取平均值, 共得到 30 份眼动数据, 包含 5 类任务场景。眼动仪记录了被试者的各项眼动指标, 通过分析眼动数据, 统计被试者在 5 类不同任务条件下的平均注视时长, 利用 SPSS 27 软件进行异常值的剔除和数据分析整理, 具体数据如表 2 所示, 图 3 为不同任务的被试者平均注视时长箱型。

表 2　不同任务中被试者平均注视时长

Table 2　Average gaze duration of subjects in different tasks

任务	平均注视时长/ms	标准差/ms	最大值/ms	最小值/ms
任务1	608.50	127.06	774.00	470.00
任务2	713.50	254.59	1 132.00	393.00
任务3	646.17	210.35	885.00	377.00
任务4	707.67	173.02	948.00	446.00
任务5	724.33	377.64	1 151.00	214.00

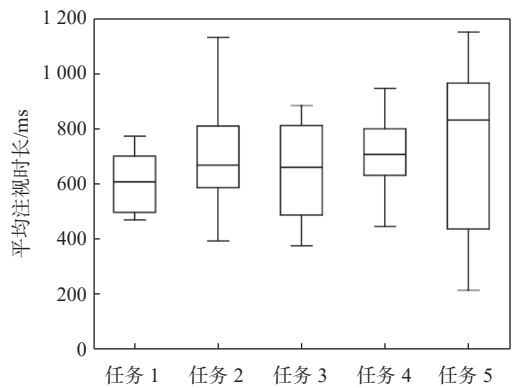


图 3　不同任务被试者平均注视时长箱型

Fig. 3　Box plots of subjects' mean gaze duration across tasks

当人眼对信息注视时间超过 50 ms 时, 便会形成注视点。注视时间越长, 代表人员对该信息更感兴趣或信息更复杂。由表 2 可知, 5 类任务下被试者的平均注视时长均值 600～730 ms。对实验样本进行统计分析, 设样本自由度 $df = 25$, 首先, 进行方差齐性检验, 莱文检验原假设: 各组方差相等, 符合方差齐性, 此时平均注视时长均值的 P 值(0.079)>0.05, 无法拒绝原假设, 即数据符合方差齐性。其次, 对 5 类任务下的眼动数据进行单因素方差分析, 得出检验统计量 F=0.255, 显著性 P=0.904>0.05, 以上数据分析表明, 不同任务下的被试者平均注视时长无明显差异, 即任务类型的改变对被试者的平均注视时长数据设有显著影响。

通过对不同被试者的平均注视时长的实验数据进行统计分析, 如表 3 所示, 样本自由度 $df = 24$, 首先, 进行方差齐性检验, 莱文检验原假设: 各组方差相等, 符合方差齐性, 此时平均注视时长均值的

表3 不同被试者的平均注视时长

Table 3 Average gaze duration of different subjects

被试者	平均注视时长/ms	标准差/ms	最大值/ms	最小值/ms
1	720.80	210.48	1 003.00	477.00
2	851.60	187.30	1 151.00	660.00
3	672.20	106.30	806.00	555.00
4	400.00	61.22	470.00	314.00
5	692.00	173.31	858.00	434.00
6	743.60	347.42	1 132.00	214.00

P值(0.124)>0.05,无法拒绝原假设,即数据符合方差齐性。对6位被试者的眼动数据进行单因素方差分析,得出检验统计量F = 2.784,显著性P = 0.04<0.05,数据表明,不同被试者平均注视时长存在显著差异。

通过对各被试者的多重比较,如表4所示,通常取显著性水平为0.05和95%的置信区间,利用SPSS 27软件计算P值。仅被试者4与其他5位被试者的平均注视时长均存在显著差异,其他5位被试者之间则不存在显著差异,个体差异性较小。

通过上述数据分析可知,不同任务类型和不同被试者对指控系统任务执行中的平均注视时长无显著影响,这表明用户平均注视时间相近,可以用于增加KLM参数。结合表5对平均注视时长的总体描述统计结果(样本量=30),为避免实验数据中极端值的影响,本文将采用中位数,即第50个百分位数下的平均注视时长作为一次视觉搜索处理用时的统计特征,即单次人员视觉搜索处理时间为0.668 5 s。在为后续实验中方便计算,取值为0.67 s。

2.3.6 KLSM定义

构建KLSM,先要定义操作单元。KLSM的操作单元除沿用KLM的点击、指向和心理操作3个操作符外,还包括视觉搜索操作符。基于2.3节实验数据分析,用户查看一个信息点所用的视觉搜索处理时间约为0.67 s,视觉搜索操作符量化时间S确定为0.67 s。

完成任务执行总时间可以表示为

$$T_{\mathrm{KLSM}} = T_K + T_P + T_S + T_M + \cdots \tag{2}$$

式中:T_{KLSM}为预估的总时间;T_S为视觉搜索的总耗时。

表6为改进后KLSM的基本参数。

表4 各被试者多重比较

Table 4 Multiple comparisons for each subject

被试者	比较被试者	注视时长平均值差值/ms	P值	95%置信区间	
				下限	上限
1	2	−130.800	0.316	−394.66	133.06
	3	48.600	0.707	−215.26	312.46
	4	320.800*	0.019	56.94	584.66
	5	28.800	0.824	−235.06	292.66
	6	−22.800	0.860	−286.66	241.06
2	1	130.800	0.316	−133.06	394.66
	3	179.400	0.173	−84.46	443.26
	4	451.600*	0.002	187.74	715.46
	5	159.600	0.224	−104.26	423.46
	6	108.000	0.407	−155.86	371.86
3	1	−48.600	0.707	−312.46	215.26
	2	−179.400	0.173	−443.26	84.46
	4	272.200*	0.044	8.34	536.06
	5	−19.800	0.878	−283.66	244.06
	6	−71.400	0.582	−335.26	192.46
4	1	−320.800*	0.019	−584.66	−56.94
	2	−451.600*	0.002	−715.46	−187.74
	3	−272.200*	0.044	−536.06	−8.34
	5	−292.000*	0.032	−555.86	−28.14
	6	−343.600*	0.013	−607.46	−79.74
5	1	−28.800	0.824	−292.66	235.06
	2	−159.600	0.224	−423.46	104.26
	3	19.800	0.878	−244.06	283.66
	4	292.000*	0.032	28.14	555.86
	6	−51.600	0.690	−315.46	212.26
6	1	22.800	0.860	−241.06	286.66
	2	−108.000	0.407	−371.86	155.86
	3	71.400	0.582	−192.46	335.26
	4	343.600*	0.013	79.74	607.46
	5	51.600	0.690	−212.26	315.46

注:*表示平均值差值的显著性水平为0.05。

表5 平均注视时长总体描述统计结果

Table 5 Overall descriptive statistical results for mean gaze duration

平均值/ms	标准差/ms	最小值/ms	最大值/ms	百分位数下的平均注视时长/ms		
				第25个	第50个	第75个
680.03	231.148	214	1 151	475.25	668.50	840.00

表6 KLSM基本参数

Table 6 Basic parameters of KLSM

操作	典型值/s	含义
K	0.2	敲击键盘中一个键所用的时间
P	1.1	用户指向显示器某个位置所用的时间
S	0.67	用户查看一个信息点所用的时间
M	1.35	用户进入下一步所需心理准备时间

3　KLSM 有效性的验证

操作人员在使用指控系统执行任务的过程中, 通常需要进行点击鼠标、操作键盘、查看界面信息等一系列操作。本文选取与 2.3 节不同的某指控系统的 5 类典型任务作为实验任务, 使用 KLM 和 KLSM 预测量化任务操作时间, 并将预测的任务操作时间与实际操作时间进行对比, 以验证 KLSM 的有效性。被试者要求与人员情况与 2.3.3 节相同。

3.1　任务步骤

以 5 类典型任务中的某任务(任务 1)为例进行说明, 其可分为开启界面、规划任务、修改信息、绑定站点、集群申请 5 个任务流程, 将成员分配至指定站台, 并实现站点绑定。图 4 为任务总体流程。



图 4　任务总体流程

Fig. 4　Overall task process

3.2　实验数据分析

使用 KLM 和 KLSM 对系统中的任务操作步骤进行详细分解, 人员在每进入下一个步骤之前都会有心理操作, 对于无法预测或需要处理的信息也需要一定的心理操作时间, 通常取值为 M。将任务 1 进行任务拆解, 在各具体操作步骤中分别计算 2 类模型的单元分解数, 具体如表 7 所示。

表 7　2 类模型的任务步骤分解

Table 7　Decomposition of task steps for two types of models

任务流程	具体操作步骤	KLM单元分解	KLSM单元分解
开启界面	①双击系统图标, 点击下拉框, 选择相应席位, 点击“登录”	$2P+5K+1M$	$2P+5K+1M+3S$
	②点击“集群”“任务管理”, 弹出任务界面	$1P+2K$	$1P+2K+2S$
规划任务	③在任务列表中选择一个任务, 点击“任务规划”, 弹出窗口	$1M+1P+2K$	$1M+1P+2K+2S$
	④点击“任务规划调整”进行2位成员的分组修改, 点击“提交修改”	$2P+4K+1M$	$2P+4K+1M+4S$
修改信息	⑤选择通道, 点击“成员信息”“信道调整”, 弹出对话框, 选中成员, 拖动信道调整	$1M+2P+3K$	$1M+2P+3K+6S$
	⑥点击“提交”, 弹出“保存成功”提示框, 点击确定	$1P+2K$	$1P+2K+3S$
绑定站点	⑦点击“集群”“规划”弹出对话框。选中信息, 点击“绑定”	$1M+1P+4K$	$1M+1P+4K+4S$
集群申请	⑧在信息框中选择可用站点, 点击“申请”	$1M+1P+2K$	$1M+1P+2K+2S$
	⑨弹出确认弹窗, 先点击对应标识, 再点击“确定”, 完成申请	$1P+2K$	$1P+2K+2S$

结合表 7 计算出模型评估时间:

$$\begin{aligned}T_{KLM} &= T_1+T_2+T_3+T_4+T_5 = \\ &(1M+3P+7K)+(2M+3P+6K)+\\ &(1M+3P+5K)+(2M+1P+4K)+\\ &(1M+2P+4K)=6M+12P+26K=26.50\ \text{s}\end{aligned} \quad (3)$$

$$\begin{aligned}T_{KLMS} &= T_1+T_2+T_3+T_4+T_5 = \\ &(1M+3P+7K+6S)+(2M+3P+6K+6S)+\\ &(1M+3P+5K+9S)+(1M+1P+4K+4S)+\\ &(1M+2P+4K+4S)=6M+12P+\\ &26K+29S=45.93\ \text{s}\end{aligned} \quad (4)$$

由式(3)和式(4)计算可得出, KLM 预测的任务操作时长为 26.50 s, KLSM 预测的任务操作时长为 45.93 s。

本实验选取 6 位被试者, 按照任务步骤进行操作并对任务进行计时, 各被试者的任务操作时长如表 8 所示。

表 8　任务操作总时长

Table 8　Total task operation time

被试者	操作总时长/s
1	52.508
2	45.031
3	76.704
4	59.386
5	115.252
6	42.060

剔除被试者 5 的极端异常值后, 计算得出其他 5 名被试者的平均任务操作时长为 55.138 s, 即为实际任务时长。

按照上述方法, 对 5 类典型任务分解的操作单元情况如表 9 所示, KLM 预测时长、KLSM 预测时长及实际任务时长数值如表 10 所示。

表 9　5 类任务操作单元分解对比

Table 9　Comparison of task decomposition under five types of tasks

任务	KLM模型	KLSM模型
任务1	$6M+12P+26K$	$6M+12P+26K+29S$
任务2	$8M+13P+28K$	$8M+13P+28K+30S$
任务3	$4M+10P+22K$	$4M+10P+22K+19S$
任务4	$15M+17P+40K$	$15M+17P+40K+10S$
任务5	$5M+11P+23K$	$5M+11P+23K+22S$

3.3　验证结果

实验通过对某指控系统的 5 类典型任务进行模型有效性验证得出, 视觉搜索处理时间即为 KLSM 预测时长与 KLM 预测时长的差值, 被试者

表 10 5类任务下模型预测时长对比

Table 10 Comparison of model prediction duration under five types of tasks

任务	KLM预测时长/s	KLSM预测时长/s	视觉搜索处理时间/s	实际任务时长/s
任务1	26.50	45.93	19.43	55.138
任务2	30.70	50.80	20.10	60.365
任务3	20.80	33.53	12.73	40.267
任务4	46.95	53.65	6.70	58.602
任务5	23.45	38.19	14.74	44.820

在5类典型任务的视觉加工处理步骤花费的平均时间为14.74 s。

结果表明，增加了用户视觉搜索时间的KLSM的预测时长，比KLM的预测时长更接近实际任务时长，因此KLSM对任务完成时间预测评估更为准确。

4 结 论

1）被试者执行5类典型任务的平均注视时长为600～730 ms，执行不同类型的任务时，被试者的视觉搜索处理时间无显著差异。这说明被试者充分熟悉系统后，任务类型不会对其视觉搜索处理时间产生显著影响；不同的被试者之间的平均注视时间基本不存在明显的个体差异性。本文得到的搜索操作单元参数的典型时间值为0.67 s，可以将该数值作为评估模型的基本参数。

2）相较于其他简单系统软件界面，指控系统包含了更多复杂专业的态势信息，被试者需要花费较长时间进行搜索和加工处理。而增加了用户视觉搜索时间的KLSM能够有效反映真实的任务处理过程，克服了KLM的不足，其预测的任务时间更接近任务实际操作时长，从而提升了预测的准确性。本文的KLSM可用于指导指控系统界面设计评估、分析与优化。

KLSM继承了KLM易学易用、效率高、良好的信度和效度等优点，但同时也存在无差错执行、序列操作、操作符线性叠加等前提假设带来的误差。此外，恰当的系统响应时间等外界环境也是影响模型预测准确度的一个重要方面，这为未来的进一步研究指明了方向。

参考文献（References）

[1] 王璐, 霍其恩, 李青山, 等. 基于并行搜索优化的指控系统自适应决策方法[J]. 软件学报, 2022, 33(5): 1774-1799.
WANG L, HUO Q E, LI Q S, et al. Self-adaptation decision-making based on parallel search optimization for command and control information system[J]. Journal of Software, 2022, 33(5): 1774-1799(in Chinese).

[2] 周海瑞, 李皓昱, 介冲. 美军联合作战指挥体制及其指挥控制系统[J]. 指挥信息系统与技术, 2016, 7(5): 10-18.
ZHOU H R, LI H Y, JIE C. Military joint operation command system and its C^2 systems[J]. Command Information System and Technology, 2016, 7(5): 10-18(in Chinese).

[3] 王颖, 牟苏斌. 基于GOMS模型的机载指控系统人机界面优化研究[J]. 计算机与数字工程, 2016, 44(7): 1378-1382.
WANG Y, MOU S B. Human computer interface optimization based on GOMS model for airborne command and control system[J]. Computer & Digital Engineering, 2016, 44(7): 1378-1382(in Chinese).

[4] 陈建明, 金传洋, 魏琨. 装甲指控软件的可用性问题研究[J]. 计算机测量与控制, 2013, 21(8): 2278-2281.
CHEN J M, JIN C Y, WEI K. Study on usability problem of command and control software for armored vehicle[J]. Computer Measurement& Control, 2013, 21(8): 2278-2281(in Chinese).

[5] REN F J, BAO Y W. A review on human-computer interaction and intelligent robots[J]. International Journal of Information Technology & Decision Making, 2020, 19(1): 5-47.

[6] CARD K S. The psychology of human-computer interaction[M]. New York: CRC Press, 2018: 12-48.

[7] SCHREPP M. GOMS analysis as a tool to investigate the usability of web units for disabled users[J]. Universal Access in the Information Society, 2010, 9(1): 77-86.

[8] JUNG K, JANG J. Development of a two-step touch method for website navigation on smartphones[J]. Applied Ergonomics, 2015, 48: 148-153.

[9] HUANG P H, CHIU M C. Integrating user centered design, universal design and goal, operation, method and selection rules to improve the usability of DAISY player for persons with visual impairments[J]. Applied Ergonomics, 2016, 52: 29-42.

[10] RAMKUMAR A, STAPPERS P J, NIESSEN W J, et al. Using GOMS and NASA-TLX to evaluate human–computer interaction process in interactive segmentation[J]. International Journal of Human-Computer Interaction, 2017, 33(2): 123-134.

[11] ROSE J A, BEARMAN C. Making effective use of task analysis to identify human factors issues in new rail technology[J]. Applied Ergonomics, 2012, 43(3): 614-624.

[12] CHRISTOU G, RITTER F E, JACOB R J K. Codein—A new notation for GOMS to handle evaluations of reality-based interaction style interfaces[J]. International Journal of Human-Computer Interaction, 2012, 28(3): 189-201.

[13] CARD S K, MORAN T P, NEWELL A. The psychology of human-computer interaction[M]. Mahwah: Lawrence Erlbaum Associates Inc, 1983: 12-33.

[14] FINDLATER L, FROEHLICH J E, FATTAL K, et al. Age-related differences in performance with touchscreens compared to traditional mouse input[C]//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: ACM, 2013: 343-346.

[15] 赵欣, 陈凯翔, 谷贺瑾, 等. 基于触屏移动操作界面的KLM模型改良[J]. 包装工程, 2018, 39(16): 31-35.
ZHAO X, CHEN K X, GU H J, et al. Evolving KLM for touchscreen based mobile interface[J]. Packaging Engineering, 2018, 39(16): 31-35(in Chinese).

[16] 张楠, 苏玉婷, 廖斌. 时间压力视角下人格特征与视觉搜索策略

及绩效的关系[J]. 人类工效学, 2023, 29(3): 25-29.
ZHANG N, SU Y T, LIAO B. A study on the relationship between personality trait and visual search strategy and performance in a time pressure perspective[J]. Chinese Journal of Ergonomics, 2023, 29(3): 25-29(in Chinese).
[17] 田晨智, 宋敏, 田继伟, 等. 基于眼动追踪的指挥控制系统人机交互效能评估方法[J]. 航空兵器, 2023, 30(4): 57-66.
TIAN C Z, SONG M, TIAN J W, et al. Evaluation method of human-computer interaction effectiveness of command and control system based on eye tracking[J]. Aero Weaponry, 2023, 30(4): 57-66(in Chinese).
[18] GOLDBERG J H, KOTVAL X P. Computer interface evaluation using eye movements: Methods and constructs[J]. International Journal of Industrial Ergonomics, 1999, 24(6): 631-645.
[19] POOLE A, BALL L J. Eye tracking in HCI and usability research [J]. Medicine, 2005, 1: 211-219.

Considering visual search characteristics on command and control system KLSM

GUO Zhong[1, 2, *], ZHAO Gang[1], FU Bin[2], LIU Jiaye[3]

(1. Systems Engineering Institute, Academy of Military Science, PLA, Beijing 100101, China;
2. Unit 93216 of PLA, Beijing 100085, China;
3. Unit 93131 of PLA, Beijing 100038, China)

Abstract: The traditional keystroke level model (KLM) focuses on the operation analysis of the keyboard and mouse, which is well-suited for efficiently quantifying job completion time in straightforward software interfaces. However, a significant error occurs when the software interface of the command and control system is transferred to evaluate complicated information characteristics. This paper introduces a physical operation time unit index for the command and control system, based on the visual search processing time of users. The index is derived from experiments that determine the typical parameters of user visual search processing time. Additionally, a more precise task operation analysis model called keystroke level and visual search model (KLSM) is established, which takes into account the impact of visual search characteristic factors. The comparative test examines the accuracy of the KLSM model in forecasting and evaluating the task completion time by using 5 typical command and control tasks as examples.

Keywords: visual search characteristics; command and control system; software interface; KLSM; prediction and evaluation

Received: 2024-03-04; **Accepted**: 2024-03-15; **Published Online**: 2024-03-25 12:17
URL: link.cnki.net/urlid/11.2625.V.20240325.0956.001
Foundation item: National Social Science Foundation of China (2022-SKJJ-B-043)
* **Corresponding author**. E-mail: 18803806@qq.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0746

基于机器学习的高超声速飞行器双曲率前缘气动热预测方法

杨帆[1,2]，林明月[1,2]，胡宗民[1,2,*]，罗长童[1,2]

(1. 中国科学院力学研究所 高温气体动力学国家重点实验室，北京 100190；2. 中国科学院大学 工程科学学院，北京 100049)

摘 要：高超声速气动热预测技术是高超声速飞行器发展的关键技术之一，气动热环境的精准预测对飞行器热防护系统设计及气动布局优化具有重要意义。为快速获得高超声速飞行器表面的热流分布情况，缩短飞行器设计周期，基于具有广义可分离特性、可实现强非线性数据快速建模的多层分块(MBB)算法，提出一种针对高超声速飞行器双曲率前缘气动热分布的快速预测方法。通过数值计算获得双曲率前缘驻点区的气动热分布作为训练集数据，基于MBB算法提出预测热流分布的显式表达式，对表达式预测结果的统计分析显示，表达式预测值与测试集数据的偏差低于2%，这表明其具有较高的预测精度；将驻点区热流分布表达式进行外推，验证了机器学习公式在不同几何外形下的适用性。在双曲率前缘构型的防热设计及气动外形优化阶段，所提表达式可实现气动热环境的精准、快速预测。

关 键 词：高超声速流动；气动热；多层分块算法；快速预测方法；驻点热流；双曲率前缘
中图分类号：V221+.3; TB553
文献标志码：A **文章编号**：1001-5965(2024)09-2826-09

为进一步提升进入空间、探索空间、利用空间的能力，发展新一代高超声速飞行器，实现便捷、经济、安全的飞行是高超声速飞行器发展的必然趋势[1-3]。飞行器高速飞行时，空气的剧烈压缩和摩擦使得驻点区域产生极为严酷的气动加热，加热量随着飞行速度的增加呈指数级上升，飞行器面临着棘手的“热障”问题[4-5]。采用大钝头外形可以有效缓解驻点气动热，但这是以极大牺牲飞行器气动性能为代价的。发展以吸气式发动机为动力的升力体外形要求采用低阻力的尖化前缘[6-7]，驻点区域面临更加严酷的气动热问题，从而对热防护系统的设计提出了更高的要求。因此，高超声速飞行器前缘驻点区气动热环境的准确预测对飞行器热防护系统的精准设计及气动布局优化具有重要意义。

对高超声速飞行器气动热环境的精准预测十分困难[8-10]，特别是前缘驻点区极端热问题突出，是气动热问题研究的重点及难点。自20世纪50年代起，针对驻点区域气动热问题展开了一系列理论[10-13]和数值[14-18]研究。在理论方面，驻点区热流的研究主要基于边界层的自相似解，进行了一系列的推广和修正后，获得了计算驻点热流的理论公式。1958年，Fay和Riddell[11]利用曼格勒变换将高温气体边界层偏微分方程转化为常微分方程，推导出至今仍广泛应用的Fay-Riddell驻点热流公式，其可用于直接求解球头和圆柱的驻点热流值。1957年，Kemp和Riddell[12]对卫星再入的实际轨道进行了大量计算，并将来流流动参数与正激波后流动参数相结合，给出了Kemp-Riddell驻点热流公式。除此之

收稿日期：2022-08-30；录用日期：2022-12-05；网络出版时间：2023-02-14 14:11
网络出版地址：link.cnki.net/urlid/11.2625.V.20230214.0911.001
基金项目：国家自然科学基金 (12172365,12072353,12132017)；国家重点研发计划 (2019YFA0405204)
*通信作者. E-mail：huzm@imech.ac.cn

引用格式：杨帆，林明月，胡宗民，等. 基于机器学习的高超声速飞行器双曲率前缘气动热预测方法[J]. 北京航空航天大学学报，2024，50(9)：2826-2834. YANG F，LIN M Y，HU Z M，et al. Prediction method of aero-heating of hypersonic vehicle bi-curvature leading edge based on machine learning[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2826-2834 (in Chinese).

外，较为重要的驻点热流公式还有 Lees 公式[19]、Scala 公式[20] 等，不同公式的适用范围不同，用于工程计算时，要根据具体流动情况选用公式。理论公式可以清楚地揭示气动热与各物理因素之间的内在联系，并给出气动热的快速估计。但现有的理论公式只能计算球头和圆柱这类简单外形，对于高超声速飞行器的尖化前缘外形，即使是较为简单的介于圆柱和球头之间的双曲率外形，也没有较为精准的热流计算公式。随着计算流体力学（computational fluid dynamics, CFD）技术的迅猛发展，采用数值模拟方法可以更精准、细致地刻画流场，这也成为气动热预测的重要研究手段[8-9]。研究表明，数值格式和网格分布都会影响气动热的数值计算[15-16,21-23]，特别是近壁面法向网格间距对数值计算结果的影响较大。1988 年，Klopfer 和 Yee[22] 发现钝体驻点热流值取决于网格及格式的选取，要得到准确的热流值，网格雷诺数要小于 3。1998 年，Lee 和 Rho[23] 在钝头体气动热计算中探讨了网格相关性问题，发现在其他条件均保持不变的情况下，壁面网格雷诺数扩大一倍，热流计算偏差值能达到 20% 以上。马崇立和刘景源[24] 应用理论和数值模拟方法，提出一种基于壁面网格雷诺数及钝体特征长度的来流雷诺数的网格划分方法，结果表明，对于工程上采用的能有效抑制数值黏性的格式，壁面网格雷诺数可供参考的取值为 20～80。

目前，利用数值模拟方法研究气动热问题，并与地面实验互为验证补充，是对气动热问题的主流研究手段。但在工程设计的初期阶段，开展大规模数值模拟要付出昂贵的代价，否则将难以满足对气动热精准快速计算的需求。而常用于工程设计初期的经验公式，往往具有一定的局限性，且常难以满足精度的要求。因此，发展一种与 CFD 具有相当精度的高超声速飞行器气动热环境的快速分析预测方法，可以有效缩短气动热的设计周期。随着人工智能技术的发展，机器学习方法的兴起为气动热的计算与分析提供了一种新的、强有力的研究手段[25-33]。Chen 等[27] 开展了气动热降阶模型研究，用于实现快速精准的气动热预测，并通过数值模拟进行了验证。聂春生等[28] 利用本征正交分解（proper orthogonal decomposition, POD）方法对数值计算得到的数据库进行降阶处理，建立了基于 POD 方法的复杂外形飞行器表面热流计算代理模型。张智超等[29] 利用数值计算获得的飞行器表面热流数据，基于径向基网络模型，开展了高超声速飞行器气动热快速预测的代理模型方法的研究。这些研究利用数据驱动的方法降低了气动热问题的分析难度，提高了热流计算的效率。Chen 等[30] 提出可以实现快速分析建模气动热数据的符号回归方法——多层分块 (multi-level block building, MBB) 算法，即以实验或数值模拟数据为基础，利用符号回归的方法逼近真实模型，从而得到简洁的显式表达式，进而探索其中的物理规律。除此之外，MBB 算法还能够处理大规模、多变量、强非线性的问题，并且具有良好的外推能力。

本文基于 MBB 算法提出高超声速飞行器双曲率前缘表面热流分布的快速预测公式，以实现设计初期阶段对飞行器气动热环境的快速预测。

1　数值方法与机器学习方法

1.1　控制方程与数值方法

本文涉及的马赫数为 5～10，总温约 2 000 K，采用仅考虑气体分子振动激发且处于振动平衡状态的热完全气体模型，数值计算的控制方程为三维可压缩的 Navier-Stokes 方程组，具体形式如下：

$$\frac{\partial \boldsymbol{U}}{\partial t}+\frac{\partial (\boldsymbol{F}-\boldsymbol{F}_{\mathrm{vis}})}{\partial x}+\frac{\partial (\boldsymbol{G}-\boldsymbol{G}_{\mathrm{vis}})}{\partial y}+\frac{\partial (\boldsymbol{H}-\boldsymbol{H}_{\mathrm{vis}})}{\partial z}=\boldsymbol{0} \tag{1}$$

式中：$\boldsymbol{U}$ 为守恒变量矢量；$\boldsymbol{F}$、$\boldsymbol{G}$、$\boldsymbol{H}$ 分别为 x、y、z 方向上的对流通量矢量；$\boldsymbol{F}_{\mathrm{vis}}$、$\boldsymbol{G}_{\mathrm{vis}}$、$\boldsymbol{H}_{\mathrm{vis}}$ 分别为 x、y、z 方向上的黏性通量矢量。

$$\boldsymbol{U}=\begin{bmatrix}\rho u\\ \rho u\\ \rho v\\ \rho w\\ E\end{bmatrix} \tag{2a}$$

$$\boldsymbol{F}=\begin{bmatrix}\rho u\\ \rho u^2+p\\ \rho uv\\ \rho uw\\ (E+p)u\end{bmatrix} \tag{2b}$$

$$\boldsymbol{F}_{\mathrm{vis}}=\begin{bmatrix}0\\ \tau_{xx}\\ \tau_{xy}\\ \tau_{xz}\\ u\tau_{xx}+v\tau_{xy}+w\tau_{xz}+q_x\end{bmatrix} \tag{2c}$$

$$\boldsymbol{G}=\begin{bmatrix}\rho v\\ \rho \mathrm{u}v\\ \rho v^2+p\\ \rho vw\\ (E+p)v\end{bmatrix} \tag{2d}$$

$$\boldsymbol{G}_{\mathrm{vis}}=\begin{bmatrix}0\\ \tau_{yx}\\ \tau_{yy}\\ \tau_{yz}\\ u\tau_{yx}+v\tau_{yy}+w\tau_{yz}+q_y\end{bmatrix} \tag{2e}$$

$$\boldsymbol{H}=\begin{bmatrix}\rho w\\ \rho uw\\ \rho vw\\ \rho w^2+p\\ (E+p)w\end{bmatrix} \tag{2f}$$

$$\boldsymbol{H}_{\text{vis}}=\begin{bmatrix}0\\ \tau_{zx}\\ \tau_{zy}\\ \tau_{zz}\\ u\tau_{zx}+v\tau_{zy}+\tau_{zz}+q_z\end{bmatrix} \tag{2g}$$

式中：ρ、p 分别为气体的密度、压力；u、v、w 分别为 x、y、z 方向上气体的速度；q_x、q_y、q_z 分别为 x、y、z 方向上的热通量密度；τ 为黏性应力；E 为单位体积气体的总能量，表示为

$$E=\rho h-p+\frac{1}{2}\rho\left(u^2+v^2+w^2\right) \tag{3}$$

其中：h 为比焓。

考虑气体分子振动激发，比焓 h 与定压比热容 c_p 都是温度 T 的函数，可以分别通过 NASA[34] 的 7 系数多项式拟合出来，如下：

$$\begin{cases}\dfrac{c_{pi}(T)}{R_i}=a_{i1}+a_{i2}T+a_{i3}T^2+a_{i4}T^3+a_{i5}T^4\\ \dfrac{h_i(T)}{R_iT}=a_{i1}+a_{i2}\dfrac{T}{2}+a_{i3}\dfrac{T^2}{3}+a_{i4}\dfrac{T^3}{4}+a_{i5}\dfrac{T^4}{5}+\dfrac{b_{i1}}{T}\end{cases} \tag{4}$$

式中：i 表示空气的不同组分；系数 a_{i1}～a_{i5} 和 b_{i1} 可查阅参考文献；R 为气体常数。

本文数值计算采用的是有限体积法（finite volume method, FVM）。对流项采用多维多项式插值的二阶总变差不增 (total variation diminishing, TVD) 格式对界面值进行重构，并用近似黎曼求解器 (Harten-Lax-van Leer contact, HLLC) 对界面处的数值通量进行计算，HLLC 格式能够准确地捕捉激波，接触间断和稀疏波，减少计算量、降低耗散、提高分辨率[35]；黏性项采用多维多项式插值方法进行计算，精度同样为二阶；时间推进采用隐式双时间步长法进行积分计算。

1.2 机器学习方法

对于机器学习方法与流体力学相结合的研究，依据所使用的模型，大概可分为 2 种类型：①预先给定模型结构，通过相应的策略优化模型参数，可称为代理模型；②模拟人工推导过程，建立具有显式表达式的模型，如基于遗传规划 (genetic programming, GP) 的符号回归方法。

符号回归方法是一种基于数据驱动的回归分析方法，在实际应用时，不需要对目标模型的结构（具体的函数形式和系数）强加假设，其目标是在给定的函数算子和运算算子所定义的函数空间中进行遍历式搜索，以实现最佳逼近，并获得最优的回归模型。符号回归方法从数据中获得模型，不仅可以给出显式的表达式，同时还具有外推和预测能力。

GP[36] 是基于数据驱动函数优化的最典型的一类方法，其核心思想是达尔文的自然进化理论。理论上，只要提供足够长的计算时间，GP 就可以得到最佳拟合数据集的函数模型。然而，对于具有多变量的大规模问题，GP 的收敛速度可能会非常慢，从而限制了其在工程中的实际应用。为此，Chen 等[37] 提出块构建编程 (block building programming, BBP) 算法，将原始目标函数分成几个块，并进一步分解为多个因子，通过减小搜索空间来提高计算效率，但由于变量不能出现在不同的子函数中，因此子函数没有被完全分离。陈辰[38] 进一步提出了一种广义可分离模型 (generalized separable model, GSM)，将变量分为重复变量和非重复变量，使变量可以出现在不同的子函数中，以进一步分离子函数，从而可以减小搜索空间，提高计算效率。为识别具有可能 GSM 特征的函数的结构，Chen 等[30] 提出基于 BBP 改进的 MBB 算法，将目标模型分解为更高层次的块和因子，直到被确认为最小块和因子，再对最小块和因子进行适当组装，即形成目标函数。与大多数传统 GP 或其他非进化算法相比，MBB 算法中的最小因素更容易确定，因此，MBB 算法更加实用和高效，在工程上具有重要的应用价值，可以为气动数据的分析与建模提供精准快速的算法支持。文献 [31-33] 利用 MBB 算法，分别给出了不同气体模型下，激波干扰流场结构的几何特征及产生极端载荷的位置，并对其精度及适用性进行了评估，取得了良好效果。图 1 为 MBB 算法的示意[30]。

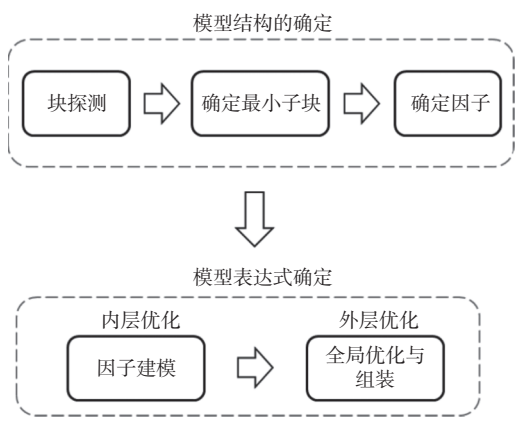


图 1 MBB 算法的示意[30]

Fig. 1 Flowchart of MBB algorithm [30]

目前，球、柱等简单外形已有驻点区热流计算公式，但对于介于球柱之间、广泛存在于高超声速飞行器特别是乘波体前缘的双曲率构型，还没有计

算公式来实现对不同几何外形下驻点区热流的精准快速预测。因此, 利用具有广义可分离特性、可实现强非线性数据快速建模的 MBB 算法, 对双曲率前缘构型的气动热数据进行分析建模, 从而获得最优的回归模型, 并给出双曲率构型驻点区热流分布的显式表达式。算法中主要参数设置如下: 函数模块解析矩阵高度为 $5d$, 决定了单个函数模块的复杂度, 其中, d 为自变量个数。种群规模为 $10d$, 最大进化代数为 $20d$, 降维概率为 0.5, 当残差下降至 10^{-6} 时, 算法停止运行。

1.3　算例验证

为验证数值计算结果的可靠性, 选取具有详细实验数据的钝锥[39] 进行算例验证, 其前缘流动情况与本文模型相似。实验[39] 中, 钝锥模型的长度为 447 mm, 头部曲率半径为 27.94 mm, 半锥角为 15°。实验来流条件为 $Ma_\infty = 10.6$, $T_\infty = 47.3$ K, $Re_\infty = 3.937\times 10^6$/m, 壁面为等温壁 $T_w = 294.44$ K。图 2 为数值计算的热流分布与实验值[39] 对比, 横坐标为轴向距离与头部半径的比值, 纵坐标为无量纲热流。

Fay-Riddell 公式为[11]

$$q = 0.763Pr^{-0.6}(\rho_s\mu_s)^{0.4}(\rho_w\mu_w)^{0.1}\sqrt{\frac{\mathrm{d}u_e}{\mathrm{d}x}}(h_s - h_w) \quad (5)$$

式中: Pr 为普朗特数, ρ_s、μ_s 和 h_s 分别为驻点处的密度、黏性系数和焓值; ρ_w、μ_w 和 h_w 分别为壁面处的密度、黏性系数和焓值; u_e 为驻点边界层外缘的速度。

由式 (5) 计算出驻点热流值 q_{ref} 归一化, 其中, q_{ref} 的值为 2.16×10^5 W/m^2。

由图 2 可知, 本文数值计算的钝锥驻点区热流分布与实验值[39] 吻合较好, 且计算的驻点热流值与 Fay-Riddell 公式[11] 的理论值仅偏差 3.7%, 这表明本文使用的数值计算方法可以较为精确地获得高超声速飞行器前缘驻点区的热流分布情况。

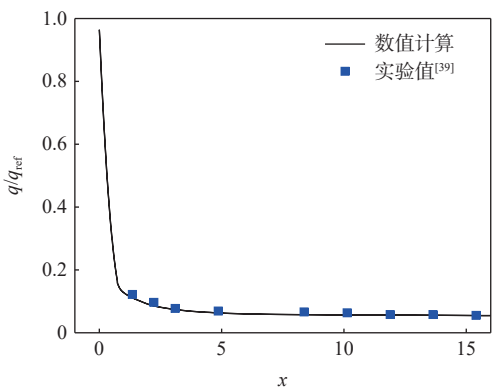


图 2　钝锥表面热流分布的数值计算结果与实验值[39] 对比

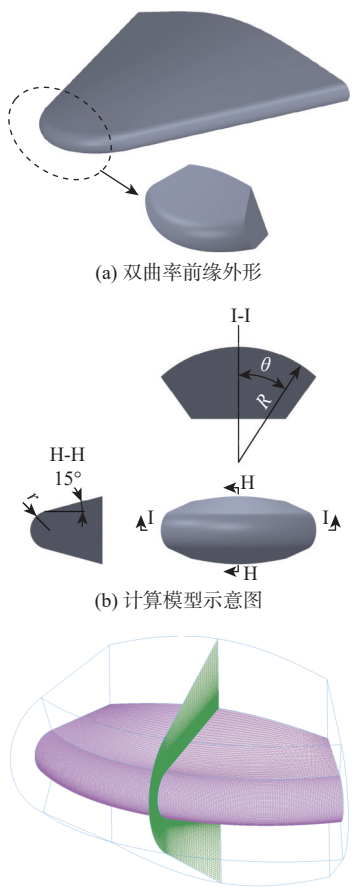
Fig. 2　Comparison of distributions of heat flux on the surface of cone obtained by numerical simulation and experimental value[39]

2　热流分布的机器学习公式

2.1　双曲率前缘模型及训练集数据构成

本文的高超声速飞行器双曲率前缘外形及模型的网格划分如图 3 所示。图 3(a) 为双曲率前缘外形示意图, 为减少计算量, 本文截取高热流集中的驻点区域 (见图 3(b)) 为计算域。如图 3(b) 所示, 计算模型的横向 (I-I 剖面) 与纵向对称面 (H-H 剖面) 分别为半径 R 与 r 的圆, 其中, R 为前缘倒圆半径, r 为钝化半径, 扩张角取固定值 15°。头部倒圆半径 R 与钝化半径 r 之比为本文表征双曲率前缘外形的特征量: 当 $R_0 = 1$ 时, 该模型退化为球头钝锥模型; 当 R_0 趋近于无限大时, 则趋于圆柱模型。θ 为模型横向截面 (I-I 剖面) 中的周向角。

模型的网格划分如图 3(c) 所示, 热流的计算对网格具有很强的依赖性, 主要体现在壁面法向的网格间距上[14-17,22-23]。为保证数值计算气动热数据的准确性, 本文对网格雷诺数 Re_{cell} 分别为 1、20、40 的网格进行收敛性验证, 结果如图 4 所示。当网格雷诺数小于 20 时, 热流的计算结果趋于收敛。



(a) 双曲率前缘外形

(b) 计算模型示意图

(c) 计算域网格划分

图 3　计算模型及网格划分

Fig. 3　Schematic diagram of computational modeling and meshing

本文使用的样本库数据均按此方法进行无关性检验，将获得的热流计算结果的近似收敛值作为机器学习的输入值。

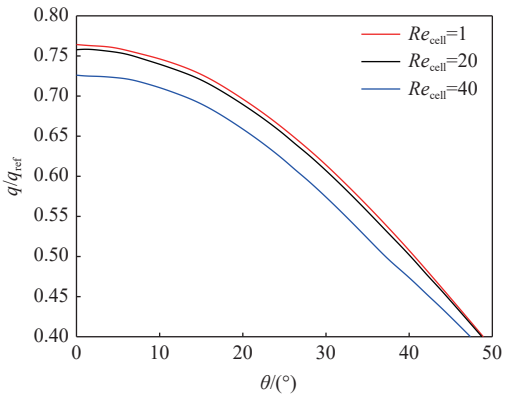


图 4 网格无关性检验

Fig. 4 Grid independence verification

对于 $R_0=1$ 的球头钝锥模型与 R_0 趋向于无穷的圆柱模型，驻点区的热流分布情况已有较为成熟的理论[10-13]公式进行预测。但对于双曲率尖化前缘 (见图 3(a)) 外形，则缺少驻点区热流分布的快速预测公式[40-43]。因此，本文通过数值模拟获得了不同来流条件和几何外形下的驻点区热流分布情况作为建立预测公式的样本数据库，并利用 MBB 算法进行分析与建模。数值计算的来流参数选取高空 $H=30$ km 的来流条件 ($p_\infty=1\,172$ Pa, $T_\infty=226.6$ K)，参数设置如表 1 所示。

表 1 数值模拟的参数设置

Table 1 Settings of parameters for numerical simulation

Ma	r/mm	R/mm	R_0
6, 7, 8, 9, 10	2	2, 4, 6, ⋯, 18, 20	1, 2, 3, ⋯, 9, 10

飞行器表面热流与来流条件、几何外形密切相关，在机器学习的过程中，输入的自变量为 Ma、R_0、$\sin\theta$，本文选取的计算域 (见图 3(b)) 中 θ 的为 0°～20°。q/q_{ref} 为无量纲热流值，q_{ref} 为相同来流条件下半径为 r 的球头钝锥模型的驻点热流值。机器学习的样本库共包含 1 500 余组数据，该数据通过 5×10 = 50 组算例的数值计算获得。在 MBB 算法中，含有优化的部分可分为内层优化和外层优化。内层优化主要解决函数结构的优化及因子内部系数优化；外层优化为全局参数的优化及函数结构的组装。损失函数在内部迭代中使用的是残差均方根，外部迭代中使用的是 $1-R^2$，其中，R^2 为拟合优度。

2.2 MBB 算法热流分布公式

基于以上 50 组算例的数值计算结果得到的双曲率前缘热流分布预测公式如表 2 所示。热流分布公式的准确性可由拟合优度 R^2 来衡量。通常，R^2 可用于评估表达式与样本库数据的符合程度，定义为回归平方和 S_{SR} 与总离差平方和 S_{ST} 的比值，$R^2=S_{SR}/S_{ST}$。该值介于 0～1，0 表示完全不符合，1 表示完全符合。一般认为，超过 0.8 时模型的拟合优度较高[44]。表 2 列出的是拟合优度最高的 4 个表达式，R^2 值均超过 0.99，这表明模型与训练集的回归拟合效果较好。从物理意义上讲，无量纲热流值与自变量是一一对应的关系，表达式应该是唯一的。但依靠现有的机器学习手段很难直接获取这个具有物理意义的准确表达式，只能得到数学上与其逼近的表达式。选取表 2 中编号为 1 的拟合优度最高的表达式进行分析，即

$$q/q_{ref}=0.803\,735-0.037\,163\,3\ln(Ma)-0.128\,599R_0^{-2}+0.398\,309R_0^{-1}-0.645\,404(\sin\theta)^2 \tag{6}$$

2.3 热流预测公式精度分析及简化

图 5 为不同 Ma 和几何外形下的热流分布。可以看出，本文公式准确预测了驻点区的热流分布趋势，数值模拟结果基本分布在预测公式附近，这表明式 (6) 对前缘驻点区的热流预测具有较高的精度。

进一步观察表 2 中的公式可以发现，Ma 对 q/q_{ref} 的影响仅存在于单独的一项中，且不与其他变量耦合。由图 5 中的热流分布曲线可知，不同 Ma 下的热流分布曲线近乎平行且彼此之间非常接近，Ma 从 6 变化到 10 的过程中，无量纲热流值仅改变了 1.9%，其对无量纲热流值的影响很小且 Ma 单调变化。基于此，对式 (6) 进行简化。对式 (6) 中含有 Ma 的一项，取 Ma 为 6～10 的中值替代此项，获得一个仅与几何外形 R_0 有关的热流分布公式，如下：

表 2 MBB 算法热流分布公式

Table 2 Formula of heat flux distributions obtained by MBB algorithm

序号	q/q_{ref}	R^2
1	$0.803\,735-0.037\,163\,3\ln(Ma)-0.128\,599R_0^{-2}+0.398\,309R_0^{-1}-0.645\,404(\sin\theta)^2$	0.993 480 17
2	$0.782\,952-0.014\,027\,4Ma^{2/3}-0.128\,64R_0^{-1}+0.398\,357R_0^{-1}-0.645\,377(\sin\theta)^2$	0.993 094 38
3	$0.743\,498-0.026\,448\,9Ma^{1/2}+0.324\,038R_0^{2/3}+0.012\,534\,3R_0^{1/3}-0.641\,078(\sin\theta)^2$	0.992 062 82
4	$0.706\,337-0.004\,669\,15Ma+0.324\,044R_2^{2/3}+0.012\,532\,5R_0^{1/3}-0.641\,065(\sin\theta)^2$	0.991 746 52

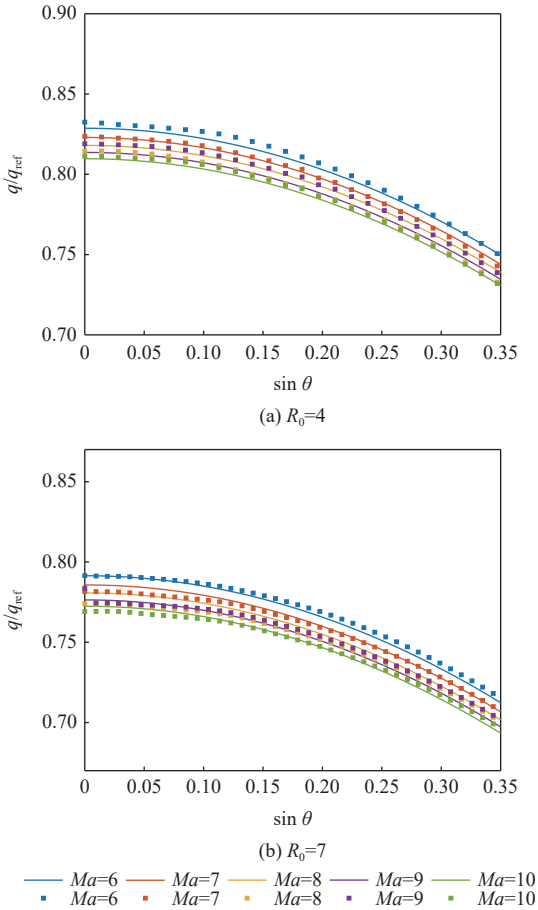


图 5　不同 *Ma* 和几何外形下的热流分布

Fig. 5　Distributions of heat flux at different *Ma* and geometries

$$q/q_{\text{ref}} = 0.727\,655\,3 - 0.128\,599R_0^{-2} + 0.398\,309R_0^{-1} - 0.645\,404(\sin\theta)^2 \quad (7)$$

简化后的式 (7) 与式 (6) 相比偏差保持在 1% 以内。

将 1 500 余组训练集数据与简化后的热流分布式(7)的计算结果进行统计分析, 并将机器学习公式计算出的无量纲热流与训练集数据的偏差值记为 e_{rror}。图 6 为不同 *Ma* 和几何外形下偏差值的分布, 由于简单地采用中值代替含有 *Ma* 的项, 因此 *Ma*=6、*Ma*=10 热流数据的偏差值较大。总体来看, 在不同的几何外形下, 热流偏差值的分布基本一致。由统计结果可知, 偏差的平均值为 0.67%, 所有偏差值均分布在 2% 以下, 超过 73.8% 的大部分偏差值分布在 1% 以下, 由此可见热流分布的机器学习公式具有较高的精度。

2.4　热流分布公式尺度量的外推

在用于机器学习的训练集数据中, 代表几何外形的变量 $R_0 = R/r$, 其中, r 在之前的研究中取固定值为 2 mm, 通过改变 R 来获得代表不同几何外形的 R_0 值。实际上, r 相当于双曲率前缘大小的尺度量。当 R_0 一定时, 改变 r, 钝头体会等比例放大或缩小。本节保证 R_0 不变, 改变 r, 以验证相同几何外形、不同尺度下的热流公式是否成立。新增算例的设置如表 3 所示。

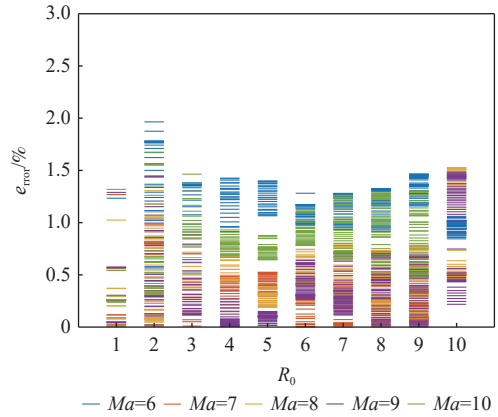


图 6　训练误差分布

Fig. 6　Distribution of training error

表 3　新增算例参数设置

Table 3　Settings of parameters for new cases

Ma	*r*/mm	R_0
6, 7, 8, 9, 10	3,4	5

图 7 为新增验证算例的数值计算结果与热流分布的机器学习公式的结果对比。可以看出, 对于 r=3,4 mm 的情况, 数值计算结果与热流分布公式吻合度很高。进行进一步统计分析后, 图 7 中的蓝色阴影区域为热流分布式(7)2% 的偏差带, 可以看出, 数值模拟结果全部落在偏差带内, 这说明本文机器学习得到的公式外推至相似的几何外形下也是成立的。这是由于本文在机器学习阶段时, 输入的 q/q_{ref} 为双曲率前缘驻点热流值与半径为 r 的球头钝锥的驻点热流值之比, 这实际上已将几何尺度 r 的影响包含其中。当几何外形等比例变化时, 弓形激波形状与驻点附近的流动具有相似性, 因而热流的分布也具有相似性。

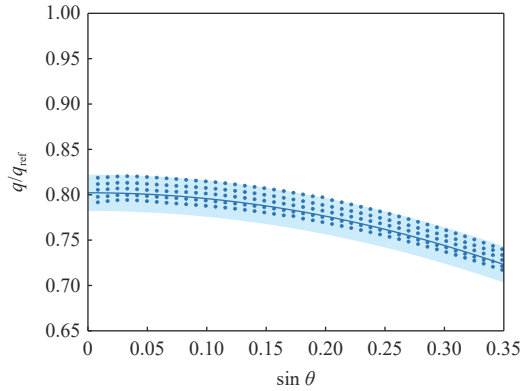
2.5　驻点热流值几何外形量的外推

在式 (7) 中, 当 θ=0°时, 预测公式计算得到的即为双曲率前缘的驻点热流值, 如下:

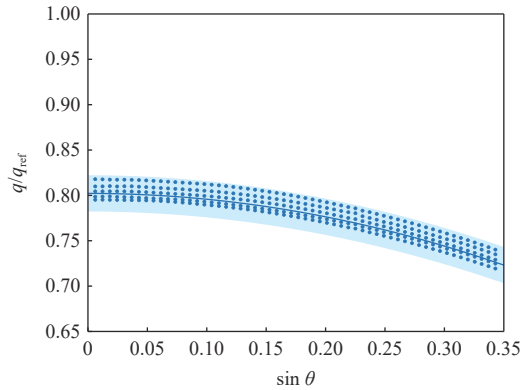
$$q/q_{\text{ref}} = 0.727\,655\,3 - 0.128\,599R_0^{-2} + 0.398\,309R_0^{-1} \quad (8)$$

图 8 为无量纲驻点热流值及其导数随几何外形量 R_0 的变化情况。由图 8 可知, 前缘几何外形特征对驻点热流值的影响集中在小 R_0 范围内。对于 R_0 较大的双曲率前缘外形, 继续增大 R_0 对于降低驻点区气动加热作用极小。

在机器学习的过程中, 训练集包含的是几何外形量 R_0=1～10 的气动热数据, 对应的 $1/R_0 \in [0.1,1]$。当 $1/R_0$ 趋于 0 和 1 时, 双曲率前缘分别趋向于圆柱



(a) r=3 mm, R=15 mm



(b) r=4 mm, R=20 mm

2% 偏差带 式(7)曲线 数值模拟结果

图 7 2% 偏差带下公式与数值计算结果的对比

Fig. 7 Comparison between formula under 2% deviation band and numerical results

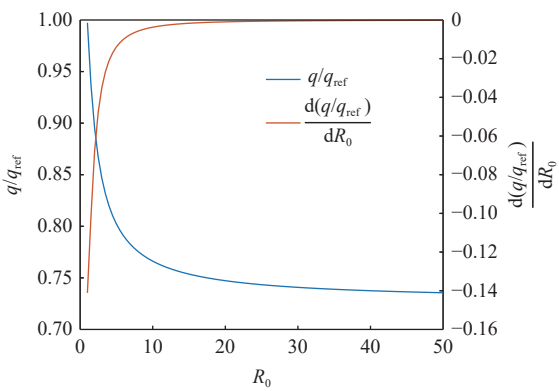


图 8 驻点热流值及其导数值随 R_0 的变化

Fig. 8 The variations of stagnation heat flux and its derivative with the values of R_0

和球面，此时可利用 Fay-Riddell 理论公式[11] 求解驻点热流值。图 9 为 R_0 外推至整个范围内的公式预测值在 2 个极限值附近与 Fay-Riddell 公式[11] 的符合情况，并给出了与文献 [45] 中的经验公式对比，此经验公式是由二维驻点区的热传导系数进行修正获得了双曲率构型的热传导系数，从而获得的驻点热流值。由图 9 可知，与已有的经验公式[45] 相比，本文给出的机器学习公式在 1/ R_0 趋于 0 时与 Fay-

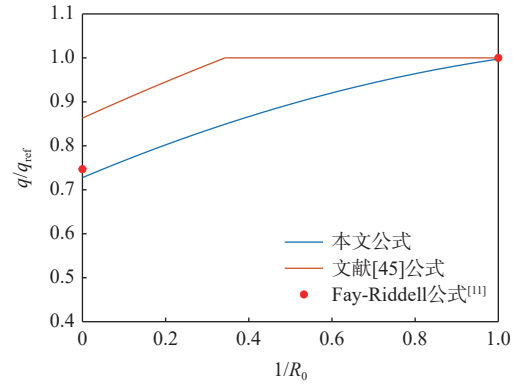


图 9 热流预测公式对比

Fig. 9 Comparisons of formulas of heat flux

Riddell 公式的预测值更为吻合，可见此公式对几何外形量 R_0 具有良好的泛化能力。在 1/ R_0 介于 0 与 1 之间时，即介于球头与圆柱之间任意几何外形的双曲率构型，均可由此公式实现对驻点区热流的精准快速预测。

3 结 论

1）本文提出的高超声速飞行器双曲率前缘热流分布公式能够实现驻点区热流分布的快速准确估计，公式的拟合优度超过 0.99，可实现较好的回归拟合效果。

2）针对参数的敏感性分析发现，公式中 Ma 项对预测热流值的影响较小，以此为基础对公式进一步简化，得到了仅与双曲率前缘几何外形有关的热流预测公式。简化后的公式预测值与训练集数据偏差值均保持在 2% 以下，具有较高的预测精度。

3）在保持前缘外形几何相似时，验证了本文给出的机器学习公式对不同尺度的前缘驻点区热流分布的预测能力。新增数值算例的结果全部落入热流分布公式 2% 的偏差带内，表明本文公式在不同几何尺度的适用性。

4）将简化后的双曲率外形前缘驻点区热流预测公式外推至球体和柱体的极限情况，预测值与 Fay-Riddell 公式给出的理论值非常接近，且趋势上优于大部分的经验公式，这表明本文公式具有良好的泛化能力，不同几何外形下的双曲率前缘构型均可适用，从而弥补了介于球头与圆柱之间双曲率构型驻点区热流计算公式的空缺。

参考文献（References）

[1] LIU Q L, BACCARELLA D, LEE T H. Review of combustion stabilization for hypersonic airbreathing propulsion[J]. Progress in Aerospace Sciences, 2020, 119: 100636.

[2] FLAHERTY K W, ANDREWS K M, LISTON G W. Operability benefits of airbreathing hypersonic propulsion for flexible access to

space[J]. Journal of Spacecraft and Rockets, 2010, 47(2): 280-287.

[3] 梁伟, 金华, 孟松鹤, 等. 高超声速飞行器新型热防护机制研究进展[J]. 宇航学报, 2021, 42(4): 409-424.
LIANG W, JIN H, MENG S H, et al. Research progress on new thermal protection mechanism of hypersonic vehicles[J]. Journal of Astronautics, 2021, 42(4): 409-424(in Chinese).

[4] ANDERSON J D. Hypersonic and high-temperature gas dynamics [M]. 2nd ed. Reston: A2AA, 2006.

[5] 卞荫贵, 钟家康. 高温边界层传热[M]. 北京: 科学出版社, 1986.
BIAN Y G, ZHONG J K. High temperature boundary layer heat transfer[M]. Beijing: Science Press, 1986(in Chinese).

[6] 高莹莹, 杨凯威, 孔维萱, 等. 高超声速飞行器尖化前缘气动热环境研究[J]. 导弹与航天运载技术, 2019(6): 112-116.
GAO Y Y, YANG K W, KONG W X, et al. Research on hypersonic sharp leading edge heating environment[J]. Missiles and Space Vehicles, 2019(6): 112-116(in Chinese).

[7] 孙健, 刘伟强. 尖化前缘高导热材料防热分析[J]. 航空学报, 2011, 32(9): 1622-1628.
SUN J, LIU W Q. Analysis of sharp leading-edge thermal protection of high thermal conductivity materials[J]. Acta Aeronautica et Astronautica Sinica, 2011, 32(9): 1622-1628(in Chinese).

[8] 喻成璋, 刘卫华. 高超声速飞行器气动热预测技术研究进展[J]. 航空科学技术, 2021, 32(2): 14-21.
YU C Z, LIU W H. Research status of aeroheating prediction technology for hypersionic aircraft[J]. Aeronautical Science & Technology, 2021, 32(2): 14-21(in Chinese).

[9] 彭治雨, 石义雷, 龚红明, 等. 高超声速气动热预测技术及发展趋势[J]. 航空学报, 2015, 36(1): 325-345.
PENG Z Y, SHI Y L, GONG H M, et al. Hypersonic aeroheating prediction technique and its trend of development[J]. Acta Aeronautica et Astronautica Sinica, 2015, 36(1): 325-345(in Chinese).

[10] KARIMI M S, OBOODI M J. Investigation and recent developments in aerodynamic heating and drag reduction for hypersonic flows[J]. Heat and Mass Transfer, 2019, 55(2): 547-569.

[11] FAY J, RIDDELL F R. Theory of stagnation point heat transfer in dissociated air[J]. Journal of the Aerospace Sciences, 1958, 25: 73-85.

[12] KEMP N H, RIDDELL F R. Heat transfer to satellite vehicles reenters the atmosphere[J]. Jet Propulsion, 1957, 27(2): 132-137.

[13] VAN DRIEST E R. The problem of aerodynamic heating[J]. Aeronautical Engineering Review, 1956, 15(10): 26-41.

[14] KITAMURA K, ROE P, ISMAIL F. Evaluation of Euler fluxes for hypersonic flow computations[J]. AIAA Journal, 2009, 47(1): 44-53.

[15] FENG Q, CHEN J J, DI S, et al. A grid strategy for predicting the space plane's hypersonic aerodynamic heating loads[J]. Aerospace Science and Technology, 2019, 86: 659-670.

[16] 阎超, 禹建军, 李君哲. 热流 CFD 计算中格式和网格效应若干问题研究[J]. 空气动力学学报, 2006, 24(1): 125-130.
YAN C, YU J J, LI J Z. Scheme effect and grid dependency in CFD computations of heat transfer[J]. Acta Aerodynamica Sinica, 2006, 24(1): 125-130(in Chinese).

[17] 潘沙, 冯定华, 丁国昊, 等. 气动热数值模拟中的网格相关性及收敛[J]. 航空学报, 2010, 31(3): 493-499.
PAN S, FENG D H, DING G H, et al. Grid dependency and convergence of hypersonic aerothermal simulation[J]. Acta Aeronautica et Astronautica Sinica, 2010, 31(3): 493-499(in Chinese).

[18] HOFFMANN K, SIDDIQUI M, CHIANG S. Difficulties associated with the heat flux computations of high speed flows by the Navier-Stokes equations[C]//Proceedings of the 29th Aerospace Sciences Meeting. Reston: AIAA, 1991: 467.

[19] LEES L. Laminar heat transfer over blunt-nosed bodies at hypersonic flight speeds[J]. Jet Propulsion, 1956, 26(4): 259-269.

[20] SCALA S M. A study of hypersonic ablation[C]//Proceedings of the 10th International Congress. Berlin: Springer, 1959: 790.

[21] ZHAO R, WEN C Y, TIAN X D, et al. Numerical simulation of local wall heating and cooling effect on the stability of a hypersonic boundary layer[J]. International Journal of Heat and Mass Transfer, 2018, 121: 986-998.

[22] KLOPFER G, YEE H. Viscous hypersonic shock-on-shock interaction on blunt cowl lips[C]//Proceedings of the 26th Aerospace Sciences Meeting. Reston: AIAA, 1988: 233.

[23] LEE J H J, RHO O. Accuracy of AUSM+ scheme in hypersonic blunt body flow calculations[C]//Proceedings of the 8th AIAA International Space Planes and Hypersonic Systems and Technologies Conference. Reston: AIAA, 1998: 1538.

[24] 马崇立, 刘景源. 网格对高超声速钝头体表面热流数值模拟结果的影响[J]. 航空学报, 2023, 44(5): 68-81.
MA C L, LIU J Y. Effect of grid strategy on numerical simulation results of aerothermal heating loads over hypersonic blunt bodies[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(5): 68-81(in Chinese).

[25] 张伟伟, 寇家庆, 刘溢浪. 智能赋能流体力学展望[J]. 航空学报, 2021, 42(4): 524689.
ZHANG W W, KOU J Q, LIU Y L. Prospect of artificial intelligence empowered fluid mechanics[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(4): 524689(in Chinese).

[26] 陈正. 高超声速气动热实验数据的多层学习方法[D]. 北京: 中国科学院大学, 2021.
CHEN Z. Multi-layer learning method for hypersonic aerothermal experimental data[D]. Beijing: University of Chinese Academy of Sciences, 2021(in Chinese).

[27] CHEN X, CAO Z W, FU B, et al. Adding-point strategy for surrogate-based reduced-order hypersonic aerothermodynamic modeling based on fuzzy clustering[J]. Journal of Spacecraft and Rockets, 2021, 58(1): 244-253.

[28] 聂春生, 黄建栋, 王迅, 等. 基于 POD 方法的复杂外形飞行器热环境快速预测方法[J]. 空气动力学学报, 2017, 35(6): 760-765.
NIE C S, HUANG J D, WANG X, et al. Fast aeroheating prediction method for complex shape vehicles based on proper orthogonal decomposition[J]. Acta Aerodynamica Sinica, 2017, 35(6): 760-765 (in Chinese).

[29] 张智超, 高太元, 张磊, 等. 基于径向基神经网络的气动热预测代理模型[J]. 航空学报, 2021, 42(4): 524167.
ZHANG Z C, GAO T Y, ZHANG L, et al. Aeroheating agent model based on radial basis function neural network[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(4): 524167(in Chinese).

[30] CHEN C, LUO C T, JIANG Z L. A multilevel block building algorithm for fast modeling generalized separable systems[J]. Expert Systems with Applications, 2018, 109: 25-34.

[31] PENG J, LUO C T, HAN Z J, et al. Parameter-correlation study on shock-shock interaction using a machine learning method[J]. Aerospace Science and Technology, 2020, 107: 106247.

[32] 李帅, 彭俊, 罗长童, 等. 基于多层分块算法的激波干扰流场预测[J]. 力学学报, 2021, 53(12): 3284-3297.
LI S, PENG J, LUO C T, et al. Prediction of shock interference flow field structure based on the multi-level block building algorithm[J]. Chinese Journal of Theoretical and Applied Mechanics, 2021,

53(12): 3284-3297(in Chinese).

[33] 彭俊. 强激波相互作用及其极端热载荷诱发机制研究[D]. 北京: 中国科学院大学, 2021.
PENG J. Study on strong shock wave interaction and its induced mechanism of extreme thermal load[D]. Beijing: University of Chinese Academy of Sciences, 2021(in Chinese).

[34] MCBRIDE B J. NASA Glenn coefficients for calculating thermodynamic properties of individual species: NASA/TP-2002-211556[R]. Washington, D. C. : NASA, 2002.

[35] 傅林, 高正红, 左英桃. 基于 HLL-HLLC 的高阶 WENO 格式及其应用研究[J]. 计算力学学报, 2014, 31(1): 128-134.
FU L, GAO Z H, ZUO Y T. High order WENO scheme based on HLL-HLLC solver and its application[J]. Chinese Journal of Computational Mechanics, 2014, 31(1): 128-134(in Chinese).

[36] KOZA J R. Genetic programming: On the programming of computers by means of natural selection[M]. 5th ed. Cambridge: MIT Press, 1992.

[37] CHEN C, LUO C T, JIANG Z L. Block building programming for symbolic regression[J]. Neurocomputing, 2018, 275: 1973-1980.

[38] 陈辰. 快速建模算法及其在流体力学中的应用[D]. 北京: 中国科学院大学, 2017.
CHEN C. Fast modeling algorithm and its application in fluid mechanics[D]. Beijing: University of Chinese Academy of Sciences, 2017(in Chinese).

[39] CLEARY J W. Effects of angle of attack and blunt-ness on laminar heating-rate distributions of a 15° cone at a Mach number of 10.6: NASA TN D-5450[R]. Washington, D. C. : NASA, 1969.

[40] CHEN X Q, HOU Z X, LIU J X, et al. Bluntness impact on performance of waverider[J]. Computers and Fluids, 2011, 48(1): 30-43.

[41] LIU J X, HOU Z X, DING G H, et al. Numerical and experimental study on waverider with blunt leading edge[J]. Computers & Fluids, 2013, 84(19): 203-217.

[42] LIU J X, HOU Z X, CHEN X Q, et al. Experimental and numerical study on the aero-heating characteristics of blunted waverider[J]. Applied Thermal Engineering, 2013, 51(1-2): 301-314.

[43] LI S B, WANG Z G, HUANG W, et al. Aerodynamic performance investigation on waverider with variable blunt radius in hypersonic lows[J]. Acta Astronautica, 2017, 137: 362-372.

[44] 叶博嘉, 鲍序, 刘博, 等. 基于机器学习的航空器进近飞行时间预测[J]. 航空学报, 2020, 41(10): 324136.
YE B J, BAO X, LIU B, et al. Machine learning for aircraft approach time prediction[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(10): 324136(in Chinese).

[45] MAHULIKAR S P. Theoretical aerothermal concepts for configuration design of hypersonic vehicles[J]. Aerospace Science and Technology, 2005, 9(8): 681-685.

Prediction method of aero-heating of hypersonic vehicle bi-curvature leading edge based on machine learning

YANG Fan[1, 2], LIN Mingyue[1, 2], HU Zongmin[1, 2, *], LUO Changtong[1, 2]

(1. State Key Laboratory of High-Temperature Gas Dynamics, Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190, China;
2. School of Engineering Science, University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: The prediction technology of hypersonic aero-heating is one of the key technologies for the development of high-speed vehicles. Creating an efficient method for predicting the hypersonic thermal conditions is highly important for designing thermal protection systems and optimizing aerodynamics. In order to obtain the heat flux distribution on the surface of hypersonic vehicles quickly and to shorten the vehicle design cycle, a fast prediction method for the aerothermal environment of the bi-curvature leading edge of hypersonic vehicles is proposed based on the multi-level block building (MBB) algorithm. The MBB algorithm is distinguished by its generalized separability, which enables it to efficiently represent highly nonlinear data. First, numerical simulations are conducted to obtain the database composed of the aero-heating data of the bi-curvature leading edges of the vehicles in the training set. Based on the MBB algorithm, an explicit expression for predicting the distributions of heat flux is given. The statistical analysis results demonstrate that the discrepancy between the estimated value and the observed value is below 2%, suggesting that the formula given in this study exhibits a high level of predictive precision. Further, an extrapolation of the formula is performed to verify its applicability for different geometric shapes. At the stage of thermal design and aerodynamic optimization of the bi-curvature leading edge configuration, the formulation proposed in this paper enables accurate and rapid prediction of the aerodynamic thermal environment.

Keywords: hypersonic flow; aero-heating; multi-level block building algorithm; fast prediction method; heat flux of stagnation point; bi-curvature leading edge

Received: 2022-08-30; **Accepted**: 2022-12-05; **Published Online**: 2023-02-14 14:11

URL: link.cnki.net/urlid/11.2625.V.20230214.0911.001

Foundation items: National Natural Science Foundation of China (12172365,12072353,12132017); National Key Research and Development Program of China (2019YFA0405204)

* **Corresponding author**. E-mail: huzm@imech.ac.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0701

复合坐标系跨临界液滴表面 Kelvin-Helmholtz 不稳定增长高效预测方法

崔海港[1]，闫啸宇[2]，高岩飞[3]，陈飞[1]，杨世春[1,*]

(1. 北京航空航天大学 交通科学与工程学院，北京 100191；2. 北京航空航天大学 航空发动机研究院，北京 100191；
3. 山东交通学院 汽车工程学院，济南 250399)

摘 要：针对混合动力专用内燃机实际运行的跨临界条件，传统方法对于主导可燃混合气形成质量的液滴表面 Kelvin-Helmholtz 不稳定性分析方法难以实现跨临界宽范围的有效预测。采用复合坐标系方法，建立跨临界强对流环境的等效液滴不稳定性高效预测模型，分别采用全局扩散动力学模型、热力学控制模型及局部液滴表面切向不稳定模型。在全局坐标系模型可以通过模态坐标变换求得解析解，局部坐标系引入势函数等，提高求解速度。而对于可实用性流体计算仿真模型，可通过无量纲物理参数描述不同跨临界条件下空气动力、惯性力、黏性力、表面张力、加速度等控制因素的变化规律，以及其对液滴表面切向和法向不稳定波增长的影响规律。结果表明：液滴表面的空气动力仍然支配着跨临界条件液滴表面切向不稳定波的发展；压力增大时，液相雷诺数控制的流动特性减小与欧尼索控制的液滴黏性力减弱基本抵消；环境温度增加时，雷诺数表征的流动特性对 Kelvin-Helmholtz 波增长的贡献越来越小，液滴的惯性力影响也越来越弱。

关 键 词：复坐标系；跨临界；液滴；表面 Kelvin-Helmholtz 不稳定；高效预测
中图分类号：U469
文献标志码：A **文章编号**：1001-5965(2024)09-2835-08

为实现绿色低碳发展、交通领域碳排放尽早达峰，发动机的应用场景越来越复杂多变，导致汽车能耗与排放极度依赖于非平衡工作过程[1]。特别是高效内燃机及混合动力专用内燃机等，为提高燃烧质量，实际燃烧室环境压力、温度等条件不断被改善，在复杂行驶工况下，这些条件通常处于从远离亚临界至超临界的宽范围的跨临界状态。而在跨临界区域，液滴表面不稳定主导的蒸发及破碎决定着可燃混合气形成过程及浓度场分布，且直接影响发动机燃烧质量与排放水平[2-4]。因此，建立描述跨临界液滴表面不稳定波发展的高效快速计算模型，具有极其重要的意义，既可以实现计算流体力学(computational fluid dynamics, CFD)仿真环境大液滴群跨临界液滴表面切向不稳定的行为表征，也可以统计大尺度下优化控制复杂工况时发动机的运行过程。

当前，仿真环境普遍采用的喷雾液滴表面皿不稳定波破碎理论，仅适用于中低压环境计算，未考虑跨临界高传热传质影响、气相溶解、气液界面热力学、热物理性质的非理想性等[5-9]，且均采用数据拟合方式计算蒸发焓等。而精细大涡模拟则需要巨大的计算资源，因而难以实现工程上的应用[10]。许多研究者对二次雾化中导致液滴破裂的物理过程进行了大量的试验和数值模拟研究，得出许多重

收稿日期：2022-08-10；录用日期：2022-11-12；网络出版时间：2023-05-05 15:59
网络出版地址：link.cnki.net/urlid/11.2625.V.20230505.1354.005
基金项目：国家自然科学基金 (T2241003)
* 通信作者. E-mail：yangshichun@buaa.edu.cn

引用格式：崔海港，闫啸宇，高岩飞，等. 复合坐标系跨临界液滴表面 Kelvin-Helmholtz 不稳定增长高效预测方法 [J]. 北京航空航天大学学报，2024，50(9)：2835-2842. CUI H G，YAN X Y，GAO Y F，et al. Efficient prediction method for Kelvin-Helmholtz instability growth on transcritical droplet surface in composite coordinate system[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2835-2842 (in Chinese).

要结论：袋状破碎、多模式破碎和液膜变薄破碎模式。Dai 和 Faeth[11] 研究表明，当 Weber 数达到剪切破碎区的边缘时，液滴核心的体积分数占主导地位，对于多模式破碎区的总破碎时间，在 Weber 数为 40 时，破碎时间的当地最大值发生在袋/羽和羽/剪切破碎模式的过渡阶段。Reitz[12] 提出开尔文-亥姆霍兹（Kelvin-Helmholtz, K-H）模型，针对加压蒸发与非蒸发性条件，采用散射光的可视化与片光测量、相位多普勒测速仪（phase Doppler anemometry, PDA）、激光多普勒测速（laser Doppler velocimetry, LDV）方法对该联合模型进行了验证。跨临界密度和比热容等热物理性质将不再是常数，而是随着流体状态的变化而变化[13-14]。超临界流体同时具有类似液体和气体的热物理性质，如具有类似液滴的高密度，还具有类似气体的黏度，而不再是单纯的液体或气体。流体的密度会变得对温度的扰动或微小变化非常敏感，密度梯度也会变得非常陡峭[15-19]。

由于临界点附近流体热力学性质的显著变化，所以很难利用 CFD 来模拟超临界流体的行为。此外，由于温度和压力的扰动或微小变化对在临界点附近的流体的结构有很大影响，因此喷雾流场当地局部的物性值也非常重要。

本文建立适用于跨临界强对流环境的等效液滴蒸发与破碎的高效模型，采用复合坐标系方法，分别建立全局扩散动力学模型、热力学控制模型及局部液滴表面切向不稳定模型。在全局坐标系下，模型可以通过模态坐标变换求得解析解，也可以通过局部坐标系引入势函数等，以提高求解速度。对于实现可实用性流体计算仿真模型，可通过无量纲物理参数描述不同跨临界对流边界（环境压力、温度、气液相对速度等）条件下空气动力、惯性力、黏性力、表面张力、加速度等控制因素的变化规律，以及其对液滴表面切向和法向不稳定波增长的影响规律。这对提高复杂多变工况下汽车高性能发动机可燃混合气燃烧质量、降低尾气排放具有重要意义，也为交通系统微观尺度的碳排放控制策略及优化设计提供了理论基础。

1 跨临界 Kelvin-Helmholtz 不稳定模型

当气液存在相对速度时，在 2 种不同流体沿速度切向的界面处，气液相之间相互作用将导致 K-H 不稳定表面波的出现。在 Reitz 模型[12] 中，热物性（如密度、黏度、表面张力等）均采用常压下的恒定值。国内外学者围绕 K-H 不稳定性预测开展了大量研究，如晶格-玻尔兹曼方法 (lattice-Boltzmann method, LBM) [20-21] 通过耦合分子间的相互作用正确模拟界面的运动；相场法[22-23] 可以模拟多组分流体的 K-H 模型的不稳定性，能够自动处理多相物理界面拓扑结构的改变，计算的时间步长也可以更大，但计算精度有待提高；界面追踪法 (front tracking method, FTM)[24-25] 能够以高精度高分辨率捕捉界面结构拓扑变化，可以观测到 K-H 不稳定性的界面形态变化过程。本文在此基础上，采用复合坐标系，即全局主坐标系和局部子坐标系，将液滴分解为稳定尺度和不稳定尺度，如图 1 所示。在稳定大尺度的主坐标系，取随液滴一起运动的柱坐标系统，建立气液相质量、动量守恒方程，满足全局能量、动量和组分守恒；而在不稳定微尺度子坐标系（即表面坐标系），描述液滴表面 K-H 切向不稳定波的增长。对热物性进行跨临界热力学和物理学参数修正，即热物性参数均视为压力、温度和组分的多元函数，使其扩展到跨临界环境中。通过液体稳定性分析得出液滴表面在沿速度方向上的扰动增长机制。

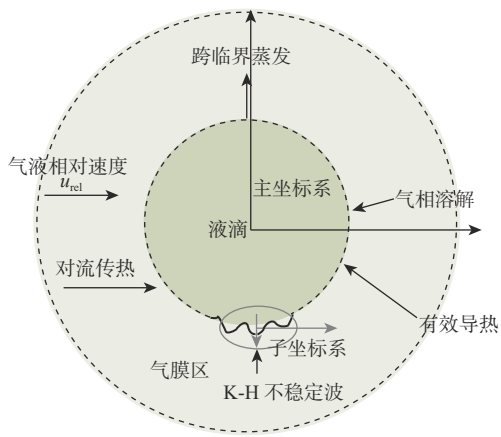


图 1 跨临界复合坐标系统液滴切向不稳定模型示意

Fig. 1 Tangential instability model of droplet in transcritical composite coordinate system

1.1 全局主坐标系模型

跨临界强对流环境中的液滴，在表面发生不稳定波的同时，气液相处于高传热传质过程中，模型示意如图 1 所示。在主坐标系，液滴在全局时空范围的热物理动态特性由界面热力学平衡和相内扩散动力学机制控制。

1.1.1 相内扩散动力学模型

1）液相质量守恒方程。

$$\frac{\partial \rho}{\partial t}+\frac{1}{r^2}\cdot\frac{\partial}{\partial r}\left(r^2\rho u\right)=0 \tag{1}$$

式中：r 为球坐标半径；u 为径向速度；t 为时间；ρ 为密度。

2）气液相组分守恒方程。假定气相在液滴中的溶解仅限于表面层，则气相组分守恒方程为

$$W_{\mathrm{F}}=-\frac{\mathrm{d}m}{\mathrm{d}t} \tag{2}$$

$$W_{\mathrm{F}}=Y_{\mathrm{F,sf}}(W_{\mathrm{F}}+W_{\mathrm{N}})-4\pi r_{\mathrm{d}}^{2}K_{\mathrm{FN}}^{*}\left(Y_{\mathrm{F,sf}}-Y_{\mathrm{F},\infty}\right) \tag{3}$$

式中: W_{F} 和 W_{N} 分别为燃料组分和氮气组分的质量通量; m 为液滴质量; Y 为气相组分质量分数; 下标 sf 表示液滴表面; 下标∞表示远方环境; 下标 d 表示液滴; 下标 F、N 分别表示燃料组分、氮气组分; K_{FN}^{*}为修正的传质速率。

3) 气液相能量守恒方程。对于对流环境中的液滴, 忽略辐射传热, 采用有效导热的方法处理对流传热:

$$4\pi r_{\mathrm{d}}^{2}k^{\mathrm{e}}\left.\frac{\partial T}{\partial r}\right|_{r=r_{\mathrm{d}}}=\left(\overline{H}_{\mathrm{F}}^{\mathrm{g}}-\overline{H}_{\mathrm{F}}^{\mathrm{l}}\right)\frac{\mathrm{d}m}{\mathrm{d}t}+4\pi r_{\mathrm{d}}^{2}\left[h_{\mathrm{g}}^{*}(T_{\infty}-T_{\mathrm{sf}})\right] \tag{4}$$

式中: k^{e}为有效导热系数; T 为温度; $\overline{H}_{\mathrm{F}}$为偏摩尔焓; 上标 l 和 g 分别表示液相态和气相态; h_{g}^{*}为修正气相传热系数。

4) 液相内部能量方程。

$$C_{p}\frac{\partial T}{\partial t}-\frac{1}{\rho r^{2}}\cdot\frac{\partial}{\partial r}\left(r^{2}k\frac{\partial T}{\partial r}\right)=0 \tag{5}$$

式中: C_p为定压比热容; k 为导热系数。

5) 动量守恒方程。在计算中忽略压力的变化, 假设气相域和液相域的压力均匀。动量守恒方程为

$$-p_{\mathrm{l}}+2\mu_{\mathrm{l}}\frac{\partial u_{\mathrm{l}}}{\partial r}+p_{\mathrm{g}}-2\mu_{\mathrm{g}}\frac{\partial u_{\mathrm{g}}}{\partial r}+\frac{2\sigma_{\mathrm{mix}}}{r_{\mathrm{d}}}=0 \tag{6}$$

式中: p 为压力; μ为黏度; σ为表面张力; 下标 mix 表示混合物。

1.1.2　界面热力学平衡控制模型

液滴气液两相界面由混合物热力学平衡控制, 燃料和气体组分均满足化学热力学吉布斯自由能 G 守恒, 即

$$\begin{cases}G_{\mathrm{F}}^{\mathrm{l}}=\left[G_{\mathrm{F},0}^{\mathrm{l}}+RT\ln\left(X_{\mathrm{F}}^{\mathrm{l}}\right)\right]X_{\mathrm{F}}^{\mathrm{l}}=G_{\mathrm{F}}^{\mathrm{g}}=\left[G_{\mathrm{F},0}^{\mathrm{g}}+RT\ln\left(X_{\mathrm{F}}^{\mathrm{g}}\right)\right]X_{\mathrm{F}}^{\mathrm{g}}\\ G_{\mathrm{N}}^{\mathrm{l}}=\left[G_{\mathrm{N},0}^{\mathrm{l}}+RT\ln\left(X_{\mathrm{N}}^{\mathrm{l}}\right)\right]X_{\mathrm{N}}^{\mathrm{l}}=G_{\mathrm{N}}^{\mathrm{g}}=\left[G_{\mathrm{N},0}^{\mathrm{g}}+RT\ln\left(X_{\mathrm{N}}^{\mathrm{g}}\right)\right]X_{\mathrm{N}}^{\mathrm{g}}\end{cases} \tag{7}$$

式中: 下标 0 表示标准态; R 为气体常数; X 为摩尔分数。

燃料和气体组分的化学势 e_{F}、e_{N} 分别为

$$\begin{cases}e_{\mathrm{F}}^{\mathrm{l}}=G_{\mathrm{F},0}^{\mathrm{l}}+RT\ln\left(X_{\mathrm{F}}^{\mathrm{l}}\right)\\ e_{\mathrm{F}}^{\mathrm{g}}=G_{\mathrm{F},0}^{\mathrm{g}}+RT\ln\left(X_{\mathrm{F}}^{\mathrm{g}}\right)\\ e_{\mathrm{N}}^{\mathrm{l}}=G_{\mathrm{N},0}^{\mathrm{l}}+RT\ln\left(X_{\mathrm{N}}^{\mathrm{l}}\right)\\ e_{\mathrm{N}}^{\mathrm{g}}=G_{\mathrm{N},0}^{\mathrm{g}}+RT\ln\left(X_{\mathrm{N}}^{\mathrm{g}}\right)\end{cases} \tag{8}$$

在临界点处, 各组分化学势(吉布斯自由能对成分的偏微分)在相界面两侧相等, 液滴表面消失, 蒸发焓减为 0, 气液相没有进行区分。

1.2　局部子坐标系模型

假设流体是不可压缩的, 并忽略气相黏度系数, 则液相以速度 u_{rel} 进入静态环境中。在液滴表面采用局部子坐标系, 如图 2 所示。

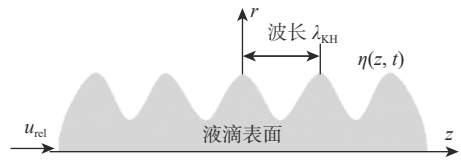


图 2　局部表面坐标系 K-H 不稳定表面波示意

Fig. 2　K-H unstable surface waves in local surface coordinate system

任意时刻界面函数为 $r=\eta(z,t)$, 无扰动时 $r=0$。其中, z 为轴向坐标。

气、液两相各自的控制方程如下。

连续性方程:

$$\frac{\partial\tilde{u}}{\partial z}+\frac{1}{r}\cdot\frac{\partial(r\tilde{v})}{\partial r}=0 \tag{9}$$

z 方向动量方程:

$$\frac{\partial\tilde{u}}{\partial t}+U_{i}\frac{\partial\tilde{u}}{\partial z}+\tilde{v}\frac{\mathrm{d}\tilde{u}}{\mathrm{d}r}=-\frac{1}{\rho}\cdot\frac{\partial\tilde{p}}{\partial z}+\frac{\mu}{\rho}\left[\frac{\partial^{2}\tilde{u}}{\partial z^{2}}+\frac{1}{r}\cdot\frac{\partial}{\partial r}\left(r\frac{\partial\tilde{u}}{\partial r}\right)\right] \tag{10}$$

r 方向动量方程:

$$\frac{\partial\tilde{v}}{\partial t}+U_{i}\frac{\partial\tilde{v}}{\partial z}=-\frac{1}{\rho}\cdot\frac{\partial\tilde{p}}{\partial r}+\frac{\mu}{\rho}\left[\frac{\partial^{2}\tilde{v}}{\partial z^{2}}+\frac{1}{r}\cdot\frac{\partial}{\partial r}\left(r\frac{\partial\tilde{v}}{\partial r}\right)\right] \tag{11}$$

式中: $\tilde{u}$为气液相相对于主体运动的 z 方向脉动速度; $\tilde{v}$为气液相相对于主体运动的 r 方向脉动速度; $\tilde{p}$为气液相压力脉动; U_i=U_{l}、U_{g} 分别为液相、气相主流速度。假设气相主流速度 $U_{\mathrm{g}}=0$, 则液相主流速度 $U_{\mathrm{l}}=u_{\mathrm{rel}}$, u_{rel} 为气液相对速度。

界面的运动边界条件:

$$\tilde{v}=\frac{\partial\eta}{\partial t}+U_{i}\frac{\partial\eta}{\partial z} \tag{12}$$

切向应力边界条件:

$$\frac{\partial\tilde{u}_{\mathrm{l}}}{\partial r}+\frac{\partial\tilde{v}_{\mathrm{l}}}{\partial z}=0 \tag{13}$$

法向应力边界条件:

$$-\tilde{p}_{\mathrm{l}}+2\mu_{\mathrm{l}}\frac{\partial\tilde{u}_{\mathrm{l}}}{\partial z}+\tilde{p}_{\mathrm{g}}-2\mu_{\mathrm{g}}\frac{\partial\tilde{v}_{\mathrm{g}}}{\partial r}-\sigma_{\mathrm{mix}}\frac{\partial^{2}\eta}{\partial z^{2}}=0 \tag{14}$$

在高温、高压下, 由于气相在液相中的溶解, 因此可将液滴表面视为混合物。

1.3　求解方法

为避免数值求解复杂非线性偏微分方程, 提高模型预测的时效性, 在相对稳定的主坐标系界面, 对密度 ρ、温度 T、速度 u 采用模态转换方法求得解

析解。其中，密度模态坐标转换方法如下：

$$\rho(r,t)\approx\sum_{j=0}^{\infty}\rho_j^*(t)\varphi_j(z) \tag{15}$$

$$\rho_j^*(t)=\int\psi_j(r)\rho(r,t)\mathrm{d}r \qquad j=0,1,2,\cdots \tag{16}$$

式中：$\varphi(\cdot)$和$\psi(\cdot)$为权重函数。

在局部子坐标系，由于界面为极不稳定过程，因此引入速度势函数和流函数求解表面波长和频率：

$$\phi_\mathrm{l}=\Phi_\mathrm{l}(r)\mathrm{e}^{\mathrm{i}\left(\frac{2\pi}{\lambda}\right)z+\omega t} \tag{17}$$

$$\psi_\mathrm{l}=\Psi_\mathrm{l}(r)\mathrm{e}^{\mathrm{i}\left(\frac{2\pi}{\lambda}\right)z+\omega t} \tag{18}$$

$$\psi_\mathrm{g}=\left[U_\mathrm{g}-\mathrm{i}\left(\frac{\omega\lambda}{2\pi}\right)\right]\eta_0\mathrm{e}^{\mathrm{i}\left(\frac{2\pi}{\lambda}\right)z+\omega t}f(r) \tag{19}$$

式中：η_0为初始不稳定波的振幅；λ为波长；ω为频率；$f(r)$满足 Orr-Sommerfeld 方程；$\Phi(\cdot)$和$\Psi(\cdot)$分别为速度势和流函数的幅值。

2 数值结果分析

利用式(1)～式(19)求解正辛烷燃料在跨临界氮气环境中的表面不稳定动态过程，并将计算结果与同条件定容单进行试验验证。为了增强模型对于不同燃料等条件的普适性，本文采用无量纲参数，其中，对比压力为环境压力与燃料临界压力之比($P_\mathrm{r}=P/P_\mathrm{c}$)，对比温度为温度与燃料临界温度之比($T_\mathrm{r}=T/T_\mathrm{c}$)。初始液滴温度为0.53倍临界温度，环境温度为0.95倍临界温度，气体速度u_rel=80 m/s时，液滴迎风面破碎高速摄影原始影像如图3所示。表1为数值预测结果与试验结果，可以看出，二者吻合很好，这验证了本文提出的复坐标系跨临界液滴不稳定过程预测方法的有效性。

图4为不同环境压力P_r下气相混合物韦伯数$We_\mathrm{g}=r_\mathrm{d}\rho_\mathrm{g}u_\mathrm{rel}^2/\sigma_\mathrm{mix}$、液相雷诺数$Re_\mathrm{l}=2r_\mathrm{d}\rho_\mathrm{l}u_\mathrm{rel}/\mu_\mathrm{l}$、气液密度比、欧尼索数$Oh=\sqrt{\rho_\mathrm{l}We_\mathrm{g}/\left(\rho_\mathrm{g}Re_\mathrm{l}^2\right)}$随无量纲时间$\tau$的变化曲线。其中，初始液滴温度$T_\mathrm{rdrop0}$=0.56，环境温度$T_\mathrm{ramb}$=1.0，初始液滴速度$u_\mathrm{rel0}$=40 m/s。

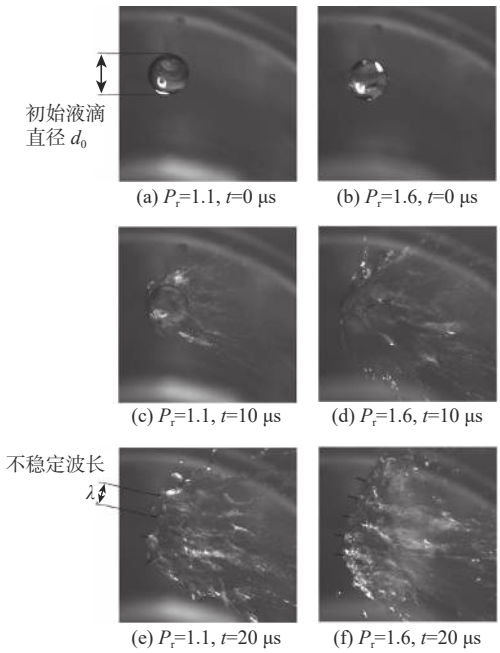


图3 液滴迎风面破碎高速摄影图像

Fig. 3 High-speed photographic image of droplets breaking up on the windward side

与图4相同的条件下，不同环境压力下最大K-H(表面切向)增长率Ω_KH随时间的变化如图5所示。可以发现，随着时间的推移，最大K-H(表面切向)增长率越来越小；环境压力越大，则最大K-H增长率越大，随压力增长的幅度也越大。将图5与图4进行比较，可以发现，在整个蒸发破碎过程中，气相韦伯数对液滴表面切向K-H不稳定波增长具有显著影响，由此说明液滴表面的空气动力是支配液滴表面切向不稳定波发展的决定性因

表1 不稳定波长与初始粒径比的试验与数值模拟结果

Table 1 Experimental and numerical simulation results of ratio of unstable wavelength to initial particle size

对比态环境压力P_r	对比态初始液滴温度T_rdrop0	对比态环境温度T_ramb	气流速度u_rel/(m·s^{-1})	λ/d_0计算值	λ/d_0试验值
1.1	0.53	0.95	80	0.53	0.6
1.6	0.53	0.95	80	0.4	0.4
1.0	0.59	0.95	80	0.62	0.7
1.0	0.78	0.95	80	0.3	0.3
1.0	0.53	0.66	80	0.60	0.7
1.0	0.53	1.15	80	0.56	0.6
1.0	0.53	0.95	80	0.59	0.6
1.0	0.53	0.95	100	0.43	0.5

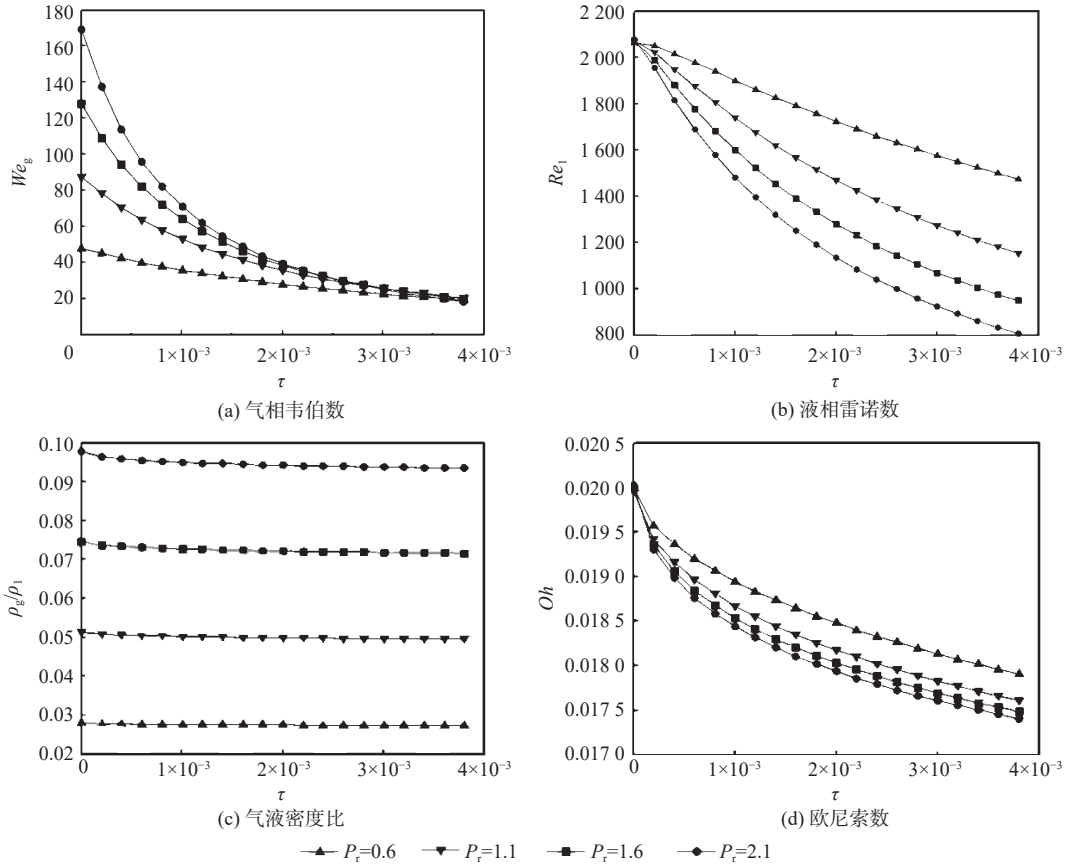


图 4　不同环境压力下各参数随无量纲时间的变化曲线

Fig. 4　Variation of variables with dimensionless time under different ambient pressures

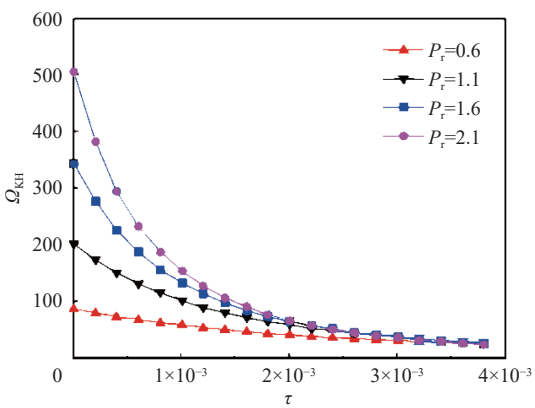


图 5　不同环境压力下最大 K-H 增长率随无量纲时间的变化曲线

Fig. 5　Variation of maximum K-H growth rate with dimensionless time under different environmental pressures

素。在整个蒸发过程中, 最大 K-H 增长率因空气动力作用的减弱而越来越低。但是, 在蒸发初始阶段, 气液密度比对表面切向破碎影响较大, 最大不稳定波增长率也会随气液密度比的增大而增大。还可以发现, 在蒸发后期, 虽然液相雷诺数 Re_l 和欧尼索数 Oh 变化较大, 但是对 K-H 不稳定波影响不大。这说明随着蒸发的进行, 一方面, 环境压力的增大使液相雷诺数减小, 流场流动特性减弱, 液滴表面切向不稳定性减弱; 另一方面, 环境压力的增大使欧尼索数减小, 液滴黏性力减弱, 液滴表面切向不稳定性增强; 而在跨临界蒸发过程中, 压力增大则会导致这 2 种影响基本抵消。

图 6 为不同环境温度下气相韦伯数 We_g、液相雷诺数 Re_l、气液密度比、欧尼索数 Oh 随无量纲时间的变化曲线。其中, 环境压力 P_r=1.0, 初始液滴温度 T_{rdrop0}=0.56, 初始液滴速度 u_{rel0}=40 m/s。与图 6 相同的条件下, 不同环境温度 T_{ramb} 下最大 K-H(表面切向)增长率随时间的变化如图 7 所示。将图 7 和图 6 进行比较可以发现, 气相韦伯数控制着 K-H 波增长率, 即空气动力控制着液滴表面切向方向不稳定波增长。在蒸发过程中, 一方面, 液相雷诺数不断减小, 惯性力随之减小, 气液密度比减小, 气液的相互作用也减小, 抑制了 K-H 波的增长; 另一方面, 欧尼索数在蒸发过程中不断减小, 黏性力减小, 从而抑制了 K-H 波的增长, 也抵消了惯性力和气液密度比的影响。此外, 还可以看出, 环境温度越高, 液相雷诺数变化越平缓, 气液密度比和欧尼索数的变化也越来越快。这说明随着环境温度的增加, 雷诺数对 K-H 波增长的贡献越来越小, 液滴的惯性力影响也越来越弱。

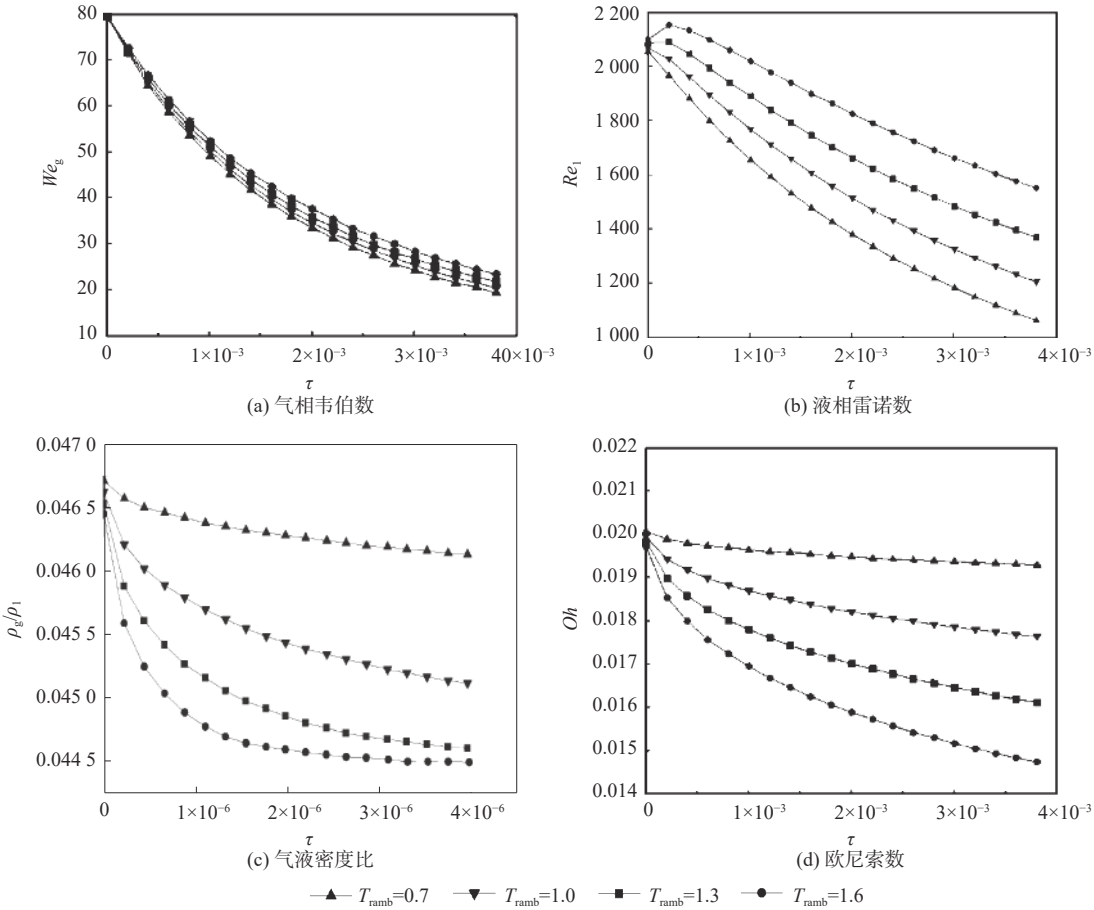


图 6 不同环境温度下各参数随无量纲时间的变化曲线

Fig. 6 Variation of variables with dimensionless time under different ambient temperatures

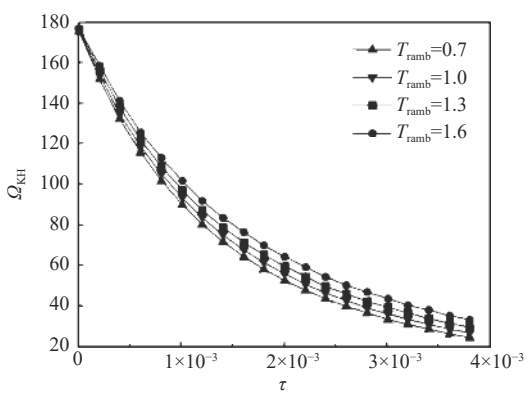


图 7 不同环境温度下最大 K-H 增长率随无量纲时间的变化曲线

Fig. 7 Variation of maximum K-H growth rate with dimensionless time under different ambient temperatures

3 结 论

1）仿真与试验结果对比分析表明，本文方法的不稳定波长预测与试验数据具有较好的一致性，因此，本文方法具有较高的有效性，可用于描述不同跨临界条件下空气动力、惯性力、黏性力、表面张力、加速度等控制因素的变化规律，以及其对液滴表面不稳定波增长的影响规律分析。

2）液滴表面的空气动力仍然支配着跨临界条件液滴表面切向不稳定波的发展。压力增大时，液相雷诺数控制的流动特性的减弱与欧尼索控制的液滴黏性力的减弱基本抵消。

3）在蒸发的初始阶段，气液密度比对表面切向破碎影响较大，最大不稳定波增长率随气液密度比的增大而增大。在蒸发后期，环境压力对流场流动特性和液滴黏性力的影响基本抵消。

4）随着环境温度的增加，液相雷诺数表征的液相惯性力对 K-H 波增长的贡献也越来越小。

参考文献（References）

[1] GAO Y F, YANG J, WANG X B, et al. Fuel consumption and exhaust emissions under varying road condition considering effects of vehicles on other lanes[J]. International Journal of Modern Physics C, 2021, 32(10): 1-22.

[2] 陆大勇, 贾和坤, 叶泽. 高压条件下椭圆孔喷雾破碎雾化分析[J]. 江苏大学学报 (自然科学版), 2021, 42(3): 291-297.
LU D Y, JIA H K, YE Z. Analysis of spray's breakup and atomization of elliptical diesel nozzle under high injection pressure[J]. Journal of Jiangsu University (Natural Science Edition), 2021, 42(3): 291-297(in Chinese).

[3] 王字满, 戴晓宇, 李佳峰, 等. 柴油喷雾初始阶段喷孔内流及油束破碎特性[J]. 汽车安全与节能学报, 2019, 10(2): 241-248.

WANG Z M, DAI X Y, LI J F, et al. Nozzle internal flow and primary breakup characteristics of diesel spray during initial stage[J]. Journal of Automotive Safety and Energy, 2019, 10(2): 241-248(in Chinese).

[4] 仇滔, 王凯欣, 雷艳, 等. 基于动态识别柴油喷雾破碎特征的 ELSA 改进算法[J]. 内燃机学报, 2021, 39(2): 153-158.
QIU T, WANG K X, LEI Y, et al. Dynamic identification of diesel spray breakup model feature based on improved ELSA algorithm[J]. Transactions of CSICE, 2021, 39(2): 153-158(in Chinese).

[5] SULA C, GROSSHANS H, PAPALEXANDRIS M V. Assessment of droplet breakup models for spray flow simulations[J]. Flow, Turbulence and Combustion, 2020, 105(3): 889-914.

[6] ASHKEZARI A Z, DIVSALAR K, MALMIR R, et al. Emission and performance analysis of DI diesel engines fueled by biodiesel blends via CFD simulation of spray combustion and different spray breakup models: A numerical study[J]. Journal of Thermal Analysis and Calorimetry, 2020, 139(4): 2527-2539.

[7] 裴毅强, 王琨, 王同金, 等. GDI 喷油器超高压乙醇喷雾微观特性试验[J]. 天津大学学报 (自然科学与工程技术版), 2018, 51(2): 181-189.
PEI Y Q, WANG K, WANG T J, et al. Experimental on microscopic spray characteristics of GDI injector fueled with ethanol under ultra-high injection pressure[J]. Journal of Tianjin University (Science and Technology), 2018, 51(2): 181-189(in Chinese).

[8] 施红辉, 师顺, 刘晨, 等. 超声速条件下亚毫米液滴的变形破碎模态[J]. 航空动力学报, 2020, 35(10): 2017-2027.
SHI H H, SHI S, LIU C, et al. Deformation and fracture patterns of sub-millimeter droplets under supersonic conditions[J]. Journal of Aerospace Power, 2020, 35(10): 2017-2027(in Chinese).

[9] 雷基林, 李建微, 刘懿, 等. 液滴高速撞击低温壁面的动态特性及破碎机理研究[J]. 实验流体力学, 2022, 36(5): 96-101.
LEI J L, LI J W, LIU Y, et al, Experimental study on spreading and breaking mechanism of droplet impinging on low temperature wall at high speed[J]. Journal of Experiments in Fluid Mechanics, 2022, 36(5): 96-101(in Chinese).

[10] BOURGOIN A, GUILLOU S S, THIÉBOT J, et al. Use of large-eddy simulation for the bed shear stress estimation over a dune[J]. International Journal of Sediment Research, 2021, 36(6): 687-695.

[11] DAI Z, FAETH G M. Temporal properties of secondary drop breakup in the multimode breakup regime[J]. International Journal of Multiphase Flow, 2001, 27(2): 217-236.

[12] REITZ R D. Modeling atomization processes in high-pressure vaporizing sprays[J]. Atomization and Sprays Technology, 1987, 3(309): 309-337.

[13] 刘旻昀, 黄彦平, 唐佳, 等. 超临界流体物性畸变特性的多尺度研究[J]. 原子能科学技术, 2021, 55(11): 1921-1929.
LIU M Y, HUANG Y P, TANG J, et al. Multi-scale study on property distortion mechanism of supercritical fluid[J]. Atomic Energy Science and Technology, 2021, 55(11): 1921-1929(in Chinese).

[14] 熊勃, 张浩, 郭鹏越, 等. 边界条件对椭球颗粒在超临界水中受力与传热特性影响的数值模拟[J]. 材料与冶金学报, 2021, 20(4): 297-303.
XIONG B, ZHANG H, GUO P Y, et al. Numerical investigation on the effect of boundary condition on momentum and heat transfer of spheroids in supercritical water[J]. Journal of Materials and Metallurgy, 2021, 20(4): 297-303(in Chinese).

[15] CHEN H, HU Y, LIU J W, et al. Surface characteristics analysis of fractures induced by supercritical CO_2 and water through three-dimensional scanning and scanning electron micrography[J]. Journal of Rock Mechanics and Geotechnical Engineering, 2021, 13(5): 1047-1058.

[16] 王彦红, 陆英楠, 李洪伟, 等. 竖直螺旋管中超临界 RP-3 航空煤油换热数值研究[J]. 北京航空航天大学学报, 2023, 49(5): 1108-1115.
WANG Y H, LU Y N, LI H W, et al. Numerical study on heat transfer of supercritical RP-3 aviation kerosene in vertical helical tubes[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(5): 1108-1115(in Chinese).

[17] YAO H X, WEI X Z, YE H. Supercritical carbon dioxide as a new working medium for pneumatic launch: A theoretical study[J]. Defence Technology, 2021, 17(4): 1296-1306.

[18] 范世望, 朱郁波, 周璐, 等. 基于开源求解工具的超临界二氧化碳印刷电路板式换热器流动传热特性数值研究[J]. 科学技术与工程, 2021, 21(27): 11587-11594.
FAN S W, ZHU Y B, ZHOU L, et al. Numerical study of the heat transfer in printed circuit heat exchanger with supercritical CO_2 using open-source tool chain[J]. Science Technology and Engineering, 2021, 21(27): 11587-11594(in Chinese).

[19] 刘志军, 孙海礼, 董超, 等. 超临界乙醇法制备超薄碳层包覆磷酸铁锂的研究[J]. 现代化工, 2021, 41(11): 86-90.
LIU Z J, SUN H L, DONG C, et al. Study on synthesis of ultrathin carbon coated lithium iron phosphate via supercritical ethanol method [J]. Modern Chemical Industry, 2021, 41(11): 86-90(in Chinese).

[20] LI X, DONG Z Q, LI Y, et al. A fractional-step lattice Boltzmann method for multiphase flows with complex interfacial behavior and large density contrast[J]. International Journal of Multiphase Flow, 2022, 149: 103982.

[21] AWASTHI M K, ASTHANA R, UDDIN Z. Nonlinear study of Kelvin-Helmholtz instability of cylindrical flow with mass and heat transfer[J]. International Communications in Heat and Mass Transfer, 2016, 71: 216-224.

[22] KIM K S, KIM M H. Simulation of the Kelvin-Helmholtz instability using a multi-liquid moving particle semi-implicit method[J]. Ocean Engineering, 2017, 130: 531-541.

[23] LEE H G, KIM J. Two-dimensional Kelvin-Helmholtz instabilities of multi-component fluids[J]. European Journal of Mechanics-B, 2015, 49: 77-88.

[24] 尚文强, 张莹, 陈昭奇, 等. 基于 FTM 方法的二维 Kelvin-Helmholtz 不稳定性数值模拟[J]. 计算力学学报, 2018, 35(4): 424-430.
SHANG W Q, ZHANG Y, CHEN Z Q, et al. Numerical simulation of two-dimensional Kelvin-Helmholtz instabilities using a front tracking method[J]. Chinese Journal of Computational Mechanics, 2018, 35(4): 424-430(in Chinese).

[25] 过海龙, 张莹, 李文彬, 等. 基于界面追踪法的 Kelvin-Helmholtz 不稳定性直接数值模拟[J]. 内燃机工程, 2021, 42(3): 16-25.
GUO H L, ZHANG Y, LI W B, et al. Direct numerical simulation of Kelvin-Helmholtz instability based on front tracking method[J]. Chinese Internal Combustion Engine Engineering, 2021, 42(3): 16-25 (in Chinese).

Efficient prediction method for Kelvin-Helmholtz instability growth on transcritical droplet surface in composite coordinate system

CUI Haigang[1]，YAN Xiaoyu[2]，GAO Yanfei[3]，CHEN Fei[1]，YANG Shichun[1,*]

(1. School of Transportation Science and Engineering，Beihang University，Beijing 100191，China；
2. Research Institute of Aero-engine，Beijing University，Beijing 100191，China；
3. Scool of Automotive Engineering, Shandong Jiaotong University，Jinan 250399，China)

Abstract：Traditional methods for analyzing the Kelvin-Helmholtz instability on the surface of droplets that dominate the formation mass of combustible mixtures are difficult to achieve effective predictions over a wide range of critical conditions for the actual operation of hybrid internal combustion engines. In this paper, the composite coordinate system method was used to establish an efficient prediction model of equivalent droplet instability in a transcritical strong convection environment. The global diffusion dynamics model, thermodynamics control model, and tangential instability model of local droplet surface were used, respectively. The analytical solution could be obtained by the modal coordinate transformation in the global coordinate system model, and the potential function was introduced in the local coordinate system to improve the solution speed. To realize a practical fluid computational simulation model, the dimensionless physical parameters were used to describe the changes in aerodynamic force, inertial force, viscous force, surface tension, and other control factors under different transcritical conditions and analyze their influence on the growth of tangential and normal unstable waves on the droplet surface. The results show that the aerodynamic forces on the droplet surface still dominate the development of tangential unstable waves on the droplet surface under transcritical conditions. When the pressure increases, the decrease in the flow characteristics controlled by Reynolds number in liquid phase and the decrease in the viscous force of the droplet controlled by Ornisol number basically offset. With the increase in ambient temperature, the contribution of flow characteristics controlled by Reynolds number to the growth of the Kelvin-Helmholtz wave becomes small, and the influence of the inertial force of the droplet becomes weak.

Keywords：composite coordinate system；transcritical；droplet；Kelvin-Helmholtz instability on surface；efficient prediction

Received：2022-08-10；**Accepted**：2022-11-12；**Published Online**：2023-05-05 15：59
URL：link.cnki.net/urlid/11.2625.V.20230505.1354.005
Foundation item：National Natural Science Foundation of China (T2241003)
* **Corresponding author**. E-mail：yangshichun@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0739

扫频雷达双频连续波伪信号干扰效应

赵宏泽，魏光辉*，潘晓东，杜雪，卢新福
(陆军工程大学 电磁环境效应国家重点实验室，石家庄 050003)

摘 要：为掌握典型雷达双频电磁辐射伪信号干扰的效应规律，揭示双频伪信号干扰效应机理，以某型Ku波段扫频测距雷达为试验对象，开展双频连续波伪信号干扰效应试验，分别揭示了二阶互调射频干扰、二阶互调低频干扰和三阶交互调低频干扰效应机理，并确定了伪信号电平随干扰场强的变化规律，解释明了伪信号的形态特征和产生位置。试验结果表明：当干扰频差为540～660 MHz时，带外双频二阶互调信号会造成脉冲型伪信号干扰，伪信号产生的位置随机；当干扰频差小于5 MHz时，双频二阶互调信号和三阶交互调信号均会造成冲激型伪信号干扰，伪信号产生的位置固定。

关 键 词：扫频雷达；双频干扰；电磁辐射；伪信号干扰；二阶互调
中图分类号：O441; TN972; TN95
文献标志码：A **文章编号**：1001-5965(2024)09-2843-09

随着信息技术的高速发展，电磁战概念逐渐成为各国研究的重点[1-3]，未来信息化条件下的一体化联合作战，用频设备被广泛应用，电磁空间日趋拥挤，加之敌方实施的有意电磁干扰，会使战场电磁环境更加恶劣，因此复杂电磁环境下雷达的电磁环境效应成为被广泛研究的问题[4-7]。近年来，国内外诸多学者对典型用频设备(如无人机数据链、导航接收机等)在单频及多频连续波电磁辐射的阻塞效应规律和建模评估方法开展了研究[8-11]，揭示了连续波电磁辐射的损伤规律和阻塞效应机理，为复杂电磁环境下评价电子信息设备的电磁适应能力和提高用频设备的电磁防护水平奠定了坚实的基础。

雷达指利用无线电对目标体进行探测和测距的电磁传感器，相当于人类所说的“千里眼”，其被广泛应用于军用及民用领域[12]。现有对雷达伪信号干扰的研究主要着重于以信息型干扰为背景的有意干扰，而以非信息型干扰为背景的无意干扰的相关研究较少[13-14]。扫频连续波测距雷达作为典型的用频设备，容易受到带内外电磁辐射干扰信号的影响，尤其是带内(指受试雷达工作频段内)连续波信号，因干扰信号形式与雷达有用信号形式相近，在干扰场强很弱的条件下就可在雷达显示界面形成伪信号干扰，因此在干扰场强较强时，伪信号强度甚至超过真实回波信号强度，从而造成雷达误判[15]。文献[16]指出了干扰频率位于雷达工作频带内的单频信号可对雷达造成虚警干扰，但未对虚警干扰的目标特征进行分析。文献[17]从雷达方程出发，分析了雷达噪声干扰和多假目标干扰的产生机理，推导建立了雷达噪声模型，试验验证了机理分析的正确性。文献[18]认为单频干扰可使调频连续波雷达显示界面产生伪信号干扰带。但上述文献鲜有涉及雷达带内外多频伪信号干扰，尤其是多频信号干扰频率在某些特定频率组合下产生的互调信号或与有用信号作用产生的交调信号[19]，上述信号对雷达造成的伪信号强度甚至远远超过单频伪信号强度，但国内外鲜见相关研究。

收稿日期：2022-08-24；**录用日期**：2022-12-16；**网络出版时间**：2023-01-04 19:28
网络出版地址：link.cnki.net/urlid/11.2625.V.20230104.1423.002
基金项目：电磁环境效应国家重点实验室基金 (6142205200301)
* **通信作者**. E-mail：wei-guanghui@sohu.com

引用格式：赵宏泽，魏光辉，潘晓东，等. 扫频雷达双频连续波伪信号干扰效应 [J]. 北京航空航天大学学报，2024，50（9）：2843-2851.
ZHAO H Z，WEI G H，PAN X D，et al. Dual-frequency continuous wave pseudo-signal interference effect in swept-frequency radar [J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50（9）：2843-2851 （in Chinese）.

Ku波段连续波雷达天线物理尺寸较小，具有良好的便携和测距功能，其被广泛应用于近程跟踪及制导[20]。本文以某型Ku波段扫频连续波测距雷达为试验对象，针对雷达带内与带内临近频段开展单频及双频连续波电磁辐射伪信号干扰效应试验，得到受试雷达带内单频伪信号干扰敏感频段，发现带外双频脉冲型伪信号干扰和带内外双频冲激型伪信号干扰现象，研究得到双频伪信号干扰效应规律，揭示了双频伪信号干扰效应与成像机制，为开展雷达复杂电磁环境下的伪信号干扰效应评估及防护奠定了基础。

1 受试雷达单频伪信号干扰敏感特性

选择某型Ku波段扫频测距雷达作为受试对象，雷达工作带宽为200 MHz，跳频时间为0.05 ms，扫频阶梯为10 kHz，最大探测距离可达5 000 m，设置距离约8.25 m的喇叭天线作为雷达探测目标，试验配置如图1所示。因雷达在实际工作过程中需探测的目标距离大多会远于试验中所设置的距离，而近距离的目标回波强度较大，为使试验结果具有普适性，在雷达发射天线的连接端口处增加不同程度的衰减器，用于模拟距离较远处的探测目标回波强度。参照GJB 8848—2016[21]，采用等效的差模注入方法代替辐照法开展测试，单频干扰信号通过监测注入模块直接注入雷达接收端口，频谱分析仪监测单频电磁辐射注入功率值，而后换算为单频电磁辐射在雷达接收天线处的干扰场强值。设备的主要参数如下：单频干扰源型号为罗德&施瓦茨SMR20，可产生1～20 GHz微波信号。频谱分析仪型号为Ceyear公司的4204G，频率为0.009 kHz～44 GHz。目标天线采用BBHA-9120D型双脊宽带喇叭天线，频率为1～18 GHz，增益为6.3～18 dBi。具有监测功能的注射耦合模块是由课题组自行研发的。

带内单频电磁辐射可对雷达造成脉冲型伪信号干扰，选取单次雷达显示受扰界面如图2所示，参考文献[15]选择伪信号干扰电平6 dBmV作为伪信号干扰敏感判据，可得到单频伪信号干扰敏感频段，其临界干扰场强变化曲线如图3所示。

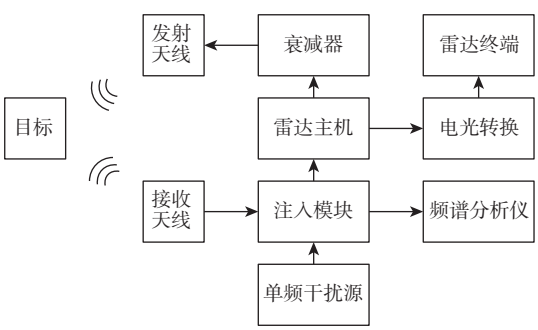


图1 单频连续波伪信号干扰试验配置

Fig. 1 Configuration of single-frequency continuous wave pseudo-signal interference test

由图3可知，受试雷达伪信号干扰敏感频段为(f_0±60) MHz，f_0为雷达中心工作频率，敏感频段均呈“U”型曲线，沿中心频点基本对称，当辐射频偏（电磁辐射干扰频率与受试雷达工作频段中心频率）处于±50～60 MHz时，临界干扰场强随辐射频偏剧烈变化。由此可知，单频伪信号干扰敏感频段明显窄于雷达工作频段，约为工作频段的一半，单频伪信号干扰的防护重点靠近雷达工作中心频点。

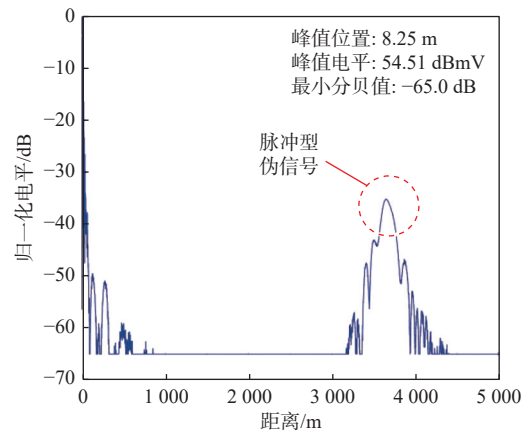


图2 单频脉冲型伪信号干扰示意图

Fig. 2 Schematic diagram of single-frequency pulse pseudo-signal interference

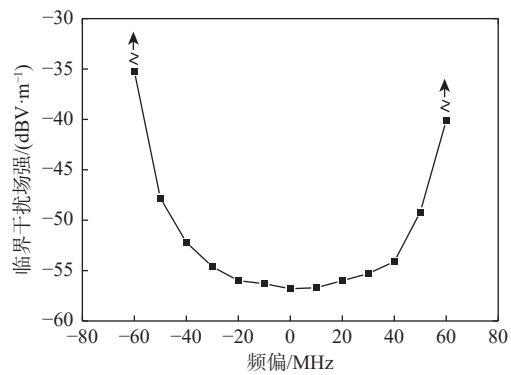


图3 单频伪信号干扰临界干扰场强变化曲线

Fig. 3 Single-frequency pseudo-signal interference critical interference field strength variation curve

2 受试雷达双频伪信号干扰敏感现象

为探明扫频雷达在复杂电磁环境下的伪信号干扰规律，对其进行双频电磁辐射伪信号干扰效应研究。试验配置如图4所示，采用2个信号源产生不同频率的连续波干扰信号，通过合路器与定向耦

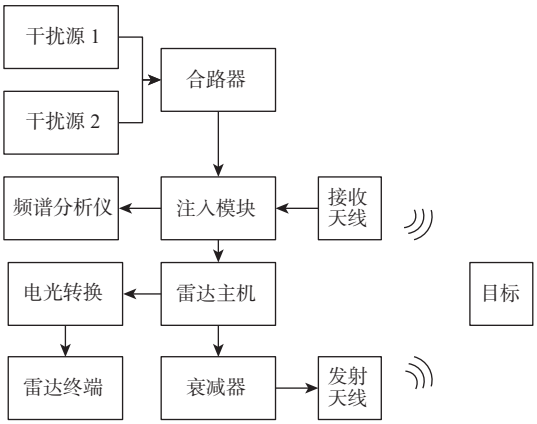


图 4　双频连续波伪信号干扰试验配置

Fig. 4　Configuration of dual-frequency continuous wave pseudo-signal interference test

合注入模块注入雷达接收端口，其他试验配置与单频干扰试验配置保持一致。

2.1　带内双频脉冲型伪信号干扰现象

当双频连续波信号干扰频率随意组合且都位于受试雷达单频伪信号干扰频带内时，其干扰特性基本与单频连续波伪信号干扰特性相同，区别仅为雷达显示界面产生 2 个脉冲型伪信号，而伪信号中心位置均随机产生，选取双频干扰下雷达显示界面如图 5 所示。

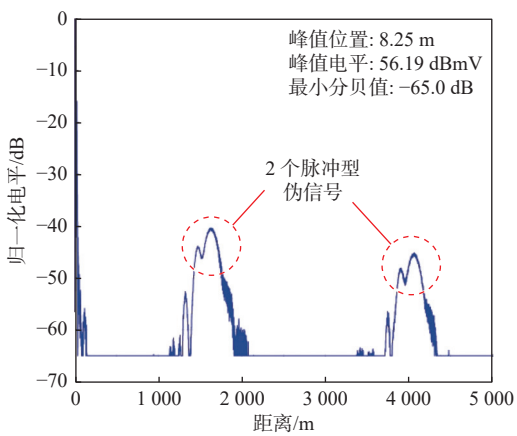


图 5　带内双频脉冲型伪信号干扰示意

Fig. 5　Schematic diagram of in-band dual-frequency pulse pseudo-signal interference

2.2　带外双频脉冲型伪信号干扰现象

在双频电磁辐射试验中发现，若改变双频干扰的频率组合，使双频干扰频率的频差较大，为 540～660 MHz 时，仍会对雷达造成脉冲型伪信号干扰。以干扰频率设为 f_0−300 MHz 和 f_0+300 MHz 为例，任一单频干扰信号均不会对受试雷达产生伪信号干扰，但两干扰信号同时注入雷达时，雷达显示界面出现了如图 6 所示的脉冲型伪信号干扰。

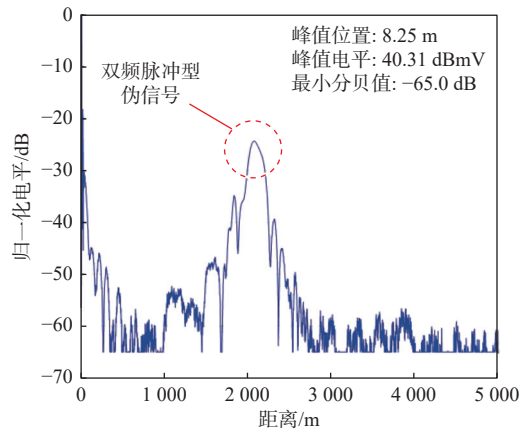


图 6　带外双频脉冲型伪信号干扰示意

Fig. 6　Schematic diagram of out-of-band dual-frequency pulse pseudo-signal interference

对于这一新的脉冲型伪信号干扰敏感现象，不能用非互调迭加机理解释，而需从互调的角度探寻这种伪信号干扰效应的本质原因。

2.3　带内外双频冲激型伪信号干扰现象

在上述试验的基础上，继续改变双频干扰频率的频差组合，试验发现，不论干扰频率位于受试雷达的带内还是带外，当双频干扰频差小于 5 MHz 时，均会对受试雷达造成冲激型伪信号干扰，冲激型伪信号的形态远远窄于脉冲型伪信号干扰，与真实目标信号相似，且冲激型伪信号干扰电平远高于脉冲型伪信号干扰电平，甚至在干扰场强较低时，其伪信号电平值已超过真实回波信号电平值，从而造成雷达误判。这一干扰现象应当引起重视。通过改变双频干扰信号的强度，发现冲激型伪信号有时为 1 个，有时为 2 个，为探明这一现象产生的原因，选取某次雷达受扰界面如图 7 所示，其中，冲激型伪信号干扰细节放大如图 8 所示。

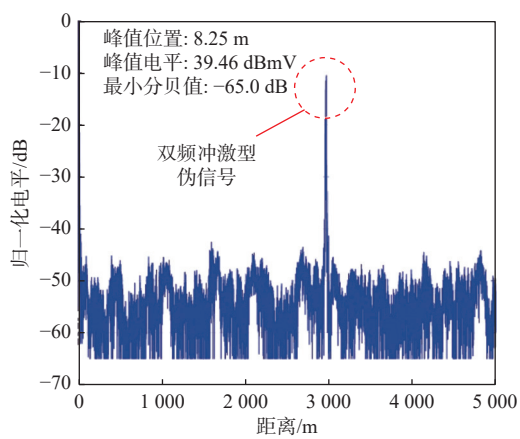


图 7　双频冲激型伪信号干扰示意

Fig. 7　Schematic diagram of dual-frequency impulse pseudo-signal interference

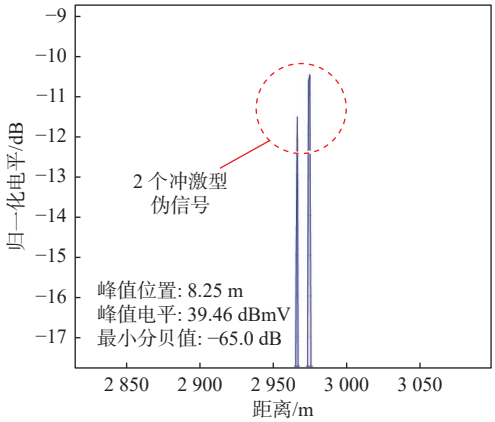


图8 双频冲激型伪信号干扰细节放大

Fig. 8 Enlarged detail of dual-frequency impulse pseudo-signal interference

3 双频连续波伪信号干扰效应机理

针对上述伪信号干扰敏感现象，结合扫频测距雷达的工作原理进行进一步分析，阐明了带外双频脉冲型伪信号干扰的效应机理为二阶互调射频干扰，带内外双频冲激型伪信号干扰的效应机理为二阶互调低频干扰和三阶交互调低频干扰，从而解释了上述敏感现象，带内双频脉冲型伪信号干扰仅仅是单频伪信号干扰的简单叠加，本文不再赘述。

受试扫频测距雷达的原理结构如图9所示（一般扫频连续波雷达原理结构图基本相同）。雷达一本振定频信号与二本振扫频信号混频后经滤波放大后由发射天线发出，到达有用目标后即折返接收天线处，真实回波信号经一级放大先与一本振定频信号混频进行下变频，将较高频段的回波扫频信号降低至中频扫频信号，经过带通滤波放大后与二本振扫频信号混频，得到含有时延信息的零中频信号，而后经低通滤波器滤波，再次放大后的信号经过采样与处理得到目标距离信息。

由图9可设发射天线发出的信号为

$$S(t)=\sum_{i=-N/2}^{N/2}E_{si}\left\{\mathrm{rect}\left[\frac{t-T_{\mathrm{R}}/2-(i+N/2)T_{\mathrm{R}}}{T_{\mathrm{R}}}\right]\cdot \exp\left[\mathrm{j}(2\pi(f_0+i\Delta f)t+\varphi_i)\right]\right\} \tag{1}$$

式中：N为扫频阶梯数；E_{si}为各扫频阶梯内信号场强幅值；T_{R}为单个扫频阶梯跳频时间；f_0为雷达工作频段的中心频率；Δf为扫频阶梯；φ_i为各扫频阶梯信号初相位。设雷达工作频段的带宽为$2f_{\mathrm{H}}$，则$N=2f_{\mathrm{H}}/\Delta f$。

设f_{a}为雷达接收机的一本振频率，则雷达的一本振定频信号可表示为

$$u_{\mathrm{a}}(t)=2\exp(\mathrm{j}2\pi f_{\mathrm{a}}t) \tag{2}$$

设f_{b}为雷达接收机的二本振中心频率，则雷达

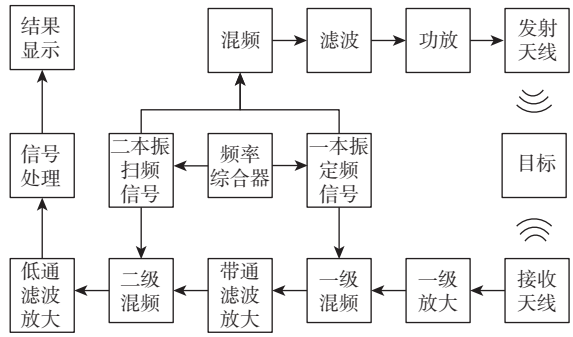


图9 扫频雷达原理结构框图

Fig. 9 Block of swept-frequency radar principle structure

的二本振扫频信号可表示为

$$u_{\mathrm{b}}(t)=2\sum_{i=-N/2}^{N/2}\left\{\mathrm{rect}\left[\frac{t-T_{\mathrm{R}}/2-(i+N/2)T_{\mathrm{R}}}{T_{\mathrm{R}}}\right]\cdot \exp\left[\mathrm{j}(2\pi(f_{\mathrm{b}}+i\Delta f)t+\varphi_i)\right]\right\} \tag{3}$$

可得

$$f_0=f_{\mathrm{a}}+f_{\mathrm{b}} \tag{4}$$

距离为R的真实目标回波信号被雷达接收后可表示为

$$u_{\mathrm{c}}(t)=\sum_{i=-N/2}^{N/2}A_{ci}E_{ci}\left\{\mathrm{rect}\left[\frac{t-T_{\mathrm{R}}/2-(i+N/2)T_{\mathrm{R}}-2R/c}{T_{\mathrm{R}}}\right]\cdot \exp\left[\mathrm{j}(2\pi(f_0+i\Delta f)(t-2R/c)+\varphi_i)\right]\right\} \tag{5}$$

式中：c为光速；A_{ci}为第i个扫频阶梯信号场路耦合到雷达接收机非线性器件前的等效传递函数；E_{ci}为真实目标回波信号在雷达接收天线处的场强幅值。

设干扰频率为f_k的单频连续波干扰信号为

$$S_k(t)=E_k\exp\left[\mathrm{j}(2\pi f_kt+\theta_k)\right] \tag{6}$$

式中：E_k为干扰信号在接收天线处的场强幅值；θ_k为干扰信号初相位。

设雷达接收机非线性器件前接收到的干扰信号为

$$u_k(t)=A_kE_k\exp\left[\mathrm{j}(2\pi f_kt+\theta_k)\right] \tag{7}$$

式中：A_k为干扰信号场路耦合到雷达接收机非线性器件前的等效传递函数。

为进一步释明上述双频伪信号的成像机理，结合受试雷达的工作原理及结构框图，假设：①接收天线接收到工作频段外的双频电磁辐射干扰信号，频率分别为f_1和f_2，令$f_{\mathrm{a}}-f_{\mathrm{H}}<f_2-f_1<f_{\mathrm{a}}+f_{\mathrm{H}}$；②接收天线接收到雷达工作频段附近的双频电磁辐射干扰信号，频率分别为f_3和f_4，令f_4-f_3小于受试雷达低通滤波器的滤波范围，则上述干扰信号进入受试扫频雷达各结构后输出信号各分量的频率如表1

表 1　双频干扰信号进入扫频雷达各结构后输出分量频率

Table 1　Output component frequencies of dual-frequency interfering signals entering each structure of the swept-frequency radar

结构	带外双频脉冲型	带内双频冲激型
天线接收	F_0, f_1, f_2	F_0, f_3, f_4
一级放大	$F_0, f_1, f_2, F_0 \pm f_1, F_0 \pm f_2, f_2 \pm f_1, F_0 \pm (f_2 - f_1)$	$F_0, f_3, f_4, F_0 \pm f_3, F_0 \pm f_4, f_4 \pm f_3, F_0 \pm (f_4 - f_3)$
一级混频(f_a)	$F_0, f_1, f_2, F_0 \pm f_1, F_0 \pm f_2, f_2 \pm f_1, F_0 \pm (f_2 - f_1)$, $F_0 \pm f_a, f_1 \pm f_a, f_2 \pm f_a, F_0 \pm f_1 \pm f_a, F_0 \pm f_2 \pm f_a$, $f_2 \pm f_1 \pm f_a, F_0 \pm (f_2 - f_1) \pm f_a$	$F_0, f_3, f_4, F_0 \pm f_3, F_0 \pm f_4, f_4 \pm f_3, F_0 \pm (f_4 - f_3)$, $F_0 \pm f_a, f_3 \pm f_a, f_4 \pm f_a, F_0 \pm f_3 \pm f_a, F_0 \pm f_4 \pm f_a$, $f_4 \pm f_3 \pm f_a, F_0 \pm (f_4 - f_3) \pm f_a$
带通滤波($f_b \pm f_H$)	$F_0 - f_a, f_2 - f_1$	$F_0 - f_a, f_3 - f_a, f_4 - f_a, F_0 \pm (f_4 - f_3) - f_a$
二级混频 (F_b)	$F_0 - f_a, f_2 - f_1, F_0 - f_a \pm F_b$, $f_2 - f_1 \pm F_b, F_0 - f_a \pm (f_2 - f_1)$	$F_0 - f_a, f_3 - f_a, f_4 - f_a, F_0 \pm (f_4 - f_3) - f_a$, $F_0 - f_a \pm F_b, f_3 - f_a \pm F_b, f_4 - f_a \pm F_b$, $F_0 \pm (f_4 - f_3) - f_a \pm F_b, f_4 \pm f_3, F_0 \pm f_3, F_0 \pm f_4$
低通(窄带)滤波	$F_0 - f_a - F_b$(真实目标回波信号) $f_2 - f_1 - F_b$(二阶互调脉冲型伪信号)	$F_0 - f_a - F_b$(真实目标回波信号) $F_0 \pm (f_4 - f_3) - f_a - F_b$(三阶交互调冲激型伪信号) $f_4 - f_3$(二阶互调冲激型伪信号) $f_3 - f_a - F_b$(一阶脉冲型伪信号) $f_4 - f_a - F_b$(一阶脉冲型伪信号)

所示。其中, 忽略三阶以上分量的影响, 干扰信号频偏受试雷达工作频率不远才能输出表中所示的频率分量。其中, 设雷达工作扫频信号为 F_0, 设 $f_{i0} = f_0 + i\Delta f, i = -\frac{N}{2}, -\frac{N}{2}+1, \cdots, \frac{N}{2}-1, \frac{N}{2}$, 则 $F_0 = \left\{ f_{-\frac{N}{2}0}, f_{(-\frac{N}{2}+1)0}, \cdots, f_{(\frac{N}{2}-1)0}, f_{\frac{N}{2}0} \right\}$; 设雷达二本振扫频信号为 F_b, 设 $f_{ib} = f_b + i\Delta f, i = -\frac{N}{2}, -\frac{N}{2}+1, \cdots, \frac{N}{2}-1, \frac{N}{2}$, 则 $F_b = \left\{ f_{-\frac{N}{2}b}, f_{(-\frac{N}{2}+1)b}, \cdots, f_{(\frac{N}{2}-1)b}, f_{\frac{N}{2}b} \right\}$。

表 1 中, 当 f_3 和 f_4 落入单频脉冲型伪信号干扰敏感频带内时, 才会产生一阶脉冲型伪信号。由表 1 可知, 产生带外双频脉冲型伪信号干扰的根本原因为: 雷达接收机一级放大处产生了二阶互调信号 $f_2 - f_1$, 该信号可顺利通过一本振后的带通滤波器, 从而对雷达造成干扰; 产生带内外双频冲激型伪信号干扰的根本原因为: 雷达接收机一级放大处产生的三阶交互调信号 $F_0 \pm (f_4 - f_3)$ 因频率与雷达回波信号相近, 可顺利通过各级滤波, 最终形成包含真实目标回波信息的三阶交互调冲激型伪信号 $F_0 \pm (f_4 - f_3) - f_a - F_b$; 在雷达后端, 二级混频处产生的二阶互调信号 $f_4 - f_3$ 也可通过后端的低通滤波器, 从而对雷达造成二阶互调冲激型伪信号干扰。

为探明雷达双频伪信号干扰的形态特征, 可结合式 (6), 则表 1 中干扰频率为 f_1、f_2、f_3、f_4 的干扰信号可设为

$$\begin{cases} S_1(t) = E_1 \exp\left[\mathrm{j}(2\pi f_1 t + \theta_1)\right] \\ S_2(t) = E_2 \exp\left[\mathrm{j}(2\pi f_2 t + \theta_2)\right] \\ S_3(t) = E_3 \exp\left[\mathrm{j}(2\pi f_3 t + \theta_3)\right] \\ S_4(t) = E_4 \exp\left[\mathrm{j}(2\pi f_4 t + \theta_4)\right] \end{cases} \tag{8}$$

结合式 (7), 则雷达接收机非线性器件前接收到的干扰信号可表示为

$$\begin{cases} u_1(t) = A_1 E_1 \exp\left[\mathrm{j}(2\pi f_1 t + \theta_1)\right] \\ u_2(t) = A_2 E_2 \exp\left[\mathrm{j}(2\pi f_2 t + \theta_2)\right] \\ u_3(t) = A_3 E_3 \exp\left[\mathrm{j}(2\pi f_3 t + \theta_3)\right] \\ u_4(t) = A_4 E_4 \exp\left[\mathrm{j}(2\pi f_4 t + \theta_4)\right] \end{cases} \tag{9}$$

设扫频雷达在每个扫频阶梯内的采样时刻 $t = t_0 + (i + N/2)T_R$, $i = -N/2, -N/2+1, -N/2+2, \cdots, N/2$。(受篇幅限制, 推导公式未全部列出)。真实回波信号经混频滤波放大采样后可得到采样序列:

$$u_c(t)u_a(t)u_b(t) \rightarrow A_{ci}E_{ci}\exp\left[\mathrm{j}2\pi f_b(-2R/c)\right] \cdot \exp\left[\mathrm{j}2\pi i\Delta f(-2R/c)\right] \tag{10}$$

带外二阶互调脉冲型伪信号的采样序列为

$$\begin{aligned} u_1(t)u_2(t)u_b(t) \rightarrow & \frac{1}{2}A_1A_2E_1E_2 \cdot \\ & \exp\left[\mathrm{j}(2\pi(f_2 - f_1 - f_b - f_H)t_0 + (\theta_2 - \theta_1))\right] \cdot \\ & \exp\left[\mathrm{j}2\pi((f_2 - f_1 - f_b - f_H)T_R - \Delta f t_0)(i + N/2)\right] \cdot \\ & \exp\left[-\mathrm{j}2\pi\Delta f T_R(i + N/2)^2\right] \end{aligned} \tag{11}$$

二阶互调冲激型伪信号的采样序列为

$$\begin{aligned} u_1(t)u_2(t) \rightarrow & \frac{1}{2}A_3A_4E_3E_4\exp\left[\mathrm{j}(2\pi(f_4 - f_3)t_0 - (\theta_4 - \theta_3))\right] \cdot \\ & \exp\left[\mathrm{j}2\pi(f_4 - f_3)T_R(i + N/2)\right] \end{aligned} \tag{12}$$

三阶交互调冲激型伪信号的采样序列为

$$\begin{aligned} u_1(t)u_2(t)u_c(t)u_a(t)u_b(t) \rightarrow & \frac{1}{4}A_3A_4A_{ci}E_3E_4E_{ci} \cdot \\ & \exp\left[\mathrm{j}(2\pi(f_4 - f_3)t_0 + (\theta_4 - \theta_3))\right] \cdot \\ & \exp\left[\mathrm{j}2\pi\left((f_4 - f_3)T_R - \Delta f\frac{2R}{c}\right)(i + N/2)\right] \end{aligned} \tag{13}$$

式 (10)～式 (13) 中不含 i 的指数项仅对目标信号电平幅值起作用; 含有 i 的一次相位的指数项经

快速逆傅里叶变换与门限判决可得到目标回波距离信息；式 (11) 中指数项 $\exp[-\mathrm{j}2\pi\Delta f T_{\mathrm{R}}(i+N/2)^2]$ 因含有 i 的二次相位，所以使得伪信号能量分散，波形展宽，最终呈现为脉冲型伪信号，而式 (10)、式 (12) 和式 (13) 因不含有 i 的二次相位，故能量集中，冲激型伪信号的波形形态与真实回波目标形态几乎相同。

由此可得到真实回波目标的距离为

$$R=\frac{ck}{2N\Delta f} \tag{14}$$

带外二阶互调脉冲型伪信号的距离为

$$R_{21}=\frac{\left[(f_2-f_1-f_{\mathrm{b}}-f_{\mathrm{H}})T_{\mathrm{R}}-\Delta f t_0\right]c}{2\Delta f} \tag{15}$$

因雷达每次探测时 t_0 为随机量，所以带外二阶互调脉冲型伪信号位置也是随机的。

二阶互调冲激型伪信号的距离为

$$R_{43}=\frac{(f_4-f_3)T_{\mathrm{R}}c}{2\Delta f} \tag{16}$$

三阶交互调冲激型伪信号的距离为

$$R_{\mathrm{c}43}=\frac{(f_4-f_3)T_{\mathrm{R}}c}{2\Delta f}+R \tag{17}$$

由式 (16) 和式 (17) 可知，当干扰信号频差固定时，二阶互调冲激型伪信号和三阶交互调冲激型伪信号的位置固定，且两者之间的间距为真实回波信号距离。

4 试验验证

为验证机理分析的正确性，采用图 4 的试验配置设定干扰源 1 和干扰源 2 的干扰频率 $f_1=f_0-300$ MHz和 $f_2=f_0+300$ MHz的干扰信号，易知干扰信号频差为 $\Delta f_{21}=600$ MHz，在受试雷达正常工作状态下开启干扰信号，在雷达发射端口分别设置 30 dB 和 40 dB 的衰减，改变双频信号的干扰场强，试验结果如表 2 所示，由此绘制带外双频脉冲型伪信号干扰电平随干扰场强的变化曲线，如图 10 所示。

由表 2 和图 10 可知，当双频干扰场强较低时，受试雷达工作于线性区，伪信号电平值随干扰场强的增大而呈近似线性增长，当干扰场强增大至受试雷达工作于强非线性区时，伪信号电平值基本不随干扰场强的增大而改变。改变有用信号强度对伪信号干扰电平值几乎无影响，伪信号干扰电平值仅与两干扰信号场强有关，最大值可达到 16 dBmV 左右，证明带外双频伪信号干扰的本质为带外二阶互调干扰，这验证了理论分析的正确性。带外双频二阶互调脉冲型伪信号的中心距离随机，原因为式 (15) 中含有随机量 t_0。

表 2 带外双频二阶互调脉冲型伪信号试验结果

Table 2 Test results of out-of-band dual-frequency second-order intermodulation pulse pseudo-signals

发射信号衰减程度/dB	干扰场强1/(dBV·m^{-1})	干扰场强2/(dBV·m^{-1})	伪信号电平/dBmV	伪信号中心距离/m
30	−16	−16	3.08	2 893
	−13	−13	7.66	4 011
	−10	−10	12.38	329.4
	−7	−7	15.66	1 892
	−4	−4	15.70	2 939
	−1	−1	15.84	224.5
40	−15	−15	3.91	3 741
	−12	−12	9.59	3 462
	−9	−9	13.69	660.4
	−6	−6	15.33	1 548
	−3	−3	15.59	1 254
	−1	−1	15.73	2 143

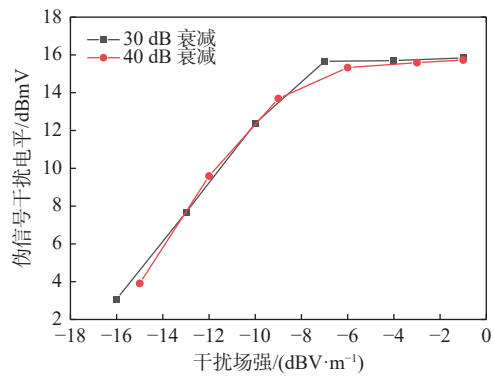


图 10 带外双频二阶互调脉冲型伪信号干扰电平随干扰场强的变化

Fig. 10 Out-of-band dual-frequency second-order intermodulation pulse pseudo-signal level variation with critical interference field strength

与前述试验配置相同，改变干扰源频率 $f_1=f_0$、$f_2=f_0+1$ MHz的干扰信号，易知干扰信号频差为 $\Delta f_{21}=1$ MHz，改变双频信号的干扰场强，试验结果如表 3 所示，由此绘制双频冲激型伪信号干扰电平随临界干扰场强的变化曲线，如图 11 所示。

由表 3 和图 11 可知，改变有用信号强度仅会对三阶交互调冲激型伪信号干扰电平值产生影响。2 种伪信号距离值为固定值，基本保持不变，依据式 (16) 和式 (17)，由干扰信号频率的频差可算出两伪信号的理论距离值为 750 m 和 758.3 m，两伪信号距离差值为真实回波信号距离，该试验结果也证明了理论分析的正确性。当双频干扰场强较低时，受试雷达工作于线性区，二阶互调和三阶交互调冲激型伪信号电平值随干扰场强的增加而呈近似线性增长；当干扰场强继续增大，受试雷达工作于强非

表 3　双频二阶互调与三阶交互调冲激型伪信号试验结果

Table 3　Test results of two-frequency second-order intermodulation and third-order intermodulation impulse pseudo-signal

发射信号衰减程度/dB	干扰场强1/(dBV·m^{-1})	干扰场强2/(dBV·m^{-1})	二阶互调伪信号电平/dBmV	三阶交互调伪信号电平/dBmV	二阶互调伪信号距离/m	三阶交互调伪信号距离/m	距离差/m
30	−60	−60	2.91	13.27	755.9	764.1	8.3
	−55	−55	7.26	19.99	759.7	767.8	8.2
	−50	−50	13.23	26.69	758.9	767.1	8.3
	−45	−45	17.5	31.24	758.2	766.3	8.2
	−40	−40	21.77	33.74	762.7	770.8	8.2
	−35	−35	25.92	32.97	762.7	770.8	8.2
	−30	−30	28.45	29.99	746.9	755.1	8.3
	−25	−25	28.78	27.52	749.2	757.3	8.2
	−20	−20	27.43	24.25	746.2	754.4	8.3
	−15	−15	27.54	20.98	745.4	753.6	8.3
	−10	−10	28.18	17.81	748.4	756.6	8.3
	−5	−5	28.22	14.39	752.2	760.3	8.2
40	−60	−60	2.21	4.77	749.5	757.6	8.2
	−55	−55	8.53	10.83	746.5	754.7	8.3
	−50	−50	13.25	16.52	751	759.1	8.2
	−45	−45	16.58	20.3	751	759.1	8.2
	−40	−40	21.13	21.92	753.2	761.4	8.3
	−35	−35	25.05	19.99	753.2	761.4	8.3
	−30	−30	27.51	18.03	748	756.1	8.2
	−25	−25	28.74	15.14	748.7	756.9	8.3
	−20	−20	27.35	12.67	751.7	759.8	8.3
	−15	−15	27.55	10.74	749.5	757.6	8.2
	−10	−10	27.32	8.13	750.2	758.4	8.3
	−5	−5	28.31	5.91	751	759.1	8.2

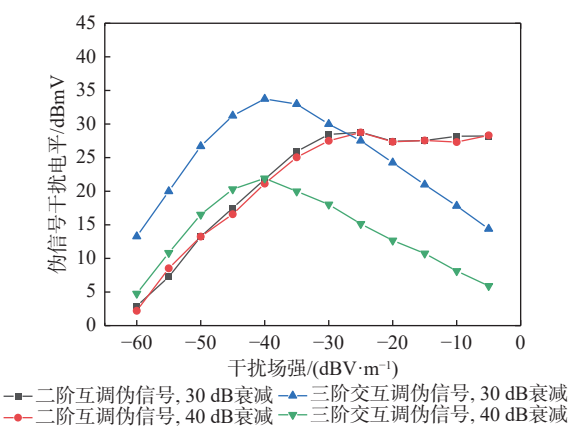


图 11　双频冲激型伪信号干扰电平随干扰场强的变化

Fig. 11　Dual-frequency impulse pseudo-signal level variation with critical interference field strength

线性区, 二阶互调伪信号电平值逐渐增大至最大值后保持恒定, 三阶交互调伪信号电平值达到最大值后逐渐降低, 原因为: 真实目标回波信号在雷达强非线性区被压制, 影响了三阶交互调伪信号的强度, 而二阶互调信号却不受影响。雷达工作于线性区时, 若目标较近, 则真实回波信号较强, 三阶交互调伪信号干扰电平更高, 为主要干扰信号; 若目标较远, 则真实回波信号较弱, 二阶互调伪信号干扰电平更高, 为主要干扰信号; 雷达工作于强非线性区时, 三阶交互调伪信号逐渐消失, 二阶互调伪信号为主要干扰信号。

5　结　论

本文基于受试雷达接收电路频率分析揭示了雷达双频伪信号干扰效应机理, 结论如下:

1) 受试扫频测距雷达单频脉冲型伪信号干扰的敏感频段为 (f_0±60) MHz, 敏感频段呈“U”型曲线, 沿中心频点基本对称, 单频伪信号干扰敏感频段明显窄于雷达工作频段。

2) 受试雷达双频脉冲型伪信号干扰的效应机理为二阶互调射频干扰, 该类伪信号出现的条件为: 双频干扰频差处于 540～660 MHz; 双频冲激型

伪信号干扰的效应机理为二阶互调低频干扰和三阶交互调低频干扰，该类伪信号出现的条件为双频干扰频差小于5 MHz。

3）二阶互调脉冲型伪信号采样序列的距离项中包含采样时刻t_0，因该值为随机值，所以其产生位置也是随机的；二阶互调冲激型伪信号产生位置固定，且由双频干扰频差决定；三阶交互调冲激型伪信号产生位置固定，且由双频干扰频差和有用目标距离共同决定。

4）冲激型伪信号干扰电平的最大值（35 dBmV以上）远大于脉冲型伪信号干扰电平最大值（约16 dBmV）。若探测目标距离较近，则防护应侧重于三阶交互调伪信号干扰；若干扰场强较强，则防护应侧重于二阶互调伪信号干扰。

参考文献（References）

[1] 许彤, 陈亚洲, 王玉明, 等. 无人机数据链带内连续波电磁干扰效应研究[J]. 北京理工大学学报, 2021, 41(10): 1084-1094.
XU T, CHEN Y Z, WANG Y M, et al. Research on in-band continuous wave electromagnetic interference effect of unmanned aerial vehicle data link[J]. Transactions of Beijing Institute of Technology, 2021, 41(10): 1084-1094(in Chinese).

[2] 刘尚合, 武占成, 张希军. 电磁环境效应及其发展趋势[J]. 国防科技, 2008, 29(1): 1-6.
LIU S H, WU Z C, ZHANG X J. Electromagnetic enviroment effect and its development trends[J]. National Defense Science & Technology, 2008, 29(1): 1-6(in Chinese).

[3] 孙国至, 刘尚合, 陈京平, 等. 战场电磁环境效应对信息化战争的影响[J]. 军事运筹与系统工程, 2006, 20(3): 43-47.
SUN G Z, LIU S H, CHEN J P, et al. The influence on information-based war of battle field electromagnetic environment effects[J]. Military Operations Research and Systems Engineering, 2006, 20(3): 43-47(in Chinese).

[4] 毛二可, 龙腾, 韩月秋. 频率步进雷达数字信号处理[J]. 航空学报, 2001, 22(S1): 16-25.
MAO E K, LONG T, HAN Y Q. Digital signal processing of stepped frequency radar[J]. Acta Aeronautica et Astronautica Sinica, 2001, 22(S1): 16-25(in Chinese).

[5] TIKOLAM M, SAMCOVIC A, NIKOLIC D, et al. An algorithm for detection of electromagnetic interference in high frequency radar range-Doppler images caused by LEDs[J]. IEEE Access, 2019, 7: 84413-84419.

[6] 郭栋, 王世军, 魏冬峰. 高功率微波武器对雷达的威胁及其防护[J]. 指挥信息系统与技术, 2011, 2(5): 74-77.
GUO D, WANG S J, WEI D F. Threat of high power microwave weapon on radars and its countermeasures[J]. Command Information System and Technology, 2011, 2(5): 74-77(in Chinese).

[7] SEIFI Z, GHORBANI A, ABDIPOUR A. Analysis and experimental study of radiative microwave pulses effects on the nonlinear performance of a low-noise amplifier[J]. IEEE Transactions on Plasma Science, 2021, 49(3): 1105-1114.

[8] 张冬晓, 陈亚洲, 程二威, 等. 无人机信息链路电磁干扰效应规律研究[J]. 北京理工大学学报, 2019, 39(7): 756-762.
ZHANG D X, CHEN Y Z, CHENG E W, et al. Effects of electromagnetic interference (EMI) on information link of UAV[J]. Transactions of Beijing Institute of Technology, 2019, 39(7): 756-762(in Chinese).

[9] 赵宏泽, 魏光辉, 杜雪, 等. 卫星导航接收机三阶互调阻塞效应分析[J]. 系统工程与电子技术, 2022, 44(4): 1336-1342.
ZHAO H Z, WEI G H, DU X, et al. Analysis of third-order intermodulation blocking effect for satellite navigation receiver[J]. Systems Engineering and Electronics, 2022, 44(4): 1336-1342(in Chinese).

[10] OLIVIER K, CILLIERS J, DU PLESSIS M. Design and performance of wideband DRFM for radar test and evaluation[J]. Electronics Letters, 2011, 47(14): 824-825.

[11] LI Y, CHEN D, ZHANG X L. Research on EMI radiation source model based on battlefield EMC prediction analysis[J]. IOP Conference Series: Materials Science and Engineering, 2019, 569(3): 032075.

[12] 丁国胜. 舰船大功率雷达电磁辐射场强分析和改善技术研究[D]. 南京: 南京理工大学, 2012.
DING G S. Study on analysis of electromagnetic radiation field strength from ship-borne high power radar and improvement technology[D]. Nanjing: Nanjing University of Science and Technology, 2012(in Chinese).

[13] ZHAO B, ZHOU F, BAO Z. Deception jamming for squint SAR based on multiple receivers[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 8(8): 3988-3998.

[14] ZHAO B, HUANG L, ZHANG J H. Single channel SAR deception jamming suppression via dynamic aperture processing[J]. IEEE Sensors Journal, 2017, 17(13): 4225-4230.

[15] ZHAO H Z, WEI G H, PAN X D, et al. Pseudo-signal interference regularity of single-frequency electromagnetic radiation to stepped-frequency radar[J]. Electronics, 2022, 11(17): 2768.

[16] DU X, WEI G, ZHAO H Z, et al. Research on continuous wave electromagnetic effect in swept frequency radar[J]. Mathematical Problems in Engineering, 2021, 2021: 4862451.

[17] 李文臣, 黄烽, 杨会民, 等. 雷达噪声干扰和多假目标干扰效能分析[J]. 中国电子科学研究院学报, 2013, 8(4): 403-406.
LI W C, HUANG F, YANG H M, et al. Efficiency analysis of radar noise jamming and multiple false target jamming[J]. Journal of China Academy of Electronics and Information Technology, 2013, 8(4): 403-406(in Chinese).

[18] 徐飞虎. 系统非线性对高频雷达目标探测的影响[D]. 哈尔滨: 哈尔滨工业大学, 2010.
XU F H. The impact on target detection of high frequency radar as system's nonlinearity[D]. Harbin: Harbin Institute of Technology, 2010(in Chinese).

[19] BIONDI A, VALLOZZI L, ROGIER H, et al. Electromagnetic compatibility aware design and testing of intermodulation distortion under multiple co-located sources illumination[J]. IET Science, Measurement & Technology, 2012, 6(2): 105-112.

[20] 杜文越. Ku波段导航雷达系统设计与实现[D]. 北京: 北京理工大学, 2019.
DU W Y. Design and implementation of Ku band navigation radar sys-

tem[D]. Beijing: Beijing Institute of Technology, 2019(in Chinese).

[21] 中央军委装备发展部. 系统电磁环境效应试验方法: GJB 8848—2016[S]. 北京: 中央军委装备发展部, 2016: 51-53.
Equipment Development Department of the Central Military Commission. Electromagnetic environmental effects test methods for systems: GJB 8848—2016[S]. Beijing: Equipment Development Department of the Central Military Commission, 2016: 51-53(in Chinese).

Dual-frequency continuous wave pseudo-signal interference effect in swept-frequency radar

ZHAO Hongze, WEI Guanghui*, PAN Xiaodong, DU Xue, LU Xinfu

(National Key Laboratory on Electromagnetic Environment Effects, Army Engineering University, Shijiazhuang 050003, China)

Abstract: To fully understand the impact of dual-frequency electromagnetic radiation pseudo-signal interference on radar and uncover the underlying mechanism behind this interference effect. Taking a Ku-band swept-frequency ranging radar as the test object, a dual-frequency continuous wave pseudo-signal interference effect test was carried out, and the effect mechanism of second-order intermodulation RF interference, second-order intermodulation low-frequency interference and third-order intermodulation low-frequency interference was revealed respectively. The variation rule of the pseudo-signal level with the interference field intensity was determined, and the morphological characteristics and generation position of the pseudo-signal were explained. The test results indicate that when the difference in interference frequency is approximately between 540 MHz and 660 MHz, the out-of-band dual-frequency second-order intermodulation signal will generate pseudo-signal interference of the wide-pulse type, with a random location for the pseudo-signal. On the other hand, when the difference in interference frequency is less than 5 MHz, both the dual-frequency second-order intermodulation signal and the third-order intermodulation signal will produce pseudo-signal interference of the impulse type, with a fixed location for the pseudo-signal.

Keywords: swept-frequency radar; dual-frequency interference; electromagnetic radiation; pseudo-signal interference; second-order intermodulation

Received: 2022-08-24; **Accepted**: 2022-12-16; **Published Online**: 2023-01-04 19:28

URL: link.cnki.net/urlid/11.2625.V.20230104.1423.002

Foundation item: National Key Laboratory Fund of Electromagnetic Environment Effects (6142205200301)

* **Corresponding author**. E-mail: wei-guanghui@sohu.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0702

大行程磁阻驱动微定位平台及其轨迹跟踪控制

张旭[1]，陆晓伟[2]，赖磊捷[1,*]

(1. 上海工程技术大学 机械与汽车工程学院，上海 201620；2. 华立科技股份有限公司，杭州 310023)

摘 要：麦克斯韦磁阻驱动器克服了传统压电陶瓷驱动器行程小及音圈电机推力密度和效率低等难题，在大行程高速微纳米定位中具有极大的应用潜力。设计了一种基于麦克斯韦磁阻驱动的大行程二自由度柔顺微定位平台，并对其进行了高性能轨迹跟踪控制。该定位平台由2个永磁偏置的麦克斯韦磁阻驱动器和二自由度并联解耦柔性导向机构组成，平台利用磁阻驱动器的非线性负刚度来部分补偿柔性导向机构的弹性恢复力，有效提高了运动行程和能量传递效率，使其在±2 mm行程范围内的所需推力从±120 N减少至±24 N。轨迹跟踪控制方面，为补磁阻驱动器偿驱动器软磁材料冲磁带来的磁滞非线性，利用逆Prandtl-Ishlinskii磁滞模型构建了率相关磁滞补偿控制器并将其置于前馈回路中。为解决平台低阻尼谐振及动力学模型不一致等问题，设计了含分数阶相位超前环节的PI反馈控制器以完成对系统开环频率特性的灵活调整，实现了对平台的高精度轨迹跟踪控制，有效减小了系统的跟踪误差。对2 mm幅值、1 Hz和10 Hz频率的三角波信号进行轨迹跟踪所得到的均方根误差分别为0.013 mm和0.017 mm。

关 键 词：微纳米定位；麦克斯韦磁阻驱动器；柔性机构；磁滞；分数阶控制；轨迹跟踪控制
中图分类号：TM352
文献标志码：A **文章编号**：1001-5965(2024)09-2852-10

大行程超精密微纳米定位通常指能够在毫米级 (≥1 mm)行程范围内实现亚微米或纳米级(0.1～100 nm)分辨率或定位精度的技术[1-2]。目前，广泛使用的压电陶瓷驱动器通常具有纳米级精度，但其行程一般只有几十微米[3]。另一种常见的驱动器——音圈电机，存在推力密度和效率低等缺点，难以完成高频响快速跟踪任务[4]。近年来，文献[5-6]先后研制了多种基于变磁阻正应力(又称麦克斯韦力)驱动的磁阻驱动器。有别于音圈电机的洛伦兹力驱动原理，磁阻驱动器依靠磁阻最小原理产生变磁阻正应力，其大小与气隙处磁感应强度平方成正比，且励磁线圈集中绕制在定子上，因此，该类驱动器具有力密度大、惯量低、响应速度快、结构简单及热负荷能力强等优点[7]。

基于上述驱动原理，同时利用柔性机构在微纳米定位中所具有的无摩擦和无间隙等优点，国内外学者设计了多种基于麦克斯韦力磁阻驱动的柔性纳米定位平台。例如，文献[8]设计了一种基于复合磁阻驱动器的大行程快速纳米定位平台，其工作行程达2 mm，一阶谐振为178 Hz；文献[9]设计了结合麦克斯韦力电磁致动和柔性放大机构的一维纳米定位平台，该平台能在±95 μm的行程内实现743 Hz的一阶固有频率，同时该团队还提出一种轴对称布置的磁轭结构，该结构使用单个永磁体为双电枢提供偏置磁通，可实现54.4 μm的运动行程和782 Hz的一阶谐振频率；文献[10-12]基于C型磁轭结构设计了多种麦克斯韦力电磁致动的快速刀具伺服(fast tool servo, FTS)系统。尽管磁阻驱动器

收稿日期：2022-08-10；**录用日期**：2022-11-12；**网络出版时间**：2022-11-25 13:53
网络出版地址：link.cnki.net/urlid/11.2625.V.20221124.1443.002
基金项目：国家自然科学基金(U2013211)；上海市自然科学基金(21ZR1426000)；机械系统与振动国家重点实验室课题资助项目(MSV202210)
*通信作者. E-mail：lailj@sues.edu.cn

引用格式：张旭，陆晓伟，赖磊捷. 大行程磁阻驱动微定位平台及其轨迹跟踪控制[J]. 北京航空航天大学学报，2024，50(9)：2852-2861. ZHANG X，LU X W，LAI L J. Large-stroke microposition stage driven by reluctance actuator and its trajectory tracking control[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2852-2861 (in Chinese).

相较于压电驱动器和音圈电机等具有明显优势, 但上述平台大多为中小行程(几十至几百微米), 磁阻驱动器在大气隙下存在的磁滞、漏磁及动力学模型不一致等问题, 严重限制了其在大行程微纳米定位平台上的应用, 因此, 需要设计有效的控制和磁滞补偿算法来解决以上问题。

针对磁滞造成的非线性误差, 利用磁滞逆模型进行前馈补偿是一种常规而有效的手段, 即利用反演技术来获得 Preisach、Bouc-Wen 和 Prandtl-Ishlinskii (P-I) 等模型的逆模型, 并将其作用在前馈或反馈路径上, 以完成对磁滞非线性的补偿[13-15]。针对微定位平台的低阻尼谐振和动力学特性不一致等问题, 国内外学者也开展了相关研究。例如, 文献 [16] 在闭环系统中模型前方串联陷波滤波器, 使得系统模型与陷波滤波器的零极点对消, 从而抑制了系统谐振。文献 [17] 提出 PID 模型预测控制算法, 结合对期望轨迹的前馈控制, 能够实现优于普通 PID 和传统模型预测控制的定位性能。当前诸如陷波滤波、动力学逆模型前馈和主动阻尼控制等主流控制器较多依赖于系统的建模精度, 由于大行程磁阻驱动微定位平台在不同位置的动力学模型的不一致性, 因此诸多方法存在局限性。

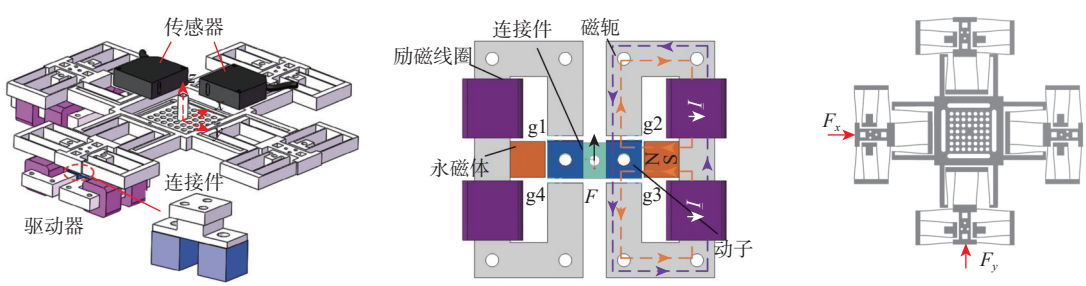
基于上述文献, 本文设计了包含 2 个麦克斯韦磁阻驱动器和二自由度并联解耦柔性导向机构的大行程微定位平台, 并利用综合分数阶相位超前的 PI(fractional order phase lead PI, FOPLPI) 控制器和率相关磁滞补偿控制器的复合控制方法, 实现了大行程微定位平台的高速高精度轨迹跟踪控制。首先, 对麦克斯韦磁阻驱动器和二维解耦柔性机构的工作原理进行说明, 阐述了驱动器的负刚度特性及其对增加平台行程和能量效率的作用。利用直接逆磁滞前馈补偿方法设计了率相关磁滞补偿控制器, 完成了对微定位平台磁滞特性的补偿。其次, 为解决系统低阻尼谐振、谐振点漂移及动力学模型不一致等复杂动力学问题, 设计了 FOPLPI 控制器。最后, 对微定位平台进行了跟踪实验, 将综合了 FOPLPI 控制器和率相关磁滞补偿控制器的复合控制方法与其他控制方法进行对比, 证明了本文提出的微定位平台和控制方法的有效性。本文的主要创新点为: 充分利用了磁阻驱动器的非线性负刚度来部分补偿柔性机构的弹性恢复力, 有效提高了平台运动行程和能量传递效率, 同时, 利用 FOPLPI 控制器解决了低阻尼谐振及动力学模型不一致等问题, 并与率相关磁滞补偿控制器共同构成了前馈+反馈的复合控制回路。

1　磁阻驱动微定位平台

图 1 为磁阻驱动微定位平台的整体布局, 该平台由二自由度并联解耦柔性导向机构和 2 个磁阻驱动器组成。图 1(b) 为磁阻驱动器结构, 由硅钢片压叠制作而成的 C 型磁轭组成了驱动器的定子, 由汝铁硼(Nd-Fe-B)烧结而成的永磁体构成永磁偏置磁路用于改善驱动器的非线性[10]。励磁线圈安装在磁轭固定位置处, 提供驱动器运动所需的励磁磁场。g1、g2、g3 和 g4 分别代表了驱动器的不同气隙。驱动器磁场主要由 2 部分组成: 永磁体产生的偏置磁场、向线圈施加电流 I 时产生的励磁磁场。由麦克斯韦电磁正应力计算公式及磁场的可叠加性可知, 当偏置磁场和励磁磁场叠加作用时, 动子左右两侧磁感应强度就会出现差值, 即动子受到差动的麦克斯韦电磁正应力, 可产生与输入电流 I 呈良好线性关系的驱动力 F。

如图 1(c) 所示, 柔性导向机构采用冗余对称约束结构, 2 个驱动支链及其对应辅助支链两两对称分布于定位平台的 x 和 y 方向, 以提升平台的固有频率和负载能力, 同时有效消除轴间的耦合位移。平台支链充分利用内部的两级复合双平行四边形机构在 xy 方向上的刚度特性和解耦功能, 实现了输入输出全解耦的二自由度平动并联解耦微定位平台[18]。该机构材料为 7075 铝合金, 采用线切割电火花加工技术制成。

为对麦克斯韦磁阻驱动微定位平台的轨迹跟踪性能进行测试, 搭建了如图 2 所示的基于实时仿真机(LinksBox-02)和数据采集卡(PCI-6251)的硬



(a) 微定位平台　(b) 麦克斯韦磁阻驱动器　(c) 柔性导向机构

图 1　麦克斯韦磁阻驱动微定位系统

Fig. 1　Microposition system driven by Maxwell reluctance actuator

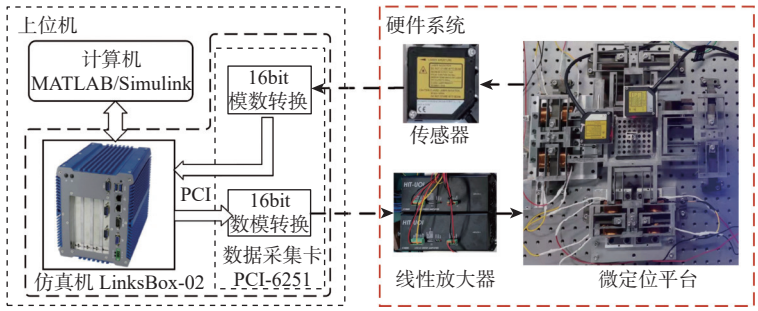


图2 微定位平台硬件布局

Fig. 2 Hardware layout of microposition stage

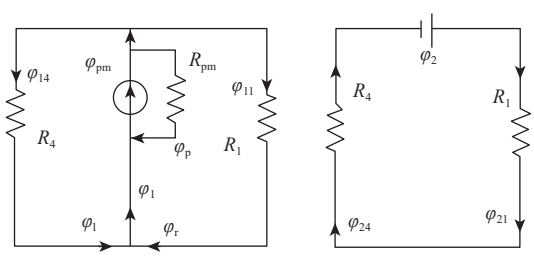
件在环实时仿真系统。其中，线性放大器(CH808)将数据采集卡(PCI-6251)的输出电压(−10～+10 V)线性转换为驱动器输入电流(−8～+8 A)。平台位移由激光位移传感器(HL-G103-S-J，量程为±4 mm)进行测量，并转换为−10～+10 V的模拟量输出，由数据采集卡对其进行实时采集。

2 系统等效模型建立

随着气隙宽度的不断增加(本文为2 mm)，气隙中漏磁及磁场分布不均匀的程度增加，磁阻驱动器推力密度(电机常数)随着气隙宽度的增加而减小，且呈现严重的非线性特性。为解释磁阻定位平台的工作原理并分析大气隙条件下平台非线性特性的变化规律，需要建立磁阻驱动器的磁路及平台的整体动力学模型。

图3为偏置磁场和励磁磁场在不考虑漏磁情况下的理想等效磁路。图3(a)中，R_{pm}为永磁体的磁阻，φ_p为通过R_{pm}的磁通，φ_{pm}为永磁体内部磁场强度等于0时所产生的总磁通，φ_1为流入永磁体的总偏置工作磁通，$\varphi_1=\varphi_{14}$和$\varphi_r=\varphi_{11}$分别为流经左右气隙磁阻$R_4=(x_0+x)/(\mu_0 S)$和$R_1=(x_0-x)/(\mu_0 S)$的磁通量。图3(b)中，气隙磁阻R_1和R_4串联，因此，流经磁阻R_1和R_4的磁通量φ_{21}和φ_{24}相等，且等于磁路中的总磁通量φ_2。当永磁体单独作用时，由电磁学理论计算公式可得

$$\begin{cases}\varphi_{pm}=B_r S\\ R_{pm}=L_{pm}/\mu_0 S\end{cases} \tag{1}$$



(a) 偏置磁路 (b) 励磁磁路

图3 磁阻驱动器理想等效磁路

Fig. 3 Ideal equivalent magnetic circuit of reluctance actuator

式中：μ_0为真空磁导率；S为垂直于磁通量流向的磁轭的横截面积；B_r为永磁体剩磁；L_{pm}为永磁体高度。

联立磁阻计算公式R_1和R_4，此时，与之对应的磁感应强度B_{11}和B_{14}分别为：$(L_{pm}B_r)/(x_0-x)$和$(L_{pm}B_r)/(x_0+x)$。同时，对于图3(b)所示的励磁磁路，如下公式成立：

$$\begin{cases}\varphi_2=\dfrac{2NI}{R_1+R_4}=\dfrac{\mu_0 NIS}{x_0}\\ B_{21}=B_{24}=\dfrac{\mu_0 NI}{x}\end{cases} \tag{2}$$

式中：x_0为初始气隙宽度，N=200为线圈匝数。

此时，由麦克斯韦电磁正应力公式可得，驱动器输出力F_{id}可定义为

$$F_{id}=\frac{S}{2\mu_0}\left[(B_{11}+B_{21})^2-(B_{14}+B_{24})^2\right] \tag{3}$$

基于本节分析，实际工作中考虑三维漏磁系数的驱动器推力模型可表达为

$$F_a=2\frac{S}{2\mu_0}\left(B_{g1}^2-B_{g4}^2\right)=k_m I+k_a x \tag{4}$$

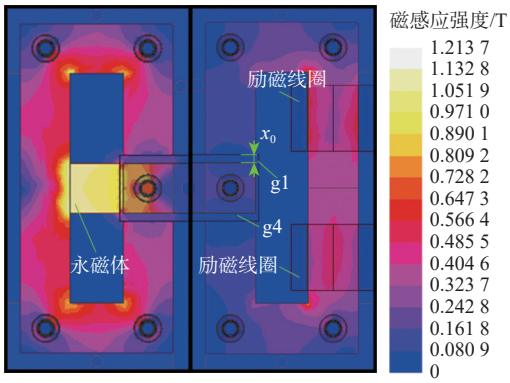
式中：k_m和k_a分别为驱动器电机常数和电机刚度，且均受漏磁系数的影响。

本文利用有限元仿真软件Ansys Electronic Desktop中的Maxwell 3D模块对磁阻驱动器进行了三维磁场分析，获得了电机常数k_m和电机刚度k_a，结果如图4所示。其中，图4(a)为仿真过程中磁阻驱动器励磁及永磁磁路中的漏磁情况，电机刚度及电机常数通过后处理计算获得，结果如图4(b)和图4(c)所示。

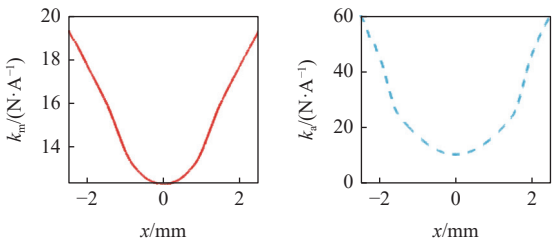
为探究磁阻驱动器非线性电机常数及电机刚度对推力的影响，参考二阶系统集总质量模型得到其运动学方程：

$$F=k_m I+k_a x=m\ddot{x}+c\dot{x}+kx \tag{5}$$

式中：m为平台运动部分的质量；c为阻尼，其通常比临界阻尼小得多；k为柔性机构弯曲刚度，k=60 N/mm，通过$k=2Ewh^3/l^3$进行计算，E=71.7 GPa为弹性模量，其由柔性机构材料决定，梁长度l、宽度



(a) 磁场云图



(b) 电机常数 k_m　　(c) 电机刚度 k_a

图 4　磁阻驱动器电磁仿真结果

Fig. 4　Electromagnetic simulation results of reluctance actuator

w 和厚度 h 分别为 63 mm、15 mm 和 2 mm。如图 5 所示，在受到 x 及 y 方向大小为 130 N 的推力时，x-x 和 y-y 方向位移均为 2.167 mm。另外，柔性导向机构在正交方向的耦合位移 x-y 和 y-x 分别为 0.014 mm 和 0.016 mm，耦合度分别为 0.7% 和 0.8%。且当平台处于准静态时，式 (5) 可改写为

$$I = F_f(x)/k_m(x) = x(k-k_a)/k_m \tag{6}$$

式中：$F_f = (k-k_a)x$ 为随着平台位置的变化需要对其总刚度进行补偿的力。

如图 6 所示，当不考虑 k_a 时，平台由初始位置移动至正负极限位置(±2 mm)需要约±120 N 的力来平衡柔性导向机构的刚度力。然而，参考 k_a 对柔性导向机构刚度进行设计之后，平台仅需±24 N 便

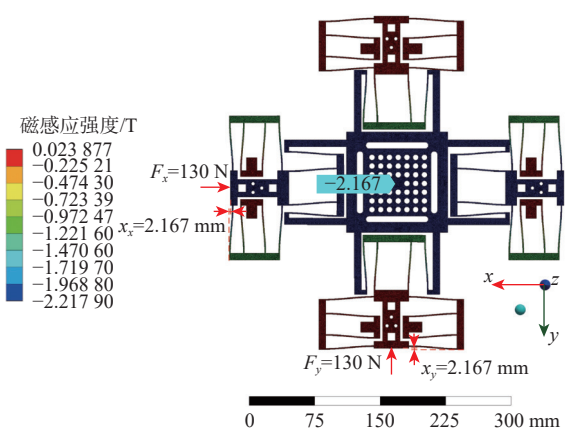


图 5　柔性导向机构静力学仿真结果

Fig. 5　Static simulation results of flexible mechanism

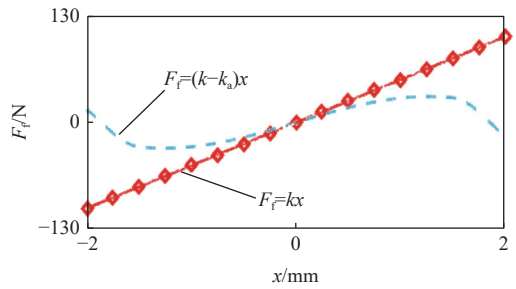


图 6　考虑 k_a 前后的总刚度力 F_f

Fig. 6　Total stiffness force F_f with and without k_a considered

可运动到极限位置，此时励磁电流减小，平台能效增加。总之，分析和仿真结果表明，本文所设计的微定位平台能够以较小的电流和较高的能效完成大行程范围的运动。

3　率相关 Prandtl-Ishlinskii 模型

磁阻驱动器材料内部磁场强度与磁感应强度之间非线性及大气隙下漏磁增加等因素导致驱动器控制电压与平台输出位移之间存在强烈的非对称磁滞特性。因此，要实现微定位平台的高速、高精控制，需要进行磁滞补偿。本文采用直接逆磁滞补偿的方法，即通过实验数据直接对平台磁滞曲线的逆曲线进行辨识，同时利用粒子群算法优化逆磁滞模型参数，直接设计逆磁滞补偿控制器，并将其置于前馈回路中补偿磁滞非线性。

磁滞补偿的原理如图 7 所示。$y_d(t)$ 为期望轨迹，磁阻驱动器输出为 $y(t)$，此时，逆磁滞模型将产生一个作用于大行程微定位平台的输入信号 $v(t)$。由于大行程微定位平台的模型由率相关磁滞子模型 H 和线性动力学子模型 G_p 串联组成，如果输入信号 $v(t)$ 由较低频率的信号组成，系统动力学 G_p 可以忽略不计，磁滞模型输出 $w(t)$ 近似于 $y(t)$。此时，如果逆磁滞模型是理想模型，输出 $y(t)$ 应遵循预期轨迹 $y_d(t)$，即 $y(t)$=$y_d(t)$。因此，逆模型的输入输出关系可以通过 $v(t)$ 和 $y(t)$ 直接获得[19]。

本文选择 P-I 模型来构建逆磁滞模型。考虑到磁阻驱动器磁滞非线性具有不对称性，本文采用了改进的 P-I 模型[20]，该模型由 Play 算子和多项式输入函数组成，定义为

$$v(t) = g(y_d(t)) + \int_0^r p(r)F_r[y_d](t)\mathrm{d}r \tag{7}$$

式中：$g(y_d(t)) = a_1y_d^3(t)+a_2y_d(t)$ 为以常数 a_1 和 a_2 作为系数的多项式；$p(r)$ 为根据实验数据计算的密度函数；$F_r[y_d](t)$ 为双侧齿隙算子，表示为

$$\begin{cases} F_r[y_d](0) = \max\{y_d(0)-r, \min\{y_d(0), 0\}\} \\ F_r[y_d](t) = \max\{y_d(t)-r, \min\{y_d(t)+r, F_r[y_d](t-T_s)\}\} \end{cases} \tag{8}$$

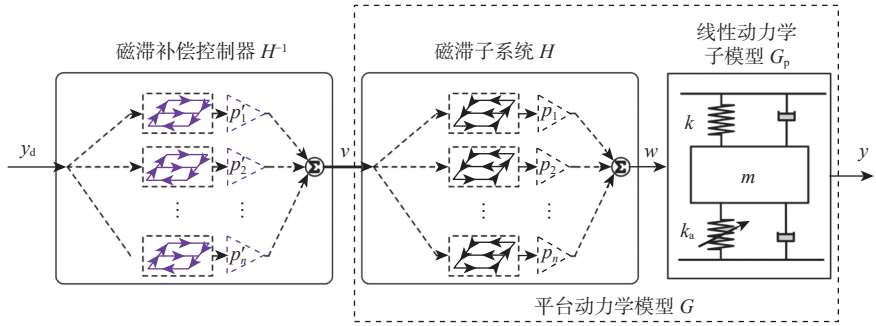


图 7 逆磁滞补偿原理

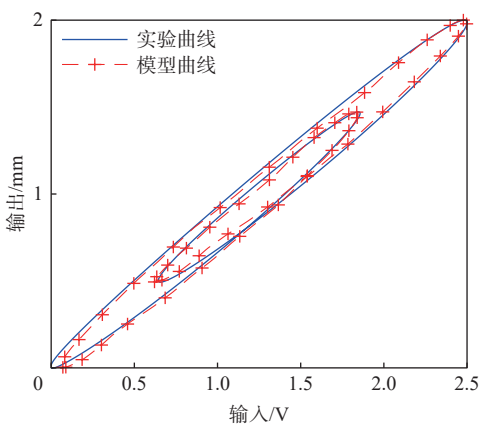
Fig. 7 Inverse hysteresis compensation principle

其中：r 为 Play 算子的阈值；T_s 为采样周期。

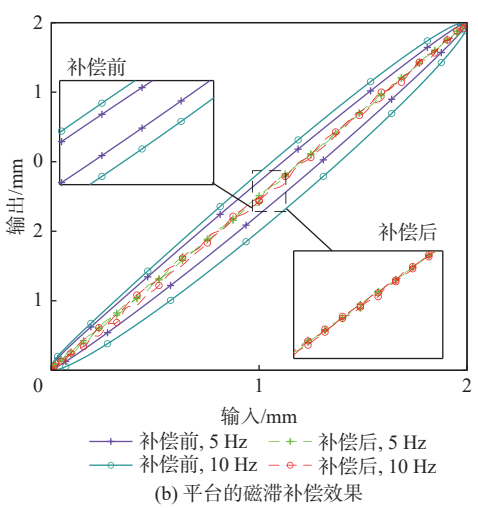
事实上，为在应用中更好地描述磁滞特性，方便后续控制器设计，通常采用有限个 Play 算子，即

$$y(t)=g(y_d(t))+\sum_{i=1}^{n}p_iF_r[y_d](t) \tag{9}$$

式中：p_i 为带有阈值 r_i 的 Play 算子的加权值；N 为 Play 算子的数目。该模型的参数可通过粒子群算法辨识获得。所得的逆磁滞模型可直接用于磁滞补偿，无须复杂的反演计算。图 8 为大行程微定位平台的磁滞拟合及开环磁滞补偿结果。



(a) 模型训练结果



(b) 平台的磁滞补偿效果

图 8 磁滞拟合与补偿效果

Fig. 8 Hysteresis fitting and compensation effect

4 分数阶反馈控制器设计

为设计平台的线性反馈控制器，需进行动力学辨识，获得其开环频率响应。本文选择频率为 0～600 rad/s 的 Chrip 信号为输入信号，通过调整该信号的偏置电压，将微定位平台的位移限制在 0，±1，±2 mm 等不同的位置进行系统动力学模型的辨识，利用傅里叶变换对采集到的输入输出信号进行处理，得到平台在不同工作位置的开环频率响应，利用傅里叶变换对采集到的输入输出信号进行处理，得到平台在 0 mm、±1 mm、±2 mm 处的开环频率响应 G_0、G_1 及 G_2，结果如图 9 所示。在电机刚度 k_a 和电机常数 k_m 的影响下，平台从初始位置 x=0 mm 移动到极限位置时，平台的线性化总刚度逐渐减小，谐振频率下降，驱动器在不同位置处的动力学呈现非一致的特性。

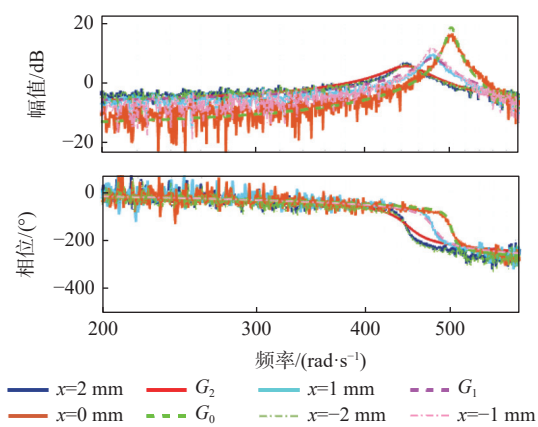


图 9 大行程微定位系统不同位置处开环频率响应

Fig. 9 Open-loop frequency response at different positions of large-stroke microposition system

为克服磁阻驱动柔顺微定位平台低阻尼谐振及动力学模型非一致性等复杂动力学问题，本文选择 FOPLPI 控制器作为平台的反馈控制器。同时，将基于 P-I 模型的率相关磁滞补偿控制器置于前馈回路中，这补偿了微定位平台的磁滞特性，实现了磁阻驱动微定位平台的高精度轨迹跟踪控制。图 10 为平台控制原理。

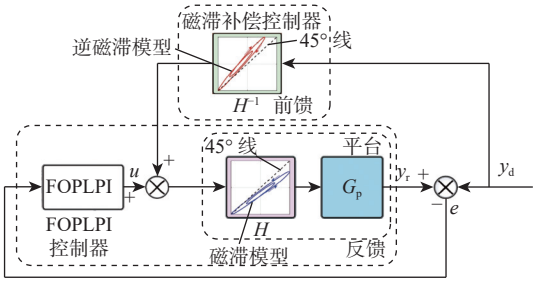


图 10　复合前馈控制原理

Fig. 10　Compound feedforward control principle

FOPLPI 控制器由分数阶相位超前（fractional order phase lead, FOPL）控制器和分数阶 PI（fractional order PI, FOPI）控制器组成。FOPI 控制器的结构可以描述为

$$G_{\mathrm{FOPI}}=\frac{K_{\mathrm{P}}s^{\tau}+K_{\mathrm{I}}}{s^{\lambda}} \tag{10}$$

式中：K_{P}、K_{I}、τ 和 λ 分别为控制器比例增益、积分增益、零点阶数和极点阶数。

由于 FOPI 控制器增加了极点阶数 λ 和零点阶数 τ，所以可以更灵活地调整系统的幅值裕度和相位裕度。与 FOPI 控制器相比，FOPL 控制器主要用于调整系统相位，其结构为

$$G_{\mathrm{FOPL}}=k_1\frac{(Ts)^{\alpha}+1}{(\gamma Ts)^{\beta}+1} \tag{11}$$

式中：k_1 为控制器增益，其由系统稳态速度要求决定；T、γ 决定了控制器转角频率；FOPL 及 FOPI 控制器零极点阶数 β、λ、α、τ 均在区间 [0,2] 之间。此时，式 (11) 表示有界输入有界输出（bounded input and bounded output, BIBO）稳定传递函数。

将 s=jω 代入式 (11) 可得

$$G_{\mathrm{FOPL}}(\mathrm{j}\omega)=k_1\left[(T\omega)^{\alpha}\cos\left(\frac{\alpha\pi}{2}\right)+(T\omega)^{\alpha}\mathrm{j}\sin\left(\frac{\alpha\pi}{2}\right)+1\right]\Big/\left[(\gamma T\omega)^{\beta}\cos\left(\frac{\beta\pi}{2}\right)+(\gamma T\omega)^{\beta}\mathrm{j}\sin\left(\frac{\beta\pi}{2}\right)+1\right] \tag{12}$$

通过计算式 (12) 相角可知，FOPL 控制器的相角在 ω=0 和∞时均为 0。由 $\omega^{-(\alpha+\beta)}(\gamma-1)=T^2\left(\gamma^2-\gamma\right)$ 可得控制器的相角导数 $\nabla\angle G_{\mathrm{FOPL}}(\mathrm{j}\omega)=0$。即当 $\gamma\neq1$ 时，$\nabla\angle G_{\mathrm{FOPL}}(\mathrm{j}\omega)=0$ 在区间 $(0,\infty)$ 内有唯一解：

$$\omega_{\mathrm{m}}=\left(\gamma T^2\right)^{-1/(\alpha+\beta)} \tag{13}$$

可见，与整数阶相位超前控制器相比，FOPLPI 控制器的幅值、相位及穿越频率均受到零点阶数 α 和极点阶数 β 的影响，这种改变赋予了 FOPLPI 控制器更强的调节能力，使其可以对目标系统的幅值和相位进行灵活的调整，从而增强了目标系统的稳定性、快速性及鲁棒性。

图 11 为不同零极点阶数对 FOPI 和 FOPL 控制器的幅值及相位的影响。可以看出，FOPI 和 FOPL 控制器的幅值曲线斜率和相位极值随着零点阶次 $\alpha(\tau)$ 和极点阶次 $\beta(\lambda)$ 的变化而变化，针对不同的控制对象，可以灵活选择。当系统对降低高频噪声没有严格要求，且需要足够多的相位裕度以提供充足的鲁棒性，可以选择 $\alpha>\beta(\tau>\lambda)$。如果系统已经足够稳定，需要对噪声进行抑制，则可在系统保持稳定性的前提下，谨慎选择 $\alpha<\beta(\tau<\lambda)$。当没有特殊要求时，选择 $\alpha=\beta(\tau=\lambda)$ 即可获得所需的控制性能，同时可以大大简化设计步骤。因此，本文 FOPI 控制器的零点阶数 τ 与极点阶数 λ 相等。

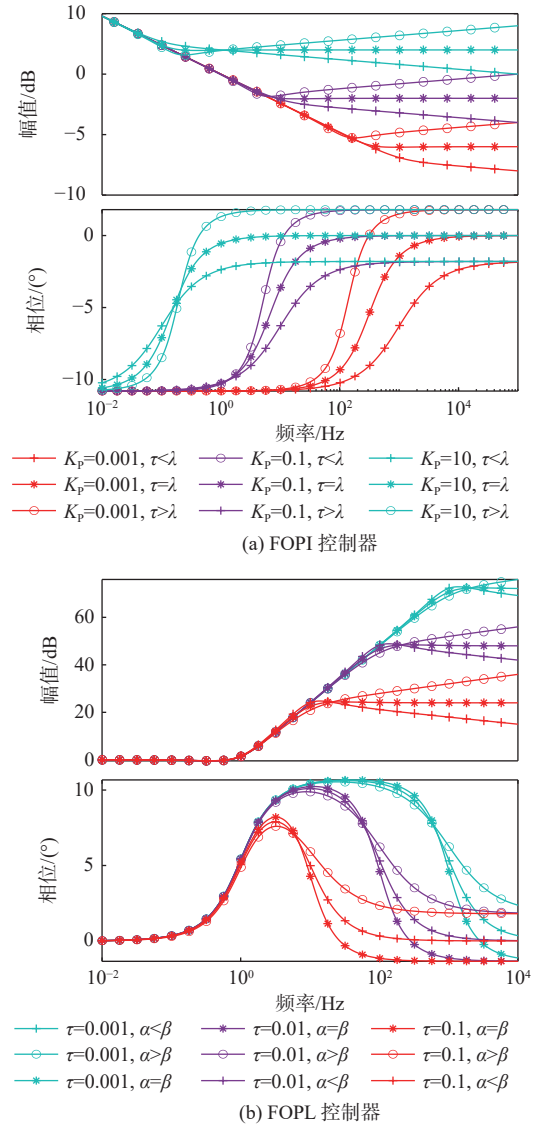


图 11　不同零极点阶数对控制器幅值及相位的影响

Fig. 11　Effects of different zero pole orders on amplitude and phase of controller

为消除平台的稳态误差，并对系统相位进行调整，FOPLPI 控制器由 FOPL 和 FOPI 控制器串联组成：

$$G_{\mathrm{fo}}=k_1\frac{(Ts)^{\alpha}+1}{(\gamma Ts)^{\beta}+1}\cdot\frac{K_{\mathrm{P}}s^{\tau}+K_{\mathrm{I}}}{s^{\lambda}} \tag{14}$$

对于上述控制器的参数，常以系统动力学模型为依据并转换为解带约束的非线性最优化问题[21]：

$$\begin{cases}\min\left|\left(\dfrac{\mathrm{d}(A_{\mathrm{rg}}[G_{\mathrm{fo}}(\mathrm{j}\omega)G(\mathrm{j}\omega)])}{\mathrm{d}\omega}\right)_{\omega=\omega_{\mathrm{c}}}\right| \\ \text{s.t.}\begin{cases}\left|G_{\mathrm{fo}}(\mathrm{j}\omega_{\mathrm{c}})G(\mathrm{j}\omega_{\mathrm{c}})\right|_{\mathrm{dB}}=1 \\ A_{\mathrm{rg}}[G_{\mathrm{fo}}(\mathrm{j}\omega_{\mathrm{c}})G(\mathrm{j}\omega_{\mathrm{c}})]=-\pi+\varphi_{\mathrm{m}}\end{cases}\end{cases} \tag{15}$$

式中：A_{rg}为复数的轴角主值。

由图 9 所示的平台开环频率响应及对开环稳定性的分析可知，当x_0=0 mm时，系统传递函数的极点均在复平面的左半平面，平台在中心位置处具有较强的稳定性，能够保证平台动力学模型的辨识精度及控制器参数计算的准确性和有效性。如图 9 所示，x=0 mm时，G_{p}=G_0，如下：

$$G_{\mathrm{p}}=1.893\times10^{3}/\left(s^{4}+776.3s^{3}+6.6\times10^{5}s^{2}+1.98\times10^{8}s+1\times10^{11}\right) \tag{16}$$

此时，基于粒子群算法及式 (14) 优化得到FOPLPI控制器的初始参数，并通过反复实验调整参数，最终获得控制器为

$$G_{\mathrm{fo}}=11.23\frac{(0.009\,0s)^{1.17}+1}{(0.012\,5s)^{1.36}+1}\cdot\frac{0.061s^{1.08}+41}{s^{1.08}} \tag{17}$$

图 12 为平台添加控制器后的开环频率响应。FOPLPI控制器实现了对控制器幅值及相位的灵活控制，摆脱了平台复杂动力学特性对控制器设计的影响，尽管不同位置的动力学模型有一定差距，但系统开环传递函数的幅值裕度和相位裕度分别处于[4.2, 8.31] dB和[31°, 37°]这2个稳态区间，从而保证了平台的稳定性。图 13 为系统使用FOPLPI控制器时的灵敏度函数及最终的闭环频率响应，由图 13(b) 可知，闭环系统幅频曲线与−3 dB刻度线相

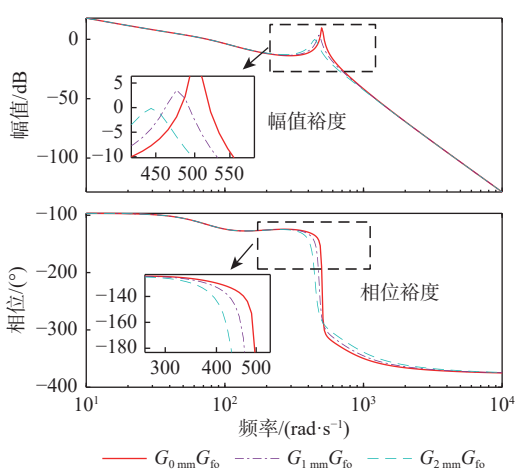


图 12 添加FOPLPI控制器后系统开环频率响应

Fig. 12 Open-loop frequency response of system after controller introduction

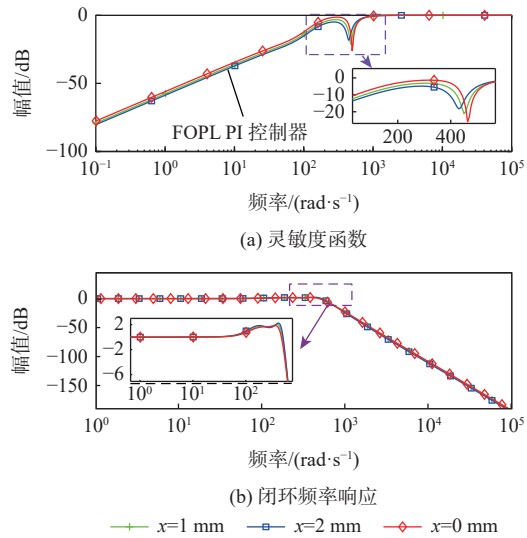


图 13 系统灵敏度函数及闭环频率响应

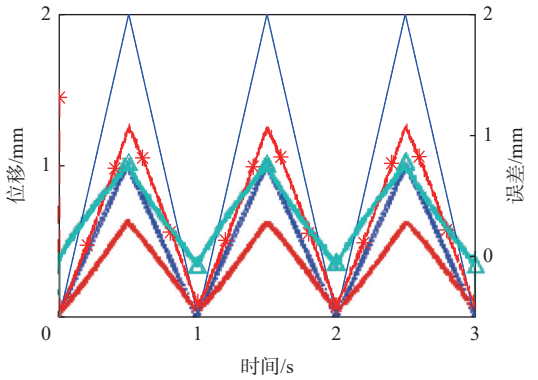
Fig. 13 System sensitivity function and closed-loop frequency response

交定义的有效带宽在560 rad/s附近，图 13(a) 中的灵敏度函数同样展现了控制器在低频段的抗干扰能力。

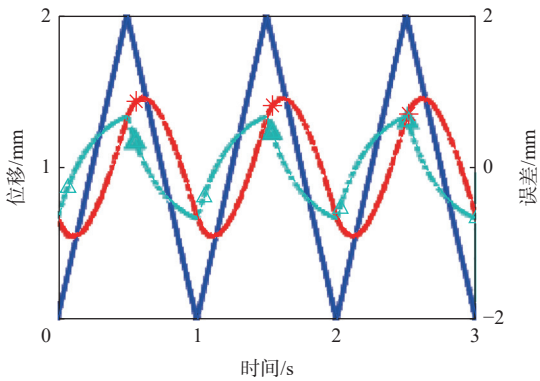
5 轨迹跟踪实验

为验证本文所设计的FOPLPI控制器的有效性，分别使用PID控制器、FOPLPI控制器+磁滞前馈补偿、传递函数逆模型前馈控制、FOPLPI控制器+磁滞前馈补偿4种控制方法，对频率为1 Hz，幅值为2.5 V(对应行程为2 mm)的三角波信号进行跟踪对比实验，结果如图 14 所示。

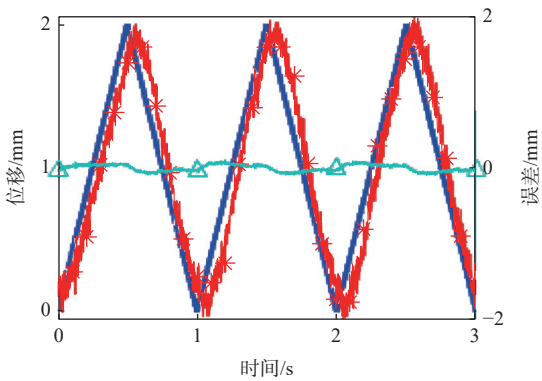
由图 14 可知，由于本文中大行程微定位平台复杂的动力学特性，PID控制器+磁滞前馈补偿的跟踪误差最大值高达0.78 mm，均方根误差为0.41 mm，无法保证平台定位精度及速度。对于可以提供一定相位裕量的整数阶相位超前PI控制器，尽管其可以对平台处于某个固定位置的动力学特性进行相位补偿，然而由于驱动器变化的电机常数k_{m}及刚度k_{a}的影响，所以无法对整个行程内的动力学特性进行调整，其最大跟踪误差如图 14(b) 所示，为0.68 mm，均方根误差为0.43 mm。由图 14(b) 可知，磁阻驱动器的磁滞非线性对平台定位精度误差影响较大，仅使用反馈控制器无法对其进行有效补偿。如图 14(c) 所示，当加入动力学前馈控制器后，受辨识精度及模型不确定性的影响，所得动力学模型误差较大，前馈控制的稳定性无法保证，此时其最大误差及均方根误差分别为0.95 mm和0.11 mm。相较于本文提出的FOPLPI控制器+磁滞前馈补偿复合控制方法，其他3种控制方法的均方



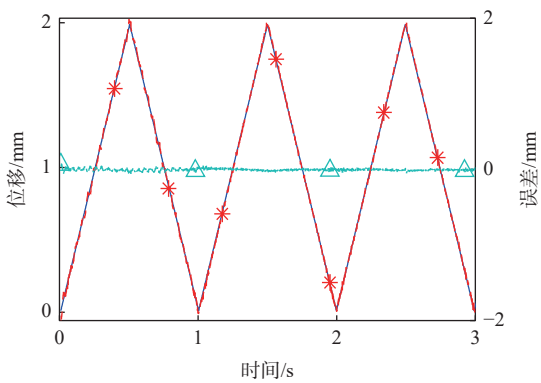
(a) PID 控制器+磁滞前馈补偿



(b) 整数阶相位超前 PI 控制器



(c) 整数阶相位超前 PI 控制器+动力学前馈控制



(d) FOPLPI 控制器+磁滞前馈补偿

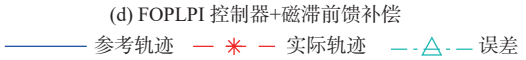
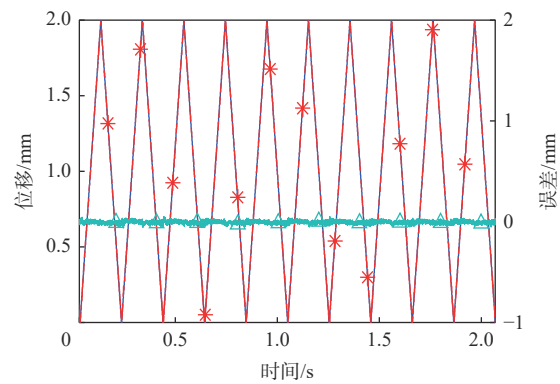


图 14　不同控制器的轨迹跟踪结果对比

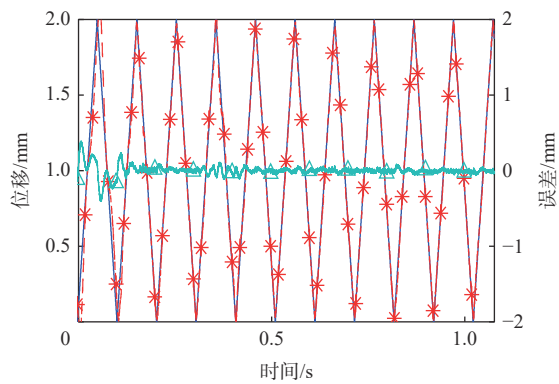
Fig. 14　Comparison of trajectory tracking results of different controllers

根误差分别为本文方法均方根误差(0.013 mm)的 31.5 倍、33.1 倍、8.46 倍。因此，FOPLPI 控制器通过灵活调整稳态裕度，保证了系统在全行程范围内的稳定性，为系统提供了足够的跟踪精度及跟踪速度。

为进一步验证 FOPLPI 控制器在大行程微定位平台上的强鲁棒性及高精度特性，对幅值为 2.5 V (2 mm)，频率分别为 5 Hz、10 Hz 的三角波信号进行跟踪，轨迹跟踪结果如图 15 所示。可知，在轨迹跟踪过程中，误差均保持稳定，在跟踪频率为 5 Hz 和 10 Hz 时，其最大误差分别为 0.055 mm 和 0.39 mm，均方根误差分别为 0.013 mm 和 0.072 mm。在跟踪 10 Hz 三角波信号时，其最大误差在一个周期后从 0.23 mm 迅速收敛至 0.053 mm，仅为整体行程的 2.65%，均方根误差收敛至 0.017 mm，由此证明了本文方法的稳定性及鲁棒性。



(a) 5 Hz



(b) 10 Hz

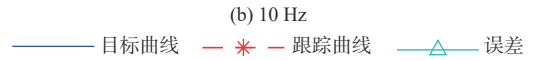


图 15　不同频率三角波轨迹跟踪结果

Fig. 15　Trajectory tracking results of triangular waves with different frequencies

6　结　论

1）本文设计了一种基于麦克斯韦磁阻驱动的大行程二自由度柔顺微定位平台，并充分利用了磁阻驱动器的非线性负刚度来部分补偿柔性导向机构的弹性恢复力，能够有效提高其运动行程和能量

传递效率，使得平台在±2 mm行程范围内的所需推力从±120 N减少至±24 N。

2）利用FOPLPI控制+磁滞前馈补偿完成了对系统开环频率特性的灵活调整，解决了微定位系统低阻尼谐振、谐振点漂移及动力学模型不一致等复杂动力学问题，相较于PID控制+磁滞前馈补偿、整数阶相位超前PI控制、整数阶相位超前PI控制+传递函数逆模型前馈控制等方法，跟踪均方根误差分别减少了31.5倍、33.1倍和8.46倍，实现了大行程微纳米定位的高速高精控制。

参考文献（References）

[1] LEE C, LEE J W, RYU S G, et al. Optimum design of a large area, flexure based *XYθ* mask alignment stage for a 12-inch wafer using grey relation analysis[J]. Robotics and Computer-Integrated Manufacturing, 2019, 58: 109-119.

[2] 刘庆玲. 柔性对称微位移放大机构性能分析方法的研究[J]. 工程设计学报, 2013, 20(4): 344-347.
LIU Q L. Study on the performance analysis methods of the compliant symmetric micro-displacement magnifying mechanism[J]. Chinese Journal of Engineering Design, 2013, 20(4): 344-347(in Chinese).

[3] 王念峰, 张志远, 张宪民, 等. 三种两自由度柔顺精密定位平台的性能对比与分析[J]. 机械工程学报, 2018, 54(13): 102-109.
WANG N F, ZHANG Z Y, ZHANG X M, et al. Performance comparison and analysis of three 2-DOF compliant precision positioning stages[J]. Journal of Mechanical Engineering, 2018, 54(13): 102-109(in Chinese).

[4] 李玄, 周双武, 路松, 等. 基于二级杠杆机构的二自由度微定位平台设计与分析[J]. 工程设计学报, 2020, 27(4): 533-540.
LI X, ZHOU S W, LU S, et al. Design and analysis of two-DOF micro-positioning platform based on two-level lever mechanism[J]. Chinese Journal of Engineering Design, 2020, 27(4): 533-540(in Chinese).

[5] KATALENIC A, DE BOEIJ J, BUTLER H, et al. Linearization of a current-driven reluctance actuator with hysteresis compensation[J]. Mechatronics, 2013, 23(2): 163-171.

[6] LU X D, TRUMPER D L. Ultrafast tool servos for diamond turning[J]. CIRP Annals, 2005, 54(1): 383-388.

[7] HIEMSTRA D B, PARMAR G, AWTAR S. Performance tradeoffs posed by moving magnet actuators in flexure-based nanopositioning[J]. IEEE/ASME Transactions on Mechatronics, 2014, 19(1): 201-212.

[8] ITO S, TROPPMAIR S, LINDNER B, et al. Long-range fast nanopositioner using nonlinearities of hybrid reluctance actuator for energy efficiency[J]. IEEE Transactions on Industrial Electronics, 2019, 66(4): 3051-3059.

[9] ZHU Z W, CHEN L, TO S. A novel direct drive electromagnetic *XY* nanopositioning stage[J]. CIRP Annals, 2021, 70(1): 415-418.

[10] FANG Y N, PU X, TO S, et al. Normal-stressed electromagnetic triaxial fast tool servo for microcutting[J]. IEEE Transactions on Industrial Electronics, 2023, 70(7): 7131-7140.

[11] NIE Y H, FANG F Z, ZHANG X D. System design of Maxwell force driving fast tool servos based on model analysis[J]. The International Journal of Advanced Manufacturing Technology, 2014, 72(1): 25-32.

[12] FANG F Z, NIE Y H, ZHANG X D. Design of fast tool servo system based on magnetic field analysis[J]. Nanotechnology and Precision Engineering, 2011, 9(6): 539-544.

[13] NASER M F M, AL-HDAIBAT B, GUMAH G, et al. On the consistency of local fractional semilinear Duhem model[J]. International Journal of Dynamics and Control, 2020, 8(3): 723-729.

[14] RAKOTONDRABE M. Bouc-Wen modeling and inverse multiplicative structure to compensate hysteresis nonlinearity in piezoelectric actuators[J]. IEEE Transactions on Automation Science and Engineering, 2011, 8(2): 428-431.

[15] AL JANAIDEH M, RAKHEJA S, SU C Y. An analytical generalized Prandtl-Ishlinskii model inversion for hysteresis compensation in micropositioning control[J]. IEEE/ASME Transactions on Mechatronics, 2011, 16(4): 734-744.

[16] CSENCSICS E, SCHLARP J, SCHITTER G. High-performance hybrid-reluctance-force-based tip/tilt system: Design, control, and evaluation[J]. IEEE/ASME Transactions on Mechatronics, 2018, 23(5): 2494-2502.

[17] XU Q S. New flexure parallel-kinematic micropositioning system with large workspace[J]. IEEE Transactions on Robotics, 2012, 28(2): 478-491.

[18] LAI L J, GU G Y, ZHU L M. Design and control of a decoupled two degree of freedom translational parallel micro-positioning stage[J]. Review of Scientific Instruments, 2012, 83(4): 045105.

[19] JANOCHA H, KUHNEN K. Real-time compensation of hysteresis and creep in piezoelectric actuators[J]. Sensors and Actuators A: Physical, 2000, 79(2): 83-89.

[20] GU G Y, YANG M J, ZHU L M. Real-time inverse hysteresis compensation of piezoelectric actuators with a modified Prandtl-Ishlinskii model[J]. Review of Scientific Instruments, 2012, 83(6): 065106.

[21] 薛定宇. 分数阶微积分学与分数阶控制[M]. 北京: 科学出版社, 2018.
XUE D Y. Fractional calculus and fractional-order control[M]. Beijing: Science Press, 2018(in Chinese).

Large-stroke microposition stage driven by reluctance actuator and its trajectory tracking control

ZHANG Xu[1]，LU Xiaowei[2]，LAI Leijie[1,*]

(1. School of Mechanical and Automotive Engineering，Shanghai University of Engineering Science，Shanghai 201620，China；
2. Huali Technology Co.，Ltd.，Hangzhou 310023，China)

Abstract：Maxwell reluctance actuators overcome the problems of small strokes of traditional piezoelectric actuators and low thrust density and efficiency of voice coil motors and have great application potential in large-stroke and high-speed micro/nanoposition. In this paper, a large-stroke two-degree-of-freedom flexible microposition stage driven by a Maxwell reluctance actuator was designed, and its high-performance trajectory tracking control was carried out. The position stage was composed of two Maxwell reluctance actuators with permanent magnet bias and a two-degree-of-freedom decoupled and parallel flexible guiding mechanism. The stage made full use of the nonlinear negative stiffness of the reluctance actuator to partially compensate for the elastic restoring force of the flexible mechanism, which could effectively improve its motion range and energy transfer efficiency and reduce the required thrust of the stage within the range of ± 2 mm from ±120 N to ±24 N. In terms of trajectory tracking control, firstly, in order to compensate for the hysteresis nonlinearity caused by the impulse of soft magnetic material of the actuator, a rate-dependent hysteresis compensator was constructed by using the inverse Prandtl-ishlinskii hysteresis model and placed in the feedforward loop. In addition, in order to solve the low damping resonance and the inconsistency of the dynamic model of the stage, a PI feedback controller with a fractional order phase-lead link was designed to flexibly adjust the open-loop frequency characteristics of the system, realize the high-precision trajectory tracking control of the stage, and effectively reduce the tracking error of the system. The root mean square errors obtained by tracking the trajectory of triangular wave signals with amplitude of 2 mm and frequencies of 1 Hz and 10 Hz are 0.013 mm and 0.017 mm, respectively.

Keywords：micro/nanoposition；Maxwell reluctance actuator；flexible mechanism；hysteresis；fractional order control；trajectory tracking control

Received：2022-08-10；**Accepted**：2022-11-12；**Published Online**：2022-11-25 13：53

URL：link.cnki.net/urlid/11.2625.V.20221124.1443.002

Foundation items：National Natural Science Foundation of China (U2013211); Natural Science Foundation of Shanghai (21ZR1426000); State Key Laboratory of Mechanical System and Vibration Subjects Funded Projects (MSV202210)

* **Corresponding author**. E-mail：lailj@sues.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0707

多策略融合改进的飞蛾火焰优化算法

何加文，许贤泽*，高波
(武汉大学 电子信息学院，武汉 430072)

摘 要：针对飞蛾火焰优化算法容易出现局部最优解、接近全局最优时开发能力不足等问题，提出一种多策略融合改进的飞蛾火焰优化(RGMFO)算法。在每次迭代开始时，使用随机反向学习策略获得高质量飞蛾种群；利用高斯变异将较差的火焰个体替换为优秀个体；使用阿基米德螺线、权重因子和围绕最优火焰飞行3种方式改进飞蛾更新机制。为验证所提算法的有效性，与11个不同类型的基准函数进行寻优测试，基准函数实验结果与秩和检验表明：相较于其他优化算法及其他MFO改进算法，所提算法具有更好的全局搜索能力与更高的寻优精度。将所提算法应用于减速器设计和槽形舱壁设计的实际工程场景中，以进一步验证算法的实用性与可行性。

关 键 词：飞蛾火焰优化算法；多策略融合；随机反向学习；高斯变异；智能优化
中图分类号：TP301
文献标志码：A **文章编号**：1001-5965(2024)09-2862-10

飞蛾火焰优化(moth-flame optimization, MFO)算法是一种新型启发式群体优化智能算法[1]，其通过模拟自然界中飞蛾绕火焰沿螺线轨迹运动的行为来求解目标函数。该算法具有结构参数简单、计算成本小、计算速度快等优点，其在求解具有约束和未知搜索空间的复杂问题时，有较好的全局寻优能力和较高的寻优精度[2]，现已成功应用于许多科研场景中。

当前，国内外学者为追求更好的寻优效果与计算精度，对MFO算法做了诸多改进，且部分已应用于实际场景中。文献[3]针对MFO算法存在的早熟收敛和易陷入局部最优的问题，提出了一种基于多策略扰动机制的MFO算法，该算法通过莱维飞行和布朗运动的随机选择对飞蛾进行重新定位，以克服种群后期多样性的严重下降，可应用在乙烯聚合过程中的动力学参数估计上；文献[4]提出了一种Kent混沌动态惯性权值的改善MFO算法，引入Kent混沌映射搜索策略帮助当前最优解跳出局部最优，采用基于适应度值和迭代次数的动态惯性权值策略来平衡算法的开发和探索能力，以进一步提升算法性能；文献[5]针对传统群体智能优化算法在复杂环境下求解无人机突防过程中路径搜索能力不足、易陷入局部最优、搜索时间长等问题，在MFO算法中引入交叉算子、高斯变异算子及自适应权重，提出了一种基于改进飞蛾扑火优化算法的无人机智能突防方法；文献[6]针对MFO算法在求解最优化问题时存在的寻优精度低等问题，提出一种基于混沌初始化和高斯变异的改进飞蛾火焰优化算法(moth-flame optimization algorithm based on chaotic initialization and Gaussian mutation, CGMFO)，使飞蛾更均匀地分布于搜索空间，从而提升了算法对未知领域的探索能力；文献[7]为增强MFO算法的全局搜索能力、加快收敛速度，提出了一种定义探索与开发之间混合阶段的改进MFO算法(improved moth-flame optimization, IMFO)，该算法有很好的搜索能力与较快的收敛速度；文献[8]根据

收稿日期：2022-08-11；**录用日期**：2022-09-02；**网络出版时间**：2022-11-08 10:42
网络出版地址：link.cnki.net/urlid/11.2625.V.20221108.0859.002
基金项目：国家自然科学基金(51975422)
*通信作者. E-mail：xxz@whu.edu.cn

引用格式：何加文，许贤泽，高波. 多策略融合改进的飞蛾火焰优化算法[J]. 北京航空航天大学学报，2024，50(9)：2862-2871.
HE J W，XU X Z，GAO B. Improved moth-flame optimization algorithm with multi-strategy integration[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2862-2871 (in Chinese).

MFO 算法缺乏高质量火焰、飞蛾种群缺乏多样性、算法全局搜索能力不足及由于火焰减少机制和飞蛾的简单横向飞行方式的问题, 提出了一种改进火焰产生机制和局部搜索机制组成的改进飞蛾火焰优化 (moth-flame optimization based on opposition-based learning, differential evolution, shuffled frog leaping algorithm and death mechanism, ODSFMFO) 算法, 这不仅使算法更容易跳出局部最优, 而且进一步提高了收敛速度。

为提升 MFO 算法的全局搜索能力与计算精度, 本文提出了一种多策略融合改进的飞蛾火焰优化 (moth-flame optimization with multi-strategy integration including random opposition-based learning and Gaussian mutation, RGMFO) 算法。在测试函数上与遗传算法 (genetic algorithm, GA)、差分进化 (differential evolution, DE) 算法、粒子群 (particle swarm optimization, PSO) 算法[9]、正弦余弦算法 (sine cosine algorithm, SCA)[10]、灰狼优化 (grey wolf optimizer, GWO) 算法[11] 和 MFO 算法及其改进算法 (CGMFO、IMFO、ODSFMFO) 进行对比实验, 结果表明, RGMFO 算法有着更好的全局搜索能力及寻优精度。同时, 将 RGMFO 算法应用于减速器设计和槽形舱壁设计问题中, 能得到最优结果。

1　MFO 算法

在 MFO 算法中, 每只飞蛾在解空间中的位置表示一个候选解, 飞蛾在多维空间中绕火焰不断迭代寻优, 最终得到全局最优解。其中, 飞蛾的位置集合用 n 行 d 列的矩阵 $\boldsymbol{M}$ 表示, n 为飞蛾数量, d 为解空间的维度。每一只飞蛾都有一个对应的适应度值。火焰位置矩阵 $\boldsymbol{F}$ 与飞蛾类似, 有 N 行 d 列, 其中, N 为火焰数量, 火焰数量随着迭代过程按式 (1) 逐渐减小。

$$N=\mathrm{round}\left(N_0-k\frac{N_0-1}{K}\right) \tag{1}$$

式中: round 为取整函数; k 为当前迭代次数; N_0=n 为最大火焰个数; K 为最大迭代次数。

飞蛾位置更新的机制使用的是对数螺线, MFO 算法的位置更新定义:

$$S(\boldsymbol{M}_i,\boldsymbol{F}_j)=\boldsymbol{D}_i\mathrm{e}^{bt}\cos(2\pi t)+\boldsymbol{F}_j \tag{2}$$

式中: $\boldsymbol{M}_i$ 和 $\boldsymbol{F}_j$ 分别为第 i 个飞蛾与第 j 个火焰的位置; $\boldsymbol{D}_i$=pos($\boldsymbol{F}_j$−$\boldsymbol{M}_i$), 其中函数 pos() 表示对向量中各元素值取非负数; b 为定义对数螺线形状的常量; t 为 [r,1] 中的随机数, r=−1−k/K。

MFO 算法在探索和开发上有很好的平衡性, 但由于探索范围较宽泛等, 所以开发精细程度不够, 并且由于火焰减少机制的存在, 导致 MFO 算法过早收敛, 易陷入局部最优。因此, 需要对 MFO 算法进行有针对性的改进, 以得到更好的寻优效果。

2　RGMFO 算法

鉴于上述分析, 本文从 3 个方面对 MFO 算法进行改进。

2.1　随机反向学习策略

为提高搜索能力, 提升种群多样性水平, Long 等在反向学习[12] 的基础上提出了随机反向学习 (random opposition-based learning, ROL) 策略[13], 该方法增强了反向解的随机性, 扩大了搜索范围, 进一步提高了种群的多样性, 能够有效地避免局部最优, 其计算公式如下:

$$\tilde{x}_i=a_i+b_i-q_1x_i \qquad i=1,\cdots,n \tag{3}$$

式中: a_i 和 b_i 分别为解空间在第 i 维的下界和上界; x_i 和 $\tilde{x}_i$ 分别为当前解和反向解在第 i 维的位置; q_1 为 [0,1] 之间的随机数。

每次迭代开始时, 都使用 ROL 策略以获得更优质的飞蛾种群, 以提高算法的全局探索能力, 跳出局部最优。

2.2　高斯变异

在 MFO 算法中, 飞蛾围绕火焰沿螺线更新其坐标, 虽然该区域范围较大, 但对火焰周围的探索不足, 目前对最优解的利用能力有限。因此, 每次火焰位置更新完成后, 都要对火焰使用高斯变异[14], 搜索火焰附近的局部区域, 对最优值附近做进一步开发, 从而提高算法的收敛速度及计算精度。

计算变异后火焰的适应度值, 与原适应度值进行比较, 若新的适应度值较为优秀, 则使用新位置替换原位置。

2.3　飞蛾更新机制的改进

为增强算法在探索和开发之间的匹配与衔接, 引入了能量密度 E[15], 根据 E 值使用的 3 种不同方式对飞蛾位置进行更新。E 的变化范围随着迭代次数的增加而逐渐缩小:

$$E=2E_0\left(1-\frac{k}{K}\right) \tag{4}$$

式中: E 为飞蛾的能量; E_0 为能量初始状态, 每次迭代时在区间 (−1,1) 内随机变化。

设置 2 个阈值 e_1 和 e_2 (0<e_1<e_2<2), 若|E|≥e_2, 则采用方式 1 进行更新; 若 e_1≤|E|<e_2, 则采用方式 2 进行更新; 若|E|<e_1, 则采用方式 3 进行更新。

E 随迭代次数变化的示意如图 1 所示。其中, 折线图表示 E 值的变化规律, 散点图则反映了 E 值

落在不同区域而采用相应方法的分布。根据 E 值的大小采用阿基米德螺线(方式 1)、围绕最优火焰飞行(方式 2)和增加权重因子 ω(方式 3)这 3 种方式更新飞蛾有利于保持开发与探索的平衡。

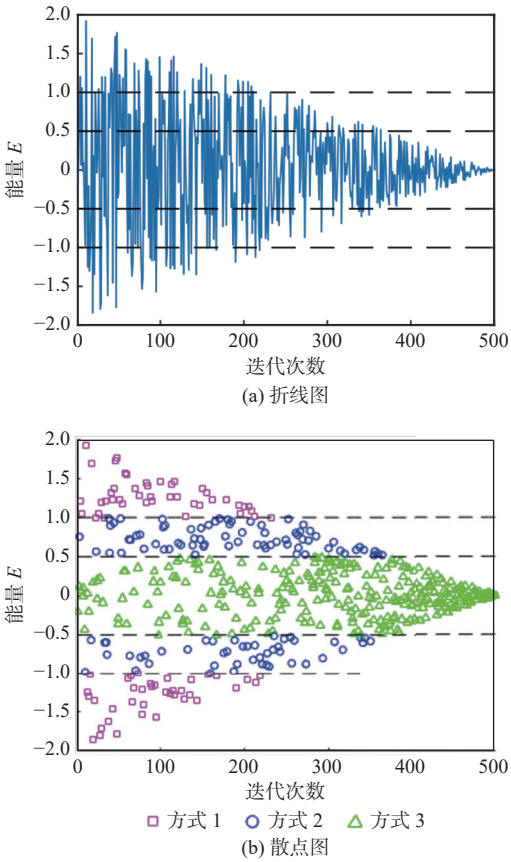


图 1 E 值折线图及散点图(500 次迭代)

Fig. 1 E-value line chart and scatter chart (500 iterations)

2.3.1 阿基米德螺线

在 MFO 算法中,飞蛾的飞行轨迹对算法的搜索能力有着很大的影响,对数螺线的轨迹集中在中心点附近,这限制了飞蛾的搜索范围。因此,该方式侧重于对整个解空间进行探索,并使用阿基米德螺线来扩大飞蛾的搜索范围,这有利于提高算法的探索能力。因此,设定飞蛾位置的更新公式为

$$S(\boldsymbol{M}_i,\boldsymbol{F}_j)=\boldsymbol{D}_i\mathrm{e}^{bt}(rt)\sin(2\pi\beta(t-r))+\boldsymbol{F}_j \tag{5}$$

式中:β=10。

2.3.2 增加权重因子 ω

为避免算法的过度开发、陷入局部最优解,本文增加一种过渡方式,既不丢失种群多样性,又能使飞蛾更快偏向全局最优解。该方式是对全局探索与局部开发的整合,飞蛾根据当前火焰及全局最优位置使用式 (6) 来更新位置:

$$S(\boldsymbol{M}_i,\boldsymbol{F}_j)=\boldsymbol{D}_i\mathrm{e}^{bt}\cos(2\pi t)+\omega\boldsymbol{F}_j+(1-\omega)\boldsymbol{M}_{\text{best}} \tag{6}$$

定义权重因子 ω 如下:

$$\omega=\begin{cases}\left|\dfrac{f(\boldsymbol{M}_{\text{best}})}{f(\boldsymbol{M}_i^k)}\right| & f(\boldsymbol{M}_{\text{best}})\geqslant 0且f(\boldsymbol{M}_i^k)>0\\ 0.5 & f(\boldsymbol{M}_{\text{best}})<0且f(\boldsymbol{M}_i^k)\geqslant 0\\ \left|\dfrac{f(\boldsymbol{M}_i^k)}{f(\boldsymbol{M}_{\text{best}})}\right| & f(\boldsymbol{M}_{\text{best}})<0且f(\boldsymbol{M}_i^k)<0\end{cases} \tag{7}$$

式中,$\boldsymbol{M}_{\text{best}}$ 为当前全局最优位置;$\boldsymbol{M}_i^k$ 为第 i 只飞蛾在第 k 次迭代时的位置,$f(\boldsymbol{M}_{\text{best}})$ 和 $f(\boldsymbol{M}_i^k)$ 分别为其对应的适应度函数值。因为只考虑最小化的问题,所以,ω 处于 [0,1] 内。若当前飞蛾位于最佳位置,即 $\boldsymbol{M}_i^k=\boldsymbol{M}_{\text{best}}$ 时,则 ω=1。反之,如果飞蛾位于很不好的位置,即 $f(\boldsymbol{M}_i^k)\gg f(\boldsymbol{M}_{\text{best}})$ 时,则 ω 趋近 0。

增加权重因子 ω,在保证探索能力的前提下具有一定的开发能力,从而实现了探索与开发的折中与过渡。

2.3.3 围绕最优火焰飞行

为提升算法的收敛水平,提高计算精度,围绕最优位置进行飞行。迭代时,在火焰中选择适应度最优秀的 σ 个火焰,使所有飞蛾依次绕这 σ 个火焰飞行。其中,$\sigma=\lceil q_2\cdot N\rceil$,$q_2$ 表示 (0,1) 中的随机值,N 为当前迭代轮次的火焰数量,$\lceil\rceil$表示向上取整。

飞蛾绕火焰飞行的公式为

$$S(\boldsymbol{M}_i,\boldsymbol{F}_\delta)=\boldsymbol{D}_i\mathrm{e}^{bt}\cos(2\pi t)+\boldsymbol{F}_\delta \tag{8}$$

第 i 只飞蛾绕第 δ 个火焰进行飞行,其中:

$$\delta=\begin{cases}i & i\leqslant\sigma\\ \mathrm{mod}(i,\sigma) & i>\sigma且\mathrm{mod}(i,\sigma)\neq 0\\ \sigma & i>\sigma且\mathrm{mod}(i,\sigma)=0\end{cases} \tag{9}$$

飞蛾围绕少数最优解飞行,大大增强了对最优位置附近的开发,既提高了开发效率,又不完全丢失其他火焰个体,还能够保持在具备开发精度与避免产生局部最优上的平衡。

2.4 RGMFO 算法步骤

RGMFO 算法流程如图 2 所示,其步骤如下:

步骤 1 初始化参数。飞蛾数量 n、维度 d、最大迭代次数 K 和螺线形状参数 b,初始化飞蛾种群。

步骤 2 使用 ROL 策略获得优质飞蛾种群,当 k=1 时,将优质种群按适应度值升序排列;当 $k\neq$1 时,将本次与上次迭代的优质飞蛾种群按适应度值升序排列。

步骤 3 通过式 (1) 计算火焰数量,将排序后优秀的 N 个位置赋给火焰,对更新后的火焰使用高斯变异。

步骤 4 通过式 (4) 计算 E 值,根据 E 值大小选用 3 种方式中的 1 种更新飞蛾位置。

步骤 5 若满足迭代次数,则算法结束;否则,返回步骤 2。

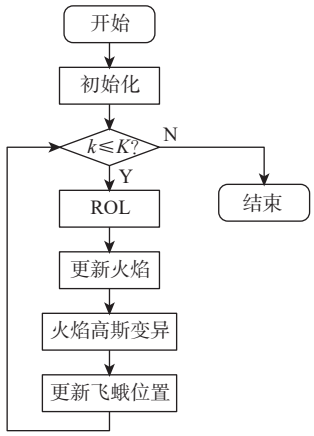


图 2　RGMFO 算法流程

Fig. 2　RGMFO algorithm flow

3　算法性能测试与分析

3.1　实验设计与测试函数选取

实验采用 Intel(R) Core(TM) i5-8250U CPU @ 1.60 GHz 1.80 GHz, 内存 8.00 GB, Windows10 系统和 MATLAB R2019a 的环境进行测试。选用 MFO 算法、其他 5 个优化算法 GWO、SCA、PSO、DE、GA 及 3 个 MFO 改进算法 CGMFO、IMFO、ODSF-MFO 与本文算法进行对比实验。

实验中, 所有算法的种群规模为 30, 最大迭代次数为 500 次, 各算法参数设置如表 1 所示, 所有算法的参数均经过实验验证, 均具有最好的优化效果。

为验证 RGMFO 算法在寻优能力与计算精度上的优越性, 利用 3 种类型的基准函数进行优化对比测试实验, 测试函数的详细介绍如表 2 所示, 其中 d 表示测试函数的维度。

表 1　各算法参数设置

Table 1　Parameter setting of each algorithms

算法	参数
RGMFO	b=1, e_1=0.5, e_2=1, β=10
IMFO	b=1, A=0.5, B=0.6, c_{rmin}=0.55, c_{rmax}=0.65 σ_{min}=0.1, σ_{max}=0.15
CGMFO	b=2, β=8
ODSFMFO	m=6, γ=5, P_c=0.5, a_{alpha}=1, b_{beta}=1.5, l=10, b=1
MFO	b=1
GWO	a从2线性递减至0, r_1、r_2∈[0,1]
SCA	a=2, r_1从a线性递减至0
PSO	$V_{\max}$=8, $V_{\min}$=−8, w=0.2, c_1=3, c_2=3
DE	p_c=0.5
GA	p_c=0.7, p_m=0.1

表 2　基准测试函数

Table 2　Benchmark functions

函数类型	基准测试函数	维度	搜索区间	最优值
高维单峰	$F_1(x)=\sum_{i=1}^{d}x_i^2$	30	$[-100,100]^d$	0
	$F_2(x)=\sum_{i=1}^{d}\lvert x_i\rvert+\prod_{i=1}^{d}\lvert x_i\rvert$	30	$[-10,10]^d$	0
	$F_3(x)=\sum_{i=1}^{d}\left(\sum_{j=1}^{i}x_j\right)^2$	30	$[-100,100]^d$	0
	$F_4(x)=\max_i\{\lvert x_i\rvert,1\leqslant i\leqslant d\}$	30	$[-100,100]^d$	0
	$F_5(x)=\sum_{i=1}^{d-1}[100(x_{i+1}-x_i)^2+(x_i-1)^2]$	30	$[-30,30]^d$	0
	$F_6(x)=\sum_{i=1}^{d}i\cdot x_i^4+\text{random}[0,1)$	30	$[-1.28,1.28]^d$	0
高维多峰	$F_7(x)=\sum_{i=1}^{d}[x_i^2-10\cos(2\pi x_i+10)]$	30	$[-5.12,5.12]^d$	0
	$F_8(x)=-20\exp\left(-0.2\sqrt{\frac{1}{d}\sum_{i=1}^{d}x_i^2}\right)-\exp\left(\frac{1}{d}\sum_{i=1}^{d}\cos(2\pi x_i)\right)+20+\text{e}$	30	$[-32,32]^d$	0
	$F_9(x)=\frac{1}{4\,000}\sum_{i=1}^{d}x_i^2-\prod_{i=1}^{d}\cos\left(\frac{x_i}{\sqrt{i}}\right)+1$	30	$[-600,600]^d$	0
	$F_{10}(x)=0.1\left\{\sin^2(3\pi x_1)+\sum_{i=1}^{d}(x_i-1)^2[1+\sin^2(3\pi x_1+1)]+(x_d-1)^2[1+\sin^2(2\pi x_d)]\right\}+\sum_{i=1}^{d}u(x_i,5,100,4)$	30	$[-50,50]^d$	0
低维多峰	$F_{11}(x)=-\sum_{i=1}^{10}[(X-\boldsymbol{a}_{\text{SHi}})(X-\boldsymbol{a}_{\text{SHi}})^{\text{T}}+\boldsymbol{c}_{\text{SHi}}]^{-1}$	4	$[0,10]^4$	−10.536

3.2　RGMFO 算法与其他算法对比分析

为得到更准确的实验结果, 10 种算法分别在 11 个测试函数上进行 30 次独立实验, 结果如表 3～表 5 所示。可以看出, 在相同条件下, RGMFO 算法

表3 RGMFO算法及其他算法高维单峰函数寻优结果

Table 3 High-dimensional single-peak function optimization results of RGMFO algorithm and other algorithms

算法	平均值						标准差					
	F_1	F_2	F_3	F_4	F_5	F_6	F_1	F_2	F_3	F_4	F_5	F_6
RGMFO	0	1.68×10^{-208}	0	4.19×10^{-208}	1.98×10^{-5}	1.07×10^{-4}	0	0	0	0	3.07×10^{-5}	1.2×10^{-4}
IMFO	2.68×10^{-7}	5.13×10^{-4}	5 850	17.6	15 700	0.117	5.15×10^{-7}	6.74×10^{-4}	3 440	4.85	11 000	0.044 7
CGMFO	1.66×10^{-23}	7.32×10^{-14}	1.81×10^{-17}	7.50×10^{-15}	1.46×10^{-3}	0.026 8	7.45×10^{-23}	2.92×10^{-13}	9.93×10^{-17}	1.66×10^{-14}	2.92×10^{-3}	0.021 8
ODSFMFO	1.96×10^{-24}	4.98×10^{-13}	1.40×10^{-14}	8.87×10^{-14}	0.017	0.056 3	8.54×10^{-24}	1.55×10^{-12}	6.08×10^{-14}	2.12×10^{-13}	0.045 9	0.040 3
MFO	4 970	20.3	59 600	22.6	494	0.824	2 350	4.54	58 600	6.94	2 690	0.439
GWO	9.68×10^{-28}	1.01×10^{-16}	2.09×10^{-5}	5.60×10^{-7}	27.9	2.18×10^{-3}	1.65×10^{-27}	1.19×10^{-16}	4.99×10^{-5}	4.36×10^{-7}	0.851	1.48×10^{-3}
PSO	12.2	0.026 9	10 500	37.2	7 480	0.1	17.8	0.041	8 450	9.88	26 100	0.083 9
DE	0.399	0.402	1 120	0.179	36.2	7.43×10^{-3}	0.344	0.312	2 940	0.124	7.33	3.91×10^{-3}
GA	2.12×10^{-5}	2.61×10^{-4}	18 500	14.1	752	0.314	1.31×10^{-5}	6.69×10^{-5}	3 530	4.79	460	0.067 9

表4 RGMFO算法及其他算法高维多峰函数寻优结果

Table 4 High-dimensional multi-peak function optimization results of RGMFO algorithm and other algorithms

算法	平均值				标准差			
	F_7	F_8	F_9	F_{10}	F_7	F_8	F_9	F_{10}
RGMFO	0	8.88×10^{-16}	0	7.58×10^{-6}	0	0	0	1.10×10^{-5}
IMFO	69.9	1.16	8.86×10^{-3}	7.16	19.5	0.246	9.33×10^{-3}	5.34
CGMFO	0	1.60×10^{-14}	0	0.022 6	0	4.58×10^{-14}	0	0.030 9
ODSFMFO	0	1.92×10^{-14}	0	0.023 1	0	4.16×10^{-14}	0	0.021 4
MFO	113	2.58	36.2	3.87×10^{6}	25	0.321	13.8	3.51×10^{6}
GWO	4.09	0.021 5	3.73×10^{-3}	0.653	6.06	0.056 2	8.01×10^{-3}	0.26
PSO	39.9	1.07	0.876	1.90×10^{5}	30.6	0.529	0.357	8.09×10^{5}
DE	33.2	0.395	0.591	2.78	35.6	0.387	0.296	0.639
GA	27.4	0.631	1.33×10^{-4}	0.358	6.85	0.087 4	3.17×10^{-4}	0.965

表5 RGMFO算法及其他算法低维多峰函数寻优结果

Table 5 Low-dimensional multi-peak function optimization results of RGMFO algorithm and other algorithms

算法	平均值	标准差
RGMFO	−10.5	4.47×10^{-9}
IMFO	−8.2	3.43
CGMFO	−10.5	4.34×10^{-4}
ODSFMFO	−8.19	2.72
MFO	−8.72	2.88
GWO	−9.99	2.06
PSO	−3.89	1.65
DE	−10.5	6.09×10^{-3}
GA	−10.4	0.987

的寻优能力高于其他MFO改进算法和其他优化算法。对于高维单峰函数，RGMFO算法有着更高的收敛精度，在平均值和标准差方面也远高于其他算法；对于高维多峰函数，RGMFO算法与CGMFO算法、ODSFMFO算法在F_7与F_9函数上都能找到最优解，而在其他函数上，RGMFO算法的收敛精度却能超越其他算法几个甚至十几个数量级；对于低维多峰函数，虽然所选算法已经较为优秀，但RGMFO算法依然能得到更精确的结果。由此可见，RGMFO算法比其他算法有着更好的搜索能力，以及更好的稳定性与鲁棒性，既能对解空间进行充分探索，又具备很高的收敛精度。

为进一步验证RGMFO算法的优秀性能，在α=0.05的显著性水平下，对30次独立运算下的RGMFO算法与其他9种算法的寻优结果进行Wilcoxon符号秩检验，得到的结果如表6所示。Wilcoxon符号秩检验的p值表示2组数据对比的相关性，若小于0.05，则认为其有明显差异。表6中的数值大多远小于0.05，这说明RGMFO算法的

表 6　Wilcoxon 秩和检验 p 值

Table 6　*P*-value for Wilcoxon rank sum test

函数	RGMFO-IMFO	RGMFO-CGMFO	RGMFO-ODSFMFO	RGMFO-MFO	RGMFO-GWO	RGMFO-SCA	RGMFO-PSO	RGMFO-DE	RGMFO-GA
F_1	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}
F_2	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}
F_3	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}
F_4	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}
F_5	3.02×10^{-11}	2.44×10^{-9}	3.69×10^{-11}	5.57×10^{-10}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}
F_6	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}
F_7	1.21×10^{-12}			1.21×10^{-12}	1.94×10^{-9}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}
F_8	1.21×10^{-12}	2.82×10^{-5}	2.16×10^{-6}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}
F_9	1.21×10^{-12}			1.21×10^{-12}	0.011	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}	1.21×10^{-12}
F_{10}	3.02×10^{-11}	3.34×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.01×10^{-7}	3.02×10^{-11}
F_{11}	0.025 4	3.47×10^{-10}	2.30×10^{-7}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	3.02×10^{-11}	2.20×10^{-10}	3.02×10^{-11}

性能在绝大多数测试函数上明显优于其他算法, 仅在 2 个测试函数上与 2 个 MFO 改进算法持平。

表 7 展示了 10 种算法在 11 个测试函数上的运算时间。可以看出, PSO 算法、MFO 算法和 GA 算法的运算时间较短(平均用时分别为 0.122 0 s、0.129 4 s、0.130 2 s), 速度较快。与其他 MFO 改进算法相比, RGMFO 算法的运算时间(平均用时 0.679 4 s)介于效果较好的 CGMFO 算法(平均用时 0.251 s)与 ODSFMFO 算法(平均用时 1.335 9 s)之间, 且在大部分测试函数上与 IMFO 算法(平均用时 0.464 5 s)相差不多。因此, 考虑到 RGMFO 算法的强探索能力与高收敛精度, 所提算法所付出的时间代价是可以接受的。

通过绘制基准测试函数迭代收敛曲线, 可以更加直观地看出算法的收敛性能和跳出局部最优解的能力。图 3 绘制了 10 种算法在 3 个测试函数上的迭代收敛曲线。可以看出, RGMFO 算法不仅收敛精度更高, 而且收敛到相同精度下所需的迭代次数比其他算法少, 这表明 RGMFO 算法有较快的开发速度和较高的局部逃逸能力, 能兼顾全局探索与局部开发。

3.3　消融实验分析

引入多项改进策略后, RGMFO 算法的寻优能力大幅提升, 为具体分析每种策略对算法改进的影响, 对 RGMFO 算法进行了消融实验。在原始 MFO 算法的基础上, 将引入火焰高斯变异策略的算法定义为 GMFO 算法, 将引入 ROL 策略、改进飞蛾更新机制的算法定义为 RMFO 算法。使用 GMFO 算

表 7　RGMFO 算法及其他算法寻优运算时间

Table 7　Optimization operation time of RGMFO algorithm and other algorithms　　s

函数	RGMFO	IMFO	CGMFO	ODSFMFO	MFO	GWO	SCA	PSO	DE	GA
F_1	0.257 9	0.163 1	0.084 2	0.785 5	0.067 5	0.240 6	0.172 9	0.057 9	0.185 8	0.085 0
F_2	0.469 5	0.363 9	0.193 7	1.020 6	0.071 6	0.233 1	0.176 7	0.060 6	0.122 3	0.075 0
F_3	0.503 8	0.389 3	0.222 8	1.036 6	0.090 5	0.247 6	0.194 8	0.076 1	0.140 4	0.083 9
F_4	0.535 0	0.373 7	0.198 1	1.023 0	0.070 2	0.225 7	0.181 5	0.113 6	0.120 6	0.072 9
F_5	0.567 6	0.357 8	0.183 3	1.012 7	0.064 8	0.236 9	0.168 5	0.056 9	0.145 5	0.065 6
F_6	0.980 2	0.747 9	0.366 2	1.608 0	0.215 3	0.396 3	0.393 3	0.209 3	0.236 1	0.132 0
F_7	0.718 6	0.426 6	0.200 4	1.324 6	0.096 1	0.239 5	0.187 8	0.090 1	0.137 8	0.077 9
F_8	0.732 1	0.466 9	0.230 5	1.494 8	0.100 1	0.253 1	0.213 1	0.092 4	0.140 1	0.080 0
F_9	0.727 0	0.605 8	0.320 0	1.995 9	0.167 2	0.314 8	0.264 4	0.127 1	0.209 6	0.137 1
F_{10}	1.172 2	0.740 3	0.415 1	1.930 5	0.327 7	0.457 1	0.477 0	0.320 5	0.686 7	0.418 4
F_{11}	0.809 2	0.473 8	0.346 4	1.463 1	0.152 7	0.171 1	0.160 7	0.137 9	0.342 7	0.204 1

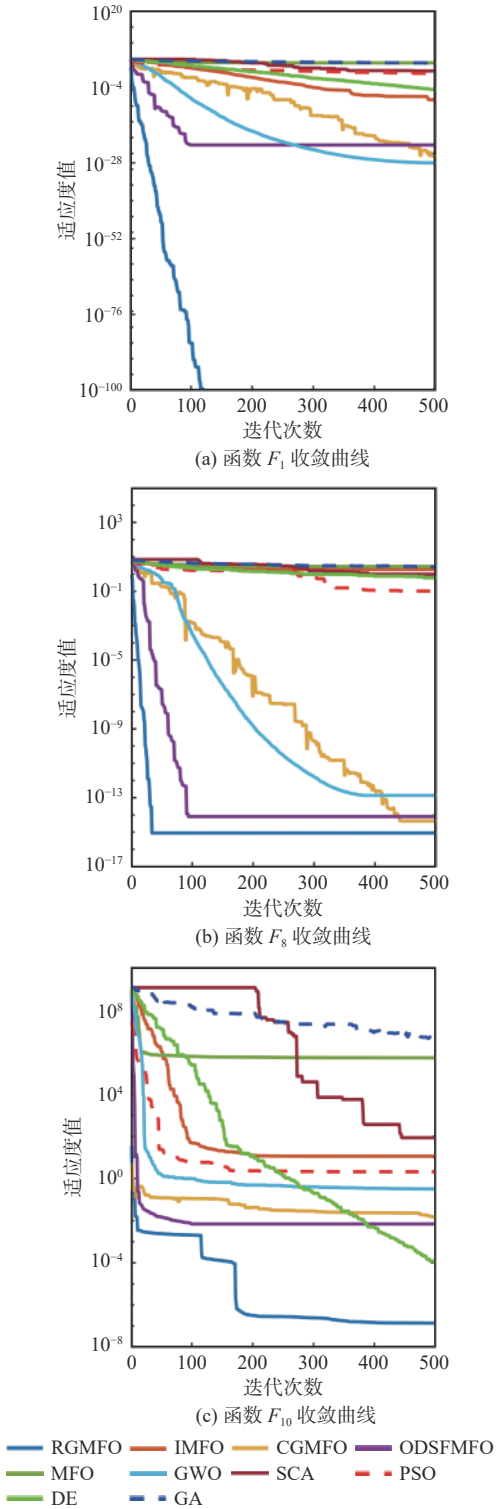


图3 RGMFO算法与对比算法迭代收敛曲线

Fig. 3 Iterative convergence curves of RGMFO algorithm and comparative algorithms

法和RMFO算法在11个测试函数上进行30次独立运算，将计算结果与MFO算法、RGMFO算法进行对比，如表8～表10所示。

从表8～表10可以看出，对于高维单峰函数F_1～F_4、高维多峰函数F_7～F_9，GMFO算法和RMFO算法的寻优能力相比MFO算法均有提升，且GMFO算法的结果显著优秀；二者结合后的RGMFO算法，除GMFO算法在F_7～F_9上已经找到最优解外，RGMFO算法在其他测试函数上均还会有进一步的提升。

对于函数F_5和F_{10}，GMFO算法和RMFO算法的寻优能力比MFO算法有所提升，且RMFO算法的结果较为优秀，二者结合后的RGMFO算法能够更进一步地提高寻优能力，找到更优秀的结果。

对于函数F_6，RMFO算法比MFO算法的寻优能力提升不大，而对于函数F_{11}，GMFO算法和RMFO算法的寻优结果不如MFO算法，但二者结合后的RGMFO算法，寻优结果能够明显优于MFO算法。

因此，火焰高斯变异对RGMFO算法优化能力的提升十分关键，ROL策略和改进飞蛾更新机制可以使RGMFO算法的寻优能力达到更好的效果。

RGMFO算法从整体探索、个体更新和最优值进化3个角度对MFO算法进行了改进，更加具有针对性，也能够更好地弥补MFO算法存在的短板与不足，从而实现更高效、更精确的运算。ROL策略在很大程度上提高了算法的全局探索能力；使用高斯变异对最优位置附近进行更进一步的探索，大大推动了对极值点的开发，提高了运算精度；飞蛾位置采用3种更新方式可以提高飞蛾对未知区域的探索能力，并且在探索中融入开发，探索到多个局部最优位置并进行比较择优，由此有效避免全局最优解的丢失。

综上所述，RGMFO算法在11个基准函数上的收敛精度、全局搜索能力和局部最优逃逸能力显著高于其他优化算法及其他MFO改进算法；通过消融实验可以证明，提出的每一种策略都对RGMFO算法寻优能力的提升有所帮助。多策略融合改进下的RGMFO算法具有更强的寻优能力，也能够得

表8 高维单峰函数消融实验结果

Table 8 High-dimensional single-peak function experimental results of model ablation

算法	平均值						标准差					
	F_1	F_2	F_3	F_4	F_5	F_6	F_1	F_2	F_3	F_4	F_5	F_6
RGMFO	0	1.68×10^{-208}	0	4.19×10^{-208}	1.98×10^{-5}	1.07×10^{-4}	0	0	0	0	3.07×10^{-5}	1.20×10^{-4}
GMFO	2.00×10^{-210}	6.63×10^{-115}	2.59×10^{-165}	3.66×10^{-110}	2.66	0.022 5	0	3.16×10^{-114}	0	1.86×10^{-109}	3.23	0.016 4
RMFO	0.263	1.66	7.38	0.178	5.72×10^{-5}	0.62	0.144	0.823	5.47	0.088 8	2.15×10^{-4}	1.02
MFO	4 970	20.3	59 600	22.6	494	0.824	2 350	4.54	58 600	6.94	2 690	0.439

表 9　高维多峰函数消融实验结果

Table 9　High-dimensional multi-peak function experimental results of model ablation

算法	平均值				标准差			
	F_7	F_8	F_9	F_{10}	F_7	F_8	F_9	F_{10}
RGMFO	0	8.88×10^{-16}	0	7.58×10^{-6}	0	0	0	1.10×10^{-5}
GMFO	0	8.88×10^{-16}	0	0.145	0	0	0	0.061 5
RMFO	0.477	0.014 7	0.010 3	2.37×10^{-5}	0.947	0.016	6.74×10^{-3}	4.97×10^{-5}
MFO	113	2.58	36.2	3.87×10^{6}	25	0.321	13.8	3.51×10^{6}

表 10　低维多峰函数消融实验结果

Table 10　Low-dimensional multi-peak function experimental results of model ablation

算法	平均值	标准差
RGMFO	−10.5	4.47×10^{-9}
GMFO	−7.72	2.66
RMFO	−5.13	1.83×10^{-12}
MFO	−8.72	2.88

到更优秀的寻优效果。

4　RGMFO 算法的实际工程应用

本文选用 2 种常见的工程设计优化问题对 RGMFO 算法的实际应用效果进行可行性检验, 分别为减速器设计问题和槽形舱壁设计问题。

4.1　减速器设计问题

在机械系统中, 减速器是齿轮箱的重要部件之一, 可适用于多种应用。在该优化问题中, 减速器的质量将在 11 个约束条件下实现最小化。该问题有 7 个变量, 分别为齿面宽 b_g、齿轮模数 m、小齿轮中的齿数 z、轴承之间第 1 根轴的长度 l_1、轴承之间第 2 根轴的长度 l_2、第 1 根轴的直径 d_1 和第 2 根轴的直径 d_2。$\boldsymbol{X}$ 表示待优化变量的向量, $\boldsymbol{X}$=[x_1, x_2, ⋯, x_7] 这 7 个变量分别对应待优化变量 x_1～x_7。对于该优化问题, 其目标函数为

$$f(\boldsymbol{X})=0.785\,4x_1x_2^2\left(3.333\,3x_3^2+14.933\,4x_3-43.093\,4\right)-1.508x_1(x_6^2+x_7^2)+7.477\,7\left(x_6^3+x_7^3\right)+0.785\,4\left(x_4x_6^2+x_5x_7^2\right)\quad(10)$$

约束条件如式 (11)～式 (21) 所示:

$$g_1(\boldsymbol{X})=\frac{27}{x_1x_2^2x_3}-1\leqslant0\quad(11)$$

$$g_2(\boldsymbol{X})=\frac{397.5}{x_1x_2^2x_3^2}-1\leqslant0\quad(12)$$

$$g_3(\boldsymbol{X})=\frac{1.93x_4^3}{x_2x_6^4x_3}-1\leqslant0\quad(13)$$

$$g_4(\boldsymbol{X})=\frac{1.93x_5^3}{x_2x_7^4x_3}-1\leqslant0\quad(14)$$

$$g_5(X)=\frac{\sqrt{(745x_4/(x_2x_3))^2+16.9\times10^6}}{110x_6^3}-1\leqslant0\quad(15)$$

$$g_6(X)=\frac{\sqrt{(745x_5/(x_2x_3))^2+157.5\times10^6}}{85x_7^3}-1\leqslant0\quad(16)$$

$$g_7(X)=\frac{x_2x_3}{40}-1\leqslant0\quad(17)$$

$$g_8(X)=\frac{5x_2}{x_1}-1\leqslant0\quad(18)$$

$$g_9(X)=\frac{x_1}{12x_2}-1\leqslant0\quad(19)$$

$$g_{10}(X)=\frac{1.5x_6+1.9}{x_4}-1\leqslant0\quad(20)$$

$$g_{11}(X)=\frac{1.1x_7+1.9}{x_5}-1\leqslant0\quad(21)$$

边界约束如表 11 所示。

表 11　减速器设计变量边界约束

Table 11　Reducer design variable boundary constraints

变量	下界	上界
x_1	2.6	3.6
x_2	0.7	0.8
x_3	17	28
x_4	7.3	8.3
x_5	7.8	8.3
x_6	2.9	3.9
x_7	5.0	5.5

在 MATLAB 上使用 RGMFO 算法和其他算法对上述问题进行寻优实验, 得到的结果如表 12 所示。

表 12　减速器设计结果

Table 12　Reducer design results

算法	x_1	x_2	x_3	x_4	x_5	x_6	x_7	适应度值
RGMFO	3.50	0.70	17.00	7.30	7.80	3.35	5.29	2 996.41
IMFO	3.50	0.70	17.00	7.30	7.80	3.35	5.29	2 996.41
CGMFO	3.50	0.70	17.00	7.60	7.80	3.80	5.29	3 130.09
ODSFMFO	3.50	0.70	17.00	7.30	7.80	3.35	5.29	2 996.41
MFO	3.50	0.70	17.00	7.30	7.80	3.35	5.37	3 047.56
GWO	3.50	0.70	17.00	7.39	7.85	3.36	5.29	3 003.19
SCA	3.60	0.70	17.00	8.30	8.30	3.45	5.37	3 133.94
PSO	3.60	0.70	17.00	7.30	8.28	3.35	5.29	3 046.37
DE	3.50	0.70	17.00	7.30	7.80	3.35	5.29	2 996.41
GA	3.57	0.70	17.16	7.67	7.82	3.43	5.30	3 096.49

从表 12 可以看出, RGMFO、IMFO 和 ODSFMFO 这 3 种算法在减速器的设计中得到了同样的

最优解，计算出了最小适应度值，且得到的优化变量的值也相等，这说明在此问题的求解中，RGMFO算法能够合理选择待优化变量，并得到最优解。

4.2 槽形舱壁设计问题

槽形舱壁设计问题是一个结构工程设计问题，该问题旨在使化学品罐车波槽形舱壁的质量最小化，其中，设计变量为宽度 x_1、深度 x_2、长度 x_3 和板厚 x_4。该优化问题的目标函数为

$$f(\boldsymbol{X})=\frac{5.885x_4(x_1+x_3)}{x_1+\sqrt{|x_3^2-x_2^2|}} \tag{22}$$

条件约束与边界约束如式(23)～式(30)所示：

$$g_1(\boldsymbol{X})=-x_4x_2\left(0.4x_1+\frac{x_3}{6}\right)+8.94\left(x_1+\sqrt{|x_3^2-x_2^2|}\right)\leqslant 0 \tag{23}$$

$$g_2(\boldsymbol{X})=-x_4x_2\left(0.2x_1+\frac{x_3}{12}\right)+2.2\left[8.94\left(x_1+\sqrt{|x_3^2-x_2^2|}\right)\right]^{4/3}\leqslant 0 \tag{24}$$

$$g_3(X)=-x_4+0.015\,6x_1+0.15\leqslant 0 \tag{25}$$

$$g_4(X)=-x_4+0.015\,6x_3+0.15\leqslant 0 \tag{26}$$

$$g_5(X)=-x_4+0.15\leqslant 0 \tag{27}$$

$$g_6(X)=-x_3+x_2\leqslant 0 \tag{28}$$

$$0\leqslant x_i\leqslant 100 \qquad i=1,2,3 \tag{29}$$

$$0\leqslant x_4\leqslant 5 \tag{30}$$

在MATLAB上进行寻优实验，得到的结果如表13所示。

在槽形舱壁的设计中，RGMFO算法与IMFO算法、DE算法得到了比其他算法更好的设计结果，得到了最优解。由此可以看出，在多约束优化问题中，RGMFO算法有着很好的寻优能力，说明了该算法在实际工程应用中的可行性，具有良好的应用前景。

表13 槽形舱壁设计结果

Table 13 Trough bulkhead design results

算法	x_1	x_2	x_3	x_4	适应度值
RGMFO	37.51	33.11	37.51	0.74	5.89
IMFO	37.83	33.19	37.83	0.74	5.89
CGMFO	31.34	32.48	37.88	0.78	6.23
ODSFMFO	38.89	33.46	38.89	0.76	5.90
MFO	37.28	33.23	38.13	0.74	5.90
GWO	37.69	33.20	37.85	0.74	5.90
SCA	32.71	35.53	46.11	0.90	6.70
PSO	40.91	33.30	39.36	0.79	6.02
DE	37.49	33.11	37.49	0.73	5.89
GA	90.70	33.53	75.91	1.67	10.29

5 结 论

1）根据对MFO算法计算过程和算法结构的不足的分析，针对飞蛾位置更新过程和火焰变异过程等方面提出了相应的改进措施，大大提高了算法的全局探索能力和局部开发精度。

2）使用RGMFO算法对基准函数进行寻优，能够得到比其他算法更好的精度和效果，收敛速度更快，稳定性与鲁棒性相较其他算法更强。

3）RGMFO算法在工程中具有良好的应用，对于工程设计优化问题的求解有着很好的寻优效果，进一步验证了算法的优越性。

参考文献（References）

[1] MIRJALILI S. Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm[J]. Knowledge-Based Systems, 2015, 89: 228-249.

[2] 张淑清, 段晓宁, 张立国, 等. Tsne降维可视化分析及飞蛾火焰优化ELM算法在电力负荷预测中应用[J]. 中国电机工程学报, 2021, 41(9): 3120-3130.
ZHANG S Q, DUAN X N, ZHANG L G, et al. Tsne dimension reduction visualization analysis and moth flame optimized ELM algorithm applied in power load forecasting[J]. Proceedings of the CSEE, 2021, 41(9): 3120-3130(in Chinese).

[3] 安许锋, 田洲, 钱锋. 基于多策略扰动机制的飞蛾火焰算法及在乙烯聚合过程动力学参数估计中的应用[J]. 高校化学工程学报, 2018, 32(4): 918-925.
AN X F, TIAN Z, QIAN F. Moth-flame optimization algorithm based on multi-strategy disturbance and its application in kinetic parameter estimation of ethylene polymerization[J]. Journal of Chemical Engineering of Chinese Universities, 2018, 32(4): 918-925 (in Chinese).

[4] 田鸿, 陈国彬, 刘超. 新型飞蛾火焰优化算法的研究[J]. 计算机工程与应用, 2019, 55(16): 138-143.
TIAN H, CHEN G B, LIU C. Research on new moth-flame optimization algorithm[J]. Computer Engineering and Applications, 2019, 55(16): 138-143(in Chinese).

[5] 黄鹤, 吴琨, 王会峰, 等. 基于改进飞蛾扑火算法的无人机低空突防路径规划[J]. 中国惯性技术学报, 2021, 29(2): 256-263.
HUANG H, WU K, WANG H F, et al. Path planning of UAV low altitude penetration based on improved moth-flame optimization[J]. Journal of Chinese Inertial Technology, 2021, 29(2): 256-263(in Chinese).

[6] 刘倩, 冯艳红, 陈嶷瑛. 基于混沌初始化和高斯变异的飞蛾火焰优化算法[J]. 郑州大学学报 (工学版), 2021, 42(3): 53-58.
LIU Q, FENG Y H, CHEN Y Y. Moth-flame optimization algorithm based on chaotic initialization and Gaussian mutation[J]. Journal of Zhengzhou University (Engineering Science), 2021, 42(3): 53-58(in Chinese).

[7] PELUSI D, MASCELLA R, TALLINI L, et al. An improved moth-flame optimization algorithm with hybrid search phase[J]. Know-

ledge-Based Systems, 2020, 191: 105277.

[8] LI Z F, ZENG J H, CHEN Y Q, et al. Death mechanism-based moth-flame optimization with improved flame generation mechanism for global optimization tasks[J]. Expert Systems with Applications, 2021, 183: 115436.

[9] 刘宝宁, 章卫国, 李广文, 等. 一种改进的多目标粒子群优化算法[J]. 北京航空航天大学学报, 2013, 39(4): 458-462.
LIU B N, ZHANG W G, LI G W, et al. Improved multi-objective particle swarm optimization algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2013, 39(4): 458-462(in Chinese).

[10] MIRJALILI S. SCA: A sine cosine algorithm for solving optimization problems[J]. Knowledge-Based Systems, 2016, 96: 120-133.

[11] MIRJALILI S, MIRJALILI S M, LEWIS A. Grey wolf optimizer [J]. Advances in Engineering Software, 2014, 69: 46-61.

[12] TIZHOOSH H R. Opposition-based learning: a new scheme for machine intelligence[C]//Proceedings of the International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce. Piscataway: IEEE Press, 2006, 1: 695-701.

[13] LONG W, JIAO J J, LIANG X M, et al. A random opposition-based learning grey wolf optimizer[J]. IEEE Access, 2019, 7: 113810-113825.

[14] XU Y T, CHEN H L, LUO J, et al. Enhanced moth-flame optimizer with mutation strategy for global optimization[J]. Information Sciences, 2019, 492: 181-203.

[15] HEIDARI A A, MIRJALILI S, FARIS H, et al. Harris Hawks optimization: algorithm and applications[J]. Future Generation Computer Systems, 2019, 97: 849-872.

Improved moth-flame optimization algorithm with multi-strategy integration

HE Jiawen，XU Xianze*，GAO Bo

(Electronic Information School，Wuhan University，Wuhan 430072，China)

Abstract: An enhanced moth-flame optimization algorithm with multi-strategy integration (RGMFO) was proposed in order to address the issues that the moth-flame optimization algorithm was prone to falling into the local optimum and that its exploitation ability was insufficient while it was approaching the global optimum. To generate high-quality moth populations, a random opposition-based learning strategy was applied at the start of each iteration. To generate high-quality moth populations, a random opposition-based learning strategy was applied at the start of each iteration. Gaussian mutation was then used to swap out subpar flame individuals with superior ones. Archimedes spirals, weighting factors, and the surrounding of a small number of optimal flames were employed to enhance the moth update mechanism. The proposed algorithm was tested on 11 benchmark functions of different types. The test results and rank sum test show that RGMFO has better global search ability and higher search accuracy. Lastly, RGMFO is applied to the engineering scenarios of reducer design and trough bulkhead design, which further verifies the practicability and feasibility of the algorithm.

Keywords: moth-flame optimization algorithm; multi-strategy integration; random opposition-based learning; Gaussian mutation; intelligent optimization

Received：2022-08-11；**Accepted**：2022-09-02；**Published Online**：2022-11-08 10：42

URL：link.cnki.net/urlid/11.2625.V.20221108.0859.002

Foundation item：National Natural Science Foundation of China (51975422)

* **Corresponding author**. E-mail：xxz@whu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0721

动态场景中基于特征点几何约束的3D SLAM算法

刘铭坚[1]，罗景文[1,*]，秦世引[2]
(1. 云南师范大学 信息学院，昆明 650500；2. 北京航空航天大学 自动化科学与电气工程学院，北京 100191)

摘 要：针对动态场景中动态物体会导致机器人在进行位姿估计时引入大量动态误差的问题，提出一种利用特征点间几何约束来剔除动态特征点的移动机器人3D同步定位与地图构建(SLAM)算法。利用当前帧的ORB特征点与上一帧特征点生成的地图点进行投影匹配，通过引入Delaunay三角剖分法构建能够表示2帧间、匹配地图点间几何关系的三角网。利用相邻2帧地图点的几何关系变化检测出动态特征点，考虑到静态特征点可能被误检测为动态特征点而导致特征点缺失的情况，在相邻2帧匹配时提取更多的特征点以实现静态特征点的补偿，进而剔除动态特征点，实现对移动机器人位姿的精确估计。在此基础上，通过引入滑动窗口提取关键帧并完成闭环检测，从而构建出精确的3D稠密地图。在多组公开数据集上进行仿真实验及室内动态场景下的实验，结果表明，所提算法能够有效剔除动态特征点，提高移动机器人在动态场景中位姿估计的精度和地图的一致性。

关 键 词：动态场景；同步定位与地图构建；移动机器人；几何约束；三角剖分
中图分类号：TP242.6; TB553
文献标志码：A **文章编号**：1001-5965(2024)09-2872-13

近年来，随着智能机器人技术的发展及机器人在各行各业的推广应用，面向动态场景的机器人自主作业成为了具有挑战性的研究热点。而同步定位与地图构建(simultaneous localization and mapping, SLAM)是机器人自主导航的关键支撑技术，该技术可使机器人利用自身携带的传感器通过环境的感知信息对自身位姿进行实时估计，并构建全局一致的地图，以提高机器人的环境适应能力，满足自主作业的应用需求。

为简化移动机器人的位姿求解问题，多数SLAM算法会将场景视为静态，导致动态场景中的机器人将动态物体视为与背景环境一样的静态，因此，从动态物体中提取到的特征会被视为静态特征点，从而出现大量动态误差，导致机器人的最终定位结果不准确，所构建的地图也会产生严重偏差，甚至算法不收敛。虽然可以利用随机抽样一致(random sample consensus, RANSAC)算法[1]和光束平差(bundle adjustment, BA)法[2]来剔除部分动态特征点，解决未知环境中动态物体引入误差的问题，但当属于动态物体的动态特征点数量过多时，这2种方法便会失效。因此，如何有效区分动态特征点和静态特征点，提高SLAM系统在动态场景中的鲁棒性和可靠性是确保移动机器人高效、稳健自主作业的关键。

目前，多数针对动态场景的SLAM解算方案通常是在现有的优秀的开源SLAM系统架构基础上引入动态目标检测的，其中，如何精准地分割出图像中的动态目标成为需要解决的关键问题，主要包括2类方法：

1) 基于几何的方法，即依赖几何约束技术来处理未知环境中的动态成分。Zou和Tan[3]利用相邻2帧的重投影特征距离误差区分出动态和静态特征，从而实现动态特征的分割，获取准确的相机位姿。张慧娟等[4]利用ORB[5]点特征和直线段检测

收稿日期：2022-08-17；**录用日期**：2022-09-21；**网络出版时间**：2022-10-08 18:33
网络出版地址：link.cnki.net/urlid/11.2625.V.20221006.1321.002
基金项目：国家自然科学基金(62063036)；云南师范大学博士科研启动项目(01000205020503115)
*通信作者. E-mail：by1503117@buaa.edu.cn

引用格式：刘铭坚，罗景文，秦世引. 动态场景中基于特征点几何约束的3D SLAM算法[J]. 北京航空航天大学学报，2024，50(9)：2872-2884. LIU M J，LUO J W，QIN S Y. 3D SLAM algorithm based on geometric constraints of feature points in dynamic scenarios[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2872-2884 (in Chinese).

(line segment detector, LSD)算法[6]线特征进行相机位姿的估计, 可减少环境中动态成分对 SLAM 算法的影响, 但容易在非结构化的未知环境中因线特征不足导致跟踪失败。Sun 等[7]通过对连续 RGB 图像进行深度图聚类, 分割出动态物体像素点, 从而获取场景中的静态部分, 但分割动态成分耗时太长, 导致实时性不足。艾青林等[8]利用严格的几何约束和多帧的观测信息将未知状态点区分为静态特征点、动态特征点和错误匹配点, 并引入卡尔曼滤波筛选, 该方法有较好的动态检测效果和定位精度, 但由于根据地图点的性质分成了 2 个地图, 因而降低了地图的一致性。Li 和 Lee[9]提出了一种基于关键帧中边缘点检测及对其进行静态加权的算法, 从而减少了动态物体对 SLAM 系统中位姿估计的影响。

2) 基于光流的方法, 即通过估计相邻 2 帧的像素运动来分割动态物体, 从而实现精确的位姿估计。高成强等[10]在半直接法视觉里程计的基础上通过双单高斯模型(single Gaussian model, SGM)来构建图像块的灰度值变化模型, 根据图像块方差的变化区分前景和背景, 并检测出动态目标, 消除动态特征, 从而构建出基于截段符号距离函数(truncated signed distance function, TSDF)模型的网格状地图。Derome 等[11]使用预测图像与双目相机观察图像之间的残差来计算光流, 再通过检测残差场中的点来观察动态物体。Li 和 Xu[12]利用光流与图像的超像素分割实现对场景中动态目标的检测。Jaimez 等[13]根据图像的几何和彩色信息进行 K 均值聚类得出图像块, 再用直接法估计相机位姿, 然后根据估计的位姿对图像块进行动态物体的判断。然而, 基于光流的分割动态物体的方法计算量较大, 较难实现实时计算, 且其受环境光照的影响较大。

基于上述分析, 针对动态场景下移动机器人 3D SLAM 的应用需求, 本文在提取的 ORB 特征点基础上引入 Delaunay 三角剖分法构建出能够表示相邻 2 帧匹配地图点几何关系的三角网, 并利用当前帧的特征点反投影生成 3D 点, 与三角网中表示特征点几何关系的边向量进行向量距离的比较, 检测出发生动态变化的边, 进一步确定并剔除动态特征点, 从而有效地提高移动机器人在动态场景中位姿估计的精度。

1　动态场景下移动机器人 3D SLAM 算法设计

本文算法采用基于图优化的视觉 SLAM 算法架构, 主要分为前端和后端 2 个部分。首先, 前端特征点的提取采用 ORB 算法[5], 该算法在 FAST 角点的基础上构建了图像金字塔以实现特征的尺度不变性, 并且通过计算图像灰度质心实现了特征的旋转不变性, 因此, 其已被广泛应用于各种视觉 SLAM 方案前端部分的特征提取。其次, 通过引入 Delaunay 三角剖分法构建地图点间的几何关系来剔除属于动态物体的特征点, 即根据相邻 2 帧的匹配地图点进行三角剖分, 从而构建出能够表示匹配地图点间几何关系的 3D 三角网, 再利用当前帧的特征点反投影方法获取并剔除动态特征点。最后, 输出静态特征点。为获得高质量的关键帧, 提高系统的实时性, 在前端引入一个滑动窗口来提取局部最优关键帧, 从而在动态场景中实现移动机器人位姿的精确估计和 3D 稠密地图的构建。算法架构如图 1 所示。

1.1　表示匹配地图点几何关系的三角网构建

本文算法中, 动态特征点的检测是基于相邻帧

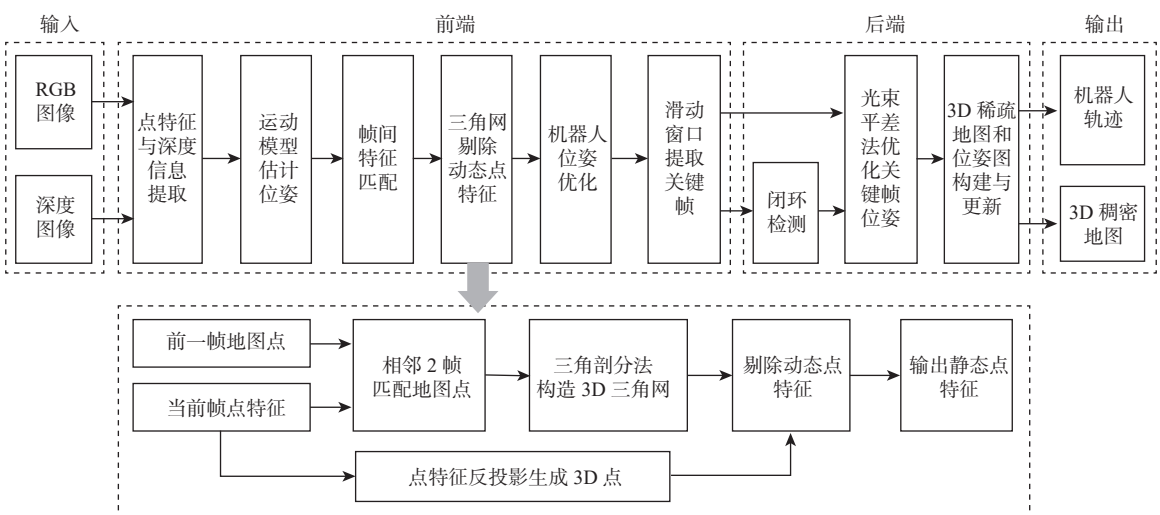


图 1　基于三角剖分法剔除动态特征点的 3D SLAM 算法架构

Fig. 1　Architecture of 3D SLAM algorithm based on triangulation method to eliminate dynamic feature points

的匹配特征点对应地图点之间的几何关系的变化来实现的。通常,未知环境中的静态特征点之间的几何关系不会发生变化,当环境中存在动态物体时,属于动态物体的动态特征点之间的几何关系也不会发生变化,但随着动态物体的运动,动态特征点与静态特征点之间的几何关系会相应地发生变化。文献[14]使用Delaunay三角剖分法[15]将图像中的特征点连结成三角形,并通过比较相邻2帧分别对应的三角形的边,将属于动态物体的特征点从三角剖分法构成的连通图中剔除出来,从而实现了图像动态区域的有效剔除。由于Delaunay三角剖分法生成的三角网数据冗余度小、存储效率高,且其构建的三角网能够表示地图点之间的几何关系,因此为避免建立出冗余度较高的特征点两两连接的关系网,本文考虑引入Delaunay三角剖分法来构建三角网用于表示地图点间的几何关系,并通过相邻帧的匹配特征的对应地图点几何关系的变化来检测动态特征点。

在对相邻图像特征点进行跟踪时,利用恒速运动模型获取当前帧的位姿,并将前一帧的匹配地图点投影到当前帧中,投影方程为

$$\begin{cases} u = f_x \dfrac{x}{z} + c_x \\ v = f_y \dfrac{y}{z} + c_y \end{cases} \tag{1}$$

式中:(u,v)为特征点的像素坐标;$P_c=(x,y,z)$为地图点在相机参考系下的坐标;f_x、f_y、c_x和c_y为RGB-D相机的内参。通过投影方程获取当前图像相对前一帧的匹配地图点后,再对获取的匹配地图点进行有效的Delaunay三角剖分,从而构建出能够表示相邻2帧的3D匹配地图点间几何位置关系的Delaunay三角网。

如图2所示,假设第i–1帧图像的特征点p_1、r_1、q_1的匹配地图点为P、Q、R,将P、Q、R这3个地图点投影到第i帧图像中,匹配到的特征点分别为p_2、r_2、q_2,接下来特征点p_1、r_1、q_1通过地图点P、Q、R与特征点p_2、r_2、q_2建立了匹配关系,因此,引入Delaunay三角剖分法构建由P、Q、R组成的Delaunay三角网来表示3个地图点间的几何关系,该三角网包括3个向量$\overrightarrow{PQ}$、$\overrightarrow{PR}$和$\overrightarrow{QR}$,由于向量具有长度和方向属性,因而能够表示出三角网中P、Q、R之间的距离和方向关系。

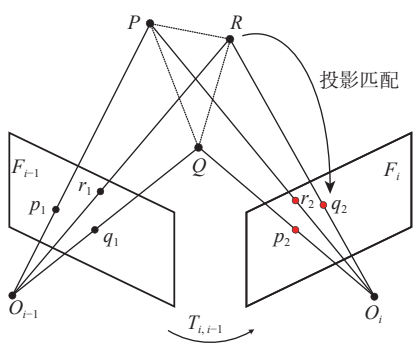


图2 投影模型

Fig. 2 Projection model

1.2 基于三角网向量边的动态特征点检测

完成对匹配特征点所对应地图点的三角网的构建后,将当前帧的匹配特征点反投影到相机坐标系中生成对应的3D点,其中,当前帧的特征点反投影方程为

$$\begin{cases} x = \dfrac{u - c_x}{f_x} z \\ y = \dfrac{v - c_y}{f_y} z \\ z = d \end{cases} \tag{2}$$

式中:(u,v)为当前帧的特征点像素坐标;d为从RGB-D深度相机的深度图像中获取的该特征点的深度信息;$P_c=(x,y,z)$为反投影生成的相机参考系下3D点的坐标。

由于构成三角网的向量边能够表示向量边连接的2个地图点间的几何关系,因此利用这2个地图点所对应的当前帧的特征点反投影生成的3D点连接形成的新向量边进行动态向量检测,如图3所示。假设第i帧的特征点(p_2, r_2, q_2)与前一帧的特征点(p_1, r_1, q_1)通过地图点P、Q、R投影完成特征匹配,并假设P、Q属于静态物体,R属于动态物体,第i帧的特征点p_2、r_2、q_2反投影生成的3D点为P'、Q'、R',由于P、Q地图点属于静态物体,而R属于动态物体,所以反投影生成的3D点P'、Q'分别与P、Q重合,R'不与R重合,同时由于投影误差的存在,即使是对属于静态物体的特征点进行反投影,也不会与匹配地图点完全重合,因此,可利用能够表示地图点间几何关系的向量来进行动态特征点的检测。动态向量检测函数ψ可定义为

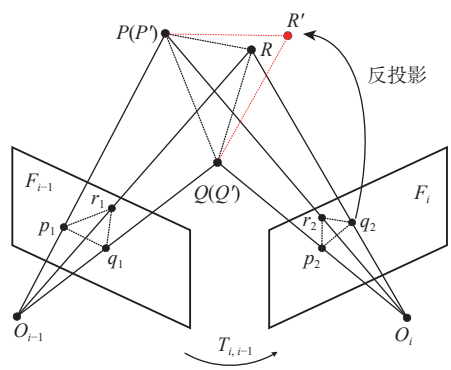


图3 算法原理

Fig. 3 Algorithm principle

$$\psi \to \psi\left(\overrightarrow{PR}, \overrightarrow{P'R'}\right) \tag{3}$$

式中: $\overrightarrow{PR}$为三角网的向量边; $\overrightarrow{P'R'}$为当前帧特征点反投影生成的 3D 点连接而成的向量边。

遍历基于匹配地图点建立的 Delaunay 三角网的向量边, 该向量边连接着 2 个匹配特征点对应的地图点, 表示的是这 2 个地图点的几何关系, 利用 Delaunay 三角网的向量边对该向量边连接的 2 个顶点所对应的当前帧特征点进行反投影生成的 3D 点连接的向量边进行动态评估, 通过计算 2 个向量边的相似度来确定该向量边的动态程度, 若该向量边被动态向量检测函数 ψ 判定为动态向量边, 则增加该边连接的 2 个地图点的动态权重, 以完成动态地图点动态权重的累积。

假设建立的 Delaunay 三角网存在如图 4 所示的三角形 $\triangle PQR$, 则属于该三角形的 3 条向量边分别为$\overrightarrow{PQ}$、$\overrightarrow{PR}$和$\overrightarrow{QR}$, 当前帧的特征点反投影生成如图 3 所示的三角形 $\triangle P'Q'R'$, 属于该三角形的 3 条向量边分别为$\overrightarrow{P'Q'}$、$\overrightarrow{P'R'}$和$\overrightarrow{Q'R'}$, 显然, $\overrightarrow{PQ}$与对应的当前帧特征点反投影生成的$\overrightarrow{P'Q'}$重合, 这说明动态物体的运动没有改变 P、Q 2 个地图点之间的几何关系。相反地, 由于 R 点的运动, 导致向量边$\overrightarrow{PR}$和向量边$\overrightarrow{QR}$都发生了改变。为获取向量边的动态程度, 引入能够表示向量距离的 L_2 范数及余弦相似度来计算向量相似度, 从而确定向量边的动态程度。其中, L_2 范数定义为

$$\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2} \tag{4}$$

余弦相似度定义为

$$\cos\theta = \frac{\overrightarrow{PR}\cdot\overrightarrow{P'R'}}{\left\|\overrightarrow{PR}\right\|\left\|\overrightarrow{P'R'}\right\|} \tag{5}$$

L_2 范数的意义在于比较 2 个向量边之间的长度变化, 而余弦相似度的意义在于比较 2 个向量边之间的角度变化。因此, 引入基于 L_2 范数的判定条件:

$$\frac{\left\|\overrightarrow{PR}\right\|}{\left\|\overrightarrow{P'R'}\right\|} \notin (\varepsilon_{\mathrm{dmin}}, \varepsilon_{\mathrm{dmax}}) \tag{6}$$

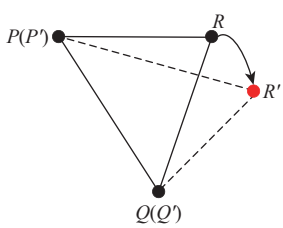


图 4　特征点反投影

Fig. 4　Back projection of feature points

式中: $\varepsilon_{\mathrm{dmax}}$ 为向量二范数的最大比值; $\varepsilon_{\mathrm{dmin}}$ 为向量二范数的最小比值。取 $[\varepsilon_{\mathrm{dmin}}, \varepsilon_{\mathrm{dmax}}]$=[0.999, 1.001], 限定向量边之间的距离比值变化位于 [0.999, 1.001]。同时, 为避免向量边不发生长度变化而发生角度变化, 引入了基于余弦相似度的判定条件:

$$\frac{\overrightarrow{PR}\cdot\overrightarrow{P'R'}}{\left\|\overrightarrow{PR}\right\|\left\|\overrightarrow{P'R'}\right\|} \leqslant \varepsilon_{\cos} \tag{7}$$

式中: $\varepsilon_{\cos}$ 为余弦相似度阈值, 取 0.999 848(对应角度为 1°), 小于此阈值, 这表明向量的方向向量发生了大于 1°的偏转。

引入 2 种向量相似度的判定条件来判定动态向量:

$$\psi\left(\overrightarrow{PR}, \overrightarrow{P'R'}\right) = \begin{cases} \text{true} & \dfrac{\left\|\overrightarrow{PR}\right\|}{\left\|\overrightarrow{P'R'}\right\|} \notin [\varepsilon_{\mathrm{dmin}}, \varepsilon_{\mathrm{dmax}}] \text{ 或} \\ & \dfrac{\overrightarrow{PR}\cdot\overrightarrow{P'R'}}{\left\|\overrightarrow{PR}\right\|\left\|\overrightarrow{P'R'}\right\|} \leqslant \varepsilon_{\cos} \\ \text{false} & \text{否则} \end{cases} \tag{8}$$

若向量$\overrightarrow{PR}$被标记为动态向量, 则增加该动态向量连结的 2 个地图点的动态权重:

$$\begin{cases} w_P = w_P + 1 \\ w_R = w_R + 1 \end{cases} \tag{9}$$

式中: w_P 和 w_R 分别为向量 $\overrightarrow{PR}$的 2 个端点的动态权重。

如图 4 所示, 假设向量$\overrightarrow{PR}$和向量$\overrightarrow{P'R'}$的相似度超过了设定阈值, 则将当前帧特征点反投影生成的向量$\overrightarrow{P'R'}$视为动态向量, 并将向量$\overrightarrow{PR}$连接的 2 个地图点的动态权重增加。同理, $\overrightarrow{Q'R'}$也被视为动态向量。经过三角形向量边的遍历后, 该三角形中 P 地图点和 Q 地图点的动态权重均为 1, R 地图点的动态权重为 2, 因此, 可将该三角形中属于动态物体的 R 地图点检测出来, 从而实现对相邻 2 帧动态特征点的检测。

1.3　静态特征点补偿和动态特征点剔除

动态目标具有 2 种性质: 静态性质和动态性质。其中, 静态性质指动态目标自身具有的性质, 与其运动状态无关, 可分为刚体和非刚体。对动态物体的定义是相对静态背景环境而言的, 对于环境中的动态刚体目标, 属于动态刚体目标的所有特征点相对静态背景的位姿变换的旋转是一致的, 但由于动态刚体目标特征点分布在刚体的表面, 并不位于同一空间三维点, 其位姿变换的平移不是一致的, 因此, 动态刚体目标的特征点的位姿变换是不一致的。同理, 动态非刚体目标上不同特征点的位

姿变换也是不一致的。因此，无论动态目标是刚体还是非刚体，属于该动态目标的特征点的相对静态背景都是动态的，只要动态目标相对静态背景运动，无论其是属于动态刚体目标的特征点还是属于动态非刚体目标的特征点，或者是具有变形特性的物体，均能在前端部分被引入的基于 Delaunay 三角网剖分法构成 Delaunay 三角网，只要后续物体继续运动或发生变形，属于动态物体的特征点之间构成的向量边就会被基于动态向量边动态特征点检测方法检测出来。因此，物体的刚体与非刚体性质差异对本文算法的性能影响不大。

除了动态物体自身具有的静态性质外，其运动时所赋予的动态性质也会对本文算法造成一定影响。当动态目标的速度较快时，其在图像中的投影可能就会比较模糊，导致系统无法从动态目标中提取出有效的特征点，从而规避了动态特征点的检测和剔除，如果系统能够从速度较快的动态目标中检测出特征点，那么图像中的特征点构成的 Delaunay 三角网中连接静态特征点和动态特征点的向量边变化幅度会较大，就容易被基于三角网向量边的动态特征点检测方法检测到。但当动态目标速度较慢时，特征点构成的 Delaunay 三角网中连接静态特征点和动态特征点的向量边变化幅度会较小，从而可能无法有效检测出动态向量。

当未知动态环境中属于动态物体的动态特征点数量过多时，引入的动态特征点检测方法剔除过多的特征点，就会导致相邻 2 帧跟踪失效，因此，需要进行静态特征点的补偿。本文在当前帧跟踪上一帧地图点前，利用上一帧未被 3D 化的特征点生成新的临时地图点，并与当前帧的特征点进行投影匹配，从而产成更多的匹配特征点，以确保匹配特征点中静态特征点的数量。

另外，由于动态特征点检测方法也会存在误检测的情况，即将静态特征点检测为动态特征点，所以如果将该部分被误检测的静态特征点剔除，就会导致前端部分的特征点的跟踪鲁棒性降低。考虑到未知环境中属于动态物体的特征点深度相近，对检测出的动态特征点进行深度值归类，并将检测出的动态特征点的深度值区间进行分为 n 段，每个区间为

$$s_i=(i-1,i) \tag{10}$$

从而深度值区间为一个区间集合：

$$S=\{s_1,s_2,\cdots,s_n\} \tag{11}$$

因此，可通过遍历该区间集合得出动态特征点最多的动态区间，完成对动态特征点的剔除。

如图 5 所示，将动态深度值分成 8 个区间，存在较多动态特征点的深度区间段 [i−1,i] 作为动态物体特征点存在的区间段，只对在该区间点的动态特征点作剔除处理，从而完成对动态特征点的剔除。

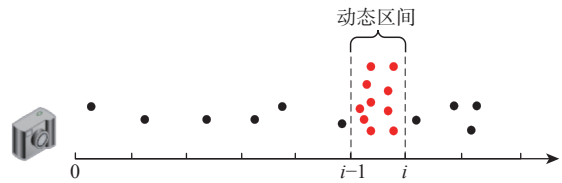


图 5 动态区间分段

Fig. 5 Segmentation of dynamic interval

1.4 移动机器人位姿估计

经过对动态特征点的处理，可以得到有效特征点集合 X 及有效地图点集合 P，由于环境噪声及相机观测噪声的存在，3D 空间中地图点的投影与图像的特征点位置并不重合，因此可以得到有效特征点集合 X 与有效地图点集合 P 的观测误差，并通过最小化重投影误差方法优化相机位姿：

$$\hat{\boldsymbol{T}}_{\mathrm{cw}}=\arg\min_{\hat{\boldsymbol{T}}_{\mathrm{cw}}}\frac{1}{2}\sum_{\boldsymbol{x}_i\in X,\boldsymbol{P}_i\in P}\|\boldsymbol{x}_i-\pi(\boldsymbol{T}_{\mathrm{cw}}\boldsymbol{P}_i)\|^2 \tag{12}$$

式中：$\boldsymbol{x}_i$ 为图像上的特征点；$\boldsymbol{P}_i$ 为特征点对应的 3D 地图点；π 为投影方程；$\boldsymbol{T}_{\mathrm{cw}}$ 为机器人的位姿。同时，由于特征匹配过程中可能发生误匹配而导致投影误差较大，因此可在位姿优化过程中引入鲁棒核函数，以进一步提高相机位姿估计的精度。

1.5 关键帧选择

在 SLAM 系统中，选择关键帧的目的在于筛选出冗余度低的高质量关键帧，并利用关键帧及其跟踪的特征点对移动机器人的位姿进行优化，从而构建出简洁高效的全局一致地图。目前，SLAM 开源方案通常采用当前的最优方法来挑选关键帧，之后往往在一段时间或一定空间距离内不再提取新的关键帧。然而，在动态场景中提取关键帧后的几帧内往往会出现更加适合作为关键帧的普通帧，这就无法有效保证关键帧的质量。当移动机器人在动态环境中自主作业时，沿用 SLAM 系统原有的基于静态环境下的关键帧选择策略虽然能够保证构建出冗余度低的地图，但由于动态环境下动态物体的存在，系统能够获取到的静态特征点整体比静态环境下的数量更少且质量更低，从而降低了关键帧的质量，也降低了地图精度。为降低动态环境下关键帧之间的冗余度及后端优化的计算的复杂度，提高 3D SLAM 系统在动态环境下的实时性和地图的一致性，本文在前端引入一个滑动窗口来提取局部最优关键帧。

滑动窗口的起始帧为符合基于静态环境下的

关键帧选择策略的第 1 个普通帧, 其终止帧是根据系统当前帧的动态程度判断得出的, 因此, 滑动窗口大小的设置不是固定的, 而是取决于移动机器人实际作业环境的动态程度, 其目的是避免选择跟踪到特征点整体质量较低的关键帧。为便于说明滑动窗口选择关键帧方法的具体执行策略, 设定起始帧与终止帧之间的窗口尺寸为 4, 如图 6 所示。

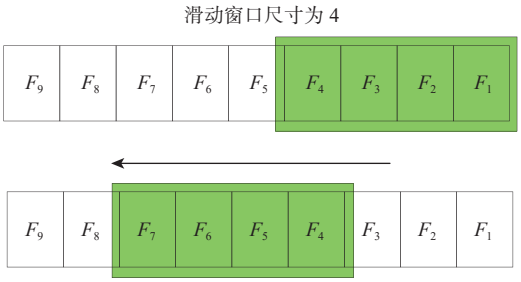


图 6　尺寸为 4 的滑动窗口

Fig. 6　Sliding window with a size of 4

滑动窗口中存在 4 帧普通帧, 从帧间运动距离及与前一关键帧的共视程度和普通帧的动态程度出发, 对处于滑动窗口内的 F_1、F_2、F_3、F_4 这 4 个普通帧的质量作出评估, 先出该 4 帧相对前一个关键帧的旋转向量 $\boldsymbol{R}$ 和平移向量 $\boldsymbol{t}$, 再求出滑动窗口内所有普通帧相对前一个关键帧的相对运动距离:

$$D = \|\boldsymbol{t}\| + \min(2\pi - \|\boldsymbol{R}\|, \|\boldsymbol{R}\|) \tag{13}$$

设置运动距离的上、下界为 $D_{\min}$、$D_{\max}$, 在 4 帧中挑选出满足条件 $D_{\min}<D<D_{\max}$ 且运动距离最大、与前一关键帧共同观测到的特征数量较少的普通帧作为候选局部最优的关键帧; 同时, 当移动机器人在动态环境中运动时, 如果某一帧的动态特征点比例过高, 由于系统可能存在动态特征点误检测及静态特征点误删除的情况, 因此会导致整体特征点质量低下的情况更加严重, 从而使该普通帧的质量降低。利用基于 Delaunay 三角剖分法的动态特征点剔除法剔除的特征点对应所有特征点比例, 引入对普通帧动态程度的检测, 得出拥有较高质量的特征点的普通帧作为关键帧, 并利用滑动窗口的局部特性, 将局部最优的关键帧插入 3D 稀疏地图中, 在当前滑动窗口的局部最优关键帧被挑选出之后, 将其普通帧的后一帧设定为新的滑动窗口的起始帧, 进入下一个关键帧的挑选流程。

随着移动机器人的持续运动, 系统的定位误差及建图误差不可避免地累积, 导致机器人的轨迹和构建的稀疏地图与实际的运动轨迹和环境地图发生了较大偏离, 影响系统的精确性。引入高效的基于词袋模型(bag of word model, BOW)的闭环检测方法来对输入图像和历史图像进行相似度计算, 以完成系统的闭环检测。利用 BOW 计算图像相似度时, 需要利用图像中提取到的 ORB 特征点的 BRIEF 描述子与视觉词典中作为聚类中心的节点来计算汉明距离, 并将这些描述子划分为词典中与其距离最近的叶子节点。根据具体的划分情况, 将输入图像转化为词袋向量 $\boldsymbol{v}$, 形式如下:

$$\boldsymbol{v} = \left[(w_1,\eta_1),(w_2,\eta_2),\cdots,(w_n,\eta_n)\right] \tag{14}$$

式中: w_i 为叶子节点; η_i 为叶子节点的权重; n 为词袋向量的维度, 即视觉词典中词汇的个数。

在进行图像之间的相似度计算时, 仅需要与具有共同视觉词汇的图像进行比较。假设已经给定了 2 幅图像的词袋向量 $\boldsymbol{v}_{\mathrm{a}}$ 和 $\boldsymbol{v}_{\mathrm{b}}$, 则 2 个词袋向量的 L_1-范数形式的距离为

$$s(\boldsymbol{v}_{\mathrm{a}},\boldsymbol{v}_{\mathrm{b}}) = 1 - \frac{1}{2}\left|\frac{\boldsymbol{v}_{\mathrm{a}}}{|\boldsymbol{v}_{\mathrm{a}}|} - \frac{\boldsymbol{v}_{\mathrm{b}}}{|\boldsymbol{v}_{\mathrm{b}}|}\right| \tag{15}$$

通过上述闭环检测图像间距离的计算, 选取与当前关键帧距离最小的历史关键帧作为闭环帧, 并利用该闭环帧优化构建的 3D 稀疏地图, 完成 3D SLAM 系统的闭环检测与 3D 稀疏地图的优化。

1.6　3D 稠密地图构建

完成关键帧选择后, 通过利用关键帧位姿及关键帧跟踪到的地图点构建出 3D 稀疏地图, 但稀疏地图的形式不能直观展现室内场景, 因此, 本文在移动机器人的精确位姿估计和滑动窗口生成的关键的帧基础上进一步引入点云库来构建可直观展现场景的 3D 点云地图。

仅使用关键帧位姿及其跟踪到的 3D 地图点构建稠密 3D 点云地图, 容易引入原始关键帧中属于动态物体的动态点云。同时, 如果在提取特征点和剔除动态特征点的过程中仅考虑符合 ORB 算法的特征点, 则对于原始关键帧中属于动态物体且不符合 ORB 算法的特征点及其对应的 3D 地图点, 需要进行处理以避免其出现在构建的稠密 3D 点云地图中, 从而导致不能精确地构建基于静态背景的稠密 3D 点云地图。因此, 本文算法在剔除动态特征点时即生成动态区间, 并整体剔除该区间的动态点云, 以消除动态物体对稠密 3D 点云地图的影响。由于稠密 3D 点云地图的构建是一个持续递进的过程, 因此即使当前关键帧下该动态区间的静态点云被误剔除, 随着动态物体离开该区间或在该区间暂时静止, 利用其他关键帧构建地图时也会补充该部分的静态点云, 从而顺利完成动态场景下 3D 稠密点云地图的构建。

2　仿真实验与结果分析

为验证本文算法的性能, 在公开数据集 TUM RGB-

D及室内真实动态场景中对算法进行仿真和实验测试。笔记本电脑配置为: CPU i5-4300 m, 内存8 GB, 运行环境为64位Ubuntu 16.04系统, 并安装机器人操作系统(robot operating system, ROS)。

2.1 TUM RGB-D数据集实验

TUM RGB-D数据集[16]为慕尼黑工业大学的实验场景数据集, 其主要针对室内场景, 并提供了由相机获取的RGB图像、深度图像及高精度传感器获取的相机真实轨迹。由于该数据集中walking系列的图像序列能够很好地描述动态程度较高的场景, 因此本文选用其图像序列作为输入来验证算法在动态场景下的性能。该系列的图像序列包括walking_xyz、walking_halfsphere、walking_rpy、walking_static 4个数据集, 分别代表沿*xyz*轴、半球轨迹、偏航-俯仰-滚转和静止4种不同类型的相机运动。除相机的自身运动外, 场景中还存在走动的行人, 这对SLAM系统的精度和鲁棒性构成了较大的挑战。TUM RGB-D数据集仿真实验分为2部分: 第1部分对ORB-SLAM2算法[17]与本文算法的动态特征点剔除效果进行比较分析, 第2部分对现有动态场景中的SLAM算法与本文算法的性能进行综合比较分析。

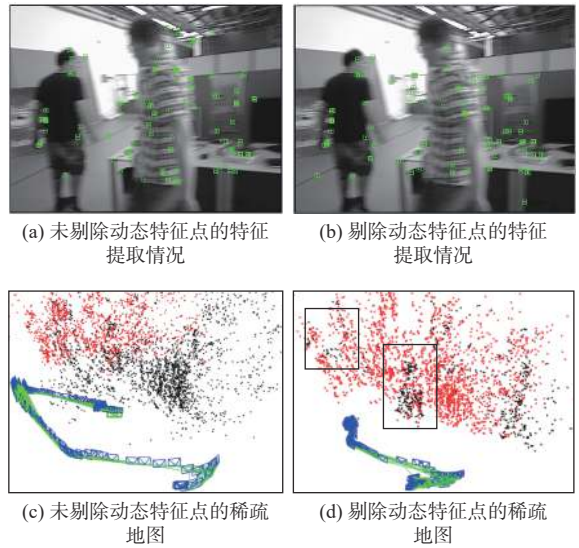
2.1.1 动态特征点剔除效果分析

为测试本文算法剔除动态特征点的性能, 使用walking_rpy图像序列作为输入, 分别在机器人相对静态背景运动及相对静态背景静止2种情况下测试算法对属于动态物体的动态特征点的剔除效果。

图7和图8分别为ORB-SLAM2算法和本文算法在第288帧(帧id为275)、421帧(帧id为405)的运行情况。其中, 第288帧为相机相对静态背景运动的动态场景, 第421帧为相机相对静态背景静止的动态场景。如图7(b)所示, 即使在第1种场景下, 前端也会从动态物体上提取特征点, 但由于本文算法能够有效检测出动态特征点, 如图7(d)所示, 黑色方框中被标记为黑色的离群点(红色为内点, 黑色为离群点), 因而完成了动态特征点的有效检测。而在第2种场景下, 如图8(b)所示, 由于前端部分避免了从动态物体中提取特征点, 所以有效检测出动态特征点。综上, 在2种动态场景下, 本文算法均能成功地将属于动态物体的地图点转换为离群点, 从而有效检测并剔除了动态特征点。

2.1.2 动态场景下算法性能分析

为验证本文算法在动态场景中的综合性能, 利用Grupp[18]开发的SLAM评估工具EVO, 以TUM RGB-D数据集中存在高动态物体的walking系列的图像序列作为SLAM系统的输入, 并且将绝对轨迹误差(absolute trajectory error, ATE)[16]和相对轨迹误差(relative pose error, RPE)[16]作为评价标准。

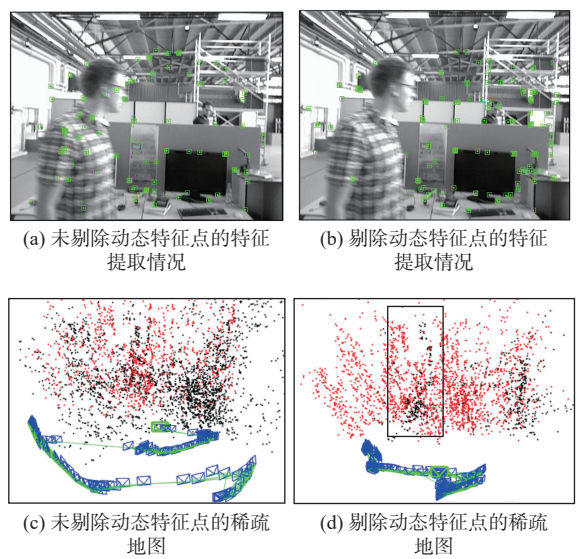


(a) 未剔除动态特征点的特征提取情况 (b) 剔除动态特征点的特征提取情况

(c) 未剔除动态特征点的稀疏地图 (d) 剔除动态特征点的稀疏地图

图7 机器人运动时的动态特征点剔除效果

Fig. 7 Effect of eliminating dynamic feature points when the robot is moving



(a) 未剔除动态特征点的特征提取情况 (b) 剔除动态特征点的特征提取情况

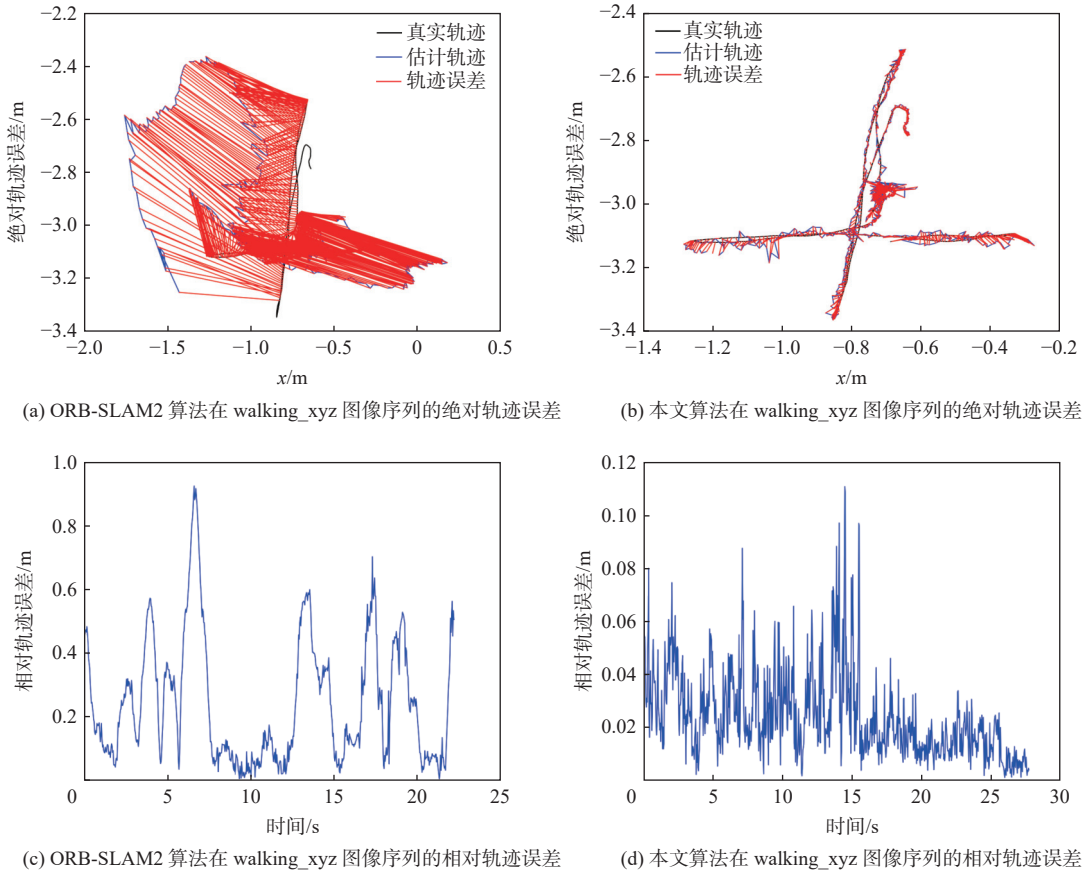
(c) 未剔除动态特征点的稀疏地图 (d) 剔除动态特征点的稀疏地图

图8 机器人静止时的动态特征点剔除效果

Fig. 8 Effect of eliminating dynamic feature points when the robot remains static

图9为ORB-SLAM2算法与本文算法以walking_xyz图像序列为输入的轨迹误差。从ORB-SLAM2算法和本文算法的绝对轨迹误差可以得出, 本文算法能够有效提高移动机器人在动态场景下的定位精度; 同时, 相对轨迹误差的结果也可以表明, 本文算法能够有效剔除动态特征点, 从而降低动态物体对特征连续跟踪时相机相对位姿变换的影响。

为利用绝对轨迹误差和相对轨迹误差2种评价标准更全面地评估本文算法的精度, 使用TUM RGB-D数据集walking系列的所有图像序列进行实验, 对比包括DVO (dense visual odometry and SLAM)[19]



(a) ORB-SLAM2 算法在 walking_xyz 图像序列的绝对轨迹误差　(b) 本文算法在 walking_xyz 图像序列的绝对轨迹误差

(c) ORB-SLAM2 算法在 walking_xyz 图像序列的相对轨迹误差　(d) 本文算法在 walking_xyz 图像序列的相对轨迹误差

图 9　以 walking_xyz 图像序列为输入的轨迹误差

Fig. 9　Trajectory error with walking_xyz image sequences as input

和 BaMVO (effective background model-based RGB-D dense visual odometry)[20] 等算法的相对轨迹误差, 以及与 ORB-SLAM2、半直接法[10]、Detect-SLAM[21] 和 DVO+MR[7] 等算法的绝对轨迹误差, 对比对象包括基于 TUM RGB-D 数据集 walking 系列图像序列的实验数据及近年相关文献的数据, 如表 1～表 3 所示。

表 1、表 2 定量比较了 DVO、BaMVO、线特征法和本文算法以具有高动态场景的 walking 图像系列为输入的相对轨迹误差, 包括旋转和平移均方根误差 (root mean square error, RMSE)。可以看出, DVO 算法是基于静态背景的假设, 但未剔除属于动态物体的动态特征点, 当处在高动态环境中时, 由于动态物体的存在, 导致其相对轨迹误差较大; 在动态环境中, 本文算法明显优于 DVO 算法, 相较于 DVO 算法, 本文算法的平移 RMSE 平均降低了 79.8%, 旋转 RMSE 平均降低了 76.1%。但是, 由于 walking_halfsphere 图像序列代表的沿着半球轨迹的运动降低了本文算法中动态区间的稳定性, 且运动速度过快导致无法获取足够的特征点, 因此, 跟踪失败, 最终无法获取有效的轨迹误差。与 DVO 算法不同的是, BaMVO 算法根据运动物体边缘特征点的深度值差异来完成动态特征点的检测, 与 DVO 算法相比, BaMVO 算法实现了对部分动态特征的剔除, 在动态场景下的位姿估计精度有了一定提升, 但在高动态环境下, 位姿估计仍然存在较大的误差; 而本文方法的动态特征点剔除方法能够剔除大部分的动态特征点, 因此, 相较于 DVO 算法和 BaMVO 算法, 本文算法在 walking_static、walking_rpy 和 walking _xyz 这 3 组图像序列中均优于这 2

表 1　不同算法的相对平移轨迹误差对比

Table 1　Comparison of relative translation trajectory error among different algorithms　m

图像序列	平移RMSE					
	DVO	BaMVO	线特征法	半直接法	DVO+MR	本文算法
walking_static	0.381 8	0.133 9	0.023 4	0.010 2	0.084 2	0.042 4
walking_xyz	0.436 0	0.232 6	0.243 3	0.032 0	0.121 4	0.042 6
walking_rpy	0.430 8	0.358 4	0.156 0		0.175 1	0.171 7
walking_halfsphere	0.262 8	0.173 8	0.135 1	0.047 6	0.167 2	

表2 不同算法的相对旋转轨迹误差对比

Table 2 Comparison of relative rotation trajectory error among different algorithms (°)

图像序列	旋转RMSE					
	DVO	BaMVO	线特征法	半直接法	DVO+MR	本文算法
walking_static	6.350 2	2.083 3	1.854 7	0.252 5	2.048 7	0.600 5
walking_xyz	7.666 9	4.391 1	6.911 6	0.686 9	3.234 6	0.856 6
walking_rpy	7.066 2	6.339 8	5.580 9		4.375 5	3.602 1
walking_halfsphere	5.217 9	4.286 3	4.641 2	1.045	5.010 8	

表3 不同算法的绝对轨迹误差对比

Table 3 Comparison of absolute trajectory error among different algorithms m

图像序列	绝对轨迹误差				
	ORB-SLAM2	半直接法	DVO+MR	Detect-SLAM	本文算法
walking_static	0.430 0	0.008 0	0.065 6		0.007 9
walking_xyz	0.620 2	0.037 1	0.093 2	0.024 1	0.023 1
walking_rpy	0.668 9		0.133 3	0.295 9	0.317 4
walking_halfsphere	0.323 1	0.040 9	0.125 2	0.051 4	

种算法。另外，线特征法的平移RMSE比本文算法稍好，但由于其对旋转变换的鲁棒性不足，旋转误差较大，因此，本文算法的旋转RMSE明显优于线特征法。相对半直接法，虽然在walking_xyz和walking_static图像序列下，本文算法的精度稍弱于该算法，但二者是在同一个数量级下的，而在剩余2个图像序列中，由于本文算法及半直接法均无法获取有效的相对轨迹误差，因此无法直接进行比较。较于DVO+MR算法，得益于本文算法在剔除动态特征点方面的有效性，除了不能有效获取walking_halfsphere图像序列的相对轨迹误差，在其他3组图像序列中，本文算法的旋转RMSE明显优于线特征法。

在walking图像系列所表示的高动态场景中，属于动态物体的动态特征点较多，即使ORB-SLAM2算法在位姿优化时使用了鲁棒核函数，但由于其在位姿估计时将属于动态物体的特征点当作静态特征点处理，因此估计的位姿也会产生较大的误差。由于本文算法在前端部分实现了动态特征点的检测与剔除，因此相较于ORB-SLAM2算法，本文算法精度平均提升了82.3%，如表3所示。

相较于半直接法和DVO+MR算法，由于本文算法获取动态区间的局限性，因此非xyz相机运动类型与xyz相机运动类型的误差数量级过大，导致其在walking_rpy及walking_halfsphere这2组图像序列中的效果不佳。但是，由于在walking系列中的walking_static数据集和walking_xyz数据集中，且相机运动为*xyz*方向的运动，所以本文算法能够获取稳定的动态区间，从而能够有效剔除属于动态物体的动态特征点。由表3可以看出，相较于其他4种算法，本文算法在walking_static和walking_xyz数据集中表现出最佳效果，而walking_halfsphere图像序列中相机沿着半球轨迹运动速度过快，导致无法获取足够的特征点，因此跟踪失败，最终无法获取有效的轨迹误差。

2.2 动态场景下移动机器人实验测试

为验证本文算法在实际动态场景中的可行性和有效性，利用“三宝”移动机器人搭载深度相机Letv Pro Xtion开展实验测试，该相机尺寸为18 cm×3.5 cm×5 cm，包括RGB摄像头、红外结构光发射器和接收器3个主要部件，有效测距范围为0.8～3.5 m，RGB-D摄像机视场角的水平视场角为58°、垂直视场角为45°、对角视场角为70°。实验中，深度相机采集的RGB-D图像信息通过USB接口传输到电脑端作为实时输入，后输出为移动机器人的运动轨迹和构建的环境3D地图。场景中，行人被用于模拟动态物体，速度约为0.5 m/s，并在机器人静止和运动2种状态下验证算法的性能。

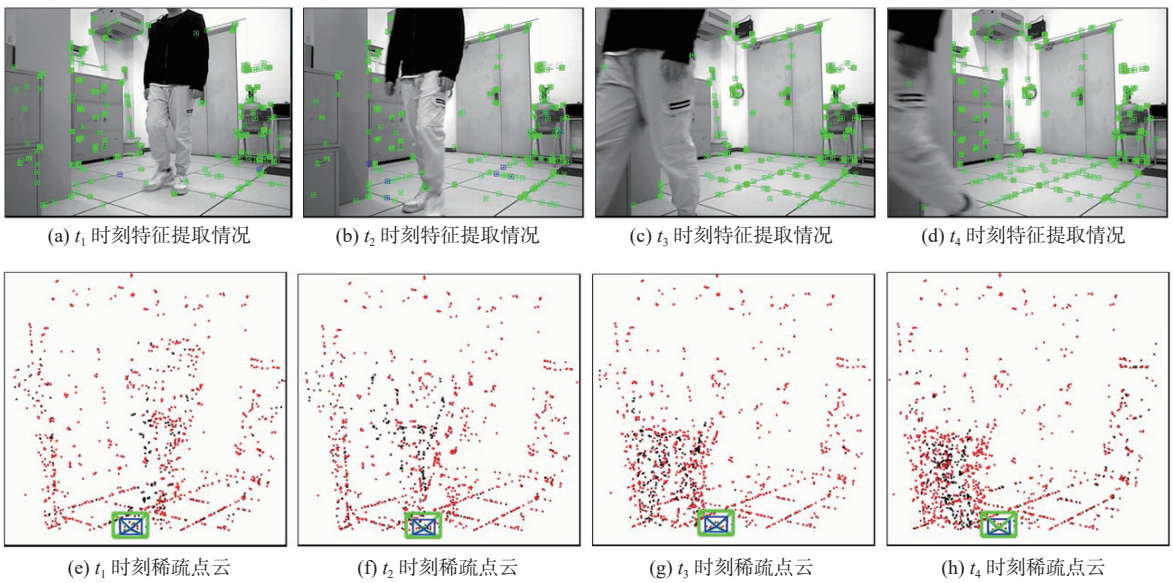
当机器人静止时，图10(a)～图10 (d)分别表示机器人静止时的t_1～t_4时刻的特征提取情况，可以看出，前端部分没有提取到运动物体的特征，并且动态物体没有对相机的位姿产生影响，如图10(e)～图10 (h)所示，稀疏地图中表示当前帧的绿色方框的位姿没有发生明显变化。当机器人运动时，图11(a)～图11 (d)分别表示机器人运动时的t_5～t_8时刻的特征提取情况，可以看出，除图11(c)外，前端部分都没有提取到运动物体的特征，即使图11(c)提取到了少量的特征点，本文算法仍能将动态物体上特征点对应的地图点剔除，如图11(g)所示，稀疏地图中，动态物体中的特征点对应地图点被本文算法检

测出来, 并被标记为黑色的离群点。因此, 本文算法能够实时剔除动态场景中的动态特征, 从而提高了 SLAM 系统在动态场景中的精度和鲁棒性。

为进一步验证本文算法在实际动态场景中的综合性能, 分别针对本文算法和 ORB-SLAM2 算法在 6.4 m×8.4 m 的办公室进行移动机器人运动回环实验, 如图 12 所示, 其中, 红色虚线回路为设定的大小为 1.2 m×4.2 m 的参考运动轨迹。实验中, 通过行人在机器人在视距范围内来回走动来模拟动态环境。

图 13 比较了 2 种算法的输出轨迹。由于 ORB-SLAM2 算法不能检测并剔除动态特征点, 因此当机器人视距范围内有行人走过时, 移动机器人的位姿估计过程会引入较大的误差, 从而导致其输出轨迹与设定的移动机器人参考运动轨迹产生较大的误差。本文算法则能够有效剔除动态特征点, 预估的移动机器人运动轨迹能够呈现出与设定参考轨迹相似度更高的矩形轮廓, 因此, 本文算法得到的定位的精度更高。

图 14 对比了上述实验中本文算法和 ORB-SLAM2 算法跟踪每帧图像的平均耗时。可以看出, ORB-SLAM2 算法的平均耗时为 53.35 ms, 本文算法的平均耗时为 64.27 ms, 由于本文算法在前端加入了动态特征点剔除, 虽然比 ORB-SLAM2 算法

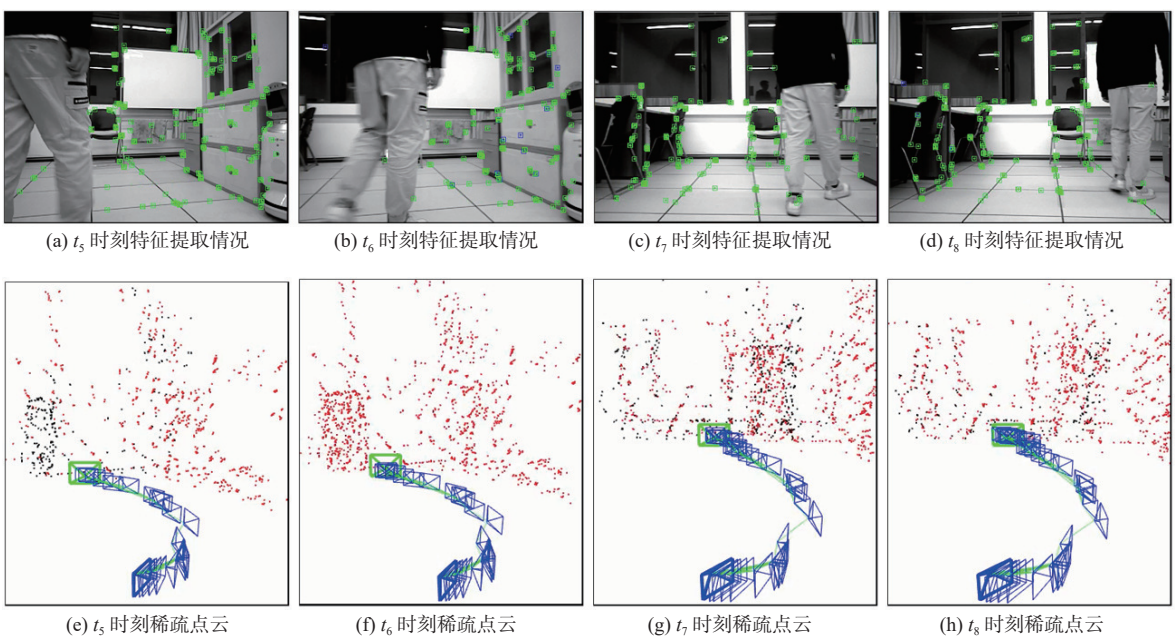


(a) t_1 时刻特征提取情况　(b) t_2 时刻特征提取情况　(c) t_3 时刻特征提取情况　(d) t_4 时刻特征提取情况

(e) t_1 时刻稀疏点云　(f) t_2 时刻稀疏点云　(g) t_3 时刻稀疏点云　(h) t_4 时刻稀疏点云

图 10　真实场景下机器人静止时的动态特征点剔除效果

Fig. 10　Effect of eliminating dynamic feature points when the robot remains static in real environment



(a) t_5 时刻特征提取情况　(b) t_6 时刻特征提取情况　(c) t_7 时刻特征提取情况　(d) t_8 时刻特征提取情况

(e) t_5 时刻稀疏点云　(f) t_6 时刻稀疏点云　(g) t_7 时刻稀疏点云　(h) t_8 时刻稀疏点云

图 11　真实场景下机器人运动时的动态特征点剔除效果

Fig. 11　Effect of eliminating dynamic feature points when the robot is moving in real environment

在跟踪图像时耗时较长，但系统运行平稳，所以仍然能够满足实时性的要求。

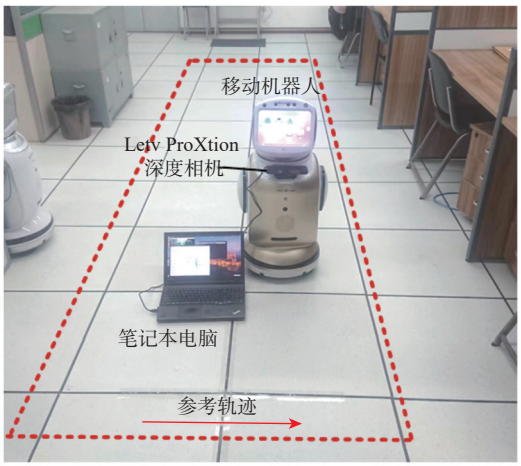


图 12 实验场景及机器人运动参考轨迹

Fig. 12 Experimental scenario and reference trajectory of mobile robot

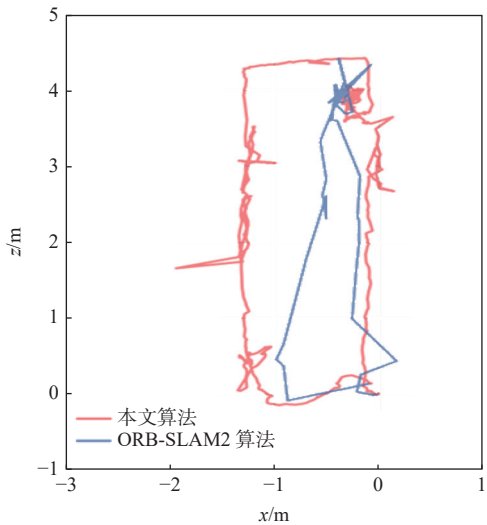


图 13 动态场景下 2 种算法的轨迹比较

Fig. 13 Comparison of trajectory for two algorithms in dynamic scenarios

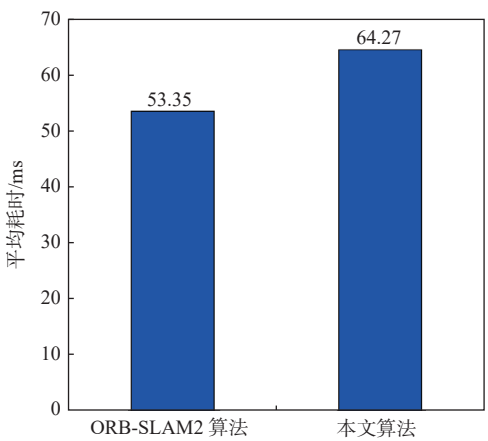


图 14 ORB_SLAM2 算法与本文算法的实时性比较

Fig. 14 Real-time performance comparison between ORB_SLAM2 algorithm and the proposed algorithm

为展示 3D 地图的构建效果，利用生成的关键帧及其对应的位姿进行点云的拼接，构建了动态场景的 3D 稠密地图，如图 15 所示。图中，红色方框为利用关键帧及其位姿进行点云拼接时，利用生成的关键帧标记的动态区间，以剔除属于动态物体的点云后残留的动态点云。可以看出，经过本文算法的处理后，属于动态物体的大部分点云被成功剔除，整个地图与真实环境一致，这进一步说明本文算法能够有效获取存在动态特征点的动态区间，因此，可在构建稠密地图时剔除关键帧中位于动态区间的点云，从而构建出基于静态背景的稠密地图。

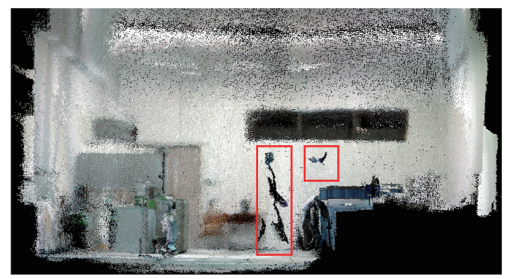


图 15 动态场景下的稠密地图

Fig. 15 Dense map in dynamic scenarios

3 结 论

1）本文提出了一种动态环境下基于特征点几何约束的移动机器人 3D SLAM 算法，在特征跟踪阶段引入 Delaunay 三角剖分法构建出能够表示相邻 2 帧匹配地图点几何关系的三角网，并利用当前帧特征点反投影生成的 3D 点来检测出动态特征点。

2）为避免误检测情况的发生，考虑到属于同一动态物体的动态特征点通常位于同一个深度区间，分段，确定动态物体的深度区间，并实现动态特征点的剔除，从而有效提高移动机器人在动态场景中位姿估计的精度。

3）在前端引入一个滑动窗口，并利用局部特性提取动态环境中局部最优的关键帧，以减少稀疏地图中关键帧的数量、降低关键帧之间的冗余度及后端优化的计算复杂度、构建高效的稀疏地图，提高了系统的实时性。

考虑到未知环境中结构化的场景存在大量线段特征，下一步工作中，将在点特征的基础上引入线段特征用作特征提取、位姿匹配及后端优化部分，以进一步提高本文算法在动态场景中的鲁棒性和精度。

参考文献（References）

[1] FISCHLER M A, BOLLES R C. Random sample consensus: A paradigm for model fitting with applications to image analysis and

automated cartography[M]//FISCHLER M A, FIRSCHEIN O. Readings in computer vision. Amsterdam: Elsevier, 1987: 726-740.

[2] STRASDAT H, MONTIEL J M M, DAVISON A J. Visual SLAM: Why filter?[J]. Image and Vision Computing, 2012, 30(2): 65-77.

[3] ZOU D P, TAN P. CoSLAM: Collaborative visual SLAM in dynamic environments[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(2): 354-366.

[4] 张慧娟, 方灶军, 杨桂林. 动态环境下基于线特征的 RGB-D 视觉里程计[J]. 机器人, 2019, 41(1): 75-82.
ZHANG H J, FANG Z J, YANG G L. RGB-D visual odometry in dynamic environments using line features[J]. Robot, 2019, 41(1): 75-82(in Chinese).

[5] RUBLEE E, RABAUD V, KONOLIGE K, et al. ORB: An efficient alternative to SIFT or SURF[C]//Proceedings of the International Conference on Computer Vision. Piscataway: IEEE Press, 2011: 2564-2571.

[6] VON GIOI R G, JAKUBOWICZ J, MOREL J M, et al. LSD: A line segment detector[J]. Image Processing On Line, 2012, 2: 35-55.

[7] SUN Y X, LIU M, MENG M Q H. Improving RGB-D SLAM in dynamic environments: A motion removal approach[J]. Robotics and Autonomous Systems, 2017, 89: 110-122.

[8] 艾青林, 刘刚江, 徐巧宁. 动态环境下基于改进几何与运动约束的机器人 RGB-D SLAM 算法[J]. 机器人, 2021, 43(2): 167-176.
AI Q L, LIU G J, XU Q N. An RGB-D SLAM algorithm for robot based on the improved geometric and motion constraints in dynamic environment[J]. Robot, 2021, 43(2): 167-176(in Chinese).

[9] LI S L, LEE D. RGB-D SLAM in dynamic environments using static point weighting[J]. IEEE Robotics and Automation Letters, 2017, 2(4): 2263-2270.

[10] 高成强, 张云洲, 王晓哲, 等. 面向室内动态环境的半直接法 RGB-D SLAM 算法[J]. 机器人, 2019, 41(3): 372-383.
GAO C Q, ZHANG Y Z, WANG X Z, et al. Semi-direct RGB-D SLAM algorithm for dynamic indoor environments[J]. Robot, 2019, 41(3): 372-383(in Chinese).

[11] DEROME M, PLYER A, SANFOURCHE M, et al. Real-time mobile object detection using stereo[C]//Proceedings of the 13th International Conference on Control Automation Robotics & Vision. Piscataway: IEEE Press, 2014: 1021-1026.

[12] LI X Z, XU C L. Moving object detection in dynamic scenes based on optical flow and superpixels[C]//Proceedings of the IEEE International Conference on Robotics and Biomimetics. Piscataway: IEEE Press, 2015: 84-89.

[13] JAIMEZ M, KERL C, GONZALEZ-JIMENEZ J, et al. Fast odometry and scene flow from RGB-D cameras based on geometric clustering[C]//Proceedings of the IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2017: 3992-3999.

[14] DAI W C, ZHANG Y, LI P, et al. RGB-D SLAM in dynamic environments using point correlations[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(1): 373-389.

[15] BARBER C B, DOBKIN D P, HUHDANPAA H. The quickhull algorithm for convex hulls[J]. ACM Transactions on Mathematical Software, 1996, 22(4): 469-483.

[16] STURM J, ENGELHARD N, ENDRES F, et al. A benchmark for the evaluation of RGB-D SLAM systems[C]//Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2012: 573-580.

[17] MUR-ARTAL R, TARDÓS J D. ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras[J]. IEEE Transactions on Robotics, 2017, 33(5): 1255-1262.

[18] GRUPP M. EVO: Python package for the evaluation of odometry and SLAM[EB/OL]. (2017-09-14)[2022-04-28]. https://michaelgrupp.github.io/evo/.

[19] KERL C, STURM J, CREMERS D. Robust odometry estimation for RGB-D cameras[C]//Proceedings of the IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2013: 3748-3754.

[20] KIM D H, KIM J H. Effective background model-based RGB-D dense visual odometry in a dynamic environment[J]. IEEE Transactions on Robotics, 2016, 32(6): 1565-1573.

[21] ZHONG F W, WANG S, ZHANG Z Q, et al. Detect-SLAM: Making object detection and SLAM mutually beneficial[C]//Proceedings of the IEEE Winter Conference on Applications of Computer Vision. Piscataway: IEEE Press, 2018: 1001-1010.

3D SLAM algorithm based on geometric constraints of feature points in dynamic scenarios

LIU Mingjian[1], LUO Jingwen[1,*], QIN Shiyin[2]

(1. School of Information Science and Technology, Yunnan Normal University, Kunming 650500, China;
2. School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China)

Abstract: The dynamic objects will cause a large number of dynamic errors in the pose estimation of robots in dynamic scenarios. To address this issue, a 3D simultaneous localization and mapping (SLAM) algorithm for mobile robots was presented by using geometric constraints between feature points to eliminate dynamic feature points. First, the ORB feature points of the current frame and the map points generated by feature points of the previous frame were used for projection matching, and the Delaunay triangulation method was introduced to construct a triangle net that could represent the geometric relationship between the matching map points of the two frames. Then, the dynamic feature points were detected according to the geometric relationship changes of the map points in the adjacent two frames. Since the static feature points may be incorrectly detected as dynamic feature points, which thus brings about the loss of feature points, more feature points were extracted during the matching of the adjacent two frames, so as to compensate for static feature points. Then, the dynamic feature points were eliminated, and the pose of the mobile robots was estimated accurately. On this basis, a sliding window was introduced to extract key frames and complete closed-loop detection, and thus an accurate 3D dense map was constructed. The results of simulation experiments on multiple sets of public datasets and the experiments in the indoor dynamic scenarios show that the proposed algorithm in this paper can effectively eliminate the dynamic feature points and improve the accuracy of the pose estimation of mobile robots in dynamic scenarios and the consistency of the map.

Keywords: dynamic scenarios; simultaneous localization and mapping; mobile robot; geometric constraints; triangulation

Received: 2022-08-17; **Accepted**: 2022-09-21; **Published Online**: 2022-10-08 18:33

URL: link.cnki.net/urlid/11.2625.V.20221006.1321.002

Foundation items: National Natural Science Foundation of China (62063036); Yunnan Normal University Doctoral Research Initiation Project (01000205020503115)

* **Corresponding author**. E-mail: by1503117@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0741

大规模卫星星座组网的码分多址干扰分析

徐可笛[1,2]，徐兆斌[1,2,*]，郭晓旭[1]，杨嘉[1]，成恒飞[1]，金仲和[1,2]

(1. 浙江大学 微小卫星研究中心，杭州 310027；2. 浙江大学 浙江省微纳卫星研究重点实验室，杭州 310027)

摘　　要：码分多址(CDMA)的通信体制是大规模卫星星座组网的主要部分，针对星间多通道通信存在多址干扰的情况，对其接收链路建立模型，描述灵敏度与比特能量噪声密度比(BENR)、设备采样信噪比(SNR)的关系，并分析多址干扰和加性白噪声对SNR的影响。在MATLAB仿真不同接收、干扰功率时的BENR与灵敏度，仿真结果表明，信干比(SIR)大于0 dB时，干扰信号对灵敏度影响较小，而维持高灵敏度时可抵抗−6 dB干扰。建立3个发射器和1个接收器的实验平台，由Chipscope、MATLAB软件处理数据，得到符合实际应用需求的接收灵敏度，实验结果表明，SIR大于0 dB时，接收灵敏度可达−107 dBm，高灵敏度下可抵抗−9 dB干扰，误差影响因素为电路噪声。理论分析与实测相符，为实际低轨卫星星座组网的设计奠定了基础。

关 键 词：星座；卫星通信；码分多址；多址干扰；灵敏度

中图分类号：V11

文献标志码：A　　**文章编号**：1001-5965(2024)09-2885-08

大规模低轨卫星星座组网由按照预定构型运行在不同轨道平面、高度和相位的数百、数千甚至上万颗卫星构成，凭借其全球覆盖、低时延、强信号、分布式载荷和成本优势等特点，将会成为未来全球网络和6G的关键技术[1-2]。国内外针对低轨卫星通信发展提出了较多方案，当下实际运营并真正发射的主要是铱、全球星和北斗卫星导航系统等[3]。未来，全球范围内的星座组网计划有Starlink、OneWeb、Lightspeed、Kuiper和Ether等。卫星组网已上升为中国的国家战略性工程，代表性的有科研单位的“鸿雁”“虹云”“行云”和“天象”等计划，商用公司的“银河”“天基互联网”和“微厘空间”等计划[1]。

在大规模卫星星座组网中，星间的频谱资源受限，星上频率协调复杂[4]，根据国际电信联盟“先登先占”的原则，同类星座频率的申报将越来越困难[5]，传统的“一对一”星间通信已不能适应大规模星座编队的应用需求，集成度高且适用于大规模低轨卫星星座组网的多通道通信系统亟待开发。

为实现多星编队的高效接入，将现有的接入方式分为竞争类[6-7]、非竞争类[8]和混合类[9]。竞争类有各节点自主访问信道的ALOHA协议[10-11]和采用IEEE 802.11协议分布式协调工作模式的无线局域网[12]；混合类有竞争与时分复用混合的LDMA协议[13]和竞争与码分多址(code division multiple access, CDMA)混合的通信系统[14]；实时性要求高的业务通常采用非竞争类协议，如CDMA[15-16]、时分多址、频分多址和正交频分复用[17]。

大规模星座通常采用CDMA的通信方式[18-19]，如全球星，其轨道高度约1 400 km，工作频段为L/S[20]，接收机并行接收多个卫星信号，相较于其他接入方式，该方式可使用较少的物理层信道带宽，无须时间同步，能避免时隙划分带来的逻辑信道带宽减少，具有抗截获能力。采用CDMA通信方式的优势有：星间链路通信设备较为对称，实现该方案的成本较低；测量和通信可同时进行；兼容中国

收稿日期：2022-08-29；**录用日期**：2022-11-19；**网络出版时间**：2022-12-15 17:22
网络出版地址：link.cnki.net/urlid/11.2625.V.20221214.1447.006
基金项目：国家自然科学基金面上项目(62073289)；国家自然科学基金(U21A20443)
*通信作者. E-mail：zjuxzb@zju.edu.cn

引用格式：徐可笛，徐兆斌，郭晓旭，等. 大规模卫星星座组网的码分多址干扰分析[J]. 北京航空航天大学学报，2024，50(9)：2885-2892. XU K D，XU Z B，GUO X X，et al. Analysis of CDMA interference in large-scale satellite constellation[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2885-2892 (in Chinese).

测控网，且功率谱密度低。但是，编队内最大卫星节点个数受CDMA系统自干扰特性的限制。

多颗卫星信号同时接收时，各卫星接收信噪比(signal-to-noise ratio, SNR)随其至接收机的距离方位、传输时间而变化，对于某码组的卫星，其余码组信号均为干扰，需规避多颗卫星信号的干扰，以保持接收的灵敏度。本文对星座组网通信过程中多路卫星信号间的干扰进行分析，建立低轨信道接收链路模型，阐述多址干扰和加性白噪声对SNR的影响，反映灵敏度与比特能量噪声密度比(bit energy to noise density ratio, BENR)、设备采样SNR的关系，为低轨卫星编队组网的实际设计奠定了基础。

1 卫星通信的码分多址原理

一般的大规模星座多通道通信系统由多块通信板组成，每块通信板均采用CDMA方式，接收多路卫星信号，通信接收机则通过ADC采样扩频信号，输入多组不同卫星Gold码组的接收模块，解扩、解调并译码，如图1所示。

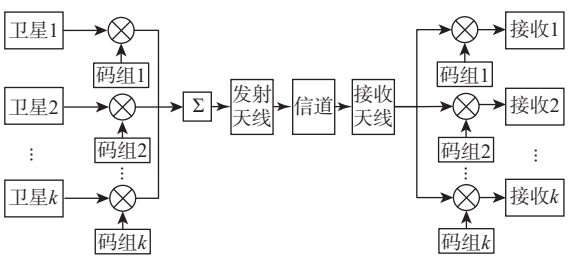


图1 码分多址系统结构

Fig. 1 System structure of code division multiple access

在大规模低轨卫星星座组网中，主要的信道特性包含自由空间损耗、多普勒频偏和多径衰落等。自由空间损耗是指电磁波在传播过程中没有任何障碍的空间，不存在吸收、折射、反射和散射等现象时产生的损耗，比如在星间链路中，电离层闪烁会造成附加的路径损耗；多普勒频偏是卫星发射信号和接收信号的频率之差，卫星同轨时频偏较低，异轨时相对速度较快，在L和S波段的频偏分别约为±50 kHz和70 kHz，因此，通信机需满足动态的多普勒范围[21]；多径衰落是电磁波在星间传输时，受星体散射、绕射、折射等的影响，信号沿各种不同的路径传输，单个码元到达接收端的时刻不同，造成的相邻码元的串扰现象[22]。

通信机的接收分为伪码搜索、伪码捕获确认和系统锁定3个状态，如图2所示，本文主要针对伪码的搜索和捕获过程展开分析。

星座编队的多通道通信机大多采用载波伪码并行搜索的软件结构，如图3所示。伪码搜索状态时，信号经下混频器下变频到基带信号，低通滤波器滤除其混频产生的二倍频分量，与本地伪码关联

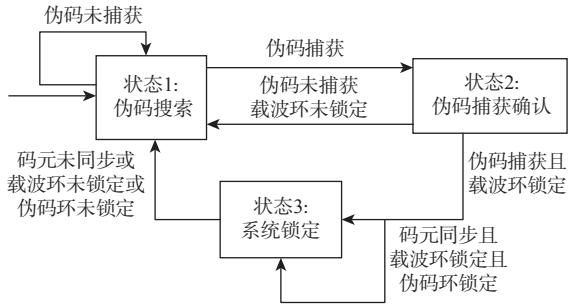


图2 接收端的状态转换

Fig. 2 State transition of receiving end

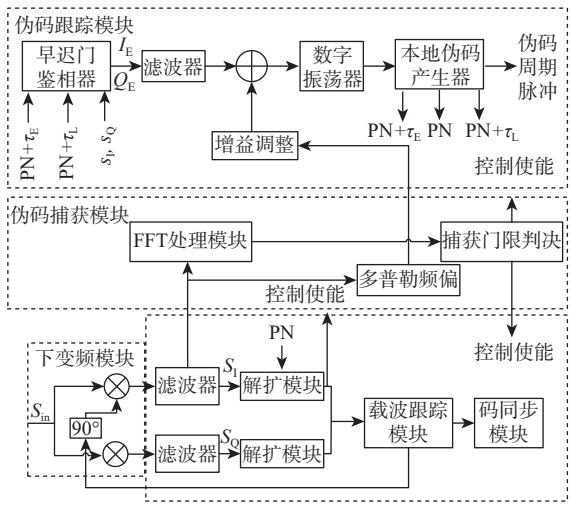


图3 接收端的软件结构

Fig. 3 Software structure of receiving end

后，经积分抖落滤波器将数据速率降至快速傅里叶变换(fast Fourier transform, FFT)处理速率。对数据作FFT等相关运算，并与阈值比较，若未超过阈值，则移动本地伪码相位，继续前述步骤运算，若超过门限则认定为伪码捕获。伪码捕获后，载波跟踪环和伪码跟踪环闭合。伪码跟踪环包括早迟门鉴相器、环路滤波器、锁定检测器、本地伪码产生器，分别用超前、滞后的两路伪码与信号相关鉴相，在本地恢复出伪码后给解扩模块。解扩信号进入载波跟踪环，载波跟踪环锁定后，码同步模块开启，给出码元同步标志和时钟。

2 理论分析

2.1 捕获灵敏度

接收信号的捕获灵敏度P_s的计算公式如下：

$$P_s = -174 + N_f + P_r + B_n + L_w + L_c + L_p \tag{1}$$

式中：−174 dBm/Hz为大自然基地噪声；N_f为电路噪声系数；P_r为设备接收SNR；B_n为噪声带宽；L_w为FFT扇贝损失，即频率检测点的误差；L_c为伪码相关损失；L_p为降速率滤波损失。

B_n的表达式为

$$B_n = 10\lg f_\Delta \tag{2}$$

式中: f_Δ为伪码搜索阶段 FFT 的分辨率, 即伪码搜索间隔频率。

L_c表示伪码多普勒频偏引起的时延损失, 如下:

$$L_c = 20\lg\left(1-\frac{\xi}{2}-\frac{f_b T_c|D_f|}{2}\right) \tag{3}$$

式中: ξ为相位搜索间隔比例系数; f_b为伪码符号速率; T_c为积分相关时间; D_f为多普勒频偏系数。

L_p的表达式如下:

$$L_p = 20\lg\left(\frac{\sin(N_p f_d\pi/f_b)}{N_p\sin(f_d\pi/f_b)}\right) \tag{4}$$

式中: N_p为降速率滤波器所取点数; f_d为多普勒频偏。

2.2　单通道接收链路模型

一般的大规模星座编队均采用近似共面轨道, 微波信号在近真空的自由空间内传输, 可用 BENR 描述信道质量, 并对其接收链路建立模型, 如下:

$$\frac{E_b}{N_0}=\frac{C/R}{kT_s}=\frac{P_t L_l G_t(\alpha_1,\beta_1) L_s L_a G_r(\alpha_2,\beta_2)}{kT_s R} \tag{5}$$

式中: E_b为比特能量; N_0为噪声谱密度; C为接收载波功率; R为码元传输速率; k为玻尔兹曼常数; T_s为系统噪声温度; P_t为发射功率; L_l为功放到发射天线的路径损耗; α为天线的方位角; β为天线的俯仰角; $G_t(\alpha_1,\beta_1)$为发射天线的增益; L_s为自由空间传播的损耗; L_a为传播路径的损耗; $G_r(\alpha_2,\beta_2)$为接收端天线的增益。

L_s与自由空间传输路径和频率有关, 如下:

$$L_s=\left(\frac{c}{4\pi d f}\right)^2 \tag{6}$$

式中: d为通信距离; f为信号载波频率; c为常数。

设备接收端的 SNR 需要达到一定值才能满足实际应用对误码率(bit error rate, BER)的要求。一般, 接收机 BENR 与 SNR 的关系如下:

$$\frac{E_b}{N_0}=P_r\frac{B_R}{R} \tag{7}$$

式中: B_R为射频载波带宽。

通信链路中接收机的 SNR 与收发天线的增益和通信距离等因素有关, 而大规模星座编队中收发天线的增益、通信速率和通信距离等相对固定, 因此, 多址干扰将对 BENR 产生主要影响。

2.3　多通道接收链路模型

假设有 K 个通道同时接入, 环路信号如下:

$$s_i(t)=\sum_{i=1}^{K}A_i b_{i,n}c_i(t-nT_b-\tau_i)\cos(\omega t+\theta_i)+n(t) \tag{8}$$

式中: A_i为i通道的信号幅值; $b_{i,n}$为通道i调制的第n个码元符号; $c_i(t-nT_b-\tau_i)$为通道i的波形, T_b为码元持续时间, τ_i为传输时延; $n(t)$为加性白噪声, 其双边带功率谱密度为$N_0/2$; ω为载波频率; θ_i为传输延时引起的该通道载波相移。

当通道 k 的信号被成功捕获, 经过正交下混频和滤除倍频项后得到的信号形式为

$$s_I=\frac{1}{2}A_k b_{k,n}c_k(t-nT_b-\tau_k)\cos\theta_e+\frac{1}{2}\sum_{\substack{i=1\\ i\neq k}}^{K}A_i b_{i,n}c_i(t-nT_b-\tau_i)\cos\varphi_i+n_I \tag{9}$$

$$s_Q=\frac{1}{2}A_k b_{k,n}c_k(t-nT_b-\tau_k)\sin\theta_e+\frac{1}{2}\sum_{i=1,i\neq k}^{K}A_i b_{i,n}c_i(t-nT_b-\tau_i)\sin\varphi_i+n_Q \tag{10}$$

式中: n_I、n_Q为正交混频后的 I/Q 路噪声; θ_e为当前通道k载波环锁定时的相位误差, 接近 0; φ_i为通道i相位θ_i与通道k相位θ_k之差。

若信号在伪码环中经过了一个码元延时T_b, 则在第$n-1$个码元相关时间内, 超前路的信号经相关积分、归一化后, 超前 I/Q 支路的相关器输出I_E、Q_E为

$$I_E=\frac{1}{T_b}\int_{(n-1)T_b}^{nT_b}s_I(t-T_b)c_k(t-nT_b-\tau_k+\tau_E)\mathrm{d}t \tag{11}$$

$$Q_E=\frac{1}{T_b}\int_{(n-1)T_b}^{nT_b}s_Q(t-T_b)c_k(t-nT_b-\tau_k+\tau_E)\mathrm{d}t \tag{12}$$

式中: τ_E为本通道超前的相关时间, 取伪码码片的持续时间T_b的一半。

化简后可得

$$I_E=a_k+\sum_{i=1,i\neq k}^{K}\xi_{ik}+n_{IE} \tag{13}$$

$$n_{IE}=\frac{1}{T_b}\int_{(n-1)T_b}^{nT_b}n(t)c_k(t-(n+1)T_b-\tau_k+\tau_E)\cos(\omega t)\mathrm{d}t \tag{14}$$

$$a_k=\frac{1}{2}A_k b_{k,n}R_{kk}(\tau_E)\cos\theta_e \tag{15}$$

$$R_{kk}(\tau_E)=\frac{1}{T_b}\int_{(n-1)T_b}^{nT_b}c_k(t)c_k(t+\tau_E)\mathrm{d}t \tag{16}$$

$$\xi_{ik}=\frac{1}{2}A_i\alpha_i(\tau_E+\Delta\tau_i)\cos\varphi_i \tag{17}$$

式中: n_{IE}为相关积分I_E路噪声; a_k为有用信号; $R_{kk}(\tau_E)$为通道k归一化自相关结果; ξ_{ik}为通道i对通道k的干扰项; $\Delta\tau_i$为通道i和通道k之间由传播时延引起的伪码相位差$\tau_i-\tau_k$; $\alpha_i(\tau_E+\Delta\tau_i)$为与信息码元符号有关的互相关积分结果, 是多址干扰的来源。

由于客观存在不同通道间的时延, 所以通道i的相邻码元各有一段会进入通道k的积分区间, 如图 4 所示, 则对干扰通道i的积分结果为

$$\alpha_i(\tau_E+\Delta\tau_i)=\frac{1}{T_b}\int_{(n-1)T_b}^{\Delta\tau_i+(n-1)T_b}b_{i,n-2}c_i(t-(n+2)T_b-\tau_k+\tau_E)\mathrm{d}t+\frac{1}{T_b}\int_{\Delta\tau_i(n-1)T_b}^{nT_b}b_{i,n-2}c_i(t-(n+2)T_b-\tau_k+\tau_E)\mathrm{d}t \tag{18}$$

若$b_{i,n-2}$与$b_{i,n-1}$的符号相同, 则积分结果由两通道周期性互相关所得, 即$R_{ik}(\tau_E+\Delta\tau_i)$。若符号相

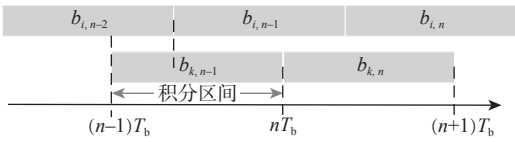


图4 多址干扰示意图

Fig. 4 Schematic diagram of multiple access interference

反，则积分结果为两通道分段后分别与正反序列相关所得循环相关函数$R_{ik}(\tau_E+\Delta\tau_i)+R_{ik}(\tau_E+\Delta\tau_i)$。

2.3.1 多址干扰分析

星座组网中的接收机多采用Gold码序列，伪码同步后互相关峰值有3种可能[23]：

$$R_{ik}\in\left\{\frac{\beta(n)-2}{N},\frac{-1}{N},\frac{-\beta(n)}{N}\right\} \tag{19}$$

$$\beta(n)=1+2^{\lfloor(n+2)/2\rfloor} \tag{20}$$

式中：n为优选m序列级数；N为伪码周期所包含的码片数；$\lfloor a\rfloor$为小于等于a的最大整数。

若$\Delta\tau_i$在一个码元周期之内均匀分布，则不同码组的互相关峰值最大为$(\beta(n)-2)/N$，且R_{ik}三值分布概率为0.125, 0.75, 0.125，可得

$$E\left(R_{ik}(\tau_E+\Delta\tau)\right)=\frac{-1}{N} \tag{21}$$

$$D\left(R_{ik}(\tau_E+\Delta\tau)\right)=\frac{1}{8}\left(\frac{\beta(n)-1}{2^n-1}\right)^2+\frac{1}{8}\left(\frac{-\beta(n)+1}{2^n-1}\right)^2=\frac{1}{4}\left(\frac{2^{(n+2)/2}}{2^n-1}\right)^2=\frac{2^n}{(2^n-1)^2} \tag{22}$$

式中：$E(R_{ik}(\tau_E+\Delta\tau))$为$R_{ik}(\tau_E+\Delta\tau)$均值；$D(R_{ik}(\tau_E+\Delta\tau))$为$R_{ik}(\tau_E+\Delta\tau)$均方差。由于$n$较大，$2^n\gg1$，循环自相关的方差可近似为$1/2^n=1/N$。由于大样本数据仿真得到的正负互相关和周期互相关的统计特性一致，均值为0，方差为$1/N$，则

$$D(\alpha_i(\tau_E+\Delta\tau))=\frac{1}{N} \tag{23}$$

伪码周期长、N较大时，$E(R_{ik}(\tau_E+\Delta\tau))$为0，则

$$E(\alpha_i^2(\tau_E+\Delta\tau_i))=D(\alpha_i(\tau_E+\Delta\tau_i))+E(\alpha_i(\tau_E+\Delta\tau_i))^2=\frac{1}{N} \tag{24}$$

载波相位差φ_i一般在T_b内服从$[0,2\pi]$的均匀分布，$E(\cos\varphi_i)=0$，$E(\cos^2\varphi_i)=D(\cos\varphi_i)=1/2$。故干扰项的方差为

$$D(\xi_{ik})=E(\xi_{ik}^2)=\frac{A_i^2}{4}E(\alpha_i^2(\tau_E+\Delta\tau_i))E(\cos^2\varphi_i)=\frac{A_i^2}{8N} \tag{25}$$

此时，对I_E取期望得到的有用信号幅值：

$$E(I_E)=E\left(\frac{1}{2}A_kb_{k,n-1}R_{kk}(\tau_E)\cos\theta_e\right)+E\left(\sum_{i=1,i\neq k}^{K}\xi_{ik}\right)+E(n_{IE})=\frac{1}{2}A_kb_{k,n-1}E(R_{kk}(\tau_E))E(\cos\theta_e) \tag{26}$$

由于载波和伪码之间相互独立，因此$E(\cos\varphi_i)=0$时，$E(\xi_{ik})=0$，且加性白噪声均值为0。假设载波环已经跟踪，θ_e很小，则$E(\cos\theta_e)=1$。相关器间隔为T_{corr}，超前自相关的输出期望$R_{kk}(\tau_E)=0.5$，化简可得

$$E(I_E)=\frac{1}{4}A_kb_{k,n-1} \tag{27}$$

联立式(25)和式(27)，可将其他通道对通道k的干扰，记作S/I：

$$\frac{S}{I}=\frac{E(I_E)^2}{\sum_{i=1,i\neq k}^{K}D(\zeta_i)}=\frac{N}{2}\cdot\frac{A_k^2}{\sum_{i=1,i\neq k}^{K}A_i^2} \tag{28}$$

式中：ζ_i为第i个干扰信号。

2.3.2 加性白噪声分析

中心极限定理求出$E(n_{IE})=0$，其方差为

$$D(n_{IE})=E\left(n_{IE}^2\right)=E\left(\frac{1}{T_b^2}\int_{(n-1)T_b}^{nT_b}n^2(t)c_k^2(t)\cos^2(\omega t)\mathrm{d}t\right)=\frac{N_0}{2T_b^2}\int_{(n-1)T_b}^{nT_b}c_k^2(t_1)\cos^2(\omega_ct_1)\mathrm{d}t_1=\frac{N_0}{4T_b} \tag{29}$$

式中：$N_0/4T_b$为加性白噪声的自相关函数。

由式(27)和式(29)可得有用信号和加性噪声间的比例关系：

$$\frac{S}{N}=\frac{E(I_E)^2}{D(n_{IE})}=\frac{A_k^2T_b}{4N_0}=\frac{A_k^2}{4N_0R} \tag{30}$$

由式(28)和式(30)可得多通道接收机的总SNR为

$$\left(\frac{S}{N+I}\right)^{-1}=\left(\frac{S}{I}\right)^{-1}+\left(\frac{S}{N}\right)^{-1}=\frac{2\sum_{i=1,i\neq k}^{K}A_i^2+4NN_0R}{NA_k^2} \tag{31}$$

由式(7)和式(31)可得多通道时接收端BENR即比特能量E_b与噪声谱密度N_0+N_I之比为

$$\left(\frac{E_b}{N_0+N_I}\right)^{-1}=\left(\frac{C/R}{kT_s}\right)^{-1}+\left(\frac{C/R}{I/B_R}\right)^{-1} \tag{32}$$

式中：C为当前通道接收载波功率；I为接收其他通道的总功率。

根据二进制相移键控(binary phase shift keying, BPSK)的BER公式可得多通道接收BER：

$$P_e=\frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{E_b}{N_0+N_I}}\right) \tag{33}$$

3 仿真分析

根据一般大规模低轨卫星组网的情况，系统采用S频段，通信机多普勒频偏取±115 kHz以内，参数中伪码序列级数n取10，伪码周期所包含的码片数N为1 023，码片速率f_b采用10.23 MHz。利用MATLAB生成分析模型，满足一般低轨卫星快速

捕获要求的接收机典型参数设置如表 1 所示。表中, f_s 为系统时钟频率, f_D 为多普勒频偏, f_Δ 为频率偏移量。考虑到系统的通用性, 接收机的前端数字信号处理速率 f_s 采用 40 MHz, 伪码相关后, 积分抖落滤波和补零操作将数据采样率降为 1.22 kHz, 即 FFT 处理速率, 此时相关积分时间就是其倒数 0.409 6 ms。

表 1　接收机参数

Table 1　Receiver parameter

f_s/MHz	f_b/MHz	f_D/kHz	T_{corr}/ms	N_f/dBm	L_w/dB	B_R/kHz	f_Δ/kHz
40	10.23	115	0.409 6	6	0.9	20	1.22

代入分析后可得 B_n 为 33.9 dBm, L_c 为 3.8 dB, L_p 为 0.5 dBm, 一般捕获所需 SNR 取 16 dB, 由此得到捕获灵敏度为−112.8 dBm。

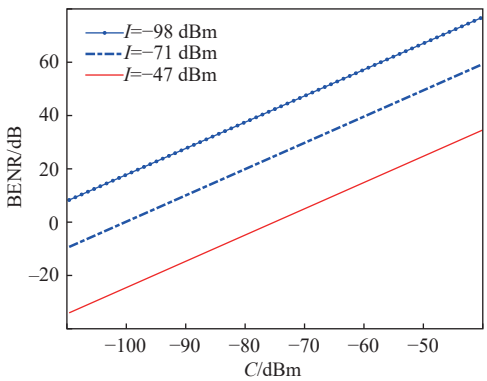
接收链路的典型参数设置如表 2 所示。一般情况下, 发射功率 P_t 为−4 dBm, 发射天线的增益 $G_t(\alpha_1,\beta_1)$ 为−9 dB, 接收端天线的增益 $G_r(\alpha_2,\beta_2)$ 为 35 dB, 天线噪声温度 T_s 为 300 K, 可求得单通道的 BENR 为 21.6 dB, 满足大于 10.8 dB 的实际应用的要求。

表 2　接收链路参数

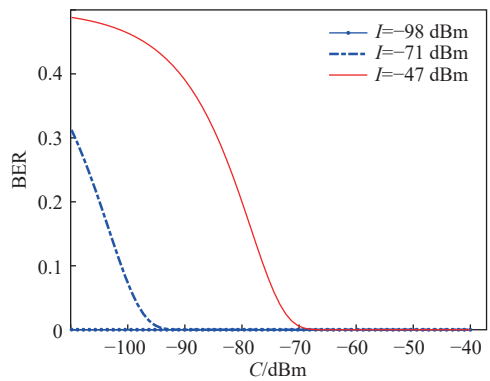
Table 2　Receiving link parameter

P_t/ dBm	L_f/ dB	L_s/ dB	L_a/ dB	$G_t(\alpha_1,\beta_1)$/ dB	$G_r(\alpha_2,\beta_2)$/ dB	T_s/ K	k/ $(\text{dBm}\cdot\text{Hz})^{-1}$
−4	−1	−164	−2	−9	35	300	−228.6

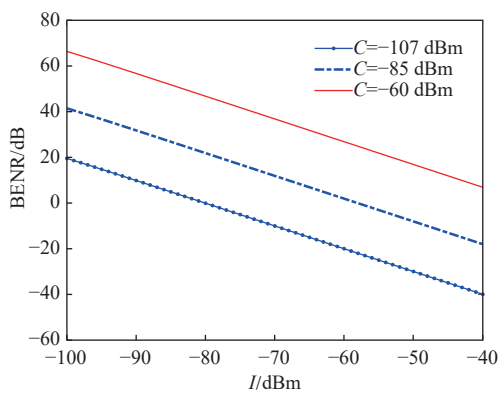
根据式(1)可由灵敏度得当前 SNR, 由式(32)得当前 BENR, 由式(33)得 BER, 仿真得 BENR、BER 与 C、I 的关系如图 5 所示。BENR 与 BER 随 C 的增大而增大, 随 I 的增大而减小; 干扰通道总功率



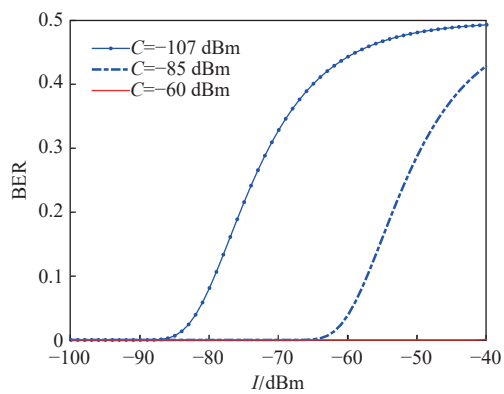
(a) 不同干扰功率时接收功率对BENR的影响



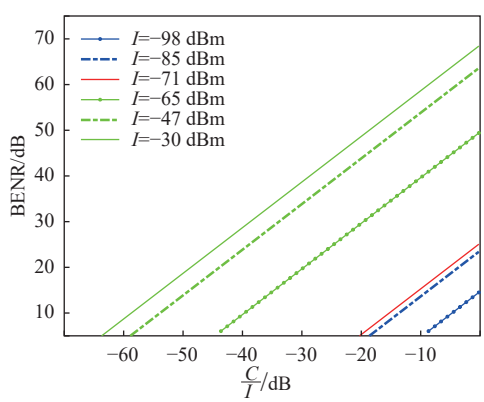
(b) 不同干扰功率时接收功率对BER的影响



(c) 不同接收功率时干扰功率对BENR的影响



(d) 不同接收功率时干扰功率对BER的影响



(e) 不同干扰功率时信干比对BER的影响

图 5　当前通道接收功率、干扰功率对接收情况的影响

Fig. 5　Influence of current channel reception power and interference power on reception

越大，BENR 损耗越大，BER 越高。经理论计算，当干扰总功率 I 为−98 dBm、−71 dBm 和−47 dBm，当前通道的接收灵敏度即满足 BER 要求的接收功率 C 可达−104 dBm、−87 dBm 和−63 dBm，此时信干比 (signal-to-interference-ratio, SIR) 分别达到−6 dB、−16 dB 和−16 dB。若期望保持接收灵敏度为−85 dBm 和−60 dBm，则 I 需低于−69 dBm 和−44 dBm，即 SIR 需高于−16 dB。

4 实验验证

测试设备的连接如图 6 所示，主要由控制模块、频谱仪、衰减器、电脑和示波器组成。电脑与 FPGA 通信，采集数字中频接收链路信号数据，控制 3 块模拟板发射 3 路不同伪码扩频生成的卫星信号，通过可调衰减器模拟太空中存在干扰时星间链路的信号强弱，并用频谱仪观测接收链路的信号频率和功率。在 Chipscope 软件中实时观察单路和多路信号的捕获灵敏度。实际应用中，常要求 BENR 大于 10.8 dB，对应的 BER 小于 10^{-6}，使用串口和 MATLAB 软件处理接收的 1.64×10^{7} 位数据，得到符合实际要求的信号接收电平，即为信号接收灵敏度。

实测单通道信号捕获灵敏度为−112 dBm，信号接收灵敏度为−107 dBm。C 弱于 I 时，多通道的捕获灵敏度和信号接收灵敏度与单通道相同。C 强于 I 时，BER 的实测结果如表 3 所示，接收灵敏度与 I 和 SIR 有关。I 为−98 dBm、SIR 为−9 dB 时，接收灵敏度仍可保持−107 dBm；I 为−72 dBm、SIR 为−13 dB 时，接收灵敏度可达−85 dBm；I 为−47 dBm、SIR 为−13 dB 时，接收灵敏度可达−60 dBm。C 弱于 I 时，接收机灵敏度受影响较小；C 强于 I、$I<-70$ dBm 时，SIR 越大，受干扰越强，对应的 EBNR 越大，接收灵敏度越高，$I>-70$ dBm 时，接收灵敏度主要与 I 的数值有关。在维持高接收灵敏度−107 dBm 的条件下，最多可抵抗−9 dB 的码间干扰，较理论分析多 3 dB，主要影响因素可能为接收机的噪声系数。实测结果符合仿真理论分析，因此可用于大规模低轨卫星编队组网的实际设计。

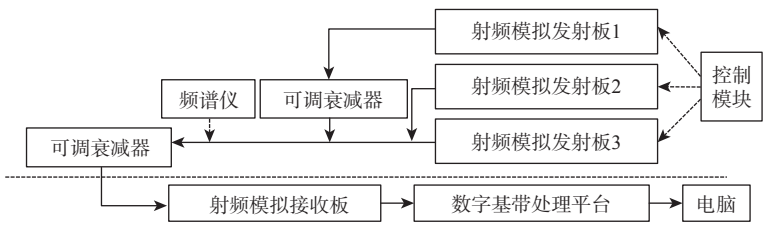


图 6 验证实验实施框图

Fig. 6 Verification test implementation block diagram

表 3 不同接收功率和干扰功率时的 BER 情况

Table 3 BER of different reception power and interference power

C/dBm	I/dBm	BER
−107	−104	0
−107	−99	3.05×10^{-7}
−107	−98	3.66×10^{-7}
−107	−97	1.65×10^{-6}
−107	−94	1.05×10^{-4}
−107	−89	3.82×10^{-2}
−85	−82	0
−85	−72	9.76×10^{-7}
−85	−71	7.80×10^{-6}
−85	−70	3.82×10^{-5}
−60	−57	0
−60	−47	3.05×10^{-7}
−60	−46	5.73×10^{-6}
−60	−45	4.01×10^{-5}

5 结 论

1）相较于传统的一对一星间通信，大规模低轨星座编队的多通道通信系统可实现一种接收机并行接收多个卫星信号的功能。

2）捕获灵敏度为−112.8 dBm，BENR 与 BER 随 C 的增大而增大，随 I 的增大而减小，干扰通道总功率越大，BENR 损耗越大，BER 越高。其他通道信号对当前通道干扰小于 6 dB 时，理论接收 BER 可满足 BENR 大于 10.8 dB 的要求。

3）实测得到多通道与单通道接收机的灵敏度相差较小，捕获灵敏度为−112 dBm，信号接收灵敏度为−107 dBm，其他通道信号对当前通道的干扰小于 9 dB 时，接收 BER 能够满足实际应用要求。

实验结果表明，多通道干扰理论分析可满足多卫星的实际应用需求，已应用于浙江大学正在研制的某型大规模卫星星座组网。

参考文献（References）

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

[2] 姜逸飞, 吴树范, 莫乾坤. 基于大规模卫星星座的 6G 组网部署及核心网节点动态规划技术[C]//推动网络演进促进应用创新—5G 网络创新研讨会. 北京: TD 产业联盟, 2021: 408-411.
JIANG Y F, WU S F, MO Q K. 6G networking deployment and core network node dynamic planning technology based on large-scale satellite constellation [C]//Proceedings of the Promoting Network Evolution and Promoting Application Innovation-Proceedings of 5G Network Innovation Seminar. Beijing: Telecommunication Development Industry Alliance, 2021: 408-411(in Chinese).

[3] 孙建亮, 崔国刚, 周英庆. 发展我国低轨卫星通信星座系统的思考[J]. 通信技术, 2020, 53(9): 2224-2227.
SUN J L, CUI G G, ZHOU Y Q. Thoughts on development of China's low-orbit satellite communication constellation system[J]. Communications Technology, 2020, 53(9): 2224-2227(in Chinese).

[4] 陈全, 杨磊, 郭剑鸣, 等. 低轨巨型星座网络: 组网技术与研究现状[J]. 通信学报, 2022, 43(5): 177-189.
CHEN Q, YANG L, GUO J M, et al. LEO mega-constellation network: Networking technologies and state of the art[J]. Journal on Communications, 2022, 43(5): 177-189(in Chinese).

[5] 方芳, 吴明阁. 全球低轨卫星星座发展研究[J]. 飞航导弹, 2020(5): 88-92.
FANG F, WU M G. Research on the development of global LEO satellite constellation[J]. Aerodynamic Missile Journal, 2020(5): 88-92(in Chinese).

[6] XU X Y, WANG C H, JIN Z H. A MAC protocol for distributed satellite network[C]//Proceedings of the IEEE 4th International Conference on Computer and Communication Engineering Technology. Piscataway: IEEE Press, 2021: 327-331.

[7] EDMONSON W, GEBREYOHANNES S, DILLION A, et al. Systems engineering of inter-satellite communications for distributed systems of small satellites[C]//Proceedings of the Annual IEEE Systems Conference Proceedings. Piscataway: IEEE Press, 2015: 705-710.

[8] TANEN BAUM A S. 计算机网络[M]. 4 版. 潘爱民, 译. 北京: 清华大学出版社, 2004.
TANENBAUM A S. Computer networks[M]. 4th ed. PAN A M, translated. Beijing: Tsinghua University Press, 2004(in Chinese).

[9] 李倩. 低轨卫星通信系统随机接入技术研究[D]. 北京: 北京邮电大学, 2019.
LI Q. Research on low-orbit satellite random access technology[D]. Beijing: Beijing University of Posts and Telecommunications, 2019 (in Chinese).

[10] 韦芬芬, 刘晓旭, 谢继东, 等. 低轨卫星物联网多址接入方式研究[J]. 计算机技术与发展, 2019, 29(5): 116-120.
WEI F F, LIU X X, XIE J D, et al. Research on multiple access methods for LEO satellite IoT[J]. Computer Technology and Development, 2019, 29(5): 116-120(in Chinese).

[11] 黄长文, 朱立东. 低轨卫星系统扩频 ALOHA 多址接入性能分析[C]//第十三届卫星通信学术年会. 北京: 中国通信学会, 2017: 86-93.
HUANG C W, ZHU L D. Performance analysis of spread spectrum ALOHA multiple access for low orbit satellite systems [C]//Proceedings of the 13th Annual Conference on Satellite Communications. Beijing: China Institute of Communications, 2017: 86-93(in Chinese).

[12] NGUYEN D T, PARK Y. Performance improvement of optical satellite communications by interleaved IEEE 802.11 LDPC[C]//Proceedings of the Tenth International Conference on Ubiquitous and Future Networks. Piscataway: IEEE Press, 2018: 575-579.

[13] CHEN B C, YU L. Design and implementation of LDMA for low earth orbit satellite formation network[C]//Proceedings of the IFIP 9th International Conference on Embedded and Ubiquitous Computing. Piscataway: IEEE Press, 2011: 409-413.

[14] 窦中兆, 雷湘. WCDMA 系统原理与无线网络优化[M]. 北京: 清华大学出版社, 2009.
DOU Z Z, LEI X. WCDMA system principle and wireless network optimization[M]. Beijing: Tsinghua University Press, 2009(in Chinese).

[15] MENG E T, BU X Y. Two-dimensional joint acquisition of Doppler factor and delay for MC-DS-CDMA in LEO satellite system[J]. IEEE Access, 2020, 8: 148203-148213.

[16] MENG E T, LI R D, YU J H, et al. Joint CCI mitigation and power control for MC-DS-CDMA in LEO satellite networks[J]. IEEE Internet of Things Journal, 2022, 9(18): 17627-17639.

[17] 程鹏仇, 佩亮, 周玥, 等. 802.20 OFDMA 系统的上行接入同步算法[J]. 浙江大学学报 (工学版), 2008, 42(11): 1894-1899.
CHENG P Q, PEI L, ZHOU Y, et al. Uplink access synchronization algorithm for 802.20 OFDMA system[J]. Journal of Zhejiang University (Engineering Edition), 2008, 42(11) : 1894-1899(in Chinese).

[18] 谢展腾. LEO 卫星 CDMA 上行链路功率控制算法研究[D]. 哈尔滨: 哈尔滨工业大学, 2017.
XIE Z T. The research of CDMA uplink power control algorithms for LEO satellites[D]. Harbin: Harbin Institute of Technology, 2017 (in Chinese).

[19] ILCEV S D. Analyses of code division multiple access (CDMA) schemes for global mobile satellite communications (GMSC)[J]. TransNav: The International Journal on Marine Navigation and Safety of Sea Transportation, 2020, 14(4): 805-810.

[20] NEINAVAIE M, KHALIFE J, KASSAS Z M. Doppler stretch estimation with application to tracking globalstar satellite signals [C]//Proceedings of the IEEE Military Communications Conference. Piscataway: IEEE Press, 2021: 647-651.

[21] 朱国富. 扩频体制低轨卫星通信信号捕获与跟踪系统设计[J]. 电讯技术, 2022, 62(5): 576-584.
ZHU G F. Design of a spread spectrum LEO satellite signal acquisition and tracking system[J]. Telecommunication Engineering, 2022, 62(5): 576-584(in Chinese).

[22] 牛翼. 基于 5G 的低轨卫星系统下行频偏估计算法研究[D]. 西安: 西安电子科技大学, 2020.
NIU Y. Research on frequency offset estimation algorithm for downlink of LEO satellite system based on 5G[D]. Xi'an: Xidian University, 2020(in Chinese).

[23] KAPLAN E D. GPS 原理与应用[M]. 2 版. 寇艳红, 译. 北京: 电子工业出版社, 2012.
KAPLAN E D. Principles and applications of GPS[M]. 2nd ed. KOU Y H, translated. Beijing: Electronic Industry Press, 2012(in Chinese).

Analysis of CDMA interference in large-scale satellite constellation

XU Kedi[1, 2], XU Zhaobin[1, 2, *], GUO Xiaoxu[1], YANG Jia[1], CHENG Hengfei[1], JIN Zhonghe[1, 2]

(1. Micro-Satellite Research Center, Zhejiang University, Hangzhou 310027, China;
2. Key Laboratory of Micro-Nano Satellite Research, Zhejiang Province, Zhejiang University, Hangzhou 310027, China)

Abstract: Code division multiple access (CDMA) communication system is the main part of large-scale satellite constellation network. In the case of multiple access interference in multi-channel communication between satellites, the receiving link is modeled. This study elucidates the correlation between bit energy to noise density ratio (BENR) and device received signal-to-noise ratio (SNR), while also investigating the impacts of additive white noise and multiple access interference on SNR. The BENR and sensitivity are simulated in MATLAB with different receiving and interference power. The simulation demonstrates that the interference signal exerts a diminished impact on the sensitivity when the signal-to-interference ratio (SIR) exceeds 0 dB. The receiver can withstand −6 dB interference while maintaining high sensitivity. An experimental platform for three transmitters and one receiver is established. Data is processed by Chipscope and MATLAB software to obtain the received sensitivity that meets the actual application requirements. Experiments show that when SIR is greater than 0 dB, the received sensitivity can reach −107 dBm. The receiver can resist −9 dB interference at high sensitivity. Error is affected by circuit noise. The congruence between the measured data and the theoretical analysis provides a fundamental basis for the development of the real low-orbit satellite constellation network.

Keywords: constellation; satellite communication; code division multiple access; multiple access interference; sensitivity

Received: 2022-08-29; **Accepted**: 2022-11-19; **Published Online**: 2022-12-15 17:22

URL: link.cnki.net/urlid/11.2625.V.20221214.1447.006

Foundation items: National Natural Science Foundation of China General Program (62073289); National Natural Science Foundation of China(U21A20443)

* **Corresponding author**. E-mail: zjuxzb@zju.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0742

基于近似计算技术的多级逻辑电路面积优化

赵维凯[1]，王伦耀[1,*]，岑旭梦[2]，夏银水[1]，储著飞[1]

(1. 宁波大学 信息科学与工程学院，宁波 315211；2. 宁波大学科学技术学院 信息工程学院，宁波 315300)

摘 要：针对现有多级逻辑近似优化算法在大型电路优化时无法较好兼顾优化效果与算法速度的问题，提出一种基于常量替换的多级逻辑电路面积优化算法。该算法通过在与非图(AIG)中引入节点输出距离、常量传输距离等参数，并结合提出的多扇出节点常量替换过程中错误率控制方法，在全局范围内提取合适替换的候选节点集和每个候选节点输出的常量替换值；同时，通过根据电路规模选取不同的错误率计算方法等策略，实现了错误率约束下的多级逻辑电路面积近似优化并使算法速度得到提高。所提算法用C语言和ABC内置命令编程实现，使用EPFL及MCNC电路进行测试。实验结果表明：所提算法与已提出的常量替换方法相比，面积优化效果提升48.77%；相较于近似关心集重代换的近似逻辑综合(ALSRAC)优化算法，所提算法在面积优化和运算时间上分别有1.28%和60.91%的提升。

关 键 词：多级电路；近似计算；误差约束；常量替换；逻辑优化

中图分类号：TN431.2

文献标志码：A **文章编号**：1001-5965(2024)09-2893-09

随着晶体管尺寸的越来越小，集成电路设计变得愈加复杂，通过传统的设计方法来提升电路的性能和降低能量消耗也越发艰难[1]；同时，近年来出现的许多诸如人工智能、音频图像处理、大数据推送等应用具有的容错特性，使得其在一定范围内的误差不影响实际应用功能。近似计算技术就是利用这些容错特性，通过降低计算精度，进而获得电路功耗、延迟和面积的优化的。目前基于近似计算技术的电路设计已经成为一种新的电路设计范式[2]。利用近似计算技术，在给定的近似度约束条件下，如错误率、错误距离、平均误差[3]等，可设计出更高性能的电路。

近似计算技术在电路优化上主要集中在电路设计和逻辑综合两方面。在电路设计上，目前主要集中在加法器、乘法器等算术电路的近似实现中[4]。逻辑综合则侧重于近似计算技术在电路优化方面的应用，如二级逻辑和多级逻辑的综合与优化。本文主要研究多级电路的近似逻辑综合。现有研究工作表明，利用近似计算技术，在引入较小计算误差的情况下，能显著提升电路的性能[5-13]。

在通过近似计算技术实现多级电路逻辑综合研究方面，文献[7]提出了替换并简化(substitute-and-simplify, SASIMI)的优化方法，该方法通过寻找2个功能近似的信号，用一个替换另一个，以实现电路性能的优化。文献[8]提出了批量误差估计法，并对文献[7]的误差估算部分进行了优化，获得了更好的电路优化效果。文献[9]通过用小面积或低功耗门电路替换大面积或高功耗门电路的方式来实现面积和功耗的优化。文献[10]提出基于布尔重写，针对切割集用常量节点替换的方式来实现电路性能的优化。从实验结果来看，该方法在用常量替换处理切割集时，容易导致电路中过量节点被

收稿日期：2022-08-29；录用日期：2022-09-16；网络出版时间：2022-11-01 09:52
网络出版地址：link.cnki.net/urlid/11.2625.V.20221031.1449.003
基金项目：国家自然科学基金(U1709218,61471211,61871242)
*通信作者. E-mail：wanglunyao@nbu.edu.cn

引用格式：赵维凯，王伦耀，岑旭梦，等. 基于近似计算技术的多级逻辑电路面积优化[J]. 北京航空航天大学学报，2024，50(9)：2893-2901. ZHAO W K，WANG L Y，CEN X M，et al. Area optimization of multilevel logic circuits using approximate computing[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2893-2901 (in Chinese).

删除，造成错误率激增而突破错误率约束，进而影响优化结果。文献 [11] 提出近似逻辑综合 (approximate logic synthesis by resubstitution with approximate care, ALSRAC) 的优化方法，通过搜索、获取近似候选集，并对近似候选集中的节点进行近似变换，从而实现面积优化。但该方法存在大电路优化时间过长的问题。

为解决文献 [10-11] 存在的问题，本文提出如下改进方法：

1）不同于文献 [10] 仅局限于关键路径上的切割集搜索，本文提出在全电路搜索适合常量替换节点的方法，以及对候选节点采用何种常量替换方式的评估方法。

2）针对多扇出节点常量替换过程容易发生突破错误率约束的问题，提出对应的错误率控制方法，实现错误约束下电路结构的尽可能简化。

3）提出一种结合 Miter 电路[14] 和二元决策图 (binary decision diagram, BDD)[15] 的多级电路错误率计算方法。

本文多级电路近似优化方法是结合 ABC 开源工具实现的。实验结果表明：相较于文献 [10-11]，本文方法在相同错误率约束下，可以获得更好的面积优化效果。

1 与非图

由于“与”“非”运算可以构成逻辑运算完备集，因此，任何一个逻辑表达式都可以表示为“与”“非”运算构成的表达式。与非图 (and-inverter-graph, AIG) 是一种仅包含“与”“非”运算，用来表示逻辑功能的有向无环图[16]。AIG 由二输入与门、反相器、有向边和输入输出组成。其是多级逻辑电路的一种通用表示形式，也是多级逻辑优化中重要的研究对象。以图 1 所示的 C17 电路对应的 AIG 为例，图中正三角表示原始输入 (primary inputs, PIs)，倒三角表示原始输出 (primary outputs, POs)。圆圈表示与门。AIG 中的有向连线代表各个节点之间存在信号输出输入关系，其中，虚线表示对该节点的输出信号取反，相当于接入反相器。为方便描述，通常给 AIG 中的节点、PIs 和 POs 进行编号。其中，编号 1～5 为 PIs，编号 8～13 为节点，编号 22 和 23 为 POs。

一般情况下，AIG 中 PI 到 PO 之间存在多条路径，其中，包含节点最多的路径称为关键路径，关键路径中的信号延迟也最大。由于 AIG 中的每个节点对应一个“与”运算，因此，AIG 中包含的节点数越多，对应的电路面积也越大。这也是基于 AIG 的逻辑优化中常利用 AIG 节点数和关键路径长度作为电路面积和延迟评估标准[17] 的原因。本文也利用 AIG 节点数量作为面积优化的衡量标准。

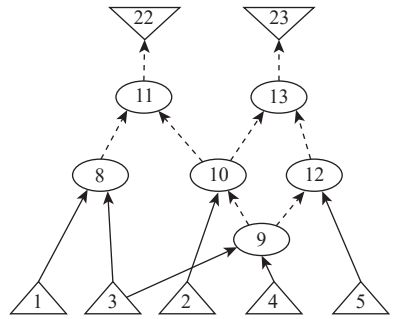


图 1 C17 电路 AIG

Fig. 1 AIG of C17 circuit

2 错误率计算

采用近似计算技术的逻辑优化都必须满足给定的近似度约束。近似度有多种衡量方式，错误率便是比较常用的一种方式[3]。近似逻辑函数的错误率的定义为导致近似函数与原函数输出不同的输入组合占所有输入组合的比例。

对于一个 n 输入，m 输出的逻辑函数 $f\left(\vec{X}\right)$，令 $\vec{X}=\{x_0,x_1,\cdots,x_{n-1}\}$ 为输入变量集合，$f_{\mathrm{ap}}\left(\vec{X}\right)$ 为 $f\left(\vec{X}\right)$ 的近似函数，E 为错误率，可表示为

$$E=\frac{M\left[f\left(\vec{X}\right)\neq f_{\mathrm{ap}}\left(\vec{X}\right)\right]}{2^n}\times 100\% \tag{1}$$

式中：$M\left[f\left(\vec{X}\right)\neq f_{\mathrm{ap}}\left(\vec{X}\right)\right]$ 表示在所有输入组合中，使得 $f\left(\vec{X}\right)\neq f_{\mathrm{ap}}\left(\vec{X}\right)$ 的输入组合数。

由于当 $f\left(\vec{X}\right)\neq f_{\mathrm{ap}}\left(\vec{X}\right)$，有 $f\left(\vec{X}\right)\oplus f_{\mathrm{ap}}\left(\vec{X}\right)=1$，因此，满足 $f\left(\vec{X}\right)\oplus f_{\mathrm{ap}}\left(\vec{X}\right)=1$ 输入的组合数与 $M\left[f\left(\vec{X}\right)\neq f_{\mathrm{ap}}\left(\vec{X}\right)\right]$ 相等，可将式 (1) 表示为

$$E=\frac{M\left[f\left(\vec{X}\right)\oplus f_{\mathrm{ap}}\left(\vec{X}\right)=1\right]}{2^n}\times 100\% \tag{2}$$

式中：$M[f\left(\vec{X}\right)\oplus f_{\mathrm{ap}}\left(\vec{X}\right)=1]$ 的值可以利用如图 2 所示的 Miter 电路求得。图 2 中，近似电路 C_{app} 对应 $f_{\mathrm{ap}}\left(\vec{X}\right)$，电路 C_{org} 对应 $f\left(\vec{X}\right)$，则所有使 Miter 电路输出为 1 的输入组合数量等于 $M\left[f\left(\vec{X}\right)\oplus f_{\mathrm{ap}}\left(\vec{X}\right)=1\right]$。

满足图 2 中 Miter 电路输出为 1 的输入组合，即 $M\left[f\left(\vec{X}\right)\oplus f_{\mathrm{ap}}\left(\vec{X}\right)=1\right]$，可以借助 BDD 来求解。具体方法包含如下 3 个基本步骤：

步骤 1 通过开源工具 ABC 的内置函数，构建图 2 中的 Miter 电路，并通过 ABC 将 Miter 电路对应的逻辑功能用 BDD 进行描述。

步骤 2 统计 BDD 中所有从根节点到取值为

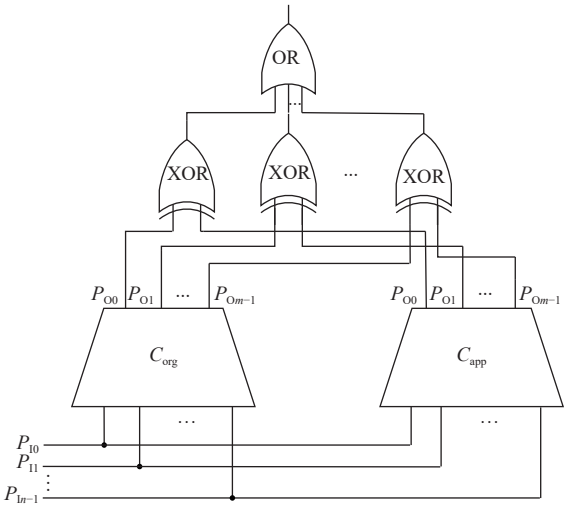


图 2　Miter 电路

Fig. 2　Miter circuit

1 的叶子节点路径数量。

步骤 3　所有取值为 1 的路径对应的乘积项所包含最小项的数量既是 $M\left[f\left(\vec{X}\right)\oplus f_{ap}\left(\vec{X}\right)=1\right]$ 的值, 最终利用式 (2) 得到近似电路 C_{app} 的错误率。

当 C_{app} 和 C_{org} 规模较大时, Miter 电路对应的 BDD 规模也将变得很大, 进而影响 BDD 中取值为 1 的路径的统计效率。因此, 在大电路近似优化时, 本文使用蒙特卡罗模拟法来估算错误率[8]。该方法不是对所有输入情况进行模拟得到精确的错误率, 而是随机产生一定数量的测试输入, 再比较输出的结果, 最后通过式 (2) 获得对应的错误率。

虽然该方法得到的结果有误差, 但能有效降低错误率计算的复杂度, 从而减少了大电路近似优化时间。因此, 蒙特卡罗模拟法也是一种通用的错误率估算方法。由式 (2) 可得, 当输入变量数 n 比较大时, 如 $n>16$, 式 (2) 中分母大小是十万级, 则计算 $M\left[f\left(\vec{X}\right)\oplus f_{ap}\left(\vec{X}\right)=1\right]$ 过程中引入的误差对式 (2) 的计算结果的影响有限, 即大电路对错误率敏感程度较低, 这也是大电路在错误率计算时可以采用估算方法的一个重要原因。

3　多级逻辑电路面积优化算法

本文的多级逻辑电路面积优化利用 AIG 来实现, 并用 AIG 的节点数量来衡量逻辑电路的面积, 通过减少 AIG 节点数量以实现电路面积优化。

3.1　基于常量替换的 AIG 简化方法

假设 $N_i(a_i,b_i)$ 为 AIG 中的 1 个节点, a_i 和 b_i 为该节点的 2 个输入, 为节点输出。则当 $a_i=1$ 时, 有 $y=b_i$, 即可以通过将节点 $N_i(a_i,b_i)$ 中的 1 个输入用常数 1 替代, 来实现对节点 $N_i(a_i,b_i)$ 的删除。当 AIG 节点 $N_i(a_i,b_i)$ 中某个输入为 0 时, 如 $a_i=0$, 则 $y=N_i(a_i,b_i)=0$; 若 y_i 不是 AIG 的 P_O, $y_i=0$ 会重复 $a_i=0$ 过程继续向输出端传输。只有当扇出信号接入反相器时, “0” 信号的传递才会结束。反之, 当替换节点的扇出存在反相器时, $a_i=1$ 会不断向输出端传输, 直至扇出信号不存在反相器, “1” 信号的传递才结束。从本节分析可知, 通过对 AIG 节点输出用常量替换可以实现 AIG 节点的删除, 进而实现电路面积优化。

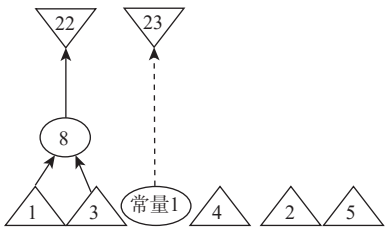
多数情况下, 节点 $N_i(a_i,b_i)$ 的某一个输入来自另一个节点的输出, 因此, 实现 AIG 简化的另一个途径是直接将某一个节点 j 的输出设置为常量, 这样以节点 j 为输出节点的子 AIG 就可以用一个常量来代替, 从而让整个 AIG 将得到简化。以图 1 中节点 9 为例, 假设节点 9 的输出用常量 1 代替, 则不但以节点 9 为输出节点的子 AIG 将用常量 1 代替, 且又因节点 9 的输出为 1, 同时与节点 9 相连的节点 10、12 均存在反相器, 使得输出由 1 翻转为 0, 根据与门的性质, 节点 10 和节点 12 可以用常数 0 代替, 并进一步引起节点 13 的 2 个输入均为 1, 节点 11 的 1 个输入为 1, 这使得编号为 23 的输入为常量 1, 节点 11 可以删除。最终图 1 中的 AIG 可简化为如图 3(a) 所示的结构。

在 AIG 中都会存在一定数量的多扇出节点, 由于一个多扇出节点与多个节点相连, 因此, 当多扇出节点的输出被设置为常量时会导致较多数量的节点被删除, 使得错误率超过约束值。因此, 本文在处理多扇出节点时先进行逻辑复制, 后对复制得到的单输出节点再逐个用常量替换。以图 1 所示的 AIG 为例, 在对多扇出节点 9 的 2 个扇出进行常量替换时, 可以先将以节点 9 为扇出的子 AIG 进行逻辑复制, 得到 2 个子 AIG, 其结构一致, 扇出节点分别为节点 9a 和节点 9b; 其中, 节点 9a 的扇出为节点 10 的一个扇入, 而节点 9b 的扇出为节点 12 的 1 个扇入; 将节点 9b 扇出用常量 1 替换, 得到如图 3(b) 所示的结果。与图 3(a) 相比, 对多扇出节点先进行逻辑复制, 再进行常量替换可以有效控制节点删除规模, 进而实现 AIG 优化过程中对错误率的控制。

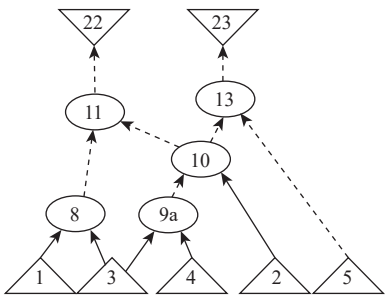
3.2　输出用常量替换的节点搜索策略

由于 AIG 中的每个节点都表示一个二输入的逻辑 “与” 运算, 而当逻辑 “与” 运算的一个输入为逻辑 0 时, 无论其他输入发生什么变化, 其输出都恒为 0。换言之, 即使存在一些逻辑的错误输出, 当这些错误的输出汇聚到一个输入为 0 节点时, 这些错误也会因无法传递到电路输出端而对整个电路的错误率无法产生影响。因此, 在 AIG 中, 离 PO

越远的节点产生的错误被传输到输出端的概率就越低，即删除那些离POs越远的节点，对整个电路错误率的影响越小。另外，在AIG中，层级越高的节点，与其有关联的变量就越多，导致此类节点常量替换后，引起错误输出的输入组合越多，即引起的错误率也会越高。为控制错误率增加，本文将那些远离POs的节点作为常量替换的优选节点，并且用输出距离(output distance, OD)来表示节点与POs的距离，如算法1所示，该算法1用于介绍如何计算每一个节点到输出端的距离。



(a) 节点9常量1替换结果



(b) 节点9逻辑复制后处常量1替换结果

图3 节点9扇出的不同常量替换结果

Fig. 3 Results of the different constant replacement for node 9 fanout

算法1 OD_Comp。

输入: 电路$A(m,n,k)$。

输出: 每个节点的OD。

1. FOR each Node DO

2. $N_{\mathrm{PO}}(O_{\mathrm{D}})=0$, $N(O_{\mathrm{D}})=K$; //初始化，N_{PO}为与输出相连节点，N为内部节点

3. FOR each node in AIG DO //从PO端向PI端搜索

4. $N_b=\mathrm{fanin}(N_a)$;

5. $O_{\mathrm{D}}(N_b)=\min(O_{\mathrm{D}a}+1,O_{\mathrm{D}b})$

6. IF$(N_b\in S(N,O_{\mathrm{D}}))$

7. $O_{\mathrm{D}}(N_b)=\min(O_{\mathrm{Dnew}},O_{\mathrm{Dold}})$

8. $S(N,O_{\mathrm{D}})\leftarrow\{N_b,O_{\mathrm{D}b}\}$;//将所获得的节点和OD值保存

9. END FOR

10. return $S(N,O_{\mathrm{D}})$

在AIG中，若某个节点N_i与AIG第j个输出PO_j存在信号通路L_{ji}，则N_i与PO_j之间的$O_{\mathrm{D}}(N_i)$定义为L_{ji}中除N_i外经过的节点数。以图1为例，节点11和节点13的OD值为0，节点8、节点10和节点12的OD值为1，节点9的OD值为2。一般而言，由于节点N_i与输出之间存在多条信号通路，因此，N_i存在多个OD值，本文以N_i所有OD值中的最小值作为节点N_i的OD值。对于具有m个输出、n个输入、K个节点的AIG，其每个节点的OD值可以用算法OD_Comp来确定：将与PO端相连的节点OD值设置为0，将内部节点OD值设置为K；以每一个与PO相连的节点为根节点通过深度优先的方式开始向下探索；通过获得当前节点对应的扇入节点，对2个节点的OD值进行对比，取其较小的OD值作为该节点的OD值并保存，若该节点已经存在于节点集中，则比较两者的OD值，取其较小的OD值进行保存，最终基于此遍历所有节点并记录其对应的OD值。

图4为M_C17电路的AIG，该AIG有2个输出、4个输入、7个节点。初始化后，节点10和节点11的OD值为0，其他节点的OD值为7。

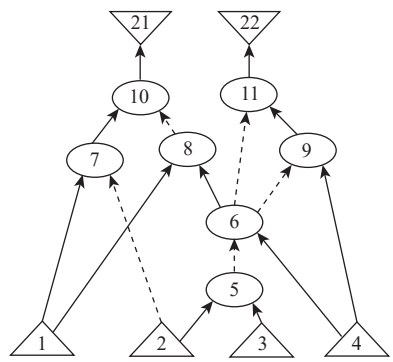


图4 M_C17电路的AIG

Fig. 4 AIG of benchmark M_C17

以编号为21的PO为出发点，逆AIG连线方向找到一条与PI相连的通路，并假设为编号10、8、6、5节点组成的通路。由于节点10的OD值为0，因此，节点8的OD值为$0+1=1$，考虑到节点8原来的OD值为7(初始为7)，因此，$\min(1,7)=1$，即节点8的OD值为1。同理可得节点6和节点5的OD值分别为2和3。用同样的方法可得节点7的OD值为1。

在编号为22的PO与PI之间，存在如下节点构成的通路(11,9,6,5)、(11,6,5)、(11,9,6)、(11,6)和(11,9)。其中，利用通路(11,6,5)计算节点6的OD值为1，而节点6原来的OD值为2，因此，更新节点6的OD值为1。同理得到节点5的OD值为2，节点9的OD值为1。最后，经过算法OD_Comp后，图4中节点5、6、7、8、9、10、11的OD值分别为：2、1、1、1、1、0、0。

3.3　节点输出替换的常量确定

候选节点确定后, 其输出的常量取值 0 或 1, 这对 AIG 的简化影响也很大。如图 4 所示, 当节点 5 的输出设为 1 时, 编号为 22 的 PO 直接与编号为 4 的 PI 相连, 编号为 21 的 PO 直接与节点 7 相连, 此时图 4 中只剩下节点 7。反之, 当节点 5 的输出设为 0 时, $\mathrm{PO22}=(\mathrm{PI4})\cdot\overline{(\mathrm{PI4})}=0$, 相当于编号直接与常量 0 相连, 而节点 7、8、10 将保留。从上述分析可知, 节点输出的常量设定将直接影响 AIG 中的节点删除效果, 进而影响近似简化后得到的 AIG 的错误率。因此, 在节点输出常量选择上引入常量传输距离 TD(transmission distance) 值来判断选用常量 0 还是 1 进行近似替换。

常量传输距离 TD 值定义为: 假设节点 i 的 OD 值为δ, 节点 i 的输出被替换为常量 (0 或 1) 时, 将导致α个节点被删除, 设C_{OD}表示被删除的节点的 OD 值的集合, $O_{\mathrm{D}j}$表示第 j 个被删除节点的 OD 值, 则 TD 值表达式为

$$T_D|_p=\alpha\delta-\sum_{j=1}^{\alpha}O_{\mathrm{D}j}\qquad p\in\{0,1\};\quad O_{\mathrm{D}j}\in C_{\mathrm{OD}}\qquad(3)$$

式中: p 的取值为 0 或 1, 表示节点 i 的输出设为 0 或 1。

以图 4 所示 AIG 为例, 当节点 5 的输出设为 1 时, 节点 6、8、9、10、11 被删除, 根据 3.2 节的例子得到每个节点对应的 OD 值, 即$T_{\mathrm{D}}|_{p=1}=5\times2-1-1-1-0-0=7$, 当节点 5 的输出设为 0 时, $T_{\mathrm{D}}|_{p=0}=4$, 一般而言, TD 值越大, 被删除更多节点数量的可能性更高。本文利用 TD 值来确定节点输出常量用 1 或 0 来替换。

3.4　基于 AIG 的逻辑电路面积近似优化算法

由于基于 AIG 的电路面积近似优化的核心是在错误率的约束下尽可能多地删除 AIG 节点, 而能否充分利用错误率约束提供的优化空间是提高电路面积优化效果的关键。同时考虑到近似优化后的电路面积由 AIG 的节点数来表示, 引入的错误率小的节点应该是首选的待删除对象。因此, 本文提出的面积近似优化算法包括 3 个部分: 1) 通过引入节点距离, 确定常量替换候选节点集合, 即待删节点集合; 2) 通过引入常量传输距离, 确定某一节点选取何种常量进行替换, 以实现节点删除; 3) 根据各个待删节点的优化效果, 引入优化权重, 最终确定本轮迭代中哪些节点的输出用哪个常量替换, 从而实现节点删除, 如算法 2 所示, 该算法主要用于介绍本文近似面积优化的思路, Area_opt 表示面积优化。

算法 2　Area_opt。

输入: 电路$A_{\mathrm{org}}(m,n,K)$, 错误率约束E_{th}。

输出: 近似优化后电路$A_{\mathrm{app}}(m,n,K_{\mathrm{app}})$, 优化后的错误率$E_{\mathrm{R}_0}$。

初始化: 待优化节点集合$C_N=\varnothing$, 错误率$E_{\mathrm{app}}=0$, $C_W=\varnothing$。

1. $A_{\mathrm{app}}=\mathrm{AIGCopy}(A_{\mathrm{org}})$, $\lambda=K$
2. WHILE ($E_{\mathrm{app}}<E_{\mathrm{th}}$) DO
3. $A_{\mathrm{c}}=\mathrm{AIGCopy}(A_{\mathrm{app}})$
4. $C_{\mathrm{OD}}=\mathrm{OD_Comp}(A_{\mathrm{c}}(m,n,\lambda))$;
5. $C_N=\mathrm{Sort}(C_{\mathrm{OD}})$
6. $C'_N=\mathrm{Mo_So}(C_N)$
7. FOR each node in C'_N,$N_i\in C'_N$, DO
8. $(N_i,p_i)=\mathrm{GetMaxTD}(N_i)$
9. $A_i(m,n,\lambda_i)=\mathrm{LAC}(A_{\mathrm{c}},N_i,p_i)$
10. $E_i=\mathrm{GetER}(A_{\mathrm{org}},A_i)$
11. $W_i=\dfrac{(\lambda-\lambda_i)}{(E_i-E_{\mathrm{app}})}$
12. $C_W=C_W\cup\{W_i\}$
13. END FOR
14. $(N_{\mathrm{bst}},p_{\mathrm{bst}})=\mathrm{BestSel}(C_W)$, $C_W=\varnothing$
15. $A_{\mathrm{app}}(m,n,\lambda_1)=\mathrm{LAC}(A_{\mathrm{c}},N_{\mathrm{bst}},p_{\mathrm{bst}})$
16. $A_{\mathrm{app}}(m,n,\lambda)=\mathrm{Merg}(A_{\mathrm{app}})$
17. $E_{\mathrm{app}}=\mathrm{GetER}(A_{\mathrm{org}},A_{\mathrm{app}})$
18. END WHILE
19. return $A_{\mathrm{app}}(m,n,K_{\mathrm{app}})$, E_0

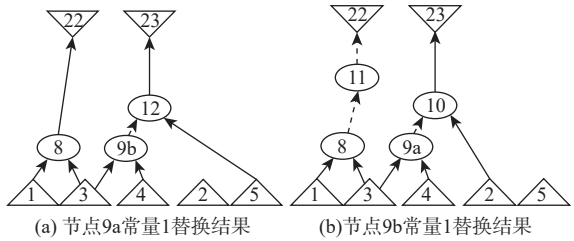
算法 2 的基本思想值: AIG 中远离输出端的节点(即 OD 值大)的删除, 会导致其引入的错误变小; 每个节点输出的常数值“0”或“1”的选择由 TD 值决定, TD 值大的对应的输出常数替代删除的节点数越多, 因此, 可以根据各节点 OD 值的大小, 优选部分 OD 值大的节点作为候选节点, 然后结合各候选节点的输出设置为常数“0”或“1”后, 引起错误率增加和 AIG 节点删除情况, 决定优先删除哪个 AIG 节点。第 1 步通过 AIGCopy 函数, 将待简化的 AIG 进行备份, 用于少节点删除后的错误率。第 2～17 步为在候选节点中选取最佳节点及其输出常数的确定过程, 其中, 第 4～6 步确定本轮候选节点。第 4 步中用算法 1 计算当前 AIG 各节点对应的 OD 值, C_{OD}为所有节点 OD 值的集合; 第 5 步对C_{OD}中 OD 值进行排序, 并取较大的前 N 个节点得到集合C_N; 第 6 步用逻辑复制技术, 将C_N中多扇出节点转化为多个单输出节点, 得到节点集合为C'_N。第 7～13 步用来确定C'_N哪个节点的输出设置为常数“0”或“1”可使得在引入错误率和 AIG 节点减少方面获得最好的结果。其中, 第 7 步确定C'_N中的某个节点N_i的最大 TD 值, 以及对应的输出为常数的取值p_i; 第 9 步将节点N_i的输出设置为p_i,

得到简化的A_i，该AIG包含的节点数为λ_i；第10步计算AIGi的错误率为ERi；第11步中引入错误率和节点减少的评价指标W_i，W_i值越大代表对应节点越适合进行常数替代。C'_N中每个节点的W_i存储在C_W。

第14～17步完成C'_N中最佳节点选择，选中的节点的输出被设置为常数p_{bst}，根据常数将对应的节点用0或者1替换，电路会由于常量值得到优化，生成新的优化后的AIG电路，算法会同时计算对应的错误率。其中，第14步完成在C_W中选取具有最大W_i值的节点，该节点为N_{bst}，其输出被预置的常数为p_{bst}；第15步将节点N_{bst}的输出置为p_{bst}，得到简化后的AIG为AIGapp，其中，A_{app}中包含的节点数为λ_1；第16步将A_{app}中等效节点合并为多扇出节点，实现节点数的减少，最后，A_{app}的节点数为λ。第17步计算AIGapp的错误率。

根据电路输入变量数的大小，本文采用2种处理模式。对于PIs < 17的电路，因其输入组合数量至多只有十万级别，所以本文通过构建Miter电路和BDD来获得精确的错误率。当PIs ⩾ 17时，则采用文献[8]中提出的蒙特卡罗方式估算电路的错误率。从而可在保持一定估算精度的情况下提高算法运行速度。

以图1为例来说明算法2实现AIG的简化过程。由于该电路节点较少，所以假设错误率约束不作限制。先分别计算出节点8、9、10、11、12和13对应的OD值为1、2、1、1、0和0。图5以OD值最大的节点9作为候补集来进行说明。由于节点9有2个扇出，则将节点9拆分为节点9a和节点9b。对于节点9a，$\text{TD}|_{p=1}=3$，$\text{TD}|_{p=0}=1$，由此选择将节点9a输出设置1接入节点10中，得到的优化电路如图5(a)所示，计算得到错误率为$\frac{11}{32}$，节点优化数量为3，即$W_1=8.73$，将W_1存储到优化权重



(a) 节点9a常量1替换结果　(b)节点9b常量1替换结果

图5　对C17中节点9不同扇出优化结果

Fig. 5　Optimization results for different fanouts of node 9 in C17

合集C_W中；对于节点9b，$\text{TD}|_{p=1}=3$，$\text{TD}|_{p=0}=1$，由此选择将节点9b输出设置1接入节点12中，得到的优化电路如图5(b)所示，计算得到错误率为$\frac{7}{32}$，节点优化数量为2，即$W_2=9.14$。据此比较2次替换尝试W值的大小，可以得到对节点9b进行常量1的替换在本次面积优化中结果更优，由此将其电路结构及错误率作为最终输出结果。

4　实验结果分析

本文算法使用C语言编程实现并结合逻辑综合工具ABC完成电路面积优化。运行环境为：处理器频率为2.3 GHz，内存为8 GB，操作系统为Ubuntu-20.04。由于初始输入电路存在冗余结构，因此，对于每个输入电路都先用ABC中的“resyn2”命令进行预处理，对待近似优化电路则进行精确简化。表1为本次实验用到的北卡罗莱纳州微电子中心(MCNC)和洛桑联邦理工学院(EPFL)测试电路，映射的工艺库取自MCNC，其中，初始面积为通过ABC精确优化后再进行工艺库映射所得到的面积，在计算大电路的错误率时，由于大电路对电路错误率敏感程度较低，因此本文采用与文献[11]相同的蒙特卡罗模拟方法，通过随机生成测试输入进行仿真获得当前电路的错误率，以减少程序计算错误率的运行时间。

表1　本文使用的MCNC和EPFL测试电路参数

Table 1　Parameters of MCNC and EPFL benchmarks used in this paper

MCNC	电路输入端口数量	电路输出端口数量	初始面积	EPFL	电路输入端口数量	电路输出端口数量	初始面积
cm163	16	5	61	arbiter	25	129	23 366
z4ml	7	4	61	cavlc	10	11	1 124
alu2	10	6	636	ctrl	7	26	204
frg1	28	3	179	dec	8	256	649
alu4	14	8	1 158	i2c	147	142	2 103
unreg	36	16	163	int2float	11	7	379
x2	10	7	74	mem_ctrl	1 204	1 231	79 153
count	35	16	203	priority	128	8	1 270
term1	34	10	276	router	60	30	316
				voter	1 001	1	16 240

本文实验中面积优化程度用式 (4) 所示的面积优化率 (area reduction rate, AR) 来表示:

$$A_R = \frac{A_{org} - A_{app}}{A_{org}} \quad (4)$$

式中: A_{org} 为原始电路的面积; A_{app} 为采用近似方法后得到的电路面积。

表 2 为本文算法与文献 [10] 进行对比的结果, 测试的电路与错误率设置与文献 [10] 一样, 将错误率设为 25%。不同之处在于文献 [10] 的运行环境, 其处理器频率为 3.4 GHz, 内存为 32 GB, 操作系统为 Linux。由表 2 可知, 对于 MCNC 测试电路, 在 25% 错误率的约束下, 本文算法的平均错误率可达 24.08%, 且绝大多数电路的错误率均非常接近 25%。文献 [10] 的平均错误率为 14.00%, 与限定的 25% 之间还有较大的提升空间。由表 2 可知, 在多数情况下, 文献 [10] 的方法不能很好地逼近 25% 这个设定的阈值, 尤其在处理 z4ml、unreg 电路时表现更为明显, 导致无法在 25% 的错误率约束下对上述 2 个电路进行任何近似优化。其原因是文献 [10] 提供的 AIG 节点删除策略无法实现对节点的精细删除。如在多扇出节点的处理上, 文献 [10] 采用所有扇出均一次性用常量替代的方法, 这种替代方法很容易造成错误率突破约束值, 导致无法充分利用错误率约束提供的优化空间。本文采用逐个节点的精细优化策略, 在处理多扇出节点时, 先将多扇出节点分解成多个单扇出节点, 然后逐个用常量替代, 这样的处理使得每次删除节点的数量不是很大, 错误率不易大幅增加而超出约束, 进而在逼近错误率的同时尽可能删除节点, 以获得更好的面积优化效果。此外, 从表 2 数据可以发现, 在面积优化的同时, 电路的延迟也得到了显著的优化, 这是由于在面积优化的过程中, 会删除关键路径上的节点, 从而优化电路延迟。与文献 [10] 相比, 本文算法对延迟优化也具有明显优势。

表 2　本文算法与文献 [10] 的比较结果

Table 2　Comparison results of the proposed algorithms and Ref.[10]　%

电路名称	面积优化		延时优化		错误率	
	文献[10]	本文算法	文献[10]	本文算法	文献[10]	本文算法
cm163	53.85	68.85	47.37	67.48	21.00	24.82
z4ml	0.00	42.62	0.00	27.72	0.00	24.99
alu2	9.09	57.70	0.00	46.02	24.00	24.43
frg1	2.49	98.32	0.00	95.58	16.00	24.32
alu4	6.01	50.09	0.00	26.91	22.00	24.76
unreg	0.00	34.36	0.00	0.00	0.00	24.99
x2	10.81	72.27	1.72	38.46	12.00	24.25
count	57.85	58.13	79.45	80.62	24.00	24.77
term1	46.73	93.12	27.03	78.72	7.00	19.28

由于文献 [11] 提供了算法的开源代码, 因此为保证算法速度的可比较性, 将本文算法和文献 [11] 算法放在在同一运算环境中进行测试, 结果如表 3 所示。由于错误率的增加会极大地延长文献 [11] 算法的运行时间, 因此, 表 3 中将文献 [11] 的错误率约束设置为 1%。同时考虑到文献 [11] 算法对同一电路的优化结果具有随机性, 因而表格中给出的数据为文献 [11] 算法 3 次测试结果的平均值。

由表 3 可知, 总体面积优化上, 本文算法的优化效果要略优于文献 [11]。在运算时间上, 两者在小规模电路上的耗时接近, 但在处理 arbiter、mem_ctrl 和 voter 等规模较大电路时, 文献 [11] 算法效率较低。这是由于文献 [11] 每次优化迭代获得的面积减少量, 对待处理的大电路来说微乎其微, 因此, 需要大量的迭代优化, 且每次迭代耗时又较长, 尤其是错误率设置比较大时, 会导致整个算法运行时间变得很长。这也是本文设置错误率为 1% 的原因。

从表 3 可以看出, 对于 priority 和 router 电路,

表 3　本文算法与文献 [11] 的比较结果

Table 3　The comparison results of the algorithms of this paper and Ref.[11]

电路名称	面积优化/%		耗时/s	
	文献[11]	本文算法	文献[11]	本文算法
arbiter	90.15	90.79	9 015.06	8 593.68
cavlc	5.70	7.47	375.57	67.05
ctrl	3.80	4.90	2.24	3.67
dec	0.62	0.62	8.45	10.06
i2c	31.59	37.95	189.50	176.49
int2float	16.30	23.48	10.86	51.71
mem_ctrl		62.66	>86 400.00	2 278.56
priority	96.92	98.27	14.95	36.89
router	100.00	98.73	2.00	0.18
voter	5.72	0.09	13 504.23	177.04

在错误率约束范围内，优化程度很大，几乎可以将电路内部节点全部删除。其原因是 priority 和 router 电路的 POs 只对少数几种输入情况敏感，而其他输入情况不会影响电路的输出结果，因此，在大量删除内部节点的前提下，错误率也没有超出约束。总体来说，相较于文献 [11]，即使在不考虑 mem_ctrl 测试电路的情况下，在面积优化提升了 1.28% 的情况下，本文算法平均耗时也减少了 60.91%。

5 结 论

1）本文针对多扇出节点常量替换过程容易发生的突破错误率约束的问题，提出多扇出节点常量替换过程中的错误率控制算法。同时，针对大规模电路，采用蒙特卡罗模拟的方式以较小的计算误差大幅减少了程序耗时，达到了兼顾算法速度和错误率计算精度的目的。

2）本文通过逻辑复制的方式，可以精细地控制常量替换节点的位置，从而使得该方式能够充分利用错误率约束下的优化空间，得到比对比算法更优秀的面积优化结果。

3）相较于 ALSRAC 优化算法，由于本文采取了蒙特卡罗模拟的错误率计算方式，因此本文算法在保证优化效果相近的前提下，令程序运行时间降低了 60%。

在逻辑综合领域中，电路面积上的优化固然重要，但延迟上的优化也是很多设计人员所要关注的重点。基于本文提出的多级逻辑电路近似优化模型，在今后的研究工作中，会以此开展对电路延迟优化方面的研究。

参考文献（References）

[1] WALDROP M M. The chips are down for Moore's law[J]. Nature, 2016, 530(7589): 144-147.

[2] HAN J, ORSHANSKY M. Approximate computing: An emerging paradigm for energy-efficient design[C]//Proceedings of the 18th IEEE European Test Symposium. Piscataway: IEEE Press, 2013: 1-6.

[3] SCARABOTTOLO I, ANSALONI G, CONSTANTINIDES G A, et al. A formal framework for maximum error estimation in approximate logic synthesis[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2022, 41(4): 840-853.

[4] SCARABOTTOLO I, ANSALONI G, CONSTANTINIDES G A, et al. Approximate logic synthesis: A survey[J]. Proceedings of the IEEE, 2020, 108(12): 2195-2213.

[5] PASANDI G, NAZARIAN S, PEDRAM M. Approximate logic synthesis: A reinforcement learning-based technology mapping approach[C]//Proceedings of the 20th International Symposium on Quality Electronic Design. Piscataway: IEEE Press, 2019: 26-33.

[6] ZHOU Z Z, YAO Y, HUANG S Y, et al. DALS: Delay-driven approximate logic synthesis[C]//Proceedings of the IEEE/ACM International Conference on Computer-Aided Design. Piscataway: IEEE Press, 2018: 1-7.

[7] VENKATARAMANI S, ROY K, RAGHUNATHAN A. Substitute-and-simplify: A unified design paradigm for approximate and quality configurable circuits[C]//Proceedings of the Design, Automation & Test in Europe Conference & Exhibition. Piscataway: IEEE Press, 2013: 1367-1372.

[8] SU S B, WU Y, QIAN W K. Efficient batch statistical error estimation for iterative multi-level approximate logic synthesis[C]//Proceedings of the 55th ACM/ESDA/IEEE Design Automation Conference. Piscataway: IEEE Press, 2018: 1-6.

[9] PASANDI G, PETERSON M, HERRERA M, et al. Deep-PowerX: A deep learning-based framework for low-power approximate logic synthesis[C]//Proceedings of the ACM/IEEE International Symposium on Low Power Electronics and Design. New York: ACM, 2020: 73-78.

[10] CHANDRASEKHARAN A, SOEKEN M, GROßE D, et al. Approximation-aware rewriting of AIGs for error tolerant applications[C]//Proceedings of the 35th International Conference on Computer-Aided Design. New York: ACM, 2016: 1-8.

[11] MENG C, QIAN W K, MISHCHENKO A. ALSRAC: Approximate logic synthesis by resubstitution with approximate care set[C]//Proceedings of the 57th ACM/IEEE Design Automation Conference. Piscataway: IEEE Press, 2020: 1-6.

[12] WU Y, QIAN W K. ALFANS: Multilevel approximate logic synthesis framework by approximate node simplification[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2020, 39(7): 1470-1483.

[13] SU S B, MENG C, YANG F, et al. VECBEE: A versatile efficiency-accuracy configurable batch error estimation method for greedy approximate logic synthesis[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2022, 41(11): 5085-5099.

[14] BRAND D. Verification of large synthesized designs[C]//Proceedings of the International Conference on Computer Aided Design. Piscataway: IEEE Press, 1993: 534-537.

[15] BRYANT. Graph-based algorithms for Boolean function manipulation[J]. IEEE Transactions on Computers, 1986, C-35(8): 677-691.

[16] MISHCHENKO A, CHATTERJEE S, BRAYTON R. DAG-aware AIG rewriting: A fresh look at combinational logic synthesis [C]//Proceedings of the 43rd ACM/IEEE Design Automation Conference. Piscataway: IEEE Press, 2006: 532-535.

[17] BRAYTON R, MISHCHENKO A. ABC: An academic industrial-strength verification tool[C]//International Conference on Computer Aided Verification. Berlin: Springer, 2010: 24-40.

Area optimization of multilevel logic circuits using approximate computing

ZHAO Weikai[1], WANG Lunyao[1,*], CEN Xumeng[2], XIA Yinshui[1], CHU Zhufei[1]

(1. Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China;
2. School of Information Engineering, College of Science & Technology Ningbo University, Ningbo 315300, China)

Abstract: According to the reported multi-level logic approximation optimization algorithm which can not balance the optimization effect and working speed well for largecircuits. It is proposed to use constant substitution in conjunction with and-inverter-graph (AIG) to optimize the logic at multiple levels for circuit areas. A set of candidate nodes to be replaced and a constant value for the node to be replaced are determined by employing the proposed error rate control technique of constant substitution of multiple fan-out nodes and a number of parameters known as node output distance, constant transmission distance, and so forth. Furthermore, according to the circuit size, different error rate calculation methods are selected to improve the speed of the algorithm. The proposed algorithm is programmed in C and implemented with the ABC tool and tested with EPFL and MCNC bench-marks. The experimental results show, compared with the report's similar methods, the proposed algorithm can save 48.77% area. When compared to approximate logic synthesis by resubstitution with approximate care (ALSRAC), this method achieves a 1.28% increase in area optimization and a 60.91% reduction in running time.

Keywords: multi-level circuit; approximate computing; error constraint; constant substitution; logic optimization

Received: 2022-08-29; **Accepted**: 2022-09-16; **Published Online**: 2022-11-01 09:52

URL: link.cnki.net/urlid/11.2625.V.20221031.1449.003

Foundation items: National Natural Science Foundation of China (U1709218,61471211,61871242)

* **Corresponding author**. E-mail: wanglunyao@nbu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0709

基于竞争情景的双枢纽航线网络优化模型

何湘妮，李艳华*，阳杰
(北京交通大学 交通运输学院，北京 100044)

摘　　要：针对双枢纽机场不属于同一机场集团的航线网络重合问题，以机场为视角优化航线网络布局，提出竞争情景下的双枢纽航线网络优化模型。在P枢纽中位问题航线网络构建思想的基础上，以旅客计划延误成本和航线重合成本最小为目标，考虑双枢纽机场直飞和中转航线、旅客需求、中转航程等因素，建立了竞争情景的无容量限制严格多分配0-1整数规划双枢纽航线网络优化模型，并使用交互式的线性和通用优化求解器（LINGO）对模型进行求解。选取成渝机场群成都双流机场和重庆江北机场作为双枢纽进行实例分析，通过不同航线重合成本研究航线网络优化问题。结果表明：优化后的目标函数值下降明显，计划延误成本下降率随航线重合成本的升高而减小，其中，航线重合成本为2 000元和3 000元时，目标函数值的下降率最高，分别为32.8%和32.6%，此时航线网络计划延误成本和航线重合成本的平衡效果最好。可见，所提模型可以在充分考虑双枢纽机场竞争关系、兼顾计划延误成本的同时降低航线重合率，从而减少航线冗余导致的资源浪费，为存在竞争关系且追求协同发展的双枢纽合理调整航线网络提供指导。

关 键 词：航线网络；双枢纽；航线重合成本；计划延误成本；P枢纽中位；多目标整数规划
中图分类号：F560; V351; U8
文献标志码：A　　**文章编号**：1001-5965(2024)09-2902-10

航线网络是航班计划的制作基础，是运行控制的面向对象，也是收益管理的根本来源。航线网络对航空公司来说是其他一切工作的基础，对机场来说是吸引客流的重要支撑，所以优化航线网络是航空运输业研究的重中之重。随着航空运输业的发展，机场群的建设，尤其是机场群、一市两场及双枢纽运营模式的逐步发展，各机场之间对航线网络优化提出了新的要求。目前大多数对于多机场航线网络优化的都是以合作为条件，以协同发展为目标进行的，但对实际而言，大多数机场群内各枢纽机场都隶属于不同的机场集团，如成渝机场群内成都双流机场和重庆江北机场、粤港澳大湾区机场群内广州白云机场和深圳宝安机场都不属于同一机场集团，不同枢纽机场之间体现出充分的竞争关系，若以合作为前提对其进行航线网络优化不符合实际，故需要针对竞争模式下的双枢纽进行航线网络优化研究，以实现在竞争中的协同发展。

目前国内外对航线网络优化的研究通常从航空公司和机场2个角度进行。从航空公司的角度研究航线网络主要解决的是枢纽机场的选址问题（hub location problems, HLP），O'Kelly [1] 从航空公司的角度对航线网络进行研究，建立了一个面对枢纽中位问题的单分配非线性整数规划模型；Campbell[2] 提出多分配的HLP，并利用多分配P枢纽中位问题的解提供下限给单分配P枢纽中位数问题最优解；Soylu和Katip[3] 提出一个多目标的、无容量限制的多分配P枢纽中位问题，采用寻找帕累托边界的精确和启发式算法求解，最大限度地降低网络运输总

收稿日期：2022-08-13；**录用日期**：2022-10-05；**网络出版时间**：2022-12-14 15:51
网络出版地址：link.cnki.net/urlid/11.2625.V.20221214.1330.004
基金项目：北京市社会科学重点项目 (22JCB032)；中央高校基本科研业务费专项资金 (2021RC262)
* **通信作者**. E-mail：yh.li@bjtu.edu.cn

引用格式：何湘妮，李艳华，阳杰. 基于竞争情景的双枢纽航线网络优化模型 [J]. 北京航空航天大学学报，2024，50（9）：2902-2911.
HE X N，LI Y H，YANG J. Optimization model for dual hub airline networks based on competition scenario [J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50（9）：2902-2911（in Chinese）.

成本, 并尽可能减少换乘次数以提高旅客满意度; Andrade de Araújo 等[4] 考虑了 2 家公司间的竞争, 提出了 2 种混合整数线性规划公式, 采用精确的分支定界法解决多分配的枢纽质心问题; Mohri 等[5] 将机队规划与 HLP 结合, 提出了设计最佳枢纽网络和机队配置的见解。姜涛等[6] 对中枢辐射航线网络进行规划设计时考虑了成本和需求的不确定性, 采用了鲁棒优化的方法, 并将最短路算法加以改进得到精确算法, 最终得出的鲁棒解可以适用于多种可能的情景; 柏明国[7] 建立了无容量限制的非严格多分配 P 枢纽中位问题的混合整数线性规划模型, 提出一种结合了最短路算法的新的禁忌搜索算法; 黄佳和王庆云[8] 采取多目标优化的遗传算法对枢纽航线网络结构进行了研究; 高荣环和杨芳[9] 研究了竞争环境下的航空公司枢纽位置问题, 建立了相应模型并设计了求解算法; 乐美龙等[10] 针对机场容量包络曲线构建了随机需求下非严格多分配的两阶段混合整数随机规划模型。从机场角度对航线网络的研究起步较晚, Santos 和 Antunes[11] 针对机场网络容量扩展的问题, 提出以系统总吞吐量最大化为目标的优化模型, 并为航空当局的长期决策提供建议。陈梵驿等[12] 提出基于航班准点率、航空公司份额和功能定位的机场群航线优化原则, 构建了目标为机场群总体效益最大化的机场群航线网络优化模型; Zhang 等[13] 提出一个基于多机场系统内部协作的航线网络优化模型, 可促进多机场系统的协同发展, 充分发挥各机场的优势; 徐涛等[14] 考虑枢纽容量限制和拥堵问题, 在非严格的枢纽航线网络基础上建立了优化模型。吕宗磊等[15] 设计了面向空铁联运的枢纽航线网络优化模型, 在非严格的枢纽航线网络中加入高铁站作为非枢纽节点, 以有效地减缓枢纽机场拥堵问题。现有的研究在以航空公司为角度, 为以机场间充分合作为条件下的航线网络优化问题提供了解决方法, 但在以机场为角度、考虑机场间竞争关系的航线网络研究很少, 而是通常不考虑机场间的竞争关系, 或在模型的前提条件中假设机场群统一归同一个机场管理集团, 机场之间协作制度健全、行政阻碍少, 且暂未有研究考虑以航线重合成本最小为优化目标函数, 然而这与实际情况有出入。

本文以双枢纽机场为研究对象, 从机场视角出发, 在已有航线网络优化研究的基础上, 进一步探索枢纽中位问题的解决方案, 借鉴 P 枢纽中位问题的航线网络构建思想, 以旅客计划延误成本和航线重合成本最小为目标, 构建竞争情景的无容量限制严格多分配 0-1 整数规划双枢纽航线网络优化模型, 并将成渝机场群成都双流机场和重庆江北机场作为双枢纽进行实例分析。提出的航线网络优化模型可以作为对单枢纽航线网络优化研究和机场群航线网络优化研究的补充, 选择双枢纽作为研究对象, 既考虑了单个机场自身吸引客流的能力, 又具备整体规划的思路。根据双枢纽竞争关系构建模型, 引入航线重合成本, 又加入对降低航线重合率的思考, 为竞争情景下的双枢纽航线网络优化提供了可借鉴的理论和方法, 对实践也有一定指导意义。

1　问题描述

双枢纽机场航线网络是由某一特定区域内客源范围有所重叠、功能定位相近、运行指标相差不大的两座枢纽机场的航线组合形成的航线网络。

京津冀、长三角、粤港澳大湾区是中国较为发达的几个地区, 迅速的社会发展和活跃的经济活动使得这些区域在航空运输方面产生了巨大的需求。基于此, 中国的区域多机场系统逐步建设并发展起来。其中, 地理区位、功能定位、运行指标都相近的 2 座在一定程度上可归为枢纽机场的机场, 即“双枢纽机场”。双枢纽机场该如何协调发展, 以促成良好的竞争合作关系, 逐渐成为民航领域的学者需要重点关注的问题。

目前主要的航空网络运营模式是中枢辐射式航线网络[5], 因为其特点是枢纽机场之间直接相连, 非枢纽机场可以与枢纽机场相连但不与其他非枢纽机场相连, 即非枢纽机场之间的客流必须经过枢纽机场中转, 而点对点航线网络的特点是所有存在需求客流的机场均已开通航线。因此, 与点对点航线网络相比, 中枢辐射式航线网络更能体现规模经济, 在实践中的应用也更广。

中枢辐射式航线网络根据非枢纽点和枢纽点是否唯一连接, 分为单分配枢纽航线网络和多分配枢纽航线网络, 每个非枢纽点只能选择一个枢纽点连接为单分配, 一个非枢纽点可以连接多个枢纽点为多分配; 根据非枢纽点是否必须经过枢纽点连接, 分为严格和非严格的枢纽航线网络, 非枢纽点可以直接相连为严格, 否则为非严格; 根据是否存在机场或航路的容量限制, 分为有容量限制和无容量限制的枢纽航线网络。

HLP 提供了基于真实世界运输和电信系统具有挑战性模型的丰富来源。从广义上讲, HLP 包括定位枢纽设施, 其设计了枢纽网络和确定通过网络的流量路由, 同时优化了基于成本或服务的目标[16]。根据 Alumur 和 Kara 的综述[17], HLP 可分为 P 枢纽

中位问题、固定成本的枢纽选址问题、P枢纽中心问题和枢纽覆盖问题，其中，P枢纽中位问题的目标是最小化为给定的一组客流量提供服务所需的总运输成本(时间、距离等)，给定 n 个需求节点，起点-终点 (origin-destination, OD) 对之间的客流量及要定位的枢纽数量 p。本文所述问题正是从 n 个备选机场中选出 p 个作为体现中转功能的枢纽机场，再确定各机场节点的具体连接方式，这符合P枢纽中位问题的描述。P枢纽中位问题中，基于枢纽点的航线网络优化设计方法的步骤总共分为3步：①确定所需枢纽个数；②确定具体枢纽位置；③通过设置约束条件来确定枢纽点与非枢纽点、非枢纽点相互之间的连接形式。

本文基于枢纽点即枢纽机场，借鉴P枢纽中位问题的航线网络构建思想，结合机场间的竞争关系，构建中枢辐射式航线网络优化模型，旨在为双枢纽机场的航线网络布局提供更加合理的方案。

2 模型建立及求解

2.1 模型假设

针对上述问题，构建了无容量限制的严格多分配双枢纽航线网络优化模型。模型假设如下：

1) 不考虑机场和航线上的容量限制。

2) 一个非枢纽机场可与一个或多个枢纽机场之间建立航线。

3) 旅客需要在机场开放时段内服从均匀分布。

4) 不考虑航空公司之间的影响。

5) 系统外部的非枢纽机场可与双枢纽机场建立直飞航线，也可建立通过系统外部枢纽机场中转至双枢纽机场的中转航线，非枢纽机场之间不可建立航线。

2.2 模型建立

机场之间的竞争是基于客源区域展开的。旅客对机场的认可和选择决定了其客源区域面积、市场份额、收入和利润。若双枢纽之间充分竞争将会导致双枢纽航线网络同质化严重，也不利于双枢纽地区的协同发展。航线同质化主要体现为航线重合成本，为减弱航线同质化，添加航线重合成本最小为目标函数，并根据双枢纽机场间可通过地面运输中转的特点，考虑双枢纽机场直飞和中转航线、旅客需求、中转航程等限制因素，建立无容量限制严格多分配0-1整数规划双枢纽航线网络优化模型。

2.2.1 目标函数

以通过双枢纽外的枢纽机场中转的计划延误成本、航线重合成本和直飞或通过双枢纽中转的计划延误成本、直飞的航线重合成本为目标建立目标函数为

$$\min(Z)=\sum_{i\in M}\sum_{j\in N}\sum_{k\in M}w_{ij}c_{ijk}^{\mathrm{sd}}x_{ijk}+\sum_{i\in M}\sum_{k\in M}c^{\mathrm{ol}}x_{x1k}x_{x2k}+\sum_{i\in M}\sum_{j\in N}\sum_{l\in M}w_{ij}c_{ijl}^{\mathrm{sd}}x_{ijl}+\sum_{i\in M}c^{\mathrm{ol}}x_{x11}x_{x22} \tag{1}$$

式中：M和N分别为双枢纽系统外部和内部的机场集合；w_{ij}为机场i至机场j的OD流；c_{ijk}^{sd}为从机场i经双枢纽外部枢纽机场k到双枢纽内部机场j的单位旅客计划延误成本，包括从机场i到机场k和从机场k到机场j的单位旅客计划延误成本，即$c_{ijk}^{\mathrm{sd}}=t_{ik}^{\mathrm{sd}}c_{ik}^{\mathrm{sd}}+t_{kj}^{\mathrm{sd}}c_{kj}^{\mathrm{sd}}$，$t_{ik}^{\mathrm{sd}}$和$t_{kj}^{\mathrm{sd}}$分别为从机场$i$到外部枢纽机场$k$和从外部枢纽机场$k$到机场$j$的计划延误时间，$c_{ik}^{\mathrm{sd}}$和$c_{kj}^{\mathrm{sd}}$分别为从机场$i$到外部枢纽机场$k$和从外部枢纽机场$k$到机场$j$的单位旅客单位时间计划延误成本；$x_{ijk}$、$x_{ijl}$为0-1决策变量，当开通$i-k/l-j$的航线时为1，否则为0；$c^{\mathrm{ol}}$为航线重合成本；$c_{ijl}^{\mathrm{sd}}$为从双枢纽外部机场$i$经双枢纽内部枢纽机场$l$到另一个双枢纽机场$j$的单位旅客计划延误成本，由于双枢纽机场内部机场$l$至另一个双枢纽机场$j$的距离较近，故从机场$l$至机场$j$采取地面运输方式，所以$c_{ijl}^{\mathrm{sd}}$包括从机场$i$到机场$l$的单位旅客计划延误成本和从机场$l$到机场$j$的单位地面运输成本，即$c_{ijl}^{\mathrm{sd}}=t_{il}^{\mathrm{sd}}c_{il}^{\mathrm{sd}}+c_{lj}^{\mathrm{g}}$。若$l=j$，那么$c_{lj}^{\mathrm{g}}=0$，$t_{il}^{\mathrm{sd}}$为从机场$i$到内部枢纽机场$l$的计划延误时间，$c_{il}^{\mathrm{sd}}$为从机场$i$到内部枢纽机场$l$的单位旅客单位时间计划延误成本，$c_{lj}^{\mathrm{g}}$为从机场$l$到机场$j$的地面运输成本。

航空运输的计划延误被定义为旅客的首选出发时间和最近的可用出发时间之间的差异。根据文献[18]，每名旅客的计划延误是平均航班间隔的$\frac{1}{4}$，即$t_{ik}^{\mathrm{sd}}=T/4f_{ik}$。同理，$t_{kj}^{\mathrm{sd}}=T/4f_{kj}$，$t_{il}^{\mathrm{sd}}=T/4f_{il}$，$T$为旅客具有出行意向的时间段，$f_{ik}$、$f_{kj}$和$f_{il}$分别为从机场$i$到外部枢纽机场$k$、从外部枢纽机场$k$到机场$j$和从机场$i$到内部枢纽机场$l$的航班频率。

在文献[19]中，将时间折算为费用的系数为VOT，用于表示旅客在单位时间内出行的费用支付意愿。本文的计划延误成本可解释为一种机会成本，即在计划延误的时间里本可以获得却没有获得的收入。航线重合成本可解释为由于航线冗余但客流不足而导致的资源浪费，具体包括空域资源、时刻资源的占用成本和设施设备的重复建设成本，以及双枢纽机场或地方政府为引流而给予航空公司的航线补贴。

2.2.2 约束条件

综合考虑双枢机场的特点和运行规划要求，约束条件如下：

$$H_k \in \{0,1\}, \forall k \in M \tag{2}$$

$$H_l \in \{0,1\}, \forall l \in N \tag{3}$$

$$\sum_{k \in M} H_k = P_1 \tag{4}$$

$$\sum_{l \in N} H_l = P_2 = 2 \tag{5}$$

$$x_{ijk} \in \{0,1\}, \forall i,k \in M, i \neq k, \forall j \in N \tag{6}$$

$$x_{ijl} \in \{0,1\}, \forall i \in M, \forall j,l \in N \tag{7}$$

$$x_{ijk} \leqslant H_k, \forall i,k \in M, i \neq k, \forall j \in N \tag{8}$$

$$\sum_{k \in M} x_{ijk} + \sum_{l \in N} x_{ijl} = 1, \forall i \in M, i \neq k, \forall j \in N \tag{9}$$

$$\left(d_{ik} + d_{kj} - 1.5 d_{ij}\right) x_{ijk} \leqslant 0, \forall i,k \in M, i \neq k, \forall j \in N \tag{10}$$

式中: H_k和H_l为 0-1 变量, 表示机场节点, 当k或l是枢纽机场时为 1, 否则为 0; P_1和P_2分别为双枢纽系统外部和双枢纽系统内部枢纽机场个数, 其中, 双枢纽系统内部有$P_2 = 2$个枢纽; x_{ijk}和x_{ijl}为 0-1 变量, 当开通i—k/l—j的航线时为 1, 否则为 0; d_{ij}为机场i至机场j的距离; OD 流只能通过枢纽机场中转, 即只有k是枢纽机场时, $i-k-j$航线才能开通; 所有 OD 流都应被满足, 且只能在直达和中转之间选择一种方式; 中转总航程不得大于直飞航程的 1.5 倍。

3　实例分析

成都双流机场和重庆江北机场是成渝机场群中势均力敌的 2 大重量级机场。根据 2022 年 3 月民航发〔2022〕15 号[20] 的规定, 指出要以成都、重庆国际航空枢纽功能建设为牵引, 加快构建高效协同发展体系。根据要求, 成都双流机场和重庆江北机场的合作是主基调, 同时由于二者隶属于不同的机场集团公司, 因此在遵循机场群协同发展的基础上, 客观上存在一定竞争关系。本文选取成渝机场群内成都双流机场和重庆江北机场作为验证竞争情形下的双枢纽航线网络优化模型的实例进行验证分析。

3.1　参数设定及原始情况

选取中国各地区共 30 个机场(长沙黄花 CSX[1]、青岛胶东 TAO、杭州萧山 HGH、石家庄正定 SJW、乌鲁木齐地窝堡 URC、福州长乐 FOC、泉州晋江 JJN、海口美兰 HAK、厦门高崎 XMN、深圳宝安 SZX、珠海金湾 ZUH、宁波栎社 NGB、温州龙湾 WNZ、上海虹桥 SHA、沈阳桃仙 SHE、哈尔滨太平 HRB、拉萨贡嘎 LXA、长春龙嘉 CGQ、大连周水子 DLC、黔江武陵山 JIQ、兰州中川 LHW、武汉天河 WUH、毕节飞雄 BFJ、遵义新舟 ZYI、信阳明港 XAI、宜春明月山 YIC、银川河东 INC、阿坝红原 AHJ、济南遥墙 TNA、襄阳刘集 XFN)作为双枢纽(成都双流 CTU、重庆江北 CKG)之外的机场。结合飞常准软件所查的航线数据及《从统计看民航 2020》中的航司机队配置和 2019 年中国主要航段的正班客座率, 由 150×0.8×航班频率保守计算得到客流量数据[21]。根据国家统计局 2019 全年全国居民人均可支配收入[22], 取单位旅客单位时间计划延误成本c_{ij}^{sd} = 15 元/(人·h)。根据成都双流机场至重庆江北机场的火车、高铁动车、公交地铁(含机场穿梭车)的票价和自驾车的费用, 取单位地面运输成本为c_{lj}^{g} = 150 元/人, 航程数据来自机场间距离的查询网站。取双枢纽外的枢纽个数$P_1 = 10$, 旅客具有出行意向的时间段$T = 18$ h。优化前的航线情况, 如图 1 所示。此时若不考虑航线重合成本, 目标函数值等于计划延误成本, 即 619 723 元, 则重合航线占所有航线的 68.1%。

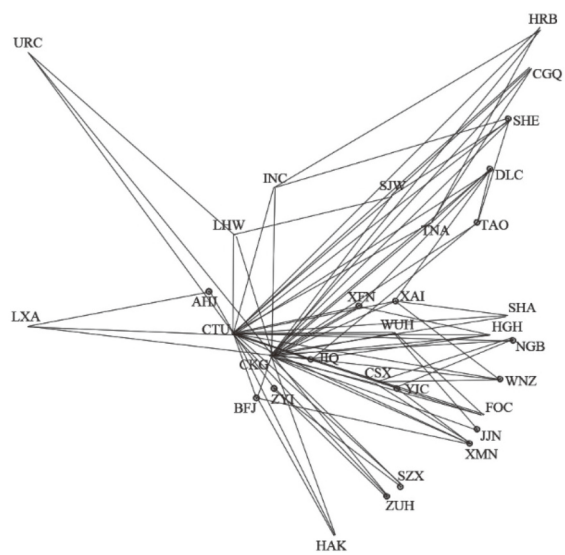


图 1　原航线网络

Fig. 1　Original airline network

此处的重合航线是指通达成都双流和重庆江北的所有相同的出发机场及所有出发及中转相同的机场连成的航线, 如哈尔滨太平—成都双流和哈尔滨太平—重庆江北、哈尔滨太平—银川河东—成都双流和哈尔滨太平—银川河东—重庆江北均属于重合航线; 哈尔滨太平—银川河东—成都双流、哈尔滨太平—济南遥墙—重庆江北、沈阳桃仙—银川河东—成都双流和哈尔滨太平—银川河东—重庆江北、沈阳桃仙—重庆江北、银川河东—重庆江北均不属于重合航线。

接下来将针对不同的航线重合成本c^{ol}进行求解算法设计及航线优化分析。

3.2 求解算法设计

由于交互式的线性和通用优化求解器(linear interactive and general optimizer, LINGO)语言在建立了集合的基础上,能够高效表达目标函数和各种约束条件,为求解以上模型并对上述航线进行优化分析,故以LINGO为求解工具设计求解算法:

1)定义集合部分。客流量数据为30×2的矩阵,距离和计划延误时间数据为30×32的矩阵,中转的计划延误成本和决策变量x_{ijk}为30×2×30的矩阵,直达的计划延误成本和决策变量x_{ijl}为30×2×2的矩阵,首先,定义一个包含1、2、…、30的一维集合M,属性有变量H_k的集合H,一个包含1、2的一维集合N,一个包含1、2、…、32的一维集合A;其次,定义包含M和N的派生二维集合B,属性有变量w_{ij}的集合W,包含M和A的派生二维集合C,属性有变量d_{ij}的集合d和变量t_{ij}^{sd}的集合t,包含M、N、M的派生三维集合K,属性有变量c_{ijk}^{sd}的集合E和决策变量x_{ijk}的集合I,包含M、N、N的派生三维集合L,属性有变量c_{ijl}^{sd}的集合F和决策变量x_{ijl}的集合J。

2)变量赋值部分。使用ole函数输入客流量和计划延误时间到W和t,再输出H、I和J的结果。

3)逻辑部分。使用for函数和sum函数表示目标函数和约束条件。

3.3 优化结果分析

为验证模型的有效性,针对上述参数设定,对川渝双枢纽航线网络进行优化,并讨论不同航线重合成本对航线网络优化情况的影响。

3.3.1 航线重合成本

利用LINGO求解优化模型,当航线重合成本c^{ol}为0时,全局最优解对应的航线优化情况如表1和表2所示,据此绘制航线网络如图2所示。

表1 $c^{ol}=0$优化中转航线网络

Table 1 Optimized transit airline network at $c^{ol}=0$

出发机场	到达机场	中转机场	出发机场	到达机场	中转机场
青岛胶东TAO	成都双流CTU	上海虹桥SHA	哈尔滨太平HRB	成都双流CTU	杭州萧山HGH
青岛胶东TAO	重庆江北CKG	上海虹桥SHA	哈尔滨太平HRB	重庆江北CKG	杭州萧山HGH
泉州晋江JJN	成都双流CTU	武汉天河WUH	长春龙嘉CGQ	成都双流CTU	杭州萧山HGH
海口美兰HAK	成都双流CTU	深圳宝安SZX	长春龙嘉CGQ	重庆江北CKG	济南遥墙TNA
宁波栎社NGB	成都双流CTU	深圳宝安SZX	大连周水子DLC	成都双流CTU	杭州萧山HGH
温州龙湾WNZ	成都双流CTU	深圳宝安SZX	大连周水子DLC	重庆江北CKG	杭州萧山HGH
沈阳桃仙SHE	重庆江北CKG	杭州萧山HGH			

表2 $c^{ol}=0$优化直飞航线网络

Table 2 Optimized direct airline network at $c^{ol}=0$

出发机场	到达机场	出发机场	到达机场	出发机场	到达机场
长沙黄花CSX	成都双流CTU	深圳宝安SZX	重庆江北CKG	武汉天河WUH	重庆江北CKG
长沙黄花CSX	重庆江北CKG	珠海金湾ZUH	成都双流CTU	毕节飞雄BFJ	成都双流CTU
杭州萧山HGH	成都双流CTU	珠海金湾ZUH	重庆江北CKG	毕节飞雄BFJ	重庆江北CKG
杭州萧山HGH	重庆江北CKG	宁波栎社NGB	重庆江北CKG	遵义新舟ZYI	成都双流CTU
石家庄正定SJW	成都双流CTU	温州龙湾WNZ	重庆江北CKG	信阳明港XAI	成都双流CTU
石家庄正定SJW	重庆江北CKG	上海虹桥SHA	成都双流CTU	信阳明港XAI	重庆江北CKG
乌鲁木齐地窝堡URC	成都双流CTU	上海虹桥SHA	重庆江北CKG	宜春明月山YIC	成都双流CTU
乌鲁木齐地窝堡URC	重庆江北CKG	沈阳桃仙SHE	成都双流CTU	宜春明月山YIC	重庆江北CKG
福州长乐FOC	成都双流CTU	拉萨贡嘎LXA	成都双流CTU	银川河东INC	成都双流CTU
福州长乐FOC	重庆江北CKG	拉萨贡嘎LXA	重庆江北CKG	银川河东INC	重庆江北CKG
泉州晋江JJN	重庆江北CKG	黔江武陵山JIQ	成都双流CTU	阿坝红原AHJ	成都双流CTU
海口美兰HAK	重庆江北CKG	黔江武陵山JIQ	重庆江北CKG	济南遥墙TNA	成都双流CTU
厦门高崎XMN	成都双流CTU	兰州中川LHW	成都双流CTU	济南遥墙TNA	重庆江北CKG
厦门高崎XMN	重庆江北CKG	兰州中川LHW	重庆江北CKG	襄阳刘集XFN	成都双流CTU
深圳宝安SZX	成都双流CTU	武汉天河WUH	成都双流CTU	襄阳刘集XFN	重庆江北CKG

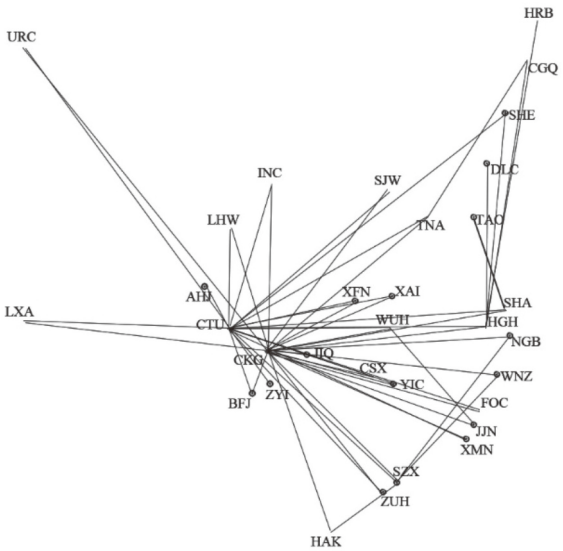


图 2　$c^{ol}=0$ 优化航线网络

Fig. 2　Optimized airline network at $c^{ol}=0$

此时的计划延误成本为 423 759 元, 较原计划下降了 31.6%, 目标函数值与计划延误成本相同。但航线重合率为 75.9%, 高于原航线重合率。为降低航线重合率, 下面讨论航线重合成本取 1 000 元～8 000 元的优化情况, 以航线重合成本为 4 000 元和 8 000 元的情况为例来具体说明。

取 c^{ol} 为 4 000 元, 全局最优解对应的航线优化情况如表 3 和表 4 所示, 据此绘制航线网络如图 3 所示。

此时的计划延误成本为 478 151 元, 较原计划延误成本下降了 22.8%, 航线重合率为 37.9%, 目标函数值为 522 151 元, 较原目标函数值 747 723 元下降了 30.2%。针对航线重合成本为 0 的情况, 石家庄正定、福州长乐、厦门高崎、珠海金湾、信阳明港至成都双的流航线分别从直飞改为从兰州中川、上海虹桥、武汉天河等机场进行中转, 乌鲁木齐地窝堡、上海虹桥至重庆江北的直飞航线分别改为在银川河东、厦门高崎进行中转, 黔江武陵山至成都双流的航线从直飞改为在重庆江北通过地面运输进行中转, 而青岛胶东中转上海虹桥至重庆江北的航线改

表 3　$c^{ol}=4\,000$ 元优化中转航线网络

Table 3　Optimized transit airline network at $c^{ol}=4\,000$ RMB

出发机场	到达机场	中转机场	出发机场	到达机场	中转机场
青岛胶东TAO	成都双流CTU	上海虹桥SHA	上海虹桥SHA	重庆江北CKG	厦门高崎XMN
石家庄正定SJW	成都双流CTU	兰州中川LHW	沈阳桃仙SHE	重庆江北CKG	杭州萧山HGH
乌鲁木齐地窝堡URC	重庆江北CKG	银川河东INC	哈尔滨太平HRB	成都双流CTU	杭州萧山HGH
福州长乐FOC	成都双流CTU	上海虹桥SHA	哈尔滨太平HRB	重庆江北CKG	济南遥墙TNA
泉州晋江JJN	成都双流CTU	武汉天河WUH	长春龙嘉CGQ	成都双流CTU	杭州萧山HGH
海口美兰HAK	成都双流CTU	深圳宝安SZX	长春龙嘉CGQ	重庆江北CKG	济南遥墙TNA
厦门高崎XMN	成都双流CTU	武汉天河WUH	大连周水子DLC	成都双流CTU	武汉天河WUH
珠海金湾ZUH	成都双流CTU	武汉天河WUH	大连周水子DLC	重庆江北CKG	杭州萧山HGH
宁波栎社NGB	成都双流CTU	深圳宝安SZX	信阳明港XAI	成都双流CTU	武汉天河WUH
温州龙湾WNZ	成都双流CTU	深圳宝安SZX	黔江武陵山JIQ	成都双流CTU	重庆江北CKG

表 4　$c^{ol}=4\,000$ 元优化直飞航线网络

Table 4　Optimized direct airline network at $c^{ol}=4\,000$ RMB

出发机场	到达机场	出发机场	到达机场	出发机场	到达机场
长沙黄花CSX	成都双流CTU	珠海金湾ZUH	重庆江北CKG	毕节飞雄BFJ	重庆江北CKG
长沙黄花CSX	重庆江北CKG	宁波栎社NGB	重庆江北CKG	遵义新舟ZYI	成都双流CTU
青岛胶东TAO	重庆江北CKG	温州龙湾WNZ	重庆江北CKG	信阳明港XAI	重庆江北CKG
杭州萧山HGH	成都双流CTU	上海虹桥SHA	成都双流CTU	宜春明月山YIC	成都双流CTU
杭州萧山HGH	重庆江北CKG	沈阳桃仙SHE	成都双流CTU	宜春明月山YIC	重庆江北CKG
石家庄正定SJW	重庆江北CKG	拉萨贡嘎LXA	成都双流CTU	银川河东INC	成都双流CTU
乌鲁木齐地窝堡URC	成都双流CTU	拉萨贡嘎LXA	重庆江北CKG	银川河东INC	重庆江北CKG
福州长乐FOC	重庆江北CKG	黔江武陵山JIQ	重庆江北CKG	阿坝红原AHJ	成都双流CTU
泉州晋江JJN	重庆江北CKG	兰州中川LHW	成都双流CTU	济南遥墙TNA	成都双流CTU
海口美兰HAK	重庆江北CKG	兰州中川LHW	重庆江北CKG	济南遥墙TNA	重庆江北CKG
厦门高崎XMN	重庆江北CKG	武汉天河WUH	成都双流CTU	襄阳刘集XFN	成都双流CTU
深圳宝安SZX	成都双流CTU	武汉天河WUH	重庆江北CKG	襄阳刘集XFN	重庆江北CKG
深圳宝安SZX	重庆江北CKG	毕节飞雄BFJ	成都双流CTU		

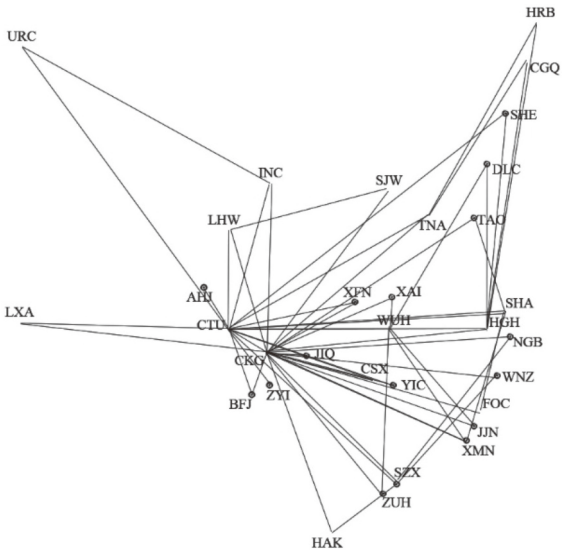


图3 $c^{ol}=4\,000$ 元优化航线网络

Fig. 3 Optimized airline network at $c^{ol}=4\,000$ RMB

为直飞。

取c^{ol}为8 000元，全局最优解对应的航线优化情况如表5和表6所示，据此绘制航线网络如图4所示。

此时的计划延误成本为637 065元，较原计划上涨了2.8%，因此，不再继续讨论航线重合成本升高的情况。航线重合率为13.8%，目标函数值为669 065元，较原目标函数值875 723元下降了23.6%。针对航线重合成本为4 000元的情况，杭州萧山至重庆江北、深圳宝安至成都双流、济南遥墙至成都双流航线从分别直飞改为在厦门高崎、武汉天河、银川河东进行中转，同时兰州中川、毕节飞雄、宜春明月山、襄阳刘集至重庆江北航线从直飞改为在成都双流通过地面运输进行中转。

3.3.2 优化结果讨论

不同航线重合成本c^{ol}（c^{ol}=[0,8 000]元）的航线

表5 $c^{ol}=8\,000$ 元优化中转航线网络

Table 5 Optimized transit airline network at $c^{ol}=8\,000$ RMB

出发机场	到达机场	中转机场	出发机场	到达机场	中转机场
青岛胶东TAO	成都双流CTU	上海虹桥SHA	哈尔滨太平HRB	成都双流CTU	杭州萧山HGH
杭州萧山HGH	重庆江北CKG	厦门高崎XMN	哈尔滨太平HRB	重庆江北CKG	济南遥墙TNA
石家庄正定SJW	成都双流CTU	兰州中川LHW	长春龙嘉CGQ	成都双流CTU	杭州萧山HGH
乌鲁木齐地窝堡URC	重庆江北CKG	银川河东INC	长春龙嘉CGQ	重庆江北CKG	济南遥墙TNA
福州长乐FOC	成都双流CTU	上海虹桥SHA	大连周水子DLC	成都双流CTU	武汉天河WUH
泉州晋江JJN	成都双流CTU	武汉天河WUH	大连周水子DLC	重庆江北CKG	杭州萧山HGH
海口美兰HAK	成都双流CTU	深圳宝安SZX	信阳明港XAI	成都双流CTU	武汉天河WUH
厦门高崎XMN	成都双流CTU	武汉天河WUH	济南遥墙TNA	成都双流CTU	银川河东INC
深圳宝安SZX	成都双流CTU	武汉天河WUH	黔江武陵山JIQ	成都双流CTU	重庆江北CKG
珠海金湾ZUH	成都双流CTU	武汉天河WUH	兰州中川LHW	重庆江北CKG	成都双流CTU
宁波栎社NGB	成都双流CTU	深圳宝安SZX	毕节飞雄BFJ	重庆江北CKG	成都双流CTU
温州龙湾WNZ	成都双流CTU	深圳宝安SZX	宜春明月山YIC	重庆江北CKG	成都双流CTU
上海虹桥SHA	重庆江北CKG	厦门高崎XMN	襄阳刘集XFN	重庆江北CKG	成都双流CTU
沈阳桃仙SHE	重庆江北CKG	杭州萧山HGH			

表6 $c^{ol}=8\,000$ 元优化直飞航线网络

Table 6 Optimized direct airline network at $c^{ol}=8\,000$ RMB

出发机场	到达机场	出发机场	到达机场	出发机场	到达机场
长沙黄花CSX	成都双流CTU	珠海金湾ZUH	重庆江北CKG	武汉天河WUH	重庆江北CKG
长沙黄花CSX	重庆江北CKG	宁波栎社NGB	重庆江北CKG	毕节飞雄BFJ	成都双流CTU
青岛胶东TAO	重庆江北CKG	温州龙湾WNZ	重庆江北CKG	遵义新舟ZYI	成都双流CTU
杭州萧山HGH	成都双流CTU	上海虹桥SHA	成都双流CTU	信阳明港XAI	重庆江北CKG
石家庄正定SJW	重庆江北CKG	沈阳桃仙SHE	成都双流CTU	宜春明月山YIC	成都双流CTU
乌鲁木齐地窝堡URC	成都双流CTU	拉萨贡嘎LXA	成都双流CTU	银川河东INC	成都双流CTU
福州长乐FOC	重庆江北CKG	拉萨贡嘎LXA	重庆江北CKG	银川河东INC	重庆江北CKG
泉州晋江JJN	重庆江北CKG	黔江武陵山JIQ	重庆江北CKG	阿坝红原AHJ	成都双流CTU
海口美兰HAK	重庆江北CKG	兰州中川LHW	成都双流CTU	济南遥墙TNA	重庆江北CKG
厦门高崎XMN	重庆江北CKG	武汉天河WUH	成都双流CTU	襄阳刘集XFN	成都双流CTU
深圳宝安SZX	重庆江北CKG				

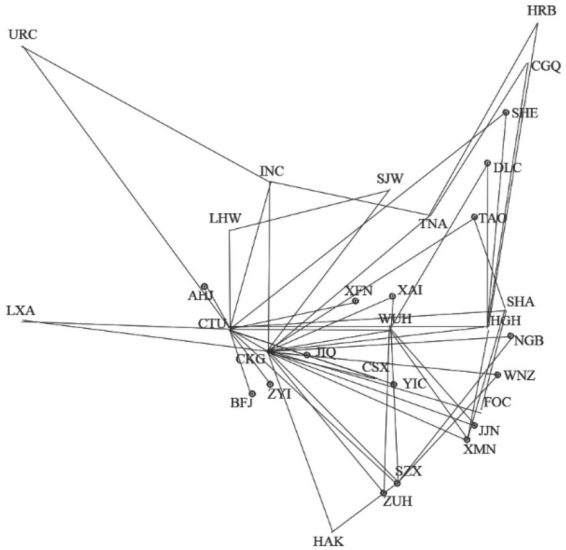


图 4　c^{ol} = 8 000 元优化航线网络

Fig. 4　Optimized airline network at c^{ol} = 8 000 RMB

优化结果如表 7 所示。

结果表明, 随着航线重合成本c^{ol}的增长, 计划延误成本下降率降低, 这意味着计划延误成本越来越高, 同时航线重合率降低, 也意味着航线重合情况逐步得到改善, 两者变化趋势如图 5 所示。

综合考虑计划延误成本和航线重合成本的目标函数值, 不同航线重合成本对应的目标函数值及下降率变化趋势如图 6 所示。

通过分析可知, 当航线重合成本为 0 时, 目标函数值的下降幅度最大, 但航线高度重合。为降低重合率, 将原本重合的中转航线改为直飞航线或改变中转机场, 原本重合的直飞航线改为中转航线。航线重合成本增长到一定程度后会出现通过地面运输中转的航线。当航线重合成本为 8 000 元时, 航线重合情况最轻, 但计划延误成本过高。当航线重合成本为 2 000～3 000 元时, 目标函数值降低幅度最大, 分别为 32.8% 和 32.6%, 意味着此时计划延误成本和航线重合成本的平衡效果最好。

如果要将重合的直飞航线改为中转航线, 那么武汉天河机场是最佳选择。在目标函数值下降幅度最大、航线重合成本取 2 000 元时, 7 条有改动的航线里有 4 条选择在武汉天河机场中转。这是因为武汉天河机场位于川渝和东部沿海较发达地区的中间位置, 比较适合作为中转机场。

综上, 构建的模型可以在降低航线重合率的同时兼顾计划延误成本, 使优化后的航线网络总成本

表 7　不同航线重合成本航线优化结果

Table 7　Optimization results of overlapped airline cost for different airlines

状态	c^{ol}	计划延误成本/元	计划延误成本下降率/%	航线重合率/%	目标函数现值/元	目标函数原值/元	目标函数下降率/%
优化前		619 723		68.1			
	0	423 759	31.6	75.9	423 759	619 723	31.6
	1 000	429 245	30.7	51.7	459 245	651 723	29.5
	2 000	429 245	30.7	51.7	459 245	683 723	32.8
	3 000	440 238	29.0	48.3	482 238	715 723	32.6
优化后	4 000	478 151	22.8	37.9	522 151	747 723	30.2
	5 000	532 151	14.1	27.6	572 151	779 723	26.6
	6 000	554 290	10.6	24.1	596 290	811 723	26.5
	7 000	606 465	2.1	17.2	641 465	843 723	24.0
	8 000	637 065	−2.8	13.8	669 065	875 723	23.6

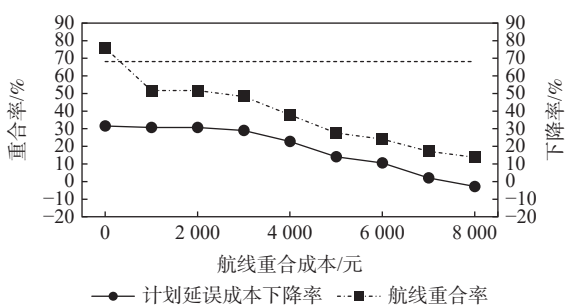


图 5　计划延误成本下降率及航线重合率变化趋势

Fig. 5　Trend of decrease rate of planned delay cost and overlapped airline rate

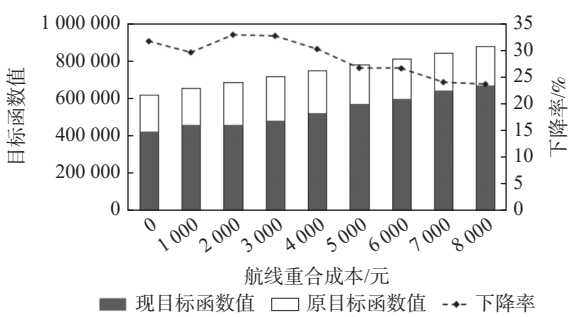


图 6　目标函数值变化趋势

Fig. 6　Trend of objective function value

最小，也可使航线重合与计划延误达到平衡。

4 结 论

本文主要解决竞争情境中双枢纽机场航线重合造成资源浪费时的航线网络优化问题，引入航线重合成本概念及降低航线重合率建立了双枢纽机场航线网络优化模型，并通过实例分析验证了模型对于降低航线重合率的有效性。

1）不同的航线网络成本所引起的旅客计划延误成本不同，航线网络布局决策者可选择不同航线的重合成本，以优化网线网络的总成本，使计划延误成本和航线重合成本达到平衡。

2）计划延误成本下降率随着航线重合成本的增长而降低，计划延误成本随之增高；反之，航线重合率减小，航线重合情况逐步得到改善。通过调整参数得到航线重合成本在2 000元～3 000元时目标函数值降低的幅度最大，牺牲一些计划延误成本来保证航线重合率不过高，能够更加合理地指导双枢纽航线网络优化。

3）旅客类型的不同所导致计划延误时间的不同，同时考虑到时效特征不同的航空货物，计划延误成本也会有所不同。因此未来可以结合分类不同的旅客出行和航空货物因素对双枢纽航线网络优化进行进一步的深入研究。

参考文献（References）

[1] O'KELLY M E. A quadratic integer program for the location of interacting hub facilities[J]. European Journal of Operational Research, 1987, 32(3): 393-404.

[2] CAMPBELL J F. Location and allocation for distribution systems with transshipments and transportion economies of scale[J]. Annals of Operations Research, 1992, 40(1): 77-99.

[3] SOYLU B, KATIP H. A multiobjective hub-airport location problem for an airline network design[J]. European Journal of Operational Research, 2019, 277(2): 412-425.

[4] ANDRADE DE ARAÚJO A C, ROBOREDO M C, PESSOA A A, et al. Exact methods for the discrete multiple allocation (r| p) hub-centroid problem[J]. Computers & Operations Research, 2020, 116: 104870.

[5] MOHRI S S, NASROLLAHI M, PIRAYESH A, et al. An integrated global airline hub network design with fleet planning[J]. Computers & Industrial Engineering, 2021, 164: 107883.

[6] 姜涛, 朱金福, 覃义. 基于最短路的中枢辐射航线网络鲁棒优化方法[J]. 系统工程, 2007, 25(1): 53-59.
JIANG T, ZHU J F, QIN Y. Robust optimization of hub-and-spoke airline network design based on shortest path algorithm[J]. Systems Engineering, 2007, 25(1): 53-59(in Chinese).

[7] 柏明国. 基于禁忌搜索算法的枢纽航线网络优化设计研究[J]. 经济数学, 2008, 25(4): 412-417.
BAI M G. Research on hub-and-spoke network design based on tabu search algorithm[J]. Mathematics in Economics, 2008, 25(4): 412-417(in Chinese).

[8] 黄佳, 王庆云. 基于多目标遗传算法的枢纽航线网络的鲁棒优化方法[J]. 交通运输系统工程与信息, 2009, 9(3): 86-92.
HUANG J, WANG Q Y. Robust optimization of hub-and-spoke airline network design based on multi-objective genetic algorithm[J]. Journal of Transportation Systems Engineering and Information Technology, 2009, 9(3): 86-92(in Chinese).

[9] 高荣环, 杨芳. 基于鲁棒性的竞争环境下枢纽航线网络优化设计[J]. 交通运输系统工程与信息, 2014, 14(5): 161-167.
GAO R H, YANG F. Optimal design of hub network in competitive environment based on robust[J]. Journal of Transportation Systems Engineering and Information Technology, 2014, 14(5): 161-167(in Chinese).

[10] 乐美龙, 郑文娟, 吴明功, 等. 不确定需求下航空公司枢纽网络优化设计[J]. 北京航空航天大学学报, 2020, 46(4): 674-682.
LE M L, ZHENG W J, WU M G, et al. Airline hub network optimization design under uncertain demand[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(4): 674-682(in Chinese).

[11] SANTOS M G, ANTUNES A P. Long-term evolution of airport networks: Optimization model and its application to the United States[J]. Transportation Research Part E: Logistics and Transportation Review, 2015, 73: 17-46.

[12] 陈梵驿, 杨新湜, 翟文鹏, 等. 基于决策树C4.5算法的京津冀机场群航线网络优化[J]. 中国科技论文, 2017, 12(7): 798-801.
CHEN F Y, YANG X S, ZHAI W P, et al. Optimization of the airline network of Beijing Tianjin Hebei region multi-airport system based on the decision tree C4.5 algorithm[J]. China Sciencepaper, 2017, 12(7): 798-801(in Chinese).

[13] ZHANG L, LI C, XU K, et al. A collaborative route optimization model for regional multi-airport system[J]. IEEE Access, 2019, 7: 152158-152165.

[14] 徐涛, 吴志帅, 卢敏, 等. 面向拥堵问题的枢纽航线网络优化模型[J]. 系统工程与电子技术, 2020, 42(11): 2553-2559.
XU T, WU Z S, LU M, et al. Optimization model of hub-and-spoke network for congestion problem[J]. Systems Engineering and Electronics, 2020, 42(11): 2553-2559(in Chinese).

[15] 吕宗磊, 吴志帅, 徐涛, 等. 面向空铁联运的枢纽航线网络优化模型[J]. 计算机工程与设计, 2021, 42(4): 1188-1194.
LYU Z L, WU Z S, XU T, et al. Hub-spoke airline network optimization model for air-high speed rail intermodal transportation[J]. Computer Engineering and Design, 2021, 42(4): 1188-1194(in Chinese).

[16] ALUMUR S A, CAMPBELL J F, CONTRERAS I, et al. Perspectives on modeling hub location problems[J]. European Journal of Operational Research, 2021, 291(1): 1-17.

[17] ALUMUR S, KARA B Y. Network hub location problems: The state of the art[J]. European Journal of Operational Research, 2008, 190(1): 1-21.

[18] 姜思露, 朱金福, 孔明星, 等. 基于旅客计划延误的航班频率优化研究[J]. 武汉理工大学学报 (交通科学与工程版), 2019, 43(1): 136-140.
JIANG S L, ZHU J F, KONG M X, et al. Research of flight frequency optimization based on passenger schedule delay[J]. Journal

of Wuhan University of Technology (Transportation Science & Engineering), 2019, 43(1): 136-140(in Chinese).

[19] 胡文君, 周溪召, 沈四林. 基于计划延误成本的出发时间选择模型[J]. 交通信息与安全, 2018, 36(2): 99-105.
HU W J, ZHOU X Z, SHEN S L. A choice model for departure time based on schedule delay costs[J]. Journal of Transport Information and Safety, 2018, 36(2): 99-105(in Chinese).

[20] 中国民用航空局. 民航局关于加快成渝世界级机场群建设的指导意见：民航发〔2022〕15 号[A/OL]. (2022-02-22)[2024-06-30]. https://www.gov.cn/zhengce/zhengceku/2022-03/15/content_5679156.htm.
Civil Aviation Administration of China. Guiding opinions of Civil Aviation Administration on accelerating the construction of Chengdu-Chongqing world-class airport cluster: CAAC[2022]No. 15 [A/OL]. (2022-02-22)[2024-06-30]. https://www.gov.cn/zhengce/zhengceku/2022-03/15/content_5679156.htm(in Chinese).

[21] 中国民用航空局. 从统计看民航 2020[M]. 北京: 中国民航出版社有限公司, 2021: 54-69.
Civil Aviation Administration of China. Civiaviation in statistical terms 2020[M]. Beijing: China Civil Aviation Publishing House, 2021: 54-69(in Chinese).

[22] 国家统计局. 2019 年居民收入和消费支出情况[EB/OL]. (2019-01-07) [2024-06-30]. https://www.stats.gov.cn/sj/zxfb/202302/t20230203_1900600.html.
National Bureau of Statistics. Personal income and consumption expenditure in 2019[EB/OL]. (2019-01-07)[2024-06-30]. https://www.stats.gov.cn/sj/zxfb/202302/t20230203_1900600.html(in Chinese).

Optimization model for dual hub airline networks based on competition scenario

HE Xiangni，LI Yanhua*，YANG Jie

(School of Traffic and Transportation，Beijing Jiaotong University，Beijing 100044，China)

Abstract：In view of the overlapped airline network of dual hub airports that do not belong to the same airport group, an optimization model of dual hub airline networks in a competition scenario was proposed by optimizing the airline network layout of the airports. Based on the airline network construction idea in the median of p hub, the goal was to minimize the planned delay cost of the travelers and overlapped airline cost, and factors such as direct and transit airlines in dual hub airports, passenger demand, and transit flight were considered. In addition, an optimization model for dual hub airline networks with strict multi-allocation 0–1 integer programming without capacity constraints in competitive scenarios was established, and LINGO was used to solve the model. Chengdu Shuangliu Airport and Chongqing Jiangbei Airport in Chengdu-Chongqing Airport Group were selected as dual hubs for case analysis, and the optimization of the airline network was studied based on the overlapped airline cost. The results show that the optimized objective function value decreases significantly, and the decrease rate of planned delay cost decreases with the increase in overlapped airline cost. When the overlapped airline costs are 2 000 RMB and 3 000 RMB, the decrease rate of the objective function value is the highest, which are 32.8% and 32.6%, respectively. In this case, the balance effect of the planned delay cost and the overlapped airline cost of the airline network is the best. It can be seen that the model can fully consider the competition relationship between dual hub airports, take into account the planned delay cost, lower the overlapped airline rate, reduce the waste of resources caused by airline redundancy, and provide guidance for the reasonable adjustment of the airline network of dual hub airports in competition and collaborative development.

Keywords：airline network；dual hub；overlapped airline cost；planned delay cost；median of P hub；multi-objective integer programming

Received：2022-08-13；**Accepted**：2022-10-05；**Published Online**：2022-12-14 15：51

URL：link.cnki.net/urlid/11.2625.V.20221214.1330.004

Foundation items：Beijing Social Science Project (22JCB032); The Fundamental Research Funds for the Central Universities (2021RC262)

* **Corresponding author**. E-mail：yh.li@bjtu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0712

双弯曲旋转超声椭圆振动加工系统设计

邵壮壮，郑侃*，董松，孙振文，孙连军，严嘉文
(南京理工大学 机械工程学院，南京 210094)

摘 要：由于旋转超声单一振动模式加工效率提升有限，因此采用超声振动激励源耦合法设计了一种双弯曲旋转超声椭圆振动加工系统。推导弯曲波动方程，计算的弯曲超声振动系统参数与仿真结果的理论分析误差仅为5%。耦合双弯曲旋转超声椭圆振动加工系统2个不同方向的振动，控制弯振频率偏差低于1%。对超声刀柄的法兰位置与形状进行有限元分析发现，开设减振槽有利于增大弯振振幅，并确定了法兰的最佳位置。对双弯曲旋转超声椭圆振动加工系统的振动频率和振幅进行测试，结果表明：所设计系统在工作频率和2个方向的振幅均满足设计预期。

关 键 词：双弯曲超声椭圆振动；激励源耦合；超声刀柄；弯振频率；弯振振幅
中图分类号：TH161
文献标志码：A **文章编号**：1001-5965(2024)09-2912-07

近年来，鉴于在加工过程中的降低切削力、降低切削热、提高表面质量、抑制颤振等方面表现出的优异性能[1-3]，旋转超声加工技术被广泛应用于加工陶瓷、复合材料及高性能金属材料等难加工材料上[4-6]，常规超声振动依靠单一方向的高频高振幅，实现刀具与被加工表面的分离，从而提升加工性能。但是，过高的频率和振幅增加了装置设计的难度，同时单一方向的振动模式使其应用的领域具有局限性。相较于常规超声振动，超声椭圆振动应用范围更为广泛。张德远教授团队验证了二维超声椭圆振动加工对加工性能的提升较为显著[7-11]。Gao和Altintas[12]研究了三维椭圆超声振动系统的设计与控制，指出三维超声椭圆振动对加工性能的影响。此外，旋转超声椭圆振动加工具有独特的微观结构，可显著提升零件某方面的性能[13]。Wang等[14]利用刀具与工作台间振动相互耦合的方式实现了超声椭圆振动，2个系统间的振动耦合限制了超声椭圆振动系统的应用范围。Pang等[15]、Chen等[16]和顾磊等[17]采用了模态转换的形式将2种超声振动进行耦合，基于有限元法在变幅杆上增加螺旋槽，设计了纵振与扭振耦合的变幅杆，主要是将换能器产生的部分纵振波经过变幅杆上螺旋槽的折射转换成扭振，最终在刀具上形成纵扭振动。将换能器的能量分散成2波分别作用，2个方向的振动耦合相互影响，这种模态转换方式获得的不同方向的振幅很难分别控制，并且一般转换后的振幅值较小。Wang等[18]采用2个纵振系统进行耦合，2个纵振系统分别激励不同方向上的纵振，在刀具处采用柔性结构将2种振动形式耦合成椭圆振动输出在刀具末端。在2个方向采用了2个系统耦合的形式实现了刀具超声椭圆振动，同时减少了2个方向的超声振动在工作时的相互干扰。Bai等[19]采用纵振与弯振耦合将超声椭圆振动应用到车削领域，超声椭圆振动表面粗糙度达到6 nm。Du等[20]和王晓雪等[21]采用2个换能器在一个系统上激励纵振和弯曲振动，2个方向的振动通过变幅杆将振幅放大传输至刀具末端，使不同方向的振动系统耦合后在刀具末端输出椭圆振动，并将此方式应用到旋转

收稿日期：2022-08-13；**录用日期**：2022-10-07；**网络出版时间**：2023-01-06 08:58
网络出版地址：link.cnki.net/urlid/11.2625.V.20230104.2053.007
基金项目：国家自然科学基金(52075265)
*通信作者. E-mail：zhengkan@njust.edu.cn

引用格式：邵壮壮，郑侃，董松，等. 双弯曲旋转超声椭圆振动加工系统设计[J]. 北京航空航天大学学报，2024，50(9)：2912-2918.
SHAO Z Z，ZHENG K，DONG S，et al. Design of double bending rotary ultrasonic elliptical vibration machining system[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2912-2918 (in Chinese).

超声加工领域。

相较于更改变幅杆结构实现椭圆振动的方式, 双激励源模式更具优势: ①更有利于实现超声场能的可控再分配; ②能够同时满足椭圆振动 2 个方向上的大振幅。针对双激励模式的换能器结构与法兰位置难以确定的问题, 本文设计了一种双弯曲旋转超声椭圆振动加工系统, 利用换能器的特殊结构激励 2 个方向的弯曲振动进行耦合, 在刀具末端实现高频稳定的超声椭圆振动, 具有较好的适用性。

1　双弯曲旋转超声椭圆振动加工系统结构设计

旋转超声振动加工系统主要包括换能器、变幅杆、电路传输部分和超声电源, 如图 1 所示。其中, 换能器和变幅杆是超声振动系统的重要组成部分, 本文通过解析法和有限元解析对换能器和变幅杆进行设计, 形成双弯曲旋转超声椭圆振动加工系统。

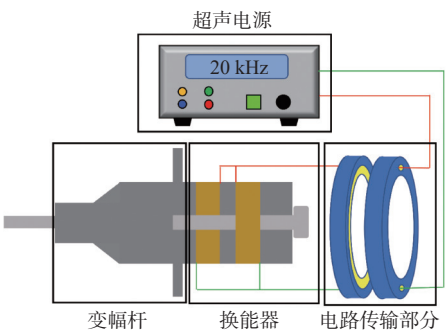


图 1　超声振动加工系统原理

Fig. 1　Principle of ultrasonic vibration machining system

1.1　双弯曲超声振动换能器结构设计

双弯曲超声椭圆振动加工系统采用 2 组压电陶瓷堆分别激励 2 个方向的弯曲超声振动, 2 组压电陶瓷耦合后形成双弯曲超声振动换能器。弯曲超声振动换能器的压电陶瓷结构采用半环型压电陶瓷, 具体结构如图 2 所示, 左右 2 个半环型压电陶瓷处在同一电场中, 但其之间具有间隙, 可使其在工作时互不干扰。另外, 在装配中保证左右 2 个半环型压电陶瓷的极化方向与交变电场的极化方向总是一个同向, 另一个反向, 当交变电场接通高频电压以后, 电场中左右 2 个半圆环形压电陶瓷便会呈现出一个伸长、另一个收缩的形式, 从而激励一个方向弯曲振动。双弯曲超声振动换能器存在上下 2 组弯曲压电陶瓷堆激励结构, 并且 2 组压电陶瓷的间隙呈现 90°分别激励 x 方向和 y 方向的弯振。2 组压电陶瓷堆中间利用钢隔开使其分别为最佳激励位置, 以减少相互间的影响。

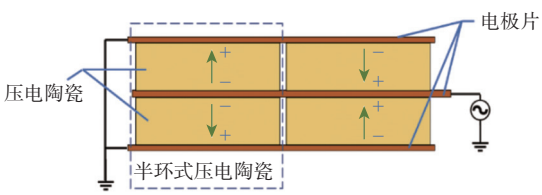


图 2　一组弯振压电陶瓷堆结构示意图

Fig. 2　Structural diagram of a set of piezoelectric ceramic stack

双弯曲旋转超声椭圆振动加工系统的弯曲换能器激励 x 和 y 方向上的弯振, 并在刀具末端进行了振动耦合形成椭圆振动的原理, 如图 3 所示。图 3(a)和图 3(b)为 y 方向的压电陶瓷组激励变幅杆在 y 方向呈现出弯振, 图 3(c)和图 3(d)为 x 方向的压电陶瓷组激励变幅杆在 x 方向呈现弯振, 最终在 xy 平面内刀具末端的振动轨迹呈现为椭圆形。

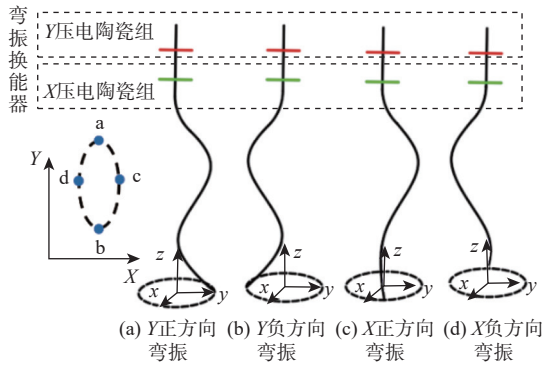


图 3　双弯曲旋转超声振动换能器椭圆振动工作原理

Fig. 3　Working principle of elliptical vibration of double-bending rotary ultrasonic vibration transducer

1.2　双弯曲旋转超声椭圆振动系统尺寸计算

Timoshenko 弯曲梁理论考虑了梁在弯曲过程中的剪切模量, 相较于 Euler-Bernoulli 理论, 该理论更适合推导弯曲波动方程, 并进行双弯曲旋转超声椭圆振动系统的整体尺寸设计。

等截面的弯曲自由振动微分方程[19]:

$$EI\frac{\partial^4 Y}{\partial x^4}+\rho A\frac{\partial^2 Y}{\partial t^2}-\rho I\left(1+\frac{E}{K'G}\right)\frac{\partial^4 Y}{\partial x^2\partial t^2}+\frac{\rho^2 I}{K'G}\cdot\frac{\partial^4 Y}{\partial t^4}=0 \tag{1}$$

式中: E 为材料的弹性模量; I 为材料的转动模量; Y 为位移; G 为材料的剪切模量; ρ 为材料的密度; K'为剪切系数, 一般圆形截面取 6/5; A 为横截面面积; l 为杆的长度。

考虑到该系统的放大倍数, 因此其变幅杆的设计采用阶梯型, 如图 4 所示。

该系统各部分的边界条件如图 4 所示, 其中, 第 1 部分和第 3 部分的边界条件相同。考虑到过大的直径过渡会产生较大的应力集中, 于是采用了过渡部分即第 2 部分。第 1 部分和第 3 部分由双弯曲旋转超声椭圆振动加工系统的边界条件可以计算得出。代入 Timoshenko 理论, 经 Timoshenko 弯

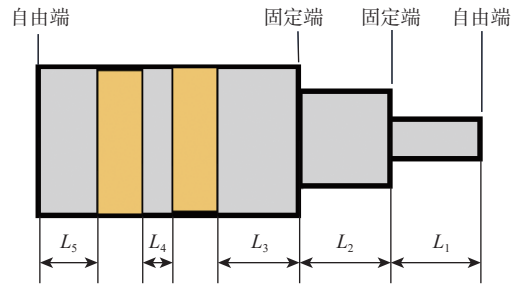


图4 双弯曲变幅杆结构

Fig. 4 Double bending horn structure

曲梁理论公式推导得出弯振的频率方程：

$$\tanh(n_1 l_{13}) = \frac{n_1 \phi_2}{|n_2| \phi_1} \tan(|n_2| l_{13}) \tag{2}$$

式中：

$$\begin{cases} \phi_1 = -\dfrac{\dfrac{\omega^2}{c_0^2} + n_1^2}{c_S^2 C} \\ \phi_2 = \dfrac{-\dfrac{\omega^2}{c_0^2} + |n_2|^2}{c_S^2 C} \\ n_1 = \omega M \left(-1 + N \left(1 + \dfrac{a^2}{\omega^2} \right)^{1/2} \right)^{1/2} \\ n_2 = \mathrm{j}\omega M \left(1 + N \left(1 + \dfrac{a^2}{\omega^2} \right)^{1/2} \right)^{1/2} \end{cases} \tag{3}$$

$$\begin{cases} C = \dfrac{A\rho}{EI} \\ a = \dfrac{2\sqrt{C}}{c_S^{-2} - c_0^{-2}} \\ N = \dfrac{c_0^2 - c_S^2}{c_0^2 + c_S^2} \\ M = \dfrac{1}{\sqrt{2}} \left(\dfrac{1}{c_S^2} + \dfrac{1}{c_0^2} \right)^{1/2} \\ c_0 = \left(\dfrac{E}{\rho} \right)^{1/2} \\ c_S = \left(\dfrac{K'G}{\rho} \right)^{1/2} \\ G = \dfrac{E}{2(1+v)} \end{cases} \tag{4}$$

其中：ω为角频率；v为泊松比。为增大系统的整体振幅，压电陶瓷采用PZT-8，其为外径为50 mm，内孔直径为17 mm，厚度为6.5 mm，且中间存在间隙的半环型压电陶瓷，由于刀具末端采用ER20的弹簧夹头进行夹持，所以变幅杆末端直径设计为24 mm。双弯曲旋转超声椭圆振动加工系统的变幅杆采用的材料为45#钢，其材料属性如表1所示，若该系统的设计频率为20 kHz，则由式(1)推导出L_1、L_2和L_3。将换能器部分2个压电陶瓷堆分别放置在波腹位置，每个压电陶瓷堆厚度为13 mm，可由式(1)得出L_4和L_5。

表1 材料属性

Table 1 Material properties

材料	密度/(kg·m^{-3})	杨氏模量/GPa	泊松比
45#钢	7 850	208	0.28
PZT-8	7 650	90	0.3

2 双弯曲旋转超声振动加工系统模态与谐响应分析

2.1 构建有限元模型

由于2个弯曲振动在弹性体内的传播速度相同，并且在同一频率下会同时出现2个方向的弯曲振动，因此，考虑将同一频率下的2个弯振进行耦合，但是耦合后的节面呈点状，传统解析法难以推算出具体的法兰位置，双弯曲旋转超声椭圆振动加工系统基于有限元分析方法分析法兰结构对其性能的影响，法兰的引入不会影响其性能。

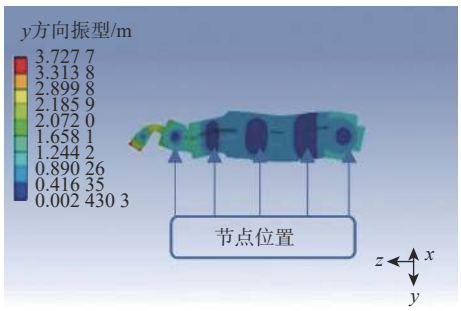
依据理论推导，建立双弯曲旋转超声椭圆振动加工系统参数模型，并对其进行网格划分，网格的节点数量为34 556个，足以满足仿真的精度要求。将网格模型导入有限元分析软件workbench中，赋予各个零部件的材料属性如表2所示，对整体模型进行模态分析，模态分析结果云图如图5所示，第5阶2个方向弯振频率分别为22 127 Hz和22 049 Hz，误差仅为0.35%，因此可以忽略不记。验证了理论推导的正确性，并且2个方向上的弯曲频率相差较小，可满足设计要求。

表2 超声振动系统各零件有限元分析材料属性

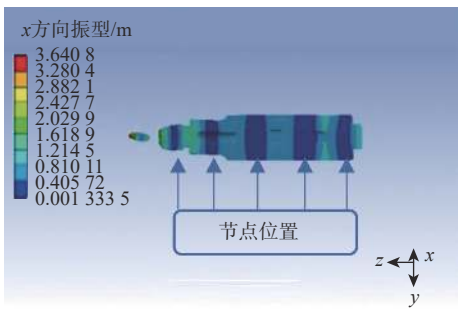
Table 2 Material properties of ultrasonic vibration system by finite element analysis

零件	材料	密度/(kg·m^3)	弹性模量/Pa	泊松比
后盖板	45#钢	7 800	2.08×10^{11}	0.28
前盖板	45#钢	7 800	2.08×10^{11}	0.28
压电陶瓷	PZT-8	7 800	8.2×10^{10}	0.29
预紧螺栓	40cr	7 900	2.14×10^{11}	0.3
刀具	硬质合金YG8	14 000	1.9×10^{11}	0.3

针对本文系统的同轴度问题，以及方便2个方向的弯曲超声振动可以较好的耦合，本文选用圆环型法兰进行设计。考虑到有限元模型与实际的工况的相符程度及刀具对系统整体的影响，刀具采用硬质合金YG8，再划分双弯曲旋转超声椭圆振动加工系统网格，划分结果如图6所示，节点数量为32 674个，网格数量为15 376个，网格数量满足仿真的精度要求。将网格模型导入有限元分析软件workbench中，阻尼设为0.02，并对双弯曲旋转超声椭圆振动加工系统进行模态分析及谐响应分析。各个部分的材料赋予如表2所示。



(a) *y*方向振型, 频率为 22 127 Hz



(b) *x*方向振型, 频率为 22 049 Hz

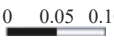


图 5　无法兰超声系统模态云图

Fig. 5　Modal cloud atlas of ultrasonic system with unmovable flange

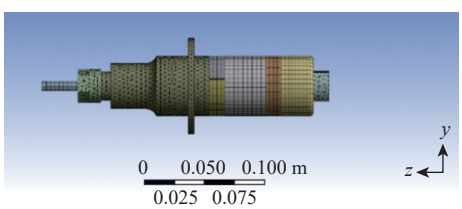


图 6　双弯曲旋转超声椭圆振动加工系统网格

Fig. 6　Double-bending rotary ultrasonic elliptical vibration machining system grid

2.2　仿真结果对比

2.2.1　法兰位置对频率的影响

双弯曲超声椭圆振动加工系统首先应参考无法兰状态下的模态云图(图 5)。考虑到实际加工需求将法兰初步定位到变幅杆的右边第 1 个节点位置, 粗略给出在距离变幅杆大端面 14 ～16 mm 的位置, 然后逐个分析法兰加入后对整体的影响进而确定其具体位置。

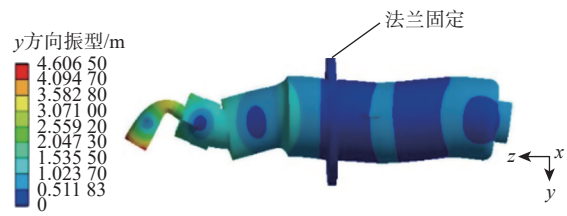
由有限元仿真结果分析不同位置的法兰引入后对 *x* 和 *y* 方向弯曲振动频率的影响, 仿真结果表明, 法兰插入后对 *x* 和 *y* 方向弯曲振动频率偏差影响都较小, 频率偏差均在 104 Hz 左右。这说明圆环型的法兰插入对整体 *x* 和 *y* 方向的频率得偏差方面的影响可以忽略。

由表 3 对比法兰插入后固定法兰与不固定法兰频率变化, 可以看出, 法兰固定以后频率变化较大, 主要原因是弯曲振动的节点呈点状, 如图 5 所示。圆环状的法兰导致了其不可能与节点完全重合, 所以固定法兰结构后双弯曲旋转超声椭圆振动加工系统的频率会发生明显的偏移, 这又导致法兰部分对超声能量的消耗增加。由有限元分析得到法兰距离变幅杆大端面为 15 mm 时其频率偏移量最小, 此时法兰与其节点位置重合较好, 如图 7 所示。

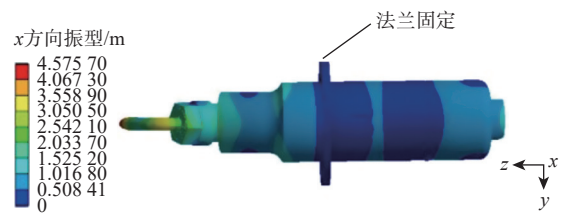
表 3　双弯曲旋转超声椭圆振动加工系统弯曲频率

Table 3　Bending frequency of double bending rotary ultrasonic elliptical vibration machining system

法兰距离/mm	*x*方向自由频率/Hz	*y*方向自由频率/Hz	*x*方向固定频率/Hz	*y*方向固定频率/Hz
14	22 410	22 414	23 480	23 528
14.5	22 301	22 405	23 619	23 569
15	22 596	22 693	23 535	23 574
15.5	22 281	22 385	23 761	23 812
16	22 269	22 375	23 856	23 914



(a) *y*方向振型, 频率为 23 574 Hz



(b) *x*方向振型, 频率为 23 535 Hz

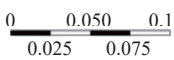


图 7　双弯曲旋转超声椭圆振动加工系统振型云图

Fig. 7　Modal cloud atlas of dual-bending rotary ultrasonic elliptical vibration machining system

2.2.2　法兰位置对振幅的影响

如图 8 所示, *x* 和 *y* 这 2 个方向的超声振动幅值随着距离的增加, 总体都呈先增大后减小的形式, 并且在法兰距离变幅杆大端面为 15 mm 时, 其振幅值最大, *y* 方向最大值为 8.8 μm, *x* 方向最大值为 6.4 μm。上述现象说明: 法兰距离变幅杆大端面 15 mm 的位置时, 对超声振动系统的振动振幅性能影响较小, 且此时法兰所在平面与超声振动系统的截面位置重合度较高。

法兰距离变幅杆大端面 15 mm 时, 谐响应分析结果如图 9 所示, 当谐振频率为 23 535 Hz 时, *x* 和 *y* 方向振幅分别为 8.87 μm 和 6.48 μm, 由仿真结果可以看出设计的可行性, 法兰位置也满足设计需求。

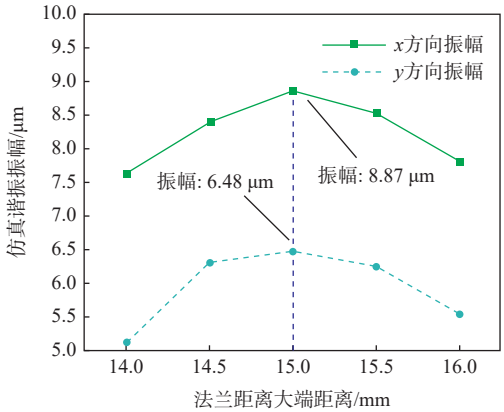


图 8 法兰位置与振幅关系

Fig. 8 Relationship between flange position and amplitude

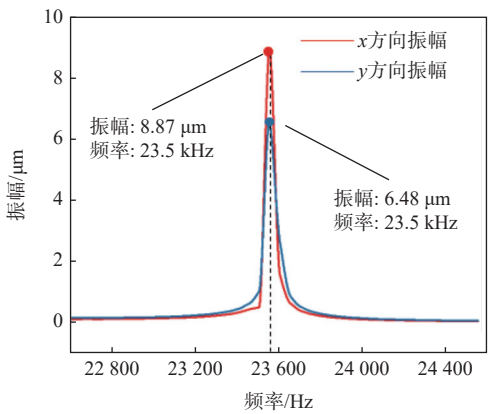


图 9 双弯曲旋转超声椭圆振动加工系统弯曲振幅

Fig. 9 Bending amplitude of double bending rotary ultrasonic elliptical vibration machining system

此外，本文还设计了双弯曲旋转超声椭圆振动加工系统的整体尺寸并确定了法兰的具体位置，综合考虑双弯曲旋转超声椭圆振动加工系统加工过程中的影响因素，在法兰上采用锥面定位，定位时主要考虑双弯曲旋转超声椭圆振动加工系统的同轴度问题，由仿真结果可知：增加 1∶10 的锥面，并不会影响双弯曲旋转超声椭圆振动加工系统整体频率的变化，振幅在仿真时会略微下降 0.2 μm 左右。超声振动系统振幅下降的主要原因是法兰的平面具有一定厚度，而双弯曲旋转超声椭圆振动加工系统的节点在理论上是分散的点状，因此，二者始终无法重合，从而导致小部分能量在法兰部分被消耗掉了。

3 系统性能测试

3.1 阻抗分析

实验设备采用阻抗分析仪对夹持直径为 6 mm 的刀具的双弯曲旋转超声椭圆振动加工系统进行测量，测量结果如图 10 所示，图中，lg$Z$ 为阻抗的对数函数，其中，Z 为阻抗。由图可知，该系统谐振频率为 24 245 Hz，与有限元仿真值 22 127 Hz 和 22 049 Hz，存在一定的误差，误差值为 5%，误差产生的原因可能是建模与实际存在偏差。在仿真中各个接触面为理想的品面，而实际情况比较复杂，并且在仿真过程中，没有考虑到螺栓预紧力对系统整体的影响。机械品质因数反映了机械能的转换效率，一般要求在 300～800[22]，双弯曲旋转超声椭圆振动加工系统的机械品质因数 Q_m=552.7，符合要求。另外，一般动态电阻要求越小越好，但对于半圆环压电陶瓷，其装配工艺的限制导致其动态电阻值比单一纵振换能器大一些，一般为 150 Ω 以下，该系统的动态电阻 R_1=126 Ω，符合设计要求。由图 10 可知，阻抗圆的形状较好，这也表明其阻抗性能较好。

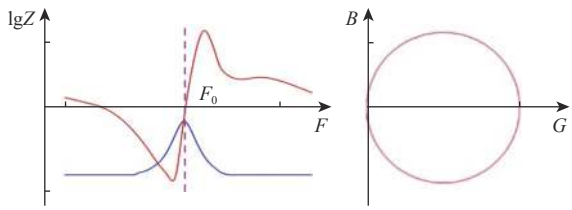
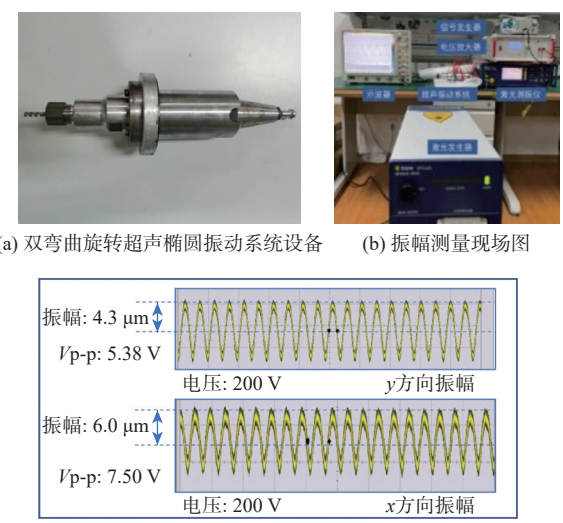


图 10 阻抗分析

Fig. 10 Impedance analysis

3.2 振幅测量

超声振幅是双弯曲旋转超声椭圆振动加工系统的重要评判标准，一般采用激光测振仪进行测量，因激光测振仪采用非接触式的测量，所以结果较为精确。测量该系统的振幅结果如图 11 所示，其中，V 为峰值电压。目前已知的振动形式主要有纵振、扭振、弯振及复合振动，当刀具发生纵振时，z 方向振动较为明显；当刀具发生扭振时 x 和 y 方向振动较为明显，并且 x 和 y 方向振幅相同；当刀具



(a) 双弯曲旋转超声椭圆振动系统设备 (b) 振幅测量现场图

(c) 振幅测量结果

图 11 x 和 y 方向的振动幅值测量及结果

Fig. 11 Vibration amplitude measurement and results in different directions

发生弯振时, 弯振方向上的振幅较为明显, x 和 y 方向上的振幅不同。如图 12 所示, 刀具在 x 和 y 方向上的振幅较为明显, 且 2 个方向的振幅并不相同, 而在 z 方向上的振幅比 x 和 y 方向的振幅小, 所以刀具发生了 x 和 y 方向的弯曲振动, 并在可以 xy 平面上形成椭圆振动。

测量不同电压下双弯曲旋转超声椭圆振动加工系统 2 个方向上的振幅, 结果如图 12 所示。由图 12 可知, 随着电压的增大, 2 个方向的超声振幅均增加, 但是二者增加的速率不同, x 方向上的振幅增加更为明显, 所以为获得该系统的大振幅, 可以适当增加电压。将电压提升至 200 V, 其振幅如图 11 所示, 可以看出, 刀具在 2 个方向输出稳定的简谐振动信号, 2 个方向的振幅分别为 4.3 μm 和 6.0 μm。

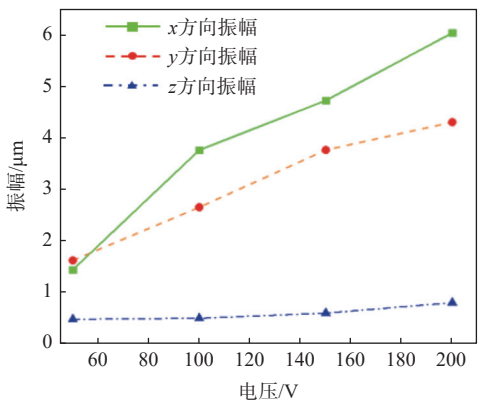


图 12　不同方向的电压与振幅关系

Fig. 12　Relationship between voltage and amplitude in different directions

4　结　论

本文通过 2 个弯曲换能器激励不同方向上的振动进行耦合, 并设计了一种双弯曲旋转超声椭圆振动系统, 结合理论设计和系统性能测量结果可得到如下结论:

1) 双弯曲旋转超声椭圆振动加工系统的实际工作频率为 24 245 Hz, 与理论设计值偏差在 5% 以内, 这验证了本文系统的可行性与有效性;

2) 振动幅值测量结果验证了刀具振动轨迹是椭圆振动, 200 V 电压下 x 和 y 方向上的振幅分别为 6.0 μm 和 4.3 μm;

3) 双弯曲旋转超声椭圆振动加工系统的法兰位置在距离变幅杆大端 15 mm 处频率偏差较小, 2 个方向耦合振幅最大, 由此证明, 本文系统具有较好的适用性。

参考文献 (References)

[1] 郑侃, 廖文和, 孙连军, 等. 机器人纵振与纵扭超声铣削稳定性对比研究[J]. 机械工程学报, 2021, 57(7): 10-17.
ZHENG K, LIAO W H, SUN L J, et al. Comparative study on stability of robotic longitudinal vibration and longitudinal-torsional ultrasonic milling[J]. Journal of Mechanical Engineering, 2021, 57(7): 10-17(in Chinese).

[2] 廖文和, 郑侃, 孙连军, 等. 大型复杂构件机器人加工稳定性研究进展[J]. 航空学报, 2022, 43(1): 026061.
LIAO W H, ZHENG K, DONG S, et al. Burr characteristics of robotic rotary ultrasonic drilling aluminum alloy stacked components[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(1): 026061(in Chinese).

[3] 胡力闯, 郑侃, 董松, 等. 机器人旋转超声钻削铝合金叠层构件毛刺特性[J]. 北京航空航天大学学报, 2020, 46(2): 164-183.
HU L C, ZHEN K, DONG S, et al. Burr characteristics of robotic rotary ultrasonic drilling aluminum alloy stacked components[J]. Journal of Beijing University of Aeronautics and Astronantics, 2020, 46(2): 164-183(in Chinese).

[4] YANG Z C, ZHU L D, ZHANG G X, et al. Review of ultrasonic vibration-assisted machining in advanced materials[J]. International Journal of Machine Tools and Manufacture, 2020, 156: 103594.

[5] SHU J, LIAO W H, ZHENG K, et al. Effect of rotary ultrasonic machining on surface structure and ablation resistance of C/C composite[J]. Ceramics International, 2022, 48(13): 18246-18256.

[6] XUE F, ZHENG K, LIAO W H, et al. Investigation on fiber fracture mechanism of c/sic composites by rotary ultrasonic milling[J]. International Journal of Mechanical Sciences, 2021, 191: 106054.

[7] ZHANG M L, ZHANG D Y, GENG D X, et al. Effects of tool vibration on surface integrity in rotary ultrasonic elliptical end milling of Ti–6Al–4V[J]. Journal of Alloys and Compounds, 2020, 821: 153266.

[8] 张德远, 彭振龙, 耿大喜, 等. 高效低损伤高速波动式超声加工技术进展[J]. 航空制造技术, 2022, 65(8): 22-33.
ZHANG D Y, PENG Z L, GENG D X, et al. Development of high-speed wave-motion ultrasonic machining technology with high efficiency and low damage[J]. Aeronautical Manufacturing Technology, 2022, 65(8): 22-33(in Chinese).

[9] 彭振龙, 张翔宇, 张德远. 航空航天难加工材料高速超声波动式切削方法[J]. 航空学报, 2022, 43(4): 525587.
PENG Z L, ZHANG X Y, ZHANG D Y. High-speed ultrasonic vibration cutting for difficult-to-machine materials in aerospace field[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(4): 525587(in Chinese).

[10] 高泽, 张德远, 李哲, 等. 高速超声椭圆振动铣削腹板表面质量研究[J]. 机械工程学报, 2019, 55(7): 249-256.
GAO Z, ZHANG D Y, LI Z, et al. Research on surface quality of titanium alloy webs via high-speed ultrasonic elliptical vibration milling[J]. Journal of Mechanical Engineering, 2019, 55(7): 249-256(in Chinese).

[11] 刘佳佳, 姜兴刚, 高泽, 等. 高速旋转超声椭圆振动侧铣削振幅对钛合金表面完整性影响的研究[J]. 机械工程学报, 2019, 55(11): 215-223.
LIU J J, JIANG X G, GAO Z, et al. Investigation of the effect of vibration amplitude on the surface integrity in high-speed rotary ultrasonic elliptical machining for side milling of Ti-6Al-4V[J]. Journal of Mechanical Engineering, 2019, 55(11): 215-223(in Chinese).

[12] GAO J, ALTINTAS Y. Chatter stability of synchronized elliptical vibration assisted milling[J]. CIRP Journal of Manufacturing Science and Technology, 2020, 28: 76-86.

[13] LU H, ZHU L D, YANG Z C, et al. Research on the generation mechanism and interference of surface texture in ultrasonic vibration assisted milling[J]. International Journal of Mechanical Sciences, 2021, 208: 106681.

[14] WANG H, PEI Z J, CONG W L. A feeding-directional cutting force model for end surface grinding of CFRP composites using rotary ultrasonic machining with elliptical ultrasonic vibration[J]. International Journal of Machine Tools and Manufacture, 2020, 152: 103540.

[15] PANG Y, FENG P F, ZHANG J F, et al. Frequency coupling design of ultrasonic horn with spiral slots and performance analysis of longitudinal-torsional machining characteristics[J]. The International Journal of Advanced Manufacturing Technology, 2020, 106(9): 4093-4103.

[16] CHEN P, TONG J L, ZHAO J S, et al. A study ofthe surface microstructure and tool wear of titanium alloys after ultrasonic longitudinal-torsional milling[J]. Journal of Manufacturing Processes, 2020, 53: 1-11.

[17] 顾磊, 郑侃, 董松. 纵扭共振旋转超声加工系统设计[J]. 南京理工大学学报, 2020, 44(2): 127-133.
GU L, ZHENG K, DONG S. Design of miniature longitudinal-torsional ultrasonic machining system[J]. Journal of Nanjing University of Science and Technology, 2020, 44(2): 127-133(in Chinese).

[18] WANG J J, LIAO W H, GUO P. Modulated ultrasonic elliptical vibration cutting for ductile-regime texturing of brittle materials with 2-D combined resonant and non-resonant vibrations[J]. International Journal of Mechanical Sciences, 2020, 170: 105347.

[19] BAI W, WANG K, DU D X, et al. Design of an ultrasonic elliptical vibration device with two stationary points for ultra-precision cutting[J]. Ultrasonics, 2022, 120: 106662.

[20] DU P F, HAN L, QIU X, et al. Development of a high-precision piezoelectric ultrasonic milling tool using longitudinal-bending hybrid transducer[J]. International Journal of Mechanical Sciences, 2022, 222: 107239.

[21] 王晓雪, 张臣, 王生才. 纵弯椭圆振动辅助钻削装置的设计与优化研究[J]. 航空制造技术, 2020, 63(12): 70-79.
WANG X X, ZHANG C, WANG S C. Design and optimization of longitude-bend complex mode elliptical vibration assisted drilling device[J]. Aeronautical Manufacturing Technology, 2020, 63(12): 70-79(in Chinese).

[22] 张若愚. 用于旋转超声加工的半波谐振超声振子设计及实验研究[D]. 杭州: 浙江大学, 2020: 50-51.
ZHANG R Y. Design and experimental study of half-wave resonant ultrasonic oscillator for rotary ultrasonic machining[D]. Hangzhou: Zhejiang University, 2020: 50-51(in Chinese).

Design of double bending rotary ultrasonic elliptical vibration machining system

SHAO Zhuangzhuang, ZHENG Kan*, DONG Song, SUN Zhenwen, SUN Lianjun, Yan Jiawen

(College of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: The processing efficiency of the rotary ultrasonic single vibration mode is limited. Therefore, an elliptical vibration processing system with double-bending ultrasonic vibrations was designed using the ultrasonic vibration excitation source coupling method. The bending wave equation was derived, and the calculated parameters of the bending ultrasonic vibration system showed only a 5% error compared to the simulation results in theoretical analysis. The coupled double-bending rotary ultrasonic elliptical vibration processing system combines vibrations in two different directions, controlling the bending vibration frequency deviation to be less than 1%. Finite element analysis was performed on the flange position and shape of the ultrasonic tool holder, revealing that the inclusion of damping grooves helps increase the bending vibration amplitude and determines the optimal position of the flange. Finally, the vibration frequency and amplitude of the double-bending rotary ultrasonic elliptical vibration processing system were tested, and the results indicated that the operating frequency and amplitudes in both directions met the design expectations.

Keywords: double-bend ultrasonic elliptical vibration; excitation source coupling; ultrasonic shank; bending vibration; bending vibration amplitude

Received: 2022-08-13; **Accepted**: 2022-10-07; **Published Online**: 2023-01-06 08:58

URL: link.cnki.net/urlid/11.2625.V.20230104.2053.007

Foundation item: National Natural Science Foundation of China (52075265)

* **Corresponding author**. E-mail: zhengkan@njust.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0713

柱塞泵配流盘卸荷槽开度对瞬态流场特征影响

刘珊珊[1]，雷欣瑞[2]，宋宇宇[1]，杨凯[1]，黄家海[1,*]

(1. 太原理工大学 机械与运载工程学院，太原 030024；2. 太重集团榆次液压工业有限公司，晋中 030600)

摘 要：空化是影响轴向柱塞泵性能的关键因素，柱塞缸体表面损坏是一种常见的失效形式，但目前成因尚不清晰。建立柱塞泵三维流场仿真模型，设计可观察单个柱塞腔与卸荷槽导通状态下流场特征的实验装置，并采用流体仿真和流场可视化测量的方法揭示了其成因。研究结果表明：射流角随卸荷槽开度的增大而增大，并伴随漩涡的产生与消失，射流中心速度大于20 m/s时漩涡产生；柱塞腔与卸荷槽由导通至完全截止的转变过程中，射流角由0°逐渐增大至60°，当射流角由0°向20°变化时，射流方向主要集中于柱塞缸体与配流盘腰型槽接触的表面处，导致柱塞缸体表面和配流盘吸油腰型槽上部产生射流破坏和气蚀破坏；当射流角由20°向60°变化时，射流方向将向配流盘腰型槽内部扩展，导致配流盘破坏区域由上部移向内部。上述研究结果有助于厘清柱塞缸体摩擦表面气蚀成因，为配流盘结构优化设计提供了约束条件，对提升柱塞泵的综合性能也有重要意义。

关 键 词：轴向柱塞泵；空化；配流盘；缸体；卸荷槽；气蚀

中图分类号：TH137.5

文献标志码：A **文章编号**：1001-5965(2024)09-2919-11

随着液压技术的发展，对轴向柱塞泵压力、转速和排量提出了更高的要求，但其在高转速工况下容易出现气穴现象[1]，不但降低了容积效率[2]、诱发振动与噪声[3]，甚至会引起气蚀损坏[4]。液压元件气蚀分为蒸汽气蚀和空蚀[5]，前者是流体局部压力低于蒸汽压力而引起；后者是由于当压力降至饱和水平以下，溶解于油液中的气体析出，被输送到高压区时内爆，导致材料腐蚀、振颤和噪声[6]。为减小配流冲击、降低噪声，柱塞泵在吸、排油转换区域必须进行预升压和预卸压，由于减振三角槽两端存在较大压差，因此当压力低于油液空气分离压时，会形成空化[7-8]。

国内外学者在液压系统空化和空蚀方面取得了较大进展[9-12]，实验和仿真成为空化研究的主要手段。利用激光在静水中产生单气泡，采用高速摄像机记录气泡的瞬态变化过程，成为普遍的实验研究方法[13-16]。Zhang等[17]通过改变泵速或进口压力来抑制空化。当泵高速运行时，需要足够的进口压力以避免气穴[16]。郭伟伟[18]采用传统理论分析和数值模拟相结合的方法，研究轴向柱塞泵内部流场分布与空化，解释了淹没式射流造成的金属剥蚀现象。築地徹浩等[19]将CFD仿真和可视化试验结合，分析并预测空化现象的产生和产生区域。文献[20-21]使用高速摄像机对实际轴向柱塞泵中的气穴进行了数据采集，观察到空化的产生及脱落过程。文献[22-23]通过增大减振三角槽的射流角或在槽中加工阻尼孔改变气穴发生位置，减小了气蚀破坏程度。

综上所述，研究人员已对柱塞泵配流盘表面预卸荷槽和预升压槽内流场特征和空化现象进行了较为详细的仿真模拟，但受柱塞泵机械结构的限制，无法对柱塞泵内部流场进行实验观察，仅有少数研究搭建了简易模拟装置[19-20]进行实验。

柱塞泵吸、排油过程诱发的气蚀现象主要集中

收稿日期：2022-08-13；录用日期：2022-11-09；网络出版时间：2022-11-24 09:48
网络出版地址：link.cnki.net/urlid/11.2625.V.20221122.1343.002
基金项目：国家重点研发计划 (2021YFB2011903)
*通信作者. E-mail：huangjiahai@tyut.edu.cn

引用格式：刘珊珊，雷欣瑞，宋宇宇，等. 柱塞泵配流盘卸荷槽开度对瞬态流场特征影响[J]. 北京航空航天大学学报，2024，50(9)：2919-2929.
LIU S S，LEI X R，SONG Z Y，et al. Influence of unloading groove opening of port plate of plunger pumps on transient flow field characteristics[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2919-2929 (in Chinese).

在柱塞腔内壁，以及配流盘吸、排油转变区域和柱塞缸体摩擦表面。研究人员对柱塞腔内壁和配流盘吸、排油转变区域的气蚀成因做了大量研究，并提出了解决方案，但针对柱塞缸体摩擦表面气蚀成因的研究较少。由于缸体摩擦表面气蚀会导致摩擦副外密封带失效，使柱塞泵容积效率降低，甚至可造成柱塞泵功能失效，因此，开展柱塞缸体摩擦表面气蚀发生机制的研究工作具有实际工程意义。

本文采用流场仿真和可视化实验等技术手段开展研究，分析卸荷槽开度与射流角演变、漩涡产生与消失、空化发生等关联性，揭示了卸荷槽开度与柱塞缸体表面气蚀的内在关系；基于柱塞泵吸、排油周期的变化特点，研制出可模拟单个柱塞腔与卸荷槽导通状态下流场特征的装置，并验证了仿真结果；继而利用理论研究成果对柱塞缸体摩擦表面气蚀的形成原因进行解释与分析。

1 仿真建模与试验系统简介

1.1 仿真建模

1.1.1 物理模型

图1为轴向柱塞泵典型结构。当柱塞处于吸油和排油的过渡阶段时，配流盘阻尼槽附近的负压梯度容易引发高速射流，导致局部压力低于饱和蒸气压，形成气穴[5,24]、产生噪声和振动、甚至是气蚀破坏，从而影响柱塞泵的可靠性和寿命。图2是排量为250 mL/rad的柱塞泵球面配流盘示意，在吸油和排油的过渡区域分别加工了预升压槽和预卸压槽，实现了高低压的平稳过渡。本文将着重分析柱塞腔通过卸荷槽时的流场特征演变情况。

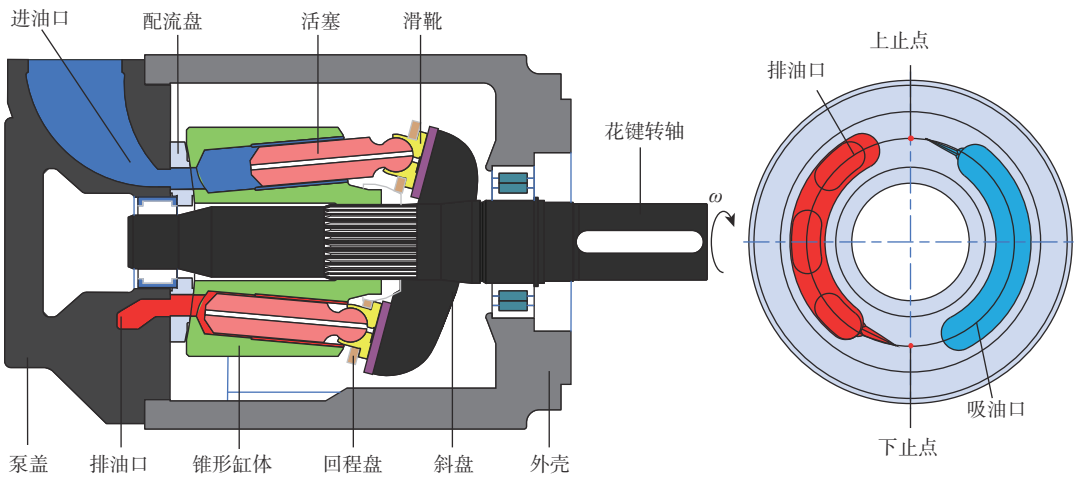


图1 斜盘式轴向柱塞泵示意

Fig. 1 Swash plate axial plunger pump

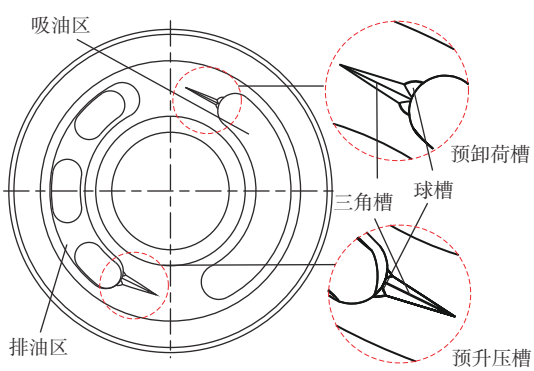


图2 配流盘结构示意

Fig. 2 Port plate structure

1.1.2 流体域网格划分

高质量的网格是精确计算的前提。为提高网格的精度和质量，对不同流体域进行网格单独划分，如图3所示，对配流副球形表面网格、预升压槽和预卸荷槽的网格进行了精化处理。在仿真软件中设置多组网格参数进行预仿真，以进、出口流量误差值为依据对网格敏感性进行评判。理论上，仿真结果应当符合质量守恒原则，当工作介质为不可压缩或弱可压缩流体时，仿真模型的流入、流出流量是近似相等的，若存在较大误差，则说明网格数量较少或质量较差。自定义流量误差ΔQ如下：

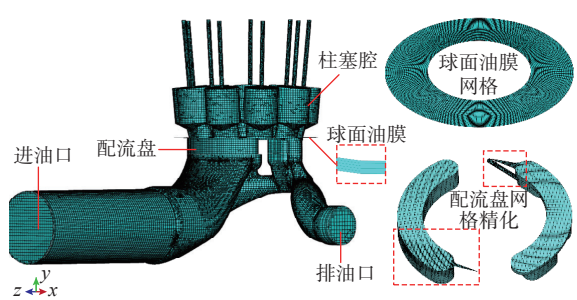


图3 计算域网格

Fig. 3 Computational domain grid

$$\Delta Q=\frac{|Q_{\text{in}}-Q_{\text{out}}|}{Q_{\text{out}}} \tag{1}$$

式中: Q_{in} 为流入流量; Q_{out} 为流出流量。

表 1 为 7 组不同网格数量下的计算结果, 流量误差随网格数量的增加而降低, 但当网格数量增加到一定值后, 流量误差的降低不再明显。经过权衡计算时间和计算精确度的利弊, 选择第 5 组网格划分数量进行仿真计算。

表 1　网格敏感性

Table 1　Grid sensitivity

组别	关键边界角度	曲率分辨率	最大单元格	边界单元格尺寸	网格数目	ΔQ/%
1	30	35	0.03	0.015	188 105	7.13
2	30	35	0.025	0.012 5	214 750	4.76
3	30	35	0.02	0.01	225 204	2.26
4	25	30	0.02	0.01	232 352	1.42
5	15	20	0.02	0.01	367 691	1.12
6	15	20	0.02	0.005	518 554	0.97
7	15	20	0.01	0.005	897 371	0.98

1.1.3　边界条件与模型参数

液体在圆管中无量纲雷诺数 Re 为

$$Re=\frac{vd\rho}{\mu}=\frac{vd}{\upsilon} \tag{2}$$

式中: v 为流速; d 为管径; ρ 为流体密度; μ 为流体黏度; υ 为运动黏度。

流体在非圆管道中无量纲雷诺数 Re 为

$$Re=\frac{4vR_{\text{H}}}{\upsilon} \tag{3}$$

式中: R_{H} 为流通界面水力半径。

模型选用 46 号液压油, 根据 46GB/T265 标准[25], 其在 40 ℃ 时的运动黏度为 46 mm^2/s, 泵输入最大转速为 1 500 r/min, 斜盘倾角为 15°时吸油口的最大流速为 1.415 m/s, 根据文献 [18] 可求出柱塞泵各部位雷诺数, 如表 2 所示, 已知临界雷诺数为 2 000～2 320。

表 2　柱塞泵内流态情况

Table 2　Flow pattern in plunger pump

位置	雷诺数	流态
柱塞泵吸油口	$Re=\frac{vd_{\text{in}}}{\upsilon}=\frac{1.415\times10^{-3}\times75}{46\times10^{-6}}\approx2\,307<2\,320$	层流
柱塞泵排油口	$Re=\frac{vd_{\text{out}}}{\upsilon}=\frac{3.4\times10^{-3}\times40}{46\times10^{-6}}\approx2\,957>2\,320$	湍流
三角槽处	$Re=\frac{4vR_{\text{H}}}{\upsilon}=\frac{4\times186\times10^{-3}\times0.481}{46\times10^{-6}}\approx7\,780>2\,320$	湍流
柱塞腔内	$Re=\frac{vd_{\text{pis}}}{\upsilon}=\frac{10\times10^{-3}\times31}{46\times10^{-6}}\approx6\,739>2\,320$	湍流

由表 2 可知, 泵内部流动处于湍流状态; 为减小吸油口边界和排油口边界对柱塞泵内部流场的影响, 在图 3 所示的仿真模型中分别增加了吸油管道和排油管道的长度, 管道长度均为管径的 10 倍。为较真实地模拟流场湍流和空化特征, 仿真模型选用半隐式压力耦合方程算法和雷诺平均 Navier-Stokes 方程控制湍流; 考虑到近壁面处空化演变情况, 选用 RNG k-ε 湍流模型及全空化模型来进行描述[8]。

由于采用独立网格划分法, 所以需要在不同流体域之间创建交互面, 以实现不同流体域之间的数据传输。除进、出口边界和交互面外, 其他边界设置为实体壁面。采用线性求解器, 收敛容差为 0.05, 松弛因子为 0.5。仿真模型其他参数如表 3 所示。

表 3　仿真参数的设置

Table 3　Settings of simulation parameters

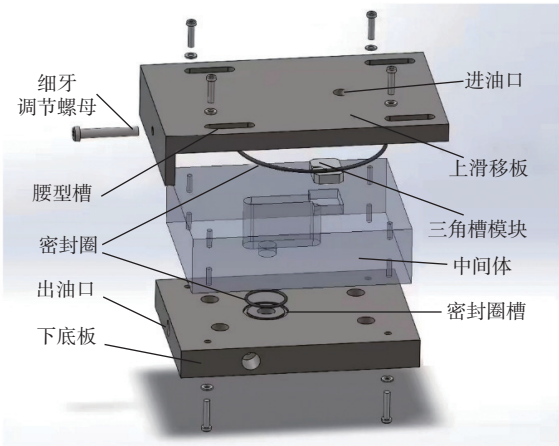
斜盘倾角/(°)	工作压力/MPa	油膜泄漏压力/MPa	球面副曲率半径/mm	液压油密度/(kg·m^{-3})	空气分离压/MPa	饱和蒸汽压/MPa	动力黏度/(Pa·s)	额定转速/(r·min^{-1})	体积弹性模量/MPa	入口压力/MPa	饱和压力/MPa	空化蒸汽系数	冷凝系数	工作温度/K
15	35	2	396±0.3	865	0.004	0.003 5	0.039 79	1 500	1.5×10^3	0.1	0.04	0.02	0.01	313

1.2　流场可视化观测的实验装置

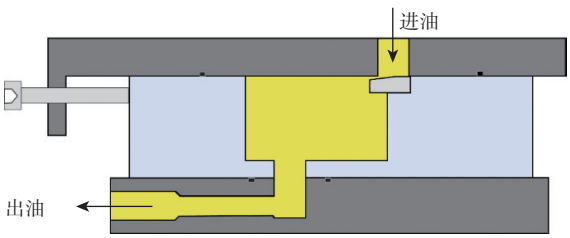
轴向柱塞泵中的 9 个柱塞腔是均匀布置的, 任意单个柱塞腔与卸荷槽构成的通流面积均是周期变化的, 即轴向柱塞泵旋转一周时, 柱塞腔与卸荷槽构成的过流面积呈 9 个周期重复变化, 所以流场特征也将呈周期性变化。从理论上说, 单个柱塞腔与卸荷槽导通状态下的流场特征具有一定的代表性, 能够从一定程度上体现真实柱塞泵卸荷槽的流场特征。基于上述原则, 本文提出如图 4 所示的实验装置, 用于模拟单个柱塞腔与卸荷槽导通状态下的流场特征。该装置主要包括上滑移板、下底板、中间体和三角槽模块, 上滑移板为 L 型, 具有 4 个腰型槽通孔和进油孔, L 型板一侧安装了细牙调节螺母, 使上滑移板可在中间体表面移动, 以调节进油孔和三角槽相对位置。需要注意的是, 当上板腰型槽左端与螺钉接触时, 进油孔将恰好与卸荷槽接触, 但尚未导通。

中间体由透明亚克力材料制成, 内部截面如图 4(b) 所示, 内部流道包括进油口、出油口、卸荷槽和观测区等, 且与真实 250 柱塞泵的相关零部件尺寸一致。由于亚克力材料坚硬且脆, 无法加工卸荷槽, 所以将加工好的卸荷槽模块嵌入亚克力槽内, 这种分体设计不但解决了卸荷槽难以加工的问题, 而且还可以更换卸荷槽模块, 以便于对不同结构形状的卸荷槽进行测试。

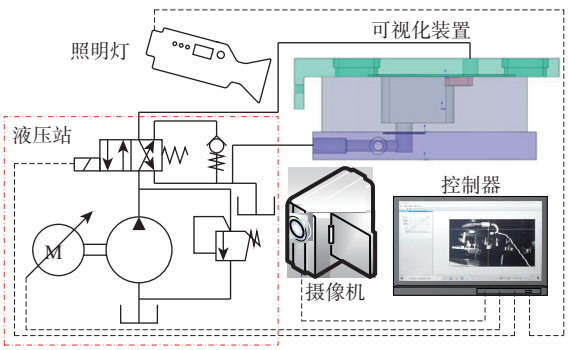
中间体通过螺钉和下底板联接, 且两者之间无



(a) 三维结构



(b) 内部流道剖面



(c) 实验原理

图 4 可视化观测的实验装置

Fig. 4 Experimental device for visual observation

相对运动;上滑动顶板通过细牙调节螺母可在中间体表面滑动,模拟真实柱塞泵内任意柱塞腔在卸荷槽表面滑移运动。本文为简化实验,仅测试了若干固定位置下的流场流动情况。

图 4(c) 为实验原理,包括可视化装置、数据采集装置及辅助装置。使用千眼狼 5KF20 高速摄像机,图像采集速度为 120 000 帧/s、分辨率为 1 920 像素×1 080 像素、像元尺寸为 14 μm×14 μm、动态范围为 60 dB、精度为 10 bit、曝光时间为 1 us,配备 1 kW 金属卤化物灯。高倍聚光镜将光线汇聚到可视化窗口处,调节镜头光圈和曝光时间使图像更加清晰。

2 结果与分析

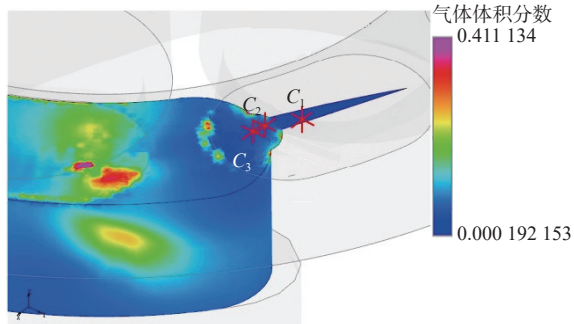
2.1 卸荷槽内压力脉动

图 5 为卸荷槽的压力脉动,其中,C_1、C_2 和 C_3 分别位于三角槽、球槽和卸荷槽外,位置如图 5(a) 所示。图 5(b) 为上述 3 点在 0°～360°内的压力变化情况,共有 9 个周期,分别对应 9 个柱塞经过卸荷槽时所产生的压力脉动。图 5(c) 为一个周期的放大,当 $\theta=15°$～$18°$时,柱塞腔到达 C_1 点,压力瞬时上升至 35 MPa,当 $\theta=18°$～$26°$时,随着通流面积的增大,泄压效果增强,C_1 点压力大幅下降,直至不发生变化:当 $\theta=21°$时,柱塞腔与 C_2 点连通,压力上升,当 $\theta=23°$时达到峰值 8 MPa,随后下降至平稳;C_3 点压力变化相对比较稳定。

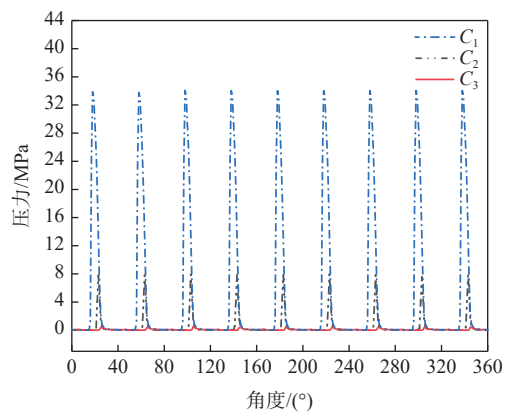
2.2 卸荷槽开度对漩涡、射流角和空化的影响

2.2.1 流体总压分布特征

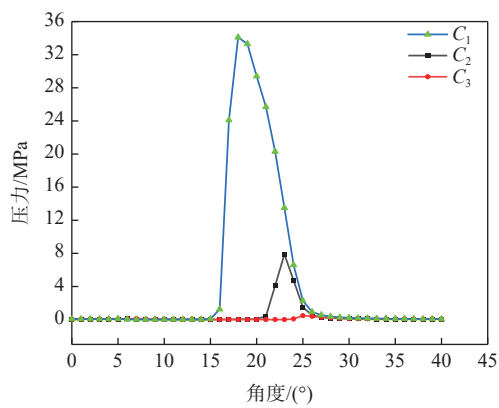
图 6 为流体总压周期变化特征,图 6(a)～图 6(j)表示柱塞 P1 接触卸荷槽到下一个柱塞 P2



(a) 取点位置



(b) 压力脉动全周期



(c) 压力脉动放大

图 5 卸荷槽的压力脉动

Fig. 5 Pressure pulsation of unloading groove

接触卸荷槽的周期过程。图 6(a)～图 6(d)阶段，腰型槽内压力逐渐升高，θ=19°时高压区域面积达到最大。图 6(d)向图 6(j)过渡时，腰型槽内高压区逐渐消失，这是柱塞与卸荷槽连通状态下不同转角对应卸荷槽不同开度所致。当 θ=1°时柱塞 P1 刚接触但未连通卸荷槽，仅在卸荷槽端部出现微小的局部高压区；当 θ 由 1°增大到 19°时，柱塞与卸荷槽接触面积逐渐增大、压力逐步升高；当 θ=20°时，柱塞 P1 与卸荷槽球槽区连通，过流面积突增，柱塞腔内静压降低，腰型槽内压力升高，压差减小。值得注意的是，在图 6(d)和图 6(e)中，即使卸荷槽处静压已经降低，但中间区域流体还存在较高动能，这不但直接增大了射流冲击力，也间接加剧了气蚀现象。

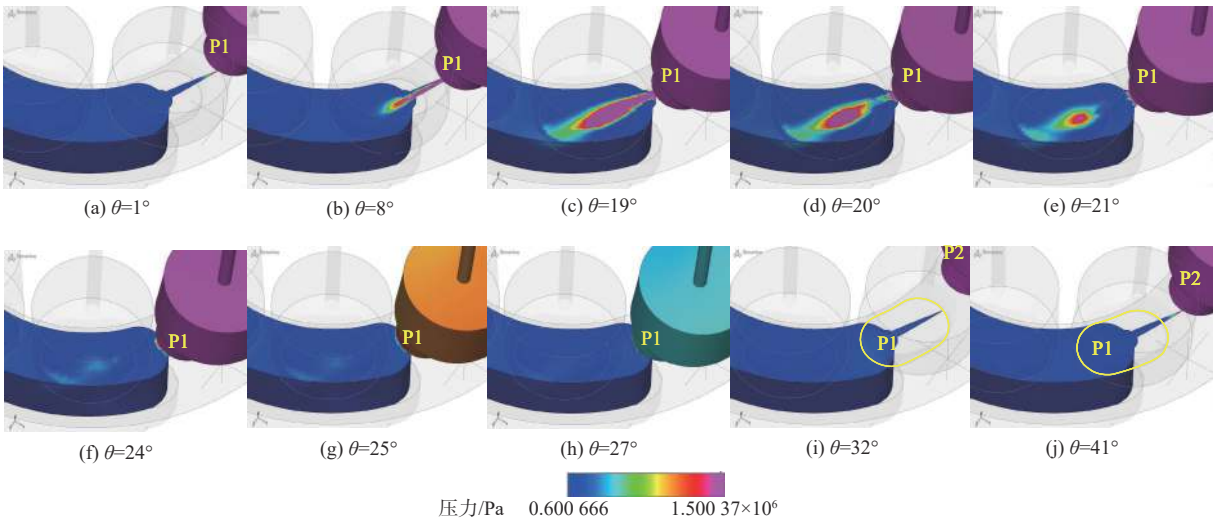


图 6　总压周期

Fig. 6　Total pressure period

2.2.2　卸荷槽开度对空化的影响

图 7 和图 8 分别为静压和气体体积分数的演变过程。由图 7 可知，图 7(a)～图 7(h)阶段，卸荷槽内低压区域逐渐增大，当 θ=22°时低压区域面积达到最大。图 7(h)向图 7(j)过渡时，卸荷槽内低压区逐渐消失。对应图 8(a)～图 8(h)阶段，当 θ=22°时气体体积分数到达 99 % 的区域达到最大值。图 8(h)～图 8(j)阶段，气体区逐渐消失。

之所以出现上述情况，是由于柱塞以不同转角与卸荷槽连通所对应的卸荷槽开度不同。当 θ 由 1°过渡到 25°时，油液在收缩静脉的压力凹陷内，压差变化较大，压力降至液体饱和蒸汽压以下。当 θ 由 25°过渡到 41°时，随着接触面积的增大，柱塞与配流盘的压差大幅减小，压力逐渐回升，气体不再产生并随着流体向下游运动。值得注意的是，在图 8(g)中低压区达到最大值，柱塞 P1 完全与卸荷槽球槽连通，达到柱塞压力大幅下降的临界值，此时压差将达到最大，对应气体体积分数也达到 99%。

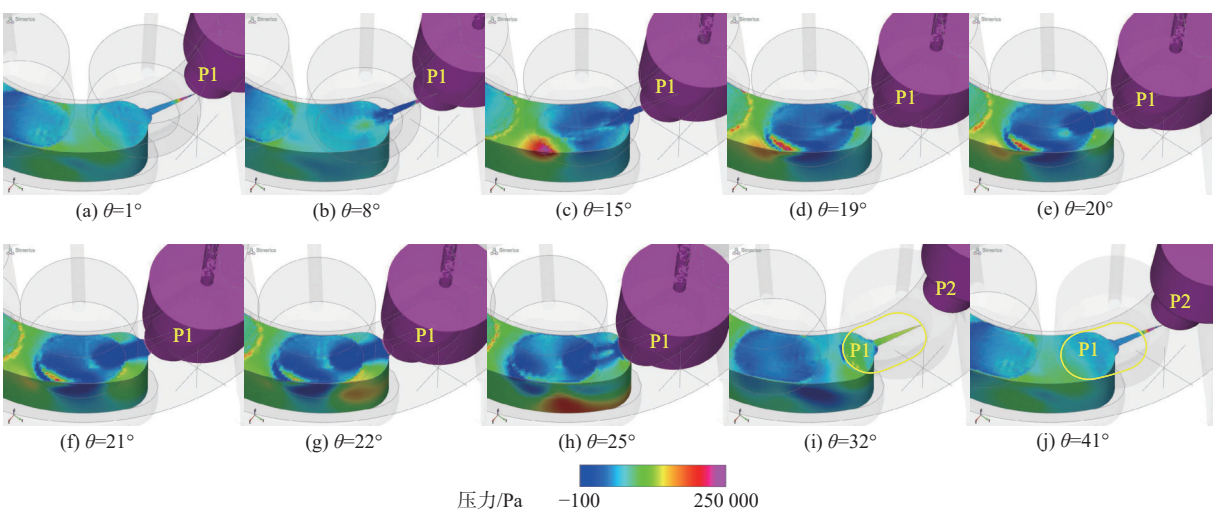


图 7　静压周期

Fig. 7　Static pressure period

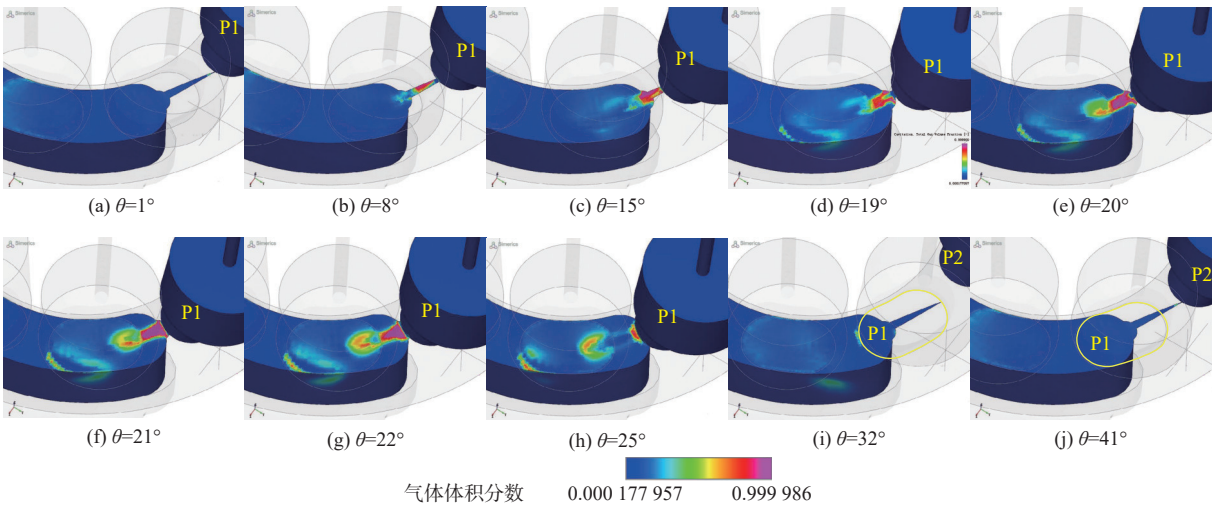


图 8 气体体积分数周期

Fig. 8 Gas volume fraction period

2.2.3 卸荷槽开度对漩涡的影响

图 9 为卸荷槽开度变化过程中漩涡的演变情况。柱塞腔残余压力卸荷过程中，配流盘内部流场极其复杂，存在漩涡周期性地出现与消失的现象。当 θ 由 1°向 19°过渡时，卸荷槽开度逐渐增大，流经预卸压槽后的流体中心速度大于 20 m/s，速度漩涡逐渐产生。当 $\theta \geqslant 20°$时，柱塞腔接触到球槽，随着通流面积的增大，流体中心速度降低到 20 m/s 以下，速度漩涡逐渐消失。因此，流体中心速度对漩涡有影响，流速大于 20 m/s 时，速度漩涡产生，流速小于 20 m/s 时，速度漩涡逐渐消失。

图 10 为图 9(b)中 θ=7°开度下的流场状态的变化过程，可清晰地发现气泡的产生和消失，以及剪切带和旋涡区的发展情况；随着时间的变化，流场达到稳态，气泡和漩涡区随之消失。由此可知，气泡和漩涡在瞬态下产生并在稳态流动下消失。但对柱塞泵而言，每隔一个周期就会产生气泡和漩涡，这是一个瞬态过程，所以气泡和漩涡在柱塞泵中一直存在，并对吸油效率产生影响，从而加大了吸空概率。

产生上述现象的原因是流束在节流槽内收缩加剧，导致流速急剧升高，并形成高速剪切射流区。剪切流周围的油液与周边油液产生流动差，引起差速流动，产生了漩涡流：一个是在高速剪切射流末端形成含有大量空化泡的漩涡空化区；另一个是由于剪切射流的边缘速度在壁面效应的作用下，流体各液层流速不等，所以形成夹带少量气泡的弱流漩涡区。

2.2.4 卸荷槽开度对射流角的影响

图 11 为卸荷槽切面射流角速度云图，当 θ 由 1°向 20°过渡时，由速度云图可观察到，射流角从 φ_1=0°向 φ_4=20°过渡，其速度方向与配流盘/柱塞缸体的摩

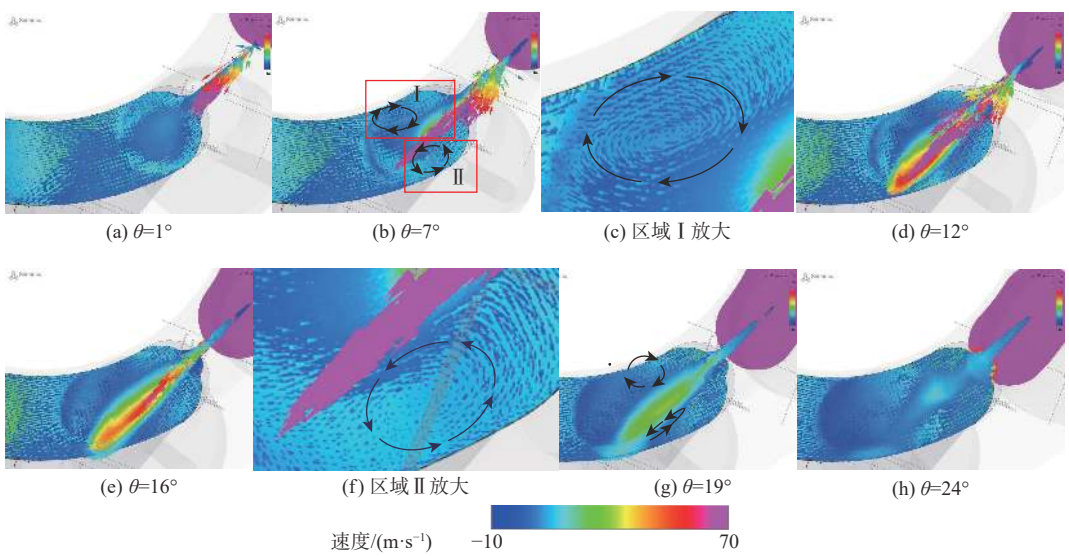
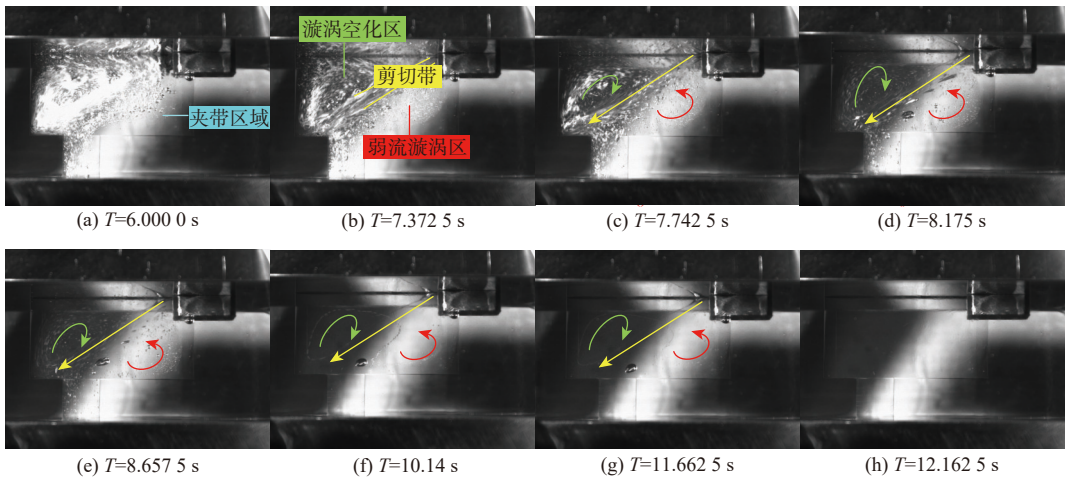


图 9 漩涡周期

Fig. 9 Vortex period

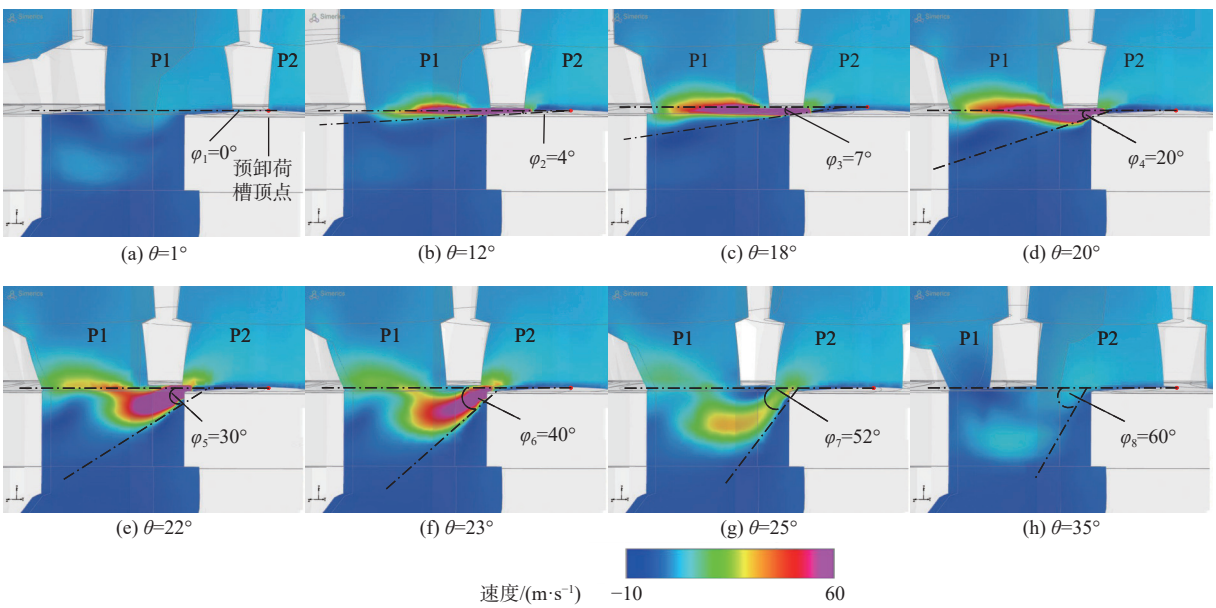


(a) T=6.000 0 s　(b) T=7.372 5 s　(c) T=7.742 5 s　(d) T=8.175 s

(e) T=8.657 5 s　(f) T=10.14 s　(g) T=11.662 5 s　(h) T=12.162 5 s

图 10　可视化过程

Fig. 10　Visualization process



(a) θ=1°　(b) θ=12°　(c) θ=18°　(d) θ=20°

(e) θ=22°　(f) θ=23°　(g) θ=25°　(h) θ=35°

图 11　切面射流角速度云图

Fig. 11　Angular velocity nephogram of jet in section

擦表面接近。当 θ 由 21°过渡到 35°时, 柱塞腔与配流盘接触面积不断增大, 射流角从 φ_4=20°向 φ_8=60°不断增大, 速度方向明显偏离摩擦表面。随着射流角的增大, 射流剪切作用向外扩散, 气泡被流体携带以不同射流角进行射流撞击, 造成了不同区域的射流破坏。

造成以上现象的原因是, 随着卸荷槽开度的变化, 流速流经卸荷槽, 其后过流面积的变化使射流角发生偏移。由图可知, 柱塞 P2 与配流盘接通并保持射流角小于 20°, 大量气泡随着流体射向配流盘/柱塞缸体的摩擦表面附近, 这不仅影响了吸油效率, 还会对柱塞缸体表面造成射流破坏及空蚀。射流角大于 20°将对配流盘内部造成射流破坏及空蚀。

图 12 为不同开度下的可视化, φ_1 表示较小开度下的射流角情况, 与图 11(b) 对应。φ_2 表示柱塞全覆盖卸荷槽下的射流角情况, 与图 11(e) 对应, 即试验与仿真对应。由此可知, 射流角随卸荷槽开度的变化而变化。

由以上分析可知, 当卸荷槽开度 θ<20°时, 对应射流角 φ<20°, 此时卸荷槽内气体体积分数高达 99%, 射流中心速度超过 60 m/s, 射流方向靠近配流盘/缸体摩擦表面, 流体携带气泡以高速射流对其表面产生射流破坏, 同时气泡高速撞击壁面溃灭后产生强烈冲击压造成气蚀破坏, 如图 13(a) 所示。当卸荷槽开度 θ 由 20°向 25°过渡时, 射流角 φ 由 20°向 60°过渡, 对应气体体积分数为 99% 的区域开始减小但仍存在, 射流中心速度超过 35 m/s, 射流方向偏离配流盘/柱塞缸体摩擦表面, 并向配流盘腰型槽内部扩展, 从而造成对配流盘腰型槽内部壁面的破坏, 如图 13(b) 所示。

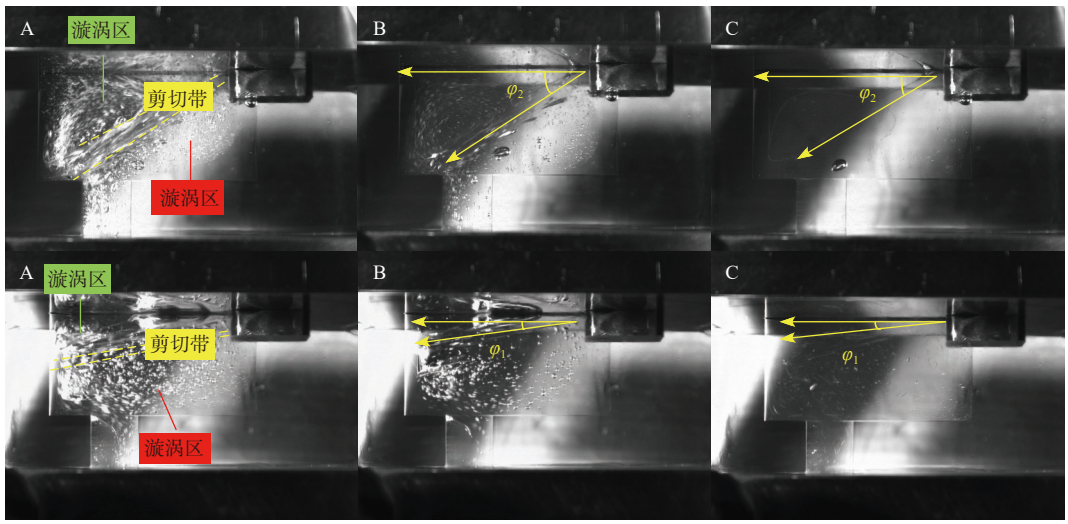
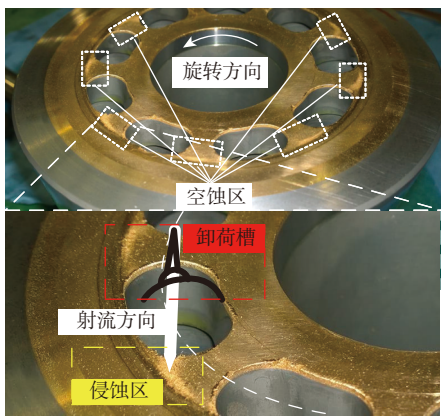
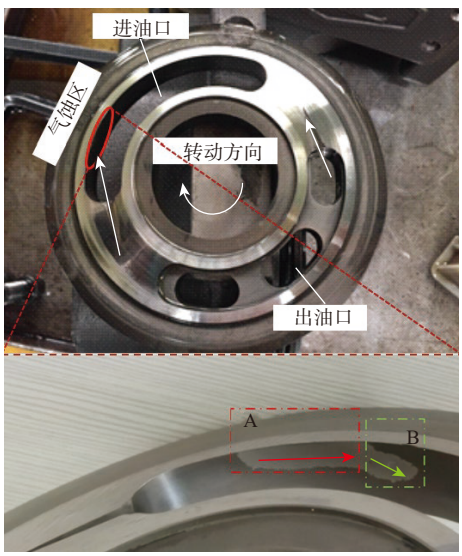


图 12 不同开度下的可视化

Fig. 12 Visualization under different openings



(a) 配流盘/缸体摩擦表面空蚀损伤



(b) 配流盘吸油腰型槽内侧壁空蚀损伤

图 13 侵蚀区域实物照片

Fig. 13 Photos of cavitation damage area

2.3 真实柱塞泵内配流副空蚀现象分析

对轴向柱塞泵样机(与仿真模型一致)进行测试,运行参数如下:转速为 1 500 r/min,斜盘倾角处于最大位置,46 号液压油,吸油口为大气压力,排油压力为 30 MPa。运行一段时间后发现了在配流盘/缸体摩擦表面和配流盘吸油槽腰型槽内侧壁的侵蚀现象,如图 13 所示。

图 13(a) 为配流盘/缸体摩擦表面损伤区域,由图可知,缸体摩擦表面冲蚀痕迹方向与旋转方向相同,冲蚀痕迹在缸体外边缘位置,与卸荷槽射流方向对应,由此可得,柱塞缸体摩擦表面气蚀与卸荷槽有关。造成上述现象的原因有:由图 11 所述的柱塞腔连通卸荷槽时流体流速的变化情况可知,连通过程中流体中心速度基本保持在 60 m/s,即流体以高速射击配流盘/缸体摩擦表面,使摩擦表面遭到强烈的射流破坏;由图 8 所示的气体体积分数可知,流体高速射流的同时还会携带气泡运动,部分气泡将在极短时间内高速撞击壁面,气泡破裂所产生的冲击压作用在壁面产生气蚀破坏;图 11 中射流角的演变过程中,射流角 φ<20°的射流区域包括配流盘/缸体摩擦表面,如图 14 所示的小射流角所对应的射流区域,因此,在气泡、射流角、流体速度

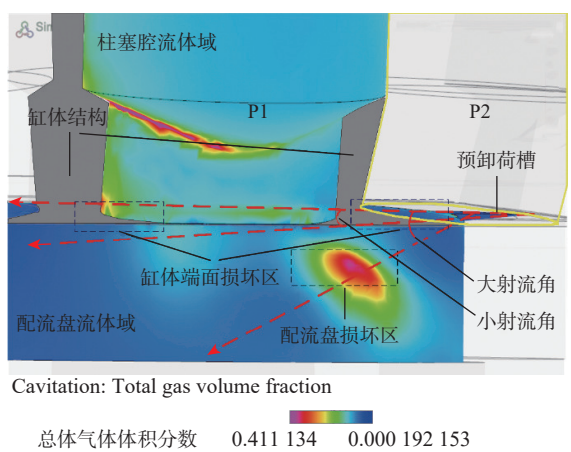


图 14 侵蚀区域

Fig. 14 Cavitation damage area

的共同作用下产生射流破坏和气蚀破坏; 又由图 13(b)中的侵蚀区域 A 可知, 气蚀区域在配流盘吸油腰型槽内壁面上半部分表现不完整, 即不完整损伤区域只可能出现在与配流盘接触的柱塞缸体摩擦表面。由此推断配流盘吸油腰型槽内壁和柱塞缸体摩擦表面的侵蚀正是由卸荷槽喷射出的流体造成的射流破坏和空化对其冲击导致的气蚀破坏引起的。

图 13(b) 为配流盘吸油腰型槽内壁面损伤区域, 由图可知, 球面配流副处侵蚀痕迹的方向沿着转子转动方向, 侵蚀位置位于远离转子中心的外壁面。由图 14 中的射流角所对应的射流区域可知, 小射流角对 A 区域造成破坏, 大射流角对 B 区域造成破坏。由此可得, 配流内壁面的破坏范围与射流角的大小有关。

造成上述现象的原因有: 与上述配流盘/缸体摩擦表面损伤机制相似, 连通过程中的流体中心速度基本保持在 60 m/s, 流体以高速射击配流盘吸油区腰型槽壁面, 使壁面遭到强烈的射流破坏; 另外, 流体携带气泡在极短时间内高速撞击壁面, 气泡破裂所产生的冲击压作用在壁面产生气蚀破坏; 在预卸荷过程中, 当柱塞腔与配流盘连通开度较小时, 射流角区域位于近摩擦表面端靠上部分, 与图 13(b)中 A 侵蚀区域刚好是吻合的, 随着三角槽开度的增大, 柱塞腔与三角槽接触面积增大, 当射流角 $\theta \geqslant 20°$时, 射流区将位于远摩擦表面靠下部分, 与 B 区域刚好吻合。由此可得, 实验与仿真所得冲蚀区域位置吻合。

通过以上分析发现, 当射流角 $\theta<20°$时, 会造成柱塞缸体摩擦表面区域及配流盘内壁靠近缸体接触面的部分损伤; 当射流角 $\theta \geqslant 20°$时, 会造成配流盘内壁下部分的损伤。

3　结　论

本文分析了配流盘卸荷槽与柱塞缸体摩擦表面气蚀的内在关系, 综合利用流体仿真、流场可视化和轴向柱塞泵型式试验等技术手段开展了研究工作, 并获得以下结论:

1) 轴向柱塞泵吸、排油过程中的流体流动具有周期性, 基于该原则设计的卸荷槽流场可视化观察系统是有效的, 能够观察到流体通过卸荷槽时的主要流动特征, 但受有机玻璃材料力学特性的限制, 该系统仅可观察到低压状态下的流场特征。

2) 当柱塞腔周期性接通配流盘, 卸荷槽开度 φ 由 0°向 35°过渡时, 射流角 θ 由 0°向 60°变化。在这个变化过程中, 射流角会影响射流位置, 射流位置从柱塞缸体和配流盘摩擦表面向配流盘内壁偏离。射流角 $\theta<20°$时, 射流方向靠近配流盘/柱塞缸体摩擦表面, 这不仅会将大量气泡射入吸油口, 影响吸油效率, 还会对缸体摩擦表面和配流盘内壁面形成破坏; 射流角 $\theta \geqslant 20°$时, 射流方向向配流盘腰型槽内部扩展, 从而导致配流盘吸油腰型槽内部破坏区域扩大。

3) 在预卸荷阶段, 柱塞腔内残余高压与配流盘卸荷槽连通时将产生高速射流, 并在处于吸油阶段的上一个柱塞下方形成速度漩涡, 漩涡使部分空化产生的气泡形成涡滞留, 导致柱塞吸油受阻。速度漩涡的产生、消失与流体中心速度有关, 以流体中心速度 20 m/s 为界限, 超过此界限时, 速度漩涡将产生, 小于时该速度漩涡逐渐消失。因此, 对于卸荷槽的设计, 建议以最大流速 20 m/s 为一个约束条件。

参考文献（References）

[1] CHAO Q, XU Z, TAO J F, et al. Cavitation in a high-speed aviation axial piston pump over a wide range of fluid temperatures[J]. Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy, 2022, 236(4): 727-737.

[2] CHAO Q, ZHANG J H, XU B, et al. Effects of inclined cylinder ports on gaseous cavitation of high-speed electro-hydrostatic actuator pumps: a numerical study[J]. Engineering Applications of Computational Fluid Mechanics, 2019, 13(1): 245-253.

[3] GULLAPALLI S, MICHAEL P, KENSLER J, et al. An investigation of hydraulic fluid composition and aeration in an axial piston pump[C]//Proceedings of the ASME/BATH Symposium on Fluid Power and Motion Control. Sarasota: American Society of Mechanical Engineers, 2017: 4259.

[4] ZHAO B, GUO W W, QUAN L. Cavitation of a submerged jet at the spherical valve plate/cylinder block interface for axial piston pump[J]. Chinese Journal of Mechanical Engineering, 2020, 33(1): 67.

[5] HARRIS R M, EDGE K A, TILLEY D G. The suction dynamics of positive displacement axial piston pumps[J]. Journal of Dynamic Systems, Measurement, and Control, 1994, 116(2): 281-287.

[6] SINGHAL A K, ATHAVALE M M, LI H Y, et al. Mathematical basis and validation of the full cavitation model[J]. Journal of Fluids Engineering, 2002, 124(3): 617-624.

[7] 苑士华, 周俊杰, 罗先伟, 等. 轴向柱塞泵空化时气相动态演进过程及影响[J]. 兵工学报, 2015, 36(3): 559-565.
YUAN S H, ZHOU J J, LUO X W, et al. Dynamic evolution and effects of gas phase in cavitation of axial piston pump[J]. Acta Armamentarii, 2015, 36(3): 559-565(in Chinese).

[8] 潮群. EHA 轴向柱塞泵高速化若干关键技术研究[D]. 杭州: 浙江大学, 2019.
CHAO Q. Research on some key technologies of EHA axial piston pump with high speed[D]. Hangzhou: Zhejiang University, 2019(in Chinese).

[9] PENG K W, TIAN S C, LI G S, et al. Bubble dynamics characteristics and influencing factors on the cavitation collapse intensity for

self-resonating cavitating jets[J]. Petroleum Exploration and Development, 2018, 45(2): 343-350.

[10] LIANG J, LUO X H, LIU Y S, et al. A numerical investigation in effects of inlet pressure fluctuations on the flow and cavitation characteristics inside water hydraulic poppet valves[J]. International Journal of Heat and Mass Transfer, 2016, 103: 684-700.

[11] SHEN X, ZHANG D S, XU B, et al. Comparative study of tip leakage vortex trajectory and cavitation in an axial flow pump with various tip clearances[J]. Journal of Mechanical Science and Technology, 2022, 36(3): 1289-1302.

[12] CHO I S. A study on the optimum design for the valve plate of a swash plate-type oil hydraulic piston pump[J]. Journal of Mechanical Science and Technology, 2015, 29(6): 2409-2413.

[13] GIOVANNESCHI P, DUFRESNE D. Experimental study of laser-induced cavitation bubbles[J]. Journal of Applied Physics, 1985, 58(2): 651-652.

[14] AKHATOV I, LINDAU O, TOPOLNIKOV A, et al. Collapse and rebound of a laser-induced cavitation bubble[J]. Physics of Fluids, 2001, 13(10): 2805-2819.

[15] BRUJAN E A, KEEN G S, VOGEL A, et al. The final stage of the collapse of a cavitation bubble close to a rigid boundary[J]. Physics of Fluids, 2002, 14(1): 85-92.

[16] ROBINSON P B, BLAKE J R, KODAMA T, et al. Interaction of cavitation bubbles with a free surface[J]. Journal of Applied Physics, 2001, 89(12): 8225-8237.

[17] ZHANG Y, DONG Z, HU X. Cavitation limiting strategies for pumping system: US, 20180040226A1[P]. 2018-02-08.

[18] 郭伟伟. 轴向柱塞泵淹没空化射流特征研究[D]. 太原: 太原理工大学, 2020.
GUO W W. Study on characteristics of submerged cavitation jet in axial piston pump[D]. Taiyuan: Taiyuan University of Technology, 2020(in Chinese).

[19] 築地徹浩, 陈卓, 陈晶晶. 轴向柱塞泵内部空化流的可视化分析[J]. 液压与气动, 2015(2): 1-7.
Tetsuhiro T, CHEN Z, CHEN J J. Visualized analysis of cavitation inside axial piston pump[J]. Chinese Hydraulics & Pneumatics, 2015(2): 1-7(in Chinese).

[20] YAMAGUCHI A, TAKABE T. Cavitation in an axial piston pump[J]. Bulletin of JSME, 1983, 26(211): 72-78.

[21] ITO K, INOUE K, SAITO K. Visualization and detection of cavitation in v-shaped groove type valve plate of an axial piston pump[J]. Proceedings of the JFPS International Symposium on Fluid Power, 1996, 1996(3): 67-72.

[22] QI G N, TIAN T, HE Z Q, et al. Cavitation phenomenon of high-pressure piston pump and control[J]. Machine Tool & Hydraulics, 2021, 49(9): 187-191.

[23] LIU W, WANG A L, SHAN X W, et al. Valve plate for piston pump cavitation problem with the damp groove structural optimization[J]. Applied Mechanics and Materials, 2014, 543-547: 154-157.

[24] KOLLEK W, KUDŹMA Z, STOSIAK M, et al. Possibilities of diagnosing cavitation in hydraulic systems[J]. Archives of Civil and Mechanical Engineering, 2007, 7(1): 61-73.

[25] GB/T 265-1988, 石油产品运动粘度测定法和动力粘度计算法[S]. 北京: 中国标准出版社, 1990.
GB/T 265-1988, Petroleum products-determination of kinematic viscosity and calculation of dynamic viscosity[S]. Beijing: Standards Press of China, 1990(in Chinese).

Influence of unloading groove opening of port plate of plunger pumps on transient flow field characteristics

LIU Shanshan[1]，LEI Xinrui[2]，SONG Ziyu[1]，YANG Kai[1]，HUANG Jiahai[1,*]

(1. College of Mechanical and Carrier Engineering，Taiyuan University of Technology，Taiyuan 030024，China；
2. Taiyuan Heavy Machinery Group Yuci Hydraulic Industry Co.， Ltd.，Jinzhong 030600，China)

Abstract：Cavitation in axial plunger pumps is a main factor affecting its comprehensive performance. The damage in the surface of plunger cylinder is a common failure mode, but the cause is not clear. In view of this damage reason, a three-dimensional CFD model of the pump was established, and a visualization device was designed to observe the flow field characteristics of single plunger cavity and unloading groove on state, so the cause relationship was revealed by simulation and visual measurement of the flow field. The results show that the jet angle increases with the increase in unloading groove opening, accompanied by the generation and disappearance of the vortex, the vortex is generated when the central velocity of the jet is greater than 20 m/s. The simulation results further show that during the transition of the plunger and unloading groove from on state to complete cut-off, the jet angle increases from 0° to 60° gradually. When the jet angle changes from 0° to 20°, the jet direction mainly concentrates on the contact surface between the plunger type cylinder and the waistband groove of the port plate, which leads to jet damage and cavitation damage on the friction surface of the plunger type cylinder and the upper part of the oil absorption waistband groove of the port plate. When the jet angle changes from 20° to 60°, the jet angle expands to the inside of the waistband groove of the port plate, which causes the damaged area of the port plate to move from the upper part to the inside. The formation and disappearance of the vortex are related to the central velocity of the jet. The above research results are helpful in clarifying the causes of cavitation damage on the friction surface of plunger type cylinders, and they provide constraints for the optimization design of port plate structure and have important significance for improving the comprehensive performance of plunger pumps.

Keywords：axial plunger pump; cavitation; valve plate; cylinder block; unloading trough; cavitation damage

Received：2022-08-13；**Accepted**：2022-11-09；**Published Online**：2022-11-24 09：48

URL：link.cnki.net/urlid/11.2625.V.20221122.1343.002

Foundation item：National Key Research and Development Program of China (2021YFB2011903)

* **Corresponding author**. E-mail：huangjiahai@tyut.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0738

空间站舱内悬浮实验台制导控制方法

卢力铮[1,2]，刘伟[1,*]，张皓[1]

(1. 中国科学院太空应用重点实验室 空间应用工程与技术中心，北京 100094；2. 中国科学院大学，北京 100049)

摘 要：针对中国空间站舱内悬浮实验台微重力水平最优的控制问题，考虑舱内运动行程空间约束及各类不确定性因素的影响，以不同时长下的最优微重力水平为目标，提出基于凸优化的空间站舱内悬浮实验台自主轨迹规划方法和基于误差反馈控制的轨迹跟踪方案。介绍悬浮实验台的结构，建立空间相对运动动力学模型，并将其线性化；通过凸优化的方法求解微重力水平最优的运动轨迹及相应的控制力序列；设计基于误差反馈的PID跟踪控制律，兼顾微重力水平的同时实现对优化轨迹的快速跟踪控制。数值仿真结果表明：当微重力科学实验时长需求为1 h时，最优微重力水平可达9.121×10^{-6} m/s^2。基于凸优化的空间站舱内悬浮实验台自主轨迹规划方法和基于误差反馈控制的轨迹跟踪方案，能够实现指定时长需求的最优微重力水平，给当前中等时长（1 h）最优微重力水平问题提供了解决思路，提升了中国空间站舱内悬浮实验台的应用价值。

关 键 词：凸优化；悬浮实验台；轨迹规划；相对运动；最优微重力水平
中图分类号：V448.2
文献标志码：A **文章编号**：1001-5965(2024)09-2930-09

中国载人航天三期工程如火如荼地进行，航天医学、航天技术、空间生命科学等领域一系列的空间科学实验[1-4]不断开展，对良好的空间微重力环境提出了更高的要求。然而，航天器在轨运行时不仅会受到大气阻力、太阳辐射光压等外部扰动的影响，其自身也会被航天员动作、航天器变轨、柔性部件颤振等内部扰动所影响，导致其微重力水平难以达到预期[5]。因此，需要悬浮实验台来隔离航天器本体的振动干扰。

当前中国空间站舱内悬浮实验台采用舱内无拖曳控制和两级隔振控制2种控制策略[6]。前者能够提供最优的微重力水平，但是受到空间站实验舱空间的约束，其无拖曳运动时间受限，适用于微重力水平需求高且时间短的科学实验；后者通过PID控制使悬浮实验台内体移动到外体的行程空间的中心位置并保持稳定，在提供较高微重力水平的同时确保实验的长期进行，但也损失了一定的微重力水平，而提供中等时长(1 h)高微重力水平的方案一直有待探索。

以有限时长的最优微重力水平为出发点，本文对悬浮实验台内体的轨迹规划方案进行了深入研究。内体的姿态可以采用闭环控制的方式保持稳定，避免与舱壁发生碰撞，方法简单高效的同时也不会影响内体的微重力水平；内体的飞行轨迹则需要进行优化，充分利用舱内运动行程空间，探索不同时长下微重力水平最优的运动轨迹。

多项式制导法[7]是一种便于工程应用的近似燃料最优制导方法，在阿波罗登月计划中首次运用。考虑到燃耗最优性，孙军伟等[8]采用开环最优制导方法对月球软着陆的多项式制导控制律进行设计，得出一种易于实现的次优闭环实时制导控制方法。文献[9-10]提出使用无损凸优化方法解决

收稿日期：2022-08-24；录用日期：2023-01-17；网络出版时间：2023-02-07 00:00
网络出版地址：link.cnki.net/urlid/11.2625.V.20230206.1635.005
基金项目：国家自然科学基金(62003325)
*通信作者. E-mail：liuwei@csu.ac.cn

引用格式：卢力铮，刘伟，张皓. 空间站舱内悬浮实验台制导控制方法[J]. 北京航空航天大学学报，2024，50(9)：2930-2938.
LU L Z，LIU W，ZHANG H. Guidance and control method of levitation experiment facility inside China's Space Station[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2930-2938 (in Chinese).

动力下降过程中的非凸约束问题, 引发了基于凸优化制导算法的技术创新。自此, 越来越多的学者利用凸优化方法解决动力下降过程中的轨迹优化问题, 并开展了相关的实验测试。针对火箭在地球的垂直软着陆问题, Blackmore[11] 团队利用凸优化计算的高速性, 实现了在箭载计算机上数百毫秒内生成火箭动力高精度着陆轨迹。邓云山等[12] 针对扰动环境下火星着陆轨迹重规划问题, 使用序列凸优化方法实现了在终端时间自由条件下的自主轨迹规划。王劲博等[13] 将凸优化技术与伪谱离散方法结合, 研究出了高精度快速轨迹优化算法, 并将其成功运用于垂直起降可重复使用运载火箭子级返回着陆问题。以上研究结果充分证明了凸优化方法具有计算效率高、收敛性好等特点, 能有效应对各类过程、终端约束, 是实时快速处理轨迹规划问题的优质方法。

综上所述, 目前空间站舱内悬浮实验台在轨迹规划方面研究较少, 不能提供给定时长下最优的微重力水平。本文针对悬浮实验台内体的飞行轨迹, 提出了一种基于凸优化理论的空间站舱内悬浮实验台自主轨迹跟踪制导方法。建立空间相对运动动力学模型, 并将其线性化; 采用凸优化方法求解不同时间长度的微重力最优化的运动轨迹以及最优的控制力序列; 研究内体跟随微重力最优化轨迹的控制算法, 克服建模误差和各项不确定性因素的影响, 兼顾控制力大小与微重力水平的同时, 实现轨迹跟踪控制。

1　悬浮实验台动力学建模

1.1　悬浮实验台简介

中国空间站核心舱安置了医学样本分析与高微重力科学实验柜, 简称高微柜(high microgravity level research rack, HMLR), 如图 1 所示[14]。高微柜致力于隔离空间站本体的振动干扰, 以提供更优的微重力水平。目前, 航天医学、基础物理、空间生命科学等领域一系列的微重力科学实验已在空间站陆续开展。



(a) 悬浮实验台　　(b) 高微柜

图 1　悬浮实验台和高微柜示意

Fig. 1　LEF and HMLR

悬浮实验台(levitation experiment facility, LEF)是 HMLR 的核心部分, 是一个双层悬浮平台, 包括内体和外体[15] 两部分。外体是装置的支撑单元, 内体是科学实验载荷的实验平台, 内体与外体之间采用无线传能和无线通信的方式进行供电和数据通信。内体采用电磁作动器进行磁悬浮控制, 外体采用喷气推力器进行喷气悬浮控制。

悬浮实验台有柜内、柜外工作 2 种工作模式: 在柜内工作模式下, 悬浮实验台安装在高微柜内部, 相对高微柜的运动行程范围约束为±50 mm, 行程空间很小且不利于进行轨迹规划研究; 在柜外工作模式下, 航天员将悬浮实验台从高微柜中取出, 将其在空间站核心舱的指定区域内释放, 此时悬浮实验台的运动行程范围可达±500 mm。因此, 本文针对柜外工作模式, 探索在行程空间约束下内体微重力水平最优的运动轨迹规划方法, 实现不同时长下悬浮实验台提供最优的微重力水平。

1.2　空间相对运动动力学模型

以中国空间站舱内搭载的悬浮实验台为研究对象, 建立坐标系如图 2 所示。

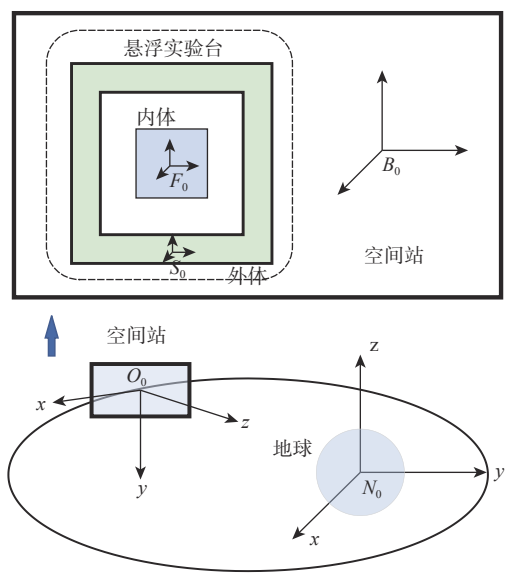


图 2　坐标系示意

Fig. 2　Coordinate system

1) 地心赤道惯性坐标系, 简称 N 系, 坐标原点 N_0 为地心; x 轴在赤道平面内, 指向 J2000.0 时刻的平春分点; z 轴垂直于赤道面, 指向北极; y 轴与 x 轴和 z 轴垂直, 构成右手直角坐标系。

2) 质心轨道坐标系, 简称 O 系, 坐标原点 O_0 为卫星质心; x 轴沿轨道平面与当地水平面的交线指

向前进方向；z轴沿当地垂线指向地心；y轴垂直于轨道平面，与另两轴构成右手直角坐标系。该坐标系在空间以卫星的轨道角速度绕y轴旋转，且旋转方向与y轴的正方向相反。

3）空间站本体坐标系，简称B系，坐标系原点B_0为空间站核心舱质心，与空间站固联。三轴方向的选取与O系相同。

4）外体坐标系，简称S系，坐标系原点S_0为悬浮实验台外体质心，与外体固联。S系三轴方向的选取与B系相同。S系是内体相对外体运动的参考坐标系。

5）内体坐标系，简称F系，坐标系原点F_0为内体质心，与内体固联。F系三轴方向的选取与S系相同。

根据以上坐标系定义，可得坐标变化矩阵。为建立空间站与悬浮实验台内体的轨道动力学模型，进行受力分析，空间站和悬浮实验台内质心受力分析如表1所示。

表1 空间站和悬浮实验台内质心受力分析

Table 1 Centroid force analysis of CSS and LEF inner body

类型	符号	说明
空间站	$\boldsymbol{f}_{\mathrm{graE_B}}$	空间站所受地球球形引力加速度
	$\boldsymbol{f}_{\mathrm{gra_B}}$	空间站所受地球非球形引力摄动加速度和其他天体引力加速度，包括日、地、月三体引力摄动加速度等
	$\boldsymbol{f}_{\mathrm{env_B}}$	环境摄动加速度，包括大气阻力、太阳辐射光压等引起的摄动加速度
	$\boldsymbol{f}_{\mathrm{ctr_B}}$	对空间站施加的控制加速度
	$\boldsymbol{f}_{\mathrm{oth_B}}$	其他扰动加速度，包括载荷动作、风扇转动等引起的振动加速度
悬浮实验台	$\boldsymbol{f}_{\mathrm{graE_F}}$	悬浮实验台所受地球球形引力加速度
	$\boldsymbol{f}_{\mathrm{gra_F}}$	悬浮实验台所受地球非球形引力摄动加速度和其他天体引力加速度，包括日、地、月三体引力摄动加速度等
	$\boldsymbol{f}_{\mathrm{env_F}}$	环境摄动加速度，包括空气流动、温度波动等引起的摄动加速度
	$\boldsymbol{f}_{\mathrm{ctr_F}}$	对悬浮实验台内体施加的控制加速度
	$\boldsymbol{f}_{\mathrm{oth_F}}$	其他扰动加速度，包括测量扰动、电磁扰动等与外体耦合的扰动加速度

基于牛顿第二定律，建立空间站质心与悬浮实验台内体质心的轨道动力学模型：

$$\ddot{\boldsymbol{r}}_{\mathrm{NB}}|_{\mathrm{N}}=\boldsymbol{f}_{\mathrm{graE_B}}+\boldsymbol{f}_{\mathrm{gra_B}}+\boldsymbol{f}_{\mathrm{env_B}}+\boldsymbol{f}_{\mathrm{ctr_B}}+\boldsymbol{f}_{\mathrm{oth_B}} \tag{1}$$

$$\ddot{\boldsymbol{r}}_{\mathrm{NF}}|_{\mathrm{N}}=\boldsymbol{f}_{\mathrm{graE_F}}+\boldsymbol{f}_{\mathrm{gra_F}}+\boldsymbol{f}_{\mathrm{env_F}}+\boldsymbol{f}_{\mathrm{ctr_F}}+\boldsymbol{f}_{\mathrm{oth_F}} \tag{2}$$

式中：$\boldsymbol{f}_{\mathrm{graE_B}}=-\mu_{\mathrm{E}}\cdot\boldsymbol{r}_{\mathrm{NB}}/r_{\mathrm{NB}}^3$，$\boldsymbol{r}_{\mathrm{NB}}$为空间站质心在N系中的位置矢量，$r_{\mathrm{NB}}$为$\boldsymbol{r}_{\mathrm{NB}}$的模值；$\mu_{\mathrm{E}}$为地球引力常数，$\mu_{\mathrm{E}}=3.986\,005\times10^{14}\ \mathrm{m^3/s^2}$；$\boldsymbol{f}_{\mathrm{graE_F}}=-\mu_{\mathrm{E}}\cdot\boldsymbol{r}_{\mathrm{NF}}/r_{\mathrm{NF}}^3$，$\boldsymbol{r}_{\mathrm{NF}}$为悬浮实验台内体质心在N系中的位置矢量，$r_{\mathrm{NF}}$为$\boldsymbol{r}_{\mathrm{NF}}$的模值。

在柜外工作模式下，由于悬浮实验台的行程空间中心位置通常无法设置在空间站的质心位置，因此计算悬浮实验台的微重力加速度时要考虑偏离质心的附加微重力加速度。悬浮实验台在空间站核心舱中的运动空间的几何中心点为P，其到核心舱质心的位置矢量为${}^{(\mathrm{B})}\boldsymbol{r}_{BP}$。根据刚体两点运动关系，由航天器质心运动计算$P$点的运动方程为

$$\begin{cases}\boldsymbol{r}_{\mathrm{N}P}=\boldsymbol{r}_{\mathrm{NB}}+\boldsymbol{r}_{BP}\\ \dot{\boldsymbol{r}}_{\mathrm{N}P}|_{\mathrm{N}}=\dot{\boldsymbol{r}}_{\mathrm{NB}}|_{\mathrm{N}}+{}^{\mathrm{B}}\boldsymbol{\omega}^{\mathrm{N}}\times\boldsymbol{r}_{BP}\\ \ddot{\boldsymbol{r}}_{\mathrm{N}P}|_{\mathrm{N}}=\ddot{\boldsymbol{r}}_{\mathrm{NB}}|_{\mathrm{N}}+{}^{\mathrm{B}}\boldsymbol{\omega}^{\mathrm{N}}\times\left({}^{\mathrm{B}}\boldsymbol{\omega}^{\mathrm{N}}\times\boldsymbol{r}_{BP}\right)+{}^{\mathrm{B}}\boldsymbol{\alpha}^{\mathrm{N}}\times\boldsymbol{r}_{BP}\end{cases} \tag{3}$$

式中：$\boldsymbol{r}_{\mathrm{N}P}$、$\dot{\boldsymbol{r}}_{\mathrm{N}P}|_{\mathrm{N}}$、$\ddot{\boldsymbol{r}}_{\mathrm{N}P}|_{\mathrm{N}}$分别为空间站上$P$点相对于惯性系的位置、速度、加速度，$P$点相对于核心舱质心的位置矢量$\boldsymbol{r}_{BP}=\left(r_{BP_x},r_{BP_y},r_{BP_z}\right)$；${}^{\mathrm{B}}\boldsymbol{\omega}^{\mathrm{N}}$、${}^{\mathrm{B}}\boldsymbol{\alpha}^{\mathrm{N}}$分别为空间站本体角速度与角加速度在惯性系中的列阵。

联立式(1)～式(3)，根据几何关系以及刚体复合运动规律得空间站核心舱P点与悬浮实验台内体之间的空间相对运动动力学模型为

$$\begin{aligned}&\ddot{\boldsymbol{r}}_{FP}|_{\mathrm{B}}+2{}^{\mathrm{N}}\boldsymbol{\omega}^{\mathrm{B}}\times\dot{\boldsymbol{r}}_{FP}|_{\mathrm{B}}+{}^{\mathrm{N}}\boldsymbol{\alpha}^{\mathrm{B}}\times\boldsymbol{r}_{FP}+{}^{\mathrm{N}}\boldsymbol{\omega}^{\mathrm{B}}\times\left({}^{\mathrm{N}}\boldsymbol{\omega}^{\mathrm{B}}\times\boldsymbol{r}_{FP}\right)=\\ &\quad\boldsymbol{f}_{\mathrm{graE_F}}+\boldsymbol{f}_{\mathrm{gra_F}}+\boldsymbol{f}_{\mathrm{env_F}}+\boldsymbol{f}_{\mathrm{ctr_F}}+\boldsymbol{f}_{\mathrm{oth_F}}-\\ &\quad\boldsymbol{f}_{\mathrm{graE_B}}-\boldsymbol{f}_{\mathrm{gra_B}}-\boldsymbol{f}_{\mathrm{env_B}}-\boldsymbol{f}_{\mathrm{ctr_B}}-\boldsymbol{f}_{\mathrm{oth_B}}-\\ &\quad{}^{\mathrm{B}}\boldsymbol{\omega}^{\mathrm{N}}\times({}^{\mathrm{B}}\boldsymbol{\omega}^{\mathrm{N}}\times\boldsymbol{r}_{BP})-{}^{\mathrm{B}}\boldsymbol{\alpha}^{\mathrm{N}}\times\boldsymbol{r}_{BP}\end{aligned} \tag{4}$$

式中：$\boldsymbol{r}_{FP}=\left(r_x,r_y,r_z\right)$；$\dot{\boldsymbol{r}}_{FP}=\left(v_x,v_y,v_z\right)$；${}^{\mathrm{N}}\boldsymbol{\omega}^{\mathrm{B}}$、${}^{\mathrm{N}}\boldsymbol{\alpha}^{\mathrm{B}}$分别为空间站本体角速度与角加速度在本体系中的列阵。

式(4)表示的是非线性的空间相对运动动力学模型，为动力学分析以及轨迹规划方法的运用带来一定的困难。为建立线性化的空间相对运动动力学模型，提出如下合理假设：

1）中国空间站的飞行轨道近似理想圆轨道并且三轴对地定向；悬浮实验台的工作时间避开舱段对接、轨道维持等大幅动作过程，对空间站施加的控制加速度可忽略不计。

2）空间站质心与悬浮实验台质心受到的非球形引力摄动加速度和其他天体引力加速度近似。

3）为屏蔽舱内空气等对内体的扰动，悬浮实验台外体跟随内体运动，因此内体受到的环境扰动加速度和其他扰动加速度可忽略不计。

由此可得空间相对运动的线性化方程：

$$\begin{cases}\dot{\boldsymbol{X}}=\boldsymbol{A}\boldsymbol{X}+\boldsymbol{B}\boldsymbol{u}+\boldsymbol{C}\\ \boldsymbol{X}=[r_x\quad r_y\quad r_z\quad v_x\quad v_y\quad v_z]^{\mathrm{T}}\\ \boldsymbol{u}=[u_x\quad u_y\quad u_z]^{\mathrm{T}}\end{cases} \tag{5}$$

式中：$\boldsymbol{X}$为由相对位置与相对速度组成的状态变量；$\boldsymbol{A}$、$\boldsymbol{B}$、$\boldsymbol{C}$分别为状态方程系统矩阵、状态仿真

输入矩阵、偏离质心产生的附加微重力加速度导致的矩阵。

$$\boldsymbol{A}=\begin{bmatrix}0&0&0&1&0&0\\0&0&0&0&1&0\\0&0&0&0&0&1\\0&0&0&0&0&2\omega\\0&-\omega^2&0&0&0&0\\0&0&3\omega^2&-2\omega&0&0\end{bmatrix}$$

$$\boldsymbol{B}=\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\\1&0&0\\0&1&0\\0&0&1\end{bmatrix},\quad \boldsymbol{C}=\begin{bmatrix}0\\0\\0\\0\\-r_{BP_y}\cdot\omega^2\\3\cdot r_{BP_z}\cdot\omega^2\end{bmatrix}$$

其中, ω为空间站的轨道角速度; $\boldsymbol{u}$ 为控制量, $\boldsymbol{u}=\boldsymbol{f}_{\mathrm{ctr_F}}$。

2　轨迹规划与跟踪控制方法设计

当前中国空间站舱内悬浮实验台采用的无拖曳控制策略与两级隔振控制策略, 未能充分利用舱内空间, 不能满足中等时长科学实验的最优微重力水平需求, 如图 3 所示。因此, 本文研究的关键在于充分利用舱内空间、实时规划内体微重力水平最优的运动轨迹, 以及在尽量减少对微重力水平破坏的原则下跟踪微重力最优的运动轨迹。

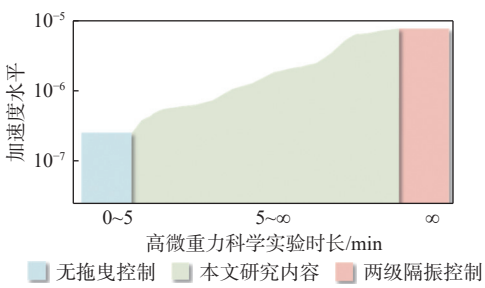


图 3　不同控制策略下的微重力加速度水平

Fig. 3　Microgravity acceleration level of different control strategies

本文采用的方案分以下两步: 第一步是通过凸优化方法得到指定时长内体微重力水平最优的运动轨迹; 第二步是对优化轨迹的跟踪控制, 通过设计合适的跟踪控制律, 采用较小的误差反馈控制力以减少对微重力水平的破坏, 实现存在扰动等不确定性情况下对于最优轨迹的跟踪, 最终实现高微重力水平。本文的整体框架如图 4 所示。

2.1　基于凸优化的自主轨迹规划

由于空间站舱内悬浮实验台面向不同时长的空间科学实验, 因此需要具备实时自主最优微重力水平轨迹规划的能力, 这为轨迹规划方法提出了新的挑战。本文采用凸优化的方法, 实时求解空间站舱内悬浮实验台内体的轨迹规划问题。

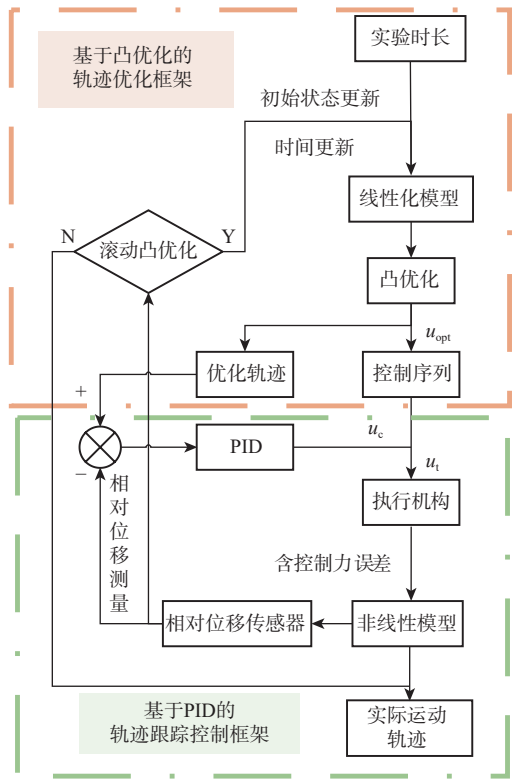


图 4　本文整体框架

Fig. 4　Overall framework of this paper

本文的目标是提供指定时长最优的微重力水平即使悬浮实验台内体控制力最小, 来减少对微重力水平的破坏。因此, 轨迹优化的目标函数可以描述为内体控制力加速度模值的最大值为

$$\min(\max(u(t)))\quad \forall t\in[0,t_{\mathrm{f}}] \tag{6}$$

$$u(t)=\|\boldsymbol{u}(t)\|_2=\sqrt{u_x(t)^2+u_y(t)^2+u_z(t)^2} \tag{7}$$

式中: t 为当前实验时刻; t_{f}为终端时间。

根据式 (5) 可得空间相对运动动力学模型约束。其中, $\boldsymbol{X}=[r_x\quad r_y\quad r_z\quad v_x\quad v_y\quad v_z]^{\mathrm{T}}$为状态向量, $\boldsymbol{u}=[u_x\quad u_y\quad u_z]^{\mathrm{T}}$为控制量。

考虑到悬浮实验台执行机构的控制力大小有一定的限制, 对控制力加速度施加约束:

$$0\leqslant\|\boldsymbol{u}(t)\|_2\leqslant f_{\max}\quad \forall t\in[0,t_{\mathrm{f}}] \tag{8}$$

式中: $f_{\max}$为执行机构提供的最大控制力加速度。

考虑到内体优化轨迹与实际轨迹存在偏差, 为防止悬浮实验台与舱壁发生碰撞, 设计安全裕度参数对悬浮实验台的行程空间进行约束。设悬浮实验台在核心舱的行程空间为$a\times b\times c$, 在实际问题中, 安全裕度参数$\eta\in(0,1)$为

$$\begin{cases}-0.5\cdot a\cdot\eta\leqslant r_x(t)\leqslant 0.5\cdot a\cdot\eta\\-0.5\cdot b\cdot\eta\leqslant r_y(t)\leqslant 0.5\cdot b\cdot\eta\\-0.5\cdot c\cdot\eta\leqslant r_z(t)\leqslant 0.5\cdot c\cdot\eta\end{cases}\quad\forall t\in[0,t_{\mathrm{f}}] \tag{9}$$

式中：$r_x(t)$、$r_y(t)$与$r_z(t)$为悬浮实验台内体相对于空间站舱内P点的位置在轨道坐标系下的投影分量。

针对式(7)中非凸形式的等式约束，无法直接运用凸优化方法求解。基于凸优化理论，对其进行松弛，可转化为

$$u(t) \leqslant \sqrt{u_x(t)^2+u_y(t)^2+u_z(t)^2} \quad \forall t\in[0,t_f] \tag{10}$$

综上所述，空间站舱内悬浮实验台微重力水平最优的优化问题可表示如下：

$$\begin{cases} \min(\max(u(t))) \quad \forall t\in[0,t_f] \\ \text{s.t.} \quad \text{式(5)},\text{式(8)},\text{式(9)},\text{式(10)} \end{cases} \tag{11}$$

轨道动力学模型是连续且时变的，而凸优化只能解决有限维度的优化问题，因此对其进行离散化处理。考虑到实时精确轨迹规划，需要兼顾计算精度与效率，本文采用龙格库塔法进行离散化，保证求解精度的同时，形式也并不复杂。定义终端时间为t_f，间隔时间为Δt。在实验时间段$[0,t_f]$内取等间隔离散时间序列$\{t_0,t_1,\cdots,t_n\}$，将式(8)~式(10)中的时间段t离散化。同时，采用四阶龙格库塔法将式(5)离散化，可得

$$\begin{cases} \boldsymbol{X}_{k+1}=\boldsymbol{X}_k+\boldsymbol{k}_1/6+\boldsymbol{k}_2/3+\boldsymbol{k}_3/3+\boldsymbol{k}_4/6 \\ \boldsymbol{k}_1=\Delta t\cdot(\boldsymbol{A}\boldsymbol{X}_k+\boldsymbol{B}u_k+\boldsymbol{C}) \\ \boldsymbol{k}_2=\Delta t\cdot[\boldsymbol{A}(\boldsymbol{X}_k+\boldsymbol{k}_1/2)+\boldsymbol{B}(\boldsymbol{u}_k+\boldsymbol{u}_{k+1})/2+\boldsymbol{C}] \\ \boldsymbol{k}_3=\Delta t\cdot[\boldsymbol{A}(\boldsymbol{X}_k+\boldsymbol{k}_2/2)+\boldsymbol{B}(\boldsymbol{u}_k+\boldsymbol{u}_{k+1})/2+\boldsymbol{C}] \\ \boldsymbol{k}_4=\Delta t\cdot[\boldsymbol{A}(\boldsymbol{X}_k+\boldsymbol{k}_3)+\boldsymbol{B}\boldsymbol{u}_{k+1}+\boldsymbol{C}] \end{cases} \tag{12}$$

式中：下标k为第k个离散点。

此时，可采用凸优化方法进行数值求解，以解决有限维度离散时间凸优化问题，得到微重力水平最优轨迹序列$\boldsymbol{X}_{\text{opt}}$及优化控制序列$\boldsymbol{u}_{\text{opt}}$。

2.2 基于误差反馈的轨迹跟踪控制

由于线性化模型误差、扰动不确定性、测量误差和执行误差的存在，悬浮实验台内体的实际飞行轨迹与微重力水平最优轨迹存在一定的偏差，需要通过实时闭环反馈控制进行修正。

式(2)表明内体的微重力水平由控制力、内外体间的耦合力、空气扰动力、测量扰动力、电磁扰动力、温度波动扰动等共同决定。其中，占主导部分的是对内体的控制力。由此可见，在优化控制力确定的情况下，跟踪控制力越小，内体的微重力水平越接近最优。

悬浮实验台一方面要实现对优化轨迹的跟踪，避免与空间站舱壁发生碰撞；另一方面要约束跟踪控制力，减少对微重力水平的破坏。高精度跟踪轨迹则需要较大的控制力实现快速跟踪，较大的控制力会降低微重力水平，这两个目标是矛盾的。因此，本文的研究重点在于合理设计控制器参数，保证轨迹跟踪精度的同时减少对微重力水平的破坏。

本文采用基于实时误差反馈的PID反馈控制，PID控制器参数搜索方案是离线进行的，当所有计算完成后将控制器参数值从地面上传至悬浮实验台。根据在轨测试结果，可以在地面重新计算控制器参数值并再次上传，不会影响科学实验的正常进行。

基于PID控制器的反馈控制力$\boldsymbol{F}_{\text{ctr}}$为

$$\boldsymbol{F}_{\text{ctr}}=m_{\text{F}}\boldsymbol{u}_{\text{c}}=\boldsymbol{K}_{\text{p}}\boldsymbol{r}_{\text{e}}+\boldsymbol{K}_{\text{d}}\dot{\boldsymbol{r}}_{\text{e}}+\boldsymbol{K}_{\text{i}}\int\boldsymbol{r}_{\text{e}}\text{d}t \tag{13}$$

式中：m_{F}为悬浮实验台内体的质量；$\boldsymbol{u}_{\text{c}}$为反馈控制力加速度；$\boldsymbol{K}_{\text{p}}$、$\boldsymbol{K}_{\text{d}}$、$\boldsymbol{K}_{\text{i}}$分别为PID控制器的比例项系数、微分项系数、积分项系数；$\boldsymbol{r}_{\text{e}}$为内体实际位置相对优化轨迹的实时位移误差，$\dot{\boldsymbol{r}}_{\text{e}}$为内体实际速度相对优化速度的实时速度误差，计算式分别为

$$\begin{cases} \boldsymbol{r}_{\text{e}}=\boldsymbol{r}_{\text{PF}}-\boldsymbol{r}_{\text{PFd}} \\ \dot{\boldsymbol{r}}_{\text{e}}=\dot{\boldsymbol{r}}_{\text{PF}}-\dot{\boldsymbol{r}}_{\text{PFd}} \end{cases} \tag{14}$$

式中：$\boldsymbol{r}_{\text{PF}}$、$\dot{\boldsymbol{r}}_{\text{PF}}$分别为内体的实际相对位置、相对速度；$\boldsymbol{r}_{\text{PFd}}$、$\dot{\boldsymbol{r}}_{\text{PFd}}$分别为内体相对位置、相对速度的控制目标，即2.1节计算出的最优轨迹序列$\boldsymbol{X}_{\text{opt}}$。

与常规设计PID控制器参数的方法不同，本文采用一种全新的参数整定策略，流程如图5所示。

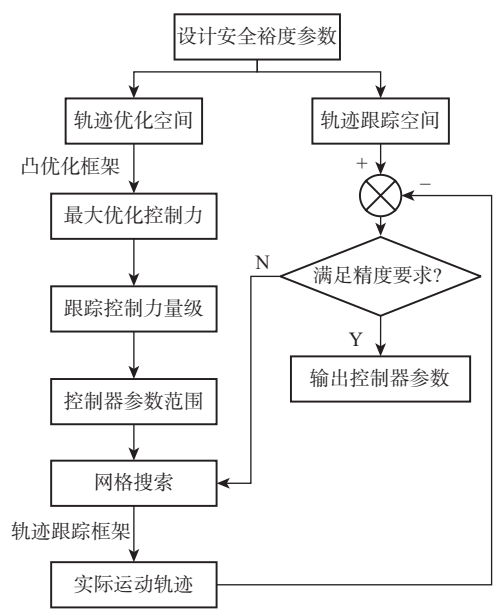


图5 控制器参数确定流程

Fig. 5 Flow chart of determining controller parameters

由于控制目标需要兼顾跟踪精度与微重力水平，式(9)中引入合适的安全裕度参数η很有必要。一方面，不充分利用行程空间会导致微重力水平降低，因此轨迹优化的空间仍要占主要部分；另一方面，预留的轨迹跟踪空间要保证悬浮实验台不与舱壁发生碰撞。综合考虑这两方面因素，参数η取0.9较为合适。

根据凸优化求解出的最大优化控制力, 确定轨迹跟踪控制力量级在最大优化控制力的 1/10 左右, 通过式 (13) 计算出控制器参数的合适范围。

采用网格搜索法, 选取合适的步长, 对 $\boldsymbol{K}_{\mathrm{p}}$、$\boldsymbol{K}_{\mathrm{d}}$进行搜索, 求解不同控制器参数下悬浮实验台内体的实际运动轨迹, 并判断跟踪精度是否满足避免碰撞的要求, 最终确定出控制效果最佳的$\boldsymbol{K}_{\mathrm{p}}$、$\boldsymbol{K}_{\mathrm{d}}$参数。若实际运动轨迹与优化轨迹存在稳态误差, 则根据误差的大小设计参数$\boldsymbol{K}_{\mathrm{i}}$对其进行校正。

最终, 基于此 PID 控制器实时计算出控制序列 $\boldsymbol{u}_{\mathrm{c}}$, 并将优化控制序列$\boldsymbol{u}_{\mathrm{opt}}$与跟踪控制序列 $\boldsymbol{u}_{\mathrm{c}}$ 相加得出完整的控制序列 $\boldsymbol{u}_{\mathrm{t}}$, 其是最终对悬浮实验台内体施加的控制力, 即动力学模型中的$\boldsymbol{f}_{\mathrm{ctr_F}}$。

3　数值仿真与结果分析

3.1　仿真工具与仿真参数

本文采用 MATLAB 中的凸优化求解工具箱对提出的微重力水平最优的轨迹规划问题进行求解。仿真采用的计算机基本配置如下: CPU 为酷睿 i7-10875H, 主频为 2.30 G, 内存 8 G。

空间站运动仿真参数如表 2 所示, 其中包括航天器参数、轨道根数以及航天器所受摄动力相关的参数。仿真中, 空间站微重力加速度为实测数据。

表 2　航天器运动仿真参数

Table 2　Simulation parameters of spacecraft motion

参数	数值
太阳帆板面积/m²	230
核心舱行程中心位置${}^{(\mathrm{B})}\boldsymbol{r}_{BP}$/m	$[-8.5 \quad 1.8 \quad -2.6]^{\mathrm{T}}$
悬浮实验台内体质量/kg	30
轨道偏心率e	0
轨道高度h/km	400
轨道倾角i/(°)	42
升交点赤经Ω/(°)	120
近地点幅角ω/(°)	30
过近地点时刻t_{p}/s	25
航天器大气阻力系数C_{d}	3
大气密度	NRLMSISE 2 000
航天器太阳光压反射系数C_{P}	1.5
大气阻力偏心矢量${}^{(B)}\boldsymbol{L}_{\mathrm{drag}}$/m	$[1 \quad -3 \quad 2]^{\mathrm{T}}$
太阳辐射光压偏心矢量${}^{(B)}\boldsymbol{L}_{\mathrm{spr}}$/m	$[-2 \quad 1 \quad 3]^{\mathrm{T}}$

真实轨道采用四阶龙格库塔法进行动力学积分计算, 积分步长为 5 ms。凸优化问题离散化的时间间隔为 1 s。仿真中, PID 控制器的详细参数见表 3。

表 3　PID 控制器仿真参数

Table 3　Simulation parameters of PID controller

参数	数值
比例项系数$\boldsymbol{K}_{\mathrm{p}}$	$[2.7\times10^{-3} \quad 2.7\times10^{-3} \quad 2.7\times10^{-3}]$
积分项系数$\boldsymbol{K}_{\mathrm{i}}$	$[6.0\times10^{-6} \quad 6.0\times10^{-6} \quad 6.0\times10^{-6}]$
微分项系数$\boldsymbol{K}_{\mathrm{d}}$	$[2.4\times10^{-1} \quad 2.4\times10^{-1} \quad 2.4\times10^{-1}]$
控制器频率/Hz	1

仿真考虑的不确定性包括动力学模型不确定性、测量不确定性以及执行机构输出不确定性。其中, 动力学模型不确定性包括: 悬浮实验台内体质量不确定性、质心位置测量不确定性。不确定性采用以下形式表征:

$$X_{\mathrm{real}} = X_{\mathrm{measure}} + \Delta X_{\mathrm{const_error}} + \Delta X_{\mathrm{noise_error}} \tag{15}$$

式中: X_{real}、X_{measure}分别为变量X的真实值、测量值; $\Delta X_{\mathrm{const_error}}$、$\Delta X_{\mathrm{noise_error}}$分别为常值测量误差、测量噪声误差。

仿真中, 各种不确定性参数如表 4 所示。

表 4　不确定性参数

Table 4　Uncertain parameters

参数	数值
$\Delta m_{\mathrm{const_error}}$不确定性/kg	0.1
$\Delta \boldsymbol{r}_{\mathrm{PF_nosie_error}}$不确定性/mm	±1
$\Delta \boldsymbol{v}_{\mathrm{PF_noise_error}}$不确定性/(mm·s^{-1})	±0.1
$\Delta \boldsymbol{F}_{\mathrm{ctr_noise_error}}$不确定性/(m·N)	±0.01

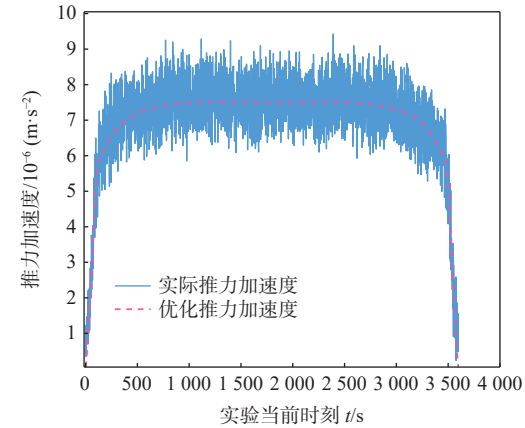
3.2　仿真结果与分析

当实验时长为 1 h 时, 采用本文提出的轨迹规划方案求解悬浮实验台内体最优的微重力水平, 凸优化的计算时长为 2.88 s, 仿真结果如图 6 所示。

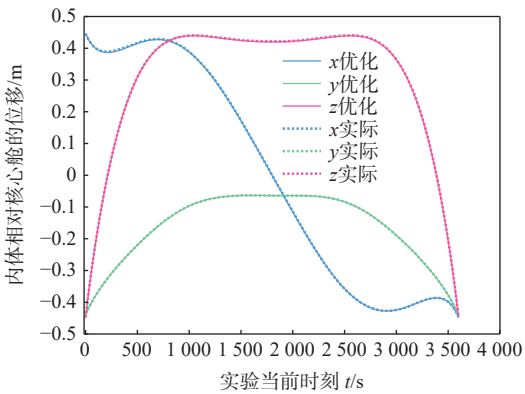
由图 6(a) 可知, 实际推力加速度最大值为 7.519×10^{-6} m/s^2。由于悬浮实验台内体所受的其他扰动加速度远小于控制力加速度, 因此优化微重力加速度近似于内体所受的控制力加速度, 最大微重力加速度为 9.121×10^{-6} m/s^2。由图 6(a) 还可知, 跟踪控制力加速度的大小相比优化加速度为小量, 很好地减少了对微重力水平的破坏。

图 6(b) 与图 6(c) 为悬浮实验台内体的实际运动轨迹与优化运动轨迹对比, 相对位移误差如图 7 所示。由图 7 可知, 本文采用的轨迹跟踪控制算法能够保证在 5 mm 左右的精度, 远小于为防止碰撞而设计的安全裕度参数, 保证了悬浮实验台的运行安全。仿真结果证明, 该轨迹跟踪算法合理可靠, 能够实现高精度、高效率的轨迹跟踪。

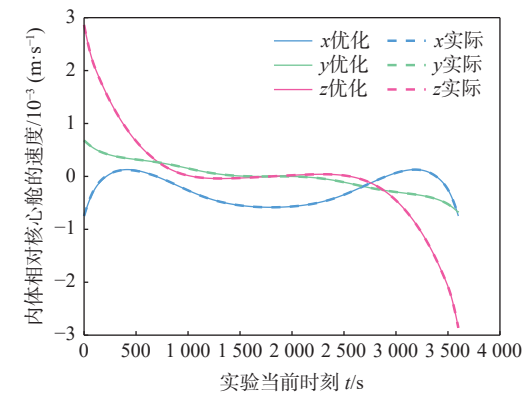
时域、频域下核心舱和悬浮实验台内体微重力



(a) 推力加速度



(b) 相对位移



(c) 相对速度

图 6 仿真结果

Fig. 6 Simulation result

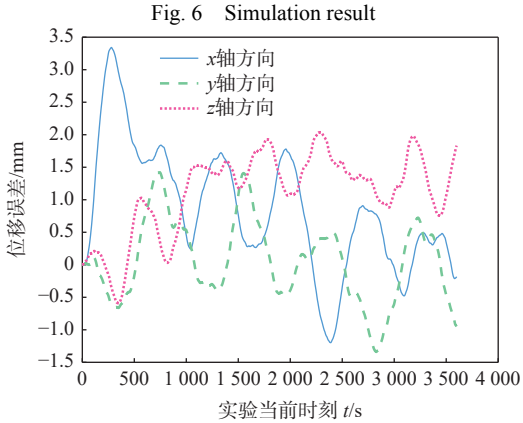
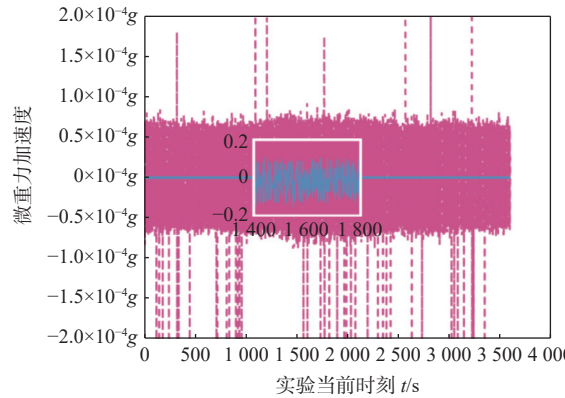


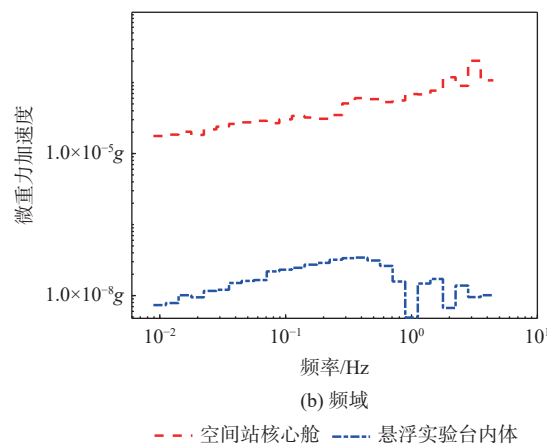
图 7 相对位移误差曲线

Fig. 7 Curve of relative displacement error

加速度如图 8 所示。由图 8 可知，核心舱在时域内的最大微重力加速度为 $2\ 135\times10^{-6}g$，内体的最大微重力加速度为 $0.9\times10^{-6}g$。频域中的微重力加速度用三分之一倍频程的均方根表示，证明本文所提出的轨迹规划与跟踪控制方案在时域与频域均可给悬浮实验台内体提供高微重力水平。



(a) 时域



(b) 频域

图 8 时域、频域下核心舱和悬浮实验台内体的微重力加速度

Fig. 8 Microgravity accelerations of core module and LEF inner body in time domain and frequency domain

4 结 论

针对空间站舱内悬浮实验台不同时间长度微重力水平最优的控制问题，本文提出了一种基于凸优化理论的轨迹规划与跟踪控制方法，该方法能够实现指定时长需求的最优微重力水平，可得到以下结论：

1）基于凸优化的轨迹规划方案考虑了轨道动力学、执行机构输出力上限、运动空间范围等约束条件，实现对微重力水平最优的飞行轨迹和控制序列的实时规划。

2）基于 PID 控制器的轨迹跟踪算法采用全新的控制器参数整定策略，能有效解决跟踪精度与高微重力水平之间的矛盾。当跟踪时长为 1 h 时，微重力水平最优运动轨迹的最大相对位移误差为

3.2 mm, 能保证悬浮实验台运行的安全性。

3）所提出的制导控制方法能够实时、高效规划并跟踪微重力水平最优的运动轨迹, 给悬浮实验台内体提供指定时长的最优微重力水平。在需求实验时长为 1 h 的条件下, 最优微重力水平为 9.121×10^{-6} m/s^2。

参考文献（References）

[1] 胡文瑞. 微重力科学概论[M]. 北京: 科学出版社, 2010.
HU W R. Introduction to microgravity science[M]. Beijing: Science Press, 2010(in Chinese).

[2] 吴季, 孙丽琳, 尤亮, 等. 2016—2030 年中国空间科学发展规划建议[J]. 中国科学院院刊, 2015, 30(6): 707-720.
WU J, SUN L L, YOU L, et al. Prospect for Chinese space science in 2016—2030[J]. Bulletin of Chinese Academy of Sciences, 2015, 30(6): 707-720(in Chinese).

[3] 陈善广, 陈金盾, 姜国华, 等. 我国载人航天成就与空间站建设[J]. 航天医学与医学工程, 2012, 25(6): 391-396.
CHEN S G, CHEN J D, JIANG G H, et al. Achievements of manned space program and construction of space station in China[J]. Space Medicine & Medical Engineering, 2012, 25(6): 391-396(in Chinese).

[4] 胡文瑞, 龙勉, 康琦, 等. 中国微重力流体科学的空间实验研究[J]. 科学通报, 2009, 54(18): 2615-2626.
HU W R, LONG M, KANG Q, et al. Research on space experiment of microgravity fluid science in China[J]. Chinese Science Bulletin, 2009, 54(18): 2615-2626(in Chinese).

[5] DELOMBARD R, HROVAT K, KELLY E, et al. Interpreting the international space station microgravity environment[C]// Proceedings of the 43rd AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2005.

[6] LIU W, GAO Y, DONG W B, et al. Flight test results of the microgravity active vibration isolation system in China's Tianzhou-1 mission[J]. Microgravity Science and Technology, 2018, 30(6): 995-1009.

[7] KLUMPP A R. Apollo lunar descent guidance[J]. Automatica (Journal of IFAC), 1974, 10(2): 133-146.

[8] 孙军伟, 崔平远. 月球软着陆多项式制导控制方法[J]. 宇航学报, 2007, 28(5): 1171-1174,1218.
SUN J W, CUI P Y. Polynomial guidance law for lunar soft landing[J]. Journal of Astronautics, 2007, 28(5): 1171-1174,1218(in Chinese).

[9] ACıKMESE B, CARSON J M, BLACKMORE L. Lossless convexification of nonconvex control bound and pointing constraints of the soft landing optimal control problem[J]. IEEE Transactions on Control Systems Technology, 2013, 21(6): 2104-2113.

[10] ACIKMESE B, PLOEN S R. Convex programming approach to powered descent guidance for Mars landing[J]. Journal of Guidance Control Dynamics, 2007, 30(5): 1353-1366.

[11] GERALDINK K T. People in control: Lars Blackmore[J]. IEEE Control Systems Magazine, 2016, 36(6): 24-26.

[12] 邓云山, 夏元清, 孙中奇, 等. 扰动环境下火星精确着陆自主轨迹规划方法[J]. 航空学报, 2021, 42(11): 524834.
DENG Y S, XIA Y Q, SUN Z Q, et al. Autonomous trajectory planning method for Mars precise landing in disturbed environment[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(11): 524834(in Chinese).

[13] 王劲博, 崔乃刚, 郭继峰, 等. 火箭返回着陆问题高精度快速轨迹优化算法[J]. 控制理论与应用, 2018, 35(3): 389-398.
WANG J B, CUI N G, GUO J F, et al. High precision rapid trajectory optimization algorithm for launch vehicle landing[J]. Control Theory & Applications, 2018, 35(3): 389-398(in Chinese).

[14] 中国载人航天办公室. 中国空间站科学实验资源手册[M]. 北京: 中国载人航天办公室, 2019.
China Manned Space Office. China's Space Station science experiment resource manual[M]. Beijing: China Manned Space Office, 2019(in Chinese).

[15] 刘伟. 空间高微重力悬浮平台控制技术研究[D]. 北京: 中国科学院大学, 2019.
LIU W. Research on control technology of space high microgravity suspension platform[D]. Beijing: University of Chinese Academy of Sciences, 2019(in Chinese).

Guidance and control method of levitation experiment facility inside China's Space Station

LU Lizheng[1, 2], LIU Wei[1, *], ZHANG Hao[1]

(1. Technology and Engineering Center for Space Utilization, Key Laboratory of Space Utilization, Chinese Academy of Sciences, Beijing 100094, China;
2. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: To optimize the microgravity level of the levitation experiment facility (LEF) inside China's Space Station (CSS), the limited intra-vehicular travel distance and various uncertainties were considered, so as to achieve the optimal microgravity level under different time scales. An autonomous trajectory planning method based on convex optimization and a trajectory tracking scheme based on error feedback control were proposed. Firstly, the structure of the LEF was introduced, and a spatial relative motion model was provided and linearized. Secondly, the motion trajectory and corresponding controlling force sequence with the optimal microgravity level were solved by convex optimization. Finally, the proportion-integration-differentiation (PID) tracking control law based on error feedback was designed to realize the fast tracking control of the optimized trajectory while guaranteeing the microgravity level. Numerical simulation results show that the optimal microgravity level can be 9.121×10^{-6} m/s^2 when the microgravity science experiment lasts for one hour. The autonomous trajectory planning method based on convex optimization and the trajectory tracking scheme based on error feedback control can achieve the optimal microgravity level for the specified duration. It provides a new solution to the current problem of the optimal microgravity level for a medium duration (such as one hour) and enhances the application value of the LEF in the CSS.

Keywords: convex optimization; levitation experiment facility; trajectory planning; relative motion; optimal microgravity level

Received: 2022-08-24; **Accepted**: 2023-01-17; **Published Online**: 2023-02-07 00:00
URL: link.cnki.net/urlid/11.2625.V.20230206.1635.005
Foundation item: National Natural Science Foundation of China (62003325)
* **Corresponding author**. E-mail: liuwei@csu.ac.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0745

面向区域覆盖的多无人机动态通信资源分配方法

卢毛毛[1,2]，刘春辉[3,*]，董赞亮[4]

(1. 西安电子科技大学 通信工程学院，西安 710071；2. 中国电子科技集团公司第五十四研究所，石家庄 050299；
3. 北京航空航天大学 无人系统研究院，北京 100191；4. 北京航空航天大学 电子信息工程学院，北京 100191)

摘 要：针对多无人机区域覆盖任务中的机间通信资源分配问题，提出了一种基于强化学习的多智能体动态通信资源分配模型。利用多智能体生成树覆盖方法生成任务区域内各个无人机的覆盖航线，对无人机与地面基站及无人机之间的通信链路进行建模。由于飞行环境的不确定性，将长期的资源分配问题建模为随机博弈模型，将无人机间的空-空链路视作一个智能体，每个智能体采取的动作包含选择工作频段和发送端的传输功率。在此基础上，基于双深度Q网络（DDQN）设计多智能体强化学习(MARL)模型，使得每个智能体通过奖励函数的反馈学习到最优通信资源分配策略。仿真结果表明：MARL模型能够在动态航迹下自适应选择最佳通信资源分配策略，提高时延约束下的负载交付成功率，同时降低空-空链路对空地下行链路的干扰并增大信道总容量。

关 键 词：多无人机区域覆盖；动态通信资源分配；强化学习；双深度Q网络；多智能体
中图分类号：V243.5; TN911.72
文献标志码：A **文章编号**：1001-5965(2024)09-2939-12

得益于无人机(unmanned aerial vehicle, UAV)高机动性、灵活部署、低成本和小型化等特点，其应用领域已逐渐由国防建设拓展到诸多民用新兴领域，如大气监测、森林火灾探测、交通控制、货物运输、紧急搜救、中继通信等[1]。其中，对地覆盖侦察任务是上述应用的重要组成部分，无人机在执行覆盖侦察任务时需要在指定区域内进行不间断的搜索以获取目标区域内的完整信息[2]。特别是针对全方位、大范围的复杂任务，由多架无人机组成的无人机集群系统展现了巨大潜力，成为国内外无人机应用领域的热门趋势[3-4]。然而随着空中交通的日益拥堵，无人机集群作为新的空中用户设备接入蜂窝网络仍面临许多挑战，主要包括：频谱资源高度竞争，高可靠性和严格服务质量(quality of service, QoS)要求，以及动态变化的信道环境等。因此，合理设计无人机集群内的通信资源分配机制，满足任务环境要求，对研究更加智能化的多无人机协同资源分配技术具有重要意义。

多无人机系统的通信资源分配技术旨在提高通信网络容量并增强网络覆盖范围[5-6]，现有研究工作包括：无人机作为中继缓存平台，以提高物联网设备吞吐量[7-9]；无人机作为辅助地面传感器的信息采集平台，以节约传感器传输能耗[10]；无人机作为中继通信平台，以满足地面移动用户体验质量(quality of experience, QoE)的同时提高系统能效[11-13]；无人机作为辅助移动边缘计算平台，以降低系统能耗[14-15]。然而大多数资源分配策略主要关注空-地信道，通过优化无人机的航迹为地面通信系统服务[8-10,12-13,15]，而针对动态航路下无人机集群网络的通信资源调度的研究工作较少。文献[16]基于启发式和贪婪式的图着色算法，从时域、频域、空域3个维度对大规模无人机网络的资源分配问

收稿日期：2022-08-30；**录用日期**：2022-11-04；**网络出版时间**：2022-12-02 10:22
网络出版地址：link.cnki.net/urlid/11.2625.V.20221201.1504.003
基金项目：2020年度科技创新2030——“新一代人工智能”重大项目(2020AAA0108200)
* **通信作者**. E-mail：liuchunhui2134@buaa.edu.cn

引用格式：卢毛毛，刘春辉，董赞亮. 面向区域覆盖的多无人机动态通信资源分配方法[J]. 北京航空航天大学学报，2024，50(9)：2939-2950. LU M M，LIU C H，DONG Z L. Dynamic communication resource allocation for multi-UAV area coverage[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2939-2950 (in Chinese).

题进行优化，然而其信道环境建模较为简单，没有考虑网络拓扑连续动态变化场景。文献[17]基于干扰消除的无人机集群实时频谱分配问题建模为势能博弈模型，利用并行最优反应算法对势能博弈求解，实现动态航迹下的多无人机频谱资源分配方案。然而其优化变量单一，缺少时延限制下传输可靠性等约束条件，且采用的数学解析法求解复杂，泛化能力较差。

传统的资源分配优化方法往往是将复杂的非凸优化问题分解为多个近似于凸优化的子问题，然后通过优化算法交替迭代求解一个次优方案[9,12-13,15]。随着人工智能技术的大规模应用，基于机器学习的资源分配算法展现了巨大潜力[18-19]，能够通过智能方式提供自主有效的解决方案，以增强无人机通信网络性能。如采用无监督学习框架解决大规模异构设备到设备（device-to-device, D2D）网络的功率控制问题[20]；采用深度强化学习算法针对通信、缓存和计算融合的未来车联网进行多维资源分配，以提高虚拟网络运营商的总效能[21]；利用基于高斯混合模型和加权期望最大化算法的机器学习框架来预测潜在的网络拥塞，以优化无人机的空中部署并用最小的功率满足用户下行通信需求[22]。然而上述工作的贡献集中在D2D网络、车联网以及无人机作为空中基站的部署和轨迹设计上，对于实际场景中的多无人机通信网络的动态资源分配策略鲜有研究。

在实践中，由于无人机的高速移动特性[23-24]，实时掌握无人机在动态飞行过程中的通信环境难度较高，这对设计可靠的多无人机无线通信网络带来了巨大的挑战。本文采用的深度强化学习算法（deep reinforcement learning, DRL）结合了深度学习的感知能力和强化学习的决策能力，以分布式方案视角实现无人机通信网络的智能资源管理，能够在无人机难以获取全局信道信息的情形下，对智能体之间的局部交互进行建模和求解，为复杂多无人机系统的感知决策问题提供了解决思路。

1 任务环境建模

1.1 区域覆盖航迹部署

无人机集群执行的区域覆盖任务是指多架无人机在飞行过程中，所搭载的侦察设备扫描地面区域直至逐渐覆盖全部任务区域的过程[25]。针对形状规则的目标区域，多无人机覆盖侦察常用区域分割[26]等方法，将目标区域分割为与无人机数目相等的子区域，然后在各分区内计算生成平行、直线的扫描航路[27]。因此，对任务区域的处理流程如图1所示，需要对整体的区域进行划分，无人机对划分后的每一个子区域进行扫描，直至所有无人机完成子区域全覆盖即实现最终任务目标。对于每个子区域由数量为$L_x \times L_y$个最小覆盖单元组成，表示为矩形$ABCD$，其中传感器视场在地面的投影近似为以无人机为圆心、半径R为监视距离的圆，A点到B点的长度表示无人机单位时间内的运动距离。

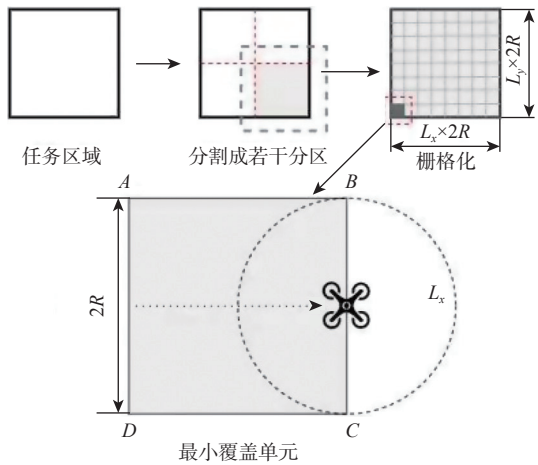


图1 任务区域处理流程

Fig. 1 Processing flow in task area

航迹规划属于动态通信资源分配训练环境中的关键要素，需要事先将其部署于强化学习的观测环境中，之后无人机按此进行航向的自动选择，并能根据相对位置自适应调节通信资源分配策略。这里选用多智能体生成树覆盖（multi-agent spanning tree coverage, MSTC）[28]航路规划方法离线生成覆盖航路。MSTC基于单个智能体的生成树覆盖（spanning tree coverage, STC）算法[29]，将无人机的运动方向离散化为前、后、左、右4个方向，具体规划流程如下：

步骤1 在目标区域用边长为$2R \times 2R$的正方形网格单元表示。

步骤2 对得到的网格中心点建立一棵或几棵生成树，用无向图$G=(V, E)$表示。其中，V为节点集，取自每个网格单元的中心点；E为边集，表示相邻网格单元中心的连线。覆盖所有连线则组成了G的一个最小生成树T。

步骤3 从任一中心节点V出发，按逆时针方向沿最小生成树T则形成了覆盖整个目标区域的STC路径L。

步骤4 根据k架无人机的初始位置将路径L划分成k段，每架无人机以分布式方式跟随分配的子路径沿逆时针方向环绕生成树一圈即可完成区域全覆盖任务。

1.2 通信环境建模

多无人机区域覆盖任务中的通信场景建模如

图 2 所示, 其中无人机对地的空-地链路(UAV to infrastructure, U2I)为高容量的正交信道, 无人机对无人机的空-空链路(UAV to UAV, U2U)复用 U2I 下行链路以提高频谱利用率。对于涉及的空-地链路和空-空链路, 3GPP 工作组基于已有城市/郊区环境中的测量和验证结果, 形成了 TR 36.777[30] 和 TR 38.901[31] 两类技术规范, 并结合 5G 应用场景, 进一步形成了无人机系统技术规范 TS 22.125[32]。基于以上标准文件[30-32], 设置基站天线高度为 25 m, 无人机高度为 100 m, 载波频率为 2 GHz, 基于市区宏基站(urban macro base station with aerial vehicle, UMa-AV)场景进行以下信道模型描述。

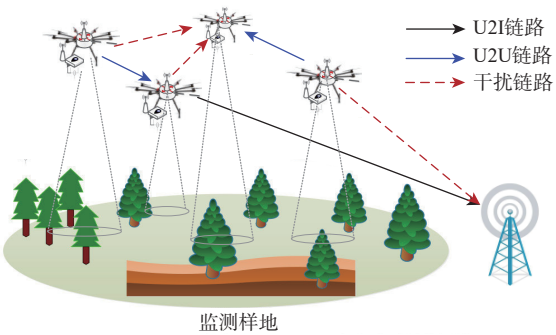


图 2　多无人机通信场景

Fig. 2　Multi-UAV communication scenarios

1.2.1　空地信道模型

根据无人机的高度和仰角不同, 其空地信道模型可分为直视空地信道模型和非直视空地信道模型。综合考虑无人机高度以及接收机和发射机相对位置动态变化的特点, 概率型直视、非直视空地信道模型更适用于无人机作为蜂窝网络的空中用户执行区域覆盖任务的场景。

依据 3GPP[30] 的统计模型, 无人机与地面基站间的空-地链路在正常情况下需要考虑路径损耗、阴影衰落与小尺度衰落的影响。其中概率型空地信道模型中的路径损耗L定义为

$$L = P_{\mathrm{LOS}}(L_{\mathrm{LOS}}) + P_{\mathrm{NLOS}}(L_{\mathrm{NLOS}}) \tag{1}$$

式中: P_{LOS}、P_{NLOS}分别为当前位置下的直视径概率、非直视径概率, 由 3GPP 文件[30] 中附录 B-1 公式计算得到; L_{LOS}、L_{NLOS}分别为直视径下的路损模型、非直视径下的路损模型, 由 3GPP 文件[30] 中附录 B-2 公式计算得到。

对于阴影衰落, 通常建模为服从对数正态分布的随机变量, 其均值为零, 标准差在不同场景下的取值不同。这里借鉴 3GPP 文件[30] 中附录 B-3 的统计模型计算直视径和非直视径下的阴影衰落标准差, 从而产生服从对数正态分布的直视径下阴影衰落Ω_{LOS}和非直视径下阴影衰落Ω_{NLOS}。

在无人机对地信道中, 高增益天线导致接收信号中存在一部分较强的直视分量, 因此将小尺度衰落建模为服从莱斯分布的随机变量。其中莱斯因子K参考文献 [33] 中的实际数据, 取值在 10～20 dB 之间, 则莱斯衰落可表示为

$$\Phi = 10\lg\sqrt{\frac{\left(x+\sqrt{2K}\right)^2 + y^2}{2(K+1)}} \tag{2}$$

式中: Φ为莱斯衰落, dB; x、y为均值分别是 0、方差是σ^2的高斯随机变量, 且K和σ^2的关系为$K = 10\lg(A^2/2\sigma^2)$, A为主信号的幅度峰值。

综上所述, 不考虑信道干扰等外在因素情况下, 空地信道模型依据直视径和非直视径概率可能受到的总干扰η为

$$\eta = P_{\mathrm{LOS}}(L_{\mathrm{LOS}} + \Omega_{\mathrm{LOS}} + \Phi) + P_{\mathrm{NLOS}}(L_{\mathrm{NLOS}} + \Omega_{\mathrm{NLOS}} + \Phi) \tag{3}$$

1.2.2　空空信道模型

由于空中往往不存在遮挡直视径的障碍物, 因此常将空空信道建模为直视型链路[34-37], 且文献 [34] 认为当无人机高度较低时, 可忽略大尺度阴影衰落的影响, 因此信道衰落主要来源于路径损耗与小尺度阴影衰落。

对于路径损耗, 实测中空空信道的路径损耗大多数情况下接近于自由空间衰落模型[36-37]

$$\overline{L} = 32.45 + 10\overline{n}\lg d + 20\lg f_{\mathrm{c}} \tag{4}$$

式中: $\overline{L}$为空-空链路路径损耗, dB; d为两点通信的直线距离, km; f_{c}为载波频率, MHz; $\overline{n}$为路径损耗指数, 自由空间模型中$\overline{n}=2$。

对于小尺度阴影衰落, 通常将直视型链路建模为服从莱斯分布的随机变量$\overline{\Phi}$。根据文献 [34] 的测量结论, 无人机在低空场景中, 空空信道的莱斯因子K取值在 3.533～10.048 dB 之间, 且K随着无人机高度的升高而增大, 因此本文按照一般情况取K=10 dB。

1.3　无人机通信资源分配

对于多无人机通信场景, 包括 U2I 链路和 U2U 链路。U2I 链路可以利用蜂窝网络的接口连接无人机与地面基站进行高数据速率传输, 由于基站没有关于 U2U 链路的信息, 因此 U2I 链路的通信资源管理应独立于其资源分配过程。于是将蜂窝网络分配给无人机的频段离散化后, 提前为每条 U2I 链路分配固定传输功率的正交频谱, 因此主要任务是针对 U2U 链路设计高效的频谱共享策略, 包括选择频谱子带和控制合适的发送功率, 使得 U2I 和 U2U 链路在无人机高速移动环境下, 以最小的信令开销实现各自的服务需求。

考虑如图2所示的在离散时间轴中运行的多无人机链路，其中包括m架单天线无人机与基站形成U2I链路，链路集合表示为$M_{\mathrm{U2I}}=\{1,\cdots,m\}$，以及无人机与相邻无人机形成的$k$对U2U链路，链路集合表示为$K_{\mathrm{U2U}}=\{1,\cdots,k\}$，无人机可操作的总带宽被划分为$m$条正交子信道分配给U2I链路。此外，假设所有无人机在没有中央控制器的协助下进行通信，并且不具备无线通信环境的全局信息，即U2U链路的信道状态信息仅在本地是已知的。基于无人机的机动性，该假设在实际中是合理的[23-24]。

由于基站处的干扰更可控，为了充分利用可用资源，U2U链路共享U2I下行链路频谱，并且每对U2U链路在当前时隙只能占用一条子信道，但是不同的U2U链路占用的子信道可能相互重叠。因此U2I链路受到的干扰由背景噪声和同一子信道上运行的其他无人机产生。假设第k和k'条U2U链路都共享第m条子带，则第m条U2I链路可能受到的干扰来自第k和k'对U2U链路的发射端到基站信道的干扰。

由上述分析可知，第m条U2I链路的信干噪比$\gamma_m^i[m]$为

$$\gamma_m^i[m]=\frac{P_m^i\tilde{g}_{m,B}[m]}{n^2+\sum_{k\in K}\rho_k[m]P_k^u[m]g_{k,B}[m]} \tag{5}$$

式中：$m\in K_{\mathrm{U2I}}$；P_m^i为U2I链路中无人机m到基站的传输功率；$\tilde{g}_{m,B}[m]$为第m条U2I链路的信道增益；$\rho_k[m]$为子带复用因子，$\rho_k[m]=1$表示第k对U2U链路复用第m条U2I链路子信道，不复用则$\rho_k[m]=0$；$P_k^u[m]$为第m条子带上第k对U2U链路的传输功率；$g_{k,B}[m]$为第k对U2U链路发送端到基站B间链路的信道增益，且第k对U2U链路共享第m条子带；n^2为噪声功率，由式(3)中总干扰η和随机白噪声组成。

同理，当第k和k'对U2U链路共享第m条子带时，第k对U2U链路可能受到的干扰源有：来自第m条U2I链路的发射端到第k对U2U链路的接收端信道的干扰，将该链路的信道增益表示为$\tilde{g}_{m,k}[m]$，以及来自第k'对U2U链路的发射端到第k对U2U链路的接收端信道的干扰，将该链路的信道增益表示为$g_{k',k}[m]$。

因此，占用第m条子信道的第k对U2U链路的信干噪比$\gamma_k^u[m]$为

$$\gamma_k^u[m]=\frac{P_k^u[m]g_k[m]}{\overline{\eta}^2+I_k[m]} \tag{6}$$

式中：$g_k[m]$为第m条子信道上第k对U2U链路的信道增益；$\overline{\eta}^2$为空-空链路的背景噪声，包含$\overline{L}$、$\overline{\Phi}$以及随机白噪声；$I_k[m]$为第m条子信道上第k对U2U链路受到的所有干扰之和，计算式为

$$I_k[m]=P_m^i\tilde{g}_{m,k}[m]+\sum_{k'\neq k,\,k\in K_{\mathrm{U2U}}}\rho_{k'}[m]P_{k'}^u[m]g_{k',k}[m] \tag{7}$$

式中：用$\rho_{k'}[m]=1$来标记第k'对U2U链路也同时选择了第m条子信道，否则为0，且第k'对U2U链路的传输功率为$P_{k'}^u[m]$。

根据香农定理，结合式(5)和式(6)，则第m条U2I链路的信道容量$C_m^i[m]$、占用第m条子信道的第k对U2U链路的信道容量$C_k^u[m]$分别为

$$C_m^i[m]=W\log_2(1+\gamma_m^i[m]) \tag{8}$$

$$C_k^u[m]=W\log_2(1+\gamma_k^u[m]) \tag{9}$$

式中：W为带宽。

由上述分析可知，带宽一定的条件下，如果要增大U2U链路的吞吐量以保证信息的可靠传输，当提高发送端的传输功率$P_k^u[m]$时，会导致对所占用的U2I链路的干扰增大，且暂且没有考虑对同一子信道上其他U2U链路的干扰。因此对无人机集群系统通信资源分配方案的设计，关键在于如何权衡U2U链路负载的可靠交付以及对U2I链路吞吐量的影响。这也是设计强化学习奖励函数的关键指标，以此引导智能体学习到全局最优决策。

2 基于强化学习的通信资源分配

2.1 多无人机通信资源分配的随机博弈模型

强化学习(reinforcement learning, RL)是机器学习的一个重要分支，被广泛用于解决马尔科夫决策过程(Markov decision process, MDP)问题。对多智能体强化学习而言，其本质是一个随机博弈过程[38]，其中，每个智能体都满足马尔科夫特性，即智能体的奖励仅取决于当前状态和动作[39]。本文将每对U2U链路视为一个智能体，因此需要根据环境条件、所优化的决策、优化目标来设计状态集、动作集以及奖励函数。

对于状态集中的变量，表示智能体能获知的信息，通过这些信息计算得到下一步决策动作。这里设计智能体k在每个时隙t的状态集s_k^t为

$$\mathrm{s}_k^t=\left\{g_k[m],\tilde{g}_{m,B}[m],N_{t-1},I_{t-1},L_t,U_t\right\} \tag{10}$$

式中：$m\in M_{\mathrm{U2I}}$；$g_k[m]$为该智能体k的信道增益，且选择的是子信道m；$\tilde{g}_{m,B}[m]$为该智能体k所占用的第m条U2I链路的信道增益；N_{t-1}、I_{t-1}分别为智能体前一时隙选择的子信道、受到的总干扰；L_t、U_t分别为智能体当前的剩余负载、剩余传输时间。

对于动作集，则是强化学习算法需要优化的决策，其直接影响了通信资源分配的最终结果。因

此, 根据 U2U 链路传输的可靠性要求以及 U2I 链路的高容量要求, 为智能体k建立包含子信道选择策略及功率选择策略$p_k(t)$的二维动作集a_k^t, 计算式为

$$a_k^t = \{\overline{\rho}_k(t), p_k(t)\} \tag{11}$$

式中: $\overline{\rho}_k(t)$为智能体k在时隙t所占用的子信道标号$\overline{\rho}_k(t) \in M_{\text{U2I}}$; $p_k(t)$为智能体k在时隙t选择的发送功率$p_k(t) \in \{P_1, \cdots, P_n\}$, 其中$P_1, \cdots, P_n$为$n$个级别的备选功率。

由此可见, 每个智能体采取的动作都会影响整个无人机系统的通信性能。因此, 需要定义一个函数来评判每个智能体的动作对所学目标的正负影响, 这个函数被称为奖励函数。对于本文的多智能体通信资源分配问题, 学习目标在于多无人机能协同决策以提高 U2U 链路传输可靠性, 同时减少对 U2I 链路的干扰。对于 U2I 链路, 信道干扰的减少表现为信道总容量增大, 因此结合式 (8) 中信道容量$C_m^i[m]$, 奖励函数的组成部分r_1为

$$r_1 = \sum_{m \in M_{\text{U2I}}} C_m^i[m] \tag{12}$$

对于 U2U 链路, 其传输可靠性具体体现为有限时延内负载的完全传输, 因此奖励函数还包含了所有智能体剩余负载的大小r_2, 结合式 (9) 可表示为

$$r_2 = \sum_{k \in M_{\text{U2U}}} \left(D - \hat{t} C_k^u[m]\right) \tag{13}$$

式中: D为传输时延下的负载; $\hat{t} = 1$ ms为计算时间间隔。

综合式 (12) 和式 (13), 并加权$\lambda \in [0,1)$后得到每个时间步t的奖励函数R_{t+1}为

$$R_{t+1} = \lambda r_1 - (1-\lambda) r_2 \tag{14}$$

但是式 (14) 仅代表了当前动作带来的即时奖励, 为了预估该动作的长期影响, 强化学习定义折扣累计奖赏(也称为 Q 函数[40])来衡量当前状态和动作下的长期收益。当初始状态为s_0、初始动作为a_0时, Q 函数的计算式为

$$Q(s,a) = E\left[\sum_{t=0}^{+\infty} \delta^t R_{t+1} \,\middle|\, s_0 = s, a_0 = a\right] \tag{15}$$

式中: $Q(s,a)$为以s和a为变量的 Q 函数, s为所有智能体在当前时刻的状态集合, $s = \{s_1^t, s_2^t, \cdots, s_k^t\}$, a为所有智能体在当前状态下采取的动作策略, $a = \{a_1^t, a_2^t, \cdots, a_k^t\}$; $\delta \in [0,1)$为折扣率, δ取值与未来奖励对最优决策的影响程度有关, δ值越接近 0, 表示决策更看重近期收益, 反之越接近 1, 则说明未来奖励占有更大权重, 需要做出更富有远见的决策; $E[\cdot]$为均值函数。

本文设定所有智能体共享同一奖励函数R_{t+1}, 且经理论[40]证明, 式 (15) 由于折扣率δ的存在, 当迭代达到一定步数后, Q 函数将趋于一个恒定值, 为最终期望收益。因此, 每个智能体的学习目标便是预估一个全局最优的决策, 使得式 (15) 中的折扣累计奖赏最大。

2.2　DDQN 网络模型

DDQN 是深度 Q 网络(deep Q-network, DQN)的改进版本, 核心思想是结合深度神经网络来拟合 Q 函数的迭代计算过程。DDQN 具体算法流程如图 3 所示, 其中包含一个全连接的当前值神经网络 Q, 用于输入每架无人机的状态, 输出估计的 Q 值, 并为目标值网络$\hat{\text{Q}}$提供最大 Q 值对应的动作。而目标值神经网络$\hat{\text{Q}}$用于固定神经网络参数, 神经网络结构与当前值神经网络相同, 输出损失函数中的目标 Q 值。DDQN 中的损失函数定义为

$$L = \left(y_{\text{DDQN}} - Q(s,a;\theta)\right)^2 \tag{16}$$

式中: θ用于标记当前值网络参数; y_{DDQN}为目标 Q 值, 计算式为

$$y_{\text{DDQN}} = R_{t+1} + \delta \hat{Q}\left(s_{t+1}, \text{argmax} a_{t+1} Q(s_{t+1}, a_{t+1}; \theta); \hat{\theta}\right) \tag{17}$$

式中: $\hat{\theta}$为目标值神经网络参数, 而回放记忆库的作用则是为当前值神经网络和目标值网络提供独立同分布的训练样本。

2.3　多智能体 DDQN 模型

对于无人机集群区域覆盖任务, 要求每架无人机在地理上既可以分散、独立地执行任务, 同时又可以相互通信, 互相协调, 共同决策通信资源分配方案。因此, 结合 1.1 节中离线部署的覆盖航迹结果, 本节主要研究航迹确定情况下, 无人机集群的通信资源分配网络。

传统的多智能体强化学习的训练模式主要有集中式学习和独立式学习 2 种, 集中式学习是指所有智能体都把观测信息传送给中央控制器, 中央控制器指导所有智能体的动作及奖励, 智能体本身不做决策, 仅仅是动作的执行者。独立式学习是指每个智能体作为一个独立个体, 独立地和环境交互, 并且用自己的观测和奖励来更新自己的策略, 各个智能体之间不互通信息。

然而集中学习方式在处理大规模多智能体环境时效率低下, 并且一旦某个智能体学到有用的策略后, 其他智能体会选择懒惰的策略避免阻碍已经学会的智能体, 从而导致全局奖励下降。而独立学习方式会由于单个智能体的学习破坏其他智能体环境的平稳性[41]。

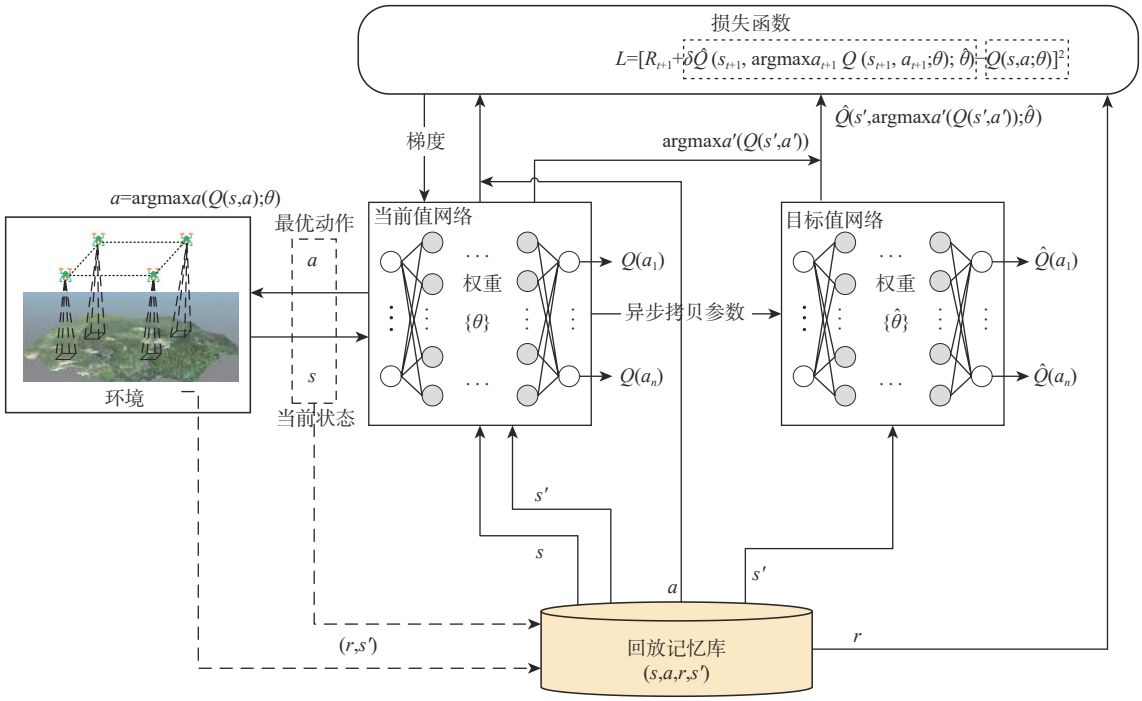


图3 DDQN算法流程

Fig. 3 Processing flow of the DDQN algorithm

本文基于DDQN网络改进独立式学习的训练方式，以解决环境非平稳性问题。多智能体DDQN（multi-agent DDQN, MDDQN）训练流程如图4所示，其中所有智能体共用一套DDQN网络结构，但是为每个智能体训练不同的网络参数。在每一时间步，智能体首先同步输入当前的观测状态，用A_t代表网络输出的所有智能体动作的集合，然后同步更新每架无人机的位置坐标和信道状态信息，据此计算动作集合A_t所能得到的奖励值R_{t+1}。MDDQN中，由于所有智能体的状态是同步改变的，并且奖励函数依据所有智能体的联合动作，因此对于单个智能体，其在同一时间步内的观测状态是相对稳定的，避免智能体陷入无止境的探索中。最后将每个智能体的当前状态、执行动作、得到的奖励值与执行动作后的新状态一同作为一条训练样本存入回访记忆库，之后从中采样的经验才能作为有用的经验，从而引导独立的智能体学习到“合作式”的策略。

由于训练时奖励函数会依据所有智能体的状态和动作给出一个整体的评判，因此MDDQN以一种“合作”学习的方式进行训练，但是训练好的模型可分布式实现于每架无人机上，根据无人机接收的观测数据，直接输出估计的最优决策，实现分布式部署。

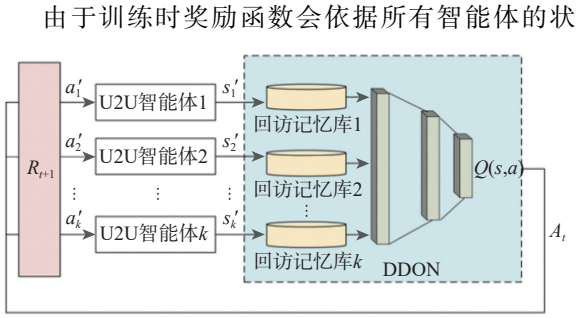


图4 MDDQN训练流程

Fig. 4 Training process of the MDDQN algorithm

3 仿真分析

3.1 离线部署覆盖航迹

为了能使多无人机在动态航迹下自适应决策通信资源分配策略，且结合区域覆盖的任务要求，因此需要将无人机的飞行轨迹作为环境的一部分进行离线部署。无人机在每个时间步沿部署好的航迹飞行，同时多无人机间相对距离的改变也影响了信道衰落模型，在此基础上训练动态航迹下的通信资源分配策略，使得多无人机系统既能满足区域覆盖任务要求，又能提高整体通信性能。

本文针对4架无人机的区域覆盖任务，利用1.1节中MSTC航路规划方法[28]，生成航迹如图5所示。其中，任务区域设置为400 m×400 m的规则矩形，并分成不同颜色的4个子区域，对应4架无人机。无人机随机初始化起始点，并用红色实心圆表示，生成的轨迹可保证对目标区域的完全覆盖，且无重复区域。

3.2 基于MDDQN的通信资源分配

依据图3的DDQN算法流程和图4的MDDQN训练流程，结合图5中4架无人机的覆盖航迹，进行多智能体通信资源分配网络的训练，伪代码见算

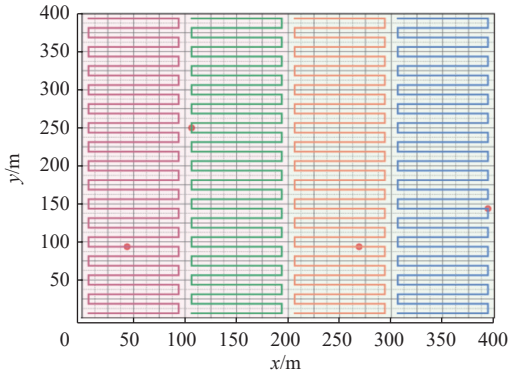


图 5　MSTC 航迹

Fig. 5　Track map of the MSTC algorithm

法 1。为了探索到更多的状态-动作空间以避免陷入局部最优, 这里对网络预测的动作采用衰减 ε-greedy 策略执行, 用1-ε值代表选择最大 Q 值对应动作的概率, 否则以ε概率随机选取动作, 随着训练轮数的增加, 智能体的经验越来越丰富, ε也随之减小。最后, 智能体共享同一奖励函数来评价整体决策的好坏, 则累计折扣奖励(Q 值)收敛时, 可认为所有智能体的联结策略达到纳什均衡。

算法 1

1: 初始化 MSTC 覆盖航迹和无人机运动步数 s、实验中无人机数目 n_{w}

2: 建立区域覆盖动态环境模型, 初始化信道模型

3: 随机初始化网络参数 $\widehat{\theta}=\theta$

4: for 周期 $e=1:e^{\max}$

5: 更新无人机位置和大尺度衰落

6: 初始化所有智能体的传输负载 $L_0=L$和传输时延 $U_0=T$

7: for $s=1:N$

8: for U2U $n_{\mathrm{w}}=1:k$

9: 获取观测状态集合 s_k^t

10: 按照衰减ε-greedy 策略选择动作a_k^t

11: end for

12: 更新小尺度信道衰落

13: 同时更新所有智能体动作

14: 计算奖励函数R_{t+1}

15: for U2U $n_{\mathrm{w}}=1:k$

16: 获取下一时刻状态集 s_k^{t+1}

17: 将经验 $\left[s_k^t,a_k^t,R_{t+1},s_k^{t+1}\right]$作为训练样本存入记忆库

18: end for

19: end for

20: for U2U $n_{\mathrm{w}}=1:k$

21: 从记忆库随机均匀采样 B大小的训练样本

22: RMSprop 优化式 (16) 定义的损失函数

23: if step % $c==c-1$

24: 复制当前值网络参数给目标值网络$\widehat{\theta}=\theta$

25: end if

26: end for

27: end for

实验的主要环境参数及信道参数依据 3GPP TR 36.777[30] 中的 UMa-AV 场景设置, 如表 1 所示。DDQN 网络参数如表 2 所示。

表 1　环境参数

Table 1　1Environmental parameters

参数	数值
U2U智能体数目k/个	4
带宽W/MHz	4
载频f_{c}/GHz	2
基站天线高度/m	25
无人机高度/m	100
无人机速度/(m·s^{-1})	[16,20]
传输时延T/ms	100
传输负载L/Byte	2×1 060
U2U传输功率等级/dBm	[23,10,5,−100]
U2I传输功率/dBm	23
覆盖区域大小/(m×m)	400×400

表 2　DDQN 网络参数

Table 2　2Parameters of the DDQN network

参数	数值
全连接层数/层	3
每层神经元个数/个	[500,250,120]
最小贪婪系数ε	0.02
折扣率δ	1.0
训练轮数$e^{\max}$/轮	3 000
每轮训练步数N/步	100
记忆库容量/条	200 000
采样大小B/条	100
更新目标值网络步数c/步	400

图 6 为所有智能体损失函数随训练轮数的变化情况。由图 6 可知, 4 个智能体 DDQN 网络的损失函数在 600 轮之后便全部收敛, 说明网络模型基本拟合了 Q 值函数, 但是并不能代表策略最优。图 7 为所有智能体在一轮训练中的平均奖励随训练轮数的变化。由图 7 可知, 随着训练轮数增长, MDDQN 模型在 2 000 轮左右达到收敛, 说明多无人机系统已学习到近似最优的策略(轻微震荡是由于无人机继续移动导致信道衰落变化), 之后的通信资源分配实验结果都是基于此 MDDQN 模型进行测试。

3.2.1　U2I 信道总容量

本研究内容将训练得到的多智能体深度强化学习网络模型(MDDQN)与 2 种方法进行了对比: 1) 独立学习的 DDQN 模型。文献 [41] 中以独立 Q 学习 (independent Q-learning, IQL) 方式训练的网络, 为了保证算法的一致性, 将原文中的 DQN 强化

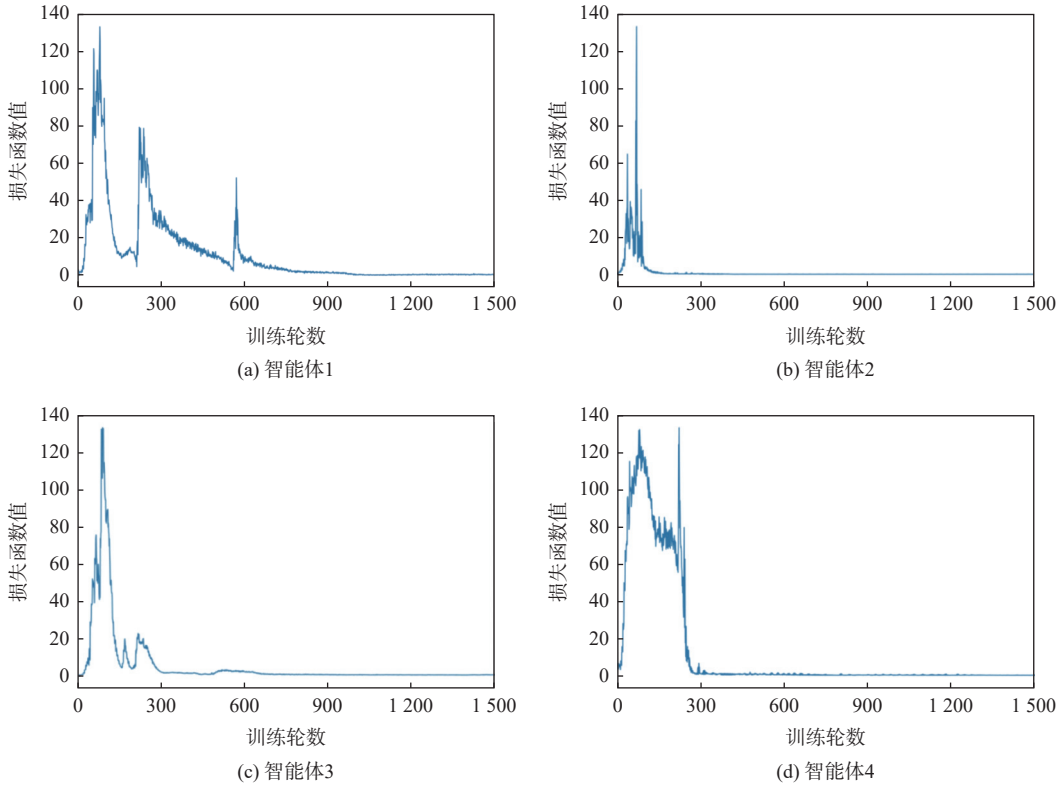


图 6 损失函数随训练轮数变化情况

Fig. 6 Change of loss with the training episodes

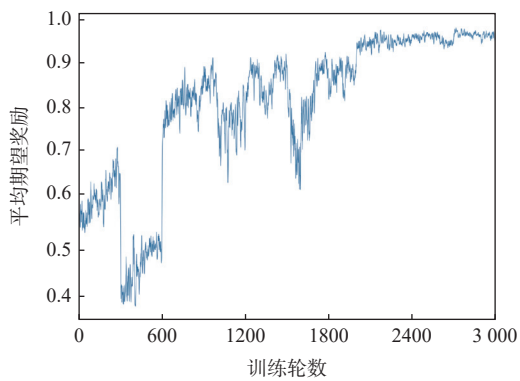


图 7 每轮平均奖赏的变化

Fig. 7 Changes of average reward per episode

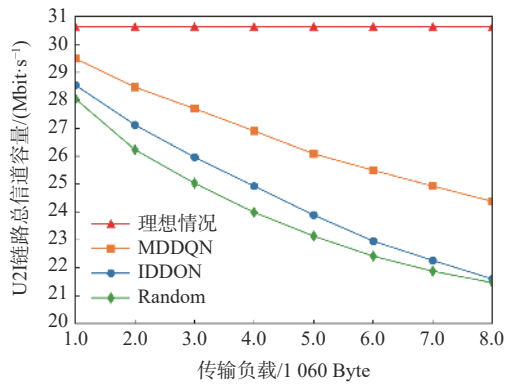


图 8 不同方法下的 U2I 信道总容量随传输负载变化情况

Fig. 8 Variation of the total U2I channel capacity with the transmission load for different methods

学习算法替换为本文中的 DDQN 算法，得到 IDDQN 模型。与 MDDQN 不同的是，IDDQN 下每个智能体都是在解决一个单智能体任务，其他智能体都被视为环境的一部分。一个智能体在获得预测动作后会立即更新状态和奖励函数，而另一架无人机的观测状态会随着前一架无人机状态的改变而改变，因此观测状态的不稳定导致该模型较难收敛。2）随机分配模型（Random）。其中每个智能体都采用随机选择传输功率和频谱子带的策略。

不同方法下的 U2I 信道总容量随传输负载变化情况如图 8 所示，图 8 中红色的基准线代表没有 U2U 信息传输时，U2I 链路受到 U2U 链路的信道干扰为 0 dB，则以该信道容量作为理想情况下的最大信道容量。由图 8 可知，随着传输负载的增大，多智能体 DDQN 模型（MDDQN）、独立学习的 DDQN 模型（IDDQN）以及随机分配模型（Random）策略下的 U2I 总信道容量都呈下降趋势，原因是负载的增大会导致传输时间延长，使智能体偏向于选择更大的传输功率以提高 U2U 传输成功率，从而造成对 U2I 链路的干扰增大。但即使如此，MDDQN 策略下的总信道容量仍然高于另外 2 种方法。值得注意的是，虽然训练时是以 2×1 060 Byte 大小作为传输负载，但测试时分别测试了负载 $L=(1,2,\cdots,8)\times$ 1 060 Bytes 下的表现情况，随着 U2U 传输负载的增大，MDDQN 的泛化性能也优于另外 2 种方法。

3.2.2 U2U 传输成功率

传输成功率指的是负载能在一轮（100 ms）内传

完的次数与总轮数的比值, 该指标反映出智能体每一时间步(1 ms)选择的发送功率是否合理可靠。图 9 为不同方法下 U2U 传输成功率随传输负载变化的情况。由图 9 可知, 虽然随着负载的增加, 传输成功率都呈下降趋势, 但是本文提出的 MDDQN 模型在负载为 4×1 060 Byte 之内都保持了较好的鲁棒性, 不仅传输成功率高于另外 2 种方法, 且负载为 2×1 060 Byte 以内时, 几乎全部成功。之后下降较明显的原因可能在于只有大功率才能满足负载的完全传输, 然而若所有无人机都选择最大功率传输势必会造成对 U2I 链路的恶劣影响, 因此智能体应采取折中策略, 通过牺牲传输成功率来维护整体信道容量。

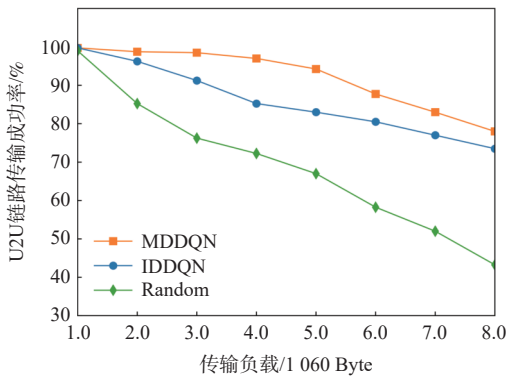


图 9　不同方法下 U2U 传输成功率随传输负载变化情况

Fig. 9　Variation of U2U transmission success rate with transmission load for different methods

3.2.3　功率分配策略

智能体自主选择不同传输功率概率随剩余传输时间变化情况(100 ms 时延)如图 10 所示, 横坐标从右至左代表剩余传输时间逐渐减少。由图 10 可知, 当剩余传输时间多于 83 ms 时, 大部分智能体倾向于选择以小功率 5 dB 传输, 因为时间充裕且能减少信道间干扰。之后随着剩余时间的减少, 智能体优先选择最大功率 23 dB 传输的概率增加; 当剩余传输时间不足 67 ms 时, 选择功率为 23 dB 的概率开始下降; 而选择中等功率 10 dB 的概率持续增加, 原因可能是智能体判断没必要一直以最大功率传输, 这样反而会增加对其他链路的干扰; 当剩余时间不足 10 ms 时, 已经有智能体完成传输任务, 而没有完成的智能体更多地选择以 10 dB 的功率继续传输。

3.2.4　剩余负载变化情况

MDDQN 策略下 4 个智能体剩余负载随时间变化情况(随机选择两轮结果)如图 11 所示。由图 11 可知, 在该策略下智能体依次完成传输任务, 一条链路最短在 10 ms 左右便能传输成功, 4 个智能体整体传输完成时间不超过 40 ms。

IDDQN 策略下 4 个智能体剩余负载随时间变化情况(随机选择两轮结果)如图 12 所示。由图 12 可知, 该策略下单条 U2U 链路最快完成传输的时间大于 10 ms 但不超过 20 ms, 整体传输完成时长大于 MDDQN 策略。可能原因在于智能体 4 的观测状态受到其他智能体动作影响, 导致其难以收敛到一个稳定的策略, 进而影响了无人机集群整体的链路传输时间, 因此没达到多智能体协同最优策略。

随机策略(Random)下 4 个智能体剩余负载随时间变化的情况(随机选择两轮结果)如图 13 所

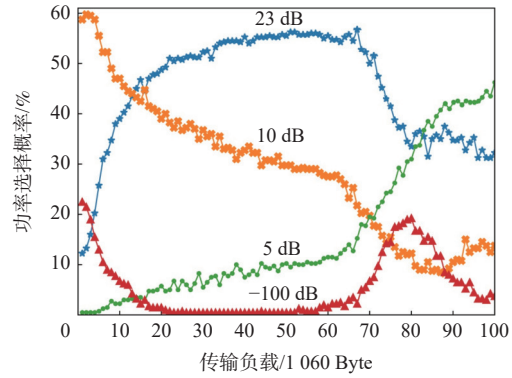
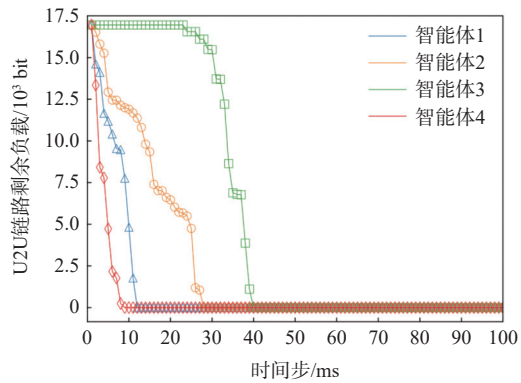
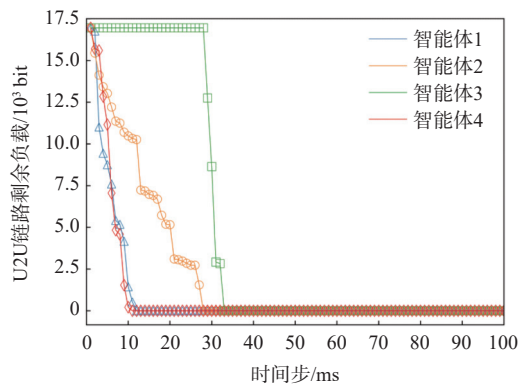


图 10　选择传输功率概率随剩余传输时间变化情况

Fig. 10　Changes of transmission power probability with remaining transmission time



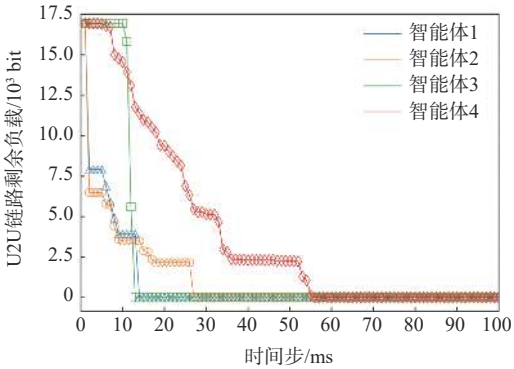
(a) 第1轮结果



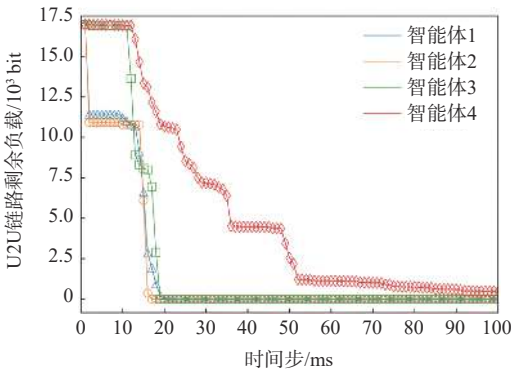
(b) 第2轮结果

图 11　MDDQN 模型下智能体剩余负载随时间变化情况

Fig. 11　Changes of agent residual load with time for the MDDQN model



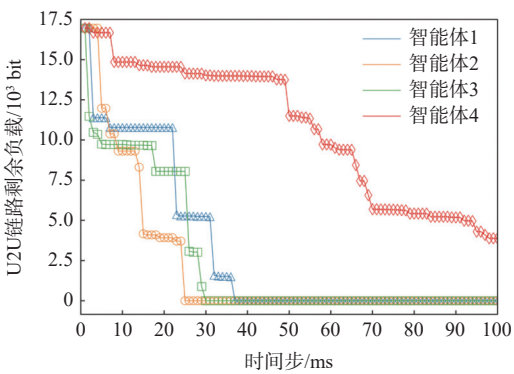
(a) 第1轮结果



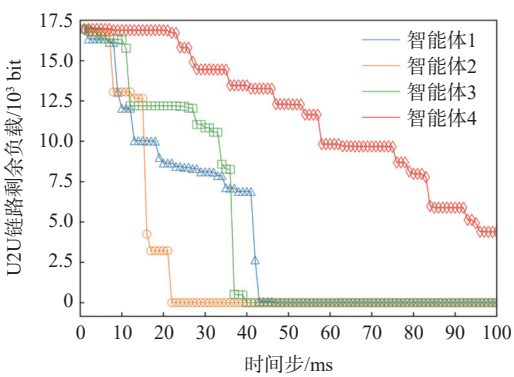
(b) 第2轮结果

图 12 IDDQN 模型下智能体剩余负载随时间变化情况

Fig. 12 Changes of agent residual load with time for the IDDQN model



(a) 第1轮结果



(b) 第2轮结果

图 13 随机资源分配下智能体剩余负载随时间变化情况

Fig. 13 Variation of residual load of agents with time for the random resource allocation

示。由图 13 可知，随机分配频谱和发送功率的情况下，不能保证无人机集群所有链路的成功传输，图中仅有 3 条链路能成功完成传输任务，且最早传输成功的智能体所需时间也超过 20 ms。

4 结 论

本文基于 DDQN 深度强化学习框架研究了区域覆盖任务场景下，多无人机间的通信资源分配策略优化问题，可得出以下结论：

1）通过解耦覆盖路径规划问题与通信资源分配问题，实现了多无人机动态飞行过程中的通信资源分配策略的优化。

2）随机博弈模型定义了多智能体的观测状态与动作集合，设计的奖励函数以整体通信性能为导向，有效指导了多无人机分布式状态下，达成一种“合作式”策略，通过最大化长期收益，提高了整体信道容量及链路传输可靠性。

3）将单智能体强化学习算法 DDQN 扩展用于多智能体场景，并对其训练方式进行改进，仿真验证了改进后的 MDDQN 策略下能获得更高的信道容量与传输成功率，并在特定负载范围内保持了一定的鲁棒性。

4）从功率分配及剩余负载变化情况的角度展现了各个无人机依据不同信道状态作出的智能决策，体现出 MDDQN 模型下的自适应决策能力。

本文采用的 DDQN 算法尚不适用于处理连续动作，针对实际场景中连续的功率与频谱选择问题，在后续研究中可通过增加动作空间维度或采用新的强化学习算法加以解决。

参考文献（References）

[1] HANSCOM A F B, BEDFORD M A. Unmanned aircraft system (UAS) service demand 2015-2035: Literature review & projections of future usage[R]. Washington: Department of Transportation, 2013.

[2] 李艳庆. 基于遗传算法和深度强化学习的多无人机协同区域监视的航路规划[D]. 西安: 西安电子科技大学, 2018.
LI Y Q. Cooperative path planning for region surveillance of multi-UAV based on genetic algorithm and deep reinforcement learning [D]. Xi'an: Xidian University, 2018(in Chinese).

[3] 赵林, 张宇飞, 姚明昕, 等. 无人机集群协同技术发展与展望[J]. 无线电工程, 2021, 51(8): 823-828.
ZHAO L, ZHANG Y F, YAO M W, et al. Development and trend of UAV swarm cooperative techniques[J]. Radio Engineering, 2021, 51(8): 823-828(in Chinese).

[4] 刘立辉, 赵彦杰, 赵小虎, 等. 一种无人集群系统仿真平台设计[J]. 中国电子科学研究院学报, 2017, 12(5): 506-512.
LIU L H, ZHAO Y J, ZHAO X H, et al. Design of a simulation platform of unmanned swarm system[J]. Journal of China Academy

of Electronics and Information Technology, 2017, 12(5): 506-512(in Chinese).

[5] BUCAILLE I, HÉTHUIN S, MUNARI A, et al. Rapidly deployable network for tactical applications: aerial base station with opportunistic links for unattended and temporary events ABSOLUTE example[C]//2013 IEEE Military Communications Conference. Piscataway: IEEE Press, 2013: 1116-1120.

[6] XIAO Z Y, XIA P F, XIA X G. Enabling UAV cellular with millimeter-wave communication: potentials and approaches[J]. IEEE Communications Magazine, 2016, 54(5): 66-73.

[7] JIANG B, YANG J C, XU H F, et al. Multimedia data throughput maximization in internet-of-things system based on optimization of cache-enabled UAV[J]. IEEE Internet of Things Journal, 2019, 6(2): 3525-3532.

[8] MARDANI A, CHIABERGE M, GIACCONE P. Communication-aware UAV path planning[J]. IEEE Access, 2019, 7: 52609-52621.

[9] ZHAN C, HU H, SUI X F, et al. Joint resource allocation and 3D aerial trajectory design for video streaming in UAV communication systems[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31(8): 3227-3241.

[10] 胡欣颖. 面向信息采集场景的无人机轨迹规划与通信资源分配研究[D]. 北京: 北京邮电大学, 2021: 8-19.
HU X Y. Trajectory design and resource allocation of uav for data collection in WSNs[D]. Beijing: Beijing University of Posts and Telecommunications, 2021: 8-19(in Chinese).

[11] MANZOOR A, KIM D H, HONG C S. Energy efficient resource allocation in UAV-based heterogeneous networks[C]// 2019 20th Asia-Pacific Network Operations and Management Symposium. Piscataway: IEEE Press, 2019: 1-4.

[12] ZENG F Z, HU Z Z, XIAO Z, et al. Resource allocation and trajectory optimization for QoE provisioning in energy-efficient UAV-enabled wireless networks[J]. IEEE Transactions on Vehicular Technology, 2020, 69(7): 7634-7647.

[13] LU W D, SI P Y, HUANG G X, et al. Interference reducing and resource allocation in UAV-powered wireless communication system[C]// 2020 International Wireless Communications and Mobile Computing. Piscataway: IEEE Press, 2020: 220-224.

[14] 陶丽佳, 赵宜升, 徐新雅. 无人机协助边缘计算的能量收集 MEC 系统资源分配策略[J]. 南京邮电大学学报 (自然科学版), 2022, 42(1): 37-44.
TAO L J, ZHAO Y S, XU X Y. Resource allocation strategy for UAV-assisted edge computing in energy harvesting MEC system[J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2022, 42(1): 37-44(in Chinese).

[15] QIN X T, SONG Z Y, HAO Y Y, et al. Joint resource allocation and trajectory optimization for multi-UAV-assisted multi-access mobile edge computing[J]. IEEE Wireless Communications Letters, 2021, 10(7): 1400-1404.

[16] 龚广伟, 谢添, 赵海涛, 等. 基于图着色的大规模无人机群三维网络资源分配算法[J]. 信号处理, 2022, 38(8): 1693-1702.
GONG G W, XIE T, ZHAO H T, et al. A three-dimension resources allocation algorithm for large-scale UAV network based on graph coloring[J]. Journal of Signal Processing, 2022, 38(8): 1693-1702(in Chinese).

[17] 吴启晖, 陈佳馨, 吴盛君, 等. 一种基于动态航迹的无人机群频谱资源分配方法: CN108632831A[P]. 2021-08-10.
WU Q H, CHEN J X, WU S J, et al. A dynamic trajectory-based spectrum resource allocation method for UAV swarms: CN108632831A[P]. 2021-08-10(in Chinese).

[18] CHEN M Z, SAAD W, YIN C C. Liquid state machine learning for resource allocation in a network of cache-enabled LTE-U UAVs[C]//2017 IEEE Global Communications Conference. Piscataway: IEEE Press, 2017: 1-6.

[19] CHEN J X, WU Q H, XU Y H, et al. Distributed demand-aware channel-slot selection for multi-UAV networks: a game-theoretic learning approach[J]. IEEE Access, 2018, 6: 14799-14811.

[20] ZHANG X C, ZHAO H T, XIONG J, et al. Scalable power control/beamforming in heterogeneous wireless networks with graph neural networks[C]// 2021 IEEE Global Communications Conference. Piscataway: IEEE Press, 2021: 1-6.

[21] 贺颖. 基于深度强化学习的无线网络多维资源分配技术研究[D]. 大连: 大连理工大学, 2018.
HE Y. Research on multi-dimensional resource allocation of wireless networks based on deep reinforcement learning[D]. Dalian: Dalian University of Technology, 2018(in Chinese).

[22] ZHANG Q Q, MOZAFFARI M, SAAD W, et al. Machine learning for predictive on-demand deployment of uavs for wireless communications[C]// 2018 IEEE Global Communications Conference. Piscataway: IEEE Press, 2018: 1-6.

[23] SUN N, WU J X. Minimum error transmissions with imperfect channel information in high mobility systems[C]//2013 IEEE Military Communications Conference. Piscataway: IEEE Press, 2013: 922-927.

[24] CHUA M Y K, YU F R, LI J, et al. Medium access control for unmanned aerial vehicle (UAV) ad-hoc networks with full-duplex radios and multipacket reception capability[J]. IEEE Transactions on Vehicular Technology, 2013, 62(1): 390-394.

[25] 任博, 刘圣宇, 崔连柱, 等. 多无人机侦察航路规划[C]// 2014(第五届) 中国无人机大会论文集. 北京: 航空工业出版社, 2014: 793-797.
REN B, LIU S Y, CUI L Z, et al. Multi-UAV reconnaissance route planning[C]//Proceedings of China Drone Conference. Beijing: Aviation Industry Press, 2014: 793-797(in Chinese).

[26] ACEVEDO J J, ARRUE B C, MAZA I, et al. Distributed cooperation of multiple UAVs for area monitoring missions[M]//CARBONE G, GOMEZ-BRAVO F. Motion and Operation Planning of Robotic Systems. Cham: Springer, 2015: 471-494.

[27] 张睿文, 宋笔锋, 裴扬, 等. 复杂任务场景无人机集群自组织侦察建模与仿真[J]. 航空工程进展, 2020, 11(3): 316-325,343.
ZHANG R W, SONG B F, PEI Y, et al. Modelling and simulation of UAV swarm self-organized surveillance in complex mission scenarios[J]. Advances in Aeronautical Science and Engineering, 2020, 11(3): 316-325,343(in Chinese).

[28] HAZON N, KAMINKA G A. On redundancy, efficiency, and robustness in coverage for multiple robots[J]. Robotics and Autonomous Systems, 2008, 56(12): 1102-1114.

[29] GABRIELY Y, RIMON E. Spanning-tree based coverage of continuous areas by a mobile robot[C]// Proceedings 2001 ICRA. IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2001: 1927-1933.

[30] 3GPP. Study on enhanced LTE support for aerial vehicles: TR 36.777[S]. Nice, France: ETSI, 2017.

[31] 3GPP. 5G; Study on channel model for frequencies from 0.5 to 100 GHz: TR 38.901[S]. Nice, France: ETSI, 2020.

[32] 3GPP. 5G; Unmanned aerial system (UAS) support in 3GPP: TS 22.125[S]. Nice, France: ETSI, 2020.

[33] RICE M, DYE R, WELLING K. Narrowband channel model for aeronautical telemetry[J]. IEEE Transactions on Aerospace and Electronic Systems, 2000, 36(4): 1371-1376.

[34] GODDEMEIER N, WIETFELD C. Investigation of air-to-air channel characteristics and a UAV specific extension to the rice model[C]// 2015 IEEE Globecom Workshops. Piscataway: IEEE Press, 2015: 1-5.

[35] AHMED N, KANHERE S S, JHA S. On the importance of link characterization for aerial wireless sensor networks[J]. IEEE Communications Magazine, 2016, 54(5): 52-57.

[36] ABUALHAOL I Y, MATALGAH M M. Performance analysis of multi-carrier relay-based UAV network over fading channels[C]// 2010 IEEE Globecom Workshops. Piscataway: IEEE Press, 2010: 1811-1815.

[37] GODDEMEIER N, ROHDE S, WIETFELD C. Experimental validation of RSS driven UAV mobility behaviors in IEEE 802.11s networks[C]// 2012 IEEE Globecom Workshops. Piscataway: IEEE Press, 2012: 1550-1555.

[38] NOWÉ A, VRANCX P, DE HAUWERE Y M. Game theory and multi-agent reinforcement learning[M]// WIERING M, VAN OTTERLO M. Reinforcement Learning. Berlin: Springer, 2012: 441-470.

[39] NEYMAN A. From Markov chains to stochastic games[M]// Stochastic Games and Applications. Dordrecht: Springer Netherlands, 2003: 9-25.

[40] SUTTON R S, BARTO A G. Reinforcement learning: an introduction[M]. Cambridge: MIT Press, 1998.

[41] TAMPUU A, MATIISEN T, KODELJA D, et al. Multiagent cooperation and competition with deep reinforcement learning[J]. PLoS One, 2017, 12(4): e0172395.

Dynamic communication resource allocation for multi-UAV area coverage

LU Maomao[1, 2], LIU Chunhui[3, *], DONG Zanliang[4]

(1. School of Communication Engineering, Xidian University, Xi'an 710071, China;
2. 54th Research Institute of China Electronics Technology Group Corporation, Shijiazhuang 050299, China;
3. Institute of Unmanned System, Beihang University, Beijing 100191, China;
4. School of Electronics and Information Engineering, Beihang University, Beijing 100191, China)

Abstract: This study presents a reinforcement learning-based multi-agent dynamic communication resource allocation model that addresses the issue of communication resource allocation in multi-UAV area coverage tasks. We first generate the coverage route of each UAV in the mission area by the multi-agent spanning tree coverage (MSTC) method, and model the communication link between the UAV and ground base station as well as UAV pairs. The uncertainty inherent in the flight environment motivates the modeling of the long-term resource allocation problem as a random game. T Considered an agent, the air-to-air connection between UAVs entails receiver, subchannel, and transmission power selection, among other modifications. We then design a multi-agent reinforcement learning (MARL) model based on the double deep Q-network (DDQN), where each agent learns the optimal communication resource allocation strategy through the feedback of the reward function. As shown by simulation results, the proposed MARL method can increase the overall channel capacity, decrease interference from air-to-ground uplink, and optimize communication resource allocation strategies under dynamic trajectories and delay constraints, while also improving the success rate of load delivery.

Keywords: multi-UAV area coverage; dynamic communication resource allocation; reinforcement learning; double deep Q-network; multi-agent

Received: 2022-08-30; **Accepted**: 2022-11-04; **Published Online**: 2022-12-02 10:22

URL: link.cnki.net/urlid/11.2625.V.20221201.1504.003

Foundation item: Science and Technology Innovation 2030—Key Project of "New Generation Artificial Intelligence" (2020AAA0108200)

* **Corresponding author**. E-mail: liuchunhui2134@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0747

空间飞网系统动力学建模与仿真

陈鹏旭[1,2]，吴晨晨[1,3,*]，倪智宇[4]

(1. 中国科学院 沈阳自动化研究所，沈阳 110017；2. 沈阳工业大学 机械工程学院，沈阳 110870；
3. 南京理工大学 物理学院，南京 210094；4. 沈阳航空航天大学 航空宇航学院，沈阳 110136)

摘 要：空间飞网系统捕获带有自旋的空间碎片后，碎片会对捕获网与卫星间的连接绳施加扭矩，而目前绳索相关动力学模型无法完整描述其扭转行为。基于集中质量法，对绳索进行二次离散，建立了可以描述绳索扭转行为的动力学模型。通过与传统绳索动力学模型的对比及扭转实验，验证了该模型能够描述绳索在拉伸、弯曲及扭转载荷下的行为，发现绳索的扭转会产生额外的牵引力。针对飞网展开过程进行了仿真分析与实验，验证该动力学模型适用于飞网系统，并对飞网与空间碎片的组合体进行了仿真分析，发现因为绳索扭转产生的牵引力，会对主体卫星的稳定性产生一定的影响，忽略绳索的扭转行为会降低动力学仿真模型的准确性。

关 键 词：绳索动力学；飞网；空间碎片；集中质量法；多体动力学
中图分类号：V414.9; TB122
文献标志码：A **文章编号**：1001-5965(2024)09-2951-12

空间碎片日益增多，已经威胁到在轨航天器的安全运行，并且空间碎片之间会相互碰撞使碎片数量进一步增加(斯凯勒效应[1])，开展空间碎片主动清除技术研究势在必行。NASA[2]指出，每年需要清除5～10个高风险空间碎片，从而保证近地轨道的环境安全。空间碎片清除中的核心技术是2020年中国航天大会发布的十大“宇航领域科学问题和技术难题”之一。在过去的几十年，学术界相继提出了多种空间碎片清除方法，例如机械臂[3]、飞矛[4]、飞爪[5]及激光清除等。其中，飞网捕获技术具有成本低、质量轻、体积小及捕获距离远等优点，成为空间碎片主动清除的主要技术手段之一。自2002年欧空局在地球静止轨道清理机器人ROGER[5]计划中首次提出飞网捕获的概念，各国均开展了相关技术的深入研究，英国萨里大学萨里空间中心于2018年在RemoveDEBRIS任务[6]中首次在轨演示了飞网捕获过程，完成了飞网捕获技术的在轨试验验证，但该任务没能在轨验证系统的拖曳离轨技术。

飞网捕获技术涉及多个学科，其中，绳索动力学建模被认为是飞网理论研究的重点方向之一。绳索能够承受较大的拉力，但其抗弯能力较弱，且在内部发生扭转时会对绳索的两端产生拉力。绳索动力学常用的建模方法主要包括集中质量法[7-9]和绝对节点坐标法(absolute nodal coordinate formulation, ANCF)[10-13]。其中，集中质量法是最经典且应用最广泛的方法[14-17]，其近年来依然处于不断完善和改进的状态。集中质量法将绳索离散为多个质点或圆柱段(也称为有限段模型)[18]，并采用弹簧阻尼器连接离散体，具有建模简洁、计算速度快、精度高[19]等特点。Hovell[20]等在集中质量法的基础上将绳索刚度系数修改为非线性变量，通过引入非线性刚度系数，使模型可以更好地描述绳索的非线性行为。Hou[21]提出了离散杆与节能积分算法结合的集成方法，使飞网仿真计算速度得到提高。

收稿日期：2022-08-30；**录用日期**：2023-01-10；**网络出版时间**：2023-02-08 10:36
网络出版地址：link.cnki.net/urlid/11.2625.V.20230207.1705.003
基金项目：国家自然科学基金(51905527)；中国科学院沈阳自动化研究所自主部署项目；中央高校基本科研业务费专项资金；辽宁省教育厅基本科研项目(JYTMS20230254)；国家自然科学基金基础科学中心项目(62388101)
* **通信作者**. E-mail：wuchenchen@njust.edu.cn

引用格式：*陈鹏旭，吴晨晨，倪智宇. 空间飞网系统动力学建模与仿真[J]. 北京航空航天大学学报，2024，50(9)：2951-2962.*
CHEN P X，WU C C，NI Z Y. Dynamic modelling and simulation of a tethered-net in space[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2951-2962 (in Chinese).

ANCF是由Shabana[10]提出的，其是通过单元节点的绝对位置坐标及其导数来描述节点运动及变形，常用于求解大变形、大位移等非线性问题，其特点是精度高但计算耗时[22]，同条件下绝对节点坐标法的单次迭代用时是集中质量法的7.5倍[23]。

在空间碎片移除的相关研究中，往往需要考虑空间碎片的自旋问题。受残余角动量、重力梯度、光压摄动等影响，多数空间碎片往往伴随着自由翻滚运动[24]，使飞网的连接绳产生扭转。连接绳的扭转行为会对主体卫星的姿控产生影响，然而，现有的绳索动力学模型无法精确描述绳索的扭转行为，主要是因为：1）集中质量法将绳索离散成质点，一维的质点模型无法精确描述绳索的形态变化；2）有限段模型将绳索离散成圆柱形离散体，虽然可以描述绳索的位形，但无法描述绳索在松弛状态下对扭转的反馈以及绳索因扭转而产生拉力的现象；3）绝对节点坐标法在建模时忽略了剪切、扭转等变形[25]。针对上述问题，本文通过二次离散建立了绳索动力学理论模型。该模型可以描述绳索受到弯曲、扭转和拉伸载荷时的位形变化，同时可以描述绳索扭转与内力变化的联系，进而更好地描述绳索对卫星姿态控制的影响。

1 动力学建模

空间飞网系统主要由发射机构、连接绳、飞网与牵引块（收口机构）组成。捕获过程（见图1）中，通过发射机构将牵引块弹出卫星，牵引块带动飞网展开，并在捕获目标后收紧网口，通过连接绳实现飞网与卫星的连接。

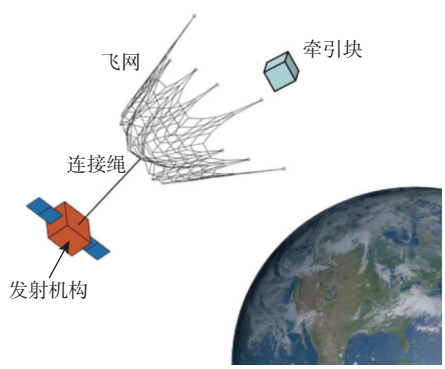


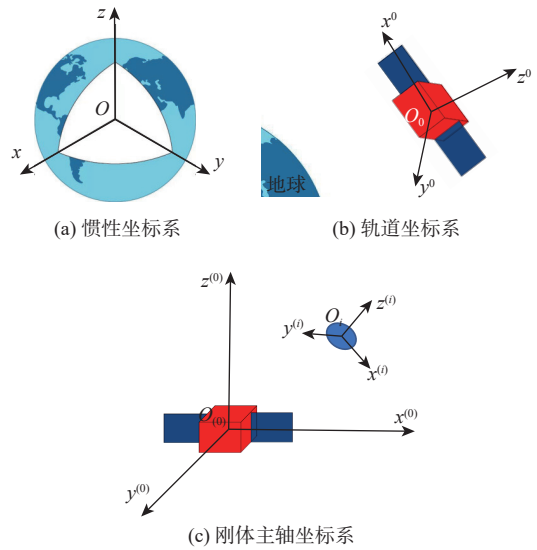
图1 空间飞网示意

Fig. 1 Schematic diagram of a tethered-net system

本文建立了如图2所示的坐标系：

1）惯性坐标系$Oxyz$，其原点位于地心，z轴指向北极；x轴指向春分点；y轴与x轴，z轴构成右手系。

2）航天器的轨道坐标系$O_{(0)}xyz^{(0)}$，其原点位于航天器的质心，$z^{(0)}$轴由地心指向航天器；$x^{(0)}$轴在航天器轨道面内与$z^{(0)}$轴垂直，并指向卫星移动速度方向；$y^{(0)}$轴与$x^{(0)}$、$z^{(0)}$构成右手系。



(a) 惯性坐标系 (b) 轨道坐标系

(c) 刚体主轴坐标系

图2 坐标系示意

Fig. 2 Schematic diagram of the coordinate system

3）刚体的主轴坐标系$O_ixyz^{(i)}$，其原点位于刚体质心，三轴为刚体的惯量主轴构成的右手系。

1.1 柔性绳索动力学描述

绳索发生扭转后，其内部纤维会呈螺旋状排列从而导致绳索长度缩短。使用集中质量法对绳索动力学建模时，将绳索离散为质点的建模方式无法描述绳索的形态变化，而将绳索离散为圆柱段的建模方式又无法表述“绳索长度因扭转而缩短”的现象。考虑到上述问题，本文通过二次离散绳索构建了可以描述绳索扭转行为的动力学模型。模型选取薄圆片作为一次离散体、质点为二次离散体，具体离散过程如图3所示。

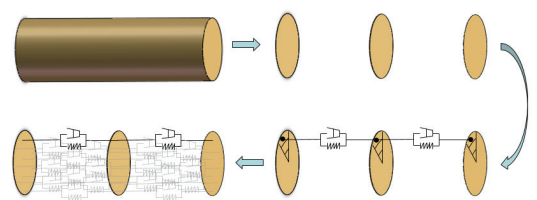


图3 离散过程示意

Fig. 3 Schematic diagram of discrete process

1）将绳索划分为若干个单元绳段，将绳段单元的质量集中在两端，得到如图3所示的二维匀质薄圆片；

2）将圆片划分为若干个扇形区域，将质量集中于扇形区域质心，得到的质点群为二次离散体；

3）采用弹簧阻尼器将相邻薄圆片上的对应质点相连接。

考虑到模型计算量，绳索的位形由每个薄圆片B_i的位形来确定，B_i的位形由质心O_{ci}相对$(O_0\text{-}e^{(0)})$的笛卡尔坐标（x_i, y_i, z_i）及B_i相对$(O_0\text{-}e^{(0)})$姿态的

3 个欧拉角(α_i, β_i, γ_i)来表达, 两者的坐标向量$\boldsymbol{O}_{ci}$、$\boldsymbol{\theta}_i$可以表示为

$$\begin{cases}\boldsymbol{O}_{ci}=\begin{bmatrix}x_i & y_i & z_i\end{bmatrix}^{\mathrm{T}}\\ \boldsymbol{\theta}_i=\begin{bmatrix}\alpha_i & \beta_i & \gamma_i\end{bmatrix}\end{cases} \tag{1}$$

假设将绳索离散成 N 个薄圆片 ($1\leqslant i\leqslant N$) 以及 n个质点, 则 N 与n取值越大, 模型整体精度越高。因质点间的内力之和以积分形式计算, n取值趋于无穷, 此时薄圆片与质点群转动惯量等效。根据刚体动力学[26], 薄圆片的位形计算如下:

$$m_i\ddot{\boldsymbol{O}}_{ci}=\boldsymbol{T}_i+\boldsymbol{F}_{\mathrm{G}}+\boldsymbol{F}_{\mathrm{e}} \tag{2}$$

$$\boldsymbol{J}_i\cdot\dot{\boldsymbol{\omega}}_i+\boldsymbol{\omega}_i\times(\boldsymbol{J}_i\cdot\boldsymbol{\omega}_i)=\boldsymbol{M}_i+\boldsymbol{M}_{\mathrm{G}}+\boldsymbol{M}_{\mathrm{e}} \tag{3}$$

式中, $\boldsymbol{\omega}_i$为薄圆片 B_i的角速度; $\boldsymbol{J}_i$为转动惯量。因为薄圆片的厚度可忽略不计, 且B_i的旋转轴为惯量主轴, 所以转动惯量$\boldsymbol{J}_i$的取值为

$$\boldsymbol{J}_i=\begin{bmatrix}\dfrac{m_ih^2}{4} & & \\ & \dfrac{m_ih^2}{4} & \\ & & \dfrac{m_ih^2}{2}\end{bmatrix} \tag{4}$$

式中, h为绳索处于张紧临界状态时的半径。m_i为薄圆片 B_i的质量, 其为相邻绳段质量的一半, 考虑到B_i在绳索端点和非端点的区别, m_i的取值有以下两种情况:

$$m_i=\begin{cases}\dfrac{1}{2}\rho Al & i\in[1,N]\\ \rho Al & i\notin[1,N]\end{cases} \tag{5}$$

式中, ρ为绳索的密度; A为绳索的横截面积; l为两个薄圆片之间的初始距离, 其表达式如下:

$$l=\frac{L}{N-1} \tag{6}$$

其中, L为绳索总长。

式 (2)～式 (3) 中的$\boldsymbol{F}_{\mathrm{G}}$和$\boldsymbol{M}_{\mathrm{G}}$分别为重力和重力梯度力矩, $\boldsymbol{F}_{\mathrm{e}}$和 $\boldsymbol{M}_{\mathrm{e}}$分别为其余外力和外力矩。$\boldsymbol{\omega}_i$为$B_i$的角速度, 由于$B_i$的位形由$\boldsymbol{\theta}_i$表达, 因此计算时需要对$\boldsymbol{\theta}_i$与$\boldsymbol{\omega}_i$进行转换, $\boldsymbol{\theta}_i$的导数$\dot{\boldsymbol{\theta}}_i$与$\boldsymbol{\omega}_i$可利用方向余弦矩阵进行转换。$\boldsymbol{F}_{\mathrm{G}}$和$\boldsymbol{M}_{\mathrm{G}}$的表达式及$\dot{\boldsymbol{\theta}}_i$与$\boldsymbol{\omega}_i$直接的变换关系如下:

$$\boldsymbol{F}_{\mathrm{G}}=-\frac{\mu m_i}{|\boldsymbol{R}_i|^3}\boldsymbol{R}_i \tag{7}$$

$$\boldsymbol{M}_{\mathrm{G}}=-\frac{3\mu\boldsymbol{r}_i}{|\boldsymbol{R}_i|^3|\boldsymbol{r}_i|}\times\boldsymbol{J}_i\frac{\boldsymbol{r}_i}{|\boldsymbol{r}_i|} \tag{8}$$

$$\begin{bmatrix}\dot{\alpha}_i\\ \dot{\beta}_i\\ \dot{\gamma}_i\end{bmatrix}=\begin{bmatrix}1 & \sin\alpha_i\tan\beta_i & \cos\alpha_i\tan(\beta_i)\\ 0 & \cos\alpha_i & -\sin\alpha_i\\ 0 & \sin\alpha_i/\cos\beta_i & \cos\alpha_i/\cos\beta_i\end{bmatrix}\boldsymbol{\omega}_i \tag{9}$$

式中, μ是引力常数; $\boldsymbol{R}_i$为B_i在惯性坐标系下的位置矢量; $\boldsymbol{r}_i$为B_i在刚体坐标系下的位置矢量。

$\boldsymbol{T}_i$和$\boldsymbol{M}_i$分别为B_i所受的内力和内力矩, 其数值为B_i的二次离散体的受力总和。$\boldsymbol{T}_i$和$\boldsymbol{M}_i$的计算公式如下:

$$\boldsymbol{T}_i=\sum_{j=1}^{n}\boldsymbol{T}_{i,j}+\sum_{j=1}^{n}\boldsymbol{T}_{i+1,j} \tag{10}$$

$$\boldsymbol{M}_i=\sum_{j=1}^{n}\boldsymbol{M}_{i,j}+\sum_{j=1}^{n}\boldsymbol{M}_{i+1,j} \tag{11}$$

式 (10) 和式 (11) 中$\boldsymbol{T}_{i,j}$为 $p_{i,j}$与 $p_{i+1,j}$间的内力、$\boldsymbol{M}_{i,j}$为 $p_{i,j}$与 $p_{i+1,j}$间的内力作用在 B_i 上的内力矩, $p_{i,j}$表示在薄圆片B_i上的第 j个质点(见图 4)。$\boldsymbol{T}_{i,j}$可根据 Kelvin-Voigt 模型计算[14], $\boldsymbol{M}_{i,j}$根据力元矢量$\boldsymbol{r}_{i,j}$与$\boldsymbol{T}_{i,j}$计算, 如式 (12)～式 (14) 所示:

$$\boldsymbol{l}'=\boldsymbol{R}_{i,j}-\boldsymbol{R}_{i+1,j} \tag{12}$$

$$\boldsymbol{T}_{i,j}=\begin{cases}\left(k_{i,j}(|\boldsymbol{l}'|-l)+D_{i,j}\left|\boldsymbol{V}_{i,j}-\boldsymbol{V}_{i+1,j}\right|\right)\dfrac{\boldsymbol{l}'}{|\boldsymbol{l}'|} & \|\boldsymbol{l}'\|\geqslant l\\ 0 & \|\boldsymbol{l}'\|\leqslant l\end{cases} \tag{13}$$

$$\boldsymbol{M}_{i,j}=\begin{cases}\boldsymbol{r}_{i,j}\times\boldsymbol{T}_{i,j} & \|\boldsymbol{l}'\|\geqslant l\\ 0 & \|\boldsymbol{l}'\|\leqslant l\end{cases} \tag{14}$$

力元矢量$\boldsymbol{r}_{i,j}$在刚体坐标系框架内并由$\boldsymbol{O}_{c_i}$指向 $p_{i,j}$, $\boldsymbol{R}_{i,j}$和$\boldsymbol{V}_{i,j}$分别为质点在轨道坐标系中的位置和速度。$k_{i,j}$和$D_{i,j}$分别代表质点间的弹簧系数和阻尼系数, 计算公式如下:

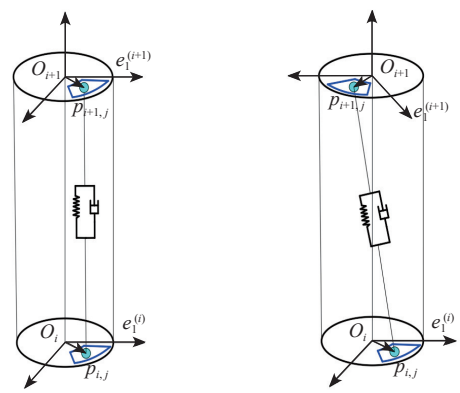
$$k_{i,j}=\frac{E\cdot S}{l} \tag{15}$$

$$D_{i,j}=\xi\sqrt{\rho E\cdot S^2} \tag{16}$$

式中, S 为质点所处离散区域的面积; E 为杨氏模量; ξ为阻尼比。

式 (12)～(13) 中, $\boldsymbol{R}_{i,j}$和$\boldsymbol{R}_{i+1,j}$根据$\boldsymbol{r}_{i,j}$、$\boldsymbol{r}_{i+1,j}$、$\boldsymbol{O}_{ci}$及$\boldsymbol{\theta}_i$计算:

$$\boldsymbol{R}_{i,j}=\boldsymbol{r}_{i,j}^{(0)}+\boldsymbol{O}_{ci} \tag{17}$$



(a) 初始位置　　(b) 扭转后

图 4　模型参数示意

Fig. 4　Schematic diagram of model parameters

$\boldsymbol{r}_{i,j}$和$\boldsymbol{r}_{i+1,j}$分别属于$O_ixyz^{(i)}$与$O_{i+1}xyz^{(i+1)}$坐标系，因此需要将$\boldsymbol{r}_{i,j}$和$\boldsymbol{r}_{i+1,j}$变换至轨道坐标系进行计算，$\boldsymbol{r}_{i,j}^{(0)}$为$\boldsymbol{r}_{i,j}$在轨道坐标系中的表达形式。考虑到绳索在姿态变换过程中幅度较大，使用欧拉角构成的方向余弦矩阵进行姿态变换会引起万向锁问题，因此$\boldsymbol{r}_{i,j}$至$\boldsymbol{r}_{i,j}^{(0)}$的变换由四元数实现：

$$\begin{bmatrix} 0 \\ \boldsymbol{r}_{i,j}^{(0)} \end{bmatrix} = q^* \begin{bmatrix} 0 \\ \boldsymbol{r}_{i,j} \end{bmatrix} q_i \tag{18}$$

式中，q_i为轨道坐标系至刚体坐标系$O_ixyz^{(i)}$的四元数，q_i^*为q_i的共轭四元数。定义四元数为$q_i=[(u_1,u_2,u_3)w]$，由于B_i的位形定义由欧拉角构成，因此四元数采用欧拉角的形式表达如下：

$$\begin{cases} u_1 = \sin\dfrac{\beta_i}{2}\sin\dfrac{\gamma_i}{2}\cos\dfrac{\alpha_i}{2} + \cos\dfrac{\beta_i}{2}\cos\dfrac{\gamma_i}{2}\sin\dfrac{\alpha_i}{2} \\ u_2 = \sin\dfrac{\beta_i}{2}\cos\dfrac{\gamma_i}{2}\cos\dfrac{\alpha_i}{2} + \cos\dfrac{\beta_i}{2}\sin\dfrac{\gamma_i}{2}\sin\dfrac{\alpha_i}{2} \\ u_3 = \cos\dfrac{\beta_i}{2}\sin\dfrac{\gamma_i}{2}\cos\dfrac{\alpha_i}{2} - \sin\dfrac{\beta_i}{2}\cos\dfrac{\gamma_i}{2}\sin\dfrac{\alpha_i}{2} \\ w = \cos\dfrac{\beta_i}{2}\cos\dfrac{\gamma_i}{2}\cos\dfrac{\alpha_i}{2} - \sin\dfrac{\beta_i}{2}\sin\dfrac{\gamma_i}{2}\sin\dfrac{\alpha_i}{2} \\ q_i = [(u_1,u_2,u_3)w] \end{cases} \tag{19}$$

根据离散方式可知$\boldsymbol{r}_{i,j}=\boldsymbol{r}_{i+1,j}$，将式(17)带入式(12)得到式(20)：

$$\boldsymbol{l}' = \boldsymbol{r}_{i,j}^{(0)} - \boldsymbol{r}_{i+1,j}^{(0)} + \boldsymbol{O}_{ci} - \boldsymbol{O}_{ci+1} \tag{20}$$

当二次离散体数量n取值趋于无穷时，式(10)和式(11)可以变换成二重积分形式：

$$\boldsymbol{T}_i = \iint \boldsymbol{T}_{i,j} + \iint \boldsymbol{T}_{i+1,j} \tag{21}$$

$$\boldsymbol{M}_i = \iint \boldsymbol{M}_{i,j} + \iint \boldsymbol{M}_{i+1,j} \tag{22}$$

以积分形式计算时，二次离散体所处区域的形状与图4中的蓝色扇形区域一致，扇形区域的面积S可由以下公式计算：

$$S = r\cdot \mathrm{d}\theta \mathrm{d}r \tag{23}$$

式中，$\mathrm{d}\theta$为离散体所处的扇形区域的圆心角；$\mathrm{d}r$为扇形区域弧线的间距。定义r为$p_{i,j}$与B_i质心的距离，θ为$\boldsymbol{r}_{i,j}$与$O_ixyz^{(i)}$的x轴的夹角，则有：

$$\boldsymbol{r}_{i,j} = \begin{bmatrix} r\cos\theta \\ r\sin\theta \\ 0 \end{bmatrix} \tag{24}$$

将式(13)、式(23)和式(24)带入式(21)和式(22)，则$\boldsymbol{T}_i$与$\boldsymbol{M}_i$的计算公式如下：

$$\boldsymbol{T}_i = \int_0^h \int_{\varphi_{i,s}}^{\varphi_{i,j}} \left[\frac{E}{l}\left(\left|\boldsymbol{R}_{i,j}-\boldsymbol{R}_{i+1,j}\right|-l\right) + \xi\sqrt{\rho E}\left|\boldsymbol{V}_{i,j}-\boldsymbol{V}_{i+1,j}\right| \right] \frac{c_{\mathrm{rope}} r\left(\boldsymbol{R}_{i,j}-\boldsymbol{R}_{i+1,j}\right)}{\left|\boldsymbol{R}_{i,j}-\boldsymbol{R}_{i+1,j}\right|}\mathrm{d}\theta\mathrm{d}r + \int_0^h \int_{\varphi_{i-1,s}}^{\varphi_{i-1}} \left[\frac{E}{l}\left(\left|\boldsymbol{R}_{i-1,j}-\boldsymbol{R}_{i,j}\right|-l\right) + \xi\sqrt{\rho E}\left|\boldsymbol{V}_{i-1,j}-\boldsymbol{V}_{i,j}\right| \right] \frac{c_{\mathrm{rope}} r\left(\boldsymbol{R}_{i-1,j}-\boldsymbol{R}_{i,j}\right)}{\left|\boldsymbol{R}_{i-1,j}-\boldsymbol{R}_{i,j}\right|}\mathrm{d}\theta\mathrm{d}r \tag{25}$$

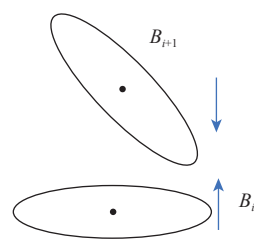
$$\boldsymbol{M}_i = \int_0^h \int_{\varphi_{i,s}}^{\varphi_{i,j}} \boldsymbol{r}_{i,j} \times \left[\frac{E}{l}\left(\left|\boldsymbol{R}_{i,j}-\boldsymbol{R}_{i+1,j}\right|-l\right) + \xi\sqrt{\rho E}\left|\boldsymbol{V}_{i,j}-\boldsymbol{V}_{i+1,j}\right| \right] \frac{c_{\mathrm{rope}} r\left(\boldsymbol{R}_{i,j}-\boldsymbol{R}_{i+1,j}\right)}{\left|\boldsymbol{R}_{i,j}-\boldsymbol{R}_{i+1,j}\right|}\mathrm{d}\theta\mathrm{d}r + \int_0^h \int_{\varphi_{i-1,s}}^{\varphi_{i-1}} \boldsymbol{r}_{i-1,j} \times \left[\frac{E}{l}\left(\left|\boldsymbol{R}_{i-1,j}-\boldsymbol{R}_{i,j}\right|-l\right) + \xi\sqrt{\rho E}\left|\boldsymbol{V}_{i-1,j}-\boldsymbol{V}_{i,j}\right| \right] \frac{c_{\mathrm{rope}} r\left(\boldsymbol{R}_{i-1,j}-\boldsymbol{R}_{i,j}\right)}{\left|\boldsymbol{R}_{i-1,j}-\boldsymbol{R}_{i,j}\right|}\mathrm{d}\theta\mathrm{d}r \tag{26}$$

式中，$\varphi_{i,s}$与$\varphi_{i,l}$为薄圆片B_i的圆心角角度；c_{rope}为判断绳索拉伸程度的系数，当$\|\boldsymbol{l}'\|\geqslant l$时$c_{\mathrm{rope}}$取1，$\|\boldsymbol{l}'\|\leqslant l$时$c_{\mathrm{rope}}$取0。计算时，$\varphi_{i,s}$与$\varphi_{i,l}$的取值分为2种情况：第1种情况为相邻圆片的质心距离大于初始距离，此时可直接令$\varphi_{i,s}$与$\varphi_{i,l}$分别等于0与2π；第2种情况为相邻圆片的质心距离小于等于初始距离且相邻圆片不平行，此时可将圆片划分出无内力区域，若区域内的质点均满足$l'\leqslant l$的条件，则可以不参与计算。无内力区域划分的原理是圆片向内偏转一侧的质点间距会更小且相邻圆片间距小于初始距离，因此判定$l'\leqslant l$。

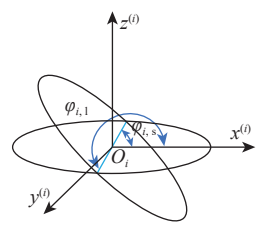
$\varphi_{i,s}$与$\varphi_{i,l}$的取值根据无内力区域的分界线来定义，分界线可以通过假设B_i与B_{i+1}的质心重合来获取，由于一次离散体为薄圆片，因此其重合得到的切线即为分界线，如图5所示。一个空间旋转可以由互为相反数的两个四元数表示，因此相反的两个四元数的虚轴即为B_i与B_{i+1}重合后的切线，可以通过$q_{i+1,i}$的虚轴确定$\varphi_{i,s}$至$\varphi_{i,l}$的取值范围，$q_{i+1,i}$表示从$O_ixyz^{(i)}$至$O_{i+1}xyz^{(i+1)}$的四元数。四元数具有如下性质：

$$q_{ca} = q_{cb}q_{ba} \tag{27}$$

式中，q_{ca}为坐标系a到坐标系c的四元数，q_{cb}为坐标



(a) 薄圆片初始位形



(b) 薄圆片质心重合后位形

图5 $\varphi_{i,s}$、$\varphi_{i,l}$及切线示意(蓝色为相交切线)

Fig. 5 Schematic about $\varphi_{i,s}$, $\varphi_{i,l}$ and tangents (intersecting tangent lines in blue)

系 b 到坐标系 c 的四元数, q_{ba} 为坐标系 a 到坐标系 b 的四元数。由四元数的性质可得:

$$q_{i+1,i} = q_{i+1}q_i^* \tag{28}$$

定义 $q_{i+1,i}$ 的虚轴与 $O_i xyz^{(i)}$ 的 x 轴的夹角为 $\varphi_{i,s}$, 虚轴相反的共线向量与 x 轴的夹角为 $\varphi_{i,l}$, 如图 5 所示。将 $\varphi_{i,s}$、$\varphi_{i,l}$ 及式 (25) 带入式 (26)～式 (27) 即可得到 T_i、M_i。

1.2　牵引块与捕获物动力学模型

在飞网系统(包括飞网和连接绳)中, 牵引块与网体直接连接, 将带动网体展开, 且其动力学模型中的位姿坐标定义与绳索节点一致, 本文不做重复叙述。牵引块动力学方程如下:

$$m_{\mathrm{b}}\boldsymbol{O}_{\mathrm{cb}} = \boldsymbol{F}_{\mathrm{b}} + \boldsymbol{F}_{\mathrm{G}} + \boldsymbol{F}_{\mathrm{e}} \tag{29}$$

$$\boldsymbol{J}_{\mathrm{b}}\cdot\dot{\boldsymbol{\omega}}_{\mathrm{b}} + \boldsymbol{\omega}_{\mathrm{b}}\times(\boldsymbol{J}_{\mathrm{b}}\cdot\boldsymbol{\omega}_{\mathrm{b}}) = \boldsymbol{M}_{\mathrm{b}} + \boldsymbol{M}_{\mathrm{G}} + \boldsymbol{M}_{\mathrm{e}} \tag{30}$$

式中, $\boldsymbol{F}_{\mathrm{b}}$ 和 $\boldsymbol{M}_{\mathrm{b}}$ 分别为牵引块受到绳索的拉力和力矩, 其大小为所有的绳索与牵引块连接点的拉力和力矩总和。设 s 为牵引块与绳索连接点的编号, 则有:

$$\boldsymbol{F}_{\mathrm{b}} = \sum_{s=1,2}\boldsymbol{T}_s \tag{31}$$

$$\boldsymbol{M}_{\mathrm{b}} = \sum_{s=1,2}\boldsymbol{M}_s \tag{32}$$

捕获物的动力学方程描述如下:

$$m_{\mathrm{c}}\boldsymbol{O}_{\mathrm{cc}} = \boldsymbol{F}_{\mathrm{N}} + \boldsymbol{F}_{\mathrm{G}} + \boldsymbol{F}_{\mathrm{e}} \tag{33}$$

$$\boldsymbol{J}_{\mathrm{c}}\cdot\dot{\boldsymbol{\omega}}_{\mathrm{c}} + \boldsymbol{\omega}_{\mathrm{c}}\times(\boldsymbol{J}_{\mathrm{c}}\cdot\boldsymbol{\omega}_{\mathrm{c}}) = \boldsymbol{M}_{\mathrm{N}} + \boldsymbol{M}_{\mathrm{G}} + \boldsymbol{M}_{\mathrm{e}} \tag{34}$$

式中, $\boldsymbol{F}_{\mathrm{N}}$ 为捕获物受到来自绳网的接触力与摩擦力的总和, $\boldsymbol{M}_{\mathrm{N}}$ 为来自绳网的接触力与摩擦力作用在捕获物上的力矩总和。在第 3 章仿真中, 由于假设绳网与捕获物紧密贴合, 因此在此条件下可以对捕获物与绳网的组合进行一定程度的简化, 简化后的方程与式 (29)～式 (32) 一致, 这里不做重复叙述。

2　柔性绳索动力学仿真校验

本文通过二次离散得到的绳索动力学模型(以下简称二次离散模型)与传统绳索模型的区别在于, 该模型可以表述绳索受到扭转后离散体间的位形和内力变化。为了验证二次离散模型的有效性, 本节对该模型与绳索的有限段模型和质点模型进行了对比。首先通过仿真确定了离散体的数量, 随后观察对绳索施加扭矩后, 各个节点在两种不同边界条件和载荷(算例 1、2)下的反馈。仿真假设过程中无重力和空气阻力等外力影响。为便于分析与实验观测, 绳索参数选取如下: 绳索长度设置为 9 cm, 绳索直径为 1 cm, 弹性模量设置为 2 GPa, 密度为 1 440 kg/m^3。本章仿真与实验中的绳索材料参数均相同。

2.1　离散体数量选取

模型的整体计算时长与精度取决于离散体的数量选取是否恰当, 离散体数量过少则精度低, 离散数量过多则计算量会增加。从第 1 章的动力学公式可知, 二次离散体的受力根据积分计算, 因此二次离散模型仅需确定一次离散体的数量 N。为确定 N, 本文对一段 9 cm 长的绳索进行了 6 组数值仿真, 每组离散体的数量选取依次为 2、4、6、8、10、12。仿真条件为绳索一端完全固定, 另一端固定 3 个移动自由度并施加 36°的扭转角度, 仿真结果如图 6 所示。6 组绳索受扭后均对固定端产生了拉力, 拉力的数值在离散体数量大于 8 后趋于平稳, 由此可知, 9 cm 长的绳索的离散体数量选取 8 以上可以保证计算精度。考虑到计算时长, 本章离散体的数量取值为 10, 并将每个离散体作为观测节点记录数据, 节点编号如图 7 所示。

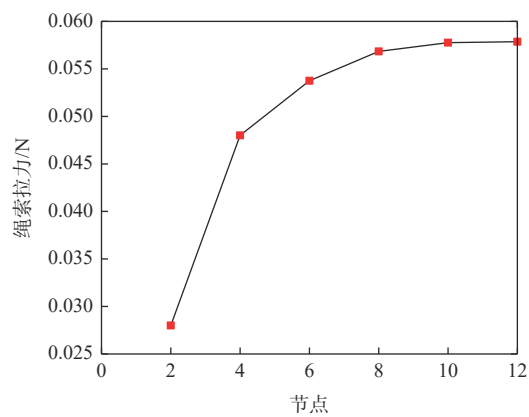


图 6　拉力与离散体数量的关系

Fig. 6　Relationship between tension and discrete quantities

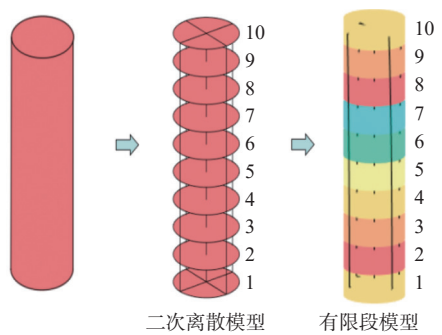


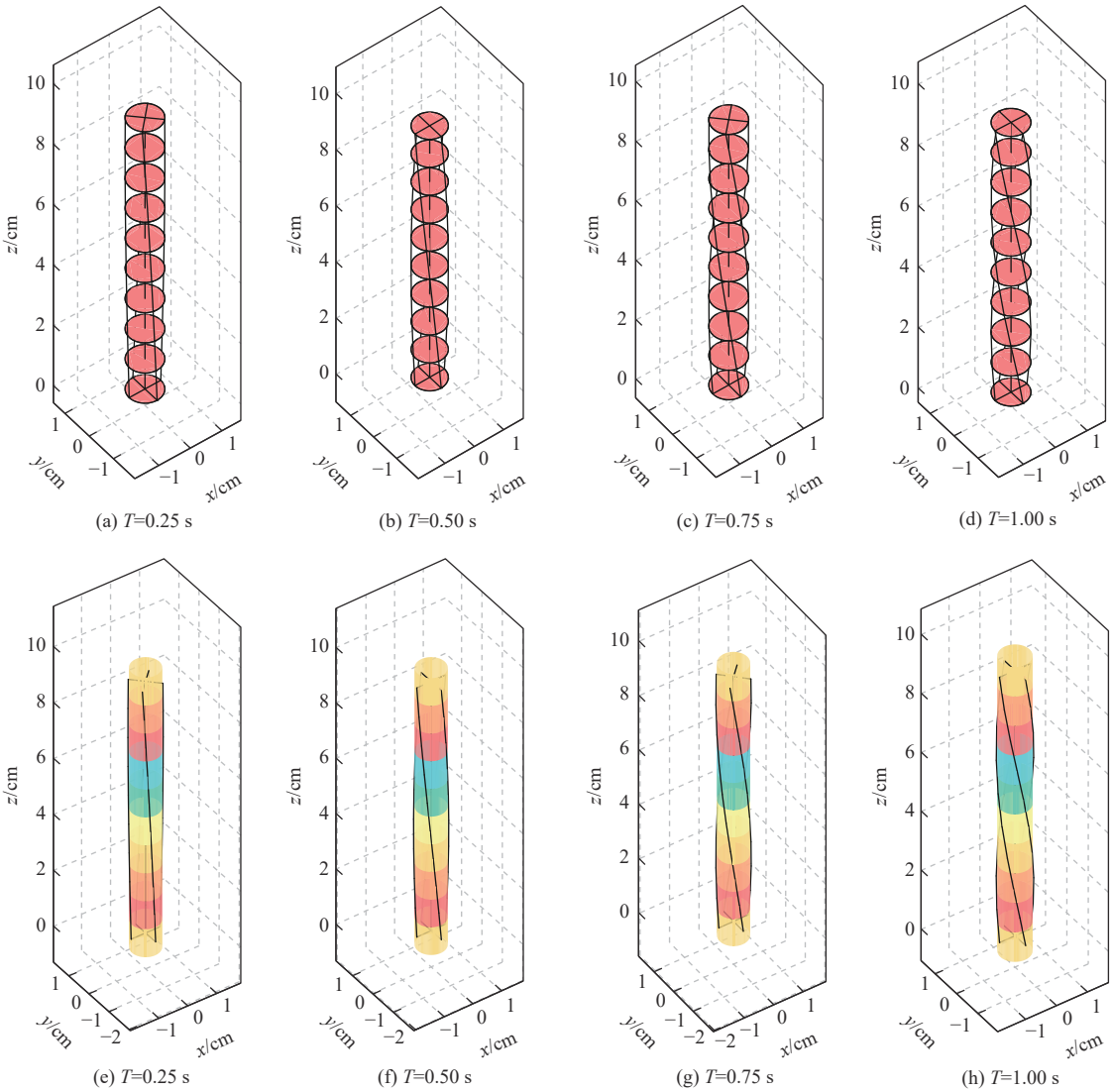
图 7　绳索观测节点示意

Fig. 7　Schematic diagram of tether observation node

2.2　扭转分析

算例 1 旨在考虑扭转对绳索的影响, 将节点 1 完全固定, 对二次离散模型和有限段模型中的节

点10施加0.5 r/s转速，仿真分析步长为0.002 s，仿真得到绳索在0 s至1 s内不同时刻的位形变化，在T=0.25 s、0.50 s、0.75 s、1.00 s时结果如图8和图9所示。



(a) T=0.25 s (b) T=0.50 s (c) T=0.75 s (d) T=1.00 s

(e) T=0.25 s (f) T=0.50 s (g) T=0.75 s (h) T=1.00 s

图8 两种绳索模型对扭转运动的反馈(图(a)～图(d)为二次离散模型，图(e)～图(h)为有限段模型)

Fig. 8 Feedback on torsional motion from two rope models (Fig. (a)～Fig. (d) for quadratic discrete model, Fig. (e)～Fig. (h) for model with line segments)

如图10和图11所示，一方面，两种模型中各节点的转动角度基本一致；另一方面，二次离散模型中的绳索总长度会随着绳索的扭转而不断缩减，有限段模型中的绳索总长并不会因扭转而发生变化。由此可知，二次离散模型对绳索扭转的表述更加符合实际。

绳索长度随扭转而缩减的原因是，二次离散模型中的各个节点在z轴方向产生了拉力。以节点1为例，从图11中可以看出节点1在z轴方向的受力变化，节点1的受力在0.018 s前快速增长，0.147 s后开始缓慢减小。

二次离散模型的仿真结果与传统绳索模型相比更加贴合实际，实际绳索发生扭转后长度同样会缩减。其原理是绳索发生扭转后，绳索内部的纤维会呈螺旋状排列，螺旋状的结构导致绳索总长发生缩减。

2.3 拉伸及弯曲分析

算例2旨在验证模型是否能够描述绳索受拉伸和弯曲的情况。对二次离散模型和质点模型中的节点1、节点10分别施加了3 cm/s的移动速度，速度方向沿x轴正向，仿真迭代步长选取0.002 s，仿真得到绳索位形如图12所示。通过对比二次离散模型中各个节点的质心坐标与集中质量法中的质点坐标，两种模型得到的仿真结果基本一致，区别在于二次离散模型包含绳索各节点弯曲程度，如图12(a)所示，随着绳索两端弯曲，内部各节点均产生了不同程度的弯曲。对比以上仿真结果可知，二次离散模型能够准确描述绳索承受拉伸和弯曲载

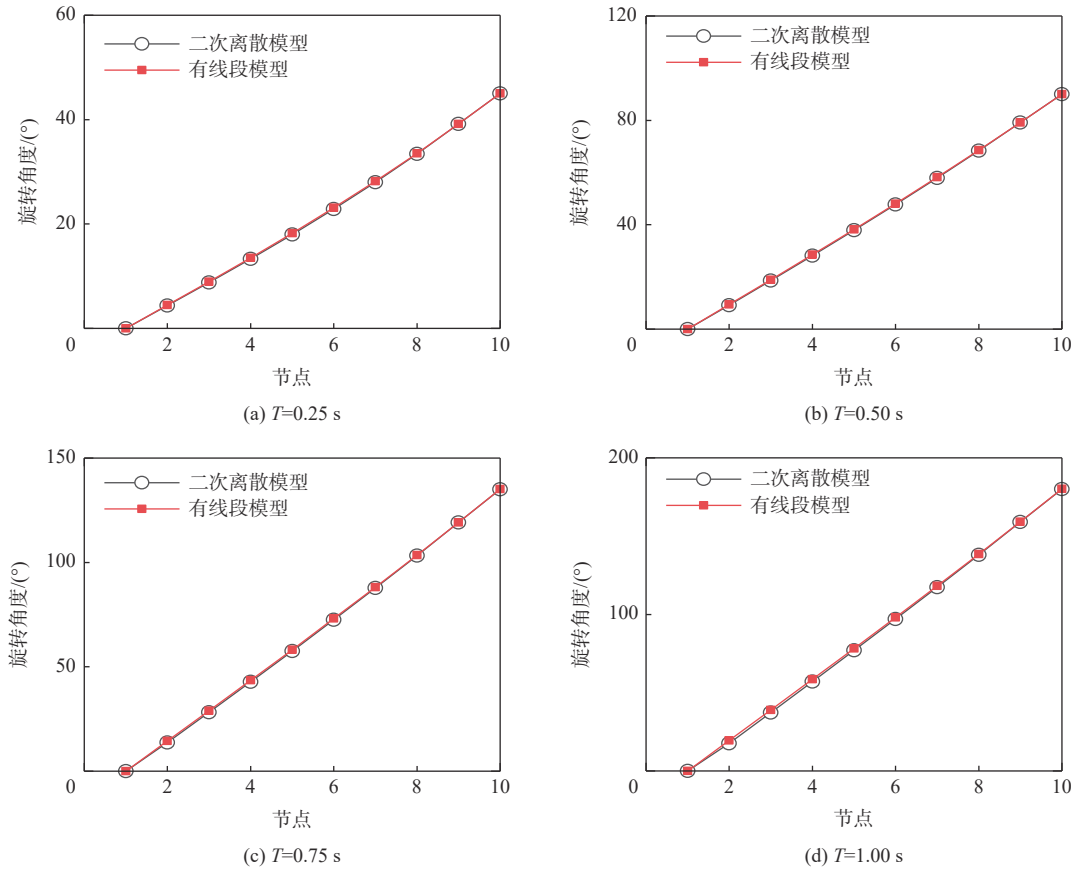


图 9　各节点绕 z 轴的转角

Fig. 9　Rotation angle around z axis for each node

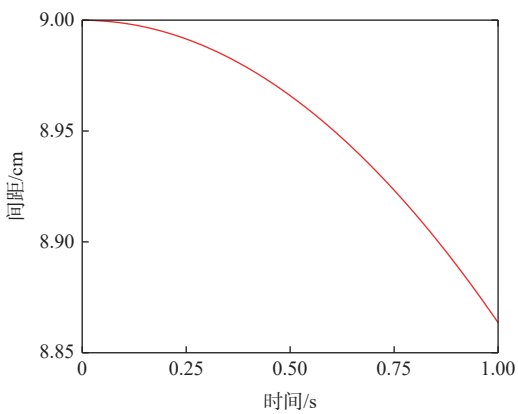


图 10　节点 1 与节点 10 的间距随时间变化情况

Fig. 10　Distance between node 1 and node 10 varies with time

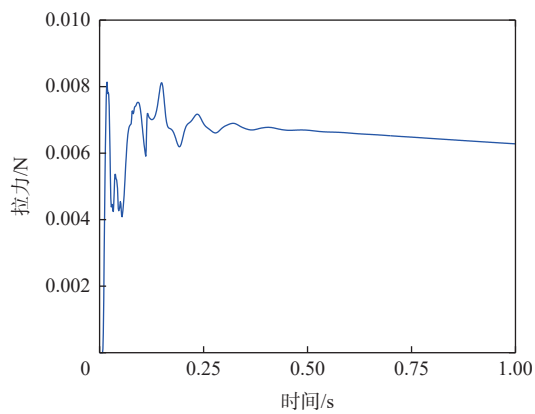


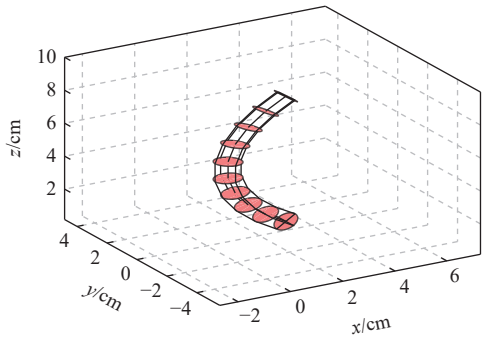
图 11　节点 1 在 z 轴方向上的拉力变化情况

Fig. 11　Tension in z direction varies with time for node 1

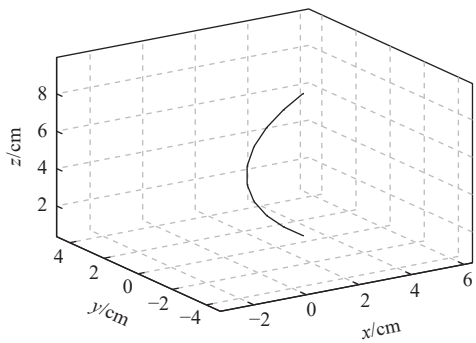
荷时的动力学响应。

2.4　绳索扭转实验

为了验证模型数值计算结果是否正确, 本节进行了绳索扭转的仿真与实验对比。实验平台由固定器、拉力测量仪及绳索组成, 其中绳索一端由固定器固定、另一端与测量仪连接。实验过程: 将绳索拉伸成直线, 转动测量仪端的绳索, 每转动 18°施加外力使绳索恢复初始长度, 随后对测量仪数值进行记录(如图 13(a))。仿真条件: 节点 1 完全固定, 固定节点 2～10 的 3 个移动自由度, 固定节点 10 的 3 个移动自由度, 并对节点 10 施加与实验对应的初始角度, 记录节点 1 的受力变化。仿真与实验误差对比如图 13(b) 所示, 绳索产生的拉力初始变化较小, 其增速不断提高。考虑到测量本身存在一定的误差, 这个误差是可以接受的, 该模型能够描述绳索承受扭转载荷时的动力学响应。从图 13(b) 中也可以看出, 绳索在端点固定时, 其产生的拉力会随着绳索扭转而迅速增加。



(a) 二次离散模型在 T=1 s



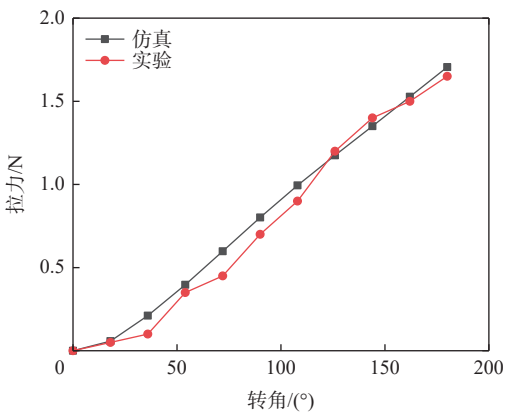
(b) 集中参数模型在 T=1 s

图 12 两种绳索模型受拉伸和弯曲载荷时的反馈

Fig. 12 Feedback for tensile and bending loads for two rope models



(a) 测量结果



(b) 仿真与实验误差

图 13 扭转实验验证

Fig. 13 Torsion experiment verification

3 飞网系统动力学仿真及实验验证

本章节对飞网展开过程进行了仿真与实验分析，旨在验证绳索二次离散模型在飞网系统动力学仿真中的有效性。飞网系统完成捕获后进入拖曳阶段，此时飞网系统与捕获物会通过连接绳构成组合体，本章对飞网系统捕获后的组合体进行了仿真分析，仿真假设空间碎片具有一定的旋转速度，以此观察空间碎片的自旋行为对飞网系统整体稳定性的影响。

3.1 展开阶段仿真与实验

实验系统由气动网枪及相机组成，牵引块的位置信息根据摄像机位置及参照物尺寸计算。网体展开程度通过对角牵引块之间的水平距离 D_v 评定[19]，计算公式如下：

$$D_v = \frac{C_v}{C_r} D_r \tag{35}$$

式中，D_r 为参照物的实际大小，C_v、C_r 分别为图片中的对角牵引块的间距和参照物大小。飞网展开仿真中不计算空气阻力影响。绳网材料为尼龙，单根绳索直径为 0.1 cm，飞网系统仿真参数见表 1。

表 1 飞网展开过程仿真参数

Table 1 Simulation parameters of tether-net deployment process

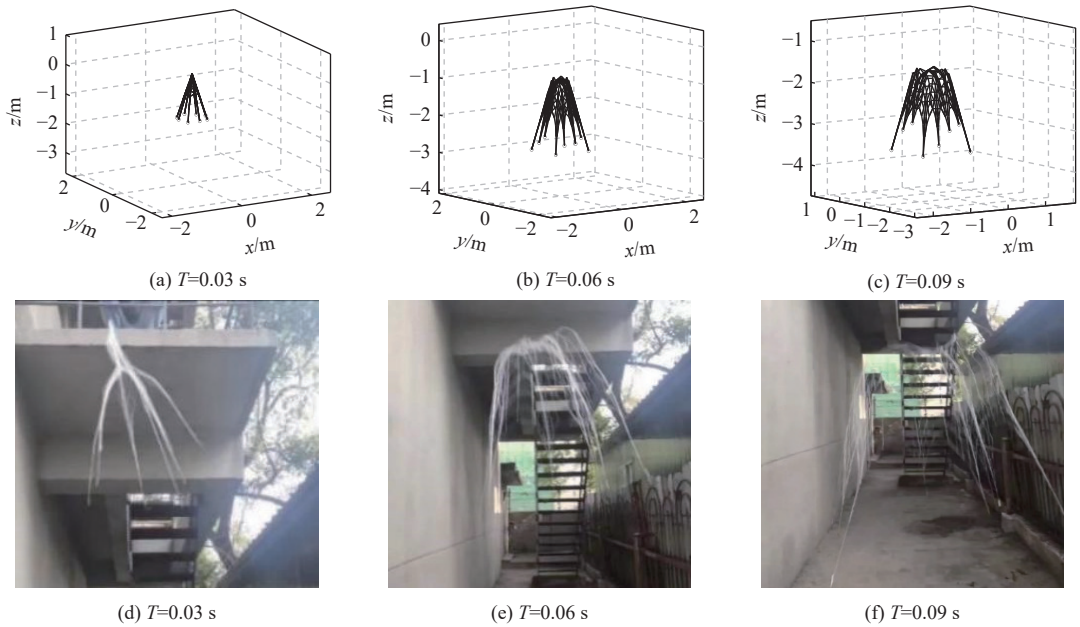
飞网网体大小/m	发射角度/(°)	发射速度/(m·s^{-1})	牵引块质量/kg	牵引块数量
3×3	30	50	0.3	8

仿真与实验对比结果如图 14 和图 15 所示。图 14 显示了实验与仿真过程中网体的展开情况，从图中可以看出，仿真中的网体位形与实验结果基本一致。分析图 15 可知，实验和仿真得到的 D_v 最大差距为 0.068，该误差可能来源于空气阻力、风阻及测量精度等。因该误差大小与集中质量法的误差接近，故认为该模型适用于飞网的动力学仿真模拟，能够适用于组合体的拖曳仿真模拟。

3.2 组合体拖曳仿真分析

本节仿真模拟了飞网系统捕获目标后的拖曳阶段，此阶段飞网与空间碎片（捕获物）通过连接绳构成组合体，仿真分析了自旋目标对主体卫星的影响，仿真参数设置如表 2 所示，仿真结果如图 16 和图 17 所示。

从图 17(a) 中可以看出，主体卫星与捕获物的间距不断缩短，缩短速度在初始时较快，随后趋于匀速，随着间距的缩短主体卫星与捕获物将在 1 591.07 s 后发生碰撞。主体卫星与捕获物的间距缩短的主要原因是绳索的扭转运动（绳索扭转导致长度缩减的现象已在第 2 章中证明）。绳索扭转产生的拉力如图 17(b) 所示，拉力在 0.3 s 所示达



(a) T=0.03 s　(b) T=0.06 s　(c) T=0.09 s

(d) T=0.03 s　(e) T=0.06 s　(f) T=0.09 s

图 14　飞网展开过程仿真与实验对比 ((a)～(c) 为仿真, (d)～(f) 为实验)

Fig. 14　Comparison between simulation and experiment of tether-net deployment process ((a)～(c) for simulation, (d)～(f) for experiment)

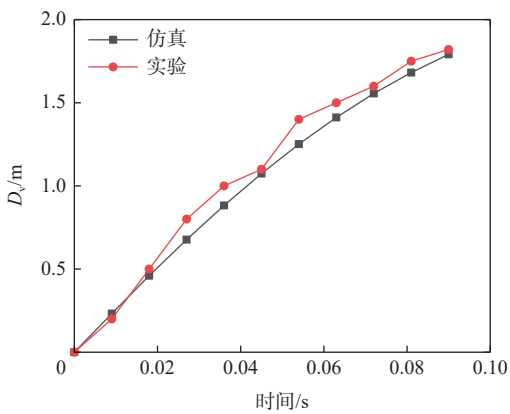


图 15　仿真与实验的D_v变化

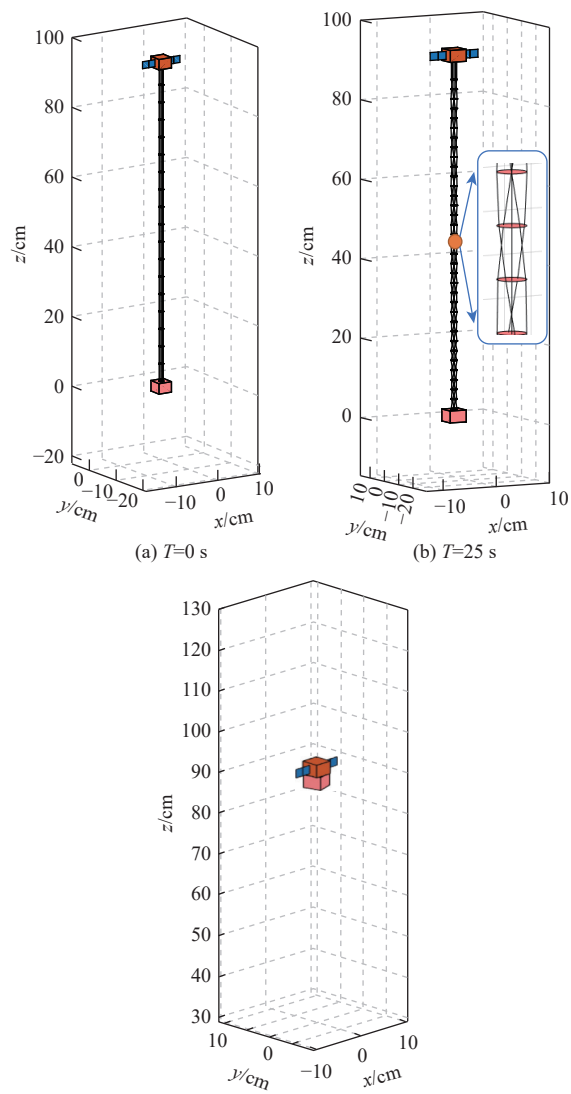
Fig. 15　Variation of D_v from simulation and experiment

表 2　组合体仿真参数

Table 2　Simulation parameters of combination

绳索材料	绳索直径/cm	拖曳物重量/kg	连接绳长度/cm	碎片转速/(r·s^{-1})
尼龙	0.5	2	90	2

到最大, 后因间距的缩短拉力减小至 0, 随着扭转角度的累加, 拉力进行了 3 次波动, 最终在 5.64 s 后消失。图 17(c) 为承载飞网系统的卫星受绳索施加的扭矩随时间的变化情况。图 17(d) 为空间碎片旋转速度随时间的变化情况, 从图中可以看出捕获物的旋转速度随时间的延长不断降低, 在 25 s 时其的旋转速度趋近于 0, 由此可以证明飞网系统本身具备一定的消旋能力。根据以上结果可知, 绳索的扭转行为对飞网系统的稳定性具有较大的影响, 忽略绳索的扭转行为将会增大仿真结果的误差。

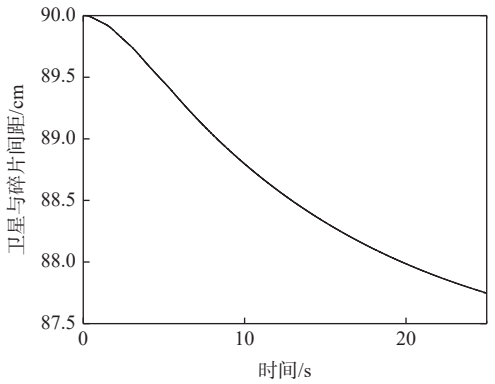


(a) T=0 s　(b) T=25 s

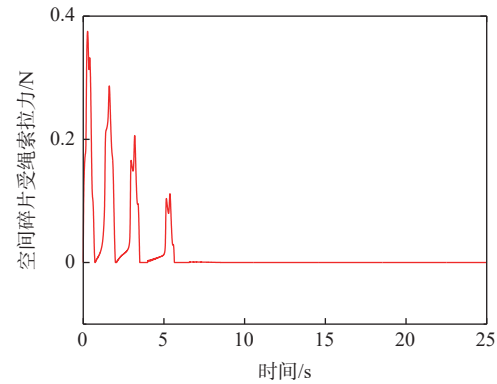
(c) T=1 591.07 s

图 16　组合体仿真 3D

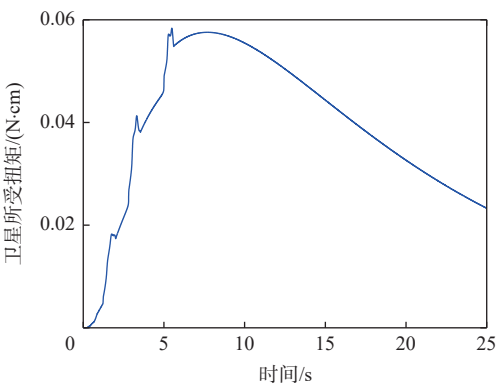
Fig. 16　Composite simulation 3D



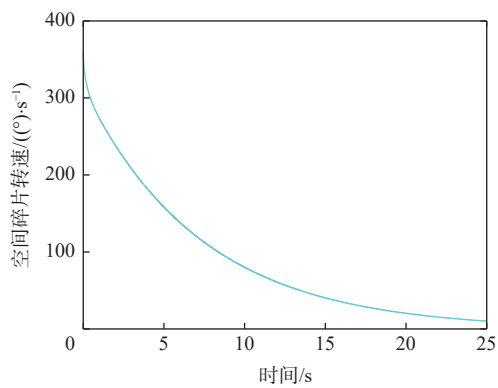
(a) 卫星与空间碎片间距随时间的变化情况



(b) 空间碎片受绳索拉力随时间的变化情况



(c) 卫星所受绳索施加的扭矩随时间的变化情况



(d) 空间碎片旋转速度随时间的变化情况

图 17 组合体仿真数值结果

Fig. 17 Numerical results of combination simulation

4 结 论

本文通过二次离散的方式构建了考虑绳索扭转行为的动力学方程，该模型可以完整地表达绳索各节点的位形信息。通过对单段绳索、飞网网体及捕获目标后的组合体进行仿真分析，并将仿真结果与传统绳索动力学模型或实验进行对比，得出以下结论：

1）二次离散模型可以将绳索的转动与绳索内力相关联，并且可以表述绳索长度因扭转而缩减的现象，这是传统绳索动力学所不具备的。

2）该模型对飞网展开过程进行仿真的误差为6.8%，误差大小与飞网传统动力学模型的误差接近，因此二次离散模型适用于飞网系统仿真。

3）二次离散模型能够准确描述绳索的扭转、拉伸及弯曲行为。

4）应用二次离散模型对组合体仿真的结果证明，带有自旋行为的空间碎片会通过连接绳索对卫星施加一定的扭矩，扭转的绳索会产生一定的拉力，从而导致卫星与捕获物的间距缩短，且在无其他设备辅助的情况下，飞网本身具有一定的消旋能力，忽略绳索的扭转行为会导致仿真结果与实际情况产生较大偏差。

参考文献（References）

[1] KESSLER D J, COUR-PALAIS B G. Collision frequency of artificial satellites: the creation of a debris belt[J]. Journal of Geophysical Research: Space Physics, 1978, 83(A6): 2637-2646.

[2] LIOU J C. An active debris removal parametric study for LEO environment remediation[J]. Advances in Space Research, 2011, 47(11): 1865-1876.

[3] FLORES-ABAD A, MA O, PHAM K, et al. A review of space robotics technologies for on-orbit servicing[J]. Progress in Aerospace Sciences, 2014, 68: 1-26.

[4] AGLIETTI G S, TAYLOR B, FELLOWES S, et al. The active space debris removal mission RemoveDebris. Part 2: in orbit operations[J]. Acta Astronautica, 2020, 168: 310-322.

[5] BISCHOF B, KERSTEIN L. ROGER-robotic geostationary orbit restorer[C]// 34th COSPAR Scientific Assembly. Florida Miami: NASA Astrobiology Institute General, 2003.

[6] AGLIETTI G S, TAYLOR B, FELLOWES S, et al. RemoveDEBRIS: an in-orbit demonstration of technologies for the removal of space debris[J]. The Aeronautical Journal, 2020, 124(1271): 1-23.

[7] LAVAGNA M, ARMELLIN R, BOMBELLI A, et al. Debris removal mechanism based on thethered nets[C]// International Symposium on Artificial Intelligence. Piedmon Torino: Robotics and Automation in Space, 2012.

[8] 陈钦, 杨乐平, 张青斌. 空间飞网发射动力学建模仿真研究与地面试验[J]. 国防科技大学学报, 2009, 31(3): 16-19.
CHEN Q, YANG L P, ZHANG Q B. Dynamic model and simulation of orbital net casting and ground test[J]. Journal of National University of Defense Technology, 2009, 31(3): 16-19(in Chinese).

[9] SHAN M H, GUO J, GILL E. Deployment dynamics of tethered-net

for space debris removal[J]. Acta Astronautica, 2017, 132: 293-302.

[10] SHABANA A A, YAKOUB R Y. Three dimensional absolute nodal coordinate formulation for beam elements: theory[J]. Journal of Mechanical Design, 2001, 123(4): 606-613.

[11] BERZERI M, SHABANA A A. Development of simple models for the elastic forces in the absolute nodal co-ordinate formulation[J]. Journal of Sound Vibration, 2000, 235(4): 539-565.

[12] OMAR M A, SHABANA A A. A two-dimensional shear deformable beam for large rotation and deformation problems[J]. Journal of Sound Vibration, 2001, 243(3): 565-576.

[13] STADNYK K, ULRICH S. Validating the deployment of a novel tether design for orbital debris removal[J]. Journal of Spacecraft and Rockets, 2020, 57(6): 1335-1349.

[14] BENVENUTO R, LAVAGNA M, CINGOLI A, et al. Multibody dynamics simulation tool to support the GNC design for active debris removal with flexible elements[C]// 9th International ESA Conference on Guidance Navigation and Control Systems. Haute-Garonne Toulouse: European Space Agency, 2014.

[15] SI J Y, PANG Z J, DU Z H, et al. Dynamics modeling and simulation of a net closing mechanism for tether-net capture[J/OL]. International Journal of Aerospace Engineering. 2021(2021-08-06) [2022-08-18]. https://doi.org/10.1016/j.asr.2019.08.006.

[16] ENDO Y, KOJIMA H, TRIVAILO P M. Study on acceptable offsets of ejected nets from debris center for successful capture of debris[J]. Advances in Space Research, 2020, 66(2): 450-461.

[17] ZHANG G B, ZHANG Q B, FENG Z W, et al. A simplified model for fast analysis of the deployment dynamics of tethered-net in space[J]. Advances in Space Research, 2021, 68(4): 1960-1974.

[18] 潘冬, 张越, 魏承, 等. 空间大型末端执行器绳索捕获动力学建模与仿真[J]. 振动与冲击, 2015, 34(1): 74-79.
PAN D, ZHANG Y, WEI C, et al. Dynamic modeling and simulation on rope capturing by space large end effector[J]. Journal of Vibration and Shock, 2015, 34(1): 74-79(in Chinese).

[19] 高庆玉. 空间绳网系统展开动力学与优化设计[D]. 长沙: 国防科技大学, 2017.
GAO Q Y. Deployment dynamics and optimization design of space net system[D]. Changsha: National University of Defense Technology, 2017(in Chinese).

[20] HOVELL K, ULRICH S. Postcapture dynamics and experimental validation of subtethered space debris[J]. Journal of Guidance Control Dynamics, 2018, 41(2): 519-525.

[21] HOU Y S, LIU C, HU H Y, et al. Dynamic computation of a tether-net system capturing a space target via discrete elastic rods and an energy-conserving integrator[J]. Acta Astronautica, 2021, 186: 118-134.

[22] 张国斌. 新型空间柔性捕获机构动力学研究[D]. 长沙: 国防科技大学, 2018.
ZHANG G B. Dynamics of the novel space capture mechanism[D]. Changsha: National University of Defense Technology, 2018(in Chinese).

[23] SHAN M, GUO J, GILL E. An analysis of the flexibility modeling of a net for space debris removal[J]. Advances in Space Research, 2020, 65: 1083-1094.

[24] GOMEZ N O, WALKER S. Earth's gravity gradient and eddy currents effects on the rotational dynamics of space[J]. Debris Objects: Envisat Case Study. Advances in Space Research, 2015, 56(3): 494-508.

[25] 沈剑, 韩峰, 陈放, 等. 基于绝对节点坐标法的柔性绳索体建模仿真研究[J]. 科学技术与工程, 2016, 16(31): 1-7.
SHEN J, HAN F, CHEN F, et al. Modeling and simulation of flexible cable-body by absolute nodal coordinate formulation[J]. Science Technology and Engineering, 2016, 16(31): 1-7(in Chinese).

[26] 刘延柱, 潘振宽, 戈新生. 多体系统动力学[M]. 2 版. 北京: 高等教育出版社, 2014: 33-48.
LIU Y Z, PAN Z K, GE X S. Dynamics of multibody systems[M]. 2nd ed. Beijing: Higher Education Press, 2014: 33-48(in Chinese).

Dynamic modelling and simulation of a tethered-net in space

CHEN Pengxu[1, 2], WU Chenchen[1, 3, *], NI Zhiyu[4]

(1. Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110017, China;
2. School of Mechanical Engineering, Shenyang University of Technology, Shenyang 110870, China;
3. School of Physics Nanjing University of Science and Technology, Nanjing 210094, China;
4. School of Aeronautics and Astronautics Shenyang Aerospace University, Shenyang 110136, China)

Abstract: Once the space tether-net captures the spinning space debris, a torque will be exerted on the connecting tether between the capture net and the satellite. However, the current tether-related dynamic model can't fully describe its twisting behavior. In order to address this issue, a quadratic discrete model is constructed using the lumped parameter method to accurately depict the twisting behaviors of the tether. Comparison with the traditional dynamic model and a torsional experiment for the tether verify that the quadratic discrete model can describe the responses of the tether subjected to tension, bending or torsion. It is found that the torsion of the tether will produce additional traction. Furthermore, the deployment procedure of the tether-net is subjected to simulation analysis and experimentation, which confirms the suitability of the dynamic model for the tether-net system. The simulated examination of the tether-net and space debris combo reveals that the torsional force generated by the rope's twisting will affect the stability of the primary satellite to some extent. Ignoring the twisting behavior of the tether will reduce the accuracy of the dynamic simulation model.

Keywords: tether dynamics; tether-net; space debris; lumped parameter method; multi-body dynamics

Received: 2022-08-30; **Accepted**: 2023-01-10; **Published Online**: 2023-02-08 10:36

URL: link.cnki.net/urlid/11.2625.V.20230207.1705.003

Foundation items: National Natural Science Foundation of China (51905527); Funding of Shenyang Institute of Automation, Chinese Academy of Sciences; Fundamental Research Funds for the Central Universities; Basic Scientific Research Project of Education Department of Liaoning Province (JYTMS20230254); Basic Science Center Program of the National Natural Science Foundation of China (62388101)

* **Corresponding author**. E-mail: wuchenchen@njust.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0755

多网络控制器 FC-AE-1553 动态带宽调度机制

李心茹[1]，张馨月[1]，李铮[2,3]，何锋[1,*]

(1. 北京航空航天大学 电子信息工程学院，北京 100191；2. 中国科学院 空天信息创新研究院，北京 100190；
3. 中国科学院 空间信息处理与应用系统技术重点实验室，北京 100190)

摘 要：由于航电系统通信需求的日趋复杂化，对基于FC-AE-1553组网的通信系统需要采用多网络控制器（NC）构建分布式网络环境，如何实施多个控制器之间的控制权分配，并解决跨域业务冲突的问题是其中的关键。为实现多NC协同调度，提出了一种主从NC协作的多NC动态带宽调度机制，采用并发通信方式提高网络带宽利用率，同时提升各类业务的传输保障能力。对于周期性业务采用固定时隙进行传输，对于强时效性突发业务采用抢占式带宽分配，对于一般突发业务按照“先跨域后域内”分别进行带宽分配。通过OMNeT++搭建FC-AE-1553网络仿真平台，结果表明：该调度机制可以满足多NC条件下不同类型通信业务的协同带宽分配和调度，对于周期性业务和强时效性突发业务，可以保障其实时性通信要求；对于一般突发业务，在网络负载不超过57%时也能为其提供合适的带宽进行传输。与单NC集中式相比，当网络负载大于0.5或跨域/域内业务比值大于0.8时，多NC控制具有较明显的时延优势。

关 键 词：FC-AE-1553；多网络控制系统；动态调度；控制权分配；跨域传输
中图分类号：V243.1
文献标志码：A **文章编号**：1001-5965(2024)09-2963-12

随着空间探索和应用的不断发展，航天综合电子系统的规模和复杂性不断增大，对网络通信系统的带宽、可靠性、实时性和确定性等方面提出了更高的要求。传统的联合式电子系统采用MIL-STD-1553总线协议[1-3]，传输速率1.0 Mbit/s，可同时连接31个远程终端，具有较高的可靠性和较强的实时性。但其传输速度和组网规模严重制约了航天综合电子系统的发展。光纤通道（fibre channel，FC）的高传输速率、低延时、低误码、开放标准等特点，使其能够很好地满足新一代航天综合电子系统互联的要求，并有专门针对机载环境的FC-AE协议增强子集，FC-AE-1553便是其中的一种，具备大带宽和多点通信能力[4-7]，在综合电子系统领域有着广泛的应用前景。

随着航天综合电子系统各种入网设备的增多，FC-AE-1553一级交换网络远不能满足大规模场景下的需求，网络结构逐渐呈现多级、分布式的特点。在FC-AE-1553多级交换网络中，可以存在多个网络控制器，终端之间可以并行传输数据[8]。网络中的业务通信不仅存在于同一个交换域的不同网络终端之间，也会存在于不同交换域的两个网络终端之间，每个交换域都可能存在自身域内的网络控制节点。相比于一级交换网络，多级交换网络的业务传输和控制更为复杂，需要考虑多个网络控制节点控制权分配以及域间、域内业务传输冲突问题，因此需要一种多网络控制器下的多级交换网络的优化调度方案。

传统的FC-AE-1553总线式网络，采用大小周

收稿日期：2022-09-02；**录用日期**：2023-02-10；**网络出版时间**：2023-03-13 16:46
网络出版地址：link.cnki.net/urlid/11.2625.V.20230310.2238.004
基金项目：国家自然科学基金 (62071023)
*通信作者. E-mail：robinleo@buaa.edu.cn

引用格式：李心茹，张馨月，李铮，等. 多网络控制器FC-AE-1553动态带宽调度机制[J]. 北京航空航天大学学报，2024，50(9)：2963-2974. LI X R，ZHANG X Y，LI Z，et al. FC-AE-1553 dynamic bandwidth scheduling mechanism for multi network controllers[J]. Journal of Beijing University of Aeronautics and Astronautics，2024，50(9)：2963-2974 (in Chinese).

期调度机制，需要根据数据通信任务的最大和最小发送周期以及允许的业务传输延迟来确定调度表的安排，大小周期方法没有充分利用FC-AE-1553可以并发传输的特性，不能提高网络的吞吐量，网络带宽利用率较低[9-11]。存在一些改良方案[12]，基于大小周期机制，文献[12]提出了一种数据传输控制方法，与传统的大小周期机制相比，带宽利用率有所提高，但同样没有利用FC-AE-1553可并发传输的特性。

针对FC-AE-1553一级交换式网络，文献[13-14]提出了一种“并发交换式”动态带宽分配(dynamic bandwidth allocation, DBA)机制，一般突发业务采用并发传输方式动态分配带宽，周期性业务采用固定带宽分配，强时效性业务抢占式分配带宽，在一定程度上提高了网络吞吐量。但突发业务传输阶段，每个网络终端(network terminal, NT)节点在当前DBA周期只产生去往一个目的节点的业务请求，带宽利用率较低。类似的，文献[15]提到了一种带宽等比例分配(LE-PA)机制，在每个DBA周期，网络控制器(network controller, NC)轮询各节点，节点仅上报去往一个目的节点的一般突发业务请求。为解决上述问题，文献[15-16]针对一般突发业务提出了一种动态带宽调度机制，可以在每个NT同时面向多个目标节点进行业务传输的情况下，提高网络性能，同时避免链路访问冲突。针对无源光网络技术下的FC-AE-1553网络，文献[17]提出了基于信用值排序的动态带宽分配算法，发挥了FC-AE-1553高带宽传输优势，并确保了服务的平均时延，文献[18]提出了静态动态结合的带宽分配方式，为网络优化提供了有效的解决方案。文献[19-20]提出了支持IP服务的FC-AE-1553调度机制，保证了多协议下的IP数据传输效率。

针对FC-AE-1553多级交换式网络，文献[21]设计了一种优先分配交换域间业务传输的集中控制调度机制，区分域间和域内的数据传输，可以降低域间数据交换时延，提高网络吞吐量。整个多级交换网络仅有一个NC，不存在NC控制权冲突的问题，不适用于网络中存在多个NC的场景。

针对多NC控制的FC-AE-1553多级并发交换式网络，本文提出了一种多NC控制权分配方式，依照交换式网络结构，将NC划分为主NC和从NC，网络业务划分为域内传输和跨域传输两类，根据网络业务类型，由相应NC的进行合理调度分配，主从NC相互配合，实现多NC控制下的网络带宽分配无冲突。与单NC集中式带宽分配相比，多NC控制下的业务调度可以减轻单个NC集中调度的压力，降低业务时延，顺应当前网络架构分布式发展方向。

1 多级FC交换网络多NC调度机制

1.1 交换式网络架构

针对一级FC-AE-1553交换网络不能满足大规模组网需求的问题，构建了具有多个网络控制节点的FC-AE-1553多级交换网络拓扑架构，如图1所示。

该多级网络由多个NC、主/分交换机以及NT构成，由主、分交换机为核心单元构成“核心-接入”多级FC-AE交换式网络，由不同的网络控制器控制不同的FC-AE交换域。其中，主交换机1和主交换机2共同构成核心FC交换网络，两者互为备份关系；分交换机自身构成多个接入FC交换网络。每个交换机和与它直接相连的NC、NT构成一个交换域，在结构上各交换域相互独立。

航空电子系统的业务数据类型可以分为周期性业务(periodic traffic, PT)、一般突发业务(com-

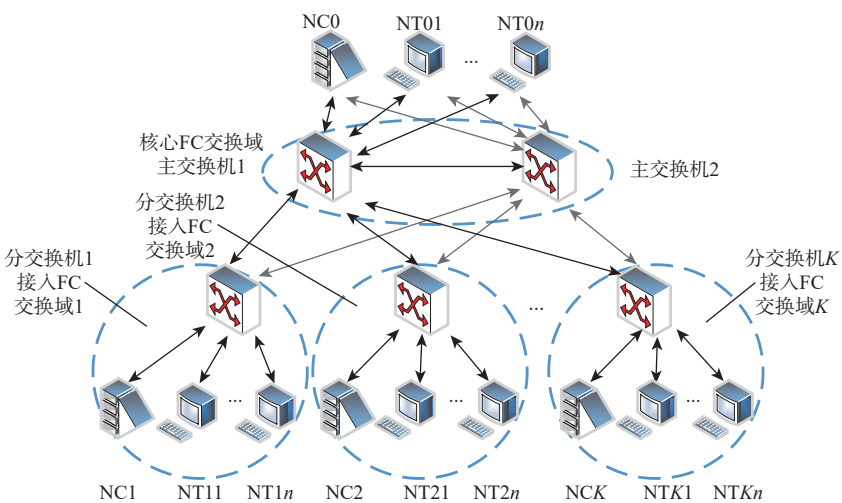


图1 多级FC交换网络多NC架构

Fig. 1 Multi-level FC switching network Multi-NC architecture

mon burst traffic, CBT)和强时效性突发业务(time constraint burst traffic, TCBT)3 种。周期性业务在 NC 和 NT 之间周期性传输, 如遥测数据、供电参数等; 一般突发业务指网络中某时间段内产生的大量突发数据; 强时效性突发业务在网络中突发产生, 且具有很强的实时性要求, 例如由网络控制器产生的各种应用指令。

在该多级交换式网络中, 任意两个节点之间均可以进行数据传输, 且均通过 NC 的命令响应式控制机制完成。不失一般性[13-17,19,21], 周期性业务仅存在于每个 NC 与自身交换域内的 NT 节点之间; 一般突发业务存在于任意两个 NT 节点之间, 例如视频类信息传递、遥测数据传输、实验数据转发存储等[21]; 强时效性突发业务由任意 NC 产生, 发送给域内任意 NT 节点。通过多个 NC 节点合理调度, 多对节点之间可以进行数据的并行传输, 即多个数据传输任务可以同时执行, 从而提高网络的带宽利用率。

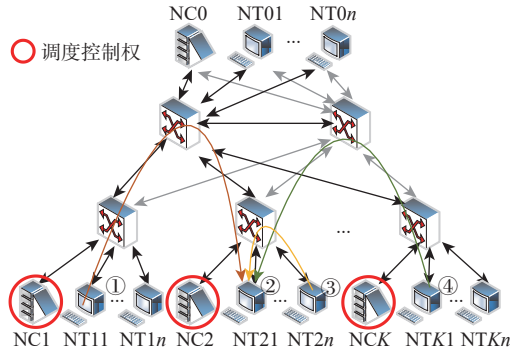
基于单 NC 网络业务定长时隙调度算法可知: 周期性业务可以分配固定时隙, 强时效性突发业务进行抢占式调度, 一般突发业务采用基于首次适应装箱算法的调度机制[21]。对于图 1 的多 NC 交换网络, 周期性业务和强时效性突发业务发生在各交换域内, 可由各 NC 独立进行带宽分配, 互不干扰, 与单 NC 网络相似。一般突发业务涉及跨域问题, 需要考虑如何解决网络多 NC 控制下的控制权分配以及域间和域内业务冲突问题。

1.2　多 NC 控制权分配

根据网络中业务传输涉及的交换域, 可以将其分为跨域传输和域内传输两种。由于各交换域在结构上相互独立, 因此各个域内传输可以并行进行。在交换域内, 各 NC 独立控制自身交换域内的业务, 发送节点和接收节点均不相同的业务可以并发传输, 而不产生业务冲突。

对于跨域业务, 多个 NC 同时进行调度时无法及时获取其他交换域的业务传输情况, 极易产生数据冲突, 如图 2 所示。在图 2(a) 中, 由于各 NC 均拥有域内节点调度控制权, 因此业务调度相互独立。如 1.1 节所提到的, 节点①、③、④可能在同一时段被域内 NC 调度, 向节点②进行视频类或其他类型业务传输, 占用大量带宽, 产生数据冲突。

时隙占用情况如图 2(b) 所示, 发送时间轴的色块表示相应节点传输业务所用时隙, 接收时间轴的色块表示目的节点接收业务所用时隙。由图 2(b) 可知, 节点②的接收时间轴在时隙 0 和时隙 1 产生



(a) 业务传输

时隙	发送时间轴				接收时间轴			
	①	②	③	④	①	②	③	④
0								
1								
2								
3								
4								
5								

(b) 时隙冲突

图 2　跨域及域内业务冲突示意

Fig. 2　Diagram of cross-domain and intra-domain service conflicts

业务冲突。

为解决这一问题, 选取图 1 中与核心 FC 交换网络直接相连的 NC 节点作为该多级网络的主 NC(master network controller, MNC), 与接入 FC 交换网络直接相连的 NC 作为从 NC(slave network controller, SNC), 如图 3(a) 所示。

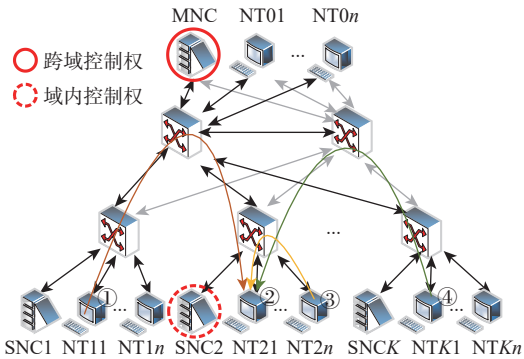
各 SNC 向 MNC 上报自身交换域内的跨域业务请求, 由 MNC 进行跨域数据传输的统一调度和带宽分配, 避免目的节点相同的跨域业务在相同时隙发送, 造成目的节点处的业务冲突。SNC 依据 MNC 生成的跨域调度表, 控制域内业务数据带宽分配和调度。

由此生成的全网业务调度时隙分配表如图 3(b) 所示, 相同目的节点的业务发送时间相互错开, 接收时间轴无冲突产生。

1.3　域间及域内调度表生成

采用 MNC 统一调度跨域业务, SNC 分别调度域内业务的方式, 将网络中的业务数据传输分为跨域传输和域内传输, 全网会存在两张调度表: MNC 生成的全网跨域业务调度表以及各 SNC 生成的自身交换域业务调度表。两张调度表若独立生成, 则会存在跨域及域内业务冲突问题。

综合考虑全网的链路共享率, 在业务类型复杂以及流量需求较大的情况下, 全网各链路带宽大小相同时, 多级交换网络的业务传输能力受到交换域



(a) 控制权分配

时隙	发送时间轴				接收时间轴			
	①	②	③	④	①	②	③	④
0								
1								
2								
3								
4								
5								
6								

(b) 时隙分配情况

图 3 业务调度示意

Fig. 3 Service scheduling diagram

间各交换机互连的链路带宽资源大小限制[21]。针对多级交换网络跨域及域内业务带宽分配冲突问题，采用优先分配跨域业务传输的方式，并将 MNC 跨域调度表下发给各个 SNC，SNC 在跨域调度的基础上选择空闲时隙进行自身域内业务调度，以此来避免域内和跨域业务冲突。

在每个 DBA 周期开始后，将首先按照固定时隙传输 PT 业务，随后进入 CBT 业务请求上报和时隙分配阶段。全网 NC 通过轮询的方式收集交换域内 NT 节点的 CBT 请求，每个 NT 节点可以上报去往所有 NT 节点的 CBT 请求，同一发送节点到同一目的节点的带宽请求最多上报 1 个。各 NC 将收集到的 CBT 带宽请求根据目的节点划分为域内或跨域请求，SNC 将跨域请求上报给 MNC。MNC 收集全网跨域请求，进行带宽分配，并将带宽分配结果下发给各 SNC。各主从 NC 基于跨域带宽分配的结果进行各自域内请求带宽分配，从而完成该轮 DBA 周期的 CBT 带宽分配。该多 NC 控制下的 CBT 带宽分配机制流程如图 4 所示。

综合考虑 PT 业务、CBT 业务和 TCBT 业务，一个完整的由 MNC 控制跨域传输、各 NC 独立控制

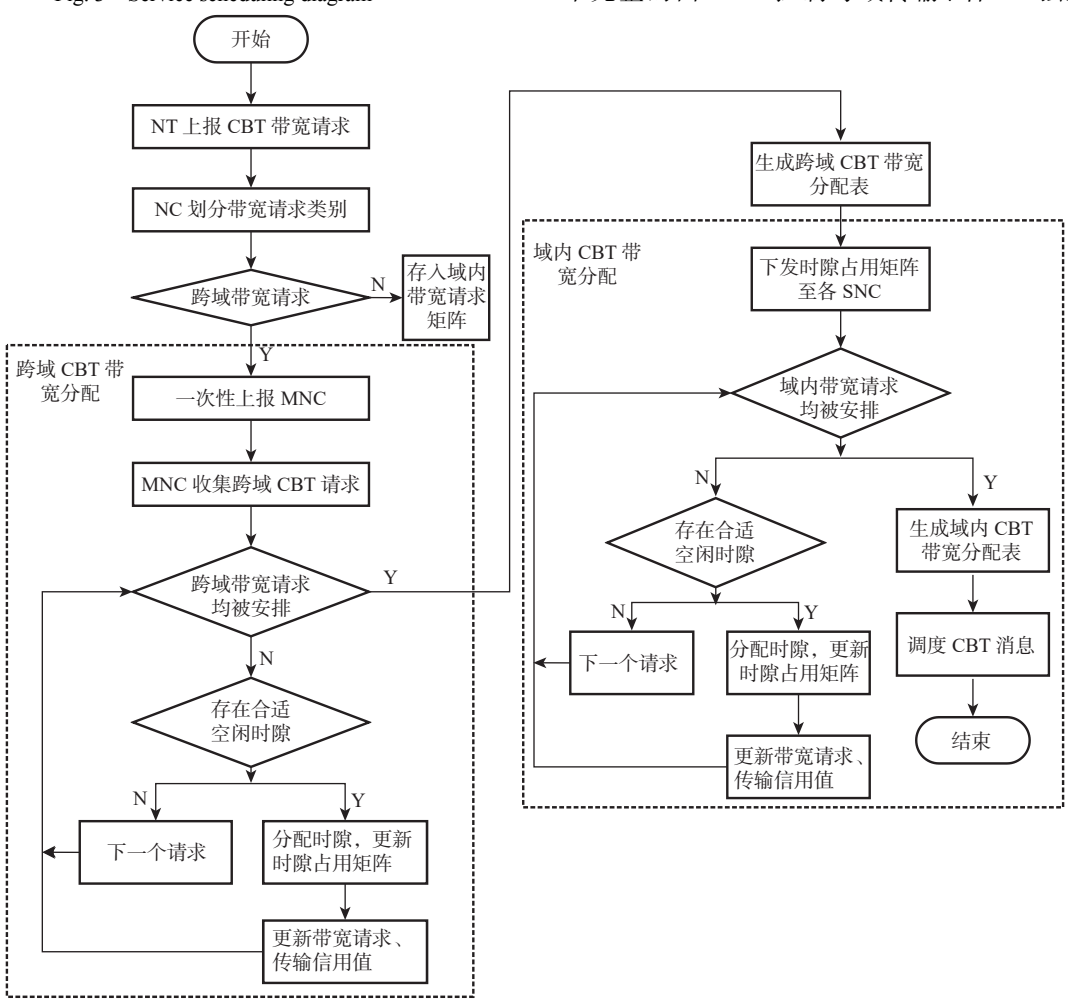


图 4 CBT 带宽分配机制流程

Fig. 4 CBT bandwidth scheduling mechanism flowchart

域内传输的 DBA 周期性业务调度机制时序如图 5(a)所示。在图 5(a) 的周期性业务时段, 各 NC 均拥有控制权, 控制自身交换域内的周期性业务进行无冲突传输; 在 CBT 业务传输阶段, MNC 具有跨域调度控制权, 负责全网的跨域业务以及自身域内业务调度, 各 SNC 具有域内调度控制权, 基于 MNC 跨域业务调度表控制域内业务传输。在最大化保障可调度性的同时, 规避了控制权分配不明以及业务调度冲突问题。图 5(b) 给出了两个交换域之间业务传输的收发节点时隙对应示意, 其余任意两个交换域之间的业务传输同理, 相同颜色表示同一业务信息。

1.4　多 NC 下的业务调度机制

明确了 NC 控制权分配以及域间域内业务冲突避免方式, 接下来需要考虑业务调度的具体方式。根据业务类别, 分别介绍如下。

1.4.1　周期性业务

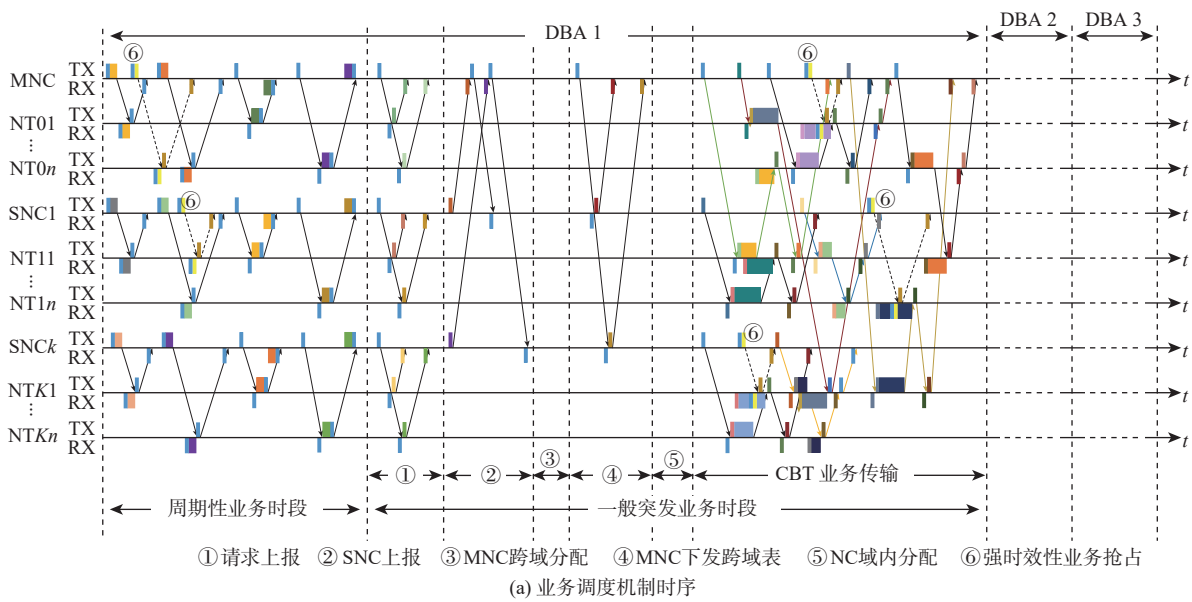
由于周期性业务在网络中周期产生, 且业务传输需要保证较为严格的周期性, 因此对于周期业务, 采用固定时隙的方式进行带宽分配。在每个动态带宽调度周期内, 为周期业务预留固定的时隙, 且为了保证业务的严格周期性不被干扰, 其他类型的业务不允许占用该段预留时隙[13-14,16,21]。

由于周期性业务仅存在于每个 NC 与自身交换域内的 NT 节点之间[13-15,17,19,21], 在当前 DBA 周期调度开始时, NC 节点首先进行自身发往 NT 节点的周期业务传输, 各交换域内可并行进行。随后, NT 节点不需要进行周期性业务请求上报, NC 节点可以依据已知的周期性业务需求, 向相应的 NT 节点发送周期性业务命令帧, 进行 NT 节点到 NC 节点的周期性业务传输。

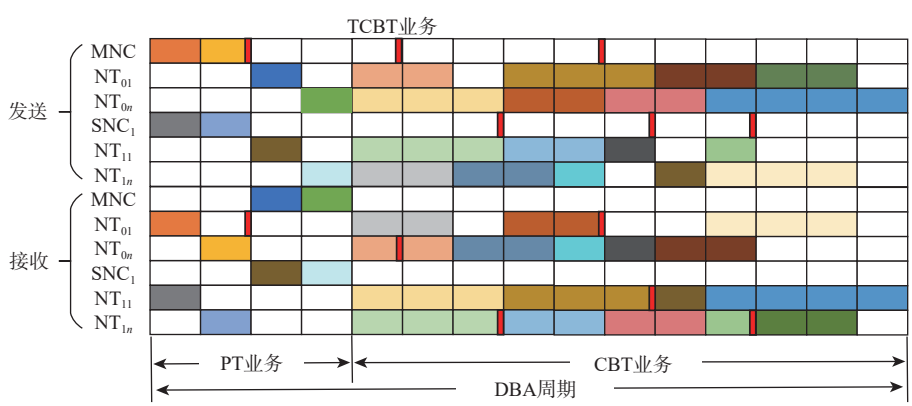
1.4.2　强时效性突发业务

强时效性突发业务的产生具有突发性强、业务量小和实时性要求高等特点, 且仅由 NC 节点产生, 发往其域内 NT 节点[13-14,16-17,21]。为了保证强时效性突发业务的实时性, 同时考虑到其业务大小, 可以采用抢占式的带宽分配方式。

对于周期性业务, 为了保证其严格的周期性, 仅可在两个周期业务之间预留一小段保护时隙, 用



(a) 业务调度机制时序



(b) 收发时隙对应示意

图 5　业务调度机制

Fig. 5　Service scheduling mechanism

来进行TCBT业务的发送。当NC节点产生TCBT业务后，会将该业务请求插入当前发送时隙的保护时隙。

对于一般突发业务，可将在其传输时段内产生的TCBT业务插入其帧间间隔内，不打断当前正在进行的数据帧传输。

因此，在进行网络调度的时隙大小设计时，需要考虑为TCBT业务预留保护时隙，以确保正在进行的数据传输不会因为TCBT业务插入而被干扰。

1.4.3 一般突发业务

一般突发业务是网络中占比较大的业务类型，需要进行合理设计，生成CBT业务调度表，在保证无CBT业务冲突的情况下最大化进行带宽分配。由于采用并发方式传输业务，网络中会同时存在多个点对点请求，可以将节点的收发时间轴按照定长时隙进行划分，类比装箱问题对带宽资源进行分配。

设全网节点数为N，则带宽请求矩阵$\boldsymbol{B}_{\mathrm{R}}$可以表示为

$$\boldsymbol{B}_{\mathrm{R}}=\begin{bmatrix} r_{00} & \cdots & r_{0(N-1)} \\ \vdots & & \vdots \\ r_{(N-1)0} & \cdots & r_{(N-1)(N-1)} \end{bmatrix} \tag{1}$$

式中：r_{ij}为发送节点i发往接收节点j的带宽请求，$i,j\in[0,N-1]$。

一个完整的DBA周期包括PT业务调度、CBT请求上报及分配、CBT业务调度3个阶段，将DBA周期除去前两个阶段，即可得到用于CBT业务分配的调度时隙长度。

设节点用于CBT业务调度的收发时间轴时隙占用矩阵为$\boldsymbol{S}_{\mathrm{T}}$，时隙数为$M$。对于一级交换式网络，仅需考虑域内消息交互，其域内业务的时隙占用矩阵$\boldsymbol{S}_{\mathrm{T}}$可表示为

$$\boldsymbol{S}_{\mathrm{T}}=\begin{bmatrix} t_{00} & \cdots & t_{0(N-1)} & t_{0N} & \cdots & t_{0(2N-1)} \\ \vdots & & \vdots & \vdots & & \vdots \\ t_{(M-1)} & \cdots & t_{(M-1)(N-1)} & t_{(M-1)N} & \cdots & t_{(M-1)(2N-1)} \end{bmatrix} \tag{2}$$

由于光纤通道可以同时接收和发送数据，因此网络中的链路均可以视为单向的发送轴和接收轴的组合，发送轴和接收轴可以独立进行时隙分配，对于N个节点，其发送时间轴和接收时间轴共$2N$条。即在矩阵$\boldsymbol{S}_{\mathrm{T}}$中，当$0\leqslant n\leqslant N-1$时，$t_{mn}$表示节点$n$的发送时间轴第$m$个时隙；当$N\leqslant n\leqslant 2N-1$时，$t_{mn}$表示节点$n$的接收时间轴第$m$个时隙，同一节点发送和接收数据互不干扰。对于$\forall m,n$，$0\leqslant m\leqslant M-1$，$0\leqslant n\leqslant 2N-1$，$t_{mn}$满足：

$$t_{mn}=\begin{cases}0, & \text{该时隙未被占用}\\ 1, & \text{该时隙被占用}\end{cases} \tag{3}$$

假设多级交换式网络有K个接入FC交换网络，若考虑跨域消息交互，则需要考虑交换域之间的链路空闲情况，可以将两个交换机之间的直连收发链路，视为分别以两个交换机为源端的发送链路的组合。即在图1中，核心FC交换域和接入FC交换域1之间的链路，可以看作“核心FC交换域→接入FC交换域1”与“接入FC交换域1→核心FC交换域”的组合。若接入交换域的个数为K，则发送链路总数为$2K$。跨域业务的时隙占用矩阵$\boldsymbol{S}_{\mathrm{T}}$可表示为

$$\boldsymbol{S}_{\mathrm{T}}=\begin{bmatrix} t_{00} & \cdots & t_{0(N-1)} & t_{0N} & \cdots & t_{0(2N-1)} & t_{0(2N)} & \cdots & t_{0(2N+K-1)} & t_{0(2N+K)} & \cdots & t_{0(2N+2K-1)} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ t_{(M-1)0} & \cdots & t_{(M-1)(N-1)} & t_{(M-1)N} & \cdots & t_{(M-1)(2N-1)} & t_{(M-1)(2N)} & \cdots & t_{(M-1)(2N+K-1)} & t_{(M-1)(2N+K)} & \cdots & t_{(M-1)(2N+2K-1)} \end{bmatrix} \tag{4}$$

当$0\leqslant n\leqslant N-1$时，$t_{mn}$表示节点$n$的发送时间轴第$m$个时隙；当$N\leqslant n\leqslant 2N-1$时，$t_{mn}$表示节点$n$的接收时间轴第$m$个时隙；当$2N\leqslant n\leqslant 2N+K-1$时，$t_{mn}$表示核心FC交换网络与接入FC交换网络相连的发送时间轴第m个时隙；当$2N+K\leqslant n\leqslant 2N+2K-1$时，$t_{mn}$表示接入FC交换网络与核心FC交换网络相连的发送时间轴第m个时隙。对于$\forall m,n$，$0\leqslant m\leqslant M-1$，$0\leqslant n\leqslant 2N+2K-1$，$t_{mn}$满足：

$$t_{mn}=\begin{cases}0 & \text{该时隙未被占用}\\ 1 & \text{该时隙被占用}\end{cases} \tag{5}$$

对于域内业务，其在当前DBA周期可被调度的条件是：存在合适的时隙，满足发送节点i的发送时间轴和接收节点j的接收时间轴同时空闲，即当$0\leqslant i\leqslant N-1$，$0\leqslant j\leqslant N-1$，对于$\exists m$时隙，$0\leqslant m\leqslant M-1$，使得

$$t_{mi}+t_{m(j+N)}=0 \tag{6}$$

对于跨域业务，还需要再考虑交换域之间的发送轴是否空闲，可以分为以下3种情况：

①发送节点位于核心交换网络，接收节点位于接入交换网络k，则t_{mn}需满足：

$$t_{mi}+t_{m(j+N)}+t_{m(2N+k-1)}=0 \tag{7}$$

②发送节点位于接入交换网络k, 接收节点位于核心交换网络, 则t_{mn}需满足:

$$t_{mi}+t_{m(j+N)}+t_{m(2N+K+k-1)}=0 \tag{8}$$

③发送节点位于接入交换网络k_1, 接收节点位于接入交换网络k_2, 则t_{mn}需满足:

$$t_{mi}+t_{m(j+N)}+t_{m(2N+K+k_1-1)}+t_{m(2N+k_2-1)}=0 \tag{9}$$

式中: $k,k_1,k_2\in[0,K-1]$。

对于多级交换式网络, 每个 DBA 周期内业务调度所用时隙个数可表示为

$$t=\sum_{m=0}^{M-1}\sum_{n=0}^{2N+2K-1}t_{mn} \tag{10}$$

式中: 对于 $\forall m(0\leqslant m\leqslant M-1)$以及 $\forall n(0\leqslant n\leqslant 2N+K-1)$, $t_{mn}=1$或$t_{mn}=0$。

时隙占用个数越多, 表明当前 DBA 周期带宽利用率越高, 在每个 DBA 周期进行业务调度的目标, 就是使t尽可能趋近最大值。综合考虑优化目标以及算法复杂性, 选择基于首次适应的装箱算法为业务分配时隙。

且为了保证不同节点间 CBT 业务调度的公平性, 设置传输信用矩阵$\boldsymbol{T}_c$, 用来记录当前节点对之间被分配的业务带宽与保证带宽之间的差距, 差值越大, 表明在之前的 DBA 周期内分配给该对节点的带宽越小, 在后续的 DBA 周期中其应当被优先考虑调度安排。每个周期内生成的调度信息被存入带宽资源分配表$\boldsymbol{L}$中, 分配表包含 6 种信息: 调度序列号N_{seq}, 源节点地址D_{src}, 目的节点地址D_{dest}, 开始时隙索引S_{begin}, 结束时隙索引S_{end}以及传输数据量f。

综上所述, 主 NC 的跨域业务调度机制实现伪代码如下。其中, 设置跨域 CBT 带宽请求收集矩阵为$\boldsymbol{B}_{\text{sort}}$, 根据主 NC 收集的全网跨域 CBT 带宽请求进行初始化, 跨域 CBT 带宽请求数为N_{CBT}, 发送节点和接收节点所在交换域分别为i_{send}和j_{rcv}, 接入 FC 交换网络个数为K, 网络节点数为N, 跨域 CBT 带宽资源分配表为$\boldsymbol{L}$。

输入: $\boldsymbol{B}_{\text{sort}}$, N_{CBT}, i_{send}, j_{rcv}, K, N。

输出: $\boldsymbol{L}$。

1. FOR $x\leftarrow 0$ TO $N_{\text{CBT}}-1$ DO

2. FOR $k_i\leftarrow 0$ TO $\boldsymbol{B}_{\text{sort}}[2][x]$ DO //根据带宽请求的大小选择时隙

3. //核心交换网络发送, 接入交换网络接收

4. $t_1\leftarrow \boldsymbol{T}[k+k_i][i]+\boldsymbol{T}[k+k_i][j+N]+\boldsymbol{T}[k+k_i][2N+j_{\text{rcv}}-1]$

5. //接入交换网络发送, 核心交换网络接收

6. $t_2\leftarrow \boldsymbol{T}[k+k_i][i]+\boldsymbol{T}[k+k_i][j+N]+\boldsymbol{T}[k+k_i][2N+K+i_{\text{send}}-1]$

7. //接入交换网络发送, 接入交换网络接收

8. $t_3\leftarrow \boldsymbol{T}[k+k_i][i]+\boldsymbol{T}[k+k_i][j+N]+\boldsymbol{T}[k+k_i][2N+j_{\text{rcv}}-1]+\boldsymbol{T}[k+k_i][2N+K+i_{\text{send}}-1]$

9. IF ($i_{\text{send}}=0$ AND $t_1\neq 0$) OR ($j_{\text{rcv}}=0$ AND $t_2\neq 0$) OR $t_3\neq 0$ THEN

10. Flag ← false //时隙被占用

11. BREAK

12. ELSE

13. Flag ← true //时隙空闲

14. END IF

15. END FOR

16. IF Flag THEN

17. 生成一个 CBT 带宽调度任务, 存入 CBT 分配表$\boldsymbol{L}$, 修改相关参数矩阵

18. END IF

19. END FOR

2　网络仿真及结果分析

该多 NC 仿真系统采用如图 1 所示的网络架构, 采用 OMNeT++作为仿真平台[22]。该网络架构包括 1 个核心 FC 交换网络和 8 个接入 FC 交换网络, 每个交换网络包含 1 个 NC 和 16 个 NT。网络终端分为计算存储类节点和实验类节点, 区别在于: 计算存储类节点之间存在跨域业务传输, 实验类节点仅产生域内业务。业务传输采用并发交换, 由多个 NC 协同进行动态带宽分配。

2.1　仿真参数及相关指标

网络中存在的业务类型及相关指标如表 1 所示[21]。其他的固定仿真参数设置如下。

表 1　业务类型及相关指标

Table 1　Service type and related metrics

业务类型	业务流方向	发送周期/ms	最大允许时延/ms	调度方式
PT	NC→NT(域内) NT→NC(域内)	1.5～10	1.5～10	固定时隙
CBT	NT→NT	突发	100	动态分配
TCBT	NC→NT(域内)	突发	0.05	抢占式

1) 链路速率 4.25 Gbit/s;

2) 仿真时间 20 s, 时隙大小 20 μs;

3) 周期性业务的周期与 DBA 周期保持一致, 业务大小为 512 Byte;

4) 相同类型 NT 的一般突发业务负载相同, 产生时间间隔服从指数分布, 业务大小为 1 024 Byte;

5) 强时效性突发业务产生间隔服从均值为 2 ms 的指数分布, 业务大小为 100 Byte。

单NC系统参数设置与多NC仿真系统相同，仅网络控制节点个数为1个，每个DBA周期内，所有的节点向NC节点上报业务请求，由单NC进行统一的分配和调度[21]。

2.2 DBA周期对业务时延影响

图6、图7展示了当网络负载为0.3，跨域业务量与域内业务量之比为1∶2时，PT业务时延抖动、TCBT业务时延和CBT业务时延随DBA周期变化的情况。其中，小图表示对应曲线的局部放大，后续图中同理。

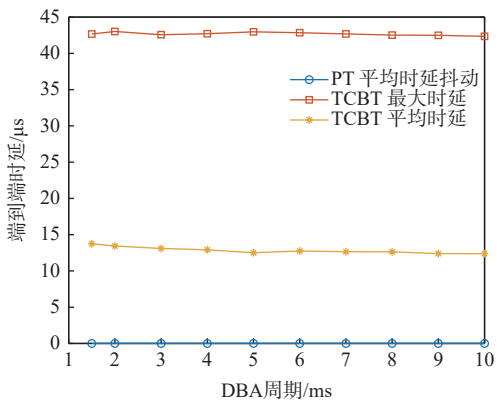
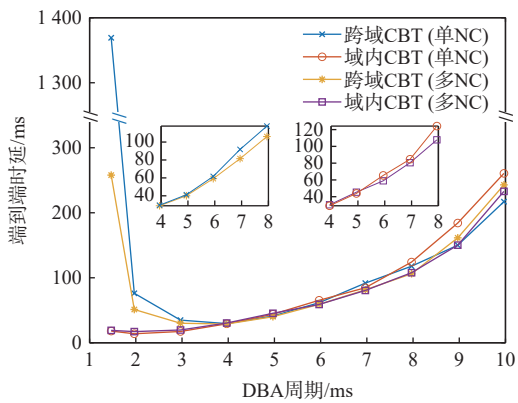
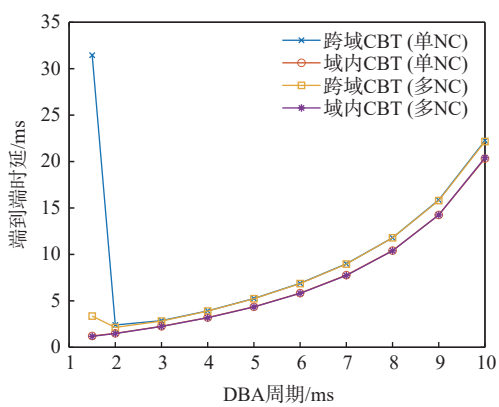


图6 DBA周期与PT、TCBT业务

Fig. 6 DBA period and PT, TCBT service



(a) 最大时延



(b) 平均时延

图7 DBA周期与CBT业务端到端时延

Fig. 7 DBA period and CBT service end-to-end delay

如图6所示，当DBA周期变化时，由于PT业务采用固定时隙调度，其时延抖动为0；TCBT业务采用抢占式调度，端到端时延几乎不受DBA周期变化的影响，最大时延小于45 μs，符合表1中TCBT业务的最大允许时延要求。

图7的仿真结果表明，多NC控制下的网络随着DBA周期的增加，CBT业务端到端时延总体呈现先下降后上升趋势。当DBA周期长度设置较短时，PT业务时段与CBT业务调度分配时段占比较大，CBT业务实际传输可用时隙数量较少，大量CBT业务产生且没有被及时调度，使业务时延较大。适当增加DBA周期长度，CBT业务可用时隙与业务量相匹配，时延会逐渐减小。而当DBA周期设置较大时，随着周期增加，CBT业务被问询的频率降低，业务得不到及时的带宽分配，导致其时延持续上升。且由于跨域业务调度条件与域内业务相比更为严苛，因此跨域业务时延整体高于域内业务。由图7(a)可知，与单NC控制相比，多NC控制下的CBT业务最大时延均不大于单NC情况，且随着DBA周期的增大或减小，时延差距逐渐增加，当DBA周期小于3 ms或大于8 ms时，最大时延差距较为明显。图7(b)中的平均业务时延两者差距不大。在DBA周期小于8 ms时，CBT业务最大时延及平均时延均能满足表1指标要求。

图8给出了当DBA周期变化时，多NC调度下3种业务调度数量的变化情况。对于PT业务，每个周期内业务总数相同，随着DBA周期的增加，调度轮数减少，导致PT业务可调度数量减少；对于CBT业务，DBA周期增加使业务上报请求次数减少，相同时间内调度量有所降低；TCBT业务调度与DBA周期无关，业务数量正常波动，并无明显降低。

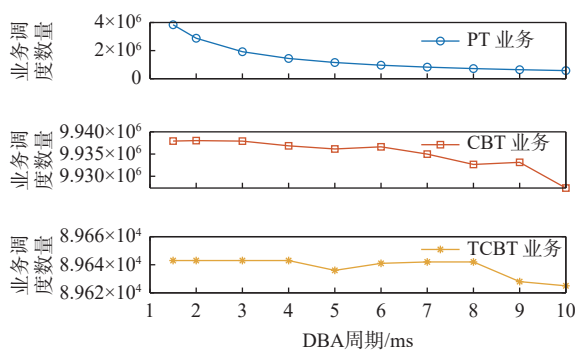


图8 DBA周期与业务调度数量

Fig. 8 DBA period and scheduled services number

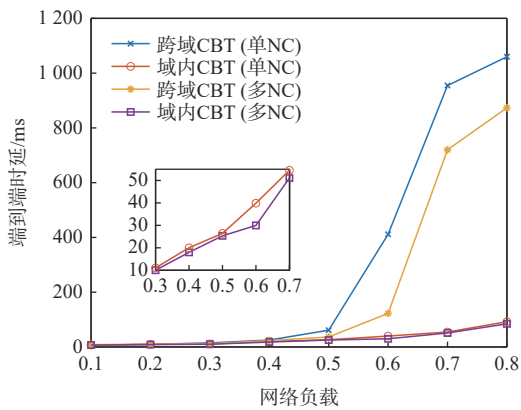
2.3 网络负载对业务时延影响

由图7可知，当DBA周期为2 ms时，CBT跨域业务平均时延与跨域业务最大时延达到最小值，且跨域业务最大时延和域内业务平均时延较低，因此

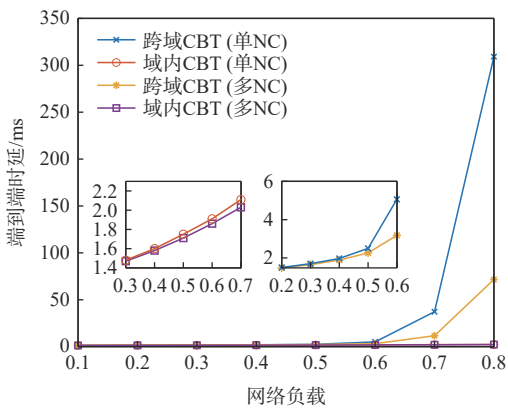
可选择 2 ms 作为后续仿真的 DBA 周期长度。

图 9 展示了当 DBA 周期为 2 ms, 跨域业务量与域内业务量之比为 3∶10 时, CBT 业务时延随网络负载变化的情况。

由图 9 可知, 随着网络负载的增加, 跨域以及域内 CBT 业务的端到端时延均有所增长。由于跨域业务调度不仅要保证发送节点和接收节点与自身域内交换机相连时间轴存在符合条件的空闲时隙, 还需保证交换域之间的收发时间轴存在相应空闲时隙, 与域内业务相比, 跨域业务调度条件更加严苛。因此当网络负载较重时, 跨域业务的时延最大值以及平均值涨幅更加明显。在多 NC 控制下, 当网络负载处于 0.57 及以下时, 跨域以及域内 CBT 业务最大时延均符合相关指标要求; 对于 CBT 业务平均时延, 在图 9 所给出的网络负载下均能满足指标要求。图 9(a)、图 9(b) 中, 多 NC 控制下的 CBT 业务最大时延及平均时延均小于单 NC 情况, 且时延差距随着网络负载的增加而增加, 这是因为多 NC 控制下的业务调度可以通过主从 NC 并行进行, 随着网络负载的增加, 相比于单 NC, 主从 NC 并行调度效率更高。当网络负载逐渐增大至 0.5 时, 多 NC 调度有较为明显的优势。



(a) 最大时延



(b) 平均时延

图 9　网络负载与 CBT 业务端到端时延

Fig. 9　Network load and CBT service end-to-end delay

图 10 给出了网络负载增加时, 多 NC 调度下业务数量的变化情况。可以看出, 由于 DBA 周期没有改变, 因此相同时间内调度的 PT 业务数量相同, 网络负载的增加也未对抢占式调度的 TCBT 业务产生明显影响。由于未产生丢包现象, 因此 CBT 业务调度数量随着网络负载的增加呈线性增长。

2.4　跨域/域内业务比对业务时延影响

由于跨域业务需要考虑交换域之间的时隙空闲情况, 相比于域内业务, 其变化对网络整体业务调度的影响更大, 图 11、图 12 展示了当 DBA 周期为 2 ms, 网络负载为 0.2 时, 保持域内业务量不变, 仅增加跨域业务量对各类业务调度的影响。

如图 11 所示, 多 NC 调度时, 由于采用固定时隙分配, PT 业务的时延抖动依然为 0; TCBT 业务的最大时延和平均时延略有增加, 最大时延小于 45 μs, 满足指标要求, 跨域 CBT 业务的增加对 TCBT 业务影响不大。

在图 12 中, 随着跨域业务量的增多, 跨域 CBT 的最大时延和平均时延均呈上升趋势。由于更多时隙被跨域业务占用, 相应域内 CBT 的最大时延也存在一定的增长。当跨域业务与域内业务比值为 1.2 时, 跨域业务的交换域间时隙分配冲突加剧,

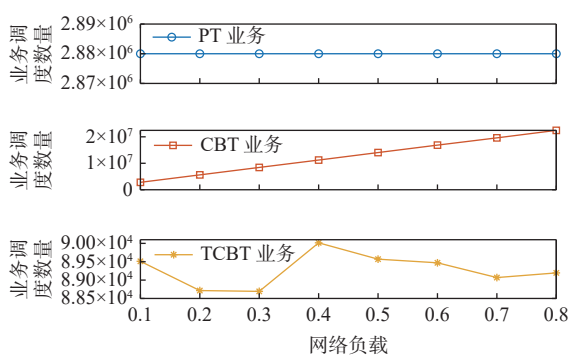


图 10　网络负载与业务调度数量

Fig. 10　Network load and scheduled services number

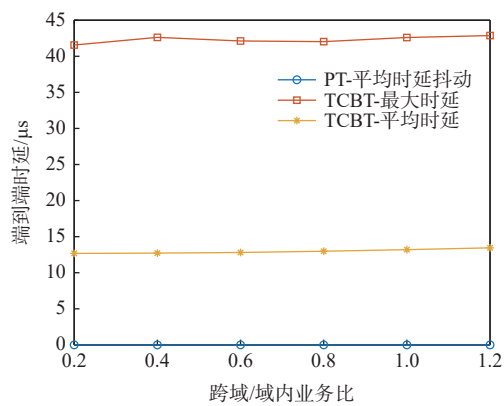


图 11　跨域/域内业务比与 PT、TCBT 业务

Fig. 11　Cross-domain/Intra-domain service ratio and PT, TCBT service

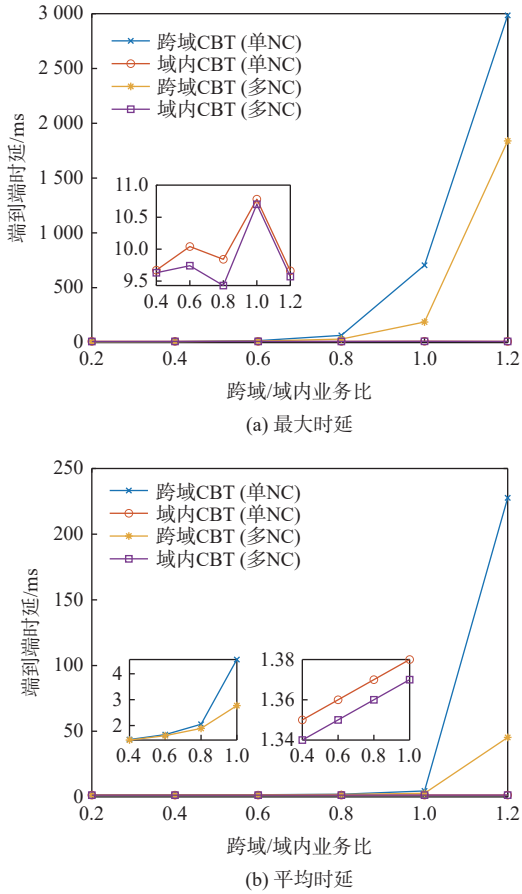


图 12 跨域/域内业务比与 CBT 端到端时延

Fig. 12 Cross-domain/Intra-domain service ratio and CBT service end-to-end delay

跨域 CBT 的最大时延以及平均时延均大幅度上升，大量跨域业务未得到及时的带宽分配；同时，更多域内节点之间的空闲时隙被用来进行域内业务传输，使域内 CBT 的最大时延略有降低。在多 NC 控制下，当跨域/域内业务比小于 1 时，CBT 业务最大时延与平均时延均能满足相关指标要求。与单 NC 相比，多 NC 控制下的 CBT 业务时延较小，且随着跨域/域内业务比的增加，多 NC 控制的时延优势逐渐明显。当跨域/域内比值增加至 0.8 时，多 NC 业务时延明显小于单 NC。

3 结 论

随着空间应用系统规模不断增大，网络结构逐渐走向分布式，单 NC 集中式调度方式存在调度不灵活、最大带宽利用率受限等问题，针对分布式网络架构，本文结合 FC-AE-1553 协议，所做主要工作如下：

1）提出了一种多网络控制节点协同作用下的动态带宽分配机制，支持 PT、CBT 和 TCBT 的混合并发传输。

2）通过理论分析和 OMNeT++仿真，验证了该方案的可行性。仿真结果表明，PT 和强 TCBT 调度几乎不受网络条件的影响。

3）CBT 调度时延会受到周期长度、网络负载和跨域业务占比等因素影响，在网络负载不超过 57% 时，该调度机制也能提供合适的带宽进行传输。

4）与单 NC 集中式控制下的 FC 交换网络相比，采用多 NC 进行业务带宽分配与控制可以提高 CBT 业务的传输效率，降低业务传输时延，当网络负载大于 0.5 或跨域/域内业务比值大于 0.8 时，多 NC 控制下的时延优势更加明显。

5）该调度机制解决了多网络控制节点调度冲突问题，为多节点控制下的业务带宽分配提供了一种解决方案。

参考文献（References）

[1] 何锋. 机载网络技术基础[M]. 北京: 国防工业出版社, 2018: 74-82.
HE F. Fundamentals of airborne network[M]. Beijing: National Defense Industry Press, 2018: 74-82(in Chinese).

[2] 林强, 熊华钢, 张其善. 光纤通道中的 1553 总线技术[J]. 航空电子技术, 2004, 35(1): 1-5.
LIN Q, XIONG H G, ZHANG Q S. The usage of 1553 in the fibre channel[J]. Avionics Technology, 2004, 35(1): 1-5(in Chinese).

[3] SOURAV B, NAIDU C D, SAI Y P, et al. Design and implementation of remote terminal for MIL-STD-1553 B[C]//2017 IEEE 7th International Advance Computing Conference (IACC). Piscataway: IEEE Press, 2017: 270-274.

[4] 于晓磊, 陈绍炜, 周虎. FC-AE-1553B 通信结构的仿真研究[J]. 信息安全与通信保密, 2010, 8(5): 52-54.
YU X L, CHEN S W, ZHOU H. Study and smulation of FC-AE-1553 network[J]. Information Security and Communications Privacy, 2010, 8(5): 52-54(in Chinese).

[5] 徐亚军, 熊华钢. 未来航电系统 FC 互连的拓扑结构研究[J]. 电光与控制, 2004, 11(4): 17-20, 23.
XU Y J, XIONG H G. Study on topology of FC interconnections in future avionic system[J]. Electronics Optics & Control, 2004, 11(4): 17-20, 23(in Chinese).

[6] 鞠铭阳, 张利洲, 王世奎. FC-AE-1553 协议分析与研究[J]. 现代电子技术, 2016, 39(11): 21-23.
JU M Y, ZHANG L Z, WANG S K. Analysis and research of FC-AE-1553 protocol[J]. Modern Electronics Technique, 2016, 39(11): 21-23(in Chinese).

[7] WANG X, WANG C, YU B W, et al. Communication module of FC-AE-1553 interface[C]// 2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC). Piscataway: IEEE Press, 2015: 1369-1373.

[8] 曹素芝, 张善从. FC-AE-1553 高级特性介绍[J]. 光通信技术, 2010, 34(2): 49-51.
CAO S Z, ZHANG S C. Analysis on advanced features of FC-AE-1553[J]. Optical Communication Technology, 2010, 34(2): 49-51(in Chinese).

[9] FANG L, ZHAO G H, CAO S Z. Design of heterogeneous FC-AE-

1553 network[C]// 2014 IEEE International Conference on Control Science and Systems Engineering. Piscataway: IEEE Press, 2014: 130-134.

[10] MCGARRY M P, REISSLEIN M. Investigation of the DBA algorithm design space for EPONs[J]. Journal of Lightwave Technology, 2012, 30(14): 2271-2280.

[11] 高敏, 朱岩, 喻芳, 等. 具有优先级消息的 FC-AE-1553 建模仿真分析[J]. 电子设计工程, 2014, 22(9): 145-148.
GAO M, ZHU Y, YU F, et al. Modeling and performance analysis of FC-AE-1553 with priority between messages[J]. Electronic Design Engineering, 2014, 22(9): 145-148(in Chinese).

[12] 丁凡, 熊华钢, 宋丽茹. FC-AE-1553 网络的建模仿真研究[J]. 计算机工程与应用, 2008, 44(31): 20-24.
DING F, XIONG H G, SONG L R. Modeling and simulation for FC-AE-1553 network[J]. Computer Engineering and Applications, 2008, 44(31): 20-24(in Chinese).

[13] 吴少俊, 赵光恒, 王立芊, 等. FC-AE-1553 网络的“并发交换式”动态带宽调度机制[J]. 北京航空航天大学学报, 2016, 42(12): 2579-2586.
WU S J, ZHAO G H, WANG L Q, et al. Dynamic bandwidth allocation mechanism with parallel and switching for FC-AE-1553 network[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(12): 2579-2586(in Chinese).

[14] ZHAN Y Y, WANG L Q, WU S J, et al. Dynamic bandwidth allocation for switching FCAE-1553 network in avionics system[C]// 2018 20th International Conference on Advanced Communication Technology (ICACT). Piscataway: IEEE Press, 2018: 1.

[15] HE Y, WANG L Q, ZHAN Y Y, et al. Dynamic bandwidth scheduling algorithm for space applications in FC-AE-1553 switching network[C]//2018 Asia Communications and Photonics Conference (ACP). Piscataway: IEEE Press, 2018: 1-3.

[16] WANG J, WANG L Q, WU S J, et al. A hybrid bandwidth allocation algorithm based on parallel scheduling for FC-AE-1553 switching network[C]// Asia Communications and Photonics Conference 2016. Washington, D. C. : OSA, 2016: 1-3.

[17] WU S J, ZHAN Y Y, QIAO K Y, et al. Scheduling mechanism of FC-AE-1553 network based on credit ranking[C]// 2018 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob). Piscataway: IEEE Press, 2018: 1-8.

[18] ZHAN Y Y, WANG L Q, LI J J, et al. Static dynamic bandwidth allocation for PON FC-AE-1553 network[C]//2016 15th International Conference on Optical Communications and Networks (ICOCN). Piscataway: IEEE Press, 2016: 1-3.

[19] WU S J, ZHAN Y Y, QIAO K Y, et al. Dynamic bandwidth scheduling for multi protocol convergence fiber channel network [C]//2018 IEEE/CIC International Conference on Communications in China (ICCC Workshops). Piscataway: IEEE Press, 2018: 54-59.

[20] ZHONG Z F, WANG L Q, LI J J, et al. FC-AE-1553 switching network supporting IP services based on parallel scheduling strategy [C]//2018 23rd Opto-Electronics and Communications Conference (OECC). Piscataway: IEEE Press, 2018: 1-2.

[21] 何玥. FC-AE-1553 多级网络多业务调度机制研究[D]. 北京: 北京邮电大学, 2019: 33-40.
HE Y. Research on multi-service scheduling mechanism of FC-AE-1553 multi-stage network[D]. Beijing: Beijing University of Posts and Telecommunications, 2019: 33-40(in Chinese).

[22] BROTHERS T J, MUKNAHALLIPATNA S, HAMANN J C. Fibre channel switch modeling at fibre channel-2 level for large fabric storage area network simulations using OMNeT: preliminary results[C]//32nd IEEE Conference on Local Computer Networks (LCN 2007). Piscataway: IEEE Press, 2007: 191-202.

FC-AE-1553 dynamic bandwidth scheduling mechanism for multi network controllers

LI Xinru[1], ZHANG Xinyue[1], LI Zheng[2, 3], HE Feng[1, *]

(1. School of Electronic Information Engineering, Beihang University, Beijing 100191, China;
2. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100190, China;
3. Key Laboratory of Technology in Geo-Spatial Information Processing and Application Systems, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: Due to the increasingly complex communication requirements of avionics system, the communication system based on FC-AE-1553 networking needs to use multiple network controllers (NC) to build a distributed network environment. The key lies in the distribution of control rights across various controllers and the resolution of cross-domain service conflicts. In order to realize multi NC collaborative scheduling, this paper proposes a multi NC dynamic bandwidth scheduling mechanism based on master-slave NC collaboration, which uses concurrent communication to improve network bandwidth utilization and enhance the transmission guarantee capability of various services. For periodic traffic, fixed time slots are used for transmission, for time constraint burst traffic, preemptive bandwidth allocation is used, and for common burst traffic, bandwidth allocation is based on "inter-domain and then intra-domain". The results indicate that the scheduling mechanism is capable of satisfying the collaborative bandwidth allocation and scheduling of various communication services in the presence of multiple network controllers. Furthermore, it can ensure the real-time usage remains below 57%. Multi-NC control offers a clear advantage in terms of latency compared to single NC centralized control, specifically when the network load exceeds 0.5 or the ratio of cross-domain/intra-domain services is higher than 0.8.

Keywords: FC-AE-1553; multi network control system; dynamic scheduling; control rights allocation; cross-domain transmission

Received: 2022-09-02; **Accepted**: 2023-02-10; **Published Online**: 2023-03-13 16:46
URL: link.cnki.net/urlid/11.2625.V.20230310.2238.004
Foundation item: National Natural Science Foundation of China (62071023)
* **Corresponding author**. E-mail: robinleo@buaa.edu.cn

《北京航空航天大学学报》征稿简则

《北京航空航天大学学报》是北京航空航天大学主办的以航空航天科学技术为特色的综合性自然科学学术期刊(月刊)。本刊以反映航空航天领域研究成果与动态、促进学术交流、培养科技人才和推动科技成果向社会生产力转化为办刊宗旨。本刊为中国自然科学技术核心期刊,并被 Ei Compendex 等国内外权威文献检索系统收录。本刊向国内外公开发行,为进一步提高办刊质量和所刊出文章的学术水平,特制定本简则。

1 论文作者及内容

1.1 本刊面向海内外所有学者。

1.2 主要刊载与航空航天科学技术有关的材料科学及工程、飞行器设计与制造、宇航科学与工程、信息与电子技术、控制技术和自动化工程、流体力学和动力工程、计算机科学及应用技术、可靠性工程与失效分析等领域的研究文章。航空航天科学技术民用方面以及具有航空航天工程背景的应用数学、应用物理、应用力学和工程管理等方面的文章也在本刊优先考虑之列。

2 来稿要求

2.1 论文应具有创新性、科学性、学术性和可读性。

2.2 论文为原创作品,尚未公开发表过,并且不涉及泄密问题。若发生侵权或泄密问题,一切责任由作者承担。

2.3 主题明确,数据可靠,图表清晰,逻辑严谨,文字精练,公式符号标准规范,标点符号正确。

2.4 文稿撰写顺序:中文题名(一般不超过 20 个汉字),作者中文姓名、单位、所在城市、邮政编码(作者顺序始终不得更改),中文摘要(包括目的、方法、结果及结论),中文关键词(5 ~ 8 个),中图分类号,引言,正文,参考文献,英文题名,作者英文姓名、单位、所在城市、邮政编码、国别,英文摘要,英文关键词。首、末页下角注明基金项目名称及编号,作者信息。

2.5 作者请登录本刊网页进行在线投稿。

3 稿件的审核、录用与版权

3.1 来稿须经专家两审和主编、编委讨论后决定刊用与否。

3.2 若来稿经审查后认定不宜在本刊发表,将及时告知作者。如果在投稿满 3 个月后仍未收到本刊任何通知,作者有权改投其他刊。在此之前,请勿一稿多投,否则一切后果自负。

3.3 来稿一经作者签字并在本刊刊出,即表明所有作者都已经认可其版权转至本刊编辑部。本刊在与国内外文献数据库或检索系统进行交流及合作时,不再征询作者意见。

邮寄地址:100191 北京市海淀区学院路 37 号 《北京航空航天大学学报》编辑部

办公地点:北京航空航天大学办公楼 405,407 房间

电　　话:(010)82316698,82316981,82315499,82317448

E-mail: jbuaa@ buaa. edu. cn

http://bhxb. buaa. edu. cn

http://www. buaa. edu. cn

《北京航空航天大学学报》
第五届编辑委员会

主　任（主　编）：赵沁平

（以下按姓氏笔画为序）

副主任（副主编）：丁希仑　王少萍　孙志梅　李秋实　李焕喜　杨嘉陵
苗俊刚　相　艳　徐立军　钱德沛　曹晋滨

编　委：马殿富　王　琪　王　聪　邓小燕　王青云　王荣明　刘　宇
刘　红　江　洁　刘　强　闫　鹏　朱天乐　刘铁钢　齐铂金
陈万春　邹正平　苏东林　杨世春　沈成平　邱志平　宋知人
杨树斌　张晓林　杨晓奕　杨继萍　李惠峰　吴新开　张瑞丰
杨照华　宋凝芳　周　锐　林宇震　林贵平　战　强　姚仰平
胡庆雷　赵秋红　段海滨　赵巍胜　席　平　郭　宏　徐　洁
徐世杰　郭洪波　康　锐　翟　锦　熊华钢

北京航空航天大学学报

Beijing Hangkong Hangtian Daxue Xuebao

（原《北京航空学院学报》）

（月刊　1956年创刊）

第50卷 第9期　2024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　100191
地　　址　北京市海淀区学院路37号
印　　刷　北京科信印刷有限公司
发　　行　《北京航空航天大学学报》编辑部
发行范围　国内外发行
联系电话　(010) 82316698 82316981
82315499 82317448
电子信箱　jbuaa@ buaa.edu.cn

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS (JBUAA)

(Monthly, Started in 1956)

Vol.50 No.9　September 2024

Administrated by　Ministry of Industry and Information Technology of the People's Republic of China
Sponsored by　Beijing University of Aeronautics and Astronautics (BUAA)
(Beijing 100191, P. R. China)
Chief Editor　ZHAO Qinping
Edited and Published by　Editorial Board of JBUAA
Printed by　Beijing Kexin Printing Co., Ltd.
Distributed by　Editorial Board of JBUAA
Telephone　(010) 82316698 82316981
82315499 82317448
E-mail　jbuaa@buaa.edu.cn
http://bhxb.buaa.edu.cn

中国标准连续出版物号：ISSN 1001-5965 / CN 11-2625/V

ISSN 1001-5965

国内定价：50.00元／期

微信公众号　超星电子书　QQ群