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

考虑城市低空风场的小型无人机路径规划方法

赵嶷飞　谷瑞嘉　任新惠

6 | 2026
Vol.52

北 京 航 空 航 天 大 学 学 报

第 52 卷　第 6 期　（总第 400 期）　2026 年 6 月

目　　次

考虑城市低空风场的小型无人机路径规划方法 …… 赵嶷飞，谷瑞嘉，任新惠（1777）

核函数约束的 S 模式信号脉冲结构简化及 TOA 精确度研究 …… 宫峰勋，刘焘（1789）

基于改进滑模观测器的电磁直线执行器无位移传感器控制 ……
…… 葛文庆，李德同，宋亚东，谭草，李波（1803）

融合混合搜索算子和竞争学习的海洋捕食者算法及应用 ……
…… 徐中辉，饶振远，马艳丽，汤泽京，黄晓东（1810）

基于改进贝叶斯网络的 A320 空调系统故障传播路径识别方法 … 陈玖圣，余卓洋，郭润夏，吴军（1827）

基于 NOMA 的无人机 ISAC 系统轨迹和资源联合优化 …… 唐菁敏，胡丞，宋耀莲，虞贵财（1839）

基于数字孪生和 Transformer-LSTM-XGBoost 的液压机器人故障诊断方法 ……
…… 李亚洁，吴瑞龙，李炜（1850）

脉冲星导航试验卫星核心载荷工作性能稳定性分析 ……
…… 闫林丽，张健康，周庆勇，雷耀虎，樊少娟（1869）

垂直起降机场运行任务规划模型与容量估算方法 …… 魏志强，肖鑫隆（1880）

基于增强逐点图卷积网络的民航短文本组合分类方法 …… 刘晓琳，宋营营，李卓（1890）

基于改进 YOLOv7 的风机叶片缺陷检测 …… 汤占军，张朝杰，王健，陆鹏，刘汇塬，寨洪（1903）

飞行员工作负荷定量实时预测模型 …… 何雪丽，王黎静，湛希，苗冲冲，谷宗辉（1915）

面向直升机吊挂飞行品质改善的模糊减摆控制 …… 王洛烽，陈仁良，赵宇（1924）

基于迁移学习的自动调制分类方法 …… 王栋，崔天舒，姬丽彬，黄永辉，朱岩（1935）

基于知识图谱的在轨卫星与卫星网络资料匹配关系预测方法 ……
…… 徐聪，石梦鑫，王洪锋，贾清玉，智佳，杨甲森（1944）

基于多源图像融合的无人机航拍舰船目标识别方法 …… 姜杰，闫文君，刘凯，张立民（1955）

翼身融合组合动力运载器耦合抑制方法 …… 熊知越，张曙光，丁嘉元，廖雨舟（1965）

DoS 攻击下基于 H_∞的移动机器人系统鲁棒控制 …… 季旭东，陈友东，魏洪兴（1982）

可重复使用运载器连接结构损伤容限分析 …… 李配缘，蔡巧言，黎增山，冯家赫（1990）

一体化加力燃烧室燃烧特性模拟与机器学习燃烧性能预测 ……
…… 张宇，王奉明，王彦红，穆林，东明（2000）

低雷诺数柔性拍动平板非定常气动特性 ……
…… 曹梦达，郑孟宗，苏冠廷，潘天宇，李志平，李秋实（2011）

基于牛顿-欧拉递推算法的高置信度斜铰旋翼动力学模型 …… 杨毅松，李建波，段登燕（2024）

基于空气品质及代偿损失的客舱回风比例优化 …… 高阳，林家泉（2034）

提高存算一体化系统性能的静态计算划分方法 …… 薛洪宇，徐晟，罗乐，闫亮，邹兴奇（2042）
鹊桥二号中继星热控设计特点与飞行验证 ……
…… 杨昌鹏，任红艳，孙腾飞，周强，黄兴，张立华，熊亮，孙骥，程文龙（2054）
碳基材料烧蚀传热问题的不确定性分析方法 …… 刘挺，刘骁，郭磊，曾磊，国义军（2064）
变推力、变充液比、贮箱倾斜状态下液体晃动分析 …… 何家伟，吕敬，王天舒，刘颖杰（2074）
基于多策略融合蚁狮优化算法的 MPRM 逻辑电路面积优化方法 ……
…… 潘家义，何振学，赵晓君，何俊才，周宇豪，王翔（2083）
飞艇抗随机干扰控制 …… 常家铭，李素兰，段学超，张伟，汪晨阳（2092）
一种高效 Swin Transformer 加速器设计 …… 蔡青竹，刘强（2102）
基于错误率分配的逻辑电路面积近似优化 …… 范海媚，吴乾火，王伦耀（2114）
基于超级块嵌套的动态二进制翻译性能优化方法 …… 金帅，柴志雷，周浩杰（2123）
考虑构型的旋转碟状飞行器水漂动力学特性 …… 王迪，罗剑桥，于宁，孟军辉（2133）
基于结构重参数化的深度可分离卷积神经网络 …… 陈红，闫建国，杨华，张靖，李伟，杨靖（2145）
基于模仿学习的机动目标拦截强化学习制导律 ……
…… 任乐亮，鲜勇，刘振宇，张大巧，李冰，李少朋（2156）
强阵风下大柔性机翼的非线性流固耦合响应分析 …… 苏进鑫，席梓严，戴玉婷（2172）
基于序列反转的高速机动目标雷达相干积累方法 …… 王睿铮，李世强（2184）
联合语义分割和边缘纹理的人脸图像修复 …… 石计亮，张乾，周遵富，杨思红（2194）
多尺度感知与红外特征增强的 RGB-T 人群计数方法 …… 郑棣文，石洋宇，谢承杰，卢树华（2208）
基于音视频特征优化与跨模态 Transformer 的多模态情感分析 …… 林宜山，左景，卢树华（2219）

期刊基本参数：CN 11-2625/V * 1956 * m * A4 * 452 * zh * P * ￥50.00 * 400 * 40 * 2026-06

（**编　辑**　张　嵘　孙　芳　万玉敏　卞欢欢　王　茜　李亚泰）

JOURNAL OF BEIJING UNIVERSITY OF
AERONAUTICS AND ASTRONAUTICS

Vol. 52 No. 6 （Sum 400） June 2026

CONTENTS

Small unmanned aerial vehicles path planning method considering urban low-altitude wind fields
…… ZHAO Yifei，GU Ruijia，REN Xinhui （1777）

Simplified S-mode signal pulse structure with kernel function constraints and TOA accuracy study
…… GONG Fengxun，LIU Tao （1789）

Displacement sensorless control of electromagnetic linear actuator based on an improved sliding mode observer
…… GE Wenqing，LI Detong，SONG Yadong，TAN Cao，LI Bo （1803）

Marine predator algorithm incorporating hybrid search operators and competitive learning and applications
…… XU Zhonghui，RAO Zhenyuan，MA Yanli，TANG Zejing，HUANG Xiaodong （1810）

A fault propagation path identification method for A320 air conditioning system based on improved Bayesian network
…… CHEN Jiusheng，YU Zhuoyang，GUO Runxia，WU Jun （1827）

NOMA-based joint optimization of trajectory and resources for UAV-enable integrated sensing and communication system
…… TANG Jingmin，HU Cheng，SONG Yaolian，YU Guicai （1839）

Fault diagnosis method for hydraulic robots based on digital twin and Transformer-LSTM-XGBoost
…… LI Yajie，WU Ruilong，LI Wei （1850）

Stability analysis of core payload operational performance on the XPNAV-1 satellite
…… YAN Linli，ZHANG Jiankang，ZHOU Qingyong，LEI Yaohu，FAN Shaojuan （1869）

Vertiport operational task planning model and capacity estimation method
…… WEI Zhiqiang，XIAO Xinlong （1880）

Civil aviation short text combined classification method based on enhanced point-wise graph convolutional networks
…… LIU Xiaolin，SONG Yingying，LI Zhuo （1890）

Wind turbine blade defect detection based on improved YOLOv7
…… TANG Zhanjun，ZHANG Chaojie，WANG Jian，LU Peng，LIU Huiyuan，JIAN Hong （1903）

Quantitative real-time prediction model of pilot workload
…… HE Xueli，WANG Lijing，ZHAN Xi，MIAO Chongchong，GU Zonghui （1915）

Fuzzy anti-swing controller for improving handling quality of helicopter slung-load operation
…… WANG Luofeng，CHEN Renliang，ZHAO Yu （1924）

Automatic modulation classification method based on transfer learning
…… WANG Dong，CUI Tianshu，JI Libin，HUANG Yonghui，ZHU Yan （1935）

Prediction method for matching between in-orbit satellites and satellite network filings based on knowledge graph
…… XU Cong，SHI Mengxin，WANG Hongfeng，JIA Qingyu，ZHI Jia，YANG Jiasen （1944）

Ship target recognition method based on multi-source image fusion for unmanned aerial vehicle aerial photography
…… JIANG Jie，YAN Wenjun，LIU Kai，ZHANG Limin （1955）

Coupling suppression method of blended-wing-body combined-cycle-power vehicle
…… XIONG Zhiyue，ZHANG Shuguang，DING Jiayuan，LIAO Yuzhou （1965）

Robust control of mobile robots based on H_∞ under DoS attack
…… JI Xudong，CHEN Youdong，WEI Hongxing （1982）

Damage tolerance analysis on reusable launch vehicle connection structures
…… LI Peiyuan，CAI Qiaoyan，LI Zengshan，FENG Jiahe （1990）

Simulation of combustion characteristics and prediction of combustion performance using machine learning in an integrated afterburner …… ZHANG Yu，WANG Fengming，WANG Yanhong，MU Lin，DONG Ming （2000）

Unsteady aerodynamic characteristics of flexible flapping plate at low Reynolds numbers
…… CAO Mengda，ZHENG Mengzong，SU Guanting，PAN Tianyu，LI Zhiping，LI Qiushi （2011）

Dynamic model of high confidence tilt-hinge rotor based on Newton-Euler recursion algorithm
…… YANG Yisong，LI Jianbo，DUAN Dengyan （2024）

Optimization of cabin return air ratio based on air quality and compensation loss
…… GAO Yang，LIN Jiaquan （2034）

Enhancing performance through static computing partitioning approach in processing-in-memory systems
…… XUE Hongyu，XU Sheng，LUO Le，YAN Liang，ZOU Xingqi （2042）

Thermal control design features and flight validation of Queqiao-2 satellite
…… YANG Changpeng，REN Hongyan，SUN Tengfei，ZHOU Qiang，HUANG Xing，ZHANG Lihua，XIONG Liang，SUN Ji，CHENG Wenlong （2054）
Uncertainty analysis methods for heat transfer ablation in carbon-based materials
…… LIU Ting，LIU Xiao，GUO Lei，ZENG Lei，GUO Yijun （2064）
Analysis of liquid sloshing under variable thrust, variable filling liquid ratio, and tilted tank state
…… HE Jiawei，LYU Jing，WANG Tianshu，LIU Yingjie （2074）
Area optimization approach for MPRM logic circuits based on multi-strategy synergy ant lion optimization algorithm
…… PAN Jiayi，HE Zhenxue，ZHAO Xiaojun，HE Juncai，ZHOU Yuhao，WANG Xiang （2083）
Anti-stochastic disturbance control of airship
…… CHANG Jiaming，LI Sulan，DUAN Xuechao，ZHANG Wei，WANG Chenyang （2092）
A high-efficient Swin Transformer accelerator design
…… CAI Qingzhu，LIU Qiang （2102）
Area approximate optimization of logic circuits based on error rate allocation
…… FAN Haimei，WU Qianhuo，WANG Lunyao （2114）
Superblock nesting based on performance optimization for dynamic binary translation
…… JIN Shuai，CHAI Zhilei，ZHOU Haojie （2123）
Dynamic characteristics of rotating disk-shaped flight vehicle during skipping considering configuration
…… WANG Di，LUO Jianqiao，YU Ning，MENG Junhui （2133）
Deep separable convolutional neural networks based on structural reparameterization
…… CHEN Hong，YAN Jianguo，YANG Hua，ZHANG Jing，LI Wei，YANG Jing （2145）
Reinforcement learning guidance law for maneuvering target interception based on imitation learning
…… REN Leliang，XIAN Yong，LIU Zhenyu，ZHANG Daqiao，LI Bing，LI Shaopeng （2156）
Nonlinear fluid-structure interaction response analysis of a large flexible wing under strong gusts
…… SU Jinxin，XI Ziyan，DAI Yuting （2172）
Radar coherent integration method for high-speed maneuvering targets based on sequence reversal
…… WANG Ruizheng，LI Shiqiang （2184）
Face image inpainting combining semantic segmentation and edge texture
…… SHI Jiliang，ZHANG Qian，ZHOU Zunfu，YANG Sihong （2194）
RGB-T crowd counting method with multi-scale perception and infrared feature enhancement
…… ZHENG Diwen，SHI Yangyu，XIE Chengjie，LU Shuhua （2208）
A multimodal sentiment analysis based on audio and video features optimization and cross-modal Transformer
…… LIN Yishan，ZUO Jing，LU Shuhua （2219）

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0281

考虑城市低空风场的小型无人机路径规划方法

赵嶷飞[1,*]，谷瑞嘉[1]，任新惠[2]
(1. 中国民航大学 空中交通管理学院，天津 300300；2. 中国民航大学 经济与管理学院，天津 300300)

摘 要：随着无人机(UAV)在城市物流配送中的广泛应用，无人机航路运行安全逐渐引发关注。针对城市风场引发的无人机安全与效率问题，结合物流无人机运行特点，使用计算流体力学(CFD)方法估计城市风场，从飞行安全、无人机速度变化、建筑物周边湍流区等角度分析风场对无人机的影响，构建路径规划模型，并提出改进的Theta*求解算法。以深圳市龙岗区中心商务区为研究场景，基于历史主导风进行城市风场仿真，划设飞行风险区与障碍物影响范围，并使用所提方法在不同风场中规划无人机路径。结果显示：构建的路径规划模型有效识别了空域内的高风速区与湍流区，在考虑风场对无人机速度影响并避开飞行禁区的基础上，相比于以最小化距离为目标的无人机路径规划模型降低了9.27%的飞行时间；所提算法在路径距离与飞行时间优于传统算法的同时，减少了31.78%的拐点数量与62.63%的计算时间。

关 键 词：无人机；城市风场；路径规划；物流配送；城市空中交通
中图分类号：U8; V279
文献标志码：A **文章编号**：1001-5965(2026)06-1777-12

近年来，无人机(unmanned aerial vehicle, UAV)在城市内的应用日益增多。其中，轻小型无人驾驶航空器以其机动性强、灵活度高的特点，在物流配送领域相比于传统运输方式具有独特优势。伴随着市场需求的扩大，美团、京东、顺丰等企业在中国多地开展了无人机物流配送的应用试点工作，无人机物流配送场景不断拓宽。城市环境中，地面人口密集且高大建筑物较多，无人机物流运输需保证运行安全。为支持大规模无人机安全高效运行，路径规划是其中的关键技术。无人机路径规划可在飞行前有效降低运行风险并优化航路效率，被广泛研究[1]。现有研究多聚焦于避障[2]、噪声[3]、导航信号[4]与能耗[5]。然而，对于城市内低高度运行的无人机，风场是重要的考虑因素，因为其可能导致无人机失控并丧失稳定性[6]。阵风可使无人机偏离其预期航路，甚至发生碰撞[7]。风场还会影响无人机飞行速度，导致飞行时间增长或飞行能耗增加[8]。此外，城市峡谷与单一建筑物均可产生等级显著的湍流，危及航空器[9]。因此，在路径规划阶段考虑城市低空风场对无人机的影响，对提升无人机运行安全水平具有重要意义。

现有研究从城市风场估计、风场对无人机影响及考虑风场的无人机路径规划3个方面展开。城市风场估计多使用计算流体动力学(computational fluid dynamics, CFD)方法。Sutherland使用CFD研究了单建筑物与城市峡谷的风场估计方法，并测试了建筑物附近的风速与湍流动能(turbulence kinetic energy, TKE)[10]。Raza使用大涡模拟(large eddy simulation, LES)模型模拟了单建筑物附近的风场，并对建筑物屋顶回流区、后方尾部涡流区及两侧加速区等位置的风速分量进行了评估与比较[11]。Liu[12]、Kang[13]、Yang[14]等使用雷诺平均纳维-斯托

收稿日期：2024-05-06；**录用日期**：2024-07-12；**网络出版时间**：2024-07-23 10:34
网络出版地址：link.cnki.net/urlid/11.2625.V.20240722.1809.002
基金项目：国家重点研发计划(2022YFB4300904)
*通信作者. E-mail：yfzhao@cauc.edu.cn

引用格式：*赵嶷飞，谷瑞嘉，任新惠. 考虑城市低空风场的小型无人机路径规划方法[J]. 北京航空航天大学学报，2026，52(6)：1777-1788.*
ZHAO Y F，GU R J，REN X H. Small unmanned aerial vehicles path planning method considering urban low-altitude wind fields[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1777-1788 (in Chinese).

克斯(Reynolds-averaged Navier-Stokes, RANS)模型模拟了建筑物群影响下的风场,表明不同的城市结构与风向会对风场产生影响。

基于城市风场分析,风场对无人机运行的影响被进一步研究。Sutherland[10]与Raza[11]测试了风场下无人机在建筑物周围的位置保持能力,结果表明,较高的风速及湍流扰动会使无人机产生较大的位置偏差。Jung等研究了风场对无人机飞行路径的影响,发现阵风扰动下控制算法无法保证准确的导航与避障[15]。Jeong等使用深度学习预测了无人机在风场下的飞行风险区[16]。Giersch等评估了城市环境内风场对无人机运行安全影响,选取指标包括平均风速与湍流动能等[17]。此外,一些风险评估研究将风速列为了考虑因素。Rubio-Hervas等将不安全事件发生概率表示为随风速的提升而增加[18]。Altun等在用于先进空中交通(advanced air mobility, AAM)的风险评估方法中指出,风速超出航空器抗风等级后,其失控概率将会提升[19]。

目前,考虑风场的无人机路径规划研究较少。Ware等构建了风场下多旋翼无人机的能耗模型,结果表明,基于风场信息进行路径规划可有效提升无人机的平均地速,并减少航行所需能耗[20]。Rubio-Hervas等以风速表征意外事故发生概率,以最小化事故概率为目标进行无人机路径规划[18]。Thanellas等基于风场信息与无人机动力学模型,使用A*算法规划航路,缩短了路径航行所需时间[21]。Hong等使用神经网络预测无人机在给定风场下的能耗,并使用强化学习生成能量高效的无人机航路[22]。

总之,城市风场研究多使用CFD方法分析风场特性。风场对无人机的影响则从飞行安全、无人机速度变化与能耗改变的角度分析。现有考虑风场的无人机路径规划研究聚焦于降低高风速区所引发的风险,或主要关注减少路径飞行时间与能耗,缺乏对风场内风速与湍流状况所带来的飞行安全问题、无人机速度变化及城市环境内建筑物周边风场特性的综合考虑。基于此,本文提出了风场下无人机路径规划方法,分析了风场影响,以最小化飞行时间为目标进行路径规划。本文主要创新点如下:

1)在风场影响分析中,除风速外,添加了对建筑物周边湍流区域的考虑。分析了湍流对无人机飞行的影响,探讨了湍流区域对路径规划的影响,并选取湍流动能为指标衡量空域内的湍流强度。

2)在风场处理时,引入平均风速、最大风速与最大湍流动能。平均风速用于计算风场下无人机飞行速度变化;最大风速用于输入无人机抗风能力限制;最大湍流动能用于输入湍流区飞行限制。

1 城市风场影响分析

城市风场指城市边界层内由于建筑物布局、街道走向及地形等因素产生的特殊风流模式[23]。城市风场受2类因素影响:主导风与城市环境特征。主导风会因城市结构的影响被重新导向,结果表现为不同区域风速的增加或降低,以及部分位置的风向改变或回流。分析城市低空风场对无人机飞行的影响应从风速与湍流2个方面着手。一方面,低空流场中增加风的影响后,流场内部可能出现部分区域风速的大幅提升,超过无人机抗风能力,从而形成无法进入的飞行禁区;即便在风速较小的区域,无人机地速和能耗也会受到影响。另一方面,按照卡门涡流原理,风场中的建筑物周围会出现不稳定的湍流区域,对其中飞行的无人机产生混沌扰动,是无法进入的飞行禁区。风场对无人机飞行影响如图1所示。

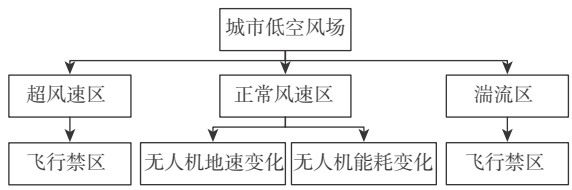


图1 城市低空风场对无人机运行影响

Fig. 1 Impact of urban low-attitude wind fields on UAV operations

1)无人机抗风能力限制。无人机抗风能力与机型相关,部分无人机制造商公布其无人机可承受最大风速为10 m/s或更高,而城市内建筑物两侧或城市峡谷间的风流加速可使得局部风速超出无人机抗风等级,导致无人机偏离航路甚至发生碰撞,危及无人机飞行安全。超风速区域应被划设为飞行禁区避免进入,在路径规划时进行绕行。

2)无人机地速变化影响。无人机飞行速度约为10~20 m/s,而城市内瞬时阵风可达8.0 m/s甚至更高,两者非常接近,由此导致城市风对无人机地速影响很大。顺风、逆风、侧风飞行时,无人机航行相同距离所需时间差异显著。无人机空速$\boldsymbol{v}_a$、地速$\boldsymbol{v}_g$、风速$\boldsymbol{v}_w$三者之间的关系如图2所示。本文假设,无人机在风场中飞行时,$\boldsymbol{v}_g$始终指向下一位置点;$\boldsymbol{v}_a$大小保持不变,方向由地速与风速共同决定。

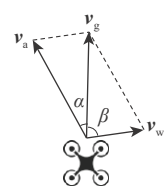


图2 无人机空速、地速与风速关系

Fig. 2 Relationship between UAV airspeed, ground speed, and wind speed

3) 无人机能耗变化影响。无人机飞行的主要能耗来源于驱动旋翼旋转所需的功率, 该功率与无人机空速$\boldsymbol{v}_{\mathrm{a}}$密切相关, 而由于风速在无人机自身飞行速度中占较大比重, 无人机飞行的能耗受风场影响较大。在尾风、顶风及侧风情况下, 无人机航行相同距离的能耗差异显著。考虑到研究风场对无人机航行能耗影响的复杂性与工作量, 本文的路径规划研究主要考虑风场对无人机飞行速度而非能耗的影响。

4) 湍流区飞行限制。由于尺寸较小、飞行高度较低且飞行速度较低, 无人机相比于传统航空器对小尺度的大气湍流及风切变更为敏感, 城市低空由建筑物诱导产生的显著湍流可导致无人机失控。湍流区应被划设为飞行禁区避免进入, 在路径规划时进行绕行。可用于衡量湍流风场对无人机运行安全影响的指标包括平均风速、风速变化量、湍流动能及瞬时风速, 因风速信息已在“无人机抗风能力限制”中分析, 本文选择湍流动能作为指标衡量空域内的湍流强度。湍流动能增加将导致无人机飞行更加不稳定, 从而需要与障碍物保持更大的间距。本文选取城市风场内较大的湍流动能数值 1.5 $\mathrm{m^2/s^2}$, 将建筑物周边湍流动能超出此阈值的区域划设为飞行禁区。

2　风场影响下无人机路径规划方法

2.1　空域建模

路径规划前需进行空域建模。以栅格表征飞行环境[24], 设空域为 *OABC-O′A′B′C′*立方体区域, 长、宽、高分别为 *Q*、*W*、*E*, 以 *O* 为原点建立三维直角坐标系, 如图 3 所示。

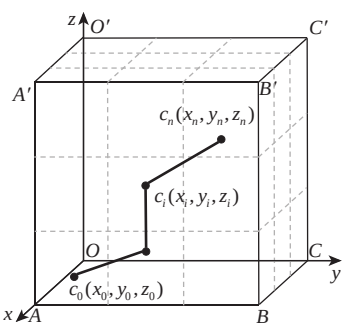


图 3　空域栅格化

Fig. 3　Airspace gridification

依据栅格尺寸l_{grid}, 将空域划分为$u\times v\times w$个立体栅格, 其中, $u=\mathrm{int}(Q/l_{\mathrm{grid}})$, $v=\mathrm{int}(W/l_{\mathrm{grid}})$, $w=\mathrm{int}(E/l_{\mathrm{grid}})$, int(·)为取整函数。路径规划中, 以$c_0$表示起始点, c_n表示目标点, c_i表示途经路径点, 其中, $i\in(1,2,\cdots,n-1)$。

栅格建模如图 4 所示, 其尺寸l_{grid}取值应考虑无人机机型、运输任务类型及 CFD 仿真网格精度等因素。在计算开销可接受的前提下, l_{grid}数值应与无人机尺寸及 CFD 网格尺寸相匹配。由于 CFD 仿真会在建筑物周边及尾流区域进行网格精细化, 这些位置的空域栅格可能包含多个 CFD 网格的风速与湍流动能信息。对于空域模型 *OABC-O′A′B′C′*, 其内每个栅格c_i所含信息包括坐标$\boldsymbol{p}_i(x_i,y_i,z_i)$、风速$\boldsymbol{v}_{\mathrm{w},i}(v_{x,i},v_{y,i},v_{z,i})$、湍流动能$k_i$, 即

$$c_i=\{\boldsymbol{p}_i(x_i,y_i,z_i),\boldsymbol{v}_{\mathrm{w},i}(v_{x,i},v_{y,i},v_{z,i}),k_i\} \tag{1}$$

式中: $v_{x,i}$、$v_{y,i}$、$v_{z,i}$分别为x、y、z方向的风速分量。

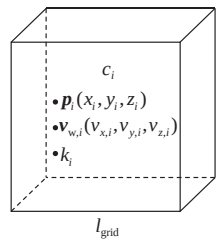


图 4　空域栅格建模

Fig. 4　Airspace grid modelling

当栅格c_i含多个 CFD 网格时, 使用栅格c_i内风速集合$\{\boldsymbol{v}_{\mathrm{w},i,1},\cdots,\boldsymbol{v}_{\mathrm{w},i,n}\}$计算栅格的平均风速与最大风速, 湍流动能集合$\{k_{i,1},\cdots,k_{i,n}\}$计算栅格的最大湍流动能, 如下:

$$\boldsymbol{v}_{\mathrm{w},i,\mathrm{mean}}=\frac{\sum_{j=1}^{n}\boldsymbol{v}_{\mathrm{w},i,j}(v_{x,i,j},v_{y,i,j},v_{z,i,j})}{n} \tag{2}$$

$$\boldsymbol{v}_{\mathrm{w},i,\max}=\max_{j=1,\cdots,n}\boldsymbol{v}_{\mathrm{w},i,j}(v_{x,i,j},v_{y,i,j},v_{z,i,j}) \tag{3}$$

$$k_{i,\max}=\max_{j=1,\cdots,n}k_{i,j} \tag{4}$$

式中: $\boldsymbol{v}_{\mathrm{w},i,\mathrm{mean}}$为$c_i$内的平均风速; $\boldsymbol{v}_{\mathrm{w},i,\max}$为$c_i$内的最大风速; $k_{i,\max}$为c_i内的最大湍流动能; j为c_i内的第j个 CFD 网格。$\boldsymbol{v}_{\mathrm{w},i,\mathrm{mean}}$用于路径规划中计算风速影响下无人机的地速; $\boldsymbol{v}_{\mathrm{w},i,\max}$与$k_{i,\max}$用于判断空域栅格是否可以进入, 标准如下:

1) 无人机抗风能力限制。使用c_i内最大风速大小$|\boldsymbol{v}_{\mathrm{w},i,\max}|$判断该空域栅格是否可以进入。当$|\boldsymbol{v}_{\mathrm{w},i,\max}|$超出无人机可承受最大风速$v_{\mathrm{w,res}}$时, 该空域栅格为不可进入栅格, 如下:

$$\text{If } |\boldsymbol{v}_{\mathrm{w},i,\max}|>v_{\mathrm{w,res}} \qquad c_i\in C_{\text{no-fly}} \tag{5}$$

式中: $v_{\mathrm{w,res}}$为无人机可承受最大风速, 取值为 10.7 m/s; $C_{\text{no-fly}}$为无人机禁飞区栅格集合。

2) 湍流区飞行限制。使用c_i内最大湍流动能$k_{i,\max}$判断该空域栅格是否可以进入。当$k_{i,\max}$超出

城市风场中湍流动能阈值k_{res}，且该栅格距空间内建筑物的最小距离r_i小于或等于湍流动能检测半径r_k，该空域栅格为不可进入栅格，如下：

$$\text{If } k_{i,\max} > k_{\text{res}} \text{ and } r_i \leqslant r_k \qquad c_i \in C_{\text{no-fly}} \tag{6}$$

式中：k_{res}为湍流动能阈值，取值为1.5 m^2/s^2；r_k为建筑物周边的湍流动能检测半径，取值为20 m。

3）建筑物限制。除空域中的禁飞区外，由建筑物占据的栅格同样不可进入。此外，建筑物周围半径r_{bld}内的区域作为安全缓冲区，划入建筑物栅格集合内，即

$$\text{If building in } c_i \text{ or } r_i \leqslant r_{\text{bld}} \qquad c_i \in C_{\text{bld}} \tag{7}$$

式中：r_{bld}为建筑物安全缓冲区半径，取值为5 m；C_{bld}为建筑物栅格集合。

2.2 目标函数

目标函数应依据影响因素设置。由第1节分析可知，本文重点研究无人机抗风能力限制、无人机飞行速度变化，以及湍流区飞行限制对路径规划的影响。由于无人机抗风能力限制与湍流区飞行限制均指出应划设飞行禁区，两者将在约束条件中以硬性约束（无法进入）的形式出现，而非在目标函数中作为惩罚因子。对于无人机飞行速度变化，在风场中通过无人机地速计算飞行时间以衡量路径成本相比于通过距离衡量更为合理，因此，本文选取飞行时间作为目标函数考虑因素。无人机地速推导过程及目标函数的确定如下：

考虑风场对无人机飞行速度影响，对任意栅格c_i，推导其空速与地速变化。已知栅格c_i处风速大小为$|\boldsymbol{v}_{\text{w},i}|$，风速方向为$\boldsymbol{v}_{\text{w},i}/|\boldsymbol{v}_{\text{w},i}|$，设无人机在$c_i$处地速方向为其移动方向，空速大小为$|\boldsymbol{v}_{\text{a},i}|$，如图2所示，由于$\boldsymbol{v}_{\text{g}} = \boldsymbol{v}_{\text{a}} + \boldsymbol{v}_{\text{w}}$，有

$$|\boldsymbol{v}_{\text{a},i}| \sin\alpha = |\boldsymbol{v}_{\text{w},i}| \sin\beta \tag{8}$$

式中：α为空速矢量$\boldsymbol{v}_{\text{a},i}$与地速矢量$\boldsymbol{v}_{\text{g},i}$的夹角；$\beta$为风速矢量$\boldsymbol{v}_{\text{w},i}$与地速矢量$\boldsymbol{v}_{\text{g},i}$的夹角。

由式(8)可求得空速与地速夹角α及空速方向，则栅格c_i处无人机地速如下：

$$|\boldsymbol{v}_{\text{g},i}| = |\boldsymbol{v}_{\text{a},i}| \cos\alpha + |\boldsymbol{v}_{\text{w},i}| \cos\beta \tag{9}$$

式中：$|\boldsymbol{v}_{\text{g},i}|$为无人机在栅格$c_i$处的地速大小。

以T表示无人机沿航路由c_0航行至c_n所需的飞行时间，则目标函数表达式为

$$\min T = \sum_{i=1}^{n} \frac{\sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 + (z_i - z_{i-1})^2}}{|\boldsymbol{v}_{\text{g},i-1}|} \tag{10}$$

式中：$|\boldsymbol{v}_{\text{g},i-1}|$为无人机由路径点$c_{i-1}$航行至$c_i$的地速大小。

2.3 约束条件

1）最小路径段长度。相邻路径点距离不能小于最小路径段长度，约束为

$$l_i \geqslant l_{\min} \qquad i = 1,2,\cdots,n \tag{11}$$

$$l_i = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 + (z_i - z_{i-1})^2} \tag{12}$$

式中：l_i为各路径段长度；$l_{\min}$为最小路径段长度。

2）航程约束。无人机路径距离必须小于其最大航程，约束为

$$\sum_{i=1}^{n} l_i \leqslant L_{\max} \tag{13}$$

式中：$L_{\max}$为无人机最大航程。

3）转弯角约束。无人机转弯时，满足约束：

$$0 \leqslant \beta_i \leqslant \beta_{\max} \qquad i = 1,2,\cdots,n \tag{14}$$

式中：$\beta_{\max}$为最大转弯角。

4）飞行高度约束。无人机飞行高度的限制约束为

$$H_{\min} \leqslant z_i \leqslant H_{\max} \qquad i = 1,2,\cdots,n \tag{15}$$

式中：$H_{\min}$、$H_{\max}$分别为最高、最低飞行高度，取决于无人机类型与空域政策。

5）栅格限制。任意路径点c_i不可进入禁飞区或建筑物栅格集合，如下：

$$\forall \boldsymbol{p}_i(x_i,y_i,z_i), c_i \notin C_{\text{no-fly}} \text{ and } c_i \notin C_{\text{bld}} \tag{16}$$

式中：$\boldsymbol{p}_i(x_i,y_i,z_i)$为路径中任意节点$c_i$的坐标。

2.4 求解算法

路径规划中常用A*算法，其在移动时仅可选择相邻节点，因而路径中包含非必要拐点，增加了路径长度，并降低了路径效率。Theta*算法通过使用视线函数连接非相邻节点，避免了多余的转向，从而生成距离更短且更加高效的路径。为求解风场下的路径规划问题，本文改进了Theta*算法，如下：

1）算法成本函数改进。算法的成本函数为

$$f(c_0,c_k) = g(c_0,c_k) + h(c_k,c_n) \tag{17}$$

式中：$f(\cdot)$为规划至c_k时，由路径点c_0至c_n的总成本；$g(\cdot)$为c_0至c_k的路径成本；$h(\cdot)$为c_k至c_n的启发式成本。

改进$g(\cdot)$与$h(\cdot)$如下：

$$g(c_0,c_k) = \sum_{i=1}^{k} \frac{\text{dist}(c_{i-1},c_i)}{|\boldsymbol{v}_{\text{g},i-1}|} \tag{18}$$

$$h(c_k,c_n) = \sum_{i=k+1}^{n} \frac{\text{dist}(c_{i-1},c_i)}{v_{\text{g,est}}} \tag{19}$$

式中：$\text{dist}(c_{i-1},c_i)$为路径点c_{i-1}至c_i的距离；$|\boldsymbol{v}_{\text{g},i-1}|$为无人机由$c_{i-1}$航行至$c_i$的地速，由式(8)与式(9)计算；$v_{\text{g,est}}$为无人机航行的预估地速。

2）节点选择改进。为避免非必要节点产生，使用视线函数检查非相邻节点间是否存在直线连通路径。设风速变化阈值为$\Delta v_{w,max}$，风向变化阈值为$\Delta\theta_{w,max}$，若非相邻节点直线连接间任意处风速与风向变化不超出各自阈值，且连线上不包含不可进入栅格，则两点间存在连通路径并进行直线连接。

应用上述改进，算法以含风场信息的空域栅格为输入，在起点与终点间规划路径，并输出航路节点与航路所需飞行时间，流程如图 5 所示。

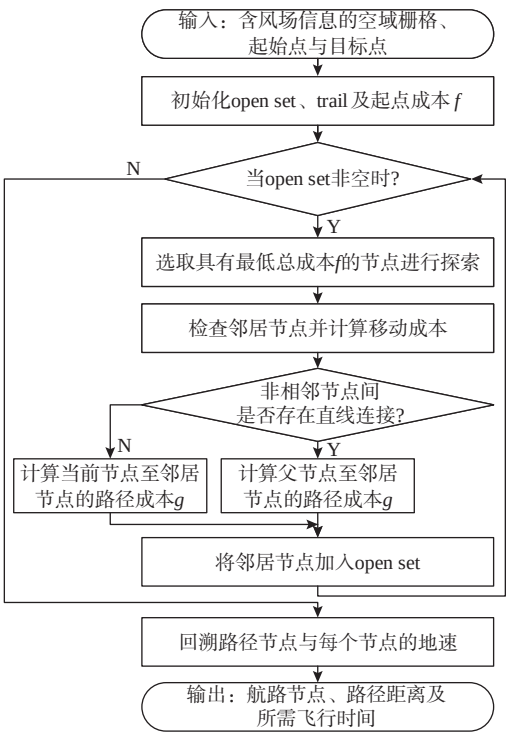


图 5　Theta*算法流程

Fig. 5　Theta* algorithm flowchart

步骤 1　初始化，包括优先队列 open set 与列表 trail。优先队列 open set 用于存储待探索节点及其总成本值 f，列表 trail 用于记录到达每个节点的最优路径中的父节点。总成本计算见式 (17)～式 (19)。

步骤 2　主循环。从起点开始，当优先队列非空时，选择具有最低总成本 f 的节点作为当前节点进行探索，迭代处理直到到达目标节点。到达目标节点后跳转至最终输出步骤 5。

步骤 3　路径节点选择。对于当前节点，检查所有邻居节点并计算移动成本，相邻节点的移动成本为两点间航行所需时间。禁飞区与建筑物栅格无法作为邻居节点被选择。对于非相邻节点，使用视线函数检查是否存在直线连通路径。若存在，则计算父节点至邻居节点的临时路径成本；否则，计算当前节点至邻居节点的临时路径成本。

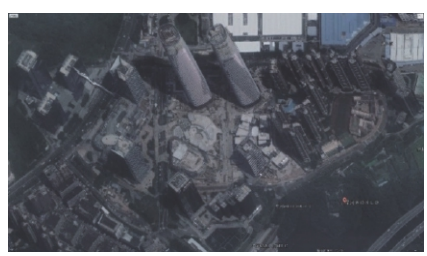
步骤 4　成本更新。若临时路径成本低于已记录的路径成本，则更新列表 trail 及邻居节点成本，并将其加入 open set 用于后续探索。

步骤 5　最终输出。到达目标节点后，算法将回溯列表 trail 并输出最优路径，累加路径中各点的飞行时间以获得路径总飞行时间。

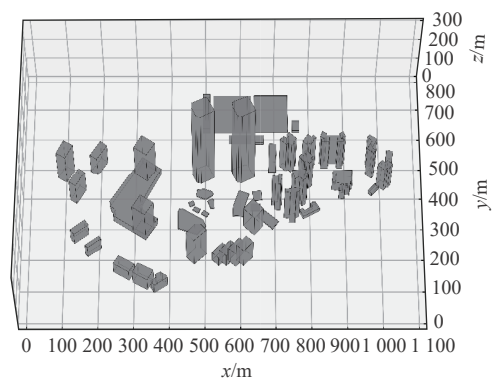
3　无人机路径规划实例

3.1　研究区域

本节选取对无人机运行挑战较大的城市中心商务区进行实例研究。以深圳市龙岗区中心商务区为研究场景，区域范围约为 1.1 km（东西向）×0.8 km（南北向），其间建筑物密集且高大建筑物较多，多数高度超 100 m。研究区域现配备 1 个无人机枢纽机场与 6 个子机场，用于小型无人机物流运输。子机场均位于枢纽机场直线距离 1 km 范围内。小型无人机飞行高度位于 60～120 m 之间，因而需避开大量地面障碍物。该配送网络日均订单可达 200 单，无人机在楼宇间航行频繁。目前的执飞机型为美团第三代无人机，机身质量 4 kg，物流配送载重可达 2.4 kg。该无人机为六旋翼无人机，抗风等级为 5 级风。对研究区域构建城市模型，以 OSM 为主要数据源，从中提取区域内关键建筑物的坐标、尺寸与高度信息，并添加无人机机场位置。使用 Python 语言，以 PyCharm 为集成开发环境，构建研究区域的三维城市模型，如图 6 所示。



(a) 卫星地图 (来源: Google Earth)



(b) 三维城市模型

图 6　研究区域地图与模型

Fig. 6　Map and model of study area

3.2　主导风分析

为了解该地区主导风情况，从 ERA5 气象数据

库提取研究区域的地表主导风数据。通过指定经纬度，提取区域内2014—2023年地表月均风场数据，包括100 m高度风速分量及10 m高度瞬时阵风。使用风速分量计算主导风风速、风向并统计出现频率，得到表征地表100 m高度的主导风统计图，如图7所示。研究区域地表主导风以东北风、东风及南风为主，主导风风速主要为3级风(3.4～5.4 m/s)，在部分风向下可达4级风(5.5～7.9 m/s)。研究区域瞬时阵风在部分时段可达5级风(8.0 m/s)。综上所述，选取有代表性风场9组进行CFD风场仿真，如表1所示。

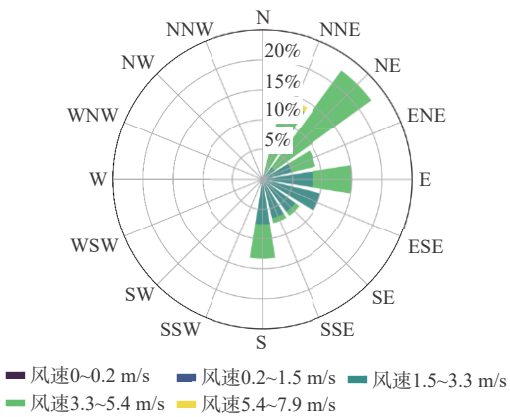


图7 研究区域主导风风速、风向特征

Fig. 7 Characteristics of prevailing wind speed and direction in study area

表1 CFD风场仿真输入风速、风向

Table 1 Input wind speed and direction for CFD wind field simulation

主导风风向	主导风风速/(m·s^{-1})	风力级数	场景编号
东北风	4.3	3级风	场景1
	6.7	4级风	场景2
	8.0	5级风	场景3
东风	4.3	3级风	场景4
	6.7	4级风	场景5
	8.0	5级风	场景6
南风	4.3	3级风	场景7
	6.7	4级风	场景8
	8.0	5级风	场景9

3.3 CFD风场仿真

依据主导风分析设置风场仿真入口边界条件。风向选择东北风、东风与南风，风速使用平均风速剖面公式描述，如下：

$$U(z)=U_{\mathrm{r}}\left(\frac{z}{z_{\mathrm{r}}}\right)^{\gamma} \tag{20}$$

式中：$U(z)$为离地表参考高度z处的平均风速；U_{r}为离地表参考高度z_{r}处的平均风速；γ为地面粗糙指数，其数值与地面粗糙度类别有关。此处，z_{r}取100 m，γ取0.4。仿真风速、风向如表1所示，共计9种场景。将入口边界条件输入至CFD求解器，湍流模型选择RANS模型。

仿真完成后，对流体域内风速分量进行后处理与可视化。选取9种场景下100 m高度层风速展示，如图8所示，以热力图呈现风场。整体而言，主导风风向决定了区域内整体的风场流向，区域内风速随主导风风速的增加而提升。此外，建筑物群会改变来流风速、风向，使得城市低空风场进一步复杂化。具体地，对比主导风为东北风、东风与南风时的风场仿真结果，分别如场景2、场景5、场景8所示。东北风时，北侧建筑物间出现了明显的风流加速，风速接近8.0 m/s，且影响范围较大。其中，枢纽机场北侧城市峡谷造成的风速提升区域位于枢纽机场上方，对路径规划有较大影响。主导风为东风时，区域中心风速较低，因为位于东侧的密集高大建筑物群阻碍了来风。主导风为南风时，北侧2座较高建筑物(369 m)的两侧与城市峡谷间风速大幅提升，超出8.0 m/s，在规划北侧投递柜路径时需避开此区域。对比不同主导风风速时的风场仿真结果，以场景1～场景3为例。3级风(4.3 m/s)时，区域内整体风速较低，建筑物后方出现了大面积的低风速区域。主导风达到4级风(6.7 m/s)时，区域内风速明显提升，高风速区域主要位于建筑物上方与城市峡谷之间。当主导风为5级风(8.0 m/s)时，区域内风速显著增大，部分位置甚至超出10.7 m/s，严重影响无人机飞行。

3.4 飞行禁区确定

确定飞行禁区前应进行空域建模。本文关注的研究区域空域大小为1 100 m×800 m×600 m，考虑的小型多旋翼无人机轴距约为0.96 m，运输航线距离为1 km以内(单程)，CFD网格在建筑物周围尺寸较小，最小尺寸不足1 m。考虑到计算开销与空域栅格模型精度之间的权衡，本文选取2.5 m作为空域栅格尺寸。基于空域栅格模型，依据风场仿真结果，对9种场景划设飞行风险区与障碍物影响范围。以100 m高度层为例，如图9所示。

1) 飞行风险区。将风速大于10.7 m/s的风速超标区域划设为飞行风险区，如图9中红色区域所示。可以看出，飞行风险区主要出现在场景3与场景9中(主导风风速8.0 m/s)。场景3中，飞行风险区跨越北侧城市峡谷与枢纽机场东侧，路径规划时需进行大幅度绕行。场景9中，飞行风险区位于北侧建筑物之间，北侧投递柜的路径规划受到较大影响。

2) 障碍物影响范围。将建筑物周边半径20 m

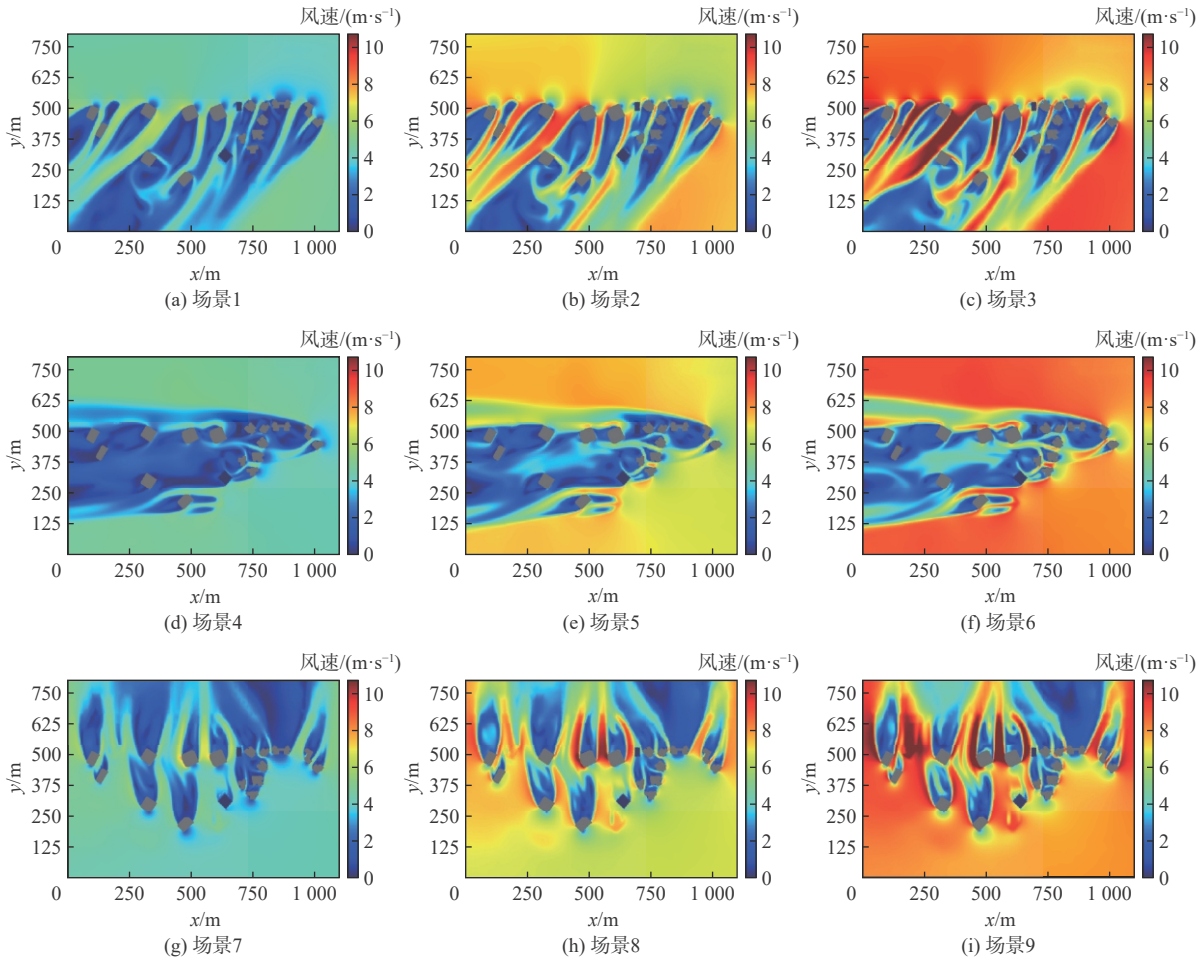


图 8　不同主导风下 100 m 高度风速热力图

Fig. 8　Heatmap of wind speeds at 100 m altitude under different prevailing winds

内湍流动能大于 1.5 m^2/s^2 的湍流区域及距建筑物 5 m 以内的区域划设为障碍物影响范围。其中, 20 m 为部分物流配送无人机使用的建筑物安全距离。障碍物影响范围如图 9 所示, 图中深灰区域为建筑物, 浅灰区域为风场下障碍物影响范围。可以看出, 障碍物影响范围与主导风风向、风速密切相关。风速较高时, 建筑物附近会出现强烈的风切变区域, 导致湍流扰动的增强。此外, 湍流动能较高的区域主要位于建筑物周围及其尾流位置, 因此, 风向也会改变障碍物影响范围。值得注意的是, 部分建筑物间的区域被划设为障碍物影响范围, 表明这些城市峡谷间不再适于无人机的安全飞行。此外, 一些障碍物影响范围内并不包含建筑物, 表明此类区域位于建筑物上方, 湍流扰动较强或距建筑物较近。

3.5　路径规划

不同风场下, 60 ～120 m 高度层, 对枢纽机场与物流配送投递柜间进行往返路径规划, 路径规划模型参数如表 2 所示。运行仿真的计算机处理器为 Intel Core i9-13900HX, 内存 16 GB。本文聚焦于低高度空域内航路的水平飞行阶段, 以 9 种场景下 100 m 高度层路径规划结果为例展示, 如图 10 所示。可以看出, 路径有效避开了飞行风险区与障碍物影响范围, 受风速、风向的影响而变化。以飞行时间为目标函数, 算法合理地利用了风场, 面对顶风趋向于绕行, 面对顺风则加以利用以减小飞行时间。

对比不同主导风下的路径规划结果, 如表 3 所示。主导风为东北风时(场景 1～场景 3), 路径规划主要影响来源于北侧城市峡谷间的风流加速, 且影响程度随着主导风风速的增加而上升。主导风风速较低时(场景 1), 路径比较平滑, 平均路径距离较短, 且路径的平均飞行时间(指水平阶段飞行时间)较小。主导风风速增高后(场景 2), 障碍物影响范围有所扩大, 路径绕行更加明显, 平均路径距离与所需飞行时间均有所增加。当主导风达 5 级风时(场景 3), 大面积的风速超标区及建筑物周边的湍流区域导致路径需进行大范围的绕行, 平均路径距离与飞行时间均大幅提升, 为 9 种场景下最高。此时, 路径规划已变得较为困难, 只能采取绕行策略, 甚至取消配送任务。此外, 在风场影响下, 往返

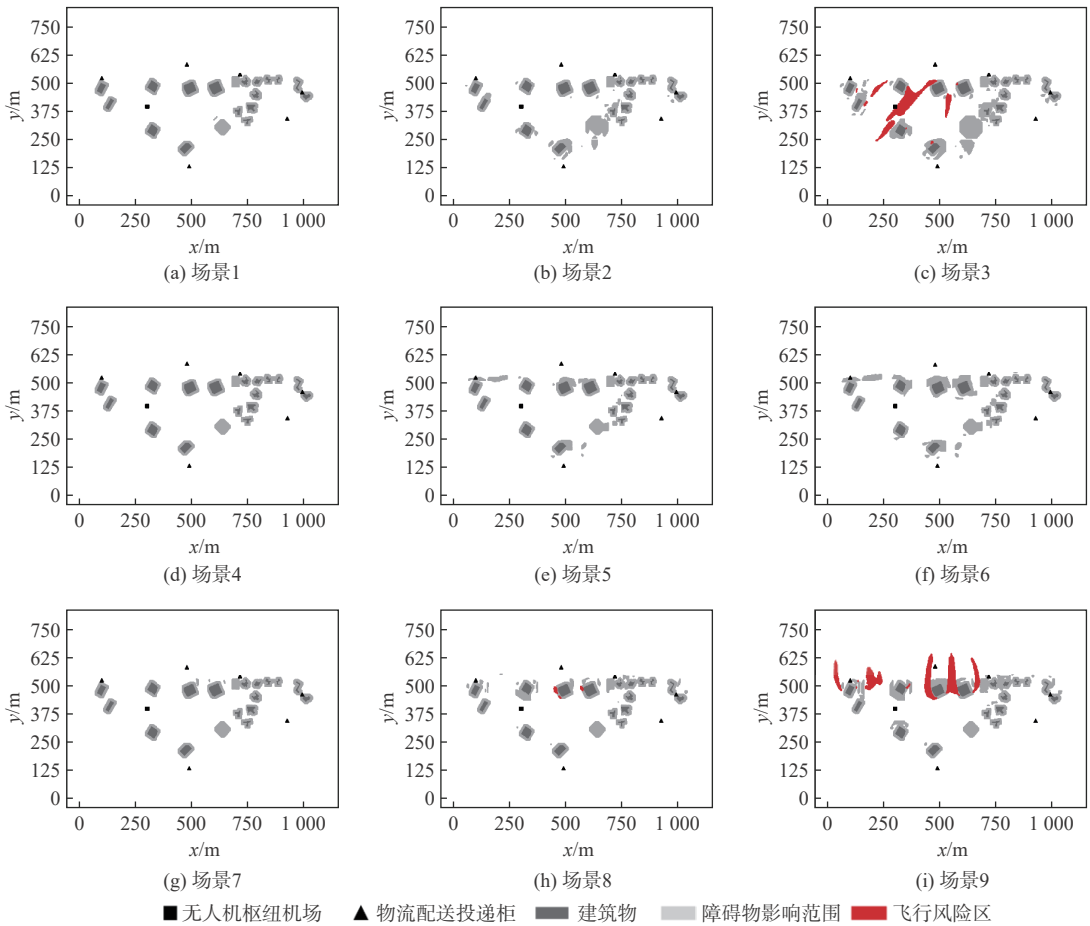


图 9 不同主导风下 100 m 高度飞行风险区与障碍物影响范围

Fig. 9 Hazardous flight regions and obstacle impact zones at 100 m altitude under different prevailing winds

表 2 路径规划模型参数

Table 2 Path planning model parameters

参数	取值
风场在c_i处风速$\boldsymbol{v}_{\mathrm{w},i}$	CFD仿真结果
无人机在c_i处空速$\boldsymbol{v}_{\mathrm{a},i}$	大小为12.5 m/s
无人机在c_i处地速$\boldsymbol{v}_{\mathrm{g},i}$	由式(8)与式(9)计算
风场在c_i处湍流动能k_i	CFD仿真结果
最小路径段长度$l_{\min}$/m	2.5
无人机最大航程$L_{\max}$/m	3 000
无人机最大转弯角$\beta_{\max}$/rad	π/2
无人机最低飞行高度$H_{\min}$/m	60
无人机最高飞行高度$H_{\max}$/m	120
无人机可承受最大风速$v_{\mathrm{w,res}}$/(m·s^{-1})	10.7
湍流动能阈值k_{res}/(m^2·s^{-2})	1.5
湍流动能检测半径r_k/m	20
建筑物安全缓冲区半径r_{bld}/m	5
空域模型栅格大小l_{gird}/m	2.5
建筑物栅格集合C_{bld}	三维城市模型
无人机禁飞区栅格集合$C_{\mathrm{no\text{-}fly}}$	空域栅格模型

路径通常存在差异。由于物流配送投递柜多位于枢纽机场东侧或北侧，东北风时配送路径距离与飞行时间均显著高于返回路径距离与飞行时间。

主导风为东风时(场景 4～场景 6)，由于东侧较高的密集建筑物群的遮挡，风场对路径规划的影响较其他场景而言较小，平均路径距离较低且平均飞行时间较短。因为 6 个物流配送投递柜中 5 个位于枢纽机场东侧，在东风下配送路径距离与飞行时间均高于返回路径距离与飞行时间。主导风为南风时(场景 7～场景 9)，路径间差异因风速的提升而变大。主导风风速较低时(场景 7)，风场对路径规划影响较小。主导风风速增加后(场景 8)，平均路径距离有所增加且往返路径间差异更加明显。主导风达 5 级风后(场景 9)，北侧较高建筑物间出现了多处风速超标区，与场景 3 类似，路径需大幅绕行以避开飞行风险区与障碍物影响范围。南风影响下，配送路径所需飞行时间小于返回路径的所需飞行时间。

3.6 结果分析

为验证算法有效性，本节对比以最小化路径时间为目标的改进 Theta*与改进 A*算法，以及以最小化路径距离为目标的基础 Theta*与基础 A*算法所规划航路差异，并从安全与效率的角度分析。在 9 种场景下规划路径，结果如表 4 所示。以场景 2 为例，不同算法在 100 m 高度规划路径如图 11 所示。

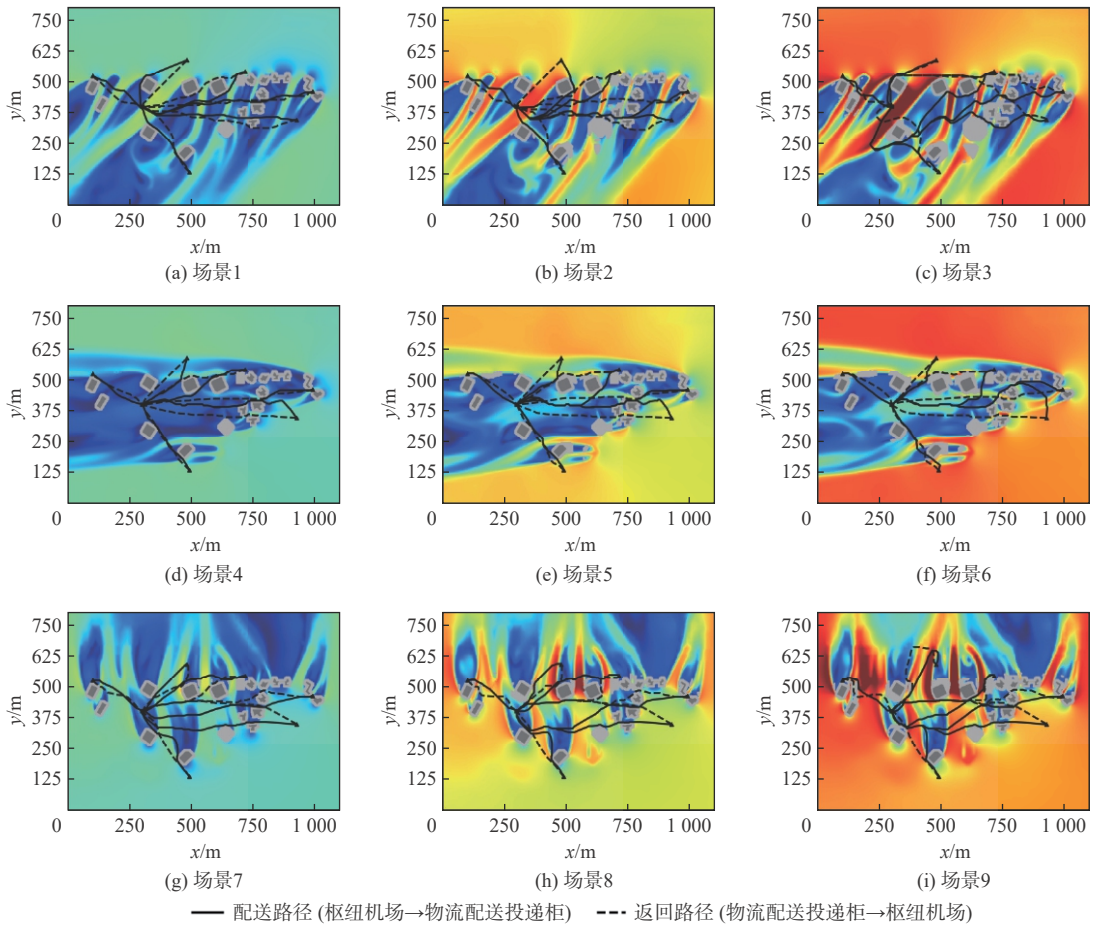


图 10　不同主导风下 100 m 高度路径规划结果

Fig. 10　Path planning results at 100 m altitude under different prevailing winds

表 3　不同主导风下航路数据

Table 3　Route data under different prevailing winds

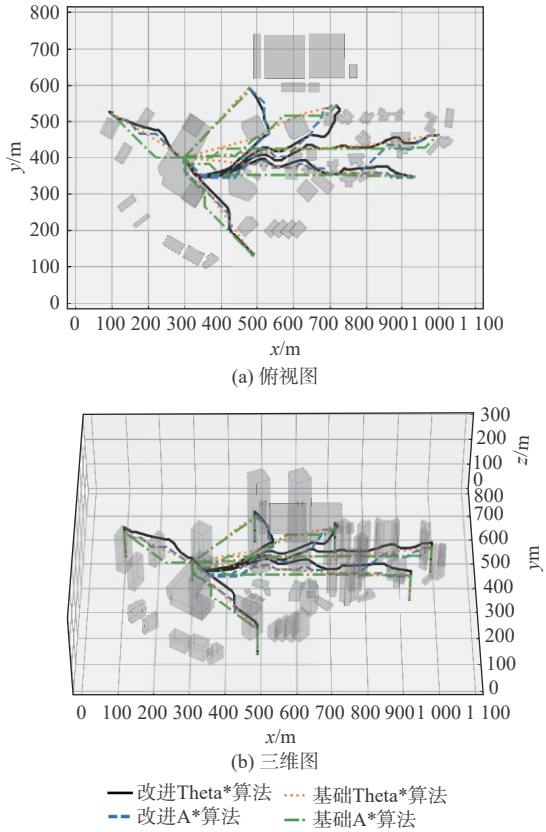
主导风风向	场景编号	平均路径距离/m	配送/返回路径距离/m	平均飞行时间/s	配送/返回飞行时间/s
东北风	场景1	452.55	459.32/445.79	36.19	39.25/33.12
	场景2	469.19	480.15/458.23	38.15	42.93/33.38
	场景3	630.47	694.95/565.99	48.18	56.82/39.54
东风	场景4	447.23	455.80/438.66	35.21	38.51/31.91
	场景5	455.53	468.09/442.97	35.64	40.98/30.30
	场景6	479.07	497.12/461.01	36.92	42.98/30.86
南风	场景7	451.47	452.24/450.69	35.98	35.81/36.16
	场景8	489.72	487.92/491.53	39.22	38.07/40.37
	场景9	563.36	560.47/566.24	44.58	42.78/46.39

表 4　不同算法路径规划结果对比

Table 4　Comparison of path planning results using different algorithms

算法	路径距离/m	飞行时间/s	飞行风险区/m	障碍物影响范围/m	拐点数	计算时间/s	不可通行路径条数
改进Theta*算法	482.81	38.45			11.25	1.85	0
改进A*算法	495.42	39.09			16.49	4.95	0
基础Theta*算法	432.07	42.38	7.46	41.00	0.22	0.08	3
基础A*算法	450.28	43.98	6.40	37.42	6.67	0.07	3

对比改进 Theta*算法与改进 A*算法路径规划结果, 两者均考虑风场对无人机飞行速度影响, 并避开飞行风险区与障碍物影响范围, 所规划路径整体上呈一致趋势。相比于改进 A*算法, 改进 Theta*



(a) 俯视图

(b) 三维图

图 11 不同算法路径规划结果

Fig. 11 Path planning results using different algorithms

算法所得路径距离与飞行时间更短，分别减少了2.55%与1.64%，且单条路径平均拐点数明显少于改进A*算法。此外，改进Theta*算法规划单一路径所需计算时间为1.85 s，相比改进A*算法减少了62.63%。与改进Theta*算法及改进A*算法不同，基础Theta*算法与A*算法未考虑飞行风险区与障碍物影响范围，因而路径距离普遍更短。基础Theta*算法可以任意角度规划路径，路径距离为4种算法中最低，且单一路径拐点数量接近于0。但是，基础Theta*算法进入了飞行风险区与障碍物影响范围，长度分别为7.46 m与41.00 m。此外，由于未依据风向调整航向，基础Theta*算法所规划路径的飞行时间高于改进Theta*与改进A*算法。在面对飞行风险区时（对应场景3与场景9），出现了无人机无法以12.5 m/s的空速继续沿航路向目标点飞行的情况，因而规划结果含3条不可通行路径。基础A*算法的路径距离与所需飞行时间略高于基础Theta*算法，但拐点数较多。类似地，基础A*算法规划结果中包含3条不可通行路径。

整体而言，改进的Theta*算法与A*算法的路径距离高于基础Theta*算法与A*算法，但飞行时间更低，且未包含不可通行路径。就运行效率而言，基础Theta*算法与A*算法以距离为目标，计算开销显著低于改进Theta*算法与A*算法，因而计算时间更低。但是，本文研究的路径规划为飞行前的战略或预战术操作，因此单一路径平均1.85 s的计算时间完全处于可接受范围内。与改进A*算法相比，改进Theta*算法在路径距离、飞行时间、拐点数量及计算时间方面均效果更优。

4 结 论

本文基于城市风场对无人机影响分析构建了路径规划模型，使用CFD仿真估计城市风场，并使用所提算法求解路径规划问题。主要结论如下：

1）所提算法有效识别了飞行风险区与障碍物影响范围，在绕行避开飞行禁区的基础上，路径距离有所增加，但相比于基础Theta*算法减少了9.27%的飞行时间。所提算法在提升航路安全性的同时优化了路径效率。

2）使用了ERA5气象数据对研究区域内近10年地表主导风风速、风向与出现频率及地表阵风风速进行分析。基于主导风分析与城市模型进行CFD仿真，可为风场下的无人机路径规划提供可靠的仿真环境。

3）所提算法的路径距离与飞行时间均低于改进A*算法，且航路拐点数量减少了31.78%，计算时间减少了62.63%，算法效率大幅提升。

未来研究可针对不同机型进行测试，以更加细化地评估飞行风险区与障碍物影响范围。风场对无人机飞行速度影响可通过动力学模型或加入控制算法进行模拟。风场对无人机能耗影响则有待进一步研究并加入路径规划模型。此外，还可研究动态变化风场对无人机路径规划的影响。

参考文献（References）

[1] AGGARWAL S, KUMAR N. Path planning techniques for unmanned aerial vehicles: a review, solutions, and challenges[J]. Computer Communications, 2020, 149: 270-299.

[2] DAI W, PANG B Z, LOW K H. Conflict-free four-dimensional path planning for urban air mobility considering airspace occupancy[J]. Aerospace Science and Technology, 2021, 119: 107154.

[3] PANG B Z, HU X T, DAI W, et al. UAV path optimization with an integrated cost assessment model considering third-party risks in metropolitan environments[J]. Reliability Engineering & System Safety, 2022, 222: 108399.

[4] DELAMER J A, WATANABE Y, CHANEL C P C. Safe path planning for UAV urban operation under GNSS signal occlusion risk[J]. Robotics and Autonomous Systems, 2021, 142: 103800.

[5] HE X Y, HE F, LI L S, et al. A route network planning method for urban air delivery[J]. Transportation Research Part E: Logistics and Transportation Review, 2022, 166: 102872.

[6] MURRAY C W A, IRELAND M L, ANDERSON D. On the res-

ponse of an autonomous quadrotor operating in a turbulent urban environment[C]//Proceedings of the AUVSI's Unmanned Systems Conference. Orlando: AUVSI, 2014.

[7] MCGRATH B, CYBYK B Z, FREY T M. Environment-vehicle interaction modeling for unmanned aerial system operations in complex airflow environments[J]. Johns Hopkins APL Technical Digest, 2012, 32(2): 115-131.

[8] THIBBOTUWAWA A, NIELSEN P, ZBIGNIEW B, et al. Energy consumption in unmanned aerial vehicles: a review of energy consumption models and their relation to the UAV routing[C]//Proceedings of 39th International Conference on Information Systems Architecture and Technology. Berlin: Springer, 2019: 173-184.

[9] GALWAY D, ETELE J, FUSINA G. Modeling of the urban gust environment with application to autonomous flight: AIAA 2008-6565[R]. Reston: AIAA, 2008.

[10] SUTHERLAND M. Urban wake field generation using LES for application to quadrotor flight[D]. Ottawa: Carleton University, 2015.

[11] RAZA S A. Autonomous UAV control for low-altitude flight in an urban gust environment[D]. Ottawa: Carleton University, 2015: 40-147.

[12] LIU S M, PAN W X, ZHANG H, et al. CFD simulations of wind distribution in an urban community with a full-scale geometrical model[J]. Building and Environment, 2017, 117: 11-23.

[13] KANG G, KIM J J, CHOI W. Computational fluid dynamics simulation of tree effects on pedestrian wind comfort in an urban area[J]. Sustainable Cities and Society, 2020, 56: 102086.

[14] YANG L, LIU X D, QIAN F, et al. Research on the wind environment and air quality of parallel courtyards in a university campus[J]. Sustainable Cities and Society, 2020, 56: 102019.

[15] JUNG J, DSOUZA S N, JOHNSON M A, et al. Applying required navigation performance concept for traffic management of small unmanned aircraft systems[C]//Proceedings of the Congress of the International Council of the Aeronautics Sciences. Daejeon: Korean Society for Aeronautical and Space Sciences, 2016.

[16] JEONG S, YOU K, SEOK D. Hazardous flight region prediction for a small UAV operated in an urban area using a deep neural network[J]. Aerospace Science and Technology, 2021, 118: 107060.

[17] GIERSCH S, EL GUERNAOUI O, RAASCH S, et al. Atmospheric flow simulation strategies to assess turbulent wind conditions for safe drone operations in urban environments[J]. Journal of Wind Engineering and Industrial Aerodynamics, 2022, 229: 105136.

[18] RUBIO-HERVAS J, GUPTA A, ONG Y S. Data-driven risk assessment and multicriteria optimization of UAV operations[J]. Aerospace Science and Technology, 2018, 77: 510-523.

[19] ALTUN A T, XU Y, INALHAN G, et al. Comprehensive risk assessment and utilization for contingency management of future AAM system: AIAA 2023-3687[R]. Reston: AIAA, 2023.

[20] WARE J, ROY N. An analysis of wind field estimation and exploitation for quadrotor flight in the urban canopy layer[C]//Proceedings of the IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2016: 1507-1514.

[21] THANELLAS G A, MOULIANITIS V C, ASPRAGATHOS N A. A spatially wind aware quadcopter (UAV) path planning approach[J]. IFAC-PapersOnLine, 2019, 52(8): 283-288.

[22] HONG D, LEE S, CHO Y H, et al. Energy-efficient online path planning of multiple drones using reinforcement learning[J]. IEEE Transactions on Vehicular Technology, 2021, 70(10): 9725-9740.

[23] MIKHAILUTA S V, LEZHENIN A A, PITT A, et al. Urban wind fields: Phenomena in transformation[J]. Urban Climate, 2017, 19: 122-140.

[24] 张洪海, 李翰, 刘皞, 等. 城市区域物流无人机路径规划[J]. 交通运输系统工程与信息, 2020, 20(6): 22-29.
ZHANG H H, LI H, LIU H, et al. Path planning for logistics unmanned aerial vehicle in urban area[J]. Journal of Transportation Systems Engineering and Information Technology, 2020, 20(6): 22-29(in Chinese).

Small unmanned aerial vehicles path planning method considering urban low-altitude wind fields

ZHAO Yifei[1,*], GU Ruijia[1], REN Xinhui[2]

(1. School of Air Traffic Management, Civil Aviation University of China, Tianjin 300300, China;
2. School of Economics and Management, Civil Aviation University of China, Tianjin 300300, China)

Abstract: The operational safety of drone routes has gained more attention as unmanned aerial vehicles (UAVs) are used extensively in urban logistics. Addressing the safety and efficiency issues caused by urban wind fields, this paper integrates the operational characteristics of logistics drones and employs computational fluid dynamics (CFD) to estimate urban wind fields. The impact of wind fields on UAVs is analyzed from the perspectives of flight safety, variations in drone speed, and turbulence zones around buildings. A path planning model is constructed, and an improved Theta* algorithm is proposed. The study takes the central business district of Longgang District, Shenzhen, as the research scenario. Simulations of urban wind fields are conducted based on historical prevailing winds to delineate hazardous flight regions and the influence range of obstacles within the wind field. The proposed method is then used to plan drone routes under different wind field conditions. Results show that the constructed path planning model effectively identifies high wind speed and turbulent areas within the airspace. When compared to routes designed to minimize distance, the drone route saves flying time by 9.27% by taking into account how wind fields affect drone speed and avoiding no-fly zones. Additionally, the proposed algorithm outperforms traditional algorithms in terms of route distance and flight time, and reduces the number of turns by 31.78% and computation time by 62.63%.

Keywords: unmanned aerial vehicle; urban wind field; path planning; logistics distribution; urban air mobility

Received: 2024-05-06; **Accepted**: 2024-07-12; **Published Online**: 2024-07-23 10:34

URL: link.cnki.net/urlid/11.2625.V.20240722.1809.002

Foundation item: National Key Research and Development Program of China (2022YFB4300904)

* **Corresponding author**. E-mail: yfzhao@cauc.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0204

核函数约束的S模式信号脉冲结构简化及TOA精确度研究

宫峰勋*，刘焘
(中国民航大学 电子信息与自动化学院，天津 300300)

摘　　要：民航S模式应答信号空时结构复杂，易受到传输链路非线性影响和临近电磁干扰，严重影响到达时间(TOA)高精确提取。在民航S模式应答信号空时结构与频谱分析基础上，提出S模式信号脉冲结构简化的准S模式信号模型，根据准S模式信号、弱非线性系统特性反演约束的核函数，构建简化Volterra级数模型。仿真结果表明，基于准S模式信号的简化Volterra级数模型的核函数收益倍数不低于27倍。采用3阶简化Volterra级数模型与高次频点频谱补偿组合方式，对实测信噪比约为13 dB的受干扰S模式信号的准S模式信号进行恢复，实现了用约8.8%的准S模式信号模型的计算量得到约1.54%波形恢复误差，且恢复信号比受干扰信号提取TOA的准确度提升超过73%。基于核函数约束的准S模式信号简化Volterra级数模型，对民航S模式应答信号TOA的准确估计、定位精确度提升具有重要理论价值。

关　键　词：核函数；准S模式信号；简化Volterra级数；频谱补偿；到达时间；多点定位
中图分类号：V351.3; TN851
文献标志码：A　　**文章编号**：1001-5965(2026)06-1789-14

传输链路非线性导致空管二次雷达应答S模式信号波形失真，严重影响到达时间(time of arrival, TOA)估计的精确度及目标定位精确度[1-4]。在有线信道链路传输过程中，接收机前端非线性器件的大范围使用是导致传输波形失真的主要原因。Lenhard等采用功率放大器预失真回退方法保证波形线性变化，但大规模组件回退不利于波形上升沿波峰的描述[5]。Vansebrouck等根据传递函数对各个模块线性补偿和记忆校正得到失真前波形，但一般不能预先知道非线性传递函数[6]。若对整个非线性接收机系统进行建模研究，Maiti等采用反馈模型线性化修正，但失真波形收敛速度会影响TOA估计[7]。曹韬等提出包络增强型多项式预失真模型表征幅度非线性，但对于波形上升沿包络，跟踪吻合度不高[8]。Obais采用非线性相关变换和线性化恢复，但时延滞后状态波形不利于TOA的提取[9]。因此，以上方案对空管二次雷达设备传输S模式信号失真波形的描述存在不足，用于上升前沿TOA估计性能较差。

通常，空管二次雷达接收机是弱非线性记忆系统，利用频域Volterra级数对频域核的非参数识别进行建模[10]，可以用任意序列对系统进行非线性逼近，进而描述传输波形失真变化，准确估计无失真S模式上升前沿。但为避免建模阶数上升和记忆存储长度增长导致待辨识核函数的大幅增多，一般最多只用3阶传统Volterra级数建模[11]，故简化Volterra级数与提升模型精度的方法涌现，如截断法[12]、记忆多项式模型[13]、正交三角分解[14-15]等。在系统“近对角线”结构中移除长记忆分量和高阶分量，降低待辨识核函数数目的同时也会降低Volterra级

收稿日期：2024-04-11；**录用日期**：2024-07-26；**网络出版时间**：2024-09-09 16:39
网络出版地址：link.cnki.net/urlid/11.2625.V.20240909.1304.004
基金项目：国家重点研发计划(2018YFC0809500)
*通信作者. E-mail：fxgong@cauc.edu.cn

引用格式：宫峰勋，刘焘. 核函数约束的S模式信号脉冲结构简化及TOA精确度研究[J]. 北京航空航天大学学报，2026，52(6)：1789-1802. GONG F X，LIU T. Simplified S-mode signal pulse structure with kernel function constraints and TOA accuracy study[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1789-1802 (in Chinese).

数对非线性系统的表示能力；或将 Volterra 核函数表示为基函数的核回归问题，但由于系统弱非线性项的阶数、记忆长度及非零项的位置和个数都是未知的，难以选定合适的数值。

因此，只考虑未知非线性系统本身简化 Volterra 级数较为困难。但 Volterra 表征非线性过程是通过核函数与输入信号组合频率分量的乘积得到输出波形，如果知道系统输入的先验信号，对不同输入信号进行差异化表征，就可能实现描述波形精度高且用到的核函数少。Faifer 等分别提出基于准正弦信号输入的简化 Volterra 级数方法[16]和只考虑输入信号谐波失真的简化 Volterra 级数的系统非线性建模方法[17]，其核心是利用已知输入信号频谱特性求取核函数实现系统非线性的简化 Volterra 级数表征，效果显著，但对不同于准正弦的 S 模式输入信号的建模误差较大。此外，若用频域 Volterra 级数描述 S 模式信号输入情况，复杂 S 模式信号过多的频谱分量也会对 Volterra 级数的维数造成极大干扰，导致待辨识核函数过多，不利于建模。

针对以上问题，在 S 模式信号频谱特征分析基础上，本文设计基于核函数约束的 S 模式前导脉冲简化态结构，改善原频谱分量数量，以便频域 Volterra 级数建模；基于该结构基频规律简化 Volterra 级数，达到模型精确度与维数的均衡；对于表征精确度不足的高次频点引入频谱补偿模型，以最终实现前导脉冲波形在二次雷达接收机传输链路的有效非线性表征。仿真结果表明，本文模型可有效提升 S 模式信号前导脉冲识别精确度，降低建模计算量，有效应对系统非线性造成 TOA 估计不准的问题。

1 准 S 模式信号及简化级数

1.1 准 S 模式信号

民航机载应答 S 模式基带信号的数学模型 $u_s(t)$ 可表示为[2]

$$u_s(t)=\sum_{q=1}^{128/240}u(q)p(t-qT) \tag{1}$$

式中：T 为脉冲宽度；$p(t)$ 为 S 模式矩形脉冲，$p=\{1,0,1,0,0,0,0,1,0,1,0,0,0,0,0,0\}$；$u(q)$ 为脉冲信号。

S 模式基带信号波形如图 1(a) 所示。基于到达时间差定位算法中多考虑从前导脉冲中提取 TOA 参数，因此，为研究方便，可将 S 模式基带信号简化为包含 4 个脉冲、周期为 8 μs 的脉冲序列及包含 4 个脉冲、周期为 7 μs 的脉冲序列，其波形如图 1(b) 和图 1(c) 所示结构。

由图 1(a) 可见，S 模式应答信号由 4 个时空结构固定的前导脉冲和 1 个 56/112 位的数据块组成，

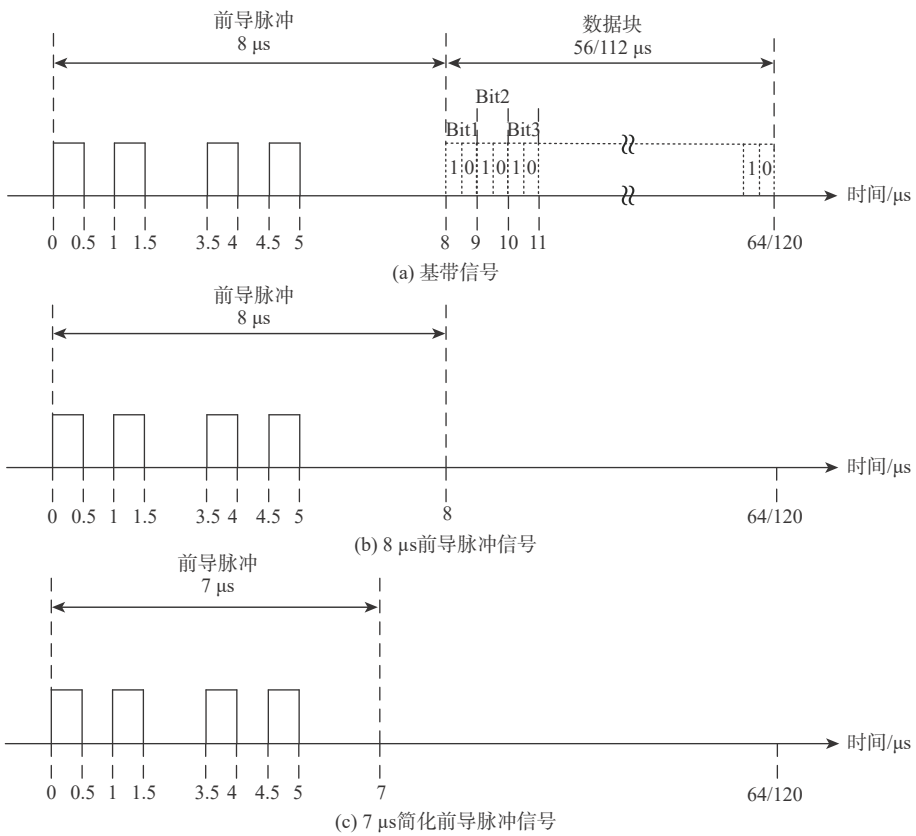


图 1 S 模式基带信号及简化前导脉冲波形结构

Fig. 1 S-mode baseband signal and simplified leading pulse waveform structure

其频谱特征如图 2(a)、(b) 所示。8 μs 前导脉冲的连续谱及以 8 μs 为周期的离散谱如图 2(c)、(d) 所示，其中，离散谱间隔为 1/8 MHz。7 μs 前导脉冲的连续谱及以 7 μs 为周期的离散谱如图 2(e)、(f) 所示，其中，离散谱间隔为 2/7 MHz。

由图 2(a) 和图 2(b) 可见，S 模式信号频谱谐波能量较直流分量很小，难以分辨，且机载应答 S 模式信号不是周期信号，无法得到离散频谱。又由图 2(b) 和图 2(c) 可见，S 模式信号与 8 μs 的 S 模式前导脉冲的连续谱波形非常接近，且 8 μs 的 S 模式前导脉冲能得到频谱幅值易于提取的离散谱，如图 2(d) 所示，因此，基于时差定位算法只用考虑前导脉冲，将前导脉冲独立提出利用离散谱进行频域 Volterra 级数分析是可行的。故以下分析所用 S 模式信号均指图 1(b) 所示 8 μs 的 S 模式前导脉冲。

如图 2(c) 和图 2(e)、图 2(d) 和图 2(f) 所示，8 μs 前导脉冲与 7 μs 前导脉冲的连续谱和离散谱都非常接近。此外，图 2(e) 所对应的 7 μs 前导脉冲从结构上可以看成 2 个 3.5 μs 的波形相加，以 7 μs 为周期可以得到更少的离散谱线。图 2(f) 谱线间隔为 2/7 MHz，图 2(d) 谱线间隔为 1/8 MHz，间隔更大、分量更少的离散频谱，利于降低频域 Volterra 级数维数。因此，设计与原 S 模式信号频谱相似但频谱更为简单的脉冲结构用频域 Volterrra 级数表征非线性传输，是一种有利于简化建模的约束 Volterra 核函数数量的方式。因此，采用图 1(c) 所示周期为 7 μs 的基带 4 脉冲信号代替 S 模式应答信号研究分

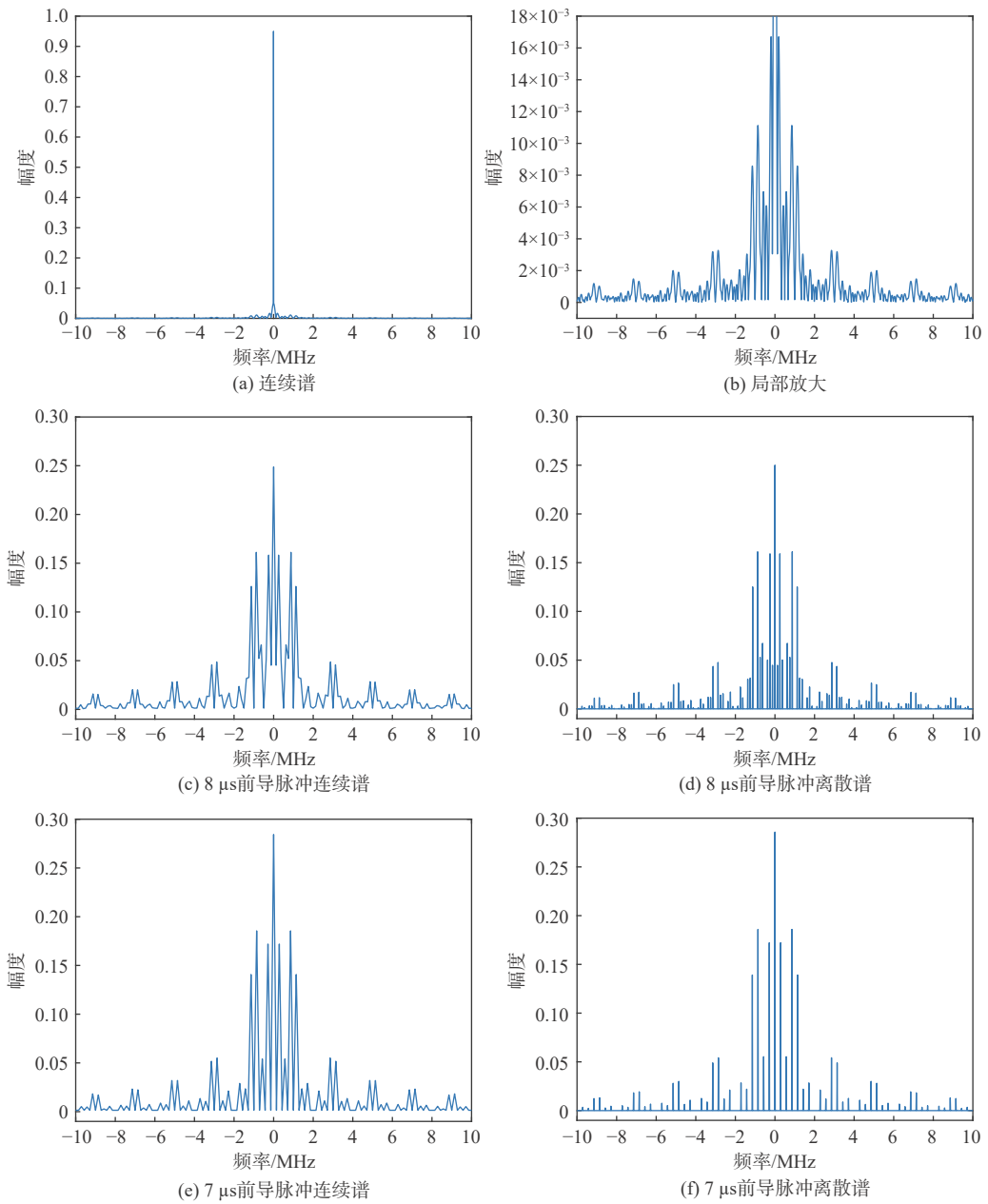


图 2　前导脉冲频谱比较

Fig. 2　Leading pulse spectrum comparison

析其简化 Volterra 级数，源于 S 模式应答信号的周期为 7 μs 的基带 4 脉冲序列简称准 S 模式信号，其数学模型$u_{s'}(t)$的可表示为

$$u_{s'}(t)=\sum_{q=1}^{14}u(q)p(t-qT) \tag{2}$$

式中：$p=\{1,0,1,0,0,0,0,1,0,1,0,0,0,0\}$。

1.2 核函数

设输入信号$u(t)$和输出信号$i(t)$的频域表示分别为$U(n)$和$I(m)$，其中，n和m分别为输入和输出信号频谱分量的谐波次数，则弱非线性系统广义频率响应函数为

$$\begin{cases}I(m)=\sum_{d=1}^{D}I^d(m)\\ I^d(m)=\sum_{-N\leqslant n_1,\cdots,n_d\leqslant N}H^d(n_1,\cdots,n_d)\prod_{p=1}^{d}U(n_p)\end{cases} \tag{3}$$

由式 (3) 第 2 个式子可得，输出信号$I(m)$的第m次谐波数是d阶子系统函数$I^d(m)$的输入信号谐波次数的总和，即$m=n_1+n_2+\cdots+n_d$。输入信号$U(n)$的d个频谱分量乘积与H^d加权得到$I^d(m)$，其d个频谱分量乘积被称为组合频率产物，加权系数H^d被称为核函数。

由式 (3) 可得，前D阶所需计算的核函数总数$C_{\text{tot,ind}}$为[16]

$$C_{\text{tot,ind}}=\sum_{d=1}^{D}\frac{(N+d-1)!}{(N-1)!d!} \tag{4}$$

由式 (4) 可见，信号谐波项数N越大，则弱非线性系统建模所需的阶数D值越高，待辨识核函数的个数将会巨大，导致建模维数灾难。

计算带宽分别为 10 MHz、20 MHz 时，S 模式信号和简化结构准 S 模式信号作为输入的传统 3、8 阶 Volterra 级数模型所需的核函数总数，如表 1 所示。可知，基于核函数约束的简化态准 S 模式信号比 S 模式信号输入的传统 Volterra 级数模型所需核函数的数量大幅度减少，因此，对复杂频谱输入信号简化结构得到频谱分量较少的准输入信号有利于频域 Volterra 级数建模。由表 1 可知，传统 Volterra 级数待辨识核函数依旧较多，需要依据输入信号的基频规律进一步简化 Volttera 级数，减少待辨识核函数。

1.3 简化 Volterra 级数

传统频域 Volterra 级数的核函数不仅有对称项的冗余，还有对输出特性表征较弱项，全部算出易造成维数灾难，故根据输入信号基频特征找到能差异化体现输出信号频谱包络的核函数，是进一步简化 Volterra 级数的方式。

表 1 传统 Volterra 级数模型所需核函数总数

Table 1 Total number of kernel functions required for traditional Volterra series model

建模参数		输入信号	核函数数目
阶数	带宽/MHz		
3	10	S模式信号	12 340
		准S模式信号	1 329
3	20	S模式信号	91 800
		准S模式信号	8 435
8	10	S模式信号	3.77×10^8
		准S模式信号	1.56×10^6
8	20	S模式信号	6.42×10^{10}
		准S模式信号	1.45×10^8

依据简化 Volterra 级数原理[16]，若对于准正弦信号的输入情况，准正弦信号的基波幅值最大，通过非线性系统后，含基波成分的组合频率分量对输出信号中各谐波分量的贡献较大，故可将式 (3) 第 2 个式子所示具有基波幅值最大特征的输入信号$U(n_p)$简化为通过弱非线性系统的各组合频率分量之积的形式：

$$\prod_{p=1}^{d}U(n_p)\approx U(1)^{i_p}U(-1)^{i_m}U(n) \tag{5}$$

式中：$U(1)$和$U(-1)$分别为基波项和共轭基波项；$U(n)$为通用谐波项；$i_p+i_m+1=d$，$i_p-i_m+n=m$。

弱非线性系统的核函数亦可简化为

$$H^d(n_1,\cdots,n_d)\approx H^d\left(\underbrace{1,\cdots,1}_{i_p},\underbrace{-1,\cdots,-1}_{i_m},n\right) \tag{6}$$

弱非线性系统广义频率响应函数可简化为

$$\begin{cases}I(m)=\sum_{d=1}^{D}I^d(m)\\ I^d(m)=\sum_{-N\leqslant n\leqslant N}H^d\left(\underbrace{1,\cdots,1}_{i_p},\underbrace{-1,\cdots,-1}_{i_m},n\right)\cdot\\ \qquad U(1)^{i_p}U(-1)^{i_m}U(n)\end{cases} \tag{7}$$

式 (7) 第 2 个式子即为准正弦输入的弱非线性系统的简化 Volterra 级数，其前D阶所需计算的核函数总数$C_{\text{tot,sin}}$为

$$C_{\text{tot,sin}}=\sum_{d=1}^{D}N\frac{(2+(d-1)-1)!}{(2-1)!(d-1)!}=N\sum_{d=1}^{D}d=N\frac{d(d+1)}{2} \tag{8}$$

若输入为准 S 模式前导脉冲，如图 2(f) 所示，周期为 7 μs 的准 S 模式前导脉冲的频谱包络呈现奇对称状态，且第 1 对频谱包络的幅值最大，对非线性失真 (如互调失真等) 的贡献最大[16]。故用周期为 7 μs 的准 S 模式信号第 1 对频谱 (n−1/7, n+1/7) MHz

与 (−2,2) MHz 内的频谱非线性组合形成的互调频率分量为

$$\prod_{p=1}^{d} U(n_p) \approx U(p_1)U(p_2)\cdots U(p_{d-1})U(x) \tag{9}$$

满足:

$$\begin{cases} n-1<(p_1+p_2+\cdots+p_{d-1}+n=m)<n+1 \\ -2<p_1,p_2,\cdots,p_{d-1}<2 \\ x=n\pm 1/7 \end{cases}$$

式 (9) 所示互调频率分量所对应核函数为

$$H^d(n_1,\cdots,n_d) \approx H^d(p_1,p_2,\cdots,p_{d-1},x) \tag{10}$$

于是, 基于准 S 模式信号输入的弱非线性系统简化 Volterra 级数为

$$\begin{cases} I(m)=\sum\limits_{d=1}^{D} I^d(m) \\ I^d(m)=\sum\limits_{-N\leqslant n\leqslant N} H^d(p_1,p_2,\cdots,p_{d-1},x)\cdot \\ \qquad U(p_1)U(p_2)\cdots U(p_{d-1})U(x) \end{cases} \tag{11}$$

满足:

$$\begin{cases} n-1<(p_1+p_2+\cdots+p_{d-1}+n=m)<n+1 \\ -2<p_1,p_2,\cdots,p_{d-1}<2 \\ x=n\pm 1/7 \end{cases}$$

如图 2(f) 所示, 在每个奇数频点附近有 3 根频谱的幅度较高, 即 n+1/7、n+3/7、n+5/7 处, 其对非线性失真贡献大, 故需要计算恢复。由式 (11) 第 2 个式子可得, 准 S 模式信号输入的简化 Volterra 级数前D阶所需要计算的核函数总数$C_{\text{tot,quasi}}$为

$$C_{\text{tot,quasi}}=\sum_{d=1}^{D}\frac{N(7+(d-1)-1)!}{7(7-1)!(d-1)!}=\sum_{d=1}^{D}\frac{N(5+d)!}{7!(d-1)!} \tag{12}$$

同样, 依据简化 Volterra 原理可以得到 S 模式信号输入的简化 Volterra 级数前D阶所需要计算的核函数总数$C_{\text{tot,S}}$为

$$C_{\text{tot,S}}=\sum_{d=1}^{D}\frac{N(16+(d-1)-1)!}{16(16-1)!(d-1)!}=\sum_{d=1}^{D}\frac{N(14+d)!}{16!(d-1)!} \tag{13}$$

计算带宽分别为 10 MHz、20 MHz 时, 不同输入信号类型下的简化 3 阶 Volterra 级数模型所需的核函数总数, 如表 2 所示, 其中, 删除对称核冗余。

表 2　简化 Volterra 级数模型所需核函数总数

Table 2　Total number of kernel functions required to simplify Volterra series model

建模参数		输入信号	核函数数目
阶数	带宽/MHz		
3	10	准正弦信号	56
		准S模式信号	213
		S模式信号	983
3	20	准正弦信号	116
		准S模式信号	498
		S模式信号	5 679

简化 Volterra 级数与传统 Volterra 级数模型相比较, 即表 2 与表 1 相比, 用已知输入信号频谱特性求取核函数实现系统非线性的简化 Volterra 级数表征, 可以取得维数和记忆深度的减少, 利于简化建模提升弱非线性系统引起信号失真导致的输出信号 TOA 分辨能力。此外, 基于准 S 模式信号输入的简化 Volterra 级数模型所需核函数的数量比基于 S 模式信号输入的核函数的数量大幅度减少, 特别是在宽带模型时核函数减少效果更加突出, 但比基于准正弦信号输入时的简化 Volterra 级数模型所需核函数的数量要多, 再次证明输入信号的简化态对于构建描述弱非线性系统特性的频域 Volterra 级数模型具有显著优势, 且输入信号的简化度越高, 构建级数模型所需核函数的数量就越少。以此简化级数就可以描述以准 S 模式信号为输入的二次雷达接收机的弱非线性记忆特征。

2　简化 Volterra 级数模型输出补偿

基于准 S 模式输入的弱非线性系统的简化 Volterra 级数模型构建过程如图 3 所示。如路线 1

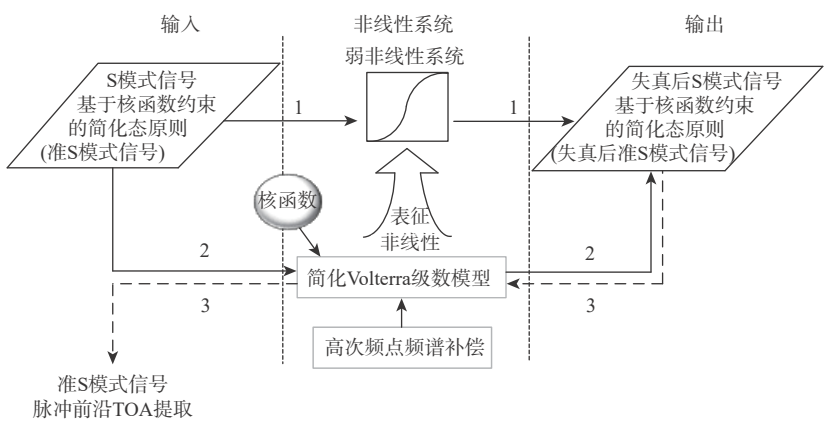


图 3　准 S 模式输入的弱非线性系统建模流程

Fig. 3　Modeling process for weakly nonlinear systems with quasi-S mode inputs

所示，原S模式信号基于核函数约束的简化态原则得到准S模式信号，信号输入弱非线性系统得到失真后波形。如路线2所示，向弱非线性系统输入准S模式信号先验数据$\boldsymbol{W}$得到输出$\boldsymbol{I}=\boldsymbol{W}^T\boldsymbol{H}$，通过$\boldsymbol{H}=(\boldsymbol{W}\boldsymbol{W}^T)^{-1}\boldsymbol{W}\boldsymbol{I}$计算核函数，根据式(11)第2个式子得到简化Volterra级数模型对此非线性系统进行表征。如路线3所示，已知失真输出信号，则可反演输出信号的频率组合分量及其对应的核函数，得到所对应输入准S模式信号再进行TOA提取，修正非线性干扰带来的误差。另外，高次频点频谱补偿是对简化Volterra级数模型精确度不足的频点的补偿。基于准正弦输入的弱非线性系统的简化Volterra级数模型构建过程同理。

2.1 简化Volterra级数模型误差分析

为讨论方便，将弱非线性系统等效为含非线性器件的串联电路，以便从输入输出信号的时频变化层面验证简化Volterra级数模型的精确性。设接收机弱非线性系统等效串联电路的输入电压与输出电流的关系函数为$F(Z)=F(V,J)$：

$$\begin{cases} J=0.139\,2\varphi-16.358\varphi^3+2\,260\varphi^5 \\ V=\dfrac{\mathrm{d}\varphi}{\mathrm{d}t}+6.962\,3\varphi-817.9\varphi^3+113\,000\varphi^5 \end{cases} \tag{14}$$

式中：φ为非线性电路中电感磁通量。

按照简化Volterra级数模型构建过程，计算核函数，得到输入为准正弦信号、表征此含参串联电路的3阶简化Volterra级数模型。其中，式(6)所示的部分3阶核函数$H(n_1,n_2,3)$数值图像如图4所示，核函数数值的显著程度与所对应输出频谱的包络相关。

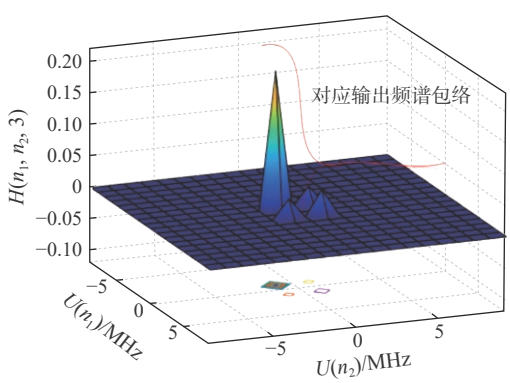


图4 准正弦信号输入时3阶核函数$H(n_1,n_2,3)$

Fig. 4 3rd order kernel function $H(n_1,n_2,3)$ for quasi-sin signal inputs

同样，按照简化Volterra级数模型构建过程，计算核函数，得到输入为准S模式信号、表征此含参串联电路的3阶简化Volterra级数模型。其中，式(10)所示的部分3阶核函数$H(p_1,p_2,20/7)$数值图像如图5所示，核函数数值的显著程度与所对应输出频谱的包络相关。

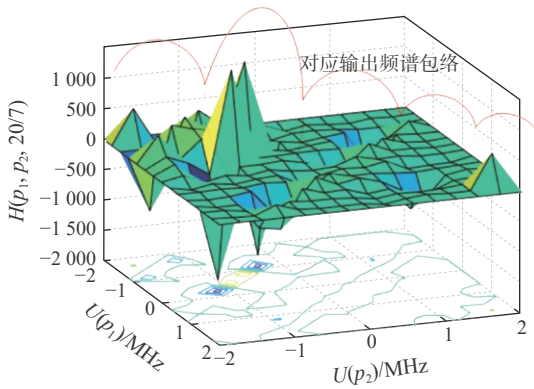


图5 准S模式信号输入时3阶核函数$H(p_1,p_2,20/7)$

Fig. 5 3rd order kernel function $H(p_1,p_2,20/7)$ for quasi-S-mode signal inputs

如图4、图5所示，根据输入信号频谱特征简化Volterra级数，求取的核函数数值与其所对应的输出频谱包络相关，得以在有限受约束核函数数量情况下，较为优越地表示出特定输入信号所对应非线性输出特征，精确构建起此串联电路简化Volterra级数非线性模型。

为验证简化Volterra级数模型性能，串联电路输入准正弦信号，其频谱前10 MHz内谐波谱如图6(a)所示，输出电流前10 MHz内谐波频谱如图6(b)所示；输入如式(2)所示准S模式信号，其频谱前10 MHz内谐波谱如图6(c)所示，输出电流前10 MHz内谐波频谱如图6(d)所示。

用所建立的3阶简化Volterra级数非线性模型表征准正弦信号、准S模式信号非线性传输过程，并评估Volterra核函数收益。设准输入信号简化Volterra级数模型频谱估计精确度为$1-E_s(n)$，所用核函数为N_s；原输入信号简化Volterra级数模型频谱估计精确度为$1-E_c(n)$，所用核函数为N_c，则准输入信号简化Volterra级数模型的核函数收益(以下简称为核收益)提升数量为

$$\mathrm{EARN}(n)=\frac{(1-E_s(n))/N_s}{(1-E_c(n))/N_c} \tag{15}$$

由式(15)可知，核收益主要与估计精确度和核函数约束个数密切相关，由表2可以看出，准S模式信号输入的N_s/N_c为4倍以上，大于准正弦输入的情况。得到基于输入信号的简化Volterra级数的3阶非线性模型的串联电路系统(简称为仿真系统)各频点的核收益如图7所示。

如图7所示，以准正弦信号输入仿真系统时，简化Volterra级数核函数收益是原输入信号简化Volterra级数的核函数收益约10倍以上，以准S模式信号输入仿真系统时，简化Volterra级数核函数

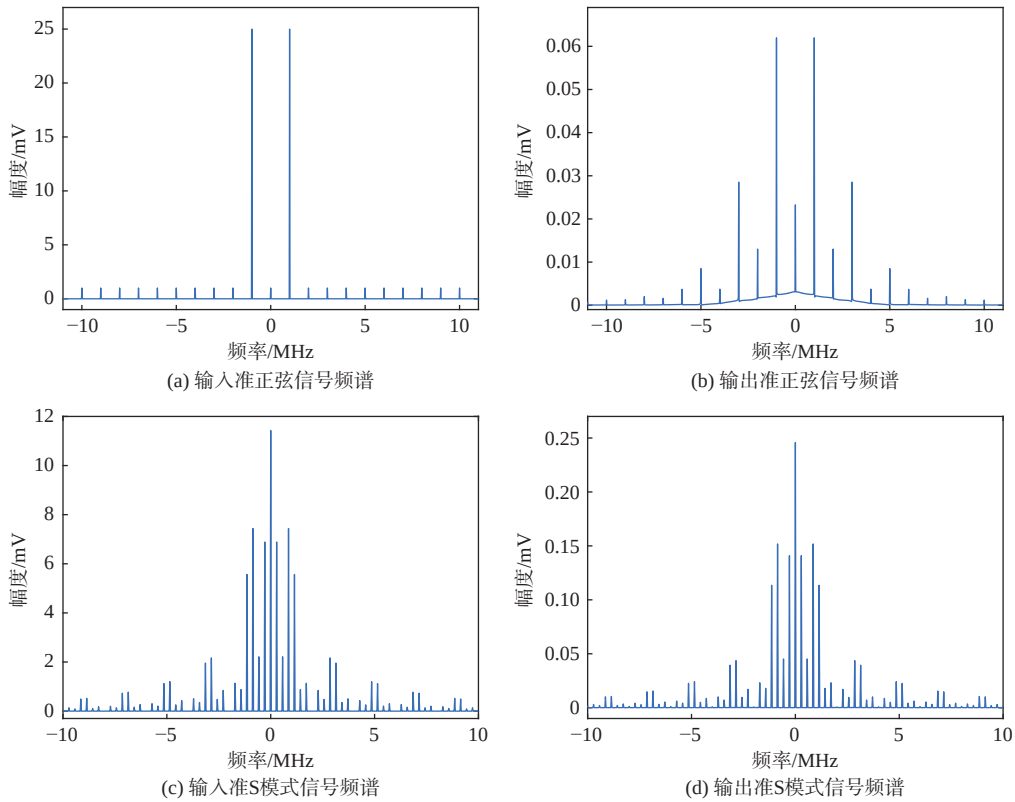


图 6　串联电路的输入与输出信号频谱

Fig. 6　Spectrum of input and output signals of series circuits

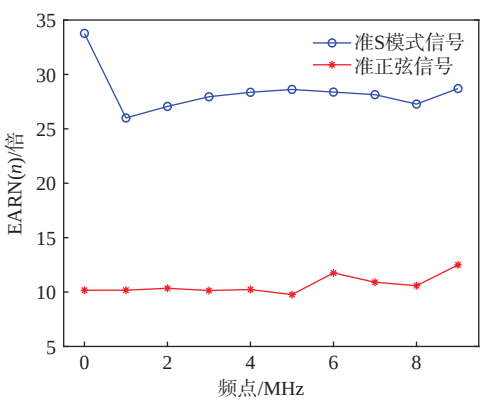


图 7　仿真系统输出信号的核收益

Fig. 7　Nuclear gain of output signal of simulation system

收益是原 S 模式信号简化 Volterra 级数的约 27 倍以上。比较可得, 越是复杂信号的简化态形式, 越有利于简化 Volterra 级数模型的核函数收益的提升。可见, 复杂输入信号的简化态(如准正弦信号/准 S 模式信号化等)引入对于弱非线性系统的简化 Volterra 级数模型构造具有重要理论和经济价值。

除核函数收益外, 模型的描述精确度也是评估模型的另外一个指标。设实际谱为 $U^{\mathrm{act}}(n)$, 估计谱为 $U^{\mathrm{est}}(n)$, 则 3 阶非线性模型描述输出信号频点幅度误差为

$$\mathrm{TVE}(n)=\frac{|U^{\mathrm{est}}(n)-U^{\mathrm{act}}(n)|}{|U^{\mathrm{act}}(n)|}\times 100\% \qquad (16)$$

得到基于输入信号的简化 Volterra 级数的 3 阶非线性模型的串联电路系统(简称仿真系统)各频点的频谱输出误差如图 8 所示。可知, 以准正弦信号输入仿真系统时, 简化 Volterra 级数能达到原输入信号简化 Volterra级数相近的估计精确度。以准 S 模式信号输入仿真系统时, 在频点小于 4 MHz 之前, 简化 Volterra级数对各频点频谱的估计精确度优于 S 模式信号简化 Volterra 级数; 在频点大于 4 MHz 之后, 简化 Volterra 级数对各频点频谱的估计精确度劣于 S 模式信号简化 Volterra 级数。结合图 7, 虽然基于准 S 模式信号的核收益较高, 但 3 阶

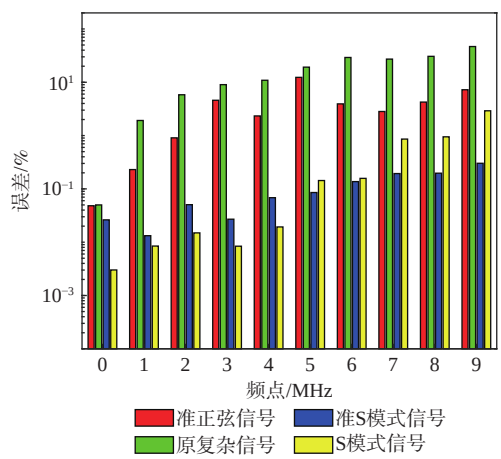


图 8　仿真系统各频点输出误差

Fig. 8　Output error at each frequency point of simulation system

简化Volterra级数对于高次频点误差的控制不如低次频点好。因此，增加少量对高次频点频谱描述影响较大的核函数，对高次频点频谱失真补偿是必要的。

2.2 高次频点频谱补偿方法

高次频点频谱的估计可用线性建模和谐波失真补偿的方法[17]，但该方法对非线性失真建模精确性不足。设实际对应高次频点频谱输出估计值$I_{\mathrm{K}}(m)$为

$$I_{\mathrm{K}}(m)=I_{\mathrm{volterra},m}+I_{\mathrm{compensate},m} \tag{17}$$

式中：$I_{\mathrm{compensate},m}$为第$m$次频点频谱补偿量。

$U^{l_+}(1)U^{l_-}(-1)$为输入信号的基波项产生的谐波失真，$K_r(l_+,l_-)$为补偿项的系数，l_+与l_-分别表示第m次频点的基波项和共轭基波项所出现的次数，r为补偿项的$U(1)$和$U(-1)$个数，则第m次频点频谱补偿量模型为

$$I_{\mathrm{compensate},m}=\sum_{r=2}^{R}\sum_{\substack{l_+-l_-=m\\ l_++l_-=r}}K_r(l_+,l_-)U^{l_+}(1)U^{l_-}(-1) \tag{18}$$

由于组合频率分量$U^l(1)U^l(-1)$是对直流项的贡献，输出频率为$m=r-2l$谐波的组合频率分量为$U^{r-2l}(1)U^l(1)U^l(-1)$，即$U^{r-2l}(1)(U(1))^{2l}$。则可将式(18)所示频谱补偿量模型改写为

$$I_{\mathrm{compensate},m}=\sum_{l=0}^{L(m)}B_m(l)U^m(1)(U(1))^{2l} \tag{19}$$

式中：$U^m(1)(U(1))^{2l}$为频率补偿项；$L(m)$表示对频率为m谐波的补偿项长度；$B_m(l)$为补偿项系数。

输入a组先验数据可得仿真系统输出信号的线性方程组，简化Volterra级数模型表示为$\boldsymbol{I}_{\mathrm{v,volterra},m}=\boldsymbol{W}_{\mathrm{v}}^{\mathrm{T}}\boldsymbol{H}_{\mathrm{v}}$，补偿项部分表示为$\boldsymbol{I}_{\mathrm{c,compensate},m}=\boldsymbol{C}_m^{(L)}\boldsymbol{P}_m^{(L)}$。其中，补偿项的系数向量$\boldsymbol{P}_m^{(L)}$和补偿项矩阵$\boldsymbol{C}_m^{(L)}$为

$$\begin{cases}\boldsymbol{P}_m^{(L)}=\begin{bmatrix}B_m(0)\\ \vdots\\ B_m(L)\end{bmatrix}\\ \boldsymbol{C}_m^{(L)}=\boldsymbol{E}_m\boldsymbol{\Phi}^{(L)}\end{cases} \tag{20}$$

式中：

$$\boldsymbol{E}_m=\mathrm{diag}\left(\left(U^{[1]}(1)\right)^{r-2l},\left(U^{[2]}(1)\right)^{r-2l},\cdots,\left(U^{[a]}(1)\right)^{r-2l}\right) \tag{21}$$

$$\boldsymbol{\Phi}^{(L)}=\begin{bmatrix}1 & |U^{[1]}(1)|^2 & \cdots & |U^{[1]}(1)|^{2(L-1)}\\ 1 & |U^{[2]}(1)|^2 & \cdots & |U^{[2]}(1)|^{2(L-1)}\\ \vdots & \vdots & & \vdots\\ 1 & |U^{[a]}(1)|^2 & \cdots & |U^{[a]}(1)|^{2(L-1)}\end{bmatrix} \tag{22}$$

补偿项是对应于系统衰减特征的高次谐波失真量，通过补偿项系数向量与补偿项矩阵将问题转化为参数估计问题，可在最小二乘意义下得到补偿项与实际误差项在复平面的最小均方根距离。对有误差波动的频点m直接更新$\boldsymbol{C}_m^{(L)}$，从而轻松找到对应于每个m次谐波误差的最佳补偿项长度L来满足简化Volterra级数模型精确度需求。对矩阵$\boldsymbol{\Phi}^{(t)}$求逆，即得简化Volterra级数的核函数和$B_m(t)$补偿项系数[17]。

当已知弱非线性系统输入信号的输出信号，由弱非线性系统的输出特性函数可推得输出信号的频率组合分量、补偿项及对应的核函数和补偿系数。于是，理论上可精确反演对应的输入信号。因此，依托核函数和补偿模型，在较低计算复杂度下可有效改善简化Volterra级数模型精确度。

2.3 高次频点频谱补偿方法仿真验证

如图9(a)所示，3阶准S模式信号简化Volterra级数模型所估计高次频点频谱项的误差较高，即32/7, 44/7, 46/7…等频点位置的误差较高。引入不同长度补偿项的频谱补偿量模型，高次频点频谱估计精确度可获得超过10^{-1}%以上幅度的提升，如图9(b)所示。增加频谱补偿项对于原来3阶简化Volterra级数精度提升的核收益量如图9(c)所示。核收益量中存在负值是由于补偿，有些频谱误差增大所导致。综合考虑核收益量和核函数增加带来的计算量增加，选取合适长度补偿项的频谱补偿量模型，且平均补偿项收益量是Volterra核收益量的1.45倍，得到简化Volterra级数模型的高次频点频谱补偿结果如图9(d)所示。若用线性建模和谐波失真补偿的方法对这些高次频点失真描述，得到结果如图9(e)所示，对比图9(b)可以看出精度不足。

简化Volterra级数结合高次频点频谱补偿组合模型、简化Volterra级数模型表征弱非线性系统传输特征的模型参数如表3所示。可见，简化Volterra级数结合高次频点频谱补偿组合模型相较于简化Volterra级数模型，在算法运算时间增加27%情况下，核函数收益增加45%，估计误差降低51.6%。动态自适应调整补偿项参数个数，可以提高其利用新数据估计的适应性和准确性。此外，在多点定位应用场景中，算法处理时间对于TOA估计的及时性与精确性影响重大。模型在简化准则下的计算复杂度和运算时间得到降低，将促进系统运行效率的提升。

综上，依托简化Volterra级数和高次频点频谱补偿组合模型表征弱非线性系统传输特征，当得到失真S模式信号后就可以精确恢复得到对应失真前波形，进而提取TOA。

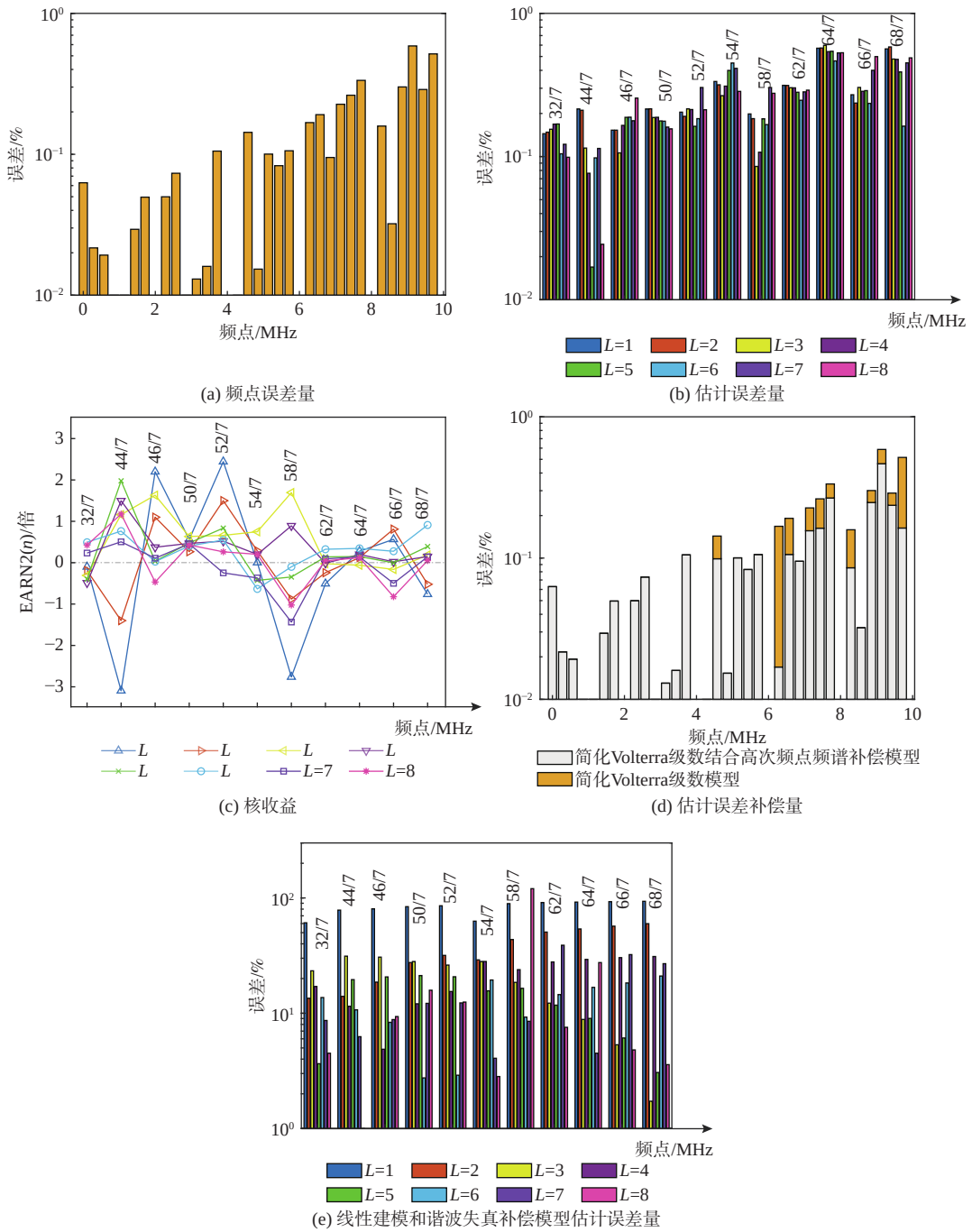


图 9　简化 Volterra 级数误差补偿项长度的影响

Fig. 9　Effect of length of simplities Volterra series error compensation term

表 3　模型参数

Table 3　Model parameters

模型	阶数	带宽/MHz	补偿项个数	样本数	运算时间/s	误差率/%
简化Volterra级数模型	3	10	0	300	0.343	0.31
简化Volterra级数结合高次频点频谱补偿组合模型	3	10	42	300	0.436	0.15

3　准 S 模式信号波形 TOA 提取

输入为准 S 模式信号, 非线性系统通过 Volterra 级数结合高次频点频谱补偿组合模型表征, 输出是经过组合模型非线性系统的准 S 模式信号。已知失真输出波形即可得到对应的失真前输入信号, 提取 TOA 并对比提取性能。此外, 仿真噪声和非线性参数对模型性能的影响。

3.1　TOA 提取

准 S 模式信号由 4 个脉冲组成, 只要准确得到这前 4 个脉冲即可实现 TOA 的提取。一个理想、上升前沿为 0.07 μs 的 S 模式信号输入弱非线性仿真系统, 依据图 1(c) 准 S 模式信号脉冲结构得到归一化输入与输出波形, 如图 10 所示。

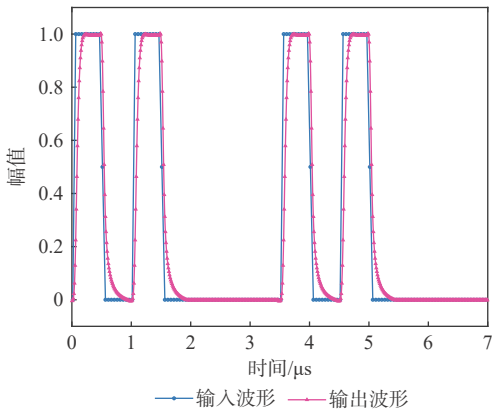


图 10 准 S 模式信号脉冲的输入与输出波形

Fig. 10 Input and output waveforms of quasi-S mode signal pulse

采用图 5、图 9(d) 所示 3 阶简化 Volterra 级数结合高次频点频谱补偿组合模型表征弱非线性系统传输特征，得到失真后波形所对应失真前 S 模式信号和准 S 模式信号脉冲的时域恢复波形的第 1 个脉冲上升前沿波峰部分，如图 11 所示。可知，通过 3 阶简化 Volterra 级数及高次频点频谱补偿组合模型，可准确描述非线性效应带来的波形失真。通过模型恢复输入波形上升前沿与真实失真前信号波形的上升前沿基本相当，且波形整体误差小于 0.15%。

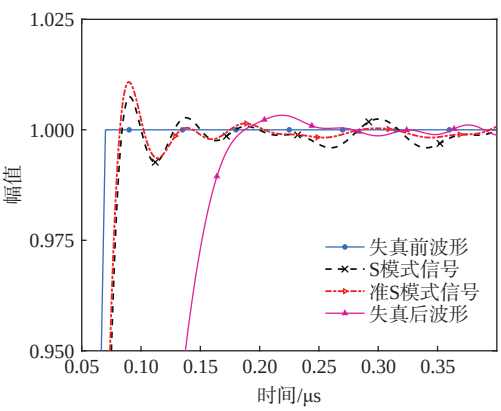


图 11 模型恢复波形与失真前输入信号波形

Fig. 11 Model recovery and pre-distortion input signal waveform

采用门限法提取仿真系统各个信号波形第 1 个脉冲的 TOA 参数。脉冲信号 TOA 提取误差与仿真系统的非线性程度有关，非线性程度越大，即脉冲上升沿的过渡时间越长，TOA 提取误差越大[18]。门限法提取信号 TOA 原理如图 12 所示，T_τ 越宽，即非线性程度越大，ΔT 也越大，即用失真信号算得 TOA 误差越大。

得到图 11 恢复前后波形的第 1 个前导脉冲在 100 MHz 采样率条件下估计的到达时间戳标记采样点时间误差 ΔT，如表 4 所示。由失真波形和恢复

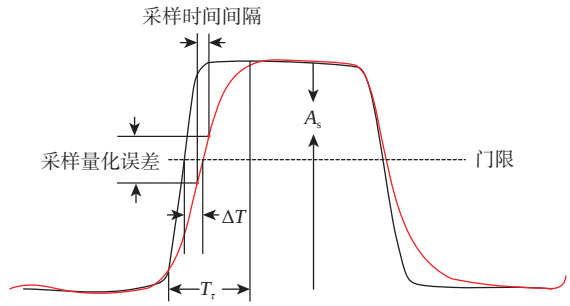


图 12 门限法提取信号 TOA 原理

Fig. 12 Principle of signal TOA extraction by threshold method

波形的 TOA 仿真实验结果(表 4)可见，简化 Volterra 级数模型的恢复波形修正了时间戳标记采样点的偏差，从而降低 TOA 估计误差。恢复准 S 模式信号略优于恢复 S 模式信号提取的 TOA，因此，基于准 S 模式信号的简化 Volterra 级数模型可用于 S 模式信号前导脉冲失真恢复及提升 TOA 提取精确度。

表 4 恢复前后波形 TOA 误差比较

Table 4 Comparison of TOA error in waveforms before and after restoration

波形	ΔT/ns
失真后波形	36
S模式信号	9
准S模式信号	8

3.2 噪声干扰

要想实际应用模型进行 TOA 准确估计，就必须考虑模型在电磁干扰和接收基站内部噪声干扰下的性能[18]。输入准 S 模式信号和 S 模式信号叠加 0～20 dB 噪声得到待验证信号，通过简化 Volterra 级数结合谐波补偿组合模型表征的仿真系统反演得到失真前输入波形。通过 1 000 次蒙特卡罗实验比较各波形相较于实际输入信号的波形误差，如图 13(a) 所示。通常，S 模式信号在机场附近或航路上的信噪比一般为 15 dB 以上，在 15 dB 信噪比时，1 000 次蒙特卡罗实验得到的波形误差散点图如图 13(b) 所示。

由图 13(a) 看出，信噪比大于 0 dB 时，恢复波形误差率均小于 13.67%，随信噪比的上升，波形恢复误差会降低，且整体恢复输入准 S 模式信号比 S 模式信号误差率低。由图 13(b) 看出，准 S 模式和 S 模式输入信号的模型平均误差率分别为 2.74% 和 2.24%，误差标准差分别为 0.74% 和 0.81%。可将图 13(b) 结果通过式 (11) 第 2 个式子表示为式 (23)，设噪声条件下误差 $\boldsymbol{\xi}$ 满足零均值、方差为 σ^2 的正态分布 $N(\boldsymbol{0},\sigma^2\boldsymbol{A})$，其中，$\boldsymbol{A}$ 为噪声的协方差矩阵。

$$\boldsymbol{I}_{\mathrm{n}}=\boldsymbol{W}_{\mathrm{n}}^{\mathrm{T}}\boldsymbol{H}_{\mathrm{n}}+\boldsymbol{\xi} \tag{23}$$

误差分布可反映模型稳定性，实质可以通过建

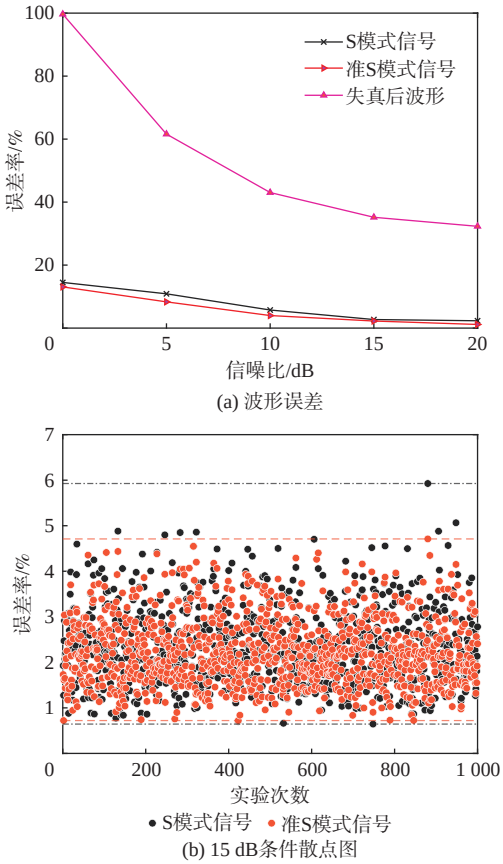


图 13　带噪声模型性能

Fig. 13　Model performance with noise

模所需 Volterra 核函数的稳定性分析得到:

$$\hat{\boldsymbol{H}}_{\mathrm{n}}=\left(\boldsymbol{W}_{\mathrm{n}}\boldsymbol{W}_{\mathrm{n}}^{\mathrm{T}}\right)^{-1}\boldsymbol{W}_{\mathrm{n}}\boldsymbol{I}_{\mathrm{n}} \tag{24}$$

噪声环境下核函数$\hat{\boldsymbol{H}}_{\mathrm{n}}$的方差-协方差矩阵为

$$\mathrm{Var}(\hat{\boldsymbol{H}}_{\mathrm{n}})=\mathrm{Var}\left[\left(\boldsymbol{W}_{\mathrm{n}}\boldsymbol{W}_{\mathrm{n}}^{\mathrm{T}}\right)^{-1}\boldsymbol{W}_{\mathrm{n}}\boldsymbol{I}_{\mathrm{n}}\right]=$$
$$\mathrm{Var}\left[\left(\boldsymbol{W}_{\mathrm{n}}\boldsymbol{W}_{\mathrm{n}}^{\mathrm{T}}\right)^{-1}\boldsymbol{W}_{\mathrm{n}}\left(\boldsymbol{W}_{\mathrm{n}}^{\mathrm{T}}\boldsymbol{H}_{\mathrm{n}}+\boldsymbol{\xi}\right)\right] \tag{25}$$

又 $\mathrm{Var}(\boldsymbol{H}_{\mathrm{n}})=0$(因为$\boldsymbol{H}_{\mathrm{n}}$不是随机变量), $\boldsymbol{W}_{\mathrm{n}}\boldsymbol{\xi}$的方差为$\sigma^2\boldsymbol{W}_{\mathrm{n}}\boldsymbol{W}_{\mathrm{n}}^{\mathrm{T}}$, 故

$$\mathrm{Var}(\hat{\boldsymbol{H}}_{\mathrm{n}})=\left(\boldsymbol{W}_{\mathrm{n}}\boldsymbol{W}_{\mathrm{n}}^{\mathrm{T}}\right)^{-1}\sigma^2\boldsymbol{W}_{\mathrm{n}}\boldsymbol{W}_{\mathrm{n}}^{\mathrm{T}}\left(\boldsymbol{W}_{\mathrm{n}}\boldsymbol{W}_{\mathrm{n}}^{\mathrm{T}}\right)^{-1}=\sigma^2\left(\boldsymbol{W}_{\mathrm{n}}\boldsymbol{W}_{\mathrm{n}}^{\mathrm{T}}\right)^{-1} \tag{26}$$

由式 (26) 可以看出, 模型的稳定性受建模先验数据矩阵和噪声共同影响。未经简化的 S 模式信号进行 Volterra 级数建模, 待估计参数的增加会导致$\boldsymbol{W}_{\mathrm{n}}\boldsymbol{W}_{\mathrm{n}}^{\mathrm{T}}$条件数(即最大奇异值与最小奇异值之比)增大, 这意味着系数估计的方差会增大, 即估计的不确定性增加, 系统稳定性降低。此外, 由于噪声的随机性, 可能增加模型发散的风险。但在 0 dB 以上信噪比应用环境, 噪声功率谱密度较低, 且系统具有足够的衰减特性, 因此, Volterra 级数依然收敛。

综上, 基于核函数约束的准 S 模式输入信号简化 Volterra 级数模型消除了内部大量的非线性耦合, 系数数量得到简化, 系统行为变得简单, 从而确定模型参数更加容易, 避免了非线性项造成 Volterra 核函数参数不佳, 从而在噪声干扰下有更好的精确性、稳定性和收敛性[19]。

3.3　非线性参数变化

用本文模型恢复失真前信号精度主要与接收机带宽和波形上升沿过渡时间有关, 即非线性模型参数对精确度有影响, 考虑如下: ①接收机接收信号的带宽影响恢复失真信号的可用频谱, 当运用不同宽度频谱表示不同阶及记忆长度的 Volterra 级数来表征弱非线性系统恢复失真前 S 模式信号波形, 波形精度会受影响。②根据规定, 基带 S 模式应答信号的脉冲前沿过渡时间在 0.05～0.1 μs 之间, 波形失真恢复精度会受到上升沿过渡时间影响。③多点定位算法的前导脉冲检测主要关心的是脉冲前沿和前导 4 脉冲位置。综上, 仿真分析接收机带宽和上升沿过渡时间对恢复 S 模式信号的前导脉冲前沿和整体波形的影响。仿真条件如下: 带宽由 5～30 MHz 变化, 上升沿过渡时间分别为 0 μs、0.05 μs、0.1 μs, 蒙特卡罗实验 1 000 次, 结果如图 14 所示。

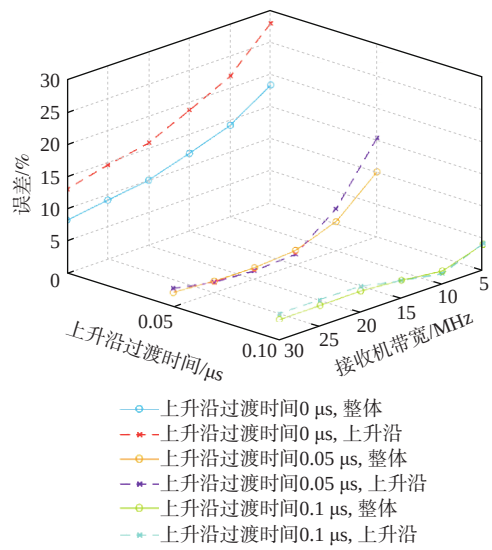


图 14　带宽和上升沿过渡时间对于恢复波形误差的影响

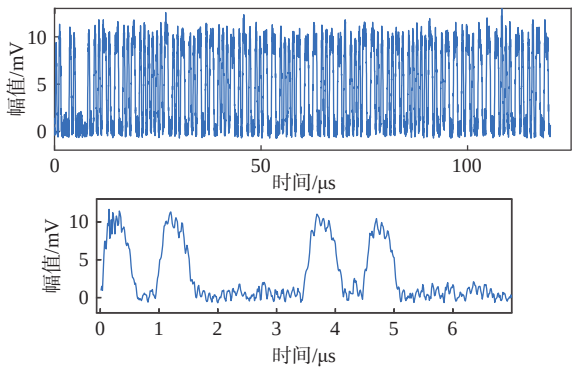
Fig. 14　Influence of bandwidth and rising edge transition time on waveform recovery error

由图 14 可见: ①接收机带宽的增加对于信号恢复误差降低有效, 但在达到一定水平之后, 恢复误差变化不大甚至反而略有增大。②信号的上升沿过渡时间越长, 信号越容易在较窄接收机带宽条件下达到较好的恢复精度。总体来看, 对于典型带宽接收机, 模型可以满足前导脉冲上升沿为 0.05～0.1 μs 的准 S 模式信号恢复误差小于 3.26%。

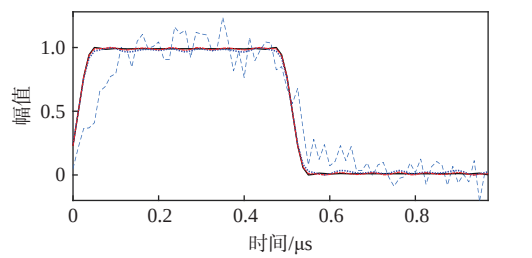
4　实测 S 模式信号验证分析

民航机场区域多点定位系统是通过测量机载

应答系统或 ADS-B 系统发射的 1 090 MHz 信号的到达时间差、主站汇算得到飞机位置坐标。到达时间差的精确度直接决定多点定位系统的定位精确度指标。目前，民航 1 090 MHz 射频信号以 S 模式形式最为常见。在天津滨海国际机场终端区设置一个信号采样频率为 80 MHz 的 ADS-B 接收机，实时接收的一段 S 模式信号基带数据及其前导脉冲基带结构如图 15(a) 所示，信噪比约为 13 dB。该信号包括 8 μs 的前导脉冲和 112 μs 的数据块，前导脉冲由 4 个分别位于 0 μs、1 μs、3.5 μs、4.5 μs 的子脉冲构成，脉冲宽度为 0.5 μs。接收机非线性传输链路导致前导脉冲的上升前沿宽度大于 0.1 μs，−3 dB 带宽大于 14 MHz，不能满足多点定位信号参数要求[18]。

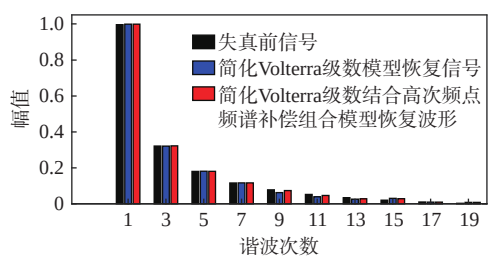


(a) S模式信号的前导脉冲基带结构



简化Volterra级数结合高次频点频谱补偿组合模型恢复波形
13 dB待验证信号
失真前波形
简化Volterra级数模型恢复波形

(b) 第1个前导脉冲恢复波形



(c) 各谐波的幅值

图 15 实测 S 模式基带信号波形及处理参数

Fig. 15 Measured S-mode baseband signal waveform and processing parameters

通常，机场区域多点定位系统的弱非线性模型容易获得，式 (14) 所示弱非线性模型不失一般性，因此，依旧以式 (14) 等效多点定位接收基站的弱非线性系统，构建的 3 阶简化 Volterra 级数模型和 3 阶简化 Volterra 级数及高次频点频谱补偿组合模型如图 5 和图 9(d) 所示。进而实测 S 模式基带信号对应输入的准 S 模式信号波形等得到恢复，其波形如图 15(b) 所示，恢复信号各频点归一化频率幅度如图 15(c) 所示。以式 (1) 所示 S 模式基带信号频谱为基准，分析失真 S 模式信号、简化 Volterra 级数模型恢复信号、简化 Volterra 级数及高次频点频谱补偿组合模型恢复信号等 3 种信号波形，可得其波形误差分别为 34.53%、2.23% 和 1.54%，且仅用约 8.8% 的准 S 模式信号模型的计算量。

可见，简化 Volterra 级数及高次频点频谱补偿组合模型所恢复信号波形的准确度最高。相比于简化 Volterra 级数模型，虽然计算量稍微增加，但收益量也相应增加，因此，对于不同失真恢复要求，综合考虑简化 Volterra 级数模型或组合模型使用。

失真 S 模式信号、简化 Volterra 级数模型恢复信号、简化 Volterra 级数及高次频点频谱补偿组合模型恢复信号相对于 10 MHz 带宽限制的失真前信号、无带宽限制失真前信号的波形用门限法提取 TOA，得到表 5。2 种模型恢复信号提取的 TOA 比失真 S 模式信号提取的 TOA 的精度提升均超过 73%。

表 5 模型恢复信号的 TOA 误差比较

Table 5 Comparison of TOA error in model recovery signals

模型	Δ*T*/ns	
	10 MHz带宽	无限制带宽
失真后信号	34	42
简化Volterra级数模型恢复信号	3	11
简化Volterra级数结合高次频点频谱补偿组合模型恢复信号	2	10

因此，基于准 S 模式信号的简化 Volterra 级数模型，以及简化 Volterra 级数结合高次频点频谱补偿组合模型均可应用于机场区域多点定位系统。

综合实测 S 模式信号与准 S 模式信号的对比仿真实验结果表明：

1）基于 S 模式信号频谱特征构建的准 S 模式信号，构建一种适用于多点定位系统的简化 Volterra 级数模型，能够有效降低失真 S 模式信号的处理复杂度。

2）采用简化 Volterra 级数模型恢复的实测 S 模式信号波形，可提升 S 模式信号受干扰导致前导脉冲失真时的 TOA 提取精确度。

3）利用简化 Volterra 级数模型恢复的 S 模式信号前导脉冲波形，在对脉冲上升前沿识别上具有一定的抗干扰能力。

5 结　论

弱非线性系统频域简化 Volterra 级数的阶数及核函数数量与输入信号复杂性密切相关, 选择一个与拟输入信号时频特征类似的准输入信号可显著减少频域简化 Volterra 级数核函数数量, 从而显著降低弱非线性系统简化 Volterra 级数复杂度, 提高对弱非线性系统的模拟仿真分析质量。简化准则不仅在多点定位 S 模式信号应用案例得到有效验证, 并且对于简化 Volterra 级数建模具有通用性, 对非线性传输建模效率和结果准确性都具有重要意义。

1) 理论分析与实测信号仿真实验表明, 构建的准 S 模式信号与简化 Volterra 级数模型可以满足多点定位等弱非线性系统应用需求。

2) 拟输入信号频域特征的简化, 可大幅度减少弱非线性系统频域 Volterra 级数核函数数量, 为频域 Volterra 级数的简化提供理论基础。

3) 拟输入信号频域特征的简化, 应以确保该输入信号包括第 1 对频谱在内的前 3 对频谱的恢复为基础, 确保拟输入信号的失真度较小。

4) 理论仿真与实际接收 S 模式信号的仿真结果表明, 基于输入信号简化的简化 Volterra 级数的核函数数量大幅度减少, 且对应的简化 Volterra 级数的计算量显著下降, 输出信号波形的恢复误差等指标达到要求。

因此, 拟输入信号频域特征的简化对降低频域 Volterra 级数的复杂度与计算量具有重要理论与应用价值。

参考文献 (References)

[1] LE NEINDRE F, FERRE G, DALLET D, et al. A successive interference cancellation-based receiver for secondary surveillance radar [J]. IEEE Transactions on Aerospace and Electronic Systems, 2023, 59(2): 805-816.

[2] 宫峰勋, 曹雅茹. S 模式前导脉冲过零点提取与 TOA 时间戳精确度研究[J]. 南京航空航天大学学报, 2020, 52(5): 701-707.
GONG F X, CAO Y R. Zero-crossing extraction of S-mode preamble pulse and TOA timestamp accuracy[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2020, 52(5): 701-707(in Chinese).

[3] 李松亭, 颜盾. 射频集成电路校准技术综述[J]. 电子与信息学报, 2022, 44(11): 4058-4074.
LI S T, YAN D. An overview on calibration techniques for radio frequency integrated circuits[J]. Journal of Electronics & Information Technology, 2022, 44(11): 4058-4074(in Chinese).

[4] 黄家露, 王文涛, 周莲, 等. 基于 LS-SVM 的宽带接收前端非线性补偿算法[J]. 电子学报, 2023, 51(6): 1500-1509.
HUANG J L, WANG W T, ZHOU L, et al. Nonlinearity mitigation method based on LS-SVM for wide-band receiver[J]. Acta Electronica Sinica, 2023, 51(6): 1500-1509(in Chinese).

[5] LENHARD K, BAUMGARTNER A, SCHWARZMAIER T. Independent laboratory characterization of NEO HySpex imaging spectrometers VNIR-1600 and SWIR-320m-e[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(4): 1828-1841.

[6] VANSEBROUCK R, JABBOUR C, JAMIN O, et al. Fully-digital blind compensation of non-linear distortions in wideband receivers [J]. IEEE Transactions on Circuits and Systems I: Regular Papers, 2017, 64(8): 2112-2123.

[7] MAITI D, BRANDT-PEARCE M. Modified nonlinear decision feedback equalizer for long-haul fiber-optic communications[J]. Journal of Lightwave Technology, 2015, 33(18): 3763-3772.

[8] 曹韬, 刘友江, 杨春, 等. 高效宽带包络跟踪系统电路性能优化及非线性行为校正[J]. 电子与信息学报, 2020, 42(3): 787-794.
CAO T, LIU Y J, YANG C, et al. Circuits optimization and system linearization for high efficiency and wideband envelope tracking architecture[J]. Journal of Electronics & Information Technology, 2020, 42(3): 787-794(in Chinese).

[9] OBAIS A M. Linearizing of low noise power amplifier using 5.8 GHz double loop feedforward linearization technique[J]. Journal of University of Babylon for Engineering Sciences, 2017, 2(25): 786-791.

[10] DJENNOUNE S, BETTAYEB M, AL-SAGGAF U M. Modulating function-based fast convergent observer and output feedback control for a class of non-linear systems[J]. IET Control Theory & Applications, 2019, 13(16): 2681-2693.

[11] ZHANG S, ZHOU Z C, ZHENG W X, et al. Frequency-domain Volterra kernel-based adaptation: formulations and algorithms[J]. Signal Processing, 2024, 216: 109295.

[12] CHIERCHIE F, AASE S O. Volterra models for digital PWM and their inverses[J]. IEEE Transactions on Circuits and Systems I: Regular Papers, 2015, 62(10): 2606-2616.

[13] CENTURELLI F, MONSURRÒ P, SCOTTI G, et al. Methods for model complexity reduction for the nonlinear calibration of amplifiers using Volterra kernels[J]. Electronics, 2022, 11(19): 3067.

[14] POURDARVISH A, SAYEVAND K, MASTI I, et al. Orthonormal Bernoulli polynomials for solving a class of two dimensional stochastic Volterra-Fredholm integral equations[J]. International Journal of Applied and Computational Mathematics, 2022, 8(1): 31.

[15] 王云海, 韩景龙, 张兵, 等. 空气动力二阶核函数辨识方法[J]. 航空学报, 2014, 35(11): 2949-2957.
WANG Y H, HAN J L, ZHANG B, et al. Identification method of second-order kernels in aerodynamics[J]. Acta Aeronautica et Astronautica Sinica, 2014, 35(11): 2949-2957(in Chinese).

[16] FAIFER M, LAURANO C, OTTOBONI R, et al. Definition of simplified frequency-domain Volterra models with quasi-sinusoidal input[J]. IEEE Transactions on Circuits and Systems I: Regular Papers, 2018, 65(5): 1652-1663.

[17] FAIFER M, LAURANO C, OTTOBONI R, et al. Compensating the harmonic distortion introduced by instrument transformers: an improved method based on frequency-domain polynomials[C]//Proceedings of the IEEE 11th International Workshop on Applied Measurements for Power Systems. Piscataway: IEEE Press, 2021: 1-6.

[18] 宫峰勋, 第五瑶光. 低信噪比 S 模式基带信号到达时间估计联合

算法[J]. 北京航空航天大学学报, 2025, 51(2): 380-388.
GONG F X, DIWU Y G. Joint algorithm for time of arrival estimation of S-mode baseband signals with low SNR[J]. Journal of Beijing University of Aeronautics and Astronautics, 2025, 51(2): 380-388(in Chinese).

[19] 张家树, 肖先赐. 混沌时间序列的自适应高阶非线性滤波预测[J]. 物理学报, 2000, 49(7): 1221-1227.
ZHANG J S, XIAO X C. Prediction of chaotic time series by using adaptive higher-order nonlinear Fourier infrared filter[J]. Acta Physica Sinica, 2000, 49(7): 1221-1227(in Chinese).

Simplified S-mode signal pulse structure with kernel function constraints and TOA accuracy study

GONG Fengxun*, LIU Tao

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

Abstract: The very accurate time-of-arrival (TOA) extraction is significantly impacted by the complicated space-time structure of the civil aviation S-mode transponder signal, which is also vulnerable to neighboring electromagnetic interference and transmission link nonlinearities. Based on the air-time structure and spectral analysis of the civil aviation S-mode transponder signal, a quasi-S-mode signal definition model is proposed to simplify the pulse structure of the S-mode signal, and a simplified Volterra level model is constructed based on the quasi-S-mode signal, the kernel function of the inversion constraints of the weak nonlinear system characteristics. Simulation results show that the kernel function gain multiplier of the simplified Volterra level model based on the quasi-S-mode signal is not less than 27 times. Finally, the quasi-S-mode recovery of an interfered S-mode signal with a measured signal-to-noise ratio of about 13 dB is achieved by using a combination of the 3rd-order simplified Volterra level model and spectral compensation at high sub-frequency points, and the waveform recovery error of about 1.54% is obtained with the computation of about 8.8% of the standard S-mode signal model. Additionally, the accuracy of the recovered signal is improved by more than 73% over the TOA extracted from the disturbed signal. Consequently, the reduced Volterra level model of quasi-S-mode signal based on kernel function constraints has significant theoretical significance for improving localization accuracy and accurately estimating the arrival time of S-mode signals in civil aviation.

Keywords: kernel function; quasi S-mode signal; simplified Volterra series; spectrum compensation; time-of-arrival; multilateration

Received: 2024-04-11; **Accepted**: 2024-07-26; **Published Online**: 2024-09-09 16:39
URL: link.cnki.net/urlid/11.2625.V.20240909.1304.004
Foundation item: National Key Research and Development Program of China (2018YFC0809500)
* **Corresponding author**. E-mail: fxgong@cauc.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0291

基于改进滑模观测器的电磁直线执行器无位移传感器控制

葛文庆[1]，李德同[1,2]，宋亚东[1]，谭草[1]，李波[1,*]

(1. 山东理工大学 交通与车辆工程学院，淄博 255000；2. 潍柴动力股份有限公司，潍坊 261061)

摘 要：针对电磁直线执行器使用位移传感器带来的安装困难、成本增加和稳定性降低等问题，提出一种基于BP神经网络的改进超螺旋滑模观测器的位移估计方法，结合自适应积分鲁棒控制算法实现了电磁直线执行器无位移传感器控制。为削弱超螺旋滑模观测器的抖振现象，改善位移估计性能，在观测器结构方面，设计一种非奇异全局快速终端滑模面，并以连续的双曲正切函数作为切换函数；在观测器参数调整方面，设计BP神经网络，以动子速度为输入动态调整超螺旋滑模观测器增益。搭建电磁直线执行器运动控制性能测试平台，对位移估计与反馈控制结果进行分析，结果表明：相比于超螺旋滑模观测器，改进的滑模观测器在阶跃工况下的位移估计最大误差减小了16.22%，频率为2 Hz的正弦工况下的位移估计最大误差减小了9.10%；无位移传感器控制与有位移传感器控制的控制性能相当，两者在8 mm阶跃工况下的稳态误差约为0.03 mm，频率为2 Hz的正弦工况下，估计位移与实际位移的最大误差为0.43 mm。证明了基于改进超螺旋滑模观测器的电磁直线执行器无位移传感器控制的有效性与实用性。

关 键 词：电磁直线执行器；超螺旋滑模观测器；神经网络；双曲正切函数；无位移传感器控制

中图分类号：V221+.3；TB553

文献标志码：A **文章编号**：1001-5965(2026)06-1803-07

电磁直线执行器相比于旋转电机加运动转换装置的方案，在质量、复杂度与响应速度等方面有着明显的优势，广泛应用于汽车、航空航天及现代机床等领域[1-4]。然而，位移传感器的使用将带来安装困难、成本增加和可靠性降低等问题，因此，无位移传感器技术成为了研究热点[5]。

电磁直线执行器的位移估计方法主要包括直接计算法、参数关系拟合法、状态观测器法等[6-7]。直接计算法直接通过执行器动态方程推导出位移信号，在动态方程建立过程中，通常将机电常数等关键参数视为常数，精度往往得不到保证[8]。参数关系拟合法是通过拟合出位移信号与执行器其他参数之间的函数关系，以此间接获得执行器位移信号的方法，适用于电磁场变化剧烈的执行器，但难以适应复杂的工作环境[9]。状态观测器法根据测量得到的系统输入(控制量)和系统输出(部分状态变量或状态变量的函数)来确定系统内部所有状态信息，基于状态观测器法的电磁直线执行器位移估计方法有着较好的发展前景[10]。

全阶及降阶观测器虽能实现对动子位移的估计，但对系统的精确数学模型具有较大依赖且设计流程复杂[11]。卡尔曼滤波作为一种特殊的状态

收稿日期：2024-05-08；录用日期：2024-05-24；网络出版时间：2024-06-13 11:29
网络出版地址：link.cnki.net/urlid/11.2625.V.20240612.1716.002
基金项目：国家自然科学基金(52375105)；山东省自然科学基金优秀青年基金项目(ZR2022YQ51)；山东省自然科学基金(ZR2023ME178, ZR2023ME177)；山东省高等学校“青创科技计划”团队项目(2022KJ232)

* 通信作者. E-mail：njustlibo@126.com

引用格式：葛文庆，李德同，宋亚东，等. 基于改进滑模观测器的电磁直线执行器无位移传感器控制[J]. 北京航空航天大学学报，2026，52(6)：1803-1809. GE W Q，LI D T，SONG Y D，et al. Displacement sensorless control of electromagnetic linear actuator based on an improved sliding mode observer[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1803-1809 (in Chinese).

观测器，是一种最优的滤波器，能够有效消除噪声和抖动，但其对非线性系统的估计效果不佳[12-13]。滑模观测器由于其结构简单、鲁棒性强等优点，在位移估计领域具有较好的应用前景[14]。然而，滑模观测器在电磁直线执行器位移估计过程中存在剧烈抖振现象，严重影响位移估计精度，尚需进一步研究。

为削弱滑模观测器的抖振现象，合理调整参数是一种有效的方法。一般情况下，滑模增益参数是不变的，但面对复杂工况的控制对象，固定的滑模增益会严重影响控制效果[15]。人工智能方法由于其自学习、自组织和自适应的特性，能够在一定程度上模仿人脑的思维过程，因此，在参数调整方面得到广泛应用。神经网络具有强大的非线性逼近特性，其在控制参数调整领域具有较好的应用潜力[16]。

本文提出一种基于BP神经网络的改进超螺旋滑模观测器（back propagation-improved super-twisting sliding mode observer, BP-ISTSMO）位移估计方法，结合自适应积分鲁棒控制算法实现了电磁直线执行器无位移传感器控制，搭建电磁直线执行器运动控制性能测试平台，对位移估计与反馈控制结果进行分析，证明所提方法的有效性与实用性。

1 电磁直线执行器原理与模型

本文以动圈式电磁直线执行器为研究对象，主要由永磁体、线圈绕组、内外磁轭和线圈骨架等结构构成，永磁体和内外磁轭为定子，线圈绕组绕制于线圈骨架上为动子，其结构如图1所示，具体原理详见参考文献[17]。

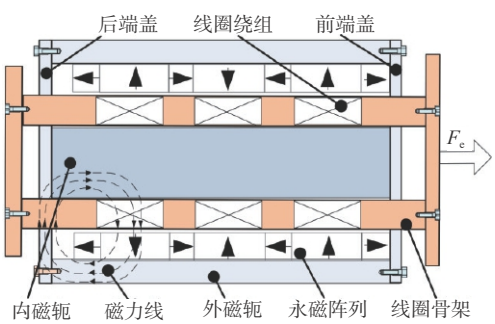


图1 电磁直线执行器结构

Fig. 1 Structure of electromagnetic linear actuator

电磁直线执行器作为一种能将电能转化为机械能的能量转换装置，是机械、电路和磁场相互耦合的复杂系统，因此，在建模时需要分别对机械子系统、磁路子系统和电路子系统建立数学模型，并充分考虑3个子系统之间的耦合关系，其状态空间方程为

$$\begin{cases} \dfrac{\mathrm{d}i(t)}{\mathrm{d}t} = -\dfrac{R}{L}i - \dfrac{K_e}{L}v + \dfrac{U}{L} \\ \dot{v} = \dfrac{-F_f - cv}{M} + \dfrac{K_m}{M}i \\ \dot{x} = v \end{cases} \tag{1}$$

式中：U为电源电压；R为整个线圈的电阻；L为整个线圈产生的电感；i为通过线圈的电流；v为线圈动子的运动速度；K_e为反电动势系数；K_m为电磁力常数；M为电机动子的质量；x为动子的位移；F_f为摩擦力；c为阻尼系数。

2 位移估计方法设计

2.1 超螺旋滑模观测器设计

超螺旋滑模观测器（super-twisting sliding mode observer, STSMO）不仅继承了传统滑模算法简便、鲁棒性强的特点，而且可以明显削弱抖振现象，提高控制精度。典型的超螺旋算法如下：

$$\begin{cases} \dot{s} = -k_1|s|^{1/2}\mathrm{sign}(s) + V \\ \dot{V} = -k_2\mathrm{sign}(s) \end{cases} \tag{2}$$

式中：s为滑模变量；k_1、k_2为滑模增益。

电磁直线执行器的电流状态方程为

$$\frac{\mathrm{d}i}{\mathrm{d}t} = -\frac{R}{L}i + \frac{U}{L} - \frac{E}{L} \tag{3}$$

所设计的超螺旋滑模观测器为

$$\frac{\mathrm{d}\hat{i}}{\mathrm{d}t} = -\frac{R}{L}\hat{i} + \frac{U}{L} - \frac{Z_c}{L} \tag{4}$$

式中：$\hat{i}$为i的估计值。

电流的观测误差为

$$\tilde{i} = i - \hat{i} \tag{5}$$

选择电流误差作为线性滑模面$s = \tilde{i} = i - \hat{i}$。超螺旋滑模控制率可设计为

$$Z_c = -k_1|\tilde{i}|^{1/2}\mathrm{sign}(\tilde{i}) - k_2\int \mathrm{sign}(\tilde{i})\mathrm{d}t \tag{6}$$

当超螺旋滑模观测器系统处于稳定状态时，$\tilde{i} = 0$，反电动势观测值近似等于实际值，此时的超螺旋滑模正在滑模平面上运行。超螺旋滑模观测器观测的电磁直线执行器反电动势为

$$E = Z_c = -k_1|\tilde{i}|^{1/2}\mathrm{sign}(\tilde{i}) - k_2\int \mathrm{sign}(\tilde{i})\mathrm{d}t \tag{7}$$

为进一步提高超螺旋滑模观测器的观测精度，消除等效反电动势中由于开关函数带来的高频干扰，需要在观测器后加一个低通滤波器（low-pass filter, LPF），消除高次谐波，其位移观测过程如图2所示。低通滤波器为

$$\hat{E} = \frac{\omega_c}{s + \omega_c}Z_c \tag{8}$$

式中：ω_c为截止角频率。

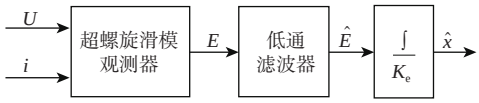


图 2　超螺旋滑模观测器的位移观测原理

Fig. 2　Displacement observation principle of super-twisting sliding mode observer

由电磁直线执行器的反电动势估计值得出位移估计值为

$$\hat{x}=\int\frac{\hat{E}}{K_{\mathrm{m}}}\mathrm{d}t \tag{9}$$

2.2　改进超螺旋滑模观测器设计

为削弱超螺旋滑模观测器的抖振, 改善位移估计性能, 从观测器滑模面和切换函数 2 个方面提出改进措施。

2.2.1　非奇异全局快速终端滑模面

为解决传统滑模面存在的各类问题, 结合各滑模面的优势, 本文提出一种非奇异全局快速终端滑模面。定义电流跟踪误差为式 (5)。

非奇异全局快速终端滑模面为

$$s=\tilde{i}+\alpha\tilde{i}^{p/q}+\beta\dot{\tilde{i}}^{p/q} \tag{10}$$

式中: $1<p/q<2$, 且 p、q 为正奇数; α、β 大于 0。

由于式 (10) 中存在非线性部分 $\tilde{i}^{p/q}$、$\dot{\tilde{i}}^{p/q}$, 当系统运动点离平衡状态较远时, 系统速度收敛主要由非线性部分决定, 且距离越远, 收敛速度越快; 而当系统状态靠近平衡状态时, 线性部分 $\tilde{i}$ 决定了系统的收敛速度, 此时其收敛速度与线性滑模控制的收敛速度类似。这种滑模面的设计可以使得系统全局快速收敛到平衡状态。由于 $(p-q)-1>0$, 从而解决了快速终端滑模控制存在的奇异性问题。

2.2.2　切换函数的设计

设计双曲正切函数代替不连续的符号函数, 可进一步降低超螺旋滑模控制中的抖振现象。双曲正切函数具有连续光滑的特性, 在需要对切换函数求导时, 还可以大幅降低系统惯性的影响。双曲正切函数如图 3 所示。

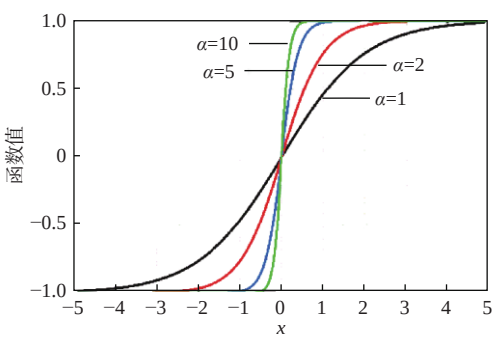


图 3　双曲正切函数的图像

Fig. 3　Image of hyperbolic tangent function

双曲正切函数的具体表达式为

$$\tanh(x)=\frac{\mathrm{e}^{\alpha x}-\mathrm{e}^{-\alpha x}}{\mathrm{e}^{\alpha x}+\mathrm{e}^{-\alpha x}} \tag{11}$$

综上, 将式 (6) 的线性滑模面改为非奇异快速终端滑模面, 将切换函数改为双曲正切函数, 改进超螺旋滑模观测器的公式为

$$Z_{\mathrm{c}}=-k_1|s|^{1/2}\mathrm{sign}(s)-k_2\int\mathrm{sign}(s)\mathrm{d}t \tag{12}$$

2.3　基于 BP 神经网络的参数调整

超螺旋滑模观测器的性能与 2 个增益 k_1、k_2 的数值密切相关。通常情况下, 2 个增益是定值, 在较小的速度范围内, 固定滑模增益可以有效抑制滑模抖振, 实现动子位移的精确估计。在速度较大时, 为有效抑制抖振, 滑模增益一般设定较大; 但动子速度下降时, 较大的滑模增益将会加剧系统的抖振。同样, 在速度较大时, 如果滑模增益较小, 也会削弱系统的稳定性。因此, 适当调整滑模增益对抑制滑模抖振和提高系统的动态响应具有重要意义。本文以 BP-ISTSMO 的 2 个增益进一步减小超螺旋滑模观测器的抖振现象。

本文设计的 BP 神经网络为输入层、隐藏层和输出层的 3 层网络, 基本结构如图 4 所示。以 Sigmoid 函数作为激活函数, 以电磁直线执行器的速度作为输入, 超螺旋滑模观测器的 2 个增益作为输出, 即输入层节点为 1, 输出层节点为 2[18]。在实际应用中, 隐藏层的节点数先默认一个数值, 根据训练完成后网络的拟合优度再进行更改, 直至出现满意的网络精度, 最终确认隐藏层节点数为 15。将得到的 500 组数据导入 MATLAB 神经网络工具箱中进行训练, 70% 的数据用于训练, 15% 的数据用于验证, 15% 的数据用于预测。

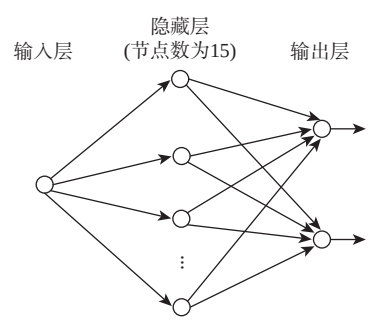


图 4　BP 神经网络的基本结构

Fig. 4　Basic structure of BP neural network

数据训练结果如图 5 所示。神经网络迭代 660 次, 均方误差趋于零, 说明拟合的神经网络具有较高的线性度。以动子速度为输入、超螺旋滑模观测器的增益值为输出的 BP 神经网络具有有效性, 能完成超螺旋滑模观测器增益适当调整的功能。

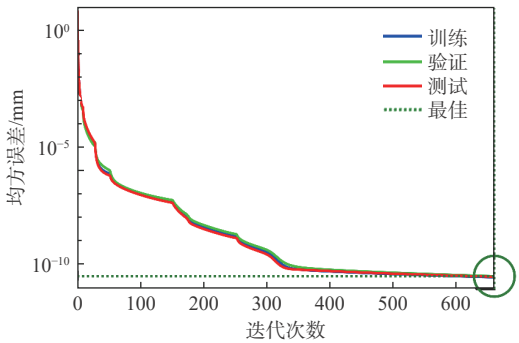


图 5 神经网络数据训练结果

Fig. 5 Data training results of neural network

3 结果与分析

3.1 实验平台

为验证电磁直线执行器无位移传感器控制的有效性，搭建了实验平台，如图 6 所示，其主要包括电磁直线执行器、上位机及控制器与驱动器 3 部分。该实验平台通过实时数字控制系统 RTU-BOX 实现控制算法与位移估计算法，其主处理器的主频为 300 MHz；驱动器采用双 H 桥的直驱电机驱动模块，该驱动模块可以用于一般大功率直流电机的驱动，其最大电流可达 100 A，性能稳定，具有比一般驱动器更为优越的性能。在控制过程中，电压、电流和动子位移等关键参数均通过传感器采集。执行器主要参数如表 1 所示，驱动器主要参数如表 2 所示。

针对电磁直线执行器系统内部参数扰动、外部不确定性干扰及观测器估计位移带来的误差对控制系统的影响，提出一种自适应积分鲁棒控制算

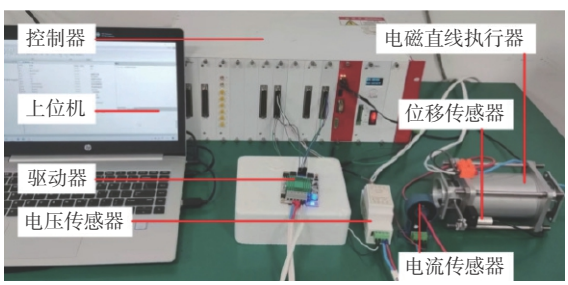


图 6 实验平台

Fig. 6 Experimental platform

表 1 电磁直线执行器参数值

Table 1 Electromagnetic linear actuator parameter values

参数	数值
执行器工作行程/mm	10
动子质量/kg	0.13
线圈电感/mH	1.3
线圈电阻/Ω	1.1
电磁力常数/(N·A^{-1})	25.83

表 2 驱动器参数值

Table 2 Driver parameter values

参数	数值
工作电压/V	15～30
最大额定电流/A	100
外观尺寸/(mm×mm×mm)	70×56×18
导通电阻/Ω	0.003
开关频率/kHz	60

法，提高了控制系统的稳定性和抗干扰能力，控制器的具体设计详见文献 [19]，其控制参数选取为 k=500、k_r=4、μ=10。基于改进滑模观测器实现了电磁直线执行器无位移传感器控制，控制系统结构如图 7 所示。改进超螺旋滑模观测器以电压、电流信号为输入，输出位移估计信号，以 BP 神经网络在动子速度的输入下，适当调整改进超螺旋滑模观测器的 2 个增益，以达到减小滑模抖振的效果；以自适应积分鲁棒控制作为控制器完成反馈控制。其中，k_1、k_2 为超螺旋滑模观测器的 2 个增益，其初始值为 k_1=18.5、k_2=15，U、i 为电压、电流信号，v 为执行器动子速度，$\hat{x}$为位移的估计值，x^*为唯一的期望值，u 为生成的控制量。

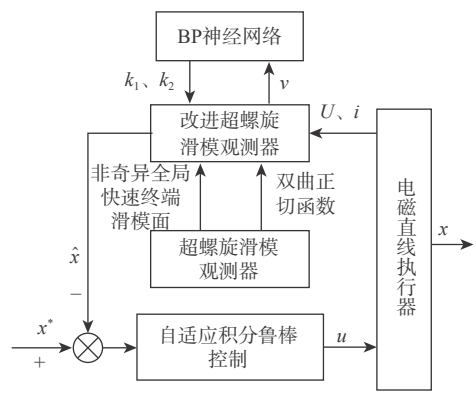


图 7 电磁直线执行器无位移传感器控制系统框图

Fig. 7 Block diagram of electromagnetic linear actuator control system of displacement sensorless

3.2 位移估计性能分析

3.2.1 阶跃实验验证

为验证 BP-ISTSMO 在阶跃工况下的位移估计性能，选择 8 mm 阶跃目标位移进行实验验证，结果如图 8 所示。在 8 mm 阶跃位移工况下，BP-ISTSMO 均能稳定估计实际位移，且其滑模抖振现象得到明显削弱。在 8 mm 阶跃工况下，估计位移与实际位移的最大误差不超过 0.31 mm，稳态误差保持在 0.03 mm 左右，其均方根误差（root mean square error，RMSE）值为 2.98×10^{-4} mm，位移估计准确度（准确度是用 1 减去整体数据的相对误差，常用来表示测得值与真值之间的符合程度）不低于 99.21%。

8 mm 阶跃工况下, 2 种观测器关键指标如表 3 所示, 说明 BP-ISTSMO 的位移估计性能相比于超螺旋滑模观测器得到了明显提升, 位移估计准确度提高了 0.9%, RMSE 指标减小了 13.12%, 最大误差减小了 16.22%。

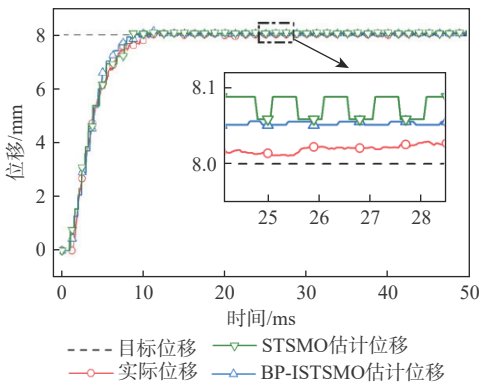


图 8　8 mm 阶跃工况位移估计效果

Fig. 8　Displacement estimation effect for 8 mm step condition

表 3　8 mm 阶跃工况下 2 种观测器关键指标对比

Table 3　Comparison of key metrics of two observers under 8 mm step condition

观测器类型	准确度/%	RMSE/10^{-4}mm	最大误差/mm
STSMO	98.35	3.43	0.37
BP-ISTSMO	99.21	2.98	0.31

3.2.2　正弦实验验证

轨迹跟踪能力是评价电磁直线执行器控制系统性能好坏的重要指标, 为验证 BP-ISTSMO 在正弦工况下的位移估计性能, 选择 2 Hz 正弦目标位移进行实验验证, 结果如图 9 所示。在全局位移跟踪估计中, BP-ISTSMO 始终能稳定跟踪估计正弦目标, 具有良好的位移跟踪估计性能, 且滑模抖振现象得到明显削弱。估计位移与实际位移的最大误差为 0.43 mm, RMSE 为 5.91×10^{-4} mm。相比于 STSMO, BP-ISTSMO 的轨迹跟踪估计能力得到了明显提升, RMSE 指标减小了 13.32%, 最大误差减小了 9.10%。

3.3　无位移传感器控制性能分析

3.3.1　阶跃实验验证

8 mm 阶跃工况下, 电磁直线执行器无位移传感器与有位移传感器控制效果对比如图 10 所示。在 8 mm 阶跃工况下, 无位移传感器控制在响应阶段无超调, 在稳态阶段控制效果稳定, 验证了无位移传感器控制的有效性。在 8 mm 阶跃工况下, 相对于目标位移的稳态误差在 0.03 mm 左右, 与有位移传感器的稳态误差相当。不同阶跃工况下有无位移传感器控制性能指标如表 4 所示。可得, 在不同阶跃工况下, 电磁直线执行器无位移传感器控制

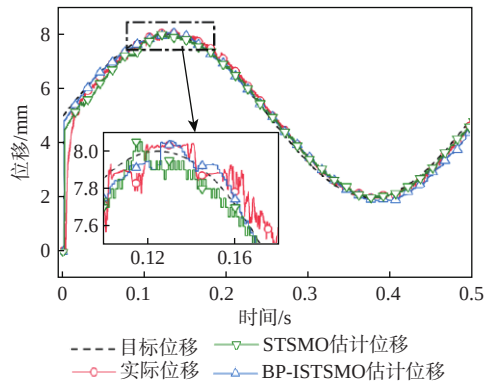


图 9　2 Hz 正弦工况位移估计效果

Fig. 9　Displacement estimation effect for 2 Hz sinusoidal condition

与有位移传感器控制的稳态误差数值相近, 均表现出了良好的性能, 验证了无位移传感器控制在不同阶跃工况下的有效性。

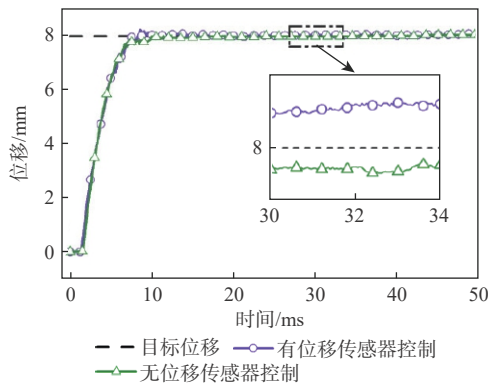


图 10　8 mm 阶跃工况下无位移传感器控制效果

Fig. 10　Control effect of displacement sensorless under 8 mm step condition

表 4　不同阶跃工况下无位移传感器控制性能指标

Table 4　Performance indexes of displacement sensorless control under different step conditions

控制类型	目标位移/mm	稳态误差/mm
无位移传感器控制	3	0.02
	5	0.03
	8	0.03
有位移传感器控制	3	0.02
	5	0.02
	8	0.03

3.3.2　正弦实验验证

2 Hz 正弦目标工况下, 电磁直线执行器无位移传感器与有位移传感器控制效果对比如图 11 所示。在全局连续目标位移跟踪过程中, 无位移传感器控制均能稳定跟随目标位移, 表现出了良好的控制效果。在 2 Hz 正弦工况下, 有位移传感器控制和无位移传感器控制的 RMSE 指标分别为 5.67×10^{-4} mm 和

6.09×10^{-4} mm，最大误差分别为0.46 mm和0.49 mm，第1个峰值处的滞后时间分别为0.01 s和0.02 s。不同频率正弦工况下有无位移传感器控制性能指标如表5所示。可得，在较低频率的正弦工况下，无位移传感器控制的RMSE指标与最大误差和有位移传感器控制的2种指标在数值上相近，无位移传感器控制表现出了良好的性能，证明了电磁直线执行器无位移传感器控制在正弦工况下的有效性。

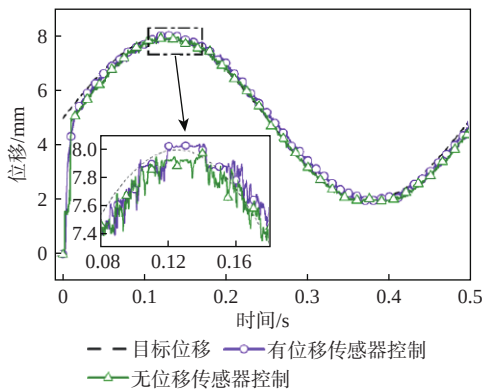


图11 2 Hz正弦工况下无位移传感器控制效果

Fig. 11 Control effect of displacement sensorless under 2 Hz sinusoidal condition

表5 不同频率正弦工况下无位移传感器控制性能指标

Table 5 Performance indexes of displacement sensorless control under sinusoidal conditions with different frequencies

控制类型	频率/Hz	RMSE/10^{-4} mm	最大误差/mm
无位移传感器控制	1	4.12	0.44
	2	6.09	0.49
有位移传感器控制	1	3.89	0.42
	2	5.67	0.46

4 结 论

1）本文将神经网络的自适应参数调整特性与滑模观测器的鲁棒性相结合，优化了电磁直线执行器的位移估计与无传感控制性能，突破了传统滑模观测器在抖振和参数适配方面的局限性。

2）采用非奇异滑模面与连续切换函数优化观测器结构，结合自适应参数调节，显著提升了位移估计精度与系统抗干扰能力。

未来将进一步研究高频正弦、变加速度、随机轨迹等复杂工况下方法的优化，提升观测器和控制系统的宽工况适配能力。

参考文献（References）

[1] 谭草, 葛文庆, 李波, 等. 双稳态线性力执行器静动态特性[J]. 电机与控制学报, 2020, 24(6): 119-126.
TAN C, GE W Q, LI B, et al. Static and dynamic characteristics of a bi-stable liner force actuator[J]. Electric Machines and Control, 2020, 24(6): 119-126(in Chinese).

[2] 谭草, 李波, 于鹏, 等. 电磁直驱静液作动器多学科建模与优化[J]. 北京理工大学学报, 2022, 42(10): 1080-1088.
TAN C, LI B, YU P, et al. Multidisciplinary modeling and optimization of direct-driving electro-hydrostatic actuator[J]. Transactions of Beijing Institute of Technology, 2022, 42(10): 1080-1088(in Chinese).

[3] 李波, 江金昊, 葛文庆, 等. 融合磁势矢模型和等效面电流模型的电磁直线驱动装置解析建模方法[J]. 机械工程学报, 2024, 60(20): 261-270.
LI B, JIANG J H, GE W Q, et al. Analytical modeling method of electromagnetic linear drive device combining magnetic potential vector model and equivalent surface current model[J]. Journal of Mechanical Engineering, 2024, 60(20): 261-270(in Chinese).

[4] 沈艳霞, 罗昌茜. 基于超螺旋滑模观测器的永磁同步直线电机无模型控制[J]. 电力系统保护与控制, 2023, 51(18): 62-69.
SHEN Y X, LUO C X. Model-free control of a permanent magnet linear synchronous motor based on a super-twisting sliding mode observer[J]. Power System Protection and Control, 2023, 51(18): 62-69(in Chinese).

[5] 于杰, 李波, 葛文庆, 等. 基于LSSVM的电磁直线执行器位移估算研究[J]. 传感器与微系统, 2023, 42(6): 58-61.
YU J, LI B, GE W Q, et al. Research on displacement estimation of electromagnetic linear actuator based on LSSVM[J]. Transducer and Microsystem Technologies, 2023, 42(6): 58-61(in Chinese).

[6] 李强. 基于支持向量机的直线驱动装置位置预测方法研究[D]. 淄博: 山东理工大学, 2020.
LI Q. Research on position prediction method of linear actuator based on support vector machine[D]. Zibo: Shandong University of Technology, 2020(in Chinese).

[7] 谭草, 黎德祥, 葛文庆, 等. 改进LuGre模型的电磁直线执行器自适应鲁棒控制[J]. 电机与控制学报, 2022, 26(10): 130-138.
TAN C, LI D X, GE W Q, et al. Adaptive robust control for electromagnetic linear actuator with improved LuGre model[J]. Electric Machines and Control, 2022, 26(10): 130-138(in Chinese).

[8] 李玉娇. 电磁直线驱动装置无传感控制技术研究[D]. 淄博: 山东理工大学, 2019.
LI Y J. Sensorless control technology study of electromagnetic linear driving device[D]. Zibo: Shandong University of Technology, 2019(in Chinese).

[9] DÜLK I, KOVÁCSHÁZY T. Parameter estimation in linear electromagnetic devices[J]. IEEE Transactions on Industrial Electronics, 2015, 62(6): 3619-3628.

[10] LIU K, YANG P, WANG R J, et al. Observer-based adaptive fuzzy finite-time attitude control for quadrotor UAVs[J]. IEEE Transactions on Aerospace and Electronic Systems, 2023, 59(6): 8637-8654.

[11] 丰富, 胡海林, 钟德鸣, 等. 基于改进互联型全阶观测器的直线感应电机在线参数辨识[J]. 西南交通大学学报, 2024, 59(4): 776-785.
FENG F, HU H L, ZHONG D M, et al. Online parameter identification of linear induction motors based on improved interconnected full-order observer[J]. Journal of Southwest Jiaotong University, 2024, 59(4): 776-785(in Chinese).

[12] MERCORELLI P. A two-stage augmented extended Kalman filter as an observer for sensorless valve control in camless internal combustion engines[J]. IEEE Transactions on Industrial Electronics, 2012, 59(11): 4236-4247.

[13] KHODARAHMI M, MAIHAMI V. A review on Kalman filter models[J]. Archives of Computational Methods in Engineering, 2023, 30(1): 727-747.

[14] 杨洋, 苏伟. 基于改进滑模观测器的四旋翼无人机轨迹控制[J]. 电光与控制, 2024, 31(7): 20-26.
YANG Y, SU W. Trajectory control of quadrotor UAVs based on improved sliding mode observer[J]. Electronics Optics & Control, 2024, 31(7): 20-26(in Chinese).

[15] 张会林, 王国强, 杨海马. 神经网络优化二阶滑模观测器的PMSM无感控制[J]. 控制工程, 2022, 29(11): 2075-2081.
ZHANG H L, WANG G Q, YANG H M. Sensorless control of PMSM based on second-order sliding mode observer optimized by neural network[J]. Control Engineering of China, 2022, 29(11): 2075-2081(in Chinese).

[16] WU Y C, FENG J W. Development and application of artificial neural network[J]. Wireless Personal Communications, 2018, 102(2): 1645-1656.

[17] 谭草, 鲁应涛, 葛文庆, 等. 直驱式永磁直线电机深度模糊滑模-自抗扰控制[J]. 西安交通大学学报, 2023, 57(1): 185-194.
TAN C, LU Y T, GE W Q, et al. Depth fuzzy sliding-mode active disturbance rejection control method of permanent magnet linear motor for direct drive system[J]. Journal of Xi'an Jiaotong University, 2023, 57(1): 185-194(in Chinese).

[18] WLAS M, KRZEMINSKI Z, GUZINSKI J, et al. Artificial-neural-network-based sensorless nonlinear control of induction motors[J]. IEEE Transactions on Energy Conversion, 2005, 20(3): 520-528.

[19] 谭草, 于鹏, 李波, 等. 直驱泵阀协同式线控制动单元压力串级控制[J]. 北京航空航天大学学报, 2025, 51(4): 1163-1171.
TAN C, YU P, LI B, et al. Pressure cascade control of brake-by-wire unit based on direct drive pump-valve cooperative[J]. Journal of Beijing University of Aeronautics and Astronautics, 2025, 51(4): 1163-1171(in Chinese).

Displacement sensorless control of electromagnetic linear actuator based on an improved sliding mode observer

GE Wenqing[1], LI Detong[1, 2], SONG Yadong[1], TAN Cao[1], LI Bo[1, *]

(1. School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo 255000, China;
2. Weichai Power Co., Ltd., Weifang 261061, China)

Abstract: A displacement estimation technique based on BP neural networks is suggested to enhance the super-twisting sliding mode observer in order to address the installation challenges, cost rise, and stability reduction of electromagnetic linear actuators induced by the usage of displacement sensors. Combined with an adaptive integral robust control algorithm, the displacement sensorless control of the electromagnetic linear actuator is realized. A non-singular fast terminal sliding mode surface is designed with the continuous hyperbolic tangent function as the switching function in order to reduce the buffeting phenomenon and enhance the displacement estimation performance of the super-twisting sliding mode observer; in terms of the observer's parameter adjustment, the BP neural network is designed to dynamically adjust the super-twisting sliding mode observer's gain using the input of the mover speed. The motion control performance test platform of the electromagnetic linear actuator is established, and the displacement estimation and feedback control results are analyzed. The results show that the maximum displacement estimation error of the improved sliding mode observer is reduced by 16.22% under the step condition and 9.10% under the sine condition with the frequency of 2 Hz compared with the super-twisting sliding mode observer; the control performance of displacement sensorless control is equivalent to that of displacement sensor control. The steady state error of the two is 0.03 mm under the 8 mm step condition, and the maximum error is 0.43 mm under the sinusoidal condition with the frequency of 2 Hz. This proves the effectiveness and practicability of the displacement sensorless control of the electromagnetic linear actuator based on an improved super-twisting sliding mode observer.

Keywords: electromagnetic linear actuator; super-twisting sliding mode observer; neural network; hyperbolic tangent function; displacement sensorless control

Received: 2024-05-08; **Accepted**: 2024-05-24; **Published Online**: 2024-06-13 11:29

URL: link.cnki.net/urlid/11.2625.V.20240612.1716.002

Foundation items: National Natural Science Foundation of China (52375105); Outstanding Youth Fund Project of Shandong Provincial Natural Science Foundation (ZR2022YQ51); Shandong Provincial Natural Science Foundation (ZR2023ME178,ZR2023ME177); Innovation Team Project of "Qing-Chuang Science and Technology Plan" of Colleges and Universities in Shandong Province (2022KJ232)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0243

融合混合搜索算子和竞争学习的海洋捕食者算法及应用

徐中辉[1]，饶振远[1,2]，马艳丽[3]，汤泽京[1]，黄晓东[3,*]

(1. 江西理工大学 信息工程学院，赣州 341000；2. 景德镇市卫生学校，景德镇 333000；
3. 江西理工大学 经济管理学院，赣州 341000)

摘　　要：针对经典海洋捕食者算法(MPA)在迭代后期种群多样性丧失、求解精度不高且难以跳出局部最优解等问题，提出了一种融合混合搜索算子和竞争学习的海洋捕食者算法PSMPA。引入随机动态重心反向学习机制，增强算法后期的种群多样性，扩展搜索空间，提升算法逃离局部最优解并加速收敛的能力；融合动态随机搜索和模式搜索作为混合搜索算子，增强算法的局部搜索能力;将竞争学习行为模式引入捕食者当中，改善种群平均适应度值，有效促进算法的快速收敛，显著提高解的质量。选用12个CEC2017测试函数进行仿真实验，结果表明：PSMPA在寻优性能、收敛速度和稳定性方面均取得了较大程度的改善。在太阳能光伏模型参数优化设计问题上的应用进一步验证了PSMPA在实际工程优化问题中的应用价值和有效性。

关 键 词：海洋捕食者算法；混合搜索算子；随机动态重心反向学习；竞争学习；工程优化问题

中图分类号：TP301.6

文献标志码：A　　**文章编号**：1001-5965(2026)06-1810-17

海洋捕食者算法(marine predators algorithm, MPA)是由Faramarzi等于2020年提出的群智能(swarm intelligence, SI)优化算法[1]。MPA以海洋捕食者觅食策略为灵感，利用捕食者和猎物之间最优相遇率策略，相较于遗传算法(genetic algorithm, GA)[2]、粒子群算法(particle swarm optimization, PSO)[3]、布谷鸟搜索算法(cuckoo search, CS)[4]，其具有简单灵活和对算法参数进行重大调整较少等优点[1]，因此，MPA已被用于解决许多实际工程问题，如目标分配或多目标任务分配、车间调度等领域[5]。尽管MPA已在多个工程优化领域取得了广泛应用，但其求解复杂工程优化问题时，仍存在搜索能力较弱、求解精度不高及容易陷入局部最优等不足[6]，因此，许多学者从不同角度对经典MPA进行了改进。

马驰等[7]引入混沌映射与对立学习相结合，生成高质量初始种群，同时引入自适应t分布变异因子，保证了MPA探索和开发之间的平衡性，提高了种群多样性，避免MPA陷入局部最优，并对更新后的种群根据适应度划分为精英、学习2组，进一步提高种群质量和搜索精度。张磊等[8]提出了融合精英反向学习(opposition-based learning, OBL)机制，提升了原算法的初始种群质量，提高了全局搜索范围，并通过加入黄金正弦策略，改进了捕食者猎物方式，提高了局部开发能力。Shaheen等[9]针对MPA在解决非线性、非凸、连续或非连续目标函数

收稿日期：2024-04-23；录用日期：2024-08-09；网络出版时间：2024-10-12 11:20
网络出版地址：link.cnki.net/urlid/11.2625.V.20241012.0858.001
基金项目：国家重点研发计划项目子课题(2020YFB1713700,2020YFB1713705)
*通信作者. E-mail：huangmail3@126.com

引用格式：徐中辉，饶振远，马艳丽，等. 融合混合搜索算子和竞争学习的海洋捕食者算法及应用[J]. 北京航空航天大学学报，2026，52(6)：1810-1826. XU Z H，RAO Z Y，MA Y L，et al. Marine predator algorithm incorporating hybrid search operators and competitive learning and applications[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1810-1826 (in Chinese).

中存在的全局搜索能力不足的问题, 提出了一种在迭代前期用高尾威布尔分布(Weibull distribution, WD)策略取代经典 MPA 中的布朗运动, 以提高全局搜索能力。Aydemir [10] 在 MPA 中引入了基于泰勒的最优领域策略, 提升了最优个体更新效率, 并降低了陷入局部最优的可能性, 同时采用动态选项学习的非对称搜索策略, 提高了种群达到最优的概率。Kusuma 等[11] 将 MPA 与隐藏对线游戏算法相结合, 提升了 MPA 的全局优化能力, 并在减少每个解的归一化分数方面优化了计算过程。

综上所述, 现有对 MPA 的改进措施主要集中在改善种群初始分布、调整捕食者运动方式及改善捕食者最佳位置等方面, 但在种群信息优化和平均种群适应度提升等方面研究相对较少。为此, 本文从优化种群信息、提升局部搜索能力和提升种群平均适应度 3 个方面出发, 提出了一种融合混合搜索算子和竞争学习的海洋捕食者算法 PSMPA。采用随机动态重心反向学习 (stochastic dynamic centroid opposition-based learning, SCOBL) 机制, 加入算法后期位置更新中, 解决了迭代后期种群多样性不足导致算法易陷入局部最优的问题, 增强了后期种群信息交流和多样性保持; 融合动态随机搜索和模式搜索作为混合搜索算子, 避免了单一搜索算子局部搜索过强导致局部最优陷阱, 提高了算法局部搜索能力; 引入海洋内捕食者相互竞争学习的行为, 提高了种群平均适应度, 加快了算法的收敛速度和收敛精度。

1　海洋捕食者算法

MPA 是一种模拟海洋捕食行为的群智能优化算法, 其模拟了海洋生态系统中捕食者和被捕食者之间的关系。猎物位置初始化均匀分布在搜索空间上, 最适解作为精英捕食者用于构建精英矩阵 $\boldsymbol{E}$, 表示为

$$\boldsymbol{E}=\begin{bmatrix} X^{\mathrm{I}}_{1,1} & X^{\mathrm{I}}_{1,2} & \cdots & X^{\mathrm{I}}_{1,d} \\ X^{\mathrm{I}}_{2,1} & X^{\mathrm{I}}_{2,2} & \cdots & X^{\mathrm{I}}_{2,d} \\ \vdots & \vdots & & \vdots \\ X^{\mathrm{I}}_{n,1} & X^{\mathrm{I}}_{n,2} & \cdots & X^{\mathrm{I}}_{n,d} \end{bmatrix}_{n\times d} \tag{1}$$

式中: $\boldsymbol{X}^{\mathrm{I}}$为最佳捕食者向量, 被复制$n$次构建精英矩阵; n为种群的数量; d为个体维度数量。迭代结束时, 最佳捕食者被更好的捕食者取代, 精英则被更新。

捕食者根据猎物矩阵来更新自己的位置。初始化阶段创建了最初的猎物, 其中, 最适合的一个猎物(捕食者)用来构建精英。猎物矩阵表示为

$$\boldsymbol{P}=\begin{bmatrix} X_{1,1} & X_{1,1} & \cdots & X_{1,d} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,d} \\ \vdots & \vdots & & \vdots \\ X_{n,1} & X_{n,2} & \cdots & X_{n,d} \end{bmatrix}_{n\times d} \tag{2}$$

MPA 优化过程分为如下 3 个阶段:

1) 阶段 1。迭代前期, 猎物速度快($V \geqslant 10$)。此时, 捕食者为停滞状态, 数学模型如下:

$$\text{While } t < T/3 \quad i = 1,2,\cdots,n$$

$$\boldsymbol{S}_i = \boldsymbol{R}_{\mathrm{B}} \otimes (\boldsymbol{E}_i - \boldsymbol{R}_{\mathrm{B}} \otimes \boldsymbol{P}_i) \tag{3}$$

$$\boldsymbol{P}_i = \boldsymbol{P}_i + P_{\mathrm{c}} \boldsymbol{R} \otimes \boldsymbol{S}_i \tag{4}$$

式中: $\boldsymbol{R}_{\mathrm{B}}$为布朗运动的正态分布随机向量; P_{c}为0.5的常数; $\boldsymbol{R}$为[0,1]间均匀随机数构成的向量; t为当前迭代次数; T为最大迭代次数。

2) 阶段 2。迭代中期, 捕食者和猎物速度相等($V = 1$)。此时, 种群的一半用于勘探, 另一半用于开采。在该阶段中, 猎物进行莱维运动, 而捕食者则进行布朗运动。前一半种群更新公式如下:

$$\text{While } T/3 < t < 2T/3 \quad i = 1,2,\cdots,n/2$$

$$\boldsymbol{S}_i = \boldsymbol{R}_{\mathrm{L}} \otimes (\boldsymbol{E}_i - \boldsymbol{R}_{\mathrm{L}} \otimes \boldsymbol{P}_i) \tag{5}$$

$$\boldsymbol{P}_i = \boldsymbol{P}_i + P_{\mathrm{c}} \boldsymbol{R} \otimes \boldsymbol{S}_i \tag{6}$$

后一半种群更新公式如下:

$$\text{While } T/3 < t < 2T/3 \quad i = n/2,\cdots,n$$

$$\boldsymbol{S}_i = \boldsymbol{R}_{\mathrm{B}} \otimes (\boldsymbol{E}_i - \boldsymbol{R}_{\mathrm{B}} \otimes \boldsymbol{P}_i) \tag{7}$$

$$\boldsymbol{P}_i = \boldsymbol{E}_i + P_{\mathrm{c}} c \otimes \boldsymbol{S}_i \tag{8}$$

$$c = \left(1 - \frac{t}{T}\right)^{\frac{2t}{T}} \tag{9}$$

式中: $\boldsymbol{S}_i$为第 i 个个体的更新步长向量; $\boldsymbol{R}_{\mathrm{L}}$为基于莱维运动的随机数的向量; c为控制捕食者运动步长的自适应参数。

3) 阶段 3。迭代后期, 捕食者速度快($V = 0.1$), 此时捕食者的最佳策略是进行莱维运动, 数学模型如下:

$$\text{While } t > 2T/3 \quad i = 1,2,\cdots,n$$

$$\boldsymbol{S}_i = \boldsymbol{R}_{\mathrm{L}} \otimes (\boldsymbol{R}_{\mathrm{L}} - \boldsymbol{E}_i \otimes \boldsymbol{P}_i) \tag{10}$$

$$\boldsymbol{P}_i = \boldsymbol{E}_i + P_{\mathrm{c}} c \otimes \boldsymbol{S}_i \tag{11}$$

涡流形成或鱼类聚集装置(fish aggregating device, FAD)环境情况会影响捕食者, FAD 的存在可以避免算法陷入局部最优解, 且可以有效增加搜索的广度和深度, 提高算法的全局搜索能力和优化性能。数学模型如下:

$$\boldsymbol{P}_i=\begin{cases}\boldsymbol{P}_i+c(\boldsymbol{X}_{\min}+\boldsymbol{R}\otimes(\boldsymbol{X}_{\max}-\boldsymbol{X}_{\min}))\otimes\boldsymbol{U} & r\leqslant p_{\mathrm{f}}\\ \boldsymbol{P}_i+\left[p_{\mathrm{f}}(1-r)+r\right]\left(\boldsymbol{P}_{r_1}-\boldsymbol{P}_{r_2}\right) & r>p_{\mathrm{f}}\end{cases} \tag{12}$$

式中：$p_{\mathrm{f}}=0.2$为FAD效应影响概率；$\boldsymbol{U}$为二进制随机向量数组，区间为[0,1]，如果数组小于0.2，则将其修改为0，如果大于或等于0.2，则将其修改为1；r为[0,1]中的均匀随机数；$\boldsymbol{X}_{\min}$和$\boldsymbol{X}_{\max}$分别为包含维度的下限和上限的向量；r_1和r_2为猎物矩阵的随机索引。

海洋记忆是用来更新猎物并实施FAD效果后，评估该矩阵是否适合更新精英的一种方式。将当前迭代的每个解的适合度与其在先前迭代中的等效值进行比较，如果更适合，则当前解将替换该解。

2 PSMPA算法

2.1 混合搜索算子

为实现增强MPA的局部开发能力，进而提升在优化问题上的收敛性能，引入动态随机搜索和模式搜索作为混合搜索算子，在当前最优解的邻域内进行精细搜索，提升最优解的质量，同时混合搜索算子避免了单一局部算子较强的局部搜索能力导致处理部分函数时易陷入局部最优的问题。

2.1.1 动态随机搜索

动态随机搜索是由Hamzacebi等[12]提出的一种算法，其由一般搜索和局部搜索2部分组成。局部搜索算子是在当前解的领域中展开随机搜索，且搜索步长逐渐减小。动态随机搜索技术的核心参数是L和N，根据实验经验判断，这2个参数的设置会影响算法的最终结果，一般取值为$L=100$，$N=10$，可达到对当前解附近搜索空间的充分开发，具体步骤如下：

步骤1 设置局部搜索迭代次数为L，初始搜索步长为α_0，局部搜索迭代计数器为e=0，k=0，$\boldsymbol{x}_{\mathrm{current}}=\boldsymbol{x}_{\mathrm{best}}$。

步骤2 重置迭代计数器$n=0$。

步骤3 生成随机向量$\boldsymbol{d}_x$，且满足$-\alpha_k\leqslant d_{x,j}\leqslant\alpha_k, j=1,2,\cdots,d$。

步骤4 更新e，e=e+1。

步骤5 f_{new}=$f(\boldsymbol{x}_{\mathrm{current}}+\boldsymbol{d}_x)$。若$f_{\mathrm{new}}<f_{\mathrm{best}}$，则$f_{\mathrm{best}}=f_{\mathrm{new}}$，$\boldsymbol{x}_{\mathrm{bset}}=\boldsymbol{x}_{\mathrm{current}}+\boldsymbol{d}_x$，$n=n+1$，转步骤7；若$f_{\mathrm{new}}<f_{\mathrm{current}}$，则$f_{\mathrm{current}}=f_{\mathrm{new}}$，$\boldsymbol{x}_{\mathrm{current}}=\boldsymbol{x}_{\mathrm{current}}+\boldsymbol{d}_x$，$n=n+1$，转步骤7。

步骤6 f_{new}=$f(\boldsymbol{x}_{\mathrm{current}}-\boldsymbol{d}_x)$。若$f_{\mathrm{new}}<f_{\mathrm{best}}$，则$f_{\mathrm{best}}=f_{\mathrm{new}}$，$\boldsymbol{x}_{\mathrm{bset}}=\boldsymbol{x}_{\mathrm{current}}-\boldsymbol{d}_x$，$n=n+1$，转步骤7；若$f_{\mathrm{new}}<f_{\mathrm{current}}$，则$f_{\mathrm{current}}=f_{\mathrm{new}}$，$\boldsymbol{x}_{\mathrm{current}}=\boldsymbol{x}_{\mathrm{current}}-\boldsymbol{d}_x$，$n=n+1$，转步骤7。

步骤7 若$n<N$，则转步骤3。

步骤8 $k=k+1$。

步骤9 $\alpha_k=\alpha_{k-1}\times0.5$。

步骤10 若e=L，则算法终止，否则转步骤2。

其中，e为局部搜索迭代计数器；$\boldsymbol{x}_{\mathrm{best}}$为算法当前最优解；$f(\cdot)$为计算适应度值的函数；$f_{\mathrm{best}}$、$f_{\mathrm{new}}$、$f_{\mathrm{current}}$为解的适应值。

2.1.2 模式搜索

模式搜索[13]是一种无需优化问题的梯度信息的数值优化方法，其无导数和快速收敛性在此后的高维空间中搜索时，被诸多学者证明是可靠有效的优化方法。

模式搜索主要由探测和模式2种移动过程组成。探测移动沿坐标轴方向移动，模式移动沿相邻2个探测点连线方向移动。2种移动交替进行，具体步骤如下：

步骤1 给定初始点$\boldsymbol{x}^0$，初始步长δ，加速系数$\alpha>0$，收缩系数$\beta>0$，精度$\varepsilon>0$，$k=1$，$j=1$，令$\boldsymbol{y}^1=\boldsymbol{x}^1$。

步骤2 从$\boldsymbol{y}$出发，依次作平行于单位矢量$\boldsymbol{e}_j(j=1,2,\cdots,n)$的轴向探测移动，若$f\left(\boldsymbol{y}^j+\delta\boldsymbol{e}_j\right)<f(\boldsymbol{y}^j)$，则$\boldsymbol{y}^{j+1}=\boldsymbol{y}^j+\delta\boldsymbol{e}_j$，转步骤4；否则，执行步骤3。

步骤3 若$f\left(\boldsymbol{y}^j-\delta\boldsymbol{e}_j\right)\leqslant f(\boldsymbol{y}^j)$，则$\boldsymbol{y}^{j+1}=\boldsymbol{y}^j-\delta\boldsymbol{e}_j$，执行步骤4；否则，$\boldsymbol{y}^{j+1}=\boldsymbol{y}^j$，执行步骤4。

步骤4 若j<n，则j=j+1，转到步骤2；否则，转到步骤5。

步骤5 若$f\left(\boldsymbol{y}^{n+1}\right)\leqslant f(\boldsymbol{x}^k)$，则执行步骤6；否则，执行步骤7。

步骤6 $\boldsymbol{x}^{k+1}=\boldsymbol{y}^{n+1}$，令$\boldsymbol{y}^1=\boldsymbol{x}^{k+1}+\alpha(\boldsymbol{x}^{k+1}-\boldsymbol{x}^k)$，$k=k+1$，$j=1$，转步骤2。

步骤7 若$\delta\leqslant\varepsilon$，则停止迭代，得到点$\boldsymbol{x}^k$，否则，$\delta=\beta\delta$，$\boldsymbol{y}^1=\boldsymbol{x}^k$，$\boldsymbol{x}^{k+1}=\boldsymbol{x}^k$，$k=k+1$，$j=1$，转步骤2。

每次使用混合搜索算子后，对搜索后的结果进行边界控制，保证其处在搜索域内。之后借助贪婪选择，判断得到的结果是否比当前最优解更好，否则，保持当前最优解不变。边界控制可表示为

$$X_{\mathrm{temp}}=\begin{cases}l_{\mathrm{b}} & X_{\mathrm{temp},j}<l_{\mathrm{b}}\\ u_{\mathrm{b}} & X_{\mathrm{temp},j}>u_{\mathrm{b}}\end{cases} \tag{13}$$

式中：$\boldsymbol{X}_{\mathrm{temp}}$为搜索后的结果；$u_{\mathrm{b}}$、$l_{\mathrm{b}}$分别为上、下界。

2.2 随机动态重心反向学习

MPA在探索和开发阶段通常只利用最佳种群个体进行迭代搜索，而未充分利用种群的有效信息，导致迭代后期种群多样性降低和容易陷入局部最优等缺陷。在MPA探索和开发阶段后，位置更

新阶段引入 SCOBL 策略, 以此维持种群多样性和跳出局部最优。

反向学习是一种优化机制, 最初由 Tizhoosh 提出[14], 其主要思想是: 利用反向解扩大搜索范围, 通过评价当前解和反向解找到更优解来解决优化问题。但标准的反向学习也存在单调性和局部优化的风险, 没有利用整个种群的搜索信息。由此, 为考虑种群间的信息, Rahnamayan 等提出重心反向学习 (centroid opposition-based learning, COBL)[15], 该策略引入种群重心与重心的反向点。重心及基于重心的反向点定义如下。

定义 1　设 $(Y_1,Y_2,\cdots,Y_n)$ 是 D 维搜索空间中带有单位质量的 n 个点, 整体重心如下:

$$M_j=\frac{1}{n}\sum_{i=1}^{D}y_{i,j}\qquad i=1,2,\cdots,D \tag{14}$$

定义 2　若离散均匀的整体的重心为 M, 则该整体中的某个点的反向点如下:

$$\overline{Y}_i=2M-Y_i\qquad i=1,2,\cdots,n \tag{15}$$

本文在 COBL 策略上引入随机概念, 提出 SCOBL 策略。选取种群中的前 K 个最优个体(K 为 1 到种群规模数之间的随机数), 计算前 K 个个体重心反向点, 反向点通过式(13)边界控制后借助贪婪选择最优反向点。

SCOBL 策略避免了 COBL 中重复点和重复评估, 同时扩展了反向搜索范围, 提升了算法的收敛速度和收敛精度。图 1 为 SCOBL 策略示意图。

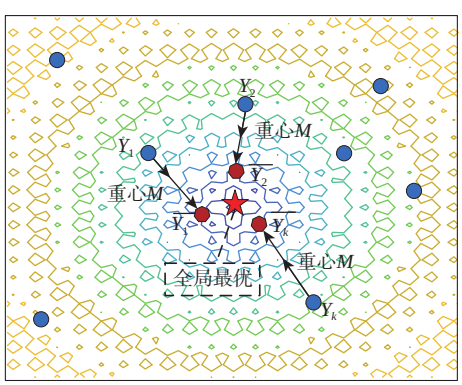


图 1　SCOBL 策略示意图

Fig. 1　Schematic diagram of SCOBL

2.3　竞争学习策略

除了根据外部条件进行选择和转换外, 为能在海洋中生存下去, 捕食者还可以通过内部机制进行学习和竞争, 以提高生存几率。位置差的捕食者需要向位置好的捕食者学习捕食本领进化, 同时, 位置差的捕食者之间需要竞争学习进化。基于该思想, 提出了一种竞争学习策略。将 FAD 影响后的猎物按照适应度排序, 将种群分为精英个体和学习个体 2 类, 适应度最优的捕食者称为精英个体, 其余适应度差的个体称为学习个体。进化过程分为“精英学习”和“竞争学习”2 个时段。在“精英学习”时段, 学习个体主动向精英个体学习, 从而达到提升学习个体的位置, 同时结合种群均值的作用, 能有效融合精英个体与学习个体之间相互促进作用。在“竞争学习”时段, 随机选取种群的 2 个个体, 相互竞争比较其优劣, 劣者向优者学习进化。

2.3.1　精英学习

学习个体通过精英个体和学习个体平均值之间的差异来向精英个体学习, 计算出当前种群的平均值 p_{mean}, 利用式 (16) 计算精英个体与学习个体的差异并更新位置, 具体过程如下:

$$p_{\text{new}}^{i}=p_{\text{old}}^{i}+s(p_{\text{elite}}-Fp_{\text{mean}}) \tag{16}$$

式中: p_{elite}为种群中精英个体; $p_{\text{mean}}=\dfrac{1}{Q}\sum_{i=1}^{Q}p^{i}$为学习个体的平均值, Q 为种群个体数; $s=\text{rand}(1,m)$为学习步长; F为学习因子, 随机取值为 1 或 2。

图 2 为精英学习阶段示意图。每个学习个体依据学习效果决定对当前位置是否更新, 若优于当前位置进行更新, 否则, 保留其原有位置。

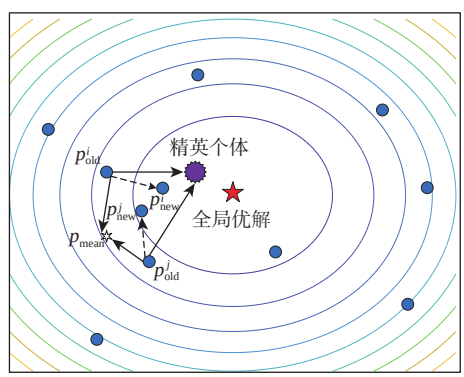


图 2　精英学习示意图

Fig. 2　Schematic diagram of elite learning

2.3.2　竞争学习

对种群中每个学习个体 p^i随机选取竞争对象 $p^j(i\neq j)$进行竞争学习, 通过计算 2 个学习个体间的目标函数适应度值来进行竞争学习更新位置。图 3 为竞争学习示意图, 具体过程如下:

$$p_{\text{new}}^{i}=\begin{cases}p_{\text{old}}^{i}+s\left(p^{i}-p^{j}\right) & f\left(p^{j}\right)<f\left(p^{i}\right)\\ p_{\text{old}}^{i}+s\left(p^{j}-p^{i}\right) & f\left(p^{i}\right)<f\left(p^{j}\right)\end{cases} \tag{17}$$

式中: $f(p^i)$和$f(p^j)$分别为学习个体 p^i和 p^j的目标函数适应度值。

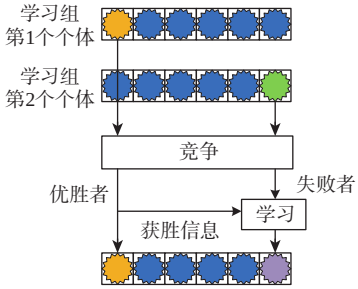


图3 竞争学习示意图

Fig. 3 Schematic diagram of competitive learning

每个学习个体依据学习效果决定对当前位置是否更新，若优于当前位置进行更新，否则，保留其原有位置。

竞争学习策略保留了种群最优个体的优势，同时学习个体通过向精英个体学习和相互之间竞争学习，提升了种群适应度，提高收敛精度和收敛速度。

2.4 PSMPA算法实现步骤流程

步骤1 初始化种群规模 N、目标函数维度 D，初始值上下界 l_b、u_b，最大迭代次数 T。

步骤2 计算适应度，构造精英矩阵 $\boldsymbol{E}$。

步骤3 根据式(3)～式(11)更新位置和步长。

步骤4 使用SCOBL策略(式(15))更新猎物位置。

步骤5 使用混合搜索算子更新猎物。

步骤6 应用FAD作用，根据式(12)更新种群。

步骤7 根据适应度值大小将种群分为精英个体和学习个体。

步骤8 学习个体根据式(16)向精英进行学习进化，根据式(17)进行竞争学习。

步骤9 判断算法是否满足结束条件，不满足则跳转至步骤2。

3 实验及结果分析

为全面验证PSMPA的性能，包括收敛速度、求解精度、鲁棒性和稳定性，仿真实验分为以下3个部分进行：

1）与经典群智能优化算法对比。将PSMPA与经典群智能优化算法在部分CEC2017测试函数上进行性能对比，验证PSMPA的有效性和稳定性。

2）与近期新型变体算法对比。将PSMPA与近期新型的MPA变体算法及其他变体算法在部分CEC2017测试函数上进行对比，并通过Wilcoxon秩和检验分析结果，验证PSMPA的优化性能、鲁棒性和显著性。

3）可扩展性分析。将CEC2017测试函数维度扩展至50维和100维进行寻优对比，验证PSMPA的可扩展性和高纬复杂问题上的卓越性。

本文使用CEC2017中12个测试函数进行实验验证。对于12个测试函数，F_1为单峰测试函数，F_2～F_4为多峰测试函数，F_5～F_8为混合测试函数，F_9～F_{12}为组合测试函数，表1给出了函数名及其全局最优点。

表1 CEC2017测试函数

Table 1 CEC2017 test functions

类型	标号	函数名称	最优值
单峰函数	F_1	Shifted and Rotated Bent Cigar Function	100
简单多峰函数	F_2	Shifted and Rotated Zakharov Function	300
	F_3	Shifted and Rotated Rosenbrock's Function	400
	F_4	Shifted and Rotated Schwefel's Function	1 000
混合函数	F_5	Hybrid Function 2(N=3)	1 200
	F_6	Hybrid Function 3(N=3)	1 300
	F_7	Hybrid Function 5(N=4)	1 500
	F_8	Hybrid Function 6(N=4)	1 600
组合函数	F_9	Hybrid Function 6(N=6)	2000
	F_{10}	Composition Function 1 (N=3)	2 100
	F_{11}	Composition Function 2 (N=3)	2 200
	F_{12}	Composition Function 8 (N=6)	2 800

实验对于所有算法参数设置与原文献使用参数相同，如表2[1,16-25]所示。算法最大迭代次数设置为600，种群数量设置为30，分别对每个函数独立运行30次，计算平均值、标准差和排名。实验环境为AMD Ryzen™ 7 7735HS CPU @3.20 GHz，并采用MATLAB2022a进行仿真。

3.1 与经典群智能优化算法性能对比

为比较PSMPA与其他经典群智能优化算法的优劣，选取早期经典的最优引导人工蜂群算法(gbest-guided artificial bee colony algorithm, GABC)[16]、

表 2　算法参数[1,16-25]

Table 2　Algorithm parameters[1,16-25]

算法	参数设置
MPA[1]	FAD效应概率为0.2, 步长控制因子为0.5
FA[17]	光吸收系数为1, 最大吸引力为2, 随机扰动因子为0.2, 扰动衰减因子为0.98
GABC[16]	位置更新步长因子为0.5, 均值学习权重为0.5
CLPSO[19]	种群规模为100, 认知因子为1, 社会因子为2, 惯性权重为[0.4,0.9], 惯性权重下限为[0.4,0.9]
SaDE[18]	学习周期为20
PSMPA	FAD效应概率为0.2, 步长控制因子为0.5
NMPA[21]	FAD效应概率为0.2, 步长控制因子为0.5
QQLMPA[20]	FAD效应概率为0.2, 步长控制因子为0.8, 局部搜索权重为0.8
RDPSO[22]	种群规模为100, 认知因子为2, 社会因子为2, 惯性权重下限为[0.4,0.9]
PSO_sono[23]	种群规模为100, 惯性权重下限为[0.4,0.9], 收敛控制因子为0.1×问题维度/种群规模, 社会学习因子为0.5
SHADE[24]	交叉概率均值为0.5, 缩放因子均值为0.5
LMHHO[25]	长期记忆限制为10

萤火虫算法(firefly algorithm, FA)[17]、策略自适应差分进化算法(differential evolution algorithm with strategy adaptation, SaDE)[18]和综合学习粒子群算法(comprehensive learning particle swarm optimizer, CLPSO)算法[19], 并使用表 1 中的 12 个 CEC2017 测试函数(30 维)中的随机 9 个进行寻优结果对比, 实验结果如表 3 所示。

从表 3 可得出, 在单峰测试函数中, PSMPA 的平均值为第一, 显著优于其他算法, 同时标准差为所有算法中最小, 超过其他经典算法多个量级。在多峰测试函数中, PSMPA 的平均值和标准差均排名第一, 显示了其卓越的全局搜索能力和多样性保持能力。在混合测试函数中, 除函数 F_8 外, PSMPA 的平均值和标准差仍排名第一。在函数 F_8 中, 虽然标准差排名第三, 略低于 SaDE 和 GABC, 但平均值最优。在组合测试函数中, PSMPA 的平均值和标准差保持第一。综上可得, PSMPA 在单峰、多峰、混合及组合函数仿真实验中, 相较于其他经典群智能优化算法, 拥有更优异的寻优性能和鲁棒性。

3.2　与近期新型变体算法性能对比

将 PSMPA 与经典 MPA 算法、近期发表的 2 种改进 MPA 变体算法 QQLMPA16[20] 和 NMPA[21]、2 种改进 PSO 算法 RDPSO[22] 和 PSO_sono[23]、SHADE[24] 及改进的哈里斯鹰优化算法(Harris hawks optimization, HHO)LMHHO[25], 以此判断 PSMPA 是否具有优于其他类似算法的性能及寻优能力。

3.2.1　测试函数统计结果分析

采用本文改进的 PSMPA 与 MPA、NMPA、QQLMPA、RDPSO、PSO_sono、SHADE、LMHHO 进行对比测试, 使用平均值和标准偏差分别对表 1 中 CEC2017 测试函数(30 维)进行优化求解实验测试, 表 4 展示了不同算法在 12 个测试函数 30 维度上的实验结果。

从表 4 分析可得, F_1 为单峰测试函数, 用于评价算法的局部搜索能力和收敛速度。在函数 F_1 中, PSMPA 的平均值显著优于其他算法, 更接近测试函数的全局最优值, 同时标准差最小。这表明了混合搜索算子在当前最优解周围进行精细搜索, 提升了最优解的质量, 证明了 PSMPA 在局部开发能力方面的卓越表现。

F_2～F_4 为多峰测试函数, 用于评估算法的全局搜索能力和多样性保持。在函数 F_2、F_3 中, PSMPA 的平均值排名第一。SCOBL 策略利用了种群的动态有效信息, 提升了全局搜索能力, 并确保迭代后期种群多样性的维持, 避免陷入局部最优。混合搜索算子的应用避免了单一算子局部搜索能力过强导致陷入局部最优, 提升了解的质量。在函数 F_4 中, PSMPA 的平均值和 MPA 几乎相当, 但寻优效果远超其他 MPA 变体及其他 4 种算法, 证明了 PSMPA 具有较强的全局搜索和跳出局部最优的能力。

F_5～F_8 为混合测试函数, 用于评估算法处理复杂问题和鲁棒性。在函数 F_5～F_7 中, PSMPA 均获得了最优的平均值和标准差。在 F_8 上, 尽管标准差表现不及经典 MPA、2 种 MPA 变体及 SHADE, 但 PSMPA 具有最优的寻优精度, 表明了 PSMPA 在处理复杂、多特性优化问题时, 具有更好的寻优精度和鲁棒性。

F_9～F_{12} 为组合测试函数, 用于全方位评估算法高维空间综合表现。在函数 F_9～F_{12} 中, PSMPA 均获得了最优的平均值, 并且在函数 F_9 和 F_{11} 中, 标准差也为最优。在函数 F_{10} 中, 标准差仅次于 SHADE, 排名第二。表明了 PSMPA 在处理复杂特

表3 经典算法测试函数结果统计

Table 3 Statistical results of test functions for classical algorithms

函数	算法	平均值	标准差
F_1	PSMPA	**1.00×10^{2}**	**2.45×10^{-1}**
	FA	7.19×10^{3}	6.78×10^{3}
	GABC	1.53×10^{9}	6.41×10^{8}
	CLPSO	3.60×10^{3}	2.22×10^{3}
	SaDE	3.18×10^{3}	2.99×10^{3}
F_2	PSMPA	**3.00×10^{2}**	**1.18×10^{-6}**
	FA	3.26×10^{2}	1.37×10^{2}
	GABC	1.69×10^{5}	4.13×10^{4}
	CLPSO	6.96×10^{4}	1.40×10^{4}
	SaDE	2.35×10^{4}	7.86×10^{3}
F_5	PSMPA	**1.43×10^{4}**	**3.50×10^{4}**
	FA	1.10×10^{5}	1.02×10^{5}
	GABC	8.61×10^{7}	4.99×10^{7}
	CLPSO	9.25×10^{5}	5.36×10^{5}
	SaDE	5.22×10^{5}	7.35×10^{5}
F_6	PSMPA	**1.99×10^{3}**	**3.09×10^{2}**
	FA	2.03×10^{4}	1.99×10^{4}
	GABC	7.32×10^{5}	5.56×10^{5}
	CLPSO	1.47×10^{4}	1.07×10^{4}
	SaDE	1.62×10^{4}	1.40×10^{4}
F_7	PSMPA	**1.59×10^{3}**	**4.12×10^{1}**
	FA	8.41×10^{3}	8.16×10^{3}
	GABC	1.39×10^{5}	2.51×10^{5}
	CLPSO	1.47×10^{4}	1.11×10^{4}
	SaDE	2.33×10^{3}	1.19×10^{3}
F_8	PSMPA	**2.29×10^{3}**	2.49×10^{2}
	FA	2.34×10^{3}	3.32×10^{2}
	GABC	3.82×10^{3}	2.23×10^{2}
	CLPSO	2.63×10^{3}	2.55×10^{2}
	SaDE	2.74×10^{3}	**1.93×10^{2}**
F_9	PSMPA	**2.18×10^{3}**	**7.13×10^{1}**
	FA	2.29×10^{3}	1.68×10^{2}
	GABC	3.04×10^{3}	2.16×10^{2}
	CLPSO	2.46×10^{3}	1.41×10^{2}
	SaDE	2.41×10^{3}	1.10×10^{2}
F_{10}	PSMPA	**2.35×10^{3}**	**1.34×10^{1}**
	FA	2.37×10^{3}	2.07×10^{1}
	GABC	2.57×10^{3}	2.42×10^{1}
	CLPSO	2.43×10^{3}	3.58×10^{1}
	SaDE	2.44×10^{3}	1.51×10^{1}
F_{11}	PSMPA	**2.30×10^{3}**	**1.37×10^{1}**
	FA	2.66×10^{3}	1.11×10^{3}
	GABC	9.56×10^{3}	2.11×10
	CLPSO	2.53×10^{3}	1.10×10^{3}
	SaDE	4.35×10^{3}	2.95×10^{3}

注: 黑体数据表示最佳结果。

表 4　近期变体算法测试函数结果统计

Table 4　Statistics results of test function for recent variants algorithms

函数	算法	平均值	标准差	排名	函数	算法	平均值	标准差	排名
F_1	PSMPA	$\mathbf{1.001\times10^{2}}$	$\mathbf{3.919\times10^{-1}}$	1	F_7	PSMPA	$\mathbf{1.617\times10^{3}}$	$\mathbf{4.251\times10^{1}}$	1
	MPA	3.306×10^{5}	3.283×10^{5}	3		MPA	1.701×10^{3}	4.967×10^{1}	2
	NMPA	9.765×10^{7}	7.434×10^{7}	7		NMPA	1.696×10^{3}	4.377×10^{1}	3
	QQLMPA	5.083×10^{7}	5.473×10^{7}	6		QQLMPA	1.728×10^{3}	5.939×10^{1}	4
	PSO_sono	4.380×10^{6}	6.529×10^{6}	4		PSO_sono	1.135×10^{4}	8.577×10^{3}	6
	RDPSO	1.120×10^{7}	1.543×10^{7}	5		RDPSO	1.024×10^{4}	1.214×10^{4}	7
	LMHHO	5.421×10^{10}	6.351×10^{9}	8		LMHHO	4.611×10^{8}	3.300×10^{8}	8
	SHADE	4.120×10^{2}	3.709×10^{2}	2		SHADE	1.979×10^{3}	6.128×10^{2}	5
F_2	PSMPA	$\mathbf{3.000\times10^{2}}$	$\mathbf{1.059\times10^{-6}}$	1	F_8	PSMPA	$\mathbf{2.262\times10^{3}}$	3.001×10^{2}	1
	MPA	4.010×10^{3}	2.292×10^{3}	2		MPA	2.301×10^{3}	2.254×10^{2}	2
	NMPA	5.633×10^{3}	1.995×10^{3}	4		NMPA	2.444×10^{3}	2.405×10^{2}	3
	QQLMPA	1.117×10^{4}	6.931×10^{3}	5		QQLMPA	2.806×10^{3}	2.315×10^{2}	6
	PSO_sono	2.139×10^{4}	6.702×10^{3}	6		PSO_sono	2.527×10^{3}	3.037×10^{2}	5
	RDPSO	2.948×10^{5}	1.599×10^{5}	8		RDPSO	3.197×10^{3}	4.939×10^{2}	7
	LMHHO	8.865×10^{4}	2.452×10^{3}	7		LMHHO	6.399×10^{3}	1.481×10^{3}	8
	SHADE	5.114×10^{3}	3.303×10^{3}	3		SHADE	2.519×10^{3}	$\mathbf{2.222\times10^{2}}$	4
F_3	PSMPA	$\mathbf{4.360\times10^{2}}$	4.590×10^{1}	1	F_9	PSMPA	$\mathbf{2.164\times10^{3}}$	$\mathbf{8.774\times10^{1}}$	1
	MPA	4.997×10^{2}	1.940×10^{1}	3		MPA	2.225×10^{3}	1.080×10^{2}	2
	NMPA	5.246×10^{2}	3.145×10^{1}	6		NMPA	2.285×10^{3}	1.247×10^{2}	3
	QQLMPA	5.146×10^{2}	2.398×10^{1}	5		QQLMPA	2.394×10^{3}	8.926×10^{1}	6
	PSO_sono	5.805×10^{2}	5.618×10^{1}	7		PSO_sono	2.389×10^{3}	1.250×10^{2}	5
	RDPSO	5.066×10^{2}	4.683×10^{1}	4		RDPSO	2.610×10^{3}	2.954×10^{2}	7
	LMHHO	1.447×10^{4}	2.742×10^{3}	8		LMHHO	3.214×10^{3}	1.530×10^{2}	8
	SHADE	4.846×10^{2}	$\mathbf{1.882\times10^{1}}$	2		SHADE	2.339×10^{3}	1.380×10^{2}	4
F_4	PSMPA	4.301×10^{3}	8.085×10^{2}	2	F_{10}	PSMPA	$\mathbf{2.350\times10^{3}}$	1.248×10^{1}	1
	MPA	$\mathbf{4.300\times10^{3}}$	6.234×10^{2}	1		MPA	2.354×10^{3}	6.429×10^{1}	2
	NMPA	4.799×10^{3}	6.658×10^{2}	4		NMPA	2.389×10^{3}	7.870×10^{1}	5
	QQLMPA	6.358×10^{3}	8.571×10^{2}	6		QQLMPA	2.423×10^{3}	4.602×10^{1}	6
	PSO_sono	5.294×10^{3}	7.823×10^{2}	5		PSO_sono	2.362×10^{3}	1.535×10^{1}	3
	RDPSO	9.179×10^{3}	$\mathbf{3.448\times10^{2}}$	8		RDPSO	2.371×10^{3}	2.892×10^{1}	7
	LMHHO	8.436×10^{3}	4.809×10^{2}	7		LMHHO	2.797×10^{3}	6.066×10^{1}	8
	SHADE	4.472×10^{3}	3.837×10^{2}	3		SHADE	2.369×10^{3}	$\mathbf{1.162\times10^{1}}$	4
F_5	PSMPA	$\mathbf{8.999\times10^{3}}$	$\mathbf{1.290\times10^{4}}$	1	F_{11}	PSMPA	$\mathbf{2.301\times10^{3}}$	$\mathbf{1.185\times10^{0}}$	1
	MPA	4.662×10^{5}	2.080×10^{6}	3		MPA	2.309×10^{3}	4.311×10^{0}	2
	NMPA	2.061×10^{6}	1.878×10^{6}	5		NMPA	2.371×10^{3}	4.826×10^{1}	4
	QQLMPA	1.142×10^{6}	1.613×10^{6}	4		QQLMPA	2.337×10^{3}	2.332×10^{1}	3
	PSO_sono	6.307×10^{6}	1.437×10^{7}	7		PSO_sono	2.412×10^{3}	4.995×10^{2}	5
	RDPSO	5.139×10^{6}	4.596×10^{6}	6		RDPSO	1.009×10^{4}	1.529×10^{3}	8
	LMHHO	1.534×10^{10}	3.777×10^{9}	8		LMHHO	1.004×10^{4}	6.296×10^{2}	7
	SHADE	4.523×10^{4}	4.530×10^{4}	2		SHADE	3.638×10^{3}	1.798×10^{3}	6
F_6	PSMPA	$\mathbf{1.980\times10^{3}}$	$\mathbf{3.467\times10^{2}}$	1	F_{12}	PSMPA	$\mathbf{3.111\times10^{3}}$	3.266×10^{1}	1
	MPA	3.731×10^{3}	6.768×10^{2}	3		MPA	3.236×10^{3}	2.136×10^{1}	3
	NMPA	3.257×10^{3}	7.251×10^{2}	2		NMPA	3.266×10^{3}	2.716×10^{1}	5
	QQLMPA	4.737×10^{3}	1.210×10^{3}	5		QQLMPA	3.257×10^{3}	2.528×10^{1}	4
	PSO_sono	2.102×10^{4}	1.896×10^{4}	6		PSO_sono	3.315×10^{3}	5.774×10^{1}	7
	RDPSO	2.334×10^{4}	3.136×10^{4}	7		RDPSO	3.299×10^{3}	$\mathbf{3.860\times10^{0}}$	6
	LMHHO	9.444×10^{9}	4.915×10^{9}	8		LMHHO	7.194×10^{3}	6.037×10^{2}	8
	SHADE	4.307×10^{3}	6.049×10^{3}	4		SHADE	3.219×10^{3}	4.205×10^{1}	2

注: 黑体数据表示最佳结果。

性组合问题时，相较于其他对比算法，拥有更优的局部和全局搜索能力及鲁棒性。

综上所述，PSMPA 在单峰、多峰、混合及组合函数仿真实验中均表现出色，证明了其在不同类型优化问题上的卓越性能和鲁棒性。

3.2.2 算法优化求解的分布特征分析

为分析对比各算法所求解的分布特性，图 4 展示了根据算法 30 次独立求解的结果绘制的箱形图。每种算法每个框的中心标记为 30 次求解函数结果的中位数，框的底部和框的顶部边缘分别为一、三等分点，符号“+”为不在箱子内的坏值。

从图 4 分析可得，在函数 F_1～F_3 中，PSMPA 在求解测试函数时没有异常值，说明 PSMPA 拥有出色的稳定性，并且所求解分布集中。在函数 F_4 中，PSMPA 相比除 QQLMPA 外的其他算法箱体和分布范围较大，但中位数更小，表明 PSMPA 的解的质量相对更优。在函数 F_5～F_7 中，PSMPA 表现稳定，解的中位数更小，进一步证明了解的质量和稳定

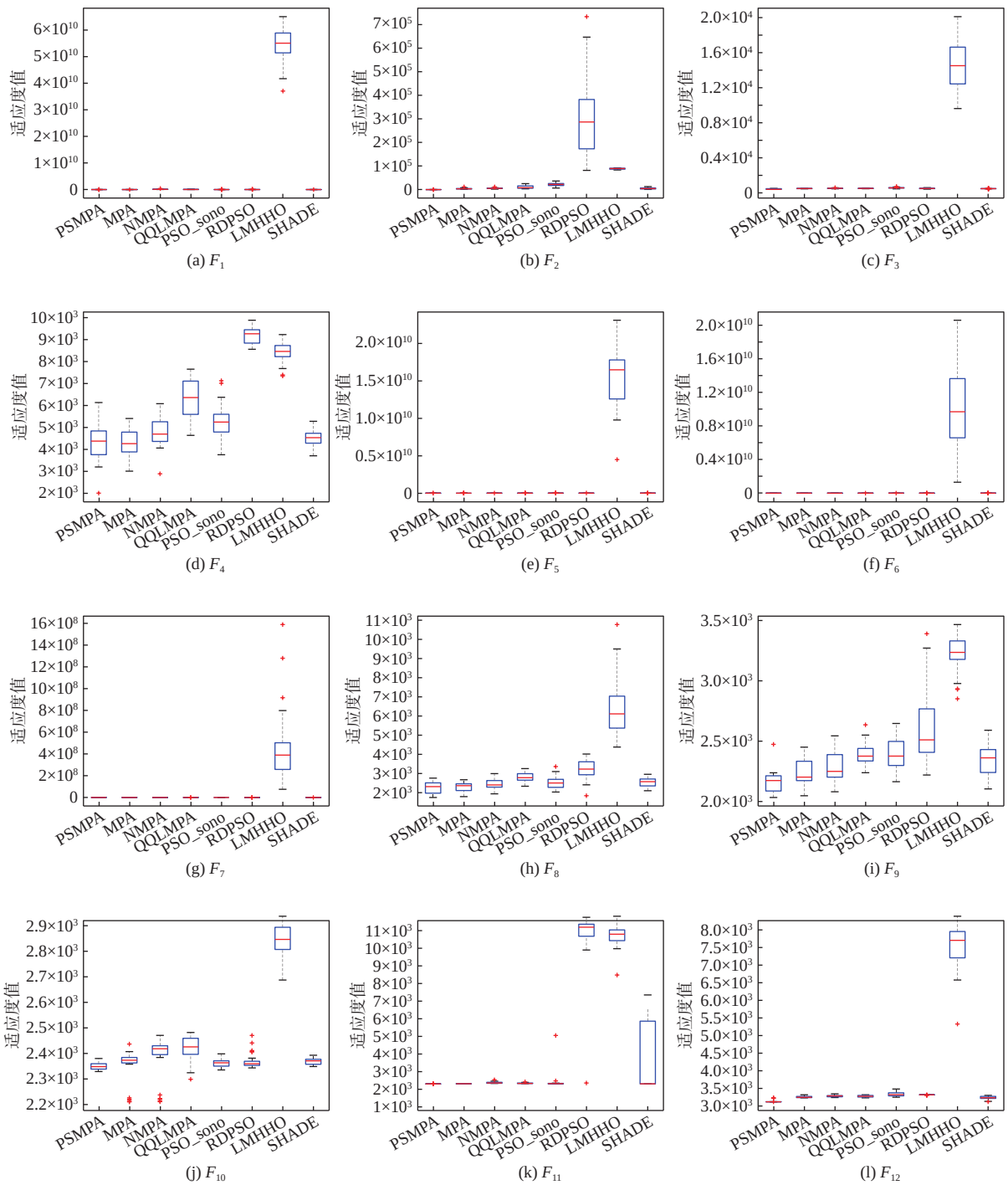


图 4 近期变体算法测试函数箱式统计图

Fig. 4 Box statistics plots of test function for recent variants algorithms

性。对于函数 F_8, 虽然 PSMPA 箱体相较于其他算法略长, 但其中位数依然较小, 说明 PSMPA 的解质量依然较优。在函数 F_9～F_{12} 中, PSMPA 均表现出最小的箱体和最低的中位数, 显示出在处理这些复杂问题时的高质量解和稳定性。

综上所述, PSMPA 在各类测试函数上的综合表现均优于其他算法, 不仅能够在少异常值的情况下稳定运行, 还在大多数测试函数上获得更优的解, 显示出其在解质量、鲁棒性和稳定性方面的显著优势。

3.2.3　算法收敛精度评价

为展示算法在各个函数上的表现, 将适应度值以 10 为底的对数计量的适应度值来作为每个算法收敛精度的评价标准。图 5 为 8 种算法在每个测试函数所展现出的进化曲线和收敛精度, 展示了逃离局部最优的能力。

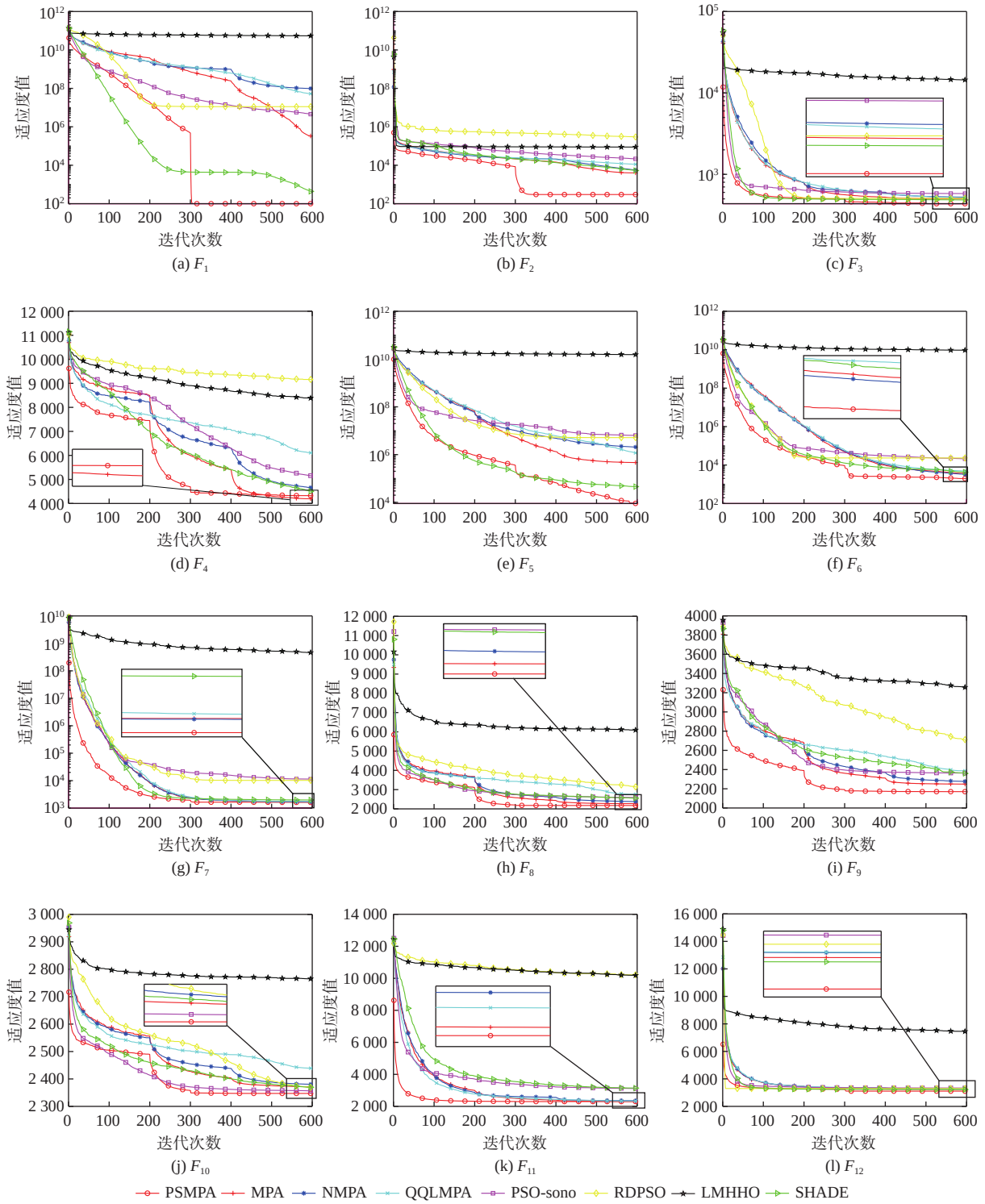


图 5　近期变体算法测试函数迭代曲线

Fig. 5　Iteration curves of test function for recent variants algorithms

在单峰测试函数 F_1 中，PSMPA 的收敛速度明显快于其他对比算法，并且稳定性最高，这表明 PSMPA 在局部开发能力上相较于其他算法具有显著优势，能快速找到全局最优解。混合搜索算子的嵌入，使得算法具有优秀的局部开发能力。

在多峰测试函数 F_2～F_4 中，PSMPA 在收敛速度和收敛精度均优于其他算法，并且在整个迭代过程中保持优秀的稳定性，排名第一，证明了 PSMPA 不仅具有强大的全局搜索能力，还能够有效保持种群多样性，避免陷入局部最优解。由于 SCOBL 机制的引入，提升了 PSMPA 在全局搜索能力，同时保持了种群多样性。

在混合测试函数 F_5～F_8 中，PSMPA 仍然保持着最优收敛速度和收敛精度，与其他对比算法相比，PSMPA 在迭代后期没有陷入局部最优解，证明了算法在复杂问题上的搜索能力、稳定性和跳出局部最优能力。在前 300 次迭代中，PSMPA 能够快速逼近全局最优值，且在迭代后期没有陷入停滞期，继续开发。竞争学习策略的引入，使每个解通过直接竞争和学习机制不断寻求改进，逐步接近最优解的高适应度特征，从而推动整个种群解向更优秀的解发展，有效促进了算法的快速收敛及提高解的质量。

在组合测试函数 F_9～F_{12} 中，PSMPA 在函数 F_9 中，收敛速度和收敛精度遥遥领先其他对比算法。在函数 F_{10} 和 F_{12} 中，虽然在迭代前 200 次收敛速度和收敛精度未达到最优，但在迭代 300 次时，收敛速度和收敛精度达到第一。这也是因为竞争学习策略和 SCOBL 策略提升了算法解的质量和收敛速度。混合搜索算子使得算法在迭代 300 次时也并未停滞，继续进行局部搜索，最终达到最优解。在函数 F_{11} 中，PSMPA 在仅 200 次迭代就达到了最优解，收敛速度远超其他对比算法，最终收敛精度也保持第一。

3.2.4 Wilcoxon 结果分析

根据平均值比较算法性能是不充分的，为避免测试中的偶然性，采用 Wilcoxon 秩和检验对 PSMPA 在 12 个测试函数 30 维下进行验证，通过统计检验判断与其他算法是否具有显著性改进优势。当 p<0.05 时，2 种算法之间存在显著性差异，否则，代表对比算法寻优能力整体相同。NaN 表示 2 个对比算法均达到最佳值时“不适用”，即不能进行显著性判断。表 5 总结了显著水平为 0.05 下的秩和检验结果，R 为显著性判断结果，“+”、“−”、“=”分别表示 PSMPA 性能优于、劣于、相当于对比算法。

表 5 Wilcoxon 秩和检验结果

Table 5 Wilcoxon rank sum test results

函数	MPA		NMPA		QQLMPA		PSO_sono		RDPSO		LMHHO		SHADE	
	p	R	p	R	p	R	p	R	p	R	p	R	p	R
F_1	3.02×10^{-11}	+	3.02×10^{-11}	+	3.02×10^{-11}	+	3.02×10^{-11}	+	3.02×10^{-11}	+	3.02×10^{-11}	+	3.69×10^{-11}	+
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}	+
F_3	2.00×10^{-6}	+	9.26×10^{-9}	+	5.09×10^{-8}	+	1.96×10^{-10}	+	2.15×10^{-6}	+	3.02×10^{-11}	+	1.25×10^{-4}	+
F_4	8.65×10^{-1}	−	1.56×10^{-2}	+	6.72×10^{-10}	+	2.28×10^{-5}	+	3.02×10^{-11}	+	3.02×10^{-11}	+	4.73×10^{-1}	−
F_5	3.69×10^{-11}	+	3.02×10^{-11}	+	3.34×10^{-11}	+	3.69×10^{-11}	+	3.02×10^{-11}	+	3.02×10^{-11}	+	3.35×10^{-8}	+
F_6	6.07×10^{-11}	+	9.76×10^{-10}	+	4.08×10^{-11}	+	3.02×10^{-11}	+	1.09×10^{-10}	+	3.02×10^{-11}	+	1.29×10^{-6}	+
F_7	4.11×10^{-7}	+	2.15×10^{-6}	+	1.60×10^{-7}	+	3.02×10^{-11}	+	4.50×10^{-11}	+	3.02×10^{-11}	+	3.16×10^{-10}	+
F_8	5.79×10^{-1}	−	2.61×10^{-2}	+	1.56×10^{-8}	+	4.64×10^{-3}	+	9.26×10^{-9}	+	3.02×10^{-11}	+	1.17×10^{-3}	+
F_9	3.78×10^{-2}	+	9.21×10^{-5}	+	3.47×10^{-10}	+	5.97×10^{-9}	+	1.78×10^{-10}	+	3.02×10^{-11}	+	7.22×10^{-6}	+
F_{10}	9.21×10^{-5}	+	9.51×10^{-6}	+	1.43×10^{-8}	+	2.75×10^{-3}	+	2.25×10^{-4}	+	3.02×10^{-11}	+	2.00×10^{-6}	+
F_{11}	2.37×10^{-10}	+	3.02×10^{-11}	+	3.02×10^{-11}	+	8.10×10^{-10}	+	3.02×10^{-11}	+	3.02×10^{-11}	+	5.61×10^{-5}	+
F_{12}	4.50×10^{-11}	+	3.02×10^{-11}	+	3.34×10^{-11}	+	3.02×10^{-11}	+	3.02×10^{-11}	+	3.02×10^{-11}	+	6.72×10^{-10}	+

从表 5 中可以分析得出，在大部分测试函数上，p 值均小于 0.05，仅在个别测试函数上寻优性能出现性能相近的情况，表明 PSMPA 与其他 7 种算法均存在显著性差异，证明了 PSMPA 的寻优性能显著和算法独特性。

3.3 可扩展性分析

为验证 PSMPA 在高纬函数上的寻优效果，将 12 个 CEC2017 测试函数的变量维度扩展至 50 维和 100 维，并与 3.1 节中其他算法进行比较，算法参数和表 2 保持一致。表 6 展示了各算法处理 50 维度

表 6　可扩展性分析结果统计

Table 6　Statistics on results of scalability analysis

函数	算法	50维		100维		函数	算法	50维		100维	
		平均值	标准差	平均值	标准差			平均值	标准差	平均值	标准差
F_1	PSMPA	**3.528×10^{3}**	**2.586×10^{3}**	**9.475×10^{3}**	**5.908×10^{3}**	F_7	PSMPA	7.431×10^{3}	6.688×10^{3}	**3.781×10^{3}**	**2.112×10^{3}**
	MPA	4.413×10^{7}	2.724×10^{7}	3.621×10^{9}	1.275×10^{9}		MPA	3.761×10^{3}	**1.423×10^{3}**	5.665×10^{4}	3.002×10^{4}
	NMPA	1.688×10^{9}	1.090×10^{9}	2.198×10^{10}	5.534×10^{9}		NMPA	**3.632×10^{3}**	1.651×10^{3}	5.331×10^{6}	3.410×10^{6}
	QQLMPA	1.077×10^{9}	8.709×10^{8}	2.412×10^{10}	1.670×10^{10}		QQLMPA	3.869×10^{3}	8.206×10^{2}	1.088×10^{6}	8.724×10^{5}
	PSO_sono	2.765×10^{8}	3.648×10^{8}	4.326×10^{9}	3.658×10^{9}		PSO_sono	1.437×10^{4}	8.904×10^{3}	3.411×10^{4}	2.164×10^{4}
	RDPSO	4.574×10^{9}	2.362×10^{9}	1.170×10^{11}	2.987×10^{10}		RDPSO	8.495×10^{4}	1.530×10^{5}	1.069×10^{7}	1.212×10^{7}
	LMHHO	1.045×10^{11}	7.368×10^{9}	2.557×10^{11}	7.525×10^{9}		LMHHO	7.694×10^{9}	2.075×10^{9}	2.650×10^{10}	3.869×10^{9}
	SHADE	5.992×10^{3}	8.188×10^{3}	7.043×10^{6}	6.060×10^{6}		SHADE	5.485×10^{3}	4.463×10^{3}	5.522×10^{3}	3.890×10^{3}
F_2	PSMPA	**4.241×10^{2}**	**1.502×10^{2}**	**7.288×10^{4}**	1.155×10^{4}	F_8	PSMPA	**2.687×10^{3}**	3.133×10^{2}	**5.093×10^{3}**	6.557×10^{2}
	MPA	4.229×10^{4}	9.827×10^{3}	2.149×10^{5}	2.338×10^{4}		MPA	3.067×10^{3}	3.994×10^{2}	5.774×10^{3}	8.139×10^{2}
	NMPA	4.673×10^{4}	1.010×10^{4}	2.255×10^{5}	1.743×10^{4}		NMPA	3.246×10^{3}	3.650×10^{2}	6.925×10^{3}	6.222×10^{2}
	QQLMPA	6.874×10^{4}	1.438×10^{4}	2.491×10^{5}	2.292×10^{4}		QQLMPA	3.840×10^{3}	4.836×10^{2}	9.162×10^{3}	1.116×10^{3}
	PSO_sono	1.217×10^{5}	3.301×10^{4}	3.211×10^{5}	4.582×10^{4}		PSO_sono	3.325×10^{3}	3.807×10^{2}	6.231×10^{3}	7.215×10^{2}
	RDPSO	1.238×10^{6}	6.525×10^{5}	7.936×10^{6}	5.217×10^{6}		RDPSO	5.450×10^{3}	7.098×10^{2}	1.226×10^{4}	7.306×10^{2}
	LMHHO	2.663×10^{5}	3.055×10^{4}	3.620×10^{5}	**8.152×10^{3}**		LMHHO	9.951×10^{3}	2.149×10^{3}	2.554×10^{4}	2.986×10^{3}
	SHADE	6.177×10^{4}	4.587×10^{4}	3.169×10^{5}	6.728×10^{4}		SHADE	3.515×10^{3}	**2.778×10^{2}**	7.377×10^{3}	**4.900×10^{2}**
F_3	PSMPA	**4.408×10^{2}**	5.482×10^{1}	**5.577×10^{2}**	1.117×10^{2}	F_9	PSMPA	**2.609×10^{3}**	3.124×10^{2}	**4.591×10^{3}**	4.973×10^{2}
	MPA	6.316×10^{2}	5.690×10^{1}	1.438×10^{3}	1.917×10^{2}		MPA	2.752×10^{3}	**1.964×10^{2}**	4.700×10^{3}	3.455×10^{2}
	NMPA	7.913×10^{2}	8.247×10^{1}	3.065×10^{3}	6.870×10^{2}		NMPA	2.833×10^{3}	2.725×10^{2}	5.104×10^{3}	3.625×10^{2}
	QQLMPA	7.409×10^{2}	1.047×10^{2}	2.598×10^{3}	1.318×10^{3}		QQLMPA	3.306×10^{3}	1.971×10^{2}	6.292×10^{3}	5.526×10^{2}
	PSO_sono	8.578×10^{2}	1.770×10^{2}	2.214×10^{3}	8.400×10^{2}		PSO_sono	3.117×10^{3}	3.186×10^{2}	5.186×10^{3}	5.874×10^{2}
	RDPSO	1.051×10^{3}	2.280×10^{2}	1.462×10^{4}	4.245×10^{3}		RDPSO	4.284×10^{3}	3.812×10^{2}	8.342×10^{3}	4.666×10^{2}
	LMHHO	3.500×10^{4}	4.117×10^{3}	1.030×10^{5}	1.041×10^{4}		LMHHO	4.476×10^{3}	2.627×10^{2}	8.021×10^{3}	**3.604×10^{2}**
	SHADE	5.523×10^{2}	**5.424×10^{1}**	7.432×10^{2}	**5.248×10^{1}**		SHADE	3.299×10^{3}	2.035×10^{2}	6.318×10^{3}	3.726×10^{2}
F_4	PSMPA	**6.753×10^{3}**	9.441×10^{2}	**1.408×10^{4}**	1.183×10^{3}	F_{10}	PSMPA	**2.420×10^{3}**	2.410×10^{1}	**2.662×10^{3}**	**5.785×10^{1}**
	MPA	6.926×10^{3}	5.894×10^{2}	1.618×10^{4}	1.381×10^{3}		MPA	2.493×10^{3}	3.603×10^{1}	2.898×10^{3}	5.965×10^{1}
	NMPA	8.805×10^{3}	9.060×10^{2}	2.183×10^{4}	1.621×10^{3}		NMPA	2.590×10^{3}	3.761×10^{1}	3.217×10^{3}	9.728×10^{1}
	QQLMPA	1.155×10^{4}	1.505×10^{3}	2.766×10^{4}	2.135×10^{3}		QQLMPA	2.631×10^{3}	6.070×10^{1}	3.257×10^{3}	1.096×10^{2}
	PSO_sono	9.186×10^{3}	1.209×10^{3}	2.356×10^{4}	3.102×10^{3}		PSO_sono	2.450×10^{3}	3.485×10^{1}	3.034×10^{3}	1.020×10^{2}
	RDPSO	1.593×10^{4}	5.371×10^{2}	3.413×10^{4}	**7.713×10^{2}**		RDPSO	2.495×10^{3}	4.477×10^{1}	3.286×10^{3}	1.689×10^{2}
	LMHHO	1.516×10^{4}	8.697×10^{2}	3.233×10^{4}	1.057×10^{3}		LMHHO	3.326×10^{3}	1.161×10^{2}	5.053×10^{3}	1.854×10^{2}
	SHADE	8.898×10^{3}	**4.419×10^{2}**	2.372×10^{4}	9.486×10^{2}		SHADE	2.504×10^{3}	**1.847×10^{1}**	3.120×10^{3}	6.968×10^{1}
F_5	PSMPA	**2.029×10^{5}**	**4.414×10^{5}**	**5.435×10^{6}**	1.679×10^{7}	F_{11}	PSMPA	**4.338×10^{3}**	2.976×10^{3}	**1.028×10^{4}**	8.180×10^{3}
	MPA	2.338×10^{7}	1.704×10^{7}	5.443×10^{8}	2.520×10^{8}		MPA	6.153×10^{3}	3.178×10^{0}	1.944×10^{4}	1.412×10^{3}
	NMPA	1.157×10^{8}	6.637×10^{7}	2.624×10^{9}	1.192×10^{9}		NMPA	7.298×10^{3}	4.065×10^{3}	2.386×10^{4}	5.707×10^{3}
	QQLMPA	7.038×10^{7}	5.181×10^{7}	2.307×10^{9}	1.730×10^{9}		QQLMPA	1.074×10^{4}	4.017×10^{3}	2.900×10^{4}	4.624×10^{3}
	PSO_sono	5.911×10^{7}	6.068×10^{7}	7.337×10^{8}	4.766×10^{8}		PSO_sono	1.077×10^{4}	1.583×10^{3}	2.573×10^{4}	2.831×10^{3}
	RDPSO	2.658×10^{8}	2.703×10^{8}	1.221×10^{10}	4.671×10^{9}		RDPSO	1.764×10^{4}	**3.756×10^{2}**	3.623×10^{4}	6.093×10^{2}
	LMHHO	8.692×10^{10}	1.581×10^{10}	1.991×10^{11}	1.682×10^{10}		LMHHO	1.742×10^{4}	6.857×10^{2}	3.490×10^{4}	**9.788×10^{2}**
	SHADE	1.191×10^{6}	1.043×10^{6}	1.574×10^{7}	**1.151×10^{7}**		SHADE	1.034×10^{4}	1.624×10^{3}	2.612×10^{4}	1.247×10^{3}
F_6	PSMPA	**4.764×10^{3}**	**3.238×10^{3}**	**9.421×10^{3}**	**5.318×10^{3}**	F_{12}	PSMPA	**3.296×10^{3}**	3.243×10^{1}	3.383×10^{3}	5.182×10^{1}
	MPA	5.149×10^{4}	3.775×10^{4}	3.225×10^{5}	4.482×10^{5}		MPA	3.423×10^{3}	4.764×10^{1}	4.685×10^{3}	5.006×10^{2}
	NMPA	8.226×10^{4}	7.141×10^{4}	8.662×10^{7}	1.232×10^{8}		NMPA	3.645×10^{3}	1.883×10^{2}	6.396×10^{3}	8.676×10^{2}
	QQLMPA	8.509×10^{4}	5.927×10^{4}	3.817×10^{7}	1.990×10^{7}		QQLMPA	3.531×10^{3}	1.308×10^{2}	6.539×10^{3}	1.858×10^{3}
	PSO_sono	4.731×10^{4}	2.871×10^{4}	1.314×10^{6}	4.375×10^{6}		PSO_sono	3.834×10^{3}	3.138×10^{2}	7.005×10^{3}	1.543×10^{3}
	RDPSO	1.430×10^{6}	1.291×10^{6}	1.403×10^{8}	1.120×10^{8}		RDPSO	3.300×10^{3}	**1.853×10^{0}**	**3.300×10^{3}**	**1.297×10^{-4}**
	LMHHO	4.842×10^{10}	1.538×10^{10}	4.720×10^{10}	5.695×10^{9}		LMHHO	1.336×10^{4}	9.151×10^{2}	2.960×10^{4}	1.694×10^{3}
	SHADE	1.299×10^{4}	1.015×10^{4}	1.238×10^{4}	8.830×10^{3}		SHADE	3.303×10^{3}	1.698×10^{1}	3.519×10^{3}	4.364×10^{1}

注: 黑体数据表示最佳结果。

和100维度的测试函数结果。

从表6的分析结果可以看出，无论在50维还是100维的测试函数中，PSMPA在大多数情况下均表现出色，其平均值在大多数测试函数中排名第一。同时，PSMPA的标准差在大多数测试函数中也优于经典MPA和2种MPA变体算法。这表明PSMPA在解决高维优化问题时，具有更高的稳定性和精确性，能有效避免陷入局部最优解，从而提升整体优化性能。综上所述，PSMPA在高维复杂问题中的优越性得到了充分验证。

4 PSMPA优化太阳能光伏模型参数

在光伏系统中，光伏模型的参数是用来描述光伏电池和组件的性能和非线性行为的关键要素，准确识别这些参数对于仿真、评估和控制光伏系统至关重要。

应用中主要有2种广泛使用的模型来描述了太阳能电池和光伏组件的电流-电压特性，最常用的是单二极管模型和双二极管模型。

将PSMPA每个实验独立运行30次，其中，种群数为30，迭代次数为1 000次，统计实验的最小值、最大值、平均值、标准差及均方根误差(root mean square error, RMSE)，并与第3节中其他算法进行比较，所有算法参数与表2一致。

4.1 单二极管模型

单二极管模型[26]因其简单准确而被广泛用于表示太阳能电池的静态特性。模型包括一个与二极管并联的电流源、一个表示泄漏电流的分流电阻器和一个表示负载电流损失的串联电阻器，该模型的输出电流计算如下:

$$I_{\mathrm{L}}=I_{\mathrm{ph}}-I_{\mathrm{sd}}\left[\exp\left(\frac{q(V_{\mathrm{L}}+R_{\mathrm{S}}I_{\mathrm{L}})}{nkT_{\mathrm{env}}}\right)-1\right]-\frac{V_{\mathrm{L}}+R_{\mathrm{S}}I_{\mathrm{L}}}{R_{\mathrm{sh}}} \tag{18}$$

式中：I_{L}为太阳能电池的输出电流;I_{ph}为太阳能电池产生的总电流;I_{sd}为二极管的反向饱和电流;R_{S}和R_{sh}分别为串联电阻和并联电阻;V_{L}为电池输出电压;n为二极管的理想因子;k为玻尔兹曼常数，取值为1.38×10^{-23} J/K;q为电子电荷，取值为1.6×10^{-18} C;T_{env}为环境温度。

综上，单二极管模型需要估计5个未知参数$(I_{\mathrm{ph}},I_{\mathrm{sd}},R_{\mathrm{S}},R_{\mathrm{sh}},n)$。

4.2 双二级管模型

双二极管模型[26]是通过考虑耗尽区再组合电流损失的影响而开发的。在该模型中，有2个与电流源并联的二极管和1个分流电阻。根据双二极管的等效电路，输出电流可描述如下：

$$I_{\mathrm{L}}=I_{\mathrm{ph}}-I_{\mathrm{d1}}-I_{\mathrm{d2}}-I_{\mathrm{sh}}=I_{\mathrm{ph}}-I_{\mathrm{sd1}}\left[\exp\frac{q(V_{\mathrm{L}}+R_{\mathrm{S}}I_{\mathrm{L}})}{n_1kT_{\mathrm{env}}}-1\right]-I_{\mathrm{sd2}}\left[\exp\frac{q(V_{\mathrm{L}}+R_{\mathrm{S}}I_{\mathrm{L}})}{n_2kT_{\mathrm{env}}}-1\right]-\frac{V_{\mathrm{L}}+R_{\mathrm{S}}I_{\mathrm{L}}}{R_{\mathrm{sh}}} \tag{19}$$

式中：I_{d1}、I_{d2}、I_{sh}为光伏电池的内部能量损耗项，与光生电流I_{ph}共同决定实际输出电流I_{L};I_{sd1}和I_{sd2}分别为扩散电流和饱和电流;n_1和n_2分别为扩散和复合二极管的理想因子。

对于双二极管模型，需要识别7个未知参数$(I_{\mathrm{ph}},I_{\mathrm{sd1}},I_{\mathrm{sd2}},R_{\mathrm{S}},R_{\mathrm{sh}},n_1,n_2)$，才能获得太阳能电池的实际行为。

4.3 光伏组件模型

光伏组件模型[27]由若干串联和或并联的太阳能电池组成的单二级管光伏组件模型，输出电流可以描述如下：

$$\frac{I_{\mathrm{L}}}{N_{\mathrm{P}}}=I_{\mathrm{ph}}-I_{\mathrm{sd}}\left[\exp\left(\frac{q(V_{\mathrm{L}}/N_{\mathrm{S}}+R_{\mathrm{S}}I_{\mathrm{L}}/N_{\mathrm{P}})}{nkT_{\mathrm{env}}}\right)-1\right]-\frac{V_{\mathrm{L}}/N_{\mathrm{S}}+R_{\mathrm{S}}I_{\mathrm{L}}/N_{\mathrm{P}}}{R_{\mathrm{sh}}} \tag{20}$$

式中：N_{P}为并联的太阳能电池数量;N_{S}为串联的太阳能电池数量。

对于光伏组件模型，需要识别5个未知参数$(I_{\mathrm{ph}},I_{\mathrm{sd}},R_{\mathrm{S}},R_{\mathrm{sh}},n)$。

4.4 目标函数

从上述3个模型中可以得出，单二极管模型需要识别5个参数：$I_{\mathrm{ph}},I_{\mathrm{sd}},R_{\mathrm{S}},R_{\mathrm{sh}},n$;双二极管模型需要识别7个参数：$I_{\mathrm{ph}},I_{\mathrm{sd1}},I_{\mathrm{sd2}},R_{\mathrm{S}},R_{\mathrm{sh}},n_1,n_2$;光伏组件模型需要识别5个参数：$I_{\mathrm{ph}},I_{\mathrm{sd}},R_{\mathrm{S}},R_{\mathrm{sh}},n$。在实验中，将参数提取问题转化为非线性最优化问题，可建立优化目标函数。方程定义的RMSE被用作量化实验和模拟电流数据之间总体差异的目标函数，该目标函数在文献中得到了广泛应用，RMSE越小，所提取的参数越准确。优化目标函数可描述为

$$\mathrm{RMSE}(\boldsymbol{x})=\sqrt{\frac{1}{N}\sum_{k=1}^{N}f_k(V_{\mathrm{L}},I_{\mathrm{L}},\boldsymbol{x})^2} \tag{21}$$

式中：$\boldsymbol{x}$为未知参数组成的解向量;N为实验数据的数量;$f_k(V_{\mathrm{L}},I_{\mathrm{L}},\boldsymbol{x})$为仿真电流与实验电流的差值。

实验太阳能电池和太阳能模块的数据均来自33℃、57 mm直径的商业硅R.T.C法国太阳能电池和45℃的36个串联多晶硅电池组成的名为Photowatt-PWP201太阳能模块，所有待识别的参数范围如表7所示。

表 7　待识别参数统计

Table 7　Statistics of parameters to be identified

模型	I_{ph}/A		I_{sd}, I_{sd1}, I_{sd2}/μA		R_S/Ω		R_{sh}/Ω		n, n_1, n_2	
	下限	上限	下限	上限	下限	上限	下限	上限	下限	上限
单二极管/双二极管	0	1	0	1	0	0.5	0	100	1	2
光伏组件	0	2	0	50	0	2	0	2 000	1	50

表 8～表 11 统计了所有算法在单二极管、双二极管和光伏组件模型的测试结果。

在单二极管模型中, 与其他算法相比, PSMPA、PSO_sono 和 SHADE 表现出了优秀的寻优效果, 寻优效果并列第一, 取得了最小的 RMSE 值($9.860\,2\times10^{-4}$)。而 PSMPA 取得的最大值、平均值和标准差均仅次于 SHADE, 均优于经典 MPA 和 2 种 MPA 变体算法。

表 8　不同算法在单二极管模型求解结果

Table 8　Solution results of different algorithms in single diode model

算法	I_{ph}/A	I_{sd}/μA	R_S/Ω	R_{sh}/Ω	n	RMSE
PSMPA	0.760 9	$9.259\,6\times10^{-7}$	0.031 7	92.245 0	1.595 51	$9.860\,2\times10^{-4}$
MPA	0.762 7	$9.223\,0\times10^{-7}$	0.031 5	56.096 2	1.595 1	$9.862\,2\times10^{-4}$
NMPA	0.766 5	$7.900\,0\times10^{-7}$	0.028 5	9.930 6	1.576 9	$1.457\,6\times10^{-3}$
QQLMPA	0.763 2	$2.892\,6\times10^{-7}$	0.035 7	23.853 0	1.475 9	$1.492\,6\times10^{-3}$
PSO_snono	0.760 8	$3.230\,2\times10^{-7}$	0.036 4	53.718 5	1.481 19	$9.860\,2\times10^{-4}$
RDPSO	0.761 1	$3.314\,4\times10^{-7}$	0.036 3	53.417 2	1.483 74	$1.009\,4\times10^{-3}$
LMHHO	0.762 7	$9.223\,0\times10^{-7}$	0.031 5	56.096 2	1.595 1	$2.804\,2\times10^{-3}$
SHADE	0.760 3	$5.022\,2\times10^{-7}$	0.034 7	81.207 8	1.526 85	$9.860\,2\times10^{-4}$

表 9　不同算法在双二极管模型求解结果

Table 9　Solution results of different algorithms in double diode model

算法	I_{ph}/A	I_{sd1}/μA	R_S/Ω	R_{sh}/Ω	n_1	I_{sd2}/μA	n_2	RMSE
PSMPA	0.757 5	$3.331\,2\times10^{-7}$	0.034 1	41.923 7	1.497 5	$5.435\,5\times10^{-7}$	1.785 1	$9.826\,1\times10^{-4}$
MPA	0.760 4	$6.684\,8\times10^{-7}$	0.033 1	89.211 6	1.558 7	$7.114\,8\times10^{-8}$	2.000 0	$1.043\,2\times10^{-3}$
NMPA	0.773 4	$1.000\,0\times10^{-6}$	0.026 8	11.017 5	1.616 1	$5.619\,4\times10^{-7}$	1.907 9	$1.393\,3\times10^{-3}$
QQLMPA	0.757 5	$3.331\,2\times10^{-7}$	0.034 1	41.923 7	1.497 5	$5.435\,5\times10^{-7}$	1.785 1	$1.446\,7\times10^{-3}$
PSO_snono	0.760 4	$8.700\,6\times10^{-7}$	0.036 5	63.969 2	1.985 8	$2.291\,0\times10^{-7}$	1.453 6	$1.023\,2\times10^{-3}$
RDPSO	0.760 7	$4.974\,9\times10^{-7}$	0.034 7	78.173 9	1.525 8	$1.000\,0\times10^{-9}$	2.000 0	$1.366\,7\times10^{-3}$
LMHHO	0.761 8	$6.757\,0\times10^{-7}$	0.031 9	32.774 6	1.573 5	$8.303\,2\times10^{-7}$	1.914 0	$7.580\,6\times10^{-3}$
SHADE	0.762 7	$6.085\,7\times10^{-7}$	0.032 0	48.466 2	1.560 9	$6.329\,0\times10^{-7}$	1.847 8	$9.824\,8\times10^{-4}$

表 10　不同算法在光伏组件模型求解结果

Table 10　Solution results of different algorithms in PV module models

算法	I_{ph}/A	I_{sd}/μA	R_S/Ω	R_{sh}/Ω	n	RMSE
PSMPA	1.039 2	$4.897\,9\times10^{-6}$	1.156 3	766.771 4	50.000 0	$2.425\,1\times10^{-3}$
MPA	1.041 4	$4.974\,7\times10^{-6}$	1.166 1	1 676.965 4	50.000 0	$2.454\,6\times10^{-3}$
NMPA	1.022 2	$4.872\,4\times10^{-6}$	1.159 4	2 000.000 0	50.000 0	$2.690\,6\times10^{-3}$
QQLMPA	1.026 5	$3.534\,9\times10^{-6}$	1.235 1	1 430.521 2	48.694 1	$2.602\,7\times10^{-3}$
PSO_snono	1.030 5	$3.482\,3\times10^{-6}$	1.201 3	981.982 2	48.642 9	$2.425\,1\times10^{-3}$
RDPSO	1.031 1	$4.302\,6\times10^{-6}$	1.174 7	993.081 1	49.475 5	$2.568\,4\times10^{-3}$
LMHHO	1.028 7	$4.854\,8\times10^{-6}$	1.166 6	1 557.999 4	49.948 7	$1.101\,3\times10^{-2}$
SHADE	1.036 1	$4.851\,1\times10^{-6}$	1.130 3	622.467 1	50.000 0	$2.425\,1\times10^{-3}$

在双二极管模型中, PSMPA(RMSE 值为 $9.826\,1\times10^{-4}$) 寻优效果仅次于 SHADE(RMSE 值为 $9.824\,8\times10^{-4}$), 并且最大值、最小值、平均值和标准差均仅次于 SHADE, 其余算法的参数识别结构均不够

表 11 不同算法在光伏组件模型结果统计

Table 11 Results statistics of different algorithms in PV module model

模型	算法	RMSE			
		最小值	平均值	最大值	标准差
单二极管模型	PSMPA	**$9.860\,2\times10^{-4}$**	$1.195\,9\times10^{-3}$	$1.979\,1\times10^{-3}$	$2.978\,7\times10^{-4}$
	MPA	$9.862\,2\times10^{-4}$	$3.077\,2\times10^{-3}$	$8.655\,1\times10^{-3}$	$1.966\,1\times10^{-3}$
	NMPA	$1.457\,6\times10^{-3}$	$4.583\,4\times10^{-3}$	$1.366\,8\times10^{-2}$	$2.973\,2\times10^{-3}$
	QQLMPA	$1.492\,6\times10^{-3}$	$1.441\,2\times10^{-2}$	$4.114\,7\times10^{-2}$	$1.313\,4\times10^{-2}$
	PSO_snono	**$9.860\,2\times10^{-4}$**	$1.792\,3\times10^{-3}$	$2.448\,0\times10^{-3}$	$6.451\,6\times10^{-4}$
	RDPSO	$1.009\,4\times10^{-3}$	$1.259\,1\times10^{-2}$	$1.168\,9\times10^{-1}$	$2.258\,3\times10^{-2}$
	LMHHO	$2.804\,2\times10^{-3}$	$3.366\,1\times10^{-2}$	$8.988\,9\times10^{-2}$	$2.361\,5\times10^{-2}$
	SHADE	**$9.860\,2\times10^{-4}$**	**$9.860\,2\times10^{-4}$**	**$9.860\,2\times10^{-4}$**	**$3.443\,1\times10^{-17}$**
双二极管模型	PSMPA	$9.826\,1\times10^{-4}$	$1.262\,1\times10^{-3}$	$4.744\,6\times10^{-3}$	$7.425\,3\times10^{-4}$
	MPA	$1.043\,2\times10^{-3}$	$2.584\,1\times10^{-3}$	$6.468\,4\times10^{-3}$	$1.312\,6\times10^{-3}$
	NMPA	$1.393\,3\times10^{-3}$	$3.480\,7\times10^{-3}$	$8.924\,8\times10^{-3}$	$1.636\,2\times10^{-3}$
	QQLMPA	$1.446\,7\times10^{-3}$	$1.167\,3\times10^{-2}$	$5.420\,1\times10^{-2}$	$1.525\,6\times10^{-2}$
	PSO_snono	$1.023\,2\times10^{-3}$	$3.152\,7\times10^{-3}$	$3.948\,8\times10^{-2}$	$6.890\,2\times10^{-3}$
	RDPSO	$1.366\,7\times10^{-3}$	$6.550\,7\times10^{-3}$	$3.218\,1\times10^{-2}$	$6.834\,3\times10^{-3}$
	LMHHO	$7.580\,6\times10^{-3}$	$3.844\,1\times10^{-2}$	$1.936\,8\times10^{-1}$	$3.414\,6\times10^{-2}$
	SHADE	**$9.824\,8\times10^{-4}$**	**$9.865\,0\times10^{-4}$**	**$1.038\,5\times10^{-3}$**	**$1.124\,9\times10^{-5}$**
光伏组件模型	PSMPA	**$2.425\,1\times10^{-3}$**	**$2.552\,1\times10^{-3}$**	$4.876\,8\times10^{-3}$	$4.465\,0\times10^{-4}$
	MPA	$2.454\,6\times10^{-3}$	$2.685\,9\times10^{-3}$	**$4.141\,3\times10^{-3}$**	**$2.900\,1\times10^{-4}$**
	NMPA	$2.690\,6\times10^{-3}$	$5.322\,3\times10^{-3}$	$1.240\,7\times10^{-2}$	$2.650\,4\times10^{-3}$
	QQLMPA	$2.602\,7\times10^{-3}$	$2.658\,1\times10^{-2}$	$8.739\,9\times10^{-2}$	$2.868\,4\times10^{-2}$
	PSO_snono	$2.425\,1\times10^{-3}$	$2.077\,7\times10^{-2}$	$2.742\,5\times10^{-1}$	$6.890\,4\times10^{-2}$
	RDPSO	$2.568\,4\times10^{-3}$	$8.355\,9\times10^{-3}$	$4.855\,9\times10^{-2}$	$1.188\,5\times10^{-2}$
	LMHHO	$1.101\,3\times10^{-2}$	$2.027\,9\times10^{-1}$	$3.188\,3\times10^{-1}$	$1.156\,3\times10^{-1}$
	SHADE	**$2.425\,1\times10^{-3}$**	$1.148\,6\times10^{-2}$	$2.742\,5\times10^{-1}$	$4.962\,8\times10^{-2}$

注: 黑体数据表示最佳结果。

准确。

在光伏组件模型中, PSMPA 与 SHADE 寻优效果并列第一, 继续展现了最小的 RMSE 值($2.425\,1\times10^{-3}$)。PSMPA 在最小值和平均值中排名第一, 最大值和标准差仅略次于 MPA, PSMPA 寻优效果显著优于其他 MPA 变体算法。

图 6 展示了 30 次独立运行的平均 RMSE 收敛曲线。PSMPA 在单二极管和双二极管模型中的收敛速度仅次于 SHADE, 且明显优于其他对比算法; 在光伏组件模型中表现出最快的收敛速度。

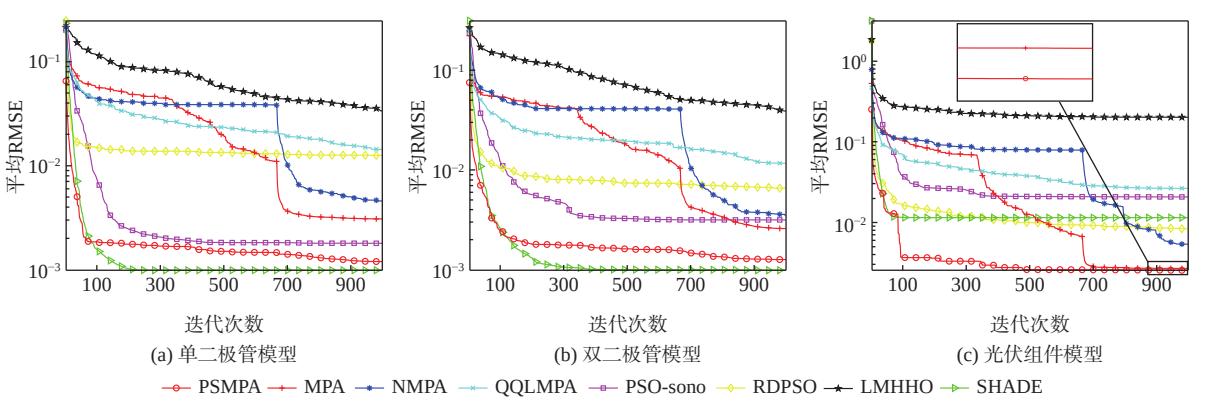


图 6 太阳能光伏模型参数曲线迭代

Fig. 6 Iteration of solar PV model parameter curves

综上所述, 充分证明了 PSMPA 在太阳能光伏模型参数识别设计问题中的卓越性能和优越能力。

5 结　论

1) PSMPA 引入 SCOBL 机制, 避免了重复点和重复评估, 扩展了反向搜索范围, 丰富了种群多样性, 增强了全局搜索能力; 将混合搜索算子嵌入 MPA 的局部搜索中, 显著提升了算法的局部开发能力; 引入海洋捕食者相互学习竞争的行为模式, 提高了解的平均适应度, 进一步加快了算法的收敛速度和提升了收敛精度。

2) 在 12 个 CEC2017 测试函数上对 PSMPA 与其他算法进行了优化求解对比测试, 结果表明, PSMPA 在收敛速度、收敛精度、跳出局部最优及鲁棒性方面均显著优于对比算法, 排名第一; Wilcoxon 秩和检验结果进一步验证了 PSMPA 的独特性和优越性。

3) 在可扩展性分析测试中, PSMPA 收敛精度排名第一, 验证了其在高维问题上的优越性。太阳能光伏参数优化问题进一步验证了 PSMPA 在实际复杂工程应用中的实用性和有效性。

未来的研究工作需进一步考虑 PSMPA 的时间复杂度。在确保算法性能提升的同时, 应进一步优化算法结构, 降低其时间复杂度, 以提升其在实际应用中的效率。

参考文献 (References)

[1] FARAMARZI A, HEIDARINEJAD M, MIRJALILI S, et al. Marine predators algorithm: a nature-inspired metaheuristic[J]. Expert Systems with Applications, 2020, 152: 113377.

[2] HOLLAND J H. Genetic algorithms[J]. Scientific American, 1992, 267(1): 66-73.

[3] KENNEDY J, EBERHART R. Particle swarm optimization[C]// Proceedings of the International Conference on Neural Networks. Piscataway: IEEE Press, 2002: 1942-1948.

[4] YANG X S, DEB S. Cuckoo search: recent advances and applications[J]. Neural Computing and Applications, 2014, 24(1): 169-174.

[5] RAI R, DHAL K G, DAS A, et al. An inclusive survey on marine predators algorithm: variants and applications[J]. Archives of Computational Methods in Engineering, 2023, 30(5): 3133-3172.

[6] HELMI A M, AL-QANESS M A A, DAHOU A, et al. Human activity recognition using marine predators algorithm with deep learning[J]. Future Generation Computer Systems, 2023, 142(C): 340-350.

[7] 马驰, 曾国辉, 黄勃, 等. 融合混沌对立和分组学习的海洋捕食者算法[J]. 计算机工程与应用, 2022, 58(22): 271-283.
MA C, ZENG G H, HUANG B, et al. Marine predator algorithm based on chaotic opposition learning and group learning[J]. Computer Engineering and Applications, 2022, 58(22): 271-283(in Chinese).

[8] 张磊, 刘升, 高文欣, 等. 精英反向黄金正弦海洋捕食者算法[J]. 计算机工程与科学, 2023, 45(2): 355-362.
ZHANG L, LIU S, GAO W X, et al. An elite opposition-based golden sine marine predator algorithm[J]. Computer Engineering & Science, 2023, 45(2): 355-362(in Chinese).

[9] SHAHEEN M A M, YOUSRI D, FATHY A, et al. A novel application of improved marine predators algorithm and particle swarm optimization for solving the ORPD problem[J]. Energies, 2020, 13(21): 5679.

[10] AYDEMIR S B. Enhanced marine predator algorithm for global optimization and engineering design problems[J]. Advances in Engineering Software, 2023, 184: 103517.

[11] KUSUMA P D, ADIPUTRA D. Hybrid marine predator algorithm and hide object game optimization[J]. Engineering Letters, 2023, 31(1): 262-270.

[12] HAMZACEBI C, KUTAY F. Continuous functions minimization by dynamic random search technique[J]. Applied Mathematical Modelling, 2007, 31(10): 2189-2198.

[13] 吴建辉, 王博华, 张小刚, 等. 基于模式搜索法的云模型粒子群算法[J]. 控制与决策, 2017, 32(11): 2076-2080.
WU J H, WANG B H, ZHANG X G, et al. Cloud model particle swarm optimization algorithm based on pattern search method[J]. Control and Decision, 2017, 32(11): 2076-2080(in Chinese).

[14] TIZHOOSH H R. Opposition-based reinforcement learning[J]. Journal of Advanced Computational Intelligence and Intelligent Informatics, 2006, 10(4): 578-585.

[15] RAHNAMAYAN S, JESUTHASAN J, BOURENNANI F, et al. Computing opposition by involving entire population[C]//Proceedings of the IEEE Congress on Evolutionary Computation. Piscataway: IEEE Press, 2014: 1800-1807.

[16] ZHU G P, KWONG S. Gbest-guided artificial bee colony algorithm for numerical function optimization[J]. Applied Mathematics and Computation, 2010, 217(7): 3166-3173.

[17] YANG X S. Firefly algorithms for multimodal optimization[C]// Proceedings of the International Symposium on Stochastic Algorithms. Berlin: Springer, 2009: 169-178.

[18] QIN A K, HUANG V L, SUGANTHAN P N. Differential evolution algorithm with strategy adaptation for global numerical optimization[J]. IEEE Transactions on Evolutionary Computation, 2009, 13(2): 398-417.

[19] LIANG J J, QIN A K, SUGANTHAN P N, et al. Comprehensive learning particle swarm optimizer for global optimization of multimodal functions[J]. IEEE Transactions on Evolutionary Computation, 2006, 10(3): 281-295.

[20] ZHAO S R, WU Y L, TAN S, et al. QQLMPA: a quasi-opposition learning and Q-learning based marine predators algorithm[J]. Expert Systems with Applications, 2023, 213: 119246.

[21] SADIQ A S, DEHKORDI A A, MIRJALILI S, et al. Nonlinear marine predator algorithm: a cost-effective optimizer for fair power allocation in NOMA-VLC-B5G networks[J]. Expert Systems with Applications, 2022, 203: 117395.

[22] SUN J, PALADE V, WU X J, et al. Solving the power economic dispatch problem with generator constraints by random drift particle swarm optimization[J]. IEEE Transactions on Industrial Informatics, 2014, 10(1): 222-232.

[23] MENG Z Y, ZHONG Y X, MAO G J, et al. PSO-sono: a novel

PSO variant for single-objective numerical optimization[J]. Information Sciences, 2022, 586: 176-191.

[24] SUN X P, JIANG L S, SHEN Y, et al. Success history-based adaptive differential evolution using turning-based mutation[J]. Mathematics, 2020, 8(9): 1565.

[25] HUSSAIN K, ZHU W, SALLEH M N M. Long-term memory Harris' hawk optimization for high dimensional and optimal power flow problems[J]. IEEE Access, 2019, 7: 147596-147616.

[26] JIANG L L, MASKELL D L, PATRA J C. Parameter estimation of solar cells and modules using an improved adaptive differential evolution algorithm[J]. Applied Energy, 2013, 112: 185-193.

[27] YU K J, LIANG J J, QU B Y, et al. Parameters identification of photovoltaic models using an improved JAYA optimization algorithm[J]. Energy Conversion and Management, 2017, 150: 742-753.

Marine predator algorithm incorporating hybrid search operators and competitive learning and applications

XU Zhonghui[1], RAO Zhenyuan[1,2], MA Yanli[3], TANG Zejing[1], HUANG Xiaodong[3,*]

(1. School of Information Engineering, Jiangxi University of Science and Technology, Ganzhou 341000, China;
2. Jingdezhen Health School, Jingdezhen 333000, China;
3. School of Economics and Management, Jiangxi University of Science and Technology, Ganzhou 341000, China)

Abstract: The predator search marine predators algorithm PSMPA is a new algorithm that incorporates hybrid search operators and competitive learning to address the problems of population diversity loss, low solution precision, and difficulty escaping local optima in the later iterations of the classic marine predator algorithm (MPA). The introduction of a stochastic dynamic centroid opposition-based learning mechanism enhances population diversity in the later stages of the algorithm, expands the search space, and improves the algorithm's ability to escape local optima and accelerate convergence. By combining random search and pattern search as hybrid search operators, the algorithm's local search capability is enhanced. The average population fitness is increased when predators engage in competitive learning behavior, which effectively promotes rapid convergence and greatly enhances solution quality. Simulations using 12 CEC2017 benchmark functions demonstrate that PSMPA achieves substantial improvements in optimization performance, convergence speed, and stability. Furthermore, its application in optimizing parameters for solar photovoltaic models further validates PSMPA's practical value and effectiveness in solving real-world engineering optimization problems.

Keywords: marine predator algorithm; hybrid search operators; stochastic dynamic centroid opposition-based learning; competitive learning; engineering optimization problems

Received: 2024-04-23; **Accepted**: 2024-08-09; **Published Online**: 2024-10-12 11:20

URL: link.cnki.net/urlid/11.2625.V.20241012.0858.001

Foundation items: Sub-themes of National Key Research and Development Program of China (2020YFB1713700,2020YFB1713705)

* **Corresponding author**. E-mail: huangmail3@126.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0224

基于改进贝叶斯网络的A320空调系统故障传播路径识别方法

陈玖圣[1,*]，余卓洋[1]，郭润夏[1]，吴军[2]

(1. 中国民航大学 电子信息与自动化学院，天津 300300；2. 中国民航大学 航空工程学院，天津 300300)

摘 要：飞机空调系统是控制飞机内部环境的关键机载系统之一，其内部结构呈现多闭环和多冗余等复杂特征，导致故障在各组件之间传播。贝叶斯网络可以有效推理开环系统的故障传播路径，但无法适应闭环结构。为解决这一适应性问题，提出一种基于贝叶斯网络的改进模型，用于识别A320空调系统特定结构下的故障传播路径。分析空调系统的功能行为和物理结构，综合复杂网络理论，构建系统拓扑有向图；结合网络结构和信息传递定义边影响强度，对边重要性进行度量，提高边重要性度量的准确性；通过基于边影响强度的解环策略，将闭环结构转换为开环结构，以获得最佳的贝叶斯故障传播网络结构，从而精确识别多闭环结构下的故障传播路径。通过A320空调系统案例研究验证了所提方法的有效性。

关 键 词：空调系统；闭环系统；边重要性评估；贝叶斯网络；故障传播路径

中图分类号：V245

文献标志码：A **文章编号**：1001-5965(2026)06-1827-12

飞机空调系统(air conditioning system, ACS)为机组人员、乘客和设备提供了舒适的环境，同时在保障飞行安全方面起着不可替代的作用。ACS是一个复杂且高度耦合的多闭环反馈控制系统[1]，通过复杂的电气线路和气体管道将飞机的驾驶舱、前客舱和后客舱等紧密连接。这种系统结构创造了潜在的故障传播路径，即使是局部微小的故障也会通过气体管道和电气线路在系统内不断传播、积累和放大[2]，可能导致ACS性能下降甚至直接失效，进而威胁飞机的飞行安全，甚至引发灾难性后果。因此，ACS的故障诊断和准确找出故障传播路径是确保飞机飞行安全的关键。

目前，故障传播分析的方法主要分为两大类：基于复杂网络的方法和数据驱动的方法。基于复杂网络的方法利用图论将复杂系统抽象成网络，以表征组件之间的拓扑特性，捕获工业过程中的相关性和因果关系，进而描述复杂系统中的故障传播过程[3-5]。然而，这些方法在故障传播问题的分析中过于依赖工业知识和专家经验，并且过度关注网络拓扑特性，忽略了实际网络中组件和链路的可靠性特征。

与基于复杂网络的方法不同，数据驱动方法需要收集监测点的运行数据。在大多数有关故障传播分析的研究中，数据驱动方法广泛应用于分析因果关系或诊断根本原因，如互相关分析(cross-correlation analysis, CCA)[6]、格兰杰因果关系(Granger causality, GC)[7]、K近邻(K-nearest neighbor, KNN)[8]。这些方法在故障的根本原因诊断和故障传播识别方面均表现出令人满意的效果。以上研究关注于多设备之间的故障传播，需要大量实时的过程数据。

贝叶斯网络(Bayesian network, BN)以图形式定性地表示变量间关系，同时以概率形式定量表示关

收稿日期：2024-04-16；**录用日期**：2024-06-07；**网络出版时间**：2024-07-02 10:23

网络出版地址：link.cnki.net/urlid/11.2625.V.20240701.1638.001

基金项目：国家自然科学基金(52375557, 62173331)；天津市自然科学基金(23JCYBJC00060)；天津市教委科研项目(2023KJ222)；中央高校基本科研业务费专项资金(3122023PY06, 3122023044, 3122024052, KJZ53420210113)

* **通信作者**. E-mail：jschen@cauc.edu.cn

引用格式：*陈玖圣，余卓洋，郭润夏，等. 基于改进贝叶斯网络的A320空调系统故障传播路径识别方法[J]. 北京航空航天大学学报，2026，52(6)：1827-1838. CHEN J S，YU Z Y，GUO R X，et al. A fault propagation path identification method for A320 air conditioning system based on improved Bayesian network[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1827-1838 (in Chinese).*

联强度，用于确定过程运行状态、诊断故障传播路径和根本原因，在处理不确定性和数据稀缺性问题时具有强大的推理能力[9]。随着贝叶斯理论和相关研究的迅速发展，基于贝叶斯网络的方法逐渐成为复杂系统中过程故障诊断最为广泛应用的方法[10-12]。然而，在基于贝叶斯网络的方法中，由于贝叶斯网络的非循环性质，无法考虑到 A320 飞机 ACS 中由于反馈控制原因而普遍存在的组件和过程变量之间的闭环因果关系，导致了不准确的根本原因诊断和故障传播路径识别[13-14]。因此，处理闭环结构对于准确识别故障传播路径至关重要[15]。

本文针对 A320 飞机 ACS 的多闭环特性，提出了一种基于改进贝叶斯网络的识别故障传播路径模型。具体来说，针对 A320 飞机 ACS 的故障传播识别，采用知识与数据相结合的方法解决了贝叶斯网络无法适应闭环结构的问题。本文的主要创新点如下：①从系统机理和物理结构出发，通过分析组件间的耦合关系来建立 ACS 拓扑有向图模型。②基于复杂网络理论，分析系统网络结构特性，提出一种结合网络结构和信息传递度量边重要性的指标，为闭环处理提供了评估准则。③提出一种基于边影响强度 (edge influence strength, EIS) 的闭环处理策略，实现对多闭环反馈控制系统故障传播路径的识别。

1 A320 飞机 ACS 拓扑有向图模型

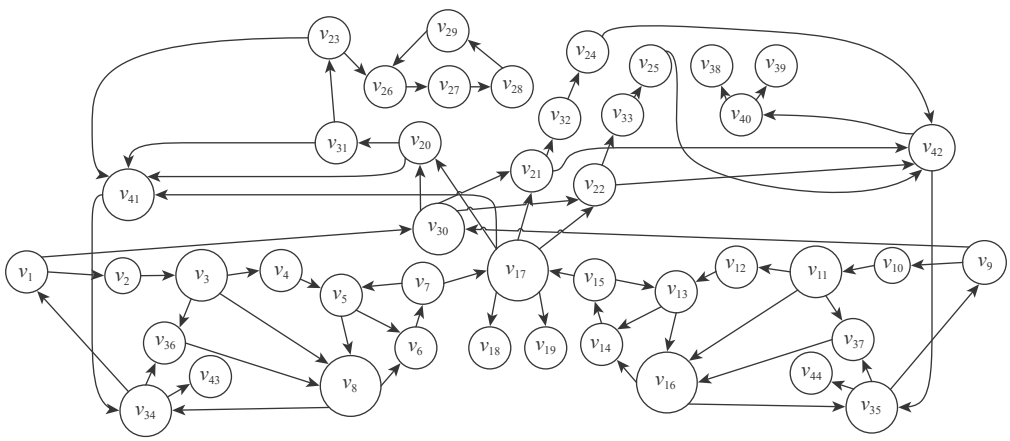
1.1 A320 飞机 ACS 系统描述

A320 飞机 ACS 由区域温度控制、压力控制、电子电气设备通风、货舱通风系统构成[16]。首先，湿热空气由引气系统供应，通过流量控制活门进入主热交换器，将空气中的热量排出到周围的压缩空气中，使其温度降低。其次，冷却后的空气进入压气机进行空气压缩，这种高能量的空气通过次热交换器将热量排出到周围的压缩空气中，经过再加热器和冷凝器。至此，Pack 组件将高温高压气体调节为具有合适流量、温度和压力的冷却空气。然后，经过配平空气活门，冷空气与热空气歧管中的热空气结合。最后，来自引气系统、空调组件和客舱的混合空气通过低压导管分配到驾驶舱、前客舱和后客舱等区域，满足了飞机温度和压力调节的要求。

Pack 组件是 ACS 的核心制冷组件，包括主、次热交换器、压气机、涡轮、水分离器和冷凝器。Pack1 与 Pack2 这 2 套组件相互独立，并由各自的组件控制器进行控制。每个 Pack 组件都能独立维持整个 ACS 的正常运行。当温度需求发生变化时，ACS 控制器通过电气线路迅速调整相关的活门和部件，以确保系统能够及时适应新的需求，维持舱内温度的稳定性，保证乘客舒适度。同时，系统会实时监测部件的温度、压力、空气流量等相关参数，并反馈给控制器。随着 ACS 功能的不断完善，其内部的电气线路和气体管道日益复杂且相互交织，导致组件之间建立了紧密的联系，形成了多闭环、多冗余的复杂系统。

1.2 基于有向图的 ACS 拓扑模型

如图 1 所示，将 ACS 中组件抽象为网络节点，将组件之间的关系抽象为网络的有向边，构建 A320 飞机 ACS 拓扑有向图 $G=(V,E,W)$。其中，



v_1:流量控制活门 1；v_2:次热交换器 1；v_3:压气机 1；v_4:主热交换器 1；v_5:再加热器 1；v_6:冷凝器 1；v_7:水分离器 1；v_8:涡轮 1；v_9:流量控制活门 2；v_{10}:次热交换器 2；v_{11}:压气机 2；v_{12}:主热交换器 2；v_{13}:再加热器 2；v_{14}:冷凝器 2；v_{15}:水分离器 2；v_{16}:涡轮 2；v_{17}:混合组件；v_{18}:再循环风扇 1；v_{19}:再循环风扇 2；v_{20}:配平空气活门 1；v_{21}:配平空气活门 2；v_{22}:配平空气活门 3；v_{23}:驾驶舱压力传感器；v_{24}:前客舱压力传感器；v_{25}:后客舱压力传感器；v_{26}:鼓风机；v_{27}:电子电气设备；v_{28}:排气扇；v_{29}:蒙皮热交换器；v_{30}:热空气压力调节活门；v_{31}:驾驶舱风管；v_{32}:前客舱风管；v_{33}:后客舱风管；v_{34}:Pack 控制器 1；v_{35}:Pack 控制器 2；v_{36}:旁通活门 1；v_{37}:旁通活门 2；v_{38}:前货舱温度传感器；v_{39}:后货舱温度传感器；v_{40}:货舱通风控制器；v_{41}:ACS 控制器 1；v_{42}:ACS 控制器 2；v_{43}:冲压空气进口活门 1；v_{44}:冲压空气进口活门 2。

图 1 A320 飞机空调系统拓扑有向图

Fig. 1 A320 air conditioning system topology directed graph

$V=\{v_1,v_2,\cdots,v_N\}$为组件的有限集合, N为节点的总数; $E=\{e_{12},\cdots,e_{ij}\}$为具有因果关系节点的边集合, e_{ij}为节点v_i到节点v_j的有向边; $W=\{w_{12},\cdots,w_{ij}\}$为边的权重集合, 其元素表示节点间的信息传递量。将组件与传感器所得的过程变量进行一对一关联, 给定$\boldsymbol{X}=\left(\boldsymbol{x}^1,\boldsymbol{x}^2,\cdots,\boldsymbol{x}^N\right)^{\mathrm{T}}=(\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_T)\in\mathbf{R}^{N\times T}$为节点集合$V$对应的特征矩阵, T为采样点数, 本文用$\boldsymbol{x}^n=\left(x_1^n,x_2^n,\cdots,x_T^n\right)^{\mathrm{T}}\in\mathbf{R}^T$来表示节点$v_n$对应长度为$T$的特征向量, 并用$\boldsymbol{x}_t=\left(x_t^1,x_t^2,\cdots,x_t^N\right)^{\mathrm{T}}\in\mathbf{R}^N$表示$t$时刻所有节点的采样值。$w_{ij}=T_{ij}$为节点$v_i$到节点$v_j$的边权重, 用传递熵值来衡量节点间信息传递的强度。

2　边重要性度量方法

2.1　ACS 边重要性度量指标

有向边表示网络中节点间存在的因果关系, 边重要性用于衡量边对网络连通性的贡献程度及承载信息能力的强度。

1) 边连通关键度。为综合考虑节点对之间最短路径经过边、边连接的节点的重要性及边所处的派系, 本文引入了 2 个顶点的介数信息, 结合了边的中介中心性和派系, 构建边连通关键度指标, 定义如下:

$$Z_{ij}=\left(B_i\sum_{\substack{v_s,v_t\in V\\ e_{ij}\in E\\ (i,j)\neq(s,t)}}\frac{\sigma\left(v_s,v_t\left|e_{ij}\right.\right)}{\sigma(v_s,v_t)}B_j\right)\Bigg/\left(\sum_{k=3}^{N}C_{ij}^k+1\right)\tag{1}$$

式中: $\sigma(v_s,v_t)$为节点v_s到节点v_t的最短路径的数目; $\sigma\left(v_s,v_t\left|e_{ij}\right.\right)$为节点$v_s$到节点$v_t$的最短路径中经过边$e_{ij}$的数目; B_i和B_j分别为节点v_i和节点v_j的节点介数; C_{ij}^k表示经过边e_{ij}且节点个数为k的完全图的数量; $\sum_{k=3}^{N}C_{ij}^k$表示经过边e_{ij}且节点个数大于 3 的所有完全图的数量。完全图指图中每对不同的节点之间都存在一条边, 即每个节点都直接连接到其他所有的节点。

2) 信息熵。网络中的信息熵反映了其信息量, 而边的信息熵则衡量了边在网络中的信息传输能力。结合节点度、节点强度指标及边的自身属性, 综合考虑边两端节点对边的支撑作用, 构建边信息熵表达式为

$$H_{ij}=w_{ij}+\frac{k_i}{k_i+k_j}S_j+\frac{k_j}{k_i+k_j}S_i\tag{2}$$

式中: k_i和k_j分别为节点v_i和节点v_j的度值; $S_i=\sum_{v_j\in A_i}w_{ij}$为节点$v_i$的强度, A_i为节点v_i的相邻节点集合。

2.2　基于边连通关键度和信息熵的度量方法

边连通关键度从网络连通性的角度识别网络中的重要边, 信息熵度量边的信息传输能力。为了将网络结构与信息传递结合起来, 提出了一种基于边连通关键度和信息熵的边重要性度量方法。

$$I_{ij}=\alpha Z_{ij}+(1-\alpha)H_{ij}\tag{3}$$

式中: I_{ij}为边影响强度, 表示边e_{ij}的重要性; α用于调整 2 个指标的权重。

此外, Z_{ij}表示边的全局信息, H_{ij}表示边的局部信息。通过 2 种方式度量边的重要性, 同时考虑点与边之间的相互作用, 从而提高边重要性度量的准确性。

3　ACS 贝叶斯故障传播模型

3.1　基于 EIS 的 ACS 闭环处理策略

传统贝叶斯网络无法解析闭环结构, 因此, 需要对 ACS 的多闭环拓扑有向图进行解环处理。本文提出了一种基于 EIS 的 ACS 闭环处理策略, 步骤如下:

步骤 1　构建 ACS 多闭环拓扑有向图。依据图 1, 抽取节点集合V和有向边集合E, 构建 ACS 拓扑有向图, 通过计算节点间的传递熵值作为 ACS 有向图的边权重W, 最终构建 ACS 拓扑有向图$G=(V,E,W)$。

步骤 2　边重要性评估。根据式 (3) 计算得到有向加权图G中每条边的I_{ij}值。

步骤 3　贪婪搜索算法。采用贪婪搜索算法对初始 ACS 多闭环有向加权图中的边进行搜索, 并移除相应组合的边。通过循环结构检验, 得到候选有向无环图集合$G_{\mathrm{greedy}}=\{G_1,G_2,\cdots,G_m,\cdots,G_K\}$, 其中, $G_m=(V_m,E_m,W_m)$为第m个有向无环图, K为候选的有向无环图总数。

步骤 4　计算基于 EIS 的得分函数。对于每个由贪婪搜索算法得到的有向无环图, 需要对结构的合理性进行评估。每个结构的得分如下:

$$\Gamma_m=\sum_{e_{ij}\in E_m}I_{ij}-\beta\lg n_m\tag{4}$$

式中: Γ_m为第m个有向无环图G_m基于I_{ij}的分数; n_m为G_m中有向边的数量; E_m为G_m的有向边集合; β为用于惩罚网络复杂性的标量参数。

对每个结构打分后, 根据式 (5), 选择得分最高的结构作为最终的贝叶斯网络结构。

$$G_{\mathrm{best}}=\arg\max_{G_m\in G_{\mathrm{greedy}}}\Gamma_m\tag{5}$$

3.2　ACS 贝叶斯参数学习

将网络中的先验信息与观测数据结合, 以学习

ACS贝叶斯网络的参数。节点之间的传递关系由计算后验概率得到,其转移关系公式如下:

$$P(v_i|\Pi_i)=\frac{P(\Pi_i|v_i)P(v_i)}{P(\Pi_i)} \tag{6}$$

式中:Π_i为节点v_i的父节点;$P(v_i)$和$P(\Pi_i)$分别为节点v_i和其父节点的先验概率;$P(\Pi_i|v_i)$为似然函数,采用极大似然估计来计算,统计给定节点v_i在特定状态下观察到Π_i特定状态的占比情况。

为使网络中各节点的条件概率更接近实际情况,通过统计ACS的运行数据和维护记录,并以组件平均故障间隔时间的倒数作为各节点的故障率来确定网络节点的先验概率。

利用链式法则,根据条件概率分布,可进一步计算出所需节点集的联合概率分布如下:

$$P(v_1,v_2,\cdots,v_N)=\prod_{i=1}^{N}P(v_i|\Pi_i) \tag{7}$$

3.3 ACS故障传播路径推理

引入基于蚁群算法的优化机制,将BN模型中各节点间故障发生条件概率及其影响强度集成到故障传播路径推理中。

定义第u个蚂蚁从节点v_i迁移到节点v_j的概率为

$$P_{ij}^{u}(t)=\frac{\left[\tau_{ij}(t)\right]^{\theta_1}\left[I_{ij}\right]^{\theta_2}}{\sum_{j\in F_i^u}\left[\tau_{ij}(t)\right]^{\theta_1}\left[I_{ij}\right]^{\theta_2}} \tag{8}$$

式中:$\tau_{ij}(t)$为时刻t在边e_{ij}上铺设的信息素数量;由于故障优先沿影响强度高的边缘传播,定义启发式信息值为I_{ij};F_i^u为第u个蚂蚁在节点v_i的可行邻域的节点集合;θ_1和θ_2分别控制信息素和启发式信息的相对重要性。

根据贝叶斯网络故障传播网络中节点的条件概率和蚁群算法中个体在选择下一节点时的转移概率,每条传播路径最终的发生概率P_{op}可依次算出。

在每次迭代后,每个路径上的信息素根据式(9)进行更新:

$$\tau_{ij}=(1-\rho)\tau_{ij}+\Delta\tau_{ij} \tag{9}$$

$$\Delta\tau_{ij}=\sum_{u=1}^{m}\Delta\tau_{ij}^{u} \tag{10}$$

$$\Delta\tau_{ij}^{u}=\begin{cases}Q/l_u & \text{当蚂蚁经过边}e_{ij}\\ 0 & \text{其他}\end{cases} \tag{11}$$

式中:$\rho(0<\rho<1)$为信息素的挥发度;m为蚂蚁的数量;$\Delta\tau_{ij}^{u}$为第u只蚂蚁在边e_{ij}上铺设的信息素的浓度;Q为常数;l_u为第u只蚂蚁的行程长度。

在A320飞机ACS中,大多数单个故障都存在多条传播路径,通过推理可以确定哪条故障传播路径最可能发生,从而制定相应的安全应急预案。本文提出的改进贝叶斯网络模型的逐步过程逻辑流程如图2所示。

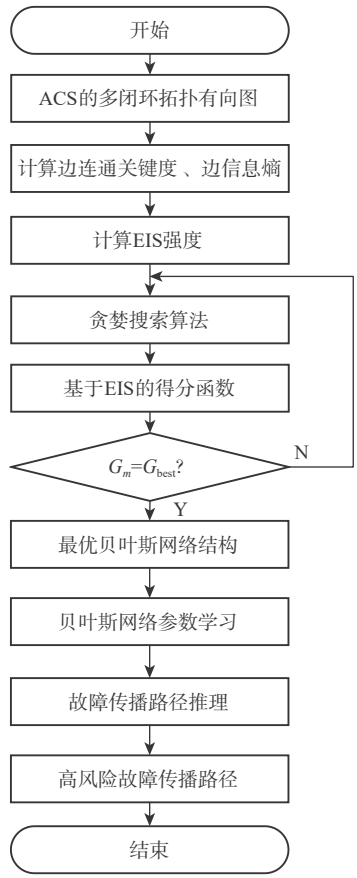


图2 改进贝叶斯网络模型流程

Fig. 2 Flow diagram of improved Bayesian network model

4 实验结果与分析

本节比较EIS方法与其他相关的边重要性度量指标在A320飞机ACS网络中的性能,推断A320飞机ACS的故障传播路径,并将本文提出的改进贝叶斯网络模型与其他4种模型进行比较。实验平台的计算机配置如下:Intel Xeon E5-1630v4,主频为3.70 GHz, RAM为16 GB。

4.1 评估指标

4.1.1 重要性度量评估指标

本文使用边连通系数δ和网络效率下降率μ来评估EIS指标,挖掘相对重要边的总体结构。在网络连通性度量中,使用边连通系数来评价方法的性能。边连通系数δ定义为

$$\delta=\phi_{ij}/n_0 \tag{12}$$

式中:ϕ_{ij}为边e_{ij}被移除后网络极大连通子图所包含的连边数量;n_0为原有网络中所包含的有向边数量。

网络效率表示去掉部分边后的网络连通性强度,可以描述为

$$\eta = \frac{1}{N(N-1)} \sum_{v_i, v_j \in V} \frac{1}{d_{ij}} \tag{13}$$

式中: N为网络节点总数; d_{ij}为节点v_i与节点v_j之间的最短路径。

为了更直观地分析边移除对网络效率的影响, 采用网络效率下降率μ, 定义为

$$\mu = 1 - \frac{\eta}{\eta_0} \tag{14}$$

式中: η_0为原有的网络效率; η为移除边后的网络效率。

4.1.2　故障传播路径识别性能指标

本文使用路径准确率、运行时长、路径数量 3 个指标来评估故障路径识别的性能。路径准确率计算公式如下:

$$\psi = \varphi_r / \varphi \tag{15}$$

式中: ψ为路径准确率; φ为所有故障传播路径数量; φ_r为分析正确的故障传播路径数量, 即故障传播符合实际运行状况的路径数量。

运行时长为故障传播路径总数识别所耗时间总和; 路径数量为识别的故障传播路径总数量。

4.2　ACS 边重要性度量方法参数计算与分析

4.2.1　EIS 值量化分析

根据 A320 飞机 ACS 拓扑有向图, 可得网络节点的邻接矩阵, 进而统计得到经过边的最短路径和完全图个数, 依据式 (1) 计算可得边连通关键度Z_{ij}, 结果如表 1 所示。结合边的权重和式 (2), 可计算边信息熵值H_{ij}, 结果如表 2 所示。

根据表 1、表 2 和式 (3), 调整参数设置为$\alpha = 0.4$, 计算 ACS 拓扑有向图中各边的影响强度值I_{ij}, 如图 3 所示。可以看出, 边重要性度量方法所定义的 EIS 值会随着边的连通关键度和边的信息熵而变化。边$e_{41,34}$和边$e_{42,35}$具有较高的边介数, 通过该边的路径数较多, 边对网络连通性的贡献较大, 其对应的 EIS 值为 1.664, 优于其他边。虽然边$e_{7,17}$和边

表 1　边连通关键度值

Table 1　Value of edge connectivity key

e_{ij}	Z_{ij}	e_{ij}	Z_{ij}	e_{ij}	Z_{ij}	e_{ij}	Z_{ij}	e_{ij}	Z_{ij}	e_{ij}	Z_{ij}
(1,2)	0.011 5	(8,6)	0.034 8	(15,13)	0.001 1	(21,32)	0.001 5	(30,20)	0.115 7	(35,44)	0.000 0
(1,30)	0.086 6	(8,34)	0.021 2	(15,17)	0.115 2	(21,42)	0.121 9	(30,21)	0.054 4	(36,8)	0.011 1
(2,3)	0.003 9	(9,10)	0.013 2	(16,14)	0.044 7	(22,33)	0.001 5	(30,22)	0.014 7	(37,16)	0.013 7
(3,4)	0.000 3	(9,30)	0.140 2	(16,35)	0.024 8	(22,42)	0.015 6	(31,23)	0.011 8	(40,38)	0.000 0
(3,8)	0.006 7	(10,11)	0.003 9	(17,18)	0.000 0	(23,26)	0.006 4	(31,41)	0.007 3	(40,39)	0.000 0
(3,36)	0.000 1	(11,12)	0.000 3	(17,19)	0.000 0	(23,41)	0.006 2	(32,24)	0.000 4	(41,34)	0.203 8
(4,5)	0.000 2	(11,37)	0.000 1	(17,41)	0.065 7	(24,42)	0.006 7	(33,25)	0.000 4	(42,35)	0.319 1
(5,6)	0.002 7	(11,16)	0.007 2	(17,20)	0.060 4	(25,42)	0.006 7	(34,1)	0.140 3	(42,40)	0.012 2
(5,8)	0.001 5	(12,13)	0.000 2	(17,21)	0.073 5	(26,27)	0.002 7	(34,36)	0.019 8		
(6,7)	0.045 4	(13,14)	0.003 6	(17,22)	0.017 2	(27,28)	0.000 7	(34,43)	0.000 0		
(7,5)	0.001 0	(13,16)	0.001 8	(20,31)	0.032 9	(28,29)	0.000 0	(35,9)	0.226 4		
(7,17)	0.097 2	(14,15)	0.061 9	(20,41)	0.061 0	(29,26)	0.000 0	(35,37)	0.025 8		

表 2　边信息熵值

Table 2　Value of edge information entropy

e_{ij}	H_{ij}	e_{ij}	H_{ij}	e_{ij}	H_{ij}	e_{ij}	H_{ij}	e_{ij}	H_{ij}	e_{ij}	H_{ij}
(1,2)	0.031 6	(8,6)	0.214 4	(15,13)	0.185 2	(21,32)	0.059 3	(30,20)	0.307 6	(35,44)	0.005 8
(1,30)	0.027 5	(8,34)	0.017 4	(15,17)	0.012 2	(21,42)	0.042 7	(30,21)	0.392 1	(36,8)	0.282 4
(2,3)	0.038 0	(9,10)	0.034 8	(16,14)	0.216 7	(22,33)	0.063 4	(30,22)	0.418 3	(37,16)	0.291 5
(3,4)	0.081 0	(9,30)	0.028 6	(16,35)	0.017 9	(22,42)	0.045 3	(31,23)	0.096 2	(40,38)	0.023 4
(3,8)	0.056 7	(10,11)	0.043 5	(17,18)	0.010 2	(23,26)	0.012 1	(31,41)	0.028 3	(40,39)	0.023 4
(3,36)	0.109 8	(11,12)	0.096 7	(17,19)	0.010 8	(23,41)	0.020 8	(32,24)	0.116 2	(41,34)	0.174 5
(4,5)	0.043 1	(11,37)	0.127 1	(17,41)	0.028 9	(24,42)	0.014 7	(33,25)	0.121 1	(42,35)	0.198 6
(5,6)	0.016 3	(11,16)	0.065 0	(17,20)	0.090 4	(25,42)	0.015 2	(34,1)	0.102 4	(42,40)	0.152 5
(5,8)	0.043 4	(12,13)	0.045 9	(17,21)	0.042 3	(26,27)	0.006 5	(34,36)	0.017 4		
(6,7)	0.007 9	(13,14)	0.016 4	(17,22)	0.044 7	(27,28)	0.005 1	(34,43)	0.005 8		
(7,5)	0.161 0	(13,16)	0.044 6	(20,31)	0.064 1	(28,29)	0.011 3	(35,9)	0.106 7		
(7,17)	0.010 3	(14,15)	0.007 9	(20,41)	0.025 9	(29,26)	0.106 0	(35,37)	0.018 6		

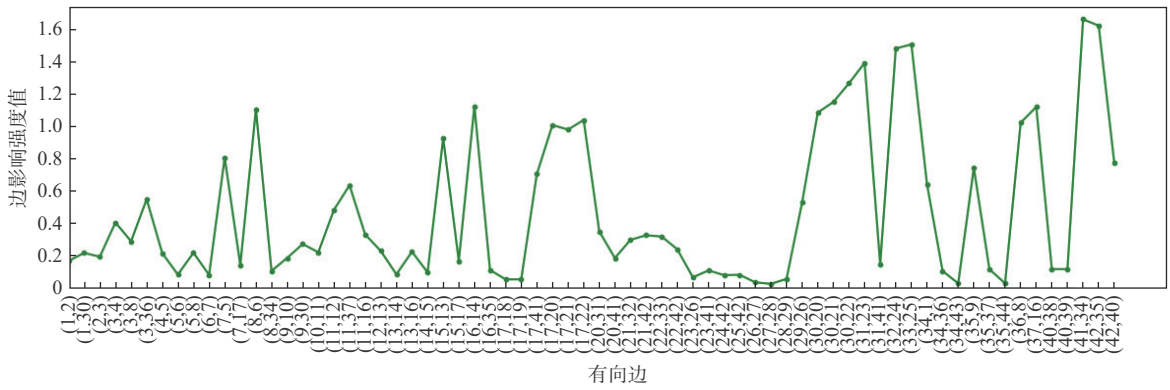


图3 边影响强度值 I_{ij}

Fig. 3 Value of edge influence intensity I_{ij}

$e_{15,17}$连接了网络中重要的节点17，但由于这2条边参与组成的完全图较多，形成许多类似的连通结构，$e_{7,17}$和$e_{15,17}$的边重要性降低。由于在A320飞机ACS中，配平空气活门的开度将直接影响客舱内的温度，节点间信息传递能力强，所对应的边$e_{30,21}$、边$e_{30,22}$的EIS值较高。由此可见，基于EIS的边重要性度量方法能够有效反映网络中真实的边重要性排序。

4.2.2 不同边重要性度量指标的性能分析

为验证本文方法的有效性，采用边的拓扑重叠（topological overlap, TO）[17]、边端点度乘积（vertex degree of edge, ED）[18]、边端点的介数乘积（vertex betweeness of edge, EB）[19]、边扩散强度（diffusion intensity of edge, DI）[20]和本文指标分别进行仿真对比分析。

边的连通系数下降速度越快，表示度量指标对网络连通性的影响越显著，其边的重要性度量精度越高。如图4所示，当边的移除比例为0.3时，TO、ED、EB、DI和EIS这5种指标对应的边的连通系数值分别为0.166、0.138、0.151、0.181和0.092。这表明当移除前30%的边时，采用EIS指标网络连通性受到的影响程度比采用TO、ED、EB、DI等指标更高。EIS指标的性能相较于第二位的ED指标提高了8.6%。

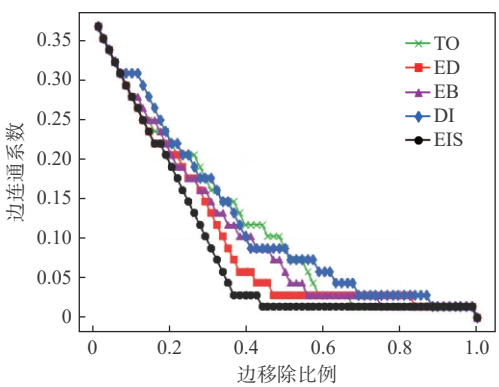


图4 不同指标下的边连通系数δ

Fig. 4 Edge connection coefficients δ for different metrics

按照重要度从大到小的顺序移除关键边时，网络效率下降率的变化情况如图5所示。随着重要边移除比例的增加，网络效率的下降率呈上升趋势。μ值越大，网络效率下降的幅度越大，相应度量指标精度越高。EIS指标对网络效率下降影响最快，最先将网络效率降低到原有的40%。以上分析结果表明，EIS指标在大多数情况下能够快速分解网络，并有效度量边的重要性。

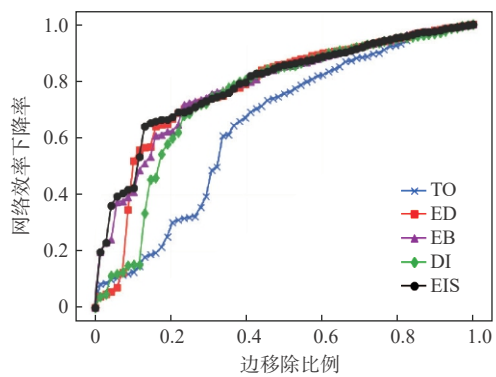


图5 不同指标下的网络效率下降率μ

Fig. 5 Network efficiency degradation rate μ for different metrics

4.3 故障传播路径分析

4.3.1 基于EIS的ACS闭环处理

在度量每条边的影响强度后，应用基于EIS的ACS闭环处理策略，根据式(4)计算每个搜索得到的有向无环图的得分，最高得分为28.409 1。图6为最终的ACS贝叶斯故障传播网络结构，图中，虚线表示已移除的边。

4.3.2 贝叶斯参数学习

通过统计ACS的运行数据和维护记录，确定网络节点的先验概率，如图7所示。以v_8（涡轮1）为例，根据v_3、v_5的状态和式(6)、式(7)，可以计算出v_8在不同父节点状态组合下的条件概率。例如，当v_3、v_5均处于正常状态时，v_8正常的条件概率为

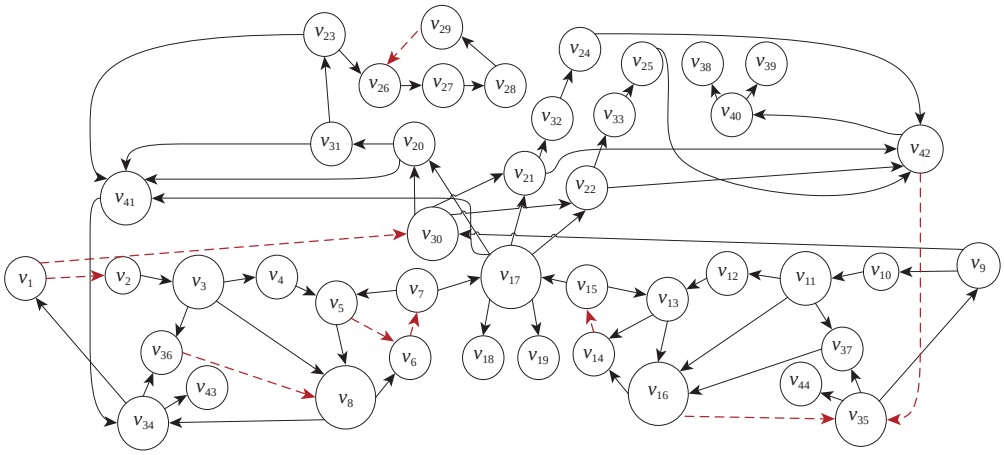


图 6　A320 飞机空调系统故障传播有向无环图

Fig. 6　Directed acyclic graph of A320 air conditioning system fault propagation

0.985 1, 故障概率为 0.014 9; 当 v_3 故障、v_5 正常时, v_8 正常的条件概率为 0.885 2, 故障概率为 0.114 8。其余状态组合下的条件概率可采用相同方法计算。同样, A320 飞机 ACS 故障传播网络中其余节点的条件概率也可以依此确定。至此, 构建了完整的 ACS 贝叶斯故障传播网络, 如图 8所示。

4.3.3　故障传播路径推理

在求解故障传播路径前, 先设置蚁群算法中的相关参数, 如表 3 所示。

利用式 (8)～式 (11), 结合网络中各节点间的条件概率和影响强度进行定量推理故障传播概率。当蚁群算法在迭代循环中趋于稳定时, 可以计算出算法中各故障传播路径的最终发生概率。当前阶段的关键在于确定受当前节点影响的相关节点的路径, 以及确定对当前节点有较高概率影响的路径。根据状态空间蚁群搜索结果和基于 A320 飞机 ACS 贝叶斯故障传播网络, 故障传播路径及其发生概率 P_{op} 如表 4 所示。其中, 风险最高的故障传播路径为 3→4→5→8, 即压气机→主热交换器→再加热器→涡轮, 如图 9 所示。图中, 空调冷却系统的核心组件压气机采用了空气轴承。在高空环境下, 外部空气压力较低, 轴承存在磨损情况, 更易使压气机出现异常运行。当压气机突然卡阻或严重磨损时, 其效率会下降, 导致主热交换器和再加热器的冷热路径温差减小, 进而影响工作效率。此外, 这种情况还会显著削弱涡轮的冷却作用, 导致 Pack 组件出口温度升高, 偏离正常范围。因此, 该路径准确反映了故障的实际传播情况, 据此可以针对实际维修任务制定相应的安全应急控制策略。当压气机运行异常时, 不仅需要检查空气轴承区域的磨损情况并更换压气机内的部件, 还需监测主热交换器、再加热器及涡轮的温度等参数, 以避免小故障的累积和放大, 从而影响 ACS 的稳定运行。利用改进贝叶斯网络模型的输出, 可以建立完整的故障传播关系模型, 计算故障的转移概率, 确定故障传播路径, 有效发现和消除系统中的脆弱点, 并制定相应的安全管理策略, 实现多闭环反馈控制系统的全面健康管理。

4.4　对比实验

本文对比了小世界网络[21]、有符号有向图 (signed directed graph, SDG)[22]、Petri 网[23]、基于最大支撑树 (most weight supported tree, MWST) 的贝叶斯网络[24] 和改进贝叶斯网络等故障传播路径识别模型。文献 [21] 通过分析小世界网络的特性定义了节点的故障传播强度, 并进行了故障攻击和最大负荷攻击, 根据蚁群算法得出高风险的故障传播路径。文献 [22] 将单层 SDG 模型转换为分层模型, 根据系统的初始响应在该模型中逐层搜索故障的传播路径。文献 [23] 根据网络节点的失效信息, 将改进的 Petri 网与有向图理论相结合, 并通过建立的方程来观察状态转移, 当网络节点状态转移时, 通过状态方程得到故障传播路径的所有可能概率。文献 [24] 在 MWST 构造过程中, 通过断开将互信息作为权值的最小边, 结合蚁群算法进行结构学习, 进而得出故障传播路径的发生概率。

本节将所提出的改进贝叶斯网络模型与小世界网络、SDG、Petri 网和 MWST+BN 模型进行对比, 分别对各网络模型故障传播路径识别的路径准确率、运行时长和路径数量 3 个指标进行评估和分析。在评估中, 采用相同的 A320 飞机 ACS 运行数据和蚁群算法参数, 根据机电系统可靠性相关手册设定小世界网络、SDG 和 Petri 网模型中全局阈值 $P_{Gate} = 10^{-8}$ 和限环阈值 $P_{Circle} = 10^{-6}$[25]。计算得到 A320 飞机 ACS 网络平均路径长度 $L = 2.364$, 平均聚类系数 C=0.077 3, 表明该网络结构具有小世界网络的性质。实验结果如表 5 所示。可知, 本文模型

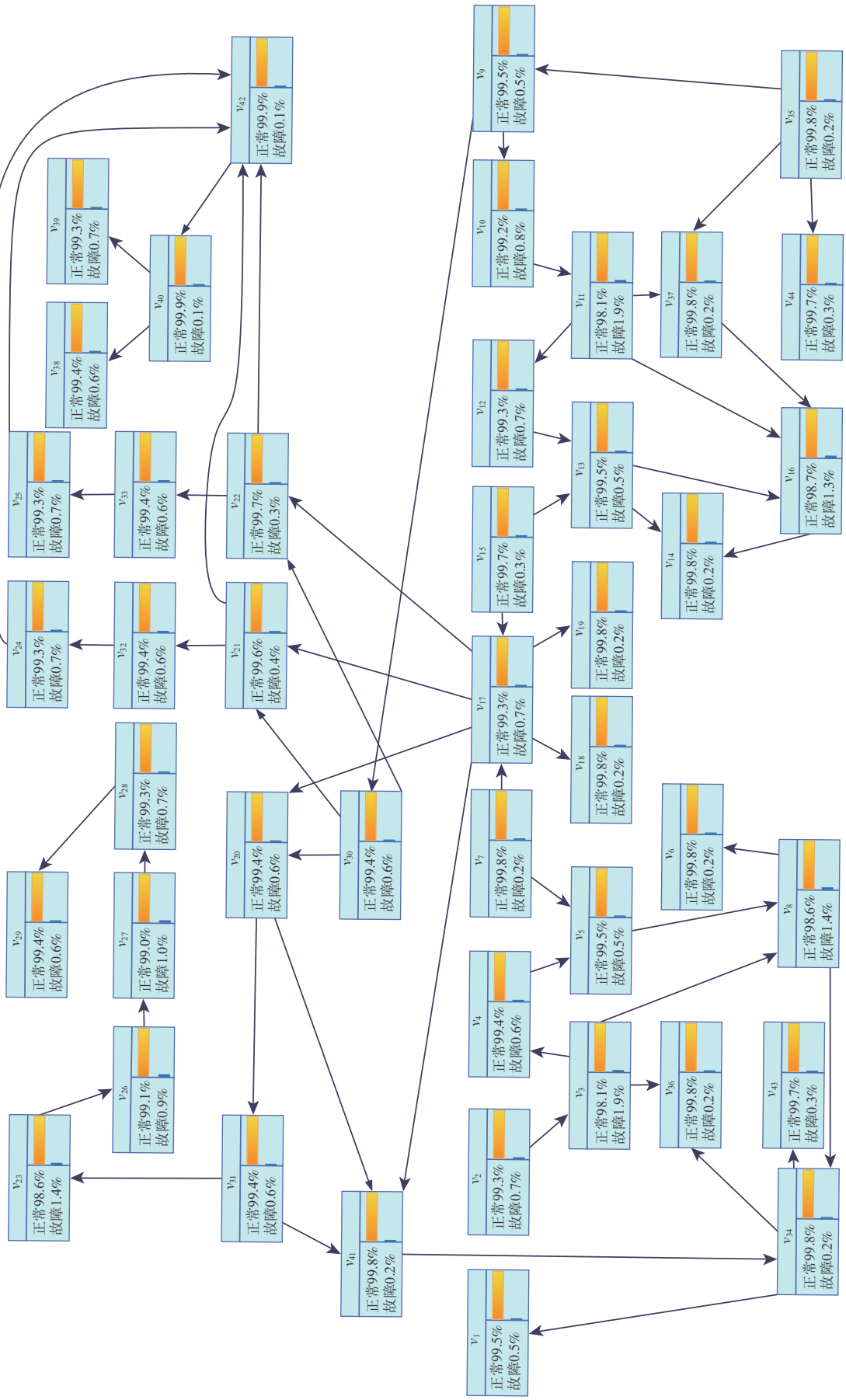


图7 ACS贝叶斯网络节点状态的先验概率

Fig. 7 Prior probability of node states in ACS Bayesian network

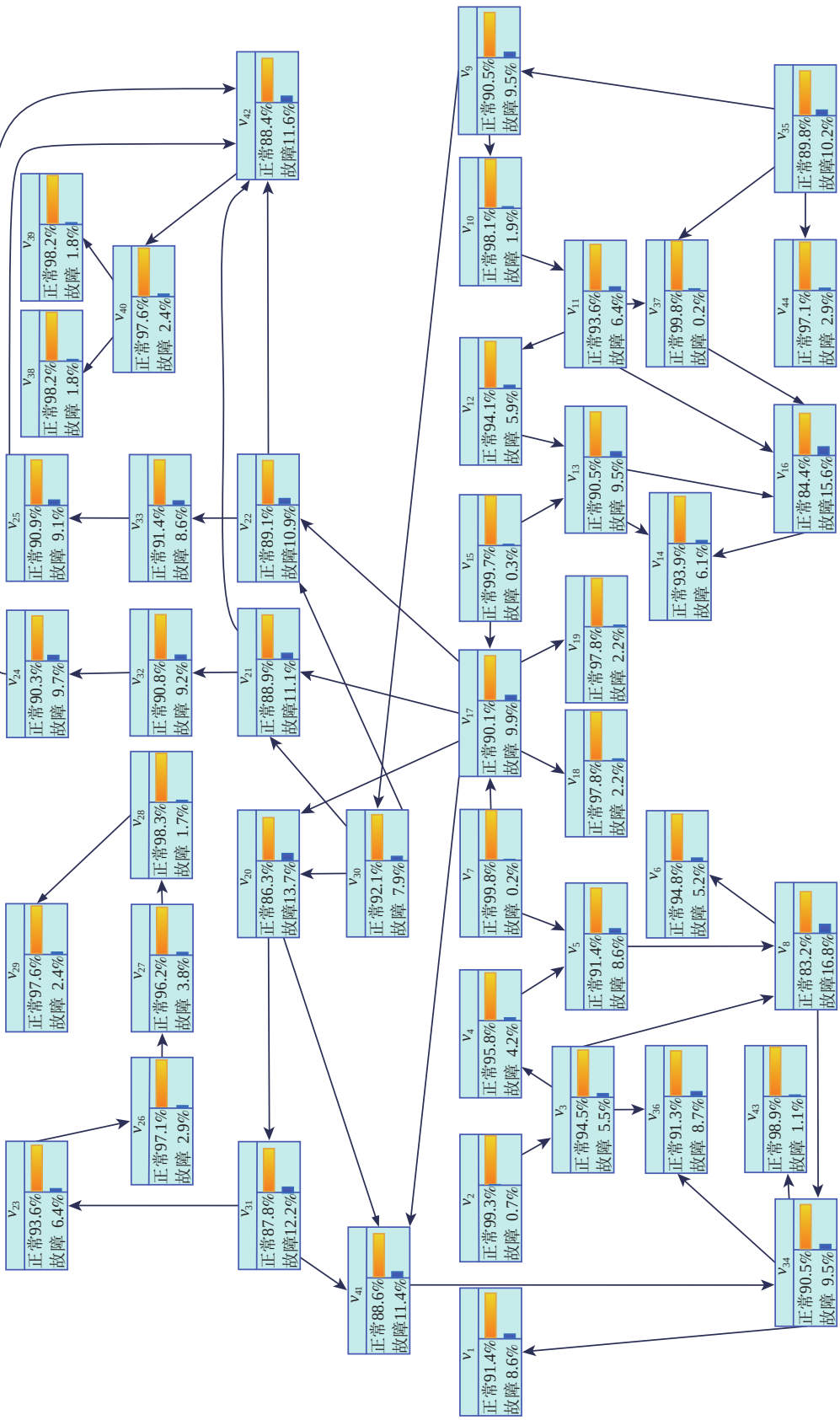


图 8　ACS 贝叶斯故障传播网络

Fig. 8　ACS Bayesian fault propagation network

表3 蚁群算法参数

Table 3 Ant colony algorithm parameters

m	θ_1	θ_2	ρ	Q
44	1	2	0.7	50

识别A320飞机ACS故障传播路径的准确率比小世界网络、Petri网和MWST+BN模型分别高出2.12%、2.93%、1.39%和0.54%；运行时长比其他4种模型分别减少了8.88%、32.06%、26.69%和6.25%，证明了本文模型在识别故障传播路径过程中具有较高的精度和效率。本文模型在较优的网络结构中，结合边的传播强度和节点间的条件概率计算故障传播概率，确定故障传播路径。比较结果表明，本文提出的改进贝叶斯网络模型能高效且准确地识别A320飞机ACS的故障传播路径，为分析A320飞机ACS的故障传播问题提供了可行方案，有助于保障飞机的安全运行。

表4 A320飞机空调系统故障传播路径及发生概率(部分)

Table 4 Fault propagation paths and occurrence probability in A320 air conditioning system (extracted)

序号	路径	P_{op}	序号	路径	P_{op}
1	3→4→5→8	0.386 1	12	30→20→31	0.192 3
2	9→10→11	0.384 7	13	17→22→33	0.167 5
3	11→12→13	0.314 5	14	22→33→25	0.149 1
4	17→21→32	0.294 3	15	20→31→23	0.145 2
5	26→28→29	0.277 2	16	11→12→13→16	0.121 7
6	9→30→20	0.251 1	17	21→32→24→42	0.106 4
7	8→34→43	0.246 7	18	17→21→32→24	0.091 3
8	21→32→24	0.245 2	19	26→27→28→29	0.088 7
9	7→5→8	0.238 8	20	10→11→16→14	0.074 2
10	15→13→16	0.221 9	21	2→3→8→6	0.070 1
11	11→37→16	0.208 4	22	10→11→16→14	0.049 4

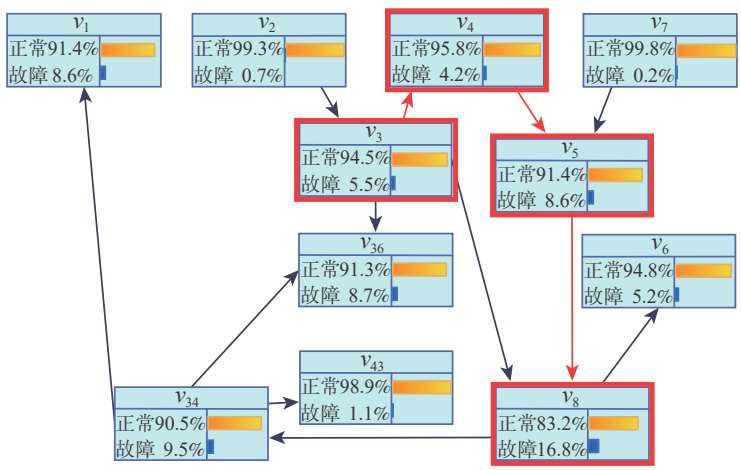


图9 ACS高风险故障传播路径

Fig. 9 ACS high-risk fault propagation path

表5 空调系统故障传播路径识别对比实验结果

Table 5 Comparative experimental results of identifying fault propagation paths in air conditioning system

模型	路径识别准确率/%	路径数量/条	运行时长/s
小世界网络	92.55	1 496	214
SDG	91.74	1 536	287
Petri网	93.28	1 432	266
MWST+BN	94.13	1 371	208
改进贝叶斯网络	94.67	1 248	195

5 结 论

1）针对A320空调系统多闭环结构特性，提出了一种能够高效解决故障传播随机性和复杂性问题的改进贝叶斯网络。

2）引入了结合网络结构和传输性能的边重要性度量指标，对边的影响强度进行量化，同时提高了边重要性度量的精度。

3）设计了处理闭环结构的策略，采用基于边影响强度的得分函数，获取最优的贝叶斯故障传播网络，有效解决贝叶斯网络无法适应闭环结构的问题。

4）使用A320空调系统数据，将本文模型与4种现有的先进模型进行对比实验，结果显示，其路径识别准确率和运行时间均优于这4种模型，证明所提出的改进贝叶斯网络在多闭环反馈控制系统中具有一定的优越性和高效性。

参考文献（References）

[1] 康文文, 李浩敏. 基于模型的飞机系统架构多视图表达方法[J]. 系统工程与电子技术, 2021, 43(11): 3266-3277.
KANG W W, LI H M. Multi-view representation method of aircraft system architecture based on model[J]. Systems Engineering and Electronics, 2021, 43(11): 3266-3277(in Chinese).

[2] GABBAR H A. Improved qualitative fault propagation analysis[J].

Journal of Loss Prevention in the Process Industries, 2007, 20(3): 260-270.

[3] HUANG S, LU N, JIANG B, et al. Fault propagation analysis of computer numerically controlled machine tools[J]. Journal of Manufacturing Systems, 2023, 70: 149-159.

[4] LI Y, LU N Y, SHI J T, et al. A quantitative causal diagram based optimal sensor allocation strategy considering the propagation of fault risk[J]. Journal of the Franklin Institute, 2021, 358(1): 1021-1043.

[5] JIANG Z, ZHAO T, WANG S, et al. A novel risk assessment and analysis method for correlation in a complex system based on multi-dimensional theory[J]. Applied Sciences, 2020, 10(9): 3007.

[6] YANG B, LI H G, WEN B. A dynamic time delay analysis approach for correlated process variables[J]. Chemical Engineering Research and Design, 2017, 122: 141-150.

[7] MA L, DONG J, PENG K X. A practical root cause diagnosis framework for quality-related faults in manufacturing processes with irregular sampling measurements[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 3511509.

[8] LANDMAN R, JÄMSÄ-JOUNELA S L. Fault propagation analysis by implementing nearest neighbors method using process connectivity[J]. IEEE Transactions on Control Systems Technology, 2019, 27(5): 2058-2067.

[9] KITSON N K, CONSTANTINOU A C, GUO Z G, et al. A survey of Bayesian network structure learning[J]. Artificial Intelligence Review, 2023, 56(8): 8721-8814.

[10] KUMARI P, LEE D, WANG Q S, et al. Root cause analysis of key process variable deviation for rare events in the chemical process industry[J]. Industrial & Engineering Chemistry Research, 2020, 59(23): 10987-10999.

[11] YE T Y, LIU L L, PANG H W, et al. Bayesian networks based approach to enhance GO methodology for reliability modeling of multi-state consecutive-k-out-of-n: F system[J]. Reliability Engineering & System Safety, 2023, 229: 108828.

[12] CHU K Y, LIU R, DUAN G J. A gray correlation based Bayesian network model for fault source diagnosis of multistage process–small sample manufacturing system[J]. Advanced Engineering Informatics, 2023, 56: 101918.

[13] CASTELLANO-QUERO M, FERNÁNDEZ-MADRIGAL J A, GARCÍA-CEREZO A. Improving Bayesian inference efficiency for sensory anomaly detection and recovery in mobile robots[J]. Expert Systems with Applications, 2021, 163: 113755.

[14] KUMARI P, BHADRIRAJU B, WANG Q S, et al. A modified Bayesian network to handle cyclic loops in root cause diagnosis of process faults in the chemical process industry[J]. Journal of Process Control, 2022, 110: 84-98.

[15] 马亮, 彭开香, 董洁. 工业过程故障根源诊断与传播路径识别技术综述[J]. 自动化学报, 2022, 48(7): 1650-1663.
MA L, PENG K X, DONG J. Review of root cause diagnosis and propagation path identification techniques for faults in industrial processes[J]. Acta Automatica Sinica, 2022, 48(7): 1650-1663(in Chinese).

[16] 孙见忠, 解志峰, 闫洪胜, 等. 民机 PHM 预测维修模式在空调系统的应用[J]. 南京航空航天大学学报, 2021, 53(6): 952-964.
SUN J Z, XIE Z F, YAN H S, et al. Application of PHM predictive maintenance on aircraft air conditioning system[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2021, 53(6): 952-964(in Chinese).

[17] ONNELA J P, SARAMÄKI J, HYVÖNEN J, et al. Structure and tie strengths in mobile communication networks[J]. Proceedings of the National Academy of Sciences of the United States of America, 2007, 104(18): 7332-7336.

[18] HOLME P, KIM B J, YOON C N, et al. Attack vulnerability of complex networks[J]. Physical Review E, Statistical, Nonlinear, and Soft Matter Physics, 2002, 65(5 Pt 2): 056109.

[19] GIRVAN M, NEWMAN M E J. Community structure in social and biological networks[J]. PNAS, 2002, 99(12): 7821-7826.

[20] LIU Y, TANG M, ZHOU T, et al. Improving the accuracy of the k-shell method by removing redundant links: from a perspective of spreading dynamics[J]. Scientific Reports, 2015, 5: 13172.

[21] XU Y F, WANG Z, JIANG Y, et al. Small-world network analysis on fault propagation characteristics of water networks in eco-industrial parks[J]. Resources, Conservation and Recycling, 2019, 149: 343-351.

[22] YANG F, XIAO D Y, SHAH S L. Signed directed graph-based hierarchical modelling and fault propagation analysis for large-scale systems[J]. IET Control Theory & Applications, 2013, 7(4): 537-550.

[23] WANG Y H, SHI H, JIA L M, et al. Multi paths fault propagation model for network modeled system[C]//Proceeding of the 11th World Congress on Intelligent Control and Automation. Piscataway: IEEE Press, 2015: 5915-5920.

[24] GUO R X, WANG Z H. A framework for modeling fault propagation paths in air turbine starter based on Bayesian network[J]. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 2022, 236(6): 1078-1095.

[25] WANG Y H, LI M, SHI H. A method of searching fault propagation paths in mechatronic systems based on MPPS model[J]. Journal of Central South University, 2018, 25(9): 2199-2218.

A fault propagation path identification method for A320 air conditioning system based on improved Bayesian network

CHEN Jiusheng[1,*], YU Zhuoyang[1], GUO Runxia[1], WU Jun[2]

(1. School of Electronic Information and Automation, Civil Aviation University of China, Tianjin 300300, China;
2. School of Aeronautical Engineering, Civil Aviation University of China, Tianjin 300300, China)

Abstract: The air conditioning system of airplanes is one of the critical airplane-mounted systems used to control the internal environments. It is highly complex in structure, having multiple closed loops and much redundancy, which causes the faults to spread among its components. The Bayesian network can be used to deduce the fault transmission path of open-looped systems, but it is unfit for close-looped structures. In order to address this issue, an enhanced model based on the Bayesian network for determining the fault transmission path of the air conditioning system in A320 aircraft in a particular structure is presented in this research. Firstly, the functional behavior and physical structure of this air conditioning system are analyzed. By using the complex network theory, a topologically directed graph of this system is constructed. Additionally, in order to quantify the relevance of sides, the intensity of side influence is established based on the network structure and message transit, which increases the measurement's accuracy.Then, through the loop-opening strategy based on the intensity of side influence, a close-looped structure is changed into an open-looped structure. In this way, the optimal Bayesian fault transmission network structure is obtained to precisely identify the path of fault transmission in a close-looped structure. Finally, a case study is conducted on the A320 air conditioning system to validate the proposed model.

Keywords: air condition system; closed-loop systems; edge importance assessment; Bayesian network; fault propagation path

Received: 2024-04-16; **Accepted**: 2024-06-07; **Published Online**: 2024-07-02 10:23

URL: link.cnki.net/urlid/11.2625.V.20240701.1638.001

Foundation items: National Natural Science Foundation of China (52375557,62173331); Tianjin Natural Science Foundation (23JCYBJC00060); Tianjin Municipal Education Commission Research Project (2023KJ222); The Fundamental Research Funds for Central Universities (3122023PY06,3122023044,3122024052,KJZ53420210113)

* **Corresponding author**. E-mail: jschen@cauc.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0275

基于NOMA的无人机ISAC系统轨迹和资源联合优化

唐菁敏[1]，胡丞[1]，宋耀莲[1,*]，虞贵财[2]

(1. 昆明理工大学 信息工程与自动化学院，昆明 650500；2. 宜春学院 物理科学与工程技术学院，宜春 336000)

摘 要：针对当前无线频谱资源紧张和多用户信号间的干扰，探讨了在非正交多址接入(NOMA)技术支持下，无人机通信感知一体化(ISAC)系统中的感知目标选择、波束成形及无人机飞行轨迹和速度的联合优化问题，旨在提高系统的吞吐量。针对该问题的非凸特性，采用块坐标下降(BCD)法，将复杂问题拆分为3个子问题。求解子问题时，引入松弛变量、一阶泰勒近似和连续凸近似(SCA)技术，简化计算过程，并提高求解效率。通过交替迭代以上子问题，得到原问题的近似最优解。仿真验证了所提算法的有效性，可有效提高系统最大平均吞吐量，并展示了良好的收敛性。

关 键 词：通信感知一体化；无人机；非正交多址接入；轨迹优化；资源优化

中图分类号：TN929.5; V279

文献标志码：A **文章编号**：1001-5965(2026)06-1839-11

近年来，下一代无线网络B5G/6G的发展受到了各界的广泛关注，无线通信系统的应用场景也越来越广。同时，在许多对于B5G/6G网络的设想中，一个共识是传感功能将发挥更重要的作用，特别是对于常见的位置/环境感知场景[1]。因此，在B5G/6G网络中，需要同时具备通信和感知能力。针对这一趋势，科研人员提出了将雷达感知和无线通信融合在一起，共享同一频谱和基础设施的通信感知一体化(integrated sensing and communication, ISAC)技术[2]。

由于ISAC具有频谱共享、成本降低等优点，其已被公认为迈向5G及以上无线网络的关键技术之一。然而，要实现ISAC的目标，需要对系统的各个方面进行全面设计，包括传输波形、接收信号的处理和波束形成设计。因此，关于ISAC的波束形成设计受到越来越多的关注。例如，文献[3]提出了以每个基站服务一个用户的方式，研究多输入多输出(multiple input multiple output, MIMO)系统中的联合通感波束的设计。文献[4-5]通过对联合通感波束进行设计，在保证通信用户最低的传输速率情况下，基于传感误差对波束成形矢量进行设计，在通信和感知之间的性能进行权衡，以达到ISAC的目标。文献[6]考虑了是否具有感知干扰消除功能的2种类型的通信接收机，并分别为2种类型的接收机设计了最佳波束成形。文献[7-8]提出了联合发射波束成形模型来优化雷达发射波束方向图，同时满足每个通信用户的信号干扰加噪声比的要求。然而，传感性能通常取决于目标和收发器之间的视距链路(line of sight, LoS)，而非视距链路(non-line of sight, NLoS)被视为对目标传感的不利干扰。对于远离基站或被障碍物阻挡的潜在目标，由于回波信号严重的路径损耗，传感性能严重下降。因此，以地面基站作为平台进行ISAC的设计面临着固有的感知限制，只适用于固定范围内提供感知和通信服务。

收稿日期：2024-04-29；**录用日期**：2024-06-28；**网络出版时间**：2024-07-17 16:49

网络出版地址：link.cnki.net/urlid/11.2625.V.20240715.1905.002

基金项目：国家自然科学基金(62261056)；江西省自然科学基金(20212BAB202029)

* **通信作者**. E-mail：39217149@qq.com

引用格式：唐菁敏，胡丞，宋耀莲，等. 基于NOMA的无人机ISAC系统轨迹和资源联合优化[J]. 北京航空航天大学学报，2026，52(6)：1839-1849. TANG J M，HU C，SONG Y L，et al. NOMA-based joint optimization of trajectory and resources for UAV-enable integrated sensing and communication system[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1839-1849 (in Chinese).

随着无人机 (unmanned aerial vehicle, UAV) 技术和多天线技术的发展，UAV 作为空中基站可以有效快速地实现应急通信[9]，为用户提供低成本高效益的无线通信连接；同时，搭载多天线可获得更多分集增益。文献 [10] 设计了 UAV 轨迹和发射波束形成，发送组合信息和感知信号与多个用户通信，同时感知地面上的潜在目标，这是将 UAV 作为空中 ISAC 平台。文献 [11] 在 UAV 赋能系统中提出了一种新颖的自适应 ISAC 机制，其中，UAV 在通信过程中按需进行感知，并且感知持续时间根据应用需求灵活配置，在保证通信和感知的服务质量要求下最大化系统吞吐量。文献 [12] 通过部署多架 UAV 协同执行任务，可以实现 ISAC 网络的更大覆盖范围。文献 [13] 提出了 ISAC 与无线蜂窝网络连接后，可以用于监视和定位的可疑 UAV 目标，以保护物理层安全。然而，UAV 搭载多天线的方案在增加了系统增益的同时，也会造成通信用户间的信道干扰，使得通信性能下降，不能满足通信需求。

为解决上述通信用户间干扰，引入非正交多址接入 (non-orthogonal multiple access, NOMA) 技术进行处理，NOMA 技术以串行干扰删除 (successive interference cancellation, SIC) 通过对信号功率进行排序，逐步减去最大信号功率用户的干扰，实现正确解调。文献 [14] 提出了基于 NOMA 的 ISAC 框架，叠加的 NOMA 通信信号被同时用于目标感测，以此最大化通信速率和感知功率。文献 [15] 采用 NOMA 来实现多播和单播信号的双频谱共享，这 2 种信号在功率域中叠加，叠加的通信信号也被用作雷达探测波形，使得信号满足组播速率要求和雷达波束方向图预定义精度的前提下，达到单播速率最大化。文献 [16] 将旋翼 UAV 作为固定空中基站来进行基于 NOMA 的 ISAC 系统的研究，在满足最低通信速率的前提下对波束形成设计，进而提高 ISAC 系统的性能。在 UAV 的工作过程中，能量恒定，但上述关于 NOMA UAV 通信感知场景并未考虑 UAV 的能耗问题。

综上所述，当前基于 NOMA 的 UAV ISAC 系统的研究未充分考虑 UAV 的能耗问题，未充分发挥 UAV 自身具有的高机动性、高灵活性特点，缩小了 UAV 进行通信与感知的范围。因此，本文针对 UAV ISAC 系统总能耗受限的情况进行研究，对于受地质灾害所致的应急救援场景，基站等基础通信设施遭到破坏，通过 UAV ISAC 系统的快速部署，对灾区的环境状况和受困人员的位置进行感知，为救援人员提供实时信息。本文主要研究工作如下：

1) 建立 UAV ISAC 系统，在保证用户最低传输速率和感知要求的条件下，设计联合优化方案，以最大化系统的平均吞吐量，该方案包括感知目标选择、波束成形、UAV 轨迹规划及速度设计。

2) 由于问题存在多变量的耦合，具有非凸性。将原问题通过块坐标下降(block coordinate descent, BCD)法拆分为 3 个相关子问题，利用一阶泰勒展开和连续凸近似(successive convex approximation, SCA)方法，并引入松弛变量，将问题转换为凸形式进行求解。通过交替迭代优化方法，获得了问题的近似最优解。

3) 实验表明，在满足感知性能和最低传输速率的前提下，所提联合优化算法能显著提高最大平均吞吐量，并具有良好的收敛性。

1 系统模型

本文考虑了如图 1 所示的由 1 架固定翼多天线 UAV、M个感知目标和K个单天线通信用户共同组成的 UAV ISAC 系统。UAV 作为空中基站搭载了 $N_{\mathrm{s}}=N_{\mathrm{t}}\times N_{\mathrm{r}}$均匀平面天线阵列；通信用户 $k\in\{1,\cdots,K\}$, $K\geqslant 1$；感知目标$m\in\{1,\cdots,M\}$, $M\geqslant 1$。

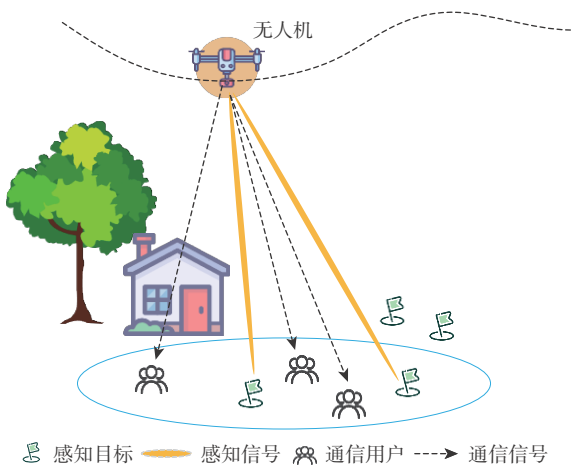


图 1 UAV ISAC 系统模型

Fig. 1 UAV ISAC system model

UAV 的高度固定为H，将任务周期T均分为N个时隙，每个时隙长度为$\delta_t=T/N$，其中，δ_t足够小。假设 UAV 的位置在一个时隙内不发生改变，以方便 ISAC 的轨迹和波束形成设计。UAV 在第n个时隙的水平坐标表示为$\boldsymbol{q}[n]=(x[n],y[n]),n=1,2,\cdots,N$，假设第$k$个用户的水平坐标位置表示为$\boldsymbol{g}_k=(x_k,y_k)$；UAV 最大飞行速度为$v_{\max}$，在第$n$个时隙速度为$\boldsymbol{v}[n]$。UAV 在实现通信感知任务时应满足以下约束：

$$\|\boldsymbol{q}[n+1]-\boldsymbol{q}[n]\|\leqslant v_{\max}\delta_t \qquad 1\leqslant n\leqslant N-1 \tag{1}$$

$$\boldsymbol{q}[1]=\boldsymbol{q}_{\mathrm{I}} \tag{2}$$

$$\boldsymbol{q}[N]=\boldsymbol{q}_{\mathrm{F}} \tag{3}$$

式中: $\boldsymbol{q}_{\mathrm{I}}$和$\boldsymbol{q}_{\mathrm{F}}$分别为 UAV 飞行的始末位置。

UAV 为用户传输信息信号, 同时提供感知信号, 可以为传感提供额外的自由度。假设 UAV 在第n个时隙内对第k个用户发送的信号为$s_k[n]$, $\boldsymbol{w}_k[n]\in\mathbf{C}^{N_{\mathrm{t}}\times 1}$表示相应的发射波束形成矢量, $\boldsymbol{s}_0[n]$表示第n个时隙发射的感知信号, 假设通信信号$s_k[n]$是均值和单位方差为零的独立圆对称复高斯(circularly symmetric complex Gaussian, CSCG)随机变量, 即$s_k[n]\sim\mathrm{CN}(0,1)$, 感知信号$\boldsymbol{s}_0[n]$均值为零。因此, UAV 在第$n$个时隙的发射信号为

$$\boldsymbol{x}_k[n]=\sum_{k=1}^{K}\boldsymbol{w}_k[n]s_k[n]+\boldsymbol{s}_0[n] \tag{4}$$

1.1　通信与感知模型

1.1.1　通信模型

由于 UAV 配置了均匀平面天线阵列, 在第n个时隙中, UAV 与第k个用户之间的天线阵列响应矢量为

$$\boldsymbol{a}_k[n]=\left[1,\mathrm{e}^{\frac{-\mathrm{j}2\pi d_{\mathrm{t}}\sin(\theta_k[n])}{\lambda_{\mathrm{c}}}},\cdots,\mathrm{e}^{\frac{-\mathrm{j}2\pi d_{\mathrm{t}}(N_{\mathrm{t}}-1)\sin(\theta_k[n])}{\lambda_{\mathrm{c}}}}\right]\otimes\left[1,\mathrm{e}^{\frac{-\mathrm{j}2\pi d_{\mathrm{r}}\sin(\phi_k[n])}{\lambda_{\mathrm{c}}}},\cdots,\mathrm{e}^{\frac{-\mathrm{j}2\pi d_{\mathrm{r}}(N_{\mathrm{r}}-1)\sin(\phi_k[n])}{\lambda_{\mathrm{c}}}}\right] \tag{5}$$

式中: $\otimes$表示 kronecker 积; d_{t}和d_{r}分别为均匀平面阵列中不同方向相邻天线之间的距离; λ_{c}为载波波长; $\theta_k[n]$和$\phi_k[n]$分别为UAV 到用户k的离开仰角和离开方位角, 具体为

$$\sin(\theta_k[n])=\frac{x_k-x[n]}{\sqrt{\|\boldsymbol{q}[n]-\boldsymbol{g}_k\|^2+H^2}} \tag{6}$$

$$\sin(\phi_k[n])=\frac{y_k-y[n]}{\sqrt{\|\boldsymbol{q}[n]-\boldsymbol{g}_k\|^2+H^2}} \tag{7}$$

UAV 与地面通信用户间的通信链路为视距链路。在第n个时隙中, UAV 和第k个用户的信道增益表示为

$$\boldsymbol{h}_k[n]=\sqrt{\frac{\beta_0}{\|\boldsymbol{q}[n]-\boldsymbol{g}_k\|^2+H^2}}\boldsymbol{a}_k[n] \tag{8}$$

式中: β_0为 UAV 与用户距离 1 m 时的信道增益。

因此, 在第n个时隙中, 第 k个用户接收到的信号为

$$y_k[n]=\boldsymbol{h}_k^{\mathrm{H}}[n]\boldsymbol{x}_k[n]+n_i \tag{9}$$

式中: n_i为加性高斯白噪声。

由于采用 NOMA 技术对用户间的干扰进行消除, 假设用户的顺序相对于信道强度由小到大呈递增趋势, 对于用户k, 采用 SIC 技术依次检测并消除所有满足$j<k$的用户干扰信号, 并在用户k处进行解码[14], 从而使得用户k处的吞吐量可表示为

$$R_{k,k}[n]=\log_2\left(1+\frac{\left\|\boldsymbol{h}_k^{\mathrm{H}}[n]\boldsymbol{w}_k[n]\right\|^2}{\sum\limits_{i<k}^{K}\left\|\boldsymbol{h}_k^{\mathrm{H}}[n]\boldsymbol{w}_i[n]\right\|^2+\sigma_n^2}\right) \tag{10}$$

式中: σ_n^2表示噪声。

用户k的信号也需要在用户j处进行解码, 从而保证串行干扰消除的正确执行, 其在用户j处的吞吐量可表示为[14]

$$R_{k,j}[n]=\log_2\left(1+\frac{\left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_k[n]\right\|^2}{\sum\limits_{i>k}^{K}\left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_i[n]\right\|^2+\sigma_n^2}\right) \tag{11}$$

进一步, 可得到用户k在第n个时隙的吞吐量为

$$R_k[n]=\min(R_{k,k}[n],R_{k,k+1}[n],\cdots,R_{k,K}[n]) \tag{12}$$

则用户k的吞吐量可以表示为

$$R_k=\sum_{n=1}^{N}R_k[n] \tag{13}$$

因此, 系统的平均吞吐量可以表示为

$$R_{\mathrm{c}}=\frac{1}{N}\sum_{k=1}^{K}R_k \tag{14}$$

1.1.2　感知模型

在针对应急救灾场景的 UAV ISAC 系统中, 考虑对任务区域潜在目标的感知, 需要对目标位置方向进行检测, 以获得更大的增益。假设 UAV 的目标是探测地面上M个潜在目标, 其中, 第m个目标的水平坐标为$\boldsymbol{g}_m=(x_m,y_m)$。在感知过程中, 由于 UAV 发射的通信信号是已知的, 目标反射的通信信号也可用于系统中的目标参数估计, 通信信号$s_k[n]$也可用于感知。在第n个时隙内, 通过可联合通信信号$s_k[n]$和感知信号$\boldsymbol{s}_0[n]$获得从 UAV 到目标m方向的发射波束方向图增益为

$$\vartheta(\boldsymbol{q}[n],m)=E[\|\boldsymbol{a}_{\mathrm{s}}^{\mathrm{H}}[n]\boldsymbol{x}_k[n]\|^2] \tag{15}$$

式中: $\boldsymbol{a}_{\mathrm{s}}[n]$为对应方向的导向向量。

据式 (15) 可知, $E(\|\boldsymbol{x}_k[n]\|^2)=\sum\limits_{k=1}^{K}\|\boldsymbol{w}_k[n]\|^2+\mathrm{tr}(\boldsymbol{R}_{\mathrm{s}}[n])$, 其中, $\boldsymbol{R}_{\mathrm{s}}[n]$为感知信号$\boldsymbol{s}_0[n]$的协方差矩阵。因此, 可得到

$$\vartheta(\boldsymbol{q}[n],m)=\boldsymbol{a}_{\mathrm{s}}^{\mathrm{H}}[n]\left(\sum_{k=1}^{K}\boldsymbol{w}_k[n]\boldsymbol{w}_k^{\mathrm{H}}[n]+\boldsymbol{R}_{\mathrm{s}}[n]\right)\boldsymbol{a}_{\mathrm{s}}[n] \tag{16}$$

从式 (15) 中可得有效感知功率$P_{\mathrm{s}}=\boldsymbol{a}_{\mathrm{s}}^{\mathrm{H}}[n]\boldsymbol{R}_{\mathrm{s}}[n]\cdot\boldsymbol{a}_{\mathrm{s}}[n]$, 进一步得到第$n$个时隙时第$m$个目标反射信号功率为

$$\Gamma(\boldsymbol{q}[n],\boldsymbol{g}_m)=\sum_{k=1}^{K}\|\boldsymbol{w}_k[n]\|^2+\mathrm{tr}(\boldsymbol{R}_{\mathrm{s}}[n]) \tag{17}$$

基于感知任务的实时性要求[17], 在多用户和多

目标场景之间的通信和感知需要进行权衡。为了在有限的资源和时隙内最大化系统的整体性能，在每个时隙内，UAV需要根据当前的飞行轨迹和波束形成，从所有潜在的感知目标中选择一个最优的目标进行感知，更好地平衡通信和感知之间的资源分配，提高系统的整体性能。例如，当UAV靠近通信用户时，可以优先选择靠近的目标进行感知，以减少路径损耗，增强感知信号的强度，在保证通信用户需求的同时，最大化感知目标的覆盖范围。假设在每个时隙UAV仅感知一个目标，$\alpha_m[n]=1$表示UAV在第n个时隙感知目标m；否则，$\alpha_m[n]=0$。基于上述讨论，以下条件成立：

$$\sum_{m=1}^{M}\alpha_m[n]\leqslant 1 \qquad \alpha_m[n]\in\{0,1\} \tag{18}$$

为确保感知性能，UAV在朝向目标的特定方向上发射的波束图增益必须满足一定的阈值，以确保感知信号能够被有效接收和处理。因此，在第n个时隙，UAV到感知目标m的发射波束图增益满足以下约束：

$$\alpha_m[n]\vartheta(\boldsymbol{q}[n],m)\geqslant\alpha_m[n]d_{\text{uav}\to m}^2\varsigma \tag{19}$$

式中：$d_{\text{uav}\to m}$为UAV与感知目标m之间的距离；ς为保证感知的阈值。

1.2 能耗模型

在任务周期T内，UAV的总能耗由感知能耗、通信能耗和飞行能耗组成，具体表示为

$$E_{\text{sum}}=E_{\text{com}}+E_{\text{sen}}+E_{\text{fly}} \tag{20}$$

1）通信能耗E_{com}。UAV飞行过程期间，对地面通信用户进行数据传输所消耗的能量表示为

$$E_{\text{com}}=\sum_{k=1}^{K}\sum_{n=1}^{N}\boldsymbol{w}_k^{\text{H}}[n]\boldsymbol{w}_k[n] \tag{21}$$

2）感知能耗E_{sen}。UAV在飞行过程期间，对潜在目标感知消耗的能量表示为

$$E_{\text{sen}}=\sum_{n=1}^{N}\delta_t\text{tr}(\boldsymbol{R}_{\text{s}}[n]) \tag{22}$$

3）飞行能耗E_{fly}。保证UAV飞行消耗的能量，在第n个时隙UAV飞行消耗的功率为

$$P(\boldsymbol{v}[n])=a\left(1+\frac{3\|\boldsymbol{v}[n]\|^2}{U_{\text{tip}}}\right)+b\left(\sqrt{1+\frac{\|\boldsymbol{v}[n]\|^4}{4v_1^4}}-\frac{\|\boldsymbol{v}[n]\|^2}{2v_1^2}\right)^{\frac{1}{2}}+\frac{1}{2}D\rho_1\rho_2F\|\boldsymbol{v}[n]\|^3 \tag{23}$$

式中：a和b为与UAV自身硬件材料相关的参数；U_{tip}为UAV转子的叶尖速度；v_1为平均转子速度；D为机身阻力比；ρ_1和ρ_2分别为空气密度和转子密度；F为转子盘面积。

因此，整个飞行周期内的UAV飞行总能耗可表示为

$$E_{\text{fly}}=\sum_{n=1}^{N}\delta_tP(\boldsymbol{v}[n]) \tag{24}$$

2 问题的建立与解决

为确保系统的通信与感知功能，通过联合优化UAV的感知目标选择、波束成形、轨迹和速度，在保证用户的最低传输速率的前提下最大化平均吞吐量，该联合优化问题P1可以表示为

$$\text{P1}:\max_{\{\boldsymbol{w}_k[n],\alpha_m[n],\boldsymbol{q}[n],\boldsymbol{v}[n]\}}R_{\text{c}}$$
$$\text{s.t.}\begin{cases}\text{C1}:\|\boldsymbol{q}[n+1]-\boldsymbol{q}[n]\|\leqslant v_{\max}\delta_t,1\leqslant n\leqslant N-1\\ \text{C2}:\boldsymbol{q}[1]=\boldsymbol{q}_{\text{I}}\\ \text{C3}:\boldsymbol{q}[N]=\boldsymbol{q}_{\text{F}}\\ \text{C4}:v_{\min}\leqslant\|\boldsymbol{v}[n]\|\leqslant v_{\max}\\ \text{C5}:R_k\geqslant R_{\min}\\ \text{C6}:\alpha_m[n]\vartheta(\boldsymbol{q}[n],m)\geqslant\alpha_m[n]d_{\text{uav}\to m}^2\varsigma\\ \text{C7}:\sum_{m=1}^{M}\alpha_m[n]\leqslant 1,\forall n\\ \text{C8}:\alpha_m[n]\in\{0,1\}\\ \text{C9}:E_{\text{com}}+E_{\text{sen}}+E_{\text{fly}}\leqslant P_{\max}T\end{cases} \tag{25}$$

式中：C1为轨迹约束；C2和C3为飞行始末位置；C4表示UAV的速度约束；C5表示至少满足用户k最低的吞吐量，$R_{\min}$为每个用户的最小吞吐量；C6为保证感知性能的约束；C7和C8保证正确感知目标的调度约束；C9为UAV的能耗约束，$P_{\max}$为UAV的最大发射功率。由于P1中多变量耦合、非凸问题，直接求解最优解比较困难。

为解决原问题，本文提出了一种基于BCD的优化迭代算法。通过固定变量策略，逐一优化各个子问题，并通过交替迭代直至算法收敛。联合优化问题P1被有效地分解成3个子问题：感知目标选择优化、波束形成优化及UAV的轨迹和速度优化。为方便处理，定义$\boldsymbol{\alpha}=[\alpha_m[n],\forall m,n]$，$\boldsymbol{w}=[\boldsymbol{w}_k[n],\forall k,n]$，$\boldsymbol{q}=\left[\boldsymbol{q}[n],\forall n\right]$，$\boldsymbol{v}=[\boldsymbol{v}[n],\forall n]$。

2.1 感知目标选择

固定飞行轨迹$\boldsymbol{q}$、速度$\boldsymbol{v}$和波束成形$\boldsymbol{w}$，问题P1转变为关于变量$\boldsymbol{\alpha}$的子问题P2，即

$$\text{P2}:\max_{\boldsymbol{\alpha}}R_{\text{c}}$$
$$\text{s.t.}\begin{cases}\text{C1}:R_k\geqslant R_{\min}\\ \text{C2}:\alpha_m[n]\vartheta(\boldsymbol{q}[n],m)\geqslant\alpha_m[n]d_{\text{uav}\to m}^2\varsigma\\ \text{C3}:\sum_{m=1}^{M}\alpha_m[n]\leqslant 1,\forall n\\ \text{C4}:\alpha_m[n]\in\{0,1\}\end{cases} \tag{26}$$

在子问题 P2 中, 由于 C4 中存在一个二元约束, 可将$\alpha_m[n]$转换成连续变量的形式, 即$0 \leqslant \alpha_m[n] \leqslant 1$。子问题 P2 是一个线性规划问题, 因此, 可以直接使用凸优化工具 CVX 进行求解。

2.2　波束成形优化

由 2.1 节可知, 感知目标选择$\boldsymbol{\alpha}$, 同时固定飞行轨迹$\boldsymbol{q}$和速度$\boldsymbol{v}$, 原问题转变为关于波束成形矢量$\boldsymbol{w}$的子问题 P3, 即

$$\mathrm{P3}: \max_{\boldsymbol{w}} \frac{1}{N}\left(\sum_{k=1}^{K} R_k\right)$$
$$\text{s.t.}\begin{cases} \mathrm{C1}: R_k \geqslant R_{\min} \\ \mathrm{C2}: \alpha_m[n]\vartheta(\boldsymbol{q}[n], m) \geqslant \alpha_m[n] d_{\mathrm{uav}\to m}^2 \varsigma \\ \mathrm{C3}: E_{\mathrm{com}} \leqslant P_{\max} T - E_{\mathrm{fly}} - E_{\mathrm{sen}} \end{cases} \tag{27}$$

由于 P3 的R_k同时在目标函数和约束条件 C1 中, 存在耦合且是非凸形式, 为解决R_k的非凸性, 通过 SCA 技术[18] 处理。根据式 (8)、式 (9) 和式 (11), 可将约束条件 C1 改写为

$$R_{k,j}[n] = \log_2\left(1 + \frac{\left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_k[n]\right\|^2}{\sum\limits_{i>k}^{K}\left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_i[n]\right\|^2 + \sigma_n^2}\right) =$$
$$R_1 - R_2 \geqslant R_{\min} \tag{28}$$

式中:

$$\begin{cases} R_1 = \log_2\left(\sum\limits_{i\geqslant k}^{K}\left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_i[n]\right\|^2 + \sigma_n^2\right) \\ R_2 = \log_2\left(\sum\limits_{i>k}^{K}\left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_i[n]\right\|^2 + \sigma_n^2\right) \end{cases} \tag{29}$$

R_2为相对于$\left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_i[n]\right\|^2$的凹函数, 将$R_2$在任意点进行一阶泰勒展开, 得到全局上界。因此, 在第l次迭代中, R_2的全局上界可以表示为

$$R_2 \leqslant A[n]\left(\left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_i[n]\right\|^2 - \left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_i^{(l)}[n]\right\|^2\right) +$$
$$\sum_{i>k}^{K}\left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_i^{(l)}[n]\right\|^2 + \sigma_n^2 = R_2^{\mathrm{up}} \tag{30}$$

式中: $\boldsymbol{w}_i^{(l)}[n]$为在第$l$次迭代中给定的波束形成向量; $A[n]$具体表达为

$$A[n] = \frac{\log_2 e}{\sum\limits_{i>k}^{K}\left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_i^{(l)}[n]\right\|^2 + \sigma_n^2} \tag{31}$$

$\left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_i[n]\right\|^2$为关于$\boldsymbol{w}_i[n]$的凸函数, 因此, 对$\left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_i[n]\right\|^2$进行一阶泰勒展开, 得到其在定点$\boldsymbol{w}_i^{(l)}[n]$的下界为

$$\left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_i[n]\right\|^2 \geqslant \left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_i^{(l)}[n]\right\|^2 +$$
$$2\mathrm{Re}\left\{\boldsymbol{w}_i^{(l)\mathrm{H}}\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{h}_j[n]\left(\boldsymbol{w}_i[n] - \boldsymbol{w}_i^{\mathrm{u}}[n]\right)\right\} \tag{32}$$

式中: Re表示取复数的实数部分。

问题 P3 涉及非凸约束, 可以通过上述处理将 P3 转换为更容易解决的子问题 P3.1, 具体如下:

$$\mathrm{P3.1}: \max_{\boldsymbol{w}} \frac{1}{N}\left(\sum_{k=1}^{K}\sum_{n=1}^{N} R_1 - R_2^{\mathrm{up}}\right)$$
$$\text{s.t.}\begin{cases} \mathrm{C1}: \sum\limits_{n=1}^{N} R_1 - R_2^{\mathrm{up}} \geqslant R_{\min} \\ \mathrm{C2}: \alpha_m[n]\vartheta(\boldsymbol{q}[n], m) \geqslant \alpha_m[n] d_{\mathrm{uav}\to m}^2 \varsigma \\ \mathrm{C3}: E_{\mathrm{com}} \leqslant P_{\max} T - E_{\mathrm{fly}} - E_{\mathrm{sen}} \\ \mathrm{C4}: \left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_i[n]\right\|^2 \geqslant \left\|\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{w}_i^{(l)}[n]\right\|^2 + \\ \quad 2\mathrm{Re}\{\boldsymbol{w}_i^{(l)\mathrm{H}}\boldsymbol{h}_j^{\mathrm{H}}[n]\boldsymbol{h}_j[n](\boldsymbol{w}_i[n] - \boldsymbol{w}_i^{\mathrm{u}}[n])\} \end{cases} \tag{33}$$

子问题 P3.1 是标准的凸优化问题, 可以使用 CVX 直接进行求解。

2.3　UAV 飞行轨迹和速度优化

已知最佳波束成形$\boldsymbol{w}$和感知目标调度$\boldsymbol{\alpha}$, 关于变量$\boldsymbol{q}$、$\boldsymbol{v}$的子问题 P4 表述为

$$\mathrm{P4}: \max_{\boldsymbol{q},\boldsymbol{v}} R_{\mathrm{c}}$$
$$\text{s.t.}\begin{cases} \mathrm{C1}: \|\boldsymbol{q}[n+1] - \boldsymbol{q}[n]\| \leqslant \boldsymbol{v}_{\max}\delta_t, 1 \leqslant n \leqslant N-1 \\ \mathrm{C2}: \boldsymbol{q}[1] = \boldsymbol{q}_{\mathrm{I}} \\ \mathrm{C3}: \boldsymbol{q}[N] = \boldsymbol{q}_{\mathrm{F}} \\ \mathrm{C4}: v_{\min} \leqslant \|\boldsymbol{v}[n]\| \leqslant v_{\max} \\ \mathrm{C5}: R_k \geqslant R_{\min} \\ \mathrm{C6}: \alpha_m[n]\vartheta(\boldsymbol{q}[n], m) \geqslant \alpha_m[n] d_{\mathrm{uav}\to m}^2 \varsigma \\ \mathrm{C7}: E_{\mathrm{com}} \leqslant P_{\max} T - E_{\mathrm{fly}} - E_{\mathrm{sen}} \end{cases} \tag{34}$$

在子问题 P4 中, 约束条件 C5、C6 是非凸的, 利用基于信赖域的 SCA 方法处理。信赖域方法定义了在当前迭代一个区域内, 信赖模型可以充分地表示目标函数, 选择合适的迭代方向和步长作为模型在该区域中的最小值的近似。实质上, 信赖域同时选择方向和步长。如果迭代不可接受, 将会减小域的大小, 再找一个新的最小值。一般来说, 每当信赖域的大小改变时, 迭代步的方向就会改变, 通过给定的信赖域半径得到该问题的一个局部最优点, 若使目标函数值增加, 则减小信赖域半径继续迭代, 再次进行求解。重复迭代过程, 直到满足设定阈值条件, 保证收敛。

定义$\boldsymbol{T}_1[n] = \sum\limits_{k=1}^{K}\boldsymbol{w}_i[n]\boldsymbol{w}_i^{\mathrm{H}}[n] + \boldsymbol{R}_{\mathrm{s}}[n]$, $\boldsymbol{W}_i[n] = \boldsymbol{w}_i[n]\cdot\boldsymbol{w}_i^{\mathrm{H}}[n]$, 并将$\boldsymbol{W}_i[n]$、$\boldsymbol{T}_1[n]$、$\boldsymbol{R}_{\mathrm{s}}[n]$在第$p$行第$q_1$列的大小表示为$[\boldsymbol{W}_i[n]]_{p,q_1}$, $[\boldsymbol{T}_1[n]]_{p,q_1}$, $[\boldsymbol{R}_{\mathrm{s}}[n]]_{p,q_1}$。

下面对 C5 进行化简, 先将式 (7) 代入式 (10), 可得

$$R_k[n] = \log_2\left(1 + \frac{\dfrac{\beta_0|\boldsymbol{a}_k^{\mathrm{H}}[n]\boldsymbol{w}_k[n]|^2}{\|\boldsymbol{q}[n] - \boldsymbol{g}_k\|^2 + H^2}}{\sum\limits_{i>k}^{K}\dfrac{\beta_0|\boldsymbol{a}_k^{\mathrm{H}}[n]\boldsymbol{w}_k[n]|^2}{\|\boldsymbol{q}[n] - \boldsymbol{g}_k\|^2 + H^2} + \sigma_n^2}\right) \tag{35}$$

对式(35)进一步化简可得

$$\overset{*}{R}_k[n]=\log_2\left(\sum_{i\geqslant k}^{K}\mu+\frac{\sigma_n^2}{\beta_0}\|\boldsymbol{q}[n]-\boldsymbol{g}_k\|^2\right)-\log_2\left(\sum_{i>k}^{K}\mu+\frac{\sigma_n^2}{\beta_0}\|\boldsymbol{q}[n]-\boldsymbol{g}_k\|^2\right) \tag{36}$$

式中：

$$\mu=\sum_{\alpha_1=1}^{N_t}[\boldsymbol{W}_i[n]]_{\alpha_1,\alpha_1}+2\sum_{p=1}^{N_t}\sum_{q_1=p+1}^{N_t}|\boldsymbol{W}_i[n]_{p,q_1}|\cdot\cos\left(\theta_{p,q_1}^{\boldsymbol{W}_i}+2\pi\frac{d}{\lambda}(q_1-p)\frac{H}{\sqrt{\|\boldsymbol{q}[n]-\boldsymbol{g}_k\|^2}}\right) \tag{37}$$

式中：$\theta_{p,q_1}^{\boldsymbol{W}_i}$为$[\boldsymbol{W}_i[n]]_{p,q_1}$对应的相角；$d$为阵列中相邻天线之间的距离；$\lambda$为载波波长。

对于式(36)，$\overset{*}{R}_k[n]$是关于$\boldsymbol{q}[n]$的非凸函数，但相对于$\|\boldsymbol{q}[n]-\boldsymbol{g}_k\|^2$是凸函数，根据凸函数的性质可得，$\overset{*}{R}_k[n]$在给定点$\boldsymbol{q}^{(l)}[n]$的全局下界为

$$\overset{*}{R}_k[n]\geqslant B_k^{(l)}[n]-C_k^{(l)}[n]\left\|\boldsymbol{q}[n]-\boldsymbol{q}^{(l)}[n]\right\| \tag{38}$$

式中：$B_k^{(l)}[n]$为$\overset{*}{R}_k[n]$在给定点$\boldsymbol{q}^{(l)}[n]$的值；$C_k^{(l)}[n]$为$\overset{*}{R}_k[n]$在给定点$\boldsymbol{q}^{(l)}[n]$处关于$\boldsymbol{q}[n]$的一阶导数，具体展开为

$$C_k^{(l)}[n]=\log_2\left(\frac{\|\boldsymbol{q}^{(l)}[n]-\boldsymbol{g}_k\|^2+\sum_{i\geqslant k}^{K}\mu_1}{\left(\sum_{i\geqslant k}^{K}\mu+\frac{\sigma_n^2}{\beta_0}\|\boldsymbol{q}^{(l)}[n]-\boldsymbol{g}_k\|^2\right)\ln 2}\right)-\log_2\left(\frac{\|\boldsymbol{q}^{(l)}[n]-\boldsymbol{g}_k\|^2+\sum_{i>k}^{K}\mu_1}{\left(\sum_{i>k}^{K}\mu+\frac{\sigma_n^2}{\beta_0}\|\boldsymbol{q}^{(l)}[n]-\boldsymbol{g}_k\|^2\right)\ln 2}\right) \tag{39}$$

式中：

$$\mu_1=2\sum_{p=1}^{N_t}\sum_{q_1=p+1}^{N_t}|\boldsymbol{W}_i[n]_{p,q_1}|\frac{2\pi dH(q_1-p)}{\lambda\|\boldsymbol{q}^{(l)}[n]-\boldsymbol{g}_k\|^3}\cdot\sin\left(\theta_{p,q_1}^{\boldsymbol{W}_i}+2\pi\frac{d}{\lambda}(q_1-p)\frac{H}{\sqrt{\|\boldsymbol{q}^{(l)}[n]-\boldsymbol{g}_k\|^2}}\right) \tag{40}$$

根据Hessian矩阵可知，$\hat{R}_k$相对$\boldsymbol{q}[n]$是凸的。

对于约束条件C6，将式(17)代入后，类似式(38)，将非凸约束重新表示为

$$\alpha_m[n]\left\{\sum_{\alpha_1=1}^{M}[\boldsymbol{T}_1[n]]_{\alpha_1,\alpha_1}+2\sum_{p=1}^{M}\sum_{q_1=p+1}^{M}|\boldsymbol{T}_1[n]_{\alpha_1,\alpha_1}|\cdot\cos\left(\theta_{p,q_1}^{\boldsymbol{T}_1}[n]+\frac{2\pi d(q_1-p)H}{\lambda\sqrt{d_{\text{uav}\to m}^2}}\right)\right\}\geqslant\alpha_m[n]d_{\text{uav}\to m}^2\boldsymbol{\varsigma} \tag{41}$$

对于式(41)中的非凸约束，可以基于一个局部给定点$\boldsymbol{q}^{(l)}[n]$进行一阶泰勒展开来近似式(41)左边部分，即将式(41)重新表示为

$$D_m^{(l)}[n]+Z_m^{(l)}[n]\left\|\boldsymbol{q}^{(l)}[n]-\boldsymbol{q}[n]\right\|\geqslant\alpha_m[n]d_{\text{uav}\to m}^2\boldsymbol{\varsigma} \tag{42}$$

式中：

$$D_m^{(l)}=\sum_{\alpha_1=1}^{M}[\boldsymbol{T}_1[n]]_{\alpha_1,\alpha_1}+2\sum_{p=1}^{M}\sum_{q_1=p+1}^{M}|\boldsymbol{T}_1[n]_{\alpha_1,\alpha_1}|\cdot\cos\left(\theta_{p,q_1}^{\boldsymbol{T}_1}[n]+\frac{2\pi d(p-q_1)H}{\lambda\sqrt{\|\boldsymbol{q}[n]-\boldsymbol{g}_m\|^2}}\right)\alpha_m[n] \tag{43}$$

$$Z_m^{(l)}[n]=\sum_{p=1}^{M}\sum_{q_1=p+1}^{M}|\boldsymbol{T}_1[n]_{p,q_1}|\cdot\frac{4\pi dH(q_1-p)}{\lambda\|\boldsymbol{q}^{(l)}[n]-\boldsymbol{g}_k\|^3}\cdot\sin\left(\theta_{p,q_1}^{\boldsymbol{W}_i}+\frac{2\pi H(q_1-p)d}{\lambda\sqrt{\|\boldsymbol{q}^{(l)}[n]-\boldsymbol{g}_k\|^2}}\right)\alpha_m[n] \tag{44}$$

综合以上分析，将问题P4中的非凸约束C5和C6，用式(38)和式(42)代替，得到近似的凸约束。此外，为保证近似精度，对原问题施加一个信任约束：

$$\|\boldsymbol{q}^{(l)}[n]-\boldsymbol{q}^{(l-1)}[n]\|\leqslant\xi^{(l)}\qquad\forall n\in N \tag{45}$$

式中：ξ为信赖域的半径。若该半径下未得到正确的目标值，则将半径缩小为$\xi/2$。从理论上讲，如果信赖域的半径ξ选择得足够小，那么迭代的收敛性总是可以确保的。

对于非凸约束C4引入一个松弛变量β_n，使得

$$\|\boldsymbol{v}[n]\|^2\geqslant\beta_n^2 \tag{46}$$

$$\beta_n\geqslant v_{\min} \tag{47}$$

根据一阶泰勒展开原理可得$\boldsymbol{v}[n]$的下界为

$$\|\boldsymbol{v}[n]\|^2\geqslant\|\bar{\boldsymbol{v}}[n]\|^2=\|\boldsymbol{v}^{(l)}[n]\|^2+2\boldsymbol{v}^{(l)}[n](\boldsymbol{v}[n]-\boldsymbol{v}^{(l)}[n]) \tag{48}$$

将松弛变量β代入第n个时隙UAV飞行消耗的功率，对于约束C7引入另一个松弛变量$\kappa_n^2\geqslant\sqrt{1+\frac{\beta_n^4}{4v_1^4}}-\frac{\beta_n^2}{2v_1^2}$，经化简可得：$\kappa_n^2+\frac{\beta_n^2}{v_1^2}\geqslant\frac{1}{\kappa_n^2}$，对于任意给定点$\{\kappa_n^{(l)},\beta_n^{(l)}\}$，将$\kappa_n^2+\beta_n^2/v_1^2$经过一阶泰勒展开可得相对于$\kappa$和$\beta$都是凸的下界，可表示为

$$\Gamma_n=(\kappa_n^{(l)})^2+\kappa_n^{(l)}(\kappa_n-\kappa_n^{(l)})+\frac{\beta_n^{(l)}}{v_1^2}+\frac{2\beta_n^{(l)}}{v_1^2}(\beta_n-\beta_n^{(l)}) \tag{49}$$

因此，可将UAV的飞行功耗近似为

$$\hat{E}_{\text{fly}}=\sum_{n=1}^{N}\delta_t\left[a\left(1+\frac{3\beta_n^2}{U_{\text{tip}}}\right)+b\kappa_n+\frac{1}{2}D\rho_1\rho_2F\beta_n^3\right] \tag{50}$$

经上述化简，可以将原问题P4重新表述为子问题P4.1：

$$\text{P4.1}: \max_{\boldsymbol{q},\boldsymbol{v}} R_{\text{sum}} = \frac{1}{N}\sum_{k=1}^{K}\sum_{n=1}^{N}\hat{R}_k[n]$$

$$\text{s.t.}\begin{cases} \text{C1}: \|\boldsymbol{q}[n+1]-\boldsymbol{q}[n]\| \leqslant \boldsymbol{v}_{\max}\delta_t, 1\leqslant n\leqslant N-1 \\ \text{C2}: \boldsymbol{q}[1]=\boldsymbol{q}_{\text{I}} \\ \text{C3}: \boldsymbol{q}[N]=\boldsymbol{q}_{\text{F}} \\ \text{C4}: \|\boldsymbol{v}[n]\| \leqslant v_{\max} \\ \text{C5}: \|\overline{\boldsymbol{v}}[n]\| \geqslant \beta_n^2 \\ \text{C6}: \sum_{n=1}^{N}\hat{R}_k[n] \geqslant R_{\min} \\ \text{C7}: \hat{E}_{\text{fly}} \leqslant P_{\max}T - E_{\text{com}} - E_{\text{sen}} \\ \text{C8}: \begin{cases} D_m^{(l)}[n] + Z_m^{(l)}[n]\left\|\boldsymbol{q}^{(l)}[n]-\boldsymbol{q}[n]\right\| \geqslant \\ \qquad \alpha_m[n]d_{\text{uav}\to m}^2\varsigma \\ \|\boldsymbol{q}^{(l)}[n]-\boldsymbol{q}^{(l-1)}[n]\| \leqslant \xi^{(l)} \qquad \forall n\in N \end{cases} \end{cases} \tag{51}$$

子问题 P4.1 是标准的凸优化问题, 可以直接使用凸优化工具 CVX 进行求解。具体的 UAV 轨迹优化方法详述于算法 1 中。

算法 1　UAV 轨迹优化算法。

1. 初始化: $\boldsymbol{\alpha}^{\cdot}, \boldsymbol{w}^{\cdot}, \xi$; 迭代次数 $l=1$; 设定阈值 Δ; $\boldsymbol{q}^{(l-1)}[n]=\boldsymbol{q}^{\cdot}[n]$。

2. 重复。

3. 解决问题 P4.1 得 $\boldsymbol{q}^{(l)}[n]$ 和 $\boldsymbol{v}^{(l)}[n]$。

4. If 问题 P4 中 R_{sum} 增加, 则 $\boldsymbol{q}^{(l)}[n]=\boldsymbol{q}^{(l)\cdot}[n]$, $l=l+1$

5. else $\xi^{(l)}=\xi^{(l)}/2$

6. End if

Until $\xi^{(l)} \leqslant \Delta$

7. 更新 $\boldsymbol{q}^{\cdot}[n]=\boldsymbol{q}^{(l)}[n]$。

8. Until $R_{\text{sum}}^{(l)} - R_{\text{sum}}^{(l-1)} \leqslant \Delta$, 结束并输出结果。

3　算法设计与分析

对式 (26)、式 (33)、式 (51) 分别求得问题 P2、P3、P4 的解, 交替迭代优化算法通过循环迭代不断更新求得 $\boldsymbol{\alpha}$、$\boldsymbol{w}$、$\boldsymbol{q}$、$\boldsymbol{v}$ 以至收敛。基于 BCD 的联合轨迹优化和波束形成方案的交替迭代优化如算法 2 所示。

算法 2　基于 BCD 的联合轨迹和波束形成算法。

初始化: 设置收敛精度 ε; 最大迭代次数 $Z_{\max}$; 迭代次数 $z=0$。

1. 根据式 (14) 计算 R_{c}, 定义 $R_{\text{c}}^{(-1)}=-\infty$。

2. 当 $R_{\text{c}}^{(z)} - R_{\text{c}}^{(z-1)} > \varepsilon$ 或 $z < Z_{\max}$。

3. 根据给定的 $\boldsymbol{q}^{(z-1)}$、$\boldsymbol{w}^{(z-1)}$、$\boldsymbol{v}^{(z-1)}$、$\boldsymbol{\alpha}^{(0)}$、$\boldsymbol{q}^{(0)}$、$\boldsymbol{w}^{(0)}$、$\boldsymbol{v}^{(0)}$, 解决式 (26) 得到子问题 P2 的感知目标选择解 $\boldsymbol{\alpha}^{(z)}$。

4. 根据给定的 $\boldsymbol{q}^{(z-1)}$、$\boldsymbol{v}^{(z-1)}$、$\boldsymbol{\alpha}^{(z)}$, 解决式 (33) 得到子问题 P3 的波束成形解 $\boldsymbol{w}^{(z)}$。

5. 根据给定的 $\boldsymbol{w}^{(z)}$、$\boldsymbol{\alpha}^{(z)}$, 解决式 (51) 得到子问题 P4 的 UAV 轨迹和速度解 $\boldsymbol{q}^{(z)}$、$\boldsymbol{v}^{(z)}$。

6. 根据式 (14) 计算 $R_{\text{c}}^{(z)}$。

7. 重复步骤 2, 当不满足步骤 2 要求时, 执行步骤 8。

8. 结束并输出结果。

9. 返回 $\boldsymbol{\alpha}^{\cdot}=\boldsymbol{\alpha}^{(z)}, \boldsymbol{w}^{\cdot}=\boldsymbol{w}^{(z)}, \boldsymbol{q}^{\cdot}=\boldsymbol{q}^{(z)}, \boldsymbol{v}^{\cdot}=\boldsymbol{v}^{(z)}$。

证明　算法 2 可得到 P1 的近似最优解并保证收敛。在第 z 次迭代过程中, 在已知波束成形 $\boldsymbol{w}^{(z-1)}$、UAV 的飞行轨迹 $\boldsymbol{q}^{(z-1)}$ 和速度 $\boldsymbol{v}^{(z-1)}$ 的基础上, 通过解决式 (26) 的线性规划问题, 确定感知用户选择 $\boldsymbol{\alpha}^{(z)}$, 且存在:

$$\Lambda(\boldsymbol{\alpha}^{(z)}, \boldsymbol{w}^{(z-1)}, \boldsymbol{q}^{(z-1)}, \boldsymbol{v}^{(z-1)}) \geqslant \Lambda(\boldsymbol{\alpha}^{(z-1)}, \boldsymbol{w}^{(z-1)}, \boldsymbol{q}^{(z-1)}, \boldsymbol{v}^{(z-1)})$$

式中: $\Lambda(\cdot)$ 为对应的系统平均吞吐量。

在已知 UAV 的飞行轨迹 $\boldsymbol{q}^{(z-1)}$ 和速度 $\boldsymbol{v}^{(z-1)}$、用户选择 $\boldsymbol{\alpha}^{(z)}$ 的情况下, 通过采用求解式 (33) 得到最优的波束成形方案 $\boldsymbol{w}^{(z)}$, 且存在:

$$\Lambda(\boldsymbol{\alpha}^{(z)}, \boldsymbol{w}^{(z)}, \boldsymbol{q}^{(z-1)}, \boldsymbol{v}^{(z-1)}) \geqslant \Lambda(\boldsymbol{\alpha}^{(z)}, \boldsymbol{w}^{(z-1)}, \boldsymbol{q}^{(z-1)}, \boldsymbol{v}^{(z-1)})$$

在最佳用户选择 $\boldsymbol{\alpha}^{(z)}$ 和最佳波束成形 $\boldsymbol{w}^{(z)}$ 已知情况下, 通过一阶泰勒展开、引入松弛变量, 基于信赖域的 SCA 等方法求解式 (51), 得到 UAV 的飞行轨迹 $\boldsymbol{q}^{(z)}$ 和速度 $\boldsymbol{v}^{(z)}$, 此时有

$$\Lambda(\boldsymbol{\alpha}^{(z)}, \boldsymbol{w}^{(z)}, \boldsymbol{q}^{(z)}, \boldsymbol{v}^{(z)}) \geqslant \Lambda(\boldsymbol{\alpha}^{(z)}, \boldsymbol{w}^{(z)}, \boldsymbol{q}^{(z-1)}, \boldsymbol{v}^{(z-1)})$$

因此, 在第 z 迭代中, 可以得到

$$\Lambda(\boldsymbol{\alpha}^{(z)}, \boldsymbol{w}^{(z)}, \boldsymbol{q}^{(z)}, \boldsymbol{v}^{(z)}) \geqslant \Lambda(\boldsymbol{\alpha}^{(z-1)}, \boldsymbol{w}^{(z-1)}, \boldsymbol{q}^{(z-1)}, \boldsymbol{v}^{(z-1)})$$

由此可见, 在给定精度阈值 ε 的前提下, 问题 P1 存在上界, 因此, 可以保证算法 1 的收敛性。

由于子问题是标准的非凸问题, 可以使用多项式复杂度内的内点法进行求解。其中, 算法 1 的复杂度为 $O\{L_1(KN)^{3.5}\}$, L_1 为算法 1 最大迭代次数。由算法 2 可知计算过程的复杂度集中在步骤 3～步骤 5。步骤 3 使用 CVX 求解, 计算复杂度为 $O\{(KN)^{25}\}$; 步骤 4 使用 CVX 和 SCA 方法求解, 其复杂度为 $O\{L_2(KN)^{3.5}\}$; 步骤 5 使用 CVX 和基于信赖域的 SCA 求解, 其计算复杂度为 $O\{L_3(L_1(KN)^{3.5})+KN\}$, L_2 和 L_3 分别为步骤 4 和步骤 5 中进行 SCA 处理的次数。因此, 算法 2 总的复杂度为 $O\{(L_3L_1+L_2+1)(KN)^{3.5}+KN\}\ln(1/\varepsilon)$, ε 为给定的收敛精度。同时, 如果使用穷举法解决所提问题, 则穷举法的复杂度为 $O\{(K+M)Q^N[(L_3L_1+L_2+1)(KN)^{3.5}+KN]\}[\ln(1/\varepsilon)]^2$, 其中, Q 为每个时隙内生成分布位置的可能情况, 由于通信用户和感知目标的位置在一定范围内随机生成, 可知 $Q \gg 1$, 对比可知, 所提方案计算复杂度得到有效降低。

4 仿真分析

为验证本文方案的有效性，设置如下实验场景：配备4×4均匀平面天线阵列的UAV为分布在300 m×300 m区域内的4个单天线用户提供通信服务，并同时进行了对该区域内2个目标的感知任务。通信用户与感知目标的位置通过随机产生，用户坐标分别为[38,151] m, [107,200] m, [243,109] m, [281,228] m, 感知目标的坐标为[75,100] m, [230,180] m。其中，$H=100$ m, $P_{\max}=30$ dBm, $Z_{\max}=30$, $v_{\max}=30$ m/s, $d_{\mathrm{t}}=d_{\mathrm{r}}=\lambda_{\mathrm{c}}/2$, $\varepsilon=10^{-2}$, 接收信号的噪声功率$\sigma_n^2=-110$ dBm, 相对参考距离的信道增益$\beta_0=-60$ dB。

为验证本文方案的优越性，与另外3种基准方案作对比：

1）本文方案。联合优化感知目标选择$\boldsymbol{\alpha}$、波束成形$\boldsymbol{w}$、UAV轨迹$\boldsymbol{q}$和飞行速度$\boldsymbol{v}$的方案。

2）方案1。随机波束成形和感知目标选择情况下，联合优化UAV轨迹$\boldsymbol{q}$和速度$\boldsymbol{v}$的方案。

3）方案2。随机感知目标选择和UAV轨迹，UAV以低于$v_{\max}$的某一速度$\boldsymbol{v}$开始飞行，并以恒定的加速度加速飞行，当$|\boldsymbol{v}|=v_{\max}$时，以$v_{\max}$速度飞行至终点，在整个飞行过程仅优化波束成形$\boldsymbol{w}$的方案。

4）方案3。随机感知目标选择$\boldsymbol{\alpha}$、波束成形$\boldsymbol{w}$，UAV由起始位置径直飞向终止位置，在整个周期T内以飞行速度v_0匀速飞行的方案。

图2给出了$T=100$ s时不同方案下的UAV飞行轨迹。可以看出，相比于仅优化波束的方案2，在联合优化的本文方案和仅优化轨迹的方案1时，UAV飞行的轨迹都更倾向于靠近通信用户的分布。由于UAV轨迹优化算法仅可以得到某个时隙内的最优轨迹位置，但通过交替迭代优化算法可以得到更合理的位置选择，本文方案的飞行轨迹更加平滑且更加接近每个用户，可以保证用户的通信需求并获得尽可能大的最大平均吞吐量，更适用于UAV的实际运动。方案2仅对波束形成进行优化会导致UAV轨迹与用户位置出现较大的偏差，无法到达最佳通信位置为用户提供通信服务，从而影响系统的整体性能。

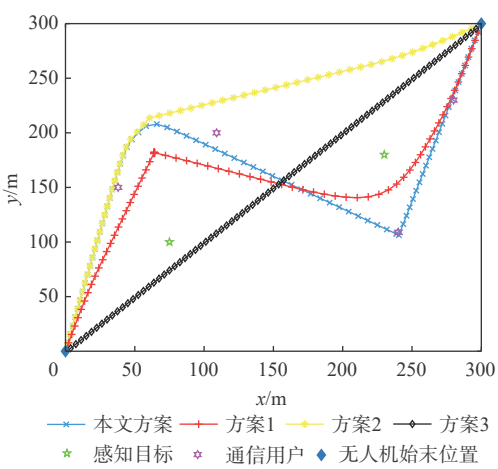


图2 UAV飞行轨迹

Fig. 2 UAV flight trajectory

图3给出了$T=100$ s时不同方案下最大平均吞吐量和迭代次数的关系。可以观察到，在4种不同的方案中，随着迭代次数的增加，系统的最大平均吞吐量也随之增加，在大约7次迭代之后，最大平均吞吐量趋于平稳。这一结果说明了本文方案的良好收敛性。同时，本文方案的最大平均吞吐量始终优于另外3种方案，性能分别提升了10.29%、13.80%、21.38%。

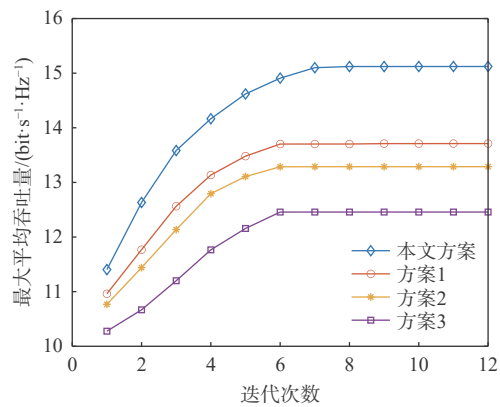


图3 不同方案下迭代次数与最大平均吞吐量的关系

Fig. 3 Relationship between number of iterations and maximum average throughput under different schemes

图4给出了$T=100$ s时4种方案下UAV最大发射功率与最大平均吞吐量的对比关系。随着发射功率增加，4种方案的最大平均吞吐量持续增长。本文方案的系统性能效果明显优于其他3种方案，因为通过联合UAV轨迹设计和感知目标选择优化，在对用户通信时保证每个时隙仅对一个目

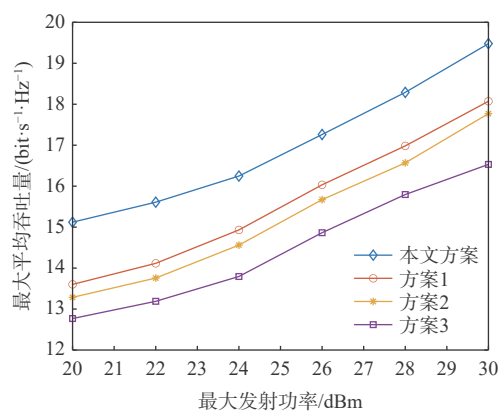


图4 不同方案下最大发射功率与最大平均吞吐量关系

Fig. 4 Relationship between maximum transmit power and maximum average throughput under different schemes

标进行感知, 节约了感知能量的消耗, 从而提高了通信性能。方案 1 的最大平均吞吐量高于方案 2 和方案 3, 说明通过 UAV 飞行轨迹和速度的设计, 可以节约飞行能耗, 提高系统的最大平均吞吐量。

图 5 给出了不同飞行周期下能耗的关系。可知, 随着飞行周期的延长, 系统的飞行能耗和感知能耗均呈现持续上升的趋势, 而通信能耗则在初期增加后趋于减少, 且飞行能耗远大于其他能耗。当飞行周期较小时, UAV 飞行速度快, 很难飞行至最佳位置完成通信和感知工作, UAV 的通信和感知能耗较小; 随着飞行周期的增加, UAV 可以飞行至更靠近用户和目标的位置, 通信与感知的能耗增加; 当飞行周期设置过长时, 由于感知任务持续时间增加, 感知能耗也会相应增加, 而 UAV 在此周期内可以以较低的速度更靠近通信用户完成通信任务, 减少通信过程中因信号衰减等因素造成的额外通信能耗, 将更多的能量用于目标感知。

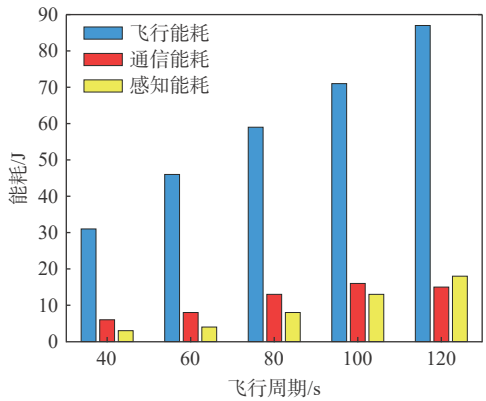


图 5　不同飞行周期下能耗的关系

Fig. 5　Relationship of energy consumption under different flight cycles

图 6 给出了 $T = 50$ s时不同方案下获得的发射波束图。当 $T = 50$ s时, 各方案下的 UAV 的位置、发射波束预计角度和最大峰值角度如表 1 所示。

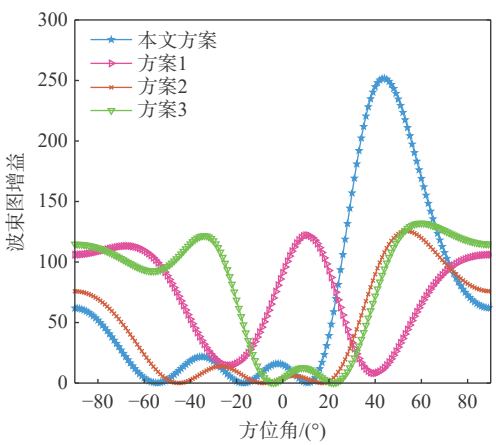


图 6　T=50 s 时的发射波束图

Fig. 6　Transmitted beam pattern obtained of T=50 s

表 1　不同方案获得的发射波束图角度

Table 1　Angle of transmitting beam pattern obtained by different schemes

方案	UAV位置/m	期望角度/(°)	实际角度/(°)	最大增益
本文方案	[142,165,100]	43.02	43.37	251.06
方案1	[80,177,100]	37.60	10.30	122.13
方案2	[90,223,100]	51.17	52.30	125.53
方案3	[150,150,100]	41.93	59.35	132.05

可知, 本文方案与方案 2 可在预计方向上基本实现发射波束图的主峰, 且本文方案峰值远高于方案 2, 这说明了本文所提联合优化方案在感知性能方面更具优势。方案 1 和方案 3 未对波束进行优化, 由于随机信道具有多径传播和信号衰减等复杂的信道特性, 会导致信号的干扰和衰减在不同方向和时间上表现出随 机性。因此, 未经优化的波束往往不能很好地适应信道的变化, 导致波束图呈现不规律的形状, 感知性能不佳。

图 7 给出了 UAV ISAC 系统中通信性能和感知性能的权衡。其中, 感知性能以有效感知功率为衡量指标。4 种方案均随着有效感知功率逐渐下降, 最大平均吞吐量不断增加。在相同的有效感知功率下, 本文方案的最大平均吞吐量始终大于其他方案, 且随着有效感知功率的减小, 4 种方案的最大平均吞吐量间的差值逐渐增大。这是因为其他 3 种方案在相同空间自由度情况下, 为了满足感知性能的需求, UAV 无法到达最佳的通信位置或获得最佳波束形成, 从而需要更多的资源来满足感知需求, 导致通信性能有限。

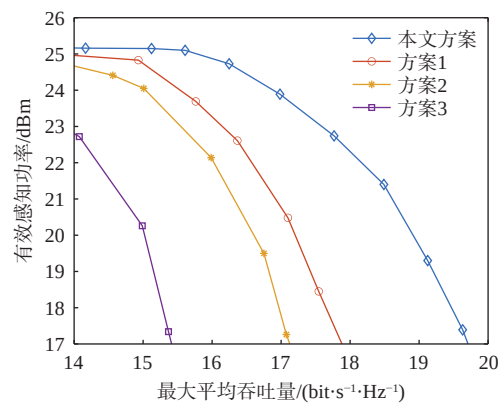


图 7　不同方案下有效感知功率与最大平均吞吐量的关系

Fig. 7　Relationship between effective sensing power and maximum average throughput under different schemes

图 8 给出了本文方案在不同用户数目和不同感知目标数的情况下最大平均吞吐量与最大发射功率的关系。可见, 相同用户数目的情况下, 随着感知目标数量增加, 最大平均吞吐量随之下降, 这

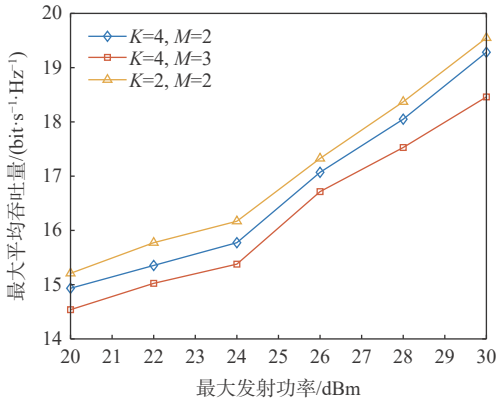


图8 不同用户数目和感知目标数下最大平均吞吐量与最大发射功率的关系

Fig. 8 Relationship between maximum average throughput and maximum transmit power under different number of users and perceived target number

是因为过多的感知目标数会增加系统的负担，导致资源竞争加剧，引起系统性能下降；而在感知目标数相同的情况下，通信用户数目减少会使得最大平均吞吐量增加，这是因为较少的用户数可以集中分配到更多的系统资源，从而提高各用户的吞吐量，用户数多反而会分散资源。

5 结 论

1）构建无人机与多用户多目标之间的通信感知一体化系统，所提算法具有良好的收敛性，可以7次迭代后收敛，并且得到的最大平均吞吐量优于其他方案。

2）在不同最大发射功率、不同用户数目和不同感知目标数的情况下，所提方案始终优于其他方案，可以提高系统的最大平均吞吐量。

3）在满足通信需求的同时，所提方案在满足系统通信需求的前提下，可以对潜在的目标实现正确感知，说明了本文所提系统的对通感性能权衡的有效性。

参考文献（References）

[1] BEKKERMAN I, TABRIKIAN J. Target detection and localization using MIMO radars and sonars[J]. IEEE Transactions on Signal Processing, 2006, 54(10): 3873-3883.

[2] LIU F, CUI Y H, MASOUROS C, et al. Integrated sensing and communications: toward dual-functional wireless networks for 6G and beyond[J]. IEEE Journal on Selected Areas in Communications, 2022, 40(6): 1728-1767.

[3] LIU X, HUANG T Y, SHLEZINGER N, et al. Joint transmit beamforming for multiuser MIMO communications and MIMO radar[J]. IEEE Transactions on Signal Processing, 2020, 68: 3929-3944.

[4] FORTUNATI S, SANGUINETTI L, GINI F, et al. Massive MIMO radar for target detection[J]. IEEE Transactions on Signal Processing, 2020, 68: 859-871.

[5] LIU F, ZHOU L F, MASOUROS C, et al. Toward dual-functional radar-communication systems: optimal waveform design[J]. IEEE Transactions on Signal Processing, 2018, 66(16): 4264-4279.

[6] HUA H C, XU J, HAN T X. Optimal transmit beamforming for integrated sensing and communication[J]. IEEE Transactions on Vehicular Technology, 2023, 72(8): 10588-10603.

[7] LIU F, MASOUROS C, PETROPULU A P, et al. Joint radar and communication design: applications, state-of-the-art, and the road ahead[J]. IEEE Transactions on Communications, 2020, 68(6): 3834-3862.

[8] CHEN L, LIU F, WANG W D, et al. Joint radar-communication transmission: a generalized Pareto optimization framework[J]. IEEE Transactions on Signal Processing, 2021, 69: 2752-2765.

[9] WANG B W, SUN Y J, SUN Z, et al. UAV-assisted emergency communications in social IoT: a dynamic hypergraph coloring approach[J]. IEEE Internet of Things Journal, 2020, 7(8): 7663-7677.

[10] LYU Z H, ZHU G X, XU J. Joint maneuver and beamforming design for UAV-enabled integrated sensing and communication[J]. IEEE Transactions on Wireless Communications, 2023, 22(4): 2424-2440.

[11] DENG C L, FANG X M, WANG X B. Beamforming design and trajectory optimization for UAV-empowered adaptable integrated sensing and communication[J]. IEEE Transactions on Wireless Communications, 2023, 22(11): 8512-8526.

[12] CHEN X, FENG Z Y, WEI Z Q, et al. Performance of joint sensing-communication cooperative sensing UAV network[J]. IEEE Transactions on Vehicular Technology, 2020, 69(12): 15545-15556.

[13] WEI Z Q, LIU F, NG D W K, et al. Safeguarding UAV networks through integrated sensing, jamming, and communications[C]//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE Press, 2022: 8737-8741.

[14] WANG Z L, LIU Y W, MU X D, et al. NOMA empowered integrated sensing and communication[J]. IEEE Communications Letters, 2022, 26(3): 677-681.

[15] MU X D, LIU Y W, GUO L, et al. NOMA-aided joint radar and multicast-unicast communication systems[J]. IEEE Journal on Selected Areas in Communications, 2022, 40(6): 1978-1992.

[16] 张睿智, 厉俊, 沈晓峰. 基于非正交多址接入的无人机通感一体化系统[J]. 信息与控制, 2023, 52(6): 701-710.
ZHANG R Z, LI J, SHEN X F. UAV integrated sensing and communication system based on non-orthogonal multiple access[J]. Information and Control, 2023, 52(6): 701-710(in Chinese).

[17] MENG K T, WU Q Q, MA S D, et al. Throughput maximization for UAV-enabled integrated periodic sensing and communication [J]. IEEE Transactions on Wireless Communications, 2023, 22(1): 671-687.

[18] RAZAVIYAYN M. Successive convex approximation: analysis and applications[D]. Minneapolis: University of Minnesota, 2014.

NOMA-based joint optimization of trajectory and resources for UAV-enable integrated sensing and communication system

TANG Jingmin[1]，HU Cheng[1]，SONG Yaolian[1,*]，YU Guicai[2]

(1. Faculty of Information Engineering and Automation，Kunming University of Science and Technology，Kunming 650500，China；
2. School of Physical Science and Engineering Technology，Yichun University，Yichun 336000，China)

Abstract：This study examines the difficulties of target selection, beamforming, and the combined optimization of UAV trajectory and speed inside the integrated sensing and communication (ISAC) architecture in order to address the shortage of wireless spectrum resources and the interference among multi-user signals. By leveraging non-orthogonal multiple access (NOMA) technology, the overall system throughput is enhanced. Given the non-convex nature of the problem, this research employs the block coordinate descent (BCD) method to decompose the intricate problem into three manageable subproblems. Utilizing relaxation variables, first-order Taylor approximation, and successive convex approximation (SCA), the method reduces computational complexity and improves efficiency. Alternating iterations of these subproblems yield an approximately optimal solution to the original problem. The suggested algorithm's effectiveness is confirmed by simulation results, which show how it may greatly increase maximum average throughput and exhibit strong convergence.

Keywords：integrated sensing and communication; unmanned aerial vehicle; non-orthogonal multiple access; trajectory optimization; resource optimization

Received：2024-04-29；**Accepted**：2024-06-28；**Published Online**：2024-07-17 16:49

URL：link.cnki.net/urlid/11.2625.V.20240715.1905.002

Foundation items：National Natural Science Foundation of China (62261056)；Jiangxi Provincial Natural Science Foundation (20212BAB202029)

* **Corresponding author**. E-mail：39217149@qq.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0815

基于数字孪生和 Transformer-LSTM-XGBoost 的液压机器人故障诊断方法

李亚洁[1,2]，吴瑞龙[1,2]，李炜[1,2,*]

(1. 兰州理工大学 自动化与电气工程学院，兰州 730050；2. 兰州理工大学 有色冶金智能机器人创新中心，兰州 730050)

摘 要：随着液压机器人的应用范围日益广泛，其健康诊断与维护需求也愈发迫切。通过将数字孪生与深度学习技术相结合，提出一种基于 Transformer、长短期记忆网络 (LSTM) 与极端梯度提升 (XGBoost) 算法相融合的液压机器人智能故障诊断方法。提出基于数字孪生技术的液压机器人故障诊断架构与方案；建立涵盖三维模型与属性模型且具备虚实同步能力的液压机器人数字孪生体，并研究了提升孪生模型精度的校准方法。针对常见的液压系统泄漏、阀门卡滞、阻尼孔堵塞、过滤器堵塞 4 种典型故障，在所建数字孪生体上进行故障机理分析与演化过程模拟，构建覆盖正常工况及多类故障工况的高质量数据集，用于支撑数据驱动建模与智能诊断。提出融合 Transformer 与 LSTM 的深度特征提取模型，以捕获多维特征间的全局依赖关系与时序动态特征，并引入 XGBoost 分类器实现多故障模式识别。实验验证结果表明：所提方法在多种液压系统故障诊断中可获得 96.6% 的诊断准确率，在不同噪声扰动条件下仍保持稳定的诊断性能，所提方法具有良好的鲁棒性与泛化能力，可为液压机器人故障诊断与维护提供有效的技术支撑。

关 键 词：数字孪生；液压机器人；故障诊断；深度学习；机器学习

中图分类号：TH165+.3; TP181; V245.1

文献标志码：A **文章编号**：1001-5965(2026)06-1850-19

随着智能制造水平的不断提升，液压机器人因其功率密度高、负载能力强、动态响应快、适应能力优等特点，已被广泛应用于工业制造、航空航天、深海探测和应急救援等领域[1]。然而，液压机器人长期运行于复杂多变的工况环境中，频繁遭遇负载持续波动、外部因素叠加干扰等挑战，导致系统元件磨损加剧、系统性能渐进退化，甚至会发生突发性的执行机构失效等问题[2]。随着液压机器人的驱动系统向轻量化、小型化、高压化及大功率化方向逐步发展，其结构日益复杂，元件集成度高，对系统运行的安全性、稳定性及智能运维能力提出了更高要求[3]。作为典型的液-机-电耦合系统，液压机器人的故障演变表现出高度的随机性、复杂性与隐蔽性，设备一旦发生故障，往往面临排查难、定位难、处置难等问题[4]。因此，改善机器人驱动系统的长期运行稳定性和故障诊断精准性，已然成为提升液压机器人系统性能亟需攻克的关键技术难题。

在液压机器人的故障诊断领域，传统方法大多依赖于传感监测与信号处理技术，常见做法包括傅里叶变换、小波分析、经验模态分解等信号处理手段，用以提取压力、流量、振动等参数中的故障特征[5-6]。除此之外，还可通过对油液污染度、颗粒度的物理化学特性进行分析，进而评估系统的健康状态[7]。虽然上述方法在实践中验证了其可靠性，然

收稿日期：2025-11-25；**录用日期**：2026-02-05；**网络出版时间**：2026-03-03 13:20

网络出版地址：link.cnki.net/urlid/11.2625.V.20260302.1522.002

基金项目：国家自然科学基金 (62163022)；甘肃省产业支撑计划项目 (2024CYZC-18)；甘肃省科技重大专项计划 (23ZDGA007)；甘肃省联合科研基金 (24JRRA827)；兰州市科技计划项目 (2024-2-5)

* **通信作者**. E-mail：liwei@lut.edu.cn

引用格式：李亚洁，吴瑞龙，李炜. 基于数字孪生和 Transformer-LSTM-XGBoost 的液压机器人故障诊断方法 [J]. 北京航空航天大学学报，2026，52（6）：1850-1868. LI Y J，WU R L，LI W. Fault diagnosis method for hydraulic robots based on digital twin and Transformer-LSTM-XGBoost [J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（6）：1850-1868（in Chinese）.

而，当其应对非线性、时变和多噪声等复杂工况时却存在一定局限性，特别是对液压机器人动态故障模式的诊断精度和实时性支持不足。因此，基于数据驱动的智能化故障诊断方法应运而生，其能够有效提高诊断精度和实时响应速度，成为解决此类问题的可靠手段。

近年来，随着人工智能技术的迅速发展，基于数据驱动的智能诊断方法逐渐成为机器人液压驱动系统故障诊断领域的重要研究方向。卷积神经网络 (convolutional neural networks, CNN)、循环神经网络（recurrent neural networks, RNN）、长短期记忆网络（long short-term memory networks, LSTM）等模型被广泛应用于故障特征的自动提取和模式识别[8]。Peng 等[9] 提出基于噪声深度 CNN 的旋转矢量减速器（rotary vector reducer, RV）多工况诊断方法；姜少飞等[10] 针对传统机器学习故障诊断精度不足的问题，提出基于极端梯度提升（extreme gradient boosting, XGBoost）的隐含特征提取与支持向量机（support vector machine, SVM）多故障模式识别方法；郑彪等[11] 提出基于深度神经网络（deep neural networks, DNN）与胶囊网络相结合的液压泵故障程度诊断方法，利用 DNN 提取特征并通过胶囊网络动态路由实现故障模式识别，有效提升了对液压泵多种故障的诊断精度；Sun 等[12] 提出基于多尺度一维 CNN 的智能故障诊断方法，进一步增强了液压系统的故障识别能力。尽管上述方法在特征提取与模式识别方面取得了显著进展，但其性能依赖于完整且高质量的训练数据，而在实际工程中，真实故障数据往往稀缺且难以获取，通过破坏性实验获取故障数据的成本偏高[13]。

液压机器人安全稳定性的提升与系统故障精准诊断构成了一个多维耦合的复杂系统工程，其突破性进展的取得将高度依赖于机械控制、流体传动、智能算法与状态监测等技术的深度交叉融合与协同创新。而传统的设计优化过程往往需要耗费大量的人力、物力和时间资源，在实际应用中面临兼顾效率与成本的双重压力[14]。

在此背景下，数字孪生（digital twin, DT）技术应运而生，凭借其独特优势为液压机器人的故障诊断提供了全新的技术路径。数字孪生通过构建与物理实体高度一致的虚拟信息模型，结合实时传感数据与历史运行数据，实现物理实体与虚拟模型的双向映射与动态交互，极大提升了对系统的综合管控水平[15]。与传统方法相比，数字孪生不仅能够实时监测设备状态，还能在虚拟环境中模拟典型故障场景，从而生成多源、多样化的高质量故障数据集，以此来弥补实际故障数据不足的难题[16]。此外，数字孪生还可用于揭示虚实差异，通过数据补偿与模型优化提升模型精度，使其在复杂工况下的预测与诊断能力更为可靠。已有研究表明，基于数字孪生的智能诊断框架能够在模型不确定性强与故障数据获取难的约束下显著提升诊断性能[17]。Liu 等[18] 提出基于数字孪生的 AI 仿真框架，帮助解决数据不足的问题。李亚洁等[19] 提出一种基于数字孪生流场云图的离心泵叶轮机械故障智能诊断方法，通过学习训练 Yolov5 算法，对离心泵叶轮进行故障诊断。

近年来，学者们在机器人液压驱动系统数字孪生建模与智能化故障诊断领域取得了诸多进展。李志祥等[20] 开发了基于数字孪生技术的液压机器人运动仿真与状态监控系统，实现了物理实体与虚拟系统的协同仿真。简正豪等[21] 提出基于数字孪生的液压回路仿真系统，借助精确虚拟建模和实时数据同步，实现了系统运行数据的采集、存储与三维可视化等功能。赵大伟等[22] 将数字孪生与知识图谱技术结合，构建了液压系统的知识模型，促进了知识的重用与可视化管理。Sun 等[23] 针对液压系统故障诊断过程中模型参数不确定性强和故障诊断精度低的问题，提出基于数字孪生技术的液压系统故障诊断方法。Wang 等[24] 提出一种将虚拟模型与物理实体相结合的数字孪生框架来解决液压系统的故障问题，为系统故障诊断提供了一种更为可行的解决方案。

尽管上述方法在特征提取与故障诊断方面取得了显著进展，但仍存在以下不足：①其性能依赖于大量标注数据，而在实际工程中真实故障数据稀缺、获取成本高，难以满足模型训练需求；②部分方法只能针对单一工况或有限故障类型，难以适应液压机器人在实际运行环境下所面临的多工况、多故障模式的复杂特性；③多数方法在多源传感信号的特征提取与融合建模方面仍存在不足，尤其未能有效捕捉不同信号间的高耦合特征关系与动态交互效应，使得模型在复杂环境下的稳定性和鲁棒性仍有较大的提升空间。

综上所述，针对传统方法在液压机器人典型故障识别中存在的难题，本文依托 Automation Studio™（AS）平台，面向五自由度机器人的液压驱动系统，提出了一种基于数字孪生与 Transformer-LSTM-XGBoost 相融合的智能化故障诊断方法，进而构建具有虚实同步能力的数字孪生模型，并基于孪生模型进行故障诊断方法的研究。本文的主要创新点总结如下：

1）本文提出一种基于数字孪生技术的液压机

器人故障诊断框架。通过构建支持虚实同步的数字孪生模型，实现了液压机器人在多种故障工况下的高保真虚拟仿真与动态行为模拟。为提升模型精度，本文还引入了基于实测数据的补偿机制，利用实际机器人运行数据对数字孪生模型进行动态修正与误差补偿，确保虚实同步的精度，从而在动态与稳态阶段准确反映实体机器人的运行特性。

2）本文依托机器人的数字孪生体，模拟了液压系统泄漏、阀门卡滞、阻尼孔堵塞和过滤器堵塞4种典型液压故障场景，避免了传统破坏性实验在获取故障数据时存在的高经济成本和实施风险问题。构建覆盖多工况、多类别的高质量故障样本数据集，显著提升了模型训练的泛化能力和诊断可靠性。

3）本文提出了一种融合Transformer、LSTM和XGBoost的智能故障诊断方法。在深度特征提取阶段，Transformer模块用于捕获多源信号的全局依赖关系与特征交互信息，LSTM模块则提取时序特征的动态演化规律，从而获得兼具全局与时序关联的深层特征表示；在决策阶段，XGBoost被用于对融合特征进行分类，精准识别多类别液压系统故障。实验结果表明，本文方法在复杂工况下具有较高的故障识别准确率及较强的鲁棒性，验证了该方法在液压机器人故障诊断中的有效性与工程适用性。

1 基于数字孪生的液压机器人故障诊断架构与方案

1.1 液压机器人数字孪生架构的建立

为实现液压机器人的精准监测、故障诊断与智能运维，本文基于数字孪生五维模型，设计并构建了由物理层、数据层、孪生层、连接层和服务层组成的液压机器人数字孪生五维架构[25]，如图1所示。

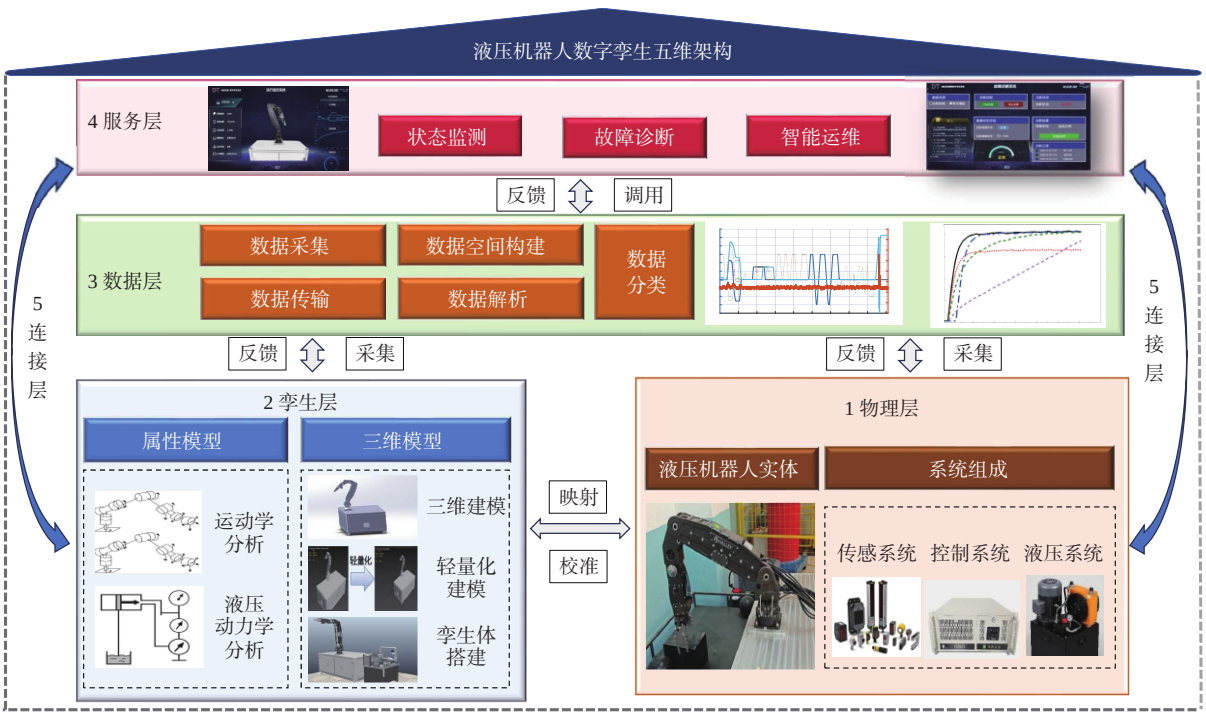


图1 液压机器人数字孪生五维架构

Fig. 1 Five-dimensional digital twin architecture of hydraulic robot

如图1所示，本文构建的液压机器人数字孪生故障诊断系统架构分为5个层次。

1）物理层。作为数字孪生系统的基础，其涵盖机器人液压驱动系统的实体设备、液压传动组件、控制器及各类传感器节点和数据采集模块。该层负责实时获取关键部件的运行状态与环境参数，是实现机器人数字孪生的核心数据来源。

2）孪生层。作为数字孪生系统的核心，其基于数字孪生技术构建与物理系统高度一致的虚拟仿真模型，综合考虑了三维建模、关节运动学分析与液压动力学分析等要素。

3）数据层。主要负责物理层采集数据的接收、存储、预处理与统一管理。该层集成数据库管理系统、数据清洗和同步模块，并向上层模型提供统一、高效的数据访问接口，有效保障数据的实时性与准确性。

4）服务层。作为数字孪生系统与用户间的交互界面，其主要负责展示机器人液压驱动系统的运行状态与故障诊断结果，并提供故障报警、趋势预测、远程维护建议等服务功能，为运维人员的科学决策

提供支持，是数字孪生技术的核心应用价值所在。

5）连接层。其负责各层之间的双向数据交互与同步，通过工业标准通信协议，实现物理层与孪生层、数据层与服务层之间的高效数据传输，确保虚拟模型准确反映真实系统状态。

1.2　基于数字孪生技术液压机器人故障诊断方案的提出

数字孪生技术为液压机器人运行状态的实时感知与行为分析提供了有力支撑。当液压机器人出现故障时，通常会导致其运动性能和工作状态发生改变，从而引发液压系统及其执行机构的压力、流量、位移等关键物理参数发生明显波动。依托数字孪生技术对这些实时数据进行动态监测与虚实对比分析，可为故障的快速识别与精准定位提供数据基础和理论支撑。

基于上述分析，本文提出一种基于液压机器人数字孪生技术的智能故障诊断方案，该方案依托液压系统实时监测数据与虚实同步仿真模型，结合状态监测、信号处理与深度学习等方法，构建系统化的故障诊断流程。具体方案如图 2 所示，主要包括数字孪生建模、故障数据模拟、数据采集处理、故障智能诊断及孪生系统实现 5 个部分。

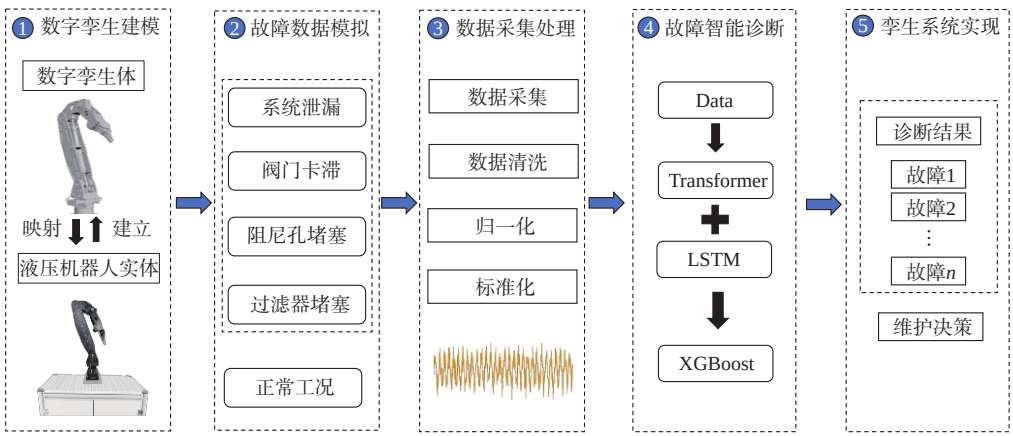


图 2　液压机器人故障诊断方案

Fig. 2　Fault diagnosis scheme for hydraulic robot

1）数字孪生建模。基于液压机器人的物理结构与液压系统特性，构建对应的数字孪生模型，实现对机器人动态行为的精准模拟。通过仿真分析对液压系统进行虚拟调试与性能测试，并结合真实运行数据对模型参数进行校准与优化，确保数字孪生模型能够准确反映实际系统的动态特性。该过程为故障诊断提供可对比的基准模型，有助于识别和提取潜在的故障特征。

2）故障数据模拟。为构建覆盖典型运行状态的多类别故障诊断模型，本文在数字孪生系统中设计并模拟了多种常见的机器人液压驱动系统故障场景，包括液压系统泄漏、阀门卡滞、阻尼孔阻塞、滤油器堵塞故障。在故障模拟过程中，系统持续采集各关键部件的模拟传感器数据，完整记录故障的发生、发展过程中的动态响应轨迹，构建多种运行工况下的高质量故障数据集。

3）数据采集处理。通过在液压机器人关键部位布设压力传感器与位移传感器，实时获取各运动构件的动态运行数据；同时，结合数字孪生模型在不同典型故障情况下生成的模拟数据，形成“物理实测+虚拟仿真”相融合的多源数据基础。为确保数据的准确性与可靠性，对采集到的原始信号进行预处理，包括噪声抑制、异常值剔除与数据校正等。

4）故障智能诊断。采用 Transformer-LSTM-XGBoost 融合架构，对采集的多源特征数据进行训练与建模，结合历史故障样本与已知故障模式，构建高精度智能故障诊断模型。该模型能够对液压机器人运行过程中的特征信号进行实时分析，准确识别故障类型并提前发出预警，从而实现对多类别液压故障的高效、可靠诊断。

5）孪生系统实现。在故障诊断模型完成实时分析并输出诊断结果后，系统通过服务层向操作员提供故障报警、故障定位、维护建议等信息，并生成详细的故障诊断报告。其中，维护建议指系统能够基于诊断模型的分类结果，结合历史故障模式和运维经验，为运维人员制定针对性的维护策略，从而提高液压机器人整体运行的可靠性与安全性。

2　液压机器人数字孪生模型的建立与故障机理分析

2.1　液压机器人的数字孪生模型建立

为构建高精度的液压机器人数字孪生模型，本文从三维建模、属性建模及模型校准 3 个方面进行

数字孪生模型建立。首先，利用 SolidWorks 进行三维几何建模，精确构建液压机器人结构及关键部件布局，为运动学与动力学分析提供基础。其次，基于改进的 D-H 参数法，建立关节间的运动学映射关系，实现末端执行器与关节运动参数的精确关联，确保高精度运动学分析。最后，采用基于实测数据的误差补偿方法，实时校准模型，通过动态计算修正误差，提升模型精度与稳定性。

2.1.1 液压机器人数字孪生三维建模

本文研究的液压机器人包含 5 个运动关节，末端配置有独立控制的夹爪机构，可支持多姿态抓取与高精度搬运任务。本文研究的液压驱动五自由度机器人的实体结构如图 3 所示，其基本参数如表 1 所示。

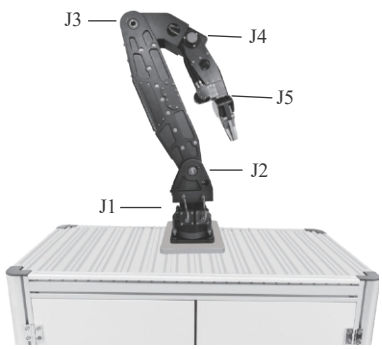


图 3 液压机器人实体图

Fig. 3 Physical hydraulic robot

表 1 液压机器人的基本参数

Table 1 Basic parameters of hydraulic robot

参数	数值
质量/kg	40
水平伸展范围/mm	1 330
向上伸展范围/mm	1 640
向下伸展范围/mm	293
抓举质量/kg	30
肩部转向范围/(°)	180
肩部俯仰范围/(°)	110
肘部俯仰范围/(°)	100
腕部俯仰范围/(°)	100
腕部回转范围/(°)	310
钳口张开范围/mm	100

机器人采用集中式液压泵站作为主动力源，通过高压液压管路为执行元件提供能量。液压油作为能量传递介质，通过比例伺服阀与控制单元协调，实现多关节的同步协同与动态响应。泵站采用电机与油泵直联结构，减少传动损耗，并集成换向阀、压力控制阀和流量阀，确保对系统压力和流量的精确调节。模块化油路设计降低了泄漏风险和系统复杂度，回油经过滤器净化后返回油箱，并通过冷却装置控制油温，保障系统稳定运行。

三维模型是数字孪生系统构建的几何基础，其建模精度与计算效率直接影响后续系统仿真验证、控制策略部署与虚拟联动调试的效率。为实现结构精度与运行效率的统一，本文的三维建模流程如图 4 所示。

图 4 液压机器人建模流程

Fig. 4 Hydraulic robot modeling process

1）为确保数字孪生系统在几何层面高度还原真实设备，使用 SolidWorks，依据液压机器人的设计参数和结构图纸，依次对各关节模块、臂段单元、末端夹爪进行精细建模，确保模型的几何尺寸和结构特性与实际设备高度一致。在完成各部件建模后，通过装配模块设定运动副及几何约束，构建同实体自由度与结构拓扑关系高度一致的完整装配体模型，具体模型如图 5 所示。

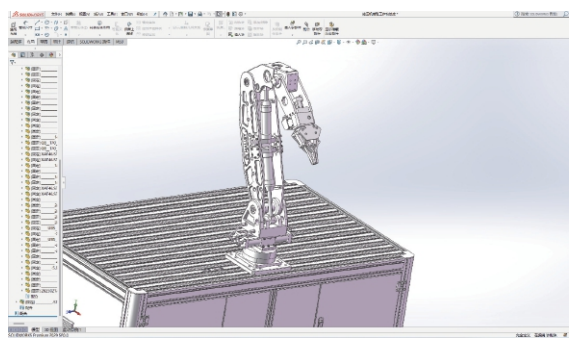
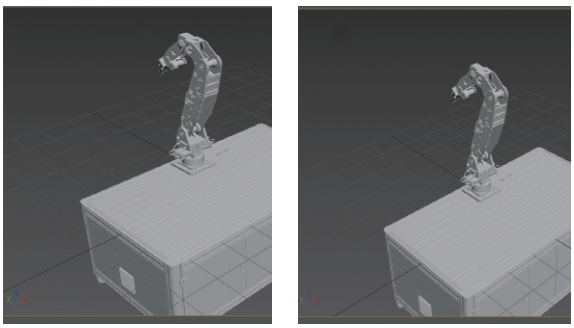


图 5 液压机器人 SolidWorks 模型

Fig. 5 SolidWorks model of hydraulic robot

2）为提升虚拟仿真计算效率并避免因模型过于复杂造成的计算负载过大和响应迟滞，将初始高精度三维模型导入 3ds Max 平台，利用其优化工具对冗余边角、微小曲面等非关键几何信息进行简化和压缩处理。轻量化处理后的模型在后续仿真过程中更具易用性，可有效兼顾模型精度与仿真效率

间的折中平衡, 进一步优化后, 模型的多边形数量减少至 1 765 026, 顶点数量减少至 885 987, 减少了约 13%, 上述优化操作有助于减少计算负担, 但又没有过度简化, 确保了模型在保证仿真效率的前提下, 依然能真实反映机械臂的结构特征和运动表现。优化结果如图 6 所示。



(a) 轻量化前的机器人模型　　(b) 轻量化后的机器人模型

图 6　3ds Max 模型轻量化处理

Fig. 6　Lightweight processing of 3ds Max model

3) 为实现结构运动验证与液压驱动过程间的协同仿真, 将处理后的轻量化三维模型导入 CoppeliaSim 仿真平台, 通过添加旋转关节、设定初始位姿及搭建层级结构, 完成机器人运动树的构建, 实现关节间的层级联动与精准控制, 如图 7 所示。

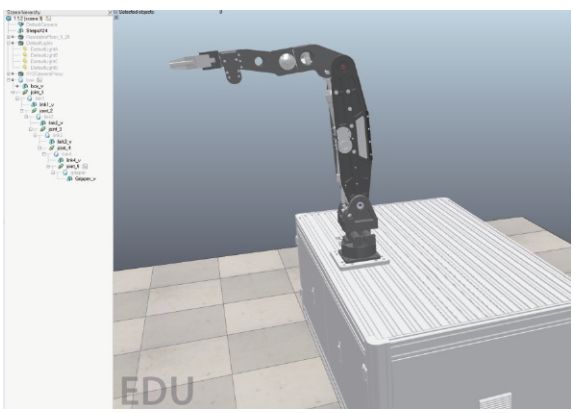


图 7　CoppeliaSim 机器人树状图

Fig. 7　CoppeliaSim robot tree diagram

2.1.2　液压机器人数字孪生属性建模

1) 机器人运动学建模

为实现液压机器人末端执行器在三维空间中的精确定位与姿态控制, 需对其进行运动学建模。本文采用改进 D-H 参数法, 按照标准坐标系建立规则, 对各连杆之间的几何与运动关系进行系统化描述。与传统 D-H 法相比, 该方法在坐标系分配及参数取值上进行了适配性调整, 可有效降低相邻关节参数冗余引起的建模误差, 尤其适用于液压驱动系统的高精度运动学建模需求[26]。

首先, 依据机器人结构特性与关节连接关系, 依次为各连杆建立关节坐标系, 如图 8 所示, 其中, a_{i-1}表示沿着x_{i-1}轴, 从z_{i-1}移动到z_i的距离; d_i表示沿着z_i轴, 从x_{i-1}移动到x_i的距离; θ_i表示绕着z_i轴, 从x_{i-1}旋转到x_i的角度。

在确定各连杆坐标系后, 通过分析相邻连杆间的空间几何关系, 得到改进 D-H 参数表, 如表 2 所示。

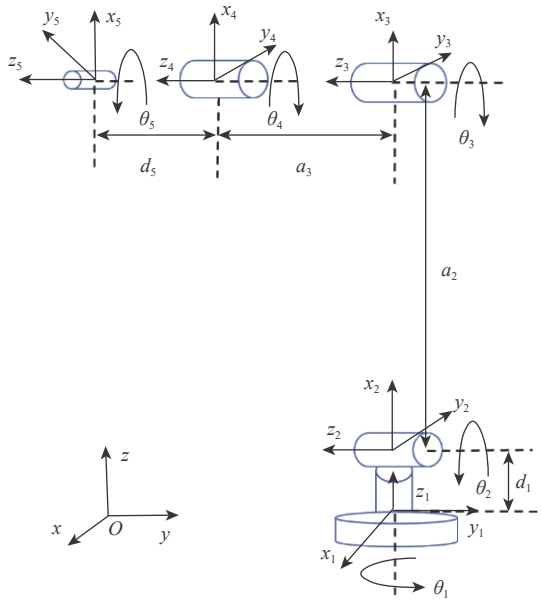


图 8　机器人的连杆坐标系

Fig. 8　Link coordinate system of robot

表 2　液压机器人改进 D-H 参数表

Table 2　Modified D-H parameters of hydraulic robot

连杆i	α_{i-1}/(°)	a_{i-1}/mm	d_i/mm	θ_i/(°)
1	0	0	286.923	θ_1
2	−90	0	0	θ_2
3	0	646.901	0	θ_3
4	0	244.866	0	θ_4
5	−90	0	461.049	θ_5

注: α_{i-1}表示绕着x_{i-1}轴, 从z_{i-1}旋转到z_i的角度。

基于表 2 的参数, 可推导出相邻连杆坐标系之间的齐次变换矩阵, 并通过链式法求得末端执行器相对于基座坐标系的位姿[27]。为多自由度液压机器人中串联的每个关节建立连杆坐标系。利用表 2 中的 4 个关键变量, 可推导出第i个连杆的坐标系相对于第$i-1$个连杆的坐标系的变换矩阵。其中, 第i个连杆和第$i-1$个连杆为 2 个相邻的连杆, 其变换矩阵表示为

$$\boldsymbol{A}_i^{i-1}=\boldsymbol{T}_x(a_{i-1})\times\boldsymbol{R}_x(\alpha_{i-1})\times\boldsymbol{R}_z(\theta_i)\times\boldsymbol{T}_z(d_i)=\begin{bmatrix}\cos\theta_i & -\sin\theta_i & 0 & a_{i-1}\\ \sin\theta_i\cos\alpha_{i-1} & \cos\theta_i\cos\alpha_{i-1} & -\sin\alpha_{i-1} & -d_i\sin\alpha_{i-1}\\ \sin\theta_i\sin\alpha_{i-1} & \cos\theta_i\sin\alpha_{i-1} & \cos\alpha_{i-1} & d_i\cos\alpha_{i-1}\\ 0 & 0 & 0 & 1\end{bmatrix}\tag{1}$$

式中：$\boldsymbol{R}_x(\alpha_{i-1})$为沿$x$轴旋转$\alpha_{i-1}$角度；$\boldsymbol{T}_x(a_{i-1})$为沿$x$轴移动$a_{i-1}$距离；$\boldsymbol{R}_z(\theta_i)$为沿$z$轴旋转$\theta_i$角度；$\boldsymbol{T}_z(d_i)$为沿$z$轴移动$d_i$距离。

将相邻坐标系间的变换依次相乘，即可得到机器人的正运动学方程为

$$\boldsymbol{T}_0^5=\boldsymbol{A}_0^1\times\boldsymbol{A}_1^2\times\boldsymbol{A}_2^3\times\boldsymbol{A}_3^4\times\boldsymbol{A}_4^5=\begin{bmatrix}n_x & o_x & m_x & p_x\\ n_y & o_y & m_y & p_y\\ n_z & o_z & m_z & p_z\\ 0 & 0 & 0 & 1\end{bmatrix} \tag{2}$$

式中：n_x、n_y、n_z为机器人末端坐标系x轴在基坐标系中的方向余弦分量；o_x、o_y、o_z为机器人末端坐标系y轴在基坐标系中的方向余弦分量；m_x、m_y、m_z为机器人末端坐标系z轴在基坐标系中的方向余弦分量；p_x、p_y、p_z为机器人末端在基坐标系中的位置。

根据表1中提供的改进D-H参数，可以求得其所在空间内末端相对于基座的具体位置和姿态，推导出坐标系5相对于坐标系0的变换矩阵为

$$\boldsymbol{T}_0^5=\begin{bmatrix}\cos(\theta_2+\theta_3+\theta_4)\cos\theta_5+\sin\theta_5 & -\cos(\theta_2+\theta_3+\theta_4)\sin\theta_5+\sin(\theta_2+\theta_3+\theta_4)\\ \sin(\theta_2+\theta_3+\theta_4)\cos\theta_5-\sin\theta_5\cos(\theta_2+\theta_3+\theta_4) & -\sin(\theta_2+\theta_3+\theta_4)\sin\theta_5-\cos\theta_5\cos(\theta_2+\theta_3+\theta_4)\\ \sin(\theta_2+\theta_3+\theta_4)\sin\theta_5 & -\cos\theta_5\cos(\theta_2+\theta_3+\theta_4)\\ 0 & 0\end{bmatrix}$$
$$\begin{bmatrix}-\sin\theta_5\cos(\theta_2+\theta_3+\theta_4) & \cos_1(a_2\cos\theta_2+a_3\cos\theta_3-d_3\sin(\theta_2+\theta_3+\theta_4))-d_1\sin\theta_1\\ -\sin(\theta_2+\theta_3+\theta_4) & \sin_1(a_2\mathrm{c}\sin\theta_2+a_3\mathrm{c}\sin\theta_3-d_3\sin(\theta_2+\theta_3+\theta_4))-d_1\cos\theta_1\\ \cos(\theta_2+\theta_3+\theta_4) & -a_2\sin\theta_2+a_3\sin\theta_2+\theta_3-d_2\cos\theta_1\\ 0 & 1\end{bmatrix}$$

液压机器人的逆运动学分析旨在已知末端执行器位姿矩阵的前提下，求解各个关节所对应的旋转角度。对于五自由度的液压机器人，考虑其关节间的几何关系和约束，代数法因其解析性强、求解精度高，成为逆运动学求解的常用方法之一。因此，本文采用代数法对液压机器人进行逆运动学求解，最终得到各关节角度的解析表达式为

$$\theta_1=\operatorname{atan2}(p_y,p_x)-\operatorname{atan2}\left(-d_1,\pm\sqrt{p_x^2+p_y^2-d_2^2}\right) \tag{3}$$

$$\theta_2=\operatorname{atan2}(\sin\theta_2,\cos\theta_2) \tag{4}$$

$$\theta_3=-\operatorname{acos}\left(\frac{\delta+\zeta}{2a_2a_3}\right) \tag{5}$$

$$\theta_4=\operatorname{atan2}(\sin(\theta_2+\theta_3+\theta_4),\cos(\theta_2+\theta_3+\theta_4))-\theta_2-\theta_3 \tag{6}$$

$$\theta_5=\operatorname{atan2}(n_x\sin\theta_1-n_y\cos\theta_1,o_x\sin\theta_1-o_y\cos\theta_1) \tag{7}$$

式中：$\zeta=(p_z-d_1-d_5m_z)^2-a_2^2-a_3^2$；$\delta=[p_x\cos\theta_1+p_y\sin\theta_1-a_1-d_5(m_x\cos\theta_1+m_y\sin\theta_1)]^2$。

2）液压系统的特性建模

液压系统作为机械臂的主要动力源，其建模精度直接影响控制系统的响应速度与运行稳定性。然而，由于液压系统结构复杂，包含多种非线性元件及耦合特性，若采用全量级、高保真方式进行整体动力学建模，将显著增加模型的计算复杂度，不利于仿真分析与后续控制策略的设计与实现。为兼顾建模精度与仿真效率，本文引入模块化建模方法，将液压系统划分为若干功能子模块分别进行建模，并通过标准化接口实现各子模块间的耦合与数据交互，从而构建完整的液压系统模型。

在建模过程中，主要针对常温、低速、层流工况下的运行状态，着重考虑内泄漏、液压容腔压缩性对系统动态性能的影响。为突出执行器动态响应特性，暂时忽略高阶效应如油液温升、黏度变化及湍流扰动等因素的影响。

本文研究的液压系统，其基本运行机制遵循流体力学基本原理。以液体为工作介质，通过液压元件串联构建典型控制回路，实现驱动动作的转换与输出。系统行为满足帕斯卡定律[28]且能量转化过程可由伯努利方程描述，作为后续分析液压回路中压力、流量及能耗变化的理论基础。伯努利方程的数学表达式为

$$C=P+\frac{1}{2}\rho v^2+\rho gh \tag{8}$$

式中：P为流体静压力；ρ为液压油密度；v为流体速度；g为重力加速度；h为液体高度。

为实现液压系统的模块化动力学建模，需对其中的关键元件进行精确建模，包括液压泵、液压缸、换向阀、溢流阀、过滤器。

①液压泵。

液压泵作为系统的动力源，其主要作用是将电动机的机械能转化为液压能，为系统提供流体动力。当液压驱动系统处于稳定工况时，液压泵的实际流量Q_{P}为理论流量Q_{t}与内泄漏流量Q_{l}的差，表达式为

$$Q_{\mathrm{P}}=Q_{\mathrm{t}}-Q_{\mathrm{l}}=D_{\mathrm{P}}\omega_{\mathrm{P}}-C_{\mathrm{inP}}(P_{\mathrm{P}}-P_0) \tag{9}$$

式中：D_{P}为液压泵的排量，$\mathrm{m}^3/\mathrm{rad}$；$\omega_{\mathrm{P}}$为泵的转速；

C_{inP}为泵的内泄露系数；P_P为泵出口压力；P_0为泵的入口压力。液压泵的工作性能直接影响整个系统的稳定性和动态响应。当液压泵发生内泄漏时，部分液压油未能有效输出至系统负载，而是在泵体内部泄回，从而导致泵的实际输出流量减小，系统压力难以维持在设定值。

②液压缸。

液压缸作为执行元件，其动力学行为直接决定机器人的关节运动效果。液压缸无杆腔的有效面积为

$$A_1 = \frac{\pi}{4}D^2 \tag{10}$$

式中：D为活塞直径。

液压缸有杆腔的有效面积为

$$A_2 = \frac{\pi}{4}(D^2 - d^2) \tag{11}$$

式中：d为液压杆直径。

在油液进入腔体后，液压缸腔中油液的连续性方程为

$$Q_1 = A\dot{x} + \frac{V}{\beta}\cdot\frac{\mathrm{d}P}{\mathrm{d}t} + C_{inlink}(P_1 - P_2) \tag{12}$$

式中：A为液压缸腔体的有效面积；$\dot{x}$为活塞移动速度；V为腔体容积；C_{inlink}为液压缸内泄露系数。液压缸的内泄漏会引起流量损失和压力波动。

③换向阀。

液压系统中采用三位四通换向阀作为主控制元件，负责调节液流方向，从而实现液压缸的伸出、回缩或保持。此类换向阀具有 4 个油口，3 种工作位置。此类换向阀是液压执行系统中最常见、应用最广泛的控制元件，其流量模型常采用孔口流量方程描述：

$$Q_v = Q_t - Q_z = C_d A_3\sqrt{\frac{2\Delta P}{\rho}} - C_{inp}\Delta P \tag{13}$$

式中：Q_v为阀口流量；C_d为流量系数；A_3为阀口开度面积；ΔP为阀门入口与出口的压力差。

④溢流阀。

溢流阀作为液压系统中的关键压力调节元件，其内部结构较为复杂。为维持系统压力的稳定，尤其是在湿式离合器液压系统中，由于工作频率较高、开关切换频繁，随着运行时间的延长，溢流阀容易出现主阀阻尼孔堵塞或阀门卡滞等退化性故障。

在正常状态下，流经溢流阀阀口处的流量计算式为

$$Q_B = C_d A_3\sqrt{\frac{2\Delta P}{\rho}} \tag{14}$$

当主阀阻尼孔堵塞时，阀口开度面积减小，导致流量无法正常响应系统压力变化。

⑤过滤器。

过滤器通过拦截液压油中的颗粒杂质、水分和其他污染物，防止其进入液压元件，从而降低系统磨损、延长设备寿命、提高控制精度。过滤器流量方程为

$$Q_C = kA_4\Delta P/\mu \tag{15}$$

式中：k为过滤器通油能力系数；A_4为过滤器有效流通面积；μ为液压油的黏度。一旦过滤器发生堵塞，液压油的有效流通面积将减小，导致系统流量不足，进而影响执行元件的响应速度与输出力，造成系统无法建立正常压力等现象。

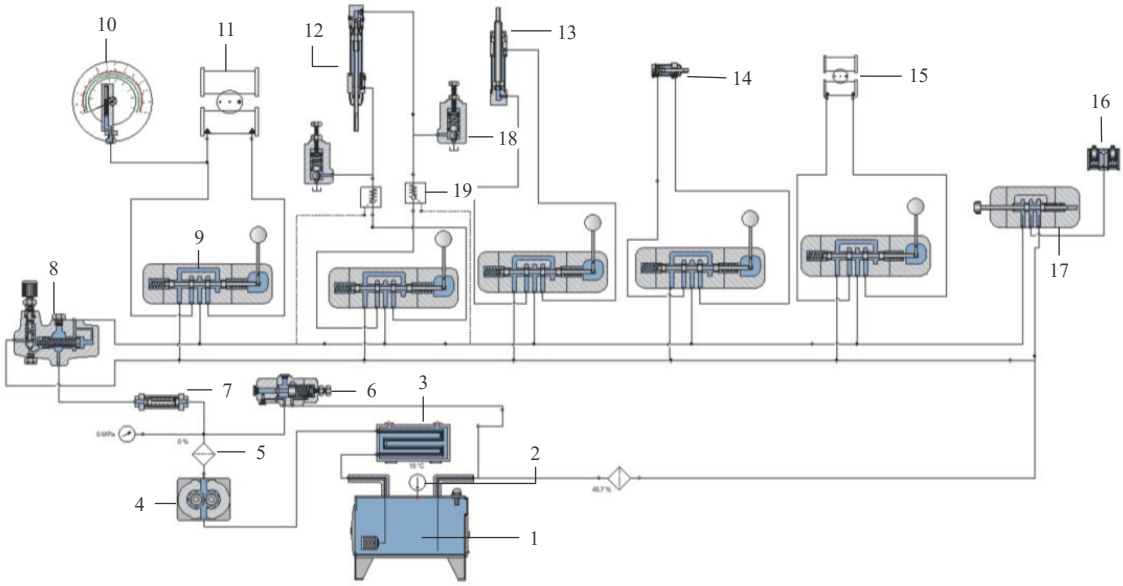
3）液压机器人仿真平台搭建

本文采用 CoppeliaSim 和 AS 构建了液压机器人数字孪生仿真平台，分别用于构建机器人的运动学数字孪生和建模与仿真液压驱动系统。CoppeliaSim 是一款功能丰富的机器人仿真软件，广泛应用于机器人建模、运动规划和虚拟调试。其提供了灵活的场景配置、基于脚本的控制及直观的三维可视化。通过 simRemoteApi 接口，CoppeliaSim 与 MATLAB 连接，MATLAB 从物理机器人读取关节测量数据并驱动虚拟机器人同步运动。同时，CoppeliaSim 实时收集关节角度和末端执行器位姿数据，作为后续故障诊断模型训练和验证的重要数据来源。

由 Famic Technologies开发的 Automation Studio™ 被采用为液压驱动系统的建模与仿真平台，用于分析液压回路在正常工作条件下和各种典型故障场景下的动态响应。AS 是一款高度集成的工业系统仿真软件，广泛应用于液压、气动、电气控制等自动化领域，具备强大的跨领域建模能力和直观的可视化功能。本文建立的压驱动系统故障仿真模型如图 9 所示。

在仿真模型的建立过程中，本文采用模块化设计思路，将液压机器人驱动系统划分为动力源模块、控制模块、执行模块、负载模块、传感模块和信号反馈模块等功能单元。各模块间通过标准液压接口实现互联，构建成闭环液压控制系统。

在此基础上，AS 平台可实时采集液压缸的位移、速度、负载反力等关键变量，有效再现液压系统的能量转换与动态执行过程，并与机器人运动学特性相结合，形成具备实时响应能力的液压机器人数字孪生系统。搭建完成的系统支持多种典型故障模式的注入，可用于分析液压机器人在异常工况下的动态行为，系统化评估其稳定性、响应特性及控制策略。



1.液压油箱; 2.油箱温度计; 3.风冷式冷却器; 4.油泵; 5.过滤器; 6, 8, 18.溢流阀; 7.流量计; 9, 17.换向阀; 10.压力表; 11, 15.摆动缸; 12, 13, 14, 16.直线缸; 19.止回阀。

图9 液压驱动系统的模型

Fig. 9 Model of hydraulic drive system

同时,采用AS平台与CoppeliaSim对液压机械臂进行建模与动态仿真,增强液压驱动过程的物理逼真度。通过上述两平台的协同支持,搭建集成结构建模、控制验证与液压响应模拟于一体的虚拟仿真环境,为后续数字孪生系统的开发与故障诊断分析提供平台环境。

2.1.3 数字孪生模型的校准

在数字孪生系统运行过程中,虚拟模型与实际机器人之间不可避免地存在误差,具体误差来源分析如下。

1)在启动与停止阶段,关节加速度变化较大,实际机器人与数字孪生体之间的动态响应特性存在差异。尤其是液压系统的非线性行为、关节摩擦力变化等因素,可能导致数字孪生体未能完全捕捉实际系统的瞬态特性,从而引发偏差。

2)传感器噪声也是误差的重要来源,尤其在运动的起始与结束阶段,信号波动较大,噪声影响更为显著,从而对角度测量精度造成干扰。

针对上述误差来源,本文未对数字孪生模型结构进行大幅改动,而是引入基于实测数据的误差补偿方法。在仿真运行过程中,实时采集实际机器人各关节的角度数据,并与虚拟模型的预测输出进行比对,得到瞬时误差序列。随后,采用滑动窗口平均法对误差进行平滑处理,并结合自适应矩估计(adaptive moment estimation, Adam)算法来动态计算补偿量,最终将其反馈至模型输出端,实现对仿真角度的动态修正。补偿算法的具体步骤如下:

1)误差计算。假设实际机器人在第t时刻的关节角度为$\boldsymbol{\theta}_{\text{real}}(t)$,数字孪生模型的角度为$\boldsymbol{\theta}_{\text{mod}}(t)$,则误差$\boldsymbol{e}(t)$为

$$\boldsymbol{e}(t)=\boldsymbol{\theta}_{\text{real}}(t)-\boldsymbol{\theta}_{\text{mod}}(t) \tag{16}$$

2)平滑处理。为减少误差中的高频噪声,采用滑动窗口平均法对误差进行平滑处理。设定窗口大小为M,则平滑误差$\hat{\boldsymbol{e}}(t)$为

$$\hat{\boldsymbol{e}}(t)=\frac{1}{M}\sum_{c=t-M+1}^{t}\boldsymbol{e}(t) \tag{17}$$

式中:$\boldsymbol{e}(t)$为时间步t时刻的误差;M为滑动窗口的大小,通常是一个固定的常数,决定了平滑处理的时间范围。

3)动态补偿。Adam算法用于动态调整补偿量$\boldsymbol{C}(t)$,以修正虚拟模型输出的角度。Adam算法的更新步骤如下:

步骤1 初始化。初始化一阶矩$\boldsymbol{m}(0)$和二阶矩$\boldsymbol{v}(0)$为$\boldsymbol{0}$,并设置学习率ϕ、一阶矩衰减率β_1、二阶矩衰减率β_2,稳定项ε及最大迭代次数N。

步骤2 梯度定义。以平滑误差的平方为优化目标函数$J(t)=\frac{1}{2}\hat{\boldsymbol{e}}^2(t)$,则关于补偿量$\boldsymbol{C}(t)$的梯度可定义为

$$\boldsymbol{g}(t)=\frac{\partial J(t)}{\partial \boldsymbol{C}(t)}=-\hat{\boldsymbol{e}}(t) \tag{18}$$

步骤3 一阶矩估计。在每个时间步t,计算一阶矩$\boldsymbol{m}(t)$为

$$\boldsymbol{m}(t)=\beta_1\boldsymbol{m}(t-1)+(1-\beta_1)\boldsymbol{g}(t) \tag{19}$$

步骤 4　二阶矩估计。二阶矩$\boldsymbol{v}(t)$为

$$\boldsymbol{v}(t)=\beta_2\boldsymbol{v}(t-1)+(1-\beta_2)\boldsymbol{g}(t)^2 \tag{20}$$

步骤 5　偏差修正。由于$\boldsymbol{m}(t)$和$\boldsymbol{v}(t)$初始为 **0**，因此，需要进行偏差修正：

$$\begin{cases}\hat{\boldsymbol{m}}(t)=\boldsymbol{m}(t)/(1-\beta_1^t)\\ \hat{\boldsymbol{v}}(t)=\boldsymbol{v}(t)/(1-\beta_2^t)\end{cases} \tag{21}$$

式中：β_1^t为t时刻的一阶矩衰减率；β_2^t为t时刻的二阶矩衰减率。

步骤 6　更新补偿量。根据修正后的矩估计，更新补偿量$\boldsymbol{C}(t)$为

$$\boldsymbol{C}(t)=\boldsymbol{C}(t-1)-\phi\frac{\hat{\boldsymbol{m}}(t)}{\sqrt{\hat{\boldsymbol{v}}(t)+\boldsymbol{\varepsilon}}} \tag{22}$$

步骤 7　基于补偿量的模型更新。根据计算得到的$\boldsymbol{C}(t)$，更新虚拟模型的输出角度为

$$\boldsymbol{\theta}_{\mathrm{mod}}(t)_{\mathrm{adj}}=\boldsymbol{\theta}_{\mathrm{mod}}(t)+\boldsymbol{C}(t) \tag{23}$$

步骤 8　迭代更新。该补偿过程在每次仿真迭代中持续进行，实时计算误差并调整补偿量，直到仿真模型与实际机器人之间的误差降至可接受范围内，或达到最大迭代次数N。具体校准流程如图 10 所示。

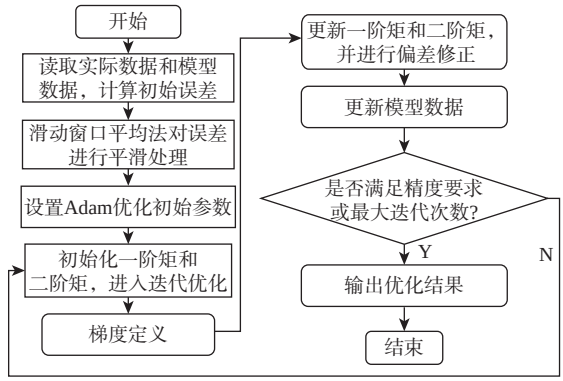


图 10　数字孪生模型校准流程

Fig. 10　Digital twin model calibration process

2.2　液压机器人的故障机理分析

为实现液压机器人的智能化故障诊断与预测性维护等功能，须深入理解其常见故障的物理机理，并构建可复现、多场景的故障仿真模型，进而将其注入至液压机器人数字孪生模型中，以获取具有代表性的系统动态响应数据。

故障机理是指引发液压系统中零部件或整体功能失效的物理、化学、电气或机械过程，其本质在于揭示系统故障形成的内在逻辑与本源机制。在进行故障机理分析时，需重点关注以下 3 个基本要素。

1）内因。其指故障元件自身的结构特性及运行状态，这些内在因素决定其对外部扰动的敏感性。液压机器人中的关键元件，包括液压泵、控制阀、液压缸等，其设计参数的合理性、选型匹配度、安装规范性及密封结构的完整性，均直接影响设备的抗干扰能力和故障发生概率。

2）外因。其指作用于故障载体的外部因素，此类因素可加速系统失效。常见诱因包括系统压力和流量异常、杂质堵塞、管路泄漏、人为操作失误及密封件老化等，上述外部因素通常是故障发生的直接触发点。

3）表现。其指在外部诱因作用于载体后所表现出的故障现象，如系统压力下降、执行机构动作失灵、响应时间延长、零部件过度磨损，甚至导致整个系统停机。上述具体表现构成了故障的外在特征。

结合液压机器人实际运行状况，其典型故障原因、故障类型和行为表征如图 11 所示。

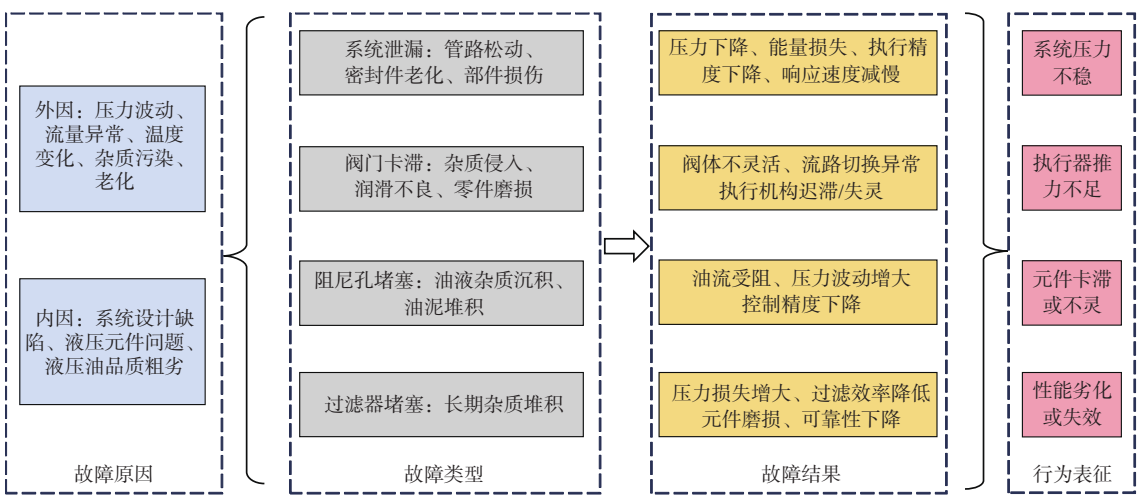


图 11　机器人故障形成分析

Fig. 11　Analysis of robot fault formation

3 基于 Transformer-LSTM-XGBoost 融合的液压机器人故障诊断方法

为实现液压机器人在复杂工况下的智能化故障诊断，本文提出了一种基于 Transformer-LSTM-XGBoost 融合的液压机器人智能故障诊断方法。该方法将深度特征提取与集成学习分类相结合，以实现从多源传感信号到故障类型识别的诊断模型建立。首先，利用 Transformer 模块捕获传感数据间的全局依赖关系，借助 LSTM 网络建模系统的时序动态特征；最终，通过 XGBoost 分类器实现对液压机器人驱动系统典型故障的诊断。

3.1 Transformer-LSTM 特征提取模块

为有效挖掘液压机器人多源信号中的空间关联性与时序依赖性，本文构建了 Transformer-LSTM 串联式特征提取模块，用于深层特征融合与时间序列建模。

1）Transformer 模块

Transformer 模块负责捕获液压机器人运行期间多源信号之间的全局依赖关系，利用注意力机制实现多源时间序列特征的全局建模与深层表示[29]。与 RNN 不同，Transformer 摒弃了顺序递归结构，采用并行化的自注意力机制，能够同时关注输入序列中不同位置点特征间的相关性，从而显著提升特征提取的效率与全局表达能力[29]。

Transformer 模块的核心结构为编码器，由多层堆叠的多头注意力机制与前馈神经网络组成。

在自注意力机制中，输入特征首先被映射为查询 $\boldsymbol{Q}$、键 $\boldsymbol{K}$ 和值 $\boldsymbol{V}$ 这 3 个矩阵，通过计算矩阵 $\boldsymbol{Q}$ 与 $\boldsymbol{K}$ 之间的点积，得到其之间的相似度。为避免点积值过大导致数值不稳定，通常会对点积结果进行缩放处理，具体做法是将点积除以查询向量维度的平方根 $\sqrt{d_{\boldsymbol{K}}}$。接着，对缩放后的点积应用 SoftMax 操作将相似度转换为概率分布，得到每个查询与各个键之间的权重。最终，这些权重用于对值矩阵 $\boldsymbol{V}$ 进行加权求和，从而生成自注意力机制的输出特征，其数学表达式为

$$\mathrm{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\mathrm{SoftMax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_{\boldsymbol{K}}}}\right)\boldsymbol{V} \tag{24}$$

式中：$d_{\boldsymbol{K}}$ 为键向量的维度，用于防止点积结果过大；SoftMax 为获得值的权重的函数。

在此基础上，多头注意力机制通过并行计算多个注意力子空间并进行拼接融合，进一步增强了模型的特征捕获能力，其计算式为

$$\mathrm{MH}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\mathrm{Concat}(h_1,h_2,\cdots,h_j)\boldsymbol{W}_o \tag{25}$$

$$h_j=\mathrm{Attention}(\boldsymbol{Q}\boldsymbol{W}_j^{\boldsymbol{Q}},\boldsymbol{K}\boldsymbol{W}_j^{\boldsymbol{K}},\boldsymbol{V}\boldsymbol{W}_j^{\boldsymbol{V}}), \tag{26}$$

式中：j 为注意力头数；$\boldsymbol{W}_j^{\boldsymbol{Q}}$、$\boldsymbol{W}_j^{\boldsymbol{K}}$、$\boldsymbol{W}_j^{\boldsymbol{V}}$、$\boldsymbol{W}_{\mathrm{o}}$ 为可训练参数矩阵。

2）LSTM 模块

本文采用 LSTM 模块进一步对 Transformer 提取的高维特征进行时间递推建模。基于输入门、遗忘门和输出门的门控机制，LSTM 能够选择性地保留关键特征，抑制冗余信息，从而有效刻画液压系统的动态演化过程及长期依赖特性[30-31]。

在 LSTM 单元结构中，主要包括输入门、遗忘门和输出门 3 个部分，分别用于控制信息的输入、更新与输出。其数学表达式为

$$\boldsymbol{f}_t=\sigma\left(\boldsymbol{W}_f[\boldsymbol{h}_{t-1},\boldsymbol{x}_t]+\boldsymbol{b}_f\right) \tag{27}$$

$$\boldsymbol{i}_t=\sigma(\boldsymbol{W}_i[\boldsymbol{h}_{t-1},\boldsymbol{x}_t]+\boldsymbol{b}_i) \tag{28}$$

$$\boldsymbol{o}_t=\sigma(\boldsymbol{W}_o[\boldsymbol{h}_{t-1},\boldsymbol{x}_t]+\boldsymbol{b}_o) \tag{29}$$

$$\tilde{\boldsymbol{C}}_t=\tanh(\boldsymbol{W}_C[\boldsymbol{h}_{t-1},\boldsymbol{x}_t]+\boldsymbol{b}_C) \tag{30}$$

$$\boldsymbol{C}_t=\boldsymbol{f}_t\boldsymbol{C}_{t-1}+\boldsymbol{i}_t\tilde{\boldsymbol{C}}_t \tag{31}$$

$$\boldsymbol{h}_t=\boldsymbol{o}_t\tanh(\boldsymbol{C}_t) \tag{32}$$

式中：$\boldsymbol{x}_t$ 为当前时刻输入；$\boldsymbol{h}_{t-1}$ 为前一时刻隐藏状态；$\boldsymbol{C}_t$ 为当前细胞状态；$\boldsymbol{f}_t$、$\boldsymbol{i}_t$、$\boldsymbol{o}_t$ 分别为遗忘门、输入门和输出门的门控向量；$\boldsymbol{W}$ 和 $\boldsymbol{b}$ 为权重和偏置。

3）Transformer-LSTM 耦合模型构建

在液压机器人故障诊断中，传感器采集的数据通常表现出复杂的时序依赖性和空间相关性，单一模型难以全面捕捉这些特征。为充分挖掘多源传感器信号中的有效信息，本文设计了一种基于 Transformer 与 LSTM 相结合的特征提取模块。

与传统的机器学习方法（如 CNN、LSTM）相比，CNN 在空间特征提取方面表现优异，但对时间序列依赖的建模能力有限，难以有效处理动态变化的故障信号。Transformer-LSTM 耦合结构可以同时捕获全局信息与时序依赖，显著提升对复杂信号的理解和表示能力。其中，Transformer 模块通过自注意力机制和并行计算，有效建模长时间序列的全局依赖关系，并能够自动关注故障诊断的关键特征，缓解了传统 RNN 和 LSTM 在处理长序列时容易出现的梯度消失和信息丢失问题。LSTM 模块则专注于局部时间序列特征的建模，能够捕捉故障信号的瞬时变化和动态特征，从而补充 Transformer 在局部依赖建模上的不足。通过两者的耦合，提升了长序列处理效率，还增强了对复杂、多源传感器信号的识别能力，可为液压机器人故障诊断提供更加可靠和精准的数据驱动解决方案[32]。

3.2　XGBoost 故障诊断模块

为实现液压机器人在复杂工况下的多类别故障精准诊断，本文在特征提取阶段获得高维特征基础上，构建了 XGBoost 故障诊断模块。该模块以梯度提升决策树（gradient boosting decision tree,GBDT）为核心，通过集成多棵弱分类树构建强分类器，实现对多源特征的非线性映射与故障类别诊断[33]。

XGBoost 算法本质上是一种基于梯度提升思想的集成学习算法。模型以加法形式逐步叠加弱分类器，每一轮迭代均在前一轮模型预测残差的基础上进行优化，从而不断提高整体预测性能。XGBoost 的目标函数可表示为

$$O_{\mathrm{bj}}=\sum_{i=1}^{m} l(y_i,\hat{y}_i^{(t)})+\sum_{K=1}^{t}\Omega(f_K) \tag{33}$$

式中：$l(y_i,\hat{y}_i^{(t)})$为损失函数；$\Omega(f_K)$为正则化项，用于约束第K棵树的复杂度。通过对目标函数进行二阶泰勒展开，可得第t轮迭代的近似优化目标：

$$O_{\mathrm{bj}}^{(t)}\approx\sum_{i=1}^{m}\left(g_i f_t(x_i)+\frac{1}{2}h_i f_t^2(x_i)\right)+\Omega(f_t) \tag{34}$$

式中：g_i和h_i分别为一阶与二阶梯度。模型通过最小化式 (34) 所示目标函数确定最优分裂节点与叶节点权重，实现每轮残差的最优拟合。

3.3　基于 Transformer-LSTM-XGBoost 的故障诊断方法

针对液压机器人运行环境复杂、信号噪声干扰强及多源数据特征耦合度高等问题，本文提出了一种基于 Transformer-LSTM-XGBoost 融合架构的故障诊断方法。

首先，通过 Transformer-LSTM 模块对液压机器人多源动态信号进行深度特征提取，Transformer 编码层利用自注意力机制捕捉信号间的全局相关性；其次，借助 LSTM 层进一步挖掘时间序列的长期依赖关系，从而获得兼具全局与时序特征的高维特征表示；再次，将提取的深层特征作为 XGBoost 模型的输入，利用其基于 GBDT 的增量学习机制实现特征的非线性映射与最优决策边界构建；最后，XGBoost 模块通过迭代优化损失函数，逐步增强弱分类器的判别能力，从而提升模型在复杂工况下的故障诊断精度与鲁棒性。该故障诊断方法的整体框架如图 12 所示。

具体诊断方法的步骤描述如下。

1）原始数据采集与预处理。从液压机器人数字孪生系统中采集多源时序信号数据（包括正常与故障工况），并对其进行噪声滤波、归一化及时间对齐处理，以保证输入特征的统一性与可比性。

2）Transformer 模块的全局特征建模。先将处理后的时序信号输入 Transformer 编码层，进行多头自注意力机制处理。信号被转换为$\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{V}$矩阵，计算查询与键的相似度，得到注意力权重。通过 SoftMax 对权重归一化，得到各时间步间的关联强度，并与值矩阵相乘，得到加权特征。多个注意力头并行计算，最后拼接输出并通过线性变换，生成全局特征表示，用于后续的时序建模和故障诊断。

3）LSTM 模块的时序建模。首先，将 Transformer 编码层输出的全局特征序列输入 LSTM层；然后，

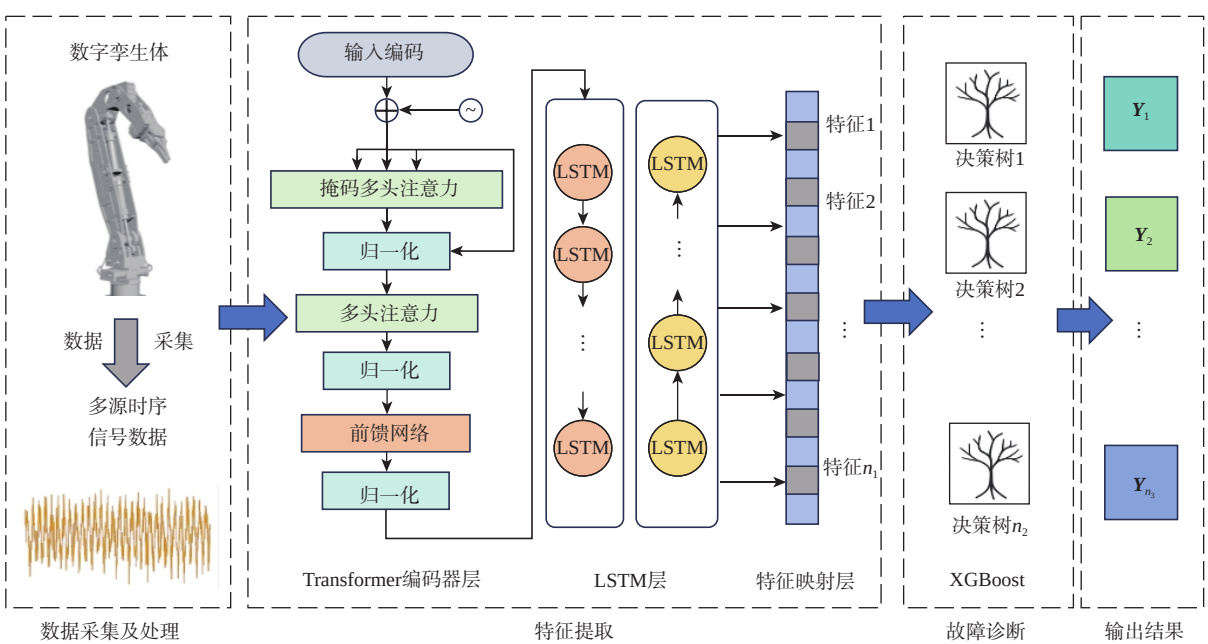


图 12　Transformer-LSTM-XGBoost 的液压机器人故障诊断方法框架

Fig. 12　Transformer-LSTM-XGBoost-based framework for hydraulic robot fault diagnosis method

LSTM 逐步处理每个时间步的输入，利用其内部的遗忘门、输入门和输出门来控制信息的传递和记忆，通过这些门控机制，LSTM 能够根据当前输入和历史信息动态调整状态，捕捉信号的短期变化和长期依赖性，每个时间步的隐藏状态和细胞状态会更新；最后，生成时序特征表示，这些表示包含了液压机器人多源信号的时间依赖关系，并为后续处理提供时序信息。

4）深层特征向量化处理。Transformer-LSTM 模块输出的高维时序特征通常以张量形式存在，不利于传统机器学习分类器的直接输入。为此，本文对该输出进行特征向量化处理，将多维时序特征映射为定长的二维特征矩阵，以实现从深度网络特征空间到传统分类模型输入格式的结构转换，为后续 XGBoost 模块的训练提供了高质量的特征输入。

5）基于 XGBoost 的故障诊断决策。XGBoost 通过构建一系列的决策树来进行分类，每棵树通过学习样本的残差来逐步改进模型。每次迭代，XGBoost 会计算当前预测值与真实值之间的误差，并基于该误差构建新的决策树。决策树的分裂通过计算每个特征的增益来选择最优的划分点，同时通过正则化项控制树的复杂度，避免过拟合。最后，所有树的输出会被加权汇总，得到最终的故障诊断结果和相应的置信度指标。

6）输出故障诊断结果。将分类结果输出至诊断模块，得到液压机器人在不同运行工况下的故障类型与置信度指标。通过与真实标签数据的比对，可计算出准确率、召回率及 F_1 值等性能指标，用以评估模型在不同数据集和工况下的诊断精度与鲁棒性。

4 仿真实验与结果分析

4.1 液压机器人数据的采集

为建立高精度的液压机器人数字孪生系统并支撑故障诊断研究，需采集多维状态变量数据，系统在故障工况下的典型特征。数据来源包括两部分：来自液压机器人物理实体的实测数据及基于液压机器人孪生体生成的故障数据。

1）实测数据采集。依托液压机器人物理实体，通过部署的位移、压力和温度传感器，采集实际运行过程中的关节位姿、角度变化、系统内部压力波动及油液温度变化等信息，用于液压机器人数字孪生模型的运行驱动与校准修正。

2）故障数据采集。基于液压机器人的数字孪生体，在正常工况及 4 类典型故障工况下（系统泄漏、阀门卡滞、阻尼孔堵塞、滤油器堵塞）进行多轮仿真实验，生成故障数据集。

利用孪生系统内置的实时监测与数据导出接口，采集液压缸的位移、速度及液压系统中油液的压力分布与流量变化，所有仿真均采用固定时间步长记录，确保数据时序一致性与可比性。通过在不同工况下获取的数据，可为故障诊断模型提供具有代表性的特征样本，并可作为孪生模型性能验证与优化的重要依据。图 13 为部分不同故障类型下的液压机器人压力信号的变化趋势。

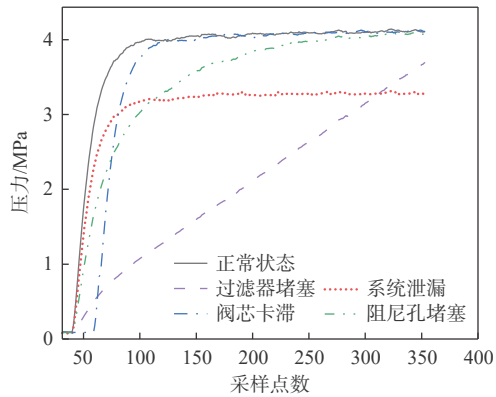


图 13 不同故障类型下压力信号的变化趋势

Fig. 13 Pressure signal variations under different fault types

由图 13 可得，不同故障工况下信号表现出显著的特征差异。在正常工况下，系统压力迅速上升并稳定于约 4.0 MPa；当发生系统泄漏故障时，压力上升速度较慢，稳定压力低于正常状态，约维持在 3.3 MPa 左右；当发生阀门卡滞故障时，初期压力变化不明显，随后出现快速上升，最终稳定压力与正常状态接近；当发生阻尼孔堵塞故障时，压力逐渐上升并稳定于约 3.8 MPa，略低于正常运行水平；当发生滤油器堵塞故障时，受滤芯阻塞影响，油液流动受限，压力上升明显放缓。

4.2 液压机器人数字孪生模型的验证

为验证液压机器人数字孪生模型的精确性，在正常工况下将机器人孪生体输出的预测数据与液压机器人的物理实体数据进行对比分析，图 14 和图 15 分别为液压机器人 5 个关节的实测角度与数字孪生预测角度的对比曲线，以及各关节角度误差随时间的变化趋势。

表 3 为数字孪生模型的精度评价结果，包括平均绝对误差（mean absolute error, MAE）、均方根误差（root mean square error, RMSE）、偏差（bias ratio, BIAS）及决定系数（coefficient of determination, R^2）。

由图 14、图 15 及表 3 可知，本文构建的数字孪生模型整体精度较高。然而，在机器人各关节依次运动至目标角度及从目标角度返回零位的过程中，误差呈阶段性增大；而在关节稳定运动阶段（即非启动与停止阶段），误差显著减小。这一现象表明

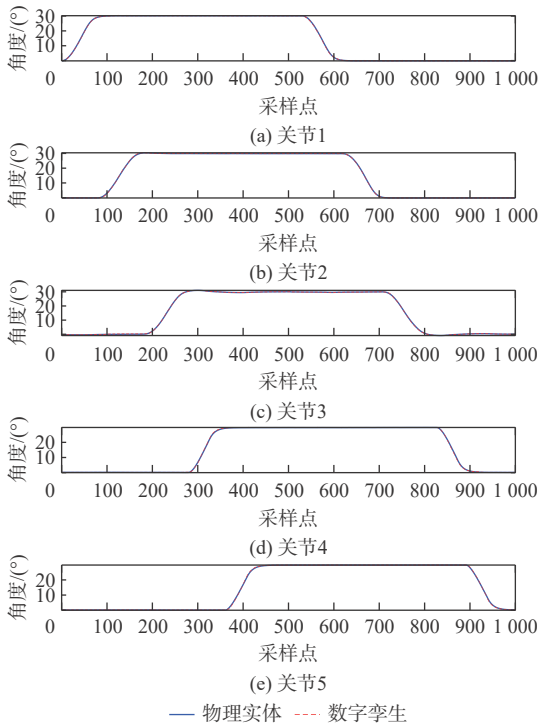


图 14　关节角度对比

Fig. 14　Comparison of joint angles

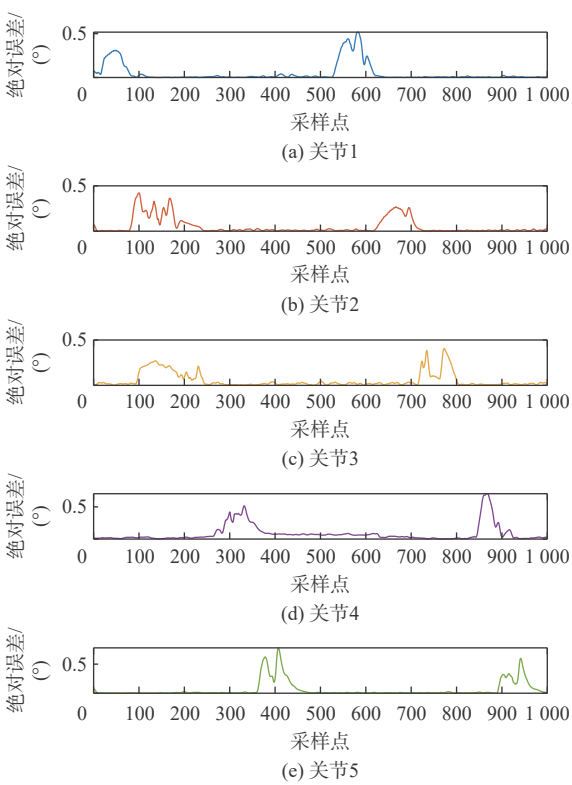


图 15　关节角度误差对比

Fig. 15　Comparison of joint angle errors

表 3　数字孪生模型精度评价结果

Table 3　Accuracy evaluation results of digital twin model

MAE/(°)	RMSE/(°)	BIAS/(°)	R^2/%
0.26	0.32	−0.16	98.3

模型在动态工况下的响应精度存在一定不足，而在稳态条件下能够达到较高的拟合精度。基于 2.1.3 节中的方法，在不增加模型计算复杂度的前提下，抑制启动与停止阶段的瞬时偏差，具体的补偿结果如图 16 所示。

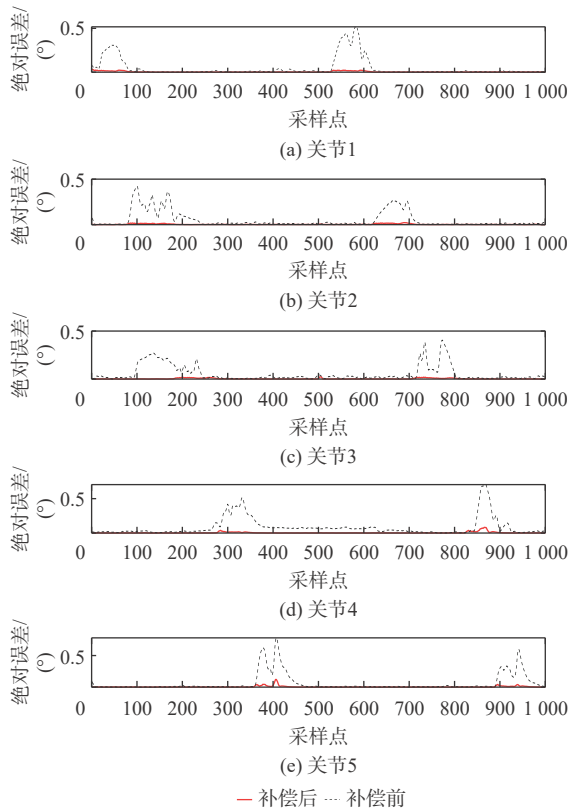


图 16　补偿后的关节误差对比

Fig. 16　Comparison of joint errors after compensation

表 4 为补偿后的数字孪生模型的精度评价结果，包括 MAE、RMSE、BIAS、R^2。

表 4　补偿后的数字孪生模型精度评价结果

Table 4　Accuracy evaluation results of compensated digital twin model

MAE/(°)	RMSE/(°)	BIAS/(°)	R^2/%
0.059	0.127	0.07	99.9

由图 16 可得，经过补偿后，关节角度的误差显著降低。在机器人启动与停止阶段，经补偿后的误差曲线波动显著减弱。在机器人稳定运行阶段，优化后的误差进一步趋近于 0，表明数字孪生体的稳态精度得到明显提升。综合来看，补偿后的数字孪生模型在关节角度的预测精度方面表现优异。由表 4 可见，优化后模型的 MAE、RMSE 和 BIAS 均有明显下降，R^2 接近 100%，表明补偿后的关节角度预测结果与实测值高度一致。

4.3　样本数据集的处理

本文采集的液压机器人多源传感信号数据经

过去噪、缺失值填补和 *Z*-score 标准化处理后，构建了符合 Transformer-LSTM-XGBoost 模型输入要求的高质量样本数据集。

为验证本文故障诊断方法的有效性与泛化能力，本文将处理后的数据集按 80%、10%、10% 的比例随机划分为训练集、验证集和测试集。训练集用于模型参数学习与特征提取；验证集用于模型超参数调优与防止过拟合；测试集用于评估模型在未知数据上的故障识别性能。各类别样本标签及样本数量的分布情况如表 5 所示。

表 5　样本分布

Table 5　Sample distribution

故障类别	类别标签	样本数/条
滤油器堵塞	1	392
阻尼孔堵塞	2	392
阀门卡滞	3	392
系统泄漏	4	392
正常工况	5	779

数据集说明：采样频率为 50 Hz，确保能够捕捉到细微的动态变化用于故障诊断；每类样本长度为 2 000，确保能够捕捉液压机器人动态行为的变化；窗口长度为 200，通过实验验证，该窗口长度在计算效率与特征提取之间可实现良好的折中平衡。

4.4　故障诊断方法验证与结果分析

1）基本参数设置

本文基于 MATLAB R2024a 平台，在 Intel Core i5-1135G7 处理器与 RTX 4 060 Laptop GPU 环境下完成模型训练。

Transformer 模块。采用单层 Transformer 编码结构，注意力头数设为 2，嵌入维度为 64。该模块通过多头自注意力机制实现对多源信号全局特征的捕捉与加权建模，以增强特征表达能力。

LSTM 模块。设置双层长短期记忆网络，每层隐藏单元数为 16，最大迭代轮数为 50，批量大小为 32。该模块用于建模液压机器人多源信号的时序依赖特征。

XGBoost 模块。以 CART 回归树为基学习器，学习率设为 0.1，最大树深度为 6，子采样率为 0.8，迭代次数为 200，正则化参数 $\lambda = 1$,。该模块通过二阶梯度优化损失函数并引入正则化项抑制过拟合，实现多类别故障的非线性判别与分类决策。

2）有效性验证与结果分析

为验证基于液压机器人数字孪生模型的 Transformer-LSTM-XGBoost 故障诊断方法的有效性与实用性，本节设计了多组对比实验，涵盖了多个经典与现代机器学习模型，包括：LSTM[31]、Transformer[32]、XGBoost[33]、Transformer-LSTM、LSTM-XGBoost、Transformer-XGBoost 这 6 种消融实验模型，以及 SVM[34]、CNN[30]、BiLSTM[31]、门控循环单元（gated recurrent unit,GRU）[35] 和 Transformer-LSTM-RF[36] 这 5 种对照模型。所有模型均使用相同的液压机器人多源时序数据集，并且在一致的训练与验证条件下进行实验。本节设计确保了对比结果的可靠性与可比性，避免了外部变量的干扰，进一步保障了结论的适用性。具体的验证情况如图 17、表 6 和图 18 所示。

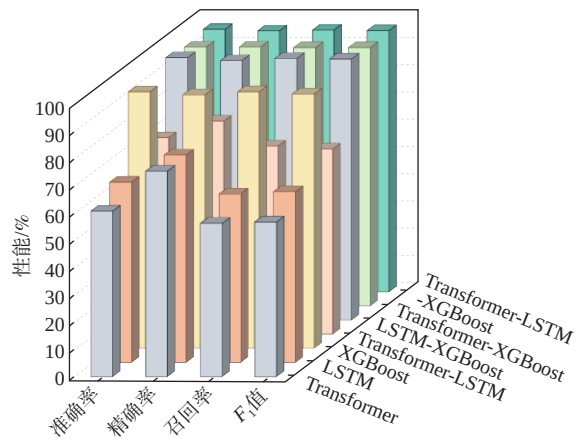


图 17　性能指标对比

Fig. 17　Performance metrics comparison

表 6　多模型性能指标对比

Table 6　Multi-Model performance comparison chart

模型	准确率/%	精确率/%	召回率/%	F_1值/%
SVM[34]	52.8	60.8	49.6	45.3
CNN[30]	81.7	83.6	78.6	80.0
BiLSTM[31]	68.0	77.7	63.4	66.2
GRU[35]	62.1	70.6	56.7	58.1
Transformer[32]	60.9	75.6	56.4	56.9
LSTM[31]	66.4	76.3	62.1	62.9
XGBoost[33]	94.4	93.3	94.4	93.6
Transformer-LSTM	72.3	78.4	69.2	68.1
LSTM-XGBoost	96.6	95.5	96.3	96.1
Transformer-XGBoost	95.3	95.3	95.1	95.1
Transformer-LSTM-RF[36]	96.2	94.8	96.2	95.1
Transformer-LSTM-XGBoost	96.6	96.0	96.3	96.1

由图 17、表 6 和图 18 可知，不同模型在液压系统故障的识别精度上存在明显差异。具体而言，Transformer-LSTM-XGBoost 融合模型在多类别故障诊断中表现出最高的准确率和最优的诊断效果。传统的单一模型（如 SVM、CNN、GRU 等）在准确率和其他性能指标上明显低于融合模型，尤其在准确率、精确率和召回率等关键指标上表现的更为明显。

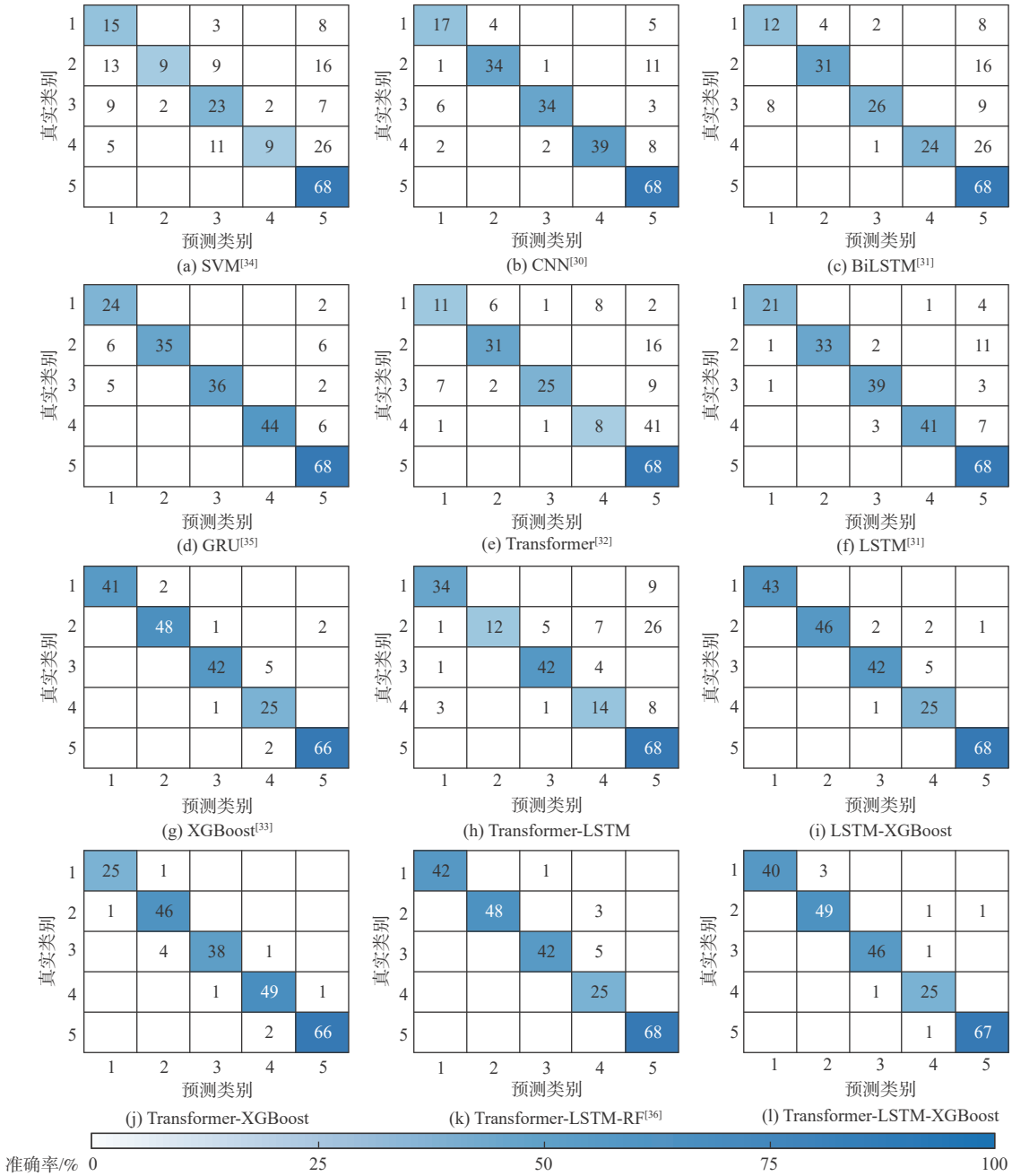


图 18　不同模型的混淆矩阵

Fig. 18　Confusion matrices of different models

在消融实验中，单一模型的表现欠佳，主要原因在于单一模型难以同时捕捉到空间结构特征和时序变化规律，从而导致其在多类别故障诊断中的精确度和鲁棒性较低。而两两复合模型，如 Transformer-XGBoost、LSTM-XGBoost，尽管表现较好，但相比 Transformer-LSTM-XGBoost 融合模型，仍有一定差距。进一步证明了多模型融合在液压系统故障诊断中的有效性及准确性。

为评估本文故障诊断方法在噪声环境中的鲁棒性，在测试集上分别加入了不同强度的高斯噪声，以模拟液压系统中的实际噪声干扰。训练集和验证集则保持原始数据，不进行噪声干扰，以确保模型的学习过程不受外部因素的干扰，从而能够从干净的数据中有效学习相关特征。

在本实验中，噪声强度是通过高斯噪声的标准差来控制的。噪声的强度分为：低强度噪声（标准差为 0.01，模拟轻微的电气或传感器噪声）、中等强度噪声（标准差为 0.05，模拟液压系统中可能出现的中等程度的噪声）和高强度噪声（标准差为 0.1，模拟液压系统在较为恶劣的工作环境中所遇到的强噪声）。

通过向验证集数据中逐步添加不同强度的高斯噪声，模拟噪声对故障诊断性能的影响，实验结果验证了 Transformer-LSTM-XGBoost 融合模型在

噪声干扰下的鲁棒性与稳定性，如表7所示。

从表7可以看出，随着噪声强度的增加，模型的准确率、精确率、召回率和F_1值均呈下降趋势。在无额外噪声条件下，模型准确率为96.6%；当噪声强度增加至中强度噪声时，模型准确率下降至93.1%，但仍保持较好的诊断性能；当噪声进一步增强至高强度噪声时，模型准确率降至89.8%，精确率、召回率和F_1值也相应降低。总体来看，虽然噪声增强会导致模型性能下降，但Transformer-LSTM-XGBoost融合模型在不同噪声强度下仍能维持较高的诊断精度，尤其在中低强度噪声条件下表现出较好的鲁棒性。

表7 不同噪声强度下的模型性能对比

Table 7 Performance comparison under different noise intensities

噪声强度	准确率/%	精确率/%	召回率/%	F_1值/%
无额外噪声	96.6	96.0	96.3	96.1
低强度噪声	95.5	95.1	95.3	95.2
中强度噪声	93.1	93.7	94.0	93.8
高强度噪声	89.8	91.0	89.5	88.8

5 结 论

针对基于传统方法的液压机器人典型故障数据获取难，故障诊断精度低等问题，本文提出了一种基于数字孪生技术融合Transformer-LSTM-XGBoost的液压机器人故障智能诊断方法，并通过实验验证了该方法的有效性，得到如下结论。

1）建立了支持系统状态精准感知与虚实同步的液压机器人数字孪生模型，实现了液压驱动系统在多工况下的高保真建模与动态行为模拟。通过引入基于实测数据的补偿机制，保证了孪生模型在动态与稳态阶段均能准确反映实体机器人的运行特性，为后续故障诊断提供了可靠的虚拟实验环境。

2）依托所建立的液压机器人数字孪生模型，对液压系统泄漏、阀门卡滞、阻尼孔堵塞和过滤器堵塞等典型液压故障进行了精准模拟，构建了具有代表性的高质量故障样本数据集，并经归一化等预处理操作，显著提升了数据集的适应性与模型的训练效果。

3）本文提出的Transformer-LSTM特征提取模块能够有效挖掘多源时序信号中的空间特征、时序依赖关系及关键特征区域；结合XGBoost分类器对深度融合特征进行高效决策，从而实现了多类别故障的准确诊断。实验结果表明，Transformer-LSTM-XGBoost融合模型的准确率达到了96.6%，性能上明显优于基线模型。此外，在鲁棒性验证中，该融合模型依然能够在中强度噪声条件下维持93.1%的准确率。因此，本文故障诊断方法不仅在正常工作环境下具有优异的诊断能力，而且在实际应用中也能应对噪声干扰，具有较好的鲁棒性。

下一步的研究方向将结合液压机器人关键部件的退化机理与历史运行数据，在数字孪生的技术框架下开展全生命周期状态监测与寿命预测研究，为液压设备的预测性维护和健康管理提供技术支持。

参考文献（References）

[1] DI LALLO A, YU S Y, SLIGHTAM J E, et al. Untethered fluidic engine for high-force soft wearable robots[J]. Advanced Intelligent Systems, 2024, 6(11): 2400171.

[2] 朱明君, 张鹏飞, 黄德民, 等. 液压系统的故障诊断与健康管理研究综述[J]. 润滑油, 2023, 38(3): 7-12.
ZHU M J, ZHANG P F, HUANG D M, et al. A research review of fault diagnosis and health management of hydraulic systems[J]. Lubricating Oil, 2023, 38(3): 7-12(in Chinese).

[3] 陈斌, 王占林, 裘丽华. 机载液压系统的主要发展趋势[J]. 航空学报, 1998, 19(1): 2-7.
CHEN B, WANG Z L, QIU L H. Main developmental trend of aircraft hydraulic systems[J]. Acta Aeronautica et Astronautica Sinica, 1998, 19(1): 2-7(in Chinese).

[4] REN X, GUO Q, LI T S, et al. Fault detection and tolerant control with a variable-bandwidth extended state observer for electro-hydraulic servo systems with lumped disturbance and measurement noise[J]. ISA Transactions, 2025, 162: 287-300.

[5] DAI J Y, TANG J, HUANG S Z, et al. Signal-based intelligent hydraulic fault diagnosis methods: review and prospects[J]. Chinese Journal of Mechanical Engineering, 2019, 32: 75.

[6] PENG Z K, CHU F L. Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography[J]. Mechanical Systems and Signal Processing, 2004, 18(2): 199-221.

[7] 周汝胜, 焦宗夏, 王少萍. 液压系统故障诊断技术的研究现状与发展趋势[J]. 机械工程学报, 2006, 42(9): 6-14.
ZHOU R S, JIAO Z X, WANG S P. Current research and developing trends on fault diagnosis of hydraulic systems[J]. Chinese Journal of Mechanical Engineering, 2006, 42(9): 6-14(in Chinese).

[8] 夏婷, 曹戈, 朱加升, 等. 工程机械故障诊断技术研究综述及展望[J]. 工程机械与维修, 2024(4): 6-9.
XIA T, CAO G, ZHU J S, et al. Review and prospect of fault diagnosis technology for construction machinery[J]. Construction Machinery & Maintenance, 2024(4): 6-9(in Chinese).

[9] PENG P, WANG J G. NOSCNN: a robust method for fault diagnosis of RV reducer[J]. Measurement, 2019, 138: 652-658.

[10] 姜少飞, 邬天骥, 彭翔, 等. 基于XGBoost特征提取的数据驱动故障诊断方法[J]. 中国机械工程, 2020, 31(10): 1232-1239.
JIANG S F, WU T J, PENG X, et al. Data driven fault diagnosis

method based on XGBoost feature extraction[J]. China Mechanical Engineering, 2020, 31(10): 1232-1239(in Chinese).

[11] 郑彪, 高丙朋, 程静. 基于 DNN-CapsNet 的液压泵故障程度诊断方法[J]. 液压与气动, 2023, 47(8): 41-49.
ZHENG B, GAO B P, CHENG J. Fault degree diagnosis of hydraulic pump based on DNN-CapsNet[J]. Chinese Hydraulics & Pneumatics, 2023, 47(8): 41-49(in Chinese).

[12] SUN J C, DING H, LI N, et al. Intelligent fault diagnosis of hydraulic system based on multiscale one-dimensional convolutional neural networks with multiattention mechanism[J]. Sensors, 2024, 24(22): 7267.

[13] LIU Y J, LI W H, LIN S Y, et al. Hydraulic system fault diagnosis of the chain jacks based on multi-source data fusion[J]. Measurement, 2023, 217: 113116.

[14] 袁亮. 基于数字孪生的机械臂抓取作业仿真与监控系统[D]. 上海: 东华大学, 2023: 5-15.
YUAN L. Simulation and monitoring system for mechanical arm grasping operations based on digital twins[D]. Shanghai: Donghua University, 2023: 5-15(in Chinese).

[15] TAO F, ANWER N, LIU A, et al. Digital twin towards smart manufacturing and industry 4.0[J]. Journal of Manufacturing Systems, 2021, 58: 1-2.

[16] 蔡宝, 康亮, 顾鸿良. 工业机器人数字孪生系统研究[J]. 自动化与仪表, 2025, 40(8): 62-66.
CAI B, KANG L, GU H L. Research on digital twin of industrial robots[J]. Automation & Instrumentation, 2025, 40(8): 62-66(in Chinese).

[17] 丁珈, 周炜, 张永, 等. 数字孪生驱动的装备全寿命周期智能运维技术进展与展望[J]. 系统工程理论与实践, 2025, 45(6): 1828-1845.
DING J, ZHOU W, ZHANG Y, et al. Digital twin-driven full lifecycle intelligent equipment maintenance: progress and prospects[J]. Systems Engineering—Theory & Practice, 2025, 45(6): 1828-1845 (in Chinese).

[18] LIU X R, DAVID I. AI simulation by digital twins: systematic survey, reference framework, and mapping to a standardized architecture[J]. Software and Systems Modeling, 2025: 1-26.

[19] 李亚洁, 刘强, 李炜. 基于离心泵数字孪生流场云图的叶轮故障诊断方法与应用[J]. 北京航空航天大学学报, 2025, 51(1): 193-201.
LI Y J, LIU Q, LI W. Impeller fault diagnosis method and application based on digital twin flow field contour of centrifugal pump[J]. Journal of Beijing University of Aeronautics and Astronautics, 2025, 51(1): 193-201(in Chinese).

[20] 李志祥, 唐宏宾, 王永胜, 等. 基于数字孪生的液压机械臂运动仿真与状态监控系统设计[J]. 计算机集成制造系统, 2025, 31(12): 4382-4391.
LI Z X, TANG H B, WANG Y S, et al. Design of motion simulation and condition monitoring system for hydraulic manipulator based on digital twin[J]. Computer Integrated Manufacturing Systems, 2025, 31(12): 4382-4391(in Chinese).

[21] 简正豪, 陈卓贤, 刘云鸿, 等. 基于数字孪生的液压回路虚拟实验系统[J]. 实验技术与管理, 2024, 41(9): 111-117.
JIAN Z H, CHEN Z X, LIU Y H, et al. Digital twin-based virtual reality experiment system for a hydraulic circuit[J]. Experimental Technology and Management, 2024, 41(9): 111-117(in Chinese).

[22] 赵大伟, 张太华, 卢尧, 等. 基于数字孪生体的液压系统知识模型构建及应用[J]. 现代制造工程, 2022(6): 17-24.
ZHAO D W, ZHANG T H, LU Y, et al. Construction and application of hydraulic system knowledge model based on digital twin[J]. Modern Manufacturing Engineering, 2022(6): 17-24(in Chinese).

[23] SUN J S, YANG Y S, SHI L, et al. Research on fault diagnosis method of the hydraulic system based on digital twin[J]. Journal of Physics: Conference Series, 2024, 2770(1): 012022.

[24] WANG L T, LIU Y C, YIN H, et al. Fault diagnosis and predictive maintenance for hydraulic system based on digital twin model[J]. AIP Advances, 2022, 12(6): 065213.

[25] LIU H T, ZHANG J C, CHENG Y, et al. Digital twin shop-floor: a complex system-oriented construction method and operation mechanism[J]. Procedia CIRP, 2023, 119: 52-57.

[26] ZHOU D X, XIE M Z, XUAN P C, et al. A teaching method for the theory and application of robot kinematics based on MATLAB and V-REP[J]. Computer Applications in Engineering Education, 2020, 28(2): 239-253.

[27] 宋增来, 陈振华. N 自由度机器人仿真的实现[J]. 计算机工程与设计, 2006, 27(24): 4813-4816.
SONG Z L, CHEN Z H. Simulation of N-dof robot[J]. Computer Engineering and Design, 2006, 27(24): 4813-4816(in Chinese).

[28] XU B, SHEN J, LIU S H, et al. Research and development of electro-hydraulic control valves oriented to industry 4.0: a review[J]. Chinese Journal of Mechanical Engineering, 2020, 33: 29.

[29] 赵洪利, 杨佳强. 基于融合卷积 Transformer 的航空发动机故障诊断[J]. 北京航空航天大学学报, 2025, 51(4): 1117-1126.
ZHAO H L, YANG J Q. Aero-engine fault diagnosis based on fusion convolutional Transformer[J]. Journal of Beijing University of Aeronautics and Astronautics, 2025, 51(4): 1117-1126(in Chinese).

[30] ZHAO J F, MAO X, CHEN L J. Speech emotion recognition using deep 1D & 2D CNN LSTM networks[J]. Biomedical Signal Processing and Control, 2019, 47: 312-323.

[31] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.

[32] HUSSAN U, WANG H Z, PENG J C, et al. Transformer-based renewable energy forecasting: a comprehensive review[J]. Renewable and Sustainable Energy Reviews, 2026, 226: 116356.

[33] LEI Y F, JIANG W L, JIANG A Q, et al. Fault diagnosis method for hydraulic directional valves integrating PCA and XGBoost[J]. Processes, 2019, 7(9): 589.

[34] 肖健华, 樊可清, 吴今培, 等. 应用于故障诊断的 SVM 理论研究[J]. 振动、测试与诊断, 2001, 21(4): 258-262.
XIAO J H, FAN K Q, WU J P, et al. A study on SVM for fault diagnosis[J]. Journal of Vibration, Measurement & Diagnosis, 2001, 21(4): 258-262(in Chinese).

[35] ENCALADA-DÁVILA Á, MOYÓN L, TUTIVÉN C, et al. Early fault detection in the main bearing of wind turbines based on gated recurrent unit (GRU) neural networks and SCADA data[J]. IEEE/ASME Transactions on Mechatronics, 2022, 27(6): 5583-5593.

[36] MU H S, ZHAI X D, YIN D B, et al. A method of remaining useful life prediction of multi-source signals aero-engine based on RF-Transformer-LSTM[C]//Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics. Piscataway: IEEE Press, 2022: 2502-2507.

Fault diagnosis method for hydraulic robots based on digital twin and Transformer-LSTM-XGBoost

LI Yajie[1, 2], WU Ruilong[1, 2], LI Wei[1, 2, *]

(1. School of Automation and Electrical Engineering, Lanzhou University of Technology, Lanzhou 730050, China;
2. Innovation Center of Intelligent Robotics for Nonferrous Metallurgy, Lanzhou University of Technology, Lanzhou 730050)

Abstract: Hydraulic robots are increasingly used in industrial systems, which raises the need for reliable health diagnosis and maintenance under complex operating conditions. An intelligent fault diagnosis method was developed by integrating digital twin technology with deep learning and combining a Transformer, a long short-term memory networks (LSTM), and extreme gradient boosting (XGBoost). A diagnosis architecture based on digital twins was built, and cyber-physical synchronization was made possible by a digital twin with an attribute model and a three-dimensional model. Calibration improved twin accuracy and virtual-physical consistency. Fault mechanism analysis and fault evolution simulation were conducted for four typical hydraulic system faults: leakage, valve sticking, damping orifice blockage, and filter blockage. The simulations generated a dataset covering normal and fault conditions for data-driven modeling and diagnosis. A Transformer-LSTM feature extractor captured global dependencies and temporal dynamics in multi-dimensional time-series data, and XGBoost performed multi-fault classification. The experimental verification results show that the proposed method demonstrate steady performance under various noise disturbances and 96.6% diagnostic accuracy across many problems, suggesting high resilience and generalization and supporting hydraulic robot fault detection and maintenance in industrial applications.

Keywords: digital twin; hydraulic robot; fault diagnosis; deep learning; machine learning

Received: 2025-11-25; **Accepted**: 2026-02-05; **Published Online**: 2026-03-03 13:20

URL: link.cnki.net/urlid/11.2625.V.20260302.1522.002

Foundation items: National Natural Science Foundation of China (62163022); Gansu Provincial Industry Support Program (2024CYZC-18); Gansu Provincial Major Science and Technology Project (23ZDGA007);Gansu Provincial Joint Research Fund (24JRRA827);Lanzhou Science and Technology Plan Project (2024-2-5)

* **Corresponding author**. E-mail: liwei@lut.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0273

脉冲星导航试验卫星核心载荷工作性能稳定性分析

闫林丽[1]，张健康[1]，周庆勇[2,3,*]，雷耀虎[4]，樊少娟[1]

(1. 安徽建筑大学 数理学院，合肥 230601；2. 西安测绘研究所，西安 710054；3. 地理信息工程国家重点实验室，西安 710054；
4. 深圳大学 物理与光电工程学院 光电子器件与系统教育部/广东省重点实验室，深圳 518060)

摘 要：中国首颗脉冲星导航试验卫星XPNAV-1的核心载荷为一款结构精密的聚焦型X射线脉冲星望远镜(FXPT)。在发射振动、空间辐射、器件寿命等因素的综合作用下，FXPT的关键性能参数易发生微小漂移，亟须对其在轨性能开展长期监测，为天文计时分析与导航解算提供高精度、高可靠性的性能参数支撑。为监测XPNAV-1卫星FXPT的性能变化，提出了一种融合光子计数率统计分布和脉冲轮廓形状特征的综合分析方法；基于该卫星2016—2019年对Crab脉冲星的在轨观测数据，完成了对FXPT在轨工作稳定性的定量分析。结果表明：FXPT探测获取的Crab脉冲星光子计数率的统计分布近似服从高斯分布；即使存在少量计数率异常时间段，不同时期的脉冲轮廓仍保持较高的相似度。基于FXPT观测建立的计时模型可折叠得到高信噪比脉冲轮廓，其与Jodrell Bank天文台发布的星历所折叠出的脉冲轮廓相似度达99.99%，充分验证了XPNAV-1卫星核心载荷FXPT具有良好的在轨工作稳定性。

关 键 词：XPNAV-1卫星；X射线脉冲星导航；计时分析；聚焦型X射线脉冲星望远镜；稳定性分析

中图分类号：V557+.4

文献标志码：A **文章编号**：1001-5965(2026)06-1869-11

脉冲星是一类具有极高自转速度和超强磁场的中子星，其辐射的周期脉冲信号非常稳定，且位置能够被精确测定。借助现代天文观测设备对脉冲星进行持续高精度观测，科学家们通过脉冲星计时分析推断出太阳系行星的质量[1]，并建立了基于脉冲星自转频率的时间标尺[2-4]。此外，毫秒脉冲星不仅可作为引力波探测器的关键载荷[5]，还能为深空航天器提供自主导航与授时服务[6-7]。X射线脉冲星自主导航是一种新型的天文自主导航技术，其通过探测脉冲星辐射的X射线光子信息，融合脉冲星天体测量信息及航天器轨道动力学，为深空航天器提供位置、姿态和时间等参数，从而实现深空航天器持续安全自主导航。该技术具备高度自主性与安全性等突出优势，且导航误差不会随作用距离增加而累积[8-9]，是当前国际航天领域的研究热点。精确测定X射线脉冲星的脉冲轮廓、辐射特性及自转周期等信息，是开展脉冲星导航应用的前提条件[10]。但脉冲星信号精确测定受X射线望远镜在轨性能的影响，高精度脉冲星导航服务的实现，通常要求X射线望远镜具备精确的时间测量精度、良好的能量分辨率等性能。由于绝大多数X射线脉冲星辐射流量较弱，且空间弥散高能射线和带电粒子，不可避免地会在X射线望远镜上产生本底噪声，同时望远镜电子学噪声也会进一步增加读出噪声水平，因此，需要采用性能稳定、探测能力优异的X射线望远镜对脉冲星开展长时间观测，以获得高精度导

收稿日期：2024-04-29；**录用日期**：2024-06-18；**网络出版时间**：2024-06-20 17:07
网络出版地址：link.cnki.net/urlid/11.2625.V.20240620.1538.002
基金项目：国家自然科学基金(11903001,42004004)；安徽省教育厅中青年教师培养行动项目(YQZD2023053)；安徽建筑大学校博士启动基金项目(2019QDZ14,2019QDZ31)；安徽省高校优秀青年科研项目(2024AH030011)；安徽建筑大学大学生创新创业训练计划(202310878057)
* **通信作者**. E-mail：zjlzqy1986@163.com

引用格式：*闫林丽，张健康，周庆勇，等. 脉冲星导航试验卫星核心载荷工作性能稳定性分析[J]. 北京航空航天大学学报，2026，52(6)：1869-1879. YAN L L，ZHANG J K，ZHOU Q Y，et al. Stability analysis of core payload operational performance on the XPNAV-1 satellite[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1869-1879 (in Chinese).*

航所需的X射线标准脉冲轮廓，为脉冲星导航解算提供基本的观测信息。

欧美一直高度重视X射线脉冲星导航技术的研究。2004年，美国发布了X射线脉冲星导航计划XNAV，其目标是追求宇宙全空域优于30 m的导航精度，该计划技术实现难度特别大，执行一年后被暂停，相关研究成果被NASA继承并转变为SEXTANT项目的前身。同时，欧洲航天局（European Space Agency, ESA）主要成员国德国和西班牙也积极推动脉冲星导航技术的发展。2004年，ESA发布了基于脉冲星计时信息的航天器导航可行性研究技术报告；2015年，ESA在国际上首次组织了以脉冲星导航为主题的国际会议，德国Becker教授持续在该领域进行脉冲星数据库构建、脉冲星导航探测器论证、导航算法等研究工作。2015年6月，脉冲星自主导航和X射线通信技术作为导航通信领域“革命性概念”被NASA列入空间发展规划（2015—2035年）[11]。2018年，美国利用安装在国际空间站上的中子星内部组成探测器NICER首次进行了在轨毫秒脉冲星自主导航空间试验，通过序贯观测4颗毫秒脉冲星，将国际空间站位置精度收敛到16 km，最佳情况下可达5 km[10]，展现出X射线脉冲星导航技术良好的技术应用前景。X射线脉冲星导航技术作为星际空间唯一的航天器自主导航手段，能够节省地面宝贵的测控资源。此外，借助对毫秒脉冲星的观测能实现对时间基准长期稳定性的监测[10]。基于这一独特优势，脉冲星计时导航技术已成为各国航天领域竞相布局的战略方向，也引起了中国航天机构的高度关注与重点部署。

2000年以来，中国一直积极推动脉冲星自主导航技术的发展，扩大航天器自主运行的范围，以满足导航定位和授时体系（positioning, navigation & timing, PNT）深空场景的建设需求。中国同样高度重视X射线脉冲星导航技术的工程化实践，并于2016年首次发射了脉冲星导航试验卫星XPNAV-1。XPNAV-1卫星由中国空间技术研究院研制，是中国首颗专门用于验证脉冲星导航技术的试验卫星。XPNAV-1卫星自投入运行以来，累积了大量空间观测数据，国内研究人员开展了大量脉冲星信号处理和计时等方面的工作。通过对这些数据进行深入分析，成功探测到了Crab脉冲星（PSR B0531+21）的信号，并初步开展了脉冲星导航体制验证工作。黄良伟博士等研究了XNAVP-1卫星计时数据和轮廓恢复算法[12]，并利用XPNAV-1卫星85 d观测数据进行了脉冲星导航解算试验，通过将Crab脉冲星的计时观测量作为控制点，成功抑制了卫星轨道传播误差，在控制点处的平均导航误差为38.4 km[13]；李连升博士等实现了在轨观测数据的轮廓恢复，实现了国产首款聚焦X射线望远镜成功探测到脉冲信号的目标[14]；帅平研究员等通过分析最初4个月的观测数据，给出脉冲星观测数据的处理方法及脉冲星参数的拟合过程[15]；周庆勇和闫林丽博士等对XPNAV-1卫星的聚焦型X射线脉冲星望远镜（focusing X-ray pulsar telescope, FXPT）性能进行了在轨测试[9]，且利用XPNAV-1卫星3 d的观测数据实现基于最小二乘算法的脉冲星导航试验解算，得到卫星轨道确定精度约为56.93 km[16]。

FXPT是XPNAV-1卫星的主载荷，属于一种结构精密的科学仪器。当FXPT随卫星被运载火箭发射升空时，火箭强烈的冲击力可能导致其聚焦光学系统镜片结构与探测器之间产生微小位移，破坏地面装调保持的高精度。而且，在轨的恶劣热环境也可能导致X射线聚焦镜片面型与结构发生微小变化，从而影响FXPT的聚焦性能。此外，随着时间推移，复杂的空间电磁与粒子辐射会加速FXPT内部元器件老化，特别是焦平面X射线硅漂移探测器（silicon drift detector, SDD），进而导致望远镜性能参数发生退化，如能量分辨率呈现逐渐略化的趋势。因此，望远镜在轨工作稳定性需开展长期持续监测，并需及时对其性能参数进行标定与校正，以确保脉冲星观测及后续数据分析结果的准确性。在性能监测方面，大面积X射线望远镜通常会配备标定放射源，以对望远镜部分被遮挡区域进行标定。然而，XPNAV-1卫星空间载荷容量有限，难以携带标定装置，且缺乏对具有特征能谱辐射的超新星遗迹的持续观测数据。笔者前期已提出一种基于望远镜本征特征能谱与超上阈信号联合判别其在轨稳定性的监测方法[17]，但受限于望远镜本征特征能谱分辨率低、超上阈信号微弱，且其产生机理尚不明确、误差来源复杂多样等问题，监测精度与可靠性难以满足长期在轨评估需求。因此，需探索新型、高效的监测方法，以精准表征在轨卫星核心载荷的性能演化与工作稳定性。在天文学领域，Crab脉冲星常被用作探测器性能标定源，其X射线辐射非常稳定。因此，通过观测Crab脉冲星辐射特性的变化，可直接反映望远镜在轨性能的变化。FXPT的探测效率、有效面积与噪声水平等主要指标的一个重要表象是脉冲星的光子流量，因此可依托光子流量开展FXPT工作性能分析。此外，基于长时间观测Crab脉冲星得到的脉冲轮廓与标准轮廓的差异，同样能够反映FXPT的在轨工作稳定性。尽管上述2种方法对FXPT工作稳定性的判断存在一定时间滞后性，但可有效反映观测期间的长

期运行稳定性。本文基于 XPNAV-1 卫星 2016—2019 年的 Crab 脉冲星观测数据, 从光子计数率统计分布与脉冲轮廓形状 2 个方面, 综合评估了 FXPT 的在轨工作稳定性。

1　XPNAV-1 卫星及其科学数据

1.1　XPNAV-1 卫星

XPNAV-1 卫星于 2016 年 11 月 9 日发射升空, 其轨道为太阳同步近圆形轨道, 高度为 500 km, 质量为 270 kg, 结构如图 1 所示[18]。XPNAV-1 卫星提供 4 种观测模式: 自检模式、扫描模式、默认目标观测模式和任意目标指向模式。自检模式需要在镜头被罩住的情况下工作, 以便测量仪器噪声。扫描模式通过旋转卫星扫描空间带, 以评估背景噪声。默认目标观测模式使 X 射线仪器指向其中一个默认目标进行观测。任意目标指向模式类似于默认目标观测模式, 但目标角位置参数从地面上传。XPNAV-1 卫星采用了集成电子设计, 将载荷数据处理、跟踪遥测与指令、姿态确定和控制、GNSS 导航及电源控制等功能集成在一个电子组件中, 采用零动量三轴姿态控制系统, 具有高精度的惯性指向能力, 采用三轴稳定的姿态控制模式, 能够根据需求快速指向任意惯性位置, 精度可达 2′, 能够提供长达 90 min 的持续观测。

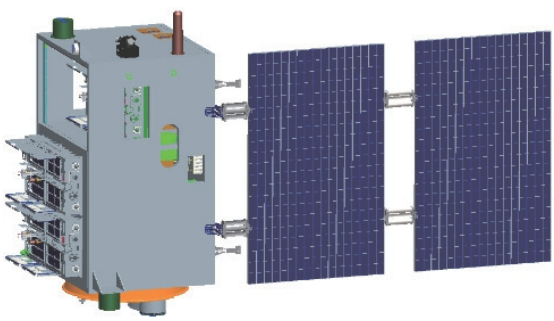


图 1　XPNAV-1 卫星结构[18]

Fig. 1　XPNAV-1 satellite structure[18]

XPNAV-1 卫星的主要任务是通过自主研发的 FXPT 测试软 X 射线脉冲星观测技术, 开展初步的脉冲星导航空间试验。卫星工程的 3 个试验目标为: ①测试 X 射线仪器在外层空间的功能和性能; ②探测典型 X 射线脉冲星的辐射光子并获取脉冲轮廓, 以验证 X 射线脉冲星观测的能力; ③长时间积累 X 射线观测数据, 通过 X 射线计时来测量脉冲星的参数。通过 3 个目标的实施, 促进脉冲星物理特性及导航体制的研究。XPNAV-1 卫星主要载荷 FXPT 是一台 Wolter-I 型 X 射线望远镜, 其主要性能参数如表 1 所示[18]。FXPT 聚焦结构采用 4 层镜片嵌套, 采集面积为 30 cm^2, 视场为 15′, 聚焦结构能将视场内的 X 射线光子聚焦到面积 25 mm^2 的 SDD 上[18]。当 X 射线光子与 SDD 探测器相互作用时, 会产生电子-空穴对, 空穴被附近的阴极电极捕获, 而电子则在电场驱动下向阳极漂移, 被阳极收集并转换为电信号, 通过测量这些电信号的幅值, 可以确定相应光子的能量。当电信号幅值超过预设阈值时, 将触发记录一个特定时刻, 并利用星载铷钟标记该时刻, 即光子到达时间, 也就是单个 X 射线光子事例到达探测器的时间。

表 1　XPNAV-1 卫星上 FXPT 的参数指标[18]

Table 1　Parameters and specifications of FXPT on the XPNAV-1 satellite[18]

参数	指标
探测元件	SDD
聚焦光学系统	Wolter-I
探测能段/keV	0.5～10
光学视场/(′)	15
1 keV时的有效面积/cm^2	3
时间分辨率/μs	≤1.5
5.9 keV处能量分辨率/eV	≤180

在后续的科学数据分析中, 除了需要上述的 X 射线光子信息外, 还必须明确卫星的运行状态参数, 如卫星的位置速度、姿态、指向等信息, 这些信息在观测中被保存在数据文件中, 其数据格式本文不详述。

1.2　观测数据筛选

尽管 XPNAV-1 卫星的设计寿命仅 1 年, 但截至目前其已在轨运行超过 9 年时间, 仍正常工作。由于 FXPT 有效探测面积相对较小, 仅适于观测 X 射线信号流量较高的脉冲星。与其他脉冲星相比, Crab 脉冲星辐射的 X 射线光子流量较强, 达到 1.54 counts·cm^{-2}·s^{-1}。Crab 脉冲星在 J2000.0 坐标系下的赤经、赤纬分别为 05 h 34 min 31. 972 s、+22° 00′52″.07, 距离地球约 2 kpc (约 6.17×10^{19} m)。Crab 脉冲星是一颗年轻的脉冲星, 特征年龄约 1 100 年, 位于超新星遗迹中心, 被蟹状星云所包裹, 其具有较强的宽波段辐射能力, 常被选为空间 X 射线望远镜的首选观测目标和标定信号源。

XPNAV-1 卫星对 Crab 脉冲星进行了高频次的观测, 累积了大量空间观测数据。本文主要筛选了 2016 年 11 月至 2019 年 12 月之间的数据进行统计分析。数据中, 光子到达时间文件记录了光子到达望远镜焦平面探测器的时间和能量信息, 包括来自脉冲星及其星云的 X 射线光子、空间粒子噪声与 SDD 本身的热流噪声等。需要注意的是, XPNAV-1 卫星观测脉冲星期间, 在多个观测时间段内出现了

光子流量分布涨落较大的现象，如图 2 所示，这可能与卫星姿态、太阳辐射、宇宙射线有关。以2016年的观测为例，部分观测时间段内，光子流量随时间呈现较大的涨落变化。经分析发现，这种变化与卫星轨道区域、观测指向等因素密切相关。由于本文研究目标是检验 XPNAV-1 卫星上 FXPT 的整体观测稳定性，对观测数据进行数据质量较高的时间段筛选时，剔除了受空间环境干扰的短时间高流量观测数据。

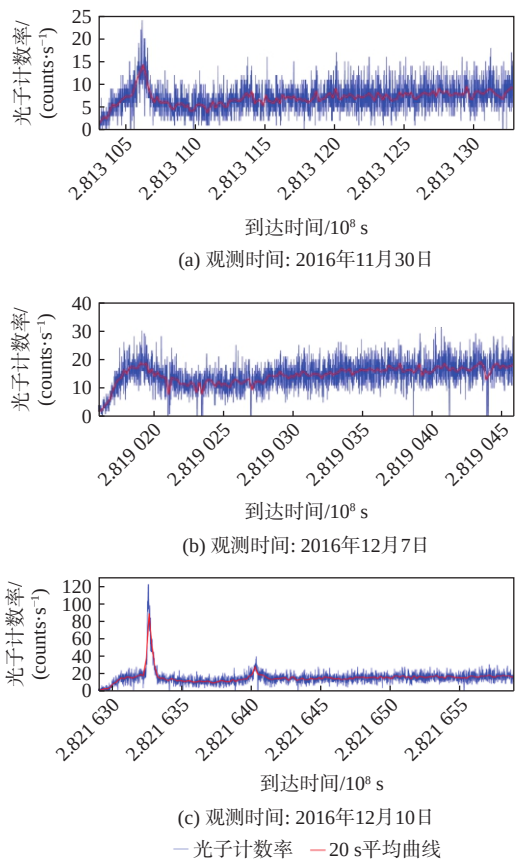


图 2 XPNAV-1 卫星上 FXPT 在 3 个不同时间段内的光子计数率随时间的变化曲线

Fig. 2 Temporal variation of photon count rate of FXPT on the XPNAV-1 satellite during three different time periods

XPNAV-1 卫星轨道周期约 90 min，由于能源配置等原因，卫星在观测脉冲星一轨后，需停止观测，进行一轨充电操作。受地球遮挡的影响，每轨最长能观测脉冲星约 50 min。本文共收集整理 XPNAV-1 卫星对 Crab 脉冲星的 551 个数据文件，包含 1 455 轨观测，发现 1 d 内最多约有 9 h 用于 Crab 脉冲星观测，即每天约有 6 轨 Crab 脉冲星观测数据。经分析发现，光子计数率超 100 counts/s的观测共计 29 轨（类似图 2(c) 的情况），仅占比 1.99%，观测时长占比更低。可见，异常数据的剔除不影响对 FXPT 性能长期稳定性的判断。

地球上，南大西洋异常区（South Atlantic anomaly, SAA）的磁场强度显著低于其他区域，区内带电粒子流量远高于同轨道高度的其他空间位置。低轨道航天器穿越该区域时，会直接穿过辐射带并遭受高能粒子的持续轰击，星载电子设备的工作状态易受到干扰。SAA 的空间分布范围为：南纬−30°～−10°、西经 90°～东经 30°。本文依据卫星轨道文件，计算出图 2(c) 中对应时间段内卫星的位置坐标，结果如图 3 所示。图 2(c) 中，光子计数率最大时刻卫星的天球坐标为 (336.18°, 60.85°)，光子计数率次大时刻的天球坐标为 (185.66°, 68.39°)，而卫星在这段时间内并未经过 SAA 区域。由此表明，图 2(c) 出现的流量突变与卫星是否经过 SAA 无关，其具体物理成因仍有待进一步深入探究。

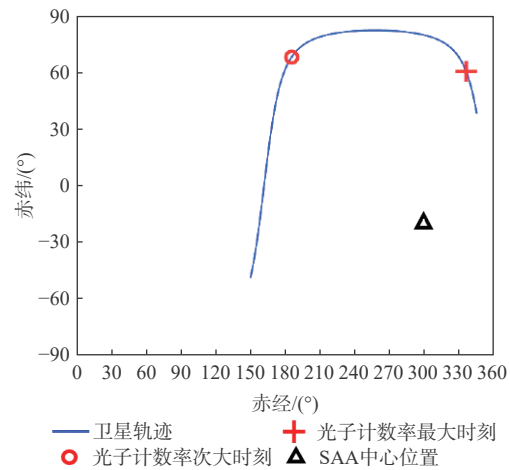


图 3 卫星在 2016 年 12 月 10 日内一段观测时长对应的天球坐标

Fig. 3 Celestial coordinates of satellite during a segment of observation on December 10, 2016

1.3 光子到达时间的太阳系质心修正和相位计算

卫星在观测过程中持续运动，其所处的引力场不断变化，星载时钟记录的光子到达时间需要转换到惯性坐标系中的某一参考点。以太阳系质心 (solar system barycenter, SSB) 为中心的天球参考坐标系是准惯性坐标系，通常选取 SSB 为坐标系原点，将探测器记录的脉冲星光子到达时间 $t_{\rm obs}$ 转换成到达 SSB 的时间 $t_{\rm SSB}$[19]，这一步骤称为太阳系质心修正。本文对 X 射线光子的质心修正公式为

$$t_{\rm SSB}=t_{\rm obs}+\frac{\boldsymbol{n}\cdot\boldsymbol{r}}{c}-\frac{1}{2cR_0}[(\boldsymbol{r})^2-(\boldsymbol{n}\cdot\boldsymbol{r})^2]-\frac{1}{c}(\boldsymbol{n}\cdot\boldsymbol{v})\Delta t+\frac{1}{cR_0}[(\boldsymbol{r}\cdot\boldsymbol{v})-(\boldsymbol{n}\cdot\boldsymbol{r})(\boldsymbol{n}\cdot\boldsymbol{v})]\Delta t+\Delta t_{\rm c}+\frac{2}{c^3}\sum_{k=1}^{N}\mu_k\ln(|\boldsymbol{r}_k|+\boldsymbol{n}\cdot\boldsymbol{r}_k)-\frac{4\mu_{\rm S}^2}{c^5|\boldsymbol{r}_{\rm S}|\tan\psi\sin\psi} \quad (1)$$

式中：$t_{\rm SSB}$ 和 $t_{\rm obs}$ 分别为子到达 SSB 和探测器的时间；$\boldsymbol{n}$为脉冲星在太阳系质心坐标系的单位矢量；

$\boldsymbol{r}$为望远镜相对于质心位置矢量; c 为光速; $\boldsymbol{v}$为脉冲星相对于 SSB 的速度矢量; $\Delta t = t - t_0$为观测时刻 t 与参考历元t_0之间的差值; R_0为参考历元下脉冲星的距离; Δt_c为望远镜时钟修正至质心坐标时(barycenter coordinate time, TCB)尺度的校正项, 包括原子钟钟差和各种爱因斯坦时延; μ_k为太阳系第k个天体的引力常数; $\boldsymbol{r}_k$为太阳系第k个天体相对于望远镜的位置矢量; N为在太阳系内考虑的行星天体数量; μ_S为太阳引力常数; $|\boldsymbol{r}_S|$为太阳相对于望远镜位置矢量的模; ψ为太阳和脉冲星相对于望远镜的张角。式 (1) 等号右边第 2～5 项分别为几何真空时延、视差、脉冲星的径向速度引起的时延和一阶 Röemer 时延; 第 7 项和第 8 项分别为太阳系各种引力时延和太阳二阶引力时延。第 7 项中, 太阳系各种引力时延仅考虑八大行星的引力时延。对于空间 X 射线观测数据处理来说, 由于 X 射线的波长极短且能量很大, 其色散时延可忽略不计。

脉冲星辐射信号具有周期性, 可以统计整个观测光子数量在一个周期内的统计分布, 该分布函数被称为脉冲星的脉冲轮廓。利用泰勒级数可将光子的 SSB 到达时间转化成相位, 公式如下:

$$\phi(t) = \phi_0 + F_0(t-t_0) + \frac{1}{2}F_1(t-t_0)^2 + \frac{1}{6}F_2(t-t_0)^3 + \cdots \tag{2}$$

式中: ϕ_0为参考历元t_0时刻的绝对相位; F_0为该时刻的自转频率; F_1、F_2分别为自转频率的一阶、二阶导数。这些参数共同构成脉冲星星历参数。根据式(2), 可以获得星历有效时间段内任一时刻脉冲星的自转相位$\phi(t)$。自 1985 年起, 英国 Jodrell Bank 天文台利用其射电望远镜的观测, 定期发布 Crab 脉冲星的星历。利用这些星历, 可对 XPNAV-1 卫星观测到的 X 射线光子到达时间进行相位换算。表 2 列出了 Jodrell Bank 天文台发布的 Crab 脉冲星部分星历, 本文主要利用该星历来实现对 X 射线光子相位的精确计算。

表 2　Jodrell Bank 天文台发布的 Crab 脉冲星星历 (MJD 58 027～58 318)

Table 2　Ephemeris of the Crab pulsar released by the Jodrell Bank Observatory (MJD 58 027−58 318)

T_1/MJD	T_2/MJD	$T_{0,\mathrm{geo}}$/MJD	F_0/s^{-1}	$F_1/10^{-10}$ s^{-2}	$F_2/10^{-20}$ s^{-3}
58 027	58 058	58 042.000 000 104	29.637 435 208 764 4	−3.68636	1.93
58 058	58 063	58 060.000 000 217	29.636 861 925 018 0	−3.68612	6.46
58 065	58 071	58 068.000 000 228	29.636 621 566 643 3	−3.69981	−0.081 0
58 071	58 089	58 080.000 000 333	29.636 237 075 846 7	−3.70847	0.294
58 088	58 120	58 104.000 000 331	29.635 468 845 132 7	−3.70141	0.281
58 119	58 150	58 134.000 000 229	29.634 510 198 632 8	−3.69604	0.176
58 150	58 178	58 164.000 000 327	29.633 552 635 302 0	−3.69318	9.27
58 178	58 209	58 194.000 000 208	29.632 595 645 123 2	−3.69147	3.44
58 209	58 238	58 224.000 000 223	29.631 639 010 174 4	−3.69007	3.16
58 239	58 270	58 255.000 000 322	29.630 650 839 376 5	−3.68916	4.60
58 270	58 300	58 285.000 000 004	29.629 694 714 016 6	−3.68847	2.63
58 300	58 318	58 309.000 000 366	29.628 929 903 171 1	−3.68799	7.36

注: $[T_1, T_2]$为自转参数有效的时间段, 用MJD时间表示; $T_{0,\mathrm{geo}}$为单个射电脉冲到达地心的时间, 其整数部分为射电脉冲到达太阳系质心的时间, 即各阶自转参数F_0、F_1、F_2对应的时间参考点。

针对本文所用的脉冲星光子到达时间数据, 为提升数据处理结果精度, 删除由于遥测遥控遗失的数据, 去除含有错误能量信息及重复记录的 X 射线光子数据。同时, 为降低空间背景粒子噪声的干扰, 剔除光子计数率大于 50 counts/s 的时段内的观测数据。另外, 本文只选取了在 Jodrell Bank 天文台发布的星历覆盖时间段内的光子数据。综合分析对 Crab 脉冲星共 551 次观测的数据, 共计约 6 000 万个光子, 总观测时长约 1 100 d, 总曝光时间约为 4.4×10^6 s, 所选取的数据覆盖了 Jodrell Bank 天文台射电星历中的 28 个时间段。

2　FXPT 稳定性分析

基于 XPNAV-1 卫星获取的脉冲星光子到达时间信息, 统计分析了 X 射线光子流量统计分布函数的时间稳定性, 以及脉冲轮廓形状的时间稳定性。

2.1　光子流量的统计分布

脉冲星光子流量随时间的变化能够反映 X 射线望远镜有效面积、噪声抑制能力等关键性能指标。在对 Jodrell Bank 天文台发布的星历所覆盖的 28 个月期间观测数据分析时, 综合统计了每个月内光子计数率随时间的统计分布, 如图 4 所示(图中

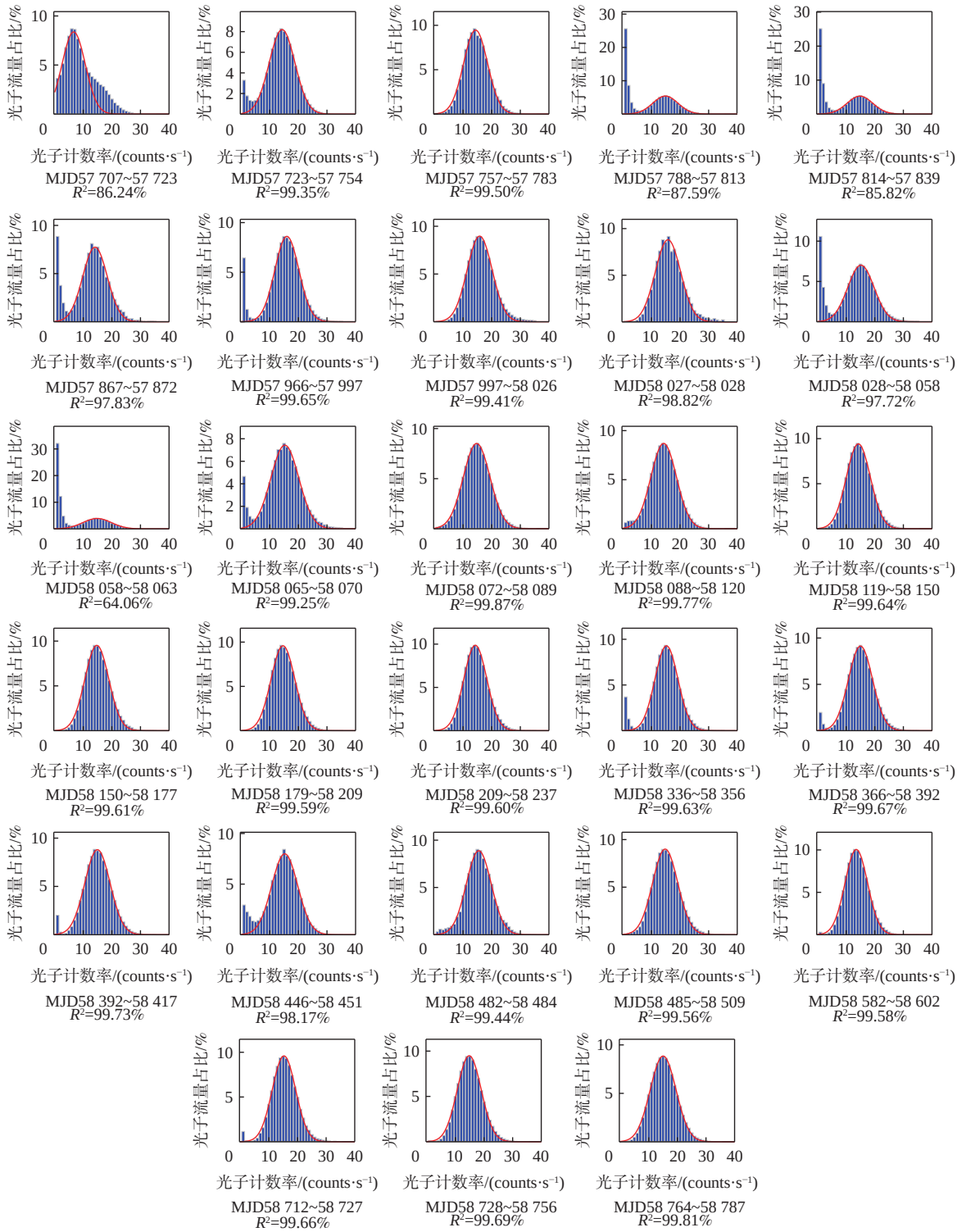


图 4　28 个月的观测期中每个月光子计数率的统计分布及其高斯拟合曲线

Fig. 4　Statistical distribution of photon count rates for each month during the 28-month observation period along with their Gaussian fitting curves

标示了每个月观测光子数的 MJD 时间范围)，可见，光子计数率的统计分布近似服从高斯分布。在对光子流量进行统计时，进行归一化处理，纵坐标为光子流量的占比，各光子计数率时间段的光子流量占比之和为 100%。为定量描述光子计数率的时间稳定性，采用高斯函数 $f(x)=N_0\mathrm{e}^{-(x-x_0)^2/2\sigma^2}$ 对数据进行拟合，其中，N_0为归一化因子，x_0为高斯分布的中心值，σ为高斯分布的标准偏差，分布的半高全宽 (full width at half maximum, FWHM) 可表示为 $w=2\sqrt{2\ln 2}\sigma$，光子计数率统计分布的高斯函数拟合结果如图 4 中红色曲线所示。同时，用决定系数 R^2 衡量高斯函数拟合效果，其计算公式为

$$R^2 = 1 - \frac{\sum_i \left(x_i - x_{\mathrm{m},i}\right)^2}{\sum_i \left(x_i - \bar{x}\right)^2} \tag{3}$$

式中: x_i为光子计数率的统计分布的观测值; $\bar{x}$为平均值; $x_{\mathrm{m},i}$为模型预测值。当R^2接近 1 时, 表明模型很好地解释了数据的变异性, 即拟合效果好; 反之, 当R^2接近 0 时, 则代表拟合效果差。为排除低光子计数率光子噪声的影响, 并分析 FXPT 整体工作稳定性, 计算了光子计数率统计分布在 3～35 counts/s 范围内的决定系数值, 标记在图 4 中。结果显示, 只有 4 个月的观测数据的决定系数低于 90%, 而针对其他月份的观测, 均可以用高斯函数很好地拟合计数率分布曲线的形状, 决定系数均在 97% 以上, 表明了 FXPT 的工作性能具有较高的稳定性。同时发现, 部分月份的低光子计数率时间段所占比例较大, 导致这种情况的原因有很多。例如, 某些月份可能有较长的观测时间段受到卫星姿态调整的影响, 从而导致较低的光子计数率; 另外, 空间环境(如空间天气、地磁场扰动等)在不同月份可能存在较大变化, 从而导致部分月份的低光子计数率时间段较多; 在数据处理和校正过程中, 某些月份的数据可能存在更多的异常或不完整情况, 需要进行更多的数据剔除和校正, 导致这些月份显示出更多的低光子计数率时间段; 同时, 探测器的性能(如探测效率、背景噪声等)可能会随着时间发生变化, 尤其是受到空间辐射、温度变化等因素的影响, 这些变化在某些月份可能更加明显, 从而导致低光子计数率时间段的增加。具体原因需要结合观测日志、空间环境数据、探测器状态记录等进行更详细的分析才能确定, 而当前获取这些数据存在特别多的困难。

用N_0、x_0及w描述光子计数率统计分布曲线的形状, 来进一步定量说明各个形状参数随时间的变化趋势, 如图 5 所示。综合来看, 除第 1 个月进行观测模式调整导致数据异常外, 统计分布曲线的中心值随时间表现非常稳定, 均值为 14.87 counts/s, 宽度值稳定在 7.26 counts/s, 但存在一定的涨落变化, 标准偏差约为 0.42 counts/s。由于部分月份的观测存在一些计数率极小和极大的时间段, 导致统计分布曲线的归一化因子N_0也存在涨落变化。整体来看, 这 3 个参数的涨落变化随时间趋于稳定, 与平均值的偏离程度越来越小, 表明 FXPT 的工作性能逐渐趋于稳定。结合图 4 和图 5, 在 28 个月的观测期间, 即使存在一些低光子计数率的观测时间段, Crab 脉冲星光子计数率的统计分布整体上仍符合高斯分布。

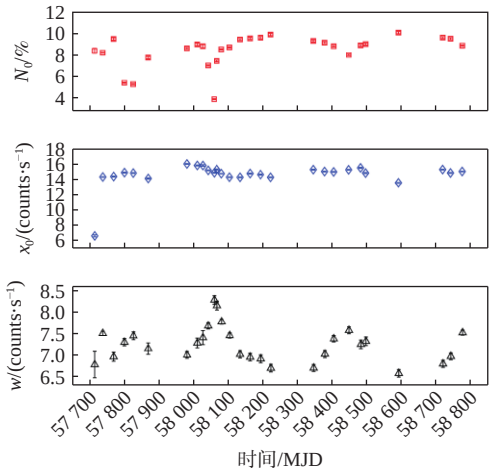


图 5　光子计数率统计分布曲线的形状参数随时间的变化

Fig. 5　Variation of shape parameters of photon count rate statistical distribution curve over time

2.2　脉冲轮廓形状演化

脉冲轮廓是光子数在一个脉冲星旋转周期内的累积分布曲线, 是标识脉冲星辐射特性的特征。利用式 (2) 计算出 X 射线光子相对于参考时刻的相位, 并将相位归算到一个周期内, 即可获得 Crab 脉冲星的脉冲轮廓, 28 个月的观测期中, 每个月的脉冲轮廓形状如图 6 所示。习惯上将先到达的脉冲称为主峰, 后到达的脉冲称为次峰, 图 6 中将主峰强度进行了归一化处理, 每个月的脉冲轮廓平均信噪比为 70。

Jain 等从峰值幅度比、峰值相位差和脉冲宽度 3 个方面研究了 Crab 脉冲星的 X 射线脉冲轮廓变化趋势, 这 3 个指标能够描述脉冲轮廓形状的局部变化[20]; 孙海峰等采用 Pearson 相关系数、标准偏差和功率谱熵 3 个指标量化脉冲轮廓的稳定性, 并利用罗西 X 射线计时探测器对 Crab 脉冲星 11 年的观测数据, 系统分析了脉冲轮廓的稳定性[21-22]。在以上研究中, 进行稳定性分析之前均要求对脉冲轮廓进行噪声消除的预处理。本文采用相似度 S 量化脉冲轮廓形状的时间稳定度, 该方法需要提供一个模板轮廓用于比较, 但不需要对轮廓进行降噪处理。将图 6 中 28 个月内的轮廓相位对齐后叠加, 得到一个模板轮廓, 如图 7(a) 所示。比较每个月内的脉冲轮廓与标准轮廓之间的相似度。相似度S的计算方法如下:

$$S = \frac{\int_0^T P(t)P_0(t)\mathrm{d}t}{\int_0^T P^2(t)\mathrm{d}t \int_0^T P_0^2(t)\mathrm{d}t} \tag{4}$$

式中: $P(t)$为观测脉冲轮廓; $P_0(t)$为模板轮廓; T 为脉冲星自转周期。

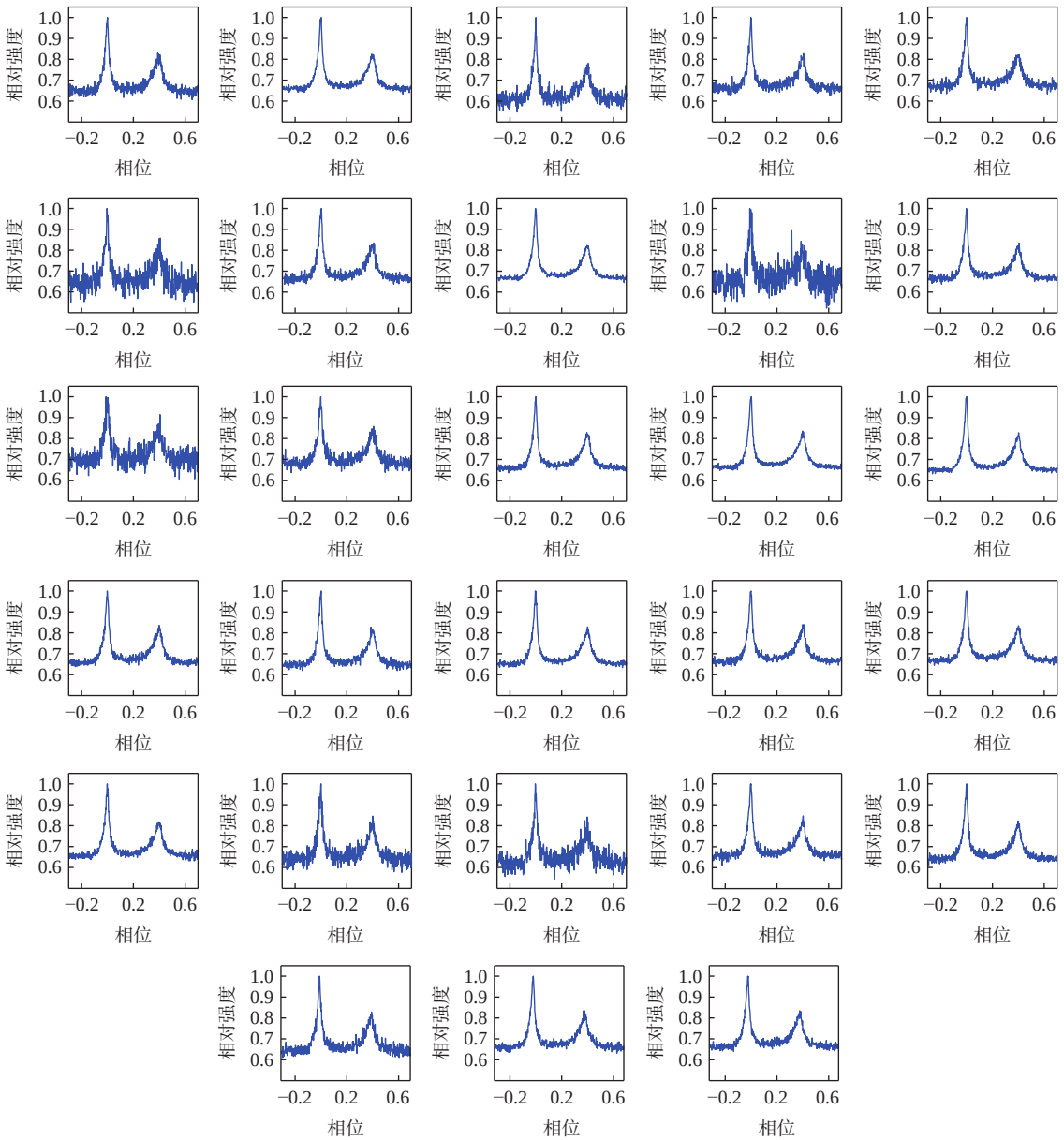


图 6 28 个月的观测期中每个月观测的脉冲轮廓

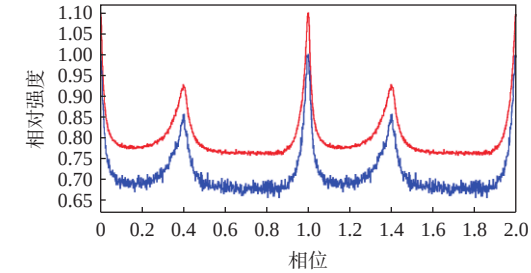
Fig. 6 Pulse profiles observed for each month during the 28-month observation period.

比较结果示于图 8(a)。相似度无法直接分辨脉冲轮廓的本底强度 (即非脉冲部分的强度) 变化，只能用来比较轮廓形状的相对起伏高度是否相似。从图形上看，尽管不同时期观测本底强度不一致，但其轮廓形状却高度相似。从图 8(b) 中可以发现，不同月份观测的脉冲轮廓非脉冲部分的辐射强度各不相同，如第 11 个月的非脉冲辐射强度最高，这主要是因为当月观测低光子计数率时间段较多，如图 4 所示。因此，非脉冲辐射强度也应作为表征脉冲轮廓形状时间稳定度的一个重要指标。

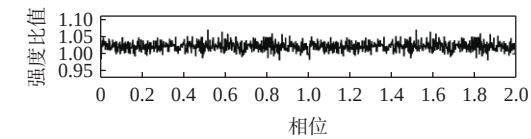
如图 8(b) 所示，脉冲轮廓的非脉冲辐射强度稳定在 0.66 附近，但个别月份的非脉冲辐射强度涨落幅度较大，尽管如此，脉冲轮廓的相似度仍均高于 99.7%，表明轮廓形状具有高度相似性。因此，可推断光子计数率异常的时间段内的光子并非来自脉冲星，而是来自空间噪声抖动。随着卫星运行时间加长，脉冲轮廓相似度和非脉冲辐射强度这 2 项特征参数的涨落幅度逐渐降低，并趋于稳定，这也表明 FXPT 性能是稳定的。

2.3 计时模型参数拟合

应用 XPNAV-1 卫星观测数据进行导航分析的前提是获取准确的脉冲星自转参数。通过对 X 射线观测数据进行计时分析，建立 Crab 脉冲星的计时模型，获取脉冲星 X 射线频段的星历，能够预测有效时间段内任意时刻脉冲星的自转信息。Tempo2 是专门用于处理和分析脉冲星计时数据的分析软件，可以通过拟合脉冲星的到达时间，构建精确的脉冲星计时模型，这些模型包括脉冲星的自



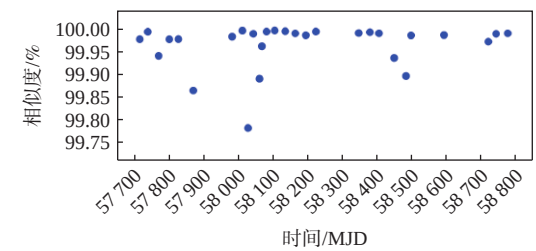
(a) MJD 58 074~58 101期间的脉冲轮廓与模板轮廓的对比



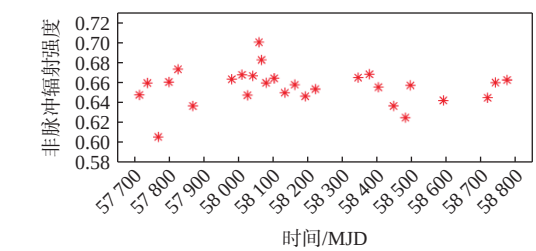
(b) 脉冲轮廓与模板轮廓的相对强度比值

图 7　Crab 脉冲星模板轮廓和 MJD 58 074～58 101 期间的脉冲轮廓及其比值

Fig. 7　Crab pulsar's template profile and pulse profiles during MJD 58 074−58 101, and their ratio



(a) 每月观测脉冲轮廓与模板轮廓的相似度



(b) 脉冲轮廓的非脉冲辐射强度

图 8　每月观测脉冲轮廓与模板轮廓的相似度和非脉冲辐射强度

Fig. 8　Similarity between monthly observed pulse profiles and the template profile, and non-pulsed radiation intensity

转频率及其高阶导数、位置、运动速度和轨道参数等[23]。本文选择 2016—2019 年内观测最密集时间段(MJD 58 074～58 101)中的 38 次观测, 计算每次观测的脉冲到达时间, 输入 Tempo2 中, 拟合获取该时间段内的计时模型。每次观测的脉冲轮廓主峰在一个周期内出现的相对时刻与此次观测的第一个 X 射线光子的质心到达时间相加, 记为此次观测脉冲的到达时间。采用 X 射线计时分析方法[24], 用 Tempo2 拟合得到的计时残差分布如图 9 所示。脉冲星在 t_0= 58 080 MJD 时刻的自转频率 F_0、自转频率一阶导数 F_1、二阶导数 F_2 列于表 3 中, 其中, F_0、F_1、

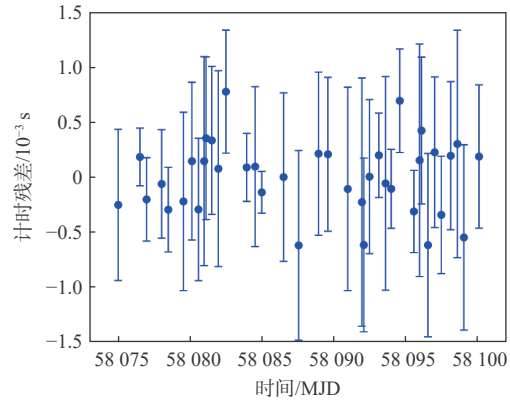


图 9　Crab 脉冲星的计时残差

Fig. 9　Timing fit residuals of Crab pulsar

F_2 参数中括号内的数字表示误差, 如 0.123 456 78 (9) 表示 0.123 456 78±9×10^{-8}。从图 9 中可以看出, 拟合残差在 0 附近均匀分布, 其均方根值为 281.826 μs, 该自转模型只在时间段 MJD 58 074～58 101 内有效。根据计时模型折叠出的脉冲轮廓形状(见图 7(a) 中蓝色线)所示, 信噪比达到 148, 与由 Jodrell Bank 天文台提供的星历折叠得到的模板轮廓相似度高达 99.99%, 这表明利用 FXPT 可以获得高质量的脉冲星信号。

表 3　Crab 脉冲星在 MJD 58 074～58 101 期间的计时模型

Table 3　Timing model of Crab pulsar during MJD 58 074−58 101

参数	数值
t_0/MJD	58 080
T_1/MJD	58 074
T_2/MJD	58 101
F_0/s^{-1}	29.636 237 078 5(64)
F_1/s^{-2}	−3.709 2(3)×10^{-10}
F_2/s^{-3}	4.79(48) ×10^{-19}
均方根值/μs	281.826

将 MJD 58 074～58 101 期间的脉冲轮廓与模板轮廓相除, 分析其比值在相位上的分布, 如图 7(b) 所示, 2 个轮廓的比值分布在 1.02 附近, 而非 1 附近。在不同的观测时间段内, 观测条件和背景噪声等因素可能会有所变化, 从而导致脉冲轮廓的本底强度有所不同。但图 7(b) 显示观测时间段内的脉冲轮廓与模板轮廓的比值在整个周期内基本一致, 这表明 FXPT 探测器在不同观测时段均具有一致的性能表现, 没有随运行时间的加长而导致显著的探测误差。需要注意的是, 图 9 中单次观测脉冲到达时间的误差仍然较大, 这是因为 FXPT 的有效面积较小, 接收 X 射线光子数较少导致统计表征误差大, 但 FXPT 对 X 射线光子的到达时间记录准确,

因此能够获得高信噪比的脉冲轮廓。

3 结 论

1）利用XPNAV-1卫星核心载荷FXPT对Crab脉冲星2016—2019年观测数据，采用统计学方法，从光子计数率统计分布曲线和脉冲轮廓形状2个维度分析了卫星的在轨稳定性。分析表明，各月计数率分布均符合高斯分布，中心值稳定在14.87 counts/s。在剔除计数率异常月份后，其余月份的决定系数均高于97%；拟合高斯函数的半高全宽稳定在7.26 counts/s，标准偏差约为0.42 counts/s。光子计数率存在一定涨落，主要受空间背景噪声影响。

2）从脉冲轮廓相似度和非脉冲辐射强度2个方面分析脉冲轮廓特征参数变化。结果表明，各月脉冲轮廓与标准轮廓的相似度均高于99.7%，且非脉冲辐射强度稳定在0.66附近，说明光子计数率异常时间段内的光子很可能并非来自脉冲星。

3）利用FXPT观测数据建立的计时模型能够折叠出高信噪比脉冲轮廓，与根据Jodrell Bank天文台发布的天文星历折叠得到的脉冲轮廓相似度达到99.99%，表明FXPT记录的X射线光子到达时间具有较高精度。

综合以上分析结果表明，XPNAV-1卫星FXPT在2016—2019年在轨工作性能稳定，能够持续获取可靠的X射线脉冲星观测数据，可满足后续科学分析和导航体制验证的需求。

参考文献（References）

[1] CHAMPION D J, HOBBS G B, MANCHESTER R N, et al. Measuring the mass of solar system planets using pulsar timing[J]. The Astrophysical Journal, 2010, 720(2): L201-L205.

[2] HOBBS G, COLES W, MANCHESTER R N, et al. Development of a pulsar-based time-scale[J]. Monthly Notices of the Royal Astronomical Society, 2012, 427(4): 2780-2787.

[3] PETIT G, TAVELLA P. Pulsars and time scales[J]. Astronomy and Astrophysics, 1996, 308(1): 290-298.

[4] 童明雷, 杨廷高, 赵成仕, 等. 脉冲星计时模型参数的测量精度分析与估计[J]. 中国科学: 物理学 力学 天文学, 2017, 47(9): 103-112.
TONG M L, YANG T G, ZHAO C S, et al. Analyses and estimates on measurement precisions of model parameters in pulsar timing[J]. Scientia Sinica (Physica, Mechanica & Astronomica), 2017, 47(9): 103-112(in Chinese).

[5] HOBBS G, ARCHIBALD A, ARZOUMANIAN Z, et al. The International Pulsar Timing Array project: using pulsars as a gravitational wave detector[J]. Classical and Quantum Gravity, 2010, 27(8): 084013.

[6] SHEIKH S I. The use of variable celestial X-ray sources for spacecraft navigation[D]. College Park: University of Maryland, 2005: 15-70.

[7] BECKER W, BERNHARDT M G, JESSNER A. Autonomous spacecraft navigation with pulsars[J]. Acta Futura, 2013(7): 11-28.

[8] 郑伟, 王奕迪, 汤国建, 等. X射线脉冲星导航理论与应用[M]. 北京: 科学出版社, 2015: 1-28.
ZHENG W, WANG Y D, TANG G J, et al. X-ray pulsar-based navigation: theory and applications[M]. Beijing: Science Press, 2015: 1-28(in Chinese).

[9] 周庆勇. 脉冲星计时数据的处理理论与方法研究[D]. 郑州: 战略支援部队信息工程大学, 2020: 120-146.
ZHOU Q Y. Research on theory and methods of pulsar timing data[D]. Zhengzhou: Information Support Force Engineering University, 2020: 120-146(in Chinese).

[10] MITCHELL J W, WINTERNITZ L B, HASSOUNEH M A, et al. SEXTANT X-ray pulsar navigation demonstration: initial on-orbit results[C]//Proceedings of the 41st Annual Guidance and Control Conference of American Astronautical Society. Piscataway: IEEE Press, 2018: 1-12.

[11] NASA. 2015 NASA technology roadmaps TA5 communications navigation and orbital debris tracking and characterization systems[EB/OL]. (2015-08-10)[2021-12-26]. https://www.lpi.usra.edu/sbag/goals/capability_inputs/2015_Tech_5_communication_navigation.pdf.

[12] 黄良伟, 帅平, 张新源, 等. 脉冲星导航试验卫星时间数据分析与脉冲轮廓恢复[J]. 中国空间科学技术, 2017, 37(3): 1-10.
HUANG L W, SHUAI P, ZHANG X Y, et al. XPNAV-1 satellite timing data analysis and pulseprofile recovery[J]. Chinese Space Science and Technology, 2017, 37(3): 1-10(in Chinese).

[13] HUANG L W, SHUAI P, ZHANG X Y, et al. Pulsar-based navigation results: data processing of the X-ray pulsar navigation-I telescope[J]. Journal of Astronomical Telescopes, Instruments, and Systems, 2019, 5(1): 018003.

[14] 李连升, 梅志武, 吕政欣, 等. 掠入射聚焦型X射线脉冲星望远镜及在轨数据分析[J]. 兵器装备工程学报, 2017, 38(12): 175-179.
LI L S, MEI Z W, LV Z X, et al. Grazing incidence focusing X-ray pulsar telescope and analysis of in-orbit observation data[J]. Journal of Ordnance Equipment Engineering, 2017, 38(12): 175-179(in Chinese).

[15] 帅平, 张新源, 黄良伟, 等. 脉冲星导航试验卫星科学观测数据分析[J]. 空间控制技术与应用, 2017, 43(2): 1-6.
SHUAI P, ZHANG X Y, HUANG L W, et al. X-ray pulsar navigation test satellite science data analysis[J]. Aerospace Control and Application, 2017, 43(2): 1-6(in Chinese).

[16] 周庆勇, 张健康, 贾小林, 等. 脉冲星导航试验卫星空间计时数据的初步分析[J]. 全球定位系统, 2023, 48(4): 63-68.
ZHOU Q Y, ZHANG J K, JIA X L, et al. Space observation data preliminary analysis of XPNAV-1 satellite[J]. GNSS World of China, 2023, 48(4): 63-68(in Chinese).

[17] 周庆勇, 闫林丽, 李连升, 等. XPNAV-1卫星聚焦型X射线脉冲星望远镜在轨稳定性分析[J]. 航空学报, 2023, 44(3): 138-148.
ZHOU Q Y, YAN L L, LI L S, et al. On-orbit stability analysis of FXPT on XPNAV-1[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(3): 138-148(in Chinese).

[18] ZHANG X Y, SHUAI P, HUANG L W, et al. Mission overview and initial observation results of the X-ray pulsar navigation-I satellite[J]. International Journal of Aerospace Engineering, 2017, 2017: 8561830.

[19] LYNE A, GRAHAM-SMITH F. Pulsar astronomy[M]. Cambridge: Cambridge University Press, 2012: 61-66.

[20] JAIN C, PAUL B. Pulse profile stability of the Crab pulsar[J]. Research in Astronomy and Astrophysics, 2011, 11(10): 1134-1142.

[21] 孙海峰, 包为民, 方海燕, 等. X 射线脉冲轮廓稳定性对导航精度的影响[J]. 物理学报, 2014, 63(6): 069701.
SUN H F, BAO W M, FANG H Y, et al. Effect of stability of X-ray pulsar profiles on range measurement accuracy in X-ray pulsar navigation[J]. Acta Physica Sinica, 2014, 63(6): 069701(in Chinese).

[22] SUN H F, BAO W M, FANG H Y, et al. Effect of X-ray energy band on the X-ray pulsar based navigation[J]. Aerospace Science and Technology, 2016, 58: 150-155.

[23] HOBBS G B, EDWARDS R T, MANCHESTER R N. TEMPO2, a new pulsar-timing package–I. an overview[J]. Monthly Notices of the Royal Astronomical Society, 2006, 369(2): 655-672.

[24] 闫林丽, 葛明玉, 庹攸隶, 等. 5 颗导航用 X 射线脉冲星计时分析[J]. 航空学报, 2023, 44(3): 170-185.
YAN L L, GE M Y, TUO Y L, et al. Timing analysis for five navigation X-ray pulsars[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(3): 170-185(in Chinese).

Stability analysis of core payload operational performance on the XPNAV-1 satellite

YAN Linli[1], ZHANG Jiankang[1], ZHOU Qingyong[2, 3, *], LEI Yaohu[4], FAN Shaojuan[1]

(1. School of Mathematics and Physics, Anhui Jianzhu University, Hefei 230601, China;
2. Xi'an Institute of Surveying and Mapping, Xi'an 710054, China;
3. State Key Laboratory of Geographic Information Engineering, Xi'an 710054, China;
4. Key Laboratory of Optoelectronic Devices and Systems of Ministry of Education and Guangdong Province, College of Physics and Optoelectronic Engineering, Shenzhen University, Shenzhen 518060, China)

Abstract: A carefully engineered focusing X-ray pulsar telescope (FXPT) is the main payload of China's first X-ray pulsar navigation test satellite XPNAV-1. Factors such as launch vibrations, space radiation, and device lifespan can cause slight variations in the performance parameters of the FXPT. Therefore, long-term monitoring of its performance is necessary to provide more reliable observational data for astronomical timing analysis and navigation calculations. To monitor the performance changes of the FXPT on the XPNAV-1 satellite, a comprehensive analysis method combining the statistical distribution of photon count rates and the characteristics of pulse profile shapes was proposed. Using observation data of the Crab pulsar from 2016 to 2019 obtained by the XPNAV-1 satellite, the in-orbit operational stability of the FXPT was analyzed. The results show that the statistical distribution of the Crab pulsar photon count rates obtained by the FXPT approximately follows a Gaussian distribution. Even with a few periods of anomalous count rates, the pulse profile shapes still exhibit high similarity over different periods. High signal-to-noise ratio pulse profiles that are 99.99% comparable to pulse profiles folded using the ephemeris supplied by Jodrell Bank Observatory can be produced by the timing model generated from FXPT data. These results confirm the in-orbit operational stability of XPNAV-1's core payload, the FXPT.

Keywords: XPNAV-1 satellite; X-ray pulsar navigation; timing analysis; focusing X-ray pulsar telescope; stability analysis

Received: 2024-04-29; **Accepted**: 2024-06-18; **Published Online**: 2024-06-20 17:07

URL: link.cnki.net/urlid/11.2625.V.20240620.1538.002

Foundation items: National Natural Science Foundation of China (11903001,42004004); Youth Teacher Training Program of Anhui Provincial Department of Education (YQZD2023053); Doctor Foundation of Anhui Jianzhu University (2019QDZ14,2019QDZ31); Anhui Province University Excellent Youth Scientific Research Project (2024AH030011); Innovation and Entrepreneurship Training Program for College Students of Anhui Jianzhu University (202310878057)

* **Corresponding author**. E-mail: zjlzqy1986@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0249

垂直起降机场运行任务规划模型与容量估算方法

魏志强*，肖鑫隆
(中国民航大学 空中交通管理学院，天津 300300)

摘 要：针对电动垂直起降飞行器(eVTOL)在垂直起降机场的运行任务规划与容量估算问题，通过对垂直起降机场布局结构的梳理，绘制出运行任务网络图；考虑eVTOL在垂直起降机场内的起降、滑行与保障任务，基于多商品流模型，引入网络内各节点进入时刻、离开时刻，构建出运行任务规划模型，以时间区间内最优任务规划为优化目标，将节点占用时间作为约束；以典型起降机场为例使用Python+Gurobi进行编程求解，对机场容量及影响因素进行计算分析。结果表明：所构建出的任务规划模型具有可行性。在影响垂直起降机场容量因素方面，当起降坪利用率低时，增加1个登机口能增加垂直起降机场3架次容量，并能为垂直起降机场提供100 s容量韧性；而当登机口利用率低时，增加1个起降坪，能为垂直起降机场提供7架次容量；通过缩短30 s起降时长能为垂直起降机场提供平均3架次容量。研究结果有助于未来估算垂直起降机场容量，并为eVTOL在垂直起降机场的运行调度提供一种建模思路。

关 键 词：城市空中交通；垂直起降机场；多商品流模型；任务规划；容量估算
中图分类号：V355
文献标志码：A **文章编号**：1001-5965(2026)06-1880-10

随着电动垂直起降飞行器(electric vertical take-off and landing, eVTOL)相关技术的不断成熟，亟需开展与其配套的垂直起降机场规划建设研究工作。对垂直起降机场内的网络模型构建、运行任务规划与容量估算是科学合理规划垂直起降机场的基础[1-4]。

目前由于eVTOL为新兴产业，如何将其运用于城市空中交通(urban air mobility, UAM)仍处在研究阶段，在运行方面，Su等[5]提出一种垂直起降机场的奖励适应机制，通过构建多目标模型，在考虑垂直起降机场最大化效益的同时，最小化eVTOL之间的拥堵情况。Suzuki等[6]在城市空中交通运行方面，设计一种航路规划工具，通过专家评审方式对此工具的合理性进行评估。Xiang等[7]着重于eVTOL运行过程中的降落部分，对垂直起降机场上的识别标识、eVTOL精确探测器着陆算法方面及eVTOL多阶段着陆策略方面展开研究。Kleinbekman等[8]主要研究eVTOL降落排序，综合考虑eVTOL剩余电量及垂直起降机场容量，从而构建混合整数线性规划模型，求解出最优降落方案。Conrad等[9]针对eVTOL运行中的数据处理问题，提出一种智能数据处理方法，用于处理eVTOL在运行中存在的大量数据，并以此为基础，对eVTOL运行进行模拟仿真。Pradeep等[10]以最短着陆完成时长为目标函数，结合混合整数规划及时间优先级构建数学模型进行算例仿真，从而得到其最优降落规划。Preis等[11]在已有的垂直起降机场运行数学模型基础上，对其进行智能体模拟仿真，通过专家评审方式，分析影响模型的主要影响参数。由上述文献可知，eVTOL运行的相关研究，仍可进一步探索，可以针对eVTOL于垂直起降机场的运行设计运行规划模型，用于求

收稿日期：2024-04-23；**录用日期**：2024-08-30；**网络出版时间**：2024-09-11 09:48
网络出版地址：link.cnki.net/urlid/11.2625.V.20240910.1642.004
基金项目：国家自然科学基金(U2133210)；天津市科技计划项目(23JCZDJC00580)；激光雷达与器件技术重点实验室开放课题(LLD2023-008)
* **通信作者**. E-mail：weizhiqia@sina.com

引用格式：*魏志强，肖鑫隆. 垂直起降机场运行任务规划模型与容量估算方法[J]. 北京航空航天大学学报，2026，52(6)：1880-1889.*
WEI Z Q，XIAO X L. Vertiport operational task planning model and capacity estimation method[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1880-1889 (in Chinese).

解其最优规划。

eVTOL 运行需要有基础设施作为保障，如何设计机场，从而使其容量需求满足未来可能产生的 eVTOL 流量需求。在这方面，Brunelli 等[12] 对近些年来垂直起降机场的布局设计、选址和容量做总结综述，从而分析总结出机场选址、布局的影响因素。Preis [13] 则侧重于垂直起降机场内的布局及 eVTOL 的运行方式，对近些年来的研究进行归纳总结，并对未来垂直起降机场研究方向提供参考意见。Preis[11] 考虑目前有关 eVTOL 设计，从而提出一种基于 eVTOL 构型的容量评估方法。Ahn 等[14] 研究垂直起降机场中功能区的设计标准，并以此设计为标准对垂直起降机场进行容量影响因素分析。Schweiger 等[15] 则从垂直起降机场空侧角度，考虑空中风对垂直起降机场空中交通流的影响，通过某机场运行中空中风数据，对其进行分析。Park 等[16] 分析航空器于垂直起降机场内的平均等待时间，用于评估垂直起降机场的容量问题。上述研究中，主要是针对 eVTOL 运行中的某一因素作为基础并展开研究，而如何从多因素方面进行考虑，从而考虑多个因素下的机场容量，仍可进一步研究。

eVTOL 于垂直起降机场的任务规划问题，属于典型的多商品流问题。在多商品流问题方面，Niu 等[17] 以经典多商品流模型为基础，针对其求解方法，利用量子退火算法进行求解，与原求解方法进行对比，最后得出此算法适用于大规模运输问题。Khanal 等[18] 从多商品流网络进行考虑，分别构建静态多商品流问题及动态带储藏点的多商品流问题模型，最后以多项式和伪多项式时间算法对以上 2 个问题分别进行求解。Nishi 等[19] 提出结合开始结束点容量限制的多商品流模型，通过与传统多商品模型进行比较，从而证明此模型的合理性。Narayan 等[20] 在多商品流问题基础上对原问题进行改进，提出带随机路径权重的多商品流网络问题。Liu 等[21] 在民用机场的运行规划方面，将停机位分配和滑行道规划结合起来，对多商品流问题进行改进，构建离散的时空网络模型，并根据模型研究机场最优滑行与停机位分配。Vascik 等[22] 也对多商品流模型进行改进，构建出离散时间的整数规划模型，并根据模型对影响因素进行分析，分析出影响机场容量的最主要因素，得到机场的最优起降坪与登机口数量搭配。在上述研究中，主要针对的是模型求解方面，而在具体模型改进方面，对如何改进网络，从而使之成为对应研究场景的运行规划模型，仍有进一步研究的空间。

因此，本文系统研究和构建了 eVTOL 在垂直起降机场内的运行任务描述与规划模型。本文通过对多商品流问题模型的改进，将 eVTOL 于垂直起降机场内执行任务抽象为任务节点，绘制任务网络；同时引入开始任务时刻、结束任务时刻变量，从而构建出 eVTOL 于垂直起降机场内的任务调度模型。本文提出基于任务规划模型的垂直起降机场容量估算方法，定量分析和识别出起降机场容量的关键影响因素。

1　垂直起降机场运行任务网络

1.1　运行任务分析

垂直起降机场内，航空器的起飞、降落、滑行、周转和等待是航空器于机场内运行所执行的基本任务，具体运行概念如图 1 所示。

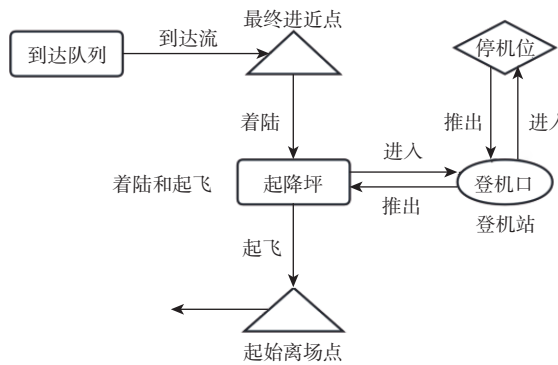


图 1　运行概念图

Fig. 1　Operation concept diagram

首先，eVTOL 到达最终进近点；接着，着陆至起降坪；然后，进入登机口进行周转，根据其此次任务类型，若继续起飞则周转完后推出至起降坪；最后是起飞。若任务是进入停机位修整，最后则是进入停机位进行等待。

垂直起降机场由 3 部分构成：起降坪（Pads）、登机口（Gates）和停机位（Stands），滑行道用于连接此 3 部分。起降坪用于 eVTOL 执行起降任务；登机口用于 eVTOL 周转；停机位用于存放 eVTOL，eVTOL 在机场内运行时，可于停机位中等待，从而缓解运行冲突。

垂直起降机场的任务规划，需要绘制出对应的任务网络，并在任务网络的基础上，构建对应的任务规划模型，任务规划模型需要确定其目标函数、约束条件和决策变量。并基于此模型得出其最优目标函数下多个 eVTOL 于机场内起飞、滑行、周转、等待和降落的最优任务安排。

1.2　运行任务过程网络的绘制

为绘制运行任务网络，将 eVTOL 于垂直起降机场所执行的任务抽象为节点，不同任务之间的连接抽象为弧，如图 2 所示。

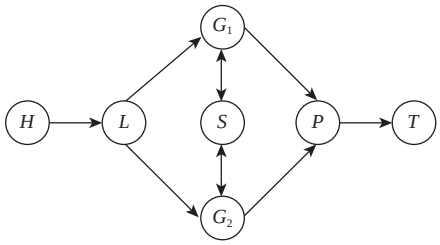


图2 运行任务网络

Fig. 2 Operate mission networks

eVTOL在起降机场常见的运行程序有:等待、起飞、周转、滑行和降落,故设置节点H为盘旋等待任务,L为降落任务,G_1、G_2为滑行周转任务,S为滑行及等待任务,P为起飞前的滑行任务,T为起飞任务。

定义网络图$G(V, E)$,其中,V表示任务点集合,E表示各任务点之间的弧集。假设eVTOL于此垂直起降机场内任务为起降任务,eVTOL将执行从任务H至任务T的网络流程。假设eVTOL集合为K,$K=\{k_1,k_2,\cdots,k_m\}$;分析容量问题中需要考虑登机口数量,因此,再假设登机口集合为G_a,$G_a=\{G_{a1},G_{a2},\cdots,G_{an}\}$。

2 运行任务调度模型构建

为实现多个eVTOL于垂直起降机场内运行,本文以多商品流模型为基础,引入各个节点开始执行任务时刻和结束任务时刻,综合考虑eVTOL于此任务网络中可能遇到的冲突,从而对模型进行改进,使之成为解决eVTOL于垂直起降机场内的运行任务调度模型。在模型中针对的多商品流对象为eVTOL。

2.1 决策变量

为解决运行任务调度问题,本文在此引入$x_{i,j,k}$、$a_{i,k}$、$d_{i,k}$、o_k和l_k作为决策变量。

$x_{i,j,k}$为0-1决策变量,含义为k执行完任务i,且执行的下一个任务为j时为1,否则为0。

$a_{i,k}$和$d_{i,k}$分别为用于记录k于任务i的开始时刻和结束时刻的连续变量,其满足式(1)~式(6)。

k的开始等待时间如下:

$$a_{H,k}=t_{\text{arrive}}k \qquad k\in K \tag{1}$$

式中:t_{arrive}为预降时间间隔。

eVTOL之间各任务点的执行任务满足先来先服务,如下:

$$a_{i,k_1}\leqslant a_{i,k_2} \qquad i\in V,k_1\in K,k_2\in K,k_1<k_2 \tag{2}$$

各任务点满足结束任务时刻大于其开始任务时刻,如下:

$$0\leqslant a_{i,k}\leqslant d_{i,k} \qquad \forall i\in V,\forall k\in K \tag{3}$$

式中:$d_{i,k}$为k离开i点时刻。

各eVTOL于各节点的开始任务时刻及结束任务时刻,应保证先后关系,后开始及结束的任务,其时刻应大于先开始及结束的任务,约束如下:

$$a_{i,k}-M(1-x_{i,j,k})\leqslant a_{j,k} \qquad \forall(i,j)\in E,k\in K \tag{4}$$

$$d_{i,k}-M(1-x_{i,j,k})\leqslant d_{j,k} \qquad \forall(i,j)\in E,k\in K \tag{5}$$

式中:M为取值很大的常数。

eVTOL上一任务结束时刻,等于其开始下一任务时刻,约束如下:

$$x_{i,j,k}(d_{i,k}-a_{j,k})=0 \qquad \forall(i,j)\in E,k\in K \tag{6}$$

定义o_k、l_k分别为用于判断eVTOL起飞时刻和着陆时刻是否在规定时段内的0-1变量,在时段内o_k和l_k分别为1,否则为0,其满足:

$$d_{L,k}\geqslant t-Mo_k \qquad k\in K \tag{7}$$

$$d_{L,k}\leqslant t+M(1-o_k) \qquad k\in K \tag{8}$$

$$d_{T,k}\geqslant t-Ml_k \qquad k\in K \tag{9}$$

$$d_{T,k}\leqslant t+M(1-o_k) \qquad k\in K \tag{10}$$

式中:t为规定的时间段,表示若eVTOL在时段t内结束起飞及降落任务,其对应的决策变量o_k和l_k为1,否则为0。

本文模型需要得到在规定时间段内,在给定起降权重下,垂直起降机场所能达到的最优起降规划,分别假设eVTOL起飞价值系数$c_{\text{take-off}}$和航空器降落价值系数c_{land},从而可得目标函数:

$$O_{\text{bj}}=\max\sum_{k\in K}(c_{\text{take-off}}o_k+c_{\text{land}}l_k) \tag{11}$$

式(11)表示eVTOL在此起飞价值系数和降落价值系数下的最优起降架次,其中,起飞价值系数和降落架次系数,主要用于区分在起降运行规划,其可区分侧重执行起飞、侧重执行降落或两者都有侧重。

2.2 任务网络约束限制

网络限制分为网络流限制和任务网络执行时间限制,网络流正常流动,需要满足基本网络约束,且由于构建的网络是任务网络,于任务节点上需要满足一定的执行任务时间约束。

2.2.1 网络流的平衡

为保证eVTOL于此任务网络内执行任务,k必须执行完等待任务H后执行降落任务L,并必须执行完起飞前的滑行任务P后执行起飞任务,应满足如下约束:

$$x_{H,L,k}=1 \qquad k\in K \tag{12}$$

$$x_{P,T,k}=1 \qquad k\in K \tag{13}$$

其中间节点 i 必须满足流入等于流出，eVTOL 必须连续执行 j_1、i、j_2 任务，从而保证其连续性，应满足如下约束：

$$\sum_{j_1:(j_1,i)\in E} x_{j_1,i,k}=\sum_{j_2:(i,j_2)\in E} x_{i,j_2,k} \qquad i\in V/\{H\cup T\},k\in K \tag{14}$$

2.2.2　网络任务安排

假设 eVTOL 降落时长、滑行时长、周转时长、和起飞时长分别为 t_{land}、t_{taxi}、$t_{\text{turnaround}}$ 和 $t_{\text{take-off}}$。eVTOL 于任务点 H 执行等待任务，于任务点 L 执行着陆任务，于任务点 G 执行周转与从起降坪至登机口的滑行任务，于任务点 S 执行停机位等待与从停机位至登机口的滑行任务，于任务点 P 执行起飞前从登机口至起降坪的滑行任务，最后于任务点 T 执行起飞任务，可得如下任务时刻关系约束：

$$d_{H,k}-a_{H,k}\geqslant 0 \qquad k\in K \tag{15}$$

$$d_{L,k}-a_{L,k}=t_{\text{land}} \qquad k\in K \tag{16}$$

$$d_{i,k}-a_{i,k}\geqslant t_{\text{turnaround}}+t_{\text{taxi},G} \qquad i\in G,k\in K \tag{17}$$

$$d_{S,k}-a_{S,k}\geqslant t_{\text{taxi},S} \qquad k\in K \tag{18}$$

$$d_{P,k}-a_{P,k}=t_{\text{taxi},G} \qquad k\in K \tag{19}$$

$$d_{T,k}-a_{T,k}=t_{\text{take-off}} \qquad k\in K \tag{20}$$

2.3　运行冲突约束限制

eVTOL 于垂直起降机场进行任务规划，最关键的是避免冲突。经过本文模型测试发现冲突主要集中在任务点冲突、滑行冲突和起降冲突上，针对这些冲突，分别构建约束模型。

2.3.1　任务点冲突

引入 0-1 变量 u_{i,k_1,k_2}，用于判断 2 架 eVTOL 执行任务时刻前后，当 k_1 开始 i 任务时刻晚于 k_2 时 u_{i,k_1,k_2} 为 1，否则为 0。当 k_1 开始任务 i 时刻晚于 k_2 时，则 k_1 开始执行任务时刻应大于 k_2 结束任务时刻（见式(21)）；同理，当 k_1 开始任务 i 时刻早于 k_2 时，则 k_2 开始执行任务时刻应大于 k_1 结束任务时刻（见式(22)），保证不同 eVTOL 之间所执行任务时刻相互错开，从而避免冲突。

$$a_{i,k_1}-d_{i,k_2}\geqslant -M(1-u_{i,k_1,k_2}) \qquad i\in V,i\notin S,k_1\in K,k_2\in K,k_1\neq k_2 \tag{21}$$

$$a_{i,k_2}-d_{i,k_1}\geqslant -Mu_{i,k_1,k_2} \qquad i\in V,i\notin S,k_1\in K,k_2\in K,k_1\neq k_2 \tag{22}$$

2.3.2　登机口对头滑行冲突

eVTOL 在起降坪与登机口之间滑入滑出时会存在对头冲突，还会出现功能区之间的任务重叠，因此，引入一个 0-1 变量 q_{i,k_1,k_2}，用于判断不同 eVTOL 滑入滑出时刻的先后关系，当 k_1 于任务 i 执行滑入任务时刻晚于 k_2 于任务 P 执行滑出任务时刻时，q_{i,k_1,k_2} 为 1，否则为 0。假设 y 用于设置时间间隔，从而避免出现功能区冲突的情况。当 2 架 eVTOL 从相同登机口执行滑入与滑出任务时，根据其滑入滑出时刻先后关系，需要满足 k_1 开始执行滑入任务时刻大于 k_2 结束执行滑出任务时刻；或 k_1 结束滑入任务时刻小于 k_2 开始滑出任务时刻。可得如下约束：

$$a_{i,k_1}-d_{P,k_2}\geqslant -M(2-x_{L,i,k_1}-x_{i,P,k_2})-M(1-q_{i,k_1,k_2}) \qquad i\in G,k_1\in K,k_2\in K,k_1\neq k_2 \tag{23}$$

$$a_{P,k_2}-(a_{i,k_1}+t_{\text{taxi}})\geqslant -M(2-x_{L,i,k_1}-x_{i,P,k_2})-Mq_{i,k_1,k_2} \qquad i\in G,k_1\in K,k_2\in K,k_1\neq k_2 \tag{24}$$

2 架 eVTOL 结束降落和结束起飞任务时刻，满足时间间隔 y，可得如下约束：

$$|d_{L,k_1}-d_{P,k_2}|\geqslant y-M(1-x_{i,P,k_2}) \qquad i\in G,k_1\in K,k_2\in K,k_1\neq k_2 \tag{25}$$

2.3.3　起降坪起降冲突

由于 eVTOL 于起降坪起降时，需要满足独立性，即在起飞时不得有其他 eVTOL 降落，反之亦然。因此，引入 0-1 变量 h_{k_1,k_2}，用于判断不同 eVTOL 起降的先后顺序，若 k_1 开始降落时刻晚于 k_2 结束起飞时刻时，h_{k_1,k_2} 则为 1，否则为 0。当 k_1 开始降落时刻晚于 k_2 结束起飞时刻时，k_1 开始执行降落任务时刻应大于 k_2 结束起飞任务时刻，反之也是如此。从而可得如下约束：

$$a_{L,k_1}-d_{T,k_2}\geqslant -M(1-h_{k_1,k_2}) \qquad k_1\in K,k_2\in K,k_1\neq k_2 \tag{26}$$

$$a_{T,k_1}-d_{L,k_2}\geqslant -Mh_{i,k_1,k_2} \qquad k_1\in K,k_2\in K,k_1\neq k_2 \tag{27}$$

3　模型验证与分析

3.1　基于典型场景的模型验证

假设垂直起降机场的布局类型为线性布局，且满足单起降坪、双登机口布局，投入 4 架 eVTOL。参数假设如表 1 所示。

表 1　测试参数值

Table 1　Test parameter values

$c_{\text{take-off}}$	c_{land}	t/s	t_{arrive}/s	t_{land}/s	$t_{\text{take-off}}$/s	$t_{\text{turnaround}}$/s	t_{taxi}/s
1/−1	5	200	15	15	15	40	5

通过 Python+Gurobi 进行求解，表 2 为求解出的 eVTOL 任务规划。

表 2 eVTOL 任务规划

Table 2 eVTOL mission scheduling

任务	$a_{i,k}$/s				$d_{i,k}$/s			
	k_1	k_2	k_3	k_4	k_1	k_2	k_3	k_4
H	0	15	30	45	0	15	70	115
L	0	15	70	115	15	30	85	130
G_1	15		85		50		140	
G_2		30		130		95		165
P	50	95	140	165	55	100	145	170
T	55	100	145	170	70	115	160	185

从连续性角度，所求数据满足其连续性要求，eVTOL 结束某任务时刻与开始执行下一任务时刻相等。

从避免冲突角度，冲突主要有任务点冲突、登机口对头滑行冲突和起降坪起降冲突。由表 2 可知，eVTOL 各自在功能区执行任务时刻相互错开，未发生任务点冲突。此外，在滑行冲突方面，不同 eVTOL 执行任务 G_1、G_2 中的滑入任务与在任务点 P 执行的滑出任务时刻相互错开，未发生滑行冲突。在起降冲突方面，不同 eVTOL 于任务点 L 与任务点 T 执行任务时刻也相互错开，因此，未发生起降冲突。

3.2 垂直起降机场容量估算方法

本文中垂直起降机场的容量指根据垂直起降机场受垂起降坪、登机口、停机位、eVTOL 性能和进场流量影响垂直起降机场所能提供的最大起降架次。

容量包线可作为容量可视化方法，以登机口数量 4 为例，分别假设 $c_{\text{take-off}}$ 为 1 和−1，c_{land} 为 5，通过调整输入 eVTOL 流数量，从而得到 eVTO 容量包线如图 3 所示。

当 $c_{\text{take-off}}$ 为 1 时，而 c_{land} 为 5 时，则模型需要满足起降架次，通过不断加大网络流，从而得到容量包络线的上限，即图中的最大吞吐量。而当 $c_{\text{take-off}}$ 为−1，而 c_{land} 为 5 时，主要侧重于降落，此时不断加大网络流，从而得到容量包线的下限，即图中的最大着陆架次。当交通流加大到一定程度，上下限交汇于一个点，从而绘制出垂直起降机场容量包线如图 3 所示，此点的横纵坐标之和为垂直起降机场的容量。

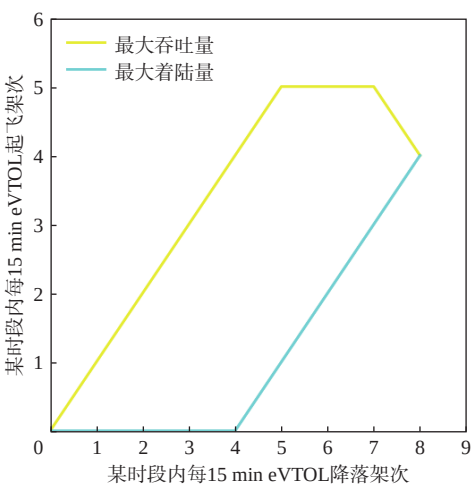


图 3 容量包线绘制方法

Fig. 3 Capacity envelope drawing method

3.3 垂直起降机场容量估算合理性验证

为验证此容量估算方法的可靠性，在此引入文献 [23] 机场的容量估算方法，用于估算垂直起降机场容量。

首先，根据进场容量架次和时间计算机场地面进场容量如下：

$$C_{\text{Arr}} = N_{\text{Arr}}/(T_{\text{MaxArr}} - T_{\text{MinArr}}) \tag{28}$$

式中：C_{Arr} 为机场到达容量；N_{Arr} 为进场架次；T_{MaxArr} 为 eVTOL 最晚到达时刻；T_{MinArr} 为 eVTOL 最早到达时刻。

其次，根据进场容量架次和时间计算机场地面进场容量如下：

$$C_{\text{Dep}} = N_{\text{Dep}}/(T_{\text{MaxDep}} - T_{\text{MinDep}}) \tag{29}$$

式中：C_{Dep} 为机场离场容量；N_{Dep} 为进场架次；T_{MaxDep} 为 eVTOL 最晚离场时刻；T_{MaxDep} 为 eVTOL 最早离场时刻。

最后，可得垂直起降机场地面总容量如下：

$$C_{\text{Air}} = C_{\text{Arr}} + C_{\text{Dep}} \tag{30}$$

在单起降坪四登机口场景下，通过任务规划计算出来结果可知参数如表 3 所示

表 3 垂直起降机场运行参数

Table 3 Vertiport operation values

15 min C_{Arr}/架	15 min C_{Dep}/架	T_{MaxArr}/s	T_{MinArr}/s	T_{MaxDep}/s	T_{MinDep}/s
8	4	900	60	630	1 290

由表 3 可知，在单起降坪登机口数量为 4 情况下的容量从而可得起飞容量为 15 min 8 架，而降落架次为 15 min 4 架，于是总容量为 15 min 12 架。

而在本文容量估算模型中所得的单起降坪登机口数量为 4 情况下的绘制出的容量包络线为图 3 所示，其容量为 15 min 12 架，与式(30)容量评估模型仅差 1 架次起降 15 min 起降容量，因此，本文模型具有可靠性。

与式(30)容量评估方法不同，本文容量估算方法通过结合运行规划模型从而得出结果，能够通过修改任务规划模型参数分析容量影响因素，而文献 [23] 容量评估模型通过已知的运行数据计算容

量，缺少运行规划部分。

3.4　垂直起降机场容量因素分析

假定 eVTOL 运行模式为 eVTOL 执行完降落任务后，在垂直起降机场内执行周转、等待、滑行任务后，然后执行新的起飞任务。基于以上假设，对影响因素进行计算分析。

3.4.1　登机口的影响

对相关参数的假设如表 4 所示。

表 4　eVTOL 运行参数

Table 4　eVTOL operation values

$c_{\text{take-off}}$	c_{land}	t/s	t_{arrive}/s	t_{land}/s	$t_{\text{take-off}}$/s	$t_{\text{turnaround}}$/s	t_{taxi}/s
1/−1	5	900	60	60	60	300	15

分别绘制登机口数量 m 为 1~6 时，对应的垂直起降机场容量包络线如图 4 和表 5 所示。

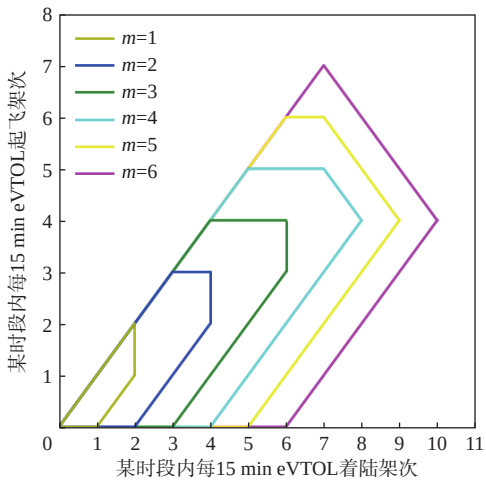


图 4　容量包线

Fig. 4　Capacity envelope

表 5　功能区利用率

Table 5　Functional area utilization rate

登机口数量	起降坪利用率/%	登机口平均利用率/%	登机口等待率/%
1	27	70	0
2	47	71	15
3	67	60	17
4	87	48	8
5	97	47	23
6	100	53	53

图 4 为垂直起降机场容量包线，总体上看，随着登机口数量增加，垂直起降机场容量包线面积逐步增加。从其增长幅度来看，从垂直起降机场机场容量上看，当登机口数量小于等于 4 时，垂直起降机场机场容量有着较为显著的增加平均每增加 1 个登机口垂直起降机场容量增加 3 架次；而当登机口数量超过 4 时，增长趋势放缓，从原先增加 1 个登机口增加 3 架次，减至增加 1 个登机口增加 1 架次。其原因可能是随着登机口数量增加，登机口数量不再是影响垂直起降机场容量主要因素导致。

表 5 为其功能区利用率，总体上，起降坪利用率不断增加，登机口的平均利用率总体呈现下降趋势，而登机口等待率呈现上升趋势。在起降坪利用率变化方面，当登机口数量小于 4 时，随着登机口数量增加，起降坪利用率有着显著上升，当超过 4 时，变化幅度不大；在登机口平均利用率变化方面，当登机口数量在 2～4 时，登机口平均利用率有显著下降，而当登机口数量超过 4 时，同样也变化不大；在登机口等待率变化方面，登机口等待率在登机口数量由 3～6 时有着较为明显的上下浮动。以上原因，可能是随着登机口数量增加，垂直起降机场能容纳的 eVTOL 更多，从而能容纳更多 eVTOL 起降，因此，起降坪利用率逐步上升；表中的登机口数量 4 可作为一个分界线，当登机口数量超过 4 时，其登机口等待率上升，从而使登机口的平均利用率涨幅较慢，此时影响垂直起降机场容量的主因是起降坪数量。

3.4.2　起降坪的影响

如表 4 所示进行假设，考虑双起降坪，根据起降坪功能的不同绘制出如图 5 和表 6 所示。

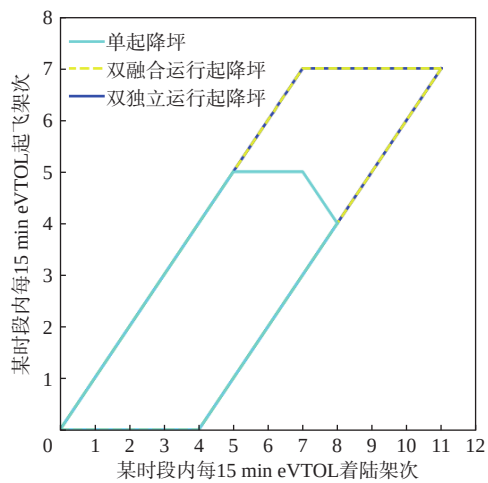


图 5　起降坪影响容量包线

Fig. 5　Capacity envelope of influence on pads

表 6　不同起降坪数量下的功能区利用率

Table 6　Functional area utilization rate of different pads number

起降坪数量	起降坪平均利用率/%	登机口平均利用率/%	登机口等待率/%
1	57	66	17
2(融合运行)	70	59	0
2(独立运行)	70	59	0

图 5 为不同起降坪数量和不同起降坪功能下的容量包线图。由图可知，在起降坪数量方面，通

过增加起降坪数量能够增加垂直起降机场容量包线面积，增加一个起降坪能为垂直起降机场增加7架次容量；而在不同起降坪功能方面，双起降坪可分为单降落坪、单起飞坪或双起降混合坪，起降坪功能不同并不会影响垂直起降机场容量包线变化。以上变化可能由于增加起降坪，能为更多的eVTOL服务起降；而在起降坪功能方面，双起降混合坪包含了单起降坪、单起飞坪功能，所以在容量包线上没有变化。

表6为不同起降坪数量下各功能区之间的利用率。从表中可以看出，通过增加起降坪数量，能够增加起降坪利用率，减少登机口平均利用率，并减少eVTOL于登机口的等待率。这可能是通过增加起降坪数量，垂直起降机场能够容纳更多eVTOL，从而缓解了eVTOL于登机口的利用情况，并减少了eVTOL于登机口上的等待。

3.4.3 起降时长的影响

在已有的基础上调整eVTOL起降时长由原先60 s缩短至30 s，结果如图6与表7所示。

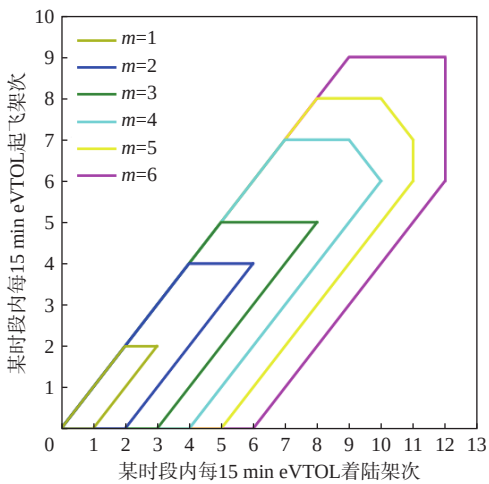


图6 缩短起降时长容量包线

Fig. 6 Capacity envelope of shorten take-off and land time

表7 缩短起降时长功能区利用率

Table 7 Functional area utilization rate of shorten take-off and land time

登机口数量	起降坪利用率/%	登机口平均利用率/%	登机口等待率/%
1	17	80	10
2	30	72	0
3	43	53	2
4	57	66	17
5	63	68	22
6	70	68	37

由图6可知，当登机口数量增加，其机场容量包线面积总体呈增加趋势；从增长幅度上来看，当登机口数量小于4时，有着大幅度增加，而登机口数量大于4时，增长幅度放缓。与图4相对比可知，更短的起降时长，能够增加垂直起降机场的容量，通过缩短30 s起降时长，能为垂直起降机场容量增加平均3架次，且随着登机口数量越多，其增加架次越大。上述原因可能是在登机口数量多的垂直起降机场，其起降坪占用紧张，通过减少起降时长，能够更有效的提高垂直起降机场的容量。

表7为其功能区利用率，总体上看，起降坪利用率呈上升趋势，而登机口平均利用率先下降后上升，登机口等待率也是下降后上升；在增长幅度上，起降坪利用率在登机口数量为4之后放缓，而登机口利用率在登机口数量为4之后趋于平缓，但登机口等待率却持续上升。对比表4可知，缩短后的起降时长，能够减少起降坪的利用率，在提高登机口的平均利用率的同时，减少eVTOL于登机口的等待率。以上可能是由于缩短起降时长后，eVTOL于起降坪上执行任务时间缩短，从而减少了起降坪上的利用率；当登机口数量足够多，投入垂直起降机场的eVTOL数量足够多时，垂直起降机场上的登机口使用情况更加紧张，但又由于起降时长短，因此，在登机口上等待率更小。

3.4.4 预降时刻的影响

分别假设eVTOL降落时刻为30 s、60 s、90 s、120 s和150 s，如图7所示，总体上eVTOL预计到达时刻间隔越长，垂直起降机场的最大吞吐量呈下降趋势；从不同登机口数量上看，登机口数量越多，受eVTOL预计到达时刻影响程度越深。这可能是登机口数量多的垂直起降机场能够在一段时间内处理更多的eVTOL起降，而不频繁的eVTOL降落，会浪费垂直起降机场的起降能力，因此受到预降时刻影响更深。

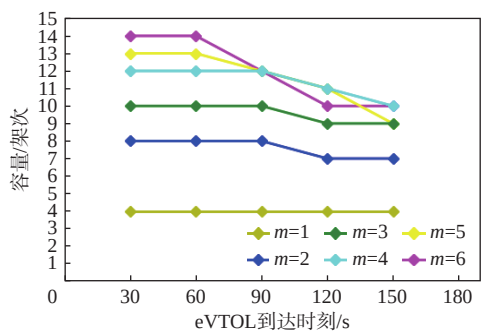


图7 eVTOL预降时刻与容量关系

Fig. 7 Relationship of estimated arrival time and capacity

3.4.5 周转时长的影响

假设周转时长为50 s、100 s、200 s、300 s、400 s、500 s和600 s，如图8所示。

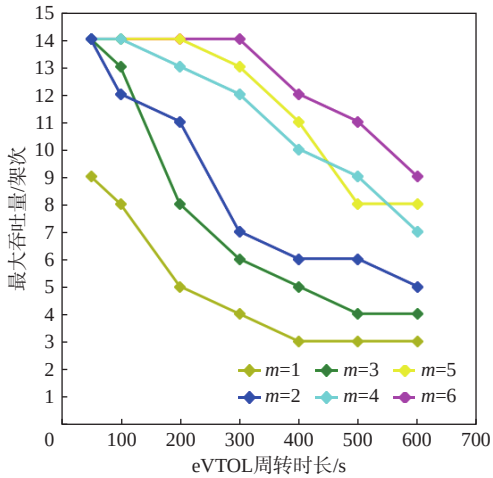


图 8　周转时长与容量关系

Fig. 8　Relationship of turnaround time and capacity

由图 8 可知，不同登机口数量下的垂直起降机场容量皆受到周转时长的影响，呈现明显的下降趋势；而在变化程度上，登机口数量多的垂直起降机场，其容量受周转时长影响下降趋势相较于登机口数量少的垂直起降机场较为平缓，通过增加 1 个登机口能够为其增加周转时长 100 s 内垂直起降机场容量的韧性。上述结果可能是因为更多的登机口，能够为更多的 eVTOL 提供周转服务，从而当单个登机口周转时长发生变化时，仍有其他登记口能提供服务，从而提高了垂直起降机场容量的韧性。

3.4.6　滑行时长的影响

假设 eVTOL 滑行时长为 5 s、30 s、45 s、60 s、90 s，结果如图 9 所示。

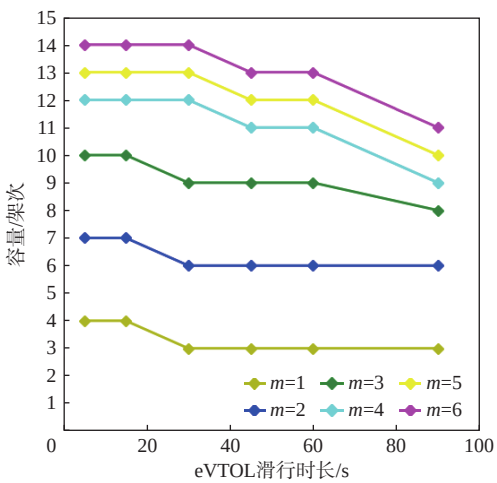


图 9　滑行时长与容量关系

Fig. 9　Relationship of taxi time and capacity

由图 9 可知，不同登机口数量下的垂直起降机场容量，都随着 eVTOL 滑行时长呈现下降趋势；而登机口数量多的垂直起降机场，其容量受滑行时长的影响变化越为明显。上述原因可能是由于滑行道用于连接各功能区，特别在登机口数量多的垂直起降机场，由于其具有起降更多 eVTOL 的能力，更长的滑行时长容易造成垂直起降机场内拥堵。

4　结　论

1）以任务网络为基础，对多商品流模型进行改进，从而证明本文构建出的垂直起降机场内的 eVTOL 任务规划模型具有可行性。

2）在垂直起降机场布局方面，起降坪数量和登机口数量为影响容量的关键因素。在单起降坪情况下，当起降坪利用率低时，增加 1 个登机口数量，能为垂直起降机场提供 3 架次容量；而在登机口数量为 4 的情况下，增加起降坪数量能为垂直起降机场提供 7 架次容量。

3）eVTOL 任务时长方面，eVTOL 的起降时长和周转时长为关键因素。在起降时长方面，通过缩短 30 s 起降时长能为垂直起降机场提供平均 3 架次容量；而在周转时长上，通过增加 1 个登机口能为垂直起降机场多提供 100 s 的垂直起降机场容量韧性。

参考文献（References）

[1] 李诚龙, 屈文秋, 李彦冬, 等. 面向 eVTOL 航空器的城市空中运输交通管理综述[J]. 交通运输工程学报, 2020, 20(4): 35-54.
LI C L, QU W Q, LI Y D, et al. Overview of traffic management of urban air mobility (UAM)with eVTOL aircraft[J]. Journal of Traffic and Transportation Engineering, 2020, 20(4): 35-54(in Chinese).

[2] 张洪海, 邹依原, 张启钱, 等. 未来城市空中交通管理研究综述[J]. 航空学报, 2021, 42(7): 75-99.
ZHANG H H, ZOU Y Y, ZHANG Q Q, et al. Future urban air mobility management: review[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(7): 75-99(in Chinese).

[3] 廖小罕, 屈文秋, 徐晨晨, 等. 城市空中交通及其新型基础设施低空公共航路研究综述[J]. 航空学报, 2023, 44(24): 028521.
LIAO X H, QU W Q, XU C C, et al. A review of urban air mobility and its new infrastructure low-altitude public routes[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(24): 028521(in Chinese).

[4] LAARMANN L, THOMA A, MISCH P, et al. Automotive safety approach for future eVTOL vehicles[J]. CEAS Aeronautical Journal, 2023, 14(2): 369-379.

[5] SU P Y, MAHESHWARI C, TUCK V M, et al. Incentive-compatible vertiport reservation in advanced air mobility: an auction-based approach[C]//Proceedings of the IEEE 63rd Conference on Decision and Control. Piscataway: IEEE Press, 2025: 7720-7727.

[6] SUZUKI A S, DAO Q V. A flight replanning tool for terminal area urban air mobility operations[C]//Proceedings of the IEEE/AIAA 41st Digital Avionics Systems Conference. Piscataway: IEEE Press, 2022: 1-7.

[7] XIANG S W, YE M X, ZHU S Q, et al. A multi-stage precision landing method for autonomous eVTOL based on multi-marker joint localization[C]//Proceedings of the IEEE International Confer-

ence on Robotics and Biomimetics. Piscataway: IEEE Press, 2023: 1-6.

[8] KLEINBEKMAN I C, MITICI M A, WEI P. eVTOL arrival sequencing and scheduling for on-demand urban air mobility[C]//Proceedings of the IEEE/AIAA 37th Digital Avionics Systems Conference. Piscataway: IEEE Press, 2018: 1-7.

[9] CONRAD C, XU Y, PANDA D, et al. Intelligent vertiport traffic flow management for scalable advanced air mobility operations[C]//Proceedings of the IEEE/AIAA 42nd Digital Avionics Systems Conference. Piscataway: IEEE Press, 2023: 1-10.

[10] PRADEEP P, WEI P. Heuristic approach for arrival management of eVTOLs in on-demand urban air mobility[C]//Proceedings of the IEEE/AIAA 37th Digital Avionics Systems Conference. Piscataway: IEEE Press, 2018: 1-10.

[11] PREIS L, HORNUNG M. Vertiport operations modeling, agent-based simulation and parameter value specification[J]. Electronics, 2022, 11(7): 1071.

[12] BRUNELLI M, DITTA C C, POSTORINO M N. New infrastructures for Urban Air Mobility systems: a systematic review on vertiport location and capacity[J]. Journal of Air Transport Management, 2023, 112: 102460.

[13] PREIS L. Estimating vertiport passenger throughput capacity for prominent eVTOL designs[J]. CEAS Aeronautical Journal, 2023, 14(2): 353-368.

[14] AHN B, HWANG H Y. Design criteria and accommodating capacity analysis of vertiports for urban air mobility and its application at gimpo airport in Korea[J]. Applied Sciences, 2022, 12(12): 6077.

[15] SCHWEIGER K, SCHMITZ R, KNABE F. Impact of wind on eVTOL operations and implications for vertiport airside traffic flows: a case study of hamburg and Munich[J]. Drones, 2023, 7(7): 464.

[16] PARK B T, KIM H, KIM S H. Vertiport performance analysis for on-demand urban air mobility operation in Seoul metropolitan area[J]. International Journal of Aeronautical and Space Sciences, 2022, 23(5): 1065-1078.

[17] NIU C C, RASTOGI P, SOMAN J, et al. Applying quantum computing to solve multicommodity network flow problem[J]. IEEE Transactions on Intelligent Transportation Systems, 2025, 26(4): 5094-5101.

[18] KHANAL D P, PYAKUREL U, DHAMALA T N. Maximum multicommodity flow with intermediate storage[J]. Mathematical Problems in Engineering, 2021, 2021(1): 5063207.

[19] NISHI T, AKIYAMA S, HIGASHI T, et al. Cell-based local search heuristics for guide path design of automated guided vehicle systems with dynamic multicommodity flow[J]. IEEE Transactions on Automation Science and Engineering, 2020, 17(2): 966-980.

[20] NARAYAN O, SANIEE I. An analytical solution to the multicommodity network flow problem with weighted random routing[J]. Applied Network Science, 2021, 6(1): 45.

[21] LIU J M, GUO Z, YU B. Optimising Gate assignment and taxiway path in a discrete time–space network: integrated model and state analysis[J]. Transportmetrica B: Transport Dynamics, 2023, 11(1): 1-23.

[22] VASCIK P D, HANSMAN R J. Development of vertiport capacity envelopes and analysis of their sensitivity to topological and operational factors[C]//Proceedings of the AIAA Scitech 2019 Forum. Reston: AIAA, 2019.

[23] 田勇, 付建军, 王艳军. 机场地面容量评估研究[J]. 南京航空航天大学学报, 2006, 38(5): 619-622.
TIAN Y, FU J J, WANG Y J. Research on airport ground capacity evaluation[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2006, 38(5): 619-622(in Chinese).

Vertiport operational task planning model and capacity estimation method

WEI Zhiqiang*，XIAO Xinlong

(College of Air Traffic Management，Civil Aviation University of China，Tianjin 300300，China)

Abstract：To solve the problem of electric vertical take-off and landing（eVTOL）operational planning and capacity estimation in a vertiport. Firstly, the operation task network was constructed by sorting out the layout structure of the vertiport. Then, based on the multi-commodity flow model and the entry and departure times of each node were set to construct the operational task planning model. The best task planning was taken as the optimization objective, and the node occupation time was taken as a constraint. Following the model's construction, a standard vertiport was utilized as an example, the problem was programmed and solved using Python and Gurobi, and the vertiport's maximum capacity as well as its influencing factors were computed and examined. The results show that the constructed task planning model is feasible. Regarding the variables influencing vertical take-off and landing airport capacity, when pad utilization is low, adding one gate can increase the capacity of three flights and provide a maximum capacity resilience of 100 seconds for vertiport; when boarding gate utilization is low, adding one pad can provide the capacity of seven flights for vertiport, and by reducing the take-off and landing time by thirty seconds, it can provide an average of three flights of capacity for vertiport. It can be seen that the research results are helpful for the estimation of vertiport capacity and provide a modeling idea for operational planning of eVTOL in vertiport.

Keywords：urban air mobiliy；vertiport；multi-commodity flow model；task planning；capacity estimation

Received：2024-04-23；**Accepted**：2024-08-30；**Published Online**：2024-09-11 09：48

URL：link.cnki.net/urlid/11.2625.V.20240910.1642.004

Foundation items：National Natural Science Foundation of China (U2133210); Tianjin Municipal Science and Technology Projects (23JCZDJC00580); Open Research Fund of Key Laboratory of LiDAR and Device Technology (LLD2023-008)

* **Corresponding author**. E-mail：weizhiqia@sina.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0223

基于增强逐点图卷积网络的民航短文本组合分类方法

刘晓琳[1,*]，宋营营[1]，李卓[2]

(1. 中国民航大学 电子信息与自动化学院，天津 300300；2. 中国农业大学 信息与电气工程学院，北京 100083)

摘 要：目前，大多数短文本分类方法存在信息挖掘不充分和局部信息关注度不足的问题，致使分类精度无法得到提升。鉴于此，提出一种融合增强语义信息和句法信息的逐点图卷积网络(ESS-PWGCN)少样本半监督民航短文本组合分类模型。筛选训练集高置信度关键词汇信息，丰富和增强民航短文本中关键信息的表达能力，扩大模型的应用领域；结合逐点卷积和图卷积网络(GCN)，并引入多头注意力机制，学习民航短文本的语义-句法信息，同时平衡文本图中全局-局部信息的影响权重；采用全连接层融合获取到的信息，输出分类结果；利用民航数据集和其他领域的公开数据集进行实验。结果表明：ESS-PWGCN模型与当前最先进的自训练文本图卷积网络(ST-TextGCN)模型相比，不仅分类的准确率和F_1值分别提高了4.59%和6.53%，而且具有更高的鲁棒性和泛化性。

关 键 词：短文本分类；深度学习；图卷积网络；逐点卷积；注意力机制；长短期记忆网络
中图分类号：V221$^+$.3; TB553
文献标志码：A **文章编号**：1001-5965(2026)06-1890-13

随着民航业及其智能终端技术的蓬勃发展，该行业积累了大量的短文本数据。这些数据涵盖了整个航班的飞行情况，包括飞行事故、设备事故、气象条件、人为失误等非结构化描述性信息，蕴含着关键的安全信息和潜在的致险因素。然而，由于人力、物力、机场资源有限，民航短文本数据的挖掘和分类效率低下，导致安全隐患的发现和处理常常滞后，从而影响整个民航系统的运行效率和客户满意度[1]。同时，相较于长文本数据，短文本数据的长度通常在20个词左右，缺乏足够的上下文信息，存在一词多义、一义多词、用词不规范等现象，这将干扰对短文本数据语义的准确解读，最终引起其底层特征到其高层语义之间的信息传输偏差。因此，对民航短文本数据进行深入挖掘和精确分类，不仅能够及时识别和处理其中潜在的安全隐患，而且对全面提高民航业的安全性、可靠性、运行效率及乘客满意度，具有十分重要的意义。

民航短文本分类以该领域中的事故案例、文本记录、维修日志、工作手册等数据为基础，利用特征工程、机器学习算法或深度学习算法对其进行分类。特征工程是最早被应用的一种短文本分类方法。常用的特征工程分类方法有词袋模型[2]、词频-逆文本频率(term frequency–inverse document frequency, TF-IDF)[3]、N-grams[4]等。这些方法大多建立在规则[5]和统计[6]的基础上，不仅需要额外的知识和繁琐的过程，而且会消耗大量的人力、物力、财力。此外，受人为主观因素的影响，分类的效率和长期准确率无法得到有效保证。为克服这些限制因素，有效捕获词语本身的语义特征，实现短文本自动分类。在以往的研究过程中，机器学习算法常被应用于此领

收稿日期：2024-04-16；**录用日期**：2024-07-12；**网络出版时间**：2024-08-13 12:34
网络出版地址：link.cnki.net/urlid/11.2625.V.20240812.1643.002
基金项目：天津市自然科学基金(17JCYBJC18200)
* **通信作者**. E-mail：caucyanjiusheng@163.com

引用格式：刘晓琳，宋营营，李卓. 基于增强逐点图卷积网络的民航短文本组合分类方法[J]. 北京航空航天大学学报，2026，52(6)：1890-1902. LIU X L，SONG Y Y，LI Z. Civil aviation short text combined classification method based on enhanced point-wise graph convolutional networks[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1890-1902 (in Chinese).

域, 如随机森林、最近邻居、朴素贝叶斯、支持向量机等[7]。尽管这些算法在一定程度上提高了短文本分类的计算效率, 但仅限于浅层次的特征提取, 未能完全捕获上下文深层次语义信息。此外, 因其通过改进特征工程予以实现, 所以需要大量的有标签数据保证模型训练的充分性[8], 且无法解决特征向量高度稀疏性问题[9]。

近年来, 随着词向量[10-11]的出现, 深度学习算法被广泛用于短文本分类领域, 其中, 图卷积网络 (graph convolutional networks, GCN) 在半监督文本分类上的出色表现, 引起了学术界和工业界许多研究人员的广泛关注[12]。尽管已有研究聚焦于民航数据集的应用, 但涉及 GCN 及其衍生模型的研究相对较少, 本文旨在探索 GCN 在民航短文本分类领域的应用潜力, 尤其是在挖掘深层文本特征和解决内容稀疏性问题方面的优势[8]。Yao 等[13] 提出的 TextGCN 模型, 运用 GCN 进行文本分类, 以文档和单词为节点构建语料库文本图, 将文本分类问题转化为节点分类问题。Liu 等[14] 受 TextGCN 模型启发, 通过构建一种张量 GCN(TensorGCN) 分类模型, 分别获取语义信息图、句法依赖关系图和局部序列上下文信息图来优化模型。Yang 等[15] 在 TextGCN 模型的基础上, 提出异构图注意力网络 (heterogeneous graph attention networks, HGAT) 分类模型, 利用短文本、主题和实体构建异构信息网络, 并引入双重注意力机制区分不同节点的重要性, 以提高分类精度。Wu[16] 为缓解短文本的内容稀疏性, 引入与短文本相关的先验知识, 构建包含多样化知识的异构信息网。Ji 等[17] 构建了主动半监督学习异构图注意力网络 (active semi-suprivised learning heterogeneous graph attention networks, ASSL-HAGAT), 在外部知识库作为先验知识构建知识图谱的基础上, 引入不确定感知机改进的主动学习和双层注意力机制进一步优化异构信息网, 提高模型整体性能。

尽管现有的图卷积网络算法兼顾了短文本的语义和句法信息, 但仍有不足之处: ①对有限的有标签文本数据利用不充分, 主要集中在对短文本语义信息和句法信息的获取及扩展上, 限制了算法在不同应用场景中的泛化能力; ②在短文本分类的过程中, 现有的算法更倾向于强调全局信息的重要性, 而忽视了从局部信息中挖掘和利用细粒度特征的潜力, 导致在跨类别特征提取及同类别特征学习方面存在显著的不足。

鉴于此, 本文根据短文本的特点, 提出一种融合增强语义信息和句法信息的逐点图卷积网络 (ehanced semantic-syntactic point-wise graph convolutional networks, ESS-PWGCN) 少样本半监督民航短文本组合分类模型, 通过多种方式提升该模型对民航短文本类别特征的捕捉和学习能力, 同时扩展其在不同应用领域中的适用性。首先, 通过筛选训练集中的关键词汇信息, 丰富和增强民航短文本中关键信息的表达能力, 增大其信息密度; 其次, 结合双向长短期记忆 (bi-directional long short-term memory, Bi-LSTM) 网络和多头自注意力 (multi-head self-attention, MHSA) 机制, 深入挖掘民航短文本底层语义特征, 增强模型的语义理解能力; 然后, 在 GCN 中引入逐点卷积 (point-wise convolution, PWC), 精确提取并分析民航短文本的句法结构, 同时有效平衡整个文本图中的全局与局部信息; 此外, 采用多头交互注意力 (multi-head interactive-attention, MHIA) 机制, 根据文本图中各节点的相对重要性自动分配合适的注意力权重, 并将获取的信息并行融合到 Softmax 层, 输出分类结果; 最后, 在民航 ASRS 短文本数据集上进行大量的分类对比实验, 结果表明: ESS-PWGCN 模型的分类效果在准确率和 F_1 值方面均优于基线模型, 有效提高了民航短文本分类精度。与传统的民航短文本模型相比, 该模型的主要创新点如下:

1) 采用词汇置信度和类别相对频率筛选出训练集中的高置信度关键词汇及其标签, 并将其作为伪标签数据补充到训练集中用于模型训练。与其他模型相比, 本文模型仅使用训练集中少量标注的关键词汇信息, 就能有效丰富和增强民航短文本中关键信息的表达能力, 同时扩展其在各个应用领域中的实用性和有效性, 提高模型整体性能。

2) 在 GCN 中融入 PWC, 有效平衡文本图中的全局信息与局部信息。一方面, 深入挖掘文本图中节点的全局特征, 细致区分不同类别数据的特征差异; 另一方面, 精细学习每个节点的局部特征, 捕捉同类数据的内在相似性。这种全局与局部信息的双重视角, 使本文模型在进行分类决策时能够更加全面和准确地评估不同节点和子图的重要性和贡献度。

3) 在最后的全连接层中, 引入 MHIA 机制将短文本和训练集关键词汇的语义-句法信息并行融合得到分类特征。这不仅使本文模型以多层次、全方位的方式, 对民航短文本中的丰富语义深入挖掘和理解, 而且通过其并行处理能力, 高效捕捉民航短文本中的显式语义信息及隐含的句法关系和语义联接, 增强模型对复杂语义结构的理解和表达能力, 从而确保提炼的特征在语义上的准确性和一致

性，提升本文模型在民航短文本特征提取过程中潜在歧义的处理能力。

1 本文模型

作为信息传递的媒介，短文本网络是一个典型的复杂语言系统。短文本之间的关联和内部的语义结构构成非线性聚合的复杂空间语义特征，给民航短文本分类任务带来了挑战。因此，综合考虑语义信息和句法信息的动态影响，本文设计的ESS-PWGCN模型结构如图1所示，其中，$\boldsymbol{e}_n^{\mathrm{c}}$、$\boldsymbol{e}_m^{\mathrm{w}}$分别表示第$n$文本和第$m$个关键词汇的向量序列；$\boldsymbol{h}_n^{\mathrm{c}}$、$\boldsymbol{h}_m^{\mathrm{w}}$分别表示第$n$文本和第$m$个关键词汇的隐藏状态向量；$\boldsymbol{D}_n$、$\boldsymbol{W}_m$分别表示在GCN中第$n$文本和第$m$个关键词汇的节点信息；$\boldsymbol{h}_n^{\mathrm{cs}}$、$\boldsymbol{h}_m^{\mathrm{ws}}$分别表示第$n$文本和第$m$个关键词汇的语义信息向量；$\boldsymbol{h}_{m+n}^{\mathrm{gs}}$表示第$n$文本和第$m$个关键词汇的句法信息向量；$\boldsymbol{h}_n^{\mathrm{cg}}$、$\boldsymbol{h}_m^{\mathrm{wg}}$分别表示第$n$文本上下文感知和第$m$个关键词汇上下文感知的句法信息向量。该模型由关键词汇筛选模块、语义信息编码模块、句法信息编码模块和信息融合分类模块4部分组成。其中，关键词汇筛选模块，根据词汇置信度和类别相对频率，筛选出训练集中的关键词汇及其类别信息；语义信息编码模块，引入Bi-LSTM网络和MHSA机制，充分学习民航短文本和训练集关键词汇的上下文语义信息；句法信息编码模块，利用GCN和PWC，编码句法信息，同时通过多头自注意力权重的分配，有效平衡文本图中的全局-局部信息；信息融合分类模块，通过MHIA机制将获取的民航短文本和训练集关键词汇的语义-句法信息并行融合到Softmax分类层，输出分类结果，完成分类任务。

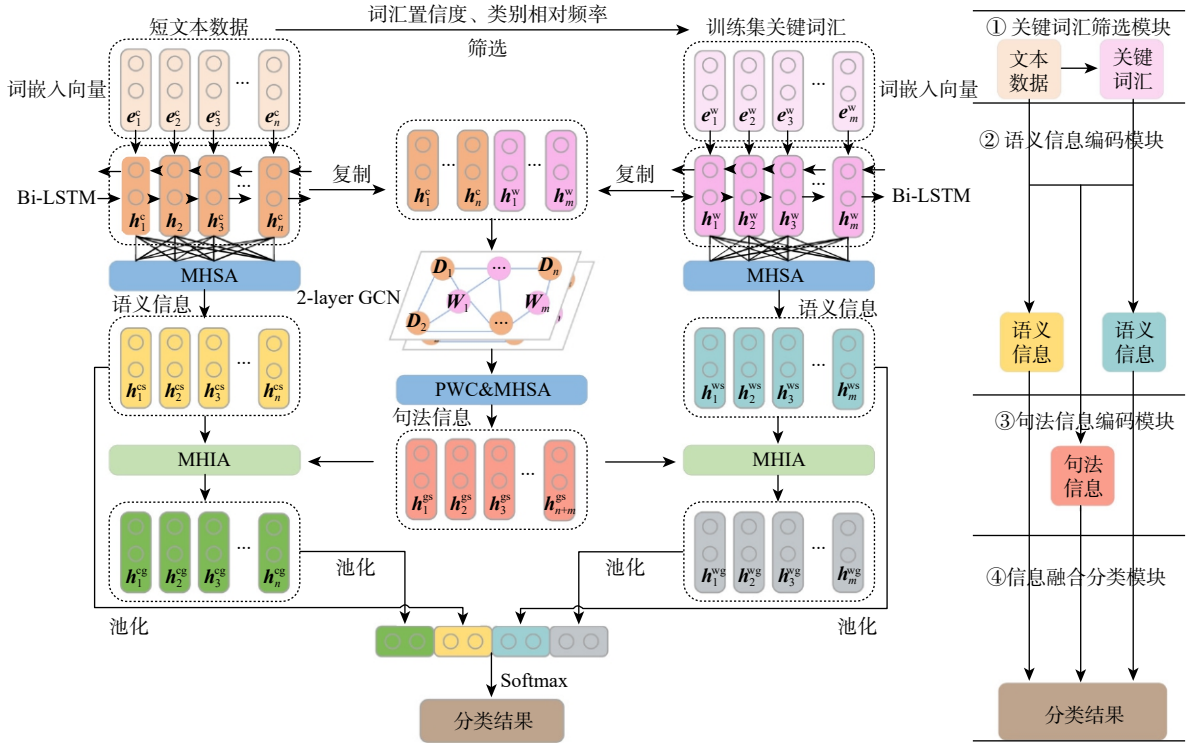


图1 ESS-PWGCN模型结构

Fig. 1 Model architecture of ESS-PWGCN

1.1 短文本网络拓扑图构建

为充分研究短文本之间的类别关系，根据文本拓扑结构，将整个短文本数据集网络定义成不规则的拓扑图，可表示为

$$G_t = (V, E, \boldsymbol{A}) \tag{1}$$

式中：V为拓扑图中的节点集合，满足$V=\{v_i|i=1,2,\cdots,n+m\}$，$n+m$为短文本和训练集关键词汇的总量；$E$为节点构成的边的集合；$\boldsymbol{A}$为刻画节点之间关系的邻接矩阵且$\boldsymbol{A}\in\mathbf{R}^{(n+m)\times(n+m)}$，$\boldsymbol{A}$中的元素$\boldsymbol{A}_{ij}$为节点$v_i$和节点$v_j$之间类别特征的相互影响程度。其中，节点间类别特征相互影响程度的衡量是通过多头交互注意力机制在全局范围内动态评估和优化各节点间的影响权重实现的。

在短文本分类的过程中，拓扑图上观测到的n个短文本分类特征为

$$\boldsymbol{X}_t = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_{n+m}] \in \mathbf{R}^{(n+m)\times d} \tag{2}$$

式中：d为短文本的特征向量维度；n为短文本数据总量；m为训练集中的关键词汇总量。将整个数据集构建拓扑图进行分类如图2所示，其中，C_k表示文本$\boldsymbol{D}_c$的类别k的文档数。民航短文本分类问题可

以表述为: 给出短文本数据集、训练集关键词汇和短文本网络拓扑图G_t, 寻找最优的模型可学习权重参数函数$F_w(\cdot)$, 输出测试集的分类结果。训练集和测试集之间的映射关系可表示为

$$\{\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_c\}=F_w\{\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_{n+m-c};G_t\} \tag{3}$$

式中: c为测试集文本总量。

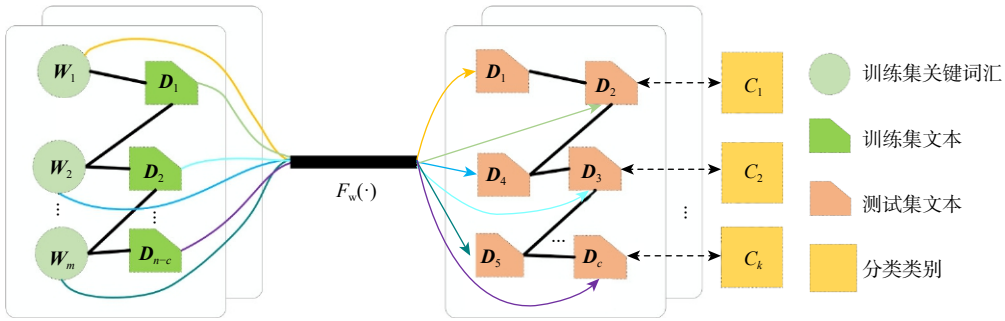


图 2　图结构分类

Fig. 2　Classification of graph structures

以短文本和训练集关键词汇为节点构成的不规则拓扑图结构, 采用欧几里得标准卷积操作很难捕捉数据的空间特征。为解决这一问题, 将图卷积操作延伸到文本数据处理的拓扑图结构中。根据图谱理论原理, 拓扑图的性质由其拉普拉斯矩阵及特征值矩阵决定。为避免多层卷积后, 输入与输出之间的范围差距越来越大, 确保信息传播的稳定性, 对图的拉普拉斯矩阵对称归一化为

$$\boldsymbol{L}=\boldsymbol{D}-\boldsymbol{A}=\boldsymbol{I}-\hat{\boldsymbol{D}}^{-\frac{1}{2}}\hat{\boldsymbol{A}}\hat{\boldsymbol{D}}^{-\frac{1}{2}} \tag{4}$$

$$\hat{\boldsymbol{A}}=\boldsymbol{A}+\boldsymbol{I} \tag{5}$$

式中: $\boldsymbol{I}$为单位对角阵; $\boldsymbol{D}$为邻接矩阵$\boldsymbol{A}$的度矩阵, 且有$\boldsymbol{D}=\sum\boldsymbol{A}_{ij}$; $\hat{\boldsymbol{A}}$为加入自身影响后的邻接矩阵; $\hat{\boldsymbol{D}}$为$\hat{\boldsymbol{A}}$的度矩阵, 且有$\hat{\boldsymbol{D}}=\sum\hat{\boldsymbol{A}}_{ij}$。对半正定实对称矩阵$\boldsymbol{L}$进行拉普拉斯矩阵谱分解为

$$\boldsymbol{L}=\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{-1}=\boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^{\mathrm{T}} \tag{6}$$

式中: $\boldsymbol{\Lambda}$为特征值对角阵; $\boldsymbol{U}$为单位特征 (列向量) 矩阵, 且$\boldsymbol{U}$为正交矩阵, 满足$\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}=\boldsymbol{E}$。通过正交矩阵$\boldsymbol{U}$将卷积操作转换到傅里叶域中进行。频谱域的图卷积定义为输入的图信号$\boldsymbol{x}\in\mathbf{R}^N$与图形滤波器$\boldsymbol{g}_{\boldsymbol{\theta}}=\mathrm{diag}(\boldsymbol{\theta})(\boldsymbol{\theta}\in\mathbf{R}^N)$在傅里叶域的乘积, 故图结构的频谱卷积操作为

$$\boldsymbol{g}_{\boldsymbol{\theta}}*\boldsymbol{x}=\boldsymbol{U}\boldsymbol{g}_{\boldsymbol{\theta}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{x} \tag{7}$$

式中: $\boldsymbol{g}_{\boldsymbol{\theta}}$为可学习卷积核参数; $*$为图卷积操作运算符。

然而, 式 (7) 在图谱卷积操作中, 计算量很大。为降低计算的复杂度, Defferrard 等[18] 使用切比雪夫多项式 (Chebyshev polynomial)$T_k(x)$的k阶截断来获取对$\boldsymbol{g}_{\boldsymbol{\theta}}(\boldsymbol{\Lambda})$的近似为

$$\boldsymbol{g}_{\boldsymbol{\theta}'}(\boldsymbol{\Lambda})\approx\sum_{k=0}^{K}\boldsymbol{\theta}'_kT_k(\tilde{\boldsymbol{\Lambda}}) \tag{8}$$

式中: $\tilde{\boldsymbol{\Lambda}}$为$\boldsymbol{L}$的最大特征值 (即谱半径) 缩放后的特征向量矩阵, 满足$\tilde{\boldsymbol{\Lambda}}=\frac{2}{\lambda_{\max}}\boldsymbol{\Lambda}-\boldsymbol{I}$; $\boldsymbol{\theta}_k'\in\mathbf{R}^k$为切比雪夫向量; T_k为切比雪夫多项式。

故简化后的图卷积操作为

$$\boldsymbol{g}_{\boldsymbol{\theta}}*\boldsymbol{x}\approx\sum_{k=0}^{K}\boldsymbol{\theta}'_k\boldsymbol{U}T_k(\tilde{\boldsymbol{\Lambda}})\boldsymbol{U}^{\mathrm{T}}\boldsymbol{x}\approx\sum_{k=0}^{K}\boldsymbol{\theta}'_k\boldsymbol{U}T_k(\tilde{\boldsymbol{L}})\boldsymbol{x} \tag{9}$$

式中: $T_k(\tilde{\boldsymbol{\Lambda}})$为$\boldsymbol{\Lambda}$的$k$阶多项式, 且满足$\boldsymbol{U}\tilde{\boldsymbol{\Lambda}}^k\boldsymbol{U}^{\mathrm{T}}=(\boldsymbol{U}\tilde{\boldsymbol{\Lambda}}\boldsymbol{U}^{\mathrm{T}})^k=\tilde{\boldsymbol{L}}^k$, 其中, $\tilde{\boldsymbol{L}}=\frac{2}{\lambda_{\max}}\boldsymbol{L}-\boldsymbol{I}$。

在此基础上, 为进一步简化计算, 采用一阶多项式近似, 简化后的图谱卷积近似为

$$\boldsymbol{g}_{\boldsymbol{\theta}}*\boldsymbol{x}\approx\boldsymbol{\theta}(\boldsymbol{I}+\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{D}^{-\frac{1}{2}})\boldsymbol{x} \tag{10}$$

由于$\boldsymbol{I}+\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{D}^{-\frac{1}{2}}$的特征值范围为 [0, 2]。当网络非常深时, 重复此操作会引起梯度消失或梯度爆炸。为避免此问题, 进行归一化操作为

$$\boldsymbol{A}'=\hat{\boldsymbol{D}}^{-\frac{1}{2}}\hat{\boldsymbol{A}}\hat{\boldsymbol{D}}^{-\frac{1}{2}} \tag{11}$$

结合式(10)和式(11), 根据图卷积网路的层级传导规则, 隐藏层的非线性函数为

$$f(\boldsymbol{H}^l,\boldsymbol{A})=\sigma(\hat{\boldsymbol{D}}^{-\frac{1}{2}}\hat{\boldsymbol{A}}\hat{\boldsymbol{D}}^{-\frac{1}{2}}\boldsymbol{H}^l\boldsymbol{W}^l)=\sigma(\boldsymbol{A}'\boldsymbol{H}^l\boldsymbol{W}^l) \tag{12}$$

$$\boldsymbol{H}^0=\boldsymbol{X} \tag{13}$$

$$\boldsymbol{H}^{l+1}=f(\boldsymbol{H}^l,\boldsymbol{A}) \tag{14}$$

式中: σ为激活函数; $\boldsymbol{H}^l$为第l卷积层的输出; $\boldsymbol{W}^l$为作线性变换的权重矩阵, 为待训练参数; $\boldsymbol{X}$为初始节点特征矩阵。

1.2　关键词汇筛选模块

为充分利用和挖掘短文本中的有标签数据信息, ESS-PWGCN 模型根据词汇置信度和类别相对频率, 筛选出训练集中的类别关键词汇, 并将这些关键词汇作为伪标签数据添加到训练集中, 用于缓解短文本的内容稀疏性, 同时丰富和增强短文本中

关键信息的表达能力，以提高分类的效果。训练集关键词汇的筛选如下：

1）词汇置信度。即词汇在最常见类别中出现的概率，可表示为

$$D_r = \mathrm{Max}(C_0, C_1, \cdots, C_l) \Big/ \sum_{k=0}^{l} C_k \tag{15}$$

式中：D_r为训练集文本中的第r个词汇的置信度；l为类别总数。

2）类别相对频率。即词汇在最常见类别中出现的频率相对于总频率，可表示为

$$F_i = lC_K \Big/ \sum_{k=0}^{l} C_K \tag{16}$$

$$P_r = f_v(l + \mathrm{Max}(F_i))/l \tag{17}$$

式中：f_v为训练集词汇的最小频次；C_K为训练集中第K类的文档总数。

1.3 语义信息编码模块

为充分学习上下文序列信息，ESS-PWGCN模型引入Bi-LSTM网络和MHSA机制学习短文本和训练集关键词汇的语义编码信息。

1.3.1 双向长短期记忆网络

ESS-PWGCN模型基于8.4×10^{10}个单词的大型语料库预训练Glove模型[10]来获取300维短文本词嵌入向量，并引入Bi-LSTM来捕获上下文序列信息。假设短文本上下文序列为$\boldsymbol{T} = [\boldsymbol{e}_1^{\mathrm{c}}, \boldsymbol{e}_2^{\mathrm{c}}, \cdots, \boldsymbol{e}_n^{\mathrm{c}}]$，训练集关键词汇序列为$\boldsymbol{W} = [\boldsymbol{e}_1^{\mathrm{w}}, \boldsymbol{e}_2^{\mathrm{w}}, \cdots, \boldsymbol{e}_m^{\mathrm{w}}]$，则通过Bi-LSTM得到短文本和训练集关键词汇的隐藏状态向量分别为$\boldsymbol{H}^{\mathrm{t}} = [\boldsymbol{h}_1^{\mathrm{c}}, \boldsymbol{h}_2^{\mathrm{c}}, \cdots, \boldsymbol{h}_n^{\mathrm{c}}]$和$\boldsymbol{H}^{\mathrm{w}} = [\boldsymbol{h}_1^{\mathrm{w}}, \boldsymbol{h}_2^{\mathrm{w}}, \cdots, \boldsymbol{h}_m^{\mathrm{w}}]$，其中$(\boldsymbol{h}_i^{\mathrm{c}}, \boldsymbol{h}_i^{\mathrm{w}}) \in \mathbf{R}^{2d_h}$，$d_h$为单向LSTM隐藏向量维度。

1.3.2 多头自注意力机制

ESS-PWGCN模型对短文本和训练集关键词汇的隐藏状态，应用MHSA机制学习其语义编码信息。首先，定义一个Key序列$\boldsymbol{k} = [\boldsymbol{k}_1, \boldsymbol{k}_2, \cdots, \boldsymbol{k}_n]$和一个Query序列$\boldsymbol{q} = [\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_n]$，通过Key计算注意力分配并附加到Value中的注意力得分。然后，利用注意力函数，输出Key和Query的分数序列为

$$\boldsymbol{A}(\boldsymbol{k}, \boldsymbol{q}) = \mathrm{Softmax}(f_m(\boldsymbol{k}, \boldsymbol{q}))\boldsymbol{k} \tag{18}$$

式中：f_m为$\boldsymbol{q}_i$和$\boldsymbol{k}_j$之间的语义相关性函数，其表达式为

$$f_m(\boldsymbol{q}_i, \boldsymbol{k}_j) = \tanh([\boldsymbol{q}_i; \boldsymbol{k}_j] \cdot \boldsymbol{W}_{\mathrm{A}}) \tag{19}$$

式中：$\boldsymbol{W}_{\mathrm{A}}$为学习权重矩阵，其中，$\boldsymbol{W}_{\mathrm{A}} \in \mathbf{R}^{2d_H}$，$d_H$为注意力隐藏层维度；[·;·]表示向量拼接。多头注意力(multi-head attention, MHA)机制能够并行学习注意力机制中的不同分数，并通过一个全连接层将每一层的结果进行连接，投影到特定的隐藏层维度上，最终的输出为

$$\mathrm{MHA}(\boldsymbol{k}, \boldsymbol{q}) = [\boldsymbol{O}^1 \oplus \boldsymbol{O}^2 \oplus \cdots \oplus \boldsymbol{O}^{n_{\mathrm{h}}}] \cdot \boldsymbol{W}_{\mathrm{A}} \tag{20}$$

$$\boldsymbol{O}^{\mathrm{h}} = \boldsymbol{A}^{\mathrm{h}}(\boldsymbol{k}, \boldsymbol{q}) \tag{21}$$

式中：$\oplus$为连接层；$\boldsymbol{W}_{\mathrm{A}} \in \mathbf{R}^{d_H \times d_H}$为权重矩阵；$\boldsymbol{O}^a = \{\boldsymbol{O}_1^a, \boldsymbol{O}_2^a, \cdots, \boldsymbol{O}_n^a,\}$为第$a$个自注意力层的输出，且满足$a \in [1, n_{\mathrm{h}}]$。给定上下文的隐藏状态向量$\boldsymbol{H}^{\mathrm{t}}$和训练集关键词汇的隐藏状态向量$\boldsymbol{H}^{\mathrm{w}}$。然后利用MHSA机制推出上下文和训练集关键词汇的语义编码信息$\boldsymbol{H}^{\mathrm{Y}}$、$\boldsymbol{H}^{\mathrm{y}}$为

$$\boldsymbol{H}^{\mathrm{Y}} = \mathrm{MHSA}(\boldsymbol{H}^{\mathrm{t}}, \boldsymbol{H}^{\mathrm{t}}) \tag{22}$$

$$\boldsymbol{H}^{\mathrm{y}} = \mathrm{MHSA}(\boldsymbol{H}^{\mathrm{w}}, \boldsymbol{H}^{\mathrm{w}}) \tag{23}$$

式中：$\boldsymbol{H}^{\mathrm{Y}} = \{\boldsymbol{h}_1^{\mathrm{Y}}, \boldsymbol{h}_2^{\mathrm{Y}}, \cdots, \boldsymbol{h}_n^{\mathrm{Y}}\} \in \mathbf{R}^{d_H \times n}$；$\boldsymbol{H}^{\mathrm{y}} = \{\boldsymbol{h}_1^{\mathrm{y}}, \boldsymbol{h}_2^{\mathrm{y}}, \cdots, \boldsymbol{h}_m^{\mathrm{y}}\} \in \mathbf{R}^{d_H \times m}$。

1.4 句法信息编码模块

为在捕获文本图全局信息的同时，兼顾局部信息的学习，本文提出了GCN和PWC的融合算法，对短文本句法依存树信息进行编码的同时，平衡文本图中的全局信息和局部信息。

1.4.1 句法依存树

本文利用Spacy库中的强大语法依存解析功能建立句法依存树，遍历每个节点及其子节点，当节点与其子节点之间存在语法关系时，将两者之间的依存关系设置为1，同时也将子节点与此节点之间的依存关系设置为1，表示依存关系是双向的，否则两者之间的依存关系设置为0。最终得到整个语料库句法依存关系的自适应二进制邻接矩阵。句法依存树关系分析如图3所示，其中，VERB表示动词；ADV表示副词；ADJ表示形容词；NOUN表示名词；ADP表示介词。compound表示复合词；nsubj表示名词性主语；dobj表示直接宾语；advmod表示状语；amod表示形容词修饰语；probj表示介词宾语。

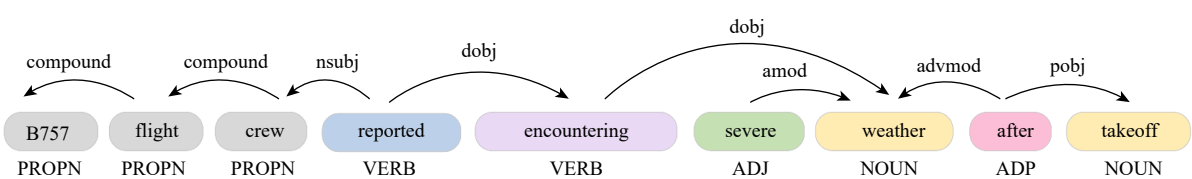


图3 句法依存树关系分析

Fig. 3 Analysis of syntactic dependency tree relationships

1.4.2　图卷积网络

图卷积网络擅长处理包含大量节点和丰富关系的图结构数据。给定一个图, 有k个节点, 则邻接矩阵$\boldsymbol{A}\in\mathbf{R}^{k\times k}$。为方便起见, 假设 GCN 有$L$层, 其中, $l\in[1,2,\cdots,L]$, $\boldsymbol{H}_i^l$为节点i的最终状态。节点的图卷积表示为

$$\boldsymbol{H}_i^l=\sigma\left(\sum_{j=1}^{k}\boldsymbol{A}_{ij}\boldsymbol{W}^l\boldsymbol{H}_j^{l-1}+\boldsymbol{b}^l\right)\tag{24}$$

式中: $\boldsymbol{W}^l$为线性变换权重矩阵; $\boldsymbol{b}^l$为偏移向量; σ为非线性激活函数。由于 GCN 的层数为 2 时, 效果最佳[19], 故本文使用双层 GCN, 第 1 层使用 ReLU 激活函数, 第 2 层使用 Softmax 激活函数。当上一层的隐藏状态$\boldsymbol{H}_i^{l-1}$发生线性变换后, 将与之直接相关节点的隐藏状态进行求和, 得到当前层的隐藏状态$\boldsymbol{H}_i^l$。

Zhang 等[20] 提出句法依存树上的图卷积。在图卷积网络架构中, 随着层数L的增加, 图中的节点将受到其L阶邻域节点的连续影响。因此, 通过对短文本的句法依存树进行卷积操作, 为目标节点添加语法约束并确定相应的句法距离, 能够在保留非连续词汇信息的同时, 有效整合这些词汇的特征, 处理句子复杂的语法结构。基于此, 本文提出一种将句法依存树上的图卷积和逐点卷积相结合的方法来分析短文本数据句法依存树上的依存关系。此外, 引入 MHSA 机制对经过卷积处理后的句法关系进一步编码, 得到最终的句法编码信息。在将$\boldsymbol{H}^{\mathrm{c}}$反馈给连续的下一层 GCN 之前, 对$\boldsymbol{H}_i^l$进行位置感知变换, 以增强模型对句法结构位置信息的感知能力, 从而提升句法依存分析的准确性和效率, 可表示为

$$\boldsymbol{S}_i^l=F(\boldsymbol{H}_i^l)\tag{25}$$

式中: $F(\cdot)$为位置权重分布函数[21-22], 用于强调靠近特定目标词汇的重要性, $F(\cdot)$函数如下:

$$F(\boldsymbol{H}_i^l)=e_i\boldsymbol{H}_i^l\tag{26}$$

$$e_i=\begin{cases}1-\dfrac{\tau+1-i}{n} & 1\leqslant i\leqslant\tau+1\\ 0 & \tau+1\leqslant i\leqslant\tau+m\\ 1-\dfrac{i-\tau-m}{n} & \tau+m\leqslant i\leqslant n\end{cases}\tag{27}$$

式中: e_i为第 i 个词汇对应的位置。

首先, 根据句法依存树中的语法关系, 得到短文本邻接矩阵$\boldsymbol{A}\in\mathbf{R}^{k\times k}$。接着, 遵循自循环的思想, 确保每个词汇与其自身相邻, 将$\boldsymbol{A}$的对角线值设置为 1。同时, 考虑依存结构的双向性, 即若存在节点i到节点j的边, 则$|\boldsymbol{A}_{ij}|=1$和$|\boldsymbol{A}_{ji}|=1$; 否则$|\boldsymbol{A}_{ij}|=0$, $|\boldsymbol{A}_{ji}|=0$。最后, 引入带有归一化因子的图卷积更新每个节点的表示为

$$\tilde{\boldsymbol{H}}_i^l=\sum_{j=1}^{n}\boldsymbol{A}_{ij}\boldsymbol{W}^l\boldsymbol{S}_j^{l-1}\tag{28}$$

$$\boldsymbol{H}_i^l=\mathrm{ReLu}\left(\frac{\tilde{\boldsymbol{H}}_i^l}{\boldsymbol{d}_i+\boldsymbol{I}}+\boldsymbol{b}^l\right)\tag{29}$$

式中: $\boldsymbol{S}_j^{l-1}\in\mathbf{R}^{2d_H}$为第 j个词汇表示, 同时也是前一个 GCN 层的输出; $\boldsymbol{H}_i^l\in\mathbf{R}^{2d_h}$为当前 GCN 层的输出; $\boldsymbol{d}_i=\sum_{j=1}^{n}\boldsymbol{A}_{ij}$为句法依存树中第$i$个词汇的度矩阵; $\boldsymbol{W}^l$和$\boldsymbol{b}^l$均为可训练参数矩阵。

1.4.3　逐点卷积

模型导出前层的输出$\boldsymbol{H}^l=[\boldsymbol{H}_1^l,\boldsymbol{H}_2^l,\cdots,\boldsymbol{H}_n^l]$后, 对其进行 PWC 操作。逐点表示卷积核的大小为 1。给定输入序列$\boldsymbol{h}$, PWC 的定义为

$$P(\boldsymbol{h})=\sigma(\boldsymbol{h}*\boldsymbol{W}_P+\boldsymbol{b}_P)\tag{30}$$

式中: σ为激活函数 ReLU; $*$为逐点卷积操作运算符; $\boldsymbol{W}_P\in\mathbf{R}^{2d_H\times 2d_H}$为逐点卷积核的学习权重矩阵; $\boldsymbol{b}_P\in\mathbf{R}^{2d_H}$为逐点卷积核的偏置项。通过对$\boldsymbol{H}^l$进行逐点卷积操作, 得到 GCN 当前层的输出$\boldsymbol{H}^{l_P}$为

$$\boldsymbol{H}^{l_P}=P(\boldsymbol{H}^l)\tag{31}$$

第L层的最终输出为$\boldsymbol{H}^{L_P}=[\boldsymbol{H}_1^{L_P},\boldsymbol{H}_2^{L_P},\cdots,\boldsymbol{H}_n^{L_P}]$, $\boldsymbol{H}_i^{L_P}\in\mathbf{R}^{2d_H}$。通过 MHSA 机制对第$L$层输出进行编码, 得到句法编码信息$\boldsymbol{H}^{\mathrm{G}}$为

$$\boldsymbol{H}^{\mathrm{G}}=\mathrm{MHSA}(\boldsymbol{H}^{L_P},\boldsymbol{H}^{L_P})\tag{32}$$

式中: $\boldsymbol{H}^{\mathrm{G}}=[\boldsymbol{H}_1^{\mathrm{G}},\boldsymbol{H}_2^{\mathrm{G}},\cdots,\boldsymbol{H}_n^{\mathrm{G}}]\in\mathbf{R}^{d_H\times n}$; d_H为 MHSA 机制的维度。

1.5　信息融合分类模块

ESS-PWGCN 模型采用 MHIA 机制将得到的训练集关键词汇信息、语义编码信息和句法编码信息并行融合, 以提高分类效果。

1.5.1　多头交互注意力机制

多头交互注意力是$e_i\neq k$的常见形式, 给定句法编码信息$\boldsymbol{H}^{\mathrm{G}}$、上下文语义编码信息$\boldsymbol{H}^{\mathrm{Y}}$和训练集关键词汇语义编码信息$\boldsymbol{H}^{\mathrm{y}}$, 得到上下文感知的句法信息$\boldsymbol{H}^{\mathrm{S}}$和训练集关键词汇语义感知的句法信息$\boldsymbol{H}^{\mathrm{s}}$分别为

$$\boldsymbol{H}^{\mathrm{S}}=\mathrm{MHIA}(\boldsymbol{H}^{\mathrm{Y}},\boldsymbol{H}^{\mathrm{G}})\tag{33}$$

$$\boldsymbol{H}^{\mathrm{s}}=\mathrm{MHIA}(\boldsymbol{H}^{\mathrm{y}},\boldsymbol{H}^{\mathrm{G}})\tag{34}$$

式中: $\boldsymbol{H}^{\mathrm{S}}=[\boldsymbol{H}_1^{\mathrm{S}},\boldsymbol{H}_2^{\mathrm{S}},\cdots,\boldsymbol{H}_n^{\mathrm{S}}]\in\mathbf{R}^{d_H\times n}$; $\boldsymbol{H}^{\mathrm{s}}=[\boldsymbol{H}_1^{\mathrm{s}},\boldsymbol{H}_2^{\mathrm{s}},\cdots,\boldsymbol{H}_m^{\mathrm{s}}]\in\mathbf{R}^{d_H\times m}$。

1.5.2　信息融合

将上下文语义编码信息$\boldsymbol{H}^{\mathrm{Y}}$、训练集关键词汇语义编码信息$\boldsymbol{H}^{\mathrm{y}}$、上下文感知的句法信息$\boldsymbol{H}^{\mathrm{S}}$和训

练集关键词汇语义感知的句法信息$\boldsymbol{H}^{\mathrm{s}}$进行平均池化操作，然后，将其串联起来得到最终的分类特征$\boldsymbol{u}$为

$$\boldsymbol{u}=[\boldsymbol{H}^{\mathrm{Y}}_{\mathrm{avg}}\oplus\boldsymbol{H}^{\mathrm{y}}_{\mathrm{avg}}\oplus\boldsymbol{H}^{\mathrm{S}}_{\mathrm{avg}}\oplus\boldsymbol{H}^{\mathrm{s}}_{\mathrm{avg}}] \tag{35}$$

$$\boldsymbol{H}^{\mathrm{Y}}_{\mathrm{avg}}=\sum_{i=1}^{n}\frac{\boldsymbol{H}^{\mathrm{Y}}_{i}}{n} \tag{36}$$

$$\boldsymbol{H}^{\mathrm{y}}_{\mathrm{avg}}=\sum_{i=1}^{m}\frac{\boldsymbol{H}^{\mathrm{y}}_{i}}{m} \tag{37}$$

$$\boldsymbol{H}^{\mathrm{S}}_{\mathrm{avg}}=\sum_{i=1}^{n}\frac{\boldsymbol{H}^{\mathrm{S}}_{i}}{n} \tag{38}$$

$$\boldsymbol{H}^{\mathrm{s}}_{\mathrm{avg}}=\sum_{i=1}^{n}\frac{\boldsymbol{H}^{\mathrm{s}}_{i}}{m} \tag{39}$$

式中：表示注意力连接层；[] 表示将这些信息通过注意力连接层串联起来得到最后的分类特征。

1.5.3 数据分类

根据最终得到分类特征$\boldsymbol{u}$，将其输入到 Softmax 层，得到短文本在不同类别中的概率分布为

$$\boldsymbol{x}=\tilde{\boldsymbol{W}}_{\boldsymbol{u}}^{\mathrm{T}}\boldsymbol{u}+\boldsymbol{b}_{\boldsymbol{u}} \tag{40}$$

$$\hat{\boldsymbol{y}}=\mathrm{Softmax}(\boldsymbol{x})=\exp(\boldsymbol{x})\Big/\sum_{k=1}^{l}\exp(\boldsymbol{x}) \tag{41}$$

式中：$\hat{\boldsymbol{y}}\in\mathbf{R}^{l}$为短文本的预测标签；$l$为类别总量；$\tilde{\boldsymbol{W}}_{\boldsymbol{u}}^{\mathrm{T}}\in\mathbf{R}^{1\times l}$和$\boldsymbol{b}_{\boldsymbol{u}}\in\mathbf{R}^{l}$分别为可训练权重矩阵和偏置项。

1.6 模型训练

模型引入分类交叉熵和 L2-正则化之和作为损失函数，并采用反向传播更新权重和参数，具体计算式为

$$L_{\mathrm{oss}}=-\sum_{i}\sum_{j=1}^{l}y_i^j\log_{\hat{y}_i}j+\lambda\|\boldsymbol{\gamma}\|_2 \tag{42}$$

式中：L_{oss}为训练损失函数；i为训练集中的第i个样本；j为类别目录中的第j类别；y为短文本的真实标签；λ为正则化参数；$\boldsymbol{\gamma}$为可训练参数矩阵。

2 实验分析

2.1 数据集介绍

本文基于 ASRS 数据集及其他领域公开数据集 Ohsumed[23]、MR[24]、Laptop[25] 和 Snippets[23] 进行了一系列实验，旨在验证本文模型的可行性和优越性。实验设计包括模型性能对比实验、参数调优实验、模块消融实验和不同数据集泛化对比实验，并对实验结果进行了深入分析。为探究 ESS-PWGCN 模型在少样本半监督短文本数据上的分类效果，实验按照 90∶9∶1 的比例将数据集划分为测试集、训练集和验证集，以确保实验的科学性和可信性。数据集的详细信息如表 1 所示，数据集的介绍如下。

表 1 数据集信息

Table 1 Dataset information

数据集	类别总量	数据总量/个	测试集/个	训练集/个	验证集/个	平均词汇个数
ASRS	10	8 865	7 979	798	88	17.468 9
Ohsumed	23	7 400	6 660	666	74	11.929 6
Snippets	8	7 370	6 633	664	73	15.543 8
MR	2	8 000	7 200	720	80	20.998 9
Laptop	3	3 532	3 179	318	35	20.334 7

ASRS。ASRS 数据集源自全球民航安全事件报告数据库，涵盖了飞行员、空管人员、技术维修人员等其他工作人员描述和记录发生的不安全事件、潜在的安全隐患及相关经验教训等，旨在防止类似事件的再次发生。本文选取该数据库中近年来近一万条事件的简要概述和主要致险因素用于实验，整个数据集分为 10 类。

Ohsumed。Ohsumed[23] 数据集与医学主题有关的语料库，整个数据集分为 23 类。

Snippets。Snippets[23] 数据集由网络引擎返回的片段组成，与科技、体育、娱乐等方面有关，整个数据集分为 8 类。

MR。MR[24] 数据集与电影评论有关，整个数据集分为积极和消极 2 类。

Laptop。Laptop[25] 数据集与笔记本电脑产品评论有关，整个数据集分为 3 类。

2.2 实验环境和参数设置

实验环境如下：处理器为 Intel(R) Core(TM) i5-8250U；CPU 频率为 1.80 GHz；机带 RAM 内存为 8.00 GB；GPU 为 Intel(R)UHD Graphics 620；操作系统为 Windows 10(64 位)；程序设计语言为 Python 3.6；所采用的深度学习框架为 PyTorch 1.10.2。

模型基本参数设置如下：最大迭代次数 (epoch) 为 100；词汇置信度阈值 (vocab_confidence) 为 0.9；训练集词汇的最小频次 (vocab_freq) 为 3；词汇移除频次 (remove_word_freq) 为 2；注意力头数 (head) 为

3; 词嵌入维度 (embid_dim) 和隐藏层维度 (hidden) 均为 300; 批处理量 (batch_size) 为 128; 学习率 (lr) 为 0.001; L2 为 0; dropout 为 0.5; 早停机制 (early_stopping) 为 20。

基线模型的基本参数如下: 最大迭代次数为 100; 词嵌入维度和隐藏层维度均为 300; 批处理量为 128; 学习率为 0.001; L2 为 0; dropout 为 0.5; 早停机制为 20。

2.3　模型评价指标介绍

本文选取文本分类领域常用的 2 个度量标准: 准确率 (accuracy, ACC) 和 F_1 值。ACC 是在所有测试样本中被正确分类的样本数量与测试样本总数的比值, F_1 值是精确率 P_{recision} 和召回率 R_{ecall} 的加权平均, 具体的计算式为

$$A_{\text{CC}}=\frac{T_{\text{P}}+T_{\text{N}}}{T_{\text{P}}+T_{\text{N}}+F_{\text{P}}+F_{\text{N}}} \tag{43}$$

$$P_{\text{recision}}=\frac{T_{\text{P}}}{T_{\text{P}}+F_{\text{P}}} \tag{44}$$

$$R_{\text{ecall}}=\frac{T_{\text{P}}}{T_{\text{P}}+F_{\text{N}}} \tag{45}$$

$$F_1=\frac{2P_{\text{recision}}R_{\text{ecall}}}{P_{\text{recision}}+R_{\text{ecall}}} \tag{46}$$

式中: T_{P} 是实际为正, 预测为正的样本量; T_{N} 是实际为负, 预测为负的样本量; F_{P} 是实际为负, 预测为正的样本量; F_{N} 是实际为正, 预测为负的样本量。

2.4　模型性能分析实验

为全面评估 ESS-PWGCN 模型的性能, 将 ESS-PWGCN 模型与以下 8 个经典的基线模型进行对比。具体的实验结果如表 2 所示。

表 2　ASRS 数据集不同模型实验结果对比

Table 2　Comparison of experimental results of different models on ASRS datasets

模型	分类结果										A_{CC}	F_1	T/s
	Equipment_Tooling	Weather	Manuals	Human_Factors	Company_Policy	Procedure	Ambiguous	Airport	Chart_Or_Publication	Environment-Non_Weather_Related			
ESS-PWGCN	273	626	125	496	672	278	419	487	471	438	0.537 0	0.530 4	18.617 4
ST-TextGCN[23]	0	674	0	420	865	15	197	699	542	116	0.442 2	0.343 7	1.392 8
Mean[26]	0	740	0	333	212	0	5	222	432	152	0.262 7	0.244 2	1.159 6
Transformer[27]	105	624	110	401	476	345	231	757	333	160	0.443 9	0.354 3	23.925 4
TextCNN[28]	50	684	82	351	551	260	89	431	355	181	0.380 2	0.334 1	5.015 4
TextRNN[29]	29	726	125	328	540	203	109	703	264	70	0.388 1	0.375 0	9.469 6
TextING[30]	173	662	67	409	464	205	338	757	160	97	0.420 9	0.382 4	64.567 0
TextGCN[13]	32	618	36	531	626	182	255	620	551	249	0.463 7	0.407 7	4.873 5

1) Mean[26] 模型通过捕捉短文本中的关键信息, 学习文本的平均池化表示, 以完成文本分类任务。

2) Transformer[27] 模型通过自注意力机制, 对整个输入序列进行操作, 完成分类任务。

3) TextCNN[28] 模型基于卷积神经网络 (convolutional neural networks, CNN), 使用卷积层来捕捉文本中不同长度的关键信息。

4) TextRNN[29] 模型基于循环神经网络 (recurrent neural networks, RNN), 通过引入循环连接, 捕获序列中的顺序信息。

5) TextGCN[13] 模型基于 GCN, 根据词共现和文档词关系构建文本图, 将文档分类问题转化为节点分类问题。

6) TextING[30] 模型基于图神经网络 (graph neural networks, GNN), 将每个文档构建成单独的图, 学习文本级别的信息。

7) 自训练文本图卷积网络 (self-training text graph convolution networks, ST-TextGCN)[23] 模型基于 GCN, 加入自训练算法, 提高短文本分类效果。

8) ESS-PWGCN 模型基于 GCN, 融入训练集关键词汇信息和短文本语义-句法信息, 并采用逐点图卷积网络在深度和广度上挖掘短文本信息特征, 提高分类性能。

基于 ASRS 数据集, ESS-PWGCN 模型与其他基线模型的分类准确率和 F_1 值随迭代次数的变化曲线分别如图 4 和图 5 所示, ASRS 训练集类别分布情况如图 6 所示。

实验设置了早停机制, 通过监控验证集损失来防止模型过拟合, 保证模型具有一定的泛化能力。本次实验设置了相同的最大迭代次数和早停机制。由分类结果可知, 首先, 从计算成本来看, ESS-PWGCN 模型的训练时间并不是最长的, 特别是与 TextING 模型相比, 该模型显示出更短的训练时间和更高的计算效率。其次, 从模型性能来看, ESS-PWGCN 模型在 23 次左右就达到了最佳分类效果, 平均每个 epoch 的训练时间 T 为 18.62 s, 学习效率

高，且与目前最优的ST-TextGCN模型相比，准确率提高了9.48%左右，F_1值提高了18.67%左右。最后，从性能平衡性来看，ESS-PWGCN模型能够在复杂性和计算成本之间达到良好的平衡，且不易受训练集类别分布不均衡影响，具有较高的鲁棒性。

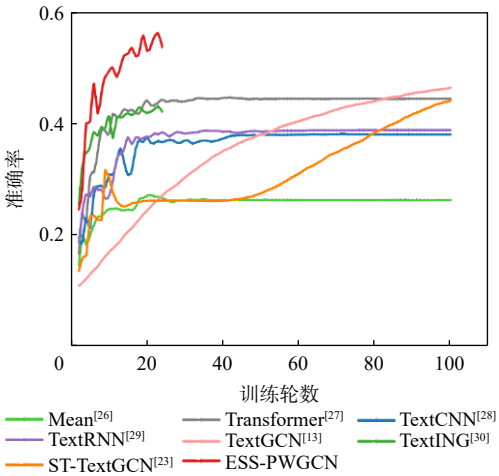


图4 分类准确率随迭代次数变化曲线

Fig. 4 Classification accuracy curves with number of iterations

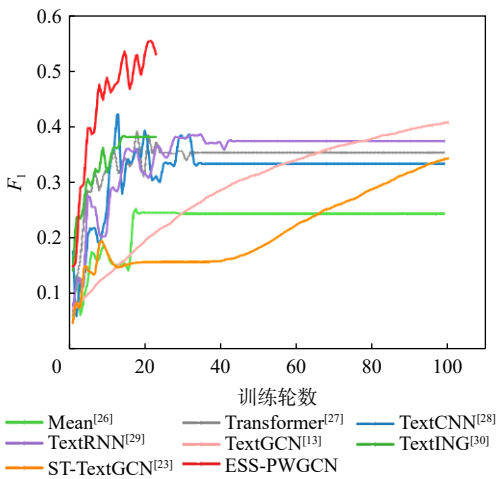


图5 分类F_1值随迭代次数变化曲线

Fig. 5 Classification F_1 value curves with number of iterations

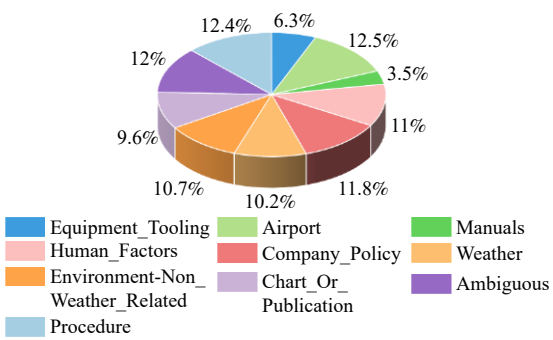


图6 训练集类别分布

Fig. 6 Category distribution of training set

2.5 参数优化实验

本文ESS-PWGCN模型包含关键词汇筛选模块、语义信息编码模块、句法信息编码模块和信息融合分类模块，这些模块中的词汇置信度阈值和注意力头数的设置影响着模型的分类性能，故分别对其进行参数调优实验，以求得模型的最优参数和最佳分类效果。

2.5.1 词汇置信度阈值最优化

词汇置信度阈值是ESS-PWGCN模型中的重要参数之一，其不仅决定着训练集关键词汇筛选的质量，而且还影响着文本图的构建效果。为获取最佳的词汇置信度阈值，使用枚举法对其大小进行调整。实验以ASRS数据集为例，选取了0.7、0.8、0.9、1.0、1.1这5个阈值进行分类结果验证。调优实验的结果如图7所示。

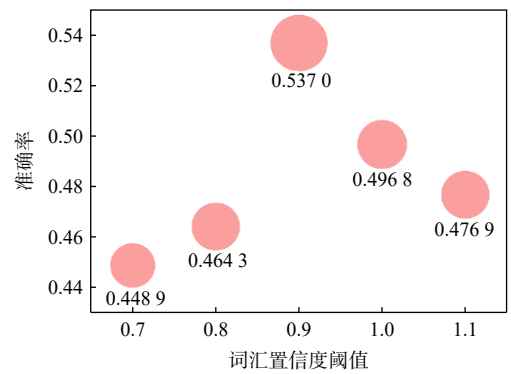


图7 词汇置信度阈值对分类效果的影响

Fig. 7 Effect of lexical confidence thresholds on classification effectiveness

气泡越大，F_1值越大。由图7可知，随着词汇置信度阈值的增加，分类的准确率和F_1值先增高后降低，当词汇置信度阈值设置为0.9时，气泡最大且准确率最高，ESS-PWGCN模型达到最佳分类效果。一方面，较高的阈值有助于选取更有质量的训练集关键词汇从而提高分类性能；另一方面，词汇置信度过高会导致模型对短文本中的重要信息关注不全面，使模型难以迭代到最佳分类效果。因此，在ESS-PWGCN模型训练过程中，词汇置信度阈值设置在0.9时，其分类效果最佳。

2.5.2 注意力头数最优化

注意力头数直接影响着模型对短文本数据的理解深度和不同信息的融合效率，是ESS-PWGCN模型中的关键参数。为获取注意力头数的最优值，实验以ASRS数据集为例，将注意力头数设置为1、2、3、4、5，并分析其对分类效果的影响。注意力头数的优化实验结果如图8所示。

由图8可知，当注意力头数为3时，气泡最大且准确率最高，ESS-PWGCN模型达到了最优分类性能。注意力头数过多或过少，模型的分类性能都

会下降。这主要是因为注意力头数过少时, 模型对短文本理解深度欠佳且对不同信息的融合效率低; 注意力头数增大到某个门槛时, 模型所需要计算的参数爆炸使得计算难度及复杂度迅速增加, 导致计算机资源消耗和模型过拟合风险增加。因此, 为实现 ESS-PWGCN 模型的最佳分类性能, 多头注意力机制的头数应设置为 3。

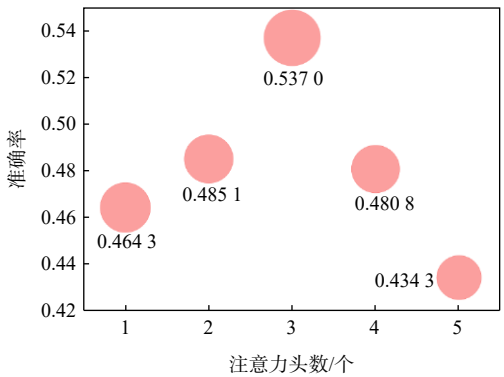


图 8　注意力头数对分类效果的影响

Fig. 8　Effect of number of attention heads on classification effect

2.6　模块消融实验

为验证关键词汇筛选模块和逐点卷积模块在 ESS-PWGCN 模型短文本分类实验中的有效性, 分别对其进行消融实验。

2.6.1　关键词汇筛选模块评估

为综合评估训练集关键词汇在短文本分类过程中的作用, 构建了无关键词汇筛选模块的 NW-ESS-PWGCN 模型。NW-ESS-PWGCN 模型与 ESS-PWGCN 模型的对比实验结果如图 9 所示。

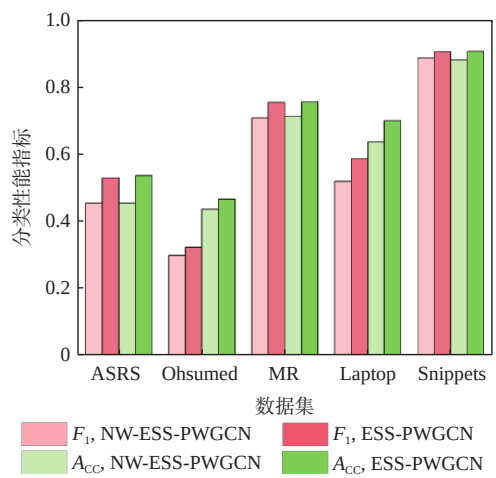


图 9　关键词汇筛选模块消融实验结果

Fig. 9　Results of ablation experiment of key vocabulary screening module

由图 9 可知, 在无关键词汇筛选模块的情况下, 模型的准确率会下降 4.72% 左右, F_1 值会下降 0.468 左右。由此说明, 关键词汇筛选模块提取的训练集关键词汇信息, 能够有效缓解短文本数据的内容稀疏性, 丰富和增强其关键信息的表达能力。这不仅提高了本文模型对短文本数据的理解深度, 还拓展了其在不同应用场景中的适用性和灵活性, 证明了关键词汇筛选模块的有效性和必要性。

2.6.2　逐点卷积操作模块评估

为研究逐点卷积操作模块在整个模型中的作用, 构建了无逐点卷积操作的 NP-ESS-PWGCN 模型。ESS-PWGCN 模型与 NP-ESS-PWGCN 模型的对比实验结果如图 10 所示。

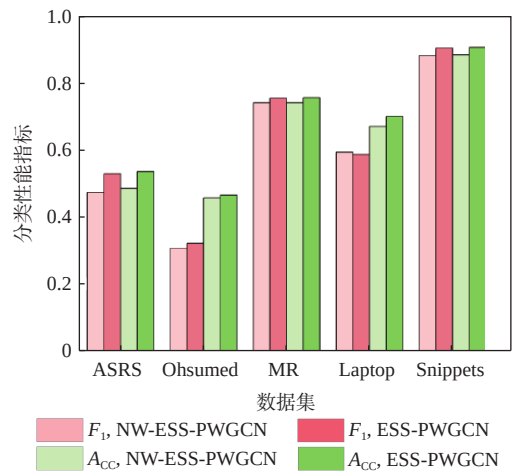


图 10　逐点卷积模块消融实验结果

Fig. 10　Ablation results of point-by-point convolution module

由图 10 可知, 在无逐点卷积操作时, 模型的准确率会下降 2.47% 左右, F_1 值会下降 1.97% 左右。由此说明, 逐点卷积操作能够深入挖掘短文本数据内部的细节特征, 提升模型对短文本内容的洞察力和解读效率, 使得模型能够精确捕捉到短文本中的关键类别特征。这种细粒度的特征学习方法, 不仅提升了模型对短文本中细微语义差异的响应能力, 而且增强了本文模型在处理多样化和复杂化数据时的鲁棒性和准确性, 证明了逐点卷积操作模块引入的创新性和重要性。

2.7　泛化实验

为全面评估 ESS-PWGCN 模型的泛化能力, 并验证其在不同领域的适用性, 通过一系列的实验, 深入研究 ESS-PWGCN 模型在面对不同语义结构和领域特征时的适应性和鲁棒性, 进而验证模型广泛的适用性和优越的泛化性能。在 ASRS 数据集的基础上, 引入了其他领域的公开数据集 Ohsumed、MR、Laptop、Snippets, 用以进行 ESS-PWGCN 模型的泛化性能实验。在模型训练的过程中, 所有数据集均按照 90: 9: 1 的比例划分为测试集、训练集和

验证集，以确保模型训练和评估的一致性。模型的其他超参数与 2.2 节实验环境和参数设置中的模型基本参数一致，以确保实验结果的可比性。模型仍选择准确率和 F_1 值，这 2 个度量指标来评估模型在不同领域数据集上的分类性能。各模型在不同领域数据集上的分类效果如表 3、图 11 和图 12 所示。其中，表 3 中的加粗部分为最优分类结果，下划线部分是次优分类结果。

表 3　不同数据集不同模型泛化实验结果对比

Table 3　Comparison of the results of generalization experiments of different models on different datasets

模型	A_{CC}					F_1				
	ASRS	Ohsumed	MR	Laptop	Snippets	ASRS	Ohsumed	MR	Laptop	Snippets
Mean[26]	0.297 0	0.179 7	0.673 5	0.540 8	0.655 5	0.100 8	0	0.558 9	0.260 6	0.587 1
Transformer[27]	0.490 9	0.291 0	0.720 7	0.649 9	0.817 3	0.384 7	0.100 0	0.639 0	0.498 1	0.759 8
TextCNN[28]	0.433 4	0.230 3	0.730 6	0.615 0	0.802 2	0.286 4	0.100 0	0.751 9	0.386 4	0.744 6
TextRNN[29]	0.434 8	0.235 6	0.730 7	0.626 6	0.799 3	0.226 0	0	0.666 7	0.420 6	0.755 1
TextGCN[13]	0.465 1	0.303 6	0.651 7	0.617 2	0.709 6	0.427 1	0.148 9	0.651 7	0.486 8	0.703 2
TextING[30]	0.438 3	0.409 6	0.718 6	0.663 1	0.841 5	0.410 6	0.243 1	0.728 2	0.525 4	0.840 2
ST-TextGCN[23]	0.475 0	0.420 9	0.702 1	0.662 5	0.877 7	0.394 0	0.269 6	0.702 1	0.533 0	0.876 7
ESS-PWGCN	**0.537 0**	**0.465 8**	**0.757 5**	**0.700 5**	**0.907 1**	**0.530 4**	**0.322 6**	**0.755 9**	**0.587 4**	**0.905 6**

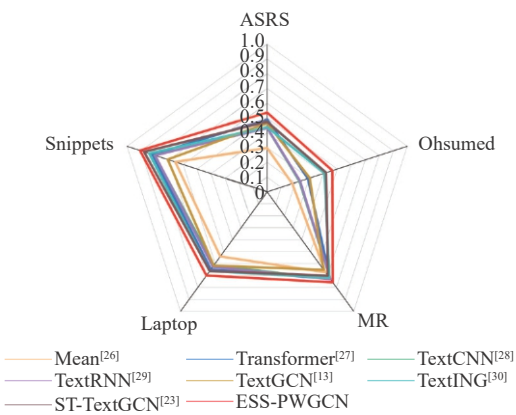


图 11　不同模型不同数据集准确率结果对比

Fig. 11　Comparison of accuracy results of different models and different datasets

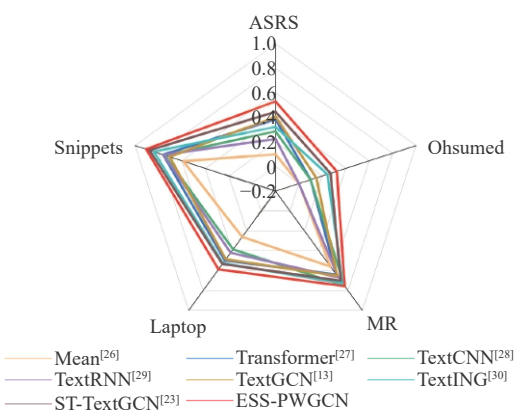


图 12　不同模型不同数据集 F_1 值结果对比

Fig. 12　Comparison of F_1 values of different models and different datasets

由图 11 和图 12 可知，ESS-PWGCN 模型在其他短文本集上也有良好的分类效果，与目前最先进的 ST-TextGCN 模型相比，准确率提高了 4.59% 左右，F_1 值提高了 6.53% 左右。

3　结　论

本文针对短文本信息利用不充分及局部信息关注度不足问题，构建了融入训练集关键词汇信息的逐点图卷积网络模型，提高了民航短文本分类模型的精确性和泛化性，得到的主要结论如下：

1）本文将训练集关键词汇信息整合到民航短文本数据集中，直接避免对外部知识库的依赖，降低本文模型对外部资源的需求，有效解决了民航短文本固有的内容稀疏性问题。更重要的是，仅精选少量的训练集关键词汇信息，本文模型就能够准确地捕捉和理解民航短文本中的核心内容，从而提升分类性能。

2）本文在图卷积网络中加入的逐点卷积操作，为每个节点信息的处理引入差异化权重机制。这种细粒度的信息处理方式，有助于充分学习文本图中最相关的类别特征信息，有效提升了本文模型对民航短文本数据内在特征的挖掘能力。

3）本文在全连接层中引入的多头交互注意力机制，实现了短文本和训练集关键词汇及其语义-句法信息的高效并行融合，并通过精细的注意力权重分配，得到了综合的分类特征，提高了本文模型对民航短文本细节的敏感度，也为模型赋予了更好的泛化能力和鲁棒性。

4）本文在民航数据集和其他领域的公开数据集上对本文模型进行实验。结果表明：基于不同领域的数据集，ESS-PWGCN 模型在分类准确率和 F_1 值方面均达到了最优效果，且学习效率高，具有

较高的鲁棒性。由此可见，该模型不仅比其他基线模型具有一定的优势，而且对不同领域的数据集具有一定的泛化性。

由于当前数据集中的文本长度只有 20 个词左右，一旦训练样本的长度较大时，本文的 ESS-PWGCN 模型的文本分类性能表现会有所下降。因此，未来将进行相应的研究，优化和扩展模型架构以应对更长的文本类别特征提取等挑战。

参考文献（References）

[1] 李博涵, 向宇轩, 封顶, 等. 融合知识感知与双重注意力的短文本分类模型[J]. 软件学报, 2022, 33(10): 3565-3581.
LI B H, XIANG Y X, FENG D, et al. Short text classification model combining knowledge aware and dual attention[J]. Journal of Software, 2022, 33(10): 3565-3581(in Chinese).

[2] ZHANG Y, JIN R, ZHOU Z H. Understanding bag-of-words model: a statistical framework[J]. International Journal of Machine Learning and Cybernetics, 2010, 1(1): 43-52.

[3] BLEI D M, NG A Y, JORDAN M I. Latent dirichlet allocation[J]. Journal of Machine Learning Research, 2003, 3: 993-1022.

[4] JOULIN A, GRAVE E, BOJANOWSKI P, et al. Bag of tricks for efficient text classification[EB/OL]. (2016-07-06)[2024-04-05]. https://arxiv.org/abs/1607.01759.

[5] LIANG H, SUN X, SUN Y L, et al. Text feature extraction based on deep learning: a review[J]. EURASIP Journal on Wireless Communications and Networking, 2017, 2017(1): 211.

[6] XU Y, HONG K, TSUJII J, et al. Feature engineering combined with machine learning and rule-based methods for structured information extraction from narrative clinical discharge summaries[J]. Journal of the American Medical Informatics Association, 2012, 19(5): 824-832.

[7] 贾宝惠, 姜番, 王玉鑫, 等. 基于民机维修文本数据的故障诊断方法[J]. 航空学报, 2023, 44(5): 253-267.
JIA B H, JIANG F, WANG Y X, et al. Fault diagnosis method based on civil aircraft maintenance text data[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(5): 253-267(in Chinese).

[8] 蒋云良, 王青朋, 张雄涛, 等. 基于门控双层异构图注意力网络的半监督短文本分类[J]. 模式识别与人工智能, 2023, 36(7): 602-612.
JIANG Y L, WANG Q P, ZHANG X T, et al. Semi-supervised short text classification based on gated double-layer heterogeneous graph attention network[J]. Pattern Recognition and Artificial Intelligence, 2023, 36(7): 602-612(in Chinese).

[9] 万家山, 吴云志. 基于深度学习的文本分类方法研究综述[J]. 天津理工大学学报, 2021, 37(2): 41-47.
WAN J S, WU Y Z. Review of text classification research based on deep learning[J]. Journal of Tianjin University of Technology, 2021, 37(2): 41-47(in Chinese).

[10] RAMACHANDRAN D, PARVATHI R. Enhanced classification of crisis related tweets using deep learning models and word embeddings[J]. International Journal of Web Engineering and Technology, 2021, 16(2): 158-186.

[11] WU Y, ZHAO S, LI W. Phrase2Vec: phrase embedding based on parsing[J]. Information Sciences, 2020, 517: 100-127.

[12] HUANG Y R, CHEN J J, ZHENG S M, et al. Hierarchical multi-attention networks for document classification[J]. International Journal of Machine Learning and Cybernetics, 2021, 12(6): 1639-1647.

[13] YAO L, MAO C S, LUO Y. Graph convolutional networks for text classification[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2019, 33(1): 7370-7377.

[14] LIU X E, YOU X X, ZHANG X, et al. Tensor graph convolutional networks for text classification[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2020, 34(5): 8409-8416.

[15] YANG T C, HU L M, SHI C, et al. HGAT: heterogeneous graph attention networks for semi-supervised short text classification[J]. ACM Transactions on Information Systems(TOIS), 2021, 39(3): 1-29.

[16] WU M Q. Commonsense knowledge powered heterogeneous graph attention networks for semi-supervised short text classification[J]. Expert Systems with Applications, 2023, 232: 120800.

[17] JI Z Y, KONG D Y, YANG Y Y, et al. ASSL-HGAT: active semi-supervised learning empowered heterogeneous graph attention network[J]. Knowledge-Based Systems, 2024, 290: 111567.

[18] DEFFERRARD M, BRESSON X, VANDERGHEYNST P. Convolutional neural networks on graphs with fast localized spectral filtering[EB/OL]. (2022-04-04)[2024-04-06]. https://arxiv.org/pdf/1606.09375v1.

[19] 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.

[20] ZHANG C, LI Q C, SONG D W. Aspect-based sentiment classification with aspect-specific graph convolutional networks[EB/OL]. (2019-09-08)[2024-04-06]. https://arxiv.org/abs/1909.03477.

[21] 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. Kerrville: Association for Computational Linguistics, 2017: 452-461.

[22] LI X, BING L D, LAM W, et al. Transformation networks for target-oriented sentiment classification[EB/OL]. (2018-03-03)[2024-04-06]. https://arxiv.org/abs/1805.01086.

[23] CUI H Y, WANG G K, LI Y X, et al. Self-training method based on GCN for semi-supervised short text classification[J]. Information Sciences, 2022, 611: 18-29.

[24] PANG B, LEE L. Seeing stars: exploiting class relationships for sentiment categorization with respect to rating scales[C]//Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL'05). Kerrville: Association for Computational Linguistics, 2005: 115-124.

[25] LIN Y X, MENG Y X, SUN X F, et al. BertGCN: transductive text classification by combining GCN and BERT[EB/OL]. (2021-03-12)[2024-04-06]. https://arxiv.org/abs/2105.05727.

[26] SUBAKTI A, MURFI H, HARIADI N. The performance of BERT as data representation of text clustering[J]. Journal of Big Data, 2022, 9(1): 15.

[27] DAI Z H, YANG Z L, YANG Y M, et al. Transformer-XL: attentive language models beyond a fixed-length context[EB/OL]. (2019-01-09)[2024-04-06]. https://arxiv.org/abs/1901.02860.

[28] KIM Y. Convolutional neural networks for sentence classification[EB/OL]. (2014-08-25)[2024-04-06]. https://arxiv.org/abs/1408.5882.

[29] LIU P F, QIU X P, HUANG X J. Recurrent neural network for text classification with multi-task learning[EB/OL] . (2016-03-17)[2024-04-06]. https://arxiv.org/abs/1605.05101.

[30] ZHANG Y F, YU X L, CUI Z Y, et al. Every document owns its structure: inductive text classification via graph neural networks[EB/OL] . (2020-04-22)[2024-04-06]. https://arxiv.org/abs/2004.13826.

Civil aviation short text combined classification method based on enhanced point-wise graph convolutional networks

LIU Xiaolin[1,*], SONG Yingying[1], LI Zhuo[2]

(1. College of Electronic Information and Automation, Civil Aviation University of China, Tianjin 300300, China;
2. School of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China)

Abstract: The improvement of classification accuracy is currently hampered by the fact that most short text classification approaches suffer from inadequate information mining and insufficient attention to local information. In light of this, an enhanced semantic-syntactic point-wise graph convolutional network (ESS-PWGCN) short-text combination classification model with few samples and semi-supervised civil aviation was proposed. Firstly, the model selects training set high-confidence keyword information to enrich and enhance the expression of key information within civil aviation short texts, thereby broadening the applicability of the model. Secondly, it balances the influence weights of global-local information within the textual graph structure while learning the semantic-syntactic information of civil aviation short texts by combining point-wise convolution with graph convolutional networks (GCN) and multi-head attention mechanisms.Then, a fully connected layer is employed to amalgamate the acquired information for outputting classification results. Finally, experiments conducted on aviation datasets and other public domain datasets demonstrate that the ESS-PWGCN model not only surpasses the current state-of-the-art self-training text graph convolution networks(ST-TextGCN) model in terms of accuracy and F_1 score by 4.59% and 6.53%, respectively, but also exhibits superior robustness and generalizability.

Keywords: short text classification; deep learning; graph convolutional networks; pointwise convolution; attention mechanisms; long short-term memory networks

Received: 2024-04-16; **Accepted**: 2024-07-12; **Published Online**: 2024-08-13 12:34

URL: link.cnki.net/urlid/11.2625.V.20240812.1643.002

Foundation item: Natural Science Foundation of Tianjin, China (17JCYBJC18200)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0228

基于改进 YOLOv7 的风机叶片缺陷检测

汤占军[1]，张朝杰[1]，王健[2,*]，陆鹏[3]，刘汇塬[2]，蹇洪[1]

(1. 昆明理工大学 信息工程与自动化学院，昆明 650500；2. 昆明理工大学 电力工程学院，昆明 650500；
3. 云南龙源新能源有限公司，昆明 650228)

摘　　要：为实现对风力发电机叶片缺陷的实时精准检测，针对风机叶片缺陷尺度变化大、背景复杂的特点，提出一种基于 YOLOv7 网络模型的风机叶片缺陷检测模型 EPW-YOLOv7。设计 CSM 注意力模块并添加于主干网络中，以抑制复杂背景干扰，提高重要特征的提取效率；设计轻量化的 PWK 模块替换原有的高效层聚合网络 (ELAN) 模块，减少冗余参数量及计算量，加快网络的检测速度；引入双向特征金字塔网络 (BiFPN) 特征融合模块，促使网络更精确地识别多尺度缺陷特征；采用 Wise-交并比 (WIoU) 损失函数优化网络，提高 EPW-YOLOv7 模型的整体检测性能。在风机叶片缺陷数据集上进行实验，实验结果表明：EPW-YOLOv7 模型的平均准确率可达 88.4 %，相较于 YOLOv7-tiny 模型提高了 7.1 %，且帧率达到 66 帧/s，满足实时检测需求。此外，与当前先进的目标检测算法相比，EPW-YOLOv7 模型在风机叶片缺陷的检测精度和速度上更具优势，证明所提模型更适用于风机叶片缺陷实时检测定位任务。

关　键　词：风机叶片；缺陷检测；深度学习；注意力机制；特征融合
中图分类号：TP391.4; V267
文献标志码：A　　**文章编号**：1001-5965(2026)06-1903-12

近年来，随着全球能源需求不断增长，传统能源结构转型日益迫切，风力发电以其清洁、可再生的优势，迅速崛起为全球能源供应的重要力量[1]。风机叶片作为风力发电机组的关键部件，其运行状态对发电效率具有显著的影响。然而，由于风力发电机组通常部署在偏远且环境恶劣的地区，长期受到极端天气条件的挑战，风机叶片极易出现各种缺陷[2]。因此，对风机叶片进行定期、高效的巡查与检测，及时发现并消除潜在隐患，对于保障风力发电机的安全稳定运行、降低运维成本及减少发电量损失具有至关重要的意义。

目前，风机叶片故障检测主要依赖于人工巡检的方式。然而，随着大型风电机组的快速发展，风机容量的不断增加使得风机叶片长度逐渐增大，这进一步加大了叶片缺陷检测的难度，导致人工巡检的检测效率显著降低。为克服这一挑战，近年来有学者提出其他缺陷检测技术。如 Lei 等[3] 通过利用端到端的长短时记忆神经网络对采集到的振动信号进行分类，实现了对叶片故障的有效检测；Choung 等[4] 提出一种自适应超声相控阵系统用于检测风电机组叶片的内部缺陷；邹宜金等[5] 通过结合聚类和中值收敛的周期性音频切割方法，对风机叶片声纹进行有效的切割，随后采用稳态差异法检测故障；Xu 等[6] 基于小波包分解对风机叶片波形进行特征提取，并根据提取的特征进行损伤模式识别和奇异信号检测的聚类分析，以检测叶片是否存在故障；Tian 等[7] 提出一种基于光纤布拉格光栅的非基线损伤检测模型，以准确检测风力发电机叶片

收稿日期：2024-04-17；**录用日期**：2024-08-03；**网络出版时间**：2024-08-07 15:32
网络出版地址：link.cnki.net/urlid/11.2625.V.20240807.1253.001
基金项目：国家自然科学基金 (82160347)；云南省基础研究专项-青年项目 (202401AU070148)；国家能源集团科技项目 (CSIEKJ230700104)
* **通信作者**. E-mail：Jianwangzx@163.com

引用格式：*汤占军，张朝杰，王健，等. 基于改进 YOLOv7 的风机叶片缺陷检测 [J]. 北京航空航天大学学报，2026，52（6）：1903-1914.*
TANG Z J，ZHANG C J，WANG J，et al. Wind turbine blade defect detection based on improved YOLOv7[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（6）：1903-1914（in Chinese）.

缺陷；Hwang 等[8]则通过分析激光激励后热波在风机叶片内部的传导情况，对风机叶片内部分层和表面损伤进行了简单区分。尽管上述检测方法在一定程度上缓解了人工巡检低效的问题，但其仍面临采集数据处理难度大、传感器等设备维护成本高、无法有效检测早期的小缺陷等挑战。因此，进一步研究和开发高效、准确且经济的风机叶片缺陷检测技术依然具有迫切的需求和重要的实践意义。

随着计算机视觉技术的不断发展，其逐渐被学者应用于风机叶片缺陷检测领域，如 Deng 等[9]利用 Haar-like 算子提取叶片裂纹特征，并以此训练得到一个用于裂纹检测的级联分类器；徐灵鑫[10]采用灰度插值、灰度增强、Robert 算子和形态学分析共同作用去除图像噪声等一系列方法对风机叶片图像进行预处理，随后通过 8 连通链码法提取每个缺陷的外轮廓特征进行缺陷标记，最后设计了融合多特征参数的叶片表面缺陷模式识别分类器以检测不同类别的风机叶片缺陷；Wang 等[11]首先通过对风机叶片图像依次进行灰度化、图像插值、图像去噪等预处理，接着，采用基于局部粒子群优化（local particle swarm optimization, LPSO）算法改进的 Log-Gabor 滤波器对边缘特征进行提取，最后，使用方向梯度直方图特征结合支持向量机的方法完成了对不同风机叶片缺陷的分类。虽然以上传统的目标检测算法在一些简单场景下展现出较好的检测效果，然而，其仍存在因手工设计特征导致的检测精度受限、实时性较差及背景越复杂识别效果下降越明显等问题。

相较于传统的目标检测算法，基于深度学习的目标检测算法凭借强大的特征提取能力和泛化能力，使其在目前缺陷检测领域取得了较好的实验结果并得到广泛应用，如 Reddy 等[12]利用无人机图像对卷积神经网络进行训练，实现了对风机叶片缺陷的分类检测；Zhang 等[13]提出 Mask-MRNet 目标检测模型网络，该网络通过将基于掩码区域的卷积神经网络（mask region-based convolutional neural networks, Mask RCNN）与 MRNet 进行堆叠，显著减少了推理时间，从而提高了风机叶片缺陷检测效率；Lv 等[14]设计一种高效准确的损伤检测器，该检测器以单激发多框探测器 (single shot multibox detector, SSD) 算法为基础，采用增强的残差网络作为骨干网络，有效提升了风机叶片缺陷的检测精度；Hu 等[15]提出 YOLOx-Nano 目标检测模型，该模型在保持较快推理速度的同时，显著提高了风机叶片缺陷的检测精度；Liu 等[16]通过采用卷积层滤波器剪枝和知识提炼等模型压缩方法，减小了 YOLOv3 模型的体积，使其检测速度近乎翻倍；Zhang 等[17]提出 SOD-YOLO 模型，其在 YOLOv5 模型的基础，利用 K-means 算法重新聚类锚点，并在各特征融合层中引入了卷积块注意力模块（convolutional block attention module, CBAM），从而能够快速有效地检测小目标缺陷和其他缺陷；Hu 等[18]提出 YOLO-Tiny-attention 模型，该模型通过结合空间-通道注意力模块，以及额外添加检测头，使网络能够更高效地检测多尺度缺陷目标。上述基于深度学习的目标检测算法虽在一定程度上提高了风机叶片缺陷检测效率，然而，其仍存在一些不足，如部分改进模型虽然模型体积降低，但计算量依然较大，不利于嵌入式部署；部分改进模型不能很好地实现检测精度与速度之间的平衡。

针对风机叶片缺陷检测任务实时高效且易于嵌入式部署的实际检测需求，本文在 YOLOv7 的基础上，提出了 EPW-YOLOv7 网络模型，该模型首先设计并引入了 CSM 注意力机制，通过合理分配权重，提高模型对于重要缺陷特征的提取能力；其次，设计 PWK 卷积模块替换原模型 Backbone 部分高效层聚合网络（efficient layer aggregation networks, ELAN）模块，以减少模型参数量及计算量，使模型更易于实现嵌入式部署；接着，采用双向特征金字塔网络（bidirectional feature pyramid networks, BiFPN）替换路径聚合网络（path aggregation networks, PANet），促使多尺度特征信息更好地融合，提高模型检测精度；最后，引入 Wise-交并比（wise-intersection over union, WIoU）损失函数替换完整交并比（complete-intersection over union, CIoU）损失函数，加快模型训练过程中的收敛速度，提升模型整体性能。在风机叶片缺陷数据集上进行实验，实验结果表明，本文改进模型可以快速、准确地检测出风机叶片缺陷。

1 EPW-YOLOv7 网络模型

1.1 YOLOv7-Tiny 模型

YOLO 系列模型是当前目标检测领域的典型算法，其凭借较快的检测速度而被广泛应用于缺陷检测、机器人视觉、无人驾驶等领域[19]。YOLOv7 模型[20]作为 YOLO 系列的基本模型，在帧率为 5～160 帧/s 范围内，其检测精度及推理速度超越了其他 YOLO 系列模型。因此，本文针对风力发电机叶片缺陷的巡检需求，选用 YOLOv7 模型中的轻量化版本 YOLOv7-Tiny 作为基础模型，并在此基础上提出了 EPW-YOLOv7 网络模型。

YOLOv7-Tiny 模型由输入（Input）、骨干网络（Backbone）、颈部网络（Neck）、头部网络（Head）这

4 部分组成, 如图 1 所示, 其中, Input 部分沿用了 YOLOv5 模型结构, 采用 Mosaic 数据增强技术对输入图片进行预处理; Backbone 部分则主要用于目标特征信息的提取, 相较于 YOLOv5 模型, 其创新性引入了 ELAN 模块和最大池化卷积(max pooling convolution, MPConv)模块。ELAN 模块通过数个卷积层、归一化层及 SiLU 激活函数层的多分支堆叠, 有效控制了梯度路径, 使深层网络能够学习到更丰富的有效特征。MPConv 模块结合了最大池化层与多个卷积层, 形成双分支结构, 以确保模型进行下采样操作时, 通道数保持不变; Neck 及 Head 部分主要用于融合 Backbone 层的输出特征, 并基于融合后的特征生成先验框及类别预测。Neck 部分沿用 PANet 结构, 同时引入了重参数化卷积结构模块, 以降低模型推理时的复杂度, 从而提高模型的检测效率。Head 部分在 YOLOv5 模型的基础上引入了辅助检测头, 通过在网络中增加额外的输出层, 获取更多的多尺度特征信息, 进而提升了模型检测的准确性。为不影响模型的推理速度, 该辅助检测头只出现在训练过程中。

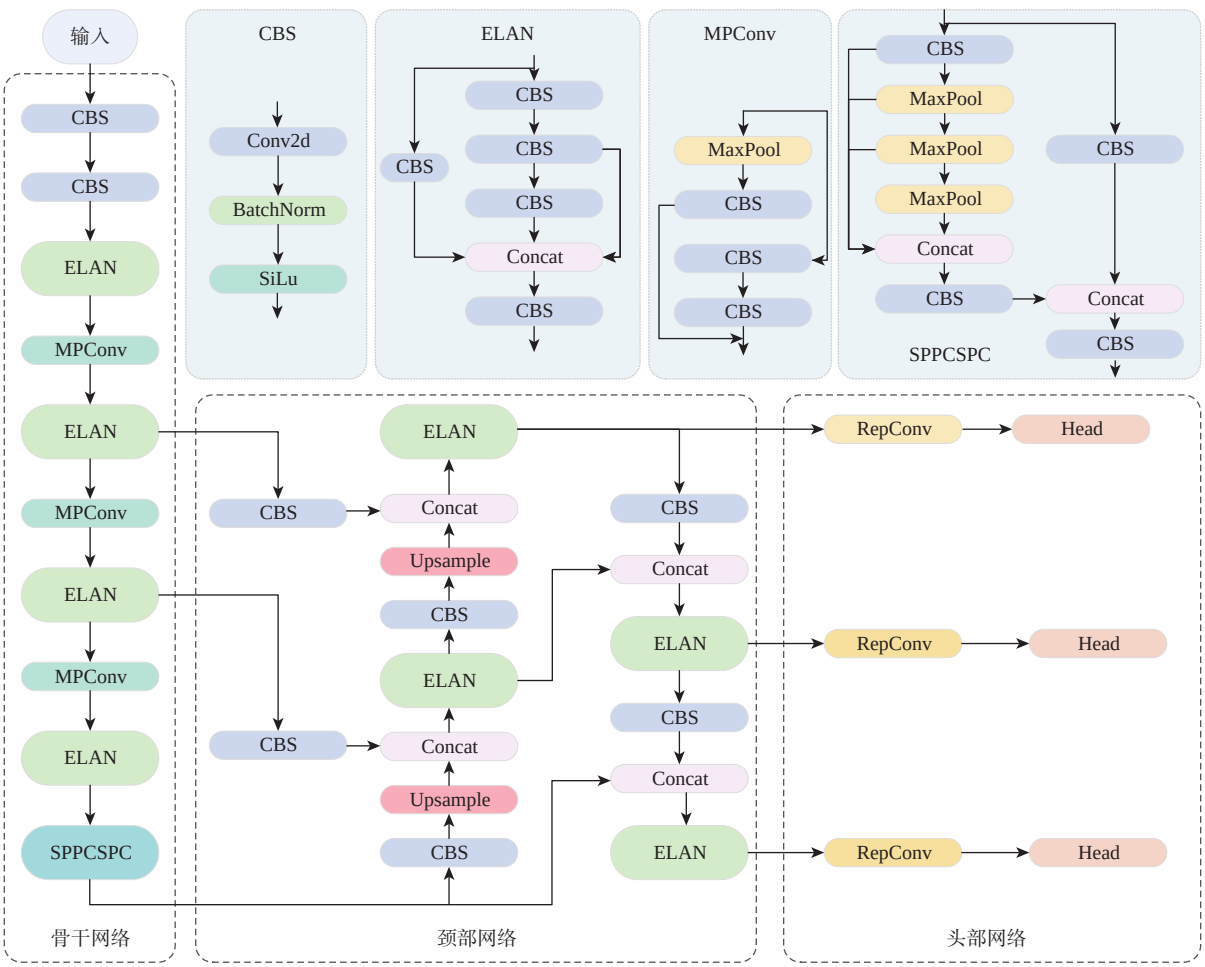


图 1　YOLOv7-Tiny 模型结构

Fig. 1　YOLOv7-Tiny model structure

1.2　YOLOv7-Tiny 模型改进

针对风机叶片缺陷尺寸变化大、背景复杂的特点, EPW-YOLOv7 在保证实时性检测的前提下, 实现了模型检测精度的提高, 其结构如图 2 所示。

1.2.1　CSM 注意力模块

为提高 EPW-YOLOv7 模型特征提取能力, 本文在 YOLOv7 的 Backbone 网络部分中插入了 CSM 注意力模块。该 CSM 模块由通道注意力模块(channel attention module, CAM)、空间注意力模块(spatial attention module, SAM)及多头自注意力(multi-head self-attention, MHSA)模块[21]这 3 部分组成。其中, 通道注意力模块, 如图 3 所示, 通过捕获特征不同通道间的交互信息, 生成通道注意力权重, 以此使网络能够聚焦于具有较高重要性的通道, 同时抑制不重要的通道, 进而提高特征的表征能力和判别能力。

CAM 首先对输入大小为 $H\times W\times C$ 的输入特征图 $\boldsymbol{X}$ 进行平均池化操作与最大池化操作; 然后, 将池化后得到的 2 个一维向量送入一维卷积层进行计算; 最后, 经 Sigmoid 函数激活处理生成通道注意

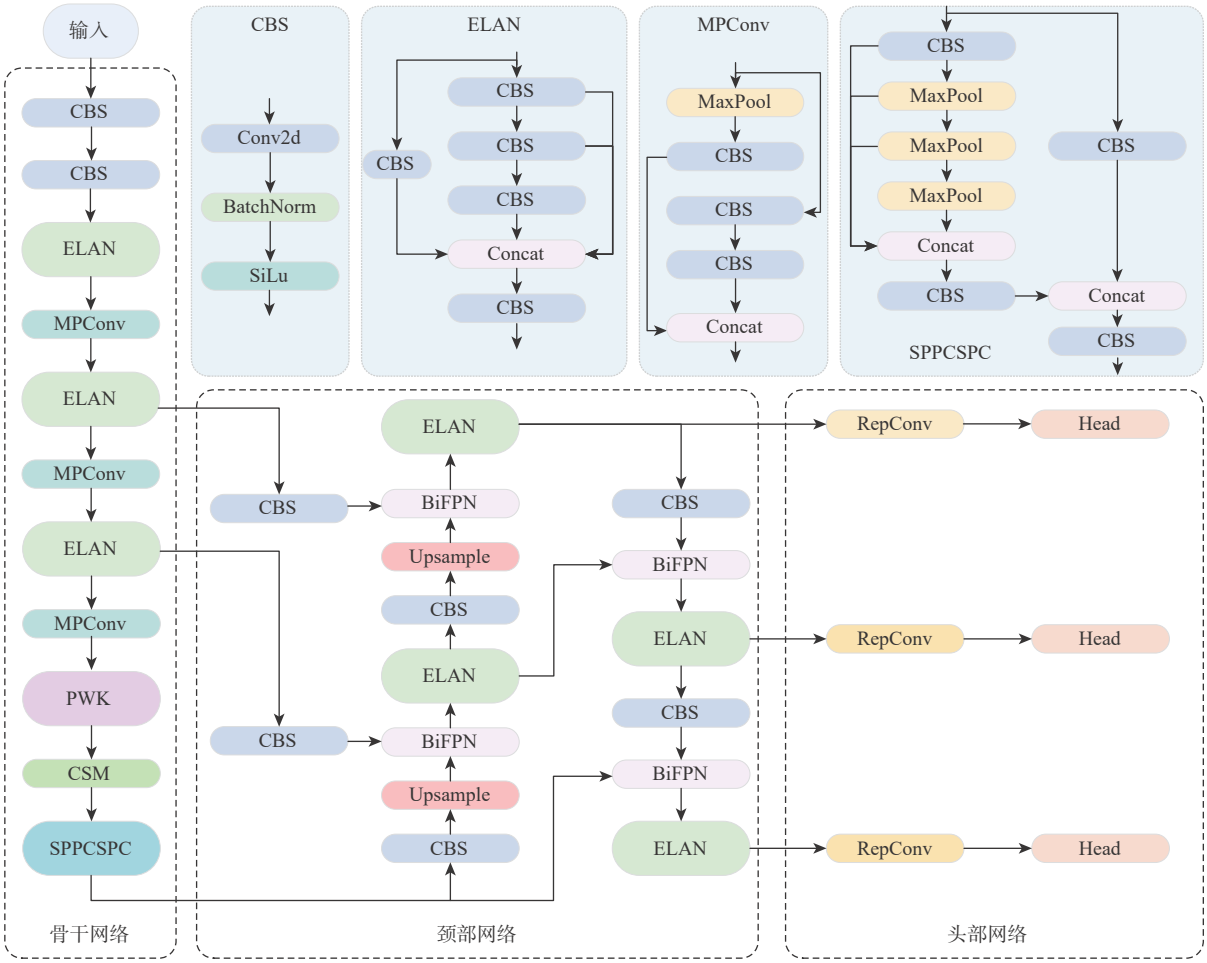


图 2 EPW-YOLOv7 模型结构

Fig. 2 EPW-YOLOv7 model structure

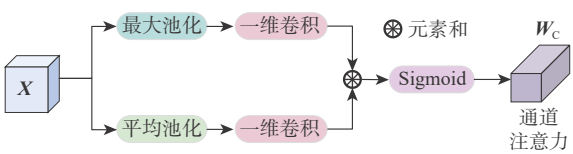


图 3 CAM 模块结构

Fig. 3 CAM module structure

力权重$\boldsymbol{W}_{\mathrm{C}}$，其定义如下：

$$\boldsymbol{W}_{\mathrm{C}}=\mathrm{Sigmoid}(\mathrm{Conv1d}(\mathrm{AvgPool}(\boldsymbol{X}))+\mathrm{Conv1d}(\mathrm{MaxPool}(\boldsymbol{X}))) \quad (1)$$

式中：$\boldsymbol{X}$为输入特征图；AvgPool为平均池化操作；MaxPool为最大池化操作；Conv1d为一维卷积层，该一维卷积的卷积核大小为k，其值代表局部跨通道交互的覆盖率，可通过交叉验证进行确定；Sigmoid(·)为激活函数；$\boldsymbol{W}_{\mathrm{C}}$为CAM输出权重。该CAM未采用全连接层去捕获所有通道间的交互信息，而采用一维卷积捕获相关性更大的局部通道间的交互信息，避免了使用全连接层所产生的特征信息丢失及大量的参数计算等副作用，从而使网络表达能力得到提高的同时，降低了模块的计算成本。

空间注意力模块通过捕获输入特征图不同空间位置之间的交互信息，生成空间注意力权重，以此使网络能够聚焦于具有较高重要性的空间位置，同时抑制不重要的位置，进而提高模型对于目标的定位精度和区分能力，该SAM结构如图4所示。

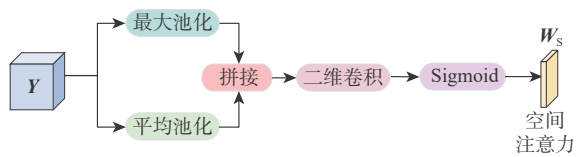


图 4 SAM 结构

Fig. 4 SAM structure

SAM首先对特征图$\boldsymbol{Y}$进行平均池化操作与最大池化操作；然后，将池化后得到的2个二维向量送入7×7的二维卷积层计算；最后，经Sigmoid函数激活处理生成空间注意力权重$\boldsymbol{W}_{\mathrm{S}}$，其定义如下：

$$\boldsymbol{W}_{\mathrm{S}}=\mathrm{Sigmoid}(\mathrm{Conv2d}(\mathrm{AvgPool}(\boldsymbol{Y}),\mathrm{MaxPool}(\boldsymbol{Y}))) \quad (2)$$

式中：Conv2d为7×7二维卷积；$\boldsymbol{Y}$为基于通道注意力权重加权生成的特征图。

多头自注意力模块由多个如图5所示的自注意力层构成，其通过并行地捕获不同区域之间特征

的依赖信息, 生成自注意力权重, 以此使网络能够更好地整合上下文信息, 进而提高模型在复杂背景中的目标检测精度。

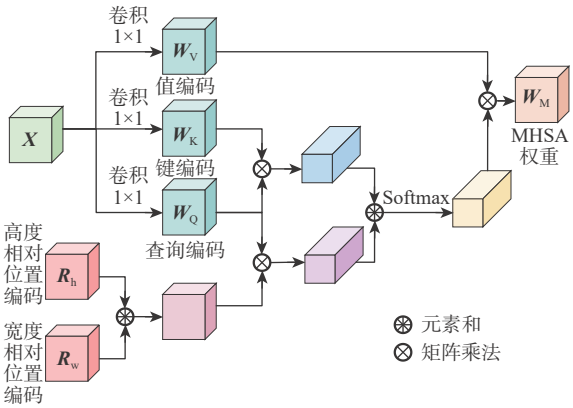


图 5　MHSA 模块结构

Fig. 5　MHSA module structure

MHSA 模块首先将输入特征图 $\boldsymbol{X}$ 进行线性变换得到 $\boldsymbol{W}_{\mathrm{Q}}$、$\boldsymbol{W}_{\mathrm{K}}$、$\boldsymbol{W}_{\mathrm{V}}$、$\boldsymbol{R}_{\mathrm{h}}$、$\boldsymbol{R}_{\mathrm{w}}$ 这 5 个矩阵; 然后, 利用这些矩阵计算出点积注意力矩阵 $\boldsymbol{H}$, 并行重复以上过程 i 次(i 为 MHSA 模块头数); 最后, 将 i 个头的点积注意力矩阵进行拼接, 生成多头自注意力权重 $\boldsymbol{W}_{\mathrm{M}}$, 其定义如下:

$$\boldsymbol{W}_{\mathrm{M}}=\mathrm{Concat}(\boldsymbol{H}_1,\boldsymbol{H}_2,\cdots,\boldsymbol{H}_i)\qquad i=1,2,\cdots,n \tag{3}$$

$$\boldsymbol{H}=\mathrm{Softmax}(\boldsymbol{W}_{\mathrm{Q}}\times\boldsymbol{W}_{\mathrm{K}}+\boldsymbol{W}_{\mathrm{Q}}\times(\boldsymbol{R}_{\mathrm{h}}+\boldsymbol{R}_{\mathrm{w}}))\boldsymbol{W}_{\mathrm{V}} \tag{4}$$

式中: $\boldsymbol{W}_{\mathrm{Q}}$ 为查询编码; $\boldsymbol{W}_{\mathrm{K}}$ 为键编码; $\boldsymbol{W}_{\mathrm{V}}$ 为值编码; $\boldsymbol{R}_{\mathrm{h}}$ 为高度相对位置编码; $\boldsymbol{R}_{\mathrm{w}}$ 为宽度相对位置编码; Softmax(·)为归一化指数函数; Concat 为拼接操作。

另外, 为进一步提高模型对有效特征信息的提取能力, 削弱无关背景信息的影响, 本文提出了 CSM 结构, 如图 6 所示, 该结构采用拼接的方式将 CAM 和 SAM 串行连接后输出的特征图与 MHSA 输出特征图结合, 使三者产生的特征信息能够互补地融合在一起, 以此生成了基于多维度注意力权重加权的输出特征图, 其定义如下:

$$\boldsymbol{Z}_{\mathrm{out}}=\mathrm{Concat}(\boldsymbol{W}_{\mathrm{S}}\otimes(\boldsymbol{W}_{\mathrm{C}}\times\boldsymbol{X}),\boldsymbol{W}_{\mathrm{M}}\otimes\boldsymbol{X}) \tag{5}$$

式中: $\boldsymbol{Z}_{\mathrm{out}}$ 为模块最终输出特征图; ⊗为加权操作。CSM 注意力模块采取将 CAM、SAM 串行连接, 而 MHSA 模块与两者并行连接的方式, 使得模型不但能够从空间和通道维度获取到更多的有效特征信息, 而且弥补了 CAM、SAM 在处理长距离依赖信息时存在的潜在局限性问题, 借此进一步加强模型对整个特征图信息的全局理解, 提高模型的泛化能力。CSM 注意力模块作为轻量高效的模块可以无缝插入到 YOLOv7 网络结构中, 进行端到端的训练。

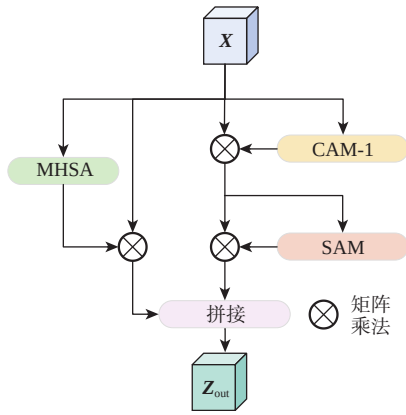


图 6　CSM 模块结构

Fig. 6　CSM module structure

1.2.2　PWK 卷积模块

ELAN 模块虽然通过堆叠大量的卷积层显著提高了特征提取能力, 但同时其也导致了参数量及计算量的增加, 进而使得模型的检测效率降低。针对该问题, 本文提出了一种轻量化的 PWK 卷积模块, 用以替代模型中的部分 ELAN 模块, 以使模型在保证检测精度的前提下, 能够实现检测速度的提高和更易于嵌入式部署。

PWK 卷积模块结构如图 7 所示, 通过在 ELAN 模块的基础上引入部分卷积(partial convolution, PConv)[22]、深度膨胀卷积 (depthwise dilation convolution, DW-D-Conv)[23]、逐点卷积(pointwise convolution, PWConv)替代 ELAN 模块中部分 3×3 卷积, 使得模型能够获取到局部通道信息和空间维度的适应性。其中, PConv 只对部分输入通道进行特征提取, 其余通道保持不变, 以此可以更有效地提取局部空间特征信息; 深度膨胀卷积通过扩大卷积核大小, 获得更大的感受野, 以此捕捉全局空间特征信息; PWConv 通过 1×1 卷积对所有输入通道进行聚合, 以此获取通道维度特征信息。

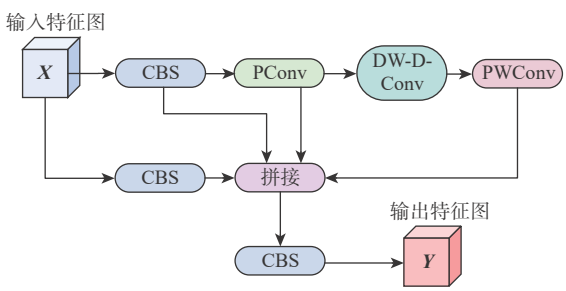


图 7　PWK 卷积模块结构

Fig. 7　PWK convolution module structure

PWK 卷积模块通过采用轻量化的卷积模块减少输入通道数量, 有效降低了整体模块的参数量及计算量, 同时实现了长距离依赖及多维度的适应性。因此, 本文结合风机叶片缺陷存在小目标缺陷

及细长缺陷等特点，采用PWK卷积模块替代部分ELAN模块能够在有效提升模型的检测性能的同时，使得模型更加轻量化。

1.2.3 BiFPN特征融合模块

由于主干网络在特征提取时存在尺度差异，这使得浅层大尺度特征图通常分辨率更高，有利于目标精确定位，而深层小尺度特征图具有更丰富的语义信息，有利于识别目标。因此，为提高对于不同尺度目标的检测能力，目标检测模型往往会通过特征融合网络对多尺度特征信息进行整合。在YOLOv7模型中，采用PANnet作为特征融合结构，其在特征金字塔网络（feature pyramid networks, FPN）基础上额外添加自下而上的路径聚合网络，实现了浅层较强定位信息和深层较强语义信息的有效融合。然而，PANet虽然相较于FPN而言，获得的特征信息更为丰富，但同时也带来了更多的参数量和计算量，进而降低了模型的检测速度，此外，PANet在融合不同分辨率的特征要素时，采取了将其调整为相同分辨率并进行简单相加的方法，未充分考虑不同分辨率的输入特征对输出特征的贡献不相等的问题，这在模型进行多尺度目标检测时可能产生不利影响。针对以上问题，Tan等[24]对PANnet进行改进，提出了BiFPN特征融合网络，具体结构对比如图8所示。

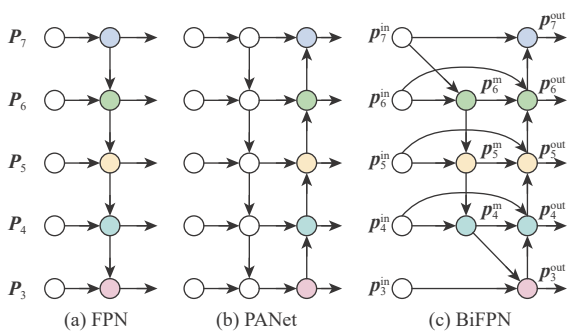


图8 特征融合网络结构对比[24]

Fig. 8 Comparison of feature fusion network structure[24]

BiFPN在PANnet的基础上首先移除对融合不同特征贡献较小的节点，简化双向网络结构；接着，对位于同一水平的输入节点和输出节点之间额外增加一条输入边，用少量的成本增加换取更多的特征融合；最后，采用加权特征融合机制，保留更有效的特征信息。BiFPN节点定义如下：

$$\boldsymbol{p}_i^{\mathrm{m}}=\mathrm{Conv}\left(\frac{w_1\boldsymbol{p}_i^{\mathrm{in}}+w_2\mathrm{Res}\left(\boldsymbol{p}_{i+1}^{\mathrm{in}}\right)}{w_1+w_2+\epsilon}\right) \tag{6}$$

$$\boldsymbol{p}_i^{\mathrm{out}}=\mathrm{Conv}\left(\frac{w_1\boldsymbol{p}_i^{\mathrm{in}}+w_2\boldsymbol{p}_i^{\mathrm{m}}+w_3\mathrm{Res}\left(\boldsymbol{p}_{i-1}^{\mathrm{out}}\right)}{w_1+w_2+w_3+\epsilon}\right) \tag{7}$$

式中：$\boldsymbol{p}_i^{\mathrm{in}}$为自下而上第$i$层的输入特征层；$\boldsymbol{p}_i^{\mathrm{m}}$为第$i$层的输出特征层；$\boldsymbol{p}_i^{\mathrm{out}}$为自下而上第$i$层的输出特征层；$w_1$、$w_2$、$w_3$为可学习权重；Conv为卷积操作；Res为采样操作；ϵ为0.000 1。

本文通过引入BiFPN特征融合网络替代PANet，使得模型能够在不显著增加计算成本的情况下融合更多有效特征，同时缓解了多尺度目标缺陷特征信息丢失对于模型检测精度的影响。

1.2.4 WIoU损失函数

在目标检测任务中损失函数常用于判断训练数据与实际数据之间是否有差距，合适的损失函数可以使模型在训练时更快收敛，有利于提升模型的检测性能。原始YOLOv7模型采用CIoU损失函数作为边界框损失函数，其定义如下：

$$L_{\mathrm{CIoU}}=1-I+\frac{\rho^2(b,b^{\mathrm{gt}})}{d^2}+\alpha\nu \tag{8}$$

$$\alpha=\frac{\nu}{(1-I)+\nu} \tag{9}$$

$$\nu=\frac{4}{\pi^2}\left(\arctan\frac{w^{\mathrm{gt}}}{h^{\mathrm{gt}}}-\arctan\frac{w}{h}\right)^2 \tag{10}$$

式中：I为预测边界框与真实边界框的交并比；b和b^{gt}分别为预测边界框和真实边界框的中心点；ρ为2个中心点之间的距离；d为同时包含2个框的最小外接框的对角线距离；α为权重函数；ν为纵横比的相似度；w^{gt}、h^{gt}、w、h分别为真实边界框和预测边界框的宽和高。CIoU通过构造惩罚项将2个中心点的距离归一化，加速预测框收敛，这使得回归时其能够更好地描述重叠信息，但其存在长宽比的惩罚项在预测框和真实框的长宽比相同时恒为0、预测框的高和宽不能同时增大或减小等不合理的情况，以及因几何因素导致的低质量训练样本惩罚加重的问题。因此，本文采用WIoU损失函数[25]代替CIoU损失函数，其定义如下：

$$L_{\mathrm{WIoUv3}}=\frac{\beta}{\delta\alpha^{\beta}}R_{\mathrm{WIoU}}L_I \tag{11}$$

$$R_{\mathrm{WIoU}}=\exp\left(\frac{(x-x^{\mathrm{gt}})^2+(y-y^{\mathrm{gt}})^2}{(W_{\mathrm{g}}^2+H_{\mathrm{g}}^2)^*}\right) \tag{12}$$

$$L_I=1-I \tag{13}$$

式中：x^{gt}和y^{gt}为真实边界框中心点；W_{g}和H_{g}分别为同时包含预测框的和真实框的最小外接框的宽和高；上标*表示分离操作，即将W_{g}、H_{g}从计算图中分离出；δ和α为超参数；β为离群度。

L_{WIoUv3}损失函数基于距离度量构建出距离注意力，其根据锚框和目标框重合程度不同，调整几何度量的惩罚，解决了CIoU长宽比在个别情况下恒为0的问题；同时通过分离最小外接框的高和宽，

削弱了几何因素对于低质量训练样本惩罚的影响。此外, L_{WIoUv3} 定义了离群度β, 其值与分配的梯度增益大小成正比, 若β值越小, 则锚框质量越高, 分配的梯度增益越小, 以此使边界框回归聚焦到普通质量的锚框上, 反之, 若β值越大, 则锚框质量越低, 分配的梯度增益越大, 借此防止低质量锚框产生有害梯度。β具体定义为

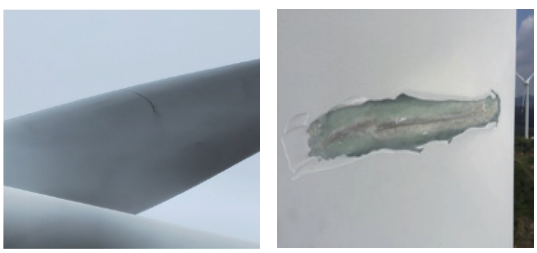
$$\beta = L_I^* / \overline{L_I} \tag{14}$$

式中: L_I^*为单调聚焦系数; $\overline{L_I}$为动量 m 的滑动平均值。其中, L_I作为动量, 使得β值大小始终在动态更新, 锚框质量的划分标准在不停调整, 这意味着 L_{WIoUv3} 能根据当前情况制定出对应的策略, 从而实现对梯度增益的动态分配。L_{WIoUv3} 运用动态非单调分配机制对梯度增益动态调整, 使模型聚焦于普通质量锚框, 提高了模型的泛化能力。

2　实验与分析

2.1　风机叶片缺陷数据集

本文首先通过无人机拍摄等方式获取了西南某风力发电场风机叶片原始图像 604 张, 如图 9(a)所示。接着, 为提高实验鲁棒性, 在原始图像的基础上, 对已有图像采用 Mixup、Mosaic 等数据增强技术扩充样本数, 如图 9(b)所示。最后, 利用 LabelImg 标注软件对已有样本进行标注, 构建了包含裂缝缺陷、表皮脱落缺陷等常见风机叶片缺陷的风机叶片缺陷数据集。该数据集包含 2 345 张图片, 其中, 缺陷图片 1 501 张, 无缺陷图片 844 张。将风机叶片缺陷数据集按 8:1:1 的比例随机划分训练集、验证集和测试集。



(a) 风机叶片缺陷样本

(b) 数据增强风机叶片缺陷图像

图 9　风机叶片缺陷数据集

Fig. 9　Dataset of wind turbine blade defects

2.2　评价指标

根据风机叶片缺陷检测实时精确的检测需求, 本文采用每秒帧数对模型的检测速度进行评估; 采用浮点运算速度及模型参数量对模型嵌入式部署能力进行评价; 采用准确率 P(precision), 召回率 R(recall), 平均精度均值(mean average precision, mAP)对检测精度进行评估, 一般在 IoU=0.5:0.95 的阈值条件下计算 mAP 值, 即 mAP_0.5, 具体定义如下:

$$P = \frac{T_P}{T_P + F_P} \tag{15}$$

$$R = \frac{T_P}{T_P + F_N} \tag{16}$$

$$A_P = \int_0^1 P(R)\mathrm{d}R \tag{17}$$

$$m_{AP} = \frac{\sum A_P}{N} \tag{18}$$

式中: T_P 为正样本被正确预测为正样本; F_P 为负样本被错误预测为正样本; F_N 为正样本被错误预测为负样本; A_P 为平均精确度; $P(R)$为基于准确率和召回率的曲线; N 为所有类别数量。

2.3　实验环境搭建及网络训练

本文实验硬件配置采用 Intel(R) Xeon(R) Platinum 8255C CPU 和 NVIDIA GeForce RTX 3 080(10 GB) GPU, 软件环境配置采用 Windows10 操作系统, 编程语言为 Python 3.8, 使用 Pytorch 1.8.1 深度学习框架构建模型, CUDA 为 11.1。模型训练时的初始学习率为 0.01, 训练轮数 epoch 为 400, 批处理大小为 16, 学习率动量为 0.937, 权重衰减系数为 0.000 5, 使用 SGD 优化器对学习率进行调整。

在相同实验条件下, EPW-YOLOv7 及 YOLOv7-tiny 训练曲线对比如图 10 所示。由图 10 可知, EPW-YOLOv7 采用 WIoU 损失函数替代了 CIoU 损失函数作为边界框回归损失函数, 相较于 YOLOv7-tiny, EPW-YOLOv7 收敛速度更快。同时由 mAP_0.5 对比曲线可以看到, EPW-YOLOv7 模型明显具有更高的检测精度, 验证了本文模型的可行性。

2.4　实验结果分析

2.4.1　注意力模块对比实验

为体现本文 CSM 的优越性, 在相同实验条件下, 选择同一风机叶片缺陷数据集, 将 CSM 及其他主流注意力模块分别引入 YOLOv7-tiny 模型中进行添加注意力模块对比实验。实验结果如表 1 所示, 最优结果加粗显示。

通过在 YOLOv7-tiny 模型中添加 CSM 注意力

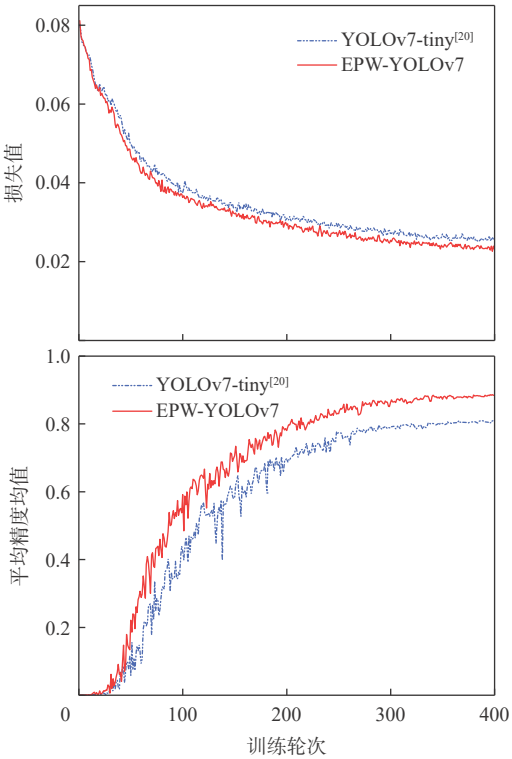


图 10 模型训练曲线对比

Fig. 10 Comparison of model training curves

模块可以在风机叶片缺陷数据集上实现 83.8% 的 mAP_0.5, 相较于添加其他注意力模块的对比模型, 其在同一数据集中表现最好。这不但证明了 CSM 注意力模块的有效性, 还凸显了该模块出色的特征提取能力。

表 1 添加注意力模块对比实验

Table 1 Comparison experiment of adding attention modules

模型	mAP_0.5/%	浮点运算速度/10^9 s^{-1}	模型参数量
YOLOv7-tiny[20]	81.2	**13.2**	**6.02×10^6**
YOLOv7-tiny+CBAM	83.1	**13.2**	6.03×10^6
YOLOv7-tiny+MHSA	83.2	13.8	6.80×10^6
YOLOv7-tiny+ECA	82.3	13.3	6.03×10^6
YOLOv7-tiny+SimAm	82.4	**13.2**	**6.02×10^6**
YOLOv7-tiny+CSM	**83.8**	13.4	6.22×10^6

2.4.2 消融实验

为验证本文模型对于风机叶片缺陷检测网络性能的影响, 本文以 YOLOv7-tiny 为基线模型进行消融实验。该实验将 YOLOv7-tiny 网络作为基线模型, 称为模型 1; 将 CSM 注意力模块加入到模型 1 主干网络中, 得到的模型称为模型 2; 将 PWK 卷积模块加入到模型 2 主干网络中, 得到的模型称为模型 3; 将模型 3 的损失函数 CIoU 替换为 WIoU 损失函数, 得到的模型称为模型 4; 将 BiFPN 特征融合模块替换模型 4 原有的 PANet 特征融合模块, 得到的模型称为模型 5, 其实验结果如表 2 所示, 最优结果加粗显示。具体实验步骤如下:

表 2 消融实验

Table 2 Ablation experiment

模型编号	CSM	PWK	WIoU	BiFPN	mAP_0.5/%	帧率/(帧·s^{-1})	浮点运算速度/10^9 s^{-1}	模型参数量
1					81.2	64	13.2	6.02×10^6
2	√				83.8	67	13.4	6.22×10^6
3	√	√			84.9	**71**	**12.7**	**5.22×10^6**
4	√	√	√		86.2	68	**12.7**	**5.22×10^6**
5	√	√	√	√	**88.4**	66	**12.7**	**5.22×10^6**

步骤 1 将 CSM 注意力模块引入主干网络中, 增加了从多维度获取特征信息的操作, 这虽导致了模型参数量及计算量的略微增加, 但也使模型 2 相对模型 1 的 mAP_0.5 提升了 2.6 % 且帧率为 67 帧/s, 在未影响检测速度的同时, 有效提高了检测精度。

步骤 2 采用 PWK 卷积模块替换 Backbone 中的部分 ELAN 模块, 通过将 ELAN 模块中运算量较大的 3×3 卷积替换为轻量化的深度卷积, 降低了整体模型的参数量及计算量, 同时提高了模型识别多尺度缺陷目标的准确率, 这使模型 3 相较于模型 2 在更轻量化的前提下, mAP_0.5 从 83.8 % 提高到了 84.9 % 且使帧率提升到了 71 帧/s。

步骤 3 采用 WIoU 损失函数替换 CIoU 损失函数, 通过采取高效的梯度增益分配策略, 使模型更关注普通质量的锚框, 有效提高了模型的泛化能力, 在模型参数量及计算量保持不变且检测速度未明显下降的情况下, 模型 4 相较于模型 3, mAP_0.5 提高了 1.3 %, 实现了模型整体性能的提高。

步骤 4 引入 BiFPN 特征融合模块替换 Neck 部分的 PANet 特征融合模块, 通过为不同尺度的特征分配权重, 增强了多尺度目标的检测能力, 这使模型 5 相较于模型 4 的 mAP_0.5 提高了 2.2 %, 达到 88.4 %。

最终得到的 EPW-YOLOv7 模型相较于原有 YOLOv7 模型在检测准确率及推理速度上均具有较大的优势, 具体检测效果对比如图 11 所示。

(a) 原始图片　(b) YOLOv7-tiny[20]　(c) EPW-YOLOv7

图 11　检测效果对比

Fig. 11　Comparison of detection effects

2.4.3　对比实验

为进一步验证本文模型在同类算法模型中的优越性, 本文在相同的配置环境和初始训练参数一致的条件下, 将改进模型与快速区域卷积神经网络(faster region-based convolutional neural networks, Faster RCNN)[26]、单激发多框目标检测模型 (single shot multibox detector, SSD)[27]、DETR(detection transformer)[28] 及 YOLO 系列[29-30] 模型进行对比实验, 实验结果如表 3 所示, 最优结果加粗显示。

表 3　对比实验

Table 3　Comparative experiments

模型	mAP_0.5/%	帧率/(帧·s^{-1})	浮点运算速度/10^9 s^{-1}	模型参数量
Faster-RCNN[26]	81.6	19	370.2	137.09×10^6
DETR[27]	78.5	42	42.0	52.01×10^6
SDD[28]	79.2	**104**	60.9	23.87×10^6
YOLOv3-tiny[29]	79.7	53	12.9	8.66×10^6
YOLOv5s[30]	80.4	81	15.8	7.02×10^6
YOLOv5l[30]	82.5	47	107.7	46.14×10^6
YOLOv5x[30]	84.2	45	203.8	86.10×10^6
YOLOv7[20]	83.4	46	36.5	103.80×10^6
YOLOv7-tiny[20]	81.2	64	13.2	6.02×10^6
EPW-YOLOv7	**88.4**	66	**12.7**	**5.22×10^6**

由表 3 可知, Faster RCNN 目标缺陷平均检测精度虽可达到 81.6%, 但其推理速度明显劣于其他模型, 且参数量及计算量过大, 不利于嵌入式部署; DETR 通过将 Transformer 结构整合到目标检测模型中, 借此结合全局信息, 从而更好地实现了端到端的目标检测, 但引入该结构同时也造成了整体模型针对小目标缺陷的检测效果较差, 这导致该模型的 mAP_0.5 仅为 78.5%, 低于其他目标检测模型; SDD 作为一阶段目标检测模型, 采用多尺度卷积特征图进行检测, 以此兼顾大目标及小目标缺陷的检测, 这使得该模型检测性能优于 Faster-RCNN、DETR 等模型; YOLOv3 模型在保持 YOLO 系列模型检测速度优势的前提下, 通过加强对小目标缺陷的识别能力, 从而使 mAP_0.5 达到 79.7%, 且帧率为 53 帧/s; YOLOv5 算法模型作为目前目标检测领域中较为流行的算法之一, 在 YOLOv3 的基础上使用了更多的卷积层及数据增强技术, 从而更好地实现了检测精度与检测速度的平衡。YOLOv5s 作为 YOLOv5 的轻量化版本, 有着较快的推理速度及较少的参数量、计算量, 然而, 由于其网络深度与 YOLOv5l、YOLOv5x 相比明显较低, 这使得部分目标缺陷的特征信息未能成功提取, 导致 YOLOv5s 在风机叶片缺陷数据集中的 mAP_0.5 仅为 80.4%, 略低于 YOLOv5l、YOLOv5x 的平均检测精度, 但其帧率为 81 帧/s, 明显优于其他大多数目标检测模型; YOLOv7 模型采用了更加高效的 ELAN 结构, 该结构通过控制最长和最短的梯度路径, 以此使模型能够获取到更多的目标缺陷特征信息, 这使得 YOLOv7 在该风机叶片缺陷检测任务中的检测效

率相较于其他目标检测模型更高；本文的EPW-YOLOv7模型，针对风机叶片缺陷的mAP_0.5达到了88.4%，明显高于其他目标检测模型，检测效果对比如图11所示，且其帧率为66帧/s，浮点运算速度为12.7 ×10^9/s及模型参数量为5.22×10^6，这使得该模型能够很好地满足风机叶片缺陷检测任务高实时性、高准确率、易嵌入式部署的需求。

2.4.4 泛化性实验

为进一步验证本文改进模型的有效性，本文在保持相同实验参数的前提下，采用东北大学NEU-DET钢材表面缺陷数据集进行泛化性实验。NEU-DET数据集包含：裂纹、夹杂物、斑块、点蚀表面、氧化皮和划痕等6种缺陷，共计包含1 800张图片。实验结果如表4所示，最优结果加粗显示。

表4 泛化性对比实验

Table 4 Generalization comparison experiment

模型	AP_0.5/%						mAP_0.5/%
	裂纹	夹杂物	斑块	点蚀表面	氧化皮	划痕	
YOLOv7-tiny[20]	42.6	76.8	89.4	94.3	53.4	81.2	72.9
EPW-YOLOv7	**47.7**	**82.2**	**92.5**	**95.4**	**63.6**	**86.4**	**77.9**

由表4可知，在NEU-DET钢材表面缺陷数据集中，EPW-YOLOv7模型平均检测精度为77.9%，相较于基准模型YOLOv7-tiny，其平均检测精度提升了5.0%，且各类缺陷的检测精度均实现了不同程度的提高，这充分证明EPW-YOLOv7模型具有较好的泛化性，适用于不同的缺陷检测数据集。

3 结 论

为实现风机叶片缺陷检测的实时精准检测，保证风力发电机安全高效运行，本文提出一种基于改进YOLOv7的风机叶片缺陷检测模型EPW-YOLOv7，主要结论如下：

1）在Backbone部分引入CSM注意力机制，经过捕捉多维度的特征信息，使网络更关注风机叶片缺陷特征；利用PWK模块替换Backbone中的部分ELAN模块，使整体模块的参数量及计算量降低；采用BiFPN特征融合模块替换FPN特征融合模块，使多尺度特征更好地融合；使用WIoU损失函数代替CIoU损失函数，使模型能够快速收敛。

2）本文模型的mAP_0.5可以达到88.4%，且帧率为66帧/s，可以实时准确检测风机叶片缺陷。相比Faster-RCNN、SDD、DETR、YOLOv3、YOLOv5s、YOLOv5l、YOLOv5x、YOLOv7-tiny，EPW-YOLOv7等模型平均检测精度最高，同时检测速度较快，满足风机叶片缺陷检测实际需求。

参考文献（References）

[1] ROGA S, BARDHAN S, KUMAR Y, et al. Recent technology and challenges of wind energy generation: a review[J]. Sustainable Energy Technologies and Assessments, 2022, 52: 102239.

[2] DU Y, ZHOU S X, JING X J, et al. Damage detection techniques for wind turbine blades: a review[J]. Mechanical Systems and Signal Processing, 2020, 141: 106445.

[3] LEI J H, LIU C, JIANG D X. Fault diagnosis of wind turbine based on long short-term memory networks[J]. Renewable Energy, 2019, 133: 422-432.

[4] CHOUNG J, LIM S, LIM S H, et al. Automatic discontinuity classification of wind-turbine blades using A-scan-based convolutional neural network[J]. Journal of Modern Power Systems and Clean Energy, 2021, 9(1): 210-218.

[5] 邹宜金, 连应华, 黄新宇, 等. 基于声纹的高泛化性风机叶片异常检测方法研究[J]. 电子科技大学学报, 2021, 50(5): 795-800.
ZOU Y J, LIAN Y H, HUANG X Y, et al. High generalization in anomaly detection of wind turbine generator based on voiceprint[J]. Journal of University of Electronic Science and Technology of China, 2021, 50(5): 795-800(in Chinese).

[6] XU D, LIU P F, CHEN Z P. Damage mode identification and singular signal detection of composite wind turbine blade using acoustic emission[J]. Composite Structures, 2021, 255: 112954.

[7] TIAN S H, YANG Z B, CHEN X F, et al. Damage detection based on static strain responses using FBG in a wind turbine blade[J]. Sensors, 2015, 15(8): 19992-20005.

[8] HWANG S, AN Y K, YANG J, et al. Remote inspection of internal delamination in wind turbine blades using continuous line laser scanning thermography[J]. International Journal of Precision Engineering and Manufacturing-Green Technology, 2020, 7(3): 699-712.

[9] DENG L W, GUO Y G, CHAI B R. Defect detection on a wind turbine blade based on digital image processing[J]. Processes, 2021, 9(8): 1452.

[10] 徐灵鑫. 风力发电机叶片表面缺陷检测的研究[D]. 杭州: 中国计量学院, 2015.
XU L X. Study on surface defect detection of wind turbine blades[D]. Hangzhou: China University of Metrology, 2015(in Chinese).

[11] WANG L, ZHANG Z J. Automatic detection of wind turbine blade surface cracks based on UAV-taken images[J]. IEEE Transactions on Industrial Electronics, 2017, 64(9): 7293-7303.

[12] REDDY A, INDRAGANDHI V, RAVI L, et al. Detection of cracks and damage in wind turbine blades using artificial intelligence-based image analytics[J]. Measurement, 2019, 147: 106823.

[13] ZHANG C, WEN C B, LIU J H. Mask-MRNet: a deep neural network for wind turbine blade fault detection[J]. Journal of Renewable and Sustainable Energy, 2020, 12(5): 053302.

[14] LV L, YAO Z Y, WANG E M, et al. Efficient and accurate damage detector for wind turbine blade images[J]. IEEE Access, 2022, 10: 123378-123386.

[15] HU C S, ZHAO Y, CHENG F J, et al. Multi-object detection algorithm in wind turbine nacelles based on improved YOLOX-nano[J]. Energies, 2023, 16(3): 1082.

[16] LIU Y J, WANG Z, WU X L, et al. Cloud-edge-end cooperative detection of wind turbine blade surface damage based on lightweight deep learning network[J]. IEEE Internet Computing, 2023, 27(1): 43-51.

[17] ZHANG R, WEN C B. SOD-YOLO: a small target defect detection algorithm for wind turbine blades based on improved YOLOv5[J]. Advanced Theory and Simulations, 2022, 5(7): 2100631.

[18] HU Y Z, WANG L W, KOU T H, et al. YOLO-tiny-attention: an improved algorithm for fault detection of wind turbine blade[C]//Proceedings of the 8th International Conference on Intelligent Computing and Signal Processing. Piscataway: IEEE Press, 2023: 1228-1232.

[19] DAI L Y, LIU G, HUANG L, et al. Feature transfer method for infrared and visible image fusion via fuzzy lifting scheme[J]. Infrared Physics & Technology, 2021, 114: 103621.

[20] WANG C Y, BOCHKOVSKIY A, LIAO H M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2023: 7464-7475.

[21] VASWANI A, RAMACHANDRAN P, SRINIVAS A, et al. Scaling local self-attention for parameter efficient visual backbones[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 12889-12899.

[22] CHEN J R, KAO S H, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2023: 12021-12031.

[23] CHOLLET F. Xception: deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 1800-1807.

[24] TAN M X, PANG R M, LE Q V. EfficientDet: scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 10778-10787.

[25] TONG Z J, CHEN Y H, XU Z W, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[EB/OL]. (2023-04-08)[2024-04-10]. https://arxiv.org/abs/2301.10051.

[26] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[27] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]//Proceedings of the Computer Vision–ECCV. Berlin: Springer, 2016: 21-37.

[28] Zhu X Z, SU W J, LU L W, et al. Deformable DETR: deformable transformers for end-to-end object detection[EB/OL]. (2021-03-18)[2024-04-10]. https://arxiv.org/abs/2010.04159.

[29] REDMON J, FARHADI A. YOLOv3: an incremental improvement[EB/OL].(2018-04-08)[2024-04-10].https://arxiv.org/abs/1804.02767.

[30] JOCHER G, CHAURASIA A, STOKEN A, et al. Ultralytics/yolov5: v6. 2-yolov5 classification models, apple M1, reproducibility, clearML and deci. ai integrations[EB/OL]. (2022-08-17)[2024-04-10]. https://ui.adsabs.harvard.edu/abs/2022zndo...7002879J/abstract.

Wind turbine blade defect detection based on improved YOLOv7

TANG Zhanjun[1], ZHANG Chaojie[1], WANG Jian[2,*], LU Peng[3], LIU Huiyuan[2], JIAN Hong[1]

(1. School of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China;
2. School of Electric Power Engineering, Kunming University of Science and Technology, Kunming 650500, China;
3. Yunnan Longyuan New Energy Company, Kunming 650228, China)

Abstract: A wind turbine blade defect detection technique, EPW-YOLOv7, based on the YOLOv7 network model, is suggested for the features of wind turbine blade defects with large scale variation and complicated backdrop in order to achieve real-time accurate identification of wind turbine blade flaws. Firstly, the CSM attention module is designed and added to the backbone network to suppress the complex background interference and to improve the efficiency of extracting important features. Second, the lightweight PWK module is designed to replace the original (ELAN) module to reduce the amount of redundant parameters and computation, and to accelerate the detection speed of the network. Then, the bidirectional feature pyramid networks(BiFPN) feature fusion module is introduced to motivate the network to recognize multi-scale defect features more accurately. Finally, the wise-intersection over union(WIoU) loss function is used to optimize the network and improve the overall detection performance of the improved model. The wind turbine blade defect dataset is used for experiments, and the results demonstrate that the average accuracy of the EPW-YOLOv7 model can reach 88.4%, which is 7.1% higher than that of the YOLOv7-tiny model. Additionally, the frame rate reaches 66 frames/s, satisfying the requirement for real-time detection. In addition, compared with the current state-of-the-art target detection algorithms, the EPW-YOLOv7 model can detect defects of wind turbine blades with high accuracy and faster. In addition, compared with the current advanced target detection algorithms, the EPW-YOLOv7 model has more advantages in detecting the defects of wind turbine blades in terms of accuracy and speed, which demonstrates that the proposed algorithm is more suitable for the real-time detection and localization of wind turbine blade defect.

Keywords: wind turbine blade; defect detection; deep learning; attention mechanism; feature fusion

Received: 2024-04-17; **Accepted**: 2024-08-03; **Published Online**: 2024-08-07 15:32

URL: link.cnki.net/urlid/11.2625.V.20240807.1253.001

Foundation items: National Natural Science Foundation of China (82160347) ; Yunnan Basic Research Project-Youth Project (202401AU070148); National Energy Group Technology Project (CSIEKJ230700104)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0279

飞行员工作负荷定量实时预测模型

何雪丽[1,*]，王黎静[2]，湛希[1]，苗冲冲[1]，谷宗辉[3]
(1. 中国航空综合技术研究所，北京 100028；2. 北京航空航天大学 航空科学与工程学院，北京 100191；
3. 天津杰普逊国际飞行学院有限公司，北京 122000)

摘 要：基于驾驶舱环境下飞行员执行飞行任务时的操作，进行飞行员工作负荷预测评估，提出一种考虑影响因素、任务单元任务负荷、任务冲突系数、多任务系数的飞行员工作负荷定量实时预测模型。选择DA 42公务机的正常五边飞行作为预测计算的实例，设计任务场景，确定飞行操作程序，进行任务分解，通过所提模型计算出起飞、爬升、巡航、下降、进近5个飞行阶段下的飞行工作负荷实时预测值，并通过综合计算式计算得到不同飞行阶段下的平均工作负荷预测值。在DA 42公务机飞行模拟器上进行验证实验，采用NASA-TLX、Bedford、Overall工作负荷、SWAT这4种主观评价量表对不同飞行阶段的飞行工作负荷进行主观评价，得到的主观评价值与采用所提模型预测得到的工作负荷值存在非常高的相关性，值的差异性也非常小，从而验证了所提模型的准确性和可用性。

关 键 词：飞行员工作负荷；飞行任务；任务分解；主观评价；预测模型
中图分类号：V7; R857.1
文献标志码：A **文章编号**：1001-5965(2026)06-1915-09

基于任务分析法建立工作负荷预测模型是研究时间较长、研究经费较多、应用最广的一类工作负荷预测方法[1-2]。因为在飞机设计早期样机还未制造出来时，只能通过飞机人机界面及飞行操作程序进行飞行操作任务预测，从而预测飞行员工作负荷，所以在飞机设计阶段，只能通过分析预测手段进行工作负荷的评估[3-4]。对于飞行操作系统来说，其是一个随着时间动态变化的操作系统，建立一个动态的实时飞行员工作负荷预测模型非常重要。

由于工作预测模型研究的时期及应用领域的不同，已有多个基于任务的工作负荷预测模型出现。20世纪60年代，Siegel等[5]通过建立时间压力模型(完成任务所需要的时间与给定的完成任务的时间之比)进行工作负荷的预测评估。Pei等[6]基于Petri网和多资源模型理论，从常规资源占用、额外资源占用和时间占用3个维度构建了时间通道动作单元网络模型，用于预测飞行员空中加油脑力负荷。此类方法均是通过一段时间内的操作任务时间紧迫性来预测工作负荷，并没有考虑任务本身的负荷值。同时也只能对一段时间内的工作负荷平均值进行预测，不能对工作负荷进行实时预测。Bradley等[7]、Yan等[8]、Yu等[9]通过对飞行任务进行分析预测飞行员工作负荷，该类研究主要从任务的复杂程度单一维度进行工作负荷的预测。此类通过任务直接预测工作负荷的方法未考虑人的操作的特点，与任务操作实际存在差异。

20世纪90年代，美国陆军研究所建立了考虑任务负荷并可以实时预测工作负荷模型，如任务分析/工作负荷(task analysis/workload,TAWL)[10]、VACP模型[11-12]、装甲车任务分析和工作负荷(armored vehicle crews task analysis and workload, AVTAWL)[13]，其中，VACP为TAWL的一个裁剪版，AVTAWL为

收稿日期：2024-05-06；录用日期：2024-08-17；网络出版时间：2024-09-04 11:43
网络出版地址：link.cnki.net/urlid/11.2625.V.20240903.1744.001
*通信作者. E-mail：15210988283@163.com

引用格式：何雪丽，王黎静，湛希，等. 飞行员工作负荷定量实时预测模型[J]. 北京航空航天大学学报，2026，52(6)：1915-1923.
HE X L，WANG L J，ZHAN X，et al. Quantitative real-time prediction model of pilot workload[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1915-1923 (in Chinese).

TAWL模型的延伸。这些模型均基于任务分析模型将操作任务分解为任务单元,通过综合计算任务单元负荷从而预测操作人员工作负荷,此类工作负荷预测模型均基于VACP模型。在VACP模型中,单个任务的注意力需求用0～7分的量表进行评估各个资源维度的负荷值,包含视觉、听觉、认知、操作,通过统计各个资源维度的负荷值作为总负荷值。这些模型是为军用飞机的设计而建立的,并应用在AH-64、MH47、UH-60等飞机的设计中[14],Liang等[15]提出McCrackeneAldrich技术适用于作为一个脑力工作负荷预测分析的工具。虽然这些模型能进行工作负荷的实时预测,但这些模型在综合计算时,得到的是视觉(包含需要辅助和不需要辅助)、听觉、运动、认知和心理方面每个操作任务模块下各自的负荷值,而不是总的负荷值;同时,对于双任务情况也只是用相加的方法进行计算,而没有考虑任务的冲突性。

Honeywell公司的North等[16]建立的W/INDEX与20世纪90年代初建立的MIDAS TLM[17]也是在VACP模型的基础上建立的,其在多资源模型[18-20]的基础上考虑了任务冲突的情况,在提出的W/INDEX中明确了一个工作负荷的实时计算式。Li等[21]提出适用于设计初期的用于评估单个飞行员的工作负荷预测模型方法论,该模型的建立基于任务分析、层次任务分析、多资源理论,该模型参考了North等[16]提出的W/INDEX计算式,在此基础上进行了更改和验证工作。但W/INDEX中没有基于任务属性对任务进行分类,而是对具体的任务进行注意力分析,其任务冲突负荷也是基于具体的任务进行赋值,没有通用性;MIDAS TLM中,所有的任务需求和任务冲突值都是直接给负荷值进行赋值,赋值只出现在其报告中,未正式公布,其直接应用的可行性有待验证。

本文在多资源模型、W/INDEX和MIDAS TLM模型研究的基础上,考虑了任务本身属性和人操作属性,进行了优化考虑,提出一种考虑影响因素、任务单元、冲突任务、多任务的工作负荷预测模型,并采用矩阵计算的数学表达方式体现,通过简单的模拟飞行实验对模型进行验证实验。

该模型的综合计算式为矩阵形式,可以对工作负荷的实时过程进行直观地描述和表达,该数据模型也可为计算模型的实现提供很好的数据基础。这种应用于设计阶段的工作负荷预测,可将预测评估结果应用于早期的设计中,为驾驶舱人机界面设计提出指导,减少不必要的反复设计修改,在一定程度上缩短设计周期和设计成本。

1 工作负荷预测模型

1.1 模块分析

基于任务的飞行员工作负荷预测模型包含4个模块:影响因素模块、任务单元模块、冲突任务模块、多任务模块。本文将其分为影响因素模块和任务模块(任务单元模块、冲突任务模块、多任务模块)进行分析。

1.1.1 影响因素模块

Meshkati等[22]提出脑力负荷影响因素分为任务、环境、操作者特征等。van Tonder[23]揭露脑力负荷的影响因素包括人、人机界面、环境等因素。在这2个研究基础上,与专家飞行员面谈后,本文提出了详细的飞行员工作负荷的影响因素,主要分为4个方面:飞行任务、飞行环境、驾驶舱人机界面、个体因素。其中,飞行任务包括:飞行状态(正常/非正常)、时间压力、飞行任务的重要性、飞行任务的熟练程度、飞行阶段;飞行环境包括:天气(晴天/雷雨天)、时间(白天/夜晚)、地形环境、机场环境、飞机类型(真机/模拟);驾驶舱人机界面包括:界面的视觉整洁性、界面的操作逻辑合理性、显示界面的可读性程度、操纵界面的可操作性;个体因素包括:性格、心情(当时)、疲劳、飞行时间(总)。

1.1.2 任务模块

在飞行操作任务中,除了大量的单任务,也有大量的多任务同时操作的情况,多任务操作是必须考虑的情况之一。多任务操作机理研究是建立可靠工作负荷预测模型的一个重要关键技术。Welford[24]认为在信息加工系统中存在一个瓶颈,使得操作者很难或几乎不可能同时对2个不同刺激做出恰当的反应。Wickens[18]总结了相关证据后做出结论,2个任务彼此干扰的程度取决于其享用相同的感觉通道(如视觉或听觉)、涉及的加工阶段(如输入、内部加工和输出)及依赖相关的记忆编码(如言语或视觉编码)。Wickens[18]的多资源理论认为,人有多个不同的资源,而不是一个中心资源,每种资源有自己的一些特性。当不同的工作需要不同的资源时,其之间的干扰就小些,如果其所需要的资源相同或相近,干扰就大些。也有研究结论认为当且只有当不同通道的信息有维度重叠而且是协同的时候,2个通道的信息才能平行加工[25-27]。Heyde[28]提出的模特法(简称MODAPTS)中对操作动作的同时性,得出以下结论:双手的终结动作均不需要注意力时,可以同时动作;只有一只手的终结动作需要注意力时,可以同时动作;双手的终结动作均需注意力时,不可能同时动作。在多任务同时发生时,既有

发生干扰出现冲突的情况，也有可以同时操作的任务。所以在计算模型中，除了考虑单任务顺序发生，还应该考虑多任务同时发生时的冲突性或同时性。

通过以上分析，本文将同通道任务同时发生称为冲突任务；不同通道任务同时发生称为多任务。将在基于任务的工作负荷预测模型中考虑单任务负荷，同通道任务的冲突性，不同通道任务的并行操作性。

1.2　计算模型

基于任务的飞行员工作负荷预测模型应该包含 4 个模块：影响因素模块、任务单元模块、冲突任务模块、多任务模块。该模型在任务单元负荷的基础上，考虑任务冲突的冲突系数，多任务操作时多任务负荷系数值，然后采用矩阵计算方法进行表达。为方便计算，计算式包含 5 个模块，增加了任务单元存在性矩阵：

$$\boldsymbol{W}_t = K_t \boldsymbol{A}_t \boldsymbol{B}_t \boldsymbol{C}_t \boldsymbol{D}_t \tag{1}$$

式中：K_t 为在 t 时刻下的影响因素系数；$\boldsymbol{A}_t$ 为存在性矩阵；$\boldsymbol{B}_t$ 为单任务负荷值矩阵；$\boldsymbol{C}_t$ 同通道任务冲突系数值矩阵；$\boldsymbol{D}_t$ 为多任务负荷系数矩阵。

K_t 值的确定主要通过设计单一变量，如飞行阶段、飞机类型（真机/模拟）、时间压力等。在实验中设计 2 组实验，采用 1～10 的工作负荷评价量表进行工作负荷值的评估，对单一影响因素下的工作负荷值进行回归分析，获得系数值。

任务单元存在性矩阵为

$$\boldsymbol{A}_t = \begin{bmatrix} a_{\mathrm{v}} & a_{\mathrm{a}} & a_{\mathrm{c}} & a_{\mathrm{k}} & a_{\mathrm{s}} \end{bmatrix} \tag{2}$$

式中：a 表示各通道任务种类的存在性，a 取 1 表示存在，0 表示不存在。按照 Wickens 的多资源及信息处理模型，将任务按通道分为视觉感知（v）、听觉感知（a）、感知（c）、动力输出（k）及语言输出（s）5 类。所以下标 v 表示视觉输入任务，a 表示听觉输入任务，c 表示认知任务，k 表示动作输出任务，s 表示语言输出任务。

任务单元负荷值矩阵为

$$\boldsymbol{B}_t = \begin{bmatrix} b_{\mathrm{v}} & 0 & 0 & 0 & 0 \\ 0 & b_{\mathrm{a}} & 0 & 0 & 0 \\ 0 & 0 & b_{\mathrm{c}} & 0 & 0 \\ 0 & 0 & 0 & b_{\mathrm{k}} & 0 \\ 0 & 0 & 0 & 0 & b_{\mathrm{s}} \end{bmatrix} \tag{3}$$

式中：b 为任务单元负荷值，存在的任务单元负荷值根据实际的负荷进行填写，而不存在的任务单元负荷 b 用 1 代替。基于操作任务进行单任务负荷的测量，通过模拟飞行实验让被试对单任务负荷进行主观评价，采用 1～10 的工作负荷评价量表进行工作负荷值的评估，建立单任务负荷数据库。

任务冲突矩阵为

$$\boldsymbol{C}_t = \begin{bmatrix} c_{\mathrm{v}} & 0 & 0 & 0 & 0 \\ 0 & c_{\mathrm{a}} & 0 & 0 & 0 \\ 0 & 0 & c_{\mathrm{c}} & 0 & 0 \\ 0 & 0 & 0 & c_{\mathrm{k}} & 0 \\ 0 & 0 & 0 & 0 & c_{\mathrm{s}} \end{bmatrix} \tag{4}$$

式中：c 为任务冲突系数，存在冲突的通道任务冲突系数按实际的填写，而不存在冲突的通道任务处 c 用 1 代替。冲突值通过基础实验进行同通道任务冲突系数的测量。采用 1～10 的工作负荷评价量表进行工作负荷值的评估。

多任务系数矩阵为

$$\boldsymbol{D}_t = \begin{bmatrix} d_{\mathrm{v}} & d_{\mathrm{a}} & d_{\mathrm{c}} & d_{\mathrm{k}} & d_{\mathrm{s}} \end{bmatrix}^{\mathrm{T}} \tag{5}$$

式中：d 表示 2 个及 2 个不同通道任务同时发生时，多任务负荷与任务单元负荷之间的系数，存在多任务系数按实际的填写，而不存在多任务的通道任务处 d 用 1 代替。d 也是通过基础实验测量得到，采用 1～10 的工作负荷评价量表进行工作负荷值的评估。

结合式（1）～式（5），工作负荷预测模型中的计算模型为

$$\boldsymbol{W}_t = K_t \begin{bmatrix} a_{\mathrm{v}} & a_{\mathrm{a}} & a_{\mathrm{c}} & a_{\mathrm{k}} & a_{\mathrm{s}} \end{bmatrix} \begin{bmatrix} b_{\mathrm{v}} & 0 & 0 & 0 & 0 \\ 0 & b_{\mathrm{a}} & 0 & 0 & 0 \\ 0 & 0 & b_{\mathrm{c}} & 0 & 0 \\ 0 & 0 & 0 & b_{\mathrm{k}} & 0 \\ 0 & 0 & 0 & 0 & b_{\mathrm{s}} \end{bmatrix} \begin{bmatrix} c_{\mathrm{v}} & 0 & 0 & 0 & 0 \\ 0 & c_{\mathrm{a}} & 0 & 0 & 0 \\ 0 & 0 & c_{\mathrm{c}} & 0 & 0 \\ 0 & 0 & 0 & c_{\mathrm{k}} & 0 \\ 0 & 0 & 0 & 0 & c_{\mathrm{s}} \end{bmatrix} \begin{bmatrix} d_{\mathrm{v}} & d_{\mathrm{a}} & d_{\mathrm{c}} & d_{\mathrm{k}} & d_{\mathrm{s}} \end{bmatrix}^{\mathrm{T}} \tag{6}$$

对于飞行操作任务，将飞行操作子任务分为视觉感知、听觉感知、感知、动力输出及语言输出 5 类后，对于实验飞行操作中，有将飞行操作分为 3 类：单任务操作、同通道多任务操作和不同通道多任务操作。

由式（6）可知，当飞行操作只有单任务时，基于单任务的种类对矩阵 $\boldsymbol{A}_t$ 中对应的任务种类的 a 赋值 1，其余赋值为 0，此时矩阵 $\boldsymbol{B}_t$ 中对应种类的单任务负荷值赋予相应的负荷值，而其他值为 1，矩阵 $\boldsymbol{C}_t$ 为单位矩阵，矩阵 $\boldsymbol{D}_t$ 全为 1；当飞行操作为处于同一通道的双任务或多个任务同时操作构成时，对矩阵 $\boldsymbol{A}_t$ 中对应的任务种类的 a 赋值 1，其余赋值为 0，此时矩阵 $\boldsymbol{B}_t$ 中对应种类的单任务负荷值赋予相应的负荷值，而其他值为 1，矩阵 $\boldsymbol{C}_t$ 中对应任务种类赋予同通道任

务操作的任务负荷冲突系数，矩阵 $\boldsymbol{D}_t$ 全为1；当飞行操作为处于不同通道的双任务或多个任务同时操作构成时，种类对矩阵 $\boldsymbol{A}_t$ 中对应的任务种类的 a 赋值1，其余赋值为0，此时矩阵 $\boldsymbol{B}_t$ 中对应种类的单任务负荷值赋予相应的负荷值，而其他值为1，矩阵 $\boldsymbol{C}_t$ 为单位矩阵，矩阵相应通道下的系数进行赋值。

在时间 T 内，工作负荷的平均负荷值 W_{ave}：

$$W_{\text{ave}} = \int_0^T W_t \mathrm{d}t / T_{\text{total}} \tag{7}$$

式中：T_{total} 为任意时间段，如起飞阶段、巡航阶段，也可以是整个飞行阶段。但是如果时间过长，就应该对预测模型加入更多的影响因素（如疲劳因素），才能保证模型预测的精度。

1.3 验证关系

工作负荷的测量方法主要有3类：生理测量、主观测量、绩效测量。主观测量是用于操作人员感受到的“工作负荷”多少的最准确的方法，因为工作负荷的表述最终还是主观性的。在众多的主观评价方法中，工作负荷测量研究常用的方法包括：美国国家航空航天局任务负荷指数（NASA Task Load Index, NASA-TLX）[29]，主观负荷评估技术（subjective workload assessment technique, SWAT）[30]，修订的库柏-哈柏量表（modified Cooper-Harper scale, MCH）方法[30]，贝德福德量表（Bedford Scale）方法[30]，综合 Overall 工作负荷方法[30]。NASA-TLX 方法被应用在工业工效学领域最早的工作负荷研究中，也被认为是最好的多维度测量方法[31]，该方法包含脑力负荷、体力负荷、任务需求、绩效、努力程度、受挫程度等维度。Hamilton 等[10]将 NASA-TLX 作为工作负荷分析模型（TAWL）的验证标尺，用于预测评估 UH-60A 飞行仿真中的典型任务下的飞行员工作负荷。Lin 等[32]采用 NASA-TLX 对通讯环境下及不同飞行阶段的飞行工作负荷进行了测量，用于研究通讯和飞行阶段对飞行工作负荷的影响。

工作负荷预测值与主观评价值之间高度关联，式（8）的线性回归式用于表达工作负荷预测值与 NASA-TLX 评估值之间的关系。

$$f_{\text{NASA-TLX, mean}} = aW_{\text{ave}} + b \tag{8}$$

式中：a 和 b 为工作负荷预测值与主观评价值线性关系系数。

Corwin 等[33]在美国空军和美国联邦航空管理局的资助下，为飞机的最小机组航认证中的机组工作负荷评估技术研究中，采用了 SWAT、NASA-TLX、Bedford Scale 及 Overall 工作负荷4种主观评价方法对飞行任务场景下不同飞行阶段的飞行工作负荷进行了测量，并进行了检验分析，结果表明这4种测量方法都可用于飞行员工作负荷的测量评估。

所以本文将采用 NASA-TLX、Bedford Scale、Overall 工作负荷、SWAT 这4种主观评价方法对飞行工作负荷预测值进行验证。

2 预测分析实验

2.1 飞行任务

DA 42 公务机的正常五边飞行，飞行任务示意图如图1所示。飞行操作手册[34]中的正常飞行操作程序与设计的飞行任务有一定的差异，所以本文采用结合飞行操作手册[34]，并与飞行教员面谈的方式，确定该飞行任务下各个飞行阶段下的飞行操作程序，飞行任务可分为5个飞行阶段：起飞、爬升、巡航、下降、进近。

2.2 飞行操作工作负荷预测值

2.2.1 实时负荷值

通过研究 DA 42 飞行操作程序，并进行任务分

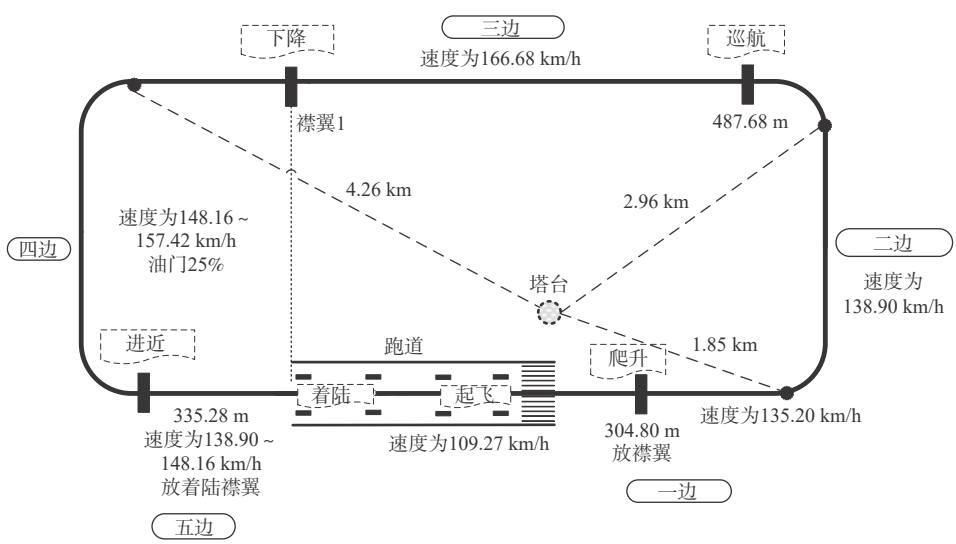


图1 飞行任务示意图

Fig. 1 Schematic diagram of flight task

析, 逐级分解到飞行任务单元, 最后采用式(7)计算得到各个飞行阶段下的飞行工作负荷实时预测值, 如图 2 所示。

飞行操作工作负荷预测的基础研究建立的数据基础: 影响因素系数 K_i、任务单元负荷 b、多任务系数 c、任务冲突值 d。

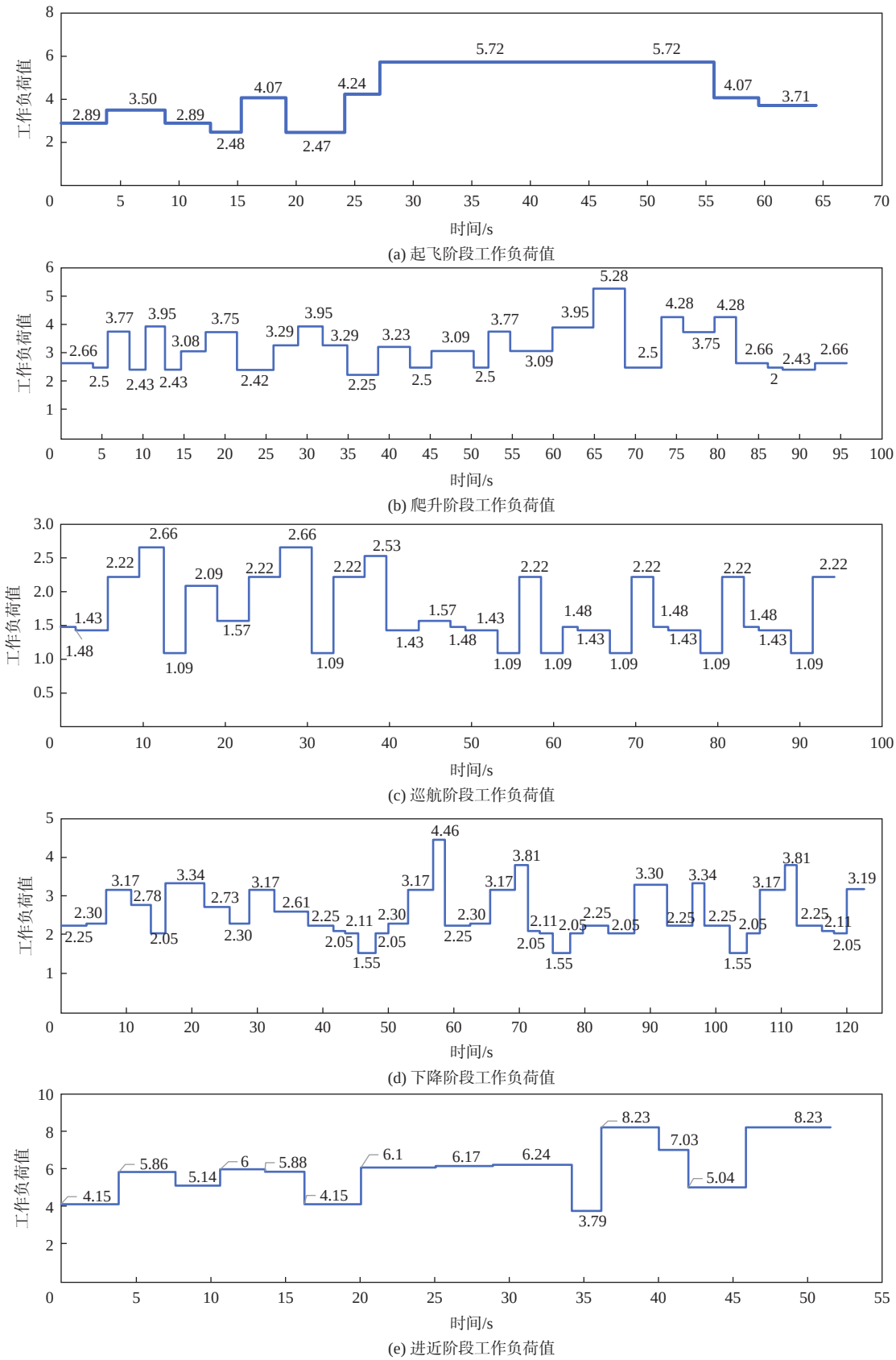


图 2　飞行工作负荷预测值

Fig. 2　Predictive value of pilot workload

影响因素系数 K_t。该验证实验是在模拟飞行环境下进行验证，所以影响因素只考虑“飞行阶段”，飞行阶段影响系数为飞行阶段起飞系数为 0.834，爬升系数为 0.769，巡航系数为 0.454，下降系数为 0.650，进近系数为 1.257[35]。

任务单元负荷值 b。主观评价得到的操作任务单元负荷值数据库、任务单元操作时间数据库[36-37]，如监控高度负荷值为 3.73，操纵驾驶杆负荷值为 4.26。

多任务负荷值 c。通过多任务操作总负荷与单任务负荷回归分析得到。如“监控高度+驾驶杆”和“监控姿态仪+驾驶杆”双任务在与单任务负荷线性回归系数中，视觉输入任务（监控高度和监控姿态仪）回归系数均为 0.60 左右，操作输出任务（操纵驾驶杆）的回归系数为 0.7 左右。“监控速度+操纵油门杆”和“监控垂直速度+操纵油门杆”双任务的回归系数中，操作输出任务（操纵油门杆）的回归系数为 0.65 左右，监控速度的回归系数为 0.65 左右，监控垂直速度的回归系数为 0.55 左右。

任务冲突值 d。通过设计“双视觉、双听觉、视觉+听觉”冲突任务实验，得到视觉冲突值为 1.14，听觉冲突任务系数为 1.27。

2.2.2 工作负荷综合计算预测结果

根据计算式 $W_{\text{ave}}=\int_0^T W_t\mathrm{d}t/T_{\text{total}}$，计算各个飞行阶段的工作负荷值，其中，$T_{\text{total}}$ 由飞行经验确定。通过对多组模拟飞行记录的飞行时间确定，起飞的总时间为 80 s，爬升的总时间为 110 s，巡航的总时间为 120 s，下降的总时间为 90 s，进近总时间为 70 s。

所以在考虑了飞行阶段系数的情况下得到的飞行工作负荷预测值如表 1 所示。

表 1 不同飞行阶段下的工作负荷预测值

Table 1 Predictive value of pilot workload under different flight phases

飞行阶段	工作负荷预测值
起飞阶段	3.55
爬升阶段	2.78
巡航阶段	1.34
下降阶段	3.50
进近阶段	4.42

3 验证实验及结果

3.1 实验设置

实验设备为 DA 42 模拟飞行器，能真实模拟 DA 42 飞机飞行过程界面和飞行场景。

被试者为 12 名飞行学院大四即将毕业的学生，已取得飞行驾照，飞行时间为 250 h 左右。

DA 42 公务机的正常五边飞行，飞行任务示意图如图 1 所示。由于最后的验证实验采用的是模拟飞行实验，而模拟飞行中无法对着陆后的情况进行模拟，所以任务不包含着陆阶段，分为起飞、爬升、巡航、下降、进近 5 个飞行阶段。

实验流程设计如下：首先，给被试者介绍飞行任务；其次，介绍量表的意义和填写方式；最后，让被试者进行练习至少 2 次以上，当熟练掌握了飞行任务要求后，即可开始正式实验。实验结束后，让被试者基于 5 个飞行阶段进行 NASA-TLX 量表、Bedford 量表、Overall 工作负荷量表和 SWAT 量表工作负荷主观评价工作。

3.2 主观评价工作负荷值及相关性分析

将 12 名飞行员在模拟飞行实验中，采用 4 种主观评价量表评价得到的不同飞行阶段下的飞行工作负荷主观评价值进行平均化处理，如表 2 所示。

表 2 不同飞行阶段下工作负荷主观评价值

Table 2 Subjective value of pilot workload under different flight phases

主观评价量表	工作负荷				
	起飞阶段	爬升阶段	巡航阶段	下降阶段	进近阶段
NASA-TLX	6.27	5.50	4.13	5.73	7.87
Bedford	3.60	2.82	1.89	3.30	4.32
Overall	51.54	40.77	24.23	45.00	62.08
SWAT	1.32	1.19	1.01	1.50	1.80

采用 Pearson 相关性对不同主观评价方法评价得到的工作负荷值进行相关性分析，结果如表 3 所示。

表 3 不同主观评价负荷值相关性分析结果

Table 3 Correlations analysis between different subjective evaluation load value

主观评价量表	Pearson 相关性	显著性（双侧）
NASA-TLX	0.953*	0.012
Bedford	0.989**	0.001
Overall	0.984**	0.002
SWAT	0.925*	0.024

注：** 表示在0.01 水平（双侧）上显著相关；* 表示在 0.05 水平（双侧）上显著相关。

由表 3 可知，通过 NASA-TLX、Bedford、Overall 工作负荷、SWAT 这 4 种主观评价量表得到的主观评价值与采用预测模型预测得到的工作负荷值存在非常高的相关性。

3.3　对比分析

将 NASA-TLX、Bedford、Overall 工作负荷这 3 种主观评价量表得到的不同飞行阶段下的飞行工作负荷值进行单位换算到同一数值范围内, 方便数值上的对比分析, 将 NASA-TLX 量评价值和 Overall 工作负荷评价值都处理为最大负荷值为 10 的标准中, 将 NASA-TLX 主观评价负荷值都除 2, Overall 工作负荷主观评价值除 10, 工作负荷值的对比示意图如图 3 所示。

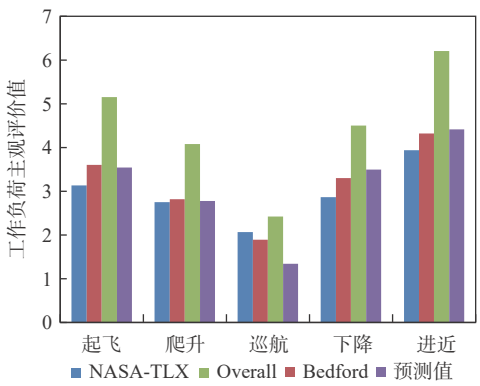


图 3　工作负荷主观评价值与预测值对比

Fig. 3　Contrast analysis of subjective value and predictive value

由图 3 的工作负荷值对比可知, NASA-TLX 和 Bedford 主观评价得到的值与本文模型预测得到的工作负荷值之间的差值非常小, 方差为 0.3 左右。Overall 工作负荷量表评价得到的负荷值明显高于本文模型预测值, 方差为 5.5 左右。

通过 4 种主观评价方法得到的飞行工作负荷值与采用预测模型预测评估得到的工作负荷值的相关性分析和数值的对比分析可以表明, 本文模型具有很高的准确性和可行性, 可用于预测飞行员工作负荷。

4　讨　论

研究结果表明本文工作负荷预测模型可以详细而直观地描述飞行操作任务下的通道使用情况, 并且可以很好地实时预测飞行员工作负荷。基于这些结果, 一些问题需要进行以下讨论。

第 1 个问题是预测模型的有效性分析, 本文采用 Pearson 相关性系数对该预测模型的有效性进行检验, Bedford Scale 工作负荷值, Overall 工作负荷量表, NASA-TLX 量表模型与工作负荷预测结果的相关性分析结果表明本文模型预测的飞行负荷值与 NASA-TLX 等主观评价量表评价的结果存在很高的相关性, 所以本文模型能有效地预测工作负荷。

第 2 个问题是工作负荷预测模型的敏感性, 采用配对 T 检验对起飞、爬升、巡航 3 个阶段的工作负荷预测结果进行显著性检验, 检验结果如表 4 所示。各个飞行阶段的工作负荷值存在显著性差异 (p<0.05)。

表 4　成对样本检验

Table 4　Sensitive measure test

成对样本	差分的 95% 置信区间下限	差分的 95% 置信区间上限	显著性
起飞-爬升	0.67	0.96	0
爬升-巡航	0.58	1.24	0.002
巡航-起飞	−2.18	−1.27	0

第 3 个问题是预测值与真实值之间的差异性, 基于任务的工作负荷预测模型为保障预测值最大程度与真实值相同, 考虑了所有因素, 本文提出了影响因素库, 由于影响因素较多, 做了系列的研究工作, 包括飞行阶段、时间压力、飞行时间(飞行经验)等, 其中, 飞行时间(飞行经验)影响不大, 飞机类型(真机) 影响系数为 1.6[35], 后续根据工作需要开展影响因素的实验, 构建影响因素系数数据库, 最大程度得到真实飞行操作工作负荷值。

5　结　论

本文提出一个考虑了影响因素、单任务负荷、多任务负荷、任务冲突等要素的飞行员工作负荷实时预测模型及其阶段负荷综合计算方法。

1) 本文提出工作负荷实时预测模型及其阶段综合计算方法, 并通过飞行实验, 采用主观评价的方法对本文模型进行了验证, 结果表明该模型可用于飞行工作负荷值的预测, 且准确性较高。

2) 本文模型可敏感预测由各种通道任务产生的工作负荷, 因为其考虑了同通道任务同时发生的冲突任务和不同通道任务同时发生的多任务。而且实时获得的是工作负荷绝对值大小, 该值可与多种工作负荷主观评价方法进行对比分析, 所以本文模型预测出的结果即可表现工作负荷真实大小, 为尽量保障预测值趋近于真实值, 后续将完善影响因素数据库。

3) 本文模型是一个实时预测模型及其阶段综合计算方法, 验证实验主要是对综合计算处理后的数据进行了验证, 采用 NASA-TLX(公认的验证标尺)等主观评价量表进行验证, 为减少验证评估工作对飞行任务的影响, 提高评估的准确性, 选择只对 5 个阶段进行验证, 验证结果对综合计算式得到的值进行了验证, 验证了综合计算模型的准确性, 在一定程度也验证了实时预测数据的准确性。

4) 在模型验证之前, 通过大量基础实验研究,

构建了数据库,包括影响系数(飞行阶段、飞行时间)、任务单元负荷数据、多任务负荷数据、冲突系数,本文模型的数据来源为飞行操作任务,其可较好地应用在不同飞机设计阶段,通过飞机界面设计预测工作负荷,从而达到对人机交互界面设计进行分析评估的目的。

参考文献(References)

[1] XIE B, SALVENDY G. Review and reappraisal of modelling and predicting mental workload in single-and multi-task environments[J]. Work & Stress, 2000, 14(1): 74-99.

[2] MEGAW T. The definition and measurement of mental workload[M]//3rd ed. Evaluation of Human Work. Boca Raton: CRC Press, 2005: 525-551.

[3] PARKS D L, BOUCEK J P. Workload prediction, diagnosis, and continuing challenges[M]//Applications of Human Performance Models to System Design. Berlin: Springer, 1989: 47-63.

[4] PATEL U H, SALVENDY G, GEDDES L A, et al. An electrical-circuit model for predicting mental workload in computer-based tasks[J]. Journal of the Chinese Institute of Industrial Engineers, 2002, 19(1): 1-15.

[5] SIEGEL A I, WOLF J J. Man-machine simulation models[M]. New York: John Wiley & Sons, 1969.

[6] PEI H N, MA Y J, LI W H, et al. Mental workload evaluation model of receiver aircraft pilots based on multiple resource theory[J]. Human Factors and Ergonomics in Manufacturing & Service Industries, 2024, 34(2): 159-173.

[7] BRADLEY R, MACDONALD C A, BUGGY T W. Quantification and prediction of pilot workload in the helicopter/ship dynamic interface[J]. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 2005, 219(5): 429-443.

[8] YAN X F, CHEN R L. Augmented flight dynamics model for pilot workload evaluation in tilt-rotor aircraft optimal landing procedure after one engine failure[J]. Chinese Journal of Aeronautics, 2019, 32(1): 92-103.

[9] YU Z M, YAN X F, CHEN R L. Prediction of pilot workload in helicopter landing after one engine failure[J]. Chinese Journal of Aeronautics, 2020, 33(12): 3112-3124.

[10] HAMILTON D B, BIERBAUM C R. Task analysis/workload (TAWL): a methodology for predicting operator workload[J]. Proceedings of the Human Factors Society Annual Meeting, 1990, 34(16): 1117-1121.

[11] MCCRACKEN J H, ALDRICH T B. Analysis of selected lhx mission functions: implications for operator workload and system automation goals: DTIC ADA232330[R]. Washington, D. C. : Army Research Institute Aviation Research and Development Activity, 1984.

[12] ALDRICH T B, SZABO S M, BIERBAUM C R. The development and application of models to predict operator workload during system design[M]//Applications of Human Performance Models to System Design. Berlin: Springer, 1989: 65-80.

[13] RAYMOND C S. Data elements for workload analysis of armored vehicle crews: DTIC ADA228422[R]. Washington, D. C. : United States Army Research Institute Aviation Research and Development Activity. 1990.

[14] EDGAR M J. Task analysis of the UH-60 mission and decision rules for developing a UH-60 workload prediction model: DTIC ADA201317[R]. Washington, D. C. : Army Research Institute Aviation Research and Development Activity, 1989.

[15] LIANG S M, RAU C L, TSAI P F, et al. Validation of a task demand measure for predicting mental workloads of physical therapists[J]. International Journal of Industrial Ergonomics, 2014, 44(5): 747-752.

[16] NORTH R A, RILEY V A. W/INDEX: a predicted model of operator workload[M]. New York: Plenum, 1989: 81-90.

[17] STAVELAND L E. MIDAS TLM: man-machine integrated design and analysis system task loading model[C]//Proceedings of the IEEE International Conference on Systems, Man and Cybernetics. Piscataway: IEEE Press, 1991: 1219-1224.

[18] WICKENS C D. Processing resources in attention[M]. New York: Academy Press, 1984: 63-101.

[19] WICKENS C D. Multiple Resource time sharing model, handbook of human factors and ergonomics methods[M]. Boca Raton: CRC Press, 2005.

[20] WICKENS C D. Multiple resources and performance prediction[J]. Theoretical Issues in Ergonomics Science, 2002, 3(2): 159-177.

[21] LI M, WANG M, WANG G Q, et al. Predictive mental workload modeling methodology for single-pilot operations system design[J]. Journal of Aerospace Information Systems, 2024, 21(3): 249-262.

[22] MESHKATI N, LOEWENTHAL A. The effects of individual differences in information processing behavior on experiencing mental workload and perceived task difficulty: a preliminary investigation[J]. Advances in Psychology, 1988, 52: 269-288.

[23] VAN TONDER J A. Certificate in industrial ergonomics[D]. Pretoria: University of Pretoria, 1999.

[24] WELFORD A T. The ‘psychological refractory period’ and the timing of high-speed performance: a review and a theory[J]. British Journal of Psychology General Section, 1952, 43(1): 2-19.

[25] DONOHUE S E, APPELBAUM L G, PARK C J, et al. Cross-modal stimulus conflict: the behavioral effects of stimulus input timing in a visual-auditory stroop task[J]. PLoS One, 2013, 8(4): e62802.

[26] YUVAL-GREENBERG S, DEOUELL L Y. The dog’s meow: asymmetrical interaction in cross-modal object recognition[J]. Experimental Brain Research, 2009, 193(4): 603-614.

[27] YUVAL-GREENBERG S, TOMER O, KEREN A S, et al. Transient induced gamma-band response in EEG as a manifestation of miniature saccades[J]. Neuron, 2008, 58(3): 429-441.

[28] HEYDE G C. The concept and history "MODAPTS plus"[R]. Sydney: Australian Association for Predetermined Time Standards and Research, 1982.

[29] HART S G, STAVELAND L E. Development of NASA-TLX (Task Load Index) results of empirical and theoretical research[M]. Amsterdam: North-Holland, 1988: 139-183.

[30] 杜俊敏, 何雪丽, 康卫勇. 飞行员工作负荷[M]. 北京: 北京航空航天大学出版社, 2021: 35-74.
DU J M, HE X L, KANG W Y. Pilot workload[M]. Beijing: Beijing University of Aeronautics & Astronautics Press, 2021: 35-74(in Chinese).

[31] SARAH M. Literature review workload measures: RTO-TR-HFM-121-Part-II [R]. Iowa: The University of IOWA. 2001.

[32] LIN C J, LIN P H, CHEN H J, et al. Effects of controller-pilot communication medium, flight phase and the role in the cockpit on pilots' workload and situation awareness[J]. Safety Science, 2012, 50(9): 1722-1731.

[33] CORWIN W H, SANDRY-GARZA D L, BIFERNO M H, et al. Assessment of crew workload measurement methods, techniques, and procedures: process, methods and results: WRDC-TR-89-7006 [R]. Wright-Patterson AFBOH: Wright Research and Development Center, 1989.

[34] Diamond Aircraft. Airplane flight manual DA 42[EB/OL].(2004-04-29)[2024-04-27]. https://www.manualslib.com/manual/1422391/Diamond-Da-42.html?page=3#manual.

[35] HE X L, DING L, MIAO C C, et al. Effects of flight environment on pilot workload in simulated and actual flight[C]//Proceedings of the Asia-Pacific International Symposium on Aerospace Technology. Berlin: Springer, 2019: 3040-3051.

[36] HE X L, DING L, MIAO C C, et al. Decomposition and classification of flight operation tasks of civil aircraft[C]//Proceedings of the Man–Machine–Environment System Engineering. Berlin: Springer, 2017: 789-801.

[37] HE X L, DING L, MIAO C C, et al. The research on task unit workload of civil aircraft flight operation[C]//Proceedings of the Man–Machine–Environment System Engineering. Berlin: Springer, 2018: 249-259.

Quantitative real-time prediction model of pilot workload

HE Xueli[1,*]，WANG Lijing[2]，ZHAN Xi[1]，MIAO Chongchong[1]，GU Zonghui[3]

(1. AVIC China Aero-Polytechnology Establishment，Beijing 100028，China；
2. School of Aeronautic Science and Engineering，Beihang University，Beijing 100191，China；
3. Tianjin Jeppesen International Flight College Co.，Ltd，Beijing 122000，China)

Abstract：Quantitative real-time prediction model of workload that take into account influencing factors, task unit load, task conflict coefficient, and multi-task coefficient are proposed based on the operational tasks of pilots conducting flying activities in the cockpit environment. Selecting the normal pentagonal flight of the DA42 business jet as an example for prediction calculation, designing task scenarios, determining flight operation procedures, decomposing tasks, and calculating the real-time predicted flight workload values for different flight stages through a comprehensive calculation formula. In order to subjectively assess the flight workload under various flying stages, a validation experiment was carried out on the DA42 business jet flight simulator utilizing four subjective assessment scales: NASA-TLX, Bedford, Overall workload, and SWAT. The subjective evaluation values obtained were highly correlated with the workload values predicted by the prediction model. This verified the correctness and usability of the proposed model of workload proposed in this study.

Keywords：pilot workload; flight tasks; decomposing tasks; subjective evaluation; prediction model

Received：2024-05-06；**Accepted**：2024-08-17；**Published Online**：2024-09-04 11：43

URL：link.cnki.net/urlid/11.2625.V.20240903.1744.001

* **Corresponding author**. E-mail：15210988283@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0265

面向直升机吊挂飞行品质改善的模糊减摆控制

王洛烽[1]，陈仁良[1,*]，赵宇[2]

(1. 南京航空航天大学 航空学院 直升机动力学全国重点实验室，南京 210016；
2. 中国空气动力研究与发展中心 旋翼空气动力学重点实验室，绵阳 621000)

摘 要：针对有人直升机吊挂飞行面临的吊挂减摆效果和指令跟踪性能冲突导致飞行品质低的问题，提出一种融合飞行员意图的模糊减摆控制方法。构建高置信度非线性人-机-吊挂闭环系统模型，分析基于飞控系统减摆面临的吊挂减摆效果和指令跟踪性能相互冲突问题，建立飞行员意图模糊推理系统并与吊索摆角反馈控制律融合。仿真结果表明：所提方法能够合理反映飞行员意图，将其与减摆控制器融合能够以符合飞行员意图的方式动态调整指令跟踪性能和吊挂减摆效果，避免了飞行员与减摆控制器的相互冲突引起飞行员诱发振荡，且鲁棒性较好。

关 键 词：直升机；直升机吊挂；模糊推理；飞行员意图；吊索摆角反馈；飞行员诱发振荡
中图分类号：V221.4
文献标志码：A **文章编号**：1001-5965(2026)06-1924-11

直升机外吊挂飞行是低空经济的重要组成部分，在远程伐木、输电塔安装、应急救援等方面用途广泛。然而，直升机外吊挂飞行安全面临吊挂摆动振荡威胁，当摆动幅度过大时，飞行员不得不抛弃吊挂物以保证直升机飞行安全，带来了巨大损失。

研究者已针对吊挂减摆方法进行了很多研究，主要可分为直接和间接2种，目前以间接法为主。直接减摆方法使用安装在吊挂物上的控制机构，直接将控制力或力矩施加到吊挂物或吊索上，而不考虑直升机的运动，例如，直升机腹部的主动吊钩[1]和吊挂物后部的尾翼[2]，但是直接减摆方法需要额外装置，导致质量成本增加，且适用面窄。相比之下，间接减摆方法使用直升机运动来减少吊挂摆动，如将吊索摆角反馈至旋翼变距[3]或使用输入整形技术[4]等。由于不需要加装额外装置，间接减摆方法的经济性更好，而且可有效适应不同吊挂配置，通用性更强，所以有很多学者研究了在飞行控制系统中集成间接减摆方法[5−8]。

然而，间接减摆方法面临指令跟踪性能和吊挂减摆效果之间的固有冲突，导致直升机吊挂飞行品质难以提高。由于间接减摆方法本质上是通过调节直升机的运动响应，使直升机按照能够抵消吊挂摆动的方式移动来减少吊挂摆动，在此过程中，直升机的运动由吊挂摆动状态决定，可能与飞行员的指令输入不一致，甚至导致飞行员诱发振荡[9]，不利于飞行品质提高。

动态调整减摆控制器响应模式是改善直升机吊挂飞行品质的有效方法，但如何调整还没有定论。Ivler等[3]提出一种基于飞行状态的控制器模式切换策略，可以根据飞行状态自动调整控制器模式，实现吊挂减摆效果和指令跟踪性能的动态变化。类似地，Patterson等[10]提出自动吊挂阻尼系统，控制器根据飞行员控制杆偏转量和持续

收稿日期：2024-04-29；**录用日期**：2024-08-16；**网络出版时间**：2024-09-02 16:33
网络出版地址：link.cnki.net/urlid/11.2625.V.20240902.1521.001
基金项目：直升机动力学全国重点实验室基金(2023-HA-LB-067-01)；中国博士后科学基金(2023M741678)；旋翼空气动力学重点实验室基金(RAL202302-4)；江苏省卓越博士后计划；江苏高校优势学科建设工程资助项目
*通信作者. E-mail：crlae@nuaa.edu.cn

引用格式：王洛烽，陈仁良，赵宇. 面向直升机吊挂飞行品质改善的模糊减摆控制[J]. 北京航空航天大学学报，2026，52(6)：1924-1934.
WANG L F，CHEN R L，ZHAO Y. Fuzzy anti-swing controller for improving handling quality of helicopter slung-load operation[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1924-1934 (in Chinese).

时间自动判断是否应该调整控制器模式。然而, 上述控制器模式动态调整规则均为预先给定, 可能与飞行员意图不一致, 尤其是直升机外吊挂飞行可应用于各种现实场景, 上述预设规则是否能够保证各种场景下的飞行员-减摆控制器有效协调还有待讨论。

现代控制系统的发展趋势是自动化、智能化, 吊挂减摆控制器应能够根据飞行员意图动态调整吊挂减摆效果和指令跟踪性能, 使吊挂减摆控制器与飞行员操纵协调。现有涉及直升机飞行员意图的研究主要是美国陆军航空应用技术局的飞行员辅助计划[11], 目标是开发智能化的人-机交互界面, 但并未将飞行员意图融合到飞行控制系统中。近年来, D'Intino 等[12] 提出基于交互式多模型滤波器的直升机飞行员意图识别方法, 通过预测直升机飞行员的期望轨迹并借助触感系统辅助飞行员完成飞行任务, 但仍未实现飞行员意图与飞行控制系统的深度融合。除此以外, 还有学者通过模糊推理[13-14]、隐式马尔可夫模型[15] 等方法识别飞行员意图并用于改善人-机交互。

本文旨在设计并验证一种融合飞行员意图的吊挂减摆控制器, 实现直升机吊挂系统闭环响应特性动态调整, 始终保持与飞行员意图一致。首先, 建立人-机-吊挂闭环系统模型并进行了验证。然后, 提出融合飞行员意图的减摆控制器, 基于 ADS-33E-PRF[16] 定性分析了吊挂飞行品质中稳定性和操纵性的矛盾。最后, 通过方波输入响应分析及横侧定位科目数值模拟分析验证了本文融合飞行员意图的减摆控制器的有效性。

1　人-机-吊挂耦合闭环系统模型

如图 1 所示, 人-机-吊挂耦合闭环系统模型由直升机吊挂飞行动力学模型、飞控系统和飞行员组成。由于线性模型不能满足大幅值操纵输入和大范围吊挂摆动的仿真要求, 本文基于非线性直升机吊挂飞行动力学模型建立人-机-吊挂系统模型, 操纵输入是座舱中的总距杆量、纵/横向驾驶杆量和

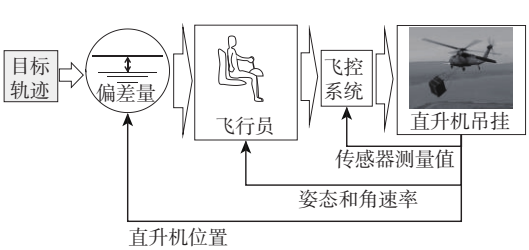


图 1　人-机-吊挂耦合闭环系统模型

Fig. 1　Pilot-helicopter-slung-load coupled closed-loop system model

脚蹬位移量, 考虑了从座舱操纵到自动倾斜器偏转的作动器、助力器等中间环节动态特性。飞控系统根据测量的状态反馈补偿直升机操纵量。飞行员通过视觉感知直升机位置与目标轨迹的误差, 再根据仪表和生理感知的姿态和角速率进行操纵。

1.1　直升机吊挂飞行动力学模型

本文建立的直升机吊挂系统模型由 UH-60A 直升机和 CONEX 集装箱组成[17]。直升机质量为 6 450 kg, 吊挂物质量为 1 862 kg, 吊索长度为 4.57 m。直升机飞行动力学模型经过很多年发展, 目前已经较为成熟[18-19], 本文建立的直升机模型考虑了机身刚体动力学、单片桨叶刚性挥舞/摆振动力学、三维动态入流、基于桨盘理论的尾桨模型及直升机气动干扰。吊挂物是刚体, 其上表面 4 个角与直升机机腹吊点通过 4 根线弹性吊索连接。吊挂物气动模型基于非线性风洞试验数据插值。具体建模方法可见文献 [20-21], 上述直升机吊挂系统模型的状态空间形式如下:

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

式中: $\boldsymbol{x}$为直升机吊挂系统模型的状态向量(见式(2)); $\boldsymbol{\delta}$为控制输入向量(见式(3)); t为时间。

$$\begin{aligned}\boldsymbol{x} = [&\dot{\beta}_1,\dot{\beta}_2,\dot{\beta}_3,\dot{\beta}_4,\beta_1,\beta_2,\beta_3,\beta_4,\dot{\zeta}_1,\dot{\zeta}_2,\dot{\zeta}_3,\dot{\zeta}_4,\zeta_1,\zeta_2,\\ &\zeta_3,\zeta_4,v_0,v_{1C},v_{1S},v_{0TR},u,v,w,p,q,r,\phi,\\ &\theta,\psi,x,y,z,u_L,v_L,w_L,p_L,q_L,r_L,\phi_L,\theta_L,\\ &\psi_L,x_L,y_L,z_L]^{\mathrm{T}}\end{aligned} \tag{2}$$

$$\boldsymbol{\delta} = [\delta_{\mathrm{col}},\delta_{\mathrm{lat}},\delta_{\mathrm{lon}},\delta_{\mathrm{ped}}]^{\mathrm{T}} \tag{3}$$

式中: β_i、$\dot{\beta}_i$、ζ_i、$\dot{\zeta}_i$分别为旋翼第i片桨叶的挥舞/摆振角和角速度; v_0、v_{1c}、v_{1s}为旋翼动态入流状态量; v_{0TR}为尾桨入流状态量; u、v、w为直升机的速度; p、q、r为直升机的角速度; ϕ、θ、ψ为直升机的姿态角; x、y、z为直升机在地轴系下的位置; 角标 L 表示吊挂物的对应状态量; δ_{col}、δ_{lat}、δ_{lon}、δ_{ped}分别为直升机的总距操纵、横向操纵、纵向操纵和脚蹬操纵。

上述非线性直升机吊挂系统模型在刚体动力学的基础上增加了旋翼动态特性, 考虑了旋翼挥舞后退型模态与直升机滚转运动的不利耦合, 以及旋翼挥舞运动滞后对带宽/相位延迟的不利影响, 基于该模型设计模糊减摆控制器, 可提高与实际飞行情况的一致性。

为便于后续研究, 本文根据直升机和吊挂物的相对运动定义吊索摆角。如图 2 所示, 在与地轴系$O_{\mathrm{E}}x_{\mathrm{E}}y_{\mathrm{E}}z_{\mathrm{E}}$方向一致的吊钩坐标系$O_{\mathrm{K}}x_{\mathrm{K}}y_{\mathrm{K}}z_{\mathrm{K}}$中定义了纵向摆角$\alpha_{\mathrm{c}}$和横向摆角$\beta_{\mathrm{c}}$。原点$O_{\mathrm{K}}$是直升机上的

吊点平均位置。原点O_K在体轴$O_Bx_By_Bz_B$中的相对位置矢量定义为$\underline{\boldsymbol{r}}_K^B$。$\underline{\boldsymbol{R}}_L^K$表示吊挂物重心在吊钩坐标系中的位置矢量。为得到摆角，将$\underline{\boldsymbol{R}}_L^K$分别用摆角和直升机吊挂系统状态量表示并建立其之间的关系为

$$\underline{\boldsymbol{R}}_L^K=\begin{bmatrix}x_L-x\\ y_L-y\\ z_L-z\end{bmatrix}-\underline{\boldsymbol{A}}_{KB}\underline{\boldsymbol{r}}_K^B=\begin{bmatrix}l_c\cos\beta_c\sin\alpha_c\\ -l_c\sin\beta_c\\ l_c\cos\beta_c\cos\alpha_c\end{bmatrix} \tag{4}$$

$$\begin{cases}l_c=\left|\underline{\boldsymbol{R}}_L^K\right|\\ \beta_c=\arcsin\left(-R_{L,y}^K/l_c\right)\\ \alpha_c=\arcsin\left(R_{L,x}^K/l_c\cos\beta_c\right)\end{cases} \tag{5}$$

式中：$\underline{\boldsymbol{A}}_{KB}$为从体轴系到吊钩坐标系的转换矩阵；$l_c$为直升机吊钩到吊挂物重心的距离。

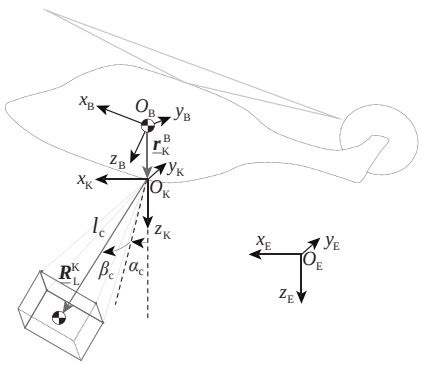


图2 吊索摆角定义

Fig. 2 Cable swing angle definition

图3为UH-60A+COENX吊挂物的横向通道频响特性验证结果，该频响曲线基于非线性直升机吊挂飞行动力学模型的扫频仿真结果辨识得到，与试验值的频响曲线求解方法一致。可发现频率在1～2 rad/s频段内，计算结果能够体现吊挂摆动模态引起的幅值衰减，且在其他频率范围内与试验值吻合良好，说明本文建立的非线性直升机吊挂飞行动力学模型精度较高，能够合理反映真实吊挂飞行特性。

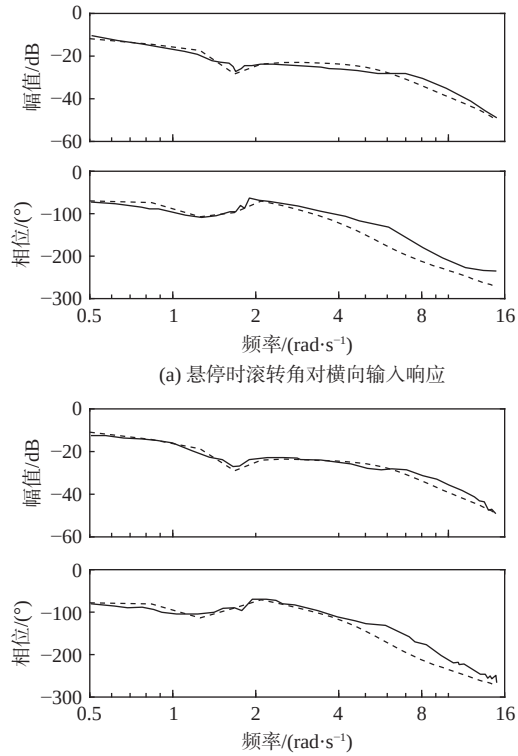
1.2 直升机飞行控制系统模型

直升机飞行控制系统采用现代直升机主流的显模型跟踪控制系统(model-following control system, MFCS)，以横向通道为例，闭环直升机吊挂系统横向通道传递函数Y_C为

$$Y_C=G_M\left(G_P^{-1}+\tau_PG_H\right)\frac{G_P\tau_P}{1+G_P\tau_PG_H}=G_M\tau_P \tag{6}$$

式中：G_P为直升机模型；G_M为指令模型；τ_P为直升机高阶环节的等效时间延迟；G_P^{-1}为逆模型；G_H为反馈回路。

MFCS的一大优点是系统的响应特性主要由显式指令模型G_M确定，若存在理想逆模型G_P^{-1}，则被控对象的动态特性可被完全抵消。在ADS-33E-PRF[16]的要求下，通过设计指令模型可得到理想的直升机飞行动力学特性和飞行品质。



(a) 悬停时滚转角对横向输入响应

(b) 前飞时滚转角对横向输入响应

——试验值 ---- 计算值

图3 UH-60A+CONEX横向通道频响特性

Fig. 3 Frequency response of UH-60A carrying CONEX container

1.3 飞行员模型

本文基于Hess理论建立结构化飞行员模型[22-23]，图4为横向通道频响特性，飞行员模型的输入为位置误差，模型内部的反馈回路包括本体感受反馈、前庭反馈和视觉反馈。

图4中，y_c为目标横向位置，Y_P为横向位置误差补偿，飞行员基于横向位置误差确定期望滚转角ϕ_c，再根据当前看到的直升机滚转角与期望滚转角间的误差，经过姿态视觉补偿Y_E和中枢系统时间延迟τ_0后确定预期操纵量U_E，再经过飞行员神经肌肉环节Y_{NM}和直升机座舱操纵力感系统Y_{FS}得到直升机横向操纵δ_{lat}，输入到闭环直升机吊挂系统Y_C，得到直升机运动响应，包括横向位置y、滚转角ϕ和滚转角速度$\dot{\phi}$。Y_{PF}为本体感受环节，飞行员根据当前的座舱操纵量对操纵量进行修正，对应的修正量是U_{PF}。Y_{VF}为前庭系统，飞行员根据生理感知的角速度对操纵量进行修正，对应的修正量为U_{VF}。表1为本文使用的横向通道结构化飞行员模型参数取值。

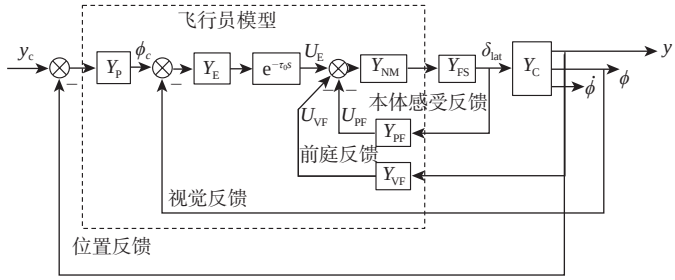


图 4　横向通道结构化飞行员模型

Fig. 4　Structural pilot model for lateral passages

表 1　横向通道结构化飞行员模型参数

Table 1　Parameter for horizontal channel structured pilot model

名称	取值
外环补偿传递函数Y_P	$-0.1s+0.02$
内环补偿传递函数Y_E	20
中枢延迟传递函数τ_0	0.2
直升机力感系统传递函数Y_{FS}	$\dfrac{25^2}{s^2+2\times0.707\times25s+25^2}$
神经肌肉系统传递函数Y_{NM}	$\dfrac{10^2}{s^2+2\times0.707\times10s+10^2}$
本体感受系统传递函数Y_{PF}	$\dfrac{2}{0.2s+1}$
前庭系统传递函数Y_{VF}	1.8

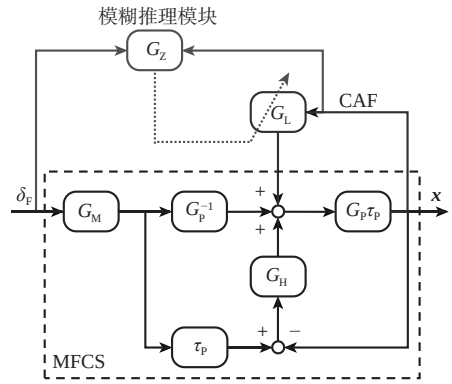


图 5　融合飞行员意图的模糊减摆控制器

Fig. 5　Fuzzy anti-swing controller that incorporates pilot intention

2　模糊减摆控制器设计

直升机的显模型跟踪控制系统中的逆模型是基于无吊挂的直升机模型得到的, 可以有效消除直升机本身的动态特性, 使直升机闭环系统响应与显模型一致。然而, 当直升机吊挂飞行时, 逆模型无法消除吊挂物的动态特性, 既会导致闭环响应与指令模型产生误差, 又会导致吊挂摆动容易被激发且难以收敛, 所以需要采取减少吊挂摆动的额外措施。

众多减摆方法中, 在飞控系统中集成减摆控制律是一种通用且经济的间接减摆技术, 但是面临指令跟踪与吊挂减摆之间的冲突问题。为此, 如图 5 所示, 本文在直升机 MFCS 的基础上, 结合吊索摆角反馈(cable angle feedback, CAF)与模糊推理模块, 提出融合飞行员意图的模糊减摆控制器, 基于飞行员意图模糊推理结果动态调整减摆控制器模式, 有助于解决指令跟踪性能与吊挂减摆效果之间的冲突问题。

如图 5 所示的模糊控制器基于显模型跟踪控制系统, 在式 (6) 基础上增加了吊索摆角反馈回路G_L和模糊推理系统G_Z, 实现直升机吊挂飞行的人-机协同控制。

2.1　基本的吊索摆角反馈减摆控制器

吊索摆角反馈减摆控制器已经飞行试验验证能够有效减少吊挂摆动, 但飞行员反应该方法会使吊挂物变得更重, 导致直升机更难操纵。吊索摆角反馈是将吊索摆角和角速度反馈至直升机操纵输入的间接减摆方法, 具体控制律如下:

$$\delta_{lat,L}=K_\beta\beta_c\frac{s}{s+0.5}+K_{\dot\beta}\dot\beta_c \tag{7}$$

$$\delta_{lon,L}=K_\alpha\alpha_c\frac{s}{s+0.5}+K_{\dot\alpha}\dot\alpha_c \tag{8}$$

式中: $\delta_{lat,L}$和$\delta_{lon,L}$分别为吊索摆角反馈引起的横向和纵向变距; K_β和$K_{\dot\beta}$分别为吊索横向摆角和角速度的反馈增益; K_α和$K_{\dot\alpha}$分别为吊索纵向摆角和角速度的反馈增益。此外, 使用截止频率为 0.5 rad/s 的高通滤波器消除配平下的摆角。

然而, CAF 同样面临间接减摆方法共有的指令跟踪与吊挂减摆冲突引起的飞行品质问题。目前的 ADS-33E-PRF[16] 中还没有明确的直升机吊挂飞行品质指标, 而最新的 ADS-33F[24] 还有待讨论, 所以本文使用 ADS-33E-PRF[16] 中的直升机飞行品质指标定性地评估集成 CAF 的直升机吊挂飞行品质。图 6 和图 7 分别为 CAF 横向增益变化对直升机飞行品质中稳定性和指令跟踪性能的影响, 图中展示了直升机飞行品质规范的不同等级边界。图 6 中, 当横向摆动角速度增益$K_{\dot\beta}$逐渐正向增加时, 吊挂横向摆动模态会从稳定模态变成不稳定模态。但图 7 中, 若$K_{\dot\beta}$为负值时会导致吊挂横向摆动频率处

的相位滞后太大，导致该频率处存在极大时间延迟。上述特点清楚地显示了间接减摆控制器无法同时满足指令跟踪性能和吊挂减摆要求，是导致直升机吊挂飞行品质难以提高的重要原因。对于横向摆角增益K_β，随着K_β负向增加，吊挂横向摆动模态的频率会逐渐增加，同时带宽逐渐增加（$K_\beta<0$）或逐渐减小（$K_\beta>0$），但并不会导致吊挂摆动不稳定性或飞行品质降级。

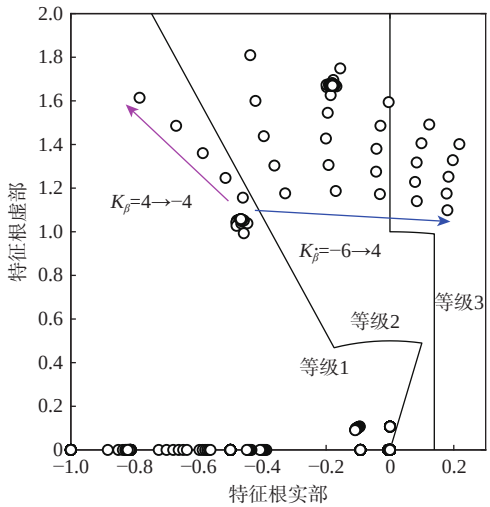


图6 稳定性指标（横航向振荡要求）

Fig. 6 Stability criterion (lateral-directional oscillatory requirements)

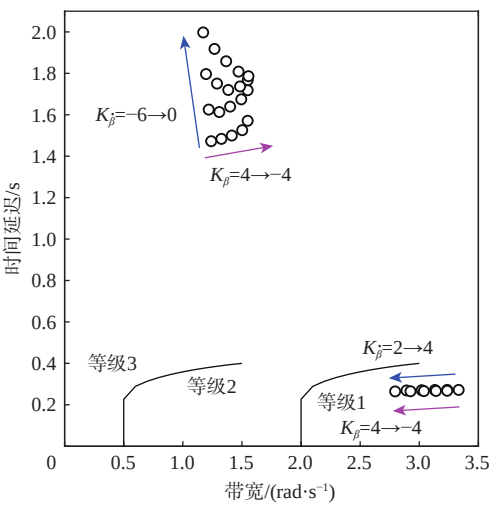


图7 指令跟踪指标（小幅滚转姿态变化要求）

Fig. 7 Controllability criterion (small-amplitude roll attitude change requirements)

由图6和图7的飞行品质结果可知，横向摆动角速度增益对飞行品质的影响比横向摆角增益更为显著，所以CAF的关键参数是横向摆动角速度增益K_β。当$K_\beta=4$时，吊挂摆动稳定性最好，但指令跟踪性能很差；而当$K_\beta=0$时，即不引入吊索摆角反馈，采用直升机原有飞行控制系统，指令跟踪性能较好，但会出现吊挂摆动不稳定性。

2.2 飞行员意图的模糊推理系统

为有效改善直升机吊挂飞行品质，本文基于模糊推理模块在线识别飞行员意图，再根据推理结果动态调整减摆控制器模式，使减摆控制器能够符合飞行员意图地进行吊挂减摆或跟踪指令，实现人性化的直升机吊挂飞行减摆控制。

由于飞行员意图通常是模糊且不确定的，本文基于模糊推理系统，根据飞行员瞬时输入幅值与历史平均输入幅值的差异，以及当前吊挂摆动能量大小判断飞行员意图，即倾向于让直升机跟踪指令还是需要尽快使吊挂摆动稳定，然后基于飞行员意图实时调整减摆控制器内部参数实现控制模式切换。

图8为飞行员意图模糊推理系统，控制器的输入为：横向摆角β_c和角速度$\dot{\beta}_c$及飞行员横向操纵$\delta_{\mathrm{lat,p}}$。通过对这些参数进行预处理可以得到模糊推理系统所需的飞行员瞬时/历史平均输入幅值的差异，以及当前吊挂摆动能量，再经过模糊推理系统得到飞行员意图，最后，根据飞行员意图实时调整CAF减摆控制器中的横向摆动角速度增益$K_{\dot{\beta}}$实现飞行品质动态优化。具体步骤如下：

步骤1 输入预处理。为确定飞行员意图，需要对模糊推理系统的输入进行预处理。瞬时输入幅值即飞行员横向操纵$\delta_{\mathrm{lat,p}}$的绝对值$\left|\delta_{\mathrm{lat,p}}\right|$；历史平均输入幅值取过去$T_{\mathrm{lat}}$内输入时间历程绝对值的平均值$\left|\overline{\delta}_{\mathrm{lat,p}}\right|$（见式(9)）；吊挂摆动能量通过等效摆角$\beta_{\mathrm{c,e}}$定义，即对应摆动角速率为0的最大摆角，同时为避免摆角速度突然变化引起的摆动能量突然增加，对$\beta_{\mathrm{c,e}}$取过去T_β的平均值$\overline{\beta}_{\mathrm{c,e}}$（见式(10)）。将上述参数输入模糊推理系统，完成输入的预处理部分。

$$\left|\overline{\delta}_{\mathrm{lat,p}}\right|=\frac{1}{T_{\mathrm{lat}}}\int_{t-T_{\mathrm{lat}}}^{t}\left|\delta_{\mathrm{lat,p}}\right|(\tau)\,\mathrm{d}\tau \tag{9}$$

$$gl_{\mathrm{c}}(1-\cos\beta_{\mathrm{c}})+\frac{1}{2}\left(l_{\mathrm{c}}\dot{\beta}_{\mathrm{c}}\right)^2=gl_{\mathrm{c}}\left(1-\cos\beta_{\mathrm{c,e}}\right)\Rightarrow$$

$$\beta_{\mathrm{c,e}}=\begin{cases}\dfrac{\pi}{2} & \cos\beta_{\mathrm{c}}-\dfrac{l_{\mathrm{c}}}{2g}\dot{\beta}_{\mathrm{c}}^2\leqslant 0\\ \arccos\left(\cos\beta_{\mathrm{c}}-\dfrac{l_{\mathrm{c}}}{2g}\dot{\beta}_{\mathrm{c}}^2\right) & \text{其他}\end{cases}\Rightarrow$$

$$\overline{\beta}_{\mathrm{c,e}}=\frac{1}{T_\beta}\int_{t-T_\beta}^{t}\beta_{\mathrm{c,e}}(\tau)\,\mathrm{d}\tau \tag{10}$$

步骤2 模糊推理系统。模糊推理系统由模糊化、模糊规则和去模糊化3步组成。先通过模糊化将等效摆角和飞行员操纵转化为模糊集；然后，将模糊规则引用于模糊集，得到飞行员意图；最后，通过去模糊化得到与飞行员意图一致的吊挂摆动角速度增益。

本文定义了2个模糊输入集和一个模糊输出集，如图9所示。2个模糊输入集是横向摆动角度

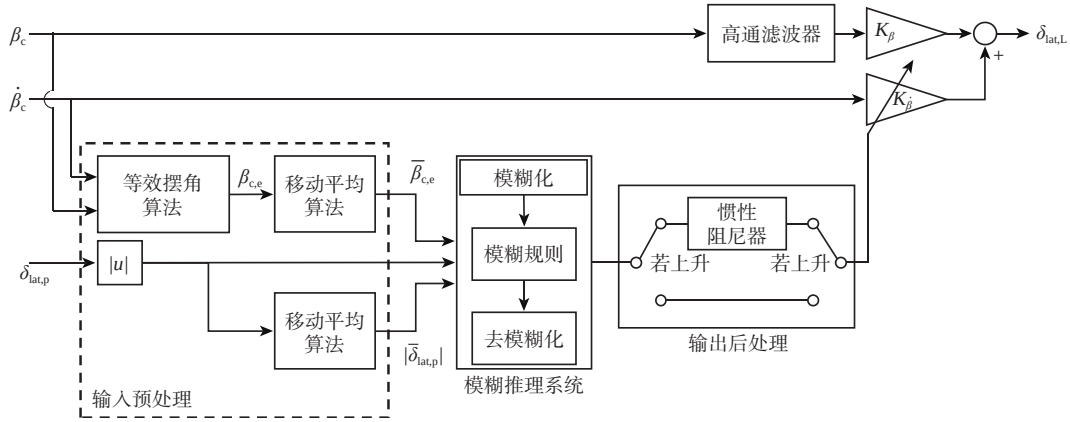


图 8　飞行员意图模糊推理系统

Fig. 8　Pilot intention fuzzy reasoning system

和飞行员横向输入, 而模糊输出集是横向摆动角速度增益。摆动角度被归类为 VS(非常小)、S(小)、M(中等)、L(大)、VL(非常大)。同样, 飞行员输入被归类为 S(小)、M(中等)和 L(大)。摆动角速度增益被归类为增强指令跟踪(command, CMD)、平衡(balanced, BLC)和稳定吊挂摆动(stability, STB), 对应飞行员的 3 种意图。

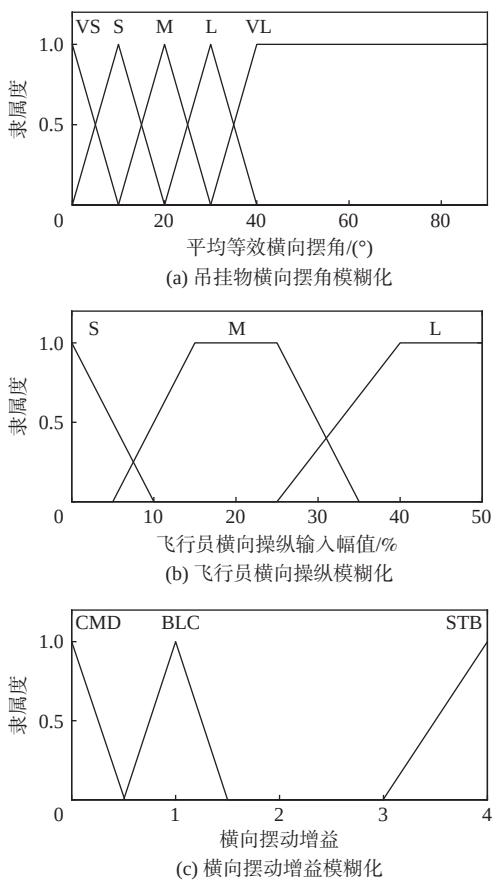


图 9　模糊化

Fig. 9　Fuzzification

上述模糊集的成员函数一部分来自专家知识和经验。例如, 超过 30°的摆动角度被视为大, 而低于 10°的摆动角度被视为小。对于飞行员输入, 超过±25%的输入被视为大, 而低于±10%的输入被视为小。需要注意的是, 本节取值是本文给出的名义值, 实际中需要统计飞行员的评价并确定与之匹配的模糊集。另一部分来自于飞控增益对直升机吊挂飞行特性的影响分析, 根据图 6 和图 7 的飞行品质分析结果可知, 当吊挂摆动角速度增益 $K_{\dot{\beta}}=4$时, 吊挂摆动稳定性最好, 但指令跟踪性能很差; 而当 $K_{\dot{\beta}}=0$时, 指令跟踪性能很好, 但会出现吊挂摆动不稳定性。因此, 若飞行员意图是增强指令跟踪, 则 $K_{\dot{\beta}}<0.5$; 若飞行员意图是稳定吊挂摆动, 则 $K_{\dot{\beta}}>3$。

在本节模糊集的基础上, 模糊推理系统根据吊挂等效摆角和飞行员输入, 使用隶属度函数确定它们是“小”还是“大”, 再应用模糊规则确定驾驶员的意图。在此之前, 首先明确关于人类飞行员的 2 个基本原则:

1) 飞行员对操纵性的需求随着输入强度的增加而增加。当飞行员施加更大操纵输入时, 通常反映了飞行员想要精确控制直升机运动的意图。

2) 飞行员对稳定性的要求随着吊挂摆动变大而增加。且当吊挂剧烈摆动时, 无论此时飞行员输入如何, 其意图都是使吊挂尽快稳定。

基于以上基本原则, 确定飞行员意图模糊规则如表 2 所示。举例而言, 当吊挂物基本上没有摆动而飞行员操纵杆量突然变大时, 飞行员更倾向指令跟踪性能; 而当吊挂物大幅摆动且飞行员操纵杆量减少时, 飞行员更倾向吊挂减摆。飞控系统根据上述意图模糊推理结果实时调整 $K_{\dot{\beta}}$, 当飞行员更倾向指令跟踪性能时, 将 $K_{\dot{\beta}}$ 调整至接近 CMD 模式; 当飞行员更倾向吊挂减摆时, 将 $K_{\dot{\beta}}$ 调整至接近 STB 模式。

通过应用模糊规则, 可以确定飞行员意图, 再根据图 9(c) 中的输出模糊集可确定 $K_{\dot{\beta}}$, 如下:

$$K_{\dot{\beta}}=\sum_{i=1}^{N} v_i S_i \Big/ \sum_{i=1}^{N} S_i \tag{11}$$

式中: N、S_i 和 v_i 分别为模糊等级、隶属度和模糊值。

表 2 飞行员意图模糊规则

Table 2 Fuzzy rule of pilot intention

时均操纵	瞬时操纵	摆动角速度增益分类				
		VS	S	M	L	VL
S	S	STB	STB	STB	STB	STB
	M	CMD	CMD	BLC	BLC	STB
	L	CMD	CMD	CMD	CMD	STB
M	S	STB	STB	STB	STB	STB
	M	CMD	CMD	BLC	STB	STB
	L	CMD	CMD	CMD	BLC	STB
L	S	STB	STB	STB	STB	STB
	M	CMD	BLC	BLC	STB	STB
	L	CMD	CMD	BLC	BLC	STB

步骤 3 输出后处理。为防止吊挂摆动角速度增益K_β突然增加对飞行员的反直觉影响，在模糊输出后加入了增益上升阻尼器，为一阶惯性环节，以实现K_β的缓慢增加。相比之下，若K_β减少，则直接提供给控制器，而不经过阻尼环节，以快速地实现指令跟踪。

3 人-机-吊挂闭环响应分析

为验证第 2 节建立的模糊减摆控制器的有效性，基于本文建立的人-机-吊挂耦合闭环系统模型，对模糊减摆控制器进行方波输入响应分析和横侧定位科目数值模拟响应分析。采用的控制器参数即第 2 节给出的参数，由于这组控制器参数基于悬停的平衡状态设计，所以适用于悬停和低速飞行科目及状态。

3.1 方波输入响应分析

为解释融合飞行员意图的模糊减摆控制器的作用机理，图 10 为飞行员输入持续 5 s 的横向方波操纵下的响应，以及模糊减摆控制器的输入-输出信号。图中，BSL 表示指令跟踪性能最好的控制系统，即不引入吊挂摆角反馈，采用直升机原有飞行控制系统；CAF 表示仅仅引入了吊挂摆动角速度反馈的控制器；FUZZY 表示本文提出的融合飞行员意图的模糊减摆控制器。

由图 10(a) 和图 10 (b) 中的直升机滚转响应可知，尽管 BSL 可以跟上期望的滚转响应，但会产生持续横向摆动振荡。CAF 尽管能够成功地抑制横向摆动振荡，但相应的滚转响应与预期响应偏离较大，导致指令跟踪性能较差。相比之下，本文的 FUZZY 既能跟踪期望的响应，又能抑制吊挂摆动。

如图 10(c) 和图 10((d) 所示，FUZZY 的优势是通过根据飞行员的瞬时输入和移动平均输入及平均摆角推断出飞行员意图，再实时调整横向摆动角速率增益来实现的。当飞行员输入在 10 s 附近从

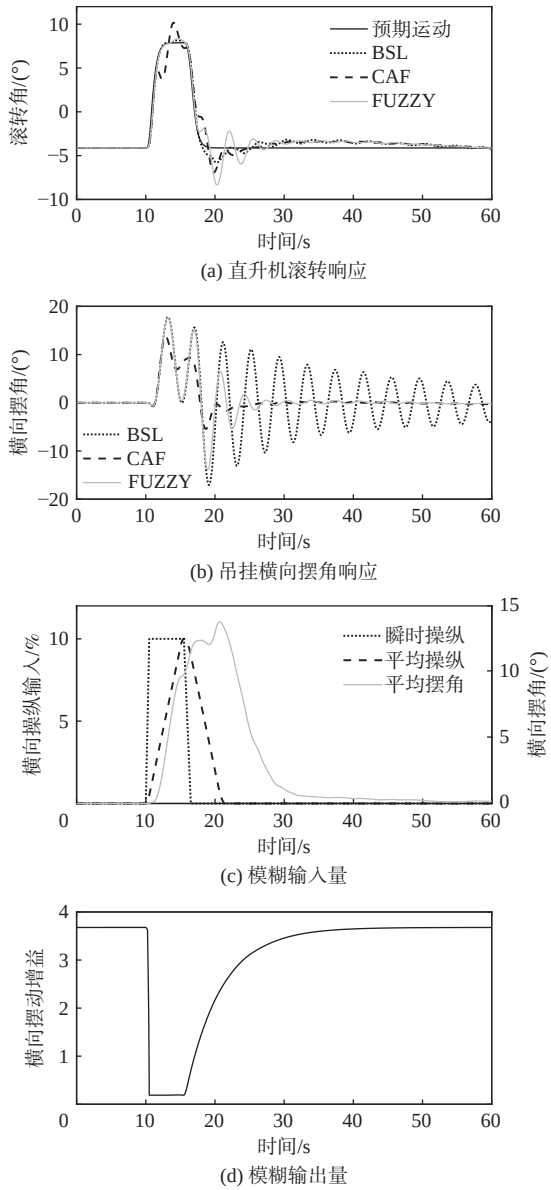


图 10 方波输入下的响应和模糊输入/输出

Fig. 10 Response and fuzzy input/output time history under square wave input

0% 增加到 10% 时，横向摆动角速率增益迅速降低至接近 0，以提供更好的跟踪性能，此时移动平均输入和移动平均摆角随时间逐渐增大。当飞行员输入从 10% 减少到 0% 时，横向摆动角速率增益开始缓慢增加，在操纵完成后稳定吊挂物，与飞行员意图一致。此处横向摆动角速率增益缓慢上升而非迅速上升，可以让飞行员逐渐适应随增益增加而逐渐变差指令跟踪性能。

然而，尽管 FUZZY 具有 CAF 和 BSL 的优点，但是与 CAF 相比，FUZZY 稳定横向摆角所需时间延长(见图 10(b))。而且，当输入减小到 0% 后，直升机滚转角的最大超调量更大(见图 10(a))。其原因是：一旦飞行员增加输入，横向增益就会迅速且积极地减小，并如 BSL 一样快速激发吊挂摆动，导

致操纵结束后的吊挂摆动幅值更大, 所以需要更多的时间来稳定; 并且, 由于吊挂摆动抑制需要通过直升机滚转运动, 由于吊挂摆动更大, 也就需要更大的滚转超调来使之稳定。此外, 由于增益上升阻尼器的存在, 在输入减小到 0% 后, 横向增益缓慢增长, 使得吊挂摆动阻尼增长变慢, 所需的吊挂摆动稳定时间增加了 5～10 s。

总而言之, FUZZY 的不利响应特征只会持续较短时间, 且与 BSL 引起持续振荡和 CAF 存在的指令跟踪性能差的缺点相比, FUZZY 优势明显。

3.2　横侧定位科目数值模拟响应分析

本文基于人-机-吊挂耦合闭环系统模型开展横侧定位科目的数值模拟。该科目是将直升机从悬停位置移动到 122 m 的横向位置。本文调整飞行员模型增益为高增益激进型以充分暴露潜在的飞行品质缺陷。根据图 11(a) 可知, 直升机最终都能够维持在目标的横向位置。然而, 如图 11(b) 的吊挂摆角响应所示, BSL 结果中一直存在幅值约为 5°的吊挂摆动振荡, 难以收敛, 该响应特征与方波输入结果类似; 在 CAF 结果中, 在抵达目标位置时(约 30 s)吊挂摆动已收敛, 但随着时间推进, 摆动幅值不断变大; 相比之下, 在 FUZZY 结果中, 当直升机抵达目标位置后吊挂摆动振荡一直较小, 说明本文的融合飞行员意图的模糊减摆控制器在吊挂减摆方面更具优势。

进一步对比 BSL、CAF 和 FUZZY 的仿真结果可发现, 在仿真开始时, FUZZY 仿真结果与 BSL 类似, 以符合飞行员在施加操纵时希望指令跟踪性能好的意图; 自 15 s 后, FUZZY 仿真结果更加接近 CAF 的结果, 以符合飞行员在抵达目标位置后试图抑制吊挂物摆动的意图; 自 35 s 后, CAF 的结果中出现了振荡逐渐增加的现象, 结合图 11(c) 和图 11 (d) 中的直升机滚转角和飞行员输入可知, 该现象是因为飞行员输入与吊挂摆动同步放大, 出现了飞行员诱发振荡, 导致直升机滚转与吊挂物横向摆动耦合振荡, 但 FUZZY 的仿真结果中没有出现飞行员诱发振荡。

以下分析模糊控制器与飞行员意图的一致性。由图 11(e) 和图 11 (f) 可知, 在 5 s 时刻, 飞行员瞬时操纵幅值超过了 10% 且大于平均操纵, 同时吊挂平均摆角较小, 这意味着飞行员需要更强的指令跟踪性能, 于是横向摆动增益相应地快速调整至接近 0 以提高指令跟踪效果; 从 10 s 开始, 瞬时操纵小于平均操纵, 且平均摆角较大, 意味着飞行员需要更强的吊挂摆动抑制效果, 所以横向摆动增益缓

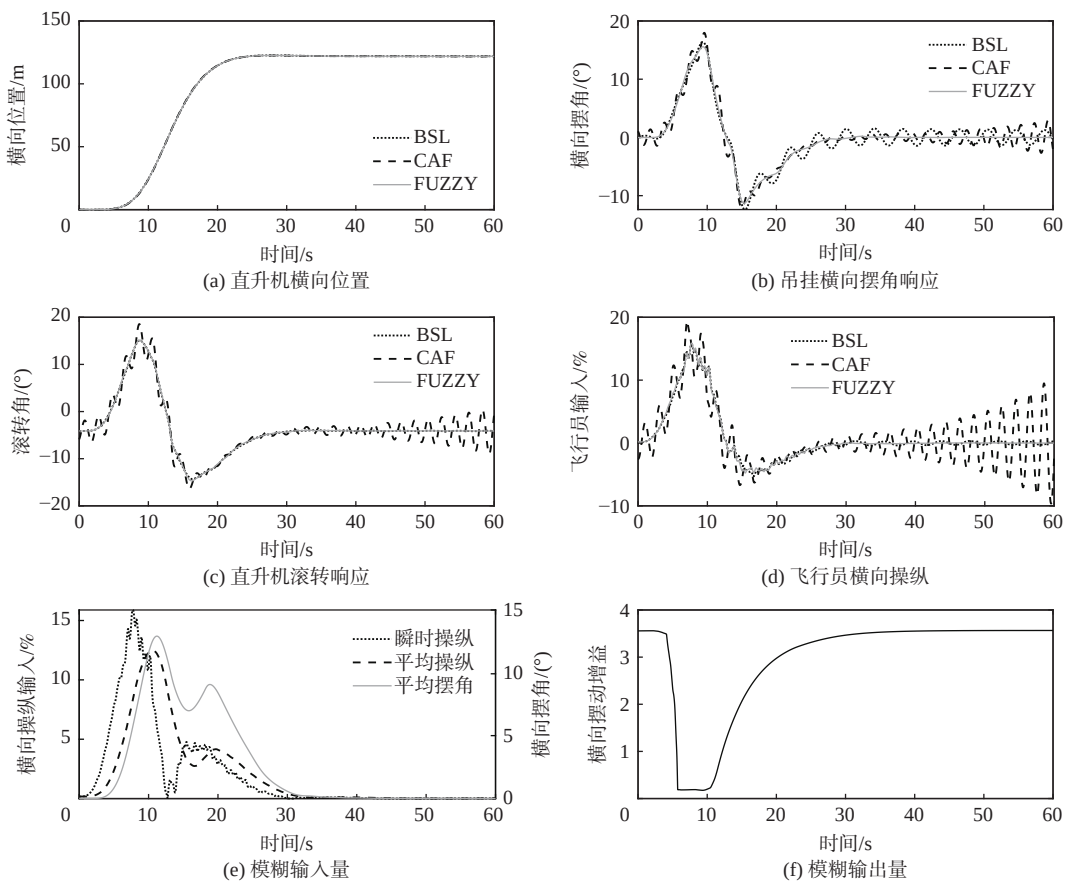


图 11　横侧定位科目响应和操纵时间历程

Fig. 11　Response and control histories of lateral reposition task

慢增加，在约 20 s 超过 3 并逐渐趋于稳定，且最大值约为 3.5，比 CAF 的横向摆动增益小，所以没有出现飞行员诱发振荡。

此外，CAF 的仿真结果中，飞行员在整个飞行任务过程中持续修正操纵输入，幅值约为 5%，极大增加了飞行员的操纵负荷。相比之下，FUZZY 的仿真结果中，飞行员操纵较为光滑，与 BSL 的操纵历程类似，证明了本文 FUZZY 控制器同时具有消除吊挂摆动和指令跟踪性能好的优势。

为验证融合飞行员意图的模糊减摆控制器对于不同飞行员均有良好的吊挂减摆效果和指令跟踪性能，本文采用蒙特卡罗方法，对表 1 中的飞行员模型参数进行随机调整，幅值为±25%，仿真结果如图 12 所示，图中 1 条曲线对应 1 次仿真，可发现吊挂横向摆动均得到了有效抑制，而且未出现飞行员诱发振荡，表明本文方法具有鲁棒性。

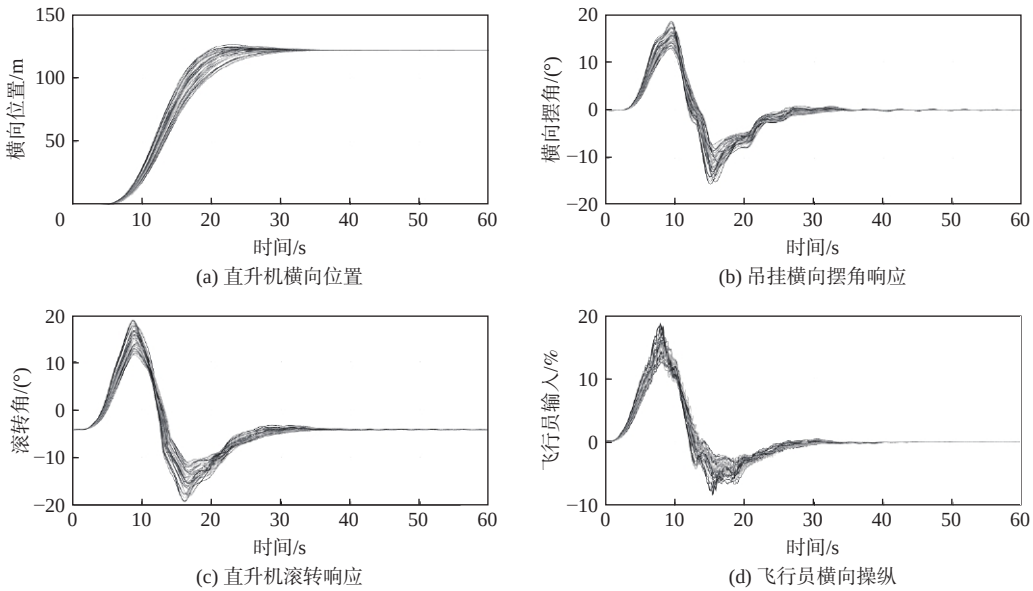


图 12 蒙特卡罗鲁棒性验证结果

Fig. 12 Monte-Carlo robustness verification results

为验证融合飞行员意图的模糊减摆控制器对于测量时间延迟的鲁棒性，计算了不同摆角测量延迟对仿真结果的影响，如图 13 所示，可发现当测量延迟达到 1 000 ms 时会出现无法衰减的振荡，但延迟小于 500 ms 就可将吊索摆角抑制在±0.5 °幅度内。由于传感器测量延迟至多在 100 ms 量级，本文模糊减摆控制器对于传感器测量延迟也具有鲁棒性。

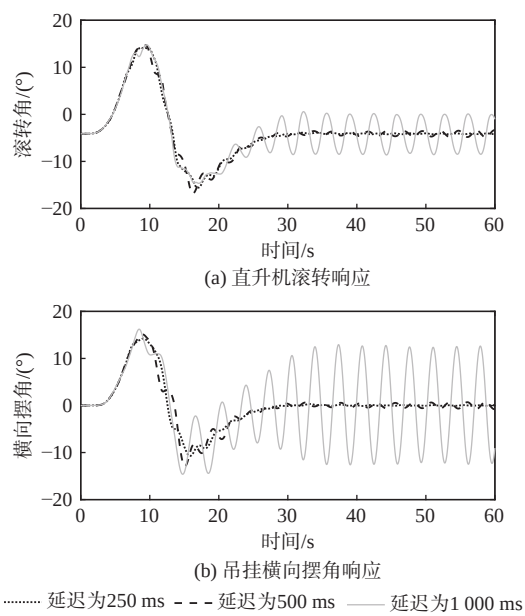


图 13 摆角测量延迟对仿真的影响

Fig. 13 Effects of swing angle measurement delay on simulation results

4 结 论

本文构建了非线性的人-机-吊挂耦合闭环系统模型，提出一种融合飞行员意图的模糊减摆控制器，解决了目前的吊挂减摆控制器难以兼顾吊挂减摆效果和指令跟踪性能的问题，有助于提高有人直升机吊挂飞行品质。主要结论如下：

1）融合飞行员意图的模糊减摆控制器能够以符合飞行员预期的方式动态调整指令跟踪性能和吊挂摆动抑制效果，且鲁棒性较好；

2）基于模糊推理得到的飞行员意图能够合理反映飞行员在吊挂飞行任务中对于指令跟踪性能和吊挂摆动抑制效果的倾向性；

3）与单纯的吊挂减摆控制器相比，融合飞行员意图的模糊减摆控制器需要相对更长的时间使吊挂摆动稳定，但可以避免飞行员与减摆控制器的相互冲突引起的飞行员诱发振荡现象。

参考文献（References）

[1] ENCIU J, SINGH A, HORN J F. Stabilization of external loads in high-speed flight using an active cargo hook[J]. Journal of the American Helicopter Society, 2020, 65(2): 1-12.

[2] RAZ R, ROSEN A, CARMELI A, et al. Wind tunnel and flight evaluation of passive stabilization of a cargo container slung load[J]. Journal of the American Helicopter Society, 2010, 55(3): 32001.

[3] IVLER C M, POWELL J D, TISCHLER M B, et al. Design and flight test of a cable angle feedback flight control system for the RASCAL JUH-60 helicopter[J]. Journal of the American Helicopter Society, 2014, 59(4): 1-15.

[4] ADAMS C, POTTER J, SINGHOSE W. Input-shaping and model-following control of a helicopter carrying a suspended load[J]. Journal of Guidance, Control, and Dynamics, 2014, 38(1): 94-105.

[5] 张松, 马钊, 张辉. 直升机吊挂飞行中载荷摆动控制方法研究[J]. 航空工程进展, 2022, 13(2): 57-63.
ZHANG S, MA Z, ZHANG H. Research on slung-load swing control method in helicopter suspension flight[J]. Advances in Aeronautical Science and Engineering, 2022, 13(2): 57-63(in Chinese).

[6] REN Y, ZHAO Z J, AHN C K, et al. Adaptive fuzzy control for an uncertain axially moving slung-load cable system of a hovering helicopter with actuator fault[J]. IEEE Transactions on Fuzzy Systems, 2022, 30(11): 4915-4925.

[7] LIU W, CHEN M. An adaptive anti-swing control for the helicopter slung-load system based on trajectory planning and neural network[J]. International Journal of Adaptive Control and Signal Processing, 2022, 36(5): 1116-1140.

[8] LIU L J, CHEN M, LI T. Disturbance observer-based LQR tracking control for unmanned autonomous helicopter slung-load system[J]. International Journal of Control, Automation and Systems, 2022, 20(4): 1166-1178.

[9] ENCIU J, ROSEN A. Simulation of coupled helicopter-slung load-pilot dynamics[J]. Journal of the American Helicopter Society, 2017, 62(1): 1-13.

[10] PATTERSON B, ENNS R, KING C, et al. Design and flight test of a hybrid external load stabilization system for an H-6 helicopter testbed[C]//Proceedings of the Vertical Flight Society 71st Annual Forum. Virginia Beach: The Vertical Flight Society, 2015: 1-16.

[11] ROBERTSON G. Flight demonstration of an associate system: a rotorcraft pilot's associate example[C]//Proceedings of the American Helicopter Society 56 rd Annual Forum. Virginia Beach: The Vertical Flight Society, 2000: 431-445.

[12] D'INTINO G, OLIVARI M, BÜLTHOFF H H, et al. Haptic assistance for helicopter control based on pilot intent estimation[J]. Journal of Aerospace Information Systems, 2020, 17(4): 193-203.

[13] LUO Y, WEI Y H, KAN Y Z, et al. Control strategy for electric startup of P2.5-PHEV based on slope memory and driver's startup intention[J]. IEEE Access, 2021, 9: 77044-77057.

[14] 程建锋, 董新民, 薛建平, 等. 飞机-驾驶员闭环系统模糊预见控制器设计[J]. 航空学报, 2014, 35(3): 807-820.
CHENG J F, DONG X M, XUE J P, et al. Fuzzy preview controller design for aircraft-pilot closed loop system[J]. Acta Aeronautica et Astronautica Sinica, 2014, 35(3): 807-820(in Chinese).

[15] TRAN D, DU J H, SHENG W H, et al. A human-vehicle collaborative driving framework for driver assistance[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(9): 3470-3485.

[16] United States Army Aviation and Missile Command. Aeronautical design standard performance specification handling qualities requirements for military rotorcraft: ADS-33E-PRF[S]. Washington, D. C. : United States Army Aviation and Missile Command, 2000.

[17] CICOLANI L S, MCCOY A H, SAHAI R, et al. Flight test identification and simulation of a UH-60A helicopter and slung load[J]. Journal of the American Helicopter Society, 2001, 46(2): 140-160.

[18] HOWLETT J J. UH-60A black hawk engineering simulation program: volume I-mathematical model: NASA-CR-166309[R]. Washington, D. C. : National Aeronautics and Space Administration, 1981.

[19] CHEN R L, YUAN Y, THOMSON D. A review of mathematical modelling techniques for advanced rotorcraft configurations[J]. Progress in Aerospace Sciences, 2021, 120: 100681.

[20] WANG L F, CHEN R L, YAN X F. Trajectory optimization of aerial slung load release for piloted helicopters[J]. Chinese Journal of Aeronautics, 2021, 34(2): 229-239.

[21] WANG L F, CHEN R L. Nonlinear helicopter rigid-elastic coupled modeling with its applications on aeroservoelasticity analysis[J]. AIAA Journal, 2021, 60(1): 102-112.

[22] TIAN Y, WANG L F, ZHOU Z L, et al. Nonlinear rigid-elastic coupled modeling and oscillation mechanism analysis of rotor-body-slung-load system[J]. Aerospace, 2023, 10(10): 1-14.

[23] HESS R A. Unified theory for aircraft handling qualities and adverse aircraft-pilot coupling[J]. Journal of Guidance, Control, and Dynamics, 1997, 20(6): 1141-1148.

[24] BLANKEN C L, TISCHLER M B, LUSARDI J A, et al. Proposed revisions to aeronautical design standard-33E (ADS-33E-PRF) toward ADS-33F-PRF: SR-FCDD-AMV-19-01[R]. Washington, D. C.: United States Army Aviation Development Directorate, 2019: 148.

Fuzzy anti-swing controller for improving handling quality of helicopter slung-load operation

WANG Luofeng[1]，CHEN Renliang[1,*]，ZHAO Yu[2]

(1. National Key Laboratory of Science and Technology on Rotorcraft Aeromechanics，College of Aerospace Engineering，Nanjing University of Aeronautics and Astronautics，Nanjing 210016，China；
2. Key Laboratory of Rotor Aerodynamics，China Aerodynamics Research and Development Center，Mianyang 621000，China)

Abstract：In order to solve the problem of compromised flight quality in helicopter sling load operations—which arises from the trade-off between load stabilization and command responsiveness—this paper presented a unique fuzzy control solution that incorporates pilot intention. By developing a high-fidelity nonlinear model of the pilot-helicopter-slung-load system, this method tackles the inherent conflict between reducing load swing and maintaining precise control that is common in current anti-swing controllers. The proposed method can reasonably reflect the pilot's intent, and a cable angle feedback control law has been demonstrated through simulation to dynamically adjust control parameters in line with pilot expectations. With good resilience, this integration reduces pilot-induced oscillations and conflicts between pilot actions and the control system.

Keywords：helicopter；helicopter slung-load；fuzzy logic；pilot intention；cable angle feedback；pilot induced oscillation

Received：2024-04-29；**Accepted**：2024-08-16；**Published Online**：2024-09-02 16：33

URL：link.cnki.net/urlid/11.2625.V.20240902.1521.001

Foundation items：Fund of National Key Laboratory of Science and Technology on Rotorcraft Aeromechanics(2023-HA-LB-067-01); China Postdoctoral Science Foundation(2023M741678); Fund of Key Laboratory of Rotor Aerodynamics(RAL202302-4); Jiangsu Funding Program for Excellent Postdoctoral Talent; A Project Funded by Priority Academic Program Development of Jiangsu Higher Education Institutions

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0231

基于迁移学习的自动调制分类方法

王栋[1,2]，崔天舒[3]，姬丽彬[1,2]，黄永辉[1]，朱岩[1,*]

(1. 中国科学院国家空间科学中心 复杂航天系统电子信息技术重点实验室，北京 100190；2. 中国科学院大学 计算机科学与技术学院，北京 100049；3. 中国航天科技创新研究院，北京 100190)

摘 要：深度学习在调制分类上的应用取得了很大进展，但受到训练数据和测试数据分布一致性的限制。为解决源域和目标域采样率不同导致的数据分布不一致的调制分类问题，提出一种新颖的多尺度注意力迁移学习架构(MATLA)来解决跨域识别问题，该架构利用3个具有不同尺度的并行卷积核来提取不同粒度级别特征。此外，该过程还融入了多尺度注意力机制，这一机制通过强化重要特征的权重，增强了提取具有辨别力的特征的能力。为实现源域与目标域特征的有效对齐，使用多核最大均值差异(MK-MMD)来度量再生核希尔伯特空间(RKHS)中2个域的特征分布的差异程度。为有效减少源域特征内部的变化，增强特征的一致性和稳定性，提出多核中心损失(MKCL)。实验结果表明：MATLA在信噪比大于0 dB时，识别准确率达到83.42%，优于其他几种网络模型和域适应模型。

关 键 词：迁移学习；注意力机制；深度学习；调制分类；采样频率
中图分类号：TN911.3; V443.1
文献标志码：A **文章编号**：1001-5965(2026)06-1935-09

作为信号接收和解调之间的中间步骤，自动调制分类(automatic modulation classification, AMC)是军用和民用通信中的一项关键技术[1-2]。其可应用于电子战、认知无线电和频谱监控，有助于识别调制方案并据此对接收信号进行解调。因此，可靠有效的AMC方案至关重要。

传统的自动调制分类方法包括基于最大似然假设检验的方法[3]和基于特征的方法[4-5]。前者需要大量的先验知识，并且其性能可能会受到参数估计偏差或模型与实际信道特性之间不匹配的显著影响。后者涉及手动选择特征，如小波变换[4]、高阶统计量[5]和循环谱[6]等。然而，在复杂的通信环境中，很难确定哪种特征组合对于分类最有效。因此，传统方法的通用性有限。

近年来，随着深度学习在语音[7]、图像[8]等领域的巨大成功，将深度学习应用于自动调制分类已成为研究热点。与传统手工设计的特征难以捕捉复杂的数据模式相比，深度学习方法具有更好的表示学习能力。文献[9]使用卷积神经网络(convolutional neural networks,CNN)从输入的原始IQ数据中提取逐层特征，在RadioML2016.10a数据集[10]上，当信噪比(signal-to-noise ratio, SNR)大于0 dB时，平均识别准确率大于78%。此外，文献[11]中使用了残差网络(ResNet)。当信噪比超过5 dB时，表现出良好的分类效果。进一步地，文献[12]使用CNN和循环神经网络(recurrent neural networks, RNN)相结合的卷积长短期深度神经网络(convolutional long short-term deep neural networks,CLDNN)架构。类似地，文献[13]提出CNN与长短期记忆(long short-term memory, LSTM)网络级联的空时融合网络(space-time fusion networks,STF-Net)，目的是同时提取空域和时域的特征。此外，文献[14]通过将IQ信号转换

收稿日期：2024-04-18；录用日期：2024-07-30；网络出版时间：2024-08-02 14:29
网络出版地址：link.cnki.net/urlid/11.2625.V.20240802.1334.001
*通信作者. E-mail：zhuyan@nssc.ac.cn

引用格式：王栋，崔天舒，姬丽彬，等. 基于迁移学习的自动调制分类方法[J]. 北京航空航天大学学报，2026，52(6)：1935-1943.
WANG D，CUI T S，JI L B，et al. Automatic modulation classification method based on transfer learning[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1935-1943 (in Chinese).

为幅度和相位表示，然后将其输入LSTM网络，实现了显著的识别精度。文献[15]将原始IQ信号转换为星座图，利用AlexNet和GoogLeNet来执行分类任务。文献[16]采用2个独立的CNN，第1个网络利用原始IQ信号作为输入并识别7种调制方式，而第2个网络处理输入的正交幅度调制（quadrature amplitude modulation, QAM）星座图以识别16QAM和64QAM。文献[17]基于非对称卷积和残差连接的思想设计网络来从时频图中提取特征。

然而，值得注意的是，以上深度网络架构的性能改进依赖于训练数据的分布与测试数据的分布相同的假设。当训练和测试数据分布不匹配时，网络模型的性能将出现下降[18]。当训练数据和测试数据以不同的速率采样时，这种现象在调制识别任务中尤其明显。然而，在现实世界的无线通信环境中，获取带标签的变采样率数据来重新训练模型以解决分布不匹配的问题通常具有挑战性。

鉴于上述挑战，迁移学习已成为解决这些问题的有效方法[19]。其目的是通过利用相关领域的标记或未标记数据或利用领域不变的特征表示来最小化训练和测试数据之间的分布差异。通过使模型适应目标域从不同但相关的领域迁移知识，迁移学习增强了模型的泛化能力并提高了模型对于未知分布的测试数据的性能。文献[20]基于模型无关的元学习算法，首先，在辅助数据集上进行预训练；然后，保留网络的特征提取部分，同时设置2个可学习的参数对预训练的模型进行缩放和偏移；最后，使用少量的目标域数据进行模型的微调。文献[21]设计一种迁移学习机制，通过迁移同领域数据集的知识，显著提升了CNN的特征提取能力，从而大幅提高了信号聚类性能。文献[22]改进了传统识别网络，并加入了域分类子网络。同时在训练代价中加入了域分类代价，目的是让网络能够同时适应目标域和源域的数据。文献[23]使用参数迁移方法来解决2个域采样率不一致的问题，首先，利用源域数据训练模型，然后，将模型的前几层参数复制到目标网络，再利用带标记的目标域样本重新训练模型。文献[24]提出了自动调制分类的对抗性迁移学习架构，该架构采用对抗性训练，来挖掘源域中的先验知识，以减小不同采样率信号之间的分布差异。然而，这些方法仍然需要一部分标记的目标域数据来再次训练模型或微调模型。为解决无标签目标域中的域适应问题，无监督的领域自适应方法被提出。文献[25]提出了最大分类器差异（maximum classifier discrepancy, MCD），通过最大化2个分类器的差异来获得鲁棒的判别特征以解决域分布差异的问题。文献[26]提出了最小类混淆（minimum class confusion, MCC），其试图最小化未标记目标类对象的熵以增加预测的置信度。文献[27]在网络中引入了一个适应层，通过最小化该层源域和目标域之间的最大均值差异（maximum mean discrepancy, MMD）来减少2个域之间的分布差异。在此基础上，文献[28]在神经网络中实现了多层适应，该文献在AlexNet的最后3个全连接层中引入了域间分布差异的度量，同时确保条件概率分布保持不变。

受深度域适应方法的启发，本文提出了一种迁移学习架构来解决未标记目标域中变采样率的调制分类问题。主要贡献如下：

1）本文提出多尺度注意力迁移学习架构（multi-scale attention transfer learning architecture, MATLA），其为基于多尺度注意力网络和设计的损失函数的有效组合。该架构可以提取源域信号和目标域信号之间的深层特征来进行对齐。损失函数包括多核最大均值差异（multiple kernel maximum mean discrepancy, MK-MMD）损失来减少2个域之间的分布差异，以及多核中心损失（multiple kernel center loss, MKCL）以鼓励学习紧凑又具有区分性的特征。

2）本文提出多尺度注意力网络（multi-scale attention networks, MANet）来从2个域中提取特征。该网络结合了不同尺度的注意力机制，专注于更具代表性的特征，同时减少噪声或不相关信息的影响，确保可以提取更稳健和丰富的特征表示。

实验结果证明了本文MATLA的有效性，在未标记的目标域上实现了出色的识别精度。

1 问题描述

接收端收到经过信道作用的信号，可以表示为

$$x(n)=M(s(n))*h(n)+w(n)\qquad n=1,2,\cdots,N \tag{1}$$

式中：$s(n)$为基带信号，由发射机通过调制函数M调制后发射；$h(n)$为信道效应；$w(n)$为噪声；“*”表示卷积。N为采样序列的长度。调制识别的目的是分析接收到的$x(n)$并确定调制函数M的类型。

对于无监督域适应的跨域调制识别问题，源域代表了所获取的知识并具有丰富的标签信息；目标域的知识需要利用源域的知识来学习，并且没有标签信息。源域和目标域中的任务相同，但数据分布不同。由于接收端的硬件设备和要求不同，接收端会以不同的采样率对信号进行采样。即本文中源域和目标域的采样率不同，标签空间一致。源域数据集$D_s=\{X_{s,i},Y_{s,i}\}_{i=1}^{N_s}$具有足够的信号样本$X_s$和标签$Y_s$，采样率为$F_s$。对于未标记的目标域数据集

$D_t=\{X_{t,i}\}_{i=1}^{N_t}$, 采样率为 F_t, 只有未标记的数据可用。需要从标记的源域数据中提取有用的信息来构建分类模型, 以实现对未标记的目标域的正确分类。

2　方法介绍

图 1 为本文 MATLA, 该架构利用 MANet 从源域信号和目标域信号中提取特征, 然后, 将这些特征投影到再生核希尔伯特空间 (reproducing kernel Hilbert space, RKHS) 中以进行域对齐。图 1 中, GAP 表示全局平均池化, FC 表示全连接层。本节将详细解释 MANet 的结构和损失函数的设计。

2.1　多尺度注意力网络

MANet 由 3 个核心模块组成: 多尺度注意力卷积模块、深层特征提取模块和分类器模块。输入大小为 2×128 的二维时域原始 IQ 信号, 包含来自源域的标记数据和来自目标域的未标记数据。使用 3 个不同大小的卷积核并行处理原始 IQ 信号以提取特征, 核大小不同的卷积可以从序列信号中提取具有多种分辨率的局部信息, 因为其代表了各种感受野。每个卷积层后均有批归一化层及非线性激活层, 卷积后产生同一信号的 3 个不同层次的特征图 $\boldsymbol{F}_1$、$\boldsymbol{F}_2$和 $\boldsymbol{F}_3$。多尺度卷积独立地在各自的通道上执行, 假设所有通道都同样重要。然而, 并非所有卷积通道都提供同样有用的信息。为捕获不同尺度和层次结构的信号特征, 同时智能地挑选并关注信息最丰富的部分, 从而减轻冗余信息的影响, 提出了多尺度注意力机制来实现这一目标。图 2 为多尺度注意力模块中 a_1 的计算过程。

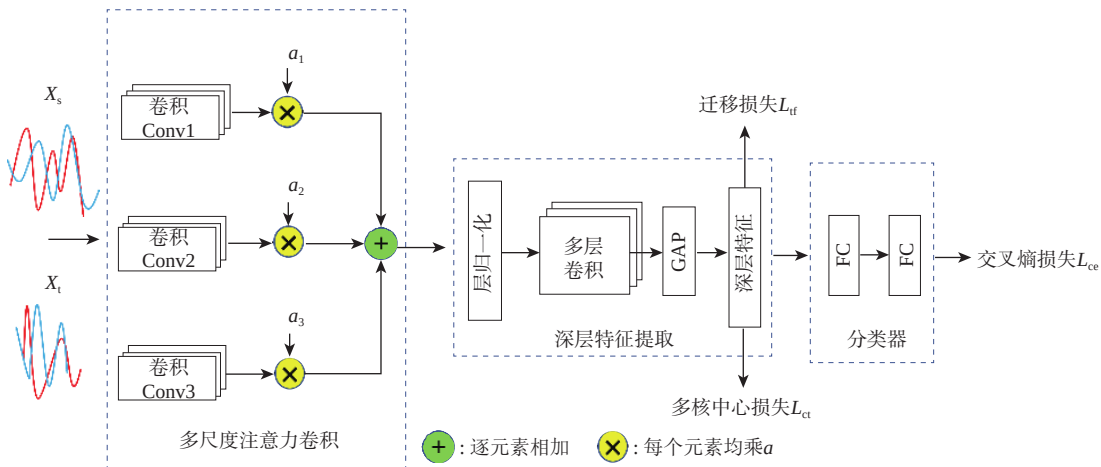


图 1　多尺度注意力迁移学习架构

Fig. 1　Multi-scale attention transfer learning architecture

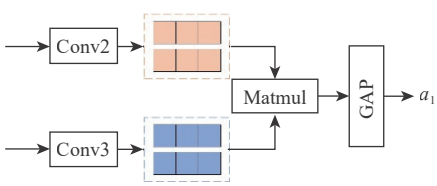


图 2　多尺度注意力模块

Fig. 2　Multi-scale attention module

注意力权重计算如下:

$$a_1=\mathrm{GAP}(\boldsymbol{F}_2\cdot\boldsymbol{F}_3) \tag{2}$$

$$a_2=\mathrm{GAP}(\boldsymbol{F}_1\cdot\boldsymbol{F}_3) \tag{3}$$

$$a_3=\mathrm{GAP}(\boldsymbol{F}_1\cdot\boldsymbol{F}_2) \tag{4}$$

不同特征图中特征向量之间的点积操作, 代表其之间的相关程度, 因此, 得到的注意力权重捕获了信号的局部关联性。通过 GAP 实现信息压缩后, 将注意力权重乘以相应的卷积特征图:

$$\boldsymbol{F}_1'=a_1\boldsymbol{F}_1 \tag{5}$$

$$\boldsymbol{F}_2'=a_2\boldsymbol{F}_2 \tag{6}$$

$$\boldsymbol{F}_3'=a_3\boldsymbol{F}_3 \tag{7}$$

通过逐元素相加的方式对不同尺度空间特征进行融合, 增强模型表达能力, 得到最终的多尺度卷积融合特征图。

在深层特征提取模块中, 采用 LayerNorm[29] 技术来有效地保证特征内的时序关系。LayerNorm 层后采用卷积核大小为 3 的 1D 卷积 (包括 3 个卷积层) 对融合特征执行深层特征提取过程, 并使用 GAP 来提取全局特征表示, 有助于减少模型的参数量和计算量, 同时增强模型的泛化性能。最后, 深层次的全局特征表示被投影到 RKHS 中, 促进源域和目标域信号的对齐。在分类器模块中, 使用 2 个全连接层进行分类。所有卷积层均采用 LeakyRelu 激活函数。

通过利用不同大小的卷积核, MANet 有效地捕获了深层次特征。较小的卷积核有利于提取详细

的局部特征，而较大的卷积核有助于提取采样点和全局结构信息之间的长期依赖性。此外，MANet中采用的多尺度注意力机制带来了许多优势。首先，其使网络能够自主地专注于信号中信息最丰富的部分，同时丢弃冗余信息，从而实现更高效、更准确的特征提取；其次，注意力机制增强了模型的判别力，从而提高了其在IQ时间序列信号分析中的整体性能；最后，MANet中的深层特征提取过程有助于从不同域得到最具代表性的特征，从而促进其之间更有效的知识迁移。

2.2 损失函数

本文采用MK-MMD[30]来量化2个域分布间的差异。MK-MMD计算源域和目标域在RKHS中概率分布P_{s}和P_{t}之间的期望值之差的平方范数，表示为$d_k(P_{\mathrm{s}},P_{\mathrm{t}})$。MK-MMD距离的平方为

$$d_k^2(P_{\mathrm{s}},P_{\mathrm{t}}) \triangleq \|E_{P_{\mathrm{s}}}[\phi(\boldsymbol{x}_{\mathrm{s}})]-E_{P_{\mathrm{t}}}[\phi(\boldsymbol{x}_{\mathrm{t}})]\|_{\mathcal{H}}^2 \tag{8}$$

式中：$\boldsymbol{x}_{\mathrm{s}}$和$\boldsymbol{x}_{\mathrm{t}}$分别为源域和目标域的深层特征；$E_P[\cdot]$为RKHS中域分布的数学期望；$\phi(\boldsymbol{x})$为映射函数；$\mathcal{H}$为希尔伯特空间。

本文将源域和目标域之间的MK-MMD距离作为迁移损失，表达式为

$$L_{\mathrm{tf}} = d_k^2(\boldsymbol{x}_{\mathrm{s}},\boldsymbol{x}_{\mathrm{t}}) \tag{9}$$

对于所有的样本特征$\boldsymbol{x}_{\mathrm{s}}$、$\boldsymbol{x}_{\mathrm{t}}$，映射函数$\phi(\boldsymbol{x})$将输入特征空间映射到希尔伯特空间，核函数$k(\boldsymbol{x}_{\mathrm{s}},\boldsymbol{x}_{\mathrm{t}})$满足内积$k(\boldsymbol{x}_{\mathrm{s}},\boldsymbol{x}_{\mathrm{t}})=\langle\phi(\boldsymbol{x}_{\mathrm{s}}),\phi(\boldsymbol{x}_{\mathrm{t}})\rangle$，即特征核与特征映射$\phi$相关，那么迁移损失的表达式为

$$\begin{aligned}L_{\mathrm{tf}} &= \left\|\frac{1}{n}\sum_{i=1}^{n}\phi(\boldsymbol{x}_{\mathrm{s},i})-\frac{1}{m}\sum_{j=1}^{m}\phi(\boldsymbol{x}_{\mathrm{t},j})\right\|_{\mathcal{H}}^2 = \\ &\left\|\frac{1}{n^2}\sum_{i=1}^{n}\sum_{i'=1}^{n}\phi(\boldsymbol{x}_{\mathrm{s},i})\phi(\boldsymbol{x}_{\mathrm{s},i'})-\frac{2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\phi(\boldsymbol{x}_{\mathrm{s},i})\phi(\boldsymbol{x}_{\mathrm{t},j})+\frac{1}{m^2}\sum_{j=1}^{m}\sum_{j'=1}^{m}\phi(\boldsymbol{x}_{\mathrm{t},j})\phi(\boldsymbol{x}_{\mathrm{t},j'})\right\|_{\mathcal{H}} = \\ &\left\|\frac{1}{n^2}\sum_{i=1}^{n}\sum_{i'=1}^{n}k(\boldsymbol{x}_{\mathrm{s},i},\boldsymbol{x}_{\mathrm{s},i'})-\frac{2}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}k(\boldsymbol{x}_{\mathrm{s},i},\boldsymbol{x}_{\mathrm{t},j})+\frac{1}{m^2}\sum_{j=1}^{m}\sum_{j'=1}^{m}k(\boldsymbol{x}_{\mathrm{t},j},\boldsymbol{x}_{\mathrm{t},j'})\right\|_{\mathcal{H}}\end{aligned} \tag{10}$$

多核被定义为一系列半正定核的线性组合[28]，其形式为

$$K \triangleq \left\{k=\sum_{u=1}^{v}\beta_u k_u:\ \sum_{u=1}^{v}\beta_u=1\right\} \qquad \beta_u\geqslant 0,\forall u \tag{11}$$

式中：β_u为多核的权重。本文中采用高斯核函数，其表达式为

$$k(\boldsymbol{x}_{\mathrm{s}},\boldsymbol{x}_{\mathrm{t}}) = \exp\left(-\frac{\|\boldsymbol{x}_{\mathrm{s}}-\boldsymbol{x}_{\mathrm{t}}\|^2}{2\sigma^2}\right) \tag{12}$$

式中：σ为高斯核函数的宽度。

为使得提取的不同类别的特征具有明显的区分度，促使同类的特征紧凑地聚在一起，提高网络模型的分类能力，进一步最小化RKHS中源域数据的类内变化，同时确保类间语义特征的良好分离，本文提出了MKCL，表达式为

$$L_{\mathrm{ct}} = \frac{1}{2N_{\mathrm{b}}}\sum_{i=1}^{N_{\mathrm{b}}}\|\phi(\boldsymbol{x}_{\mathrm{s},i})-\boldsymbol{c}_{Y_{\mathrm{s},i}}\|_2^2 \tag{13}$$

式中：N_{b}为批大小；$\boldsymbol{c}_{Y_{\mathrm{s},i}}$为样本$i$所对应的类别的所有样本特征的特征中心。MKCL通过优化特征的表示，将同一类别内的特征映射到其中心位置，最小化每个样本与对应类别中心的距离，从而实现类内特征的聚合。

此外，使用两层全连接层作为分类器来实现分类。多尺度卷积模块和深层特征提取模块实现了从信号空间到高层次特征空间的直接映射，从原始信号中提取的特征向量$\boldsymbol{x}_{\mathrm{s},i}$通过简单的线性映射函数$f(\cdot)$得到预测的类分布$q_{\mathrm{s},i}$。分类器将识别源域数据的标签，并利用交叉熵损失(cross entropy loss, CE)来控制分类器的误差。交叉熵损失通常用于迭代更新网络模型的权值，以最小化输出类分布与真实类分布之间的差异，定义为

$$L_{\mathrm{ce}} = -\frac{1}{N_{\mathrm{b}}}\sum_{i=1}^{N_{\mathrm{b}}}Y_{\mathrm{s},i}\ln q_{\mathrm{s},i} \tag{14}$$

式中：Y_{s}为实际源域标签；q_{s}为模型预测的源域标签。

因此，总损失函数结合交叉熵损失、迁移损失和多核中心损失，可得

$$L = L_{\mathrm{ce}}+\lambda_{\mathrm{tf}}L_{\mathrm{tf}}+\lambda_{\mathrm{ct}}L_{\mathrm{ct}} \tag{15}$$

式中：权重λ_{tf}和λ_{ct}用来平衡3个损失函数。当$\lambda_{\mathrm{tf}}=0$及$\lambda_{\mathrm{ct}}=0$时，该函数就会退化成常规的交叉熵损失函数。λ权重因子用来平衡不同损失函数的重要性，交叉熵损失用于从标记数据中提取信息模式，而迁移损失旨在最小化源域和目标域特征分布之间的差异。此外，MKCL减少了类内变化，同时保证了不同类之间的明显分离。因此，选择合适的权重因子，对于网络特征识别能力的提升是至关重要的。

3　实验结果与分析

3.1　数据集和实验设置

本文将 RadioML2016.10a 数据集[10] 作为基准数据集, 验证本文 MATLA 的有效性。该数据集包含 8PSK、AM-DSB、AM-SSB、BPSK、CPFSK、GFSK、PAM4、QAM16、QAM64、QPSK、WBFM 这 11 种调制样式。考虑到实际调制识别场景中输入信号的信噪比一般保持在 0 dB 以上, 本文选取信噪比为 0～18 dB 的样本进行模型训练和性能评估。每个信噪比条件下, 每种调制样式由 1 000 个原始 IQ 信号样本表示, 每个样本大小为 2×128。本文将每个信噪比下每种调制类型的 500 个样本构成的数据集称为 RML1。对另外 500 个样本以 2 的下采样因子进行下采样构造的数据集称为 RML2。在本文实验中, 如不特别说明, 均以 RML1 作为源域数据, RML2 作为目标域数据进行实验。

本文将 MATLA 与 3 种深度神经网络模型 CNN[9]、CLDNN[12]、ResNet[11] 进行了比较。CLDNN 合并 CNN 和 LSTM 网络进行调制识别。ResNet 是近年来先进的深度网络, 具有出色的识别性能, 应用于波形识别[11], 该网络包含 8 个残差模块, 每个残差模块包含 2 层卷积。同时, 将 ResNet 与本文的总损失函数结合起来作为对比方法。此外, MATLA 与具有代表性的深度域混淆 (deep domain confusion, DDC)[27]、深度自适应网络 (deep adaptation networks, DAN)[28]、最大分类器差异 (maximum classifier discrepancy, MCD)[25] 和最小类混淆 (minimum class confusion, MCC)[26] 等无监督域适应模型进行了比较。

在训练中, 使用随机梯度下降 (stochastic gradient descent, SGD) 优化器训练网络, 批大小为 32。所有的实验都在 Nvidia GeForce GTX 1 080 GPU 上执行, 所有模型都使用 PyTorch 深度学习框架实现。

3.2　实验结果分析

实验 1　一维卷积核大小对识别性能的影响。本实验研究了多尺度注意力卷积模块内的一维卷积核大小对识别性能的影响。为此, 使用多组不同大小的卷积核进行了实验。每组包含 2 层, 每层包含 16 个滤波器。实验结果汇总如表 1 所示。从表中可以看出, 在 3 个并行卷积中使用较小的卷积核可以获得更高的识别精度, 特别是当核之间的大小差异较小时, 当卷积核大小为 1、3、5 时, 准确率达到 83.42%。这一结果归因于 2 个主要因素。首先, 较小的卷积核更有效地捕获相邻采样点之间的局部关系, 但可能不足以捕捉足够的全局信息; 再大一些的卷积核(如 1×5)可以提供较广泛的上下文信息, 但可能忽略一些细节, 采取多尺度卷积提取特征既能够提供足够的上下文信息, 又不会过于忽略细节。但同时随着卷积核尺寸的增大, 跨越的采样点的时间序列越远, 这种关系就越不明显, 从而导致更多噪声的积累。1×1 的卷积核有助于学习 IQ 通道间的相关性, 而 1×3 和 1×5 卷积核则能够捕获局部空间特征及上下文信息。这种设计使得 MATLA 在能够保持参数数量和计算复杂度相对较低的同时, 有效地提取多尺度特征。后续的融合操作具有一定的平滑效果, 有助于降低单个尺度卷积核中可能存在的噪声和局部波动。通过结合局部和全局信息, 这有助于更全面地了解数据。因此, 在后续的实验中, 将 3 路并行的多尺度卷积核的大小分别设置为 1、3、5。

表 1　不同卷积核大小下的识别性能

Table 1　Recognition performance under different convolution kernel sizes

卷积核大小	准确率/%
1, 3, 5	83.42
1, 5, 9	81.18
3, 7, 11	80.32
5, 7, 9	81.07
7, 9, 11	80.29

实验 2　MATLA 与其他方法的识别性能进行对比。本实验将本文 MATLA 与几种网络架构以及域适应模型进行了全面的比较。表 2 为 RML1 为源域, RML2 为目标域的域适应任务 RML1→RML2, 以及 RML2 为源域, RML1 为目标域的域适应任务 RML2→RML1 的分类结果。表 2 中还包含每个批次的训练时间。MATLA 的训练时间多于 CNN、CLDNN、ResNet、DDC、DAN 和 MCC, 但结果清楚地表明, 本文 MATLA 在所有评估的方法中达到了最高的识别精度, 在 RML1→RML2 的任务中, 与其他方法中识别精度最高的 MCD 相比, 识别精度提高了 2.78%。在 RML2→RML1 任务中, 与其他方法中识别精度最高的 ResNet+L 相比, 识别精度提高了 4.30%。值得注意的是, 包括 CNN、CLDNN 和 ResNet 在内的网络架构在识别源域数据方面表现出色, 但迁移到目标域时遇到了显著的性能下降。用 ResNet 网络替换本文 MANet 进行特征提取, 同时使用本文提出的损失函数 L 来进行迁移学习(表 2 中的 ResNet+L 方法)。将 ResNet 网络提取的特征进行迁移, 有助于 ResNet 识别精度的提高。在 RML1→RML2 的任务中, 其准确率从 55.21% 提高到 76.63%, 并超过了域适应模型 DDC; 在 RML2→RML1 的任务中, 其准确率从 41.78% 提高

到 71.11%, 超过了域适应模型 DDC、DAN、MCD 和 MCC。但与 MATLA 相比, 由于 MANet 出色的特征提取能力及损失函数的设计, MATLA 获得了最好的识别性能。DAN 等模型只关注对齐全连接层内的特征, 而忽略了特定网络层特征重要的迁移能力及类内特征的约束。相比之下, MATLA 利用多尺度特征融合来获得细粒度特征, 并且在类内的紧凑性和类间的分离性之间取得平衡, 这些特征表现出强大的可迁移性。因此, MATLA 优于 DDC、DAN、MCD 和 MCC 等模型, 在实验中取得了最好的结果。

图 3 为 RML1→RML2 任务下信噪比为 10 dB 的识别结果的混淆矩阵。为可视化地表示目标域中信号样本的特征分布, 图 4 为使用 t 分布式随机邻居嵌入 (t-distributed Stochastic Neighbor Embedding, t-SNE) 降维后深层特征的分布。值得注意的是, MK-MMD 和 MKCL 的结合使得大多数类的样本点簇分离。这种清晰的聚类分离证实了 MATLA 在将知识从源域转移到目标域方面的有效性。然而, 图 3 和图 4 仍然揭示了某些挑战。具体来说, WBFM 信号很容易被误认为 AM-DSB, 原因可能在模拟信号静默期中只有一个载波频率。此外, 由于 QAM64 和 QAM16 在星座图上的幅度和相位点相似, 以及其在 IQ 数据上表现出相似的特征, 因此, 可以观察到其之间的混淆。总体而言, MATLA 以较高的精度识别出大多数调制样式。

表 2　不同模型识别准确率对比

Table 2　Comparison of recognition accuracy of different models

模型	训练时间/s	准确率/%	
		RML1→RML2	RML2→RML1
CNN[9]	0.004	44.09	31.90
CLDNN[12]	0.006	46.77	45.46
ResNet[11]	0.013	55.21	41.78
DDC[27]	0.012	63.24	52.24
DAN[28]	0.013	80.25	64.46
MCD[25]	0.089	80.64	62.46
MCC[26]	0.008	79.28	61.79
ResNet+*L*	0.027	76.63	71.11
MATLA	0.015	83.42	75.41

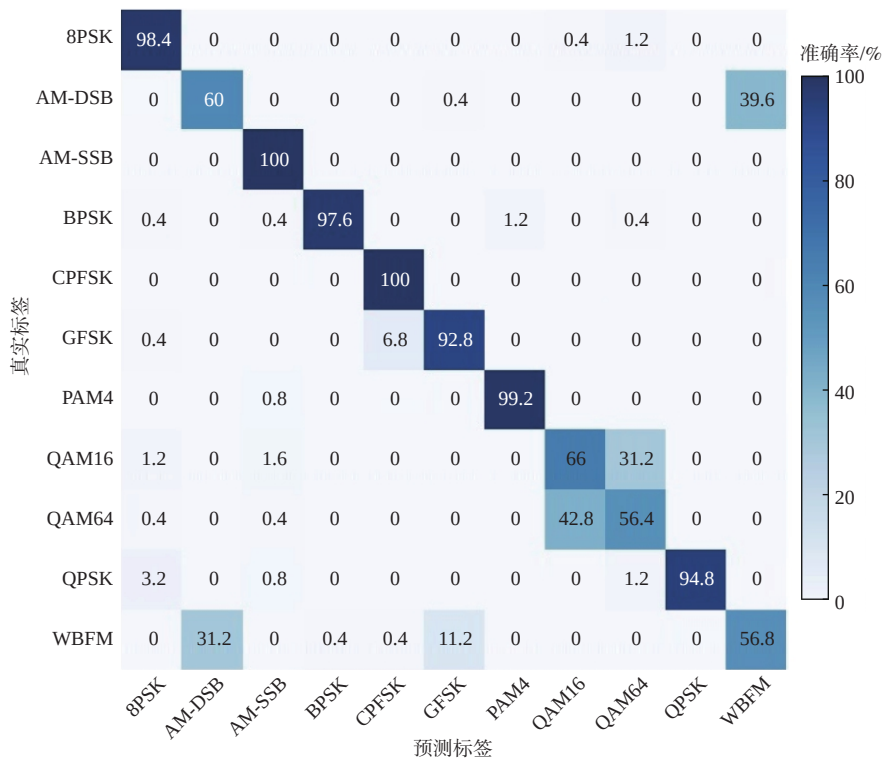


图 3　混淆矩阵 (信噪比为 10 dB)

Fig. 3　Confusion matrix (signal-to-noise ratio is 10 dB)

实验 3　消融实验。图 5 为在有和没有注意机制 2 种情况下, 准确率随信噪比的变化。从结果可以看出, 在所有信噪比水平上, 具有注意机制的模型几乎始终优于无注意机制的模型。这证明了本文提出的多尺度注意机制在提高识别性能方面的有效性。将多尺度注意力纳入模型有几个好处。首先, 其使模型能够关注局部上下文信息, 并在不同尺度下捕获相关特征, 有利于从时间序列信号样

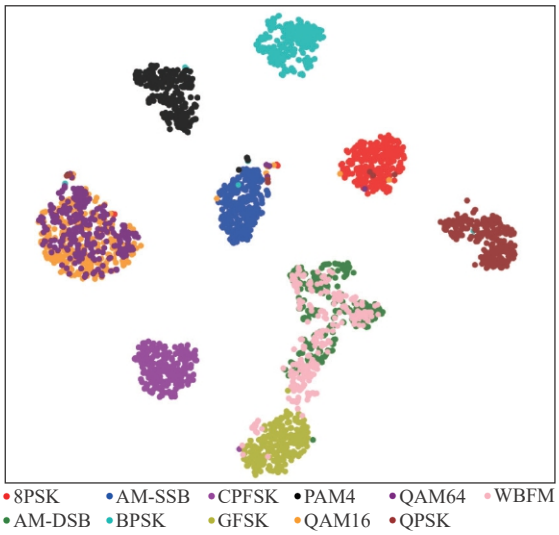


图 4　特征可视化

Fig. 4　Feature visualization

本中提取更精细、更全面的特征; 其次, 注意力机制有效降低了噪声和不相关信息的影响, 增强了鲁棒性和泛化性。因此, 多尺度注意卷积模块使模型能够做出更准确、更可靠的预测。此外, 多尺度注意卷积模块的通用性是另一个优点。除在当前识别任务中表现出的有效性外, 该模块还可以无缝地适应各种其他分类和识别任务, 将其适用性扩展到更广泛的领域。

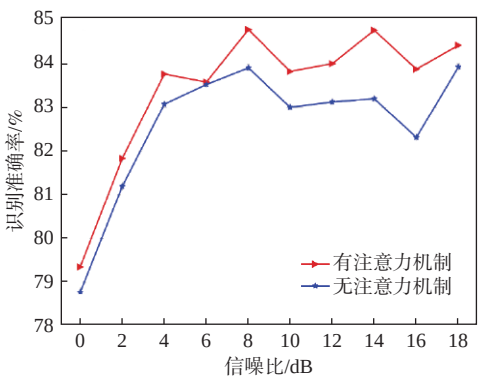


图 5　有/无多尺度注意力的识别性能

Fig. 5　Recognition performance with/without multi-scale attention

为验证本文的组合损失函数的必要性, 表 3 显示了不含 MK-MMD、不含 MKCL、不同权重因子条件下模型的识别性能。可以发现, 单独使用这 2 个损失函数中的任何一个时, 模型的性能都会下降。此外, 不同权重因素的组合对精度也有影响。当 2 种损失函数的权重相等时, 与单独使用 MK-MMD 损失相比, 识别精度分别提高了 0.33% 和 0.40%。这证明了组合损失函数在增强模型从不同类别学习到的特征之间的区分性方面的重要性。

表 3　不同权重因子作用下的识别性能

Table 3　Recognition performance under different weight factors

λ_{tf}	λ_{ct}	准确率/%
0	0	75.98
0	0.5	76.59
0.5	0	82.83
0.5	0.5	83.16
0.5	1.0	83.22
1.0	0	83.02
1.0	0.5	83.20
1.0	1.0	83.42
1.0	1.5	84.07

实验 4　源域数据集大小对模型性能的影响。通过对不同比例的源域训练集进行实验, 本实验探索了源域数据集的大小对模型性能的影响。具体来说, 总的源域样本数量为 N_s, 实际参与训练的源域样本数量为 N_{st}, 则参与训练的源域数据比例 $\eta = N_{st}/N_s$。图 6 为参与训练的源域数据比例 η 范围为 0.1～0.9 的情况下, MATLA 性能随信噪比的变化情况。当比例为 0.1 时, 识别准确率达到 79.31%, 仍高于其他几种对比模型; 当比例为 0.8 时, 识别准确率上升到 83.42%。实验结果表明, 即使源域训练集的比例很小, MATLA 也能获得良好的性能。在实际场景中, 如果获得带标签的源域数据越多, 那么模型就可能学习到越多的知识, 因此, 迁移效果越好。但是带标签的数据标注比较困难, 如非合作通信情况下, 需要依靠专家知识耗费大量的人力来进行标注。因此, 在实际场景中可用的标签样本很少, 如本实验中设置带标签的源域样本比例为 0.1。而 MATLA 在有较少的带标签源域数据参与训练的情况下识别准确率仍高于其他对比模型。由此可知, MATLA 在拥有少量带标签源域数据情况下, 仍可以有效地进行识别。

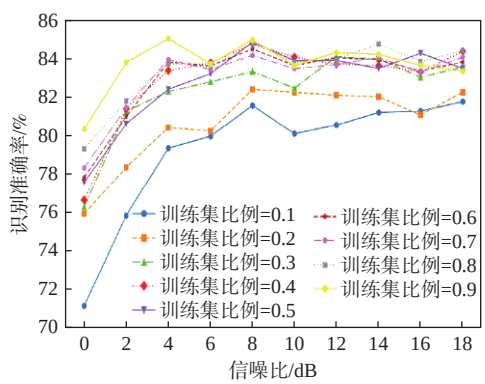


图 6　MATLA 在不同训练集比例下的识别性能

Fig. 6　Recognition performance of MATLA under different proportions of training sets

4 结 论

本文针对变采样率调制识别问题，提出了一种新颖的多尺度注意力迁移学习架构。主要结论如下：

1）本文设计了一种多尺度注意力机制来提取多层次的特征信息，利用深层特征来有效地对齐源域和目标域。此外，MKCL 和 MK-MMD 的有效结合使得同一类内的表达更加紧凑，不同类之间的可分离性更好。

2）实验评估了本文 MATLA 的有效性，对比实验结果表明，MATLA 超越了对比的网络模型和无监督域适应模型，识别准确率达到 83.42%。因此，在解决变采样率调制识别的挑战方面表现出卓越的性能。

通过与其他方法进行对比及一系列消融实验，本文 MATLA 为解决变采样率的调制识别问题提供了一种有效的解决方案。

参考文献（References）

[1] 张海燕, 闫文君, 张立民, 等. 通信信号调制识别综述[J]. 海军航空大学学报, 2022, 37(1): 126-132.
ZHANG H Y, YAN W J, ZHANG L M, et al. An overview of communication signal modulation recognition[J]. Journal of Naval Aviation University, 2022, 37(1): 126-132(in Chinese).

[2] 林冲, 闫文君, 张立民, 等. 通信信号调制识别综述[J]. 中国电子科学研究院学报, 2021, 16(11): 1074-1085.
LIN C, YAN W J, ZHANG L M, et al. An overview of communication signals modulation recognition[J]. Journal of China Academy of Electronics and Information Technology, 2021, 16(11): 1074-1085 (in Chinese).

[3] WEI W, MENDEL J M. Maximum-likelihood classification for digital amplitude-phase modulations[J]. IEEE Transactions on Communications, 2000, 48(2): 189-193.

[4] HASSAN K, DAYOUB I, HAMOUDA W, et al. Automatic modulation recognition using wavelet transform and neural networks in wireless systems[J]. EURASIP Journal on Advances in Signal Processing, 2010, 2010(1): 532898.

[5] ORLIC V D, DUKIC M L. Automatic modulation classification algorithm using higher-order cumulants under real-world channel conditions[J]. IEEE Communications Letters, 2009, 13(12): 917-919.

[6] YANG W C, REN K, DU Y, et al. Modulation recognition method of mixed signals based on cyclic spectrum projection[J]. Scientific Reports, 2023, 13: 21459.

[7] MALIK M, MALIK M K, MEHMOOD K, et al. Automatic speech recognition: a survey[J]. Multimedia Tools and Applications, 2021, 80(6): 9411-9457.

[8] GUO M H, XU T X, LIU J J, et al. Attention mechanisms in computer vision: a survey[J]. Computational Visual Media, 2022, 8(3): 331-368.

[9] O'SHEA T J, CORGAN J, CLANCY T C. Convolutional radio modulation recognition networks[C]//Proceedings of the Engineering Applications of Neural Networks. Berlin: Springer, 2016: 213-226.

[10] O'SHEA T, WEST N E. Radio machine learning dataset generation with gnu radio[C]//Proceedings of the GNU Radio Conference. Boulder: GNU Radio, 2016, 1(1): 1-6.

[11] O'SHEA T J, ROY T, CLANCY T C. Over-the-air deep learning based radio signal classification[J]. IEEE Journal of Selected Topics in Signal Processing, 2018, 12(1): 168-179.

[12] WEST N E, O'SHEA T. Deep architectures for modulation recognition[C]//Proceedings of the IEEE International Symposium on Dynamic Spectrum Access Networks. Piscataway: IEEE Press, 2017: 1-6.

[13] 哈晖, 高翔, 姚秀娟, 等. 基于 STF-Net 的信号调制波形识别方法[J]. 北京航空航天大学学报, 2025, 51(9): 3150-3160.
HA H, GAO X, YAO X J, et al. Signal modulation waveform recognition method based on STF-Net[J]. Journal of Beijing University of Aeronautics and Astronautics, 2025, 51(9): 3150-3160(in Chinese).

[14] RAJENDRAN S, MEERT W, GIUSTINIANO D, et al. Deep learning models for wireless signal classification with distributed low-cost spectrum sensors[J]. IEEE Transactions on Cognitive Communications and Networking, 2018, 4(3): 433-445.

[15] PENG S L, JIANG H Y, WANG H X, et al. Modulation classification based on signal constellation diagrams and deep learning[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(3): 718-727.

[16] WANG Y, LIU M, YANG J, et al. Data-driven deep learning for automatic modulation recognition in cognitive radios[J]. IEEE Transactions on Vehicular Technology, 2019, 68(4): 4074-4077.

[17] 姬丽彬, 朱岩, 崔天舒, 等. 基于时频重排算法的 LPI 雷达信号识别[J]. 北京航空航天大学学报, 2025, 51(4): 1324-1331.
JI L B, ZHU Y, CUI T S, et al. LPI radar signal recognition based on time-frequency reassignment algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2025, 51(4): 1324-1331 (in Chinese).

[18] WONG L J, MICHAELS A J. Transfer learning for radio frequency machine learning: a taxonomy and survey[J]. Sensors, 2022, 22(4): 1416.

[19] PAN S J, YANG Q. A survey on transfer learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10): 1345-1359.

[20] 庞伊琼, 许华, 张悦, 等. 基于迁移元学习的调制识别算法[J]. 兵工学报, 2023, 44(10): 2954-2963.
PANG Y Q, XU H, ZHANG Y, et al. Modulation recognition algorithm based on transfer meta-learning[J]. Acta Armamentarii, 2023, 44(10): 2954-2963(in Chinese).

[21] 李晓慧, 陈壮志, 徐东伟, 等. 基于知识迁移的深度学习无线电信号聚类方法[J]. 高技术通讯, 2023, 33(11): 1172-1180.
LI X H, CHEN Z Z, XU D W, et al. A deep learning radio signal clustering method based on knowledge transfer[J]. Chinese High Technology Letters, 2023, 33(11): 1172-1180(in Chinese).

[22] 毛红霞. 无监督领域自适应调制识别方法[J]. 探测与控制学报, 2021, 43(4): 117-122.
MAO H X. Unsupervised domain adaptation based automatic modulation recognition[J]. Journal of Detection & Control, 2021, 43(4): 117-122(in Chinese).

[23] WANG Q, DU P F, YANG J Y, et al. Transferred deep learning

based waveform recognition for cognitive passive radar[J]. Signal Processing, 2019, 155: 259-267.

[24] BU K, HE Y, JING X J, et al. Adversarial transfer learning for deep learning based automatic modulation classification[J]. IEEE Signal Processing Letters, 2020, 27: 880-884.

[25] SAITO K, WATANABE K, USHIKU Y, et al. Maximum classifier discrepancy for unsupervised domain adaptation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 3723-3732.

[26] JIN Y, WANG X M, LONG M S, et al. Minimum class confusion for versatile domain adaptation[C]//Proceedings of the Computer Vision-ECCV. Berlin: Springer, 2020: 464-480.

[27] TZENG E, HOFFMAN J, ZHANG N, et al. Deep domain confusion: maximizing for domain invariance[EB/OL]. (2014-12-10)[2024-04-05]. https://arxiv.org/abs/1412.3474.

[28] LONG M S, CAO Y, WANG J M, et al. Learning transferable features with deep adaptation networks[EB/OL]. (2015-02-10)[2024-04-05]. https://arxiv.org/abs/1502.02791.

[29] BA J L, KIROS J R, HINTON G E. Layer normalization[EB/OL]. (2016-07-21)[2024-04-05]. https://arxiv.org/abs/1607.06450.

[30] GRETTON A, SRIPERUMBUDUR B, SEJDINOVIC D, et al. Optimal kernel choice for large-scale two-sample tests[C]//Proceedings of the 26th International Conference on Neural Information Processing Systems. New York: ACM, 2012: 1205-1213.

Automatic modulation classification method based on transfer learning

WANG Dong[1, 2], CUI Tianshu[3], JI Libin[1, 2], HUANG Yonghui[1], ZHU Yan[1, *]

(1. Key Laboratory of Electronic and Information Technology for Space Systems, National Space Science Center, Chinese Academy of Science, Beijing 100190, China;
2. School of Computer Science and Technology, University of Chinese Academy of Science, Beijing 100049, China;
3. China Academy of Aerospace Science and Innovation, Beijing 100190, China)

Abstract: Deep learning has achieved considerable progress in the field of modulation classification; however, it is often limited by the consistency of the distribution between training and testing data. A novel multi-scale attention transfer learning architecture (MATLA) has been designed to assist cross-domain identification in order to overcome the problem of modulation classification with inconsistent data distribution caused by variable sample rates of source and target domains. This framework employs three parallel convolutional kernels with varying scales to extract features at different granularities. Moreover, the architecture integrates a multi-scale attention mechanism, which bolsters the extraction of discriminative features by emphasizing the weights of salient features. To effectively align features from the source and target domains, multiple kernel maximum mean discrepancy (MK-MMD) is utilized to measure the divergence of feature distributions in the reproducing kernel Hilbert space (RKHS) between the two domains. Additionally, to mitigate internal variation within the source domain features and enhance their consistency and stability, multiple kernel center loss (MKCL) is proposed. According to experimental data, the suggested approach outperforms a number of different network models and domain adaption techniques, achieving a recognition accuracy of 83.42% when the signal-to-noise ratio surpasses 0 dB.

Keywords: transfer learning; attention mechanism; deep learning; modulation classification; sampling rate

Received: 2024-04-18; **Accepted**: 2024-07-30; **Published Online**: 2024-08-02 14:29
URL: link.cnki.net/urlid/11.2625.V.20240802.1334.001

* **Corresponding author**. E-mail: zhuyan@nssc.ac.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0217

基于知识图谱的在轨卫星与卫星网络资料匹配关系预测方法

徐聪[1,2]，石梦鑫[3]，王洪锋[4]，贾清玉[1]，智佳[1]，杨甲森[1,*]

(1. 中国科学院国家空间科学中心 复杂航天系统电子信息技术重点实验室，北京 100190；
2. 中国科学院大学 计算机科学与技术学院，北京 100049；3. 东南大学 信息科学与工程学院，南京 210096；
4. 中国卫星发射测控系统部，北京 100094)

摘 要：在轨卫星与国际电信联盟(ITU)卫星网络申报资料的精确匹配，对于卫星频率和轨道的设计、选取、申报、协调具有重要意义。针对传统人工匹配方法领域知识要求高、匹配效率低等问题，提出一种基于知识图谱的在轨卫星与卫星网络资料匹配关系无监督预测(UPMIS)方法。该方法在构建匹配关系预测指标体系、在轨卫星知识图谱、卫星网络资料知识图谱的基础上，基于业务知识融合与图划分的思想，设计由时间参数模块、数值型轨道参数模块和字符型社会参数模块组成的三级滤波结构框架，实现在轨卫星与卫星网络资料的快速精确匹配。实验结果表明：UPMIS在真实数据集上的排名前10命中率(H_{10})为0.854 2，优于其他对比模型，且平均运行时间达到毫秒级。此外，实验给出过滤数量、聚合深度等参数选取的推荐值，对后续匹配关系预测具有参考价值。

关 键 词：关系预测；知识图谱；知识融合；无监督学习；图划分；在轨卫星；卫星网络资料

中图分类号：V419; TP391.1

文献标志码：A **文章编号**：1001-5965(2026)06-1944-11

有限不可再生的频率轨道资源是卫星开展空间无线电业务的基础[1-2]。根据国际电信联盟规则[3]，卫星在发射前必须提前向国际电联登记卫星网络资料。近年来，Starlink[4-5]、Oneweb[6-8]、Telsat[9]等巨型低轨通信星座的兴起加速了空间频轨资源的紧张程度[10-11]。为竞争有限的频轨资源，登记在国际电联的卫星网络资料数量远远大于实际在轨运行的卫星数量[12]，这种仅停留在卫星网络纸质材料层面而没有真正启用的卫星被称为“纸卫星”[13-14]。

“纸卫星”问题极大增加了卫星系统选频用频的难度，严重阻碍后续卫星网络的建设和使用[15]。研究在轨卫星与卫星网络资料匹配关系是解决“纸卫星”难题、有效进行卫星系统频率轨道设计的必要条件。目前，开展在轨卫星与卫星网络资料匹配关系预测研究主要存在以下问题：

1）人工模式匹配方法效率低。当前匹配工作主要通过人工分析申报与协调资料的方式实现，要求从业人员掌握卫星频轨资源分配机制、卫星网络申报协调与登记维护、星间干扰分析等专业知识，涉及频轨资源管理、空间通信、深空探测等跨学科多个领域，受人员的专业水平和知识占有影响，往往精度不高。此外，根据空间无线电通信台站(space radio communications stations, SRS)2 996期数据库统计，目前全球有效的卫星网络资料数量为13 819份，

收稿日期：2024-04-15；录用日期：2024-05-31；网络出版时间：2024-06-17 11:22
网络出版地址：link.cnki.net/urlid/11.2625.V.20240616.1430.001
基金项目：国家自然科学基金(62341106)；北京市自然科学基金(L222003)
*通信作者. E-mail：jsy@nssc.ac.cn

引用格式：徐聪，石梦鑫，王洪锋，等. 基于知识图谱的在轨卫星与卫星网络资料匹配关系预测方法[J]. 北京航空航天大学学报，2026，52(6)：1944-1954. XU C，SHI M X，WANG H F，et al. Prediction method for matching between in-orbit satellites and satellite network filings based on knowledge graph[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1944-1954 (in Chinese).

总计划发射卫星数超过 100 万颗[12], 而截至 2023 年 5 月 1 日在轨卫星数量为 7 560 颗, 在轨卫星与卫星网络资料数量的严重失衡进一步加剧了人工模式匹配的时间成本。

2) 卫星网络资料与在轨卫星数据集指标粒度与量不统一, 图同构算法难以直接适用。卫星网络资料数据集以 SRS 2 996 期数据库为例, 其详细覆盖了卫星网络通知资料、卫星网络轨道等多方面的上千个指标。在轨卫星数据集以忧思科学家联盟(Union of Concerned Scientists, UCS)卫星数据集为例, 其覆盖了在轨卫星的物理特性、目的、轨道参数等 26 个指标。显然, 二者的指标在数量与粒度上有显著差异, 为异构图。此外, 在轨卫星与卫星网络资料的对应关系为“多对一”, 即多颗卫星匹配于同一份卫星网络资料[16], 无法适用基于“一一对应”的图同构算法。如何构建在轨卫星与卫星网络资料指标体系是进行匹配的前提和基础, 如何设计准确合适的算法是进行匹配的关键。

3) 神经网络方法可解释性差, 模型结构冗余。图神经网络(graph neural networks, GNN)作为一种处理图结构数据的神经网络模型, 广泛应用于推荐系统[17]、智能问答[18] 等领域。然而, 受限于神经网络模型框架的限制, 其中大量的参数和非线性结构使得 GNN 很难被彻底解释, 因此,难以判断模型的学习能力是普遍的还是在特定数据集上的过度学习[19]。随着图注意力网络[20]、图匹配网络[21]、联合学习[22] 等模型提出, 模型结构的复杂化不可避免地增加了算法的时空复杂度, 有研究指出, 关系感知的双图卷积网络(relation-aware dual graph convolutional network, RDGCN)[20] 的运行时间是 Vanilla GCN[23] 的 10 倍以上[24]。卫星领域具有敏感性, 对算法的可解释性和响应速度有更严格的要求。

知识图谱以结构化知识表示现实中复杂关系, 在关联性分析上具有更好的可解释性[25]。为解决以上问题, 本文提出一种基于知识图谱的在轨卫星与卫星网络资料匹配关系无监督预测(unsupervised prediction of matching between in-orbit satellites and satellite network filings, UPMIS)方法, 将 2 个独立知识图谱的匹配关系预测问题转化为排序问题。首先, 设计在轨卫星与卫星网络资料匹配关系预测指标体系, 进一步构建在轨卫星知识图谱和卫星网络资料知识图谱, 随后, 基于业务知识融合与图划分的思想设计由时间参数模块、数值型轨道参数模块和字符型社会参数模块组成的 3 级滤波结构框架, 解决在轨卫星与卫星网络资料匹配关系预测问题, 实现二者的关联与融合。

1　问题描述与知识图谱构建

知识图谱是由点和边组成的语义图[26], 结构化地客观描述了现实世界的知识[27], 已成为自然语言理解[28]、推荐系统[17]、智能问答[18] 等人工智能应用中实现人类智能的关键课题[29]。关系预测作为知识图谱的下游任务之一, 构建合适的知识图谱是实现匹配关系预测的必要条件。本节对在轨卫星与卫星网络资料匹配关系预测任务的输入输出与模型假设进行梳理, 对匹配关系预测指标体系和在轨卫星与卫星网络资料知识图谱构建方法展开研究。

1.1　问题描述

1) 问题定义。已知在轨卫星数据集与卫星网络资料数据集, 首先, 设计一种面向匹配关系预测的知识图谱构建方法将数据集转换为知识图谱形式, 得到在轨卫星知识图谱 $K^{\text{sat}}=\{E^{\text{sat}},R^{\text{sat}},T^{\text{sat}}\}$, 卫星网络资料知识图谱 $K^{\text{net}}=\{E^{\text{net}},R^{\text{net}},T^{\text{net}}\}$, 其中, E^{sat} 和 E^{net} 分别表示在轨卫星与卫星网络资料的实体集合, R^{sat} 和 R^{net} 分别表示在轨卫星与卫星网络资料的关系集合, T^{sat} 和 T^{net} 分别表示在轨卫星与卫星网络资料的三元组集合, $T^{\text{sat}}=\{(h,r,t)|h\in E^{\text{sat}},r\in R^{\text{sat}},t\in T^{\text{sat}}\}$, $T^{\text{net}}=\{(h,r,t)|h\in E^{\text{net}},r\in R^{\text{net}},t\in T^{\text{net}}\}$, h表示头实体, r表示关系, t表示尾实体。匹配关系预测任务旨在对于在轨卫星实体 h_{s}, 找到唯一的卫星网络资料 t_{n}^{*}使得三元组 $(h_{\text{s}},r_{\text{m}},t_{\text{n}}^{*})$ 成立, 其中, r_{m}表示匹配关系, $h_{\text{s}}\in E^{\text{sat}},t_{\text{n}}^{*}\in E^{\text{net}}$。

假设在轨卫星实体 h_{s} 和卫星网络资料实体 t_{n} 经过嵌入模型 $f(\cdot)$ 生成的嵌入式向量分别为 $\boldsymbol{e}^{\text{sat}}$ 和 $\boldsymbol{e}^{\text{net}}$, 其表达式分别为

$$\boldsymbol{e}^{\text{sat}}=f\left(h_{\text{s}},K^{\text{sat}}\right) \tag{1}$$

$$\boldsymbol{e}^{\text{net}}=f\left(t_{\text{n}},K^{\text{net}}\right) \tag{2}$$

可通过向量相似度函数得到在轨卫星实体与多个卫星网络资料实体的相似度, 相似度值排序即可选出最优卫星网络资料 t_{n}^{*}作为在轨卫星 h_{s}的匹配卫星网络资料。因此, 如何构建知识图谱 K^{sat} 和 K^{net} 是匹配关系预测的基础, 如何基于知识图谱设计嵌入模型 $f(\cdot)$ 是匹配关系预测的关键。

2) 输入与输出。输入: 卫星目录号是美国航天司令部按识别顺序分配给所有地球轨道卫星的连续 5 位数字, 具有唯一性。因此, 本问题的输入 h_{s} 为在轨卫星的唯一标识符“卫星目录号”。输出: 实际实验中输出为候选卫星网络资料 t^{*} 的排序列表。

3) 假设。输入的在轨卫星 h_{s} 一定在 K^{sat} 中, 预测的卫星网络资料 t^{*} 一定在 K^{net} 中。

1.2 知识图谱构建

采用自上而下的方式构建知识图谱。对于匹配关系预测任务,所需的卫星信息与卫星网络资料信息并非是越多越好,因为过多无关信息反而会对结果的有效性产生干扰。一般认为,事物间的相似部分能够反映出二者的共性与关联。基于此,设计关联指标提取的规则如下。

定义 1 (关联指标)设在轨卫星指标集合为R^{sat},卫星网络资料指标集合为R^{net},则2个指标集合的交集对于在轨卫星匹配关系的预测具有更强的判断依据:

$$R^{*}=R^{\text{sat}}\cap R^{\text{net}} \tag{3}$$

基于以上规则分析在轨卫星数据集UCS-Satellite-Database-5-1-2023[30]与卫星网络资料数据集SRS 2 996期数据库,整理出在轨卫星数据集与卫星网络资料数据集的指标关联表如表1所示,相似的指标分为轨道类、组织类、时间类和标志类4大类。

表1 在轨卫星与卫星网络资料关联指标

Table 1 Indicators for associating on-orbit satellites with satellite network filings

指标范围	在轨卫星指标	卫星网络资料指标	指标关联度
轨道类	近地点	近地点	完全一致
	远地点	远地点	完全一致
	倾角	轨道倾角	完全一致
	周期	轨道周期	完全一致
组织类	国家/UN登记组织	主管部门代码	关联度高
	操作者/拥有者		
	操作者/拥有者的国家		
	制造商		
	制造商的国家		
标志类	卫星官方名称	卫星网络名称	关联度低
	卫星别名		
时间类	发射时间	接收时间	关联度高

根据关联指标提取的定义,轨道类指标在2个数据集中完全一致,即R^{*}为轨道类关系,同时数据易获取、置信度高,因此,轨道是在轨卫星与卫星网络资料匹配关系预测的关键特征。卫星网络资料的“主管部门代码”为递交卫星网络申报资料的主管部门的国家符号,与在轨卫星数据集中“操作者/拥有者的国家”“制造商的国家”和“国家/UN登记组织”的范围均为国家,指标关联度高。为进一步加深关联指标的深度,将“操作者/拥有者的国家”和“制造商的国家”的下一级指标“操作者/拥有者”“制造商”也纳入组织类在轨卫星与卫星网络资料关联指标中。标志类指标指卫星网络和在轨卫星的名称,由于“卫星网络资料名称”与在轨卫星的官方名称或别名存在一定差异,如Starlink卫星的卫星名称一般为“Starlink+编号”格式,而其卫星网络资料的名称一般为“USASAT-NGSO+编号”,因此,关联度较低。卫星的部署严格遵守先申请后发射的原则,因此,时间类指标处于高关联度。

知识图谱在关联性分析时具有更好的可解释性,将数据集格式转换为知识图谱形式:在轨卫星知识图谱$K^{\text{sat}}=\{E^{\text{sat}},R^{\text{sat}},T^{\text{sat}}\}$,卫星网络资料知识图谱$K^{\text{net}}=\{E^{\text{net}},R^{\text{net}},T^{\text{net}}\}$。

2 UPMIS方法

UPMIS方法由时间参数模块、数值型轨道参数模块(本文简称轨道参数模块)、字符型社会参数模块(本文简称社会参数模块)组成。本节介绍UPMIS方法的框架和每个模块的详细结构。

2.1 方法框架

UPMIS方法框架如图1所示,包含5个部分。①时间过滤:根据业务时间知识初步筛选出符合时间逻辑的卫星网络资料。②图划分:将K^{sat}和K^{net}根据关系类型和数据类型分割为数值型轨道类子图和字符型社会类子图。③数值型轨道子图过滤:根据轨道距离函数进一步筛选出距离相近的候选卫星网络资料。④字符型社会子图聚合:基于图卷积多层聚合字符类社会邻居信息生成在轨卫星与候选卫星网络资料的嵌入向量。⑤层集成:使用每层实体嵌入向量来构造每层的相似度向量,这些相似度向量聚合为最终相似度向量,多个输入卫星的相似度堆叠构成在轨卫星与候选卫星网络资料的相似度矩阵,根据相似度大小排序预测出匹配卫星网络资料候选列表。①和③分别为时间参数模块和轨道参数模块,④、⑤构成社会参数模块。

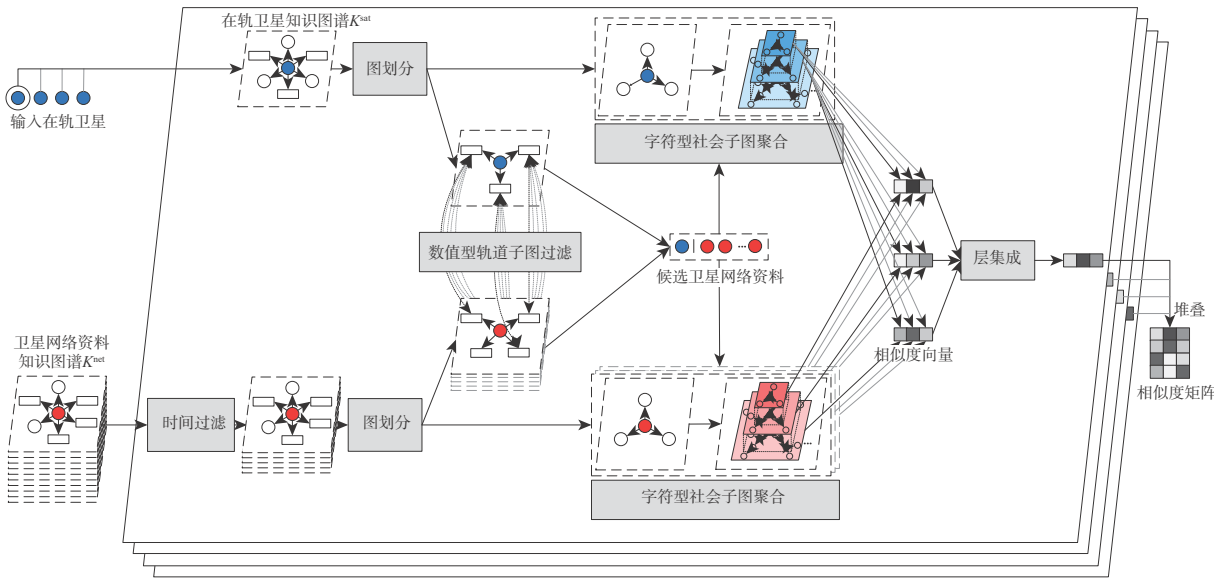


图 1　UPMIS 方法框架

Fig. 1　Framework of UPMIS method

2.2　时间过滤

国际电信联盟规定, 卫星在发射前需要提前向国际电信联盟登记卫星网络资料[31-32]。卫星网络资料主要分为规划类和非规划类两大类。非规划业务的应用范围更广, 竞争态势更为激烈。因此, 本文瞄准非规划业务。非规划类卫星网络资料有 3 类: 提前公布资料为申请时的提前公布资料, 简称 A 资料; 协调资料描述卫星网络的各项申报参数, 如轨道参数、频率指配参数、地面站参数, 以及业务和业务区等, 简称 C 资料; 通知资料进一步细化卫星网络资料参数, 同时涉及协调完成状态等, 简称 N 资料[33-34]。

根据卫星网络资料申报流程, A 资料最先报送至国际电信联盟[35], 因此, 卫星发射日期一定在卫星网络 A 资料接收日期之后。根据这一时间逻辑, 在算法的最初阶段比较实际卫星发射日期与卫星网络 A 资料接收日期, 保留 A 资料接收日期早于实际卫星发射日期的卫星网络资料。

2.3　图划分

关系和实体有不同的类型, 如字符型 Ψ {卫星, 使用者, 中国}和数值型 Ω {卫星, 近地点, 5 000 km}。不同类型关系有不同的相似度评价方式, 如数值型实体可直接通过数值差异比较, 字符型实体可通过语义相似度评价。因此, 将知识图谱 $K=\{E,R,T\}$ 中的关系 R 按照类型的不同分割为数值型子图 K_1 和字符型子图 K_2, 有针对性地进行相似度度量:

1) 数值型子图 $K_1=\{E_1,R_1,T_1\}$, $T_1=\{(h,r,t)|r\in R_1$ 且 $r\in\Omega, r\in h,t\in E_1\}$。

2) 字符型子图 $K_2=\{E_2,R_2,T_2\}$, $T_2=\{(h,r,t)|r\in R_2$ 且 $r\in\Psi, r\in h,t\in E_2\}$。其中, $K=K_1\cup K_2$, $E=E_1\cup E_2$, $T=T_1\cup T_2$, $R=R_1\cup R_2$ 且 $R_1\cap R_2=\varnothing$。

将在轨卫星知识图谱 K^{sat} 和卫星网络资料知识图谱 K^{net} 分别按照关系类型分为数值型子图 K_1^{sat}、K_1^{net} 和字符型子图 K_2^{sat}、K_2^{net}。根据表 1 中的分类, 数值型关系有轨道类信息, 字符型关系有组织类, 标志类等社会信息。由于时间类指标在时间过滤模块被使用, 图划分等价于将 K^{sat} 和 K^{net} 按照关系类型分割为数值型轨道类子图和字符型社会类子图。

2.4　数值型轨道子图过滤

考虑到轨道类特征是在轨卫星与卫星网络资料匹配关系预测的关键特征, 也是在轨卫星数据集与卫星网络资料数据集中指标的完全交集, 基于在轨卫星与卫星网络资料的轨道类信息的相似程度从 K^{net} 中过滤出 k_{orbit} 个相似度高的候选卫星网络资料。

由表 1 可知, 轨道类参数有近地点、远地点、倾角、周期。此时, 式 (3) 中 $R^*=R^{sat}\cap R^{net}=R_{orbit}$, 在轨卫星对应的轨道参数向量 $\boldsymbol{O}^{sat}=\left[o^{sat}_{perigee},o^{sat}_{apogee},o^{sat}_{inclination},o^{sat}_{period}\right]$, 第 i 份卫星网络资料对应的轨道参数向量 $\boldsymbol{O}_i^{net}=\left[o^{net}_{i,perigee},o^{net}_{i,apogee},o^{net}_{i,inclination},o^{net}_{i,period}\right]$。卫星运动具有实时性, 各轨道参数会随着卫星移动位置而变化, 由于轨道参数之间存在物理特性差异, 不同轨道参数对轨位变化程度的影响也不一样。本文设计结合物理特性的轨道距离函数。

对于倾角, 卫星自发射后倾角几乎不变, 因此, 赋予其较高的权重, 距离差异用指数函数表示 $\left(D_{i,inclination}\geqslant 0\right)$:

$$D_{i,inclination}=e^{\alpha\left|o^{sat}_{i,inclination}-o^{net}_{i,inclination}\right|}-1 \tag{4}$$

对于周期, 与倾角同理, 距离差异用指数函数表示 $\left(D_{i,period}\geqslant 0\right)$:

$$D_{i,\text{period}} = \mathrm{e}^{\beta\left|o_{i,\text{period}}^{\text{sat}} - o_{i,\text{period}}^{\text{net}}\right|} - 1 \tag{5}$$

为区分二者的不同权重，根据业务知识一般倾角的稳定性更高，因此，取系数$\alpha = 20, \beta = 10$。

对于近地点和远地点，根据当前基于轨道高度的轨道分类[36]，80～1 700 km为低地球轨道，1 700～35 700 km为中地球轨道，35 700 km以上为高地球轨道，不同类型的轨道高度数量级从成百上千千米到上万千米不等。近地点和远地点随卫星的运动稳定性较差，即使小幅度百分比偏差也会产生较大的绝对偏差，采用欧几里得距离表示二者距离差异：

$$D_{i,\text{pa}} = \sqrt{\left|o_{i,\text{perigee}}^{\text{sat}} - o_{i,\text{perigee}}^{\text{net}}\right|^2 + \left|o_{i,\text{apogee}}^{\text{sat}} - o_{i,\text{apogee}}^{\text{net}}\right|^2} \tag{6}$$

输入的在轨卫星与第i份卫星网络资料的轨道距离函数是式(4)～式(6)之和：

$$\begin{aligned} D_i = {} & D_{i,\text{inclination}} + D_{i,\text{period}} + D_{i,\text{pa}} = \\ & \mathrm{e}^{\alpha\left|p_{i,\text{inclination}}^{\text{sat}} - o_{i,\text{inclination}}^{\text{net}}\right|} + \mathrm{e}^{\beta\left|p_{i,\text{period}}^{\text{sat}} - o_{i,\text{perior}}^{\text{net}}\right|} + \\ & \sqrt{\left|o_{i,\text{perigee}}^{\text{sat}} - o_{i,\text{perigee}}^{\text{net}}\right|^2 + \left|o_{i,\text{apogee}}^{\text{sat}} - o_{i,\text{apogee}}^{\text{net}}\right|^2} - 2 \end{aligned} \tag{7}$$

式中：$i = 1, 2, \cdots, k_{\text{time}}$。根据结果排序进一步筛选出最小的前$k_{\text{orbit}}$个候选卫星网络资料作为轨道参数模块的卫星网络资料过滤数量。

2.5 字符型社会子图聚合

在字符型社会子图上对于在轨卫星和轨道参数模块过滤出的k_{orbit}个候选卫星网络资料进行特征传播，生成卫星和卫星网络资料的嵌入向量。本文设计一种新的邻接矩阵定义方法，采用矩阵向量乘积实现图卷积。

以vanilla GCN为例，传统神经网络的更新方式为[37]

$$\boldsymbol{H}^{l+1} = \sigma\left(\boldsymbol{A}\boldsymbol{H}^{l}\boldsymbol{W}^{l}\right) \tag{8}$$

式中：σ表示激活函数；$\boldsymbol{A}$为邻接矩阵；$\boldsymbol{H}$为输入特征；$\boldsymbol{W}$为变换矩阵；l为层数。由于输入特征$\boldsymbol{H}$和变换矩阵$\boldsymbol{W}$是随机初始化的，可以被简化为一个矩阵[22,38-39]：

$$\boldsymbol{H}^{l+1} = \sigma\left(\boldsymbol{A}\boldsymbol{H}^{l}\right) \tag{9}$$

本文基于代数的方式实现图卷积，实体嵌入向量迭代公式为

$$\boldsymbol{H}(l+1) = \boldsymbol{A}\boldsymbol{H}(l) \tag{10}$$

由于拉普拉斯矩阵等传统邻接矩阵只考虑图的结构，忽略了关系种类和数量。例如，随着国际空间合作日益紧密，美国国家航空航天局主导的登月项目“阿尔忒弥斯计划”的参与国家足有28个，则与其相关的卫星“使用者”这一关系类型就有28个实体。因此，本文定义了一种新的邻接矩阵$\boldsymbol{A}_{\text{rel}}$。

一般认为频率越低的关系类型代表更重要的信息，相应的权重应该更高。不失一般性，令r_{ij}表示实体e_i与实体e_j的关系，设实体e_i有N_i个一阶邻域实体，记实体e_i的一阶邻域实体集合为$M(e_i)$，c_{ij}表示关系r_{ij}在实体e_i一阶邻域关系中出现的次数。除邻居以外，实体自身的信息也很重要，因此，将实体自身也纳入$M(e_i)$中，即邻接矩阵包含环。因此，定义$\boldsymbol{A}_{\text{rel}} = \left[a_{ij}\right]_{n\times n}$如下：

$$a_{ij} = \begin{cases} \left(N_i/c_{ij}\right) \Big/ \left(\sum_{j=1}^{N_i}\left(N_i/c_{ij}\right)\right) & e_j \in M(e_i) \\ 0 & e_j \notin M(e_i) \end{cases} \tag{11}$$

设实体e_i的一阶邻域实体集合$M(e_i)$对应k_i种关系类型。式(11)中，$e_j \in M(e_i)$条件下的分母$\sum_{j=1}^{N_i}\left(N_i/c_{ij}\right) = k_i N_i$，此时式(11)简化为

$$a_{ij} = 1/\left(k_i c_{ij}\right) \tag{12}$$

实体特征的初始化方式$\boldsymbol{H}(0)$见3.2节。

2.6 层集成

使用每层聚合的实体嵌入向量来构造相似度矩阵。设对输入的每个在轨卫星预测k_{society}个候选匹配卫星网络资料，每层聚合的相似度向量$\boldsymbol{s}^l = \left[s_1^l,\ s_2^l,\ \cdots,\ s_{k_{\text{society}}}^l\right]$如下：

$$s_j^l = \cos\left(\boldsymbol{e}^{\text{sat},l}, \boldsymbol{e}_j^{\text{net},l}\right) \tag{13}$$

式中：$\boldsymbol{e}^{\text{sat},l}$为在轨卫星在第$l$层的实体嵌入向量；$\boldsymbol{e}_j^{\text{net},l}$为第$j$个卫星网络资料在第$l$层的实体嵌入向量，$l = 0, 1, \cdots, L$，$j = 1, 2, \cdots, k_{\text{society}}$。假设每层有相同的重要性，最终相似度向量$\boldsymbol{s}^*$集成为所有层的平均值，按相似度由大到小排序得到预测的匹配卫星网络资料列表。

式(14)为预测1个在轨卫星的匹配卫星网络资料结果，若输入为m个在轨卫星，最终相似度矩阵为m个相似度向量$\boldsymbol{s}^*$堆叠而成。

$$\boldsymbol{s}^* = \sum_{l=0}^{L} \boldsymbol{s}^l / L \tag{14}$$

3 实验结果与分析

本节进行各模块消融实验、运行时间对比实验和补充实验3组实验。采用人工整理的真实验证数据集通过消融实验和对比实验验证UPMIS方法的有效性，补充实验给出UPMIS方法中超参数的选取推荐值。

3.1 数据集

真实验证数据。通过人工整理筛选出中国关注度高的48颗在轨NGSO卫星与匹配卫星网络资

料的真实实体对, 作为 UPMIS 方法的真实验证数据。

在轨卫星知识图谱 K^{sat}。数据来源自 UCS-Satellite-Database-5-1-2023[30], 记录了目前绕地球运行的 7 560 颗卫星信息。由于在时间参数模块不涉及知识图谱, 表 1 中时间类信息不在 K^{sat} 的考虑范围。为便于实验将输入卫星的唯一标识符“卫星目录号”指标纳入 K^{sat}。

卫星网络资料知识图谱 K^{net}。数据来源自 ITU SRS 2 996 数据库[40]。与 K^{sat} 同理, 不包含时间类信息。为与 K^{sat} 对应, 增加了 SRS 数据库中卫星网络资料的唯一标识符“卫星网络通知资料编号”。

如 2.2 节所述卫星网络 N 资料是申报流程中的最终通知资料, 鉴于真实验证数据的范围为 NGSO 卫星, 将卫星网络资料的范围限定为 NGSO 卫星网络 N 资料。K^{sat} 与 K^{net} 的规模如表 2 所示。

表 2　K^{sat} 与 K^{net} 统计数据

Table 2　Statistics of K^{sat} and K^{net}

数据集	三元组个数	实体个数	关系个数	在轨卫星/卫星网络资料个数
K^{sat}	82 979	20 307	11	7 548
K^{net}	6 258	2007	6	628

3.2　实验设置

1) 评价指标。采用 H_k 作为评估指标, 表示真实的卫星网络资料在候选卫星网络资料列表中的排名小于 k 的比例。

$$H_k = \frac{1}{m}\sum_{i=1}^{m} I(r_{\text{ank}i} < k) \tag{15}$$

式中: $I(\cdot)$ 为指示函数, 若条件为真则函数值为 1, 否则为 0; $r_{\text{ank}i}$ 为第 i 个在轨卫星的真实匹配卫星网络资料在候选匹配卫星网络资料列表中的排名, $i = 1,2,\cdots,m$, m 为待预测匹配关系的在轨卫星数量。由于社会参数模块是 UPMIS 方法的终止模块, 该模块的评价指标 $H_{k_{\text{society}}}$ 为 UPMIS 方法的最终评价指标, 轨道参数模块的评价指标 $H_{k_{\text{orbit}}}$ 为中间评价指标。

2) 初始实体嵌入特征 $\boldsymbol{H}(0)$。受限于词汇表规模的限制, 通过词汇表获得的单词级特征存在严重的词汇表外(out-of-vocabulary, OOV)问题[41]。本文采取实体对齐领域[28,42] 常用的方法[39], 不仅考虑单词级特征, 也将实体名称的字符二元组作为字符级特征考虑进来, 二者拼接作为初始实体嵌入特征。

3) 超参数。作为一种无监督方法, UPMIS 不需要任何训练参数, 只使用轨道参数模块卫星网络资料过滤数量 k_{orbit}、预测的候选卫星网络资料数量 k_{society} 和聚合深度 L 这 3 个超参数。除额外说明, 实验统一使用以下默认设置: $k_{\text{orbit}} = 30$, $k_{\text{society}} = 10$, $L = 4$。

3.3　主要实验

3.3.1　各模块消融实验

表 3 为在轨卫星与卫星网络资料匹配关系预测的消融实验结果。黑色加粗字体表示所有模型的最好结果。实验条件为搭载 NVIDIA GeForce RTX 2080 Ti GPU 和 Intel(R) Core(TM) i9-10900X CPU @ 3.70GHz 的工作站。

表 3　消融实验结果

Table 3　Ablation experimental results

方法	模块			$k = k_{\text{orbit}}$	$k = k_{\text{society}}$	$H_{k_{\text{orbit}}}$	$H_{k_{\text{society}}}$	H_{10}
	时间参数模块	轨道参数模块	社会参数模块					
①	√		√					0.583 3
②			√					0.458 3
③		√		10		0.687 5		0.687 5
④			√		10		0.312 5	0.312 5
⑤	√	√		10		0.791 7		0.791 7
⑥	√		√		10		0.416 7	0.416 7
⑦		√	√	30	10	0.687 5	0.666 7	0.666 7
⑧(UPMIS)	√	√	√	30	10	0.937 5	0.854 2	**0.854 2**

表 3 中, 方法①和方法②为不对卫星网络资料知识图谱和在轨卫星知识图谱进行图划分, 而使用全图聚合条件下的消融实验, 评价指标取 H_{10}。方法③～方法⑧为考虑图划分条件下的消融实验。为进一步研究各模块的重要性, 在考虑图划分的条件下分别进行使用不同数量模块的消融实验。H_{10} 为各方法的最终评价指标, 对于最后一个模块为轨道参数模块的方法③和方法⑤, $H_{10} = H_{k_{\text{orbit}}}$, 为便与其他方法对比, 取轨道参数模块卫星网络资料过滤数量 $k_{\text{orbit}} = 10$。对于最后一个模块为社会参数的方法④和方法⑥～方法⑧, $H_{10} = H_{k_{\text{society}}}$。

1) 子图与全图对比。图划分与选择显著影响

实验正确率。考虑全部子图的方法⑧在所有对比方法中取得了最优的结果，H_{10}高达 0.854 2，较全图聚合方法①和方法②分别提升 0.270 9 和 0.395 9。对于考虑子图的场景，设计不同模块的消融实验，UPMIS 的预测准确率领先于只考虑 1 个模块的方法和 2 个模块的方法。通过关系类型将知识图谱分割为子图能有效对各类三元组进行建模，丰富在轨卫星与卫星网络资料的关联信息，以及对关联信息的分而治之是提高在轨卫星匹配关系预测正确率的有效手段。

2）轨道参数模块>时间参数模块>社会参数模块。与效果最优的 UPMIS 方法相比，删除轨道参数模块（方法⑥）的H_{10}大幅下降 0.437 5，删除时间参数模块（方法⑦）下降 0.187 5，而删除社会参数模块（方法⑤）的H_{10}下降最少。因此，对于在轨卫星匹配关系预测领域来说，3 个模块的贡献度排序为轨道参数>时间参数>社会参数，这一点从方法⑤～方法⑦的对比中也可看出。下面对每个模块的重要性进行进一步解释。

3）时间参数是必要的。对于只考虑全图聚合的方法②来说，添加时间参数模块后的方法①H_{10}提高了 0.125；对于只考虑单一子图的轨道参数模块方法③或社会参数模块方法④来说，添加时间参数模块后的方法⑤和方法⑥H_{10}均提高了 0.104 2。卫星的发射时间与卫星网络资料的接收时间严格受限于国际电信联盟相关规则的约束，业务知识的融入明确降低了挖掘卫星与卫星网络资料之间关系的复杂程度。

4）轨道参数是关键。仅考虑轨道子图（方法③）的H_{10}高达 0.687 5，加入时间参数模块后的方法⑤预测精度在对比模型中仅次于 UPMIS 模型，提升至 0.791 7。反映出数值型轨道类子图对匹配关系预测的贡献度明显高于字符型社会类子图，同理可见方法⑤与方法⑥。由于卫星轨道具有动态稳定性，一般不会出现在轨卫星大幅偏离原定轨道的情况，因此，轨道参数是卫星与卫星网络资料的重要特征。

值得注意的是，UPMIS 在轨道参数模块的$H_{k_{\text{orbit}}}$高达 0.937 5，只有 3 颗在轨卫星的真实匹配卫星网络资料没有在轨道参数模块所预测的候选卫星网络资料中，原因在于其真实匹配卫星网络资料的相似轨道过多。统计分析发现 3 颗在轨卫星在其轨道向量的正负 10% 偏移区间内分别有：106 份卫星网络资料、223 个卫星网络资料轨道面；101 份卫星网络资料、223 个卫星网络资料轨道面；137 份卫星网络资料、245 个卫星网络资料轨道面，而轨道参数模块卫星网络资料过滤数量k_{orbit}设置为 30，过多相似的轨道导致了轨道参数模块的预测误差。

5）社会参数是辅助。虽然只考虑社会子图（方法④）的H_{10}在所有对比模型最低，为 0.312 5，但社会组织类信息蕴含了卫星和卫星网络资料的背景及来源，社会参数模块多级邻域聚合的结构进一步加大了挖掘深度，新定义的邻接矩阵丰富了图卷积感知能力。UPMIS 与删除社会参数模块的方法⑤相比H_{10}提升 0.062 5。说明在时间信息和轨道信息的基础上，社会组织类信息的融入对于卫星网络资料匹配关系同样具有不可忽视的判定作用。

3.3.2 运行时间对比实验

图 2 分析了在不同环境下平均每个卫星各模块的平均运行时间，实验结果为 5 次重复实验的平均值。UPMIS 在 GPU 和 CPU 的环境下均为毫秒级的运算速度。从 3 个模块的对比结果来看，UPMIS 算法的主要运行时间消耗在轨道参数模块，这是因为时间参数模块通过条件判别就可以实现，社会参数模块的图传播虽然耗时但并不会随着算法动态更新而改变，因此，可以作为已知量提前生成，故此模块耗时最少。而轨道参数模块中需要计算输入的在轨卫星与待对比的全部卫星网络资料的距离，耗时占比最大。

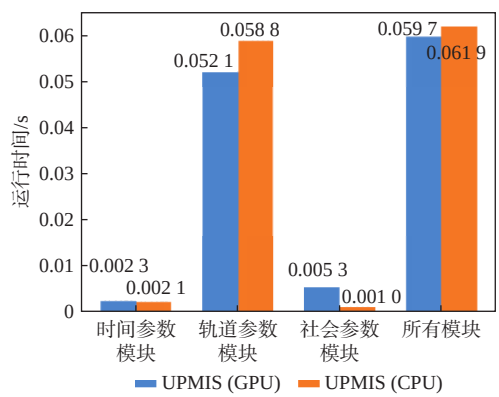


图 2 UPMIS 在 GPU 和 CPU 下的运行时间

Fig. 2 Time costs of UPMIS on GPU and CPU

为进一步探究时间参数模块节省的时间成本，将删除时间参数模块后的 UPMIS 作为对比，实验如图 3 所示，结果为 5 次重复实验的平均值。UPMIS 的运行时间仅为与不考虑时间参数模块的方法的 17%，说明融入业务时间逻辑的方法不仅能提高匹配关系预测的正确率，也能大幅提升算法的运行时间。

总的来说，UPMIS 是一种不需要任何训练参数的无监督机器学习方法。不同于在每个批处理均

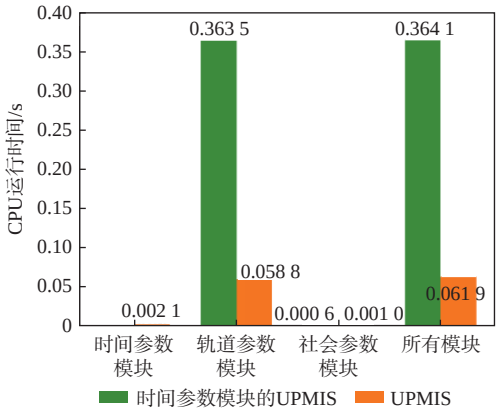


图 3　时间参数模块对 UPMIS 运行时间的影响

Fig. 3　Impact of time module on runtime of UPMIS

需要前向传播且经过上百轮批处理才会收敛的基于 GNN 类模型, UPMIS 没有任何训练参数, 只需要在其他类子图聚合模块中的一次前向传播即可实现上述效果, 不仅预测正确率高而且实现毫秒级运算, 具有高效性。

3.3.3　超参数实验

为探索不同情况下 UPMIS 的性能, 设计下列新的评价指标与实验, 给出超参数的推荐值。

1) 新评价指标 H_{H}。提出新的评价指标调和命中率(harmonic hits, HH)。由于 UPMIS 在轨道参数模块和社会参数模块均计算了 H_k, $k=k_{\mathrm{orbit}}$, $k=k_{\mathrm{society}}$, 且二者之间存在关联: 轨道参数模块评价指标 $H_{k_{\mathrm{orbit}}}$ 是社会参数模块评价指标 $H_{k_{\mathrm{society}}}$ 的上确界, 即 $H_{k_{\mathrm{society}}} \leqslant H_{k_{\mathrm{orbit}}}$。考虑到 $H_{k_{\mathrm{orbit}}} \leqslant 1$, $H_{k_{\mathrm{society}}}/H_{k_{\mathrm{orbit}}} \leqslant 1$, 借鉴分类任务[43-46]中 F_1 值的定义, 设计新的评价指标 H_{H} 如下:

$$H_{\mathrm{H}}=\frac{2}{1/H_{k_{\mathrm{orbit}}}+H_{k_{\mathrm{orbit}}}/H_{k_{\mathrm{society}}}} \tag{16}$$

式中: H_{H} 为 $H_{k_{\mathrm{orbit}}}$ 与 $H_{k_{\mathrm{society}}}/H_{k_{\mathrm{orbit}}}$ 的调和平均数。

2) 轨道距离函数 Dis。为验证在轨道参数模块中设计的轨道距离函数 Dis 的有效性, 使用传统均方误差 MSE 作为对比, 结果如表 4 所示。Dis 在 3 个评价指标中均实现最优的性能, 表中加粗显示。

表 4　不同轨道距离函数的性能

Table 4　Performances with different orbit distance functions

距离函数	$H_{k_{\mathrm{orbit}}}$	$H_{k_{\mathrm{society}}}$	H_{H}
MSE	0.916 667	0.812 5	0.901 3
Dis	**0.937 500**	**0.854 2**	**0.924 1**

3) 邻接矩阵 $\boldsymbol{A}_{\mathrm{rel}}$。为验证在社会参数模块中设计的邻接矩阵 $\boldsymbol{A}_{\mathrm{rel}}$ 的有效性, 表 5 为 UPMIS 在不同邻接矩阵下的预测结果。$\boldsymbol{A}$ 为标准邻接矩阵, $\boldsymbol{A}_{\mathrm{L}}=\boldsymbol{D}-\boldsymbol{A}$ 为拉普拉斯矩阵, 其中, $\boldsymbol{D}$ 为度矩阵。由于邻接矩阵对轨道参数模块没有影响, 评价指标为社会参数模块中的 $H_{k_{\mathrm{society}}}$ 和新评价指标 H_{H}。$\boldsymbol{A}_{\mathrm{rel}}$ 在所有评价指标中均最优, 表中加粗显示。

表 5　UPMIS 在不同邻接矩阵下的性能

Table 5　Performances of UPMIS under different adjacency matrices

邻接矩阵	$H_{k_{\mathrm{society}}}$	H_{H}
$\boldsymbol{A}$	0.395 8	0.582 2
$\boldsymbol{A}_{\mathrm{L}}$	0.750 0	0.863 3
$\boldsymbol{A}_{\mathrm{rel}}$	0.854 2	**0.924 1**

4) 轨道参数模块卫星网络资料过滤数量 k_{orbit}。图 4 展示了选取最优 k_{orbit} 的实验结果。显而易见, 随着 k_{orbit} 的增加轨道参数模块评价指标 $H_{k_{\mathrm{orbit}}}$ 稳步上升, 并在 $k_{\mathrm{orbit}}=50$ 后收敛到 1, 说明只要在轨卫星与卫星网络之间的距离约束地足够宽松, 其匹配卫星网络资料必在候选集合中。随着 k_{orbit} 的增加社会参数模块评价指标 $H_{k_{\mathrm{society}}}$ 先上升后下降, 说明通过轨道子图过滤出的卫星网络资料的数量并非越多越好, 当卫星网络资料数量多到一定数目后反而对社会参数模块的结果造成干扰, 降低最终预测精度。UPMIS 的 H_{H} 在 $k_{\mathrm{orbit}}=30$ 时达到最大值, 因此, 取 $k_{\mathrm{orbit}}=30$。

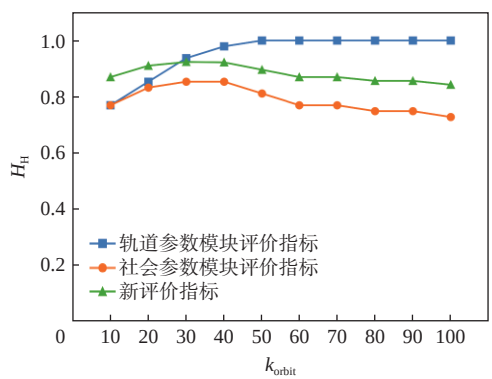


图 4　轨道参数模块卫星网络资料过滤数量 k_{orbit} 对 UPMIS 性能的影响

Fig. 4　Impact of filtered satellite network data sets k_{orbit} on UPMIS performance

5) 聚合深度 L。对于聚合深度 L 的结果如图 5 所示。由于聚合深度 L 只影响社会参数模块, 因此, 更改 L 并不会影响轨道参数模块评价指标 $H_{k_{\mathrm{orbit}}}$。特别地, $L=0$ 时等价于只考虑实体本身而不考虑其邻居信息。一般来说, 随着聚合深度的增加模型效果在初期会有大幅提升, 说明卫星之间、卫星网络资料之间、卫星与卫星网络之间的关系并非是浅层的, 需要深度聚合才能实现更多信息的有效关联。与图神经网络方法相同, UPMIS 在后期也受到过平滑影响。由图 5 可知, UPMIS 在 L=4 时取得最优性能, 因此, 取 L=4。

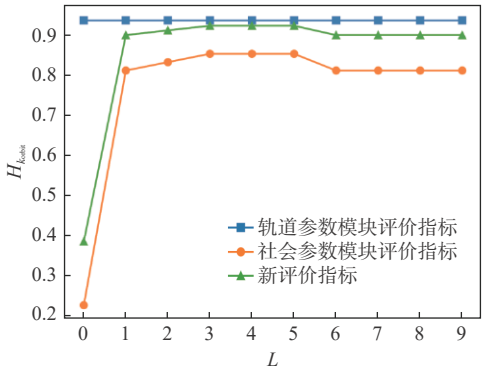


图5 聚合深度L对UPMIS性能的影响

Fig. 5 Impact of aggregation depths L on UPMIS performance

4 结 论

1）本文面向在轨卫星与卫星网络资料匹配关系预测任务提出UPMIS方法，实现在轨卫星匹配关系高效预测。

2）UPMIS方法由业务时间参数模块、数值型轨道参数模块和字符型社会参数模块组成，对于每个模块针对性地设计过滤规则。

3）时间参数模块严格排除不符合业务时间逻辑的卫星网络资料；轨道参数模块充分考虑不同轨道参数的重要特征设计了新型轨道距离函数；社会参数模块改进了邻接矩阵的构造方式，加强模型对邻域的感知能力。

4）作为一种无神经网络的无监督方法，实验证明，UPMIS方法领先对比方法并实现了毫秒级的运算效果，具高效性、高速性和稳定性。

参考文献（References）

[1] 袁俊, 鲍晓月, 孙茜, 等. 巨型低轨星座频率轨道资源趋势分析及启示建议[J]. 空间碎片研究, 2021, 21(1): 48-57.
YUAN J, BAO X Y, SUN Q, et al. Analysis and suggestions on orbit and spectrum resources trend for mega LEO constellations[J]. Space Debris Research, 2021, 21(1): 48-57(in Chinese).

[2] 田伟, 陈珊珊, 林源. 空间频率轨道资源法规标准建设情况初探[J]. 中国航天, 2019(8): 54-62.
TIAN W, CHEN S S, LIN Y. Preliminary analysis on the regulations and standards for space spectrum and orbit resources[J]. Aerospace China, 2019(8): 54-62(in Chinese).

[3] International Telecommunication Union. Radio regulations[EB/OL]. (2023-03-16)[2024-01-25]. https://www.itu.int/pub/R-REG-RR.

[4] 王羽, 李清, 李克军, 等. Starlink星座应用现状及分析[J]. 天地一体化信息网络, 2023, 4(2): 93-102.
WANG Y, LI Q, LI K J, et al. Application status and analysis of starlink constellation[J]. Space-Integrated-Ground Information Networks, 2023, 4(2): 93-102(in Chinese).

[5] MCDOWELL J C. The low earth orbit satellite population and impacts of the SpaceX starlink constellation[J]. The Astrophysical Journal Letters, 2020, 892(2): L36.

[6] 崔潇潇, 王琦, 范青松. 从“一网”(OneWeb)卫星 探究卫星互联网面临的风险与挑战[J]. 国防科技工业, 2022(6): 56-58.
CUI X X, WANG Q, FAN Q S. Exploring the risks and challenges faced by satellite Internet from the perspective of “OneWeb” satellites[J]. Defence Science & Technology Industry, 2022(6): 56-58(in Chinese).

[7] 王学宇, 武坦然. OneWeb低轨道卫星系统及其军事应用分析[J]. 航天电子对抗, 2022, 38(4): 59-64.
WANG X Y, WU T R. Analysis on OneWeb LEO satellite system and its military application[J]. Aerospace Electronic Warfare, 2022, 38(4): 59-64(in Chinese).

[8] RADTKE J, KEBSCHULL C, STOLL E. Interactions of the space debris environment with mega constellations: using the example of the OneWeb constellation[J]. Acta Astronautica, 2017, 131: 55-68.

[9] 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).

[10] ZHANG J R, CAI Y F, XUE C B, et al. LEO mega constellations: review of development, impact, surveillance, and governance[J]. Space: Science & Technology, 2022: 9865174.

[11] DAEHNICK C, KLINGHOFFER I, MARITZ B, et al. Large LEO satellite constellations: will it be different this time?[EB/OL]. (2020-05-04)[2024-01-25]. https://www.mckinsey.com.br/industries/aerospace-and-defense/our-insights/large-leo-satellite-constellations-will-it-be-different-this-time.

[12] FALLE A, WRIGHT E, BOLEY A, et al. One million (paper) satellites[J]. Science, 2023, 382(6667): 150-152.

[13] 王琦. “纸卫星”: 卫星产业面临的一道难题[J]. 卫星与网络, 2010(6): 62-64.
WANG Q. “Paper satellite”: a difficult problem faced by satellite industry[J]. Satellite & Network, 2010(6): 62-64(in Chinese).

[14] 田伟. 关注“纸卫星”[J]. 卫星与网络, 2014(4): 68-72.
TIAN W. Pay attention to “paper satellite”[J]. Satellite & Network, 2014(4): 68-72(in Chinese).

[15] ALLISON A L. The advent of paper satellites[C]//Proceedings of the ITU and Managing Satellite Orbital and Spectrum Resources in the 21st Century. Berlin: Springer, 2014: 25-29.

[16] Space. com Staff. SpaceX’s starlink broadband service will begin in 2020: report [EB/OL]. (2019-10-24) [2024-01-24]. https://www.15minutenews.com/article/2019/10/24/172984982/spacexs-starlink-broadband-service-will-begin-in-2020-report/.

[17] BUNGLAWALA W, SHAH J, PARMAR D. Recommender system using knowledge graph and ontology: a survey[C]//Proceedings of the Expert Clouds and Applications. Berlin: Springer, 2022: 237-251.

[18] PEREIRA A, TRIFAN A, LOPES R P, et al. Systematic review of question answering over knowledge bases[J]. IET Software, 2022, 16(1): 1-13.

[19] LUO S X, YU S. An accurate unsupervised method for joint entity alignment and dangling entity detection[C]//Proceedings of the Findings of the Association for Computational Linguistics. Kerrville: Association for Computational Linguistics, 2022: 2330-2339.

[20] WU Y T, LIU X, FENG Y S, et al. Relation-aware entity alignment for heterogeneous knowledge graphs[C]//Proceedings of the 28th International Joint Conference on Artificial Intelligence. California: International Joint Conferences on Artificial Intelligence Organization. 2019: 5278-5284.

[21] XU K, WANG L W, YU M, et al. Cross-lingual knowledge graph alignment via graph matching neural network[EB/OL]. ((2019-07-29)[2024-03-04]. https://arxiv.org/abs/1905.11605.

[22] CAO Y X, LIU Z Y, LI C J, et al. Multi-channel graph neural network for entity alignment[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Stroudsburg: Association for Computational Linguistics. 2019: 1452–1461.

[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] ZHAO X, ZENG W X, TANG J Y, et al. An experimental study of state-of-the-art entity alignment approaches[J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 34(6): 2610-2625.

[25] GAUR M, FALDU K, SHETH A. Semantics of the black-box: can knowledge graphs help make deep learning systems more interpretable and explainable?[J]. IEEE Internet Computing, 2021, 25(1): 51-59.

[26] WU X D, WU J, FU X Y, et al. Automatic knowledge graph construction: a report on the 2019 ICDM/ICBK contest[C]//Proceedings of the IEEE International Conference on Data Mining. Piscataway: IEEE Press, 2020: 1540-1545.

[27] STOKMAN F N, DE VRIES P H. Structuring knowledge in a graph[C]//Proceedings of the Human-Computer Interaction. Berlin: Springer, 1988: 186-206.

[28] WANG Z C, YANG J J, YE X J. Knowledge graph alignment with entity-pair embedding[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg: Association for Computational Linguistics, 2020: 1672-1680.

[29] ZHONG L F, WU J, LI Q, et al. A comprehensive survey on automatic knowledge graph construction[J]. ACM Computing Surveys, 2024, 56(4): 1-62.

[30] Union of Concerned Scientists . Satellite database | union of concerned scientists[EB/OL]. (2024-01-02) [2024-01-24]. https://www.ucs.org/resources/satellite-database.

[31] HENRI Y, MATAS A, DOS SANTOS J M S. Regulation of telecommunications by satellites: ITU and space services[M]. Abingdon: Routledge, 2016: 129-163.

[32] ROTHBLATT M A. ITU regulation of satellite communication[J]. Stanford Journal of International Law, 1982, 18: 1.

[33] DODEL I B H. The ITU filing of satellite systems[M]. Abingdon: Routledge, 2011.

[34] ALLISON A L. Requirements for obtaining spectrum and of orbital approvals for small satellite constellations[C]//Proceedings of the Handbook of Small Satellites: Technology, Design, Manufacture, Applications, Economics and Regulation. Berlin: Springer, 2020: 1-23.

[35] LOO C C, WANG X. Navigating the spectrum: an overview of ITU's regulatory process for small satellites[C]//Proceedings of the 38th Annual Small Satellite Conference. Logan: Utah State University, 2024: 134.

[36] ANSELMO L. The use of space and satellites: problems and challenges[C]//Proceedings of the Technology and International Relations. Berlin: Springer, 2021: 109-131.

[37] KIPF T N, WELLING M. Semi-supervised classification with graph convolutional networks[EB/OL]. (2017-02-22)[2024-03-04]. https://arxiv.org/abs/1609.02907.

[38] MAO X, WANG W T, XU H M, et al. MRAEA: an efficient and robust entity alignment approach for cross-lingual knowledge graph[C]//Proceedings of the 13th International Conference on Web Search and Data Mining. New York: ACM, 2020: 420-428.

[39] MAO X, WANG W T, WU Y B, et al. From alignment to assignment: frustratingly simple unsupervised entity alignment[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg: Association for Computational Linguistics, 2021: 2843-2853.

[40] International Telecommunication Union. SRS[EB/OL]. (2023-03-16)[2024-03-04]. https://www.itu.int/en/ITU-R/software/Pages/srs.aspx.

[41] GARCIA-BORDILS S, MAFLA A, BITEN A F, et al. Out-of-vocabulary challenge report[EB/OL]. (2022-09-14)[2024-03-04]. https://arxiv.org/abs/2209.06717.

[42] LIU Z Y, CAO Y X, PAN L M, et al. Exploring and evaluating attributes, values, and structures for entity alignment[C]//Proceedings of the Conference on Empirical Methods in Natural Language Processing. Stroudsburg: Association for Computational Linguistics , 2020: 6355-6364.

[43] LI X, WANG Y Y, ACERO A. Learning query intent from regularized click graphs[C]//Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2008: 339-346.

[44] CHINCHOR N. MUC-3 evaluation metrics[C]//Proceedings of the 3rd Conference on Message Understanding. New York: ACM, 1991: 17-24.

[45] FANG J Y, WANG X B, MENG Z Q, et al. MANNER: a variational memory-augmented model for cross domain few-shot named entity recognition[C]//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Stroudsburg: Association for Computational Linguistics, 2023: 4261-4276.

[46] ZHENG C M, FENG J H, CAI Y, et al. Rethinking multimodal entity and relation extraction from a translation point of view[C]//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). Stroudsburg: Association for Computational Linguistics , 2023: 6810-6824.

Prediction method for matching between in-orbit satellites and satellite network filings based on knowledge graph

XU Cong[1, 2], SHI Mengxin[3], WANG Hongfeng[4], JIA Qingyu[1], ZHI Jia[1], YANG Jiasen[1, *]

(1. Key Laboratory of Electronics and Information Technology for Space Systems, National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China;
2. School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing 100049, China;
3. School of Information Science and Engineering, Southeast University, Nanjing 210096, China;
4. China Satellite Launch and Tracking Control General, Beijing 100094, China)

Abstract: Matching in-orbit satellites with International Telecommunication Union (ITU) satellite network declaration filings is crucial for the design, selection, declaration, and coordination of satellite frequencies and orbits. Due of their low matching efficiency and high domain knowledge requirements, traditional manual matching algorithms frequently encounter difficulties. To address these issues, we propose an unsupervised prediction of matching between in-orbit satellites and satellite network filings (UPMIS) method. This method establishes a prediction indicator system and knowledge graphs for both in-orbit satellites and satellite network filings. By integrating domain knowledge and graph partitioning, we design a three-tier filtering framework comprising a time parameter module, a numerical orbit parameter module, and a character-based societal parameter module. This framework enables fast and accurate matching between in-orbit satellites and satellite network filings. Experimental results demonstrate that UPMIS achieves a H_{10} score of 0.854 2 on real datasets, outperforming other comparative models. Additionally, the average runtime reaches millisecond-level efficiency. Additionally, the trials offer helpful references for future matching relationship predictions by recommending values for parameters like filtering quantity and aggregation depth.

Keywords: matching prediction; knowledge graph; knowledge fusion; unsupervised learning; graph partition; in-orbit satellites; satellite network filings

Received: 2024-04-15; **Accepted**: 2024-05-31; **Published Online**: 2024-06-17 11:22

URL: link.cnki.net/urlid/11.2625.V.20240616.1430.001

Foundation items: National Natural Science Foundation of China (62341106); Beijing Municipal Natural Science Foundation (L222003)

* **Corresponding author**. E-mail: jsy@nssc.ac.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0289

基于多源图像融合的无人机航拍舰船目标识别方法

姜杰，闫文君，刘凯，张立民*
(海军航空大学 航空作战勤务学院，烟台 264001)

摘 要：针对无人机航拍采集到的多源舰船图像，提出一种多源舰船图像融合识别方法，面对现实场景中存在诸多干扰，采用像素级融合方式，通过对红外、可见光舰船图像进行融合，再进行目标识别，相比于特征级识别方法，既可以降低网络对样本的依赖，又可以增强方法的可解释性。所提方法聚焦于解决因传感器参数不同存在的像素偏移情况，为避免传统图像配准容易产生的畸变、伪影等问题，将图像配准转化为端到端下的特征对齐，提出一种多源舰船图像融合识别网络，网络共由交叉调制特征提取模块、特征动态对齐模块、多粒度特征细化模块和金字塔特征融合模块组成，可以充分融合不同模态图像的特征和纹理细节，有效提升对舰船目标的识别性能。经实验验证：所提方法对多源图像的融合效果好、可解释性强，对舰船目标的识别准确度高、鲁棒性强。

关 键 词：无人机；多源舰船图像；融合识别；像素级融合；像素偏移
中图分类号：V391.4
文献标志码：A **文章编号**：1001-5965(2026)06-1955-10

当今随着无人机技术的不断发展，其在各个领域的应用也愈加广泛，特别是在军事侦察方面扮演着越来越重要的角色，通过配备的各类传感器无人机可以实时地将侦测到的信息回传给指控系统，这对于掌握战场态势、迅速打击敌方目标有着重要意义，因此，基于无人机航拍图像处理技术是目前研究的重点，其中，通过无人机配备的多源传感器实现图像信息的融合识别是亟须解决的热点问题[1-2]。海军航空大学何友院士曾指出当今军事信息技术的高速发展，使得“信息”已成为夺取战争主动权的关键要素之一，信息感知与融合技术则是夺取战场信息优势的基石与先导[3]。多源信息融合通过融合来自不同源的信息，可以提高整体信息的质量和可靠性，减少噪声和误差的影响，融合后的信息可以提供多角度、更全面的信息，有助于提高决策的准确性和可信度，增强系统的鲁棒性[4-5]。

实现多源图像融合面临的难点问题主要包括：①数据的异质性问题，来自不同传感器或模态的图像数据通常具有不同的分辨率、噪声水平、亮度和对比度等特征，需要进行有效的归一化和融合处理[6]；②特征的提取与选择问题，不同源的图像可能包含大量冗余信息，如何有效地提取并选择最具代表性的特征进行融合是一个难点[7]；③融合规则和权重确定问题，确定不同源图像之间的融合规则和权重分配是一个挑战，需要考虑不同源之间的相关性、重要性及噪声的影响[8]；④实时性和计算复杂度问题，对于实时应用，多源图像融合需要在有限的时间内完成，同时需要考虑计算复杂度和资源消耗[9-10]。针对上述问题的解决方法目前共分为3类：①像素级融合识别，文献[11]提出一种基于小波变换的像素融合方法，通过将图像分解成不同频率的

收稿日期：2024-05-07；**录用日期**：2024-06-21；**网络出版时间**：2024-08-15 13:53
网络出版地址：link.cnki.net/urlid/11.2625.V.20240815.0751.001
基金项目：国家自然科学基金(62371465)；山东省青创团队(2022kj084)；山东省自然科学基金(ZR2020QF010)
*通信作者. E-mail：iamzlm@hotmail.com

引用格式：姜杰，闫文君，刘凯，等. 基于多源图像融合的无人机航拍舰船目标识别方法[J]. 北京航空航天大学学报，2026，52(6)：1955-1964. JIANG J，YAN W J，LIU K，et al. Ship target recognition method based on multi-source image fusion for unmanned aerial vehicle aerial photography[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1955-1964 (in Chinese).

子图像然后采用不同的融合策略重构出新的图像，文献[12]提出一种基于潜在低阶表示的自适应尺度像素融合方法，将红外和可见光图像分解成2个尺度表示来构造自适应权值用以图像重建；②特征级融合识别，文献[13]提出一种基于边缘卷积滤波与合成孔径雷达(synthetic aperture rader, SAR)及光学多特征分类的检测方法，文献[14]设计了一种基于不变特征的风格转换网络，利用多源数据间的共享知识学习不变特征实现信息互补，再采用通用网络完成目标检测；③决策级融合识别，文献[15]提出一种采用有监督和无监督分类器相结合的方法进行信息提取，再通过支持向量机完成决策的选择，文献[16]提出利用社会网络分析多模态生物特征的模式，根据多源图像的类间特征和类内特征进行决策融合。

上述方法一定程度上实现了对多源图像的融合识别，但是海洋环境下舰船目标识别任务相比于其他目标具有更多需要解决的问题，具有更加复杂的背景，存在诸如海浪、云雾等复杂的视觉干扰因素，且舰船之间存在相互遮挡，不同角度和光照条件下的舰船外形变化很大，特别是航拍视角下的舰船目标大小不一，小目标占比高，以上问题都大大增加了舰船目标识别的难度[17-20]。对于上述情况，本文提出一种基于多源图像融合的无人机航拍舰船目标识别方法，采用像素级融合的方式，将无人机采集到的红外、可见光图像融合，再完成舰船目标的识别，相比于直接将多源图像输入网络进行训练输出识别结果的方式，本文方法既可以得到融合后的图像，依据战场形势需要辅助人工判别；又可以直观地感受到方法融合后的效果，特别是在处理现实场景中，目标存在遮挡、光照条件不足、海面干扰较大等情形，本文方法可解释性更高、鲁棒性更强，不依赖于样本数量和识别网络的性能，在现实应用中更加便捷。

1 多源舰船图像融合识别方法

多源舰船图像融合识别(mutisource ship image fusion recognition, MSIFR)方法，聚焦于解决因传感器参数不同导致所采集到的图像存在像素偏移的情况，在传统方法中通常采用先对采集到的不同模态的图像进行配准再融合的方式，这种方式严重依赖于配准方法是否精准，既增加了工作量也会存在配准错误的情况，并且融合后的图像也因红外、可见光图像结构上的差异出现部分的扭曲变形和伪影，这将大大加大对目标识别的难度[21]，本文采用端对端的架构结合卷积神经网络可以实现非严格对齐的红外可见光图像的融合。MSIFR方法流程如图1所示，首先，用特征对齐替代像素配准，以可见光特征作为参考，在特征对齐后添加交叉调制(cross modulation, CM)策略[22]，利用空间感知相互间的潜在关系，从而生成调制参数来对特征进行调制，平移的相关性可以被自适应测量，平移特征也会被动态提取，在特征动态对齐模块之后，通过自适应学习卷积核让红外图像在空间上和可见光特征更加一致，最后，在特征重塑阶段，多粒度特征细化模块根据局部相似性进行设计，以加强纹理相关性信息。

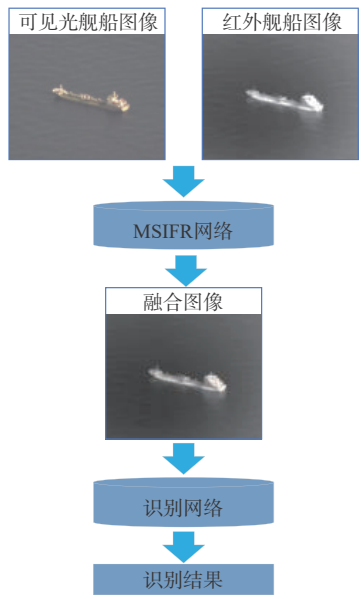


图1 MSIFR方法流程

Fig. 1 MSIFR method flow

1.1 网络框架

红外可见光融合网络由交叉调制特征提取模块(cross-modulation feature extraction module, CMFEM)、特征动态对齐模块(feature dynamic alignment module, FDAM)、多粒度特征细化模块(multi-grained feature refinement module, MGFRM)和金字塔特征融合模块(pyramid feature fusion module, PFFM)组成。在CMFEM中，嵌入了交叉调制策略，旨在提取2个源图像特征之间的潜在关系，生成调制参数进行调制，根据调制后的特征，通过FDAM为红外图像生成动态卷积核，然后，将红外特征与可见光特征进行空间对齐，在MGFRM之后，通过增强可见光形态中的纹理信息，并基于相似性度量抑制红外形态中的非肌理信息，进一步细化了在多尺度下获得的特征。最后，融合多个层次特征，分别通过PFFM和残差块获得融合图像，该图像没有伪影，具有丰富的纹理。网络模型如图2所示，其中，DKG表示动态内核生成器，RB为标准残差块。

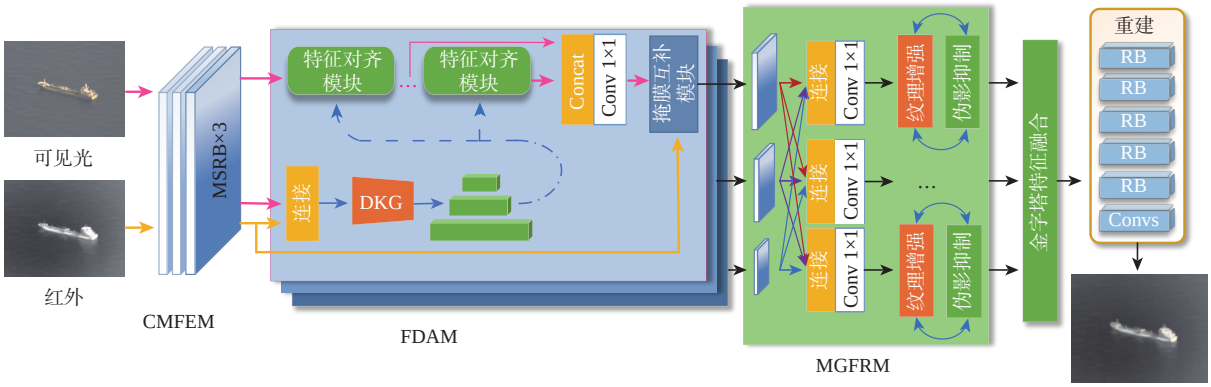


图 2　网络模型

Fig. 2　Network model

1.2　交叉调制特征提取模块

对于源图像的配准采用高级特征对齐的方式, 由于任何一对图像在位置和方向上都满足不同的移动像素, 因此, 具有固定参数的编码器无法对这些不同的情况进行建模。为解决这个问题, 一种合理的方法是根据其在空间中的潜在关系动态提取特征, 因此, 引入了一种交叉调制策略, 旨在揭示潜在关系然后生成一组动态调制参数。

浅层特征具有空间细节, 而深层特征获得更多的语义信息, 为利用其互补信息, 设计了基于多尺度残差块(multi scale residual block, MSRB)的金字塔特征提取器。MSRB 将不同尺度的特征结合起来, 以增强特征的语义信息, 模块由 3 个部分组成, 第 1 部分由 3 个 MSRB 和 1 个交叉调制策略组成, 第 2 部分和第 3 部分都包含 2 个 MSRB 和 1 个交叉调制策略, 根据这个模块, 可以获得 2 个源图像的特征。在数学上, 将未对准的可见光和红外图像分别表示为 $\boldsymbol{x}_{\mathrm{vi}} \in \mathbf{R}^{H\times W}$ 和 $\boldsymbol{x}_{\mathrm{ir}} \in \mathbf{R}^{H\times W}$, 其中, H 和 W 分别为源图像的高度和宽度, 通过第 1 部分将其转发给 3 个 MSRB, 因此, 可以得到

$$\begin{cases} \boldsymbol{F}_{1,\mathrm{vi}} = D_1(\boldsymbol{x}_{\mathrm{vi}}) \\ \boldsymbol{F}_{1,\mathrm{ir}} = D_2(\boldsymbol{x}_{\mathrm{ir}}) \end{cases} \tag{1}$$

式中: D_1、D_2 由第 1 部分的 MSRB 组成; $\boldsymbol{F}_{1,\mathrm{vi}}$ 为可见光图像特征; $\boldsymbol{F}_{1,\mathrm{ir}}$ 为红外图像特征。

基于动态参数, 特征 $\boldsymbol{F}_{1,\mathrm{vi}}$、$\boldsymbol{F}_{1,\mathrm{ir}}$ 被交叉调制, 极大地利于后续步骤中特征的完美对齐。交叉调制模块架构如图 3 所示, 其中, 包含卷积层和激活函数, 将 $\boldsymbol{F}_{l,\mathrm{vi}}$、$\boldsymbol{F}_{l,\mathrm{ir}}(l=\{1,2,3\})$ 表示为上述 3 个部分的 MSRB 的输出:

$$\boldsymbol{F}^{c}_{l,\mathrm{vi+ir}} = \mathrm{concat}(\mathrm{conv}(\boldsymbol{F}_{l,\mathrm{vi}}, 3\times 3, 2), \mathrm{conv}(\boldsymbol{F}_{l,\mathrm{ir}}, 3\times 3, 2)) \tag{2}$$

式中: 3×3 为卷积核大小, 2 为步长。

1.3　特征动态对齐模块

特征映射是通过加权方式对接受域中的元素

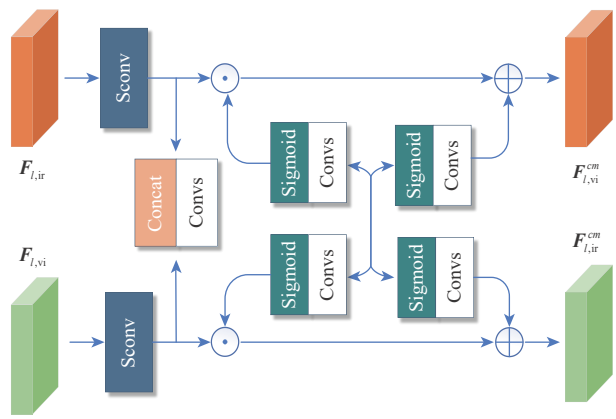


图 3　CMFEM 架构

Fig. 3　Architecture of CMFEM

进行总结而获得的, 因此, 卷积核在特征对齐中起着重要作用, 如果模型已经过训练, 现有的工作会遇到固定的内核, 不同的配对会遇到不同的平移错位情况。具有固定参数的卷积核无法满足要求, 因此, 本文提出一种基于移位感知的动态核学习策略, 不学习具有固定参数的内核, 这样可以根据不同的调制特征对自动估计核中的参数。

DKG[23] 架构如图 4 所示, 不同尺度下 3 个动态内核由 3 个分支生成, 对于平移不对齐的多重情况具有鲁棒性。通过每个分支的 B_k(每个 B_k 包含有一个卷积层、全局平均池化(global average pooling, GAP)算子和一个完全连接层)转发 $\boldsymbol{F}^{\mathrm{align}}_{l,\mathrm{ir}}$ 的连接, 相关动态内核的计算式为

$$\boldsymbol{\theta}_k(2k+1) = B_k(\mathrm{convs}(\mathrm{concat}(\boldsymbol{F}^{cm}_{l,\mathrm{vi}}, \boldsymbol{F}^{cm}_{l,\mathrm{ir}}))) \tag{3}$$

式中: $\boldsymbol{\theta}_k(2k+1)$ 表示动态内核的大小为 $(2k+1)\times(2k+1)$; c、m 为卷积层参数。

根据学习到的动态核对红外特征进行卷积, 红外特征可以很好地与可见光特征对齐。特别是, 假设内核为 $\boldsymbol{\theta}_k$, 其参数为 $\boldsymbol{\theta}_k(2k+1)$, 通过将 3 个动态卷积核得到的特征图连接起来, 利用 Conv(1×1)融合

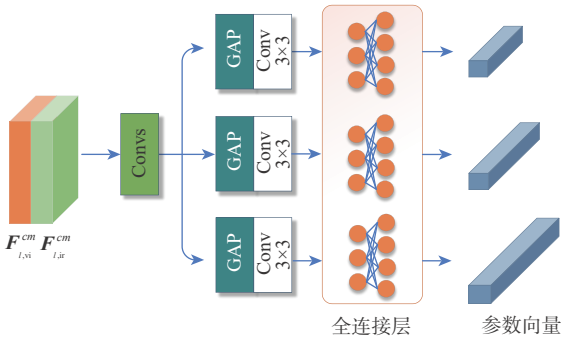


图 4 DKG 架构

Fig. 4 Architecture of DKG

多尺度特征，然后，将融合后的特征对通道注意力模块（channel attention module, CAM）[24] 进行归一化，提高了不同信道的重要性。此外，采用多个特征对齐块，逐步细化红外特征，获得更好的对齐，最后，采用可见光模态作为参考，得到对齐的红外特征 $\boldsymbol{F}_{l,\mathrm{ir}}^{\mathrm{align}}$ 及其对应的重建图像 $\boldsymbol{x}_{l,\mathrm{ir}}^{\mathrm{align}}$，利用 Conv（1×1）的输入绘制来自不同特征对齐块的多个对齐特征的组合。为实现 $\boldsymbol{x}_{l,\mathrm{ir}}^{\mathrm{align}}$ 与 ground-truth $(2k+1)\times(2k+1)$ 的完全对齐，引入 L_1 损失如下：

$$L_{\mathrm{align}}=\sum_{l=1}^{3}\left\|\hat{\boldsymbol{x}}_{l,\mathrm{ir}}-\boldsymbol{x}_{l,\mathrm{ir}}^{\mathrm{align}}\right\|_1 \tag{4}$$

在得到 $\boldsymbol{F}_{l,\mathrm{vi}}^{cm}$ 及其对准红外特征 $\boldsymbol{F}_{l,\mathrm{ir}}^{\mathrm{align}}$ 后，特征融合应该遵循同时包含文本信息和热性能信息来获得特征，一种简单的方法是连接或添加 2 种类型的特征，虽然这种策略很容易实现，但源图像中的冗余互信息也会混淆网络的学习，这限制了融合图像的质量，因此，提出一个互补掩模融合（complementary mask fusion, CMF）[25] 策略，将感兴趣的信息从 2 种模态转换为融合图像。在数学上，如图 5 所示，先将 $\boldsymbol{F}_{l,\mathrm{vi}}^{cm}$ 和 $\boldsymbol{F}_{l,\mathrm{ir}}^{\mathrm{align}}$ 连接起来，然后，通过 3 个 3×3 卷积层和 1 个 Sigmoid 激活函数将它们输入，从而得到一个注意力图 $\boldsymbol{A}_l$。得到了不同尺度下的融合特征：

$$\boldsymbol{F}_{l,\mathrm{vi+ir}}=\boldsymbol{F}_{l,\mathrm{ir}}^{\mathrm{align}}\odot\boldsymbol{A}_l+\boldsymbol{F}_{l,\mathrm{vi}}^{cm}\odot(\boldsymbol{I}-\boldsymbol{A}_l) \tag{5}$$

式中：$\boldsymbol{I}$ 为单位矩阵；⊙表示融合操作。

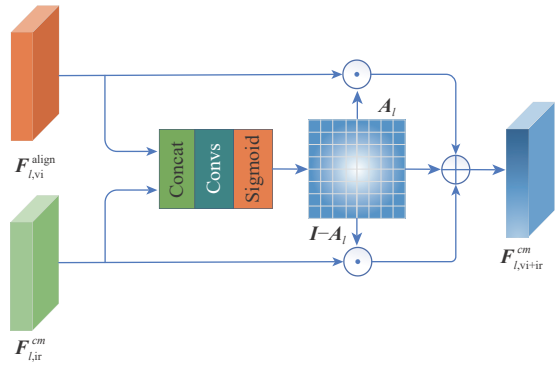


图 5 CMF 架构示意图

Fig. 5 Architecture of CMF

1.4 多粒度特征细化模块

多粒度特征细化模块可以测量源图像特征与局部区域融合特征之间的相似性，从而进一步去除红外图像中的伪影，并增强可见图像的纹理信息。为进一步结合不同层次融合特征的细节和语义信息，首先，融合特征 $\{\boldsymbol{F}_{1,\mathrm{vi+ir}},\boldsymbol{F}_{2,\mathrm{vi+ir}},\boldsymbol{F}_{3,\mathrm{vi+ir}}\}$，得到 $\boldsymbol{F}_{l,\mathrm{vi+ir}}^{f}\in\{\boldsymbol{F}_{1,\mathrm{vi+ir}}^{f},\boldsymbol{F}_{2,\mathrm{vi+ir}}^{f},\boldsymbol{F}_{3,\mathrm{vi+ir}}^{f}\}$，这样 $\boldsymbol{F}_{l,\mathrm{vi+ir}}^{f}$ 可以作为局部特征细化（local feature refinement, LFR）[26] 策略的特征图输入，f 为融合特征参数。如图 6 所示，LFR 分别由可见光和红外成像相关的 2 个分支组成，以可见光图像为基准，其不仅指导红外特征的对齐，还为融合特征提供纹理信息。对于可见光模态，基于可见光特征和融合特征之间的区域相似性重建融合特征，以增强纹理相关特征；对于红外模态，由于原始红外图像的边缘与伪影高度相关，利用红外图像的梯度图通过测量其梯度特征和融合特征之间的相似性来抑制融合图像中的伪影。

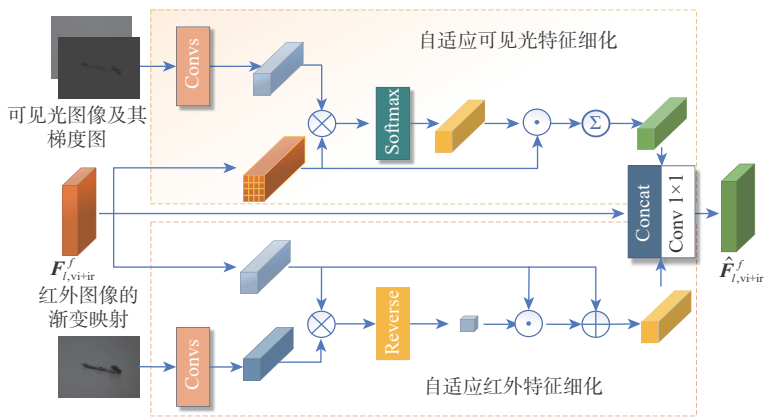


图 6 LFR 架构

Fig. 6 Architecture of LFR

对于自适应可见光特征细化分支, 通过测量可见光特征和融合特征之间的局部相似性以更好地重建。将可见光图像及其梯度图通过 3 个 3×3 卷积层获得的特征表示为 $\boldsymbol{F}^{\mathrm{e}}_{l,\mathrm{vi}}$, 以及每个属于空间位置$(i,j)$的特征向量 $\boldsymbol{F}^{\mathrm{e}}_{l,\mathrm{vi}}$表示为 $\boldsymbol{f}^{ij}_{l,\mathrm{vi}}$, 为测量局部相似性, 从融合的特征 $\boldsymbol{F}^{f}_{l,\mathrm{vi+ir}}$找到相应的以相同位置$(i,j)$为中心的$k\times k$个补丁, 该补丁的每个特征向量 $\boldsymbol{f}^{ij}_{l,\mathrm{vi+ir}}$相似度计算式为

$$\boldsymbol{s}^{ij}_{l}(x,y)=\frac{(\boldsymbol{f}^{ij}_{l,\mathrm{vi+ir}}(i+x,j+y))^{\mathrm{T}}\boldsymbol{f}^{ij}_{l,\mathrm{vi}}}{\left\|\boldsymbol{f}^{ij}_{l,\mathrm{vi+ir}}(i+x,j+y)\right\|_2\left\|\boldsymbol{f}^{ij}_{l,\mathrm{vi}}\right\|_2} \tag{6}$$

经过$k\times k$次测量后, 可以得到相似度映射在每个位置(i,j), 这是$k\times k$的大小。由于丰富的纹理在可见模态中, 相似度值越大, 表示 $\boldsymbol{f}^{ij}_{l,\mathrm{vi+ir}}(i+x,j+y)$拥有更多的纹理信息。重构补丁时提供在位置$(i,j)$中向量的权值为

$$\boldsymbol{w}^{ij}_{l}=\mathrm{softmax}(\boldsymbol{s}^{ij}_{l}) \tag{7}$$

根据权值图, 可以将每个位置(i,j)处的特征向量重构为

$$\overline{\boldsymbol{f}}^{ij}_{l,\mathrm{vi+ir}}=\sum_{x,y}\boldsymbol{w}^{ij}_{l}(x,y)\boldsymbol{f}^{ij}_{l,\mathrm{vi+ir}}(i+x,j+y) \tag{8}$$

这意味着每个位置(i,j)处的重构特征向量是由$k\times k$补丁中的质心为(i,j)加权得到的, 而斑块中的纹理特征向量越多, 对重建的贡献就越大。

对于自适应红外特征细化分支, 将红外图像的梯度图中的特征表示为 $\boldsymbol{F}^{\mathrm{e}}_{l,\mathrm{ir}}$, 与可见光模态不同的是, 因为执行红外特征与可见特征的空间对齐, 使得 $\boldsymbol{f}^{ij}_{l,\mathrm{vi+ir}}$与 $\boldsymbol{f}^{ij}_{l,\mathrm{ir}}$之间的相似性越大, 则表示 $\boldsymbol{f}^{ij}_{l,\mathrm{vi+ir}}$与该位置的伪影相关性越高, 因此, 使用 1 减去相似度作为 $\boldsymbol{f}^{ij}_{l,\mathrm{vi+ir}}$的权重:

$$\tilde{\boldsymbol{f}}^{ij}_{l,\mathrm{vi+ir}}=\left(1-\frac{(\boldsymbol{f}^{ij}_{l,\mathrm{ir}})^{\mathrm{T}}\boldsymbol{f}^{ij}_{l,\mathrm{vi+ir}}}{\left\|\boldsymbol{f}^{ij}_{l,\mathrm{ir}}\right\|_2\left\|\boldsymbol{f}^{ij}_{l,\mathrm{vi+ir}}\right\|_2}\right)\boldsymbol{f}^{ij}_{l,\mathrm{vi+ir}}+\boldsymbol{f}^{ij}_{l,\mathrm{vi+ir}} \tag{9}$$

1.5　金字塔特征融合模块

为避免大规模降采样导致的信息缺失, 并引入大规模降采样导致的冗余信息, 引入了 PFFM 如图 7 所示。$\hat{\boldsymbol{F}}_{2,\mathrm{vi+ir}}$分别与 $\hat{\boldsymbol{F}}_{1,\mathrm{vi+ir}}$和 $\hat{\boldsymbol{F}}_{3,\mathrm{vi+ir}}$融合, 并进一步融合其融合特征以实现最终融合 $\boldsymbol{F}_{\mathrm{vi+ir}}$:

$$\boldsymbol{F}_{\mathrm{vi+ir}}=\mathrm{convs}(\mathrm{concat}(n(\boldsymbol{F}^{2,3}_{\mathrm{vi+ir}}),\boldsymbol{F}^{1,2}_{\mathrm{vi+ir}}),3\times 3) \tag{10}$$

式中: n表示最近邻上采样算子。

最终的$\boldsymbol{F}_{\mathrm{vi+ir}}$再通过 6 个残差块、1 个最近邻上采样算子和 3 个 3×3 的卷积层, 最终得到重建图像$\boldsymbol{x}_{\mathrm{vi+ir}}$。

1.6　损失函数

损失函数除式(4)中的对准损失外, 同时采用了像素损失和感知损失[27-28]。

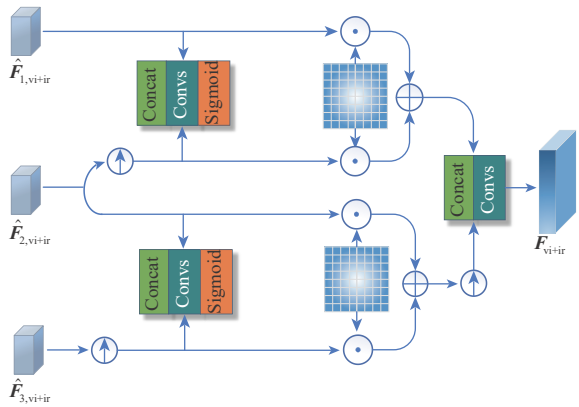


图 7　PFFM 架构

Fig. 7　Architecture of PFFM

像素损失计算如下:

$$l_{\mathrm{pixel}}=(1-\alpha)\|\boldsymbol{x}_{\mathrm{vi+ir}}-\boldsymbol{x}_{\mathrm{vi}}\|_1+\alpha\|\boldsymbol{x}_{\mathrm{vi+ir}}-\hat{\boldsymbol{x}}_{\mathrm{ir}}\|_1 \tag{11}$$

式中: $\hat{\boldsymbol{x}}_{\mathrm{ir}}$为与$\boldsymbol{x}_{\mathrm{vi}}$严格配准的红外图像; α为预定义的非负性参数。

感知损失是为获取图像的显著性信息, 计算如下:

$$\begin{aligned}l_{\mathrm{per}}=&(1-\beta)\|\mathrm{VGG16}(\boldsymbol{x}_{\mathrm{v+ir}})-\mathrm{VGG16}(\boldsymbol{x}_{\mathrm{vi}})\|_F+\\&\beta\|\mathrm{VGG16}(\boldsymbol{x}_{\mathrm{v+ii}})-\mathrm{VGG16}(\hat{\boldsymbol{x}}_{\mathrm{ir}})\|_F\end{aligned} \tag{12}$$

式中: VGG16()表示在 ImageNet 上预训练得到的 VGG16 网络[29-30]损失函数; β为预定义的非负性参数。最终整体损失计算如下:

$$l_{\mathrm{tol}}=l_{\mathrm{pixel}}+l_{\mathrm{per}}+\lambda l_{\mathrm{align}} \tag{13}$$

式中: λ为预定义的非负性参数。

2　实验验证

2.1　数据集构建

实验数据集通过操作民用无人机航拍采集的不同地点不同场景下多源舰船视频影像制作而成, 首先, 使用 OpenCV 工具包将视频分解成一系列的图像帧, 并按照设定的时间间隔对视频帧进行提取, 在分解视频帧生成图像帧的过程中, 根据每个图像帧的帧列序号进行命名, 有助于保证整个数据集的顺序和组织, 像素大小统一设定为 640 像素×512 像素, 最终共得到红外和可见光舰船图像 30 506 张, 分别存放于 2 个文件夹, 其中, 相同图像序号为同一时间同一场景采集, 但是因传感器的差异, 视场角大小不同, 故图像并非严格对齐, 使用 Labelimg 工具对样本进行标注, 保存为 VOC 数据格式。图 8 为所制作的多源舰船图像数据集架构。

数据集中舰船类别包括: 军舰、帆船、邮轮、散货船、渔船、游艇、集装箱船 7 种类型, 其具体数量如表 1 所示。

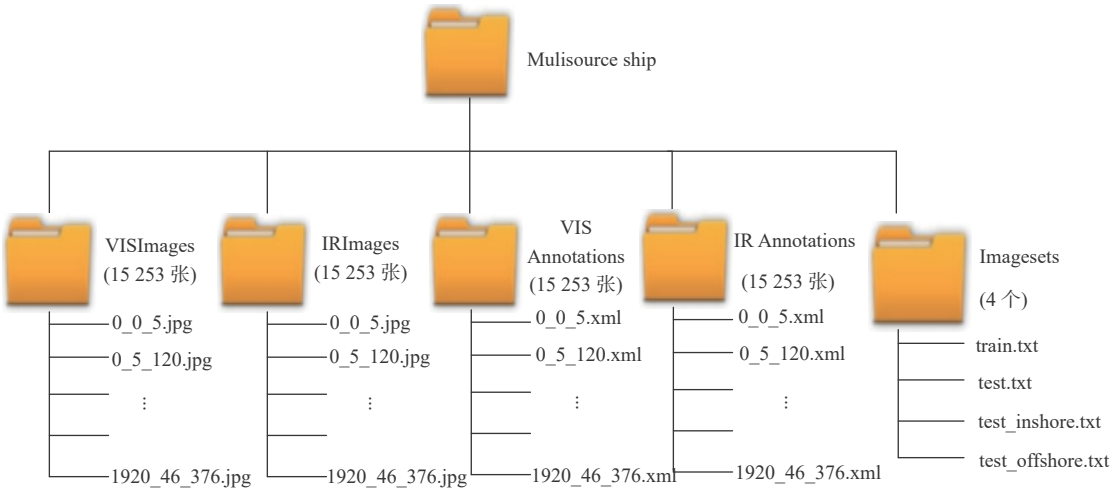


图 8 多源舰船数据集架构

Fig. 8 Architecture of multi-source ship dataset

表 1 舰船类别统计

Table 1 Ship category statistics

类型	数目
军舰	517
帆船	7 642
邮轮	436
散货船	678
渔船	4 251
游艇	21 578
集装箱船	863

2.2 实验环境及评价指标

实验操作系统为 Ubuntu20.04, 深度学习框架为 Pytorch1.9.0, 开发工具为 Visual Studio, 硬件环境 CPU 为 Inter(R) Core(R)i7, GPU 为 NVIDIA(R) GTX(R) 3080。

为评估融合后的图像质量, 共采用 8 种广泛使用的度量评价指标, 共包括: 交叉熵互信息(cross entropy, CE)、互信息(mutual information, MI)、边缘相似性度量 $Q_{AB/F}$、信息熵(entropy, EN)、Chen Blum 度量 Q_{CB}、Chen-Varshney 度量 Q_{CV}、标准偏差(standard deviation, SD)、结构相似性指数度量(structural similarity index measure, SSIM)。其中, 除 CE 和 Q_{CV}, 其他指标遵循数值越高, 性能越好。

对于识别结果的评价指标采用平均准确率(overall accuracy, OA)和混淆矩阵(confusion matrix, CM), OA 可以反映总体的识别精度, 其定义为

$$O_A = \frac{t_p + t_n}{t_p + t_n + f_p + f_n} \times 100\% \tag{14}$$

式中: t_p表示真正例; t_n表示真反例; f_p表示假正例; f_n表示假反例。CM 以矩阵形式将数据按真实类别和预测类别进行汇总, 每行之和表示真实类别样本数量, 列之和表述预测类别样本数量, 被正确分类的样本数量沿着矩阵的对角线分布。

2.3 消融实验

为验证 CM、FDAM、CMF、PFFM、LFR 不同模块及策略对融合图片质量的提升, 共设计了 5 组消融实验进行对比, 其中, MI 用来判定融合图像与源图像之间的相关性, $Q_{AB/F}$用来评定融合图像保留的边缘和细节数量, SSIM 从强度、对比度、结构 3 个方面衡量融合图像和源图像的相似度, EN 用来反映融合图像中的信息丰富程度, 结果如表 2 所示。由表可知, 各个模块和策略的增加使最终的融合图像质量有了明显的提升, 对不同模块产生的影响进行分析总结。

表 2 消融实验结果

Table 2 Ablation experimental results

模型	MI	$Q_{AB/F}$	SSIM	EN
Baseline	1.242	0.356	1.052	5.945
+CM	1.565	0.376	1.095	6.031
+FDAM	1.605	0.403	1.241	6.156
+CMF	1.689	0.467	1.264	6.347
+PFFM	1.897	0.478	1.279	6.779
+LFR	1.920	0.513	1.341	7.056

1) 采用 CM 策略是为提取 2 个源图像之间的关系, 从而计算 2 种模态的调制参数, 有利于 FDAM 的动态特征提取, 因此, 对 MI 影响较大, 使其增加了 0.323。

2) 采用 FDAM 模块学习动态核函数, 是为将红外特征与可见光特征在空间上对齐, 这对于 SSIM 的提升较为显著, 增加了 0.146。

3) 为从源图像获得更多有价值的互补信息, 采用 CMF 策略可以保存更为详细的纹理信息, 因此,

使$Q_{AB/F}$、EN 分别增加了 0.064、0.191

4) 采用 PFFM 模块可以利用金字塔结构融合多尺度特征, 从而保留更多有价值的信息, MI 和 EN 分别增加了 0.208 和 0.432。

5) LFR 策略由 2 个分支组成, 一个是纹理增强, 另一个是伪影去除, 其中, 任一分支都会影响到融合图像边缘是否模糊和存在伪影, 因此, 对$Q_{AB/F}$、SSIM 影响较明显, 分别增加了 0.035、0.062。

2.4　融合性能对比

为定量测量融合方法性能, 选取当前主流的 DDcGAN、FusionGAN、U2Fusion 等 8 种图像融合方法进行比较分析, 结果如表 3 所示, 表中加粗数值为最佳得分。由表可知, 本文方法在 CE、MI、$Q_{AB/F}$、Q_{CB}、SSIM、Q_{CV}这 6 项指标上优于其他方法, 分别达到了 0.454、1.920、0.513、0.561、1.341、516.6, SD、EN 这 2 个指标上次优于其他方法, 分别为 41.38 和 7.056, 与最优的 DDcGAN 相差不大, 从各项参数结果明显可以看出, 本文方法图像融合效果较好。

表 3　融合性能对比

Table 3　Comparison of fusion performance

方法	CE	MI	$Q_{AB/F}$	Q_{CB}	SSIM	Q_{CV}	SD	EN
DDcGAN	0.951	1.023	0.411	0.491	1.108	1 125.2	**48.08**	**7.500**
FusionGAN	2.041	1.344	0.248	0.433	1.170	1 149.1	23.78	6.892
NestFuse	0.910	1.916	0.497	0.527	1.304	573.2	41.34	7.024
PMG1	0.978	1.230	0.426	0.484	1.251	610.6	32.13	6.745
SDNet	1.076	0.997	0.434	0.517	1.227	902.4	25.30	6.868
U2Fusion	0.591	1.066	0.446	0.553	1.307	632.7	22.56	6.456
GTF	0.751	1.354	0.352	0.437	1.170	1 561.7	25.46	6.347
SwinFusion	0.472	1.895	0.481	0.554	1.225	553.1	38.13	6.779
本文方法	**0.454**	**1.920**	**0.513**	**0.561**	**1.341**	**516.6**	41.38	7.056

2.5　识别性能对比

为验证本文方法对目标识别性能的提升, 共采用了 ResNet、DenseNet、Inception、Xception、SqueezeNet 等 7 种当前主流的识别网络进行验证, 结果如表 4 所示。由表可知, 融合后的图像识别结果均优于单一模态, 并且在不同识别方法上均表现出良好的结果, 融合图像的平均识别准确度均到达了 0.923 以上, 相比可见光提升了至少 6%, 相比于红外至少提升了 8%, 提升效果较为明显, 反映出融合图像不仅有效保留了源图像信息, 并且充分合并了 2 种模态的目标特征, 获得了更为丰富的目标信息。

表 4　识别性能对比

Table 4　Comparison of recognition performance

图像	识别准确度/%						
	VGG16	ResNet50	Dense-Net201	Inception-v4	Squeeze-Net-v1.1	Xception	Alex-Net
融合图像	0.923	0.931	0.932	0.930	0.927	0.925	0.928
可见光图像	0.865	0.870	0.871	0.867	0.864	0.862	0.868
红外图像	0.848	0.853	0.844	0.846	0.841	0.837	0.839

不同模态下的混淆矩阵识别结果如图 9 所示。由图可知, 融合图像对目标的识别性能具有较明显的提升, 不论是整体识别准确度还是单一类别识别准确度均达到了 91% 以上, 反映出本文方法的鲁棒性较强。

2.6　可视化分析

在现实场景中存在许多干扰目标识别的因素, 如在无人机拍摄过程中, 受气候及光照条件

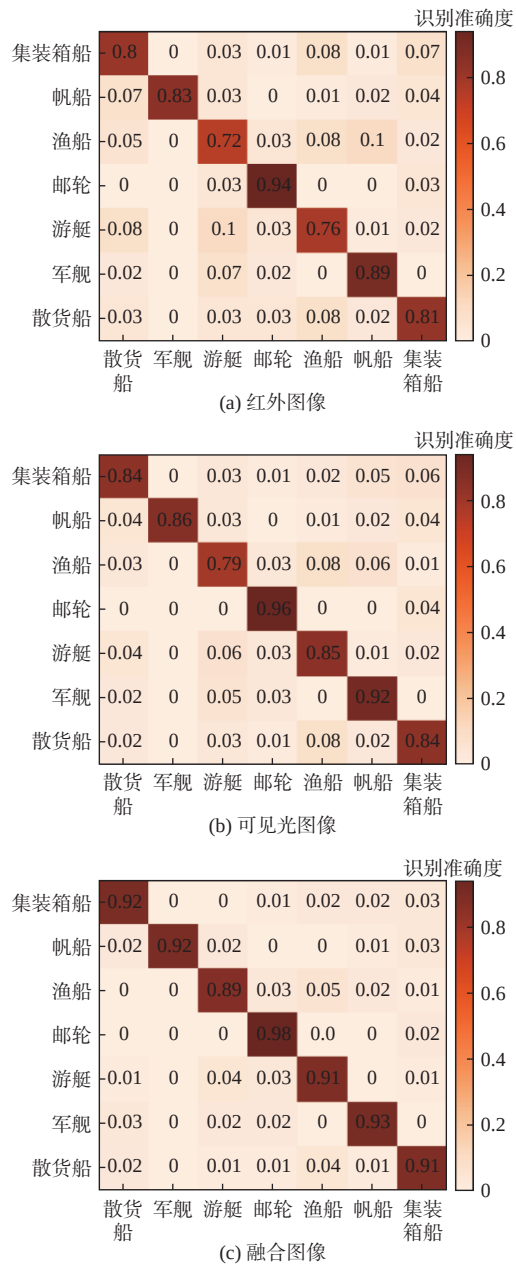
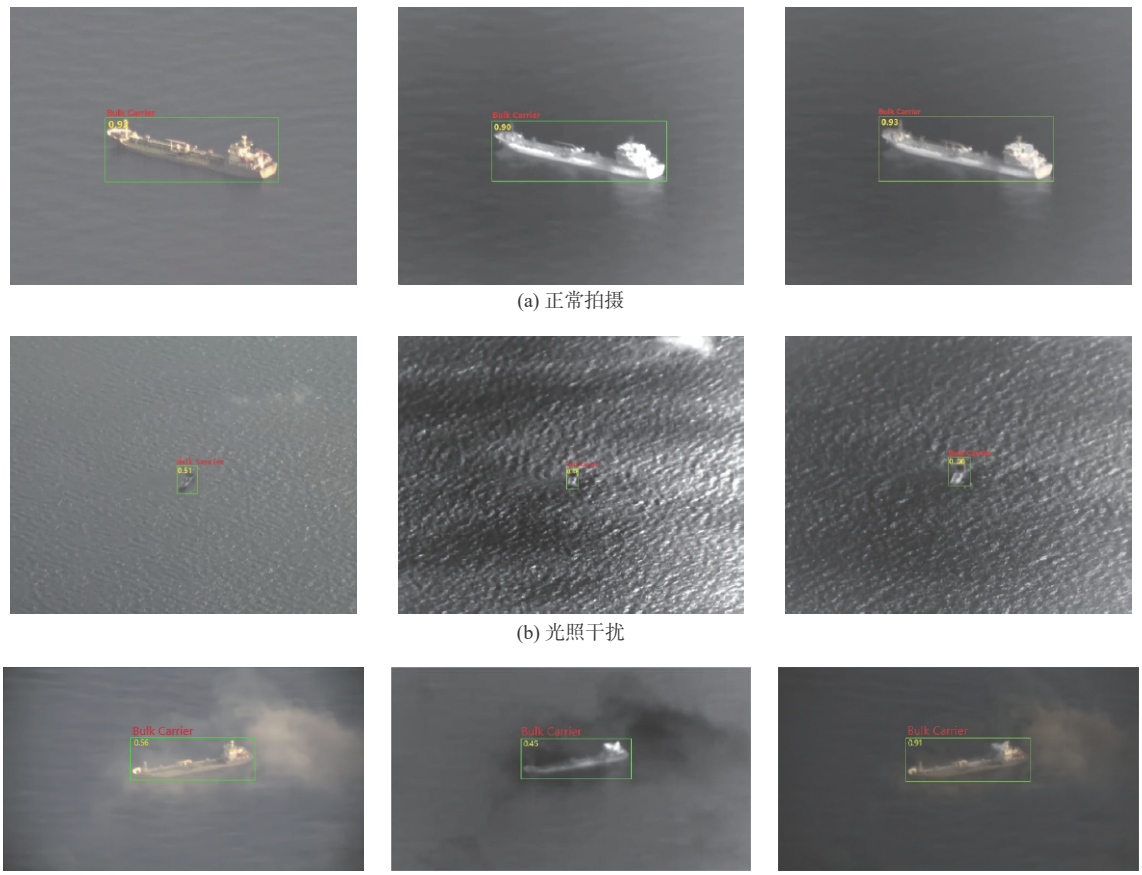


图 9　混淆矩阵结果

Fig. 9　Confusion matrix results

影响较大，此时单源识别方法的准确度难以得到保证，采用多源融合识别可以有效提升准确度，鲁棒性更强，为验证本文方法在真实场景中的图像融合质量及目标识别效果，选取实测数据中的部分样本进行测试，并通过可视化分析，用以评估本文方法在实际应用中的真实效果，结果如图 10 所示。图 10(a)为正常拍摄下，融合后的图像相比于源图像的纹理细节更加丰富，识别准确度得到提升；图 10(b)为海面反射光线较强且拍摄距离较远的情况，舰船目标较小且图像特征不明显，这对于单源图像识别而言，难度较大，经融合后，舰船轮廓明显清晰，海浪和光线造成的干扰明显降低，目标的图像信息更加丰富，识别准确度明显提升。图 10(c)为存在云雾干扰的情形，该情况在拍摄过程中极为普遍，云雾遮挡对于单源识别方法而言是极大的干扰因素，会对识别准确度产生很大影响，经图像融合后，可以发现云雾的遮挡明显降低，识别准确度大幅提升，有效验证了本文方法的有效性及实用性，对于处理现实问题具有重要意义。



(a) 正常拍摄

(b) 光照干扰

(c) 云雾干扰

图 10 可视化分析

Fig. 10 Visual analysis

3 结 论

本文提出一种多源舰船图像融合识别方法，主要解决了以下问题：

1）针对传感器采集到多源图像存在像素偏差问题，设计了一种端对端网络 MSIFR，根据 2 幅源图像之间的空间关系，共同学习调制参数和动态卷积核，使其模态在高级特征空间中对齐，实现了可见光和红外图像的配准和融合；

2）在网络模型中，设计了交叉调制策略、互补掩膜融合策略、局部特征细化策略以保留更多的图像信息，提升图像的融合效果；

3）构建了多源舰船图像数据集，用以验证图像的融合效果和对目标的识别性能。

经实验对比验证和可视化分析，结果表明，本文方法对多源舰船图像融合效果好、鲁棒性强，对目标的识别精度高、抗干扰能力强，能够较好地满足现实任务需要。下一步将探索研究更多有关无人机航拍多源图像的融合识别方法。

参考文献（References）

[1] 江波, 屈若锟, 李彦冬, 等. 基于深度学习的无人机航拍目标检测研究综述[J]. 航空学报, 2021, 42(4): 524519.

JIANG B, QU R K, LI Y D, et al. Object detection in UAV im-

agery based on deep learning: review[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(4): 524519(in Chinese).

[2] 苑玉彬, 吴一全, 赵朗月, 等. 基于深度学习的无人机航拍视频多目标检测与跟踪研究进展[J]. 航空学报, 2023, 44(18): 28334.
YUAN Y B, WU Y Q, ZHAO L Y, et al. Research progress of UAV aerial video multi-object detection and tracking based on deep learning[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(18): 28334(in Chinese).

[3] 何友, 熊伟, 刘俊, 等. 海上信息感知与融合研究进展及展望[J]. 火力与指挥控制, 2018, 43(6): 1-10.
HE Y, XIONG W, LIU J, et al. Review and prospect of research on maritime information perception and fusion[J]. Fire Control & Command Control, 2018, 43(6): 1-10(in Chinese).

[4] ZHANG P F, LI T R, WANG G Q, et al. Multi-source information fusion based on rough set theory: a review[J]. Information Fusion, 2021, 68: 85-117.

[5] 何友, 王国宏, 关欣, 等. 信息融合理论及应用[M]. 北京: 电子工业出版社, 2010.
HE Y, WANG G H, GUAN X, et al. Information fusion theory and applications [M]. Beijing: Publishing House of Electronics Industry, 2010.

[6] QIU X H, LI M, ZHANG L Q, et al. Deep convolutional feature fusion model for multispectral maritime imagery ship recognition[J]. Journal of Computer and Communications, 2020, 8(11): 23-43.

[7] LIU J M, CHEN H, WANG Y. Multi-source remote sensing image fusion for ship target detection and recognition[J]. Remote Sensing, 2021, 13(23): 4852.

[8] TANG L F, YUAN J T, ZHANG H, et al. PIAFusion: a progressive infrared and visible image fusion network based on illumination aware[J]. Information Fusion, 2022, 83-84: 79-92.

[9] LI Y S, WEI F Y, ZHANG Y J, et al. HS2P: hierarchical spectral and structure-preserving fusion network for multimodal remote sensing image cloud and shadow removal[J]. Information Fusion, 2023, 94: 215-228.

[10] ZHANG Y, XIAO Q L, DENG X Y, et al. A multi-source information fusion method for ship target recognition based on Bayesian inference and evidence theory[J]. Journal of Intelligent & Fuzzy, 2022, 42(3): 2331-2346.

[11] LI M J, DONG Y B, WANG X L. Pixel level image fusion based the wavelet transform[C]//Proceedings of the 6th International Congress on Image and Signal Processing. Piscataway: IEEE Press, 2014: 995-999.

[12] HAN X Y, LV T, SONG X Y, et al. An adaptive two-scale image fusion of visible and infrared images[J]. IEEE Access, 2019, 7: 56341-56352.

[13] YOU T T, TANG Y. Visual saliency detection based on adaptive fusion of color and texture features[C]//Proceedings of the 3rd IEEE International Conference on Computer and Communications. Piscataway: IEEE Press, 2018: 2034-2039.

[14] 杨曦, 张鑫, 郭浩远, 等. 基于不变特征的多源遥感图像舰船目标检测算法[J]. 电子学报, 2022, 50(4): 887-899.
YANG X, ZHANG X, GUO H Y, et al. Invariant features based ship detection model for multi-source remote sensing images[J]. Acta Electronica Sinica, 2022, 50(4): 887-899(in Chinese).

[15] WANG A L, JIANG J N, ZHANG H Y. Multi-sensor image decision level fusion detection algorithm based on D-S evidence theory[C]//Proceedings of the Fourth International Conference on Instrumentation and Measurement, Computer, Communication and Control. Piscataway: IEEE Press, 2014: 620-623.

[16] PAUL P P, GAVRILOVA M L, ALHAJJ R. Decision fusion for multimodal biometrics using social network analysis[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2014, 44(11): 1522-1533.

[17] ZHANG T W, ZHANG X L, LI J W, et al. SAR ship detection dataset (SSDD): official release and comprehensive data analysis[J]. Remote Sensing, 2021, 13(18): 3690.

[18] REN Y M, YANG J, ZHANG Q N, et al. Ship recognition based on Hu invariant moments and convolutional neural network for video surveillance[J]. Multimedia Tools and Applications, 2021, 80(1): 1343-1373.

[19] ER M J, ZHANG Y N, CHEN J, et al. Ship detection with deep learning: a survey[J]. Artificial Intelligence Review, 2023, 56(10): 11825-11865.

[20] SALI S M, MANISHA N L, KING G, et al. A review on object detection algorithms for ship detection[C]//Proceedings of the 7th International Conference on Advanced Computing and Communication Systems. Piscataway: IEEE Press, 2021: 1-5.

[21] MA J Y, JIANG J J, ZHOU H B, et al. Guided locality preserving feature matching for remote sensing image registration[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(8): 4435-4447.

[22] AYMERICH F, STASZEWSKI W J. Experimental study of impact-damage detection in composite laminates using a cross-modulation vibro-acoustic technique[J]. Structural Health Monitoring, 2010, 9(6): 541-553.

[23] KLUS S, NÜSKE F, HAMZI B. Kernel-based approximation of the Koopman generator and Schrödinger operator[J]. Entropy, 2020, 22(7): 722.

[24] QIN Z Q, ZHANG P Y, WU F, et al. FcaNet: frequency channel attention networks[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2022: 763-772.

[25] LI J X, GUO X B, LU G M, et al. DRPL: deep regression pair learning for multi-focus image fusion[J]. IEEE Transactions on Image Processing, 2020, 29: 4816-4831.

[26] LI J Y, ZHA S, CHEN C, et al. Attention guided global enhancement and local refinement network for semantic segmentation[J]. IEEE Transactions on Image Processing, 2022, 31: 3211-3223.

[27] JOHNSON J, ALAHI A, LI F F. Perceptual losses for real-time style transfer and super-resolution[C]//Proceedings of the Computer Vision – ECCV. Berlin: Springer, 2016: 694-711.

[28] SONG H H, XU W J, LIU D, et al. Multi-stage feature fusion network for video super-resolution[J]. IEEE Transactions on Image Processing, 2021, 30: 2923-2934.

[29] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2015-04-10) [2024-03-01]. https://arxiv.org/abs/1409.1556.

[30] RUSSAKOVSKY O, DENG J, SU H, et al. ImageNet large scale visual recognition challenge[J]. International Journal of Computer Vision, 2015, 115(3): 211-252.

Ship target recognition method based on multi-source image fusion for unmanned aerial vehicle aerial photography

JIANG Jie，YAN Wenjun，LIU Kai，ZHANG Limin*

(School of Aviation Operations Service，Naval Aviation University，Yantai 264001，China)

Abstract：A multi-source ship image fusion recognition method is proposed for unmanned aerial vehicle aerial photography of multi-source ship images. In the face of many interferences in real scenes, pixel level fusion is adopted to fuse infrared and visible light ship images, and then perform target recognition. It can improve the algorithm's interpretability and lessen the network's reliance on samples as compared to feature level recognition techniques. This article focuses on solving the pixel offset caused by different sensor parameters. To eliminate the distortion and artifacts that traditional picture registration can readily cause, image registration has been turned into end-to-end feature alignment. A multi-source ship image fusion recognition network is proposed, which consists of a cross modulation feature extraction module, a feature dynamic alignment module, a multi granularity feature refinement module, and a pyramid feature fusion module. It can fully integrate the features and texture details of different modal images, effectively improving the recognition performance of ship targets. The approach described in this study has demonstrated great interpretability for multi-source images, good fusion performance, and high accuracy and robustness in recognizing ship targets through experimental verification.

Keywords：unmanned aerial vehicle；multi source ship images；fusion recognition；pixel level fusion；pixel offset

Received：2024-05-07；**Accepted**：2024-06-21；**Published Online**：2024-08-15 13：53

URL：link.cnki.net/urlid/11.2625.V.20240815.0751.001

Foundation items：National Natural Science Foundation of China (62371465); Shandong Province Youth Innovation Team(2022kj084); Shandong Provincial Natural Science Foundation(ZR2020QF010)

* **Corresponding author**. E-mail: iamzlm@hotmail.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0270

翼身融合组合动力运载器耦合抑制方法

熊知越[1]，张曙光[1,*]，丁嘉元[2]，廖雨舟[1]
(1. 北京航空航天大学 交通科学与工程学院，北京 100191；2. 中国运载火箭技术研究院，北京 100076)

摘 要：采用翼身融合、飞发一体化设计技术的水平起降组合动力运载器，具有可快速部署、高效运行、重复执行空天往返任务的应用前景。但是，这类布局飞行器的飞行动力学受气动非线性、气动-动力-结构弹性耦合等影响，呈现纵-横向运动耦合、飞行姿态-轨迹耦合等特性，给控制设计带来挑战。对此，从建立描述翼身融合组合动力运载器的多种耦合特性的飞行动力学模型着手，研究其飞行动力学特性，并提出多回路控制方案，引入动态逆回路进行耦合抑制，结合常规PID控制，易于设计实现。通过大量拉偏仿真表明:所提控制方案在抑制耦合的基础上能够实现良好的姿态和轨迹跟踪性能，在模型、环境、信号测量等不确定性的影响下具有良好的鲁棒性和抗干扰性。

关 键 词：高超声速运载器；翼身融合；机身弹性；动力学耦合；多回路控制；轨迹跟踪
中图分类号：V448.2
文献标志码：A **文章编号**：1001-5965(2026)06-1965-17

翼身融合高超声速运载器具有较高的升阻比、较强的机动能力，而且往往可进行跑道降落，从而具备空天往返和可重复使用能力，通常归为空天飞行器范围[1]。其能够突破传统消耗性火箭和航天飞机的局限，实现快速部署、从大气层到低轨空间自由高速穿梭[2]，具有广阔应用前景，因而逐渐成为全球航空航天研究重点[3]。高超声速运载器的先驱者有美国X-15验证机[4]、美国航天飞机[5]、前苏联暴风雪号航天飞机等。这些运载器方案采用火箭助推，一般为垂直或携带起飞，油耗较高、可重复使用率低，难以实现临近空间长时间巡航。20世纪80年代，随着以吸气式动力为代表的高超声速关键技术取得突破性进展，天地往返可重复使用运载器的研究热潮重新掀起[6]。如德国最早提出采用两级入轨(two-stage to orbit, TSTO)、冲压/火箭组合动力的桑格尔(Sänger)方案[6]，英国提出了采用单级入轨(single-stage to orbit, SSTO)、RB545组合发动机的水平起降(horizontal take-off and landing, HOTOL)运载器方案[7]等。其中，TSTO方案能够缓解高、低速设计矛盾，有助于进一步节能减重，技术可实现性更强[3]。然而受限于吸气式组合动力的技术成熟度，这些方案多止步于初步设计阶段。20世纪90年代以后，组合动力持续受到国内外重视，基于火箭、冲压、涡轮及预冷技术进行有机组合，形成了火箭基组合循环(rocket based combined cycle, RBCC)[8]、涡轮基组合循环(turbine based combined cycle, TBCC)[9]、协同吸气式火箭发动机(synergistic air-breathing rocket engine, SABRE)[10]、预冷空气涡轮火箭(precooled air turbo rocket, PATR)[11]等组合动力方式。各国也纷纷开展基于组合动力的运载器方案研究，如英国基于SABRE发动机的云霄塔(Skylon)方案[12]、美国空军研究实验室(Air Force Research Lab, AFRL)与REL(Reaction Engines, Ltd.)公司合作提出的基于SABRE的TSTO方案[13]、欧洲航天局基于预冷涡轮冲压动力的LAPCAT-II计划[14]及欧洲多国近年来提出的基于双模冲压和火箭动力的平流层高超声速民机Horizon 2020 STRATOFLY计划[15]。超燃冲压方式可提升冲压动力的工作马赫数上限，针对这一推进方式，美国X-43A[16]和X-51A[17]验证机分别创造了试飞速度和时长纪录。

收稿日期：2024-04-29；录用日期：2024-06-11；网络出版时间：2024-06-20 13:32
网络出版地址：link.cnki.net/urlid/11.2625.V.20240619.1628.003
*通信作者. E-mail：gnahz@buaa.edu.cn

引用格式：熊知越，张曙光，丁嘉元，等. 翼身融合组合动力运载器耦合抑制方法[J]. 北京航空航天大学学报，2026，52(6)：1965-1981. XIONG Z Y，ZHANG S G，DING J Y，et al. Coupling suppression method of blended-wing-body combined-cycle-power vehicle[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1965-1981 (in Chinese).

采用翼身融合布局的组合动力运载器(以下简称运载器)在具有廉价、快速、可靠、便捷等众多优势的同时[3],也给控制带来了新的问题和挑战。大长细比、翼身融合的设计导致运载器存在明显的气动强非线性[18]、纵-横向气动耦合和静不稳特性,也易产生多重不确定性[19];加之追求轻质化的设计导致机身较低的刚度特征,从而产生结构弹性问题[20]。不同于常规飞行器的气动和动力相对解耦的关系,组合动力运载器的飞发一体化设计有助于减轻结构质量、提升大气层内续航能力和任务灵活性,但也导致明显的气动-动力-弹性耦合(飞-发-弹性耦合),以及区别于传统火箭的飞行姿态-轨迹耦合特性[18],需通过制导-控制一体化设计予以重视。另外,不同动力和级间控制模式切换问题也不能忽视。

面对翼身融合组合动力运载器在由低速、亚跨声速到高超声速的宽域任务下的诸多控制问题,需要深入研究被控对象的精确模型。目前国内外针对翼锥体和乘波体这 2 种构型运载器的飞行动力学建模研究已有进展。由于运载器普遍采用大长细比的设计,许多研究为降低控制设计难度,主要针对纵向特性进行建模和控制设计,而对横侧向的考虑相对初步[19]。最具代表性的运载器模型包括美国国家航空航天局(National Aeronautics and Space Administration, NASA)兰利实验室基于其风洞数据建立的翼锥体模型[21]、美国加州大学洛杉矶分校的 Mirmirani 等基于计算流体动力学(computational fluid dynamics, CFD)技术建立的考虑飞-发耦合特性的 CSULA-GHV 模型[22]及 AFRL 基于 X-43A 试验数据和机理分析建立的第一原理模型(first principal model, FPM)[23]。其中,FPM 考虑了飞-发-弹性耦合特性,成为近年来运载器飞行控制设计的研究热点之一[19]。基于 FPM,曹林分析其动力学耦合特性,设计了高阶滑模和干扰解耦鲁棒控制器[19];安昊设计了基于反步法的抗干扰、抗饱和预设性能控制器[24]。基于 CFD 数据建模方式,Li 等为高超声速飞行器纵向弹性体模型设计了考虑舵面容错的 L1 自适应控制律[25]。

在控制设计方面,传统运载火箭和火箭动力运载器的轨迹和姿态动力学一般解耦、控制与制导律的设计可分别独立进行。例如,长征运载火箭[26]和美国航天飞机[5]均采用姿态回路 PID 控制、轨迹回路标称轨迹和预测校正制导的控制方案。而对于具有较大升力面的翼身融合组合动力运载器,其姿态-轨迹回路的高度耦合特性对解耦控制提出了额外需求;而不确定性、飞-发-弹性耦合等特性也对控制鲁棒性提出了较高要求。另外,运载器在临近空间巡航和转场机动方面的优势,表明了需要关注其三维制导控制问题。目前可用于运载器的姿态控制技术包括反馈线性化[27-29]、反步法[30-31]、滑模控制[32-33]、智能控制[34-36]等,可用于轨迹制导的技术包括标称轨迹跟踪[37]、在线轨迹规划[38-39]、预测校正制导[40]等。这类研究或仅在纵向平面内对运载器进行控制设计,或未能考虑其飞-发-弹性耦合特征,或仅考虑姿态运动回路控制设计。

综上所述,具有优异气动性能的翼身融合运载器显现出较强非线性、纵-横耦合特性,与组合动力等先进推进方式结合,更易凸显飞-发-弹性耦合等问题,导致基于常规方法的姿态-轨迹解耦飞行控制性能下降,甚至无法实现预期控制目标。本文针对这些耦合、非线性问题展开研究,从建立描述翼身融合组合动力运载器的多种耦合特性的飞行动力学模型入手,设计适应于姿态-轨迹耦合的多层耦合抑制控制器。

1 运载器飞行动力学建模

1.1 目标运载器总体特征

本文参考桑格尔方案[6]构建目标翼身融合组合动力运载器模型,采用两级入轨,其几何外形如图 1 所示。

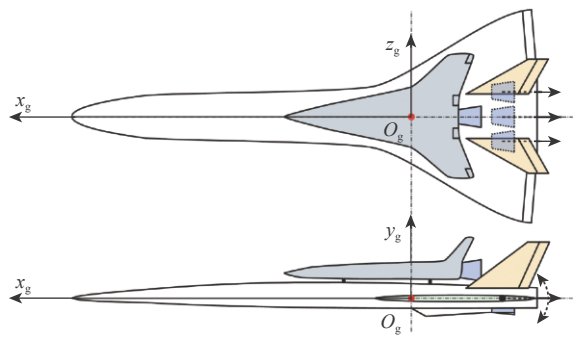


图 1 目标运载器几何外形

Fig. 1 Geometric shape of target vehicle

其中,二子级采用火箭动力,且工作高度在 50 km 以上,气动效应可忽略,控制方法与传统火箭类似,因此,本文不作重点研究。一子级采用尖前缘、大长细比、翼身融合布局设计,能够适应从亚跨声速、低超声速到高超声速的宽速域任务,其与二子级的组合体是本文的重点研究对象。参考 SABRE 组合动力特性建立发动机模型[41],机身内部后方整合有 3 台组合发动机,喷口位于机尾下表面,具备吸气式和火箭 2 种模式,前者工作范围为:海拔高度为 0～35 km、马赫数为 0～5.5,后者无限制,且具有推力矢量控制能力。组合体具有 2 片升降副翼、2 片方向舵。目标运载器组合体构型的部分总体参数如表 1 所示。

表 1　目标运载器总体特征

Table 1　Overall characteristics of target vehicle

满油总质量/kg	燃油质量/kg	机身长度/m	机翼翼展/m^2	吸气模式总推力(台架推力)/kN	火箭模式总推力(推力恒定)/kN
18 750	8 750	28.5	4.5	125	125

该外形具有较强的飞-发耦合特性。这一方面由于机体前段下表面可用作进气道的压缩面, 进气流量与飞行迎角密切相关, 进而影响推力特性; 另一方面由于后段下表面的气流流速和压强受发动机喷流作用而增大, 进而影响气动特性。

1.2　气动和推进特性

本文依据航天领域常用的第 3 类坐标系[42], 定义箭体坐标系 $O_g x_g y_g z_g$(体轴系)、当地铅垂系 $O_d x_d y_d z_d$(地面系)和地心旋转系 $O_e x_e y_e z_e$。

运载器的气动和推进特性基于 CFD 技术获取。其中, 体轴系下气动特性和推进特性表达为

$$C_i = C_{i.\mathrm{bas}}(\alpha, Ma, \beta) + C_{i.\mathrm{inc}}(\alpha, Ma, \beta, \delta_{e1}, \delta_{e2}, \delta_{r1}, \delta_{r2}, \Phi) \quad i \in \{x, y, z, m_x, m_y, m_z\} \tag{1}$$

$$T = 3T_{\mathrm{single}}(\alpha, Ma, h, \Phi) \tag{2}$$

$$e = 3e_{\mathrm{single}}(\alpha, Ma, h, \Phi) \tag{3}$$

式中: C_i 为体轴系气动力和力矩系数; $C_{i.\mathrm{bas}}$ 为无舵偏基准态的气动系数; $C_{i.\mathrm{inc}}$ 为舵偏引起的气动系数增量; m_x、m_y、m_z 分别为 x、y、z 方向的力矩系数; T 和 T_{single} 分别为推力和单发推力; e 和 e_{single} 分别为耗油率和单发耗油率; α 为迎角, Ma 为马赫数; β 为侧滑角; Φ 为油门开度; h 为海拔高度; δ_{e1}、δ_{e2} 分别为右、左侧升降副翼偏角; δ_{r1}、δ_{r2} 分别为右、左侧方向舵偏角。

目标运载器组合体构型具有区别于传统航空航天器的气动和推进基本特性, 主要体现在:

1) 目标运载器纵向静不稳定(尤其对于高超声速), 且具有较强的纵向气动非线性。以基准态俯仰力矩系数 $C_{mz.\mathrm{bas}}$ 为例, 如图 2 所示, 随着 α 增大, $C_{mz.\mathrm{bas}}$ 在亚声速时变化很小, 跨声速时先减后增, 高超声速时增大, 且斜率不均匀。

2) 目标运载器航向静不稳定。如图 3 所示, 滚转力矩系数 $C_{mx.\mathrm{bas}}$ 相对于侧滑角 β 的斜率为负, 具有一定的横向静稳定性, 但高超声速时稳定性较弱; 在低速和高超声速时, 偏航力矩系数 $C_{my.\mathrm{bas}}$ 相对于侧滑角 β 的斜率为正, 航向静不稳(尤其对于高超声速), 因此, 横航向总体呈现静不稳定性。

3) 目标运载器的纵-横气动耦合特性不可忽视, 且具有显著的纵-横操纵耦合特性。对于气动耦合, 以基准态俯仰力矩系数 $C_{mz.\mathrm{bas}}$、滚转力矩系数 $C_{mx.\mathrm{bas}}$ 为例, 如图 4 所示, 在较大迎角(15°)、亚跨声速时, β 对 $C_{mz.\mathrm{bas}}$ 的影响较大; 而在较大侧滑角(5°)、宽速域下, α 对 $C_{mx.\mathrm{bas}}$ 的影响较大, 且较大 α 能适当提升横向静稳定性。对于操纵耦合, 以滚转操纵 δ_x 引起的法向气动力系数增量 $C_{y.\mathrm{inc}}$、偏航力矩系数增量 $C_{my.\mathrm{inc}}$ 为例, 如图 5 所示, 在低速和高超声速时, 纵-横操纵耦合较大; 而在亚跨声速时, 横-航向操纵耦合较大。其中, 运载器的滚转操纵由升降副翼差动实现:

$$\delta_x = (\delta_{e1} - \delta_{e2})/2 \tag{4}$$

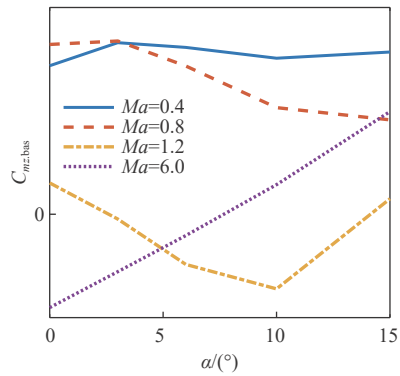


图 2　基准态俯仰力矩系数与迎角的关系

Fig. 2　Relationship between pitching moment coefficient and angle of attack

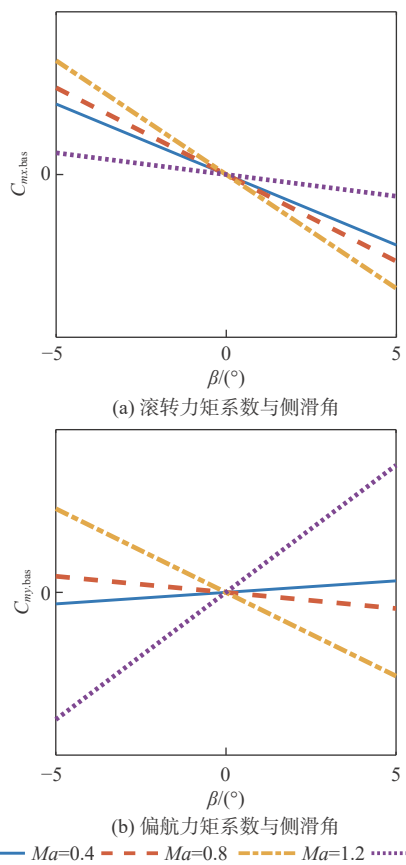
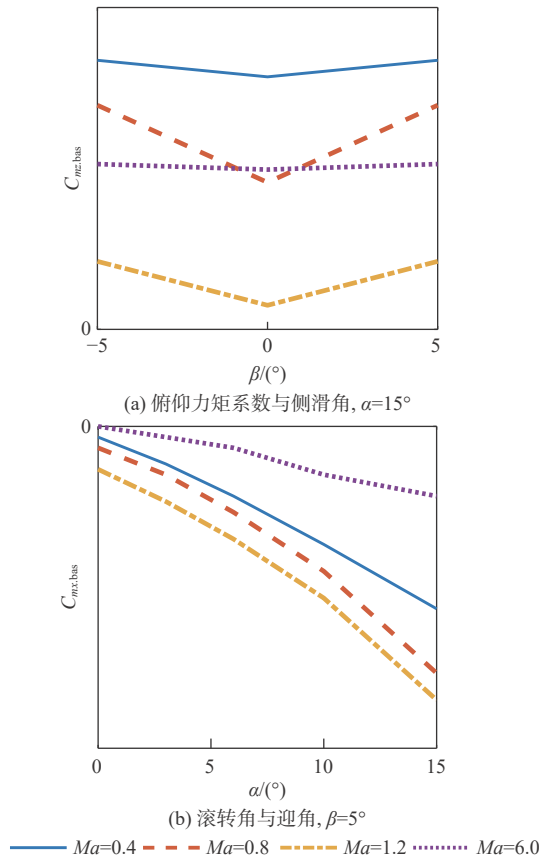


图 3　基准态滚转、偏航力矩系数与侧滑角的关系

Fig. 3　Relationship between rolling, yawing moment coefficient and sideslip angle

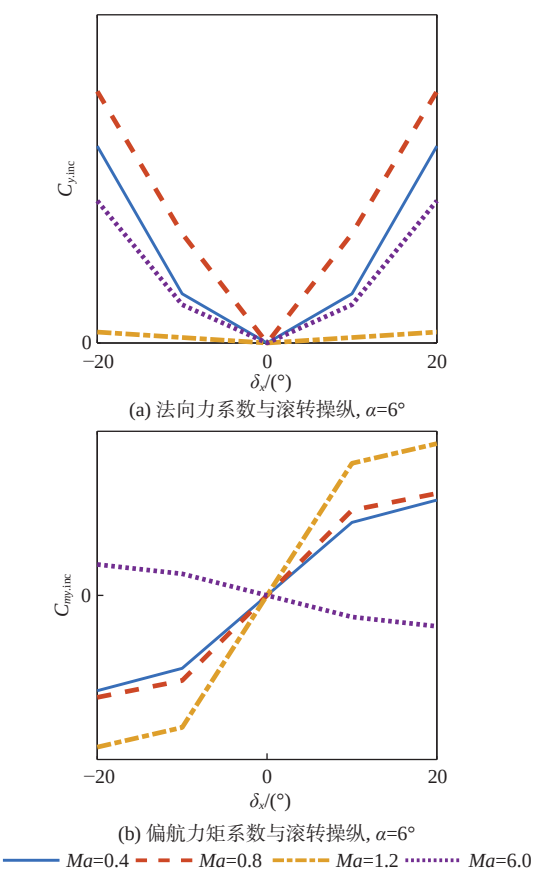


(a) 俯仰力矩系数与侧滑角, α=15°

(b) 滚转角与迎角, β=5°

图 4 纵-横气动耦合特性

Fig. 4 Longitudinal-lateral cross coupling characteristics



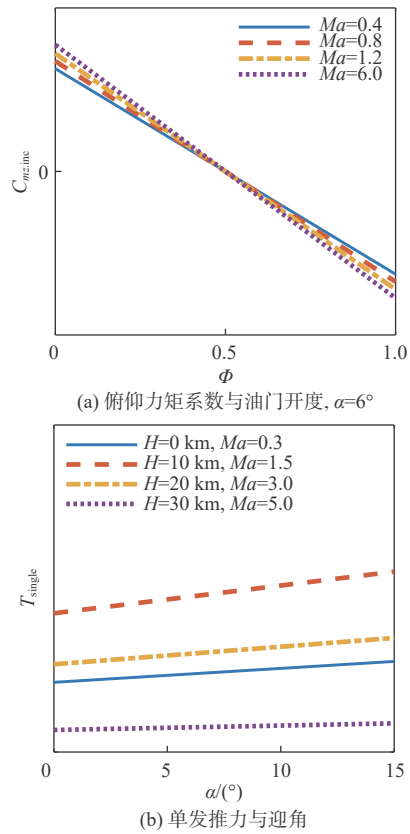
(a) 法向力系数与滚转操纵, α=6°

(b) 偏航力矩系数与滚转操纵, α=6°

图 5 纵-横操纵耦合特性

Fig. 5 Longitudinal-lateral control coupling characteristics

4) 目标运载器具有明显的飞-发耦合特性。以俯仰力矩系数和吸气模式单发推力为例, 如图 6 所示, 随着油门开度Φ的增大, $C_{mz.inc}$减小, 这源于机体后段下表面的气动载荷受发动机喷流作用而增大, 进而减阻增升, 产生附加低头力矩; 而随着迎角α增大, 发动机进气流量增大, 推力随之增大。



(a) 俯仰力矩系数与油门开度, α=6°

(b) 单发推力与迎角

图 6 飞-发耦合特性

Fig. 6 Aero-engine coupling characteristics

综上, 目标运载器具有以下需要在飞行控制设计中重点考虑的动力学特性, 主要包括较强的气动非线性、高超声速下尤为显著的纵-航向静不稳定性、明显的纵-横气动和操纵耦合特性、明显的飞-发耦合特性。

1.3 考虑机身弹性的飞行动力学方程

1.3.1 机身弹性影响

运载器高速飞行引起的大动压使机体产生弹性变形, 影响气动和推力特性, 引起弹性和刚性动力学耦合。基于以下几点简化考虑, 可视运载器为准二维构型, 即主要考虑机身纵向弹性变形影响而忽略横向变形: ① 机身为大长细比, 除首尾以外, 质量和刚度等物理特性沿横侧向分布基本均匀; ②机身前缘楔角较小, 斜激波波角也较小[43], 表明其气动特性沿横侧向分布基本均匀; ③运载器在飞行中纵向姿态变化比横航向剧烈地多, 且侧滑角基本保持为 0。

机身弹性对运动的影响主要分 2 部分[43]: ①静

弹性变形, 以全机质心为参考点, 可认为质心处无变形; 气动载荷作用于机身各处, 若机身刚度不够大, 则前、后段会分别产生不可忽略的弹性变形, 具体来说, 若前段合力向上则上翘、迎角增大, 后段合力向上则迎角减小, 进而影响各段气动和推力特性, 最终影响全机运动; ② 刚-弹运动耦合, 在静弹性变形的基础上, 机身弹性变形运动与刚性运动存在耦合, 其耦合动力学可通过含弹性耦合参数的方程描述[43], 这种耦合随机身刚性的减小而增大。

1.3.2　目标运载器机身弹性数据

基于目标运载器准二维构型假设, 依据文献 [44], 可将机身等效为图 7 中两段固定于质心处的悬臂梁, 只考虑机体沿 y_g 方向的变形, 且质心位置不受变形影响, 图中, τ_f 和 τ_a 分别为前、后梁的变形转角, $\Delta M_{z.\mathrm{inc}}$ 为升降副翼由于机身附加俯仰力矩而产生的附加俯仰操纵力矩, $\Delta M_{z.\mathrm{bas}}$ 为机身变形至实线位置时产生的附加俯仰力矩。表 2 列出了各段梁的弯曲刚度和满油时沿 x_g 轴的线密度, 参数取值依据运载器质量进行了缩比。

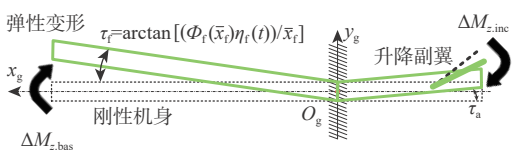


图 7　机身等效悬臂梁及其弹性变形示意

Fig. 7　Elasticity deformation of fuselage equivalent beams

表 2　机身弹性数据

Table 2　Elastic data of fuselage

弯曲刚度 E_1/(N·m^2)	机体前段等效梁线密度 $\overline{m}_f$/(kg·m^{-1})	机体后段等效梁线密度 $\overline{m}_a$/(kg·m^{-1})	机体前段等效梁长度$\bar{x}_f$/m	机体后段等效梁长度$\bar{x}_a$/m	弹性振动阻尼比ζ
1.121×10^9	301.2	1 437.1	19.55	8.95	0.05

悬臂梁具有多阶弹性振动模态, 不同模态对机身刚性运动的影响程度也不同。其模态固有频率可根据悬臂梁振动偏微分方程 (5), 结合梁挠度解析式 (6)、模态固有频率定义式 (7)、边界条件式 (8) 进行求解[43,45]:

$$E_1\frac{\partial^4 w_i(x,t)}{\partial x^4}+\overline{m}_i\frac{\partial^2 w_i(x,t)}{\partial t^2}=0 \qquad i\in\{\mathrm{f,a}\} \tag{5}$$

$$w_i(x,t)=\sum_{k=1}^{\infty}\Phi_{i,k}(x)\eta_{i,k}(t) \qquad i\in\{\mathrm{f,a}\} \tag{6}$$

$$\frac{\mathrm{d}^2\eta_{i,k}(t)}{\mathrm{d}t^2}+\omega_{i,k}^2\eta_{i,k}(t)=0 \qquad i\in\{\mathrm{f,a}\} \tag{7}$$

$$\begin{cases}\Phi(0)=0\\ \Phi'(0)=0\\ \Phi''(\bar{x}_i)=0\\ \Phi'''(\bar{x}_i)=0\end{cases} \qquad i\in\{\mathrm{f,a}\} \tag{8}$$

式中: $w_i(x,t)$ 为梁在 x 位置、t 时刻的沿 y_g 方向的挠度; f 和 a 分别表示前段梁和后端梁; $\Phi_{i,k}(x)$、$\eta_{i,k}(t)$ 和 $\omega_{i,k}$ 分别为梁的 k 阶振型函数、弹性广义坐标和模态固有频率。

由表 2 数据, 依据式 (5)～式 (8) 可算出满油时运载器等效梁的各阶模态固有频率和振型函数。表 3 列出其中前 4 阶固有频率, 图 8 给出飞行高度为 30 km、马赫数为 5 时的一阶振型函数, 横轴方向与图 7 保持一致。由于二阶以上的弹性模态频率与目标运载器刚性运动模态频率相去甚远, 可与后者自然解耦, 因此, 本文仅考虑一阶弹性模态。

表 3　机身等效梁的固有频率

Table 3　Natural frequencies of equivalent beams

阶数k	前段梁ω_f/ (rad·s^{-1})	后段梁ω_a/(rad·s^{-1})
1	8.87	19.39
2	55.62	121.49
3	155.73	340.17
4	305.17	666.58

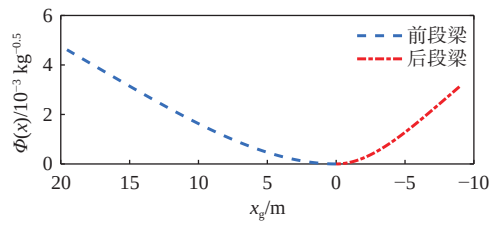


图 8　机身等效梁的一阶振型函数

Fig. 8　The first order vibration mode of equivalent beams

1.3.3　弹性飞行动力学建模

运载器的弹性变形会给运动引入沿 y_g 方向的弹性自由度, 一般以变形前的刚性体轴系 $O_g x_g y_g z_g$ 为基准, 使用前、后段梁的弹性广义坐标 η_f 和 η_a 来度量[43]。机身结构弹性对飞行的影响表现在弹性引起的附加惯性力和气动力; 机身两段等效梁的变形, 受到等效的集中气动力、推力等影响。因此, 考虑圆地球自转, 体轴系下飞行动力学和弹性运动方程为[46]

$$\begin{aligned}\begin{bmatrix}\dot{u}\\ \dot{v}\\ \dot{w}\end{bmatrix}=&-\left(\boldsymbol{L}_{\mathrm{gd}}\frac{1}{R}\begin{bmatrix}V_{z\mathrm{d}}\\ V_{x\mathrm{d}}\tan\phi_{\mathrm{c}}\\ -V_{x\mathrm{d}}\end{bmatrix}+\begin{bmatrix}p\\ q\\ r\end{bmatrix}\right)\times\begin{bmatrix}u\\ v\\ w\end{bmatrix}+\\ &\frac{G_{\mathrm{M}}}{R^2}\begin{bmatrix}-\sin\theta\\ -\cos\theta\cos\phi\\ \cos\theta\sin\phi\end{bmatrix}+\boldsymbol{L}_{\mathrm{gd}}\omega_{\mathrm{E}}^2R\begin{bmatrix}0\\ \cos^2\phi_{\mathrm{c}}\\ \sin\phi_{\mathrm{c}}\cos\phi_{\mathrm{c}}\end{bmatrix}+\\ &\frac{1}{m}\begin{bmatrix}F_x\\ F_y\\ F_z\end{bmatrix}-2\boldsymbol{L}_{\mathrm{gd}}\omega_{\mathrm{E}}\begin{bmatrix}0\\ \sin\phi_{\mathrm{c}}\\ -\cos\phi_{\mathrm{c}}\end{bmatrix}\times\begin{bmatrix}u\\ v\\ w\end{bmatrix}+\\ &\frac{1}{m}\begin{bmatrix}\dot{r}(\lambda_{\mathrm{f}}\eta_{\mathrm{f}}+\lambda_{\mathrm{a}}\eta_{\mathrm{a}})+2r(\lambda_{\mathrm{f}}\dot{\eta}_{\mathrm{f}}+\lambda_{\mathrm{a}}\dot{\eta}_{\mathrm{a}})\\ r^2(\lambda_{\mathrm{f}}\eta_{\mathrm{f}}+\lambda_{\mathrm{a}}\eta_{\mathrm{a}})-(\lambda_{\mathrm{f}}\ddot{\eta}_{\mathrm{f}}+\lambda_{\mathrm{a}}\ddot{\eta}_{\mathrm{a}})\\ 0\end{bmatrix}\end{aligned} \tag{9}$$

$$\begin{cases}\dot{p}=\left\{I_{yy}M_x+I_{xy}M_y-I_{xy}\left(I_{xx}+I_{yy}-I_{zz}\right)pr-\right.\\ \qquad\left.\left[I_{yy}\left(I_{zz}-I_{yy}\right)-I_{xy}^2\right]qr\right\}/\left(I_{xx}I_{yy}-I_{xy}^2\right)\\ \dot{q}=\left\{I_{xx}M_y+I_{xy}M_x-I_{xy}\left(I_{zz}-I_{xx}-I_{yy}\right)qr-\right.\\ \qquad\left.\left[I_{xx}\left(I_{xx}-I_{zz}\right)+I_{xy}^2\right]pr\right\}/\left(I_{xx}I_{yy}-I_{xy}^2\right)\\ \dot{r}=\left[M_z+I_{xy}\left(p^2-q^2\right)+\left(I_{xx}-I_{yy}\right)pq+\right.\\ \qquad(\dot{u}-rv)(\lambda_{\mathrm{f}}\eta_{\mathrm{f}}+\lambda_{\mathrm{a}}\eta_{\mathrm{a}})-2r(\eta_{\mathrm{f}}\dot{\eta}_{\mathrm{f}}+\eta_{\mathrm{a}}\dot{\eta}_{\mathrm{a}})-\\ \qquad\left.(\psi_{\mathrm{f}}\ddot{\eta}_{\mathrm{f}}+\psi_{\mathrm{a}}\ddot{\eta}_{\mathrm{a}})\right]/\left(I_{zz}+\eta_{\mathrm{f}}^2+\eta_{\mathrm{a}}^2\right)\end{cases}\tag{10}$$

$$\begin{cases}\ddot{\eta}_{\mathrm{f}}=N_{\mathrm{f}}+(\dot{v}+ru)\lambda_{\mathrm{f}}+\dot{r}\psi_{\mathrm{f}}-\\ \qquad2\zeta\omega_{\mathrm{f}}\dot{\eta}_{\mathrm{f}}-\left(\omega_{\mathrm{f}}^2-r^2\right)\eta_{\mathrm{f}}\\ \ddot{\eta}_{\mathrm{a}}=N_{\mathrm{a}}+(\dot{v}+ru)\lambda_{\mathrm{a}}+\dot{r}\psi_{\mathrm{a}}-\\ \qquad2\zeta\omega_{\mathrm{a}}\dot{\eta}_{\mathrm{a}}-\left(\omega_{\mathrm{a}}^2-r^2\right)\eta_{\mathrm{a}}\\ \dot{\eta}_{\mathrm{f}}=\mathrm{d}\eta_{\mathrm{f}}/\mathrm{d}t\\ \dot{\eta}_{\mathrm{a}}=\mathrm{d}\eta_{\mathrm{a}}/\mathrm{d}t\end{cases}\tag{11}$$

其中，方程式 (9) 涵盖了地表曲率、机体转动、地球引力、受结构弹性影响的气动力和推力、弹性惯性力、地球自转哥氏力和牵连惯性力的影响。u、v、w分别为体轴系下运载器相对于地面的轴向、法向和侧向速度；V_{xd}、V_{zd}分别为地面系下运载器的绝对轴向、侧向速度；p、q、r分别为体轴系下运载器的绝对滚转、偏航、俯仰角速率；ϕ、ψ、θ分别为滚转角、偏航角、俯仰角；R、λ_{c}、ϕ_{c}分别为运载器到地心的距离、经度、纬度；$\boldsymbol{L}_{\mathrm{gd}}$为地面系到体轴系的转换矩阵[46]；$G_{\mathrm{M}}$为地心引力常数；$m$为运载器质量；$I_{xx}$、$I_{yy}$、$I_{zz}$和$I_{xy}$分别为运载器沿$x$、$y$、$z$方向的惯性矩和沿$z$方向的惯性积；$F_x$、$F_y$、$F_z$、$M_x$、$M_y$、$M_z$为气动与推力合力(矩)在体轴系上的分量；$\omega_{\mathrm{E}}$为地球自转角速率。$\lambda_{\mathrm{f}}$、$\lambda_{\mathrm{a}}$、$\psi_{\mathrm{f}}$、$\psi_{\mathrm{a}}$为弹性耦合参数，计算方法如下[43]：

$$\begin{cases}\lambda_{\mathrm{f}}=\overline{m}_{\mathrm{f}}\int_0^{\overline{x}_{\mathrm{f}}}\Phi_{\mathrm{f}}(x)\,\mathrm{d}x\\ \lambda_{\mathrm{a}}=\overline{m}_{\mathrm{a}}\int_0^{\overline{x}_{\mathrm{a}}}\Phi_{\mathrm{a}}(x)\,\mathrm{d}x\\ \psi_{\mathrm{f}}=\overline{m}_{\mathrm{a}}\int_0^{\overline{x}_{\mathrm{f}}}x\Phi_{\mathrm{f}}(x)\,\mathrm{d}x\\ \psi_{\mathrm{a}}=\overline{m}_{\mathrm{a}}\int_0^{\overline{x}_{\mathrm{a}}}x\Phi_{\mathrm{a}}(x)\,\mathrm{d}x\end{cases}\tag{12}$$

其中，N_{f}、N_{a}分别为前、后梁所受的沿y_{g}方向的弹性广义力：

$$N_i(t)=\int_0^{\overline{x}_i}\Phi_i(x)\,p(x,t)\,\mathrm{d}x+\sum_{j=1}^{n}\Phi_i\left(x_j\right)P_j(t)\qquad i\in\{\mathrm{f},\mathrm{a}\}\tag{13}$$

其中：$p(x,t)$为均布载荷，如机身气动力；$P_j(t)$为集中载荷，如升降副翼面的增量气动力、推力矢量产生的法向力等。

运载器运动学方程为[46]

$$\begin{bmatrix}\dot{R}\\ \dot{\lambda}_{\mathrm{c}}\\ \dot{\phi}_{\mathrm{c}}\end{bmatrix}=\begin{bmatrix}V_{\mathrm{yd}}\\ V_{\mathrm{xd}}/(R\cos\phi_{\mathrm{c}})\\ -V_{\mathrm{zd}}/R\end{bmatrix}\tag{14}$$

$$\begin{bmatrix}\dot{\phi}\\ \dot{\psi}\\ \dot{\theta}\end{bmatrix}=\begin{bmatrix}p_{\mathrm{r}}+\tan\theta(r_{\mathrm{r}}\sin\phi-q_{\mathrm{r}}\cos\phi)\\ (r_{\mathrm{r}}\sin\phi-q_{\mathrm{r}}\cos\phi)/\cos\theta\\ r_{\mathrm{r}}\cos\phi+q_{\mathrm{r}}\sin\phi\end{bmatrix}\tag{15}$$

式中：V_{yd}为地面系下运载器的绝对法向速度；p_{r}、q_{r}、r_{r}分别为体轴系下运载器相对于地面的滚转、偏航和俯仰角速率，满足：

$$\begin{bmatrix}p_{\mathrm{r}}\\ q_{\mathrm{r}}\\ r_{\mathrm{r}}\end{bmatrix}=\begin{bmatrix}p\\ q\\ r\end{bmatrix}-\boldsymbol{L}_{\mathrm{gd}}\begin{bmatrix}-\dot{\phi}_{\mathrm{c}}\\ \left(\omega_{\mathrm{E}}+\dot{\lambda}_{\mathrm{c}}\right)\sin\phi_{\mathrm{c}}\\ -\left(\omega_{\mathrm{E}}+\dot{\lambda}_{\mathrm{c}}\right)\cos\phi_{\mathrm{c}}\end{bmatrix}\tag{16}$$

机身弹性对运动的影响体现于动力学方程式 (9) 和式 (10) 中。一方面，弹性引起的附加惯性力项显含于$\dot{u}$、$\dot{v}$和$\dot{r}$的方程中，从而作用于飞行动力学，即刚性-结构弹性运动耦合(刚-弹耦合)。另一方面，弹性变形通过改变机身梁的转角、迎角和挠度，产生附加气动力并隐含于方程的气动力项，同样作用于飞行动力学。

在计算由结构弹性引起的形变和附加气动力时，考虑到弹性变形为小变形，结合图 8 的一阶振型形状，可假设前、后段梁在变形时仍分别保持平直。基于方程 (11) 解算出的弹性广义坐标η_{f}和η_{a}，结合挠度解析式 (6) 中的一阶部分，可以计算出机身轴向各处的变形后转角、迎角和挠度：

$$\begin{cases}\tau_{\mathrm{f}}=\arctan\left(\Phi_{\mathrm{f}}(\overline{x}_{\mathrm{f}})\eta_{\mathrm{f}}(t)/\overline{x}_{\mathrm{f}}\right)\\ \tau_{\mathrm{a}}=-\arctan\left(\Phi_{\mathrm{a}}(-\overline{x}_{\mathrm{a}})\eta_{\mathrm{a}}(t)/\overline{x}_{\mathrm{a}}\right)\\ \alpha_{\mathrm{f}}^{*}=\alpha+\tau_{\mathrm{f}}\\ \alpha_{\mathrm{a}}^{*}=\alpha+\tau_{\mathrm{a}}\\ y_{\delta_{\mathrm{e}}}^{*}=w_{\mathrm{a}}\left(x_{\delta_{\mathrm{e}}},t\right)\\ y_{\mathrm{e}}^{*}=w_{\mathrm{a}}\left(x_{\mathrm{e}},t\right)\end{cases}\tag{17}$$

式中：α_{f}^{*}、α_{a}^{*}分别为前、后梁变形后迎角；$y_{\delta_{\mathrm{e}}}^{*}$、$y_{\mathrm{e}}^{*}$分别为升降副翼、推力作用点的$y_{\mathrm{g}}$向变形挠度；$x_{\delta_{\mathrm{e}}}$、$x_{\mathrm{e}}$分别为升降副翼、推力作用点的坐标。基于机身等效悬臂梁的等刚度、等截面积特性，由升力对俯仰力矩系数的影响关系，可将这 2 个气动系数分配给 2 段梁。依据变形后迎角计算 2 段梁的实际气动力和推力，然后，根据变形转角和挠度转换为刚性体轴系下的气动力和推力，最后，参与方程式 (9) 和式 (10) 运算。

2 目标运载器飞行动力学特征

2.1 刚性特征

首先，分析目标运载器的刚性飞行动力学模型的配平性能。鉴于飞行中存在着大量油门管理(加减速调度)，因而实际的配平需求是在所需油门条件下保持升力和俯仰力矩平衡。表 4 为各配平状态参数。表 5 给出按表 4中选定高度、马赫数条件下，2 种典型油门开度对应的满足升力、俯仰力矩平衡的参数。由结果可知，在同一高度和速度下，

当油门开度不同时, 源于目标运载器的飞-发耦合特性, 升降副翼和迎角配平值均有差异, 推力增大引起了附加低头力矩(需要附加上偏升降副翼), 并影响配平迎角。

表 4　配平状态

Table 4　Trimming status

状态	高度/km	马赫数
1	0	0.3
2	10	1.2
3	30	5.0
4	26	8.0

表 5　目标运载器刚性/弹性模型的配平参数

Table 5　Trim parameters of rigid/flexible target vehicle models

状态	模型	油门开度	迎角/(°)	升降副翼/(°)	弹性前、后段广义坐标/($kg^{0.5}$·m)
1	刚性	0.27	12.00	7.38	
		1.00	10.80	6.16	
	弹性	1.00	10.35	7.79	22.8、6.3
2	刚性	0.33	2.46	0.38	
		1.00	2.44	−0.73	
	弹性	1.00	1.97	0.98	26.6、4.5
3	刚性	0.67	7.07	−4.73	
		1.00	7.11	−9.42	
	弹性	1.00	6.77	−4.89	22.8、2.8
4	刚性	0	3.36	−19.52	
		0.28	3.51	−24.35	
	弹性	0.28	2.99	−18.94	38.0、−1.2
4（文献[43]）	刚性	0.26	3.70	10.10	
	弹性	0.28	2.99	11.22	6.6、3.7

作为对比, 表 5 列出了文献 [43] 中具有相近布局特征刚性飞行器在状态 4(高度为26 km、马赫数为 8)的配平参数, 可以见到, 该结果与本文目标运载器同一状态的结果可比拟, 但在升降副翼偏角方面有很大差异, 文献 [43] 飞行器具有明显更大的零升力矩。后文将继续对 2 种模型特性进行对比。

随后, 利用特征根分析目标运载器的刚性模态。表 6 列举了表 4 各状态的不稳定模态特征根。分析结果与 1.2 节相一致: 随着马赫数增大, 俯仰短时域和横侧振荡模态由不稳转为稳定、再转为不稳, 这表明目标运载器在低速和高超声速时俯仰、偏航静不稳, 跨声速时具有一定静稳定性。横、航向静稳定性的综合影响产生了偏航发散模态。

作为对比, 表 6 也列出了文献 [43]飞行器刚性模型在状态 4 的纵向模态特征根。可以见到, 该结果具有与本文运载器相似的俯仰短时域模态, 但其俯仰静稳定性稍强。

表 6　目标运载器刚性/弹性部分模态特征根

Table 6　Partial modal eigenvalues of rigid/flexible vehicle models

状态	模型	弹性模态	俯仰短时域	横侧振荡	偏航发散
1	刚性		−0.66, 0.21	0.17 ± 0.82i	3.0×10^{-5}
	弹性	−2.2 ± 28.4i −0.53 ± 9.9i	−0.70, 0.24	0.18 ± 0.80i	2.9×10^{-5}
2	刚性		−0.25 ± 1.4i	−0.078 ± 0.83i	2.6×10^{-5}
	弹性	−2.3 ± 28.9i −0.53 ± 9.9i	−0.25 ± 1.3i	0.091 ± 0.78i	2.5×10^{-5}
3	刚性		−1.54, 1.48	0.70, 0.026	1.7×10^{-4}
	弹性	−2.3 ± 30.7i −0.59 ± 10.7i	−1.49, 1.43	0.70, 0.029	1.7×10^{-4}
4	刚性		−2.41, 2.28	1.74, −0.007 2	7.5×10^{-4}
	弹性	−2.3 ± 29.5i −0.56 ± 10.0i	−2.68, 2.55	1.77, −0.003 6	7.4×10^{-4}
4（文献[43]）	刚性		−1.25, 1.20		
	弹性	−2.7 ± 49.3i −0.38 ± 16.2i	−1.42, 1.35		

2.2　弹性影响

对目标运载器的弹性飞行动力学模型, 同样在表 4 各状态进行配平和模态特征分析。将分析结果与刚性模型结果进行对比, 列于表 5 和表 6。可以看出:

1) 当油门开度相同时, 弹性变形会影响迎角、升降副翼配平值。以状态 3 为例, 如图 7 所示, 受到机身升力的主导影响, 弹性模型的机身前、后段弹性广义力N_f、N_a均向上而发生上翘变形。相比于刚性模型, 机身前段实际迎角增大、后段减小, 产生附加抬头力矩。由表 6 弹性广义坐标与式 (6) 和式 (17) 计算可得, 机身前段形变大于后段, 因此, 总升力增大、所需配平迎角减小, 但迎角减小所产生的低头力矩小于形变产生的附加抬头力矩, 因此, 配平升降副翼增大。

2) 机身弹性会给模态特性引入 2 对与刚性特征根相差很大的高频弹性特征根, 分别代表前、后段梁的弹性模态。刚-弹耦合会导致弹性模态频率比表 3 给出的机身等效梁固有频率更大, 并使俯仰短时域模态更加不稳。这种耦合效应在低速时较强, 高速时目标运载器质量减小导致弹性模态频率增大、耦合减弱。耦合对横航向模态影响较小。

作为对比, 表 5 和表 6 同样列出了文献 [43] 飞行器弹性模型在状态 4 的配平参数和纵向模态特征根。可知, 其弹性模态结构与本文目标运载器一致, 刚-弹耦合效应同样导致其俯仰短时域模态更加静不稳。区别在于目标运载器的配平升降副翼为上偏, 导致后段机身升力减小、向下弯曲。另外, 由于机身刚度和质量的差异, 二者弹性模态频率有所不同。

综上所述，本文建立的运载器弹性飞行动力学模型具有静不稳、飞-发耦合、刚-弹耦合等特性。相比于文献[43]所建立的相近布局特征飞行器模型，本文模型能更为细腻地描述油耗、横航向模态、机身弹性变形附加气动力等特征。分析可知，机身等效刚度和质量的组合产生明显的弹性模态，在低速时与刚性运动存在耦合；而在高超声速时其模态频率较高，与刚性运动模态频率差距较大，此时在控制设计中可与刚性运动自然解耦。

3 姿态-轨迹耦合控制设计

3.1 多回路PID控制

第2节分析表明了目标运载器具有较强的俯仰和偏航静不稳、飞-发耦合特征。因此，在控制设计时应先考虑在内回路引入迎角、侧滑角反馈以增稳，再通过角速率、姿态角反馈以实现不同的指令跟踪效果。按照上述思路，本文针对目标运载器弹性动力学模型，采用常规PID控制方法设计了俯仰角、滚转角跟踪控制律。

以表4中静稳定性较差的高速状态3为例，将控制增稳后模型的特征根与开环特征根对比如图9所示。由图可知，通过PID控制设计，模型的特征根全部调至稳定。

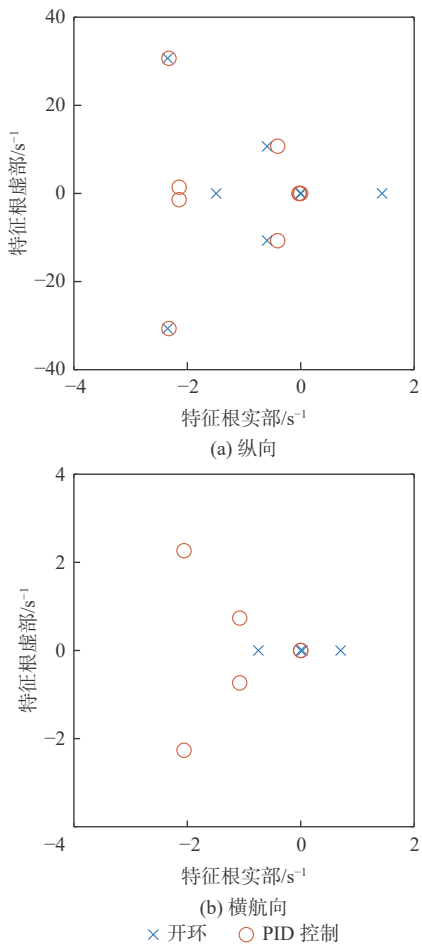


图9 PID控制对模态特征根的影响(状态3)

Fig. 9 Effect of PID controllers on eigenvalues at status 3

向增稳后的模型分别施加俯仰θ、滚转阶跃指令θ_{cmd}，结果如图10和图11所示。由图可知，模型的指令跟踪效果较好，但俯仰角跟踪会引起航迹的明显变化，刚-弹耦合会导致升降副翼偏角高频振荡，并引起法向过载的较明显振荡；滚转角跟踪也会引起迎角变化。

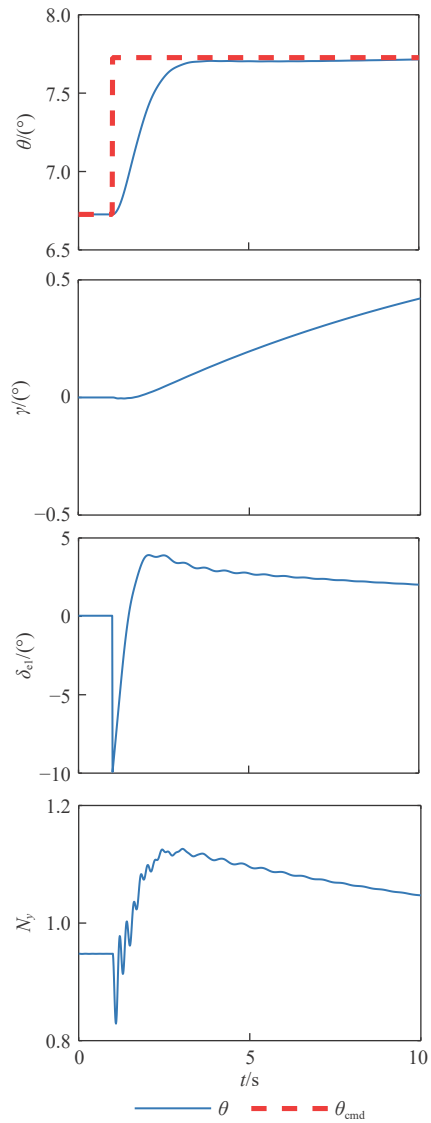


图10 常规PID控制俯仰阶跃响应

Fig. 10 Pitching step response of classical PID control

由此可知，目标运载器模型呈现明显纵-横向运动耦合、飞行姿态-轨迹耦合、刚-弹运动耦合特性，仅通过常规PID控制不易解决。

3.2 基于动态逆方法补偿

本文引入非线性动态逆(nonlinear dynamic inversion, NDI)，构建目标运载器模型的逆系统以抵消模型非线性和耦合特征。考虑到运载器的飞行走廊比较窄，在特定状态配平后，模型气动、

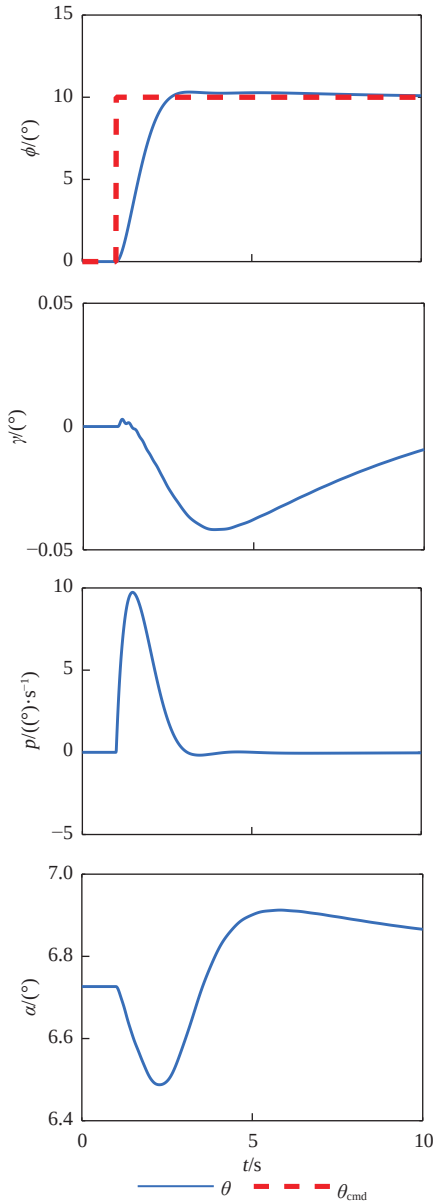


图 11　常规 PID 控制滚转阶跃响应

Fig. 11　Rolling step response of classical PID control

推进数据在该状态周围, 随高度、马赫数、迎角变化幅度较小这一特性, 因此, 可在构建逆系统时将非线性气动系数、推力系数线性化为大导数形式:

$$\begin{aligned} C_i + C_{Ti} = D_{i\delta}\Delta\delta = &[(C_i(\delta+\Delta\delta) - C_i(\delta))/\Delta\delta + \\ &(C_{Ti}(\delta+\Delta\delta) - C_{Ti}(\delta))/\Delta\delta]\Delta\delta \\ i \in &\{x,y,z,mx,my,mz\}, \delta \in \{\alpha, Ma, \beta, \delta_{e1}, \delta_{e2}, \\ &\delta_{r1}, \delta_{r2}, \cdots, \Phi, \delta_{tv1}, \delta_{tv2}, \delta_{tv3}\} \end{aligned} \tag{18}$$

式中: C_i为体轴系气动力和力矩系数; C_{Ti}为体轴系推力和力矩系数, 由推力式 (2) 计算得出; $D_{i\delta}$为气动和推力合力的线化大导数; δ为迎角、马赫数、舵偏等自变量。在此基础上, 结合 PID 进一步消除控制误差。这种处理方式有利于适应更大空、速域范围的不确定性。

下面将研究如何对回路重新划分, 解决多回路下的控制设计问题。

3.2.1　回路划分

将圆地球假设下飞行动力学方程式 (9) 和式 (10) 中与刚性运动相关的 12 个状态量u、v、w、p、q、r、ϕ、ψ、θ、R、λ_c、ϕ_c转换为一般平地球假设下的状态量α、β、V、p、q、r、μ、χ、γ、h、x、z。按照变化快慢, 将其由内到外分为 4 个控制回路: ①角速率快变量: 滚转角速率p、偏航角速率q、俯仰角速率r; ②姿态慢变量: 迎角α、侧滑角β、速矢滚转角μ; ③航迹很慢变量: 地速V、航向角χ、航迹角γ; ④位置极慢变量: 高度h、纵程x、横程z。

每层回路的指令输入对控制输出的相对阶为 1[47], 有助于简化 NDI 控制设计。

考虑到模型的不确定性主要通过气动、推力小导数及质量、参考面积等参数, 影响气动力和推力, 从而主要作用于角速率和航迹回路。因此, 具体实施时, 在这 2 个回路使用增量非线性动态逆 (incremental nonlinear dynamic inversion, INDI)控制, 可以避免建立完整的逆系统, 从而减少不确定性的影响。对于姿态和位置回路取决于解析化的运动学关系, 因而使用标准的 NDI 建立逆关系。下面具体描述各回路控制结构。

1) 角速率回路补偿操纵耦合

该回路的被控对象为目标运载器飞行动力学, 目的使其p、q、r动态能够跟踪目标角速率指令, 从而抑制纵-横-航向之间的操纵耦合。指令输入为角速率:

$$\boldsymbol{y} = [p, q, r]^{\mathrm{T}} \tag{19}$$

参考一般亚声速民机的三通道指令形式, 令控制指令为

$$\boldsymbol{u}_c = [\delta_{xc}, \delta_{yc}, \delta_{zc}]^{\mathrm{T}} \tag{20}$$

式中: δ_{xc}、δ_{yc}和δ_{zc}分别为滚转、偏航和俯仰指令。

对于目标运载器, 从三通道控制指令$\boldsymbol{u}_c$到舵偏指令的转换关系为

$$\begin{cases} \delta_{ec1} = \delta_{xc} + \delta_{zc} \\ \delta_{ec2} = -\delta_{xc} + \delta_{zc} \\ \delta_{rc1} = \delta_{rc2} = \delta_{yc} \\ \delta_{tvc1} = \delta_{tvc2} = \delta_{tvc3} = \begin{cases} 0 \\ \delta_{zc}/k \end{cases} \end{cases} \tag{21}$$

式中: δ_{ec1}、δ_{ec2}分别为右、左侧升降副翼偏角指令, 下偏为正; δ_{rc1}、δ_{rc2}分别为右、左侧方向舵偏角指

令，右偏为正；$\delta_{\mathrm{tvc1}}\sim\delta_{\mathrm{tvc3}}$为 3 台发动机在火箭模式下的推力矢量偏角指令，下偏为正，吸气模式时为 0，火箭模式时为δ_{zc}/k，k为俯仰指令与推力矢量偏角的比例系数。

由于δ_x等三通道控制量和δ_{e1}等实际舵偏量之间也满足式 (21) 的关系，因此，INDI 控制分布矩阵可写为[47]

$$\boldsymbol{B}(\boldsymbol{x})=\boldsymbol{J}^{-1}\bar{q}S_{\mathrm{ref}}L_{\mathrm{ref}}\tilde{\boldsymbol{D}} \tag{22}$$

式中：$\boldsymbol{J}$为惯性矩矩阵；$\bar{q}$为动压；S_{ref}和L_{ref}分别为参考面积和参考长度；$\tilde{\boldsymbol{D}}$为气动和推力合力矩小导数矩阵。对于吸气模式：

$$\tilde{\boldsymbol{D}}=\begin{bmatrix} D_{mx\delta_{\mathrm{e1}}}-D_{mx\delta_{\mathrm{e2}}} & D_{mx\delta_{\mathrm{r1}}}+D_{mx\delta_{\mathrm{r2}}} & 0\\ D_{my\delta_{\mathrm{e1}}}-D_{my\delta_{\mathrm{e2}}} & D_{my\delta_{\mathrm{r1}}}+D_{my\delta_{\mathrm{r2}}} & 0\\ D_{mz\delta_{\mathrm{e1}}}-D_{mz\delta_{\mathrm{e2}}} & D_{mz\delta_{\mathrm{r1}}}+D_{mz\delta_{\mathrm{r2}}} & D_{mz\delta_{\mathrm{e1}}}+D_{mz\delta_{\mathrm{e2}}} \end{bmatrix} \tag{23}$$

对于火箭模式：

$$\tilde{\boldsymbol{D}}=\begin{bmatrix} D_{mx\delta_{\mathrm{e1}}}-D_{mx\delta_{\mathrm{e2}}} & D_{mx\delta_{\mathrm{r1}}}+\boldsymbol{D}_{mx\delta_{\mathrm{r2}}} & 0\\ D_{my\delta_{\mathrm{e1}}}-D_{my\delta_{\mathrm{e2}}} & D_{my\delta_{\mathrm{r1}}}+D_{my\delta_{\mathrm{r2}}} & 0\\ D_{mz\delta_{\mathrm{e1}}}-D_{mz\delta_{\mathrm{e2}}} & D_{mz\delta_{\mathrm{r1}}}+D_{mz\delta_{\mathrm{r2}}} & D_{mz\delta_{\mathrm{e1}}}+D_{mz\delta_{\mathrm{e2}}}+D_{mz\delta_{\mathrm{tv1}}}+D_{mz\delta_{\mathrm{tv2}}}+D_{mz\delta_{\mathrm{tv3}}} \end{bmatrix} \tag{24}$$

由此可设计控制指令增量：

$$\Delta\boldsymbol{u}_{\mathrm{c}}=\begin{bmatrix} c\Delta\delta_{\mathrm{xc}}\\ \Delta\delta_{\mathrm{yc}}\\ \Delta\delta_{\mathrm{zc}} \end{bmatrix}=\boldsymbol{B}^{-1}(\boldsymbol{x}_0)\left(-\begin{bmatrix} \dot{p}_0\\ \dot{q}_0\\ \dot{r}_0 \end{bmatrix}+\begin{bmatrix} v_p\\ v_q\\ v_r \end{bmatrix}\right) \tag{25}$$

式中：$\dot{p}_0$、$\dot{q}_0$、$\dot{r}_0$为角加速率反馈信号；v_p、v_q、v_r为角速率期望输出动态(伪控制量)。

角速率跟踪响应带宽在 4 个回路中一般为最大，本文拟设计为 4/s，与弹性模态频率最接近(10～30/s)。因此，为避免刚-弹耦合，在角速率、角加速率和舵面位置反馈通道引入陷波器：

$$G_{\mathrm{b}}(s)=\frac{s^2+\omega_{\mathrm{b}}^2}{s^2+2\zeta_{\mathrm{b}}\omega_{\mathrm{b}}s+\omega_{\mathrm{b}}^2} \tag{26}$$

式中：陷波中心频率ω_{b}=9.9/s，为目标运载器低速状态下的机身前段弹性模态频率；阻尼比ζ_{b}=0.3，以达到过滤弹性模态、并获得良好角速率响应的效果。

图 12 给出了俯仰角速率跟踪结果。可知，引入陷波器后，升降副翼和机身弹性变形的振荡减弱，表明刚-弹耦合得到抑制；角速率响应迅速，调节时间约 2 s。

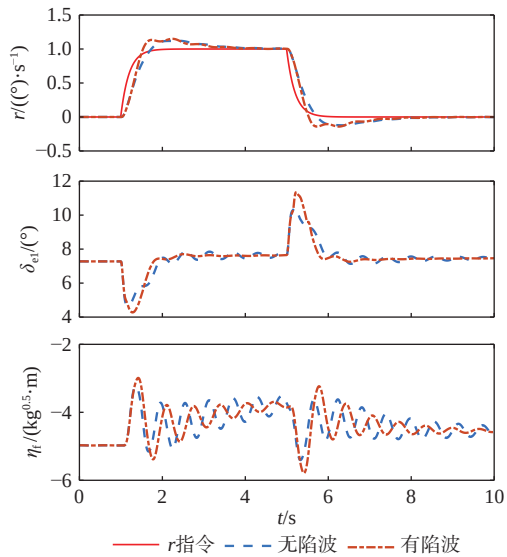


图 12 陷波器对俯仰角速率跟踪的影响

Fig. 12 Effect of band stop filter on pitch rate tracking

2）姿态回路补偿运动学耦合

该回路的被控对象为角速率回路，控制输出为角速率指令，控制目的是使运载器的μ、β、α动态能够跟踪目标姿态指令，从而抑制纵-横向运动耦合。该回路有

$$\boldsymbol{u}_{\mathrm{c}}=\left[p_{\mathrm{c}},q_{\mathrm{c}},r_{\mathrm{c}}\right]^{\mathrm{T}} \tag{27}$$

$$\boldsymbol{y}=\left[\mu,\beta,\alpha\right]^{\mathrm{T}} \tag{28}$$

式中：p_{c}、q_{c}、r_{c}分别为滚转、偏航和俯仰角速率指令。根据文献 [47] 中的姿态角微分方程，可得控制矩阵为

$$\boldsymbol{B}(\boldsymbol{x})=\begin{bmatrix} \cos\alpha\cos\beta & \sin\alpha & 0\\ -\sin\alpha\cos\beta & \cos\alpha & 0\\ \sin\beta & 0 & 1 \end{bmatrix}^{-1} \tag{29}$$

状态矩阵为[47]

$$\boldsymbol{A}(\boldsymbol{x})=-\boldsymbol{B}(\boldsymbol{x})\boldsymbol{L}_{\mathrm{gk}}\begin{bmatrix} \dot{\chi}\sin\gamma\\ \dot{\chi}\cos\gamma\\ \dot{\gamma} \end{bmatrix} \tag{30}$$

式中：$\boldsymbol{L}_{\mathrm{gk}}$为航迹系到体轴系的转换矩阵，航向角$\chi$、航迹角$\gamma$的微分为[47-48]

$$\begin{cases} \dot{\chi}=-\left(n_{xg}g\sin\alpha\sin\mu+n_{yg}g\cos\alpha\sin\mu+\right.\\ \qquad\left.n_{zg}g\cos\mu\right)/(V\cos\gamma)\\ \dot{\gamma}=\left(n_{xg}g\sin\alpha\cos\mu+n_{yg}g\cos\alpha\cos\mu-\right.\\ \qquad\left.n_{zg}g\sin\mu-g\cos\gamma\right)/V \end{cases} \tag{31}$$

其中：n_{xg}、n_{yg}、n_{zg}为体轴系的质心过载。由此得控制指令为

$$\boldsymbol{u}_{\mathrm{c}}=\begin{bmatrix}p_{\mathrm{c}}\\ q_{\mathrm{c}}\\ r_{\mathrm{c}}\end{bmatrix}=\boldsymbol{B}^{-1}(\boldsymbol{x})\left(-\boldsymbol{A}(\boldsymbol{x})+\begin{bmatrix}v_{\mu}\\ v_{\beta}\\ v_{\alpha}\end{bmatrix}\right)=\boldsymbol{L}_{\mathrm{gk}}\begin{bmatrix}\dot{\chi}\sin\gamma\\ \dot{\chi}\cos\gamma\\ \dot{\gamma}\end{bmatrix}+\boldsymbol{B}^{-1}(\boldsymbol{x})\begin{bmatrix}v_{\mu}\\ v_{\beta}\\ v_{\alpha}\end{bmatrix}\tag{32}$$

其中: v_{μ}、v_{β}、v_{α}为姿态角伪控制量。

3) 航迹回路补偿飞-发耦合

该回路的被控对象为姿态回路, 控制输出为姿态角和油门指令, 控制目标是使运载器的航迹和速度动态能够跟踪目标指令, 从而抑制由飞-发耦合导致的飞行航迹-速度耦合。该回路有

$$\boldsymbol{u}_{\mathrm{c}}=\left[\mu_{\mathrm{c}},\alpha_{\mathrm{c}},\Phi_{\mathrm{c}}\right]^{\mathrm{T}}\tag{33}$$

$$\boldsymbol{y}=\left[\chi,\gamma,V\right]^{\mathrm{T}}\tag{34}$$

令

$$\begin{cases}\boldsymbol{u}_1=\left[\mu_{\mathrm{c}}\right]\\ \boldsymbol{u}_2=\left[\Phi_{\mathrm{c}},\alpha_{\mathrm{c}}\right]^{\mathrm{T}}\end{cases}\tag{35}$$

$$\begin{cases}\boldsymbol{y}_1=\left[\chi\right]\\ \boldsymbol{y}_2=\left[V,\gamma\right]^{\mathrm{T}}\end{cases}\tag{36}$$

对于航向角通道, 由式 (31) 可设计μ_{c}控制指令:

$$\mu_{\mathrm{c}}=\arctan\frac{v_{\chi}V\cos\gamma+n_{zg}\cos\mu}{v_{\gamma}V-n_{zg}\sin\mu-g\cos\gamma}\tag{37}$$

式中: v_{χ}、v_{γ}为航向角、航迹角伪控制量。对于速度和航迹角通道, 控制矩阵为[47]

$$\boldsymbol{B}(\boldsymbol{x})=\frac{\overline{q}S_{\mathrm{ref}}}{mV}\begin{bmatrix}V\left(D_{x\Phi}\cos\alpha-D_{y\Phi}\sin\alpha\right) & V\left(D_{x\alpha}\cos\alpha-D_{y\alpha}\sin\alpha\right)\\ D_{x\Phi}\sin\alpha\cos\mu+D_{y\Phi}\cos\alpha\cos\mu & D_{x\alpha}\sin\alpha\cos\mu+D_{y\alpha}\cos\alpha\cos\mu\end{bmatrix}\tag{38}$$

由此可设计控制指令增量:

$$\Delta\boldsymbol{u}_{\mathrm{c}}=\begin{bmatrix}\Delta\alpha_{\mathrm{c}}\\ \Delta\Phi_{\mathrm{c}}\end{bmatrix}=\boldsymbol{B}^{-1}(\boldsymbol{x}_0)\left(-\begin{bmatrix}\dot{V}_0\\ \dot{\gamma}_0\end{bmatrix}+\begin{bmatrix}v_V\\ v_{\gamma}\end{bmatrix}\right)\tag{39}$$

式中: $\dot{V}_0$、$\dot{\gamma}_0$分别为速度、航迹角变化率反馈信号; v_V为速度伪控制量。

4) 位置回路实现轨迹跟踪

该回路的被控对象为航迹回路, 控制输出为航迹指令, 控制目标是使运载器的横程和高度动态能够跟踪目标指令。该回路有

$$\boldsymbol{u}_{\mathrm{c}}=\left[\chi_{\mathrm{c}},\gamma_{\mathrm{c}}\right]^{\mathrm{T}}\tag{40}$$

$$\boldsymbol{y}=\left[z,h\right]^{\mathrm{T}}\tag{41}$$

根据地速与航迹之间的运动学关系, 可设计如下控制指令[47]:

$$\boldsymbol{u}_{\mathrm{c}}=\begin{bmatrix}\chi_{\mathrm{c}}\\ \gamma_{\mathrm{c}}\end{bmatrix}=\begin{bmatrix}-\arcsin\left[v_z/(V\cos\gamma)\right]\\ \arcsin(v_h/V)\end{bmatrix}\tag{42}$$

式中: v_z、v_h为位置参数伪控制量。

3.2.2 PID 指令跟踪设计

通过动态逆将运载器的非线性动力学方程转化为解耦的伪线性系统后, 基于 PID、利用状态量的期望动态和实际动态误差, 设计各回路伪控制量:

$$\begin{bmatrix}v_p\\ v_q\\ v_r\end{bmatrix}=\left(K_{\mathrm{P1}}+\int K_{\mathrm{I1}}\mathrm{d}t\right)\begin{bmatrix}p_{\mathrm{des}}-p\\ q_{\mathrm{des}}-q\\ r_{\mathrm{des}}-r\end{bmatrix}\tag{43}$$

$$\begin{bmatrix}v_{\mu}\\ v_{\beta}\\ v_{\alpha}\end{bmatrix}=\omega_2\begin{bmatrix}\mu_{\mathrm{des}}-\mu\\ \beta_{\mathrm{des}}-\beta\\ \alpha_{\mathrm{des}}-\alpha\end{bmatrix}-\boldsymbol{\xi}_2\begin{bmatrix}\dot{\mu}\\ \dot{\beta}\\ \dot{\alpha}\end{bmatrix}\tag{44}$$

$$\begin{bmatrix}v_V\\ v_{\chi}\\ v_{\gamma}\end{bmatrix}=\boldsymbol{\omega}_3\begin{bmatrix}V_{\mathrm{des}}-V\\ \chi_{\mathrm{des}}-\chi\\ \gamma_{\mathrm{des}}-\gamma\end{bmatrix}-\boldsymbol{\xi}_3\begin{bmatrix}\dot{V}\\ \dot{\chi}\\ \dot{\gamma}\end{bmatrix}\tag{45}$$

$$\begin{bmatrix}v_z\\ v_h\end{bmatrix}=\left(K_{\mathrm{P4}}+\int K_{\mathrm{I4}}\mathrm{d}t+K_{\mathrm{D4}}\frac{\mathrm{d}}{\mathrm{d}t}\right)\begin{bmatrix}z_{\mathrm{des}}-z\\ h_{\mathrm{des}}-h\end{bmatrix}\tag{46}$$

式中: 下标des表示状态量的期望动态; $K_{\mathrm{P1}}\sim K_{\mathrm{D4}}$、$\omega_2\sim\boldsymbol{\xi}_3$为 PID 设计参数。由于内回路控制律的设计带宽会影响外回路的指令跟踪能力, 因此, 综合考虑目标运载器的舵效、过载、可用迎角范围等限制, 将上述参数设计为 K_{P1}=4/s, K_{I1}=4/s^2, K_{P4}=0.13/s, K_{I4}=0.001/s^2, K_{D4}=1; 吸气模式下: $\omega_2=2/\mathrm{s},\xi_2=1$, $\boldsymbol{\omega}_3=\begin{bmatrix}0.5&0&0\\0&0.4&0\\0&0&0.8\end{bmatrix}\Big/\mathrm{s}$, $\boldsymbol{\xi}_3=\begin{bmatrix}1&0&0\\0&0.1&0\\0&0&0.5\end{bmatrix}$; 火箭模式下: $\omega_2=1.8/\mathrm{s}$, $\boldsymbol{\xi}_2=\begin{bmatrix}2.7&0&0\\0&1.8&0\\0&0&0.9\end{bmatrix}$, $\boldsymbol{\omega}_3=\begin{bmatrix}0.3&0&0\\0&0.6&0\\0&0&1\end{bmatrix}\Big/\mathrm{s},\boldsymbol{\xi}_3=\begin{bmatrix}0.5&0&0\\0&3&0\\0&0&1\end{bmatrix}$。

以纵向为例, 本节多回路控制方案的具体结构如图 13 所示。其中, 4 个红色虚线框表示 4 层回路, 每个外回路的控制输出经过灰色矩形的指令参考模块, 作为内回路指令输入。每个回路的红色矩形表示 PID 控制器, 而绿色矩形表示建立逆系统并解算控制输出。2 个绿色虚线框表示油门和升降副翼作动器。

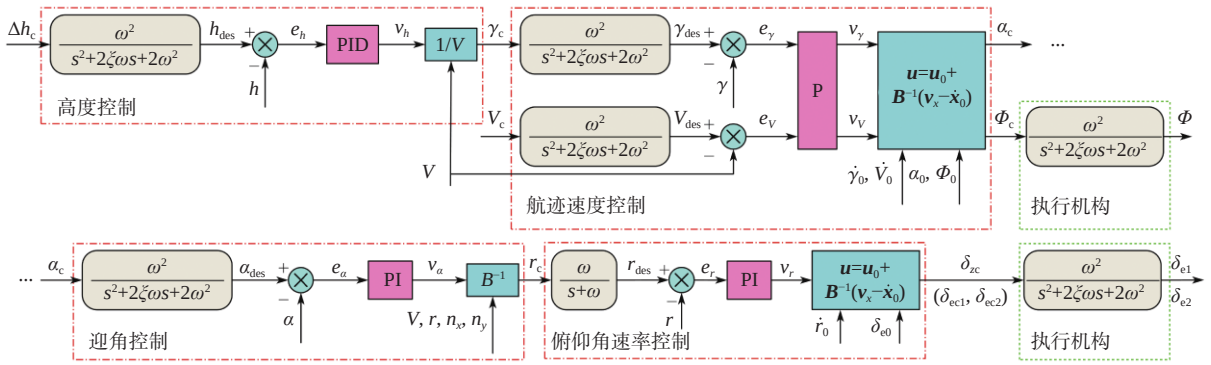


图 13 纵向通道控制结构

Fig. 13 Structure of longitudinal control channel

4 综合仿真评估

4.1 姿态指令跟踪仿真

图 14 给出所设计的多回路 NDI 与 PID 控制方案的姿态回路跟踪仿真效果。在表 4 飞行状态 3 下，向目标运载器弹性飞行动力学模型依次施加迎角α、速矢滚转角μ和侧滑角β的指令输入(红色虚线)。可以看见，姿态角的响应时间约 7 s，无稳态误差；横航向运动几乎不改变迎角；机身弹性广义坐标η_f也在 5 s 内收敛。该结果表明，目标运载器的纵-横耦合、刚-弹耦合特性在控制中得以抑制。

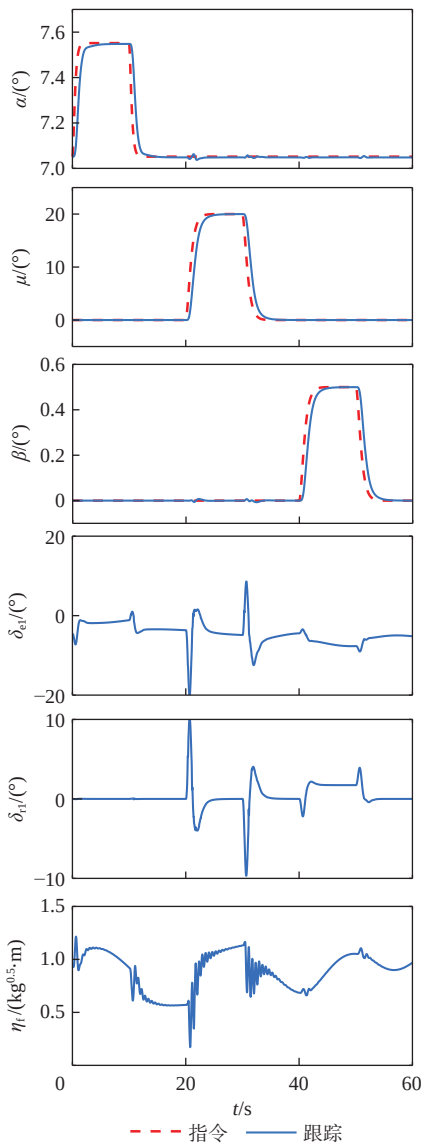


图 14 姿态指令跟踪仿真结果

Fig. 14 Simulation result of attitude tracking

4.2 上升段标称轨迹跟踪控制仿真

设计如图 15 中红色实线所示的运载器上升段目标任务剖面，具体要求如表 7 所示。该任务综合考虑了满足俯仰平衡、动压、过载、热流率等软、硬约束，以及动力由吸气式转火箭模式的窗口。使用目标运载器弹性飞行动力学模型，应用第 3 节中的姿态-轨迹耦合控制方案，选择该任务剖面的高度和速度数据作为制导指令，进行轨迹跟踪仿真。设定仿真结束条件为高度、速度与目标终端状态的总偏差率最小。

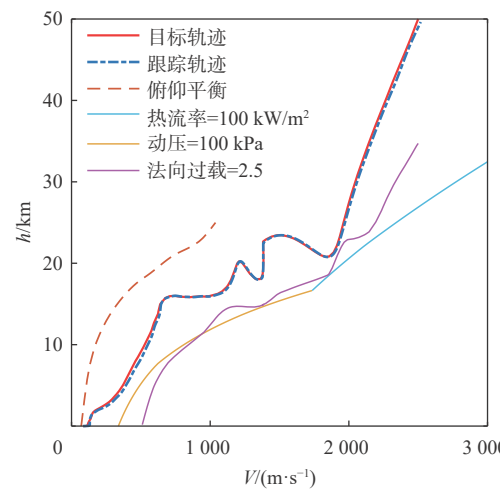


图 15 运载器上升段目标任务和跟踪轨迹

Fig. 15 Ascending task and tracking trajectory of vehicle

图 16 对比了仿真结果和目标轨迹的高度、地速、航迹角、迎角，并给出升降副翼偏角、机身弹性变形广义坐标数据，图中，t'为总飞行时间。由

表 7　上升段任务参数要求

Table 7　Parameter requirements of ascending task

初始速度/(m·s^{-1})	初始高度/km	转模式速度/(m·s^{-1})	转模式高度/km	终端速度/(m·s^{-1})	终端高度/km	终端航迹角/(°)
100	0	1 380	22.6	2 500	50	5.7

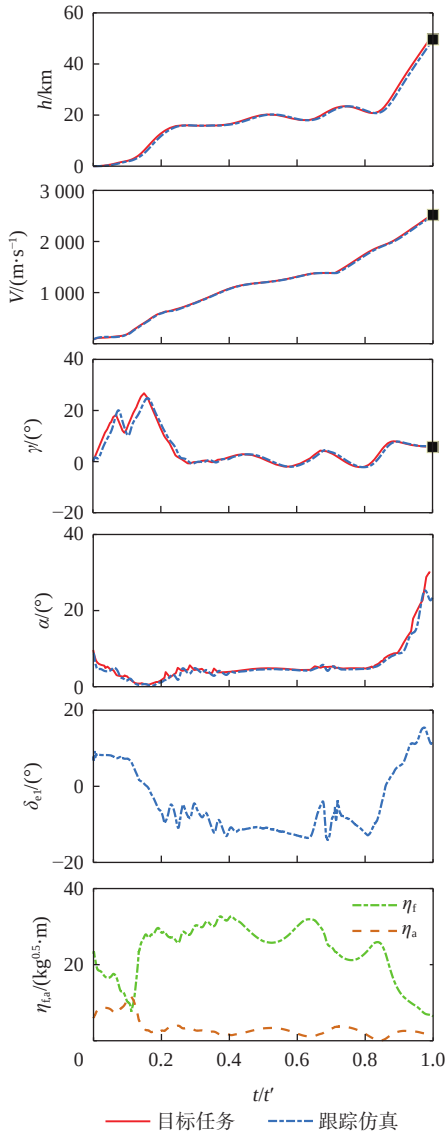


图 16　上升段轨迹跟踪仿真结果

Fig. 16　Simulation result of ascent trajectory guidance

图 15 和图 16 可知, 对于位置、航迹外回路, 目标运载器较好地跟踪了目标轨迹的高度、速度和航迹角指令, 完成了上升段飞行任务。终端高度 49.6 km 和速度 2 520 m/s 与任务目标的偏差小于 1%。对于姿态内回路, 运载器也实现了对迎角的跟踪。机身弹性未产生明显刚-弹耦合振荡现象。

4.3　不确定性分析

本文综合考虑实际飞行中存在的模型参数、大气数据和信号测量等误差, 拟对表 8 中列出的不确定性进行蒙特卡罗(Monte Carlo, MC)仿真分析, 验证运载器的轨迹跟踪控制性能。

图 17 给出了 5 000 次 MC 随机仿真的高度-速度、剩余质量-速度剖面。由图可知, 所有仿真均成功完成了目标轨迹跟踪任务, 表明所设计的控制律在表 8 不确定性的影响下, 仍具有良好的鲁棒性和抗干扰性。

为独立分析表 8 中每种不确定性偏离对任务性能的影响, 对其分别进行 100 次随机拉偏分析, 以终端速度 V_f、油耗 m_f 为性能指标。使用直线拟合各不确定性与仿真性能指标的数学关系, 计算拟合斜率如表 9 所示(将不确定性偏离无量纲化至 [−1,1] 的尺度)。图 18 给出了推力偏差 $\Delta T/T^*$、法向气动力系数偏差 $\Delta C_y/C_y^*$ 对 V_f、m_f 的影响分布, 这 2 种不确定性对任务性能影响较大, T^* 和 C_y^* 分别表示标称状态下的推力和法向气动力系数。

由表 9 可知, 按照影响程度排序, 对任务性能影响较大的不确定性依次为:

1) 推力。较大的推力有助于 V_f 增加, 并使达到目标状态所需的油耗 m_f 减少。

2) 结构质量和质心轴向位置。较大的结构质量导致油耗增大; 而较靠前的质心位置导致低头力矩增大, 升降副翼所需的上偏量增大, 产生附加阻力系数, 从而导致 V_f 减小、m_f 增大。

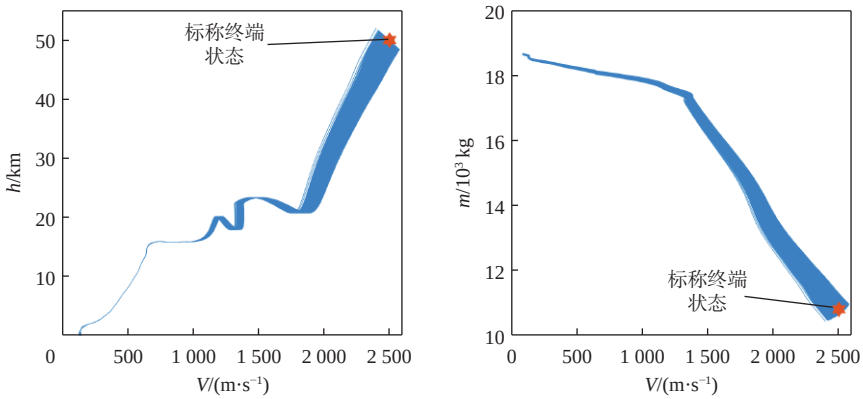
3) 轴向气动力系数 C_x、法向气动力系数 C_y 和大气密度。较大的 C_x 导致阻力增大, 与减小推力效果类似; 较大的 C_y 导致升力增大, 所需迎角减小, 导致阻力也减小, 与增大推力效果类似; 而较大的大气密度同时增大升力和阻力, 由表 9 中 C_x 和 C_y 的拟合直线斜率可知, C_x 对 m_f 的影响稍大、C_y 对 V_f 的影响较大, 因此, 最终导致 V_f 和 m_f 均增大。

其余不确定性对任务性能影响相对较小。

表 8　不确定性及其偏差

Table 8　Uncertainties and deviations

结构质量/kg	质心轴向偏移/mm	质心法向偏移/mm	俯仰转动惯量/%	推力/%	推力线法向偏移/mm	推力线俯仰偏斜/(')	大气密度/%	轴向气动力系数/%	法向气动力系数/%	俯仰角加速率测量偏差/%	迎角测量偏差/%
±200	±200	±80	±10	±5	±4	±30	±5	±5	±5	±10	±10



(a) 高度-速度　(b) 剩余质量-速度

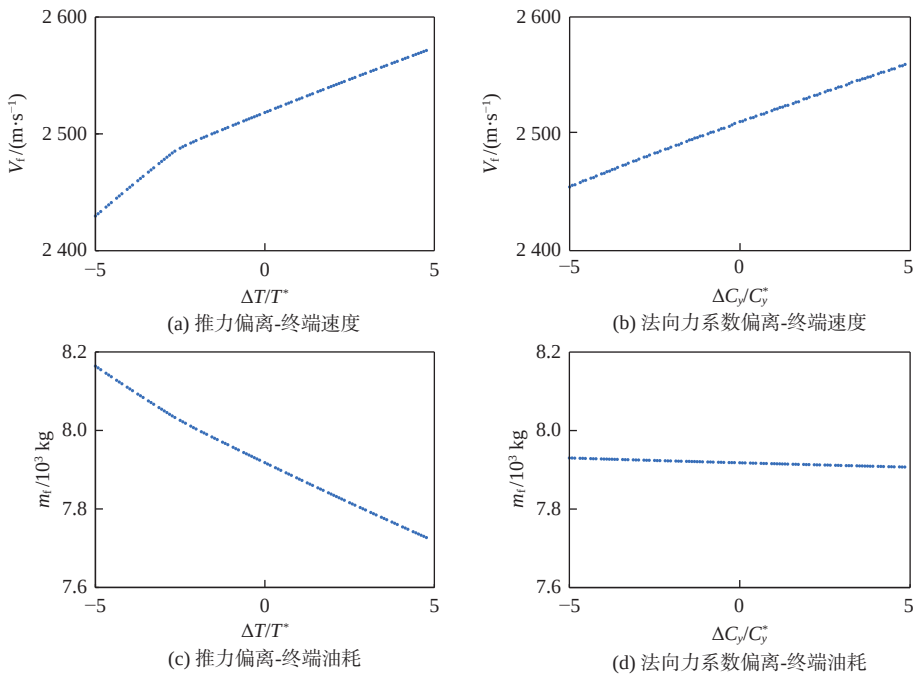
图 17　上升段轨迹跟踪 MC 仿真结果

Fig. 17　MC simulation results of ascent task trajectory

表 9　不确定性对任务性能的影响

Table 9　Effect of uncertainties on task performance

不确定性	结构质量/kg	质心轴向偏移/mm	质心法向偏移/mm	俯仰转动惯量/%	推力/%	推力线法向偏移/mm	推力线俯仰偏斜/(′)	大气密度/%	轴向气动力系数/%	法向气动力系数/%	俯仰角加速率测量偏差/%	迎角测量偏差/%
与V_t的拟合直线斜率	−18	−3.1	−0.047	0.024	67	0.009 3	0.57	1.0	−2.4	4.0	≈0	0.049
与m_t的拟合直线斜率	0.054	0.022	0.001 4	≈0	−0.22	5.1×10^{-4}	7.4×10^{-4}	0.016	0.025	−0.012	≈0	-1.4×10^{-4}



(a) 推力偏离-终端速度　(b) 法向力系数偏离-终端速度

(c) 推力偏离-终端油耗　(d) 法向力系数偏离-终端油耗

图 18　不确定性对任务性能的影响分布示例

Fig. 18　Exemplary distributions of uncertainties and task performance

综合以上分析结果可知，本文结合 NDI 与 PID 所设计的多回路控制律能够抑制运载器飞行中的纵-横耦合、飞-发耦合和刚-弹耦合特性，并能适应其姿态-轨迹耦合特性，实现姿态跟踪、轨迹跟踪等控制目标，在不确定性影响下具有良好的鲁棒性。

5　结　论

本文针对翼身融合组合动力运载器易凸显的气动非线性、纵-横向耦合、气动-动力-结构弹性耦合和飞行姿态-轨迹耦合等特性，建立其弹性飞行动力学模型，设计了能够抑制耦合的控制方案。研究结论如下：

1）本文建立目标运载器考虑多种耦合特性的飞行动力学模型，并通过悬臂梁简化模型、从弹性变形和运动耦合两方面入手，描述机身弹性对运动的影响。与参考文献中具有相近布局特征的飞行

器模型进行对比, 结果表明二者具有可比拟的静不稳、飞-发耦合、刚-弹运动耦合特性, 且本文模型对质量变化、横航向运动、弹性变形附加气动力等特征的描述更为细腻。

2) 目标运载器具有较强的气动非线性; 质心靠后, 具有高速时尤为显著的纵、航向静不稳定性, 导致较大的俯仰和横侧不稳定模态特征根; 翼身融合布局导致明显的纵-横向气动和操纵耦合特性; 飞发一体化外形产生显著的飞-发耦合特性, 体现在升降副翼和油门的控制耦合; 机身较低的刚度产生高频弹性模态, 在低速时与刚性运动耦合明显, 而在高速时可与之自然解耦。这些特性难以通过传统 PID 控制进行抑制。

3) 将运载器飞行状态量按照快慢分为 4 个回路, 引入动态逆, 在角速率回路补偿纵-横向操纵耦合、姿态回路补偿运动耦合、航迹回路补偿飞-发耦合、位置回路实现轨迹跟踪, 再结合 PID 进行指令跟踪控制设计。大量拉偏仿真表明, 控制律在抑制耦合的同时, 具有良好的角速率、姿态和轨迹跟踪性能, 在模型、环境等不确定性影响下具有良好的鲁棒性和抗干扰性。

参考文献 (References)

[1] 陈柏屹. 空天飞行器面向控制一体化的建模及多系统关联特性分析[D]. 南京: 南京航空航天大学, 2019.
CHEN B Y. Control-oriented modeling and associated characteristic analysis for integrated design of aerospace vehicles[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2019(in Chinese).

[2] SZIROCZAK D, SMITH H. A review of design issues specific to hypersonic flight vehicles[J]. Progress in Aerospace Sciences, 2016, 84: 1-28.

[3] 佘文学, 刘凯, 乔鸿. 组合动力空天飞行器制导技术发展分析[J]. 战术导弹技术, 2020(5): 52-65.
SHE W X, LIU K, QIAO H. Development analysis of guidance technology for aerospace vehicle based on combination engine[J]. Tactical Missile Technology, 2020(5): 52-65(in Chinese).

[4] JENKINS D R. X-15 extending the frontiers of flight[M]. Washington, D. C. : NASA, 2007: 1-10.

[5] DOYLE J, LENZ K, PACKARD A. Design examples using μ-synthesis: space shuttle lateral axis FCS during reentry[C]//Proceedings of the Modelling, Robustness and Sensitivity Reduction in Control Systems. Berlin: Springer, 1987: 127-154.

[6] 佘文学, 刘晓鹏, 刘凯. 桑格尔空天飞行器技术途径分析与思考[J]. 火箭推进, 2021, 47(6): 11-20.
SHE W X, LIU X P, LIU K. Analysis and thinking on technical approach of Sanger aerospace vehicle[J]. Journal of Rocket Propulsion, 2021, 47(6): 11-20(in Chinese).

[7] BURNS B A. HOTOL space transport for the twenty-first century[J]. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 1990, 204(2): 101-110.

[8] ZHANG T T, WANG Z G, HUANG W, et al. The overall layout of rocket-based combined-cycle engines: a review[J]. Journal of Zhejiang University: Science A, 2019, 20(3): 163-183.

[9] ZUO F Y, MÖLDER S. Hypersonic wavecatcher intakes and variable-geometry turbine based combined cycle engines[J]. Progress in Aerospace Sciences, 2019, 106: 108-144.

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

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

[12] LONGSTAFF R, BOND A. The SKYLON project[C]//Proceedings of the 17th AIAA International Space Planes and Hypersonic Systems and Technologies Conference. Reston: AIAA, 2011.

[13] HELLMAN B M, BRADFORD J E, ST GERMAIN B D, et al. Two stage to orbit conceptual vehicle designs using the SABRE engine[C]//Proceedings of the AIAA SPACE 2016. Reston: AIAA, 2016.

[14] STEELANT J, VARVILL R, WALTON C, et al. Achievements obtained for sustained hypersonic flight within the LAPCAT-II project[C]//Proceedings of the 20th AIAA International Space Planes and Hypersonic Systems and Technologies Conference. Reston: AIAA, 2015.

[15] RODRÍGUEZ-SEGADE M, HERNÁNDEZ S, DÍAZ J. Multibubble scheme and structural analysis of a hypersonic stratospheric flight vehicle[J]. Aerospace Science and Technology, 2022, 124: 107514.

[16] MARSHALL L, CORPENING G, SHERRILL R. A chief engineer's view of the NASA X-43A scramjet flight test[C]//Proceedings of the AIAA/CIRA 13th International Space Planes and Hypersonics Systems and Technologies Conference. Reston: AIAA, 2005.

[17] HANK J, MURPHY J, MUTZMAN R. The X-51A scramjet engine flight demonstration program[C]//Proceedings of the 15th AIAA International Space Planes and Hypersonic Systems and Technologies Conference. Reston: AIAA, 2008: 2540.

[18] 郭建国, 胡冠杰, 郭宗易. 组合动力空天飞行器控制问题研究[J]. 空天技术, 2022(2): 27-41.
GUO J G, HU G J, GUO Z Y. Research on control problems of combined power aerospace vehicle[J]. Aerospace Technology, 2022 (2): 27-41(in Chinese).

[19] 曹林. 吸气式高超声速飞行器轨迹跟踪控制方法研究[D]. 西安: 西北工业大学, 2018.
CAO L. Research on trajectory tracking control method of air-breathing hypersonic vehicle[D]. Xi'an: Northwestern Polytechnical University, 2018(in Chinese).

[20] 佘文学, 刘凯, 刘晶. 空天飞行器制导控制技术发展思考[J]. 战术导弹技术, 2017(4): 1-10.
SHE W X, LIU K, LIU J. Thoughts on the development of guidance and control technology for aerospace vehicle[J]. Tactical Missile Technology, 2017(4): 1-10(in Chinese).

[21] SHAUGHNESSY J D, PINCKNEY S Z, MCMINN J D, et al. Hypersonic vehicle simulation model: winged-cone configuration:

NASA-TM-102610[R]. Washington, D. C. : NASA, 1990: 1-18.

[22] MIRMIRANI M, WU C, CLARK A, et al. Airbreathing hypersonic flight vehicle modeling and control, review, challenges, and a CFD-based example[C]//Proceedings of the Workshop on Modeling and Control of Complex Systems. Laxenburg: IFAC Press, 2005: 1-15.

[23] BOLENDER M, DOMAN D. A non-linear model for the longitudinal dynamics of a hypersonic air-breathing vehicle[C]//Proceedings of the AIAA Guidance, Navigation, and Control Conference and Exhibit. Reston: AIAA, 2005: 6255.

[24] 安昊. 吸气式高超声速飞行器控制方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2017: 11-77.
AN H. Research on control methods of air-breathing hypersonic vehicles[D]. Harbin: Harbin Institute of Technology, 2017: 11-77(in Chinese).

[25] LI Z Y, SHI S. L_1 adaptive loss fault tolerance control of unmanned hypersonic aircraft with elasticity[J]. Aerospace, 2021, 8(7): 176.

[26] 宋征宇, 潘豪, 王聪, 等. 长征运载火箭飞行控制技术的发展[J]. 宇航学报, 2020, 41(7): 868-879.
SONG Z Y, PAN H, WANG C, et al. Development of flight control technology of long March launch vehicles[J]. Journal of Astronautics, 2020, 41(7): 868-879(in Chinese).

[27] UR REHMAN O, PETERSEN I R, FIDAN B. Feedback linearization-based robust nonlinear control design for hypersonic flight vehicles[J]. Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering, 2013, 227(1): 3-11.

[28] GAO G, WANG J Z. Observer-based fault-tolerant control for an air-breathing hypersonic vehicle model[J]. Nonlinear Dynamics, 2014, 76(1): 409-430.

[29] 闫斌斌, 林泽淮, 刘双喜, 等. 基于动态逆控制的高超声速飞行器飞/发一体化控制方法研究[J]. 西北工业大学学报, 2023, 41(5): 878-886.
YAN B B, LIN Z H, LIU S X, et al. Research on integrated aircraft-engine control method of hypersonic vehicle based on dynamic inversion control[J]. Journal of Northwestern Polytechnical University, 2023, 41(5): 878-886(in Chinese).

[30] ZHANG S, WANG Q, YANG G, et al. Anti-disturbance backstepping control for air-breathing hypersonic vehicles based on extended state observer[J]. ISA Transactions, 2019, 92: 84-93.

[31] YU C J, JIANG J, ZHEN Z Y, et al. Adaptive backstepping control for air-breathing hypersonic vehicle subject to mismatched uncertainties[J]. Aerospace Science and Technology, 2020, 107: 106244.

[32] 林伟. 吸气式高超声速飞行器抗动态饱和自适应滑模控制研究[D]. 北京: 北京航空航天大学, 2020.
LIN W. Research of adaptive sliding mode control for air-breathing hypersonic vehicles with dynamic saturation[D]. Beijing: Beihang University, 2020(in Chinese).

[33] DING Y B, WANG X G, BAI Y L, et al. An improved continuous sliding mode controller for flexible air-breathing hypersonic vehicle[J]. International Journal of Robust and Nonlinear Control, 2020, 30(14): 5751-5772.

[34] WANG Y Y, YANG X X, YAN H C. Reliable fuzzy tracking control of near-space hypersonic vehicle using aperiodic measurement information[J]. IEEE Transactions on Industrial Electronics, 2019, 66(12): 9439-9447.

[35] CHAI R Q, TSOURDOS A, SAVVARIS A, et al. Six-DOF spacecraft optimal trajectory planning and real-time attitude control: a deep neural network-based approach[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(11): 5005-5013.

[36] 武天才, 王宏伦, 任斌, 等. 考虑规避与突防的高超声速飞行器智能容错制导控制一体化设计[J]. 航空学报, 2024, 45(15): 329607.
WU T C, WANG H L, REN B, et al. Learning-based integrated fault-tolerant guidance and control for hypersonic vehicles considering avoidance and penetration[J]. Acta Aeronautica et Astronautica Sinica, 2024, 45(15): 329607(in Chinese).

[37] PAN L, PENG S C, XIE Y, et al. 3D guidance for hypersonic reentry gliders based on analytical prediction[J]. Acta Astronautica, 2020, 167: 42-51.

[38] 许东欢, 徐桂甲, 叶赛仙, 等. 基于伪谱法的空天飞行器上升段非线性闭环最优制导算法[J]. 战术导弹技术, 2017(4): 57-65.
XU D H, XU G J, YE S X, et al. Nonlinear closed-loop optimal ascent guidance for aerospace vehicle based on pseudospectral method[J]. Tactical Missile Technology, 2017(4): 57-65(in Chinese).

[39] ZHOU H Y, WANG X G, SHAN W Z, et al. Ascent guidance law for a horizontal take-off vehicle with a multi-combined cycle engine[J]. Advances in Space Research, 2020, 65(1): 379-391.

[40] LI M M, HU J, HUANG H. A segmented and weighted adaptive predictor-corrector guidance method for the ascent phase of hypersonic vehicle[J]. Aerospace Science and Technology, 2020, 106: 106231.

[41] 周建兴, 张浩成, 高启滨, 等. 基于SABRE技术的高超声速预冷飞行器应用分析[J]. 推进技术, 2018, 39(10): 2196-2206.
ZHOU J X, ZHANG H C, GAO Q B, et al. Analysis of vehicle applications propelled by SABRE-based precooling hypersonic engine[J]. Journal of Propulsion Technology, 2018, 39(10): 2196-2206 (in Chinese).

[42] 中华人民共和国国家质量监督检验检疫总局, 中国国家标准化管理委员会. 航天飞行器常用坐标系 : GB/T 32296-2015[S]. 北京: 中华人民共和国国家质量监督检验检疫总局, 中国国家标准化管理委员会, 2015-12-31.
General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. General coordinates for space vehicle: GB/T 32296-2015[S]. Beijing: General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China, 2015-12-31(in Chinese).

[43] BOLENDER M A, DOMAN D B. Nonlinear longitudinal dynamical model of an air-breathing hypersonic vehicle[J]. Journal of Spacecraft and Rockets, 2007, 44(2): 374-387.

[44] CHAVEZ F R, SCHMIDT D K. Analytical aeropropulsive-aeroelastic hypersonic-vehicle model with dynamic analysis[J]. Journal of Guidance, Control, and Dynamics, 1994, 17(6): 1308-1319.

[45] MEIROVITCH L. Analytical methods in vibrations[M]. New York: Macmillan, 1967: 161-166.

[46] 肖业伦. 航空航天器运动的建模: 飞行动力学的理论基础[M]. 北京: 北京航空航天大学出版社, 2003: 38-51.
XIAO Y L. Modeling of aerospace vehicle motion: theoretical basis of flight dynamics[M]. Beijing: Beijing University of Aeronautics & Astronautics Press, 2003: 38-51(in Chinese).

[47] LU P, VAN KAMPEN E J, DE VISSER C, et al. Aircraft fault-tolerant trajectory control using Incremental Nonlinear Dynamic Inversion[J]. Control Engineering Practice, 2016, 57: 126-141.

[48] 陈海兵, 张曙光, 方振平. 加速度反馈的隐式动态逆鲁棒非线性控制律设计[J]. 航空学报, 2009, 30(4): 597-603.
CHEN H B, ZHANG S G, FANG Z P. Implicit NDI robust nonlinear control design with acceleration feedback[J]. Acta Aeronautica et Astronautica Sinica, 2009, 30(4): 597-603(in Chinese).

Coupling suppression method of blended-wing-body combined-cycle-power vehicle

XIONG Zhiyue[1], ZHANG Shuguang[1,*], DING Jiayuan[2], LIAO Yuzhou[1]

(1. School of Transportation Science and Engineering, Beihang University, Beijing 100191, China;
2. China Academy of Launch Vehicle Technology, Beijing 100076, China)

Abstract: Horizontal takeoff and landing vehicles with integrated aero engines, blended wing bodies, and combined cycle power have the potential to be swiftly deployed and utilized in repeated aerospace round-trip missions with great transportation efficiency. However, their flight dynamics are affected by strong nonlinearity, aero-engine-elastic coupling, longitudinal-lateral kinematic coupling, and flight attitude-trajectory coupling, making control design challenging. In this regard, a flight dynamics model is established in this paper, which describes various coupling characteristics of a general blended-wing-body combined-cycle-power vehicle. Then, its flight dynamic characteristics are analyzed. By combining traditional PID control with dynamic inversion to suppress couplings, a multi-loop control law that is simple to apply is created. The multiple stochastic simulation results show that the control law not only helps suppress couplings but also achieves good tracking performance in attitude and trajectory, as well as good control robustness and anti-interference ability under uncertainties such as modeling, environment, and sensing.

Keywords: hypersonic vehicle; blended-wing-body; airframe elasticity; dynamic coupling; multi-loop control; trajectory tracking

Received: 2024-04-29; **Accepted**: 2024-06-11; **Published Online**: 2024-06-20 13:32
URL: link.cnki.net/urlid/11.2625.V.20240619.1628.003

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0267

DoS攻击下基于H_∞的移动机器人系统鲁棒控制

季旭东，陈友东*，魏洪兴
(北京航空航天大学 机械工程及自动化学院，北京 100191)

摘 要：移动机器人系统入网逐渐应用到多种场景中，然而接入网络会使系统面临网络攻击的风险，进而可能导致系统状态失稳，甚至发生功能失效的情况。针对移动机器人系统在受拒绝服务(DoS)攻击下的系统状态稳定性，提出一种基于H_∞控制的移动机器人系统的鲁棒策略。将DoS攻击对系统的影响建模为具有伯努利分布的随机丢包，采用基于状态观测器的反馈补偿对移动机器人系统受到攻击时进行鲁棒控制；利用李雅普诺夫稳定性理论，得到满足闭环系统指数均方稳定和H_∞控制的充分条件；基于此条件求解线性矩阵不等式约束，得到了观测器和控制器的增益矩阵，进而实现系统鲁棒控制，达到抗干扰的效果。通过仿真对比实验验证了所提策略能够有效降低系统受DoS攻击的影响，保证系统状态稳定性。

关 键 词：移动机器人；DoS攻击；H_∞控制；观测器；鲁棒控制
中图分类号：TP242.6; X959; V249
文献标志码：A **文章编号**：1001-5965(2026)06-1982-08

移动机器人技术作为当前无人驾驶、智慧物流[1]、智能交通、智能建筑[2]等民生和工业领域技术发展的关键支持[3-4]，扮演着举足轻重的角色[5]。

工业互联网的发展使原来孤立的移动机器人系统升级为互联的智能设备，移动机器人从单机走向互联。当移动机器人通过开放和共享的通信网络在智能传感器、执行器和电气控制单元之间交换信息时，恶意攻击者可以通过执行网络攻击来破坏机器人系统的完整性，从而降低机器人系统的性能[6-7]。攻击可以在不同的模块中执行，如传感器/执行器、通信网络和物理接口。例如，在文献[8]中攻击者通过欺骗了游艇的导航系统GPS使游艇偏离航线；Maggi等详细描述了机器人如何受到高级黑客的攻击[9]。这些网络攻击可大致分为2类：代理的虚假数据注入(false data injection, FDI)攻击[10-11]和节点之间通信的拒绝服务(denial-of-service, DoS)攻击[12]。DoS攻击是最频繁和最容易实现的攻击模式，攻击者只需要掌握网络层的通信协议就可以发动DoS攻击。对于移动机器人而言，这类攻击可能导致机器人任务偏离期望[13]、降低机器人工作效率，可能损坏设备，甚至伤害人类[14]。

鲁棒控制能够解决此类攻击引起的控制系统中信号延迟或丢失对系统产生的不利影响，以保证设备受恶意攻击和干扰时能够正常运行。DoS攻击下网络控制系统鲁棒控制的研究已经有很多[15-16]。但是针对DoS攻击下移动机器人控制系统的鲁棒控制目前研究较少，这主要是目前的移动机器人的通信网络大部分是局域网。但是网络进一步开放导致系统受DoS攻击的风险增加。Tang等研究移动机器人受DoS攻击导致传感器数据丢失的跟踪控制问题[17]，赵宁等针对遭受DoS攻击的网络化控制系统，提出基于观测器输出的反馈控制方法[18]，都缺乏对控制输入数据端受DoS攻击的研究。Geng等基于马尔可夫模型提出一种混合动态控制器，保证

收稿日期：2024-04-29；**录用日期**：2024-07-05；**网络出版时间**：2024-07-26 09:15
网络出版地址：link.cnki.net/urlid/11.2625.V.20240725.1539.003
基金项目：国家重点研发计划(2022YFB4702102)
*通信作者. E-mail：chenyd@buaa.edu.cn

引用格式：季旭东，陈友东，魏洪兴. DoS攻击下基于H_∞的移动机器人系统鲁棒控制[J]. 北京航空航天大学学报，2026，52(6)：1982-1989. JI X D，CHEN Y D，WEI H X. Robust control of mobile robots based on H_∞ under DoS attack[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1982-1989 (in Chinese).

DoS 攻击下移动机器人轨迹跟踪控制问题[19], 虽然该研究考虑了移动机器人系统控制输入数据端和传感器端同时受攻击的情况, 但是并未考虑系统干扰。因此, 有必要对移动机器人在受 DoS 攻击鲁棒控制策略展开研究, 提高其安全性和可控性。

本文旨在研究移动机器人系统在受 DoS 攻击下如何实现控制系统状态稳定, 建立一个针对 DoS 攻击的鲁棒控制策略。该策略将 DoS 攻击的随机模型带入到含有状态观测器的鲁棒控制策略中, 并跟据系统满足 H_∞性能水平条件的条件得到系统适应于 DoS 攻击的控制增益, 从而保证系统可控性。

1　系统模型与控制策略

1.1　预备知识

考虑线性移动机器人系统。通常由状态向量表示系统中机器人物理状态, 而机器人的物理状态一般通过传感器周期性或事件触发性采集获得, 故而系统动力学模型采用离散模型。机器人当前时刻的状态取决于: 上一时刻机器人的状态和上一时刻系统的控制输入, 线性移动机器人系统的动力学可描述为

$$\begin{cases}\boldsymbol{x}(k+1)=\boldsymbol{A}\boldsymbol{x}(k)+\boldsymbol{B}\boldsymbol{u}(k)+\boldsymbol{D}\boldsymbol{w}(k) & k\in\mathbf{Z}_+\\ \boldsymbol{y}(k)=\boldsymbol{C}\boldsymbol{x}(k)\end{cases} \tag{1}$$

式中: $\boldsymbol{x}(k)$为机器人的状态向量, $\boldsymbol{x}(k)\in\mathbf{R}^n$, n 为状态维度; $\boldsymbol{u}(k)$为机器人的控制向量, $\boldsymbol{u}(k)\in\mathbf{R}^m$, m 为控制输入维度; k 为离散化的时间; $\boldsymbol{y}(k)\in\mathbf{R}^d$为机器人的测量值, d为传感器测系统状态量的维度; $\boldsymbol{w}(k)$为不确定干扰扰动; $\boldsymbol{A}$、$\boldsymbol{B}$、$\boldsymbol{C}$ 和 $\boldsymbol{D}$ 为系统给定的常量矩阵。

假设　线性系统是稳定的并且是可观察的。

H_∞控制理论属于鲁棒控制的一种, 能够使系统干扰至误差的传递函数的 H_∞范数最小, 从而将有限功率谱的干扰对系统误差的影响降到最低限度。

引理 1[20]　让 $V(\boldsymbol{\eta}(k))$作为一个李亚普诺夫函数。如果存在实数标量 $\lambda\geqslant 0,\mu>0,\nu>0,0<\varphi<1$, 使得

$$\mu\left\|\boldsymbol{\eta}(k)\right\|^2\leqslant V(\boldsymbol{\eta}(k))\leqslant\nu\left\|\boldsymbol{\eta}(k)\right\|^2 \tag{2}$$

并且

$$E\{V(\boldsymbol{\eta}(k+1)|\boldsymbol{\eta}(k)\}-V(\boldsymbol{\eta}(k))\leqslant\lambda-\varphi V(\boldsymbol{\eta}(k)) \tag{3}$$

那么 $\boldsymbol{\eta}(k)$满足:

$$E\left\{\left\|\boldsymbol{\eta}(k)\right\|^2\right\}\leqslant\frac{\nu}{\mu}\left\|\boldsymbol{\eta}(0)\right\|^2(1-\varphi)^k+\frac{\lambda}{\mu\varphi} \tag{4}$$

引理 2[21]　对于列满秩矩阵 $\boldsymbol{X}$, 总存在一个奇异值分解, 如:

$$\boldsymbol{X}=\boldsymbol{U}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{\Sigma}\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{V}^{\mathrm{T}}=\begin{bmatrix}\boldsymbol{U}_1\\ \boldsymbol{U}_2\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{V}^{\mathrm{T}} \tag{5}$$

式中: $\boldsymbol{U}\in\mathbf{R}^{n\times n}$、$\boldsymbol{V}\in\mathbf{R}^{m\times m}$为 2 个正交矩阵; $\boldsymbol{\Sigma}=\mathrm{diag}\{b_1,b_2,\cdots,b_m\}$; $\boldsymbol{U}=\begin{bmatrix}\boldsymbol{U}_1^{\mathrm{T}} & \boldsymbol{U}_2^{\mathrm{T}}\end{bmatrix}^{\mathrm{T}}$; b_i为矩阵 $\boldsymbol{X}$ 的非奇异值。

引理 3[22]　对于列满秩矩阵 $\boldsymbol{X}$, 如果存在 2 个正定矩阵 $\boldsymbol{P}_1\in\mathbf{R}^{m\times m}$, $\boldsymbol{P}_2\in\mathbf{R}^{(n-m)\times(n-m)}$使得

$$\boldsymbol{P}=\boldsymbol{U}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{P}_1 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{P}_2\end{bmatrix}\boldsymbol{U}=\boldsymbol{U}_1^{\mathrm{T}}\boldsymbol{P}_1\boldsymbol{U}_1+\boldsymbol{U}_2^{\mathrm{T}}\boldsymbol{P}_2\boldsymbol{U}_2 \tag{6}$$

则存在一个非奇异的矩阵 $\overline{\boldsymbol{P}}\in\mathbf{R}^{m\times m}$满足 $\boldsymbol{P}\boldsymbol{X}=\boldsymbol{X}\overline{\boldsymbol{P}}$。

Schur 补理[23]　对于给定的对称矩阵 $\boldsymbol{S}=\begin{bmatrix}\boldsymbol{S}_{11} & \boldsymbol{S}_{12}\\ \boldsymbol{S}_{21} & \boldsymbol{S}_{22}\end{bmatrix}$, 其中, $\boldsymbol{S}_{11}$为方阵, 以下 3 个条件是等价的: ①$\boldsymbol{S}<\boldsymbol{0}$; ②$\boldsymbol{S}_{11}<\boldsymbol{0},\boldsymbol{S}_{22}-\boldsymbol{S}_{12}^{\mathrm{T}}\boldsymbol{S}_{11}^{-1}\boldsymbol{S}_{12}<\boldsymbol{0}$; ③$\boldsymbol{S}_{22}<\boldsymbol{0}$, $\boldsymbol{S}_{11}-\boldsymbol{S}_{12}\boldsymbol{S}_{22}^{-1}\boldsymbol{S}_{12}^{\mathrm{T}}<\boldsymbol{0}$。

1.2　DoS 攻击的移动机器人模型

移动机器人系统受 DoS 攻击模型如图 1 所示[24], 由传感器、上位机、移动机器人组成, 通过无线网络进行数据传输。系统状态为过程量, 可以是机器人的位置、速度、加速度或它们的组合。网络被用来连接传感器到上位机和上位机到机器人通道。

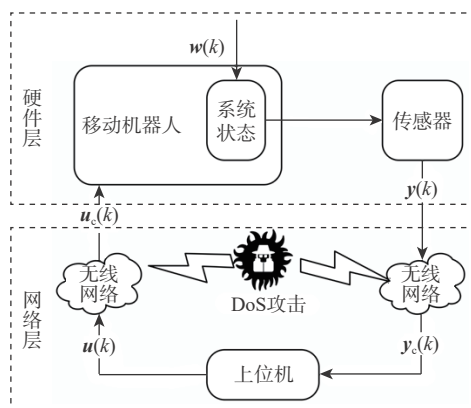


图 1　DoS 攻击下的移动机器人模型

Fig. 1　Mobile robot model under DoS attack

本文考虑的移动机器人系统是由上位机、移动机器人、传感器 3 部分构成的闭环系统, 上位机接收由传感器发来的数据, 移动机器人收到上位机发送的指令并执行, 系统状态发生改变。在数据交互过程中可能受 DoS 攻击的影响而无法传输。

网络中可能存在 DoS 攻击者, 发起攻击以阻止网络通信。当受到 DoS 攻击时, 网络的传输能力会急剧下降, 并随机造成数据包丢失。系统还可能受一个加性干扰影响。

本文考虑DoS攻击的总持续时间不能超过一定上限。如果系统多次受到DoS攻击,且攻击持续时间不受限制,可能会导致系统一直处于不稳定状态,本文暂不考虑。

为安全起见,上位机和移动机器人采用了攻击补偿机制,用最后收到的数据保持移动机器人运行。补偿后的系统状态测量值$\boldsymbol{y}_{\mathrm{c}}(k)$和系统控制输入$\boldsymbol{u}_{\mathrm{c}}(k)$可以表示为

$$\begin{cases}\boldsymbol{y}_{\mathrm{c}}(k)=(1-\delta_1)\boldsymbol{y}(k)+\delta_1\boldsymbol{y}(k-1)\\ \boldsymbol{u}_{\mathrm{c}}(k)=(1-\delta_2)\boldsymbol{u}(k)+\delta_2\boldsymbol{u}(k-1)\end{cases} \tag{7}$$

式中:δ_1、δ_2用于描述DoS攻击是否成功,写为

$$\delta_i=\begin{cases}0\\1\end{cases}\quad i=1,2 \tag{8}$$

其中:$\delta_1=1$或$\delta_2=1$表示测量(控制)数据不能成功传输到上位机(机器人);否则,数据传输成功。

本文δ_i遵循伯努利分布,满足:

$$\begin{cases}P\{\delta_1=1\}=E\{\delta_1\}=\bar{\delta}_1\\ P\{\delta_1=0\}=1-E\{\delta_1\}=1-\bar{\delta}_1\\ P\{\delta_2=1\}=E\{\delta_2\}=\bar{\delta}_2\\ P\{\delta_2=0\}=1-E\{\delta_2\}=1-\bar{\delta}_2\\ E\left\{\left(\bar{\delta}_1-\delta_1\right)^2\right\}=\bar{\delta}_1\left(1-\bar{\delta}_1\right)\\ E\left\{\left(\bar{\delta}_2-\delta_2\right)^2\right\}=\bar{\delta}_2\left(1-\bar{\delta}_2\right)\\ E\left\{\left(\bar{\delta}_1-\delta_1\right)\left(\bar{\delta}_2-\delta_2\right)\right\}=0\end{cases} \tag{9}$$

式中:$P\{\cdot\}$表示事件"$\cdot$"发生的概率;$E\{\cdot\}$表示变量"$\cdot$"的期望;$\bar{\delta}_1\in[0,1]$和$\bar{\delta}_2\in[0,1]$分别为随机变量δ_1和δ_2的已知常数预期值。

1.3 观测器和鲁棒控制策略设计

在设计鲁棒控制策略时,考虑到机器人在执行任务过程中状态量多、易变化等,在计算时采用相对状态量$\boldsymbol{x}(k)$(k时刻系统状态值-系统任务已规划的理论状态值)作为系统状态量代入式(1)中的线性模型。

由于实际系统中并非所有系统状态都是可测量的,选择状态观测器能够依据系统的控制模型与传感器的状态测量值比较准确的估计出系统的当前状态值,从而给出相应的系统控制输入$\boldsymbol{u}(k)$。状态观测器的形式描述如下:

$$\begin{cases}\bar{\boldsymbol{x}}(k+1)=\boldsymbol{A}\bar{\boldsymbol{x}}(k)+\boldsymbol{B}\boldsymbol{u}_{\mathrm{c}}(k)+\boldsymbol{G}(\boldsymbol{y}_{\mathrm{c}}(k)-\bar{\boldsymbol{y}}(k))\\ \bar{\boldsymbol{y}}(k)=(1-\bar{\delta}_1)\boldsymbol{C}\bar{\boldsymbol{x}}(k)+\bar{\delta}_1\boldsymbol{C}\bar{\boldsymbol{x}}(k-1)\end{cases} \tag{10}$$

式中:$\boldsymbol{G}\in\mathbf{R}^{n\times d}$为后续要计算的观测器增益。$\bar{\boldsymbol{x}}(k)$和$\bar{\boldsymbol{y}}(k)$分别为观测器的状态估计和状态测量估计值。

带观测器的移动机器人系统受DoS攻击模型如图2所示。

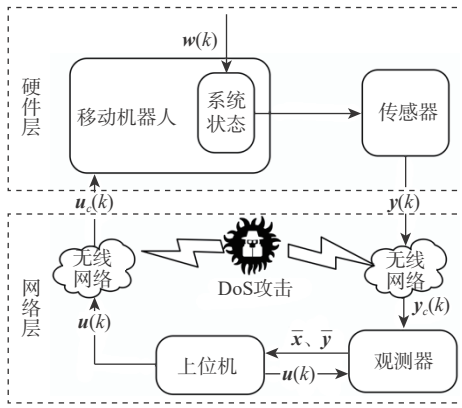


图2 含观测器的移动机器人受DoS攻击模型

Fig. 2 DoS attack model of mobile robot with observer

依据鲁棒控制理论,系统的鲁棒控制策略被设计为

$$\boldsymbol{u}(k)=\boldsymbol{K}\times\bar{\boldsymbol{x}}(k) \tag{11}$$

式中:$\boldsymbol{K}\in\mathbf{R}^{n\times m}$为控制增益矩阵。

将观测误差定义为

$$\boldsymbol{e}(k)=\boldsymbol{x}(k)-\bar{\boldsymbol{x}}(k) \tag{12}$$

结合式(1)、式(7)、式(10)~式(12),DoS攻击下的移动机器人系统被描述为

$$\begin{cases}\boldsymbol{x}(k+1)=[\boldsymbol{A}+(1-\delta_2)\boldsymbol{B}\boldsymbol{K}]\boldsymbol{x}(k)-(1-\delta_2)\boldsymbol{B}\boldsymbol{K}\boldsymbol{e}(k)+\\ \qquad\boldsymbol{D}\boldsymbol{w}(k)+\delta_2\boldsymbol{B}\boldsymbol{K}\boldsymbol{x}(k-1)-\delta_2\boldsymbol{B}\boldsymbol{K}\boldsymbol{e}(k-1)\\ \boldsymbol{e}(k+1)=[\boldsymbol{A}-(1-\bar{\delta}_1)\boldsymbol{G}\boldsymbol{C}]\boldsymbol{e}(k)-\bar{\delta}_1\boldsymbol{G}\boldsymbol{C}\boldsymbol{e}(k-1)+\\ \qquad\boldsymbol{D}\boldsymbol{w}(k)+(\delta_1-\bar{\delta}_1)\boldsymbol{G}\boldsymbol{C}\boldsymbol{x}(k)-(\delta_1-\bar{\delta}_1)\boldsymbol{G}\boldsymbol{C}\boldsymbol{x}(k-1)\end{cases} \tag{13}$$

定义$\boldsymbol{\eta}(k)=\mathrm{col}\left[\boldsymbol{x}^{\mathrm{T}}(k)\quad\boldsymbol{e}^{\mathrm{T}}(k)\quad\boldsymbol{x}^{\mathrm{T}}(k-1)\quad\boldsymbol{e}^{\mathrm{T}}(k-1)\right]$,那么式(13)可以描述为

$$\boldsymbol{\eta}(k+1)=\boldsymbol{A}_1\boldsymbol{\eta}(k)+\overline{\boldsymbol{D}}\boldsymbol{w}(k) \tag{14}$$

式中:

$$\overline{\boldsymbol{D}}=\left[\boldsymbol{D}\quad\boldsymbol{D}\quad\boldsymbol{0}\quad\boldsymbol{0}\right]^{\mathrm{T}}$$

$$\boldsymbol{A}_1=\begin{bmatrix}\boldsymbol{A}_1^{11}&\boldsymbol{A}_1^{12}&\boldsymbol{A}_1^{13}&-\boldsymbol{A}_1^{13}\\ \boldsymbol{A}_1^{21}&\boldsymbol{A}_1^{22}&-\boldsymbol{A}_1^{21}&\boldsymbol{A}_1^{24}\\ \boldsymbol{I}&\boldsymbol{0}&\boldsymbol{0}&\boldsymbol{0}\\ \boldsymbol{0}&\boldsymbol{I}&\boldsymbol{0}&\boldsymbol{0}\end{bmatrix}$$

$$\boldsymbol{A}_1^{11}=\boldsymbol{A}+(1-\delta_2)\boldsymbol{B}\boldsymbol{K}$$
$$\boldsymbol{A}_1^{12}=(1-\delta_2)\boldsymbol{B}\boldsymbol{K}$$
$$\boldsymbol{A}_1^{13}=\delta_2\boldsymbol{B}\boldsymbol{K}$$
$$\boldsymbol{A}_1^{21}=(\delta_1-\bar{\delta}_1)\boldsymbol{G}\boldsymbol{C}$$
$$\boldsymbol{A}_1^{22}=\boldsymbol{A}-(1-\bar{\delta}_1)\boldsymbol{G}\boldsymbol{C}$$
$$\boldsymbol{A}_1^{24}=-\bar{\delta}_1\boldsymbol{G}\boldsymbol{C}$$

2 增益矩阵计算与稳定性证明

应用H_∞理论研究移动机器人系统在受DoS攻击和不确定扰动下的充分稳定性条件,并基于稳定性条件求出1.3节中鲁棒控制增益矩阵。

使闭环系统式(14)在 1.3 节所提的控制输入下同时满足以下要求:

1)如果存在标量$\alpha>0,\tau\in(0,1),\boldsymbol{\eta}(1)\in\mathbf{R},\ k\in\mathbf{Z}_+$使得

$$E\left\{\left\|\boldsymbol{\eta}(k)\right\|^2\right\}\leqslant\alpha\tau^k E\left\{\left\|\boldsymbol{\eta}(1)\right\|^2\right\}$$

则该系统是指数均方稳定的, 并且$\boldsymbol{w}(k)=\mathbf{0}$。

2)在初始条件下, 对于任何$\boldsymbol{w}(k)\neq\mathbf{0}$, 如果输出$\boldsymbol{y}(k)$满足如下条件, 则系统式(14)具有$H_\infty$性能水平$\gamma$:

$$E\left\{\sum_{k=0}^{\infty}\|\boldsymbol{y}(k)\|^2\right\}\leqslant\gamma^2E\left\{\sum_{k=0}^{\infty}\|\boldsymbol{w}(k)\|^2\right\}$$

式中: γ为一个给定的非负标量, 也是干扰的衰减水平。

可得到如下定理。

定理 1　给定控制器增益矩阵$\boldsymbol{K}$和观测器增益矩阵$\boldsymbol{G}$, 如果干扰为$\mathbf{0}$, 存在满足以下条件的正定矩阵$\boldsymbol{P}_1$、$\boldsymbol{P}_2$、$\boldsymbol{S}_1$、$\boldsymbol{S}_2$满足:

$$\begin{bmatrix}\boldsymbol{\Delta}_1&\mathbf{0}&\mathbf{0}&\mathbf{0}&\boldsymbol{\Delta}_2&\mathbf{0}&\boldsymbol{\Delta}_3&\boldsymbol{\Delta}_4\\ \mathbf{0}&\boldsymbol{\Delta}_5&\mathbf{0}&\mathbf{0}&\boldsymbol{\Delta}_6&\boldsymbol{\Delta}_7&\boldsymbol{\Delta}_8&\mathbf{0}\\ \mathbf{0}&\mathbf{0}&\boldsymbol{\Delta}_9&\mathbf{0}&\boldsymbol{\Delta}_{10}&\mathbf{0}&\boldsymbol{\Delta}_{11}&\boldsymbol{\Delta}_{12}\\ \mathbf{0}&\mathbf{0}&\mathbf{0}&\boldsymbol{\Delta}_{13}&\boldsymbol{\Delta}_{14}&\boldsymbol{\Delta}_{15}&\boldsymbol{\Delta}_{16}&\mathbf{0}\\ \boldsymbol{\Delta}_2&\boldsymbol{\Delta}_6&\boldsymbol{\Delta}_{10}&\boldsymbol{\Delta}_{14}&\boldsymbol{\Delta}_{17}&\mathbf{0}&\mathbf{0}&\mathbf{0}\\ \mathbf{0}&\boldsymbol{\Delta}_7&\mathbf{0}&\boldsymbol{\Delta}_{15}&\mathbf{0}&\boldsymbol{\Delta}_{18}&\mathbf{0}&\mathbf{0}\\ \boldsymbol{\Delta}_3&\boldsymbol{\Delta}_8&\boldsymbol{\Delta}_{11}&\boldsymbol{\Delta}_{16}&\mathbf{0}&\mathbf{0}&\boldsymbol{\Delta}_{19}&\mathbf{0}\\ \boldsymbol{\Delta}_4&\mathbf{0}&\boldsymbol{\Delta}_{12}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\boldsymbol{\Delta}_{20}\end{bmatrix}<\mathbf{0}\tag{15}$$

式中: $\boldsymbol{\Delta}_1=\boldsymbol{P}_2-\boldsymbol{P}_1$; $\boldsymbol{\Delta}_2=\boldsymbol{A}+(1-\bar{\delta}_2)\boldsymbol{BK}$; $\boldsymbol{\Delta}_3=\boldsymbol{BK}$; $\boldsymbol{\Delta}_4=\boldsymbol{GC}$; $\boldsymbol{\Delta}_5=\boldsymbol{S}_2-\boldsymbol{S}_1$; $\boldsymbol{\Delta}_6=-(1-\bar{\delta}_2)\boldsymbol{BK}$; $\boldsymbol{\Delta}_7=\boldsymbol{A}+(1-\bar{\delta}_1)\boldsymbol{GC}$; $\boldsymbol{\Delta}_8=-\boldsymbol{BK}$; $\boldsymbol{\Delta}_9=-\boldsymbol{P}_2$; $\boldsymbol{\Delta}_{10}=\bar{\delta}_2\boldsymbol{BK}$; $\boldsymbol{\Delta}_{11}=-\boldsymbol{BK}$; $\boldsymbol{\Delta}_{12}=-\boldsymbol{GC}$; $\boldsymbol{\Delta}_{13}=-\boldsymbol{S}_2$; $\boldsymbol{\Delta}_{14}=-\bar{\delta}_2\boldsymbol{BK}$; $\boldsymbol{\Delta}_{15}=-\bar{\delta}\boldsymbol{GC}$; $\boldsymbol{\Delta}_{16}=\boldsymbol{BK}$, $\boldsymbol{\Delta}_{17}=-\boldsymbol{P}_1^{-1}$; $\boldsymbol{\Delta}_{18}=-\boldsymbol{S}_1^{-1}$; $\boldsymbol{\Delta}_{19}=-\left[\left(1-\bar{\delta}_2\right)\bar{\delta}_2\right]^{-1}\boldsymbol{P}_1^{-1}$; $\boldsymbol{\Delta}_{20}=-\left[\left(1-\bar{\delta}_1\right)\bar{\delta}_1\right]^{-1}\boldsymbol{S}_1^{-1}$。

那么闭环系统式(14)是指数均方稳定的。

证明　建立的李雅普诺夫函数为

$$\begin{aligned}V(\boldsymbol{\eta}(k))=&\boldsymbol{x}^{\mathrm{T}}(k)\boldsymbol{P}_1\boldsymbol{x}^{\mathrm{T}}(k)+\boldsymbol{x}^{\mathrm{T}}(k-1)\boldsymbol{P}_2\boldsymbol{x}^{\mathrm{T}}(k-1)+\\&\boldsymbol{e}^{\mathrm{T}}(k)\boldsymbol{S}_1\boldsymbol{e}(k)+\boldsymbol{e}^{\mathrm{T}}(k-1)\boldsymbol{S}_2\boldsymbol{e}(k-1)\end{aligned}\tag{16}$$

当干扰$\boldsymbol{w}(k)=\mathbf{0}$和任务期望为 0 时, 通过式(14)可得

$$\begin{aligned}&E\{V(\boldsymbol{\eta}(k+1))\,|\,\boldsymbol{x}(k)\cdots\boldsymbol{x}(0),\boldsymbol{e}(k)\cdots\boldsymbol{e}(0)\}-V(\boldsymbol{\eta}(k))=\\&\quad E\left\{\boldsymbol{x}^{\mathrm{T}}(k+1)\boldsymbol{P}_1\boldsymbol{x}^{\mathrm{T}}(k+1)+\boldsymbol{x}^{\mathrm{T}}(k)\boldsymbol{P}_2\boldsymbol{x}^{\mathrm{T}}(k)+\right.\\&\quad\left.\boldsymbol{e}^{\mathrm{T}}(k+1)\boldsymbol{S}_1\boldsymbol{e}(k+1)+\boldsymbol{e}^{\mathrm{T}}(k)\boldsymbol{S}_2\boldsymbol{e}(k)\right\}-\\&\quad\boldsymbol{x}^{\mathrm{T}}(k)\boldsymbol{P}_1\boldsymbol{x}^{\mathrm{T}}(k)-\boldsymbol{x}^{\mathrm{T}}(k-1)\boldsymbol{P}_2\boldsymbol{x}^{\mathrm{T}}(k-1)-\\&\quad\boldsymbol{e}^{\mathrm{T}}(k)\boldsymbol{S}_1\boldsymbol{e}(k)-\boldsymbol{e}^{\mathrm{T}}(k-1)\boldsymbol{S}_2\boldsymbol{e}(k)=\\&\quad\boldsymbol{\eta}^{\mathrm{T}}(k)\left(\boldsymbol{\Pi}_1+\boldsymbol{\Pi}_2^{\mathrm{T}}\boldsymbol{\Pi}_3\boldsymbol{\Pi}_2\right)\boldsymbol{\eta}(k)=\boldsymbol{\eta}^{\mathrm{T}}(k)\boldsymbol{\Omega}\boldsymbol{\eta}(k)\end{aligned}\tag{17}$$

式中:

$$\boldsymbol{\Pi}_1=\mathrm{diag}\{\boldsymbol{P}_2-\boldsymbol{P}_1,\boldsymbol{S}_2-\boldsymbol{S}_1,-\boldsymbol{P}_2,-\boldsymbol{S}_2\}$$

$$\boldsymbol{\Pi}_3=\mathrm{diag}\{\boldsymbol{P}_1,\boldsymbol{S}_1,(1-\bar{\delta}_2)\bar{\delta}_2\boldsymbol{P}_1,-(1-\bar{\delta}_1)\bar{\delta}_1\boldsymbol{S}_1\}$$

$$\boldsymbol{\Pi}_2=\begin{bmatrix}\boldsymbol{\Pi}_2^{11}&-(1-\bar{\delta}_2)\boldsymbol{BK}&\bar{\delta}_2\boldsymbol{BK}&-\bar{\delta}_2\boldsymbol{BK}\\ \mathbf{0}&\boldsymbol{A}-(1-\bar{\delta}_1)\boldsymbol{GC}&\mathbf{0}&-\bar{\delta}_1\boldsymbol{GC}\\ \boldsymbol{BK}&-\boldsymbol{BK}&-\boldsymbol{BK}&\boldsymbol{BK}\\ \boldsymbol{GC}&\mathbf{0}&-\boldsymbol{LC}&\mathbf{0}\end{bmatrix}$$

$$\boldsymbol{\Pi}_2^{11}=\boldsymbol{A}+(1-\bar{\delta}_2)\boldsymbol{BK}$$

由 Schur 补理, $\boldsymbol{\Omega}<\mathbf{0}$等价于式(15), 所以

$$\begin{aligned}&E\{V(\boldsymbol{\eta}(k+1))\,|\,\boldsymbol{x}(k)\cdots\boldsymbol{x}(0),\boldsymbol{e}(k)\cdots\boldsymbol{e}(0)\}-\\&\quad V(\boldsymbol{\eta}(k))<\lambda_{\min}\boldsymbol{\Omega}\boldsymbol{\eta}^{\mathrm{T}}(k)\boldsymbol{\eta}(k)<-\alpha\boldsymbol{\eta}^{\mathrm{T}}(k)\boldsymbol{\eta}(k)\end{aligned}\tag{18}$$

式中:

$$0<\alpha<\min\{\lambda_{\min}(-\boldsymbol{\Omega}),\theta\}$$

$$\theta=\max\{\lambda_{\max}(\boldsymbol{P}_1),\lambda_{\max}(\boldsymbol{P}_2),\lambda_{\max}(\boldsymbol{S}_1),\lambda_{\max}(\boldsymbol{S}_2)\}$$

通过计算式(18)可得

$$\begin{aligned}&E\{V(\boldsymbol{\eta}(k+1))\,|\,\boldsymbol{x}(k)\cdots\boldsymbol{x}(0),\boldsymbol{e}(k)\cdots\boldsymbol{e}(0)\}-\\&\quad V(\boldsymbol{\eta}(k))<\alpha\boldsymbol{\eta}^{\mathrm{T}}(k)\boldsymbol{\eta}(k)<-\frac{\alpha}{\theta}V(\boldsymbol{\eta}(k))=-\varphi V(\boldsymbol{\eta}(k))\end{aligned}\tag{19}$$

通过引理 1, 系统式(14)是指数均方稳定。

证毕

定理 2　考虑系统式(14)具有$\boldsymbol{w}(k)\neq\mathbf{0}$的情况, 对于给定的常数标量$\gamma$, 如果存在对称的正定矩阵$\boldsymbol{P}_1$、$\boldsymbol{P}_2$、$\boldsymbol{S}_1$、$\boldsymbol{S}_2$及$\boldsymbol{M}$、$\boldsymbol{N}$使以下不等式成立:

$$\begin{bmatrix}\tilde{\boldsymbol{\Delta}}_1&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\tilde{\boldsymbol{\Delta}}_2&\mathbf{0}&\tilde{\boldsymbol{\Delta}}_3&\tilde{\boldsymbol{\Delta}}_4&\tilde{\boldsymbol{\Delta}}_6\\ \mathbf{0}&\tilde{\boldsymbol{\Delta}}_6&\mathbf{0}&\mathbf{0}&\mathbf{0}&\tilde{\boldsymbol{\Delta}}_7&\tilde{\boldsymbol{\Delta}}_8&\tilde{\boldsymbol{\Delta}}_9&\mathbf{0}&\mathbf{0}\\ \mathbf{0}&\mathbf{0}&\tilde{\boldsymbol{\Delta}}_{10}&\mathbf{0}&\mathbf{0}&\tilde{\boldsymbol{\Delta}}_{11}&\mathbf{0}&\tilde{\boldsymbol{\Delta}}_{12}&\tilde{\boldsymbol{\Delta}}_{13}&\mathbf{0}\\ \mathbf{0}&\mathbf{0}&\mathbf{0}&\tilde{\boldsymbol{\Delta}}_{14}&\mathbf{0}&\tilde{\boldsymbol{\Delta}}_{15}&\tilde{\boldsymbol{\Delta}}_{16}&\tilde{\boldsymbol{\Delta}}_{17}&\mathbf{0}&\mathbf{0}\\ \mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\tilde{\boldsymbol{\Delta}}_{18}&\tilde{\boldsymbol{\Delta}}_{19}&\tilde{\boldsymbol{\Delta}}_{\boldsymbol{20}}&\mathbf{0}&\mathbf{0}&\mathbf{0}\\ \tilde{\boldsymbol{\Delta}}_2&\tilde{\boldsymbol{\Delta}}_7&\tilde{\boldsymbol{\Delta}}_{11}&\tilde{\boldsymbol{\Delta}}_{15}&\tilde{\boldsymbol{\Delta}}_{19}&\tilde{\boldsymbol{\Delta}}_{21}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}\\ \mathbf{0}&\tilde{\boldsymbol{\Delta}}_8&\mathbf{0}&\tilde{\boldsymbol{\Delta}}_{16}&\tilde{\boldsymbol{\Delta}}_{\boldsymbol{20}}&\mathbf{0}&\tilde{\boldsymbol{\Delta}}_{22}&\mathbf{0}&\mathbf{0}&\mathbf{0}\\ \tilde{\boldsymbol{\Delta}}_3&\tilde{\boldsymbol{\Delta}}_9&\tilde{\boldsymbol{\Delta}}_{12}&\tilde{\boldsymbol{\Delta}}_{17}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\tilde{\boldsymbol{\Delta}}_{23}&\mathbf{0}&\mathbf{0}\\ \tilde{\boldsymbol{\Delta}}_4&\mathbf{0}&\tilde{\boldsymbol{\Delta}}_{13}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\tilde{\boldsymbol{\Delta}}_{24}&\mathbf{0}\\ \tilde{\boldsymbol{\Delta}}_6&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\mathbf{0}&\tilde{\boldsymbol{\Delta}}_{25}\end{bmatrix}\tilde{\boldsymbol{\Delta}}_2<\mathbf{0}\tag{20}$$

则闭环系统式(14)是指数均方稳定的。式中: $\tilde{\boldsymbol{\Delta}}_1=\boldsymbol{P}_2-\boldsymbol{P}_1$; $\tilde{\boldsymbol{\Delta}}_2=\boldsymbol{P}_1\boldsymbol{A}+(1-\bar{\delta}_2)\boldsymbol{B}$; $\tilde{\boldsymbol{\Delta}}_3=\boldsymbol{B}$; $\tilde{\boldsymbol{\Delta}}_4=\boldsymbol{NC}$; $\tilde{\boldsymbol{\Delta}}_5=\boldsymbol{C}$; $\tilde{\boldsymbol{\Delta}}_6=\boldsymbol{S}_2-\boldsymbol{S}_1$; $\tilde{\boldsymbol{\Delta}}_7=-(1-\bar{\delta}_2)\boldsymbol{BM}$; $\tilde{\boldsymbol{\Delta}}_8=\boldsymbol{S}_1\boldsymbol{A}+(1-\bar{\delta}_1)\boldsymbol{NC}$; $\tilde{\boldsymbol{\Delta}}_9=\boldsymbol{BM}$; $\tilde{\boldsymbol{\Delta}}_{10}=-\boldsymbol{P}_2$; $\tilde{\boldsymbol{\Delta}}_{11}=\bar{\delta}_2\boldsymbol{BM}$; $\tilde{\boldsymbol{\Delta}}_{12}=-\boldsymbol{BM}$; $\tilde{\boldsymbol{\Delta}}_{13}=-\boldsymbol{NC}$; $\tilde{\boldsymbol{\Delta}}_{14}=-\boldsymbol{S}_2$; $\tilde{\boldsymbol{\Delta}}_{15}=-\bar{\delta}_2\boldsymbol{BM}$; $\tilde{\boldsymbol{\Delta}}_{16}=-\bar{\delta}\boldsymbol{NC}$; $\tilde{\boldsymbol{\Delta}}_{17}=\boldsymbol{BM}$; $\tilde{\boldsymbol{\Delta}}_{18}=-\gamma\boldsymbol{I}^2$; $\tilde{\boldsymbol{\Delta}}_{19}=\boldsymbol{P}_1\boldsymbol{D}$; $\tilde{\boldsymbol{\Delta}}_{20}=\boldsymbol{S}_1\boldsymbol{D}$; $\tilde{\boldsymbol{\Delta}}_{21}=-\boldsymbol{P}_1^{-1}$; $\tilde{\boldsymbol{\Delta}}_{22}=-\boldsymbol{S}_1^{-1}$; $\tilde{\boldsymbol{\Delta}}_{23}=-\left[(1-\bar{\delta}_2)\bar{\delta}_2\right]^{-1}\boldsymbol{P}_1^{-1}$; $\tilde{\boldsymbol{\Delta}}_{24}=-\left[(1-\bar{\delta}_1)\bar{\delta}_1\right]^{-1}\boldsymbol{S}_1^{-1}$; $\tilde{\boldsymbol{\Delta}}_{25}=-\boldsymbol{I}$; $\boldsymbol{BP}=\boldsymbol{P}_1\boldsymbol{B}$; $\boldsymbol{S}_1\boldsymbol{G}=\boldsymbol{N}$; $\boldsymbol{PK}=\boldsymbol{M}$。

证明　当$\boldsymbol{w}(k)\neq\mathbf{0}$时, 从式(14)～式(17), 可得

$$E\{V(\boldsymbol{\eta}(k+1))\}-E\{V(\boldsymbol{\eta}(k))\}+E\left\{\boldsymbol{y}^{\mathrm{T}}(k)\boldsymbol{y}(k)\right\}-\gamma^2E\left\{\boldsymbol{w}^{\mathrm{T}}(k)\boldsymbol{w}(k)\right\}=E\left\{\begin{bmatrix}\boldsymbol{\eta}(k)\\ \boldsymbol{w}(k)\end{bmatrix}^{\mathrm{T}}\cdot\begin{bmatrix}\boldsymbol{\Omega}+\overline{\boldsymbol{C}}^{\mathrm{T}}\overline{\boldsymbol{C}} & \boldsymbol{B}_2^{\mathrm{T}}\boldsymbol{P}_1\boldsymbol{B}_2+\boldsymbol{B}_3^{\mathrm{T}}\boldsymbol{S}_1\boldsymbol{B}_4\\ \boldsymbol{D}^{\mathrm{T}}\boldsymbol{P}_1\boldsymbol{B}_2+\boldsymbol{D}^{\mathrm{T}}\boldsymbol{S}_1\boldsymbol{B}_3 & \boldsymbol{D}^{\mathrm{T}}(\boldsymbol{P}_1+\boldsymbol{S}_1)\boldsymbol{D}-\gamma^2\boldsymbol{I}\end{bmatrix}\times\begin{bmatrix}\boldsymbol{\eta}(k)\\ \boldsymbol{w}(k)\end{bmatrix}\right\}=E\left\{\overline{\boldsymbol{\eta}}^{\mathrm{T}}(k)\boldsymbol{\Omega}\overline{\boldsymbol{\eta}}(k)\right\}\tag{21}$$

式中：$\overline{\boldsymbol{C}}=\begin{bmatrix}\boldsymbol{C} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\end{bmatrix}$；$\boldsymbol{B}_2=\begin{bmatrix}\boldsymbol{A}+(1-\overline{\delta}_2)\boldsymbol{BK} & -(1-\overline{\delta}_2)\boldsymbol{BK} & \overline{\delta}_2\boldsymbol{BK} & -\overline{\delta}_2\boldsymbol{BK}\end{bmatrix}$；$\boldsymbol{B}_3=\begin{bmatrix}\boldsymbol{0} & \boldsymbol{A}-(1-\overline{\delta}_1)\boldsymbol{GC} & \boldsymbol{0} & -\overline{\delta}_1\boldsymbol{GC}\end{bmatrix}$。

通过Schur补理，由不等式$\overline{\boldsymbol{\Omega}}<\boldsymbol{0}$可得

$$E\{V(\boldsymbol{\eta}(k+1))\}-E\{V(\boldsymbol{\eta}(k))\}+E\left\{\boldsymbol{y}^{\mathrm{T}}(k)\boldsymbol{y}(k)\right\}-\gamma^2E\left\{\boldsymbol{w}^{\mathrm{T}}(k)\boldsymbol{w}(k)\right\}<\boldsymbol{0}$$

将k推广到∞，可得

$$\sum_{k=0}^{\infty}E\{\|\boldsymbol{y}(k)\|^2\}+E\{V(\boldsymbol{\eta}(\infty))\}<\gamma^2\sum_{k=0}^{\infty}E\{\|\boldsymbol{w}(k)\|^2\}+E\{V(\boldsymbol{\eta}(0))\}\tag{22}$$

又因为$\boldsymbol{\eta}(0)=\boldsymbol{0}$，所以闭环系统式(14)是指数均方稳定的。

然而，式(19)由Schur补理获得的方程不是线性矩阵不等式(LMI)的形式，将其转化为LMI形式。

令$\boldsymbol{\Gamma}=\mathrm{diag}\{\boldsymbol{I},\boldsymbol{I},\boldsymbol{I},\boldsymbol{I},\boldsymbol{I},\boldsymbol{P}_1,\boldsymbol{S}_1,\boldsymbol{P}_1,\boldsymbol{S}_1,\boldsymbol{I}\}$,让不等式左右两边分别乘$\boldsymbol{\Gamma}$，可得式(18)。

证毕

定理3 考虑到DoS攻击下的系统式(14)中的非零不确定干扰$\boldsymbol{w}(k)\neq\boldsymbol{0}$，对于给定的标量$\gamma>0$，如果存在正定矩阵$\boldsymbol{P}_{11}$、$\boldsymbol{P}_{22}$、$\boldsymbol{S}_1$、$\boldsymbol{S}_2$(其中$\boldsymbol{P}_1=\boldsymbol{U}_1^{\mathrm{T}}\boldsymbol{P}_{11}\boldsymbol{U}_1+\boldsymbol{U}_2^{\mathrm{T}}\boldsymbol{P}_{22}\boldsymbol{U}_2$)使控制器增益和观测器增益矩阵满足以下要求：

$$\begin{cases}\boldsymbol{K}=\boldsymbol{V}\boldsymbol{\Sigma}^{-1}\boldsymbol{P}_{11}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{M}\\ \boldsymbol{G}=\boldsymbol{S}_1^{-1}\boldsymbol{N}\end{cases}\tag{23}$$

那么闭环系统式(14)是指数均方稳定的，并满足非零不确定需求的H_∞性能。

证明 通过引理3，得到一个非奇异矩阵$\boldsymbol{P}$满足$\boldsymbol{BP}=\boldsymbol{P}_1\boldsymbol{B}$，这样可得

$$\boldsymbol{P}_1\boldsymbol{B}=\boldsymbol{P}\boldsymbol{U}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{\Sigma}\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{V}^{\mathrm{T}}=\boldsymbol{U}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{\Sigma}\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{P}\tag{24}$$

式中：$\boldsymbol{U}=\mathrm{col}\begin{bmatrix}\boldsymbol{U}_1^{\mathrm{T}} & \boldsymbol{U}_2^{\mathrm{T}}\end{bmatrix}$。通过引理2和引理3，结合式(22)可得

$$\boldsymbol{U}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{P}_{11} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{P}_{22}\end{bmatrix}\boldsymbol{V}^{\mathrm{T}}=\boldsymbol{U}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{\Sigma}\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{P}\tag{25}$$

所以

$$\boldsymbol{P}^{-1}=\boldsymbol{V}\boldsymbol{\Sigma}^{-1}\boldsymbol{P}_{11}^{-1}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}\tag{26}$$

进而可得

$$\begin{cases}\boldsymbol{K}=\boldsymbol{V}\boldsymbol{\Sigma}^{-1}\boldsymbol{P}_{11}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathrm{T}}\boldsymbol{M}\\ \boldsymbol{G}=\boldsymbol{S}_1^{-1}\boldsymbol{N}\end{cases}\tag{27}$$

证毕

3 实验结果与分析

为验证本文鲁棒控制策略的可行性，采用了一个数值模拟实验作为实例。由于移动机器人在执行任务过程中最重要的一个状态量是位置量，在速度控制模式下为一阶线性系统，假设移动机器人在二维空间中的位置被视作系统状态。传感器测量的状态输出则是由视觉相机在固定空间内定位获取的机器人位置，具体表现为平面坐标系的坐标。而系统控制输入则代表移动机器人的速度。视觉相机捕获的位置信息通过网络传输至上位机，上位机在完成处理后，通过网络向机器人发送控制指令。

由于在执行任务过程中移动机器人是运动的，系统状态为相对位置(相对位置=测量的位置−理论规划的位置)，系统稳定的情况下相对状态为0。实验目标是通过上位机实现系统的鲁棒控制，使系统的相对位置保持在所需的设置。

考虑系统选择1 s为控制和采样周期，DoS攻击成功概率为$\overline{\delta}_1=0.5,\overline{\delta}_2=0.5$，$H_\infty$参数$\gamma=1$，并在时间为1 s时假设受到一个$|\boldsymbol{w}(k)|=0.7/k$($k$为时刻)大小的干扰。

移动机器人i在k时刻位置有二维空间2个方向的状态量，并且在移动机器人速度控制模式下，控制输入为二维空间2个方向的速度量，并设定传感器测量的位置量与实际位置量比例为3.1，干扰针对的是系统第2个方向位置量，故有系统的参数矩阵设计如下：

$$\boldsymbol{A}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix},\boldsymbol{B}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix},\boldsymbol{C}=\begin{bmatrix}3.1 & 0\\ 0 & 3.1\end{bmatrix},\boldsymbol{D}=\begin{bmatrix}0 & 1\end{bmatrix}。$$

此外，初始条件选择为

$$\boldsymbol{x}(0)=\begin{bmatrix}0.8 & 0\end{bmatrix}^{\mathrm{T}},\overline{\boldsymbol{x}}(0)=\begin{bmatrix}0 & 0\end{bmatrix}^{\mathrm{T}},\boldsymbol{u}(0)=\begin{bmatrix}0 & 0\end{bmatrix}^{\mathrm{T}}。$$

$\boldsymbol{x}(0)$中第1个状态分量不为0说明在一开始设定系统第1个状态就有偏差值，其他量为0说明系统一开始其他状态分量平衡，输入为0。

实验选用基于观测器的鲁棒控制策略作为文章基于H_∞的鲁棒控制的对比。基于观测器的鲁棒

控制策略依据经验设定增益矩阵为固定值:

$$\boldsymbol{K}^* = \begin{bmatrix} -0.5 & 0 \\ 0 & -0.5 \end{bmatrix}$$

系统控制输入为

$$\boldsymbol{u}(k) = \boldsymbol{K}^* \times \bar{\boldsymbol{x}} \tag{28}$$

其中:

$$\bar{\boldsymbol{x}}(k+1) = \boldsymbol{A}\bar{\boldsymbol{x}}(k) + \boldsymbol{B}\boldsymbol{u}_c(k) + \boldsymbol{G}^* \times (\boldsymbol{y}_c(k) - \bar{\boldsymbol{y}}(k))$$

$\boldsymbol{G}^*$的计算只需要满足 $\boldsymbol{A}-\boldsymbol{G}\boldsymbol{C}$ 的所有特征根具有负实部即可, 通过计算选取一组为

$$\boldsymbol{G}^* = \begin{bmatrix} 0.3 & 0 \\ 0 & 0.3 \end{bmatrix}$$

为验证本文鲁棒控制策略也能够适应无 DoS 攻击下的移动机器人系统稳定与安全控制, 设计了 2 组对比实验。

3.1　无 DoS 攻击下 2 个策略对比实验

第 1 组实验, 在不受 DoS 攻击环境下的鲁棒控制, 基于 H_∞的鲁棒策略计算增益矩阵如下:

$$\boldsymbol{K} = \begin{bmatrix} -0.2896 & 0 \\ 0 & -0.4388 \end{bmatrix}, \boldsymbol{G} = \begin{bmatrix} 0.3417 & 0 \\ 0 & 0.3649 \end{bmatrix}。$$

并将系统状态$\boldsymbol{x}(k)$、系统控制输入$\boldsymbol{u}_c(k)$随时间变化如图 3 所示, 图中, $\boldsymbol{G}$ 为基于观测器的鲁棒控制, $\boldsymbol{H}$ 为基于 H_∞与观测器结合的鲁棒控制。

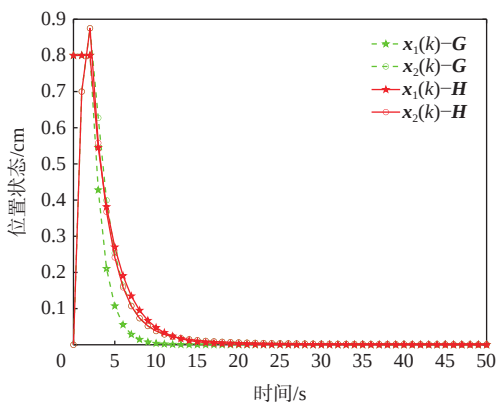


图 3　无 DoS 攻击下的系统状态

Fig. 3　System state without DoS attack

第 1 组对比实验展示了无 DoS 攻击 2 种策略的抗干扰能力。图 3 说明了在抑制干扰方面, 2 种策略让系统从 0 时刻的异常状态到稳定状态, 所用时间都是大约 15 s; 不同的是观测器的策略让系统 x_1先到稳定状态。图 4 说明了 2 种策略的系统控制输入(相对速度)也都在 k = 15 s时, 相对速度为 0, 达到要求的系统稳定。

3.2　DoS 攻击下 2 个策略对比实验

第 2 组实验, 在 DoS 攻击下的鲁棒控制, 通过基于 H_∞理论计算增益矩阵如下:

$$\boldsymbol{K} = \begin{bmatrix} -0.157\,3 & 0 \\ 0 & -0.185\,3 \end{bmatrix}, \boldsymbol{G} = \begin{bmatrix} 0.197\,8 & 0 \\ 0 & 0.149\,5 \end{bmatrix}。$$

DoS 攻击序列遵循伯努利分布, 选用了一组概率为 0.5 的序列如图 5 所示, 图中, 丢包值为 1 表示攻击成功, 0 表示攻击失败。

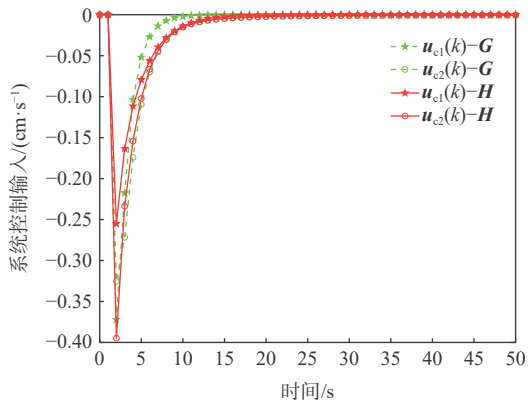


图 4　无 DoS 攻击的系统控制输入$\boldsymbol{u}(k)$

Fig. 4　System input $\boldsymbol{u}(k)$ without DoS attack

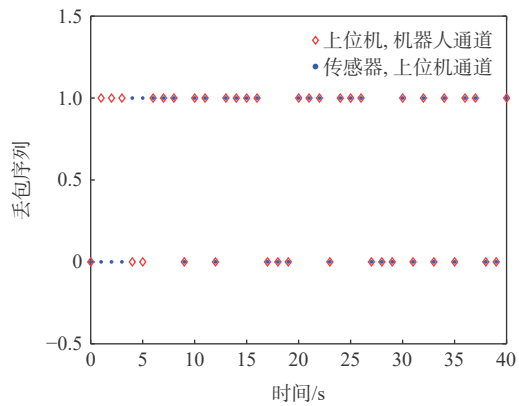


图 5　DoS 攻击分布

Fig. 5　DoS attack distribution

系统状态$\boldsymbol{x}(k)$、受攻击之后系统控制输入$\boldsymbol{u}_c(k)$随时间变化如图 6 所示。

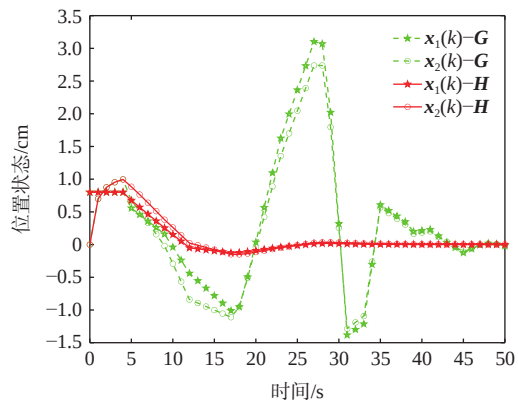


图 6　DoS 攻击下的系统状态

Fig. 6　System state under DoS attack

第 2 组对比实验说明了 DoS 攻击下 2 种鲁棒控制策略能够实现系统状态稳定, 并且展示了抗干

扰效果。图6直观地展现了，基于观测器的策略在传感数据丢失后，系统的状态$\boldsymbol{x}_1$、$\boldsymbol{x}_2$在$k=10$ s之后有一个振幅较大的衰减振荡；而基于H_∞的鲁棒控制让系统状态干扰平稳的抑制到0。从抑制干扰所需的时间角度看，后一种策略优于前一种。图7说明了后一种策略在抑制干扰的过程中速度变化小，而前一种策略让系统状态产生振荡，速度变化大，可能对移动机器人系统设备造成损害。

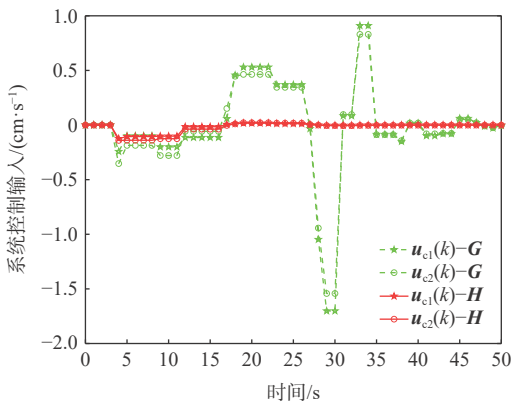


图7 DoS攻击的系统控制输入$\boldsymbol{u}_c(k)$

Fig. 7 System input $\boldsymbol{u}_c(k)$ for DoS attacks

实验结果验证了，基于H_∞鲁棒控制策略能够有效抵抗DoS攻击引起的数据丢失所导致的系统状态振荡，有效解决DoS攻击模型下移动机器人系统的稳定问题，使系统状态平稳收敛至稳定状态。本文提出的鲁棒策略虽然只是在二维空间一阶线性系统进行了仿真验证，但是可以推广到机器人系统多维状态变量的稳定控制，甚至可推广到高阶线性系统，能一定程度上解决移动机器人在接入互联网后遭受DoS攻击的控制安全问题。

4 结 论

1）在随机DoS攻击与系统不确定扰动共同作用下，本文研究的移动机器人闭环系统满足均方意义下的稳定性条件，证明了网络攻击环境中系统稳定性的可保证性。

2）通过引入鲁棒性能约束，系统在外部扰动输入存在时仍满足给定的H_∞性能指标，实现了扰动抑制与稳定控制的统一。

3）对比仿真结果验证了理论推导所得稳定性判据的有效性，系统在随机攻击发生情况下未出现发散现象，且状态响应具有良好的收敛特性。

4）研究结果说明，将随机网络攻击建模引入控制系统稳定性分析框架，可为网络化移动机器人安全控制问题提供可分析、可验证的理论依据。

参考文献（References）

[1] ZHU X Y, ZHANG H, CAO D P, et al. Robust control of integrated motor-transmission powertrain system over controller area network for automotive applications[J]. Mechanical Systems and Signal Processing, 2015, 58-59: 15-28.

[2] WEI H H, ZHANG Y T, SUN X M, et al. Intelligent robots and human-robot collaboration in the construction industry: a review[J]. Journal of Intelligent Construction, 2023, 1(1): 1-12.

[3] YIN T, LU N, GUO G S, et al. Knowledge and data dual-driven transfer network for industrial robot fault diagnosis[J]. Mechanical Systems and Signal Processing, 2023, 182: 109597.

[4] YAN Z, OUYANG B, LI D, et al. Network intelligence empowered industrial robot control in the F-RAN environment[J]. IEEE Wireless Communications, 2020, 27(2): 58-64.

[5] 张振国, 毛建旭, 谭浩然, 等. 重大装备制造多机器人任务分配与运动规划技术研究综述[J]. 自动化学报, 2024, 50(1): 21-41.
ZHANG Z G, MAO J X, TAN H R, et al. A review of task allocation and motion planning for multi-robot in major equipment manufacturing[J]. Acta Automatica Sinica, 2024, 50(1): 21-41(in Chinese).

[6] OLIVATO M, COTUGNO O, BRIGATO L, et al. A comparative analysis on the use of autoencoders for robot security anomaly detection[C]//Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2020: 984-989.

[7] CLARK G W, DORAN M V, ANDEL T R. Cybersecurity issues in robotics[C]//Proceedings of the IEEE Conference on Cognitive and Computational Aspects of Situation Management. Piscataway: IEEE Press, 2017: 1-5.

[8] MIT Technology Review. Spoofers use fake GPS signals to knock a yacht off course[EB/OL]. (2013-08-14)[2024-03-01]. https://www.technologyreview.com/2013/08/14/177015.

[9] MAGGI F, QUARTA D, POGLIANI M, et al. Rogue robots: testing the limits of an industrial robot's security[EB/OL]. (2017-05-03)[2024-03-01]. https://documents.trendmicro.com/assets/wp/wp-industrial-robot-security.pdf.

[10] SARGOLZAEI A, YAZDANI K, ABBASPOUR A, et al. Detection and mitigation of false data injection attacks in networked control systems[J]. IEEE Transactions on Industrial Informatics, 2020, 16(6): 4281-4292.

[11] WU C W, YAO W R, LUO W S, et al. A secure robot learning framework for cyber attack scheduling and countermeasure[J]. IEEE Transactions on Robotics, 2023, 39(5): 3722-3738.

[12] AMULLEN E M, SHETTY S, KEEL L H. Model-based resilient control for a multi-agent system against Denial of Service attacks[C]//Proceedings of the World Automation Congress. Piscataway: IEEE Press, 2016: 1-6.

[13] GUERRERO-HIGUERAS Á M, DECASTRO-GARCÍA N, MATELLÁN V. Detection of Cyber-attacks to indoor real time localization systems for autonomous robots[J]. Robotics and Autonomous Systems, 2018, 99: 75-83.

[14] PU H Y, HE L, CHENG P, et al. Security of industrial robots: vulnerabilities, attacks, and mitigations[J]. IEEE Network, 2023, 37(1):

111-117.

[15] ZHANG D, SHEN Y P, ZHOU S Q, et al. Distributed secure platoon control of connected vehicles subject to DoS attack: theory and application[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2021, 51(11): 7269-7278.

[16] CAI X, SHI K B, SHE K, et al. Performance degradation estimation mechanisms for networked control systems under DoS attacks and its application to autonomous ground vehicle[J]. IEEE Transactions on Cybernetics, 2024, 54(5): 2992-3002.

[17] TANG Y, ZHANG D D, HO D W C, et al. Event-based tracking control of mobile robot with denial-of-service attacks[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2020, 50(9): 3300-3310.

[18] 赵宁, 张会焱, 石碰. 拒绝服务攻击下网络化控制系统的基于观测器输出反馈控制[J]. 南京信息工程大学学报 (自然科学版), 2020, 12(3): 279-283.
ZHAO N, ZHANG H Y, SHI P. Observer-based output feedback control for networked control systems under denial-of-service attacks[J]. Journal of Nanjing University of Information Science & Technology (Natural Science Edition), 2020, 12(3): 279-283(in Chinese).

[19] GENG Q, ZHAO L, LI L, et al. A dynamic controller design for trajectory tracking control of wheeled mobile robot under stochastic denial of service attacks[J]. IEEE Transactions on Circuits and Systems II: Express Briefs, 2022, 69(8): 3560-3564.

[20] TARN T J, RASIS Y. Observers for nonlinear stochastic systems[J]. IEEE Transactions on Automatic Control, 1976, 21(4): 441-448.

[21] WANG Z D, YANG F W, HO D W C, et al. Robust H_{infty} control for networked systems with random packet losses[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 2007, 37(4): 916-924.

[22] WU J, CHEN T W. Design of networked control systems with packet dropouts[J]. IEEE Transactions on Automatic Control, 2007, 52(7): 1314-1319.

[23] SUH Y S. Stability and stabilization of nonuniform sampling systems[J]. Automatica, 2008, 44(12): 3222-3226.

[24] HALE M T, EGERSTEDTY M. Differentially private cloud-based multi-agent optimization with constraints[C]//Proceedings of the American Control Conference. Piscataway: IEEE Press, 2015: 1235-1240.

Robust control of mobile robots based on H_∞ under DoS attack

JI Xudong，CHEN Youdong*，WEI Hongxing

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

Abstract：Mobile robot systems are increasingly being integrated into various scenarios. However, connecting these systems to networks exposes them to the risk of cyber-attacks, potentially leading to functional failures and system destabilization. This paper focuses on the stability of mobile robot systems under denial-of-service (DoS) attacks and proposes a robust control strategy based on H_∞ control. In the strategy, the impact of DoS attacks on the system is first modeled as random packet loss with a Bernoulli distribution. When the mobile robot system is attacked, robust control is achieved by using feedback compensation based on a state observer. Sufficient conditions for exponential mean-square stability and H_∞ control of the closed-loop system are derived using Lyapunov stability theory. n order to acquire the observer and controller gain matrices, which allow for robust system control and anti-interference effects, these criteria are used to solve linear matrix inequality constraints. Finally, through simulation experiments, it is demonstrated that the proposed strategy effectively mitigates the impact of DoS attacks on the system, ensuring the stability of the system.

Keywords：mobile robot; DoS attack; H_∞ control; observer; robust control

Received：2024-04-29；**Accepted**：2024-07-05；**Published Online**：2024-07-26 09：15

URL：link.cnki.net/urlid/11.2625.V.20240725.1539.003

Foundation item：National Key Research and Development Program of China (2022YFB4702102)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0221

可重复使用运载器连接结构损伤容限分析

李配缘[1]，蔡巧言[2,*]，黎增山[3]，冯家赫[3]
(1. 北京宇航系统工程研究所，北京 100076；2. 中国运载火箭技术研究院，北京 100076；
3. 北京航空航天大学 航空科学与工程学院，北京 100191)

摘 要：针对可重复使用运载器冷结构中应用的复合材料蒙皮-金属骨架结构，开展了含初始损伤的金属-复材连接结构的损伤容限试验，并建立对应结构的损伤容限性能预测模型。研究结果表明：含孔边裂纹的金属结构在疲劳载荷作用下会发生裂纹扩展，而且由于连接不对称的影响，上下表面的裂纹扩展速率不一致，致使从外表面进行裂纹观测会造成对损伤尺寸的保守统计；含预置分层的复合材料层合板在挤压载荷作用下出现一定的分层扩展，其主要损伤还是由于复材的挤压疲劳破坏；复材孔边分层会急剧降低金属-复材连接结构的寿命；金属-复材连接结构的损伤容限设计应该重点关注复材孔边分层损伤；建立的金属-复材结构损伤容限性能预测模型预测结果与试验结果吻合良好，能很好的预测结构的损伤扩展过程和扩展寿命，可用于指导该类结构的损伤容限设计。

关 键 词：可重复使用运载器；结构；损伤容限；有限元法；损伤检测
中图分类号：V415.5; TB333
文献标志码：A **文章编号**：1001-5965(2026)06-1990-10

可重复使用运载器任务剖面多样，飞行中载荷环境复杂，结构形式特殊。由于可重复使用性的技术要求，结构设计时须在传统航天运载器静强度、刚度和稳定性设计的基础上，在主结构设计中引入损伤容限的概念，以保证结构系统安全可靠。可重复使用运载器部分冷结构采用框梁蒙皮连接，复合材料蒙皮与金属框梁通过紧固件连接。由于复合材料整体壁板成型及钻孔过程中可能导致现有无损检测技术难以捕捉的分层，在连接孔周围局部挤压过程中分层逐渐扩展，降低结构承载能力。金属孔边在制孔及使用过程中易形成微小疲劳源逐渐扩展成裂纹，对结构完整性造成破坏。

对于单一的金属材料或复合材料的损伤容限，许多学者已经做了大量的研究工作。目前对于金属损伤容限的研究多集中在对等先进合金[1-3]和复杂结构裂纹扩展的分析方法[4]等领域。复材方面主要从各尺度对含损结构分层扩展特性[5-6]、疲劳特性[7]及材料损伤后的剩余性能研究[8-9]，并建立多层级积木式试验体系，考察工艺分散性、环境、结构不连续、损伤和特定的设计工艺细节问题。苏玲等[10]将含孔的复合材料层合板模拟受冲击后的结构，进行剩余强度分析，并将该种损伤容限分析方法应用于工程实例。杨钧超等[11]基于改进的破坏行为准则建立了包含凹坑、分层损伤边界等几何信息的复合材料加筋板渐进损伤有限元分析模型，用于预测屈曲模态、失效过程及破坏模式。

对于金属-复合材料混合连接结构，目前研究主要集中在连接与载荷分配[12-13]、疲劳寿命、刚度匹配[14]、热载荷[15]、工艺因素[16-17]引起的局部应力分配影响等领域的研究。目前认为金属材料-复合材料的混合结构中，结构的静强度取决于复合材料，疲劳寿命则取决于金属结构。由于混合结构中复合材料结构与金属结构具有不同的疲劳特性，难以在同一个试验件上同时验证这2种材料结构的耐

收稿日期：2024-04-16；**录用日期**：2024-06-21；**网络出版时间**：2024-06-27 10:16
网络出版地址：link.cnki.net/urlid/11.2625.V.20240626.1114.001
基金项目：中国运载火箭技术研究院自主研发重点项目资金支持
*通信作者. E-mail：aseexplorer@163.com

引用格式：李配缘，蔡巧言，黎增山，等. 可重复使用运载器连接结构损伤容限分析[J]. 北京航空航天大学学报，2026，52(6)：1990-1999. LI P Y，CAI Q Y，LI Z S，et al. Damage tolerance analysis on reusable launch vehicle connection structures[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：1990-1999 (in Chinese).

久性, 为此可选用下述方法之一验证[18]: ①采用全尺寸试验以下的积木式验证方法, 对复合材料结构部分进行耐久性/损伤容限验证, 而全尺寸结构试验件则用于证实金属结构的耐久性/损伤容限; ②采用 2 个相同的全尺寸结构试验件, 分别证实金属结构和复合材料结构的耐久性/损伤容限。

静载和疲劳载荷作用下, 混合结构中复材层合板与金属板的失效顺序和失效模式会发生变化。对于混合结构中复材和金属的疲劳损伤耦合问题, 侯赤等[19]认为混合结构多钉连接中存在局部损伤与刚度的耦合作用, 从而影响层合板疲劳损伤规律, 在混合结构多钉连接疲劳分析中不能将金属和层合板独立计算, 必须考虑二者疲劳损伤的耦合才能保证分析结果的正确性。

对于可重复使用航天、高空高速低频次使用飞行器的损伤容限设计研究, 国内外相关研究较少且尚未形成统一方法及标准, 目前已有成果多集中在针对防隔热材料损伤容限的研究[20]。由于可重复使用航天器结构服役主要是承受高低温交变, 重复加载, 振动及疲劳载荷, 因此, 渐进损伤是结构性能演化和衰变的根本原因。SpaceX 的技术人员根据检修维护的频率和程度将全箭检测分为 3 个等级, 并针对复杂的大型复合材料空间结构提出一种实用损伤容限设计方法[21]: ①使用经试验验证的分析技术确定可接受的缺陷尺寸, 同时确保这些缺陷大于与所选无损评估 (nondestructive examination, NDE) 技术相对应的最小可检测缺陷尺寸; ②实施验证测试程序, 以验证工艺; ③使用严格的接受/拒绝标准进行预验证和后验证无损评估。

综上所述, 目前针对可重复使用运载器含损伤的金属-复材连接结构的损伤容限特性研究尚没有公开报道。本文针对可重复使用运载器广泛采用金属骨架-复材蒙皮结构, 分别开展了含孔边裂纹 (金属) 和孔边分层 (复材) 2 类损伤的金属-复材双钉横向连接结构的损伤容限特性试验, 并建立了对应的损伤容限特性预测模型。

1　试　验

1.1　试验件

本文采用的试验件由 7050-T7451 铝合金板件和复合材料层合板连接组成, 如图 1 所示。复合材料层合板选用国产 T800 级碳纤维斜纹织物/603A 预浸料铺成, 单层厚度为 0.2 mm。以试验件纵向拉伸方向为 0°, 铺层顺序为 $[0°/90°/0°/90°/0°/\overline{90°}]_S$, 共 11 层, 名义厚度为 2.2 mm, 材料力学性能如表 1 所示, 表中, E_{11} 和 E_{22} 分别为纵向和横向拉伸模量, ν_{12} 为泊松比, X_T 和 X_C 分别为纵向拉伸和压缩强度, Y_T 和 Y_C 分别为横向拉伸和压缩强度, S_{12} 为横向剪切强度。

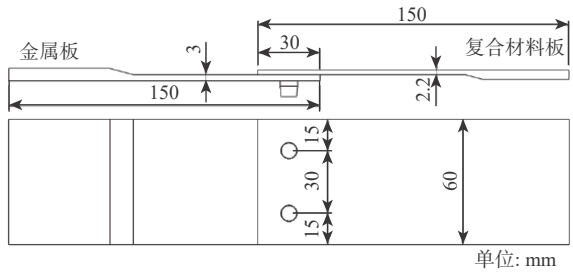


图 1　双钉横向连接试验件及主要尺寸

Fig. 1　Double-nail transverse connection test piece and main dimensions

表 1　TG800-6K-xw/603A 复合材料力学性能

Table 1　Mechanical properties of TG800-6K-xw/603A composite materials

E_{11} /GPa	E_{22} /GPa	ν_{12}	X_T /MPa	X_C /MPa	Y_T /MPa	Y_C /MPa	S_{12} /MPa
87	83.7	0.08	1 158	720	1 046	707	81.4

紧固件选用直径为 5 mm 的 100°沉头抗剪型钛合金高锁螺栓 (GJB 3 378-1998/2), 接头附近结构形式如图 2 所示。以与紧固件齐平一面为上表面, 以与金属板件贴合一面为下表面。

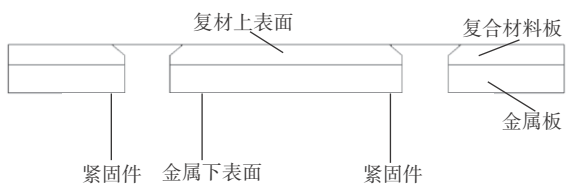


图 2　接头结构剖面

Fig. 2　Cross-section of joint structure

试验件总共包括 3 种: ①在复合材料接头区域布置层间预置分层损伤; ②在金属一个连接孔处设置单边孔边径向裂纹; ③复合材料和金属均没有损伤。

复合材料板件内的预置分层区域如图 3 所示, 在连接孔周围区域布置有 20 mm×20 mm 的预置分层, 预置分层通过在铺层时夹入聚四氟乙烯薄膜实现。分层布置中性面, 即约 1/2 厚度处, 第 6 层和第 7 层之间。由于实际生产工艺误差, 预置分层区域与理论位置有部分偏差。孔边初始裂纹尖端距离孔边总长度为 3 mm, 如图 4 所示。

1.2　试验程序

试验加载采用 MTS-370 型液压静力-疲劳试验机进行加载, 试验机精度为 1 级, 最大量程为 500 kN。试验程序如下:

1) 首先, 采用不含损伤的试验件进行静力试验, 确定静力破坏载荷。

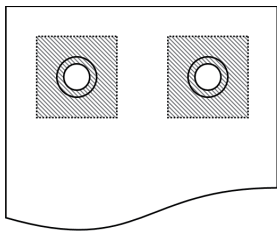


图3 预置分层范围示意图(阴影部分为预置分层区域)

Fig. 3 Schematic diagram of preset delamination damage (rangewith shaded areas indicating preset stratification regions)

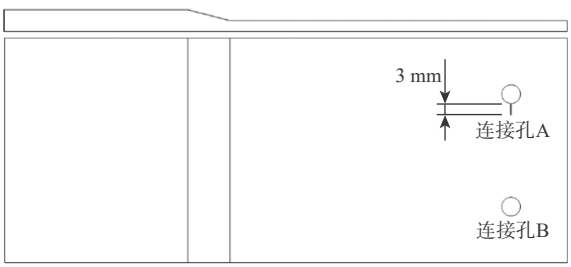


图4 孔边裂纹示意图

Fig. 4 Schematic diagram of metal plate with preset cracks

2) 根据静力试验结果,初步选取50%进行含损伤连接件的疲劳寿命试验(应力比为0.1),如果寿命大于2.0×10^5次,则增大试验载荷,如果寿命小于5 000次,则减小试验载荷;在5 000~2.0×10^5次之间设置3个不同量级的载荷进行试验,并进行损伤扩展观测,其中,针对含分层损伤试验件采用超声C扫的方式进行损伤观测;含金属裂纹试验件采用裂纹观测仪进行裂纹长度观测。

1.3 试验结果

1.3.1 静力试验结果

静力试验结果如表2所示。

表2 静强度试验结果

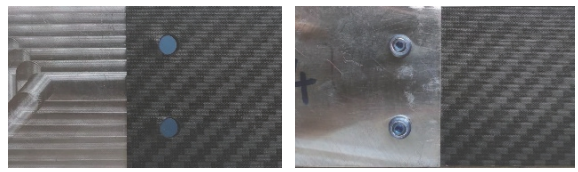
Table 2 Static strength test results

第1件破坏载荷/kN	第2件破坏载荷/kN	第3件破坏载荷/kN	平均值/kN	离散系数/%
15.09	14.92	14.84	14.95	0.70

破坏如图5所示。静强度试验中的失效模式主要是由于紧固件对复合材料层合板孔边的挤压破坏引起的。随着拉载的增大,复合材料层合板孔边受压一侧在局部压力的作用下发生纤维与基体压缩破坏,钉头沿着锪窝斜面发生滑动,楔进孔边,造成掉载,此时金属连接孔附近区域未观察到损伤产生。

1.3.2 含预置分层试验结果

经过载荷调试,最终确定疲劳试验载荷如表3所示,表中,σ_b为静强度试验破坏强度。剩余寿命试验结果如表4所示,N_1~N_6为不同实验件的发生破坏时的试验加载次数。损伤扩展观测结果表明,直至最终破坏,损伤扩展并不明显,损伤观测结果



(a) 连接失效处正面 (b) 连接失效处背面

图5 试验件静强度拉伸失效照片

Fig. 5 Static strength tensile failure of test piece

如图6所示,试验件最终失效模式为复材板孔边挤压破坏,紧固件滑出。

表3 含孔边预置分层试验件加载

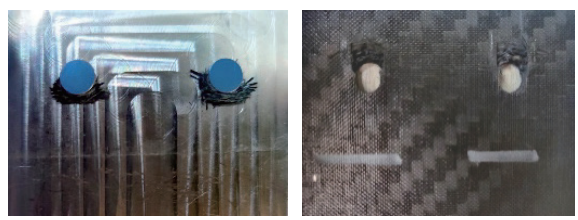
Table 3 Loading of test pieces with preset composite partial delamination

组别	载荷水平	峰值载荷/kN	应力比	频率/Hz	试件数量
第1组	45%σ_b	6.7	0.1	6	5
第2组	50%σ_b	7.5	0.1	6	5
第3组	55%σ_b	8.2	0.1	6	5

表4 含孔边预置分层复材件剩余寿命试验结果

Table 4 Test results of test pieces with preset delamination in composite materials

组别	N_1	N_2	N_3	N_4	N_5	N_6	试验寿命平均值	对数寿命离散系数/%
第1组	175 333	88 750	220 792	24 718	265 513	50 125	137 539	7.37
第2组	34 698	1 013	19 218	16 175	5 046	27 417	17 261	13.10
第3组	2 154	4 126	2 702	15 263	12 079	4 678	6 834	8.45



(a) 连接失效处(金属板) (b) 连接失效处(复材板)

图6 含孔边分层试验件疲劳破坏照片

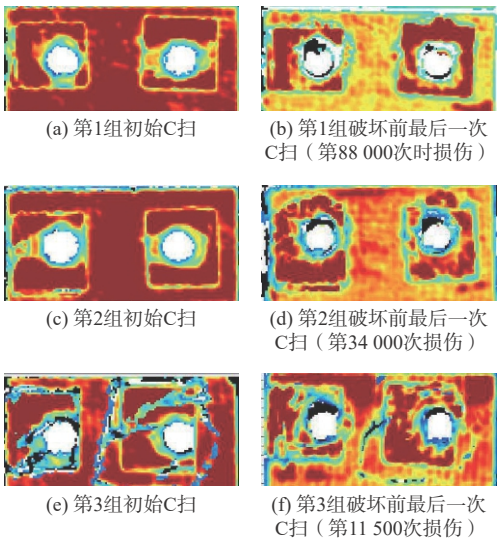
Fig. 6 Fatigue failure of test pieces containing delamination at the edge of the hole

试验件起始和破坏前最后一次超声C扫图像如图7所示。

1.3.3 含孔边裂纹试验结果

经过载荷调试,最终确定疲劳试验载荷如表5所示,剩余寿命试验结果如表6所示,N_1~N_5为不同实验件发生破坏时的试验加载次数。

在进行裂纹扩展观测时,75%σ_b和80%σ_b载荷水平加载的试验件每4 000次循环测量一次裂纹长度,83%σ_b载荷加载的试验件每隔2 000次循环测量一次裂纹长度,并进行数据记录与处理。各组试样的$\Delta a-\lg N$测量数据如图8所示。



(a) 第1组初始C扫　(b) 第1组破坏前最后一次C扫（第88 000次时损伤）

(c) 第2组初始C扫　(d) 第2组破坏前最后一次C扫（第34 000次损伤）

(e) 第3组初始C扫　(f) 第3组破坏前最后一次C扫（第11 500次损伤）

图 7　双钉横向连接试验件超声 C 扫描图像

Fig. 7　Ultrasonic C-scan image of double-nail transverse connection test piece

表 5　含预置裂纹金属疲劳裂纹扩展试验加载方案

Table 5　Loading scheme for fatigue crack growth test of metal with preset cracks

组别	载荷水平	峰值载荷/kN	裂纹截面应力/MPa	应力比	频率/Hz	有效数量
第1组	75%σ_b	11.2	79.43	0.1	5	5
第2组	80%σ_b	12.0	85.11	0.1	5	5
第3组	83%σ_b	12.4	87.94	0.1	5	5

表 6　含预置裂纹金属件剩余寿命试验结果

Table 6　Test result of remaining life of metal parts with pre-cracks

组别	N_1	N_2	N_3	N_4	N_5	试验寿命平均值	对数寿命离散系数/%
第1组	37 356	77 089	78 330	39 236	76 578	61 718	3.14
第2组	18 821	18 209	57 063	44 236	58 264	53 188	1.15
第3组	11 655	14 859	23 655	24 372	30 201	23 272	2.58

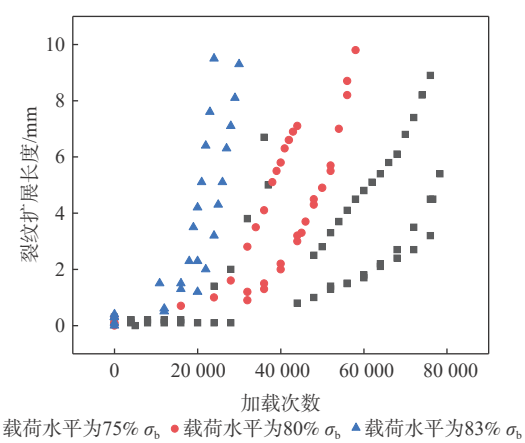
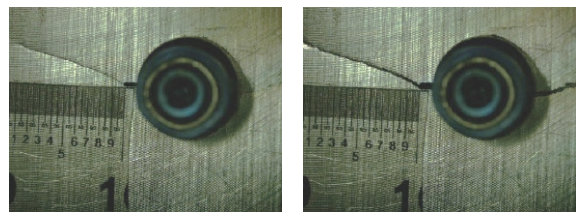


图 8　试验测量数据点

Fig. 8　Test measurement data points

从试验中发现，金属裂纹扩展方向总体指向另外一个连接孔，当裂纹达到一定长度时，试验件突然断裂，如图 9 所示。



(a) 76 000次加载　(b) 77 089次加载（结构破坏）

图 9　连接孔边金属裂纹的扩展

Fig. 9　Propagation at the edge of metal cracks

2　仿　真

2.1　含孔边裂纹试验件有限元仿真计算

本节计算是基于断裂力学及扩展有限元法，采用商用有限元分析软件 ABAQUS 低周疲劳分析步完成模拟。建立的有限元模型如图 10 所示，边界条件为复材层板端部固支，金属端头限制轴向之外的位移，并施加拉-拉循环载荷。紧固件与复材孔、与金属连接受压的半侧接触为 tie 约束，其余均为默认接触约束。单元为 8 节点减缩积分单元（C3D8R）。

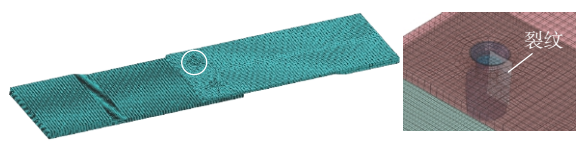


图 10　有限元分析模型

Fig. 10　Finite element analysis model

ABAQUS 中定义裂纹萌生的条件与扩展行为，以及断裂准则等可参见文献 [22-24]。

2.1.1　模型参数确定

对于双钉混合连接试验件的结构不对称性，单个连接孔处由于单剪螺栓倾斜引起的弯曲效应，导致金属板厚度方向产生相对应力梯度，进而导致金属板上表面扩展速率大于下表面。上表面与复材贴合，无法直接观测得到，但是主导了裂纹扩展行为。采用目视检测裂纹长度时，容易造成对损伤尺寸的保守估计。近似认为上表面扩展速率是下表面的 2 倍，有

$$\frac{\mathrm{d}a}{\mathrm{d}N}=2\frac{a_{i+1}-a_i}{N_{i+1}-N_i}\qquad i=1,2,\cdots,n-1 \tag{1}$$

式中：a 为裂纹长度；N 为载荷循环加载次数。

平面应力状态下，有

$$\Delta G=\left(K_{\max}^2-K_{\min}^2\right)/E \tag{2}$$

式中：$K_{\max}$ 和 $K_{\min}$ 分别为每次循环中峰值载荷和

谷值载荷对应的应力强度因子值；E为弹性模量。通过对裂纹开始发生时的次数N与每次加载中的应变能释放率之差ΔG拟合，可得断裂表征参数c_1和c_2值；对于断裂表征参数c_3和c_4值，采用独立试验中7 050-T7451铝合金平板在拉-拉试验中的da/dN-ΔG数据拟合。数据点及拟合结果如图11和图12所示。

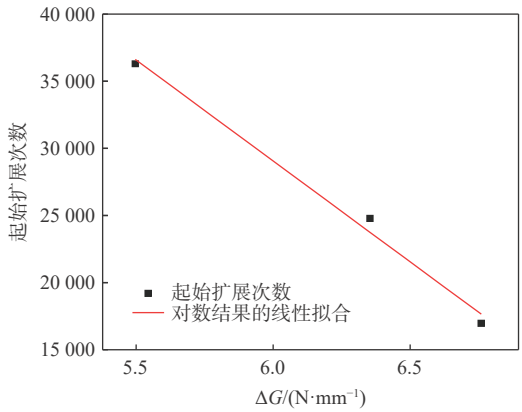


图11 金属裂纹起始扩展次数N-ΔG拟合结果

Fig. 11 Fitting results of crack initial expansion times N and ΔG

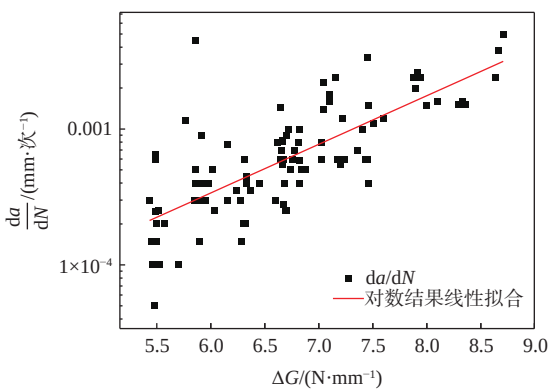


图12 金属裂纹da/dN-ΔG数据点及拟合结果

Fig. 12 da/dN-ΔG data points and results of crack

通过本节计算，确定的模型断裂参数如表7所示，表中，$G_{\text{thresh}}/G_{\text{C}}$和$G_{\text{pl}}/G_{\text{C}}$分别为Paris公式中使用的能量释放率阈值与等效临界能量释放率的比值和能量释放率上限与等效临界能量释放率的比值，默认值分别为0.01和0.85，a_{m}、a_{n}、a_0为指数，本处取默认值1。

表7 材料断裂表征参数

Table 7 Material fracture performance parameters

c_1	c_2	c_3	c_4	$G_{\text{C}}/(\text{N}\cdot\text{mm}^{-1})$	$\frac{G_{\text{thresh}}}{G_{\text{C}}}$	$\frac{G_{\text{pl}}}{G_{\text{C}}}$
2.7×10^{8}	−5	1.1×10^{-7}	4.25	9	0.01	0.85

2.1.2 结果对比

各组仿真结果与试验结果统计与对比如表8和图13所示。

表8 仿真结果对比

Table 8 Comparison of simulation results

组别	试验寿命平均值	有限元计算值	仿真与试验的相对误差/%
第1组	61 718	69 135	12.0
第2组	53 188	45 334	−14.8
第3组	23 272	28 066	20.6

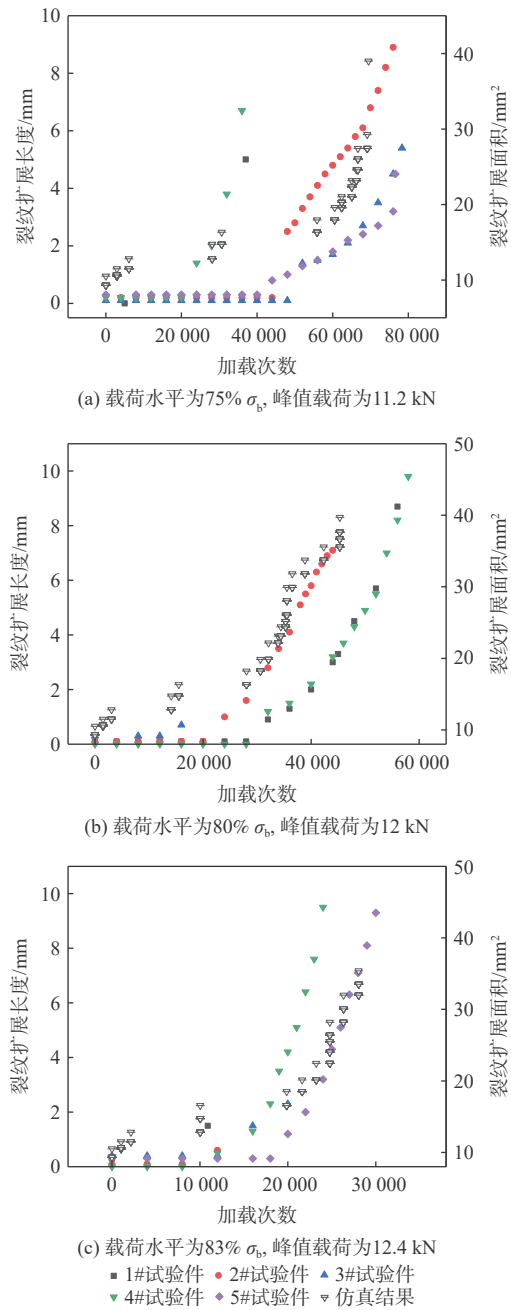


图13 试验及仿真结果

Fig. 13 Test and simulation results

2.2 含孔边分层试验件有限元仿真计算

建立的有限元模型如图14所示，边界条件为复材层板端部固支，金属端头限制轴向之外的位移，并施加拉-拉循环载荷。各独立部件均接触。

采用 ABAQUS 自带的 cohesive 单元模拟分层扩展。在建模时, 分层位置与实际位置一致。复材和金属部分均为 8 节点减缩积分单元(C3D8R)。

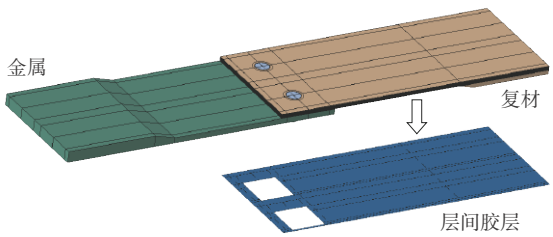


图 14　复材疲劳寿命分析有限元模型与界面单元

Fig. 14　Finite element model and interface element for composite fatigue life analysis

复合材料的失效涉及到层内失效及层间失效, 因此, 在进行有限元模拟时分别针对层内失效和层间失效分别采用不同的失效准则。

2.2.1　层内损伤准则

复合材料层内损伤使用根据三维 Hashin 静力失效准则[25]改进的疲劳失效准则。层内损伤的材料性能退化参考 Camanho[26]给出的材料退化方法, 具体如表 9 所示, 表中, E_{ij} (i,j =1,2,3) 为拉伸模量, G_{ij} (i,j =1,2,3) 为剪切模量。

表 9　层内损伤退化准则

Table 9　Intralayer damage degradation criterion

失效模式	材料退化准则
经向拉伸/剪切失效	E_{11},G_{12},G_{13}退化到初始值的0.01
经向压缩失效	E_{11},G_{12},G_{13}退化到初始值的0.01
纬向拉伸/剪切失效	E_{22},G_{13},G_{23}退化到初始值的0.01
纬向压缩失效	E_{22},G_{13},G_{23}退化到初始值的0.01

2.2.2　层间损伤准则

针对分层损伤的考虑是基于双线性内聚力本构关系, 考虑疲劳加载对损伤累积过程的影响[27], 层间强度随着疲劳进程逐渐衰减。层间性能属性如表 10 所示, 表中, T_n、T_s、T_t 分别为界面承受的法向应力和 2 个切向应力, $G_{\mathrm{I C}}$、$G_{\mathrm{II C}}$、$G_{\mathrm{III C}}$ 分别为界面单元 Ⅰ 型、Ⅱ 型、Ⅲ 型裂纹对应的临界应变能释放率。

表 10　层间界面单元属性

Table 10　Interlayer interface element properties

T_n/MPa	T_s/MPa	T_t/MPa	$G_{\mathrm{I C}}$/(N·mm^{-1})	$G_{\mathrm{II C}}$/(N·mm^{-1})	$G_{\mathrm{III C}}$/(N·mm^{-1})
81	81	60	0.2	1.0	1.0

在单元未发生损伤时, 对复合材料层内及层间的强度性能进行逐步退化, 在单元发生损伤后, 对复合材料层内及层间的刚度性能进行直接折减, 因此, 结构的承载能力不断下降。当结构整体强度降低至疲劳载荷谱的最大疲劳应力水平 σ_{max} 以下, 就发生了结构整体失效, 此时结构承载能力突然下降, 通过有限元模型输出的加载端载荷/位移-时间曲线的突变作为结构失效的判定准则, 即认为此时的加载次数 n 为结构的疲劳寿命 N_f。

复合材料结构疲劳有限元分析流程如图 15 所示。

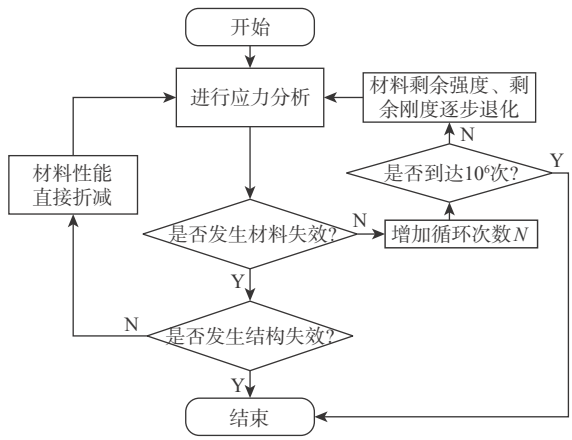


图 15　有限元分析流程

Fig. 15　Finite element analysis flow

复合材料的性能折减通过编写用户自定义场变量子程序 VUSDFLD 实现。

2.2.3　仿真结果分析

各组载荷作用下, 预置分层损伤的复材-金属连接试验件仿真结果和试验结果对比情况如表 11 所示。仿真计算得到的结构寿命与试验平均值吻合程度良好。

表 11　仿真结果对比

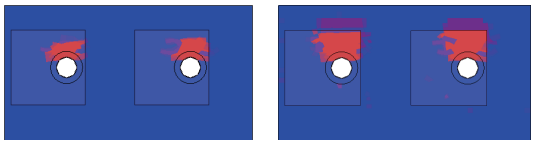
Table 11　Comparison of simulation results

组别	试验寿命平均值	有限元计算值	仿真与试验的相对误差/%
第1组	137 539	131 825	−4.15
第2组	17 261	17 378	0.70
第3组	6 834	6 309	−7.68

从损伤模拟结果来看, 复材 2 个钉孔的的挤压区域的损伤水平相当, 均为在挤压载荷作用下的挤压应力导致的疲劳损伤破坏, 同时层间也出现了一些损伤起始, 但是没有完全损伤。层内和层间损伤的扩展结果如图 16 和图 17 所示。

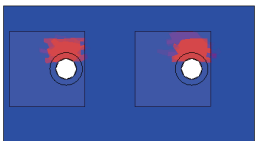
如图 17 所示, 在疲劳载荷加载过程中出现了一定程度的预置分层的层间损伤扩展行为。主要表现为在局部压载作用下向着压向逐渐扩展,

由于挤压应力分布导致有扩展成圆形分层区域的趋势。

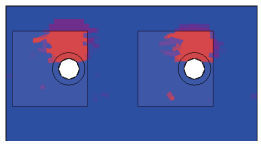


(a) 45% σ_b (6.7 kN)加载下的损伤扩展（第2 200次）

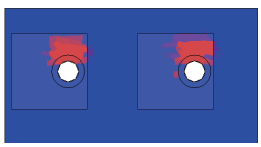
(b) 45% σ_b (6.7 kN)加载下的损伤扩展（第131 825次，发生破坏）



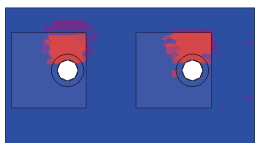
(c) 50% σ_b (7.5 kN)加载下的损伤扩展（第2 200次）



(d) 50% σ_b (7.5 kN)加载下的损伤扩展（第17 378次，发生破坏）



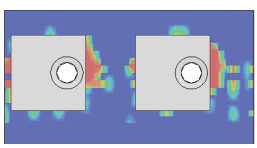
(e) 55% σ_b (8.2 kN)加载下的损伤扩展（第2 200次）



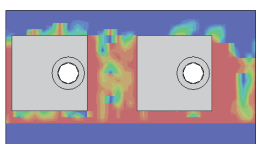
(f) 55% σ_b (8.2 kN)加载下的损伤扩展（第6 309次，发生破坏）

图 16　损伤扩展仿真结果

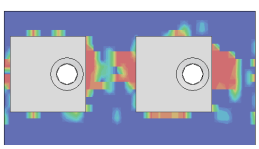
Fig. 16　Damage propagation simulation results



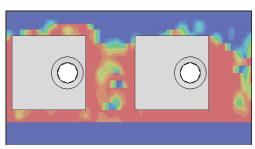
(a) 45% σ_b (6.7 kN)加载下的损伤扩展（第2 200次）



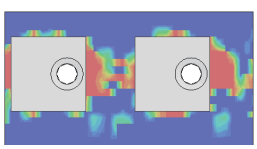
(b) 45% σ_b (6.7 kN)加载下的损伤扩展（第131 825次，发生破坏）



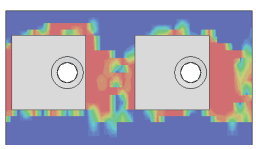
(c) 50% σ_b (7.5 kN)加载下的损伤扩展（第2 200次）



(d) 50% σ_b (7.5 kN)加载下的损伤扩展（第17 378次，发生破坏）



(e) 55% σ_b (8.2 kN)加载下的损伤扩展（第2 200次）



(f) 55% σ_b (8.2 kN)加载下的损伤扩展（第6 309次，发生破坏）

图 17　层间损伤的仿真结果

Fig. 17　Simulation results of interlayer damage

3　随机谱下的含损伤连接结构剩余寿命分析

本节采用可重复使用运载器 2 种典型工况下的载荷谱进行含损伤连接结构的损伤扩展和剩余寿命进行分析，如图 18 所示。

1）含金属孔边裂纹金属-复材连接结构。分别计算 2 个工况下含 3 mm 和 8 mm 这 2 种长度的初始裂纹结构在 500 次目标寿命下的裂纹扩展行为，结果如图 19 所示。

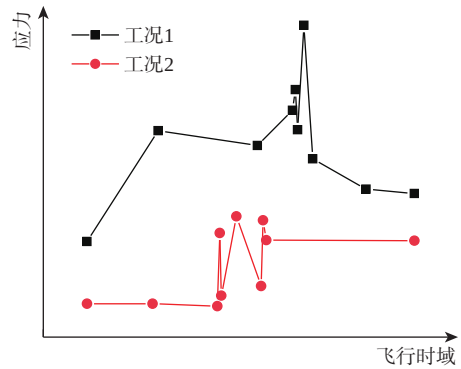


图 18　应力时域信号

Fig. 18　Stress time domain signal

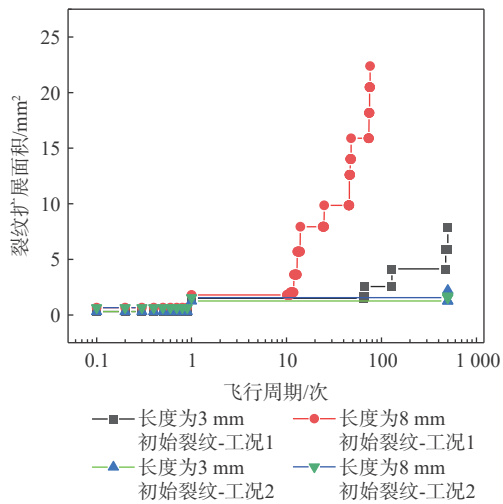


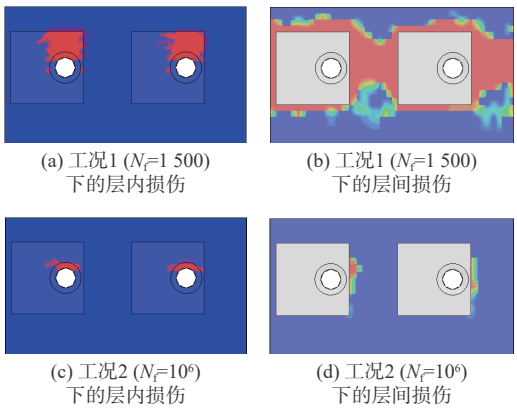
图 19　预置裂纹在 2 种载荷下扩展情况

Fig. 19　Preset crack propagation under two loads

可见，在工况 1 下，3 mm 初始裂纹在 90 次飞行周期开始扩展，在 500 次飞行周期发生轻微扩展，8 mm 初始裂纹在 10 次飞行周期后开始显著扩展，在约 75 次飞行周期后达到断裂。在工况 2 的载荷下，2 种初始裂纹长度结构在 500 次飞行周期内均保持缓慢扩展。

2）含复材孔边分层金属-复材连接结构。计算了含复材孔边分层（见图 3）的连接结构在 2 种工况下的剩余寿命。

2 种工况下的复合材料损伤及寿命如图 20 所示。在工况 1 载荷作用下结构寿命 N_f 约为 1 500 次循环，层内与层间均发生较严重的破坏。在工况 2 载荷作用下，结构可在经历 1×10^6 次循环加载下相对保持结构完整，孔边受压一侧发生轻微层内损伤，由于结构弯曲产生的层间剪切作用，预置分层区域的侧向发生轻微分层扩展。



(a) 工况1 (N_f=1 500)下的层内损伤　(b) 工况1 (N_f=1 500)下的层间损伤

(c) 工况2 (N_f=10^6)下的层内损伤　(d) 工况2 (N_f=10^6)下的层间损伤

图 20　复合材料层板 2 种载荷循环作用下的损伤

Fig. 20　Damage of laminates under two load cycles

4　讨　论

从含预置金属孔边穿透裂纹试验件结果来看, 载荷水平越高, 裂纹扩展速率越大, 剩余寿命越小。在裂纹开始扩展后随着加载次数的增加, 裂纹长度不断增大。从扩展曲线来看, 试验并没有体现出一般远端加载的金属试验件中典型的从裂纹“出现-缓慢扩展-急速扩展”的过程, $\Delta a-\lg N$为非连续曲线, 即在金属外表面观测到裂纹后, 裂纹很快进入快速扩展阶段, 其原因主要是由于结构特性引起的偏心效应[28], 导致金属件的上表面裂纹扩展长度要比下表面的长, 由于试验中内表面无法观测, 只能观测到下表面的裂纹扩展情况, 因此, 对金属-复材连接结构, 采用目视检测金属表面裂纹长度来表征结构的损伤程度是非常保守的。

试验过程中裂纹的扩展方向与载荷方向并不垂直其主要原因可能是由于金属和复材接触表面存在一定的摩擦, 导致载荷方向发生偏转。

1) 含初始复材孔边分层的金属-复材连接结构在拉-拉疲劳载荷作用下的损伤是以孔边的挤压破坏为主, 同时孔边分层边缘也会产生一定的损伤起始, 但是在最终破坏前, 损伤扩展并不明显,疲劳破坏过程比较突然。

2) 对比含不同损伤的 2 种连接结构的剩余寿命结果来看, 不含孔边分层的复材部分在较高的疲劳载荷作用下也不会明显损伤, 说明孔边分层会明显降低金属-复材连接结构的剩余寿命。

3) 对比含金属裂纹和复材孔边分层 2 种试验件测试结果, 说明复材孔边分层的影响明显大于金属孔边裂纹的影响, 建议针对金属-复材连接结构进行损伤容限设计时, 重点考虑复材孔边分层损伤的影响。

4) 根据随机谱下的分析结果可以看出, 工况 1 载荷下含 8 mm 初始裂纹结构会发生明显扩展, 工况 2 载荷作用则基本不扩展; 含 3 mm 裂纹结构基本都不扩展, 说明裂纹扩展受载荷大小和初始长度共同影响。

5　结　论

1) 含初始孔边分层(复材)的金属-复材的双钉横向连接结构, 在拉-拉疲劳载荷作用下, 复合材料部分的孔边分层会在挤压载荷作用下发生一定扩展; 受挤压一侧逐渐发生挤压疲劳损伤扩展, 但是最终破坏前损伤观测都不明显。

2) 含初始孔边穿透裂纹(金属)的金属-复材双钉横向连接结构, 在拉-拉疲劳载荷作用下, 金属内表面的裂纹扩展速率快于外表面。对于此类连接结构, 仅采用表面裂纹长度来表征金属损伤程度可能导致保守估计。

3) 复材孔边分层会明显降低金属-复材连接结构的寿命和承载能力; 金属-复材连接结构的损伤容限设计应该重点关注复材孔边分层损伤。

4) 本文建立的含损伤金属-复材单排双列的连接结构疲劳预测模型, 能够预测此类结构的剩余寿命和损伤扩展行为, 可用于指导此类结构的损伤容限设计和评估。

参考文献（References）

[1] WANHILL R J H, BRAY G H. Aluminum-lithium alloys: processing, properties, and applications [M]. Boston : Elsevier Science, 2014: 381-413.

[2] 秦剑波, 宁宇, 周颜, 等. 2E12 铝合金机身壁板损伤容限分析与试验验证[J]. 南京航空航天大学学报, 2023, 55(1): 73-79.
QIN J B, NING Y, ZHOU Y, et al. Analysis and test verification of damage tolerance of 2E12 aluminum alloy fuselage panel[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2023, 55(1): 73-79(in Chinese).

[3] 赵永庆, 吴聪, 王欢. 1200 MPa 级高强高韧损伤容限钛合金强韧性研究进展[J]. 稀有金属材料与工程, 2022, 51(12): 4389-4397.
ZHAO Y Q, WU C, WANG H. Advance in relationship between tensile strength and toughness for 1200 MPa high strength and high toughness Ti-alloy with damage tolerance[J]. Rare Metal Materials and Engineering, 2022, 51(12): 4389-4397(in Chinese).

[4] 孙侠生, 苏少普, 孙汉斌, 等. 国外航空疲劳研究现状及展望[J]. 航空学报, 2021, 42(5): 40-65.
SUN X S, SU S P, SUN H B, et al. Current status and prospect of overseas research on aeronautical fatigue[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(5): 40-65(in Chinese).

[5] 付晨. 纤维增强复合材料层合板分层损伤机理与预测方法研究[D]. 北京: 北京交通大学, 2022.
FU C. Research on the damage mechanisms and the prediction method of delamination in fiber-reinforced composite laminates[D]. Beijing: Beijing Jiaotong University, 2022(in Chinese).

[6] 顾志平. 基于内聚力理论的复合材料分层失效机理研究[D]. 杭州: 浙江大学, 2016.

GU Z P. Study on delamination failure mechanism of composite laminates based on cohesive theory[D]. Hangzhou: Zhejiang University, 2016(in Chinese).

[7] 王小玲. 含损伤复合材料层合板疲劳特性研究[D]. 天津: 中国民航大学, 2009.
WANG X L. Fatigue behavior study of damaged composite laminates[D]. Tianjin: Civil Aviation University of China, 2009(in Chinese).

[8] AOKI R, HIGUCHI R, YOKOZEKI T. Progressive damage and residual strength of open-hole thin-ply CFRP laminates under tensile fatigue loading[J]. Composite Structures, 2023, 314: 116973.

[9] 汪厚冰, 邓凡臣, 魏宏艳, 等. 含开孔与含裂缝的复合材料层压板拉伸剩余强度评估[J]. 复合材料科学与工程, 2024(2): 27-34.
WANG H B, DENG F C, WEI H Y, et al. Residual strength evaluation of laminates with open hole or crack under tensile load[J]. Composites Science and Engineering, 2024(2): 27-34(in Chinese).

[10] 苏玲, 徐倩. 含应力集中复合材料结构损伤容限分析方法研究与应用[J]. 导弹与航天运载技术, 2018(2): 105-109.
SU L, XU Q. Research and application of damage tolerance analysis method for composite structure with stress concentration[J]. Missiles and Space Vehicles, 2018(2): 105-109(in Chinese).

[11] 杨钧超, 王雪明, 陈向明, 等. 低速冲击损伤对复材加筋板压缩性能的影响[J]. 航空学报, 2023, 44(20): 228498.
YANG J C, WANG X M, CHEN X M, et al. Effect of low-velocity impact damage on compressive properties of composite stiffened panels[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(20): 228498(in Chinese).

[12] ZHENG Y P, ZHANG C Y, TIE Y, et al. Tensile properties analysis of CFRP-titanium plate multi-bolt hybrid joints[J]. Chinese Journal of Aeronautics, 2022, 35(3): 464-474.

[13] 陈龙. 复合材料-钛合金混合结构多钉连接分析与试验研究[D]. 南京: 南京航空航天大学, 2020.
CHEN L. Testing and numerical analysis of composite-to-titanium hybrid structure with multiple-bolted joints[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2020(in Chinese).

[14] 方志晓. 钛合金与复合材料壁板连接结构群钉载荷分配仿真与试验研究[D]. 哈尔滨: 哈尔滨工业大学, 2016.
FANG Z X. Research on the load distribution of mechanical connection structure of titanium and composite panel by simulation and experiment[D]. Harbin: Harbin Institute of Technology, 2016(in Chinese).

[15] 桂林. 温度对复合材料-金属机械连接结构静强度影响研究[D]. 哈尔滨: 哈尔滨工业大学, 2021.
GUI L. Research on the influence of temperature on the static strength of composite material-metal mechanical connection structure[D]. Harbin: Harbin Institute of Technology, 2021(in Chinese).

[16] CHEN Y W, YANG X J, LI M J, et al. Mechanical behavior and progressive failure analysis of riveted, bonded and hybrid joints with CFRP-aluminum dissimilar materials[J]. Thin-Walled Structures, 2019, 139: 271-280.

[17] 贾云超, 关志东, 宋晓君. 复合材料-金属机械连接性能研究[J]. 玻璃钢/复合材料, 2015(4): 66-70.
JIA Y C, GUAN Z D, SONG X J. Study on performance of composite-metal mechanical joints[J]. Fiber Reinforced Plastics/Composites, 2015(4): 66-70(in Chinese).

[18] 沈真, 柴亚南, 杨胜春, 等. 复合材料飞机结构强度新规范要点评述[J]. 航空学报, 2006, 27(5): 784-788.
SHEN Z, CHAI Y N, YANG S C, et al. The outline of new specifications on strength of composite aircraft structures[J]. Acta Aeronautica et Astronautica Sinica, 2006, 27(5): 784-788(in Chinese).

[19] 侯赤, 周银华, 全泓玮, 等. 混合结构中金属疲劳对层合板损伤的影响[J]. 西北工业大学学报, 2018, 36(1): 74-82.
HOU C, ZHOU Y H, QUAN H W, et al. The effect of metal on composites fatigue damage in mixed structure[J]. Journal of Northwestern Polytechnical University, 2018, 36(1): 74-82(in Chinese).

[20] DI CAPRIO F, RUSSO A, MANSERVIGI C, et al. Damage tolerance evaluation of a C/C-SiC composite body flap of a re-entry vehicle[J]. Composite Structures, 2021, 274: 114341.

[21] BABUSKA P, GOYAL V K, MAQUEDA I, et al. Damage tolerance approach for composite space structures with curved bends[C]// Proceedings of the AIAA Scitech 2020 Forum. Reston: AIAA, 2020.

[22] 蔡巧言, 李配缘, 黎增山, 等. 可重复使用运载器舵轴裂纹扩展寿命分析[J]. 强度与环境, 2023, 50(6): 1-8.
CAI Q Y, LI P Y, LI Z S, et al. The analysis of the rudder shaft crack growth life analysis of the reusable vehicle[J]. Structure & Environment Engineering, 2023, 50(6): 1-8(in Chinese).

[23] Dassault Systèmes Simulia Corp. Abaqus/standard 6.14 user's manual[M]. Providence: Dassault Systèmes, 2014.

[24] 航空航天工业部科学技术研究院. 美国空军损伤容限设计手册: 飞机结构损伤容限分析和设计指南[M]. 西安: 西北工业大学出版社, 1989.
Science and Technology Research Institute of the Ministry of Aerospace Industry. USAF damage tolerant design handbook: guidelines of the analysis and design of damage tolerant aircraft structures[M]. Xi'an: Northwestern Polytechnical University Press, 1989(in Chinese).

[25] SHOKRIEH M M, LESSARD L B. Progressive fatigue damage modeling of composite materials, part I: modeling[J]. Journal of Composite Materials, 2000, 34(13): 1056-1080.

[26] CAMANHO P P, DAVILA C G, DE MOURA M F. Numerical simulation of mixed-mode progressive delamination in composite materials[J]. Journal of Composite Materials, 2003, 37(16): 1415-1438.

[27] 万傲霜, 朱飞扬, 云新尧, 等. 平面机织复合材料疲劳分层数值分析方法研究[J]. 复合材料学报, 2024, 41(8): 4418-4433.
WAN A S, ZHU F Y, YUN X Y, et al. Investigation on numerical analysis method of fatigue delamination damage of plane woven composites[J]. Acta Materiae Compositae Sinica, 2024, 41(8): 4418-4433(in Chinese).

[28] 刘峰, 周建国, 陈威杨, 等. 维修扩孔对碳纤维复合材料层合板与金属板双钉连接强度的影响分析[J]. 宇航材料工艺, 2022, 52(3): 94-102.
LIU F, ZHOU J G, CHEN W Y, et al. Analysis of maintenance reaming effect on the strength of double-bolted joints of carbon fiber composite laminates and metal plates[J]. Aerospace Materials & Technology, 2022, 52(3): 94-102(in Chinese).

Damage tolerance analysis on reusable launch vehicle connection structures

LI Peiyuan[1], CAI Qiaoyan[2,*], LI Zengshan[3], FENG Jiahe[3]

(1. Beijing Institute of Astronautical Systems Engineering, Beijing 100076, China;
2. China Academy of Launch Vehicle Technology, Beijing 100076, China;
3. School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China)

Abstract: A damage tolerance test of the metal-composite connection structure with initial damage was conducted for the composite skin-metal skeleton structure used in the reusable launch vehicle, and a damage tolerance performance prediction model of the corresponding structure was created. The research results show that metal structures containing initial cracks at the edge of the connection hole will undergo crack expansion under fatigue load. Due to the geometric asymmetry of the connection structure, the crack expansion rates on the upper and lower surfaces are inconsistent, resulting in crack observation from the outer surface causing an error in the damage size. Extrusion load causes some delamination expansion in composite laminates with preset delaminations; extrusion fatigue damage to the composite materials is the primary cause of the damage. Delamination around composite connection holes will drastically reduce the life of the metal-composite connection structure. The damage tolerance design of the metal-composite connection structure should focus on composite hole edge delamination damage. The prediction results of the established metal-composite structure damage tolerance performance prediction model are in good agreement with the test results. It can well predict the damage expansion process and extended life of the structure, and can be used to guide the damage tolerance design of this type of structure.

Keywords: reusable launch vehicle; structure; damage tolerance; finite element method; damage detection

Received: 2024-04-16; **Accepted**: 2024-06-21; **Published Online**: 2024-06-27 10:16

URL: link.cnki.net/urlid/11.2625.V.20240626.1114.001

Foundation item: Financial Support for Key Independent Research and Development Projects of China Academy of Launch Vehicle Technology

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0213

一体化加力燃烧室燃烧特性模拟与机器学习燃烧性能预测

张宇[1]，王奉明[2,3]，王彦红[4]，穆林[1]，东明[1,3,*]

(1. 大连理工大学 能源与动力学院，大连 116024；2. 南京航空航天大学 能源与动力学院，南京 210016；
3. 先进航空动力创新工作站，北京 101300；4. 东北电力大学 能源与动力工程学院，吉林 132012)

摘 要：针对航空发动机提高推重比的问题，提出涡轮后框架支板、径向火焰稳定器、喷油杆一体化的加力燃烧室设计方案，开展不同油气比和涵道比下加力燃烧室热态流场和燃烧特性大涡模拟研究。揭示燃烧室温度、速度、压力的分布特征及其对燃烧过程的影响。探究燃油液滴、氧气、二氧化碳和水的分布状况。建立机器学习模型，以内涵进气压力、进气温度、喷油量、涵道比、轴向距离为输入变量，对总压恢复系数、温度均匀性系数、燃烧效率进行了预测。结果表明：该一体化加力燃烧室的整体燃烧性能良好，仅在火焰下方观察到局部燃烧弱化区。燃烧室前部存在3个低速回流区，分别位于强火焰区、喷油杆中下部和中心锥尾部。随着油气比增大，燃烧效率下降；随着涵道比增加，燃烧效率提高。所建机器学习模型训练集和测试集的校正决定系数大于0.788，预测效果较好。

关 键 词：一体化加力燃烧室；热态流场；回流区；大涡模拟；机器学习；燃烧性能预测
中图分类号：V235.11
文献标志码：A **文章编号**：1001-5965(2026)06-2000-11

随着航空发动机性能需求的不断提高，需要克服传统加力燃烧室喷油杆、火焰稳定器、混合器置于高温燃气环境下流动损失大、部件容易烧蚀、燃烧效率低等缺陷和不足[1-3]。通过将喷油和稳燃等装置进行一体化设计，以加力燃烧室结构优化增加推重比、减少发动机质量、提高燃烧性能并进行良好冷却防护是现今航空发动机发展的前沿方向[4-6]。一体化加力燃烧室的热态流场、燃烧特性及燃烧性能预测受到广泛的重视。

国内外学者对一体化加力燃烧室的冷态热态流场、燃烧特性和燃烧性能预测进行了一定的研究。Jia等[7]研究表明，外涵冷空气对隔热屏和一体化结构均起到冷却作用。随着涵道比增加，沿程热混合效率不断增大。Li等[8]把双壁冷却结构应用于一体化加力燃烧室，观察到螺旋涡流和回流涡流，一方面起到稳定燃烧的作用，另一方面回流涡流压力促使冷空气以薄膜流的方式覆盖壁面，可以有效地冷却壁面。韦裕恒等[9]提出前缘冲击和气膜孔复合冷却方案，冲击板结构改变了冷气流涡状态和气膜流量分配，对外涵气流与内涵气流的掺混和燃烧室的防护极为有利。Huang等[10]研究发现，非均匀进口流速具有偏心的速度剖面，不对称的双涡结构起到促进掺混燃烧作用。尹成茗等[11]提出截尾支板与凹腔分流环的一体化加力燃烧室方案，其具有3个回流区，可起到稳焰和联焰作用，总压恢复系数和燃烧效率相比传统加力燃烧室更高。

收稿日期：2024-04-12；录用日期：2024-05-24；网络出版时间：2024-07-08 09:32
网络出版地址：link.cnki.net/urlid/11.2625.V.20240705.1209.003
基金项目：先进航空动力创新工作站 (依托中国航空发动机研究院) 资助 (HKCX2022-01-017)
* 通信作者. E-mail：dongming@dlut.edu.cn

引用格式：张宇，王奉明，王彦红，等. 一体化加力燃烧室燃烧特性模拟与机器学习燃烧性能预测 [J]. 北京航空航天大学学报，2026，52（6）：2000-2010. ZHANG Y，WANG F M，WANG Y H，et al. Simulation of combustion characteristics and prediction of combustion performance using machine learning in an integrated afterburner [J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（6）：2000-2010（in Chinese）.

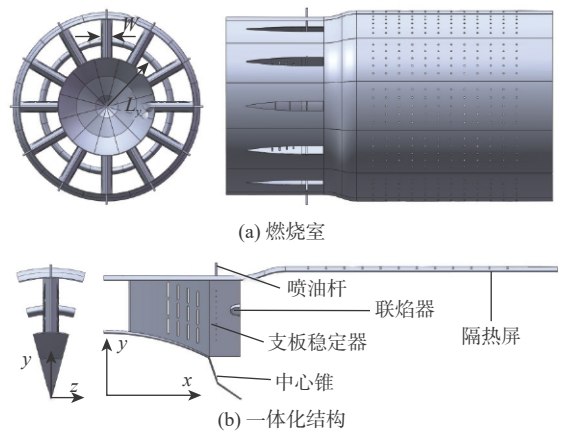
谭云川等[12]实验探讨了等面积法设计和等氧含量法设计 2 种燃油分配对加力燃烧室热态流场和燃烧性能的影响, 等氧含量法设计的燃油方案具有更高的燃烧效率。王伟龙等[13]提出带内突扩中心锥的稳定器方案, 该一体化结构相比钝体稳定器对总压恢复系数和燃烧效率具有突出的改善作用, 波瓣混合器的燃烧室性能最好。付垚等[14]研究发现, 喷油杆多喷嘴分散方案相比喷油杆少喷嘴集中方案的燃烧室燃油分布更加均匀, 燃烧性能更好。通过机器学习对一体化加力燃烧室的燃烧性能预测也值得重视[15]。安健等[16]提出机器学习可以简化燃烧反应动力学机理实现计算加速, 辅助湍流燃烧大涡模拟建模, 进行流场和燃烧性能预测。基于火箭基燃烧室, An 等[17]把雷诺时均模拟数据进行自组织映射和反向传播神经网络处理, 通过该方法重新构建甲烷燃烧化学反应机理, 获得燃烧室温度和燃烧产物云图, 预测燃烧效率, 计算精度较高且速度更好。Yao 等[18]通过机器学习建立大涡模拟耗散率亚网格模型, 修正的大涡模拟模型对喷雾燃烧模拟性能更好。Wong 等[19]通过机器学习建立发动机燃烧模型, 对燃烧室流场、排气温度、燃烧过程和污染物排放进行了有效预测。已有研究不足体现在 3 个方面: 实验研究不能准确反映一体化加力燃烧室的详细结构; 数值研究多采用常规湍流模型, 相比这些湍流模型, 大涡模拟可以更好地反映火焰结构和温度场分布状况; 机器学习在一体化加力燃烧室燃烧研究还较少应用, 需要通过机器学习对总压恢复系数、温度均匀性系数及燃烧效率进行预测。

本文针对隔热屏气膜孔冷却的一体化加力燃烧室, 其将支板稳定器、联焰器、喷油杆进行一体化设计, 探究油气比和涵道比对一体化燃烧室温度场、压力场和流场分布的影响, 揭示燃油液滴分布和燃烧过程机制, 评估不同工况下的总压恢复系数、温度均匀性系数和燃烧效率并通过机器学习进行预测, 可以为加力燃烧室的一体化设计和燃烧组织提供理论基础和应用方案。

1　数值模型与数值设置

1.1　物理模型

图 1 为一体化加力燃烧室的设计方案。内涵面积和外涵面积比值约为 3:1, 该燃烧室总长度为 1 000 mm, 燃烧室进口到喷油杆的距离为 300 mm, 中心锥半径为 150 mm, 燃烧室近似为圆筒, 圆筒半径为 300 mm。外涵冷气流经过隔热屏的气膜孔(周向 60 个×轴向 12 排)对燃烧室壁面进行冷却。对支板稳定器、联焰器、喷油杆进行一体化设计, 喷油杆处于内部, 与高温环境不接触。联焰器位置 L_y=220 mm, 支板宽度 W=32 mm。喷油杆有 10 排喷口, 每排 2 个喷口, 同时使用外涵气流冷却一体化结构。加力燃烧室旋转对称, 选取一个完整的一体化结构, 即周向 1/12 的扇形作为计算域。外涵和内涵均为质量流量进口, 内涵进气流量为 4.5 kg/s, 进口马赫数为 0.3, 涵道比 B 的范围为 0.30～0.35。外涵进气温度设置为 400 K, 内涵进气压力范围为 0.2～0.6 MPa, 内涵进气温度范围为 900～1 200 K。燃油进口为离散相, 喷油量分别为 0.15 kg/s、0.20 kg/s、0.25 kg/s、0.30 kg/s, 定义油气比为喷油量与加力燃烧室内外涵进口总质量流量之比, 则对应的油气比 M 依次为 2.5%、3.3%、4.2%、5.0%。加力燃烧室出口设置为压力出口, 侧面为旋转周期边界, 流固交界面设置为耦合面, 其余壁面为绝热壁面。



(a) 燃烧室

(b) 一体化结构

图 1　一体化加力燃烧室

Fig. 1　Integrated afterburner

1.2　数值方法

对计算域划分四面体非结构化网格, 最大网格尺度为 4 mm。流固交界面进行节点融合, 对一体化部件进行局部网格加密。通过网格无关性分析确定合理的网格方案, 选取 5 种网格数量, 分别为 9.43×10^6、1.146×10^7、1.449×10^7、1.608×10^7、1.823×10^7。表 1 给出了不同网格数量下加力燃烧室出口温度和出口速度情况。喷油量为 0.20 kg/s, 涵道比 B 为 0.33, 内涵进气压力为 0.5 MPa, 内涵进气温度为 1 000 K。可以看到, 计算域网格数量为 1.449×10^7 时, 继续增加网格数量对出口温度和出口速度基本无影响, 满足网格无关性的要求。因此, 确定计算域的网格数量为 1.449×10^7。网格划分如图 2 所示。

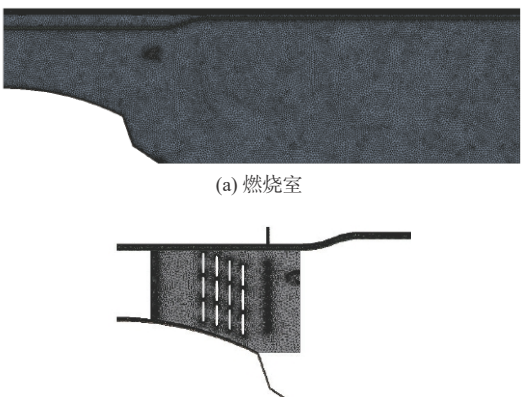
通过 Fluent 进行三维数值计算, 求解稳态的连续性方程、动量守恒方程、能量守恒方程和组分输运方程, 采用大涡模拟处理湍流问题。将大涡模拟所得数据进行时均处理, 相较于雷诺时均模型可以

更好地反应火焰的分布状况。燃油 $C_{12}H_{23}$ 燃烧反应取自文献 [20] 提出的 16 组分 23 步机理模型。对气相湍流燃烧采用非预混燃烧模型，其假设反应受混合速率控制，只求解混合物分数及其方差的输运方程，无需求解组分的输运方程，且考虑湍流与化学反应的相互作用，选择的是稳态扩散小火焰模型。气液两相流动采用拉格朗日法，通过离散相模型考虑液滴与流体之间的相互作用。燃油液滴服从 Rosin-Rammler 分布函数，液滴直径范围设置为 0.02～0.04 mm，符合文献 [21] 提出的 10^{-2} 数量级。液滴追踪采用随机轨道模型。压力速度耦合采用 Coupled 算法，基于二阶迎风格式离散控制方程。

表 1 网格无关性分析

Table 1 Grid independence analysis

网格数量	出口温度/K	出口速度/(m·s^{-1})
9.43×10^{6}	2 057.72	258.14
1.146×10^{7}	2 061.48	260.61
1.449×10^{7}	2 064.75	263.45
1.608×10^{7}	2 064.91	263.84
1.823×10^{7}	2 065.24	264.59



(a) 燃烧室

(b) 一体化结构

图 2 计算域网格

Fig. 2 Computational domain grids

1.3 模型验证

目前，缺乏一体化加力燃烧室的运行和试验数据，通过文献 [22] 简化的矩形加力燃烧室燃烧特性试验数据进行模型验证。矩形加力燃烧室的长度为 1 480 mm，宽度为 125 mm，详细结构和计算参数见文献 [22]。表 2 给出了数值计算的燃烧室出口平均温度 T_{out} 和总压恢复系数 σ 的数值数据与试验结果对比情况。数值计算包括文献 [22] 采用 Realizable k-ε 湍流模型和本文大涡模拟的结果。通过该表可以看到，数值结果与实验数据符合较好，大涡模拟可以更好地反映加力燃烧室的燃烧性能，数值结果的误差主要源于大涡模拟对小涡的模化处理，可以运用到复杂加力燃烧室结构的数值研究。

表 2 试验数据与数值结果对比

Table 2 Comparison of experimental data and numerical results

数据来源	T_{out}/K	σ/%
试验数据	1 564	99.1
文献[22]数值结果	1 616	99.4
本文数值结果	1 598	98.8

2 数值结果与分析

2.1 油气比对燃烧的影响

本节讨论油气比 M 对燃烧特性的影响，M 的变化范围为 2.5%～5.0%，涵道比为 0.33，内涵进气压力为 0.5 MPa，内涵进气温度为 1 000 K。

图 3 为不同油气比下燃烧室温度 T 分布情况。由图可知，M=2.5% 时一体化结构后方观察到明显的高温火焰，高温火焰趋向燃烧室顶部，火焰燃烧沿轴向传播到整个燃烧室直至尾喷口。除隔热屏附近为外涵冷空气以外，高温燃气占据燃烧室大部分空间，整体的热态性能较好。燃烧室发现 2 个局部低温区，一个是高温火焰下方的椭圆形低温区，另一个是高温燃气包覆下中心锥后方的尖角形低温区。尾喷口截面为 2 个紧邻的圆形高温区。随着油气比增加，燃烧室温度明显下降，热态性能变差，相应的燃烧性能也减弱。高油气比下椭圆形低温区不断拓展甚至下方又出现新的低温区，中心锥后方的狭长低温区受到压缩。M=5.0% 时，燃烧室中心对称轴观察到低温燃气层。加力燃烧室平均温度由 M=2.5% 的 1 956 K 下降到 M=5.0% 的 1 752 K，温度场的径向均匀度也显著下降。文献 [11] 标准 k-ε 湍流模型给出的是连续高温温度场，观察不到火焰形态，而通过大涡模拟可以反映火焰的分布状况。

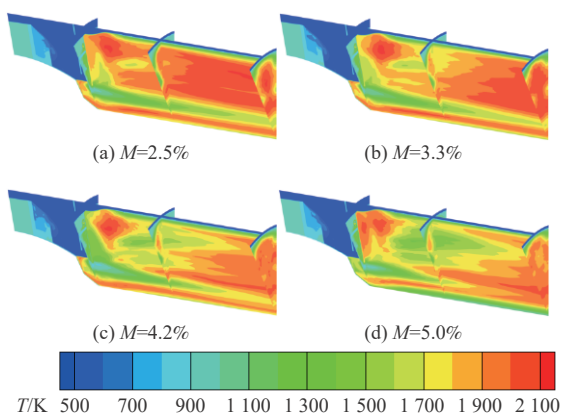


图 3 不同油气比下加力燃烧室温度分布情况

Fig. 3 Temperature distribution of afterburner under various kerosene-air ratios

图 4 为不同油气比下燃烧室压力 p 分布情况。可以看到, 一体化结构后方出现强燃烧形成的火焰形低压区, 喷油的雾化、蒸发、掺混、燃烧集中在该区间。内涵进气经一体化结构的局部突缩, 压力提高, 促使火焰趋向顶部, 造成喷油杆下方燃油与内涵进气不能有效掺混, 形成高压气体隔绝层。核心燃烧区气流回流到中心锥, 进入隔绝层下方出现低压层。随着油气比增加, 火焰受到压缩, 火焰区的压力提高。因为设定相同进口流量, 受燃烧室燃烧影响进口区压力具有一定的差别。

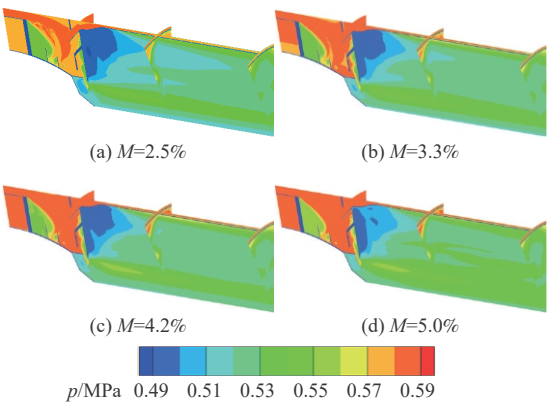


图 4　不同油气比下加力燃烧室压力分布情况

Fig. 4　Pressure distribution of afterburner under various kerosene-air ratios

图 5 给出了不同油气比下燃烧室流场(速度 u 和流线)分布情况。可以看到, 燃烧室存在 3 个典型的低速区, 分别位于强火焰区、喷油杆中下部和中心锥尾部。火焰区低流速有利于燃油的雾化、蒸发和掺混, 而足够的停留时间有助于燃烧的充分发展。强火焰区的核心发现局部旋涡回流, 该回流卷吸高温燃气促进油气的掺混和着火过程。喷油杆中下部有来自外涵冷空气的向下冲刷, 内涵热气流经中心锥突扩向上偏移, 形成逆时针回流, 回流区油气混合后沿旋涡传输到上方强火焰区, 该回流区是造成火焰偏向顶部的原因。中心锥尾部出现顺时针回流形成的高压低温层, 回流还将高温燃气输运到弱燃烧区下方形成高温层。油气比改变对燃烧室流场的影响较为微弱。中心锥尾部回流区在文献 [11] 和文献 [23] 出现过, 而其他 2 个回流区源于本文一体化结构对流场的影响。

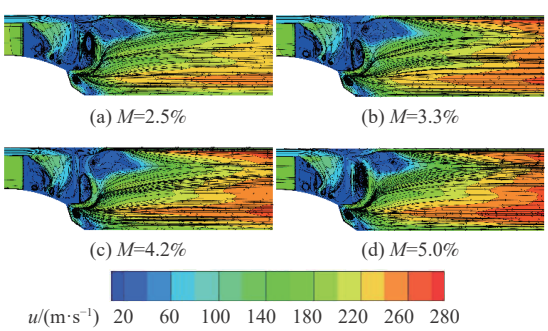


图 5　不同油气比下加力燃烧室流速和流线分布情况

Fig. 5　Velocity and streamline distributions of afterburner under various kerosene-air ratios

图 6 为不同油气比下支板截面燃油液滴分布情况。可以看到, 燃油液滴从喷油杆喷射, 除联焰器后方的空间以外, 在燃烧室充满度较好, 基本延伸到燃烧室中后部, 支撑中后部的燃烧过程。回流区裹挟作用使中心锥后方也布有液滴。

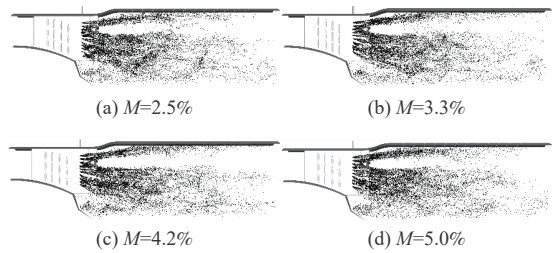


图 6　不同油气比下支板截面燃油液滴分布情况

Fig. 6　Fuel droplets distributions of strut section under various kerosene-air ratios

图 7 为不同油气比下氧气、二氧化碳和水的质量分数分布情况。可以看到, 中心锥后方存在一条狭长的富氧带, 对应的二氧化碳和水蒸气含量较少, 说明该区域燃烧较弱。这是因为虽然中心锥后方油滴较为丰富, 但是油气掺混不足, 内涵进气主要起冷却作用, 形成富氧燃烧弱化层。随着油气比增加, 燃烧室氧气含量下降, 抑制了二氧化碳和水蒸气的生成过程, 说明油气比越大, 更多的生成一氧化碳, 越不利于燃烧过程。

通过总压恢复系数 σ[23] 和温度均匀性系数 ξ[14] 定量描述加力燃烧室的热态压力场和温度场, 两者的定义如下:

$$\sigma = p_x / p_{\text{in}} \tag{1}$$

式中: p_x 为沿 x 轴的径向截面总压力; p_{in} 为进口总压力。

$$\xi = 1 - \frac{\sum_{i=1}^{N} |T_i - T_{\text{m}}| (\rho_i u_i A_i)}{2|T_{\text{m}}| \sum_{i=1}^{N} (\rho_i u_i A_i)} \tag{2}$$

式中: N 为截面节点数; T_i 为第 i 个节点的温度; ρ 为混合气体密度; u_i 为第 i 个节点的速度大小; A 为面积; T_{m} 为平均温度。

图 8 为不同油气比下湍动能 k、总压恢复系数 σ和温度均匀性系数 ξ的沿程分布情况。图中, x=300 mm 为喷油杆对应的截面。可以看到, 喷出的燃油与空气掺混, 通过燃烧产生大量小分子碳氢产物, 流动加速, 湍动能增加, 油气掺混和气体剧烈

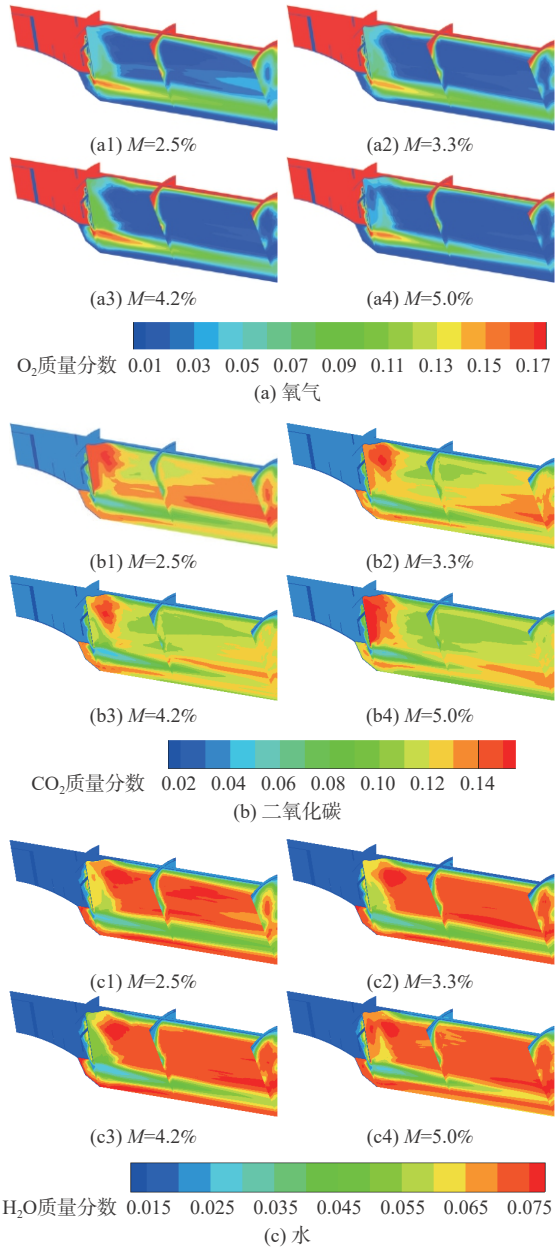


图7 不同油气比下氧气、二氧化碳和水的分布情况

Fig. 7 Distributions of oxygen, carbon dioxide and water under various kerosene-air ratios

膨胀带来突出的总压损失，总压恢复系数急剧下降。径向的强燃烧区和弱燃烧区及回流区导致温度均匀性系数陡然下降。随着沿程燃烧的逐渐平稳，燃烧过程不再剧烈，湍动能不断下降，燃烧膨胀效应带来的压损减小，温度不均匀性略有改善。油气比增加对湍动能和总压恢复系数的影响较弱，除燃烧室后部以外表现为喷油量越大温度越均匀。

通过燃烧效率[6]表征一体化加力燃烧室的燃烧性能，其定义如下：

$$\eta=\frac{m_x c_{p,x} T_x - m_{\text{in}} c_{p,\text{in}} T_{\text{in}} - m_{\text{f}} c_{p,\text{f}} T_{\text{f}}}{m_{\text{f}} H_{\text{f}}} \tag{3}$$

式中：m_x为沿x轴的径向截面的质量流量；m_{in}为进口质量流量；m_{f}为燃油质量流量；$c_{p,x}$为沿x轴的径向截面的定压比热容；$c_{p,\text{in}}$为进口定压比热容；$c_{p,\text{f}}$为燃油定压比热容；T_x为沿x轴的径向截面的温度；T_{in}为进口温度；T_{f}为燃油温度；H_{f}为燃油低位热值，H_{f}=43 348 kJ/kg。

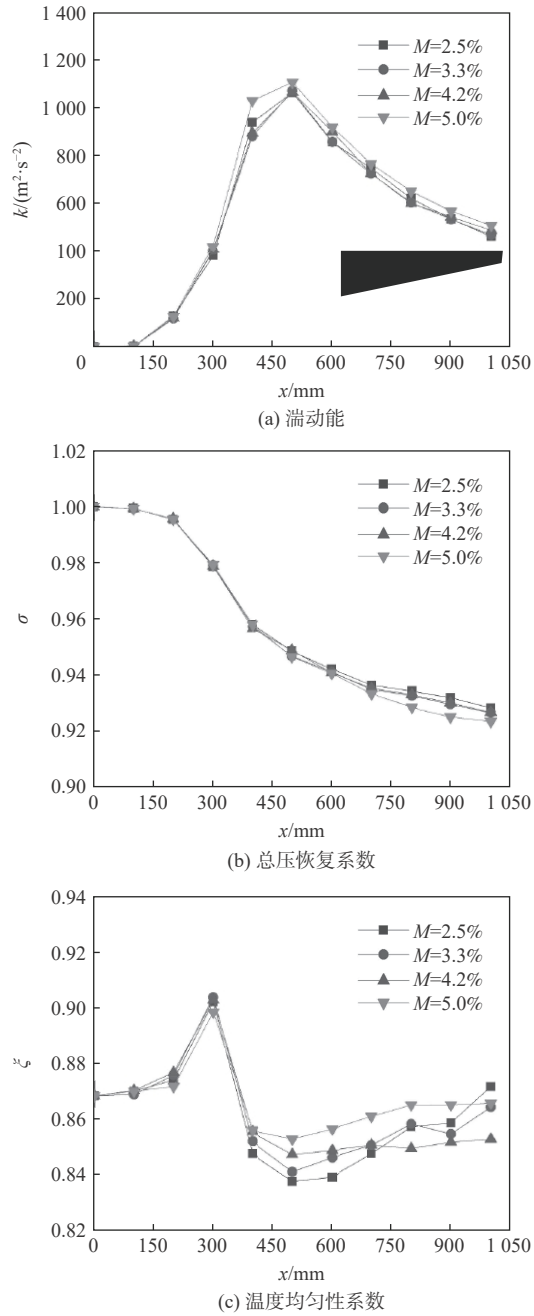


图8 不同油气比下湍动能、总压恢复系数和温度均匀性系数的轴向变化情况

Fig. 8 Turbulent kinetic energy, total pressure recovery coefficient, and temperature uniformity coefficient axial distributions under various kerosene-air ratios

图9给出了不同油气比下燃烧效率η的沿程变化情况。由图可知，随着沿程距离增加，燃烧效率不断提高。燃烧室x=300～700 mm之间，燃烧室存在低速回流区，气流速度低和回流效应有利于燃油

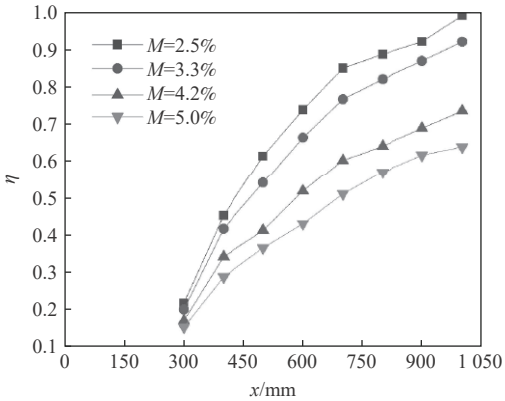


图 9　不同油气比下燃烧效率的轴向变化情况

Fig. 9　Axial distribution of combustion efficiency under various kerosene-air ratios

的蒸发和掺混, 产生强烈火焰, 燃烧效率迅速增大。之后燃烧过程平稳, 燃烧效率的上升斜率下降。随着油气比增加, 燃烧效率显著下降, M 为 2.5%、3.3%、4.2%、5.0%, 尾喷口截面的燃烧效率依次为 99.0%、92.0%、73.4%、63.7%。这是因为燃油量增加导致燃烧室局部燃油过富, 致使燃烧受到削弱。加力燃烧室燃烧性能主要通过燃烧效率评定, 因此, 油气比为 2.5% 时对应最佳的喷油量, 燃烧性能最好。

2.2　涵道比对燃烧的影响

接着讨论涵道比对燃烧特性的影响, B 的变化范围为 0.30～0.35, 喷油量为 0.20 kg/s, 内涵进气压力为 0.5 MPa, 内涵进气温度为 1 000 K。

外涵冷空气通过隔热屏对燃烧室壁面形成冷却作用, 之后与燃烧室顶部燃油混合继续组织燃烧过程。涵道比增大, 外涵冷空气的流量增大。图 10 为不同涵道比下燃烧室温度和压力分布情况。由图可知, 涵道比增大, 燃烧室中后部高温区范围不断压缩, 说明过多的外涵冷空气对燃烧室起到冷却作用, 同时外涵冷空气不利于提供燃烧所需要的活化能, 削弱了燃烧过程。涵道比对燃烧室压力分布基本没有影响。

图 11 为不同涵道比下燃烧室流场和燃油液滴分布情况。由图可知, 涵道比增大, 燃烧室中后部高流速向上游延伸, 同时燃烧室前部燃油液滴量增多, 而中后部燃油液滴量减少, 说明外涵进汽量增大增强了燃烧室前部的燃烧, 高流速向上游拓展, 而燃烧室中后部燃油量不足, 主要起到冷却作用, 致使燃烧室温度水平下降。

图 12 为不同涵道比下氧气、二氧化碳和水的质量分数分布情况。可以发现, 涵道比增大, 中心锥后方的富氧带长度缩短, 二氧化碳和水的含量增大, 这是燃烧室燃烧增强的结果。燃烧室高温燃气携

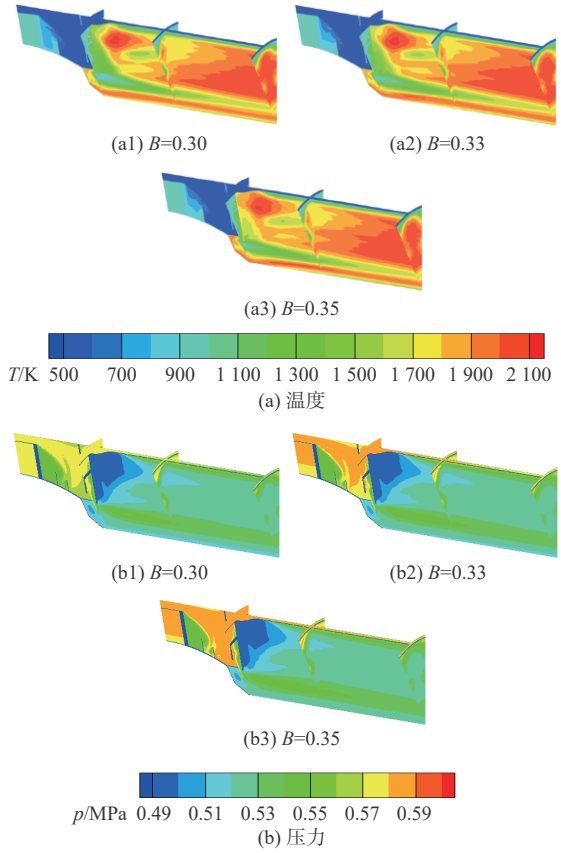


图 10　不同涵道比下加力燃烧室温度和压力分布情况

Fig. 10　Temperature and pressure distributions of afterburner under various bypass ratios

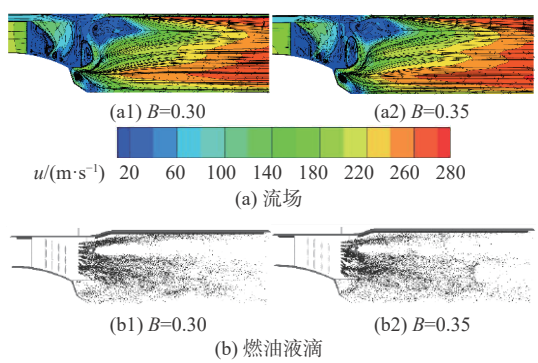


图 11　不同涵道比下加力燃烧室流场和燃油液滴分布情况

Fig. 11　Flow field and fuel droplets distributions of afterburner under various bypass ratios

带燃烧产物经中心锥回流区更多的进入富氧带下方。

图 13 给出了不同涵道比下湍动能、总压恢复系数、温度均匀性系数和燃烧效率的沿程分布情况。随着涵道比增加, 燃烧室前部燃气湍动能增大, 说明油气掺混增强, 燃烧得到增强且燃烧产物膨胀加速导致湍流强度增大。涵道比提高, 总压恢复系数下降, 其源于燃气掺混损失和燃烧热阻损失的加剧。涵道比增大加剧燃烧室前部燃烧、削弱中

后部燃烧，故温度均匀系数先随涵道比增加而增大，后随涵道比增加而减小，改变了燃烧室温度径向均匀程度。涵道比增加，燃烧室中部燃烧效率增加，而临近尾喷口燃烧效率基本不受影响。

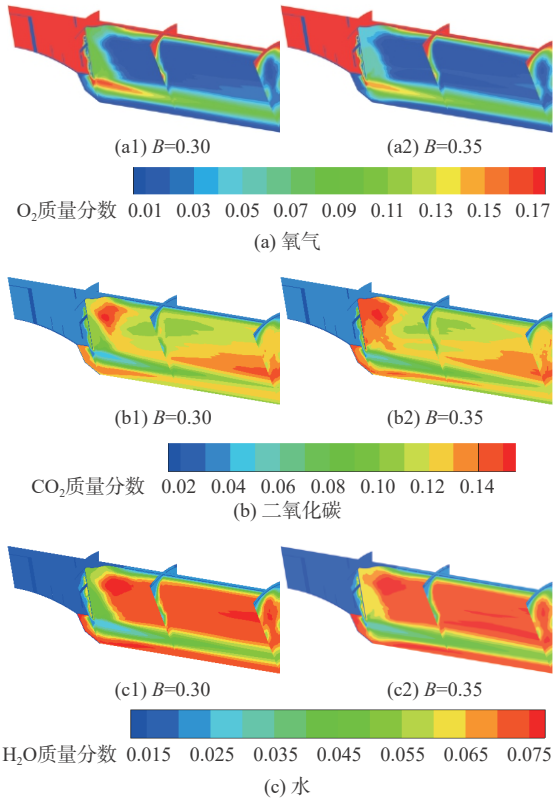


图 12 不同涵道比下氧气、二氧化碳和水的分布情况

Fig. 12 Distributions of oxygen, carbon dioxide and water under various bypass ratios

2.3 基于机器学习的燃烧性能预测

通过机器学习对一体化加力燃烧室的燃烧性能进行预测，可以在实际飞行工况中迅速完成对当前加力燃烧室燃烧状况的定量评估，可以及时调整加力燃烧室运行参数实现最优的燃烧性能，以适应航空发动机的动力需求。机器学习选择人工神经网络（artificial neural network, ANN）。人工神经网络有 2 个关键的参数：隐藏层数和隐藏层的神经元个数。根据文献 [24] 和文献 [25]，确定隐藏层数为 7，隐藏层的神经元个数为 16。加力燃烧室燃烧性能预测的人工神经网络结构如图 14 所示，输入参数为内涵进气压力 p、内涵进气温度 T、油气比 M、涵道比 B、轴向距离 x，输出参数为总压恢复系数 σ、温度均匀性系数 ξ、燃烧效率 η。将数值模拟的 136 组数据按照 75%、25% 的比例随机划分为训练集和测试集，训练集用以训练人工神经网络模型，挖掘输入输出之间的映射关系，而测试集用来检测模型的泛化能力。人工神经网络模型使用 Pytorch 搭建，损失函数选择均方误差损失（mean-square

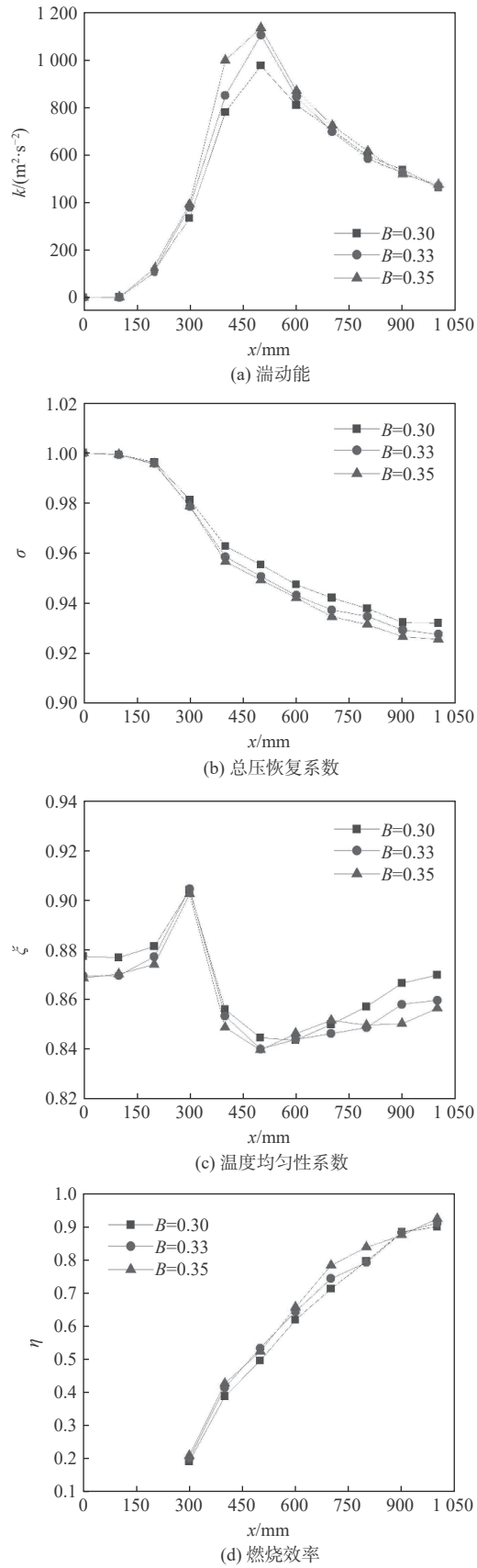


图 13 不同涵道比下湍动能、总压恢复系数、温度均匀性系数和燃烧效率的轴向变化情况

Fig. 13 Turbulent kinetic energy, total pressure recovery coefficient, temperature uniformity coefficient, and combustion efficiency axial distributions under various various bypass ratios

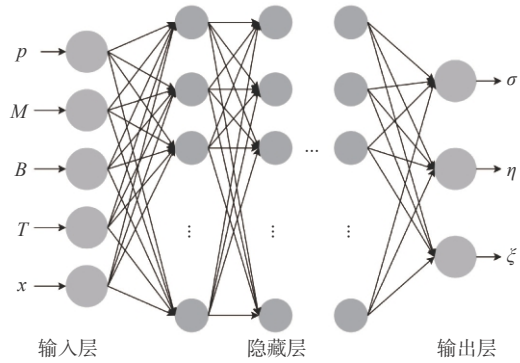


图 14　人工神经网络结构

Fig. 14　Structure of ANN

error loss, MSE loss), 激活函数为 sigmoid, 优化器选择自适应矩估计。图 15为训练过程中均方误差损失随训练次数的变化情况。第 1 081 次时 MSE loss 达到 0.01, 第 2 315 次时 MSE loss 达到 0.005。

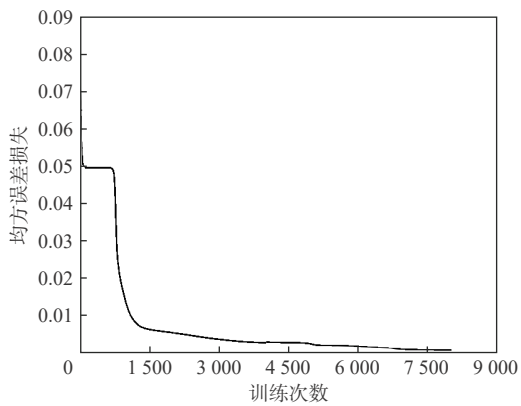


图 15　均方误差损失随训练次数的变化情况

Fig. 15　Variations of MSE loss with training times

为消除量纲单位差异的影响, 同时提高模型的训练速度和精度, 对所有数据进行归一化处理, 归一化方法如下:

$$s' = \frac{s - s_{\min}}{s_{\max} - s_{\min}} \tag{4}$$

式中: s'为归一化处理后的数据; s 为原始数据; $s_{\max}$ 为最大值; $s_{\min}$ 为最小值。

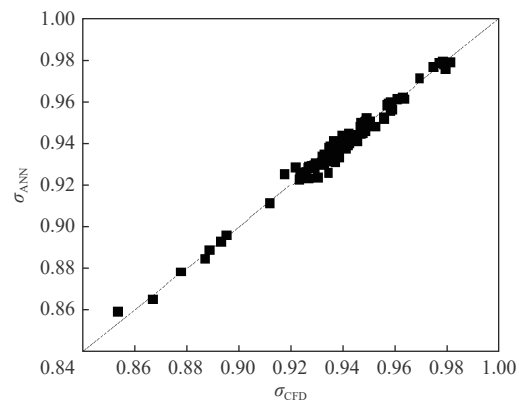
通过决定系数 R^2 和校正决定系数R^2_{Adj}共同评估训练所得的人工神经网络模型同数值结果之间的拟合优度。两者的定义如下:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - y_s)^2}{\sum_{i=1}^{n}(y_s - \overline{y_s})^2} \tag{5}$$

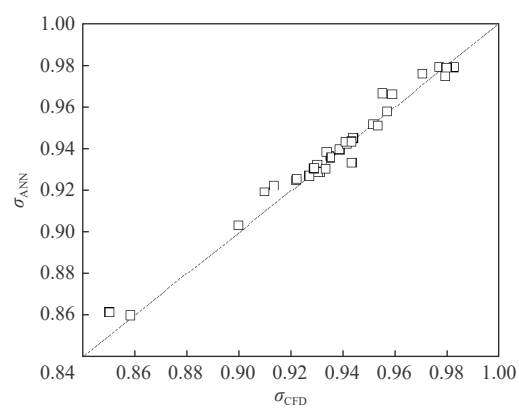
$$R^2_{\text{Adj}} = 1 - \frac{(1-R^2)(n-1)}{n-q-1} \tag{6}$$

式中: n 为样本数量; y_i 为预测值; y_s 为数值结果; $\overline{y_s}$为数值结果的平均值; q 为特征数。R^2_{Adj}正常的取值范围为 [0,1], 其值越接近 1 说明拟合效果越好, 当其值大于 0.9 时说明模型效果良好, 其值大于 0.7 时表明模型的拟合效果可以接受[26]。

图 16～图 18 依次为训练集和测试集总压恢复系数、温度均匀性系数和燃烧效率的数值结果(σ_{CFD}、ξ_{CFD}、η_{CFD})与人工神经网络预测结果(σ_{ANN}、ξ_{ANN}、η_{ANN})对比情况。可以看到, 总压恢复系数训练集和测试集的决定系数分别为 0.987 和 0.974, 对应的校正决定系数分别为 0.907 和 0.939; 温度均匀性系数训练集和测试集的决定系数分别为 0.978 和 0.905, 对应的校正决定系数分别为 0.844 和 0.788; 燃烧效率训练集和测试集的决定系数分别为 0.997 和 0.972, 对应的校正决定系数分别为 0.975 和 0.936。人工神经网络模型训练集和测试集的R^2_{Adj}均大于 0.7, 对总压恢复系数和燃烧效率这 2 个输出均大于 0.9, 对温度均匀性系数的适应性不足, 主要是温度均匀性系数的变化规律非单调, 而且样本数量不足够。整体而言, 该模型具有较好的使用价值, 可以较好用于一体化加力燃烧室的热态流场和燃烧性能预测。



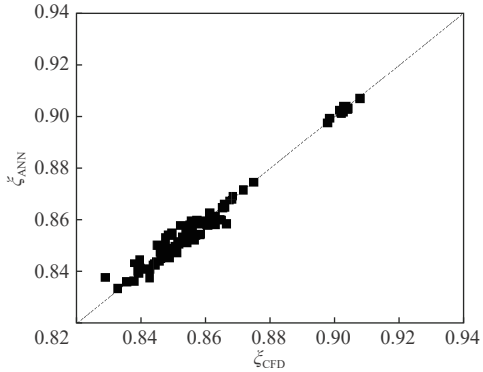
(a) 训练集, R^2=0.987, R^2_{Adj}=0.907



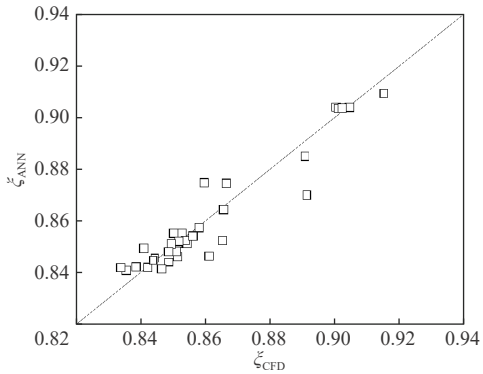
(b) 测试集, R^2=0.974, R^2_{Adj}=0.939

图 16　总压恢复系数数值结果与预测值的对比情况

Fig. 16　Comparison between numerical results and predicted values of total pressure recovery coefficient



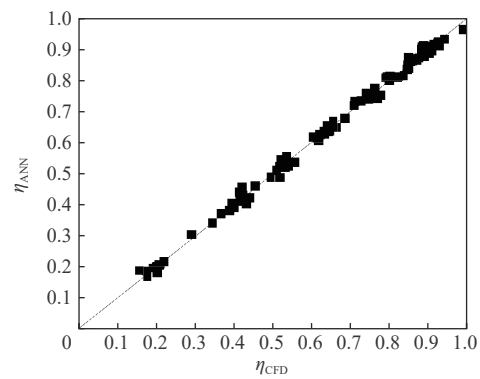
(a) 训练集, R^2=0.978, R^2_{Adj}=0.884



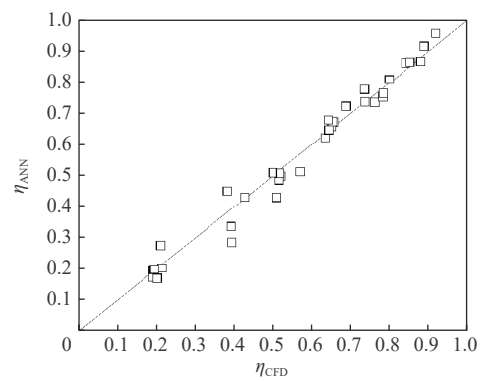
(b) 测试集, R^2=0.905, R^2_{Adj}=0.788

图 17 温度均匀性系数数值结果与预测值的对比情况

Fig. 17 Comparison between numerical results and predicted values of temperature uniformity coefficient



(a) 训练集, R^2=0.997, R^2_{Adj}=0.975



(b) 测试集, R^2=0.972, R^2_{Adj}=0.936

图 18 燃烧效率数值结果与预测值的对比情况

Fig. 18 Comparison between numerical results and predicted values of combustion efficiency

3 结 论

基于大涡模拟对一体化加力燃烧室的燃烧特性进行探究，着重分析喷油量和涵道比对燃烧的影响，通过机器学习对燃烧性能进行了预测，得到如下结论：

1）高温燃气占据燃烧室大部分空间，燃烧室观察到高温火焰下方的椭圆形低温区和中心锥后方的尖角形低温区，整体的热态性能较好。燃烧室存在,3 个典型的低速区，分别位于强火焰区、喷油杆中下部和中心锥尾部。

2）油气比增加，燃烧室温度下降，热态燃烧性能减弱。高喷油量下椭圆形低温区不断拓展甚至下方又出现附加低温区，中心锥后方的低温区受到压缩。油气比增加，燃烧室温度更加均匀，燃烧效率下降。

3）涵道比增大，燃烧室中后部高温区范围压缩且高流速向上游延伸，这是因为燃烧室前部燃油液滴量增多燃烧增强，中后部燃油液滴量减少燃烧削弱，主要起到冷却作用。随着涵道比增加，燃烧室前部燃气湍动能增大，中后部总压恢复系数下降，燃烧效率增大。

4）以内涵进气压力、内涵进气温度、喷油量、涵道比、轴向距离为输入变量，基于人工神经网络的机器学习模型建立总压恢复系数、温度均匀性系数、燃烧效率的预测模型，总压恢复系数和燃烧效率的校正决定系数大于 0.9，温度均匀性系数的校正决定系数大于 0.7，可以较好用于一体化加力燃烧室的热态流场和燃烧性能预测。

参考文献（References）

[1] 吴云柯, 李志强, 杨谦, 等. 折叠 V 形火焰稳定器原型的冷态与稳燃特性研究[J]. 推进技术, 2021, 42(6): 1321-1328.
WU Y K, LI Z Q, YANG Q, et al. Cold flow and combustion characteristics of a prototype folded-V flame holder[J]. Journal of Propulsion Technology, 2021, 42(6): 1321-1328(in Chinese).

[2] 高鹏举, 岳晨, 范育新, 等. 喷水射流预冷对加力燃烧室特性影响分析[J]. 推进技术, 2023, 44(10): 121-130.
GAO P J, YUE C, FAN Y X, et al. Effects analysis of water spray mass injection pre-compressor cooling on afterburner characteristics[J]. Journal of Propulsion Technology, 2023, 44(10): 121-130(in Chinese).

[3] 张海涛, 赵坚行. 带隔热屏加力燃烧室热态流场计算[J]. 航空动力学报, 2007, 22(8): 1241-1246.
ZHANG H T, ZHAO J X. Turbulent combustion flow field calculation for afterburner with heat shied[J]. Journal of Aerospace Power, 2007, 22(8): 1241-1246(in Chinese).

[4] 孙晓峰, 张光宇, 王晓宇, 等. 航空发动机燃烧不稳定性预测及控制研究进展[J]. 航空学报, 2023, 44(14): 1-23.

SUN X F, ZHANG G Y, WANG X Y, et al. Research progress in aero-engine combustion instability prediction and control[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(14): 1-23(in Chinese).

[5] 李安琦, 刘勇, 张祥. 加力燃烧室湍流火焰面模型优化方法[J]. 航空动力学报, 2023, 38(5): 1083-1089.
LI A Q, LIU Y, ZHANG X. Optimization method of turbulent flamelet model in afterburner[J]. Journal of Aerospace Power, 2023, 38(5): 1083-1089(in Chinese).

[6] 刘友宏, 杜力伟. 新型热混合效率公式及其在一体化加力燃烧室中的应用[J]. 推进技术, 2018, 39(7): 1515-1522.
LIU Y H, DU L W. A new thermal mixing efficiency formula and its application in integrated afterburner[J]. Journal of Propulsion Technology, 2018, 39(7): 1515-1522(in Chinese).

[7] JIA X Z, SHAN Y, TAN X M, et al. Numerical investigation of effects of cooling structure parameters on performance of flameholder in an integrated afterburner[J]. Aerospace Science and Technology, 2022, 121: 107378.

[8] LI W, TAN X M, HUANG X F, et al. Application of double-wall cooling structure in the integrated strut flame stabilizer[J]. Thermal Science and Engineering Progress, 2022, 36: 101526.

[9] 韦裕恒, 谭晓茗, 黄晓锋, 等. 带前缘冲击的一体化加力支板内外耦合传热数值研究[J]. 推进技术, 2023, 44(5): 149-162.
WEI Y H, TAN X M, HUANG X F, et al. Numerical investigation on internal and external conjugate heat transfer characteristics of integrated strut with leading edge impingement[J]. Journal of Propulsion Technology, 2023, 44(5): 149-162(in Chinese).

[10] HUANG Y K, HE X M, JIN Y, et al. Effect of non-uniform inlet profile on the combustion performance of an afterburner with bluff body[J]. Energy, 2021, 216: 119142.

[11] 尹成茗, 张荣春, 樊未军, 等. 一种一体化加力燃烧室的数值模拟[J]. 航空动力学报, 2018, 33(2): 470-476.
YIN C M, ZHANG R C, FAN W J, et al. Numerical simulation on a type of the integrated afterburner[J]. Journal of Aerospace Power, 2018, 33(2): 470-476(in Chinese).

[12] 谭云川, 王亚军, 林建府, 等. 不同燃油分配的一体化加力燃烧室性能对比[J]. 推进技术, 2022, 43(12): 249-256.
TAN Y C, WANG Y J, LIN J F, et al. Comparison of integrated afterburner performance for different fuel distribution[J]. Journal of Propulsion Technology, 2022, 43(12): 249-256(in Chinese).

[13] 王伟龙, 金捷, 井文明, 等. 改进型一体化加力燃烧室方案的数值模拟[J]. 航空动力学报, 2015, 30(5): 1119-1124.
WANG W L, JIN J, JING W M, et al. Numerical simulation on improved integrated afterburner scheme[J]. Journal of Aerospace Power, 2015, 30(5): 1119-1124(in Chinese).

[14] 付垚, 朱健, 高源, 等. 喷嘴布局对加力燃烧室燃烧性能的影响[J]. 热科学与技术, 2023, 22(1): 13-20.
FU Y, ZHU J, GAO Y, et al. Effects of the nozzle layout on the afterburner combustion performance[J]. Journal of Thermal Science and Technology, 2023, 22(1): 13-20(in Chinese).

[15] 郑健, 王海鸥, 陈果, 等. 基于机器学习的边界层湍流燃烧壁面模型[J]. 工程热物理学报, 2023, 44(9): 2600-2607.
ZHENG J, WANG H O, CHEN G, et al. Machine learning-based wall model for turbulent boundary layer combustion[J]. Journal of Engineering Thermophysics, 2023, 44(9): 2600-2607(in Chinese).

[16] 安健, 陈宇轩, 苏星宇, 等. 机器学习在湍流燃烧及发动机中的应用与展望[J]. 清华大学学报 (自然科学版), 2023, 63(4): 462-472.
AN J, CHEN Y X, SU X Y, et al. Applications and prospects of machine learning in turbulent combustion and engines[J]. Journal of Tsinghua University (Science and Technology), 2023, 63(4): 462-472 (in Chinese).

[17] AN J, HE G Q, LUO K H, et al. Artificial neural network based chemical mechanisms for computationally efficient modeling of hydrogen/carbon monoxide/kerosene combustion[J]. International Journal of Hydrogen Energy, 2020, 45(53): 29594-29605.

[18] YAO S, WANG B, KRONENBURG A, et al. Conditional scalar dissipation rate modeling for turbulent spray flames using artificial neural networks[J]. Proceedings of the Combustion Institute, 2021, 38(2): 3371-3378.

[19] WONG K I, WONG P K, CHEUNG C S, et al. Modeling and optimization of biodiesel engine performance using advanced machine learning methods[J]. Energy, 2013, 55(C): 519-528.

[20] KUNDU K, DEUR J. A simplified reaction mechanism for calculation of emissions in hydrocarbon (Jet-A) combustion[C]//Proceedings of the 29th Joint Propulsion Conference and Exhibit. Reston: AIAA, 1993.

[21] 莫毅, 陈璠, 许笑颜, 等. 航空发动机燃烧室两相湍流燃烧建模与仿真[J]. 清华大学学报 (自然科学版), 2023, 63(4): 670-680.
MO Y, CHEN F, XU X Y, et al. Modeling and simulation of two-phase turbulent combustion in aeroengine combustors[J]. Journal of Tsinghua University (Science and Technology), 2023, 63(4): 670-680 (in Chinese).

[22] 向缘酝, 李伟, 刘云鹏, 等. 一体化加力燃烧室燃烧性能数值研究[J]. 热能动力工程, 2023, 38(5): 71-79.
XIANG Y Y, LI W, LIU Y P, et al. Numerical study on combustion performance of integrated afterburner[J]. Journal of Engineering for Thermal Energy and Power, 2023, 38(5): 71-79(in Chinese).

[23] 贾翔中, 单勇, 徐兴平, 等. 一体化加力燃烧室冷态流动特性数值研究[J]. 航空动力学报, 2021, 36(7): 1472-1480.
JIA X Z, SHAN Y, XU X P, et al. Numerical study on cold flow characteristics of integrated afterburner[J]. Journal of Aerospace Power, 2021, 36(7): 1472-1480(in Chinese).

[24] 王志凯, 陈盛, 范玮. 神经网络宽度对燃烧室排放预测的影响[J]. 航空学报, 2023, 44(5): 82-92.
WANG Z K, CHEN S, FAN W. Effect of neural network width on combustor emission prediction[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(5): 82-92(in Chinese).

[25] 王超, 王贵东, 白鹏. 飞行仿真气动力数据机器学习建模方法[J]. 空气动力学学报, 2019, 37(3): 488-497.
WANG C, WANG G D, BAI P. Machine learning method for aerodynamic modeling based on flight simulation data[J]. Acta Aerodynamica Sinica, 2019, 37(3): 488-497(in Chinese).

[26] 韩雷. 基于 BP 神经网络的学术论文评价模型研究[J]. 现代情报, 2024, 44(2): 170-177.
HAN L. Research on academic paper evaluation model based on BP neural network[J]. Modern Information, 2024, 44(2): 170-177(in Chinese).

Simulation of combustion characteristics and prediction of combustion performance using machine learning in an integrated afterburner

ZHANG Yu[1]，WANG Fengming[2, 3]，WANG Yanhong[4]，MU Lin[1]，DONG Ming[1, 3, *]

(1. School of Energy and Power Engineering，Dalian University of Technology，Dalian 116024，China；
2. School of Energy and Power，Nanjing University of Aeronautics and Astronautics，Nanjing 210016，China；
3. Advance Jet Propulsion Creativity Center，Beijing 101300，China；
4. School of Energy and Power Engineering，Northeast Electric Power University，Jilin 132012，China)

Abstract: An afterburner design concept that integrated the fuel injection pipe, radial flame stabilizer, and turbine rear frame support plate was suggested as a solution to the problem of increasing the thrust-to-weight ratio of aero-engine. The thermal flow field and combustion characteristics of the afterburner under different kerosene-air ratios and bypass ratios were investigated through large eddy simulation. Distribution characteristics of the temperature, velocity, and pressure in the combustion chamber and their impacts on the combustion process were studied. The distributions of fuel droplets, oxygen, carbon dioxide, and water were analyzed. Using the intake pressure, intake temperature, fuel rate, bypass ratio, and axial distance as input variables, the machine learning method was developed and predictions were made for the total pressure recovery coefficient, temperature uniformity coefficient, and combustion efficiency. The results indicate that the overall combustion performance of the integrated afterburner is high, and only a localized combustion weakening zone is observed below the flame. There are three low-velocity recirculation zones in front of the combustion chamber, located in the strong flame zone, the middle and lower parts of the fuel injection pipe, and the tail of the central cone. As the kerosene-air ratio increases, the combustion efficiency is decreased, while an increase in bypass ratio leads to an improvement in combustion efficiency. The proposed machine learning model has a corrected determination coefficient greater than 0.788 for both the training and testing datasets, indicating good predictive performance.

Keywords: integrated afterburner; thermal flow field; recirculation zone; large eddy simulation; machine learning; combustion performance prediction

Received：2024-04-12；**Accepted**：2024-05-24；**Published Online**：2024-07-08 09：32

URL：link.cnki.net/urlid/11.2625.V.20240705.1209.003

Foundation item：Supported by Advance Jet Propulsion Creativity Center (relying on Aero Engine Academy of China) (HKCX2022-01-017)

* **Corresponding author**. E-mail：dongming@dlut.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0235

低雷诺数柔性拍动平板非定常气动特性

曹梦达[1]，郑孟宗[1]，苏冠廷[2,*]，潘天宇[1]，李志平[1]，李秋实[1,2,3]

(1. 北京航空航天大学 航空发动机研究院，北京 100191；2. 北京航空航天大学 能源与动力工程学院，北京 100191；
3. 西华大学 航空航天学院，成都 610039)

摘　　要：柔性变形会对扑翼气动力产生影响，且对于柔性扑翼问题，伴随着严重的流固耦合效应。为揭示柔性变形对非定常气动力的影响机理，开展低雷诺数柔性拍动平板流固耦合气动特性研究。将蜻蜓运动规律赋予柔性扑翼及刚性扑翼并开展流固耦合数值模拟计算，在相同运动规律下，柔性扑翼相对于刚性扑翼在整个周期的平均升力提高60.5%，在下拍初期升力下降31.4%，下拍中后期升力提高76.7%；柔性扑翼与刚性扑翼在整个周期内产生的推力几乎为0。揭示了柔性变形对气动力的时空影响机制，柔性扑翼在下拍阶段产生的展向弯曲变形能够维持扑翼表面前缘涡的展向分布而提高升力；柔性扑翼在拍动初期会产生下拍滞后，不利于前缘涡的形成而使升力降低，随着下拍进行，展向弯曲变形能够提高下拍速度进而影响前缘涡结构而提高升力。

关 键 词：柔性扑翼；蜻蜓；流固耦合；升力系数；前缘涡
中图分类号：V221^{+}.3; TB553
文献标志码：A　　**文章编号**：1001-5965(2026)06-2011-13

相对于传统大型飞行器，微型飞行器(micro aeria vehicle, MAV)因其小尺寸的优势，可以高效地在更为狭窄、复杂的空间中执行任务，因此，微型飞行器在未来数字化、信息化战场中的地位与日俱增。当前国内外兴起研制微小型飞行器的热潮，由于微型飞行器的特征尺寸急剧减小，其飞行时所处的低雷诺数环境下，空气黏性对飞行器的影响越来越明显，此时传统固定翼和旋翼已无法满足其设计要求。在自然界中，昆虫无论是尺寸，还是飞行雷诺数都与微型飞行器相当，且采用扑翼飞行方式，其通过翅膀的快速拍动、翻转，利用低雷诺数下空气的黏性在翅膀表面产生复杂的涡结构而产生气动力，进而实现其起动、前飞、悬停等飞行状态。针对昆虫扑翼飞行的气动特性的研究，Weis-Fogh[1]通过对台湾小黄蜂翅膀的开合情况进行分析，提出Weis-Fogh机制。Ellington等[2]在进行鹰蛾及其翅膀拍动的实验时提出“延迟失速机制”。Dickinson等[3]通过测量果蝇模型的气动力发现了“尾迹捕获机制”。可见扑翼飞行具有显著的非定常性。

昆虫翅膀是由翅脉、翅膜等组件构成的复杂的各向异性结构，早期研究者利用高速摄像机对蚊类、食蚜蝇、蜂类、蛾类等昆虫扑翼过程中的翅膀进行拍摄研究[4-6]，发现昆虫在扑翼过程中存在的柔性变形主要为沿展向的扭转变形和沿弦向的弯曲变形。研究发现，昆虫翅膀的柔性变形对其气动力会产生影响，Nguyen等[7]利用流固耦合方法研究发现，当前飞速度较小时，柔性翅膀相比较刚性翅膀具有更高的气动效率和更小的气动功。但随着飞行速度的增加，由于柔性翼的变形会增强下拍开始时产生的向上气流，柔性翼所需的气动功会超过刚性翼。Nakata等[8]通过耦合计算结构动力学模型及昆虫动力学飞行模拟器而建立了昆虫的柔性翅

收稿日期：2024-04-19；**录用日期**：2024-07-29；**网络出版时间**：2024-08-06 09:40
网络出版地址：link.cnki.net/urlid/11.2625.V.20240805.1816.004
基金项目：国家自然科学基金 (52206040,52006002,51976005)
* **通信作者**. E-mail：suguanting@buaa.edu.cn

引用格式：曹梦达，郑孟宗，苏冠廷，等. 低雷诺数柔性拍动平板非定常气动特性 [J]. 北京航空航天大学学报，2026，52（6）：2011-2023.
CAO M D，ZHENG M Z，SU G T，et al. Unsteady aerodynamic characteristics of flexible flapping plate at low Reynolds numbers [J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（6）：2011-2023 （in Chinese）.

流固耦合模型，对鹰蛾翅膀的柔性变形进行数值模拟，发现柔性变形能够使得尾流下洗速度提高从而延迟前缘涡的脱落，进而提高气动力。周超英等[9]针对二维柔性扑翼模型，在耦合柔性变形的基础上数值求解雷诺平均 Navier-Stokes 方程，发现适当的柔性变形可以减缓前缘涡的脱落从而使得推进效率得以提升且对扑翼产生高升力带来不利影响。孟令兵等[10]开发了一套基于计算固体力学/计算流体力学(computational structural dynamics/ computational fluid dynamics, CSD/CFD)的流固耦合数值模拟算法，用于计算蜻蜓前飞的气动力，证实了柔性变形可以提高推力而使阻力减小。陈利丽等[11]通过求解结构动力学方程及雷诺平均应力方程，发现柔性变形可以提高扑翼的推力特性。王奇等[12]通过建立柔性椭圆翼型的流固耦合模型，认为柔性变形可以延长尾缘涡脱落的时间，以及能够有效抑制尾流流场的扰动，从而降低升力扰动的振荡频率与幅值。综上，柔性变形确实能够影响扑翼气动特性，且对于柔性扑翼问题，伴随着严重的流固耦合效应。针对低雷诺数情况下流固耦合机制的研究，目前，国内外刚刚起步，这些研究还都是分散性质的研究，对象较为单一和特殊，尤其在国内更缺少针对柔性变形对非定常气动力的影响机理的探究。因此，亟需开展系统的低雷诺数流固耦合机制的数值研究。

针对上述问题，本文选择更具普适性的柔性平板模型作为研究对象，开展低雷诺数下不同柔性平板流固耦合气动特性的研究。从仿生学的角度出发，首先，依据蜻蜓翅膀设计平板扑翼模型，模型运动规律参考现实蜻蜓观测数据来模拟其悬停飞行状态，进而开展流固耦合数值模拟，将柔性扑翼产生的气动力与刚性扑翼进行对比，针对二者气动力的差别，对其附近流场进行分析，从而得出柔性变形对非定常气动力的影响机制。

1 模型与方法

1.1 扑翼模型及运动规律

蜻蜓作为自然界中的飞行冠军，相较于其他昆虫翅膀尺寸较大，约为 3～5 cm，因其具有大展弦比翅膀，拍动时柔性变形更加明显。这更有利于分析柔性变形对扑翼非定常气动特性的影响机理。蜻蜓翅膀由翅脉、翅膜、翅痣等复杂结构组成，为使结论更具有普适性，在构建扑翼模型时对蜻蜓翅膀进行简化，忽略蜻蜓翅膀的各向异性结构，而采用各向同性的平板扑翼模型来进行研究。将平板模型展长设为 4 cm，弦长设为 1 cm(蜻蜓翅膀展弦比约为 4)，厚度设为 0.02c(c 为平板模型弦长)，平板扑翼模型如图 1 所示。

图 1 平板扑翼模型

Fig. 1 Flat plate flapping wing model

接下来，对平板扑翼模型的运动规律进行定义。图 2 为欧拉角及拍动角示意图，在随体坐标系 $Oxyz$ 中，蜻蜓翅膀运动规律可通过 3 个欧拉角(拍动角α、翻转角β、偏离角γ)及拍动平面角θ确定[13]，其中，拍动平面为翅尖拍动轨迹的连线与翅根所定义的平面，拍动平面角θ为翅膀的拍动平面与水平面的夹角，拍动角α为翅尖与翅根的连线在拍动平面的投影与 z 轴的夹角，翻转角β为拍动平面与翅膀所在平面的夹角，偏离角γ为翅尖与翅根连线与其在拍动平面投影的夹角，其中，偏离角由于变化无规律且数值较小，忽略其影响。扑动频率及拍动平面角参考现实对蜻蜓进行观测得到的数据[14-16]，即扑动频率给定 35 Hz(一个拍动周期经历的时间 t_c=1/35 s)，拍动平面角为 60°。给定拍动角变化范围为 40°～−20°。故给定平板扑翼模型拍动角运动规律为

$$\alpha(T_0) = 10^\circ + 30^\circ \cos(2\pi T_0) \tag{1}$$

式中：T_0为无量纲时间，定义其为物理时间t与整个拍动周期经历的时间t_c之比。

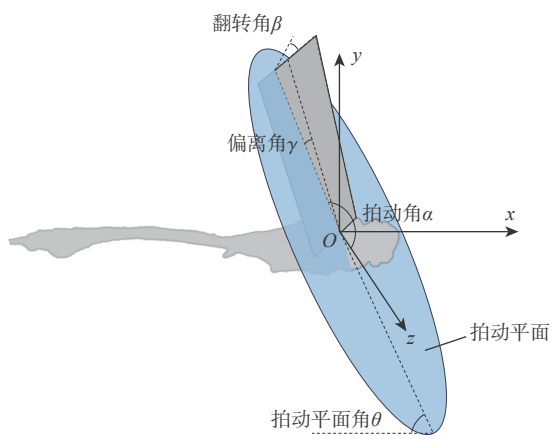


图 2 欧拉角及拍动平面角示意图

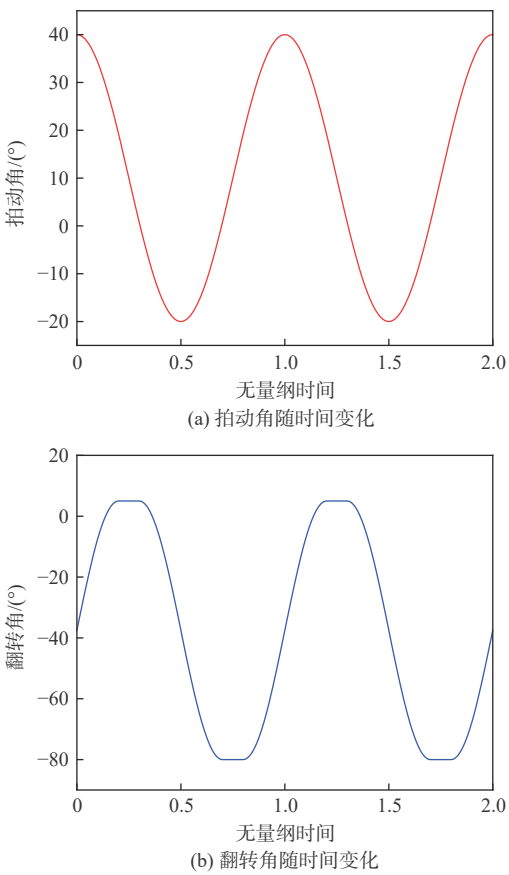
Fig. 2 Diagram of Euler angle and flapping plane angle

根据文献 [17-18] 的观测数据可知，由于翻转角在上拍及下拍过程中各有一段时间近似不变，其余时间近似以正弦函数变化，其变化规律为

$$\beta(T_0)=\begin{cases}(\beta_u+\beta_d)/2+(\beta_d-\beta_u)/2\sin(\pi T_0/\Delta T_0) & T_0\in(-\Delta T_0/2,\Delta T_0/2)\\ \beta_d & T_0\in(\Delta T_0/2,1-\Delta T_0/2)\\ (\beta_u+\beta_d)/2+(\beta_d-\beta_u)/2\sin\left(\pi\dfrac{2T+\Delta T_0-1}{2\Delta T_0}\right) & T_0\in((1-\Delta T_0)/2,(1+\Delta T_0)/2)\\ \beta_u & T_0\in((1+\Delta T_0)/2,1-\Delta T_0/2)\end{cases} \tag{2}$$

式中: 下拍翻转角$\beta_d=5°$; 上拍翻转角$\beta_u=-80°$; 无量纲翻转时间$\Delta T_0=0.4$。

经过以上定义, 拍动角、翻转角变化规律如图 3 所示。



(a) 拍动角随时间变化

(b) 翻转角随时间变化

图 3　模型运动规律

Fig. 3　Motion pattern of model

雷诺数 Re 是表征流体流动状况的无量纲数, 其在物理上可表示为惯性力与黏性力之比。本文将雷诺数定义为

$$Re=\rho U_{\mathrm{ref}}L_{\mathrm{ref}}/\mu \tag{3}$$

式中: ρ为流体密度; μ为流体动力黏性系数, U_{ref}为参考速度; L_{ref}为特征长度。选择扑翼模型的弦长 c 作为 L_{ref}。将 U_{ref}定义为平板扑翼 0.5 倍展长处的平均速度, 即: $U_{\mathrm{ref}}=2\varphi fr$ (φ为弧度制的拍动角幅度), r为平板扑翼 0.5 倍展长处的旋转半径。计算可得平板扑翼模型雷诺数 $Re=1\,505$, 而实际观测中蜻蜓悬停雷诺数为 1 000～2 000, 所得雷诺数与实际观测数值相符[19]。

在研究平板扑翼模型产生的气动力时, 主要考虑升力及推力, 将升力和推力分别定义为垂直于和平行于水平面的气动力, 将升力系数和推力系数分别定义为

$$\begin{cases}C_L=\dfrac{F_L}{0.5\rho S U_{\mathrm{ref}}^2}\\ C_T=\dfrac{F_T}{0.5\rho S U_{\mathrm{ref}}^2}\end{cases} \tag{4}$$

式中: F_L和F_T分别为升力和推力, 其分别为每个时刻平板扑翼表面压力在竖直、水平方向的积分; S为扑翼表面积。

1.2　流固耦合数值模拟方法

本文采用流体力学求解器 Xflow 2019 和结构力学求解器 Abaqus 2019 协同仿真进行双向流固耦合计算。

基于格子-玻尔兹曼方法(lattice-Boltzmann method, LBM)的流体求解器 Xflow 具有处理复杂边界和并行效率等方面的优势, 不同于大多数 CFD 求解器基于雷诺平均 Navier-Stokes 方程从宏观层面描述流体运动, Xflow 采用 LBM, 基于微观尺度模型, 将空间离散为笛卡儿分布的格子, 每个格子上沿着不同方向的速度$\boldsymbol{e}_i(i=1,2,\cdots,b)$。连续空间中 Boltzmann 输运方程:

$$\frac{\partial f(\boldsymbol{r},t)}{\partial t}+\boldsymbol{e}_i\cdot\nabla f_i(\boldsymbol{r},t)=\boldsymbol{\Omega}_i\qquad i=1,2,\cdots,b \tag{5}$$

式中: $f_i(\boldsymbol{r},t)$表示在时间t速度$\boldsymbol{e}_i$的粒子在位置$\boldsymbol{r}$处的分布; $\boldsymbol{\Omega}_i$为碰撞算子, 计算质量守恒和动量守恒下的碰撞。该方程在每个格子上沿不同方向速度$\boldsymbol{e}_i(i=1,2,\cdots,b)$将方程离散为

$$f_i(\boldsymbol{r}+\boldsymbol{e}_i\Delta t,t+\Delta t)=f_i(\boldsymbol{r},t)+\boldsymbol{\Omega}_i(f_1,f_2,\cdots,f_b) \tag{6}$$

如同在连续的 Boltzmann 方程中一样, 宏观变量, 如密度ρ和速度$\boldsymbol{v}$, 可以计算为密度函数的矩:

$$\rho=\sum_{i=1}^{b}f_i \tag{7}$$

$$\rho\boldsymbol{v}=\sum_{i}^{b}f_i\boldsymbol{e}_i \tag{8}$$

LBM 网格划分简单, 并行效率高, 可处理复杂边界问题[20]。而结构求解器 Abaqus 在模拟翅膀的力学变形方面的准确性也已得到验证[21], 其基于中心差分法, 对运动方程进行时间显式积分, 使用一个增量的运动学条件来计算下一个增量的运动

学条件，在每个增量求解开始时，需要满足动态平衡条件，即节点质量矩阵$\boldsymbol{M}$乘以节点加速度$\ddot{\boldsymbol{u}}$等于净节点力（外部施加的力$\boldsymbol{P}$和内部单元力$\boldsymbol{I}$之间的差）：

$$\boldsymbol{M}\ddot{\boldsymbol{u}} = \boldsymbol{P} - \boldsymbol{I} \tag{9}$$

因此，当增量开始时的加速度计算为

$$\ddot{\boldsymbol{u}}\big|_{(t)} = \boldsymbol{M}^{-1}(p-1)\big|_{(t)} \tag{10}$$

使用中心差分法对加速度进行时间积分，假设加速度恒定的情况下计算速度的变化。将速度变化与前一增量中间的速度增加，以确定当前增量中间的速度，即

$$\dot{\boldsymbol{u}}\left|_{\left(t+\frac{\Delta t}{2}\right)}\right. = \dot{\boldsymbol{u}}\left|_{\left(t-\frac{\Delta t}{2}\right)}\right. + \frac{\left(\Delta t\big|_{(t+\Delta t)} + \Delta t\big|_{(t-\Delta t)}\right)}{2}\ddot{\boldsymbol{u}}\big|_{(t)} \tag{11}$$

将当前增量中间的速度进行时间积分，并与增量开始时的位移相加以确定增量结束时的位移，即

$$\boldsymbol{u}\big|_{(t+\Delta t)} = \boldsymbol{u}\big|_{(t)} + \Delta t\big|_{(t+\Delta t)}\dot{\boldsymbol{u}}\left|_{\left(t+\frac{\Delta t}{2}\right)}\right. \tag{12}$$

为建立流体求解器和结构力学求解器之间的数据交换并执行流固耦合求解，需要使用 Abaqus 协同仿真引擎（co-simulation engine, CSE）来管理 Xflow 和 Abaqus 之间的耦合工作，目的是为在流体力学数值求解器和结构力学数值求解器之间建立数据交换并进行流固耦合求解。Abaqus CSE 通过指导求解器间的数据交换，在时间上同步求解器及在流/固交接网格上映射信息来处理耦合过程。在本研究中，求解器使用显式 Gauss-Seidel 算法进行耦合，显式 Gauss-Seidel 算法是一个串行过程，因此，在耦合步骤中会出现一个物理场的求解滞后于另一个物理场的情况。Abaqus CSE 的协商方法遵循主求解器，考虑优先使用 XFlow 的时间步长作进行数据交换，使 XFlow 承担主要求解任务。XFlow 将流固耦合节点处受力负荷传递到 Abaqus，并将受到相应载荷后的位移信息从 Abaqus 传递到 XFlow，实现双向流固耦合模拟。流固耦合模拟的协同仿真流程如图 4 所示，图中，$\boldsymbol{\Omega}_i^{\mathrm{MRT}}$为碰撞算子$\boldsymbol{\Omega}_i$用多重弛豫时间（multiple relaxation time, MRT）的近似算子，$\boldsymbol{M}_{ij}$为 $b \times b$ 阶转换矩阵，$\hat{\boldsymbol{S}}_{ij}$为 $b \times b$ 阶对角松弛矩阵，$\boldsymbol{\mu}_i$为统计学中密度分布函数的统计矩阵，$\boldsymbol{\mu}_i^{\mathrm{eq}}$为达到平衡态时的统计矩阵，$k$、$l$和$m$分别为 x、y 和 z 方向上矩阵的阶数，$\boldsymbol{e}_{ix}$、$\boldsymbol{e}_{iy}$ 和$\boldsymbol{e}_{iz}$分别为 x、y 和 z 方向上的粒子离散速度，矩阵的总阶数为 $k+l+m$。

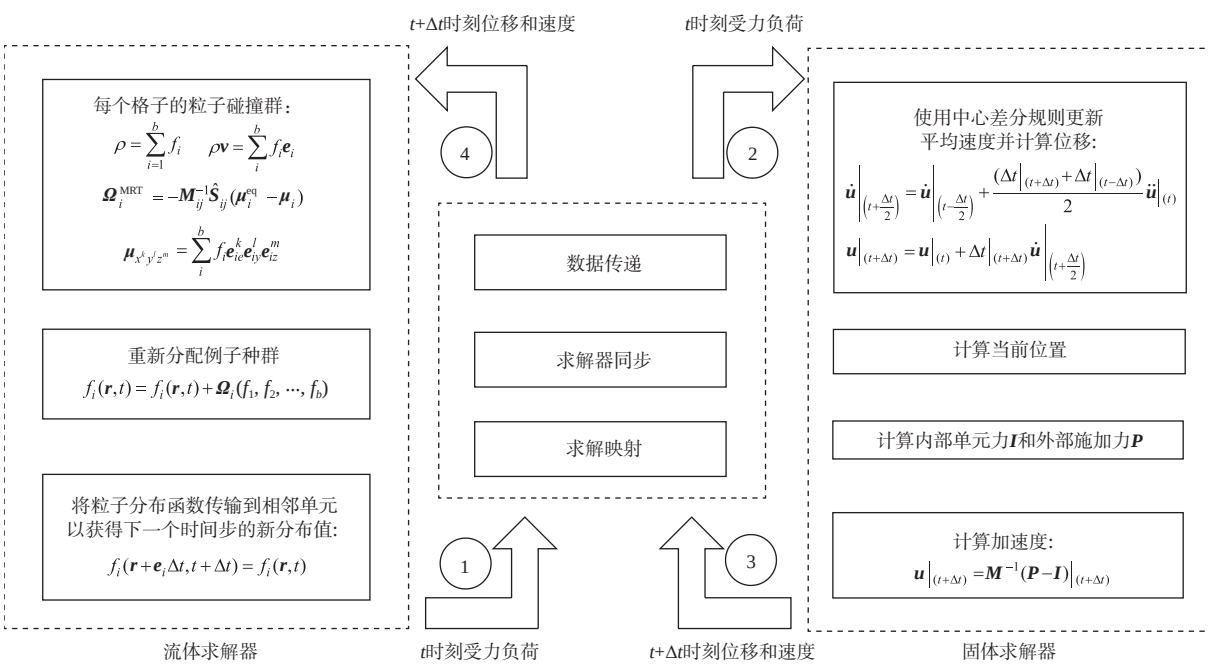


图 4 协同仿真流程

Fig. 4 Process of co-simulation

数值计算时，在固体求解器中完成对模型节点的相互作用和边界条件及模型的初始拍动位置的定义，在流体求解器中进行流体计算域、时间步、边界条件的设定及网格划分。计算域设定为 $30c \times 30c \times 15c$，周期时间步为 200 步（计算域和时间步的设定均已满足无关性验证）。将模型表面设为无滑移壁面，进口和出口设置为压力边界条件，本次算例仅对右翅进行模拟，在侧面设置 symmetric（对称边界条件）来模拟左翅的影响。计算域及边界条件如图 5 所示。

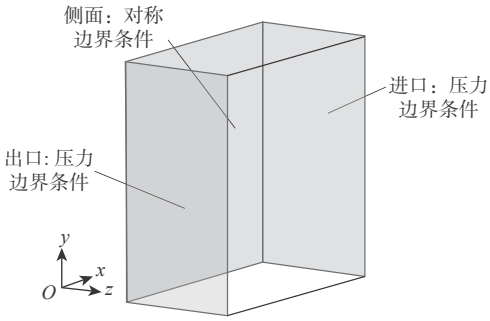


图 5　计算域及边界条件

Fig. 5　Computational domain and boundary conditions

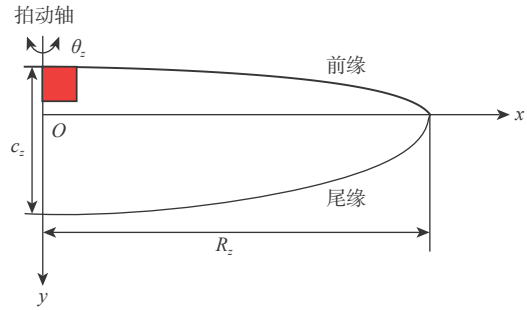


图 6　齐默尔曼机翼相关参数说明

Fig. 6　Related parameters of Zimmerman wing

1.3　流固耦合方法验证

本节将对本文所采取的流固耦合数值模拟方法的准确性进行验证。

Aono 等[22] 利用齐默尔曼机翼在一个自由度中仅用一个简单的正弦函数拍动来验证 FSI 求解器。因此, 为证明本文采用的流固耦合方法的准确性, 本节将文献 [22] 的数值结果与本求解器的数值结果进行对比。机翼平面为齐默尔曼形状, 如图 6 所示。

齐默尔曼机翼详细表达式为

$$\begin{cases}\left(\dfrac{x}{R_z}\right)^2+\left(\dfrac{y}{0.75c_z}\right)^2=1 & 0\leqslant x\leqslant R_z, 0\leqslant y\leqslant 0.75c_z\\ \left(\dfrac{x}{R_z}\right)^2+\left(\dfrac{y}{0.25c_z}\right)^2=1 & 0\leqslant x\leqslant R_z, -0.25c_z\leqslant y\leqslant 0\end{cases} \tag{13}$$

式中: c_z 为翅根处弦长, 为 25 mm; R_z 为翅膀的最大展长。翅膀厚度 h_z 为 0.4 mm, 由铝板制成, 图中, 正方形刚性区域(红色区域, 5 mm×5 mm)是翅膀自由度的约束。翅膀拍动角 θ_z 随时间的变化规律可用正弦函数 $\theta_z=\theta_A\sin(2\pi ft)$ 表示, 其中, $f=10$ Hz, θ_A 为拍动幅值。图 6 总结了翅膀的参数设置及数值验证的模拟参数。其中, 定义雷诺数 $Re=\dfrac{\theta_A f R_z c_z \rho}{\pi\mu}$, (其中, μ 为动力黏性系数, f 为拍动频率)。参考速度为翼尖平均拍动速度。

表 1 为机翼及数值模拟相关参数。

翼尖位移系数 C_{WT}(翼尖垂直位移与翼根弦长之比)随时间变化如图 7(a) 所示, 可以看出, 本文方法与文献 [22] 吻合地很好。图 7(b) 为基于翼尖速度的翅膀升力系数随时间的变化, 通过将本文方法与文献 [22] 及文献 [23] 进行对比, 可以看出, 3 种方法的计算结果差异不大, 3 种情况的平均升力系数分别为: −0.016 89、−0.016 81、−0.016 77, 一致性良好。

表 1　机翼及数值模拟相关参数

Table 1　Related parameters of wing and numerical simulation

展长 R_z/mm	弦长 c_z/mm	厚度 h_z/mm	密度/(kg·m^{-3})	泊松比	弹性模量 E/GPa	雷诺数 Re	参考速度 V_{ref}/(m·s^{-1})	动力黏性系数 μ/(Pa·s)	拍动频率 f/Hz	拍动幅值 θ_A/(°)	拍动规律
75	25	0.4	2 700	0.3	70	2 605	1.099 5	1.8×10^{-5}	10	21	正弦

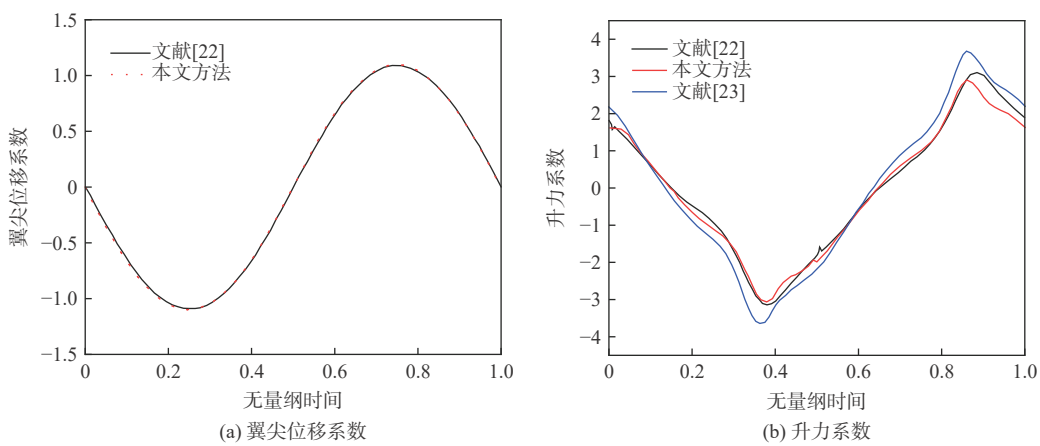


图 7　本文方法与其他方法对比

Fig. 7　Comparison between the proposed method and other methods

1.4 网格无关性验证

1.4.1 流体域网格无关性验证

流体域网格设置如图8所示。对于流体域网格进行无关性验证，最小网格尺寸选择3个尺度，从大到小分别为：0.04c、0.02c、0.01c。

以细网格的计算结果作为基准，粗网格的计算结果相对于细网格偏差8.03%；中网格的计算结果相对于细网格结果偏差2.19%，由此可见，选用粗网格计算结果偏差较大，因此，选择0.02c作为流体域的最小网格尺度，流体域网格无关性具体结果如表2和图9所示。

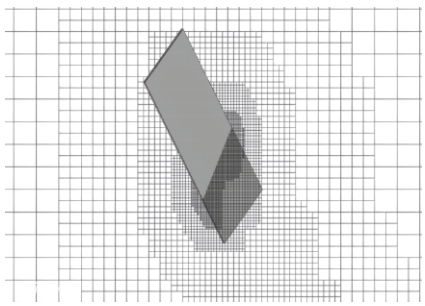


图8 流体域网格

Fig. 8 Mesh of fluid domain

表2 流体域网格尺度选择

Table 2 Selection of mesh scale in fluid domain

网格	最小网格尺度	周期平均升力系数	偏差/%
粗网格	0.04c	1.26	8.03
中网格	0.02c	1.34	2.19
细网格	0.01c	1.37	

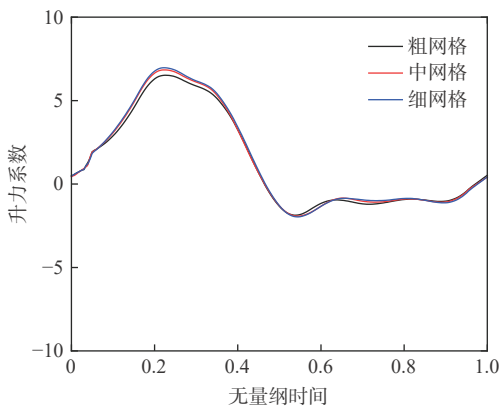


图9 流体域网格无关性验证

Fig. 9 Verification of independence of mesh in fluid domain

1.4.2 固体域网格无关性验证

固体域网格设置如图10所示。对于固体域网格进行无关性验证，网格尺寸选择3个尺度，从大到小分别为：0.06c、0.04c、0.02c。

以细网格的计算结果作为基准，粗网格的计算结果相对于细网格偏差9.19%；中网格的计算结果相对于细网格结果偏差2.57%，由此可见，选用粗网格计算结果偏差较大，因此，选择0.04c作为固体域的最小网格尺度，固体域网格无关性结果如表3和图11所示。

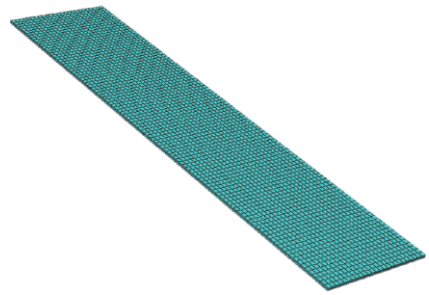


图10 固体域网格

Fig. 10 Mesh of solid domain

表3 固体域网格尺度选择

Table 3 Selection of mesh scale in solid domain

网格	最小网格尺度	周期平均升力系数	偏差/%
粗网格	0.06c	2.97	9.19
中网格	0.04c	2.65	2.57
细网格	0.02c	2.72	

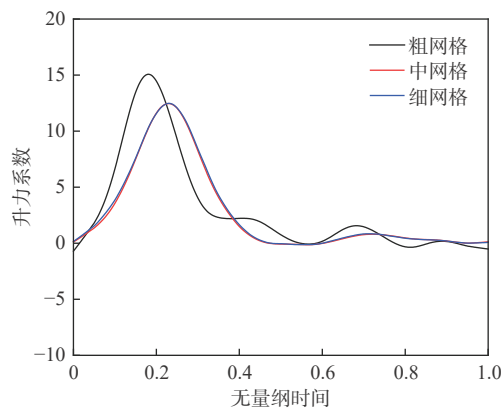


图11 固体域网格无关性验证

Fig. 11 Verification of independence of mesh in solid domain

2 柔性平板扑翼气动力

本算例为各向同性的平板模型，忽略了蜻蜓翅膀翅脉、翅膜等复杂的各向异性结构，通过调研文献[24-26]，将平板模型的密度设为1 200 kg/m^3，泊松比设为0.3，模型整体杨氏模量与翅脉刚度一致，设为40 GPa。为防止计算结果的偶然性，不改变密度和泊松比，再将平板模型刚度S_t设为80 GPa进行对照，通过计算，得出整个拍动周期内升力系数随时间变化，将柔性扑翼与刚性扑翼的升力曲线进行对比，如图12所示，其中，刚性扑翼除了将翅膀定义为刚性外，其他设定均与柔性扑翼相同。

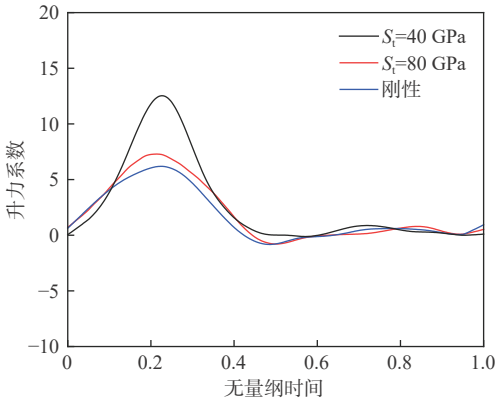


图 12　3 种扑翼模型升力系数对比

Fig. 12　Comparison of lift coefficients among three flapping wing models

平板模型从静止到下拍刚开始的过程中, 由于流场的突然变化, 会导致升力的突变, 因此, 舍弃前 4 个周期的气动力数据, 选择第 5 个周期进行分析, 在运动规律相同的情况下, 对比柔性扑翼与刚性扑翼, 在整个周期内, 柔性变形带来的影响大致分为 3 个阶段: 第 1 阶段为 T_0=0～0.12, 此时, 刚度为 40 GPa 的柔性扑翼产生的升力最小, 刚性扑翼产生的升力最大, 刚度为 80 GPa 的柔性扑翼接近且略小于刚性扑翼产生的升力, 此时柔性变形产生不利影响; 第 2 阶段为 T_0=0.12～0.5, 此时, 刚度为 40 GPa 的柔性扑翼产生的升力最大, 刚度为 80 GPa 的柔性扑翼次之, 刚性扑翼产生的升力最小, 此时柔性变形带来正向收益; 第 3 阶段为 T_0=0.5～1.0(上拍阶段), 此时三者几乎不产生升力。

图 13 为 3 种扑翼模型推力系数对比。总体来看, 柔性扑翼和刚性扑翼在下拍阶段主要产生负推力, 而在上拍阶段则主要产生正推力。计算可知, 无论是柔性扑翼还是刚性扑翼, 其周期平均推力几乎为 0, 符合悬停飞行状态。而在进行微型扑翼飞行器设计时, 升力往往是研究人员更加关注的指标, 因此, 下文主要讨论柔性变形对升力的影响。

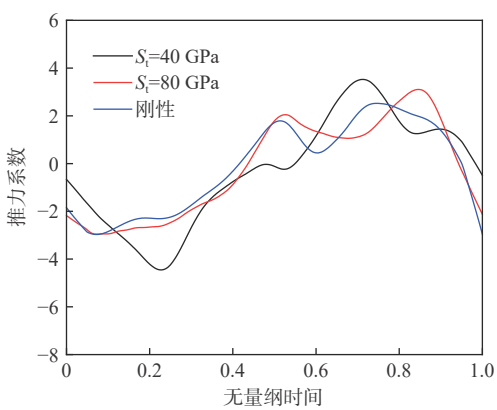


图 13　3 种扑翼模型推力系数对比

Fig. 13　Comparison of thrust coefficients among three flapping wing models

表 4 为柔性扑翼与刚性扑翼各阶段产生的平均升力系数, 当 T_0=0～0.12 时, 刚度为 40 GPa 和 80 GPa 的柔性扑翼产生的平均升力分别为 1.77 和 2.52, 相对于刚性扑翼(2.58), 分别下降了约 31.4% 和 2.3%; 当 T_0=0.12～0.5 时, 刚度为 40 GPa 和 80 GPa 的柔性扑翼产生的平均升力分别为 5.99 和 4.12, 相对于刚性扑翼(3.39), 分别提升了约 76.7% 和 21.5%; 当 T_0=0.5～1 时, 三者几乎不产生升力。对于整个周期, 刚度为 40 GPa 和 80 GPa 的柔性扑翼产生的平均升力为 2.68 和 1.97, 相对于刚性扑翼(1.67), 分别提升了约 60.5% 和 18.0%。综合来看, 柔性变形对于升力的产生具有正向收益。因此, 接下来将对柔性变形产生正向收益的气动特性进行分析。

表 4　3 种扑翼模型平均升力系数对比

Table 4　Comparison of average lift coefficient among three flapping wing models

时间	平均升力系数			柔性变形影响
	S_t=40 GPa	S_t=80 GPa	刚性	
T_0=0～0.12	1.77	2.52	2.58	不利
T_0=0.12～0.5	5.99	4.12	3.29	有利
T_0=0.5～1.0	0.30	0.17	0.20	
全周期	2.68	1.97	1.67	有利

3　柔性平板扑翼气动特性

3.1　空间维度

由于在 T_0=0.25 时刻柔性扑翼与刚性扑翼所产生的升力差别最大, 故分析 T_0=0.25 时刻涡量由翼根至翼尖的展向分布情况, 以 0.1R 为增量(R 为展长), 提取 0.1R～0.9R 不同截面的二维流场, 由于当 R<0.5 时, 柔性变形并不明显, 不便于分析, 故选择 0.5R～0.9R 截面位置的二维涡量进行分析, 如图 14 所示。

三者对比来看, 刚度为 40 GPa 的柔性扑翼表面涡量最大, 刚度为 80 GPa 的柔性扑翼次之, 刚性扑翼表面涡量最小。由于在扑翼下拍过程中, 产生升力的主要来源是附着于翅膀前缘的前缘涡(leading edge vortex, LEV)结构。对于刚度为 40 GPa 的柔性扑翼, 在 0.5R～0.7R 处, LEV 结构保持的很好,直到 0.8R 处 LEV 开始破碎, 向后缘发展并逐渐与尾缘处的尾缘涡(trailing edge vortex, TEV)发生掺混。而对于刚性扑翼, LEV 结构无法在其表面得到很好的保持, 早在 0.6R 处就已经开始破碎直至与 TEV 发生掺混。刚度为 80 GPa 的柔性扑翼相对于刚性扑翼, 其表面的 LEV 结构也相对更不容易破碎。因此, 在 T_0=0.25 时刻呈现出柔性扑翼产生的升力

远大于刚性扑翼的现象。通过以上分析可知，在运动规律相同的情况下，柔性变形会影响扑翼表面涡量的展向分布而提高升力。

T_0=0.25时刻，3种扑翼模型在0.7R～0.9R的压力云图如图15所示，可以看到，由于LEV的存在，会在扑翼上表面产生低压区，该区域会随着LEV结构的演变而变化。刚度为40 GPa的柔性扑翼表面可以明显观察到由LEV产生的低压区（见图中红框），且其沿展向保持的相对稳定；而刚度为80 GPa柔性扑翼上表面低压区强度较低，且沿展向的稳定性较差；刚性扑翼上表面的低压区强度最低。这也解释了三者此时产生的升力差异。

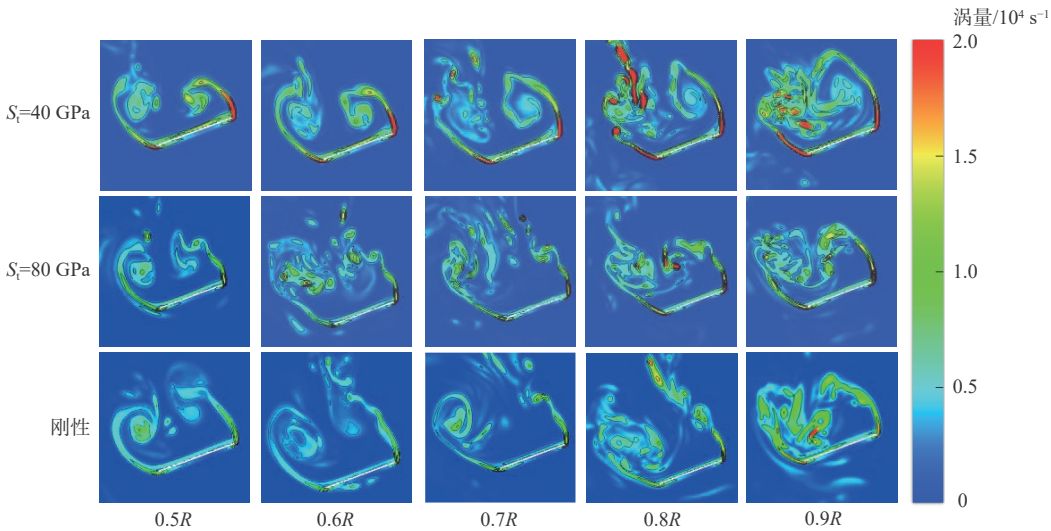


图14 T_0=0.25时刻3种扑翼模型在0.5R～0.9R的涡量对比

Fig. 14 Comparison of vorticity among three flapping wing models at 0.5R to 0.9R at T_0=0.25

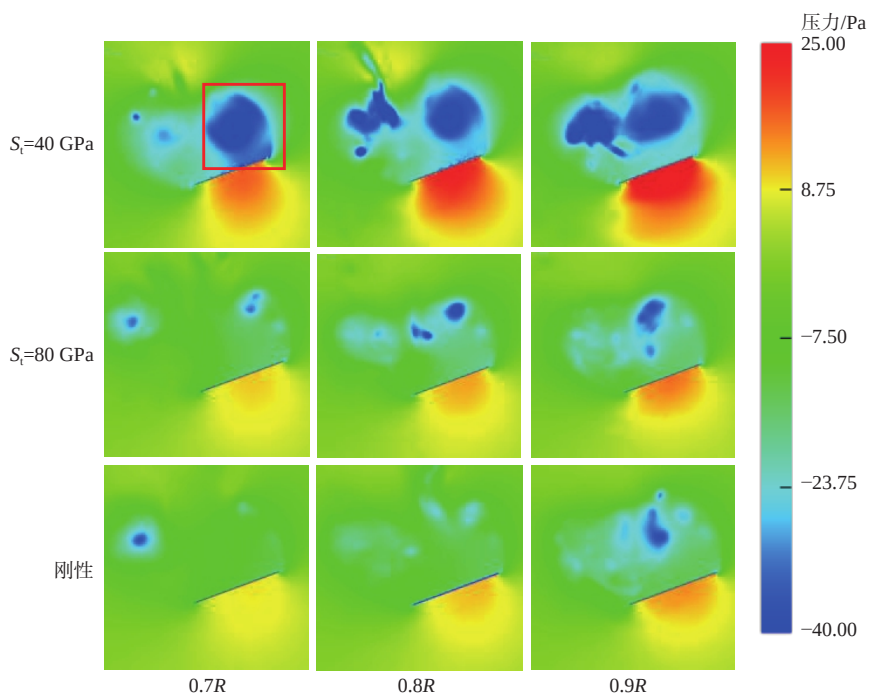


图15 T_0=0.25时刻3种扑翼模型在0.7R～0.9R的压力对比

Fig. 15 Comparison of pressure among three flapping wing models at 0.7R to 0.9R at T_0=0.25

3.2 时间维度

由图12可知，柔性变形对于气动力的影响主要发生于下拍阶段，因此，有必要对下拍阶段柔性扑翼的变形情况进行分析。柔性扑翼在下拍过程中会发生展向的弯曲变形，而其弦向的扭转变形很小可以忽略不计。引入一个新的参数弯曲角如图16所示，定义弯曲角为柔性扑翼、刚性扑翼翼根与翼尖连线的夹角。

在整个下拍阶段内取10个均匀时刻作为测点，测得弯曲角的变化情况，如图17所示，由图可

知, 对于刚度为 40 GPa 和 80 GPa 的柔性扑翼, 在下拍阶段前期, 弯曲角为正, 且先增大后减小; 下拍中期弯曲角基本为 0, 下拍阶段后期, 弯曲角为负, 且负向先增大后减小, 刚度为 40 GPa 的柔性扑翼展向弯曲变形更大。图 18 为 3 种扑翼模型下拍阶段拍动角对比, 由图可见, 柔性变形会对下拍阶段的拍动角产生影响, 但总体来说柔性扑翼的拍动角仍以正弦曲线变化。

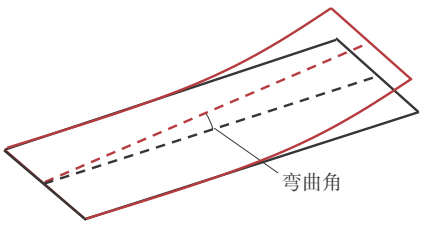


图 16　弯曲角示意图

Fig. 16　Diagram of bending angle

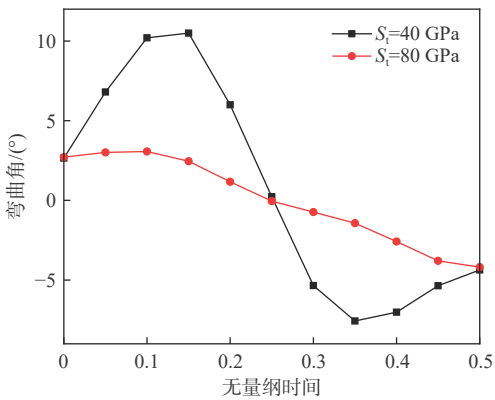


图 17　下拍阶段弯曲角随时间变化

Fig. 17　Change of bending angle during downstroke phase over time

由图 17 可知, T_0=0.25 时柔性扑翼的变形并不明显, 而图 14 显示, 在相同展向截面处柔性扑翼与刚性扑翼涡量强度存在明显差异。产生上述涡量

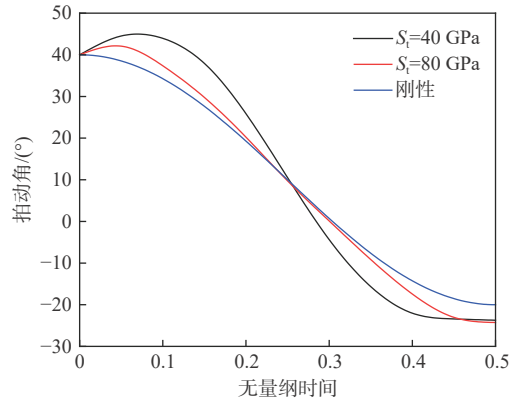


图 18　3 种扑翼模型下拍阶段拍动角对比

Fig. 18　Comparison of flapping phase angles among three flapping wing models

差异的原因必定发生在 T_0=0.25 之前, 因此, 针对 T_0=0.25 之前的时间段展开流场分析很有必要。又由于越靠近翼尖的位置, LEV 会与 TEV 掺混越严重, 无法清晰地观察出涡量的演变情况, 且 0.5R 内, 柔性扑翼和刚性扑翼流场差别不大。因此, 对 0.7R 截面流场进行分析, 以 T_0=0.05 为增量, 分析 T_0=0.05～0.25 时涡量随时间的演变情况, 如图 19 所示。由图可知, 当 T_0=0.05 时, 刚度为 40 GPa 的柔性扑翼表面涡量并不显著, 而刚度为 80 GPa 的柔性扑翼及刚性扑翼表面可以看到, 前 LEV 结构的形成。这是由于下拍阶段开始时, 柔性扑翼会产生柔性变形, 由图 17 可知, 此阶段内, 柔性扑翼会产生正向弯曲且刚度为 40 GPa 的柔性扑翼的弯曲角更大, 由于弯曲角的存在, 使得此阶段柔性扑翼产生下拍滞后而处于上拍行程(见图 18), 这不利于形成 LEV 结构。图 20 为 T_0=0.05 时 3 种扑翼表面速度云图, 可以观察到, 此时刚性扑翼表面速度最大, 刚度为 80 GPa 和 40 GPa 的柔性扑翼由于下拍滞后, 表面速度更小, 刚度为 40 GPa 的柔性扑翼表

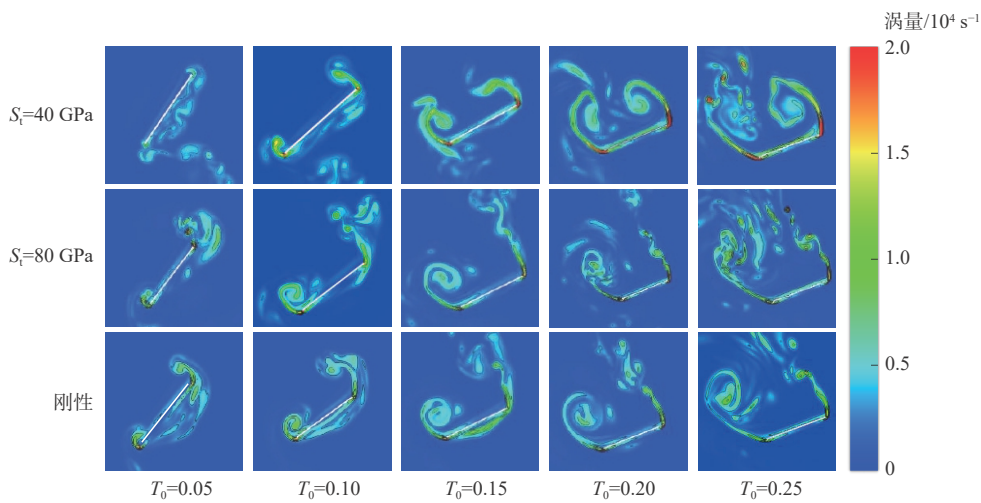


图 19　T_0=0.05～0.25 时 0.7R 处 3 种扑翼模型涡量对比

Fig. 19　Comparison of vorticity among three flapping wing models at 0.7R from T_0=0.05 to T_0=0.25

面速度最小。以上分析解释了图 12 中 T_0=0～0.05 时间内，刚度为 40 GPa 的柔性怕扑翼产生的升力最小，刚度为 80 GPa 的柔性扑翼产生的升力略小于刚性扑翼。

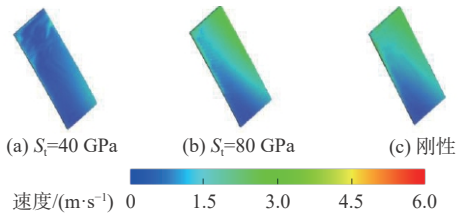


图 20 T_0=0.05 时 3 种扑翼模型表面速度对比

Fig. 20 Comparison of surface velocity among three flapping wing models at T_0=0.05

T_0=0.05～0.12 时间内，当 T_0=0.1 时，刚度为 40 GPa 的柔性扑翼表面开始产生 LEV 结构，而刚度为 80 GPa 的柔性扑翼及刚性扑翼表面的 LEV 结构开始逐渐被拉长并逐渐脱离扑翼表面。这是因为此阶段内柔性扑翼已进入下拍行程，LEV 结构随着下拍进行而逐渐形成。图 21 为 T_0=0.1 时 3 种扑翼表面速度云图，可以观察到，此时柔性扑翼与刚性扑翼表面速度已相差不大，因此，其与刚性扑翼在此阶段产生的升力差异越来越小。这很好地解释了图 12 扑翼模型升力系数对比中 T_0=0.05～0.12 时间内三者产生的升力差异。

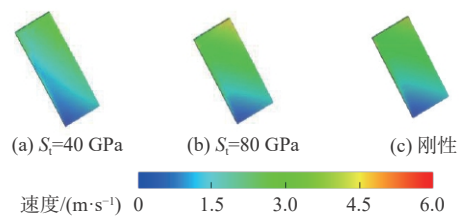


图 21 T_0=0.1 时 3 种扑翼模型表面速度对比

Fig. 21 Comparison of surface velocities among three flapping wing models at T_0=0.1

T_0=0.12～0.25 时间内，当 T_0=0.15 时，可以看到，刚度为 40 GPa 的柔性扑翼表面已经初步形成 LEV 结构，而刚度为 80 GPa 的柔性扑翼表面的 LEV 结构继续脱离，刚性扑翼表面的 LEV 结构脱落的程度更大。这是因为在此阶段内，由图 18 可知，柔性扑翼下拍拍动角变化曲线斜率的绝对值更大，下拍速度更大(刚度为 40 GPa 的柔性扑翼下拍速度最大，刚度为 80 GPa 的柔性扑翼次之，刚性扑翼下拍速度最小)。图 22 为 T_0=0.15 时三者表面速度云图，显示的结果也是如此，更大的下拍速度有利于 LEV 结构的形成，且此阶段柔性扑翼会产生正向弯曲，更容易使得 LEV 结构附着在其表面而不易脱落，因此，更有利于产生更高的升力。以上分析解释了图 12 扑翼模型升力系数对比 T_0=0.12～0.25 时间内三者产生的升力差异。

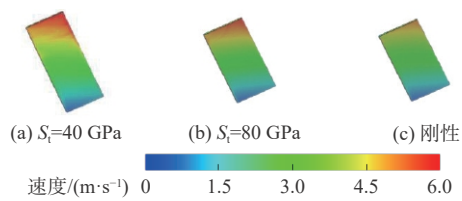


图 22 T_0=0.15 时 3 种扑翼模型表面速度对比

Fig. 22 Comparison of surface velocities among three flapping wing models at T_0=0.15

T_0=0.25～0.5 时，柔性扑翼和刚性扑翼表面的 LEV 与 TEV 在 0.5R 截面都发生了掺混，此时已不便于针对二者流场进行分析，图 23 为柔性扑翼与刚性扑翼在 T_0=0.3～0.5 表面涡量对比，可以观察到，此阶段三者后半段表面的 LEV 及 TEV 结构发生掺混并逐渐从翅膀表面脱落。由图 18 扑翼模型下拍阶段拍动角对比可知，此阶段柔性扑翼与刚性扑翼下拍拍动角变化曲线的斜率的绝对值逐渐减小，并最终趋于 0；其下拍速度的差异逐渐减小，如图 24 所示，T_0=0.4 时三者表面的速度分布差距不

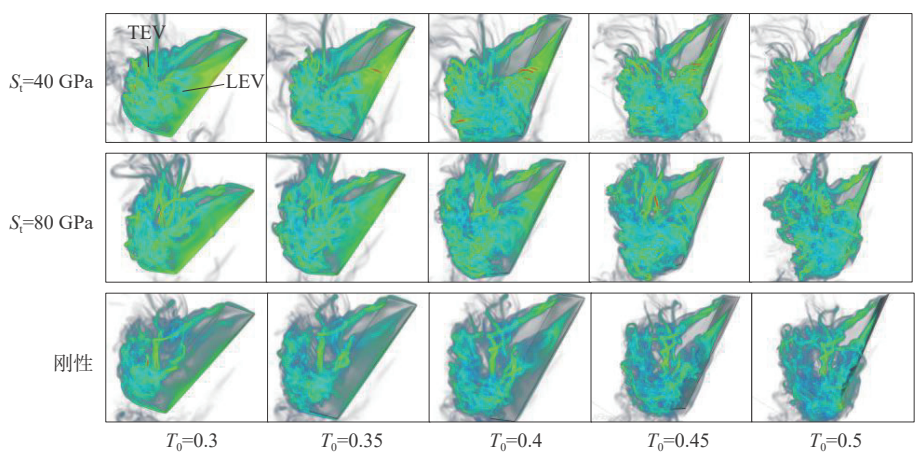


图 23 T_0=0.3～0.5 时 3 种扑翼模型三维涡量对比

Fig. 23 Comparison of three-dimensional vorticity among three flapping wing models from T_0=0.3 to T_0=0.5

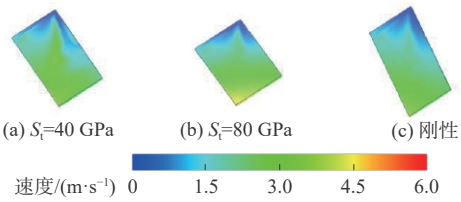


图 24　T_0=0.4 时 3 种扑翼表面速度云图

Fig. 24　Comparison of surface velocities among three flapping wing models at T_0=0.4

大, 且此阶段内柔性扑翼会发生负向弯曲, 这加剧了涡系结构从其表面的脱落。以上分析解释了图 12 扑翼模型升力系数对比中 T_0=0.25～0.5 时三者产生的升力差异。

T_0=0.5～1.0 时为上拍阶段, 图 25 为 T_0=0.6～1.0 时三者表面涡量对比, 可以看到, 此阶段无论是柔性扑翼还是刚性扑翼, 受运动规律的影响, 其表面的复杂涡系随着时间的推进而直至几乎全部脱落, 此时已几乎无法产生升力。这对应了图 12 中当 T_0=0.5～1.0 时, 柔性扑翼和刚性扑翼产生的升力几乎为 0。

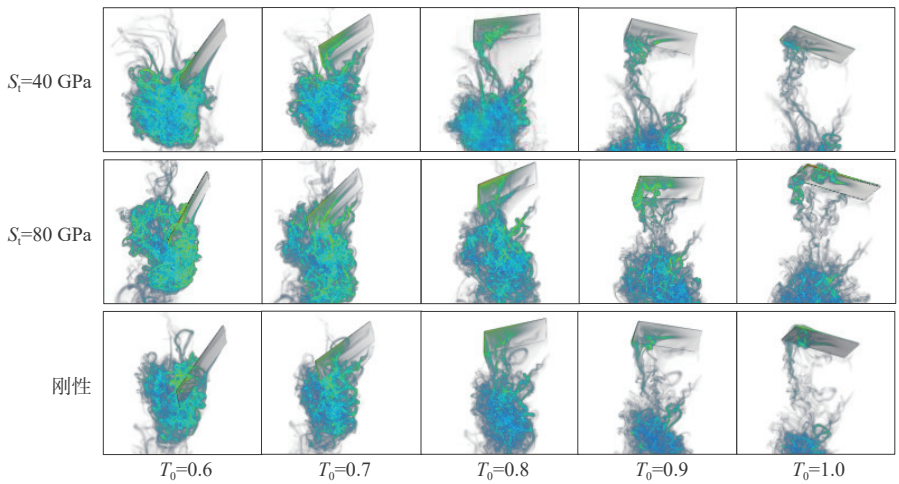


图 25　T_0=0.6～1.0 时 3 种扑翼三维涡量对比

Fig. 25　Comparison of three-dimensional vorticity among three flapping wing models from T_0=0.6 to T_0=1.0

4　结　论

本文通过流固耦合数值模拟方法开展了低雷诺数柔性平板气动特性的研究。针对对比不同柔性扑翼与刚性扑翼产生气动力的差别, 对扑翼附近流场进行分析, 从而给出柔性变形对非定常气动力的影响机制。

1) 对于升力而言, 柔性扑翼在整个拍动周期的平均升力相对于刚性扑翼升力提高了 60.5%; 而由于选择悬停飞行状态, 在整个拍动周期中, 柔性扑翼与刚性扑翼的平均推力都几乎为 0。

2) 柔性变形对于升力的影响主要发生在下拍阶段, 初始阶段(T_0=0～0.12), 柔性扑翼相对于刚性扑翼升力下降 31.4%; 此后(T_0=0.12～0.5), 柔性扑翼相对于刚性扑翼升力提高 76.7%; 上拍阶段几乎不产生升力。

3) 进一步分析流场, 揭示了柔性变形对于气动力的时空影响机制, 即在空间层面, 柔性变形能够影响 LEV 的展向分布而提高升力, 柔性变形可以使扑翼表面的 LEV 沿展向可以保持的很好而不破碎; 而在时间层面, 柔性变形产生的展向弯曲在拍动初期使得下拍行程滞后而不利于 LEV 的形成, 展向弯曲能够影响下拍速度进而影响表面的 LEV 结构而提高升力。

参考文献（References）

[1] WEIS-FOGH T. Quick estimates of flight fitness in hovering animals, including novel mechanisms for lift production[J]. Journal of Experimental Biology, 1973, 59(1): 169-230.

[2] ELLINGTON C P, VAN DEN BERG C, WILLMOTT A P, et al. Leading-edge vortices in insect flight[J]. Nature, 1996, 384(6610): 626-630.

[3] DICKINSON M H, LEHMANN F O, SANE S P. Wing rotation and the aerodynamic basis of insect flight[J]. Science, 1999, 284(5422): 1954-1960.

[4] ELLINGTON C P. The aerodynamics of hovering insect flight. IV. aeorodynamic mechanisms[J]. Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 1984, 305(1122): 79-113.

[5] NORBERG R Å. Flight characteristics of two plume moths, alucita pentadacty/a L. and orneodes hexadactyla L. (microlepidoptera)[J]. Zoologica Scripta, 1972, 1(5): 241-246.

[6] VOGEL S. Flight in Drosophila. II. Variations in stroke parameters and wing contour[J]. Journal of Experimental Biology, 1967, 46(2): 383.

[7] NGUYEN A T, HAN J H. Wing flexibility effects on the flight performance of an insect-like flapping-wing micro-air vehicle[J]. Aerospace Science and Technology, 2018, 79: 468-481.

[8] NAKATA T, LIU H. Aerodynamic performance of a hovering hawkmoth with flexible wings: a computational approach[J]. Pro-

ceedings Biological Sciences, 2012, 279(1729): 722-731.

[9] 周超英, 朱建阳, 汪超, 等. 柔性扑翼气动性能的数值研究[J]. 工程力学, 2013, 30(5): 13-18.
ZHOU C Y, ZHU J Y, WANG C, et al. Numerical study on the effect of flexiblity of a flapping wing on its aerodynamic performance[J]. Engineering Mechanics, 2013, 30(5): 13-18(in Chinese).

[10] 孟令兵, 昂海松, 肖天航. 基于 CFD/CSD 方法的蜻蜓柔性翼气动特性分析[J]. 航空动力学报, 2014, 29(9): 2063-2069.
MENG L B, ANG H S, XIAO T H. Analysis of aerodynamic characteristics of flexible wing of dragonfly based on CFD/CSD method[J]. Journal of Aerospace Power, 2014, 29(9): 2063-2069(in Chinese).

[11] 陈利丽, 宋笔锋, 宋文萍, 等. 柔性扑翼气动结构耦合特性数值研究[J]. 空气动力学学报, 2015, 33(1): 125-133.
CHEN L L, SONG B F, SONG W P, et al. Numerical aerodynamic-structural coupling research for flexible flapping wing[J]. Acta Aerodynamica Sinica, 2015, 33(1): 125-133(in Chinese).

[12] 王奇, 朱寅鑫, 牛培行, 等. 柔性扑翼翼型的气动性能仿真分析[J]. 应用数学和力学, 2022, 43(5): 586-596.
WANG Q, ZHU Y X, NIU P X, et al. Simulation of aerodynamic performances of flexible flapping wing airfoils[J]. Applied Mathematics and Mechanics, 2022, 43(5): 586-596(in Chinese).

[13] 彭连松, 郑孟宗, 潘天宇, 等. 翼间干涉效应对蜻蜓悬停气动性能的影响[J]. 航空学报, 2021, 42(7): 264-273.
PENG L S, ZHENG M Z, PAN T Y, et al. Effects of tandem-wing interactions on aerodynamics of hovering dragonflies[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(7): 264-273(in Chinese).

[14] RÜPPELL G. Kinematic analysis of symmetrical flight manoeuvres of Odonata[J]. Journal of Experimental Biology, 1989, 144(1): 13-42.

[15] NORBERG R Å. Hovering flight of the dragonfly aeschna juncea L. , kinematics and aerodynamics[C]//Proceedings of the Swimming and Flying in Nature. Berlin: Springer, 1975(2): 763-781.

[16] ÅKE NORBERG R. The pterostigma of insect wings an inertial regulator of wing pitch[J]. Journal of Comparative Physiology, 1972, 81(1): 9-22.

[17] AZUMA A, AZUMA S, WATANABE I, et al. Flight mechanics of a dragonfly[J]. Journal of Experimental Biology, 1985, 116(1): 79-107.

[18] WANG H, ZENG L J, LIU H, et al. Measuring wing kinematics, flight trajectory and body attitude during forward flight and turning maneuvers in dragonflies[J]. The Journal of Experimental Biology, 2003, 206(Pt 4): 745-757.

[19] WAKELING J M, ELLINGTON C P. Dragonfly flight: I. gliding flight and steady-state aerodynamic forces[J]. Journal of Experimental Biology, 1997, 200(3): 543-556.

[20] BHARDWAJ S, DALAL A, BISWAS G, et al. Analysis of droplet dynamics in a partially obstructed confinement in a three-dimensional channel[J]. Physics of Fluids, 2018, 30(10): 102102.

[21] RAJABI H, GORB S N. How do dragonfly wings work? A brief guide to functional roles of wing structural components[J]. International Journal of Odonatology, 2020, 23(1): 23-30.

[22] AONO H, CHIMAKURTHI S K, WU P, et al. A computational and experimental studies of flexible wing aerodynamics [C]//Proceedings of the 48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition. Reston: AIAA, 2010.

[23] SHAHZAD A, TIAN F B, YOUNG J, et al. Effects of flexibility on the hovering performance of flapping wings with different shapes and aspect ratios[J]. Journal of Fluids and Structures, 2018, 81: 69-96.

[24] WANG X S, LI Y, SHI Y F. Effects of sandwich microstructures on mechanical behaviors of dragonfly wing vein[J]. Composites Science and Technology, 2008, 68(1): 186-192.

[25] WAINWRIGHT S A. Mechanical design in organisms[M]. Princeton: Princeton University Press, 2020.

[26] VINCENT J F V, WEGST U G K. Design and mechanical properties of insect cuticle[J]. Arthropod Structure & Development, 2004, 33(3): 187-199.

Unsteady aerodynamic characteristics of flexible flapping plate at low Reynolds numbers

CAO Mengda[1], ZHENG Mengzong[1], SU Guanting[2,*], PAN Tianyu[1], LI Zhiping[1], LI Qiushi[1,2,3]

(1. Research institute of Aero-Engine, Beihang University, Beijing 100191, China;
2. School of Energy and Power Engineering, Beihang University, Beijing 100191, China;
3. School of Aeronautics and Astronautics, Xihua University, Chengdu 610039, China)

Abstract: Flexible deformation impacts the aerodynamics of flapping wings and is accompanied by severe fluid-structure interaction effects in the case of flexible flapping wings. This study examines the fluid-structure interaction aerodynamic properties of a flexible flapping plate at low Reynolds numbers in order to clarify the process by which flexible deformation influences unsteady aerodynamic forces. By applying the motion patterns of dragonflies to both flexible and rigid flapping wings and conducting fluid-structure interaction numerical simulations, it was found that, under the same motion patterns, the average lift of flexible flapping wings over an entire cycle increased by 60.5% compared to rigid flapping wings. During the initial downstroke, the lift decreased by 31.4%, but increased by 76.7% during the mid to late downstroke. The thrust generated by both flexible and rigid flapping wings throughout the cycle was nearly zero. The study revealed the spatiotemporal influence mechanism of flexible deformation on aerodynamic forces. Lift is increased by the flexible flapping wings' spanwise bending deformation during the downstroke, which keeps the leading-edge vortex's spanwise distribution on the wing surface. At the beginning of the downstroke, flexible flapping wings experience downstroke lag, which is detrimental to the formation of the leading-edge vortex, thus reducing lift. As the downstroke progresses, spanwise bending deformation increases the downstroke speed, affecting the structure of the leading-edge vortex and subsequently increasing lift.

Keywords: flexible flapping wing; dragonfly; fluid-structure interaction; lift coefficient; leading edge vortex

Received: 2024-04-19; **Accepted**: 2024-07-29; **Published Online**: 2024-08-06 09:40

URL: link.cnki.net/urlid/11.2625.V.20240805.1816.004

Foundation items: National Natural Science Foundation of China (52206040,52006002,51976005)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0230

基于牛顿-欧拉递推算法的高置信度斜铰旋翼动力学模型

杨毅松，李建波*，段登燕
(南京航空航天大学 航空学院 直升机动力学全国重点实验室，南京 210016)

摘 要：斜铰旋翼无需自动倾斜器，仅通过电机加减速变化就能实现周期变距操控，结构复杂性低，但引入斜铰后，旋翼摆振、变距相耦合，使斜铰旋翼动力学建模更加复杂。针对现有建模方法中正、反桨叶间挥舞差异揭示不充分、模型预测精度不高的问题，基于改进的Denavia-Hartenberg方法建立桨叶局部连杆坐标系，采用牛顿-欧拉递推动力学算法，在局部连杆坐标系下，对斜铰旋翼连杆系统向外递推计算各连杆的速度和加速度，向内递推计算各连杆间的相互作用力、力矩，构建斜铰旋翼动力学模型。在此基础上，通过仿真计算，进一步揭示了斜铰旋翼周期变距机理。计算结果表明：所提模型可以更加准确地预测正、反桨叶间挥舞差异；对于桨叶摆振运动幅值的预测精度提高了9.05%。

关 键 词：斜铰旋翼；无自动倾斜器；摆振-变距耦合；动力学模型；周期变距控制
中图分类号：V275+.1
文献标志码：A **文章编号**：1001-5965(2026)06-2024-10

直升机可以垂直起降、悬停作业，具有优秀的点对点运输能力，且低空机动性能优异，在军事、抢险救护、民用运输等领域取得了广泛应用[1]。近年来，随着微电子系统的蓬勃发展[2]及微小型飞行器应用场景的不断拓宽，在航拍摄影、物流运输等领域的应用取得了长足进步，使微小型飞行器吸引了许多研究人员的注意，已经成为热门的研究方向[3]。得益于直升机优异的悬停能力，将其微小型化后，应用在军事侦察、民用航拍摄影等领域的潜力巨大。这也对直升机的微小型化发展提出了新的需求。

目前，直升机微小型化面临的技术难点为微型无人直升机的结构和制造工艺[4]。传统大型直升机主要通过自动倾斜器与舵机的配合来实现桨叶的周期变距效果，虽然直接将自动倾斜器小型化后应用于小型直升机上已被证实是可行的[5-6]，但简单地将操纵机构进行等比例缩小再应用于微小型直升机上的效果并不理想。这是因为自动倾斜器具有结构复杂、质量较大、加工成本高的缺点，将其等比例缩小应用于微小型直升机上时，受限于小尺寸零件的加工水平和加工成本，这些缺点会被进一步放大。同时，自动倾斜器还需要依靠多套舵机配合来实现桨叶的周期变距，进一步制约了其在微小型直升机上的发展。

本文研究的斜铰旋翼是一种新型的周期变距装置，由于其独特的机械结构，使旋翼的摆振运动与变距运动相耦合，在无需自动倾斜器等机构的情况下，仅仅依托电机加减速的变化，就能实现旋翼的周期变距操控。斜铰旋翼可以大大简化旋翼系统的复杂性，降低成本，减轻质量，为直升机周期变距机构的微小型化提供一种简单可行的方案。

斜铰旋翼方案最早于2013年由宾夕法尼亚大学的Paulos和Yim提出，并成功运用于共轴双旋翼直升机上，从工程实现的角度对该方案进行了阐

收稿日期：2024-04-18；录用日期：2024-05-24；网络出版时间：2024-06-20 11:29
网络出版地址：link.cnki.net/urlid/11.2625.V.20240619.1210.001
*通信作者. E-mail：ljb101@nuaa.edu.cn

引用格式：杨毅松，李建波，段登燕. 基于牛顿-欧拉递推算法的高置信度斜铰旋翼动力学模型[J]. 北京航空航天大学学报，2026，52(6)：2024-2033. YANG Y S，LI J B，DUAN D Y. Dynamic model of high confidence tilt-hinge rotor based on Newton-Euler recursion algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：2024-2033 (in Chinese).

述[7], 证明了斜铰旋翼方案的有效性。在后续的飞行实验中, 该团队对搭载斜铰旋翼方案的飞行器进行了飞行性能测试, 实验表明, 由于斜铰旋翼方案相较于自动倾斜器方案的结构质量更低, 因此, 可以搭载更大的机载电池, 获得更长的续航时间[8]。得益于斜铰旋翼简单的机械构造, 其在微小型直升机领域颇具优势。在国内, 国防科技大学的尹欣繁[9]、北京邮电大学的黄灿铭[10] 开展了斜铰旋翼方案的相关跟踪研究, 并开展了工程复现实验, 实现了斜铰旋翼方案在单旋翼带尾桨、纵列式双旋翼直升机上的应用。南京航空航天大学的王彪教授团队[11-14] 开展了相关的驱动单元转速控制、有效操纵力矩生成规律、斜铰旋翼在无人直升机上的应用等方面的研究。斜铰旋翼方案已经在多种构型的直升机上实现了应用, 证明了该方案的可行性。

当前对于斜铰旋翼方案的研究中, 大部分还停留在工程实现的应用阶段, 有关斜铰旋翼的变距机理阐述及动力学模型建立方面的工作内容还相对较少, 且已经建立的动力学模型中, 对于斜铰旋翼的正、反桨挥舞差异性的揭示并不充分。文献 [15] 假设正、反桨叶近似对称, 以一侧旋翼为研究对象建立了旋翼动力学模型, 并与实验结果进行了比对, 实验结果中正、反桨叶的挥舞差异性并未在动力学模型中体现出来。南京航空航天大学的 Yu 等[16] 采用正、反桨叶对称的近似假设, 以一侧旋翼为研究对象, 建立了斜铰旋翼的非线性动力学模型, 相较于文献 [15] 中的动力学模型, 显著提升了桨叶挥舞相位的预测精度, 但该动力学模型对正、反桨叶间挥舞运动差异性的揭示并不充分, 对反桨挥舞运动的预测并不准确。

开展动力学研究工作是进行旋翼结构参数优化设计及桨叶运动仿真的基础。针对现有模型对正、反桨叶间挥舞差异预测不准确的问题, 本文通过在斜铰旋翼上建立多个局部坐标系, 借由局部坐标系可以更准确地描述由于斜铰引入而导致的复杂桨叶姿态变化, 进而通过牛顿-欧拉递推动力学算法完成旋翼的动力学建模; 设计斜铰旋翼周期变距操纵仿真实验, 以验证动力学模型的正确性; 将动力学模型仿真计算结果与文献 [15] 中的实验数据进行比对, 对动力学模型的准确性进行检验。

1　桨叶运动学

本节给出基于改进的 Denavia-Hartenberg 方法(简称 D-H 法)建立的连杆坐标系及由连杆运动学方程得到的桨叶姿态随关节角度的变化关系式。

1.1　斜铰旋翼结构及周期变距原理

如图 1 所示, 斜铰旋翼的构造较为简单, 仅由桨毂、对称布置的挥舞铰、反对称布置的倾斜铰、桨夹、桨叶组成。将倾斜轴线与桨毂转轴上侧轴线相交的一侧桨叶定义为正桨, 反之, 定义为反桨。

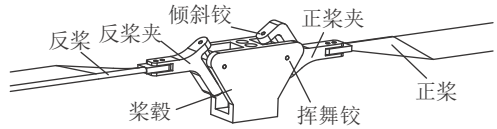


图 1　斜铰旋翼结构示意

Fig. 1　Tilt-hinge rotor structure diagram

斜铰旋翼在周期变距的实现上与自动倾斜器不同。斜铰旋翼由于斜置放置的倾斜铰使桨叶的摆振运动与变距运动相耦合。当电机做加速运动时, 由于惯性, 正、反桨叶均滞后于桨毂, 正桨桨距增加, 反桨桨距减小; 而当电机做减速运动时, 桨叶的桨距变化情况与之相反。通过调制电机转速, 使电机在特定位置进行加、减速运动, 使桨叶桨距呈周期性变化, 即可实现与自动倾斜器类似的周期变距效果。

1.2　连杆坐标系

为能以关节转角解算桨叶运动姿态, 以及后续的动力学分析, 将桨毂、桨叶简化为连杆机构。对斜铰旋翼系统进行如下简化: ①忽略桨叶弹性变形, 连杆为均匀刚体, 质量均匀分布; ②连杆结构严格对称; ③连杆重心与中心重合。

对连杆、关节进行如下命名: 将桨毂旋转轴定义为关节 1, 将正桨的挥舞铰、倾斜铰分别定义为关节 2、3, 将反桨的挥舞铰、倾斜铰分别定义为关节 2N、3N。同时, 将关节 2N 至关节 2 的桨毂部分定义为连杆 1, 将关节 2 至关节 3 部分定义为连杆 2, 将关节 2N 至关节 3N 部分定义为连杆 2N, 将正、反桨桨叶分别定义为连杆 3、3N。定义后的关节命名如图 2 所示, 其中, θ_1 为关节 1 的转动角度, θ_2、θ_3、θ_{2N}、θ_{3N} 分别为关节 2、关节 3、关节 2N、关节 3N 的转动角度。

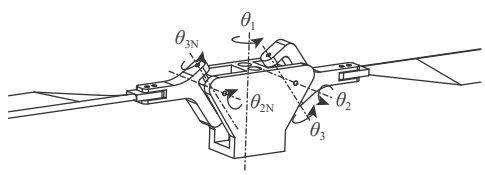


图 2　斜铰旋翼关节命名

Fig. 2　Naming of tilt-hinge rotor joints

机器人运动学中, 普遍采用 D-H 法进行坐标系的建立, 其是由 Denavia 和 Hartenberg 在 1955 年提出的用于描述低副机构的方法, 该方法在约束运动副的杆件上建立局部坐标系, 采用 4 个参数来描

述附加坐标系之间的变换矩阵。通过四阶齐次变换矩阵来描述连杆的空间几何关系。本节使用改进的D-H法进行坐标系的建立。坐标系如图3所示，以下角标N对正、反桨叶坐标系进行区分，对于坐标系中重合的轴，以逗号对其进行分隔，如$x_{2,3}$表示坐标系2、3的x轴，$x_{2N,3N}$表示坐标系2N、3N的x轴。

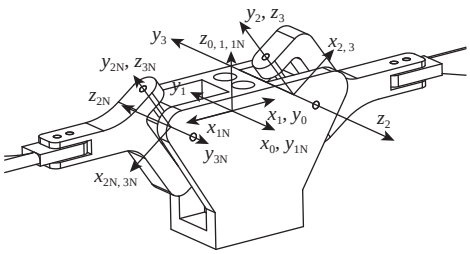


图3 斜铰旋翼连杆坐标系

Fig. 3 Tilt-hinge rotor linkage coordinate system

由图3建立的坐标系得到的连杆D-H参数如表1所示。其中，连杆长度L_{i-1}为沿x_i轴从z_i移动到z_{i+1}的距离，关节扭角α_{i-1}为绕x_i轴从z_i旋转到z_{i+1}的角度，连杆之间的距离d_i为沿z_i轴从x_{i-1}移动到x_i的距离，关节变量θ_i为绕z_i轴从x_{i-1}旋转到x_i的角度，L_1为连杆1长度，即正桨挥舞铰到反桨挥舞铰的距离，δ为倾斜铰的倾斜角度。

表1 连杆坐标系的D-H参数

Table 1 D-H parameters of linkage coordinate system

坐标系	L_{i-1}	α_{i-1}/(°)	d_i	θ_i/(°)
$0\to1$	0	0	0	$90+\theta_1$
$1\to2$	$L_1/2$	90	0	$\delta+\theta_2$
$2\to3$	0	−90	0	θ_3
$0\to1N$	0	0	0	θ_1-90
$1N\to2N$	$L_1/2$	90	0	$-\delta+\theta_{2N}$
$2N\to3N$	0	−90	0	θ_{3N}
$1N\to1$	0	0	0	180

位置转换矩阵[17]为

$$\boldsymbol{T}_i^{i-1}=\begin{bmatrix}\cos\theta_i & -\sin\theta_i & 0 & L_{i-1}\\ \sin\theta_i\cos\alpha_{i-1} & \cos\theta_i\cos\alpha_{i-1} & -\sin\alpha_{i-1} & -d_i\sin\alpha_{i-1}\\ \sin\theta_i\sin\alpha_{i-1} & \cos\theta_i\sin\alpha_{i-1} & \cos\alpha_{i-1} & d_i\cos\alpha_{i-1}\\ 0 & 0 & 0 & 1\end{bmatrix} \tag{1}$$

式中：$\boldsymbol{T}_i^{i-1}$为位置变化矩阵，描述坐标系i相对于坐标系$i-1$的位置变换关系，如$\boldsymbol{T}_2^1$为坐标系1到坐标系2的位置变换矩阵，$\boldsymbol{T}_1^2$为坐标系2到坐标系1的位置变换矩阵。

坐标系0至坐标系3下的位置变换矩阵可表示为

$$\boldsymbol{T}_3^0=\boldsymbol{T}_1^0\cdot\boldsymbol{T}_2^1\cdot\boldsymbol{T}_3^2 \tag{2}$$

定义旋转变换矩阵为位置变换矩阵的前3行前3列，因此，旋转变换矩阵表示为

$$\boldsymbol{R}_i^{i-1}=\begin{bmatrix}\cos\theta_i & -\sin\theta_i & 0\\ \sin\theta_i\cos\alpha_{i-1} & \cos\theta_i\cos\alpha_{i-1} & -\sin\alpha_{i-1}\\ \sin\theta_i\sin\alpha_{i-1} & \cos\theta_i\sin\alpha_{i-1} & \cos\alpha_{i-1}\end{bmatrix} \tag{3}$$

式中：$\boldsymbol{R}_i^{i-1}$为旋转变换矩阵，描述坐标系i相对于坐标系$i-1$的旋转变换关系，如$\boldsymbol{R}_2^1$为坐标系1到坐标系2的旋转变换矩阵，$\boldsymbol{R}_1^2$为坐标系2到坐标系1的旋转变换矩阵。

由旋转变换矩阵的性质可得

$$\boldsymbol{R}_i^{i-1}=\left(\boldsymbol{R}_{i-1}^{i}\right)^{-1}=\left(\boldsymbol{R}_{i-1}^{i}\right)^{\mathrm{T}} \tag{4}$$

如此，便建立了各级连杆间的位置变换矩阵、旋转变换矩阵，以使用关节间的转动解算桨叶的挥舞、摆振、变距运动。

1.3 桨叶运动姿态分析

依据图4将挥舞角β_p、摆振角ζ_p、变距角θ_p分别定义为桨叶绕y_1轴负方向、z_1轴正方向、x_1轴正方向的转动角度。因此，正桨挥舞角可以表示为正桨桨叶方向向量在X_1Z_1平面的投影与x_1轴的夹角。正桨摆振角可以表示为正桨桨叶方向向量在X_1Y_1平面的投影与x_1轴的夹角。而正桨变距角可以表示为y_3轴在Y_1Z_1平面的投影与y_1轴的夹角。按照上述描述，桨叶方向向量在坐标系1中的表示如图5所示。其中，红色带箭头实线为桨叶方向向量，橙色、绿色带箭头虚线分别为其在X_1Z_1、X_1Y_1平

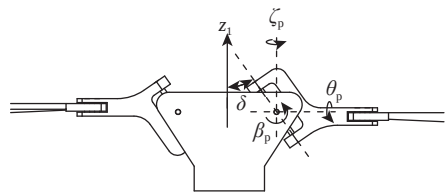


图4 桨叶挥舞角、摆振角、变距角定义

Fig. 4 Blade flapping angle, lag angle, pitch angle definition

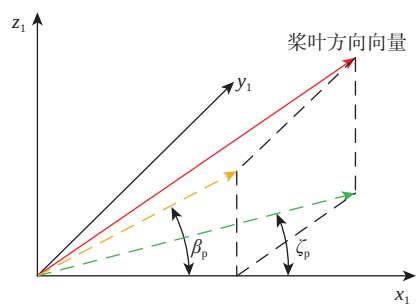


图5 坐标系1中的正桨桨叶方向向量

Fig. 5 Forward blade vector in coordinate system 1

面的投影。桨叶方向向量在X_1Z_1、X_1Y_1平面上的投影与x_1轴的夹角即为对应的挥舞角、摆振角。因此, 将坐标系 3 下的对应矢量转换到坐标系 1 下计算相应夹角, 即可得到对应的挥舞角、摆振角、变距角。

依据旋转变换矩阵得到的坐标系 1 下正桨桨叶方向向量、Y_3轴向量如下:

$$\boldsymbol{P}_{\mathrm{pj}}^{1}=\begin{bmatrix} L_3(\sin(\delta+\theta_2)\sin\delta+\cos(\delta+\theta_2)\cos\delta\cos\theta_3) \\ L_3\cos\delta\sin\theta_3 \\ L_3(\sin(\delta+\theta_2)\cos\delta\cos\theta_3-\cos(\delta+\theta_2)\sin\delta) \end{bmatrix} \tag{5}$$

$$\boldsymbol{P}_{\mathrm{py}}^{1}=\begin{bmatrix} -\cos(\delta+\theta_2)\sin\theta_3 \\ \cos\theta_3 \\ -\sin(\delta+\theta_2)\sin\theta_3 \end{bmatrix}^{\mathrm{T}} \tag{6}$$

式中: $\boldsymbol{P}_{\mathrm{pj}}^{1}$为正桨桨叶方向向量在坐标系 1 下的向量表示, 在本文中, 左上角标默认作为向量所在坐标系的标志, 后续向量不再进行赘述; $\boldsymbol{P}_{\mathrm{py}}$为$y_3$轴向量; L_3为连杆 3 长度, 即桨叶长度。

最终得到的正桨挥舞角、摆振角、变距角分别为

$$\beta_{\mathrm{p}}=\arctan\left(\frac{\tan(\delta+\theta_2)\cos\theta_3-\tan\delta}{\tan(\delta+\theta_2)\tan\delta+\cos\theta_3}\right) \tag{7}$$

$$\zeta_{\mathrm{p}}=\arctan\left(\frac{\sin\theta_3}{\sin(\delta+\theta_2)\tan\delta+\cos(\delta+\theta_2)\cos\theta_3}\right) \tag{8}$$

$$\theta_{\mathrm{p}}=\arctan\left(\frac{-\sin(\delta+\theta_2)\sin\theta_3}{\cos\theta_3}\right) \tag{9}$$

使用同样的方法可获得反桨的挥舞角β_{n}、摆振角ζ_{n}、变距角θ_{n}表达式。至此, 通过建立连杆坐标系完成了对桨叶挥舞、摆振、变距运动的描述。

2　动力学模型

牛顿-欧拉递推动力学算法是建立在以牛顿方程和欧拉方程为基础的经典刚体动力学上, 以牛顿-欧拉方法分别建立单个刚体的动力学方程, 应用递推算法来建立多刚体系统的动力学方程。递推动力学算法可分为 2 部分: 首先, 向外递推计算各连杆的速度和加速度, 由牛顿-欧拉公式计算各连杆的惯性力和力矩; 然后, 向内递推计算各连杆相互作用的力和力矩及施加在关节上的气动力矩、摩擦力矩、电机转矩。由各关节上的力矩平衡方程, 最终可得到封闭形式下的动力学方程。

2.1　外推计算

为计算作用在连杆上的惯性力, 需要计算操作臂每个连杆的角速度、线加速度和角加速度。可应用递推方法完成计算。按照$0\rightarrow1\rightarrow2\rightarrow3$、$0\rightarrow1\mathrm{N}\rightarrow2\mathrm{N}\rightarrow3\mathrm{N}$的递推方向由坐标系 0 开始向坐标系 3、3N 方向进行递推计算, 对各个连杆上的角速度、角加速度、惯性力、惯性力矩进行计算。递推计算公式为

$$\boldsymbol{\omega}_{i+1}^{i+1}=\boldsymbol{R}_{i}^{i+1}\cdot\boldsymbol{\omega}_{i}^{i}+\dot{\theta}_{i+1}\cdot\hat{\boldsymbol{Z}}_{i+1}^{i+1} \tag{10}$$

$$\dot{\boldsymbol{\omega}}_{i+1}^{i+1}=\boldsymbol{R}_{i}^{i+1}\cdot\dot{\boldsymbol{\omega}}_{i}^{i}+\boldsymbol{R}_{i}^{i+1}\cdot\boldsymbol{\omega}_{i}^{i}\times\dot{\theta}_{i+1}\cdot\hat{\boldsymbol{Z}}_{i+1}^{i+1}+\ddot{\theta}_{i+1}\cdot\hat{\boldsymbol{Z}}_{i+1}^{i+1} \tag{11}$$

$$\dot{\boldsymbol{v}}_{i+1}^{i+1}=\boldsymbol{R}_{i}^{i+1}(\dot{\boldsymbol{\omega}}_{i}^{i}\times\boldsymbol{P}_{i+1}^{i}+\boldsymbol{\omega}_{i}^{i}\times(\boldsymbol{\omega}_{i}^{i}\times\boldsymbol{P}_{i+1}^{i})+\dot{\boldsymbol{v}}_{i}^{i}) \tag{12}$$

$$\begin{aligned}\dot{\boldsymbol{v}}_{i+1\mathrm{c}}^{i+1}=&\dot{\boldsymbol{\omega}}_{i+1}^{i+1}\times\boldsymbol{P}_{i+1\mathrm{c}}^{i+1}+\\&\boldsymbol{\omega}_{i+1}^{i+1}\times(\boldsymbol{\omega}_{i+1}^{i+1}\times\boldsymbol{P}_{i+1\mathrm{c}}^{i+1})+\dot{\boldsymbol{v}}_{i+1}^{i+1}\end{aligned} \tag{13}$$

$$\boldsymbol{F}_{i+1}^{i+1}=m_{i+1}\cdot\dot{\boldsymbol{v}}_{i+1\mathrm{c}}^{i+1} \tag{14}$$

$$\boldsymbol{N}_{i+1}^{i+1}=\boldsymbol{I}_{i+1}^{\mathrm{c}i+1}\cdot\dot{\boldsymbol{\omega}}_{i+1}^{i+1}+\boldsymbol{\omega}_{i+1}^{i+1}\times\boldsymbol{I}_{i+1}^{\mathrm{c}i+1}\cdot\boldsymbol{\omega}_{i+1}^{i+1} \tag{15}$$

式中: $\boldsymbol{\omega}_{i+1}$、$\dot{\boldsymbol{\omega}}_{i+1}$分别为连杆$i+1$的角速度向量、角加速度向量; $\boldsymbol{\omega}_i$、$\dot{\boldsymbol{\omega}}_i$分别为上一连杆$i$的角速度向量、角加速度向量; $\dot{\theta}_{i+1}$、$\ddot{\theta}_{i+1}$分别为关节$i+1$的角速度、角加速度; $\hat{\boldsymbol{Z}}_{i+1}$为坐标系$i+1$的$z$轴单位向量; $\dot{\boldsymbol{v}}_{i+1}$、$\dot{\boldsymbol{v}}_i$分别为坐标系$i+1$、$i$原点处的线加速度向量; $\boldsymbol{P}_{i+1}$为由坐标系i的原点到坐标系$i+1$原点的向量; $\dot{\boldsymbol{v}}_{i+1\mathrm{c}}$为连杆$i+1$重心处的线加速度向量; $\boldsymbol{P}_{i+1\mathrm{c}}$为由坐标系$i+1$的原点到连杆$i+1$重心位置的向量; m_{i+1}为连杆$i+1$的质量; $\boldsymbol{I}_{i+1}^{\mathrm{c}i+1}$为连杆$i+1$在连杆重心坐标系下的惯性张量, 其中, 连杆重心坐标系的原点位于连杆重心处, 各坐标轴方向与原连杆坐标系$i+1$方向相同; $\boldsymbol{F}_{i+1}$、$\boldsymbol{N}_{i+1}$分别为连杆$i+1$受到的惯性力向量、惯性力矩向量。

由于基座处于固定状态, 固定坐标系即坐标系 0 的初始角速度、角加速度均为 0, 即$\boldsymbol{\omega}_0^0=[0,\,0,\,0]^{\mathrm{T}}$、$\dot{\boldsymbol{\omega}}_0^0=[0,\,0,\,0]^{\mathrm{T}}$。同时, 为考虑连杆本身质量的影响, 令$\dot{\boldsymbol{v}}_0^0=[0,\,0,\,g]^{\mathrm{T}}$, 其中, g为重力加速度, 对固定基座施加向上的重力加速度g等同于对连杆系统施加了向下的重力。给定基座初始的状态量, 借由外推关系式 (10)～式 (15) 进行递推计算, 得到各连杆的角速度、角加速度、线加速度等信息。

2.2　气动力模型

确定连杆 3、3N 的角速度后, 根据连杆运动学关系即可得到连杆上任意一点处的速度和来流迎角, 通过叶素理论计算得到桨叶的升力。

为更方便地计算、描述气动力, 在已有坐标系基础上, 新增桨叶坐标系(即坐标系 4)及来流坐标系(即坐标系 5), 分别如图 6 和图 7 所示。坐标系 4 是在坐标系 3 的基础上旋转y_3轴, 使x_3轴与连杆 3 重合, 此时得到的新坐标系即为坐标系 4。坐标系 5 是通过旋转x_4轴, 使y_4轴与来流方向重合, 此时得到的坐标系即为坐标系 5。反桨的坐标系 4N、5N

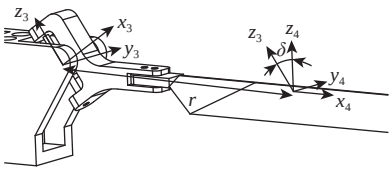


图6 桨叶坐标系

Fig. 6 Blade coordinate system

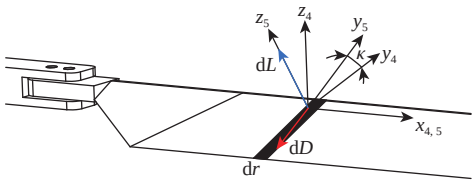


图7 来流坐标系

Fig. 7 Airflow coordinate system

建立规则与正桨一致，不再进行赘述。

在悬停状态下，通过动量叶素理论[18]得到的桨叶诱导速度表达式为

$$v_{\mathrm{ip}}=-\frac{\varpi L_{\mathrm{R}}a\sigma}{16}\left(\sqrt{1+\frac{32(\kappa_0+\theta_{\mathrm{p}})r}{a\sigma L_3}}-1\right) \tag{16}$$

式中：v_{ip}为正桨桨叶微段处的诱导速度；ϖ为旋翼的基准转速；a为翼型的升力线斜率；σ为旋翼的实度；L_{R}为桨盘半径，$L_{\mathrm{R}}=L_1/2+L_3$；r为桨叶微段距离桨叶根部处的距离；κ_0为桨叶安装角。

桨叶微段处的桨叶运动速度为

$$\boldsymbol{v}_{\mathrm{r}}^3=\boldsymbol{R}_1^3\cdot\boldsymbol{\omega}_1^1\times\boldsymbol{P}_2^1+\boldsymbol{\omega}_3^3\times\boldsymbol{R}_4^3\cdot\boldsymbol{P}_{\mathrm{r}}^4 \tag{17}$$

式中：$\boldsymbol{v}_{\mathrm{r}}$为桨叶微段处的桨叶运动速度向量；$\boldsymbol{P}_{\mathrm{r}}$为桨叶微段位置向量。

将桨叶速度取负，与诱导速度叠加，最终求得的来流速度为

$$\boldsymbol{v}_{\mathrm{L}}^4=-\boldsymbol{R}_3^4\cdot\boldsymbol{v}_{\mathrm{r}}^3+\boldsymbol{R}_1^4(v_{\mathrm{ip}}\cdot\hat{\boldsymbol{Z}}_1^1) \tag{18}$$

式中：$\boldsymbol{v}_{\mathrm{L}}$为来流速度向量。

由于$\boldsymbol{v}_{\mathrm{L}}$沿$x_4$轴方向的分量很小，且叶素理论中近似认为来流速度沿桨叶展向分量为0，因此，忽略$\boldsymbol{v}_{\mathrm{L}}$沿$x_4$轴方向的分量，得到来流速度大小及入流角为

$$V_{\mathrm{H}}=\sqrt{v_{\mathrm{L}z}^{4,2}+v_{\mathrm{L}y}^{4,2}} \tag{19}$$

$$\kappa=\arctan\left(\frac{v_{\mathrm{L}z}^4}{v_{\mathrm{L}y}^4}\right) \tag{20}$$

式中：V_{H}为来流合速度；κ为入流角度。由于计算来流速度$\boldsymbol{v}_{\mathrm{L}}$时涉及多次坐标变换，导致来流速度$\boldsymbol{v}_{\mathrm{L}}$的表达式过长，完全写出会过于复杂，以$v_{\mathrm{L}z}$、$v_{\mathrm{L}y}$表示来流速度$\boldsymbol{v}_{\mathrm{L}}$的$z$、$y$轴分量。

在桨叶微段dr处的气动升力dL、气动阻力dD为

$$\mathrm{d}L=\rho C_L c V_{\mathrm{H}}^2\mathrm{d}r/2 \tag{21}$$

$$\mathrm{d}D=\rho C_D c V_{\mathrm{H}}^2\mathrm{d}r/2 \tag{22}$$

式中：ρ为大气密度；c为桨叶微段位置的弦长；C_L、C_D分别为翼型的升力系数、阻力系数。

则在坐标系5下桨叶微段上的气动力向量d$\boldsymbol{f}_{\mathrm{q}}$为

$$\mathrm{d}\boldsymbol{f}_{\mathrm{q}}^5=\left[0,\ -\mathrm{d}D,\ \mathrm{d}L\right]^{\mathrm{T}} \tag{23}$$

如式(24)和式(25)所示，对微段上的气动力、气动力矩进行积分，最终得到施加在关节3上的气动力、气动力矩：

$$\boldsymbol{f}_{\mathrm{q}}^3=\boldsymbol{R}_4^3\left(\int_0^{L_{\mathrm{R}}}\boldsymbol{R}_5^4\cdot\mathrm{d}\boldsymbol{f}_{\mathrm{q}}^5\right) \tag{24}$$

$$\boldsymbol{n}_{\mathrm{q}}^3=\boldsymbol{R}_4^3\left(\int_0^{L_{\mathrm{R}}}\boldsymbol{P}_{\mathrm{r}}^4\times\boldsymbol{R}_5^4\cdot\mathrm{d}\boldsymbol{f}_{\mathrm{q}}^5\right) \tag{25}$$

式中：$\boldsymbol{f}_{\mathrm{q}}^3$、$\boldsymbol{n}_{\mathrm{q}}^3$分别为关节3受到的气动力、气动力矩。

使用同样方法可以得到施加在关节3N上的气动力$\boldsymbol{f}_{\mathrm{qn}}$、气动力矩$\boldsymbol{n}_{\mathrm{qn}}$。

2.3 电机模型

电机为旋翼提供克服旋翼气动阻力及斜铰旋翼进行周期变距时所需的驱动力矩，本文参考文献[15]中的无刷直流电机模型给出与之相似的电机模型：

$$\tau_{\mathrm{m}}=-c_{\mathrm{m}}\dot{\theta}_1-k_{\mathrm{m}}\theta_1+Q_0+K_{\mathrm{e}}\tilde{V}/R_{\mathrm{ohm}} \tag{26}$$

式中：τ_{m}为电机输出的驱动力矩；c_{m}、k_{m}分别为电机的阻尼系数、刚度系数；Q_0为电机的常值力矩；K_{e}、R_{ohm}分别为电机的反电动势常数、电阻；$\tilde{V}$为施加在电机上的余弦波动电压。为实现对电机转速的余弦速度调制，设定波动电压$\tilde{V}=A\cos\theta_1$，其中，A为波动电压幅值。可以通过控制波动电压幅值的大小实现对电机波动转速的控制，进而控制桨叶的挥舞、摆振幅值大小。

2.4 关节摩擦损失

在连杆相连接的关节处存在摩擦阻力，造成能量损失，因此，在旋翼动力学建模中，必须对摩擦损失进行建模。关节中的摩擦损失依据文献[15]，由摩擦阻尼系数的形式给出：

$$n_{\mathrm{m2}}=-I_\beta\varpi C_\beta\dot{\theta}_2 \tag{27}$$

$$n_{\mathrm{m3}}=-I_\beta\varpi C_\zeta\dot{\theta}_3\cos^2\delta \tag{28}$$

式中：n_{m2}、n_{m3}分别为施加在关节2、3上的摩擦力矩；I_β为桨叶的挥舞惯量；C_β、C_ζ分别为挥舞阻尼系数、摆振阻尼系数，其与铰链间的几何尺寸、摩擦

系数、挥舞和摆振幅度有关, 因此, 在不同幅值的激励电压下, 挥舞、摆振阻尼系数也不同, 两者的表达式为

$$C_\beta = \frac{6\mu_1 G_\mathrm{P}(1+e)}{\pi A_\beta (1-e)^2} \tag{29}$$

$$C_\zeta = \frac{6(\mu_1 G_\mathrm{P} + 2\mu_2 G_\mathrm{D}|\tan\delta|/3)(1+e)}{\pi A_\zeta (1-e)^2} \tag{30}$$

其中: μ_1、μ_2为摩擦系数; e为挥舞铰偏置量; G_P、G_D为铰关节的几何系数; A_β、A_ζ分别为桨叶挥舞运动幅值、桨叶摆振运动幅值, 两者均由对应条件下的实验测量结果给出。

同理, 可以得到施加在关节 2N、3N 上的摩擦力矩n_{m2N}、n_{m3N}。

2.5　内推计算

确定连杆所受到的力、力矩后, 需要计算施加在连杆上的力和力矩所对应的关节力矩, 因此, 进行内推计算。由坐标系 3、3N 向坐标系 0 方向进行内推计算。对后杆进行受力分析, 得出后杆施加给前杆的力和力矩, 通过逐级递推计算, 得到关节 1、2、3、2N、3N 处的关节力矩, 取关节力矩中 Z 方向分量, 令其为零, 可以建立 5 个以波动电压为输入量, 以关节角度、角速度、角加速度为待求解量的微分方程。

在内推计算开始时, 连杆 3 为末端连杆, 因此, 上一连杆对其施加的关节力、力矩为零。同时, 连杆 3 还受到气动力作用, 关节 3 受到摩擦力矩作用。最终, 在关节 3 处的关节力、力矩分别为

$$\boldsymbol{f}_3^3 = \boldsymbol{F}_3^3 - \boldsymbol{f}_\mathrm{q}^3 \tag{31}$$

$$\boldsymbol{n}_3^3 = \boldsymbol{N}_3^3 + \boldsymbol{P}_{3\mathrm{c}}^3 \times \boldsymbol{F}_3^3 - \boldsymbol{n}_\mathrm{q}^3 - n_{\mathrm{m3}} \cdot \boldsymbol{Z}_3^3 \tag{32}$$

式中: $\boldsymbol{f}_3$、$\boldsymbol{n}_3$分别为关节 3 处的关节力、关节力矩; $\boldsymbol{Z}_3$为坐标系 3 的 z 轴坐标向量。同理, 可得关节 3N 处的关节力$\boldsymbol{f}_{3\mathrm{N}}$和关节力矩$\boldsymbol{n}_{3\mathrm{N}}$。

计算关节 2 的关节力、关节力矩的递推表达式分别为

$$\boldsymbol{f}_2^2 = \boldsymbol{R}_3^2 \cdot \boldsymbol{f}_3^3 \tag{33}$$

$$\boldsymbol{n}_2^2 = \boldsymbol{R}_3^2 \cdot \boldsymbol{n}_3^3 - n_{\mathrm{m2}} \cdot \boldsymbol{Z}_2^2 \tag{34}$$

式中: $\boldsymbol{f}_2$、$\boldsymbol{n}_2$分别为关节 2 处的关节力、关节力矩; $\boldsymbol{Z}_2$为坐标系 2 的 z 轴坐标向量。同理, 可得关节 2N 处的关节力$\boldsymbol{f}_{2\mathrm{N}}$和关节力矩$\boldsymbol{n}_{2\mathrm{N}}$。

关节 1 处的关节力、力矩来自于关节 2、2N, 则关节 1 的关节力、关节力矩的递推表达式分别为

$$\boldsymbol{f}_1^1 = \boldsymbol{R}_2^1 \cdot \boldsymbol{f}_2^2 + \boldsymbol{R}_{2\mathrm{N}}^{1\mathrm{N}} \cdot \boldsymbol{f}_{2\mathrm{N}}^{2\mathrm{N}} + \boldsymbol{F}_1^1 \tag{35}$$

$$\boldsymbol{n}_1^1 = \boldsymbol{N}_1^1 + \boldsymbol{R}_2^1 \cdot \boldsymbol{n}_2^2 + \boldsymbol{P}_2^1 \times \boldsymbol{R}_2^1 \cdot \boldsymbol{f}_2^2 - \tau_\mathrm{m} \cdot \boldsymbol{Z}_1^1 + \boldsymbol{R}_{1\mathrm{N}}^1 (\boldsymbol{R}_{2\mathrm{N}}^{1\mathrm{N}} \cdot \boldsymbol{n}_{2\mathrm{N}}^{2\mathrm{N}} + \boldsymbol{P}_{2\mathrm{N}}^{1\mathrm{N}} \times \boldsymbol{R}_{2\mathrm{N}}^{1\mathrm{N}} \cdot \boldsymbol{f}_{2\mathrm{N}}^{2\mathrm{N}}) \tag{36}$$

式中: $\boldsymbol{f}_1$、$\boldsymbol{n}_1$分别为关节 1 处的关节力、关节力矩; $\boldsymbol{Z}_1$为坐标系 1 的 z 轴坐标向量。

本节通过内推计算, 由末端连杆向内依次进行递推计算, 对各关节处的关节力矩进行计算, 使 5 个关节力矩的 z 轴方向分量等于 0, 最终得到了 5 个以关节角度、角速度、角加速度为待求解量的微分方程。

2.6　封闭形式下的动力学方程

将关节处的力矩平衡方程进行联立, 可将其写为封闭形式下的动力学方程组:

$$\boldsymbol{\tau}_\mathrm{t}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}, \tilde{V}) = \boldsymbol{M}(\boldsymbol{\theta}) \cdot \ddot{\boldsymbol{\theta}} + \boldsymbol{V}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}) + \boldsymbol{G}(\boldsymbol{\theta}) - \boldsymbol{n}_\mathrm{m}(\dot{\boldsymbol{\theta}}) - \boldsymbol{n}_\mathrm{q}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}) \tag{37}$$

式中: $\boldsymbol{\theta} = \left[\theta_1, \theta_2, \theta_3, \theta_{2\mathrm{N}}, \theta_{3\mathrm{N}}\right]^\mathrm{T}$; $\boldsymbol{\tau}_\mathrm{t}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}, \tilde{V})$为5×1的电机转矩向量; $\boldsymbol{M}(\boldsymbol{\theta})$为连杆的5×5质量矩阵; $\boldsymbol{V}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}})$为5×1的离心力和哥氏力向量; $\boldsymbol{G}(\boldsymbol{\theta})$为5×1的重力向量; $\boldsymbol{n}_\mathrm{m}(\dot{\boldsymbol{\theta}})$为5×1的摩擦力矩向量; $\boldsymbol{n}_\mathrm{q}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}})$为5×1的气动力矩向量。

可以进一步将式(37)改写为

$$\ddot{\boldsymbol{\theta}} = \boldsymbol{M}(\boldsymbol{\theta})^{-1}(\boldsymbol{\tau}_\mathrm{t}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}, \tilde{V}) - \boldsymbol{V}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}) - \boldsymbol{G}(\boldsymbol{\theta}) + \boldsymbol{n}_\mathrm{m}(\dot{\boldsymbol{\theta}}) + \boldsymbol{n}_\mathrm{q}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}})) \tag{38}$$

可以看出, 连杆受到的气动力矩、摩擦力矩及重力、离心力、哥氏力引起的力矩仅为关节角度、角速度的函数。电机扭矩除与关节角度、角速度有关外, 还取决于激励电压。通过改变电机上的激励电压, 可以引起关节角加速度的变化, 进而引起角度、角速度的变化。

3　仿真计算

为对旋翼动力学模型的操纵规律、建模准确性进行检验, 采用文献 [15] 中的旋翼结构参数、电机特性参数、相同的 NACA0012 翼型进行仿真计算, 将仿真计算结果与文献中的实验数据进行比对, 验证模型的准确性。其中, 旋翼结构参数、电机特性参数如表 2 和表 3 所示。

3.1　旋翼周期变距规律验证

斜铰旋翼的周期变距效果来源于电机转速的周期变化, 因此, 为更好地体现模型在相同波动转速输入下的响应情况, 将电机波动转速与基准转速的比值定义为转速变化因子η, 以η作为尺度衡量同一输入下的仿真计算结果与实验数据间的差值, 进而评估动力学模型的准确性。

为验证动力学模型的正确性, 设计仿真实

表2 旋翼结构参数

Table 2 Rotor structure parameters

倾斜铰倾斜角度/(°)	桨盘半径/mm	挥舞铰偏置量	桨毂长度/mm	桨毂质量/g	桨叶半径/mm	升力线斜率	桨叶弦长/mm	桨叶安装角/(°)	旋翼实度	桨叶质量/g	大气密度/(kg·m^{-3})	挥舞惯量/(kg·m^2)	摩擦系数μ_1	摩擦系数μ_2	几何系数G_D	几何系数G_P	基准转速/(rad·s^{-1})
45	159	0.076	24.2	10.5	146.9	5.73	19.3	9	0.077 3	5.40	1.2	3.9×10^{-5}	0.20	0.07	0.012 5	0.003 3	200

表3 电机特性参数

Table 3 Motor characteristic parameter

反电动势常数/(V·(rad·s^{-1})$^{-1}$)	电阻/Ω	电机阻尼系数	电机刚度系数	电机常值力矩
0.009 54	0.305	0.001 2	0.000 94	0.014 1

验对模型的周期变距规律进行验证。在基准转速200 rad/s、转速变化因子$\eta=0.1$的情况下进行仿真实验，应用式(7)～式(9)将得到的关节转角变化曲线解算为桨叶的挥舞、摆振、变距运动变化曲线，以方位角为横坐标对不同方位角下桨叶的挥舞、摆振、变距运动变化绘制曲线。

观察图8～图10可以发现，正、反桨叶的挥舞、摆振、变距运动在电机波动转速的激励下均呈现周期性波动变化的规律。在图8中，正、反桨叶的挥舞响应近似对称分布，在方位角为90°时，正桨挥舞角达到最大值，而反桨挥舞角为最小值，在电机旋转180°后，即在方位角为−90°的情况下，反桨挥舞角取得了最大值，而正桨挥舞角为最小值。这

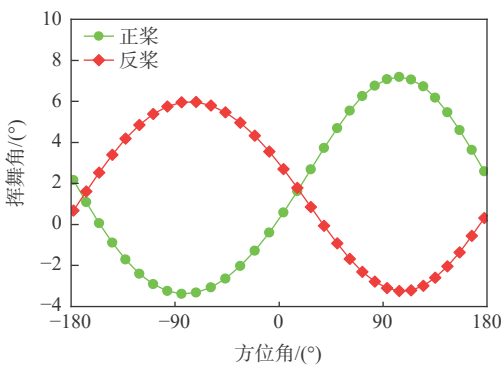


图8 桨叶挥舞角随方位角变化曲线

Fig. 8 Change curve of blade flapping angle with azimuth angle

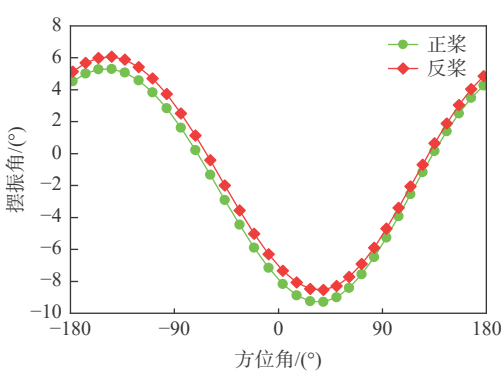


图9 桨叶摆振角随方位角变化曲线

Fig. 9 Change curve of blade swing angle with azimuth angle

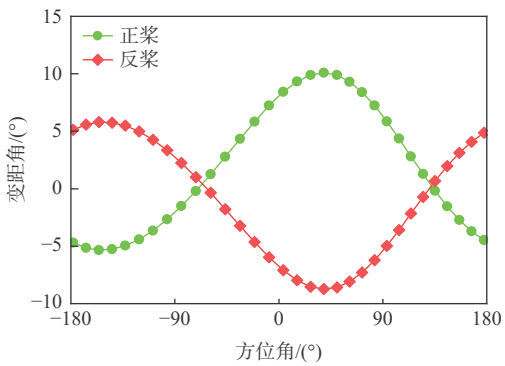


图10 桨叶变距角随方位角变化曲线

Fig. 10 Change curve of blade pitch angle with azimuth angle

种变化说明桨叶在受到波动转速的激励下，旋翼桨盘向一侧进行了倾倒。同时可以发现，正、反桨叶的挥舞曲线虽然近似对称，但反桨的挥舞运动幅度略弱于正桨，造成了正、反桨叶挥舞平面的不一致。

在图9的桨叶摆振曲线中可以发现，正、反桨叶的摆振运动近似一致，一同呈周期性变化，在方位角为45°时，正、反桨叶达到后摆的最大值，在方位角为−135°时，正、反桨叶的前摆最大。可以发现，桨叶虽然在进行前后摆振运动，但其摆振运动的均值并不等于零。正、反桨叶摆振角度的平均值都为负数，说明由于气动阻力的影响，导致正、反桨叶在恒定转速的情况下，均存在一定的稳态后摆角度。由于正、反桨叶受到的气动阻力近似相同，不同桨叶稳态后摆角度也近似一致，正、反桨叶的摆振运动呈现出一致性。

在图10的桨叶变距曲线中，正、反桨叶变距角的变化曲线呈现对称变化的规律。在45°方位角时，桨叶的后摆角度最大，此时，摆振角处于最小值，而正桨由于摆振运动与变距运动的负耦合关系，此时，正桨的变距角达到最大值，与之相反的反桨变距角由于摆振运动与变距运动的正耦合关系，取得最小值。而在方位角为−135°时，桨叶的前摆最大，正桨的变距角最小，反桨的变距角取得最大

值。这种变距角变化的差异性说明了正、反桨叶摆振-变距耦合关系的不同。同时观察变距曲线可以发现, 虽然不同桨叶上变距角的变化幅度是一致的, 但正、反桨叶变距曲线的平均值却不相同。这是由于桨叶稳态后摆角度的存在使正桨在恒定转速情况下的桨距增加, 而反桨在恒定转速下的桨距减小, 导致正、反桨叶在转速恒定时的桨距角不同, 从而使正桨变距角变化曲线的均值大于反桨变距角变化曲线的均值。

正、反桨叶的变距角均呈周期性变化, 这样的周期变距响应变化说明动力学模型计算结果呈现出了与自动倾斜器类似的周期变距效果, 验证了斜铰旋翼的周期变距操纵规律。

3.2　动力学模型准确性验证

为进一步评估斜铰旋翼动力学模型的准确性, 通过给定不同幅值下的波动电压进行动力学仿真实验, 以转速变化因子η为横坐标, 对正、反桨叶的挥舞、摆振响应幅值绘制曲线。

如图 11 所示, 本文建立的动力学模型预测得到的正、反桨叶的摆振幅值与转速变化因子η呈线性变化关系, 随着η的增大, 正、反桨叶的摆振幅值均增大。与文献 [15] 建立的动力学模型相比, 本文建立的动力学模型对摆振幅值的预测更为精准, 误差更小。在转速变化因子$\eta=0.096$的激励下, 实验测得摆振幅值为7.4°, 文献 [15]、本文动力学模型对摆振幅值的预测值分别为6.33°、7.0°。以实验测定值作为真值, 则文献 [15]中动力学模型预测摆振幅值的相对误差为14.46%。本文动力学模型预测摆振幅值的相对误差为5.41%。相较于原模型, 本文模型在大波动转速激励下的预测精度更高, 其中, 在转速变化因子$\eta=0.096$激励下, 对摆振幅值的相对误差预测精度提升了 9.05%, 应用本文动力学模型仿真计算得到的预测结果与实验测定结果间的误差较小, 且变化趋势一致。

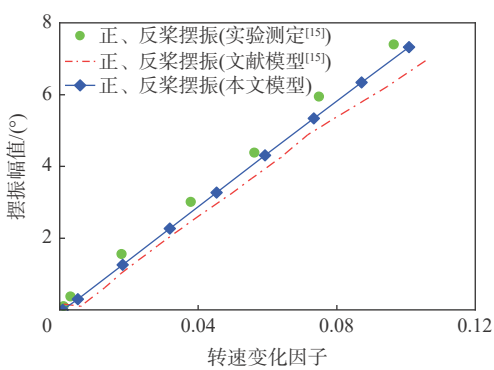


图 11　转速变化因子对摆振运动幅值的影响

Fig. 11　Influence of speed change factor on swing motion amplitude

如图 12 所示, 在波动转速为零时, 本文仿真计算得到的正、反桨叶的挥舞幅值均为零, 伴随着转速变化因子η的增大, 正、反桨叶的挥舞幅值也随之增大, 正、反桨叶的挥舞响应差异也越来越大。与文献 [15] 的预测结果相比, 本文模型对于正、反桨叶间的挥舞差异预测更为精准, 与实验结果符合更好, 说明了本文动力学模型的准确性。

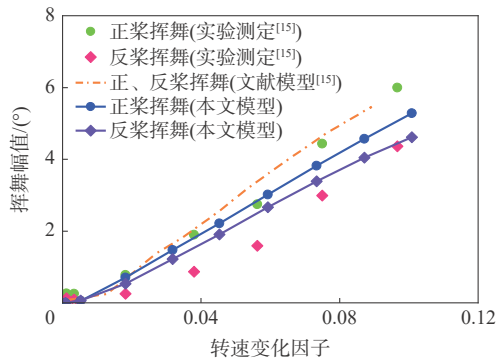


图 12　转速变化因子对挥舞运动幅值的影响

Fig. 12　Influence of speed change factor on flap motion amplitude

进一步分析正、反桨叶间挥舞运动差异性的产生原因可知, 正、反桨叶间的挥舞响应差异性来源于桨毂结构的不对称, 气动阻力的存在使桨叶与桨毂在转速恒定时存在一个固定的摆振滞后角度, 由于正、反桨叶结构的不同, 该滞后角度使正桨的变距角增加, 而反桨的变距角减小, 从而导致旋翼稳态时正桨的来流迎角大于反桨的来流迎角, 因此, 在施加同等激励的情况下, 正桨的挥舞幅值大于反桨的挥舞幅值, 这种桨间挥舞差异为旋翼带来了更剧烈的桨叶振动与桨间不平衡, 因此, 应着力减小转速恒定时桨叶的摆振滞后角度, 以减弱桨叶间的挥舞差异。

4　结　论

1) 本文采用多局部连杆坐标系方法描述斜铰旋翼复杂的桨叶摆振、变距耦合运动, 相较于原模型中采取小角度假设近似描述桨叶运动, 该方法具有更高的精准度, 对于桨叶摆振、变距运动相耦合的复杂运动现象的运动学描述效果则更为突出。

2) 在给定波动电压$\tilde{V}=A\cos\theta_1$的条件下进行周期变距仿真实验。仿真计算结果表明, 本文建立的动力学模型在周期变化的电机转速激励下, 桨叶可以产生类似于自动倾斜器的周期变距效果, 验证了本文动力学模型的正确性。

3) 本文建立的斜铰旋翼动力学模型具有较高的预测精度, 对于摆振运动幅值的预测, 在转速变化因子$\eta=0.096$的情况下, 相较于原模型的预测精

度提高了 9.05%。在桨叶挥舞运动的预测方面，与原模型相比，本文模型正确模拟了正、反桨叶间的挥舞差异性，与实验测定数据符合很好。相较于原模型，本文动力学模型可以更好地模拟桨叶挥舞、摆振、变距运动现象，具有更高的置信度。

4）本文进一步分析了造成桨间挥舞差异的原因。分析表明，正、反桨间的结构差异导致的稳态时的桨叶摆振滞后角可能是造成正、反桨叶间挥舞差异的主要原因。斜铰旋翼的这种桨间挥舞差异性增大了斜铰旋翼的应用难度，在未来的研究中，应从斜铰旋翼本身结构着手以寻求削减正、反桨叶间挥舞差异性的可行方案。

参考文献（References）

[1] 邓景辉. 直升机技术发展与展望[J]. 航空科学技术, 2021, 32(1): 10-16.
DENG J H. Development and prospect of helicopter technology[J]. Aeronautical Science & Technology, 2021, 32(1): 10-16(in Chinese).

[2] 李路明, 王立鼎. MEMS 研究的新进展: 微型系统及其发展应用的研究[J]. 光学精密工程, 1997, 5(1): 67-73.
LI L M, WANG L D. Microsystem and its development and applications[J]. Optics and Precision Engineering, 1997, 5(1): 67-73(in Chinese).

[3] 陈世适, 姜臻, 董晓飞, 等. 微小型飞行器发展现状及关键技术浅析[J]. 无人系统技术, 2018, 1(1): 38-53.
CHEN S S, JIANG Z, DONG X F, et al. Overview of micro aerial vehicles and its key technologies[J]. Unmanned Systems Technology, 2018, 1(1): 38-53(in Chinese).

[4] 尹欣繁, 张代兵, 张纪阳, 等. 微型无人直升机发展现状及关键技术[J]. 飞航导弹, 2018(5): 33-37.
YIN X F, ZHANG D B, ZHANG J Y, et al. Development status and key technologies of micro unmanned helicopter[J]. Aerodynamic Missile Journal, 2018(5): 33-37(in Chinese).

[5] KIM S, TILBURY D. Mathematical modeling and experimental identification of a model helicopter[C]//Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit. Reston: AIAA, 1998: 4357.

[6] DADKHAH N, METTLER B. System identification modeling and flight characteristics analysis of miniature coaxial helicopter[J]. Journal of the American Helicopter Society, 2014, 59(4): 36-51.

[7] PAULOS J, YIM M. An underactuated propeller for attitude control in micro air vehicles[C]//Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2014: 1374-1379.

[8] PAULOS J, YIM M. Flight performance of a swashplateless micro air vehicle[C]//Proceedings of the 2015 IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2015: 5284-5289.

[9] 尹欣繁. 无斜盘微型无人直升机设计与飞行控制研究[D]. 长沙: 国防科技大学, 2017: 5-10.
YIN X F. Design and flight control of a swashplateless micro unmanned helicopter[D]. Changsha: National University of Defense Technology, 2017: 5-10(in Chinese).

[10] 黄灿铭. 微小型无斜盘双旋翼飞行器的设计及实现[D]. 北京: 北京邮电大学, 2021: 10-20.
HUANG C M. Implementation of micro swashplateless twin rotor aircraft[D]. Beijing: Beijing University of Posts and Telecommunications, 2021: 10-20(in Chinese).

[11] 王彪, 任至楷, 张宝晴, 等. 斜铰旋翼操纵力矩生成研究[J]. 机械制造与自动化, 2021, 50(2): 102-105.
WANG B, REN Z K, ZHANG B Q, et al. Research on the moment generation of tilt-hinge rotor[J]. Machine Building & Automation, 2021, 50(2): 102-105(in Chinese).

[12] 任至楷. 电动斜铰旋翼研发及其在无人直升机上的应用[D]. 南京: 南京航空航天大学, 2020: 5-10.
REN Z K. Research and development of electric tilt-hinge rotor and its application in unmanned helicopter[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2020: 5-10(in Chinese).

[13] 徐家梁. 电动旋翼飞行器动力/推进操纵新技术研究[D]. 南京: 南京航空航天大学, 2018: 8-18.
XU J L. Research on power/propulsion control technology of electrical-driving rotor systems[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2018: 8-18(in Chinese).

[14] 张宝晴. 基于电动斜铰旋翼的纵列式双旋翼无人直升机研究[D]. 南京: 南京航空航天大学, 2021: 7-18.
ZHANG B Q. Research of an unmanned tandem-rotor helicopter based on electric tilt-hinge rotors[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2021: 7-18(in Chinese).

[15] PAULOS J, YIM M. Cyclic blade pitch control without a swashplate for small helicopters[J]. Journal of Guidance, Control, and Dynamics, 2018, 41(3): 689-700.

[16] YU J, LIU Y, YANG W D. Dynamics research on actively controlled swashplateless rotor[J]. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 2019, 233(12): 4492-4508.

[17] CRAIG J J. Introduction to robotics: mechanics and control[M]. 4th ed. London: Pearson Education Inc, 2005: 70-80.

[18] LEISHMAN J G. Principles of helicopter aerodynamics[M]. 2nd ed. New York: Cambridge University Press, 2006: 80-127.

附录 A

牛顿-欧拉递推动力学算法外推计算过程中需要用到的惯性张量为

$$\boldsymbol{I}_1^{C1}=\begin{bmatrix}0 & 0 & 0\\ 0 & \dfrac{m_1L_1^2}{48} & 0\\ 0 & 0 & \dfrac{m_1L_1^2}{48}\end{bmatrix} \tag{A1}$$

$$\boldsymbol{I}_2^{C2}=\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{bmatrix} \tag{A2}$$

$$\boldsymbol{I}_{2\mathrm{N}}^{\mathrm{C2N}}=\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix} \tag{A3}$$

$$\boldsymbol{I}_{3}^{\mathrm{C3}}=\begin{bmatrix}M_{11}&0&M_{13}\\0&M_{22}&0\\M_{13}&0&M_{33}\end{bmatrix} \tag{A4}$$

$$\boldsymbol{I}_{3\mathrm{N}}^{\mathrm{C3N}}=\begin{bmatrix}M_{11}&0&-M_{13}\\0&M_{22}&0\\-M_{13}&0&M_{33}\end{bmatrix} \tag{A5}$$

$$M_{11}=\frac{m_3L_3^2\sin^2\delta}{12} \tag{A6}$$

$$M_{13}=\frac{m_3L_3^2\sin\delta\cos\delta}{12} \tag{A7}$$

$$M_{22}=\frac{m_3L_3^2}{12} \tag{A8}$$

$$M_{33}=\frac{m_3L_3^2\cos^2\delta}{12} \tag{A9}$$

Dynamic model of high confidence tilt-hinge rotor based on Newton-Euler recursion algorithm

YANG Yisong，LI Jianbo*，DUAN Dengyan

(National Key Laboratory of Helicopter Aeromechanics，School of Aeronautics，
Nanjing University of Aeronautics and Astronautics，Nanjing 210016，China)

Abstract：The tilt-hinge rotor has a simpler structure because it doesn't need a swashplate and can provide cyclic pitch control just by accelerating and decelerating the motor. However, the tilt hinge coupled with the rotor lag and pitch motion complicates the modeling of rotor dynamics. However, the rotor lag and pitch motion are coupled with the tilt hinge, which makes the rotor dynamics modeling more complicated. The nearby linkage coordinate system of the blade was created based on the enhanced Denavia-Hartenberg approach in order to address the issues of low model prediction accuracy in the current modeling techniques and inadequate disclosure of the difference between forward and reverse blade flap-motion. The Newton-Euler recursion algorithm is used to calculate the velocity and acceleration of each linkage and the interaction force and torque of each linkage in the local linkage coordinate system. The dynamic model of the tilt-hinge rotor is established. On this basis, the mechanism of periodic pitch variation of the tilt-hinge rotor is further revealed through simulation calculation. At the same time, the calculation results show that this model can predict the flapping difference between forward and reverse blades more accurately. The prediction accuracy of blade lagging amplitude is improved by 9.05%.

Keywords：tilt-hinge rotor; swashplateless; lag-pitch coupling; dynamical model; cyclic pitch control

Received：2024-04-18；**Accepted**：2024-05-24；**Published Online**：2024-06-20 11：29

URL：link.cnki.net/urlid/11.2625.V.20240619.1210.001

* **Corresponding author**. E-mail：ljb101@nuaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0210

基于空气品质及代偿损失的客舱回风比例优化

高阳，林家泉*
(中国民航大学 电子信息与自动化学院，天津 300300)

摘　　要：为使飞机客舱获得较好的空气品质，同时提高客舱空调的节能水平，采用计算流体力学(CFD)方法建立波音737客机经济舱的仿真模型，并结合粒子图像测速(PIV)技术验证仿真模型的准确性。在此模型基础上，研究客舱空调使用不同回风比例对客舱内CO_2浓度场的影响，并采用乘客呼吸区通风效率指标评估不同回风比例对舱内空气品质的影响。此外，使用起飞总质量法计算相同送风量时不同回风比例工况下的燃油代偿损失，并拟合出燃油代偿损失、乘客呼吸区通风效率与回风比例的函数关系，采用功效系数法构建评价函数，得到客舱空调的最佳回风比例为64.864%。所提方法能够为客舱空调的回风比例控制提供依据。

关　键　词：飞机客舱；回风比例；空气品质；代偿损失；评价函数；通风效率
中图分类号：V245.3
文献标志码：A　　**文章编号：**1001-5965(2026)06-2034-08

环境控制系统是现代飞机的重要系统，用于保持客舱内适宜的压力、温度、湿度及空气质量等。目前，大多数民用客机正常飞行时都采用发动机引气作为环境控制系统的主要气源[1]。来自发动机的高温高压引气分为2路进入环境控制系统，其中一路引气经过一系列降温降压过程后成为冷空气，另一路引气作为热空气与冷空气混合，达到目标温度后进入客舱，并在舱内流动、扩散，最终部分排出舱外。

相关研究表明，在空调总供气量不变的情况下，循环利用一部分过滤净化后的客舱回风能够减少使用发动机引气量，从而节省燃油消耗[2]。Herbig等[3]研究得出，对于波音B767机型，客舱空调全部使用发动机引气供气，产生的燃油消耗占飞机总燃油消耗的1.6%，而使用50%新风和50%回风的混合风作为空调气源，产生的燃油消耗只占飞机总燃油消耗的0.8%。一方面，提高客舱回风比例能够节省燃油，减少飞行碳足迹。另一方面，提高回风比例会降低客舱内的空气品质，如增加CO_2体积分数，这是因为减少了用于稀释的外部新鲜空气。而乘客呼吸不断产生CO_2，当客舱内的CO_2体积分数较高时，会使人体血液中的碳酸体积分数增大，使乘客感到不适[4-5]。因此，从空调能耗和客舱的空气品质2个角度研究飞机客舱空调的回风比例具有重要意义。

目前，国内外对于飞机客舱空调回风已经有不少研究。Herbig等[3]研究了飞机客舱空调提高回风比例对乘客健康和舒适性的影响，发现随着回风比例的增加，乘客健康和舒适性会系统性下降；高经诚等[5]研究了飞机空调采用30%和50%回风比例对客舱内流场环境的影响，得出当回风比例为50%时，客舱内的CO_2体积分数明显升高；陈希远等[6]研究了有回风和无回风2种工况下飞机客货舱内引气污染物的扩散规律；杨建忠等[7]研究了不同送风形式下，改变空调回风比例对客舱内引气污染物的影响，发现在污染物累积阶段，提高回风比例可以降低客舱内的引气污染物体积分数。目前，对于客舱空调回风比例的研究只包含有无回风或部分回风比例工况，且飞机客舱空调通常采用50%的回

收稿日期：2024-04-10；录用日期：2024-06-07；网络出版时间：2024-07-25 11:47
网络出版地址：link.cnki.net/urlid/11.2625.V.20240725.0910.001
基金项目：工业和信息化部民机专项(2020020306)
*通信作者. E-mail：jqlin@cauc.edu.cn

引用格式：高阳，林家泉. 基于空气品质及代偿损失的客舱回风比例优化[J]. 北京航空航天大学学报，2026，52(6)：2034-2041.
GAO Y，LIN J Q. Optimization of cabin return air ratio based on air quality and compensation loss[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：2034-2041 (in Chinese).

风比例, 是否存在更合适的回风比例, 能同时满足客舱高空气品质与空调低能耗的需求, 当前文献研究较少。

本文以乘客呼吸区通风效率作为客舱空气品质的评价指标, 采用计算流体力学 (computational fluid dynamics, CFD) 方法对不同回风比例下的客舱内空气品质进行了研究, 并用起飞总质量法研究客舱空调不同回风比例对飞机燃油消耗的影响, 分别得到乘客呼吸区通风效率、燃油代偿损失与回风比例的函数关系; 采用功效系数法构造客舱空调最佳回风比例评价函数, 得到优化的回风比例。本文方法能够为客舱空调的回风比例控制提供依据。

1　研究方法

1.1　飞机客舱 CFD 模型

图 1 为经过简化后的波音 737 客舱 CFD 仿真模型, 共包含 2 个位于过道上方的天花板送风口、10 个位于行李架下方的侧壁送风口、10 个位于客舱侧壁面底部的排风口和对称分布在过道两侧的 5 排 6 列共 30 个座椅。

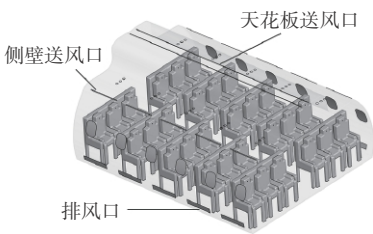


图 1　客舱 CFD 仿真模型

Fig. 1　CFD simulation model of cabin

对整个客舱模型进行四面体网格划分, 并在送风口、排风口等复杂流场区域进行局部网格加密[8-9]。划分的网格数量为 501 万个。

1.2　数值模型

使用 FLUENT 软件对客舱模型进行瞬态模拟计算。研究表明, 重整化群 (renormalization group, RNG) k-ε湍流模型可准确模拟客舱内气流组织和污染物传播规律[10]。因此, 本文选用 RNG k-ε湍流模型对客舱内气流场进行模拟, 湍流模型的控制方程如下:

$$\rho\frac{\partial\bar{\phi}}{\partial t}+\rho\bar{u}_i\frac{\partial\bar{\phi}}{\partial x_i}-\frac{\partial}{\partial x_i}\left(\Gamma_{\phi,\mathrm{eff}}\frac{\partial\bar{\phi}}{\partial x_i}\right)=S_\phi \tag{1}$$

式中: ρ为空气密度, kg/m^3; $\bar{\phi}$为流体变量均值; t为时间, s; $\Gamma_{\phi,\mathrm{eff}}$为有效扩散系数; $\bar{u}_i$为x_i方向的平均速度, m/s; S_ϕ为源项。

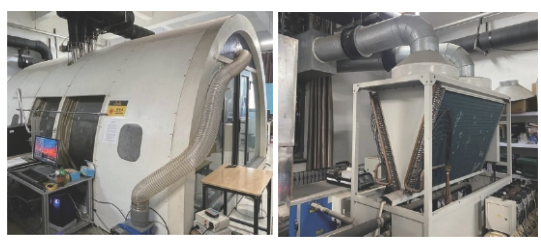
计算时采用协调一致的半隐式压力连接方程法 (semi-implicit method for pressure-linked equations consistent, SIMPLEC) 计算方案求解压力速度耦合, 压力等各参数的空间离散均选用二阶迎风格式, 最终根据残差大小及出口流量稳定情况来判定收敛。

1.3　客舱 CFD 模型验证

由于简化了客舱模型部分区域, 且使用的湍流模型中存在部分假设, 为确保后续计算的准确性, 需对客舱仿真模型进行实验验证。

采用粒子图像测速 (particle image velocimetry, PIV) 技术进行实验验证。实验参数设置和 CFD 仿真的边界条件保持一致: 送风量为 1 200 m^3/h, 送风温度为 295.15 K, 天花板送风口和侧壁送风口的送风速度分别设置为 2 m/s 和 1.5 m/s。

图 2 为波音 737 客舱实验平台, 由实验舱和空调机组构成, 通过空调控制平台可对实验舱的送风量和送风温度进行调节。PIV 系统主要由同步器、激光器、计算机和 16 M 像素的电荷耦合器件 (charge-coupled device, CCD) 相机组成。首先, 使用烟雾发生器产生平均粒径约为 1.5 μm 的示踪粒子, 从送风口送入实验舱。然后, 通过同步器控制激光器和 CCD 相机工作, 在激光器发射激光脉冲的同时, 使用 CCD 相机拍摄示踪粒子的图像。最后, 将图像信息传回计算机, 处理后得到采样区域的流场数据。



(a) 实验舱　　(b) 空调机组

图 2　波音 737 客舱实验平台

Fig. 2　Boeing 737 cabin experimental platform

如图 3 所示, 在客舱第 2 排与第 3 排中间, 共采样 4 个大小为 900 mm×600 mm 的子区域, 利用数据处理将 4 个子区域进行拼接, 得到整片区域的流场数据。如图 4 所示, 将 PIV 实验得到的流场数据与 CFD 仿真得到的流场数据进行对比, 得到两者流场的运动趋势基本一致。

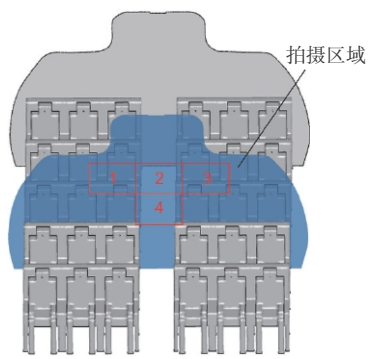


图 3　PIV 实验拍摄区域

Fig. 3　PIV experimental shooting area

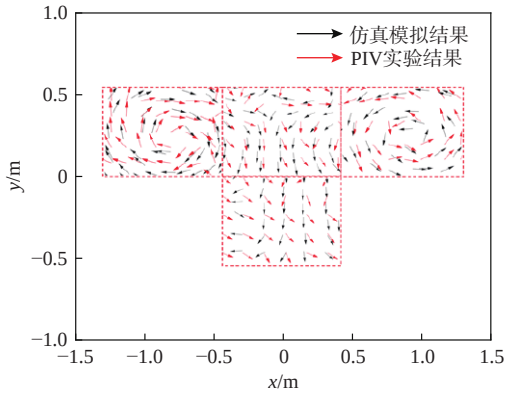


图4 PIV实验和仿真结果

Fig. 4 PIV experimental and simulation results

为进一步验证仿真模型的准确性，选取7个采样点处速度进行对比。采样点位置如图5所示。

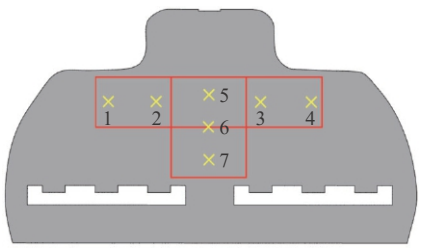


图5 采样点示意图

Fig. 5 Sampling point diagram

如图6所示，得到采样点处速度平均误差为8.62%，对比文献[11]可知，误差处于合理范围内。因此，可以验证所建立的飞机客舱仿真模型准确。

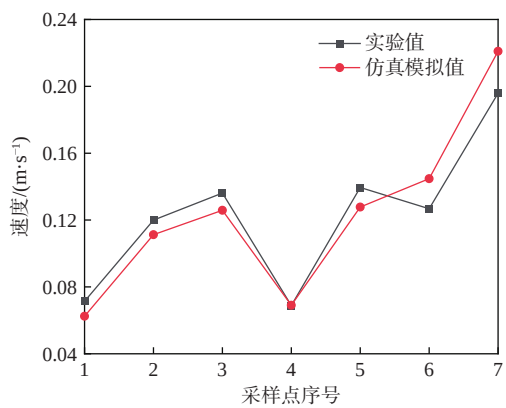


图6 PIV实验和仿真采样值对比

Fig. 6 Comparison of PIV experiment and simulation sampling values

1.4 网格独立性验证

选取数量为211万、501万和776万的网格进行网格独立性验证，模拟回风比例为50%工况下的客舱内流场。选取第3排座椅上方7个采样点的CO_2体积分数进行对比，如图7所示。可以看出，网格数量为501万和776万2种情况下的采样点数值较为接近，说明数量为501万的网格可以获得较为准确的流场信息，增加网格对计算结果无明显影响。而数量为211万的网格计算结果差异较大。因此，选用501万数量的网格进行模拟。

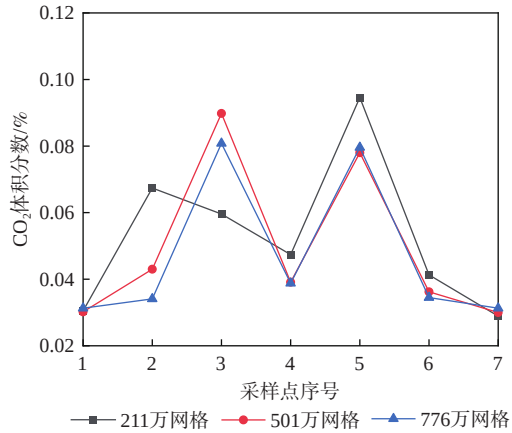


图7 网格独立性分析

Fig. 7 Grid independence analysis

1.5 仿真边界条件设定

对客舱内部进行不同回风条件下的模拟计算，通过用户自定义函数(user defined functions, UDF)功能编译函数载入FLUENT中，分别进行回风比例为30%、40%、50%、60%、70%、80%的计算。具体边界条件为：

1）回风气体组分设置。由于高效空气过滤器并不能清除客舱回风中的CO_2气体，CO_2仍会随着客舱回风重新进入客舱[12]。因此，利用UDF功能进行新风与回风混合后的CO_2组分计算，作为送风口实际的CO_2组分。

2）送风口边界条件。采用天花板及侧壁供风的混合送风方式。送风口设定为质量流量入口。根据民机客舱气流组织方案，供给客舱的新鲜空气每人最少应为0.006 kg/s[13]。因此，客舱总的送风量应为0.18 kg/s。考虑到乘客的舒适性，增大总送风量为0.36 kg/s，送风方向为垂直壁面，送风温度设置为295.15 K。

3）排风口边界条件。排风口设定为边界出口。

4）壁面边界条件。设定客舱壁面、座椅的温度为295.15 K。

5）人体模型边界条件。假设人体边界条件热流密度为25.4 W/m^2。

6）CO_2气体出口设置。在客舱内人体模型的头部截取半径为0.008 m的圆形面作为CO_2气体出口，由于一个成年人在平稳呼吸时的CO_2释放速率约为18 L/h[14]，因此，设置CO_2呼出速度为0.024 8 m/s。

2 仿真结果及分析

为研究飞机客舱空调在相同送风量时不同回风比例对客舱内空气品质的影响，分别进行回风比例为30%、40%、50%、60%、70%、80%的仿真计

算。每组模拟仿真时间为 180 s, 最终计算收敛。

为更直观地显示回风比例对客舱内 CO_2 浓度场的影响, 在客舱内截取图 8 所示的 2 个典型位置平面作为采样截面。

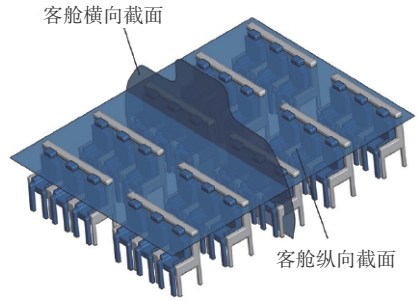


图 8　采样截面示意图

Fig. 8　Sampling cross-section diagram

2.1　客舱 CO_2 浓度场仿真

图 9 为客舱第 3 排区域横向截面的 CO_2 体积分数场对比。可知, 乘客呼出的 CO_2 随着气流流动积聚在行李架底部, 造成乘客呼吸区及上方区域 CO_2 体积分数较大, 空气品质较差, 而距离乘客呼吸位置较远的区域 CO_2 体积分数相对较低。图 10 为乘客呼吸区客舱纵向截面的 CO_2 体积分数场对比。可以看出, 由于送风口送入新鲜空气, 且过道处无座椅和人体阻挡, 因此, 过道处 CO_2 体积分数相对其他位置较低, 空气品质较好。

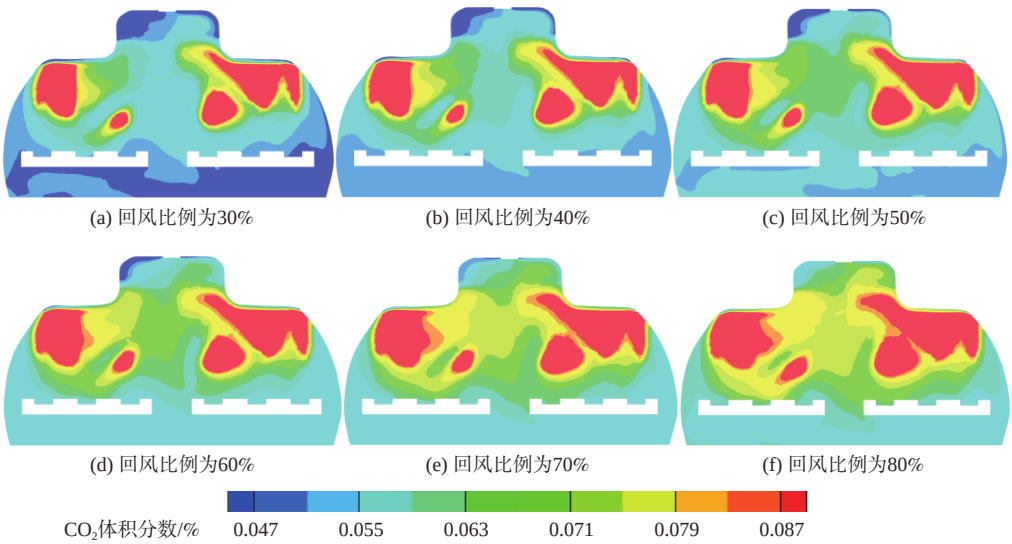


图 9　客舱横向截面的 CO_2 浓度场对比

Fig. 9　CO_2 concentration field comparison of cabin cross section

对比不同回风比例工况下的 CO_2 体积分数场可以看出, 随着回风比例的增大, 客舱内 CO_2 体积分数明显增大。这是因为 CO_2 随着客舱回风重新进入客舱, 导致送风口送出的气体中 CO_2 体积分数增大。

2.2　乘客呼吸区通风效率与回风比例的函数关系

CO_2 体积分数分布会影响飞机客舱内的空气品质, 从而对乘客的健康和舒适性产生影响。通过对 CO_2 体积分数场的分析可以得出, 客舱内乘客呼吸区 CO_2 体积分数较大, 空气品质较差, 且不同回风比例对客舱内空气品质影响较大。因此, 为更好地研究回风比例与客舱内空气品质的关系, 采用乘客呼吸区通风效率作为评价指标, 建立通风效率与客舱空调回风比例之间的函数关系, 进而优化客舱空调的回风比例。乘客呼吸区通风效率公式[15] 为

$$\varepsilon = \frac{C_1 - C_0}{C_2 - C_0} \tag{2}$$

式中: ε为通风效率; C_0为送风口 CO_2 体积分数; C_1为排风口 CO_2 体积分数; C_2为呼吸区 CO_2 体积分数。ε越大, 说明 CO_2 排出舱外效果越好, 乘客呼吸区通风效果越好。

根据对 30%、40%、50%、60%、70%、80% 回风比例下客舱流场的数值模拟结果, 得到不同回风比例下乘客呼吸区、送风口、排风口处的 CO_2 平均体积分数。根据式(2)分别求出不同回风比例下的乘客呼吸区通风效率, 如图 11 中离散点所示。

由于仅对离散的回风比例进行了数值模拟, 因此, 未能计算出 30%～80% 区间内所有回风比例下的乘客呼吸区通风效率。为研究回风比例的变化对通风效率的影响, 将计算求得的离散的通风效率进行高斯拟合, 拟合生成的曲线如图 11 所示, 乘客呼吸区通风效率与回风比例的关系为

$$\varepsilon(X) = 72.24\exp\left(-\left(\frac{X-13.64}{57.91}\right)^2\right) + 23.79\exp\left(-\left(\frac{X-71.29}{33.81}\right)^2\right) \tag{3}$$

式中: X为回风比例; $\varepsilon(X)$为回风比例为X时的通风效率。

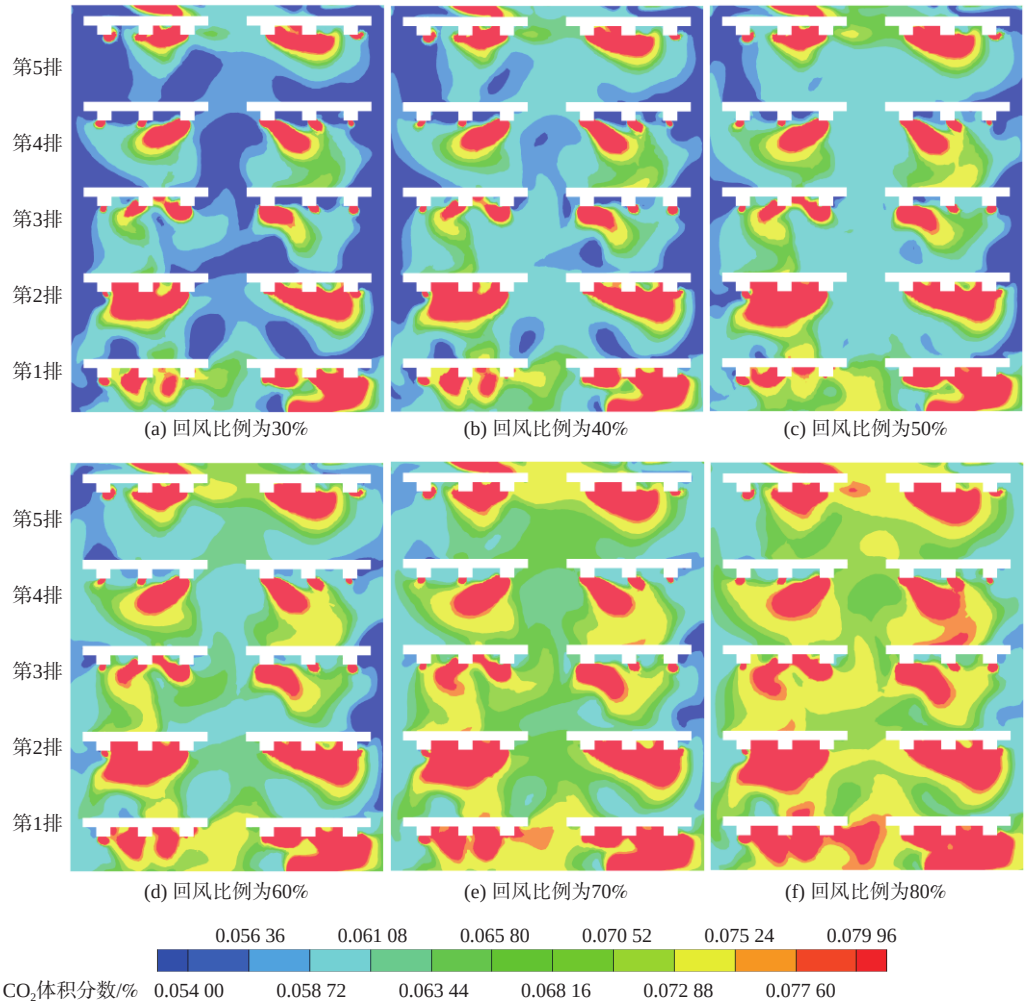


图 10 客舱纵向截面的 CO_2 浓度场对比

Fig. 10 CO_2 concentration field comparison of cabin longitudinal section

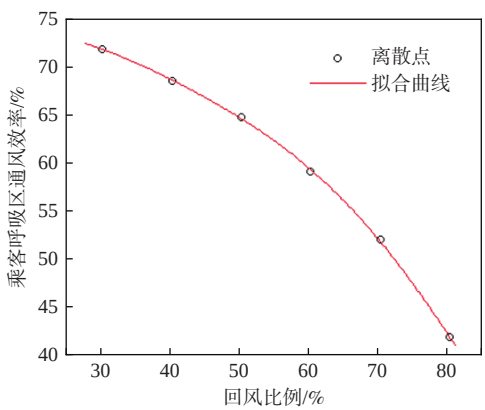


图 11 乘客呼吸区通风效率与回风比例函数关系

Fig. 11 Function relationship between ventilation efficiency and return air ratio in breathing area

由图 11 可知，随着回风比例增加，乘客呼吸区通风效率逐渐降低，且下降速率增大，可以得出，回风比例越大，客舱的通风排污效果越差。

3 燃油代偿损失计算与分析

客舱空调回风比例的改变不仅对客舱内的 CO_2 浓度场有影响，还对飞机环境控制系统的燃油消耗有一定影响。飞机客舱空调采用不同的回风比例，在总送风量不变的情况下，会改变发动机引气的使用量，从而引起系统能耗的变化。

3.1 起飞总质量法

飞机的环境控制系统运行需要消耗能量，最后一并归结到飞机燃油的消耗上。许多因素都会增加燃油的代偿损失，如发动机引气量、发动机输出轴功率等，从而增加飞机的起飞总质量。这种用起飞总质量的增加值作为评价系统的指标，估算不同系统方案的方法，称为起飞总质量法[16]。采用起飞总质量法研究不同回风比例对燃油消耗的影响，从而优化回风比例。

导致飞机起飞总质量增加的因素为发动机引气量。因此，需要计算从发动机引气时产生的附加燃油消耗量 $q_{m,f,bl}$。当从发动机压气机引气时，通过发动机涡轮的流量减少，使推动飞机的可用功率下降。为补偿这部分推力损失，需要在发动机的涡轮入口处补充加热，对此产生的附加燃油

消耗量$q_{m,f,bl}$为[17]

$$q_{m,f,bl}=q_{m,bl}C_e v+\frac{3\,600q_{m,bl}c_pT_3}{H_u\varepsilon_c} \tag{4}$$

式中: v为飞行速度, m/s; C_e为燃油比耗, kg/(N·h); $q_{m,bl}$为引气流量, kg/s; T_3为涡轮进口温度, K; c_p为燃气质量定压热容, J/(kg·K); H_u为燃油燃烧的单位热值, kJ/kg; ε_c为燃油燃烧完全系数。

而运输附加的燃油$q_{m,f,bl}$会消耗额外的燃油。因此, 总的燃油代偿损失$m_{f,bl}$可写为[17]

$$m_{f,bl}=\frac{q_{m,f,bl}K}{C_e g}\left(\exp\left(\frac{C_e\tau_0 g}{K}\right)-1\right) \tag{5}$$

式中: K为升阻比; τ_0为飞行时间, h; g为重力加速度, m/s^2。

根据式(5)计算不同回风比例下的燃油代偿损失。计算时采用的 CFM56 发动机特性参数如表 1 所示[18-19]。

表 1　CFM56 发动机特性参数

Table 1　CFM56 engine characteristic parameters

升阻比	飞行速度/(m·s^{-1})	计算飞行时间/h	燃油比耗/(kg·(N·h)$^{-1}$)	燃气质量定压热容/(J·(kg·K)$^{-1}$)	涡轮进口温度/K	燃油燃烧的单位热值/(kJ·kg^{-1})	燃油燃烧完全系数
16	272	1	0.039 5	1 297	1 577	43 000	0.98

3.2　燃油代偿损失与回风比例的函数关系

为研究回风比例的变化对燃油消耗的影响, 当空调总送风量为 0.36 kg/s 时, 由式(5)计算得到不同回风比例引起的燃油代偿损失。并对离散值进行多项式拟合, 得到 30%～80% 区间内所有回风比例对应的燃油代偿损失, 拟合曲线如图 12 所示。拟合方程为

$$m_{f,bl}(X)=-0.675\,875X+67.587\,5 \tag{6}$$

式中: $m_{f,bl}(X)$为回风比例为X时的燃油代偿损失。

由图 12 可知, 在回风比例为 30%～80% 区间内, 随着客舱回风比例的增加, 燃油代偿损失呈线性下降。当回风比例为 30% 时, 燃油代偿损失最大, 每小时为 47.31 kg; 当回风比例为 80% 时, 燃油代偿损失最小, 每小时为 13.52 kg。回风比例每增加 1%, 每小时燃油消耗减少 0.676 kg。

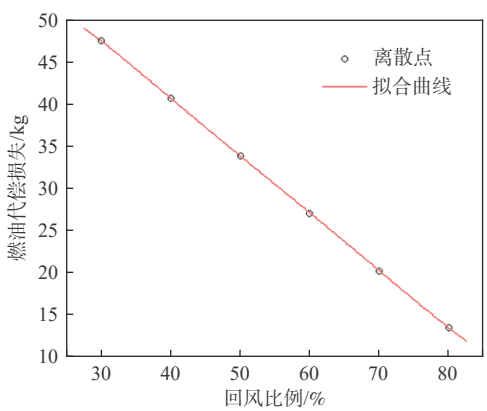


图 12　燃油代偿损失与回风比例函数关系

Fig. 12　Function relationship between fuel compensation loss and return air ratio

4　客舱空调最佳回风比例的确定

综合考虑回风比例变化对客舱内空气品质及飞机燃油消耗的影响, 从而确定客舱空调的最佳回风比例。基于乘客呼吸区通风效率和燃油代偿损失 2 个指标, 采用功效系数法构造客舱空调最佳回风比例评价函数[20]。

功效系数法根据多目标规划的原理, 首先, 分别确定 2 个评价指标子函数的函数类型, 乘客呼吸区通风效率在整个回风比例区间内为逆向指标, 其子函数属于极小型函数; 燃油代偿损失在整个回风比例区间内为正向指标, 其子函数属于极大型函数。其次, 计算 2 个评价指标子函数的功效系数。最后, 经过加权得到总功效系数$G(X)$。权重系数由改进的层次分析法[21]计算得出, 首先, 将 2 个评价指标按照重要程度进行排序, 由于本文更侧重客舱空调的节能性, 因此, 将燃油代偿损失指标记为u_1, 将乘客呼吸区通风效率指标记为u_2。然后, 对这 2 个指标的权重系数比r_1进行定量赋值, r_1的赋值可参考表 2[21]。最后, 通过计算得到 2 个评价指标的权重系数: λ_1=0.55, λ_2=0.45。构造出的评价函数为

$$G(X)=\lambda_1\left(\frac{m_{f,bl,max}-m_{f,bl}(X)}{m_{f,bl,max}-m_{f,bl,min}}\times 40+60\right)+\lambda_2\left(\frac{\varepsilon(X)-\varepsilon_{min}}{\varepsilon_{max}-\varepsilon_{min}}\times 40+60\right) \tag{7}$$

式中: ε_{max}、ε_{min}分别为乘客呼吸区通风效率的最大值、最小值; $m_{f,bl,max}$、$m_{f,bl,min}$分别为燃油代偿损失的最大值、最小值。

表 2　r_1赋值参考

Table 2　Reference values of r_1

r_1	说明
1.0	指标u_1与指标u_2同样重要
1.2	指标u_1比指标u_2稍微重要
1.4	指标u_1比指标u_2明显重要
1.6	指标u_1比指标u_2强烈重要
1.8	指标u_1比指标u_2极端重要

通过式(3)和式(6)计算得到回风比例为30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%时的乘客呼吸区通风效率和燃油代偿损失，代入式(7)得到对应的总功效系数。再将离散的总功效系数进行二阶高斯曲线拟合，如图13所示。拟合得到的函数关系式为

$$G(X)=50.6\exp\left(-\left(\frac{X-91.83}{68.21}\right)^2\right)+61.33\exp\left(-\left(\frac{X+2.091}{104}\right)^2\right) \tag{8}$$

采用式(8)求得该函数最大值对应的回风比例X=64.864%，即为在评价函数$G(X)$中满足客舱空气品质及空调节能指标的最佳回风比例。

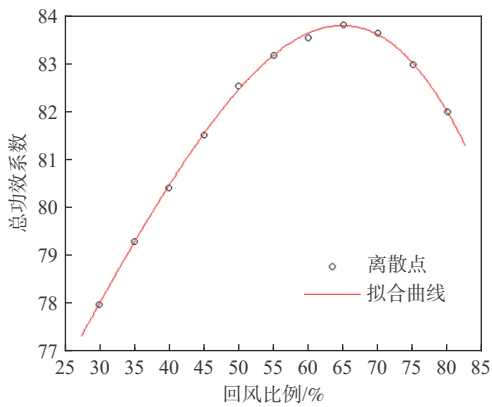


图13 总功效系数与回风比例函数关系

Fig. 13 Function relationship between total efficiency coefficient and return air ratio

5 结 论

1）在相同送风量的工况下，客舱内CO_2体积分数随回风比例增大而升高，乘客呼吸区通风效率随回风比例增大而逐渐降低，回风比例越大，客舱内通风排污效果越差。

2）在空调总送风量为0.36 kg/s的工况下，燃油代偿损失随客舱回风比例增加呈线性下降，且回风比例每增加1%，每小时燃油消耗减少0.676 kg。

3）基于乘客呼吸区通风效率和燃油代偿损失2个指标，构建总功效系数$G(X)$作为评价函数，得出客舱空调的最佳回风比例为64.864%。本文方法能够为客舱空调的回风比例控制提供依据。

参考文献（References）

[1] YANG H, ZHANG X J, WANG C, et al. Experimental and theoretical study on a novel energy-saving ECS for commercial airliners[J]. Applied Thermal Engineering, 2017, 127: 1372-1381.

[2] CHEN L, ZHANG X J, WANG C, et al. A novel environmental control system facilitating humidification for commercial aircraft[J]. Building and Environment, 2017, 126: 34-41.

[3] HERBIG B, NORREFELDT V, MAYER F, et al. Effects of increased recirculation air rate and aircraft cabin occupancy on passengers' health and well-being–Results from a randomized controlled trial[J]. Environmental Research, 2023, 216: 114770.

[4] 张英杰. 狭小空间内温度和二氧化碳对人体舒适性影响实验研究[D]. 重庆: 重庆大学, 2018: 1-5.
ZHANG Y J. The experimental research into the influence of temperature and carbon dioxide to human comfort in small space[D]. Chongqing: Chongqing University, 2018: 1-5(in Chinese).

[5] 高经诚, 陈维建, 胡万俊, 等. 民用飞机客舱CO_2分布特性分析[J]. 北京航空航天大学学报, 2023, 49(9): 2510-2517.
GAO J C, CHEN W J, HU W J, et al. Analysis of CO_2 distribution characteristics in cabin of civil aircraft[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(9): 2510-2517(in Chinese).

[6] 陈希远, 于天鹏, 欧阳建亮, 等. 考虑再循环风的飞机座舱引气污染仿真研究[J]. 科学技术与工程, 2017, 17(19): 295-300.
CHEN X Y, YU T P, OUYANG J L, et al. A study of bleed air contaminant in aircraft based on recirculation air[J]. Science Technology and Engineering, 2017, 17(19): 295-300(in Chinese).

[7] 杨建忠, 马博文, 陈希远, 等. 送风形式对飞机座舱引气污染物扩散影响[J]. 交通运输工程学报, 2019, 19(1): 108-118.
YANG J Z, MA B W, CHEN X Y, et al. Influence of air supply form on contaminat diffusion of bleed air in aircraft cabin[J]. Journal of Traffic and Transportation Engineering, 2019, 19(1): 108-118(in Chinese).

[8] 刘毓迪, 孙学德, 张存, 等. 民用飞机个人通风送风温度对人体舒适性的影响[J]. 航空学报, 2019, 40(2): 522363.
LIU Y D, SUN X D, ZHANG C, et al. Impact of individual ventilation temperature on passenger comfort of commercial aircraft[J]. Acta Aeronautica et Astronautica Sinica, 2019, 40(2): 522363(in Chinese).

[9] 林家泉, 戴仕卿. 基于排污效率和吹风感指数的客舱空调最佳送风方式[J]. 航空学报, 2022, 43(7): 93-102.
LIN J Q, DAI S Q. Optimal air supply mode of aircraft cabin based on removal effectiveness and draft rating[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(7): 93-102(in Chinese).

[10] 周立祥, 林家泉. 一种清除引气污染物的空调梯形信号送风方法[J]. 航空学报, 2023, 44(10): 77-86.
ZHOU L X, LIN J Q. Trapezoidal signal air supply method for air conditioning to remove bleed air pollutants[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(10): 77-86(in Chinese).

[11] 林家泉, 李弯弯, 王瑞婷, 等. 基于飞机客舱空气品质的桥载空调送风优化[J]. 北京航空航天大学学报, 2017, 43(11): 2259-2265.
LIN J Q, LI W W, WANG R T, et al. Optimization of air supply for bridge load air conditioning based on aircraft cabin air quality[J]. Journal of Beijing University of Aeronautics and Astronautics, 2017, 43(11): 2259-2265(in Chinese).

[12] ZHANG X, LIU J J, LIU X, et al. Performance optimization of airliner cabin air filters[J]. Building and Environment, 2021, 187: 107392.

[13] 赵澎渤. 四轮高压除水空气循环制冷系统优化与分析[D]. 南京: 南京航空航天大学, 2012: 34-39.

ZHAO P B. Optimization and analysis for four-wheel high pressure water removal air cycle refrigeration system[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2012: 34-39(in Chinese).

[14] 林家泉, 李弯弯. 空调送风速度对客舱环境影响的模拟及优化[J]. 液压与气动, 2016, 40(10): 12-17.
LIN J Q, LI W W. Simulation and optimization for the effect of air supply velocity on cabin environment[J]. Chinese Hydraulics & Pneumatics, 2016, 40(10): 12-17(in Chinese).

[15] 黄衍, 段然, 李炳烨, 等. 飞机座舱个性送风下的气态污染物传播规律实例研究[J]. 应用力学学报, 2015, 32(4): 586-592.
HUANG Y, DUAN R, LI B Y, et al. Simulation of contaminant transportation in aircraft cabin with partly gaspers on[J]. Chinese Journal of Applied Mechanics, 2015, 32(4): 586-592(in Chinese).

[16] 赵维维, 娄德仓, 钟世林. 基于烟分析和代偿损失的热管理系统性能评价方法[J]. 航空动力学报, 2023, 38(12): 2829-2836.
ZHAO W W, LOU D C, ZHONG S L. Performance evaluation method of thermal management system based on exergy analysis and compensatory loss[J]. Journal of Aerospace Power, 2023, 38(12): 2829-2836(in Chinese).

[17] 寿荣中, 何慧姗. 飞行器环境控制[M]. 北京: 北京航空航天大学出版社, 2004: 253-257.
SHOU R Z, HE H S. Spacecraft optimal control theory and method[M]. Beijing: Beihang University Press, 2004: 253-257(in Chinese).

[18] TURAN O. Exergo-economic analysis of a CFM56-7B turbofan engine[J]. Energy, 2022, 259: 124936.

[19] KORBA P, BALLI O, CALISKAN H, et al. Energy, exergy, economic, environmental, and sustainability assessments of the CFM56-3 series turbofan engine used in the aviation sector[J]. Energy, 2023, 269: 126765.

[20] 刘隽楷, 邹晓松, 袁旭峰, 等. 含功效系数法的电力系统多目标综合优化[J]. 电测与仪表, 2018, 55(24): 59-63.
LIU J K, ZOU X S, YUAN X F, et al. Multi-objective comprehensive optimization of power system containing the efficacy coefficient method[J]. Electrical Measurement & Instrumentation, 2018, 55(24): 59-63(in Chinese).

[21] 李银霞, 袁修干, 杨春信, 等. 歼击机座舱工效学评价指标权重系数的确定[J]. 航空学报, 2006, 27(3): 370-373.
LI Y X, YUAN X G, YANG C X, et al. Weight factors determination of the evaluation indexes of cockpit ergonomics[J]. Acta Aeronautica et Astronautica Sinica, 2006, 27(3): 370-373(in Chinese).

Optimization of cabin return air ratio based on air quality and compensation loss

GAO Yang，LIN Jiaquan*

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

Abstract: In order to obtain better air quality and improve the energy saving level of cabin air conditioning, computational fluid dynamics (CFD) method was used to establish a simulation model of the economy class of Boeing 737 passenger aircraft, and particle image velocimetry (PIV) technology was used to verify the accuracy of the simulation model. Based on this model, the effects of different return air ratios on the CO_2 concentration field in cabin air conditioning were studied, and the ventilation efficiency index of the passenger breathing area was used to evaluate the effects of different return air ratios on cabin air quality. The fuel compensation loss under various return air proportion conditions with the same air supply volume was also calculated using the total takeoff mass method, and a functional relationship between the fuel compensation loss, the ventilation efficiency of the passenger breathing area, and the return air proportion was fitted. The efficiency coefficient approach was used to design the evaluation function, and the ideal return air proportion for the cabin air conditioning system was 64.864%. This method can provide a basis for the proportional control of return air in cabin air conditioning.

Keywords: aircraft cabin; return air ratio; air quality; compensation loss; evaluation function; ventilation efficiency

Received：2024-04-10；**Accepted**：2024-06-07；**Published Online**：2024-07-25 11：47

URL：link.cnki.net/urlid/11.2625.V.20240725.0910.001

Foundation item：Special Program for Civil Airplane of the Ministry of Industry and Information Technology (2020020306)

* **Corresponding author**. E-mail：jqlin@cauc.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0209

提高存算一体化系统性能的静态计算划分方法

薛洪宇[1]，徐晟[1,2]，罗乐[1]，闫亮[3,4]，邹兴奇[3,*]

(1. 安徽师范大学 计算机与信息学院，芜湖 241000；2. 合肥综合性国家科学中心 人工智能研究院，合肥 230094；
3. 中国科学院 计算技术研究所，北京 100083；4. 中国科学院大学，北京 100049)

摘 要：存算一体化 (PIM) 系统通过引入存内计算单元改变冯·诺依曼结构“存算分离”模式，可有效缓解存储墙问题，但 PIM 系统与现有软件体系不匹配，性能和能效受程序的计算划分限制，甚至可能出现负优化。针对该问题，提出一种 PIM 系统中的静态计算划分方法。该方法将 PIM 系统中程序的执行过程抽象为基于注释调用图 (ACG) 的表征模型，从而将计算划分问题转化为 ACG 的最小割问题，并提出一种基于模拟退火的启发式计算划分求解算法。实验结果表明：相较于传统方法，所提方法的计算划分结果可平均提升 39% 系统性能，降低 32% 系统能耗。

关 键 词：存算一体化；计算划分；最小割；模拟退火；异构系统
中图分类号：TP303
文献标志码：A **文章编号**：1001-5965(2026)06-2042-12

随着大数据、深度学习、人工智能等新兴领域的蓬勃发展，应用的数据量不断增大，信息处理逐渐由传统的“计算密集型”向“访存密集型”转移。近十年来，随着硅基电子器件逼近物理极限，摩尔定律濒临失效，中央处理器 (central processing unit, CPU) 的性能难以高速、持续地提升。同时，传统存储器件无法兼具“高速度”和“高密度”的特点，导致现有以计算为核心的冯·诺依曼结构严重依赖层次化存储结构获得系统性能与存储容量的折中。而大数据时代下离散化、随机化的数据访问使层次化存储的效率断崖式下降，甚至濒临失效，数据频繁在 CPU 与多级存储间搬运，产生带宽受限、计算效率下降等一系列体系结构瓶颈，引发存储墙问题[1]。以个人电脑常用的第四代双倍数据率(double data rate 4, DDR4)内存为例，其外部封装双列直插式内存模块(dual in-line memory module, DIMM)接口运行在 3 200 MHz 时的理论带宽为 25.6 GB/s，而其内部带宽高达 1 TB/s，约为外部带宽的 40 倍[2]。另有研究指出，一次 CPU 与内存间的数据搬运能耗约为计算同大小浮点数能耗的 100～1 000 倍[3]。传统冯·诺依曼结构计算机正遭受严峻挑战，难以满足智能社会对计算的需求。

为缓解存储墙问题，目前有 2 种主要的技术路线：①进一步优化层次化存储的结构和效率。例如，AMD EPYC 9684X 处理器的最后一级缓存 (last-level cache, LLC) 容量已经超过 1 GB[4]。如此庞大的缓存可在更大地址范围内发掘数据的局部性，从而隐藏访存延迟，提升系统性能。②彻底打破冯·诺依曼结构中“存算分离”的固有思想，从底层架构进行突破，对计算机体系结构进行重新设计。该技术路线被称为存算一体化 (processing-in-memory, PIM)[5]。

PIM 并非最新出现的技术，其发展可以追溯到 20 世纪 60 年代[6]。顾名思义，PIM 的主要思想是在存内集成计算资源，利用存内的超高带宽和低访问延迟，实现近数据计算。近年来，3D 堆叠[7]、硅通

收稿日期：2024-04-10；**录用日期**：2024-08-23；**网络出版时间**：2024-09-13 16:17
网络出版地址：link.cnki.net/urlid/11.2625.V.20240912.1433.003
基金项目：国家自然科学基金 (62102005,62104230)；安徽省自然科学基金 (2008085QF330,2108085QF265)；安徽高校协同创新项目 (GXXT-2021-011)；安徽师范大学研究项目 (751968)
* **通信作者**. E-mail：zouxingqi@ict.ac.cn

引用格式：薛洪宇，徐晟，罗乐，等. 提高存算一体化系统性能的静态计算划分方法 [J]. 北京航空航天大学学报，2026，52（6）：2042-2053. XUE H Y，XU S，LUO L，et al. Enhancing performance through static computing partitioning approach in processing-in-memory systems [J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（6）：2042-2053 （in Chinese）.

孔 (through-silicon vias, TSVs)[8]、芯粒[9] 等制造智能技术及大数据、神经网络等“杀手级”应用的出现, 给 PIM 的发展带来了新的契机。通过在存内集成高并发、高效能的计算单元, PIM 系统可以实现部分应用 1～2 个数量级的性能加速[10], 降低约 80%[11] 的数据搬运能耗。部分新型非易失存储器件甚至可以实现计算和存储的完全融合[12]。由于数据无需通过外部接口在内存与 CPU 间频繁搬运, 而是在存内以更大并发规模就地计算, 极大提升了性能和能效。因此, PIM 体系结构已经成为缓解甚至打破存储墙的可行手段之一。

尽管 PIM 存在诸多好处, 但也面临着一系列挑战, 其中, 首先要解决的就是与现有“存算分离”软件体系不匹配的问题。由于现有软件体系都是针对传统冯·诺依曼结构设计并优化的, 若要充分发挥存算一体系统的优势和潜力, 需要根据程序的特点, 筛选出适合在存内计算单元上执行的指令或代码段, 从而最大化系统性能和能效。该问题被称为 PIM 系统的计算划分[13] 问题。

PIM 系统中, 存内引入的计算单元能耗更低、能效更高。理想情况下, PIM 系统理应具备更好的性能和能效。实际上, 引入存内计算单元有时反而导致系统性能和能效的负优化。例如, 在 PIM 系统中执行蒙特卡罗算法可能导致系统性能降低 64%[14]; 执行 xalanc 程序的能效比至多下降为原来的一半[15]。此外, Ahn 等[5] 指出即使采用相同的程序和计算划分, 使用不同数据集也会导致系统性能在 −50%～+80% 之间变化。由此可见, PIM 系统的计算划分对系统的性能和效能有巨大影响, 如何得到程序的最优计算划分, 从而最大化获得性能和能效的提升, 成为亟待解决的问题。

目前, 已有一些 PIM 系统计算划分的相关研究, 但主要侧重解决存内多计算单元间的划分, 对 CPU 与内存间的计算划分考虑较少。例如, Prometheus[16] 基于社区发现算法, 提出一种最小化混合内存立方体 (hybrid memory cube, HMC) 间数据传输的计算划分算法; Li 等[17] 提出一种图计算应用的数据分配模型, 侧面决定了计算在多个 HMC 中的执行位置, 降低了 HMC 间的数据通信。这些方法通常假设全部的计算任务都在存内执行, CPU 几乎完全闲置, 浪费了宝贵的计算资源。

基于上述分析, 本文立足于 CPU 和存内计算单元共存的场景, 提出一种 PIM 系统中最大化系统性能的静态计算划分方法, 将存算一体程序的执行过程抽象为基于注释调用图 (annotated call graph, ACG)[18] 的表征模型, 并将计算的划分转化为 ACG 的最小割问题, 从而实现可量化的计算划分。该方法将存内的计算单元视为一个整体, 综合考虑 CPU 和内存两者间的计算划分, 发挥各自优势, 从而最大化存算一体系统综合性能。此外, 本文方法还可以与现有存内多计算单元间的计算划分研究进行补充, 从而进一步提升系统性能, 降低不必要数据搬运带来的能耗和延迟开销。本文贡献如下:

1) 提出最大化 PIM 系统性能的静态计算划分方法, 针对现有划分方法对 CPU 和存内计算单元共存场景考虑较少的问题, 提出在两者间的最佳计算划分方法。

2) 给出 PIM 程序执行过程的表征模型。通过对程序的离线分析, 生成程序的 ACG, 最终将 PIM 系统中的计算划分问题转化为 ACG 的最小割问题。

3) 提出一种基于模拟退火的启发式计算划分求解算法, 赋予执行次数多的基本块更高的权重, 极大程度降低了解空间大小, 加快了算法收敛最优解的速度。

1　相关研究

计算划分是 PIM 系统中的重要环节[5], 根据划分粒度不同, 有时也被称为指令卸载[19] 或计算的卸载[20]。计算划分结果决定了程序每条指令的具体执行位置, 不同的执行位置会导致完全不同的访存和数据搬运行为, 从而极大程度影响系统性能和能耗。极端情况下, 不合理的计算划分可能会导致数据“乒乓”现象[21-22]。数据频繁在 CPU 与存内计算单元间搬运, 导致性能和能耗的负优化。

计算划分并不是 PIM 系统中独有的, 在多核系统和异构系统中已有较多研究。但 PIM 系统中存内计算单元与 CPU 间数据同步的带宽更低、延迟更大, 且存在数据同步消息对内存带宽的潜在占用问题, 因此, 需要新的针对性方法和理论。从近年来发表的文献看, 国内外对 PIM 系统中的计算划分研究尚处于起步阶段, 研究思路大致可分为 3 类[13]: 静态计算划分、动态计算划分和间接导致的计算划分。

1.1　静态计算划分

静态计算划分指在程序编写或编译时就确定了存内执行的代码片段(或指令)。被指定的代码片段或指令在存内计算单元中执行, 未被指定的部分仍由 CPU 完成运算。常见的静态计算划分方法有 3 类。

第 1 类是由编程人员基于对存算一体系统和程序特性的专业理解, 主动对需要在存内执行的代码片段进行手动标记。这种划分方法适用于存内计算单元功能单一、结构简单的情况。例如,

忆阻器天然具备逻辑蕴含和模态乘加的计算优势，因此，在 PipeLayer[23] 的研究中，将神经网络中的全部乘加操作通过代码标记的方式在存内执行；文献 [24] 在动态随机存取存储器（dynamic random access memory, DRAM）中设计了异或电路，将程序中全部的异或运算划分至存内执行。使用类似方法的研究还有文献 [25-26]。

第 2 类静态计算划分通过编译器或设计特定的编程模型完成。相比代码标记，编译器和编程模型可以提供更多软件层面的信息，从而带来更好的划分结果。例如，文献 [27] 通过建立代码分析结构模型，筛选出适合存内执行的代码片段，并通过编译器完成对代码的自动标记。Ahmed 等[28] 通过对向量访问的数据重用距离分析，给出了一种自动选择存内执行代码段的编译器方案。该类划分方法的优点是无需程序员参与，计算划分由编译器自动完成。但编译器只能从代码层面进行分析，无法感知程序的实际执行情况。

第 3 类静态计算划分基于程序的离线分析，也是本文使用的计算划分方法。该类方法需要使用指令插装工具对目标程序的执行过程进行信息抓取与建模，分析出理论上最优的计算划分。例如，Prometheus[16] 通过对程序的离线分析，建立了一种多层网络结构，从而实现代码片段在 HMC 内部计算单元间的计算划分；Jin 等[29] 通过离线分析，完成了在 GPU 与存内计算单元间的计算划分。这类方法虽然可以获得理论上最优的计算划分，但需要大量的时间进行程序的离线分析。

1.2 动态计算划分

动态计算划分指存算一体程序在执行过程中，基于当前系统的执行状态，动态决定当前指令是否在存内计算单元中执行。动态计算划分往往在静态计算划分的基础上完成。先通过静态计算划分确定在存内执行的候选代码片段，实际执行到相应位置时，再基于启发式算法动态决定。因此，并非全部被标记的指令都会在存内执行，未在存内执行的指令仍在 CPU 中执行。由于动态执行的特性，这类计算划分通常以指令作为划分粒度。例如，Ahn 等[5] 首次提出使用数据的局部性作为动态划分的准则；文献 [30] 使用温度动态决定指令是否分发等。动态计算划分往往通用性更强，但需要额外的硬件代价以对系统运行状态进行监控。

1.3 间接导致的计算划分

PIM 系统中，数据和计算是紧密相连的。计算的划分可以决定数据的流向和访问模式，反过来，数据的分布和组织结构也可以间接决定计算的位置和效率。一些存算一体设计通过设计高效率的数据结构和优化数据存放位置，实现数据的就近计算，也可以有效提升系统效率，降低数据搬运的关键路径距离。这类设计可以间接导致计算的划分。例如，Pattnaik 等[31] 通过对 GPU 内存利用率的分析，反推出 GPU 中高性能核应远离内存、简单核应与内存更近；Hsieh 等[32] 通过预测即将访问的系统页，并将对应页划分至较近的存内计算单元计算，从而提升存算一体系统性能；文献 [33] 基于 3D 堆叠内存中 TSVs 流量分析，优化计算位置，减少冗余存内流量。

总体来看，现有 PIM 计算划分的研究中，不论是静态计算划分还是动态计算划分，其研究内容多偏向于只考虑内存（或显存）模块内部多个计算单元间的计算划分，未综合考虑 CPU 和存内计算单元共存时最大化系统性能的计算划分，这也是本文解决的重点。

2 计算划分方法的实现

2.1 存算一体程序的执行模型

为在 PIM 系统中对 CPU 和存内计算单元间进行可量化的计算划分，需要先对其执行过程进行抽象，提取存算一体程序的执行模型；在执行模型的基础上进行计算划分。本文的主要思路是将存算一体程序的执行过程转化为程序的 ACG，再通过对 ACG 的图划分，实现存算一体程序的计算划分。

ACG 是一种有向无环图 (directed acyclic graph, DAG)，可表示程序在执行过程中不同子过程的执行与依赖关系[18]。注释调用图 $G=\{V, E, P\}$ 的本质是一种程序依赖图 (program dependence graph, PDG)[34]，由 3 个部分组成：节点集 V、边集 E 和属性集 P。其中，每个节点 $v \in V$ 代表程序的一个子过程或子函数。ACG 中的每一条边 $e=(u, v) \in E$，代表节点 u 和节点 v 间的调用关系。ACG 中的节点和边还拥有一些属性 $p \in P$。通过在 ACG 上利用节点和边的属性对节点进行划分，可以将计算划分问题转化为最小割问题，有效实现在 CPU 和存内计算单元间的计算划分。

为获得存算一体程序的 ACG，需要先对程序进行动态分析。高级程序语言对数据的封装程度较高，难以精确展示程序的访存行为，因此，需先将源程序转化为更贴近底层硬件的汇编语言。基于该考虑，本文使用底层虚拟机（low level virtual machine, LLVM）中间表示 (intermediate representation, IR) 作为中间语言进行程序访存的行为级表示。LLVM

在编译 IR 的过程中, 会将整个程序划分为若干个基本块。这些基本块可作为 ACG 中的节点; 基本块间的调用与依赖关系可以构建 ACG 中的有向边; 基本块本身的执行信息可以作为 ACG 中节点的属性。

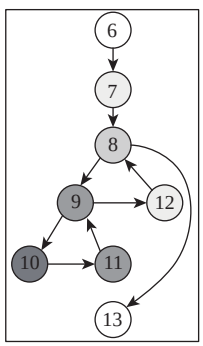
基于上述分析, 本文使用 Contech[35] 作为前端工具生成存算一体程序的 ACG, 流程如图 1 所示。Contech 是一种基于 LLVM 编译器的指令插装和动态分析工具, 本文修改了 Contech, 使其可以支持 ACG 的生成。以图 1(a) 中 PageRank(PR)[5,19] 算法标注的阴影部分为例(算法第 6～12 行), Contech 首先利用 LLVM 编译生成源代码的 IR, 如图 1(b) 所示。在图 1 中, 本文使用不同颜色对应不同的基本块, 并依据基本块与图 1(a) 中行编号的对应关系, 对 ACG 中的节点进行编号。完成 IR 的生成后, 使用 Contech 中提供的动态控制流图(dynamic control flow graph, dynamicCFG)后端工具生成程序的执行流程图, 如图 1(c) 所示。

```
1     parallel:   _for (v: graph.vertices) {
2:          v.pagerank = 1.0 / graph.num_vertices;
3:          v.next_pagerank = 0.15 / graph.num_vertices;
4:    }
5:    count = 0;
6:    do  {
7:          parallel_for (v : graph.vertices) {
8:              delta = 0.85 * v. pagerank / v.out_ degree;
9:              for (w : v.successors) (
10 :                    w.next_pagerank += delta;
11 :             }
12 :        }
13 :        diff = 0.0;
14 :        parallel_for (v : graph.vertices) {
15 :            diff += abs (v.next_ pagerank -v.pagerank) ;
16 :            v.pagerank = v. next_pagerank;
17 :            v.next_pagerank = 0.15 /
graph.num_vertices ;
18 :        }
19 :   }while (++count < max_iteration  &&  diff > e) ;
```

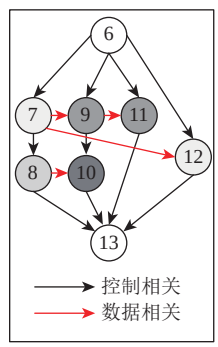
(a) PR 程序的源代码示例

```
; <label>:24:                                    ; preds = %67, %23
%25 = load i64, i64* %4, align 8
%26 = icmp ult i64 %25, 3072627
br i1 %26, label %27, label %70
; <label>:27:                                    ; preds = %24
%28 = load i64, i64* %4, align 8
%29 = load %struct.Node*, %struct.Node** @nodes, align 8
……
br label %43
; <label>:43:                                    ; preds = %51, %27
%44 = load i64, i64* %6, align 8
……
%50 = icmp ult i64 %44, %49
br i1 %50, label %51, label %67
; <label>:51:                                    ; preds = %43
%52 = load i64, i64* %6, align 8
%53 = load i64, i64* %4, align 8
%54 = load %struct.Node*, %struct.Node** , align 8
……
br label %43
; <label>:67:                                    ; preds = %43
%68 = load i64, i64* %4, align 8
%69 = add i64 %68, 1  store i64 %69, i64* %4, align 8
br label %24;
```

(b) PR 程序的中间表示



(c) 执行流程图



(d) 注释调用图

图 1　由源代码生成注释调用图的流程

Fig. 1　The working-flow of ACG generation from the source code

得到程序的执行流程图后, 采用文献 [18] 中方法生成程序的 ACG(图 1(d)), 具体生成过程如下:

输入: 程序的执行流程图 $g=\{V, E'\}$、属性集 P。

输出: 程序的注释调用图 $G=\{V, E, P\}$。

1. $g \leftarrow$ addStartandEndNode(g)
2. $g' \leftarrow$ reverse(g)
3. $T_{PDT} \leftarrow$ LengauerTarjan(g')
4. $\{S\} \leftarrow$ getEdges(g)
5. for each $e = (u, v)$ in $\{S\}$ do
6. if isNotAncestor(u, v, T_{PDT}) do
7. $L \leftarrow$ leastCommonAncestor(u, v, T_{PDT})
8. $\{p\} \leftarrow$ getPath(L, v,T_{PDT})
9. if $L = u$ do
10. remove(v, $\{p\}$)
11. end if
12. for each $n \in \{p\}$ do
13. addEdge(n, L, E)
14. end for each
15. end if
16. end for each
17. $G \leftarrow$ addDateDep(G)

由执行流程图生成 ACG 的过程, 主要是去除了执行流程图中的环, 其总体思路为: 先基于执行流程图生成程序的后序支配树 (post-dominator tree, PDT), 再基于 PDT 中的支配关系, 去除非支配边, 保留下来的就是 ACG。

1）PDT 的生成与控制依赖的建立。使用 Lengauer-Tarjan 算法[36-37]生成执行流程图的 PDT（算法中使用 T_{PDT} 表示）。首先，需使用 addStartAndEndNode(·) 函数给执行流程图添加起始和终结节点（算法第 1 行），以确保首尾节点不会因为算法的执行与其他节点断开。然后，通过 reverse(·) 函数将执行流程图中全部边反向（算法第 2 行），并使用 Lengauer-Tarjan 算法去除程序执行流程图中的非控制相关边，完成去环操作（算法第 3 行）。具体做法为：对程序执行流程图中全部基本块 v，使用 getEdges(·) 函数查找 PDT 中指向 v 的边 $e=(u, v)$。如果基本块 u 在 PDT 中是 v 的祖先，则两者不满足控制相关。为保留执行流程图中控制相关，需找到 PDT 中和 v 不满足祖先关系的基本块 u（算法第 4～6 行），使用 leastCommonAncestor(·) 函数查找两者在 PDT 中的最近共同祖先 L，并通过 getPath(·) 函数获得 PDT 中从 L 到 v 路径中的全部节点集合 $\{p\}$（算法第 7～8 行）。一旦执行流程图中存在环，则 PDT 中基本块 u 和 v 的最近共同祖先 L 就是 u 本身。利用这一特性，只需将 v 从点集合 $\{p\}$ 中去除即可（算法第 9～11 行）。经过去环处理后，点集合 $\{p\}$ 中的任意点 $n\in\{p\}$ 均与节点 L 存在控制相关，由此可以在 ACG 中构建从节点 L 至节点 n 的有向边（算法 12～14 行），至此，完成了 ACG 中控制依赖边的构建。

2）数据依赖的建立和基本块属性的获取。数据依赖边可以直接根据源代码中的变量依赖关系使用 addDateDep(·) 函数在 ACG 中添加得到（算法第 17 行）。随后，利用 Contech 中集成的 Pin[38] 工具对每个基本块进行指令插装。通过执行插装后的程序，可以获取每个基本块的执行时间、访存请求类型、访问数据量、访存地址等动态执行信息。将这些信息附加在 ACG 中的节点和边上，就完成了 PIM 程序 ACG 的生成过程。

2.2 计算划分算法的实现

在 ACG 的基础上，提出一种基于最小割的计算划分方法，其流程与结果如图 2 所示。其中，图 2(a) 定义了存算一体程序的控制流，并明确基本块执行边界，图 2(b) 在此基础上引入 CPU 与 PIM 及基本块之间的数据依赖，通过多种划分方案，直观展示不同策略对性能和能效的影响。本节以该段代码为例，介绍本文计算划分方法。

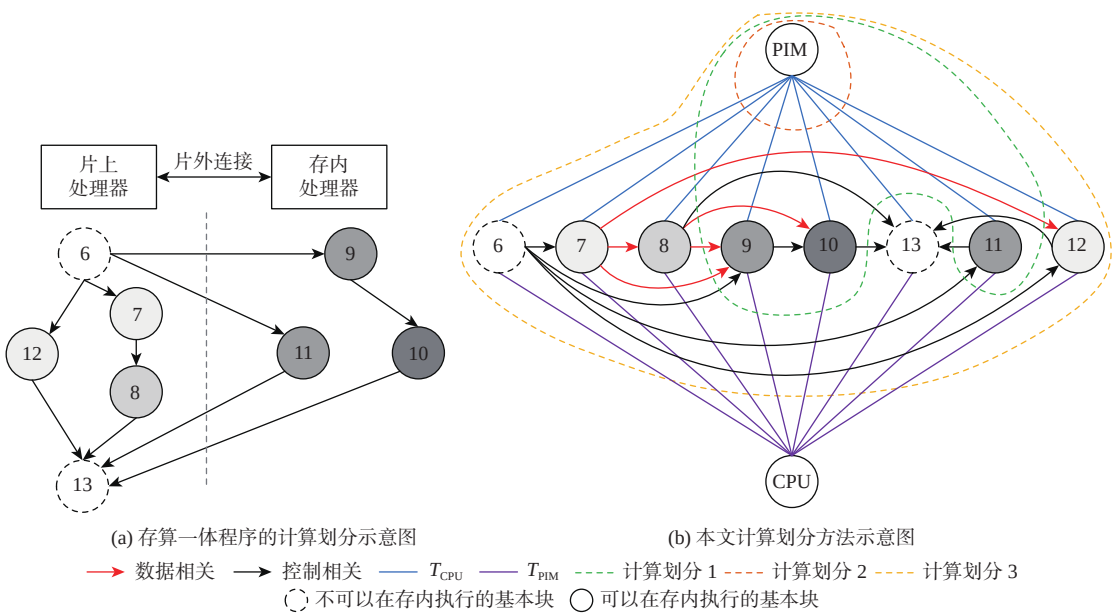


图 2 基于注释调用图的计算划分方法和对应的存算一体程序计算划分示意图

Fig. 2 ACG-based computing partitioning approach and corresponding compute-in-memory program computing partitioning diagram

为直观展示本文计算划分方法的原理，引入 2 个伪节点——CPU 和 PIM，并从两者出发分别与 ACG 中其他每个节点引出连线，如图 2(b) 中蓝色与紫色连线所示。由此，PIM 系统的计算划分问题可以转化为对图 2(b) 所示 ACG 的分割问题，即将添加过伪节点后的注释调用图 G 分割为 2 个点集 $\{S, T\}$，CPU$\in S$ 且 PIM$\in T$。点集 T 中点对应的基本块在存内计算单元上执行，点集 S 中对应的基本块在 CPU 上执行。图 2(b) 给出了 3 种不同的计算划分示例。其中，计算划分 2（橙色虚线圈内）中只包含了 PIM 伪节点，表示全部基本块都在 CPU 上执行（因为全部基本块对应点均在点集 S 中）；计算划分 3（黄色虚线圈内）中除伪节点 PIM 外，包含了全部基本块，则代表全部基本块都在存内计算单元上执行。

计算划分 2 和计算划分 3 通常只会在比较极端的情况下出现。大多数情况下，计算划分的结果应

该是一部分基本块在存内执行、一部分在 CPU 中执行, 与图 2(b) 中计算划分 1 类似。计算划分 1 中除伪节点 PIM 之外, 包含 9、10、11 这 3 个基本块, 表示这 3 个基本块对应代码段在存内计算, 其余部分在 CPU 中执行, 其对应的计算划分结果如图 2(a) 所示。

由此, 一旦确定注释调用图 $G=(V, E, P)$ 的分割 $V=\{S, T\}$, 存算一体程序的执行总时间 T_{all} 计算如下:

$$T_{\text{all}} = T_{\text{CPU}} + T_{\text{PIM}} + T_{\text{M}} \tag{1}$$

存算一体程序的执行总时间 T_{all} 由 3 部分组成: 程序在 CPU 上的执行总时间 T_{CPU}、在存内计算单元的执行总时间 T_{PIM} 和在两者间数据传输的总时间 T_{M}。其中, T_{CPU} 可由点集 S 中全部基本块 $u\in S$ 在 CPU 上的执行时间求和表示; 类似的, T_{PIM} 可由点集 T 中全部基本块 $v\in T$ 在存内计算单元上执行的总时间表示。$S_{u,v}$ 为基本块 u 和 v 之间传输的数据大小, B_{W} 为内存的读写带宽, 则数据块在 CPU 与存内计算单元间的传输时间可由 $S_{u,v}/B_{\text{W}}$ 计算。基于上述分析, 存算一体程序的执行总时间 T_{all} 可变换为

$$T_{\text{all}} = \sum_{u\in S} T_u^{\text{CPU}} + \sum_{v\in T} T_v^{\text{PIM}} + \sum_{\substack{(u,v)\in E\\ u\in S, v\in T}} \frac{S_{u,v}}{B_{\text{W}}} \tag{2}$$

ACG 的分割过程中, 割线有可能与 4 种类型的边相交: 数据相关边(红色)、控制相关边(黑色)、从 PIM 伪节点出发的边(蓝色)和从 CPU 伪节点出发的边(紫色)。若假设存在如表 1 所示的边和属性的对应关系, 则可将 ACG 的分割问题转化为图的最小割问题, 即如何确定 ACG 的分割结果 $V=\{S, T\}$, 使割线相交的边对应的属性和最小。

表 1　注释调用图中边的属性说明

Table 1　Description of the edge properties in the ACG

ACG中边的类型	对应属性	说明
数据相关边 $e=(u, v)$	$\frac{S_{u,v}}{B_{\text{W}}}$	基本块u向基本块v传输数据的时间
控制相关边		控制相关决定了基本块间的先行后续关系
从PIM伪节点出发的边 $e=(v_{\text{PIM}}, v)$	T_v^{PIM}	基本块v在存内计算单元中的执行时间, 无法在存内执行的基本块值为+∞
从CPU伪节点出发的边 $e=(v_{\text{CPU}}, v)$	T_v^{CPU}	基本块v在CPU上的执行时间

2.3　划分算法的求解

2.2 节中提出的目标问题可以等效为 0-1 整型规划问题。如式(2)所示, 优化目标为最小化存算一体程序的执行时间, 算法最终需确定每个 ACG 中的非伪节点(基本块)i 的执行位置。设 I_v 为二值变量, 表示基本块 v 的执行位置: 当 I_v 的值为 0 时, 表示基本块 v 对应的指令在 CPU 上执行; I_v 的值为 1 表示基本块 v 在存内计算单元上执行。由此, 存算一体程序的计算划分问题可转化为: 找到一个二值变量集合 $\{I_v\}(v\in V)$, 使式(2)中存算一体程序的执行总时间 T_{all} 最小。

引入变量 I_v 后, T_{CPU}、T_{PIM} 和 T_{M} 可由式 (3)～式 (5) 表示。此外, 为表征只能在 CPU 上执行、无法在存内计算单元执行的基本块(如程序的起始基本块和结束基本块受限于执行流程, 必须在 CPU 上执行), 算法给每个基本块引入二值变量 C_v, 用以限制 I_v 的取值, 如式 (6) 所示。当 C_v 的值为 1 时, 基本块 v 的执行位置由 I_v 的值决定; 否则, I_v 值只能为 0, 即只能在 CPU 上执行。

$$T_{\text{CPU}} = \sum_{u\in S} T_u^{\text{CPU}} = \sum_{v\in V} \left[(1-I_v)T_v^{\text{CPU}}\right] \tag{3}$$

$$T_{\text{PIM}} = \sum_{v\in T} T_v^{\text{PIM}} = \sum_{v\in V} (I_v T_v^{\text{PIM}}) \tag{4}$$

$$T_{\text{M}} = \sum_{\substack{(u,v)\in E\\ u\in S, v\in T}} \frac{S_{u,v}}{B_{\text{W}}} = \sum_{(u,v)\in E} \left(|I_u - I_v| \frac{S_{u,v}}{B_{\text{W}}} \right) \tag{5}$$

$$I_v \leqslant C_v \qquad \forall v \in V \tag{6}$$

综合式 (2)～式 (6), 可以求解一个给定存算一体程序的计算划分结果 $\{I_v\}(v\in V)$。根据每个基本块 v 对应的 I_v 值, 可以最终确定计算的划分。

实际情况中, 由于存算一体程序生成的 ACG 中基本块数量较多、解空间大, 其求解往往需要经过多次迭代, 才能够收敛最优解。因此, 本文提出一种基于模拟退火的启发式算法, 加速划分算法的求解和收敛速度。

输入: 注释调用图 $G=(V, E, P)$、模拟退火起始温度 T_{s}、结束温度 T_{e}、$\{C_v\}(v\in V)$、基本块执行次数阈值 T_{H}。

输出: 计算划分解集 $\{I_v\}^{\text{opt}}(v\in V)$。

1. $\{I_v\}^{\text{opt}} \leftarrow$ pickupBlockwithMostCount(G)
2. 模拟退火温度 $T' \leftarrow T_{\text{s}}$
3. while $T' > T_{\text{e}}$ do
4. $\quad \{U\} \leftarrow$ pickupCountExceed(V, T_{H}, $\{C_v\}$)
5. $\quad j \leftarrow$ randomPickup($\{U\}$)$\cdot i_{\text{d}}$
6. $\quad \{I_v\}^{\text{temp}} \leftarrow \{I_1, I_2, \cdots, I_{j-1}\}^{\text{opt}} \cup \{\text{not } (I_j)\}^{\text{opt}} \cup \{I_{j+1}, \cdots, I_{|V|}\}^{\text{opt}}$
7. $\quad \Delta P \leftarrow$ ExecutionTime($\{I_v\}^{\text{temp}}$) − ExecutionTime($\{I_v\}^{\text{opt}}$)
8. $\quad$ if $\Delta P < 0$ then
9. $\quad\quad \{I_v\}^{\text{opt}} \leftarrow \{I_v\}^{\text{temp}}$
10. $\quad$ else
11. $\quad\quad k \leftarrow$ ExecutionCount(j) / getMaxCount(G)
12. $\quad\quad$ if Exp($-k\times\Delta P/T'$) > Rand(0, 1) then

13. $\{I_v\}^{\text{opt}} \leftarrow \{I_v\}^{\text{temp}}$
14. end if
15. end if
16. $T' \leftarrow$ ReduceTemperature(T')
17. end while

首先，划分算法筛选出可在存内执行的全部基本块，并通过 pickupBlockwithMostCount(·) 函数在其中选择执行次数最高的基本块 u 的集合 $\{I_v\}^{\text{opt}}$，标记其在存内计算单元中执行（标记其 I_u 值为 1）。其余所有基本块都在 CPU 上执行。该划分结果将作为模拟退火的初始解。然后，算法根据式 (2)～式 (5) 计算按照该计算划分方案执行的总时间（算法第 1～2 行）。接着，算法根据模拟退火经典思路进行迭代求解：每次迭代过程中，先通过 pickupCountExceed(·) 函数筛选出执行次数占比超过阈值 T_{H} 且可以在存内执行的基本块集合 $\{U\}$，并通过 randomPickup(·) 函数在其中随机选择一个基本块获得其编号 j（算法第 4～5 行）。设置最低阈值避免了在存内计算单元上运行执行次数较少的基本块，因为研究表明这些基本块无法通过存内执行获得收益[39]。随后，算法在当前最优解 $\{I_v\}^{\text{opt}}$ 的基础上，取反以 j 为编号的基本块变量 I_j 形成新的可行解 $\{I_v\}^{\text{temp}}$（算法第 6 行）。通过 ExecutionTime(·) 函数计算出与当前最优计算划分对应的执行时间比较，得到时间差 ΔP。如果 ΔP 为负，说明新解 $\{I_v\}^{\text{temp}}$ 的执行时间优于当前最优解，算法会接受 $\{I_v\}^{\text{temp}}$ 作为当前最优解（算法第 7～9 行）。否则，算法会以一定概率接受新解（算法第 10～15 行）。该算法针对原有 Metropolis 准则[40]，引入修正系数 k，其值为以 j 为编号基本块的执行次数（使用 ExecutionCount(·) 计算）占所有基本块最大执行次数（使用 getMaxCount(·) 计算）的百分比。该设计使执行次数较多的基本块获得更高的更新权重（算法第 11 行），从而避免频繁尝试执行次数较少的基本块的情况。经过多次迭代，最终算法会得到收敛解 $\{I_v\}^{\text{opt}}$，即存算一体程序的性能最优计算划分。

3 实验评估

3.1 模拟平台配置

使用 PIMSim[41] 模拟器对本文方法的有效性进行验证。PIMSim 支持基于通用执行驱动多处理器（general execution-driven multiprocessor, GEM5）的全系统模拟，支持存算一体体系结构的指令级自定义、编译和执行。本文利用 PIMSim 中提供的 PIMKernel 存内执行模型，实现了基本块级的计算划分模拟，具体模拟环境配置如表 2 所示。

表 2 模拟存算一体系统配置参数

Table 2 Configuration parameters of simulated PIM system

处理单元类型	功能模块	参数配置
CPU	处理器	16核, 2 GHz频率, 乱序执行
	一级缓存	分离的指令与数据缓存, 各32 KB, 4路组相联, 块大小64 B, 修改、独占、共享、无效（modified, exclusive, shared, invalid, MESI）
	二级缓存	共享缓存, 2 MB, 8路组相联, 块大小64 B, MESI
存内处理器	内存模型	HMC V2.0, 4 GB, 256 Banks
	处理器	32个, 顺序执行, 2 GHz, 1个/vault
	一级缓存	分离的指令与数据缓存, 各32 KB, 4路组相联, 块大小64 B, MESI
	一致性协议	粗粒度锁[21-22], 锁粒度4 KB

3.2 基准设计

模拟 4 种不同的系统配置，用以验证本文计算划分方法的有效性：

1）仅 CPU。该配置不启用存内计算单元，不存在 CPU 和存内计算单元间的计算划分，全部运算由 CPU 完成。

2）仅存内计算。该配置启用存内计算单元，除程序中无法在存内执行的基本块外，其余全部基本块均在存内计算单元中执行。

3）原子操作。该配置同时启用 CPU 和存内计算单元，但仅将原子操作在存内计算单元中执行。这一计算划分方式与文献 [5] 类似，但并非基于局部性感知的动态划分，而是基于代码标记的静态划分。尽管 HMC 内存模块已经支持在存内执行部分原子操作[7]，但本配置不启用 HMC 的运算电路，而是用存内处理器完成计算，因此，本配置下可以完成的原子操作种类相较 HMC 更多。

4）本文方法。该配置使用本文提出的存算一体计算划分方法完成计算划分。

3.3 基准程序及数据集

使用 4 个经典图计算程序作为基准程序，计算核心代码来自文献 [42]，数据集来自斯坦福大学网络分析项目（Stanford network analysis project, SNAP）大型社交和信息网络数据集[43]。为体现计算划分的鲁棒性，使用 5 个不同大小的真实图数据集，如表 3 所示。

表 3 本文使用的基准程序和数据集

Table 3 Benchmarks and datasets used in this paper

基准程序	数据集
Average Teenage Follower (ATF)、Breadth-First Search (BFS)、Bellman Ford Shortest Path (SP)、PageRank (PR)	p2p-Gnutella30 (36 K点, 88 K边)、com-DBLP (317 K点, 1 M边)、com-Youtube (1.1 M点, 2.9 M边)、wiki-Talk (2.3 M点, 5 M边)、soc-LiveJournal1 (4.8 M点, 6.9 M边)

3.4　性能评估

本文计算划分方法充分考虑了 PIM 系统的执行特点, 因此, 可以最大化系统执行效率。为验证其划分效果, 本节评估计算划分所带来的性能改进, 如图 3 所示。为直观地展示结果, 图 3 中所有结果都根据 CPU-Only 进行了归一化。

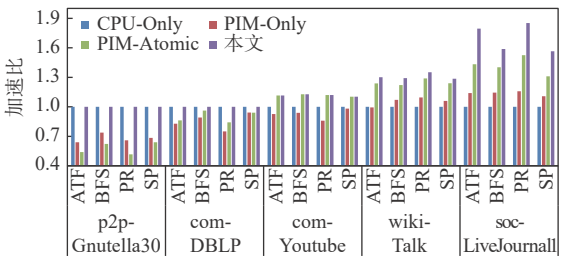


图 3　不同大小数据集和不同计算划分下的性能加速比

Fig. 3　Performance acceleration ratio under different dataset sizes and computational partitions

由图 3 可知, 在数据集较小时(如 p2p-Gnutella30 和 com-DBLP), 本文方法的性能与 CPU-Only 相同。这是由于 CPU 的缓存可以很好地发掘小数据集的数据局部性, 这种情况下, 仅使用 CPU 反而能防止不必要的数据搬运带来的延迟。随着数据集的增大(如 com-Youtube), 本文方法的划分结果与 PIM-Atomic 接近, 但相比 CPU-Only 和 PIM-Only, 性能仍分别提升 12% 和 20%。随着数据集进一步增大(如 soc-LiveJournal1), 相比 CPU-Only、PIM-Only 和 PIM-Atomic, 本文方法性能分别提升 70%、49% 和 20%, 平均提升 39%(最高提升至 1.85 倍)。上述实验数据说明, 本文方法可以有效计算出合理的计算划分, 从而提升存算一体系统性能。同时, 在数据集大小不同的情况下, 本文方法都可以获得相对最优的性能, 也证明了本文方法的鲁棒性。

3.5　指令的存内执行比例评估

为进一步探索计算划分的合理性, 评估存算一体程序在执行过程中存内执行的指令比例。由于数据集较小时, 本文方法的划分结果与 CPU-Only 和 PIM-Atomic 相同, 不具有比较性, 因此, 选取 wiki-Talk 和 soc-LiveJournal1 这 2 个数据集作为代表, 结果如图 4 所示, 其中, GM 为几何平均值。

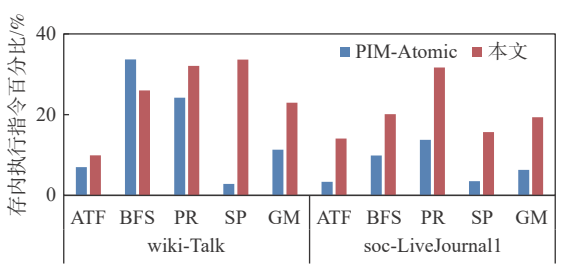


图 4　基准程序在存内执行指令百分比

Fig. 4　Percentage of instructions executed in-memory by benchmark programmes

由图 4 可见, 使用 2 种不同大小的数据集执行基准程序时, 本文方法的划分结果相比 PIM-Atomic, 在存内平均多执行了 13% 的指令。数据集较大时, 本文方法更倾向于在存内执行更多的指令。这是由于大数据集的离散化内存访问使数据局部性下降, 基本块间的数据传输开销进一步增大。因此, 算法倾向于使这类基本块划分到存内执行, 以避免数据的不必要传输。

在部分情况下(如在 wiki-Talk 上执行 BFS 算法时), 本文计算划分方法在存内执行的指令比例相比 PIM-Atomic 反而减少了 7.7%。这是由于本文使用的 BFS 基准程序以地址指针的形式手动维护了进栈出栈操作, 而未使用 C++模板提供的标准栈容器。这种实现方式原本目的是迎合 HMC 内存模块中的原子操作, 却增强了数据的局部性。这一变化被本文方法有效感知, 因此, 判断其不适合在存内执行情况, 从而出现存内执行比例下降的情况。

3.6　数据搬运能耗评估

为准确评估数据搬运导致的能耗变化, 仅对非处理器核能耗进行评估。图 5 展示了几种不同配置下的能耗分解图。其中, Caches 表示所有缓存, 包括 CPU 的缓存和存内处理器缓存的能耗, 本文使用 CACTI 6.0[44] 对缓存能耗进行模拟; DRAM 表示 HMC 中 DRAM 层的能耗, 由 PIMSim 中集成的 DRAMSim2 给出; Links 表示 HMC 中串行互联产生的能耗, 其数值根据 CoPIM[20] 中给出的数据计算得出, 大小为 13.7 pJ/bit。由图 5 可见, 不论使用哪种数据集, 本文设计都可以降低系统能耗。以最大的数据集 soc-LiveJournal1 为例, 相比 CPU-Only, 本文设计可以减少 41% 左右的系统能耗; 相比 PIM-Only 和 PIM-Atomic, 可降低 37% 和 18% 的系统能耗, 平均降低了 32%。

尽管本文方法的限制条件和优化目标与能耗无关, 但对能耗的优化也隐含其中。究其原因, 是

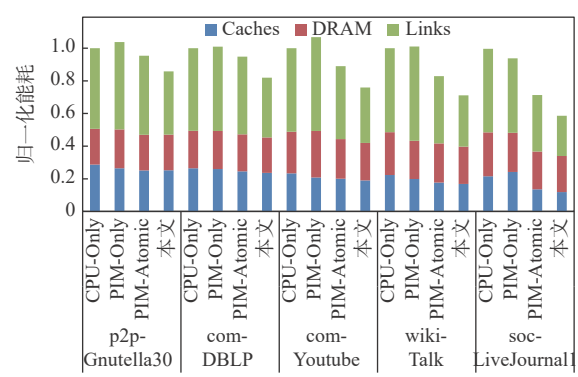


图 5　不同基准设计的能耗分解图

Fig. 5　Energy consumption breakdown for different benchmark schemes

因为 CPU 与内存间的数据传输延迟较大，合理的计算划分会大大降低两者间的数据传输量，从而减少数据的搬运能耗。由图 5 可见，3 类能耗中减少最多的就是 Links，这也从侧面反映出本文方法可以有效减少数据在 CPU 与内存间的数据传输量，降低系统能耗。

3.7 存内计算单元性能敏感度分析

为验证存内计算单元性能对本文计算划分方法的影响，对存内处理器核的频率进行敏感度分析，结果如图 6 所示。可以看出，在各存内处理器核频率下，本文计算划分方法都可以获得最佳系统性能。当存内处理器核频率为 500 MHz 和 1 GHz 时，系统性能相较 2 GHz 平均分别降低 39% 和 26%。当存内处理器核频率较低 (500 MHz) 时，本文方法给出的计算划分与 PIM-Only 结果相同。频率为 1 GHz 时，本文方法给出的计算划分与 PIM-Atomic 相同。随着存内处理器频率进一步提高，性能逐渐趋近 CPU，此时，存内处理器的近数据优势愈发凸显。因此，本文计算划分方法便倾向于在存内执行更多的指令。

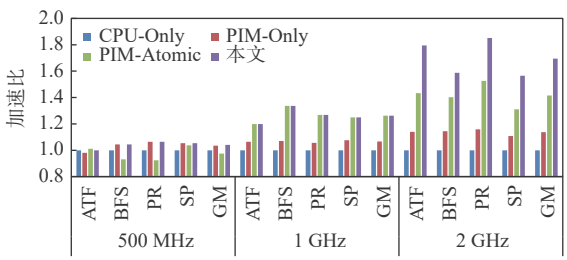


图 6 不同存内处理器核频率对系统性能的影响

Fig. 6 Impact of different in-memory processor core frequencies on system performance

3.8 局部性感知策略评估

本文计算划分方法是静态的，存内执行的代码段在源代码编写阶段已经确定。为量化静态划分和动态划分对系统性能的影响，对 PIM 系统进行局部性感知策略评估。实验借鉴文献 [5] 中使用数据局部性作为动态划分的参照，即当被标记的指令所需数据未被加载至 CPU 缓存时，对应指令才在存内执行。而当存内执行的是基本块而非单一指令时，则依据多数数据的位置确定执行位置。这种具备局部性感知的动态划分方法在图 7 中以“-L”结尾。动态划分和静态划分的计算划分范围相同，但实际在存内执行的指令比例受数据所在位置的影响并不相同。此外，由于实验是通过模拟完成的，在模拟器中可精确查找数据所在位置。现实中物理机无法以较低成本确定任意数据的存储位置，因此，本文实验中局部性感知方法的性能实际为使用局部性感知类方法的理论上限。

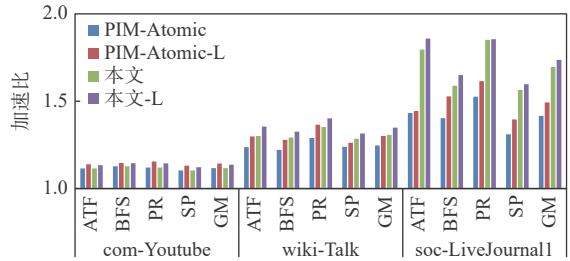


图 7 局部性感知对划分结果的影响示意图

Fig. 7 Schematic diagram of the impact of locality awareness on partitioning results

根据实验结果可知，当数据集较大时，动态划分 PIM-Atomic-L 和本文-L 对比静态划分性能，平均提升仅为 4.2% 和 2.5%。动态划分对 PIM-Atomic 的提升相对更多，但整体提升不大。当数据集较小时（如 com-Youtube），本文方法的动态划分本文-L 相比 PIM-Atomic-L 损失 0.6% 的性能，但整体区别不大，总体来看，本文-L 仍有 10% 左右的性能提升。

3.9 对非图处理应用的划分效果

为验证本文方法对非图处理应用的划分效果，选择 PrIM 基准测试集 [45] 中的 4 类应用（GEMV、Select、Unique、MLP）进行实验。4 类应用的输入数据均采用默认大小，归一化后的性能比较结果如图 8 所示，在存内执行的指令比例如表 4 所示。可见，4 类应用使用本文计算划分方法均获得了性能提示，平均加速比为 1.5 倍。这说明本文方法不仅适用于图处理应用，也适用于其他类型的应用。

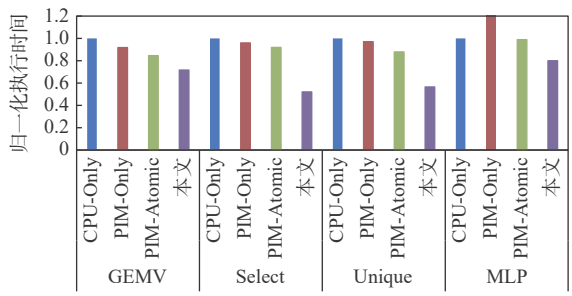


图 8 使用非图处理程序的性能加速比

Fig. 8 Performance acceleration ratio when using non-graph processing programmes

表 4 非图处理应用存内加载比例

Table 4 In-memory loading ratio for non-graph processing applications

基准程序	存内加载比例/%		
	PIM-Only	PIM-Atomic	本文方法
GEMV	100	15.44	13.75
Select	100	31.45	36.48
Unique	100	40.60	46.17
MLP	100	5.35	9.64

3.10　方法开销分析

本文计算划分方法的开销主要来源于 2 处: ①需要对程序进行离线分析, 包含程序 IR 的生成、基本块属性的获取、ACG 的生成等; ②需要使用基于模拟退火的启发式算法对计算划分进行求解。以本文实验使用的英特尔至强银牌 4 214 服务器为例 (2.2 GHz, 128 GB 内存), 对在 soc-LiveJournal1 上运行的 PR 程序进行计算划分, 程序进行离线分析的总时间约为 30.6 min。得到全部基本块信息并生成 ACG 后, 计算划分求解算法可以在 1 min 内完成执行。从离线分析到得到计算划分结果的总时间约为 31.6 min。

4　结　论

1) 本文方法可有效提升存算一体系统性能, 且具有一定的鲁棒性, 可在输入不同大小数据集或运行不同程序时均取得良好效果。在数据集较大时, 至多可取得 1.85 倍性能加速。

2) 本文提出的静态计算划分方法可与现有动态计算划分方法形成互补, 进一步提升存算一体系统的性能。实验数据表明, 运用局部性感知策略后, 可进一步提升 10% 左右的性能。

3) 本文提出的存算一体系统中的静态计算划分方法具有一定的通用性, 不仅适用于图处理应用的存算一体加速, 也适用于神经网络、数据库等领域的应用。

参考文献 (References)

[1] HUANG X H, LIU C S, JIANG Y G, et al. In-memory computing to break the memory wall[J]. Chinese Physics B, 2020, 29(7): 078504.

[2] LEE C, SHIN W, KIM D J, et al. NVDIMM-C: a byte-addressable non-volatile memory module for compatibility with standard DDR memory interfaces[C]//Proceedings of the 2020 IEEE International Symposium on High Performance Computer Architecture. Piscataway: IEEE Press, 2020: 502-514.

[3] PATTNAIK A P. Be aware of data movement: optimizing throughput processors for efficient computations[D]. State College: The Pennsylvania State University, 2019: 10-12.

[4] Advanced Micro Devices, Inc. AMD EPYC™ 9684X server processors[EB/OL]. (2023-01-16)[2024-03-12]. https://www.amd.com/zh-cn/products/processors/server/epyc/4th-generation-9004-and-8004-series/amd-epyc-9684x.html.

[5] AHN J, YOO S, MUTLU O, et al. PIM-enabled instructions: a low-overhead, locality-aware processing-in-memory architecture[C]//Proceedings of the 2015 ACM/IEEE 42nd Annual International Symposium on Computer Architecture. Piscataway: IEEE Press, 2015: 336-348.

[6] KAUTZ W H. Cellular logic-in-memory arrays[J]. IEEE Transactions on Computers, 1969, C-18(8): 719-727.

[7] JEDDELOH J, KEETH B. Hybrid memory cube new DRAM architecture increases density and performance[C]///Proceedings of the 2012 Symposium on VLSI Technology. Piscataway: IEEE Press, 2012: 87-88.

[8] VAHIDPOUR M, O'BRIEN W, WHYLAND J T, et al. Superconducting through-silicon vias for quantum integrated circuits[EB/OL]. (2017-08-07)[2024-03-12]. https://arxiv.org/abs/1708.02226.

[9] LI T, HOU J, YAN J L, et al. Chiplet heterogeneous integration technology—status and challenges[J]. Electronics, 2020, 9(4): 670.

[10] AHN J, HONG S, YOO S, et al. A scalable processing-in-memory accelerator for parallel graph processing[C]//Proceedings of the 2015 ACM/IEEE 42nd Annual International Symposium on Computer Architecture. Piscataway: IEEE Press, 2015: 105-117.

[11] IMANI M, KIM Y, ROSING T. MPIM: multi-purpose in-memory processing using configurable resistive memory[C]//Proceedings of the 2017 22nd Asia and South Pacific Design Automation Conference. Piscataway: IEEE Press, 2017: 757-763.

[12] STRUKOV D B, SNIDER G S, STEWART D R, et al. The missing memristor found[J]. Nature, 2008, 453(7191): 80-83.

[13] KHAN K, PASRICHA S, KIM R G. A survey of resource management for processing-in-memory and near-memory processing architectures[J]. Journal of Low Power Electronics and Applications, 2020, 10(4): 30.

[14] HSIEH K, EBRAHIM E, KIM G, et al. Transparent offloading and mapping (TOM): enabling programmer-transparent near-data processing in GPU systems[C]//Proceedings of the 2016 ACM/IEEE 43rd Annual International Symposium on Computer Architecture. Piscataway: IEEE Press, 2016: 204-216.

[15] TSAI P A, CHEN C P, SANCHEZ D. Adaptive scheduling for systems with asymmetric memory hierarchies[C]//Proceedings of the 2018 51st Annual IEEE/ACM International Symposium on Microarchitecture. Piscataway: IEEE Press, 2018: 641-654.

[16] XIAO Y, NAZARIAN S, BOGDAN P. Prometheus: processing-in-memory heterogeneous architecture design from a multi-layer network theoretic strategy[C]//Proceedings of the 2018 Design, Automation & Test in Europe Conference & Exhibition. Piscataway: IEEE Press, 2018: 1387-1392.

[17] LI Z R, CHEN X M, HAN Y H. Optimal data allocation for graph processing in processing-in-memory systems[C]//Proceedings of the 2022 27th Asia and South Pacific Design Automation Conference. Piscataway: IEEE Press, 2022: 238-243.

[18] RYDER B G. Constructing the call graph of a program[J]. IEEE Transactions on Software Engineering, 1979, SE-5(3): 216-226.

[19] NAI L F, HADIDI R, SIM J, et al. GraphPIM: enabling instruction-level PIM offloading in graph computing frameworks[C]//Proceedings of the 2017 IEEE International Symposium on High Performance Computer Architecture. Piscataway: IEEE Press, 2017: 457-468.

[20] YAN L, ZHANG M Z, WANG R J, et al. CoPIM: a concurrency-aware PIM workload offloading architecture for graph applications[C]//Proceedings of the 2021 IEEE/ACM International Symposium on Low Power Electronics and Design. Piscataway: IEEE Press, 2021: 1-6.

[21] BOROUMAND A, GHOSE S, PATEL M, et al. CoNDA: efficient cache coherence support for near-data accelerators[C]//Proceedings

of the 2019 ACM/IEEE 46th Annual International Symposium on Computer Architecture. Piscataway: IEEE Press, 2020: 629-642.

[22] XU S, CHEN X M, WANG Y, et al. CuckooPIM: an efficient and less-blocking coherence mechanism for processing-in-memory systems[C]//Proceedings of the 2019 24th Asia and South Pacific Design Automation Conference. Piscataway: IEEE Press, 2024: 1-6.

[23] SONG L H, QIAN X H, LI H, et al. PipeLayer: a pipelined ReRAM-based accelerator for deep learning[C]//Proceedings of the 2017 IEEE International Symposium on High Performance Computer Architecture. Piscataway: IEEE Press, 2017: 541-552.

[24] ZOU K W, WANG Y, LI H W, et al. XORiM: a case of in-memory bit-comparator implementation and its performance implications[C]//Proceedings of the 2018 23rd Asia and South Pacific Design Automation Conference. Piscataway: IEEE Press, 2018: 349-354.

[25] NAI L F, HADIDI R, XIAO H, et al. Thermal-aware processing-in-memory instruction offloading[J]. Journal of Parallel and Distributed Computing, 2019, 130: 193-207.

[26] LI S C, XU C, ZOU Q S, et al. Pinatubo: a processing-in-memory architecture for bulk bitwise operations in emerging non-volatile memories[C]//Proceedings of the 2016 53nd ACM/EDAC/IEEE Design Automation Conference. Piscataway: IEEE Press, 2016: 1-6.

[27] HADIDI R, NAI L F, KIM H, et al. CAIRO: a compiler-assisted technique for enabling instruction-level offloading of processing-in-memory[J]. ACM Transactions on Architecture and Code Optimization, 2017, 14(4): 1-25.

[28] AHMED H, SANTOS P C, LIMA J P C, et al. A compiler for automatic selection of suitable processing-in-memory instructions[C]// Proceedings of the 2019 Design, Automation & Test in Europe Conference & Exhibition. Piscataway: IEEE Press, 2019: 564-569.

[29] JIN H, CHEN D, ZHENG L, et al. Accelerating graph convolutional networks through a PIM-accelerated approach[J]. IEEE Transactions on Computers, 2023, 72(9): 2628-2640.

[30] NAI L F, HADIDI R, XIAO H, et al. CoolPIM: thermal-aware source throttling for efficient PIM instruction offloading[C]//Proceedings of the 2018 IEEE International Parallel and Distributed Processing Symposium. Piscataway: IEEE Press, 2018: 680-689.

[31] PATTNAIK A, TANG X L, JOG A, et al. Scheduling techniques for GPU architectures with processing-in-memory capabilities[C]// Proceedings of the 2016 International Conference on Parallel Architectures and Compilation. New York: ACM, 2016: 31-44.

[32] HSIEH K, KHAN S, VIJAYKUMAR N, et al. Accelerating pointer chasing in 3D-stacked memory: challenges, mechanisms, evaluation[C]//Proceedings of the 2016 IEEE 34th International Conference on Computer Design. Piscataway: IEEE Press, 2016: 25-32.

[33] LI J, WANG X, TUMEO A, et al. PIMS: a lightweight processing-in-memory accelerator for stencil computations[C]//Proceedings of the International Symposium on Memory Systems. New York: ACM, 2019: 41-52.

[34] FERRANTE J, OTTENSTEIN K J, WARREN J D. The program dependence graph and its use in optimization[C]//Proceedings of the International Symposium on Programming. Berlin: Springer, 1984: 125-132.

[35] RAILING B P, HEIN E R, CONTE T M. Contech: efficiently generating dynamic task graphs for arbitrary parallel programs[J]. ACM Transactions on Architecture and Code Optimization, 2015, 12(2): 1-24.

[36] GIBBONS P B, MUCHNICK S S. Efficient instruction scheduling for a pipelined architecture[C]//Proceedings of the 1986 SIGPLAN Symposium on Compiler Construction. New York: ACM, 1986: 11-16.

[37] BUCHSBAUM A L, KAPLAN H, ROGERS A, et al. Linear-time pointer-machine algorithms for least common ancestors, MST verification, and dominators[C]//Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing. New York: ACM, 1998: 279-288.

[38] LUK C K, COHN R, MUTH R, et al. Pin: building customized program analysis tools with dynamic instrumentation[J]. ACM SIGPLAN Notices: A Monthly Publication of the Special Interest Group on Programming Languages, 2005, 40(6): 190-200.

[39] BARBERA M V, KOSTA S, MEI A, et al. To offload or not to offload the bandwidth and energy costs of mobile cloud computing[C]//Proceedings of the 2013 Proceedings IEEE INFOCOM. Piscataway: IEEE Press, 2013: 1285-1293.

[40] DELAHAYE D, CHAIMATANAN S, MONGEAU M. Simulated annealing: from basics to applications[M]//GENDREAU M, POTVIN J Y. Handbook of metaheuristics. Berlin: Springer, 2019: 1-35.

[41] XU S, CHEN X M, WANG Y, et al. PIMSim: a flexible and detailed processing-in-memory simulator[J]. IEEE Computer Architecture Letters, 2019, 18(1): 6-9.

[42] BEAMER S, ASANOVIĆ K, PATTERSON D. The GAP benchmark suite[EB/OL]. (2017-05-16)[2024-03-12]. https://doi.org/10.48550/arXiv.1508.03619.

[43] LESKOVEC J, SOSIČ R. SNAP: a general-purpose network analysis and graph-mining library[J]. ACM Transactions on Intelligent Systems, 2016, 8(1): 1-20.

[44] MURALIMANOHAR N, BALASUBRAMONIAN R, JOUPPI N P. CACTI 6.0: a tool to model large caches: HPL-2009-85[R]. Palo Alto: HP Laboratories, 2009.

[45] GÓMEZ-LUNA J, EL HAJJ I, FERNANDEZ I, et al. Benchmarking a new paradigm: experimental analysis and characterization of a real processing-in-memory system[J]. IEEE Access, 2022, 10: 52565-52608.

Enhancing performance through static computing partitioning approach in processing-in-memory systems

XUE Hongyu[1], XU Sheng[1, 2], LUO Le[1], YAN Liang[3, 4], ZOU Xingqi[3, *]

(1. School of Computer and Information, Anhui Normal University, Wuhu 241000, China;
2. Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 230094, China;
3. Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100083, China;
4. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Processing-in-memory (PIM) systems mitigate the von Neumann "memory-wall" bottleneck by integrating in-memory computing units to break the conventional memory-computation separation paradigm. However, PIM architectures are incompatible with mainstream software stacks, their performance and energy efficiency are highly constrained by the computational partitioning of programs, and may even suffer from performance degradation or negative optimization. In this paper, we propose a static computing partitioning approach that deals with this challenge. The key insight of our work is to reframe the computing partitioning as an annotated call graph (ACG) partitioning problem and propose a simulated annealing-based algorithm to find the optimal computing partitions. In comparison to traditional methods, our trials show that our methodology can improve performance by 39% and cut energy use by an average of 32%.

Keywords: processing-in-memory; computing partitioning; min-cut; simulated annealing; heterogeneous system

Received: 2024-04-10; **Accepted**: 2024-08-23; **Published Online**: 2024-09-13 16:17

URL: link.cnki.net/urlid/11.2625.V.20240912.1433.003

Foundation items: National Natural Science Foundation of China (62102005,62104230); Anhui Provincial Natural Science Foundation (2008085QF330, 2108085QF265); The University Synergy Innovation Program of Anhui Province (GXXT-2021-011); Research Program of Anhui Normal University (751968)

* **Corresponding author**. E-mail: zouxingqi@ict.ac.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2025.0059

鹊桥二号中继星热控设计特点与飞行验证

杨昌鹏[1,2,*]，任红艳[1]，孙腾飞[1]，周强[1]，黄兴[1]，
张立华[3]，熊亮[3]，孙骥[3]，程文龙[2]

(1. 北京空间飞行器总体设计部 航天器热控全国重点实验室，北京 100094；2. 中国科学技术大学工程科学学院，合肥 230027；3. 航天东方红卫星有限公司，北京 100094)

摘 要：鹊桥二号中继星是探月工程四期任务的重要组成部分。相比嫦娥四号鹊桥中继星，鹊桥二号中继星运行轨道由地月拉格朗日L2点的Halo轨道变成环月大椭圆轨道，中继通信载荷和科学载荷配置均有较大变化。基于此，分析环月大椭圆轨道太阳热流、月球红外热流随轨道变化的规律，为鹊桥二号中继星热控系统设计提供了依据。为满足鹊桥二号中继星多任务模式、轨道姿态约束下的控温需求，热控系统采用基于环路热管技术散热能力可调节的设计方案，同时，设计搭载正温度系数(PTC)自控温加热器，并成功在轨应用。鹊桥二号中继星发射入轨后，各飞行阶段设备温控水平良好。嫦娥六号任务期间，鹊桥二号中继星前向X固放等大功率中继载荷温度控制在22～28 ℃之间，有力保障了人类首次月背采样中继通信任务，可为深空探测器高适应能力热控系统设计提供参考。

关 键 词：鹊桥二号中继星；高适应能力；热控系统设计；环路热管；正温度系数加热器
中图分类号：V221^{+}.3；TB553
文献标志码：A **文章编号**：1001-5965(2026)06-2054-10

鹊桥二号中继星是探月工程四期任务的重要组成部分，其任务目标是运行在环月大椭圆冻结使命轨道，支持月球探测任务，实现月面多次任务(嫦娥四号、嫦娥六号、嫦娥七号、嫦娥八号)、多器(着陆器、飞跃探测器、勘察巡视器等)和地面站之间的中继通信，同时，搭载科学与技术试验载荷，完成深空甚长基线干涉测量(very long baseline interferometry, VLBI)和射电观测研究等工作[1-3]。

鹊桥二号中继星采用环月太阳同步大椭圆冻结轨道作为嫦娥六号服务的使命轨道，可以满足嫦娥六号不同发射窗口的中继需求[4]，相比位于地月平动点Halo轨道的嫦娥四号鹊桥中继星，其所经历的轨道热流环境更为复杂、恶劣。鹊桥二号中继星在中继通信载荷和科学载荷配置上均有较大变化，卫星功耗大幅提升，大功率载荷设备散热需求迫切。在轨8年任务期间，由于受到地球遮挡而造成阴影的影响，鹊桥二号中继星最长阴影时长达到近5 h，面临无任何光照极端低温的考验。因此，统筹兼顾中继通信模式高功耗散热和长阴影模式低功耗保温是鹊桥二号中继星热控系统设计的关键。

深空探测器面临极端温度环境，传统的被动热控技术(如多层隔热材料)难以满足需求。可调控热控技术通过动态调控热传导、辐射或吸收特性，成为保障深空探测器安全运行的关键技术。美国国家航空航天局(National Aeronautics and Space Administration, NASA)“旅行者”号探测器采用百叶窗式热控系统，通过双金属弹簧驱动叶片开合，

收稿日期：2025-01-22；录用日期：2025-04-11；网络出版时间：2025-04-17 13:28
网络出版地址：link.cnki.net/urlid/11.2625.V.20250417.1023.001
基金项目：国家重点研发计划(2022YFC2204400)；中国探月工程四期
*通信作者. E-mail：yangchangpeng501@163.com

引用格式：杨昌鹏，任红艳，孙腾飞，等. 鹊桥二号中继星热控设计特点与飞行验证[J]. 北京航空航天大学学报，2026，52(6)：2054-2063. YANG C P，REN H Y，SUN T F，et al. Thermal control design features and flight validation of Queqiao-2 satellite[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：2054-2063 (in Chinese).

调节辐射散热面积。该技术后续应用于 NASA 的 ST-5 卫星和欧洲航天局(European Space Agency, ESA)的“罗塞塔”彗星探测器, 通常需要搭配温度传感器、驱动器、驱动元件、控制电路、供电系统等部件, 增加了航天器质量、体积及能源消耗, 也增加了相应的制造和发射成本[5-6]。环路热管是一种靠蒸发器内的毛细芯产生毛细抽力驱动回路运行, 利用工质相变来传递热量的高效传热装置, 具有长距离传输、传热能力强、逆重力及无运动部件等优点, 已在嫦娥五号返回器、高分七号、陆地探测四号 01 星等多个型号成功应用[7-9]。此外, 环路热管具有良好的热二极管特性, 可以有效阻断热量的传递过程。利用上述特性, 鹊桥二号中继星热控系统设计采用基于环路热管散热能力可调节的热控方案, 满足了中继星在轨多任务模式、轨道姿态约束下的不同控温需求, 显著提升了在轨极端热环境下的生存能力。同时, 针对深空探测器在低温真空环境中的自主补偿加热需求, 设计搭载正温度系数(positive temperature coefficient, PTC)自控温加热器, 并成功在轨应用。本文以鹊桥二号中继星热控研制任务为基础, 研究环月大椭圆轨道探测器的外热流随轨道变化的规律, 开展基于环路热管技术高适应能力热控系统设计及新型 PTC 自控温加热器搭载验证工作, 介绍了鹊桥二号中继星热控系统在轨飞行情况, 可为月球环绕、深空探测器热控设计提供参考。

1　任务分析

鹊桥二号中继星(图 1 所示)任务继承了嫦娥四号鹊桥中继星的任务特点, 分系统组成和平台设备与鹊桥中继星较为接近, 但在运行轨道、构型布局、载荷配置等方面都有了明显的变化, 表 1 给出了鹊桥二号中继星热控设计输入基线变化的对比情况[10]。

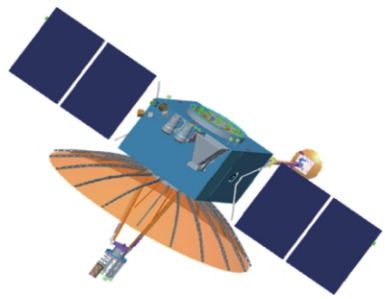


图 1　鹊桥二号中继星示意图

Fig. 1　Queqiao-2 satellite diagram

表 1　鹊桥二号中继星热控设计输入基线变化

Table 1　Input baseline changes of Queqiao-2 satellite thermal control design

中继星	轨道	构型尺寸/(mm×mm×mm)	质量/kg	整星功耗/W	寿命/a
嫦娥四号鹊桥中继星	环地月L2点Halo轨道	箱板式构型 1 400×1 400×850	425	长阴影模式146.6, 中继通信模式最大610	3
鹊桥二号中继星	环月大椭圆轨道(24 h、12 h)	承力筒构型 1 850×1 850×1 172	1 200	长阴影模式554, 中继通信模式最大1 233	8

除具备嫦娥四号鹊桥中继星工作模式多、长阴影功耗资源紧张、4.2 m 伞状抛物面天线反照影响大等技术难点外[11], 为适应鹊桥二号中继星任务需求的变化, 热控系统设计还需考虑以下因素: ①鹊桥二号中继星与鹊桥中继星相比, 轨道条件完全不同, 鹊桥中继星为位于地月平动点 Halo 轨道的卫星, 外热流主要为太阳热流, 鹊桥二号中继星使命轨道为环月大椭圆轨道, 需要分析月球对中继星轨道外热流的影响; ②鹊桥二号中继星载荷设备配套变化较大, 散热需求大幅增加, 中继通信模式和长阴影模式整星功耗的变化范围更大, 热控设计需兼顾高热耗模式下散热和低热耗模式的保温问题, 同时, 尽量减小加热回路对卫星能源的占用; ③受设备布局限制, 部分中继及科学载荷大功率设备约 210 W 布置在卫星顶板(如前向 X 固放、低温电子单元等), 由于受伞状可展开抛物面天线遮挡影响, 卫星+Z 顶板无法开设散热面, 设备散热及地面热试验验证存在困难; ④鹊桥二号中继星寿命由 3 a 提高到 8 a, 寿命末期涂层退化增加, 末期散热压力较大。

2　中继轨道外热流分析

2.1　飞行轨道与卫星姿态

鹊桥二号中继星整个飞行过程包括发射段、地月转移段、月球捕获段、24 h 中继使命段、中继轨道过渡段、12 h 中继使命段, 如图 2 所示, 典型轨道主要包括地月转移轨道、24 h 中继使命轨道和 12 h 中继使命轨道。

1) 地月转移轨道。采用能量最省的转移轨道, 转移时间为 112 h、近月点高度为 100 km、倾角为 90°附近, 为满足中继星对月球南极着陆点的中继, 要求大椭圆冻结轨道的远月点在月球的南半球, 相应地, 近月点在月球的北半球, 因此, 选择地月转移轨道降轨到达月球, 使初始的近月点位于月球的北

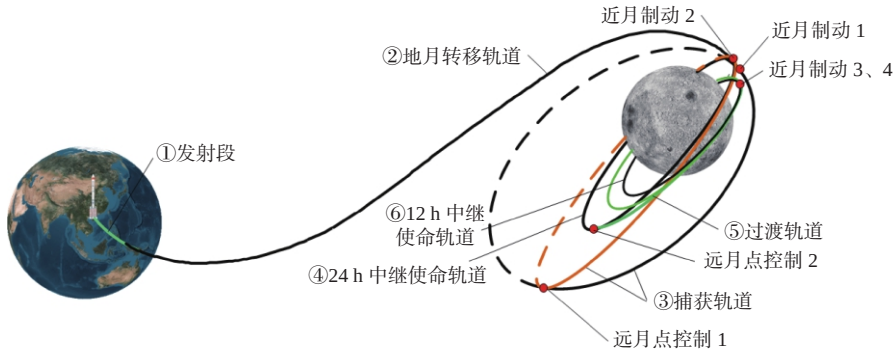


图 2 鹊桥二号中继星飞行轨道

Fig. 2 Queqiao-2 satellite flight orbit

半球，后续进入大椭圆中继轨道需要的速度增量小。

2）24 h 中继使命轨道。第 1 阶段中继星为半长轴约 9 977.2 km 的月球大椭圆轨道，近月点高度为 400 km（在约 100～1 000 km 之间变化），轨道倾角为 118°，轨道周期约为 24.8 h，支持嫦娥六号任务。

3）12 h 中继使命轨道。第 2 阶段中继星为半长轴约 6 285.3 km 的月球大椭圆轨道，近月点高度为 200 km（在约 100～1 000 km 之间变化），轨道倾角为 124°，轨道周期约为 12.4 h，支持嫦娥七号及后续任务。

鹊桥二号中继轨道最长阴影时间出现在中继轨道升交点或降交点附近，太阳在中继轨道面内时，月食期间中继轨道最长阴影时间不超过 300 min。

鹊桥二号中继星在轨运行典型姿态主要有对日定向姿态和对月中继跟踪姿态。在地月转移轨道、捕获轨道及使命轨道无中继任务阶段，中继星采用对日定向姿态，在对日定向姿态下，卫星+z 轴指向太阳，太阳帆板粘贴太阳电池的一面指向+z 轴，卫星其他表面基本不受照射。

在使命轨道有中继任务阶段，中继星采用对月中继跟踪姿态，同时，太阳帆板跟踪太阳，如图 3 所示。该姿态具体描述为：卫星+z 轴在惯性空间的指向 $\boldsymbol{T}_{Z\mathrm{I}}$，$\boldsymbol{T}_{Z\mathrm{I}}$ 为卫星指向月面目标单位矢量；卫星+y 轴或−y 轴在惯性空间的指向 $\boldsymbol{T}_{y\mathrm{I}}=\boldsymbol{T}_{z\mathrm{I}}\times\boldsymbol{S}_{\mathrm{I}}$，$\boldsymbol{S}_{\mathrm{I}}$ 为太阳矢量，双矢量定姿确定目标姿态。在中继星二维数传天线与地球不可见时，中继星绕 z 轴偏航机动 180°实现二维数传天线对地跟踪。在一轨过程中，卫星±x、±z 面均可能受照，±y 面不受照。

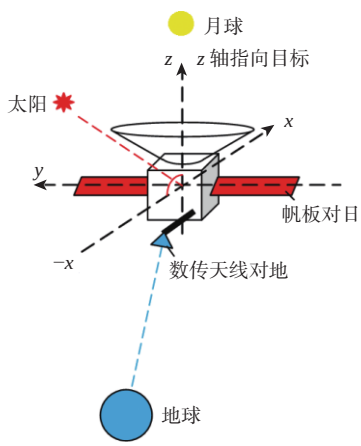


图 3 鹊桥二号中继星对月中继跟踪姿态示意图

Fig. 3 Schematic diagram of Queqiao-2 satellite's lunar relay tracking attitude

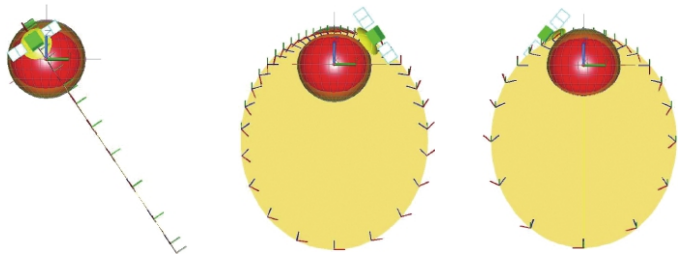
2.2 外热流分析

鹊桥二号中继星使命轨道采用 24 h 和 12 h 环月大椭圆轨道，与嫦娥一号卫星、嫦娥二号卫星、嫦娥四号鹊桥中继星均有所不同，嫦娥一号卫星使命轨道为高度 200 km 环月圆轨道[12]，嫦娥二号卫星使命轨道为高度 100 km 环月圆轨道。文献 [13] 对常规环月圆轨道外热流进行了分析计算，给出了不同高度环月圆轨道下，卫星太阳和月球红外热流随轨道的变化规律。嫦娥四号鹊桥中继星使命轨道为环地月 L2 点 Halo 轨道，外热流只有太阳热流，太阳热流在卫星 xOz 轨道平面大约 29 d 内完成一次变化循环。由于鹊桥二号中继星运行轨道的差异性及特殊姿态控制要求，其外热流变化规律与上述月球探测器有明显差异。

月球热流包括月球反照热流和月球红外热流，总体上，月球反照热流总量较小，月球红外热流随月球经纬位置和与太阳相对角度而变化。环月轨道外热流计算的难点在于不同经纬度下月面温度的处理，本文通过嫦娥月球探测任务确定的月球表面温度分布模型[14]获取月球表面温度，利用月面温度分布特征数据，结合太阳、中继星在轨道中的位置关系及卫星姿态等信息，计算得到中继星各个面到达的外热流。以 12 h 中继使命轨道为例，典型升交点赤经中继星在环月大椭圆轨道中示意如图 4 所示。

β 角为太阳矢量方向与中继星轨道面的夹角，计算中继星 12 h 中继使命轨道 β 角在−57.5°～57.5°之间的变化，图 5 给出了近月点高度为 100 km，中继星不同升交点赤经各个面的到达外热流。其中，

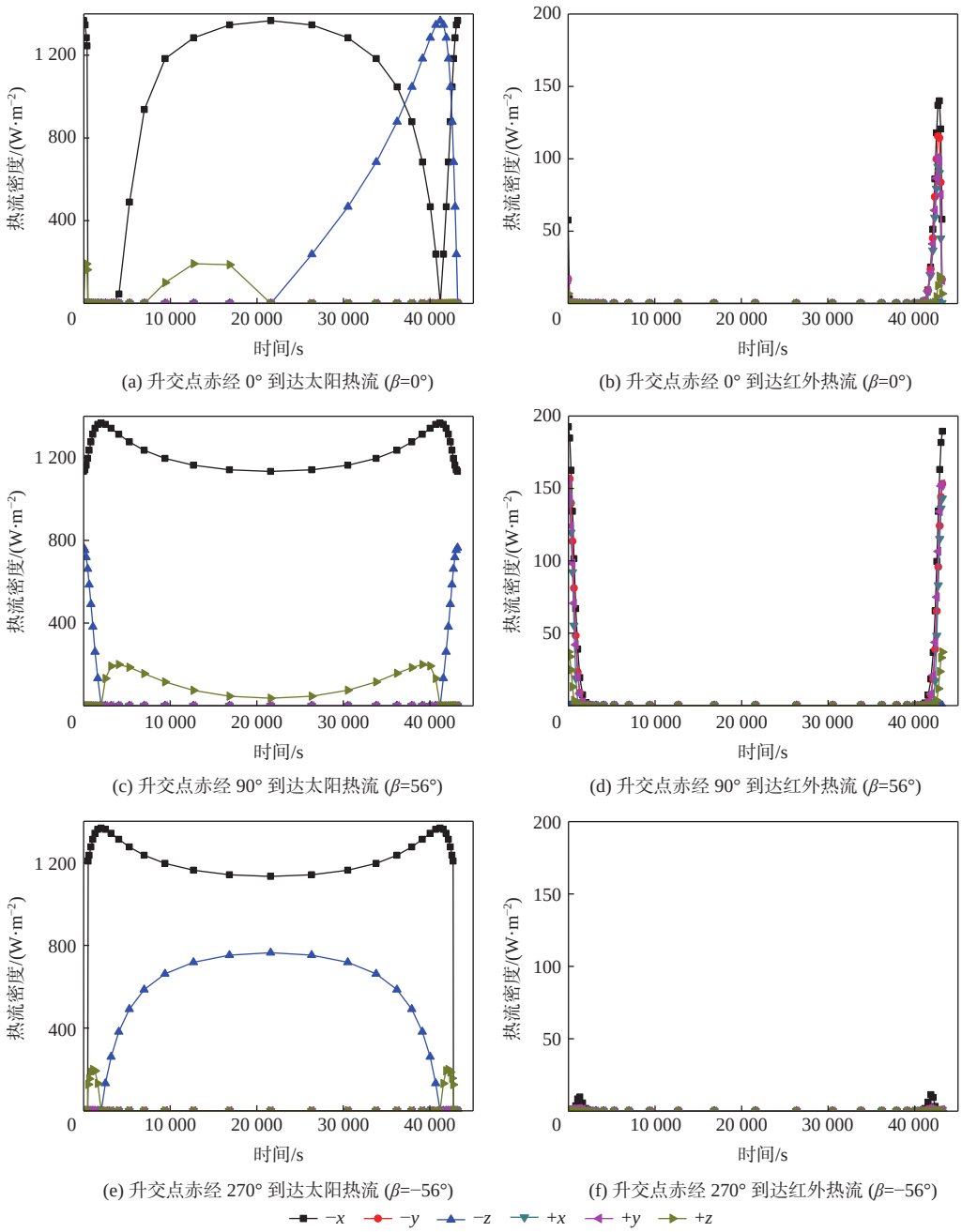
太阳热流包含太阳直射热流和月面反射热流, 红外热流指月球红外热流, 月面反照热流较小, 近月点不超过 20 W/m^2。在此算例中, 约定太阳位于中继星$-x$一侧。



(a) 升交点赤经为 0°　(b) 升交点赤经为 90°　(c) 升交点赤经为 270°

图 4　鹊桥二号中继星环月大椭圆轨道示意图

Fig. 4　Schematic diagram of the Queqiao-2 satellite ringing the Moon in a highly elliptical orbit



(a) 升交点赤经 0° 到达太阳热流 (β=0°)　(b) 升交点赤经 0° 到达红外热流 (β=0°)

(c) 升交点赤经 90° 到达太阳热流 (β=56°)　(d) 升交点赤经 90° 到达红外热流 (β=56°)

(e) 升交点赤经 270° 到达太阳热流 (β=−56°)　(f) 升交点赤经 270° 到达红外热流 (β=−56°)

图 5　鹊桥二号中继星环月大椭圆轨道各面到达外热流曲线

Fig. 5　Arrival heat flow curves of each side of the large elliptical orbit around the Moon of Queqiao-2 satellite

从图5可以看出，月球红外在近月点附近对中继星外热流有较大影响，在远月点附近，对中继星影响可以忽略。升交点赤经为90°时，中继星环月大椭圆轨道近月点月球红外热流最大为192 W/m^2，中继星经过近月点后，随着轨道高度的抬升，月球红外热流快速下降，月球红外热流在100～200 W/m^{2}时间最长不超过20 min，在升交点赤经为270°时，由于中继星近月点位于月球背阴侧，月球红外热流较小，最大为11 W/m^2。

中继星太阳热流与β角及卫星姿态密切相关，由于其在轨的特殊姿态控制要求，卫星$+y$侧板、$-y$侧板在轨基本无阳光入射，而卫星$\pm x$、$\pm z$侧板最大外热流均出现在阳光直照时。从图5可以看出，升交点赤经为90°时，中继星$-x$面太阳热流最大，升交点赤经为270°时，中继星$-z$面太阳热流最大。为保证对地数传天线指向地球，中继星在轨可能绕Z轴偏航机动180°，此时，中继星$-x$面、$-z$面外热流变化剧烈。太阳位于中继星$+x$一侧时，中继星$+x$面最大太阳热流同样出现在升交点赤经90°附近。

由环月大椭圆轨道外热流分析结果可知，月球红外热流轨道周期平均数值较小，太阳热流是中继星在轨工作期间的主要外热流，但在近月点附近，月球红外热流数值较大，对星外设备瞬时加热效应不可忽略，在星外设备热设计时应予以考虑。中继星$\pm y$侧板在轨基本不受太阳照射，因此，$\pm y$面是整星散热面最理想的选择。

3 鹊桥二号中继星热控系统设计

3.1 热控技术方案

根据鹊桥二号中继星热控系统的功能需求和卫星经历的空间热环境，热控系统设计基本思路是：统筹考虑星上资源（主要是电源）、设备温度要求和热环境等因素，采取适当的措施和手段兼顾大功耗设备散热及长阴影工况保温需求，充分可控地利用星上资源，发挥被动热控措施，尤其是热管的作用，在保证系统可靠性的基础上，使整个热控系统性能达到相对最优。

3.1.1 散热面设计

由2.2节分析可知，中继星在轨寿命期间，$\pm y$面均无太阳热流入射。卫星$+z$面由于伞状抛物面天线的遮挡，不宜选作散热面。因此，选择$\pm y$侧板作为整星的主散热面，$\pm x$侧板、$-z$底板作为辅助散热面。散热面选用低太阳吸收比、高发射率的热控涂层，其中，$\pm y$侧板散热面选用KS-ZA白漆（太阳吸收比为0.15，红外发射率为0.92），$\pm x$侧板、$-z$底板选用空间稳定性更好的铈玻璃镀银光学太阳反射镜片（太阳吸收比为0.10，红外发射率为0.79）。卫星外表面除散热面外，其余外表面均包覆15单元多层隔热材料，多层面膜为F46镀银二次表面镜，如图6所示。

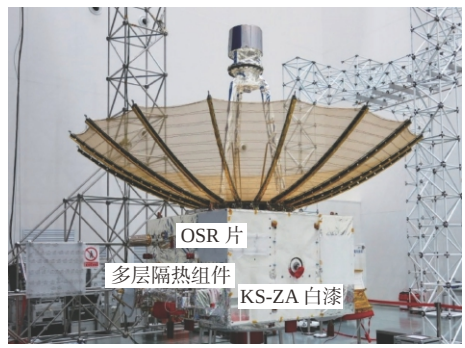


图6 鹊桥二号中继星散热面与多层隔热组件

Fig. 6 Queqiao-2 satellite radiator surface and multi-layer insulation component

3.1.2 热管设计

中继星采用“承力筒+隔板式”的结构，隔板将空间分成若干个封闭的小舱，而星体的热量通过几个散热面向空间排散，因此，会出现一些仪器设备不能直接面对散热面的情况，造成整星温差大甚至局部温度过高的情况。解决上述问题的有效办法是采用预埋热管及外贴热管改善舱内仪器设备之间的换热，以减小仪器设备与散热面之间的温差。在星内主要设备安装板都设计有预埋热管，如前向X固放A/B、电源控制器等大热耗设备及蓄电池、氢原子钟等需要独立控温的设备，另外，外贴热管与预埋热管构建了一体化散热体系，图7所示三维模型中，黄色热管为预埋热管，红色热管为外贴热管，热管设计有效提高了整星散热能力和等温化水平。

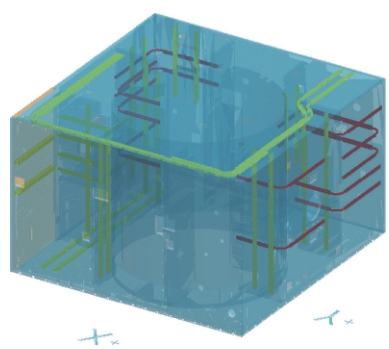


图7 鹊桥二号中继星热管布局示意图

Fig. 7 Queqiao-2 satellite heat pipe layout diagram

3.1.3 主动控温设计

为提高热控系统的调控能力，在热控设计时，采取以电加热为主的主动热控措施。整星电加热回路主要包括蓄电池加热回路、贮箱加热回路、肼

管路加热回路、星敏感器加热回路、二维抛物面天线加热回路、科学载荷加热回路及舱板补偿加热回路等。根据中继星在轨运行状态, 除正常控温模式外, 增加了进影前升温模式和阴影模式设计, 用于长阴影前整星温度的抬升及进影后关键平台设备(蓄电池、肼管路、发动机等)的保温, 有效降低了在轨极端低温工况热控系统对电功耗资源的占用。

3.2　基于环路热管可调控散热设计

鹊桥二号中继通信分系统、科学载荷分系统部分大热耗设备(热耗共计约 210 W)安装在卫星+z 顶板, 卫星+z 面由于伞状抛物面天线的遮挡, 不宜选作散热面。传统外贴热管方案利用+z 顶板与±y 侧板耦合传热, 由于外贴热管热端布局在+z 顶板, 整星热试验逆重力条件下无法顺利启动, 地面验证存在困难, 同时, 在轨长阴影模式时载荷设备关机, 顶板加热回路需补偿热耗 90 W 以上, 对电功耗资源消耗较大。

为满足顶板大热耗设备的散热需求, 同时兼顾长阴影工况保温及热平衡试验热管逆重力启动验证问题, 热控设计采用基于环路热管的可调控散热方案, 即采用预埋热管对顶板设备热量进行收集, 通过环路热管与+y 侧板辐射器耦合进行热量排散, 设计原理如图 8 所示, 三维设计模型如图 9 所示, 具体设计状态如下:

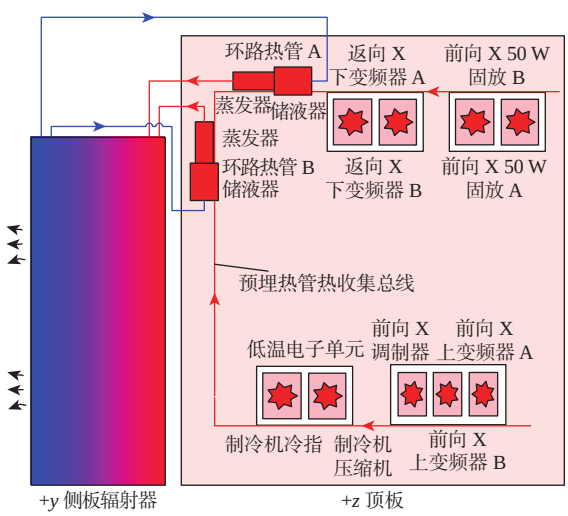


图 8　基于环路热管可调控散热设计原理

Fig. 8　Principle of controllable heat dissipation design based on loop heat pipe

1) +z 顶板为铝蜂窝板结构, 其上安装有前向 X 50 W 固放 A/B、前向 X 调制器、VLBI 低温电子单元等大功率载荷设备, 舱板内部设计预埋 4 根热管, 热管两两一组, 穿过上述设备安装面, 实现多点分散热源热量收集。

2) 设计 2 套 C18 型环路热管。其主要由毛细泵、冷凝器和管路 3 部分组成。其中, 毛细泵由鞍座、毛细芯、储液器等部分组成, 储液器安装热电冷却器(thermal electronic cooler, TEC)半导体制冷器用于辅助启动。冷凝器由管路和翅片焊接组成, 管路采用直径为 4 mm 的不锈钢管连接毛细泵和冷凝器, 单套环路热管传热能力≥210 W, 毛细泵与冷凝器间传热温差≤10 ℃。

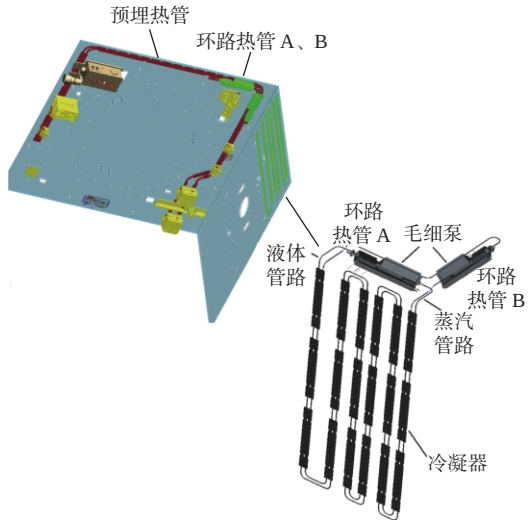


图 9　基于环路热管可调控散热设计三维模型

Fig. 9　3D model of a controllable heat dissipation design based on loop heat pipe

3) 2 套环路热管鞍座与预埋热管区域导热安装, 冷凝器与+y 侧板辐射器涂导热脂安装, 辐射器面积约为 0.64 m^2, 用于+z 侧板载荷设备散热。

2 套环路热管组成了调节中继星+z 顶板大功耗载荷设备与+y 侧板辐射器热量传输的“热开关”, 配合中继星在轨不同工作模式, 环路热管工作过程设计如下:

1) 散热模式。当卫星处于中继通信模式时, 预埋热管收集载荷设备热量后传递给环路热管蒸发器, 热量通过蒸发器鞍座传递到毛细芯中, 加热毛细芯中的氨工质, 产生饱和蒸汽, 蒸汽通过蒸汽管路到达冷凝器, 热量通过+y 侧板辐射器排散至星外, 饱和蒸汽冷凝为过冷液体, 回到储液器中, 液体在毛细泵毛细力的作用下回流到蒸发器, 继续吸热蒸发, 如此循环往复, 完成热量的传输。

2) 保温模式。当中继星载荷设备关机或进入长阴影模式时, 环路热管通过对储液器加热提高储液器压力, 阻断工质流动, 从而将顶板载荷设备至散热面的传热路径断开, 保证了载荷设备在低温工况下的温度水平。

环路热管毛细芯毛细驱动力大, 逆重力能力强, 解决了地面热试验热管逆重力启动验证问题, 同时, 具有良好的热二极管特性, 在长阴影设备关

机时，可有效阻断设备向散热面的传热，经仿真分析验证，长阴影模式热控补偿功耗降低约 90 W。基于环路热管可调控散热设计，可有效提升中继星在轨极端阴影工况下的生存能力。

3.3 PTC 自控温加热器搭载试验设计

PTC 自控温加热器集成了加热、感温和控制 3 项功能，核心部件为 PTC 材料。该材料为具有温度敏感特性的导电材料，其电阻率随自身温度升高而增大，电阻率开始发生突变的温度称为居里点，其基本特性如图 10 所示。基于此特性，PTC 自控温加热器可以根据控温对象温度实现加热功率的自适应调节，将控温对象温度保持在目标温度附近，实现恒温控制。相比传统的加热控温系统，PTC 自控温加热器使用开环控制方式，不需要增加额外的温度传感器和控制设备，即可实现一定范围的自控加热，具有简洁架构、可靠性高、鲁棒性强等优点，PTC 加热技术在深空探测、小卫星等领域具有良好的应用前景。本文利用鹊桥二号卫星平台，结合空间飞行器的温控需求和环境特点，开展 PTC 自控温加热器搭载试验研究工作。

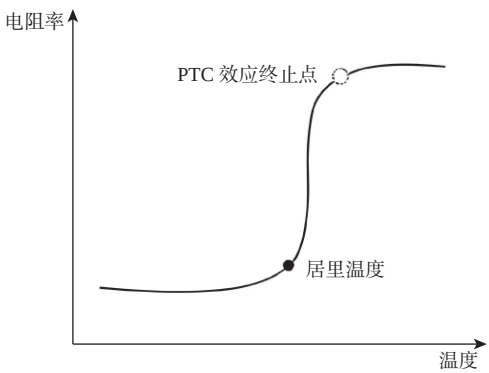


图 10 PTC 自控温加热器典型阻温曲线

Fig. 10 Typical resistance temperature curve of PTC heater

目前，国内外研究的 PTC 材料的居里点普遍较高，一般在 50～400 ℃。鹊桥二号中继星搭载的 PTC 自控温加热器组件由中国科学技术大学研制，采用陶瓷基 PTC 材料，可实现居里温度在室温 0～40 ℃、电阻率在 800～1 500 Ω·cm 之间可调[15]。PTC 材料使用 GD414 硅橡胶固封在铝合金盒体内部。PTC 自控温加热器组件外形尺寸为 66 mm×66 mm×4.8 mm，质量为 53 g，产品实物如图 11 所示。PTC 自控温加热器在各温度下的功率特性如图 12 所示（供电电压为 30.5 V），温度从 0 ℃ 变化到 40 ℃ 时，PTC 自控温加热器功率减小了 95%，即电阻增大了 18 倍。

PTC 自控温加热器使用 M4 紧固件安装在卫星+y 短隔板动量轮 X 线路附近，底部涂 GD414 硅

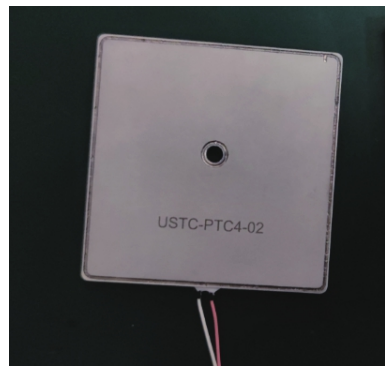


图 11 鹊桥二号中继星搭载的 PTC 自控温加热器产品

Fig. 11 PTC heater products carried on Queqiao-2 satellite

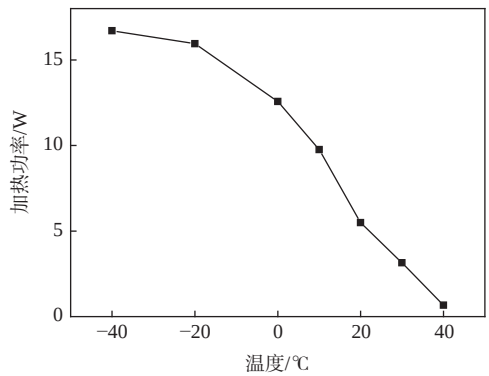


图 12 PTC 自控温加热器功率随温度的变化曲线

Fig. 12 Power versus temperature curve of the PTC heater

橡胶导热安装，如图 13 所示。中继星为 PTC 自控温加热器提供一路供电回路，电压为 30.5 V，在轨默认为开环关闭状态。在 PTC 自控温加热器本体和舱板安装面附近各粘贴一枚热敏电阻，编号分别为 TR8 和 TR7。卫星入轨后，选择合适时机开启 PTC 自控温加热器加热回路，通过热敏电阻温度遥测数据判断 PTC 自控温加热器空间环境下的实际工作性能。

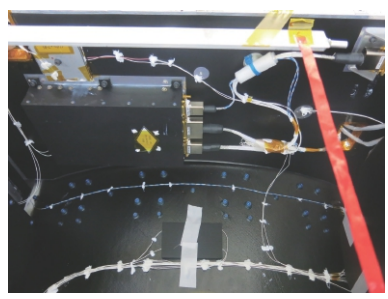


图 13 PTC 自控温加热器星上安装照片

Fig. 13 Photo of onboard installation of the PTC heater

4 飞行验证

鹊桥二号中继星于 2024 年 3 月 20 日由长征八号运载火箭发射进入地月转移轨道，经中途修正、近月制动、轨道捕获与修正后，4 月 2 日进入 24 h

环月大椭圆冻结轨道, 6 月 2—4 日支持嫦娥六号探测器完成人类首次月球背面落月采样任务, 7 月开展科学载荷在轨测试。

图 14 给出了鹊桥二号中继星地月转移段(3 月 20 日—3 月 24 日)典型平台和载荷设备(中心控制单元 TK11、S/X 双频数字应答机 A TC2、电源控制器 TN1、星务综合管理单元 TS1、科学载荷管理器 TY1、综合数据处理器 TZ1、前向 X 50 W 固放 A TZ15)温度曲线, 该阶段中继星采用+*z* 对日定向姿态, 外热流主要为太阳热流, 除+*z* 面外, 中继星其他表面基本不受照射, 外热流非常稳定。中继星平台设备处于工作状态, 载荷设备处于待机或关机状态, 因此, 入轨第 2 天(3 月 21 日), 中继星入轨工作状态设置完毕后, 设备温度即达到相对稳定的状态, 整星温度分布在−20～25 ℃ 之间, 温度水平良好。

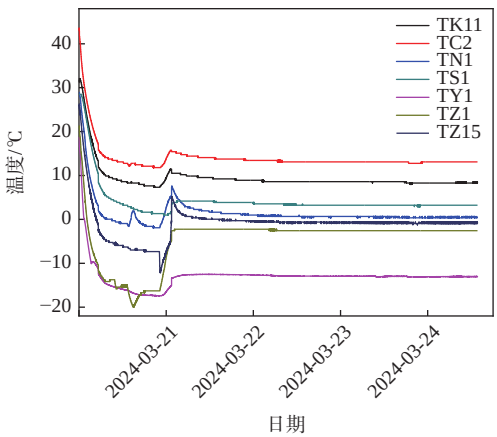


图 14　鹊桥二号中继星地月转移段典型设备温度曲线

Fig. 14　Typical equipment temperature curves of satellite on the Earth-Moon transfer section

图 15 给出了鹊桥二号中继星支持嫦娥六号落月采样任务期间(6 月 2—4 日), 典型平台和载荷设备温度曲线。该阶段中继星处于 24 h 中继使命轨道, 采用对月中继跟踪姿态, 中继通信载荷长期工作, 科学载荷处于关机状态。中继通信载荷开机后, 前向 X 50 W 固放 A(长期热耗 120 W)等大功率设备温升明显, 在基于环路热管散热系统的调控下, 最高温度控制在 28 ℃ 以下。由于运行轨道的周期性变化, 整星温度呈现较为明显的周期性变化规律。嫦娥六号任务期间, 整星温度水平控制在 −15～35 ℃ 之间, 为大功率中继载荷连续开机提供了良好的温度环境。

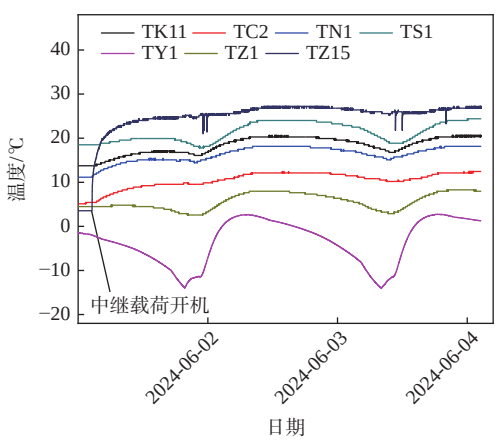


图 15　鹊桥二号中继星 24 h 中继使命轨道典型设备温度曲线

Fig. 15　Typical equipment temperature curves on 24 hour relay mission orbit of Queqiao-2 satellite

鹊桥二号中继通信载荷开机前, 环路热管 A/B 开机, 环路热管 A/B 启动与工作过程温度曲线如图 16 和图 17 所示。环路热管 A/B 开机后, 储液器 (TR16、TR22) 温度下降, 冷凝器 (TR19、TR25) 温度上升, 表明环路热管 A/B 顺利启动, 在轨工作期间, 环路热管 A/B 运行稳定, 环路热管 A 蒸发器 (TR17) 和环路热管 B 蒸发器 (TR23) 最高温度均为 16.4 ℃, 环路热管 A/B 蒸发器出口(TR18、TR24)与冷凝器入口 (TR19、TR25) 温差小于 2.5 ℃, 表明环

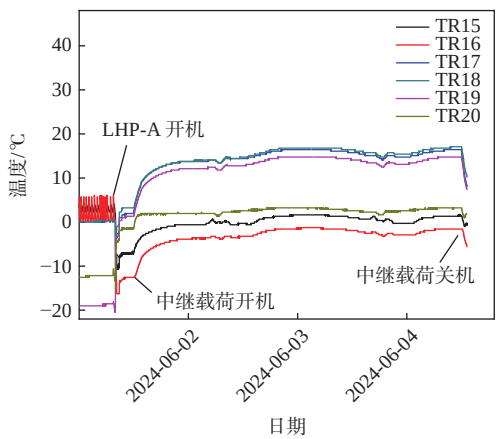


图 16　嫦娥六号任务期间环路热管 A 工作温度曲线

Fig. 16　Loop heat pipe-A working temperature curves during the Chang'e 6 mission

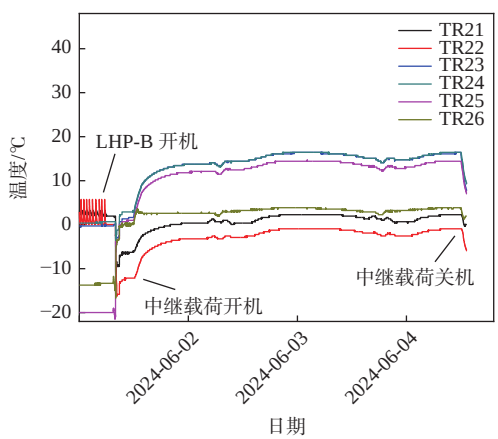


图 17　嫦娥六号任务期间环路热管 B 工作温度曲线

Fig. 17　Loop heat pipe-B working temperature curves during the Chang'e 6 mission

路热管A/B成功将卫星顶板大功率中继载荷热量输送到+y侧板辐射器排散。

鹊桥二号中继星科学载荷在轨测试期间，开展了PTC自控温加热器在轨试验验证工作，性能测试曲线如图18所示。在轨开启PTC自控温加热器加热回路后，PTC自控温加热器本体温度测点TR8从未加热时的3.9 ℃逐渐上升，最终稳定在10.5 ℃附近，PTC自控温加热器安装面温度测点TR7从未加热时的6.5 ℃逐渐上升，最终稳定在9 ℃附近，满足了被控对象设计控温指标(10±5) ℃的要求，在轨工作性能表现稳定。

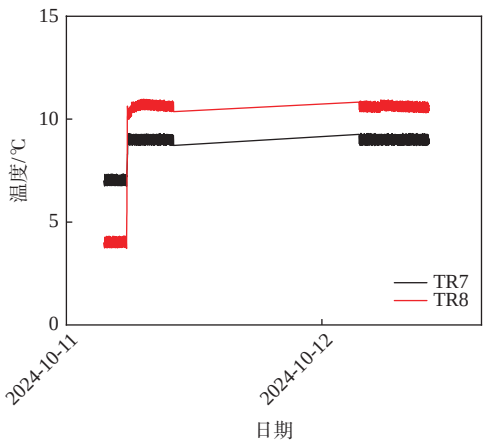


图18 PTC自控温加热器在轨测试温度曲线

Fig. 18 PTC heater temperature curves during the orbit test

5 结 论

本文结合鹊桥二号中继星任务特点和热控需求，开展了环月大椭圆轨道中继星热控系统设计和研制工作。在轨飞行数据表明，热控系统工作正常，为鹊桥二号中继星支持嫦娥六号月背采样中继通信任务提供了良好的温度保障，结论如下：

1）鹊桥二号中继星环月大椭圆轨道外热流变化规律与以往月球探测器有明显差异，需重点考虑其特殊姿态控制对热流变化的影响，同时，在近月点附近，月球红外热流数值较大，对星外设备瞬时加热效应不可忽略。

2）基于环路热管可调控散热设计很好满足了鹊桥二号中继星在轨不同任务模式下的散热与保温需求，具有调控能力强、节约控温资源、布局灵活、可靠性高等优点，拓展了环路热管在深空探测领域的应用场景。

3）室温PTC自控温加热器成功实现在鹊桥二号中继星上的搭载应用，取得了良好的控温效果，为后续基于PTC自控温加热器的自适应热控技术在空间飞行器的推广应用提供了重要的技术支撑。

参考文献（References）

[1] ZHANG L H, XIONG L, SUN J, et al. System design and validation of Queqiao-2 lunar relay communication satellite[J]. Chinese Space Science and Technology, 2024, 44(5): 23-39.

[2] 余后满, 饶炜, 张益源, 等. “嫦娥七号”探测器任务综述[J]. 深空探测学报 (中英文), 2023, 10(6): 567-576.
YU H M, RAO W, ZHANG Y Y, et al. Mission analysis and spacecraft design of Chang’e-7[J]. Journal of Deep Space Exploration, 2023, 10(6): 567-576(in Chinese).

[3] 张立华, 吴伟仁. 月球中继通信卫星系统发展综述与展望[J]. 深空探测学报, 2018, 5(6): 497-505.
ZHANG L H, WU W R. The development overview and prospect of lunar relay communication satellite system[J]. Journal of Deep Space Exploration, 2018, 5(6): 497-505(in Chinese).

[4] 周文艳, 高博宇, 董烜姗, 等. 鹊桥二号轨道与嫦娥六号任务匹配性设计和飞行实践[J]. 中国空间科学技术 (中英文), 2024, 44(6): 16-22.
ZHOU W Y, GAO B Y, DONG T S, et al. Matching design of trajectory with Chang’e-6 mission and flight pratice of Queqiao-2 satellite[J]. Chinese Space Science and Technology, 2024, 44(6): 16-22(in Chinese).

[5] OSIANDER R, FIREBAUGH S L, CHAMPION J L, et al. Microelectromechanical devices for satellite thermal control[J]. IEEE Sensors Journal, 2004, 4(4): 525-531.

[6] 金海波, 凌晨, 李静波. 航天器用可变发射率热控器件的研究进展[J]. 深空探测学报, 2018, 5(2): 188-200.
JIN H B, LING C, LI J B. Development of variable-emittance thermal control technology[J]. Journal of Deep Space Exploration, 2018, 5(2): 188-200(in Chinese).

[7] 宁献文, 苏生, 陈阳, 等. 月地高速再入返回器热控设计及实现[J]. 中国科学:技术科学, 2015, 45(2): 145-150.
NING X W, SU S, CHEN Y, et al. Design and implementation of circumlunar return and reentry spacecraft thermal control system[J]. Scientia Sinica (Technologica), 2015, 45(2): 145-150(in Chinese).

[8] 黄金印, 潘福明, 傅伟纯, 等. 高分七号卫星激光测高仪热设计及验证[J]. 航天器工程, 2020, 29(3): 138-143.
HUANG J Y, PAN F M, FU W C, et al. Thermal design and verification of laser altimeter for GF-7 satellite[J]. Spacecraft Engineering, 2020, 29(3): 138-143(in Chinese).

[9] 孟恒辉, 徐亚威, 韩东阳, 等. 基于钎焊工艺的环路热管耦合系统设计及验证[J]. 中国空间科学技术 (中英文), 2024, 44(5): 57-65.
MENG H H, XU Y W, HAN D Y, et al. Design and verification of loop heat pipe coupling heat system based on brazing technology[J]. Chinese Space Science and Technology, 2024, 44(5): 57-65(in Chinese).

[10] 张立华, 熊亮, 孙骥, 等. 嫦娥四号任务中继星“鹊桥”技术特点[J]. 中国科学:技术科学, 2019, 49(2): 138-146.
ZHANG L H, XIONG L, SUN J, et al. Technical characteristics of the relay communication satellite “Queqiao” for Chang’e-4 lunar farside exploration mission[J]. Scientia Sinica (Technologica), 2019, 49(2): 138-146(in Chinese).

[11] REN H Y, YANG C P, SUN T F, et al. Thermal design and flight validation of the Chang’e-4 relay satellite[J]. Journal of Physics:

Conference Series, 2024, 2691(1): 012048.

[12] 向艳超, 邵兴国, 刘自军, 等. 嫦娥一号卫星热控系统及其特点[J]. 航天器工程, 2008, 17(5): 42-46.
XIANG Y C, SHAO X G, LIU Z J, et al. Thermal control system and its characteristics of Chang'e-1[J]. Spacecraft Engineering, 2008, 17(5): 42-46(in Chinese).

[13] 陈建新, 向艳超, 钟奇, 等. 月球探测器外热流与散热能力分析[J]. 航天器工程, 2015, 24(5): 73-78.
CHEN J X, XIANG Y C, ZHONG Q, et al. Calculating external heating and analysing radiating surface of lunar spacecraft[J]. Spacecraft Engineering, 2015, 24(5): 73-78(in Chinese).

[14] 向艳超, 刘自军, 宁献文, 等. 我国月球探测器热控技术发展[J]. 航天器工程, 2022, 31(2): 29-34.
XIANG Y C, LIU Z J, NING X W, et al. Review of thermal control technology in Chinese lunar probes[J]. Spacecraft Engineering, 2022, 31(2): 29-34(in Chinese).

[15] 桑泽康, 赵锐, 程文龙. 低电阻率陶瓷基 PTC 材料温控特性研究[J]. 北京航空航天大学学报, 2023, 49(8): 2147-2153.
SANG Z K, ZHAO R, CHENG W L. Study on temperature control characteristics of low-resistivity ceramic-based PTC material[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(8): 2147-2153(in Chinese).

Thermal control design features and flight validation of Queqiao-2 satellite

YANG Changpeng[1, 2, *], REN Hongyan[1], SUN Tengfei[1], ZHOU Qiang[1], HUANG Xing[1], ZHANG Lihua[3], XIONG Liang[3], SUN Ji[3], CHENG Wenlong[2]

(1. National Key Laboratory of Spacecraft Thermal Control, Beijing Institute of Spacecraft System Engineering, Beijing 100094, China;
2. School of Engineering Science, University of Science and Technology of China, Hefei 230027, China;
3. DFH Satellite Co., Ltd., Beijing 100094, China)

Abstract: The Queqiao-2 satellite is a crucial part of the fourth Lunar Exploration Program mission. Compared to the Chang'e 4 Queqiao satellite, the Queqiao-2 satellite orbit changes from a Halo orbit at the Earth-Moon Lagrange L2 point to a circumlunar large elliptical orbit, with large changes in both relay communication and science load equipment. In this paper, the variation of solar heat flow and lunar infrared in a circumlunar large elliptical orbit with orbit is analyzed, which provides a basis for the design of the thermal control system of the Queqiao-2 satellite. In order to meet the temperature control requirements of the Queqiao-2 satellite in multi-mission mode and orbital attitude constraints, the thermal control system adopts a design scheme with adjustable heat dissipation capability based on loop heat pipe technology, while the design carries a positive temperature coefficient (PTC) self-controlled temperature heater, which is successfully applied in orbit. After Queqiao-2 was launched into orbit, the level of equipment temperature control during each flight phase was good. The first human lunar dorsal sample relay communications mission was substantially ensured by controlling the temperature of high-power relay loads, such as the Queqiao-2 X-band solid state amplifier, between 22 °C and 28 °C during the Chang'e 6 mission. The design method in this paper can provide a reference for the design of the highly adaptable thermal control system of a deep space exploration mission.

Keywords: Queqiao-2 satellite; high adaptability; thermal control system design; loop heat pipe; positive temperature coefficient heater

Received: 2025-01-22; **Accepted**: 2025-04-11; **Published Online**: 2025-04-17 13:28

URL: link.cnki.net/urlid/11.2625.V.20250417.1023.001

Foundation items: National Key Research and Development Program of China (2022YFC2204400); Fourth Phase of Lunar Exploration Program of China

* **Corresponding author**. E-mail: yangchangpeng501@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0301

碳基材料烧蚀传热问题的不确定性分析方法

刘挺[1]，刘骁[2,*]，郭磊[1,3]，曾磊[2]，国义军[2]

(1. 新疆大学 计算机科学与技术学院，乌鲁木齐 830000；2. 中国空气动力研究与发展中心，绵阳 621000；
3. 电子科技大学 计算机科学与工程学院，成都 610000)

摘　　要：针对烧蚀传热问题进行不确定性量化研究，烧蚀预测采用双平台模型，对烧蚀动边界下的传热过程使用有限元方法进行数值求解。使用正交试验研究参数对目标变量的影响，发现在2种典型加热条件(高热流、短时间和低热流、长时间)下，输入参数对烧蚀量的影响有所不同，而对背面温度的影响大小一致。为得到更精准的不确定性量化结果，进一步使用蒙特卡罗(MC)方法和多项式混沌(PC)方法对烧蚀传热问题进行不确定性分析，通过敏感度分析发现，在2种加热条件下，热流是影响烧蚀量较为关键的因素。而对背面温度的影响，状态1(高热流、短时间)中的热导率影响最大，状态3(低热流、长时间)中的热流影响相对较大。相比MC方法，PC方法能够有效减少计算成本，获得较为满意的计算结果。

关 键 词：碳基材料；双平台理论；不确定性；多项式混沌；正交试验
中图分类号：V19
文献标志码：A　　**文章编号**：1001-5965(2026)06-2064-10

再入式高速飞行器(如返回舱等)在大气内飞行过程中，由于飞行速度极高，会受到严重的气动加热，必须采取热防护措施，碳基材料(包括普通碳/碳复合材料、碳酚醛等)是当前用于再入式飞行器热防护非常广泛的一类高温烧蚀型防热材料。再入飞行过程中，通过该类材料的烧蚀消耗，带走大量热量，从而达到保护飞行器内部结构和设备的目的。

对于烧蚀型热防护的设计而言，烧蚀后退量和防热层内壁面温度是必须重点关注并严格控制的2个变量[1-2]，因为前者产生的外形变化有可能导致过大的落点偏差甚至失稳坠毁，后者超限则会导致内部结构及仪器设备的损坏。而表面烧蚀反应和材料内部温度分布存在复杂的相互影响机制[3-4]，因此，建立同时考虑材料表面烧蚀后退-内部传热的耦合计算方法对于烧蚀热防护设计十分重要[5-6]。同时，受限于预测手段及生产工艺，作为烧蚀传热计算必要输入条件的气动加热热流及材料热物性参数，还存在一定难以避免的系统误差，对这些误差所带来的预测结果不确定性进行量化分析，同样对热防护设计具有重要的参考价值。对于热防护设计中的不确定性问题，Dec等[7]提出使用蒙特卡罗(Monte Carlo, MC)方法来进行热防护系统安全余量概率设计。Chen等[8]在文献[7]的基础上，提出使用MC方法进行热防护概率设计及敏感度分析的通用流程，并展示了该方法在星尘号、火星巡游者及X-37热防护设计中的应用。此后，基于MC的热防护概率设计方法在飞行任务中得到了大量应用，包括火星科学实验室(Mars Science Laboratory, MSL)[9]、Orion Crew Exploration Vehicle[10]及高超声速充气式气动减速器(hypersonic inflatable aerodynamic decelerator, HIAD)[11]等。陈江涛等[12]使用多项式混沌(polynomial chaos, PC)方法开展炭化热解材料温度场预测中的不确定性分析，以提升不确

收稿日期：2024-05-09；**录用日期**：2024-07-05；**网络出版时间**：2024-08-19 15:57
网络出版地址：link.cnki.net/urlid/11.2625.V.20240816.1613.002
基金项目：国家重点研发计划(2019YFA0405202)
*通信作者. E-mail：liuxiao8146@163.com

引用格式：刘挺，刘骁，郭磊，等. 碳基材料烧蚀传热问题的不确定性分析方法[J]. 北京航空航天大学学报，2026，52(6)：2064-2073.
LIU T，LIU X，GUO L，et al. Uncertainty analysis methods for heat transfer ablation in carbon-based materials[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：2064-2073 (in Chinese).

定性分析的效率。

上述研究主要针对材料参数的不确定性对结构内部温度、应力等产生的影响开展研究, 对表面烧蚀考虑较少。针对碳基材料的表面烧蚀, 文献 [13] 提出了双平台理论, 建立了碳基材料在氧化速率控制区、过渡区、扩散控制区及碳氮反应和碳升华的通用烧蚀计算模型, 揭示了烧蚀反应控制机理。本文在其基础上, 建立烧蚀-传热耦合计算方法, 基于该方法, 对 2 种典型加热工况下的烧蚀规律及其背后的反应机理开展分析, 通过正交试验选取典型输入参数值进行计算分析结果, 综合利用 MC 方法和 PC 方法, 开展不同加热条件下的烧蚀传热问题不确定性分析研究。

1　烧蚀传热问题的数值模拟

本文的烧蚀传热问题计算基于碳基材料氧化烧蚀双平台模型[13], 该模型指出, 碳的表面氧化由 2 个基本过程组成:固体表面上的化学动力学过程和反应物向表面输送的过程, 且速率较慢的过程起主导作用。本文基于上述基本条件, 简单阐述碳基材料烧蚀模型。

在材料表面, 主要考虑如下反应[13]:

$$\mathrm{C(s)} \longrightarrow \mathrm{C_1} \tag{1}$$

$$2\mathrm{C(s)} \longrightarrow \mathrm{C_2} \tag{2}$$

$$3\mathrm{C(s)} \longrightarrow \mathrm{C_3} \tag{3}$$

$$4\mathrm{C(s)} \longrightarrow \mathrm{C_4} \tag{4}$$

$$5\mathrm{C(s)} \longrightarrow \mathrm{C_5} \tag{5}$$

$$\mathrm{C(s)}+\frac{1}{2}\mathrm{O_2} \longrightarrow \mathrm{CO}\cdots\cdots\dot{m}_{\mathrm{C_1}} \tag{6}$$

$$\mathrm{C(s)}+\frac{1}{2}\mathrm{N_2} \longrightarrow \mathrm{CN}\cdots\cdots\dot{m}_{\mathrm{C_2}} \tag{7}$$

$$2\mathrm{C(s)}+\frac{1}{2}\mathrm{N_2} \longrightarrow \mathrm{C_2N}\cdots\cdots\dot{m}_{\mathrm{C_3}} \tag{8}$$

$$\mathrm{CO}+\frac{1}{2}\mathrm{O_2} \Longleftrightarrow \mathrm{CO_2} \tag{9}$$

$$\frac{1}{2}\mathrm{O_2} \Longleftrightarrow \mathrm{O} \tag{10}$$

$$\frac{1}{2}\mathrm{N_2} \Longleftrightarrow \mathrm{N} \tag{11}$$

式 (1)～式 (5) 为碳的升华反应; 式 (6) 为非平衡反应; 式 (7)～式 (11) 为平衡反应。

按照上述反应, 反应气体混合物的平均分子质量为

$$\overline{M}_{\mathrm{w}}=\left(\sum_{j=1}^{11}\frac{C_{j,\mathrm{w}}}{M_j}\right)^{-1} \tag{12}$$

式中: $C_{j,\mathrm{w}}$为元素 j 在壁面的质量分数; M_j为元素 j 的摩尔质量; 下标 w 表示壁面。

对于式 (1)～式 (5) 的升华反应, 符合动力学升华过程, 反应消耗的碳速率为

$$\dot{m}_{\mathrm{C}_i,\mathrm{w}}=\beta_{\mathrm{C}_i}\frac{\overline{M}_{\mathrm{w}}P_{\mathrm{e}}}{\sqrt{2\pi\hat{R}M_{\mathrm{C}_i}T_{\mathrm{w}}}}\left(C_{\mathrm{C}_i,\mathrm{eq}}-C_{\mathrm{C}_i,\mathrm{w}}\right) \tag{13}$$

式中: β_{C_i}为无量纲系数, 通过查询相关数据库获得; P_{e}为边界层外缘压力; $\hat{R}$为气体常数; 下标 e 表示边界层外缘; 下标 eq 表示平衡常数; M_{C_i}为碳元素 i 的摩尔质量; T_{w}为壁面温度; C_{C_i}为组元的质量分数。

假设边界层内压力梯度为 0, 即边界层内气体压力等于边界层外缘压力, 可得

$$\dot{m}_{\mathrm{C}_i,\mathrm{w}}=\beta_{\mathrm{C}_i}\sqrt{\frac{M_{\mathrm{C}_i}}{2\pi\hat{R}T_{\mathrm{w}}}}\left(P_{\mathrm{C}_i,\mathrm{eq}}-P_{\mathrm{C}_i,\mathrm{w}}\right) \tag{14}$$

式中: P_{C_i}为碳元素 i 的分压; $P_{\mathrm{C}_i,\mathrm{eq}}=K_{\mathrm{p},\mathrm{C}_i}$为对应反应的平衡常数, 通过查询相关化学反应数据库获得。

引入无量纲质量损失系数:

$$B_{\mathrm{c}}=\frac{\dot{m}_{\mathrm{C,w}}}{\rho_{\mathrm{e}}u_{\mathrm{e}}St}=\frac{\dot{m}_{\mathrm{C,w}}}{\rho_{\mathrm{e}}u_{\mathrm{e}}C_{\mathrm{M}}}=\frac{\dot{m}_{\mathrm{C,w}}}{\psi q_{\mathrm{c}}/h_{\mathrm{r}}} \tag{15}$$

$$B_{\mathrm{p}}=\frac{\dot{m}_{\mathrm{gw}}}{\rho_{\mathrm{e}}u_{\mathrm{e}}C_{\mathrm{M}}}=\frac{\dot{m}_{\mathrm{gw}}}{\psi q_{\mathrm{c}}/h_{\mathrm{r}}} \tag{16}$$

$$B_{\mathrm{w}}=B_{\mathrm{c}}+B_{\mathrm{p}}=\frac{\dot{m}_{\mathrm{w}}}{\rho_{\mathrm{e}}u_{\mathrm{e}}C_{\mathrm{M}}}=\frac{\dot{m}_{\mathrm{w}}}{\psi q_{\mathrm{c}}/h_{\mathrm{r}}} \tag{17}$$

式中: ψ为质量引射系数, 为考虑烧蚀引射效应后对热流的修正系数, 可采用与引射相关的模型计算获得; ρ_{e}为边界层外缘密度; u_{e}为边界层外缘速度; St为斯坦顿数; C_{M}为传质系数; q_{c}为冷壁热流; h_{r}为恢复焓; $\dot{m}_{\mathrm{gw}}$为壁面气体质量流率; $\dot{m}_{\mathrm{w}}$为壁面处元素消耗速率。

组元 j 在壁面的扩散流率为

$$J_{j\mathrm{w}}=\rho_{\mathrm{e}}u_{\mathrm{e}}C_{\mathrm{M}}\left(C_{j,\mathrm{w}}-C_{j,\mathrm{e}}\right) \tag{18}$$

$$\dot{m}_{j,\mathrm{w}}=\left[C_{j,\mathrm{w}}(1+B_{\mathrm{w}})-C_{j,\mathrm{e}}\right]\psi\frac{q_{\mathrm{c}}}{h_{\mathrm{r}}} \tag{19}$$

式中: $\dot{m}_{j,\mathrm{w}}$为壁面处元素 j 消耗速率。

对于反应式 (6), 对不同的温度范围采用不同的反应动力学模型, 当$T_{\mathrm{w}}\leqslant 1\,700$ K时, 采用 1/2 级反应, 当$T_{\mathrm{w}}>1\,700$ K时, 采用 1 级反应:

$$\dot{m}_{\mathrm{C,w1}}=(1-f)\alpha_1P_{\mathrm{O_2,w}}^{1/2}+f\alpha_1'P_{\mathrm{O_2,w}} \tag{20}$$

式中: $f=0$, 采用 1/2 级反应模型, $f=1$, 采用 1 级反应模型; α_1和α_1'为反应速率系数; $P_{\mathrm{O_2}}$为 $\mathrm{O_2}$ 分压。反应动力学数据按照 Scala 提出的快反应模型给出[14]。再对 $\mathrm{O_2}$ 组元使用式 (19), 可得

$$\dot{m}_{\mathrm{O_2,w}}=\left[P_{\mathrm{O_2,w}}(1+B)-P_{\mathrm{O_2,e}}\right]\psi\frac{q_{\mathrm{c}}}{h_{\mathrm{r}}}\cdot\frac{M_{\mathrm{O_2}}}{\overline{M}_{\mathrm{w}}} \tag{21}$$

式中: B 为无因次烧蚀质量流率; $M_{\mathrm{O_2}}$为$\mathrm{O_2}$的摩尔

质量。

定义反应生成物CO与CO_2的分压比为δ，O与CO_2的分压比为γ，有

$$\dot{m}_{O_2,w}=-\frac{(2+\delta+\gamma)M_{O_2}}{(2+2\delta)M_C}\dot{m}_{C,w1} \tag{22}$$

式中：M_C为C的摩尔质量。

随后，可求得$P_{O_2,w}^{1/2}$，O元素当量质量分数为

$$\tilde{C}_{O,w}=C_{O,w}+C_{O_2,w}+\frac{M_O}{M_{CO}}C_{CO,w}+\frac{2M_O}{M_{CO_2}}C_{CO_2,w} \tag{23}$$

式中：$C_{O,w}$、$C_{O_2,w}$、$C_{CO,w}$、$C_{CO_2,w}$分别为O、O_2、CO、CO_2在壁面的质量分数；M_O、M_{CO}、M_{CO_2}分别为O、CO、CO_2的摩尔质量。

根据边界层内的传质过程：

$$P_O+2P_{O_2}+P_{CO}+2P_{CO_2}=\frac{\tilde{C}_{O,e}}{1+B_w}\cdot\frac{\overline{M}_wP_e}{M_O} \tag{24}$$

式中：P_O、P_{O_2}、P_{CO}、P_{CO_2}分别为O、O_2、CO、CO_2的分压。

根据反应式(10)的平衡条件，有

$$K_{p,O}=\frac{P_{O,w}}{P_{O_2,w}^{1/2}}=10^{3.518-13\,392/T_w} \tag{25}$$

式中：$K_{p,O}$为O的化学平衡常数。

由式(25)可求出P_O。根据反应式(9)的平衡条件，进一步可得

$$P_{CO}=\frac{\dfrac{\tilde{C}_{O,e}}{1+B_w}\cdot\dfrac{\overline{M}_wP_e}{M_O}-P_{O,w}-2P_{O_2,w}}{1+2P_{O_2,w}^{1/2}K_{p,CO_2}} \tag{26}$$

式中：K_{p,CO_2}为CO_2的化学平衡常数。

对氮元素有

$$K_{p,N}P_{N_2,w}^{1/2}+2P_{N_2,w}+K_{p,CN}P_{N_2,w}^{1/2}+K_{p,C_2N}P_{N_2,w}^{1/2}=\frac{P_e\overline{M}_w}{M_N}\cdot\frac{\tilde{C}_{N,e}}{1+B_w} \tag{27}$$

式中：$K_{p,N}$为N的化学平衡常数。

至此，壁面所有组分的分压均可求得。进一步可求出壁面气体混合物中碳元素的质量浓度为

$$\tilde{C}_{C,w}=\sum_{i=1}^{3}C_{C_i,w}+\frac{M_C}{M_{CO}}C_{CO,w}+\frac{M_C}{M_{CO_2}}C_{CO_2,w}+\frac{M_C}{M_{CN}}C_{CN,w}+\frac{2M_C}{M_{C_2N}}C_{C_2N,w} \tag{28}$$

并得到B_c和$\dot{m}_C$。

各组分质量分数与无量纲参数B_w、质量引射系数ψ及壁面气体混合物平均分子质量$\overline{M}_w$为隐式关系，通过牛顿迭代可实现求解。

基于以上阐述的双平台模型，结合烧蚀动边界技术，给出烧蚀传热问题在动坐标下的控制方程表达式为

$$\rho c_p\left.\frac{\partial T}{\partial t}\right|_{x_s,y_s,z_s}=\rho c_p\nabla T|_t\cdot v_m+\nabla(\kappa\nabla T)|_t \tag{29}$$

式中：ρ为密度；c_p为比热容；T为温度；x_s、y_s、z_s分别为动坐标下x、y、z方向；∇表示求梯度；v_m为计算单元内任意位置处的微分控制体速度；κ为热传导系数。

式(29)为对流-扩散型方程，采用一般有限元方法结合二阶Euler后差进行时间积分，并使用Newton-GMRES迭代求解大型稀疏矩阵，即可实现其求解。所采用的数值算法详见文献[15]，不再详述。

对于烧蚀热边界条件，可通过表面能量平衡计算获得。温度场计算所需的热流边界条件为

$$-\kappa\left.\frac{\partial T}{\partial|\boldsymbol{n}|}\right|_w=q_{cond}=\rho_eu_eC_H(h_r-h_w)-\varepsilon_w\sigma(T_w^4-T_\infty^4)+\dot{m}_C''h_C+\dot{m}_g''h_g \tag{30}$$

式中：q_{cond}为导热进入固体的净热流；$\boldsymbol{n}$为表面边界单元的单位法向量；C_H为传热系数；$\dot{m}_C''$、$\dot{m}_g''$分别为碳元素、气体质量流率；$\dot{m}_C''h_C$为碳烧蚀引起的能量流动；$\dot{m}_g''h_g$为流入表面的热解气体携带的能量；σ为斯蒂芬-玻尔兹曼常数；ε_w为表面发射系数；T_∞为环境温度；h_w为壁面焓；h_C为烧蚀产物的焓；h_g为热解气体的焓。

至此，本文采用的烧蚀传热问题计算方法介绍完毕。本文计算方法在以往研究中得到了验证。使用碳/碳复合材料试验件开展电弧加热器烧蚀试验。试验件为构型为ϕ15 mm×40 mm的平头圆柱，材料物性详见文献[16]。表1为碳/碳材料模型试验状态及烧蚀测量结果。

表1 碳/碳材料模型试验状态及烧蚀测量结果

Table 1 Carbon/carbon material model test status and ablation measurement results

状态	烧蚀前尺寸/mm	烧蚀后尺寸/mm	焓值/(MJ·kg^{-1})	热流/(MW·m^{-2})	烧蚀时间/s
状态1	40.12	36.5	10	30	15
状态2	40.06	38.5	5	10	15
状态3	40.05	33.5	5	10	45

计算出3种工况下的烧蚀后退量，并与试验结果对比，如图1所示。可以看出，计算值与试验值的偏差在0.11～0.53 mm范围内。考虑到烧蚀后的模型表面本身具有一定的不平整度(0.1 mm量级)，且计算给定的理想状态热环境与试验过程中实际热环境尚有一定偏差，可以认为，计算结果与试验值达到了较好的吻合度。

研究中发现，烧蚀、热流及结构温度存在复杂的相互作用关系，如图2所示。状态1和状态3分别属于“高热流、短时间”和“低热流、长时间”。可以看出，温度在持续升高，而烧蚀速率的变化并

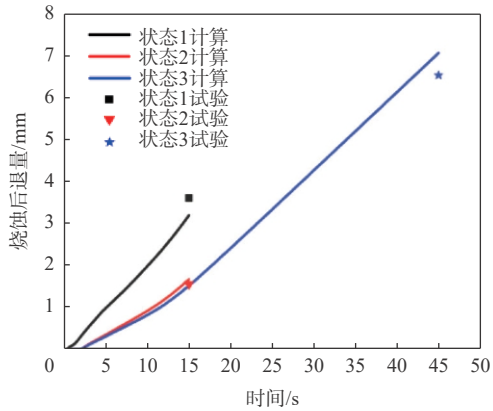


图 1　烧蚀后退量随时间变化

Fig. 1　Change in ablation retreat over time

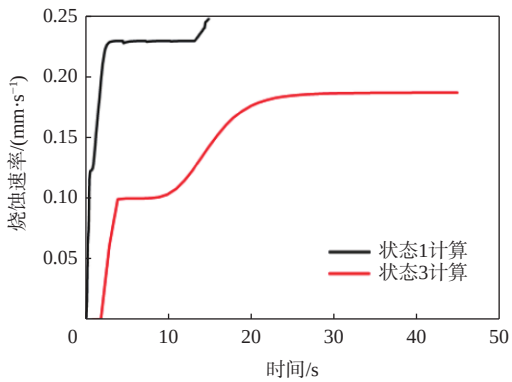


图 2　状态 1 和状态 3 烧蚀速率随时间变化

Fig. 2　The ablation velocity of state 1 and state 3 changes over time

没有规律，2 种状态下的烧蚀速率曲线呈现出比较复杂、有显著分段平台特征的变化规律。与状态 1 相比，状态 3 的烧蚀速率曲线爬升段更缓，在第 1 个平台停留的时间更长，且末端没有出现显著的爬升。由于反应温度的上升，反应速率分别受到反应动力学控制、反应动力学和反应物扩散共同控制、反应控制 3 个过程。从图 3 和图 4 可以看到，状

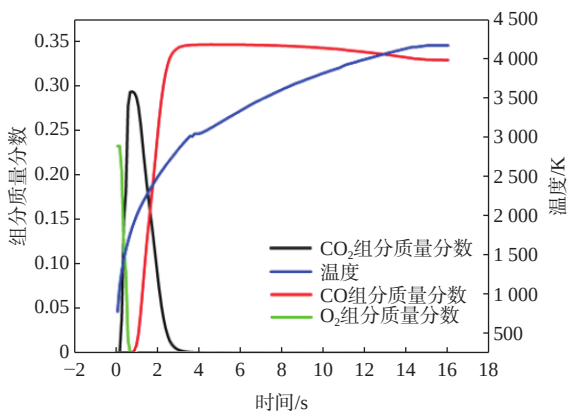


图 3　状态 1 碳-氧反应相关组分质量分数曲线及表面温升曲线

Fig. 3　Concentration curve and surface temperature rise curve of carbon oxygen reaction related components in state 1

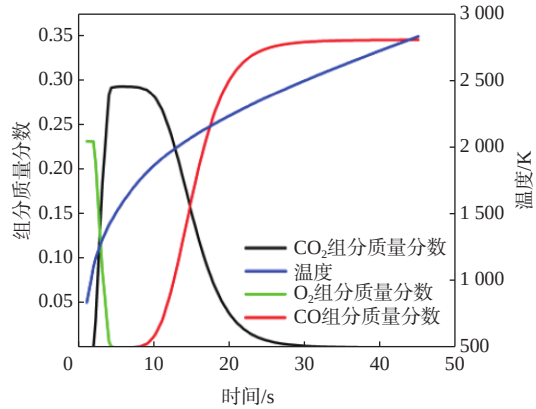


图 4　状态 3 碳-氧反应相关组分质量分数曲线及表面温升曲线

Fig. 4　Concentration curve and surface temperature rise curve of carbon oxygen reaction related components in state 3

态 3 为低热流，材料表面的温升更慢，因此，反应速率上升相较状态 1 也更慢。状态 1 反应后期温度达到 3 000 K 以上，发生碳-氮反应，致使烧蚀速率再次上升。

2　烧蚀传热问题的不确定性分析方法

由第 1 节分析可知，烧蚀量与温度有着复杂的关系，烧蚀量的变化不仅与温度有关，还受来流条件的影响。温度的变化受所用材料物性参数的影响。飞行器飞行过程中，来流条件频繁变化及材料物性由于工艺问题存在误差，因此，对于表面烧蚀传热问题，需要进行大量计算研究，进行敏感性分析，得到不同参数对输出变量的影响规律。

采用 PC 方法作为烧蚀传热问题的不确定性分析方法，其不确定性传播分析流程如图 5 所示。

通过与大量 MC 确定性计算进行对比，验证 PC 代理模型精度，本文不确定性输入参数的分布类型均为均匀分布。

PC 方法通过多项式基组合成随机空间，描述和传播随机变量的不确定性。响应变量 $Q(\boldsymbol{\xi})$，即不确定性分析中输出参数表示为

$$Q(\boldsymbol{\xi}) = q_0 I_0 + \sum_{i_1=1}^{\infty} q_{i_1} I_1(\xi_{i_1}) + \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1} q_{i_1 i_2} I_2(\xi_{i_1},\xi_{i_2}) + \sum_{i_1=1}^{\infty}\sum_{i_2=1}^{i_1}\sum_{i_3=1}^{i_2} q_{i_1 i_2 i_3} I_3(\xi_{i_1},\xi_{i_2},\xi_{i_3}) + \cdots \tag{31}$$

式中：$\boldsymbol{\xi}$为随机向量，即不确定性输入参数；I_0 和 q_0 分别为多项式混沌展开中的基准多项式和常数项系数；$I_m(\xi_{i_1},\xi_{i_2},\cdots,\xi_{i_m})$为 m 阶多元正交多项式；$q_{i_1 i_2 \cdots i_p}$为未知待求系数。由于阶数过大会造成计算量的增加，式(31)中的 PC 模型通常在 P 阶被截

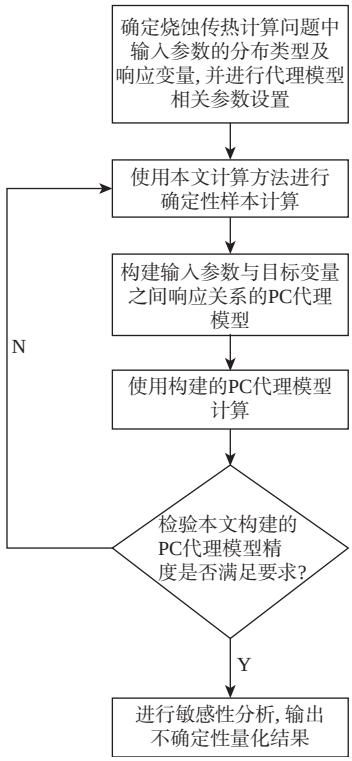


图5 不确定性传播分析流程

Fig. 5 Uncertainty propagation analysis flowchart

断。对应 P 阶 PC 模型表示为

$$Q(\boldsymbol{\xi}) \approx \sum_{i=0}^{P} b_i \Phi_i(\boldsymbol{\xi}) \tag{32}$$

式中：b_i和$\Phi_i(\boldsymbol{\xi})$对应式(31)中的$q_{i_1 i_2 \cdots i_p}$和$I_m(\xi_{i_1},\xi_{i_2},\cdots,\xi_{i_m})$。$\Phi_i(\boldsymbol{\xi})$的形式根据输入参数的概率分布类型确定，目前较为常用的是 Askey 方案[17]。关键在于 PC 待求系数b_i的求解。

采用 Galerkin 投影法进行 PC 系数求解，将式(32)两边依次投影到各正交多项式$\Phi_i(\boldsymbol{\xi})$上：

$$< Q(\boldsymbol{\xi})\Phi_i(\boldsymbol{\xi}) > = < \sum_{i=0}^{P} b_i \Phi_i(\boldsymbol{\xi})\Phi_i(\boldsymbol{\xi}) > \tag{33}$$

根据内积定义，整理得到

$$b_i = E[Q(\boldsymbol{\xi})\Phi_i(\boldsymbol{\xi})] / E\left[\Phi_i(\boldsymbol{\xi})\Phi_i(\boldsymbol{\xi})\right] \tag{34}$$

式中：“$<\cdot>$”表示内积；$E[\cdot]$表示期望。

待求 PC 系数可通过式 (34) 依次得到。式 (34) 中分母的求解通过对正交多项式求期望，分子的求解需要调用一定次数的输出响应函数，本文通过数值方法进行求解。

3 正交试验设计

正交试验设计是用来研究多因素多水平的试验方法，从全部试验中挑选出具有代表性的样本点进行试验，从而分析不同因素对输出变量的影响程度。基于此，本文选取几组具有代表性的组合，使用正交试验进行计算极差结果分析，初步对烧蚀传热问题中输入变量对目标变量的影响进行分析。

使用标准正交设计表计算第 2 节中 2 种典型状态下的结果，如表 2 所示，表中，110% 表示原有状态数值的 110%。选取一维表面烧蚀算例，长方体(1 mm×1 mm×40 mm, 100 个网格)的表面烧蚀后退量和背面(40 mm 处)温度变化作为目标变量，选取状态 1 和状态 3 这 2 种典型状态进行计算。

表 2 正交试验结果

Table 2 Orthogonal test results

分组	热导率水平/%	比热容水平/%	密度水平/%	热流水平/%	烧蚀量/mm		背面温度/K	
					状态1	状态3	状态1	状态3
1	110	110	110	110	−3.064 4	−6.368 7	910.949 7	1 429.159 3
2	110	120	120	120	−3.097 5	−6.309 6	807.627 7	1 293.738 2
3	110	130	130	130	−2.998 4	−6.255 8	725.752 9	1 174.807 5
4	120	110	120	130	−3.556 0	−7.788 1	920.596 1	1 457.903 8
5	120	120	130	110	−2.839 5	−5.188 8	790.248 6	1 241.390 9
6	120	130	110	120	−2.812 7	−5.734 4	858.475 4	1 357.617 1
7	130	110	130	120	−3.352 1	−6.612 7	895.940 7	1 395.975 0
8	130	120	110	130	−3.300 6	−7.054 7	974.781 1	1 515.508 0
9	130	130	120	110	−2.696 7	−4.700 4	829.174 3	1 287.565 8

表 3 为烧蚀量的极差分析结果。可知，在状态 1 下，比热容对烧蚀量的影响最大，其次为热流，热导率和密度的影响较小，也比较接近；在状态 3 下，对烧蚀量的影响次序分别为热流、比热容、密度、热导率。

表 4 给出在 2 个典型状态下背面温度的极差分析结果。可以看到，在状态 1 下，密度、比热容、热导率和热流对背面温度的影响依次降低；在状态 3 下，密度对背面温度的影响最大，其次为比热容、热导率，热流的影响最小。

表 3　烧蚀量的极差分析结果

Table 3　Range analysis results of ablation amount

状态	因素	k_1/mm	k_2/mm	k_3/mm	极差/mm
状态1	热导率	−3.053	−3.069	−3.116	0.063
	比热容	−3.324	−3.079	−2.836	0.488
	密度	−3.059	−3.117	−3.063	0.058
	热流	−2.867	−3.087	−3.285	0.418
状态3	热导率	−6.311	−6.237	−6.064	0.189
	比热容	−6.923	−6.184	−5.505	1.360
	密度	−6.386	−6.208	−6.019	0.367
	热流	−5.361	−6.219	−7.033	1.614

注: k_1、k_2、k_3分别为因素在110%、120%、130%水平下对结果的影响。

表 4　背面温度的极差分析结果

Table 4　Range analysis results of back temperature

状态	因素	k_1/K	k_2/K	k_3/K	极差/K
状态1	热导率	814.777	856.44	899.965	85.188
	比热容	909.162	857.552	804.468	104.694
	密度	914.735	852.466	803.981	110.754
	热流	843.458	854.015	873.71	30.252
状态3	热导率	1 299.235	1 352.304	1 375.694	100.448
	比热容	1 427.679	1 350.212	1 249.341	154.349
	密度	1 434.095	1 322.414	1 270.724	163.370
	热流	1 295.383	1 349.11	1 382.74	63.368

4　不确定性方法量化分析

由于上述正交试验未考虑因素之间相互影响, 所选样本点仅是个别因素水平的组合, 并未考虑所有可能因素水平下的可能性。为获得更加精准的敏感性分析, 采用 MC 方法和 PC 代理模型对上述算例进行烧蚀传热问题的不确定性量化分析, 表 2 中的 4 种实验因素作为不确定性输入参数, 输入参数变化范围如表 5 所示。

表 5　输入随机变量及其变化范围

Table 5　Input random variables and their variation range

输入参数	不确定度/%
热导率	±20
比热容	±10
密度	±10
热流	±30

对表 5 中输入参数构成的随机空间进行 MC 抽样计算。由于 MC 方法非常依赖样本空间大小, 使用 MC 方法分别计算 1 000 个、10 000 个和 100 000 个样本点。由此得到 100 000 次 MC 计算目标变量的最值, 如表 6 所示。

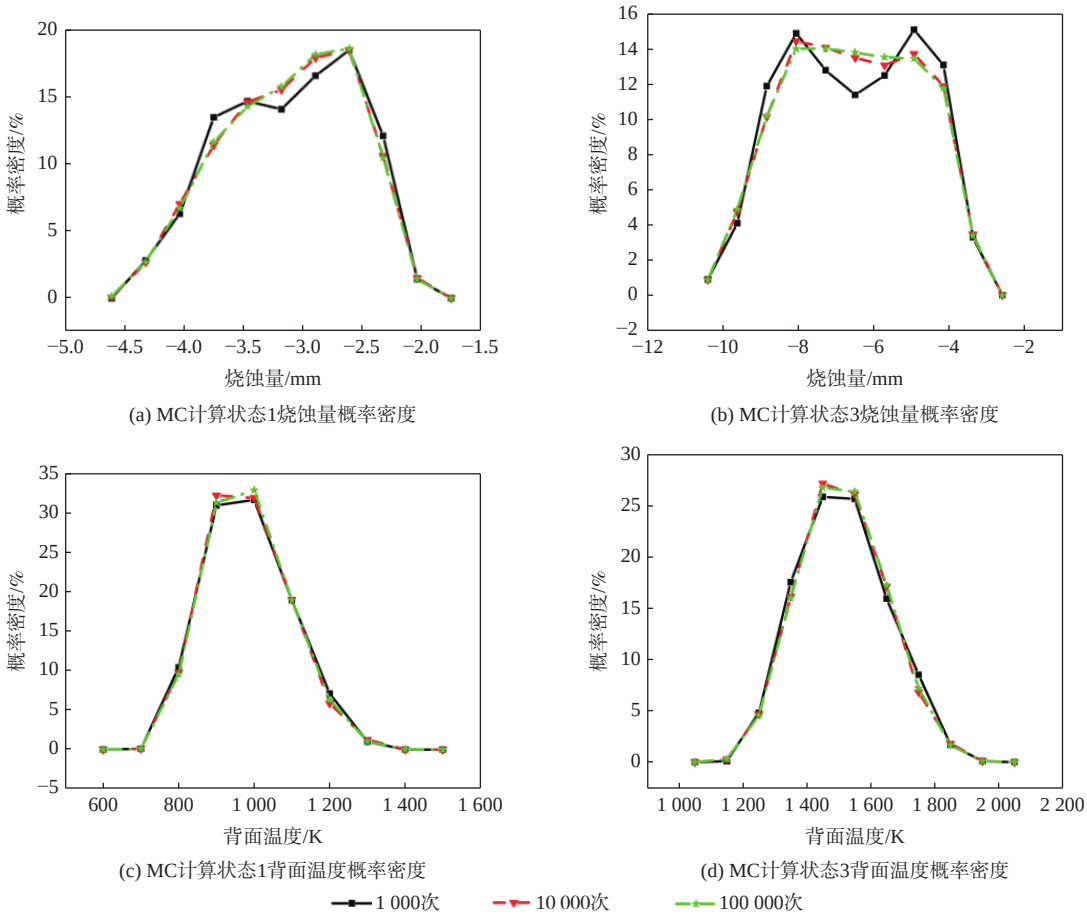
结合图 3 和图 4 可知, 尽管状态 1 的表面温度高于状态 3, 但通过表 6 可以看到, “低热流、长时间”情况下的烧蚀情况比“高热流、低时间”要严重, 在状态 3 条件下, 最大烧蚀量达到 10.73 mm, 背面温度达到 1 986.703 22 K。由此, 在长时间的烧蚀条件下, 即使飞行器遭受的气动加热热流不高, 也需要较厚的防热层, 防止被烧穿及内部热传导致防热层内壁温度过高。

表 6　100 000 次 MC 计算目标最值

Table 6　100 000 MC calculations target most value

状态	烧蚀量/mm		背面温度/K	
	最小值	最大值	最小值	最大值
状态1	−4.58	−2	716.342 7	1 404.276 78
状态3	−10.73	−2.97	1 120.521 57	1 986.703 22

进一步, 为获得 2 种典型状态下目标变量的分布情况, 计算目标变量的概率密度曲线, 如图 6 所示。可以看到, 背面温度的概率密度分布曲线重合度很高。从烧蚀量的概率密度分布来看, 1 000 次 MC 计算出的烧蚀量概率密度与 10 000 次和 100 000 次计算出的曲线存在差距, 其中, 状态 3 下差距尤为明显。总的来说, 2 种状态下 MC 计算 10 000 次的结果与 100 000 次的结果差距极小, 可以认为, 10 000 次样本数据能够较好反映烧蚀传热问题。



(a) MC计算状态1烧蚀量概率密度

(b) MC计算状态3烧蚀量概率密度

(c) MC计算状态1背面温度概率密度

(d) MC计算状态3背面温度概率密度

图 6 2种状态下不同MC计算次数下目标变量的概率密度

Fig. 6 Probability density of target variable under different MC calculation times in two states

由于MC方法计算成本较大，选取256个样本点训练PC方法构建目标变量与输入参数之间的响应关系，以分析目标变量的统计特性。同时，为验证PC方法的精度，将10 000次MC计算所得的样本点与256次样本数据训练出的PC代理模型计算10 000次结果的统计结果对比。

从表7统计结果来看，2个典型状态下，目标变量均值和标准差的差距均不大。烧蚀量均值误差为0，背面温度的均值最大误差为0.021%左右。说明PC方法能够较为精准地刻画出目标变量的统计特性。为验证PC方法的精度，在10 000个样本点上进行评估，参考文献[12]定义的决定系数R^2来评估代理模型的准确性：

$$R^2=1-\frac{\sum_{k=0}^{n_{\text{sample}}-1}(y_k-\hat{y}_k)^2}{\sum_{k=0}^{n_{\text{sample}}-1}(y_k-\bar{y})^2} \tag{35}$$

式中：y_k为MC计算的确定性样本值；$\bar{y}$为确定性样本均值；$\hat{y}_k$为PC计算的目标值；n_{sample}为样本数量。该算例中，状态1下烧蚀量的R^2=0.986 914 129，状态1下背面温度的R^2=0.983 278 774，状态3下烧蚀量的R^2=0.997 049 814，状态3下背面温度的R^2=0.989 142 669。说明本文构建的PC方法能够很好地预测样本点的目标变量。

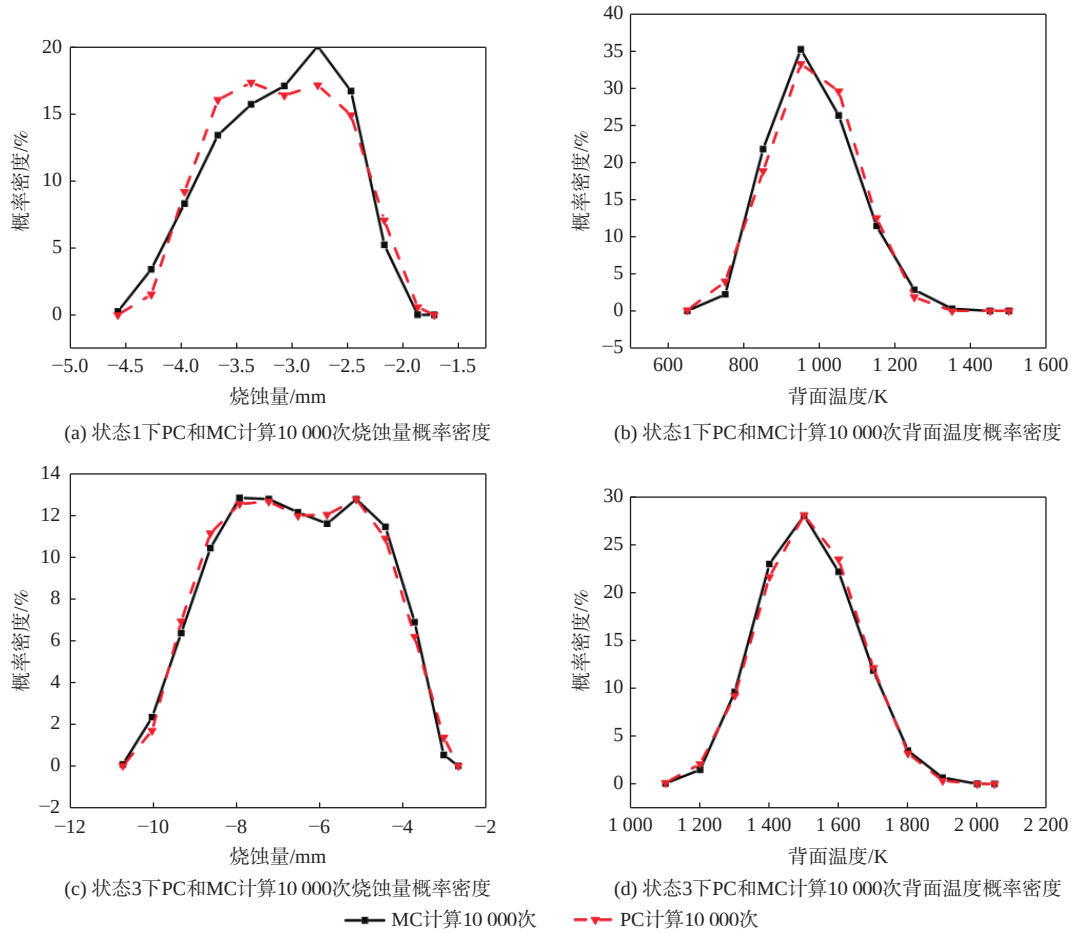
表7 目标变量统计矩

Table 7 Target variable statistical moment

方法	烧蚀量均值/mm	烧蚀量标准差/mm	背面温度均值/K	背面温度标准差/K
MC计算状态1	−3.11	0.545 479	982.902 6	106.781 9
PC计算状态1	−3.11	0.543 199	983.108 55	105.953 02
MC计算状态3	−6.56	1.72	1 512.022 63	133.025 74
PC计算状态3	−6.56	1.71	1 512.319 66	132.358 31

进一步，将PC代理模型与MC方法均计算10 000次的目标变量的概率密度进行对比，如图7所示。可以看到，除状态1下的烧蚀量概率密度曲线相差较大外，其他数据曲线吻合较好。状态1的拟合差值较大，但在数值上仅有4%左右的差距。敏感性分析作为不确定性量化中重要的一步，研究随机输入变量对目标变量的贡献度。采用Sobol指标来分析每个输入变量对输出结果的影响大小，结果如表8所示。

由表8可以看到，2种典型状态下，热流对烧蚀



(a) 状态1下PC和MC计算10 000次烧蚀量概率密度　(b) 状态1下PC和MC计算10 000次背面温度概率密度

(c) 状态3下PC和MC计算10 000次烧蚀量概率密度　(d) 状态3下PC和MC计算10 000次背面温度概率密度

图 7　2 种状态下 PC 和 MC 计算 10 000 次目标变量概率密度

Fig. 7　PC and MC calculate the probability density of target variables 10 000 times in two states

表 8　目标变量 Sobol 指标

Table 8　Target variable Sobol index

因素	烧蚀量Sobol指标		背面温度Sobol指标	
	状态1	状态3	状态1	状态3
热导率	1.41×10^{-5}	0.005 923 237	0.512 097 066	0.300 210 835
密度	0.109 554 085	0.077 289 485	0.191 548 295	0.182 521 824
比热容	1.19×10^{-5}	0.003 586 252	0.209 484 497	0.205 598 757
热流	0.890 419 977	0.913 201 026	0.086 870 142	0.311 668 585

量的影响尤为显著, 密度、热导率和比热容对烧蚀量的影响依次降低。计算背面温度 Sobol 指标, 状态 1 条件下, 热导率对背面温度的影响最大, 其次是比热容、密度和热流。而在状态 3 条件下, 热流对背面温度的影响最大, 热导率相比热流, 影响较低, 随后为比热容和密度的影响。

5　结　论

1) 在烧蚀传热问题中, 烧蚀速率与温度之间的变化并不是简单的函数关系, 在 2 种典型加热状态(高热流、短时间和低热流、长时间)下, 烧蚀速率曲线呈现出比较复杂、有显著分段平台特征的变化规律。

2) 在 2 种典型加热状态下, 分别使用正交试验设计、MC 方法和 PC 方法对烧蚀传热问题的关键输入参数进行敏感性分析, 发现在典型加热状态下, 热流对烧蚀量的影响较为突出, 而背面温度方面, “高热流、短时间”状态下, 热导率的影响较大, 而“低热流、长时间”状态下, 热流的影响相对较大。

3) 对烧蚀传热问题进行不确定性量化计算, 传统的 MC 方法需要大量样本, 而 PC 方法不仅可以减少样本数据, 节约计算时间, 而且能得到较为精准的计算结果。

参考文献（References）

[1] 杨琼梁, 史晓鸣, 许斌, 等. 烧蚀防热层与结构耦合热传导分析的交替计算法[J]. 宇航学报, 2011, 32(8): 1854-1858.

YANG Q L, SHI X M, XU B, et al. Alternative algorithm for heat transfer analysis of ablative protection layer coupled with structure[J]. Journal of Astronautics, 2011, 32(8): 1854-1858(in Chinese).

[2] 孟松鹤, 疏浩, 易法军, 等. 烧蚀防热材料在线后退测试技术研究进展[J]. 宇航学报, 2023, 44(7): 957-972.
MENG S H, SHU H, YI F J, et al. Research progress of online recession test technology for ablative thermal protection materials[J]. Journal of Astronautics, 2023, 44(7): 957-972(in Chinese).

[3] 杨驰, 刘娜, 孔维萱, 等. 高超声速再入热环境下防热复合材料烧蚀传热计算及影响因素分析[J]. 玻璃钢/复合材料, 2017(4): 35-41.
YANG C, LIU N, KONG W X, et al. Study on ablation heat transfer calculation and influencing factors of thermal protection materials under hypersonic thermal environment[J]. Fiber Reinforced Plastics/Composites, 2017(4): 35-41(in Chinese).

[4] 高铁锁, 江涛, 丁明松, 等. 再入体碳基防热材料烧蚀流场红外辐射模拟[J]. 红外与激光工程, 2023, 52(5): 20220606.
GAO T S, JIANG T, DING M S, et al. Simulation of flow field infrared radiation over reentry vehicle with ablation of carbon-based thermal protection material[J]. Infrared and Laser Engineering, 2023, 52(5): 20220606(in Chinese).

[5] 周印佳, 张志贤, 付新卫, 等. 再入飞行器烧蚀热防护一体化计算方法[J]. 航空学报, 2021, 42(7): 202-211.
ZHOU Y J, ZHANG Z X, FU X W, et al. Integrated computing method for ablative thermal protection system of reentry vehicles[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(7): 202-211 (in Chinese).

[6] 周述光, 国义军, 贺立新, 等. 再入弹头三维非对称烧蚀外形模拟[J]. 航空学报, 2017, 38(12): 114-125.
ZHOU S G, GUO Y J, HE L X, et al. Simulation of 3D asymmetric ablation shape of reentry missile[J]. Acta Aeronautica et Astronautica Sinica, 2017, 38(12): 114-125(in Chinese).

[7] DEC J A, MITCHELTREE R A. Probabilistic design of a Mars Sample Return Earth entry vehicle thermal protection system[C]//Proceedings of the 40th AIAA Aerospace Sciences Meeting & Exhibit. Reston: AIAA, 2002: 910.

[8] CHEN Y K, SQUIRE T, LAUB B, et al. Monte Carlo analysis for spacecraft thermal protection system design[C]//Proceedings of the 9th AIAA/ASME Joint Thermophysics and Heat Transfer Conference. Reston: AIAA, 2006: 2951.

[9] SEPKA S, WRIGHT M. A Monte Carlo approach to fiat uncertainties-improvements and results for MSL[C]//Proceedings of the 41st AIAA Thermophysics Conference. Reston: AIAA, 2009: 4234.

[10] DEC J A, BRAUN R D. Three-dimensional finite element ablative thermal response and design of thermal protection systems[J]. Journal of Spacecraft and Rockets, 2013, 50(4): 725-734.

[11] TOBIN S A, DEC J A. Probabilistic design demonstration of a flexible thermal protection system for a hypersonic inflatable aerodynamic decelerator[C]//Proceedings of the 53rd AIAA Aerospace Sciences Meeting. Reston: AIAA, 2015: 1895.

[12] 陈江涛, 章超, 刘骁, 等. 基于稀疏多项式混沌方法的不确定性量化分析[J]. 航空学报, 2020, 41(3): 123382.
CHEN J T, ZHANG C, LIU X, et al. Uncertainty quantification analysis with sparse polynomial chaos method[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(3): 123382(in Chinese).

[13] 国义军, 代光月, 桂业伟, 等. 碳基材料氧化烧蚀的双平台理论和反应控制机理[J]. 空气动力学学报, 2014, 32(6): 755-760.
GUO Y J, DAI G Y, GUI Y W, et al. A dual platform theory for carbon-based material oxidation with reaction-diffusion rate controlled kinetics[J]. Acta Aerodynamica Sinica, 2014, 32(6): 755-760 (in Chinese).

[14] SCALA S M. The ablation of graphite in dissociated air: part 1: theory: R62SD72[R]. Philadelphia: General Electric, 1962.

[15] 刘骁. 烧蚀热防护系统三维热响应有限元计算方法研究[D]. 绵阳: 中国空气动力研究与发展中心, 2016.
LIU X. Study on three-dimensional finite element thermal response calculation method applied to ablative thermal protection system[D]. Mianyang: China Aerodynamics Research and Development Center, 2016(in Chinese).

[16] 王臣, 梁军, 吴世平, 等. 高温烧蚀条件下 C/C 材料热力耦合场模拟[J]. 复合材料学报, 2006, 23(5): 143-148.
WANG C, LIANG J, WU S P, et al. Numerical simulation of C/C composites coupled thermo-mechanical field under the condition of high temperatures and ablation[J]. Acta Materiae Compositae Sinica, 2006, 23(5): 143-148(in Chinese).

[17] ASKEY R A, WILSON J. Some basic hypergeometric orthogonal polynomials that generalize Jacobi polynomials[M]. Providence: American Mathematical Society, 1985.

Uncertainty analysis methods for heat transfer ablation in carbon-based materials

LIU Ting[1]，LIU Xiao[2,*]，GUO Lei[1,3]，ZENG Lei[2]，GUO Yijun[2]

(1. School of Computer Science and Technology，Xinjiang University，Urumqi 830000，China；
2. China Aerodynamics Research and Development Center，Mianyang 621000，China；
3. College of Computer Science and Engineering，University of Electronic Science and Technology of China，Chengdu 610000，China)

Abstract: A quantitative study on the uncertainty of ablation heat transfer is conducted, in which a dual platform model is used for ablation prediction, and the heat transfer process under the ablation dynamic boundary is numerically solved using the finite element method. Firstly, an orthogonal experimental study was conducted to analyze the influence of parameters on the target variable. It was discovered that while the input parameters had consistent impacts on the backside temperature, they had distinct effects on the ablation quantity under two common heating situations (high heat flow, short time and low heat flow, long time). In order to obtain more accurate uncertainty quantification results, Monte Carlo (MC) and polynomial chaos (PC) methods were further used to conduct uncertainty analysis on the ablation heat transfer problem. Through sensitivity analysis, it was found that under two heating conditions, heat flux was the key factor affecting the ablation amount. The back temperature is most affected by the material's thermal conductivity in state 1 (high heat flow, short time), but heat flow in state 3 (low heat flow, long time) has a comparatively bigger effect. Compared to the MC method, the PC method can effectively reduce computational costs and obtain satisfactory computational results.

Keywords: carbon-based materials; dual platform theory; uncertainty; polynomial chaos; orthogonal test

Received：2024-05-09；**Accepted**：2024-07-05；**Published Online**：2024-08-19 15：57

URL：link.cnki.net/urlid/11.2625.V.20240816.1613.002

Foundation item：National Key Research and Development Program of China (2019YFA0405202)

* **Corresponding author**. E-mail：liuxiao8146@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0286

变推力、变充液比、贮箱倾斜状态下液体晃动分析

何家伟[1,2]，吕敬[2,*]，王天舒[3]，刘颖杰[1]

(1. 湖北省航天技术研究院总体设计所，武汉 430048；2. 北京航空航天大学 航空科学与工程学院，北京 102206；
3. 清华大学 航天航空学院，北京 100083)

摘　　要：随着航天技术的发展，新型深空探测器在起飞时会遇到2种特殊工况：发动机推力矢量的大小和方向随时间变化；发射初期，重力方向和航天器主轴线不重合。同时，飞行过程中，液体燃料不断消耗。面对上述3种工况，针对航天器贮箱内液体晃动问题，提出解决方案：在动力学建模阶段，考虑航天器平动加速度的影响，分析推力不沿航天器轴线方向的液体晃动情况；借助变等效力学模型参数方法，计算变充液比工况下的液体晃动问题；通过设定等效力学模型中单摆存在一个初始摆角，模拟贮箱倾斜时液面与贮箱轴线不垂直的状态。以Cassini贮箱为例，在半充液状态，分别计算上述3种工况下的晃动力和晃动力矩。以某型号航天器液体贮箱为例，在实际发射条件下进行晃动分析，计算结果与商用软件Flow-3D计算结果吻合较好，验证了所提计算方法的可行性。

关 键 词：等效力学模型；液体晃动；变推力；变充液比；贮箱倾斜
中图分类号：V19
文献标志码：A　　**文章编号**：1001-5965(2026)06-2074-09

随着航天事业不断发展，要求不断提高火箭的运载能力、延长卫星在轨时间，同时，需要提升其机动性，航天器液体燃料的质量占自身的比重也在不断增加。航天器发射、升空及正常运行时，在航天器外部运动激励下，液体燃料可能发生剧烈晃动，由此产生的晃动力和晃动力矩会对航天器造成重要影响，可能导致航天器运行姿态失控或贮液结构损坏等严重事故[1-2]。在工程上，经常涉及贮液系统的动力学计算问题，人们更多关注贮箱内液体运动如何影响结构响应，工程师往往需要寻求一种液体晃动等效力学模型，将复杂的流体运动用简单的刚体运动来代替[3]。从20世纪50年代开始，相关学者就对等效力学模型方法进行研究，并形成了一大批文献资料[4-6]。

常用的等效力学模型有弹簧-质量模型和单摆模型[4]2种。Graham等[7]采取线性势流理论，提出了矩形贮箱内液体晃动弹簧-质量模型，Housner[8]通过对Graham等[7]提出的模型进行近似计算，得到了液体晃动等效力学模型的简化计算公式，使其更利于实际操作。针对非矩形贮箱，Rattayya[9]提出了椭球形贮箱内液体晃动的弹簧-质量模型。同时，单摆模型也得到了长足发展。Mikishev[10]提出了在一维平动或转动激励下充液耦合系统的平面单摆模型。包光伟[11]通过深化Mikishev[10]的研究，将平面问题推广到空间问题，建立了平放柱形贮箱在三维晃动条件下的等效单摆模型。原则上，2种等效力学模型是等价的，但当重力加速度变化时，等效弹簧刚度会发生变化，而等效单摆摆长保持常数，因此，本文选择单摆模型。

在实际生产中，贮箱形状并不是标准统一的，对于复杂形状的贮箱，很难通过理论分析得到单摆模型的参数，虽然可以通过实验方法测量[12]，但实验成本高，测量难度较大。因此，一般采用有限元方法。包光伟等[13]利用有限元方法建立了任意形状贮箱内液体晃动固有频率和模态的计算方法。李青等[14]针对非对称贮箱的液体晃动问题，采用有

收稿日期：2024-05-06；录用日期：2024-08-09；网络出版时间：2024-09-27 12:10
网络出版地址：link.cnki.net/urlid/11.2625.V.20240926.1748.003
*通信作者. E-mail：lvjing@buaa.edu.cn

引用格式：何家伟，吕敬，王天舒，等. 变推力、变充液比、贮箱倾斜状态下液体晃动分析[J]. 北京航空航天大学学报，2026，52(6)：2074-2082. HE J W，LYU J，WANG T S，et al. Analysis of liquid sloshing under variable thrust, variable filling liquid ratio, and tilted tank state[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：2074-2082 (in Chinese).

限元方法计算液体晃动等效模型参数。在航天领域中, 借助等效力学模型进行晃动耦合分析, 一般以火箭推力产生的惯性力作为等效重力, 在贮箱本体坐标系中, 推力大小和方向恒定, 此外, 通常会以液面与贮箱轴线垂直状态作为贮箱内液体的初始姿态, 现代航天作业面临更加复杂的推力情况, 且航天器有时在倾斜状态下起飞, 因此, 有必要设计适用于复杂工况下的晃动分析方法。

本文采用等效力学模型法计算航天器液体贮箱内的晃动耦合作用, 将航天器的飞行姿态作为输入, 计算贮箱内液体对贮箱壁产生的作用力和作用力矩, 即晃动力和晃动力矩。航天器(特别是深空探测器)的发动机需要随时调整推力大小和方向; 同时, 贮箱内液体燃料一直处于消耗状态; 另外, 在着陆过程中, 不能保证探测器以垂直姿态降落到行星表面上, 因此, 重新起飞时, 探测器贮箱处于倾斜状态。面对上述 3 种工况, 本文提出一种解决方案, 并与商用软件 Flow-3D 计算结果比较, 以验证本文提出的液体晃动分析方法的正确性。

1　建模方法

1.1　晃动动力学方程

三维液体晃动等效力学模型将贮箱内的液体分为 2 部分: 相对于贮箱晃动的液体用 n 个 (n 为模态数) 单摆模型表示; 相对贮箱静止的液体部分用 1 个静止质量块表示。惯性坐标系下, 在贮箱平动和转动运动的外部激励作用下, 计算液体在晃动过程中对贮箱产生的作用力和作用力矩。一般情况下, 取等效力学模型的前 2 阶有效单摆进行计算, 即 n=2。充液航天器三维运动等效系统如图 1 所示。其中, O_0xyz为惯性坐标系, $O_{\mathrm{b}}xyz$为贮箱本体坐标系。

结合图 1, 对借助等效力学模型建立液体晃动

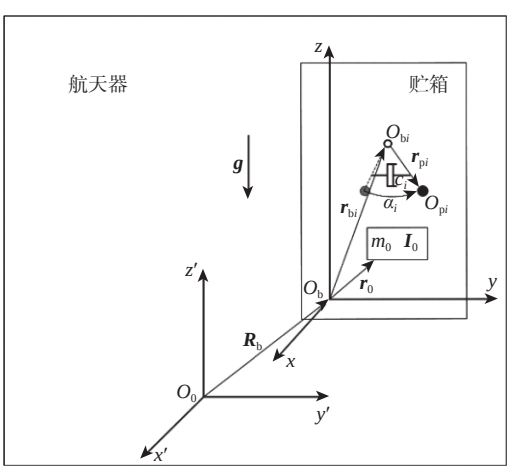


图 1　充液航天器三维运动等效系统

Fig. 1　Three-dimensional motion equivalent system for liquid filled spacecraft

动力学方程过程中所使用的各种物理量作简要说明, 其大体分为 3 类: ①液体晃动等效力学模型所包含的物理量: 重力加速度矢量$\boldsymbol{g}$、第 i 阶等效单摆质量m_i、第 i 阶等效单摆悬挂点$O_{\mathrm{b}i}$、第 i 阶等效单摆悬挂点相对于贮箱本体坐标系原点O_{b}的矢径$\boldsymbol{r}_{\mathrm{b}i}$、静止质量块质量$m_0$、静止质量块相对于质心的惯量张量$\boldsymbol{I}_0$、静止质量块质心相对于贮箱本体坐标系原点$O_{\mathrm{b}}$的矢径$\boldsymbol{r}_0$; ②以等效力学模型中的数据作为中间量推导得出的物理量: 第 i 阶等效单摆质心相对于自身悬挂点$O_{\mathrm{b}i}$的矢径$\boldsymbol{r}_{\mathrm{p}i}$、第 i 阶等效单摆质心在贮箱本体坐标系中的位置$O_{\mathrm{p}i}$、第 i 阶等效单摆摆动过程中受到的阻尼c_i、参数阻尼比ζ_i; ③贮箱受到外部激励产生的物理量: 贮箱本体坐标系原点O_{b}相对于惯性坐标系原点O_0的矢径$\boldsymbol{R}_{\mathrm{b}}$、贮箱本体坐标系相对于惯性坐标系的转动角速度$\boldsymbol{\omega}_{\mathrm{b}}$、第 i 阶等效单摆相对于贮箱本体坐标系的摆动角速度$\boldsymbol{\omega}_{\mathrm{p}i}$、第 i 阶等效单摆摆动角度α_i。

对贮箱内的液体整体进行受力分析。其受到重力$\boldsymbol{G}$、贮箱对液体产生的作用力$\boldsymbol{F}$和作用力矩$\boldsymbol{M}$、液体晃动过程中产生的阻尼力$\boldsymbol{F}_{\mathrm{d}}$、贮箱运动产生的惯性力$\boldsymbol{F}_{\mathrm{I}}$。根据牛顿第三定理, $\boldsymbol{F}$和$\boldsymbol{M}$的反作用力$-\boldsymbol{F}$和$-\boldsymbol{M}$即为液体对贮箱的晃动力和晃动力矩。将贮箱内液体在$\boldsymbol{F}_{\mathrm{I}}$、$\boldsymbol{G}$、$\boldsymbol{F}$和$\boldsymbol{F}_{\mathrm{d}}$的作用下产生的虚功率分别称为惯性力虚功率、重力虚功率、主动力虚功率和阻尼力虚功率, 分别表示为δP^{I}、δP^{g}、δP^{a}和δP^{d}。

根据虚功率原理有

$$\sum_{i=1}^{n}\left(\delta P_i^{\mathrm{I}}+\delta P_i^{\mathrm{g}}+\delta P_i^{\mathrm{d}}\right)+\delta P_0^{\mathrm{I}}+\delta P_0^{\mathrm{g}}+\delta P^{\mathrm{a}}=0 \tag{1}$$

式中: δP_i^{I}、δP_i^{g}、δP_i^{d}分别为第 i 阶等效单摆的惯性力虚功率、重力虚功率、阻尼力虚功率; δP_0^{I}、δP_0^{g}分别为静止质量块的惯性力虚功率、重力虚功率。

$$\delta P_i^{\mathrm{g}}=m_i\delta\dot{\boldsymbol{R}}_i\boldsymbol{g}=\delta\dot{\boldsymbol{R}}_i m_i\boldsymbol{g}+\delta\boldsymbol{\omega}_{\mathrm{b}}(\boldsymbol{S}_i\times\boldsymbol{g})+\delta\boldsymbol{\omega}_{\mathrm{p}i}\left(\boldsymbol{S}_{\mathrm{p}i}\times\boldsymbol{g}\right) \tag{2}$$

$$\begin{aligned}\delta P_i^{\mathrm{I}}=&-\delta\dot{\boldsymbol{R}}_{\mathrm{b}}\left(m_i\ddot{\boldsymbol{R}}_i+\dot{\boldsymbol{\omega}}_{\mathrm{b}}\times\boldsymbol{S}_i+\dot{\boldsymbol{\omega}}_{\mathrm{p}i}\times\boldsymbol{S}_{\mathrm{p}i}+\boldsymbol{Q}_{ti}\right)-\\&\delta\boldsymbol{\omega}_{\mathrm{b}}\left[\boldsymbol{S}_i\times\ddot{\boldsymbol{R}}_{\mathrm{b}}+\boldsymbol{J}_i\dot{\boldsymbol{\omega}}_{\mathrm{b}}+\boldsymbol{r}_{\mathrm{b}i}\times\left(\dot{\boldsymbol{\omega}}_{\mathrm{p}i}\times\boldsymbol{S}_{\mathrm{p}i}\right)+\boldsymbol{J}_i\dot{\boldsymbol{\omega}}_{\mathrm{p}i}+\boldsymbol{Q}_{\mathrm{r}i}\right]-\\&\delta\boldsymbol{\omega}_{\mathrm{p}i}\left[\boldsymbol{S}_{\mathrm{p}i}\times\ddot{\boldsymbol{R}}_{\mathrm{b}}+\boldsymbol{S}_{\mathrm{p}i}\times(\dot{\boldsymbol{\omega}}_{\mathrm{b}}\times\boldsymbol{r}_{\mathrm{b}i})+\boldsymbol{J}_{\mathrm{p}i}\left(\dot{\boldsymbol{\omega}}_{\mathrm{b}}+\dot{\boldsymbol{\omega}}_{\mathrm{p}i}\right)+\boldsymbol{Q}_{ii}\right]\end{aligned} \tag{3}$$

$$\delta P_i^{\mathrm{d}}=-\delta\boldsymbol{\omega}_{\mathrm{p}i}\left(2m_i\zeta_i\omega_i l_i^{\,2}\boldsymbol{p}_i\boldsymbol{p}_i\boldsymbol{\omega}_{\mathrm{p}i}\right) \tag{4}$$

$$\begin{aligned}\boldsymbol{Q}_{ti}=&\left(\boldsymbol{\omega}_{\mathrm{b}}\times\boldsymbol{\omega}_{\mathrm{p}i}\right)\times\boldsymbol{S}_{pi}+m_i\boldsymbol{\omega}_{\mathrm{b}}\times(\boldsymbol{\omega}_{\mathrm{b}}\times\boldsymbol{r}_{\mathrm{b}i})+\\&\left(\boldsymbol{\omega}_{\mathrm{b}}+\boldsymbol{\omega}_{\mathrm{p}i}\right)\times\left[\left(\boldsymbol{\omega}_{\mathrm{b}}+\boldsymbol{\omega}_{\mathrm{p}i}\right)\times\boldsymbol{S}_{\mathrm{p}i}\right]\end{aligned} \tag{5}$$

$$\begin{aligned}\boldsymbol{Q}_{\mathrm{r}i}=&\boldsymbol{r}_{\mathrm{b}i}\times\left[\left(\boldsymbol{\omega}_{\mathrm{b}}\times\boldsymbol{\omega}_{\mathrm{p}i}\right)\times\boldsymbol{S}_{\mathrm{p}i}\right]+\\&\boldsymbol{J}_{\mathrm{p}i}\left(\boldsymbol{\omega}_{\mathrm{b}}\times\boldsymbol{\omega}_{\mathrm{p}i}\right)+\boldsymbol{S}_i\times[\boldsymbol{\omega}_{\mathrm{b}}\times(\boldsymbol{\omega}_{\mathrm{b}}\times\boldsymbol{r}_{\mathrm{b}i})]+\\&\boldsymbol{r}_{\mathrm{b}i}\times\left\{\left(\boldsymbol{\omega}_{\mathrm{b}}+\boldsymbol{\omega}_{\mathrm{p}i}\right)\times\left[\left(\boldsymbol{\omega}_{\mathrm{b}}+\boldsymbol{\omega}_{\mathrm{p}i}\right)\times\boldsymbol{r}_{\mathrm{p}i}\right]\right\}+\\&\left(\boldsymbol{\omega}_{\mathrm{b}}+\boldsymbol{\omega}_{\mathrm{p}i}\right)\times\left[\boldsymbol{J}_{\mathrm{p}i}\left(\boldsymbol{\omega}_{\mathrm{b}}+\boldsymbol{\omega}_{\mathrm{p}i}\right)\right]\end{aligned} \tag{6}$$

$$\boldsymbol{Q}_{ii}=\boldsymbol{J}_{pi}\left(\boldsymbol{\omega}_{b}\times\boldsymbol{\omega}_{pi}\right)+\boldsymbol{S}_{pi}\times[\boldsymbol{\omega}_{b}\times(\boldsymbol{\omega}_{b}\times\boldsymbol{r}_{bi})]+\left(\boldsymbol{\omega}_{b}+\boldsymbol{\omega}_{pi}\right)\times\left[\boldsymbol{J}_{pi}\left(\boldsymbol{\omega}_{b}+\boldsymbol{\omega}_{pi}\right)\right] \tag{7}$$

式中：$\boldsymbol{R}_i$为第i阶晃动质量m_i在惯性坐标系下的矢径，$\boldsymbol{R}_i=\boldsymbol{R}_b+\boldsymbol{r}_{bi}+\boldsymbol{r}_{pi}$；$\boldsymbol{p}_i$为第$i$阶晃动质量$m_i$在贮箱本体坐标系下的晃动方向(晃动平面的法向)；$\boldsymbol{S}_i$和$\boldsymbol{S}_{pi}$分别为第$i$阶晃动质量$m_i$相对于贮箱本体坐标系原点$O_b$和对应单摆悬挂点$O_{bi}$的静矩；$\boldsymbol{J}_i$和$\boldsymbol{J}_{pi}$分别为第$i$阶晃动质量相对于贮箱本体坐标系原点$O_b$和单摆悬挂点$O_{bi}$的惯量张量；$\omega_i$和$l_i$分别为第$i$阶等效单摆的摆动频率和摆长。

$$\delta P_0^{g}=\delta\dot{\boldsymbol{R}}_b m_0\boldsymbol{g}+\delta\boldsymbol{\omega}_b(\boldsymbol{S}_0\times\boldsymbol{g}) \tag{8}$$

$$\delta P_0^{I}=-\delta\dot{\boldsymbol{R}}_b\left[m_0\ddot{\boldsymbol{R}}_b+\dot{\boldsymbol{\omega}}_b\times\boldsymbol{S}_0+\boldsymbol{\omega}_b\times(\boldsymbol{\omega}_b\times\boldsymbol{S}_0)\right]-\delta\boldsymbol{\omega}_b\left[\boldsymbol{S}_0\times\ddot{\boldsymbol{R}}_b+\boldsymbol{J}_0\dot{\boldsymbol{\omega}}_b+\boldsymbol{\omega}_b\times(\boldsymbol{J}_0\boldsymbol{\omega}_b)\right] \tag{9}$$

式中：$\boldsymbol{S}_0$和$\boldsymbol{J}_0$分别为静止质量块相对于贮箱本体坐标系原点O_b的静矩和惯量张量。

$$\delta P^{a}=\delta\dot{\boldsymbol{R}}_b\boldsymbol{F}+\delta\boldsymbol{\omega}_b\boldsymbol{M} \tag{10}$$

将式(2)～式(10)代入式(1)，利用变分$\delta\dot{\boldsymbol{R}}_b$、$\delta\boldsymbol{\omega}_b$、$\delta\boldsymbol{\omega}_{pi}$的独立性分别得到系统的平动方程、转动方程和第$i$阶等效模态的晃动方程：

$$\left(\sum_{i=1}^{n}m_i+m_0\right)\ddot{\boldsymbol{R}}_b+\dot{\boldsymbol{\omega}}_b\times\left(\sum_{i=1}^{n}\boldsymbol{S}_i+\boldsymbol{S}_0\right)+\sum_{i=1}^{n}\left(\dot{\boldsymbol{\omega}}_{pi}\times\boldsymbol{S}_{pi}\right)+\sum_{i=1}^{n}\boldsymbol{Q}_{ti}+\boldsymbol{\omega}_b\times(\boldsymbol{\omega}_b\times\boldsymbol{S}_0)=\left(\sum_{i=1}^{n}m_i+m_0\right)\boldsymbol{g}+\boldsymbol{F} \tag{11}$$

$$\left(\sum_{i=1}^{n}\boldsymbol{S}_i+\boldsymbol{S}_0\right)\times\ddot{\boldsymbol{R}}_b+\left(\sum_{i=1}^{n}\boldsymbol{J}_i+\boldsymbol{J}_0\right)\dot{\boldsymbol{\omega}}_b+\sum_{i=1}^{n}\left[\boldsymbol{r}_{bi}\times\left(\dot{\boldsymbol{\omega}}_{pi}\times\boldsymbol{S}_{pi}\right)+\boldsymbol{J}_{pi}\dot{\boldsymbol{\omega}}_{pi}\right]+\sum_{i=1}^{n}\boldsymbol{Q}_{ti}+\boldsymbol{\omega}_b\times(\boldsymbol{J}_0\boldsymbol{\omega}_b)=\left(\sum_{i=1}^{n}\boldsymbol{S}_i+\boldsymbol{S}_0\right)\times\boldsymbol{g}+\boldsymbol{M} \tag{12}$$

$$\boldsymbol{S}_{pi}\times\ddot{\boldsymbol{R}}_b+\boldsymbol{S}_{pi}\times(\dot{\boldsymbol{\omega}}_b\times\boldsymbol{r}_{bi})+\boldsymbol{J}_{pi}\left(\dot{\boldsymbol{\omega}}_b+\dot{\boldsymbol{\omega}}_{pi}\right)+\boldsymbol{Q}_{ii}+2m_i\zeta_i\omega_i l_i^2\boldsymbol{p}_i\boldsymbol{p}_i\boldsymbol{\omega}_{pi}=\boldsymbol{S}_{pi}\times\boldsymbol{g}\qquad i=1,2,\cdots,n \tag{13}$$

1.2 变推力工况下晃动分析

航天器发射过程中，其受到重力和火箭推力作用，随着航天技术发展的要求，火箭推力大小和方向会随时间发生变化，如图2所示。由于推力使贮箱内液体受到惯性力作用，其在重力方向会产生大于零的分量，导致航天器实际受到的重力不断变化。在小幅度晃动条件下，贮箱内液体可以等效为单摆或弹簧-质量模型，一般情况下，2种模型是等价的，但当重力变化时，等效弹簧刚度会随之变化，此时，等效模型条件不成立，而等效单摆摆长保持常数，因此，航天领域一般选择单摆模型。

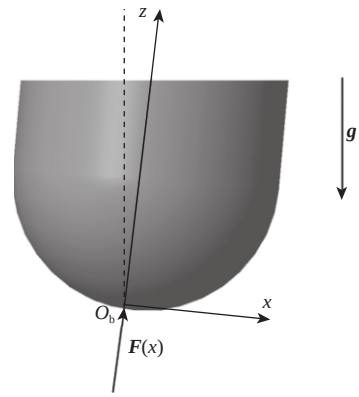


图2 变推力工况示意图

Fig. 2 Schematic diagram of variable thrust operating conditions

本文在推导液体晃动动力学方程过程中，对上述2部分作用力都有考虑：通过计算等效力学模型各部分的重力虚功率δP_i^{g}和δP_0^{g}，实现重力对液体晃动的影响；在计算等效力学模型中各部分的惯性力虚功率δP_i^{I}和δP_0^{I}时，式(3)和式(9)中引入贮箱平动加速度矢量$\ddot{\boldsymbol{R}}_b$，考虑了贮箱三维平动运动对液体晃动的影响。因此，本文推导的液体晃动动力学方程同样适用于火箭推力大小和方向不断变化的情况。

1.3 变充液比工况下晃动分析

为模拟充液比变化过程中的晃动现象，将液体晃动等效力学模型中的所有参数作为与时间相关的函数$f(x)$，随着充液比变化，等效力学模型各参数变化较为平缓，可以考虑使用线性插值法计算对应时间点的等效力学模型[15]，利用动力学方程式(11)～式(13)计算相应的晃动力和晃动力矩，具体实施步骤如图3所示。

为提高计算方法的运算效率和可操作性，使用

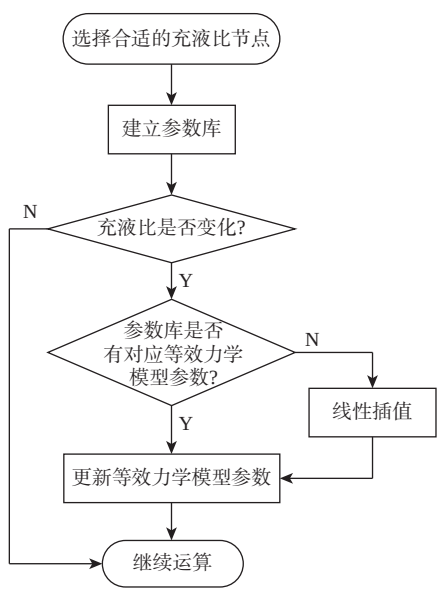


图3 变充液比工况晃动计算流程

Fig. 3 Flowchart for sloshing calculating under the variable filling liquid ratio operating condition

插值法计算贮箱内不同充液比所对应的液体等效力学模型数据时, 采用分段低次插值来保证算法的收敛性, 本文使用分段 3 次 Lagrange 插值。计算步骤如下:

1) 按照从大到小 (或从小到大) 的顺序选取 m+1 个不同充液比 $x_0, x_1, \cdots, x_m$ 作为插值节点;

2) 分别计算出插值节点对应的等效力学模型数据 $f(x_0), f(x_1), \cdots, f(x_m)$, 将其作为插值基函数;

3) 根据待计算充液比 $x(x_0 \leqslant x \leqslant x_m)$, 选择相邻的 3 个插值节点, 分别为 a_0、a_1、a_2, 计算对应的等效力学模型参数。

选择插值节点时, 将 x 与 m+1 个插值节点比较, 得到与 x 值最接近节点的下标 k。如果 k=0, 插值节点取 $a_0 = x_0, a_1 = x_1, a_2 = x_2$; 如果 k=m, 插值节点取 $a_0 = x_{m-2}, a_1 = x_{m-1}, a_2 = x_m$; 如果 0<$k$<$m$, 插值节点取 $a_0 = x_{k-1}, a_1 = x_k, a_2 = x_{k+1}$。从而, x 对应的等效力学模型参数为

$$\begin{aligned} f(x) = \sum_{r=0}^{2}\left(\prod_{\substack{j=0\\ j\neq r}}^{2}\frac{x-a_j}{a_r-a_j}\right) f(a_r) = \\ \left(\frac{x-a_1}{a_0-a_1}\right)\left(\frac{x-a_2}{a_0-a_2}\right) f(a_0) + \\ \left(\frac{x-a_0}{a_1-a_0}\right)\left(\frac{x-a_2}{a_1-a_2}\right) f(a_1) + \\ \left(\frac{x-a_0}{a_2-a_0}\right)\left(\frac{x-a_1}{a_2-a_1}\right) f(a_2) \end{aligned} \tag{14}$$

1.4　贮箱倾斜工况下晃动分析

在重力作用下, 自由液面的法线方向始终与重力方向共线。一般情况下, 为方便建立贮箱内液体的有限元模型, 贮箱的初始姿态位置选择轴线与重力方向平行, 此时, 重力在贮箱本体系中的初始方向为 $\left[0, 0, -1\right]^{\mathrm{T}}$, 对晃动方程 (13) 进行数值计算时, 等效力学模型中单摆相对于贮箱轴线的初始角位移 $\alpha_i (i = 1,2)$ 全部取 0, 而在贮箱倾斜的情况下, 需要重新计算 α_i 的值, 如图 4 所示。

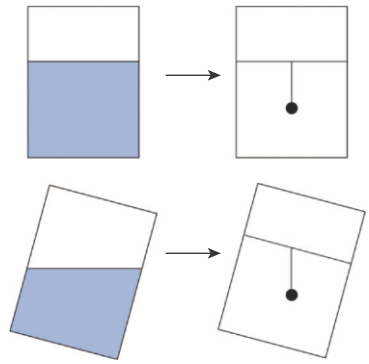


图 4　等效单摆初始摆角

Fig. 4　Equivalent initial swing angle of a pendulum

贮箱倾斜时, 在重力作用下, 等效力学模型中单摆处于贮箱相应姿态下最低点的平衡位置, 此时, 单摆在摆动方向上存在初始角位移 α_i。为求 α_i, 可以通过空间几何方法, 但计算过程复杂, 难以实现程序化计算流程; 而将重力在本体坐标系中的方向和贮箱的初始运动状态代入式 (13) 中, 通过动力学方法将原本计算过程转化为代数求解, 极大降低了运算的复杂程度。

设 t = 0 s 初始条件下, 在贮箱本体坐标系 $O_{\mathrm{b}}xyz$ 中, 重力为

$$\boldsymbol{g} = \boldsymbol{r}_{\mathrm{g}}|\boldsymbol{g}| = \left[a, b, c\right]^{\mathrm{T}}|\boldsymbol{g}| \tag{15}$$

式中: $\boldsymbol{r}_{\mathrm{g}}$ 为重力加速度在贮箱本体坐标系下的方向向量; a、b、c 为方向向量的 3 个分量。

此时, 贮箱无外部激励, 贮箱处于静止状态, 则有

$$\ddot{\boldsymbol{R}}_{\mathrm{b}} = \boldsymbol{\omega}_{\mathrm{b}} = \dot{\boldsymbol{\omega}}_{\mathrm{b}} = \boldsymbol{\omega}_{\mathrm{p}i} = \dot{\boldsymbol{\omega}}_{\mathrm{p}i} = \boldsymbol{0} \tag{16}$$

将式 (16) 代入式 (13) 得

$$\boldsymbol{S}_{\mathrm{p}i} \times \boldsymbol{g} = \boldsymbol{0} \tag{17}$$

则有

$$\begin{cases} m_i|\boldsymbol{g}|\left[x_i, y_i, 0\right]\left\{\left[\begin{bmatrix} x_i & -y_i & 0 \\ y_i & x_i & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ l_i\sin\alpha_i(t) \\ -l_i\cos\alpha_i(t) \end{bmatrix}\right] \times \begin{bmatrix} a \\ b \\ c \end{bmatrix}\right\} = 0 \\ m_i|\boldsymbol{g}|l_i\left[x_i, y_i, 0\right]\left[\begin{bmatrix} -y_i\sin\alpha_i \\ x_i\sin\alpha_i \\ -\cos\alpha_i \end{bmatrix} \times \begin{bmatrix} a \\ b \\ \mathrm{c} \end{bmatrix}\right] = 0 \\ c\left(x_i^2 - y_i^2\right)\sin\alpha_i + (bx_i - ay_i)\cos\alpha_i = 0 \end{cases} \tag{18}$$

$$\alpha_i = \arctan\left(\frac{ay_i - bx_i}{x_i^2 - y_i^2}\right) \qquad i = 1,2 \tag{19}$$

式中: x_i 和 y_i 分别为等效力学模型中第 i 阶晃动单摆的晃动方向在本体坐标系 x 轴和 y 轴上的分量。

使用龙格-库塔法对式 (13) 所示的晃动方程进行数值计算时, 将式 (19) 计算得到的 α_i 作为迭代初始值, 进行动力学方程数值求解。

2　算例分析

使用航天领域最常用的 Cassini贮箱, 分别计算

在变推力、变充液比和初始贮箱倾斜等工况下的晃动力和晃动力矩，并将计算结果与商用软件 Flow-3D 结果进行比较，验证其正确性。通过某探测器液体贮箱在实际飞行过程中的案例，验证本文计算方法在复杂工况下的有效性。

2.1 贮箱基本参数

Cassini 贮箱是航天器中常用的液体燃料贮箱，其中间段为圆柱，两端底面为半球形，是一种典型的轴对称贮箱。在本节的算例中，贮箱本体坐标系定义为原点位于贮箱下底面的顶点，z 轴与贮箱轴线重合，方向竖直向上，y 轴垂直纸面向里，与 x 轴和 z 轴成右手准则，如图 5 所示。

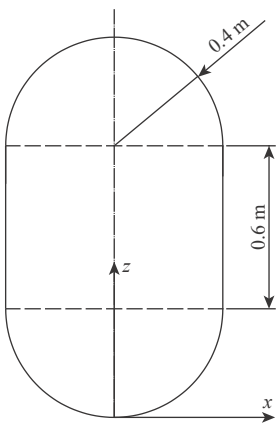


图 5 Cassini 贮箱本体坐标系定义

Fig. 5 Cassini tank coordinate system definition

贮箱具体尺寸如图 5 所示。贮箱内液体为水，密度为 1 000 kg/m^3，运动黏度为 9.29×10^{-7} m^2/s。重力加速度统一取 9.8 m/s^2，初始方向沿本体坐标系 z 轴负方向，不随贮箱姿态位置变化而改变方向。计算时，选择合适的外部激励，保证在计算过程中液体不接触贮箱顶部。

2.2 变推力工况

液体晃动动力学方程利用 C++编程实现求解过程，通过添加平动和转动激励，可以解决贮箱在任意三维运动状态下的晃动耦合问题。本节以 2.1 节中的 Cassini 贮箱为例，充液比为 0.5，为模拟探测器在着陆起飞时发动机推力产生等效重力的情况，原有的重力加速度不变，在贮箱本体系 z 轴方向施加正弦变化的加速度激励，且由于航天器受到的推力方向一直沿 z 轴正方向，加速度总是大于 0，因此，取 z 方向加速度 $a_z=2-2\sin(0.1\pi t)$ m/s^2，液体对贮箱产生的晃动力和晃动力矩分别如图 6 和图 7 所示。

根据图 6 和图 7 可知，在 z 方向施加激励时，产生的晃动力频率与激励频率相同，等效力学模型和 Flow-3D 模型计算结果吻合，其中，等效力学模型通过线性近似的方法计算液体晃动，忽略晃动过程中

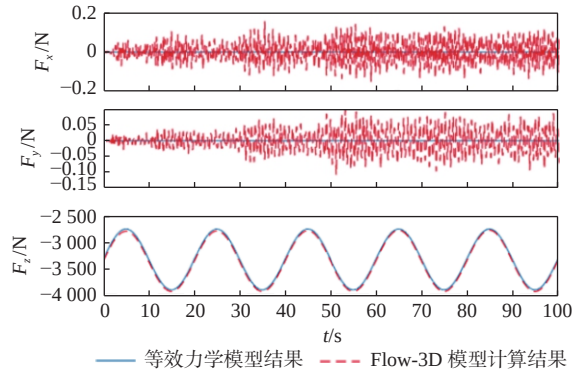


图 6 变推力工况下液体对贮箱的作用力在本体坐标系下的分量

Fig. 6 The component of the force exerted by a liquid on a storage tank under variable thrust conditions in the body coordinate system

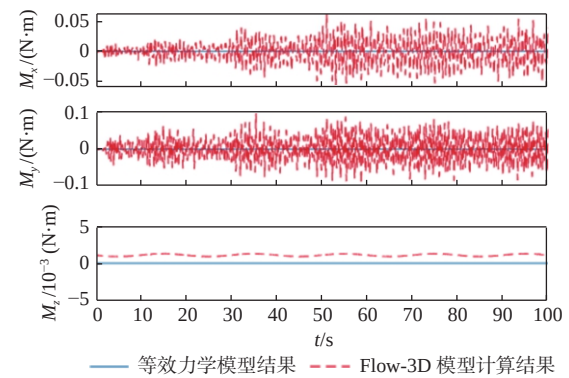


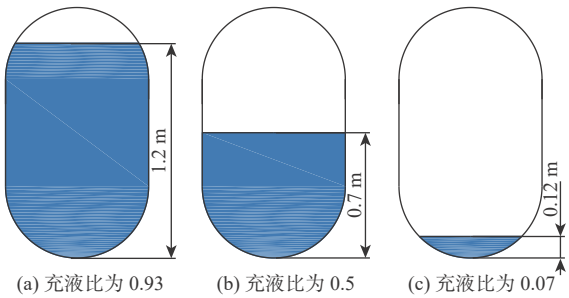
图 7 变推力工况下液体对贮箱的作用力矩在本体坐标系下的分量

Fig. 7 The component of the torque exerted by a liquid on a storage tank under variable thrust conditions in the body coordinate system

的非线性现象，而 Flow-3D 计算过程中会在 x 和 y 方向出现轻微非线性晃动，但相对 z 方向晃动力来说误差较小，符合工程实际的使用要求。晃动力在 z 方向上分量的相对误差在 1% 以内，说明变推力工况下晃动分析方法的有效性。

2.3 变充液比工况

本节进行的变充液比数值验证算例中，形成等效力学模型数据库所需要的充液比插值节点选择从大到小依次为 0.93、0.76、0.5、0.24、0.07，其中，充液比 0.76 和 0.24 所对应的位置是贮箱的上下底面与中间圆柱筒段相切的位置，为保证计算结果的准确性，插值节点中尽量包含这 2 个节点。部分充液比对应的 CAD 模型如图 8 所示。按照 1.3 节中变充液比工况的模拟方法计算各充液比对应的等效力学模型，形成后续计算所需要的数据库。在本节案例的计算时间内，贮箱充液比从 0.93～0.07 之间均匀变化。



(a) 充液比为 0.93　(b) 充液比为 0.5　(c) 充液比为 0.07

图 8　部分充液比工况 CAD 模型

Fig. 8　CAD model of partial liquid filling ratio condition

贮箱绕 x 轴平动运动方程为 $X=\sin(0.1\pi t)$ m, 同时, 绕 x 轴转动运动方程为 $\theta_x=0.1\sin(0.1\pi t)$ rad, 液体对贮箱产生的晃动力和晃动力矩分别如图 9 和图 10 所示。

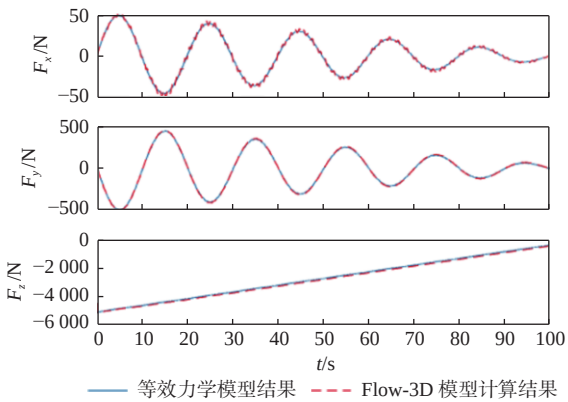


图 9　变充液比时液体对贮箱的作用力在本体坐标系下的分量

Fig. 9　The component of the force of liquid on the storage tank under the condition of variable filling ratio in the body coordinate system

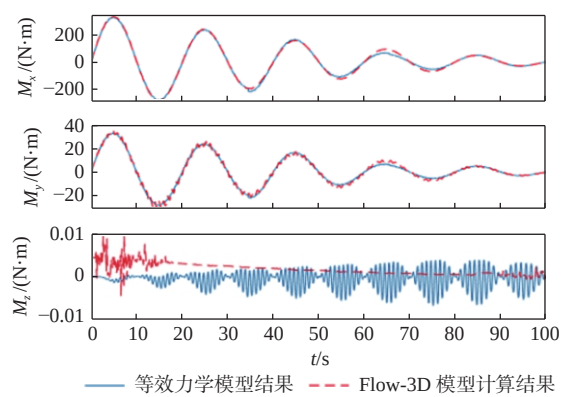


图 10　变充液比时液体对贮箱的作用力矩在本体坐标系下的分量

Fig. 10　The component of the torque of liquid on the storage tank under the condition of variable filling ratio in the body coordinate system

在使用 Flow-3D 进行变充液比条件下的数值模拟时, 采取贮箱内液体从底部均匀流出的方法来达到同样的效果。液体流速要尽可能小, 过大会导致液体流出口产生非线性现象, 此处, 液体流动速率约为 0.005 m^3/s, 对最终计算结果影响较小。在贮箱本体坐标系的同一个方向同时施加平动和转动激励, 最终计算结果与 Flow-3D 吻合较好, 其中, Flow-3D 数值模拟结果中 F_x 和 M_y 的图形曲线不光滑, 但曲线基本趋势与等效力学模型计算结果重合, 通过加密 Flow-3D 计算网格可以改善该现象; 另外, 由于贮箱的转动方向与等效力学模型中单摆摆动方向不重合, 当等效单摆由于摆动而使其质心不在贮箱中心轴时, 贮箱转动使单摆产生 z 方向的晃动力和晃动力矩, 根据计算结果来看, 上述晃动力和晃动力矩数值较小, 几乎不影响最终计算结果。

2.4　初始贮箱倾斜工况

本节算例计算方式选择定充液比为 0.5。为模拟探测器发射时主轴线与水平地面不垂直的状态, Cassini 贮箱初始姿态定义为从贮箱轴线与重力方向重合的位置绕 y 轴顺时针旋转 5°。此时, 贮箱内液体 CAD 模型如图 11 所示。

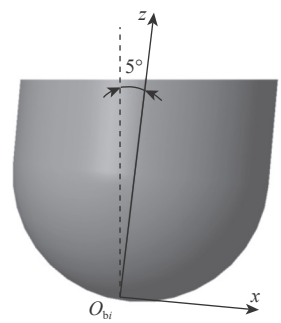


图 11　贮箱倾斜姿态下液体 CAD 模型

Fig. 11　CAD model of liquid under tilted posture of storage tank

重力在本体坐标系中的方向 $\boldsymbol{r}_g=[0.087,\ 0,\ -0.996]^T$。$t$=0 s 时, 贮箱姿态角 $\phi_b=[0.087,\ 0,\ 0]^T$。将等效力学模型数据代入式 (18), 计算求得前 2 阶单摆对应的初始摆角分别为: 第 1 阶单摆的晃动方向 $\boldsymbol{p}_1=[-0.128,\ 0.991,\ 0]$, 初始摆角 $\alpha_1=-0.011$ rad; 第 2 阶单摆的晃动方向 $\boldsymbol{p}_2=[0.991,\ 0.128,\ 0]$, 初始摆角 $\alpha_2=0.086$ rad。

计算贮箱倾斜工况下的液固耦合作用, 关键在于计算前 2 阶等效单摆的初始摆角 α_1 和 α_2, 为验证计算结果的正确性, 对本节定义的贮箱初始姿态采用一种特殊的计算案例: 从 t=0 s 时刻开始, 贮箱在 20 s 时间内, 以本体坐标系 y 轴为旋转轴, 逆时针匀速旋转 5°, 即回到贮箱主轴与重力方向共线的情况, 贮箱停止旋转, 计算总时长为 30 s, 此时, 液体对贮箱的晃动力和晃动力矩在本体系分量分别如图 12 和图 13 所示。

理论上, 此时等效力学模型中的单摆在上述计算得到的初始摆角情况下, 应处于最低点的平衡位

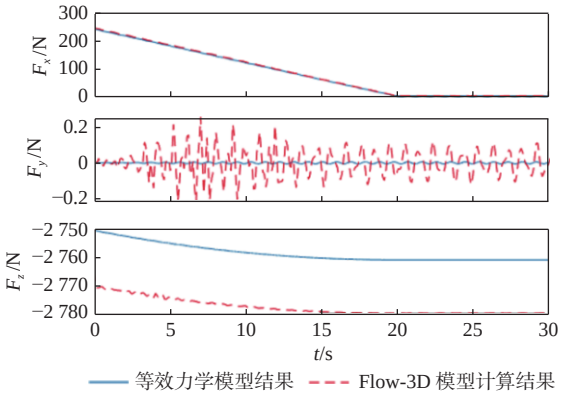


图 12 贮箱倾斜时液体对贮箱的作用力在本体坐标系下的分量

Fig. 12 The component of the force exerted by the liquid on the storage tank when the storage tank is tilted in the body coordinate system

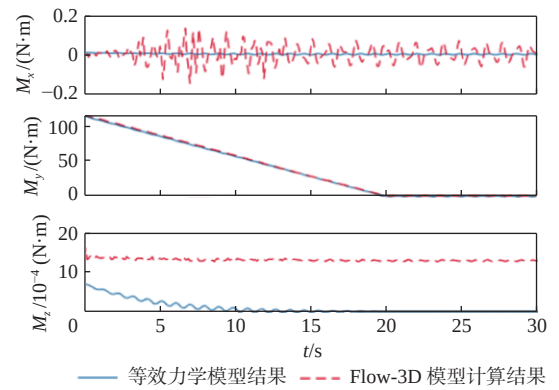


图 13 贮箱倾斜时液体对贮箱的作用力矩在本体坐标系下的分量

Fig. 13 The component of the torque exerted by the liquid on the storage tank when the storage tank is tilted in the body coordinate system

置。当贮箱在所给的转动激励下开始转动时，单摆摆角逐渐减小，在对应方向产生的晃动力和晃动力矩应该随时间逐渐降低。如果此时单摆没有在平衡点位置，则晃动力会在开始阶段少量增大，随后逐渐下降。图 12 和图 13 的 F_x 和 M_y 分量与理论分析结果符合，说明单摆初始摆角的计算正确。最终计算结果与 Flow-3D 模型对比较为一致，从而证明了本文模拟方法的有效性，其中，Flow-3D 计算的 F_y 和 M_x 出现的振荡现象由非线性晃动造成，相比于其他方向数值较小。

2.5 综合算例

以某深空探测器为例，采用飞行过程中实际受到的外部激励作用进行晃动耦合分析。探测器发射前使用推进舱推进，推力产生的加速度为 4 m/s^2，推力大小调整范围为±10%，推力方向调节范围为俯仰偏航各±10%。贮箱初始姿态设定为绕贮箱本体坐标系 x 轴顺时针旋转 5 °，发射后用 10 s 的时间调整到正常姿态，重力加速度为 9.8 m/s^2，贮箱内注入的液体为水，密度为 1 000 kg/m^3，运动黏度为 9.29×10^{-7} m^2/s，计算时长为 100 s。发射过程中，贮箱内燃料逐渐消耗，充液比变化范围为 0.92～0.14。为方便数值模拟，对上述飞行过程作相应的简化处理：①推力大小调整按照正弦变化规律(sin(0.1πt))；②推力方向保持与 O_bxz 平面呈 45°夹角的平面内转动，此时，推力在贮箱本体坐标系 x 轴和 y 轴方向上的分量相等，转动运动规律同样按照正弦变化(sin(0.1πt))；③整个计算过程中，充液比在 0.92～0.14 之间均匀变化；④贮箱在 0～10 s 时间内匀速从倾斜状态旋转到轴线与重力方向平行的状态。晃动力和晃动力矩计算结果分别如图 14 和图 15 所示。

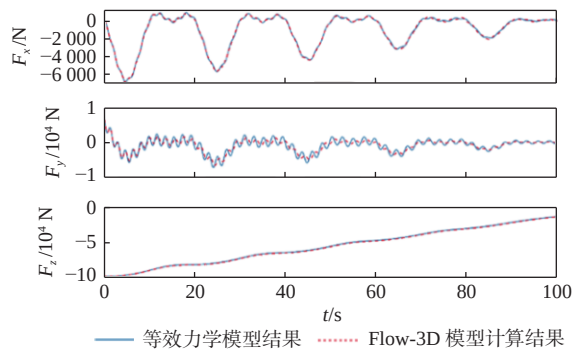


图 14 液体对贮箱的作用力在本体坐标系下的分量

Fig. 14 The component of the force of liquid on the storage tank in the body coordinate system

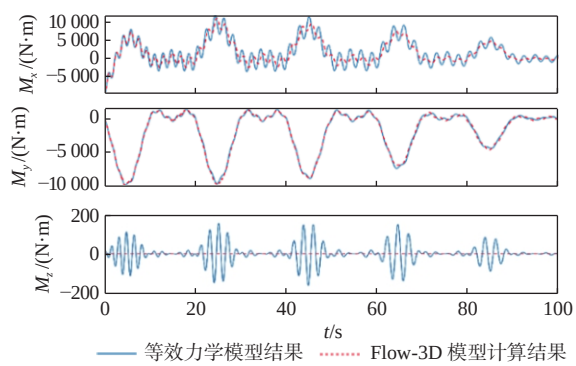


图 15 液体对贮箱的作用力矩在本体坐标系下的分量

Fig. 15 The component of the torque exerted by the liquid on the storage tank in the body coordinate system

t=0 s 时，贮箱处于绕本体坐标系 x 轴旋转了一定角度的倾斜状态，理论上，晃动力在 y 轴上的分量 F_y 不为零，而在充液比逐渐减小的过程中，晃动力和晃动力矩的幅值会随之减小，同时，在推力变化的情况下，晃动力在本体坐标系 z 轴上的分量会出现与外部激励相应的周期性起伏，且由于推力方向发生变化，晃动力在本体坐标系 x 轴和 y 轴方向

上也出现相应的分量。可见, 计算结果基本符合理论预期, 同时, 与 Flow-3D 数值模拟结果对比吻合较好。由于算例中的航天器贮箱体积较大, 对应等效力学模型中单摆的质量在 10^3 kg 数量级上, 根据 2.3 节的分析, 由于贮箱转动而产生的 z 方向上晃动力矩同样会相应增加, 虽然晃动力矩与 Flow-3D 计算结果最大误差接近 200 N·m, 但相对于其他 2 个方向的计算结果来说, 相差了接近 2 个数量级, 符合工程实际需求。另外, 在 Flow-3D 中, 为保证计算结果的收敛性, 计算过程中会添加额外的阻尼项, 导致 F_y 和 M_x 的曲线能够在很短时间内达到平滑状态, 等效力学模型对液体阻尼的计算偏小, 实际阻尼应该处于两者之间。

3　结　论

本文基于已有的液体晃动等效力学模型理论, 根据虚功率原理建立等效模型的动力学方程, 实现任意形状的航天器液体贮箱在任意三维平动和转动运动状态下的晃动力和晃动力矩计算。同时, 针对航天器在发射过程中出现的 3 种特殊工况, 提出液体燃料晃动力和晃动力矩的计算方法:

1) 利用等效单摆模型在等效重力变化时摆长保持恒定的特性, 通过在贮箱上施加大小和方向不断变化加速度激励的方式模拟航天器飞行过程中推力大小和方向变化的情况。

2) 通过插值法更新等效力学模型参数, 以适应航天器在飞行过程中液体燃料逐渐消耗的情况。

3) 通过改变等效力学模型中单摆初始摆角的方法模拟航天器在准备发射阶段贮箱主轴线与地面不垂直的情况。

使用航天领域最常见的 Cassini 贮箱和某型号航天器实际燃料贮箱, 利用本文计算方法在上述 3 种特殊工况下进行晃动力和晃动力矩计算, 并将计算结果与商业软件 Flow-3D 计算结果比较, 计算误差较小, 符合工程实际的要求。由此证明本文计算方法的正确性, 且在复杂的工况下能够保证计算结果的有效性。

参考文献 (References)

[1] ABRAMSON H N. The dynamic behavior of liquids in moving containers: with applications to space vehicle technology: NASA SP-106[R]. Washington, D.C.: Scientific and Technical Information Division, NASA, 1966.

[2] TANG Y, YUE B Z, YAN Y L. Improved method for implementing contact angle condition in simulation of liquid sloshing under microgravity[J]. International Journal for Numerical Methods in Fluids, 2019, 89(4-5): 123-142.

[3] 李遇春. 液体晃动动力学基础[M]. 北京: 科学出版社, 2017.
LI Y C. Fundamentals of liquid sloshing dynamics[M]. Beijing: Science Press, 2017(in Chinese).

[4] BUSECK R, BENAROYA H. Mechanical models for slosh of liquid fuel[C]//Proceedings of the Aerospace Design Conference. Reston: AIAA, 1993: 1093.

[5] IBRAHIM R A. Liquid sloshing dynamics[M]. Cambridge: Cambridge University Press, 2005.

[6] DODGE F T. The new "dynamic behavior of liquids in moving containers": NASA SP-106[R]. San Antonio: Southwest Research Institute, 2000.

[7] GRAHAM E W, RODRIGUEZ A M. The characteristics of fuel motion which affect airplane dynamics[J]. Journal of Applied Mechanics, 1952, 19(3): 381-388.

[8] HOUSNER G W. Dynamic pressures on accelerated fluid containers[J]. Bulletin of the Seismological Society of America, 1957, 47(1): 15-35.

[9] RATTAYYA J. Sloshing of liquids in axisymmetric ellipsoidal tanks[C]//Proceedings of the 2nd Aerospace Sciences Meeting. Reston: AIAA, 1965: 114.

[10] MIKISHEV G N. Experimental method of spacecraft dynamics[M]. Moscow: Mechanical Manufacture, 1978.

[11] 包光伟. 平放柱形贮箱内液体晃动的等效力学模型[J]. 上海交通大学学报, 2003, 37(12): 1961-1964.
BAO G W. Equivalent mechanical model of liquid sloshing in horizontal cylindrical container[J]. Journal of Shanghai Jiao Tong University, 2003, 37(12): 1961-1964(in Chinese).

[12] DODGE F T, GARZA L R. Equivalent mechanical model of propellant sloshing in the workshop configuration of the Saturn S-IVB: NASA CR-101407[R]. Washington, D.C.: NASA, 1969.

[13] 包光伟, 王政伟. 液体三维晃动特征问题的有限元数值计算方法[J]. 力学季刊, 2003, 24(2): 185-190.
BAO G W, WANG Z W. Finite element method for eigen problem of liquid 3D sloshing[J]. Chinese Quarterly of Mechanics, 2003, 24(2): 185-190(in Chinese).

[14] 李青, 马兴瑞, 王天舒. 非轴对称贮箱液体晃动的等效力学模型[J]. 宇航学报, 2011, 32(2): 242-249.
LI Q, MA X R, WANG T S. Equivalent mechanical model for liquid sloshing in non-axisymmetric tanks[J]. Journal of Astronautics, 2011, 32(2): 242-249(in Chinese).

[15] LI Q, MA X R, WANG T S. Equivalent mechanical model for liquid sloshing during draining[J]. Acta Astronautica, 2011, 68(1-2): 91-100.

Analysis of liquid sloshing under variable thrust, variable filling liquid ratio, and tilted tank state

HE Jiawei[1, 2], LYU Jing[2, *], WANG Tianshu[3], LIU Yingjie[1]

(1. The General Department of HuBei Areospace Technology Academy, Wuhai 430048, China;
2. School of Aeronautic Science and Engineering, Beihang University, Beijing 102206, China;
3. School of Aerospace Engineering, Tsinghua University, Beijing 100083, China)

Abstract: With the development of aerospace technology, new deep space probes will encounter two special operating conditions during takeoff: the magnitude and direction of the engine thrust vector changing over time; the direction of gravity during the initial launch does not coincide with the main axis of the spacecraft. During the flight, liquid fuel will be continuously consumed. This article suggests a new solution for the problem of liquid sloshing in spacecraft storage tanks to the aforementioned three working conditions: taking into account the impact of spacecraft translational acceleration during the dynamic modeling stage and then analyzing the sloshing situation of thrust not along the spacecraft axis direction; calculating the liquid sloshing problem under variable charge to liquid ratio conditions using the method of variable equivalent mechanical model parameters; and simulating the situation where the liquid level is not perpendicular to the tank axis when the tank tilts by setting an initial swing angle for a single pendulum in the equivalent mechanics model. This article takes the Cassini storage tank as an example to calculate the sloshing force and moment under the three working conditions mentioned above in a semi filled state. Finally, taking a certain type of spacecraft liquid storage tank as an example, sloshing analysis is carried out under actual launch conditions. The practicality of the calculation approach presented in this article is confirmed by the calculation results, which are in good agreement with the results of calculations made using commercial software Flow-3D.

Keywords: equivalent mechanical model; liquid sloshing; variable thrust; variable filling liquid ratio; tank tilt

Received: 2024-05-06; **Accepted**: 2024-08-09; **Published Online**: 2024-09-27 12:10
URL: link.cnki.net/urlid/11.2625.V.20240926.1748.003

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0259

基于多策略融合蚁狮优化算法的MPRM逻辑电路面积优化方法

潘家义[1,2]，何振学[1,2,*]，赵晓君[1,2]，何俊才[1,2]，周宇豪[3]，王翔[4]

(1. 河北农业大学 智能农业装备研究院，保定 071001；2. 河北农业大学 河北省农业大数据重点实验室，保定 071001；3. 同济大学 软件学院，上海 201804；4. 北京航空航天大学 电子信息工程学院，北京 100191)

摘 要：针对现有混合极性Reed-Muller (MPRM)逻辑电路面积优化方法优化效果较差的问题，提出一种多策略融合蚁狮优化(MSALO)算法。为解决蚁狮优化(ALO)算法全局搜索能力较差的问题，在算法随机游走阶段应用双策略随机游走机制；为解决算法寻优能力差的问题，针对精英蚁狮个体应用突围机制；为加快算法的收敛速度，引入基于正弦函数的自适应蚂蚁位置更新策略。提出一种基于MSALO算法的MPRM逻辑电路面积优化方法，基于北卡罗莱纳州微电子中心(MCNC)基准测试电路的实验结果表明：基于MSALO算法的MPRM逻辑电路面积优化方法平均面积节省率可在现有优秀群智能优化算法的基础上平均提高27.96%。

关 键 词：混合极性Reed-Muller；组合优化问题；群智能优化；蚁狮优化；突围机制
中图分类号：V443; TP391.72
文献标志码：A **文章编号**：1001-5965(2026)06-2083-09

如今，人们对电子产品提出了更高的要求，研发更小面积、更低功耗、更高集成度的芯片显得尤为必要。电路面积优化是实现上述需求的关键，基于XOR/AND的混合极性Reed-Muller（mixed polarity Reed-Muller, MPRM）逻辑电路在某些情况下比基于AND/OR/NOT的Boolean逻辑更优。因此，基于XOR/AND的MPRM逻辑电路面积优化成为电子设计自动化（electronics design automation, EDA）中的关键研究领域[1-3]，引起了国内外学者的广泛关注[4-5]。

任意一个n输入变量的Boolean逻辑表达式对应3^n个不同的MPRM逻辑电路表达式极性[6]。不同极性下的MPRM逻辑电路表达式所含有的XOR项和AND项的数量不同，导致电路的复杂度不同。因此，MPRM逻辑电路面积优化就是在极性优化空间中搜索一个电路表达式最简洁的最佳极性。随着电路规模的增大，对应的MPRM逻辑电路优化空间也呈指数形式增大[7]，对于中小规模电路，采用枚举法可以在有限时间内找到最佳极性；而对于中大规模电路，受限于指数型对应关系的维度爆炸现象，传统方法在该问题上失去可行性，因此，基于群智能优化算法的中大规模MPRM逻辑电路面积优化方法成为研究热点[8]。

文献[9]利用改进粒子群算法求解MPRM面积和功耗优化问题，通过改进种群中个体的学习策略提高了算法的搜索效率，但难以较好平衡局部搜索与全局开发之间的关系。文献[10]通过反向学习策略，有效提高了人工鱼群算法初始种群整体质量，增强了算法的搜索能力。文献[11]将高斯变异

收稿日期：2024-04-26；**录用日期**：2024-06-14；**网络出版时间**：2024-07-08 11:19
网络出版地址：link.cnki.net/urlid/11.2625.V.20240705.1131.002
基金项目：国家自然科学基金(62102130)；中央引导地方科技发展资金项目(226Z0201G)；河北省自然科学基金(F2024204001,F2020204003)；河北省青年拔尖人才计划项目(BJ2019008)；河北省高等学校科学技术研究项目(QN2024138)；河北省省属高等学校基本科研业务费研究项目(KY2022073)

* **通信作者**. E-mail：hezhenxue@buaa.edu.cn

引用格式：*潘家义，何振学，赵晓君，等. 基于多策略融合蚁狮优化算法的MPRM逻辑电路面积优化方法[J]. 北京航空航天大学学报，2026，52（6）：2083-2091. PAN J Y，HE Z X，ZHAO X J，et al. Area optimization approach for MPRM logic circuits based on multi-strategy synergy ant lion optimization algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（6）：2083-2091（in Chinese）.*

算子和轮盘赌选择策略融入烟花算法，对MPRM电路面积进行优化。文献[12]通过一种改进的遗传算法，提高了MPRM逻辑电路的优化效率，但仍存在稳定性较差的问题。

通过观察、模拟蚁狮在自然界中的行为，澳大利亚学者Mirjalili提出蚁狮优化(ant lion optimization, ALO)算法[13]。该算法具有原理简单、易于实现等优点，自提出以来，已成功应用于路径规划[14]、交通管理[15]、电源配置[16]等多个领域，并取得了显著的优化效果。得益于独特的游走机制，ALO具有较强的全局搜索能力，能更快在MPRM逻辑电路优化空间中定位到最优解，针对MPRM逻辑电路面积优化问题具有明显优势。

ALO在处理离散类组合优化问题时，仍存在易陷入局部最优、收敛速度慢的问题，为此，本文提出多策略融合蚁狮优化(multi-strategy synergy ant lion optimization, MSALO)算法，该算法利用双策略随机游走机制，突围策略，自适应位置更新策略，提高算法处理组合优化问题的能力，增强了算法搜索效率。同时，通过对比实验验证了本文提出的MSALO算法应用于MPRM逻辑电路面积优化问题的有效性和优越性。

1 基础知识

1.1 MPRM逻辑表达式

n输入变量位数的Boolean函数表达式为

$$f(x_{n-1},x_{n-2},\cdots,x_0)=\sum_{i=0}^{2^n-1}\alpha_i a_i \tag{1}$$

式中：α_i表示第i个最小项是否存在，当α_i取值为1时，表示该最小项存在，取0时，表示该最小项不存在；a_i为Boolean函数中的最小项。

任意输入变量位数为n的Boolean逻辑表达式都可以展开为极性为p的基于XOR/AND的MPRM逻辑表达式：

$$f^p(x_{n-1},x_{n-2},\cdots,x_0)=\oplus\sum_{i=0}^{2^n-2}\beta_i b_i \tag{2}$$

式中：p为MPRM逻辑表达式的极性；b_i表示AND项；β_i表示该AND项是否存在，取值为1时，表示该AND项存在，为0时，表示该AND项不存在；$\oplus\sum$表示XOR运算。

1.2 极 性

MPRM逻辑电路的极性取值为0、1、2，一个n输入变量的Boolean逻辑电路可展开为形如$\dot{b}_{n-1}\dot{b}_{n-2}\dot{b}_{n-3}\cdots\dot{b}_0$的MPRM逻辑电路，其中，$\dot{b}_k$的表现形式由极性$p_k$和式(2)中下标$i_k$共同决定：当$p_k=0,i_k=0$或$p_k=1,i_k=0$时，$\dot{b}_k$表现为1；当$p_k=0,i_k=1$或$p_k=2,i_k=1$时，$\dot{b}_k$表现为$b_k$；当$p_k=1,i_k=1$或$p_k=2,i_k=0$时，$\dot{b}_k$表现为$\bar{b}_k$($b_k$的逻辑反形式)。

例1 给定一个3输入的Boolean逻辑函数表达式：

$$f=x_2\bar{x}_1x_0+\bar{x}_2\bar{x}_1x_0+x_2x_1\bar{x}_0 \tag{3}$$

当极性为$0(000)_3$时，该Boolean逻辑表达式对应的MPRM逻辑表达式为

$$f=x_2x_1\oplus x_0\oplus x_2x_1x_0\oplus x_1x_0 \tag{4}$$

当极性为$26(222)_3$时，该Boolean逻辑表达式对应的MPRM逻辑表达式为

$$f=x_2\bar{x}_1x_0\oplus\bar{x}_2\bar{x}_1x_0\oplus x_2x_1\bar{x}_0 \tag{5}$$

式(4)和式(5)分别表示式(3)在极性0和极性26下的MPRM逻辑电路表达式，可知，同一个Boolean逻辑表达式在不同极性下得到的MPRM逻辑表达式复杂度不同，而不同MPRM逻辑表达式对应的电路所需门级逻辑单元不同。因此，同一个Boolean逻辑电路在不同极性下MPRM逻辑电路面积也不同。

1.3 ALO算法

ALO算法通过模拟蚁狮在沙堆中挖掘陷阱捕猎蚂蚁的行为实现对最优解的搜索。算法中，蚂蚁和蚁狮的位置是所求问题的一个可行解，通过模拟蚂蚁在蚁狮陷阱中的随机游走行为进行全局搜索和局部寻优，当蚂蚁的适应度值优于蚁狮时，蚁狮成功捕获蚂蚁，并取代当前蚂蚁的位置。ALO算法的具体流程如下。

1.3.1 蚂蚁随机游走

在ALO算法中，蚂蚁的随机游走行为表示为

$$X_s=[0,\mathrm{cumsum}(s_1),\cdots,\mathrm{cumsum}(s_T)] \tag{6}$$

式中：cumsum(·)表示计算当前蚂蚁前s_j次随机游走步长的累加和；T为最大迭代次数；s_j为第j次随机游走步长，定义为

$$s_j=\begin{cases}1 & r_{\mathrm{d}j}\geqslant 0.5\\ -1 & r_{\mathrm{d}j}<0.5\end{cases} \tag{7}$$

其中：$r_{\mathrm{d}j}$为第j次游走中的[0,1]之间的随机数。

1.3.2 蚂蚁位置标准化

为防止蚂蚁在随机游走阶段逃出蚁狮的陷阱，即超出算法解空间范围，需要通过标准化操作将X_{s_T}映射到解空间，映射后的相对位置M_j^{d}为

$$M_j^{\mathrm{d}}=\frac{(X_{s_j}-X_{\min})(u_{\mathrm{b}}-l_{\mathrm{b}})}{X_{\max}-X_{\min}}+l_{\mathrm{b}} \tag{8}$$

式中：$X_{\min}$为蚂蚁随机游走公式中的最小值；$X_{\max}$为

蚂蚁随机游走公式中的最大值; l_b为算法的初始解空间下界; u_b为算法的初始解空间上界。

1.3.3　蚁狮构造陷阱

蚁狮构造的陷阱范围如下:

$$u_b = L_k^d + \frac{l_b}{r_t} \tag{9}$$

$$l_b = L_k^d + \frac{u_b}{r_t} \tag{10}$$

式中: L_k^d为第k个蚁狮个体的第d个维度; r_t用于控制蚂蚁在陷阱内的陷落趋势, 使蚂蚁位置随迭代次数增加逐渐靠近蚁狮, 计算方法为

$$r_t = \begin{cases} \dfrac{t}{T} & t < 0.1T \\ 100\dfrac{t}{T} & 0.1T < t < 0.5T \\ 1\,000\dfrac{t}{T} & 0.5T < t < 0.75T \\ 10\,000\dfrac{t}{T} & 0.75T < t < 0.9T \\ 100\,000\dfrac{t}{T} & 0.9T < t < 0.95T \\ 1\,000\,000\dfrac{t}{T} & 0.95T < t \end{cases} \tag{11}$$

其中: t为当前迭代次数。

1.3.4　蚂蚁位置更新

在算法的每一次迭代中, 蚂蚁会根据随机游走机制更新自己的位置, 更新位置公式为

$$A_k^d = \frac{X_R^d + X_E^d}{2} \tag{12}$$

式中: A_k^d为迭代后蚂蚁位置的第d个维度; X_R^d为轮盘赌蚁狮中第d个维度对蚂蚁造成的随机扰动; X_E^d为精英蚁狮中的第d个维度对蚂蚁造成的随机扰动。

1.3.5　蚁狮位置更新

在蚂蚁随机游走后, 蚁狮会根据式 (13) 更新自身位置:

$$L_l = A_m \qquad F(A_m) > F(L_l) \tag{13}$$

式中: $F(A_m)$为第m个蚂蚁个体的适应度值; $F(L_l)$为第l个蚁狮个体的适应度值。

2　MSALO 算法

ALO 算法在处理中大规模组合优化问题时, 存在收敛性较差、易陷入局部最优的缺点。因此, 本文提出 MSALO 算法, 引入蚂蚁突围策略, 使蚂蚁在当前最优位置随机选择某个方向做出突围行为, 增强算法跳出局部最优的能力; 受哈里斯鹰算法[17]启发, 提出双策略随机游走机制, 以提高算法的全局搜索能力; 提出自适应位置更新策略, 增强算法的收敛能力。

2.1　双策略随机游走

在 ALO 算法的随机游走阶段, 蚂蚁围绕蚁狮的游走步长为定值, 不利于算法在前期的全局搜索, 导致算法收敛速度过慢。哈里斯鹰算法通过计算个体剩余体力, 当剩余体力大于设定阈值时, 采用软包围策略, 当该值小于设定阈值时, 采用硬包围策略。软包围和硬包围在算法不同阶段的使用使哈里斯鹰算法具有较好的全局搜索能力。

为提高 ALO 算法的全局搜索能力, 提出一种融合哈里斯鹰算法的双策略随机游走策略, 具体操作如下: 通过变量E判断算法当前使用哈里斯鹰算法的游走行为或 ALO 算法的游走行为, E的计算方式为

$$E = 2E_0\left(1 - \frac{t}{T}\right) \tag{14}$$

式中: E_0为范围为$(-1,1)$的随机数。

当E的绝对值小于 1 时, 采用 ALO 算法的随机游走策略, 当E的绝对值大于或等于 1 时, 采用融合哈里斯鹰算法的全局搜索策略:

$$p_{o1} = l_b + r_1(u_b - l_b) \tag{15}$$

式中: p_{o1}为精英游走系数; r_1表示范围为$(-1,1)$的随机数。

$$p_{o2} = \left|A_{r_2}^d - 2r_3A_i^d\right| \tag{16}$$

式中: p_{o2}为个体游走系数; r_2、r_3表示范围为$(-1,1)$的随机数。

$$p_{os1} = L_E - L_{avg} - r_4p_{o1} \tag{17}$$

式中: p_{os1}为精英游走步长; L_E为精英蚁狮的位置; L_{avg}为整个蚁狮种群中适应度值处于前一半的蚁狮个体位置的平均值; r_4为范围为$(-1,1)$的随机数。

$$p_{os2} = A_{r_5}^d - r_6p_{o2} \tag{18}$$

式中: p_{os2}为个体游走步长; r_5、r_6表示范围为$(-1,1)$的随机数。

$$A_j^d = \begin{cases} p_{os1} & q < 0.5 \\ p_{os2} & q \geqslant 0.5 \end{cases} \tag{19}$$

式中: q为范围为$(0,1)$的随机数。

2.2　突围机制

在 ALO 算法中, 蚁狮构造陷阱这一行为精度较差, 存在难以跳出局部最优并找到全局最优解的问题。本文提出一种突围机制, 在蚂蚁逐渐靠近全局最优, 或已经陷入局部最优的情况下, 产生突围行为: 随机选择求解问题的若干维度作为突围方

向，并参考狼群算法[18]及花授粉算法[19]得到突围距离，进而帮助算法跳出局部最优。

突围机制如下：

$$s_{\mathrm{dim}} = \mathrm{mod}\left(t, s_{\mathrm{p}}\right) + 1 \tag{20}$$

式中：s_{dim}为突围方向；$\mathrm{mod}\left(t, s_{\mathrm{p}}\right)$表示迭代次数$t$对突围参数$s_{\mathrm{p}}$取余。

策略转换因子r_{o}为

$$r_{\mathrm{o}} = \frac{t}{T} \tag{21}$$

设p_{rand}为范围为(0,1)的随机数，当$p_{\mathrm{rand}} > r_{\mathrm{o}}$时，突围距离的确定参考狼群算法的捕猎机制，表示为

$$s_{\mathrm{tru}} = 2L_{\mathrm{E}} - g_{\mathrm{p_eli}} \tag{22}$$

式中：s_{tru}为突围距离；$g_{\mathrm{p_eli}}$为每一次迭代中所有精英个体的平均值。

当$p_{\mathrm{rand}} < r_{\mathrm{o}}$时，突围距离的确定参考花授粉算法中的局部授粉机制，表示为

$$s_{\mathrm{tru}} = L_{\mathrm{E}} + r_7\left(L_{rj} - L_{rk}\right) \tag{23}$$

式中：r_7为范围为(0,1)的随机数；L_{rj}、L_{rk}表示蚁狮群中的任意2个不同个体。

根据突围机制生成的新蚁狮位置n_{L}为

$$n_{\mathrm{L}} = L_{\mathrm{E}} + s_{\mathrm{tru}} \tag{24}$$

应用突围机制后，当新生成的蚁狮个体适应度值更优时，更新精英蚁狮位置为新蚁狮位置。

突围机制示意图如图1所示。

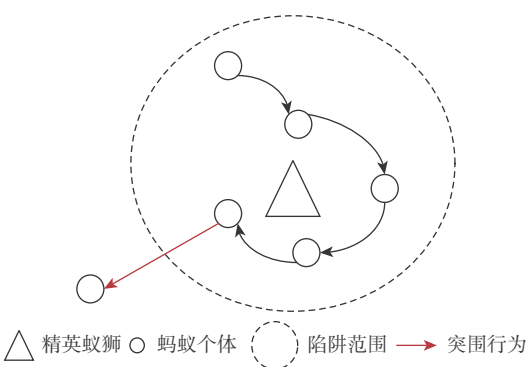


图1 突围机制示意图

Fig. 1 Schematic diagram of the breakout mechanism

2.3 自适应位置更新

在ALO算法蚂蚁围绕轮盘赌选择蚁狮和精英蚁狮的游走行为中，轮盘赌选择蚁狮和精英蚁狮的权重为1∶1，如式(12)所示，该固定权重不利于算法后期收敛。正弦函数在$(0,\pi/2)$范围内，函数值y随自变量x的增大而逐渐增大，且函数值在前期增长速度较快，随着自变量的增大，函数值的增长速度逐渐变慢，并最终趋于1[20]。因此，本文根据正弦函数的变化特点，提出一种基于正弦函数的自适应蚂蚁位置更新策略。正弦因子为

$$s_{\mathrm{group}} = \sin\left(\frac{\pi t}{2T}\right) \tag{25}$$

优化设比例参数为p_{sin}，当$s_{\mathrm{group}} < p_{\mathrm{sin}}$时，$s_{\mathrm{group}}$取定值$p_{\mathrm{sin}}$，当$s_{\mathrm{group}} > 1 - p_{\mathrm{sin}}$时，$s_{\mathrm{group}}$取定值$1 - p_{\mathrm{sin}}$。

基于正弦函数的自适应蚂蚁位置更新策略表示为

$$A_i^t = \left(1 - s_{\mathrm{group}}\right) X_{\mathrm{R}}^d + s_{\mathrm{group}} X_{\mathrm{E}}^d \tag{26}$$

将ALO算法中式(12)修改为式(26)，以提高算法的收敛速度。

2.4 算法流程

综上所述，MSALO算法的伪代码如下：

输入： MSALO算法所需参数。

输出： 全局最优解。

1. 设置初始化参数。
2. 初始化种群，计算种群中个体的适应度值。
3. 取适应度值最好的蚁狮作为精英个体。
4. WHILE true DO
5. IF $|E| < 1$ THEN
6. 执行ALO算法的随机游走行为
7. END IF
8. IF $|E| > 1$ THEN
9. 执行哈里斯鹰算法的随机游走行为
10. END IF
11. 更新蚁狮位置，并确定新的精英蚁狮
12. 针对精英蚁狮执行突围机制
13. IF 算法达到终止条件 THEN
14. 跳出WHILE循环
15. END IF
16. END WHILE
17. RETURN 全局最优解

首先，随机初始化蚂蚁和蚁狮种群，并确定精英蚁狮个体；其次，依据变量E执行双策略随机游走；最后，执行突围机制，当算法达到终止条件时，停止迭代，输出精英蚁狮个体位置为全局最优。

3 MPRM逻辑电路面积优化方法

基于XOR/AND的MPRM逻辑电路面积优化问题随着电路输入变量位数增多，其解空间呈指数形式增长，属于典型的组合优化问题。现有的MPRM逻辑电路面积优化方法存在稳定性差、易

陷入局部最优的缺点。本文提出的 MSALO 算法具有优化效果好、优化效率高的优点，因此，将 MSALO 算法用于 MPRM 逻辑电路面积优化，能有效提高优化效率。

3.1　编码方案

以极性作为蚁狮和蚂蚁个体的位置，并将位置编码为离散三进制形式，n个输入的逻辑电路解空间维度为n。

例 2　一个 6 输入变量位数的电路，蚁狮个体维度为 6，若极性为$293(101212)_3$，则蚁狮个体每一维度对应一位三进制编码，此时，该蚁狮个体位置为 101212，编码与解码流程如图 2 所示。

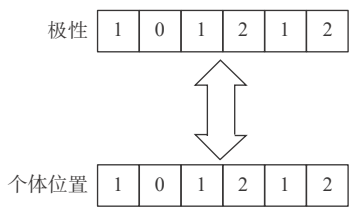


图 2　编码与解码流程

Fig. 2　Encoding and decoding flowchart

3.2　面积估算模型

将二输入 XOR 门和二输入 AND 门总数作为 MPRM 逻辑电路面积，需要将多输入 XOR 分解为二输入 XOR，并以二输入 XOR 和 OR 为计算单位，进行面积估算。二输入 XOR 门和 AND 门的计算方式为

$$a_{\mathrm{x}}=\sum_{k=0}^{2^n-1}\beta_k-1 \tag{27}$$

$$a_{\mathrm{d}}=\sum_{k=0}^{2^n-1}\left[\beta_k\left(\sum_{l=0}^{n-1}b_l-1\right)\right] \tag{28}$$

式中：a_{x}为 XOR 门数量；a_{d}为 AND 门个数；β_k为乘积项数量；b_l为输入变量数。

综上所述，MPRM 逻辑电路面积 a 估算模型为

$$a=a_{\mathrm{x}}+a_{\mathrm{d}} \tag{29}$$

例 3　式 (4) 对应的 MPRM 逻辑电路消耗 XOR 门 3 个，AND 门 4 个，则a_{x}的值为 3，a_{d}的值为 4，该电路面积a为 7。

例 4　式 (5) 对应的 MPRM 逻辑电路消耗 XOR 门 2 个，AND 门 6 个，则a_{x}的值为 2，a_{d}的值为 6，该电路面积a为 8。

3.3　方法流程

本节列出了基于 MSALO 算法的 MPRM 逻辑电路面积优化方法实现步骤，方法流程如图 3 所示。方法实现步骤如下：

1）读取 Boolean 电路逻辑网表。

2）设置 MSALO 算法最大迭代次数T、种群大小、比例参数$p_{\sin}$。

3）初始化蚁狮和蚂蚁的位置，设置当前迭代次数$t=1$，并根据面积模型计算初始蚂蚁和蚁狮的适应度值，选择适应度值最小的蚁狮为精英蚁狮。

4）更新蚂蚁位置，当$|E|<1$时，采用式 (12) 更新蚂蚁位置，否则，使用式 (15)～式 (19) 更新蚂蚁位置，并计算更新后蚂蚁适应度值。

5）根据蚂蚁适应度值更新蚁狮位置。

6）针对精英蚁狮位置应用突围机制，当$p_{\mathrm{rand}}>r_{\mathrm{o}}$时，应用式 (22) 和式 (24)；当$p_{\mathrm{rand}}<r_{\mathrm{o}}$时，应用式 (23) 和式 (24)。

7）判断当前迭代是否满足结束条件，若没有达

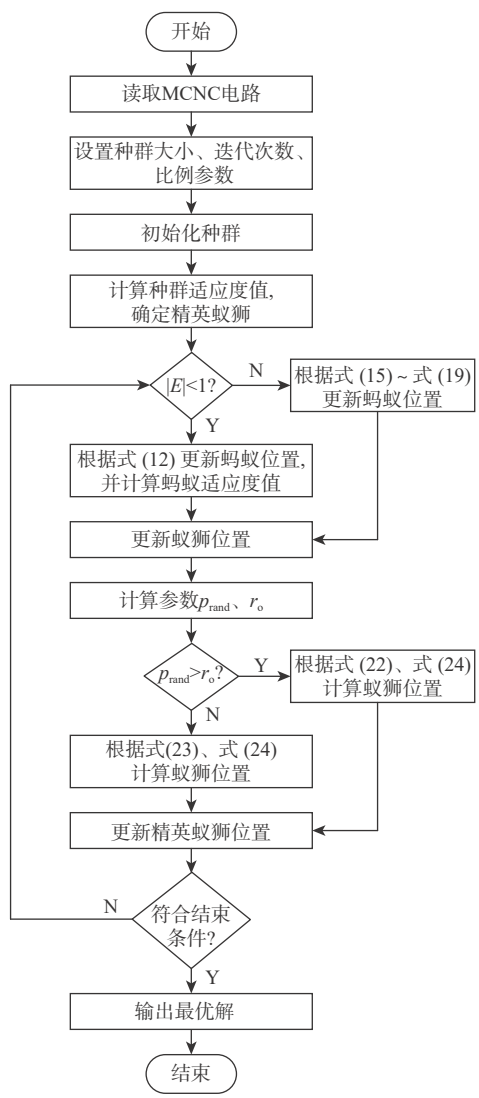


图 3　本文方法流程

Fig. 3　Flowchart of the proposed method

到，则$t=t+1$，并返回步骤4），否则输出精英蚁狮位置作为全局最优值。

4 实验结果及分析

实验运行环境基于Windows10 64位操作系统，16 GB RAM, MATLAB R2020b, Intel(R) Core(TM) i5-8250U CPU @ 1.60 GHz 1.80 GHz。本文实验方法均使用MATLAB实现。

为验证本文方法的有效性，将其分别与基于ALO算法的MPRM逻辑电路面积优化方法、基于黑猩猩优化算法(chimp optimization algorithm, ChOA)[21]的MPRM逻辑电路面积优化方法、基于雪消融优化(snow ablation optimization, SAO)算法[22]的MPRM逻辑电路面积优化方法在北卡罗莱纳州微电子中心(Microelectronics Center of North Carolina, MCNC)基准测试电路测试集中随机选取的13个电路上进行MPRM逻辑电路面积优化实验。

MSALO算法的参数通过引入三因素三水平的正交表确定，SAO算法的参数设置与文献[22]保持一致，ChOA的参数设置与文献[21]保持一致。

群智能优化算法的寻优过程是一个根据特定规则进行随机搜索的过程，因此，为避免算法随机性的干扰，并充分体现本文方法的优越性，将每个算法运行10次的实验结果进行综合分析，并给出每个方法的收敛曲线作为对比，从稳定性、收敛性、寻优能力等各个角度与其他方法进行比较，验证本文方法的有效性和优越性。

4.1 MPRM逻辑电路面积优化效果验证

平均面积节省率的计算方式为

$$A_{\text{save}}=\frac{(a_{\text{other}}-a_{\text{MSALO}})}{a_{\text{other}}}\times 100\% \tag{30}$$

式中：A_{save}为平均面积节省率；a_{other}为其他算法运行10次的平均值；a_{MSALO}为本文算法运行10次的平均值。

4种算法的逻辑电路面积优化对比数据如表1所示，其中，实验结果中的逻辑电路面积为逻辑电路门数量。可见，与其他算法相比，MSALO算法的平均电路面积节省率最小为0.64%，最大为70.14%，平均为27.96%。表明MSALO算法在搜索能力上优于ALO、ChOA、SAO算法。原因如下：①ALO、ChOA、SAO算法易陷入局部最优解，MSALO算法引入突围机制，在当前最优解附近执行更精细化的搜索，可以有效跳出局部最优解；②MSALO算法引入了双策略随机游走机制，结合哈里斯鹰算法的全局搜索能力和ALO算法的局部勘探能力，在算法迭代不同时期执行不同行为，增强了算法的全局搜索能力。

4.2 收敛性对比

随机选择6个测试电路的实验数据，绘制收敛曲线，对比本文方法与其他3种方法的收敛能力，如图4～图9所示。可以看出，由于MSALO

表1 4种算法的逻辑电路面积优化对比实验数据

Table 1 Comparison experimental data of logic circuit area optimization of four algorithms

电路名称	算法运行10次逻辑电路面积最优值				算法运行10次逻辑电路面积平均值				S_1/%	S_2/%	S_3/%
	MSALO	ALO	ChOA	SAO	MSALO	ALO	ChOA	SAO			
5xp1	37	37	37	37	37	39.7	37.4	40.8	6.80	1.07	9.31
exps	51	53	51	65	51	64.6	52.2	89.2	21.05	2.30	42.83
clip	260	260	260	260	264	265.7	268.8	271.1	0.64	17.86	2.62
br1	126	126	126	137	126	145.6	140.6	146.7	13.46	10.38	14.11
t3	113	129	123	161	119.2	155.4	139.6	176.6	23.29	14.61	32.50
alu1	7	7	7	8	7	7.4	7.1	12.2	5.41	1.41	42.62
misex3	933	1 876	1 377	2 484	933	2 469.6	1 892.2	3 124.8	62.22	50.69	70.14
table3	1 528	2 042	1 891	2 042	1 528	2 326.9	2 192.2	2 584.7	34.33	30.30	40.88
amd	309	375	346	383	323.1	415.7	374	439	22.28	13.61	26.40
gary	577	726	650	708	593.2	816.7	731.8	876.9	27.37	18.94	32.35
newtpla	35	44	47	59	35	66.5	63.4	82.1	47.37	44.79	57.37
pdc	185	185	185	230	185	259.4	252.3	297.9	28.68	26.67	37.90
table5	50	92	92	76	52.8	114	110.3	105	53.68	52.13	49.71

注：S_1、S_2、S_3分别为MSALO算法与ALO、ChOA、SAO算法相比的平均面积节省率。

算法利用正弦函数的变化趋势使蚂蚁的位置更新在算法迭代前期受轮盘赌蚁狮影响更大，在算法迭代后期受精英蚁狮影响更大。该自适应位置更新策略有效增强了算法的收敛能力，使 MSALO 算法在收敛速度、找到全局最优方面均优于 ALO、ChOA、SAO 算法。

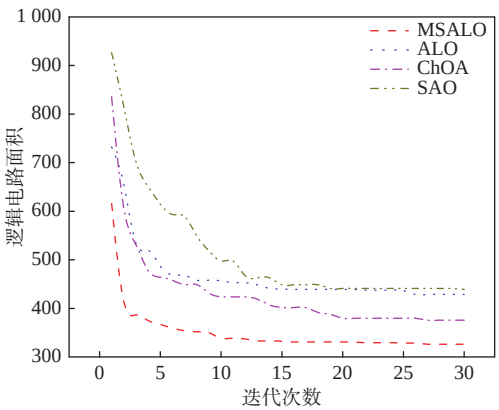


图 4　amd 电路收敛曲线

Fig. 4　Circuit convergence curves of amd

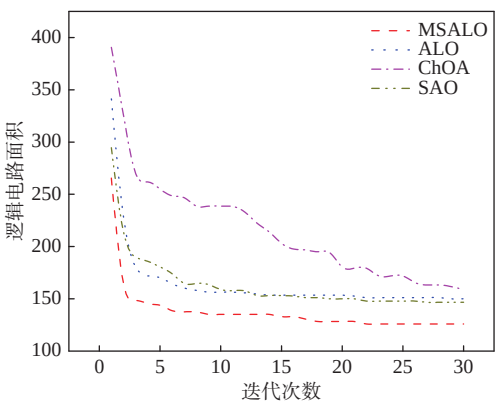


图 5　br1 电路收敛曲线

Fig. 5　Circuit convergence curves of br1

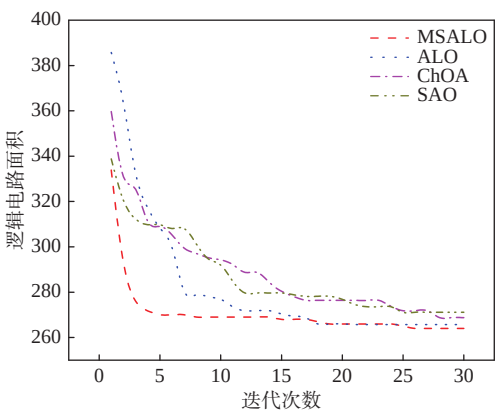


图 6　clip 电路收敛曲线

Fig. 6　Circuit convergence curves of clip

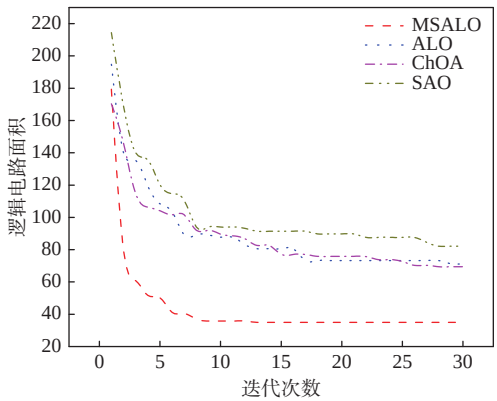


图 7　newtpla 电路收敛曲线

Fig. 7　Circuit convergence curves of newtpla

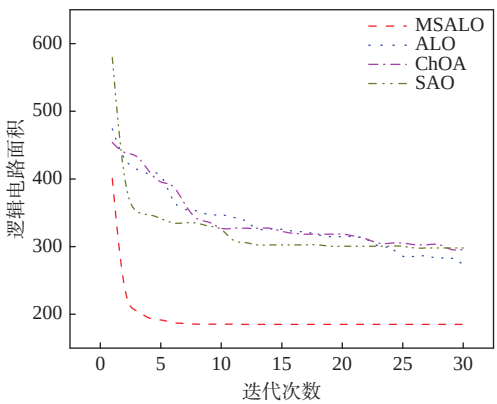


图 8　pdc 电路收敛曲线

Fig. 8　Circuit convergence curves of pdc

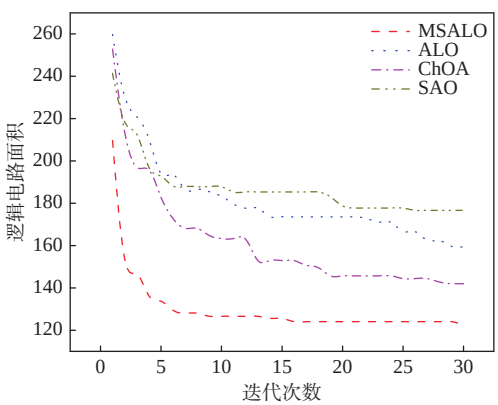


图 9　t3 电路收敛曲线

Fig. 9　Circuit convergence curves of t3

5　结　论

针对现有 MPRM 逻辑电路面积优化方法存在的问题，提出一种基于 MSALO 算法的 MPRM 逻辑电路面积优化方法，主要结论如下：

1）提出 MSALO 算法，针对 ALO 算法通过双策略随机游走机制、突围机制及基于正弦函数的自

适应蚂蚁位置更新策略来提高算法的性能。

2）将MSALO算法用于MPRM逻辑电路面积优化，可有效利用MSALO算法在MPRM逻辑电路面积优化空间中寻优。与其他优化方法的对比实验结果表明，本文方法在稳定性、寻优能力、收敛速度等方面均有较大优势。

参考文献（References）

[1] ZHOU Y H, HE Z X, CHEN C, et al. An efficient power optimization approach for fixed polarity Reed-Muller logic circuits based on metaheuristic optimization algorithm[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2022, 41(12): 5380-5393.

[2] HE Z X, PAN Y H, WANG K J, et al. Area optimization for MPRM logic circuits based on improved multiple disturbances fireworks algorithm[J]. Applied Mathematics and Computation, 2021, 399: 126008.

[3] FU Q, WANG P J, TONG N, et al. Integrated polarity optimization of MPRM circuits based on improved multi-objective particle swarm optimization[J]. Chinese Journal of Electronics, 2020, 29(5): 833-840.

[4] HE Z X, XIAO L M, GU F, et al. EDOA: an efficient delay optimization approach for mixed-polarity Reed-Muller logic circuits under the unit delay model[J]. Frontiers of Computer Science, 2019, 13(5): 1102-1115.

[5] QIN D G, HE Z X, ZHAO X J, et al. Area and power optimization for Fixed Polarity Reed-Muller logic circuits based on Multi-strategy Multi-objective Artificial Bee Colony algorithm[J]. Engineering Applications of Artificial Intelligence, 2023, 121: 105906.

[6] 周宇豪, 何振学, 梁新艺, 等. 基于BABFA的XNOR/OR电路面积优化[J]. 北京航空航天大学学报, 2022, 48(10): 2031-2039.
ZHOU Y H, HE Z X, LIANG X Y, et al. Optimization of XNOR/OR circuit area based on BABFA[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(10): 2031-2039 (in Chinese).

[7] WANG P J, LI K P, ZHANG H H. PMGA and its application in area and power optimization for ternary FPRM circuit[J]. Journal of Semiconductors, 2016, 37(1): 015007.

[8] 卜登立, 江建慧. 基于Pareto支配的MPRM电路面积与可靠性优化[J]. 电子学报, 2016, 44(11): 2653-2659.
BU D L, JIANG J H. Pareto dominance based area and reliability optimization of MPRM circuits[J]. Acta Electronica Sinica, 2016, 44(11): 2653-2659(in Chinese).

[9] 俞海珍, 汪鹏君, 张会红, 等. 基于三值多样性粒子群算法的MPRM电路综合优化[J]. 电子学报, 2017, 45(7): 1601-1607.
YU H Z, WANG P J, ZHANG H H, et al. Optimization of MPRM circuits based on ternary diversity particle swarm optimization[J]. Acta Electronica Sinica, 2017, 45(7): 1601-1607(in Chinese).

[10] 邵艺璇, 何振学, 周宇豪, 等. 基于M-AFSA的MPRM逻辑电路面积优化[J]. 北京航空航天大学学报, 2023, 49(3): 693-701.
SHAO Y X, HE Z X, ZHOU Y H, et al. Area optimization of MPRM circuits based on M-AFSA[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(3): 693-701(in Chinese).

[11] 潘钰华, 何振学, 王克俭. 基于改进烟花算法的MPRM逻辑电路面积优化[J]. 兰州大学学报 (自然科学版), 2021, 57(3): 416-422.
PAN Y H, HE Z X, WANG K J. Area optimization for MPRM logic circuits based on improved fireworks algorithm[J]. Journal of Lanzhou University (Natural Sciences), 2021, 57(3): 416-422(in Chinese).

[12] 卜登立. 基于混合遗传算法的MPRM最小化[J]. 浙江大学学报 (理学版), 2016, 43(2): 184-189.
BU D L. Hybrid genetic algorithm for MPRM minimization[J]. Journal of Zhejiang University (Science Edition), 2016, 43(2): 184-189(in Chinese).

[13] MIRJALILI S. The ant lion optimizer[J]. Advances in Engineering Software, 2015, 83: 80-98.

[14] KASHYAP A K, PARHI D R, KUMAR S, et al. Safe navigation of humanoid robot in cluttered terrain using ant lion optimizer tuned RA approach[C]//Proceedings of the Current Advances in Mechanical Engineering. Berlin: Springer, 2021: 997-1006.

[15] THAWKAR S, SHARMA S, KHANNA M, et al. Breast cancer prediction using a hybrid method based on Butterfly Optimization Algorithm and Ant Lion Optimizer[J]. Computers in Biology and Medicine, 2021, 139: 104968.

[16] KUMARI R, KUMAR R. Vehicular ant lion optimization algorithm (VALOA) for urban traffic management[C]//Proceedings of the Innovations in Computer Science and Engineering. Berlin: Springer, 2021: 99-114.

[17] ZHANG H L, CHEN M R, LI P S, et al. An improved Harris Hawks optimizer combined with extremal optimization[J]. International Journal of Machine Learning and Cybernetics, 2023, 14(3): 655-682.

[18] XU J, KANG X Y, FAN Z K, et al. Design of highly uniform magnetic field coils with wolf pack algorithm[J]. IEEE Sensors Journal, 2021, 21(4): 4412-4424.

[19] ABDULLAH M A, AB RASHID M F F, GHAZALLI Z. Optimization of assembly sequence planning using soft computing approaches: a review[J]. Archives of Computational Methods in Engineering, 2019, 26(2): 461-474.

[20] TANYILDIZI E, DEMIR G. Golden sine algorithm: a novel math-inspired algorithm[J]. Advances in Electrical and Computer Engineering, 2017, 17(2): 71-78.

[21] KHISHE M, MOSAVI M R. Chimp optimization algorithm[J]. Expert Systems with Applications, 2020, 149: 113338.

[22] DENG L Y, LIU S Y. Snow ablation optimizer: a novel metaheuristic technique for numerical optimization and engineering design[J]. Expert Systems with Applications, 2023, 225: 120069.

Area optimization approach for MPRM logic circuits based on multi-strategy synergy ant lion optimization algorithm

PAN Jiayi[1, 2], HE Zhenxue[1, 2, *], ZHAO Xiaojun[1, 2], HE Juncai[1, 2], ZHOU Yuhao[3], WANG Xiang[4]

(1. Intelligent Agricultural Equipment Research Institute, Hebei Agriculture University, Baoding 071001, China;
2. Key Laboratory of Agricultural Big Data of Hebei Province, Hebei Agriculture University, Baoding 071001, China;
3. School of Software Engineering, Tongji University, Shanghai 201804, China;
4. School of Electronic Information Engineering, Beihang University, Beijing 100191, China)

Abstract: In this paper, we propose a multi-strategy synergy ant lion optimization (MSALO) algorithm to address the problems of insufficient optimization efficacy of existing mixed polarity Reed-Muller (MPRM) circuit area optimization methods. A two-strategy random tour mechanism is used in the algorithm's random tour stage to address the issue of the ant lion optimization (ALO) algorithm's poor global search ability. A breakout mechanism is used for elite ant lion individuals to address the issue of poor local exploration ability. To expedite the convergence rate of the algorithm, an adaptive ant position update strategy based on the sine function is introduced. The MPRM circuit area optimization approach based on MSALO is proposed to search for the best polarity corresponding to the MPRM logic circuit with the smallest circuit area by using MSALO. The experimental results based on the Microelectronics Center of North Carolina (MCNC) benchmark test circuit demonstrate that the average area savings rate of the area optimization strategy based on MSALO may be increased by an average of 27.96% when compared to the current state-of-the-art swarm intelligence optimization algorithms.

Keywords: mixed polarity Reed-Muller; combinatorial optimization problem; swarm intelligent optimization; ant lion optimization; breakout mechanism

Received: 2024-04-26; **Accepted**: 2024-06-14; **Published Online**: 2024-07-08 11:19

URL: link.cnki.net/urlid/11.2625.V.20240705.1131.002

Foundation items: National Natural Science Foundation of China (62102130); Central Government Guides Local Science and Technology Development Fund Project (226Z0201G); Natural Science Foundation of Hebei Province (F2024204001,F2020204003); Hebei Youth Talents Support Project (BJ2019008); Science and Technology Research Projects of Higher Education Institutions in Hebei Province (QN2024138); Basic Scientific Research Funds Research Project of Hebei Provincial Colleges and Universities (KY2022073)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0489

飞艇抗随机干扰控制

常家铭[1]，李素兰[1,*]，段学超[1]，张伟[2]，汪晨阳[1]
(1. 西安电子科技大学 机电工程学院，西安 710071；2. 武汉第二船舶设计研究所，武汉 430064)

摘 要：飞艇具有能耗低、留空时间长等优势，被广泛应用于高低空任务中，由于工作在复杂恶劣的空中环境，经常会受到气流等随机干扰的影响，若不考虑这些干扰的作用，通常会导致飞艇性能的严重下降。针对飞艇抗随机干扰控制，结合随机系统相关理论，在动力学建模时考虑气流随机干扰项的影响，基于反步法、自适应技术及李雅普诺夫稳定性分析理论，设计了抗随机干扰的鲁棒控制器，并进行了数值仿真。仿真结果表明：在有随机气流的干扰下，所设计控制器仍能将飞艇快速部署到实验地点，并以要求的姿态悬停。

关 键 词：飞艇；动力学建模；随机干扰；随机控制；反步法
中图分类号：V249.1
文献标志码：A **文章编号**：1001-5965(2026)06-2092-10

飞艇作为一类常用的浮空器及低速飞行器，相比于其他飞行器具有载荷量大、留空时间长、覆盖范围广、能耗低、噪声污染小且安全可靠、无需跑道滑行等优势，因此，被广泛关注并应用于高空和低空的各种任务中[1-2]。对于空间太阳能电站发展规划中的浮空演示验证，飞艇作为搭载聚光器、发射天线等设备的浮空平台，需具备良好的长时间定点悬停和姿态调整能力，同时能够抵抗复杂的环境变化，如风干扰、温度变化等。因此，开展飞艇的抗扰控制研究至关重要。

已有研究都针对飞艇的确定性系统模型展开。文献 [3] 在假定飞艇模型准确已知的情况下设计了反步控制器抵抗外部干扰，并使用模糊自适应算法来逼近飞艇模型的不确定性。文献 [4] 为矢量反步滑模控制的设计和分析提供了一个通用的框架基础，并且提出了一种用于自主飞艇定位和跟踪的矢量反步滑模控制器，具有高度的鲁棒性。上述方法大多将外部干扰视作未知或已知的有界干扰，利用所设计控制器本身的鲁棒性来保证系统的稳定，也有部分研究了干扰的具体形式，如风场的扰动。文献 [5] 运用坐标变换的思想，在飞艇运动学方程中加入了风场模型，避免风超过干扰上界而导致系统不稳定的问题，但要求风速是已知且定常的，文献 [6] 在这种方法的基础上增加观测器来估计未知常值风。文献 [7] 利用飞艇的位置、姿态、模型等信息，设计了一个扩展卡尔曼滤波器来估计空气动力。文献 [8] 则进一步提出了一种无迹卡尔曼滤波器来统一对飞艇状态、模型不确定性和风扰动进行估计，性能优于扩展卡尔曼滤波器。文献 [9] 将风干扰和系统的动态耦合整体视为外部扰动，分别为俯仰通道和偏航通道设计了自抗扰控制器，性能优于传统的 PID 控制。上述研究均未考虑系统中随机干扰因素的影响，然而随着对建模准确性、控制精度要求的不断提高，考虑随机干扰影响并设计相应抗扰控制器的研究逐渐涌现。

文献 [10] 研究了非线性汽车悬架系统随机响应的求解问题，结合实际工况，建立了高斯白噪声模拟随机路面激励下时滞非线性汽车悬架系统的随机动力学模型，基于随机和确定性平均思想给出了系统解析解的计算方法。文献 [11-12] 分别以全驱动和

收稿日期：2024-06-27；录用日期：2024-08-15；网络出版时间：2024-08-26 13:06
网络出版地址：link.cnki.net/urlid/11.2625.V.20240824.1221.002
基金项目：陕西省自然科学基金 (2024JC-YBMS-451,2024JC-YBQN-0398)
*通信作者. E-mail：slli@xidian.edu.cn

引用格式：常家铭，李素兰，段学超，等. 飞艇抗随机干扰控制 [J]. 北京航空航天大学学报，2026，52（6）：2092-2101.
CHANG J M，LI S L，DUAN X C，et al. Anti-stochastic disturbance control of airship[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（6）：2092-2101 （in Chinese）.

欠驱动海上无人艇为研究对象, 考虑了海洋环境下海浪、风和洋流对无人艇带来的随机干扰影响, 基于反步法和李雅普诺夫函数设计了轨迹跟踪控制器, 仿真结果表明有效地抑制了随机干扰。文献 [13] 针对小型无人直升机系统, 考虑了随机干扰的同时还有系统不确定性、输出约束问题, 引入自适应神经网络逼近器来处理未知系统函数, 提出了一种鲁棒自适应方案。同样是对直升机系统, 文献 [14] 将基于扰动观测器的控制方法、反步控制方法和随机马尔可夫跳跃系统控制理论相结合设计了飞行控制器, 以保证直升机系统姿态和高度在合理的范围内跟踪预定值。文献 [15] 针对具有随机扰动的四旋翼无人机, 将有限时间理论与伊藤微分方程结合, 提出了一种考虑不可测状态和未知非线性函数的自适应模糊输出反馈控制方法。文献 [16] 针对具有随机扰动和未知参数不确定的临近空间飞行器, 设计了一种自适应神经网络控制器, 保证了临近空间飞行器姿态运动的跟踪性能。

以上研究表明随机抗干扰控制方法可有效提高车辆、无人机、海上无人艇等系统的控制性能, 但对飞艇系统的应用还相对有限。因此, 本文在飞艇动力学模型的基础上, 引入随机干扰项, 基于该模型设计了抗随机干扰鲁棒控制器, 通过仿真验证了控制方法的有效性和鲁棒性。

1　飞艇基础模型

本文研究的飞艇采用传统双椭球构型, 腹部对称安装 2 个多矢量推进器提供矢量推力。为便于对飞艇进行运动学和动力学分析, 建立图 1 所示坐标系。$O_e x_e y_e z_e$为地面坐标系S_e, $O_b x_b y_b z_b$为艇体坐标系S_b, $O_v x_v y_v z_v$为速度坐标系S_v, 均符合右手法则。

飞艇的运动参数包括: $\boldsymbol{\eta}_1=[x \quad y \quad z]^{\mathrm{T}}$表示飞艇的位置向量, 其中, x、y、z分别表示飞艇体积中心在地面坐标系S_e中$O_e x_e$、$O_e y_e$、$O_e z_e$方向的位移; $\boldsymbol{\eta}_2=[\phi \quad \theta \quad \psi]^{\mathrm{T}}$为飞艇的姿态角向量, 其中, ϕ为滚转角, θ为俯仰角, ψ为偏航角; $\boldsymbol{\xi}_1=[u \quad v \quad w]^{\mathrm{T}}$为飞艇在艇体坐标系$S_b$中的线速度向量; $\boldsymbol{\xi}_2=[p \quad q \quad r]^{\mathrm{T}}$为飞艇在艇体坐标系$S_b$中的角速度向量; $\boldsymbol{\eta}=[\boldsymbol{\eta}_1 \quad \boldsymbol{\eta}_2]^{\mathrm{T}}$为位姿向量; $\boldsymbol{\xi}=[\boldsymbol{\xi}_1 \quad \boldsymbol{\xi}_2]^{\mathrm{T}}$为速度向量; α为攻角; β为侧滑角。

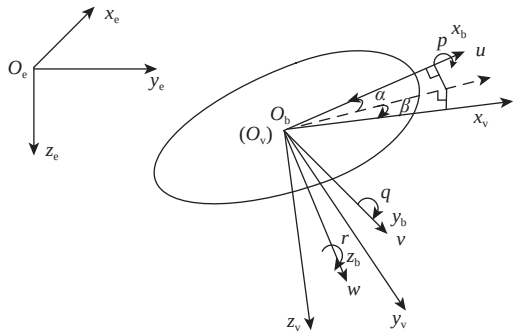


图 1　飞艇坐标系示意图

Fig. 1　Schematic diagram of the airship coordinate system

为了便于简化分析, 在后续建模及控制时作如下假设:

1) 飞艇的浮心和体心重合;

2) 飞艇的重心在纵向对称平面, 且位置在控制过程中不发生变化;

3) 飞艇为刚体, 忽略其弹性效应;

4) 2 矢量推力器可以 3 轴任意指向可控, 且全姿态可控。

1.1　飞艇运动学模型

飞艇在地面坐标系S_e和艇体坐标系S_b位移和姿态角变化的关系, 可由如下运动学方程表示:

$$\begin{bmatrix} \dot{\boldsymbol{\eta}}_1 \\ \dot{\boldsymbol{\eta}}_2 \end{bmatrix}=\begin{bmatrix} \boldsymbol{R}_1 & \boldsymbol{0}_{3\times3} \\ \boldsymbol{0}_{3\times3} & \boldsymbol{R}_2 \end{bmatrix}\begin{bmatrix} \boldsymbol{\xi}_1 \\ \boldsymbol{\xi}_2 \end{bmatrix} \tag{1}$$

即

$$\dot{\boldsymbol{\eta}}=\boldsymbol{R}(\boldsymbol{\eta})\boldsymbol{\xi}$$

式中: $\boldsymbol{R}_1=\boldsymbol{C}_b^e$为从$S_b$到$S_e$的坐标转换矩阵; $\boldsymbol{R}_2$为姿态关系矩阵; $\boldsymbol{R}(\boldsymbol{\eta})$为与位姿相关的运动学变换矩阵。

1.2　飞艇动力学模型

对飞艇进行受力分析, 运用牛顿欧拉公式可得到 6 自由度动力学模型为

$$\boldsymbol{M}\dot{\boldsymbol{\xi}}=\boldsymbol{F}_{\mathrm{f}}+\boldsymbol{F}_{\mathrm{A}}+\boldsymbol{F}_{\mathrm{G}}+\boldsymbol{F}_{\mathrm{B}}+\boldsymbol{F}_{\mathrm{T}} \tag{2}$$

式中:

$$\boldsymbol{M}=\begin{bmatrix} m+m_{11} & 0 & 0 & 0 & mz_{\mathrm{G}} & 0 \\ 0 & m+m_{22} & 0 & -mz_{\mathrm{G}} & 0 & mx_{\mathrm{G}} \\ 0 & 0 & m+m_{33} & 0 & -mx_{\mathrm{G}} & 0 \\ 0 & -mz_{\mathrm{G}} & 0 & I_x+m_{44} & 0 & -I_{xz} \\ mz_{\mathrm{G}} & 0 & -mx_{\mathrm{G}} & 0 & I_y+m_{55} & 0 \\ 0 & mx_{\mathrm{G}} & 0 & -I_{xz} & 0 & I_z+m_{66} \end{bmatrix} \tag{3}$$

其中: $\boldsymbol{M}$为质量矩阵; m为飞艇质量; $m_{11}=k_{\mathrm{m1}}m_{\mathrm{air}}$, $m_{22}=m_{33}=k_{\mathrm{m2}}m_{\mathrm{air}}$, $m_{44}=0, m_{55}=k_{\mathrm{m3}}I_y$, $m_{66}=k_{\mathrm{m3}}I_z$为附加质量, k_{m1}、k_{m2}、k_{m3}为势流理论分析得到的椭圆惯性因子, m_{air}为飞艇所在高度处排开空气的质

量，需要根据所处高度的大气密度乘以自身体积计算；$\boldsymbol{r}_{\mathrm{G}}=[x_{\mathrm{G}},y_{\mathrm{G}},z_{\mathrm{G}}]^{\mathrm{T}}$为重心坐标；$I_x$、$I_y$、$I_z$为在艇体坐标系3个坐标轴的转动惯量；$I_{xz}$为对$x_{\mathrm{b}}O_{\mathrm{b}}z_{\mathrm{b}}$平面的惯性积。

$$\boldsymbol{F}_{\mathrm{f}}=\begin{bmatrix}(m+m_{22})vr-(m+m_{33})wq+m\left[x_{\mathrm{G}}\left(q^2+r^2\right)-z_{\mathrm{G}}pr\right]\\-(m+m_{11})ur+(m+m_{33})wp-m(x_{\mathrm{G}}pq+z_{\mathrm{G}}qr)\\(m+m_{11})uq-(m+m_{22})vp+m\left[-x_{\mathrm{G}}pr+z_{\mathrm{G}}\left(p^2+q^2\right)\right]\\\left[\left(I_y+m_{55}\right)-\left(I_z+m_{66}\right)\right]qr+I_{xz}pq+m\left[z_{\mathrm{G}}(ur-wp)\right]\\\left[(I_z+m_{66})-(I_x+m_{44})\right]pr-I_{xz}\left(p^2-r^2\right)+m\left[x_{\mathrm{G}}(vp-uq)+z_{\mathrm{G}}(vr-wq)\right]\\\left[(I_x+m_{44})-\left(I_y+m_{55}\right)\right]pq-I_{xz}qr+m\left[x_{\mathrm{G}}(wp-ur)\right]\end{bmatrix}\tag{4}$$

式中：$\boldsymbol{F}_{\mathrm{f}}$为动态力项。

$$\boldsymbol{F}_{\mathrm{A}}=\begin{bmatrix}\boldsymbol{C}_{\mathrm{v}}^{\mathrm{b}}\boldsymbol{F}_{\mathrm{A0}}\\\boldsymbol{M}_{\mathrm{A}}\end{bmatrix}\tag{5}$$

式中：$\boldsymbol{F}_{\mathrm{A}}$为空气动力项。其中，$\boldsymbol{F}_{\mathrm{A0}}$和$\boldsymbol{M}_{\mathrm{A}}$分别为气动力和气动力矩，均是飞艇空速的函数，由于空速为飞艇地速与风速的矢量差，故存在风扰动时，空速会发生变化，进而导致$\boldsymbol{F}_{\mathrm{A0}}$和$\boldsymbol{M}_{\mathrm{A}}$的变化；$\boldsymbol{C}_{\mathrm{v}}^{\mathrm{b}}$为从$S_{\mathrm{v}}$到$S_{\mathrm{b}}$的坐标转换矩阵。

$$\boldsymbol{F}_{\mathrm{G}}=\left[\boldsymbol{C}_{\mathrm{e}}^{\mathrm{b}}\boldsymbol{G}_{\mathrm{E}}\quad\boldsymbol{r}_{\mathrm{G}}\times\left(\boldsymbol{C}_{\mathrm{e}}^{\mathrm{b}}\boldsymbol{G}_{\mathrm{E}}\right)\right]^{\mathrm{T}}\tag{6}$$

式中：$\boldsymbol{F}_{\mathrm{G}}$为重力项。其中，$\boldsymbol{G}_{\mathrm{E}}$为飞艇所受重力；$\boldsymbol{C}_{\mathrm{e}}^{\mathrm{b}}$为从$S_{\mathrm{e}}$到$S_{\mathrm{b}}$的坐标转换矩阵。

$$\boldsymbol{F}_{\mathrm{B}}=\left[\boldsymbol{C}_{\mathrm{e}}^{\mathrm{b}}\boldsymbol{F}_{\mathrm{BE}}\quad 0\quad 0\quad 0\right]^{\mathrm{T}}\tag{7}$$

式中：$\boldsymbol{F}_{\mathrm{B}}$项为浮力项。其中，$\boldsymbol{F}_{\mathrm{BE}}$为飞艇所受浮力。

$\boldsymbol{F}_{\mathrm{T}}$为推进系统产生的推力项，把推力表示成外部控制输入项，即$\boldsymbol{U}=\boldsymbol{F}_{\mathrm{T}}$，通过后述控制方法求出控制律$\boldsymbol{U}$以确定。

2 随机干扰控制模型

2.1 气流干扰力

飞艇常受到一定的气流扰动，可以用风对飞艇产生的力和力矩表示，其具体形式如下[17]：

$$\boldsymbol{F}_{\mathrm{W}}=\begin{bmatrix}\boldsymbol{M}_{\mathrm{Ba}}\dot{\boldsymbol{V}}_{\mathrm{w}}+\boldsymbol{\xi}_2\times\boldsymbol{M}_{\mathrm{Ba}}\boldsymbol{V}_{\mathrm{w}}\\\boldsymbol{J}_{\mathrm{Ba}}\dot{\boldsymbol{\omega}}_{\mathrm{w}}+\boldsymbol{\xi}_2\times\boldsymbol{J}_{\mathrm{Ba}}\boldsymbol{\omega}_{\mathrm{w}}\end{bmatrix}\tag{8}$$

式中：$\boldsymbol{M}_{\mathrm{Ba}}=\mathrm{diag}(m_{\mathrm{air}}+m_{11},m_{\mathrm{air}}+m_{22},m_{\mathrm{air}}+m_{33})$；$\boldsymbol{J}_{\mathrm{Ba}}=\mathrm{diag}(I_{\mathrm{xair}}+m_{44},I_{\mathrm{yair}}+m_{55},I_{\mathrm{zair}}+m_{66})$，$I_{\mathrm{xair}}$、$I_{\mathrm{yair}}$、$I_{\mathrm{zair}}$为飞艇排开空气的转动惯量；$\boldsymbol{V}_{\mathrm{w}}=[u_{\mathrm{w}}\quad v_{\mathrm{w}}\quad w_{\mathrm{w}}]^{\mathrm{T}}$和$\boldsymbol{\omega}_{\mathrm{w}}=\left[p_{\mathrm{w}}\quad q_{\mathrm{w}}\quad r_{\mathrm{w}}\right]^{\mathrm{T}}$分别为风对于艇体坐标系的速度向量和角速度向量。

整理得如下形式：

$$\boldsymbol{F}_{\mathrm{W}}=\boldsymbol{M}_{\mathrm{w}}\dot{\boldsymbol{V}}_{\mathrm{d}}+\boldsymbol{C}_{\mathrm{w}}(\boldsymbol{\xi})(\boldsymbol{M}_{\mathrm{w}}\boldsymbol{V}_{\mathrm{d}})\tag{9}$$

式中：$\boldsymbol{M}_{\mathrm{w}}=\begin{bmatrix}\boldsymbol{M}_{\mathrm{Ba}}&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{J}_{\mathrm{Ba}}\end{bmatrix}$；$\boldsymbol{C}_{\mathrm{w}}(\boldsymbol{\xi})=\begin{bmatrix}\boldsymbol{C}_1&\boldsymbol{0}\\\boldsymbol{0}&\boldsymbol{C}_1\end{bmatrix}$，$\boldsymbol{C}_1=\begin{bmatrix}0&-r&q\\r&0&-p\\-q&p&0\end{bmatrix}$；$\boldsymbol{V}_{\mathrm{d}}=[\boldsymbol{V}_{\mathrm{w}}\quad\boldsymbol{\omega}_{\mathrm{w}}]^{\mathrm{T}}$。将式(9)加入式(2)等号右侧完善飞艇的动力学模型。

2.2 随机干扰下飞艇控制模型

风速相关项可以写成确定分量和随机分量的形式，即$\dot{\boldsymbol{V}}_{\mathrm{d}}=\bar{\dot{\boldsymbol{V}}}_{\mathrm{d}}+\tilde{\dot{\boldsymbol{V}}}_{\mathrm{d}}$和$\boldsymbol{V}_{\mathrm{d}}=\bar{\boldsymbol{V}}_{\mathrm{d}}+\tilde{\boldsymbol{V}}_{\mathrm{d}}$，根据随机理论[18]，随机分量可看作高斯随机扰动，则原动力学模型可以写成

$$\begin{aligned}\dot{\boldsymbol{\xi}}=&\boldsymbol{M}^{-1}(\boldsymbol{F}_{\mathrm{f}}+\boldsymbol{F}_{\mathrm{A}}+\boldsymbol{F}_{\mathrm{G}}+\boldsymbol{F}_{\mathrm{B}})+\boldsymbol{M}^{-1}\boldsymbol{U}+\\&\boldsymbol{M}^{-1}\left(\boldsymbol{M}_{\mathrm{w}}\bar{\dot{\boldsymbol{V}}}_{\mathrm{d}}+\boldsymbol{C}_{\mathrm{w}}(\boldsymbol{\xi})\boldsymbol{M}_{\mathrm{w}}\bar{\boldsymbol{V}}_{\mathrm{d}}\right)+\\&\boldsymbol{M}^{-1}\left(\boldsymbol{M}_{\mathrm{w}}\tilde{\dot{\boldsymbol{V}}}_{\mathrm{d}}+\boldsymbol{C}_{\mathrm{w}}(\boldsymbol{\xi})\boldsymbol{M}_{\mathrm{w}}\tilde{\boldsymbol{V}}_{\mathrm{d}}\right)\end{aligned}\tag{10}$$

为便于后续控制器的设计，将式(10)中的部分项定义为

$$\begin{cases}\boldsymbol{f}\left(\boldsymbol{\xi},\boldsymbol{\eta}\right)=\boldsymbol{M}^{-1}(\boldsymbol{F}_{\mathrm{f}}+\boldsymbol{F}_{\mathrm{A}}+\boldsymbol{F}_{\mathrm{G}}+\boldsymbol{F}_{\mathrm{B}})\\\boldsymbol{U}^{*}=\boldsymbol{M}^{-1}\boldsymbol{U}\\\boldsymbol{\Phi}(\boldsymbol{\xi})=\boldsymbol{M}^{-1}[\boldsymbol{M}_{\mathrm{w}},\boldsymbol{C}_{\mathrm{w}}(\boldsymbol{\xi})\boldsymbol{M}_{\mathrm{w}}]\\\boldsymbol{\sigma}=\begin{bmatrix}\tilde{\dot{\boldsymbol{V}}}_{\mathrm{d}}&\tilde{\boldsymbol{V}}_{\mathrm{d}}\end{bmatrix}^{\mathrm{T}}\end{cases}\tag{11}$$

同时根据随机系统理论[18]，引入$\boldsymbol{\Delta}_1(t)$和$\boldsymbol{\Delta}_2(t)$分别表示随机项$\tilde{\dot{\boldsymbol{V}}}_{\mathrm{d}}$和$\tilde{\boldsymbol{V}}_{\mathrm{d}}$的强度，定义$\boldsymbol{\Delta}(t)=\mathrm{diag}(\boldsymbol{\Delta}_1(t),\boldsymbol{\Delta}_2(t))$，最终可以得到飞艇的非线性随机系统模型Ⅰ：

$$\begin{cases}\dot{\boldsymbol{\eta}}=\boldsymbol{R}\left(\boldsymbol{\eta}\right)\boldsymbol{\xi}\\\mathrm{d}\boldsymbol{\xi}=\left(\boldsymbol{f}\left(\boldsymbol{\xi},\boldsymbol{\eta}\right)+\boldsymbol{U}^{*}+\boldsymbol{\Phi}(\boldsymbol{\xi})\boldsymbol{\sigma}\right)\mathrm{d}t+\boldsymbol{\Phi}(\boldsymbol{\xi})\boldsymbol{\Delta}(t)\,\mathrm{d}\boldsymbol{\omega}\end{cases}\tag{12}$$

式中：$\boldsymbol{\omega}$为12维独立标准维纳过程向量，同时根据飞艇的物理特性及可控性，并考虑实际风干扰的强度是有限的，即存在未知正常数σ_{m}和Δ_{m}使得$\|\boldsymbol{\sigma}\|\leqslant\sigma_{\mathrm{m}}$，$\|\boldsymbol{\Delta}(t)\|\leqslant\Delta_{\mathrm{m}}$。之后将在此非线性随机系统模型Ⅰ的基础上，以定点悬停为控制目标，并且希望能够控制姿态角到期望值，即$\boldsymbol{\eta}\to\boldsymbol{\eta}_{\mathrm{d}}$，运用4次李雅普诺夫函数、反步法和自适应技术设计抗随机干扰鲁棒控制器，控制原理框图如图2所示。

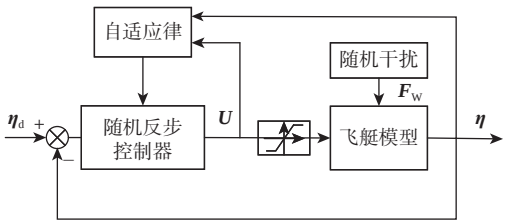


图 2　鲁棒自适应悬停控制器控制原理框图

Fig. 2　Schematic diagram of the control principle for a robust adaptive hovering controller

3　抗随机干扰鲁棒控制器设计

3.1　相关定义及引理

考虑如下伊藤型随机微分方程(stochastic differential equation, SDE)[19]:

$$\begin{cases} \mathrm{d}\boldsymbol{X} = \boldsymbol{f}(\boldsymbol{X})\mathrm{d}t + \boldsymbol{g}(\boldsymbol{X})\mathrm{d}\tilde{\boldsymbol{\omega}} \\ \boldsymbol{X}(0) = \boldsymbol{X}_0 \in \mathbf{R}^n \end{cases} \tag{13}$$

式中: $\boldsymbol{X} \in \mathbf{R}^n$为系统状态; $\tilde{\boldsymbol{\omega}}$为定义在完备概率空间上的 m 维标准维纳过程; 非线性函数 $\boldsymbol{f}(\boldsymbol{X})$: $\mathbf{R}^n \to \mathbf{R}^n$ 和 $\boldsymbol{g}(\boldsymbol{X})$: $\mathbf{R}^n \to \mathbf{R}^{n\times m}$为 Borel 可测函数且关于 $\boldsymbol{X}$满足局部 Lipschitz 条件。

定义 1[19]　对任意给定的函数 $V \in C^2(\mathbf{R}^n, \mathbf{R}_+)$及相应的 SDE, 定义无穷微分算子$\mathcal{L}$为如下形式:

$$\mathcal{L}V = \frac{\partial V}{\partial \boldsymbol{X}}\boldsymbol{f}(\boldsymbol{X}) + \frac{1}{2}\mathrm{tr}\left(\boldsymbol{g}^{\mathrm{T}}(\boldsymbol{X})\frac{\partial^2 V}{\partial \boldsymbol{X}^2}\boldsymbol{g}(\boldsymbol{X})\right) \tag{14}$$

式中: $\mathrm{tr}(\cdot)$表示矩阵的迹。

定义 2[20]　对于某个紧集 $\Gamma \in \mathbf{R}^n$及初始状态 $\boldsymbol{X}(t_0) = \boldsymbol{X}_0 \in \Gamma$, 存在一个常数$\tau > 0$和一个时间常量 $T = T(\tau, \boldsymbol{X}_0)$, 对于所有的$t > t_0 + T$, 使得$E(\|\boldsymbol{X}(t)\|^b) < \tau$, 则称随机非线性系统式 (13) 的解 $\boldsymbol{X}(t)$为 b 阶矩意义下半全局一致最终有界。

引理 1[21]　考虑上述随机非线性系统式 (13), 如果存在一个正定、径向无界、二阶连续可微的李雅普诺夫函数 $V(\boldsymbol{X})$: $\mathbf{R}^n \to \mathbf{R}_+$, 以及正常数$a_0 > 0$和 $b_0 > 0$, K_∞函数 $\bar{\alpha}_1$和 $\bar{\alpha}_2$, 使得当 $t > t_0$时, 对所有的 $\boldsymbol{X} \in \mathbf{R}^n$有

$$\bar{\alpha}_1(\|\boldsymbol{X}\|) \leqslant V(\boldsymbol{X}) \leqslant \bar{\alpha}_2(\|\boldsymbol{X}\|) \tag{15}$$

$$\mathcal{L}V(\boldsymbol{X}) \leqslant -a_0 V(\boldsymbol{X}) + b_0 \tag{16}$$

则对任意初始状态 $\boldsymbol{X}_0 \in \mathbf{R}^n$, SDE 存在唯一强解 $\boldsymbol{X}(t)$, 并且依概率有界, 同时有

$$E[V(\boldsymbol{X})] \leqslant V(\boldsymbol{X}_0)\mathrm{e}^{-a_0(t-t_0)} + b_0/a_0 \qquad \forall t \geqslant t_0 \tag{17}$$

引理 2(Young 不等式)[22]　对于$\forall(\gamma, \Theta) \in \mathbf{R}^2$, 有如下不等式成立:

$$\gamma\Theta \leqslant \frac{\varepsilon^\iota}{\iota}|\gamma|^\iota + \frac{1}{\zeta\varepsilon^\zeta}|\Theta|^\zeta \tag{18}$$

式中: 需同时满足条件$\varepsilon > 0$, $\iota > 1$, $\zeta > 1$和 $(\iota-1)\cdot(\zeta-1) = 1$。

3.2　控制器设计

首先定义跟踪误差:

$$\boldsymbol{e}_1 = \boldsymbol{\eta} - \boldsymbol{\eta}_{\mathrm{d}} \tag{19}$$

式中: $\boldsymbol{\eta}_{\mathrm{d}}$为飞艇期望位姿向量, 即要跟踪的目标。

对式 (19) 求导并结合式 (1) 有

$$\dot{\boldsymbol{e}}_1 = \boldsymbol{R}(\boldsymbol{\eta})\boldsymbol{\xi} - \dot{\boldsymbol{\eta}}_{\mathrm{d}} \tag{20}$$

同时定义跟踪误差:

$$\boldsymbol{e}_2 = \boldsymbol{\xi} - \boldsymbol{\alpha}_{\mathrm{v}} \tag{21}$$

式中: $\boldsymbol{\alpha}_{\mathrm{v}}$为虚拟控制量。

代入式 (19) 有

$$\dot{\boldsymbol{e}}_1 = \boldsymbol{R}(\boldsymbol{\eta})(\boldsymbol{\alpha}_{\mathrm{v}} + \boldsymbol{e}_2) - \dot{\boldsymbol{\eta}}_{\mathrm{d}} \tag{22}$$

设计虚拟控制律为

$$\boldsymbol{\alpha}_{\mathrm{v}} = \boldsymbol{R}(\boldsymbol{\eta})^{-1}(-\boldsymbol{K}_1\boldsymbol{e}_1 + \dot{\boldsymbol{\eta}}_{\mathrm{d}}) \tag{23}$$

式中: $\boldsymbol{K}_1$为设计的正定矩阵。

选取李雅普诺夫候选函数:

$$V_1 = \frac{1}{4}\left(\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{e}_1\right)^2 \tag{24}$$

根据定义 1 并结合式 (22) 和式 (23) 有

$$\mathcal{L}V_1 = -\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{e}_1\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{K}_1\boldsymbol{e}_1 + \boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{e}_1\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{\eta})\boldsymbol{e}_2 \tag{25}$$

结合随机动力学模型 I 对式 (21) 求微分有

$$\mathrm{d}\boldsymbol{e}_2 = \left(\boldsymbol{f}(\boldsymbol{\xi}, \boldsymbol{\eta}) + \boldsymbol{U}^* + \boldsymbol{\Phi}(\boldsymbol{\xi})\boldsymbol{\sigma} - \dot{\boldsymbol{\alpha}}_{\mathrm{v}}\right)\mathrm{d}t + \boldsymbol{\Phi}(\boldsymbol{\xi})\boldsymbol{\Delta}(t)\mathrm{d}\boldsymbol{\omega} \tag{26}$$

为了设计控制律$\boldsymbol{U}$实现整个系统随机鲁棒的目标, 考虑如下李雅普诺夫候选函数:

$$V_2 = V_1 + \frac{1}{4}\left(\boldsymbol{e}_2^{\mathrm{T}}\boldsymbol{e}_2\right)^2 + \frac{1}{2\varepsilon_1}\tilde{\sigma}^2 + \frac{1}{2\varepsilon_2}\tilde{\Delta}^2 \tag{27}$$

式中: ε_1、ε_2为设计的正常数; $\tilde{\sigma} = \sigma_{\mathrm{m}} - \hat{\sigma}$, $\tilde{\Delta} = \Delta_{\mathrm{m}} - \hat{\Delta}$为干扰参数估计误差, $\hat{\sigma}$和$\hat{\Delta}$为参数估计值。

由定义 1 对式 (27) 求无限小生成元, 有

$$\begin{aligned}\mathcal{L}V_2 = {} & -\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{e}_1\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{K}_1\boldsymbol{e}_1 + \boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{e}_1\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{\eta})\boldsymbol{e}_2 + \boldsymbol{e}_2^{\mathrm{T}}\boldsymbol{e}_2\boldsymbol{e}_2^{\mathrm{T}}\left(\boldsymbol{f}(\boldsymbol{\xi}, \boldsymbol{\eta}) + \right. \\ & \left. \boldsymbol{U}^* + \boldsymbol{\Phi}(\boldsymbol{\xi})\boldsymbol{\sigma} - \dot{\boldsymbol{\alpha}}_{\mathrm{v}}\right) - \frac{1}{\varepsilon_1}\tilde{\sigma}\dot{\hat{\sigma}} - \frac{1}{\varepsilon_2}\tilde{\Delta}\dot{\hat{\Delta}} + \\ & \frac{1}{2}\mathrm{tr}\left((\boldsymbol{\Phi}(\boldsymbol{\xi})\boldsymbol{\Delta}(t))^{\mathrm{T}}\left(2\boldsymbol{e}_2\boldsymbol{e}_2^{\mathrm{T}} + \boldsymbol{e}_2^{\mathrm{T}}\boldsymbol{e}_2\boldsymbol{I}_{6\times 6}\right)\boldsymbol{\Phi}(\boldsymbol{\xi})\boldsymbol{\Delta}(t)\right)\end{aligned} \tag{28}$$

由引理 2(即 Young 不等式)得

$$\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{e}_1\boldsymbol{e}_1^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{\eta})\boldsymbol{e}_2 \leqslant \frac{3k_1^{4/3}}{4}\|\boldsymbol{e}_1\|^4 + \frac{1}{4k_1^4}\left\|\boldsymbol{R}(\boldsymbol{\eta})\right\|^4\|\boldsymbol{e}_2\|^4 \tag{29}$$

$$\boldsymbol{e}_2^{\mathrm{T}}\boldsymbol{e}_2\boldsymbol{e}_2^{\mathrm{T}}\boldsymbol{\Phi}(\boldsymbol{\xi})\boldsymbol{\sigma} \leqslant \frac{3k_2^{4/3}}{4}\|\boldsymbol{e}_2\|^4\|\boldsymbol{\Phi}(\boldsymbol{\xi})\|^{4/3}\sigma_{\mathrm{m}} + \frac{1}{4k_2^4}\sigma_{\mathrm{m}} \tag{30}$$

根据 Frobenius 范数的定义、矩阵范数相容性和 Young 不等式有

$$\frac{1}{2}\mathrm{tr}\left((\boldsymbol{\Phi}(\boldsymbol{\xi})\boldsymbol{\Delta}(t))^{\mathrm{T}}\left(2\boldsymbol{e}_2\boldsymbol{e}_2^{\mathrm{T}}+\boldsymbol{e}_2^{\mathrm{T}}\boldsymbol{e}_2\boldsymbol{I}_{6\times6}\right)\boldsymbol{\Phi}(\boldsymbol{\xi})\boldsymbol{\Delta}(t)\right)\leqslant \frac{3}{2}\left(\frac{k_3^2}{2}\|\boldsymbol{\Phi}(\boldsymbol{\xi})\|^4\|\boldsymbol{e}_2\|^4\Delta_{\mathrm{m}}+\frac{\Delta_{\mathrm{m}}}{2k_3^2}\right) \tag{31}$$

式中：k_1、k_2、k_3为正常数。

将式(29)～式(31)代入到式(28)中，有

$$\begin{aligned}\mathcal{L}V_2\leqslant&-\left(\lambda_{\min}(\boldsymbol{K}_1)-\frac{3k_1^{4/3}}{4}\right)\|\boldsymbol{e}_1\|^4+\\&\boldsymbol{e}_2^{\mathrm{T}}\boldsymbol{e}_2\boldsymbol{e}_2^{\mathrm{T}}\left[\left(\frac{\|\boldsymbol{R}(\boldsymbol{\eta})\|^4}{4k_1^4}+\frac{3k_2^{4/3}}{4}\|\boldsymbol{\Phi}(\boldsymbol{\xi})\|^{4/3}\sigma_{\mathrm{m}}+\right.\right.\\&\left.\left.\frac{3k_3^2}{4}\|\boldsymbol{\Phi}(\boldsymbol{\xi})\|^4\Delta_{\mathrm{m}}\right)\boldsymbol{e}_2\right]+\\&\boldsymbol{e}_2^{\mathrm{T}}\boldsymbol{e}_2\boldsymbol{e}_2^{\mathrm{T}}\left[\boldsymbol{f}(\boldsymbol{\xi},\boldsymbol{\eta}_2)+\boldsymbol{U}^*-\dot{\boldsymbol{\alpha}}_{\mathrm{v}}\right]+\\&\frac{1}{4k_2{}^4}\sigma_{\mathrm{m}}+\frac{3}{4k_3^2}\Delta_{\mathrm{m}}-\frac{1}{\varepsilon_1}\tilde{\sigma}\dot{\hat{\sigma}}-\frac{1}{\varepsilon_2}\tilde{\Delta}\dot{\hat{\Delta}}\end{aligned} \tag{32}$$

式中：$\lambda_{\min}()$表示矩阵的最小特征值。

由此设计如下控制律和自适应更新律：

$$\begin{aligned}\boldsymbol{U}=\boldsymbol{M}&\left[-\boldsymbol{K}_2\boldsymbol{e}_2-\left(\frac{\|\boldsymbol{R}(\boldsymbol{\eta})\|^4}{4k_1^4}+\frac{3k_2^{4/3}}{4}\|\boldsymbol{\Phi}(\boldsymbol{\xi})\|^{4/3}\hat{\sigma}+\right.\right.\\&\left.\left.\frac{3k_3^2}{4}\|\boldsymbol{\Phi}(\boldsymbol{\xi})\|^4\hat{\Delta}\right)\boldsymbol{e}_2\right]-\boldsymbol{M}\left[\boldsymbol{f}(\boldsymbol{\xi},\boldsymbol{\eta}_2)+\dot{\boldsymbol{\alpha}}_{\mathrm{v}}\right]\end{aligned} \tag{33}$$

$$\dot{\hat{\sigma}}=\varepsilon_1\left(\frac{3k_2^{4/3}}{4}\|\boldsymbol{e}_2\|^4\|\boldsymbol{\Phi}(\boldsymbol{\xi})\|^{4/3}-\lambda_1\hat{\sigma}\right) \tag{34}$$

$$\dot{\hat{\Delta}}=\varepsilon_2\left(\frac{3k_3^2}{4}\|\boldsymbol{e}_2\|^4\|\boldsymbol{\Phi}(\boldsymbol{\xi})\|^4-\lambda_2\hat{\Delta}\right) \tag{35}$$

式中：$\boldsymbol{K}_2$为设计的正定矩阵；λ_1、λ_2为设计的正常数。

下面证明闭环系统的稳定性。

将上述控制律和自适应律代入式(32)中，有

$$\begin{aligned}\mathcal{L}V_2\leqslant&-\left(\lambda_{\min}(\boldsymbol{K}_1)-\frac{3k_1^{4/3}}{4}\right)\|\boldsymbol{e}_1\|^4-\lambda_{\min}(\boldsymbol{K}_2)\|\boldsymbol{e}_2\|^4+\\&\frac{1}{4k_2^4}\sigma_{\mathrm{m}}+\frac{3}{4k_3^2}\Delta_{\mathrm{m}}+\lambda_1\tilde{\sigma}\hat{\sigma}+\lambda_2\tilde{\Delta}\hat{\Delta}\end{aligned} \tag{36}$$

又有如下不等式成立：

$$\tilde{\sigma}\hat{\sigma}=\frac{1}{2}\left(\sigma_{\mathrm{m}}^2-\tilde{\sigma}^2-\hat{\sigma}^2\right)\leqslant\frac{1}{2}\left(\sigma_{\mathrm{m}}^2-\tilde{\sigma}^2\right) \tag{37}$$

$$\tilde{\Delta}\hat{\Delta}=\frac{1}{2}\left(\Delta_{\mathrm{m}}^2-\tilde{\Delta}^2-\hat{\Delta}^2\right)\leqslant\frac{1}{2}\left(\Delta_{\mathrm{m}}^2-\tilde{\Delta}^2\right) \tag{38}$$

将式(37)、式(38)代入到式(36)中，得

$$\begin{aligned}\mathcal{L}V_2\leqslant&-\left(\lambda_{\min}(\boldsymbol{K}_1)-\frac{3k_1^{4/3}}{4}\right)\|\boldsymbol{e}_1\|^4-\lambda_{\min}(\boldsymbol{K}_2)\|\boldsymbol{e}_2\|^4-\\&\frac{\lambda_1}{2}\tilde{\sigma}^2-\frac{\lambda_2}{2}\tilde{\Delta}^2+\frac{1}{4k_2^4}\sigma_{\mathrm{m}}+\frac{3}{4k_3^2}\Delta_{\mathrm{m}}+\\&\frac{\lambda_1}{2}\sigma_{\mathrm{m}}^2+\frac{\lambda_2}{2}\Delta_{\mathrm{m}}^2\leqslant-a_0V_2+b_0\end{aligned} \tag{39}$$

式中：

$$a_0=\min\left\{4\left(\lambda_{\min}(\boldsymbol{K}_1)-\frac{3k_1^{4/3}}{4}\right),4\lambda_{\min}(\boldsymbol{K}_2),\varepsilon_1\lambda_1,\varepsilon_2\lambda_2\right\} \tag{40}$$

$$b_0=\frac{1}{4k_2^4}\sigma_{\mathrm{m}}+\frac{3}{4k_3^2}\Delta_{\mathrm{m}}+\frac{\lambda_1}{2}\sigma_{\mathrm{m}}^2+\frac{\lambda_2}{2}\Delta_{\mathrm{m}}^2 \tag{41}$$

通过选择合适的控制参数和控制矩阵，使得$\lambda_{\min}(\boldsymbol{K}_1)-3k_1^{4/3}/4>0$，$\lambda_{\min}(\boldsymbol{K}_2)>0$，$\varepsilon_1\lambda_1>0$，$\varepsilon_2\lambda_2>0$，则由引理1可得

$$E(V_2)\leqslant V_2(0)\mathrm{e}^{-a_0t}+b_0/a_0 \tag{42}$$

因此，由式(42)及V_2的定义并结合引理1，易知系统的误差信号都是依概率有界的，同时可以得到紧集：

$$\Gamma=\left\{\boldsymbol{e}_i(i=1,2)|E\left(\|\boldsymbol{e}_i\|^4\right)\leqslant4\left(V_2(0)\mathrm{e}^{-a_0t}+\frac{b_0}{a_0}\right)\right\} \tag{43}$$

结合定义2与式(43)，可知闭环系统误差信号$\boldsymbol{e}_1$和$\boldsymbol{e}_2$最终收敛到紧集Γ，即闭环系统的误差信号4阶矩意义下半全局一致最终有界，而且通过调整所设计的控制器参数，跟踪误差信号最终能够收敛到零点的小邻域内。

对于控制参数的选取，在满足稳定性的前提下，增大$\boldsymbol{K}_1$、$\boldsymbol{K}_2$，减小k_1，会导致更快的瞬态响应但会增加控制输入量；k_2和k_3的增加会使稳态下的跟踪误差更小，同时也会增加控制输入量，在实际工程中选取时应根据闭环系统的性能需求和控制输入限制来不断调整以获得最佳的组合。

4 仿真分析

为了验证本文模型建立的正确性及所提控制方法的有效性和鲁棒性，在MATLAB/Simulink中结合S函数进行数值仿真研究，所用飞艇仿真参数部分结合项目需求背景设计得到，部分参数根据文献[23-24]方法进行估算，如表1所示。

表1 飞艇仿真模型参数

Table 1 Parameters of the airship simulation model

参数	数值	参数	数值
m/kg	1 200	z_{G}/m	2
k_{m1}	0.079	I_x/(kg·m^2)	9 024
k_{m2}	0.869	I_y、I_z/(kg·m^2)	16 900
k_{m3}	0.631	I_{xz}/(kg·m^2)	2 198
x_{G}/m	0	V_a/m^3	1 800

4.1 开环零输入响应

本节对飞艇开环零输入模型仿真，和文献中已有飞艇运动特性结论对比，验证建模的正确性，并且研究气流随机干扰对飞艇运动的影响。

仿真初始状态设置为：高度$h_0=-z=2\,800$ m，速度$u_0=5$ m/s，其余状态量初始值均为0，仿真时间设定为30 s，仿真统一为定步长，大小设置为0.01。

图3和图4分别给出飞艇的位姿状态量在受到

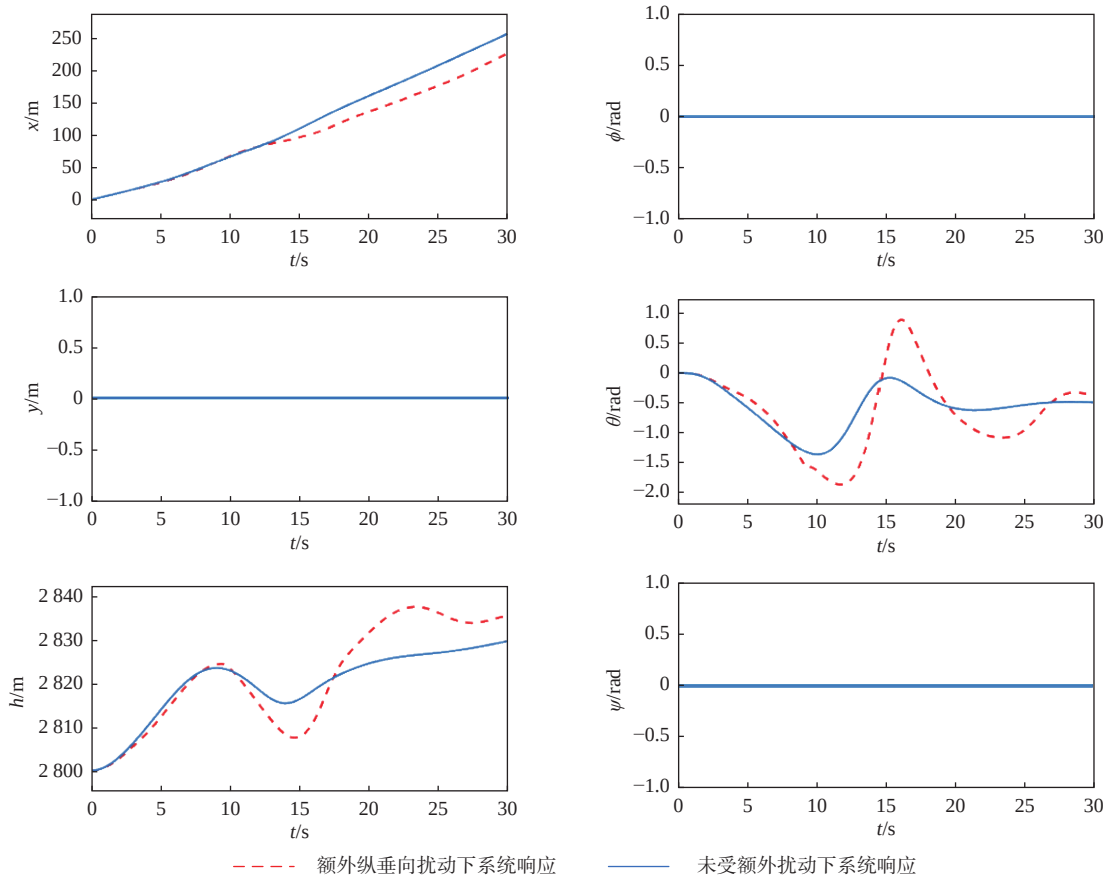


图 3　受纵垂向扰动与未受扰动飞艇位姿响应

Fig. 3　Position and attitude response of airship under longitudinal disturbance and without disturbance

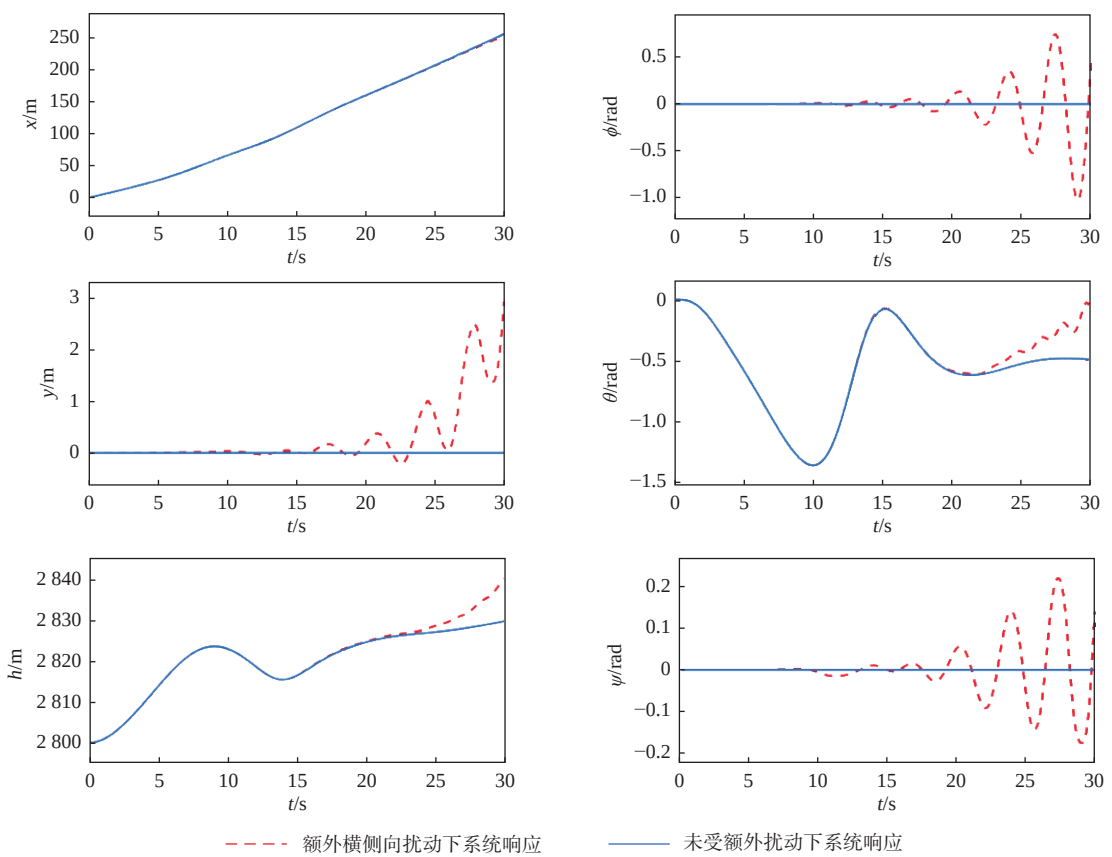


图 4　受横侧向扰动与未受扰动飞艇位姿响应

Fig. 4　Position and attitude response of the airship under directional disturbance and without disturbance

纵垂/横侧向扰动与未受到扰动时的响应曲线。可以看出，飞艇的纵垂向扰动对横向运动参数和横向扰动对纵垂向运动的参数影响都很弱，该结论符合文献 [25] 中建模分析得出的结论，说明仿真模型的正确性。同时，从仿真结果中可以看出，横侧向扰动较大时，飞艇的横向和纵垂向运动甚至会互相影响产生耦合，需要在飞艇全自由度模型基础上为其设计抗干扰控制器。

4.2 闭环控制器响应

本节对飞艇模型加入前文设计的控制律和自适应律，并且与在确定模型下的控制方法对比，来验证本文所提控制方法的效性和鲁棒性。

初始飞行速度和飞行高度分别设置为 u_0=2 m/s, h_0=−z=2 800 m, 其余状态初始值均为 0, 气流干扰确定项均值设置为 $\boldsymbol{\sigma}=[0\quad 0\quad 0\quad 0\quad 0\quad 0\quad 2\quad 1\quad 0.5\quad 0.1\quad 0.1\quad 0]^{\mathrm{T}}$, 具体含义及单位见式 (11), 随机项强度设置为 $\boldsymbol{\Delta}(t)=[10\quad 5\quad 1\quad 0.1\quad 0.01\quad 0.01\quad 0\quad 0\quad 0\quad 0\quad 0\quad 0]^{\mathrm{T}}$, 悬停位置为 $\boldsymbol{\eta}_{1\mathrm{d}}=\begin{bmatrix}x & y & -h\end{bmatrix}^{\mathrm{T}}=[10\quad 5\quad -2\,820]^{\mathrm{T}}$ m, 期望姿态角为 $\boldsymbol{\eta}_{2\mathrm{d}}=[0.2\quad 0.3\quad 0.1]^{\mathrm{T}}$rad。此外，仿真过程中的相关控制器参数经过调节设置为：$\boldsymbol{K}_1=\boldsymbol{I}_{6\times 6}$, $\boldsymbol{K}_2=0.5\boldsymbol{I}_{6\times 6}$, $k_1=1$, $k_2=k_3=0.001$, 自适应参数取：$\varepsilon_1=\varepsilon_2=0.01$, $\lambda_1=\lambda_2=0.1$, 自适应律初值都设置为 0, 仿真时间设置为 30 s, 仿真统一为定步长，大小设置为 0.01。为了更加贴合实际情况，对输入进行限幅，输入力和力矩的大小饱和值都设置为 20 kN 和 20 kN·m。

4.2.1 控制器有效性验证

飞艇位置坐标 x、y、z 和姿态角 ϕ、θ、ψ 跟踪响应曲线图分别如图 5 和图 6 所示。仿真结果表明，

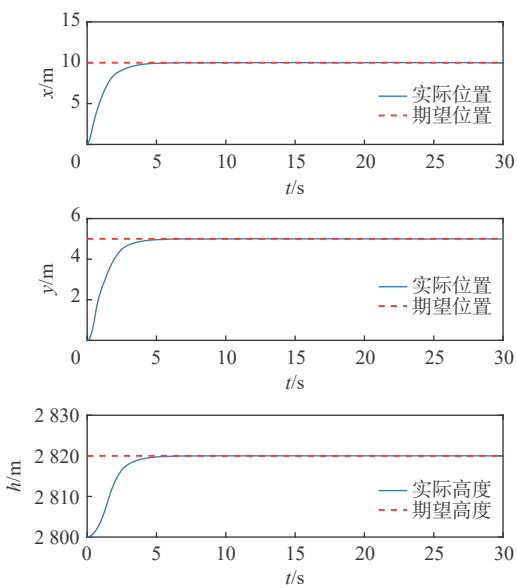


图 5 飞艇位置跟踪响应

Fig. 5 Tracking response for position of airship

在本文设计控制器作用下，由于飞艇本身初始速度 u_0=2 m/s, 飞艇输出信号能在 10 s 内快速跟踪上期望目标信号，实现定点悬停和调整姿态角的控制目标。仿真结果中，姿态角实际曲线会在开始出现反向调节的情况，是因为对控制输入进行了限幅所导致的。图 7 和图 8 分别给出了控制输入力和控制输入力矩的响应曲线。

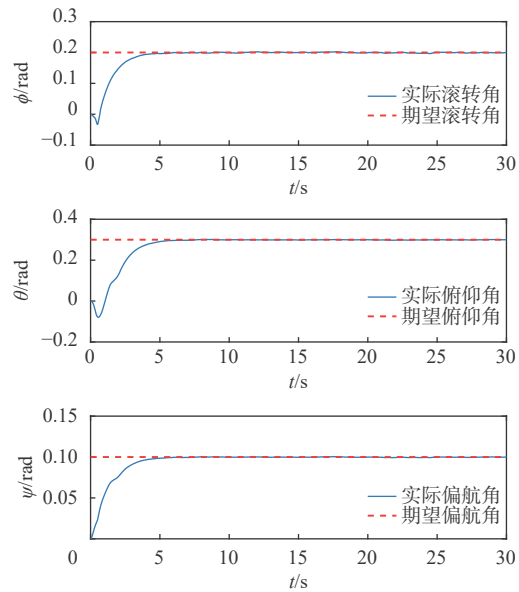


图 6 飞艇姿态角跟踪响应

Fig. 6 Tracking response for the attitude angles of the airship

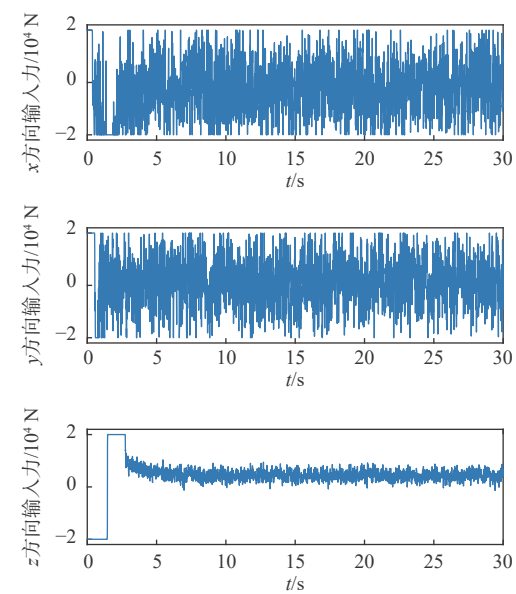


图 7 控制输入力响应

Fig. 7 Response of control input force

4.2.2 控制器鲁棒性验证

本文所设计控制器是在飞艇随机系统下，充分考虑飞艇所受到的随机干扰，引入了和干扰相关的鲁棒项，为了体现设计的控制器在干扰存在尤其是随机干扰下的鲁棒性，给出了其与在确定性系统未

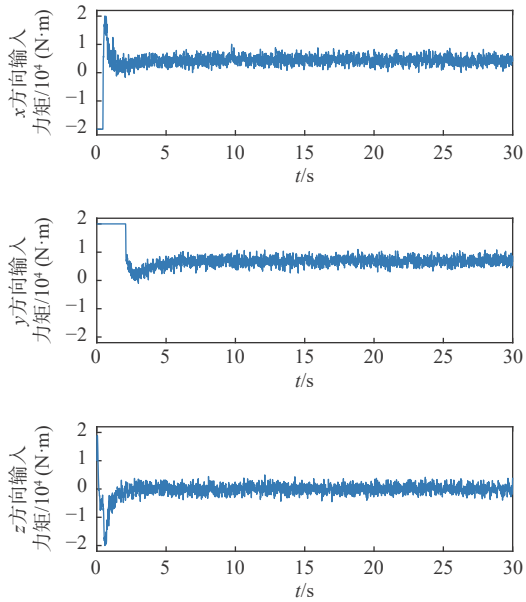


图 8　控制输入力矩响应

Fig. 8　Response of control input torque

考虑随机干扰下用反步法得出控制器的对比, 其他条件(如环境状态、初始状态、控制器参数等)均与前文仿真相同, 仿真结果如图 9 和图 10 所示。

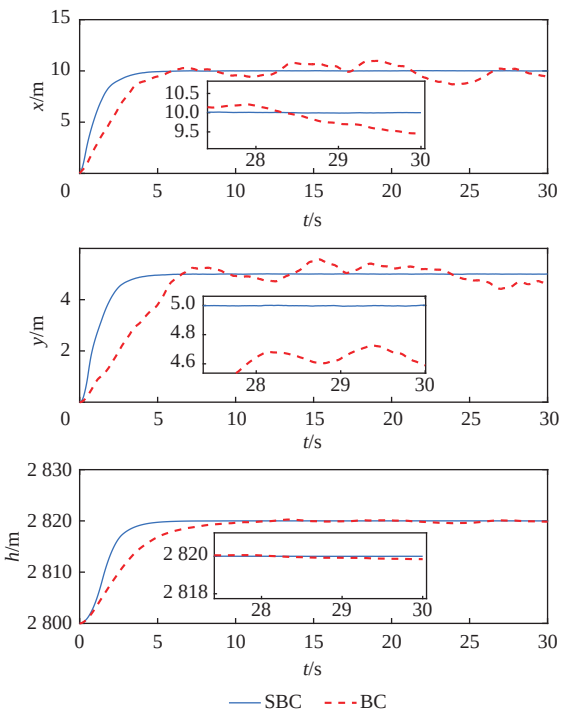


图 9　随机和确定反步法下飞艇位置跟踪响应对比

Fig. 9　Comparison of tracking response for the airship's position under stochastic and deterministic backstepping methods

图 9 和图 10 中蓝色实线表示在随机系统下得出的控制器, 即本文所述随机反步控制器(stochastic backstepping controller, SBC), 红色虚线表示在确定系统得出的基础反步控制器(backstepping controller, BC)。由于一般情况下, 风干扰给飞艇带来的影响

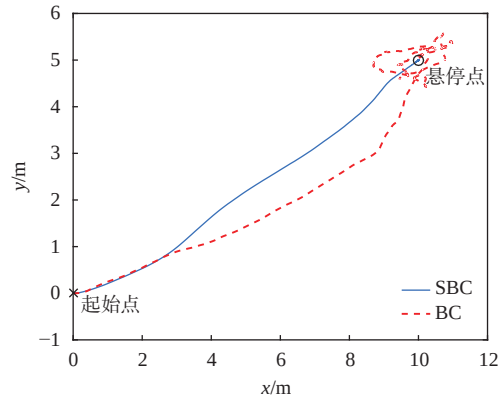


图 10　随机和确定反步法下飞艇水平位移响应对比

Fig. 10　Comparison of horizontal displacement responses of airship under stochastic and deterministic backstepping methods

较少体现在力矩上, 即姿态角受到的干扰较低, 仿真中给出的幅值也较低, 只给出了位置信息的对比。可以明显看出, 从确定系统得出的控制器由于驱动力较小, 到达跟踪状态的调节时间相较于随机反步控制器普遍偏长, 同时在持续的干扰下, 到达稳态时的控制精度也较低。表 2 中给出了图 9 中 x 方向位移在一些时刻的具体值, 从表中分析数据易知实线位置在 7 s 后就达到期望值, 虚线位置需要在 9 s 左右, 从稳态之后 29 s 和 30 s 的数据也能看出随机反步控制器的精度更高, 鲁棒性更强。

表 2　图 9 中 x 方向在一些时刻的具体值

Table 2　Specific values of x-direction at certain moments in Fig.9

时间/s	x/m	
	SBC	BC
5	9.93	9.46
6	9.97	10.05
7	10.00	10.28
8	9.98	10.14
9	9.99	9.84
10	10.00	9.57
11	10.01	9.41
29	9.99	9.72
30	10.00	9.45

4.2.3　不同环境干扰下仿真

为了验证在不同环境如强风下控制器的效果, 增大飞艇受到的干扰, 将随机项强度设置为 $\boldsymbol{\Delta}(t) = [20\ \ 8\ \ 5\ \ 0.1\ \ 0.01\ \ 0.01\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0]^{\mathrm{T}}$, 得到图 11 和图 12 以下结果。可见, 当干扰增大时, 随机系统下设计的控制器仍然能够保持精准的控制效果, 而在确定系统得出的基础反步控制器相比于干扰较小情况下控制效果更加恶劣, 难以精确

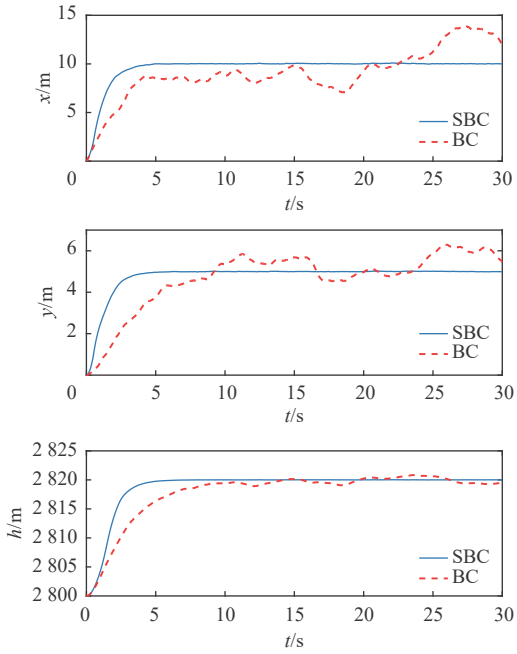


图 11 加大干扰时随机和确定反步法下飞艇位置跟踪响应对比

Fig. 11 Comparison of tracking response of airship's position under stochastic and deterministic backstepping methods when increasing disturbances

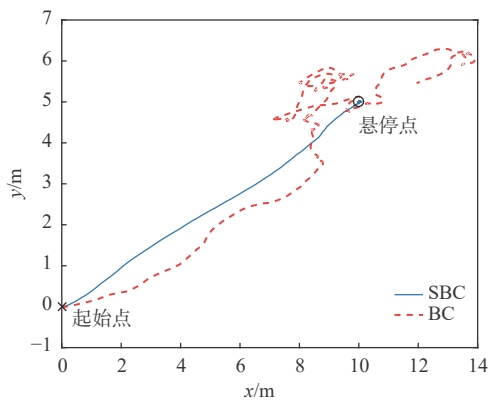


图 12 加大干扰时随机和确定反步法下飞艇水平位移响应对比

Fig. 12 Comparison of horizontal displacement responses of airship under stochastic and deterministic backstepping methods when increasing disturbances

锚定在期望位置。

5 结 论

1）本文提出的抗随机干扰鲁棒控制器可以使飞艇有效抵制随机干扰的影响，并且输出信号能快速跟踪期望目标信号，实现定点悬停和调整姿态角的控制目标。

2）本文对于随机系统的分析和控制器设计方法同样可以用于飞艇受到其他形式的随机干扰，或者其他物理对象。

参考文献（References）

[1] 赵达, 刘东旭, 孙康文, 等. 平流层飞艇研制现状、技术难点及发展趋势[J]. 航空学报, 2016, 37(1): 45-56.
ZHAO D, LIU D X, SUN K W, et al. Research status, technical difficulties and development trend of stratospheric airship[J]. Acta Aeronautica et Astronautica Sinica, 2016, 37(1): 45-56(in Chinese).

[2] 曾盛富, 梁浩全, 王岩, 等. 基于改进的组合赋权和云模型的平流层飞艇健康状态评估[J]. 北京航空航天大学学报, 2026, 52(1): 326-337.
ZENG S F, LIANG H Q, WANG Y, et al. Health status assessment of stratospheric airship based on improved combination weighting and cloud model[J]. Journal of Beijing University of Aeronautics and Astronautics, 2026, 52(1): 326-337(in Chinese).

[3] YANG Y N, WU J, ZHENG W. Station-keeping control for a stratospheric airship platform *via* fuzzy adaptive backstepping approach[J]. Advances in Space Research, 2013, 51(7): 1157-1167.

[4] VIEIRA H S, DE PAIVA E C, MORIGUCHI S K, et al. Unified backstepping sliding mode framework for airship control design[J]. IEEE Transactions on Aerospace and Electronic Systems, 2020, 56(4): 3246-3258.

[5] SAIKI H, FUKAO T, URAKUBO T, et al. Hovering control of outdoor blimp robots based on path following[C]//IEEE International Conference on Control Applications. Piscataway: IEEE Press, 2010: 2124-2129.

[6] ZHENG Z W, ZHU M, SHI D L, et al. Hovering control for a stratospheric airship in unknown wind[C]//AIAA Guidance, Navigation, and Control Conference. Reston: AIAA, 2014: 0973.

[7] WASIM M, ALI A. Estimation of airship aerodynamic forces and torques using extended Kalman filter[J]. IEEE Access, 2020, 8: 70204-70215.

[8] WASIM M, ALI A, CHOUDHRY M A, et al. Unscented Kalman filter for airship model uncertainties and wind disturbance estimation[J]. PLoS One, 2021, 16(11): e0257849.

[9] LE R C, WANG X L, DUAN D P, et al. Attitude control strategy of airship based on active disturbance rejection controller[J]. Aerospace Systems, 2021, 4(1): 7-18.

[10] WANG W, SONG Y L. Analytical computation method for steady-state stochastic response of a time-delay nonlinear automotive suspension system[J]. Mechanical Systems and Signal Processing, 2019, 131: 434-445.

[11] DO K D. Control of fully actuated ocean vehicles under stochastic environmental loads in three dimensional space[J]. Ocean Engineering, 2015, 99: 34-43.

[12] DO K D. Global robust adaptive path-tracking control of underactuated ships under stochastic disturbances[J]. Ocean Engineering, 2016, 111: 267-278.

[13] LI R, WU Q X, CHEN M. Robust adaptive control for unmanned helicopter with stochastic disturbance[J]. Procedia Computer Science, 2017, 105: 209-214.

[14] LI Y K, CHEN M, GE S S, et al. Anti-disturbance control for attitude and altitude systems of the helicopter under random disturbances[J]. Aerospace Science and Technology, 2020, 96: 105561.

[15] YU D X, MA S Z, LIU Y J, et al. Finite-time adaptive fuzzy back-

stepping control for quadrotor UAV with stochastic disturbance[J]. IEEE Transactions on Automation Science and Engineering, 2024, 21(2): 1335-1345.

[16] YAN X H, CHEN M, WU Q X, et al. Adaptive neural tracking control for near-space vehicles with stochastic disturbances[J]. International Journal of Advanced Robotic Systems, 2017, 14(3): 172988141770377.

[17] AZINHEIRA J R, DE PAIVA E C, BUENO S S. Influence of wind speed on airship dynamics[J]. Journal of Guidance, Control, and Dynamics, 2002, 25(6): 1116-1124.

[18] NÆSS A, MOAN T. Stochastic dynamics of marine structures[M]. New York: Cambridge University Press, 2013: 25-27.

[19] KHASMINSKII R. Stochastic stability of differential equations [M]. Berlin: Springer, 2012.

[20] LI B M, XIA J W, ZHANG H S, et al. Event-triggered adaptive fuzzy tracking control for stochastic nonlinear systems[J]. Journal of the Franklin Institute, 2020, 357(14): 9505-9522.

[21] WANG H Q, LIU K F, LIU X P, et al. Neural-based adaptive output-feedback control for a class of non-strict-feedback stochastic nonlinear systems[J]. IEEE Transactions on Cybernetics, 2014, 45(9): 1977-1987.

[22] WU L B, YANG G H. Adaptive output neural network control for a class of stochastic nonlinear systems with dead-zone nonlinearities[J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 28(3): 726-739.

[23] 吴雷, 李勇, 李智斌. 一类平流层飞艇质量和惯量特性的计算方法与分析[J]. 空间控制技术与应用, 2010, 36(6): 39-42.
WU L, LI Y, LI Z B. Calculation and analysis for mass and inertia characteristics of a class of stratospheric airship[J]. Aerospace Control and Application, 2010, 36(6): 39-42(in Chinese).

[24] 任一鹏, 田中伟, 吴子牛. 飞艇空气动力学及其相关问题[J]. 航空学报, 2010, 31(3): 431-443.
REN Y P, TIAN Z W, WU Z N. Some aerodynamics problems of airship[J]. Acta Aeronautica et Astronautica Sinica, 2010, 31(3): 431-443(in Chinese).

[25] 欧阳晋. 空中无人飞艇的建模与控制方法研究[D]. 上海: 上海交通大学, 2003: 18-20.
OUYANG J. Research on modeling and control of an unmanned airship[D]. Shanghai: Shanghai Jiao Tong University, 2003: 18-20 (in Chinese).

Anti-stochastic disturbance control of airship

CHANG Jiaming[1], LI Sulan[1,*], DUAN Xuechao[1], ZHANG Wei[2], WANG Chenyang[1]

(1. School of Mechano-Electronic Engineering, Xidian University, Xi'an 710071, China;
2. Wuhan Second Ship Design and Research Institute, Wuhan 430064, China)

Abstract: Due to their long endurance and low energy consumption, airships are frequently utilized in both high-altitude and low-altitude missions. Studying the airship's anti-stochastic disturbance control is essential since it operates in a complex environment and is frequently impacted by random interference like airflow. Based on the theory of stochastic systems, this paper considers the influence of stochastic disturbance term in the dynamic modeling of stochastic systems. Then, based on the backstepping method, adaptive technique and Lyapunov stability analysis theory, a robust controller against stochastic disturbance is designed. The simulation results demonstrate that the developed controller can quickly steer the airship to the trial site and hover with the desired attitude in the presence of stochastic airflow disturbance.

Keywords: airship; dynamic modeling; stochastic disturbance; stochastic control; backstepping method

Received: 2024-06-27; **Accepted**: 2024-08-15; **Published Online**: 2024-08-26 13:06

URL: link.cnki.net/urlid/11.2625.V.20240824.1221.002

Foundation items: Natural Science Foundation of Shaanxi Province of China (2024JC-YBMS-451,2024JC-YBQN-0398)

* **Corresponding author**. E-mail: slli@xidian.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0222

一种高效 Swin Transformer 加速器设计

蔡青竹，刘强*

(天津大学 微电子学院，天津 300072)

摘　　要：针对 Swin Transformer 模型在资源受限环境下的部署与执行挑战，采用自适应轮询式剪枝和搜索型无偏移量化技术，以降低模型复杂度和存储需求，同时保持模型准确度。通过自适应轮询式剪枝技术动态去除不重要的权重，减少模型的存储与计算需求。引入搜索型无偏移量化方法，优化权重与激活值的存储，进一步降低模型大小，同时尽量避免精度损失。提出了一种专为 Swin Transformer 模型优化的加速架构，通过硬件级别的优化提高数据处理速度和效率。实验结果显示：在应用剪枝和量化技术后，模型大小被压缩至原始的 14.4%，且在 ImageNet-1K 数据集上的 Top-1 准确率达到 77.4%。

关 键 词：现场可编程逻辑门阵列；神经网络；模型剪枝；模型量化；硬件加速器
中图分类号：TN431.2; V247.1+9
文献标志码：A　　**文章编号**：1001-5965(2026)06-2102-12

随着深度学习技术的快速发展，人工神经网络已在各种复杂的数据处理任务中发挥着核心作用。Transformer 模型借其独特的注意力机制，在处理大规模数据集时表现出色。Swin Transformer 作为 Transformer 架构的一个衍生体，在图像生成[1-2]、视频动作识别[3-4]、视觉自监督学习[5] 等多个视觉任务中展现了广泛的应用潜力。

Swin Transformer 通过引入窗口化的多头自注意力机制，将传统 Transformer 全局注意力的平方级计算复杂度降低至线性计算复杂度，解决了高分辨率图像下计算量爆炸的问题，有效减少了模型的推理计算成本[6]。此外，Swin Transformer 通过分层处理图像块，逐步构建层级特征映射，从而更好地适应于不同分辨率的视觉任务。

然而，将 Swin Transformer 部署在资源受限的嵌入式设备仍面临着计算和存储资源的挑战，针对这一挑战，本文首先深入分析了 Swin Transformer 的结构特点，探讨了在嵌入式平台实现这一复杂模型方面的挑战和机遇，然后提出了一系列针对 Swin Transformer 模型的优化策略，并最终在现场可编程逻辑门阵列（field programmable gate arrays, FPGA）上部署以评估加速器的性能。主要创新点包括：

1）提出了一种自适应轮询式剪枝与量化融合方法。自适应的轮询式剪枝方法引入了基于性能反馈的动态剪枝率调整策略。通过梯度敏感度分析确定剪枝顺序的优先级，并采用列级非结构化剪枝有效解决了剪枝后模型计算负载不均衡问题。同时，通过 8 位搜索型无偏移量化，显著提升了模型的压缩比与计算效率，为模型加速提供了有效途径。此外，针对剪枝后的参数特征提出了优化的稀疏矩阵存储方式。

2）优化了 Swin Transformer 模型的计算流程，设计了专用加速器架构。设计了一种矩阵乘法单元，对处理大规模稀疏矩阵和窗口化切块矩阵的乘法操作采取了差异化的优化策略。为了高效计算稀疏矩阵乘法，针对等密度压缩稀疏列存储格式开

收稿日期：2024-04-16；**录用日期**：2024-05-17；**网络出版时间**：2024-06-04 16:20
网络出版地址：link.cnki.net/urlid/11.2625.V.20240604.1119.003
基金项目：国家自然科学基金 (U21B2031)
* **通信作者**. E-mail：qiangliu@tju.edu.cn

引用格式：*蔡青竹，刘强. 一种高效 Swin Transformer 加速器设计 [J]. 北京航空航天大学学报，2026，52（6）：2102-2113.*
CAI Q Z，LIU Q. A high-efficient Swin Transformer accelerator design [J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（6）：2102-2113（in Chinese）.

发了专门的稀疏矩阵乘法计算模式。同时, 为了适应更广泛的计算需求, 设计中还包含了通用矩阵乘法模式, 以处理不适合使用稀疏存储格式的稠密矩阵乘法运算。此外, 提出了硬件友好的非线性函数高斯误差线性单元(Gaussian error linear unit, GELU)的高效实现方法。

3) 基于 ImageNet-1K 数据集的实验结果显示, 剪枝和量化将模型压缩到原始大小的 14.4%, Top-1 准确率维持在 77.416%。在 Xilinx Virtex UltraScale+ VCU1525 平台上, 本加速器的帧率为 183.9 帧/s, 与 Intel E5-2 695 CPU 相比提升了 36.42 倍; 能效比为 4.485 帧/(s·W), 达到了 Intel E5-2 695 CPU 的 8.879 倍, 是 Tesla T4 GPU 实现的 6.02 倍。

1　Swin Transformer 的工作原理及架构

在图像处理领域, 传统的 Transformer 架构采用全局自注意力机制来计算每个像素与其他所有像素之间的关系, 其计算复杂度为 $O(N^2)$, 其中, N 为图像中的像素数量。这种全局性的计算方法随着像素数量的增加而显著增长其复杂度, 从而限制了其在处理大规模或高分辨率图像的应用场景。

为解决这一问题, Swin Transformer 引入了移动窗口注意力机制。通过将图像分割成多个非重叠的窗口, 并在每个窗口内独立执行自注意力计算, 将计算复杂度降低至 $O(N)$。为进一步增强模型性能, Swin Transformer 在其连续的处理块中交替变换窗口的位置, 从而增强了模型在捕捉不同窗口间联系的能力。尽管这种策略提升了模型的整体表现, 但它也引入了窗口间元素混合的问题。

为应对窗口间的元素混合问题, 模型采用了掩码机制, 在自注意力计算中引入掩码矩阵 $\boldsymbol{M}$ 控制元素间的相互可见性:

$$\mathrm{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\mathrm{Softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d}}+\boldsymbol{M}\right)\boldsymbol{V} \tag{1}$$

式中: $\boldsymbol{Q}$ 为查询矩阵, 由输入特征通过线性变换得到, 用于衡量元素间的相关性; $\boldsymbol{K}$ 为键矩阵, 与查询矩阵共同计算注意力权重; $\boldsymbol{V}$ 为值矩阵, 用于存储输入特征的实际信息; d 为每个注意力头的特征维度; $\boldsymbol{M}$ 为窗口掩码矩阵, 对于非同一窗口的元素对, 将其对应位置的掩码值设为 $-\infty$, 从而在 Softmax 计算后将其注意力权重置为 0, 实现窗口间的信息隔离。

Swin Transformer 的架构由多个阶段组成, 如图 1 所示。其中, 4 个主要处理阶段内包含数个 Swin Transformer 块。图 2 展示了 2 个连续的 Swin

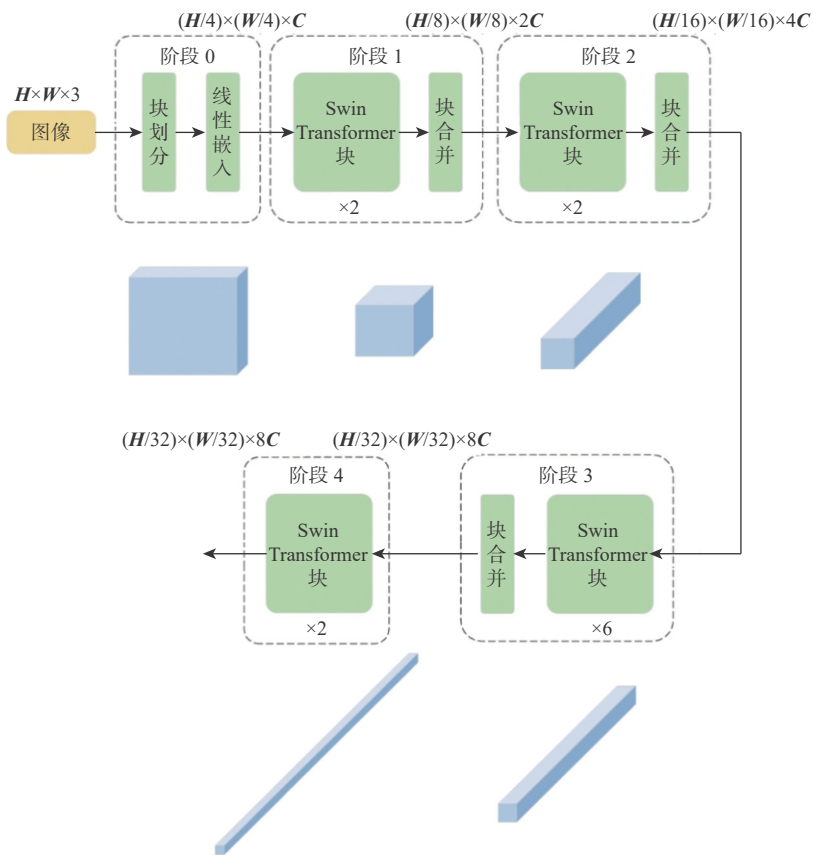


图 1　Swin Transformer 模型整体架构[6]

Fig. 1　Overall architecture of the Swin Transformer model[6]

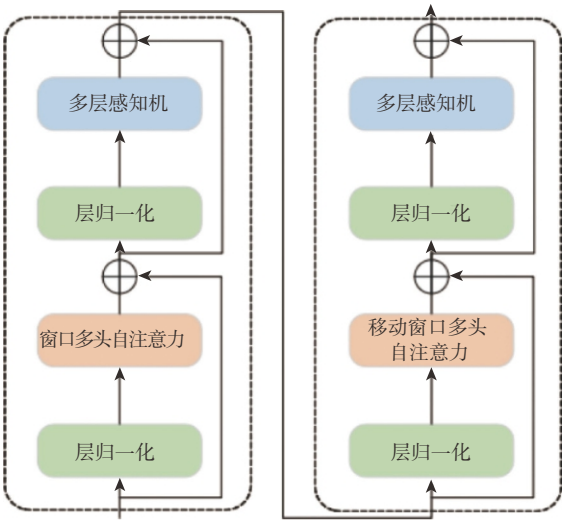


图2 2个连续的Swin Transformer块的结构[6]

Fig. 2 Structure of two consecutive Swin Transformer blocks[6]

Transformer块的结构。这些块通过层归一化、多头自注意力(multi-head self attention, MHSA)机制和多层感知机(multi-layer perceptron, MLP)的联合作用,在局部和全局层面上捕获信息,通过跨阶段的连接来增强信息流。阶段1～阶段3中,通过下采样操作逐步合并窗口,构建了层次化的特征表示结构。

尽管移动窗口自注意力机制提升了Swin Transformer处理高分辨率图像的效率和性能,其在基于FPGA的嵌入式平台上的实现却面临着一系列挑战。这些挑战主要包括因数据动态重塑而导致的处理复杂性增加,对高度并行处理与有限资源的矛盾,以及在硬件上实现非线性函数与稀疏矩阵运算的困难。

2 自适应轮询式剪枝方法

2.1 整体流程

轮询式剪枝是一种基于轮询机制的模型压缩方法,通过周期性评估并调整模型参数的剪枝率,实现了模型压缩与性能保持的平衡。轮询式剪枝过程中,剪枝的先后顺序会对剪枝效果产生影响,早期剪枝的参数组往往实现更高剪枝率,而晚期剪枝的效果则受限,因此带来了不公平性[7]。此外,反复剪枝可能引致模型性能的连续下降,即便每次剪枝造成的准确率损失微小,其累积效应仍可能导致显著性能降低。为避免剪枝过程中的不公平性和性能损失的累积效应,本文提出了一种自适应轮询式剪枝方法(adaptive polling-based pruning, APP)。首先将Swin Transformer模型的参数细分为多个基于阶段及内部结构相似性的参数组。针对每个参数组,设计了一套基于轮询机制的剪枝率动态调整算法,通过周期性评估各参数组的重要性与贡献度,以实现差异化的剪枝决策。同一阶段的不同Swin Transformer块具有相同数量的线性层,即具有内部结构相似性。如图2所示,同一阶段中的多层感知机、第1个层归一化、(移动)窗口多头注意力、第2个层归一化分别有m、n、k、q个线性层,则该阶段的线性层权重参数分为$m+n+k+q$组,其余阶段划分类似,不同阶段的参数分布在各自的参数组中。

除剪枝顺序外,仅用模型准确率变化来衡量每个参数组的剪枝表现也会引发不公平性,尤其是Swin Transformer中参数组间参数数量相差数十倍的情况下。在增加相同剪枝率的情况下,参数量大的参数组减去的参数量更多,因此不应该仅用模型准确率变化来衡量每个参数组的剪枝表现。本文针对选择剪枝目标时存在的潜在不公平性,提出了一种融合敏感度分析与参数量评估的剪枝决策方法。

敏感度分析通过定义敏感度梯度G,量化参数剪枝对模型准确率的影响,引导每轮迭代中最优剪枝参数组的选择。其计算公式为

$$G=\frac{\Delta a}{\Delta p} \tag{2}$$

式中:Δa为模型准确率的变化量;Δp为剪枝率的变化量。

随后,定义权衡系数α以综合考虑敏感度梯度G与参数量N对剪枝顺序的影响:

$$\alpha=\begin{cases} GN & G>0 \\ N & G=0 \\ G/N & G<0 \end{cases} \tag{3}$$

首先,根据G的范围进行排序,$G>0$的参数组排在最前,其次是$G=0$,最后为$G<0$的参数组。再依据α值进行排序,高值优先。这一策略在保持模型性能的同时,优化模型压缩效率。

初始阶段,所有参数组的剪枝率设为0%,保持模型的完整性,并将每轮中所有参数组剪枝率的单次增量设为ΔP。随后,采取轮询策略逐组评估参数,基于模型在验证集的Top-1准确率动态调整剪枝率。每轮迭代中,对每个参数组内非零且绝对值最小的参数部分执行剪枝,并在转至下一组前恢复上一组参数,以确保准确评估。

在迭代过程中,若参数组剪枝后模型准确率未超预设的准确率容忍阈值T,恢复该参数组全部参数并继续轮询下一个参数组;反之则将该组参数剪枝率降低ΔP并停止该组剪枝。直至所有待剪枝集合C中的参数组都经过轮询后,找出本轮最佳的唯一剪枝组进行剪枝,在此基础上将每个参数组的

轮询剪枝率继续依次提升ΔP, 进行下一次迭代。当待剪枝集合为$\varnothing$时, 以ΔP的增量进行的迭代剪枝过程已经完成。此时, 剪枝率增量减半至$\Delta P/2$, 并开始下一轮的迭代, 直至剪枝增量降至预设的最小值ΔP_{min}。

某些参数在当前轮迭代中虽然作用很小, 但是在后续轮迭代中又可能变得重要, 因此, 当剪枝率增量降至预设的ΔP_{reqair}后, 将部分被剪参数恢复至参数组, 再继续后续剪枝, 如图 3 所示。

通过引入自适应轮询式剪枝方法, 避免了传统剪枝方法中常见的性能损耗累积与决策盲目性的问题, 有助于精确剪除对模型性能贡献较少的参数, 同时保留对模型预测精度关键的参数。整体剪枝流程如下所示。

输入: 初始剪枝率增量ΔP_{init}; 最小剪枝增量ΔP_{min}; 模型原始 Top-1 准确率Δa_{init}; 开始修复的剪枝率增量ΔP_{reqair}; 准确率容忍阈值T; 恢复系数$n\%$, 表示恢复参数占被剪参数的百分比。

输出: 所有参数组的最终剪枝率。

1: $\Delta P \leftarrow \Delta P_{\text{init}}$
2: 将所有参数组的剪枝率设置为 0%
3: while $\Delta P > \Delta P_{\text{init}}$ do
4:　$C \leftarrow$所有参数组
5:　将 C 恢复至未剪枝状态
6:　if $\Delta P = \Delta P_{\text{reqair}}$ then
7:　　将所有参数组的剪枝率减少$n\%$
8:　end if
9:　while C != $\varnothing$ do
10:　for C_i in C do
11:　　将 C_i 的剪枝率增加ΔP
12:　　对 C_i 剪枝
13:　　$a_{\text{cur}} \leftarrow$准确率评估
14:　　$\Delta a \leftarrow a_{\text{cur}} - a_{\text{init}}$
15:　　$G \leftarrow \Delta a / \Delta P$
16:　　$N \leftarrow$获取 C_i 的参数量
17:　　恢复 C 至未剪枝状态
18:　　将 C_i 的剪枝率减少ΔP
19:　　if $a_{\text{cur}} < a_{\text{init}} - T$ then
20:　　仅对 C_i 剪枝
21:　　从 C 中移除 C
22:　　continue
23:　　end if
24:　　if $G > 0$ then
25:　　$\alpha \leftarrow G \times N$
26:　　Category $\leftarrow 1$
27:　　else if $G < 0$ then
28:　　$\alpha \leftarrow G/N$
29:　　Category $\leftarrow -1$
30:　　else
31:　　$\alpha \leftarrow N$
32:　　Category $\leftarrow 0$
33:　　end if
34:　　$C_i \cdot \alpha \leftarrow \alpha$
35:　　$C_i \cdot$ Category $\leftarrow$Category
36:　　end for
37:　　对C首先按 Category 降序排序, 然后在同一 Category 内按α降序排序
38:　　从排序后的列表中选择第 1 个 C_0
39:　　将 C_0 剪枝率增加ΔP
40:　end while
41:　$\Delta P \leftarrow \Delta P/2$
42: end while
43: 输出所有参数组的最终剪枝率

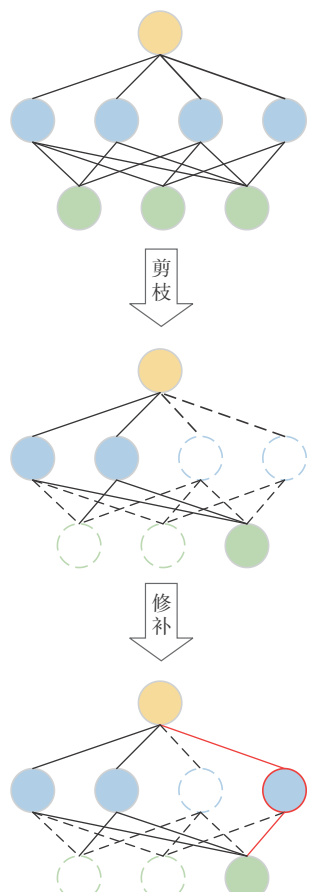


图 3　剪枝和修补过程

Fig. 3　Pruning and patching process

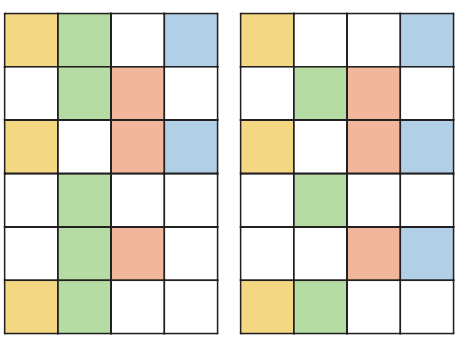
2.2　列级非结构化剪枝技术

自适应轮询式剪枝方法有效地平衡了模型的性能与参数量, 但在实际应用中仍存在如何在保持模型性能的同时进一步提升运算效率与硬件兼容性的挑战。模型剪枝的 2 种主要形式是结构化剪

枝和非结构化剪枝[8-9]。非结构化剪枝通过元素级的剔除进行稀疏化，虽然在减少模型参数方面效果显著，但由于引入的稀疏性破坏了数据的连续性，对硬件并行化处理的效率构成挑战。图4(a)展示了在并行计算架构中，传统非结构化剪枝由于其非零元素的随机分布，导致不同处理单元(processing element, PE)间存在明显的负载不平衡。此不平衡会导致部分PE过早进入空闲状态，而其他负载较重的PE仍在计算，影响整体计算效率。相对而言，结构化剪枝以网络的结构单元(如神经元、通道或层)作为剪枝单位，保持了数据的规整性和连续性，促进了硬件加速器的计算效率，但牺牲了模型压缩率[10-12]。

针对上述问题，本文提出了列级非结构化剪枝方法，此方法宏观上维持列的完整性，确保每列的剪枝率一致，而微观上允许列内部元素的非结构化剪除。这种折衷方案利用结构化剪枝的硬件友好特性，同时又不完全放弃非结构化剪枝在模型压缩率方面的优势。在剪枝过程中，每列作为一个整体，列元素按绝对值大小排序，根据设定的剪枝率逐列剪除最小元素。

列级非结构化剪枝通过确保每列非零元素数量大致相等，实现PE间负载均衡，优化并行计算过程，提高计算效率，如图4(b)所示。



(a) 传统的非结构化剪枝 (b) 列级别的非结构化剪枝

图4 50%稀疏度下2种剪枝方法的工作负载均衡效果比较

Fig. 4 Comparison of workload balancing effects of two pruning methods at 50% sparsity

2.3 稀疏矩阵的存储格式

在处理剪枝后的稀疏矩阵时，传统的稠密矩阵存储方式因大量零值元素存在而导致存储与计算效率低下。因此，适用的稀疏矩阵存储格式显得尤为重要。

压缩稀疏列(compressed sparse column, CSC)存储格式[13]通过列序存储非零元素及其行索引和列指针，可以跳过零元素连续访问非零数据，有效降低矩阵乘法的内存访问随机性，图5展示了这一过程。在CSC格式中，非零元素的值按列优先顺序存储，与之配套的行索引数组指示每个元素在原始矩阵中的行坐标。此外，列指针数组索引每列的起始非零元素位置，且其最后一个元素标示整个数组的长度，从而界定每列非零元素的范围。

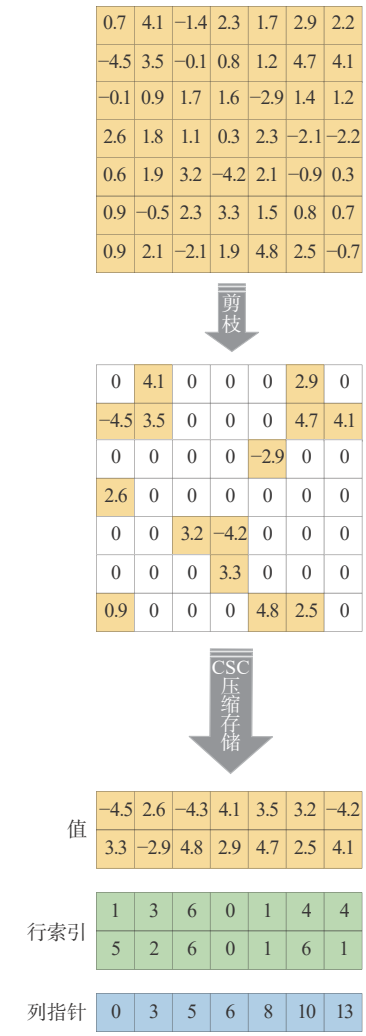


图5 CSC格式压缩存储示意图

Fig. 5 Schematic diagram of CSC format compressed storage

CSC表示法有效压缩矩阵数据，降低存储需求，并简化索引与计算。本文在此基础上进行改进，提出了与列级非结构化剪枝更适配的等密度压缩稀疏列(uniform compressed sparse column, UCSC)存储格式，如图6所示，其中，红色虚线框出的列指针表示与UCSC格式相比减少的部分。

由于经列级非结构化剪枝后每列的非零元素数量是一致的，UCSC格式可以省略CSC格式中的列指针向量，仅通过行索引向量和非零值数组即可实现稀疏矩阵的有效表示。这种格式继承了CSC的优势，而且进一步减小了存储尺寸。以图5所示矩阵为例，使用UCSC格式与CSC格式相比节省了20%的存储空间。

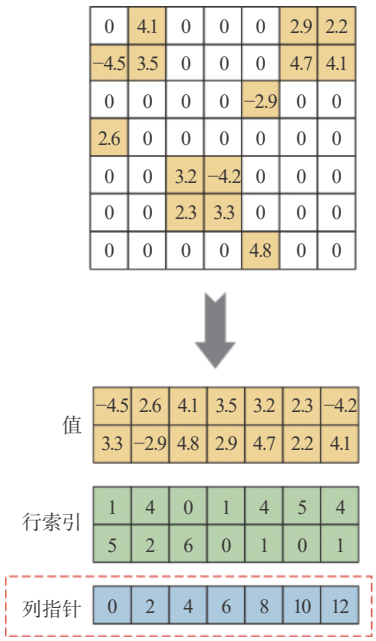


图 6　UCSC 格式压缩存储示意图

Fig. 6　Schematic diagram of UCSC format compressed storage

3　搜索型无偏移量化

自适应轮询式剪枝方法和列级非结构化剪枝技术有效地降低了模型的存储需求与计算资源消耗, 但剪枝后的模型参数仍以浮点数形式存在, 面临着效率与资源消耗的挑战。需要研究量化技术, 以进一步降低硬件资源消耗。

图 7 展示了剪枝后的 Swin Transformer 中具有代表性特征的 2 种参数分布, 可见 Swin Transformer 参数的对称分布特性, 本文采用无偏移量化[14], 其公式为

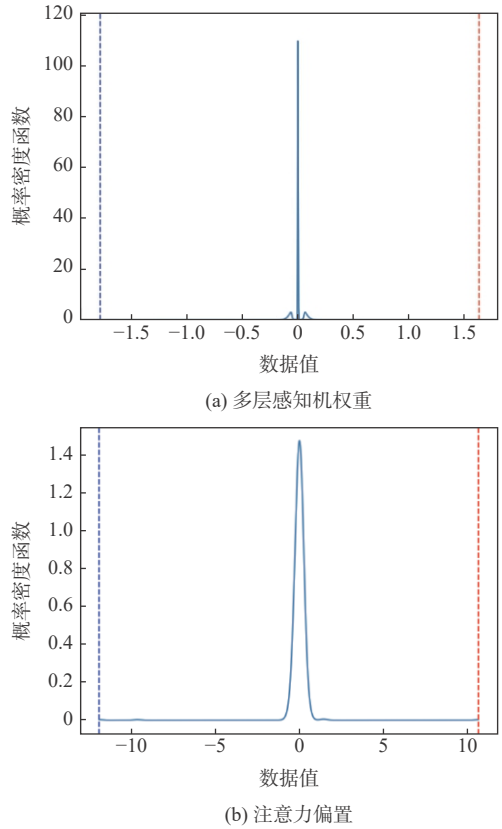
$$q = \text{clamp}\left(\text{round}\left(x/s\right), -2^{n-1}+1, 2^{n-1}-1\right) \tag{4}$$

$$\text{clamp}(x,a,b) = \begin{cases} a & x \leqslant a \\ x & a < x \leqslant b \\ b & x > b \end{cases} \tag{5}$$

$$s = \frac{c}{2^{n-1}-1} \tag{6}$$

式中: round 函数将数值调整到最近的整数; x为原始的浮点数; s为缩放因子; q 为量化后的 8 位整数输出; clamp$(\cdot,a,b)$ 为裁剪函数, 将输入值限制在区间 $[a,b]$ 内; a、b 分别为量化的下限与上限。由于本文进行的量化为 8 位量化, 因此, n为 8, 缩放因子 $s = c/127$。c为裁剪值, 决定了数据绝对值的最大界限。为最大程度保留原有分布, 通常直接使用数据的最大绝对值作为裁剪值[15]。

从图 7 可以观察到, 模型中多数的参数类型呈现长尾分布特性, 大部分数据点集中在低值区域, 仅少数极端值占据了高值区域。如果直接采用数据的最大绝对值作为裁剪值, 则位于长尾分布低值区域的大部分数据将被压缩至较小的量化范围内, 可能将接近 0 的值量化为 0, 丢失小数值信息, 从而降低准确率。相反, 较小的裁剪值虽然提高了数值区分度, 但会导致更多数据被映射到量化范围之外, 从而丢失大数值信息。



(a) 多层感知机权重

(b) 注意力偏置

图 7　Swin Transformer 量化前参数分布

Fig. 7　Distribution of parameters before quantization of Swin Transformer

为解决上述问题, 本文通过在长尾分布的特定范围内寻找最佳缩放因子来减少量化误差。从式(6)可以得出, 裁剪值与缩放因子 s成正比关系, 因而其变化趋势与 s变化趋势相同。首先确定初始缩放因子 s_0:

$$s_0 = \left\lceil \log_2 \frac{\max(|x|)}{2^{n-1}-1} \right\rceil \tag{7}$$

式中: $\lceil\cdot\rceil$为向上取整操作; max$(|x|)$为所有数据点绝对值的最大值, 避免在初始搜索时丢失极端值。在硬件实现的过程中, 用移位操作代替乘法可以使计算更为高效, 因此需要保证缩放因子为 2 的整数次幂。

从 s_0开始每次取 $s_i = s_{i-1}/2$可以得到模型准确度与缩放因子之间的关系。图 8 展示了多层感知机权重的缩放因子对模型准确度的影响。

通过观察量化后模型推理准确率随每种参数类型的缩放因子变化的趋势, 可以发现存在准确率随缩放因子的减小而单调下降和先上升后下降这

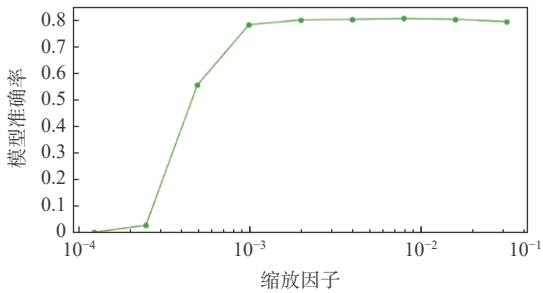


图 8 模型准确度与多层感知机权重的缩放因子的关系

Fig. 8 Relationship between model accuracy and scaling factors of multilayer perceptron weights

2 种趋势。由此可知，在指定的遍历范围内，总会存在一个 s_{opt}，使得量化准确率达到峰值。在遍历过程中，当准确度出现下降趋势时，则下降前一个缩放因子视为最佳缩放因子 s_{opt}。

4 硬件架构介绍

4.1 加速器整体架构设计

图 9 为本文设计的加速器整体架构图。加速器采取了一种基于窗口分割的并行计算框架，采用数据流控制策略优化计算效率，针对计算过程中各窗口共享权重和偏置参数的特性，将全模型参数预先加载并缓存至片上，从而降低对外部存储的频繁访问需求。切块控制单元负责处理输入特征图数据，通过预设的大小和步长将数据切分成多个窗口数据区域，将切块后的数据分别存储在多个块随机存取存储器（block random access memory, BRAM）中，用于后续的并行处理。

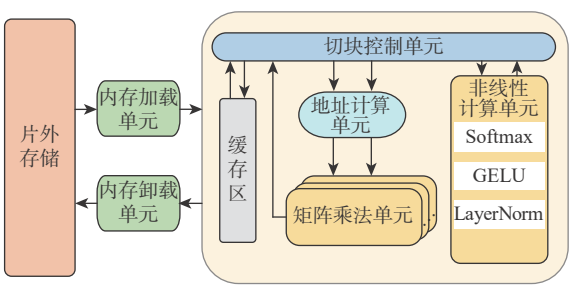


图 9 加速器整体架构图

Fig. 9 Overall architecture of the accelerator

加速器计算数据流起始于 LayerNorm 计算单元，随后数据流经过稀疏矩阵乘法单元完成全连接层运算，生成查询 $\boldsymbol{Q}$、键 $\boldsymbol{K}$ 及值 $\boldsymbol{V}$ 矩阵。切块控制单元分别将这些矩阵缓存于临时缓存区，以实现计算过程的高效数据访问。特别地，$\boldsymbol{Q}$ 矩阵在缓存前经过缩放处理，以保证注意力机制中数值的稳定性，而 $\boldsymbol{K}$ 矩阵则利用地址计算单元规避了转置带来的额外中间存储过程，直接参与后续计算。

在计算注意力分数的阶段，加速器使用了矩阵乘法单元的通用矩阵乘法模式来处理 $\boldsymbol{Q}$ 矩阵和 $\boldsymbol{K}$ 矩阵，以计算出序列中各元素间的相关度。随后，通过 Softmax 计算单元对注意力权重进行归一化处理。归一化的注意力权重结合 $\boldsymbol{V}$ 矩阵输入到矩阵乘法单元生成加权的自注意力表示。

在多头注意力机制的输出合并阶段，全连接层调用矩阵乘法单元的稀疏矩阵乘法模式，以将多头注意力的结果重新映射回原始特征空间的维度。前馈网络（feed-forward network, FFN）阶段则包含非线性计算单元及稀疏矩阵乘法，以实现多层感知机的运算，最终通过残差连接和切块控制单元将单窗口的计算结果有效回写至完整特征图数据缓存区。

为确定加速器内最合适的窗口并行计算模块数量 P，引入窗口平均任务量 W_{AW} 作为决策指标。W_{AW} 衡量的是全部计算任务所需的总计算时间与执行单个窗口计算所需基准时间的比值，有效反映了特定并行度下计算效率与资源利用的关系，其计算式为

$$W_{\text{AW}}(P)=\sum_i\left\lceil\frac{W_i}{P}\right\rceil \tag{8}$$

式中：P 为窗口并行计算模块的数量；W_i 为图 1 所示第 i 个 Swin Transformer 块中 7×7 的窗口数量。较低的 W_{AW} 意味着较高的并行度，整体计算时间减少。图 10 展示了 W_{AW} 随 P 的变化趋势。当 $P>16$ 时，W_{AW} 的下降趋势变得平缓，接近稳定，此时增加计算模块数量所带来的效率提升边际递减，引入较高的资源空闲率，导致未能充分利用计算资源。据此，综合考虑计算效率与资源利用，所有以窗口为单位的计算使用 16 个计算模块并行操作，从而实现并行性与资源空闲之间的平衡。

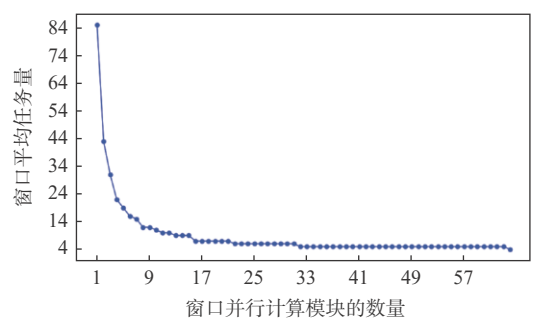


图 10 窗口并行计算资源优化策略的性能分析

Fig. 10 Performance analysis of window parallel computation resource optimization strategy

4.2 地址计算单元

在 Swin Transformer 加速器的实现中，求 $\boldsymbol{Q}$ 矩阵、$\boldsymbol{K}$ 矩阵及 $\boldsymbol{V}$ 矩阵时，包含多次数组转置、重塑操作。传统方法利用临时缓存区来暂存重塑数据，会带来额外的内存开销和读写成本。

地址计算单元采用了一种处理多维数据结构的重塑操作策略, 摒弃了传统的依赖于临时缓存区来暂存重塑数据的做法, 而是通过设计的映射函数直接得到重塑后的新索引, 从而实现了数据访问的高效化。对于目标数据中的每个索引组合, 都可以计算出其在注意力机制处理后新矩阵中的对应位置。每个索引在转换后的注意力矩阵 $\boldsymbol{Q}$、$\boldsymbol{K}$ 和 $\boldsymbol{V}$ 中的位置可以通过以下的多维映射公式计算得出:

$$i_{\text{ndex}} = b(n_{\text{heads}}N3H_{\text{DIM}}) + n(n_{\text{heads}}3H_{\text{DIM}}) + s(n_{\text{heads}}H_{\text{DIM}}) + tH_{\text{DIM}} + c \tag{9}$$

式中: i_{ndex} 表示目标索引位置, 即在转换后的注意力映射矩阵中的定位; b 表示批次索引, 用以区分不同数据批次; n_{heads} 为多头注意力机制中的头数, 表示并行处理的注意力头数量; N 为序列长度, 表示处理的数据序列的维度; 3 表示处理过程中的 3 个关键组件, 即 $\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{V}$ 矩阵; H_{DIM} 为每个注意力头的维度; n 为序列中特定元素的索引; s 为选择器, 指定读取的 3 种不同的状态: 0(读取 $\boldsymbol{Q}$ 中的元素), 1(读取 $\boldsymbol{K}$ 中的元素), 和 2(读取 $\boldsymbol{V}$ 中的元素); t 为序列中元素的索引; c 表示通道的索引。这种计算方法能够高效地遍历 $\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{V}$ 中的每个元素, 从切块控制单元中读取并通过地址计算单元直接计算出所需数据在存储器中的正确顺序位置, 从而避免了进行数据重排所带来的额外成本。

4.3　矩阵乘法单元

矩阵乘法单元有稀疏矩阵乘法和通用矩阵乘法计算 2 种模式。稀疏矩阵乘法计算模式中输入为一个用 UCSC 格式存储的稀疏矩阵和一个稠密矩阵, 而通用乘法计算模式中的输入为 2 个稠密矩阵。

稀疏矩阵乘法模式中, 权重矩阵 $\boldsymbol{W}$ 以等密度压缩稀疏列存储格式表示。计算过程中遍历 $\boldsymbol{W}$ 的每列, 使用行索引确定每列的非零元素在值数组中的位置。对于每列中的非零元素 W_{ij}, 提取其行索引 i 并从值数组中提取对应的值。随后对输入矩阵 $\boldsymbol{X}$ 的第 k 行的每个数据 X_{ki}, 计算 $X_{ki} \times W_{ij}$ 并将结果累加到结果矩阵 $\boldsymbol{Y}$ 的相应元素 Y_{kj} 中。在这一过程中, 引入缓冲区来暂存中间计算结果, 以允许并行处理多个 $\boldsymbol{X}$ 的行, 从而进一步优化性能。稀疏矩阵乘法模式的主要优势在于与 UCSC 存储结合, 仅对权重矩阵中的非零元素及其索引进行操作, 显著减少了计算量。

通用矩阵乘法模式的计算是基于窗口大小 7×7分块的矩阵进行的, 需要7×7路并行的乘累加(multiply-accumulate, MAC)操作。使用 $\boldsymbol{Q}$ 和 $\boldsymbol{K}$ 计算注意力权重及使用注意力矩阵和 $\boldsymbol{V}$ 计算注意力分布均使用此模式, 通过图 11 所示乘法单元计算引擎来实现。计算注意力权重时, 通道数为每个头(head)处理的特征向量的维度 32, 所需并行度为 32; 在计算注意力分布时, 通道数增至窗口中元素个数 49, 因而所需并行度为 49。面对 32 和 49 这 2 种不同维度的并行计算要求, 每个乘法计算单元(PE)由 32+16 个乘法器及 32+16 个加法器构成, 其内部通过寄存器实现了独立的列缓冲区, 保障了对被乘矩阵列的高效并行访问。图中以虚线框标注的组件展示了为处理不同输入数量与维度的必要扩展。

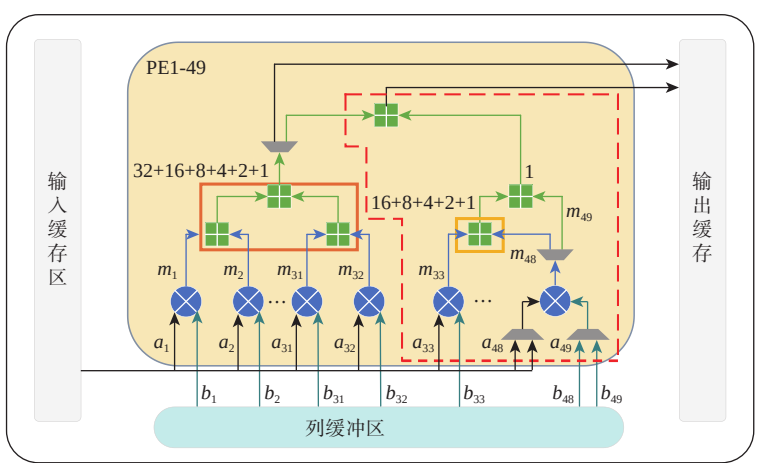


图 11　乘法单元计算引擎

Fig. 11　Multiplication unit computing engine

PE 采用层次化的设计。为了支持并行度为 32 的并行累加计算, 直接将 32 节点的完全平衡二叉树的输出作为 PE 的计算结果写入输出缓存区。为了支持并行度为 49 的并行累加计算, 鉴于 49 个数据点并非 2 的整数幂, 无法直接构建一个标准的完全平衡二叉树, 故本设计采纳了一种层次分割策略, 创建了一个 32 节点的完全平衡二叉树和一个 16 节点的完全平衡二叉树, 外加一个单独节点。在

最低层级，列间计算并行执行，通过48个乘法单元同时处理乘数矩阵的第 i 行元素 a_n 与被乘矩阵的第 j 列元素 b_n 之间的乘积 $a_n \times b_n (n \in [1,48])$，随后将结果 $m_n=a_n \times b_n$ 作为2个完全平衡二叉树的输入。由于分层计算，$a_{49} \times b_{49}$ 的计算与 $a_{48} \times b_{48}$ 的计算复用了同一个乘法器，乘积在次高层级进行累积。在计算过程中，16节点的二叉树比32节点二叉树领先一次加法计算延迟，因此被选中与单独节点相加，以确保该分支与32节点加法树在同一时间计算完毕，分别作为最高层级的2个输入，有效减少执行的总延时。

4.4 GELU计算单元

GELU函数被用在Swin Transformer模型的MHSA模块和FFN模块之间，以及FFN内部，利用其非线性特性，增强模型的表达能力[6]。通常情况下，GELU的计算式为

$$\mathrm{GELU}(x) \approx 0.5x(1+\tanh(\sqrt{2/\pi}(x+0.447\,15x^3))) \tag{10}$$

可以看出，GELU的实现依赖于复杂的双曲正切(tanh)函数计算，不利于硬件实现。为了处理此问题，本文采用了一种基于多项式逼近tanh函数的方法。由于$\tanh(x)$在x远离原点时趋近于±1，采用分段格式将函数分为3个部分进行近似：当$|x|>2$时，$\tanh(x)$近似为±1；在中心区域$-2 \leqslant x \leqslant 2$，通过最小二乘法拟合多项式。

本文采用的渐进式非线性最优化策略融合了高斯-牛顿法[16-17]的高效收敛能力与梯度下降法的广域搜索优势。在参数远离最优解时，采用梯度下降法避免局部最小值的问题，提供一个较为合适的参数估计；而在逼近最优解时转向高斯-牛顿法，利用其局部快速收敛的优势来加速求解过程。通过引入阻尼系数，调节算法在高斯-牛顿方法和梯度下降策略之间的权衡，实现了2种方法间的平滑过渡。为了进一步简化，将接近0的系数视为0，得到最终的简化多项式。

最终近似函数为

$$\tanh(x) \approx \begin{cases} -1 & x < -2 \\ 0.023\,538\,11x^5 - 0.210\,711\,95x^3 + 0.956\,837\,39x & -2 \leqslant x \leqslant 2 \\ 1 & x > 2 \end{cases} \tag{11}$$

图12展示了实际的tanh函数与使用分段5阶多项式的近似版本之间的比较，图13展示了GELU函数使用tanh近似前后的对比。

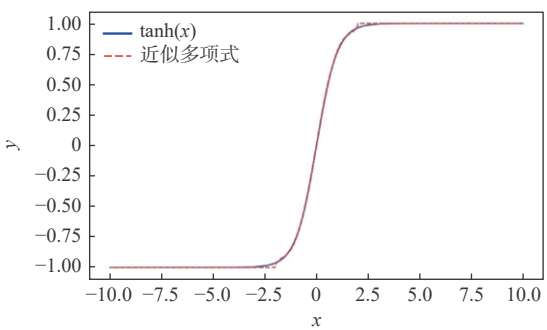


图12 tanh函数与近似五阶多项式的对比

Fig. 12 Comparison of the tanh function with an approximate fifth-order polynomial

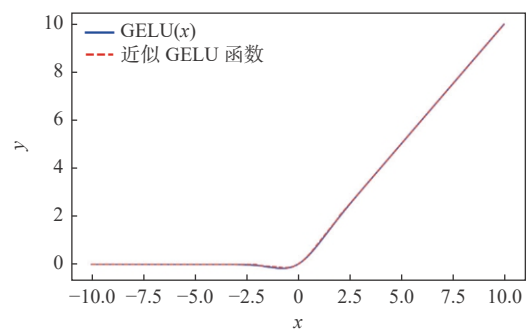


图13 tanh函数替换前后GELU函数的对比

Fig. 13 Comparison of the GELU function before and after replacement with the tanh function

由图12和图13可知，替换前后的数据显示出极高的相似性，几乎没有可观察到的差别。考虑到深度学习模型中普遍存在的参数冗余和容错特性，这种精度上的轻微牺牲通常是可接受的。频率同为100 MHz时，分段函数(120 ns)与迭代15次的CORDIC算法实现的tanh函数相比(380 ns)，单次计算的延迟仅为其31.6%。

5 实验结果与分析

5.1 仿真及实验环境

本文量化和剪枝实验所用数据集为ImageNet-1K，软件算法基于PyTorch 1.12.1框架实施，并利用CUDA 11.6环境进行计算。特别的，本文发现Swin Transformer模型中约97.0%的参数量(共27 426 816个参数)主要集中在全连接层，适合进行剪枝。然而，剩余的3.0%参数虽然比例较小，但对模型的准确率有显著的负面影响，且这部分参数与矩阵乘法操作无关。因此，余下的3%的参数未经

剪枝处理, 而是仅通过搜索型无偏移量化方法进行了优化。

加速器部署在 Xilinx UltraScale+ FPGA VCU1525 开发板上, 并运用 Vivado 2021.1 软件平台进行设计及性能评估。

5.2 自适应轮询式剪枝与搜索型无偏移量化方法评估

5.2.1 不同首轮剪枝下的实验结果

为了找到模型压缩和性能之间的最佳平衡点, 通过压缩率与 Top-1 准确率的损失的比值来评估剪枝效果。基于本文第 2.1 节提出的轮询式动态剪枝方法, 对 Swin Transformer 模型中的权重数据分类后进行动态剪枝, 尝试使用不同首次轮询剪枝率对模型中的待剪枝参数进行剪枝实验, 容忍阈值为 2% 时实验结果如图 14 所示。首轮剪枝率决定了每轮剪枝迭代的剪枝率变化步长, 随着首次轮询剪枝率的减小, 压缩率与性能损失的比值呈现先上升后下降的趋势。这是由于当首次轮询剪枝率过大时, 轮询过程中优先被选中的当前轮最佳剪枝参数组可能并非是结合其他参数组剪枝后应该选择的最佳剪枝率, 缺乏对不同参数重要性的考量, 从而误剪对模型预测性能贡献较大的关键参数; 当首次轮询剪枝率过小时, 整体剪枝步长过小, 容易陷入局部最优解。综合实验结果分析, 首次轮询剪枝率设置为 32% 时可以在最大程度上平衡模型性能损失和压缩率。

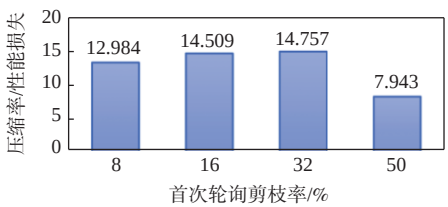


图 14　剪枝效果随首次轮询剪枝率的变化

Fig. 14　Comparison of pruning effectiveness with changes in the initial polling pruning rate

5.2.2 不同修补剪枝率起始点的实验结果

为了评估修补对剪枝效果的影响, 使用不同的修补剪枝率起始点进行测试, 图 15 展示了容忍阈值为 2%, 首次轮询剪枝率为 32% 时不同修补剪枝率起始点的对比。

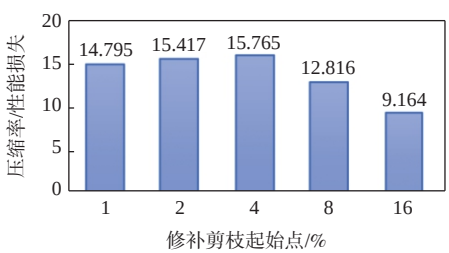


图 15　剪枝效果随修补剪枝率起始点的变化

Fig. 15　Comparison of pruning effectiveness with the starting point of repair pruning rate

与图 14 所示的压缩率/性能损失进行对比, 可以看出, 增加了修补过程后该比值普遍上升, 说明剪枝效果更优。同时, 压缩率/性能损失的比值随修补起始点的调整呈先增后减的趋势, 修补起始点设为轮询剪枝率每轮增加 4% 时, 比值最大, 剪枝效果最好, 因此, 后续实验建立在此配置的基础上。

5.2.3 不同容忍阈值下的实验结果

通过设定多个容忍阈值, 定量评估剪枝对模型压缩率和性能损失的影响, 如表 1 所示。当阈值设定为 5% 时, 模型显示出 54.920% 的压缩率, 此时超过一半的参数被剪除, 然而性能损失达 5.672%, 暴露了追求过高压缩比的潜在风险。相对地, 当阈值调整至 1% 时, 压缩率下降至 28.877%, 而性能损失减少至 1.502%。因此, 为了保持模型性能稳定性, 需要调整容忍阈值。将容忍阈值设定为 2% 时, 模型压缩率与性能损失之间达到较优的平衡状态, 压缩率为 43.802%, 性能损失减少至 2.572%。

表 1　不同容忍阈值对模型剪枝效果的影响

Table 1　Impact of different tolerance thresholds on model pruning effectiveness

容忍阈值/%	压缩率/%	准确率损失/%
5	54.920	5.672
2	43.802	2.572
1	28.877	1.502

5.2.4 不同模型量化方法准确率对比的实验结果

表 2 中的数据为先使用 2.1 节提出的自适应轮询式剪枝, 再进行量化后的效果。原始模型作为基准, 有偏移量化和最大值截断的无偏移量化的量化损失分别为 6.12% 和 2.66%, 而采用搜索型截断无偏移量化则将量化损失降低至前者的 15.1% 和 34.7%。当最大值截断的无偏移量化与剪枝相联合时, 模型准确率下降显著, 而此时搜索型无偏移量化在不牺牲压缩率的前提下, 保持了量化后的模型准确率, 在压缩效率与模型性能之间取得了更佳的平衡。

表 2　不同量化方法推理准确率对比

Table 2　Comparison of inference accuracy for different quantization methods

剪枝后量化方法	模型大小/MB	准确率/%	量化损失/%
未量化	62.4	78.34	
有偏移量化	15.6	72.22	6.120
最大值截断的无偏移量化	15.6	75.68	2.660
搜索型无偏移量化	15.6	77.42	0.924

5.3 加速器硬件评估

表 3 展示了加速器的硬件资源消耗、功耗及单张图片的推理性能评估结果，该结果由 Vivado 2021.1 完成综合、布局及布线后生成的官方报告统计得出；其中，功耗数据通过 Vivado 内置的功耗分析工具实测获取。

表 3 加速器资源消耗、功耗及性能评估

Table 3 Accelerator resource consumption, power consumption and performance evaluation

延迟/ms	BRAM/块	DSP/个	LUT/个	FF/个	URAM/块	功耗/W
5.437	1 402	3 267	508 787	682 198	33	41

表 3 中涉及的 FPGA 核心硬件资源包括：BRAM、数字信号处理单元（digital signal processing, DSP）、查找表（look-up table, LUT）、触发器（flip-flop, FF）、超高速随机存取存储器（ultra random access memory, URAM）。

从表 4 中可以看出，与 CPU 相比，本文在 Swin Transformer 模型上实现了 36.42 倍的加速，同时功耗仅为 CPU 的 41%，能效为 CPU 的 88.79 倍。与 GPU 相比，本文加速器的吞吐量为 Tesla T4 的 3.292 倍，能效比达到其 6.02 倍。E5-2695 和 Tesla T4 的数据从参考文献 [18-19] 中得到。

表 4 本文加速器与 CPU 和 GPU 实现的比较

Table 4 Comparison of accelerator in this paper with CPU and GPU implementations

型号	频率/GHz	延迟/ms	功耗/W	帧率/(帧·s^{-1})	能效/(帧·(s·W)$^{-1}$)
E5-2695[18]	3.12	198	100	5.051	0.050 51
Tesla T4[19]	1.005	17.90	75	55.86	0.745
本文加速器	0.2	5.437	41	183.9	4.485

图 16 对比了本文中提出的加速器与现有的 2 种 Swin Transformer 加速器 Vis-TOP[18] 和 HPTA[19] 在性能上的差异。Vis-TOP 并未对非线性函数进行特殊优化，而 HPTA 仅对 Swin Transformer 模型进行了量化处理，没有结合其他压缩技术。在处理时间（延迟）方面，本文加速器比现有的 2 种加速器分别提升了 2.18 倍和 1.40 倍。这归因于软硬件结合的加速策略，对模型进行剪枝和量化处理，再进行了定制化的适用于压缩模型的硬件设计，并且针对 Swin Transformer 中硬件不友好的非线性计算进行了改进。

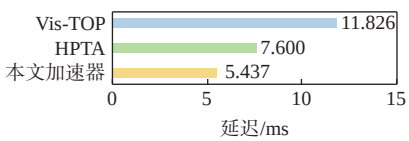


图 16 与已有相关 FPGA 加速器的推理延迟对比

Fig. 16 Comparison of inference latency with existing FPGA accelerators

6 结 论

1）本文提出的自适应轮询式剪枝与搜索型无偏移量化融合压缩方案，在 ImageNet-1K 数据集上实现模型尺寸较原始 Swin Transformer 模型压缩了 14.4% 的优化效果，同时 Top-1 准确率仍保持 77.42%，仅带来 0.924% 的量化损失。

2）针对数组转置和重塑操作中存在的资源占用及延迟问题，本文设计了地址计算单元，减少了中间数组资源占用及赋值延迟；同时，为优化稀疏矩阵与窗口切块矩阵的乘法处理，分别设计了稀疏矩阵乘法单元和通用乘法计算单元；此外，通过采用渐进式非线性最优化策略，结合高斯-牛顿法和梯度下降法的优势，并融合加法树技术，有效解决了非线性函数在硬件上的实现挑战。

3）实验结果表明，在 Xilinx Virtex UltraScale+ VCU1525 平台上，本文加速器的处理性能为 183.9 帧/s，与 Intel E5-2 695 CPU 相比提升了 36.42 倍；能效比为 4.485 帧/(s·W)，达到了 Intel E5-2 695 CPU 的 8.879 倍，是 Tesla T4 GPU 实现的 6.02 倍。

参考文献（References）

[1] ZHANG B W, GU S Y, ZHANG B, et al. StyleSwin: transformer-based GAN for high-resolution image generation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2022: 11294-11304.

[2] NGUYEN L X, TUN Y L, TUN Y K, et al. Swin transformer-based dynamic semantic communication for multi-user with different computing capacity[J]. IEEE Transactions on Vehicular Technology, 2024, 73(6): 8957-8972.

[3] WU H N, CHEN C F, LIAO L, et al. DisCoVQA: temporal distortion-content transformers for video quality assessment[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2023, 33(9): 4840-4854.

[4] LIU Z, NING J, CAO Y, et al. Video swin transformer[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2022: 3192-3201.

[5] ZHAO H, ZHANG C, ZHU B L, et al. S3T: self-supervised pre-training with swin transformer for music classification[C]//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE Press, 2022: 606-610.

[6] LIU Z, LIN Y T, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2022: 9992-10002.

[7] LIN X F, KIM S, JOO J. FairGRAPE: fairness-aware GRAdient pruning mEthod forFace attribute classification[C]//Proceedings of the Computer Vision-ECCV 2022. Berlin: Springer, 2022: 414-432.

[8] BLALOCK D, ORTIZ J J G, FRANKLE J, et al. What is the state of neural network pruning? [EB/OL]. (2020-03-06)[2024-02-16]. https://arxiv.org/abs/2003.03033.

[9] PRAGNESH T, MOHAN B R. Compression of convolution neural network using structured pruning[C]//Proceedings of the IEEE 7th International Conference for Convergence in Technology. Piscataway: IEEE Press, 2022: 1-5.

[10] LIU Z C, MU H Y, ZHANG X Y, et al. MetaPruning: meta learning for automatic neural network channel pruning[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2019: 3295-3304.

[11] CHEN H L, YANG J F, MAO S A. Convolutional layers acceleration by exploring optimal filter structures[C]//Proceedings of the IEEE International Conference on Recent Advances in Systems Science and Engineering. Piscataway: IEEE Press, 2022: 1-6.

[12] CARBALLO M V, KIL LEE B. Accuracy-aware structured filter pruning for deep neural networks[C]//Proceedings of the International Conference on Computational Science and Computational Intelligence. Piscataway: IEEE Press, 2021: 679-682.

[13] JAYAKODY S, WANG J. EMBARK: memory bounded architectural improvement in CSR-CSC sparse matrix multiplication[C]// Proceedings of the IEEE 9th International Conference on Collaboration and Internet Computing. Piscataway: IEEE Press, 2024: 8-17.

[14] WANG S, KANG Y. Gradient distribution-aware INT8 training for neural networks[J]. Neurocomputing, 2023, 541: 126269.

[15] KIM J, LEE C, CHO E, et al. Towards next-level post-training quantization of hyper-scale transformers[EB/OL]. (2024-02-14)[2024-02-16]. https://arxiv.org/abs/2402.08958.

[16] SCHABACK R. Convergence analysis of the general Gauss-Newton algorithm[J]. Numerische Mathematik, 1985, 46(2): 281-309.

[17] ARBEL M, MENEGAUX R, WOLINSKI P. Rethinking Gauss-Newton for learning over-parameterized models[C]//Proceedings of the Advances in Neural Information Processing Systems 36. San Diego: Neural Information Processing Systems Foundation, Inc., 2023: 33379-33402.

[18] HU W, XU D, FAN Z M, et al. Vis-TOP: visual transformer overlay processor[EB/OL]. (2021-10-21)[2024-02-16]. https://arxiv.org/abs/2110.10957.

[19] HAN Y T, LIU Q. HPTA: a high performance transformer accelerator based on FPGA[C]//Proceedings of the 33rd International Conference on Field-Programmable Logic and Applications. Piscataway: IEEE Press, 2023: 27-33.

A high-efficient Swin Transformer accelerator design

CAI Qingzhu，LIU Qiang*

(School of Microelectronics，Tianjin University，Tianjin 300072，China)

Abstract：This study used adaptive iterative pruning and search-based bias-free quantization approaches to minimize model complexity and storage requirements while preserving accuracy in response to the deployment and execution problems of Swin Transformer models in resource-constrained situations. Initially, Adaptive iterative pruning dynamically eliminated non-essential weights, reducing storage and computational demands. Subsequently, Search-based bias-Free quantization optimized the storage of weights and activation values, further minimizing model size with minimal accuracy loss. In terms of hardware design, an accelerator architecture tailored for Swin Transformer models was put out, improving data processing efficiency and speed via hardware-level adjustments. Experimental results indicated that after applying pruning and quantization, the model's size reduced to 14.4% of its original, achieving a Top-1 accuracy of 77.4% on the ImageNet-1K dataset.

Keywords：field programmable gate arrays；neural networks；model pruning；model quantization；hardware accelerator

Received：2024-04-16；**Accepted**：2024-05-17；**Published Online**：2024-06-04 16：20

URL：link.cnki.net/urlid/11.2625.V.20240604.1119.003

Foundation item：National Natural Science Foundation of China (U21B2031)

* **Corresponding author**. E-mail：qiangliu@tju.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0211

基于错误率分配的逻辑电路面积近似优化

范海媚，吴乾火，王伦耀*
(宁波大学 信息科学与工程学院，宁波 315211)

摘　　要：考虑到在相同错误率约束下，同一逻辑电路的二级和多级近似优化效果并不一样，且逻辑电路的二级表示与多级表示之间可以等效转化，提出一种基于错误率分配的逻辑电路近似面积优化算法。该算法包括错误率分配算法，二级错误率快速估算算法及二级/多级近似优化算法。所提算法使用C语言和面向时序逻辑综合与形式验证系统(ABC)工具内置命令编程实现，使用北卡罗莱纳微电子中心(MCNC)基准电路进行测试。实验结果表明，在错误率约束为5%情况下，所提算法在面积和延时优化上分别为42.36%和25.15%；与单纯使用多级近似优化方法相比，面积和延时的优化分别进一步提升了5.05%和9.36%。

关 键 词：近似计算；逻辑综合；近似优化；错误率分配；面积优化
中图分类号：TN431.2
文献标志码：A　　**文章编号**：1001-5965(2026)06-2114-09

随着半导体工艺尺寸进入到纳米级，集成电路在电源电压降低、集成密度不断提高的先进技术节点上对参数的波动和故障更敏感，为确保计算精确性需要在不同设计层次上进行冗余设计，从而产生显著的额外电路面积和能源开销，使得电路成本变得更高[1]；另外，多媒体、信号处理、机器学习、模式识别和数据挖掘等这些由与人类感知/认知相关的应用都表现出固有的容错特性。一定范围内的计算错误并不影响实际的应用。因此，通过降低计算精度来获得电路其他性能提升的近似计算技术已经作为一种纳米尺度电路设计的新范式受到广泛的关注[2-3]。

近似计算技术可以通过引入可接受的误差量实现电路的面积、功耗和延迟等优化。基于近似计算技术的优化也称近似优化。电路的近似优化上可分为手动近似电路设计和近似逻辑综合(approximate logic synthesis, ALS)[4]2个领域。在手动近似电路设计上，主要侧重近似算术单元电路设计，如近似加法器[5]和近似乘法器[6]。近似逻辑综合则通过借助优化算法，以自动化方式来构建任何近似电路，并控制错误引入。其基本思想是通过降低电路的计算精度，实现电路结构的简化，从而获得面积、功耗及延迟等性能的提升[7-8]。现有的近似优化技术根据被优化电路的层级可分为二级近似优化和多级近似优化2种。

逻辑功能的二级电路描述一般采用乘积项之和(sum of product, SOP)形式，在近似优化中，常通过增加或删除部分乘积项来实现逻辑函数的简化。如文献[9]通过在错误率约束下试探性地添加一些最小项来实现乘积项的进一步合并，最终实现逻辑函数包含的字母数的减少，达到电路面积优化目的。文献[10]通过修改逻辑函数真值表，使某个最小项对应的输出由“1”变为“0”或者反过来，然后，通过反复求解可满足性问题和集合覆盖问题实现函数包含的字母数的减少，进而实现错误率约束下的逻辑函数简化。

逻辑功能的多级电路描述一般采用与非图(and-inverter graph, AIG)、二元决策图(binary decision

收稿日期：2024-04-11；**录用日期**：2024-07-16；**网络出版时间**：2024-08-05 09:06
网络出版地址：link.cnki.net/urlid/11.2625.V.20240802.1441.002
基金项目：国家自然科学基金(U23A20351,U22A2013)；宁波市重点研发计划(2023Z233)
* **通信作者**. E-mail：wanglunyao@nbu.edu.cn

引用格式：*范海媚，吴乾火，王伦耀. 基于错误率分配的逻辑电路面积近似优化[J]. 北京航空航天大学学报，2026，52(6)：2114-2122.*
FAN H M，WU Q H，WANG L Y. Area approximate optimization of logic circuits based on error rate allocation[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：2114-2122 (in Chinese).

diagrams, BDD) 等, 多级电路的优化程度通常采用 AIG 或 BDD 中包含的节点数及层级数等指标来衡量。因此, 在多级电路近似优化中, 常采用节点删除或近似替代的方式实现。如文献 [11] 中结合多扇出节点常量替换的错误率控制、在全局范围内提取合适替换的候选节点集和选择每个候选节点输出的常量替换值等方法实现面积优化; 文献 [12] 提出一种基于近似关心集 (care set) 重新替换的 ALS 流程, 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, 进而实现面积优化。

由于电路的二级和多级的表示形式不一样, 并且优化程度的衡量方法也不一样, 因此, 在相同的错误率约束下, 同一个逻辑功能, 其在二级与多级的优化程度可能存在明显的差异, 即错误率优化效能在二级和多级情况下是不一样的。另外, 考虑到逻辑功能的二级和多级描述之间可以等效转换, 因此, 对于某一个逻辑电路的近似优化过程中, 通过对总的错误率约束进行适当的分配, 分别用于该电路的二级和多级近似优化, 并保留二者中优化效果更好的部分进行进一步优化, 要比完全用以往单一的二级或多级优化的效果会更好。本文把这种近似优化技术称为混合层级优化技术。相对于混合层级优化, 以往的单一的二级或多级近似优化仅仅是混合层级近似的一个特例。

本文目标是实现电路的多级近似优化, 其中, 待优化的原始电路采用可编程逻辑阵列 (programmable logic array, PLA) 格式描述的二级电路, 优化后得到的电路为多级电路。不同于以往先将二级电路转化为多级电路, 然后, 进行多级近似优化策略[13-14]。本文则先将一部分的错误率约束用于二级电路近似优化, 然后, 将近似优化后的二级电路转化为多级电路, 再实现多级近似优化。该优化策略本质上就是一种混合层级优化策略, 能较好地兼顾错误率优化效能, 进而提高最终的多级优化效果。

1　错误率及其计算方法

采用近似计算技术的逻辑综合必须满足给定的近似度约束。近似度有多种衡量方式, 如错误率、误差距离和平均误差等[15] 等。一般在讨论算术电路近似优化时通常采用误差距离, 平均误差等作为约束条件, 而对非算术电路, 通常采用错误率作为近似优化的约束条件, 本文选择错误率作为电路的近似度指标。

1.1　二级电路错误率的计算方法

近似逻辑函数的错误率定义为导致近似函数与原始函数不同输出的输入组合占所有输入组合的比例[16]。对于一个 n 输入, m 输出逻辑函数 $f_{\mathrm{org}}(X)$, 其中, $X=\{x_0,x_1,x_2,\cdots,x_{n-1}\}$为输入变量集合, $f_{\mathrm{app}}(X)$为 $f_{\mathrm{org}}(X)$的近似函数, e 为错误率, 则

$$e=\frac{\sum\left(f_{\mathrm{app}}(X)\neq f_{\mathrm{org}}(X)\right)}{2^n} \tag{1}$$

式中: $\sum\left(f_{\mathrm{app}}(X)\neq f_{\mathrm{org}}(X)\right)$为在所有输入组合中, 引起错误输出的输入组合个数。当 $f_{\mathrm{app}}(X)$和 $f_{\mathrm{org}}(X)$表示二级电路时, 一般采用 SOP 形式。

1.2　多级电路错误率的计算方法

式 (1) 中 $f_{\mathrm{app}}(X)\neq f_{\mathrm{org}}(X)$也可以等效地表示为 $f_{\mathrm{app}}(X)\oplus f_{\mathrm{org}}(X)=1$, 因此, 可以将式 (1) 变形为

$$e=\frac{\sum\left(f_{\mathrm{app}}(X)\oplus f_{\mathrm{org}}(X)=1\right)}{2^n} \tag{2}$$

在多级电路错误率计算中, 所有满足式 (2) $f_{\mathrm{app}}(X)\oplus f_{\mathrm{org}}(X)=1$的输入组合可以利用如图 1 所示的 Miter 电路求得。$f_{\mathrm{app}}(X)$对应近似电路, $f_{\mathrm{org}}(X)$对应原始电路。当 Miter 电路输出为 1 时, 表示 f_{app}和 f_{org}的输出不一致。而满足 Miter 电路输出为 1 的输入组合, 可以借助 BDD 来求解。

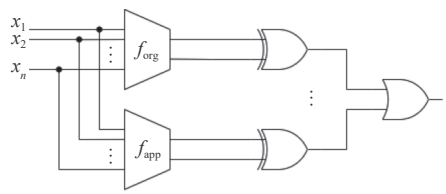


图 1　Miter 电路

Fig. 1　Miter circuit

当 $f_{\mathrm{app}}(X)$和 $f_{\mathrm{org}}(X)$规模比较大时, 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 的路径统计效率, 亦即降低了错误率的计算速度。为提高计算效率, 在大电路近似优化时, 可使用基于蒙特卡罗模拟法来估算错误率[17]。该方法通过随机产生一定数量的测试输入, 比较输出的结果, 然后通过式 (2) 获得对应的错误率。

由式 (2) 可得, 当输入变量数 n 比较大时, 如 n>20, 式 (2) 中分母大小是百万级, 使得计算 $f_{\mathrm{app}}(X)\oplus f_{\mathrm{org}}(X)=1$过程中引入的误差对式 (2) 的计算结果影响有限, 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。

2　错误率约束下的近似优化原理

2.1　二级近似优化原理

二级逻辑电路的面积常用对应的逻辑函数所包含的字母数来衡量[18], 因此, 基于近似计算技术的二级逻辑电路面积优化的主要目标是在给定错

误率约束下尽可能地减少逻辑函数所包含的字母数，而常用的方法是通过对乘积项进行添加/删除操作。

以逻辑函数 $f=\bar{x}_1\bar{x}_3x_4+\bar{x}_1\bar{x}_2x_4+x_1x_2x_3x_4$ 近似优化为例，假设错误率约束为 1/16，图 2 为 f 的卡诺图。为在错误率约束下获得更加简单的逻辑表达式，可以采用如下 2 种方式。方式 1 是添加最小项 $\bar{x}_1x_2x_3x_4$，对应图 2(b) 中标识为“A”的方格。通过添加该最小项的方式可以使逻辑函数简化为 $f_{\text{ap1}}=\bar{x}_1x_4+x_2x_3x_4$。相比于 f 包含的 10 个字母数，f_{ap1} 包含的字母数仅为 5 个。方式 2 是直接删除 f 中的最小项 $x_1x_2x_3x_4$，得到 $f_{\text{ap2}}=\bar{x}_1\bar{x}_3x_4+\bar{x}_1\bar{x}_2x_4$，$f_{\text{ap2}}$ 中包含 6 个字母数。上述 2 种优化方式中，引入的错误率均为 1/16，均引起逻辑函数表达式包含的字母数不同程度减少。若电路的面积用对应的逻辑函数包含的字母数来衡量，则上述 2 种方法在错误率约束下都实现了电路面积的近似优化。

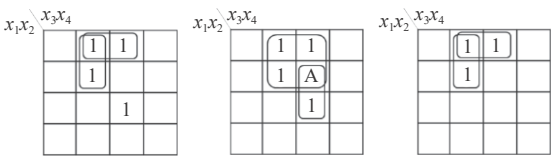


图 2 逻辑函数对应的卡诺图

Fig. 2 Karnaugh map corresponding to a logic function

2.2 多级近似优化原理

在多级逻辑优化中，逻辑电路的功能常用 AIG 或 BDD 等表示，其面积优化经常用 AIG 或 BDD 包含的节点数来衡量[13]。因此，基于多级逻辑电路面积近似优化的主要目标是在给定错误约束下尽可能地减少 AIG 或 BDD 的节点数。

以基于 AIG 的多级逻辑电路近似优化为例，在一定的错误率约束下，通过节点删除或近似替代来实现。以如图 3 卡诺图所示的 1 位全加器电路为例，2 个加数和进位输入为 x_2,x_1,x_0，z_0 和 z_1 分别表示加法器的“求和”与“进位”输出，其逻辑表达式为和 $z_0=\bar{x}_2\bar{x}_1x_0+\bar{x}_2x_1\bar{x}_0+x_2\bar{x}_1\bar{x}_0+x_2x_1x_0$，$z_1=x_2x_1+x_1x_0+x_2x_0$。2 个表达式共包含了 7 个乘积项和 18 个字母数。图 4(a) 为 1 位全加器电路逻辑功能对应的 AIG 表示。

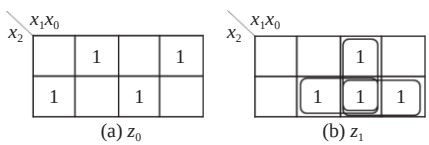


图 3 1 位全加器对应的卡诺图

Fig. 3 Karnaugh map corresponding to a 1-bit full adder

在 AIG 中的每个节点都表示 1 个 2 输入的逻辑“与”运算，虚线表示逻辑取反。考虑到逻辑“与”运算的 1 个输入为逻辑 0 时，不管其他输入发生什么变化，其输出恒为 0。因此，在基于 AIG 的近似优化中，删除那些离 POs 越远的节点，对整个电路错误率影响越小。图 4(b) 是在图 4(a) 上经过节点常量替换的近似操作后得到的。显然图 4(b) 的节点数要少于图 4(a)，但图 4(b) 付出的代价是引入了 25% 的错误率。

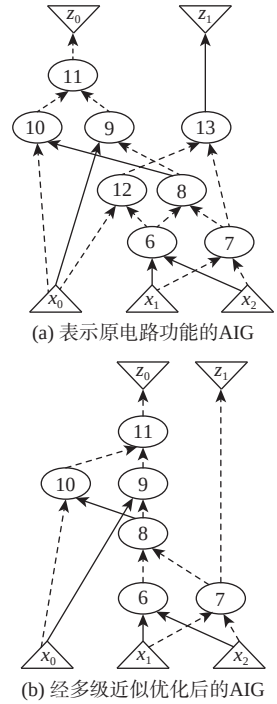


图 4 1 位全加器电路功能的 AIG 表示

Fig. 4 AIG representation of the 1-bit full adder circuit

2.3 错误率约束对近似优化的影响

在逻辑电路面积近似优化中，二级电路面积用逻辑表达式包含的字母数来衡量，多级电路的面积用 AIG 或 BDD 的节点数来衡量。考虑到表达式中字母数与表达式对应的 AIG 或 BDD 的节点数并不存在必然的正相关，即字母数的减少并不表示对应的 AIG 或 BDD 节点数的减少。因此，如果将单位错误率 Δe 先用于 f 二级电路近似优化，得到 f_{ap2L}，然后将 f_{ap2L} 转化为等效的多级电路，得到 $f_{\text{ap2L_M}}$，与直接将 Δe 用于 f 多级电路近似优化，得到 $f_{\text{ap_M}}$，最终 $f_{\text{ap2L_M}}$ 与 $f_{\text{ap_M}}$ 的复杂度可能不一样。

仍以图 4(a) 的 1 位全加器电路为例，假设错误率约束为 25%，可以先对原始电路进行二级近似预处理操作，删除乘积项 $x_2\bar{x}_1\bar{x}_0$ 和 $x_2x_1x_0$，得到近似优化后的逻辑函数表达式为 $z'_0=\bar{x}_2\bar{x}_1x_0+\bar{x}_2x_1\bar{x}_0$ 和 $z'_1=x_2x_1+x_1x_0+x_2x_0$，将上述 2 个近似逻辑函数转化为等效的多级电路，再进行多级优化后对应的 AIG 表示如图 5 所示。对比图 4(b)，图 5 中的 AIG 包含的节点数更少。

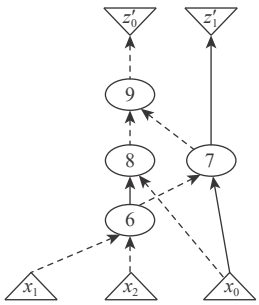


图 5　1 位全加电路经二级/多级近似优化后的 AIG 表示

Fig. 5　AIG of 1-bit full adder circuit after optimized using two-level/multi-level approximate computing

从上面例子可以发现, 相比于单一的多级近似优化, 结合二级和多级的混合层级优化可以得到更好的优化效果。因此, 在本文的混合层级近似优化中, 将总的错误率约束e_{th}分成二部分, 分别用于二级和多级近似优化。错误率分配方法如下:

$$e_{\text{alloc}} = \alpha e_{\text{th}} + \beta e_{\text{th}} \tag{3}$$

式中: α和β分别为错误率分配系数, e_{alloc}为近似优化中实际引入的错误率。考虑到在近似优化中, 多数情况下无法完全利用错误率提供的优化空间, 即$\alpha+\beta \leqslant 1$, 因此, $e_{\text{alloc}} \leqslant e_{\text{th}}$。当$\alpha$=1, β=0 时, 表示仅仅进行二级近似优化; 反之, 当α=0, β=1 时, 表示仅仅对电路进行多级近似优化。由式 (3) 可知, 单一的二级或多级近似优化仅仅是混合层级近似优化的一个特例。

3　逻辑电路面积近似优化算法

3.1　错误率分配算法

基于错误率分配的逻辑电路面积混合层级近似优化的基本思想是对于某个逻辑电路 C, 由于 C 既可以表示成二级形式也可以表示成多级形式, 且单位错误率Δe在 C 的二级和多级近似优化中效果也可能存在差异, 因此, 将总的错误率约束e_{th}分成如式 (3) 所示的αe_{th}和βe_{th}2 部分; 将αe_{th}用于 C 的二级近似优化, 得到C_{T}, 然后将C_{T}转化为等效的多级电路形式C_{M}, 并将βe_{th}用于C_{M}的多级电路的近似优化电路C_{app}。显然实现上述电路近似优化的关键问题是如何分配错误约束用于二级和多级近似优化, 亦即如何确定式 (3) 中α和β的值。对于电路 C, 可能存在多种α和β的组合, 使得 C 的混合层级近似优化结果要好于单一的二级或多级近似优化结果。

由于二级与多级表示之间可以等效转化, 因此, 对二级电路进行多级近似优化时, 如何分配错误约束用于二级和多级近似优化是关键问题, 亦即如何确定式 (3) 中α和β的值。

α和β的确定实际上是一个组合优化的问题, 通常可以采用诸如遗传算法、模拟退火等方法求解。但在本文中, 由于总的错误率约束e_{th}被设置为 5%, 并且考虑到过于小的单位错误率Δe变化对电路面积优化程度的影响不明显。因此, 在能验证错误率分配对电路面积优化的影响情况下, 同时能减少算法迭代次数, 提升算法运算效率, 本文在错误率分配中, 采用如下方法:

1) 只确定了α值, 而把剩余错误率约束留给多级优化, 即$\beta = (1-\alpha)$。

2) 采用等步长结合局部缩小步长的方法确定α值。具体做法是将用于二级近似优化的错误率αER_{th}的值以 1% 的增量从 0 到 5% 进行变化, 计算不同α值下, 电路面积的优化效果; 然后, 确定优化效果最好的αe_{th}值, 记为$\alpha_{\text{B}} e_{\text{th}}$, 并再次计算当$\alpha e_{\text{th}} = (\alpha_{\text{B}} e_{\text{th}} + 0.5\%)$和$\alpha = (\alpha_{\text{B}} - 0.5\%)$时面积优化的效果。若$\alpha_{\text{B}} e_{\text{th}} = 0$或$\alpha_{\text{B}} e_{\text{th}} = 5\%$, 则只增加那个属于 0～5% 范围内的点。最终的优化结果取所有结果中优化效果最好的值。

3.2　算法实现

算法 1 给出了基于错误率分配的电路面积近似优化算法的伪代码。

算法 1　错误率分配迭代优化算法

输入: 原始电路C, 错误率阈值e_{th}。

输出: 近似优化后的电路C_{app}。

1: $\alpha = 0$;
2: C_{app} = TM_app(C, e_{th}, α);
3: S_{app} = CountNode(C_{app});
4: $t = 0$;
5: do{
6:　α=gen(α, t);
7:　C_{temp} = TM_app(C, e_{th}, α);
8:　S_{temp} = CountNode(C_{temp});
9:　C_{app} = find_best_C($C_{\text{app}}, S_{\text{temp}}$);
10:　}while(t = 1)
11: return C_{app};

算法 1 中, 子程序TM_App(C, e_{th}, α)用来实现在确定的错误率分配情况下, 电路面积的二级/多级混合层级近似优化。其中, e_{th}为总的错误率约束, α为错误率分配系数, 用于二级电路近似优化。原始电路C为二级电路, 最终优化后得到的C_{app}为多级电路。子程序CountNode(C_{temp})用于统计C_{temp}包含的节点数, 用于衡量C_{temp}的面积。子程序α=gen(α, t)用来生成不同的α值和程序循环控制。当遍历所有α值后, α=gen(α, t)会令 t=1, 表示搜索循环的结束。子程序find_best_C($C_{\text{app}}, S_{\text{temp}}$)用来搜索优化效果最

好结果并存储为C_{app}。由于初始化时$\alpha=0$，因此，算法1初始时对电路进行的是单纯的多级近似优化。

算法1中子程序TM_App(C,e_{th},α)是由电路的二级和多级面积近似优化算法构成，其伪代码见算法2。其中原始输入电路C为二级电路，输出的是采用AIG格式表示的多级电路。

算法2 二级多级协同近似优化算法

输入： 输入电路C，错误率阈值e_{th}及错误率分配系数α。

输出： 近似优化后的电路$C_{\text{M_aig}}$。

1: C_{esp} = Espresso(C);

2: C_{T} = TL_App($C_{\text{esp}},e_{\text{th}},\alpha$);

3: $C_{\text{org_aig}}$ = CouvertAigOpt(C_{esp});

4: $C_{\text{T_aig}}$ = CouvertAigOpt(C_{T});

5: $C_{\text{M_aig}}$ = ML_App($C_{\text{org_aig}},C_{\text{T_aig}},e_{\text{th}}$);

6: return $C_{\text{M_aig}}$;

算法2中Espresso(C)指用二级优化工具Espresso对C进行二级精确优化，得到优化后的电路C_{esp}。TL_App($C_{\text{esp}},e_{\text{th}},\alpha$)表示对电路$C_{\text{esp}}$进行二级近似优化，其中错误率约束为$\alpha e_{\text{th}}$，$C_{\text{T}}$为经二级近似优化后得到的电路。程序CouvertAigOpt(C_{esp})用来将二级电路C_{esp}转化为等效的多级电路。程序ML_App($C_{\text{org_aig}},C_{\text{T_aig}},e_{\text{th}}$)用于实现电路面积的多级近似优化，其中，多级近似优化算法采用文献[11]的算法，结合多扇出节点常量替换过程的错误率控制、在全局范围提取适合替换的候选节点集和选择每个候选节点输出的常量替换值等方法实现面积优化。在ML_App($C_{\text{org_aig}},C_{\text{T_aig}},e_{\text{th}}$)中，错误率的计算采用如图1所示的Miter电路，使得在$C_{\text{T_aig}}$基础上进行的多级近似优化后得到的近似电路$C_{\text{M_aig}}$与原始电路$C_{\text{org_aig}}$之间的错误率不超过e_{th}。

从算法2不难看出，二级/多级混合层级近似优化后的最终错误率是在多级近似优化中借助Miter电路实现的，亦即二级电路近似优化中的错误率计算的精度并不会影响多级近似优化计算精度。但在以往二级近似优化中，错误率计算会占用大量的算法时间和内存，进而影响算法的运算速度。为提高混合层级近似优化算法效率，在二级近似优化中，采取了如下措施，提高二级近似优化中的运算速度。

1）二级近似优化可以通过乘积项删除实现。在乘积项删除过程中为了尽可能多地减少字母数，同时又能引入尽可能少错误，在乘积项选择上，本文优先考虑删除包含字母多的乘积项。因为这类乘积项包含最小项数量较少，删除后引入的错误也较少，有助于防止错误率的陡增，避免突破约束。同时，缓慢增加的错误率，可以更好地利用错误率约束提供的优化空间，提高优化效果。上述的乘积项选择策略可以将原先基于探测或尝试的复杂搜索过程变成按乘积项中包含的字母数多少的排序过程，大大减少了二级优化时间。

2）简化错误率计算方法。精确的错误率计算需要考虑所有可能的输入组合，而每个最小项就对应着1个输入组合，因此，随着输入变量的增加，可能的组合数呈指数级增长，使得错误率精确计算变得非常耗时。但考虑到本文最终的错误率是由多级近似优化决定，二级近似优化仅仅是1个中间过程，且二级近似优化中的错误计算精度不会影响多级近似优化中错误率的计算。因此，在二级近似优化中错误率计算中采用式(4)来估算在二级近似优化中因乘积项删除而带来的错误率。

$$e_{\text{app}}=\frac{\sum 2^{k_i}}{2^n} \tag{4}$$

式(4)中：k_i为第i个被删除的乘积项中包含的未出现的变量个数。由于被删除的乘积项中未出现的变量数统计是一个非常快的过程，因此，式(4)可以显著减少错误率计算时间，进而提高算法的运算速度。

显然，只有在所有乘积项均不相交时，由式(4)得到的错误率才等于实际错误率[19]，而当乘积项之间存在相交情况时，式(4)得到的错误率要大于实际错误率。因此，用式(4)进行错误率控制时，可以是适当松弛错误率约束。

算法2中的用于二级近似优化的子程序TL_App($C_{\text{esp}},e_{\text{th}},\alpha$)的伪代码见算法3。其中，$b_i$ = Count(p_i)用来统计乘积项p_i中未出现的变量个数，并将其结果存入集合C_b中。而子程序SortByCnt(C,C_b)则是基于乘积项中未出现的变量个数实现对原始电路C中的各个乘积项的排序操作。子程序GetMin(C_{app},C_b)和ProRemove(C_{app},p_i)分别用来选择未出现的变量个数最少的乘积项$p_{\min}$和在C_{app}中实现对选定乘积项$p_{\min}$的删除操作。

算法3 二级近似优化算法

输入： 原始电路C，错误率阈值e_{th}及错误率分配系数α。

输出： 近似优化后的逻辑函数乘积项集合C_{app}。

1: $C_{\text{app}}=\varnothing,C_b=\varnothing,m=0,e_{\text{app}}=0$;

2: for(each product p_i in C){

3: b_i = Count(p_i);

4: $C_b=C_b\cup\{b_i\}$;}

5: C_{app} = SortByCnt(C,C_b);

6: do{$p_{\min}$ = GetMin(C_{app}, C_b);

7: m+ = pow(2, $b_{\min}$);

8: $e_{\text{app}} = \dfrac{m}{\text{pow}(2,n)}$;

9: C_{app} = ProRemove(C_{app}, $p_{\min}$);

10: }while($e_{\text{app}} < \alpha * e_{\text{th}}$)

11: return C_{app};

算法 3 中, 伪代码语句 6-11 是利用乘积项删除操作, 并结合式 (4) 错误率估算方法实现电路二级近似优化。其中, 删除操作是根据排序的结果依次删除含字母多的乘积项。显然电路二级近似优化过程是一个在错误率约束下的逐次迭代的过程, 快速的错误率估算有助于算法速度的提升。

4　实验结果分析

本文算法用 C 语言编程实现并结合面向时序逻辑综合与形式验证系统(a system for sequential logic synthesis and formal verification, ABC)和 Espresso 完成逻辑电路面积优化[20]。运行环境为: 处理器频率 2.5 GHz, 主内存 64 GB, 操作系统为 CentOS7.0。表 1 为本次实验用到的北卡罗莱纳微电子中心 (Microelectronics Center of North Carolina, MCNC)[21] 测试电路, 映射的工艺库取自 MCNC, 其中的“初始面积”以及“初始延迟”为通过 ABC 的 resyn2 命令精确优化后再进行 MCNC 工艺库映射所得到的。

表 1　本文使用的 MCNC 测试电路

Table 1　The parameters of MCNC benchmarks used in this paper

电路名称	I/O	初始面积	初始延迟/ns
alu4	14/8	1 158	49.81
table3	14/14	2 836	29.9
b12	15/9	98	8.5
table5	17/9	2 563	31.6
t481	16/1	1 400	24.9
clip	9/5	193	14
sqrt8	8/4	101	11.9
rd84	8/4	301	18.4
apex4	9/19	4 685	27.4
misex1	8/7	103	9.2
5xp1	7/10	160	12.1
inc	7/9	184	10.3

本文中逻辑电路面积优化程度用面积优化率 r_{area} 来表示, 如式 (5) 所示。式 (5) 中 S_{org} 为初始面积。同理也可以得到延时优化率。

$$r_{\text{area}} = \frac{S_{\text{org}} - S_{\text{app}}}{S_{\text{org}}} \times 100\% \qquad (5)$$

表 2 给出了在错误率阈值为 5% 的情况下, 本文算法与文献 [11] 算法和文献 [12] 算法的比较结果。由于文献 [11] 算法和文献 [12] 算法的结果有一定的随机性, 因此, 表 2 的数据是各算法 3 次运算结果的平均值。

表 2　本文算法与文献 [11] 算法和文献 [12] 算法的比较结果

Table 2　Comparison results of the proposed algorithm with references [11] and [12] algorithm

电路名称	面积优化/%			延迟优化/%		
	文献[11]	文献[12]	本文算法	文献[11]	文献[12]	本文算法
alu4	11.20	24.84	22.71	7.25	19.03	46.80
table3	90.63	90.09	91.27	48.61	57.41	54.07
b12	24.49	21.77	26.53	16.47	7.84	16.47
table5	94.49	92.91	94.98	63.83	55.49	62.45
t481	87.64	93.21	96.93	36.55	40.03	66.27
clip	28.67	0	28.15	0	0	4.29
sqrt8	42.57	37.29	43.56	13.45	0	7.56
rd84	19.93	48.28	36.10	0	23.55	20.65
apex4	4.55	14.88	20.65	0	11.19	11.20
misex1	6.80	3.56	10.68	0	3.26	8.70
5xp1	16.67	0.42	16.67	3.31	0	3.31
inc	20.11	2.17	20.11	0	0	0

注: 各列平均值为37.31%, 35.79%, 42.36%, 15.79%, 18.15%, 25.15%。

在表 2 中, 相比于文献 [11] 算法, 本文算法得到的优化结果可分为如下 3 种情况: ① 面积和延时优化结果均不低于文献 [11] 算法。这类电路包括 alu4, table3, b12, t481, rd84 等大多数电路; ② 2 种算法在面积和延时优化上各有长处, 但结果比较接近。如电路 clip 和 sqrt8; ③ 2 种算法优化结果一样或非常接近, 如电路 table5, inc。

产生上述结果的主要原因是: 本文中多级优化采用的是文献 [11] 的算法, 从算法 1 中可知, 本文算法总是选择在各种错误分配中面积优化效果最好的作为最终输出, 这就是本文算法在面积优化不低于文献 [11] 的原因。但有一个例外, 就是 clip 电路。clip 的面积优化结果要稍逊于文献 [11] 算法, 这与文献 [11] 算法结果存在一定的随机性有关, 即用文献 [11] 算法对于同一电路进行优化, 并不是每次结果都一样。因此, 在每次执行本文算法时, 实际上已经选择了本次运算中面积优化效果最好的一次, 但却无法保证本次结果是所有优化中最好的, 进而导致本文算法对 clip 面积优化结果与文献 [11] 算法存在差异的原因。同理, 也是导致 table5 的结果存在小的差异的原因。考虑到, 电路 table5 和 inc 的单纯多级优化结果和二级/多级优化结果非常接近或一样。因此, table5 和 inc 不适合采

用二级/多级优化。另外，本文算法与文献 [12] 算法的结果相比，并不是所有的电路面积优化结果都优于文献 [12] 算法，如在 alu4 和 rd84 面积优化上，文献 [12] 的算法优于本文算法。但是在 t481，apex4 面积优化上，相比于单纯使用文献 [11] 算法，本文算法在面积优化效果上超过了文献 [12] 算法。其原因是本文提出的基于错误率分配的电路面积近似优化算法能有效提升电路面积的优化效果，但同时集成在本文算法中的电路面积多级优化算法，即文献 [11] 算法，对最终面积优化效果影响很大。

从表 2 中数据分析可得，除面积优化以外，多数情况下，近似计算技术同时可以实现延时的优化。本文电路的延时是由 AIG 的关键路径长度（层级数）来衡量的。在错误率阈值为 5% 的情况下，相较于仅使用多级近似优化的文献 [11] 算法和文献 [12] 算法，本文算法在平均延时优化分别提升了 9.36% 和 7.0%。与文献 [11] 算法相比，采用本文算法后，绝大多数电路的延时得到了进一步优化，有些延时无法优化的电路，如 clip, rd84, apex4 等，都得到了进一步优化。类似的情况也出现与文献 [12] 算法的对比中。总体而言，相比于单一的多级近似优化算法，对于多数电路而言，本文的混合层级近似优化算法在对电路进行多级近似优化时可以获得更好的优化效果。

在算法时间复杂度方面，由于本文优化算法采用了二级和多级相结合的优化策略，其中二级优化算法的时间复杂度与电路的输入变量数 I、输出变量数 Q 和乘积项数 N 有关，其时间复杂度表示为 $O((I+Q)N)$；多级优化算法的时间复杂度与待优化的 AIG 中包含的节点数 V，待优化候选节点数 M，算法迭代次数 K 有关，其时间复杂度表示为 $O(K(MV+V\log_2 V))$。由此根据算法 1 的错误率分配方法，本文算法的总时间复杂度为 $O(T((I+Q)N+K(MV+V\log_2 V)))$，其中，$T$ 为对算法 2 的调用次数。由于电路的输入输出数量通常远远小于其节点数。因此，多级优化算法在整体运行时间中占主导地位。

表 3 列出了在错误率阈值为 5% 的情况下，本文算法与文献 [11] 和文献 [12] 的方法各自的运行时间。因本文算法中的多级近似优化是通过多次调用文献 [11] 的算法实现，因此，本文算法的平均运行时间比文献 [11] 算法长了约 3.5 倍。另外，对于 alu4 电路，由于每次二级近似优化后会引起 AIG 节点数增加，导致 alu4 电路的后继的多级近似优化时间的增加，多次迭代后，最终导致该电路整个优化时间的大幅增加。相比文献 [12] 算法，本文算法的平均运行时间仅为文献 [12] 算法的 70% 左右。

表 3 本文算法与文献 [11] 算法和文献 [12] 算法的运行时间

Table 3 Runtime of the proposed algorithm with references [11] and [12] algorithm

电路名称	时间/s		
	文献[11]算法	文献[12]算法	本文算法
alu4	78.64	318.06	1 688.56
table3	706.42	2 224.44	2 284.05
b12	1.46	0.53	7.76
table5	542.00	632.45	1 285.45
t481	209.88	246.85	245.48
clip	8.41	0.33	43.81
sqrt8	0.81	0.53	6.19
rd84	14.37	4.83	34.12
apex4	289.43	5 573.21	926.52
misex1	0.98	0.68	4.64
5xp1	2.62	2.45	16.74
inc	2.31	1.09	13.06

注：各列平均值为154.78, 750.45, 546.36 s。

这是由于文献 [12] 算法在处理 AIG 节点数多的电路时，其每一次优化迭代获得的节点数减少量相比待处理的 AIG 包含的节点数是一个小量，因此，需要大量的迭代优化，且每次迭代耗时又较长，导致整个算法运行时间很大[11]，从而增加表 3 中文献 [12] 算法的平均运行时间。

5 结 论

1）由于错误率优化效能在二级和多级情况存在差异，且电路功能的描述在二级和多级之间可以等效地转化，本文提出了基于错误率分配的逻辑电路多级近似面积优化算法，在满足整体错误率约束下将总的错误率进行适当地切分，分别用于二级和多级近似优化，并提出了错误率分配算法。

2）为加快算法速度，提出了基于错误估算的二级近似优化算法。实验结果表明，相比于单一的多级近似优化算法，对于绝大多数测试电路，本文算法电路面积和延时上的优化效果都有所提升。

参考文献（References）

[1] XU Q, MYTKOWICZ T, KIM N S. Approximate computing: a survey[J]. IEEE Design & Test, 2016, 33(1): 8-22.

[2] SEKANINA L. Introduction to approximate computing: embedded tutorial[C]//Proceedings of the IEEE 19th International Symposium on Design and Diagnostics of Electronic Circuits & Systems. Piscataway: IEEE Press, 2016: 1-6.

[3] WU Y, CHEN C T, XIAO W H, et al. A survey on approximate multiplier designs for energy efficiency: from algorithms to circuits[J]. ACM Transactions on Design Automation of Electronic Systems, 2024, 29(1): 1-37.

[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] SEO H, KIM Y. A low latency approximate adder design based on dual sub-adders with error recovery[J]. IEEE Transactions on Emerging Topics in Computing, 2023, 11(3): 811-816.

[6] MANNEPALLI Y, KOREDE V B, RAO M. Novel approximate multiplier designs for edge detection application[C]//Proceedings of the Great Lakes Symposium on VLSI. New York: ACM, 2021: 371-377.

[7] NGUYEN D T, KIM H, LEE H J, et al. An approximate memory architecture for a reduction of refresh power consumption in deep learning applications[C]//Proceedings of the IEEE International Symposium on Circuits and Systems. Piscataway: IEEE Press, 2018: 1-5.

[8] SEN S, RAGHUNATHAN A. Approximate computing for long short term memory (LSTM) neural networks[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2018, 37(11): 2266-2276.

[9] SHIN D, GUPTA S K. Approximate logic synthesis for error tolerant applications[C]//Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (DATE 2010). Piscataway: IEEE Press, 2010: 957-960.

[10] FUJITA M. An approach to approximate computing: logic transformations for one-minterm changes in specification[C]//Proceedings of the IEEE International High Level Design Validation and Test Workshop. Piscataway: IEEE Press, 2017: 91-94.

[11] 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).

[12] 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.

[13] 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.

[14] MENG C, ZHOU Z Z, YAO Y, et al. HEDALS: highly efficient delay-driven approximate logic synthesis[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2023, 42(11): 3491-3504.

[15] 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.

[16] SU S B, ZOU C, KONG W J, et al. A novel heuristic search method for two-level approximate logic synthesis[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2020, 39(3): 654-669.

[17] 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.

[18] AMMES G, NETO W L, BUTZEN P, et al. A two-level approximate logic synthesis combining cube insertion and removal[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2022, 41(11): 5126-5130.

[19] 王伦耀, 夏银水, 储著飞. 基于近似计算技术的 FPRM 电路面积优化[J]. 电子学报, 2019, 47(9): 1868-1874.
WANG L Y, XIA Y S, CHU Z F. Area optimization of FPRM circuits using approximate computing[J]. Acta Electronica Sinica, 2019, 47(9): 1868-1874(in Chinese).

[20] BRAYTON R K, HACHTEL G D, MCMULLEN C T, et al. Logic minimization algorithms for VLSI synthesis[M]. Berlin: Springer, 1984.

[21] YANG S. Logic synthesis and optimization benchmarks user guide: version 3.0[M]. North Carolina: Microelectronics Center of North Carolina, 1991.

Area approximate optimization of logic circuits based on error rate allocation

FAN Haimei，WU Qianhuo，WANG Lunyao*

(Faculty of Electrical Engineering and Computer Science，Ningbo University，Ningbo 315211，China)

Abstract：A multi-level area optimization algorithm for logic circuits based on error rate allocation is proposed, taking into account that the approximate optimization effects of a logic function under the same error rate constraint differ when the function is expressed in two-level or multi-level forms and that there is an equivalent conversion between two-level and multi-level forms. This algorithm consists of a two-level error rate estimation technique, an error rate coefficient search method for error rate division, and a two-level/multi-level approximate optimization method. The proposed algorithm is implemented using C, the commands of a system for sequential logic synthesis and formal verification (ABC) tools and tested with Microelectronics Center of North Carolina (MCNC) benchmarks. The experimental results show that, compared to the ABC tools, with the 5% error rate, the proposed algorithm achieves 42.36% and 25.15% in area and delay optimization, respectively. Compared to those approximate methods that only use the multi-level approximation method, further improvement in area and delay can be obtained by 5.05% and 9.36%, respectively.

Keywords：approximate computing；logic synthesis；approximate optimization；error rate allocation；area optimization

Received：2024-04-11；**Accepted**：2024-07-16；**Published Online**：2024-08-05 09：06

URL：link.cnki.net/urlid/11.2625.V.20240802.1441.002

Foundation items：National Natural Science Foundation of China (U23A20351,U22A2013); Major Research Plan of Ningbo (2023Z233)

* **Corresponding author**. E-mail：wanglunyao@nbu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0298

基于超级块嵌套的动态二进制翻译性能优化方法

金帅，柴志雷*，周浩杰
(江南大学 人工智能与计算机学院，无锡 214122)

摘 要：直接块链接方法可以在动态二进制翻译过程中降低翻译器介入频率、提升目标程序运行性能，单纯的直接块链接只关注逻辑层面的连接，并未关注物理存储位置的连续。针对该问题，提出了一种基于超级块嵌套的动态二进制翻译性能优化方法。基于二进制翻译器预翻译目标代码生成的中间表示(IR)指令，通过构建嵌套结构的超级块使代码实现物理存储位置的连续，从而增加指令cache命中率。实验结果表明：以动态二进制翻译领域最经典的且最常用的SPEC2006作为测试基准，以主流的开源二进制翻译器QEMU6.0为实验框架时，相比未采用超级块嵌套的版本，所提方法可将cache命中率平均提升53.28%，最大可提升90%。目标程序运行性能平均提升3.07%，最高可提升4.58%。

关 键 词：动态二进制翻译；超级块；中间码嵌套；翻译器QEMU；指令cache
中图分类号：TP314
文献标志码：A **文章编号**：1001-5965(2026)06-2123-10

二进制翻译源于体系结构开发过程中的代码移植问题，用于将一种指令集架构(instruction set architecture, ISA)的二进制代码转换为另一种ISA的对应代码。动态二进制翻译(dynamic binary translation, DBT)技术作为一种特殊的即时编译技术，边翻译边执行，并在翻译过程中进行动态优化。由于在将源机的二进制代码翻译到目标机代码的过程中，动态二进制翻译器需要在运行时动态翻译和优化指令，因此通常会引入性能开销。系统开销通常分为两部分：翻译开销和执行目标代码的开销。事实上，目前大多数DBT系统都可以将转换成本降低到可接受的范围内。QEMU[1]为实现对多目标机和多宿主机的兼容，并未充分利用宿主机体系结构特征，未经优化的QEMU翻译效率通常只有本地编译执行的10%左右[2]。因此，主要关注的是减少目标代码的执行时间。

近年来，为了减少执行目标代码的时间，研究人员提出了多种优化方法。其中，Shen等[3]提出了翻译块链接技术，通过将多个直接跳转的翻译块连接起来，从而大幅减少翻译模块与执行模块之间的上下文切换次数，以提高性能。另外，Ebcioglu等[4]设计了一种名为VLIW的架构，能够动态编译现有指令集架构的目标代码。该架构实现了与已有架构的兼容性，并利用动态适应和运行时配置文件信息来优化翻译，从而提取指令级并行性。

Hwu等[5]提出了一种基于静态分析的优化技术，即超级块。超级块[6]是通过合并多个基本块而形成的。超级块只允许一个入口，但可以有多个出口。这种优化技术具有多个优点，例如减少翻译模块与执行模块之间的上下文切换次数、降低高速缓存的未命中率及减少冗余的转换旁视缓冲器(translation lookaside buffer, TLB)访问和降低TLB未命中率。

尽管超级块在二进制翻译中得到广泛应用，并在提高编译性能方面取得了显著效果，但也存在一些局限性：

收稿日期：2024-05-08；**录用日期**：2024-09-23；**网络出版时间**：2024-10-08 13:18
网络出版地址：link.cnki.net/urlid/11.2625.V.20241008.0955.001
基金项目：国家自然科学基金(61972180)
*通信作者. E-mail：zlchai@jiangnan.edu.cn

引用格式：金帅，柴志雷，周浩杰. 基于超级块嵌套的动态二进制翻译性能优化方法[J]. 北京航空航天大学学报，2026，52(6)：2123-2132. JIN S，CHAI Z L，ZHOU H J. Superblock nesting based on performance optimization for dynamic binary translation[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：2123-2132 (in Chinese).

1）静态分析可能无法完全准确地捕捉动态程序行为和控制流的变化，这可能导致超级块的构建和优化结果的精确性有所欠缺。

2）超级块的构建和优化通常在基本块的边界上进行，这可能导致忽略一些跨边界的优化机会。例如，某些指令之间的相关性可能无法被充分利用，从而限制了超级块的优化能力。

本文遵循超级块的主题思想，基于用户级 QEMU 设计了一种超级块嵌套的动态二进制翻译方法 SN-QEMU。该方法分为前端翻译执行模块和后端优化模块两部分，前端翻译执行模块根据程序局部热路径预测算法进行热路径选择识别，后端优化模块则通过超级块嵌套处理和条件跳转优化进行代码优化。该方法适用于基于静态分析或动态分析技术构建超级块的情况，特别是在数据预取、代码缓存管理及本地化资源管理等场景，使用 SN-QEMU 方法可以带来显著收益。

将超级块以嵌套的形式实现，可以在一定程度上解决超级块本身的不足之处。通过嵌套超级块，可以减少生成的指令数量，从而节省空间。同时，可以更精确地捕捉程序的控制流和依赖关系。通过在更细粒度的级别上进行超级块的构建和优化，可以更好地处理复杂的控制流结构，减少对静态分析的依赖，从而提高优化的准确性。同时，嵌套超级块可以跨越基本块边界，提供更多的优化机会。这样可以更好地处理跨边界的指令相关性，避免因为基本块边界限制而导致的优化机会的丧失。

本文主要从构建超级块并修改中间码空间分布的方向出发，达到提高 cache 命中率的目的。通过在 SPEC2006 基准测试下检测 QEMU 程序运行总时间和 cache_Misses 等数据可以得出系统的优化程度。数据表明，该技术可以运用到各种 SPEC 2006 基准测试中，程序平均运行总时间比未修改版本提升了 3.07%。

1 研究背景

1.1 QEMU

QEMU 作为当前主流的多平台开源二进制翻译框架，具备卓越的可移植性和可扩展性。近年来，围绕 QEMU 涌现了大量研究成果，涉及寄存器分配优化[7]、中间代码优化[8]、基本块链接优化、多线程优化[9]等领域。同时，也衍生出了众多工具，如基于底层虚拟机（low level virtual machine, LLVM）优化加速的多线程翻译器 HQEMU、符号执行与二进制翻译融合的 SymQEMU[10]、缓存共享的并行全系统模拟器 PQEMU[11]、缓存独占的并行全系统模拟器 COREMU[12]、基于分布式框架的 DQEMU[13]、兼容 Pin 的二进制插桩工具 PEMU[14]、基于多面体优化实现程序并行化的 LLPEMU[15] 及动态二进制分析工具 PANDA[16] 等。

QEMU 使用动态二进制翻译技术将目标架构的指令转换为宿主架构的指令。其能够在宿主系统上执行不同架构的代码。QEMU 首先将 GUEST 端指令反汇编生成 TCG 中间指令，之后由翻译器将小型代码生成器（tiny code generator, TCG）中间代码翻译成目标机体系结构二进制代码，并组成翻译块存放在 Tcache 缓存中，此时就得到了 HOST 端代码。宿主机上的本地代码执行模块就可以从 Tcache 中取出目标翻译块来执行。QEMU 工作流程如图 1 所示。

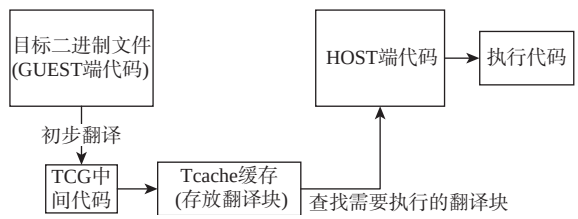


图 1 QEMU 工作流程

Fig. 1 QEMU workflow

1.2 超级块

在二进制翻译领域，超级块是一种由多个基本块组成的连续指令序列。基本块作为一次翻译和执行的基本单位，由一组连续执行的指令构成，并且只有一个入口和一个出口。一个基本块经翻译后可生成一个或多个翻译块（translated block, TB）[1]。

翻译块执行入口在函数 cpu-exec() 中，该函数在 QEMU 中的作用是执行 CPU 指令。其是 QEMU 的指令执行路径的一部分，负责调用适当的指令执行函数，并处理不同指令类型的执行逻辑。代码从 cpu-exec() 到宿主架构指令的执行过程中，存在上下文切换的过程，在这个过程中需要保留和恢复 QEMU 使用的寄存器，因此需要有一个生成环境代码区域 prologue(序言) 和恢复环境代码区域 epilogue(序尾)。翻译块执行流程如图 2 所示。

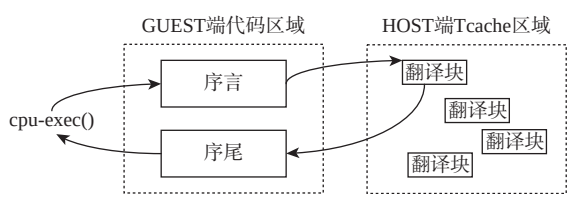


图 2 翻译块执行流程

Fig. 2 Translation block execution process

超级块的主要目标是通过减少分支开销、提高指令级并行性和利用数据局部性来优化二进制代

码的执行效率。通过将多个基本块组合成超级块, 可以减少分支指令的执行次数, 从而降低分支预测错误的概率, 提高指令级并行性。同时, 超级块内的指令序列可以更好地利用处理器的缓存层次结构, 增加数据局部性, 减少内存访问延迟。图 3 给出了一个超级块的说明示例: 左边表示一段代码的控制流图, 其中, 节点表示基本块, 节点中的数字表示该基本块的执行频率, 边表示基本块之间的控制转移关系。虚线框部分表示由该控制流图所构成的序列(Trace[17]), 共有 3 个 Trace: ABD、CE、F。图 3(b)表示超级块的最后生成状态。

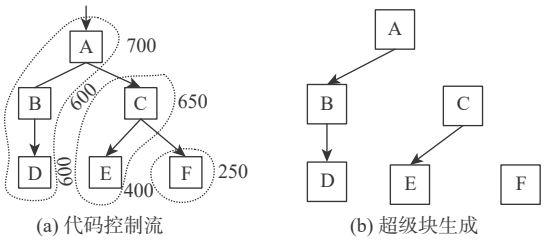


图 3　超级块构建

Fig. 3　Superblock construction

1.3　基于数据嵌套思想的二进制翻译 cache 优化

在编译领域, 研究通过数据嵌套的方法来提高 cache 命中率是一项非常重要的工作。许多研究团队和学者已经开展了一系列针对 cache 命中率的研究。文献 [18] 提出了结构分裂和字段重新排序 2 种技术, 从重组数据结构的角度利用空间局部性, 以降低 cache 冲突。该方法在 Nbench 测试集上展现了平均加速比 2%～3% 的效果。文献 [19] 提出了 Cache-Conscious-Data 框架, 通过将具有时间局部性的代码映射到虚拟地址空间中的不同高速缓存块, 以消除高速缓存冲突。文献 [20] 提出了对完全关联和集合关联高速缓存实施一些必要条件和充分条件的特性, 这些高速缓存允许在不同层次上使用不同大小的块。该框架以两级总线为基础, 实现了包含属性的缓存层次结构, 通过嵌套的缓存层次结构来提高缓存命中率和性能。

2　SN-QEMU 方法

2.1　方法设计

SN-QEMU 方法如图 4 所示, 展示了方法中搜集热点代码、构建超级块及动态优化的整体流程。该方法被划分为前端翻译执行模块和后端优化模块两部分。前端翻译执行模块管理仿真环境并对所有生成的翻译块进行头部代码插桩, 以实现对热点代码的监测。当翻译块的执行次数达到预设的阈值时, 会进行热点代码的搜集, 并构建热路径包。后端优化模块根据前端收集的热路径包, 构建超级块, 并对其进行优化。为了构建超级块, 本文采用以下方法: 在原始翻译中对基本块进行翻译时, 保留各个基本块对应的 TCG 中间码。然后, 重新翻译热路径中翻译块对应的中间码, 最终实现超级块的构建。

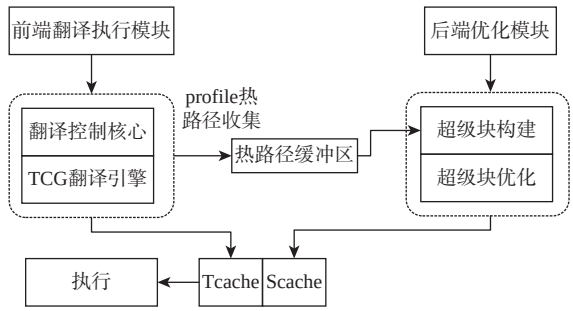


图 4　SN-QEMU 方法

Fig. 4　SN-QEMU approach

2.2　SN-QEMU 探测插桩块

由于二进制翻译器在执行翻译后的代码是以翻译块为执行单位, 各个翻译块之间相互独立。因此, 本文通过引入代码插桩, 从而更好地监测出频繁被执行的翻译块。插入的插桩块代码分为 Head_stub、Get_stub 和 Br 跳转指令 3 个部分。

1) Head_stub 用于监测该翻译块是否为头块, 即一段频繁执行的翻译块序列的第一个翻译块, 并增加引用次数。

2) Get_stub 用于收集 Head_stub 中监测出来的翻译块到一个热路径包 (tb_packet) 中, 然后根据收集到的热路径信息构建出超级块, 进而做出优化。

3) Br 跳转指令则是为了实现从翻译块到超级块的重新定位。Br 跳转指令的初始跳转地址为下一条指令, 当以某一翻译块 a 为头块构建超级块 A 后, 修改 Br 跳转指令的跳转地址为超级块 A 的位置, 下次系统再次执行到 a 时, 可以跳过翻译块 a, 直接执行超级块 A。

为了减小因为插桩带来的开销, 本文通过引入标志位 flag_profile 来控制超级块构建。当标志位为 0 时, 开始构建超级块, 当标志位为 1 时, 停止超级块的构建。同时, 后续的翻译块头部将不再生成插桩代码。

由于 QEMU 对复杂的目标指令功能不产生对应的主机代码, 而是通过调用 helper 函数来完成其具体功能。因此, 本文针对插桩代码, 采用 2 个 helper 函数去实现该功能, 即 gen_helper_head() 和 gen_helper_get()。

2.3　MCNET 热路径识别策略

在动态二进制翻译中, 常见的 profile 和热路径识别方法有 3 种[21]: 基于基本块 profile 的热路径识

别，基于边 profile 的热路径识别[22]，以及基于路径 profile 的热路径识别。这些方法在预测热路径时的准确率递增，但同时也增加了实现的难度和算法的复杂度。在实际应用中，可以根据具体情况选择其中一种方法，或者进行组合使用。

Dynamo[23] 是 HP Lab 开发的动态优化系统，采用了新颖的 NET(next execution tail) [24] 热路径预测方式。DynamoRIO[25]、Walkabout[26]、Mojo[27] 等与 Dynamo 类似的热路径选择方法在很多系统广泛应用。NET 算法是一种“先热先选择”的策略。NET 算法核心思想如下：定位基本块的头块，当头块被执行时，增加其引用计数；当头块的引用计数超出给定阈值时，将该头块及后续的基本块作为一条热路径，直到再次执行到该头块结束。NET 算法定义头块为分支跳转的目标，只将头块进行引用计数累计、阈值判断等操作，其他非头块的基本块不进行以上操作。

本文结合文献 [25, 28] 的方法，提出了一种更高效的多条件热路径网络（multi condition NET, MCNET）探测方法。在翻译块结构体内新增了一个标志位 tb_flag，用于标识该翻译块是否为头块。如果是头块，则将该标志位置为 1，否则置为 0。初始情况下，默认将 tb_flag 设置为 0。在二进制翻译器执行查询模块时，进行头块的定位。在这个基础上，增添了多条搜集终止条件：

1）统计所有构建超级块的头块和热路径包，将热路径包内的翻译块作为热路径识别的终止条件。这是因为当超级块被多次执行后，原本热路径中未达到构建超级块阈值的翻译块有可能会达到阈值。这样一来，再次构建的超级块就会与之前的超级块重复，导致不必要的开销。

2）翻译块的最后一条指令跳转到当前热路径包的起始块，构成一个循环。

3）限制了构建的超级块内包含的翻译块数量。如果数量太少，获得的优化效果将减少。

下面给出了 MCNET 热点探测算法伪代码：

```
Br $L0
set_label $L0
Head stub:
    if flag profile=true then
        if tb-> tb_flag = 1 then
        tb-> exec_cnt +=1:
        if tb->exec_ent > THREASHOLD1
    then
            super_enable =1;
        if super head count =
    THREADSHOLD2
            flag_profile =false:
    end if
Get_stub:
    if super_enable = 1 then
        if tb not eligible for hot path
collection then
        super_head[super_head_count]=
        trace_packet.tb_packet[0];
            super_head_count +=1;
        if opc ==INDEX_op_goto_tb then
            Execute superblock nesting();
        else if ope =-INDEX_op_goto_Ptr
             End superblock nesting();
        br Instruction relocation();
    else
        Add tb to trace packet.tb_packet
    end if
```

2.4 中间码收集

由于 QEMU 在将目标代码翻译为中间 IR 的过程中不保存中间码，为解决该问题，本文采取以下策略：设置一个中间码缓冲池，用于保存翻译器将目标代码转换为中间码的过程中生成的中间码。这样可以在后续的超级块构建过程中使用这些中间码。为了区分每个翻译块对应的中间码，本文引入 2 个标志位：op_start_idx 和 op_end_idx。其分别表示每个基本块对应的起始中间码和结束中间码的索引。通过这种方式，可以明确地知道每个翻译块的中间码范围。同时，为了构建超级块，本文还收集了中间变量 temp。这些中间变量在超级块的构建过程中起到重要的作用。中间码和中间变量收集伪代码如下：

中间码收集：

```
if !super_translate_enable &&
tcg_ctx->flag_profile then
    Holds elements such as opcodes();
    Holds elements such as operands();
    gen_op_buf_size++;
    oi++;
    if opc == INDEX_op_set_label then
        label processing();
```

temp 收集：

```
if !super_translate_enable &&
tcg_ctx->flag_profile then
    tb = temp;
    collection of individual parameters of
```

```
5    intermediate variables();
6        tb->op_start_idx = gen_next_op_idx;
7        oi = gen_next_op_idx; //写入操作码的
8    位置
```

2.5　前端翻译执行模块

前端翻译执行模块主要是进行翻译块的生成和执行, 同时也为后端优化模块构建所需热路径。前端在完成 QEMU 本身的翻译和执行功能的基础上, 利用 MCNET 热路径识别策略进行热点代码的查找, 将生成的热路径按照顺序依次放到热路径缓冲区并且将翻译块对应的中间码进行收集。

同时, 对于 QEMU 在 cache 的管理机制进行功能扩充。为了更好地利用构建的超级块, SN-QEMU 方法设计了分层管理 cache 机制, 将原本的 cache 分为两部分, 即 QEMU 原本的 Tcache, 以及专门存放超级块的 Scache。该分层管理机制将 cache 的查找进行分级。Scache 是第 1 层级, 系统优先查找该层级, Tcache 是第 2 层级。在查询模块中, 首先, 查找 Scache 是否有所需翻译块, 如果没有则返回 Tcache 查找。

2.6　后端优化模块

后端优化模块主要是接收前端模块发送的热路径及对应的中间码, 并对超级块进行嵌套优化。当后端优化模块接收到热路径缓冲区的热路径之后, 后端优化模块开始构建超级块。该部分将在第 3 节进行详细说明。

3　超级块嵌套设计

3.1　超级块数量限定

Duesterwald 等[24] 提出, 如果预测被推迟太久, 预测的整体效益很容易消失。超级块生成得越晚, 该超级块重用率(即潜在利益)就越低。为了降低构建超级块的开销, 需要控制生成的超级块数量。在进行超级块嵌套方法的实验测试之前, 首先对课题进行 profile, 以获得所需的超级块数量阈值和超级块的初始创建位置。一旦超级块的数量达到设定的阈值, 就停止生成超级块, 从而减少了构建超级块的开销。

3.2　嵌套超级块结构设计

在获得热路径中每个翻译块对应的中间码后, 下一步是对中间码进行重新布局, 并将其翻译为超级块。为了实现这一目标, 本文引入 2 个标志位: head_op 和 old_op。head_op 用于定位 goto_tb 指令的位置, 而 old_op 则在每次将处理后的中间码放入 QTAIL 队列后进行更新。通过使用 QTAIL 队列和标志位的引入, 本文能够对中间码进行有序的处理和存储, 从而为后续的超级块翻译提供良好的基础。

在处理热路径中间码时, 需要对 gen_op_buf[] 进行遍历。在这个过程中, 可能会出现一个或多个出口。对于无条件跳转指令, 只会产生一个出口; 而对于条件跳转指令, 根据指令的条件来确定中间码的出口位置, 可能会产生 2 个出口。当处理 A 块的 goto_tb 指令时, 首先, 根据 tb->super_next[index] 来确定当前处理的翻译块与下一个翻译块 B 是否存在连接关系。如果 2 个翻译块有连接关系, 那么需要删除 2 条跳转指令, 并使用标志位 head_op 来记录当前操作的位置。在处理完 B 块对应的所有中间码后, 将 head_op 的值赋给 old_op。这样, B 块中间码装入 QTAIL 队列时的初始位置将变成 A 块原来 goto_tb 指令的位置。通过这种方式, 实现了超级块的嵌套。

本文的嵌套超级块结构设计相比于已有工作的优势:

1) 减少控制流转移。通过删除直接跳转指令并将基本块嵌套到一个更大的翻译块中, 可以减少程序中的控制流转移。每次直接跳转都可能导致性能损失, 因为需要额外的计算和内存访问来处理跳转。该操作将多个基本块合并成一个更大的块, 减少了控制流转移的次数, 从而提高了指令流水线的效率。

2) 改善代码布局。将多个翻译块合并成一个更大的翻译块, 使得代码布局更具优化性。传统的超级块可能在合并时受到限制, 而嵌套生成的更大翻译块可以更好地调整代码的顺序, 使得指令执行顺序更连贯, 从而提升整体性能。

3) 降低开销。由于减少了跳转和控制流转移的开销, 嵌套翻译块的方案能够降低总体的执行开销。每次跳转都可能引入延迟, 而通过合并成一个大的翻译块, 可以减少这些延迟, 从而提升程序的响应速度和执行效率。

图 5 展示了超级块嵌套结构。

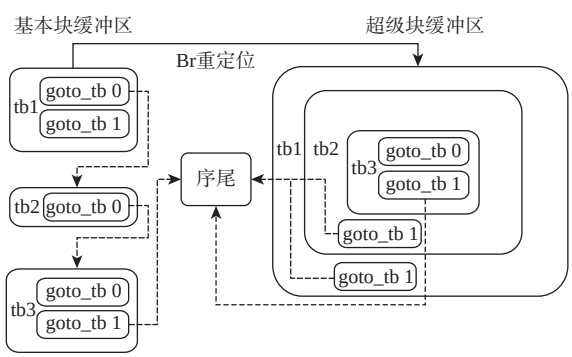


图 5　超级块嵌套结构

Fig. 5　Superblock nested architecture

3.3 嵌套超级块工作流程

如图6所示，为了构建超级块，本文在QEMU将源二进制代码翻译为中间代码TCG时，将所有的中间码收集到中间码缓冲池中。然后在进行翻译块的嵌套实现时，从中间码缓冲池中取到翻译块对应的中间码op，通过改变op的空间排布，从而将翻译块以嵌套的形式构建成超级块。

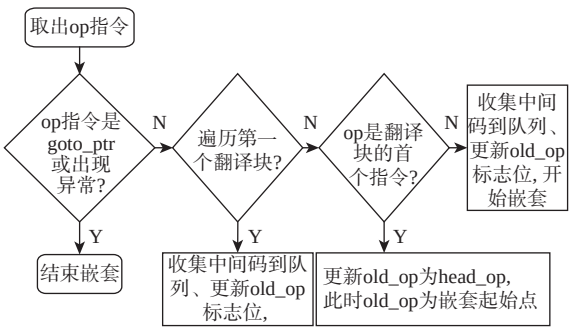


图6 超级块嵌套流程

Fig. 6 Superblock nesting flowchart

本文采用状态机数学模型来表示SN-QEMU方法中超级块嵌套处理方法的工作流。

1）状态定义

① Init：嵌套工作的初始化状态。

② Normal：当前处理指令不是goto_tb、goto_ptr、异常。

③ In_nest：当前处理指令是goto_tb指令，则进入执行嵌套状态，表示符合嵌套条件。

④ Out_nest：当前处理指令是goto_ptr指令或存在异常，则进入关闭嵌套状态，表示不符合嵌套条件。

⑤ Final：当热路径里的翻译块全部被处理完毕，则进入结束状态，表示超级块嵌套完毕。

2）条件定义

① To_nest：当读取到的指令是goto_tb指令，如果当前状态是[In_nest]或[Init]，将保持在[In_nest]；如果当前状态是[Normal]，则进入[In_nest]状态。

② To_normal：当读取到非goto_tb、goto_ptr、异常时，如果当前状态是[Init]或[Normal]，将保持在[Normal]状态；如果当前状态是[In_nest]，则进入[Normal]状态。

③ To_out_nest：当读取到goto_ptr、异常时，如果当前状态是[Init]、[Normal]或[In_nest]，则进入[Out_nest]状态。

④ To_final：当到达热路径流结束位置时，进入[Final]状态。

3）动作定义

① 状态[Init]\[Normal]->[In_nest]，表明当前翻译块嵌套下一个翻译块。

② 状态从[Init]\[In_nest]\[Normal]->[Out_nest]，表明不符合嵌套规则，结束嵌套。

③ 状态转移到[Final]，则表示热路径流已经读取完毕，嵌套工作结束。

4）状态机图

超级块嵌套状态机如图7所示。

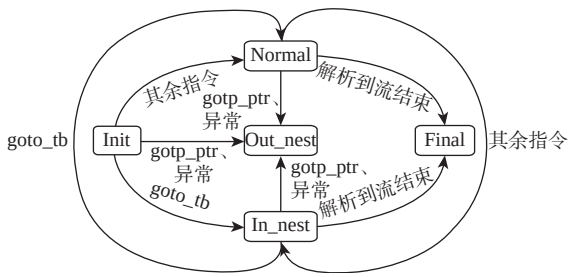


图7 超级块嵌套状态机

Fig. 7 Superblock nested state machine diagram

4 实验结果

4.1 实验环境

实验测试环境如下：实验的操作系统为Windows 11，QEMU版本为6.0，x86架构的测试平台采用的是AMD Ryzen 7 5800H，是一款八核心十六线程制程工艺7 nm的CPU。R7 5800H集成了全部8个基于Zen 3微体系结构的内核。其频率为3.2～4.4 GHz(Turbo)，3.2 GHz是保证的基本频率并支持SMT/超线程(16个线程)。

4.2 数据分析

本文测试程序选用SPEC2006程序集，所有程序均使用ref输入集进行10次实际的物理机器测试，然后取平均得到相关数据，并通过Linux性能分析工具——perf进行cache-misses的实际数据测量。

表1展示了嵌套超级块构造的相关统计数据，其中，第2、3列分别表示构造嵌套超级块所需要的基本块数量及最终生成的嵌套超级块数量，第4、5列分别表示测试程序插桩的翻译块Code Cache大小及嵌套超级块Code Cache大小，1个超级块至少

表1 SPEC2006超级块构造相关统计数据

Table 1 SPEC2006 Superblock construction related statistics

测试程序	基本块数量/个	超级块数量/个	基本块大小/KB	超级块大小/KB
410.bwaves	1 899	633	383	345
429.mcf	3 024	1 008	478	442
444.namd	10 962	3 654	5 043	4 640
456.hmmer	16 092	5 343	8 754	8 090
470.lbm	600	151	401	370
473.astar	1 400	423	656	604

需要包含 2 个翻译块。

通过引入 SN-QEMU 方法, 其所需 cache 大小相比于原版 QEMU 的翻译块呈现减小趋势, 使程序局部性得到改善。同时可以看到, 生成的嵌套超级块数量越多, 最终节省的 cache 空间就越大。

如图 8 所示, 在实现超级块嵌套翻译执行后, QEMU 6.0 在所测试的课题中展现出明显的 cache-misses 减少效果。其中, 平均减少幅度为 53.28%, 最高的减少幅度可达到 90%。

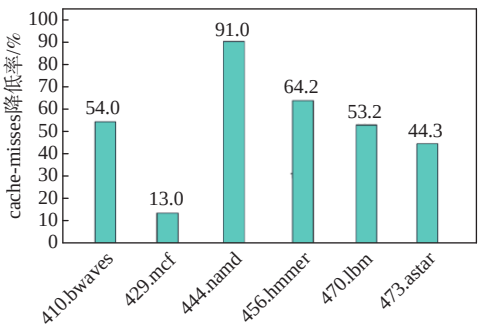


图 8　cache-misses 降低率

Fig. 8　cache-misses reduction rate

如图 9 所示, 展示了 QEMU 6.0 在实现超级块嵌套进行翻译执行后和传统构建超级块方法所测试课题的性能加速比。结果显示, 嵌套超级块平均性能加速比达到 3.07%, 最高加速比达到 4.58%。这表明, 相较于传统的超级块构建, 通过超级块嵌套的优化策略, QEMU 的翻译效率得到显著提高。该结果进一步证明, 超级块嵌套构造对于优化 QEMU 的性能具有积极的影响, 使得翻译执行过程更加高效和快速。

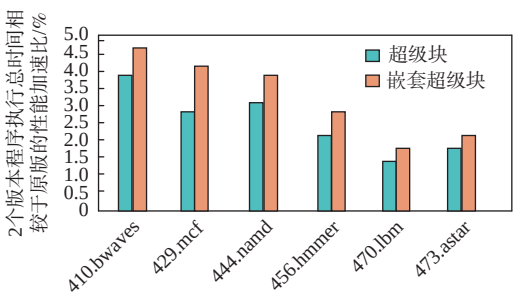


图 9　性能加速比

Fig. 9　Performance acceleration ratio

在实验过程中, 也对其他课题进行了性能加速测量, 但加速效果比较差, 如 401.bzip2 加速比为 0.42%、458.sjeng 加速比为 0.7%, 433.milc 加速比为 0.86%, 462.libquantum 加速比为 0.67% 等, 因此没在后续测试中展示。

程序的执行轨迹表面上看很复杂, 但往往会以某种嵌套的基本结构表现出来[29], 尤其是从高级结构性语言编译出来的程序更为明显。这些基本结构大致可以概括成循环和函数。

在程序执行过程中, 热路径是指频繁执行的动态指令序列, 其反映了程序动态执行的特征轨迹, 如频繁执行的循环或者函数调用。图 10 展示了所测试课题的热路径在整个程序中的占比情况。其中, 热路径占比最高达到 17.1%, 最低为 3.2%。

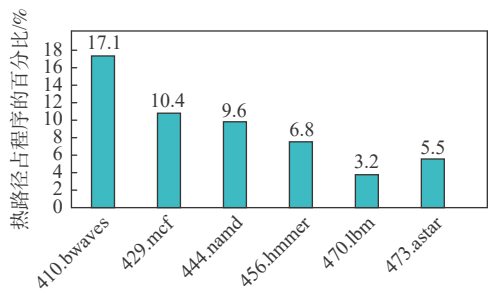


图 10　热路径占程序的百分比

Fig. 10　Percentage of hot routes to programmes

接下来, 探究 410.bwaves 课题比 470.lbm 课题的加速比提升效果更高的原因。通过 gprof 测试工具, 对 2 个课题的执行相关数据进行统计, 表 2 和表 3 分别展示了 2 个课题执行的相关统计数据, 其中第 2 列展示了函数的调用次数, 第 3 列展示了函数执行的总时间, 第 4 列展示了程序运行过程中, 函数单次执行时的循环次数。

表 2　410.bwaves 课题相关统计数据

Table 2　410.bwaves subject-related statistics

调用函数名称	调用次数/次	运行时间/s	函数单次循环次数/次
mat_times_vec_	604	21.62	3 380 000
bi_cgstab_block_	20	3.88	0
shell_	1	2.8	0
jacobian_	60	0.77	0
flux	20	0.34	0
MAIN_	1	0.01	0
main	1	0.01	2 041 520 000

通过实验数据可以发现, 410.bwaves 课题的函数调用总共为 707 次, 总循环次数为 2 041 520 000 次。470.lbm 课题函数调用总共为 919 次, 总循环次数为 5 200 000 次。410.bwaves 课题的函数调用虽然没有 470.lbm 课题多, 但是循环次数远高于 470.lbm 课题, 并且产生循环的函数是需要被多次调用的。

因此, 当测试课题中循环或函数调用占主导地位时, 超级块中间码嵌套策略的效果会更加显著。这也是测试课题中 410.bwaves 的提升效果比 470.lbm 更高的原因。

通过对程序进行超级块嵌套构造, 可以更好地

表3 470.lbm课题相关统计数据

Table 3 470.lbm subject-related statistics

调用函数名称	调用次数/次	运行时间/s	函数单次循环次数/次
LBM_performStreamCollide	300	11.70	0
LBM_initializeGrid	2	1.70	0
LBM_handleInOutFlow_	300	0.23	0
LBM_showGridStatistics_	6	0.10	0
LBM_initializeSpecialCellsForChannel	2	0.01	1 300 000
LBM_swapGrids	300	0.01	0
LBM_allocateGrid	2	0.01	0
LBM_freeGrid	2	0.01	0
LBM_loadObstacleFile	2	0.01	1 300 000
MAIN_finalize	1	0.01	0
MAIN_initialize	1	0.01	0
MAIN_parseCommandLine	1	0.01	0
MAIN_printInfo	1	0.01	0
main	1	0.01	5 200 000

利用循环和函数的结构特点,从而优化执行效率。这进一步说明,超级块嵌套策略在处理密集循环和长路径类型应用时的重要性和有效性。

5 结 论

1)本文提出了一种基于QEMU的动态二进制翻译超级块嵌套方法,即SN-QEMU。该策略能够使得翻译块在cache层面上连续分布,从而使cache-misses平均减少幅度为53.28%,最高的减少幅度可达到90%。

2)通过对跳转指令的删除,减小了构建的超级块占用空间大小。

3)该策略通过增加cache命中率,使得代码空间局部性得到改善,从而提高系统执行性能,平均性能加速比达到了3.07%,最高加速比达到了4.58%。

4)本文构建嵌套超级块目前仅考虑直接跳转的情况,后期将尝试把以间接跳转指令结尾的翻译块加入嵌套超级块中。未来工作中将尝试使用SPECCPU 2017版本进行数据实验。

参考文献(References)

[1] BELLARD F. QEMU, a fast and portable dynamic translator[C]//Proceedings of the 2005 USENIX Annual Technical Conference. Anaheim: USENIX, 2005: 41-46.

[2] 胡伟武, 汪文祥, 吴瑞阳, 等. 龙芯指令系统架构技术[J]. 计算机研究与发展, 2023, 60(1): 2-16.
HU W W, WANG W X, WU R Y, et al. Loongson instruction set architecture technology[J]. Journal of Computer Research and Development, 2023, 60(1): 2-16(in Chinese).

[3] SHEN B Y, YOU J Y, YANG W, et al. An LLVM-based hybrid binary translation system[C]//Proceedings of the 7th IEEE International Symposium on Industrial Embedded Systems. Piscataway: IEEE Press, 2012: 229-236.

[4] EBCIOGLU K, ALTMAN E, GSCHWIND M, et al. Dynamic binary translation and optimization[J]. IEEE Transactions on Computers, 2001, 50(6): 529-548.

[5] HWU W M W, MAHLKE S A, CHEN W Y, et al. The superblock: an effective technique for VLIW and superscalar compilation[C]//Proceedings of the Instruction-level Parallelism. Boston: Springer, 2011: 229-248.

[6] SCOTT K, KUMAR N, CHILDERS B R, et al. Overhead reduction techniques for software dynamic translation[C]//Proceedings of the 18th International Parallel and Distributed Processing Symposium. Piscataway: IEEE Press, 2004: 200-207.

[7] 石强. 面向国产处理器的二进制翻译关键优化技术研究[D]. 郑州: 解放军信息工程大学, 2017: 19-34.
SHI Q. Research on key optimization techniques of binary translation for domestic processors[D]. Zhengzhou: PLA Information Engineering University, 2017: 19-34(in Chinese).

[8] 李男, 庞建民. 基于中间表示规则替换的二进制翻译中间代码优化方法[J]. 国防科技大学学报, 2021, 43(4): 156-162.
LI N, PANG J M. Intermediate code optimization method for binary translation based on intermediate representation rule replacement[J]. Journal of National University of Defense Technology, 2021, 43(4): 156-162(in Chinese).

[9] HONG D Y, HSU C C, YEW P C, et al. HQEMU: a multi-threaded and retargetable dynamic binary translator on multicores[C]//Proceedings of the Tenth International Symposium on Code Generation and Optimization. New York: ACM, 2012: 104-113.

[10] POEPLAU S, FRANCILLON A. SymQEMU: compilation-based symbolic execution for binaries[C]//Proceedings of the 2021 Network and Distributed System Security Symposium. Internet Society. San Diego: DNSS, 2021: 1-13.

[11] DING J H, CHANG P C, HSU W C, et al. PQEMU: a parallel system emulator based on QEMU[C]//Proceedings of the 2011 IEEE 17th International Conference on Parallel and Distributed Systems. Piscataway: IEEE Press, 2012: 276-283.

[12] WANG Z G, LIU R, CHEN Y F, et al. COREMU: a scalable and portable parallel full-system emulator[C]//Proceedings of the 16th ACM Symposium on Principles and Practice of Parallel Programming. New York: ACM, 2011: 213-222.

[13] ZHAO Z Y, ZHANG J, LIU X M, et al. DQEMU: a scalable emulator with retargetable DBT on distributed platforms[C]// Proceedings of the 49th International Conference on Parallel Processing, Edmonton. New York: ACM, 2020: 1-11.

[14] ZENG J Y, FU Y C, LIN Z Q. PEMU: a pin highly compatible out-of-VM dynamic binary instrumentation framework[C]//Proceedings of the 11th ACM SIGPLAN/SIGOPS International Conference on Virtual Execution Environments. New York: ACM, 2015: 147-160.

[15] LI M L, PANG J M, YUE F, et al. Enhancing dynamic binary translation in mobile computing by leveraging polyhedral optimization[J]. Wireless Communications and Mobile Computing, 2021, 2021(1): 6611867.

[16] DOLAN-GAVITT B, HODOSH J, HULIN P, et al. Repeatable re-

verse engineering with PANDA[C]//Proceedings of the 5th Program Protection and Reverse Engineering Workshop. New York: ACM, 2015: 1-11.

[17] HWU W W, MAHLKE S A, CHEN W Y, et al. The superblock: an effective technique for VLIW and superscalar compilation[J]. Journal of Supercomputing, 1993, 7(1-2): 229-248.

[18] CHILIMBI T M, DAVIDSON B, LARUS J R. Cache-conscious structure definition[C]//Proceedings of the ACM SIGPLAN 1999 Conference on Programming Language Design and Implementation. New York: ACM, 1999: 13-24.

[19] CALDER B, KRINTZ C, JOHN S, et al. Cache-conscious data placement[C]//Proceedings of the Eighth International Conference on Architectural Support for Programming Languages and Operating Systems. New York: ACM, 1998: 139-149.

[20] BAER J L, WANG W H. On the inclusion properties for multi-level cache hierarchies[C]//Proceedings of the The 15th Annual International Symposium on Computer Architecture . Conference Proceedings. Piscataway: IEEE Press, 2002: 73-80.

[21] SHI H H, WANG Y, GUAN H B, et al. An intermediate language level optimization framework for dynamic binary translation[J]. ACM SIGPLAN Notices: A Monthly Publication of the Special Interest Group on Programming Languages, 2007, 42(5): 3-9.

[22] BALL T, MATAGA P, SAGIV M. Edge profiling versus path profiling: The showdown[C]//Proceedings of the 25th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages - POPL '98. New York: ACM, 1998: 134-148.

[23] BALA V, DUESTERWALD E, BANERJIA S. Transparent dynamic optimization: the design and implementation of Dynamo[R]. Hewlett-Packard Labs, 1999: HPL-1999-78.

[24] DUESTERWALD E, BALA V. Software profiling for hot path prediction: Less is more[J]. ACM SIGARCH Computer Architecture News, 2000, 35(11): 202-211.

[25] CIFUENTES C, LEWIS B, UNG D. Walkabout: a retargetable dynamic binary translation framework[J]. ACM SIGPLAN Notices, 2002, 37(7): 4-21.

[26] BRUENING D, DUESTERWALD E, AMARASINGHE S. Design and implementation of a dynamic optimization framework for Windows[C]//Proceedings of the 4th ACM Workshop on Feedback-directed and Dynamic Optimization. New York: ACM Press, 2001: 20.

[27] CHEN W K, LERNER S, CHAIKEN R, et al. Mojo: a dynamic optimization system[C]//Proceedings of the 3rd ACM Workshop on Feedback-Directed and Dynamic Optimization. New York: ACM Press, 2000: 81-90.

[28] 张世宜. 基于 QEMU 的热点代码探测与动态优化模型的研究与实现[D]. 成都: 电子科技大学, 2013: 37-39.
ZHANG S Y. University of electronic sci ence and technology of China[D]. Chengdu: University of Electronic Science and Technology of China, 2013: 37-39(in Chinese).

[29] 白童心. 动态二进制翻译与动态优化相关问题研究[D]. 北京: 中国科学院研究生院 (计算技术研究所), 2004: 33-34.
BAI T X. Two topics on dynamic binary translation and optimization[D]. Beijing: Institute of Computing Technology, Chinese Academy of Sciences, 2004: 33-34(in Chinese).

Superblock nesting based on performance optimization for dynamic binary translation

JIN Shuai，CHAI Zhilei*，ZHOU Haojie

(School of Artificial Intelligence and Computer Science，Jiangnan University，Wuxi 214122，China)

Abstract：Simple direct block linking only considers the connection at the logical level and ignores the continuity of the physical storage location, despite the fact that the "direct block linking" method can decrease the frequency of translator intervention and enhance the target program's performance during the dynamic binary translation process. Aiming at this problem, a dynamic binary translation performance optimization method based on superblock nesting is proposed, which is based on the IR instructions generated by the binary translator pre-translating the target code, and increases the instruction cache hit rate by constructing superblocks with a nested structure so that the code achieves the continuity of physical storage locations. The experimental results show that: SPEC2006, the most classic and commonly used in the field of dynamic binary translation, is used as the test benchmark, and QEMU6.0, a mainstream open-source binary translator, is used as the experimental framework. Compared with the version without superblock nesting, the proposed method can improve the cache hit rate by 53.28% on average, and the maximum increase can be 90%. The running performance of the target program can be improved by 3.07% on average and 4.58% at maximum.

Keywords：dynamic binary translation；superblock；intermediate code nesting；QEMU；cache

Received：2024-05-08；**Accepted**：2024-09-23；**Published Online**：2024-10-08 13：18

URL：link.cnki.net/urlid/11.2625.V.20241008.0955.001

Foundation item：National Natural Science Foundation of China (61972180)

* **Corresponding author**. E-mail：zlchai@jiangnan.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0288

考虑构型的旋转碟状飞行器水漂动力学特性

王迪[1,2]，罗剑桥[1,2]，于宁[1,2]，孟军辉[1,2,*]

(1. 北京理工大学 空天科学与技术学院，北京 100081；2. 北京理工大学重庆创新中心，重庆 401120)

摘 要：旋转碟状飞行器可通过模仿石块水漂有效提升海上拒止环境突防能力，保障连续水漂和降低冲击过载是其水漂动力学设计的关键。水漂动力学特性主要取决于初始运动参数，而构型参数也是影响水漂过程位姿运动和载荷分布的重要因素。针对旋转碟状飞行器水漂动力学特性影响参数耦合、作用机制不清等问题，本文提出一种旋转碟状飞行器的参数化构型，并基于任意拉格朗日-欧拉(ALE)单元法和罚函数法完成其水漂仿真分析，分别得到构型参数和初始运动参数对其水漂动力学特性的影响规律，并探究2类参数的耦合关系。结果表明，增大边缘曲率半径可有效减小水漂过程中飞行器所受过载，而增大自旋角速度可大大提高飞行器姿态稳定性。所提研究可为旋转碟状飞行器等新型水漂式反舰设计提供参考。

关 键 词：旋转碟状飞行器；水漂动力学；运动参数；构型参数；任意拉格朗日-欧拉单元法

中图分类号：TJ760.2

文献标志码：A **文章编号**：1001-5965(2026)06-2133-12

将石块以一定速度和角度投向水面能够使其在水面上反复触水弹跳，产生波浪线形的运动轨迹，这种在近水面附近滑跳的运动方式即为水漂。由日常水漂游戏的经验可知，投掷石块时给予其初始旋转运动有利于稳定其姿态，从而增加水漂运动的滑跳次数和飞行距离[1-3]。由于地球表面存在曲率，近海面区域通常是敌方雷达的探测盲区，利用水漂实现超低空掠海飞行能够有效提高拒止环境下反舰攻击突防能力[4-5]。参考石块水漂现象设计新型旋转碟状飞行器，有望通过高速自旋大幅提高水漂性能，为下一代反舰发展提供新思路。

为了实现旋转碟状飞行器连续多次安全水漂，合适的构型参数和初始运动参数至关重要。一方面，飞行器初始速度、触水角度等运动参数对于能否成功滑跳起主导作用[6-8]。若初始运动状态不合适，会使飞行器触水后直接沉入水中，或滑跳离水后不能在下次触水时保持合适的运动状态。另一方面，当飞行器构型改变时，其沾湿区域的演化历程也将发生变化，从而影响水动载荷分布和冲击过载峰值[9]。研究初始运动和构型参数对飞行器水漂动力学特性的影响规律，为保障连续水漂和降低冲击过载提供参考依据，对于旋转碟状飞行器的设计与研制意义重大。

旋转碟状飞行器水漂包含固、液、气3相耦合作用，涉及流体动力学、陀螺力学等多门学科，具有高度的非线性和非定常性[10-11]。固体以动边界形式作用于流场，产生拍击、喷溅现象，流场对固体施加反作用力，引发进动、章动等复杂动力学行为。同时，受包括速度、姿态角等初始状态与厚度、曲率等构型特征在内的多种因素共同影响，旋转碟状飞行器水漂过程影响参数众多，耦合关系复杂[12-14]，为其水漂动力学问题的研究带来挑战。

近年来，学者对运动体的水漂现象展开了研究。赵坤[15]通过理论方法研究了水漂体弹跳动力学特性的影响因素及连续弹跳机理。Clanet等[3]进行了旋转圆盘的水漂实验研究，发现当攻角为

收稿日期：2024-05-07；录用日期：2024-06-21；网络出版时间：2024-08-19 15:13
网络出版地址：link.cnki.net/urlid/11.2625.V.20240816.1916.003
基金项目：国家自然科学基金(52372346)；北京理工大学创新人才科技资助项目(2021CX01018)
*通信作者. E-mail：mengjh@bit.edu.cn

引用格式：王迪，罗剑桥，于宁，等. 考虑构型的旋转碟状飞行器水漂动力学特性[J]. 北京航空航天大学学报，2026，52(6)：2133-2144.
WANG D，LUO J Q，YU N，et al. Dynamic characteristics of rotating disk-shaped flight vehicle during skipping considering configuration[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：2133-2144 (in Chinese).

20°时，圆盘能发生更多次弹跳。华厦等[16]通过光滑粒子流体动力学方法和任意拉格朗日-欧拉(arbitrary Lagrange-Euler, ALE)方法分析了水漂运动的临界入水角和自旋角速度对滑跳距离的影响。孙士明等[17]基于计算流体力学(computational fluid dynamics, CFD)方法对3种典型的圆盘水漂运动状态进行了数值仿真，指出圆盘运动方向会受自旋影响而产生一定偏转。现有对水漂问题的研究表明，飞行器初始运动状态对于能否成功水漂起主导作用，近似于圆盘的碟状飞行器较球体和圆柱体更易发生水漂，同时，较小的俯仰角和较大的速度有利于水漂现象的发生，飞行器旋转可增强其水漂过程中姿态的稳定性。但现有研究较少分析运动体的具体构型对水漂动力学特性的影响。

现有研究的实践经验表明，数值仿真较理论推导等方法能更加精确地得到各参数变化对飞行器水漂动力学特性的影响[18-21]。ALE方法兼具Lagrange方法和Euler方法两者的优点：首先，在结构边界运动的处理上，引进了Lagrange方法的特点，因此，能够有效跟踪物质结构边界的运动；其次，在内部网格的划分上，吸收了Euler法的长处，使内部网格单元独立于物质实体而存在，但允许根据定义的参数在求解过程中适当调整位置，有利于分析大变形问题。因此，在利用数值仿真模拟旋转碟状飞行器水漂时采用ALE方法可以得到较为精确的结果[22-25]。

本文针对旋转碟状飞行器水漂机理复杂、涉及参数较多的问题，采用ALE方法建立飞行器水漂数值仿真模型，研究构型参数和初始运动参数对飞行器水漂动力学特性的影响规律。对飞行器构型进行参数化建模，从而得到3个构型参数，以研究构型对水漂动力学特性的影响。分别改变不同构型参数及初始运动参数，开展飞行器水漂数值仿真，分析仿真结果得到了不同参数对水漂动力学特性的影响规律，可为旋转碟状飞行器初步设计时构型和初始运动参数设置提供依据。

1 数值计算方法

1.1 流场控制方程

ALE描述的流场控制方程具体如下。

1）质量守恒方程：

$$\frac{\partial \rho}{\partial t}=-\rho\frac{\partial v_i}{\partial x_i}-w_i\frac{\partial \rho}{\partial x_i} \tag{1}$$

式中：ρ为质量密度；v_i为流体速度；x_i为欧拉坐标。

2）动量守恒方程：

$$\rho\frac{\partial v_i}{\partial t}=\frac{\sigma_{ij}}{\partial_{xj}}+\rho b_i-\rho w_i\frac{\partial v_i}{\partial x_j} \tag{2}$$

式中：b_i为单位体积流体受到的体力；张量分量$\sigma_{ij}=-p\delta_{ij}+\mu(v_{i,j}+v_{j,i})$，$\delta_{ij}$为单位张量分量，$\mu$为流体动力黏度，$p$为流体静压；$w_i$为相对速度，为流体速度$v_i$与网格速度$u_i$之差。

3）能量守恒方程：

$$\rho\frac{\partial E}{\partial t}=\sigma_{ij}\frac{\partial v_i}{\partial x_j}-\frac{\partial q_i}{\partial x_j}-\rho w_j\frac{\partial E}{\partial x_j} \tag{3}$$

式中：E为单位质量的内能；q_i为热通量。

4）状态方程。本文流体采用Gruneisen状态方程：

$$p=\frac{\rho_0C^2\varepsilon\left(1+(1-\gamma_0/2)\varepsilon-A\varepsilon^2/2\right)}{\left(1-(S_1-1)\varepsilon-S_2\varepsilon^2/(\varepsilon+1)-S_3\varepsilon^3/(\varepsilon+1)^2\right)^2}+(\gamma_0+A\varepsilon)w \tag{4}$$

式中：ρ_0为流体初始密度；ε为压缩度，即$\varepsilon=\rho/\rho_0-1$；$\gamma_0$为Gruneisen系数；$A$为Gruneisen系数$\gamma_0$的一阶体积修正；$w$为单位初始体积内能；$S_1$、$S_2$、$S_3$为$u_s$-$u_p$曲线的斜率系数；$C$为$u_s$-$u_p$曲线的截距。

u_s-u_p曲线即

$$u_s=C+S_1u_p+S_2\frac{u_p^2}{u_s}+S_3\frac{u_p^3}{u_s^2} \tag{5}$$

式中：u_s为冲击波波速；u_p为冲击波后质点速度。

空气采用Linear Polynomial状态方程：

$$p=c_0+c_1\varepsilon+c_2\varepsilon^2+c_3\varepsilon^3+\left(c_4+c_5\varepsilon+c_6\varepsilon^2\right)w \tag{6}$$

式中：c_0～c_6为材料参数。

假设空气为理想气体，则水和空气状态方程中各个参数的取值如表1所示。

1.2 刚体运动方程

为简化计算，在数值仿真中将飞行器视为刚体。根据牛顿第二定律，刚体质心的运动方程可以写成

$$\sum F_i=ma_i \qquad i=1,2,3 \tag{7}$$

式中：F_1、F_2、F_3分别为刚体在x轴、y轴、z轴方向上所受的合外力；m为刚体质量；a_1、a_2、a_3分别为刚体质心在x轴、y轴、z轴方向上的加速度。

表1 状态方程参数取值

Table 1 Values of parameters in state equation

$\rho_{0,w}/(\mathrm{kg\cdot m^{-3}})$	$C_w/(\mathrm{m\cdot s^{-1}})$	γ_{0w}	A_w	S_{1w}	S_{2w}	S_{3w}	$w_w/(\mathrm{J\cdot m^{-3}})$	$\rho_{0,a}/(\mathrm{kg\cdot m^{-3}})$	c_{1a}	c_{2a}	c_{3a}	c_{4a}	c_{5a}	c_{6a}	$w_a/(\mathrm{J\cdot m^{-3}})$
1 000	1 650	0.5	0.35	1.92	−0.096	0	1.9×10^5	1.225	0	0	0	0.4	0.4	0	2.1×10^5

对于刚体绕质心的转动运动, 有

$$\begin{cases} I_{11}\dot{\omega}_1 - I_{12}\dot{\omega}_2 - I_{13}\dot{\omega}_3 = M_1 + I_{13}\omega_1\omega_2 - \\ \quad I_{12}\omega_1\omega_3 + I_{23}(\omega_2^2 - \omega_3^2) + (I_{22} - I_{33})\omega_2\omega_3 \\ -I_{21}\dot{\omega}_1 + I_{22}\dot{\omega}_2 - I_{23}\dot{\omega}_3 = M_2 + I_{12}\omega_2\omega_3 - \\ \quad I_{23}\omega_2\omega_1 + I_{13}(\omega_3^2 - \omega_1^2) + (I_{33} - I_{11})\omega_1\omega_3 \\ -I_{31}\dot{\omega}_1 - I_{32}\dot{\omega}_2 + I_{33}\dot{\omega}_3 = M_3 + I_{23}\omega_3\omega_1 - \\ \quad I_{13}\omega_3\omega_2 + I_{12}(\omega_1^2 - \omega_2^2) + (I_{11} - I_{22})\omega_1\omega_2 \end{cases} \tag{8}$$

式中: I_{ij}为转动惯量; M_i为合外力矩; ω_i为角速度; $\dot{\omega}_i = \mathrm{d}\omega_i/\mathrm{d}t$, $\dot{\omega}_t$为角加速度。

考虑到本体坐标系中惯性积为 0, 式 (8) 可简化为

$$\begin{cases} I_{11}\dot{\omega}_1 = M_1 + (I_{22} - I_{33})\omega_2\omega_3 \\ I_{22}\dot{\omega}_2 = M_2 + (I_{33} - I_{11})\omega_1\omega_3 \\ I_{33}\dot{\omega}_3 = M_3 + (I_{11} - I_{22})\omega_1\omega_2 \end{cases} \tag{9}$$

在求解动力学方程时, 首先计算出刚体在某一时刻所受外力及力矩, 利用动力学方程计算出刚体此刻的加速度和角加速度, 通过显式时间推进得到下一时刻刚体的速度、角速度及姿态、位置等。

1.3　耦合算法

采用欧拉-拉格朗日耦合算法, 使流体与结构之间的边界能互相重叠, 在计算时添加约束条件, 从而使两者相互耦合, 实现动力的传递。

约束条件利用罚函数耦合的原理, 界面力根据结构与流体之间的重叠深度确定。若结构与欧拉流体未出现重叠现象, 则界面力为零; 若出现重叠, 则根据结构与流体之间的重叠深度d计算界面力F的大小。

$$F = k_i d \tag{10}$$

式中: k_i为节点质量模型的刚度系数。

采用有限体积法(finite volume method, FVM)求解 ALE 控制方程。对于流体单元, 对式 (2) 积分后, 耦合力将以源项的形式加到方程右侧。对于刚体, 耦合力以外力的形式应用于运动方程。

2　旋转碟状飞行器水漂数值模型

2.1　旋转碟状飞行器参数化构型

旋转碟状飞行器以 Pye Wacket 空空导弹为原型[26], 其外形为飞碟状的圆盘, 后端为推进装置, 通过前缘偏转在飞行中操纵导弹, 对旋转碟状飞行器外形参数进行参数化处理, 外形建模如图 1 所示。

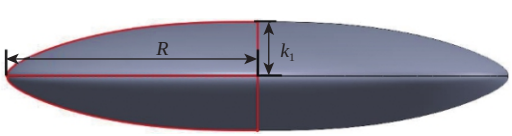


图 1　飞行器构型参数化建模

Fig. 1　Parametric modeling of aircraft configuration

飞行器半径R设置为 7 m, 截面曲线表达式采用常见的双立方多项式[27]:

$$y = k_1\left[k_2\frac{6x}{R}\left(\frac{x}{R}-1\right)^4 + \frac{k_3}{2}\left(\frac{x}{R}\right)^2 \cdot \left(\frac{x}{R}-1\right)^3 + 1 - \left(\frac{x}{R}-1\right)^4\left(\frac{4x}{R}+1\right)\right]^{\frac{1}{3}} \tag{11}$$

式中: k_1为厚度的一半; $k_2 \in [0,1]$为飞行器边缘的曲率; $k_3 \in [0,15]$为飞行器靠近中部的构型曲线曲率, 如图 2 所示。

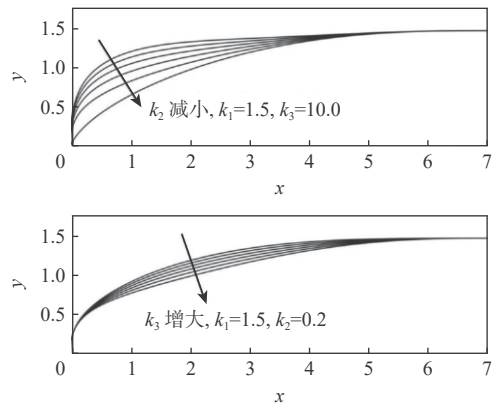


图 2　构型参数k_2、k_3的影响

Fig. 2　The impact of configuration parameters of k_2 and k_3

可以看出, 随着k_2减小, 飞行器边缘构型曲线曲率逐渐减小, 随着k_3增大, 飞行器靠近中部的构型曲线曲率逐渐减小。

通过控制构型参数k_1、k_2、k_3, 可以得到不同的飞行器外形, 为研究不同构型对于水漂动力学特性的影响提供基础。

参考目前有关设备设计飞行器初始几何参数, 如表 2 所示 I_x、I_y、I_z 为转动惯量, k_1、k_2、k_3 为构型参数。

表 2　几何模型参数

Table 2　Geometric model parameters

质量/kg	半径/m	I_x/(kg·m^{-2})	I_y/(kg·m^{-2})	I_z/(kg·m^{-2})	k_1	k_2	k_3
14 550.0	7.0	22 837.9	45 462.1	22 800.0	1.5	0.2	10

2.2　旋转碟状飞行器水漂仿真设置

由于本文研究的飞行器高速自旋, 在飞行器水漂过程中, 其质心的运动速度和绕质心转动的角速度会产生较为剧烈的变化, 因此, 采用惯性坐标系$O_0x_0y_0z_0$、平移坐标系$Ox_1y_1z_1$、连体坐标系$Ox_2y_2z_2$能更好地对其运动状态进行描述。

如图 3 所示, 惯性坐标系$O_0x_0y_0z_0$与地面固连, x_0轴与飞行器初始速度在水平面上的投影平行, y_0轴竖直向上, z_0轴依据右手定则确定。

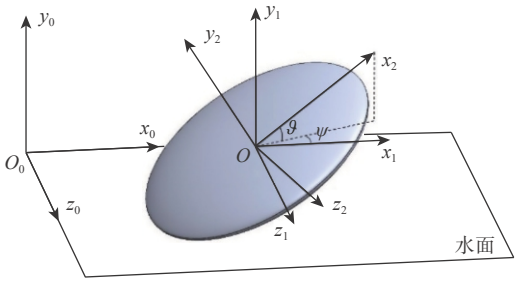


图3 飞行器坐标系示意

Fig. 3 Coordinate system diagram of the flight vehicle

平移坐标系$Ox_1y_1z_1$原点O在飞行器质心上，各轴与惯性坐标系的坐标轴平行。

连体坐标系$Ox_2y_2z_2$由平移坐标系$Ox_1y_1z_1$旋转得到：首先，绕Oy_1轴逆时针旋转偏航角ψ，使Oz_1轴到达Oz_2轴，此时原Ox_1变为Ox_1'。然后，绕Oz_2轴逆时针旋转俯仰角ϑ，使Oy_1轴到达Oy_2轴，Ox_1'轴到达Ox_2轴。此时，Ox_2轴垂直于飞行器与水面的交线，Oy_2轴为飞行器对称轴，Oz_2轴与飞行器与水面的交线平行。连体坐标系$Ox_2y_2z_2$也称为莱查坐标系，由于连体坐标系不随飞行器自转而旋转，因此，可以简化计算。

根据平移坐标系$Ox_1y_1z_1$与连体坐标系$Ox_2y_2z_2$之间的旋转关系，可得$Ox_2y_2z_2$相对于$Ox_1y_1z_1$的坐标旋转矩阵为

$$\boldsymbol{L}=\begin{bmatrix}\cos\vartheta & \sin\vartheta & 0\\ -\sin\vartheta & \cos\vartheta & 0\\ 0 & 0 & 1\end{bmatrix}\cdot\begin{bmatrix}\cos\psi & 0 & -\sin\psi\\ 0 & 1 & 0\\ \sin\psi & 0 & \cos\psi\end{bmatrix}=\begin{bmatrix}\cos\vartheta\cos\psi & \sin\vartheta & -\cos\vartheta\sin\psi\\ -\sin\vartheta\cos\psi & \cos\vartheta & \sin\vartheta\sin\psi\\ \sin\psi & 0 & \cos\psi\end{bmatrix} \tag{12}$$

定义速度倾角φ为速度矢量与水平面的夹角，速度矢量向上为正，则初始运动参数如表3所示。

表3 初始运动参数

Table 3 Initial motion parameters

速度/(m·s^{-1})	速度倾角/(°)	俯仰角/(°)	偏航角/(°)	自旋角速度/(rad·s^{-1})
100.0	−20.0	15.0	0	400.0

设置数值仿真计算域如图4所示。计算域由空气域和流体域构成，计算域长度$L=84$ m，宽度$w=30$ m，空气域高度$h_a=20$ m，流体域高度$h_w=20$ m。惯性坐标系$O_0x_0y_0z_0$原点O_0设置在水面左上角，各坐标轴指向均沿空气域边界。为保证飞行器在发生水漂时运动状态及液面波动清晰可见，且在发生水漂后的一段时间内不会飞出计算域，设置飞行器质心在惯性坐标系$O_0x_0y_0z_0$下坐标为(28 m,4 m,15 m)。为节约计算时间，将流体域设

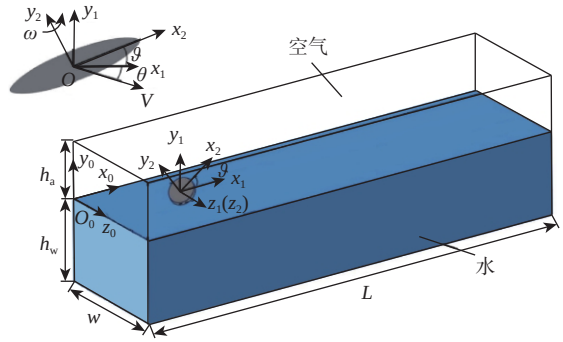


图4 数值仿真计算域设置情况

Fig. 4 Settings of simulation calculation domain

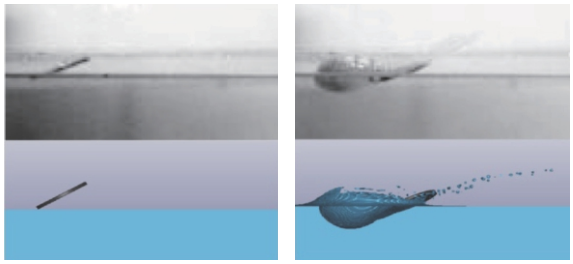
置为长方体网格，对旋转碟状飞行器划分六面体-三棱柱混合网格。保持流体与飞行器互相接触部分网格尺寸相等，以提高计算精度。

同时，为提高计算效率，保证计算准确度，对流体域网格划分进行变尺寸处理。由于飞行器与水面的碰撞发生在计算域的中间位置，且碰撞时间较短，到仿真终止时间，水面的波浪仍未到达计算域边界。将计算域靠近边界的网格设置的较粗能有效缩短仿真时间，同时保证碟状飞行器运动状态及水面波动情况的计算精度。因此，计算域的高度方面在上下两侧采用粗网格，宽度方面在左右两侧采用粗网格，而长度方面，由于飞行器与水面的碰撞位置和在水面滑行的距离可能因不同的初始条件而发生变化，因此，流体各网格长度一致。

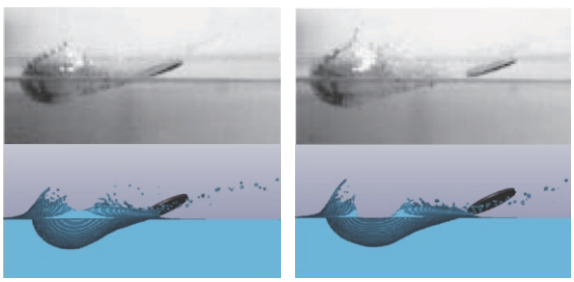
2.3 模型正确性验证

参考Rosellini等[2]所做石子水漂实验，对旋转碟状飞行器水漂模型进行正确性验证。石子刚体为半径为25 mm、高为2.75 mm的扁平圆柱体，密度为2.7×10^3 kg/m^3。计算域大小为300 mm×104 mm×140 mm，其中水域高70 mm，空气域高70 mm。与2.2节相同，对流体划分六面体网格，并作变尺寸处理，对刚体划分六面体-三棱柱混合网格，最小网格尺寸为1.4 mm×1.4 mm×1.4 mm。设置石子初始速度为3.5 m/s，自旋角速度为65 rad/s，俯仰角为20°，速度倾角为−20°。仿真与实验得到的结果对比如图5所示。

可以看出，实验中石子的姿态和运动状态与数值仿真所得到的结果吻合良好。同时，液面状态也基本相同，但由于网格尺寸限制，石子水漂时激起的水雾存在一定误差。由于本文主要研究刚体的动力学特性，可以认为所建立的数值仿真模型具有一定的稳定性和可靠性，能够为后续研究提供支撑。



(a) t=0 s　(b) t=0.013 s



(c) t=0.026 s　(d) t=0.039 s

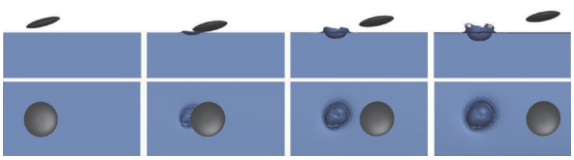
图 5　石子水漂仿真与实验结果对比

Fig. 5　Comparison of simulation and experimental results of stone skipping

3　仿真结果及分析

3.1　水漂机理分析

根据 2.2 节给定的初始运动参数及构型参数进行旋转碟状飞行器的水漂数值仿真，得到其运动状态如图 6 所示。



(a) t=0 s　(b) t=0.1 s　(c) t=0.2 s　(d) t=0.3 s

图 6　旋转碟状飞行器水漂运动状态

Fig. 6　The skimming motion state of rotating disk-shaped flight vehicle

飞行器的一次水漂运动可分为触水前、触水时、触水后 3 个状态。飞行器运动参数和姿态角随时间变化曲线如图 7 和图 8 所示。飞行器触水前，在重力作用下做斜抛运动向水面飞去，运动状态较为稳定，且姿态保持不变，即俯仰角ϑ、偏航角ψ为定值。

飞行器触水时，流体受到挤压而向四周运动，同时，给予飞行器向上的反作用力，产生向上的加速度。随着碰撞程度的加剧，飞行器所受到的反作用力和向上的加速度也逐渐增大，竖直方向速度V_{y_0}逐渐减小直至反向，此时，飞行器开始向离开水面的方向飞去。随着碰撞的继续进行，V_{y_0}逐渐增大，最终使飞行器脱离水面。飞行器在与流体的碰撞过程中，由于尾部最先触碰水面而受到反作用

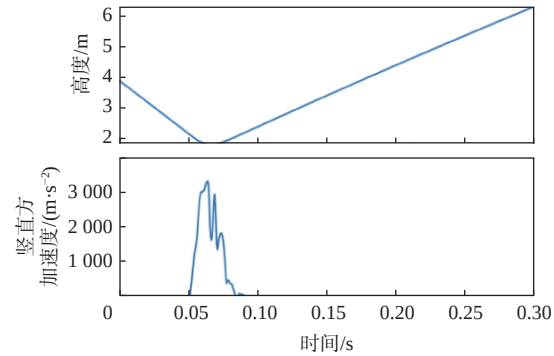


图 7　旋转碟状飞行器运动参数曲线

Fig. 7　Curves of motion parameters for rotating disk-shaped flight vehicle

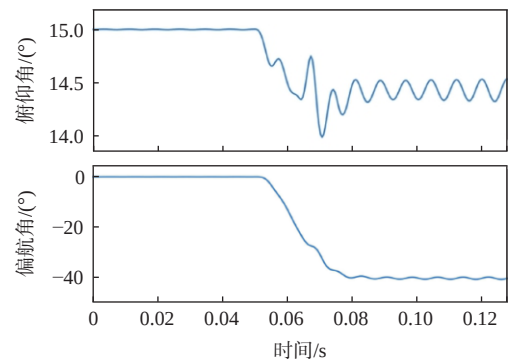


图 8　旋转碟状飞行器姿态角曲线

Fig. 8　Curves of attitude angles for rotating disk-shaped flight vehicle

力，从而产生Oz_2轴反方向的力矩，使飞行器发生进动，即产生了沿Ox_2轴反向的角加速度，使飞行器绕Ox_2轴发生旋转，偏航角ψ因此发生较大变化，由 0°快速变为−40°，飞行器俯仰角ϑ变化较小，由触水前的 15°下降为 14.4°左右。同时，飞行器产生章动现象，表现为俯仰角ϑ和偏航角ψ在运动过程中高频波动。流体在与飞行器碰撞过程中也产生了较大的变化，与飞行器最先接触的流体由于扰动而向下移动，这种扰动会向流体的四周扩散，在水面产生一个以碰撞点为中心的波浪。随着时间的推移，飞行器因进动性偏航角产生变化，靠近左侧部分先发生碰撞，受到挤压的流体单元向四周传递能量，使波浪以椭圆形向外传播。

在飞行器离开水面后，仅受到重力，因此，质心运动可以视为斜抛。但由于与水发生的碰撞，其旋转轴发生偏转，表现为俯仰角ϑ和偏航角ψ会在一个定值附近快速波动。而流体由于与飞行器的碰撞，吸收了大量能量，产生的椭圆形波浪不断向四周传播，且流体单元运动较为剧烈的区域会发生破碎和飞溅。在碰撞中心位置，水面产生较大的凹坑，且四周的流体在重力作用下开始向中心聚拢。

由此可见，旋转碟状飞行器水漂是一个以剧烈

动量交换为特征的物理现象，飞行器与流体的碰撞过程在0.05 s时刻发生，并持续相互作用0.03 s，在这一段时间内，飞行器的运动状态发生较大的变化，且受到较大过载，使其迅速改变运动方向，脱离水面。

3.2 构型参数影响规律

3.2.1 构型参数k_1影响规律

为研究旋转碟状飞行器构型对其水漂过程中动力学特性的影响规律，在2.2节设置的初始运动参数及构型参数的基础上分别改变某一构型参数进行仿真。改变构型参数k_1，分别对$k_1=0.5$、$k_1=1.5$、$k_1=2.5$这3种情况下碟状飞行器水漂进行仿真模拟。构型参数k_1对旋转碟状飞行器水漂过程中高度y、速度V和过载的影响如图9所示。

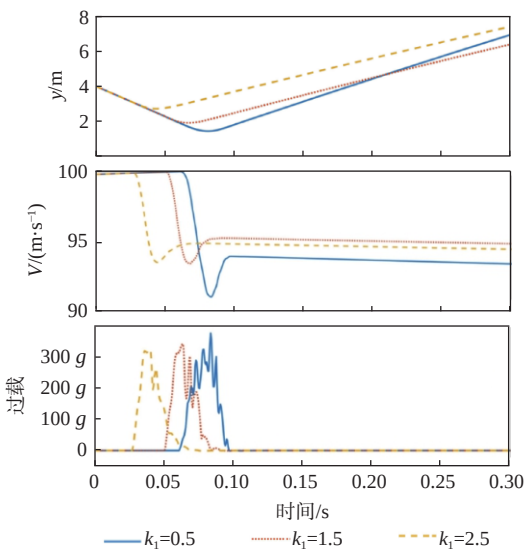


图9 不同构型参数k_1下的运动参数曲线

Fig. 9 Curves of motion parameters for different configuration parameters k_1

从图9可以看出，在3种不同情况下，飞行器触水时间有明显的不同，分别在仿真开始后0.027 s、0.05 s、0.06 s开始与流体发生作用，并改变运动轨迹。这是因为随着构型参数k_1的增大，飞行器厚度增加，更早地与水面发生接触。根据高度-时间曲线中的斜率，触水前，三者竖直方向速度相等，但离开水面后，飞行器竖直方向速度反向，$k_1=1.5$和$k_1=2.5$这2种工况斜率（即竖直方向速度V_{y_0}）基本相同，为20 m/s左右，相较于触水前减小了41.17%，而$k_1=0.5$时，触水后的竖直方向速度较大，仅减小了22.87%。因此，随着k_1增大，飞行器触水时间将会提前，且离开水面后的竖直方向速度减小，使速度倾角相较于初始状态变小，不利于连续水漂时飞行器运动状态的稳定。

在速度-时间曲线中，触水前3种情况下的飞行器速度保持一致，均在重力的作用下缓慢增加。触水时，飞行器速度因碰撞迅速减小，并在轨迹最低点取得最小值，随后，在流体反作用力下速度开始增加，直至离开水面后保持稳定，在重力作用下缓慢减小。因为构型的不同，触水时的最小速度和离开水面时的速度均有不同。对于最小速度，3种情况下分别为91.0 m/s、93.5 m/s、93.6 m/s，而对于离开水面时的速度，3种情况下分别为94.0 m/s、95.4 m/s、95 m/s，这说明与高度y的变化类似，$k_1=1.5$和$k_1=2.5$这2种情况下的最小速度并没有较大区别，仅仅在离开水面时的速度中，$k_1=1.5$的飞行器稍大，但$k_1=0.5$时，速度的波动明显增大，最小速度也显著降低。采用速度衰减率α=(初始速度−离水速度)/初始速度，可以得到3种情况下的飞行器速度衰减率分别为6%、4.6%、5%，可以发现，$k_1=1.5$的飞行器更能减少速度的损失，因此，更容易完成多次水漂。

从过载-时间曲线可以看出，飞行器在与流体碰撞过程中会承受300g以上的过载，且随着入水深度的增加，其所受过载也逐渐增加。在3种不同情况时，飞行器所受过载最大值分别为375.2g、340.9g、323.4g。相较于$k_1=0.5$时，$k_1=1.5$时飞行器所受最大过载减小了9.14%，$k_1=2.5$时减小了13.8%，由此可见，随着构型参数k_1的增大，飞行器水漂时过载减小。同时，可以看出，$k_1=0.5$时过载曲线具有较大的波动，且波动频率与飞行器自旋角速度频率一致，而当$k_1=2.5$时，曲线的波动明显减小，平滑程度明显提高。这说明，k_1增大也能减小飞行器在水漂时过载的波动。

在姿态角随构型参数k_1的变化曲线中同样如此，如图10所示，在$k_1=1.5$和$k_1=2.5$这2种情况下，飞行器俯仰角和偏航角的曲线呈现变化规律相似且离水后数值相差不大的现象，但$k_1=0.5$时出现了较大的变化。在触水时，飞行器的俯仰角在外力矩的作用下减小，飞行器表现出轻微的“低头”，但

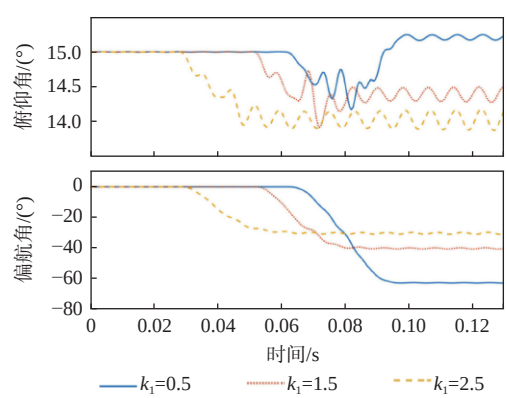


图10 不同构型参数k_1下姿态角变化曲线

Fig. 10 Curves of attitude angle variation for different configuration parameters k_1

在离开水面前迅速增加，最终超过初始的俯仰角。在 3 种情况下，偏航角因为旋转飞行器进动性产生偏转，但$k_1 = 0.5$时，这种偏转现象更为剧烈。

$k_1 = 0.5$时，旋转碟状飞行器水漂的特殊现象是由于此时飞行器厚度仅为 1 m，外形与圆形薄板类似，而另外 2 种情况厚度分别为 3 m 和 5 m，飞行器外形介于球体与圆形薄板之间。触水时，$k_1 = 0.5$的飞行器高速冲击水面，底部向前挤压流体，并在后方形成空泡，依靠飞行器上下两侧的压强差提供升力，使飞行器成功水漂。而当飞行器表面具有一定弧度，其水漂时边缘的一小部分首先接触水面，触水面积随水漂过程的推进逐渐增加，变化相对平缓，过载和速度等参数的变化幅度因此减小。

3.2.2　构型参数k_2影响规律

为研究旋转碟状飞行器边缘曲率对水漂动力学特性的影响，单独改变k_2对飞行器水漂进行数值仿真。当构型参数k_2减小时，飞行器边缘变得更为尖锐，反之则更为平滑。选取$k_2 = 0$、$k_2 = 0.2$、$k_2 = 0.5$、$k_2 = 1$这 4 种情况，研究k_2的变化对速度V、过载、偏航角ψ、俯仰角ϑ的影响，得到 4 个参数变化曲线如图 11 所示。

在速度-时间曲线中，曲线变化趋势同上文一致，在飞行器触水时下降，在要离开水面前上升，最终稳定在某一值，在重力作用下逐渐减小。对比

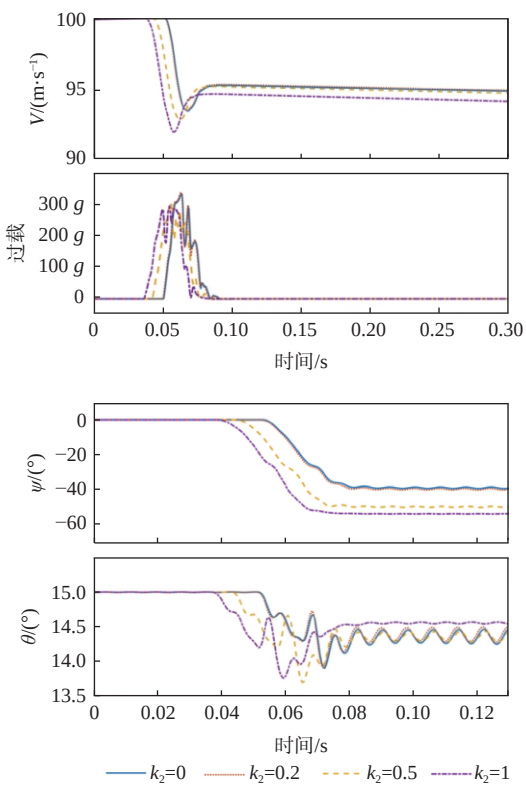


图 11　不同构型参数k_2下参数变化曲线

Fig. 11　Curves of parameter variation for different configuration parameters k_2

k_2取不同值时的速度变化情况，可以发现，随着k_2的增大，飞行器离水速度和最小速度都会减小，但$k_2 = 1$时，飞行器水漂后的速度损失相对较大，衰减率为 5.5%，而其他情况下的速度衰减率均在 5% 左右。同时，由于$k_2 = 0$和$k_2 = 0.2$时飞行器构型相差较小，两者在 4 个图中的曲线都基本重合。 在过载-时间曲线可以看出，随着k_2的增大，过载逐渐减小。相较于$k_2 = 0$，$k_2 = 0.5$时飞行器过载减小了 7%，$k_2 = 1$时过载减小了 11%。

观察偏航角-时间曲线和俯仰角-时间曲线可以看出，随着k_2增大，飞行器离开水面后的偏航角也会变大，但其波动变小。$k_2 = 1$时，虽然偏航角达到−60°，但角度的波动基本消失。而对于俯仰角变化，$k_2 = 0$、$k_2 = 0.2$、$k_2 = 0.5$这 3 种情况下，飞行器水漂后的俯仰角基本相同，且都在一个范围内不断波动，但$k_2 = 1$时，俯仰角在触水时先下降后上升，最终稳定在 14.5°左右，这使$k_2 = 1$时的飞行器相较于其他 3 种情况具有更高的稳定性，同时与偏航角相似，飞行器水漂后的俯仰角波动较弱，适合多次弹跳。

3.2.3　构型参数k_3影响规律

本节研究飞行器靠近中间部分曲率对水漂运动的动力学特性影响规律。根据 2.1 节对构型参数的讨论，随着k_3增大，飞行器靠近中间部分曲率减小。选取$k_3 = 0$、$k_3 = 5$、$k_3 = 10$、$k_3 = 15$这 4 种构型对旋转碟状飞行器进行数值仿真。

由于飞行器以一定俯仰角入水，且不会完全浸入水中，因此，在触水时靠近中部的部分并不会与流体发生大量的接触。同时，由于构型参数k_3对飞行器外形的影响较小，k_3对水漂过程中各个参数的影响也较为微弱，仅选择过载和俯仰角 2 个参数进行分析。不同构型参数k_3下两者随时间变化曲线如图 12 所示。

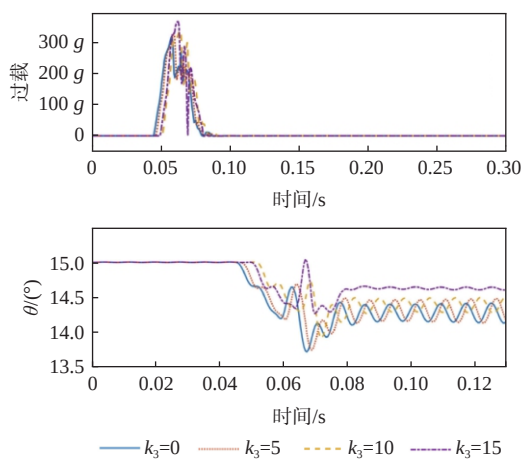


图 12　不同构型参数k_3下参数变化曲线

Fig. 12　Curves of parameter variation for different configuration parameters k_3

从过载-时间曲线中可以看出，随着k_3增大，飞行器过载也随之增加。当$k_3=15$时，飞行器过载最大，为369.5g，与此构型相比，$k_3=10$时过载下降了7.7%，$k_3=5$时下降了9.8%，$k_3=0$时下降了12.9%。在俯仰角-时间曲线中，随着k_3增大，飞行器水漂后姿态角的变化幅度减小，且在水漂后姿态角的波动也明显减弱。$k_3=15$时俯仰角减小了2.5%，为其他构型飞行器俯仰角变化的一半。

3.3 初始运动参数影响规律

为探究旋转碟状飞行器初始运动参数对水漂动力学特性的影响，分别改变飞行器初始俯仰角ϑ、速度V、速度倾角φ和自旋角速度ω，并选取$k_1=1.5$、$k_2=0.2$、$k_3=10$的飞行器构型进行数值仿真，分析不同参数变化对飞行器水漂运动的影响规律。

3.3.1 俯仰角影响规律

选取俯仰角为0°～30°、间隔为5°的共7种工况进行仿真，得到各运动参数关于俯仰角的变化情况，如图13所示。

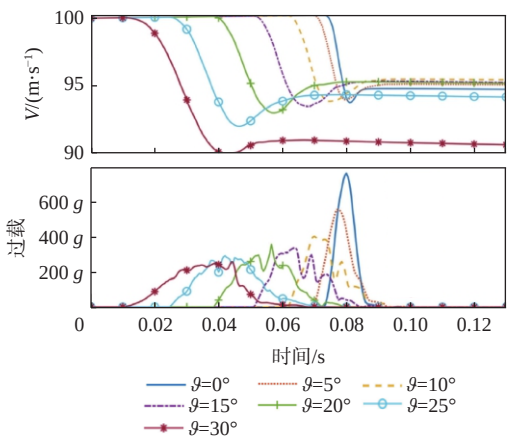


图13 不同俯仰角下参数变化曲线

Fig. 13 Curves of parameter variation for different pitch angles

从图13可以看出，俯仰角的变化对于速度的影响较小，但在$\vartheta=30°$时，飞行器最小速度及离水速度均有较大的衰减，这是因为当俯仰角增大时，飞行器受到的向后的作用力分量增大，向上的反作用力分量减小，能量因此消耗较多，离水速度也因此大量衰减。在过载-时间曲线图像中，随着ϑ增大，飞行器过载最大值降低，但降低的幅度逐渐减小。在$\vartheta=0°$时，最大过载接近800g，但过载曲线较为平滑。此时，飞行器中部触水，由于中部曲率较小，飞行器与流体接触的面积较大，与流体的碰撞和能量交换也较为剧烈，因此，$\vartheta=0°$时过载较大。由于飞行器中部触水，与流体发生的动量交换主要在竖直方向，使触水后水面产生圆形波纹。在飞行器的姿态方面，随着ϑ增大，飞行器与流体碰撞位置逐渐向边缘靠近，因此，所受Oz_2方向的外力矩也逐渐增大，根据陀螺的进动原理，其姿态的偏转也逐渐增大。

因此，要想实现飞行器连续多次成功水漂，俯仰角的选择至关重要，较小的俯仰角虽然能使飞行器具有较高的离水速度和较为稳定的姿态，但其过载很大；较大的俯仰角可以减小过载，但会使速度大量衰减，同时也不利于飞行器姿态的稳定。

3.3.2 速度和速度倾角影响规律

保持速度倾角φ及其他参数不变，分别在$V=50$ m/s、$V=100$ m/s、$V=150$ m/s、$V=200$ m/s这4种工况下对旋转碟状飞行器进行数值仿真，得到飞行器过载、俯仰角、偏航角关于速度V的变化情况，如图14所示。

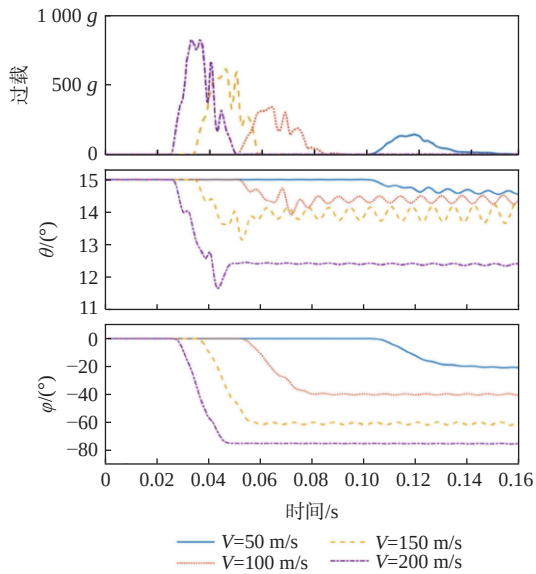
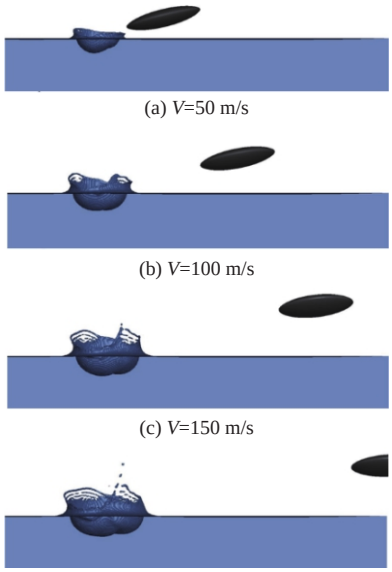


图14 不同速度下参数变化曲线

Fig. 14 Curves of parameter variation for different velocities

在过载-时间曲线中，飞行器运动过程中的最大过载随着速度增大而增大。当$V=200$ m/s时，水漂过程中过载最大为821.8g，与此种工况相比，$V=150$ m/s、$V=100$ m/s、$V=50$ m/s时，最大过载分别减小25.5%、58.5%、82.4%。可以看出，速度大小对于飞行器水漂运动过程中的过载有着巨大影响，这是由于随着飞行器速度增大，结构与流体的碰撞更为剧烈，碰撞时间更短，速度变化也更大，因此，过载更大。

这一点通过飞行器水漂后的液面状态也可以明显地观察到，飞行器在0.3 s时的液面状态如图15所示。随着飞行器初始速度的增加，流体单元由于碰撞而吸收到的能量也会增加，因此，与飞行器碰撞后的速度也会增加。表现出的液面状态即为空泡体积的增大和流体飞溅程度的上升。而且在$V=150$ m/s和$V=200$ m/s这2种工况下，碰撞中心出现明显射流，并在0.3 s时由于表面张力而破碎。



(a) V=50 m/s

(b) V=100 m/s

(c) V=150 m/s

(d) V=200 m/s

图 15　不同速度下液面状态

Fig. 15　Liquid states at various velocities

在姿态角变化方面，随着速度增加，飞行器姿态的变化也不断增大，这同样是由速度引起的碰撞剧烈程度不同导致的。因此，可以认为在一定范围内，速度的提升不利于旋转碟状飞行器连续水漂时姿态的稳定。

选取速度倾角为5°～35°、间隔为5°的共 7 种工况进行仿真，得到过载和偏航角关于速度倾角的变化规律，如图 16所示。

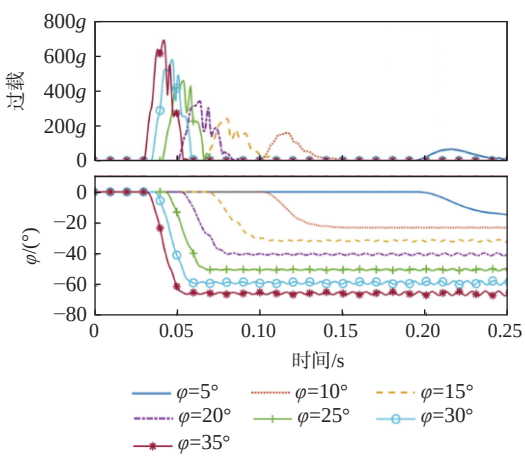


图 16　不同速度倾角下参数变化曲线

Fig. 16　Curves of parameter variation for different velocities inclination angle

可以看出，随着速度倾角减小，飞行器水漂过程中的过载有明显降低，与过载最大的$\varphi=35°$时的工况相比，$\varphi=5°$时的最大过载仅为其 9.5%。速度倾角对飞行器水漂时的过载具有较大的影响，主要是因为随着速度倾角的减小，飞行器竖直方向的速度分量也会减小，因此，在与水流碰撞直至离开水面时的速度改变量也较小。同时，由于竖直方向分量的减小，飞行器与流体的接触时间也会延长，共同作用导致过载的减小。在偏航角的变化方面，从图 16 可以看出，随着速度倾角的增大，飞行器偏航角也不断增加，且增加幅度相对均匀。

因此，为保持旋转碟状飞行器姿态稳定及所受载荷尽可能小，速度倾角应在保持飞行器成功水漂的基础上尽可能小。

3.3.3　自旋角速度影响规律

本节研究旋转碟状飞行器自旋角速度对水漂动力学特性的影响。选取$\omega=100$ rad/s、$\omega=200$ rad/s、$\omega=400$ rad/s、$\omega=600$ rad/s、$\omega=800$ rad/s、$\omega=1\,000$ rad/s、$\omega=1\,200$ rad/s这 7 种工况进行仿真，得到不同工况下过载及姿态角随时间变化曲线，如图 17 所示。在过载曲线中，由于当$\omega\geqslant800$ rad/s时的 3 种工况变化趋势相同，且变化幅度相差不大，为方便观察与对比，图 17 仅保留$\omega=100$ rad/s、$\omega=400$ rad/s、$\omega=1\,200$ rad/s这 3 种工况下的过载变化曲线。同时，由于$\omega=100$ rad/s时飞行器发生水漂后姿态变化过于剧烈，在离开水面后，飞行器连体坐标系的Ox_2轴指向后方，俯仰角、偏航角均变为负值，飞行器发生翻转，水漂失败，无法对姿态角变化进行对比，因此，在 2 个姿态角变化曲线中省略了$\omega=100$ rad/s这一工况。

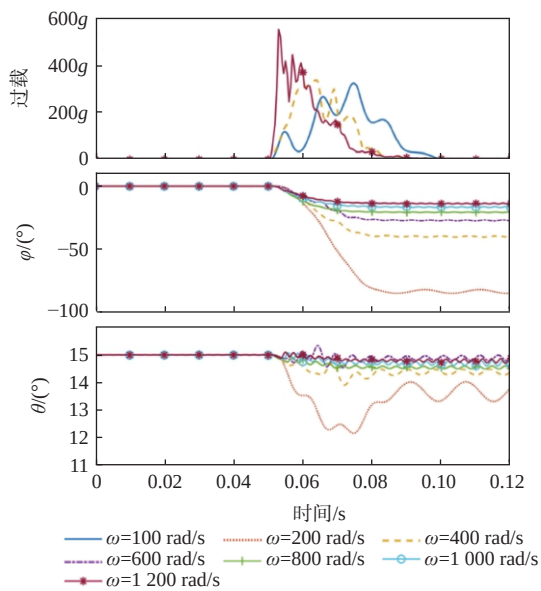


图 17　不同自旋角速度下参数变化曲线

Fig. 17　Curves of parameter variation for different spin angular velocities

从过载曲线可以发现，随着ω增大，飞行器的最大过载也逐渐增大。$\omega=1\,200$ rad/s时，最大过载为 557g，$\omega=400$ rad/s时，过载为 340.9g，下降了 39%，$\omega=100$ rad/s时，过载为 327.3g，下降了 41%。同时，在$\omega=1\,200$ rad/s的工况下，飞行器触水后，过载迅

速增大到最大值，随后伴随着波动缓慢下落，直至飞行器脱离水面。而其他工况下，飞行器过载的增速比较缓慢。

观察2个姿态角变化曲线可以看出，随着自旋角速度增大，俯仰角和偏航角在水漂结束后逐渐接近初始值。根据陀螺力学原理，随着自旋速度增大，陀螺定轴性也会不断增强，这与仿真得出的结果相同。但当飞行器自旋角速度$\omega > 800$ rad/s时，随着自旋角速度的增加，姿态角稳定性的提升不再明显。

总结仿真得到的所有结果，可汇总成构型参数及初始运动参数的影响规律，如表4所示。

表4 构型参数和初始运动参数影响规律

Table 4 The influence pattern of configuration parameters and initial motion parameters

影响因素	影响因素			
	V	过载	$\Delta\theta$	$\Delta\varphi$
k_1	\	−	+	−
k_2	+	−	−	+
k_3	+	+	−	\
θ	−	−	\	−
V	\	+	+	+
速度倾角	−	+	+	+
ω	−	+	+	−

注：首行中的V为旋转碟状飞行器离开水面后的速度大小；过载为飞行器在水漂过程中所受的最大过载；“$\Delta\theta$”“$\Delta\varphi$”为飞行器离开水面后与初始时刻的角度之差的绝对值，即角度的变化量，可以表示飞行器姿态的稳定性；“\”表示该参数不随初始参数单调变化；“+”表示正相关，“−”表示负相关。

4 结 论

本文通过对旋转碟状飞行器建立数值仿真模型，利用ALE方法进行飞行器水漂仿真，得到了飞行器水漂过程各个参数的变化情况，进而分析飞行器成功水漂的原因和运动状态。同时，分别改变不同构型参数及初始运动参数，对比得到各个初始参数变化对水漂的影响规律。

1）旋转碟状飞行器构型参数会对其水漂动力学特性产生一定影响。在其他参数不变的情况下，飞行器过载随着构型参数k_1和k_2增大而减小，随着k_3的增大而增大；俯仰角差值随着k_1增大而增大，随k_2和k_3增大而减小；偏航角差值随k_1增大而减小，随k_2增大而增大。由各构型参数的几何意义可以总结出，旋转碟状飞行器边缘曲率半径越大，所受过载越小，飞行器姿态越稳定。

2）碟状飞行器初始运动参数会对其水漂动力学特性产生一定影响。在其他参数不变的情况下，飞行器过载随着V、φ和ω的增大而增大，随ϑ增大而减小；$\Delta\theta$随着V和φ增大而增大，随ω增大而减小；$\Delta\varphi$随ϑ和ω增大而减小，随V和φ增大而增大。

参考文献（References）

[1] BOCQUET L. The physics of stone skipping[J]. American Journal of Physics, 2003, 71(2): 150-155.

[2] ROSELLINI L, HERSEN F, CLANET C, et al. Skipping stones[J]. Journal of Fluid Mechanics, 2005, 543: 137-146.

[3] CLANET C, HERSEN F, BOCQUET L. Secrets of successful stone-skipping[J]. Nature, 2004, 427(6969): 29.

[4] 李超, 吕日毅, 钱仁军, 等. 美跨介质飞行器技术发展与运用研究[J]. 战术导弹技术, 2023(6): 120-127.
LI C, LV R Y, QIAN R J, et al. Research on technology development and application of American trans-media vehicle[J]. Tactical Missile Technology, 2023(6): 120-127(in Chinese).

[5] 周正, 黄大志, 唐雯铠, 等. 潜空跨介质无人航行器研究现状与分析[J]. 科技创新与应用, 2024, 14(3): 5-7.
ZHOU Z, HUANG D Z, TANG W K, et al. Research status and analysis of underwater-aerial trans-media unmanned vehicle[J]. Technology Innovation and Application, 2024, 14(3): 5-7(in Chinese).

[6] 罗剑桥, 解春雷, 金泽华, 等. 跨介质飞行器近水面滑跳流固耦合仿真及可滑跳区域研究[J]. 航空学报, 2023, 44(21): 377-394.
LUO J Q, XIE C L, JIN Z H, et al. Water-skipping fluid-structure interaction simulation and slippable area study of trans-medium vehicle[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(21): 377-394(in Chinese).

[7] 田北晨, 刘涛涛, 吴钦, 等. 跨介质飞行器触水滑跳运动特性数值模拟[J]. 兵工学报, 2022, 43(3): 586-598.
TIAN B C, LIU T T, WU Q, et al. Numerical simulation on kinematic characteristics of trans-media aircraft during water-skipping [J]. Acta Armamentarii, 2022, 43(3): 586-598(in Chinese).

[8] 孙士明, 郁伟, 王晓辉, 等. 基于多目标优化的跨介质航行体水面滑跳初始运动参数设计[J]. 兵工学报, 2024, 45(2): 541-551.
SUN S M, YU W, WANG X H, et al. Initial kinematic parameters design of trans-media vehicle skipping over water surface based on multi-objective optimization[J]. Acta Armamentarii, 2024, 45(2): 541-551(in Chinese).

[9] 侯东伯. 运动体触水滑跳过程运动特性研究[D]. 哈尔滨: 哈尔滨工业大学, 2020.
HOU D B. Research on kinematic characteristics of moving object during water-skipping[D]. Harbin: Harbin Institute of Technology, 2020(in Chinese).

[10] 李聪慧. 圆盘近水面滑跳运动多相流动特性研究[D]. 哈尔滨: 哈尔滨工业大学, 2022.
LI C H. Study on multiphase flow characteristics in neat-watersurface skipping motion of disk[D]. Harbin: Harbin Institute of Technology, 2022(in Chinese).

[11] 付晓琴, 李阳辉, 卢昱锦, 等. 二维平板水漂运动数值模拟[J]. 航空学报, 2021, 42(6): 269-286.
FU X Q, LI Y H, LU Y J, et al. Numerical simulation of two-dimensional plate skipping[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(6): 269-286(in Chinese).

[12] 杨超, 姜宇, 吴志刚. 基于边界元法的三维结构体滑跳运动数值仿真[J]. 北京航空航天大学学报, 2022, 48(9): 1678-1691.
YANG C, JIANG Y, WU Z G. Numerical simulation of skipping motion of three-dimensional structure based on boundary element method[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(9): 1678-1691(in Chinese).

[13] 李飞, 孙凌玉, 张广越, 等. 圆柱壳结构入水过程的流固耦合仿真与试验[J]. 北京航空航天大学学报, 2007, 33(9): 1117-1120.
LI F, SUN L Y, ZHANG G Y, et al. Simulation and experiment of cylinder shell structure dropping into water based on fluid structure interaction[J]. Journal of Beijing University of Aeronautics and Astronautics, 2007, 33(9): 1117-1120(in Chinese).

[14] 任帆涛, 牛钰森, 姜毅, 等. 透镜飞行器击水弹跳运动数值模拟[J]. 兵工学报, 2024, 45(5): 1482-1496.
REN F T, NIU Y S, JIANG Y, et al. Numerical simulation of lenticular aircraft during water-skipping[J]. Acta Armamentarii, 2024, 45(5): 1482-1496(in Chinese).

[15] 赵坤. 水漂动力学建模与仿真[D]. 哈尔滨: 哈尔滨工业大学, 2014.
ZHAO K. Dynamics modling and simulation of hydroplaning[D]. Harbin: Harbin Institute of Technology, 2014(in Chinese).

[16] 华厦, 邹恒, 徐鹏, 等. 基于 SPH 方法与 ALE 方法的圆盘水漂对比研究[J]. 科学技术与工程, 2013, 13(2): 421-424.
HUA X, ZOU H, XU P, et al. Comparative study of stone skipping with SPH method and ALE method[J]. Science Technology and Engineering, 2013, 13(2): 421-424(in Chinese).

[17] 孙士明, 颜开, 陈玮琪. 圆盘水面弹跳过程运动特性数值研究[J]. 船舶力学, 2021, 25(6): 716-725.
SUN S M, YAN K, CHEN W Q. Numerical study on kinetic characteristics of disk skipping over water surface[J]. Journal of Ship Mechanics, 2021, 25(6): 716-725(in Chinese).

[18] NAGAHIRO S I, HAYAKAWA Y. Theoretical and numerical approach to “magic angle” of stone skipping[J]. Physical Review Letters, 2005, 94(17): 174501.

[19] MAKINO M. Stone skipping simulation by ALE and SPH[C]//Proceedings of the 11th International LS-DYNA Users Conference. Trento: SPHERIC, 2010: 33-38.

[20] 闫蕊. 基于 SPH 方法的结构物入水若干问题研究[D]. 西安: 西北工业大学, 2016.
YAN R. Research on some problems in structure impact with water using SPH method[D]. Xi’an: Northwestern Polytechnical University, 2016(in Chinese).

[21] YAN R, MONAGHAN J J. SPH simulation of skipping stones[J]. European Journal of Mechanics-B/Fluids, 2017, 61: 61-71.

[22] LI C H, WANG C, WEI Y J, et al. Three-dimensional numerical simulation of cavity dynamics of a stone with different spinning velocities[J]. International Journal of Multiphase Flow, 2020, 129: 103339.

[23] LI C H, WANG C, WEI Y J, et al. Hydrodynamic force and attitude angle characteristics of a spinning stone impacting a free surface[J]. Physics of Fluids, 2021, 33(12): 123309.

[24] LI C H, WANG C, WEI Y J, et al. Numerical investigation on the cavity dynamics and deviation characteristics of skipping stones[J]. Journal of Fluids and Structures, 2021, 104: 103301.

[25] 邬明. LS—DYNA 的 ALE 方法在圆盘击水滑跳中的应用[J]. 科学技术与工程, 2011, 11(33): 8247-8251.
WU M. Numerical simulation research on bounce of circular disks Base on the ALE of LS-DYNA[J]. Science Technology and Engineering, 2011, 11(33): 8247-8251(in Chinese).

[26] SURHONE L M, TENNOE M T, HENSSONOW S F. Pye Wacket[M]. Saarbrücken: Betascript Publishing, 2011.

[27] 马文朝, 孟繁敏, 马诺, 等. 跨介质飞行器头部外形优化及入水性能分析[J]. 兵工学报, 2022, 43(10): 2588-2597.
MA W C, MENG F M, MA N, et al. Head shape optimization and water entry performance analysis of trans-medium aircraft[J]. Acta Armamentarii, 2022, 43(10): 2588-2597(in Chinese).

Dynamic characteristics of rotating disk-shaped flight vehicle during skipping considering configuration

WANG Di[1, 2], LUO Jianqiao[1, 2], YU Ning[1, 2], MENG Junhui[1, 2, *]

(1. School of Aerospace Engineering, Beijing Institute of Technology, Beijing 100081, China;
2. Chongqing Innovation Center, Beijing Institute of Technology, Chongqing 401120, China)

Abstract: The rotating disk-shaped vehicle can significantly enhance penetration capabilities in denied environments by imitating stone skipping. Ensuring continuous skipping and reducing impact overload are critical dynamic design objectives. While setup settings are also important for the vehicle's posture motion and load distribution during skipping, the dynamic aspects of skipping are mostly determined by the starting motion parameters. Therefore, the parameters affecting the dynamic characteristics of the rotating disk-shaped vehicle are strongly coupled, making the mechanisms of their influences remain unclear. This paper proposes a parameterized configuration for the disk-shaped vehicle. Utilizing the arbitrary Lagrange-Euler (ALE) method and penalty function method, a simulation analysis of the rotating disk-shaped vehicle is conducted. The study investigates the coupling relationship between configuration parameters and starting motion parameters and reveals their influence patterns on the vehicle's dynamic characteristics while skipping. The results demonstrate that increasing the edge curvature radius can effectively reduce the overload on the vehicle during skipping. Additionally, a better attitude stability can be achieved by increasing the spin rate. The findings presented in this paper can serve as valuable references for the design of new types of skipping anti-ship weapons, such as a rotating disk-shaped vehicle.

Keywords: rotating disk-shaped vehicle; dynamics of skipping; motion parameters; configuration parameters; arbitrary Lagrange-Euler finite element method

Received: 2024-05-07; **Accepted**: 2024-06-21; **Published Online**: 2024-08-19 15:13

URL: link.cnki.net/urlid/11.2625.V.20240816.1916.003

Foundation items: National Natural Science Foundation of China (52372346); Innovative Talent Science and Technology Funding Project of Beijing Institute of Technology (2021CX01018)

* **Corresponding author**. E-mail: mengjh@bit.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0287

基于结构重参数化的深度可分离卷积神经网络

陈红[1]，闫建国[2]，杨华[1]，张靖[1]，李伟[3]，杨靖[1,4,*]

(1. 贵州大学 电气工程学院，贵阳 550025；2. 中国电建集团贵州工程有限公司，贵阳 550025；3. 贵州省普通高等学校工程研究中心，贵阳 550025；4. 贵州大学 贵州省“互联网+”协同智能制造重点实验室，贵阳 550025)

摘 要：针对目前卷积神经网络(CNN)模型大部分采用的单分支深度卷积方法影响模型的表达能力、占用大量参数和Flops的问题，提出一种新的轻量级CNN模型——基于结构重参数化的深度可分离卷积神经网络(DSCNN)。该模型的特征提取模块结构重参数化混合深度卷积(RepMiX)能融合不同通道和空间位置之间的信息，实现多尺度特征融合；基于结构重参数化的DSCNN中的重参数化非对称空间算子(RepASO)通过具有不同功能的多分支结构来学习不同信道特征信息，提高模型特征学习能力；RepMiX和RepASO均结合结构重参数化技术和深度可分离卷积(DS-Conv)思想，实现了训练和推理阶段的结构解耦，在减少模型参数和Flops的同时加速模型推理；在Tiny-ImageNet-200、CIFAR-10、CIFAR-100数据集及自建铝锭表面缺陷分类数据集上进行了对比实验，结果表明：基于结构重参数化的DSCNN实现了更高的浮点运算速度，并保持了具有竞争力的精度。

关 键 词：卷积神经网络；结构重参数化；深度可分离卷积；非对称卷积；结构解耦
中图分类号：V557+.1; TP751
文献标志码：A **文章编号**：1001-5965(2026)06-2145-11

卷积神经网络(convolutional neural network, CNN)在表示高维复杂函数方面表现出强大的能力，已成功应用于各种视觉场景[1]。与视觉变换器(vision Transformer, ViT)相比，CNN计算复杂度更低，效率更高，因此，在计算机视觉应用中占据主导地位。目前，为了达到更高的精度，主流的CNN体系结构复杂而庞大，对设备的计算和存储能力有很高的要求。为了在移动设备上实现实际应用，需要在保持高精度的同时，建立参数量少、Flops更少的轻量级CNN模型。

深度可分离卷积(depthwise separable convolution, DS-Conv)是在MobileNetV1[2]中提出的，与普通卷积不同，其由2个连续的部分组成：① 深度卷积(depthwise convolution, DW-Conv)，用于提取空间特征；② 点卷积(pointwise convolution, PW-Conv)，用于提取信道特征。DS-Conv减少了计算负担和模型参数，成为后续轻量级CNN[3]及现代大规模CNN[4-5]的基本设计组件。在MobileNetV1的基础上，MobileNetV2[6]引入了倒残差模块(inverted residual block, IRB)，以减轻特征损失并增强表示能力。利用IRB设计范式作为基本模块，学者提出了具有不同整体架构的各种轻量级模型。除了IRB，研究人员还设计了具有不同结构的高效基本块[7]，以增强轻量级CNN的表达能力。然而，架构的不

收稿日期：2024-05-07；**录用日期**：2024-06-25；**网络出版时间**：2024-07-03 15:00
网络出版地址：link.cnki.net/urlid/11.2625.V.20240703.0858.001
基金项目：国家自然科学基金(61640014,61963009)；贵州省教育厅创新群体(黔教合KY字[2021]012)；贵州省科技支撑计划(黔科合支撑[2022]一般017,黔科合支撑[2023]一般411,黔科合支撑[2024]一般051)；贵州省教育厅工程研究中心(黔教技[2022]040,黔教技[2022]043)；中国电建集团科技项目(DJ-ZDXM-2020-19,DJ-ZDXM-2022-44)；贵州省双碳研究院开放课题(DCRE-2023-13)；中铝国际工程股份有限公司科技重大专项(CJ2022JS-ZD02)

* **通信作者**. E-mail：jyang7@gzu.edu.cn

引用格式：陈红，闫建国，杨华，等. 基于结构重参数化的深度可分离卷积神经网络[J]. 北京航空航天大学学报，2026，52(6)：2145-2155.
CHEN H，YAN J G，YANG H，et al. Deep separable convolutional neural networks based on structural reparameterization[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：2145-2155 (in Chinese).

一致性导致结构冗余，使用不同的基本模块和架构导致不公平的容量比较，模糊了架构选择与性能提升之间的因果关系。

此外，虽然一些工作[8]将DW-Conv分解为并行的低秩分支以节省计算成本，但轻量级模型追逐更低Flops和参数的主要主干在于PW-Conv。ShuffleNet[9]进行了信道分组和洗牌，但不同分组的信道之间信息交换不足。IGCV3[10]提出了交错组卷积，其将常规卷积分解为多层组卷积，结构更为复杂。ChannelNets[11]将稀疏连接引入PW-Conv，对于每个输入通道，输出通道只关注其中的一小部分。

为解决上述问题，本文提出了一种参数和Flops更少的轻量级CNN模型，该模型在不影响精度的前提下尽可能轻量化，实现了精度和速度的平衡。本文重点讨论构建轻量级CNN模型的4个关键部分：整体架构、特征提取模块、空间算子和基本模块。主要贡献如下：

1）设计了特征提取模块结构重参数化混合深度卷积（reparameterization MiXer, RepMiX），一种由RepMixer和混合模块组成的类似Transformer的架构。RepMixer结合了多个并行DW-Conv，并结合重参数化方法来加快推理，混合模块混合信道和位置信息。

2）设计了重参数化非对称空间算子（reparameterized asymmetric spatial operator, RepASO）作为基于结构重参数化的深度可分离卷积神经网络（deep separable convolutional neural network, DSCNN）模型的空间算子。模型训练时，RepASO为并行多分支架构，以弥补可学习参数减少带来的影响，同时增强模型的表征能力。模型推理时，RepASO利用重参数化将多分支架构转化为单一分支，提高模型的推理速度。

3）基于RepMiX和RepASO模块，提出了名为基于结构重参数化的DSCNN的轻量级CNN模型。实验表明，相比现有轻量级CNN模型，基于结构重参数化的DSCNN具有更快的浮点运算速度，并保持了相当甚至更高的精度。

1 相关理论

1.1 结构重参数化

结构重参数化[2,12]是重参数化的一种代表性方法，使用某一结构转移过来的参数对新结构进行参数化。ACNet[13]在训练中使用常规卷积分支，并构建垂直和水平2个卷积分支，推理时，这些分支会转化为单分支的常规卷积。RepVGG[12]在训练期间采用多分支拓扑结构，通过结构重参数化实现训练和推理结构的解耦。MobileOne[14]利用结构重参数化，在训练期间为DW-Conv和PW-Conv多个并行分支，在推理过程中没有分支，而是采用直通式架构。结构重参数化的具体方法与RepVGG是一致的，即把批归一化（batch normalization, BN）融合到前面的卷积中，合并并行卷积参数。

总之，尽管轻量级CNN模型在减少模型参数和计算复杂度方面取得了一定的成功，但由于设计目标是轻量级，不得不在模型准确性和学习能力等方面做出妥协，面临如何在速度和准确性之间取得平衡的挑战。本文提出的基于结构重参数化的DSCNN模型，是将结构重参数化技术、DS-Conv与非对称卷积方法相结合的新方法，该方法为解决上述问题提供了新的解决方案。

1.2 深度卷积

DW-Conv是常规卷积的一种高性能变体，已被广泛采用为许多神经网络的关键构建模块。对于输入$F \in R^{m\times n\times w}$，DW-Conv应用$m$个滤波器来计算输出。每个滤波器在输入通道上的DW-Conv可以写成如下形式：

$$\hat{G}_{k,l,m}=\sum_{i,j}\hat{k}_{i,j,m}*F_{k+i-1,l+j-1,m} \tag{1}$$

式中：$\hat{k}$为大小为$D_k\times D_k\times M$的深度卷积核，核$\hat{k}$中的第m个滤波器应用于F的第m个通道，在对第m个通道进行滤波后生成输出特征图$\hat{G}$。

与普通卷积相比，DW-Conv将Flops从$m\times n\times k^2\times c^2$减少到$m\times n\times k^2\times c$。虽然DW-Conv（随后通常是PW-Conv）能有效减少Flops，但不能简单地取代常规卷积，因为这会导致精度显著下降。因此，在实际应用中，通常会增加DW-Conv的通道数或采用多分支架构来弥补精度的下降。

2 本文算法

2.1 整体架构设计

许多轻量级CNN模型具有相似的基本模块，但由于其整体架构存在差异，可能会出现不必要的冗余，这可以通过统一整合来解决。在包含不同基本模块的轻量级CNN模型中，架构设计缺乏一致性和统一标准，导致模型的整体结构各不相同，模糊了架构选择与性能提升之间的因果关系。因此，建立合理的整体架构尤为重要。参考CNN[5]和ViTs[15]的架构，本文提出了一种新的架构，称为基于结构重参数的DSCNN，图1展示了模型的整体架构。

Shalk：其为旨在使用具有多分支架构的过渡层，而不是相对较大的过渡层，在开始阶段捕捉更多细节。首先，创建并行特征提取网络，该网络

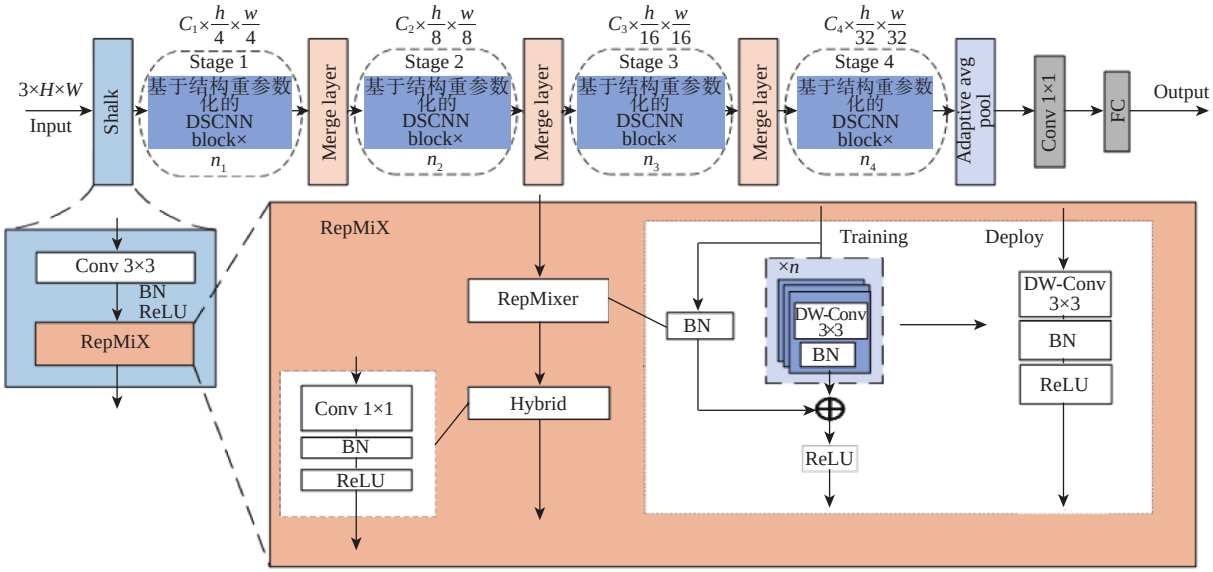


图 1　基于结构重参数化的 DSCNN 的整体架构

Fig. 1　Overall architecture of DSCNN based on structural reparameterization

由多个3×3 DW-Conv 和 BN 层组成，还采用了直连分支 (identity 分支) 来复用输入特征。其次，使用1×1卷积进行通道混合。

Stages: 基于结构重参数化的 DSCNN 共包括 4 个 Stages，每个阶段由重复的基于结构重参数化的 DSCNN 块 (包含 RepASO 算子的块) 堆叠组成，类似于 IRB。第 1～4 阶段的计算比例为 1∶2∶8∶2。本文每个阶段的基于结构重参数化的 DSCNN 块数指定为 [1, 2, 8, 2]。

Downsampling layer: 在第 1～3 阶段后，会增加一个单独的下采样层，该层使用步长为 2 的卷积核对前一阶段的输出特性进行卷积运算。该下采样层将输入分辨率降低一半，通道数增加一倍，之后是一个 BN 层以稳定训练。

Head: 其为模型的最后部分，位于第 4 阶段后。首先，使用自适应平均池化技术获得特征向量，然后，应用一组1×1卷积层进一步混合特征信息。最后，将输出的特征信息输入最后一个全连接分类器，产生最终输出。

2.2　特征提取模块 RepMiX

近年来，Transformer 架构在各种任务中取得了巨大成功。本设计从 ConvMixer[16] 和 RepViT[17] 中汲取灵感，设计了 RepMiX，其由空间混合算子 RepMixer 和通道混合算子 Hybrid 模块组成。RepMiX 能融合不同通道和空间位置间的特征信息，旨在从图像或特征数据中提取深层信息，实现多尺度特征融合[18]。

在模型训练模式下，使用具有多分支结构的转换层来捕捉更多细节，并在不同尺度上获得相应的输出结果。每个分支都会独立学习混合空间信息，并在一定范围内建立跨维交互。每个分支都包含一个3×3 DW-Conv 层和 BN 层。BN 函数表达式为

$$y_i=\gamma\frac{x_i-\mu}{\sqrt{\sigma^2+\varepsilon}}+\beta \tag{2}$$

式中：μ为均值；σ^2为方差；γ为缩放因子；β为偏移；ε为一个小数，防止分母为零。

这种设计可以确保每一层激活函数的输入在所有维度上都有相似的分布，有助于网络更全面地捕捉输入数据的结构和特征。并且，为了适应不同批大小和训练阶段的需要，在本文模型中，采用指数移动平均(exponential moving average, EMA)动态调整 BN 训练过程中均值和方差。在 BN 中，EMA 通过结合当前批次的统计量和之前累积的统计量来计算，用于更新累积的均值和方差，公式如下：

$$\begin{cases}\mu_{\text{running}}=\xi\mu_{\text{batch}}+(1-\xi)\mu_{\text{running}}\\ \sigma^2_{\text{running}}=\xi\sigma^2_{\text{batch}}+(1-\xi)\sigma^2_{\text{running}}\end{cases} \tag{3}$$

式中：μ_{batch}和σ^2_{batch}为当前批次数据的均值和方差；μ_{running}和$\sigma^2_{\text{running}}$为累积的均值和方差；$\xi$为平滑常数。

此外，RepMixer 运算器还包括可复用输入特征的直连分支。直连分支引入了跨层连接，允许前一个特征直接传递到后续层。这样，网络就能更好地捕捉低层与高层特征之间的信息流。最后，所有3×3 DW-Conv 分支的输出与直连分支的输出相加，并向前传播。

在部署模式下，为加快推理速度，合并卷积层和 BN 层。卷积计算公式如下：

$$\boldsymbol{y}=\boldsymbol{\omega}*\boldsymbol{X}+\boldsymbol{b} \tag{4}$$

式中：$\boldsymbol{\omega}$为卷积核的权重；$\boldsymbol{b}$ 为卷积层的偏置，通常在建立模型时设置为 0；$\boldsymbol{X}$为输入特征矩阵；$\boldsymbol{y}$ 为卷积层的输出，即 BN 层的输入。

将式(4)代入式(2),可得

$$\boldsymbol{y}_i=\boldsymbol{\gamma}\frac{\boldsymbol{\omega}*\boldsymbol{X}+\boldsymbol{b}-\boldsymbol{\mu}}{\sqrt{\boldsymbol{\sigma}^2+\boldsymbol{\varepsilon}}}+\boldsymbol{\beta}=\left(\boldsymbol{\gamma}\frac{\boldsymbol{\omega}}{\sqrt{\boldsymbol{\sigma}^2+\boldsymbol{\varepsilon}}}\right)*\boldsymbol{X}+\left(\boldsymbol{\gamma}\frac{\boldsymbol{b}-\boldsymbol{\mu}}{\sqrt{\boldsymbol{\sigma}^2+\boldsymbol{\varepsilon}}}+\boldsymbol{\beta}\right) \tag{5}$$

融合后,原来的卷积层和BN层可以用一个新的卷积操作来表示,其中,$\boldsymbol{\omega}_{\text{new}}=\boldsymbol{\gamma}\frac{\boldsymbol{\omega}}{\sqrt{\boldsymbol{\sigma}^2+\boldsymbol{\varepsilon}}}$,$\boldsymbol{b}_{\text{new}}=\boldsymbol{\gamma}\frac{\boldsymbol{b}-\boldsymbol{\mu}}{\sqrt{\boldsymbol{\sigma}^2+\boldsymbol{\varepsilon}}}+\boldsymbol{\beta}$。在训练阶段,BN层仍然需要计算和存储每个批次的均值、方差、尺度和偏移,训练时的内存需求和计算资源并没有减少;但在推理时,由于BN层被融合到之前的卷积中,减少了对均值、方差、尺度和偏移的计算和存储,不需要存储BN层的中间统计量,从而减少了推理时的计算负担和内存使用量。

利用结构重参数化技术,对图1中的RepMixer结构进行剪枝。在推理过程的多分支训练过程中获得的模型参数将用于单分支结构的参数化。该操作可以将多个分支融合为一个3×3卷积,直接处理输入特征图,加速模型推理。将多分支结构转换为单分支结构的数学表达式为

$$\text{bn}(\boldsymbol{M},\boldsymbol{\mu},\boldsymbol{\sigma},\boldsymbol{\gamma},\boldsymbol{\beta})_{:,i,:,:}=(W_{:,i,:,:}-\mu_i)\frac{\gamma_i}{\sigma_i}+\beta_i \tag{6}$$

$$\begin{cases}W'_{:,i,:,:}=\dfrac{\gamma_i}{\sigma_i}W_{:,i,:,:}\\ b'_i=\beta_i-\dfrac{\mu_i\gamma_i}{\sigma_i}\end{cases} \tag{7}$$

$$\begin{cases}W'_{\text{new}}=\displaystyle\sum_{j=1}^{n}W'_j\\ b'_{\text{new}}=\displaystyle\sum_{j=1}^{n}b'_j\end{cases} \tag{8}$$

式中:M为BN层的输入特征图;μ_i和σ_i分别为训练样本中单次迭代的均值和标准差;γ_i和β_i为网络学习参数;$W_{:,i,:,:}$为训练后的卷积核权重;$W'_{:,i,:,:}$和b'_i分别为转换后的权重和偏置;W'_{new}和b'_{new}分别为将多分支结构合并为单分支结构后的卷积核权重和偏置。

RepMiX利用混合算子进行信道混合,促进信道信息交换,并对上一步生成的特征图进行特征提取。整合从不同分支获得的特征图,并通过BN层和ReLU激活函数引入非线性映射,更有效地结合通道间的线性信息。

2.3 空间算子设计

构建学习特征信息的多分支空间算子,称为RepASO,如图2所示。RepASO的每个分支都有不同的作用,实现学习多尺度特征的目的[19]。

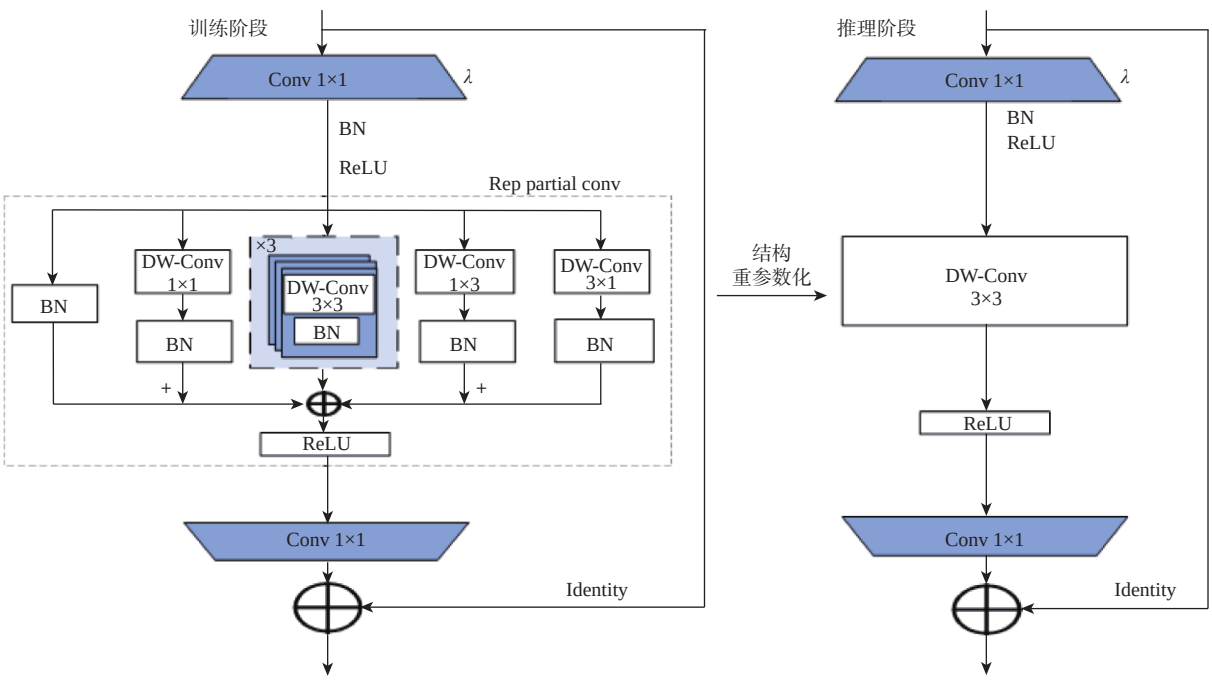


图2 重参数化非对称空间算子

Fig. 2 Reparameterized asymmetric spatial operator

卷积核的大小决定了其能观察到的输入数据的局部范围,也称为感受野。与5×5或7×7等较大的卷积核相比,3×3卷积核的感受野较小,但使用3×3卷积核可以显著减少参数。通过堆叠多层3×3卷积,网络可以学习局部和全局范围的特征。因此,本文选择参数和Flops较少的3×3卷积核。为补偿减少的特征通道,构建N个(默认情况下N=3)3×3 DW-Conv并行分支。此外,构建一个1×1 DW-Conv分支,组合1×1卷积和DW-Conv,提取中心位置[20]多个通道之间的非线性关系。为缓

解梯度消失，加速网络收敛，增加身份映射分支。

在处理具有明显水平或垂直结构的图像时，需要捕捉不同方向的特征，因此，构建水平1×3和垂直3×1 DW-Conv 并行分支，与3×3 DW-Conv 分支一起构成了非对称卷积结构。为提供更详细的解释，$\boldsymbol{F}'$为融合卷积核，$\boldsymbol{b}$ 为偏置，$\boldsymbol{F}$、$\overline{\boldsymbol{F}}$和$\hat{\boldsymbol{F}}$分别为3×3、1×3和3×1卷积核变换的输出，有

$$\boldsymbol{F}' = \frac{\gamma}{\sigma}\boldsymbol{F} \oplus \frac{\overline{\gamma}}{\overline{\sigma}}\overline{\boldsymbol{F}} \oplus \frac{\hat{\gamma}}{\hat{\sigma}}\hat{\boldsymbol{F}} \tag{9}$$

$$\boldsymbol{b} = -\frac{\gamma\mu}{\sigma} - \frac{\overline{\gamma}\overline{\mu}}{\overline{\sigma}} - \frac{\hat{\gamma}\hat{\mu}}{\hat{\sigma}} + \beta + \overline{\beta} + \hat{\beta} \tag{10}$$

式中：γ、$\overline{\gamma}$、$\hat{\gamma}$、β、$\overline{\beta}$和$\hat{\beta}$分别为 3 条卷积分支的网络学习参数；μ、$\overline{\mu}$、$\hat{\mu}$、σ、$\overline{\sigma}$和$\hat{\sigma}$分别为 3 条卷积分支训练后得到的均值和标准差。

在融合非对称卷积分支后，相当于一个新的卷积操作，其中，卷积核和偏置是多分支训练过程中获得的参数，用$\boldsymbol{y} = \boldsymbol{F}' * \boldsymbol{X} + \boldsymbol{b}$表示。图 3 直观展示了上述过程。结合上述结构，得到了一种名为 RepASO 的多分支算子，其具有七分支结构。7 个分支中的每个分支都配有一个 BN 层，每个分支的输出相加后进行前向传播。

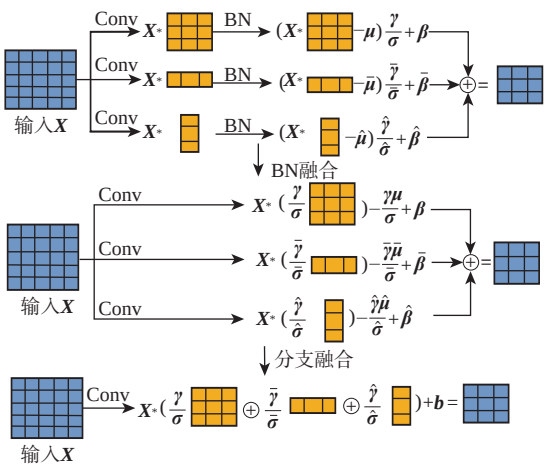


图 3　非对称卷积融合原理

Fig. 3　Principle of asymmetric convolution fusion

RepASO 是对 Inception[21] 结构的简单借鉴，其用参数和计算成本更低的 DW-Conv 取代了普通卷积，并引入非对称卷积来学习水平和垂直方向特征，但 RepASO 并不简单。根据卷积的可加性和结构重参数化原理，在推理过程中可以将这 7 个分支重参数化为一个 DW-Conv 分支，如图 2 所示，既加快了模型推理速度，又不损失模型学习能力，即 RepASO 结合了功能丰富的多分支拓扑结构和结构重参数化技术，前者在模型训练过程中提高模型学习能力，后者在模型推理过程中将多功能分支重参数化为单个分支，加快推理速度，而不仅仅是简单的 DW Inception 结构。

在 RepASO 中，每个分支都利用 DW-Conv 来有效减少 Flops，但其不能简单地取代传统的卷积，因为可能会导致模型精度下降。因此，在多分支结构之后，增加了一个 1×1 PW-Conv。通过这种 PW-Conv 操作，前一层输出特征图的通道之间进行信息交换，以弥补 DW-Conv 带来的精度损失。此外，PW-Conv 还能将前一层输出特性图的通道数量减半，从而降低模型参数和计算复杂度，提高模型效率。

2.4　合并基本块

在基于结构重参数化的 DSCNN 模型中，每个 stage 都由基于结构重参数化的 DSCNN 块堆叠而成，且每个 stage 后面都有一个合并层，该合并层的作用是进行图像块下采样操作，主要是使用大小为 2×2、步长为 2 的卷积核对上一层输出进行通道扩充和特征图压缩，作为模型的通道算子。合并基本块由空间算子和通道算子组成，即每个 stage 及其后面的合并层构成了基于结构重参数化的 DSCNN 模型的基本块。基本块是轻量级 CNN 的关键组成部分，负责提取和整合特征，通过串联和堆叠的方式形成整个模型的架构。

2.5　基于结构重参数化的 DSCNN

基于基于结构重参数化的 DSCNN 的整体架构（包括 RepMiX、RepASO 空间算子和合并层），本文提出一种名为基于结构重参数化的 DSCNN 的新型轻量级 CNN 模型，表 1 列出了基于结构重参数化的 DSCNN 的架构。实验表明，基于结构重参数化

表 1　基于结构重参数化的 DSCNN 网络架构

Table 1　Architecture of DSCNN based on structural reparameterization

名称	输出大小	模型	数量
Stalk	$\frac{h}{4}\times\frac{w}{4}$	3×3　40 RepMixer　40 FFN　40	1
Stage 1	$\frac{h}{4}\times\frac{w}{4}$	1×1　80 RepASO　80 1×1　40	1
Merge layer	$\frac{h}{8}\times\frac{w}{8}$	2×2 Conv　步长 2 BN　ReLU	1
Stage 2	$\frac{h}{8}\times\frac{w}{8}$	1×1　160 RepASO　160 1×1　80	2
Merge layer	$\frac{h}{16}\times\frac{w}{16}$	2×2 Conv　步长 2 BN　ReLU	1
Stage 3	$\frac{h}{16}\times\frac{w}{16}$	1×1　320 RepASO　320 1×1　160	8
Merge layer	$\frac{h}{32}\times\frac{w}{32}$	2×2 Conv　步长 2 BN　ReLU	1
Stage 4	7×7	1×1　640 RepASO　640 1×1　320	2
Head	1×1	7×7 AvgPool2d　320 1×1　1 280 FC　1 000	1

的DSCNN在参数量、浮点运算数和浮点运算速度方面优于现有的轻量级CNN模型，同时，保持了有竞争力的精度。

3 实 验

为验证模型的有效性，在Tiny-ImageNet-200、CIFAR-10和CIFAR-100数据集上进行了性能测试。为评估模型的泛化性能，通过实地拍摄创建了铝锭表面缺陷分类数据集，并与多个轻量级CNN模型进行比较。实验环境如表2所示。

表2 实验环境

Table 2 Experimental environment

序号	环境	版本
1	Use the system	Ubuntu 22.04 LTS
2	PyTorch	1.13.1
3	CUDA	11.6
4	GPU	NVIDIA RTX3090Ti
5	CPU	i9-12900KF
6	PyCharm	2022 Community
7	Python	3.9
8	RAM	32 GB
9	SSD	500 GB

3.1 数据集

使用Tiny-ImageNet-200、CIFAR-10和CIFAR-100数据集验证了本文模型的有效性。Tiny ImageNet-200数据集包含200个不同的类别，图像大小为224像素×224像素，共包含100 000幅图像，每类分别包含500幅训练图像、50幅验证图像和50幅测试图像。CIFAR-10数据集由10类彩色图像组成，每类包含6 000张图像，图像大小为32像素×32像素，包含60 000幅彩色图像，50 000张图像用于训练，其余用于验证。CIFAR-100数据集包含100个类别的60 000张彩色图像，图像大小为32像素×32像素，每个类别500幅图像用于训练，100幅图像测试。通过实地拍摄创建了铝锭表面缺陷分类数据集，用于对铝锭表面缺陷进行分类。铝锭表面缺陷分类数据集由约2 500张不同尺寸的图像组成，数据集按1∶9的比例分为测试集和训练集。

3.2 实验参数设置

对于Tiny-ImageNet-200、CIFAR-10和CIFAR-100图像分类任务，对数据进行随机裁剪、随机水平翻转和归一化增强[22]。使用AdamW(Adam with Weight Decay)优化器进行优化，权重衰减系数为0.005，动量为0.9。批量大小设置为240，模型训练时间为300个epoch，初始学习率设置为0.004，并采用余弦退火学习率调整策略，使用20个epoch预热，初始值为1×10^{-6}，最小学习率设为1×10^{-5}，调整公式为

$$l = l_{\min}+0.5(l_{\max}-l_{\min})\cdot(1+\cos(t_{\text{cur}}/T_{\text{total}}\pi)) \tag{11}$$

式中：l为当前训练步骤的学习率；$l_{\min}$为最小学习率；$l_{\max}$为初始学习率；t_{cur}为当前训练步骤的索引(从1开始)；T_{total}为训练步骤的总数。

使用软目标交叉熵作为损失函数，在传统的交叉熵损失函数中，目标标签以单击编码的形式表示。而使用软目标交叉熵时，目标标签是一种概率分布，具体计算公式为

$$\boldsymbol{L} = -\frac{1}{B_{\text{size}}}\sum_{j=1}^{B_{\text{size}}}\sum_{i=1}^{n}\boldsymbol{y}_{ij}\ln\hat{\boldsymbol{y}}_{ij} \tag{12}$$

式中：B_{size}为批量大小；$\boldsymbol{y}_{ij}$为目标标签的概率分布；$\hat{\boldsymbol{y}}_{ij}$为模型的预测输出。

对于每个类别，计算真实概率与预测概率之间的交叉熵，将所有类别的交叉熵相加，得出最终损失值。

在铝锭表面缺陷图像分类任务中，使用与上述相同的方法对捕获的照片数据进行缺陷增强操作。使用AdamW优化器进行优化，权重衰减系数为0.005，动量为0.9。批量大小设置为16，进行100次训练，初始学习率为0.004。采用余弦退火学习率调整策略来调整学习率，选择软目标交叉熵作为损失函数。

3.3 实验结果与分析

比较本文提出的基于结构重参数化的DSCNN模型与现有轻量级CNN模型的性能，包括MobileNeXt、FasterNet[23]、EdgeNeXt_S/XS/XXS、MobileVit_S/XS/XXS、GhostNetV1/V2、ConvNeXt_Tiny/Small、ShuffleNetV1/V2，及重量级CNN模型ResNet-50的性能。

表3列出了基于结构重参数化的DSCNN与其他轻量级CNN模型在Tiny-ImageNet-200数据集上的实验结果。本文目的是要达到模型精度与速度之间的均衡，因此，根据模型的Top-1精度，将实验数据分为4组，以比较近似精度的模型之间的速度和模型大小。通过设置不同的模型输入特征维度和基于结构重参数化的DSCNN块的堆叠数量，得到基于结构重参数化的DSCNN模型的4个不同版本DSCNN_T0/1/2/S；其中，DSCNN_T0/1/2每个stage的基于结构重参数化的DSCNN块数指定为[1, 2, 8, 2]，输入特征维度分别为40、64和96。DSCNN_S每个stage的基于结构重参数化的DSCNN块数指定为[1, 2, 13, 2]，输入特征维度为128。从表3可以得出，基于结构重参数化的DSCNN模型在各组GPU上实现了更高的吞吐量和浮点运算速度。

表 3　基于结构重参数化的 DSCNN 与其他轻量级 CNN 模型在 Tiny-ImageNet-200 数据集上的性能比较

Table 3　The performance comparison of DSCNN based on structural reparameterization and other light-weight CNN models on the Tiny-ImageNet-200 dataset

模型	参数量	浮点运算数	GPU吞吐量/(帧·s^{-1})	浮点运算速度/10^9 s^{-1}	Top-1精度/%
EdgeNeXt_XXS	1.3×10^6	0.20×10^9	5 268	29.62	63.26
MobileVit_XXS	1.3×10^6	0.37×10^9	3 472	27.21	63.81
FasterNet_T0	3.9×10^6	0.34×10^9	8 609	50.21	58.30
DSCNN_T0	3.7×10^6	0.32×10^9	6 912	44.02	60.52
EdgeNeXt_XS	2.34×10^6	0.41×10^9	3 894	32.16	66.56
MobileVit_XS	2.39×10^6	0.92×10^9	1 964	33.37	67.55
GhostNetV2	6.13×10^6	0.17×10^9	4 228	14.84	66.89
FasterNet_T1	7.61×10^6	0.85×10^9	5 364	56.17	64.35
DSCNN_T1	7.38×10^6	0.89×10^9	4 782	50.57	65.48
ConvNeXt_Tiny	28.59×10^6	3.47×10^9	1 143	43.15	71.96
ResNet-50	25.55×10^6	3.33×10^9	1 466	42.56	72.03
MobileNeXt	3.51×10^6	0.31×10^9	4 128	32.04	68.90
EdgeNeXt_S	5.59×10^6	0.97×10^9	2 598	41.93	71.61
MobileVit_S	5.64×10^6	1.79×10^9	1 535	41.07	70.30
FasterNet_T2	14.98×10^6	1.91×10^9	3 541	61.19	68.52
DSCNN_T2	14.19×10^6	1.90×10^9	2 981	55.15	69.60
Swin_T	28.3×10^6	4.51×10^9	600	36.49	78.11
DSCNN_S	29.1×10^6	4.38×10^9	1 435	63.25	78.90

从表 3 可以看出，DSCNN_T0/1/2 的浮点运算速度与 MobileViT_S/XS/XXS 模型大小相当，但在 GPU 上基于结构重参数化的 DSCNN 模型的吞吐量分别比 MobileViT_S/XS/XXS 高出 2.0 倍、2.4 倍和 1.9 倍。在基于 Top-1 精度水平的 4 组中，DSCNN_T0/1/2 在 GPU 上的吞吐量和浮点运算速度均高于其他轻量级网络，仅低于 FasterNet_T0/1/2。但值得注意的是，DSCNN_T0/1/2 的 Top-1 精度比 FasterNet_T0/1/2 高 1%～2%。此外，DSCNN_S 的 Top-1 精度达到 78.90%，与新兴的 swin_T 模型不相上下。此外，在 GPU 上，DSCNN_S 的推理吞吐量提高了 2.39 倍，浮点运算速度提高了 1.73 倍。表 3 的结果分析表明，基于结构重参数化的 DSCNN 在精度和速度之间取得了理想的平衡。图 4 显示，基于结构重参数化的 DSCNN 在精度和吞吐量之间的平衡方面明显优于其他方法。

在 CIFAR-10 和 CIFAR-100 数据集上进行对比实验，结果如表 4 所示。精度主要检验预测结果是否正确，计算公式为

$$P_{\text{recision}}=\frac{T_{\text{P}}}{T_{\text{P}}+F_{\text{P}}} \tag{13}$$

式中：T_{P}表示真阳性情况；F_{P}表示假阳性情况。

表 4 中，基于结构重参数化的 DSCNN 达到了 90.42% 的最佳精度，这验证了包括 RepASO 在内的所提组件的有效性和泛化能力。除了在 CIFAR-10 数据集上的精度比 MobileViT_XS 低 0.37% 之外，DSCNN_T0 在保持较低参数和浮点运算数（分别比 MobileViT_XS 低 1.67×10^6 和 0.6×10^9）的同时，表现优于其他方法，0.37% 的精度损失在可接受范围内。MobileVit_XXS 的浮点运算次数与 DSCNN_T0 相近，但在 CIFAR-10 和 CIFAR-100 数据集上的精度分别比 DSCNN_T0 低 6.11% 和 9.46%。此外，与 FasterNet_T0 相比，DSCNN_T0 的参数量和浮点运算次数相当或更低，但在 CIFAR-10 和 CIFAR-100 数据集上的精度分别提高了 1.01% 和 1.14%。结果表明，本文模型在降低复杂度的同时，保持了良好的表达能力和有竞争力的性能。

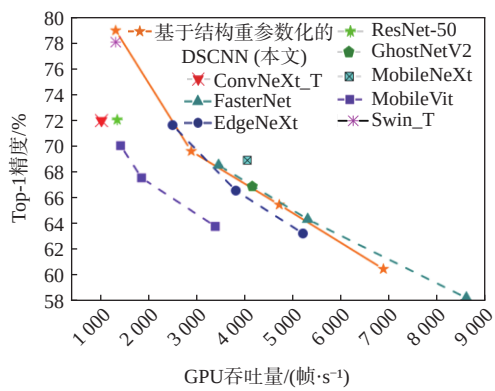


图 4　DSCNN 与其他轻量模型的性能对比

Fig. 4　Performance comparison between DSCNN and other lightweight models

铝锭表面缺陷分类数据集的样本图像如图 5

表 4 在 CIFAR-10 和 CIFAR-100 数据集上比较基于结构重参数化的 DSCNN 与其他轻量级 CNN 模型

Table 4 Comparing DSCNN based on structural reparameterization with other light-weight CNN models on CIFAR-10 and CIFAR-100 datasets

模型	参数量	浮点运算数	CIFAR-10 精度/%	CIFAR-100 精度/%
EdgeNeXt_XS	1.83×10^6	0.41×10^9	86.97	62.83
ShuffleNetV1	2.28×10^6	0.16×10^9	85.98	61.35
MobileNeXt	2.51×10^6	0.31×10^9	86.77	54.30
GhostNetV2	6.13×10^6	0.17×10^9	86.17	62.43
MobileVit_XXS	1.30×10^6	0.37×10^9	84.31	56.53
DSCNN_T0	3.45×10^6	0.32×10^9	90.42	65.99
FasterNet_T0	3.66×10^6	0.34×10^9	89.41	64.85
ConvNeXt_Tiny	27.97×10^6	3.47×10^9	89.07	64.76
ResNet-50	25.55×10^6	3.33×10^9	90.57	65.76
MobileVit_XS	5.12×10^6	0.92×10^9	90.79	65.17

和图 6 所示。在研究该数据集时，对基于结构重参数化的 DSCNN 与多个轻量级 CNN 模型的准确性、网络参数和浮点运算数进行了比较。图 7 直观展示了基于结构重参数化的 DSCNN 模型的性能。图中：圆圈大小表示浮点运算数。可以看出，与重量级 CNN 模型 ResNet-50 相比，基于结构重参数化的 DSCNN 的参数量和浮点运算数明显较低，而检测精度却高出约 6%。与 MobileNeXt、MobileVit-s、EdgeNeXt-s 和 GhostNetV2 等轻量级 CNN 模型相比，基于结构重参数化的 DSCNN 的检测精度分别提高 1.5%、3%、5.5% 和 6.6%。

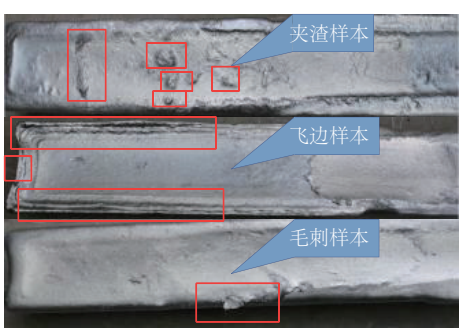


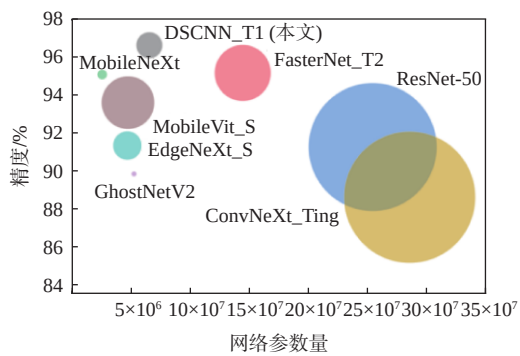
图 5 缺陷样本

Fig. 5 Defective sample

为直观地显示模型提取了哪些特征进行预测，使用铝锭表面缺陷的图像进行特征图可视化，使用 3 张不同类型缺陷的图像来生成热力图，如图 8 所示。图中：热力图中的颜色越鲜艳，说明网络对该区域越感兴趣，该区域越容易被观察到。与其他方法相比，基于结构重参数化的 DSCNN 模型能更好地覆盖样本图像中的缺陷区域，且更关注更突出、更容易观察到的位置（热力图中最亮的颜色）。即基于结构重参数化的 DSCNN 模型学会了利用目标对象

图 6 合格样本

Fig. 6 Qualified sample



圆圈大小代表 Flops

图 7 基于结构重参数化的 DSCNN 和多个轻量级 CNN 模型在铝锭表面缺陷分类数据集上的性能比较

Fig. 7 Performance comparison of DSCNN based on structural reparameterization and multiple light-weight CNN models on aluminum ingot defect detection dataset

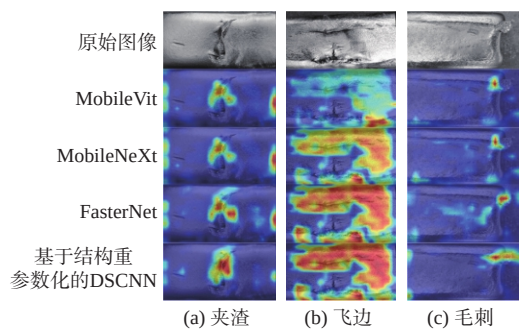


图 8 不同模型热力图的比较

Fig. 8 Comparison of heat maps of different models

区域的信息，能够关注具有不同特征和高辨别度的特征，并从中提取特征进行分类和回归。

3.4 消融实验

为探索 RepASO 空间算子的最佳配置，进行消融研究。表 5 列出了实验结果，表中的每一行代表在 RepASO 中包含不同核大小的卷积、身份映射、非对称卷积和 DW-Conv 分支的情况。实验 A、B 和 C 表明，使用大的卷积核并不能提高模型的性能。相反，其会对模型的性能产生负面影响。此外，实验 C、D、E、F、G、H、I 和 J 验证了添加非对

表 5　不同配置的 RepASO 在各个数据集上的实验结果比较

Table 5　Comparison of experimental results on various datasets with different configurations of RepASO

序号	卷积个数						身份映射	CIFAR-10 精度/%	CIFAR-100 精度/%	Tiny-ImageNet Top-1精度/%	铝锭表面缺陷分类数据集精度/%
	7×7	5×5	3×3	3×1	1×3	1×1					
A	1	0	0	0	0	0	0	88.18	61.75	58.92	92.89
B	0	1	0	0	0	0	0	88.24	62.80	59.07	92.95
C	0	0	1	0	0	0	0	88.44	63.24	59.20	93.15
D	0	0	3	0	0	0	0	88.85	63.24	59.32	93.44
E	0	0	3	0	0	1	0	88.87	63.34	59.37	93.54
F	0	0	3	0	0	0	1	88.57	63.13	59.42	93.47
G	0	0	3	1	1	0	0	88.88	63.60	59.29	93.34
H	0	0	3	1	1	1	0	88.89	63.68	60.31	94.14
I	0	0	3	1	1	0	1	88.78	63.51	59.32	94.38
J	0	0	3	1	1	1	1	90.42	65.99	60.52	96.50

称卷积分支和多个 3×3 DW-Conv 分支可以提高表示能力。

对 RepMiX 进行消融分析，获得了 RepMiX 的最佳结构形式，主要确定 RepMixer 中 DW-Conv 分支的最佳数量（用 r 表示），表 6 列出了实验结果。可以看出，每增加一条分支，训练阶段的模型参数量仅增大 0.02×10^6，当 r=4 时，模型开始出现过拟合。因此，r=3 为最佳，此时模型在 CIFAR-10 和 CIFAR-100 数据集上的准确率最高；3 条 DW-Conv 分支以增加较少的训练阶段模型参数为代价，实现了更高的识别准确率。这种设计允许用户为不同的推理设备任意选择合适的计算块，如果对识别准确率要求不高，也可以减少分支数以降低内存开销。

表 6　RepMixer 中不同 r 值的 RepMiX 在 CIFAR-10 和 CIFAR-100 数据集上的定量实验结果

Table 6　Quantitative experimental results of the RepMiX with different r values in RepMixer on CIFAR-10 and CIFAR-100 datasets

r	原始模型参数量	CIFAR-10精度/%	CIFAR-100精度/%
1	3.46×10^6	89.03	63.52
2	3.52×10^6	89.33	63.83
3	3.58×10^6	90.42	65.99
4	3.64×10^6	88.17	63.55

4　结　论

针对当前 CNN 模型较大、对设备计算能力要求较高和实时性较差的问题，本文提出了一种轻量化网络模型——基于结构重参数化的 DSCNN。

1）RepMiX 和 RepASO 在训练时的多分支并结构可融合不同信道和多尺度特征信息，并行架构避免了网络深度带来的过拟合问题。

2）在推理过程中利用训练保存的参数将 RepMiX 和 RepASO 重参数化为单分支，在保证原有性能的同时，简化模型结构，加速了模型推理。

3）模块设计时，结合 DS-Conv 替代普通卷积的思想，使用 DW-Conv 设计多分支结构，之后增加一个 PW-Conv，弥补了 DW-Conv 使用造成的精度损失，以本文使用的3×3卷积核来计算，模型参数量和计算量下降到使用普通卷积的 1/9 左右。

4）在 Tiny-ImageNet-200、CIFAR-10、CIFAR-100 数据集及自建铝锭表面缺陷分类数据集上，基于结构重参数化的 DSCNN 在保持相当精度的同时，参数和 Flops 更少，浮点运算速度更快，在速度和精度之间实现了很好的权衡。

参考文献（References）

[1] 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.

[2] DING X H, ZHANG X Y, HAN J G, et al. Scaling up your kernels to 31×31: revisiting large kernel design in cnns[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2022: 11953-11965.

[3] TAN M X, LE Q V. EfficientNetV2: smaller models and faster training[EB/OL]. (2021-06-23)[2024-01-05]. https://doi.org/10.48550/arXiv.2104.00298.

[4] LIU Z, MAO H, WU C Y, et al. A ConvNet for the 2020s[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2022: 11966-11976.

[5] YU W H, ZHOU P, YAN S C, et al. InceptionNeXt: when inception meets ConvNext[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2024: 5672-5683.

[6] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE/

CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 4510-4520.

[7] ZHANG J N, LI X T, LI J, et al. Rethinking mobile block for efficient attention-based models[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2024: 1389-1400.

[8] LIU S W, CHEN T L, CHEN X H, et al. More ConvNets in the 2020s: Scaling up kernels beyond 51×51 using sparsity[EB/OL]. (2023-03-03)[2024-01-07].https://doi.org/10.48550/arXiv.2207.03620.

[9] ZHANG X Y, ZHOU X Y, LIN M X, et al. ShuffleNet: an extremely efficient convolutional neural network for mobile devices[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 6848-6856.

[10] SUN K, LI M J, LIU D, et al. IGCV3: interleaved low-rank group convolutions for efficient deep neural networks[EB/OL]. (2018-07-20)[2024-01-08]. https://doi.org/10.48550/arXiv.1806.00178.

[11] GAO H Y, WANG Y, CAI L, et al. ChannelNets: compact and efficient convolutional neural networks via channel-wise convolutions[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(8): 2570-2581.

[12] DING X H, ZHANG X Y, MA N N, et al. RepVGG: making VGG-style ConvNets great again[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 13728-13737.

[13] DING X H, GUO Y C, DING G G, et al. ACNet: strengthening the kernel skeletons for powerful CNN via asymmetric convolution blocks[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2020: 1911-1920.

[14] VASU P K A, GABRIEL J, ZHU J, et al. MobileOne: an improved one millisecond mobile backbone[EB/OL]. (2023-03-28)[2024-01-10]. https://doi.org/10.48550/arXiv.2206.04040.

[15] LIU Z, HU H, LIN Y T, et al. Swin transformer V2: scaling up capacity and resolution[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2022: 11999-12009.

[16] TROCKMAN A, KOLTER J Z. Patches are all you need?[EB/OL]. (2022-01-24)[2024-01-12]. https://doi.org/10.48550/arXiv.2201.09792.

[17] WANG A, CHEN H, LIN Z J, et al. RepViT: Revisiting mobile CNN from ViT perspective[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2024: 15909-15920.

[18] 邱云飞, 张家欣, 兰海, 等. 融合张量合成注意力的改进 ResNet 图像分类模型[J]. 激光与光电子学进展, 2023, 60(6): 0610008.
QIU Y F, ZHANG J X, LAN H, et al. Improved ResNet image classification model based on tensor synthesis attention[J]. Laser & Optoelectronics Progress, 2023, 60(6): 0610008(in Chinese).

[19] 朱逢乐, 刘益, 乔欣, 等. 基于多尺度级联卷积神经网络的高光谱图像分析[J]. 吉林大学学报 (工学版), 2023, 53(12): 3547-3557.
ZHU F L, LIU Y, QIAO X, et al. Analysis of hyperspectral image analysis based on multi-scale cascaded convolutional neural network[J]. Journal of Jilin University (Engineering and Technology Edition), 2023, 53(12): 3547-3557 (in Chinese).

[20] 程小辉, 李钰, 康燕萍. 基于中间图特征提取的卷积网络双标准剪枝[J]. 计算机工程, 2023, 49(3): 105-112.
CHENG X H, LI Y, KANG Y P. Double standard pruning of convolution network based on feature extraction of intermediate graph[J]. Computer Engineering, 2023, 49(3): 105-112(in Chinese).

[21] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2015: 1-9.

[22] 赵利, 王雷全, 张俊三, 等. 基于双通道特征增强的高光谱图像分类[J]. 激光与光电子学进展, 2023, 60(12): 1210012.
ZHAO L, WANG L Q, ZHANG J S, et al. Hyperspectral image classification based on dual-channel feature enhancement[J]. Laser & Optoelectronics Progress, 2023, 60(12): 1210012(in Chinese).

[23] CHEN J R, KAO S H, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2023: 12021-12031.

Deep separable convolutional neural networks based on structural reparameterization

CHEN Hong[1]，YAN Jianguo[2]，YANG Hua[1]，ZHANG Jing[1]，LI Wei[3]，YANG Jing[1, 4, *]

(1. Electrical Engineering College，Guizhou University，Guiyang 550025，China；
2. China Power Construction Group Guizhou Engineering Co.，Ltd，Guiyang 550025，China；
3. Engineering Research Center of Guizhou Higher Education Institutions，Guiyang 550025，China；
4. Guizhou University，China Guizhou Provincial Key Laboratory of Internet + Intelligent Manufacturing，Guiyang 550025，China)

Abstract：A new lightweight convolutional neural network (CNN) model called the deep separable convolutional neural network (DSCNN) based on structural reparameterization is proposed, aiming at the single-branch deep convolutional approach used in the majority of the current CNN models, which not only affects the expressive ability of the model but also occupies a large number of parameters and Flops. Firstly, the feature extraction module (reparameterization MiXer, RepMiX) of the model can fuse information between different channels and spatial locations, realizing multi-scale feature fusion. Secondly, the reparameterized asymmetric spatial operator (RepASO) in the DSCNN learns different channel feature information through a multi-branch structure with different functions, which improves the model feature learning ability; meanwhile, both RepMiX and RepASO combine the structural reparameterization technique and the idea of depth wise separable convolution (DS-Conv) to realize structural decoupling in the training and inference phases, which accelerates the model inference while reducing the model parameters and Flops. Finally, comparative experiments are carried out on the Tiny-imagenet-200, CIFAR-10, and CIFAR-100 datasets in addition to a self-constructed dataset for the classification of aluminum ingot surface defects. The experimental results demonstrate that the DSCNN maintains competitive accuracy while achieving higher floating-point speeds.

Keywords：convolutional neural networks；structural reparameterization；depthwise seperable convolution；asymmetric convolution；structural decoupling

Received：2024-05-07；**Accepted**：2024-06-25；**Published Online**：2024-07-03 15：00

URL：link.cnki.net/urlid/11.2625.V.20240703.0858.001

Foundation items：National Natural Science Foundation of China (61640014, 61963009); Innovation Group of Guizhou Education Department (Qianjiaohe KY[2021]012); Guizhou Provincial Science and Technology Projects (Qiankehe Zhicheng [2022]Yiban017, Qiankehe Zhicheng [2023]Yiban411, Qiankehe Zhicheng [2024]Yiban051); Engineering Research Center of Guizhou Education Department (Qianjiaoji[2022]040, Qianjiaoji[2022]043); Science and Technology Project of Power Construction Corporation of China, Ltd (DJ-ZDXM-2020-19, DJ-ZDXM-2022-44); Guizhou Dual Carbon Research Institute Open Subjects (DCRE-2023-13); China Aluminum International Engineering Corporation Limited Science and Technology Major Project (CJ2022JS-ZD02)

*** Corresponding author**. E-mail：jyang7@gzu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0284

基于模仿学习的机动目标拦截强化学习制导律

任乐亮[1,2]，鲜勇[1,2]，刘振宇[1]，张大巧[3,*]，李冰[3]，李少朋[1,4]

(1. 火箭军工程大学 导弹工程学院，西安 710025；2. 跨域飞行交叉技术实验室，绵阳 621000；
3. 火箭军工程大学 作战保障学院，西安 710025；4. 清华大学 自动化系，北京 100084)

摘 要：机动突防技术的发展对拦截制导律性能提出了更高的要求。为提高机动目标拦截概率、降低能量消耗、提升算法鲁棒性，提出了“基于模仿学习 (IL) 末制导律模型→基于强化学习的末制导律进化模型”的梯次智能制导框架。以拦截碰撞三角为基础，建立机动目标和拦截器三维不确定对抗模型；采用 IL 方法挖掘比例导引 (PNG) 规律，为后续强化学习制导律学习提供良好的初始策略；建立了马尔可夫决策模型，并设置包含能量消耗的过程奖励和包含“过渡段”的“软”终端奖励模型，利用近端策略优化 (PPO) 算法充分探索高性能拦截策略。蒙特卡罗仿真结果表明：新型制导律鲁棒性好，稳定性高，实现了比传统制导算法更高的拦截概率和更低的能量消耗，且单次决策耗时仅 0.32 ms，具有一定的工程应用价值。

关 键 词：模仿学习；深度强化学习；机动目标拦截；近端策略优化；碰撞三角
中图分类号：V448.133
文献标志码：A **文章编号**：1001-5965(2026)06-2156-16

高速目标拦截制导律设计一直是导弹防御领域的研究热点。比例导引 (proportional navigation guidance, PNG)[1-2] 及其改进型具有结构简单、易实现的优点，广泛应用于拦截过程。PNG 假设目标无机动或者固定机动[3]，在这种线性假设下，具有较好的拦截性能[4]，但在应对采用机动突防技术[5-7]的目标时，其拦截性能显著下降[8]。增广比例导引 (augmented proportional navigation, APN)[1] 将目标机动加速度叠加到 PNG 制导指令上，在一定程度上提高了应对机动目标的拦截性能，但由于 APN 需要实时获取目标加速度信息，降低了其工程应用可能性。

针对机动目标拦截问题，学者将非线性控制[9-11]、最优控制[12-13] 和微分对策理论[14-16] 应用于制导领域，提出了一系列新型制导律。文献 [9-11] 将滑模控制理论应用于拦截机动目标制导律设计，并针对滑膜控制的奇异性[9-10]、抖振[11] 等缺点展开了研究。文献 [17-18] 分别基于最优控制研究了能量最优和时间最优制导律。利用微分对策理论，文献 [14-16] 研究了拦截弹和目标攻防对抗问题，提出了最优制导方案。

虽然前述新型制导律在拦截机动目标仿真中具有较好的效果，但仍然有一定缺点：① 大多算法对场景进行了线性化处理，因此，实际应用时的最优性难以保证[8]；② 基于最优控制的制导律需要估计目标运动轨迹、剩余飞行时间和拦截点[19]，而这些信息在实际拦截过程中难以获取；③ 基于微分对策的制导律计算复杂度高，而弹上计算能力难以满足[8]。

作为人工智能技术重要分支的深度强化学习 (deep reinforcement learning, DRL)[20-23] 在序贯决策方面应用广泛[24-27]。DRL 不需要任何监督信息，而是

收稿日期：2024-05-06；**录用日期**：2024-09-27；**网络出版时间**：2025-01-06 16:11
网络出版地址：link.cnki.net/urlid/11.2625.V.20250106.1129.001
基金项目：国家自然科学基金 (62103432)；陕西省高校科协青年人才托举计划 (20210108)；跨域飞行交叉技术实验室开放基金 (2024-KYKF-4004)
* **通信作者**. E-mail：zhangdq2012@pku.org.cn

引用格式：任乐亮，鲜勇，刘振宇，等. 基于模仿学习的机动目标拦截强化学习制导律 [J]. 北京航空航天大学学报，2026，52（6）：2156-2171. REN L L，XIAN Y，LIU Z Y，et al. Reinforcement learning guidance law for maneuvering target interception based on imitation learning [J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（6）：2156-2171（in Chinese）.

通过试错机制与环境进行不断交互, 实现复杂环境中最大化累积奖励, 进而学习到最优策略[28-29]。由于 DRL 可以直接将状态映射为制导指令, 实现端到端决策, 成为制导控制领域的研究热点, 如动力下降制导[30-32]、空间探索制导[33]、导弹制导[34-35]和交会制导[36]。

DRL 的成功应用和相关技术的不断发展, 同样为拦截制导律设计提供了新的思路, 大量学者就此开展了研究。文献 [19] 研究了单一固定场景下的 DRL 制导问题, 启发了相关研究的发展。文献 [37-38] 基于 Q 学习方法, 将动作空间进行离散化, 得到了简单场景下的强化学习制导律。文献 [39] 研究了 DRL 在导弹制导中的应用问题, 并提出了 2 种训练模式: ① 智能体从零开始学习制导指令; ② 基于先验知识学习, 即学习经典制导律的导航比。同时, 指出前者在某些初始状态下无法保证拦截目标, 而后者则可提高学习效率。文献 [40-41] 针对机动目标拦截问题, 同样以 PNG 导航比为动作, 分别研究了基于深度确定性策略梯度(deep deterministic policy gradient, DDPG)的强化学习制导律和基于 Q 学习的强化学习制导律。文献 [42] 指出, 虽然以导航比作为动作提高了智能体训练效率, 但该制导律也存在 PNG 的固有缺点。

文献 [43] 借鉴 PNG 通过抑制视线角速度完成拦截的思想, 设计了考虑视线角速度和终端脱靶量的奖励函数, 采用深度 Q 网络在三维对抗场景下对智能体进行训练。文献 [44] 进一步采用 DDPG 训练智能体, 实现了更好的拦截性能。文献 [45] 应用 DDPG 研究了大气层内拦截器的一体化制导控制问题, 直接以俯仰和偏航舵偏控制指令作为智能体输出, 并在奖励函数设计中加入收敛因子以加快收敛速度。文献 [46] 在 DRL 的目标函数中加入最小化过载正则项, 避免拦截器产生多余机动, 并采用双阶段课程学习方法对智能体进行训练, 降低了学习难度。文献 [47-48] 基于元强化学习研究了非理想环境下大气层外机动目标拦截问题, 提高了应对不确定场景下算法的适应性。文献 [3, 8, 49-50] 基于近端策略优化(proximal policy optimization, PPO)[22]、双延迟深度确定性策略梯度(twin delayed DDPG, TD3)[51]和信赖域策略优化(trust region policy optimization, TRPO)[52]系统研究了应用强化学习解决机动目标拦截问题, 通过能量消耗、视线角速度零化等因素构建奖励函数, 加快了学习收敛速度, 实现了更小的脱靶量和更稳定的拦截。

尽管前述文献已较全面地研究了强化学习在拦截制导律中的应用问题, 但仍然具有一定的提升空间。强化学习智能体由零开始学习, 在复杂、长时序场景下的学习效率不高。特别地, 对于拦截问题而言, 目标和拦截双方均处于高速运动状态, 拦截方一个时间步的失误都将增加拦截失败的可能性; 同时, 由于只有在仿真结束时才能得知是否拦截成功, 因此, 存在稀疏奖励问题, 进而导致智能体无法探索到有效拦截策略。如何提高训练效率仍然是研究的难点。文献 [3, 8, 49-50] 等采用奖励塑形的方法, 加入视线角速度等过程量引导智能体探索有效拦截目标策略, 取得了较好的效果。但上述方法一方面需要较强的专业知识设计合理的过程奖励, 对于不同场景的适应性不足, 另一方面需要进行大量调参以得到较好的超参数, 需要耗费一定的时间成本。

本文在上述文献的启发下, 参考人类逐渐学习过程及 Alpha Go[24] 和 AlphaStar[27] 中智能体均在一定经验基础上开展学习, 进一步结合行为克隆的模仿学习(imitation learning, IL)方法建立了一种拦截机动目标的强化学习制导律, 旨在实现能量消耗最少且拦截概率最大。其主要思路分为 2 步: ① 以 PNG 为基础生成大量专家样本, 智能体采用模仿学习方法学习该专家策略; ② 智能体进一步利用强化学习算法, 逐步学习到拦截机动目标的策略。本文主要贡献总结如下:

1) 以拦截碰撞三角为基础, 设计了不确定场景下拦截弹和目标初始状态求解方法, 建立了非理想状态下拦截弹和目标三维对抗场景。

2) 提出了“基于模仿学习末制导律模型→基于强化学习的末制导律进化模型”的梯次智能制导框架, 为拦截制导律设计提供了新思路。

3) 通过模仿学习挖掘 PNG 制导律规律, 为强化学习智能体提供良好的初始策略, 有效提高了智能体学习效率, 并降低了奖励函数设计难度。

4) 设计了包含能量消耗的过程奖励和“过渡段”的“软”终端奖励, 利用 PPO 算法对拦截策略进行求解, 实现了低能量消耗下高概率拦截。

1　基本模型

1.1　动力学模型

针对大气层外目标拦截问题, 可忽略空气动力影响, 在地心惯性(earth-centered inertial, ECI)坐标系下建立目标和拦截器的动力学方程(为方便描述, 如不特指下文均采用列矢量形式):

$$\begin{cases} \dot{\boldsymbol{X}}_{\mathrm{s}} = \boldsymbol{V}_{\mathrm{s}} \\ \dot{\boldsymbol{V}}_{\mathrm{s}} = \boldsymbol{a}_{\mathrm{s}}^{\mathrm{P}} + \boldsymbol{a}_{\mathrm{s}}^{\mathrm{G}} \end{cases} \tag{1}$$

式中：$\boldsymbol{X}_s$和$\boldsymbol{V}_s$分别为ECI坐标系下的位置和速度；$\boldsymbol{a}_s^P$和$\boldsymbol{a}_s^G$分别为目标或拦截器轨控发动机产生的加速度和引力加速度，可由式(2)计算得到：

$$\begin{cases}\boldsymbol{a}_s^P = \boldsymbol{P}/m_w \\ \boldsymbol{a}_s^G = -\dfrac{\mu}{r^3}\left(1+\dfrac{3}{2}J_2\left(\dfrac{a_e}{r}\right)^2\left(1-5\sin^2\varphi_s\right)\right)\boldsymbol{X}_s - \\ \qquad \dfrac{3\mu}{r^2}J_2\left(\dfrac{a_e}{r}\right)^2\sin\varphi_s\dfrac{\boldsymbol{\omega}_e}{\|\boldsymbol{\omega}_e\|}\end{cases} \tag{2}$$

式中：$\boldsymbol{P}$为轨控发动机推力矢量；m_w为质量；μ为地球引力常数；$r=\|\boldsymbol{X}_s\|$为地心距；J_2为地球形状动力学系数；a_e为地球长半轴；φ_s为地心纬度；$\boldsymbol{\omega}_e=[0,0,|\boldsymbol{\omega}_e|]^T$为ECI坐标系下的地球自转角速度矢量。

1.2 三维PNG

PNG广泛应用于拦截过程，其特点是简单、易于实现，且以目标无机动假设为基础。三维PNG的矢量形式可表示为

$$\boldsymbol{a}_{PN,0} = -N_{PN}\boldsymbol{V}_D \times \boldsymbol{\omega}_{DM} \tag{3}$$

式中：$\boldsymbol{a}_{PN,0}$为PNG的加速度指令；N_{PN}为导航比；$\boldsymbol{V}_D$为拦截器的速度矢量；$\boldsymbol{\omega}_{DM}$为视线角速度矢量，实际制导过程该矢量由导引头实时测出，仿真过程中可由式(4)计算得到：

$$\boldsymbol{\omega}_{DM} = \frac{\boldsymbol{X}_{DM}\times\boldsymbol{V}_{DM}}{\boldsymbol{X}_{DM}\boldsymbol{X}_{DM}} \tag{4}$$

式中：$\boldsymbol{X}_{DM}=\boldsymbol{X}_M-\boldsymbol{X}_D$为目标相对于拦截器的位置矢量，$\boldsymbol{X}_M$和$\boldsymbol{X}_D$分别为目标和拦截器的位置矢量；$\boldsymbol{V}_{DM}=\boldsymbol{V}_M-\boldsymbol{V}_D$为目标相对于拦截器的速度矢量，$\boldsymbol{V}_M$为目标的速度矢量。

考虑到拦截器过载存在最大值限制，因此，PNG指令修正为

$$\boldsymbol{a}_{PN} = \min\left\{\left\|\boldsymbol{a}_{PN,0}\right\|, a_{max}\right\}\cdot f_{unit}\left(\boldsymbol{a}_{PN,0}\right) \tag{5}$$

式中：a_{max}为拦截器最大过载；$f_{unit}(\cdot)$的功能为求解输入矢量对应的单位矢量。

1.3 碰撞三角模型

零控脱靶量描述了拦截器和目标在无机动情况下的拦截脱靶量。通常，拦截系统通过求解拦截弹发射诸元，可保证不考虑任何拦截误差时的零控脱靶量为0。碰撞三角模型给出了拦截器处于末制导状态时零控脱靶量为0的基本条件，即拦截器和目标状态满足碰撞三角关系时，拦截器可以成功拦截目标，如图1所示。

图1中，所有变量均表示在惯性坐标系下，$\boldsymbol{V}_{D,\perp}$和$\boldsymbol{V}_{M,\perp}$分别表示$\boldsymbol{V}_D$和$\boldsymbol{V}_M$垂直于$\boldsymbol{X}_{DM}$的分量；$\boldsymbol{V}_{D,\parallel}$和$\boldsymbol{V}_{M,\parallel}$分别表示$\boldsymbol{V}_D$和$\boldsymbol{V}_M$平行于$\boldsymbol{X}_{DM}$的分量。由图1可知，在拦截器和目标不受外力或所受外力相同情况下，当$\boldsymbol{V}_{M,\perp}=\boldsymbol{V}_{D,\perp}$时，$\|\boldsymbol{X}_{DM}\|=0$。由此可知，

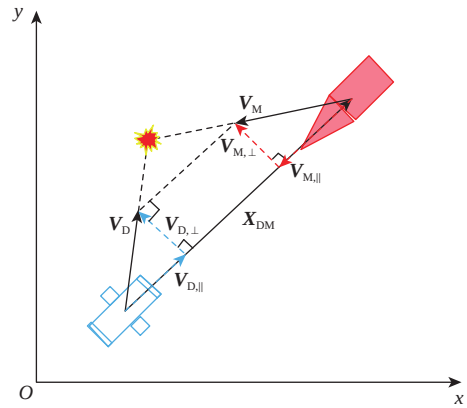


图1 碰撞三角示意

Fig. 1 Collision triangle diagram

当满足$\boldsymbol{V}_{D,\perp}=\boldsymbol{V}_{M,\perp}$关系时，拦截器零控状态下拦截目标的脱靶量为0。

1.4 深度强化学习

在强化学习中，智能体通过试错机制与环境进行不断交互，实现复杂环境中最大化累积奖励。DRL是将具有感知能力的深度学习与具有决策能力的强化学习结合起来，实现“端到端”的智能决策。交互试错过程可由马尔可夫决策过程描述，形成如图2所示的强化学习基本框架。其基本过程为：智能体感知环境状态$\boldsymbol{s}_t$，并依据当前策略π给出动作a_t，环境执行该动作得到新状态$\boldsymbol{s}_{t+1}$。同时，智能体会根据环境给出的奖励r_t计算长期回报来更新优化策略π。

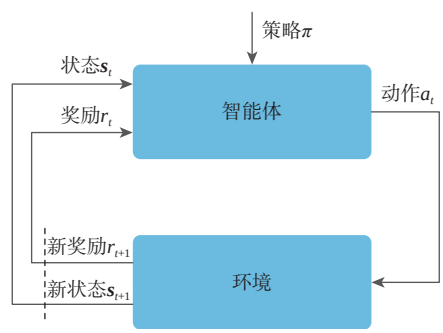


图2 强化学习智能体与环境交互示意

Fig. 2 Schematic diagram of reinforcement learning agent interaction with environment

强化学习的奖励函数在促进策略学习中具有重要作用，通常与所处状态$\boldsymbol{s}_t$和执行的动作a_t有关，可表示为$r_t(\boldsymbol{s}_t,a_t)$。定义轨迹$\tau$为智能体采用策略$\pi$形成的状态-动作序列，即$\tau=(\boldsymbol{s}_0,a_0,\boldsymbol{s}_1,a_1,\cdots)$。进而可定义轨迹上的累积奖励$R(\tau)$(即回报)：

$$R(\tau) = \sum_{t=0}^{T}\gamma^t r_t \tag{6}$$

式中：t为决策时间步；T为最大时间步；$\gamma\in[0,1]$，为

折扣因子，该因子可平衡当下奖励和未来奖励的重要程度；r_t 为 t 时刻的奖励值。

强化学习的目标是最大化期望回报，则学习最优策略π^*可表示为

$$\pi^* = \arg\max_{\pi} E_{\tau\sim\pi} R(\tau) \tag{7}$$

式中：$E_{\tau\sim\pi}R(\tau)$为轨迹τ服从策略π情况下回报的期望。

强化学习引入了状态值函数和动作值函数来评估策略好坏，并作为策略改进的基础。状态值函数$V_\pi(\boldsymbol{s})$定义为智能体以策略π为指导、处于状态$\boldsymbol{s}$时的期望回报，表示为

$$V_\pi(\boldsymbol{s}) \doteq E_{\tau_t\sim\pi}[R(\tau_t)|\boldsymbol{s}_t=\boldsymbol{s}] \tag{8}$$

式中：$\boldsymbol{s}_t$为时间步t的状态；τ_t为以状态$\boldsymbol{s}_t$为起点形成的轨迹。

动作值函数$Q_\pi(\boldsymbol{s},a)$定义为智能体以策略π为指导、处于状态$\boldsymbol{s}$采取动作a时的期望回报，表示为

$$Q_\pi(\boldsymbol{s},a) \doteq E_{\tau_t\sim\pi}[R(\tau_t)|\boldsymbol{s}_t=\boldsymbol{s},a_t=a] \tag{9}$$

式中：a_t为时间步t的动作。

$V_\pi(\boldsymbol{s})$和$Q_\pi(\boldsymbol{s},a)$可以为智能体学习最优策略提供优化方向，在 DRL 的建模过程中，通常采用神经网络对其进行拟合。

在强化学习中，有时并不需要描述智能体动作的绝对好坏，而只需要知道一个动作相对于其他动作的好坏。因此，引入强化学习中的重要概念，即优势函数，其表达式为

$$A_\pi(\boldsymbol{s},a) = Q_\pi(\boldsymbol{s},a) - V_\pi(\boldsymbol{s}) \tag{10}$$

通常，可采用广义优势估计（generalized advantage estimation, GAE）[53] 方法对优势函数进行估计：

$$A_{\pi,t}^{\mathrm{GAE}}(\boldsymbol{s},a) = \sum_{l=0}^{\infty}(\gamma\lambda)^l\delta_{t+l}^V \tag{11}$$

$$\delta_t^V = (r_t + \gamma V_\pi(\boldsymbol{s}_{t+1}) - V_\pi(\boldsymbol{s}_t)) \tag{12}$$

式中：$A_{\pi,t}^{\mathrm{GAE}}(\boldsymbol{s},a)$为 GAE 的估计值；$\delta_t^V$为优势函数的时序差分误差形式；$\lambda$为 GAE 的超参数。

2　场景设计

2.1　基本假设

本文主要研究拦截大气层外机动目标的制导指令生成问题，为方便研究，给出如下假设：① 在大气层外飞行阶段，不考虑拦截器和目标绕质心运动及姿态控制问题，仿真中采用六自由度瞬时平衡模型；② 将拦截器制导系统视为无延迟理想系统；③ 拦截器探测噪声服从正态分布。

2.2　训练场景

2.2.1　场景示意图

交互环境的设计应依据交战场景、双方性能参数进行设计，并尽可能包含交战双方所有可能的飞行状态。本文主要针对拦截器末制导问题进行研究，因此，交互环境的初始状态为拦截器转为末制导时交战双方的飞行状态，该飞行状态与拦截弹标准弹道紧密相关。拦截弹标准弹道指的是，拦截弹发射诸元对应的标准弹道，是一条能够命中预定拦截点的弹道。因此，交互环境初始状态的交战双方飞行状态应在碰撞三角对应状态附近波动，进而设计训练场景如图 3 所示。

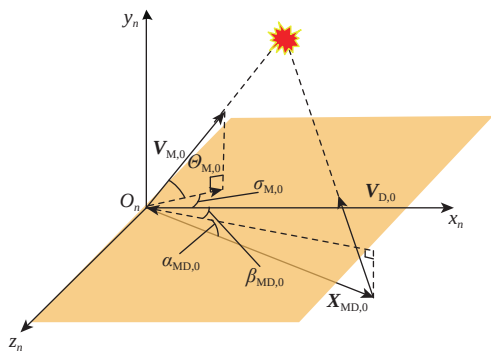


图 3　训练场景示意

Fig. 3　Training scenario diagram

图 3 描述了拦截器和目标飞行器初始三维相对运动关系，其中，变量含义如下：坐标系$x_nO_nz_n$为以目标飞行器为原点建立的北东坐标系；$\Theta_{M,0}$和$\sigma_{M,0}$分别为目标飞行器的当地速度倾角和速度方位角；$\boldsymbol{X}_{MD,0}$为拦截器相对于目标飞行器的相对位置矢量；$\alpha_{MD,0}$为$\boldsymbol{X}_{MD,0}$与平面$x_nO_nz_n$的夹角；$\beta_{MD,0}$为$\boldsymbol{X}_{MD,0}$在平面$x_nO_nz_n$投影与O_nx_n轴夹角。在训练过程中，两者初始运动状态在一定初始条件范围内随机生成，初始条件包括$\|\boldsymbol{V}_{M,0}\|$、$\Theta_{M,0}$、$\sigma_{M,0}$、$\alpha_{MD,0}$、$\beta_{MD,0}$和$\|\boldsymbol{V}_{D,0}\|$。需要说明的是，$\|\boldsymbol{X}_{MD,0}\|$和拦截器速度不在初始条件中的原因为：① 考虑到拦截器探测目标飞行器的距离一定，因此，$\|\boldsymbol{X}_{MD,0}\|$为定值；② 拦截器速度应该在碰撞三角附近，因此，交互环境仅对拦截器速度大小进行随机，而速度各分量在碰撞三角解算速度分量基础上加入随机量生成。

考虑到$\boldsymbol{X}_{MD,0}$与目标飞行器飞行方向存在一定约束，因此，引入反映$\boldsymbol{X}_{MD,0}$偏离目标飞行器飞行方向程度的 2 个变量$\Delta\alpha_{MD,0}$和$\Delta\beta_{MD,0}$，并满足如下关系：

$$\begin{cases}\alpha_{MD,0} = \Theta_{M,0} + \Delta\alpha_{MD,0}\\ \beta_{MD,0} = \sigma_{M,0} + \Delta\beta_{MD,0}\end{cases} \tag{13}$$

式中：$\Delta\alpha_{MD,0}$和$\Delta\beta_{MD,0}$为反应$\boldsymbol{X}_{MD,0}$偏离目标飞行器飞行方向程度的变量。

设置训练的初始状态边界范围如表 1 所示。

表1 训练场景参数边界

Table 1 Training scenario parameter boundaries

边界	$\|\boldsymbol{V}_{\mathrm{M},0}\|$/(m·s^{-1})	$\Theta_{\mathrm{M},0}$/(°)	$\sigma_{\mathrm{M},0}$/(°)	$\Delta\alpha_{\mathrm{MD},0}$/(°)	$\Delta\beta_{\mathrm{MD},0}$/(°)	$\|\boldsymbol{V}_{\mathrm{D},0}\|$/(m·s^{-1})
最小值	6 500.0	−5.0	0.0	−5.0	−10.0	6 000.0
最大值	7 000.0	5.0	90.0	−1.0	10.0	6 500.0

2.2.2 目标飞行器机动模型

由于拦截器捕获目标飞行器并拦截目标的时间很短，同时，受限于目标飞行器机动性能，其机动形式不宜太复杂，且工程上无法执行过于复杂的机动动作。单次方波机动形式具有形式简单、易实现、突防性能好的特点，被广泛应用于飞行器机动突防的研究中。同时，为增加拦截器拦截目标样本的丰富性，提高智能体泛化能力，随机生成目标飞行器机动时机和方向，进而得到目标飞行器机动模型为

$$\boldsymbol{a}_{\mathrm{M}}=\begin{cases}N_{\mathrm{m}}g\boldsymbol{e}_{\mathrm{M}} & T_{\mathrm{m},0}\leqslant t\leqslant T_{\mathrm{m},0}+\Delta T_{\mathrm{m}}\\ \boldsymbol{0} & \text{其他}\end{cases} \tag{14}$$

$$T_{\mathrm{m},0}\sim U\left(T_{\mathrm{m},0}^{\min},T_{\mathrm{m},0}^{\max}\right) \tag{15}$$

式中：N_{m}为目标飞行器机动过载；$T_{\mathrm{m},0}$和ΔT_{m}分别为机动开始时刻和机动时长；$\boldsymbol{a}_{\mathrm{M}}$为目标飞行器机动加速度矢量；$T_{\mathrm{m},0}^{\min}$和$T_{\mathrm{m},0}^{\max}$分别为目标飞行器可机动开始时刻的最小值和最大值；$\boldsymbol{e}_{\mathrm{M}}$为目标飞行器机动方向的单位矢量。

由于目标沿拦截器、目标飞行器连线的垂直方向施加轨控推力，对拦截器拦截效果影响最大。同时，考虑到目标轨控发动机可产生沿弹体横法向上共4个方向的推力[54]，因此，假定$\boldsymbol{e}_{\mathrm{M}}$垂直于拦截器、目标飞行器连线，且$\boldsymbol{e}_{\mathrm{M}}$在空间中指向限定为4个。训练过程中，随机选择上述4个方向中的1个作为$\boldsymbol{e}_{\mathrm{M}}$指向。

2.2.3 交互环境初始化

交互环境初始化主要完成拦截器和目标飞行器初始状态的解算。由于目标飞行器初始位置$\boldsymbol{X}_{\mathrm{M},0}$、速度$\boldsymbol{V}_{\mathrm{M},0}$和拦截器初始位置$\boldsymbol{X}_{\mathrm{D},0}$可在飞行状态范围内随机采样得到，因此，交互环境初始化的关键在于拦截器初始速度的求解。随机采样得到拦截器速度大小$|\boldsymbol{V}_{\mathrm{D},0}|$后，根据碰撞三角模型可得满足零控脱靶量为0的拦截器速度矢量$\boldsymbol{V}_{\mathrm{D},0}^{\mathrm{ct}}$为

$$\boldsymbol{V}_{\mathrm{D},0}^{\mathrm{ct}}=\frac{\boldsymbol{X}_{\mathrm{DM},0}\boldsymbol{V}_{\mathrm{M},0}}{\|\boldsymbol{X}_{\mathrm{DM},0}\|^{2}}\boldsymbol{X}_{\mathrm{DM},0}+\sqrt{V_{\mathrm{D},0}^{2}-\left(\frac{\boldsymbol{X}_{\mathrm{DM},0}\boldsymbol{V}_{\mathrm{M},0}}{\|\boldsymbol{X}_{\mathrm{DM},0}\|}\right)^{2}}\cdot f_{\mathrm{unit}}\left(\boldsymbol{X}_{\mathrm{DM},0}\times\boldsymbol{V}_{\mathrm{M},0}\times\boldsymbol{X}_{\mathrm{DM},0}\right) \tag{16}$$

式中：$\boldsymbol{X}_{\mathrm{DM},0}=\boldsymbol{X}_{\mathrm{M},0}-\boldsymbol{X}_{\mathrm{D},0}$为目标相对于拦截器的位置矢量。

预定拦截点计算、初制导、中制导等误差，导致拦截器进入末制导时的飞行状态在碰撞三角附近，如图4所示。为方便描述，定义平面C_t为$\boldsymbol{X}_{\mathrm{DM},0}$和$\boldsymbol{V}_{\mathrm{M},0}$所张成的平面，误差角$H_{\mathrm{E},\psi}\sim U\left(H_{\mathrm{E},\psi}^{\min},H_{\mathrm{E},\psi}^{\max}\right)$为$\boldsymbol{V}_{\mathrm{D},0}$与平面$C_t$的夹角，误差角$H_{\mathrm{E},\varphi}\sim U\left(H_{\mathrm{E},\varphi}^{\min},H_{\mathrm{E},\varphi}^{\max}\right)$为$\boldsymbol{V}_{\mathrm{D},0}$在平面$C_t$的投影与$\boldsymbol{V}_{\mathrm{D},0}^{\mathrm{ct}}$的夹角，其中，$H_{\mathrm{E},\psi}^{\min}$和$H_{\mathrm{E},\psi}^{\max}$分别为$H_{\mathrm{E},\psi}$的最小值和最大值，$H_{\mathrm{E},\varphi}^{\min}$和$H_{\mathrm{E},\varphi}^{\max}$分别为$H_{\mathrm{E},\varphi}$的最小值和最大值。可得拦截器初始速度矢量$\boldsymbol{V}_{\mathrm{D},0}$与满足碰撞三角速度矢量$\boldsymbol{V}_{\mathrm{D},0}^{\mathrm{ct}}$之间的关系为

$$\begin{cases}\boldsymbol{V}'_{\mathrm{D},0}=T_{\mathrm{q}}\left(H_{\mathrm{E},\varphi},f_{\mathrm{unit}}\left(\boldsymbol{X}_{\mathrm{DM},0}\times\boldsymbol{V}_{\mathrm{M},0}\right)\right)\cdot\boldsymbol{V}_{\mathrm{D},0}^{\mathrm{ct}}\\ \boldsymbol{V}_{\mathrm{D},0}=T_{\mathrm{q}}\left(H_{\mathrm{E},\psi},f_{\mathrm{unit}}\left(\boldsymbol{X}_{\mathrm{DM},0}\times\boldsymbol{V}_{\mathrm{M},0}\right)\times f_{\mathrm{unit}}\left(\boldsymbol{V}'_{\mathrm{D},0}\right)\right)\cdot\boldsymbol{V}'_{\mathrm{D},0}\end{cases} \tag{17}$$

式中：函数$T_{\mathrm{q}}(\cdot,\cdot)$的输出为旋转矩阵，第1个输入量为旋转角度，第2个输入量为旋转轴，该函数可通过四元数[55]求得。

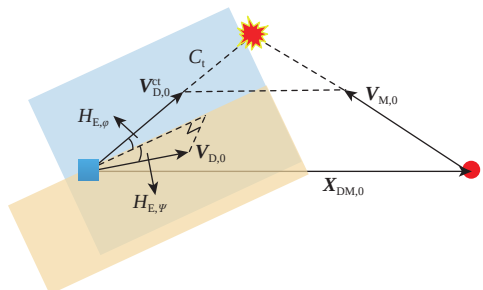


图4 拦截器初始状态误差示意

Fig. 4 Interceptor initial state error diagram

3 基于模仿学习的强化学习末制导律建模

3.1 总体框架

本文场景的拦截机动目标末制导律设计存在长序列难收敛、场景边界宽难学习及稀疏奖励问题。智能体完全从零开始学习，将带来有效奖励少、训练难以收敛、训练时间成本高等问题，因此，本文借鉴Alpha Go训练思想，提出一种基于模仿学习的PPO强化学习制导律(IL-based PPO guidance, IL-PPOG)总体框架，如图5所示，其中，V_{D}表示拦截器速度矢量。该框架主要包含模仿学习末制导模块和基于PPO的DRL末制导模块，后者以前者为基础，主要包含2个步骤：

1) 模仿学习制导(imitation learning guidance, ILG)。以PNG为专家策略，利用交互环境生成大量轨迹样本；建立智能末制导模型，采用监督学习方法对其进行离线训练，得到逼近PNG拦截性能的智能末制导模型。

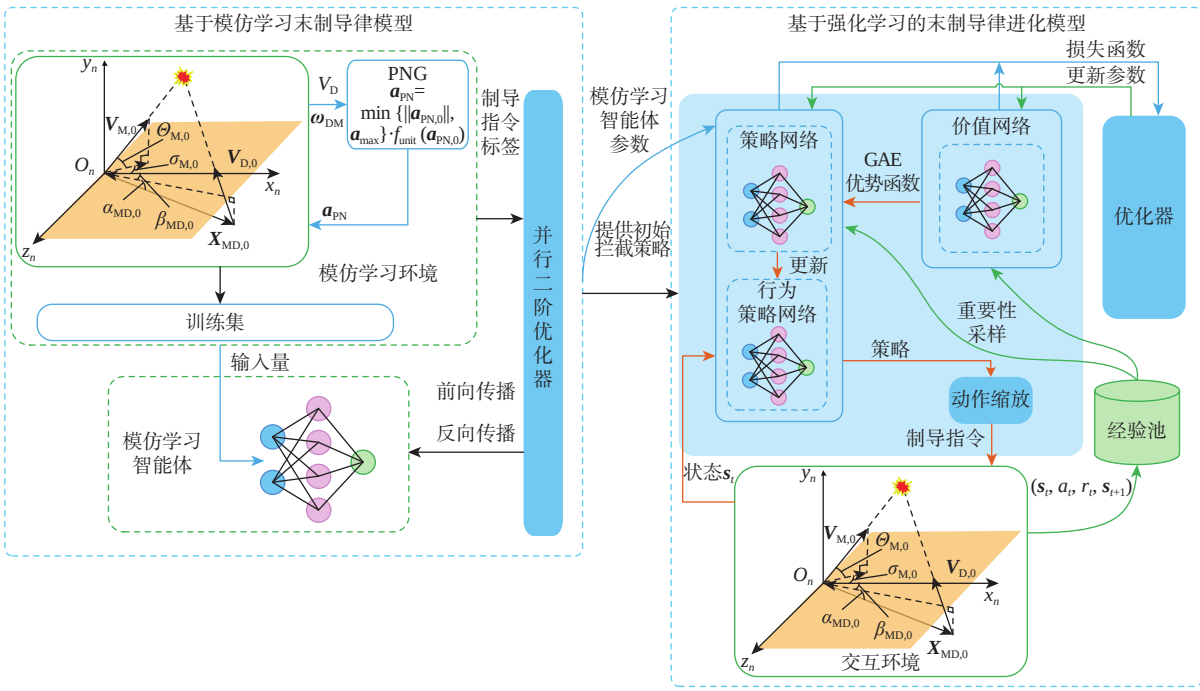


图 5　IL-PPOG 总体框架

Fig. 5　Main concept of IL-PPOG

2）PPO 强化学习制导。以步骤 1)学习得到的智能末制导模型为基础，建立强化学习执行者模块，进一步利用强化学习在序贯决策方面的优势，设置奖励函数，采用 PPO 算法对智能体进行离线训练，得到超越 PNG 拦截性能的新型制导律。训练完成后，实际在线应用时仅需执行策略网络即可。

3.2　模仿学习模块建模

3.2.1　模仿学习模型

ILG 模块的主要目的是为强化学习制导律模块提供初始策略，以降低学习难度、提高学习性能。因此，设计该模块以 PNG 指令为标签，学习 PNG 的拦截策略，可形式化表述为

$$f_{\mathrm{IL}}: \boldsymbol{s}_0^{\mathrm{IL}} \rightarrow \boldsymbol{a}_{\mathrm{PN}} \tag{18}$$

式中：$\boldsymbol{s}_0^{\mathrm{IL}}$为模仿学习模块输入量；$f_{\mathrm{IL}}$为$\boldsymbol{s}_0^{\mathrm{IL}}$到 PNG 指令$\boldsymbol{a}_{\mathrm{PN}}$的映射，模仿学习模块需要学习该映射。

模仿学习模块的核心是由神经网络f_{NN}为主体构成的制导指令生成器，需要合理设计神经网络输入输出量、激活函数和网络结构，以提高模仿学习模块性能。

3.2.2　输出量设计

考虑到末制导拦截中需要抑制视线角速度的产生，因此，令所需加速度沿视线角速度变化最大的方向，有利于节省拦截器燃料，令该方向的单位矢量为$\boldsymbol{e}_{a,\mathrm{unit}}$，则结合式 (4) 可知

$$\boldsymbol{e}_{a,\mathrm{unit}} = \arg\max \left\| \boldsymbol{X}_{\mathrm{DM}} \times \boldsymbol{e}_{a,\mathrm{unit}} \right\| \tag{19}$$

根据叉乘运算法则可知

$$\boldsymbol{e}_{a,\mathrm{unit}} = f_{\mathrm{unit}}(-\boldsymbol{X}_{\mathrm{DM}} \times \boldsymbol{\omega}_{\mathrm{DM}}) \tag{20}$$

由于拦截过程发生在大气层外，因此，可以忽略大气对拦截器姿态产生的影响，保证弹体方向始终指向目标方向，进而可产生沿$\boldsymbol{e}_{a,\mathrm{unit}}$的加速度。

为提高神经网络实时性，同时降低神经网络学习难度，在设计神经网络输出量时，应尽可能减小输出维度，并选取易于学习的变量作为输出量。因此，设计需要加速度（制导指令）沿$\boldsymbol{e}_{a,\mathrm{unit}}$方向，故神经网络输出量可设计为所需加速度大小，式 (18) 化简为

$$f_{\mathrm{IL}}: \boldsymbol{s}_0^{\mathrm{IL}} \rightarrow \|\boldsymbol{a}_{\mathrm{PN}}\| \tag{21}$$

式中：$\|\boldsymbol{a}_{\mathrm{PN}}\| \in [0, a_{\max}]$。

3.2.3　输入量设计及归一化方法

输入量$\boldsymbol{s}_0^{\mathrm{IL}}$的选取将直接影响模型学习的难易，需要反映拦截过程的关键信息[3, 49-50]。假设拦截器可获取弹目距离、接近速度和视线角速度等信息，设计输入量为

$$\boldsymbol{s}_0^{\mathrm{IL}} = \left[R_{\mathrm{DM}}, \dot{R}_{\mathrm{DM}}, \|\boldsymbol{\omega}_{\mathrm{DM}}\|\right]^{\mathrm{T}} \tag{22}$$

式中：R_{DM}和$\dot{R}_{\mathrm{DM}}$分别为弹目距离和接近速度。

为克服输入输出向量的不同维尺度不统一带来的训练效率低的问题，采用 min-max 归一化方法对$\boldsymbol{s}_0^{\mathrm{IL}}$进行归一化处理后得到$\tilde{\boldsymbol{s}}_0^{\mathrm{IL}}$：

$$\tilde{p} = 2\frac{p - p_{\min}}{p_{\max} - p_{\min}} - 1 \tag{23}$$

式中：p为变量；$p_{\min}$为变量的最小值；$p_{\max}$为变量的

最大值；$\tilde{p}$为归一化后的结果。

3.2.4 激活函数设计

为降低激活函数计算复杂度，隐藏层激活函数采用 S-Sigmoid 激活函数[56]：

$$\text{S-Sigmoid}(x)=\begin{cases}\dfrac{1}{2-x+0.5x^2} & x\leqslant 0\\ 1-\dfrac{1}{2+x+0.5x^2} & x>0\end{cases} \tag{24}$$

注释 1 激活函数的选择对神经网络而言十分重要，通常在浅层网络中采用 Sigmoid[57] 作为激活函数，该激活函数具有较强的非线性拟合能力。但由于此类激活函数中包含指数运算，计算量较大，不利于提高模型实时性。S-Sigmoid 激活函数具有与 Sigmoid 函数相近的拟合水平，但计算量小于 Sigmoid 函数。

考虑到模仿学习神经网络主要完成输入量到制导指令的映射，而制导指令存在最大值约束，如式 (5) 所示，因此，设计神经网络输出层激活函数为 tanh，其表达式为

$$\tanh(x)=\frac{\mathrm{e}^{x}-\mathrm{e}^{-x}}{\mathrm{e}^{x}+\mathrm{e}^{-x}} \tag{25}$$

结合式 (21) 可知，f_{IL}与$f_{\mathrm{IL}}^{\mathrm{NN}}$之间满足如下关系：

$$f_{\mathrm{IL}}\left(\tilde{\boldsymbol{s}}_0^{\mathrm{IL}}\right)=a_{\max}\left(f_{\mathrm{IL}}^{\mathrm{NN}}\left(\tilde{\boldsymbol{s}}_0^{\mathrm{IL}}\right)+1\right)/2 \tag{26}$$

式中：$f_{\mathrm{IL}}^{\mathrm{NN}}(\cdot)$为神经网络输出的归一化值。

3.2.5 网络结构设计

本文设计$f_{\mathrm{IL}}^{\mathrm{NN}}$为含有 2 个隐藏层的全连接神经网络，其网络结构如图 6 所示。本文选取隐藏层 1 的节点数为 6，隐藏层 2 的节点数为 5。

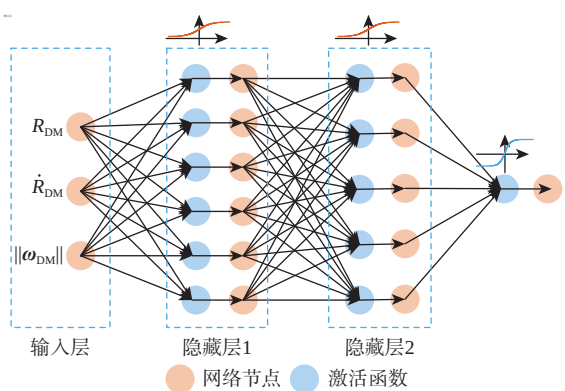


图 6 网络结构示意

Fig. 6 Network structure diagram

3.2.6 损失函数

模型学习采用均方差（mean squared error, MSE）损失函数[58]：

$$M_{\mathrm{SE}}=\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(y_{i,j}-\bar{y}_{i,j}\right)^2 \tag{27}$$

式中：M为样本数量；N为神经网络输出值的数量；$y_{i,j}$和$\bar{y}_{i,j}$分别为第i个样本的第j个输出量的预测值和标签值。

3.3 基于 DRL 的智能末制导律模块

强化学习在拦截制导领域应用的关键是如何设计合理的马尔可夫决策过程，包括状态空间、动作空间和奖励函数等，以提高学习效率和拦截制导律性能。本节给出马尔可夫决策过程模型。

3.3.1 马尔可夫决策空间建模

1）状态空间与动作空间。强化学习智能体在模仿学习的基础上，进一步探索拦截机动目标的有效策略，因此，强化学习智能体的状态空间与模仿学习的输入量一致，即式 (22)，可得状态空间为

$$\boldsymbol{s}_t^{\mathrm{RL}}=\boldsymbol{s}_0^{\mathrm{IL}}=\left[R_{\mathrm{DM}},\dot{R}_{\mathrm{DM}},\|\boldsymbol{\omega}_{\mathrm{DM}}\|\right]^{\mathrm{T}} \tag{28}$$

强化学习智能体的动作输出为制导指令，结合式 (21) 可得动作空间为

$$u_t=[0,a_{\max}] \tag{29}$$

2）奖励函数。奖励机制是保证 DRL 能够学习到较优策略的关键因素之一，因此，奖励函数设计的好坏将直接影响学习效果。在奖励函数设计过程中，需要结合研究场景所要实现的目标，保证更优策略应该获得更高的奖励，从而保证策略可靠收敛。

本文研究的拦截机动目标场景包含 2 个优化目标：最大化拦截概率和最小化能量消耗。同时，由于只有在仿真结束时才能对智能体拦截策略进行评价，因此，存在稀疏奖励问题。为缓解稀疏奖励造成智能体收敛速度慢、学习效果差的问题，且实现能量消耗小、拦截概率高，设计包含 2 部分的奖励函数为

$$r_t=r_{\mathrm{s}}+r_{\mathrm{e}} \tag{30}$$

$$r_{\mathrm{s}}=k_{\mathrm{s}}\exp\left[-(u_t/g)^2\right] \tag{31}$$

$$r_{\mathrm{e}}=\begin{cases}\dfrac{R_{\mathrm{miss}}-R_{\mathrm{miss}}^{\mathrm{th},1}}{R_{\mathrm{miss}}^{\mathrm{th},2}-R_{\mathrm{miss}}^{\mathrm{th},1}}\left(r_{\mathrm{e}}^{\mathrm{th},2}-r_{\mathrm{e}}^{\mathrm{th},1}\right)+r_{\mathrm{e}}^{\mathrm{th},1} & R_{\mathrm{miss}}^{\mathrm{th},1}\leqslant R_{\mathrm{miss}}\leqslant R_{\mathrm{miss}}^{\mathrm{th},2}\\ r_{\mathrm{e}}^{\mathrm{th},3} & R_{\mathrm{miss}}\leqslant R_{\mathrm{miss}}^{\mathrm{th},1}\\ 0 & R_{\mathrm{miss}}>R_{\mathrm{miss}}^{\mathrm{th},2}\end{cases} \tag{32}$$

式中：r_{s}和r_{e}分别为过程奖励和终端奖励，r_{s}为反映能量消耗程度的奖励值，能耗越低则正奖励越大，r_{e}为反映终端脱靶量的奖励，当拦截成功时给出一个很大的正奖励；k_{s}为大于 0 的超参数；R_{miss}为脱靶量；$R_{\mathrm{miss}}^{\mathrm{th},1}$和$R_{\mathrm{miss}}^{\mathrm{th},2}$为与脱靶量相关的超参数，$R_{\mathrm{miss}}^{\mathrm{th},1}$为成功拦截目标对应的脱靶量；$r_{\mathrm{e}}^{\mathrm{th},1}$、$r_{\mathrm{e}}^{\mathrm{th},2}$和$r_{\mathrm{e}}^{\mathrm{th},3}$为与奖励有关的超参数。$r_{\mathrm{e}}$的示意如图 7 所示。

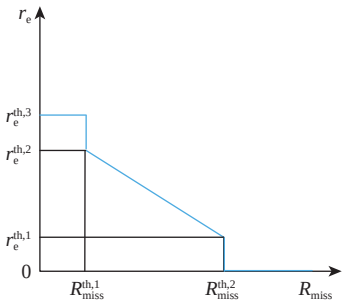


图 7　奖励r_e的示意

Fig. 7　Diagram of reward r_e

由图 7 可以看出, r_e主要分为 3 段: ① 当$R_{miss} > R_{miss}^{th,2}$时, 表示拦截失败, 智能体奖励为 0; ② 当$R_{miss}^{th,1} \leqslant R_{miss} \leqslant R_{miss}^{th,2}$时, 表示虽然拦截失败, 但脱靶量较小, 因此, 智能体仍然能够获取一定的正奖励; ③ 当$R_{miss} \leqslant R_{miss}^{th,1}$时, 表示拦截成功, 智能体获得一个很大的正奖励。不同于文献 [3, 8, 49-50] 中采用的“硬”终端奖励, 即仅当拦截成功时, 智能体才能获取一个很大的正奖励, 本文提出一种加入“过渡段”的“软”终端奖励, 在一定程度上有利于缓解稀疏奖励引起的收敛困难问题。换言之, 当智能体拦截失败时, 智能体仍然能够获取奖励, 且该奖励能够为智能体探索拦截成功策略提供方向, 进而引导智能体学习到拦截成功的策略。

注释 2　需要说明的是, 本文在设计过程奖励r_s时仅考虑了能耗奖励, 其主要原因如下:

1) 本文的强化学习智能体以模仿学习策略为基础, 已经具备一定的拦截目标能力, 因此, 不需要加入过多的过程奖励引导智能体拦截目标。

2) 通过采用“软”终端奖励形式, 在一定程度上降低了稀疏奖励影响, 进而可弱化对过程奖励的需求。

3) 当加入过多过程奖励时, 各个奖励的系数设计将很困难, 面临大量调参工作。

3.3.2　基于 PPO 算法的模型求解

PPO 算法在连续控制任务中表现出优异的性能, 被广泛应用于各个控制领域, Open AI 公司将 PPO 作为首选算法。因此, 本文采用 PPO 作为强化学习制导策略的训练算法, 并对 PPO 算法进行简要介绍, 具体可参见文献 [22]。

PPO 包含 1 个策略网络和 1 个价值网络, 其优化目标是最大化式 (33) 所给目标函数$L^{PPO}(\boldsymbol{\theta},\boldsymbol{\omega})$:

$$L^{PPO}(\boldsymbol{\theta},\boldsymbol{\omega}) = E\left[L^{CLIP}(\boldsymbol{\theta}) - c_{VF}L^{VF}(\boldsymbol{\omega}) + c_s L^{\mathcal{H}}(\boldsymbol{\theta})\right] \tag{33}$$

式中: E表示求均值; c_{VF}和c_s为超参数; $\boldsymbol{\theta}$和$\boldsymbol{\omega}$分别为策略网络和价值网络的参数; $L^{CLIP}(\boldsymbol{\theta})$为策略网络的优化目标; $L^{VF}(\boldsymbol{\omega})$为价值网络的优化目标; $L^{\mathcal{H}}(\boldsymbol{\theta}) = E_{a\sim\pi_{\boldsymbol{\theta}}}[-\ln\pi_{\boldsymbol{\theta}}(a|\boldsymbol{s})]$为策略的熵, 反映了策略的不确定性。

$$L^{CLIP}(\boldsymbol{\theta}) = E_{(\boldsymbol{s},a)\sim\pi_{\boldsymbol{\theta},old}}\left[\min\left(p(\boldsymbol{\theta})A^{GAE}_{\pi_{\boldsymbol{\theta},old}}(\boldsymbol{s},a),\ \mathrm{clip}\left(p(\boldsymbol{\theta}),1-\varepsilon,1+\varepsilon\right)A^{GAE}_{\pi_{\boldsymbol{\theta},old}}(\boldsymbol{s},a)\right)\right] \tag{34}$$

式中: $\pi_{\boldsymbol{\theta}}$和$\pi_{\boldsymbol{\theta},old}$分别为新策略和旧策略; $p(\boldsymbol{\theta})$为新旧策略对应的概率比值, 可表示为$p(\boldsymbol{\theta}) = \pi_{\boldsymbol{\theta}}(a|\boldsymbol{s})/\pi_{\boldsymbol{\theta},old}(a|\boldsymbol{s})$; $\mathrm{clip}(\cdot,\cdot,\cdot)$为裁剪函数, 可表示为

$$\mathrm{clip}(p(\boldsymbol{\theta}),1-\varepsilon,1+\varepsilon) = \begin{cases} 1-\varepsilon & p(\boldsymbol{\theta}) < 1-\varepsilon \\ p(\boldsymbol{\theta}) & 1-\varepsilon \leqslant p(\boldsymbol{\theta}) \leqslant 1+\varepsilon \\ 1+\varepsilon & p(\boldsymbol{\theta}) > 1+\varepsilon \end{cases} \tag{35}$$

式中: ε为很小的超参数, 用来保证新旧策略的差距不大。

$$L^{VF}(\boldsymbol{\omega}) = \left(V_{\pi_{\boldsymbol{\theta}}}(\boldsymbol{s};\boldsymbol{\omega}) - V^{targ}_{\pi_{\boldsymbol{\theta},old}}(\boldsymbol{s};\boldsymbol{\omega})\right)^2 \tag{36}$$

式中: $V_{\pi_{\boldsymbol{\theta}}}(\boldsymbol{s};\boldsymbol{\omega})$为$\pi_{\boldsymbol{\theta}}$的价值估计; $V^{targ}_{\pi_{\boldsymbol{\theta},old}}(\boldsymbol{s};\boldsymbol{\omega})$为价值网络的目标, 计算公式为

$$V^{targ}_{\pi_{\boldsymbol{\theta},old}}(\boldsymbol{s};\boldsymbol{\omega}) = A^{GAE}_{\pi_{\boldsymbol{\theta},old}}(\boldsymbol{s},a) + V_{\pi_{\boldsymbol{\theta},old}}(\boldsymbol{s};\boldsymbol{\omega}_{old}) \tag{37}$$

引入奖励缩放[59] 技巧, 以避免因过大或过小的奖励对价值网络的学习产生负面影响, 缩放后的价值网络的目标为

$$\tilde{V}^{targ}_{\pi_{\boldsymbol{\theta},old}}(\boldsymbol{s};\boldsymbol{\omega}) = \frac{V^{targ}_{\pi_{\boldsymbol{\theta},old}}(\boldsymbol{s};\boldsymbol{\omega})}{\sigma_{V^{targ}}} \tag{38}$$

式中: $\sigma_{V^{targ}}$为所有$V^{targ}_{\pi_{\boldsymbol{\theta},old}}(\boldsymbol{s};\boldsymbol{\omega})$的标准差。

同时, 为提高训练过程的稳定性, 引入梯度裁剪技巧[59] 求解$\boldsymbol{\theta}$的梯度$g_{\boldsymbol{\theta}}$和$\boldsymbol{\omega}$的梯度$g_{\boldsymbol{\omega}}$。

$$\begin{cases} g_{\boldsymbol{\theta}} = \min\left\{g_{max}/\left\|\nabla_{\boldsymbol{\theta}}L^{PPO}(\boldsymbol{\theta},\boldsymbol{\omega})\right\|,1.0\right\}\cdot\nabla_{\boldsymbol{\theta}}L^{PPO}(\boldsymbol{\theta},\boldsymbol{\omega}) \\ g_{\boldsymbol{\omega}} = \min\left\{g_{max}/\left\|\nabla_{\boldsymbol{\omega}}L^{PPO}(\boldsymbol{\theta},\boldsymbol{\omega})\right\|,1.0\right\}\cdot\nabla_{\boldsymbol{\omega}}L^{PPO}(\boldsymbol{\theta},\boldsymbol{\omega}) \end{cases} \tag{39}$$

式中: g_{max}为梯度裁剪值。

为更新策略网络和价值网络的参数, 采用 Adam 优化器最大化式 (33) 所给目标函数, 即

$$\begin{cases} \boldsymbol{\theta} \leftarrow \boldsymbol{\theta} + \alpha_{lr}g_{\boldsymbol{\theta}} \\ \boldsymbol{\omega} \leftarrow \boldsymbol{\omega} + \beta_{lr}g_{\boldsymbol{\omega}} \end{cases} \tag{40}$$

式中: α_{lr}和β_{lr}分别为策略网络和价值网络的学习率。

进一步, 为增强训练后期的稳定性, 引入学习率衰减技巧, 即

$$\begin{cases} \alpha_{lr} = \alpha_{lr,0}\left(1 - t_e/N_{epoch}\right) \\ \beta_{lr} = \beta_{lr,0}\left(1 - t_e/N_{epoch}\right) \end{cases} \tag{41}$$

式中: $\alpha_{lr,0}$和$\beta_{lr,0}$分别为策略网络和价值网络的初始学习率; t_e为当前训练代数; N_{epoch}为最大训练代数。

3.3.3　网络结构设计

PPO 的策略网络以模仿学习智能体网络为基础, 包含 2 个输出量, 分别是动作均值μ和动作标准差a_{std}, 其中, 动作均值μ的网络由f_{IL}^{NN}继承而来。在

训练阶段，智能体动作a_t根据分布$N\left(\mu,a_{\text{std}}^2\right)$采样后经过tanh激活函数和动作缩放后得到。本文设计PPO策略网络和价值网络节点数、激活函数如表2所示。

表2 PPO网络结构和参数

Table 2 PPO network structure and parameters

网络	层级	单元数	激活函数
策略网络	输入层	3	
	隐藏层1	6	S-Sigmoid
	隐藏层2	5	S-Sigmoid
	输出层	1	Linear
价值网络	输入层	3	
	隐藏层1	6	S-Sigmoid
	隐藏层2	5	S-Sigmoid
	输出层	1	Linear

4 仿真分析

为对本文建立的IL-PPOG算法有效性进行验证，以表1给出的参数为场景参数，建立交互环境，设计了仿真实验。本节给出超参数设计，对模仿学习性能进行验证，进一步分析强化学习制导过程，在未知场景下对算法性能进行测试。

本文仿真在Python 3.7、Pytorch 1.5、Visual Studio 2015和SWIG 4.0.1等软件工具支撑下开展。为提高交互环境运行速度，在Visual Studio 2015环境下基于C++语言建立拦截器拦截机动目标的交互环境。采用四阶龙哥库塔和阿达姆斯积分器(前者为后者提供开始4步的积分结果)，积分步长为0.02 s，由文献[60]给出的精确脱靶量计算公式完成拦截脱靶量计算，并采用SWIG 4.0.1建立与Python的数据接口。模仿学习阶段，采用的PNG律导航比为4.0，强化学习阶段，决策步长为0.1 s。仿真平台性能如下：CPU为Intel 酷睿 13900K，内存为64 GB，显卡为英伟达3080(10 GB)。

拦截过程中，观测量和动作量存在不确定性，因此，加入观测噪声($\boldsymbol{o}_t$)和动作噪声(a_t)，即

$$\begin{cases}\boldsymbol{o}_t=\boldsymbol{s}_t\cdot N\left(\boldsymbol{1}_{3\times1},\boldsymbol{\sigma}_{\boldsymbol{o}}\right)\\ a_t\leftarrow a_t+N\left(0,\sigma_a^2\right)\end{cases}\tag{42}$$

式中：$\boldsymbol{1}_{3\times1}$为元素全为1的向量；$\boldsymbol{\sigma}_{\boldsymbol{o}}=\text{diag}\left(\delta_{R_{\text{DM}}},\delta_{\dot{R}_{\text{DM}}},\delta_{\|\boldsymbol{\omega}_{\text{DM}}\|}\right)$，$\delta_{R_{\text{DM}}}$、$\delta_{\dot{R}_{\text{DM}}}$和$\delta_{\|\boldsymbol{\omega}_{\text{DM}}\|}$分别为观测量$R_{\text{DM}}$、$\dot{R}_{\text{DM}}$和$\|\boldsymbol{\omega}_{\text{DM}}\|$噪声的标准差；$\sigma_a$为动作噪声的标准差。

本文建立的IL-PPOG算法含有较多的超参数，主要涉及网络训练、奖励函数、环境设置等内容，具体参数如表3所示，其中，N_{IL}为模仿学习轨迹数量，$|\mathcal{D}|$为经验池大小，$|\mathcal{B}|$为批次大小，N_{update}为每次训练网络参数更新次数，N_{rollout}为每次训练收集的轨迹数。

表3 超参数设计

Table 3 Hyper-parameter design

超参数	数值	超参数	数值
N_{IL}	10 000	$\beta_{\text{lr},0}$	0.001
$N_{\text{epoch}}^{\text{IL}}$	1 000	N_{epoch}	300
k_{s}	0.1	N_{update}	1 000
$r_{\text{e}}^{\text{th},1}$	10.0	N_{rollout}	50
$r_{\text{e}}^{\text{th},2}$	50.0	N_{m}	4.0 g
$r_{\text{e}}^{\text{th},3}$	51.0	ΔT_{m}/s	1.0
$R_{\text{miss}}^{\text{th},1}$	1.0	$a_{\max}$	4.0 g
$R_{\text{miss}}^{\text{th},2}$	10.0	$H_{\text{E},\psi}^{\min}$/(°)	−1.5
$\|\mathcal{D}\|$	75 000	$H_{\text{E},\psi}^{\max}$/(°)	1.5
$\|\mathcal{B}\|$	256	$H_{\text{E},\varphi}^{\min}$/(°)	−1.5
γ	0.99	$H_{\text{E},\varphi}$/(°)	1.5
λ	0.95	$\|\boldsymbol{X}_{\text{MD},0}\|$/km	200.0
c_{VF}	0.25	$\delta_{R_{\text{DM}}}$/%	1.0
c_{s}	0.01	$\delta_{\dot{R}_{\text{DM}}}$/%	0.5
$g_{\max}$	0.5	$\delta_{\|\boldsymbol{\omega}_{\text{DM}}\|}$/%	3.0
ε	0.2	σ_a/%	5/3
$\alpha_{\text{lr},0}$	0.001		

4.1 模仿学习

为生成模仿学习训练样本，以PNG输出指令为动作与环境进行交互，生成10 000条拦截轨迹，每条轨迹飞行时长约为15 s，图8给出了500组对抗样本的三维轨迹。以0.1 s时间间隔对拦截轨迹进行采样，进而得到含有1.5×10^6个样本的训练集，对f_{IL}训练1 000代，得到训练曲线如图9所示，其中，y轴采用对数坐标形式。

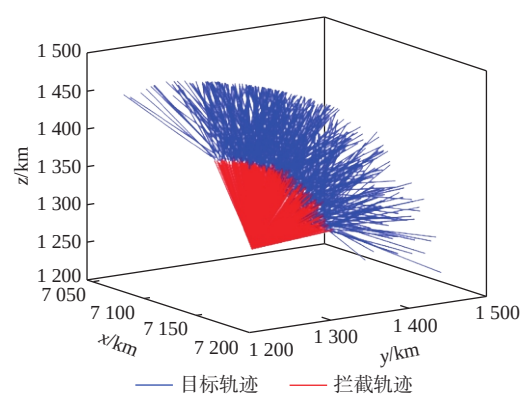


图8 模仿学习的500组样本的轨迹

Fig. 8 Trajectories of 500 samples for IL

由图9可以看出，随着训练代数的增加，训练曲线逐渐收敛，说明模仿学习智能体能够学习到PNG输出的指令。为测试基于f_{IL}的制导指令的拦截性能，在交互环境中开展了10 000次拦截仿真，得到脱靶量和能量消耗如图10所示。同时，图10给出了以PNG为指令的脱靶量和能量消耗。其

中，能量消耗性能指标J_a的定义如下：

$$J_a = \int_0^{t_f} |a_t| \mathrm{d}t \tag{43}$$

式中：t_f为拦截飞行仿真结束对应的时间。

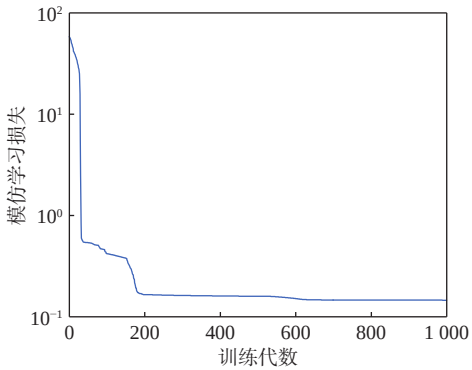


图 9　模仿学习损失随训练代数变化曲线

Fig. 9　Curve of imitation learning loss as training epoch

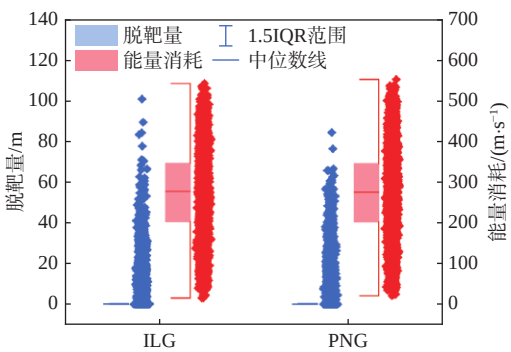


图 10　ILG 和 PNG 能量消耗和脱靶量箱线

Fig. 10　Boxplot of energy consumption and miss distance for ILG and PNG

由图 10 可以看出，基于f_{IL}的制导指令和基于 PNG 的制导指令在脱靶量和能量消耗指标上相近，说明f_{IL}较好地学习到了 PNG 指令，能够为强化学习智能体学习提供较好的初始策略。

4.2　强化学习训练

在强化学习智能体训练阶段，以f_{IL}为初始策略对强化学习智能体进行学习。假设最大训练代数为 300，经验池大小为 7 500，每次参数更新时收集轨迹数为 50，并对网络进行 1 000 次更新。为展示智能体策略性能随学习进程的变化趋势，以 5 个不同的随机种子对智能体进行训练，并采用正交初始化方式[61]初始化网络参数。每次训练间隔 10 代以智能体策略为指令开展 1 000 次蒙特卡罗拦截仿真（每次仿真均随机初始状态），得到轨迹累积奖励平均值曲线如图 11 所示，脱靶量和能量消耗量平均值如图 12 所示。需要说明的是，在图 11 和图 12 中，实线表示平均值，阴影表示标准差。

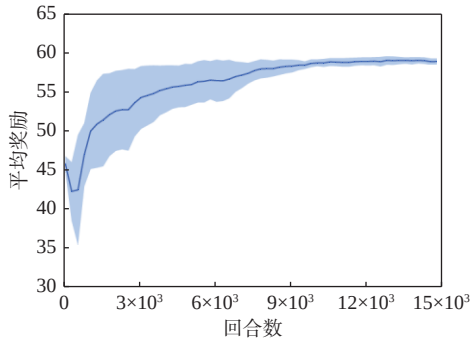


图 11　IL-PPOG 训练曲线

Fig. 11　Training curve of IL-PPOG

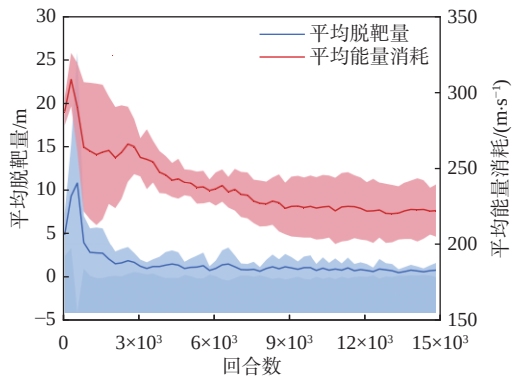


图 12　脱靶量和能量消耗随训练代数变化曲线

Fig. 12　Miss distance and energy consumption curves varing with training epoch

由图 11 可以看出，随着训练代数的增加，智能体的平均奖励曲线逐渐增大，并在回合数达到1.2×10^3时收敛；同时，由图 11 中收敛时的阴影可知，IL-PPOG 在不同随机种子下均具有较好的收敛性。由图 12 可知，由于智能体以模仿学习策略为基础，训练前期智能体已经具有较好的拦截性能，但能量消耗量较大；随着训练代数的增加，智能体的拦截脱靶量和能量消耗均有所减小，说明智能体探索到了更高效拦截策略。

为验证 IL-PPOG 算法相对于其他研究的优势，本文与文献 [61] 提出的以 PNG 导航比为动作的强化学习制导算法进行了对比分析，同时为方便描述，将文献 [39-41] 制导算法记为 LC-PPOG。图 13 和图 14 给出了 IL-PPOG 和 LC-PPOG 的训练过程中每间隔 10 代对智能体的 1 000 次测试中拦截脱靶量和能量消耗量平均值随训练代数的变化曲线。

可以看出，训练初期的 IL-PPOG 相比于 LC-PPOG，具有更小的平均拦截脱靶量；随着训练代数的增加，IL-PPOG 的平均脱靶量逐渐减小，能量消耗量逐渐降低，智能体学习到的策略实现了能量消耗最低且拦截概率更大。LC-PPOG 训练初期的拦截性能与 IL-PPOG 相近，在收敛时的平均脱靶量高于 IL-PPOG、平均能量消耗略低于 IL-PPOG，考虑

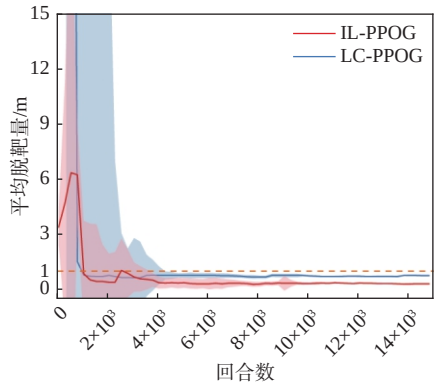


图 13 不同算法脱靶量随训练代数变化曲线

Fig. 13 Curves of miss distance versus training epoch for different algorithms

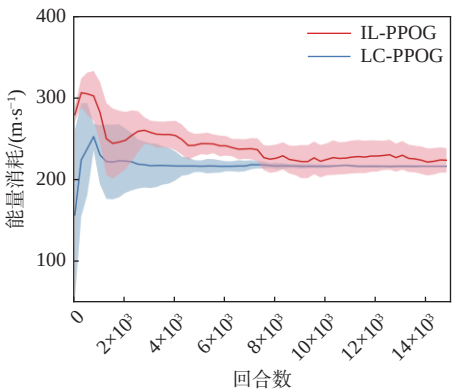


图 14 不同算法能量消耗随训练代数变化曲线

Fig. 14 Curves of energy consumption versus training epoch for different algorithms

到脱靶量是影响拦截性能的主要因素，可知 IL-PPOG 拦截性能略高于 LC-PPOG。

4.3 测试过程

为测试本文算法的拦截性能，选择 4.2 节中训练效果最好的 IL-PPOG 智能体，在以表 1 所给参数建立的包含各种误差和噪声交互环境下，目标采用方波机动形式，开展了 10 000 次蒙特卡罗仿真，得到不同算法拦截脱靶量和能量消耗结果，如图 15 所示。可以看出，IL-PPOG 在已知场景下脱靶量的

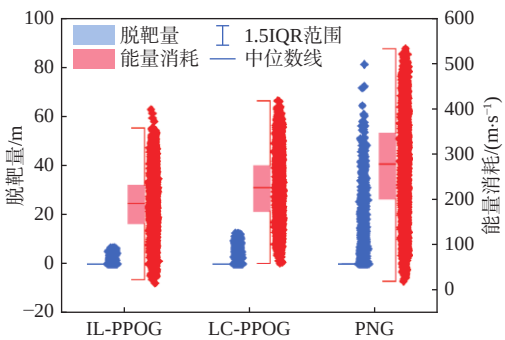


图 15 不同算法能量消耗和脱靶量箱线

Fig. 15 Boxplots of energy consumption and miss distance for different algorithms

3σ值和平均值分别为 6.11 m 和 0.34 m，小于 PNG、LC-PPOG 对应的脱靶量；同时，IL-PPOG 的能量消耗的 3σ值和平均值分别为 344.165 8 m/s 和 189.630 7 m/s，小于 PN、LC-PPOG 对应的能量消耗。这说明，IL-PPOG 对机动目标具有更好的拦截性能，在外干扰环境下具有较好的鲁棒性。

为进一步测试不同$H_{\mathrm{E},\psi}$和$H_{\mathrm{E},\varphi}$情况下不同算法的拦截性能，针对不同情况开展了 10 000 次蒙特卡罗仿真，统计拦截概率（当脱靶量小于 1.0 m 认为拦截成功）和能量消耗，如图 16 和图 17 所示。可以看出，IL-PPOG 在不同情况下，均具有更高的拦截性能，且在大偏差情况下，IL-PPOG 的优势更加明显，进一步表明算法具有较好的鲁棒性。同时，考虑到大偏差情况代表中末交班误差大，因此，IL-PPOG 可以拓宽中末交班精度约束。

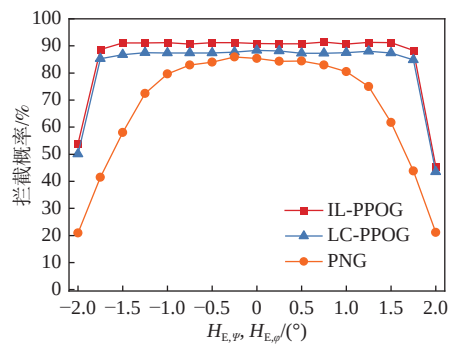


图 16 不同$H_{\mathrm{E},\psi}$和$H_{\mathrm{E},\varphi}$情况下拦截概率

Fig. 16 Interception probability under different $H_{\mathrm{E},\psi}$ and $H_{\mathrm{E},\varphi}$ conditions

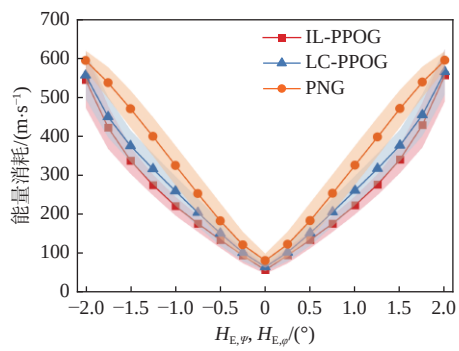


图 17 不同$H_{\mathrm{E},\psi}$和不同$H_{\mathrm{E},\varphi}$情况下能量消耗

Fig. 17 Energy consumption under different $H_{\mathrm{E},\psi}$ and different $H_{\mathrm{E},\varphi}$ conditions

在假设目标机动策略为距离拦截器 40 km 时向上机动及$H_{\mathrm{E},\psi}=H_{\mathrm{E},\varphi}=1.5°$的情况下，统计 10 个拦截过程中拦截弹拦截过载指令、目标机动指令曲线、视线角速度、拦截弹和目标飞行轨迹如图 18～图 21 所示。可以看出，由于$H_{\mathrm{E},\psi}$和$H_{\mathrm{E},\varphi}$的存在，IL-PPOG 在末制导初期给出较大的拦截过载指令，迅速降低视线角速度，并在目标机动后再次给出较

大的拦截过载，进而实现有效拦截。由图 20 可知，在拦截过程中，采用 PNG 和 LC-PPOG 制导律对应的视线角速度大于采用 IL-PPOG 制导律对应的视线角速度；由图 18 可知，在接近目标时，PNG 制导律比 IL-PPOG 制导律更早达到最大拦截过载，由此导致 PNG 的拦截失败，而 IL-PPOG 则可以拦截成功，进一步表明 IL-PPOG 具有更好的拦截性能。

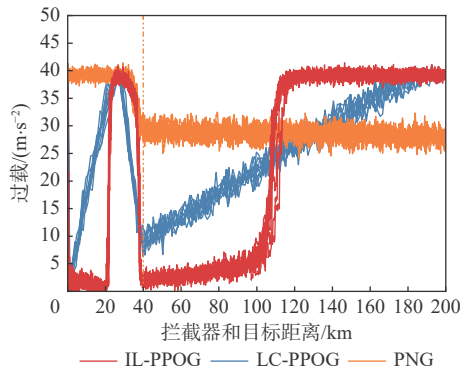


图 18　不同算法的加速度指令

Fig. 18　Acceleration commands for different algorithms

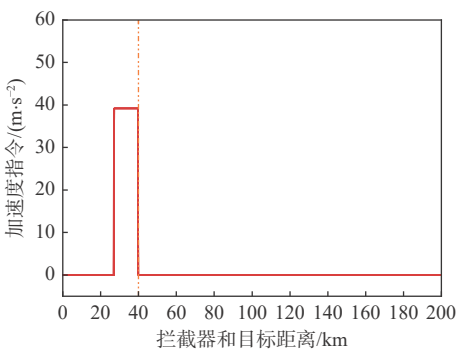


图 19　目标机动加速度指令

Fig. 19　Target maneuver acceleration command

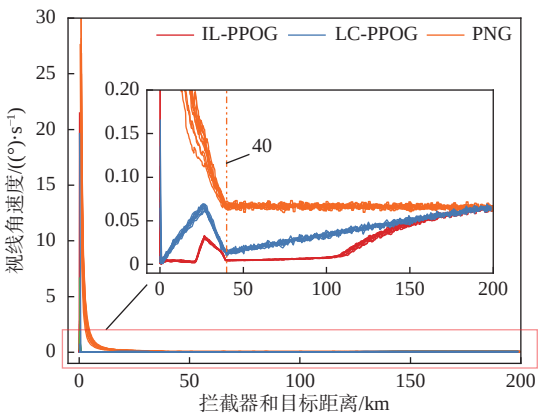


图 20　不同算法视线角速度

Fig. 20　Line-of-sight angular velocity for different algorithms

4.4　未知场景测试

为验证 IL-PPOG 对未知拦截环境的泛化能力，建立未知场景如下：将目标机动形式由方波机动形式改为正弦机动形式，如式 (44) 所示，其他场景参数不变。在目标正弦机动场景下，目标和拦截器初始状态均按照第 2 节模型随机生成，并加入观测噪声和动作噪声，开展 10 000 次蒙特卡罗仿真，统计脱靶量和能量消耗，如图 22～图 25 所示。

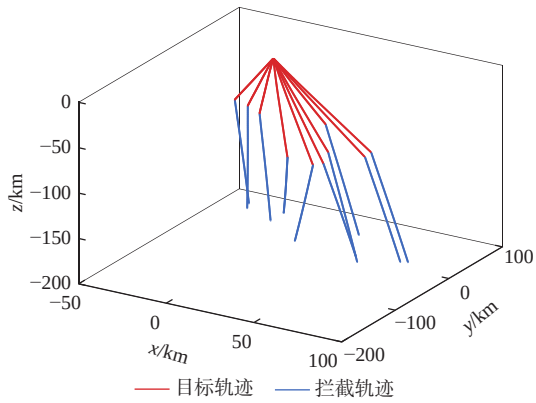


图 21　目标和拦截器三维轨迹

Fig. 21　Three-dimensional trajectories of targets and interceptors

$$\boldsymbol{a}'_{\mathrm{M}}=\begin{cases}N_{\mathrm{m}}g\sin\left(\omega_{\mathrm{s}}\left(t-T_{\mathrm{m},0}\right)\right)\cdot\boldsymbol{e}_{\mathrm{M}} & T_{\mathrm{m},0}\leqslant t\leqslant T_{\mathrm{m},0}+\Delta T_{\mathrm{m}}\\ \boldsymbol{0} & \text{其他}\end{cases}\tag{44}$$

式中：ω_{s}为正弦机动频率。

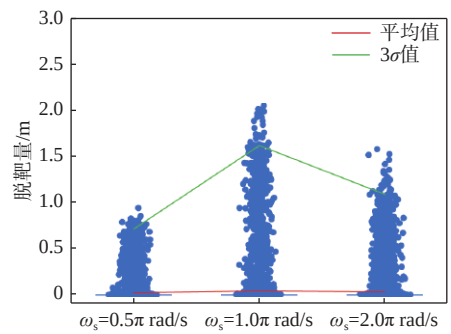


图 22　IL-PPOG 在不同ω_{s}情况下脱靶量

Fig. 22　Miss distance of IL-PPOG under different ω_{s}

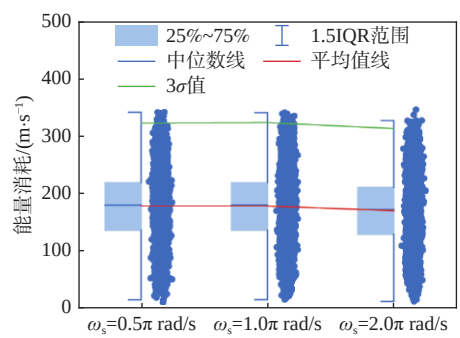


图 23　IL-PPOG 在不同ω_{s}情况下能量消耗

Fig. 23　Energy consumption of IL-PPOG under different ω_{s}

图 22 和图 23 给出了不同ω_{s}取值的正弦机动场景下，IL-PPOG 的拦截脱靶量和能量消耗量。可以看出，$\omega_{\mathrm{s}}=\pi$ rad/s 对应的拦截脱靶量相较于其他

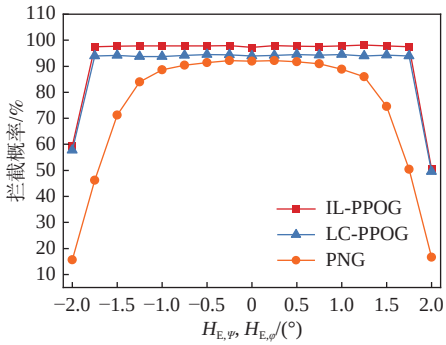


图 24 不同$H_{E,\psi}$和$H_{E,\varphi}$情况下拦截概率 ($\omega_s = \pi$ rad/s)

Fig. 24 Interception probability under different $H_{E,\psi}$ and $H_{E,\varphi}$ conditions ($\omega_s = \pi$ rad/s)

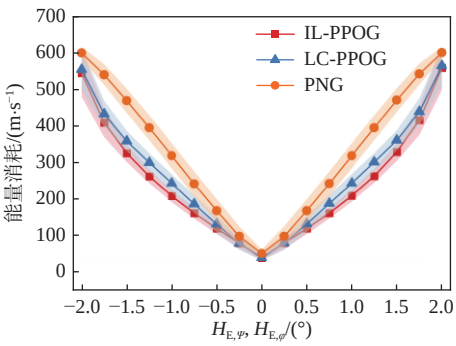


图 25 不同$H_{E,\psi}$和$H_{E,\varphi}$情况下能量消耗 ($\omega_s = \pi$ rad/s)

Fig. 25 Energy consumption under different $H_{E,\psi}$ and $H_{E,\varphi}$ conditions ($\omega_s = \pi$ rad/s)

情况更大，说明目标采用以$\omega_s = \pi$ rad/s 的正弦机动模式能够增大拦截器拦截难度。在$\omega_s = \pi$ rad/s 的情况下，IL-PPOG脱靶量的3σ值和平均值分别为1.62 m 和 0.047 m，能量消耗的3σ值和平均值分别为 322.83 m/s 和 177.91 m/s，说明拦截脱靶量很小且能量消耗较低。结合 4.3 节仿真结果可知，IL-PPOG 对于不同机动模式的目标均具有较好的拦截性能和泛化能力。在不同$H_{E,\psi}$和$H_{E,\varphi}$场景下，令正弦机动频率$\omega_s = \pi$ rad/s，针对每种情况各开展10 000 次蒙特卡罗仿真，不同算法的拦截概率和能量消耗如图 24和图 25 所示。可以看出，IL-PPOG 相比于 LC-PPOG和 PNG 具有更好的拦截性能，由此可知，IL-PPOG 具有适应外干扰未知机动模式场景的能力，有一定的工程应用潜力。

4.5 计算复杂度分析

由表 2 给出的 PPO 网络结构可知，IL-PPOG 包含 1 个策略网络和 1 个价值网络，其中，价值网络的主要作用是在离线训练阶段提供策略价值估计，在线应用阶段仅需要执行策略网络即可。因此，本文决策方法的实时性主要由 IL-PPOG 策略网络的计算复杂度决定。考虑到神经网络模型的复杂度基本由乘法量和激活函数个数来决定[62]，IL-PPOG 策略网络的乘法量为 3×6+6×5+5×1=53，激活函数个数为 11，可知决策算法复杂度较小。为进一步验证本文决策算法的实时性，以 STM32F407 单片机为实验平台，采用 double 数据类型，开展了算法耗时实验。将 IL-PPOG 策略网络烧录到 STM32F407 单片机，通过单片机计时函数统计 CPU 耗时，得到策略网络耗时约为 0.32 ms，说明实时性较好，能够降低对弹上计算能力的需求。

5 结 论

聚焦机动目标拦截制导律设计问题，本文提出了一种新型强化学习制导律，相较于传统制导律而言，实现了大偏差下更高的拦截性能。

1）以拦截碰撞三角为基础，设计了拦截弹和目标初始状态求解方法，建立了拦截弹和目标三维不确定对抗场景。

2）将模仿学习技术和强化学习技术应用于拦截弹末制导中，通过模仿学习挖掘 PNG 制导律规律，为强化学习智能体提供良好的初始策略；建立拦截智能体的马尔可夫决策模型，引入包含“过渡段”的“软”终端奖励，利用 PPO 算法对拦截策略进行求解，同时，运用奖励归一化、梯度裁剪和学习率衰减等技巧稳定训练。

3）在学习场景和未学习机动模式场景中，通过蒙特卡罗方法对 IL-PPOG 性能进行了测试。仿真结果表明，在前述 2 种场景中，IL-PPOG 均具有比传统拦截制导律更好的性能，具有更高的拦截概率和更低的能量消耗；同时，算法计算复杂度较低，在 STM32F407 上单次运行时间约为 0.32 ms。

为完善机动目标拦截模型，下一步将考虑拦截器姿态控制动态特性，研究基于 DRL 的制导控制一体化设计方法。

参考文献（References）

[1] GUELMAN M. A qualitative study of proportional navigation[J]. IEEE Transactions on Aerospace and Electronic Systems, 1971, AES-7(4): 637-643.

[2] GHAWGHAWE S N, GHOSE D. Pure proportional navigation against time-varying target manoeuvres[J]. IEEE Transactions On Aerospace and Electronic Systems, 1996, 32(4): 1336-1347.

[3] CHEN W X, GAO C S, JING W X. Proximal policy optimization guidance algorithm for intercepting near-space maneuvering targets[J]. Aerospace Science and Technology, 2023, 132: 108031.

[4] CHEN Z, SHIMA T. Nonlinear optimal guidance for intercepting a stationary target[J]. Journal of Guidance, Control, and Dynamics, 2019, 42(11): 2418-2431.

[5] HE L, YAN X D, TANG S. Spiral-diving trajectory optimization for hypersonic vehicles by second-order cone programming[J].

Aerospace Science and Technology, 2019, 95: 105427.

[6] HE L, YAN X D. Adaptive terminal guidance law for spiral-diving maneuver based on virtual sliding targets[J]. Journal of Guidance, Control, and Dynamics, 2018, 41(7): 1591-1601.

[7] YANUSHEVSKY R. Analysis of optimal weaving frequency of maneuvering targets[J]. Journal of Spacecraft and Rockets, 2004, 41(3): 477-479.

[8] QIU X Q, LAI P, GAO C S, et al. Recorded recurrent deep reinforcement learning guidance laws for intercepting endoatmospheric maneuvering missiles[J]. Defence Technology, 2024, 31: 457-470.

[9] KUMAR S R, RAO S, GHOSE D. Nonsingular terminal sliding mode guidance with impact angle constraints[J]. Journal of Guidance, Control, and Dynamics, 2014, 37(4): 1114-1130.

[10] EBRAHIMI B, BAHRAMI M, ROSHANIAN J. Optimal sliding-mode guidance with terminal velocity constraint for fixed-interval propulsive maneuvers[J]. Acta Astronautica, 2008, 62(10-11): 556-562.

[11] HE S M, LIN D F, WANG J. Robust terminal angle constraint guidance law with autopilot lag for intercepting maneuvering targets[J]. Nonlinear Dynamics, 2015, 81(1): 881-892.

[12] ASHER R B, MATUSZEWSKI J P. Optimal guidance with maneuvering targets[J]. Journal of Spacecraft and Rockets, 1974, 11(3): 204-206.

[13] CHO H, RYOO C K, TSOURDOS A, et al. Optimal impact angle control guidance law based on linearization about collision triangle[J]. Journal of Guidance, Control, and Dynamics, 2014, 37(3): 958-964.

[14] LIANG H Z, WANG J Y, WANG Y H, et al. Optimal guidance against active defense ballistic missiles via differential game strategies[J]. Chinese Journal of Aeronautics, 2020, 33(3): 978-989.

[15] SHALUMOV V. Optimal cooperative guidance laws in a multiagent target-missile-defender engagement[J]. Journal of Guidance, Control, and Dynamics, 2019, 42(9): 1993-2006.

[16] LIANG H, WANG J, LIU J, et al. Guidance strategies for interceptor against active defense spacecraft in two-on-two engagement[J]. Aerospace Science and Technology, 2020, 96: 105529.

[17] XU Z Y, CHEN Y K, XU Z X. Optimal guidance and collision avoidance for docking with the rotating target spacecraft[J]. Advances in Space Research, 2019, 63(10): 3223-3234.

[18] ZHENG Y, CHEN Z, SHAO X M, et al. Time-optimal guidance for intercepting moving targets with impact-angle constraints[J]. Chinese Journal of Aeronautics, 2022, 35(7): 157-167.

[19] GAUDET B, FURFARO R. Missile homing-phase guidance law design using reinforcement learning[C]//Proceedings of the AIAA Guidance, Navigation, and Control Conference. Minneapolis, Reston: AIAA, 2012: 4470.

[20] LILLICRAP T P, HUNT J J, PRITZEL A, et al. Continuous control with deep reinforcement learning[EB/OL]. (2019-07-15)[2024-01-10]. https://doi.org/10.48550/arXiv.1509.02971.

[21] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Playing Atari with deep reinforcement learning[EB/OL]. (2013-12-19)[2024-01-05]. https://doi.org/10.48550/arXiv.1312.5602.

[22] SCHULMAN J, WOLSKI F, DHARIWAL P, et al. Proximal policy optimization algorithms[EB/OL]. (2017-08-28)[2024-01-12]. https://doi.org/10.48550/arXiv.1707.06347.

[23] HAARNOJA T, ZHOU A, ABBEEL P, et al. Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor[EB/OL]. (2018-08-08)[2024-01-13]. https://doi.org/10.48550/arXiv.1801.01290.

[24] SILVER D, HUANG A, MADDISON C J, et al. Mastering the game of Go with deep neural networks and tree search[J]. Nature, 2016, 529(7587): 484-489.

[25] SILVER D, SCHRITTWIESER J, SIMONYAN K, et al. Mastering the game of Go without human knowledge[J]. Nature, 2017, 550(7676): 354-359.

[26] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning[J]. Nature, 2015, 518(7540): 529-533.

[27] VINYALS O, BABUSCHKIN I, CZARNECKI W M, et al. Grandmaster level in StarCraft II using multi-agent reinforcement learning[J]. Nature, 2019, 575(7782): 350-354.

[28] WANG X, WANG S, LIANG X X, et al. Deep reinforcement learning: a survey[J]. IEEE Transactions on Neural Networks and Learning Systems, 2024, 35(4): 5064-5078.

[29] ARULKUMARAN K, DEISENROTH M P, BRUNDAGE M, et al. Deep reinforcement learning: a brief survey[J]. IEEE Signal Processing Magazine, 2017, 34(6): 26-38.

[30] GAUDET B, LINARES R, FURFARO R. Deep reinforcement learning for six degree-of-freedom planetary landing[J]. Advances in Space Research, 2020, 65(7): 1723-1741.

[31] GAUDET B, LINARES R, FURFARO R. Adaptive guidance and integrated navigation with reinforcement meta-learning[J]. Acta Astronautica, 2020, 169: 180-190.

[32] FURFARO R, SCORSOGLIO A, LINARES R, et al. Adaptive generalized ZEM-ZEV feedback guidance for planetary landing via a deep reinforcement learning approach[J]. Acta Astronautica, 2020, 171: 156-171.

[33] SCORSOGLIO A, FURFARO R, LINARES R, et al. Relative motion guidance for near-rectilinear lunar orbits with path constraints via actor-critic reinforcement learning[J]. Advances in Space Research, 2023, 71(1): 316-335.

[34] GONG X P, CHEN W C, CHEN Z Y. All-aspect attack guidance law for agile missiles based on deep reinforcement learning[J]. Aerospace Science and Technology, 2022, 127: 107677.

[35] HONG D, KIM M, PARK S, et al. Study on reinforcement learning-based missile guidance law[J]. Applied Sciences, 2020, 10(18): 6567.

[36] YUAN H, LI D X. Deep reinforcement learning for rendezvous guidance with enhanced angles-only observability[J]. Aerospace Science and Technology, 2022, 129: 107812.

[37] HE X J, CHEN Z H, JIA F, et al. Guidance law based on zero effort miss and Q-learning algorithm[C]//Proceedings of the Seventh Symposium on Novel Photoelectronic Detection Technology and Applications. Bellinham: SPIE, 2021: 117.

[38] HE X J, WU M Y, CHEN Z H, et al. Guidance law based on deep Q network algorithm[J]. Journal of Physics: Conference Series, 2022, 2235(1): 012107.

[39] HE S M, SHIN H S, TSOURDOS A. Computational missile guidance: a deep reinforcement learning approach[J]. Journal of Aerospace Information Systems, 2021, 18(8): 571-582.

[40] DU M J, PENG C, MA J J. Deep reinforcement learning based missile guidance law design for maneuvering target interception [C]//Proceedings of the 2021 40th Chinese Control Conference. Piscataway: IEEE Press, 2021: 3733-3738.

[41] 张秦浩, 敖百强, 张秦雪. Q-learning 强化学习制导律[J]. 系统工程与电子技术, 2020, 42(2): 414-419.
ZHANG Q H, AO B Q, ZHANG Q X. Reinforcement learning guidance law of Q-learning[J]. Systems Engineering and Electronics, 2020, 42(2): 414-419(in Chinese).

[42] 刘扬, 何泽众, 王春宇, 等. 基于 DDPG 算法的末制导律设计研究[J]. 计算机学报, 2021, 44(9): 1854-1865.
LIU Y, HE Z Z, WANG C Y, et al. Terminal guidance law design based on DDPG algorithm[J]. Chinese Journal of Computers, 2021, 44(9): 1854-1865(in Chinese).

[43] TANG J, BAI Z H, LIANG Y G, et al. An exoatmospheric homing guidance law based on deep Q network[J]. International Journal of Aerospace Engineering, 2022, 2022(1): 1544670.

[44] LIANG Y G, TANG J, BAI Z H, et al. Homing guidance law design against maneuvering targets based on DDPG[J]. International Journal of Aerospace Engineering, 2023, 2023(1): 4188037.

[45] WANG W W, WU M Y, CHEN Z H, et al. Integrated guidance-and-control design for three-dimensional interception based on deep-reinforcement learning[J]. Aerospace, 2023, 10(2): 167.

[46] LUO W L, CHEN L, LIU K X, et al. Optimizing constrained guidance policy with minimum overload regularization[J]. IEEE Transactions on Circuits and Systems I: Regular Papers, 2022, 69(7): 2994-3005.

[47] GAUDET B, FURFARO R, LINARES R. Reinforcement learning for angle-only intercept guidance of maneuvering targets[J]. Aerospace Science and Technology, 2020, 99: 105746.

[48] GAUDET B, FURFARO R, LINARES R, et al. Reinforcement metalearning for interception of maneuvering exoatmospheric targets with parasitic attitude loop[J]. Journal of Spacecraft and Rockets, 2021, 58(2): 386-399.

[49] 邱潇颀, 高长生, 荆武兴. 拦截大气层内机动目标的深度强化学习制导律[J]. 宇航学报, 2022, 43(5): 685-695.
QIU X Q, GAO C S, JING W X. Deep reinforcement learning guidance law for intercepting endo-atmospheric maneuvering targets[J]. Journal of Astronautics, 2022, 43(5): 685-695 (in Chinese).

[50] 陈文雪, 高长生, 荆武兴. 拦截机动目标的信赖域策略优化制导算法[J]. 航空学报, 2023, 44(11): 327596.
CHEN W X, GAO C S, JING W X. Trust region policy optimization guidance algorithm for intercepting maneuvering target[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(11): 327596(in Chinese).

[51] FUJIMOTO S, VAN HOOF H, MEGER D. Addressing function approximation error in actor-critic methods[C]//Proceedings of the International Conference on Machine Learning. Cambridge: PMLR, 2018: 1587-1596.

[52] JOHN SCHULMAN S L P M. Trust region policy optimization[EB/OL]. (2017-04-20)[2024-03-15]. https://doi.org/10.48550/arXiv.1502.05477.

[53] SCHULMAN J, MORITZ P, LEVINE S, et al. High-dimensional continuous control using generalized advantage estimation[EB/OL]. (2018-08-20)[2024-03-16]. https://doi.org/10.48550/arXiv.1506.02438.

[54] XIE J W, CHEN W C. Switching logic design for divert and attitude control system of exoatmospheric kill vehicle[C]//Proceedings of the 2017 IEEE International Conference on Cybernetics and Intelligent Systems and IEEE Conference on Robotics, Automation and Mechatronics. Piscataway: IEEE Press, 2017: 194-200.

[55] GOLDMAN R. Understanding quaternions[J]. Graphical Models, 2011, 73(2): 21-49.

[56] XIAN Y, REN L L, XU Y J, et al. Impact point prediction guidance of ballistic missile in high maneuver penetration condition[J]. Defence Technology, 2023, 26: 213-230.

[57] KARLIK B, OLGAC A V. Performance analysis of various activation functions in generalized MLP architectures of neural networks[J]. International Journal of Artificial Intelligence and Expert Systems, 2011, 1(4): 111-122.

[58] YU J, HESTHAVEN J S. Flowfield reconstruction method using artificial neural network[J]. AIAA Journal, 2019, 57(2): 482-498.

[59] ENGSTROM L, ILYAS A, SANTURKAR S, et al. Implementation matters in deep policy gradients: a case study on PPO and TRPO[EB/OL]. (2020-05-25)[2024-04-20]. https://doi.org/10.48550/arXiv.2005.12729.

[60] ZIPFEL P H, SCHIEHLEN W. Modeling and simulation of aerospace vehicle dynamics[J]. Applied Mechanics Reviews, 2001, 54(6): B101-B102.

[61] SAXE A M, MCCLELLAND J L, GANGULI S. Exact solutions to the nonlinear dynamics of learning in deep linear neural networks[EB/OL]. (2014-02-19)[2024-03-22]. https://doi.org/10.48550/arXiv.1312.6120.

[62] 任乐亮, 鲜勇, 李少朋, 等. 基于改进二阶优化器并行学习的弹道导弹神经网络落点预测方法[J]. 航空学报, 2023, 44(14): 227-245.
REN L L, XIAN Y, LI S P, et al. A neural network model for impact point prediction of ballistic missile based on improved second-order optimizer with parallel learning[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(14): 227-245(in Chinese).

Reinforcement learning guidance law for maneuvering target interception based on imitation learning

REN Leliang[1, 2]，XIAN Yong[1, 2]，LIU Zhenyu[1]，ZHANG Daqiao[3, *]，LI Bing[3]，LI Shaopeng[1, 4]

(1. College of Missile Engineering，Rocket Force University of Engineering，Xi'an 710025，China；
2. Key Laboratory of Cross-Domain Flight Interdisciplinary Technology，Mianyang 621000，China；
3. College of Operational Support，Rocket Force University of Engineering，Xi'an 710025，China；
4. Department of Automation，Tsinghua University，Beijing 100084，China)

Abstract：The advancement of maneuver penetration technology necessitates an improvement in the design of interception guidance laws to a higher level. An echelon intelligent guidance framework of "terminal guidance law model based on imitation learning (IL) → terminal guidance law evolution model based on reinforcement learning" is proposed to increase the interception probability, decrease the energy consumption, and improve the robustness of the guidance law for intercepting maneuvering targets. Firstly, a three-dimensional uncertain confrontation model between the maneuvering target and interceptor is established based on the interception collision triangle. Secondly, the IL method is used to mine the proportional navigation guidance (PNG) law, which provides a good initial policy for the subsequent reinforcement learning guidance law. Finally, a Markov decision model is established, and a process reward of energy consumption and a "soft" terminal reward model, including a "transition section," are proposed. The proximal policy optimization (PPO) algorithm is used to fully explore the high-performance interception strategy. The findings of the Monte Carlo simulation show that the new guidance law is very stable and resilient, outperforming the conventional guidance algorithm in terms of interception probability and energy usage. Additionally, the single decision time is only 0.32 ms, making it of certain engineering value.

Keywords：imitation learning；deep reinforcement learning；maneuvering target interception；proximal policy optimization；collision triangle

Received：2024-05-06；**Accepted**：2024-09-27；**Published Online**：2025-01-06 16：11

URL：link.cnki.net/urlid/11.2625.V.20250106.1129.001

Foundation items：National Natural Science Foundation of China (62103432); Young Talent Fund of University Association for Science and Technology in Shaanxi, China (20210108) ； Open Fund of Key Laboratory of Cross-Domain Flight Interdisciplinary Technology (2024-KYKF-4004)

* **Corresponding author**. E-mail：zhangdq2012@pku.org.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0278

强阵风下大柔性机翼的非线性流固耦合响应分析

苏进鑫[1,2]，席梓严[1,2]，戴玉婷[1,2,*]

(1. 天目山实验室，杭州 310023；2. 北京航空航天大学 航空科学与工程学院，北京 100191)

摘 要：大柔性机翼在较大阵风来流下会产生显著变形，其中，几何非线性和气动非线性不可忽略。为研究大柔性机翼在强阵风下的响应，建立了基于几何非线性梁理论与计算流体力学(CFD)方法的流固耦合分析方法。通过对比倒置柔性板风洞试验，验证了该流固耦合仿真的准确性。在此基础上，建立了半展弦比为9的大柔性机翼流体模型和几何非线性梁模型，在雷诺数为10^5的条件下，对其进行阵风响应计算，研究了不同阵风比下(R_G=0.1～0.4)大柔性机翼的动响应特性，特别是对大阵风比(R_G=0.4)时阵风诱导的机翼动态失速气动特性进行分析。研究结果表明：随着阵风比的增大，机翼出现了显著的非线性弯曲变形，在R_G=0.4时，机翼的最大弯曲变形达63%半展长，同时，机翼的弯曲变形使前缘涡(LEV)的展向分布宽度减小，影响了气动力的分布，降低了阵风载荷。

关 键 词：大柔性机翼；非线性；流固耦合；阵风响应；动态失速
中图分类号：V221$^+$.3; TB553
文献标志码：A **文章编号**：1001-5965(2026)06-2172-12

近年来，高空长航时无人机和大型运输机在航空领域的应用价值日益凸显。这类飞机机翼柔性大、质量轻，在大气中飞行时对阵风扰动非常敏感，在强阵风作用下会产生非常大的结构变形，气动弹性问题突出。尤其当其所遇阵风速度与飞行速度相当时，流动分离更加剧烈，阵风可能会导致其结构损坏。因此，有必要对大柔性机翼遭遇强阵风情况下的动响应进行分析，以避免不利影响。

目前，对大幅阵风与机翼相互作用的研究主要集中在翼型和刚性机翼上[1-3]，不考虑机翼的结构响应。在大柔性机翼阵风响应方面也有一些相关研究，Tang等[4]通过风洞试验和解析方法研究了大柔性机翼在阵风作用下的非线性动响应问题，为后续研究提供了重要的理论和实验依据。Neumann等[5]介绍了弹性机翼的阵风响应实验，并建立了有限元模型进行流固耦合仿真，发现机翼有限元模型及风洞壁的适应性对数值分析结果有重大影响，但未系统地对流场进行分析。An等[6]建立了一个基于降阶模型(reduced order model, ROM)的非线性大展弦比机翼阵风响应的分析框架，所提方法可用于解决含有几何非线性机翼的气动弹性响应问题。Zheng等[7]基于流固耦合方法，在不同阵风比下研究了主动翼端弯曲变形对减缓阵风的作用，其研究证明，主动弯曲变形可以有效抑制阵风。上述研究都集中于大柔性机翼的阵风响应分析方法及发展阵风减缓技术。关于大柔性强阵风条件下机翼变形对气动特性和性能影响的研究十分有限。

在研究大变形结构流固耦合响应时，计算流体力学(computational fluid dynamics, CFD)和计算结构动力学(computational structural dynamics, CSD)耦合方法可以同时考虑几何和气动非线性，展现出良好的性能。陈利丽等[8]用CFD/CSD耦合方法对柔性扑翼飞行器进行了流固耦合研究。Mian等[9]开发了一种CFD/CSD耦合方法，用于研究大柔性机

收稿日期：2024-05-06；录用日期：2024-08-23；网络出版时间：2024-11-27 13:31
网络出版地址：link.cnki.net/urlid/11.2625.V.20241126.1801.004
*通信作者. E-mail：yutingdai@buaa.edu.cn

引用格式：苏进鑫，席梓严，戴玉婷. 强阵风下大柔性机翼的非线性流固耦合响应分析[J]. 北京航空航天大学学报，2026，52(6)：2172-2183. SU J X，XI Z Y，DAI Y T. Nonlinear fluid-structure interaction response analysis of a large flexible wing under strong gusts[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：2172-2183 (in Chinese).

翼的非线性静气动弹性行为。聂雪媛等[10]基于CFD/CSD耦合比较分析了柔性飞行器线性和非线性结构的静气动弹性特性。吴金洋等[11]基于CFD/CSD耦合方法探讨了二维阵风载荷对临近空间太阳能飞行器着陆状态阵风响应的影响。上述研究结果与现有实验数据在定性和定量上均取得了良好的一致性, 说明 CFD/CSD 耦合方法可以准确反映大柔性结构在非线性条件下的动响应特性。本文基于 CFD/CSD 松耦合方法主要研究不同阵风幅度下大柔性机翼的展向弯曲变形及展向变形对阵风载荷分布的影响。

1　研究方法

1.1　非线性结构动力学模型

在求解大变形梁的动响应时, 需要考虑梁结构的几何非线性, 基于绝对节点坐标公式 (absolute nodal coordinate formulation, ANFC)算法和几何精确梁理论[12-15]进行结构动力学计算。

传统的线性有限元方法中, 节点的坐标通常是相对的, ANCF 采用绝对节点坐标, 大变形和大旋转可以通过改变节点的绝对坐标来直接模拟。且梁的每个节点可以包含多个自由度, 这使模型能够直接精确地描述梁结构在空间中的实际变形和运动, 而不需要在局部坐标系之间进行转换。

如图 1 所示, 在绝对节点坐标系下, 梁上任意一点 p 的坐标可以表示为

$$\boldsymbol{r}^p(x,y,z,t)=\boldsymbol{r}(x,t)+\boldsymbol{A}(x,t)\boldsymbol{d} \tag{1}$$

式中: $\boldsymbol{r}(x,t)$为梁参考线在轴向坐标 x、时间 t 处的位置矢量。$\boldsymbol{d}=(0,y,z)^{\mathrm{T}}$定义了点在横截面参考系中的坐标; 转换矩阵$\boldsymbol{A}(x,t)$由旋转向量$\boldsymbol{\theta}(x,t)$描述:

$$\boldsymbol{A}=\boldsymbol{I}+f(\boldsymbol{\theta})\boldsymbol{\Theta}+g(\boldsymbol{\theta})\boldsymbol{\Theta}^2 \tag{2}$$

式中: $\boldsymbol{I}$为 3×3 的单位矩阵; $\boldsymbol{\Theta}$为由旋转向量$\boldsymbol{\theta}$生成的反对称矩阵。

$$\begin{cases}\boldsymbol{\theta}=\|\boldsymbol{\theta}\|=\sqrt{\theta_1^2+\theta_2^2+\theta_3^2}\\ \boldsymbol{\Theta}=\begin{bmatrix}0&-\theta_3&\theta_2\\ \theta_3&0&-\theta_1\\ -\theta_2&\theta_1&0\end{bmatrix}\\ f(\boldsymbol{\theta})=\begin{cases}\dfrac{\sin\theta}{\theta} & |\boldsymbol{\theta}|\geqslant 0.01\\ 1-\dfrac{\theta^2}{3!}+\dfrac{\theta^4}{5!}-\dfrac{\theta^6}{7!}+\ldots & |\boldsymbol{\theta}|<0.01\end{cases}\\ g(\boldsymbol{\theta})=\dfrac{1}{2}f^2\left(\dfrac{\boldsymbol{\theta}}{2}\right)\end{cases} \tag{3}$$

梁单元的广义坐标可以表示为

$$\boldsymbol{q}^{\mathrm{T}}=(\boldsymbol{r}_0^{\mathrm{T}},\boldsymbol{r}_0'^{\mathrm{T}},\boldsymbol{\theta}_0^{\mathrm{T}},\boldsymbol{\theta}_{\mathrm{m}}^{\mathrm{T}},\boldsymbol{r}_1^{\mathrm{T}},\boldsymbol{r}_1'^{\mathrm{T}},\boldsymbol{\theta}_1^{\mathrm{T}})^{\mathrm{T}} \tag{4}$$

因此, 每个三维梁单元有 21 个广义坐标。梁的控制方程可以从拉格朗日方程中推导出来:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial\dot{\boldsymbol{q}}_i}-\frac{\partial T}{\partial\boldsymbol{q}_i}+\frac{\partial V}{\partial\boldsymbol{q}_i}+\frac{\partial\boldsymbol{\Psi}^{\mathrm{T}}}{\partial\boldsymbol{q}_i}\lambda=Q_i \tag{5}$$

式中: T为梁的总动能; V为梁的总弹性势能; $\boldsymbol{\Psi}$为约束方程; λ为约束方程的拉格朗日乘子; Q为广义应用力。结构的几何非线性主要体现转换矩阵$\boldsymbol{A}$、动能T及势能V的计算中[16]。

结构动力学方程采用 Newmark 方法进行求解, 对动力学方程进行离散后, 在每个时间步长内采用 Newton-Raphson 迭代。

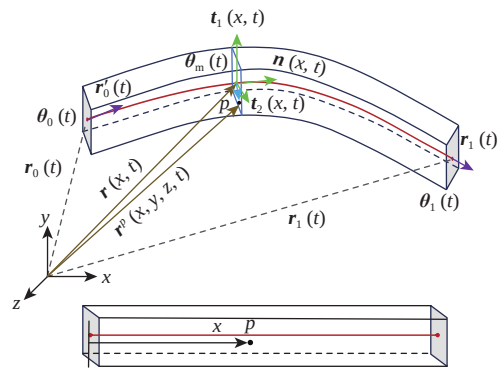


图 1　3D 梁单元广义坐标

Fig. 1　Generalized coordinates of 3D beam element

1.2　气动模型

非定常不可压缩雷诺平均纳维-斯托克斯(Reynolds-averaged Navier-Stokes equations, RANS)方程为

$$\begin{cases}\rho\dfrac{\partial\overline{\boldsymbol{U}}}{\partial t}+\rho\overline{\boldsymbol{U}}\nabla\overline{\boldsymbol{U}}=-\nabla\overline{p}+\mu\nabla^2\overline{\boldsymbol{U}}-\rho\nabla\overline{\boldsymbol{U}'\,\boldsymbol{U}'}\\ \nabla\overline{\boldsymbol{U}}=\boldsymbol{0}\end{cases} \tag{6}$$

式中: $\overline{\boldsymbol{U}}$、$\overline{\boldsymbol{U}'}$分别为速度矢量的时均分量和脉动分量; μ为空气的动力黏度系数; $\overline{p}$为压强的时均分量; 附加项$-\rho\nabla\overline{\boldsymbol{U}'\,\boldsymbol{U}'}$为雷诺应力。

本文中湍流模型采用二方程的剪切应力输运 SST k-ω 模型[17]。

采用压力基算法求解上述方程, 时间项用一阶隐式格式, 梯度和力用二阶格式, 动量、湍动能和耗散项用二阶迎风格式, 如表 1 所示。

表 1　时间和空间离散方案

Table 1　Time and space discrete schemes

设置项目	方案
瞬态	一阶隐式
梯度	基于单元的最小二乘法
压力	二阶
动量	二阶迎风
湍动能	二阶迎风
比耗散率	二阶迎风

1.3 耦合方法

流固耦合基于商用软件 Fluent 和 Adams 进行，仿真流程如图 2 所示，采用弱耦合计算方法，其中，CFD 网格变形技术采用线弹性实体光顺法，该方法有较强的网格变形能力，且变形后网格质量较高。流场与结构之间需要转换的信息包括：机翼壁面每个网格单元上的流体表面力插值到结构网格节点上、结构表面上每个单元节点的位移插值到流体网格节点上。

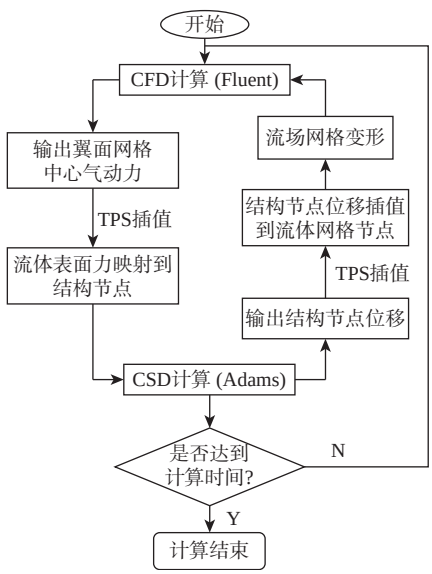


图 2 CFD/CSD 弱耦合仿真流程

Fig. 2 CFD/CSD weak coupling simulation process

由于本文中计算对象为梁式机翼，故采用薄板样条 (thin plate spline, TPS) 插值法[18] 进行力和位移的插值，TPS 插值的一般形式为

$$F(x,y,z)=a_0+a_1x+a_2y+a_3z+\sum_{i=1}^{N}A_i\phi(R_i) \tag{7}$$

式中：a_0、a_1、a_2、a_3 为常数；$R_i=\sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2}$ 表示点 (x,y,z) 到点 (x_i,y_i,z_i) 的距离；$\phi(r_i)=r_i^2\ln r_i^2$ 为薄板样条函数。通过求解线性方程组得到权重 A_i 后，就可以通过式 (6) 计算 $F(x,y,z)$ 的值。TPS 方法的核心思想就是将插值函数表示为线性和非线性 2 个部分的组合，线性部分是由平面拟合而来，非线性部分是根据薄板样条函数计算而来。

2 流固耦合算法验证

2.1 验证模型

通过 Kim 等[19] 的倒置柔性板风洞试验验证流固耦合仿真方法，该柔性板几何和材料属性如表 2 所示。

流固耦合仿真做如下假设：① 忽略重力影响；② 只考虑 XY 平面内的弯曲变形，不考虑扭转变形。

表 2 倒置柔性板参数

Table 2 Parameters of inverted flexible plate

弹性模量/GPa	泊松比	密度/(kg·m^{-3})	板长/m	板宽/m	板厚/m
2.38	0.38	1 200	0.27	0.3	0.000 8

倒置柔性板有限元模型如图 3 所示，采用四面体网格单元，平板的运动限制在 XY 平面内，平板右侧根部固支。

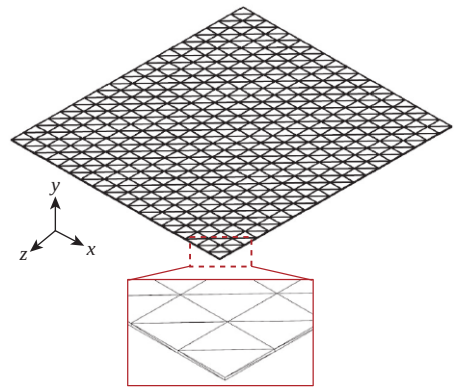


图 3 倒置柔性板有限元模型

Fig. 3 Finite element model of the inverted flexible plate

2.2 倒置柔性板流域建立及网格划分

流场域及网格划分如图 4 所示，其中，L 为倒置柔性板长度。设置速度入口、压力出口和壁面边界条件，梁的自由端靠近速度入口。采用混合网格，大变形梁附近的网格使用非结构网格，外流场采用结构网格，网格数为 33 527。

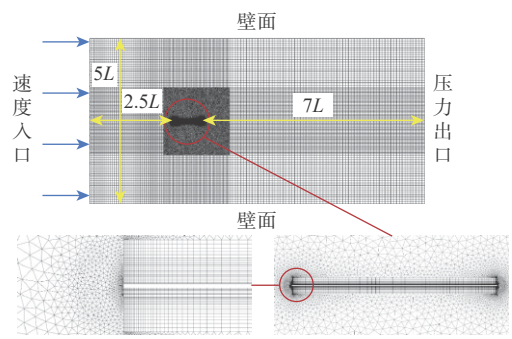


图 4 流域及网格

Fig. 4 Flow field and mesh

2.3 验证结果

当来流速度 U=4.2 m/s 时，倒置柔性板自由端位移曲线如图 5 所示。对比图中 4 个周期内仿真结果和文献 [19] 实验的结果可以看出，流固耦合仿真的自由端摆动幅值为 0.432 m，实验结果的自由端摆动幅值为 0.405 m，误差为 6.7%。流固耦合仿真的自由端摆动频率为 1.33 Hz，实验结果的自由端摆动频率为 1.29 Hz，误差为 3.1%。在 t=0.77 s 倒置柔性板弯曲变形最大时的流场网格如图 6 所示。

可知, TPS 方法结合 CFD 线弹性实体光顺法动网格变形技术在处理大变形问题时效果较好, 既不会产生负体积元素, 又能保持网格的正交性。综上, 该流固耦合仿真的准确性和适用性较高, 可用于后续大柔性机翼阵风响应仿真。

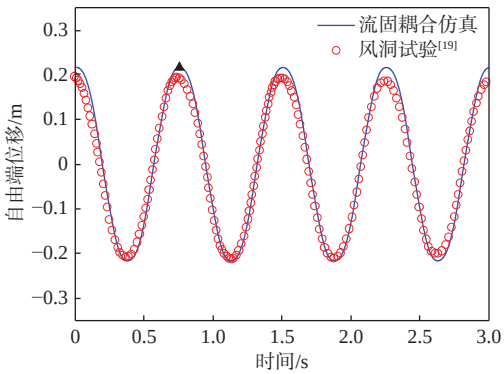


图 5　U=4.2 m/s 时仿真与实验位移对比

Fig. 5　Comparison of simulation and experimental displacements at U=4.2 m/s

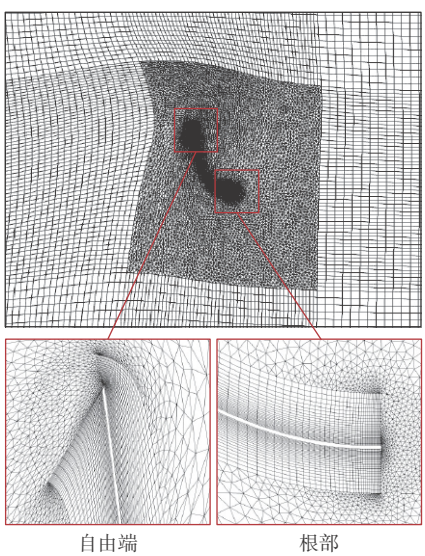


图 6　t=0.77 s 时流场网格变形示意

Fig. 6　Schematic of flow field grid deformation at t=0.77 s

不同来流速度下, 仿真所得倒置柔性板摆动时自由端位移随时间的变化结果如图 7 所示。

当来流速度小于 3.5 m/s(第 1 临界速度)时, 倒置柔性板处于直线模式, 如图 8(a) 所示; 当来流速度达到第 1 临界速度时, 倒置柔性板从直线模式进入摆动模式, 如图 8(b)～8(d) 所示; 当来流速度大于 6.2 m/s(第 2 临界速度)时, 倒置柔性板进入偏置模式, 如图 8(e) 和 8(f) 所示。仿真结果与实验规律(流速小于 3.9 m/s 时, 倒置柔性板为直线模式, 流速大于 6.7 m/s 时为偏置模式, 在 3.9～6.7 m/s 时, 倒置柔性板为摆动模式)基本一致, 进一步验证了流固耦合仿真的准确性。

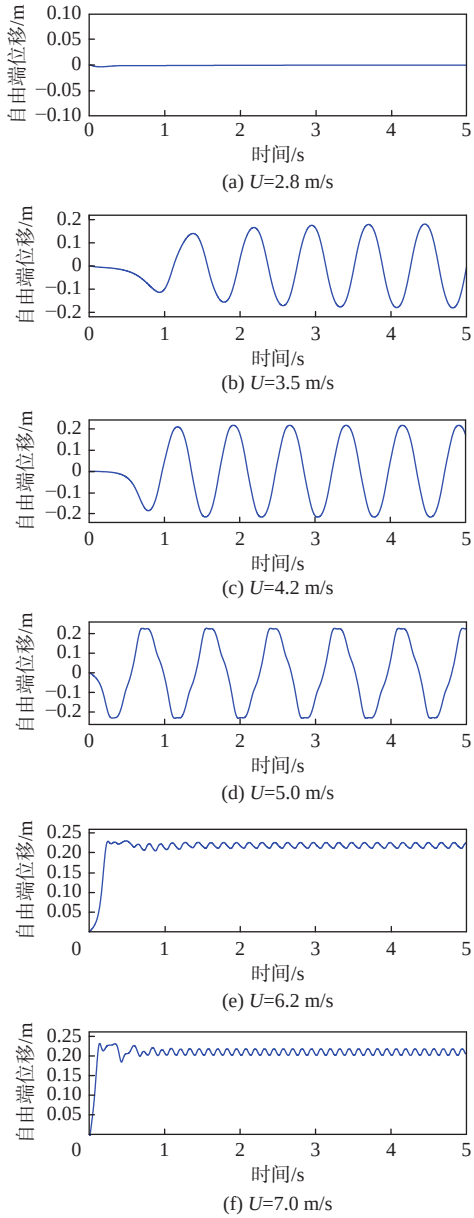


图 7　不同来流速度下倒置柔性板自由端位移曲线

Fig. 7　Displacement curves of free end of an inverted flexible plate under different incoming flow velocities

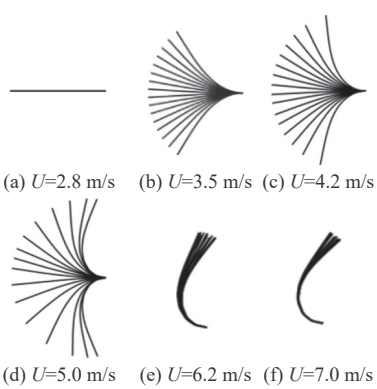


图 8　仿真所得不同流速下倒置柔性板动态轮廓叠加

Fig. 8　Dynamic contour superposition of an inverted flexible plate under different flow velocities obtained from simulation

3 大柔性机翼阵风响应分析

3.1 阵风模型及参数

阵风模型如图9所示，阵风是正弦阵风，该阵风的解析表达式为

$$w_g = \overline{w}_g \sin(2\pi f_1 t) \tag{8}$$

阵风比为

$$R_G = \frac{\overline{w}_g}{U} \tag{9}$$

式中：U为来流速度；w_g为阵风速度；$\overline{w}_g$为阵风速度幅值；f_1为阵风频率。

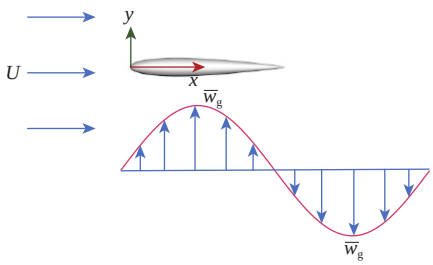


图9 阵风模型

Fig. 9 Gust model

3.2 仿真模型

在上述流固耦合验证的基础上，计算一个长直矩形机翼的阵风响应，参考Tang等[4]风洞试验模型，机翼剖面采用NACA0012翼型，机翼由柔性主梁、翼肋和翼尖配重构成，沿展向均匀布置18片翼肋。采用刚柔耦合结构，其中，翼尖配重和翼肋为刚体，翼梁采用三维21自由度大柔性梁有限元模型。机翼几何模型及参数如图10所示，材料属性如表3所示。

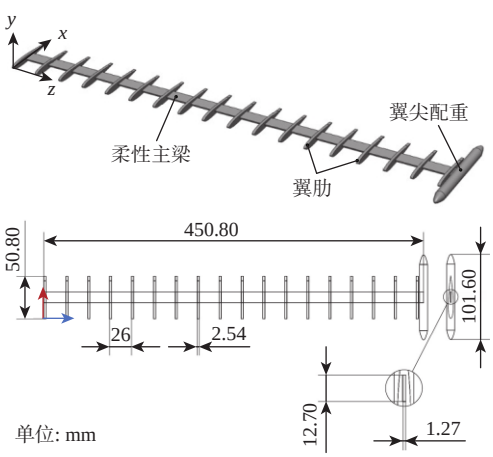


图10 机翼几何模型

Fig. 10 Geometrical model of the wing

机翼的前四阶模态振型如图11所示，各阶模态固有频率如表4所示。

表3 机翼材料属性

Table 3 Material properties of the wing

结构	密度/(kg·m^{-3})	弹性模量/GPa	泊松比
柔性体翼梁	5 800	193.1	0.38
刚体翼肋	7 125		
刚体翼尖配重	6 525		

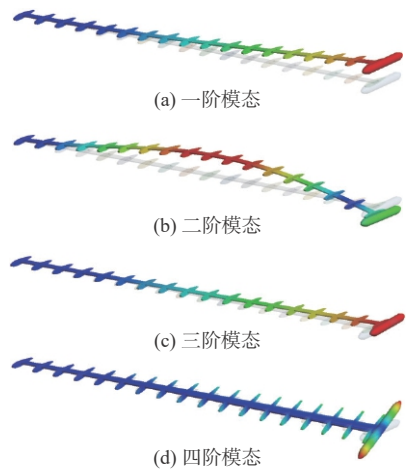


图11 机翼前四阶模态振型

Fig. 11 The first four modes shape of the wing

表4 前四阶模态频率

Table 4 Frequency of the first four modes

模态阶数	模态描述	模态频率/Hz
一	机翼面外一弯	2.46
二	机翼面外二弯	19.54
三	机翼面内一弯	23.87
四	机翼一扭	38.21

3.3 机翼流域建立及网格划分

考虑到结构变形较大，故采用重叠网格技术[20]进行流场网格划分。前景网格及翼尖附近网格如图12所示，流场网格划分忽略翼尖配重部分，为提高网格质量，对三维机翼翼尖进行了圆滑处理。网格坐标原点位于翼根附近的前缘点，即图10中的

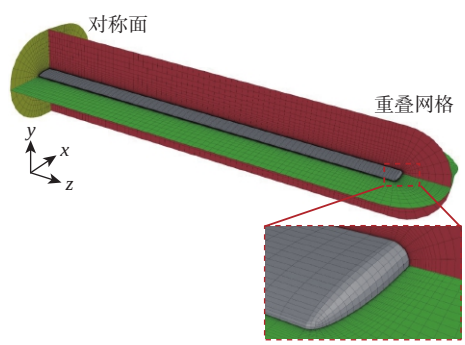


图12 机翼前景网格

Fig. 12 Foreground grid around the wing

坐标系 *Oxyz*。

图 13 显示了背景网格大小和边界条件。背景网格中与前景网格重叠的区域被加密。

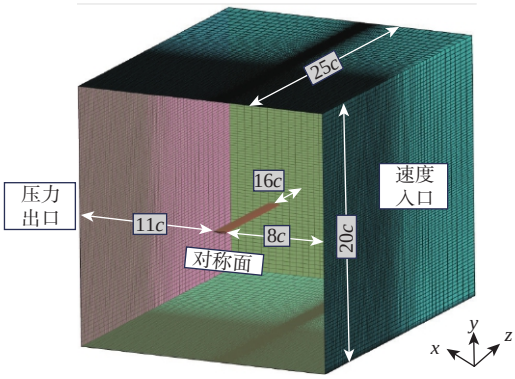


图 13　背景网格及流场域

Fig. 13　Background grid and flow field

为评估所选网格和时间步长的独立性，进行网格和时间步长收敛性验证。通过改变网格的壁面距离 Y^+值和网格增长率来验证网格收敛性，分别生成粗网格、中等网格和细网格 3 种前景网格，网格参数如表 5 所示，网格拓扑结构与图 12 所示的前景网格结构类似。在中等网格条件下，分别设置时间步长为0.5×10^{-3} s、1×10^{-4} s和0.5×10^{-4} s来验证时间步长收敛性。在来流速度 $U=25$ m/s、垂向阵风速度 $w_g=10\sin(6\pi t)$的工况下取 3 套网格计算升力，阵风脉动施加在速度入口处。一个阵风周期 T 内升力随时间的变化趋势如图 14 所示，综合考虑计算时间和精度，网格选取中等网格，时间步长取 1×10^{-4} s。

表 5　网格参数

Table 5　Parameters of the mesh

网格名称	背景网格数量	前景网格数量	壁面距离	网格增长率
粗网格	862 580	294 524	1.5	1.23
中等网格	862 580	355 460	1.0	1.20
细网格	862 580	457 020	0.5	1.17

3.4　仿真结果

当来流速度 U=25 m/s、雷诺数 Re=10^5 时，各仿真工况如表 6 所示，其中，阵风频率接近机翼的一阶弯曲模态固有频率。表中“机翼刚性”即不考虑流固耦合，直接计算刚性机翼阵风响应。

仿真时，将工况 1、3 下的刚性机翼模型作为基准模型，与工况 2、4 下的柔性机翼模型结果进行对比。

1）$R_G=0.1$。图 15 是工况 1、2 下机翼的升力随时间的变化曲线。可知，在小幅阵风作用下，柔性机翼的升力幅值小于刚性机翼，变化规律与阵风速度表达式类似，均为正弦形式。

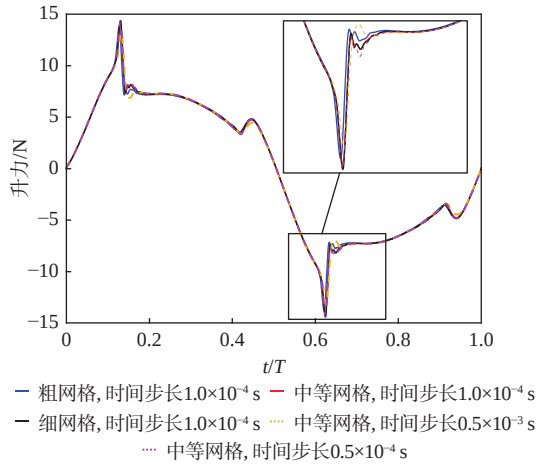


图 14　网格和时间步长收敛性曲线

Fig. 14　Grid and time-step convergence curves

表 6　仿真工况参数

Table 6　Simulated operating conditions parameters

工况编号	机翼刚柔性	机翼迎角/(°)	阵风速度幅值/(m·s^{-1})	R_G	阵风频率/Hz
1	刚性	0	2.5	0.1	3
2	柔性	0	2.5	0.1	3
3	刚性	0	10	0.4	3
4	柔性	0	10	0.4	3

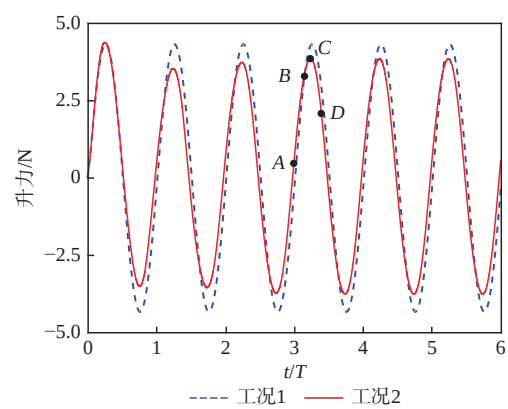


图 15　工况 1、2 升力曲线对比

Fig. 15　Comparison of lift curves for case 1 and case 2

在工况 2 中的柔性机翼模型，机翼的翼尖弯曲位移、速度和翼尖扭转响应随阵风的变化曲线如图 16 所示。其中，翼尖扭转角度是指在阵风作用下柔性机翼变形后翼尖在 *XY* 平面的投影与未发生变形前翼尖在 *XY* 平面投影的夹角，如图 17 所示，当前缘向上时翼尖扭转角为正，当前缘向下时翼尖扭转角为负。

图 16 中 $A\sim D$ 为 $3T\sim3.5T$ 半个阵风周期内 4 个不同的时刻，对应于图 15 中的 $A\sim D$，此时，阵

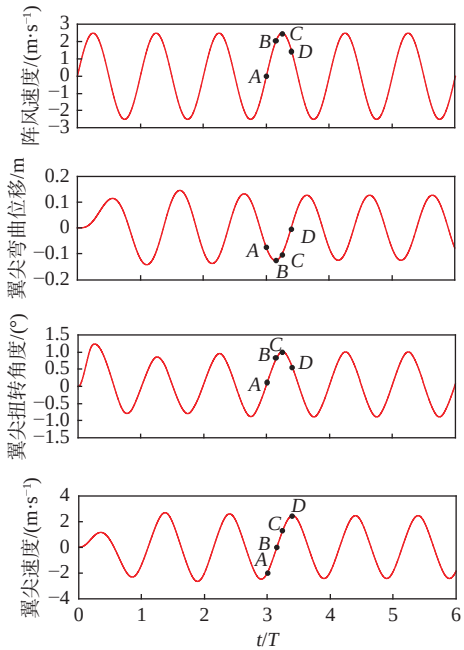


图 16 工况 2 动响应结果

Fig. 16 Dynamic response results for case 2

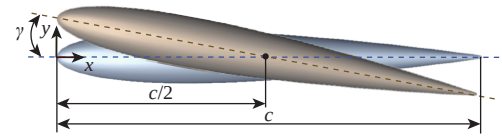


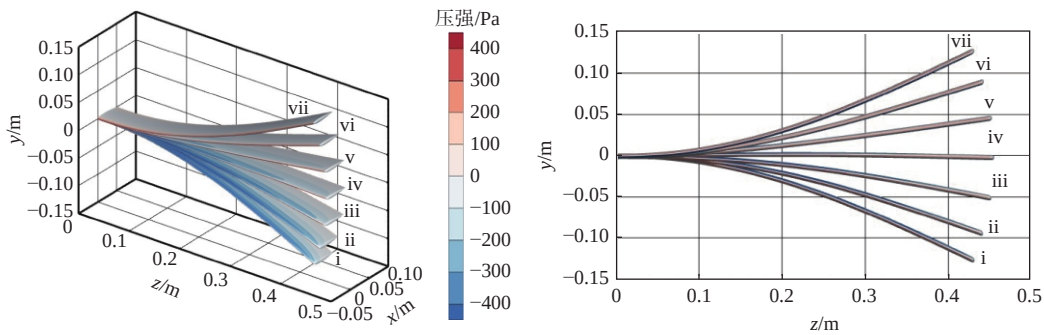
图 17 翼尖扭转角度示意

Fig. 17 Schematic of wingtip twist angle

风响应达到稳定，进行周期性变化，*A* 为阵风速度为 0 的时刻，*B* 为翼尖弯曲位移最大的时刻，*C* 为阵风速度最大的时刻，*D* 为翼尖速度最大的时刻。

阵风作用下，当机翼变形响应达到稳定周期后，选取一个阵风周期内机翼弯曲变形由最低点到最高点的 7 个典型时刻进行分析，分别记为 ⅰ～ⅶ，对应的无量纲时刻 *t*/*T* 分别为 0.150、0.270、0.338、0.398、0.458、0.525 和 0.650。各时刻机翼表面压力分布及弯曲变形如图 18 所示。由图 18 可知，工况 2 下机翼最大弯曲变形为 0.126 m，约为机翼半展长的 28%。

为进一步分析机翼柔性变形对升力的影响，



(a) 等轴侧图　　(b) 前视图

图 18 工况 2 下不同时刻压力云图及变形

Fig. 18 Pressure contours and deformation at different moments under case 2

图 19 显示了一个周期内工况 1、2 下刚柔机翼升力沿展向的分布。其中，Z_{max} 表示机翼在弯曲变形后沿 *Z* 轴方向坐标的最大值，即翼尖在 *Z* 轴方向的坐标。

由图 19 结合图 15、图 16 的结果分析可知，在 *A*～*B* 时刻，柔性机翼的总升力大于刚性机翼，柔性机翼在靠近翼根附近，升力小于靠近翼尖的升力，刚性机翼升力沿展向均匀分布。这是由于此时柔性机翼翼尖速度为负，与垂向阵风速度相反，导致阵风相对于机翼的速度变大，机翼等效迎角变大，且阵风的作用增大了机翼的扭转角度，同样使机翼迎角增大，两者的影响使其升力大于刚性机翼，而越靠近翼根的位置，机翼运动速度越小，扭转变形越小，因此，沿展向靠近翼根部分的升力小于靠近翼尖的区域。

在 *B*～*D* 时刻，柔性机翼的总升力小于刚性机翼，柔性机翼在靠近翼根附近升力大于靠近翼尖的升力，刚性机翼升力沿展向近似均匀分布，但在靠近翼尖部分升力相对较小。从 *B* 时刻开始，机翼运动速度与垂向阵风速度方向相同，均为正值，阵风相对于机翼的速度减小，机翼等效迎角减小，升力减小，越靠近翼尖，机翼的弯曲变形速度越快，因此，靠近翼尖部分升力显著小于靠近翼根部分。虽然此时机翼的扭转角度为正，但最大扭转角度仅为 1.02°，相比于阵风速度引起的迎角变化，扭转变形对升力的影响可以忽略不计。

在 *C* 时刻阵风速度达到最大，刚柔机翼总升力均达到最大值，*C* 时刻时 2 种工况下升力沿展向的变化趋势及机翼上表面压力分布如图 20 所示。其中，总升力 F_l 通过对机翼表面的压力差积分得到：

$$F_l = \int_S \Delta p \mathrm{d}S \tag{10}$$

沿展向单位长度升力为

$$l = \mathrm{d}F_l/\mathrm{d}z \tag{11}$$

式中：Δp 为机翼下表面与上表面的压力差；$\mathrm{d}S$ 为微

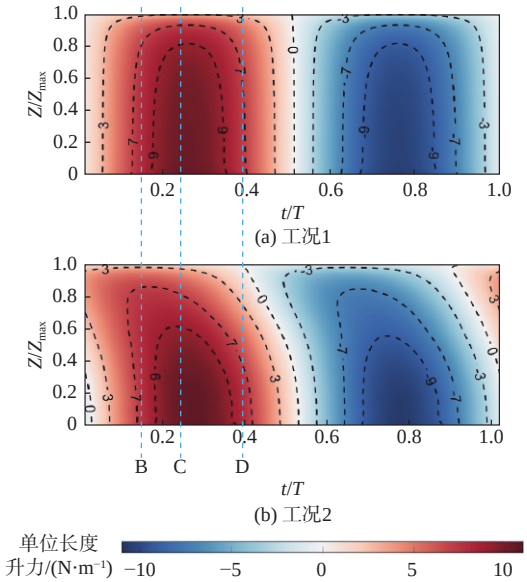


图 19　工况 1、2 截面升力时空分布

Fig. 19　Spatiotemporal distributions of the sectional lift for case 1 and case 2

元面积; z为机翼沿 z 轴方向的长度。

分析图 20 结果可知, 刚性机翼在靠近翼尖时升力大幅减小, 柔性机翼在靠近机翼根部时升力就逐渐减小, 在 Z/Z_{max}<0.20 时, 刚性机翼与柔性机翼升力分布几乎一致, 而当 Z/Z_{max}>0.20 时, 柔性机翼由于柔性变形引起升力减小。

图 21 显示了半个阵风周期内在不同时刻根据 Q 准则[21]计算出的机翼周围三维涡流等值面。

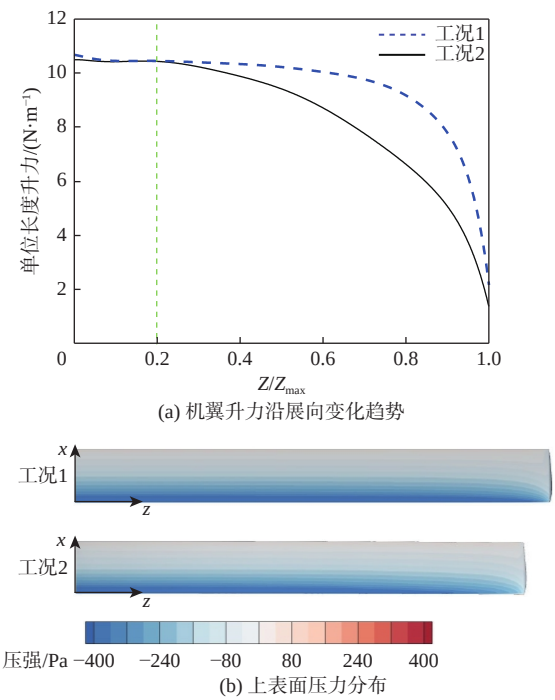


图 20　C 时刻机翼升力沿展向变化趋势及上表面压力分布

Fig. 20　Trend of wing lift along the spreading direction and pressure distribution on the upper surface at the moment C

R_G=0.1 时, 只有翼尖涡 (tip vortex, TV) 的演化。TV 的存在会使升力减小, 因此, 机翼在靠近翼尖部分的升力相对较小。在 A 时刻, 阵风速度为零, 刚性机翼只有脱落在流场中的 TV 存在, 而柔性机翼有扭转变形, 因此, 在翼尖处有 TV 形成。随着阵风的增大, TV 强度也随之增大, 在 C 时刻 TV 强度达到最大。在 t/T=3.398 时, 柔性机翼由于弯曲变形影响, TV 强度减小到最小值, 柔性机翼仅有脱落在流场中的 TV 存在, 而刚性机翼的 TV 强度依旧很大。

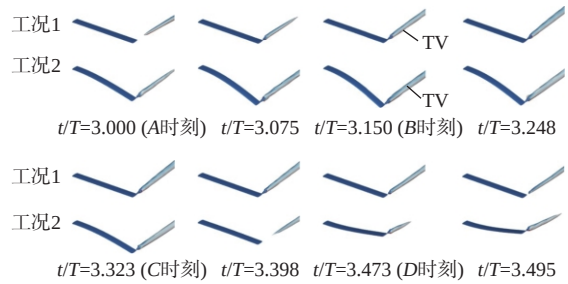


图 21　不同时刻机翼周围的 Q 准则等值面(Q = 2 000)

Fig. 21　The Q-criterion iso-surfaces around the wing at different time (Q = 2 000)

2) R_G = 0.4。工况 3、4 下机翼的升力随时间的变化曲线如图 22 所示。此时, 升力出现非线性变化, 在 R_G=0.4时, 机翼等效迎角变大, 机翼发生深度失速, 且失速后升力出现了二次峰值。而工况 4 因机翼的柔性变形改变了表面的气动力的分布, 减少了失速现象的发生。图中 a～d 为机翼升力达到极值的 4 个不同时刻。

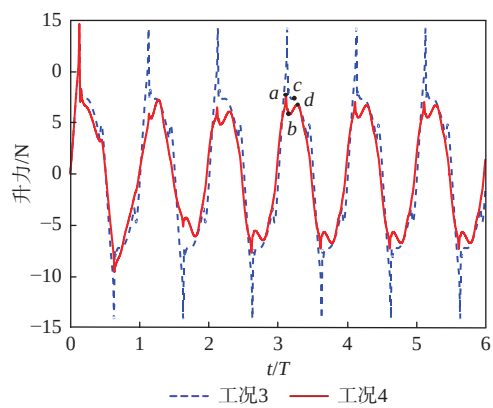


图 22　工况 3、4 升力曲线对比

Fig. 22　Comparison of lift curves for case 3 and case 4

对于工况 4 对应的柔性机翼模型, 机翼的翼尖弯曲位移、速度和翼尖扭转响应随阵风的变化曲线如图 23 所示。

图 24 给出了工况 4 下阵风响应达到稳定周期后, 机翼弯曲变形由最低点到最高点过程中 7 个典型时刻的压力云图及变形图。7 个典型时刻分别记

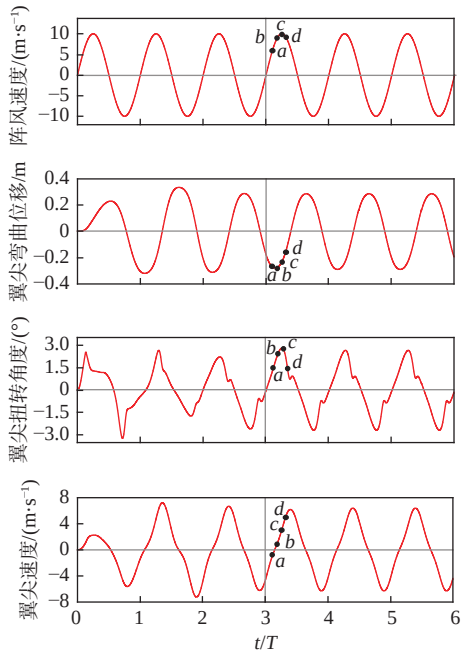


图 23 工况 4 动响应结果

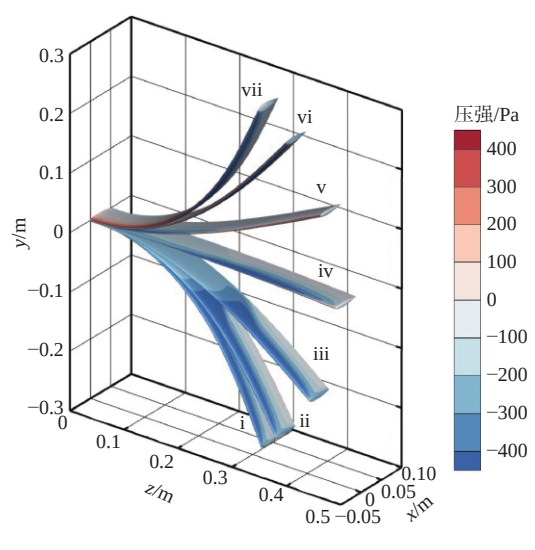
Fig. 23 Dynamic response results for case 4

为 ⅰ ～ ⅶ, 对应的无量纲时刻 t/T 分别为 0.143、0.233、0.308、0.398、0.473、0.548 和 0.643。由图 24 可知, 工况 4 下机翼发生了显著的大幅弯曲变形, 最大弯曲变形达到 0.284 m, 约为机翼半展长的 63%。

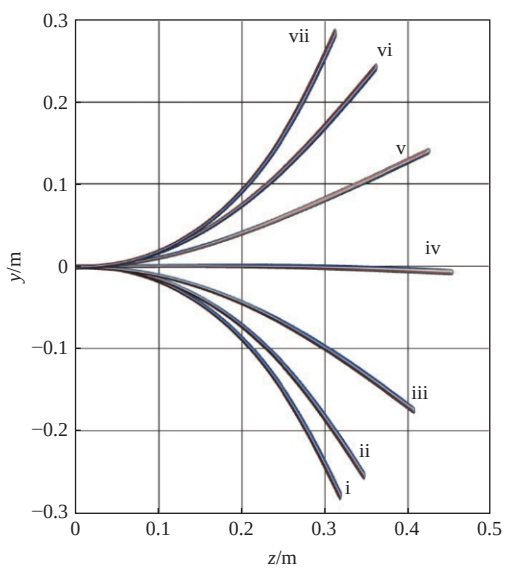
为进一步分析大柔性变形对升力的影响, 图 25 展示了一个阵风周期内工况 3、4 下刚柔机翼升力沿展向的分布。可知, 大柔性变形使机翼的展向升力分布不再均匀, 随着阵风的变化, 升力的最大值出现在机翼中部, 其主要原因是大变形影响了机翼表面的涡结构分布。

图 26 给出了 12 个不同时刻刚柔机翼三维涡流等值面, 其中包含图 22、图 23 中对应的时刻 (a～d), 可以看出, 在 t/T=3.131 时刻, 机翼表面已经形成了由阵风诱导的前缘涡 (leading-edge vortex, LEV), 流动分离发生, LEV 的存在会导致机翼上表面产生局部负压区域。a 时刻前, LEV 附着在机翼上表面, 涡致升力使机翼升力持续增加, 在 a 时刻时升力达到最大。在 a 时刻之后, LEV 开始从机翼上表面分离脱落, 在 t/T=3.128 时刻, 柔性机翼在翼根附近形成一个拱形涡 (arch vortex, AV), AV 和 LEV 的脱落使其引起的局部负压减小, 致使升力突降。在 b 时刻, 流动分离达到最大, 升力达到极小值。在 b 时刻之后, 流场涡结构演化趋势稳定, 涡对升力的变化不再起主导作用, 随着阵风的持续增加, 机翼的升力再次逐渐增大。在 c 时刻, 阵风速度达到最大, 相应的刚性机翼升力达到第 2 个峰值, 而柔性机翼升力尚未达到峰值, 柔性机翼升力在 d 时刻达到第 2 个峰值, 这是由于柔性机翼因其柔性变形, 涡会沿展向变化, 从 c～d 时刻, 机翼弯曲变形减小, LEV 沿展向分布的宽度变大, LEV 沿展向分布宽度增大引起的升力增大的效应大于阵风减小导致的升力减小。在 d 时刻之后, 随着阵风的减小, LEV 强度逐渐减小, 并在 t/T=3.450 时刻消失, 升力也随之减小。LEV 和 AV 的形成和演化导致了机翼截面升力的非线性行为。对于柔性机翼, 柔性变形使变形区域的 LEV 展向尺寸显著减小, 翼尖附近没有 LEV, 生成升力小于刚性机翼, 因此, 沿展向截面升力最大处位于机翼中部。

图 27 给出了图 25 对应 c 时刻时的展向升力分布曲线和机翼上表面压力分布, 此时, 升力分布与 R_G=0.1 时的情况截然不同。图 27(a) 的结果表明,



(a) 等轴侧图



(b) 前视图

图 24 工况 4 下不同时刻压力云图及变形

Fig. 24 Pressure contours and deformation at different moments under case 4

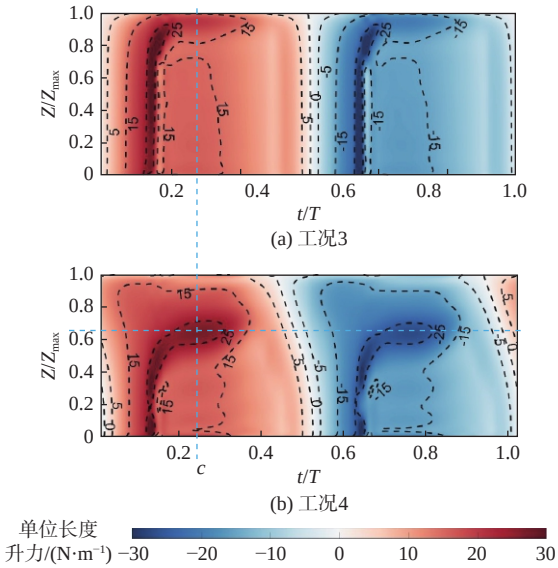


图 25　工况 3、4 截面升力时空分布

Fig. 25　Spatiotemporal distributions of the sectional lift for case 3 and case 4

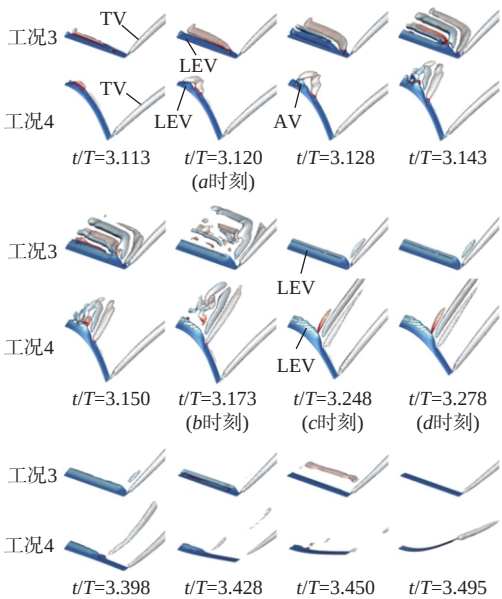


图 26　不同时刻机翼周围的 Q 准则等值面（Q = 5 000）

Fig. 26　The Q-criterion iso-surfaces around the wing at different time (Q = 5 000)

c 时刻柔性机翼升力最大处位于 Z/Z_{max}=0.60 处，而刚性机翼升力最大值位于 Z/Z_{max}=0.94 处，这是由于刚性机翼在翼尖附近 LEV 与 TV 相互影响，导致该区域 LEV 被挤压在机翼上表面附近，LEV 强度增大，进而导致局部升力增大。而柔性机翼在展向 LEV 分离区域会有类似于 TV 的涡管形成，此涡管与 LEV 相互作用导致了局部升力增大。

综上所述，机翼的大柔性变形抑制了前缘分离，减少了失速现象的发生，同时，柔性变形使 LEV 的展向面积减小，上表面的低压区域面积减小，减小了阵风载荷。

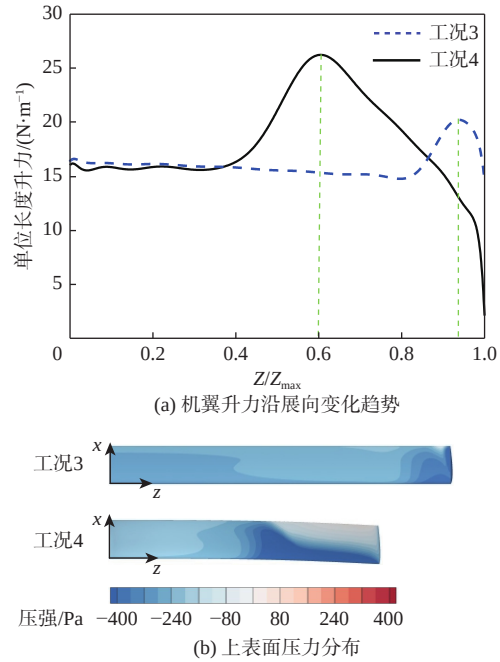


图 27　c 时刻机翼升力沿展向变化趋势及上表面压力分布

Fig. 27　Trend of wing lift along the spreading direction and pressure distribution on the upper surface at the moment c

4　结　论

基于 CFD/CSD 耦合，建立了适用于大幅阵风与大柔性梁结构的流固耦合分析方法。通过倒置柔性板算例对该方法的准确性进行了验证，并对大柔性机翼在阵风作用下的响应进行了分析，得到以下结论：

1）相比于刚性机翼，柔性机翼在阵风作用下的升力变化较小，这主要是由于柔性机翼在阵风作用下，会发生显著的弯曲和扭转变形，从而在一定程度上减弱了阵风对升力的影响。具体来说，柔性结构在阵风中的变形改变了气流的流动路径，使整体升力的波动幅度减小，这一特性有助于提高飞行稳定性。

2）在大幅阵风条件下，刚性机翼容易出现显著的失速现象，表现为升力的非线性剧烈变化。相比之下，柔性机翼通过翼尖的弯曲变形，减少了展向 LEV 的范围，从而有效减弱了分离失速所引起的升力变化。具体表现为柔性机翼能够更好地适应阵风的冲击，避免了大规模的气流分离和失速现象，从而维持较为平稳的升力输出。

综上所述，柔性机翼在大幅阵风条件下展现出了较好的气动性能，通过结构变形有效缓解了阵风对升力的冲击，减少了失速现象的发生。这些结论对于本文的研究对象具有特定的参考价值，表明柔性设计在复杂气动环境中的潜在优势，对于未来的高效飞行器设计具有重要的参考价值，但其普遍适用性有待进一步研究和验证。

参考文献（References）

[1] HU P, TANG Y, HAN Y, et al. Numerical study on self-excited forces and flow fields for a thin plate under a sinusoidal non-stationary wind condition[J]. Physics of Fluids, 2023, 35(6): 063602.

[2] BADRYA C, BILER H, JONES A R, et al. Effect of gust width on flat-plate response in large transverse gust[J]. AIAA Journal, 2021, 59(1): 49-64.

[3] CORKERY S J, BABINSKY H, HARVEY J K. On the development and early observations from a towing tank-based transverse wing-gust encounter test rig[J]. Experiments in Fluids, 2018, 59(9): 135.

[4] TANG D, DOWELL E H. Experimental and theoretical study of gust response for high-aspect-ratio wing[J]. AIAA Journal, 2002, 40: 419-429.

[5] NEUMANN J, MAI H. Gust response: simulation of an aeroelastic experiment by a fluid-structure interaction method[J]. Journal of Fluids and Structures, 2013, 38: 290-302.

[6] AN C, YANG C, XIE C C, et al. Flutter and gust response analysis of a wing model including geometric nonlinearities based on a modified structural ROM[J]. Chinese Journal of Aeronautics, 2020, 33(1): 48-63.

[7] ZHENG Y, DAI Y T, YANG C, et al. Effect of wingtip bending morphing on gust-induced aerodynamics based on fluid-structure interaction method[J]. Physics of Fluids, 2023, 35(11): 115124.

[8] 陈利丽, 宋笔锋, 宋文萍, 等. 柔性扑翼气动结构耦合特性数值研究[J]. 空气动力学学报, 2015, 33(1): 125-133.
CHEN L L, SONG B F, SONG W P, et al. Numerical aerodynamic-structural coupling research for flexible flapping wing[J]. Acta Aerodynamica Sinica, 2015, 33(1): 125-133(in Chinese).

[9] MIAN H H, WANG G, YE Z Y. Numerical investigation of structural geometric nonlinearity effect in high-aspect-ratio wing using CFD/CSD coupled approach[J]. Journal of Fluids and Structures, 2014, 49: 186-201.

[10] 聂雪媛, 黄程德, 杨国伟. 基于 CFD/CSD 耦合的结构几何非线性静气动弹性数值方法研究[J]. 振动与冲击, 2016, 35(8): 48-53.
NIE X Y, HUANG C D, YANG G W. Numerical analysis for aeroelastic with structural geometrical nonlinearity using a CFD/CSD-coupled method[J]. Journal of Vibration and Shock, 2016, 35(8): 48-53(in Chinese).

[11] 吴金洋, 吴光辉, 魏小辉, 等. 临近空间太阳能飞行器着陆状态阵风响应[J]. 海军航空大学学报, 2024, 39(1): 153-158.
WU J Y, WU G H, WEI X H, et al. Gust Loads in the landing state of solar-powered aircraft in near space[J]. Journal of Naval Aviation University, 2024, 39(1): 153-158(in Chinese).

[12] SIMO J C. A finite strain beam formulation. The three-dimensional dynamic problem. Part I[J]. Computer Methods in Applied Mechanics and Engineering, 1985, 49(1): 55-70.

[13] SIMO J C, VU-QUOC L. A three-dimensional finite-strain rod model. Part II: computational aspects[J]. Computer Methods in Applied Mechanics and Engineering, 1986, 58(1): 79-116.

[14] REN H. A computationally efficient and robust geometrically-exact curved beam formulation for multibody systems[C]//Proceedings of the 7th Asian Conference on Multibody Dynamics. Berlin: Springer, 2014: 285-296.

[15] REN H, KHUDE N, ORTIZ J, et al. A shear-locking free model of Timoshenko beams with large deformation and large rotation[C]//Proceedings of the ECCOMAS Thematic Conference. Berlin: Springer, 2013: 457-468.

[16] Hexagon AB. Adams 2020: Adams FE Part User's Guide[EB/OL]. [2024-02-10]. https://nexus.hexagon.com/documentationcenter/en-US/bundle/Adams_2020_Adams_FE_Part_User_Guide/resource/Adams_2020_Adams_FE_Part_User_Guide.pdf

[17] MENTER F R. Two-equation eddy-viscosity turbulence models for engineering applications[J]. AIAA Journal, 1994, 32(8): 1598-1605.

[18] BOOKSTEIN F L. Principal warps: thin-plate splines and the decomposition of deformations[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1989, 11(6): 567-585.

[19] KIM D, COSSÉ J, CERDEIRA C H, et al. Flapping dynamics of an inverted flag[J]. Journal of Fluid Mechanics, 2013, 736: R1.

[20] LEE Y L, BAEDER J D. Implicit hole cutting-A new approach to overset grid connectivity: AIAA-2003-4128[R]. Reston: AIAA, 2003.

[21] JEONG J, HUSSAIN F. On the identification of a vortex[J]. Journal of Fluid Mechanics, 1995, 285: 69-94.

Nonlinear fluid-structure interaction response analysis of a large flexible wing under strong gusts

SU Jinxin[1, 2], XI Ziyan[1, 2], DAI Yuting[1, 2, *]

(1. Tianmushan Laboratory, Hangzhou 310023, China;
2. School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China)

Abstract: The geometric nonlinearity and aerodynamic nonlinearity cannot be ignored as the large flexible wing undergoes significant deformation when subjected to high gusts. In order to study the response of a large flexible wing under strong gusts, a fluid-structure interaction analysis method based on geometric nonlinear beam theory and computational fluid dynamics (CFD) was established. The fluid model and geometrically nonlinear beam model of a big flexible wing with a semi-aspect ratio of 9 were constructed, and the accuracy of the fluid-structure interaction simulation approach was confirmed by comparing wind tunnel tests of inverted flexible plate. The dynamic response characteristics of large flexible airfoils were examined at various gust ratios (R_G=0.1-0.4). In particular, the dynamic stall aerodynamic characteristics of the airfoils caused by gusts at large gust ratios (R_G=0.4) were examined. The gust response analysis was conducted at a Reynolds number of 10^5. The results show that significant nonlinear bending deformation of the wing occurs with the increase of gust ratios, and the maximum bending deformation of the wing reaches 63% of the half-spread length at R_G=0.4. Meanwhile, the bending deformation of the wing reduces the width of the spreading distribution of the leading-edge vortex (LEV), which affects the distribution of aerodynamic forces and reduces the gust loads.

Keywords: large flexible wing; nonlinear; fluid-structure interaction; gust response; dynamic stall

Received: 2024-05-06; **Accepted**: 2024-08-23; **Published Online**: 2024-11-27 13:31
URL: link.cnki.net/urlid/11.2625.V.20241126.1801.004
* **Corresponding author**. E-mail: yutingdai@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0268

基于序列反转的高速机动目标雷达相干积累方法

王睿铮[1,2]，李世强[1,*]

(1. 中国科学院空天信息创新研究院，北京 100094；2. 中国科学院大学 电子电气与通信工程学院，北京 100049)

摘 要：针对高速机动目标雷达回波相干积累中的跨距离单元走动和多普勒频率徙动问题，提出基于改进时间反转变换-新距离频率反转变换-变尺度逆傅里叶变换(ITRT-NRFRT-SCIFT)和时间反转变换-新距离频率反转变换(TRT-NRFRT)的联合算法，利用慢TRT分离目标速度信息和距离信息，采用NRFRT消除多普勒频率徙动的影响。在TRT-NRFRT后，采用快速傅里叶变换(FFT)实现能量积累，而ITRT-NRFRT后则采用SCIFT消除慢时间与距离频率之间的耦合，并实现回波能量积累。经过仿真实验可以证明，算法能够在不进行任何参数搜索的条件下实现对三阶运动目标的跨距离单元走动和多普勒频率徙动校正，并进行回波能量积累，获得目标的径向距离和速度信息，且具备相对较低的运算复杂度。

关 键 词：高速机动目标；脉冲串；序列反转；距离走动；多普勒徙动

中图分类号：V221+; TN957

文献标志码：A **文章编号**：1001-5965(2026)06-2184-10

在雷达目标探测中，长时间相干积累技术能够充分利用回波中的相位信息，作为提升信噪比和弱小目标检测性能的手段，得到了广泛应用[1]。而高速机动目标的雷达回波受到目标高径向速度及高阶机动影响，可能会存在跨距离单元走动和多普勒频率徙动的问题，这些问题使传统机动目标检测(moving target detection, MTD)算法[2]失效。而如何克服跨距离单元走动和多普勒频率徙动的影响，快速实现回波能量积累与目标检测是雷达信号处理领域的一个热门话题[3]。

在现有高速机动目标相干积累算法中，最为成熟的是Keystone变换[4]和Radon-Fourier变换(Radon Fourier transform, RFT)算法[5-7]，这2种算法能够在低信噪比条件下实现有效相干积累。但Keystone变换对跨距离单元走动的校正受速度模糊的影响，需要进行模糊数搜索和估计[8]。而RFT算法需要对目标运动参数进行遍历搜索，计算量巨大[9]。同时，两者只能校正目标线性运动带来的跨距离单元走动，无法校正多普勒频率徙动。

针对适用于三阶运动模型的变加速目标，广义RFT(generalized Radon Fourier transform, GRFT)算法[10-11]可以实现变加速目标的有效相干积累，但需要多维度搜索目标运动参数，运算成本高，实时性较差[12]，工程实现比较困难。Li等[13]提出的(time reversing transform, second-order Keystone transform and LV's distribution, TRT-SKT-LVD)算法不需要参数搜索，通过时间反转变换(time reversing transform, TRT)[14]和二阶Keystone变换[15]可以实现距离走动校正，并采用吕氏分布(LV's distribution, LVD)[16]实现多普勒频率徙动校正与相干积累，但整体算法流程仍较为复杂。

基于以上研究，针对三阶运动模型，本文提出了时间反转变换-新距离频率反转变换(time reversing transform-new range frequency reversal transform,

收稿日期：2024-04-29；**录用日期**：2024-08-23；**网络出版时间**：2024-12-11 09:50
网络出版地址：link.cnki.net/urlid/11.2625.V.20241210.1555.002
基金项目：科工局民用航天技术预先研究项目(D010105)
*通信作者. E-mail：lishq@mail.ie.ac.cn

引用格式：王睿铮，李世强. 基于序列反转的高速机动目标雷达相干积累方法[J]. 北京航空航天大学学报，2026，52(6)：2184-2193.
WANG R Z，LI S Q. Radar coherent integration method for high-speed maneuvering targets based on sequence reversal[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：2184-2193 (in Chinese).

TRT-NRFRT) 和改进时间反转变换-新距离频率反转变换-变尺度逆傅里叶变换 (improved time reversing transform-new range frequency reversal transform-scale-variable inverse Fourier transform, ITRT-NRFRT-SCIFT) 算法, 通过 2 种算法的相互配合, 可以在不进行参数搜索的条件下校正线性距离走动和多普勒频率徙动。在慢时间维度通过反转和反转共轭分离目标距离与速度信息, 并通过新距离频率反转变换 (new range frequency reversal transform, NRFRT)[17] 消除多普勒频率徙动, 使用快速傅里叶变换 (fast Fourier transform, FFT) 和变尺度逆傅里叶变换 (scale-variable inverse Fourier transform, SCIFT)[18] 实现信号能量积累。同时, ITRT-NRFRT 与 TRT-NRFRT 算法可以并行实现, 进一步提升了算法实时性。

1　信号模型及问题表述

设雷达发射的线性调频信号为

$$s(\tilde{t},t_m)=\mathrm{rect}\left(\frac{\tilde{t}}{T_\mathrm{p}}\right)\exp\left[\mathrm{j}2\pi\left(f_\mathrm{c}t+\frac{k}{2}\tilde{t}^2\right)\right] \tag{1}$$

式中: f_c为信号载频; T_p为脉冲持续时间; $k=B/T_\mathrm{p}$为信号调频率, B为信号带宽; $\tilde{t}$为距离向快时间; $t_m=mT_\mathrm{r}$为慢时间, m 为第 m 个脉冲, T_r为脉冲重复周期; $t=\tilde{t}+t_m$为全时间; rect(·)为矩形函数。

设有初始距离为R_0、速度为v、加速度和加加速度分别为a_2和a_3的高速机动目标。使用三阶运动模型, 在第 m 个脉冲时对应的目标斜距大小可以表示为$R(t_m)=R_0+vt_m+a_2t_m^2/2+a_3t_m^3/6$, 此时信号回波时延为

$$\begin{aligned}\tau_m&=2\left(R_0+vt_m+\frac{1}{2}a_2t_m^2+\frac{1}{6}a_3t_m^3\right)\Big/c=\\&2\left(R_0+vmT_\mathrm{r}+\frac{1}{2}a_2m^2T_\mathrm{r}^2+\frac{1}{6}a_3m^3T_\mathrm{r}^3\right)\Big/c\end{aligned} \tag{2}$$

式中: c 为光速, 通常为 3×10^8。

雷达接收的第m个脉冲的回波信号可以表示为

$$\begin{aligned}s_\mathrm{r}(\tilde{t},t_m)=&A_\mathrm{r}\mathrm{rect}\left(\frac{\tilde{t}-\tau_m}{T_\mathrm{p}}\right)\exp\left(\mathrm{j}\pi k\left(\tilde{t}-\tau_m\right)^2\right)\cdot\\&\exp\left(\mathrm{j}2\pi f_\mathrm{c}\left(\tilde{t}-\tau_m\right)\right)\end{aligned} \tag{3}$$

式中: A_r为目标后向散射系数。

经过脉冲压缩处理后的信号可以表示为

$$s_\mathrm{c}(\tilde{t},t_m)=A_1\mathrm{sinc}\left(B\left(\tilde{t}-\tau_m\right)\right)\exp\left(-\mathrm{j}2\pi f_\mathrm{c}\tau_m\right) \tag{4}$$

式中: A_1为脉冲压缩后的回波信号幅度。

由式 (4)可知, 回波包络的峰值位于$\tilde{t}=2R(t_m)/c$处, 目标回波峰值位置随着慢时间t_m的变化而变化, 即发生了距离走动。

在雷达相干积累时间$t_m\in[-T_\mathrm{a}/2,T_\mathrm{a}/2]$内, 由目标速度、加速度和加加速度造成的一、二、三阶距离走动可以分别表示为

$$\begin{cases}\Delta R_1=\left|vt_m\right|_{t_m\in[-T_\mathrm{a}/2,T_\mathrm{a}/2]}=\left|vT_\mathrm{a}\right|\\\Delta R_2=\left|\dfrac{1}{2}a_2t_m^2\right|_{t_m\in[-T_\mathrm{a}/2,T_\mathrm{a}/2]}=\left|\dfrac{1}{8}a_2T_\mathrm{a}^2\right|\\\Delta R_3=\left|\dfrac{1}{6}a_3t_m^3\right|_{t_m\in[-T_\mathrm{a}/2,T_\mathrm{a}/2]}=\left|\dfrac{1}{24}a_3T_\mathrm{a}^3\right|\end{cases} \tag{5}$$

雷达回波信号距离分辨单元大小$\Delta R=c/2B$, 当目标运动造成的距离走动 $\Delta R_i>\Delta R(i=1,2,3)$时, 需要考虑跨距离单元走动。对于一般的高速机动目标, 二阶及以上的高阶运动造成的距离走动项ΔR_H通常不会超过一个距离单元, 即当目标的加速度和加加速度大小满足条件:

$$\Delta R_\mathrm{H}=\left|\frac{1}{8}a_2T_\mathrm{a}^2+\frac{1}{24}a_3T_\mathrm{a}^3\right|<\frac{c}{2B} \tag{6}$$

此时, 可以只考虑一阶距离走动项。

对于高速目标探测中可能会出现多普勒模糊的情况, 目标的径向速度可以重写为

$$v=v_0+N_\mathrm{b}v_\mathrm{b} \tag{7}$$

式中: $v_\mathrm{b}=\lambda T_\mathrm{r}/2$为雷达盲速, λ为雷达信号波长; v_0为模糊速度, 满足$|v_0|<v_\mathrm{b}/2$; N_b为多普勒模糊数。

则目标回波在脉冲压缩后的信号可以重写为

$$\begin{aligned}s_\mathrm{c}(\tilde{t},t_m)=&A_1\mathrm{sinc}\left\{B\left[\tilde{t}-\frac{2(R_0+vt_m)}{c}\right]\right\}\cdot\\&\exp\left[-\frac{\mathrm{j}4\pi f_\mathrm{c}}{c}\left(R_0+vt_m+\frac{1}{2}a_2t_m^2+\frac{1}{6}a_3t_m^3\right)\right]\end{aligned} \tag{8}$$

式 (8) 中, 回波信号相位会随着慢时间的变化而变化, 信号相位可以表示为

$$\begin{aligned}\Phi(t_m)=&-4\pi f_\mathrm{c}R(t_m)/c=\\&-4\pi f_\mathrm{c}\left(R_0+vt_m+\frac{1}{2}a_2t_m^2+\frac{1}{6}a_3t_m^3\right)\Big/c\end{aligned} \tag{9}$$

即式 (8) 中指数项中的二次项和三次项会造成多普勒频率徙动。目标加速度和加加速度带来的多普勒频移项分别记为Δf_2和Δf_3:

$$\begin{cases}\Delta f_2=\left|\dfrac{2a_2t_m}{\lambda}\right|_{t_m\in[T_\mathrm{a}/2,T_\mathrm{a}/2]}=\left|\dfrac{2a_2T_\mathrm{a}}{\lambda}\right|\\\Delta f_3=\left|\dfrac{a_3t_m^2}{\lambda}\right|_{t_m\in[-T_\mathrm{a}/2,T_\mathrm{a}/2]}=\left|\dfrac{a_3T_\mathrm{a}^2}{4\lambda}\right|\end{cases} \tag{10}$$

回波的多普勒分辨单元大小$\rho_{f_\mathrm{d}}=1/T_\mathrm{a}$, 当$\Delta f_i>\rho_{f_\mathrm{d}}$时, 会发生$i-1$阶的跨多普勒单元频率徙动。跨距离单元走动和多普勒频率徙动会使目标回波的能量分散在不同的距离单元和多普勒单元中, 造成相干积累处理和雷达目标检测困难。

2　快速相干积累方法

2.1　基于序列反转的高速机动目标回波积累方法

对式 (8) 进行快时间维度的 FFT 并代入式 (7),

可以得到目标回波在脉冲压缩后的距离频域表达式为

$$S_{\mathrm{c}}(f,t_m)=A_2\mathrm{rect}\left(\frac{f}{B}\right)\exp\left[-\frac{\mathrm{j}4\pi f}{c}(R_0+vt_m)\right]\cdot \exp\left[-\frac{\mathrm{j}4\pi f_{\mathrm{c}}}{c}\left(R_0+v_0t_m+\frac{1}{2}a_2t_m^2+\frac{1}{6}a_3t_m^3\right)\right] \tag{11}$$

由式(11)可知，跨距离单元走动在快时间频域表现为距离频率f与慢时间t_m之间的耦合，而高阶相位项则会引起多普勒频率徙动。为解决该问题，本文提出基于序列反转的高速机动目标回波积累方法。

设经过脉冲压缩后的回波信号$S_{\mathrm{c}}(f,t_m)$可以存储在一个大小为$(M+1)\times(N+1)$的矩阵中。由$M+1$个脉冲(M为偶数)所组成的慢时间序列$\boldsymbol{t}'_m=mT_{\mathrm{r}}$可以表示为

$$\boldsymbol{t}'_m=[-M/2T_{\mathrm{r}},(-M/2+1)T_{\mathrm{r}},\cdots,0,\cdots,(M/2-1)T_{\mathrm{r}},M/2T_{\mathrm{r}}] \tag{12}$$

式中：$m=-M/2,-M/2+1,\cdots,0,\cdots,M/2-1,M/2$。

同理，回波信号的距离频率序列$\boldsymbol{f}=n\Delta f$可以表示为

$$\boldsymbol{f}'=[-N/2\Delta f,(-N/2+1)\Delta f,\cdots,0,\cdots,(N/2-1)\Delta f,N/2\Delta f] \tag{13}$$

距离频率序列长度为$N+1$(N为偶数)，其中，$n=-N/2,-N/2+1,\cdots,0,\cdots,N/2-1,N/2$，$\Delta f$为距离频域的采样间隔。

将式(12)和式(13)代入式(11)可得到回波距离频域信号的离散表达式：

$$S_{\mathrm{c}}(n,m)=A_2\mathrm{rect}\left(\frac{n\Delta f}{B}\right)\exp\left[-\frac{\mathrm{j}4\pi n\Delta f}{c}(R_0+vmT_{\mathrm{r}})\right]\cdot \exp\left[-\frac{\mathrm{j}4\pi f_{\mathrm{c}}}{c}\left(R_0+vmT_{\mathrm{r}}+\frac{1}{2}a_2m^2T_{\mathrm{r}}^2+\frac{1}{6}a_3m^3T_{\mathrm{r}}^3\right)\right] \tag{14}$$

对式(14)中$S_{\mathrm{c}}(n,m)$进行慢时间维序列反转，得到$S_{\mathrm{c}}(n,\overleftarrow{m})$。其中，慢时间反转序列可以表示为

$$\overleftarrow{\boldsymbol{t}'_m}=[M/2T_{\mathrm{r}},(M/2-1)T_{\mathrm{r}},\cdots,0,\cdots,(-M/2+1)T_{\mathrm{r}},-M/2T_{\mathrm{r}}] \tag{15}$$

经过慢时间维序列反转的脉冲压缩后距离频域信号$S_{\mathrm{c}}(n,\overleftarrow{m})$可以表示为

$$S_{\mathrm{c}}(n,\overleftarrow{m})=A_2\mathrm{rect}\left(\frac{n\Delta f}{B}\right)\exp\left[-\frac{\mathrm{j}4\pi n\Delta f}{c}(R_0-vmT_{\mathrm{r}})\right]\cdot \exp\left[-\frac{\mathrm{j}4\pi f_{\mathrm{c}}}{c}\left(R_0-vmT_{\mathrm{r}}+\frac{1}{2}a_2m^2T_{\mathrm{r}}^2-\frac{1}{6}a_3m^3T_{\mathrm{r}}^3\right)\right] \tag{16}$$

与式(14)相比，式(16)中$S_{\mathrm{c}}(n,\overleftarrow{m})$慢时间的奇数次相位项符号相反而偶数次相位项相同。则两者相乘(即进行TRT)后，结果为

$$S_{\mathrm{c1}}(n,m)=S_{\mathrm{c}}(n,m)S_{\mathrm{c}}(n,\overleftarrow{m})=A_3\mathrm{rect}\left(\frac{n\Delta f}{B}\right)\exp\left(-\frac{\mathrm{j}8\pi n\Delta f}{c}R_0\right)\cdot \exp\left[-\frac{\mathrm{j}8\pi f_{\mathrm{c}}}{c}\left(R_0+\frac{1}{2}a_2m^2T_{\mathrm{r}}^2\right)\right] \tag{17}$$

慢TRT后的信号中只包含目标的距离和加速度信息，而径向速度和加加速度的影响则被成功消除。将$S_{\mathrm{c}}(n,m)$与$S_{\mathrm{c}}(n,\overleftarrow{m})$的共轭信号$S_{\mathrm{c}}^*(n,\overleftarrow{m})$相乘，即进行ITRT[19]，可以得到仅包含目标径向速度和加加速度信息的信号：

$$S_{\mathrm{c2}}(n,m)=S_{\mathrm{c}}(n,m)S_{\mathrm{c}}^*(n,\overleftarrow{m})=A_3\mathrm{rect}\left(\frac{n\Delta f}{B}\right)\exp\left(-\frac{\mathrm{j}8\pi n\Delta f}{c}vmT_{\mathrm{r}}\right)\cdot \exp\left[-\frac{\mathrm{j}8\pi f_{\mathrm{c}}}{c}\left(vmT_{\mathrm{r}}+\frac{1}{6}a_3m^3T_{\mathrm{r}}^3\right)\right] \tag{18}$$

至此，通过式(17)和式(18)可以成功将目标的径向距离、径向加速度、径向速度、加加速度信息分离至2组不同的信号中。且在式(17)中，通过时间反转成功消除了目标速度引起的跨距离单元走动，使回波包络峰值位置只与目标径向距离有关。

但式(17)中信号相位中由目标加速度、加加速度引起的多普勒频率徙动仍存在并影响能量聚焦。因此，引入文献[17]中的NRFRT进一步消除多普勒频率徙动的影响。设回波信号的距离频率反转序列为

$$\overleftarrow{\boldsymbol{f}'}=[N/2\Delta f,(N/2-1)\Delta f,\cdots,0,\cdots,(-N/2+1)\Delta f,-N/2\Delta f] \tag{19}$$

对$S_{\mathrm{c1}}(n,m)$和$S_{\mathrm{c2}}(n,m)$进行距离频率反转并作共轭运算，可以分别得到如下形式的信号：

$$S_{\mathrm{c1}}^*(\overleftarrow{n},m)=A_3\mathrm{rect}\left(\frac{n\Delta f}{B}\right)\exp\left(-\frac{\mathrm{j}8\pi n\Delta f}{c}R_0\right)\cdot \exp\left[\frac{\mathrm{j}8\pi f_{\mathrm{c}}}{c}\left(R_0+\frac{1}{2}a_2m^2T_{\mathrm{r}}^2\right)\right] \tag{20}$$

$$S_{\mathrm{c2}}^*(\overleftarrow{n},m)=A_3\mathrm{rect}\left(\frac{n\Delta f}{B}\right)\exp\left(-\frac{\mathrm{j}8\pi n\Delta f}{c}vmT_{\mathrm{r}}\right)\cdot \exp\left[\frac{\mathrm{j}8\pi f_{\mathrm{c}}}{c}\left(vmT_{\mathrm{r}}+\frac{1}{6}a_3m^3T_{\mathrm{r}}^3\right)\right] \tag{21}$$

将原信号与距离频率反转共轭信号相乘，执行文献[17]中的NRFRT：

$$S_{\mathrm{c3}}(n,m)=S_{\mathrm{c1}}(n,m)S_{\mathrm{c1}}^*(\overleftarrow{n},m)=A_4\mathrm{rect}\left(\frac{n\Delta f}{B}\right)\exp\left(-\frac{\mathrm{j}16\pi n\Delta f}{c}R_0\right) \tag{22}$$

$$S_{c4}(n,m)=S_{c2}(n,m)S_{c2}^{*}(\overleftarrow{n},m)=$$
$$A_4\mathrm{rect}\left(\frac{n\Delta f}{B}\right)\exp\left(-\frac{\mathrm{j}16\pi n\Delta f}{c}vmT_{\mathrm{r}}\right)\tag{23}$$

$S_{c3}(n,m)$和$S_{c4}(n,m)$通过 NRFRT 消除了二阶和三阶多普勒频率徙动。此时, $S_{c3}(n,m)$中仅有信号包络峰值位置中的目标径向距离信息。对式 (22) 执行距离向快速傅里叶逆变换(inverse fast Fourier transform, IFFT), 可以得到距离向时域信号:

$$S_{R_0}(n,m)=\mathrm{IFFT}[S_{c3}(n,m)]=A_5\mathrm{sinc}\left[B\left(t-\frac{8R_0}{c}\right)\right]\tag{24}$$

之后, 通过沿慢时间方向 FFT 可以将信号能量积累至多普勒中心位置上的峰值。此时, 通过信号峰值在距离向上的坐标即可估算出目标径向距离:

$$\mathrm{FFT}[S_{R_0}(n,m)]=A_5\mathrm{sinc}\left[B\left(t-\frac{8R_0}{c}\right)\right]\delta(f_{\mathrm{d}}+0)\tag{25}$$

式中: δ为冲击函数f_{d}为信号多普勒频率。

信号能量积累形成峰值点的距离向坐标$t=8R_0/c$, 可以据此估算目标径向距离。

对于式 (23) 中$S_{c4}(n,m)$, 距离向频率f和慢时间t_m之间的耦合并未消除, 即信号包络峰值位置仍然受到目标径向速度带来的一阶跨距离单元走动影响, 峰值位置随慢时间变化。为解决距离频率和慢时间之间的耦合, 对经过 ITRT 和 NRFRT 得到的$S_{c4}(n,m)$执行 SCIFT 以消除耦合。因为 ITRT 已经使信号包络信息中仅包含目标一阶距离走动, 故 SCIFT 无需再进行双线性自相关变换, 可以直接实行变尺度操作, 式 (23) 写成如下连续形式:

$$S_{c4}(f,t_m)=A_4\mathrm{rect}\left(\frac{f}{B}\right)\exp\left(-\frac{\mathrm{j}16\pi f}{c}vt_m\right)\tag{26}$$

对式 (26) 进行 SCIFT 可以得到

$$T_{\mathrm{S}}(t_\psi,t_m)=\int_{t_mf}S_{c4}(f,t_m)\exp(\mathrm{j}2\pi\eta t_\psi t_m f)\,\mathrm{d}(t_mf)=$$
$$\int_{t_mf}A_4\mathrm{rect}\left(\frac{f}{B}\right)\exp\left(-\frac{\mathrm{j}16\pi f}{c}vt_m\right)\cdot$$
$$\exp(\mathrm{j}2\pi\eta t_\psi t_m f)\,d(t_mf)=$$
$$A_5\mathrm{sinc}\left[B\left(t_\psi-\frac{8v}{\eta c}\right)\right]\tag{27}$$

式中: η为缩放因子; t_ψ为经过尺度变换的距离向快时间。

SCIFT 成功消除了式 (26) 信号中距离频率和慢时间之间的耦合, 信号包络峰值沿着直线$t_\psi=8v/\eta c$分布, 不再随慢时间变化。缩放因子的选择应满足经过 SCIFT 后的信号包络峰值位置不超过快时间的最大值, 即满足不等式:

$$|t_\psi|=\left|\frac{8v}{\eta c}\right|<\frac{T_{\mathrm{r}}}{2}\tag{28}$$

此时, 通过沿着慢时间方向的 FFT 即可实现信号能量聚焦, 并通过峰值检测获得目标不模糊速度。

2.2　算法多目标积累及交叉项分析

通过 2.1 节可知, 本文提出的基于 TRT-NRFRT 和 ITRT-NRFRT-SCIFT 的联合算法能够在单目标场景下顺利完成单目标回波相干积累并估算出目标距离及速度。而对于多目标的情况, 非线性的慢 TRT 和 NRFRT 会产生交叉项, 假设存在J个目标的情况, 其中, 第q个目标对应的距离、速度、加速度和加加速度分别表示为R_q、v_q、a_{q2}和a_{q3}。此时, 脉压后信号的频域表达式为

$$S_{cm}(f,t_m)=\sum_{q=1}^{J}A_q\mathrm{rect}\left(\frac{f}{B}\right)\exp\left[-\frac{\mathrm{j}4\pi f}{c}\left(R_q+v_qt_m\right)\right]\cdot$$
$$\exp\left[-\frac{\mathrm{j}4\pi f_{\mathrm{c}}}{c}\left(R_q+v_qt_m+\frac{1}{2}a_{q2}t_m^2+\frac{1}{6}a_{q3}t_m^3\right)\right]\tag{29}$$

式中: A_q表示第 q 个目标回波脉压后领域信号的信号幅度。

则多目标回波经过 TRT 及 ITRT 处理后的结果分别为

$$S_{c1m}(n,m)=S_{cm}(n,m)S_{cm}(n,\overleftarrow{m})=$$
$$\sum_{q=1}^{J}|A_q|^2\mathrm{rect}\left(\frac{f}{B}\right)\exp\left(-\frac{\mathrm{j}8\pi f}{c}R_q\right)\cdot$$
$$\exp\left[-\frac{\mathrm{j}8\pi f_{\mathrm{c}}}{c}\left(R_q+\frac{1}{2}a_{q2}t_m^2\right)\right]+$$
$$\sum_{q,l=1,q\neq m}^{J}A_{\mathrm{cross1}}\mathrm{rect}\left(\frac{f}{B}\right)\exp\left[-\frac{\mathrm{j}4\pi f}{c}\left(R_q+R_l\right)\right]\cdot$$
$$\exp\left[-\frac{\mathrm{j}4\pi f}{c}\left(v_q-v_l\right)t_m\right]\exp\left[-\frac{\mathrm{j}2\pi f_{\mathrm{c}}}{c}\left(a_{q2}+a_{l2}\right)t_m^2\right]\cdot$$
$$\exp\left[-\frac{\mathrm{j}4\pi f_{\mathrm{c}}}{c}(v_q-v_l)t_m\right]\exp\left[-\frac{\mathrm{j}2\pi f_{\mathrm{c}}}{3c}(a_{q3}-a_{l3})t_m^3\right]\tag{30}$$

式中: A_{cross1}表示多目标回波经过 TRT 处理后产生的交叉项信号幅度。

$$S_{c2m}(n,m)=S_{c}(n,m)S_{c}^{*}(n,\overleftarrow{m})=$$
$$\sum_{q=1}^{J}|A_q|^2\mathrm{rect}\left(\frac{f}{B}\right)\exp\left(-\frac{\mathrm{j}8\pi f}{c}v_qmT_{\mathrm{r}}\right)\cdot$$
$$\exp\left[-\frac{\mathrm{j}8\pi f_{\mathrm{c}}}{c}\left(v_qt_m+\frac{1}{6}a_{q3}t_m^3\right)\right]+$$
$$\sum_{q,l=1,q\neq l}^{J}A_{\mathrm{cross2}}\mathrm{rect}\left(\frac{f}{B}\right)\exp\left[-\frac{\mathrm{j}4\pi f}{c}\left(R_q-R_l\right)\right]\cdot$$
$$\exp\left[-\frac{\mathrm{j}4\pi f}{c}\left(v_q+v_l\right)t_m\right]\exp\left[-\frac{\mathrm{j}4\pi f_{\mathrm{c}}}{c}\left(v_q+v_l\right)t_m\right]\cdot$$
$$\exp\left[-\frac{\mathrm{j}2\pi f_{\mathrm{c}}}{c}\left(a_{q2}-a_{l2}\right)t_m^2\right]\exp\left[-\frac{\mathrm{j}2\pi f_{\mathrm{c}}}{3c}\left(a_{q3}+a_{l3}\right)t_m^3\right]\tag{31}$$

式中: A_{cross2}表示多目标回波经过 ITRT 处理后产生

的交叉项信号幅度。

由式(30)和式(31)可知，多目标回波经过TRT或ITRT处理后，所产生的交叉项不能像单目标及自相关项那样对目标运动参数实现分组。在TRT处理后的式(30)中，当目标的速度不同时，交叉项中仍存在跨距离单元走动。而式(31)中，ITRT处理后，交叉项的信号包络则沿着直线$\tilde{t}=2\left[\left(R_q-R_l\right)+\left(v_q+v_l\right)t_m\right]/c$分布，且交叉项存在$R_{\text{cross}}=R_q-R_l$的虚拟径向距离。对式(30)和式(31)进行NRFRT，得到结果为

$$S_{\text{c3}m}(f,t_m)=S_{\text{c1}m}(f,t_m)S_{\text{c1}m}(\overleftarrow{f},t_m)=\sum_{q=1}^{J}A_{q4}\text{rect}\left(\frac{f}{B}\right)\exp\left(-\frac{\text{j}16\pi f}{c}R_q\right)+S_{\text{c3cross1}}(f,t_m)+S_{\text{c3cross2}}(f,t_m) \tag{32}$$

$$S_{\text{c4}m}(f,t_m)=S_{\text{c2}m}(f,t_m)S_{\text{c2}m}(\overleftarrow{f},t_m)=\sum_{q=1}^{J}A_{q4}\text{rect}\left(\frac{f}{B}\right)\exp\left(-\frac{\text{j}16\pi f}{c}v_q t_m\right)+S_{\text{c4cross1}}(f,t_m)+S_{\text{c4cross2}}(f,t_m) \tag{33}$$

式中：$S_{\text{c3cross1}}(f,t_m)$和$S_{\text{c4cross1}}(f,t_m)$分别为式(30)和式(31)中的交叉项和其自身在NRFRT中形成的新交叉项；A_{q4}表示经过NRFRT处理后自相关项的信号幅度；$S_{\text{c3cross2}}(f,t_m)$和$S_{\text{c4cross2}}(f,t_m)$分别为式(30)和式(31)中的交叉项和自相关项或其他交叉项进行NRFRT得到的结果。其中，第q个目标和第l个目标形成的交叉项$S_{\text{c3cross1}}(f,t_m)$和$S_{\text{c4cross1}}(f,t_m)$可分别表示为

$$S_{\text{c3cross1}}(f,t_m)=A_{\text{cross}}\text{rect}\left(\frac{f}{B}\right)\exp\left[-\frac{\text{j}8\pi f}{c}\left(R_q+R_l\right)\right]\cdot\exp\left[-\frac{\text{j}8\pi f}{c}\left(v_q-v_l\right)t_m\right] \tag{34}$$

$$S_{\text{c4cross1}}(f,t_m)=A_{\text{cross}}\text{rect}\left(\frac{f}{B}\right)\exp\left[-\frac{\text{j}8\pi f}{c}\left(R_q-R_l\right)\right]\cdot\exp\left[-\frac{\text{j}8\pi f}{c}\left(v_q+v_l\right)t_m\right] \tag{35}$$

式中：A_{cross}表示交叉项的幅度。

当$v_q=v_l$时，TRT-NRFRT结果中的$S_{\text{c3cross1}}(f,t_m)$不再受跨距离单元走动影响，采用FFT积累后会在$\tilde{t}=4(R_q+R_l)/c$、$f_{\text{d}}=0$处产生可影响目标检测的交叉项峰值。

当交叉项与自相关项或其他交叉项进行NRFRT时，可以得到新的交叉项$S_{\text{c3cross2}}(f,t_m)$和$S_{\text{c4cross2}}(f,t_m)$，可以分别表示为

$$S_{\text{c3cross2}}(f,t_m)=A_{\text{cross}}\text{rect}\left(\frac{f}{B}\right)\exp\left[-\frac{\text{j}4\pi f}{c}\left(R_q+R_l\right)\right]\exp\left[-\frac{\text{j}4\pi f}{c}\left(v_q-v_l\right)t_m\right]\exp\left[-\frac{\text{j}4\pi f_{\text{c}}}{c}\left(v_q-v_l\right)t_m\right]\cdot\exp\left[-\frac{\text{j}2\pi f_{\text{c}}}{c}\left(a_{q2}+a_{l2}\right)t_m^2\right]\exp\left[-\frac{\text{j}2\pi f_{\text{c}}}{3c}\left(a_{q3}-a_{l3}\right)t_m^3\right]\exp\left[-\frac{\text{j}8\pi f}{c}\left(R_{\text{o}}+v_{\text{o}}t_m\right)\right]\cdot\exp\left[\frac{\text{j}8\pi f_{\text{c}}}{c}\left(R_{\text{o}}+v_{\text{o}}t_m+\frac{1}{2}a_{\text{o}2}t_m^2+\frac{1}{6}a_{\text{o}3}t_m^3\right)\right] \tag{36}$$

$$S_{\text{c4cross2}}(f,t_m)=A_{\text{cross}}\text{rect}\left(\frac{f}{B}\right)\exp\left[-\frac{\text{j}4\pi f}{c}\left(R_q-R_l\right)\right]\exp\left[-\frac{\text{j}4\pi f}{c}\left(v_q+v_l\right)t_m\right]\exp\left[-\frac{\text{j}4\pi f_{\text{c}}}{c}\left(v_q+v_l\right)t_m\right]\cdot\exp\left[-\frac{\text{j}2\pi f_{\text{c}}}{c}\left(a_{q2}-a_{l2}\right)t_m^2\right]\exp\left[-\frac{\text{j}2\pi f_{\text{c}}}{3c}\left(a_{q3}+a_{l3}\right)t_m^3\right]\exp\left[-\frac{\text{j}8\pi f}{c}\left(R_{\text{o}}+v_{\text{o}}t_m\right)\right]\cdot\exp\left[\frac{\text{j}8\pi f_{\text{c}}}{c}\left(R_{\text{o}}+v_{\text{o}}t_m+\frac{1}{2}a_{\text{o}2}t_m^2+\frac{1}{6}a_{\text{o}3}t_m^3\right)\right] \tag{37}$$

式中：R_{o}、v_{o}、$a_{\text{o}2}$和$a_{\text{o}3}$分别为与交叉项进行NRFRT的自相关项或其他交叉项的对应参数。由式(36)和式(37)可知，交叉项与自相关项或其他交叉项进行NRFRT得到的新交叉项中，跨多普勒频率单元徙动仍然存在，且多普勒相位均不为零。而本文提出算法的自相关项均位于零多普勒处，故只需在中心多普勒频率单元上进行一维峰值检测，即可排除如式(36)和式(37)的交叉项影响，而只有式(34)和式(35)交叉项之间进行NRFRT后可以成功校正交叉项的跨多普勒频率单元徙动。其中，TRT-NRFRT后的交叉项在多目标速度不同时，由于仍存在跨距离单元走动，与自相关项积累得到的峰值相比，交叉项在使用FFT积累时无法形成清晰的峰值，同样不会对后续检测造成影响。而ITRT-NRFRT后，如式(35)所示的交叉项在经过后续SCIFT处理后，会呈现余弦振荡形式，不能积累出自相关项。

2.3 基本算法流程及运算复杂度分析

通过本文算法估计高速机动目标运动参数的流程如图1所示。由于采用了ITRT-NRFRT-SCIFT和TRT-NRFRT算法的组合，其中的慢TRT和NRFRT部分相同，可以一定程度上降低算法实现所需的系统复杂度。

由图1可知，本文算法采用2组TRT和NRFRT实现回波信号距离和速度信息的分离和多普勒频

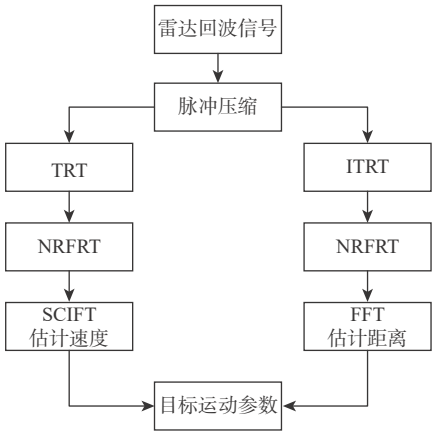


图 1　算法流程

Fig. 1　Algorithm flowchart

率徙动的校正, 在完成校正后实现能量积累, 获得目标运动参数。其中, TRT、ITRT 和 NRFRT 运算均通过一次矩阵复乘实现。设回波积累脉冲数为 M, 快时间距离门数为 N, 则一次矩阵乘法所需要的运算复杂度为 $O(MN)$。而在本文提取目标径向距离的 TRT-NRFRT 中, 能量积累所需的慢时间 FFT 运算复杂度为 $O[(MN\log_2 M)/2]$。

对于 ITRT-NRFRT-SCIFT 中用于消除速度引起的跨距离单元走动的 SCIFT, 也可以通过 chirp-z 变换实现进而减少所需的运算量。式 (27) 的 SCIFT 可以写为以下的离散表示形式:

$$\begin{aligned} T_{\mathrm{S}}(n_\psi,m)&=\sum_n S_{c4}(n,m)\exp\left(\mathrm{j}2\pi\eta n_\psi\Delta n_\psi m T_{\mathrm{r}} n\Delta f\right)=\\ &\sum_n A_4\mathrm{rect}\left(\frac{n\Delta f}{B}\right)\exp\left(-\frac{\mathrm{j}16\pi n\Delta f}{c}vmT_{\mathrm{r}}\right)\cdot\\ &\exp\left(\mathrm{j}2\pi\eta n_\psi\Delta n_\psi m T_{\mathrm{r}} n\Delta f\right)=A_1\mathrm{sinc}\left[B\left(n_\psi-\frac{8v}{\eta c}\right)\right] \end{aligned} \tag{38}$$

式中: n_ψ 为 t_ψ 的离散形式。式 (38) 可以进一步表示为

$$\begin{aligned} T_{\mathrm{S}}(n_\psi,m)&=\sum_n A_4\mathrm{rect}\left(\frac{n\Delta f}{B}\right)\exp\left(-\frac{\mathrm{j}16\pi n\Delta f}{c}vmT_{\mathrm{r}}\right)\cdot\\ &\exp\left(\mathrm{j}2\pi\eta n_\psi m T_{\mathrm{r}} n\Delta f\right)=\\ &\sum_n A_4\mathrm{rect}\left(\frac{n\Delta f}{B}\right)\exp\left(-\frac{\mathrm{j}16\pi n\Delta f}{c}vmT_{\mathrm{r}}\right)\cdot\\ &\exp\left[\mathrm{j}2\pi\eta m T_{\mathrm{r}}^2\Delta f\frac{n^2+n_\psi^2-(n-n_\psi)^2}{2}\right]=\\ &\exp\left(\frac{\mathrm{j}2\pi\eta m T_{\mathrm{r}}^2\Delta f n_\psi^2}{2}\right)\sum_n A_4\mathrm{rect}\left(\frac{n\Delta f}{B}\right)\\ &\exp\left(-\frac{\mathrm{j}16\pi n\Delta f}{c}vmT_{\mathrm{r}}\right)\exp\left(\frac{\mathrm{j}2\pi\eta m T_{\mathrm{r}}^2\Delta f n^2}{2}\right)\cdot\\ &\exp\left[-\frac{\mathrm{j}2\pi\eta m T_{\mathrm{r}}^2\Delta f(n-n_\psi)^2}{2}\right] \end{aligned} \tag{39}$$

式 (39) 中, SCIFT 可视作频域卷积, 故可以通过 IFFT 实现。此时, 实现 SCIFT 所需的计算复杂度为 $O[(3MN\log_2 N)/2]$。而由于 TRT-NRFRT 和 ITRT-NRFRT-SCIFT 算法可以并行实现, 故本文联合算法所需的计算复杂度量级为 $O[MN\log_2 M]$。

为衡量本文算法的计算复杂度, 选取 2 种前人应用于高速机动目标的经典算法: GRFT[10−11] 和 TRT-SKT-LVD[13], 并给出前人算法的计算复杂度。其中, GRFT 算法的计算复杂度为 $O[2N_{\mathrm{r}}N_{\mathrm{v}}N_{\mathrm{a}}N_{\mathrm{j}}MN]$。GRFT 算法需要在距离、速度、加速度和加加速度多维度展开逐帧搜索, 其中, N_{r} 为距离搜索次数, N_{v}、N_{a} 和 N_{j} 分别为速度、加速度和加加速度的搜索次数。

对于同样采用 TRT 且不需要进行参数搜索的 TRT-SKT-LVD 算法, 设瞬时自相关函数操作后的滞后时间样本数为 N_τ, 则运行 TRT-SKT-LVD 算法所需要的总计算复杂度为 $O[5N_\tau MN\log_2 M+N_\tau MN\log_2 N_\tau+2MN\log_2 N+M^2N+MN]$。假设 $M=N=N_\tau=N_{\mathrm{r}}=N_{\mathrm{v}}=N_{\mathrm{a}}$, 则本文算法、GRFT 算法和 TRT-SKT-LVD 算法的计算复杂度量级分别为 $O[N^2\log_2 N]$、$O[N^6]$ 和 $O[N^3\log_2 N]$, 可以说明, 本文算法与其他高阶运动目标信号积累算法相比, 具有较高的运算效率。

采用相同的计算机配置和软件环境, 对 GRFT[10−11]、TRT-SKT-LVD[13]、MTD[2] 和本文算法成功估计出目标运动参数所需的运行时间进行对比, 结果如表 1 所示。

表 1　算法运行时间比较

Table 1　Algorithm running time comparison

算法名称	运行时间/s
GRFT[10−11]	244 548.01
TRT-SKT-LVD[13]	6.57
MTD[2]	0.17
本文算法	1.30

由表 1 可见, GRFT 算法由于需要对目标运动参数进行多维度搜索, 需要花费相当长的运算时间, 在进行对比的算法中速度最慢。相对地, TRT-SKT-LVD 算法不需要进行参数搜索, 所需时间较短, 实时性更好。而本文算法成功实现仅需 1.30 s, 运算速度仅次于 MTD 算法, 更加有利于工程实现。而 MTD 算法由于仅需要一次 FFT 即可实现, 具备更快的运行速度。但其无法有效处理回波跨距离单元走动和跨多普勒单元徙动, 对快速获取目标运动信息并无帮助。

3　仿真实验

本节通过数值仿真验证本文算法的有效性。设雷达系统参数(即仿真目标运动参数)如表 2 所示。

表 2 雷达系统参数

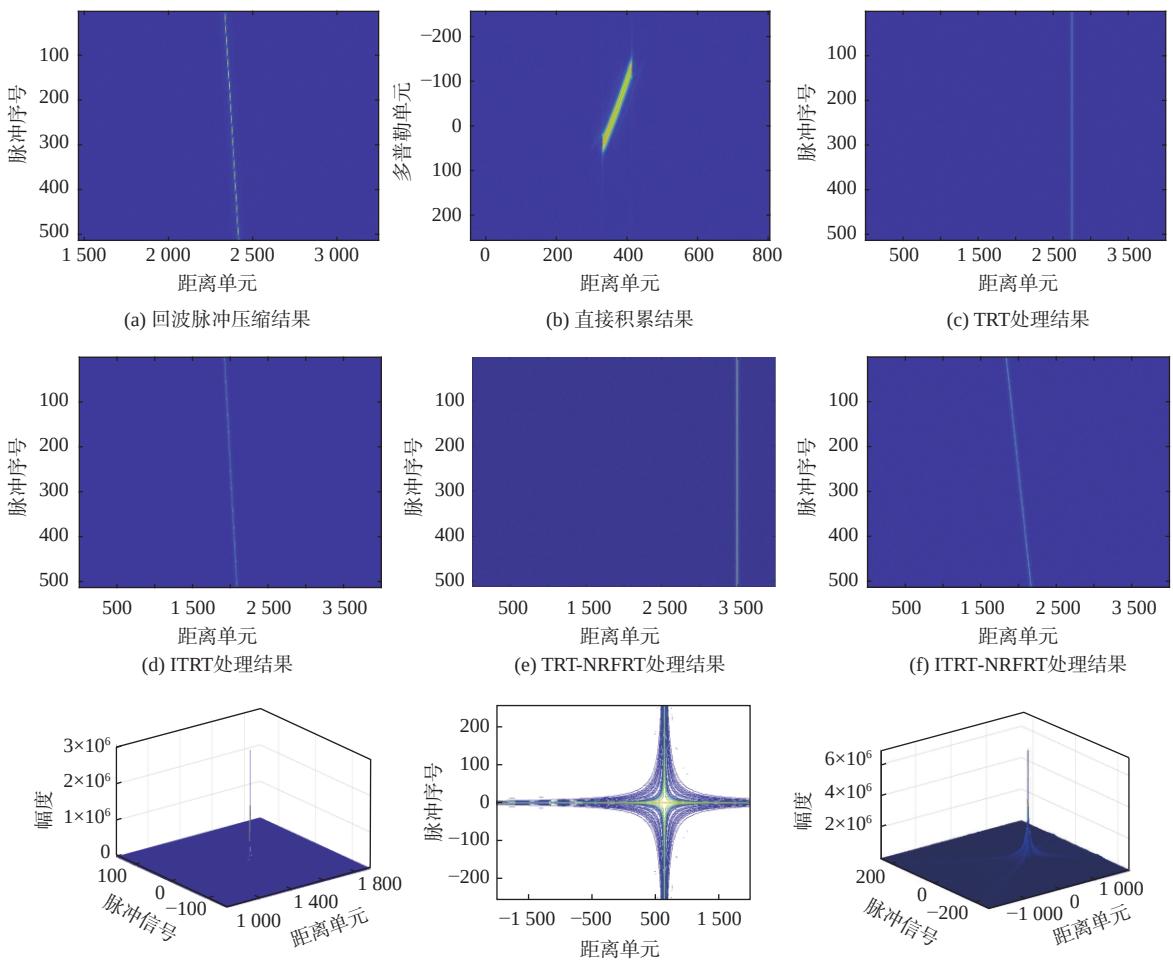
Table 2 Radar system parameters

参数	数值
载波频率/GHz	1
带宽/MHz	2
采样频率/MHz	4
脉冲重复频率/Hz	1 000
积累脉冲数	513
目标径向距离/m	14 000
目标径向速度/(m·s^{-1})	6 000
目标径向加速度/(m·s^{-2})	100
目标径向加加速度/(m·s^{-3})	10

本文算法的效果如图 2 所示。图 2(a) 为目标回波的脉冲压缩结果，可以看到，经过脉冲压缩的回波信号包络峰值随着慢时间变化而变化。当目标加速度为 100 m/s^2、加加速度为 10 m/s^3 时，在积累时间内，由二、三阶运动带来的距离走动约为 3.34 m，远小于雷达距离分辨率，高阶距离走动项几乎不造成影响，满足式 (6) 中条件。

假设在高斯白噪声环境下按照表 2 中的参数对目标回波进行处理，雷达回波在脉冲压缩压缩后的信噪比为 6 dB。可以得到如图 2 所示的仿真实验结果。

图 2(b) 为目标回波直接进行相干积累后结



(a) 回波脉冲压缩结果 (b) 直接积累结果 (c) TRT处理结果

(d) ITRT处理结果 (e) TRT-NRFRT处理结果 (f) ITRT-NRFRT处理结果

(g) TRT-NRFRT积累结果 (h) ITRT-NRFRT-SCIFT处理结果 (i) ITRT-NRFRT-SCIFT积累结果

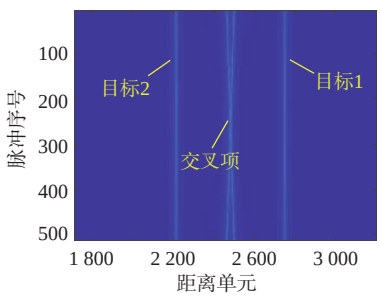
图 2 算法仿真结果

Fig. 2 Algorithm simulation results

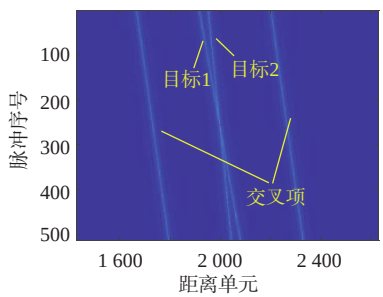
果。可以观察到，受跨距离单元走动和跨多普勒单元徙动影响，回波能量分布在不同距离和多普勒单元中，故无法直接通过 FFT 实现信号能量的有效积累。图 2(c) 和图 2(d) 分别为回波信号经过 TRT 和 ITRT 处理后得到的结果。其中，经过 TRT 处理后的信号包络峰值位置中保留了目标径向距离信息，消除了由目标径向速度带来的一阶跨距离单元走动，回波包络峰值位置沿着几乎平行慢时间轴的直线分布。而经过 ITRT 后的信号包络中仅包含由目标径向速度带来的一阶距离走动项，信号包络峰值点沿着直线 $\tau_r = 2vt_m$ 分布，其中，τ_r 为峰值点对应的快时间坐标。图 2(e) 为经过 TRT-NRFRT 处理后结果，图 2(f) 为经过 ITRT-NRFRT 处理后结果。TRT-NRFRT 处理后，信号相位中多普勒频率徙动造成

的影响被消除, 此时可以直接通过 FFT 实现能量积累, 并通过信号峰值点距离轴坐标得出目标径向距离, 积累结果如图 2(g) 所示。而经过 ITRT-NRFRT 处理后, 信号虽然不再受到多普勒频率徙动影响, 但仍需要解决距离频率与慢时间之间的耦合, 消除一阶跨距离单元走动。图 2(h) 为经过 SCIFT 操作解耦合后的信号。与图 2(f) 相比, 图 2(h) 中信号包络峰值沿着直线 $t_{\psi}=8v/\eta c$ 分布, 可以直接进行能量积累。能量积累的结果如图 2(i) 所示, 通过检测积累得到的信号峰值在距离向快时间轴上的坐标, 可以计算得到目标径向速度。且当目标速度满足 $8v/\eta c \leqslant T_a/2$ 即 $v < T_a\eta c/16$ 时, 不需要考虑目标速度模糊问题。

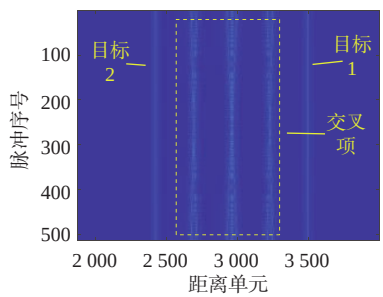
针对回波中包含多个目标的情况展开仿真实验, 雷达系统参数采用表 2 中的设置。设单目标仿真时的目标为目标 1, 假设目标 2 的径向距离 R_2=4 000 m, 径向速度 v_2=3 400 m/s, 径向加速度 a_{22}=−50 m/s^2, 加加速度 a_{23}=20 m/s^3。图 3 为多目标仿真实验结果。其中, 图 3(a) 和图 3(b) 分别为回波信号经过 TRT 和 ITRT 处理后得到的结果, 与式 (30) 和式 (31) 中的结果相符, 经过 TRT 处理后, 自相关项中的跨距离单元走动被成功校正, 而交叉项在快时间维度仍存在跨距离单元走动。而 ITRT 处理后, 自相关项沿着通过原点的直线分布, 而交叉项则存在虚拟径向距离。



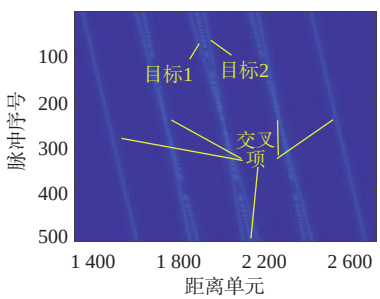
(a) TRT处理后的信号峰值包络



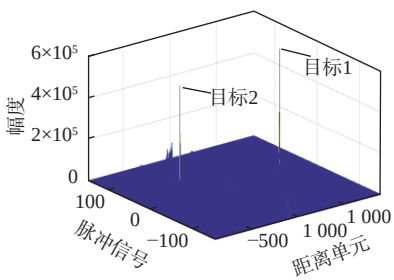
(b) ITRT处理后的信号峰值包络



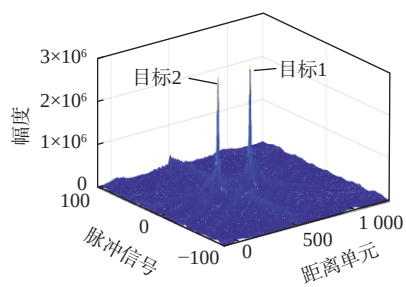
(c) TRT-NRFRT处理后的信号峰值包络



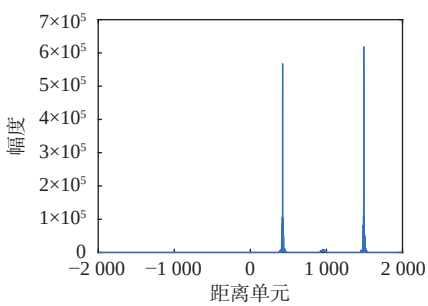
(d) ITRT-NRFRT处理后的信号峰值包络



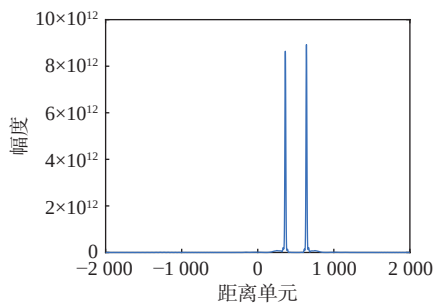
(e) TRT-NRFRT处理后的积累峰值



(f) ITRT-NRFRT-SCIFT处理后的积累峰值



(g) TRT-NRFRT处理后的积累峰值切片



(h) ITRT-NRFRT-SCIFT处理后的积累峰值切片

图 3　多目标算法仿真结果

Fig. 3　Multiple targets algorithm simulation results

图 3(c) 为经过 TRT-NRFRT 处理后的结果, 其中, 自相关项仍沿着垂直距离轴的直线分布, 无跨距离单元走动, 而交叉项的分布与式 (32)、式 (34) 和式 (36) 中的结果相符, 仍存在跨距离单元走动。图 3(d) 为 ITRT-NRFRT 处理后的结果, 自相关项和交叉项的分布与式 (33)、式 (35) 和式 (37) 中的结果相符。

图 3(e) 和图 3(f) 分别为经过 TRT-NRFRT 和 ITRT-NRFRT-SCIFT 处理后进行积累的结果, 相干积累峰值均在同一多普勒频率单元内分布。两者的零多普勒切片分别如图 3(g) 和图 3(h) 所示, 目标检测未受到交叉项影响, 测得目标距离 R_1=13 973 m, R_2=

3 979.69 m, 目标速度 v_1=6 004.69 m/s, v_2=3 398.44 m/s, 均与目标真实值相符。

为更加直观地验证本文算法的检测性能, 采用表 2 中的雷达系统参数和目标运动参数, 选择的虚警概率分别为 $P_{\mathrm{fa}}=\{10^{-10},10^{-8},10^{-6},10^{-4}\}$, 并在高斯白噪声环境下, 对[−20, −6] dB信噪比区间内的信号进行 100 次蒙特卡罗实验, 测量本文算法在不同虚警率下的检测概率, 得到的曲线如图 4 所示。

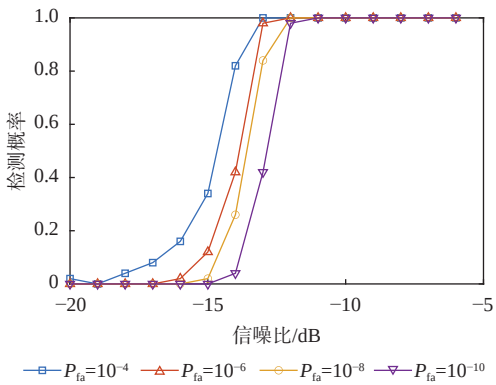


图 4 不同虚警率下的检测性能

Fig. 4 Detection performance under different false alarm rates

为对比本文算法与已有常见针对三阶运动模型的相干积累算法的能量积累性能, 设置目标检测恒虚警率 $P_{\mathrm{fa}}=10^{-6}$, 对每个信噪比进行 100 次蒙特卡罗实验, 对本文算法及 MTD[2]、TRT-SKT-LVD[13]、FR-LVD[1] 和 GRFT[10−11] 算法进行检测性能的对比仿真实验, 结果如图 5 所示。

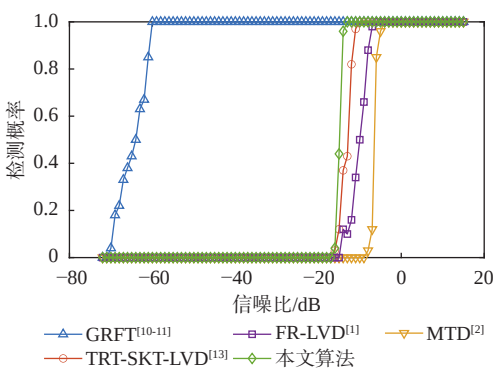


图 5 算法检测性能对比

Fig. 5 Comparison of algorithm detection performance

其中, GRFT 算法通过多维度参数搜索, 能够精准校正目标回波中的跨距离单元走动和多普勒频率徙动, 在高斯白噪声背景下作为最佳似然比检测器, 具备最好的积累性能, 但同时也需要庞大的运算量。MTD 算法在跨距离单元走动和多普勒频率徙动的影响下, 能量无法有效聚焦, 在同信噪比条件下, 相干积累检测性能最差。而 FR-LVD 算法虽然可以不经过参数搜索实现相干积累, 但其仅能够校正二阶运动引起的多普勒频率徙动, 无法校正加加速度带来的三阶多普勒频率徙动。因此, 在针对三阶运动模型时的检测性能较差, 同 MTD 算法类似, 能量无法有效聚焦。

本文算法在相同信噪比条件下的积累性能与 TRT-SKT-LVD 算法近似, 但与 TRT-SKT-LVD 算法仅获得目标加速度信息而丢失速度信息不同, 本文算法能够同时获得关键的目标距离和速度信息, 且具备相对更低的运算复杂度, 在检测性能与计算效率之间取得了良好的平衡。

4 结 论

本文针对可以近似为三阶运动的高速机动目标雷达探测, 提出一种基于序列反转的能量积累算法: 利用 TRT 和 ITRT 分离目标径向距离和速度信息, 并使用 NRFRT 消除多普勒频率徙动, 针对经过 ITRT-NRFRT 处理后回波中仍存在的由目标速度引起的跨距离单元走动问题, 采用变尺度傅里叶逆变换解耦合后实现成功积累。通过 TRT-NRFRT 和 ITRT-NRFRT-SCIFT 的联合算法, 可以估算出高速机动目标的径向距离和速度信息。

本文算法主要具有以下特点:

1) 与其他针对高速机动目标积累算法相比, 本文算法不需要参数搜索即可实现目标回波信号能量积累, 具备较好的实时性能。

2) TRT-NRFRT 和 ITRT-NRFRT-SCIFT 算法可以并行执行, 能够进一步提升运算速度。

参考文献 (References)

[1] WANG L, TAO H D, YANG A L, et al. A fast non-searching method for maneuvering weak target detection using ubiquitous radar[C]// Proceedings of the 2023 6th International Conference on Information Communication and Signal Processing. Piscataway: IEEE Press, 2023: 512-516.

[2] FROLUSHKIN V M , NOVOSELTSEV L I . Moving-target detection[J]. Izvestiya Vysshikh Uchebnykh Zavedenij. Radioelektronika, 1984, 27(7): 11-15.

[3] HUANG P H, LIAO G S, YANG Z W, et al. Long-time coherent integration for weak maneuvering target detection and high-order motion parameter estimation based on keystone transform[J]. IEEE Transactions on Signal Processing, 2016, 64(15): 4013-4026.

[4] LI Y, ZENG T, LONG T, et al. Range migration compensation and Doppler ambiguity resolution by keystone transform[C]//Proceedings of the 2006 CIE International Conference on Radar. Piscataway: IEEE Press, 2006: 1-4.

[5] XU J, YU J, PENG Y N, et al. Radon-Fourier transform for radar target detection (I): generalized Doppler filter bank[J]. IEEE Transactions on Aerospace and Electronic Systems, 2011, 47(2): 1186-1202.

[6] XU J , YU J , PENG Y N, et al. Radon-Fourier transform for radar

target detection (II): blind speed sidelobe suppression[J]. IEEE Transactions on Aerospace and Electronic Systems, 2011, 47(4): 2473-2489.

[7] YU J , XU J , PENG Y , et al. Radon-Fourier transform for radar target detection (III): optimality and fast implementations[J]. IEEE Transactions on Aerospace and Electronic Systems, 2012, 48(2): 991-1004.

[8] XU Z Y, ZHOU G J. Long-time coherent integration based on reconstructed spectrum for weak target detection using pulse-Doppler radar[J]. IEEE Sensors Journal, 2024, 24(5): 6597-6611.

[9] ZHANG Q L, ZHANG Y S, SUN B, et al. Motion parameters estimation and HRRP reconstruction of maneuvering weak target for wideband radar based on SKT-ELVD[J]. IEEE Transactions on Aerospace and Electronic Systems, 2023, 59(3): 2752-2763.

[10] XU J , XIA X G, PENG S B, et al. Radar maneuvering target motion estimation based on generalized radon-Fourier Transform[J]. IEEE Transactions on Signal Processing, 2012, 60(12): 6190-6201.

[11] MA B, ZHANG S S, JIA W K, et al. Fast implementation of generalized radon-Fourier transform[J]. IEEE Transactions on Aerospace and Electronic Systems, 2021, 57(6): 3758-3767.

[12] WAN J, HE Z Y, TAN X H, et al. Coherent integration for maneuvering target detection via fast nonparametric estimation method[J]. Signal Processing, 2023, 203: 108820.

[13] LI X L, CUI G L, YI W, et al. Fast coherent integration for maneuvering target with high-order range migration via TRT-SKT-LVD[J]. IEEE Transactions on Aerospace and Electronic Systems, 2016, 52(6): 2803-2814.

[14] HUANG P H, LIAO G S, YANG Z W, et al. An approach for refocusing of ground moving target without target motion parameter estimation[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(1): 336-350.

[15] KIRKLAND D . Imaging moving targets using the second-order keystone transform[J]. IET Radar, Sonar & Navigation, 2011, 5(8): 902-910.

[16] LV X L, BI G A, WAN C R, et al. Lv's distribution: principle, implementation, properties, and performance[J]. IEEE Transactions on Signal Processing, 2011, 59(8): 3576-3591.

[17] WAN J, KANG X H, TAN X H, et al. An efficient approach for coherent integration detection of high-speed maneuvering targets with arbitrary-order doppler frequency migrations[J]. IEEE Transactions on Aerospace and Electronic Systems, 2023, 59(5): 6729-6748.

[18] ZHENG J B, SU T, ZHU W T, et al. Radar high-speed target detection based on the scaled inverse Fourier transform[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 8(3): 1108-1119.

[19] LIN L J, SUN G H, CHENG Z Y, et al. Long-time coherent integration for maneuvering target detection based on ITRT-MRFT[J]. IEEE Sensors Journal, 2020, 20(7): 3718-3731.

Radar coherent integration method for high-speed maneuvering targets based on sequence reversal

WANG Ruizheng[1, 2], LI Shiqiang[1, *]

(1. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China;
2. School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: This paper proposes a joint algorithm based on improved time reversal transform-new range frequency reversal transform-scale-variable inverse Fourier transform (ITRT-NRFRT-SCIFT) and time reversal transform-new range frequency reversal transform (TRT-NRFRT) to solve the problems of cross-range cell migration and Doppler frequency migration in coherent integration of radar echoes from high-speed maneuvering targets. First, the slow time sequence is reversed to extract the target velocity information from the range information. Then, the NRFRT is applied to eliminate the effect of Doppler frequency migration. Fast Fourier transform (FFT) is utilized to achieve energy accumulation following TRT-NRFRT, whereas SCIFT is utilized following ITRT-NRFRT to achieve echo energy accumulation and remove the link between slow time and range frequency. Simulation experiments show that the algorithm can correct the cross-range cell migration and Doppler frequency migration of third-order moving targets without any parameter search, and obtain the radial distance and velocity information of the target with relatively low computational complexity.

Keywords: high speed maneuvering target; pulse train; sequence reversal; range migration; Doppler migration

Received: 2024-04-29; **Accepted**: 2024-08-23; **Published Online**: 2024-12-11 09:50

URL: link.cnki.net/urlid/11.2625.V.20241210.1555.002

Foundation item: Pre-research Project of Civil Aviation Technology in Administration of Science, Technology and Industry (D010105)

* **Corresponding author**. E-mail: lishq@mail.ie.ac.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0258

联合语义分割和边缘纹理的人脸图像修复

石计亮[1,2]，张乾[2,3,*]，周遵富[1,2]，杨思红[1,2]

(1. 贵州民族大学 数据科学与信息工程学院，贵阳 550025；2. 贵州省模式识别与智能系统重点实验室，贵阳 550025；3. 贵州民族大学教务处，贵阳 550025)

摘　　要：现有的图像修复方法通过预测辅助结构信息来填充逼真的补丁，但不准确的先验可能导致不合理的结构和模糊的纹理。同时，现有方法仅关注原始图像与修复图像之间的关系，未充分利用受损图像的信息。针对上述问题，提出一种端到端的Transformer人脸图像修复网络，利用语义分割和边缘纹理信息引导修复过程。其中，主修复网络包含一个RGB修复分支、2个用于语义分割和边缘纹理的辅助分支。在编码器中设计一组大核卷积上下文瓶颈(LKCCB)模块，以增加有效感受野和更好地上下文推理。为捕获遥远距离的上下文信息，提出嵌套动态辅助归一化多头注意力(NDAN-MHA)模块，其中，含有的动态辅助归一化(DAN)模块能够动态整合3个分支的结构特征，以此丰富语义一致性。此外，提出引入对比正则化(CR)网络来稳定和改进网络的训练，以生成更真实的修复图像。在CelebA-HQ和FFHQ数据集上进行定性和定量实验，结果表明：所提方法在主客观指标上均优于对比方法，能够合理地修复大面积不规则遮挡的人脸图像。

关 键 词：人脸图像修复；对比学习；大核卷积；注意力机制；动态辅助归一化
中图分类号：TP391.41
文献标志码：A　　**文章编号**：1001-5965(2026)06-2194-14

图像修复技术旨在填补图像中缺失或损坏的区域，使其内容连贯且可信。其目标是无缝地恢复图像，使修复区域与原始内容难以区分。相关研究已经取得了出色的成果，这些成果已广泛应用于实际场景，如物体移除和旧照片恢复等。

现有的图像修复方法采用编码器-解码器架构，假设损坏图像具有足够的先验知识。Nazeri等[1]提出一种2阶段图像修复模型，第1阶段根据受损图像预测缺失区域的图像结构，第2阶段将受损图像和预测的边缘图作为输入，从而更好地指导网络进行图像修复。Yang等[2]提出一种结合梯度图来辅助图像修复的多任务学习框架，训练一个共享生成器来同时完成损坏的图像修复和相应的梯度图，并通过结构嵌入方案将学习到的结构特征嵌入图像修复过程中，为图像补全提供可能的先验条件。Yamashita等[3]提出一种2阶段的图像修复方法，利用边缘图和深度图像中的边界作为图像修复的辅助线索，进一步提高了图像修复的性能。谭骏珊等[4]提出一种基于推理注意力机制的2阶段网络图像修复方法。第1阶段通过边缘生成网络预测合理的幻觉边缘信息，第2阶段通过图像补全网络完成图像的重建。该方法具有较高的图像修复精确度，生成的图像结构较为合理。Shao等[5]提出一种双流耦合网络用于图像修复任务，该网络可有效执行结构与纹理之间的双向交互。具体来说，该网络构建2个并行的编码器-解码器，分别用于重建受损的边缘图和RGB图像，并设计门控交互单元和空间通道注意的交互机制。其中，门控交互单元模块通过

收稿日期：2024-04-26；**录用日期**：2024-07-05；**网络出版时间**：2024-12-11 10:58
网络出版地址：link.cnki.net/urlid/11.2625.V.20241210.1723.003
基金项目：贵州民族大学校级科研项目(GZMUZK［2021］YB23)
*通信作者. E-mail：gzmuzq@gzmu.edu.cn

引用格式：石计亮，张乾，周遵富，等. 联合语义分割和边缘纹理的人脸图像修复[J]. 北京航空航天大学学报，2026，52(6)：2194-2207. SHI J L，ZHANG Q，ZHOU Z F，et al. Face image inpainting combining semantic segmentation and edge texture[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：2194-2207 (in Chinese).

门控机制权衡结构与纹理特征之间的相互作用, 空间通道注意机制通过空间通道注意力实现结构与纹理远距离上下文的深入交互, 从而有效地保证生成图像结构和纹理的合理性。

尽管现有修复方法在引入辅助先验(如边缘、轮廓)方面做出努力以提升修复性能, 但仍存在以下不足: ①使用预测的辅助结构来指导图像修复, 在测试阶段没有真实数据的情况下, 可能会导致修复结果出现更多的偏离和错误; ②此前的工作通常将图像修复任务分为多个阶段, 在潜在的真实应用中耗费大量时间; ③现有基于深度学习的修复方法主要关注目标图像与真实图像之间的相关性, 而受损图像的信息未被充分利用。因此, 探索受损图像与目标图像之间的关联性对于提高模型性能和修复效果至关重要。

为应对上述挑战, 本文提出一种端到端的 Transformer 人脸图像修复网络, 采用 U-Net 风格的编码器-解码器架构。首先, 提出了大核卷积上下文瓶颈(large kernel convolutional context bottlenecks, LKCCB)模块, 能够在早期增加感受野和更好地进行上下文推理。其次, 在解码阶段将编码后的特征解码为 RGB 图像、语义分割图和边缘图, 其包括 1 个 RGB 修复分支和 2 个辅助分支。与现有的直接指导方法不同, 本文的重点在于共同学习 3 个分支之间的无偏判别相互作用信息。具体来说, 设计了嵌套动态辅助归一化多头注意力(nested dynamic auxiliary normalization multi-head attention, NDAN-MHA)模块, 通过动态辅助归一化(dynamic auxiliary normalization, DAN)模块整合 3 个分支的特征, 以减少图像修复中的噪声内容。最后, 提出了一个基于对比学习的对比正则化(contrastive regularization, CR)网络, 用于图像修复。该网络将受损图像、网络生成的修复图像和真实图像分别表示为负样本、锚点和正样本, 通过对比学习将锚点图像限制在封闭边界内, 以提高修复性能。整个网络在真实 RGB 图像、语义分割图和边缘图的共同作用下以端到端的方式进行训练。本文的主要创新点和贡献如下:

1) 提出了一种端到端的 Transformer 人脸图像修复网络, 该网络结合语义分割图和边缘图来辅助网络进行图像修复。

2) 设计了一组 LKCCB 模块, 以在早期增加感受野和更好地进行上下文推理; 还设计了 NDAN-MHA 模块, 通过 DAN 模块动态地整合 3 个分支的特征, 并专注于上下文长期依赖关系, 以便捕获足够的上下文信息。

3) 将对比学习引入到图像修复任务中, 并考虑修复图像与受损图像和真实图像之间的相反关系。本文提出的基于对比学习的 CR 网络进一步提高了图像修复的性能。

4) 本文方法在多个公开人脸数据集上表现出优异的修复性能, 特别是对于大面积不规则掩码。

1　相关工作

1.1　带有辅助结构先验的图像修复

由于修复缺失区域问题本质上是不特定的, 因此, 使用额外的结构先验(如 Canny 边缘图、语义分割图和轮廓等)来辅助模型以获得更真实的修复结果。Li 等[6] 提出一种渐进重建视觉结构网络, 其中, 设计的视觉结构重建层采用部分卷积和瓶颈块来恢复缺失区域的部分边缘, 再将其与带孔的图像结合, 并通过语义上有意义的内容来逐步填补缺失区域。Xu 等[7] 使用预训练的整体嵌套边缘检测模型从受损图像中提取边缘图, 并通过边缘修复网络来填补边缘图中缺失的区域, 进一步使用边缘图来指导网络恢复受损图像。Guo 等[8] 提出一种用于图像修复的双流网络, 该网络以耦合的方式对结构约束的纹理合成和纹理引导的结构重建进行建模, 使双生成任务更好地相互促进。陈晓雷等[9] 提出结合语义先验和深度注意力残差组的图像修复方法, 其中, 语义先验网络学习缺失区域视觉元素的完整语义先验信息, 并用学习的语义信息填补缺失区域。上述方法提取结构先验信息进行图像修复任务, 取得较好的修复效果。然而大面积遮挡下提取或重建结构先验可能不准确, 导致语义边界处纹理模糊。

近期受扩散模型优异性能的启发, 研究人员提出了基于辅助结构(如文本、边缘)的图像补全解决方案, 以增强完整图像的结构和纹理。Horita 等[10] 设计了一种结构引导的扩散模型, 用于图像补全, 该模型明确地考虑使用边缘信息作为结构指导。该方法由 2 个级联的扩散概率模型组成: 结构生成器和纹理生成器。其中, 结构生成器用于生成合理的边缘图, 而纹理生成器完成基于边缘引导的纹理生成。为联合训练 2 个生成器, 设计一种利用最优贝叶斯去噪的新策略, 以在单个步骤中对结构生成器的输出进行去噪, 从而允许反向传播。该方法可以修复具有不同结构的连贯大孔, 实现更好的视觉质量, 但其复杂的联合训练方式需要更多的迭代时间。Nichol 等[11] 探讨了扩散模型在文本条件下图像合成问题上的应用, 使用无分类器指导的 35 亿参数文本条件扩散模型生成逼真的样本, 并在此模型上进行微调以执行图像修复。在微调过程中, 随

机擦除训练样本的某些区域，将剩余部分与掩码一起输入模型中，作为额外的条件信息。该模型根据给定的文本提示填充缺失区域，能够匹配图像周围上下文的信息，生成逼真的修复结果。

1.2 对比学习

在表示空间中，对于给定的锚点，对比学习旨在将锚点拉近到正样本，并将其推离负样本。Ma 等[12]提出一种自监督连体推理网络来改善上下文编码，其方法是将 2 个具有不同掩码的相同图像拉近到上下文特征空间中。由于具有不同掩码的相同图像可能存在差异，因此，该方法的策略可能不合理。为解决此问题，Zuo 等[13]有效地将对比学习应用于图像修复任务，该方法提出纹理和语义对比学习损失，通过利用鉴别器的特征表示空间来稳定和改进修复模型的训练，其中，修复图像被拉近真实图像，并被推离损坏图像，从而更好地引导修复图像从损坏图像到真实图像的转移。与文献 [13] 不同，本文选择来自相同固定预训练模型 VGG-19 的公共中间特征表示空间。

2 本文方法

给定原始图像 $\boldsymbol{I}_{\mathrm{gt}} \in \mathbf{R}^{H\times W\times 3}$ 和掩码 $\boldsymbol{M} \in \mathbf{R}^{H\times W\times 3}$，其中，$\mathbf{R}$ 为实数域，则原始图像经过退化处理后成为受损图像 $\boldsymbol{I}_{\mathrm{m}}$，定义为

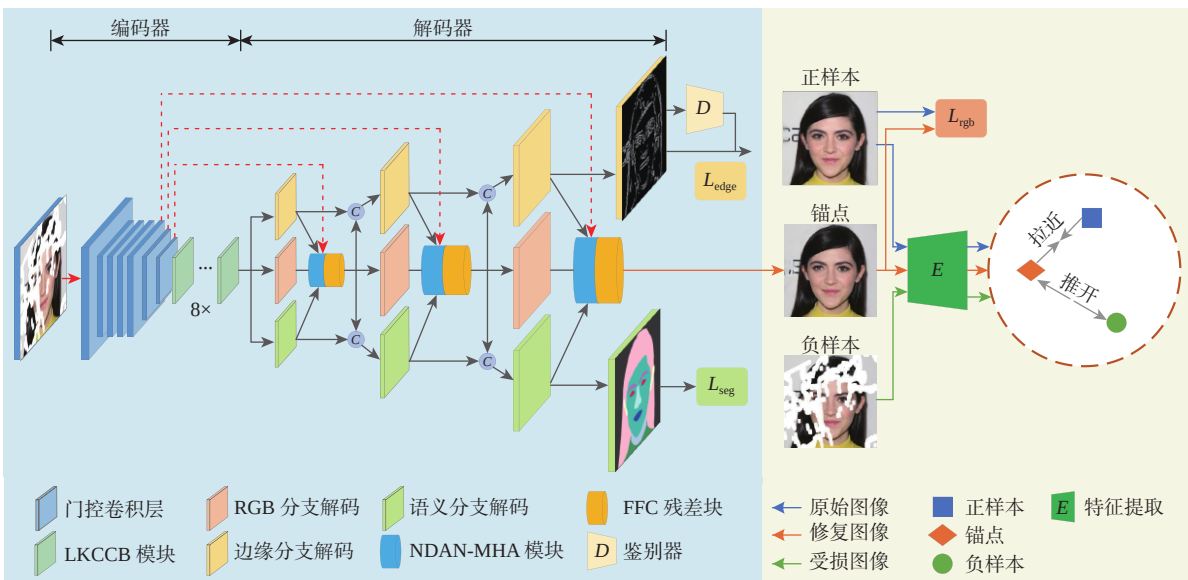
$$\boldsymbol{I}_{\mathrm{m}} = \boldsymbol{I}_{\mathrm{gt}} \odot \boldsymbol{M} \tag{1}$$

式中：$\boldsymbol{M}$ 的像素值为 0 或 1，分别表示受损图的缺失区域和有效区域；“$\odot$”表示 Hadamard 乘积。

主修复网络将受损图像 $\boldsymbol{I}_{\mathrm{m}}$ 和掩码 $\boldsymbol{M}$ 作为输入，输出修复图像的估计值 $\boldsymbol{I}_{\mathrm{out}} \in \mathbf{R}^{H\times W\times 3}$。通过将 $\boldsymbol{I}_{\mathrm{out}}$ 的掩码区域和 $\boldsymbol{I}_{\mathrm{m}}$ 的有效区域相结合得到最终修复图像，定义为

$$\boldsymbol{I}_{\mathrm{comp}} = (1-\boldsymbol{M}) \odot \boldsymbol{I}_{\mathrm{out}} + \boldsymbol{M} \odot \boldsymbol{I}_{\mathrm{m}} \tag{2}$$

本文保留 U-Net 风格的编码器-解码器架构，包括主修复网络和 CR 网络，其中，主修复网络主要包括大核卷积上下文编码器和解码器。网络整体框架如图 1 所示。



(a) 主修复网络　　(b) CR 网络

图 1 网络整体框架

Fig. 1 Overall framework of the network

2.1 大核卷积上下文编码器

为增强网络在理解图像上下文信息方面的能力，本文设计了一组 LKCCB 模块，通过增加有效感受野，有利于网络保留更多的空间结构信息，以处理图像并聚合更多的空间上下文信息，如图 2 所示。LKCCB 模块由 4 条并行路径组成，每条路径包括具有不同扩张率的卷积层和带有分解大核卷积块（decomposition large kernel convolution block, DLKCB）的权重。DLKCB 遵循文献 [14]，即将 $K\times K$ 卷积分解为带有扩张率为 d 的 $(K/d)\times(K/d)$ 深度扩张卷积（depth-wise dilated convolution, DWDConv）、$(2d-1)\times(2d-1)$ 的深度卷积（depth-wise convolution, DWConv）和 2 个 1×1 的标准卷积。此外，LKCCB 模块中还有一个门控掩码，用于加权扩张卷积。因此，编码器可以更好地理解图像中的细节和上下文信息，并根据当前的缺失区域动态选择合适的操作路径，以更好地完成图像修复任务。本文选择 $K=21$ 和 $d=3$，在消融实验中进一步说明。

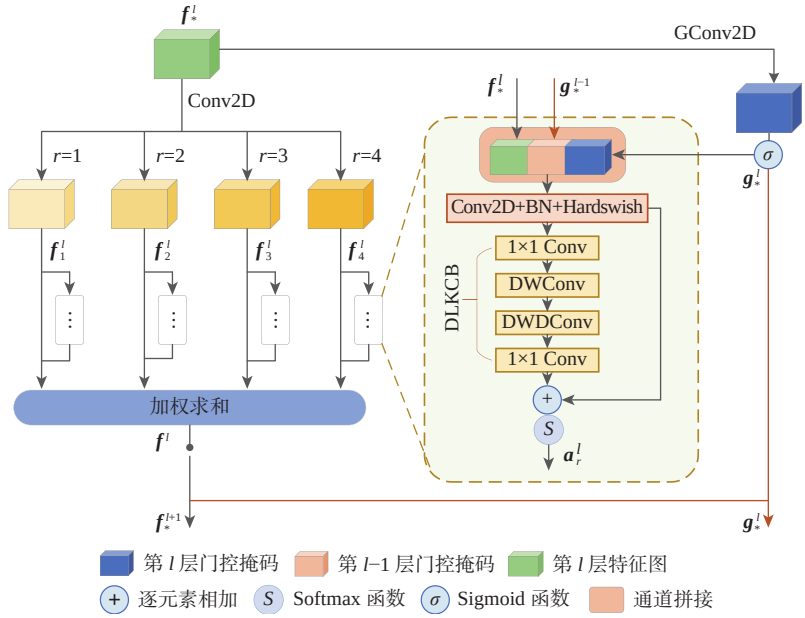


图 2　LKCCB 模块

Fig. 2　LKCCB module

将残缺图像 $\boldsymbol{I}_{\mathrm{m}}$ 经过最后一层门控卷积层的特征图记为 $\boldsymbol{f}_*^0$，门控掩码记为 $\boldsymbol{g}_*^0$。更新每一层 l 的特征 $\boldsymbol{f}_*^l$ 和门控掩码 $\boldsymbol{g}_*^l$，其中，$l=1,2,\cdots,L$。门控掩码 $\boldsymbol{g}_*^l$ 用于估计网络在第 l 层中基于特征图的信息，以确定图像中缺失区域的概率，在第 l 层的特征 $\boldsymbol{f}_*^l$ 上使用门控卷积可得到 $\boldsymbol{g}_*^l$，该过程表示为

$$\boldsymbol{g}_*^l=\mathrm{GConv}(\boldsymbol{f}_*^l) \tag{3}$$

式中：GConv(·)表示门控卷积操作。

对于第 l 层的特征 $\boldsymbol{f}_*^l$，分别使用膨胀率 r 为 1、2、3、4 的卷积层来捕获多尺度语义特征，从而得到每条路径的特征 $\boldsymbol{f}_r^l$，该过程表示为

$$\boldsymbol{f}_r^l=\mathrm{Conv}_r(\boldsymbol{f}_*^l) \tag{4}$$

对应于每条路径的权重 $\boldsymbol{a}_r^l$，将第 l 层特征 $\boldsymbol{f}_*^l$、门控掩码 $\boldsymbol{g}_*^l$ 和第 $l-1$ 层门控掩码 $\boldsymbol{g}_*^{l-1}$ 在通道维度上拼接，并通过卷积核大小为 3×3 的标准卷积、批归一化和 Hardswish 激活函数操作，得到新的门控掩码 $\boldsymbol{g}_r^l$，该过程表示为

$$\boldsymbol{g}_r^l=\delta(\mathrm{BN}(\mathrm{Conv}_{3\times3}([\boldsymbol{f}_*^l;\boldsymbol{g}_*^l;\boldsymbol{g}_*^{l-1}]))) \tag{5}$$

式中：[·;·;·]为通道拼接；$\mathrm{Conv}_{3\times3}(\cdot)$ 为 3×3 标准卷积操作；BN(·)为 BatchNorm 函数；$\delta(\cdot)$ 为 Hardswish 激活函数。

为突出来自遥远空间位置的上下文信息，使网络能够学习具有大感受野的遮挡区域的特征表示，使用 DLKCB 块对 $\boldsymbol{g}_r^l$ 进行重建，以生成带有更多空间结构信息和上下文信息的门控掩码，该过程表示为

$$\boldsymbol{a}_{tt}=\mathrm{Conv}_{1\times1}(\mathrm{DWDConv}(\mathrm{DWConv}(\mathrm{Conv}_{1\times1}(\boldsymbol{g}_r^l)))) \tag{6}$$

式中：$\boldsymbol{a}_{tt}$ 为 DLKCB 块处理后的特征图；$\mathrm{Conv}_{1\times1}(\cdot)$ 为 1×1 标准卷积；DWDConv(·)为扩张率为 3 的 7×7 深度扩张卷积；DWConv(·)为 5×5 深度卷积。

后面接着 Softmax 函数，从而得到第 l 层膨胀率为 r 的权重 $\boldsymbol{a}_r^l$，该过程表示为

$$\boldsymbol{a}_r^l=\mathrm{Softmax}(\boldsymbol{a}_{tt}+\boldsymbol{g}_r^l) \tag{7}$$

通过元素加权求和聚合 4 条并行路径的特征，得到含有丰富上下文信息的多尺度语义特征 $\boldsymbol{f}^l$，该过程表示为

$$\boldsymbol{f}^l=\sum_{r=1}^{4}\boldsymbol{a}_r^l\boldsymbol{f}_r^l \tag{8}$$

为缓解缺失区域内外输入像素值之间的差异及修复后的图像中存在颜色差异问题。在 LKCCB 块中使用新颖的门控残差连接[15]，LKCCB 块使用 $\boldsymbol{g}_*^l$ 进行加权求和，将输入特征 $\boldsymbol{f}_*^l$ 与学习到的多尺度语义特征 $\boldsymbol{f}^l$ 聚合在一起，该过程表示为

$$\boldsymbol{f}_*^{l+1}=\boldsymbol{f}_*^l\times\boldsymbol{g}_*^l+\boldsymbol{f}^l\times(1-\boldsymbol{g}_*^l) \tag{9}$$

式中：$\boldsymbol{f}_*^{l+1}$ 为第 l 层的输出特征图、第 $l+1$ 层的输入特征图。本文选择堆叠 8 个 LKCCB 块，消融实验中进一步说明。

2.2　解码器

给定编码后的特征 $\boldsymbol{f}_*$，解码器利用一系列堆叠的 Transformer 块来处理结构化信息。解码器包括 3 个分支：1 个用于修复受损图像的 RGB 修复分支及 2 个附加分支，其中，包含语义分割和边缘纹理信息。如图 1 所示，在每个 Transformer 块内，首先，通过本文设计的 NDAN-MHA 模块来计算 3 个分支

的特征和编码器的特征之间的注意力。进一步使用快速傅里叶卷积(fast Fourier convolutions, FFCs)残差块[16]增强全局特征表达能力。其次,将增强的特征进行拆分,并将其与每个分支上的特征映射相结合,以进行下一阶段的注意力计算。经过3个阶段的解码后,网络预测出所有的修复图像、边缘图和语义分割图。本节主要介绍提出的DAN模块和NDAN-MHA模块。

2.2.1 DAN模块

为了避免在融合3个分支特征的过程中丢失语义信息,设计DAN模块。与原始空间自适应归一化(spatially-adaptive normalization, SPADE)[17]模块不同,本文在DAN模块中使用像素级动态激活函数,以便调制参数可以灵活地适应局部信息。metaACON激活函数[18]根据其值自动确定神经元是否处于激活状态,相应公式为

$$f(\boldsymbol{x})=(\boldsymbol{p}_1-\boldsymbol{p}_2)\,\boldsymbol{x}\cdot\sigma(\boldsymbol{\theta}\,(\boldsymbol{p}_1-\boldsymbol{p}_2)\boldsymbol{x})+\boldsymbol{p}_2\boldsymbol{x} \tag{10}$$

式中:$\boldsymbol{p}_1$和$\boldsymbol{p}_2$为可学习的参数;$\sigma(\cdot)$为Sigmoid函数;$\boldsymbol{\theta}=\sigma(\boldsymbol{x})$。

分别使用$\boldsymbol{f}_{\text{edge}}$、$\boldsymbol{f}_{\text{seg}}$和$\boldsymbol{f}_{\text{in}}$来表示边缘分支、语义分支和RGB修复分支的输入特征映射,如图3所示。DAN模块中的第一个卷积层使用像素级动态激活函数metaACON,并构造metaACON激活函数的变体,即式(10)中的$\boldsymbol{\theta}$参数,该参数是根据2个辅助分支预测的结构信息$\boldsymbol{f}_{\text{edge}}$和$\boldsymbol{f}_{\text{seg}}$来学习得到的,如式(11)所示。另外,调制参数$\boldsymbol{\alpha}$和$\boldsymbol{\beta}$是基于辅助分支的结构特征$\boldsymbol{f}_{\text{edge}}$和$\boldsymbol{f}_{\text{seg}}$,并通过2个卷积核大小为$1\times1$的卷积层学习得到,如图3所示。通过这样的方式,多分支特征信息基于辅助结构的上下文进行合并,该过程可表示为

$$\boldsymbol{\theta}=\sigma(\text{Conv}_{1\times1}(\rho(\text{Conv}_{1\times1}([\boldsymbol{f}_{\text{edge}};\boldsymbol{f}_{\text{seg}}])))) \tag{11}$$

$$\boldsymbol{\alpha}=\text{Conv}_{1\times1}(\rho(\text{Conv}_{1\times1}([\boldsymbol{f}_{\text{edge}};\boldsymbol{f}_{\text{seg}}]))) \tag{12}$$

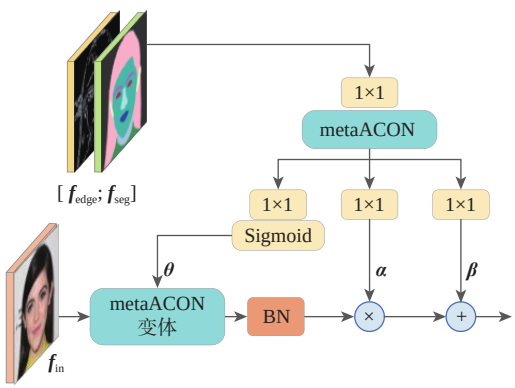


图3 DAN模块

Fig. 3 DAN module

$$\boldsymbol{\beta}=\text{Conv}_{1\times1}(\rho(\text{Conv}_{1\times1}([\boldsymbol{f}_{\text{edge}};\boldsymbol{f}_{\text{seg}}]))) \tag{13}$$

$$\text{DAN}(\boldsymbol{f}_{\text{in}}|[\boldsymbol{f}_{\text{edge}};\boldsymbol{f}_{\text{seg}}])=\boldsymbol{\alpha}\otimes\text{BN}(\rho'(\boldsymbol{f}_{\text{in}}))+\boldsymbol{\beta} \tag{14}$$

式中:$[\cdot;\cdot]$为通道拼接;$\rho(\cdot)$为metaACON激活函数操作;BN(·)为BatchNorm函数;$\rho'(\cdot)$为metaACON激活函数的变体;DAN(·)为DAN模块操作。

2.2.2 NDAN-MHA模块

线性统一嵌套注意力机制(linear unified nested attention, LUNA)[19]使用2个嵌套线性注意力函数来近似Softmax注意力,从而解决了Transformer的二次内存和计算复杂度问题。受此启发,本文设计NDAN-MHA模块,允许网络有效地建模长期依赖关系,并接受来自编码器的特征,以便在解码阶段捕获足够的上下文信息,如图4所示。

NDAN-MHA模块将编码器的特征$\boldsymbol{E}_n$作为额外的输入,并接收来自RGB修复分支解码器的特征$\boldsymbol{D}_k$和2个辅助分支解码器的特征$\boldsymbol{f}_{\text{seg}}$、$\boldsymbol{f}_{\text{edge}}$。进一步,将DAN模块嵌入到多头注意力中,称为动态辅助归一化多头注意力(dynamic auxiliary normalization multi-head attention, DAN-MHA),如图4所示。NDAN-MHA模块将编码器的特征和解码器3个分支的特征作为输入,通过DAN模块合并RGB修复分支和2个辅助分支的特征,并专注于上下文长期依赖关系,以便提供更多的上下文信息。NDAN-MHA模块中的第1个打包注意力输出$\boldsymbol{Y}_{\text{P}}$和第2个解包注意力输出$\boldsymbol{Y}_{\text{U}}$可表示为

$$\boldsymbol{Y}_{\text{P}}=\text{DAN-MHA}\,(\boldsymbol{D}_k,\boldsymbol{E}_n) \tag{15}$$

$$\boldsymbol{Y}_{\text{U}}=\text{DAN-MHA}\,(\boldsymbol{Y}_{\text{P}},\boldsymbol{D}_k) \tag{16}$$

将$\boldsymbol{Y}_{\text{P}}$和$\boldsymbol{Y}_{\text{U}}$馈送到层归一化和本文设计的辅助前馈网络(auxiliary feed-forward network, AFFN)。AFFN模块中嵌入一个DAN块,再次将辅助分支的特征$\boldsymbol{f}_{\text{edge}}$、$\boldsymbol{f}_{\text{seg}}$与主修复分支的特征合并,以便提供丰富的上下文信息。该过程可表示为

$$\boldsymbol{Y}_{\text{E}}=\text{LN}(\boldsymbol{Y}_{\text{P}}+\boldsymbol{E}_n) \tag{17}$$

$$\boldsymbol{Y}_{\text{D}}=\text{LN}(\boldsymbol{Y}_{\text{U}}+\boldsymbol{D}_k) \tag{18}$$

$$\boldsymbol{Y}'_{\text{D}}=\text{DAN}(\boldsymbol{Y}_{\text{D}}|[\boldsymbol{f}_{\text{edge}};\boldsymbol{f}_{\text{seg}}]) \tag{19}$$

$$\boldsymbol{y}=\text{DWConv}_{3\times3}(\text{Conv}_{1\times1}(\boldsymbol{Y}'_{\text{D}})) \tag{20}$$

$$\text{AFFN}(\boldsymbol{Y}_{\text{D}})=\text{Conv}_{1\times1}(\boldsymbol{y}\otimes G(\boldsymbol{y})) \tag{21}$$

式中:$\boldsymbol{Y}_{\text{D}}$为$\boldsymbol{Y}_{\text{U}}$经过LayerNorm函数操作后的特征图;$\boldsymbol{Y}_{\text{E}}$为$\boldsymbol{Y}_{\text{P}}$经过LayerNorm函数操作后的特征图;$\boldsymbol{Y}'_{\text{D}}$为DAN模块处理的特征图;LN(·)为LayerNorm函数;$G(\cdot)$为GELU激活函数操作;$\text{DWConv}_{3\times3}(\cdot)$为$3\times3$深度卷积;AFFN(·)为辅助前馈网络操作。

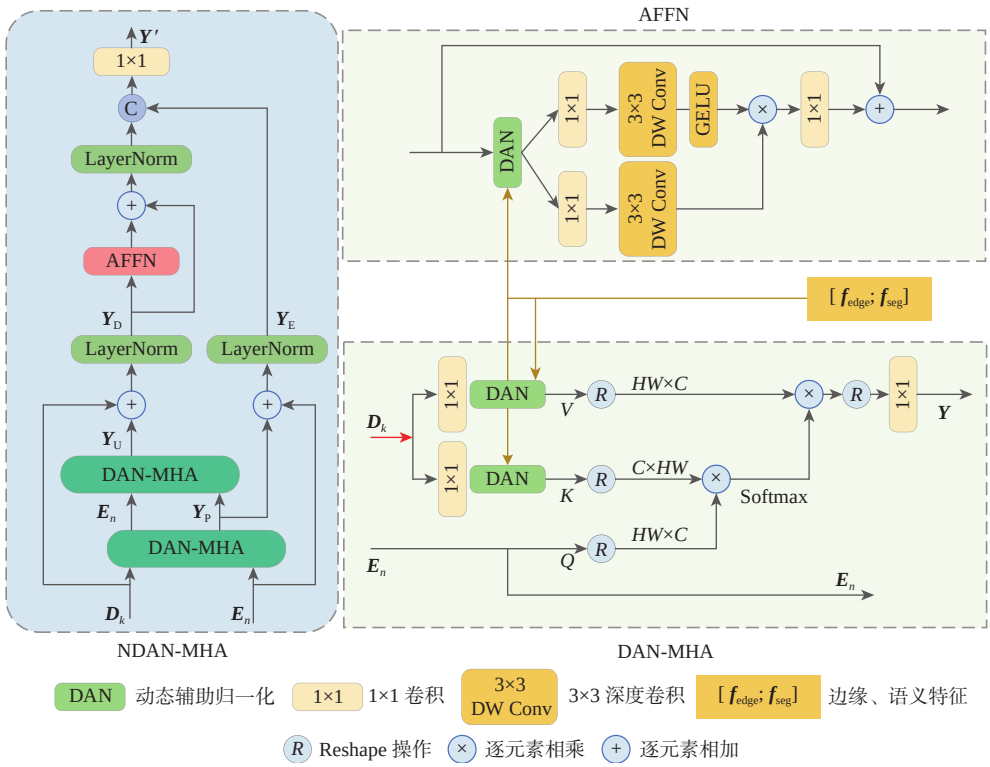


图 4　NDAN-MHA 模块

Fig. 4　NDAN-MHA module

最终得到 NDAN-MHA 模块的输出特征图 $\boldsymbol{Y}'$, 该过程可表示为

$$\boldsymbol{Y}' = \mathrm{Conv}_{1\times1}[\mathrm{LN}(\mathrm{AFFN}(\boldsymbol{Y}_\mathrm{D}) + \boldsymbol{Y}_\mathrm{D}); \boldsymbol{Y}_\mathrm{E}] \tag{22}$$

2.3　CR 网络

对比学习的目标是学习一种表示, 在嵌入空间中拉近正样本对, 并在负样本对之间推开表示。为提高修复模型的泛化能力, 考虑以下几个方面: ①确定哪些图像是正样本对和负样本对; ②选择一个潜在的特征空间; ③构建一个有效的对比约束。引入对比学习来提高和加强图像修复模型的训练效果, 并选择来自相同固定预训练模型 VGG-19 的公共中间特征表示空间, 对修复图像进行训练, 以便更好地从受损图像向真实图像过渡。此外, 分别将真实图像、残缺图像和修复图像称为正样本、负样本和锚点, 由真实图像和修复图像组成正样本对, 残缺图像和修复图像组成负样本对。使模型学习在表示空间中将修复图像与真实图像拉近, 并将其与残缺图像推开, 如图 1(b) 所示。为稳定和改进修复模型的训练效果, 使用与文献 [13] 类似的损失函数:

$$L_\mathrm{contra} = A\left[\sum_{i=1}^{n}\omega_i \cdot \frac{d(\boldsymbol{G}_i(\boldsymbol{I}_\mathrm{comp}), \boldsymbol{G}_i(\boldsymbol{I}_\mathrm{gt}))}{d(\boldsymbol{G}_i(\boldsymbol{I}_\mathrm{comp}), \boldsymbol{G}_i(\boldsymbol{I}_\mathrm{m}))}\right] \tag{23}$$

式中: A 为期望; ω_i 为权重系数; $d(\cdot,\cdot)$ 为距离度量, 采用 L_1 范数实现; $\boldsymbol{G}_i(\cdot)$ 为从固定的预训练模型 VGG-19 第 i 层提取的潜在特征。本文设置 $i = 1,3,5,9,13$, 相应的权重分别设置为 $\omega_i = 1/32, 1/16, 1/8, 1/4, 1$。

3　损失函数

为训练网络, 本文总损失函数 L_total 包括 L_rgb、L_edge、L_seg 和 L_contra, 即

$$L_\mathrm{total} = L_\mathrm{rgb} + \lambda_\mathrm{edge}L_\mathrm{edge} + \lambda_\mathrm{seg}L_\mathrm{seg} + \lambda_\mathrm{contra}L_\mathrm{contra} \tag{24}$$

式中: L_rgb 为 RGB 修复分支的损失函数, 遵循文献 [20] 的工作, 由重构损失、感知损失、风格损失和总变分损失组成; L_edge、L_seg 和 L_contra 分别为边缘分支、语义分支和 CR 网络的损失函数; λ_edge、λ_seg 和 λ_contra 分别为各分支损失函数的权重值。

与文献 [1] 类似, 使用二元交叉熵和对抗损失函数来训练边缘纹理分支, 定义为

$$L_\mathrm{edge} = w_\mathrm{BCE}L_\mathrm{BCE} + L_\mathrm{adv} \tag{25}$$

$$L_\mathrm{BCE} = \frac{1}{N}\sum_{i=1}^{N} -\left[\boldsymbol{C}_\mathrm{gt}^i \mathrm{In}\boldsymbol{C}_\mathrm{pred}^i + (1-\boldsymbol{C}_\mathrm{gt}^i)\mathrm{In}(1-\boldsymbol{C}_\mathrm{pred}^i)\right] \tag{26}$$

$$L_\mathrm{adv} = -E\left[D(\boldsymbol{C}_\mathrm{pred})\right] \tag{27}$$

式中: w_BCE 为平衡权重; $\boldsymbol{C}_\mathrm{gt}^i$ 为第 i 个样本基于 Canny 算子[1] 的真实边缘图; $\boldsymbol{C}_\mathrm{pred}^i$ 为第 i 个样本的预测边缘图; D 为 5 个卷积层组成的光谱归一化鉴别器。

对于语义分支, 使用交叉熵损失函数来训练, 定义为

$$L_{\text{seg}}=\frac{1}{N}\sum_{i=1}^{N}-\boldsymbol{S}_{\text{gt}}^{i}\ln\boldsymbol{S}_{\text{pred}}^{i} \tag{28}$$

式中：$\boldsymbol{S}_{\text{gt}}^{i}$和$\boldsymbol{S}_{\text{pred}}^{i}$分别为第$i$个样本的真实语义分割图和预测的语义分割图。

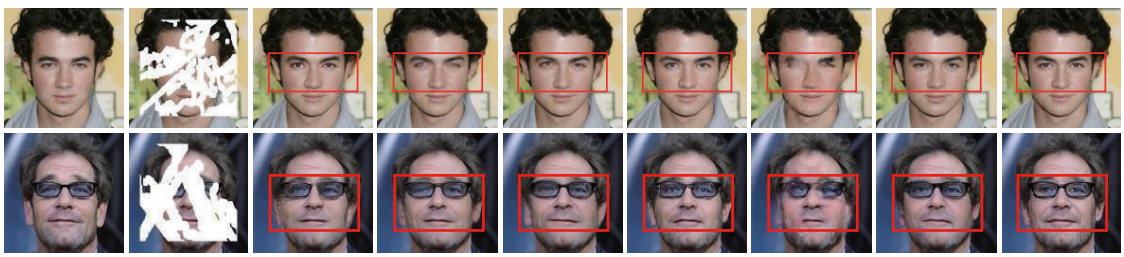
4 实验结果与分析

采用Linux系统、Nvidia GPU A100图形处理器、Pytorch 2.0.1版本进行实验。在训练过程中，使用Adam优化器对模型进行训练。遵循文献[21]的设置，模型生成器的学习率设置为2×10^{-4}，鉴别器的学习率设置为1×10^{-5}，批次大小设置为6，各分支损失函数的权重值分别设置为$\lambda_{\text{seg}}=0.5$、$\lambda_{\text{edge}}=0.5$和$\lambda_{\text{contra}}=1$，所有实验均在同一条件下训练24万次迭代。

本文使用2个公开人脸数据集和文献[20]的不规则掩码数据集。CelebA-HQ[22]数据集包含30 000张高清人脸图像，其中，每张图像都有对应19个人脸类别的语义分割掩码。本文抽取3 000张图像作为测试集，其余27 000张图像作为训练集。FFHQ[23]数据集包含70 000张高质量的PNG图像，分辨率为1 024像素×1 024像素，其中，每张图像也有对应19个人脸类别的语义分割掩码。本文抽取60 000张图像作为训练集，其余10 000张图像作为测试集。实验中的边缘图由文献[21]提供的代码得到，在训练和测试过程中，所有图像和掩码均被调整为256像素×256像素。

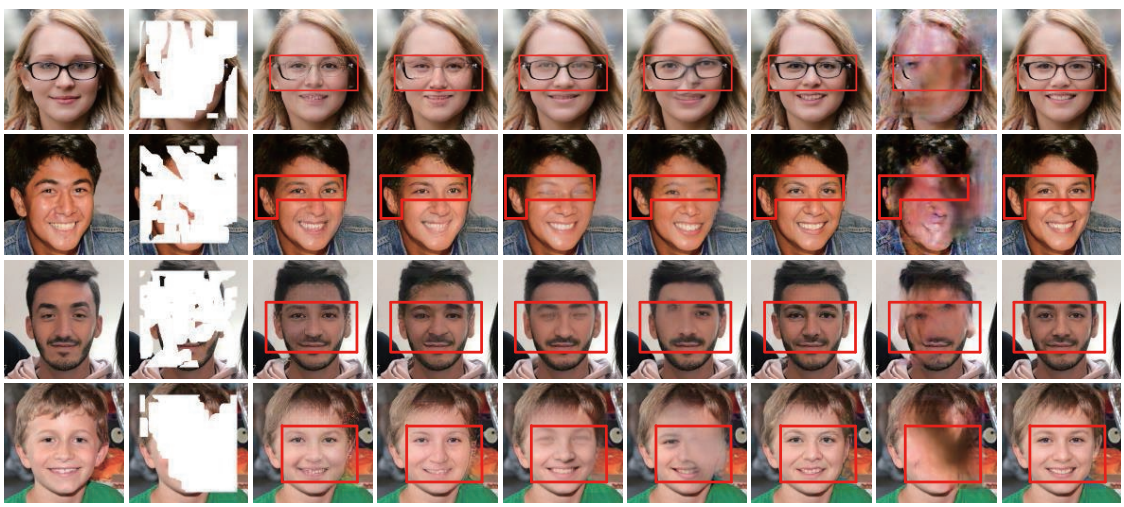
4.1 定性评价

为更好地验证本文网络修复能力，将本文方法与PRVS[6]、AOT[15]、CTSDG[8]、HAN[24]、T-former[25]、MMT[21]和HINT[26]方法进行比较，定性实验结果如图5所示。



(a1) 原图 (a2) 残损图 (a3) CTSDG[8] (a4) HAN[24] (a5) T-former[25] (a6) MMT[21] (a7) AOT[15] (a8) HINT[26] (a9) 本文方法

(a) CelebA-HQ数据集



(b1) 原图 (b2) 残损图 (b3) PRVS[6] (b4) CTSDG[8] (b5) HAN[24] (b6) T-former[25] (b7) MMT[21] (b8) AOT[15] (b9) 本文方法

(b) FFHQ数据集

图5 不同方法在2个数据集上的定性比较结果

Fig. 5 Results of qualitative comparison of different methods on two datasets

从图5(a)可以看到，AOT方法通过堆叠聚合上下文转换(aggregated contextual transformations, AOT)模块捕获丰富的远距离图像上下文信息，由于没有足够的人脸先验知识，因此修复的人脸图像结构严重扭曲(图5(a7))；CTSDG方法结合边缘图信息以恢复图像结构，但其修复的人脸图像存在模糊现象，尤其表现在耳朵和眼镜部分(图5(a3))；HAN方法和T-former方法修复结果存在全局语义不合理问题，如左右眼不一致(图5(a4)和图5(a5))；MMT方法在眼镜框部分修复效果不明显，HINT方法生成的人脸图像中耳朵部分的修复与真实图像具有一定差异(图5(a6)和图5(a8))。

对于大面积缺失区域的人脸图像, PRVS、CTSDG、HAN、T-former 和 AOT 方法修复结果出现结构扭曲、模糊及伪影, 生成的人脸图像与原始图像存在较大差异。MMT 方法联合边缘纹理和语义分割作为辅助信息, 在修复大面积缺失时避免造成扭曲结构。生成的人脸图像整体结构较为合理, 视觉效果较好, 但细节方面存在不足(图 5(b7))。而本文方法在大面积不规则掩码下的修复效果均优于其他对比方法。由于本文方法利用边缘图和语义分割图为网络提供足够的人脸先验知识, 并引入对比学习技术来提高修复模型的训练效果, 从而更好地引导修复图像在表示空间中拉近真实图像, 因此生成的人脸图像结构合理, 纹理清晰, 修复效果更符合人类视觉感知(图 5(a9) 和图 5(b9))。

4.2　定量评价

为客观比较本文方法与对比方法的修复性能, 采用的评价指标分别为峰值信噪比(peak signal-to-noise ratio, PSNR)、结构相似性参数(structural similarity index, SSIM)、L_1 损失、学习感知图像块相似度(learned perceptual image patch similarity, LPIPS)和弗雷歇特距离(Fréchet inception distance, FID)。

表 1 展示了在 CelebA-HQ 数据集上, 本文方法与对比方法的修复性能比较结果。可知, 在掩码比例为 0.01～0.2 时, 本文方法在 PSNR、SSIM、FID 和 LPIPS 指标上的值分别为 38.56 dB、0.985、0.932 和 0.015, 相较于 HINT 方法, PSNR 提升了 1.1 dB, FID 降低了 0.338。在掩码比例为 0.2～0.4 时, 本文方法的 PSNR 和 FID 的值分别为 33.47 dB 和 2.071, 相较于 CTSDG 方法, PSNR 提升了 0.29 dB, FID 降低了 0.824。在掩码比例为 0.4～0.6 时, 本文方法的 PSNR 和 FID 的值分别为 26.82 dB 和 6.318, 相较于 MMT 方法, PSNR 提升了 0.34 dB, FID 降低了 1.266。可以得到, 本文方法在 CelebA-HQ 数据集上的修复性能均优于对比方法。

表 1　不同方法在 CelebA-HQ 数据集上的定量比较结果

Table 1　Quantitative comparison results of different methods on the CelebA-HQ dataset

方法	PSNR/dB↑			SSIM↑			L_1↓			FID↓			LPIPS↓		
	掩码比例 0.01～0.2	掩码比例 0.2～0.4	掩码比例 0.4～0.6	掩码比例 0.01～0.2	掩码比例 0.2～0.4	掩码比例 0.4～0.6	掩码比例 0.01～0.2	掩码比例 0.2～0.4	掩码比例 0.4～0.6	掩码比例 0.01～0.2	掩码比例 0.2～0.4	掩码比例 0.4～0.6	掩码比例 0.01～0.2	掩码比例 0.2～0.4	掩码比例 0.4～0.6
CTSDG[8]	38.00	33.18	26.43	0.984	0.947	0.846	**0.003**	**0.008**	**0.021**	1.258	2.895	10.475	0.018	0.059	0.161
HAN[24]	35.93	31.42	26.24	0.974	0.917	0.833	0.005	0.011	0.024	1.984	4.216	8.739	0.032	0.098	0.165
T-former[25]	36.14	31.44	26.18	0.974	0.915	0.830	0.005	0.011	0.024	1.572	4.630	9.576	0.025	0.098	0.166
MMT[21]	37.71	32.77	26.48	0.982	0.942	0.844	**0.003**	**0.008**	0.022	1.164	2.700	7.584	0.018	0.058	0.141
AOT[15]	35.07	31.59	24.59	0.963	0.921	0.792	0.009	0.014	0.033	2.427	4.642	19.654	0.035	0.085	0.223
HINT[26]	37.46	31.42	25.78	0.981	0.911	0.813	**0.003**	0.010	0.025	1.270	5.635	11.435	0.020	0.113	0.193
本文方法	**38.56**	**33.47**	**26.82**	**0.985**	**0.949**	**0.852**	**0.003**	**0.008**	**0.021**	**0.932**	**2.071**	**6.318**	**0.015**	**0.049**	**0.129**

注: "↑" 表示该评价指标的值越高, 图像修复效果越好; "↓" 表示该评价指标的值越低, 图像修复效果越好; 各项指标的最优值用粗体表示。

此外, 在 FFHQ 数据集上, 本文方法与对比方法的定量结果如表 2 所示。可知, 在掩码比例区间为 0.01～0.2 时, 本文方法的 PSNR、SSIM、FID 和 LPIPS 的值分别为 37.52 dB、0.981、0.473 和 0.018, 与 MMT 方法比较, PSNR 和 SSIM 分别提升了 1.19 dB 和 0.004, FID 和 LPIPS 分别降低了 0.35 和 0.007。当掩码比例区间为 0.2～0.4 时, 相较于 AOT 方法, PSNR 和 SSIM 分别提升了 1.47 dB 和 0.03, FID 和 LPIPS 分别降低了 2.374 和 0.031。当掩码比例区间为 0.4～0.6 时, 相较于 CTSDG 方法, PSNR 和 SSIM 分别提升了 0.20 dB 和 0.009, FID 和 LPIPS 分别降低了 4.786 和 0.033。综上分析, 本文方法在不同比例不规则掩码下的修复性能均优于其他 6 种图像修复对比方法。

表 3 展示了现有方法的参数量、浮点运算速度和平均推理时间的比较结果。可以看出, 本文模型的参数量、浮点运算速度、平均推理时间分别为 61×10^6、93×10^9/s、21 ms。由于本文模型在解码阶段由 3 个分支构成, 因此, 参数量相对较多。但与其他修复方法相比, 本文修复方法在浮点运算速度和平均推理时间上仍具有竞争力。

4.3　消融实验

4.3.1　各个模块的消融实验

为验证本文提出的 LKCCB 模块、NDAN-MHA 模块和 CR 网络在图像修复任务中的有效性, 分别在 CelebA-HQ 和 FFHQ 数据集上进行消融实验, 实验均在同等条件下进行, 实验结果如表 4 和图 6 所示。表 4 和图 6 中的实验 1 表示本文网络模型移除 LKCCB、NDAN-MHA 模块和 CR 网络, 之后逐

表2 不同方法在FFHQ数据集上的定量比较结果

Table 2 Quantitative comparison results of different methods on the FFHQ dataset

方法	PSNR/dB↑			SSIM↑			L_1↓			FID↓			LPIPS↓		
	掩码比例0.01～0.2	掩码比例0.2～0.4	掩码比例0.4～0.6	掩码比例0.01～0.2	掩码比例0.2～0.4	掩码比例0.4～0.6	掩码比例0.01～0.2	掩码比例0.2～0.4	掩码比例0.4～0.6	掩码比例0.01～0.2	掩码比例0.2～0.4	掩码比例0.4～0.6	掩码比例0.01～0.2	掩码比例0.2～0.4	掩码比例0.4～0.6
PRVS[6]	31.80	26.17	22.39	0.909	0.776	0.637	0.015	0.028	0.044	29.011	42.735	58.921	0.132	0.257	0.357
CTSDG[8]	37.36	31.70	25.12	**0.981**	0.937	0.822	**0.003**	**0.008**	0.024	0.613	2.125	10.116	0.020	0.069	0.182
HAN[24]	35.40	29.77	24.73	0.971	0.905	0.806	0.005	0.013	0.028	0.930	4.161	7.849	0.031	0.104	0.181
T-former[25]	35.98	30.29	25.07	0.974	0.912	0.815	0.005	0.012	0.026	0.546	2.359	5.793	0.023	0.086	0.163
MMT[21]	36.33	30.92	24.78	0.977	0.928	0.815	**0.003**	0.009	0.025	0.823	2.938	8.845	0.025	0.080	0.177
AOT[15]	34.07	30.33	23.57	0.952	0.908	0.770	0.011	0.016	0.037	1.498	3.884	25.273	0.040	0.090	0.244
本文方法	**37.52**	**31.80**	**25.32**	**0.981**	**0.938**	**0.831**	**0.003**	**0.008**	**0.023**	**0.473**	**1.510**	5.330	**0.018**	**0.059**	**0.149**

注：“↑”表示该评价指标的值越高，图像修复效果越好；“↓”表示该评价指标的值越低，图像修复效果越好；各项指标的最优值用粗体表示。

表3 不同方法效率比较结果

Table 3 Comparative results of the efficiency of different methods

方法	参数量	浮点运算速度/($10^9 \cdot s^{-1}$)	平均推理时间/ms
PRVS[6]	56×10^6	44	14
CTSDG[8]	52×10^6	92	25
HAN[24]	19×10^6	138	16
T-former[25]	15×10^6	51	22
MMT[21]	51×10^6	98	22
AOT[15]	15×10^6	73	20
HINT[26]	72×10^6	222	112
本文方法	61×10^6	93	21

步添加NDAN-MHA模块、LKCCB模块和CR网络进行消融实验，分别得到实验2、实验3、实验4和本文方法。其中，实验4表示CR网络中不将负样本(受损图像)添加到网络中训练，即仅使用正样本(原始图像)来训练网络，而本文方法同时将正、负样本添加到CR网络中进行训练。

表4为不同模块在CelebA-HQ和FFHQ数据集上进行消融实验的定量结果。在CelebA-HQ数据集上，实验1不添加任何模块的PSNR和FID分别为27.34 dB和6.026。在此基础上添加NDAN-MHA模块得到实验2，PSNR和FID都有较大的性能提升，PSNR提升了0.2 dB，FID降低了0.079。添加

表4 不同模块在2个数据集上的定量消融结果

Table 4 Quantitative ablation results for different modules on two datasets

序号	LKCCB模块	NDAN-MHA模块	CR模块	PSNR/dB↑	SSIM↑	L_1↓	FID↓	LPIPS↓
实验1	×	×	×	27.34 / 25.84	0.865 / 0.847	0.018 / 0.020	6.026 / 7.352	0.121 / 0.151
实验2	×	√	×	27.54 / 26.06	0.868 / 0.850	0.017 / 0.020	5.947 / 6.323	0.116 / 0.145
实验3	√	√	×	27.67 / 26.13	0.870 / 0.851	0.017 / 0.020	5.697 / 6.121	0.115 / 0.143
实验4	√	√	√(无负样本)	27.98 / 26.38	0.875 / 0.856	**0.016** / 0.019	4.902 / 4.276	0.105 / 0.128
本文方法	√	√	√	**28.09** / **26.59**	**0.878** / **0.860**	**0.016** / **0.018**	**4.575** / **3.657**	**0.102** / **0.121**

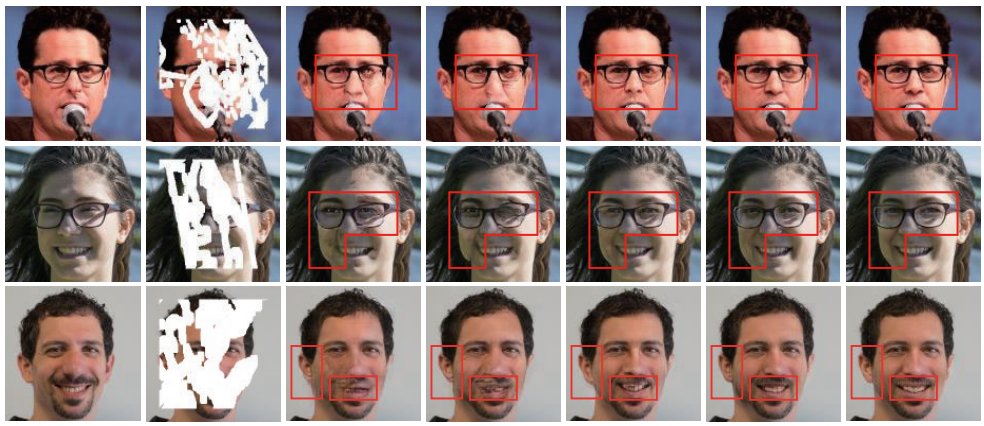
注：“√”表示使用该模块；“×”表示未使用该模块；各指标值第1个为CelebA-HQ数据集的结果；第2个为FFHQ数据集的结果。

LKCCB模块后得到实验3，PSNR提升了0.13 dB，FID降低了0.25。添加CR网络(无负样本)的PSNR和SSIM分别为27.98 dB和4.902，相较于实验3，PSNR提升了0.31 dB，FID降低了0.795。而本文方法在CR网络中添加正负样本进行训练，比实验4实现更优的PSNR和FID，分别为28.09 dB和4.575。与实验3相比，PSNR提升了0.42 dB，FID降低了1.122。这说明，在CR网络中添加正负样本进行训练的策略，模型的性能有较大的提升。

图6为不同模块在CelebA-HQ和FFHQ数据集上消融实验对比。实验1在不添加任何模块情况下，修复的人脸图像出现结构扭曲和伪影。实验2添加NDAN-MHA模块后的人脸图像有了局部改善。实验3在前者基础上添加LKCCB模块后的人脸图像纹理细节方面得到改善，如眼镜框。在此基础上添加CR网络，但不使用负样本进行训练，得到实验4，可以看到，整体修复效果均优于实验3。而本文方法的CR网络是同时使用正负样本进行训练，生成的人脸图像在语义结构和纹理细节方面明显优于实验4，如图6(g)所示。

4.3.2 LKCCB模块的消融实验

为验证LKCCB模块中核大小K对网络修复性



(a) 原图　(b) 残损图　(c) 实验 1　(d) 实验 2　(e) 实验 3　(f) 实验 4　(g) 本文方法

图 6　不同模块在 2 个数据集上的定性消融结果

Fig. 6　Qualitative ablation results of different modules on two datasets

能的影响, 选择 K 为 6、15、21、28 分别进行消融实验, 结果如表 5 所示。可知, 分解一个 21×21 的卷积要比分解一个 15×15 的卷积效果更好, 各评价指标均取得最优值。当分解一个 28×28的卷积时, 虽然 FID 和 LPIPS 分别上升了 0.023 和 0.0007, 但 PSNR 降低 0.01 dB。因此, 本文方法选择 $K=21$ 和 $d=3$。

表 5　LKCCB 中不同核大小 K 的消融实验

Table 5　Ablation experiments in LKCCB with different nucleus size K

K	d	PSNR/dB↑	SSIM↑	L_1↓	FID↓	LPIPS↓
6	2	28.08	0.877	0.016 2	4.635	0.102 8
15	3	28.05	0.877	0.016 3	4.625	0.102 6
21	3	28.09	0.878	0.016 2	4.575	0.102 3
28	4	28.08	0.878	0.016 2	4.598	0.103 0

为验证本文提出的 LKCCB 模块在修复任务中有较强的上下文推理能力, 将其与 ResNet 模块[27]和 AOT 模块[15]进行性能比较, 定量结果如表 6 所示。

表 6　CelebA-HQ 数据集上瓶颈块的定量结果

Table 6　Quantitative results on bottleneck blocks on the CelebA-HQ dataset

瓶颈块	PSNR/dB↑	SSIM↑	L_1↓	FID↓	LPIPS↓
RES@8	28.00	0.877	0.016 4	4.783	0.104 9
AOT@8	27.97	0.876	0.016 4	**4.538**	0.103 2
LKCCB@8 (-DLKCB)	28.06	0.877	0.016 3	4.702	0.103 7
LKCCB@2	28.00	0.876	0.016 4	4.654	0.103 1
LKCCB@4	28.06	0.877	0.016 3	4.576	0.102 6
LKCCB@6	28.07	0.877	**0.016 2**	4.556	**0.102 3**
LKCCB@8	**28.09**	**0.878**	**0.016 2**	4.575	**0.102 3**

注: LKCCB@L($L=2,4,6,8$)表示使用L个LKCCB模块进行实验; RES@8和AOT@8分别表示8个ResNet模块和8个AOT模块; LKCCB@8(−DLKCB)表示在8个LKCCB模块中去掉DLKCB进行实验。

从表 6 可以看出, 使用 8 个 ResNet 或 AOT 块时, 网络修复性能整体上并没有提高, 而使用 LKCCB 块时的修复性能整体较好。随着 LKCCB 块个数从 2 增加到 8, 网络修复性能逐渐提高, 可视化图如图 7 所示。而去掉 DLKCB 后的修复性能均下降, 因此, 本文选择堆叠 8 个 LKCCB 模块。

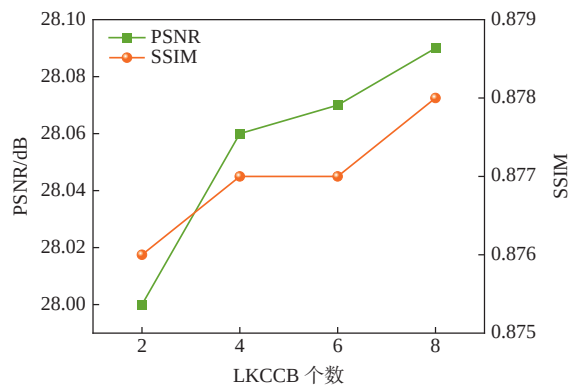


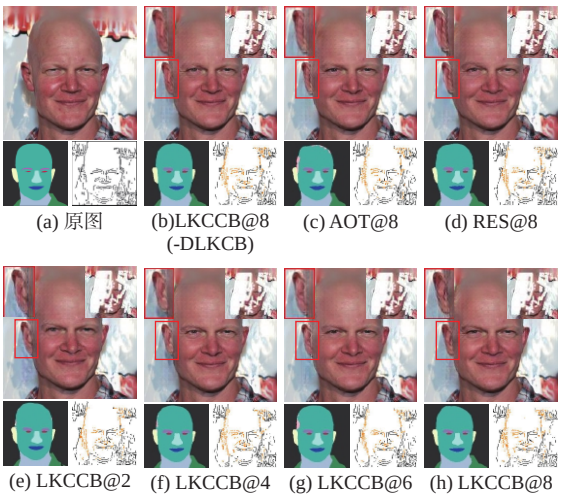
图 7　LKCCB 性能对比可视化

Fig. 7　Visualization of LKCCB performance comparison

图 8 展示了各瓶颈块的 RGB 图像修复结果及预测的语义分割图和边缘图。可以看到, 使用 AOT 模块修复的人脸图像中存在伪影, 如耳朵部分, 且预测的语义分割图不够准确。随着 LKCCB 模块个数的增加, 耳朵部分的修复得到改善, 且预测的语义分割图更加准确。

4.3.3　DAN 和 AFFN 模块的消融实验

NDAN-MHA 模块主要依赖 DAN 模块来学习 3 个分支特征的相互作用信息, 为验证本文提出的 DAN 和 AFFN 模块的有效性, 将 AFFN 与传统的前馈网络 FFN[19]进行比较。另外, 构造 3 个变体与 DAN 模块进行比较, 定量消融结果如表 7 所示, 对应的定性结果如图 9 所示。表 7 中序号 A 表示将本文提出的 AFFN 替换为传统的前馈网络, 序号 B 表示直接将来自 3 个分支的特征拼接在一起, 然后通过 2 个卷积层将其融合, 序号 C 和序号 D 分别表示将本文的 DAN(即序号 E)替换为 AdaIN[28] 和 SPADE[17],



(a) 原图 (b)LKCCB@8 (-DLKCB) (c) AOT@8 (d) RES@8

(e) LKCCB@2 (f) LKCCB@4 (g) LKCCB@6 (h) LKCCB@8

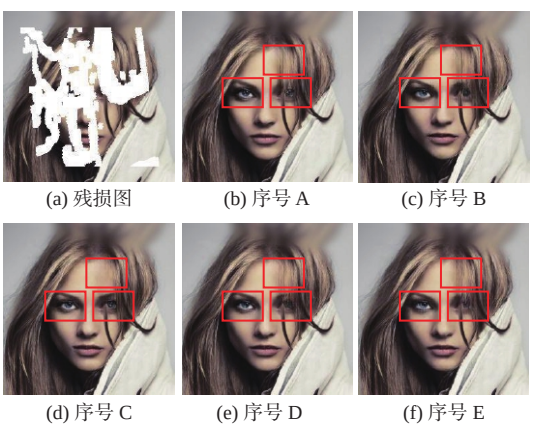
图 8 CelebA-HQ 数据集上瓶颈块的定性结果

Fig. 8 Qualitative results of bottleneck blocks on the CelebA-HQ dataset

表 7 NDAN-MHA 模块中不同组合消融实验

Table 7 Ablation experiments with different combinations in the NDAN-MHA module

序号	组合	PSNR/dB↑	SSIM↑	L_1↓
A	NDAN-MHA+FFN[19]	27.96	0.877	0.016 6
B	NDAN-MHA+拼接+Conv	27.98	0.876	0.016 4
C	NDAN-MHA+AdaIN[28]	27.92	0.874	0.016 6
D	NDAN-MHA+SPADE[17]	28.06	0.878	0.016 3
E	NDAN-MHA+DAN	28.09	0.878	0.016 2



(a) 残损图 (b) 序号 A (c) 序号 B

(d) 序号 C (e) 序号 D (f) 序号 E

图 9 NDAN-MHA 模块中不同组合定性结果

Fig. 9 Qualitative results of different combinations in the NDAN-MHA module

并基于 3 个分支的特征来学习调制参数($\boldsymbol{\alpha}$,$\boldsymbol{\beta}$)。

由表 7 可知，本文提出的 AFFN 模块在评价指标上均优于传统的前馈网络 FFN，相较于传统的前馈网络，PSNR 提升了 0.13 dB。从序号 B～D 可以看到，本文提出的 DAN 模块修复性能比 AdaIN 模块表现更好，PSNR 提升了 0.17 dB。相较于 SPADE 模块，PSNR 提升了 0.03 dB。此外，从图 9 可以看出，本文使用 DAN 模块融合多个分支特征时，保留了更多的语义信息，并提高了修复效果的细节表现。如图 9(f) 所示，本文方法在头发、眼睛部分修复更加真实。

4.3.4 不同损失函数的消融实验

为验证各种损失函数在图像修复任务中的有效性，本文在 CelebA-HQ 数据集上对 RGB 修复分支的损失函数进行消融实验，所有实验均在相同参数下进行训练，定量结果如表 8 所示，定性结果如图 10 所示。

表 8 损失函数定量消融结果

Table 8 Loss function quantitative ablation results

损失函数	FID↓	LPIPS↓
无损失函数	5.079	0.111
去除感知损失	4.874	0.106
去除风格+总变分损失	4.921	0.104
去除风格+重构损失	5.107	0.108
去除风格损失	4.834	0.102
本文方法	4.575	0.102

(a) 原图 (b) 残损图 (c) 无损失函数 (d) 去除感知损失

(e) 去除风格+总变分损失 (f) 去除风格+重构损失 (g) 去除风格损失 (h) 本文方法

图 10 损失函数定性消融结果

Fig. 10 Loss function qualitative ablation results

本文模型分别移除感知损失、风格损失和总变分损失、风格损失和重构损失及风格损失进行训练得到 4 个不同的消融模型。由表 8 可知，在本文模型上移除感知损失时，指标 FID 和 LPIPS 分别降低了 0.299 和 0.004。当在本文模型上移除风格+总变分损失时，指标 FID 降低了 0.346。在本文模型上移除风格+重构损失时，指标 FID 降低了 0.532。当移除风格损失时，指标 FID 降低了 0.259。本文方法在指标 FID 和 LPIPS 上均优于其他 4 种消融模型，表明本文使用重构损失、感知损失、风格损失和总变分损失进行联合训练的方式对图像修复效果有所提升。从图 10 可以看到，本文使用重构损失、感知损失、风格损失和总变分损失函数进行训练的修复效果均优于其他消融模型，其他消融模型

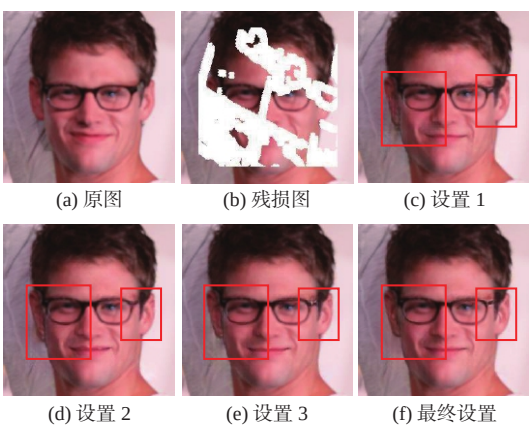
对眼镜框部分的修复效果不明显或结构扭曲，而本文方法能够合理地修复眼镜框部分。

针对边缘分支、语义分支、CR 网络中各超参数选取对修复效果的影响，本文在 CelebA-HQ 数据集上对边缘分支权重 λ_{edge}、语义分支权重 λ_{seg} 和 CR 网络权重 λ_{contra} 进行消融验证，消融定量结果如表 9 所示，相应的定性结果如图 11 所示。由表 9 可知，当 $\lambda_{edge}=0.5$、$\lambda_{seg}=0.5$ 和 $\lambda_{contra}=1$ 时，模型的修复性能取得最优值；此外，从图 11 可以看到，本文方法在眼镜框、耳朵处修复效果更加真实。

表 9　不同权重值对修复效果的定量结果

Table 9　Quantitative results of different weighting values on the effectiveness of restoration

权重设置	λ_{edge}	λ_{seg}	λ_{contra}	PSNR/dB↑	FID↓	LPIPS↓
设置1	1	1	1	27.84	5.187	0.109
设置2	0.5	0.5	0.5	27.85	5.216	0.109
设置3	0.5	0.5	2	28.03	4.741	0.104
最终设置	0.5	0.5	1	28.09	4.575	0.102



(a) 原图　(b) 残损图　(c) 设置 1

(d) 设置 2　(e) 设置 3　(f) 最终设置

图 11　不同权重值对修复效果的定性结果

Fig. 11　Qualitative results of different weighting values on the effectiveness of restoration

5　结　论

本文提出了一种端到端的用于图像修复的 Transformer 修复网络，其从共享的多种信息表示（即 RGB 图像、语义分割和边缘纹理）中丰富了耦合的空间特征。在网络编码阶段设计了一组 LKCCB 模块，通过增加有效感受野，有利于网络聚合更多的空间上下文信息，以增强上下文推理。在解码阶段提出一种 NDAN-MHA 模块，通过 DAN 模块集成来自 3 个分支的特征，并有效地捕获长期依赖关系。此外，本文提出的 CR 网络建立在对比学习的基础上，以确保恢复图像被拉近原始图像，并推离表示空间中的受损图像，从而使生成的修复图像更加真实。将本文方法与主流修复方法在 CelebA-HQ 和 FFHQ 数据集上进行定性和定量比较，实验结果表明，本文方法在大面积不规则掩码下具有更好的修复效果。然而，大面积遮挡下的修复效果与原图像仍有一定程度的差异。未来工作会将大核卷积和对比学习等技术引入图像修复任务，以进一步提高模型的修复性能。

参考文献（References）

[1] NAZERI K, NG E, JOSEPH T, et al. EdgeConnect: structure guided image inpainting using edge prediction[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision Workshop. Piscataway: IEEE Press, 2019: 3265-3274.

[2] YANG J, QI Z Q, SHI Y. Learning to incorporate structure knowledge for image inpainting[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2020, 34(7): 12605-12612.

[3] YAMASHITA Y, SHIMOSATO K, UKITA N. Boundary-aware image inpainting with multiple auxiliary cues[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE Press, 2022: 618-628.

[4] 谭骏珊, 李雅芳, 秦姣华. 基于推理注意力机制的二阶段网络图像修复[J]. 电讯技术, 2022, 62(11): 1545-1553.
TAN J S, LI Y F, QIN J H. Two-stage network image inpainting based on reasoning attention mechanism[J]. Telecommunication Engineering, 2022, 62(11): 1545-1553(in Chinese).

[5] SHAO X R, YE H L, YANG B, et al. Two-stream coupling network with bidirectional interaction between structure and texture for image inpainting[J]. Expert Systems with Applications, 2023, 231: 120700.

[6] LI J Y, HE F X, ZHANG L F, et al. Progressive reconstruction of visual structure for image inpainting[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2019: 5961-5970.

[7] XU S X, LIU D, XIONG Z W. E2I: Generative inpainting from edge to image[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2021, 31(4): 1308-1322.

[8] GUO X F, YANG H Y, HUANG D. Image inpainting via conditional texture and structure dual generation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2021: 14114-14123.

[9] 陈晓雷, 杨佳, 梁其铎. 结合语义先验和深度注意力残差的图像修复[J]. 计算机科学与探索, 2023, 17(10): 2450-2461.
CHEN X L, YANG J, LIANG Q D. Image inpainting combining semantic priors and deep attention residuals[J]. Journal of Frontiers of Computer Science and Technology, 2023, 17(10): 2450-2461 (in Chinese).

[10] HORITA D, YANG J L, CHEN D, et al. A structure-guided diffusion model for large-hole image completion[EB/OL]. (2023-09-06)[2024-05-06]. https://doi.org/10.48550/arXiv.2211.10437.

[11] NICHOL A Q, DHARIWAL P, RAMESH A, et al. GLIDE: Towards photorealistic image generation and editing with text-guided diffusion models[C]//Proceedings of the 39th International Conference on Machine Learning. New York: PMLR Press, 2022: 16784-

16804.

[12] MA X, ZHOU X Q, HUANG H B, et al. Free-form image inpainting via contrastive attention network[C]//Proceedings of the 25th International Conference on Pattern Recognition. Piscataway: IEEE Press, 2021: 9242-9249.

[13] ZUO Z W, ZHAO L, LI A L, et al. Generative image inpainting with segmentation confusion adversarial training and contrastive learning[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2023, 37(3): 3888-3896.

[14] LUO P J, XIAO G Q, GAO X B, et al. LKD-net: large kernel convolution network for single image dehazing[C]//Proceedings of the IEEE International Conference on Multimedia and Expo. Piscataway: IEEE Press, 2023: 1601-1606.

[15] ZENG Y, FU J, CHAO H, et al. Aggregated contextual transformations for high-resolution image inpainting[J]. IEEE Transactions on Visualization and Computer Graphics, 2023, 29(7): 3266-3280.

[16] SUVOROV R, LOGACHEVA E, MASHIKHIN A, et al. Resolution-robust large mask inpainting with fourier convolutions[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. Piscataway: IEEE Press, 2022: 3172-3182.

[17] PARK T, LIU M Y, WANG T C, et al. Semantic image synthesis with spatially-adaptive normalization[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2019: 2332-2341.

[18] MA N N, ZHANG X, LIU M Y, et al. Activate or not: learning customized activation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 8028-8038.

[19] MA X Z, KONG X, WANG S N, et al. Luna: linear unified nested attention[EB/OL]. (2021-11-02)[2024-05-21]. https://doi.org/10.48550/arXiv.2106.01540.

[20] LIU G L, REDA F A, SHIH K J, et al. Image inpainting for irregular holes using partial convolutions[C]//Proceedings of the 15th European Conference on Computer Vision. Berlin: Springer, 2018: 89-105.

[21] YU Y S, DU D W, ZHANG L B, et al. Unbiased multi-modality guidance for image inpainting[C]//Proceedings of the 17th European Conference on Computer Vision. Berlin: Springer Press, 2022: 668-684.

[22] KARRAS T, AILA T, LAINE S, et al. Progressive growing of GANs for improved quality, stability, and variation[EB/OL]. (2018-02-26)[2024-05-30]. https://doi.org/10.48550/arXiv.1710.10196.

[23] KARRAS T, LAINE S, AILA T. A style-based generator architecture for generative adversarial networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2019: 4401-4410. .

[24] DENG Y, HUI S Q, MENG R Y, et al. Hourglass attention network for image inpainting[C]//Proceedings of the 17th European Conference on Computer Vision. Berlin: Springer, 2022: 483-501.

[25] DENG Y, HUI S Q, ZHOU S P, et al. T-former: an efficient transformer for image inpainting[C]//Proceedings of the 30th ACM International Conference on Multimedia. New York: ACM Press, 2022: 6559-6568.

[26] CHEN S, ATAPOUR-ABARGHOUEI A, SHUM H P H. HINT: high-quality inpainting transformer with mask-aware encoding and enhanced attention[J]. IEEE Transactions on Multimedia, 2024, 26: 7649-7660.

[27] 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.

[28] HUANG X, BELONGIE S. Arbitrary style transfer in real-time with adaptive instance normalization[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 1510-1519.

Face image inpainting combining semantic segmentation and edge texture

SHI Jiliang[1, 2]，ZHANG Qian[2, 3, *]，ZHOU Zunfu[1, 2]，YANG Sihong[1, 2]

(1. School of Data Science and Information Engineering，Guizhou Minzu University，Guiyang 550025，China；
2. Key Laboratory of Pattern Recognition and Intelligent System of Guizhou，Guiyang 550025，China；
3. Academic Affairs Office，Guizhou Minzu University，Guiyang 550025，China)

Abstract：Current picture inpainting techniques use auxiliary structural information prediction to fill realistic patches, however erroneous priors can result in unrealistic structures and blurry textures. Meanwhile, existing methods only focus on the relationship between the original image and the inpainted image, and do not fully utilize the information of the damaged image. To address the above problems, an end-to-end transformer face image inpainting network is proposed, which utilizes semantic segmentation and edge texture information to guide the inpainting process. The main inpainting network includes one RGB inpainting branch and two auxiliary branches for semantic segmentation and edge texture. A set of large kernel convolutional context bottleneck (LKCCB) modules is designed in the encoder to increase the effective receptive field and better contextual reasoning. In order to capture distant contextual information, a nested dynamic auxiliary normalization multi-head attention (NDAN-MHA) module is proposed, which contains a dynamic auxiliary normalization (DAN) module that can dynamically integrate the structural features of the three branches to enrich semantic consistency. Furthermore, a contrastive regularization (CR) network is proposed to stabilize and improve the training of the network to generate more realistic inpainted images. The CelebA-HQ and FFHQ datasets were used for both qualitative and quantitative trials. The findings demonstrate that the suggested method performs better than the comparative methods in both subjective and objective measures and that it can reasonably restore huge, irregularly occluded face photos.

Keywords：face image inpainting；contrastive learning；large-kernel convolution；attention mechanism；dynamic auxiliary normalization

Received：2024-04-26；**Accepted**：2024-07-05；**Published Online**：2024-12-11 10：58

URL：link.cnki.net/urlid/11.2625.V.20241210.1723.003

Foundation item：School-level Scientific Research Projects of Guizhou Minzu University (GZMUZK［2021］YB23)

* **Corresponding author**. E-mail：gzmuzq@gzmu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0250

多尺度感知与红外特征增强的RGB-T人群计数方法

郑棣文[1]，石洋宇[1]，谢承杰[1]，卢树华[1,2,*]

(1. 中国人民公安大学 信息网络安全学院，北京 102600；2. 公安部安全防范技术与风险评估重点实验室，北京 102600)

摘 要：RGB-T人群计数旨在利用可见光与热图像的互补信息生成人群密度图，以应对低光照场景下人群计数任务。针对RGB-T人群计数在跨模态信息融合时，存在尺度变化、背景干扰等问题，提出一种基于多尺度感知与红外特征增强的RGB-T人群计数方法(MSENet)。该方法提出RGB-T特征融合机制(RTFM)，通过多分支结构实现多尺度特征提取，设计红外增强结构以充分捕捉热图像中的人群信息；利用密集连接与信息发散机制将互补特征传递到各个模态中，实现互补特征表达复用及模态特征增强。所提方法在RGBT-CC数据集和ShanghaiTechRGBD数据集上进行了对比实验。结果表明：所提方法优于现有的一些先进方法，具有较好的准确性、稳健性及良好的泛化性。

关 键 词：人群计数；红外与可见光图像；特征增强；多尺度感知；跨模态特征融合
中图分类号：TP391.4
文献标志码：A **文章编号**：1001-5965(2026)06-2208-11

人群计数旨在通过分析一幅图像或视频帧中的人群分布信息，并准确估算人群数量，在交通管理、城市规划、公共安全等领域具有广阔的应用前景[1]。近年来，随着计算机视觉和图像处理技术的不断进步，人群计数方法也得到显著改进与发展。然而，精确的人群计数仍面临诸多挑战，包括尺度变化、复杂背景、人群分布不均等[2]。

传统方法主要依赖于可见光RGB图像进行人群计数，但难以应对光照变化和严重遮挡等问题。为克服上述局限，研究人员开始探索将热图像、深度图像与RGB图像结合，提出融合可见光-热(RGB-T)图像的跨模态人群计数方法及结合可见光-深度(RGB-D)图像的跨模态人群计数方法。

现有的人群计数方法大多是利用可见光图像的光学信息进行计数，但由于可见光图像易受光照变化的影响，限制了基于可见光图像的人群计数方法在光照较低场景中的应用[3]。在低光照下，可见光图像中人群细节模糊，检测难度大，仅利用可见光图像作为人群计数任务的信息源并不完全可靠，需要挖掘更全面的信息。随着热成像仪被广泛推广，一些研究将热图像作为可见光图像的补充信息引入人群计数研究中。在低光照及相似背景干扰下，可见光图像中的人群并不清晰可见，而热图像中的人群轮廓则清晰可见。需要指出的是，部分热图像亦会突出其他热特征显著的物体(如发热壁、灯具)，易与人群信息混淆，但借助可见光图像可以有效消除。由此可见，可见光图像和热图像是高度互补的。近年来，RGB-T人群计数技术受到了广泛关注，并提出了相应的数据集和研究方法。Liu等[3]提出一种RGB-T人群计数数据集RGBT-CC，并设计了一种跨模态协同表示学习框架，具有信息聚合-分布模块(information aggregation-distribution module, IADM)，以充分捕捉不同模态的互补信息。由于不同模态信息之间存在巨大差异，文献[4-6]利用跨模态交叉注意力机制，有效地融合了不同模态之间的信息。但一些详细信息可能会

收稿日期：2024-04-24；**录用日期**：2024-09-13；**网络出版时间**：2024-09-19 15:39
网络出版地址：link.cnki.net/urlid/11.2625.V.20240919.1254.003
基金项目：中国人民公安大学双一流创新研究专项(2023SYL08)
*通信作者. E-mail：lushuhua@ppsuc.edu.cn

引用格式：郑棣文，石洋宇，谢承杰，等. 多尺度感知与红外特征增强的RGB-T人群计数方法[J]. 北京航空航天大学学报，2026，52(6)：2208-2218. ZHENG D W，SHI Y Y，XIE C J，et al. RGB-T crowd counting method with multi-scale perception and infrared feature enhancement[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(6)：2208-2218 (in Chinese).

在融合过程中丢失, 基于此, Tang 等[7] 提出一个三流自适应网络, 主流输入成对的 RGB 和热图像, 2 个辅助流分别输入 RGB 图像和热图像, 该方法能更好地提取图像特征。然而, 大多数方法集中在改进融合机制, 忽略了跨模态特征融合过程中的尺度变化问题, 文献 [8] 探索了 RGB-T 人群计数中尺度变化的负面影响, 将尺度感知的上下文学习引入跨模态特征融合过程中, 以解决 RGB-T 人群计数中的尺度变化问题。

综上所述, 由于模态间信息差异大, 在跨模态特征融合时仍存在特征融合不充分、尺度变化、背景干扰等问题。针对上述问题, 本文提出多尺度感知与红外特征增强的 RGB-T 人群计数方法 (multi-scale perception and infrared feature enhancement RGB-T crowd counting network, MSENet), 通过融合红外与可见光中高度互补特征, 以解决人群计数在光照较低场景下计数效果不佳的问题。其中, 设计多分支结构实现多尺度特征提取, 构建红外特征增强结构 (thermal enhanced structure, TES), 进一步抑制红外图像的背景干扰, 引入密集连接和信息发散机制有效整合红外与可见光图像的互补信息, 将互补信息传递到各个模态中, 实现各分支间模态信息共享, 增强各模态特征, 使特征融合更加充分, 提高计数准确性。

1　相关工作

随着深度学习的兴起, 人群计数任务在其推动下取得了较大进展[9-13]。基于深度学习的方法通过构建深层神经网络模型, 能够自动提取到深层特征, 并极大地提高了人群计数的准确性[14]。根据应用场景不同, 可将其分为可见光人群计数和跨模态人群计数, 跨模态人群计数包含 RGB-T 和 RGB-D 人群计数。

1.1　可见光人群计数

目前, 主流的人群计数方法将其视为一个回归任务, 通过输出人群密度图来得到总人数[15], 人群密度图是一个与输入图像大小相同的二维矩阵, 其中, 每个像素值表示该位置及其周围区域的人群密度, 通过对密度图的所有像素值求和, 可以估计图像中的总人数[16]。根据网络结构的不同, 可以将计数模型分为多列网络结构和单列网络结构[17]。多列卷积神经网络的人群计数方法通常采用多个并行的列来处理输入图像。每列通过不同大小的卷积核提取多尺度特征, 将其特征图融合在一起[1]。Zhang 等[18] 提出了多分支人群计数卷积神经网络 (crowd counting via multi-column convolutional neural network, MCNN), 利用多列卷积网络来解决尺度变化问题。Sam 等[19] 提出了一种基于切换卷积神经网络的人群计数方法, 通过自适应地选择合适的卷积核大小, 从而在不同密度的人群场景中获得更好的计数结果。多列卷积神经网络虽然可以提取不同尺度和语义的特征, 但结构复杂, 不同分支学习到的特征相似度高。为简化网络复杂度, 提高模型训练效率, 研究人员开发了更高效、更简单的单列网络。Li 等[20] 提出拥挤场景识别网络 (congested scene recognition network, CSRNet), 采用单分支的网络结构, 利用空洞卷积代替池化操作以扩大感受野, 在捕获多尺度特征的同时降低了网络模型的参数量。以上方法均依赖密度图作为监督信号进行模型训练, Ma 等提出了贝叶斯损失 (Bayesian loss, BL), 不再依赖从点标注生成密度图作为“伪标签”, 直接采用点标注作为监督信号, 通过构建基于点标注的密度贡献概率模型, 提高了人群计数的准确性, 基于提出的损失函数设计了人群计数方法[3,21]。

1.2　RGB-T 人群计数

如图 1(a) 所示, 在低光照条件下, 仅使用可见光图像无法有效地识别出人群并进行计数, 从而导致计数性能低下, 当使用红外图像时, 人群轮廓清晰可见, 计数性能得到明显提高, 同时使用可见光与红外图像时, 计数性能达到最优[3]。因此, 根据可见光图像和红外图像之间的互补性, 如何融合不同模态之间的信息成为关键问题。Liu 等[3] 设计了一种跨模态信息聚合发散模块, 以充分融合不同模态的互补信息, 引入 RGB-T 人群计数数据集 RGBT-CC。Zhang 等[4] 提出一种跨模态空间通道注意模块, 通过自适应通道特征聚合来优化具有空间注意力的跨模态特征。Peng 等[22] 提出一种基于无人机的 RGB-T 人群统计数据集 DroneRGBT, 并提出一种多模态人群计数网络, 具有自适应融合模块。上述方法主要集中于跨模态图像融合的研究, 考虑到信息融合过程中的尺度变化问题, Li 等[8] 设计了尺度感知通道注意聚合模块和尺度感知跨模态特征聚合模块, 用于特定模态特征提取和跨模态特征融合过程。

(a) 低光照　(b) 正常光照　(c) 热源干扰

图 1　不同光线下 RGB 图像与热图像

Fig. 1　RGB images and thermal images in different lighting conditions

2 本文方法

2.1 网络结构

本文提出 MSENet, 基于 BL 框架进行设计分为 3 个部分, 网络前端采用 VGG-16 前 10 层作为浅层特征提取器, 中部设计了 RGB-T 特征融合机制 (RGB-T feature fusion mechanism, RTFM), 后端密度图回归层利用 3 个卷积操作生成密度图与人群数量, 如图 2 所示。在图像预处理阶段, 将 RGB 与热图像由 BRG 格式转换为 RGB 格式, 读取标注文件中的人群位置点, 并过滤掉超出图像范围的点, 确保不同模态图像和标注点数据的一致性和准确性。预处理后的图像输入浅层特征提取器, 分别提取得到 RGB 特征图 $\boldsymbol{F}_{\mathrm{R}}$、红外特征图 $\boldsymbol{F}_{\mathrm{T}}$。将所提取的 $\boldsymbol{F}_{\mathrm{R}}$、$\boldsymbol{F}_{\mathrm{T}}$ 分别输入到 RTFM 中用于 RGB-T 特征融合, 通过 4 个尺度感知分支将 RGB 与红外特征送入跨模态特征融合中利用稠密连接与信息发散机制进行融合, 4 个分支融合后得到的 4 个特征图拼接成 1 个整体 $\boldsymbol{F}_{\mathrm{out}}$ 作为 RTFM 的输出。密度图回归层利用 3 个卷积操作, 将特征图 $\boldsymbol{F}_{\mathrm{out}}$ 转换为单通道的密度图, 密度图的每个像素值表示局部人群密度, 通过对密度图求和, 可以估算人群总人数。

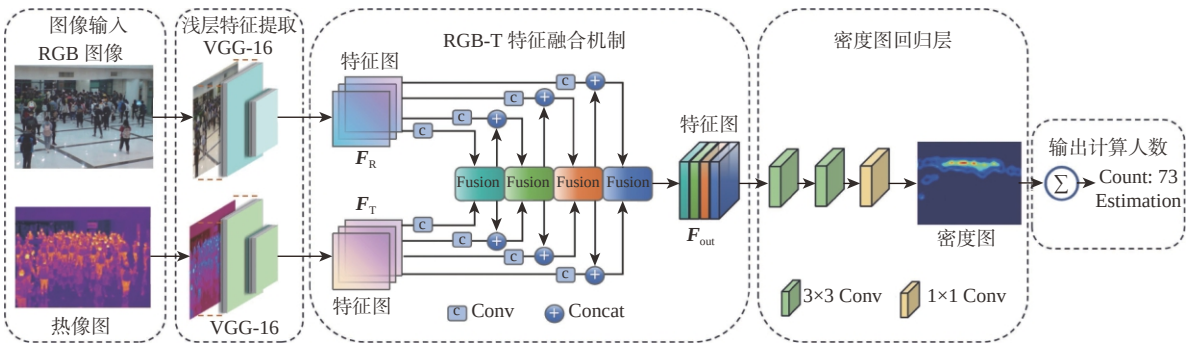


图 2 MSENet 结构

Fig. 2 Architecture of MSENet

2.2 RGB-T 特征融合机制

RTFM 结构如图 3 所示, 由 2 部分组成: 模态特征提取部分和跨模态特征融合部分。模态特征提取部分以可见光图像和红外图像作为输入, 设计多分支架构来应对尺度变化问题, 其中, 每个分支利用不同膨胀率的膨胀卷积提取对应感受野特征。跨模态特征融合部分利用密集连接与信息发散机制增强不同分支间模态特征的共享, 更好地将跨模态特征进行融合。最终将融合得到的 4 个特征图拼接在一起作为 RTFM 的输出。

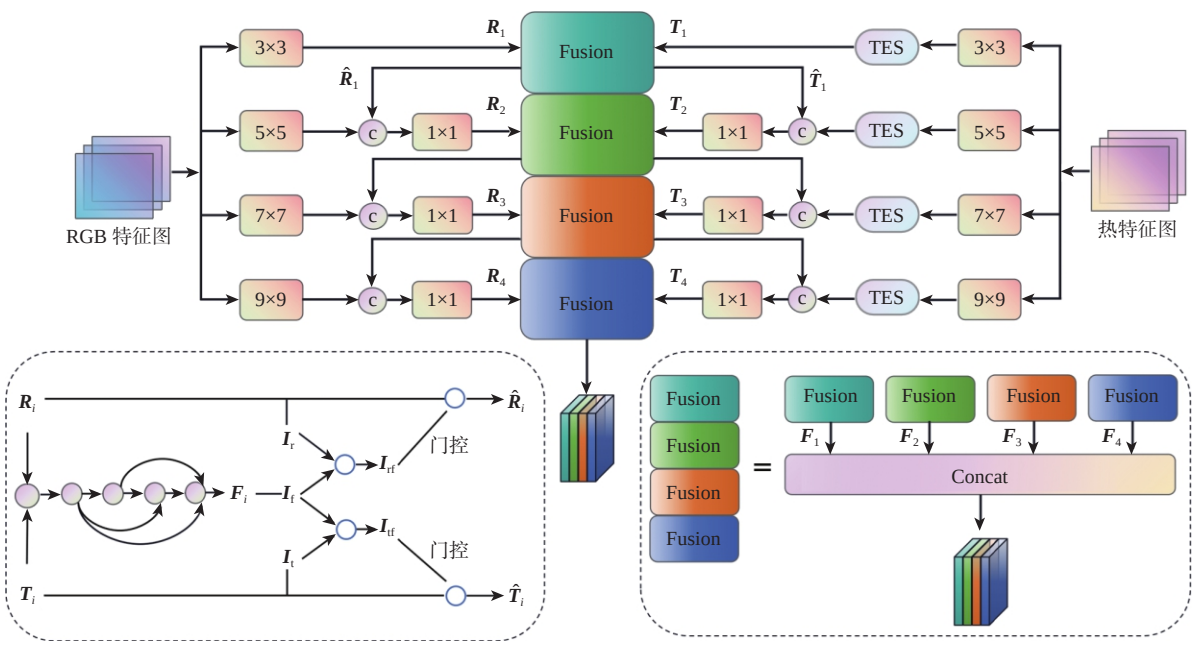


图 3 RTFM 结构

Fig. 3 Architecture of RTFM

如图 3 所示, RTFM 的左侧为 RGB 模态特征提取部分, 右侧为红外模态特征提取部分。令 $\boldsymbol{f}_i^{\mathrm{rgb}}$、$\boldsymbol{f}_i^{\mathrm{t}}$分别表示来自可见光和红外特征提取部分的第 $i(i$=1,2,3,4) 个分支经过特定卷积提取出的特征图, 可见光图像经过 4 个尺度卷积提取后, 再利用一个 1×1卷积滤波器将 $\boldsymbol{f}_i^{\mathrm{rgb}}$与来自相邻上分支的特征融合, 得到的 RGB 特征为$\{\boldsymbol{R}_i \mid i=1,2,3,4\}$。红外图像特征图经过 4 种尺度卷积提取后进入 TES 进行特征增强, 如图 3 所示, 每个 TES 都添加在卷积块之后、跨模态特征融合之前, 以抑制红外图像的背景干扰, 提高红外特征的兼容性, 再利用一个 1×1 卷积滤波器将来自相邻上分支的特征与增强后的特征整合, 得到的红外图像为$\{\boldsymbol{T}_i \mid i=1,2,3,4\}$, 跨模态融合过程可表示为

$$\boldsymbol{E}_i = \boldsymbol{R}_i + \boldsymbol{T}_i \qquad i=1,2,3,4 \tag{1}$$

$$\boldsymbol{T}_i = \boldsymbol{F}_{\mathrm{TES}}\left(\boldsymbol{f}_i^{\mathrm{t}}\right) \tag{2}$$

式中: $\boldsymbol{E}_i$ 为第 i 个分支首次融合后输出的特征; $\boldsymbol{R}_i$ 为第 i 个分支提取的 RGB 特征; $\boldsymbol{T}_i$ 为第 i 个分支提取的红外特征; $\boldsymbol{F}_{\mathrm{TES}}$ 为 TES 对应的特征变换函数, 用于对提取到的红外原始特征 $\boldsymbol{f}_i^{\mathrm{t}}$进行增强处理。

以3×3分支为例, 提取到特征图$\boldsymbol{R}_1$、$\boldsymbol{T}_1$后, 分别输入到跨模态特征融合部分中, 融合之后得到的图像为$\boldsymbol{F}_1$, 再通过信息发散机制将发散后的图像信息$\boldsymbol{I}_{\mathrm{rf}}$与$\boldsymbol{R}_1$、$\boldsymbol{I}_{\mathrm{tf}}$与$\boldsymbol{T}_1$进行拼接形成新的$\hat{\boldsymbol{R}}_1$与$\hat{\boldsymbol{T}}_1$。其余 3 个分支经过相同的操作后得到的融合特征图分别为$\boldsymbol{F}_2$、$\boldsymbol{F}_3$、$\boldsymbol{F}_4$。最终将 4 个融合后的特征图$\boldsymbol{F}_1$、$\boldsymbol{F}_2$、$\boldsymbol{F}_3$、$\boldsymbol{F}_4$进行拼接得到 $\boldsymbol{F}_{\mathrm{out}}$作为 RTFM 的输出。

为更好地提取可见光与红外图像的互补特征并进行融合, 基于稠密连接网络与信息发散机制设计了 Fusion 部分。如图 4 所示, 前半部分是一个密集连接结构, 包含 4 个密集连接层, 每层都连接到所有后续层, 以促进特征复用, 加强特征表达和传播, 提高参数效率。后半部分利用信息发散机制对每个模态的特征进行增强。经过模态特征提取部分提取的红外与可见光图像为$\boldsymbol{R}_i$与$\boldsymbol{T}_i$, 经过融合后得到的特征为$\boldsymbol{E}_i$。

获得融合特征$\boldsymbol{E}_i$后, 通过卷积核为3×3、通道数为 64 的卷积层得到$\boldsymbol{E}_{1,i}$, 通过卷积核为3×3、通道数为 16 的卷积层得到$\boldsymbol{E}_{2,i}$, 将$\boldsymbol{E}_{1,i}$与$\boldsymbol{E}_{2,i}$相加后通过3×3卷积得到$\boldsymbol{E}_{3,i}$。将$\boldsymbol{E}_{1,i}$、$\boldsymbol{E}_{2,i}$与$\boldsymbol{E}_{3,i}$提供的信息相加, 通过3×3卷积后得到$\boldsymbol{E}_{4,i}$(即输出信息$\boldsymbol{F}_i$)。上述过程可表示为

$$\boldsymbol{E}_{1,i} = \mathrm{BN}(\mathrm{ReLU}(\mathrm{Conv}_{3\times3}(\boldsymbol{E}_i))) \tag{3}$$

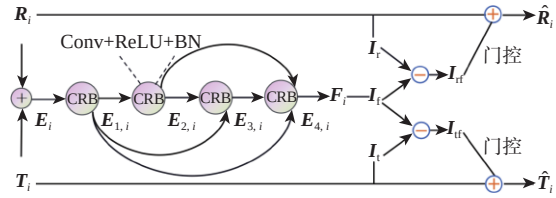


图 4　跨模态特征融合

Fig. 4　Cross-modal feature fusion

$$\boldsymbol{E}_{2,i} = \mathrm{BN}\left(\mathrm{ReLU}(\mathrm{Conv}_{3\times3}(\boldsymbol{E}_{1,i}))\right) \tag{4}$$

$$\boldsymbol{E}_{3,i} = \mathrm{BN}\left(\mathrm{ReLU}(\mathrm{Conv}_{3\times3}(\boldsymbol{E}_{1,i}+\boldsymbol{E}_{2,i}))\right) \tag{5}$$

$$\boldsymbol{E}_{4,i} = \mathrm{BN}\left(\mathrm{ReLU}(\mathrm{Conv}_{3\times3}(\boldsymbol{E}_{1,i}+\boldsymbol{E}_{2,i}+\boldsymbol{E}_{3,i}))\right) \tag{6}$$

通过融合后的$\boldsymbol{F}_i$包含 2 个模态的信息, 首先, 提取其上下文信息得到$\boldsymbol{I}_{\mathrm{f}}$, 通过$\boldsymbol{I}_{\mathrm{f}}$与各个模态的信息$\boldsymbol{I}_{\mathrm{r}}$、$\boldsymbol{I}_{\mathrm{t}}$的差值得到残差信息$\boldsymbol{I}_{\mathrm{rf}}$与$\boldsymbol{I}_{\mathrm{tf}}$, 然后, 通过 2 个门函数自适应传播互补信息以增强$\boldsymbol{R}_i$与$\boldsymbol{T}_i$, 增强后得到信息为$\hat{\boldsymbol{R}}_i$与$\hat{\boldsymbol{T}}_i$。通过信息发散机制能够充分捕捉可见光与红外图像的互补性, 以增强各模态间的信息共享, 该部分计算过程为

$$\begin{cases} \boldsymbol{I}_{\mathrm{rf}} = \boldsymbol{I}_{\mathrm{f}} - \boldsymbol{I}_{\mathrm{r}}, \cdots, \boldsymbol{I}_{\mathrm{tf}} = \boldsymbol{I}_{\mathrm{f}} - \boldsymbol{I}_{\mathrm{t}} \\ \omega_{\mathrm{rf}} = \mathrm{Conv}_{1\times1}(\boldsymbol{I}_{\mathrm{rf}}) \\ \omega_{\mathrm{tf}} = \mathrm{Conv}_{1\times1}(\boldsymbol{I}_{\mathrm{tf}}) \\ \hat{\boldsymbol{R}}_i = \boldsymbol{R}_i + \boldsymbol{I}_{\mathrm{rf}} \odot \omega_{\mathrm{rf}}, \cdots, \hat{\boldsymbol{T}}_i = \boldsymbol{T}_i + \boldsymbol{I}_{\mathrm{tf}} \odot \omega_{\mathrm{tf}} \end{cases} \tag{7}$$

2.3　红外特征增强结构

为抑制红外图像中的背景干扰问题, 充分捕捉红外图像中的人群信息, 引入 TES, 如图 5 所示, TES 包含 2 个顺序注意操作: 通道注意操作和空间注意操作[23]。TES 与卷积块注意力模块(convolutional block attention module, CBAM)相比, 在通道和空间注意力中只利用单个全局池化操作来挖掘红外图像中最关键的线索, 降低了模块的复杂性。

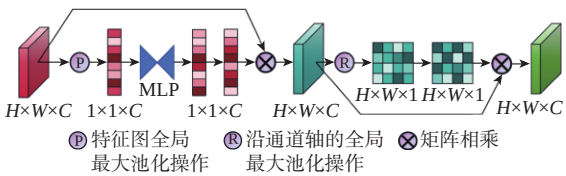


图 5　红外特征增强结构

Fig. 5　Architecture of thermal enhanced structure

经过卷积操作后的红外特征图$\boldsymbol{f}_i^{\mathrm{t}}$依次进行通道注意操作与空间注意操作。通道注意操作利用通道间的关系来选择性地增强红外图像中的关键特征。具体而言, 通道注意力通过全局最大池化操作获取每个通道的全局特征向量, 该特征向量经过 2 层感知器处理后生成通道注意力权重, 将这些权重再重新分配给原始特征图的每个通道, 实现通道特征的自适应增强。空间注意操作通过确定红外

信息携带的人群位置线索来增强特征图中的重要区域。空间注意力机制通过全局最大池化操作生成一个空间特征图，该空间特征图经过卷积层融合后生成空间注意力权重图，将这些权重重新分配给原始特征图的每个空间位置，实现空间特征的自适应增强。通过上述通道注意力和空间注意力操作，TES能够有效捕捉红外图像中的人群信息，并抑制背景干扰，从而增强RGB-T跨模态人群计数的性能。TES的运算过程定义为

$$\boldsymbol{F}_{\mathrm{TES}}(\boldsymbol{f}_i^{\mathrm{t}}) = S_{\mathrm{att}}(C_{\mathrm{att}}(\boldsymbol{f}_i^{\mathrm{t}})) \tag{8}$$

$$C_{\mathrm{att}}(\boldsymbol{f}_i^{\mathrm{t}}) = M(P_{\max}(\boldsymbol{f}_i^{\mathrm{t}})) \otimes \boldsymbol{f}_i^{\mathrm{t}} \tag{9}$$

$$S_{\mathrm{att}}(\boldsymbol{f}_i^{\mathrm{t}}) = \mathrm{Conv}(R_{\max}(\boldsymbol{f}_i^{\mathrm{t}})) \odot \boldsymbol{f}_i^{\mathrm{t}} \tag{10}$$

式中：C_{att}和S_{att}分别为通道和空间注意力；$P_{\max}$为每个特征图的全局最大池化操作；M为2层感知器；$R_{\max}$为特征图中每个点沿通道轴的全局最大池化操作。

2.4 损失函数

在现有的人群计数数据集中，通常标记每个人头部的中心点进行注释。大多数人群计数方法都是基于密度图估计，即通过高斯核将稀疏点注释转换成“地面真值”密度图，并以其作为监督信号，与模型预测得到的密度图一起用来计算相应的损失，进行密度回归估计。然而，由于目标之间高度的拥挤与重叠，通过该方法生成的“地面真值”密度图仅仅是对实际概率密度图的粗略逼近，其准确度很难保证。Ma等[21]提出了BL，不再依赖从点标注生成密度图作为伪标签，直接采用点标注作为监督信号，利用“离散”的概率密度期望值与“离散”的点标注设计损失函数并进行期望值意义上的回归估计。如式(12)所示，期望该点有人的概率是1，$E[c_n]$为每一个目标的期望计数。考虑到BL相较于其他传统方法在计数性能方面的优势，在设计网络结构时利用BL作为损失函数：

$$\mathcal{L}^{\mathrm{Bayes}} = \sum_{n=1}^{N} \mathcal{F}(1 - E[c_n]) \tag{11}$$

式中：$\mathcal{L}^{\mathrm{Bayes}}$为BL函数；$\mathcal{F}(\cdot)$为误差计算函数；$N$为输入图像中真实标注的人头目标总数量。

3 实验与结果分析

3.1 数据集及实验设置

3.1.1 RGBT-CC数据集

由Liu等[3]在CVPR2021会议发布的RGBT-CC数据集是一个RGB-T人群计数基准，包含2 030对RGB图像和热图像，每对图像分辨率均为640像素×480像素，每对图像还配有一个json格式的标注文件，该文件包含图像中行人的位置标注，记录了每张图像中所有行人的位置点，以像素坐标形式表示，标注精确到每个人头的中心点。数据集中1 013对为光线明亮环境下拍摄的图像对，其余1 017对为黑暗环境中拍摄的图像对，光线明亮环境指的是在自然光或良好人工照明条件下拍摄的图像，而光线黑暗环境则指在低光照或夜间条件下拍摄的图像[3]。数据集进一步划分为训练集1 030对、验证集200对和测试集800对。这些图像捕获于各种无约束场景，图像中的人群密度也从稀疏到密集不等，具有一定的挑战性。

3.1.2 ShanghaiTechRGBD数据集

由Lian等[24]在2019年发布的ShanghaiTech-RGBD数据集用于评估RGB-D图像在人群计数中的性能，包含2 193张图像，共标注有144 512个头部。这些图像由立体相机ZED3捕捉，有效深度为0～20 m，图像分辨率为1 080像素×1 920像素。数据集覆盖各种场景，包括繁忙的城市街道和拥挤的公园。图像中每个头部用一个点进行标注，在测试集中还提供了头部的边界框标注，以便评估头部检测性能。数据集划分为训练集1 193张图像和测试集1 000张图像，提供了丰富的多样性和挑战性，用于训练和评估人群计数模型。

3.1.3 实验设置参数

实验在Ubuntu22.04系统下进行，实验过程利用Pytorch1.10.0深度学习框架，编程语言为Python3.8，实验使用Nvidia RTX 3080 Laptop GPU进行训练，CUDA版本为12.0，显存为16 GB，电脑内存为32 GB。学习率设为固定值0.000 1，训练迭代次数为1 000，权重衰减设置为0.000 1，批次大小设置为1。权重衰减用于防止过拟合，通过在损失函数中加入权重的L2正则化项，使模型权重不至于过大，从而提高模型的泛化能力。批次大小设置为1是因为在处理高分辨率图像和复杂模型时，较小的批量大小有助于降低内存消耗。

3.2 评价指标

平均绝对误差(mean absolute error, MAE)和均方根误差(root mean square error, RMSE)是人群计数算法常用的评价指标，MAE通常用来表示模型人群数量预测值与真实值的平均绝对误差；RMSE通常用来表示人群计数模型估计值的离散程度。但MAE只能评估整张图像的全局误差，无法对图像的局部区域进行有效评估，因此，采用网格平均绝对误差(grid average mean absolute error, GAME[25])来评估不同区域的性能。具体来说，对于给定的等

级 l, 将图像划分为不重叠的 4^l 个区域, 并计算每个区域的计数误差。当 l=0 时, GAME(0) 等于 MAE, 当 l=1 时, GAME(1) 表示将图像划分为 2 像素×2 像素网格(即 4 个区域)后的人群计数误差。GAME 和 RMSE 的值越小, 表示模型性能越好, 计算方法为

$$\mu_{\mathrm{GAME}}(l)=\frac{1}{M}\sum_{i=1}^{N}\sum_{j=1}^{4^l}\left|C_i^j-G_i^j\right| \tag{12}$$

$$\mu_{\mathrm{RMSE}}=\sqrt{\frac{1}{M}\sum_{i=1}^{N}|C_i-G_i|^2} \tag{13}$$

式中: μ_{GAME}、μ_{RMSE}分别为网格平均绝对误差、均方根误差; M 为测试集图片的数量; 4^l为图像划分的区域数量; C_i和G_i分别为第 i 张图像的估计值和真值; C_i^j和G_i^j分别为第 i 张图像第 j 个区域的估计值和真值。

3.3　结果分析

3.3.1　对比实验

为验证本文方法的先进性, 在 RGBT-CC、ShanghaiTechRGBD 数据集上进行训练和测试, 并与当前诸多方法进行对比。在 RGBT-CC 数据集上的测试结果如表 1 所示, 与其他方法相比, MSENet 在所有评价指标上都获得了较好的结果, 本文方法的准确度和鲁棒性超越大部分方法。通过对比以 BL、CSRNet 为框架的 IADM 和 CSA-Net 等计数方法可以看出, 在同样的计数方法下, 基于 BL 设计的方法性能高于 CSRNet, 因此, 本文方法基于 BL 进行设计。选取同样以 BL 为框架的 IADM、CSCA、MAT 和 CSA-Net 与本文方法进行对比, 其中, IADM 和 CSCA 是较早利用跨模态进行特征融合的 RGB-T 人群计数方法, 设计了即插即用的跨模态特征融合模块, 能够较好地融合跨模态特征, MSENet 明显优于 IADM 和 CSCA, 可能是由于引入多分支可见红外特征融合结构能够更有效地捕捉和融合跨模态特征信息, 提升了多尺度感知能力, 使模型在处理密集人群和稀疏人群时都能保持较高的精度。对比结果表明, 将尺度感知引入 RGB-T 人群计数中有利于提高计数性能。选取的 CNCTrans 与 MAT 引入了 Transformer 进行跨模态特征融合, 利用跨模态注意力机制学习跨模态互补特征, 相较而言, MSENet 未利用 Transformer 进行跨模态特征融合, 结果优于 CNCTrans 和 MAT, 展示了一定的先进性和有效性。此外, 选取了同样考虑尺度变化的 CSA-Net 进行对比, MSENet 优于 CSANet, 可能因为本文方法不仅考虑到尺度变化问题, 还在特征提取过程中进一步考虑到了各模态特征的独特性, 对模态信息进行增强。通过多分支融合结构, MSENet 能够更好地提取和融合不同模态的互补特征, 从而实现更精确的人群计数。值得一提的是, 上述对比方法都是双流输入的网络结构, 本文方法与三流输入的网络结构 TAFNet 进行对比, 亦具有较好的计数性能, 展现了较强的信息处理能力。

表 1　RGBT-CC 数据集结果对比

Table 1　Comparison results on the RGBT-CC dataset

方法	框架	GAME(0)	GAME(1)	GAME(2)	GAME(3)	RMSE
IADM+BL[3]	BL	15.61	19.95	24.69	32.89	28.18
CSCA+BL[4]	BL	14.32	18.91	23.81	32.47	26.01
MAT[6]	BL	12.35	16.29	20.81	29.09	22.53
CSA-Net[8]	BL	12.45	16.46	21.48	30.62	21.64
IADM+CSRNet[3]	CSRNet	17.94	21.44	26.17	33.33	30.91
CSA-Net[8]	CSRNet	15.77	19.40	24.14	30.14	29.17
CNCTrans[5]	CNN+Transformer	13.96	17.98	23.03	31.15	24.55
TAFNet[7]	VGG16	12.38	16.98	21.86	30.19	22.45
MSENet(本文)	BL	12.21	16.32	20.69	28.94	21.59

为测试 MSENet 在不同光照条件下的性能, 将测试集分为明亮环境和黑暗环境, 并在 2 个环境中分别进行了实验。选取基于不同框架中性能较好、有代表性的方法进行对比, 结果如表 2 所示。本文方法在明亮和黑暗环境中都有稳定优异的计数表现。在明亮环境中, MSENet 在所有评价指标上均优于其他 3 种方法, 表明其在高光照条件下具有良好的特征提取和融合能力。在黑暗环境中, MSENet 也表现出色, 可能得益于其有效的跨模态融合机制, 能够充分利用 RGB 与红外图像的互补信息在低光照条件下准确计数。

将本文 RGB-T 人群计数方法 MSENet 扩展到相关的跨模态人群计数任务 RGB-D 人群计数中, 在 ShanghaiTechRGBD 数据集上的计数结果

如表3所示。本文方法在RGB-D数据集上的表现显著优于现有的多种RGB-D跨模态人群计数方法，如UC-Net、HDFNet和BBS-Net。上述方法虽然在RGB-D数据集中表现良好，但MSENet在计数精度上依然占据优势。与以BL为框架的RGB-T方法IADM和CSCA相比，本文方法有更好的计数性能，虽与CSA-Net相比结果略有降低，但本文方法仍具有一定的竞争力。MSENet不仅在RGB-T数据集中表现优异，扩展到RGB-D人群计数任务中也展示出其性能优势。上述结果表明，本文方法在不同模态的人群计数任务中具有应用价值。

表2 RGBT-CC数据集上不同光照条件下的实验结果

Table 2 The results under different illumination conditions on RGBT-CC dataset

光照	方法	GAME(0)	GAME(1)	GAME(2)	GAME(3)	RMSE
明亮	IADM+CSRNet[3]	20.36	23.57	28.49	36.29	32.57
	TAFNet[7]	15.57	20.65	26.67	36.17	24.25
	CNCTrans[5]	15.05	19.04	24.21	32.91	25.00
	MSENet(本文)	13.81	17.89	23.26	31.12	24.84
黑暗	IADM+CSRNet[3]	15.44	19.23	23.79	30.28	29.11
	TAFNet[7]	14.20	19.20	24.00	31.63	27.50
	CNCTrans[5]	13.34	17.38	21.73	29.16	24.70
	MSENet(本文)	11.62	16.19	20.06	27.63	21.27

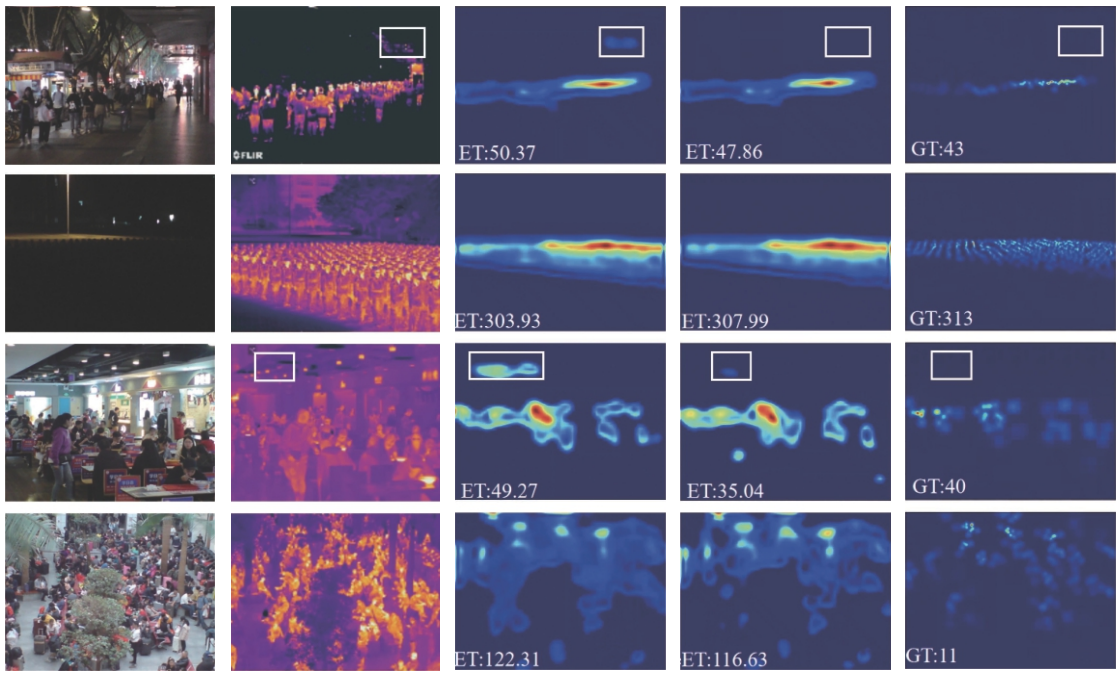
表3 ShanghaiTechRGBD数据集结果对比

Table 3 Comparison results on the ShanghaiTechRGBD dataset

方法	GAME(0)	GAME(1)	GAME(2)	GAME(3)	RMSE
UC-Net[26]	10.81	15.24	22.04	32.98	15.70
HDFNet[27]	8.32	13.93	17.97	22.62	13.01
BBS-Net[28]	6.26	8.53	11.80	16.46	9.26
IADM+BL[3]	7.13	9.28	13.00	19.53	10.27
CSCA+BL[4]	5.68	7.70	10.45	15.88	8.66
CSA-Net[8]	4.57	5.82	8.02	12.47	6.83
MSENet(本文)	4.75	6.29	8.96	12.04	7.55

由IADM+BL、CSCA+BL、CSA-Net与本文方法在RGBT-CC和ShanghaiTechRGBD数据集的结果可以看出，以上方法在ShanghaiTechRGBD数据集上的计数误差均明显小于在RGBT-CC数据集上的误差，主要是因为2个数据集场景复杂性的不同和模态信息的差异。在场景复杂性方面：ShanghaiTechRGBD数据集场景的复杂性主要体现在场景的人群密度上，数据集中图像的人群密度从稀疏到密集不等，而RGBT-CC数据集场景的复杂性不仅体现在人群密度上，同样体现在不同光照条件下的复杂环境中，如白天、夜晚、强光、弱光等，模型需要面对光照变化带来的图像质量变化。在模态信息差异方面：ShanghaiTechRGBD数据集提供的深度信息直接反映了场景中物体与摄像头之间的距离，这种直接的空间信息使模型能够更准确地区分前景（如人群）和背景及在三维空间中定位个体，对于人群计数较为有利。而RGBT-CC数据集中的热成像信息虽然在低光照条件下具有优势，但其分辨率通常较低，且缺乏细节，限制了模型在某些复杂场景中的判别能力。因此，可以推断，RGBT-CC和ShanghaiTechRGBD数据集的评价参数量级变化主要是模态信息差异和场景复杂性所导致的。

如图6所示，选取4组图像进行对比，第1行和第2行是在黑暗环境下生成的密度，第3行和第4行是在明亮环境下生成的密度。其中，图6(a)为RGB图像，图6(b)为热图像，图6(c)为CSRNet+IADM计数网络输出的密度图及计数结果，图6(d)为本文方法结果，密度图中ET表示预测值，GT表示真实值。可以看出，无论是在黑暗环境还是明亮环境下，MSENet的结果都比CSRNet+IADM更好，说明本文方法的性能在明亮和黑暗环境下都更稳定。由第1行和第3行方框标注处可以看出，本文方法相较于CSRNet+IADM能够消除一些背景干扰，提升计数性能。



(a) RGB 图像　(b) 热图像　(c) CSRNet+IADM　(d) MSENet(本文)　(e) Ground-truth

图 6　不同光照条件下生成的人群密度图

Fig. 6　Crowd density maps generated in different illumination conditions

3.3.2　消融实验

为验证 MSENet 各个组件的有效性和整体结构的合理性, 按照从整体结构到核心模块的详细设置方式, 对本文方法在 RGBT-CC 数据集上进行消融实验。考虑到 BL 相对于 CSRNet 在 RGB-T 人群计数方面的性能优势, 在本部分实验中, 将 BL 作为基础框架。

在相同的实验配置中对 MSENet 的各个部分进行消融实验, 结果如表 4 所示, 其中, Baseline 表示基础网络架构 (VGG+4 分支结构), 该基线模型的 GAME(0) 和 RMSE 分别为 15.43 和 25.22。在其他条件不变的情况下, 逐步增加 Fusion 模块、TES 结构依次进行测试。为直观展示出 MSENet 各部分的有效性, 将消融实验进行可视化, 结果如图 7 所示, 随着各部分的增加, 密度图效果也逐渐提升, 计数误差逐步下降。在 RTFM 中加入 Fusion 模块进行图像融合后, 对比图 7(c) 和图 7(d) 可以看出, 模型计数效果得到提升。如表 4 所示, 加入 Fusion 模块后, 模型的 GAME(0) 和 RMSE 分别为 12.79 和 22.05, 对比基线模型, GAME(0) 提高了 2.64。由第 2 行方框标注处可以看出, 加入 TES 进行红外增强后能有效抑制红外背景的干扰, 并且从第 1 行方框标注处可以看出, 能够更好地融合红外信息, 使模型的计数误差更小。如表 4 所示, 加入 TES 后模型的 GAME(0) 和 RMSE 分别降到 12.21 和 21.59, 准确性和稳定性方面有显著的提升。

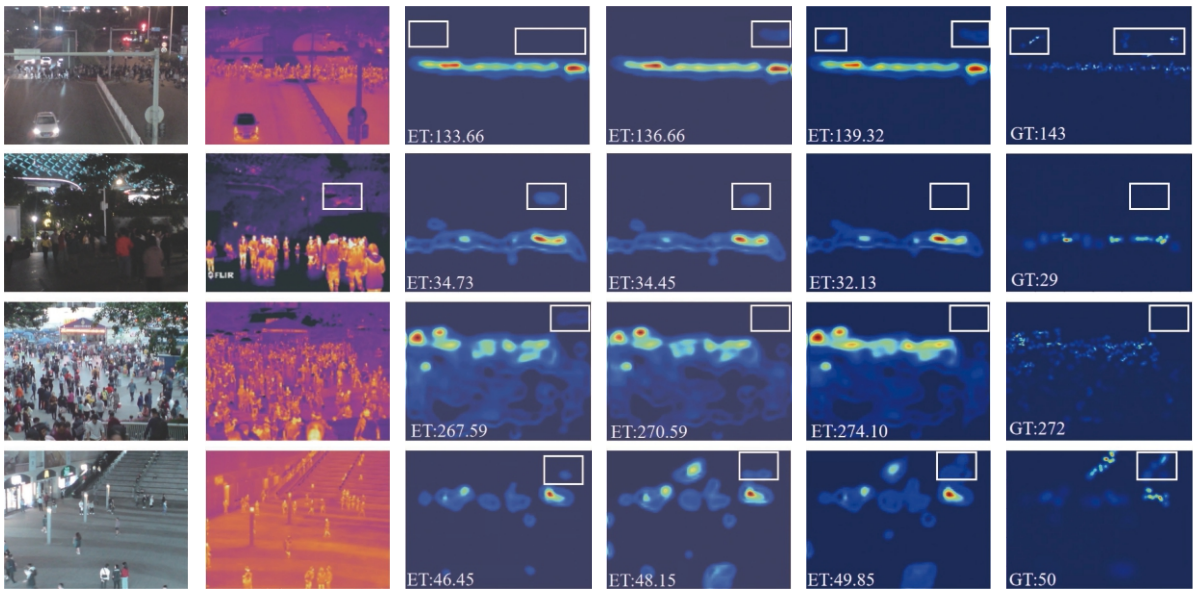
表 4　消融实验结果对比

Table 4　Comparison results of ablation experiments

方法	GAME(0)	GAME(1)	GAME(2)	GAME(3)	RMSE
Baseline	15.43	18.61	24.17	31.13	25.22
Baseline+Fusion	12.79	16.94	22.02	29.70	22.05
Baseline+Fusion+TES	12.21	16.32	20.69	28.94	21.59

3.3.3　多分支结构设计

在 RTFM 中的模态特征提取部分以实现多尺度特征表示为目标, 包含 4 个尺度分支。为探究多尺度设计对于人群计数的必要性, 设计一组探究尺度分支数量的实验。如表 5 所示, 为确定模态提取部分的分支数量, 在 RGBT-CC 数据集上进行了 6 个实验, 可以看出, 4 个尺度分支的设计使 MSENet 的性能优于 1、2、3、5、6 个尺度分支的所有设计, 更适合 RGB-T 人群计数。此外, 在可见光和红外分支中分别使用膨胀率为 1、2、3 的3×3卷积替代5×5、7×7、9×9的普通卷积, 可以在不增加参数数量和计算次数的情况下, 扩大感受野。



(a) RGB 图像 (b) 热图像 (c) Baseline (d) Baseline+Fusion (e) Baseline+Fusion+TES (f) Ground-truth

图 7 消融实验结果对比

Fig. 7 Comparison of the results of ablation experiments

表 5 RTFM 中不同分支数量在 RGBT-CC 数据集上的结果对比

Table 5 Comparison results of different branch numbers in RTFM on the RGBT-CC dataset

分支数量	GAME(0)	GAME(1)	GAME(2)	GAME(3)	RMSE
1分支	17.95	23.61	30.84	39.42	32.26
2分支	16.22	21.81	28.03	38.51	28.54
3分支	14.27	18.95	24.16	34.92	24.37
4分支	12.74	16.94	21.67	30.63	21.81
5分支	13.47	18.11	23.06	33.03	22.24
6分支	14.53	19.46	25.23	35.62	23.68

4 结 论

1）本文提出了一种用于 RGB-T 的人群计数方法，基于多尺度感知与红外特征增强设计了一个跨模态人群计数网络，将尺度感知和上下文学习引入跨模态特征融合过程中，以处理 RGB-T 人群计数中的尺度变化、模态信息差异问题。

2）在融合跨模态特征时，利用密集连接和信息发散机制来融合 RGB 和热模态的特征，增强了模态信息共享，能更充分地捕获模态互补信息。

3）本文方法在 RGBT-CC 数据集和 ShanghaiTechRGBD 数据集上进行了广泛的实验，表明在跨模态人群计数方面具有优越性和有效性。

未来，将探索在面临复杂和困难的场景时，如何进一步提高人群计数的效率和准确性。

参考文献（References）

[1] LIU C H, CHEN Y F, HE X Y, et al. A scale aggregation and spatial-aware network for multi-view crowd counting[J]. IEEE Acces, 2022, 10: 108604-108613.

[2] GAO G S, GAO J Y, LIU Q J, et al. CNN-based density estimation and crowd counting: a survey[EB/OL]. (2020-03-28)[2024-01-03]. https://doi.org/10.48550/arXiv.2003.12783.

[3] LIU L, CHEN J, WU H, et al. Cross-modal collaborative representation learning and a large-scale RGBT benchmark for crowd counting[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 4821-4831.

[4] ZHANG Y, CHOI S, HONG S. Spatio-channel attention blocks for cross-modal crowd counting[C]//Proceedings of the Computer Vision–ACCV 2022. Berlin: Springer, 2023: 22-40.

[5] ZHANG S H, WANG W, ZHAO W B, et al. A cross-modal crowd counting method combining CNN and cross-modal transformer[J]. Image and Vision Computing, 2023, 129: 104592.

[6] WU Z T, LIU L B, ZHANG Y, et al. Multimodal crowd counting with mutual attention transformers[C]//Proceedings of the IEEE International Conference on Multimedia and Expo. Piscataway: IEEE Press, 2022: 1-6.

[7] TANG H H, WANG Y, CHAU L P. TAFNet: A three-stream adaptive fusion network for RGB-T crowd counting[C]//Proceedings of the IEEE International Symposium on Circuits and Systems. Piscataway: IEEE Press, 2022: 3299-3303.

[8] LI H, ZHANG J G, KONG W H, et al. CSA-Net: cross-modal scale-aware attention-aggregated network for RGB-T crowd count-

ing[J]. Expert Systems with Applications, 2023, 213: 119038.

[9] CHENG J H, CHEN Z J, ZHANG X Y, et al. Exploit the potential of multi-column architecture for crowd counting[EB/OL]. (2020-07-28)[2024-01-30]. https://doi.org/10.48550/arXiv.2007.05779.

[10] BAI S, HE Z Q, QIAO Y, et al. Adaptive dilated network with self-correction supervision for counting[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 4593-4602.

[11] LIU L B, QIU Z L, LI G B, et al. Crowd counting with deep structured scale integration network[C]//Proceedings of the IEEE/CVF IEEE International Conference on Computer Vision Workshops. Piscataway: IEEE Press, 2020: 1774-1783.

[12] LIU L B, ZHEN J J, LI G B, et al. Dynamic spatial-temporal representation learning for traffic flow prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(11): 7169-7183.

[13] LIU W Z, SALZMANN M, FUA P. Contextaware crowd counting [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 5094-5103.

[14] DENG L J, ZHOU Q H, WANG S H, et al. Deep learning in crowd counting: a survey[J]. CAAI Transactions on Intelligence Technology, 2024, 9(5): 1043-1077.

[15] LI B, HUANG H B, ZHANG A, et al. Approaches on crowd counting and density estimation: a review[J]. Pattern Analysis and Applications, 2021, 24(3): 853-874.

[16] JIANG X L, XIAO Z H, ZHANG B C, et al. Crowd counting and density estimation by trellis encoder-decoder networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 6126-6135.

[17] 余鹰, 朱慧琳, 钱进, 等. 基于深度学习的人群计数研究综述[J]. 计算机研究与发展, 2021, 58(12): 2724-2747.
YU Y, ZHU H L, QIAN J, et al. Survey on deep learning based crowd counting[J]. Journal of Computer Research and Development, 2021, 58(12): 2724-2747(in Chinese).

[18] ZHANG Y Y, ZHOU D S, CHEN S Q, et al. Single-image crowd counting via multi-column convolutional neural network[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 589-597.

[19] SAM D B, SURYA S, BABU R V. Switching convolutional neural network for crowd counting[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 4031-4039.

[20] LI Y H, ZHANG X F, CHEN D M. CSRNet: dilated convolutional neural networks for understanding the highly congested scenes[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 1091-1100.

[21] MA Z H, WEI X, HONG X P, et al. Bayesian loss for crowd count estimation with point supervision[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2020: 6141-6150.

[22] PENG T, LI Q, ZHU P F. RGB-T crowd counting from drone: a benchmark and MMCCN network[C]//Proceedings of the Computer Vision-ACCV 2020. Berlin: Springer, 2021: 497-513.

[23] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Proceedings of the Computer Vision-ECCV 2018. Berlin: Springer, 2018: 3-19.

[24] LIAN D Z, LI J, ZHENG J, et al. Density map regression guided detection network for RGB-D crowd counting and localization[C]//Proceedings of the IEEE/CVF Conference on Computer Cision and Pattern Recognition. Piscataway: IEEE Press, 2020: 1821-1830.

[25] GUERRERO-GÓMEZ-OLMEDO R, TORRE-JIMÉNEZ B, LÓPEZ-SASTRE R, et al. Extremely overlapping vehicle counting[C]//Proceedings of the Pattern Recognition and Image Analysis. Berlin: Springer, 2015: 423-431.

[26] ZHANG J, FAN D P, DAI Y C, et al. UC-Net: uncertainty inspired RGB-D saliency detection via conditional variational autoencoders[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 8579-8588.

[27] PANG Y W, ZHANG L H, ZHAO X Q, et al. Hierarchical dynamic filtering network for RGB-D salient object detection[C]//Proceedings of the Computer Vision-ECCV 2020. Berlin: Springer, 2020: 235-252.

[28] FAN D P, ZHAI Y J, BORJI A, et al. BBS-Net: RGB-D salient object detection with a bifurcated backbone strategy network[C]//Proceedings of the Computer Vision-ECCV 2020. Berlin: Springer, 2020: 275-292.

RGB-T crowd counting method with multi-scale perception and infrared feature enhancement

ZHENG Diwen[1], SHI Yangyu[1], XIE Chengjie[1], LU Shuhua[1, 2, *]

(1. College of Information and Cyber Security, People's Public Security University of China, Beijing 102600, China;
2. Key Laboratory of Security Technology and Risk Assessment Ministry of Public Security, Beijing 102600, China)

Abstract: In order to overcome the difficulty of crowd counting in low light, RGB-T crowd counting attempts to create maps of crowd density utilizing complimentary information from visual and thermal imagery. However, existing RGB-T crowd counting methods face issues such as scale variation and background interference during cross-modality information fusion. To tackle these challenges, we propose an RGB-T crowd counting method based on multi-scale perception and infrared feature enhancement (MSENet). Our approach presents an RGB-T feature fusion mechanism (RTFM) that creates an infrared enhancement structure to completely capture crowd information in thermal images and uses a multi-branch structure for multi-scale feature extraction. Additionally, we utilize dense connections and information divergence mechanisms to transfer complementary features to each modality, achieving a reusable expression of complementary features and enhanced modality features. We evaluate our proposed method on the RGBT-CC dataset and the ShanghaiTechRGBD dataset through comparative experiments. The results demonstrate that our method outperforms existing state-of-the-art approaches on the RGBT-CC dataset, exhibiting good accuracy, robustness and good generalization.

Keywords: crowd counting; infrared and visible image; feature enhancement; multi-scale perception; cross-modal feature fusion

Received: 2024-04-24; **Accepted**: 2024-09-13; **Published Online**: 2024-09-19 15:39

URL: link.cnki.net/urlid/11.2625.V.20240919.1254.003

Foundation item: Double First-Class Innovation Research Project for People's Public Security University of China (2023SYL08)

* **Corresponding author**. E-mail: lushuhua@ppsuc.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0247

基于音视频特征优化与跨模态 Transformer 的多模态情感分析

林宜山[1]，左景[1]，卢树华[1, 2, *]

(1. 中国人民公安大学 信息网络安全学院，北京 102600；2. 公安部安全防范技术与风险评估重点实验室，北京 102600)

摘 要：针对多模态情感分析中音视频模态特征质量较差、不同模态信息交互不够充分等问题，提出一种基于音视频特征优化与跨模态 Transformer(CMT) 的多模态情感分析方法。设计了一种音视频特征优化机制 (AVFOM)，通过与文本特征的协同作用，增加音视频特征的情感信息密度，提高音频和视频特征质量；设计了 CMT 结构，以文本为主，实现文本-音频、文本-视频模态的两两充分交互，学习不同模态的一致性信息。此外，引入基于自监督学习策略的标签生成方法，实现单模态情感预测任务，学习每个模态各自的特性。所提方法在 CMU-MOSI 和 CMU-MOSEI 这 2 个公开数据集上进行了大量实验验证与测试，结果表明：所提方法超越当前诸多性能先进的方法，有效提升了多模态情感分析的准确性。

关 键 词：多模态；情感分析；Transformer 模型；自监督学习；音视频特征优化
中图分类号：TP391
文献标志码：A **文章编号**：1001-5965(2026)06-2219-10

多模态情感分析主要利用文本、音频和视频等多模态数据来分析人们的情感倾向，预测人们的情感状态，在诸多领域得到广泛应用[1]，如人机交互、电子商务、舆情监督等。当前，多模态情感分析相关研究主要关注点之一是如何有效地实现不同模态特征之间的融合[2-4]，并提出诸多融合方法。其中，根据融合过程中不同模态所占的比重，可将融合方法大致划分为 2 类：第 1 类采用三元对称结构同等比例地对文本、音频和视频 3 种模态进行融合[5-6]，上述方法在多模态情感分析中均表现良好。然而，文本、音频和视频 3 种模态中的情感信息并不均衡，文本模态相较于非文本模态含有更为丰富的情感信息，因而在多模态情感分析任务中，文本模态特征质量明显优于音频和视频模态特征质量。如果在模态融合的过程中不考虑 3 种模态之间的相对重要性，将影响最终情感预测的效果[7]。因此，第 2 类方法通过增大文本模态的权重来突出文本模态在多模态情感分析任务中的相对重要性[8-9]，有助于提高多模态情感分析的准确性。

为解决多模态情感分析中音视频特征质量较差及不同模态之间交互不够充分等问题，本文提出一种基于音视频特征优化和跨模态 Transformer (CMT)的多模态情感分析模型，以实现对音视频特征的优化，并且加强不同模态特征之间的深层次交互，从而提高多模态情感分析的准确性。本文主要贡献如下：

1) 提出了基于音视频特征优化和 CMT 的多模态情感分析模型。通过引入预训练语言模型 SentiLARE 来提取含有更丰富情感信息的文本特征。

2) 设计了一个音视频特征优化机制 (audio and video features optimizaton mechanism, AVFOM), 通过与文本特征的协同作用来增加音频和视频模态的

收稿日期：2024-04-23；**录用日期**：2024-07-05；**网络出版时间**：2024-08-15 08:10
网络出版地址：link.cnki.net/urlid/11.2625.V.20240814.1652.003
基金项目：中国人民公安大学双一流创新研究专项 (2023SYL08)
* **通信作者**. E-mail：lushuhua@ppsuc.edu.cn

引用格式：林宜山，左景，卢树华. 基于音视频特征优化与跨模态 Transformer 的多模态情感分析 [J]. 北京航空航天大学学报，2026，52（6）：2219-2228. LIN Y S，ZUO J，LU S H. A multimodal sentiment analysis based on audio and video features optimization and cross-modal Transformer [J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52（6）：2219-2228（in Chinese）.

情感信息，提高音视频特征的质量。同时，通过CMT实现文本-音频、文本-视频的两两交互，充分挖掘不同模态内部的关联性信息，有助于提高后续的多模态融合效果。

3）将本文方法在2个公开数据集上进行训练与测试，结果表明，可有效提升多模态情感分析的准确性。

1 相关工作

1.1 多模态情感分析

近年来，随着自然语言处理、计算机视觉、语音信号分析等技术的发展，多模态情感分析研究取得显著进展。2011年，Morency等[10]首次提出联合文本、音频和视频3种模态的情感分析任务。此后，多模态情感分析受到人们越来越多的关注。多模态情感分析方法可以分为特征提取、特征融合、情感预测3个部分，其中，最主要的挑战是如何有效实现多模态特征的融合，为此，人们提出了诸多融合方法，主要包括早期融合、晚期融合、张量融合、分层融合等。

较早的多模态情感分析方法主要分为早期融合和晚期融合等方式。早期融合又叫特征级融合，先将文本、音频和视频特征进行融合，然后，将融合特征输入到单个模型中进行情感预测。在早期融合的研究方面，Poria等[11]利用循环神经网络和卷积神经网络来提取视频、文本特征，使用OpenAIR来提取音频特征，进行特征融合，并将融合特征输入到多核学习分类器中进行情感预测。Zadeh等[12]设计了多注意循环网络，使用长短期混合记忆网络对每个模态中的特定视图动态建模，并在每个时间步利用注意力机制识别不同的跨视图动态。早期融合能够更快地实现不同模态特征的关联，但无法有效消除特征融合造成的信息冗余。晚期融合又称为决策级融合，对每个模态特征选用单独的分类器进行处理和分类，将每个模态分类的结果融合为一个决策向量进行情感极性的预测。在晚期融合的研究中，Kampman等[13]通过获得文本、音频和视频3种模态各自的预测分数，在决策层对3种模态的情感权重进行加权计算，得到情感预测结果。Nojavanasghari等[14]对文本、音频和视频3种模态分别使用分类器进行训练，计算3个分类器的置信度得分的平均值，以进行情感预测。晚期融合使每个模态能够使用最优的分类器或模型来学习其特征，但需要耗费更多的时间和计算资源。

现有的多模态情感分析方法包括张量融合、分层融合等方法。在张量融合研究方面，Zadeh等[15]提出张量融合网络，运用三重笛卡儿积模拟单模态、双模态、三模态之间的相互作用。但这种方式会增大特征维度，增加模型的计算量，为解决该问题，Liu等[16]利用低秩张量进行特征融合，以减少计算量。在分层融合研究方面，Majumder等[17]设计了一种分层融合网络，将文本、音频和视频3种模态特征两两融合，形成3个联合特征，将3个联合特征融合为一个多模态联合特征，借助该融合方式以减少模态融合过程中产生的冗余信息。为更有效地实现3个联合特征的融合，Mai等[18]设计了一种双向残差长短期记忆递归神经网络，以充分挖掘不同联合特征中的情感信息，从而提高多模态情感分析的准确性。

1.2 基于Transformer的方法

Transformer是一个完全基于注意力机制，摒弃了卷积和递归的网络架构[19]。在多模态情感分析领域中，基于Transformer的融合方法受到越来越多的关注，Kim等[20]通过引入单流Transformer的方法来减少对模态内特征的影响，利用多模态掩码语言建模(multimodal masked language modeling, MMLM)和对齐预测(alignment prediction, AP)学习不同模态之间的一致性信息。Ma等[21]设计了多尺度合作的Transformer，多尺度机制在跨模态交互的过程中能够充分利用每个模态不同层次的语义信息，有助于提高后续多模态融合效果。Xu等[22]利用循环翻译的方式，通过Transformer进行模态间的翻译。Wang等[23]提出了跨模态层次融合模型，采用多任务辅助模块来减少不同模态之间的差异，提高多模态联合表征的质量，增强多模态情感分析模型的鲁棒性。Fu等[24]提出了基于Transformer的跨模态预测结构来定量测定音频模态对文本和视频模态的影响，利用对比学习方法来区分模态内的情感差异，并据此消除模态间的异质性信息。

2 本文方法

2.1 方法概述

本文模型的整体架构如图1所示，主要由特征提取、AVFOM、CMT、多模态融合、单模态预测任务(unimodal prediction task, UPT)5个部分组成。模型的输入包括文本、音频和视频3种模态，可分别表示为$\boldsymbol{U}_{\mathrm{t}}\in\mathbf{R}^{N_{\mathrm{t}}d_{\mathrm{t}}}$，$\boldsymbol{U}_{\mathrm{a}}\in\mathbf{R}^{N_{\mathrm{a}}d_{\mathrm{a}}}$，$\boldsymbol{U}_{\mathrm{v}}\in\mathbf{R}^{N_{\mathrm{v}}d_{\mathrm{v}}}$。利用特征提取模块来提取文本、音频和视频的特征，通过AVFOM，提高音视频特征的情感信息密度，提高音频和视频模态的质量，将文本、音频和视频特征输入到跨模态交互Transformer中，实现多模态之间的交互，学习模态之间的一致性信息，同时，利用

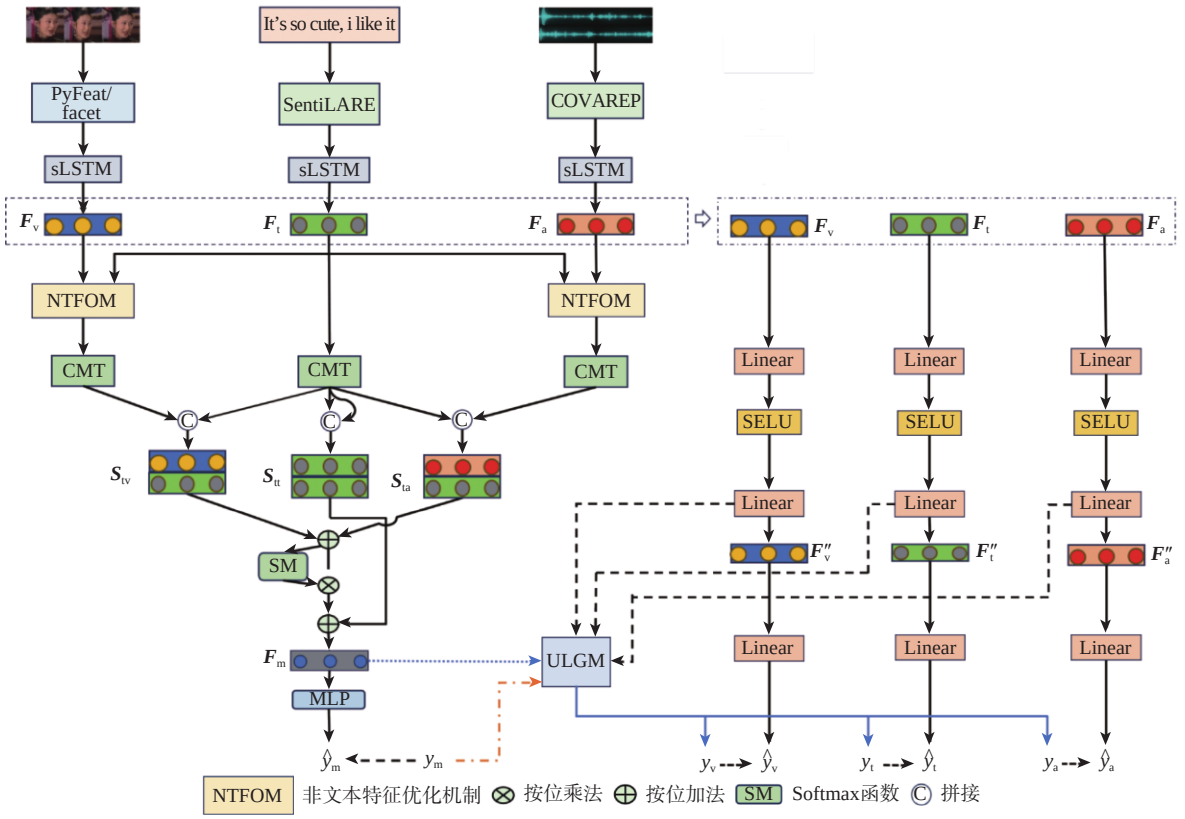


图 1　本文模型的整体结构

Fig. 1　The overall structure of the proposed model

UPT 学习单模态自身的异质性信息。

2.2　特征提取

1）文本特征提取：使用 SentiLARE[25] 作为预训练语言模型来提取文本特征，SentiLARE 在 RoBERTa[26] 的基础上进行改进，融入了情感极性和词性标签，能够提取更高质量的文本特征。

2）音频特征提取：在 CMU-MOSI[27] 和 CMU-MOSEI[28] 数据集上，选用 COVAREP[29] 来提取音频特征，音频特征向量的维数为 74。

3）视频特征提取：在 CMU-MOSI 数据集上，使用 Py-Feat[30] 来提取视频特征，视频特征向量的维数为 27。在 CMU-MOSEI 数据集上，使用 Feat 来提取视频特征，视频特征向量的维数为 35。

2.3　音视频特征优化机制

本文使用双向长短期记忆递归神经网络（Bi-directional long short-term memory, BiLSTM）对文本、音频和视频特征进行建模，提取其上下文语义信息，将 3 个模态映射到同一个特征空间，其工作机理可表示为

$$\boldsymbol{H}_m = \boldsymbol{w}_1 \mathrm{BiLSTM}(\boldsymbol{U}_m, \sigma) + \boldsymbol{b}_1 \qquad m \in \{\mathrm{t,a,v}\} \tag{1}$$

式中：$\boldsymbol{w}_1$ 为权重参数；$\boldsymbol{U}_{\mathrm{m}}$ 为原始的文本、音频和视频特征；σ 为可学习参数；$\boldsymbol{b}_1$ 为偏置项，用于调整线性变换后的特征偏移；$\boldsymbol{H}_m \in \mathbf{R}^{N_m d}$ 表示映射到同一特征空间的每个模态；t、a 和 v 分别为文本、音频和视频。

一般地，文本特征中含有更丰富的情感信息，通过文本特征的协同作用，来增大音视频特征中的情感信息，抑制音视频特征冗余信息，提高音视频特征质量。通过设计 AVFOM，以提高音视频特征的质量，有助于实现后续更高效的多模态特征融合。图 2 展示了 AVFOM 的具体结构，其工作过程描述为：文本特征和音视频特征分别通过激活函数 ReLU 对模态内的特征进行编码，经残差连接得到中间向量 $\boldsymbol{G}^{\mathrm{t}}$ 和 $\boldsymbol{G}^{\mathrm{m}}$，通过 Sigmoid 函数自适应地分配文本特征和音视频特征的比重（一般地，音视频特征比重低于文本特征比重），经多层感知器（multilayer perception, MLP）得到输出特征。AVFOM 的数学原理如下：

$$\boldsymbol{G}^{\mathrm{t}} = \mathrm{ReLU}(\boldsymbol{H}_{\mathrm{t}}) + \boldsymbol{H}_{\mathrm{t}} \tag{2}$$

$$\boldsymbol{G}^{\mathrm{m}} = \mathrm{ReLU}(\boldsymbol{H}_m) + \boldsymbol{H}_m \qquad m \in \{\mathrm{a,v}\} \tag{3}$$

$$\boldsymbol{X}_{\mathrm{g}} = \mathrm{Sigmoid}(\mathrm{Concat}(\boldsymbol{G}^{\mathrm{t}}, \boldsymbol{G}^{\mathrm{m}})) \tag{4}$$

$$\boldsymbol{\varepsilon} = \boldsymbol{G}^{\mathrm{t}} \boldsymbol{X}_{\mathrm{g}} + \boldsymbol{G}^{\mathrm{m}} (\boldsymbol{O} - \boldsymbol{X}_{\mathrm{g}}) \tag{5}$$

$$\boldsymbol{H}'_m = \mathrm{MLP}(\boldsymbol{\varepsilon} \boldsymbol{H}_m) \tag{6}$$

式中：$\boldsymbol{H}_{\mathrm{t}} \in \mathbf{R}^{N_t d}$ 为文本特征；$\boldsymbol{O}$ 为全为 1 的矩阵；$\boldsymbol{H}_m \in \mathbf{R}^{N_m d}$ 为音频或视频特征；MLP 为多层感知器；

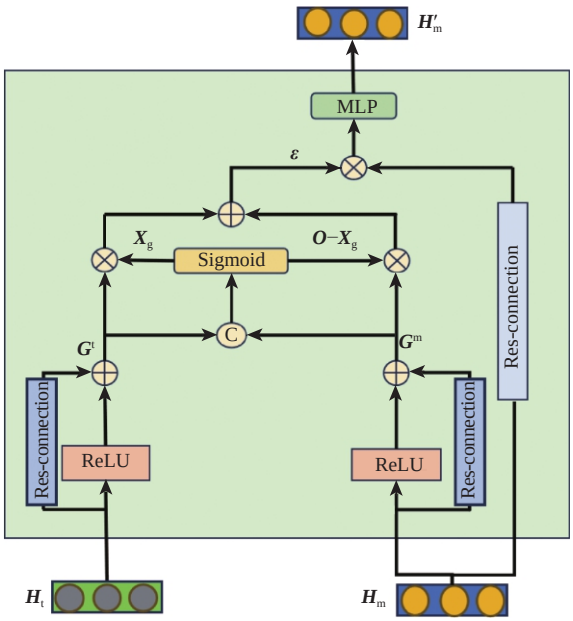


图2 音视频特征优化机制

Fig. 2 Audio and video features optimization mechanism

$\boldsymbol{H}'_m \in \mathbf{R}^{N_m d}$为输出特征。

2.4 跨模态交互 Transformer

如图3所示，由于在多模态情感分析任务中，文本模态相比音频模态和视频模态具有更高的情感贡献[31]。因此，设计了CMT机制，通过文本-音频、文本-视频模态的两两交互，对模态间的复杂依赖关系进行有效建模，捕获不同模态之间的情感信息，其工作原理简述如下。

首先，将双模态$\boldsymbol{H}'_1 \in \mathbf{R}^{N_1 d}$和$\boldsymbol{H}'_2 \in \mathbf{R}^{N_2 d}$作为CMT的模态输入，为捕获模态序列的时间和位置等信息，在每个模态特征上加入位置编码。

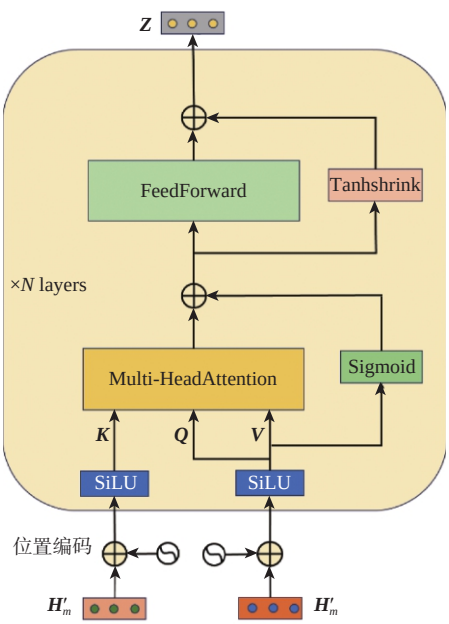


图3 跨模态Transformer

Fig. 3 Cross-modal Transformer

$$\boldsymbol{H}''_s = \boldsymbol{H}'_s + \boldsymbol{E}_{\mathrm{P}} \qquad s \in \{1,2\} \tag{7}$$

式中：$\boldsymbol{H}'_s$为经过特征优化机制增强后的音视频输出特征；$\boldsymbol{H}''_s$为加入位置编码后的特征序列；$\boldsymbol{E}_{\mathrm{P}}$为位置编码。

其次，将$\boldsymbol{H}''_1 \in \mathbf{R}^{N_1 d}$和$\boldsymbol{H}''_2 \in \mathbf{R}^{N_2 d}$这2个模态经过激活函数SiLU进行编码：

$$\boldsymbol{I}_1 = \mathrm{SiLU}(\boldsymbol{H}''_1) \tag{8}$$

$$\boldsymbol{I}_2 = \mathrm{SiLU}(\boldsymbol{H}''_2) \tag{9}$$

然后，经过多头注意力机制计算2种模态的注意力权重，实现2种模态之间的充分交互，具体过程为

$$\boldsymbol{P} = \mathrm{MultiHeadAttention} = \mathrm{Concat}(\boldsymbol{h}_1, \boldsymbol{h}_2, \cdots, \boldsymbol{h}_n) \tag{10}$$

$$\boldsymbol{h}_n = \mathrm{Attention}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V}) = \mathrm{Softmax}\left(\frac{\boldsymbol{Q}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d_K}}\right)\boldsymbol{V} = \mathrm{Softmax}\left(\frac{\boldsymbol{I}_1\boldsymbol{W}_Q\boldsymbol{W}_K^{\mathrm{T}}\boldsymbol{I}_2^{\mathrm{T}}}{\sqrt{d_K}}\right)\boldsymbol{I}_2\boldsymbol{W}_V \tag{11}$$

式中：$\boldsymbol{P}$为多头注意力机制的输出权重向量；Querys为$\boldsymbol{Q}=\boldsymbol{I}_1\boldsymbol{W}_Q$，Keys为$\boldsymbol{K}=\boldsymbol{I}_2\boldsymbol{W}_K$，Values为$\boldsymbol{V}=\boldsymbol{I}_2\boldsymbol{W}_V$，其中，$\boldsymbol{W}_Q \in \mathbf{R}^{dd_K}$，$\boldsymbol{W}_K \in \mathbf{R}^{dd_K}$，$\boldsymbol{W}_V \in \mathbf{R}^{dd_V}$是权重；$\boldsymbol{h}_n \in \mathbf{R}^{N_1 d_V}$具有与$\boldsymbol{Q}$相同的长度$N_1$和$\boldsymbol{V}$的特征维度$d_V$，从而实现了不同模态之间的交互，学习不同模态之间的信息。经前馈神经网络得到最终特征向量：

$$\boldsymbol{L} = \boldsymbol{P} + \mathrm{Sigmoid}(\boldsymbol{I}_2) \tag{12}$$

$$\boldsymbol{Z} = \mathrm{FeedForward}(\boldsymbol{L}) + \mathrm{Tanhshrink}(\boldsymbol{L}) \tag{13}$$

式中：$\boldsymbol{L}$为中间向量；$\boldsymbol{Z}$为输出的特征向量。

2.5 多模态特征融合

首先，通过分层融合的方法，将文本-音频、文本-视频进行两两融合，生成2个联合特征，借助Softmax函数来计算概率分布矩阵，通过加权的方式来增大联合特征中的情感信息。然后，将文本-文本进行融合，经残差连接得到多模态情感特征，具体过程可以表示为

$$\boldsymbol{S}_{\mathrm{ta}} = \mathrm{Concat}(\boldsymbol{Z}_{\mathrm{a}} + \boldsymbol{Z}_{\mathrm{t}}) \tag{14}$$

$$\boldsymbol{S}_{\mathrm{tv}} = \mathrm{Concat}(\boldsymbol{Z}_{\mathrm{v}} + \boldsymbol{Z}_{\mathrm{t}}) \tag{15}$$

$$\boldsymbol{S}_{\mathrm{tt}} = \mathrm{Concat}(\boldsymbol{Z}_{\mathrm{t}} + \boldsymbol{Z}_{\mathrm{t}}) \tag{16}$$

$$\boldsymbol{R} = \mathrm{Softmax}(\boldsymbol{S}_{\mathrm{ta}} + \boldsymbol{S}_{\mathrm{tv}}) \cdot (\boldsymbol{S}_{\mathrm{ta}} + \boldsymbol{S}_{\mathrm{tv}}) \tag{17}$$

$$\boldsymbol{F}'_{\mathrm{m}} = \boldsymbol{R} + \boldsymbol{S}_{\mathrm{tt}} \tag{18}$$

$$\hat{Y}_{\mathrm{m}} = \mathrm{MLP}(\boldsymbol{F}'_{\mathrm{m}}) \tag{19}$$

式中：$\boldsymbol{S}_{\mathrm{ta}}$为文本和音频融合的联合特征；$\boldsymbol{S}_{\mathrm{tv}}$为文本和视频融合的联合特征；$\boldsymbol{S}_{\mathrm{tt}}$为文本与文本的联合特征；$\boldsymbol{F}'_{\mathrm{m}}$为文本、音频和视频3种模态融合的多模态

情感特征, $\hat{Y}_{\mathrm{m}}$为最终的情感预测结果。

2.6　单模态预测任务

UPT(unimodal prediction task)用来捕获各模态自身的异质性信息, 提取单模态独特的情感特征, 以提升情感预测的准确性。其工作原理简要描述如下: 首先, 将文本、音频和视频模态特征投影到相同的特征空间中; 其次, 使用激活函数 SELU 对模态特征进行编码; 最后, 通过线性层得到单模态的预测结果。

$$\boldsymbol{F}'_c = \boldsymbol{w}_3(\mathrm{SELU}(\boldsymbol{w}_2\boldsymbol{F}_{\mathrm{c}} + \boldsymbol{b}_2)) + \boldsymbol{b}_3 \qquad c \in \{\mathrm{t,a,v}\} \tag{20}$$

$$\hat{\boldsymbol{Y}}_c = \boldsymbol{w}_4\boldsymbol{F}'_c + \boldsymbol{b}_4 \qquad c \in \{\mathrm{t,a,v}\} \tag{21}$$

式中: $\boldsymbol{F}_c$为原始的文本、音频和视频特征; $\hat{\boldsymbol{Y}}_c$为单模态预测值。

通过引入基于自监督策略的单模态标签生成模块 (unimodal label generation module, ULGM)[32] 来生成单模态标签, 其工作原理简述如下。

首先, 计算不同模态表征的积极中心和消极中心:

$$\boldsymbol{C}_{\mathrm{P}i} = \frac{\sum_{j=1}^{N} I(Y_i(j) > 0)\boldsymbol{F}'_{i,j}}{\sum_{j=1}^{N} I(Y_i(j) > 0)} \qquad i \in \{\mathrm{m,t,a,v}\} \tag{22}$$

$$\boldsymbol{C}_{\mathrm{N}i} = \frac{\sum_{j=1}^{N} I(Y_i(j) < 0)\boldsymbol{F}'_{i,j}}{\sum_{j=1}^{N} I(Y_i(j) < 0)} \qquad i \in \{\mathrm{m,t,a,v}\} \tag{23}$$

式中: $\boldsymbol{C}_{\mathrm{P}i}$为积极中心; $\boldsymbol{C}_{\mathrm{N}i}$为消极中心; $Y_i(j)$为通过自监督策略为单模态 i 中第 j 个样本生成的伪标签; $I(\cdot)$为指示函数; $\boldsymbol{F}'_{i,j}$为模态 i 中第 j 个样本的全局表征。

其次, 使用 L2 归一化来计算模态表征与积极中心、消极中心的距离:

$$D_{\mathrm{P}i} = \frac{\left\|\boldsymbol{F}'_i - \boldsymbol{C}_{\mathrm{P}i}\right\|_2^2}{\sqrt{d_i}} \qquad i \in \{\mathrm{m,t,a,v}\} \tag{24}$$

$$D_{\mathrm{N}i} = \frac{\left\|\boldsymbol{F}'_i - \boldsymbol{C}_{\mathrm{N}i}\right\|_2^2}{\sqrt{d_i}} \qquad i \in \{\mathrm{m,t,a,v}\} \tag{25}$$

式中: $D_{\mathrm{P}i}$与$D_{\mathrm{N}i}$为模态表征到积极中心、消极中心的距离; d_i为模态 i 的维度。

然后, 计算模态表征到积极中心与消极中心的相对距离:

$$\eta_i = \frac{D_{\mathrm{N}i} - D_{\mathrm{P}i}}{D_{\mathrm{P}i} + \Delta} \qquad i \in \{\mathrm{m,t,a,v}\} \tag{26}$$

式中: Δ为无限趋于 0 的数; η_i为相对距离。

$$\frac{Y_c^k}{Y_{\mathrm{m}}} \propto \frac{\hat{Y}_c^k}{\hat{Y}_{\mathrm{m}}^k} \propto \frac{\eta_c}{\eta_{\mathrm{m}}} \Rightarrow Y_c^k = \frac{\eta_c Y_{\mathrm{m}}^k}{\eta_{\mathrm{m}}} \qquad c \in \{\mathrm{t,a,v}\} \tag{27}$$

$$Y_c^k - Y_{\mathrm{m}} \propto \hat{Y}_c - \hat{Y}_{\mathrm{m}} \propto \eta_c - \eta_{\mathrm{m}} \Rightarrow Y_c^k = Y_{\mathrm{m}}^k + \eta_c - \eta_{\mathrm{m}} \tag{28}$$

$$Y_c^k = \frac{Y_{\mathrm{m}}^k \eta_c}{2\eta_{\mathrm{m}}} + \frac{Y_{\mathrm{m}}^k + \eta_c - \eta_{\mathrm{m}}}{2} = Y_{\mathrm{m}}^k + \frac{\eta_c - \eta_{\mathrm{m}}}{2} \cdot \frac{Y_{\mathrm{m}}^k + \eta_{\mathrm{m}}}{\eta_{\mathrm{m}}} = Y_{\mathrm{m}}^k + \mu_{\mathrm{cm}} \tag{29}$$

式中: μ_{cm}为通过自监督学习策略生成的单模态标签值与多模态标签值之间的偏置值; Y_{c}^k为单模态标签; Y_{m}^k为多模态标签; $\hat{Y}_{\mathrm{c}}$和$\hat{Y}_{\mathrm{m}}$为模型最终输出的情感预测结果。

2.7　损失函数

本文的损失包括单模态任务损失和多模态任务损失, 损失函数为

$$L_{\mathrm{all}} = L_{\mathrm{m}} + L_c \tag{30}$$

$$L_{\mathrm{m}} = L_{\mathrm{task}} + L_{\mathrm{strength}} + L_{\mathrm{polar}} \tag{31}$$

式中: L_{m}为多模态情感预测任务损失函数, 包含任务预测损失L_{task}、强度损失L_{strength}和极性损失L_{polar}; L_c为单模态情感预测任务损失函数。通过降低强度损失, 能够使模型的情感强度预测值更接近真实值; 通过降低极性损失, 能够提高模型对情感极性的预测准确度。

在训练阶段, 本文采用均方误差(mean squared error, MSE)损失函数来计算多模态任务预测损失L_{m}和单模态任务预测损失L_c。

$$L_{\mathrm{task}} = \frac{1}{N}\sum_{i=1}^{N}\left|Y_{\mathrm{m}}^i - \hat{Y}_{\mathrm{m}}^i\right| \tag{32}$$

$$L_c = \frac{1}{N}\sum_{j=1}^{N}\sum_{c}^{\{\mathrm{t,a,v}\}}\mathrm{Tanh}\left(\left|Y_c^j - Y_{\mathrm{m}}^j\right|\right)\left|Y_c^j - \hat{Y}_c^j\right| \tag{33}$$

在计算强度损失上, 采用相关系数距离函数:

$$D_{\mathrm{corr}}(X,Y) = 1 - \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}} \tag{34}$$

$$L_{\mathrm{strength}} = D_{\mathrm{corr}}(|Y_{\mathrm{m}}|, \left|\hat{Y}_{\mathrm{m}}\right|) \tag{35}$$

在计算极性损失上, 使用余弦相似度函数:

$$L_{\mathrm{polar}} = 1 - \alpha\mathrm{CosineSimilarity}(Y_{\mathrm{m}}^+, \hat{Y}_m^+) - \beta\mathrm{CosineSimilarity}(Y_{\mathrm{m}}^-, \hat{Y}_{\mathrm{m}}^-) \tag{36}$$

式中: α 和 β 分别为正余弦相似度和负余弦相似度的权重; "+" 和 "–" 表示正值和负值。

3　实验与结果分析

3.1　数据集

选用 2 个公开数据集 CMU-MOSI[27] 和 CMU-MOSEI[28] 来验证本文模型性能。

1）在多模态情感分析领域，CMU-MOSI数据集是非常流行的公开数据集之一，该数据集由93个Youtube的电影评论视频组成，上述视频被划分为2 199个视频片段，视频片段的情感分数为−3～3，表示非常消极、消极、有点消极、中性、有点积极、积极、十分积极。

2）CMU-MOSEI数据集相比于CMU-MOSI数据集而言，拥有更庞大的数据量，由5 000个视频所组成，包含有1 000个不同的发言者和250个不同的主题，上述视频被划分为23 453个视频片段。视频片段的情感分数为−3～3，对应从非常消极到非常积极的情感极性。

3.2 基准模型

为更好地验证本文模型的有效性，将本文模型与以下模型进行对比。

张量融合网络（tensor fusion network, TFN）[15]：通过张量外积方式对多模态间的信息进行建模，使用3种子嵌入网络的方式来提取文本、音频和视频特征。

多模态Transformer（multimodal transformer, MulT）[33]：通过双向跨模态注意力机制实现文本、音频和视频3种模态之间的交互。

自监督多任务多模态情感分析（self-supervised multi-task multimodal sentiment analysis, Slef-MM）[32]：设计了基于自监督学习策略的单模态标签生成机制来生成单模态标签用于UPT。

模态不变与模态特定表征（modality invariant and-specific attention/representations, MISA）[34]：通过模态不变子空间来学习不同模态的一致性信息，通过模态特定子空间来学习模态自身独特的信息。

多模态自适应门BERT（multimodal adaptation gate for BERT, MAG-Bert）[35]：设计了多模态自适应门来微调BERT和XLNet。

模态翻译情感分析（modality translation sentiment analysis, MTSA）[36]：通过Transformer网络将音频和视频模态翻译为本文模态，提高多模态情感分析模型的性能。

文本增强的Transformer融合网络（text enhanced transformer fusion network, TETFN）[37]：一种文本增强的Transformer网络，以文本模态为主、非文本模态为辅，通过多头注意力机制学习不同模态之间的一致性信息，并借助UPT保留每个模态自身独特的信息。

自适应模态加权多模态Transformer（multimodal transformer with adaptive modality weighting, MTAMW）[38]：自适应模态加权多模态Transformer，通过多模态自适应权重矩阵，根据文本、音频和视频模态对情感分析的贡献，为每个模态分配合适的权重。

协同空间表征交互网络（co-space representation interaction network, CRNet）[39]：将视频、音频模态分离到模态不变子空间和模态特定子空间中，并通过基于梯度的表征增强机制来增强对应的模态不变表征和模态特定表征。

基于特征恢复的动态交互网络（feature-based restoration dynamic interaction network, FRDIN）[40]：采用动态路由网络来学习不同模态之间的一致性信息和差异性信息。

多模态信息瓶颈情感分析（multimodal information bottleneck sentiment analysis, MIBSA）[41]：运用变分的方法，将文本、音频和视频3种单模态分离到模态不变子空间和模态特定子空间中，并运用信息瓶颈原理，通过最大化或最小化互信息来控制信息流。

3.3 实验细节

模型运行的操作系统为Ubuntu20.04.6LTS，处理器型号为Intel® Xeon(R),显卡型号为NVIDIA GeForce RTX3090，训练参数设置如表1所示。

表1 实验设置

Table 1 Experimental setup

数据集	学习率	批大小	注意力头	Transformer层
CMU-MOSI[27]	6.8×10^{-5}	32	16	1
CMU-MOSEI[28]	9.0×10^{-6}	32	16	1

1）在CMU-MOSI数据集上，优化器选用Adam，学习率设置为6.8×10^{-5}，批大小设置为32，注意力头设置为16，Transformer设置为1层。

2）在CMU-MOSEI数据集上，优化器选用Adam，学习率设置为9.0×10^{-6}，批大小设置为32，注意力头设置为16，Transformer设置为1层。

3.4 评价指标

采用5种评价指标来评估本文模型的性能，具体包括：平均绝对误差(mean absolute error, MAE)，用于评估预测值与真实值之间的平均绝对误差；皮尔逊相关系数(Corr)，用于衡量2个连续变量之间的连续相关性；七分类准确率(Acc-7)，范围在−3～3，对应非常消极、消极、有点消极、中性、有点积极、积极、非常积极；二分类准确率(Acc-2)，有2种不同的方法，一种是Acc-2(non-negative/negative)，另外一种是Acc-2(positive/negative)；F_1分数用来衡量模型精确度的指标。

3.5 结果与分析

3.5.1 与基准模型的对比分析

1）在CMU-MOSI数据集上的结果。如表2所示，在CMU-MOSI数据集上，与基于张量的多模态

情感分析模型 TFN 相比, 本文方法在所有的评价指标上都取得了明显进步。与基于 Transformer 的多模态情感分析模型 MulT、MAG-Bert、MTSA、TETFN 相比, 本文方法在所有的评价指标上同样取得更优的效果, 原因可能是利用 AVFOM 来增强音频特征和视频特征的质量, 以及采用 SentiLARE 预训练语言模型能够提取更高质量的文本特征, 因而能够提高多模态情感分析的准确性。与现有的先进模型 MTAMW、CRNet、FRDIN、MIBSA 相比, 本文方法在所有评价指标上均取得良好效果, 原因可能是, 利用 UPT 来学习单模态的特性, 能够关注到模态内部更为深层次的信息, 且使用 SentiLARE 预训练语言模型能够提取含有更丰富情感信息的文本特征, 通过 AVFOM 能够提高音频特征和视频特征的质量, 同时, 借助跨模态 Transfomer 能够实现不同模态之间更充分的交互, 从而提高后续情感预测的准确性。

2）在 CMU-MOSEI 数据集上的结果。如表 3 所示, 在 CMU-MOSEI 数据集上, 和基于张量的多模态情感分析模型 TFN 相比, 所提方法在所有评价指标上都取得了明显的进步。和基于 Transformer 的多模态情感分析模型 MulT、MAG-BERT、MTSA、TETFN 相比, 在大多数评价指标上, 本文方法都取得更优的效果, 但与 TETFN 方法相比, 在 (negative/non-negative)Acc-2 和 (negative/non-negative)F_1 指标上性能有所下降, 这可能是因为 CMU-MOSEI 数据集的数据量较大, 冗余信息较多, 所提模型对中性情感数据的识别不准确, 导致模型在这 2 个指标上

表 2　在 CMU-MOSI 数据集上与其他基准模型的对比结果

Table 2　The results compared with other baseline models on the CMU-MOSI dataset

模型	MAE	Corr	Acc-2/%	Acc-7/%	F_1/%
TFN_2[15]	0.901	0.698	*/80.8	34.90	*/80.7
$MulT_1$[33]	0.871	0.698	*/83.0	40.00	*/82.8
Self-MM[32]	0.713	0.798	84.0/85.98	*	84.42/85.95
MISA[34]	0.783	0.761	81.8/83.4	42.30	81.7/83.6
$MAG\text{-}Bert_2$[35]	0.731	0.798	82.5/84.3	*	82.6/84.3
MTSA[36]	0.696	0.806	*/86.8	46.40	*/86.8
TETFN[37]	0.717	0.800	84.05/86.10	*	83.83/86.07
MTAMW[38]	0.712	0.794	84.40/86.59	46.84	84.20/86.46
CRNet[39]	0.712	0.797	*/86.4	47.40	*/86.4
FRDIN[40]	0.682	0.813	85.8/87.4	46.59	85.3/87.5
MIBSA[41]	0.728	0.798	*/87.00	43.10	*/87.20
本文	**0.592**	**0.862**	**86.90/89.11**	**50.30**	**86.96/89.12**

注: $MulT_1$模型表示结果来自(AOBERT)[20], TFN_2, $MAG\text{-}Bert_2$模型表示结果来自(VLP2MSA)[42], 其他模型结果来自原论文; “*” 表示原文未提供结果; 对于Acc-2和F_1, “/” 的左边代表negative/non-negative的方法, 右边代表negative/positive的方法; 加粗数字表示最优值。

表 3　在 CMU-MOSEI 数据集上与其他基线模型对比的结果

Table 3　The results compared with other baseline models on the CMU-MOSEI dataset

模型	MAE	Corr	Acc-2/%	Acc-7/%	F_1/%
TFN_2[15]	0.593	0.700	*/82.50	50.20	*/82.10
$MulT_1$[33]	0.580	0.703	*/82.50	51.80	*/82.30
Self-MM[32]	0.530	0.765	82.81/85.17	*	82.53/85.30
MISA[34]	0.555	0.756	83.60/85.50	52.20	83.80/85.30
$MAG\text{-}Bert_2$[35]	0.543	0.755	82.51/84.82	*	82.77/84.71
MTSA[36]	0.541	0.774	*/85.50	52.90	*/85.30
TETFN[37]	0.551	0.748	**84.25**/85.18	*	**84.18**/85.27
MTAMW[38]	0.525	0.782	83.09/86.49	53.73	83.48/86.45
CRNet[39]	0.541	0.771	*/86.20	53.80	*/86.10
FRDIN[40]	0.525	0.778	83.30/86.30	54.40	83.70/86.20
MIBSA[41]	0.568	0.753	*/86.70	52.40	*/85.80
本文	**0.519**	**0.791**	83.66/**86.77**	54.41	83.24/**86.76**

注: $MulT_1$模型结果来自(AOBERT)[20], TFN_2, $MAG\text{-}Bert_2$模型结果来自(VLP2MSA)[42], 其他模型结果来自原论文, “*” 表示原文未提供结果, 对于Acc-2和F_1, “/” 的左边代表negative/non-negative的方法, 右边代表negative/positive的方法; 加粗数字表示最优值。

的性能表现较差。与现有的先进模型 MTAMW、CRNet、FRDIN、MIBSA 相比，所提方法在所有评价指标上均取得更优异的效果。

3.5.2 消融实验

如表 4 所示，为验证本文模型不同组分的有效性，在 CMU-MOSI 数据集上进行消融实验。将 SentiLARE 预训练语言模型替换为 BERT 预训练语言模型，可以看出，所有评价指标都出现了明显下降，表明 SentiLARE 预训练语言模型能够提取含有更高质量的文本特征。移除 AVFOM 后，MAE 增加了 0.025，Acc-7 降低了 1.79%，其他评价指标也均有不同程度的下降，表明 AVFOM 能够提高音频特征和视频特征的质量。删去 CMT 机制后，模型性能同样有所下降，其中，Acc-7 下降了 2.83%，表明 CMT 机制能够实现不同模态特征的充分交互。移除 UPT 后，MAE 增大了 0.021，Acc-7 下降了 2.23%，其他评价指标均有不同程度的下降，表明 UPT 能够有效提高多模态情感分析任务的准确性。分别移除音频、视频、文本模态后，所有评价指标均有不同程度的下降，表明上述 3 种模态信息均有助于提高情感分析任务的准确性，其中，移除文本模态之后，与移除音频和视频模态相比，MAE 的增大幅度更高，其他评价指标也出现显著下降，表明文本模态相比音频、视频模态对情感分析任务具有更高的贡献度。

表 4 在 CMU-MOSI 数据集上的消融实验

Table 4 The ablation experiments on the CMU-MOSI dataset

模型	MAE	Corr	Acc-2/%	Acc-7/%	F_1/%
本文	0.592	0.862	86.90/89.11	50.30	86.96/89.12
w/o SentiLARE	0.837	0.743	82.14/84.14	39.43	82.20/84.13
w/o AVFOM	0.617	0.856	85.86/88.16	48.51	85.92/88.17
w/o CMT	0.613	0.854	85.71/87.38	47.47	85.78/87.41
w/o UPT	0.613	0.855	86.01/88.16	48.07	86.12/88.22
w/o audio	0.618	0.854	86.16/88.01	46.28	86.23/88.03
w/o video	0.621	0.847	85.86/88.01	49.40	85.89/87.99
w/o text	1.426	0.012	44.70/45.98	15.60	49.83/48.46

3.5.3 特征可视化

如图 4 所示，为验证本文方法的有效性，在 CMU-MOSI 测试集上使用分布随机领域嵌入（T-distributed stochastic neighbor enbedding, T-SNE）工具对多模态特征进行可视化，可以看出，本文模型能够有效地分离积极特征和消极特征。

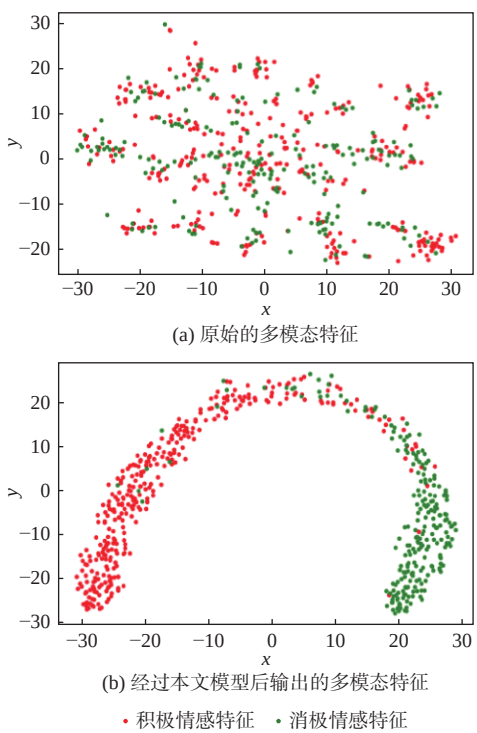


图 4 特征可视化

Fig. 4 Feature visualization

4 结 论

1）本文提出的基于音视频特征优化和 CMT 的多模态情感分析方法在 CMU-MOSI和 CMU-MOSEI 这 2 个公开数据集上和其他模型的对比结果表明，本文方法通过引入 SentiLARE 预训练语言模型，能够提取含有更丰富情感信息的文本特征，利用 AVFOM 提高音视频模态特征的质量，有助于后续多模态特征的融合。

2）通过 CMT 机制，能够对不同模态间复杂的依赖关系进行建模，实现不同模态特征更充分的交互，从而更好地完成多模态情感预测的任务。

3）此外，还借助 UPT 来学习每个模态自身的特性，提升情感预测的准确性。

参考文献（References）

[1] GANDHI A, ADHVARYU K, PORIA S, et al. Multimodal sentiment analysis: a systematic review of history, datasets, multimodal fusion methods, applications, challenges and future directions[J]. Information Fusion, 2023, 91: 424-444.

[2] HUANG C Q, ZHANG J L, WU X M, et al. TeFNA: text-centered fusion network with crossmodal attention for multimodal sentiment analysis[J]. Knowledge-Based Systems, 2023, 269: 110502.

[3] CHEN C, HONG H S, GUO J, et al. Inter-intra modal representation augmentation with trimodal collaborative disentanglement network for multimodal sentiment analysis[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2023, 31: 1476-1488.

[4] CHEN Q P, HUANG G M, WANG Y B. The weighted cross-modal attention mechanism with sentiment prediction auxiliary task for multimodal sentiment analysis[J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2022, 30: 2689-2695.

[5] YUAN Z Q, LI W, XU H, et al. Transformer-based feature reconstruction network for robust multimodal sentiment analysis[C]//Proceedings of the 29th ACM International Conference on Multimedia. New York: ACM, 2021: 4400-4407.

[6] SUN H, CHEN Y W, LIN L F. TensorFormer: a tensor-based multimodal Transformer for multimodal sentiment analysis and depression detection[J]. IEEE Transactions on Affective Computing, 2023, 14(4): 2776-2786.

[7] SUN Z K, SARMA P, SETHARES W, et al. Learning relationships between text, audio, and video via deep canonical correlation for multimodal language analysis[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(5): 8992-8999.

[8] WU Y, LIN Z J, ZHAO Y Y, et al. A text-centered shared-private framework via cross-modal prediction for multimodal sentiment analysis[C]//Proceedings of the Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021. Kerrville: Association for Computational Linguistics, 2021: 4730-4738.

[9] HU G M, LIN T E, ZHAO Y, et al. UniMSE: Towards unified multimodal sentiment analysis and emotion recognition[EB/OL]. (2022-11-21)[2024-01-10]. https://doi.org/10.48550/arXiv.2211.11256.

[10] MORENCY L P, MIHALCEA R, DOSHI P. Towards multimodal sentiment analysis: Harvesting opinions from the web[C]//Proceedings of the 13th International Conference on Multimodal Interfaces. New York: ACM, 2011: 169-176.

[11] PORIA S, CAMBRIA E, GELBUKH A. Deep convolutional neural network textual features and multiple kernel learning for utterance-level multimodal sentiment analysis[C]//Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing. Kerrville: Association for Computational Linguistics, 2015: 2539-2544.

[12] ZADEH A, LIANG P P, PORIA S, et al. Multi-attention recurrent network for human communication comprehension[C]//Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence. New York: ACM, 2018: 5642-5649.

[13] KAMPMAN O, BAREZI E J, BERTERO D, et al. Investigating audio, visual, and text fusion methods for end-to-end automatic personality prediction[EB/OL]. (2018-05-16)[2024-01-12]. https:/doi.org/10.48550//arXiv.1805.00705.

[14] NOJAVANASGHARI B, GOPINATH D, KOUSHIK J, et al. Deep multimodal fusion for persuasiveness prediction[C]//Proceedings of the 18th ACM International Conference on Multimodal Interaction. New York: ACM, 2016: 284-288.

[15] ZADEH A, CHEN M H, PORIA S, et al. Tensor fusion network for multimodal sentiment analysis[EB/OL]. (2017-07-23)[2024-01-13]. https://doi.org/10.48550/arXiv.1707.07250.

[16] LIU Z, SHEN Y, LAKSHMINARASIMHAN V B, et al. Efficient low-rank multimodal fusion with modality-specific factors[EB/OL]. (2018-05-31)[2024-01-13]. https://doi.org/10.48550/arXiv.1806.00064.

[17] MAJUMDER N, HAZARIKA D, GELBUKH A, et al. Multimodal sentiment analysis using hierarchical fusion with context modeling[J]. Knowledge-Based Systems, 2018, 161: 124-133.

[18] MAI S J, HU H F, XING S L. Divide, conquer and combine: hierarchical feature fusion network with local and global perspectives for multimodal affective computing[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Kerrville: Association for Computational Linguistics, 2019: 481-492.

[19] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of the Advances in Neural Information Processing Systems. Red Hook: Curran Associates, 2017: 5998-6008.

[20] KIM K, PARK S. AOBERT: All-modalities-in-One BERT for multimodal sentiment analysis[J]. Information Fusion, 2023, 92(C): 37-45.

[21] MA L Y, YAO Y, LIANG T, et al. Multi-scale cooperative multimodal Transformers for multimodal sentiment analysis in videos [EB/OL]. (2022-06-17)[2024-01-15]. https://doi.org/10.48550/arXiv.2206.07981.

[22] XU M, LIANG F F, SU X Y, et al. CMJRT: cross-modal joint representation Transformer for multimodal sentiment analysis[J]. IEEE Access, 2022, 10: 131671-131679.

[23] WANG L, PENG J J, ZHENG C Z, et al. A cross modal hierarchical fusion multimodal sentiment analysis method based on multi-task learning[J]. Information Processing & Management, 2024, 61(3): 103675.

[24] FU Y P, ZHANG Z Y, YANG R D, et al. Hybrid cross-modal interaction learning for multimodal sentiment analysis[J]. Neurocomputing, 2024, 571: 127201.

[25] KE P, JI H Z, LIU S Y, et al. SentiLARE: sentiment-aware language representation learning with linguistic knowledge[EB/OL]. (2020-09-24)[2024-01-17]. https://doi.org/10.48550/arXiv.1911.02493.

[26] LIU Y H, OTT M, GOYAL N, et al. RoBERTa: a robustly optimized BERT pretraining approach[EB/OL]. (2019-07-26)[2024-01-17]. https://doi.org/10.48550/arXiv.1907.11692.

[27] ZADEH A, ZELLERS R, PINCUS E, et al. MOSI: multimodal corpus of sentiment intensity and subjectivity analysis in online opinion videos[EB/OL]. (2016-08-12)[2024-01-18]. https://doi.org/10.48550/arXiv.1606.06259.

[28] BAGHER ZADEH A, LIANG P P, PORIA S, et al. Multimodal language analysis in the wild: CMU-MOSEI dataset and interpretable dynamic fusion graph[C]//Proceedings of the 56th Annual Meeting of the Association for Computational Linguistic. Kerrville: Association for Computational Linguistics, 2018: 2236-2246.

[29] DEGOTTEX G, KANE J, DRUGMAN T, et al. COVAREP: a collaborative voice analysis repository for speech technologies[C]//

Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE Press, 2014: 960-964.

[30] CHEONG J H, JOLLY E, XIE T K, et al. Py-Feat: python facial expression analysis toolbox[J]. Affective Science, 2023, 4(4): 781-796.

[31] LIN H, ZHANG P L, LING J D, et al. PS-mixer: a polar-vector and strength-vector mixer model for multimodal sentiment analysis[J]. Information Processing & Management, 2023, 60(2): 103229.

[32] YU W M, XU H, YUAN Z Q, et al. Learning modality-specific representations with self-supervised multi-task learning for multimodal sentiment analysis[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(12): 10790-10797.

[33] TSAI Y H H, BAI S J, LIANG P P, et al. Multimodal Transformer for unaligned multimodal language sequences[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Kerrville: Association for Computational Linguistics, 2019: 6558-6569.

[34] HAZARIKA D, ZIMMERMANN R, PORIA S. MISA: Modality-invariant and-specific representations for multimodal sentiment analysis[C]//Proceedings of the 28th ACM International Conference on Multimedia. New York: ACM, 2020: 1122-1131.

[35] RAHMAN W, HASAN M K, LEE S, et al. Integrating multimodal information in large pretrained Transformers [C]//Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics. Kerrville: Association for Computational Linguistics, 2020: 2359-2369.

[36] YANG B, SHAO B, WU L J, et al. Multimodal sentiment analysis with unidirectional modality translation[J]. Neurocomputing, 2022, 467: 130-137.

[37] WANG D, GUO X T, TIAN Y M, et al. TETFN: a text enhanced Transformer fusion network for multimodal sentiment analysis[J]. Pattern Recognition, 2023, 136: 109259.

[38] WANG Y F, HE J H, WANG D, et al. Multimodal Transformer with adaptive modality weighting for multimodal sentiment analysis[J]. Neurocomputing, 2024, 572: 127181.

[39] SHI H, PU Y Y, ZHAO Z P, et al. Co-space representation interaction network for multimodal sentiment analysis[J]. Knowledge-Based Systems, 2024, 283: 111149.

[40] ZENG Y F, LI Z X, CHEN Z B, et al. A feature-based restoration dynamic interaction network for multimodal sentiment analysis[J]. Engineering Applications of Artificial Intelligence, 2024, 127: 107335.

[41] LIU W, CAO S C, ZHANG S. Multimodal consistency-specificity fusion based on information bottleneck for sentiment analysis[J]. Journal of King Saud University-Computer and Information Sciences, 2024, 36(2): 101943.

[42] YI G F, FAN C H, ZHU K, et al. VLP2MSA: Expanding vision-language pre-training to multimodal sentiment analysis[J]. Knowledge-Based Systems, 2024, 283: 111136.

A multimodal sentiment analysis based on audio and video features optimization and cross-modal Transformer

LIN Yishan[1], ZUO Jing[1], LU Shuhua[1, 2, *]

(1. College of Information and Cyber Security, People's Public Security University of China, Beijing 102600, China;
2. Key Laboratory of Security Technology and Risk Assessment Ministry of Public Security, Beijing 102600, China)

Abstract: To solve problems including low-quality audio and video modal features and inadequate interaction between various modalities, a multimodal sentiment analysis approach based on cross-modal Transformer (CMT) and audio and video feature optimization is suggested. Firstly, we propose a audio and video features optimizing mechanism (AVFOM), which increases the density of sentiment information in audio and video features through synergistic interaction with textual features, thereby improving the quality of audio and video features. Secondly, in order to accomplish full interaction between text-audio and text-video modalities and learn consistent knowledge across various modalities, we construct a cross-modal Transformer structure with text as the dominant modality. Additionally, a label generation method based on the self-supervised learning strategy is introduced to perform single-modality sentiment prediction tasks, learning the characteristics of each modality separately. The proposed method is extensively validated and tested on two public datasets, CMU-MOSI and CMU-MOSEI, which surpass many currently advanced methods in terms of performance and effectively improve the accuracy of multimodal sentiment analysis.

Keywords: multimodal; sentiment analysis; Transformer; self-supervised learning; audio and video features optimization

Received: 2024-04-23; **Accepted**: 2024-07-05; **Published Online**: 2024-08-15 08:10

URL: link.cnki.net/urlid/11.2625.V.20240814.1652.003

Foundation item: Double First-Class Innovation Research Project for People's Public Security University of China (2023SYL08)

* **Corresponding author**. E-mail: lushuhua@ppsuc.edu.cn

《北京航空航天大学学报》征稿简则

《北京航空航天大学学报》是北京航空航天大学主办的以航空航天科学技术为特色的综合性自然科学学术期刊(月刊)。本刊以反映航空航天领域研究成果与动态、促进学术交流、培养科技人才和推动科技成果向社会生产力转化为办刊宗旨。本刊为中国自然科学技术核心期刊,并被 Ei Compendex 等国内外权威文献检索系统收录。本刊向国内外公开发行,为进一步提高办刊质量和所刊出文章的学术水平,特制定本简则。

1 论文作者及内容

1.1 本刊面向海内外所有学者。

1.2 主要刊载与航空航天科学技术有关的材料科学及工程、飞行器设计与制造、宇航科学与工程、信息与电子技术、控制技术和自动化工程、流体力学和动力工程、计算机科学及应用技术、可靠性工程与失效分析等领域的研究文章。航空航天科学技术民用方面以及具有航空航天工程背景的应用数学、应用物理、应用力学和工程管理等方面的文章也在本刊优先考虑之列。

2 来稿要求

2.1 论文应具有创新性、科学性、学术性和可读性。

2.2 论文为原创作品,尚未公开发表过,并且不涉及泄密问题。若发生侵权或泄密问题,一切责任由作者承担。

2.3 主题明确,数据可靠,图表清晰,逻辑严谨,文字精练,公式符号标准规范,标点符号正确。

2.4 文稿撰写顺序:中文题名(一般不超过 20 个汉字),作者中文姓名、单位、所在城市、邮政编码(作者顺序始终不得更改),中文摘要(包括目的、方法、结果及结论),中文关键词(5 ~ 8 个),中图分类号,引言,正文,参考文献,英文题名,作者英文姓名、单位、所在城市、邮政编码、国别,英文摘要,英文关键词。首、末页下角注明基金项目名称及编号,作者信息。

2.5 作者请登录本刊网页进行在线投稿。

3 稿件的审核、录用与版权

3.1 来稿须经专家两审和主编、编委讨论后决定刊用与否。

3.2 若来稿经审查后认定不宜在本刊发表,将及时告知作者。如果在投稿满 3 个月后仍未收到本刊任何通知,作者有权改投其他刊。在此之前,请勿一稿多投,否则一切后果自负。

3.3 来稿一经作者签字并在本刊刊出,即表明所有作者都已经认可其版权转至本刊编辑部。本刊在与国内外文献数据库或检索系统进行交流及合作时,不再征询作者意见。

邮寄地址:100191 北京市海淀区学院路 37 号 《北京航空航天大学学报》编辑部
办公地点:北京航空航天大学办公楼西配楼412~416房间
电　　话:(010)82316698,82316981,82315499,82317448
E-mail: jbuaa@ buaa. edu. cn
http://bhxb. buaa. edu. cn
http://www. buaa. edu. cn

集群成员期刊

《北京航空航天大学学报》
第五届编辑委员会

主　任（主　编）：赵沁平

（以下按姓氏笔画为序）

副主任（副主编）：丁希仑　王少萍　孙志梅　李秋实　李焕喜　杨嘉陵
苗俊刚　相　艳　徐立军　钱德沛　曹晋滨

编　委：马殿富　王　琪　王　聪　邓小燕　王青云　王荣明　刘　宇
刘　红　江　洁　刘　强　闫　鹏　朱天乐　刘铁钢　齐铂金
陈万春　邹正平　苏东林　杨世春　沈成平　邱志平　宋知人
杨树斌　张晓林　杨晓奕　杨继萍　李惠峰　吴新开　张瑞丰
杨照华　宋凝芳　周　锐　林宇震　林贵平　战　强　姚仰平
胡庆雷　赵秋红　段海滨　赵巍胜　席　平　郭　宏　徐　洁
徐世杰　郭洪波　康　锐　翟　锦　熊华钢

编辑部主任：张　嵘

编　　辑：孙　芳　万玉敏　王　茜　卞欢欢　李亚泰

北京航空航天大学学报

Beijing Hangkong Hangtian Daxue Xuebao

（原《北京航空学院学报》）

（月刊　1956 年创刊）

第52卷 第6期　2026年6月

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS (JBUAA)

(Monthly, Started in 1956)

Vol.52 No.6　June 2026

主管单位　中华人民共和国工业和信息化部
主办单位　北京航空航天大学
主　　编　赵沁平
编辑出版　《北京航空航天大学学报》编辑部
邮　　编　100191
地　　址　北京市海淀区学院路37号
印　　刷　北京科信印刷有限公司
发　　行　《北京航空航天大学学报》编辑部
发行范围　国内外发行
联系电话　(010) 82316698 82316981
82315499 82317448
电子信箱　jbuaa@ buaa.edu.cn

Administrated by Ministry of Industry and Information Technology of the People's Republic of China
Sponsored by Beijing University of Aeronautics and Astronautics (BUAA)
(Beijing 100191, P. R. China)
Chief Editor ZHAO Qinping
Edited and Published by Editorial Board of JBUAA
Printed by Beijing Kexin Printing Co., Ltd.
Distributed by Editorial Board of JBUAA
Telephone (010) 82316698 82316981
82315499 82317448
E-mail jbuaa@buaa.edu.cn
http: //bhxb.buaa.edu.cn

中国标准连续出版物号：ISSN 1001-5965 / CN 11-2625/V

ISSN 1001-5965

国内定价：50.00元／期

微信公众号　超星电子书