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

2023-07
Vol.49 No.7

北京航空航天大学学报

第 49 卷　第 7 期　（总第 365 期）　2023 年 7 月

目　　次

软件定义卫星技术概念及发展 …… 徐帆江，周鑫，赵军锁，吴凤鸽，林翊钧，夏玉立（1543）

民用飞机迫降后不利姿态应急撤离实验研究 …… 封文春，张桐，李伟（1553）

四旋翼的改进 PSO-RBF 神经网络自适应滑模控制 …… 唐志勇，马福源，裴忠才（1563）

液氮温区二维指向深冷环路热管设计与实验研究 …… 李楠，郭元东，许程，张红星，林贵平（1573）

电源车传感器故障检测和数据重构方法 …… 蒋栋年，把余江，李炜（1583）

特征点分段提取的时间序列模式匹配方法 …… 李正欣，刘畅，吴诗辉，郭建胜（1593）

基于角度转化的空间碎片测角数据精密定轨方法 …… 张耀，刘静（1600）

真空强电磁场环境下铝的二次电子倍增规律 ……
…… 商圣飞，杨晓宁，杨勇，毕研强，武南开，于澜涛（1606）

湿热环境下碳纤维层合板拉伸疲劳性能 …… 许名瑞，曾本银，熊欣，孟庆春，程小全（1614）

一种主瓣干扰下基于滑窗子阵的稀疏测角方法 ……
…… 周必雷，李荣锋，曾丽，陈辉，刘维建，李槟槟（1623）

高致密弹携式蜂群布局与多体分离方案 …… 富佳伟，王辰（1630）

高超声速再入飞行器气动布局多目标优化 …… 樊博璇，陈桂明，曹逸群（1639）

基于特征运动观测的蝴蝶前飞规律及样机验证 …… 张益鑫，王兴坚，王少萍，池小楷，杜韶阳（1651）

利用北斗 GEO 卫星干涉信号功率反演土壤湿度 ……
…… 汉牟田，许志超，常青，张波，王峰，洪学宝（1661）

下表面射流的超临界翼型气动性能分析 …… 王若尘，张国鑫，王翔宇，马晓平（1671）

撤机对离心血泵血液相容性影响的数值研究 …… 谢楠，唐雨萌，张岩，柳阳威（1680）

适用于弹性飞机飞行动力学仿真的气动力降阶方法 …… 师妍，万志强，吴志刚，巩轶男（1689）

高速超声振动铣削钛合金实验研究 …… 韩雄，孙哲飞，耿大喜，张德远（1707）

光纤环多极对称绕法对 Shupe 误差抑制效果仿真分析 …… 张东波，汪立新，李灿（1715）

空间大伸展并联机构的设计与性能分析 …… 赫利涛，房海蓉，陈宇飞，李寅（1722）

超声波钻探器结构参数对输出特性的影响分析 ……
…… 张兴旺，张明，于登云，曾婷，殷参，杨帅，庞勇（1735）

基于行人姿态的轨迹预测方法 …… 王瑞平，宋晓，陈凯，龚开奇，张峻凡（1743）

兔激素性股骨头坏死的造模改进及结果评价 …… 杨丰合，冯桂宇，李晋玉，牛旭锋（1755）

旋翼/机翼气动干扰对复合式直升机性能影响 …… 杨克龙，韩东（1761）
基于度量学习的机场交通态势弱监督评估 …… 杜婧涵，胡明华，张魏宁，尹嘉男（1772）
星载 GNSS-R 土壤湿度反演中开放水域的影响 ……
…… 杨文涛，徐天河，王娜子，高凡，荆丽丽，贺匀峤（1779）
考虑任务分配的无人机信息交互拓扑生成 …… 薛莹，何锋，谷晓燕（1787）
基于 LPV 的主动隔振平台单支腿鲁棒控制器设计 …… 唐育聪，朱庆华，刘付成，曹赫扬（1796）
基于 NARX 和 Kriging 的时变可靠性分析双层代理模型 …… 常泽明，李璐祎（1802）
基于可达集的无人机低空飞行冲突解脱算法 …… 杨建航，张福彪，王江（1813）
遥感图像语义分割的空间增强注意力 U 型网络 …… 宝音图，刘伟，李润生，李钦，胡庆（1828）
芯片间时间触发消息堆叠调度方法 …… 臧光界，李峭，王彤，熊华钢（1838）
CFRP 层压板脱黏缺陷红外脉冲热波层析成像检测 ……
…… 钟佳岑，徐浩军，魏小龙，韩欣珉，常怡鹏（1847）
圆周均匀多相调制相关辐射计及其灵敏度分析 …… 陈锐厚，罗景，胡岸勇，苗俊刚（1857）
基于变弯度后缘的机翼阵风响应减缓数值研究 …… 尉濛恺，戴玉婷，杨超，于思恒（1864）

期刊基本参数：CN 11-2625/V * 1956 * m * A4 * 332 * zh * P * ￥50.00 * 400 * 35 * 2023-07

（**编　辑**　张　嵘　孙　芳　苏　磊　卞欢欢　王　茜　李亚泰　李艳霞）

JOURNAL OF BEIJING UNIVERSITY OF
AERONAUTICS AND ASTRONAUTICS

Vol. 49 No. 7 （Sum 365） July 2023

CONTENTS

Conception and development of software-defined satellite technology
…… XU Fanjiang，ZHOU Xin，ZHAO Junsuo，WU Fengge，LIN Yijun，XIA Yuli （1543）

Experimental study on adverse attitude emergency evacuation of civil aircraft after crash landing
…… FENG Wenchun，ZHANG Tong，LI Wei （1553）

Improved PSO-RBF neural network adaptive sliding mode control for quadrotor systems
…… TANG Zhiyong，MA Fuyuan，PEI Zhongcai （1563）

Design and experiment of cryogenic loop heat pipe of two-dimensional pointing at liquid nitrogen zone
…… LI Nan，GUO Yuandong，XU Cheng，ZHANG Hongxing，LIN Guiping （1573）

Sensor fault detection and data reconstruction method of power supply vehicle
…… JIANG Dongnian，BA Yujiang，LI Wei （1583）

Segmentation extraction of feature points for time series pattern matching
…… LI Zhengxin，LIU Chang，WU Shihui，GUO Jiansheng （1593）

Precise orbit determination method for angle-only observation data of space debris based on angle conversion theory
…… ZHANG Yao，LIU Jing （1600）

Secondary electron multiplication of aluminum under strong vacuum electromagnetic field
…… SHANG Shengfei，YANG Xiaoning，YANG Yong，BI Yanqiang，WU Nankai，YU Lantao （1606）

Tensile fatigue properties of carbon fiber laminates in hygrothermal environments
…… XU Mingrui，ZENG Benyin，XIONG Xin，MENG Qingchun，CHENG Xiaoquan （1614）

A sparse estimation method for radar target direction with sliding-window subarray configuration in mainlobe jamming
…… ZHOU Bilei，LI Rongfeng，ZENG Li，CHEN Hui，LIU Weijian，LI Binbin （1623）

Configuration and multibody separation scheme of compact missile swarm
…… FU Jiawei，WANG Chen （1630）

Multi-objective optimization of aerodynamic layout for hypersonic reentry vehicle
…… FAN Boxuan，CHEN Guiming，CAO Yiqun （1639）

Mechanism of butterfly forward flight and prototype verification based on characteristic motion observation
…… ZHANG Yixin，WANG Xingjian，WANG Shaoping，CHI Xiaokai，DU Shaoyang （1651）

Soil moisture retrieval using Beidou GEO satellite interference signal power
…… HAN Mutian，XU Zhichao，CHANG Qing，ZHANG Bo，WANG Feng，HONG Xuebao （1661）

Aerodynamic performance analysis of supercritical airfoil with lower surface jet
…… WANG Ruochen，ZHANG Guoxin，WANG Xiangyu，MA Xiaoping （1671）

Numerical study of blood pump weaning effects on hemocompatibility of centrifugal blood pump
…… XIE Nan，TANG Yumeng，ZHANG Yan，LIU Yangwei （1680）

Aerodynamic order reduction method for elastic aircraft flight dynamics simulation
…… SHI Yan，WAN Zhiqiang，WU Zhigang，GONG Yinan （1689）

Experiment research on high-speed ultrasonic vibration milling of titanium alloy
…… HAN Xiong，SUN Zhefei，GENG Daxi，ZHANG Deyuan （1707）

Simulation analysis of reduction effect of symmetrical winding method for multi-polar fiber ring on Shupe error
…… ZHANG Dongbo，WANG Lixin，LI Can （1715）

Design and performance analysis of spatial large extension parallel mechanism
…… HE Litao，FANG Hairong，CHEN Yufei，LI Yin （1722）

Analysis of influence of ultrasonic drilling structure parameters on output characteristics
…… ZHANG Xingwang，ZHANG Ming，YU Dengyun，ZENG Ting，YIN Shen，YANG Shuai，PANG Yong （1735）

Pedestrian trajectory prediction method based on pedestrian pose
…… WANG Ruiping，SONG Xiao，CHEN Kai，GONG Kaiqi，ZHANG Junfan （1743）

Improved modeling strategy and result evaluation of steroid-induced femoral head necrosis in rabbits
…… YANG Fenghe，FENG Guiyu，LI Jinyu，NIU Xufeng （1755）

Influence of rotor/wing aerodynamic interference on performance of compound helicopters
…… YANG Kelong，HAN Dong （1761）

Weakly supervised evaluation of airport traffic situation based on metric learning
…… DU Jinghan，HU Minghua，ZHANG Weining，YIN Jianan （1772）

Influence of open water in retrieval of soil moisture by spaceborne GNSS-R
…… YANG Wentao，XU Tianhe，WANG Nazi，GAO Fan，JING Lili，HE Yunqiao （1779）

UAV information interaction topology generation considering task allocation
…… XUE Ying，HE Feng，GU Xiaoyan （1787）

Design of robust controller for single outrigger of vibration active isolation platform based on LPV
…… TANG Yucong，ZHU Qinghua，LIU Fucheng，CAO Heyang （1796）

Double-loop surrogate model for time-dependent reliability analysis based on NARX and Kriging models
…… CHANG Zeming，LI Luyi （1802）

Conflict resolution algorithms for UAV low-altitude flight based on reachable set
…… YANG Jianhang，ZHANG Fubiao，WANG Jiang （1813）

Semantic segmentation of remote sensing images based on U-shaped network combined with spatial enhance attention
…… BAO Yintu，LIU Wei，LI Runsheng，LI Qin，HU Qing （1828）

Stacking scheduling method for time-triggered messages in off-chip network
…… ZANG Guangjie，LI Qiao，WANG Tong，XIONG Huagang （1838）

Detection of debonding defect in CFRP laminates using infrared pulse thermal wave tomography
…… ZHONG Jiacen，XU Haojun，WEI Xiaolong，HAN Xinmin，CHANG Yipeng （1847）

Uniform circular multiphase modulation correlation radiometer and its sensitivity analysis
…… CHEN Ruihou，LUO Jing，HU Anyong，MIAO Jungang （1857）

Numerical study of wing gust response alleviation based on camber morphing trailing edge
…… WEI Rukai，DAI Yuting，YANG Chao，YU Siheng （1864）

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0562

软件定义卫星技术概念及发展

徐帆江[1,2]，周鑫[1,2,*]，赵军锁[1,2]，吴凤鸽[1,2]，林翊钧[1,2]，夏玉立[1]

(1. 中国科学院软件研究所，北京 100190；2. 中国科学院大学，北京 100190)

摘 要：针对在轨卫星呈现海量化、网络化、智能化发展的趋势，重点探讨了软件定义卫星技术的产生背景、概念、内容与挑战。通过历史回顾，归纳卫星系统研制的3个阶段，指出目前卫星系统正在从平台、载荷优先向算法优先演进，这一趋势推动了软件定义卫星技术的产生与发展；阐述软件定义卫星技术的概念与边界，指出研究软件定义卫星技术的目的不是研制一类新型卫星，而是利用软件定义技术实现硬件资源虚拟化、系统软件平台化、应用软件多样化，提升卫星系统的网络化与智能化；分析软件定义卫星技术的主要内容与挑战，并对未来的发展进行展望。

关 键 词：软件定义；软件定义卫星技术；算法优先；功能重构；智能卫星

中图分类号：V474; TP391

文献标志码：A **文章编号**：1001-5965(2023)07-1543-10

近年来，随着巨型低轨卫星星座、卫星互联网等计划的推进，全球卫星产业迅猛发展，在轨卫星呈现海量化、网络化、智能化趋势[1]。据忧思科学家联盟(UCS)统计，截至2021年4月，全球在轨运行卫星数量已达到4 200颗，并以每年上千颗的速度加速增长[2]。预计到2027年，仅Starlink卫星就将达到12 000颗[3]，全球卫星总数将超过2万颗。庞大的卫星系统将为飞机、船舶、车辆、个人等提供各种天基服务，产生巨大的商业价值。

商业应用如火如荼，太空领域的军事竞赛也激烈展开[4]。2018年6月，美国成立太空军，组建太空站实验室，并开展太空军事演习，不断加强对太空的军事利用[5]。美国空军、陆军均先后与太空探索技术公司(SpaceX)签署了合作协议，研究Starlink卫星[6]用于军事通信的可行性。其他国家也逐步强化太空军事计划。日本《宇宙基本法》规定太空资产要为“国家安全”服务，突破了“太空只限于和平目的”决议的禁令[7]。

然而，卫星数量的快速增长，也给卫星的研制、部署和应用带来了巨大挑战[8]。传统的“为特定任务定制卫星，为特定卫星定制载荷，为特定载荷定制软件”的卫星研制思路，导致卫星功能单一、灵活性不足、技术迭代周期长、任务满足率偏低等问题，无法适应未来卫星系统网络化、智能化、体系化发展趋势。如何加速卫星研制与技术演进，实现在轨功能重构，降低复杂卫星星座的成本，提高航天任务满足率，这些现实而严峻的问题，迫使人们重新思考卫星技术发展的重心和思路[9-11]。

从软件定义无线电(software-defined radio, SDR)到软件定义网络(software-defined network, SDN)，再到软件定义一切(software-defined anything, SDx)，软件定义已经成为一种不可逆转的趋势，在各个行业悄然兴起[12]。软件定义就是通过硬件资源标准化和功能服务软件化，在软硬件解耦与硬件资源虚拟化、标准化的基础上，充分发挥软件灵活、开放的优势，发掘系统潜力[13]。这一技术思想正在推动卫星系统从平台优先向载荷优先转变，进而向算法优先演进，软件定义技术将成为未来卫星技术发展的重要驱动力。

收稿日期：2021-09-22；**录用日期**：2021-10-18；**网络出版时间**：2021-11-02 10:31
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20211101.1556.009.html
基金项目：国家自然科学基金 (91938301,62027801)
* **通信作者**. E-mail：zhouxin@iscas.ac.cn

引用格式：*徐帆江，周鑫，赵军锁，等. 软件定义卫星技术概念及发展 [J]. 北京航空航天大学学报，2023，49（7）：1543-1552.*
XU F J，ZHOU X，ZHAO J S，et al. Conception and development of software-defined satellite technology [J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（7）：1543-1552（in Chinese）.

SpaceX公司充分利用软件定义技术的优势，实现了高速卫星互联网星座Starlink[14]、可靠的商业载人龙飞船[15]、经济高效的一箭60星与可回收火箭技术等[16]。欧洲航天局(European Space Agency, ESA)的Quantum卫星[17]和中国杨小牛院士提出的“软件星”[18-19]，则重点从软件定义无线电的角度探索通信卫星在轨功能重定义的潜力。中国科学院软件研究所牵头研制的“天智一号”卫星[20]，基于开放系统架构探索软件定义卫星技术，挖掘软件的潜力，并完成多项在轨试验，为提升卫星的功能灵活性和任务满足率提供了一种新的技术路径。

本文首先梳理了卫星系统研制的演进脉络，然后阐述了软件定义卫星技术的概念与边界，最后分析了软件定义卫星技术的主要研究内容与挑战，并对该技术的发展趋势进行展望。

1 卫星系统研制思路的演进

自1957年第一颗人造地球卫星发射以来，卫星系统对于支撑对地观测、通信保障、导航定位、空间探索等任务发挥了重要作用[21]。而任务背景的变化及需求的不断提升，促使卫星系统研制由平台优先向载荷优先转变，且正在向算法优先阶段演进。

1.1 平台优先阶段

卫星由卫星平台和有效载荷组成。其中，卫星平台部分主要包括结构系统、星务系统、能源系统、姿轨控系统、测控与数传系统等；有效载荷部分则根据卫星任务需要，选择配备雷、光、电、机等一种或多种[21]。

早期，卫星平台在整个卫星系统的研制成本中占据绝大比例，其承重、供电等能力也是有效载荷设计的重要约束。经过多年的应用验证，各大厂商均形成了谱系化产品，以满足不同的载荷质量与功率需求，如洛马公司的LM卫星平台系列(LM50/400/1000/2100)[22]、波音公司BSS-702系列(BSS-702 HP/MP/SP)[23]、国内的东方红DFH系列、CAST系列和微纳卫星平台WN系列(WN100/5000)[24]等。表1列举了主要的DFH系列卫星平台。

在平台优先阶段，卫星平台是整个卫星系统的首要约束条件，而有效载荷、业务软件等都需要适配现有的平台能力，必要时还需进行功能降级或约简。

表1 东方红DFH卫星平台

Table 1 DFH satellite platform

平台	发射质量/kg	载荷质量/kg	载荷功率/kW	寿命/a	首飞卫星与时间
DFH-3	2 320	230		8	东方红3号通信卫星,1994-11-30
DFH-3A	2 740	360	2.5	12	中星22号,2000-01-26
DFH-3B	3 800	400～450	3～4	12～15	老挝1号,2015-11-21
DFH-4	5 000～5 300	450～700	4～8	12～15	鑫诺2号,2006-10-29
DFH-4e	550～6 000	800～1 000	≥10	15	亚太6D,2020-07-09
DFH-5	6 500～9 000	1 200～2 200	4～22	12～15	实践20号,2020-01-05

1.2 载荷优先阶段

随着对地观测、通信保障等任务在分辨率、幅宽、大容量、低时延等方面的需求不断提升，极大促使了星上合成孔径雷达(synthetic aperture radar, SAR)天线、相控阵天线及可展开光学成像系统等载荷的快速发展[25]。而现有谱系化平台在尺寸、功耗等方面越来越难以满足需求，在发展新体制卫星平台的同时，围绕特定载荷制造卫星成为了主要发展趋势，卫星系统迎来载荷优先阶段，表2列举了一些典型的卫星示例。

就空间探索任务而言，作为哈勃空间望远镜的继任者，于2021年12月发射升空的詹姆斯韦伯太空望远镜(James Webb space telescope, JWST)[26]是美国国家航空航天局(National Aeronautics and Space

表2 典型载荷优先的卫星示例

Table 2 Classical payload-first satellites

序号	卫星	主要载荷	发射时间
1	哈勃空间望远镜	口径2.4 m、长度16 m的主反射镜	1990-04-24
2	詹姆斯韦伯太空望远镜	载荷质量6.2 t、口径6.5 m的主反射镜	2021-12-25
3	NROL-44	电子侦察载荷直径超100 m、展开面积超数千平方米的巨型天线	2020-12-11
4	AEHF-6	军事通信载荷调频无线电、相控阵天线、核加固等	2020-03-26
5	悟空号	粒子探测载荷塑闪阵列探测器、硅阵列探测器、BGO量能器、中子探测器	2015-12-17

Administration, NASA)、ESA、加拿大航天局(Canadian Space Agency, CSA)联合研发的新一代红外太空望远镜, 其主镜结构如图 1 所示(图片来源: https://www.nasa.gov/content/goddard/webb-telescope-image-galleries-from-nasa)。该望远镜口径达 6.5 m, 是哈勃空间望远镜[27](口径 2.4 m)的 2.7倍, 对暗光具有极高的灵敏度和出色的空间分辨率, 可观测遥远宇宙的红外线, 而这些高精密的有效载荷, 对卫星平台、发射火箭等震动因素提出了更高要求, 成为了约束卫星平台及火箭发射的主要因素。

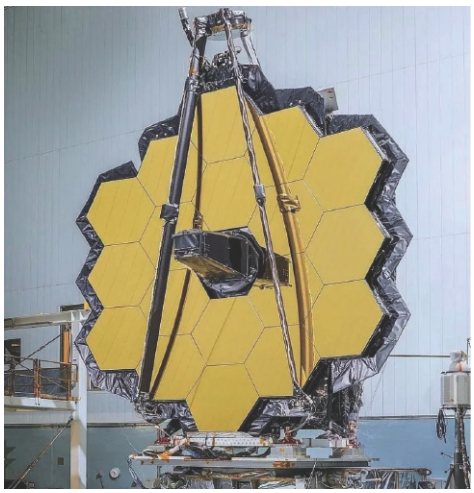


图 1　詹姆斯韦伯太空望远镜示意图

Fig. 1　Schematic of JWST

2020 年 12 月, 美国 NASA 成功使用德尔塔-Ⅳ重型火箭发射了一颗绝密卫星(编号为 NROL-44), 主要用于电子情报侦察。据悉该卫星造价超过 20 亿美元, 天线直径超过 100 m, 展开之后面积超过数千平方米[28], 给卫星平台和结构带来极大挑战。

2015 年 12 月 17 日, 中国科学院研制的暗物质粒子探测卫星“悟空号”在酒泉卫星发射中心成功发射。卫星主要通过在空间观测高能电子和伽马射线能谱, 寻找暗物质粒子存在的证据[29]。其卫星采用“以载荷为中心、以科学任务为目标的整星一体化设计”理念, 有效载荷由载荷数管与塑闪阵列等探测器构成, 载荷平台比高达 3:1。

在载荷优先阶段, 关键载荷成为整个卫星系统设计的首要约束条件, 卫星平台在必要时可以根据载荷需求进行改造, 甚至为其重新定制平台, 平台软件也需要适配该载荷。此外, 关键载荷通常采用硬件和软件一体化设计, 软件以载荷的固件形式存在, 灵活性较差, 多个载荷之间也相对独立, 载荷数据无法在轨复用, 这些都限制了硬件效能的充分发挥。

1.3　算法优先阶段

随着太空复杂环境下时敏任务保障、体系化应用等需求的提出, 卫星平台和载荷技术得以快速发展, 但仍面临存储和传输的数据量大、数据时效性较差、任务可扩展性低等问题[30]。因此, 必须大力发展在轨数据处理、自主规划、协同组网等智能算法[31], 并围绕核心算法开展卫星系统的设计与研制。

2019 年 3 月, 中国科学院的“天智一号”卫星完成了自主请求式管控试验, 通过星上算法计算, 实现自主实时定轨、高精度轨道外推、轨道超差判断和按需服务请求等功能, 将“天智一号”卫星的轨道上注周期从 24 h 延长至 72 h。该技术可有效提升中国卫星测控网的保障支撑能力[32]。

截至 2021 年 6 月, 美国 SpaceX 公司共计发射了 1 738 颗 Starlink 卫星[6], 实现全球南北纬 60°以内的无缝覆盖。该卫星系统采用动态频谱共享、自动避障等一系列先进的软件算法, 显著提高了大规模星座的应用效能和安全性[14]。

2021 年 7 月 31 日, 欧洲通信卫星公司 Eutelsat 的 Quantum 卫星成功发射入轨。该卫星采用软件无线电技术, 拥有 8 个独立运行的可重构波束, 并利用灵活的基带软件实现通信功能重构, 可以通过在轨软件升级, 在其生命周期内对不断变化的需求做出响应[17]。

与功能固定的有效载荷相比, 星载智能软件算法可以根据体系应用需求在轨调整, 在不改变现有硬件的前提下, 实现任务灵活适变、体系动态重组, 从而有效提升航天任务满足率和体系效能[33]。因此, 确立软件算法的核心地位, 而将传统的机/光/电等硬件资源作为基础保障条件, 将更能适应未来体系应用的需求。

如图 2 所示, 在算法优先阶段, 算法将成为整个卫星系统的首要约束条件。为了适应不断变化的任务需求, 星上算法需要更多的计算资源、更灵活的重构能力, 这将反过来对计算、传感等载荷的设计提出强制性约束, 并进一步对平台能力提出要求。此时, 卫星及星座的功能和灵活性都大幅提升, 在整个任务体系中, 地面端与卫星端各司其职、互相配合, 共同提高航天任务满足率。

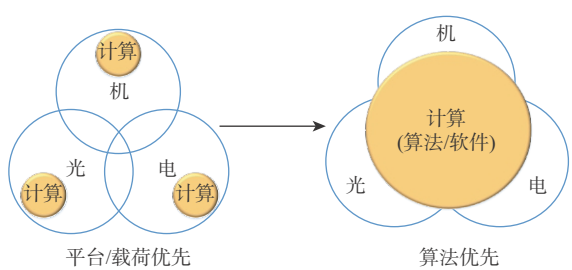


图 2　卫星平台/载荷优先向算法优先演进概念

Fig. 2　Concept of satellite evolution from platform/payload-first to algorithm-first

纵观 3 个阶段, 平台、载荷、算法始终存在、相互博弈, 博弈的核心在于卫星系统设计中谁应该被

优先保障。在发展初期，卫星技术尚不成熟，稳定、可靠的卫星平台成为首要考虑因素；随着平台技术的发展成熟，高性能载荷的地位作用更加凸显；未来，随着平台和载荷技术的成熟，多样化应用将成为核心要素，而高效、灵活的算法作为任务逻辑的载体，将逐步成为卫星设计中的首要设计约束。

当前，以美国为代表的航天大国已经实现载荷优先，正在逐渐向算法优先阶段演进。而中国仍处于平台优先向载荷优先的过渡阶段，并在算法优先方面做了一些初步探索。软件定义卫星技术正是在这样的需求背景下应运而生。

2 软件定义卫星技术的概念与边界

2.1 软件定义卫星技术的概念

“软件定义”起源于“软件无线电”。1992年，Mitola教授在无线电领域首次提出了软件定义的概念，即软件定义无线电[34]，其核心思想是尽可能早地将接收到的模拟信号数字化，以最大程度地通过软件来实现设备的各种通信功能。受当时硬件水平限制，这一设想在宽带射频天线、宽带AD/DA及高速信号处理等方面遇到了技术瓶颈，但如今都已成为现实。

在此基础上，软件定义的方法逐渐扩展到其他领域。2010年，斯坦福大学McKeown教授提出软件定义网络的概念，核心思想是将网络的控制与数据分离，根据上层应用需求，通过控制层集中对物理网络资源进行灵活配置[35]。目前，软件定义网络技术几乎成了高端网络交换机的标配，已被广泛用于Cisco、NEC、华为等国际一线品牌的网络交换机中[36]。2012年，软件巨头VMware公司提出软件定义存储（software-defined storage, SDS）的概念[37]，核心思想是将所有存储相关的控制工作都放在物理存储硬件之外的软件中完成，这一软件不是作为存储设备的固件而存在，而是运行于服务器上或者作为操作系统的一部分。软件定义存储可以管理来自不同厂商的所有物理和虚拟存储资源，并按需进行配置，因此迅速成为云计算、云存储的革命性技术。2016年，百度公司提出软件定义汽车（software-defined vehicle, SDV）[38]，其核心思想是决定未来汽车的是以人工智能为核心的软件技术，而不再是汽车的马力大小、机械性能好坏等。目前，以特斯拉为代表的新能源汽车与自动驾驶技术已经成为汽车行业的重要增长点[39]。

梅宏院士认为，软件定义的技术本质是把原先一体化的硬件设施打破，将基础硬件虚拟化并提供标准化的基本功能，通过管控软件，控制其基本功能，提供更开放、灵活、智能的管控服务[12]。倪光南院士认为，软件定义一切反映了一种技术趋势，即软件按需地赋予硬件设备各种不同的功能，但这并不意味着软件可以取代硬件，而是硬件会趋于标准化，实现高性能和低价格，而软件则趋向于个性化，实现多样性和灵活性，二者将协调使人类进入软件定义时代[40]。李培根院士认为，在“工业互联网”、“工业4.0”的发展蓝图中，软件定义将成为企业核心竞争力的战略需要，伴随着软件定义的泛化与延伸，软件将有望为物理实体定义新的功能、效能与边界[41]。可以预见，软件定义将渗透在未来各行各业的发展中，万物皆可互联，软件定义一切。

综上所述，软件定义的核心是削弱甚至消除软硬件之间的强耦合关系，最大限度发挥硬件的通用性和软件的灵活性，推动应用软件向个性化方向发展、硬件资源向标准化方向发展、系统功能向网络化与智能化方向发展。

因此，软件定义卫星技术是以星载高性能计算平台为支撑，采用开放系统架构，支持有效载荷即插即用、应用软件按需加载、系统功能快速重构、卫星系统增智赋能的相关软件定义技术的统称。

2.2 软件定义卫星技术的研究目标

研究软件定义卫星技术的目的，不是研制一类新型卫星，而是研究如何利用软件定义技术支撑智能化卫星的研制与应用，有力推动航天工业系从平台中心化向任务中心化的转变。具体而言，软件定义卫星技术，是要实现硬件资源虚拟化、系统软件平台化、应用软件多样化，用软件来定义卫星的功能、提升卫星的性能、加速卫星的生产，最大限度地发挥硬件的通用性和软件的灵活性。

硬件资源虚拟化，是指将各种实体硬件资源抽象化，打破其物理形态的不可分割性，制定组件的接口标准，泛化其访问形式，隐藏其实现细节，以便通过灵活重组、重用发挥其最大效能。

系统软件平台化，是指通过基础软件对硬件资源进行统一管控、按需分配，并通过标准化的编程接口解除上层应用软件和底层硬件资源之间的紧耦合关系，使其可以各自独立演化。

应用软件多样化，是指在成熟的平台化系统软件解决方案的基础上，应用软件不受硬件资源约束，对外提供更为灵活高效和多样化的服务。软件定义的卫星系统，将随着硬件性能的提升、算法效能的改进、应用数量的增多，逐步向智能卫星系统演变。

2.3 软件定义卫星技术的发展层次

软件定义卫星技术不是一蹴而就的，其发展层次由浅入深可以分为人在控制回路的软件缺陷修复、软件算法升级和软件功能重构，最终达到人在监督回路的软件自主演进。

1) 软件修复。指在不改变输入输出和算法核心的前提下, 通过修复软件中存在的问题缺陷, 保障卫星正常完成既定任务目标。在卫星发射入轨初期, 通过在轨测试发现卫星软件缺陷, 可通过上注更新完成软件修复。目前, 大多数新型卫星都已具备软件修复的能力。

2) 算法升级。指在保持任务目标不变的前提下, 通过算法逻辑优化、输入参数扩展、多元传感信息融合等手段, 提高算法输出的丰富性、准确性、鲁棒性, 从而更好地完成既定任务目标。初级的算法升级并不改变输入输出, 仅仅是算法逻辑的优化, 而高级的算法升级并不局限于固定的输入输出, 而是可以根据需要实现星载传感、地面遥测等数据输入的变更。为支撑算法持续升级, 应使算法软件尽量靠近后端, 而不与前端传感或控制载荷过分耦合。目前, 只有 Quantum 等极少数卫星具备较理想的算法升级能力。

3) 功能重构。指利用卫星系统现有的硬件资源, 通过软件算法或接口的组合调整与配置, 重构卫星系统的功能, 以满足新任务的要求。卫星入轨后, 用户可能会对卫星提出既定任务目标之外的功能要求, 这在传统卫星系统中是无法实现的, 但在新型卫星系统中, 可以通过功能重构实现新的任务目标, 从而不断提高在轨卫星的任务满足率。

这 3 个层次不断进行螺旋式发展迭代, 如图 3 所示。初期的软件修复—算法升级—功能重构可能仅涉及少量代码的更改, 进一步可实现软件重要模块甚至整个软件架构的更新。随着人工智能技术的发展, 这 3 个层次的发起者和执行者也会由地面控制中心逐渐转变为软件智能体本身, 卫星系统可以自主完成软件修复、算法升级和功能重构, 最终达到自主演进的目标。届时, 人将处于监督回路, 仅在必要时进行干预[42]。

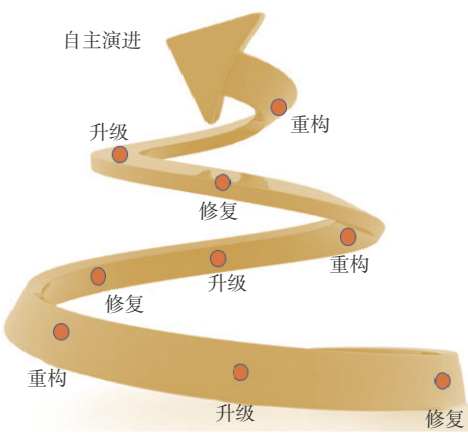


图 3　软件定义卫星技术的发展层次

Fig. 3　Development levels of software defined satellite technology

3　软件定义卫星技术的内容与挑战

软件定义卫星技术虽然名称上冠以“卫星”二字, 但并不仅限于卫星个体上的技术, 而是着眼于卫星前端、地面支撑、体系应用有机结合的全方位生态构建解决方案。如图 4 所示, 软件定义卫星技术的研究内容大体可分为 4 个方面, 分别为系统架构技术、共性基础支撑技术、智能应用技术, 以及上述技术的标准和规范制定等。

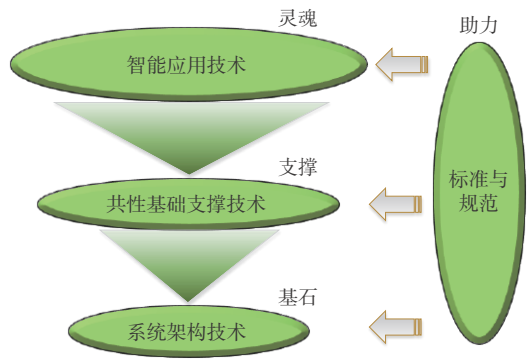


图 4　软件定义卫星技术的主要研究内容

Fig. 4　Contents of software-defined satellite technology

3.1　系统架构技术

软件定义卫星, 架构定义软件, 良好的系统架构是软件定义卫星技术的基石, 其明确了卫星系统中的部件要素及其相互联系。如图 5 所示, 软件定义卫星采用开放的系统架构, 以高性能星载计算平台为核心, 通过接入不同的有效载荷(传感器、执行器、通信载荷等), 加载不同的软件, 快速重构出具有不同功能的卫星系统。借助高性能星载计算平台强大的计算能力, 将有效载荷的部分功能通过软件实现, 并最终实现有效载荷的虚拟化。

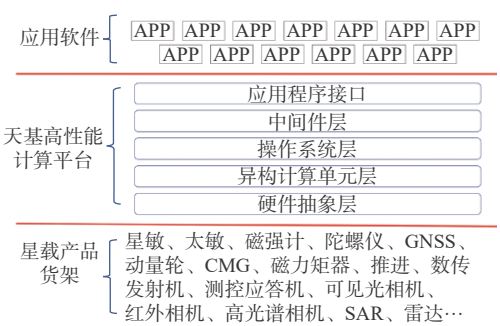


图 5　软件定义卫星系统架构

Fig. 5　System framework of software-defined satellite

对软件定义卫星系统架构技术的研究, 具体可分为 3 个方面:

1) 软硬件解耦的功能组件模块化技术。将卫星平台/载荷的封闭系统分解为一系列模块化、组件化的松耦合软硬件单元, 并根据传输需求制定标

准的、实时高速数据总线接口，完成组件间的信息交互，从而提高卫星部件的兼容性，实现技术的低成本复用和独立演进。美国洛马公司尝试将这类技术用于构建快速响应、低成本的情报监视侦查(intelligence, surveillance, and reconnaissance, ISR)卫星[43]；美国国防高级研究计划局(Defense Advanced Research Projects Agency, DARPA)将其用于设计任务管理系统Pit Boss[44]，以快速集成军用载荷和自主软件等。

2）异构高性能星载计算技术。将CPU、GPU、DSP、FPGA等不同种类的分布式计算单元“强强联合”，形成逻辑上一体化的高性能计算资源池，并可对资源进行弹性分配，从而为星载智能应用的部署和运行提供通用的、可扩展的、高效灵活的星上计算环境。美国NASA曾将其异构计算系统SpaceCube用在超光谱仪、SAR等新科学任务中，在降低功耗成本的同时提升了10～100倍计算能力[45]。

3）灵活开放的有效载荷技术。与载荷优先阶段不同，在算法优先阶段，有效载荷除了应具备更灵活的功能和更强的性能之外，还需要开放更多底层的、基本的功能服务接口，从而通过算法软件实现快速灵活的功能重构和任务响应，充分发掘系统潜力。欧洲Quantum卫星上的软件无线电载荷，可以在基带处理和天线波束2个层面对通信体制进行重新定义，具备极强的灵活性和新需求响应能力[17]。

3.2 共性基础支撑技术

共性基础支撑技术聚焦于软件定义卫星技术中的软件开发、测试、加固、网络通信、数据储备等基础性支撑技术，具体包括：

1）软硬件协同的系统可靠性与安全性加固技术。为弥补恶劣外太空环境带来的器件可靠性威胁，降低开放系统架构所导致的安全性隐患，需要软硬件协同实现可靠容错、漏洞修复、自主恢复、自主故障诊断与隔离等，保障卫星系统持续稳定与安全。目前，航天系统中主要采用冗余备份、错误检查与纠正(error-checking-and-correcting, ECC)内存、宇航级元器件、单粒子闩锁防护等硬件手段保障可靠性，采用鉴权、日志、加密等软件手段保障安全性。未来除了应该进一步完善现有技术手段外，还可以从软硬件协同的角度研究可靠性和安全性加固技术，针对体系结构和空间环境的特点，通过微重启、软件抗衰、基于容器的软件分区隔离等软件容错技术，结合总剂量防护等硬件加固手段，降低卫星的费效比。美国NASA曾尝试将图6所示(图片来源：https://spacenews.com/hpe-supercomputer/)的惠普商用超级计算机送入国际空间站，并借助软件实现功耗调整和错误恢复等可靠性加固[46]，有望大幅度降低卫星硬件成本；中国“天智一号”卫星采用通用计算平台，通过上注“抗量子攻击密钥交换软件”，保障星地数据通信安全[32]。



图6 惠普星载超级计算机原型

Fig. 6 HP Space-borne computer prototype

2）星上智能算法软件敏捷开发技术。面向广大航天算法软件开发人员，提供智能算法从设计开发到迁移部署的整套解决方案，加快算法演进速度。目前，国内外行业巨头均推出了面向神经网络应用的智能软件开发框架和云平台[47]，而卫星平台的异构性、软件的低兼容性和组件的低复用性，使得专门针对航天应用的智能算法开发平台尚是一片空白。因此，亟需针对航天智能应用特点，突破星载智能算法模型构建、模型自动训练调优、星地模型协同演进、星载异构计算平台的迁移部署等技术，形成星上智能算法软件构建、训练、迁移部署一体化平台。

3）自动化系统测试评估技术。卫星星座的动态性、星载软件的多样性和空间环境的复杂性，给星载软件正确性、完整性、安全性的鉴定评估带来了多方挑战。传统针对单个卫星、特定任务的以人为主的测评方法，在测评目标和手段上都将无法适应新的应用场景，研究高效、客观、精确和自动化的测评支撑体系，实现多层级全方位自动化测试，对提高测评效率具有重要意义。基于DevOps模式的云端软件测试[48]，以及数字孪生[49]等技术，对研发星载软件自动化测评工具具有重要借鉴意义。

4）大型星座组网与高速传输技术。面向大型星座天地组网需求，如何构建高速、可靠、灵活的天地一体化通信网络，满足全球海量、分布式、差异化的通信业务需求至关重要[50]。卫星组网与传输技术，需要针对大型星座系统下的复杂网络和电磁环境，充分发挥协议软件灵活可扩展的优势，实现面向特定任务/业务需求的灵活高效资源分配、协调与复用。目前常用的CCSDS、DVB-S、TCP/IP等组网协议体系，主要面向深空和互联网等场景[51]，而Starlink系统采用的私有组网传输技术，尚处于

探索和试验阶段。因此, 建立具有标准化接口的协议将是未来几年国内外研究的重点。

5) 多源多模态高精细样本库构建技术。基于多载荷类型, 面向多应用目标, 建立载荷源数据和样本库标准规范, 实现对多元多模态源数据和样本库的增广、索引和管理, 构建高精细样本库, 为智能算法模型的训练提供数据支撑。同时, 利用小样本学习、迁移学习等数据驱动方法, 充分挖掘数据内蕴含的知识和数据间的关联性, 以提高样本利用率、降低算法模型的构建门槛。目前, 支持智能算法模型开发训练的样本库多针对地面应用场景, 专门针对航天智能应用的样本库较少[52]。

3.3 智能应用技术

智能应用是软件定义卫星技术的灵魂。在巨型星座或深空环境下, 地面站难以实时大范围地参与到各卫星节点的信息处理回路中[53], 因此星载软件是感知环境、规划运行、控制行动的主体, 航天任务满足率受算法智能化水平的直接影响。智能应用技术具体可分为感知、控制、规划 3 个层面, 每个层面又包括很多应用场景。

1) 智能感知。智能感知是要借助星上强大计算能力, 实现面向雷达、光学、无线电等多种载荷的数据处理及智能融合算法[54-55], 提高感知时敏性, 降低星地通信传输压力。然而, 受限于星上的计算能力, 目前星上感知处理集中在预处理、目标检测等较少领域, 将来的智能感知是计算机科学、数据科学和认知科学的融合, 采用人工智能技术实现目标提取、目标认知、变化检测、热点事件检测等多领域应用。“天智一号”卫星通过星上实时智能云检测及图像质量判读, 可实现无效积云数据的预先剔除, 将下传效率提升 72%, 初步窥探了智能感知的效用。

2) 智能控制。卫星控制不仅依赖动量轮、推进器、机械臂等物理执行机构, 更需要算法的支撑。目前星上的姿控、温控等基本控制多以经典的 PID 控制算法为主, 更加高层次的任务运行则需地面控制中心通过测控链路进行管控。其中, 经典 PID 控制系统属于预置式控制系统, 对复杂、时变环境的适应能力较差; 地面遥控方式更是增加了地面运维压力, 并非常依赖测控链路的稳定性, 具有严重的安全隐患。因此, 不仅需要研究模糊系统、强化学习等智能控制理论, 在传统的姿轨控、温控、能控等方面实现自主调节[56], 还需要研究环境认知与交互控制技术, 实现更加智能化的行为, 如自我故障修复、自主防卫、太空碎片清理等[57-58]。

3) 智能规划运行。卫星的规划运行, 是在满足卫星任务约束的情况下, 为给定的一组任务生成卫星的远程控制序列, 从而达到高质量完成用户请求处理的目标。当前任务执行的控制序列主要在地面生成, 通过构建任务规划的模型并对模型中的优化问题进行求解, 从而生成卫星任务规划结果并上传至星上执行。未来的智能规划, 将根据地面站制定的宏观任务目标, 由卫星个体甚至星座整体对自身的任务序列做出规划[59], 并根据实际情况及时调整, 降低卫星运行对地面站的依赖, 缩短突发状况下的应急响应时间。

3.4 标准与规范制定

标准与规范是软件定义卫星技术发展的助力, 其以行业标准的形式统一了技术思想和框架, 降低了用户参与的门槛, 可以有力推动软件定义卫星技术的产业化。软件定义卫星技术的标准与规范涉及系统架构、共性基础支撑、智能应用等所有技术方向。在系统架构方面, 需要定义卫星的功能组件分解、统一总线接口规范、载荷功能与接口标准等; 在共性基础支撑方面, 需要制定软件加固算法标准、软件开发、测试与集成规范、通信协议规范等; 在智能应用方面, 需要对智能感知、运行、任务规划等任务进行细化, 在各自的子领域建立标准算法规范。

目前, 尚未形成关于软件定义卫星技术标准的行业共识, 但相关工作早已展开。2017 年 11 月, 软件定义卫星技术联盟成立标准化工作委员会, 致力于推动软件定义卫星技术相关标准规范的制定。近年来, 联盟在硬件功能抽象、载荷接口规范、软件开发接口、统一数据总线等方面做了很多积极探索。

4 结 束 语

航天任务需求的不断提升, 促使卫星系统研制逐渐由平台优先、载荷优先向算法优先转变。软件定义卫星技术顺应了这一历史趋势, 通过各种技术手段实现硬件资源虚拟化、系统软件平台化、应用软件多样化, 最大限度发挥硬件的通用性和软件的灵活性, 其思想已经渗入、并将更加深入地影响卫星系统的设计、研制、生产、测试、应用、维护等全生命周期。

当前, 中国正在从航天大国向航天强国迈进, “天地一体化信息网络”、“深空探测及空间飞行器在轨服务与维护系统”等国家重大工程对卫星系统网络化、智能化有迫切需求。未来应加速推进软件定义卫星技术研究, 重点构建智能卫星系统的研制与应用生态, 汇聚更多的尖端人才, 助力我国航天技术实现弯道超车。

参考文献（References）

[1] 齐心, 周思卓, 林屹立. 美军小卫星“三化”前沿技术发展探析[J]. 国际太空, 2021(3): 54-59.
QI X, ZHOU S Z, LIN Y L. An analysis of the frontier technology development of the US army's small satellites[J]. Space International, 2021(3): 54-59(in Chinese).

[2] UCS. UCS satellite database[EB/OL]. [2021-09-01]. https://www.ucsusa.org/resources/satellite-database.

[3] 王迪, 骆盛, 毛锦, 等. Starlink卫星系统技术概要[J]. 航天电子对抗, 2020, 36(5): 51-56.
WANG D, LUO S, MAO J, et al. Overview of Starlink satellite system technology[J]. Aerospace Electronic Warfare, 2020, 36(5): 51-56(in Chinese).

[4] 郭丽红, 蔡润斌, 李臻. 2020年美国太空军事力量发展综述[J]. 国际太空, 2021(5): 43-47.
GUO L H, CAI R B, LI Z. Overview of U. S. space military force development in 2020[J]. Space International, 2021(5): 43-47(in Chinese).

[5] 司耀锋, 吴林, 郝媛媛, 等. 美军作战快速响应太空计划发展综述[J]. 国际太空, 2012(1): 33-41.
SI Y F, WU L, HAO Y Y, et al. Overview of the development of the U. S. military's operational rapid response space program[J]. Space International, 2012(1): 33-41(in Chinese).

[6] SpaceX. Starlink[EB/OL]. [2021-09-01]. https://www.starlink.com.

[7] 惠仲阳. 范唯唯. 日本发布新版《宇宙基本计划》[J]. 空间科学学报, 2020, 40(6): 968.
HUI Z Y, FANG W W. Japan releases new version of “Universal Basic Plan”[J]. Chinese Journal of Space Science, 2020, 40(6): 968(in Chinese).

[8] 闵士权. 我国天基综合信息网构想[J]. 航天器工程, 2013, 22(5): 1-14.
MIN S Q. An idea of China's space-based integrated information network[J] Spacecraft Engineering, 2013, 22(5): 1-14 (in Chinese).

[9] 赵军锁, 吴凤鸽, 刘光明. 软件定义卫星技术发展与展望[J]. 卫星与网络, 2017(12): 46-50.
ZHAO J S, WU F G, LIU G M. Development and prospect of software defined satellite technology[J]. Satellite & Network, 2017(12): 46-50(in Chinese).

[10] 周渊, 张洵颖, 智永锋, 等. 软件定义卫星有效载荷技术研究[J]. 西北工业大学学报, 2020, 38(S1): 96-101.
ZHOU Y, ZHANG X Y, ZHI Y F, et al. Research on software defined satellite payload technology[J]. Journal of Northwestern Polytechnical University, 2020, 38(S1): 96-101(in Chinese).

[11] 吴启星. 软件定义卫星研究现状与技术发展展望[J]. 中国电子科学研究院学报, 2021, 16(4): 333-337.
WU Q X. State of the art and development analysis of software defined satellites[J]. Journal of China Academy of Electronics and Information Technology, 2021, 16(4): 333-337(in Chinese).

[12] 梅宏. 万物皆可互联, 一切均可编程[J]. 方圆, 2018(12): 58-59.
MEI H. Everything can be connected, everything can be programmable[J]. Fangyuan Magazine, 2018(12): 58-59(in Chinese).

[13] 梅宏. 软件定义一切: 机遇和挑战 [EB/OL]. [2021-09-01]. https://www.fx361.com/page/2018/0112/2687238.shtml.
MEI H. Software defines everything: Opportunities and challenges [EB/OL]. [2021-09-01]. https://www.fx361.com/page/2018/0112/2687238.shtml(in Chinese).

[14] 刘帅军, 徐帆江, 刘立祥, 等. Starlink第二代系统介绍[J]. 卫星与网络, 2020(12): 62-65.
LIU S J, XU F J, LIU L X, et al. Starlink second generation system introduction[J]. Satellite & Network, 2020(12): 62-65(in Chinese).

[15] 雷江利, 牟金岗, 赵广秀, 等. 新一代载人飞船试验船回收着陆系统任务特点分析[J]. 国际太空, 2020(9): 8-12.
LEI J L, MOU J G, ZHAO G X, et al. Analysis of mission characteristics of the new generation manned spacecraft test ship recovery and landing system[J]. Space International, 2020(9): 8-12(in Chinese).

[16] 王芳, 程洪玮, 彭博. “猎鹰9”运载火箭海上平台成功回收的分析及启示[J]. 装备学院学报, 2016, 27(6): 69-74.
WANG F, CHENG H W, PENG B. Analysis and enlightenment of successful recovery of “Falcon 9” carrier rocket on offshore platform[J]. Journal of Equipment Academy, 2016, 27(6): 69-74(in Chinese).

[17] Eutelsat. Quantum [EB/OL]. [2021-09-01]. https://www.eutelsat.com.

[18] 杨小牛. “软件星”概念研究[J]. 电子对抗, 2002(1): 1-5.
YANG X N. Research on the concept of software satellite[J]. Electronic Warfare, 2002(1): 1-5(in Chinese).

[19] 杨小牛. 基于“软件星”的综合一体化空间信息系统[J]. 中国电子科学研究院学报, 2004(4): 15-22.
YANG X N. Comprehensive integrated spatial information system based on software satellite[J]. Journal of China Academy of Electronics and Information Technology, 2004(4): 15-22(in Chinese).

[20] 杜莹. 软件定义卫星技术开启卫星智能新纪元[J]. 科技纵览, 2019(8): 59-61.
DU Y. Software-defined satellite technology opens a new era of satellite intelligence[J]. IEEE Spectrum, 2019(8): 59-61(in Chinese).

[21] 徐福祥. 卫星工程概论[M]. 北京: 宇航出版社, 2003.
XU F X. Introduction to satellite engineering[M]. Beijing: China Astronautic Publishing House, 2003(in Chinese).

[22] Lockheed Martin. SmartSat[EB/OL]. [2021-09-01]. https://www.lockheedmartin.com/en-us/products/satellite.html.

[23] Boeing. BOEING 702X satellites[EB/OL]. [2021-09-01]. http://www.boeing.com/space/boeing-satellite-family/702X/index.page.

[24] 中国空间技术研究院. 卫星平台[EB/OL]. [2021-09-01]. https://www.cast.cn/channel/1539.
Chinese Academy of Space Technology. Satellite platform[EB/OL]. [2021-09-01]. https://www.cast.cn/channel/1539(in Chinese).

[25] 韩天龙, 杜刚, 陆宏伟, 等. 国外通信卫星公用平台发展趋势及启示[J]. 航天工业管理, 2015(3): 35-38.
HAN T L, DU G, LU H W, el al. Development trend and enlightenment of foreign communication satellite public platform[J]. Aerospace Industry Management, 2015(3): 35-38(in Chinese).

[26] GARDNERJ P, MATHER J C, CLAMPIN M, et al. The James Webb space telescope[J]. Space Science Reviews, 2006, 123(4): 485-606.

[27] SCIVILLE N, ABRAHAM R G, AUSSEL H, et al. COSMOS: Hubble space telescope observations[J]. The Astrophysical Journal Supplement Series, 2007, 172(1): 38.

[28] Aerospace. NROL-44 launch: Delivering on mission success [EB/OL]. [2021-09-01]. https://aerospace.org/article/nrol-44-launch-

delivering-mission-success.

[29] 常进, 范一中. 悟空号: 暗物质粒子的探索者[J]. 科学, 2018, 70(3): 18-20.
CHANG J, FAN Y Z. Dark matter particle explorer (DAMPE)[J]. Science, 2018, 70(3): 18-20(in Chinese).

[30] 李德仁, 王密, 沈欣, 等. 从对地观测卫星到对地观测脑[J]. 武汉大学学报 (信息科学版), 2017, 42(2): 143-149.
LI D R, WANG M, SHEN X, et al. From earth observation satellite to earth observation brain[J]. Geomatics and Information Science of Wuhan University, 2017, 42(2): 143-149(in Chinese).

[31] 邓宝松, 孟志鹏, 义余江, 等. 对地观测卫星任务规划研究[J]. 计算机测量与控制, 2019, 27(11): 130-139.
DENG B S, MENG Z P, YI Y J, el al. Research of task scheduling of earth observing satellites[J]. Computer Measurement & Control, 2019, 27(11): 130-139(in Chinese).

[32] 李莉. 软件定义卫星, 智能引领航天[J]. 中国科技奖励, 2019(1): 26-28.
LI L. Software-defined satellites, intelligence leads aerospace[J]. China Awards for Science and Technology, 2019(1): 26-28(in Chinese).

[33] 赵军锁, 吴凤鸽, 刘光明, 等. 发展软件定义卫星的总体思路与技术实践[J]. 卫星与网络, 2018(4): 44-49.
ZHAO J S, WU F G, LIU G M, et al. General ideas and technical practices for the development of software-defined satellites[J]. Satellite & Network, 2018(4): 44-49(in Chinese).

[34] MITOLA J I. Software radio: Survey, critical evaluation and future direction[C]//Telesystems Conference. [S.l.]:[s.n.], 1992.

[35] LANTZ B, HELLER B, MCKEOWN N. A network in a laptop: Rapid prototyping for software-defined networks[C]//Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks. New York: ACM, 2010: 1-6.

[36] NUNES B A A, MENDONCA M, NGUYEN X N, et al. A survey of software-defined networking: Past, present, and future of programmable networks[J]. IEEE Communications surveys & tutorials, 2014, 16(3): 1617-1634.

[37] HOSKEN M. VWware Software-defined storage: A design guide to the policy-driven, software-defined storage era[M]. New York: John Wiley & Sons, 2016.

[38] 孟天闯, 李佳幸, 黄晋, 等. 软件定义汽车技术体系的研究[J]. 汽车工程, 2021, 43(4): 459-468.
MENG T C, LI J X, HUANG J, et al. Study on technical system of software defined vehicles[J]. Automotive Engineering, 2021, 43(4): 459-468(in Chinese).

[39] 袁博. 新能源汽车技术发展与趋势综述[J]. 现代商贸工业, 2018, 39(35): 12-16.
YUAN B. Overview of new energy vehicle technology development and trends[J]. Modern Business Trade Industry, 2018, 39(35): 12-16(in Chinese).

[40] 李越, 马珉. 自主可控的“软件定义”时代——访中国工程院院士倪光南[J]. 高科技与产业化, 2017(12): 33-36.
LI Y, MA M. Autonomous and controllable “software-defined” era —Interview with Ni Guangnan, academician of the Chinese Academy of Engineering[J]. High-Technology & Commercialization, 2017(12): 33-36(in Chinese).

[41] 李培根. 闲话“软件定义制造”[EB/OL]. [2021-09-01]. https://mp.weixin.qq.com/s/pbdRkGAQIXDijvq5U9NR4g.
LI P G. Talking about “software-defined manufacturing”[EB/OL]. [2021-09-01]. https://mp.weixin.qq.com/s/pbdRkGAQIXDijvq5U9NR4g(in Chinese).

[42] 胡晓峰. 系统集成与系统综合集成[J]. 测控技术, 1999, 18(9): 11-13.
HU X F. System integration and system meta synthesis[J]. Measurement & Control Technology, 1999, 18(9): 11-13(in Chinese).

[43] Lockheed Martin. New tactical ISR satellites provide global, persistent support for warfighters[EB/OL]. [2021-09-01]. https://news.lockheedmartin.com/tactical-isr-satellites.

[44] HU G, LI X, ZHANG G. Research on space mosaic reconfigurable system based on MAS[C]//2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC). Piscataway: IEEE Press, 2021, 4: 1977-1981.

[45] PETRICK D, ESPINOSA D, RIPLEY R, et al. Adapting the reconfigurable spacecube processing system for multiple mission applications[C]//2014 IEEE Aerospace Conference. Piscataway: IEEE Press, 2014: 1-20.

[46] Sebastian Moss. HPE’s spaceborne computer survives a year on the International Space Station[EB/OL]. [2021-09-01]. https://www.datacenterdynamics.com/en/analysis/hpes-spaceborne-computer-survives-year-iss/.

[47] 舒娜, 刘波, 林伟伟, 等. 分布式机器学习平台与算法综述[J]. 计算机科学, 2019, 46(3): 9-18.
SHU N, LIU B, LIN W W, et al. Survey of distributed machine learning platforms and algorithms[J]. Computer Science, 2019, 46(3): 9-18(in Chinese).

[48] 顾炯炯. 云计算架构技术与实践[M]. 北京: 清华大学出版社, 2014.
GU J J. Cloud computing architecture technology and practice[M]. Beijing: Tsinghua University Press, 2014(in Chinese).

[49] 李德仁. 数字孪生城市: 智慧城市建设的新高度[J]. 中国勘察设计, 2020(10): 13-14.
LI D R. Digital twin city: New height of smart city construction[J]. China Engineering Consulting, 2020(10): 13-14(in Chinese).

[50] 赵秋艳, 胡朝斌, 陈川, 等. 低轨大规模星座的机遇与挑战[J]. 空间碎片研究, 2020, 20(1): 1-9.
ZHAO Q Y, HU C B, CHEN C, et al. Opportunities and challenges of large-scale LEO constellation[J]. Space Debris Research, 2020, 20(1): 1-9(in Chinese).

[51] 王翔, 申志伟, 朱肖曼, 等. 卫星互联网组网技术研究[J]. 信息通信技术, 2021, 15(2): 36-43.
WANG X, SHEN Z W, ZHU X M, et al. Research on networking technology of satellite internet[J]. Information and Communications Technologies, 2021, 15(2): 36-43(in Chinese).

[52] 唐杰. 浅谈人工智能的下一个十年[J]. 智能系统学报, 2020, 15(1): 187-192.
TANG J. On the next decade of artificial intelligence[J]. CAAI Transactions on Intelligent Systems, 2020, 15(1): 187-192(in Chinese).

[53] 向尚, 陈盈果, 李国梁, 等. 卫星自主与协同任务调度规划综述[J]. 自动化学报, 2019, 45(2): 252-264.
XIANG S, CHEN Y G, LI G L, et al. Review on satellite autonomous and collaborative task scheduling planning[J]. Acta Automa-

tica Sinica, 2019, 45(2): 252-264(in Chinese).

[54] 刘扬阳, 吕群波, 谭政, 等. 基于软件定义微纳卫星的多模式计算光学成像技术[J]. 北京航空航天大学学报, 2018, 44(12): 2463-2469.
LIU Y Y, LYU Q B, TAN Z, et al. Multi-mode computational optical imaging technology based on software-defined micro-nano satellite[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(12): 2463-2469(in Chinese).

[55] 王密, 杨芳. 智能遥感卫星与遥感影像实时服务[J]. 测绘学报, 2019, 48(12): 1586-1594.
WANG M, YANG F. Intelligent remote sensing and remote sensing image real-time service[J]. Acta Geodaetica et Cartographica Sinica, 2019, 48(12): 1586-1594(in Chinese).

[56] 许轲, 吴凤鸽, 赵军锁. 基于深度强化学习的软件定义卫星姿态控制算法[J]. 北京航空航天大学学报, 2018, 44(12): 2651-2659.
XU K, WU F G, ZHAO J S. Software defined satellite attitude control algorithm based on deep reinforcement learning[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(12): 2651-2659(in Chinese).

[57] 石乾乾, 张艳, 赵鹏, 等. 一种高效利用天基激光能量清除空间碎片的方法[J]. 北京航空航天大学学报, 2018, 44(12): 2621-2627.
SHI Q Q, ZHANG Y, ZHAO P, et al. A method of efficiently using space-based laser energy to remove space debris[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(12): 2621-2627(in Chinese).

[58] 梁斌, 杜晓东, 李成, 等. 空间机器人非合作航天器在轨服务研究进展[J]. 机器人, 2012, 34(2): 242-256.
LIANG B, DU X D, LI C, et al. Advances in space robot on-orbit servicing for non-cooperative spacecraft[J]. Robot, 2012, 34(2): 242-256(in Chinese).

[59] 杜永浩, 王凌, 邢立宁. 空天无人系统智能规划技术综述[J]. 系统工程学报, 2020, 35(3): 416-432.
DU Y H, WANG L, XING L N. Intelligent planning technologies for unmanned aerospace system: A literature review[J]. Journal of Systems Engineering, 2020, 35(3): 416-432(in Chinese).

Conception and development of software-defined satellite technology

XU Fanjiang[1, 2], ZHOU Xin[1, 2, *], ZHAO Junsuo[1, 2], WU Fengge[1, 2], LIN Yijun[1, 2], XIA Yuli[1]

(1. Institute of Software, Chinese Academy of Sciences, Beijing 100190, China;
2. University of Chinese Academy of Sciences, Beijing 100190, China)

Abstract: Focusing on the trend of massive, networked, and intelligent development of in-orbit satellites, this paper explores the background, concept, and trends of software-defined satellite technology as well as the challenges it is faced today. First, it is presented that the satellite system is evolving from platform and load-first to algorithm-first, which promotes the emergence and development of software-defined satellite technology. Then, the concept and boundary of software-defined satellite technology are proposed, with the belief that the purpose of software-defined satellite technology is not to develop a new type of satellite, but to realize the hardware resource virtualization, system software platformization and application software diversification for various intelligent satellites. Next, the main contents of software-based satellite technology are analyzed along with the challenges to this technology. Finally, the prospect for this technology is given.

Keywords: software-defined; software-defined satellite technology; algorithm-first; function reconstruction; smart satellite

Received: 2021-09-22; **Accepted**: 2021-10-18; **Published Online**: 2021-11-02 10:31
URL: kns.cnki.net/kcms/detail/11.2625.V.20211101.1556.009.html
Foundation items: National Natural Science Foundation of China (91938301,62027801)
* **Corresponding author**. E-mail: zhouxin@iscas.ac.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0483

民用飞机迫降后不利姿态应急撤离实验研究

封文春[1,*]，张桐[2]，李伟[1]

(1. 航空工业第一飞机设计研究院 机电系统设计研究所，西安 710089；2. 北京航空航天大学 宇航学院，北京 100191)

摘 要：为研究飞机迫降后不利姿态对应急撤离的影响，基于动态客舱应急撤离实验环境，开展了200多人次的个体及群体应急撤离实验，按照自由撤离模式和快速撤离模式，分析了不利姿态下座椅段、过道区域个体撤离速度的变化规律，建立了个体平均速度衰减比模型。实验表明：迫降后的不利姿态对群体撤离时间、人员密度和前向距离均有一定影响。横滚姿态下，群体撤离速度衰减比在5°时为0.948，10°时为0.859，随横滚角增大而减小；俯仰姿态下，在−5°俯角下群体中个体撤离速度均呈现出加速的效果，末位速度达到1.35 m/s，次序越靠后撤离速度衰减比越大，−10°时，群体中个体运动呈现两面性，撤离速度与次序呈反比。通过实验揭示了迫降后不利姿态应急撤离客舱内个体运动及群体撤离的变化规律，为民用飞机应急撤离仿真提供了重要参考数据。

关 键 词：应急撤离；迫降；不利姿态；应急撤离实验；撤离速度

中图分类号：V233^{+}.2; TP319.9

文献标志码：A **文章编号**：1001-5965(2023)07-1553-10

应急撤离是民航客机发生紧急情况迫降后乘客逃生的重要手段，适航规章明确要求对客座数大于44座的飞机，需通过全尺寸地面应急撤离演示验证实验表明所有人员能够在90 s内安全撤离至地面[1]。国内外学者开展了大量的飞机应急撤离研究工作，但大部分都是对飞机处于水平姿态(适航所要求的应急撤离实验也是飞机处于起落架放下的水平姿态)时的应急撤离研究，且大部分为数值仿真或数据统计[2-4]。在实际发生的事故中，飞机在迫降过程中可能由于起落架故障或迫降时的冲击使得飞机迫降后处于不利姿态，即存在一定的倾斜角度。不利姿态环境会对人员的运动产生影响[5]，进而影响撤离效率。2019年5月俄罗斯航空公司一架SSJ100客机发生事故，在迫降后起火并在飞机处于异常姿态下应急撤离，55 s内仅37人撤离，事故造成41人死亡[6]。研究不利姿态对应急撤离造成的影响是很有必要的[7-8]。

单通道飞机是当前民航客机的主流，其客舱空间狭小且人员密集，是一种复杂的疏散场景，同时“90 s应急撤离原则”又对其撤离效率有极高的要求[9]。个体行为是影响撤离效率的重要因素之一，在应急撤离过程中个体的撤离速度主要受2个方面因素影响：①客舱复杂的撤离环境及撤离设备，如飞机倾斜方向及角度、狭窄的座椅区域、滑梯等；②撤离人群中其他人员及其行为。不利状态下的运动及撤离过程，船舶领域大量学者进行了研究，但受限于实验条件，大都局限于各倾斜状态下单通道内的单人或单列行人运动实验研究，较少对复杂场景中的单人或多人撤离过程进行研究[10-12]。对于群体行为的研究，目前仅是通过局部场景对单独的特定结构的撤离性能进行了研究，如十字路口、T字路口、座椅区域群体人员运动规律等[13-16]。由于客舱环境是上述单独特定结构的综合体，单独结构的结论无法类比客舱复杂结构对人员撤离产生

收稿日期：2021-08-23；**录用日期**：2021-10-22；**网络出版时间**：2021-12-27 14:40
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20211224.1458.001.html
基金项目：工信部民机预研专项 (MJ-2016-F-11)
* **通信作者**. E-mail：chunfeng@buaa.edu.cn

引用格式：封文春，张桐，李伟. 民用飞机迫降后不利姿态应急撤离实验研究 [J]. 北京航空航天大学学报，2023，49（7）：1553-1562.
FENG W C，ZHANG T，LI W. Experimental study on adverse attitude emergency evacuation of civil aircraft after crash landing [J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（7）：1553-1562（in Chinese）.

的综合影响。

基于航空工业第一飞机设计研究院所建设的动态客舱应急撤离实验环境，本文开展了不利姿态客舱应急撤离实验研究，分析在客舱倾斜状态下，单独个体撤离速度随俯仰角和横滚角的变化关系及群体撤离过程受俯仰角和横滚角产生的影响，为民用飞机客舱安全设计和应急撤离仿真提供基础支撑。

1 实验环境

1.1 实验平台

动态客舱应急撤离实验环境的结构如图 1 所示，实物环境如图 2 所示。实验环境由多自由度运动平台、客舱环境、照明装置、监控装置、滑梯等组成，客舱环境为全尺寸单通道 3-3 布局模拟舱，如图 3 所示。迫降后的不利姿态通过多自由度运动平台调节不同支柱实现俯仰、横滚及二者组合的不利迫降姿态。实验环境可实现俯仰角±15°和横滚角±15°调整。

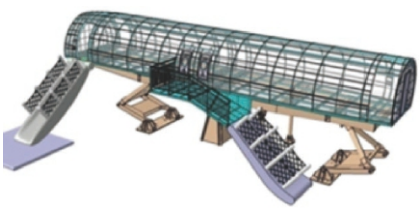


图 1 实验环境结构

Fig. 1 Structure of test environment

1.2 实验区设置

模拟客舱如图 3 所示，单通道 3-3 布局，共包含 18 排座椅，座椅间距 810 mm，腿部空间 300 mm，单边座椅宽度 1 500 mm。过道宽 510 mm，总长 13.5 m，其中一段长 4.35 m 的部分为本文实验中的测试段，两端用黄色胶布标记。过道前方包含前舱门的隔间区域为末段区域。实验中分别使用 3 台摄像机，记录人员在起始段及过道前段、测试段和末段的实验过程，具体布置及尺寸如图 4 所示。下



图 2 实验环境实物图

Fig. 2 Picture of test environment

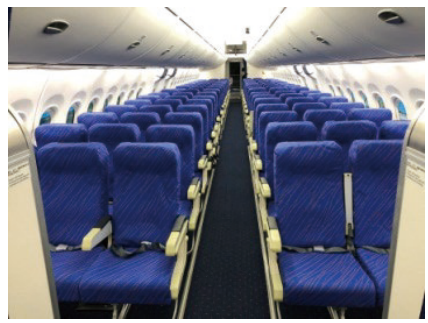


图 3 客舱内部环境

Fig. 3 Internal environment of cabin

文中的座椅区与起始段对应，直线区是指整个过道区域，包括过道前段、测试段和末段，全程是座椅区与直线区的综合。

本文实验主要目的是分析不利姿态对舱内人员撤离速度的影响，同时为了避免滑下滑梯带来其他影响，在实验中，实验人员不从滑梯撤离。

1.3 实验工况

1.3.1 个体撤离实验工况

对于民用客机迫降后可能出现的倾斜角度，可根据客机的总体外形及起落架的压缩特性计算可能的倾斜角度[17]。以 A330-200 大型客机为例，计算所得各种不利姿态角度如表 1 所示。

从表 1 中数据可以看出，客机迫降后可能出现的俯仰角及横滚角范围均在±15°之间。在预先实验中，当实验环境的仰角达到 15°时，部分实验人员

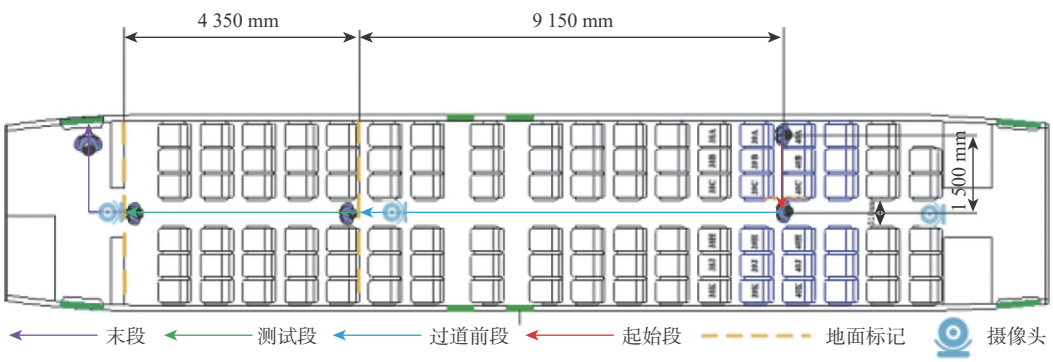


图 4 实验区布置示意图

Fig. 4 Layout of test area

表 1　A330-200 各种不利姿态角度

Table 1　Adverse attitude angle of A330-200

序号	姿态	俯仰角/(°)	横滚角/(°)
1	所有起落架放下	−1.25	0
2	双主起落架折断	+3.25	0
3	前起落架折断	−8.65	0
4	所有起落架折断	+2.48	0
5	前起及单个主起落架折断	+4.31	±10.25
6	机尾过重	+12.9	0

会明显感到头晕及严重失衡, 一些人员出现畏惧不敢运动。因此, 处于安全考虑, 实验中横滚角 γ 的取值为 0, ±5°, ±10°, 俯仰角 θ 的取值为 0, ±5°, ±10°。其中, 横滚角为正时, 机身左倾, 为负时, 机身右倾; 俯仰角为正时, 机头上仰, 即仰角, 为负时, 机头下沉, 即俯角。

实验共招募 24 名人员, 男 10 人(41.7%)、女 14 人(58.3%), 年龄 25～50 岁, 其中 50 岁的人员 9 人(37.5%), 含女性 4 人(16.7%), 35～49 岁 8 人, 含女性 5 人, 25～34 岁 7 人, 含女性 5 人。身高 150～180 cm, 体重 55～85 kg。招募的人员为科研工作者和管理人员, 其中从事科研工作的 18 人, 管理人员 6 人, 所有人员均乘坐过 2 次以上飞机, 个别人员乘坐次数超过 10 次。所有人员均没有参加或经历过应急撤离。每个实验对象以模拟舱内的第 40 排座椅 *A* 为起点, 达到客舱右前舱门并触碰舱门上的标记为结束。实验按自由撤离模式和快速撤离模式进行撤离, 以便研究快速撤离模式与自由撤离模式的差异。自由撤离模式指人员的正常行走状态; 快速撤离模式指人员模拟客机发生事故需进行应急撤离时, 在保证安全及稳定的情况下尽可能快速撤离的状态。

1.3.2　群体撤离实验工况

群体撤离实验环境与单人撤离实验环境相同, 每次实验人员 18 名, 随机就座于模拟舱内第 38 排至第 40 排的座位。实验人员需从座位处进入过道并撤离至客舱前部的黄色标记处, 如图 5 所示。群体撤离实验共招募 90 名人员, 年龄分布在 25～50 岁, 分为 5 组进行重复实验。每组的 18 人中, 性别和年龄比例按适航附录 J 要求, 女性 8 人, 含 50 岁的 3 人; 男性 10 人, 含 50 岁的 4 人。从事的职业主要是科研工作者、管理人员和学校教师, 其中科研工作者 10 人, 管理人员 5 人, 教师 3 人。所有人员均乘坐过飞机, 均没有参加或经历过应急撤离。

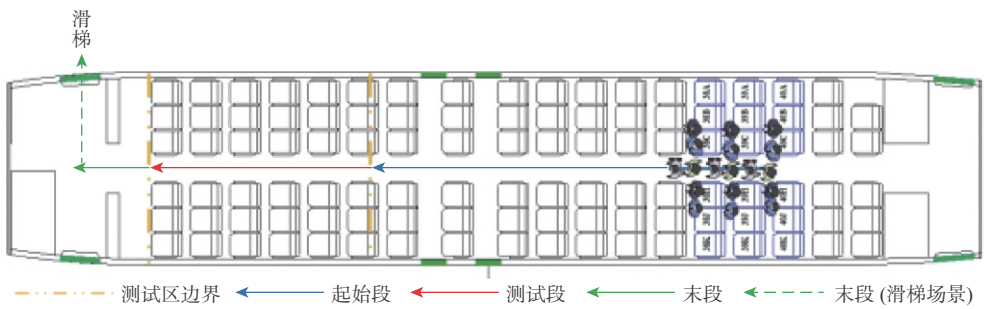


图 5　动态群体撤离实验布置示意图

Fig. 5　Arrangement of dynamic group evacuation test

1.4　测量与计算方法

1.4.1　个体撤离实验的测量与计算

个体撤离速度由行走距离除以时间获得。行走距离为实验测量区域内的有效长度 *s*; 时间取进入实验测量区域的时刻 t_{in} 与离开实验测量区域的时刻 t_{out} 之差 Δt, 通过监控装置及定位装置测量并记录。个体撤离速度 *v* 为

$$v=\frac{s}{\Delta t}=\frac{s}{t_{\text{out}}-t_{\text{in}}} \tag{1}$$

个体撤离速度的变化规律通过平均个体撤离速度分析。*n* 个实验对象的平均个体撤离速度 $\bar{v}$ 通过式(2)计算:

$$\bar{v}=\frac{\sum_{i=1}^{n}v_i}{n} \tag{2}$$

式中: v_i 为第 *i* 个实验对象的个体撤离速度。

个体撤离速度的离散程度通过撤离速度标准差来分析。每次实验个体撤离速度标准差 σ 由式(3)计算:

$$\sigma=\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(v_i-\bar{v})^2} \tag{3}$$

个体撤离速度随横滚角和俯仰角的变化通过定义个体撤离速度衰减比 $r_{b,c}^{a}$ 来分析。个体撤离速度衰减比为倾斜状态下个体撤离速度与正常(即水平)状态下个体撤离速度的比值, 由式(4)计算:

$$r_{b,c}^{a}=\frac{v_{b,c}^{a}}{v(0,0)} \tag{4}$$

式中: *a* 为倾斜角的类别, 横滚时 $a=\gamma$, 俯仰时 $a=\theta$; *b* 为运动模式类别, 正常步行时 *b* 为 w, 快速撤离

时，b 为 r；c 为区域标识，座椅区 c 为 chair，直线区 c 为 aisle；$v(0,0)$ 为客舱处于正常姿态（水平）时的平均个体撤离速度，根据所要计算的区域和撤离模式而选取相应的值。

1.4.2 群体撤离实验的测量与计算

群体撤离实验中主要测量和分析人员密度、前向距离及时间间隔等参数。文献 [18] 提供了一种如图 6 所示的累计图数据分析法。图中分别显示了 t 时刻已进入和离开实验测量区域的实验对象个数 $n_{\text{in}}(t)$ 和 $n_{\text{out}}(t)$，这 2 条曲线称为进入函数和离开函数，可分别通过计算实验对象的进入时间 t_{in}^n 和离开时间 t_{out}^n 获得[19]。

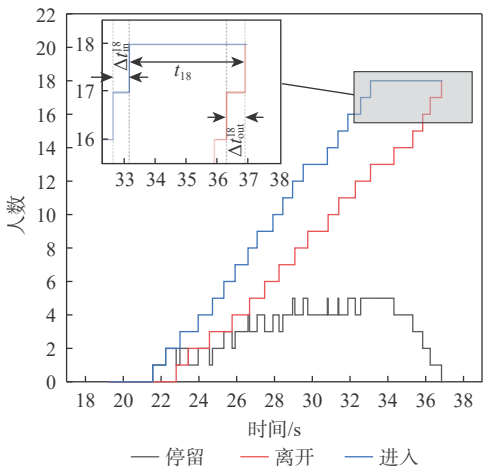


图 6 进入和离开实验测量区的实验对象累计图

Fig. 6 Aggregate test objects entering and leaving test area

在某次实验中，第 n 个实验对象与其前方的实验对象之间的进入时间间隔 Δt_{in}^n 和离开时间间隔 Δt_{out}^n 可由式（5）、式（6）计算：

$$\Delta t_{\text{in}}^n = t_{\text{in}}^n - t_{\text{in}}^{n-1} \tag{5}$$

$$\Delta t_{\text{out}}^n = t_{\text{out}}^n - t_{\text{out}}^{n-1} \tag{6}$$

第 n 个实验对象行走经过实验测量区域的时间 t_n，以及该实验对象在该时间内的个体撤离速度 v_n 可按式 (1) 计算。

第 n 个实验对象与其前方的实验对象之间的前向距离 d_n 及影响第 n 个实验对象运动的人员密度 ρ_n 可由式（7）、式（8）计算：

$$d_n = v_{n-1}\Delta t_{\text{in}}^n \tag{7}$$

$$\rho_n = \frac{1}{d_n} \tag{8}$$

2 实验结果分析

2.1 个体撤离实验

2.1.1 正常状态个体撤离速度

对所有实验对象在不同区域的个体撤离速度进行计算，得到个体撤离速度在各区域的分布情况如图 7 所示。虽然由于个体特征差异的影响，使个体撤离速度分布呈现出一定的离散，但可以发现，无论正常行走还是快速撤离，在座椅区域个体速度均表现出相似的降低。

在座椅区域，个体在自由撤离状态下撤离速度为 0.39～0.68 m/s，快速撤离模式下撤离速度为 0.49～0.89 m/s。在过道区域，个体在自由撤离状态下撤离速度为 1.21～2.06 m/s，平均速度在 1.5 m/s 左右，这与文献 [20] 在相似环境下得到的结论一致。快速撤离模式下撤离速度为 2.19～3.55 m/s。整个撤离过程，自由撤离状态下撤离速度为 0.93～1.56 m/s，快速撤离模式下撤离速度为 1.39～2.34 m/s。与自由撤离模式相比，快速撤离模式在各区域的撤离速度离散程度较高，这是由于不同性别个体的运动能力存在较大差异导致的。图 8 为快速撤离模式下不同性别的个体运动对比。可知，在快速撤离模式下，男性在各段的速度均高于女性。

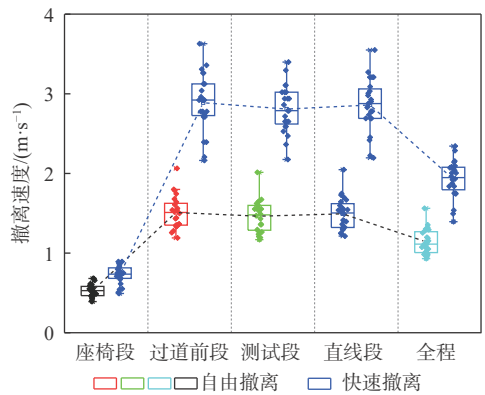


图 7 不同区域个体撤离速度的分布情况

Fig. 7 Distribution of individual walk speed in different area

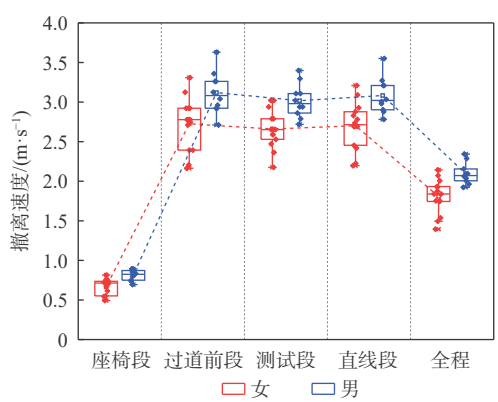


图 8 快速撤离模式下不同性别个体撤离速度

Fig. 8 Different sex evacuation speed at quick evacuation model

2.1.2 不利姿态对撤离速度的影响

横滚姿态个体撤离速度统计结果如图 9 所示。与水平姿态相比，横滚角对个体撤离速度有一

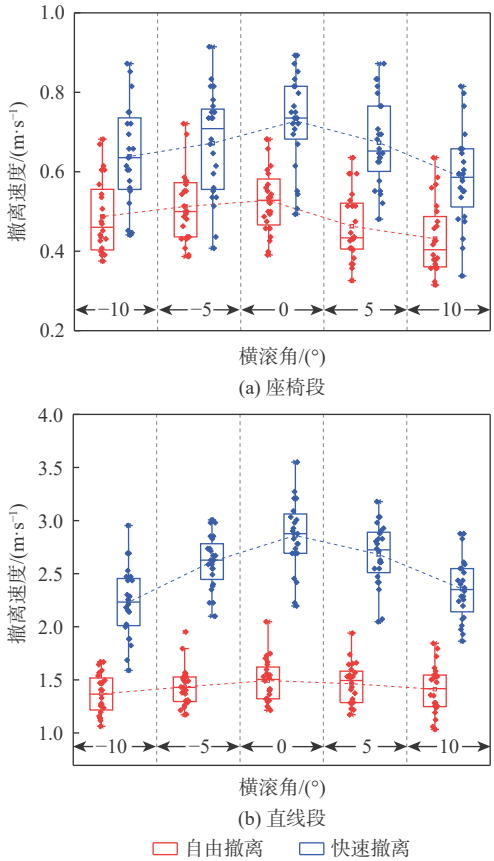


图 9　个体撤离速度随横滚角变化统计

Fig. 9　Individual evacuation speed in relation to roll angle change

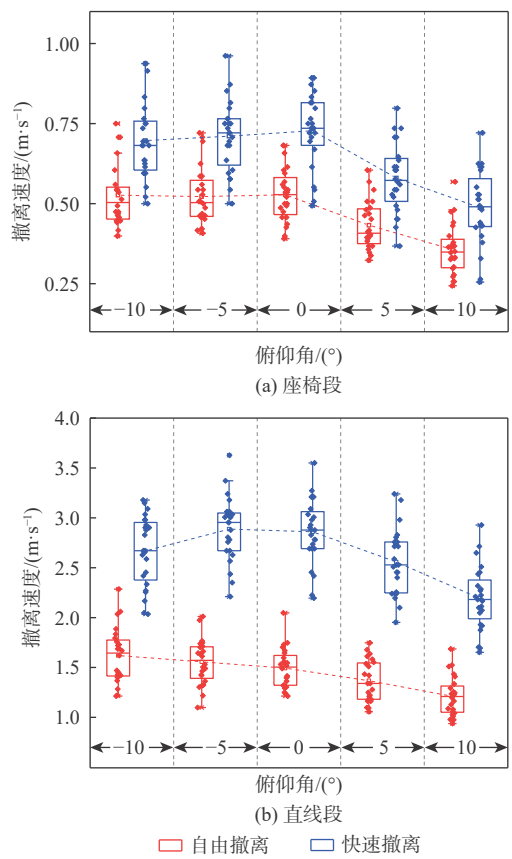


图 10　个体撤离速度随俯仰角的变化统计

Fig. 10　Individual evacuation speed in relation to pitch angle change

定的衰减作用, 且衰减程度随角度的增加而增强。由于个体差异, 速度分布呈现一定的离散分布, 且随角度增大而增大。正负横滚角对速度的影响大致呈对称状态。

俯仰状态个体撤离速度统计结果如图 10 所示。俯仰状态对个体撤离速度的影响明显与横滚状态不同。俯仰角为正时, 在座椅段及直线段, 人员撤离速度均有一定衰减, 且随角度增大而增强。俯仰角为负时, 对座椅段撤离速度几乎无影响。在直线段, 自由撤离模式下俯角会使速度提升, 且角度越大提升越高。快速撤离模式下, 小俯角 ($0<\theta<5°$), 平均速度几乎无变化, 中位数有一定增加。大俯角 ($5°<\theta<10°$), 因重力分量影响增大, 人员的撤离能力无法达到所产生的加速倾向, 撤离速度反而降低。

不同姿态下平均个体撤离速度衰减比如图 11 所示。可以看出, 撤离速度平均衰减比随横滚角的增大而减小, 近似成线性关系。在直线段正负角度影响效果几乎对称, 自由撤离模式下, 10°角平均衰减比为 0.92 左右; 快速撤离模式下, 10°角平均衰减比为 0.825 左右。在座椅区, 正负角度影响并不不对称, γ=+10°时座椅段平均衰减比为 0.825 左右, 而 γ=−10°时座椅段自由撤离和快速撤离模式下的平均衰减比分别为 0.928 与 0.882。

俯仰角 $\theta>0$ 时, 随角度的增大撤离速度衰减比减小, 对座椅段影响比直线段更大, 且仰角对于撤离速度的衰减效果明显比横滚角更大。仰角达到 10°时, 自由撤离状态下衰减比为 0.793, 快速撤离状态下速度衰减比为 0.761。而当俯仰角 $\theta<0$ 时, 自由步行模式, 直线段撤离速度衰减比随俯角度数增大而略微增大, 座椅段撤离速度衰减比几乎不变。快速撤离模式, 座椅段撤离速度衰减比随度数增大而略微减小, 直线段撤离速度衰减比在 0°～−5°几乎不变, 在−5°～−10°下降。

2.1.3　平均个体撤离速度衰减比模型

德国劳氏船级社曾根据实验提出了对船体横倾或纵倾状态下的平均个体撤离速度衰减比模型, 目前该模型已被当前广泛用于船舶设计的船舶疏散模拟软件 AENEAS 所采用[21]。基于本文所得实验数据, 针对飞机迫降后不利姿态应急撤离, 不利姿态影响下的基准撤离速度如表 2 所示, 横滚角影响下的平均个体撤离速度衰减比模型如式 (9)～式 (12) 所示, 俯仰角影响下的平均个体撤离速度衰减比模型如式 (13)～式 (16) 所示。

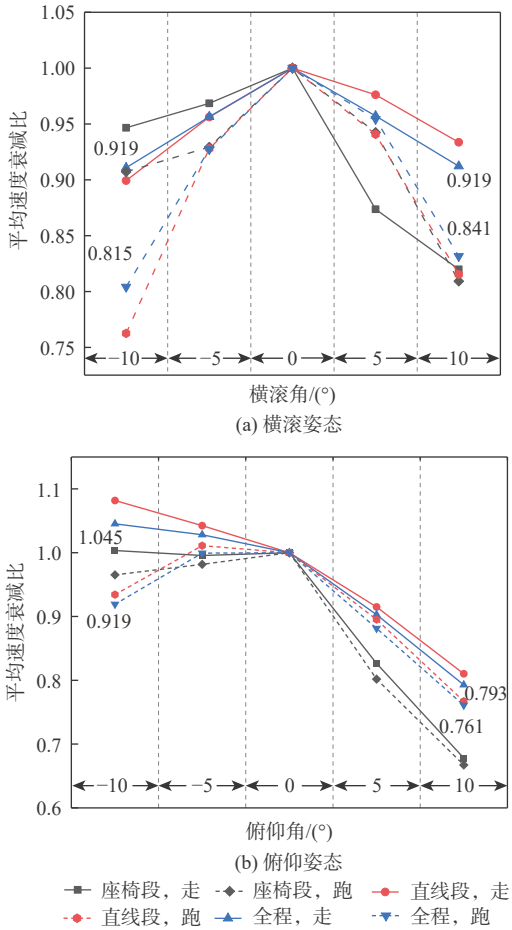


图 11 不同姿态下平均个体撤离速度衰减比

Fig. 11 Average individual evacuation speed attenuation ratio at different attitude

表 2 不利姿态下的基准撤离速度

Table 2 Reference evacuation speed under adverse attitude condition m/s

模式	性别	座椅段	直线段	全程
快速撤离	男	0.81	3.08	2.01
	女	0.67	2.70	1.80
自由撤离		0.53	1.49	1.14

横滚角影响下的平均个体撤离速度衰减比模型为

$$r^{\gamma}_{\mathrm{w,chair}}=\begin{cases}1+0.007\,2\gamma & -10\leqslant\gamma\leqslant 0\\ 1-0.018\gamma & 0<\gamma\leqslant 10\end{cases} \tag{9}$$

$$r^{\gamma}_{\mathrm{w,aisle}}=1-0.006\,75|\gamma| \qquad |\gamma|\leqslant 10 \tag{10}$$

$$r^{\gamma}_{\mathrm{r,chair}}=\begin{cases}1+0.011\,8\gamma & -10\leqslant\gamma\leqslant 0\\ 1-0.019\,2\gamma & 0<\gamma\leqslant 10\end{cases} \tag{11}$$

$$r^{\gamma}_{\mathrm{r,aisle}}=1-0.019\,4|\gamma| \qquad |\gamma|\leqslant 10 \tag{12}$$

俯仰角影响下的平均个体撤离速度衰减比模型为

$$r^{\theta}_{\mathrm{w,chair}}=\begin{cases}1 & -10\leqslant\theta\leqslant 0\\ 1-0.032\,5\theta & 0<\theta\leqslant 10\end{cases} \tag{13}$$

$$r^{\theta}_{\mathrm{w,aisle}}=\begin{cases}1-0.008\,2\theta & -10\leqslant\theta\leqslant 0\\ 1-0.019\theta & 0<\theta\leqslant 10\end{cases} \tag{14}$$

$$r^{\theta}_{\mathrm{r,chair}}=\begin{cases}1+0.003\,5\theta & -10\leqslant\theta\leqslant 0\\ 1-0.033\,3\theta & 0<\theta\leqslant 10\end{cases} \tag{15}$$

$$r^{\theta}_{\mathrm{r,aisle}}=\begin{cases}1.09+0.016\,2\theta & -10\leqslant\theta<-5\\ 1 & -5\leqslant\theta\leqslant 0\\ 1-0.023\,3\theta & 0<\theta\leqslant 10\end{cases} \tag{16}$$

式中：γ 和 θ 为倾斜角度的类别；w 和 r 为运动模式的类别，自由步行为 w，快速撤离为 r。

2.2 群体撤离实验

2.2.1 正常状态群体撤离实验

正常状态下群体 5 次撤离实验中各撤离次序的实验人员在全程及观察区域的平均撤离速度如图 12 所示。随着撤离次序的增加，全程和观察区的个体平均撤离速度均逐渐降低，且前后次序人员的速度差逐渐缩小，最终稳定于(1.22±0.23) m/s，实验结果与前人对 0.5 m 过道内单列行人撤离实验中得到的结果基本一致[20]。该结果与过道区域的人群密度变化有关。在疏散过程中，先进入过道的人员在低密度情况下拥有足够的空间进行加速。随着过道区域人群密度的不断增加，人员拥挤程度增加，移动速度降低。

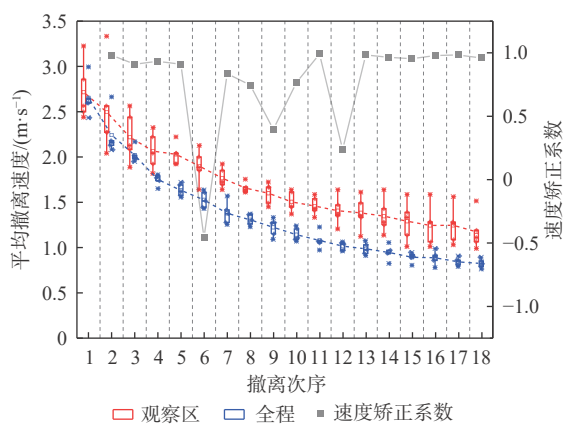


图 12 不同撤离次序个体平均撤离速度分布

Fig. 12 Individual average evacuation speed distribution in different sequence

2.2.2 不利姿态对群体撤离速度的影响

横滚姿态对个体撤离速度的影响如图 13 所示，因群体撤离实验中实验人员及路线的布置在横滚方向上有对称性，在统计时不再区分横滚角的正负。可知，群体中个体撤离速度随撤离次序的变大而逐渐降低，5°和 10°横滚角下最后 3 名速度降低至 1.075 m/s 和 0.978 m/s 左右。过道段速度衰减比均值分别为 0.898 和 0.816。

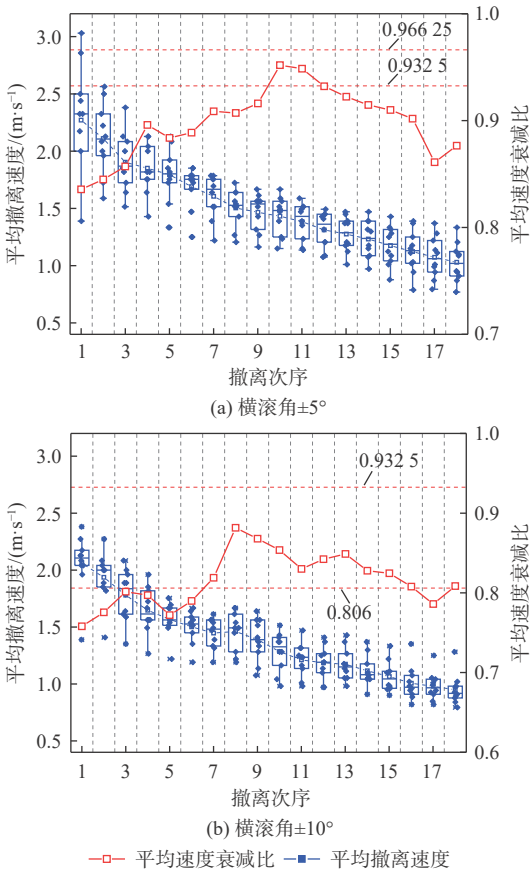


图 13　横滚姿态下平均撤离速度及其速度衰减比

Fig. 13　Average evacuation speed and speed attenuation ration at roll attitude

在同一实验中, 最前方人员由于速度最快, 其速度衰减效果更明显, 随着次序的靠后撤离速度逐渐降低, 因此, 受到横滚角衰减效果的影响也逐渐减弱。5°时整体速度的衰减比为 0.948, 10°时为 0.859, 均处于单人撤离实验中自由撤离速度衰减比和快速撤离速度衰减比之间。

俯仰姿态下平均个体撤离速度及平均撤离速度衰减比如图 14 所示。在仰角状态下, 对群体撤离中个体撤离速度的影响全部呈现出减速效果, 且大部分次序上的速度衰减比比个体实验中快速撤离模式下的速度衰减比更小, 5°和 10°时末位速度分别降低至 0.98 m/s和 0.82 m/s。

在俯角状态下, 在−5°俯角下群体中个体撤离速度均呈现出加速的效果, 末位速度达到 1.35 m/s, 总平均速度衰减比个体自由撤离速度衰减比还要大, 尤其是次序靠后的人员。而在俯仰角为−10°时, 群体中个体运动的影响也呈现出两面性: 一方面, 次序靠前、速度较快的个体受过大的倾斜角影响, 撤离速度减小; 另一方面, 次序靠后、速度较慢的人员受倾斜角造成重力分量向前的影响加速前进, 末位速度可达到 1.38 m/s。

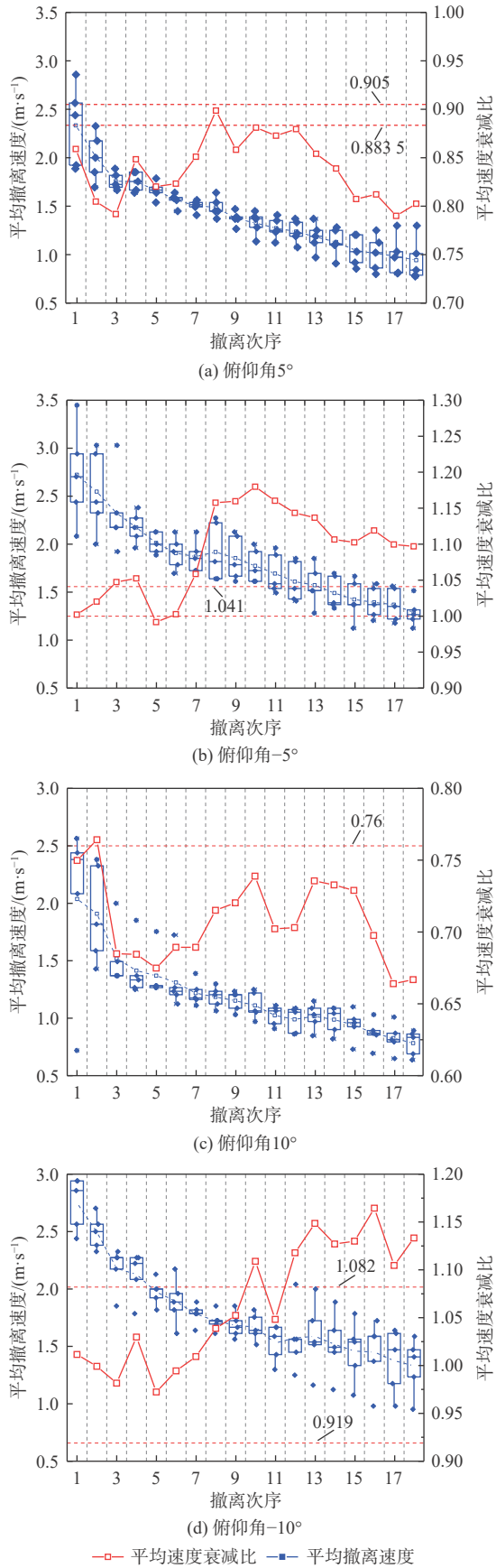


图 14　俯仰姿态下平均撤离速度及其速度衰减比

Fig. 14　Average evacuation speed and speed attenuation ratio at pitch attitude

3 不利姿态对群体撤离指标的影响

撤离时间、人员密度分布和前向距离是评价群体撤离的 3 个重要指标。

3.1 撤离时间

在相同条件下，撤离时间直接反映了撤离效率。各种不利姿态下撤离时间统计如图 15 所示。可知，不利姿态对群体撤离时间的影响与个体实验中快速撤离模式下撤离速度的影响规律类似。

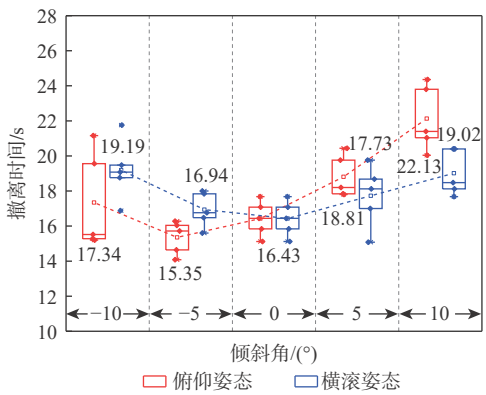


图 15 不利姿态下群体撤离时间统计

Fig. 15 Group evacuation time at adverse attitude

3.2 人员密度

单列行人运动时的个体撤离速度主要受人员密度的影响[22-23]。横滚状态下人员密度与撤离速度基本图如图 16 所示。随横滚角的增大，人员密度没有明显的上升或降低趋势。从拟合曲线也可看出，在相同的人员密度下，横滚角使人员撤离速度略微降低。当存在横滚角时，人员密度较高的区域数据点增加，即横滚角的影响使撤离速度降低，也使人员密度更高，这符合密度越高人员速度越低的趋势。

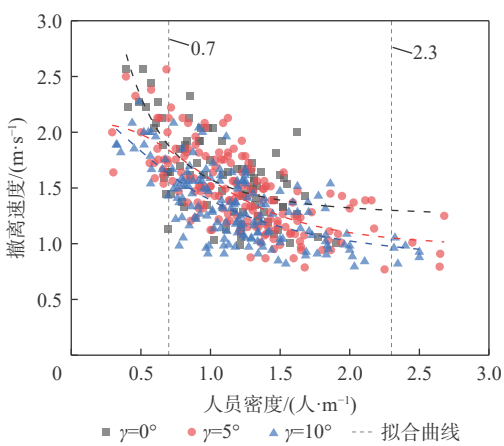


图 16 横滚姿态下的人员密度-撤离速度基本图

Fig. 16 Personnel density and evacuation speed at roll attitude

俯仰姿态下人员密度-速度基本图如图 17 所示。可以看出，数据点的分布出现了明显的分层，人员密度的分布也更加松散。在相同的密度下，仰角越大撤离速度越低。在俯角状态下，人员密度分布与 0°时近似一致，撤离速度略有提高。数据表明，在俯仰姿态，除了人员密度的影响外，由于重力分量方向与运动方向共线，重力分量对人员运动能力及群体运动规律的影响成为主要影响因素。

由实验数据可得如下结论：

1）在低密度区域 ($\rho < 0.5$ 人/m)，撤离速度的波动较大。

2）在中密度区域 (0.5 人/m$<\rho <$ 0.5 人/m)，撤离速度下降与人员密度增加几乎成线性关系。

3）在高密度区域 ($\rho > 1.5$ 人/m)，撤离速度下降曲率逐渐降低，继续增加人员密度则撤离速度几乎保持恒定。

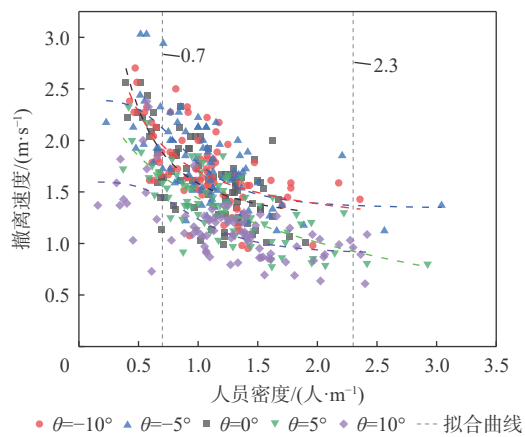


图 17 俯仰姿态下人员密度-撤离速度基本图

Fig. 17 Personnel density and evacuation speed at pitch attitude

3.3 前向距离

前向距离分布是量化单列个体运动空间分布的参数。不利姿态下撤离速度与前向距离的关系如图 18 所示，图中虚线为前向距离与撤离速度的线性拟合，其斜率表示每提高 1 m/s 速度需要多少前向空间。在横滚姿态，0°, 5°, 10°的斜率分别为 0.440 s, 0.463 s, 0.488 s, 相差不大。因此，存在横滚角时人员提升撤离速度所需要的前向距离变化不大。当俯仰角为−10°, 5°, 0°, 5°, 10°时拟合直线的斜率分别为 0.325 s, 0.373 s, 0.44 s, 0.50 s, 0.493 s, 变化明显较大，尤其是在角度小于 0, 客舱成俯角时。

因此，在俯仰姿态，个体在群体中运动所需前向距离的规律为：俯角状态，提升撤离速度所需前向空间变小，且角度越大所需前向空间越小；仰角状态，提升撤离速度所需前向距离增加，但并不随着角度增加而增大。

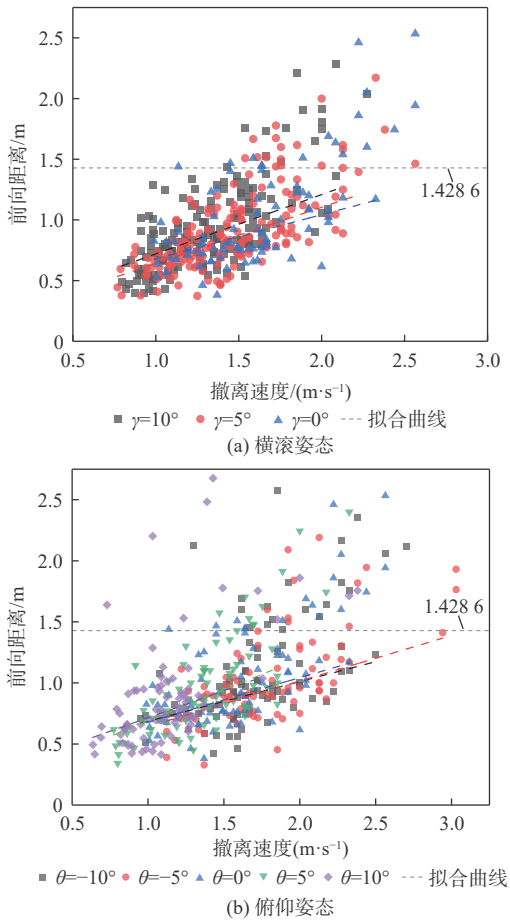


图 18　不利姿态下撤离速度与前向距离的关系

Fig. 18　Relation between evacuation speed and distance headway at adverse attitude

4　结　论

本文通过迫降后不利姿态应急撤离实验研究, 得到如下结论:

1) 不利姿态对应急撤离有一定的影响, 平均个体撤离速度会呈现出明显的衰减。在群体撤离过程中, 个体撤离速度按疏散次序逐渐减小。横滚姿态和俯仰姿态对应急撤离的影响程度不同, 与俯仰相比, 横滚对平均个体撤离速度影响相对较小。

2) 可通过撤离速度衰减比来表示不利姿态对平均个体撤离速度的影响程度。基于实验数据, 建立了不利姿态影响下的撤离速度衰减比模型。实验中俯仰和横滚均仅取了 3 种状态, 对拟合所得公式可靠度有一定影响, 但基本能反映不利姿态对撤离速度的影响。在横滚姿态, 个体的撤离速度衰减比处在自由撤离速度衰减比和快速撤离速度衰减比之间; 而在俯仰姿态, 仰角时速度衰减比比快速撤离速度衰减比更低, 俯角时速度衰减比比自由撤离速度衰减比更高。

3) 不利姿态对群体撤离中撤离速度的影响整体上与不利姿态对个体运动的影响类似。俯仰姿态除了影响个体撤离速度外, 还影响个体在群体运动中的前向距离。虽然不利姿态对撤离次序有一定的影响, 但从整体撤离次序分布来看, 即使存在较大倾斜角度, 只会对部分位置的撤离次序产生影响, 而对整体规律没有影响。

4) 实验揭示了不利姿态迫降客舱内不同个体在不同区域的运动规律和群体撤离过程中个体运动规律的变化及群体撤离行为规律, 所获得的不利姿态下个体撤离速度为民机应急撤离过程中个体运动模式提供了极具价值的基本数据, 为民用飞机应急撤离仿真提供了重要的参考依据。

参考文献 (References)

[1] XUE Z, BLOEBAUM C. A particle swarm optimization based aircraft evacuation simulation model-VacateAir: AIAA 2008-180[R]. Reston: AIAA, 2008.

[2] National Transportation Safety Board. Emergency evacuation of commercial airplanes: NTSB/SS-00/01[R]. Washington, D. C. : NTSB, 2000.

[3] YU H C, YANG H H. Cabin safety and emergency evacuation: Passenger experience of flight CI-120 accident [J]. Accident Analysis and Prevention, 2011, 43(3): 1049-1055.

[4] DENG X Q. An aircraft evacuation simulation baseline using DES for passenger path planning[D]. Daytona Beach: Embry-Riddle Aeronautical University, 2016: 12-15.

[5] VINUEZA V, MARÍA F, ZHOU K, et al. Cerebellar control of gait and interlimb coordination[J]. Brain Structure and Function, 2015, 220(6): 3513-3536.

[6] Interstate Aviation Committee. RRJ-95B RA-89098 interim report [R]. Moscow: Interstate Aviation Committee, 2019: 6-7.

[7] SUN J, GU Y, LI C, et al. An experimental study on individual walking speed during ship evacuation with the combined effect of heeling and trim[J]. Ocean Engineering, 2018, 166: 396-403.

[8] SUN J, LU S, LO S, et al. Moving characteristics of single file passengers considering the effect of ship trim and heeling[J]. Physica A:Statistical Mechanics and Its Applications, 2018, 490: 476-487.

[9] 彭浩轩, 刘小川, 白春玉, 等. 民机应急撤离实验与仿真研究进展[J]. 航空工程进展, 2020, 11(6): 759-766.
PENG H X, LIU X C, BAI C Y, et al. Advances in experiment and simulation of civil aircraft emergency evacuation[J]. Advances in Aeronautical Science and Engineering, 2020, 11(6): 759-766(in Chinese).

[10] HWANG K I. An experiment on walking speeds of freshmen unexperienced in shipboard life on a passenger ship[J]. Journal of Navigation and Port Research, 2013, 37(3): 239-244.

[11] KIM H T, LEE D K, PARK J, et al. The effect on the mobility of evacuating passengers in ship with regard to list and motion[J]. IE Interfaces, 2004, 17(1): 22-32.

[12] YOSHIKDA K, MURAYAMA M, ITAKAKI T. Study on evacuation of escape route in passenger in passenger ships by evacuation simulation and full-scale trials[J]. Journal of the Society of Naval

Architects of Japan, 1999, 1999(186): 581-589.

[13] SHI X, YE Z, SHIWAKOTI N, et al. Empirical investigation on safety constraints of merging pedestrian crowd through macroscopiduic and microscopic analysis[J]. Accident Analysis and Prevention, 2016, 95: 405-416.

[14] ZHANG J, KLINGSCH W, SCHADSCHNEIDER A, et al. Transitions in pedestrian fundamental diagrams of straight corridors and T-junctions[J]. Journal of Statistical Mechanics:Theory and Experiment, 2011, 2011(6): P06004.

[15] ZHANG T, ZHANG X, HUANG S, et al. Collective behavior of mice passing through an exit under panic[J]. Physica A:Statistical Mechanics and Its Applications, 2018, 496: 233-242.

[16] SHAH J, JOSHI G J, PARIDA P. Behavioral characteristics of pedestrian flow on stairway at railway station[J]. Procedia-Social and Behavioral Sciences, 2013, 104: 688-697.

[17] 张一山, 余江. 民航飞机擦机尾的风险防控与分析[J]. 中国民航飞行学院学报, 2021, 32(1): 9-13.
ZHANF Y S, YU J. Risk control and analysis of civil aircraft tail strike[J]. Journal of Civil Aviation Flight University of China, 2021, 32(1): 9-13(in Chinese).

[18] JOHN A, SCHADSCHNEIDER A, CHOWDHURY D, et al. Trafficlike collective movement of ants on trails: Absence of jammed phase[J]. Physical Review Letters, 2009, 102(10): 297-300.

[19] STEFFEN B, SEYFRIED A. Methods for measuring pedestrian density, flow, speed and direction with minimal scatter[J]. Physica A: Statistical Mechanics and Its Applications, 2010, 389(9): 1902-1910.

[20] HUANG S, ZHANG T, SIUMING L, et al. Experimental study of individual and single-file pedestrian movement in narrow seat aisle[J]. Physica A:Statistical Mechanics and its Applications, 2018, 509: 1023-1033.

[21] CHEN J, JIAN M, LO S. Modelling pedestrian evacuation movement on a swaying ship[EB/OL]. (2015-10-15)[2021-08-01]. https://arxiv.org/abs/1511.04686v1.

[22] JELIC A, APPERT R C, LEMERCIER S, et al. Properties of pedestrians walking in line: Fundamental diagrams[J]. Physical Review E: Statistical Nonlinear & Soft Matter Physics, 2012, 85(3): 036111.

[23] CHATTARAJ U, SEYFRIED A, CHAKROBORTY P, et al. Modelling single file pedestrian motion across cultures[J]. Procedia-Social and Behavioral Sciences, 2013, 104: 698-707.

Experimental study on adverse attitude emergency evacuation of civil aircraft after crash landing

FENG Wenchun[1,*], ZHANG Tong[2], LI Wei[1]

(1. Institute of Electromechanical Systems, The First Aircraft Institute of AVIC, Xi'an 710089, China;
2. School of Astronautics, Beihang University, Beijing 100191, China)

Abstract: Based on dynamic emergency evacuation test environment, the effects of adverse attitude on emergency evacuation were studied and more than two hundred tests, including individual evacuation and group evacuation tests, were implemented. Then the average evacuation speed of individuals at seat and aisle area was analyzed according to free walk and quick evacuation model, with the results of which models of individual average speed attenuation ratio was established. The tests show that the adverse attitude of landing has effects on evacuation time, personnel density and forward distance. Under roll condition, the speed attenuation ratio is 0.948 and 0.859 at 5 degrees and 10 degrees, respectively, declining with the increase of the roll angle. Under pitch condition, the effects of acceleration appears at −5 degrees and the evacuation speed of the last person is 1.35 meter per second, with speed attenuation ratio increasing with the evacuation sequence; while at −10 degrees, the individual movement displays two-sidedness, with the speed in inverse proportion with the evacuation sequence. The test results reveal the movement patterns of the individual and group evacuation under adverse attitude landing condition and provides important reference data for evacuation simulation of civil aircraft.

Keywords: emergency evacuation; crash landing; adverse attitude; emergency evacuation test; evacuation speed

Received: 2021-08-23; **Accepted**: 2021-10-22; **Published Online**: 2021-12-27 14:40

URL: kns.cnki.net/kcms/detail/11.2625.V.20211224.1458.001.html

Foundation item: Civil Aircraft Pre-study of Ministry of Industry and Information Technology (MJ-2016-F-11)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0477

四旋翼的改进 PSO-RBF 神经网络自适应滑模控制

唐志勇，马福源，裴忠才*
(北京航空航天大学 自动化科学与电气工程学院，北京 100191)

摘 要：针对非线性、强耦合并带有不确定性干扰的四旋翼无人机模型，提出了一种改进粒子群算法-径向基 (PSO-RBF) 神经网络自适应滑模控制器。在对 RBF 神经网络自适应滑模控制器进行控制量平滑改进的基础上，利用改进的具有全局寻优能力的 PSO 算法来调整 RBF 神经网络的拟合参数，从而进一步提升网络的拟合能力。根据实际四旋翼的模型参数，搭建四旋翼的动力学模型，通过 Lyapunov 理论验证了系统的稳定性。仿真结果表明：与 RBF 神经网络自适应滑模控制器和双闭环 PID 控制器相比，改进 PSO-RBF 神经网络自适应滑模控制器可以在一个控制周期内寻找到合适的控制量，其调节时间分别提升约 50% 和 75%；改进 PSO-RBF 神经网络自适应滑模控制器具有轨迹跟踪速度快且准、抗干扰能力强和鲁棒性好的特点。

关 键 词：四旋翼无人机；粒子群算法；径向基神经网络；自适应滑模控制；轨迹跟踪；抗干扰

中图分类号：V249.122

文献标志码：A **文章编号**：1001-5965(2023)07-1563-10

四旋翼无人机具有结构简单、易于原地起降和较强的地形适应能力等特点，在对地观测与军事领域有着广泛的应用，在科学数据监测、地质和林业植保领域也有着广泛的发展前景[1]。对四旋翼进行轨迹跟踪和抗干扰控制是四旋翼控制的重要部分。但是，外界干扰和陀螺效应的存在，使得建立准确的四旋翼动力学模型异常困难[2]。此外，四旋翼是典型的 4 输入 6 输出的欠驱动系统，具有高度的非线性和强耦合等控制难点[3]。

国内外学者提出了多种方法来设计四旋翼的控制器，常用的线性化方法有传统的内外环 PID 控制[4]和最优 LQR 控制[5]，非线性方法有自适应控制、滑模控制[6]、反步控制[7]和鲁棒 H_∞控制[8]等，智能控制方法有模糊控制[9]、神经网络控制[3]和优化算法(粒子群等)控制[10]等。

径向基 (radial basis function, RBF) 神经网络对非线性系统的逼近能力得到了国内外很多学者的关注[11]。孙瑞华等[12]采用 RBF 神经网络来优化 PID 的 3 个控制参数，提高了控制系统的快速调节能力，但其控制器中的 PID 控制部分是一种线性控制方法，对于控制非线性四旋翼模型有一定的局限性。滑模控制 (sliding mode control, SMC) 作为一种非线性控制方法具有很强的鲁棒性。韩业壮[3]、Lin[13]等采用 RBF 神经网络自适应滑模控制的方法，利用 RBF 神经网络的万能逼近特点，来逼近不精确的四旋翼模型，该方法比 RBF-PID 控制方法有更快的调节时间、更好的抗干扰能力和更强的鲁棒性，但 RBF 神经网络隐含层神经元的中心值、宽度和连接权值的选取多采用梯度类算法进行优化，易存在过拟合学习，无法保证恒收敛，缺乏可靠性[14]。粒子群算法 (particle swarm optimization, PSO) 作为一种全局寻优的优化算法，常被应用于四旋翼的路径规

收稿日期：2021-08-19；**录用日期**：2021-11-19；**网络出版时间**：2021-12-16 10:59
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20211215.1408.002.html
* **通信作者**. E-mail：peizc@buaa.edu.cn

引用格式：唐志勇，马福源，裴忠才. 四旋翼的改进 PSO-RBF 神经网络自适应滑模控制 [J]. 北京航空航天大学学报，2023，49（7）：1563-1572. TANG Z Y，MA F Y，PEI Z C. Improved PSO-RBF neural network adaptive sliding mode control for quadrotor systems [J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（7）：1563-1572 （in Chinese）.

划[15]，鲜有应用于四旋翼的稳定性和抗扰控制。有学者在对光伏系统单一控制率的实验过程中已经证实，该方法可以进一步提高RBF神经网络的逼近能力[14]，但设计的控制器存在抖振问题，应用在四旋翼上会导致电调电压频繁切换，严重缩短电机的使用寿命。

基于此，本文提出了一种四旋翼的改进PSO-RBF神经网络自适应滑模控制器。将改进PSO算法的全局寻优能力应用于寻找RBF神经网络神经元的中心值、宽度和连接权值中，进一步提高RBF神经网络对四旋翼模型的逼近速度，更快地逼近滑模控制器中四旋翼模型的不确定状态，同时合理设计了控制器以消除抖振现象，延长电机的使用寿命。

1 四旋翼无人机的动力学建模

四旋翼建模所需要的坐标系可以描述为图1所示的2个坐标系：原点固连在四旋翼机身上的机体坐标系($O_b xyz$)和原点固连在地面的大地坐标系($O_e XYZ$)。对于图1所示的四旋翼俯视图，选取四旋翼的前进方向为机体坐标系x轴正向，四旋翼的逆时针90°方向为y轴正向，z轴按照右手定则选取垂直纸面向外为正方向。同样，大地坐标系按照东北天来分别确定X、Y和Z轴的正方向。

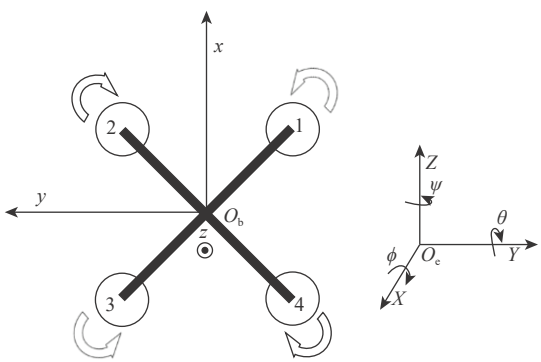


图1 坐标系和四旋翼模型

Fig. 1 Coordinate system and quadrotor model

四旋翼的姿态角定义在大地坐标系上，用来描述四旋翼的姿态。其中，滚转角ϕ定义为y轴在YO_eZ平面的投影与Y轴的夹角；俯仰角θ定义为z轴在XO_eZ平面的投影与Z轴的夹角；偏航角ψ定义为x轴在XO_eY平面的投影与X轴的夹角。

国内外很多文献已经对四旋翼进行了较为成熟的动力学建模[16]，在此不再重述建模过程，着重介绍控制器的设计过程，同时为了便于动力学建模，本文做出如下假设：

1）为了简化模型中转动惯量部分的计算，假设四旋翼飞行器是一种严格对称的刚体结构，其重心与中心相一致位于机体的几何中心。

2）四旋翼进行低空低速小角度飞行，重力为常值，且其姿态角度变化很小(通常小于5°)，这样可以简化欧拉方程。

3）由于四旋翼的相邻电机相互反转，可以抵消电机的陀螺效应的影响，四旋翼的陀螺力矩很小，可以忽略不计，同时忽略内部的各种摩擦和能量损耗。

4）根据叶素动量理论和空气动力学实验[17]，四旋翼每个旋翼产生的升力和反扭力矩都与该旋翼转速的平方成正比。

5）相关文献表明四旋翼受到的外部阻力与线速度成正比[17]。

定义U_1、U_2、U_3、U_4为四旋翼无人机4个独立的控制量，$\omega_i(i=1,2,3,4)$为4个电机的转速，则

$$\begin{cases} U_1 = K_1\omega_1^2 + K_1\omega_2^2 + K_1\omega_3^2 + K_1\omega_4^2 \\ U_2 = -K_1\omega_1^2 + K_1\omega_2^2 + K_1\omega_3^2 - K_1\omega_4^2 \\ U_3 = -K_1\omega_1^2 - K_1\omega_2^2 + K_1\omega_3^2 + K_1\omega_4^2 \\ U_4 = -K_2\omega_1^2 + K_2\omega_2^2 - K_2\omega_3^2 + K_2\omega_4^2 \end{cases} \tag{1}$$

式中：K_1为升力系数；K_2为反扭矩系数。

由牛顿-欧拉法建立的四旋翼无人机的动力学模型[16]可以表示为

$$\begin{cases} \ddot{x} = \dfrac{(\cos\phi\sin\theta\cos\psi + \sin\phi\sin\psi)U_1}{m} + d_1 \\ \ddot{y} = \dfrac{(\cos\phi\sin\theta\sin\psi - \sin\phi\cos\psi)U_1}{m} + d_2 \\ \ddot{z} = \dfrac{(\cos\phi\cos\theta)U_1}{m} - g + d_3 \\ \ddot{\phi} = \dfrac{\dot{\theta}\dot{\psi}(I_{yy} - I_{zz}) + U_2 l}{I_{xx}} + d_4 \\ \ddot{\theta} = \dfrac{\dot{\phi}\dot{\psi}(I_{zz} - I_{xx}) + U_3 l}{I_{yy}} + d_5 \\ \ddot{\psi} = \dfrac{\dot{\phi}\dot{\theta}(I_{xx} - I_{yy}) + U_4}{I_{zz}} + d_6 \end{cases} \tag{2}$$

式中：m为四旋翼的质量；I_{xx}、I_{yy}、I_{zz}为四旋翼各轴转动惯量；$d_i(i=1\sim6)$为外界对四旋翼的干扰力；g为重力加速度；l为力矩的力臂长度。

将式(2)整理为状态方程的形式：

$$\begin{cases} \dot{\boldsymbol{X}}_1 = \boldsymbol{X}_2 \\ \dot{\boldsymbol{X}}_2 = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})\boldsymbol{u} + \boldsymbol{d} \end{cases} \tag{3}$$

式中：$\boldsymbol{X}_1 = [x, y, z, \phi, \theta, \psi]^{\mathrm{T}}$，$\boldsymbol{X}_2 = [\dot{x}, \dot{y}, \dot{z}, \dot{\phi}, \dot{\theta}, \dot{\psi}]^{\mathrm{T}}$为状态向量；$\boldsymbol{u} = [1, 1, U_1, U_2, U_3, U_4]^{\mathrm{T}}$；$\boldsymbol{d} = [d_1, d_2, d_3 - g, d_4, d_5, d_6]^{\mathrm{T}}$；$\boldsymbol{g}(\boldsymbol{x}) = \left[\dfrac{U_1}{m}, \dfrac{U_1}{m}, \dfrac{\cos\phi\cos\theta}{m}, \dfrac{l}{I_{xx}}, \dfrac{l}{I_{yy}}, \dfrac{1}{I_{zz}}\right]^{\mathrm{T}}$；$\boldsymbol{f}(\boldsymbol{x}) = \left[0, 0, 0, \dfrac{(I_{yy} - I_{zz})\dot{\theta}\dot{\psi}}{I_{xx}}, \dfrac{(I_{zz} - I_{xx})\dot{\phi}\dot{\psi}}{I_{yy}}, \dfrac{(I_{xx} - I_{yy})\dot{\phi}\dot{\theta}}{I_{zz}}\right]^{\mathrm{T}}$。

2 RBF神经网络自适应滑模控制器

2.1 四旋翼控制器结构

四旋翼是4输入6输出的欠驱动、强耦合系

统, 需要设计一个控制器将四旋翼的位置/姿态跟踪误差在有限时间内收敛到零。本文控制系统采用常用的内外环控制结构。建立如图 2 所示的控制器, 外环为四旋翼的位置控制器, 内环为四旋翼的姿态控制器, 通过这种双环控制器, 可以实现 6 个通道的解耦控制, 提高系统的鲁棒性。

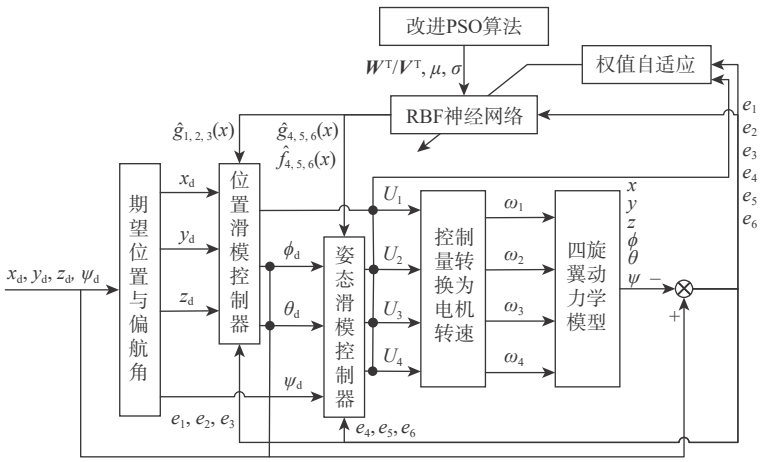


图 2　控制系统结构

Fig. 2　Control system structure

2.2　RBF 神经网络

RBF 神经网络是一种经典的三层前馈型神经网络, 合理确定神经元的参数可以逼近任意的非线性函数关系[18]。

RBF 神经网络的输出为

$$y(\boldsymbol{\zeta}) = \sum_{i=1}^{n} W_i h_i(\boldsymbol{\zeta}) \tag{4}$$

式中: $y(\boldsymbol{\zeta})$为 RBF 神经网络的输出; n 为隐含层神经元个数; W_i为隐含层神经元第 i 个神经元与输出层神经元的之间的连接权值; $h_i(\boldsymbol{\zeta})$为隐含层神经元的激活函数。

激活函数通常选取高斯函数, 输入越接近 RBF 的中心, 则输出值越大。本文中高斯函数定义为

$$\begin{cases} h_i(\boldsymbol{\zeta}) = \mathrm{e}^{\frac{-\|\boldsymbol{\zeta}-\boldsymbol{\mu}_{ni}\|^2}{2\sigma_i^2}} \\ \boldsymbol{\mu}_{ni} = \mu_i \left[\chi_1, \cdots, \chi_n\right]^{\mathrm{T}} \end{cases} \qquad \chi_n \in [-1, 1] \tag{5}$$

式中: $\boldsymbol{\zeta} = (a_1, a_2, \cdots, a_m)^{\mathrm{T}}$为输入矩阵; $\boldsymbol{\mu}_{ni}$为隐含层神经元的中心向量; σ_i为第 i 个隐含层神经元的宽度; $\|\boldsymbol{\zeta}-\boldsymbol{\mu}_{ni}\|$为输入矩阵和径向基中心之间的欧氏距离。

2.3　RBF 神经网络自适应滑模位置控制器设计

四旋翼的位置控制可以分为水平位置控制和高度位置控制, 通过式 (2) 可知, 高度 z 由 U_1控制, 下面进行该通道控制率的设计。

已知高度 z 的状态空间表达式为

$$\begin{cases} \dot{z} = \dot{z} \\ \ddot{z} = g_3(x)U_1 - g + d_3 \end{cases} \tag{6}$$

定义跟踪误差为计划高度−实际高度, 即 $e_3 = z_d - z$。

选取如下滑模面[19]:

$$s_3 = \dot{e}_3 + c_3 e_3 \quad \text{且 } c_3 > 0 \tag{7}$$

令 $\dot{s}_3 = 0$, 可设计滑模控制率 =等效控制率+切换控制率, 即

$$U_1 = \frac{1}{\hat{g}_3(x)}(\ddot{z}_d + g + c_3\dot{e}_3 + k_3\mathrm{sgn}(s_3)) \tag{8}$$

式中: $\hat{g}_3(x)$为 RBF 神经网络对$g_3(x)$的逼近结果。

定义神经网络的输入为$\boldsymbol{\zeta}_3 = (e_3, \dot{e}_3)^{\mathrm{T}}$, 根据 2.2 节结果, 可以得到

$$\begin{cases} \hat{g}_3(\boldsymbol{\zeta}_3) = \hat{\boldsymbol{W}}_{g_3}^{\mathrm{T}} \boldsymbol{h}_{g_3}(\boldsymbol{\zeta}_3) \\ g_3(\boldsymbol{\zeta}_3) = \widehat{\boldsymbol{W}}_{g_3}^{\mathrm{T}} \boldsymbol{h}_{g_3}(\boldsymbol{\zeta}_3) + \varepsilon_{g_3} \end{cases} \tag{9}$$

式中: $\widehat{\boldsymbol{W}}_{g_3}^{\mathrm{T}}$为网络逼近$g_3(x)$的理想权值; ε_{g_3}为网络的逼近误差, 且$|\varepsilon_{g_3}| < \varepsilon_{\max g(\zeta)}$。

RBF 神经网络的权值多采用梯度下降的方法来计算, 容易陷入局部最优解, 不能保证闭环系统的全局稳定, 本文采用 Lyapunov 方法来设计自适应$\hat{\boldsymbol{W}}_{g_3}^{\mathrm{T}}$。

定义闭环系统的 Lyapunov 函数为

$$L_3 = \frac{1}{2}s_3^2 + \frac{1}{2}\gamma_{g_3}\tilde{\boldsymbol{W}}_{g_3}^{\mathrm{T}}\tilde{\boldsymbol{W}}_{g_3} \quad \text{且 } \gamma_{g_3} > 0 \tag{10}$$

对式 (10) 求导, 得到

$$\begin{aligned} \dot{L}_3 = &\tilde{\boldsymbol{W}}_{g_3}^{\mathrm{T}}(s_3\boldsymbol{h}_{g_3}(\boldsymbol{\zeta}_3)U_1 + \gamma_{g_3}\dot{\hat{\boldsymbol{W}}}_{g_3}) - \\ &s_3(\varepsilon_{g_3}U_1 + k_3\mathrm{sgn}(s_3) + d_3) \end{aligned} \tag{11}$$

式中: $\tilde{\boldsymbol{W}}_{g_3}^{\mathrm{T}} = \hat{\boldsymbol{W}}_{g_3} - \widehat{\boldsymbol{W}}_{g_3}$。

设计自适应率为

$$\dot{\hat{\boldsymbol{W}}}_{g_3} = -\frac{1}{\gamma_{g_3}} s_3 \boldsymbol{h}_{g_3}(x) U_1 \tag{12}$$

此时, Lyapunov 函数变为: $\dot{L}_3 = -s_3(\varepsilon_{g_3}U_1 + d_3) - k_3|s_3|$, 逼近误差是一个很小的正实数, 取 $k_3 \geqslant$

$|\varepsilon_{g_3}U_1+d_3|$时，可以保证$\dot{L}_3\leqslant 0$。

为了减少RBF神经网络的估测误差，采用U_1来推导U_x和U_y。考虑到四旋翼的状态方程(2)，可以得到

$$\begin{cases}\ddot{x}=\dfrac{U_xU_1}{m}+d_1\\ \ddot{y}=\dfrac{U_yU_1}{m}+d_2\end{cases}\tag{13}$$

采用式(6)～式(8)相似的推导方式，可以得到x、y方向的滑模控制率U_x、U_y分别为

$$\begin{cases}U_x=\dfrac{m}{U_1}(\ddot{x}_{\mathrm{d}}+c_1\dot{e}_1+k_{U_x}\mathrm{sgn}(s_1))\\ U_y=\dfrac{m}{U_1}(\ddot{y}_{\mathrm{d}}+c_2\dot{e}_2+k_{U_y}\mathrm{sgn}(s_2))\end{cases}\tag{14}$$

由四旋翼的状态方程(2)可以得到

$$\begin{bmatrix}U_x\\U_y\end{bmatrix}=\begin{bmatrix}\sin\psi & \cos\psi\\ -\cos\psi & \sin\psi\end{bmatrix}\begin{bmatrix}\sin\phi\\ \cos\phi\sin\theta\end{bmatrix}\tag{15}$$

由图2可知，在给定参考角度ψ_{d}的情况下，由式(15)第一行公式可以得到ϕ_{d}，将所得代入第二行公式，进行简单的三角变换可以得到θ_{d}。

$$\begin{cases}\phi_{\mathrm{d}}=\arcsin(U_x\sin\psi_{\mathrm{d}}-U_y\cos\psi_{\mathrm{d}})\\ \theta_{\mathrm{d}}=\arcsin\left(\dfrac{U_x-\sin\psi_{\mathrm{d}}\cos\psi_{\mathrm{d}}}{\cos\phi_{\mathrm{d}}\cos\psi_{\mathrm{d}}}\right)\end{cases}\tag{16}$$

2.4 RBF神经网络自适应滑模姿态控制器设计

由式(2)易知，姿态控制相对独立于位置控制。以滚转角ϕ为例，设计RBF神经网络自适应滑模控制率。根据式(2)可以得到滚转角ϕ的状态方程为

$$\begin{cases}\dot{\phi}=\dot{\phi}\\ \ddot{\phi}=f_4(x)+g_4(x)U_2+d_4\end{cases}\tag{17}$$

同样给出跟踪误差的定义为期望滚转角–实际滚转角度，即$e_4=\phi_{\mathrm{d}}-\phi$，取滑模面的函数为$s_4=\dot{e}_4+c_4e_4$，且$c_4>0$。

根据2.3节的推导过程可以得到ϕ的RBF神经网络滑模控制率为

$$U_2=\frac{1}{\hat{g}_4(x)}(-\hat{f}_4(x)+\ddot{\phi}_{\mathrm{d}}+c_4\dot{e}_4+k_4\mathrm{sgn}(s_4))\tag{18}$$

定义神经网络的输入为$\boldsymbol{\zeta}_4=(e_4,\dot{e}_4)^{\mathrm{T}}$，根据2.2节结果，可以得到

$$\begin{cases}\hat{g}_4(\boldsymbol{\zeta}_4)=\hat{\boldsymbol{W}}_{g_4}^{\mathrm{T}}\boldsymbol{h}_{g_4}(\boldsymbol{\zeta}_4)\\ \hat{f}_4(\boldsymbol{\zeta}_4)=\hat{\boldsymbol{V}}_{f_4}^{\mathrm{T}}\boldsymbol{h}_{f_4}(\boldsymbol{\zeta}_4)\\ g_4(\boldsymbol{\zeta}_4)=\widehat{\boldsymbol{W}}_{g_4}^{\mathrm{T}}\boldsymbol{h}_{g_4}(\boldsymbol{\zeta}_4)+\varepsilon_{g_4}\\ f_4(\boldsymbol{\zeta}_4)=\widehat{\boldsymbol{V}}_{f_4}^{\mathrm{T}}\boldsymbol{h}_{f_4}(\boldsymbol{\zeta}_4)+\varepsilon_{f_4}\end{cases}\tag{19}$$

式中：$\widehat{\boldsymbol{W}}_{g_4}^{\mathrm{T}}$、$\widehat{\boldsymbol{V}}_{f_4}^{\mathrm{T}}$分别为$g_4(x)$、$f_4(x)$的理想权值；$\varepsilon_{g_4}$、$\varepsilon_{f_4}$为网络的逼近误差，且$|\varepsilon_{g_4}|<\varepsilon_{\max g(x)}$、$|\varepsilon_{f_4}|<\varepsilon_{\max f(x)}$。

将式(19)代入滑模面函数的导数公式，得到

$$\begin{aligned}\dot{s}_4=&(\tilde{\boldsymbol{V}}_{f_4}^{\mathrm{T}}\boldsymbol{h}_{f_4}(\boldsymbol{\zeta}_4)-\varepsilon_{f_4})+(\tilde{\boldsymbol{W}}_{g_4}^{\mathrm{T}}\boldsymbol{h}_{g_4}(\boldsymbol{\zeta}_4)-\varepsilon_{g_4})U_2-\\ &k_4\mathrm{sgn}(s_4)-d_4\end{aligned}\tag{20}$$

式中：$\tilde{f}_4(\boldsymbol{\zeta}_4)=\hat{f}_4(\boldsymbol{\zeta}_4)-f_4(\boldsymbol{\zeta}_4)=\tilde{\boldsymbol{V}}_{f_4}^{\mathrm{T}}\boldsymbol{h}_{f_4}(\boldsymbol{\zeta}_4)-\varepsilon_{f_4}$；$\tilde{g}_4(\boldsymbol{\zeta}_4)=\tilde{\boldsymbol{W}}_{g_4}^{\mathrm{T}}\boldsymbol{h}_{g_4}(\boldsymbol{\zeta}_4)-\varepsilon_{g_4}$。

同样，采用Lyapunov方法来设计自适应$\hat{\boldsymbol{V}}_{f_4}^{\mathrm{T}}$、$\hat{\boldsymbol{W}}_{g_4}^{\mathrm{T}}$，定义闭环系统的Lyapunov函数为

$$L_4=\frac{1}{2}s_4^2+\frac{1}{2}\gamma_{f_4}\tilde{\boldsymbol{V}}_{f_4}^{\mathrm{T}}\tilde{\boldsymbol{V}}_{f_4}+\frac{1}{2}\gamma_{g_4}\tilde{\boldsymbol{W}}_{g_4}^{\mathrm{T}}\tilde{\boldsymbol{W}}_{g_4}\quad\text{且}\gamma_{f_4}>0,\gamma_{g_4}>0\tag{21}$$

对式(21)求导，得到

$$\begin{aligned}\dot{L}_4=&\tilde{\boldsymbol{V}}_{f_4}^{\mathrm{T}}[s_4\boldsymbol{h}_{f_4}(\boldsymbol{\zeta}_4)+\gamma_{f_4}\dot{\hat{\boldsymbol{V}}}_{f_4}]+\tilde{\boldsymbol{W}}_{g_4}^{\mathrm{T}}[s_4\boldsymbol{h}_{g_4}(\boldsymbol{\zeta}_4)U_2+\gamma_{g_4}\dot{\hat{\boldsymbol{W}}}_{g_4}]+\\ &s[-\varepsilon_{f_4}-\varepsilon_{g_4}U_2-k_4\mathrm{sgn}(s_4)-d_4]\end{aligned}\tag{22}$$

根据式(22)可以设计RBF神经网络权重自适应律为

$$\begin{cases}\dot{\hat{\boldsymbol{V}}}_{f_4}=-\dfrac{1}{\gamma_{f_4}}s_4\boldsymbol{h}_{f_4}(\boldsymbol{\zeta}_4)\\ \dot{\hat{\boldsymbol{W}}}_{g_4}=-\dfrac{1}{\gamma_{g_4}}s_4\boldsymbol{h}_{g_4}(\boldsymbol{\zeta}_4)U_2\end{cases}\tag{23}$$

此时$\dot{L}_4$变为：$\dot{L}_4=-s_4[\varepsilon_{f_4}+\varepsilon_{g_4}U_2+d_4]-k_4|s_4|$，若选取$k_4\geqslant|\varepsilon_{f_4}+\varepsilon_{g_4}U_2+d_4|$，可以得到$\dot{L}_4\leqslant 0$。

同理可得θ、ψ的控制率U_3、U_4为

$$\begin{cases}U_3=\dfrac{1}{\hat{g}_5(\boldsymbol{\zeta}_5)}(-\hat{f}_5(\boldsymbol{\zeta}_5)+\ddot{\theta}_{\mathrm{d}}+c_5\dot{e}_5+k_5\mathrm{sgn}(s_5))\\ U_4=\dfrac{1}{\hat{g}_6(\boldsymbol{\zeta}_6)}(-\hat{f}_6(\boldsymbol{\zeta}_6)+\ddot{\psi}_{\mathrm{d}}+c_6\dot{e}_6+k_6\mathrm{sgn}(s_6))\end{cases}\tag{24}$$

相应的权值自适应率可以设计为

$$\begin{cases}\dot{\boldsymbol{V}}_{f_i}=-\dfrac{1}{\gamma_{f_i}}s_i\boldsymbol{h}_{f_i}(\boldsymbol{\zeta}_i)\\ \dot{\boldsymbol{W}}_{g_i}=-\dfrac{1}{\gamma_{g_i}}s_i\boldsymbol{h}_{g_i}(\boldsymbol{\zeta}_i)U_{i-2}\end{cases}\quad i=5,6\tag{25}$$

控制率(18)和(24)需要对期望的滚转角ϕ_{d}和期望的偏航角ψ_{d}求取一阶导数和二阶导数，常规方法难以获取，可以采用如下有限时间收敛的三阶微分器[20]实现$\dot{\phi}_{\mathrm{d}}$、$\ddot{\phi}_{\mathrm{d}}$和$\dot{\psi}_{\mathrm{d}}$、$\ddot{\psi}_{\mathrm{d}}$：

$$\begin{cases}\dot{x}_1=x_2\\ \dot{x}_2=x_3\\ \varepsilon^3\dot{x}_3=-2^{3/5}\times4[x_1-v(t)+(\varepsilon x_2)^{9/7}]^{1/3}-4(\varepsilon^2x_3)^{3/5}\end{cases}\tag{26}$$

式中：$v(t)$为待微分的输入信号；本文中$\varepsilon=0.1$；x_1为对输入信号的跟踪值；x_2为信号的一阶导数估计；x_3为信号的二阶导数的估计。微分器的初始值为$[x_1(0),x_2(0),x_3(0)]=[0,0,0]$。

针对传统滑模控制率中sgn()函数导致的抖振问题，利用边界层方法采用san()函数代替sgn()函数，即

$$\mathrm{san}(s_i)=\begin{cases}1 & s_i>\delta\\ \dfrac{1}{\delta}s_i & |s_i|\leqslant\delta\\ -1 & s_i<-\delta\end{cases} \tag{27}$$

本文中控制器取$\delta=0.1$, 采用san()函数可以减少控制率的抖振现象, 明显提升控制品质。

3　改进 PSO-RBF 神经网络自适应滑模控制器

在利用 RBF 神经网络预测控制率参数时, 发现神经网络的中心值μ_i、宽度σ_i及连接权值W_i的选取会对网络的预测效果产生显著的影响。合理选择会使预测效果良好, 使四旋翼较快跟踪上期望轨迹; 过于偏离合理值时, 会使网络的预测效果急剧变差, 不能很好地跟踪期望轨迹, 且四旋翼控制量之间的耦合关系导致调整姿态控制器的 RBF 神经网络会影响位置控制器的 RBF 神经网络预测结果。基于上述问题, 考虑到 PSO 算法的寻优能力, 决定将其应用到 RBF 神经网络的参数选择中, 以期减少人工调试时间, 同时进一步提升系统的响应速度。

3.1　改进 PSO 算法

普通 PSO 算法的全局搜索能力较差, 搜索精度低, 协调能力弱, 易陷入局部最优解。引入惯性权重可以有效改善局部最优解的问题, 常用的随机法、线性法仅可改变局部问题, 通过对 RBF 神经网络的分析发现, 粒子群的过早收敛会缺乏多样性, 使神经元感知范围变小, 系统的调节时间变长。因此, 结合 RBF 神经网络, 本文选用了改进 PSO 算法。

定义粒子群第j个粒子的位置为

$$\boldsymbol{b}_j=[b_{j,1},b_{j,2},\cdots,b_{j,D}] \tag{28}$$

第j个粒子的速度为

$$\boldsymbol{v}_j=[v_{j,1},v_{j,2},\cdots,v_{j,D}] \tag{29}$$

式中: j为全体粒子的个数; D为粒子的维数, 本文代表神经网络的中心值μ_i、宽度σ_i及连接权值W_i。

第j个粒子经t次迭代 (为提升计算速度且保证一定精度, 经多次实验后, 设定迭代次数为 100 次) 后的局部最优解, 即个体极值为

$$\boldsymbol{P}_{j\mathrm{best}}(t)=[P_{j,1}(t),P_{j,2}(t),\cdots,P_{j,D}(t)] \tag{30}$$

整个粒子群在经t次迭代后搜索到的最优位置为全局最优解, 即全局极值为

$$\boldsymbol{G}_{\mathrm{best}}(t)=[G_1(t),G_2(t),\cdots,G_D(t)] \tag{31}$$

在找到 2 个最优解的前提下, 粒子的速度和位置按照式 (32) 进行更新:

$$\begin{cases}\boldsymbol{v}_j(t+1)=\boldsymbol{H}(t)+\lambda_1 r_1[\boldsymbol{P}_{j\mathrm{best}}(t)-\boldsymbol{b}_j(t)]+\\ \qquad\lambda_2 r_2[\boldsymbol{G}_{\mathrm{best}}(t)-\boldsymbol{b}_j(t)]\\ \boldsymbol{b}_j(t+1)=\boldsymbol{b}_j(t)+\boldsymbol{v}_j(t+1)\end{cases} \tag{32}$$

式中: λ_1、λ_2为介于 [0,2] 之间的加速度常数, 通常$\lambda_1=\lambda_2$; r_1、r_2为介于 [0,1] 之间的随机常数; $\boldsymbol{H}(t)$为粒子速度的惯性权重代表粒子维持原先速度的能力。

经过适应度函数更新的粒子个体最优和全局最优值为

$$\begin{cases}\boldsymbol{P}_{j\mathrm{best}}(t+1)=\begin{cases}\boldsymbol{P}_{j\mathrm{best}}(t) & \boldsymbol{f}(\boldsymbol{b}_j(t+1))\geqslant\boldsymbol{f}(\boldsymbol{P}_{j\mathrm{best}}(t))\\ \boldsymbol{b}_j(t+1) & \text{其他}\end{cases}\\ \boldsymbol{G}_{\mathrm{best}}(t+1)=\min[\boldsymbol{f}(\boldsymbol{P}_{j\mathrm{best}}(t+1))]\end{cases} \tag{33}$$

本文选择相应控制量的$g_i(x)$或$f_i(x)$作为表示结果好坏的粒子适应度函数$f_{\mathrm{p}}(x)$。

改进粒子群的另一重点在粒子速度惯性权重$\boldsymbol{H}(t)$, 本文粒子速度更新权值定义为

$$\boldsymbol{H}(t)=\left\{\alpha+\frac{f(\boldsymbol{G}_{\mathrm{best}}(t))}{f(\boldsymbol{b}_j(t))}\right\}\{\beta^{-t}\cdot\mathrm{rand}\cdot\boldsymbol{v}_j(t)\} \tag{34}$$

式中: β为初始化常数, 且$\beta=1.000\,09$; 本文取$\alpha=2$, 作为改善粒子全局搜索能力的预定义常数; $L(t)=\{\beta^{-t}\cdot\mathrm{rand}\cdot\boldsymbol{v}_j(t)\}$用来描述策略的非线性, 定义群体的多样性, 表示粒子的聚散程度, 可进一步平衡粒子的全局搜索和局部搜索能力。

3.2　改进 PSO-RBF 神经网络自适应滑模控制算法

根据 2.2 节的 RBF 神经网络基本结构可知, 影响网络预测结果的主要因素是神经网络的中心值μ_i、宽度σ_i及连接权值W_i。因此, 结合 PSO 算法本身的特性, 设置粒子的位置为

$$\boldsymbol{b}=[\mu_1,\sigma_1,W_1,\mu_2,\sigma_2,W_2,\cdots,\mu_k,\sigma_k,W_k] \tag{35}$$

粒子的空间维数为

$$D=(2+n)k \tag{36}$$

式中: n为中心值变量个数; k为输入向量维度。

为进一步利用改进 PSO 算法使网络误差函数最小化, 以样本实际输出与网络输出之间的相对误差函数作为粒子群适应度函数$f_{\mathrm{p}}(x)$:

$$f_{\mathrm{p}}(\boldsymbol{b}_j(t))=\sqrt{\frac{1}{T}\sum_{j=1}^{T}(N_{\mathrm{d}}(t)-N(t))^2} \tag{37}$$

式中: $N_{\mathrm{d}}(t)$为期望结果; $N(t)$为网络的实际输出; T为样本数。

综上, 可以得到改进 PSO-RBF 神经网络自适应滑模控制算法的流程, 如图 3 所示。

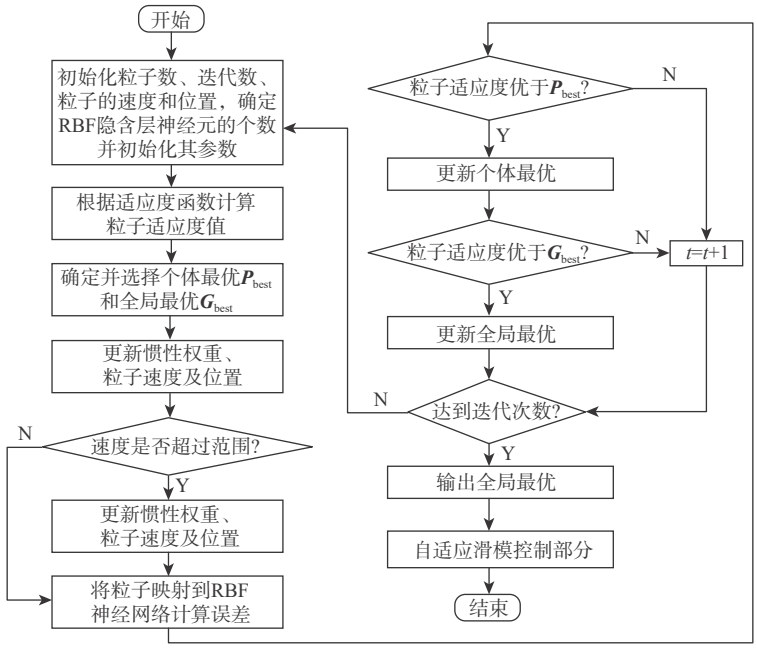


图 3 改进 PSO-RBF 神经网络自适应滑模控制流程

Fig. 3 Flow chart of improved PSO-RBF neural network adaptive sliding mode controller

4 仿真分析

对第 3 节所述控制系统进行仿真，笔者实验室四旋翼无人机的实际动力学参数如表 1 所示。

将改进 PSO-RBF 神经网络自适应滑模控制算法与 RBF 神经网络自适应滑模控制算法和传统双闭环 PID 控制算法在轨迹跟踪、抗干扰和鲁棒性等方面进行仿真实验，说明改进 PSO-RBF 神经网络自适应滑模控制算法的优越性。

表 1 四旋翼无人机的动力学参数

Table 1 Dynamic parameters of quadrotor

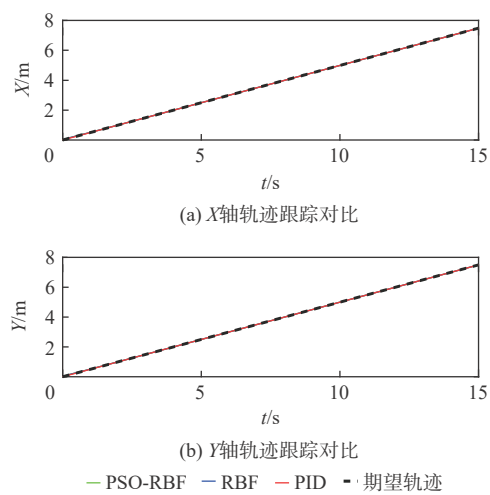
参数	数值
升力系数K_1/(N·s²·rad⁻²)	9.138×10^{-6}
反扭矩系数K_2/(N·m·s²·rad⁻²)	1.368×10^{-7}
沿x轴转动惯量I_{xx}/(kg·m²)	1.762×10^{-2}
沿y轴转动惯量I_{yy}/(kg·m²)	1.769×10^{-2}
沿z轴转动惯量I_{zz}/(kg·m²)	2.805×10^{-2}
四旋翼质量m/kg	1.311
臂长d/m	0.24

4.1 轨迹跟踪对比

在 X、Y、Z 轴设定期望的跟踪轨迹 $X_d=0.5t$，$Y_d=0.5t$，$Z_d=\begin{cases}t & 0\leqslant t<10\\ 10 & t\geqslant 10\end{cases}$ 和期望的偏航角 $\psi_d=30°$ 以此来模拟四旋翼的跟踪定高飞行过程。分别对 3 种控制器进行仿真，绘制 0～15 s的位置和姿态变化对比图。观察四旋翼在改进 PSO-RBF 神经网络自适应滑模控制器、RBF 神经网络自适应滑模控制器和传统双闭环 PID 控制器下的轨迹跟踪性能。

本文 PSO 算法加载在四旋翼的高度控制量和姿态控制量上，在 X、Y 轴的位置控制上并没有添加 PSO 算法，可以从图 4 看到，二者 (改进 PSO-RBF 神经网络自适应滑模控制器和 RBF 神经网络自适应滑模控制器) 和传统双闭环 PID 控制器的控制效果相差不多。

从图 5 看到，RBF 神经网络自适应滑模控制器和双闭环 PID 控制器在高度 (Z) 上的控制效果差于改进 PSO-RBF 神经网络自适应滑模控制器，在很短的时间内，改进 PSO-RBF 神经网络自适应滑模控制器下的四旋翼几乎可以无误差地跟踪上期望 Z 方向轨迹。在 10 s 定高的过程中，改进 PSO-RBF



(a) X轴轨迹跟踪对比

(b) Y轴轨迹跟踪对比

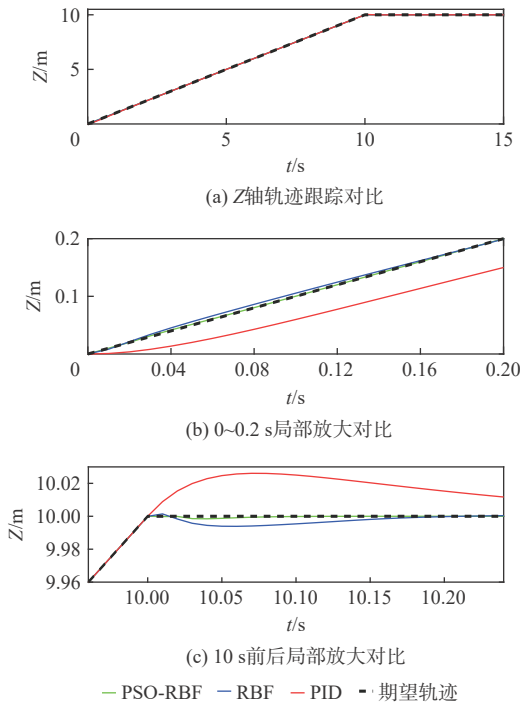
图 4 水平位置轨迹跟踪对比

Fig. 4 Comparison of horizontal position trajectory tracking

神经网络自适应滑模控制器下的四旋翼调节量少, 调节时间明显快于另外 2 种控制器。可以预想到, 本文控制器控制下的四旋翼应用于需要较高跟踪精度的场合 (如狭窄区域内穿越、搜救), 将具有更好的效果。

表 2 给出了 3 种控制器的初始轨迹跟踪调节时间和 10 s 定高飞行的调节时间, 可以更直观地对比 3 种控制器的轨迹跟踪性能。

图 6 给出了 3 种控制器在姿态角跟踪上的效果对比。可以看到, 改进 PSO-RBF 神经网络自适应滑模控制器比 RBF 神经网络自适应滑模控制器有更快的调节时间, 而双闭环 PID 控制器有明显的超调并伴随一定调节过程, 这并不是笔者所需要的。综上可以认为, 本文所设计的控制器具有良好的控制效果。



(a) Z轴轨迹跟踪对比

(b) 0~0.2 s局部放大对比

(c) 10 s前后局部放大对比

图 5　高度位置轨迹跟踪对比

Fig. 5　Comparison of height position trajectory tracking

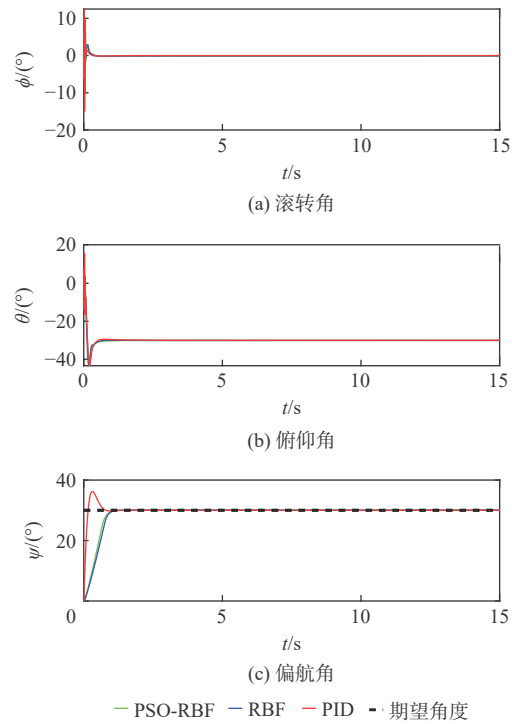
表 2　3 种控制器的调节时间

Table 2　Settling time of three kinds of controllers　s

控制器	初始轨迹调节时间	10 s定高调节时间
改进PSO-RBF神经网络自适应滑模控制器	0.08	0.1
RBF神经网络自适应滑模控制器	0.18	0.2
双闭环PID控制器	0.38	0.4

4.2　抗干扰性能对比

在 t=6 s 向四旋翼滚转轴、俯仰轴引入幅值为 0.9 N·m、周期为 0.5 Hz 的脉冲干扰, 偏航轴引入幅值为 0.09 N·m、周期为 0.5 Hz 的脉冲干扰, 在 t=10 s 时刻, 向四旋翼 x、y、z 轴引入随机白噪声模拟四旋翼飞行过程中受到的风扰等干扰, 分别对 3 种控制器进行仿真, 绘制 6～15 s 的姿态变化对比, 如图 7 所示。观察改进 PSO-RBF 神经网络自适应滑模控制器、RBF 神经网络自适应滑模控制器和双闭环



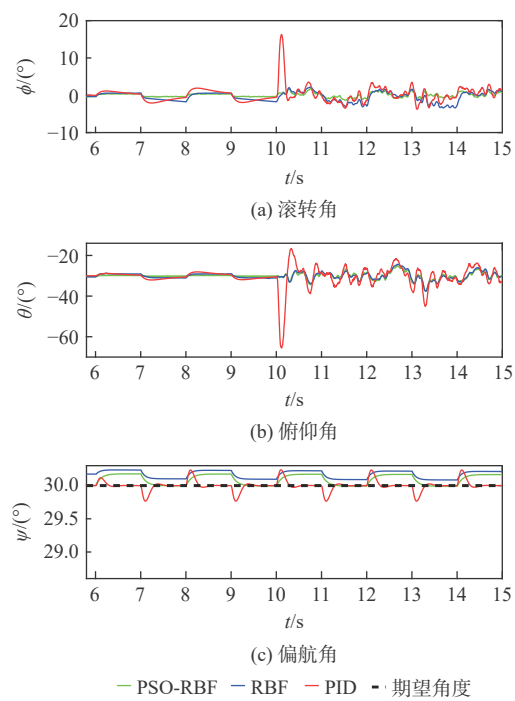
(a) 滚转角

(b) 俯仰角

(c) 偏航角

图 6　3 种控制器的姿态控制对比

Fig. 6　Comparison of attitude control of three kinds of controllers



(a) 滚转角

(b) 俯仰角

(c) 偏航角

图 7　引入干扰情况下 3 种控制器姿态控制对比

Fig. 7　Comparison of attitude control of three kinds of controllers when interference is introduced

PID 控制器的抗干扰性能。

通过图 7 可以看到，3 种控制器在有干扰的情况下都有一定的抗干扰能力，但是改进 PSO-RBF 神经网络自适应滑模控制器控制的四旋翼在应对干扰时，滚转、俯仰和偏航角度的变化都比另外 2 种控制器控制的四旋翼的角度变化幅度小。

为了验证 PSO 算法的抗干扰性，绘制了改进 PSO-RBF 神经网络自适应滑模控制器和 RBF 神经网络自适应滑模控制器控制的四旋翼姿态跟踪误差对比，如图 8 所示，观察改进 PSO 算法在四旋翼抗干扰性能上的表现。

通过图 8 可以看到，改进 PSO-RBF 神经网络自适应滑模控制器控制的四旋翼比 RBF 神经网络自适应滑模控制器控制下的四旋翼有更好的姿态角度跟踪能力，能在强干扰的情况下保持一定的跟踪精度，保证了四旋翼在轨迹跟踪上的稳定性。

因此，可以认为改进 PSO-RBF 神经网络自适应滑模控制器具有较好的抗干扰能力，保证强干扰情况下四旋翼按预定轨迹飞行的稳定性，在农业和军事侦察等方面具有一定的实用价值。

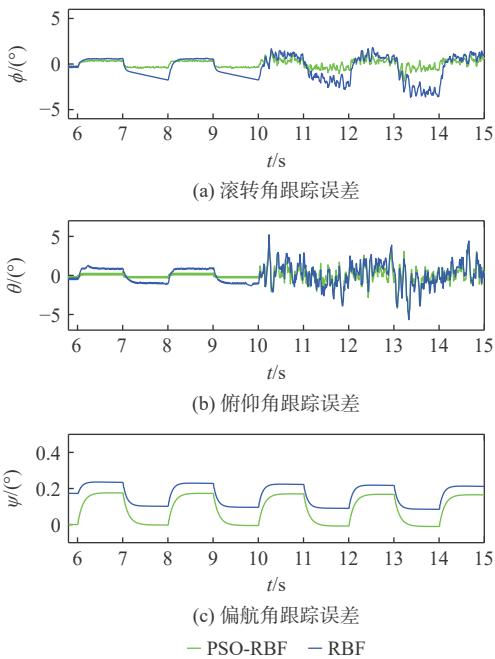


图 8 改进 PSO 算法对姿态角跟踪误差影响的对比

Fig. 8 Comparison of influence of improved PSO algorithm on attitude angle tracking error

4.3 鲁棒性对比

在 t=6 s 时刻减少 30% 的四旋翼质量来模拟四旋翼载物卸货的过程，分别对 3 种控制器进行仿真，绘制 6 s 前后 Z 轴方向高度变化对比和 6 s 前后控制量 U_1 变化对比，如图 9 所示，观察改进 PSO-RBF 神经网络自适应滑模控制器、RBF 神经网络自适应滑模控制器和双闭环 PID 控制器的鲁棒性。

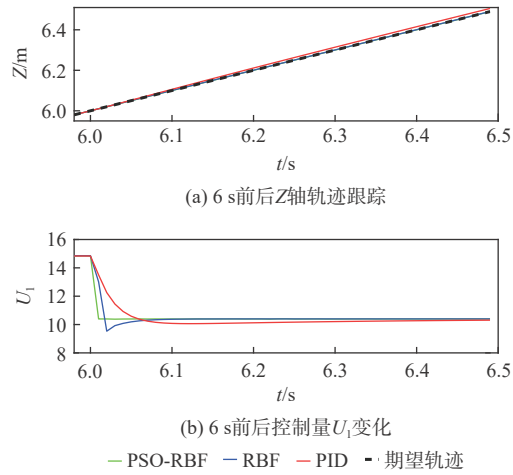


图 9 3 种控制器对 Z 轴的控制和产生的控制量 U_1 对比

Fig. 9 Control of Z-axis by three kinds of controllers and generated control quantity U_1

通过 Z 轴 (高度) 方向的对比可以看到，改进 PSO-RBF 神经网络自适应滑模控制器控制下的四旋翼由于其全局寻优能力，可以快速寻优出最佳神经网络拟合参数，在四旋翼质量突然减少的情况下也能很快减少 4 个旋翼产生的升力，稳定住四旋翼的高度。通过控制量 U_1 的对比也可以发现，改进 PSO-RBF 神经网络自适应滑模控制器比另外 2 种控制器能更快地找到合适的控制量，减少四旋翼的高度波动。

5 结 论

相对于 RBF 神经网络自适应滑模控制和双闭环 PID 控制器，改进 PSO-RBF 神经网络滑模自适应控制器可以在一个控制周期 (本文中是 0.01 s) 内寻找到合适的控制量，相对应的调节时间比 RBF 神经网络自适应滑模控制和双闭环 PID 控制器分别提升 50% 和 75% 左右，仿真结果也表明改进 PSO-RBF 神经网络滑模自适应控制方法具有调节时间短、跟踪误差小、抗干扰性能强、鲁棒性能好的优点。

这种特性使得改进 PSO-RBF 神经网络自适应滑模控制器具有在大干扰、狭小空间内通行的四旋翼上应用的前景。但是控制器仍然存在计算量较大的问题，下一步将持续探究神经网络结构和 PSO 算法的优化问题，并尝试将控制器部署到实际的四旋翼无人机上。

参考文献（References）

[1] FU J, CHAI T, SU C, et al. Motion/force tracking control of non-holonomic mechanical systems via combining cascaded design and backstepping[J]. Automatica, 2013, 49(12): 3682-3686.

[2] 李众, 沈炜皓. 四旋翼无人机的反步RBF网络自适应控制[J]. 指挥控制与仿真, 2020, 42(2): 89-94.
LI Z, SHEN W H. Backstepping sliding mode RBF network adap-

tive attitude control for quadrotor unmanned aerial vehicle[J]. Fire Control & Command Control, 2020, 42(2): 89-94(in Chinese).

[3] 韩业壮, 华容. 四旋翼飞行器的RBF网络自适应滑模控制[J]. 电光与控制, 2017, 24(11): 22-27.
HAN Y Z, HUA R. RBF neural network adaptive sliding mode control for quadrotor aerial vehicle[J]. Electronics Optics & Control, 2017, 24(11): 22-27(in Chinese).

[4] 陈骥. 基于双环PID的四旋翼飞行器自主飞行研究[J]. 机电一体化, 2018, 24(7): 22-27.
CHEN J. Research on autonomous flight of quadrotor based on double-loop PID[J]. Mechatronics, 2018, 24(7): 22-27(in Chinese).

[5] 马敏, 许中冲, 常辰飞, 等. 基于PID和LQR的四旋翼无人机控制系统研究[J]. 测控技术, 2016, 35(10): 42-45.
MA M, XU Z C, CHANG C F, et al. Research on four-rotor UAV control system based on PID and LQR[J]. Measurement & Control Technology, 2016, 35(10): 42-45(in Chinese).

[6] 刘慧博, 彭亮, 赵旭. 四旋翼飞行器姿态控制方法研究[J]. 自动化应用, 2020(12): 4-7.
LIU H B, PENG L, ZHAO X. Research on attitude control method of quadrotor[J]. Automation Application, 2020(12): 4-7(in Chinese).

[7] TAN L, LU L, JIN G. Attitude stabilization control of a quadrotor helicopter using integral backstepping[C]//International Conference on Automatic Control and Artificial Intelligence. London: IET, 2012.

[8] 姜洋. 四旋翼垂直起降机的鲁棒控制问题研究[D]. 哈尔滨: 哈尔滨工业大学, 2009.
JIANG Y. Research of four rotor vertical take-offs and landings aircraft robust control[D]. Harbin: Harbin Instituteof Technology, 2009 (in Chinese).

[9] 高福海, 胡盛斌, 王晨悦. 基于内外环的四旋翼飞行器的双模糊滑模控制[J]. 智能计算机与应用, 2020, 10(7): 158-165.
GAO F H, HU S B, WANG C Y. Double fuzzy sliding mode control of four rotor vehicle based on inner and outer rings[J]. Intelligent Computer and Applications, 2020, 10(7): 158-165(in Chinese).

[10] 王翼虎, 王思明. 基于改进粒子群算法的无人机路径规划[J]. 计算机工程与科学, 2020, 42(9): 1690-1696.
WANG Y H, WANG S M. UAV path planning based on improved particle swarm optimization[J]. Computer Engineering & Science, 2020, 42(9): 1690-1696(in Chinese).

[11] SESHAGIRI S, KHALIL H K. Output feedback control of nonlinear systems using RBF neural networks[J]. IEEE Transactions on Neural Networks, 2000, 11(1): 69-79.

[12] 孙瑞华, 郑重, 幸家正, 等. 基于RBF神经网络优化的四旋翼PID控制系统设计[J]. 自动化应用, 2018(5): 3-4.
SUN R H, ZHENG Z, XING J Z, et al. Design of quadrotor PID control system based on RBF neural networks[J]. Automation Application, 2018(5): 3-4(in Chinese).

[13] LIN X, WANG Y, LIU Y. Neural network-based robust terminal sliding-mode control of quadrotor[J]. Asian Journal of Control, 2022, 24(1): 427-438.

[14] 郝海燕, 王新军. 基于改进PSO-RBF滑模控制的光伏系统控制方法[J]. 电源技术, 2020, 44(8): 1148-1151.
HAO H Y, WANG X J. Method of tracking maximum power point of photovoltaic system based on improved PSO-RBF sliding mode control[J]. Chinese Journal of Power Sources, 2020, 44(8): 1148-1151(in Chinese).

[15] 付兴武, 胡洋. 基于改进粒子群算法的三维路径规划[J]. 电光与控制, 2021, 28(3): 86-89.
FU X W, HU Y. Three-dimensional path planning based on improved PSO algorithm[J]. Electronics Optics & Control, 2021, 28(3): 86-89(in Chinese).

[16] MAHONY R, KUMAR V, CORKE P. Multi-rotor aerial vehicles: Modeling, estimation, and control of quadrotor[J]. IEEE Robotics and Automation Magazine, 2012, 19(3): 20-32.

[17] 朱毅飞, 林德福, 莫雳, 等. 四旋翼无人机旋翼对机身非定常气动干扰特性[J]. 兵工学报, 2022, 43(2): 410-422.
ZHU Y F, LIN D F, MO L, et al. Transient aerodynamic interference between rotor and fuselage for a quadcopter[J]. Acta Armamentarii, 2022, 43(2): 410-422(in Chinese).

[18] 刘金琨. 滑模变结构控制MATLAB仿真[M]. 3版. 北京: 清华大学出版社, 2015: 15-28.
LIU J K. Sliding mode control design and MATLAB simulation[M]. 3rd ed. Beijing: Tsinghua University Press, 2015: 15-28 (in Chinese).

[19] 申珊颖. 斜十字四旋翼飞行器实验建模与控制技术研究[D]. 南京: 南京航空航天大学, 2013: 20-42.
SHEN S Y. Modelling based on experiment and control technology of an oblique cross quadrotor[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2013: 20-42(in Chinese).

[20] WANG X, CHEN Z, YANG G. Finite-time-convergent differentiator based on singular perturbation technique[J]. IEEE Transactions on Automatic Control, 2007, 52(9): 1731-1737.

Improved PSO-RBF neural network adaptive sliding mode control for quadrotor systems

TANG Zhiyong，MA Fuyuan，PEI Zhongcai*

(School of Automation Science and Electrical Engineering，Beihang University，Beijing 100191，China)

Abstract：An improved particle swarm optimization-radial basis function (PSO-RBF) neural network adaptive sliding mode controller is proposed for quadrotor systems with nonlinearity, strong coupling, and inaccurate interference. First, based on smooth improvement of the control amount of the RBF neural network sliding mode controller, an improved particle swarm optimization with global optimization capability was used to adjust the fitting parameters of the RBF neural network, thus improving the fitting ability of the network. Next, a dynamic model of quadrotor was built according to themodel parameters of actual quadrotors, the stability of which was then proved by Lyapunov theory.In contrast to the RBF neural network adaptive sliding mode controller and the double closed-loop PID controller, the improved PSO-RBF neural network adaptive sliding mode controller can find the appropriate control quantity in one control cycle, and its adjustment time is about 50% and 75% faster than that of RBF neural network adaptive sliding mode controller and double closed-loop PID controller, respectively. The simulation results show that the improved PSO-RBF neural network adaptive sliding mode controller featuresfasttrack speed with high accuracy, strong disturbance rejection and better robustness.

Keywords：quadrotor；particle swarm optimization；radial basis function neural network；adaptive sliding mode control；trajectory tracking；anti-interference

Received：2021-08-19；**Accepted**：2021-11-19；**Published Online**：2021-12-16 10：59

URL：kns.cnki.net/kcms/detail/11.2625.V.20211215.1408.002.html

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0500

液氮温区二维指向深冷环路热管设计与实验研究

李楠[1,2]，郭元东[3,*]，许程[3]，张红星[4]，林贵平[3]

(1. 南京航空航天大学 航空学院，南京 210016；2. 中国商飞民用飞机试飞中心 试飞运行部，上海 201323；
3. 北京航空航天大学 航空科学与工程学院，北京 100191；4. 北京空间飞行器总体设计部，北京 100094)

摘　　要：空间二维指向机构与红外探测器配合，有利于实现对空间目标大范围的动态追踪、指向、快速定位等功能。将深冷环路热管（CLHP）与二维指向机构耦合，可以大大降低系统机构的复杂程度，实现远距离热传输，提高探测精度和转向灵活性。为此，设计研制了液氮温区二维指向CLHP。对设计流程和部件参数进行了介绍，通过伺服电机驱动实现了俯仰、偏航±90°以上的转动。通过开展热真空实验，研究了不同工作参数对超临界启动和传热极限的影响。结果表明：所设计的系统具有最大13 W的传热能力，适当提高充装压力有利于提高系统的稳定性和传热能力，增大次蒸发器辅助载荷有助于提高最大传热能力。

关　键　词：环路热管；液氮；二维指向；超临界启动；稳态传热性能
中图分类号：V224
文献标志码：A　　**文章编号**：1001-5965(2023)07-1573-10

空间二维指向机构是一种搭载在航天设备上实现两自由度转动的专用空间结构，在各类遥感卫星与深空探测卫星上得到了广泛的应用。2020年，中国首次自主火星探测任务“天问一号”探测器被送入太空，标志着中国在深空探测领域的技术跨越，迈出了行星探测的第一步。火星环绕器上的火星矿物光谱分析仪可以通过远距离遥感获取火星表面的精细光谱图像[1]。矿物光谱分析仪中的红外传感器需要在73 K左右的深低温环境下工作。

空间二维指向机构可以实现两自由度的运动，与红外探测器配合，针对空间目标实现大范围的动态追踪、指向、快速定位等功能，在深空探测、对地探测等航天活动中发挥着日益重要的作用。传统的斯特林制冷机无法满足航天活动中快速增长的热收集要求，同时制冷机工作时引入的振动会大大影响红外探测器的工作精度，而且机械制冷机一般直接耦合在红外探测器上，大大增加了空间二维指向机构的质量。

与常温环路热管（loop heat pipe, LHP）类似，深冷环路热管（cryogenic loop heat pipe, CLHP）是一种基于毛细泵驱动流体产生气液相变原理的高效传输系统，具有传热高效、传输距离长、体积质量小等优点[2-7]。目前，关于CLHP的研究仍然处于实验室阶段，相关研究机构包括美国NASA[8-9]、法国原子能中心[10-11]、中国科学院理化技术研究所[12-15]、南京理工大学[16]、中国科学院上海技术物理研究所[17-18]、北京航空航天大学[19-25]及中国航天科技集团有限公司第五研究院[26-27]等团队，通过不同的制冷方式和工质，将CLHP的工作温区拓展到4 K以下温区[28]。

将红外探测器安装于二维指向机构上，机械制冷机安装于二维指向机构外，利用CLHP将二者连接，可以实现热量的跨转向轴传输，既可以使得红外探测器处于深低温工作温区，又可以实现低温制冷机与红外探测器的分离，有效降低了机械制冷机高频振动对红外探测器的影响，提高了探测精度与

收稿日期：2021-08-30；**录用日期**：2021-11-12；**网络出版时间**：2021-11-30 14:36
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20211129.1705.003.html
基金项目：国家自然科学基金(52106067)；卓越青年科学基金(2020-JCJQ-ZQ-042)；中国博士后科学基金(2019M660403)
*通信作者. E-mail：guoyd@buaa.edu.cn

引用格式：李楠，郭元东，许程，等. 液氮温区二维指向深冷环路热管设计与实验研究[J]. 北京航空航天大学学报，2023，49(7)：1573-1582. LI N，GUO Y D，XU C，et al. Design and experiment of cryogenic loop heat pipe of two-dimensional pointing at liquid nitrogen zone[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1573-1582 (in Chinese).

转向灵活性。关于二维指向 CLHP 的研究，仅有 Hoang 等[29]开展了初步的实验研究，传输管线长度 4.3 m，最大传热能力约为 5 W，仍然存在一定的提升空间。本文针对液氮温区的二维指向需求，设计研制了氮工质 CLHP 原理样机，研究了工作参数对超临界启动和传热极限的影响，介绍了基于 CLHP 的柔性二维指向系统设计与系统性能测试结果。

1 系统参数设计与校核

1.1 设计目标

针对二维指向 CLHP 的工程应用背景，实现红外探测器与低温制冷机的“分离”，要求设计的 CLHP 实现以下技术指标：

1）顺利实现 80～300 K 的超临界启动过程。

2）当 CLHP 运行在 80 K 温区时，传递 5 W 以上的热载荷。

3）CLHP 可以适应二维指向机构在俯仰、偏航 2 个方向上± 90°转动。

除此以外，为提高红外探测器的探测精度，所设计的 CLHP 需满足以下要求：体积小、质量轻、转动变形时所需的扭矩小；内部的工质具有较高的安全性；运行温度平稳，具有较强的适应能力。

1.2 工质选择

环路热管内的工质物性参数决定了环路热管系统的运行温度区间、传热极限、适应能力等性能指标。

首先，确定 CLHP 的目标工作温区区间。在此区间内，工质必须处于气液两相状态，即工质的临界温度应高于温区上限，工质的凝固温度应低于温区下限，否则工质可能处于超临界状态或固体状态，导致 CLHP 无法正常运行。在 80～100 K 区间内，满足气液状态条件的工质一般为氧气与氮气，而氧气的温区范围更广。

然后，需要考虑工质与部件材料的相容性。工质不能与毛细芯、管线、壳体等发生化学反应，以保证环路热管的正常运行与使用寿命。考虑材料的相容性及加工装配等因素，选择不锈钢 1Cr18Ni9Ti 作为壳体和管路材料，镍粉为原材料制作蒸发器毛细芯，鞍座选用铜鞍座。氮气作为工质，性能稳定，一般不会与金属发生反应，满足条件。

最后，需要综合考虑工质的热物理性质，使 CLHP 的传热性能与超临界启动特性更优。通常，采用 Dunbar 因数 N 来综合评价工质能否作为热管工质的指标，其根据回路内的压力平衡公式推导得到，可用来表征环路热管的传热能力，计算公式如下：

$$N=\frac{\sigma\rho_{\mathrm{v}}h_{\mathrm{fg}}^{1.75}}{\mu_{\mathrm{v}}^{0.25}} \tag{1}$$

式中：σ 为工质的表面张力，可以提高毛细芯的抽吸力；h_{fg} 为气化潜热，较高的气化潜热可以增加 CLHP 的传热量；ρ_{v} 为气体密度；μ_{v} 为气体动力黏度系数。

图 1 展示了 80～100 K 温区内氮气和氧气的 Dunbar 因数对比。

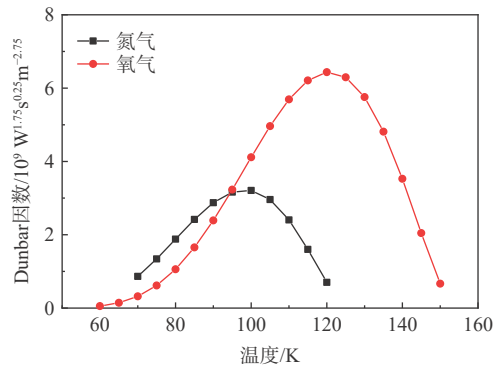


图 1 80～100 K 温区工质的 Dunbar 因数变化曲线

Fig. 1 Variation curves of Dunbar factor of different working fluids at 80−100 K

$\mathrm{d}P/\mathrm{d}T$ 是另一个评价工质性能的指标，比值越大，说明该工质饱和压力随温度变化越大，启动与运行性能更好。

$$\Delta T\left.\frac{\mathrm{d}P_{\mathrm{s}}}{\mathrm{d}T_{\mathrm{s}}}\right|_{T}=\sum\Delta P \tag{2}$$

式中：ΔT 为储液器与蒸发器毛细芯之间的蒸汽温差；ΔP 为储液器与蒸发器毛细芯之间的蒸汽压降；P_{s} 和 T_{s} 分别为饱和压力和饱和温度。

在环路热管启动的过程中，工质的 $\mathrm{d}P/\mathrm{d}T$ 比值越大，储液器与蒸发器之间可以以相同的温差产生越大压头，使工质更易循环流动。在环路热管运行时，工质的 $\mathrm{d}P/\mathrm{d}T$ 比值越大，表明工质可以运行在更低的温度，具有更低的热阻，传热性能更加优异。通过筛选对比，在 80 ～100 K 温区内，有氮气和氧气可选，由图 2 对比可知，氮气的 $\mathrm{d}P/\mathrm{d}T$ 比值高于氧气，同时 Dunbar 因数更高，综合考虑安全性能等因素，最终选取氮作为 CLHP 的工质。

1.3 部件设计

1.3.1 传输管线

传输管线主要包括气体管线、液体管线等。传输管线一般采用小管径的金属管，管径的大小对寄生漏热、传热极限有较大的影响。当减小传输管线直径时，系统的体积与质量可以减小，管线的外表面积减小，来自周围环境的寄生漏热也随之减小。但是，减小管径会引起工质在内部的流动阻力升高，降低主蒸发器的传热极限，但同时系统柔性较

强。为使 CLHP 适应二维指向运行, 实现传输管线的“柔性”功能, 传输管线结构主要包括 2 类方案: ①利用金属本身的延展性来实现弯曲变形; ②利用波纹管来实现自由收缩、弯曲变形等功能。综合考虑扭矩大小与二维指向机构的复杂程度, 采取方案①, 即将传输管线制作为螺旋盘管, 螺旋盘管的匝数越多, 可以转动的角度越大。表 1 为传输管线的结构参数, 在俯仰与偏航方向都采用螺旋盘管形式。

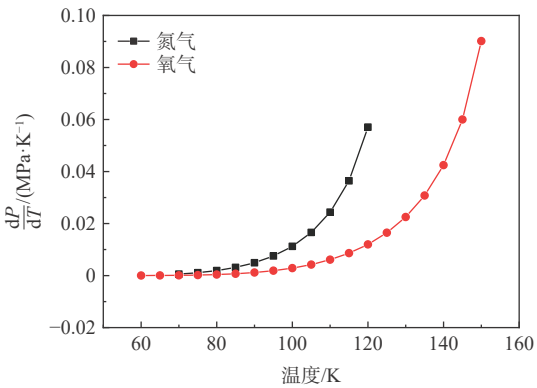


图 2　80～100 K 温区工质的dP/dT变化曲线

Fig. 2　Variation curves of dP/dT of different working fluids at 80−100 K

表 1　传输管线结构参数

Table 1　Structure parameters of transport line

部件	长度/mm	外径/mm	壁厚/mm	俯仰螺旋直径/mm	俯仰螺旋匝数	偏航螺旋直径/mm	俯仰螺旋匝数
主气体管线	10 000	3	0.5	250	6	350	4
主液体管线	10 000	3	0.5	250	6	350	4
次液体管线	10 000	3	0.5	250	6	350	4

1.3.2　蒸发器和储液器

蒸发器是 CLHP 的核心功能部件, 主要起到从热源吸收热量并为工质提供驱动力的作用。蒸发器采用圆柱构型设计, 由蒸发器壳体、毛细芯、液体引管组成。主蒸发器的尺寸主要由传热功率与工质参数确定, 次蒸发器的尺寸主要由超临界启动性能确定, 并尽可能小型化。毛细芯材质为镍粉压制而成, 最大孔径为 0.5 μm, 孔隙率为 60%, 渗透率优于 7×10^{-14} m^2。储液器位于蒸发器的一端与蒸发器直接相连, 保证对蒸发器毛细芯的液体供应。根据环路热管在不同工作状态、不同气液分布下的质量守恒设计储液器的容积。具体结构参数如表 2 所示。

1.3.3　冷凝器

冷凝器包括主冷凝器与次冷凝器, 根据现有的加工工艺与设计经验, 冷凝器的设计沿用经典的平板内嵌流道结构。冷凝器长度的设计原则为: 当

表 2　蒸发器与储液器结构参数

Table 2　Structure parameters of evaporator and compensation chamber

部件	参数	数值
主蒸发器	壳体长度×外径×厚度/(mm×mm×mm)	110×13×1
	毛细芯外径×长度/(mm×mm)	11×100
主储液器	容积/mL	5
轴向槽道	槽深×槽宽×长度/(mm×mm×mm)	0.8×0.8×80
	数量/个	4
次蒸发器	壳体长度×外径×厚度/(mm×mm×mm)	100×13×1
	毛细芯外径×长度/(mm×mm)	11×90
次储液器	容积/mL	5
轴向槽道	槽深×槽宽×长度/(mm×mm×mm)	0.8×0.8×80
	数量/个	4

CLHP 运行达到设计极限时, 工作温度依然维持在允许的温度范围内。假设冷凝器入口处为过热蒸气, 随着工质在冷凝器内冷却, 逐渐消除过热、冷凝变成过冷液体。根据能量平衡, 考虑主蒸发器载荷、次蒸发器载荷及周围环境的寄生漏热, 同时设置一定的过冷段比例, 主/次冷凝器参数最终分别为 1 160 mm×3 mm×0.5 mm 和 680 mm×3 mm×0.5 mm。

1.3.4　液氮冷板

实验测试系统的热沉由液氮冷板提供。为保证热沉冷却效果, 使冷凝器温度尽可能维持稳定与低温, 要求设计的冷板具有良好的换热效果与较小的流动阻力。为增强液氮冷板的换热效果, 材质选用导热性能优异的紫铜。对液氮冷板结构进行优化设计, 如图 3 所示, 采用液氮液池的结构, 保证冷板的温度稳定性和均匀性, 同时设置针肋以提高冷板的结构强度。经过测试, 液氮冷板具有 40 W 的稳定传热特性。

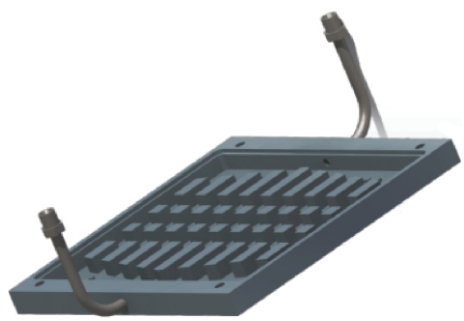


图 3　液氮冷板结构示意图

Fig. 3　Schematic diagram of liquid nitrogen cooling plate

1.3.5　储气室与充装压力

储气室与工质充装压力的基本原则为: 保证系统的安全性, 并保证 CLHP 梳理实现超临界启动。因此, 二者的设计对超临界启动过程尤为重要。结

合相关设计准则[2]，为满足供液条件，设计得到的储气室范围为$V_{cg} \geqslant 1\,000$ mL。为进一步保证系统的安全性，取保守设计，如表3所示，储气室的容积选为1 200 mL，300 K时的充装压力约为2.8～3.6 MPa。

表3 储气室容积与工质充装范围

Table 3 Gas reservoir volume and working fluid filling range

储气室容积/mL	充装质量/g	300 K时充装压力/MPa
1 200	45～65	2.8～3.6

1.3.6 U型转台与驱动部件

如图4所示，俯仰和偏航方向分别利用旋转台来实现转动，以实现准确的角度调节。电动旋转台通过电机与蜗杆机构进行驱动，可以实现在360°范围内的调节，转动精度可以达到0.01°。2个正交轴的旋转部分通过U型臂来实现。将偏航方向旋转台的转台部分固定在偏航U型臂底部，通过旋转台带动整个系统转动。俯仰U型臂安装在偏航U型臂的上半部分，而俯仰角的旋转台固定在小U型臂侧边。2个U型臂分别固定2套螺旋管线，同时考虑到旋转台和系统尺寸，俯仰U型臂主要用来固定主蒸发器，2个方向的旋转台可以分别控制系统在2个维度上进行单独转动，也可以实现2个方向上的组合运动。结果表明，系统可以实现俯仰、偏航±145°的转动。

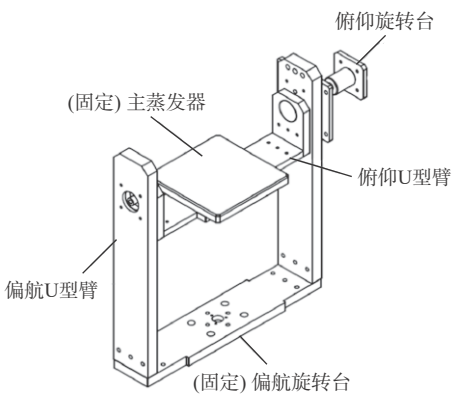


图4 二维指向机构U型臂结构示意图

Fig. 4 Schematic diagram of two-dimensional pointing mechanism U-shaped structure

2 实验系统与实验件

为模拟CLHP在空间中真空、冷黑工作环境下的工作特性，在地面热真空系统中开展测试。实验系统主要包括KM-1.5级真空舱、真空系统、热源和热沉系统、工质充装系统、数据采集系统等，系统真空度可维持在1×10^{-3} Pa以下。为减小CLHP与周围环境的辐射与分子导热换热，二维指向CLHP所有部件包覆10单元的多层隔热材料，大大降低了装置的当量热导率与当量发射率。为了减少周围环境对环路热管的漏热，热真空实验舱内设置绝热冷屏，可以有效降低周围环境的寄生漏热。热源系统采用薄膜型加热片作为模拟热源，布置于主蒸发器、次蒸发器。热沉系统的作用是为冷凝器提供足够的冷量，实验采用液氮冷板作为热沉，其中的介质温度可以稳定在85 K以下。图5为真空舱内包覆隔热材料后的实验件照片。

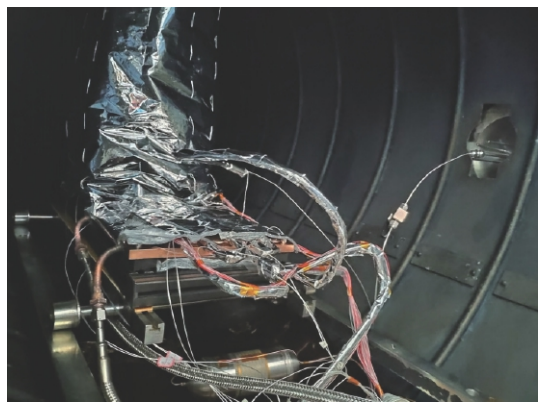


图5 实验系统与二维指向CLHP实验件

Fig. 5 Experimental system and two-dimensional pointing CLHP test unit

2.1 工质充装系统

图6为本套环路热管的工质充装系统，主要由真空泵与高压气瓶组成。在充装工质前，先对环路热管系统抽真空，通过控制阀门，将环路热管的压力调整至设计充装压力。整个充装过程在常温下进行。

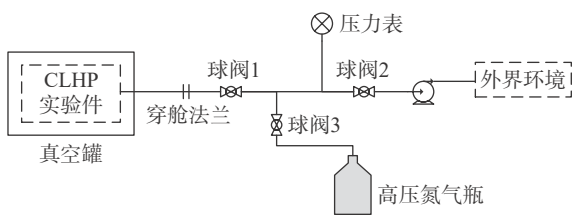


图6 环路热管工质充装系统示意图

Fig. 6 Schematic of working fluid filling system for loop heat pipe

2.2 数据采集系统

实验通过T型热电偶实时采集CLHP器件的表面温度，精度为±0.1 ℃，CLHP整体位于真空舱内，且外表面包覆有多层材料，可以忽略与周围环境的对流、辐射漏热。由于不锈钢管管壁较薄，厚度为0.5 mm，且轴向长度远远大于管壁厚度，可以将测点温度视为系统内部工质温度，温度数据通过Aglient 34970采集。热电偶具体分布如图7所示。系统压力通过数字式压力表实时采集，压力表位于储气室出口位置，将此处压力视为系统内部压力。

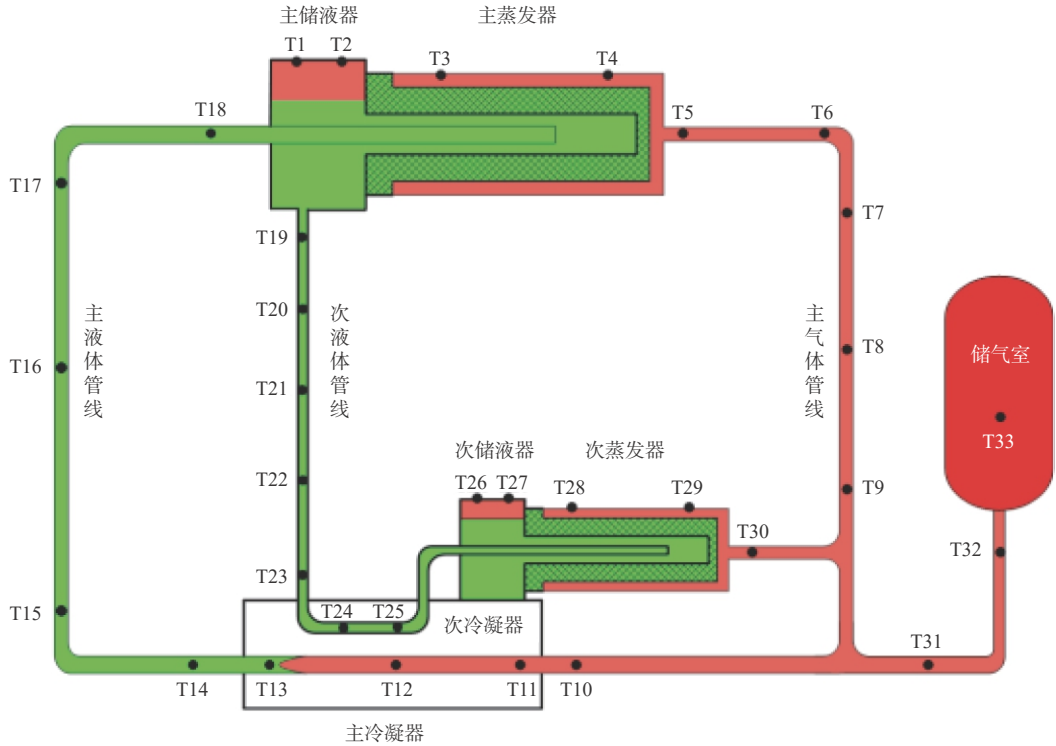


图 7　CLHP 热电偶分布与压力表位置示意图

Fig. 7　Schematic of thermocouple distribution and pressure gauge location on CLHP

压力传感器的量程为 0～4 MPa, 测量精度为 0.1 级。

3　实验结果与讨论

3.1　二维指向 CLHP 超临界启动特性

3.1.1　典型工况超临界启动过程

CLHP 的启动过程一般可以分为 3 个过程: ①冷端降温, 次蒸发器被液体完全浸润; ②热端降温, 施加辅助载荷使主蒸发器逐渐被液体浸润; ③施加主蒸发载荷主回路运行。

图 8 为向冷板通液氮后 CLHP 冷端部件主要特征点的降温过程, 包括主冷凝器、次冷凝器、次储液器、次蒸发器等, 系统的初始充装压力 P_{ch} 为 3.57 MPa。通入液氮前, 所有温度均为 297 K; 通入液氮后, 冷凝器、次蒸发器率先降温。由于冷板与液氮间巨大的温差, 液氮发生汽化, 将冷端部件的热量带走, 温度快速下降, 约 10 min 后, 冷板进出口温度及表面温度从 290 K 快速下降到 109 K, 低于氮气的临界温度。之后, 由于冷板与液氮的温差减小, 换热的主要形式是单相流动换热, 冷板温度降低趋势减缓, 100 min 内温度降低到 89 K。在冷板降温过程中, 冷板表面温度间最大温差为 15 K。CLHP 主、次冷凝器与冷板采用螺钉紧密贴合, 次储液器通过鞍座导热实现热量交换, 中间涂有低温导热脂以降低接触热阻。155 min 时, 次储液器和次蒸发器温度变化趋势趋于缓和, 分别维持在 100 K 和 105 K

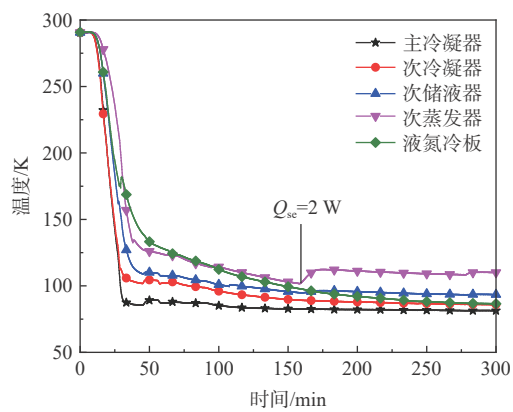


图 8　CLHP 冷端降温过程(Q_{se}=2 W, P_{ch}=3.57 MPa)

Fig. 8　Cold end cooling process of CLHP (Q_{se}=2 W, P_{ch}=3.57 MPa)

附近, 毛细芯完全被液体浸润, 次回路系统具备启动的条件。对次蒸发器施加启动载荷 Q_{se} 为 2 W, 传输管线和主蒸发器开始降温。在随后的测试中, 冷板表面温度趋于稳定, 冷板工作表面各特征点之间温差小于 5 K, 主、次冷凝器特征点之间温差小于 2 K。

图 9 为对 CLHP 次蒸发器施加载荷后主储液器、主蒸发器等热端部件的降温过程。在 CLHP 启动的过程中, 热真空实验舱内的液氮保护屏可以起到辅助降温的作用。当向液氮保护屏输送液氮时, 液氮保护屏的温度在 100 min 后降低到 85 K 并保持稳定。当次蒸发器温度低于 105 K 并维持 30 min 后, 向次蒸发器施加 2 W 载荷, 次蒸发器的温度升

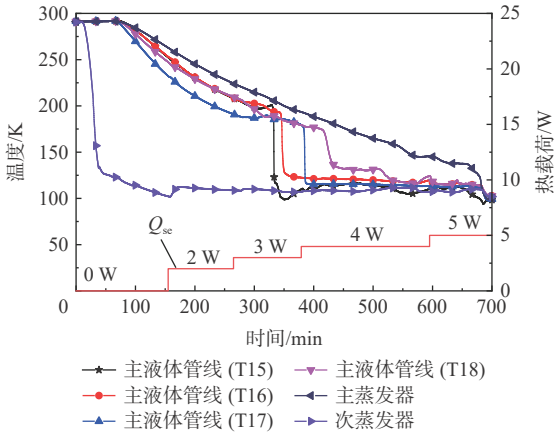


图 9 CLHP 热端降温过程(P_{ch}=3.57 MPa)

Fig. 9 Hot end cooling process of CLHP (P_{ch}=3.57 MPa)

高 10 K 到 110 K 后保持不变,但此时 CLHP 其他温度特征点未发生明显变化。表明次蒸发器载荷较小,不足以抵消环境漏热的影响,无法驱动足够的工质流动。随后增大载荷为 3 W,主冷凝器入口处温度升高,主液体管线温度突降。主冷凝器内的液体工质沿着主液体管线向主蒸发器方向移动,工质在次回路中实现了正向循环。同时表明此时次蒸发器热载荷 3 W 大于周围环境对系统的漏热,即系统整体的漏热处于 2～3 W 的水平。随着次蒸发器载荷的不断增大,沿着主液体管线从主冷凝器到主蒸发器方向的温度特征点温度依次发生突降。主冷凝器出口的工质温度约为 85 K,流经 250 K 左右的主液体管线时由于存在巨大的温差会迅速发生气化,导致管线温度快速降低。主液体管线上相邻温度特征点间距为 2.5 m,温度突降的时间间隔约为 38 min。在次蒸发器运行约 540 min 后,主蒸发器温度快速降低到 97 K 并趋于稳定,主蒸发器内的毛细芯被液氮全部浸润。

图 10 为管线降温过程中的沸腾曲线[30]。在初始降温阶段,管壁与内部工质的温差巨大,贴近管壁的液体发生剧烈沸腾,而产生的蒸汽尚未离开壁面,附着在管壁上形成一层气膜,此阶段为膜态沸腾。在此阶段中,壁面与工质之间的换热由于气膜的低导热率而阻断,换热系数小,因此管线降温缓慢。而随着管壁温度的降低,壁面过热度减小,管壁附近的气泡不断脱离壁面,此阶段为过渡沸腾。随着壁面温度进一步降低,气膜基本全部消失,壁面与液体工质直接接触,此阶段为核态沸腾。在此阶段中,换热系数高,管线的温度快速降低,表现为温度的突降过程,与实验中的现象相符。因此,常温管线的降温过程一般都会经历温度下降缓慢且时间较长的膜态沸腾阶段、温度快速下降的核态沸腾阶段,直至完全降温的单相流动阶段。

当主储液器内的液体含量较高,主蒸发器内的毛细芯被液氮全部浸润后,主回路系统就具备启动的条件。图 11 为 CLHP 在降温结束后对主蒸发器施加载荷,主回路部件主要特征点的温度变化。主蒸发器温度维持 87 K 约 30 min,对主蒸发器施加 1 W 热载荷 Q_{pe},主蒸发器温度与主气体管线温度缓慢升高后保持不变,表明 CLHP 超临界启动成功。

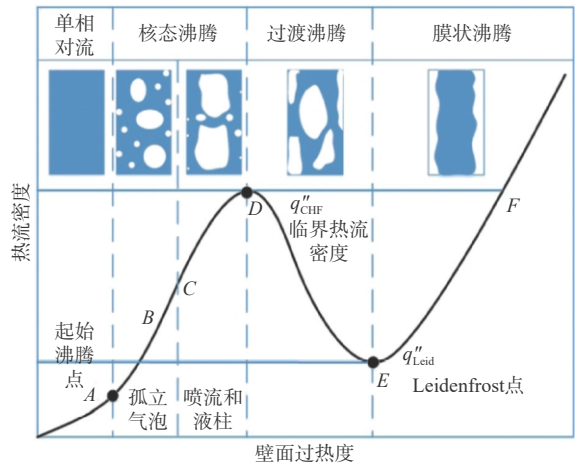


图 10 管线降温过程沸腾曲线[30]

Fig. 10 Boiling curve of pipe cooling process[30]

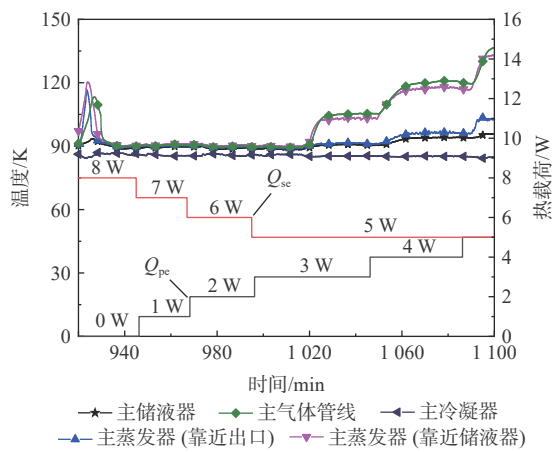


图 11 CLHP 主回路启动过程(P_{ch}=3.57 MPa)

Fig. 11 Main loop startup process of CLHP (P_{ch}=3.57 MPa)

3.1.2 充装压力的影响

CLHP 初始状态均为常温状态,研究不同充装量(3.10 MPa、3.29 MPa、3.57 MPa)对 CLHP 对超临界启动过程的影响。充装压力的范围主要保证在 CLHP 降温完成后主蒸发器、次蒸发器内充满液体的情况下,系统的压力小于临界压力,使得内部的工质可以顺利完成相变。同时,降温完成后的压力需要大于冷端部件温度对应的饱和压力,否则冷凝器内工质将为过热气体,无法产生冷凝液体。图 12 和图 13 分别表示了不同充装压力下(3.29 MPa、3.10 MPa)二维指向 CLHP 启动过程中温度变化情

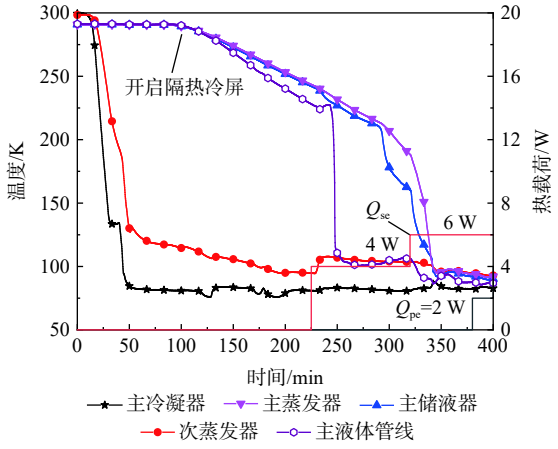


图 12　CLHP 超临界启动过程(P_{ch}=3.29 MPa)

Fig. 12　Supercritical startup of CLHP (P_{ch}=3.29 MPa)

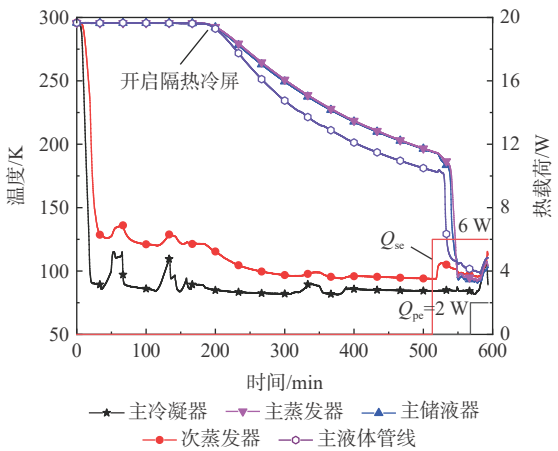


图 13　CLHP 超临界启动过程(P_{ch}=3.10 MPa)

Fig. 13　Supercritical startup of CLHP (P_{ch}=3.10 MPa)

况。从整体趋势来看, 超临界启动过程的温度测点变化规律基本相同, 均能完成启动过程。

表 4 为 CLHP 超临界启动后系统压力随不同充装压力的变化情况。结果表明, 在所测试的充装压力范围内, 系统启动压力均低于液氮临界压力同时高于热沉温度对应的饱和压力时, 均可以满足启动的基本条件。随着充装压力的提高, 主蒸发器的运行温度及系统的启动压力也会随之升高。当充装压力过高时, 主蒸发器的运行温度较高, 可能导致运行不稳定。当充装压力较低时, 系统压力可能低于冷端对应的压力, 导致冷凝器无法产生液体从而启动失败。

表 4　不同充装压力下 CLHP 启动完成时主蒸发器温度及系统压力

Table 4　Primary evaporator temperature and system pressure under different filling pressures after startup

状态	温度/K	压力/MPa
临界状态	126	3.36
冷端温度及饱和压力	85	0.228
充装压力3.10 MPa	93.1	0.279
充装压力3.29 MPa	91.9	0.499
充装压力3.57 MPa	99.4	0.567

3.1.3　逆重力状态启动

将二维指向 CLHP 处于逆重力 60 cm 状态的超临界启动过程。向次蒸发器施加 5 W 热载荷使热端开始降温。10 h 后, 主液体管线只有 3/4 完成降温且温度趋于平缓。将冷凝器内的 90 K 液氮驱动至 60 cm 高的主蒸发器时, 需要额外克服 4 kPa 重力压降。同时, 主液体管线在降温过程中, 内部为两相流动, 会产生加速压降, 增加沿程阻力。而此时次蒸发器运行温度约 115 K, 可以提供的最大毛细力相对较低, 使得 CLHP 无法实现逆重力 60 cm 高度的超临界启动。

3.2　二维指向 CLHP 稳态传热特性

主蒸发器的传热极限作为 CLHP 的重要性能指标之一, 其含义是系统可以传输的最大热量。

3.2.1　典型工况的传热极限

当 CLHP 完成超临界启动后, 对主蒸发器施加载荷 Q_{pe} 可以实现主回路的循环。逐渐增大施加主蒸发器的热载荷, 直到主蒸发器温度升高无法平衡, 即为传热极限。图 14 为充装压力为 3.57 MPa、次蒸发器辅助载荷 Q_{se} 为 10 W 时传热极限的测试曲线。在测试过程中, 主冷凝器温度基本维持在 87 K 左右, 主蒸发器施加 4 W 载荷后, 系统的压力迅速升高到 0.37 MPa, 对应饱和温度 89 K。主储液器与主冷凝器温度均提高 3 K 后达到稳定, 表明主蒸发器内的工质受热产生低温蒸气, 导致系统压力与蒸发器温度升高。待主蒸发器达到稳定后, 逐渐增大施加在主蒸发器表面的载荷, 系统压力同步升高, 随后温度与压力达到稳定。当主蒸发器施加载荷达到 14 W 后, 主蒸发器表面温度从 120 K 快速

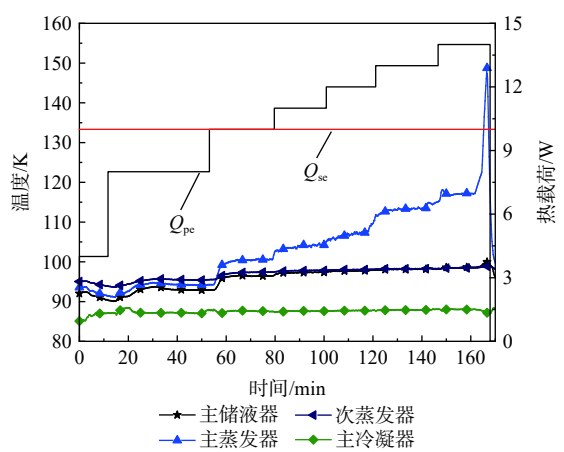


图 14　CLHP 主蒸发器传热极限(Q_{se}=10 W, P_{ch}=3.57 MPa)

Fig. 14　Heat transfer capacity of primary evaporator of CLHP (Q_{se}=10 W, P_{ch}=3.57 MPa)

升高到 150 K，系统运行失效，即充装压力 3.57 MPa、次蒸发器辅助载荷 10 W 时，系统最大传热能力约为 13 W。

3.2.2 充装压力的影响

图 15 为不同充装压力、固定次蒸发器热载荷 10 W 时的主蒸发器传热极限。可以看出，在所测试的充装压力范围内，随着充装压力增加，主蒸发器的传热极限载荷也随之增大。当充装压力较小时，主储液器内的液体含量较少，不能保证主蒸发器的液体含量。尤其是本套二维指向 CLHP 传输距离达到 10 m，液体从冷凝器传输到主蒸发器的传输时间较长，更容易出现蒸发器内液体工质含量不足的现象，最终导致环路热管运行失效。而当环路热管内的充装工质较多时，主储液器内的液体可以持续为主蒸发器内提供液体，更有利于保证系统的稳定运行。因此，在保证 CLHP 冷端降温完成时，系统压力低于临界压力，可以实现超临界启动前提下，适当增大环路热管的充装压力，保证主蒸发器的液体供应，可以提高主蒸发器的传热极限。

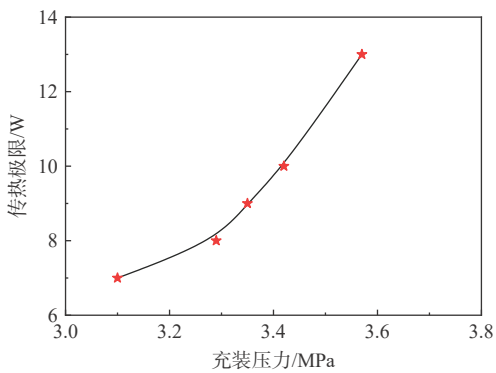


图 15 充装压力对二维指向 CLHP 主蒸发器传热极限的影响（Q_{se}=10 W）

Fig. 15 Effect of filling pressure on heat transfer capacity of CLHP primary evaporator of two-dimensional pointing (Q_{se}=10 W)

3.2.3 辅助载荷的影响

由于 CLHP 运行温度低于周围环境，系统会吸收来自周围环境的漏热。次蒸发器热载荷可以驱动次回路运行，携带低温液体抵消环境的寄生漏热，维持整体系统的稳定。针对二维指向 CLHP，图 16 为同一充装压力（3.57 MPa）、不同辅助载荷条件下的主蒸发器的传热极限测试结果。从趋势来看，当增大次蒸发器的辅助载荷，主蒸发器可以承受的极限载荷也随之增大。表明次蒸发器的辅助热载荷增大会使主冷凝器与主蒸发器内的质量流量增加，主蒸发器内的内供液更加充足，不易发生烧干或击穿，系统更加稳定，使得主蒸发器的传热极限增大。在实际空间应用中，受限于热沉制冷量和空间资源，应在满足传热需求的情况下，尽可能减小次蒸发器的辅助载荷。

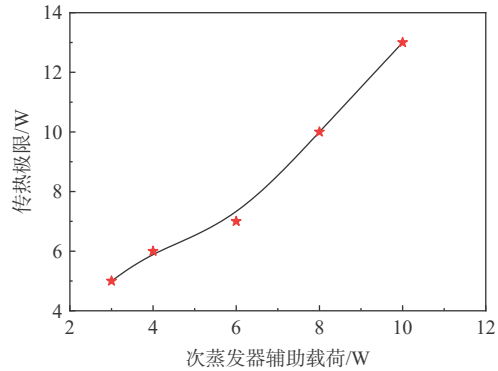


图 16 辅助载荷对二维指向 CLHP 传热极限的影响（P_{ch}=3.57 MPa）

Fig. 16 Effect of auxiliary heat load on heat transfer capacity of two-dimensional pointing CLHP (P_{ch}=3.57 MPa)

4 结 论

针对空间红外探测系统需要在轨转动的需求，将 CLHP 二维指向机构耦合，既能使红外探测器处于正常工作所需温度，又可以实现低温制冷机与红外探测器的分离，有效降低机械制冷机高频振动对红外探测器的影响，提高探测精度与转向灵活性。通过实验研究得出如下结论：

1）实验系统管线总传输距离长达 30 m，是现有文献中较长的 CLHP 样机，可实现俯仰/偏航±145°的转动。

2）长距离传输管线导致系统降温时间增加，需要设置隔热冷屏，减小周围环境的寄生漏热，漏热量在 2～3 W 的水平。

3）系统超临界启动过程中，管线逐次降温，可通过逆向流动沸腾换热曲线进行解释，与实验现象相符。

4）系统可以 13 W 左右的热量传输，适当增大充装压力和辅助载荷有利于提高系统的传热能力。

参考文献（References）

[1] 朱岩, 白云飞, 王连国, 等. 中国首次火星探测工程有效载荷总体设计[J]. 深空探测学报, 2017, 4(6): 510-514.
ZHU Y, BAI Y F, WANG L G, et al. Integral technical scheme of payloads system for Chinese Mars-1 exploration[J]. Journal of Deep Space Exploration, 2017, 4(6): 510-514(in Chinese).

[2] 刘成, 谢荣建, 王仕越, 等. 丙烯环路热管补偿器的可视化实验研究[J]. 北京航空航天大学学报, 2020, 46(5): 933-940.
LIU C, XIE R J, WANG S Y, et al. Visualization experimental study of compensation chamber of a propylene loop heat pipe[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(5): 933-940(in Chinese).

[3] 孟庆亮, 杨涛, 于志, 等. 空间遥感器用环路热管瞬态数值模拟与

在轨验证[J]. 北京航空航天大学学报, 2020, 46(11): 2045-2055.
MENG Q L, YANG T, YU Z, et al. Transient numerical simulation and on-orbit verification of loop heat pipe used for space remote sensor[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(11): 2045-2055(in Chinese).

[4] 薛好, 谢永奇, 戴华, 等. 加速度场中环路热管温度波动现象分析[J]. 北京航空航天大学学报, 2016, 42(7): 1502-1508.
XUE H, XIE Y Q, DAI H, et al. Investigation of temperature fluctuation in loop heat pipe under acceleration fields[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1502-1508(in Chinese).

[5] 郭元东. 15 K～40 K温区深冷环路热管技术的理论与实验研究[D]. 北京: 北京航空航天大学, 2019.
GUO Y D. Theoretical and experimental investigation of cryogenic loop heat pipe technology operating in 15 K～40 K[D]. Beijing: Beihang University, 2019(in Chinese) .

[6] BAI L, ZHANG L, LIN G, et al. Development of cryogenic loop heat pipes: A review and comparative analysis[J]. Applied Thermal Engineering, 2015, 89: 180-191.

[7] 何江, 苗建印, 张红星, 等. 航天器深低温热管技术研究现状及发展趋势[J]. 真空与低温, 2018, 24(1): 1-8.
HE J, MIAO J Y, ZHANG H X, et al. Current status and development trend of cryogenic heat pipe technologies in spacecraft[J]. Vacuum and Cryogenics, 2018, 24(1): 1-8(in Chinese).

[8] HOANG T T, O'CONNELL T A, KU J, et al. Performance demonstration of hydrogen advanced loop heat pipe for 20-30 K cryocooling of far infrared sensors[C]//Optics and Photonics 2005. Bellingham: SPIE, 2005: 590410.

[9] HOANG T T, O'CONNELL T A, KU J, et al. Design optimization of a hydrogen advanced loop heat pipe for space-based IR sensor and detector cryocooling[C]//SPIE's 48th Annual Meeting of Optical Science and Technology. Bellingham: SPIE, 2003: 86.

[10] MO Q, LIANG J. Operational performance of a cryogenic loop heat pipe with insufficient working fluid inventory[J]. International Journal of Refrigeration, 2006, 29(4): 519-527.

[11] GULLY P, MO Q, YAN T, et al. Thermal behavior of a cryogenic loop heat pipe for space application[J]. Cryogenics, 2011, 51(8): 420-428.

[12] ZHAO Y, YAN T, LIANG J, et al. A new way of supercritical startup of a cryogenic loop heat pipe[J]. International Journal of Heat and Mass Transfer, 2019, 145: 118793.

[13] ZHAO Y, YAN T, LIANG J. Experimental study on a cryogenic loop heat pipe with high heat capacity[J]. International Journal of Heat and Mass Transfer, 2011, 54(15-16): 3304-3308.

[14] ZHAO Y, WANG N, YAN T, et al. Experimental study on operating characteristics of a cryogenic loop heat pipe without additional power consumption[J]. Applied Thermal Engineering, 2021, 184: 116262.

[15] MO Q, LIANG J T. A novel design and experimental study of a cryogenic loop heat pipe with high heat transfer capability[J]. International Journal of Heat and Mass Transfer, 2006, 49(3-4): 770-776.

[16] 李强, 马路, 宣益民. 低温环路热管(CLHP)的实验研究[J]. 工程热物理学报, 2010, 31(1): 122-125.
LI Q, MA L, XUAN Y M. Experimental investigation of cryogenic loop heat pipe[J]. Journal of Engineering Thermophysics, 2010, 31(1): 122-125(in Chinese).

[17] 刘成志, 董德平, 杨帆. 乙烷温区低温环路热管设计与实验[J]. 低温工程, 2011, 28(6): 57-59.
LIU C Z, DONG D P, YANG F. Design and experimental investigations of ethane cryogenic loop heat pipe[J]. Cryogenics, 2011, 28(6): 57-59(in Chinese).

[18] 张畅, 谢荣建, 张添, 等. 液氮温区平板蒸发器环路热管实验研究[J]. 北京航空航天大学学报, 2019, 45(6): 1211-1217.
ZHANG C, XIE R J, ZHANG T, et al. Experimental study on a liquid nitrogen temperature region loop heat pipe with flat evaporator[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(6): 1211-1217(in Chinese).

[19] BAI L, LIN G, ZHANG H, et al. Experimental study of a nitrogen-charged cryogenic loop heat pipe[J]. Cryogenics, 2012, 52(10): 557-563.

[20] BAI L, LIN G, ZHANG H, et al. Operating characteristics of a miniature cryogenic loop heat pipe[J]. International Journal of Heat and Mass Transfer, 2012, 55(25-26): 8093-8099.

[21] GUO Y, LIN G, BAI L, et al. Experimental study on the supercritical startup of cryogenic loop heat pipes with redundancy design[J]. Energy Conversion and Management, 2016, 118: 353-363.

[22] GUO Y, LIN G, HE J, et al. Experimental study on the supercritical startup and heat transport capability of a neon-charged cryogenic loop heat pipe[J]. Energy Conversion and Management, 2017, 134: 178-187.

[23] GUO Y, LIN G, HE J, et al. Supercritical startup strategy of cryogenic loop heat pipe with different working fluids[J]. Applied Thermal Engineering, 2019, 155: 267-276.

[24] GUO Y, LIN G, BAI L, et al. Experimental study of the thermal performance of a neon cryogenic loop heat pipe[J]. International Journal of Heat and Mass Transfer, 2018, 120: 1266-1274.

[25] GUO Y, LIN G, ZHANG H, et al. Investigation on thermal behaviours of a methane charged cryogenic loop heat pipe[J]. Energy, 2018, 157: 516-525.

[26] 周顺涛, 莫青, 张红星, 等. 深冷环路热管传热性能研究[J]. 航天器工程, 2010, 19(3): 91-95.
ZHOU S T, MO Q, ZHANG H X, et al. Experimental study on heat transfer performance of cryogenic loop heat pipe[J]. Spacecraft Engineering, 2010, 19(3): 91-95(in Chinese).

[27] 周顺涛, 莫青, 张红星, 等. 深冷环路热管超临界启动实验研究[J]. 低温工程, 2010(3): 18-22.
ZHOU S T, MO Q, ZHANG H X, et al. Experimental investigation on supercritical startup of cryogenic loop heat pipe[J]. Cryogenics, 2010(3): 18-22(in Chinese).

[28] KU J, ROBINSON F. Thermal vacuum testing of a helium loop heat pipe for large area cryocooling[C]//46th International Conference on Environmental Systems. [S.l.]: [s.n.], 2016.

[29] HOANG T T, O'CONNELL T A, KHRUSTALEV D K. Development of a flexible advanced loop heat pipe for across-gimball cryocooling[C]//SPIE 48th Annual Meeting of Optical Science and Technology. Bellingham: SPIE, 2003, 5172: 68-76.

[30] SHAEFFER R, HU H, CHUNG J N. An experimental study on liquid nitrogen pipe chilldown and heat transfer with pulse flows[J]. International Journal of Heat and Mass Transfer, 2013, 67: 955-966.

Design and experiment of cryogenic loop heat pipe of two-dimensional pointing at liquid nitrogen zone

LI Nan[1, 2], GUO Yuandong[3, *], XU Cheng[3], ZHANG Hongxing[4], LIN Guiping[3]

(1. College of Aerospace Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;
2. Flight Test Operation Department, Flight Test Center of COMAC, Shanghai 201323, China;
3. School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China;
4. Beijing Institute of Spacecraft System Engineering, Beijing 100094, China)

Abstract: The cooperation of the two-dimensional pointing mechanism and the infrared detector is conducive to the realization of large-scale dynamic tracking, pointing, and rapid positioning of space targets. Coupling the cryogenic loop heat pipe (CLHP) with the two-dimensional pointing mechanism can greatly reduce system complexity, while achieving long-distance heat transmission, and improving detection accuracy and steering flexibility. A two-dimensional pointing CLHP working at the liquid nitrogen zone is designed and developed. The component parameters are introduced, and a pitch and yaw rotation of more than ±90° through the servo motor drive is realized. The thermal vacuum experiment was carried out, examining the effects of different working parameters on the supercritical start and heat transfer limit. The results show that the system designed has a maximum heat transfer capacity of 13 W. Appropriately increasing the filling pressure can help improve the stability and heat transfer capacity of the system, and increasing the auxiliary load of the secondary evaporator can increase the maximum heat transfer ability.

Keywords: loop heat pipe; liquid nitrogen; two-dimensional pointing; supercritical startup; steady-state heat transfer performance

Received: 2021-08-30; **Accepted**: 2021-11-12; **Published Online**: 2021-11-30 14:36

URL: kns.cnki.net/kcms/detail/11.2625.V.20211129.1705.003.html

Foundation items: National Natural Science Foundation of China (52106067); Outstanding Youth Science Fund of China (2020-JCJQ-ZQ-042); China Postdoctoral Science Foundation (2019M660403)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0501

电源车传感器故障检测和数据重构方法

蒋栋年[1,2]，把余江[1]，李炜[1,3,*]

(1. 兰州理工大学 电气工程与信息工程学院，兰州 730050；2. 国家电网甘肃省电力科学研究院，兰州 730050；
3. 兰州理工大学 甘肃省工业过程先进控制重点实验室，兰州 730050)

摘 要：针对电源车由于运行环境复杂而容易发生传感器故障的问题，提出了一种基于时空相关性的传感器故障检测和数据重构方法。针对单个传感器运行数据的时序关系特征，借助具有选择与遗忘机制的极限学习机(SF-ELM)建立了电源车传感器时间序列预测子模型，并据此实现对电源车传感器的故障检测；针对已检测的故障传感器，利用不同传感器之间的空间相关性，通过冗余度分析，使用改进后的互信息熵筛选出与故障传感器数据相关性较高的辅助传感器数据，实现对故障传感器失效数据的在线重构；通过仿真验证了所提方法在电源车传感器故障检测和数据重构中的可行性与有效性。

关 键 词：传感器；时空相关性；极限学习机；故障检测；数据重构
中图分类号：TP277
文献标志码：A **文章编号**：1001-5965(2023)07-1583-10

野外条件下，电源车担负着训练设备的电能供应，是重要能源补给站。在电源车运行过程中，搭载的传感器极易受工作环境的影响发生故障，使其难以正确感知电源车本体的真实状态信息，或因指示错误参与闭环控制，致使故障的影响扩散，错失对故障的及时处理，最终使系统因为电能的缺失陷入瘫痪从而贻误战机。因此，探究能够及时诊断电源车传感器故障及对故障传感器测量数据进行重构的方法，对于提升电源车系统安全性具有重要意义。

传感器故障诊断方法包括基于模型和基于数据驱动的方法。数据驱动方法由于不受复杂系统难以准确建模的限制，已成为传感器故障检测的主流方法，其中时间相关预测方法近年来备受关注[1-5]。时间相关性方法是依据传感器测量数据内部的动态趋势特征建立时序函数关系，借助传感器测量的历史数据预测下一时刻传感器测量数据的范围，利用预测值和测量值之间的偏差实现传感器故障诊断的方法。文献[6]将主成分分析法和Petri网相结合进行故障诊断，实现了高速列车多轴传感器故障检测、隔离和诊断。文献[7]在考虑传感器数据时间相关性的基础上，提出了一种数据驱动的双向递归神经网络方法，实现了传感器的故障检测和数据重构。文献[8]提出了一种基于卷积长短时记忆(long short term memory, LSTM)网络的两阶段多变量时间序列(multivariate time series, MTS)预测模型，提升了传感器数据的预测精度，增强了数据恢复能力。

由此可见，借助时间相关性可以实现对单个传感器测量数据的短时预测，进而实现对传感器故障检测的目的。然而，尽管基于机器学习的时间相关性研究已较为成熟，但该方法在对数据进行预测的过程中仅选取了系统特定的运行工况，不适合长时间的数据恢复。此外，这些研究也未从多元时间序

收稿日期：2021-08-30；**录用日期**：2021-11-20；**网络出版时间**：2021-12-15 08:56
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20211214.1415.001.html
基金项目：国家自然科学基金(62263020)；甘肃省杰出青年科学基金(20JR10RA202)；兰州理工大学红柳优秀青年人才资助计划；兰州市科技计划(2022-2-69)
* **通信作者**. E-mail：liwei@lut.edu.cn

引用格式：蒋栋年，把余江，李炜. 电源车传感器故障检测和数据重构方法[J]. 北京航空航天大学学报，2023，49(7)：1583-1592.
JIANG D N，BA Y J，LI W. Sensor fault detection and data reconstruction method of power supply vehicle[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1583-1592 (in Chinese).

列的角度去考虑传感器之间的空间相关性，制约了相关方法在多工况系统中的应用。

近年来，有学者提出了利用传感器之间的空间相关性进行故障传感器数据重构的方法[9]，该方法借助相邻传感器之间的冗余关系，利用不同传感器的相互协同实现了对故障传感器在一定时间范围内数据的重构，使得关键传感器获取了必要的容错能力。文献[10]借助自动关联神经网络(auto-associative neural network, AANN)构建基于数据降维的输入输出映射模型，从而实现了传感器误差校正和缺失数据的替换。文献[11]利用自适应核主成分分析(kernel principal component analysis, KPCA)方法，研究了时变非线性动态系统的传感器故障特性，实现了气动燃气轮机非线性动态模型中传感器故障检测和重构。

同时，有学者将时间和空间相结合，引入了一种时空相关性思想，利用时空关联数据，基于深度置信网络方法对传感器网络的单个传感器运行数据进行重构[12]。文献[13]提出了一种多尺度时空卷积深度置信网络(multiscale spatio-temporal convolutional deep belief network, MSTCDBN)来进行特征学习和分类，实现传感器故障检测和重构。文献[14]提出了一种基于加权在线序贯极限学习机(weighted online sequential extreme learning machine, WOS-ELM)的时空冗余度诊断方法，实现了航空发动机传感器故障检测和数据重构。然而，对于故障传感器利用时空相关性进行数据重构时还面临如下问题：①如何选取与故障传感器相关性最大且最优的传感器集合的问题；②如何更好地量化评价传感器测量数据之间的相关性。

鉴于此，本文借鉴时空相关性的思想，首先，借助单个传感器的时间相关性分析，利用具有选择与遗忘机制的极限学习机(selective forgetting extreme learning machine, SF-ELM)方法建立电源车传感器在线时间序列预测模型，达到传感器对未来运行状态短时预测的目的，并基于预测值与实际值之间的残差，实现单个传感器的故障检测；其次，针对已锁定的故障传感器，考虑不同传感器之间的空间相关性，通过引入互信息熵的量化评价，筛选出辅助传感器，进而基于ELM对故障传感器失效数据进行在线重构。

1 本文方法

为了实时检测运行状态，电源车安装了包括电压、电流等多种传感器，就单个传感器的数据而言，由于其反映了电源车的运行状态，从时序上来看，数据间存在着时间的关联性，而对于多个安装在不同位置的传感器，由于从不同侧面反映了电源车的机电属性，数据间又存在空间相关性。

因此，本文从传感器的时间相关性和空间相关性2个维度，进行电源车传感器故障检测和数据重构。首先，借助已采集的电源车传感器运行数据，利用SF-ELM方法对时间序列良好的跟踪特性，建立传感器时间序列预测模型，实现对传感器运行数据的短时预测，并基于此检测传感器的故障。其次，在空间维度上，针对已检测的故障传感器，将与之相关的测点传感器纳入传感器测量系统，从空间的角度分析与故障测点相关的辅助测点传感器，借助辅助变量的选择和空间相关模型的建立，达到故障传感器数据重构的目的。本文方法的主要贡献在于：可通过量化评价的角度选取与故障传感器冗余度最高的测点集合，利用时间相关性和空间相关性的有效融合，使电源车传感器具有容错能力，提高其安全运行水平。该方法的具体实现流程如图1所示。

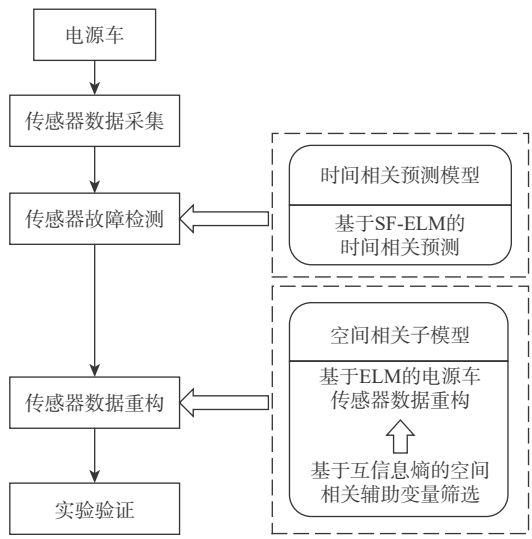


图1 电源车传感器故障检测和数据重构流程

Fig. 1 Sensor fault detection and data recovery process of power supply vehicle

2 基于时间相关性的电源车传感器故障检测

针对电源车传感器的故障检测问题，考虑传感器输出时间序列数据的时间相关性，通过有效提取数据内部的动态趋势特征，建立传感器在线时间序列预测模型，对其未来运行状态进行短时预测，通过预测值和实际值之间的残差判断传感器有无故障。

2.1 基于SF-ELM的时间序列预测模型建立

对于时间序列预测模型的建立，本文引入一种

SF-ELM 方法[15], 相比于 ELM 和 OS-ELM, SF-ELM 考虑了新旧数据对预测模型建模精度和速度的影响, 通过对新旧数据赋予不同的权值, 确保了更新后的预测模型包含更多新数据的信息, 从而使所建立的在线时间序列预测模型对时间序列的动态变化具有更强的跟踪能力, 即具有较高的在线预测精度。

考虑将传感器的N个测量数据时间序列$x_1,x_2,\cdots,x_N$转换为训练样本$(x_1,t_1),(x_2,t_2),\cdots,(x_k,t_k)$, 其中, $\boldsymbol{x}_i=[x_i,x_{i+1},\cdots,x_{i+n-1}]^{\mathrm{T}}$为预测模型的输入, $t_i=x_{i+n}$为预测模型的输出, n为嵌入维数。则一个包含L个隐层神经元函数的回归模型为

$$\sum_{i=1}^{L}\beta_i f(\boldsymbol{\alpha}_i\boldsymbol{x}_i+b_i)=t_j \tag{1}$$

式中: $j\in 1,2,\cdots,k$为训练样本数量; $\boldsymbol{\alpha}_i$为输入层神经元连接第i个隐层神经元的输入权值; β_i为连接隐层神经元和输出层神经元的输出权值; b_i为第i个隐层神经元的偏置。

将式 (1) 写成矩阵形式, 可得

$$\boldsymbol{H}_k\boldsymbol{\beta}_k=\boldsymbol{T}_k \tag{2}$$

式中: $\boldsymbol{H}_k$为神经元矩阵; $\boldsymbol{T}_k=[t_1,t_2,\cdots,t_k]^{\mathrm{T}}$为输出向量。

由于$k=N-n\geqslant L$, 求解式 (2) 可得初始输出权值为

$$\boldsymbol{\beta}_k=(\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{H}_k)^{-1}\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{T}_k \tag{3}$$

此时, 当有新的训练样本(x_{k+1},t_{k+1})产生时, 将$x_{N-n+1},x_{N-n+2},\cdots,x_N$作为输入, 对神经元矩阵$\boldsymbol{h}_{k+1}$进行更新, 从而可以得到$x_N$的一步预测值$\bar{x}_{N+1}$为

$$\bar{x}_{N+1}=\boldsymbol{h}_{k+1}\boldsymbol{\beta}_k \tag{4}$$

为了提高 SF-ELM 在线预测的精度, 对新旧数据进行加权, 当有新的训练样本(x_{k+1},t_{k+1})产生时, 输出权值可以更新为$\boldsymbol{\beta}_{k+1}$:

$$\begin{aligned}\boldsymbol{\beta}_{k+1}&=\left(\begin{bmatrix}\lambda\boldsymbol{H}_k\\ \boldsymbol{h}_{k+1}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{H}_k\\ \boldsymbol{h}_{k+1}\end{bmatrix}\right)^{-1}\begin{bmatrix}\lambda\boldsymbol{H}_k\\ \boldsymbol{h}_{k+1}\end{bmatrix}^{\mathrm{T}}\begin{bmatrix}\boldsymbol{T}_k\\ t_{k+1}\end{bmatrix}=\\ &(\lambda\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{H}_k+\boldsymbol{h}_{k+1}^{\mathrm{T}}\boldsymbol{h}_{k+1})^{-1}(\lambda\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{T}_k+\boldsymbol{h}_{k+1}^{\mathrm{T}}t_{k+1})\end{aligned} \tag{5}$$

式中: $\lambda\in(0,1)$为遗忘因子, 其作用是通过降低旧数据的影响, 提高新数据对预测模型的影响。

令

$$\boldsymbol{P}_{k+1}=(\lambda\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}_k+\boldsymbol{h}_{k+1}^{\mathrm{T}}\boldsymbol{h}_{k+1})^{-1} \tag{6}$$

将 Sherman-Morrison 求逆引理[16] 引入式 (6) 可得

$$\boldsymbol{P}_{k+1}=\frac{\boldsymbol{P}_k}{\lambda}-\frac{\boldsymbol{Q}_k\boldsymbol{Q}_k^{\mathrm{T}}}{\lambda(\lambda+\boldsymbol{h}_{k+1}\boldsymbol{Q}_k)} \tag{7}$$

式中: $\boldsymbol{Q}_k=\boldsymbol{P}_k\boldsymbol{h}_{k+1}^{\mathrm{T}}$。

为了降低计算复杂度, 实现有效的在线故障诊断, 可根据式(8)对$\boldsymbol{P}_k$进行选择性更新:

$$\boldsymbol{P}_{k+1}=\begin{cases}\dfrac{\boldsymbol{P}_k}{\lambda}-\dfrac{\boldsymbol{P}_k\boldsymbol{h}_{k+1}^{\mathrm{T}}\boldsymbol{h}_{k+1}\boldsymbol{P}_k^{\mathrm{T}}}{\lambda(\lambda+\boldsymbol{h}_{k+1}\boldsymbol{P}_k\boldsymbol{h}_{k+1}^{\mathrm{T}})} & R_N>\varepsilon\\ \boldsymbol{P}_k & R_N<\varepsilon\end{cases} \tag{8}$$

式中: $R_N=\sqrt{(\bar{x}_{N+1}-x_{N+1})^2}$为均方根误差, x_{N+1}为真实值; ε为设定的阈值。

对式 (6) 两端同时求逆可得

$$\boldsymbol{P}_{k+1}^{-1}=\lambda\boldsymbol{P}_k^{-1}+\boldsymbol{h}_{k+1}^{\mathrm{T}}\boldsymbol{h}_{k+1} \tag{9}$$

式中: $\boldsymbol{P}_k=(\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{H}_k)^{-1}$。

将式 (9) 和式 (3) 代入式 (5), 可得输出权值$\boldsymbol{\beta}_{k+1}$更新为

$$\begin{aligned}\boldsymbol{\beta}_{k+1}&=(\lambda\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{H}_k+\boldsymbol{h}_{k+1}^{\mathrm{T}}\boldsymbol{h}_{k+1})^{-1}(\lambda\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{T}_k+\boldsymbol{h}_{k+1}^{\mathrm{T}}t_{k+1})=\\ &\boldsymbol{P}_{k+1}(\lambda\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{T}_k+\boldsymbol{h}_{k+1}^{\mathrm{T}}t_{k+1})=\boldsymbol{P}_{k+1}(\lambda\boldsymbol{P}_k^{-1}\boldsymbol{\beta}_k+\boldsymbol{h}_{k+1}^{\mathrm{T}}t_{k+1})=\\ &\boldsymbol{P}_{k+1}\left[(\boldsymbol{P}_{k+1}^{-1}-\boldsymbol{h}_{k+1}^{\mathrm{T}}\boldsymbol{h}_{k+1})\boldsymbol{\beta}_k+\boldsymbol{h}_{k+1}^{\mathrm{T}}t_{k+1}\right]=\\ &\boldsymbol{\beta}_k+\boldsymbol{P}_{k+1}\boldsymbol{h}_{k+1}^{\mathrm{T}}(t_{k+1}-\boldsymbol{h}_{k+1}\boldsymbol{\beta}_k)\end{aligned} \tag{10}$$

由此可得到基于 SF-ELM 的电源车传感器在线时间序列预测模型。

2.2　基于时间序列预测模型的传感器故障检测

依据 2.1 节所建立的时间序列预测模型, 通过计算预测模型的预测值$\bar{x}_{N+1}$和传感器的实际测量值x_{k+1}之间的残差可实现传感器的故障检测。

算法 1　基于时间序列预测模型的电源车传感器故障检测算法。

步骤 1　对于已有的训练样本$(x_1,t_1),(x_2,t_2),\cdots,(x_k,t_k)$, 根据$\boldsymbol{\beta}_k=\boldsymbol{P}_k\boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{T}_k$计算初始输出权值$\boldsymbol{\beta}_k$。

步骤 2　在线生成新的训练样本(x_{k+1},t_{k+1}), 对神经网络矩阵$\boldsymbol{h}_{k+1}$进行更新, 并利用$\bar{x}_{N+1}=\boldsymbol{h}_{k+1}\boldsymbol{\beta}_k$对$x_N$进行一步预测。

步骤 3　通过式 (8) 对$\boldsymbol{P}_k$进行选择性更新。

步骤 4　利用更新后的$\boldsymbol{P}_k$根据$\boldsymbol{\beta}_{k+1}=\boldsymbol{\beta}_k+\boldsymbol{P}_{k+1}\cdot\boldsymbol{h}_{k+1}^{\mathrm{T}}(t_{k+1}-\boldsymbol{h}_{k+1}\boldsymbol{\beta}_k)$完成$\boldsymbol{\beta}_k$的更新, 其中, $t_{k+1}=x_{N+1}$。

步骤 5　在获取传感器预测值$\bar{x}_{N+1}$的基础上, 可通过传感器预测值$\bar{x}_{N+1}$和测量值x_{k+1}之间的残差进行传感器的故障检测:

$$\begin{cases}x_{k+1}-\bar{x}_{N+1}>\delta & \text{传感器发生故障}\\ x_{k+1}-\bar{x}_{N+1}<\delta & \text{传感器运行正常}\end{cases} \tag{11}$$

步骤 6　设$k=k+1$, $N=N+1$后, 当$x_{k+1}-\bar{x}_{N+1}<\delta$时返回至步骤 2 继续检测, 当$x_{k+1}-\bar{x}_{N+1}>\delta$时则检测出传感器发生故障。

式(11)中的δ为故障检测阈值。借助文献 [17] 方法, 利用传感器正常情况下的残差序列分布实现对故障检测阈值的设定, 阈值范围为$[\eta-Z\sigma,\eta+Z\sigma]$, 其中, η为传感器正常情况下残差序列均值, σ为残差的方差。

3 基于空间相关性的电源车传感器数据重构

在实现了单个传感器故障诊断的基础上，要实现故障传感器数据的数据重构，仅依靠基于时间相关性的传感器数据短时预测是难以满足电源车可靠性要求的。可行的方法是通过分析系统中所配置传感器之间的冗余度，借助不同传感器间空间相关性的量化评价，选取与故障传感器相关度最大的辅助传感器对其进行数据准确重构，并根据其正确数据确保电源车安全运行。

对于传感器空间相关性的量化评价，常见的方法有皮尔逊相关系数和互信息等。但皮尔逊相关系数只能衡量变量之间线性相关程度的量，难以满足传感器运行数据多变的复杂特性，借助于信息熵理论的互信息方法不依赖数据序列的特点，有望通过数据分布来实现相关度评价的目的。

3.1 基于信息熵理论的传感器信息值量化评价

设电源车的n个测点传感器为$S_1,S_2,\cdots,S_n$，传感器的期望输出为$\hat{\boldsymbol{x}}_1,\hat{\boldsymbol{x}}_2,\cdots,\hat{\boldsymbol{x}}_n$，实际输出数据为$\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n$，可得传感器的测量数据的残差为

$$\boldsymbol{r}_i=\boldsymbol{x}_i-\hat{\boldsymbol{x}}_i \tag{12}$$

理想情况下，当系统无故障时，由于系统不可避免地存在噪声等不确定性因素的影响，其实际输出与期望输出也会存在偏差，也即传感器残差数据在噪声的影响下偏离零值。此时，传感器的信息熵可以定义为传感器实际输出与期望输出的残差数据的概率密度函数，而残差$\boldsymbol{r}$的概率密度函数应近似于测量噪声w的概率密度函数，同时其先验概率密度函数可以估算为$f_{S_i}(\boldsymbol{r}_i)$。根据信息量的定义[18]，残差数据序列$\boldsymbol{r}_i$的信息量可以定义为$h(\boldsymbol{r}_i)=-\ln f_{S_i}(\boldsymbol{r}_i)$。为使信息的复杂程度能够被量化，可以引入信息熵进行度量：

$$H(S_i)=-\int f_{S_i}(\boldsymbol{r}_i)\ln f_{S_i}(\boldsymbol{r}_i)\mathrm{d}\boldsymbol{r}_i \tag{13}$$

由信息熵的定义可知，必须知道变量的概率密度才能够对信息熵进行计算，为更加合理地减少信息熵的计算复杂度，需要对概率密度函数进行估计。一般情况下，残差$\boldsymbol{r}_i$的分布特性可近似服从高斯分布，若概率密度函数为非高斯的单峰偏移特性，可采用BOX-COX变换矫正波形偏移[19]；若概率密度函数呈现非高斯的多峰形态，则需要运用高斯混合模型将其统一在高斯分布特性范畴之内。假定n变量的概率密度函数为

$$f_{S_i}(\boldsymbol{r}_i)=\frac{1}{\sqrt{(2\pi)^n|\boldsymbol{\Sigma}|}}\exp\left[-\frac{1}{2}(\boldsymbol{r}_i-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{r}_i-\boldsymbol{\mu})\right] \tag{14}$$

式中：$\boldsymbol{\Sigma}$为多元正态分布的协方差矩阵；$\boldsymbol{\mu}$为正态分布的均值向量。

在利用式(13)进行信息熵计算时，其非线性特性增加了计算的复杂性，考虑如下方法对其进行简化。

定理1 对于含有n个测点传感器$S_1,S_2,\cdots,S_n$的系统，其传感器S_i残差的信息熵值为

$$H(S_i)=\frac{n}{2}(1+\ln(2\pi))+\frac{1}{2}\ln|\boldsymbol{\Sigma}| \tag{15}$$

证明 已知传感器S_i的残差为$\boldsymbol{r}_i$，所含的信息量为$-\ln f_{S_i}(\boldsymbol{r}_i)$，其信息熵为

$$H(S_i)=-\int f_{S_i}(\boldsymbol{r}_i)\ln f_{S_i}(\boldsymbol{r}_i)\mathrm{d}\boldsymbol{r}_i$$

对于任意的高斯分布$\boldsymbol{r}_i\sim N(\boldsymbol{\mu},\boldsymbol{\Sigma})$，称$\boldsymbol{y}=\boldsymbol{\Sigma}^{-1/2}[\boldsymbol{r}_i-\boldsymbol{\mu}]$为马哈拉诺比斯变换[20]，其中，$\boldsymbol{y}\sim N(0,\boldsymbol{I}_k)$，即$\boldsymbol{y}$符合标准高斯分布。同时，已知式(14)为$n$变量的正态分布的概率密度函数，且$\boldsymbol{r}_i=\boldsymbol{\Sigma}^{1/2}\boldsymbol{y}+\boldsymbol{\mu}$，$R=\det\left[\dfrac{\partial\boldsymbol{r}_i}{\partial\boldsymbol{y}^{\mathrm{T}}}\right]=|\boldsymbol{\Sigma}|^{1/2}$，通过变量代换可以得到

$$f(\boldsymbol{y})=(2\pi)^{-n/2}\exp\left[-\frac{1}{2}\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}\right] \tag{16}$$

将式(14)和式(16)引入式(13)，可以得到信息熵的公式为

$$\begin{aligned}H(S_i)&=-\int f_{S_i}(\boldsymbol{r}_i)\ln f_{S_i}(\boldsymbol{r}_i)\mathrm{d}\boldsymbol{r}_i=\\&-\int f_{S_i}(\boldsymbol{r}_i)\ln\left\{(2\pi)^{-n/2}|\boldsymbol{\Sigma}|^{-1/2}\cdot\right.\\&\left.\exp\left[-\frac{1}{2}(\boldsymbol{r}_i-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{r}_i-\boldsymbol{\mu})\right]\right\}\mathrm{d}\boldsymbol{r}_i=\\&-\int f_{S_i}(\boldsymbol{r}_i)\left[\ln((2\pi)^{-n/2}|\boldsymbol{\Sigma}|^{-1/2})-\right.\\&\left.\frac{1}{2}(\boldsymbol{r}_i-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{r}_i-\boldsymbol{\mu})\right]\mathrm{d}\boldsymbol{r}_i=\\&\ln\left((2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}\right)+\frac{1}{2}\int f_{S_i}(\boldsymbol{r}_i)[(\boldsymbol{r}_i-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{r}_i-\boldsymbol{\mu})]\mathrm{d}\boldsymbol{r}_i=\\&\ln\left((2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}\right)+\frac{1}{2}\int f_{S_i}(\boldsymbol{r}_i)\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}\mathrm{d}\boldsymbol{y}=\\&\ln\left((2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}\right)+\frac{1}{2}\sum_{i=1}^{k}E[y_i^2]=\ln\left((2\pi)^{n/2}|\boldsymbol{\Sigma}|^{1/2}\right)+\frac{n}{2}=\\&\frac{n}{2}(1+\ln(2\pi))+\frac{1}{2}\ln\left|\boldsymbol{\Sigma}_{S_i}\right|\end{aligned} \tag{17}$$

通过上述定理发现，如果维数n已知，则多元分布的信息熵实际上就是关于空间协方差矩阵的函数。因此，只需要估算传感器S_i残差数据的协方差矩阵，便可以得到传感器残差的信息熵值。

3.2 基于互信息熵的传感器相关性量化评价

互信息熵可以看成是一个随机变量中包含的另一个随机变量的信息量，可以衡量不同变量之间共有信息量的大小。因此，对于传感器之间的冗余度，

即空间相关性, 可以考虑通过互信息熵来量化评价。根据互信息定义, 不同传感器之间的互信息熵为

$$I(S_i,S_j)=-\iint f(\boldsymbol{r}_i,\boldsymbol{r}_j)\ln\left(\frac{f(\boldsymbol{r}_i,\boldsymbol{r}_j)}{f_{S_i}f_{S_j}}\right)\mathrm{d}\boldsymbol{r}_i\mathrm{d}\boldsymbol{r}_j \tag{18}$$

式中: $f(\boldsymbol{r}_i,\boldsymbol{r}_j)$为传感器$S_i$和$S_j$所含信息值的联合概率分布。

图 2 为互信息熵的 V 氏图。

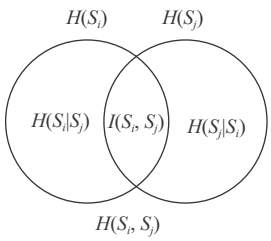


图 2　信息熵 V 氏图

Fig. 2　Information entropy V diagram

考虑到互信息熵计算式的非线性特性, 以及传感器S_i和S_j所含信息值的联合概率分布难以估计的问题, 考虑如下方法对其进行改进。

定理 2　对于含有n个测点传感器$S_1,S_2,\cdots,S_n$的系统, 其传感器S_i和传感器S_j之间的互信息熵值为

$$I(S_i,S_j)=\frac{1}{2}\ln\left(\frac{|\boldsymbol{\Sigma}_{S_i}|}{|\boldsymbol{\Sigma}_{S_i|S_j}|}\right) \tag{19}$$

式中: $\boldsymbol{\Sigma}_{S_i}$为传感器$S_i$的协方差矩阵; $\boldsymbol{\Sigma}_{S_i|S_j}$为$S_i$相对于$S_j$的条件协方差矩阵。

证明　已知传感器S_i的残差为$\boldsymbol{r}_i$, 传感器S_j的残差为$\boldsymbol{r}_j$, 则传感器S_i和传感器S_j之间的互信息熵为式 (18)。

同时, 互信息还有另一种表达形式:

$$I(S_i,S_j)=H(S_i)-H(S_i|S_j) \tag{20}$$

式中: $H(S_i|S_j)$为条件熵, 表示传感器S_i的信息值在已知传感器S_j信息值的条件下的不确定度, 能够反映变量之间的依赖程度, 其计算公式为

$$H(S_i|S_j)=H(S_i,S_j)-H(S_j) \tag{21}$$

其中: $H(S_i,S_j)$为联合熵; $H(S_j)$为传感器S_j的信息熵。

由式 (20) 可知, 互信息熵计算的关键是求取条件熵$H(S_i|S_j)$。根据 3.1 节中传感器残差数据符合高斯分布的前提下进一步假设传感器S_i和S_j的残差数据具有协方差的联合正态性:

$$\boldsymbol{\Sigma}_{S_i\cup S_j}=\begin{bmatrix}\boldsymbol{\Sigma}_{S_i} & \boldsymbol{\Sigma}_{S_i,S_j}\\ \boldsymbol{\Sigma}_{S_j,S_i} & \boldsymbol{\Sigma}_{S_j}\end{bmatrix} \tag{22}$$

式中: $\boldsymbol{\Sigma}_{S_i}$和$\boldsymbol{\Sigma}_{S_j}$分别为$S_i$和$S_j$的协方差矩阵; $\boldsymbol{\Sigma}_{S_i,S_j}$为互协方差。

根据式 (17) 和式 (21) 可以得到条件熵的另一个表达式为

$$H(S_i|S_j)=H(S_i,S_j)-H(S_j)=\frac{n}{2}(1+\ln(2\pi))+\frac{1}{2}\ln\left(\frac{|\boldsymbol{\Sigma}_{S_i\cup S_j}|}{|\boldsymbol{\Sigma}_{S_j}|}\right)=\frac{n}{2}(1+\ln(2\pi))+\frac{1}{2}\ln|\boldsymbol{\Sigma}_{S_i|S_j}| \tag{23}$$

式中: $\boldsymbol{\Sigma}_{S_i|S_j}$为$S_i$相对于$S_j$的条件协方差矩阵。

根据式 (17)、式 (20) 和式 (23) 可以得到传感器S_i和传感器S_j之间的互信息为

$$\begin{aligned}I(S_i,S_j)&=H(S_i)-H(S_i|S_j)=\frac{n}{2}(1+\ln(2\pi))+\\&\frac{1}{2}\ln|\boldsymbol{\Sigma}_{S_i}|-\left[\frac{n}{2}(1+\ln(2\pi))+\frac{1}{2}\ln|\boldsymbol{\Sigma}_{S_i|S_j}|\right]=\\&\frac{1}{2}\left(\ln|\boldsymbol{\Sigma}_{S_i}|-\ln|\boldsymbol{\Sigma}_{S_i|S_j}|\right)=\frac{1}{2}\ln\left(\frac{|\boldsymbol{\Sigma}_{S_i}|}{|\boldsymbol{\Sigma}_{S_i|S_j}|}\right)\end{aligned} \tag{24}$$

而条件协方差$\boldsymbol{\Sigma}_{S_i|S_j}$可以用式(25)求取为

$$\boldsymbol{\Sigma}_{S_i|S_j}=\boldsymbol{\Sigma}_{S_i}-\boldsymbol{\Sigma}_{S_i,S_j}\boldsymbol{\Sigma}_{S_j}^{-1}\boldsymbol{\Sigma}_{S_j,S_i} \tag{25}$$

3.3　基于互信息熵辅助变量筛选的 ELM 电源车传感器数据重构

通过 3.2 节对不同传感器之间的相关性量化评价, 本节利用不同传感器之间的信息熵值, 在引入最优特征 (best individual feature, BIF) 算法的基础上, 对所选的所有辅助变量基于评价函数$J(S)$进行评价, 再依据函数值大小选择前 k 个辅助变量作为子集 S, 评价函数 BIF 算法, 如下:

$$J(S_i)=I(S_i,c) \tag{26}$$

式中: $S_i\in F$表示辅助变量; c为主导变量。

虽然 SF-ELM 在时间序列预测中具有良好的预测能力, 但在利用电源车传感器空间冗余度进行数据重构时, ELM 算法更具优势, 因此, 借助互信息熵及 BIF 算法筛选辅助变量, 再利用 ELM 建立基于空间相关的传感器数据重构模型, 进而实现故障传感器的数据重构。

算法 2　基于互信息熵辅助变量筛选的电源车传感器数据重构算法。

步骤 1　借助算法 1 确定故障传感器输出为主导变量, 根据完好传感器确定辅助变量个数, 并借助式 (25) 和式 (26) 对辅助变量进行排序。

步骤 2　将前 k 个辅助变量作为 ELM 的输入, 主导变量输出作为 ELM 的输出。

步骤 3　利用辅助变量与主导变量正常时的历史数据, 训练获得 ELM 的输出权值。

步骤 4　以传感器实时记录的辅助变量作为 ELM 的输入, 通过计算映射即可实现故障传感器失效数据的重构。

步骤 5 通过比较不同前 k 个值所计算出的主导变量的预测值和实际值之间的预测误差，判断其是否达到传感器最低拟合误差，若实现则终止循环，否则回到步骤 2。

4 仿真实验分析

本文仿真实验所采用的数据来源于与某电源车辆研究所联合开发的“120 kW 车辆电源仿真系统”，该系统开发于 2014 年，已应用于实际工程测试中。该系统借助模块化的方式对电源车系统进行建模，将电源车模型分解为柴油机模型、同步发电机模型、调速器模型、励磁系统模型和负载模型，电源车模块关系如图 3 所示。本文中所有数据均为系统采集的数据，主要电气检测变量包括有功功率、无功功率、功率因子、定子电压、定子电流、转子速度和励磁电压。

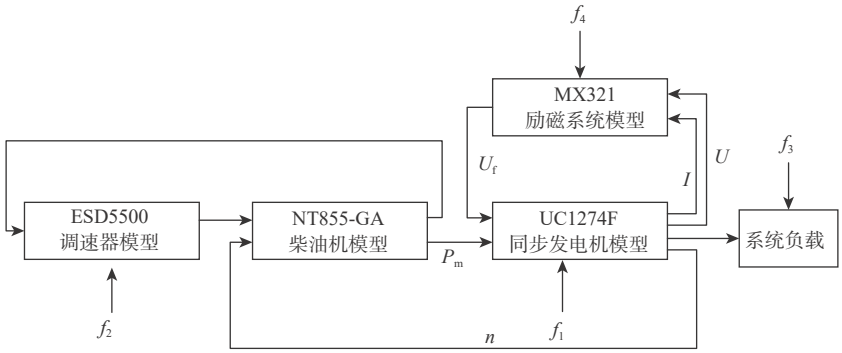


图 3 电源车模块关系

Fig. 3 Relationship of power vehicle module

4.1 基于 SF-ELM 时间序列预测模型的故障检测

1）时间序列预测模型的建立

将 SF-ELM 用于在线时间序列预测模型的建立，通过网络参数优选，得到初始训练阶段训练样本数为 200，神经元个数为 20，嵌入维数为 5，阈值 $\varepsilon=0.01$，遗忘因子 $\lambda=0.92$，激活函数则选为 sig 函数。为了验证 SF-ELM 的预测精度，与传统 OS-ELM 进行对比，图 4 为 OS-ELM 所建立的在线故障检测模型，图 5 为 SF-ELM 所建立的在线故障检测模型。

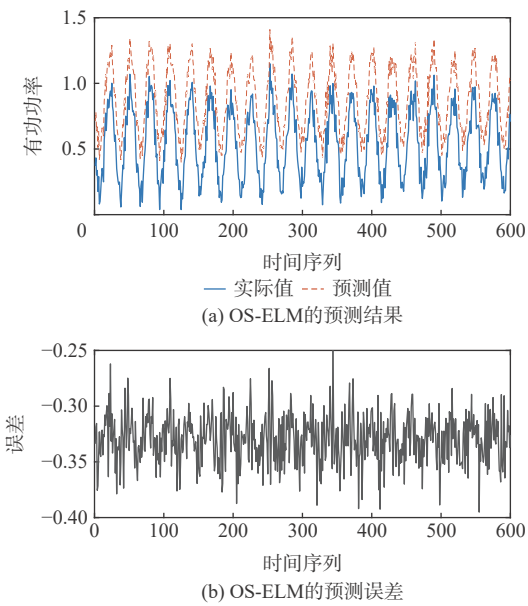


图 4 传感器正常工作时 OS-ELM 的预测结果及误差

Fig. 4 Prediction results and errors of OS-ELM during normal sensor operation

从图 4 和图 5 的对比可以看出，当传感器无故障时，由于 SF-ELM 在模型更新阶段对新旧数据赋予了不同的权值，相较于 OS-ELM，SF-ELM 方法对电源车有功功率数据具有更好的跟踪能力。表 1 为时间序列预测模型 10 次测试结果的均值。

从表 1 可以看出，由于 SF-ELM 加入了遗忘因子及对 $\boldsymbol{P}_{k+1}$ 的选择性更新，使得 SF-ELM 方法无论在预测精度还是预测时间都要优于 OS-ELM 方法，对预测精度的提升上尤为明显。因此，SF-ELM 方

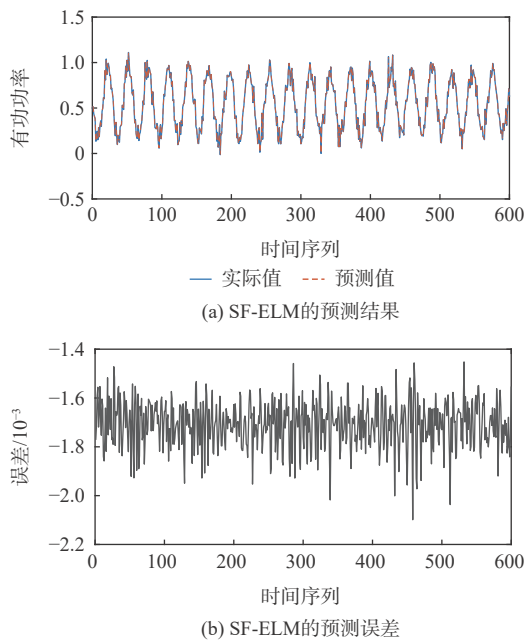


图 5 传感器正常工作时 SF-ELM 的预测结果及误差

Fig. 5 Prediction results and errors of SF-ELM during normal sensor operation

表 1　时间序列预测模型测试结果

Table 1　Test results of time series prediction model

方法	初始训练样本	相对误差/%	时间/s
OS-ELM	200	10.15	0.954 64
SF-ELM	200	1.01	0.856 89

法更适合作为电源车传感器在线时间序列的预测模型。

2) 传感器故障检测

对于故障检测阈值的选取, 在电源车传感器所含残差符合高斯分布的前提下, 通过 Lilliefors 假设检验可得显著性水平为 0.02, 说明传感器残差的分布确实可近似为正态分布。在此基础上, 可以获取传感器残差数据的均值和方差为

$$\begin{cases} \mu = \dfrac{1}{n}\sum_{i=1}^{n} T_i \\ \sigma = \dfrac{1}{n-1}\sum_{i=1}^{n}(T_i-\mu)^2 \end{cases} \tag{27}$$

式中: T 为残差序列。

若设显著性水平为 $\alpha = 0.05$, 也就是置信度为 95% 时, 相关系数则满足 $Z = 1.6$, 可得

$$p\{\bar{\eta} - Z\sigma < \eta < \bar{\eta} + Z\sigma\} = 1 - \alpha \tag{28}$$

从而可以得到阈值计算公式如下:

$$\delta = \mu \pm 1.6\sigma \tag{29}$$

参考文献 [17] 并利用式 (27) 多次循环计算后确定故障检测阈值为[−0.26, 3.44]。

电源车有功功率传感器常见故障有漂移故障和恒偏差故障, 本文以恒偏差故障为例进行仿真, 通过对比时间相关预测模型的预测值和实际值来检测传感器故障。图 6 为有功功率传感器发生恒偏差故障时的预测值及残差。

通过图 6 可以看出, 当有功功率传感器发生恒偏差故障时, 传感器残差超过了相应的故障检测阈值[−0.26, 3.44], 因此可以通过本文所建立时间序列预测模型实现电源车传感器的故障检测。

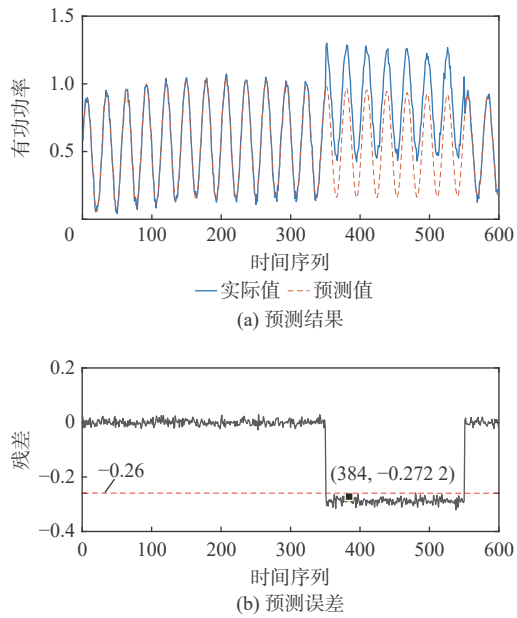


图 6　传感器发生恒偏差故障时的预测结果及残差

Fig. 6　Prediction results and residuals in case of constant bias fault of sensor

4.2　基于互信息熵筛选的 ELM 传感器数据恢复

为了对第 3 节有功功率传感器故障时的数据进行重构, 考虑利用空间相关性来进行重构信号, 其中有功功率传感器为重构主导变量, 其余传感器测量数据为辅助变量, 根据式 (24) 可以得到电源车传感器之间的互信息熵, 如表 2 所示。

表 2　电源车传感器之间的互信息熵

Table 2　Mutual information entropy between power vehicle sensors

参数	有功功率	无功功率	功率因子	定子电流	定子电压	转子速度	励磁电压
有功功率		$1.320\,1\times10^{-4}$	$3.648\,4\times10^{-5}$	$1.347\,3\times10^{-4}$	$1.554\,1\times10^{-4}$	$3.134\,3\times10^{-6}$	$3.302\,1\times10^{-5}$
无功功率	$1.320\,1\times10^{-4}$		$3.173\,8\times10^{-4}$	$5.155\,2\times10^{-5}$	$1.323\,9\times10^{-4}$	$2.576\,7\times10^{-5}$	$2.663\,7\times10^{-5}$
功率因子	$3.648\,4\times10^{-5}$	$3.173\,8\times10^{-4}$		$1.291\,0\times10^{-4}$	$3.900\,5\times10^{-4}$	$7.385\,6\times10^{-5}$	$8.571\,8\times10^{-5}$
定子电流	$1.347\,3\times10^{-4}$	$5.155\,2\times10^{-5}$	$1.291\,0\times10^{-4}$		$1.987\,2\times10^{-4}$	$3.083\,8\times10^{-5}$	$1.073\,9\times10^{-4}$
定子电压	$1.554\,1\times10^{-4}$	$1.323\,9\times10^{-4}$	$3.900\,5\times10^{-4}$	$1.987\,2\times10^{-4}$		$1.063\,5\times10^{-4}$	$1.133\,1\times10^{-4}$
转子速度	$3.134\,3\times10^{-6}$	$2.576\,7\times10^{-5}$	$7.385\,6\times10^{-5}$	$3.083\,8\times10^{-5}$	$1.063\,5\times10^{-4}$		$1.221\,0\times10^{-6}$
励磁电压	$3.302\,1\times10^{-5}$	$2.663\,7\times10^{-5}$	$8.571\,8\times10^{-5}$	$1.073\,9\times10^{-4}$	$1.133\,1\times10^{-4}$	$1.221\,0\times10^{-6}$	

通过表 2 和式 (26) 得到辅助变量与主导变量相关性大小依次为定子电压、定子电流、无功功率、功率因子、励磁电压和转子速度。选取前 k 个辅助变量个数分别为 k=2、k=3、k=4、k=5、k=6, 根据不同的 k 值, 分别采用 ELM 方法对主导变量进行重构, 得到如图 7 所示的重构误差箱线图和功率谱密度估计。

从图 7(a) 可以看到, 当辅助变量个数 k 分别为 3 和 4 时, 重构误差相近且均较小。相应的, 通过图 7(b) 也可发现, 当 k=4 时重构误差功率谱密度估

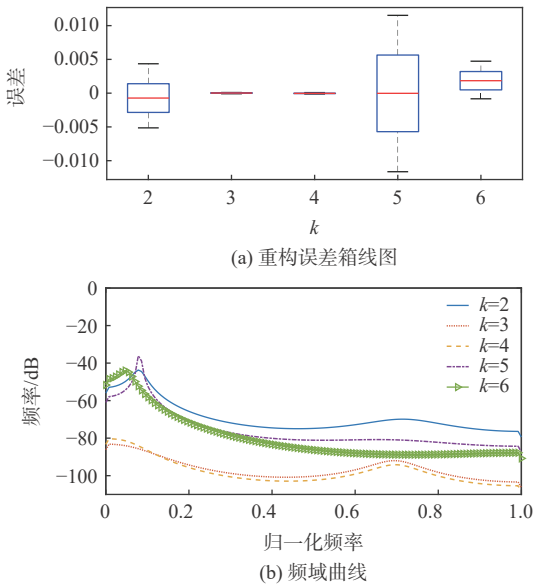


图 7 重构误差箱线图和频域曲线

Fig. 7 Reconstruction error box line plots and frequency domain curves

计值最低。由此可得，利用改进后的互信息熵方法筛选出 4 个辅助变量，依次为定子电压、定子电流、无功功率和功率因子，以此为输入可获得故障传感器数据的准确重构。

图 8 为未进行辅助变量筛选，即将 k=6 作为 ELM 的输入，通过 3.3 节中的算法 2 得到失效数据的重构值。模型重构精度利用相对误差ξ来衡量，如下：

$$\xi=\frac{\hat{y}-y}{y}\times 100\% \tag{30}$$

式中：$\hat{y}$为重构值；y为实际值。

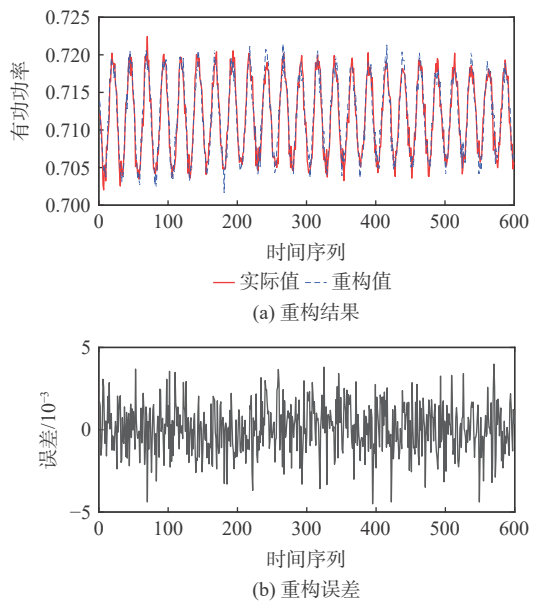


图 8 未进行辅助变量筛选时 ELM 重构结果及重构误差

Fig. 8 ELM reconstruction results and reconstruction errors without auxiliary variable screening

根据 k=6 时的传感器重构数据和式 (30) 可以得到重构数据的相对误差为 0.17%。将互信息筛选后的辅助变量个数 k=4 作为 ELM 方法的输入，得到图 9 所示的重构结果，同样通过式 (30) 得到 k=4 时的相对误差为 0.14%，进而得到表 3 所示的不同辅助变量个数的重构结果对比。

从图 8 和图 9 可以看出，经过辅助变量最优筛选后的 ELM 精度较未筛选的 ELM 具有更高的重构精度。同时从表 3 也可以看出，经过辅助变量优选后，也即辅助变量个数 k=4 时，基于 ELM 方法建立的空间相关子模型拥有更低的均方误差及更短的重构时间。说明通过互信息熵量化传感器之间冗余度的评价，并据此进行最优辅助变量的筛选，可实现故障传感器失效数据精确高效的重构。

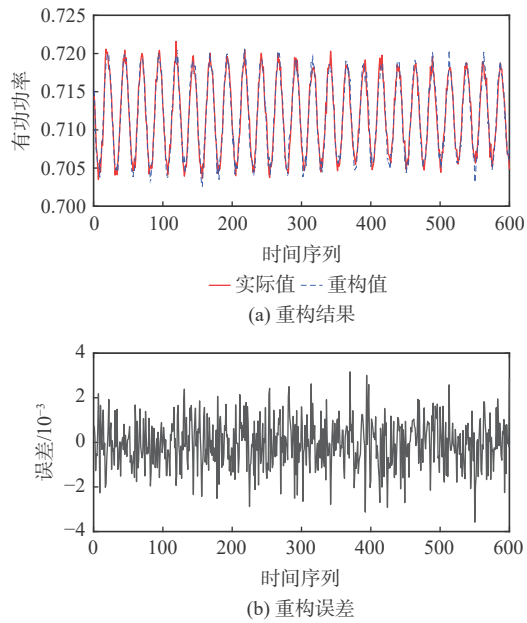


图 9 辅助变量筛选前后 ELM 的重构结果及重构误差

Fig. 9 Reconstruction results and reconstruction errors of ELM before and after auxiliary variable screening

表 3 不同辅助变量个数的重构结果对比

Table 3 Comparison of reconstruction results with different number of auxiliary variables

k	相对误差/%	时间/s
4	0.14	0.871 69
6	0.17	1.026 34

5 结 论

本文针对当前复杂的传感器故障诊断算法给系统带来较大的运算负担，较为系统的分析和研究了传感器的故障检测和数据重构，通过时间序列预测模型和冗余度概念的引入，在不增加计算量的前提下，能够在一定的范围内解决传感器的故障检测

问题, 同时提高传感器的数据重构精度。

然而, 对于现有电源车传感器数据重构方法大多是针对系统单一的运行工况下, 而没有考虑特殊极端的环境因素的影响, 因此, 在未来工作中希望能够实现电源车传感器在线数据采集和实时动态的在线重构。

参考文献（References）

[1] 李炜, 周丙相, 蒋栋年. 基于深度学习序贯检验的电源车故障诊断方法[J]. 系统仿真学报, 2020, 32(4): 638-648.
LI W, ZHOU B X, JIANG D N. Fault diagnosis method of vehicle power supply based on deep learning and sequential test[J]. Journal of System Simulation, 2020, 32(4): 638-648(in Chinese).

[2] 李炜, 周丙相, 蒋栋年, 等. 基于多状态时间序列预测学习的电源车故障预测方法[J]. 吉林大学学报, 2020, 50(4): 1532-1544.
LI W, ZHOU B X, JIANG D N, et al. Fault prediction of power supply vehicle based on multi-state time series prediction learning[J]. Journal of Jilin University, 2020, 50(4): 1532-1544(in Chinese).

[3] MANDAL S, SANTHI B, SRIDHAR S, et al. Minor fault detection of thermocouple sensor in nuclear power plants using time series analysis[J]. Annals of Nuclear Energy, 2019, 134: 383-389.

[4] MEJIA J, OCHOA-ZEZZATTI A, CRUZ-MEJÍA O, et al. Prediction of time series using wavelet Gaussian process for wireless sensor networks[J]. Wireless Networks, 2020, 26(12): 5751-5758.

[5] WU J, HU K, CHENG Y W, et al. Data-driven remaining useful life prediction via multiple sensor signals and deep long short-term memory neural network[J]. ISA Transactions, 2020, 97: 241-250.

[6] NIU G, XIONG L, QIN X, et al. Fault detection isolation and diagnosis of multi-axle speed sensors for high-speed trains[J]. Mechanical Systems and Signal Processing, 2019, 131: 183-198.

[7] JEONG S, FERGUSON M, HOU R, et al. Sensor data reconstruction using bidirectional recurrent neural network with application to bridge monitoring[J]. Advanced Engineering Informatics, 2019, 42(10): 1-14.

[8] XIAO Y, YIN H, ZHANG Y, et al. A dual-stage attention-based Conv-LSTM network for spatio-temporal correlation and multivariate time series prediction[J]. International Journal of Intelligent Systems, 2021, 36(5): 2036-2057.

[9] LIN T H, WANG T C, WU S C. Deep learning schemes for event identification and signal reconstruction in nuclear power plants with sensor faults[J]. Annals of Nuclear Energy, 2021, 154: 108-113.

[10] ELNOUR M, MESKIN N, AL-NAEMI M. Sensor data validation and fault diagnosis using auto-associative neural network for HVAC systems[J]. Journal of Building Engineering, 2020, 27: 100935.

[11] NASROLAHI S S, ABDOLLAHI F. Sensor fault detection and recovery in satellite attitude control[J]. Acta Astronautica, 2018, 145(4): 275-283.

[12] DU J, CHEN H, ZHANG W. A deep learning method for data recovery in sensor networks using effective spatio-temporal correlation data[J]. Sensor Review, 2018, 39(3): 1-11.

[13] TAYEH G B, MAKHOUL A, PERERA C, et al. A spatial-temporal correlation approach for data reduction in cluster-based sensor networks[J]. IEEE Access, 2019, 7: 50669-50680.

[14] ZHAO Z, LIU Z, SUN Y, et al. WOS-ELM-based double redundancy fault diagnosis and reconstruction for aeroengine sensor[J]. Journal of Control Science and Engineering, 2017, 2017: 1-14.

[15] 张弦, 王宏力. 具有选择与遗忘机制的极端学习机在时间序列预测中的应用[J]. 物理学报, 2011, 60(8): 74-80.
ZHANG X, WANG H L. Selective forgetting extreme learning machine and its application to time series prediction[J]. Acta Physica Sinica, 2011, 60(8): 74-80(in Chinese) .

[16] SMEJKAL T, MIKYŠKA J. Efficient solution of linear systems arising in the linearization of the VTN-phase stability problem using the Sherman-Morrison iterations[J]. Fluid Phase Equilibria, 2021, 527: 1-12.

[17] 蒋栋年, 李炜. 基于自适应阈值的粒子滤波非线性系统故障诊断[J]. 北京航空航天大学学报, 2016, 42(10): 2099-2106.
JIANG D N, LI W. Fault diagnosis of particle filter nonlinear systems based on adaptive threshold[J]. Beijing University of Aeronautics and Astronautics, 2016, 42(10): 2099-2106(in Chinese).

[18] SEKERKA R F. Entropy and information theory[M]//SEKERKA R F. Thermal physics: Thermodynamics and statistical mechanics for scientists and engineers. Berlin: Springer, 2015: 247-256.

[19] OSBORNE J W. Improving your data transformations: Applying the Box-Cox transformation[J]. Practical Assessment Research & Evaluation, 2010, 15(1): 1-10.

[20] SARMADI H, KARAMODIN A. A novel anomaly detection method based on adaptive Mahalanobis-squared distance and one-class kNN rule for structural health monitoring under environmental effects[J]. Mechanical Systems and Signal Processing, 2020, 140: 1-24.

Sensor fault detection and data reconstruction method of power supply vehicle

JIANG Dongnian[1, 2], BA Yujiang[1], LI Wei[1, 3, *]

(1. College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China;
2. State Grid Gansu Electric Power Research Institute, Lanzhou 730050, China;
3. Key Laboratory of Gansu Advanced Control for Industrial Processes, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: Aiming at the problem that the power supply vehicle is prone to sensor fault due to the complex operating environment, a sensor fault detection and data reconstruction method based on spatiotemporal correlation is proposed in this paper. Firstly, according to the time-series relationship characteristics of single sensor operation data, the fault detection sub-model of the power vehicle sensor is established with the help of the selective forgetting extreme learning machine (SF-ELM) mechanism, and the fault detection of power vehicle sensor is realized. Secondly, simultaneous interpreting the fault sensors, using the spatial correlation among different sensors, and through redundancy analysis, the improved mutual information entropy is used to screen out the auxiliary sensor data which is highly correlated with the fault sensor data, and the online reconstruction of the failure sensor data is realized. Finally, the feasibility and effectiveness of the proposed method in sensor fault detection and data reconstruction of power vehicles are verified by simulation.

Keywords: sensor; spatiotemporal correlation; extreme learning machine; fault detection; data reconstruction

Received: 2021-08-30; **Accepted**: 2021-11-20; **Published Online**: 2021-12-15 08:56

URL: kns.cnki.net/kcms/detail/11.2625.V.20211214.1415.001.html

Foundation items: National Natural Science Foundation of China (62263020); Outstanding Youth Fund of Gansu Province (20JR10RA202); Hongliu Outstanding Young Talents Support Project of Lanzhou University of Technology; Lanzhou Science and Technology Plan (2022-2-69)

* **Corresponding author**. E-mail: liwei@lut.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0546

特征点分段提取的时间序列模式匹配方法

李正欣*，刘畅，吴诗辉，郭建胜
(空军工程大学 装备管理与无人机工程学院，西安 710051)

摘　　要：常用的时间序列模式匹配方法难以平衡计算复杂度与匹配精度，针对该问题，提出了一种特征点分段提取的时间序列模式匹配方法。提取时间序列每个变量维度上的特征点，降低序列长度；将特征点序列转化为分位点矩阵，利用欧氏距离对分位点矩阵进行相似性度量；在几组时间序列数据集上对所提方法进行分类实验。结果表明：所提方法在降低计算复杂度的同时，获得了较高的匹配精度。

关　键　词：时间序列；模式匹配；特征提取；分位点矩阵；计算复杂度
中图分类号：TP391
文献标志码：A　　**文章编号**：1001-5965(2023)07-1593-07

现实世界中，存在着大量时间序列类型数据[1]，如捕捉手语动作产生的信号[2]，电商平台中用户的浏览记录[3]，股票交易中的交易量、开盘价、收盘价[4]等形成的数据。时间序列是一组记录事物状态随时间发生变化的观测值[5]，可定义为$x_t(j)$，其中，$t(t=1,2,\cdots,m)$表示t个时刻，$j(j=1,2,\cdots,n)$表示第j个变量，$x_t(j)$表示第j个变量在第t个时刻上的记录值[6]。当$n=1$时，$x_t(j)$为一元时间序列；当$n>1$时，$x_t(j)$为多元时间序列[7]。然而，随着云储存技术的快速发展，时间序列数据规模日益庞大[8]。

时间序列数据挖掘主要从规模庞大的时序数据中进行整理[9]，最终产生有效的信息[10]，其逐渐成为掌握当下现实事物规律的方法之一[11]。时间序列数据挖掘包括相似性查询[12]、关联分析[13]、分类[14]、聚类[15]、预测[16]、模式匹配[17]等内容。其中，模式匹配是描述时间序列之间相似性的方法。在开展时间序列数据挖掘其他工作时，需要先使用模式匹配方法。模式匹配又可细分为特征提取与相似性度量2个过程[18]。特征提取主要降低时间序列维数，相似性度量以降维后的特征序列为输入，由距离度量公式计算出2条时间序列的距离，距离越小说明2条时间序列相似程度越高[19]。

目前，常用的模式匹配方法在提高计算效率或提高度量准确率方面效果理想[20]，但能同时达到两者的方法不多。因此，研究既降低计算复杂度又保持较高匹配精度的模式匹配方法具有现实意义和理论价值[21]。为此，本文提出了一种特征点分段提取的时间序列模式匹配方法。首先，提出一种特征点分段提取方法，提取时间序列的特征点；然后，依据特征序列计算出分位点矩阵，利用欧氏距离对分位点矩阵进行相似性度量；最后，通过实验对本文方法进行有效性验证。

1　相关研究

常用的时间序列模式匹配方法包括奇异值分解(singular value decomposition, SVD)[22]、动态时间弯曲(dynamic time warping, DTW)[23]、自适应代价动态时间弯曲的多元时间序列相似性度量方法(adaptive cost multivariate dynamic time warping, ACM-DTW)[24]、趋势距离(trend distance, TD)[25]、欧氏距离(Euclidean distance, ED)[26]和基于点分布特征(point distribution, PD)的方法[27]等。

收稿日期：2021-09-13；录用日期：2021-12-10；网络出版时间：2022-02-15 14:35
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220215.1035.002.html
*通信作者. E-mail：lizhengxin_2005@163.com

引用格式：*李正欣，刘畅，吴诗辉，等. 特征点分段提取的时间序列模式匹配方法[J]. 北京航空航天大学学报，2023，49(7)：1593-1599.*
LI Z X，LIU C，WU S H，et al. Segmentation extraction of feature points for time series pattern matching[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1593-1599 (in Chinese).

SVD方法将时间序列不同变量放在一起处理，在特征提取阶段，建立相关系数矩阵并以线性变换的结果为输入，使用距离公式计算匹配精度。SVD的优点是能够体现不同变量的内在联系，但是会打乱时间序列的先后顺序，容易降低后续匹配精度的准确性。

DTW距离可以计算2条长度不相等的时间序列，具有较高准确性，在相似性度量中被广泛使用。但其计算代价较高，需要在众多弯曲路径中找出最短路径，在一些数据规模较大的时间序列上效果不理想。

ACM-DTW方法以DTW距离为基础，定义自适应代价函数赋给基础距离矩阵权重，在匹配过程中限制点列的使用来减少不合理匹配情况出现；ACM-DTW优化了DTW距离的弯曲路径，一定程度地降低DTW距离计算复杂度，但需要设置额外空间记录点列使用次数且参数初始化较复杂。

TD方法同样是以DTW距离为基础的方法，使用最小二乘法提取出时间序列的倾斜角与时间跨度，以此作为特征利用DTW距离完成相似性度量。TD能够有效降低DTW距离的计算代价，但需要事先设置参数且参数优化复杂。

ED距离计算复杂度低，其计算不同时间序列之间对应时刻的平方和并累积求和。ED距离要求序列等长，匹配精度也容易受到序列伸缩、弯曲等形变的影响。

PD方法以多元时间序列的局部重要点为基础，根据重要点的图像特征使用ED距离完成相似性度量。PD方法能够度量时间长度不相等的时间序列，但将不同变量和时间的观测值放在一起进行局部重要点提取，没有考虑变量之间的量纲差异，在大规模多元时间序列数据集上效果不理想。

针对以上模式匹配方法无法平衡匹配精度与计算复杂度的问题，本文提出了一种特征点分段提取的时间序列模式匹配方法。

2 特征点分段提取方法

根据时间序列的趋势特性得到序列特征，本节提出一种特征点分段提取的方法。利用滑动窗口[28]提取出时间序列各变量维度的特征点，压缩时间序列长度。时间序列$\boldsymbol{x}\in\mathbf{R}^{n\times m}$，$n$为变量维数，$m$为时间维数。通常这2个维数较高，为了降低计算复杂度，提取时间序列特征。由于时间序列不同变量代表着序列的不同属性，同时将变量和时间放在一起提取特征，则没有考虑变量之间的量纲差异，容易降低后续相似性度量结果。

在确保变量维数不变的情况下，使用滑动窗口提取特征点，提出一种特征点分段提取的方法。

输入： 时间序列$\boldsymbol{x}\in\mathbf{R}^{n\times m}$，滑动窗口因子$\Delta$。

输出： 特征序列**seg**。

步骤1 对于$\boldsymbol{x}$第j个变量$x(j)$，以第i观测值$x_i(j)$为滑动窗口中心，其中，$j=1,2,\cdots,n$，$i=1,2,\cdots,m$，滑动窗口为$\{x_{i-\Delta}(j),\cdots,x_i(j),\cdots,x_{i+\Delta}(j)\}$。

步骤2 如果观测值$x_i(j)$为当前所在滑动窗口内的最大或最小值，则观测值$x_i(j)$为特征点，并将特征点$x_i(j)$提取到序列**seg**(j)中。

步骤3 针对时间序列$\boldsymbol{x}\in\mathbf{R}^{n\times m}$所有变量，按照顺序执行步骤1、步骤2，得到一条$n$维的特征序列**seg**。

3 特征点分段提取的相似性度量

时间序列$\boldsymbol{x}\in\mathbf{R}^{n\times m}$提取出的特征点如图1所示。可以看出，不同特征序列上的特征点个数一般不相同，如果直接使用特征序列进行相似性度量，可能会因此增加后续度量的难度。

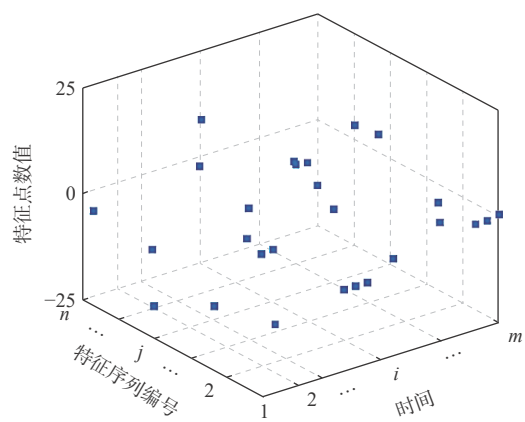


图1 特征点提取示意图

Fig. 1 Schematic diagram of feature point extraction

为此，引入统计学常用的9个分位点[29]（最小值点、5%、10%、25%、50%、75%、90%、95%、最大值点），利用特征序列构建特征矩阵。于是，提出一种分位点矩阵提取方法。首先，对第$j(j=1,2,\cdots,n)$个变量维度上的特征点序列**seg**(j)从小到大排序；然后，对排序后的序列**seg**(j)计算出9个分位点；最后，对其他特征点序列用同样方法提取9个分位点，形成分位点矩阵。

输入： 特征序列**seg**。

输出： 分位点矩阵$\boldsymbol{Q}$。

步骤1 针对**seg**的第$j(j=1,2,\cdots,n)$条特征序列**seg**(j)，对该序列上的特征点按照从小到大的方式排序。

步骤2 对排序后的特征序列，分别提取9个分位点：最小值点、5%、10%、25%、50%、75%、

90%、95%、最大值点, 构成分位点序列$\boldsymbol{Q}(j)$。

步骤 3　对于其他变量维度上的特征序列, 依次执行步骤 1、步骤 2, 最终形成分位点矩阵$\boldsymbol{Q}$。

通过以上方法, 将时间序列$\boldsymbol{x}\in\mathbf{R}^{n\times m}$转换为分位点矩阵$\boldsymbol{Q}\in\mathbf{R}^{n\times 9}$。需要注意的是, 分位点矩阵不改变时间序列的变量维度, 将时间维度压缩为 9。对于 2 条长度不同的时间序列, 其分位点矩阵结构相同。这样, 可以直接用欧氏距离度量 2 条序列的分位点矩阵。为不失一般性, 采用欧氏距离的平方刻画 2 条时间序列的距离, 如下:

$$D_{\mathrm{ED}}(\boldsymbol{Q}^x,\boldsymbol{Q}^y)=\sum_{i=1}^{n}\sum_{j=1}^{9}(q_{ij}^x-q_{ij}^y)^2 \tag{1}$$

式中: $\boldsymbol{Q}^x$、$\boldsymbol{Q}^y$分别为时间序列$\boldsymbol{x}$、$\boldsymbol{y}$的分位点矩阵; q_{ij}^x、q_{ij}^y分别为 2 个分位点矩阵第i行、第j列上的元素。显然, 式(1)的计算复杂度为$O(n)$。

为了后续实验的顺利进行, 将特征点分段提取的时间序列模式匹配方法记为 SFP-ED。

4　实验与结果分析

本节使用 6 组时间序列数据集, 对其中所有样本逐一使用 5 种模式匹配方法进行实验, 并进行匹配精度比较, 同时, 对 SFP-ED 方法的参数进行敏感性分析。

4.1　实验方法与实验数据

本文实验环境为 Intel(R) Core(TM) i5-9300H CPU, 16 GB RAM, Windows 10, MATLAB R2018a。参与实验对比的模式匹配方法有 SVD、DTW、ACM-DTW、TD 和 PD。其中, DTW 方法直接使用原始时间序列进行距离度量, 输出计算结果。

实验方法采用 k 近邻法与留一法结合的形式进行, 确定时间序列数据集的样本个数 n, 任选其中一个样本 T 作为输入序列。首先, 使用 6 种方法找到与 T 最近似的 k 个序列。然后, 找出 k 个序列中与序列 T 标签相同的样本 N。最后, 将 N 与 k 的比值作为度量结果。将时间序列数据集中的其余样本, 依次作为输入序列, 重复以上实验, 得到 n 个度量结果, 匹配精度即为这 n 个度量结果的平均值[30]。实验采用类别标签已知的 6 组公开数据集, 前 5 组为多元时间序列数据集, 后 1 组为一元时间序列数据集, 详情如表 1 所示。

表 1　实验数据集

Table 1　Detail of experimental datasets

数据集	长度	均长	变量个数	样本量	类别
ASL	47～95	59	22	216	8
LP1	15	15	6	88	4
JV	7～29	16	12	640	9
EEG	256	256	64	22	2
WR	128～1 918	368	62	44	2
Trace	275	275	1	200	4

4.2　实验结果比较

4.2.1　ASL 数据集实验结果

在 ASL 数据集上, 使用第 1 节中的模式匹配方法与 SFP-ED 方法进行实验, 其中, TD 方法(maxError=0.03、ε=0.8、λ=0.2)、PD 方法($X[i-1{:}i+1,j-1{:}j+1]$)、SFP-ED($\Delta=1$)结果如表 2 所示。

表 2　ASL 数据集上的匹配精度

Table 2　Matching accuracy on ASL dataset

k	SVD	DTW	ACM-DTW	TD	PD	SFP-ED
1	0.718	0.977	0.926	0.972	0.579	0.986
5	0.639	0.929	0.870	0.958	0.570	0.947
10	0.570	0.929	0.856	0.906	0.557	0.944

DTW、ACM-DTW、TD 和 SFP-ED 方法的匹配精度较高。其中, TD 方法提取序列的倾斜角、时间跨度作为特征, ACM-DTW 方法设置自适应损失函数提高 DTW 距离运算速率, SFP-ED 方法提取序列的特征点作为特征。DTW、ACM-DTW 和 TD 方法是以 DTW 距离为基础, 而 SFP-ED 方法以欧氏距离为基础, 使得 SFP-ED 的计算复杂度较低于 DTW 和 TD 方法。

SVD、PD 方法匹配精度较低。这是因为 SVD 方法打乱了时序顺序, 存在误判风险; PD 方法将相邻变量与时间放在一起提取局部重要点, 没有考虑变量之间的量纲差异; SFP-ED 方法针对每一个变量维度提取特征点, 能够体现不同变量之间的差异。

表 3 给出了 6 种方法在不同准确率 e 下的匹配次数。在准确率较低的情况下 (e<40%), DTW、ACM-DTW、TD 和 SFP-ED 方法对应的匹配次数比 SVD、PD 方法少; 在准确率较大的情况下 ($e\geqslant$80%), DTW、ACM-DTW、TD 和 SFP-ED 方法所对应的匹配次数比 SVD、PD 方法多。

进一步地, 将 ASL 数据集中第 87 个序列(ASL_87)作为输入(见图 2), 分别采用上述 6 种方法, 找出最相似的序列, 如图 3 所示。SVD、TD 与 SFP-ED 方法找出的序列与输入序列较为相似, 且都属于同一类(boy)。而 DTW、ACM-DTW 和 PD 方法形状差别较大。

4.2.2　其他数据集实验结果

使用 SVD、PD、TD、DTW、ACM-DTW、SFP-ED 方法在 ASL、LP1、JV、EEG、WR 多元时间序列数据集上进行实验, 计算匹配精度, 因为 SVD、PD 是

表3 6种模式匹配方法在不同准确率下的匹配次数

Table 3 Matching times of 6 pattern matching methods under different accuracy rates

e/%	SVD			DTW			ACM-DTW			TD			PD			SFP-ED		
	N=1	N=5	N=10	N=1	N=5	N=10	N=1	N=5	N=10	N=1	N=5	N=10	N=1	N=5	N=10	N=1	N=5	N=10
0	61	20	12	5	0	0	16	9	0	6	0	0	91	33	14	3	0	0
10			19			0			3			0			13			0
20		27	15		0	0		4	0		0	0		25	17		2	0
30			19			1			2			2			28			3
40		34	19		13	0		10	5		4	3		36	18		9	4
50			25			6			5			9			26			1
60		29	26		12	5		10	13		9	8		34	17		6	3
70			7			15			15			7			13			10
80		22	7		14	11		29	17		15	22		23	6		10	7
90			7			30			25			30			12			15
100	155	84	60	211	177	148	200	154	125	210	188	135	125	65	52	213	189	173

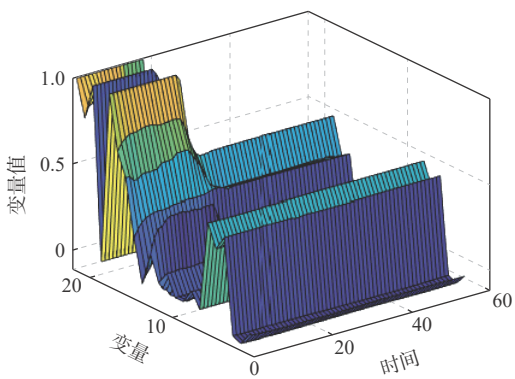


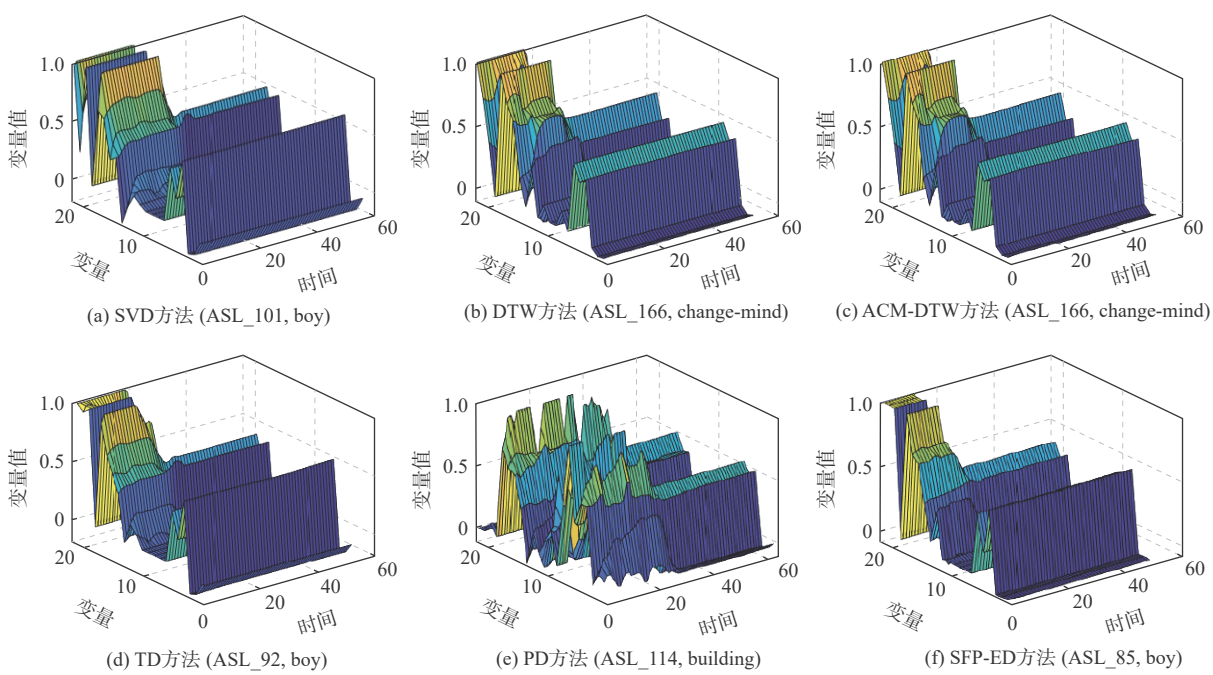
图2 输入序列 (ASL_87, boy)

Fig. 2 Input sequence (ASL_87, boy)

基于多元时间序列的模式匹配方法，所以仅使用DTW、ACM-DTW、TD、SFP-ED方法在Trace一元时间序列数据集上进行实验。实验结果为 k=1 时的匹配精度，如表4所示。其中，PD、TD、SFP-ED方法采取上述相同的参数设置[30]。DTW、ACM-DTW、TD和SFP-ED方法在6组数据集上匹配精度较高。SVD和PD方法效果不理想。

4.3 SFP-ED方法参数敏感性分析

本节分析滑动窗口因子 Δ 对SFP-ED方法的压缩程度和匹配精度的影响，实验结果如图4所示。压缩程度近似表示为



(a) SVD方法 (ASL_101, boy)

(b) DTW方法 (ASL_166, change-mind)

(c) ACM-DTW方法 (ASL_166, change-mind)

(d) TD方法 (ASL_92, boy)

(e) PD方法 (ASL_114, building)

(f) SFP-ED方法 (ASL_85, boy)

图3 不同方法在ASL数据集上寻找的最相似序列

Fig. 3 The most similar sequence found by different methods on ASL dataset

表 4　6 组数据集上的匹配精度

Table 4　Accuracy on 6 datasets

数据集	SVD	DTW	ACM-DTW	TD	PD	SFP-ED
ASL	0.718	0.977	0.926	0.972	0.578	0.986
LP1	0.477	0.886	0.909	0.636	0.830	0.921
JV	0.489	0.963	0.759	0.548	0.581	0.958
EEG	0.909	0.727	0.727	1	0.591	0.864
WR	0.977	0.977	0.955	1	1	0.977
Trace		0.995	1	0.930		0.995

$$\lambda = \frac{\bar{l}}{m} \tag{2}$$

式中: $\bar{l}$为特征序列的平均长度; m为时间序列的时间维长度。

实验结果表明, 随着参数Δ的增大, $\bar{l}$减少, λ也相应减小, 匹配精度变化不大。这表明Δ在一定范围内变化, 对 SFP-ED 方法的匹配精度影响较小。为了直观地展示Δ对匹配精度的影响, 选取 ASL 数据集中第 204 号样本, 以其第 1 个变量维度序列为分析对象, 分别给出$\Delta = 1,5,10$三种情况下特征点的提取情况, 如图 5 所示。同时, 计算出 3 种情况下对应的 9 个分位点, 如表 5 所示。Δ变大, 特征点数量减少, 而提取的最大值点和最小值点保持不变, 因此, 9 个分位点的变化不大。

4.4　计算复杂度比较

在整个实验过程中, TD、DTW 和 SFP-ED 方法性能较好。其中, TD、DTW 方法在相似性度量阶段都使用 DTW 距离进行计算, SFP-ED 方法则使用欧氏距离进行计算。TD 方法以每拟合段的斜率和

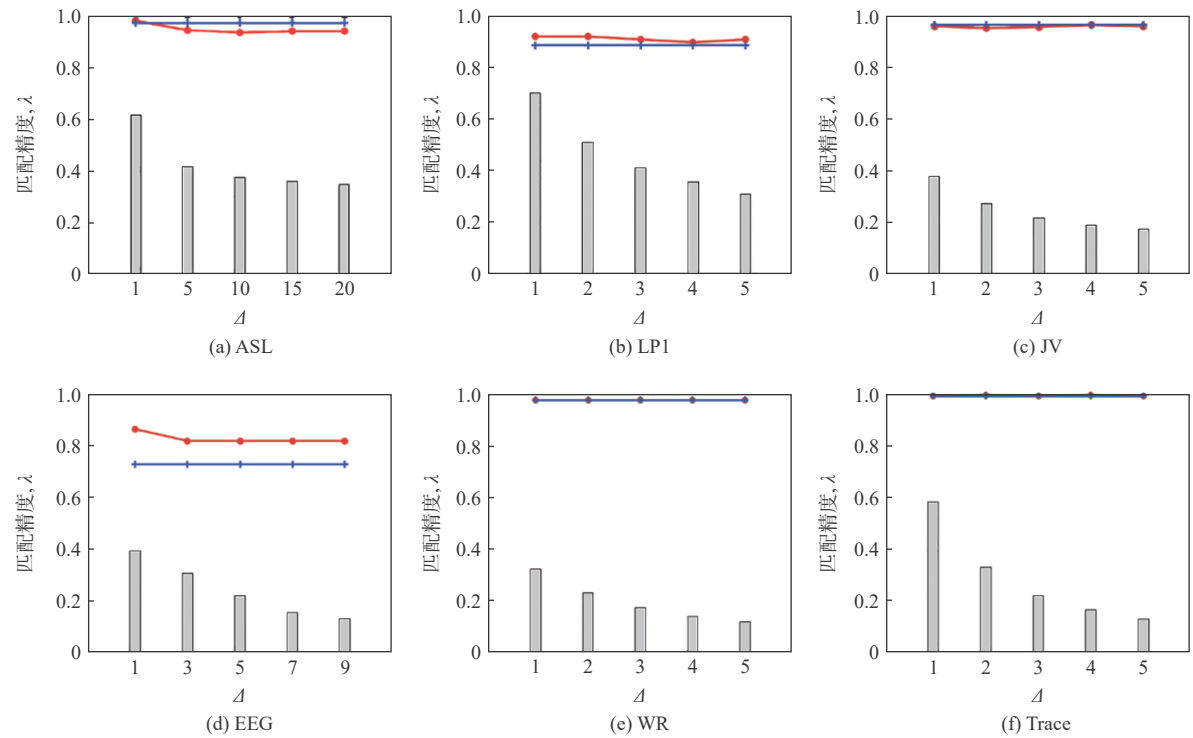


图 4　滑动窗口因子对匹配精度及计算效率的影响

Fig. 4　Influence of sliding window factor on matching accuracy and calculation efficiency

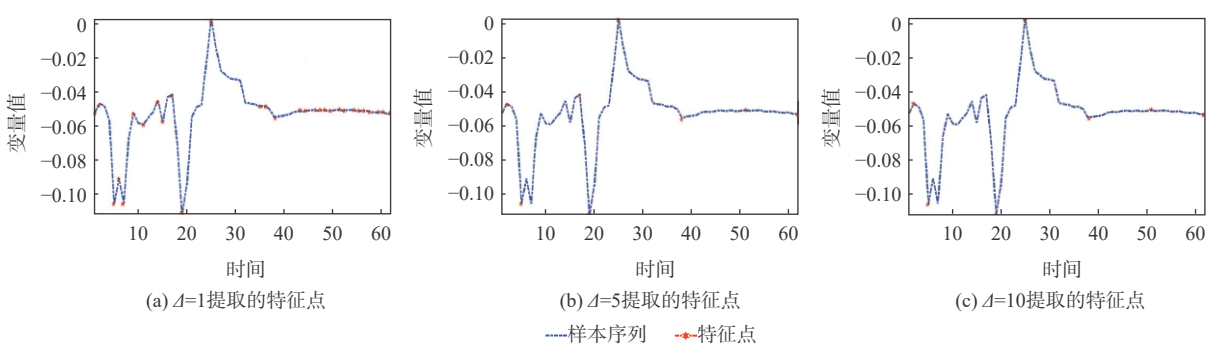


图 5　滑动窗口因子对特征点提取的影响

Fig. 5　Influence of sliding window factor on feature point extraction

表5 9个分位点提取结果

Table 5 Nine quantile results

Δ	最小值点	5%	10%	25%	50%	75%	90%	95%	最大值点
1	−0.111 1	−0.105 2	−0.097 5	−0.052 8	−0.051 2	−0.050 6	−0.046 0	−0.041 4	0.002 4
5	−0.111 1	−0.111 1	−0.109 3	−0.080 0	−0.051 6	−0.044 1	−0.010 8	0.002 4	0.002 4
10	−0.111 1	−0.111 1	−0.109 9	−0.092 6	−0.052 8	−0.047 7	−0.007 5	0.002 4	0.002 4

时间跨度为特征，使用DTW距离完成度量，SFP-ED方法以分位点矩阵为输入，使用欧氏距离完成度量。本节分析3种方法的计算复杂度。

已知DTW距离计算复杂度为$O(n(m_1 \times m_2))$，SFP-ED方法的计算复杂度为$O(n)$，其中，m_1和m_2分别为2个参与度量序列的长度，n为变量的个数。显然SFP-ED方法的计算复杂度要低于DTW和TD方法。为了进一步比较3种方法的计算复杂度，记录3种模式匹配方法的计算时间，如图6所示。可以看出，在数据集ASL、JV、EEG、WR、Trace上，SFP-ED方法的计算时间低于TD方法，在数据集JV上，SFP-ED方法的计算时间高于TD方法。

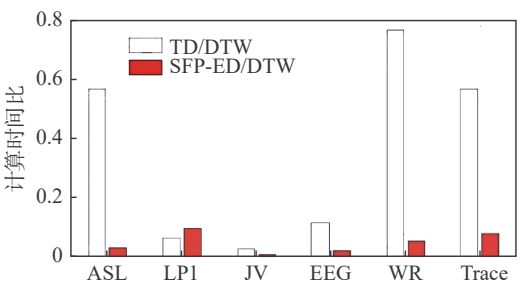


图6 TD、SFP-ED与DTW计算时间比较

Fig. 6 Calculation time of TD、SFP-ED and DTW

5 结 论

SFP-ED方法在时间序列每个变量维度上分别提取特征，体现了变量之间的差异，能够获得较高的匹配精度。SFP-ED方法将提取的特征序列进一步转化为分位点矩阵，采用欧氏距离进行相似性度量，计算复杂度较低。此外，该方法能同时适用于一元和多元时间序列。

SFP-ED方法提取特征序列时，能够保持序列的时序关系。但在相似性度量阶段，将特征序列转化为分位点矩阵，会破坏特征序列的时序关系。如何保持序列的时序关系是下一步需要继续研究的问题。

参考文献（References）

[1] BOHANNON A W, LAWHERN V J, WAYTOWICH N R, et al. The autoregressive linear mixture model: A time-series model for an instantaneous mixture of network processes[J]. IEEE Transactions on Signal Processing, 2020, 68: 4481-4496.

[2] WANG F, LI C, ZENG Z, et al. Cornerstone network with feature extractor: A metric-based few-shot model for Chinese natural sign language[J]. Applied Intelligence, 2021, 51(5): 1-12.

[3] PAN H, ZHOU H. Study on convolutional neural network and its application in data mining and sales forecasting for E-commerce[J]. Electronic Commerce Research, 2020, 20(2): 297-320.

[4] JIANG X Q, ZHANG L C. Stock price fluctuation prediction method based on time series analysis[J]. Discrete and Continuous Dynamical Systems, 2018, 12(4): 915-927.

[5] WEN S C, YANG C H. Time series analysis and prediction of nonlinear systems with ensemble learning framework applied to deep learning neural networks[J]. Information Sciences, 2021, 572: 167-181.

[6] CABRERA D, SANCHO F, CERRADA M, et al. Knowledge extraction from deep convolutional neural networks applied to cyclostationary time-series classification[J]. Information Sciences, 2020, 524: 1-14.

[7] MA Q, CHEN Z, TIAN S, et al. Difference-guided representation learning network for multivariate time series classification[J]. IEEE Transactions on Cybernetics, 2020, 3: 1-11.

[8] 李正欣, 张凤鸣, 李克武, 等. 一种支持DTW距离的多元时间序列索引结构[J]. 软件学报, 2014, 25(3): 560-575.
LI Z X, ZHANG F M, LI K W, et al. Index structure for multivariate time series under DTW distance metric[J]. Journal of Software, 2014, 25(3): 560-575(in Chinese).

[9] SOLEIMANI G, ABESSI M. DLCSS: A new similarity measure for time series data mining[J]. Engineering Applications of Artificial Intelligence, 2020, 92: 103664.

[10] ZHU H, WANG X, CHEN X, et al. Similarity search and performance prediction of shield tunnels in operation through time series data mining[J]. Automation in Construction, 2020, 114: 103178.

[11] LI H. Time works well: Dynamic time warping based on time weighting for time series data mining[J]. Information Sciences, 2021, 547: 592-608.

[12] OKAWA M. Time-series averaging and local stability-weighted dynamic time warping for online signature verification[J]. Pattern Recognition, 2021, 112: 107699.

[13] NIE C X. Nonlinear correlation analysis of time series based on complex network similarity[J]. International Journal of Bifurcation and Chaos, 2020, 14(4): 736-749.

[14] WANG B Q, JIANG T H, ZHOU X, et al. Time-series classification based on fusion features of sequence and visualization[J]. Applied Sciences, 2020, 10(12): 4124.

[15] LI H H, LIU J X, YANG Z L, et al. Adaptively constrained dynamic time warping for time series classification and clustering[J]. Information Sciences, 2020, 534: 97-116.

[16] NISHINO T, HOKUGO A. A stochastic model for time series prediction of the number of post-earthquake fire ignitions in buildings

based on the ignition record for the 2011 Tohoku Earthquake[J]. Earthquake Spectra, 2020, 36(1): 232-249.

[17] WAN Y Q, SI Y W. A hidden semi-Markov model for chart pattern matching in financial time series[J]. Soft Computing, 2018, 22: 6525-6544.

[18] SUN J, YANG Y, LIU Y, et al. Univariate time series classification using information geometry[J]. Pattern Recognition, 2019, 95: 832-843.

[19] HUANG W, YUE B, CHI Q, et al. Integrating data-driven segmentation, local feature extraction and fisher kernel encoding to improve time series classification[J]. Neural Processing Letters, 2019, 49(1): 43-66.

[20] YOU D G, LIN W W, SHI F, et al. A novel approach for CPU load prediction of cloud server combining denoising and error correction[J]. Computing, 2023, 105(3): 577-594.

[21] IOKIBE T, KOYAMA M, TANIGUCHI M. A study for complexity of chaotic time series and prediction accuracy[C]//IEEE International Fuzzy Systems Conference. Piscataway: IEEE Press, 1999: 6430833.

[22] MONTAGNON C E. Forecasting by splitting a time series using singular value decomposition then using both ARMA and a Fokker Planck equation[J]. Physica A:Statistical Mechanics and its Applications, 2020, 567(3): 536-545.

[23] THUY H, ANH D T, CHAU V. Efficient segmentation-based methods for anomaly detection in static and streaming time series under dynamic time warping[J]. Journal of Intelligent Information Systems, 2021, 56(3): 373-386.

[24] 孟晓静, 万源. 自适应代价动态时间弯曲的多元时间序列相似性度量[J]. 统计与决策, 2020, 14(2): 25-29.
MENG X J, WAN Y. Multivariate time series similarity measure for dynamic time warping of adaptive cost[J]. Statistics and Decision, 2020, 14(2): 25-29(in Chinese).

[25] 李正欣, 张凤鸣, 李克武. 多元时间序列模式匹配方法研究[J]. 控制与决策, 2011, 24(4): 565-570.
LI Z X, ZHANG F M, LI K W. Research on pattern matching method of multivariate time series[J]. Control and Decision, 2011, 24(4): 565-570(in Chinese).

[26] GUIGOU F, COLLET P, PARREND P. SCHEDA: Lightweight euclidean-like heuristics for anomaly detection in periodic time series[J]. Applied Soft Computing, 2019, 82(3): 1432-1443.

[27] 管河山, 姜青山, 王声瑞. 基于点分布特征的多元时间序列模式匹配方法[J]. 软件学报, 2009, 20(1): 67-69.
GUAN H S, JIANG Q S, WANG S R. Pattern matching method based on point distribution for multivariate time series[J]. Journal of Software, 2009, 20(1): 67-69(in Chinese).

[28] GUO X G, HONG P Y, LALEG-KIRATI T M. Calibration and validation for a real-time membrane bioreactor: A sliding window approach[J]. Journal of Process Control, 2021, 98(6): 92-105.

[29] JILANI T A, BURNEY A S, ARDIL C. A new quantile based fuzzy time series forecasting model[J]. Proceedings of World Academy of Science Engineering & Technology, 2008, 23(1): 32-41.

[30] 刘畅, 李正欣, 张晓丰, 等. 基于特征点界标过滤的时间序列模式匹配方法[J]. 计算机应用研究, 2022, 39(4): 1008-1012.
LIU C, LI Z X, ZHANG X F, et al. Time series pattern matching method based on feature point landmark filtering[J]. Application Research of Computers, 2022, 39(4): 1008-1012(in Chinese).

Segmentation extraction of feature points for time series pattern matching

LI Zhengxin*, LIU Chang, WU Shihui, GUO Jiansheng

(Equipment Management and Unmanned Aerial Vehicle Engineering College, Air Force Engineering University, Xi'an 710051, China)

Abstract: It is difficult for the common time series pattern matching methods to balance the computational complexity and matching accuracy. To solve this problem, a time series matching method based on segmented extraction of feature points is proposed. Firstly, the feature points on each variable dimension of the time series are extracted and the sequence length is compressed. Then, the quantile matrix is calculated according to the feature sequence, and the similarity of the quantile matrix is measured by Euclidean distance. Finally, the effectiveness of the proposed method is verified on the application data set. Experimental results show that the proposed method can effectively reduce the computational complexity and ensure high matching accuracy.

Keywords: time series; pattern matching; feature extraction; quantile matrix; computational complexity

Received: 2021-09-13; **Accepted**: 2021-12-10; **Published Online**: 2022-02-15 14:35

URL: kns.cnki.net/kcms/detail/11.2625.V.20220215.1035.002.html

* **Corresponding author**. E-mail: lizhengxin_2005@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0507

基于角度转化的空间碎片测角数据精密定轨方法

张耀[1, 2, 3]，刘静[1, 2, 3, *]

(1. 中国科学院国家天文台，北京 100101；2. 中国科学院大学 天文与空间科学学院，北京 100049；3. 平湖实验室，嘉兴 314200)

摘 要：空间碎片对在轨卫星的影响日益增大，持续确定和更新空间碎片轨道日益受到重视。基于角度转化思想，提出了一种空间碎片测角数据精密定轨方法。基于常用的两行根数（TLE）数据应用需求，根据差分思想建立基于SGP4/SDP4预报模型的TLE精密定轨方法；以赤经-赤纬数据为典型场景，分析原始定义法、赤经投影法的优缺点；提出将常规双元素角度资料转化为三元素角度资料的角度转化法。仿真结果表明：当观测数据较多集中于测站天顶方向时，角度转化法能够将精密定轨收敛速度提高25%，轨道精度提高2～10倍。

关 键 词：精密定轨；空间碎片；测角数据；SGP4/SDP4；两行根数

中图分类号：P135

文献标志码：A **文章编号**：1001-5965(2023)07-1600-06

随着航天活动的日益频繁，在轨空间碎片编目数量已经超过23 000颗，严重影响了在轨卫星的安全运行[1]。中国卫星每年遭遇的空间碎片近距离接近事件达到了近400次。确定和持续更新空间碎片轨道，预警空间碎片碰撞风险日益受到重视。

空间碎片轨道确定指根据雷达、望远镜及其他观测设备的观测数据，确定其轨道，一般分为初轨确定和精密定轨2种。初轨确定指在无先验信息的情况下，根据一个观测弧段确定轨道[2]。精密定轨指在初始轨道的基础上，利用多段观测数据，考虑复杂的动力学模型，对初始轨道进行改进，获取精确的轨道根数。

定轨结果的精度和收敛速度是评价精密定轨方法的重要指标。影响精密定轨结果的因素很多，如观测弧段分布、数据密度、状态转移矩阵计算、观测数据精度、轨道预报模型、定轨权重和观测量处理形式等。数据密度和观测弧段分布对定轨结果影响较大，当两者合理分布时，稀疏弧段的定轨效果可以达到全监测数据按照相同密度平均覆盖整个定轨时间段的定轨效果[3]。状态转移矩阵的计算形式有数值积分形式和分析法形式，这2种形式对精密定轨的精度不会产生实质影响，但会影响精密定轨的计算效率[4]。精密定轨的基本弧段需求为3段以上，且最好跨越2个轨道周期[5]，采用的轨道预报模型需与观测数据精度相匹配[6-9]。国外主要针对观测数据精度和轨道预报模型对精密定轨的影响进行了系统性分析，美国国家研究委员会于2012年发布的评估报告[10]重点分析了大气模型对定轨和预报的影响。目前，定轨权重的研究方向主要是基于测量误差确定权重[11-12]，针对观测量不同处理方法的研究较少。

美国联合太空作战中心采用两行根数(two-line-element, TLE)形式发布了包含绝大部分10 cm以上在轨空间碎片的轨道根数，在全球范围内应用广泛。然而，美国并未发布TLE数据的生成代码(精密定轨软件)，其他科研机构在使用SGP4/SDP4进行预报时产生了困难。韦栋和赵长印[13]推导了定轨偏微分矩阵的数学表达式，建立了基于SGP4/SDP4

收稿日期：2021-09-01；**录用日期**：2022-01-25；**网络出版时间**：2022-02-23 14:53
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220223.0859.002.html
基金项目：国家自然科学基金 (12073045,11803052)
* **通信作者**. E-mail：liujing@bao.ac.cn

引用格式：*张耀，刘静. 基于角度转化的空间碎片测角数据精密定轨方法 [J]. 北京航空航天大学学报，2023，49（7）：1600-1605.*
ZHANG Y，LIU J. Precise orbit determination method for angle-only observation data of space debris based on angle conversion theory [J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（7）：1600-1605 （in Chinese）.

的精密定轨模型, 但该模型中定轨方法与预报方法深度绑定, 不便于预报方法的便捷更换。

基于上述问题, 本文采用差分方式建立 SGP4/SDP4 定轨生成 TLE 根数的模型, 分析现有观测数据处理方法的问题, 提出角度转化精密定轨方法, 仿真分析其对定轨结果的影响。

1　TLE 精密定轨模型

精密定轨涉及到状态方程和测量方程, 如下:

$$\boldsymbol{X}_t = F(\boldsymbol{X}_0, t) \qquad t = 1, \cdots, l \tag{1}$$

$$\boldsymbol{Y}_t = G(\boldsymbol{X}_t) + \boldsymbol{\varepsilon}_t \qquad t = 1, \cdots, l \tag{2}$$

式中: $\boldsymbol{X}_0$为t_0时刻的状态量; $\boldsymbol{X}_t$为t_i时刻的状态量, 用n维向量表示; $\boldsymbol{Y}_t$为t_i时刻的观测量, 用p维向量表示; $\boldsymbol{\varepsilon}_t$为观测量$\boldsymbol{Y}_t$的测量误差。

将状态方程和测量方程在参考值$\boldsymbol{X}^*$处泰勒展开, 舍去二阶及以上高阶项, 为方便表达, 引入如下 2 个小量:

$$\boldsymbol{x}_t = \boldsymbol{X}_t - \boldsymbol{X}_t^* \tag{3}$$

$$\boldsymbol{y}_t = \boldsymbol{Y}_t - \boldsymbol{Y}_t^* \tag{4}$$

略去推导过程, 得到残差$\boldsymbol{y}_t$与轨道根数修正值$\boldsymbol{x}_t$的关系式:

$$\begin{cases} \boldsymbol{y}_t = \boldsymbol{B}\boldsymbol{x}(t_0) + \boldsymbol{\varepsilon}_t \\ \boldsymbol{B} = \boldsymbol{H}\boldsymbol{\Phi} \\ \boldsymbol{H} = \left[\dfrac{\partial G}{\partial \boldsymbol{X}}\right]_t^* \\ \boldsymbol{\Phi} = \left[\dfrac{\partial F(t)}{\partial \boldsymbol{X}(t)}\right]_{t=t_0}^* \end{cases} \tag{5}$$

式中: $\boldsymbol{H}$为测量矩阵; $\boldsymbol{\Phi}$为状态转移矩阵; $\boldsymbol{B}$为偏微分矩阵; G和$F(t)$分别为测量方程和状态方程。

由于有多组观测数据, 采用最小二乘法求解状态量的改正值, 得到

$$J(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{\varepsilon}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{\varepsilon} = \sum_{i=1}^{l}\frac{1}{2}\boldsymbol{\varepsilon}_i^{\mathrm{T}}\boldsymbol{W}_i\boldsymbol{\varepsilon}_i \tag{6}$$

式中: $\boldsymbol{W}$为权重矩阵。

根据$\partial J(\boldsymbol{x})/\partial \boldsymbol{x} = 0$, 得到

$$\boldsymbol{x} = \left(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{y} \tag{7}$$

式中: $\boldsymbol{x}$为初始状态量的修正值; $\boldsymbol{y}$为残差。将参考值加上修正值, 获得新的参考值, 重新开始进行计算, 直到满足最小二乘收法收敛条件, 最终获得定轨根数。

采用 SGP4/SDP4 预报程序进行 TLE 定轨时, 通常需要进行偏微分矩阵$\boldsymbol{B}$计算, 获取观测量相对轨道根数的测量矩阵$\boldsymbol{H}$及 7 个轨道根数(轨道 6 个根数和β^*)状态转移矩阵$\boldsymbol{\Phi}$, 计算推导过程具体如下:

$$\begin{cases} \boldsymbol{B} = \dfrac{\partial \boldsymbol{O}_{\mathrm{bs}}}{\partial \boldsymbol{\sigma}_0} = \boldsymbol{H}\boldsymbol{\Phi} \\ \boldsymbol{H} = \dfrac{\partial \boldsymbol{O}_{\mathrm{bs}}}{\partial \boldsymbol{\sigma}_t} = \dfrac{\partial \boldsymbol{O}_{\mathrm{bs}}}{\partial (a_t, e_t, i_t, \Omega_t, w_t, M_t, \beta^*)} \\ \boldsymbol{\Phi} = \dfrac{\partial \boldsymbol{\sigma}_t}{\partial \boldsymbol{\sigma}_0} = \dfrac{\partial (a_t, e_t, i_t, \Omega_t, w_t, M_t, \beta^*)}{\partial (a_0, e_0, i_0, \Omega_0, w_0, M_0, \beta^*)} \end{cases} \tag{8}$$

式中: $\boldsymbol{O}_{\mathrm{bs}}$为观测数据; a_t、e_t、i_t、Ω_t、w_t、M_t和β^*分别为轨道 6 个根数和大气参数。

采用推导方式计算偏微分矩阵, 每个观测数据对应的矩阵只需要进行 1 次简单的近似预报, 计算效率高, 但推导公式复杂, 且更换预报模型就需要重新推导偏微分矩阵。采用差分思想计算偏微分矩阵, 每个观测数据对应的偏微分矩阵至少需要进行 7 次(甚至更多次)考虑复杂力模型的轨道预报, 计算量较大, 但能够避免复杂的公式推导, 方便在定轨模型中更换不同的预报模型[14]。因此, 本文中采用差分思想建立偏微分矩阵:

$$\frac{\partial \boldsymbol{O}_{\mathrm{bs}}}{\partial \boldsymbol{\sigma}_0} = \frac{f(\boldsymbol{\sigma}_0 + \Delta\boldsymbol{\sigma}) - f(\boldsymbol{\sigma}_0)}{\Delta\boldsymbol{\sigma}} \tag{9}$$

式中: f为轨道根数和观测量之间的计算模型; $\Delta\boldsymbol{\sigma}$为初始时刻轨道根数$\boldsymbol{\sigma}_0$的微小增量, 一般取 1%～2% 为宜。

2　观测量的不同处理方法

在精密定轨过程中, 利用观测量实现对初始轨道根数的改进, 获得精密的空间碎片轨道根数。本节以光学赤经-赤纬观测数据为例分析 3 种不同处理方法对精密定轨的影响。

2.1　原始定义法

原始定义法是在站心天球坐标系中描述观测量。站心天球坐标系(见图 1)原点为测站, x 轴、y 轴和 z 轴的指向分别与 J2000 地心天球坐标系的 x 轴、y 轴和 z 轴平行, 测量值的赤经和赤纬定义如下: ①赤纬δ, 指测站-空间碎片连线和平赤道面的夹角; ②赤经α, 指测站-空间碎片连线方向在平赤道面的投影与春分点方向的夹角(逆时针方向为正); ③当前时刻空间碎片到测站的距离ρ。

当观测数据位于站心天球坐标系的天顶方向附近时, 位置发生一点变化, 就可能引起赤经发生几度至几十度的变化, 根据式(5)计算 $\boldsymbol{B}$ 矩阵及赤经残差时, 两者结果不连续, 影响精密定轨收敛效果。

2.2　赤经投影法

为解决原始定义法引起的上述问题, 刘林和汤靖师[15] 提出了赤经投影法, 作了如下改进:

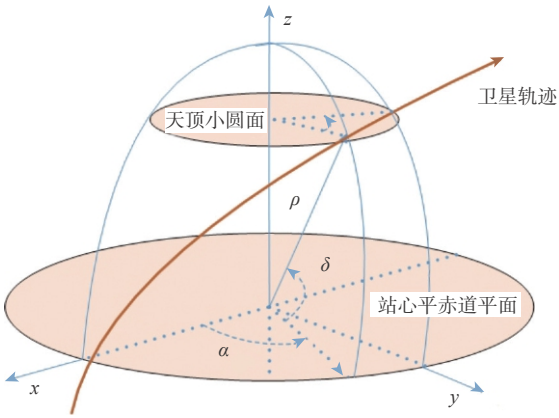


图1 赤经和赤纬的定义

Fig. 1 Definition of right ascension and declination

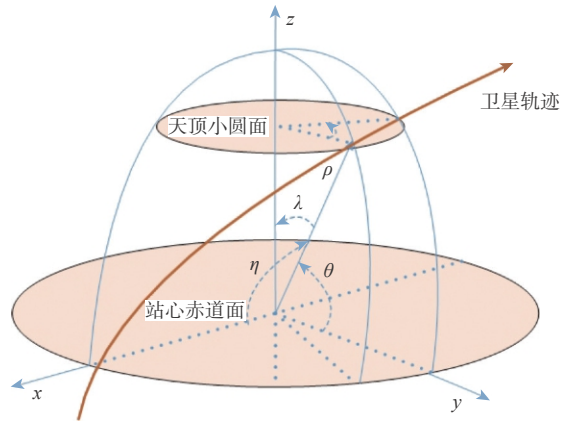


图2 新观测量的定义

Fig. 2 Definition of new angle

$$\begin{cases}\begin{bmatrix}\cos\delta\dfrac{\partial\alpha}{\partial\boldsymbol{r}}\\ \dfrac{\partial\delta}{\partial\boldsymbol{r}}\end{bmatrix}=\dfrac{1}{\rho}\begin{bmatrix}-\sin\alpha & \cos\alpha & 0\\ -\sin\delta\cos\alpha & -\sin\delta\sin\alpha & \cos\delta\end{bmatrix}\\ \boldsymbol{y}_i=\begin{bmatrix}(\alpha_i-\alpha_{\mathrm{o}})\cos\delta_i\\ \delta_i-\delta_{\mathrm{o}}\end{bmatrix}\end{cases}\tag{10}$$

式中：$\boldsymbol{r}$为当前时刻空间碎片的位置矢量；下标i表示计算值，下标o表示观测值。

在$\boldsymbol{B}$矩阵和赤经残差计算时，分别乘以赤纬的余弦值。从数学意义上来说，通过乘以$\cos\delta$，调整了空间碎片观测数据定轨中的权重，特别是降低了位于天顶方向附近的观测数据的权重。这样的处理方式避免了观测数据位于天顶附近时计算不连续的现象。

但是仍然存在如下问题：赤经值域为[−180°, 180°]，赤纬值域范围为[−90°,90°]，二者的值域不同。由于定义的原因，使得原本2个完全独立的观测量赤经和赤纬，在计算$\boldsymbol{B}$矩阵和赤经残差时，赤经相关的行列式数值$\left(\partial\alpha/\partial\boldsymbol{r}\right)\cos\delta$、$(\alpha_i-\alpha_{\mathrm{o}})\cos\delta_i$相比赤纬相关的行列式数值$\partial\delta/\partial\boldsymbol{r}$、$\delta_i-\delta_{\mathrm{o}}$，值域范围扩大了1倍，相当于增加了赤纬观测量的权重，从而影响最终定轨的收敛精度和效率。

2.3 角度转化法

为了解决赤经投影法中由于定义不同从而影响定轨的收敛精度和效率问题，本文提出角度转化法，基于权重平均的思想将观测量进行重新定义，将赤纬和和赤经确定的方向采用新的观测量η、θ和λ表示（见图2）。

η为测站-空间碎片连线与x轴方向的夹角，θ为测站-空间碎片连线与y轴方向的夹角，λ为测站-空间碎片连线与z轴方向的夹角。η、θ和λ的值域范围为[0°, 180°]。其中，赤经赤纬与新观测量的转化关系为

$$\begin{cases}\eta=\arccos(\cos\delta\cos\alpha)\\ \theta=\arccos(\cos\delta\sin\alpha)\\ \lambda=\arccos(\sin\delta)\end{cases}\tag{11}$$

采用新的观测量进行后续的精密定轨计算，可以避免观测数据位于天顶附近时，$\boldsymbol{B}$矩阵和赤经残差计算出现跳点的现象，同时3个角度值域范围相同，避免了角度值域不同对定轨结果的影响。3个新角度分别表征观测方向矢量与3个坐标轴之间的夹角，因此在精密定轨过程中认为3个观测量权重是相同的。

3 模型验证及分析

本文以SGP4（考虑带谐项J2、J3、J4和大气阻力摄动）/SDP4（考虑带谐项J2、J3、J4、大气阻力、田谐项共振效应和日月引力摄动）模型为预报器，采用观测数据生成TLE根数为例，验证上述3种观测量处理方法对定轨和预报结果的影响。具体方案如下：

1）采用精密轨道预报模型（21×21阶地球非球形摄动、光压摄动、大气阻力摄动、日月引力摄动）分别生成低轨道（low Earth orbit, LEO）、中轨道（medium Earth orbit, MEO）和大偏心率轨道（heigh eccentricity orbit, HEO）空间碎片的标称轨道，生成空间碎片相对地面测站的多段模拟观测数据（α, δ形式），并加入了正态分布的误差$N(0,9('')^2)$。

2）利用单弧段观测数据生成空间碎片的初始瞬时根数，将其转换成TLE平根数[16]，作为精密定轨的轨道根数初值。

3）分别以原始定义法（不对观测数据处理形式做改动）、赤经投影法及角度转化法3种方式，利用观测数据进行精密定轨，终止条件为残差变化率小于2%或者迭代次数大于20次。

4）分析比较3种处理方式的定轨残差和迭代

次数, 同时将精密定轨结果预报 1 000 min(1 min 1 个预报值)和标称轨道进行比较。

3 颗空间碎片的标称轨道根数如表 1 所示, 观测数据信息如表 2 所示。

表 1　3 颗空间碎片的标称轨道根数

Table 1　Nominal orbital elements of 3 debris

轨道	半长轴/km	偏心率	倾角/(°)	升交点赤经/(°)	近地点幅角/(°)	平近点角/(°)	面质比/($m^2 \cdot kg^{-1}$)
LEO	8 380.977	0.010 582	97.999	0.523	1.006	113.565	0.02
MEO	20 002.628	0.009 890	97.998	0.023	0.915	48.449	0.02
HEO	41 999.169	0.499 836	64.988	359.987	180.011	52.964	0.02

表 2　3 颗空间碎片产生的模拟观测数据

Table 2　Simulated observations from 3 debris

轨道	弧段序号	弧长/s	最低赤纬/(°)	最高赤纬/(°)
LEO	1	294.622	66.202	72.912
	2	250.162	63.409	70.409
	3	77.072	66.358	68.833
	4	46.490	64.477	64.818
MEO	1	1 623.089	43.099	71.181
	2	1 680	35.622	64.044
	3	2 825.676	65.381	89.626
	4	6 905	−3.112	89.626
	5	3 060	23.008	77.662
HEO	1	13 080	41.015	71.596
	2	8 400	31.787	53.963
	3	6 180	−17.460	−28.934
	4	2 340	−8.009	−11.759

精密定轨结果如表 3 所示。3 种不同的观测量处理方法, 残差计算和统计方式是不同的。为了统一进行比较, 在不影响各自精密定轨过程的情况下, 输出的最终定轨残差采用原始定义法进行统计。

从结果可以看出, 采用角度转化法, LEO、MEO 和 HEO 目标的定轨收敛精度最高, 同时收敛速度均小于或等于另外 2 种方法; 采用赤经投影法, 定轨收敛精度居中, 收敛速度和原始定义法相当; 原始定义法在收敛速度、收敛精度方面均较低。总体来说, 当观测数据位于测站天顶附近时(见表 3 中 LEO 和 MEO 算例), 角度转化法有较好的改进效果。针对观测数据一般分布情况(见表 3 中 HEO 算例), 角度转化法和另外 2 种方法效果相当。

表 3　精密定轨结果

Table 3　Results of precise orbit determination

轨道	观测数据处理形式	迭代次数	定轨残差
LEO	原始定义法	18	7.5×10^{-5}
	赤经投影法	18	5.0×10^{-5}
	角度转化法	14	4.0×10^{-5}
MEO	原始定义法	14	1.9×10^{-4}
	赤经投影法	15	1.4×10^{-5}
	角度转化法	10	1.0×10^{-5}
HEO	原始定义法	9	2.0×10^{-5}
	赤经投影法	9	2.0×10^{-5}
	角度转化法	9	2.0×10^{-5}

将 3 种精密定轨方法产生轨道进行点位预报, 1 min 1 个数据点, 共预报 1 000 min, 将预报值和标称轨道(作为标准)进行比较, 结果如图 3～图 5 所示。

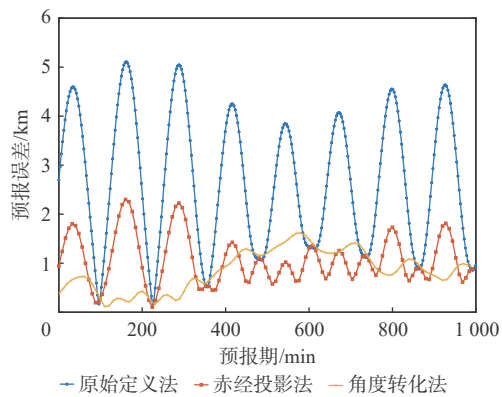


图 3　LEO 定轨结果预报误差

Fig. 3　Propagation errors of LEO orbiting results

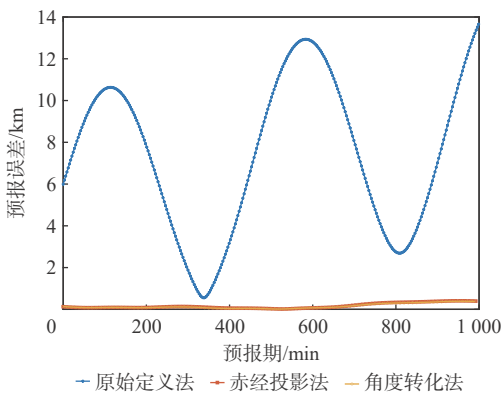


图 4　MEO 定轨结果预报误差

Fig. 4　Propagation errors of MEO orbiting results

1) LEO 轨道。原始定义法轨道平均预报误差为 2.59 km, 赤经投影法平均预报误差为 1.07 km, 角度转化法的平均预报误差为 0.85 km。

2) MEO 轨道。原始定义法轨道平均预报误差为 8.24 km, 赤经投影法平均预报误差为 0.25 km, 角度转化法平均预报误差为 0.23 km。

3) HEO 轨道。原始定义法轨道平均预报误差

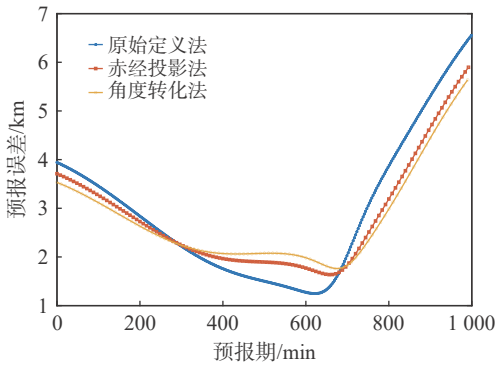


图5 HEO定轨结果预报误差

Fig. 5 Propagation errors of HEO orbiting results

为4.37 km, 赤经投影法平均预报误差为4.13 km, 角度转化法平均预报误差为4.02 km。

4 结 论

1）当观测数据较多位于测站天顶方向附近时, 相比较原始方法, 角度转化法改进效果明显, 收敛速度提高了25%左右, 收敛精度提高2～10倍, 预报精度也有明显的提升。相比较赤经投影法, 角度转化法收敛速度也提升了25%左右, 收敛精度和预报精度相当。

2）当监测数据不集中于天顶方向附近时, 3种精密定轨方法效果相当。

对于常规空间碎片编目, 每日获取数万至数十万观测数据, 进行数万次精密定轨工作, 本文研究可以提高定轨精度和加快收敛速度, 快速准确地掌握空间态势, 对于中国空间态势感知系统设计与研制具有重要参考意义。

参考文献（References）

[1] BRAUN V, LEMMENS S, REIHS B, et al. Analysis of breakup events[C]//Proceedings of the 7th European Conference on Space Debris, 2017: 1-12.

[2] PIROVANO L, PRINCIPE G, ARMELLIN R. Data association and uncertainty pruning for tracks determined on short arcs[J]. Celestial Mechanics and Dynamical Astronomy, 2020, 132: 6-30.

[3] CICALO S, BECK J, MINISCI E, et al. GOCE radar-based orbit determination for re-entry predictions and comparison with GPS-based POD[C]//Proceedings of the 7th European Conference on Space Debris, 2017: 1-14.

[4] 刘林, 张强, 廖新浩. 人卫精密定轨中的算法问题[J]. 中国科学(A辑), 1998, 28(5): 848-856.
LIU L, ZHANG Q, LIAO X H. Algorithm problems in precise orbit determination of artificial earth satellite[J]. Science in China (Series A), 1998, 28(5): 848-856(in Chinese).

[5] 程昊文. 航天器轨道理论在空间目标编目管理中的应用[D]. 南京: 南京大学, 2012: 9-11.
CHENG H W. The application of satellite orbit theory in maintaining the space object catalog[D]. Nanjing: Nanjing University, 2012: 9-11(in Chinese).

[6] MARSHALL J A, LUTHCKE S B. Modeling radiation forces acting on TOPEX/Poseidon for precision orbit determination[J]. Journal of Spacecraft and Rockets, 1994, 31(1): 99-105.

[7] BOCK H, HUGENTOBLER U, SPRINGER T A, et al. Efficient precise orbit determination of LEO satellites using GPS[J]. Advances in Space Research, 2002, 30(2): 295-300.

[8] 张宇, 孔静, 陈明, 等. CE5T拓展试验轨道精度分析[J]. 宇航学报, 2019, 40(9): 1014-1023.
ZHANG Y, KONG J, CHEN M, et al. Orbit accuracy analysis for CE5T extended mission[J]. Journal of Astronautics, 2019, 40(9): 1014-1023(in Chinese).

[9] 张宇, 周立, 孔静, 等. 长期姿控扰动情况下空间实验室轨道影响分析及建模[J]. 宇航学报, 2017, 38(12): 1273-1280.
ZHANG Y, ZHOU L, KONG J, et al. Analysis and modeling of space laboratory orbit under continuous attitude control perturbation[J]. Journal of Astronautics, 2017, 38(12): 1273-1280(in Chinese).

[10] Committee for the Assessment of the U. S. Air Force's Astrodynamics Standards. Continuing Kepler's quest: Assessing air force space command's astrodynamics standards[S]. Washington, D. C. : The National Academies Press, 2012: 21-26.

[11] GRZEGORZ B, KRZYSZTOF S, RADOSŁAW Z, et al. Determination of precise Galileo orbits using combined GNSS and SLR observations[J]. GPS Solution, 2021, 25: 11.

[12] VALLADO D A, FINKLEMAN D. A critical assessment of satellite drag and atmospheric density modeling: AIAA 2008-6442[R]. Reston: AIAA, 2008.

[13] 韦栋, 赵长印. SGP4/SDP4模型精度分析[J]. 天文学报, 2009, 50(3): 332-339.
WEI D, ZHAO C Y. Analysis on the accuracy of the SGP4/SDP4 model[J]. Acta Astronomica Sinica, 2009, 50(3): 332-339(in Chinese).

[14] VALLADO D A. Fundamental of astrodynamics and application [M]. Hawthorne: Microcosm Press, 2007: 113-115.

[15] 刘林, 汤靖师. 卫星轨道理论与应用[M]. 北京: 电子工业出版社, 2015: 427-428.
LIU L, TANG J S. Satellite orbit theory and application[M]. Beijing: Publishing House of Electronics Industry, 2015: 427-428(in Chinese).

[16] 陈磊, 韩蕾, 百显宗, 等. 空间目标轨道力学与误差分析[M]. 北京: 国防工业出版社, 2010: 93-95.
CHEN L, HAN L, BAI X Z, et al. Orbit dynamics and error analysis of space object[M]. Beijing: National Defense Industry Press, 2010: 93-95(in Chinese).

Precise orbit determination method for angle-only observation data of space debris based on angle conversion theory

ZHANG Yao[1, 2, 3], LIU Jing[1, 2, 3, *]

(1. National Astronomical Observatory, Chinese Academy of Science, Beijing 100101, China;
2. School of Astronomy and Space Science, University of Chinese Academy of Science, Beijing 100049, China;
3. Laboratory of Pinghu, Jiaxing 314200, China)

Abstract: On the basis of the angle conversion method, a brand-new accurate orbit determination technique for space debris was developed. First, based on the commonly used two-line-element (TLE) data application requirements, a TLE precise orbit determination model based on the SGP4/SDP4 forecast model was established according to the differential idea. Secondly, the advantages and disadvantages of the original definition method and the right ascension projection method were discussed by taking right ascension-declination observation as an example. Then the angle conversion method was proposed, which converts the two-element angle data into three-element angle data. Finally verifies based on the simulation that when the observation data is concentrated in the direction of the zenith of the station, the angle conversion method could increase the speed of precise orbit determination by about 25%, and increase the accuracy of orbit determination by 2−10 times.

Keywords: precise orbit determination; space debris; angle-only observation data; SGP4/SDP4; two-line-element

Received: 2021-09-01; **Accepted**: 2022-01-25; **Published Online**: 2022-02-23 14:53
URL: kns.cnki.net/kcms/detail/11.2625.V.20220223.0859.002.html
Foundation items: National Natural Science Foundation of China (12073045,11803052)
* **Corresponding author**. E-mail: liujing@bao.ac.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0514

真空强电磁场环境下铝的二次电子倍增规律

商圣飞[1,2,*]，杨晓宁[1,2]，杨勇[1,2]，毕研强[1]，武南开[1]，于澜涛[1]

(1. 北京卫星环境工程研究所，北京 100094；2. 可靠性与环境工程技术重点实验室，北京 100094)

摘 要：针对卫星表面受强电磁环境的影响导致的充放电问题，采用1D3V的粒子网格（PIC）方法对卫星表面铝材料在空间强电磁环境作用下的二次电子倍增作用规律进行研究。结果表明：星表铝材料在不同微波幅值、不同频率下的二次电子倍增效应存在“最易”倍增区间；二次电子倍增规律表现为在特定频率下，铝的二次电子倍增随着微波电场幅值的增大先增强后降低，表现出最佳倍增区间的效应；在特定幅值下，铝的二次电子倍增效应也会先增强后降低，但是整体表现出低频时倍增强，高频时抑制倍增的效应。

关 键 词：高功率微波；空间电磁环境；星表材料；二次电子倍增；放电

中图分类号：O461；O462

文献标志码：A **文章编号**：1001-5965(2023)07-1606-08

地球静止轨道卫星在轨运行期间，将面临严酷的表面带电和内带电环境[1-2]，当充电满足一定条件之后将发生静电放电，进而产生电磁脉冲信号，这些信号耦合进入在轨卫星的电子产品，将对卫星在轨稳定运行构成威胁[3]。国内卫星在轨异常统计结果显示，在轨期间发生的可恢复性异常几乎都与空间环境相关，这其中相当比例的在轨可恢复性异常是由空间电磁干扰引起，这些电磁干扰主要来源于卫星空间充放电效应[4-5]。以多颗地球静止轨道通信卫星3年内在轨发生的异常统计结果为例，约70%的异常与空间环境有关[6]。

卫星在轨运行期间会受到强电磁环境的影响，强电磁环境又可分成自然强电磁环境和人为强电磁环境。星表产品在强电磁环境，特别是高功率微波 (high-power microwave, HPM)[7-8] 作用下可能会产生二次电子倍增过程，严重的可导致产品毁伤。高功率微波通常是指峰值功率大于100 MW，频率在1～300 GHz的电磁波。目前关于微波作用下的二次电子倍增导致的沿面闪络现象已在微波输出窗有了相关的研究。当微波功率超过一定值后，容易在真空-介质[9-12]或空气-介质[13-14]界面产生击穿。国外对微波输出窗的沿面闪络的研究相对比较多，试验方面通过光学诊断方法得到了输出窗沿面放电的物理图像[15]。文献[16]采用蒙卡方法开展了二次电子倍增的敏感区间的研究，文献[17-18]最先采用1D3V的粒子网格（particle-in-cell, PIC）方法对于沿面闪络的机理进行了模拟。文献[19-22]利用1D3V静电模型PIC-蒙特卡罗模型（Monte Carlo model, MCC）粒子模拟程序研究了介质表面次级电子倍增过程，不同气压下介质界面击穿过程等。先前的研究者主要针对微波介质窗的沿面闪络进行研究，而对于空间星表材料在不同频率、不同幅值的微波作用下二次电子倍增规律少有系统的研究。

本文主要研究星表铝材料在不同电磁频率、不同微波幅值下的二次电子倍增效应的规律，得到铝在不同微波作用下的二次电子倍增演化规律特征、最易放电区间等，为卫星空间防护提供基础支持。

1 理论模型

本文采用1D3V的PIC方法研究材料在微波作

收稿日期：2021-09-05；录用日期：2021-11-22；网络出版时间：2022-01-25 08:40
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220124.1735.005.html
*通信作者. E-mail：shangshengfei@buaa.edu.cn

引用格式：商圣飞，杨晓宁，杨勇，等. 真空强电磁场环境下铝的二次电子倍增规律[J]. 北京航空航天大学学报，2023，49(7)：1606-1613.
SHANG S F，YANG X N，YANG Y，et al. Secondary electron multiplication of aluminum under strong vacuum electromagnetic field[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1606-1613 (in Chinese).

用下的二次电子倍增过程。一般情况下, 电子在强电磁波作用下有 2 种方式实现倍增: ①电子受到材料表面束缚电场作用(有电子发射导致的正电荷沉积形成的), 电子会返回轰击材料, 电子在返回过程中受到电磁波的电场加速使得电子以一定能量轰击材料, 造成二次电子发射; ②电子从微波电场中获得能量与背景气体发生电离碰撞, 形成新的电子和离子。

为了简化仿真过程, 本文采用以下 2 个假设:

1) 卫星长时间工作在真空环境中, 忽略材料释气;

2) 不考虑空间电子运动产生的电磁场对微波入射的反馈机制, 也就是说, PIC 方法解算电场时采用泊松方程, 而不是全麦克斯韦方程组。

基于这 2 个假设, 本文的电子增值过程只考虑电子轰击材料造成的二次电子发射过程。

1.1　粒子模拟动力学方程

一般情况下描述带电粒子在电磁场中的方法为全麦克斯韦方程组[23]:

$$\begin{cases}\nabla\times\boldsymbol{E}=-\dfrac{\partial\boldsymbol{B}}{\partial t}\\ \nabla\times\boldsymbol{B}=\varepsilon_0\mu_0\dfrac{\partial\boldsymbol{E}}{\partial t}+\mu_0\boldsymbol{J}\\ \nabla\cdot\boldsymbol{E}=\dfrac{\rho_q}{\varepsilon_0}\\ \nabla\cdot\boldsymbol{B}=0\end{cases}\tag{1}$$

式中: ∇为 Nabla 算子; $\boldsymbol{E}$ 为电场; $\boldsymbol{B}$ 为磁场强度; t 为时间; ε_0为真空介电常数; μ_0为磁导率; $\boldsymbol{J}$ 为传导电流密度; ρ_q为自由电荷密度。根据欧姆定律和洛伦兹力公式可得到电荷在电磁场中的运动为

$$\boldsymbol{J}=\sigma(\boldsymbol{E}+\boldsymbol{u}\times\boldsymbol{B})+\rho_q\boldsymbol{u}\tag{2}$$

$$\boldsymbol{F}=\rho_q\boldsymbol{E}+\boldsymbol{J}\times\boldsymbol{B}\tag{3}$$

式中: σ为电导率; $\boldsymbol{u}$为粒子速度; $\boldsymbol{F}$为粒子所受的力。式(3)可采用速度增量方式来表示:

$$\begin{cases}m\gamma\dfrac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t}=\boldsymbol{f}\\ \boldsymbol{u}=\dfrac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t}\end{cases}\tag{4}$$

式中: m 为粒子质量; γ为相对论因子; $\boldsymbol{x}$ 为粒子位置。则动力学方程可以表示为

$$\frac{\mathrm{d}(m\gamma\boldsymbol{u})}{\mathrm{d}t}=\boldsymbol{f}=q(\boldsymbol{E}_{\mathrm{mw}}+\boldsymbol{E}_{\mathrm{dc}}+\boldsymbol{u}\times\boldsymbol{B}_{\mathrm{mw}})\tag{5}$$

式中: 下标 mw 表示微波输入相关参数; $\boldsymbol{E}_{\mathrm{dc}}$为空间电荷产生的电场。

高功率微波的描述为

$$\begin{cases}\boldsymbol{E}_{\mathrm{mw}}=E_{\mathrm{mw0}}\sin(\omega t+\theta)\hat{\boldsymbol{x}}\\ \boldsymbol{B}_{\mathrm{mw}}=-(\mu_0E_{\mathrm{mw0}}/\eta)\sin(\omega t+\theta)\hat{\boldsymbol{y}}\end{cases}\tag{6}$$

式中: $\omega=2\pi f_{\mathrm{w}}$, f_{w}为微波频率; θ为初始相位; η为波阻抗; E_{mw0}为微波电场幅值; $\hat{\boldsymbol{x}}$和$\hat{\boldsymbol{y}}$分别为 x 和 y 方向的单位矢量。

在实际处理中, 一般为了简化计算, 常忽略感应电磁场的影响, 因此, 麦克斯韦方程组就可以用解算静电场的泊松方程计算, 此时空间电场 $\boldsymbol{E}_{\mathrm{dc}}$ 可以通过式(7)计算得到:

$$\begin{cases}\boldsymbol{E}_{\mathrm{dc}}=-\nabla\phi\\ \nabla^2\phi=-\dfrac{\rho_q}{\varepsilon_0}\end{cases}\tag{7}$$

式中: ϕ为静电势。则最后方程组为

$$\begin{cases}\dfrac{\mathrm{d}(m\gamma\boldsymbol{u})}{\mathrm{d}t}=\boldsymbol{f}=q(\boldsymbol{E}_{\mathrm{mw}}+\boldsymbol{E}_{\mathrm{dc}}+\boldsymbol{u}\times\boldsymbol{B}_{\mathrm{mw}})\\ \boldsymbol{E}_{\mathrm{mw}}=E_{\mathrm{mw0}}\sin(\omega t+\theta)\hat{\boldsymbol{x}}\\ \boldsymbol{B}_{\mathrm{mw}}=-(\mu_0E_{\mathrm{mw0}}/\eta)\sin(\omega t+\theta)\hat{\boldsymbol{y}}\\ \boldsymbol{E}_{\mathrm{dc}}=-\nabla\phi\\ \nabla^2\phi=-\dfrac{\rho_q}{\varepsilon_0}\end{cases}\tag{8}$$

1.2　二次电子倍增模型

二次电子倍增的理论当前普遍采用文献 [24] 中提出的描述二次电子发射材料特性的经验公式:

$$\sigma(\xi,\alpha)=\sigma_{\max}(\alpha)f(\xi)\tag{9}$$

式中:

$$f(\xi)=\begin{cases}[\xi\exp(1-\xi)]^{0.56} & \xi\leqslant 1\\ [\xi\exp(1-\xi)]^{0.25} & 1<\xi\leqslant 3.6\\ 1.125\xi^{-0.35} & \xi>3.6\end{cases}\tag{10}$$

$$\begin{cases}\sigma_{\max}(\alpha)=\sigma_{\max0}(1+k_{\mathrm{s}}\alpha^2/2\pi)\\ E_{\max}(\alpha)=E_{\max0}(1+k_{\mathrm{s}}\alpha^2/2\pi)\end{cases}\tag{11}$$

$$\xi=\frac{E_{\mathrm{i}}-E_{\mathrm{th}}}{E_{\max}-E_{\mathrm{th}}}\tag{12}$$

式中: σ为二次电子发射率; $\sigma_{\max0}$为正入射二次电子最大发射率; $E_{\max0}$为其对应入射能量; $\sigma_{\max}(\alpha)$为斜入射二次电子最大发射率; $E_{\max}(\alpha)$为其对应入射能量; α为入射角($\alpha=0$正入射); k_{s}为表面光滑系数($k_{\mathrm{s}}=0,1,2$分别代表粗糙、不光滑、光滑); E_{i}为碰撞能量; E_{th}为阈值; $E_{\mathrm{i}}<E_{\mathrm{th}}$不产生二次电子, 一般 $E_{\mathrm{th}}=12.5$ eV。$\sigma>1$区间($E_1<E_{\mathrm{i}}<E_2$)代表每个种子电子产生二次电子数目大于 1, 对应碰撞能量分布范围即为电子倍增区间。图 1 为典型二次电子发射系数曲线[25], 可知, 二次电子发射系数大于 1 的区间为入射电子能量在 E_1 与 E_2 之间。小于 E_1, 或者大于 E_2 均不利于二次电子的倍增。

1.3　理论模型验证

本文计算的卫星空间环境的二次电子倍增的条件为真空工况, 为验证理论模型, 选取文献 [22] 中的真空算例作为对比算例。计算条件为输入微波:

E_{mw0}=2.82 MV/m, f_w=2.85 GHz; 介质表面积为 10 cm^2。二次电子倍增参数为: E_{max0}=420 eV, σ_{max0} = 3, k_s=1。计算结果如图 2 所示, 可知, 电子在 2×10^{10}～1.7×10^{11} 个之间振荡, 振荡频率为微波频率的 2 倍, 本文计算结果与文献 [22] 中的结果吻合较好。

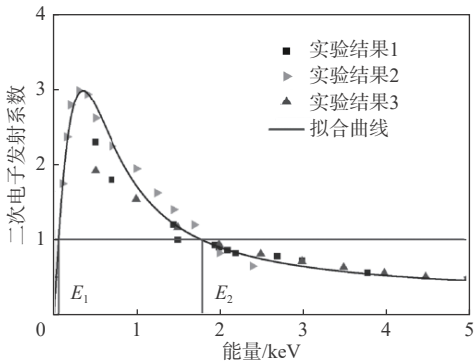


图 1 典型二次电子发射系数曲线[25]

Fig. 1 Typical secondary electron emission yield curves[25]

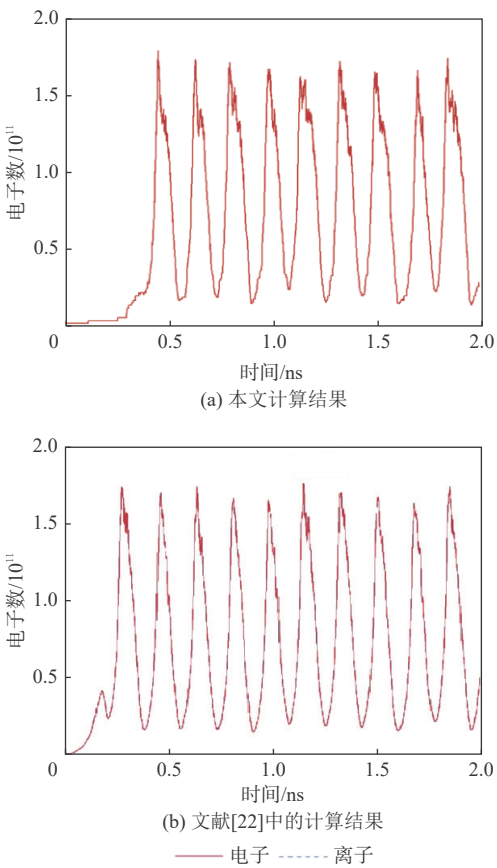


图 2 电子数在微波作用下的演化理论模型校验

Fig. 2 Verification of theoretical model for evolution of electron number evolution under presence of microwaves

2 计算输入条件

星表材料选择铝, 在微波作用下的模型及空间坐标如图 3 所示。在模型中假定: 高功率微波沿 z 方向以平面波形式传输, 电场沿 x 方向, 磁场沿 y 方向。计算网格为均匀一维网格, 尺寸长度为 1×10^{-5} m, 时间步长为 2×10^{-13} s。计算铝圆盘尺寸为直径 10 cm, 其二次电子倍增的参数参考文献 [12] 中给出的铝的参数设置: E_{max0}=350 eV, σ_{max0} = 3.5, k_s=1。

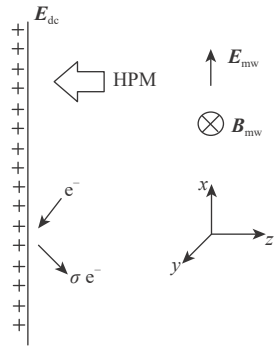


图 3 微波作用铝的模型示意图

Fig. 3 Schematic diagram of model under microwave action on aluminum

3 仿真结果及分析

3.1 不同微波电场幅值下的电子倍增规律

计算微波频率为 2 GHz 时, 不同微波电场幅值为 0.1, 0.2, 0.3, 0.5, 1, 10, 100, 500, 1 000 MV/m, 兼顾低微波场强时的衰减效应, 计算时种子电子统一设置为 1×10^{11} 个。

采用上述计算条件得到的不同微波电场幅值下空间电子数在 2 ns 内的变化规律如图 4 所示。当微波电场幅值小于 0.2 MV/m 时, 初始设定的 10^{11} 个电子会逐渐衰减消失, 此时电子碰壁后基本不会产生二次电子倍增的过程, 而是被壁面吸收。不同微波电场幅值下碰壁电子能量变化如图 5 所示, 图 5(a)为 0.1 MV/m 微波电场幅值下的碰壁电子能量。可见平均碰壁电子的能量比较低, 基本在 10 eV 以下, 小于二次电子倍增所需的 12.5 eV 的阈值, 故二次电子基本不会倍增。

当微波电场幅值达到 0.3 MV/m 时, 空间电子数呈现低水平的振荡, 电子数可以维持, 此时平均碰壁电子能量随微波的振荡呈现周期振荡特性, 能量可达几十 eV 量级如图 5(b)所示, 此时星表充放电概率将会大大增加。当大于 0.5 MV/m 一直到 10 MV/m(见图 4(d)～图 4(f))时空间电子数会维持高水平的振荡, 以 10 MV/m 工况为例, 此时平均碰壁电子能量主要在 100～2 000 eV 之间振荡, 此时星表充放电严重。

当微波电场幅值继续增高, 达到 100 MV/m 时, 此时空间电子数又出现低幅值的振荡(见图 4(g)), 平均碰壁电子能量均比较高, 一般在 1 000 eV 以上, 如图 5(d)所示。此时电子振荡只有在很短的时

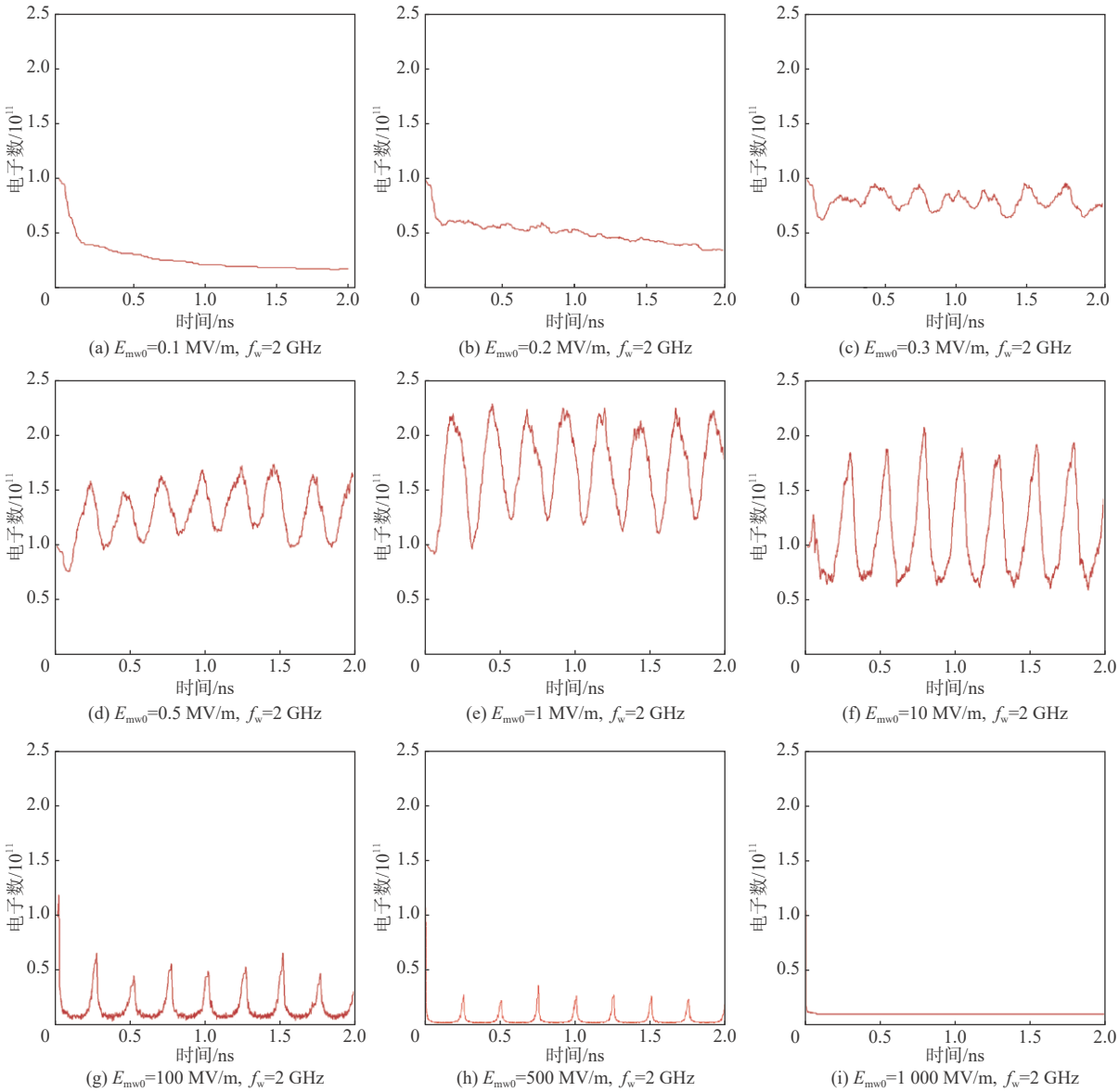


图 4　不同微波电场幅值下的空间电子数变化

Fig. 4　Variation of electron number under different amplitude of microwave electric field

间内有相对低能的电子碰壁, 产生倍增, 而高能的电子由于能量过高, 接近二次电子发射系数大于 1 的上边界 E_2, 其二次电子倍增的效率反而偏低。

当微波电场幅值达到 500 MV/m 时, 空间电子仅仅在微波周期的某时间段内有峰值, 且整体电子数量偏低, 此时碰壁电子平均能量呈现出规则的振荡, 能量一般在 100～20 000 eV 之间。当微波电场幅值到 1 000 MV/m 时, 空间电子数很快衰减为 0, 这主要是能量过高, 此时高能的电子入射到材料, 由于注入深度太深导致背散射的电子不容易到达表面形成二次电子, 二次电子不会倍增。

由此可见, 材料二次电子倍增与空间强电磁波的幅值并非单调关系, 而存在最易增值区间。一般对于铝材料而言, 空间二次电子可实现振荡的最小电磁波电场幅值约为 0.3 MV/m, 而上限约为 500 MV/m。因此, 空间星表材料充放电的危险微波电场幅值区间在 0.3～500 MV/m 之间。

3.2　不同微波频率下的电子数倍增规律

计算的 8 种工况的微波频率为 1 , 2 , 3 , 5 , 7 , 10, 30 , 50 GHz, 微波电场幅值设为 1 MV/m。其中 1～7 GHz 初始种子电子数设为 5×10^9 个, 而 10～50 GHz 的工况为了观察电子数的衰减情况, 种子电子数设置为 100×10^{11} 个。

图 6 为不同微波频率下的空间电子数变化, 图 6(a)为 8 种工况下的平均电子数, 其中 1～7 GHz 的平均电子数计算时采用电子稳定振荡后的数据平均得到, 而图 6(b)中 1 GHz 的平均电子数则是 1.1 ns 后的数据平均得到。10～50 GHz 时由于电子是衰减的, “平均电子数”则是直接采用 2 ns 时的电子数代替, 如图 6(c)中 10 GHz 的“平均电子数”选取 2 ns 时刻的值。由图 6(a)可知, 在微波频率小于 7 GHz 的相对低频区域平均二次电子数呈现出先升

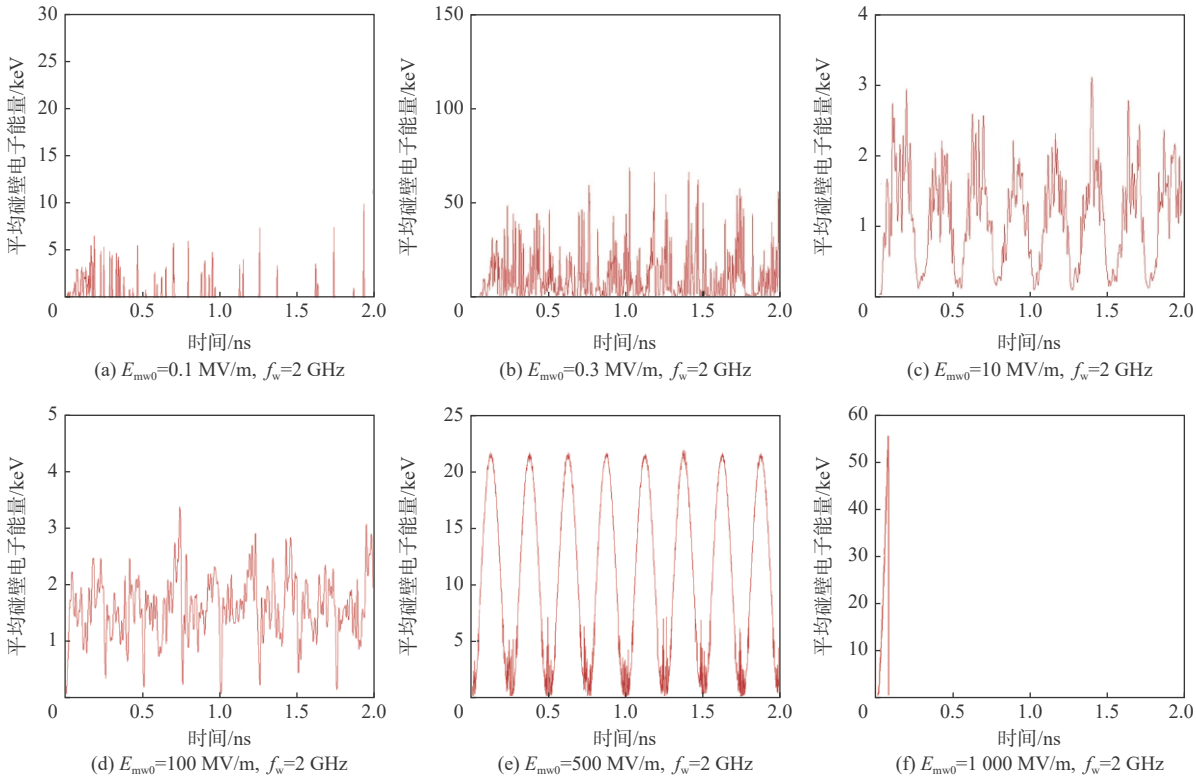


图5 不同微波电场幅值下的碰壁电子能量变化

Fig. 5 Variation of wall-impingerment electron energy under different microwave electric field amplitude

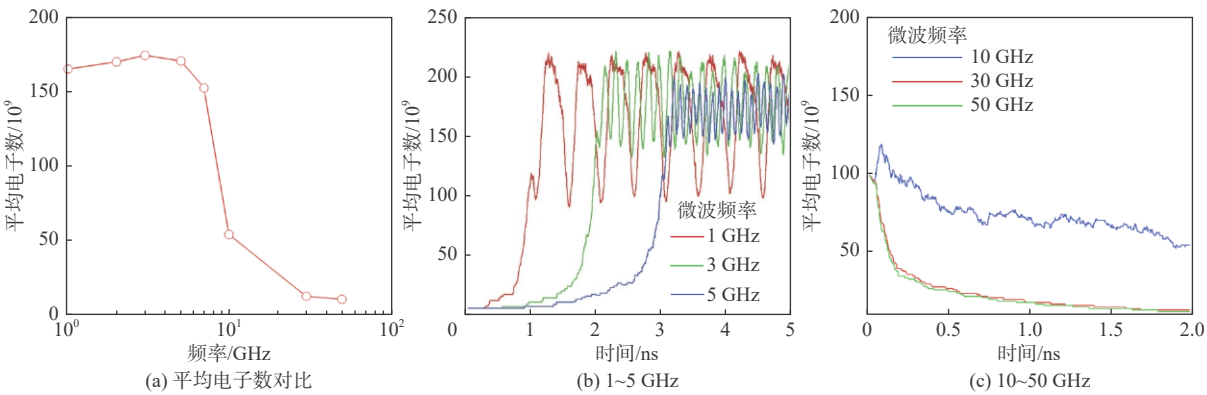


图6 不同微波频率下的空间电子数变化

Fig. 6 Variation of electron number at different microwave frequencies

高再降低的特征，但整体差别不大，在3 GHz左右达到峰值，随后随着频率的增大空间电子数迅速降低，直至发生电子消散。因此，空间电子数较高的微波频率范围为小于10 GHz的频率区间，大于10 GHz则呈现出抑制二次电子倍增的作用。

图6(b)为微波频率为1，3，5 GHz的空间电子数随时间的变化规律。以微波频率为1 GHz为例，图7为空间电子数振荡频率与微波频率的关系，可知空间电子数振荡频率是微波频率的2倍。当频率为1 GHz时，由于微波频率偏低，电子在“低幅值”空间时间较长，导致电子碰壁衰减较多、倍增少。在一定范围内(小于3 GHz)，随着频率增加，电子碰壁频率增加，导致电子在“低幅值”空间时间变短，使得电子在较高幅值上振荡。当超过3 GHz范围，虽然空间电子碰壁频率更大，但是空间电子数却在下降，这主要是频率过高时电子在空间电场的加速时间变短，导致单位电子能量变小，使得碰壁电子的二次电子发射系数下降，直至小于1，产生如图6(c)所示的电子衰减现象。

不同微波频率下碰壁电子能量变化如图8所示。可知，在同一微波电场幅值(1 MV/m)情况下，碰壁电子能量随微波的频率增加整体呈下降趋势。当在相对较低的频段，平均电子能量落在二次电子系数大于1的区间，可以产生二次电子倍增效应，但当频率增大到一定程度后，由于频率过大，导致电子在空间电场加速的时间变短，能量较小如

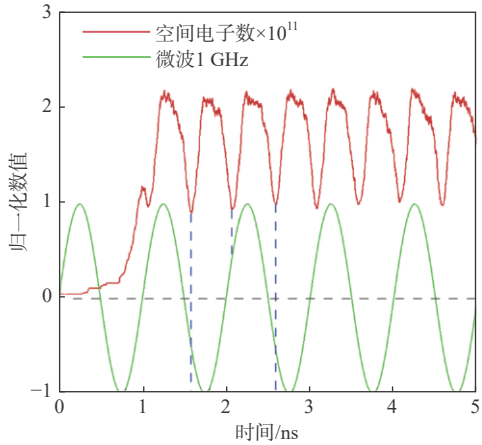


图 7　1 GHz 微波频率与空间电子数振荡频率的关系

Fig. 7　Relationship between 1 GHz microwave frequency and electron number oscillation

图 8(e)所示, 30 GHz 时平均电子能量一般在 10 eV 以下, 此时电子能量过低, 大部分电子碰壁后会被吸收, 导致电子数量逐渐衰减。

通过上述分析可见, 星表铝材料在不同微波电场幅值、不同频率下的二次电子倍增效应存在“最易”倍增区间。并且二次电子倍增规律为: ①在特定频率下, 铝的二次电子倍增随着微波电场幅值的增大先增强后降低, 表现出最佳倍增区间的效应; ②在特定幅值下, 铝的二次电子倍增效应也会随着微波频率先增强后降低, 但是整体表现出低频时倍增强, 高频抑制倍增的效应。对于典型频率如 2 GHz 条件下, 星表铝材料在微波电场幅值为 0.3～500 MV/m 之间时最易放电。在典型微波电场幅值如 1 MV/m 条件下, 星表铝材料在低频段约 1～7 GHz 之间二次电子倍增效应显著, 高频区(约大于 10 GHz)则显示出抑制二次电子倍增的效应, 并且放电最佳频率约为 3 GHz。

因此, 在需要二次电子倍增的领域, 应选择相匹配的电磁波的频率和幅值才能达到最佳的倍增效果。同样在空间强电磁环境防护方面, 应该针对特定的频率和幅值特征, 了解空间电子在微波作用下的变化规律, 有针对性的开展防护, 而不是采用最高、最低频率, 最高、最低场强等极限设计思维。

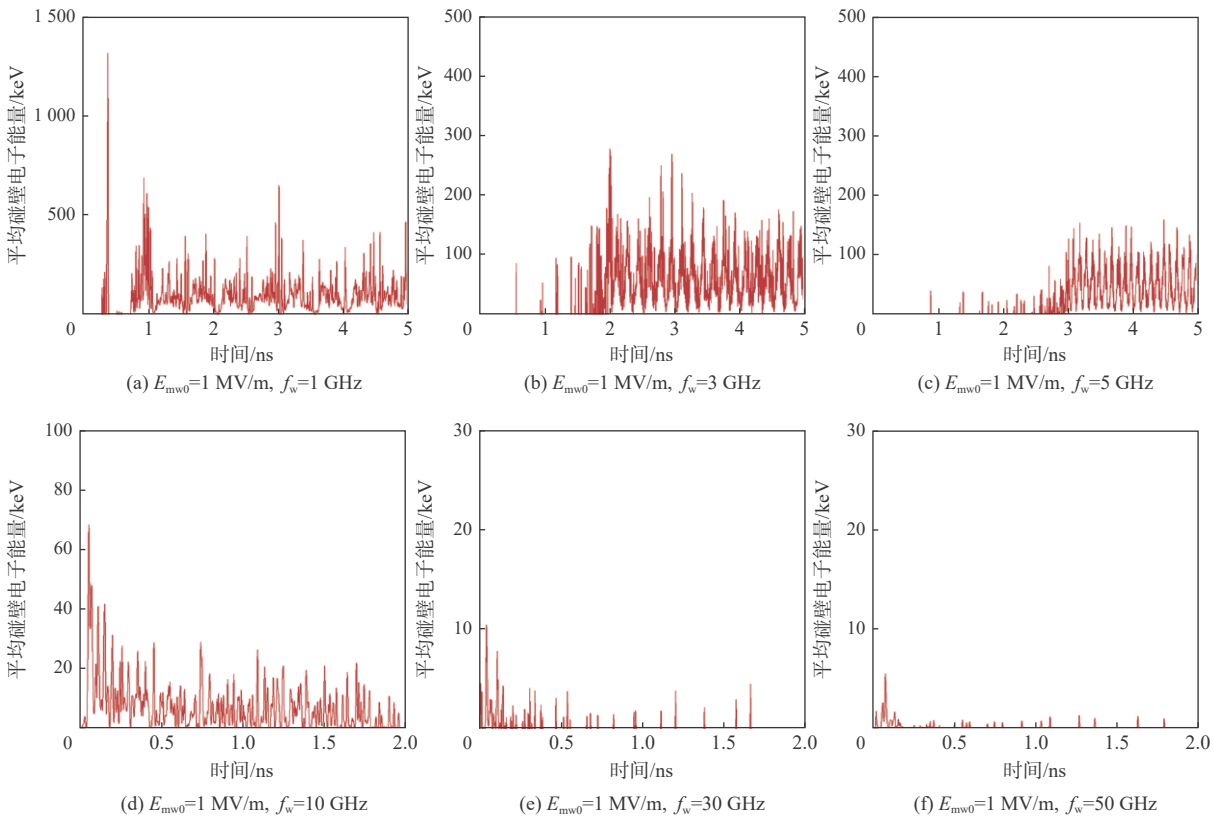


图 8　不同微波频率下碰壁电子能量变化

Fig. 8　Variation of electron energy at different microwave frequencies

4　结　论

1) 在特定频率下, 铝的二次电子倍增随着微波电场幅值的增大先增强后降低, 表现出最佳倍增区间的效应。

2) 在特定幅值下, 铝的二次电子倍增效应也会先增强后降低, 但是整体表现出低频强, 高频抑制倍增的效应。因此, 在需要二次电子倍增的领域,

应选择电磁波的频率和幅值的匹配才能达到最佳的倍增效果。同样在空间强电磁环境防护方面，应该针对特定的频率和幅值特征，了解空间电子在微波作用下的变化规律，有针对性的开展防护，而不是采用最高、最低频率，最高、最低场强等极限设计思维。

参考文献（References）

[1] 范费彬, 谢锦林, 陆全明, 等. 空间等离子体磁场重联过程地面实验装置及实验研究概述[J]. 航天器环境工程, 2019, 36(6): 655-661.
FAN F B, XIE J L, LU Q M, et al. Ground-based experimental study of magnetic reconnection in space plasma environment[J]. Spacecraft Environment Engineering, 2019, 36(6): 655-661 (in Chinese).

[2] 盛雪莲, 吴静, 张翀. 中国空间磁层线辐射现象探究[J]. 北京航空航天大学学报, 2018, 44(7): 1504-1513.
SHENG X L, WU J, ZHANG C. Space magnetospheric line radiation above China[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(7): 1504-1513(in Chinese).

[3] 唐萍, 朱光武, 秦国泰, 等. 航天器表面污染物质沉积变化和控制因子评估[J]. 北京航空航天大学学报, 2015, 41(5): 891-896.
TANG P, ZHU G W, QIN G T, et al. Changes of contamination deposition on spacecraft surface and evaluation of control factors[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(5): 891-896(in Chinese).

[4] 周建涛, 蔡伟, 武延鹏, 等. 星敏感器空间辐射效应研究[J]. 宇航学报, 2010, 31(1): 24-30.
ZHOU J T, CAI W, WU Y P, et al. Research on radiation effects of star sensors[J]. Journal of Astronautics, 2010, 31(1): 24-30(in Chinese).

[5] 韩建伟, 陈睿, 李宏伟, 等. 单粒子效应及充放电效应诱发航天器故障的甄别与机理探讨[J]. 航天器环境工程, 2021, 38(3): 344-350.
HAN J W, CHEN R, LI H W, et al. The anomalies supposed to be due to the single event effects may be caused by spacecraft charging induced electrostatic discharge[J]. Spacecraft Environment Engineering, 2021, 38(3): 344-350 (in Chinese).

[6] 陈卓, 杨晓宁, 杨勇. GEO卫星空间自然强电磁环境与防护综述[J]. 环境技术, 2019, 37(2): 68-72.
CHEN Z, YANG X N, YANG Y. Overview of GEO satellite under natural strong electromagnetic environment and protection methods[J]. Environmental Technology, 2019, 37(2): 68-72(in Chinese).

[7] BENFORD J. Space applications of high power microwaves[C]// 2007 IEEE 34th International Conference on Plasma Science (ICOPS). Piscataway: IEEE Press, 2007: 257.

[8] 周传明, 刘国治, 刘永贵. 高功率微波源[M]. 北京: 原子能出版社, 2007.
ZHOU C M, LIU G Z, LIU Y G. High power microwave source[M]. Beijing: Atomic Press, 2007 (in Chinese).

[9] 秋实. 高功率微波窗口击穿及馈源技术[D]. 西安: 西安电子科技大学, 2010.
QIU S. Study of dielectric window breakdown and feed technqies for high power microwave systems[D]. Xi'an: Xidian University, 2010 (in Chinese).

[10] CHANG C, LIU G Z, TANG C X, et al. Review of recent theories and experiments for improving high-power microwave window breakdown thresholds[J]. Physics of Plasmas, 2011, 18(5): 055702.

[11] QIU S, HAO X W, ZHANG G J, et al. Tree-like breakdown phenomena of dielectric window under X-band high power microwave in vacuum[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2010, 17(3): 971-977.

[12] 程国新. 高功率微波输出窗真空表面闪络研究[D]. 长沙: 国防科学技术大学, 2013.
CHENG G X. Study on the vacuum surface flashover of high-power microwave window[D]. Changsha: National University of Defense Technology, 2013 (in Chinese).

[13] FOSTER J, BEESON S, THOMAS M, et al. Rapid formation of dielectric surface flashover due to pulsed high power microwave excitation[J]. IEEE Transactions on Dielectrics & Electrical Insulation, 2011, 18(4): 964-970.

[14] STEPHENS J, BEESON S, DICKENS J, et al. Charged electret deposition for the manipulation of high power microwave flashover delay times[J]. Physics of Plasmas, 2012, 19(11): 112111.

[15] NEUBER A A, EDMISTON G F, KRILE J T, et al. Interface breakdown during high-power microwave transmission[J]. IEEE Transactions on Magnetics, 2007, 43(1 Pt. 2): 496-500.

[16] ZHANG P, LAU Y Y, FRANZI M, et al. Multipactor susceptibility on a dielectric with a bias dc electric field and a background gas[J]. Physics of Plasmas, 2011, 18(5): 053508.

[17] KIM H C, VERBONCOEUR J P. Time-dependent physics of a single-surface multipactor discharge[J]. Physics of Plasmas, 2005, 12(12): 123504.

[18] KIM H C, VERBONCOEUR J P, LAU Y Y. Invited paper-modeling RF window breakdown: From vacuum multipactor to RF plasma[J]. IEEE Transactions on Dielectrics and Electrical Insulation, 2007, 14(4): 774-782.

[19] 蔡利兵, 王建国, 朱湘琴. 不同气压下介质表面高功率微波击穿的数值模拟[J]. 强激光与粒子束, 2010, 22(10): 2363-2368.
CAI L B, WANG J G, ZHU X Q. Numerical simulation of high power microwave breakdown on dielectric surface at different gas pressures[J]. High Power Laser and Particle Beams, 2010, 22(10): 2363-2368(in Chinese).

[20] 蔡利兵, 王建国. 介质表面高功率微波击穿中释气现象的数值模拟研究[J]. 物理学报, 2011, 60(2): 025217.
CAI L B, WANG J G. Numerical simulation of outgassing in the breakdown on dielectric surface irradiated by high power microwave[J]. Acta Physica Sinica, 2011, 60(2): 025217(in Chinese).

[21] WANG J G, CAI L B, ZHU X Q, et al. Numerical simulations of high power microwave dielectric interface breakdown involving outgassing[J]. Physics of Plasmas, 2010, 17(6): 063503.

[22] 董烨, 周前红, 董志伟, 等. 高功率微波沿面闪络击穿机制粒子模拟[J]. 强激光与粒子束, 2013, 25(4): 950-958.
DONG Y, ZHOU Q H, DONG Z W, et al. PIC simulation of mechanism of high power microwave flashover and breakdown on dielectric surface[J]. High Power Laser and Particle Beams, 2013,

25(4): 950-958(in Chinese).

[23] 李定. 等离子体物理学[M]. 北京: 高等教育出版社, 2006.
LI D. Plasma physics[M]. Beijing: Higher Education Press, 2006 (in Chinese).

[24] VAUGHAN R M. Secondary emission formulas[J]. IEEE Transactions on Electron Devices, 1993, 40(4): 830.

[25] 张慧博, 杨建华, 程国新, 等. 刻槽结构高功率微波输出窗次级电子倍增效应[J]. 强激光与粒子束, 2013, 25(5): 1189-1194.
ZHANG H B, YANG J H, CHENG G X, et al. Investigation on multipactor of high-power microwave window with grooves[J]. High Power Laser and Particle Beams, 2013, 25(5): 1189-1194(in Chinese).

Secondary electron multiplication of aluminum under strong vacuum electromagnetic field

SHANG Shengfei[1, 2, *], YANG Xiaoning[1, 2], YANG Yong[1, 2], BI Yanqiang[1], WU Nankai[1], YU Lantao[1]

(1. Beijing Institute of Spacecraft Environment Engineering, Beijing 100094, China;
2. National key Laboratory of Science and Technology on Reliability and Environmental Engineering, Beijing 100094, China)

Abstract: This paper primarily employs the 1D3V particle-in-cell (PIC) method to investigate the rule of multifactor of aluminum material on satellite surface under the strong electromagnetic environment of space in order to study the charge and discharge effect brought on by the strong electromagnetic environment on the satellite surface. The result shows that there is an "easiest" multiplication range for the multifactor effect of aluminum materials under different microwave amplitudes and frequencies. The multifactor rules are: At a specific frequency, the multifactor of aluminum first increases and then decreases with the rising of microwave electric field amplitude, and there is an optimal secondary electron multiplication interval; At a certain amplitude, the multifactor effect of aluminum will first increase and then decrease, and it shows that the multifactor effect should be strong at low frequency and inhibited at high frequency.

Keywords: high power microwave; space electromagnetic environment; satellite catalogue materials; multifactor; discharge

Received: 2021-09-05; **Accepted**: 2021-11-22; **Published Online**: 2022-01-25 08:40
URL: kns.cnki.net/kcms/detail/11.2625.V.20220124.1735.005.html
* **Corresponding author**. E-mail: shangshengfei@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0565

湿热环境下碳纤维层合板拉伸疲劳性能

许名瑞[1,2]，曾本银[1,2]，熊欣[2]，孟庆春[2]，程小全[1,*]

(1. 北京航空航天大学 航空科学与工程学院，北京 100191；2. 中航工业中国直升机设计研究所，景德镇 333001)

摘 要：湿热环境是影响复合材料层合板力学性能的主要因素之一，研究湿热环境对复合材料结构的影响对于保证飞行结构安全具有非常重要的工程应用意义。研究碳纤维复合材料层合板在室温干态（RTD）、低温干态（CTD）和高温湿态（ETW）3种环境条件下的拉伸疲劳性能，获得3种环境下的S-N曲线与层合板疲劳破坏模式。在此基础上，建立层合板有限元分析模型，对其疲劳性能进行研究，分析讨论温度及湿度对层合板疲劳性能的影响，建立层合板疲劳寿命环境影响因子的确定方法。结果表明：湿热环境对正交层合板的拉伸疲劳性能影响很大，疲劳寿命为10^6次时，与RTD环境相比，CTD环境下层合板的疲劳强度下降了2.76%，而ETW环境下下降达到23.77%；ETW和RTD环境下破坏模式以纤维断裂和分层为主，而CTD环境下却几乎全为纤维断裂破坏；S-N曲线包括疲劳强度快速下降和缓慢下降2个阶段；温度对疲劳性能的影响要明显强于湿度，温度超过45 ℃时湿度对疲劳性能的影响进入强影响区。

关 键 词：复合材料；湿热环境；疲劳性能；S-N曲线；疲劳强度

中图分类号：V215.5$^+$2; TB332

文献标志码：A **文章编号**：1001-5965(2023)07-1614-09

碳纤维增强树脂基复合材料(carbon fibre reinforced plastic, CFRP)由于具有较高的比强度和比刚度已被大量应用于航天航空领域。文献[1-5]的研究成果显示湿热条件是影响树脂基复合材料性能的最重要的环境条件之一，湿热可能会严重降低材料的基本力学性能，湿热环境下，产生大的变形或湿热应力可能会直接导致结构产生破坏。

为了确保复合材料层合板结构的安全性，湿热行为是复合材料结构设计必须考虑的关键问题，为此很多学者就此开展了深入研究。文献[6]研究了碳纤维增强复合材料的湿热老化问题，结果表明湿热联合作用较单纯热或湿作用更加复杂，对复合材料性能的变化具有促进和抵消2种效果。文献[7]指出湿热老化导致碳纤/环氧树脂复合材料层合板中纤维与树脂基体发生界面脱黏损伤，基体性能退化，最终表现为材料的层间剪切性能和面内压缩性能降低。文献[8]给出了T300 /1 034层合板在温度范围−73～177 ℃和相对湿度0%～100%范围内的拉伸力学性能，并总结了几种材料体系层合板在湿热环境下的性能变化。文献[9]对三维编织复合材料进行了应力比为0.1、试验频率为10 Hz的拉-拉载荷作用下疲劳性能试验，疲劳试验达到10^6次停止试验并静载拉断给出剩余强度，试验结果表明，三维编织复合材料疲劳试验试件的剩余强度是静载拉伸破坏强度的110%～127%。文献[10]运用三点弯曲加载方式对3238A/EW250F复合材料的弯曲疲劳性能进行研究，利用升降法得到0°方向铺层3238A/EW250F复合材料10^6次条件下的疲劳极限为73.2 MPa。文献[11]研究了高温(100 ℃)下AS4/PEEK、T800H/Polyimide和T800H/Epoxy这3种复合材料在不同铺层角度的单向碳纤维增强复合材料的疲劳性能，结果发现AS4/PEEK复合材料在除0°铺层外均表现更强的疲劳强度，而T800H/Polyimide复合材料具有较弱的疲劳强度，除0°铺层

收稿日期：2021-09-23；录用日期：2021-11-26；网络出版时间：2021-12-15 10:13
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20211214.2023.003.html
*通信作者. E-mail：xiaoquan_cheng@buaa.edu.cn

引用格式：许名瑞，曾本银，熊欣，等. 湿热环境下碳纤维层合板拉伸疲劳性能[J]. 北京航空航天大学学报，2023，49(7)：1614-1622. XU M R，ZENG B Y，XIONG X，et al. Tensile fatigue properties of carbon fiber laminates in hygrothermal environments[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1614-1622 (in Chinese).

外, 其他铺层材料在循环次数达到 10^5 时, 疲劳强度均出现明显降低。文献 [12] 建立高周疲劳载荷作用下模拟复合材料分层扩展的损伤模型。该模型基于断裂力学和损伤力学之间的内聚力关系, 建立裂纹扩展速率和损伤变量之间的关系。文献 [13] 基于细观力学理论运用多尺度方法, 建立复合材料疲劳寿命估算的概率模型, 代表性体积单元选用正方形或六边形, 引入用于桥接宏观应力和微观应力的应力放大矩阵来计算各个组分的应力。文献 [14] 建立纤维增强复合材料考虑应力比及铺层方向影响的疲劳损伤细观模型, 考虑应力比效应提出无量纲的有效应力概念, 基于细观连续损伤力学理论建立 3 组损伤扩展方程式来分别计算疲劳载荷在轴向、横向和剪力方向作用时材料的疲劳性能。现有的研究大都集中在静力性能或者单个湿或热环境条件下的疲劳性能影响, 试验验证不够充分, 数值仿真不够深入, 缺乏对湿度、温度及两者耦合作用下的深入讨论。

本文在文献 [15] 碳纤维复合材料正交层板吸湿后拉伸疲劳性能研究基础上, 采用试验和数值模拟的方法, 进一步对复合材料层合板在湿、热环境条件下的疲劳性能进行研究分析, 建立层合板疲劳寿命环境影响因子的分析方法, 为复合材料层合结构不同环境下的寿命评估提供一种工程方法与数据支持。

1　试验件与试验

1.1　试验件

试验件为正交层合板, 其构型如图 1 所示, 铺层形式为 $[(0/90)]_n$。预浸碳布材料牌号为 CF3052/3238A, 采用热压罐工艺成型, 具体工艺如下: 室温下抽真空至 0.095 MPa, 之后按 2 ℃/min 升温至 80～90 ℃, 加压 0.45 MPa, 再按 2 ℃/min 升温至 125～130 ℃, 保温 2 h 之后自然冷却至室温(或以不大于 2 ℃/min 降至 60 ℃ 以下)出罐。

试验类型和试验件数量分配如表 1 所示, 其中, 吸湿试验件 3 件, 用于确定所有试验件饱和吸湿状态和时间。拉-拉疲劳有 3 种试验环境, 即室温干态(room temperture and dry, RTD)、低温干态(cool temperature and dry, CTD)及高温湿态(elevated temperture and wet, ETW)。为了确定疲劳试验载荷, 同时进行对应的静力试验。

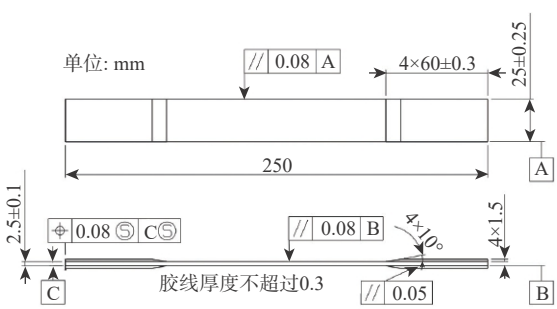


图 1　拉伸静力与疲劳试验件

Fig. 1　Tensile static force and fatigue specimens

表 1　试验类型和试验件数量

Table 1　Test type and number of test specimer

试验类型	试验件数量		
	RTD	CTD	ETW
吸湿试验		3	
拉伸静力试验	3	3	3
拉-拉疲劳试验	16	16	16

1.2　试　验

整个试验工作包含吸湿试验、拉伸静力试验及拉伸疲劳试验, ETW 试验需先在热水槽里完成平衡吸湿再进行相应力学试验, 所有力学试验均在环境箱中完成。

1.2.1　吸湿试验

吸湿试验按照标准试验方法[16-17]来进行设计, 将 ETW 试验件置于 70 ℃ 恒温水槽中进行侵泡, 吸湿首日每 2 h 进行一次称重, 之后逐渐增加称重时间间隔, 每 7 天据式 (1) 计算试验件吸湿量, 当相隔 7 天试验件吸湿量变化小于 0.02% 时认为试验件达到饱和吸湿状态。

$$M = \frac{W_i - W_0}{W_0} \times 100\% \tag{1}$$

式中: M为试验件吸湿量, %; W_0和W_i分别为吸湿前后的试验件质量, g。

1.2.2　疲劳试验

图 2 为碳纤维 CF3052/3238A 拉伸疲劳试验装



图 2　碳纤维 CF3052/3238A 拉伸疲劳试验装置

Fig. 2　Tensile fatigue test device of carbon fiber CF3052/3238A

置，疲劳试验在10 t级Instron8801液压伺服疲劳试验机上完成，RTW试验直接在实验室室温环境进行，ETW和CTD试验在环境箱中进行。由于复合材料疲劳试验中存在发热及水分流失现象，采用文献[15]使用的方法来对试验件进行保湿和表面温度监测。标准试验方法[16-17]指出加载频率应该保证试验件表面温度变化不超过10 ℃，因此，根据试验件表面温度的监测情况后加载频率选择为6 Hz，加载波形为正弦波，载荷应力比R为0.052 6。

2 疲劳寿命有限元模拟

2.1 湿热环境条件下基本性能退化模型

复合材料湿热环境条件下的基本力学性能会有退化的趋势，采用文献[15]建立的模型来描述基本力学性能变化规律，无量纲T^*考虑了温度和吸湿量对基本力学性能的影响。

$$T^* = \frac{T_{gw} - T}{T_{g0} - T_0} \tag{2}$$

式中：T为当前温度；T_0为起点温度；T_{gw}为当前吸湿量为C时树脂玻璃化转化温度；T_{g0}为干态时树脂玻璃化转化温度；T_{gw}和T_{g0}满足：

$$T_{gw} = T_{g0} - gC \tag{3}$$

式中：C为当前树脂吸湿量；g为常数。树脂吸湿量C由层合板吸湿量M、树脂密度ρ_m、层合板密度ρ和树脂体积含量V_m按式(4)计算：

$$C = \frac{\rho M}{\rho_m V_m} \tag{4}$$

得到无量纲T^*后，复合材料层合板的强度和刚度按式(5)进行衰减：

$$\begin{cases} \dfrac{E_{11}}{E_{11}^0} = \dfrac{E_{22}}{E_{22}^0} = (T^*)^a \\ \dfrac{X_T}{X_T^0} = \dfrac{Y_T}{Y_T^0} = (T^*)^b \\ \dfrac{G_{12}}{G_{12}^0} = (T^*)^c \\ \dfrac{S_{12}}{S_{12}^0} = (T^*)^d \end{cases} \tag{5}$$

式中：E_{11}和E_{22}分别为复合材料层合板当前环境下0°和90°方向弹性模量；E_{11}^0和E_{22}^0分别为层合板RTD环境下0°和90°方向弹性模量；X_T和Y_T分别为层合板当前环境下0°和90°方向的拉伸强度；X_T^0和Y_T^0分别为层合板RTD环境下0°和90°方向的拉伸强度；G_{12}和S_{12}分别为层合板当前环境下的纵横剪切模量和强度；G_{12}^0和S_{12}^0分别为层合板RTD环境下纵横剪切模量和强度；a、b、c、d为湿热环境下材料性能退化常数。

使用的材料参数及经验常数如表2和表3所示。

表2 材料参数

Table 2 Material parameters

E_{11}^0/GPa	E_{22}^0/GPa	X_T^0/MPa	Y_T^0/MPa	G_{12}^0/GPa
54.3	54.3	680	680	3.26
S_{12}^0/GPa	ρ/(g.cm^{-3})	ρ_m/(g.cm^{-3})	V_m/%	
116	1.42	1.23	45	

表3 经验常数

Table 3 Empirical constants

T_{g0}/℃	T_0/℃	g/℃	a	b	c	d
120	20	5	0.05	0.15	0.22	0.56

2.2 疲劳损伤起始判据及强度刚度退化准则

文献[18-19]显示复合材料在疲劳加载过程初始阶段，材料边界或孔边会先出现横向微裂纹，随着载荷的增大或循环次数的增加边界或孔边微裂纹会向内部扩展，达到饱和后在该处为起点会出现分层现象，到寿命后期阶段整个损伤模式以纤维断裂为主，随着纤维断裂的增多材料损伤再次扩展，直到剩余强度下降到外加载荷时，试验件完全破坏。因此，从宏观演变来看整个过程当中复合材料层合板的强度和刚度整体上都是呈“快-慢-快”退化趋势，采用文献[15]的退化模型式(6)～式(9)来对退化规律进行描述。

$$\frac{X_T(n,\sigma,R)}{X_T} = \frac{X_C(n,\sigma,R)}{X_C} = \frac{Y_T(n,\sigma,R)}{Y_T} = \frac{Y_C(n,\sigma,R)}{X_C} = 1 - \left[1 - \left(\frac{\sigma_1}{X_T}\right)\right]\left(\frac{n}{N}\right)^{1.321\,8} \tag{6}$$

$$\frac{E_{11}(n,\sigma,R)}{E_{11}} = 1 - \left[1 - \left(\frac{\sigma_1}{0.885\,7X_T}\right)^{1/1.070\,2}\right]\left(\frac{n}{N}\right)^{0.541\,8} \tag{7}$$

$$\frac{S_{12}(n,\sigma,R)}{S_{12}} = 1 - \left[1 - \left(\frac{\tau_{12}}{S_{12}}\right)\right]\left(\frac{n}{N}\right)^{9.345\,9} \tag{8}$$

$$\frac{G_{12}(n,\sigma,R)}{G_{12}} = 1 - \left[1 - \left(\frac{\tau_{12}}{4.785\,3S_{12}}\right)^{1/14.939\,3}\right]\left(\frac{n}{N}\right)^{0.262\,3} \tag{9}$$

式中：$X_T(n,\sigma,R)$、$X_C(n,\sigma,R)$、$Y_T(n,\sigma,R)$、$Y_C(n,\sigma,R)$、$S_{12}(n,\sigma,R)$和$G_{12}(n,\sigma,R)$分别为疲劳载荷加载n次后的纵向拉伸强度、纵向压缩强度、横向拉伸强度、横向压缩强度、纵向弹性模量、面内剪切强度和面内剪切模量；X_T、X_C、Y_T、Y_C和E_{11}、S_{12}、G_{12}分别为疲劳载荷加载前即初始状态下的相应强度和模量；σ_1和τ_{12}分别为正应力和12面内剪切应力；N为相应疲劳载荷下的疲劳寿命。

式(10)～式(14)为按照文献[15]建立的数学模型来描述复合材料层合板起始失效判据，失效破

坏模式包括纤维拉伸、压缩及基体-纤维剪切等5种模式。

1) 纵向纤维拉伸损伤($\sigma_1 \geqslant 0$)

$$f_{1t}^2 = \left(\frac{\sigma_1}{X_T(n,\sigma,R)}\right)^2 + \left(\frac{\tau_{12}}{S_{12}(n,\sigma,R)}\right)^2 + \left(\frac{\tau_{13}}{S_{13}(n,\sigma,R)}\right)^2 \geqslant 1 \tag{10}$$

2) 纵向纤维压缩损伤($\sigma_1 \leqslant 0$)

$$f_{1c}^2 = \left(\frac{\sigma_1}{X_C(n,\sigma,R)}\right) \geqslant 1 \tag{11}$$

3) 横向纤维拉伸损伤($\sigma_2 \geqslant 0$)

$$f_{2t}^2 = \left(\frac{\sigma_2}{Y_T(n,\sigma,R)}\right)^2 + \left(\frac{\tau_{12}}{S_{12}(n,\sigma,R)}\right)^2 + \left(\frac{\tau_{23}}{S_{23}(n,\sigma,R)}\right)^2 \geqslant 1 \tag{12}$$

4) 横向纤维压缩损伤($\sigma_2 \leqslant 0$)

$$f_{2c}^2 = \left(\frac{\sigma_2}{Y_C(n,\sigma,R)}\right) \geqslant 1 \tag{13}$$

5) 基体-纤维剪切损伤($\sigma_1 \leqslant 0$)

$$f_{12}^2 = \left(\frac{\sigma_1}{X_C(n,\sigma,R)}\right)^2 + \left(\frac{\tau_{12}}{S_{12}(n,\sigma,R)}\right)^2 + \left(\frac{\tau_{23}}{S_{23}(n,\sigma,R)}\right)^2 \geqslant 1 \tag{14}$$

式中: σ_i为单元各个方向上的正应力分量; τ_{ij}为各个面内剪应力; $S_{13}(n,\sigma,R)$和$S_{23}(n,\sigma,R)$分别为疲劳载荷下循环n次后单层板13和23平面内的剪切强度。

当单元出现初始损伤后, 需要对单元的材料性能再进行局部刚度衰减, 衰减规律按照以下模型进行, 正交各向异性复合材料单层板的本构关系为

$$\begin{bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\\ \tau_{23}\\ \tau_{31}\\ \tau_{12}\end{bmatrix} = \begin{bmatrix}C_{11} & C_{12} & C_{13} & 0 & 0 & 0\\ C_{21} & C_{22} & C_{23} & 0 & 0 & 0\\ C_{31} & C_{32} & C_{33} & 0 & 0 & 0\\ 0 & 0 & 0 & C_{44} & 0 & 0\\ 0 & 0 & 0 & 0 & C_{55} & 0\\ 0 & 0 & 0 & 0 & 0 & C_{66}\end{bmatrix} \begin{bmatrix}\varepsilon_1\\ \varepsilon_2\\ \varepsilon_3\\ \gamma_{23}\\ \gamma_{31}\\ \gamma_{12}\end{bmatrix} \tag{15}$$

写成矩阵形式为

$$\boldsymbol{\sigma} = \boldsymbol{C}\boldsymbol{\varepsilon} \tag{16}$$

式中: σ_i和τ_{ij}与式(6)～式(14)一致; ε_i为3个主方向线应变; γ_{ij}为相应面内剪应变; $\boldsymbol{C}$为三维刚度矩阵。

渐进退化模型认为当单元出现损伤后, 材料的性能会按照一定的规律逐渐衰减到最低值。引入损伤量矩阵 $\boldsymbol{D}$ 来表征材料的损伤程度, $\boldsymbol{D}$=**0**表示材料无损伤, $\boldsymbol{D}$=**1**表示材料完全破坏。出现损伤后, 对单元的刚度矩阵进行折减:

$$\widehat{\boldsymbol{C}} = (\boldsymbol{1} - \boldsymbol{D})\boldsymbol{C} \tag{17}$$

式中: $\widehat{\boldsymbol{C}}$为折减刚度矩阵, 系数$\widehat{C}_{ij}$按式(18)进行计算:

$$\begin{cases}\widehat{C}_{12} = (1-d_{1t})(1-d_{1c})(1-d_{2t})(1-d_{2c})C_{12}\\ \widehat{C}_{22} = (1-d_{2t})(1-d_{2c})C_{22}\\ \widehat{C}_{13} = (1-d_{1t})(1-d_{1c})C_{13}\\ \widehat{C}_{23} = (1-d_{2t})(1-d_{2c})C_{23}\\ \widehat{C}_{33} = C_{33}\\ \widehat{C}_{44} = (1-d_{12})C_{44}\\ \widehat{C}_{55} = C_{55}\\ \widehat{C}_{44} = C_{66}\end{cases} \tag{18}$$

式中: d_{1t}和d_{1c}分别表征织物单层板1方向拉伸和压缩的损伤程度; d_{2t}和d_{2c}分别表征织物单层板2方向拉伸和压缩的损伤程度; d_{12}表征织物单层板面内剪切的损伤程度。损伤变量d_a(a = 1t, 1c, 2t, 2c, 12)为f_a(a = 1t, 1c, 2t, 2c)的连续函数, 可由式(19)计算得出:

$$\begin{cases}d_{1t} = 1 - \dfrac{1}{f_{1t}}e^{(1-f_{1t})\left(\frac{X_T\varepsilon_t^x}{G_{1t}}\right)} = 1 - \dfrac{1}{f_{1t}}e^{(1-f_{1t})\left(\frac{X_T\varepsilon_t^x L_c}{W_{1t}}\right)}\\ d_{1c} = 1 - \dfrac{1}{f_{1c}}e^{(1-f_{1c})\left(\frac{X_C\varepsilon_c^x}{G_{1c}}\right)} = 1 - \dfrac{1}{f_{1c}}e^{(1-f_{1c})\left(\frac{X_C\varepsilon_c^x L_c}{W_{1c}}\right)}\\ d_{2t} = 1 - \dfrac{1}{f_{2t}}e^{(1-f_{2t})\left(\frac{Y_T\varepsilon_t^y}{G_{2t}}\right)} = 1 - \dfrac{1}{f_{2t}}e^{(1-f_{2t})\left(\frac{Y_T\varepsilon_t^y L_c}{W_{2t}}\right)}\\ d_{2c} = 1 - \dfrac{1}{f_{2c}}e^{(1-f_{2c})\left(\frac{Y_C\varepsilon_c^y}{G_{2c}}\right)} = 1 - \dfrac{1}{f_{2c}}e^{(1-f_{2c})\left(\frac{Y_C\varepsilon_c^y L_c}{W_{2t}}\right)}\\ d_{12} = 1 - \dfrac{\gamma(\gamma^f - \gamma_{12})}{\gamma_{12}(\gamma^f - \gamma)},\ \gamma^f = \dfrac{2G_S}{S_{12}}\end{cases} \tag{19}$$

式中: G_a和W_a(a = 1t, 2c, 2t, 2c)分别为织物单层板不同损伤形式下材料的单位体积断裂能密度和单位面积断裂能密度; L_t 和L_c分别为拉伸和压缩载荷下的单元特征长度; ε_t^x和ε_c^x分别为1方向拉伸和压缩破坏应变; ε_t^y和ε_c^y分别为2方向拉伸和压缩破坏应变; γ为破坏起始剪应变; γ^f为最终破坏剪应变; G_S为单位体积剪切断裂能。

2.3　疲劳有限元模型

图3为采用ABAQUS 6.14软件建立有限元模型, 模型分固支约束段、工作段和加载段3部分, 在厚度方向将层合板划分8层, 每层划分一层网格, 为提高计算精度单元类型选择为C3D8R, 沙漏控制选择增强模式, 共划分59 280个单元。载荷及约束条件为: 加载段上下表面的面内施加剪切力载荷; 固支段上下表面的所有节点施加固支约束。

通过对2个准静态分析步施加最大疲劳载荷来实现ABAQUS对疲劳过程的模拟。第1个分析步对层合板施加拉伸载荷, 大小与疲劳最大载荷相

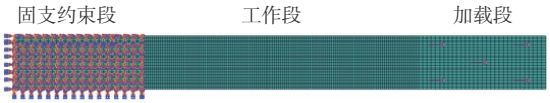


图 3 有限元模型

Fig. 3 Finite element model

同;第 2 个分析步保持上一个分析步载荷不变,设置固定增量步大小,每个增量步代表一定循环数(500 个),根据当前循环数对材料刚度、强度性能进行衰减。

使用织物复合材料为 CF3052/3238A,材料性能参数如表 4 所示。

表 4 材料弹性工程常数

Table 4 Elastic engineering constants of materials

E_1/GPa	E_2/GPa	E_3/GPa	G_{12}/GPa	G_{13}/GPa	G_{23}/GPa
54.3	54.3	3 300	3.26	2.17	2.17
v_{12}	v_{13}	v_{23}	X_T/MPa	X_C/MPa	Y_T/MPa
0.04	0.01	0.01	680	614.29	680
Y_C/MPa	S_{12}/MPa	S_{13}/MPa	S_{23}/MPa		
614.29	115.98	73.5	73.5		

采用 ABAQUS 自带的 Umat 子程序来对层合板进行疲劳分析模拟,在 Umat 子程序环境下使用 Fortran 语言来编写材料的本构刚度退化方程和初始损伤失效判据以实现整个疲劳分析的渐进损伤演化,整个疲劳分析流程图如图 4 所示。

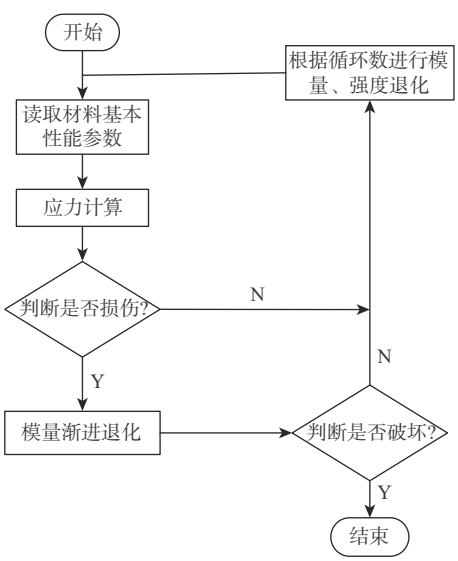


图 4 疲劳仿真分析流程图

Fig. 4 Flow chart of fatigue simulation analysis

3 结果与讨论

3.1 拉伸静力试验结果

3 种环境下各完成了静力拉伸试验件 3 件,平均拉伸强度性能如表 5 所示,对比 RTD 环境,ETW 环境拉伸强度下降约 8.71%,而 CTD 环境拉伸强度则提高了 2.16%,湿热环境对弹性模量几乎没有影响。

表 5 3 种湿热环境条件下拉伸静力试验结果

Table 5 Tensile static test results under three different hygrothermal environments

环境	最大载荷/kN	拉伸强度均值/MPa	弹性模量/GPa
RTD	48.39	679.58	54.3
CTD	50.53	694.24	54.1
ETW	45.05	620.41	52.5

3.2 疲劳试验结果

3.2.1 湿热环境对层合板疲劳性能的影响

疲劳试验结果如表 6 所示,共完成 3 种环境的疲劳试验,每种环境分 4 组试验,每组试验包含 4 个试验件,其寿命全部列在表 6 的疲劳寿命一栏中,每个环境下的应力水平是对应其拉伸强度均值的相对值。

图 5 为对数坐标下的线性曲线拟合,可知同样应力水平下疲劳寿命 N 依次为:$N_{RTD}>N_{CTD}>N_{ETW}$,选取 $N=10^6$ 次时的拟合曲线结果进行对比,此时 RTD、CTD 及 ETW 环境下的疲劳强度依次为 578.0 MPa、576.0 MPa 和 441.6 MPa。

从图 5 可知,相比 RTD 环境,CTD 及 ETW 环境下的疲劳强度依次降低了 2.76% 和 23.77%,而对比拉伸强度 CTD 环境下提高了 2.16%,ETW 环境下降低了 8.71%,可见低温对正交层板的拉伸强度和疲劳强度影响都较小,而饱和吸湿和高温的共同作用对其疲劳性能影响显著。

图 6 为线性坐标下的正规则化疲劳 S-N 曲线,采用文献 [20] 中的方法通过对横坐标和纵坐标分别引入 $N=10^6$ 和各自的拉伸强度可以获得,可以更明显看到疲劳强度与循环次数的变化趋势。

从图 6 明显看到疲劳强度下降分为 2 个阶段“A”阶段和“B”阶段,“A”阶段为快速下降阶段,整个循环次数占比 20%,而疲劳强度下降占比却超过了 80%;“B”阶段为缓慢下降阶段,循环次数占比 80%,而疲劳强度下降占比却不到 20%。

式(20)为采用文献 [21] 建立的 Wohler 疲劳 S-N 方程来对试验数据进行线性函数拟合。

$$S = A + B\lg N \tag{20}$$

式中:S为最大疲劳应力;N为疲劳寿命;A和B为曲线参数。拟合得到 3 种环境下的疲劳 S-N 曲线分别为

$$S_{RTD} = 685.82 - 17.97\lg N_{RTD} \tag{21}$$

$$S_{CTD} = 672.01 - 18.33\lg N_{CTD} \tag{22}$$

$$S_{ETW} = 718.33 - 46.29\lg N_{ETW} \tag{23}$$

表 6　3 种湿热环境下的疲劳试验结果（应力比 R=0.052 6）

Table 6　Fatigue test results under three different hygrothermal environments (Stress ratio R=0.052 6)

湿热环境	应力水平/%	最大应力/MPa	疲劳寿命/次
RTD	85.1, 83.1, 83.1, 83.1	578, 564.4, 564.4, 564.4,	3 699, 508 210, 753 116, 1 200 331
	87.1, 87.1, 85.1, 85.1	591.6, 591.6, 578, 578	43 027, 545 041, 1 200 200, 41 123
	90.1, 90.1, 90.1, 89.1	612, 612, 612, 605.2	159 724, 22 975, 67 479, 378 378
	94.1, 92.1, 92.1, 92.1	639.2, 625.6, 625.6, 625.6	1 836, 33 665, 8 491, 17 341
CTD	80	555.2	1 130 123, 1 072 928, 466 931, 760 758
	83	576.02	113 964, 141 959, 241 866, 155 203
	85	589.9	97 129, 103 678, 19 140, 3 610
	87	603.78	12 530, 211 101, 7 681, 4 500
ETW	70	555.2	1 305 108, 356 100, 98 753, 755 874
	75	576.0	133 981, 678 850, 578 040, 390 576
	84	589.9	47 578, 36 791, 11 762, 88 784
	86	603.8	36 877, 15 680, 7 160, 4 955

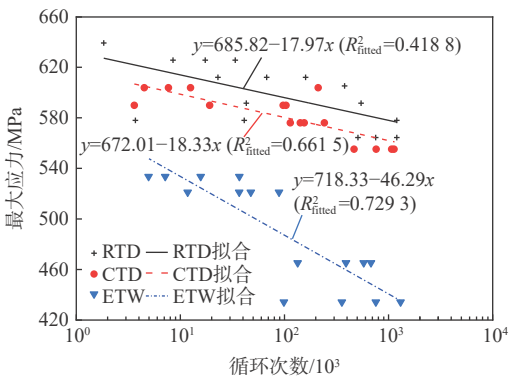


图 5　正交层合板 3 种湿热环境下的 S-N 曲线

Fig. 5　S-N curve of orthogonal laminate under three different hygrothermal environments

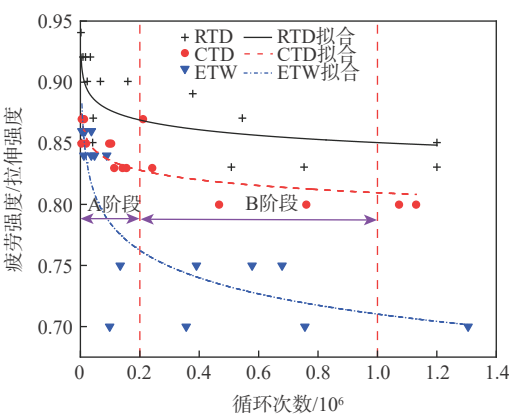
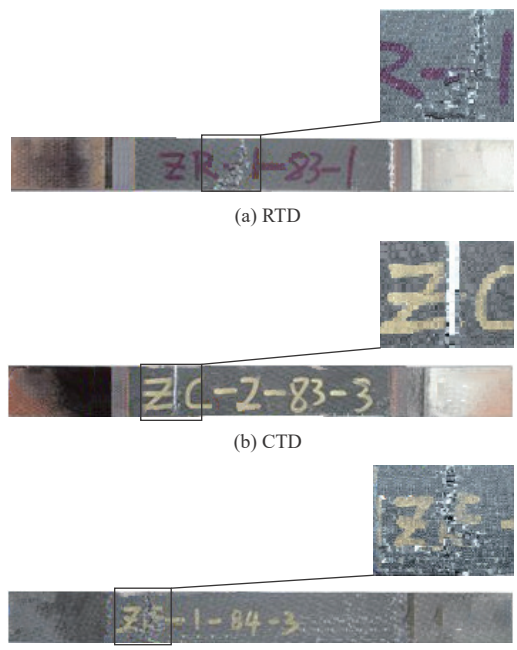


图 6　正交层合板 3 种湿热环境下的正规则化 S-N 曲线

Fig. 6　Normalized fatigue S-N curve of orthogonal laminate under three different hygrothermal environments

3.2.2　破坏模式分析

3 种环境下的疲劳破坏形貌如图 7 和图 8 所示, 可知 RTD 和 ETW 环境下伴随着纤维断裂、基体开裂和分层等破坏模式, RTD 环境下疲劳断口的分层破坏连续贯穿整个厚度方向, ETW 环境下分层破坏并不连续只在下半部分出现较多。而 CTD 环境下疲劳断口基本以纤维断裂为主, 断口形貌非常平整, 分层及基体开裂现象较少, CTD 和 RTD 环



(a) RTD

(b) CTD

(c) ETD

图 7　正交铺层层合板 3 种环境下的疲劳破坏形貌

Fig. 7　Fatigue damage morphology of orthogonal laminates under three different hygrothermal environments



(a) RTD　(b) CTD　(c) ETW

图 8　正交铺层层合板 3 种环境下沿厚度疲劳破坏形貌

Fig. 8　Fatigue damage morphology along thickness of orthogonal laminates under three different hygrothermal environments

境下疲劳性能差别较小，但是破坏模式相差较大，这是由低温导致基体脆化造成的。

3.3 有限元结果

3.3.1 模型验证

疲劳寿命最终计算结果和试验结果对比如表 7 所示，以 RTD 环境下拉伸强度的 83.1%，即 564.4 MPa 为施加疲劳最大载荷进行有限元分析，在工作段和固支约束段的交界处因为结构刚度变化大存在应力集中，因此，首先会在交界处表面 1～2 层单元出现纵向纤维拉伸损伤；随着增量步的增加，损伤在厚度方向缓慢扩展；纵向纤维拉伸损伤贯穿整个厚度方向，此时试验件发生完全破坏，可以得到有限元预测值为 635 074 次。试验测得寿命循环次数为 1 200 200、753 116、和 508 210，取均值 820 509，最终破坏位置和破坏模式如图 9 所示，有限元计算与试验结果基本一致，可见此方法可以有效预测湿热环境下的疲劳寿命。

表 7 有限元模拟预测值与试验值对比

Table 7 Comparison between predicted values of finite element simulation and test values

有限元数值		循环次数	对数寿命
试验值	1#	1 200 200	6.08
	2#	753 116	5.87
	3#	508 210	5.71
	均值	820 509	5.89
预测值		635 074	5.80

注：有限元预测值与试验值均值的循环次数误差为22.6%，对数寿命误差为1.5%。

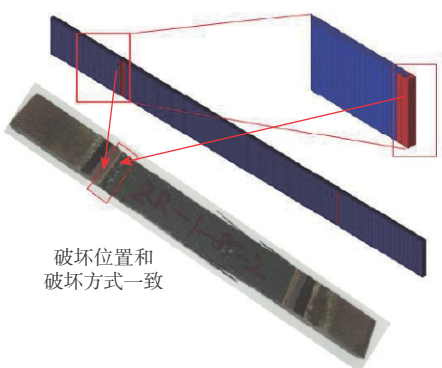


图 9 有限元计算和试验结果最终破坏模式对比

Fig. 9 Comparison of final damage modes between finite element calculation and test results

3.3.2 不同湿热环境下的疲劳寿命影响

为探究温度和湿度耦合作用对复合材料层合板疲劳寿命的影响，保持载荷为 591.5 MPa 不变，通过有限元模型计算不同温度（20 ℃、35 ℃、45 ℃、55 ℃ 和 65 ℃）和不同吸湿量（0%、0.5%、1%、1.5% 和 1.8%）共同作用下层合板的疲劳寿命。

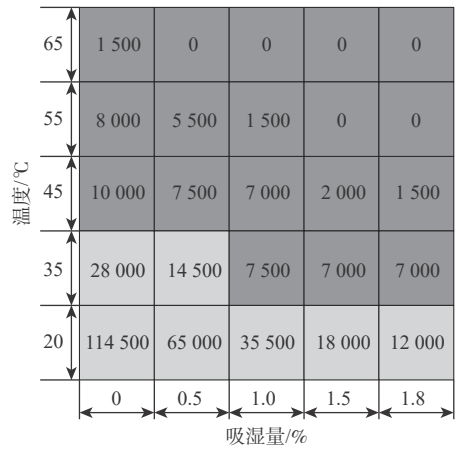


图 10 不同湿热状态下寿命循环次数

Fig. 10 Number of life cycles under different hygrothermal conditions

图 10 为针对不同湿度和温度组合，经有限元计算得到的疲劳寿命。以吸湿量为横轴，温度为纵轴，将计算得到不同湿热状态下的疲劳循环次数值画在一个宫格图中。以寿命循环次数降到室温干态下值的十分之一为临界值，将湿热环境对疲劳寿命的影响分为强影响区和弱影响区，分别以深灰色和浅灰色标记。可以看出，温度对复合材料疲劳寿命的影响更为严重，当温度达到 45 ℃ 时，所有吸湿量下的寿命均处在强影响区。

若对疲劳寿命数值取对数，再做归一化处理，则宫格图如图 11 所示。

假设温度和湿度对于复合材料疲劳寿命的影响是相互独立的，由图 11 中数据，拟合可得到复合材料疲劳寿命湿热影响系数 f 的表达式为

$$f=\frac{\lg N}{\lg N_0}=1-0.194\left(\frac{M}{M_\infty}\right)^{1.903}-0.373\left(\frac{T-T_0}{T_1-T_0}\right)^{1.004} \tag{24}$$

式中：f 为湿热影响系数；N 和 N_0 分别为湿热环境

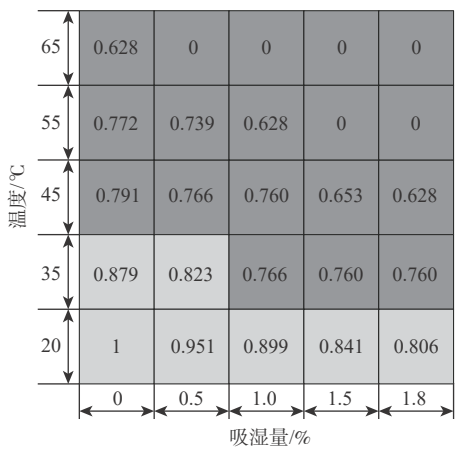


图 11 对数寿命循环次数归一化

Fig. 11 Normalization of logarithmic life cycles

和 RTD 环境下的寿命循环次数; M 和 M_∞ 分别为层合板当前吸湿量和最大吸湿量; T 为当前温度; T_0 为参考干态温度; T_1 为最高温度, 取 65 ℃。

4 结　论

1) 相对于 RTD 环境, 拉伸强度在 CTD 和 ETW 环境下分别上升了 2.16% 和下降了 8.71%, 湿热环境对拉伸模量几乎没有影响。

2) N=10^6 时, 相对于 RTD 环境, CTD 和 ETW 环境下疲劳强度分别下降了 2.76% 和 23.77%, 可见低温对疲劳性能影响较小, 饱和吸湿和高温联合作用对疲劳性能影响显著, 且远超过对拉伸强度的影响。

3) 正交层合板的 S-N 曲线可以明显分为疲劳强度快速下降阶段和缓慢下降阶段, 在快速下降阶段, 不到 20% 的循环次数其疲劳强度下降超过了整个总额的 80%。

4) RTD 和 ETW 环境下以纤维断裂、分层损伤为主, RTD 环境下的分层损伤贯穿整个厚度, ETW 环境下的分层则不连续在下半部分出现较多, CTD 环境下疲劳断口平整基本以纤维断裂为主。

5) 数值分析讨论表明温度对疲劳性能的影响要明显强于湿度, 而且当温度超过 45 ℃ 时所有湿度的影响都处于强影响区。

6) 本文提出一种复合材料疲劳寿命湿热影响系数的确定方法, 可用于不同环境条件下复合材料结构的寿命评估。

参考文献 (References)

[1] 陈浩然, 息志臣, 贺晓东. 复合材料层合板的热变形非线性分析[J]. 大连理工大学学报, 1994, 34(3): 280-286.
CHEN H R, XI Z C D, HE X D. Nonlinear analysis of thermal deformations of composite laminates under transient thermal loading[J]. Journal of Dalian University of Technology, 1994, 34(3): 280-286 (in Chinese).

[2] LIU S F, CHENG X Q, ZHANG Q, et al. An investigation of hygrothermal effects on adhesive materials and double lap shear joints of CFRP composite laminates[J]. Composites Part B:Engineering, 2016, 91: 431-440.

[3] SETHI S, RAY B C. Environmental effects on fibre reinforced polymeric composites: Evolving reasons and remarks on interfacial strength and stability[J]. Advances in Colloid and Interface Science, 2015, 217: 43-67.

[4] ZHANG Q, CHENG X Q, ZHANG J, et al. Experimental and numerical investigation of composite box joint under tensile load[J]. Composites Part B:Engineering, 2016, 107: 75-83.

[5] 沙勐, 熊欣, 许名瑞, 等. 湿热环境对复合材料疲劳性能的影响[J]. 高科技纤维与应用, 2017, 42(4): 37-43.
SHA M, XIONG X, XU M R, et al. Effect of hygrothermal environment on fatigue properties of composite materials[J]. Hi-Tech Fiber and Application, 2017, 42(4): 37-43(in Chinese).

[6] MEZIERE Y, BUNSELL A R, FAVRY Y, et al. Large strain cyclic fatigue testing of unidirectional carbon fibre reinforced epoxy resin[J]. Composites Part A:Applied Science and Manufacturing, 2005, 36(12): 1627-1636.

[7] COSTA M L, REZENDE M C, DE ALMEIDA S F M. Strength of hygrothermally conditioned polymer composites with voids[J]. Journal of Composite Materials, 2005, 39(21): 1943-1961.

[8] SHEN C H, SPRINGER G S. Effects of moisture and temperature on the tensile strength of composite materials[J]. Journal of Composite Materials, 1977, 11(1): 2-16.

[9] 李嘉禄, 杨红娜, 寇长河. 三维编织复合材料的疲劳性能[J]. 复合材料学报, 2005, 22(4): 172-176.
LI J L, YANG H N, KOU C H. Fatigue properties of three dimensional braiding composites[J]. Acta Materiae Compositae Sinica, 2005, 22(4): 172-176(in Chinese).

[10] 马丽婷, 陈新文, 邓立伟, 等. 复合材料疲劳性能研究[J]. 航空制造技术, 2013, 53(23): 73-74.
MA L T, CHEN X W, DENG L W, et al. Fatigue performance of composites[J]. Aeronautical Manufacturing Technology, 2013, 53(23): 73-74(in Chinese).

[11] KAWAI M, YAJIMA S, HACHINOHE A, et al. High-temperature off-axis fatigue behaviour of unidirectional carbon-fibre-reinforced composites with different resin matrices[J]. Composites Science and Technology, 2001, 61(9): 1285-1302.

[12] TURON A, COSTA J, CAMANHO P P, et al. Simulation of delamination in composites under high-cycle fatigue[J]. Composites Part A:Applied Science and Manufacturing, 2007, 38(11): 2270-2282.

[13] MUSTAFA G, CRAWFORD C, SULEMAN A. Fatigue life prediction of laminated composites using a multi-scale M-LaF and Bayesian inference[J]. Composite Structures, 2016, 151: 149-161.

[14] 张文姣. 纤维增强复合材料的疲劳损伤模型及分析方法[D]. 哈尔滨: 哈尔滨工业大学, 2015 : 79-90.
ZHANG W J. Fatigue damage model and analysis method of fiber reinforced composites[D]. Harbin: Harbin Institute of Technology, 2015: 79-90 (in Chinese).

[15] 张祥林, 孟庆春, 许名瑞, 等. 吸湿后碳纤维复合材料正交层板拉伸疲劳性能[J]. 材料工程, 2021, 49(8): 169-177.
ZHANG X L, MENG Q C, XU M R, et al. Tensile fatigue properties of carbon fiber reinforced composite orthogonal laminates after moisture absorption[J]. Journal of Materials Engineering, 2021, 49(8): 169-177(in Chinese).

[16] 国防科学技术工业委员会. 碳纤维树脂基复合材料层合板疲劳试验方法: GJB 2637-96[S]. 北京: 中国航空工业总公司, 1996: 1-7.
Commission of Science, Technology and Industry for National Defense. Test methods for fatigue of carbon fiber resin matrix composite laminates: GJB 2637-96[S]. Beijing: Aviation Industry Corporation of China, 1996:1-7(in Chinese).

[17] American Society of Testing Materials International. D3479 Stanard test method for tension-tension fatigue of polymer matrix composite materials: ASTM D3479/D3479M-2012[S]. West Conshohocken: American Society of Testing Materials International, 2012: 1-6

[18] REIFSNIDER K L, HENNEKE E G, STINCHCOMB W W, et al. Damage mechanics and nde of composite laminates[M]. Mechanics

of Composite Materials. Amsterdam: Elsevier, 1983: 399-420.

[19] 吴富强, 姚卫星. 纤维增强复合材料剩余强度衰减模型[J]. 南京航空航天大学学报, 2008, 40(4): 517-520.
WU F Q, YAO W X. Residual strength degradation model of fiber reinforced plastic[J]. Journal of Nanjing University of Aeronautics and Astronautics, 2008, 40(4): 517-520(in Chinese).

[20] XU M R, HUANG J J, ZENG B Y, et al. Effect of Notch on static and fatigue properties of T800 fabric reinforced composites[J]. Science and Engineering of Composite Materials, 2020, 27(1): 335-345.

[21] SPINDEL J, HAIBACH E. Some considerations in the statistical determination of the slope of SN curves[M]. Statistical Analysis of Fatigue Data. West Conshohocken: American Society of Testing Materials International, 1981.

Tensile fatigue properties of carbon fiber laminates in hygrothermal environments

XU Mingrui[1,2], ZENG Benyin[1,2], XIONG Xin[2], MENG Qingchun[2], CHENG Xiaoquan[1,*]

(1. School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China;
2. AVIC China Helicopter Research and Development Institute, Jingdezhen 333001, China)

Abstract: The hygrothermal environment is one of the main factors affecting the mechanical properties of composite laminates. To maintain the safety of the flight structures, it is crucial to study how hygrothermal conditions affect composite structures.The tensile fatigue properties of carbon fiber composite laminates (CFRP) in room temperature and dry (RTD) condition, cool temperature and dry (CTD) condition and elevated temperature and wet (ETW) condition were experimentally studied, then the S-N curves and fatigue damage modes of CFRP laminates in three different environments were obtained. This served as the foundation for the development of the laminate finite element analysis model, the study of laminate fatigue performance, the analysis and discussion of the effects of temperature and humidity on laminate fatigue performance, and the development of a method for determining the environmental factors influencing the fatigue life of laminates. The results show that the tensile fatigue properties of orthometric laminates are greatly affected by the environment of ETW condition. Compared with the RTD condition, the fatigue strength of orthometric laminates decreases by 2.76% in CTD condition when the fatigue life equates to 10^6, while the fatigue strength of orthometric laminates decreases by 23.77% ETW condition. The damage modes in RTD condition and in ETW condition are mainly fiber fracture and delamination, while the damage mode in CTD condition is almost fiber fracture. The S-N curve includes two stages of rapid and slow reduction for fatigue strength. The influence of temperature on fatigue performance is obviously stronger than that of humidity. When the temperature exceeds 45 ℃, the influence of humidity on fatigue performance enters the strong influence zone.

Keywords: composite materials; hygrothermal environment; fatigue life; S-N curve; fatigue strength

Received: 2021-09-23; **Accepted**: 2021-11-26; **Published Online**: 2021-12-15 10:13
URL: kns.cnki.net/kcms/detail/11.2625.V.20211214.2023.003.html

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0552

一种主瓣干扰下基于滑窗子阵的稀疏测角方法

周必雷[1,*]，李荣锋[2]，曾丽[3]，陈辉[1]，刘维建[1]，李槟槟[1]

(1. 空军预警学院，武汉 430000；2. 航天南湖电子信息技术股份有限公司，荆州 434000；
3. 武汉大学 电子信息学院，武汉 430000)

摘 要：雷达抗主瓣干扰（MLJ）一直是雷达领域的难点问题，针对主瓣干扰环境下的雷达目标角度测量问题，提出一种主瓣干扰下基于滑窗子阵的稀疏测角方法，对各子阵进行自适应主瓣干扰抑制处理，并利用自适应后的子阵间相位关系构建角度原子库，采用正交匹配追踪（OMP）算法估计目标角度。当目标与干扰夹角为1/2个3 dB波束宽度时（目标输入信噪比为20 dB），目标角度估计误差小于1/10倍3 dB波束宽度，所提方法无需先验信息，可同时抑制主、副瓣干扰或多个主瓣干扰，并保证较高的目标测角精度。

关 键 词：主瓣干扰；滑窗子阵；稀疏恢复；抗干扰；雷达
中图分类号：TN911.7
文献标志码：A **文章编号**：1001-5965(2023)07-1623-07

电子干扰与雷达一直以来是“矛”与“盾”的关系，相生相伴。以干扰的空间入射角的不同可以划分为副瓣干扰与主瓣干扰(mainlobe jamming, MLJ)。以干扰造成的威胁形式不同可以划分为压制干扰、欺骗干扰及复合干扰。另外，还有许多新型的干扰样式，如切片转发干扰、灵巧噪声干扰、污染谱干扰、蜂群干扰和狼群干扰等。

主瓣干扰是目前雷达领域的一个难题，其干扰能量强，且与目标空域差异小，因此难以对抗。特别是在主瓣干扰抑制后的目标波达方向(direction of arrival, DOA)估计方面，目前是理论研究与工程实现上的一大难题。主瓣干扰的表现形式主要有伴飞干扰、弹载干扰、拖曳式干扰等。

目前，大多数基于后端信号处理的主瓣干扰对抗研究集中在空域滤波方面，其中自适应波束形成技术[1]是各类方法的基础，其主要思想是在干扰方向自适应形成方向图零陷；另外一种思路是将主瓣干扰抑制作为预处理工作，主要有阻塞矩阵预处理(block matrix preprocessing, BMP)和特征投影预处理(eigen projection matrix preprocessing, EMP)这2种方法；BMP方法需要预估干扰角度，再利用相邻阵元相消法抑制干扰，但存在自由度损失问题；EMP方法是利用对阵列接收信号进行特征分解得到干扰特征矢量，继而构造特征投影矩阵，利用其与干扰子空间的正交性进行干扰抑制；同时，还有在BMP和EMP方法基础上的方向图保形研究工作[2-4]。主瓣干扰对消器(mainlobe jamming cancellation, MLC)[5]利用差波束作为辅助通道进行主瓣干扰对消；为了同时抑制主、副瓣干扰，可以将自适应数字波束形成(adaptive digital beam forming, ADBF)和MLC技术进行级联[6]；上述空域滤波方法均存在自适应综合方向图的主瓣畸变、指向偏移、副瓣电平抬高的问题，这是主瓣干扰背景下空域滤波方法无法避免的情况。上述均为主瓣干扰的抑制工作，还未涉及后续的目标DOA估计研究。

针对主瓣干扰背景下的目标DOA估计问题，

收稿日期：2021-09-14；**录用日期**：2021-10-29；**网络出版时间**：2021-11-09 14:32
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20211109.0904.001.html
基金项目：国家自然科学基金 (62001510)
* **通信作者**. E-mail：zhoubilei666888@sina.com

引用格式：*周必雷，李荣锋，曾丽，等. 一种主瓣干扰下基于滑窗子阵的稀疏测角方法[J]. 北京航空航天大学学报，2023，49(7)：1623-1629. ZHOU B L，LI R F，ZENG L，et al. A sparse estimation method for radar target direction with sliding-window subarray configuration in mainlobe jamming[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1623-1629 (in Chinese)*.

理论研究成熟且工程应用可行的一类方法主要是和差测角方法。最常见的是和差四通道方法[7]及其扩展方法[8-9]，四通道即为和通道、方位差通道、俯仰差通道和方位-俯仰联合差通道。由于方位向和俯仰向在空域上是正交的，所以可以利用方位(或俯仰)向上的和差波束对消抑制主瓣干扰同时不影响目标在俯仰(或方位)向上的和差测角。因此，当目标与干扰的方位(或俯仰)角相等时，这类方法无法测量目标俯仰(或方位)角。此外，还有采用方位(或俯仰)向的ADBF处理，然后在俯仰(或方位)向上对目标进行和差测角[10-11]，其不足之处同和差四通道类方法相同，而相对的优势是抑制主瓣干扰的自由度足够多，因此，其具备多主瓣干扰条件下的目标DOA估计能力。上述方法是基于方位、俯仰正交性的主瓣干扰下测角方法，即某一维度上的干扰抑制对另一维度的目标DOA估计不造成影响；另外一类是和差单脉冲比幅曲线约束法[12-13]，即对自适应单脉冲曲线进行约束以拟合静态单脉冲曲线，主要有两点约束法(线性区)和多点约束法(线性及非线性区)，但该类方法在主瓣区域附近的测角精度较差。

当前在抗主瓣干扰方面有2类新算法，分别是盲源分离算法与稀疏恢复算法。盲源分离的目的不是抑制干扰，而是分离目标与干扰，进而识别目标信号，其主要采用的方法有快速固定点独立成分分析法[14]和特征矩阵联合近似对角化法[15]；盲源分离技术具备抗主瓣干扰能力，但其对后续的目标DOA估计存在困难。稀疏恢复在主瓣干扰背景下发挥的主要功能是目标DOA估计，其将空域滤波与稀疏恢复紧密结合，令空域滤波后的训练样本匹配稀疏恢复算法。

本文提出的基于滑窗子阵的稀疏测角方法联合了空域滤波技术与稀疏恢复算法，其在子阵内进行空域自适应处理来抑制主瓣干扰，并以子阵滑窗的形式保证各子阵间的相位中心以满足稀疏恢复处理的需求，进而估计目标角度。本文方法在同时存在主、副瓣干扰或多个主瓣干扰环境下仍然具备目标DOA估计能力，并具备较高的测角精度。

1 信号模型

考虑N元半波长等距均匀线阵，雷达发射线性调频信号，波长为λ。环境中存在一个目标$s_0(t)$与M个干扰$[s_1(t),s_2(t),\cdots,s_M(t)]$，其采样信号分别为$s_0(n)$和$[s_1(n),s_2(n),\cdots,s_M(n)]$，$n$表示采样点。阵列接收信号可表示为

$$\boldsymbol{x}(n)=\boldsymbol{A}\boldsymbol{s}(n)+\boldsymbol{v}(n)\in\mathbf{C}^{N\times 1} \tag{1}$$

式中：$\boldsymbol{v}(n)$为高斯白噪声；$\boldsymbol{s}(n)$的表达式为

$$\boldsymbol{s}(n)=[s_0(n),s_1(n),\cdots,s_M(n)]^{\mathrm{T}} \tag{2}$$

$\boldsymbol{A}$为阵列流形矩阵，其表达式为

$$\boldsymbol{A}=[\boldsymbol{a}_0,\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m,\cdots,\boldsymbol{a}_M]\in\mathbf{C}^{N\times(M+1)} \tag{3}$$

其中：

$$\boldsymbol{a}_m=\left[1,\exp(-\mu_m),\exp(-2\mu_m),\cdots,\exp(-(N-1)\mu_m)\right]^{\mathrm{T}} \tag{4}$$

式中：$\mu_m=\mathrm{j}2\pi d\sin\theta_m/\lambda$，$\theta_m$为信源入射角，$d$为阵元间距。

2 算法原理

本文方法原理框图如图1所示。考虑等距均匀线阵，经过滑窗处理形成多个子阵，然后在子阵内进行ADBF处理抑制主瓣干扰，最后联合稀疏恢复来估计目标角度。

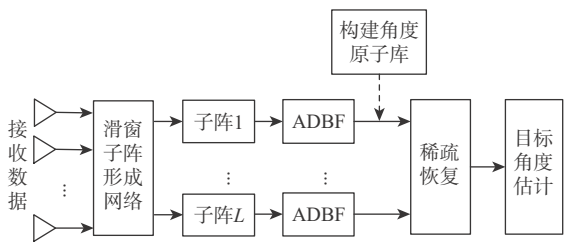


图1 本文方法原理框图

Fig. 1 Diagram of proposed method

2.1 干扰抑制

假设该线阵一共划分为L个子阵，第l个子阵的接收数据为

$$\boldsymbol{x}_l(n)=\sum_{m=0}^{M}\boldsymbol{a}_{ml}s_m(n)+\boldsymbol{v}_l(n) \tag{5}$$

式中：$\boldsymbol{a}_{ml}$为第m个信源在第l个子阵上的导向矢量，且$\boldsymbol{a}_{ml}=\boldsymbol{a}_{m1}\cdot\exp\left[-(l-1)\mu_m\right]$，式(5)可进一步表示为

$$\boldsymbol{x}_l(n)=\sum_{m=0}^{M}\exp\left[-(l-1)\mu_m\right]\boldsymbol{a}_{m1}s_m(n)+\boldsymbol{v}_l(n) \tag{6}$$

第l个子阵接收的干扰加噪声协方差矩阵可以表示为

$$\boldsymbol{R}_l=\sum_{m=1}^{M}\sigma_m^2\boldsymbol{a}_{ml}\boldsymbol{a}_{ml}^{\mathrm{H}}+\sigma_v^2\boldsymbol{I} \tag{7}$$

式中：σ_m^2为干扰能量；σ_v^2为噪声能量，式(7)可重新表示为

$$\begin{aligned}\boldsymbol{R}_l=&\sum_{m=1}^{M}\sigma_m^2\left\{\boldsymbol{a}_{m1}\cdot\exp\left[-(l-1)\mu_m\right]\right\}\cdot\\&\left\{\boldsymbol{a}_{m1}\cdot\exp\left[-(l-1)\mu_m\right]\right\}^{\mathrm{H}}+\sigma_v^2\boldsymbol{I}=\\&\sum_{m=1}^{M}\sigma_m^2\boldsymbol{a}_{m1}\boldsymbol{a}_{m1}{}^{\mathrm{H}}+\sigma_v^2\boldsymbol{I}=\boldsymbol{R}_1\end{aligned} \tag{8}$$

因此，第l个子阵的自适应权矢量可表示为

$$\boldsymbol{w}_l=\boldsymbol{R}_l^{-1}\boldsymbol{a}_{\mathrm{b}l}=\boldsymbol{R}_1^{-1}\boldsymbol{a}_{\mathrm{b}1}\exp\left[-(l-1)\mu_{\mathrm{b}}\right]=\boldsymbol{w}_1\exp\left[-(l-1)\mu_{\mathrm{b}}\right] \tag{9}$$

式中: $\boldsymbol{a}_{\mathrm{b}l}$为第 l 个子阵的波束导向矢量; $\mu_{\mathrm{b}}=\mathrm{j}2\pi\, d\sin\theta_{\mathrm{b}}/\lambda$, θ_{b}为波束指向。故输出信号可表示为

$$\begin{aligned}y_l(n)&=\boldsymbol{w}_l^{\mathrm{H}}\boldsymbol{x}_l(n)=\exp\left[(l-1)\mu_{\mathrm{b}}\right]\boldsymbol{w}_1{}^{\mathrm{H}}\cdot\left\{\sum_{m=0}^{M}\exp\left[-(l-1)\mu_m\right]\boldsymbol{a}_{m1}s_m(n)+\boldsymbol{v}_l(n)\right\}=\\&\boldsymbol{w}_1{}^{\mathrm{H}}\left\{\sum_{m=0}^{M}\exp\left[(l-1)(\mu_{\mathrm{b}}-\mu_m)\right]\cdot\boldsymbol{a}_{m1}s_m(n)+\boldsymbol{v}_l(n)\right\}=\\&\exp\left[(l-1)(\mu_{\mathrm{b}}-\mu_0)\right]\boldsymbol{w}_1{}^{\mathrm{H}}\boldsymbol{a}_{01}s_0(n)+\sum_{m=1}^{M}\exp\left[(l-1)(\mu_{\mathrm{b}}-\mu_m)\right]\boldsymbol{w}_1{}^{\mathrm{H}}\boldsymbol{a}_{m1}s_m(n)+\boldsymbol{w}_1{}^{\mathrm{H}}\boldsymbol{v}_l(n)\end{aligned} \tag{10}$$

式中: 干扰抑制部分$\sum_{m=1}^{M}\boldsymbol{w}_1{}^{\mathrm{H}}\boldsymbol{a}_{m1}s_m(n)=0$, 式(10)可改写为

$$\begin{aligned}y_l(n)&=\exp\left[(l-1)(\mu_{\mathrm{b}}-\mu_0)\right]\boldsymbol{w}_1{}^{\mathrm{H}}\boldsymbol{a}_{01}s_0(n)+\boldsymbol{w}_1{}^{\mathrm{H}}\boldsymbol{v}_l(n)=\\&y_1(n)\exp\left[(l-1)(\mu_{\mathrm{b}}-\mu_0)\right]+\tilde{v}_l(n)\end{aligned} \tag{11}$$

式中: $\tilde{v}_l(n)=\boldsymbol{w}_1{}^{\mathrm{H}}\boldsymbol{v}_l(n)$。

可见, 第 l 个子阵自适应输出y_l与第 1 个子阵自适应输出y_1之间的相位差为$\exp\left[(l-1)(\mu_{\mathrm{b}}-\mu_0)\right]$。由于波束指向$\theta_{\mathrm{b}}$是已知的, 故$\mu_{\mathrm{b}}$已知。因此, 各子阵自适应输出的相位差与输入信号的μ_0有关, 即与输入信号的角度θ_0有关, 这是后续测角的基础。

根据式 (11), L 个子阵的自适应输出可表示为

$$\begin{aligned}\boldsymbol{y}(n)&=\left[y_1(n),y_2(n),\cdots,y_L(n)\right]^{\mathrm{T}}=\\&y_1(n)\left[1,\exp(-\mu_\varepsilon),\cdots,\exp(-(L-1)\mu_\varepsilon)\right]^{\mathrm{T}}\end{aligned} \tag{12}$$

式中: $\mu_\varepsilon=\mu_0-\mu_{\mathrm{b}}$。

2.2　构造角度原子

对式(12)进行稀疏表示:

$$\boldsymbol{y}(n)=\boldsymbol{\Phi}\boldsymbol{\omega}(n)+\tilde{\boldsymbol{v}}(n) \tag{13}$$

式中: $\boldsymbol{\omega}(n)\in\mathbf{C}^{P\times1}$为稀疏系数, P 为角度原子数量;

$$\boldsymbol{\Phi}=\left[\boldsymbol{a}(\theta_{\eta1}),\boldsymbol{a}(\theta_{\eta2}),\cdots,\boldsymbol{a}(\theta_{\eta p}),\cdots,\boldsymbol{a}(\theta_{\eta P})\right]\in\mathbf{C}^{L\times P} \tag{14}$$

其中:

$$\begin{aligned}\boldsymbol{a}(\theta_{\eta p})&=\left[1,\exp\left(-\mathrm{j}2\pi\, d/\lambda\left(\sin\theta_{\eta p}-\sin\theta_{\mathrm{b}}\right)\right),\cdots,=\right.\\&\left.\exp\left(-\mathrm{j}2\pi\, d/\lambda(L-1)\left(\sin\theta_{\eta p}-\sin\theta_{\mathrm{b}}\right)\right)\right]^{\mathrm{T}}\end{aligned} \tag{15}$$

式中: 角度原子$\boldsymbol{\Phi}$需满足有限等距性质(restricted isometry property, RIP)[16]。通常情况下, 角度原子取值范围为一个波束宽度, 综合考虑运算量及稀疏恢复性能, 原子数量通常取 4～6 倍的空间自由度(即子阵数量)。

最后, 需要在角度原子库$\boldsymbol{\Phi}$中匹配得到角度$\theta_{\eta p}$拟合目标角度θ_0, 具体见 2.3 节。

2.3　基于 OMP 算法的目标 DOA 估计

令标签集$\Omega_0=\varnothing$, 初始化残差向量$\boldsymbol{r}_0=\boldsymbol{y}(n)$, 令$k=1$。求$\boldsymbol{\Phi}$中与残差向量$\boldsymbol{r}_{k-1}$的最强相关原子位置:

$$l_k=\arg\max_l|\langle\boldsymbol{\Phi}_l,\boldsymbol{r}_{k-1}\rangle| \tag{16}$$

式中: $\boldsymbol{\Phi}_l$为$\boldsymbol{\Phi}$的第l列, 更新Ω_k:

$$\Omega_k=\Omega_{k-1}\cup\{l_k\} \tag{17}$$

则第k次迭代残余更新为

$$\boldsymbol{r}_k=\boldsymbol{y}(n)-\boldsymbol{\Phi}_{\Omega_k}\boldsymbol{\omega}_k \tag{18}$$

式中: $\boldsymbol{\omega}_k$为稀疏系数, 可通过最小化迭代残余$\boldsymbol{r}_k$进行求解:

$$\boldsymbol{\omega}_k=\arg\min_{\boldsymbol{\omega}_k}\left\|\boldsymbol{y}(n)-\boldsymbol{\Phi}_{\Omega_k}\boldsymbol{\omega}_k\right\|_2 \tag{19}$$

可得

$$\boldsymbol{\omega}_k=\left(\boldsymbol{\Phi}_{\Omega_k}^{\mathrm{H}}\boldsymbol{\Phi}_{\Omega_k}\right)^{-1}\boldsymbol{\Phi}_{\Omega_k}^{\mathrm{H}}\boldsymbol{y}(n) \tag{20}$$

通过反复迭代式 (16)～式 (18) 和式 (20) 直到满足收敛条件。通常迭代终止条件有 2 个, ①已知稀疏度, 可预先设置迭代步数; ②令迭代残余小于某个预先设定值ε, 即$\|\boldsymbol{r}_k\|_2\leqslant\varepsilon$。最后, 根据稀疏系数$\boldsymbol{\omega}$的非零值位置可以估计得到目标的角参数信息。

2.4　算法流程

输入信号为阵列接收信号$\boldsymbol{x}(n)\in\mathbf{C}^{N\times1}$, 初始角度原子库$\tilde{\boldsymbol{\Phi}}\in\mathbf{C}^{L\times P}$。输出信号为稀疏系数$\boldsymbol{\omega}$。

步骤 1　子阵划分, 子阵数为 L。

步骤 2　各子阵自适应输出$\boldsymbol{y}(n)$。

步骤 3　建立稀疏模型: $\boldsymbol{y}(n)=\boldsymbol{\Phi}\boldsymbol{\omega}(n)+\tilde{\boldsymbol{v}}(n)$。

步骤 4　构造角度原子:

$\boldsymbol{\Phi}=\left[\boldsymbol{a}(\theta_{\eta1}),\boldsymbol{a}(\theta_{\eta2}),\cdots,\boldsymbol{a}(\theta_{\eta p}),\cdots,\boldsymbol{a}(\theta_{\eta P})\right]\in\mathbf{C}^{L\times P}$。

步骤 5　令标签集$\Omega_0=\varnothing$, 初始化残差向量$\boldsymbol{r}_0=\boldsymbol{y}(n)$, $k=1$。

步骤 6　求解$\boldsymbol{\Phi}$中与残差向量$\boldsymbol{r}_{k-1}$相关性最强的原子位置: $l_k=\arg\max_l|\langle\boldsymbol{\Phi}_l,\boldsymbol{r}_{k-1}\rangle|$。

步骤 7　更新Ω_k: $\Omega_k=\Omega_{k-1}\cup\{l_k\}$。

步骤 8　残差更新: $\boldsymbol{r}_k=\boldsymbol{y}(n)-\boldsymbol{\Phi}_{\Omega_k}\boldsymbol{\omega}_k$。

步骤 9　求解稀疏系数: $\boldsymbol{\omega}_k=\left(\boldsymbol{\Phi}_{\Omega_k}^{\mathrm{H}}\boldsymbol{\Phi}_{\Omega_k}\right)^{-1}\boldsymbol{\Phi}_{\Omega_k}^{\mathrm{H}}\boldsymbol{x}(n)$, 其体现为信号能量。

步骤 10　迭代终止判据: 令$k\leftarrow k+1$, 重复步骤 6～步骤 9, 直至满足某个迭代终止判据。迭代终止判据一般为某个固定的迭代步数或迭代残余能量小于某个预先设定的值ε: $\|\boldsymbol{r}_k\|_2\leqslant\varepsilon$。

3　仿真实验与性能分析

3.1　仿真实验

考虑等距均匀线阵, 详细仿真条件设置如表 1 所示, 干扰环境设置如表 2 所示。

表 1 仿真条件设置

Table 1 Simulation parameters setting

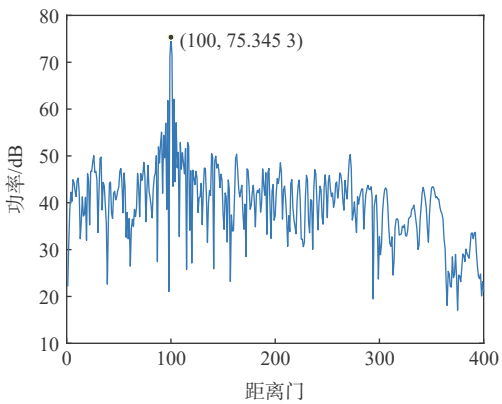
条件	设置
阵列天线	20元等距均匀线阵
阵元间距	半波长
滑窗子阵数	15
波束指向/(°)	0
$B_{0.5}$(3 dB波束宽度)/(°)	5.05
雷达发射信号	线性调频信号, 带宽5 MHz, 脉宽20 μs, 采样频率10 MHz, 阵元级SNR_e=0 dB
干扰	主瓣压制干扰
目标	位于第100个距离门, 空间位置在0°

表 2 干扰环境设置

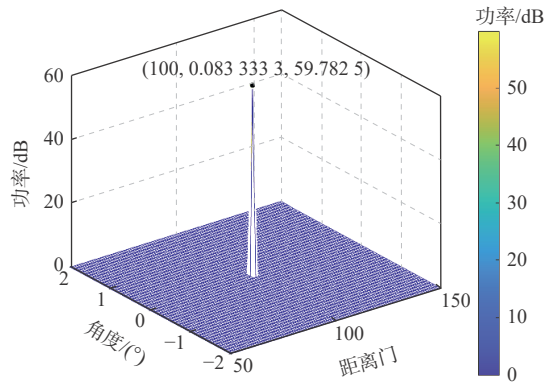
Table 2 Jamming parameters setting

干扰环境	数量	干噪比/dB	类型	干扰与目标的归一化夹角/(°)
1	1	50	主瓣噪声压制干扰	0.5 $B_{0.5}$
2	2	50, 50	主/副瓣噪声压制干扰	0.5 $B_{0.5}$,10 $B_{0.5}$
3	2	50, 50	主/主瓣噪声压制干扰	0.5 $B_{0.5}$, 0.5 $B_{0.5}$

图 2 为 3 种干扰环境下的自适应输出与目标 DOA 估计结果。3 种情况下目标角度估计结果分

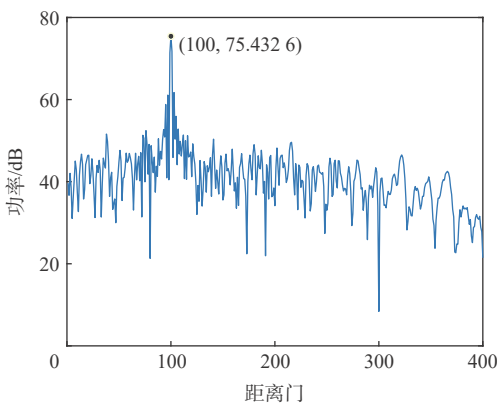


(a1) 滑窗子阵自适应输出

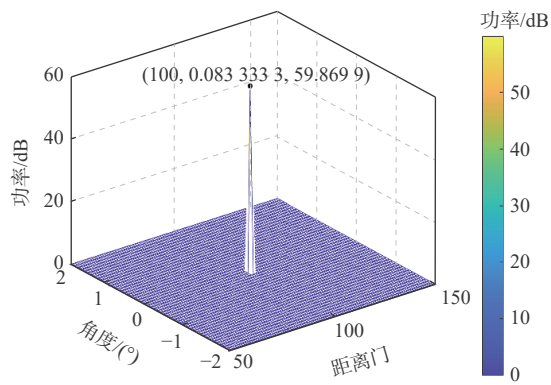


(a2) 目标DOA估计结果

(a) 干扰环境1 (1个主瓣干扰)

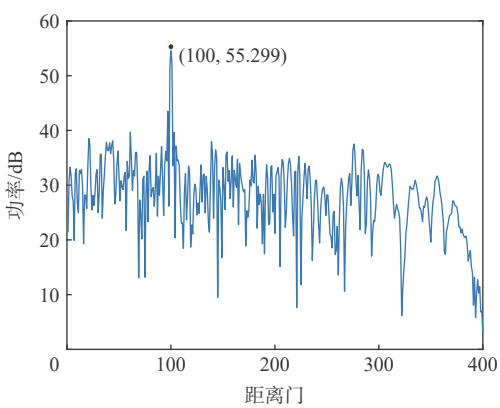


(b1) 滑窗子阵自适应输出

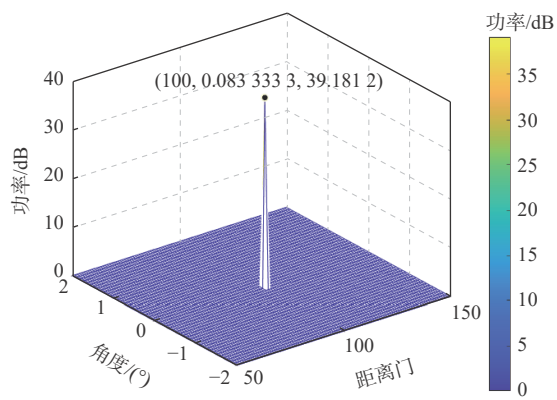


(b2) 目标DOA估计结果

(b) 干扰环境2 (1个主瓣干扰+1个副瓣干扰)



(c1) 滑窗子阵自适应输出



(c2) 目标DOA估计结果

(c) 干扰环境3 (2个主瓣干扰)

图 2 目标参数估计结果

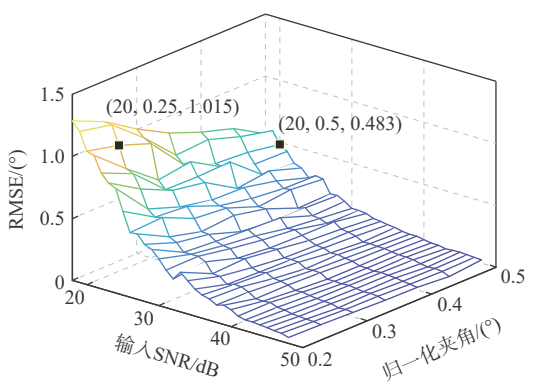
Fig. 2 Target parameter estimation results

别为 0.083 33°、0.083 33°、−0.083 33°, 与初始设置值 0°基本一致。由此分析可知, 本文方法在同时存在主、副瓣干扰或多个主瓣干扰环境下具备目标角度估计能力。

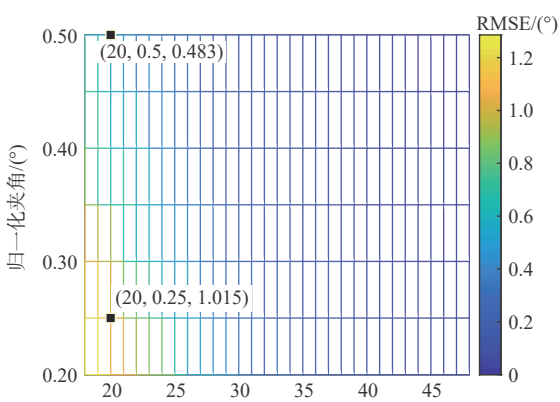
3.2　性能分析

假设电磁环境中存在 1 个干扰, 干噪比(interference-noise ratio, INR)为 50 dB, 干扰角度为 1°～2.5°, 步径为 0.25°, 目标位于 0°, 即目标与干扰的夹角变化范围为 [0.2$B_{0.5}$, 0.5$B_{0.5}$], 目标信噪比(signal-noise ratio, SNR)(波束合成及脉压后)变化区间为 18～48 dB, 其他仿真条件同表 1。仿真结果为 100 次 Monte-Carlo 仿真实验的统计平均。图 3 为目标角度估计均方根误差(root mean squared error, RMSE)曲面, 归一化夹角表示目标与干扰的夹角与 $B_{0.5}$ 比值。可见, 目标 DOA 估计精度将随着 SNR 的增加而提高, 随着归一化夹角的增大而提高。当目标输入 SNR=20 dB, 归一化夹角为 0.5°时, 目标 DOA 估计 RMSE=0.483°, 小于 1/10 倍 3 dB 波束宽度 0.505°; 当目标输入 SNR=20 dB, 归一化夹角为 0.25°时, 目标 DOA 估计 RMSE=1.015°, 近似为 1/5 倍 3 dB 波束宽度 1.01°。

假设电磁环境中存在 1 个干扰, 干噪比 INR=50 dB, 干扰位于 1.25°, 即 1/4 倍 3 dB 波束宽度, 目标角度由 − 2.5°变化至 2.5°, 步径 0.25°, 目标 SNR(波束合成及脉压后)为 20 dB, 其他仿真条件同表 1, 仿真结果为 100 次 Monte-Carlo 仿真实验的统计平均。图 4 比较了本文方法与多点约束法[12], 给出了 2 种方法的目标角度估计 RMSE 曲线。可见, 本文方法估计目标角度精度要高于多点约束法。另外, 2 种方法均存在着干扰角附近区域的目标角度估计精度下降严重的问题, 其原因在于当目标在空域上越接近干扰源时, 其干扰抑制后的目标损失将增大, 由此导致目标 DOA 估计精度降低。对此, 后续拟考虑采用多维信号处理开展主瓣干扰抑制, 寻找干扰抑制后对目标损失最小的某维度或某联合维度, 进而提高目标 DOA 估计精度。



(a) RMSE曲面



(b) 投影图

图 3　目标角度估计 RMSE

Fig. 3　RMSE of target direction estimation

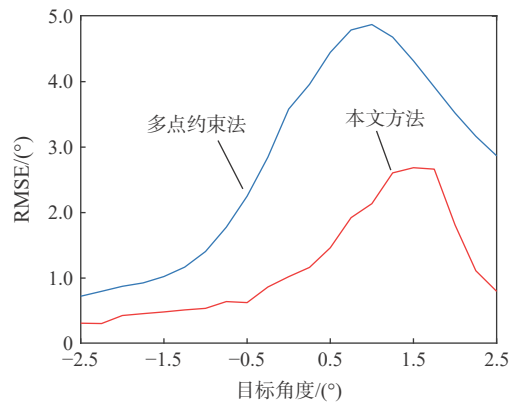


图 4　目标角度估计精度对比曲线

Fig. 4　Comparison curves of target direction estimation accuracy

4　结　论

1) 本文方法可有效抑制主、副瓣干扰或者多个主瓣干扰并测量目标角度。

2) 在测角精度方面, 当目标输入 SNR 为 20 dB(波束合成及脉压后), 目标与干扰夹角为 1/2 个 3 dB 波束宽度时, 目标 DOA 估计 RMSE 小于 1/10 倍 3 dB 波束宽度。

参考文献(References)

[1] 王永良, 丁前军, 李荣锋. 自适应阵列处理[M]. 北京: 清华大学出版社, 2009.

WANG Y L, DING Q J, LI R F. Adaptive array processing[M]. Beijing: Tsinghua University Press, 2009 (in Chinese).

[2] 李荣锋, 王永良, 万山虎. 主瓣干扰下自适应方向图保形方法的研究[J]. 现代雷达, 2002, 24(3): 50-53.

LI R F, WANG Y L, WAN S H. Research of reshaping adapted pattern under mainlobe interference conditions[J]. Modern Radar,

2002, 24(3): 50-53(in Chinese).

[3] 李荣锋, 王永良, 万山虎. 一种在主瓣干扰条件下稳健的自适应波束形成方法[J]. 系统工程与电子技术, 2002, 24(7): 61-64.
LI R F, WANG Y L, WAN S H. Robust adaptive beam forming under main lobe interference conditions[J]. Systems Engineering and Electronics, 2002, 24(7): 61-64(in Chinese).

[4] 潘帅, 张永顺, 葛启超, 等. 基于特征投影矩阵的自适应方向图改善方法[J]. 空军工程大学学报(自然科学版), 2019, 20(2): 47-52.
PAN S, ZHANG Y S, GE Q C, et al. Research on improvement of adaptive pattern based on eigen-projection matrix[J]. Journal of Air Force Engineering University (Natural Science Edition), 2019, 20(2): 47-52(in Chinese).

[5] APPLEBAUM S P, WASIEWICZ R. Main beam jammer cancellation for monopulse sensors: DTIC RADC-TR-86-267[R]. Washington, D. C.: DTIC, 1984.

[6] YU K B, MURROW D J. Combining sidelobe canceller and mainlobe canceller for adaptive monopulse radar processing: US6867726[P]. 2005-03-15.

[7] YU K B, MURROW D J. Adaptive digital beamforming for angle estimation in jamming[J]. IEEE Transactions on Aerospace and Electronic Systems, 2001, 37(2): 508-523.

[8] LI R F, RAO C, DAI L Y, et al. Combining sum-difference and auxiliary beams for adaptive monopulse in jamming[J]. Journal of Systems Engineering and Electronics, 2013, 24(3): 372-381.

[9] 周必雷, 李荣锋, 戴凌燕, 等. 和差四通道及辅助阵元联合自适应单脉冲方法[J]. 系统工程与电子技术, 2017, 39(9): 1905-1914.
ZHOU B L, LI R F, DAI L Y, et al. Adaptive monopulse algorithm combining four-channel sum-difference beam and auxiliary elements[J]. Systems Engineering and Electronics, 2017, 39(9): 1905-1914(in Chinese).

[10] CHEN X Z, SHU T, YU K B, et al. Enhanced ADBF architecture for monopulse angle estimation in multiple jammings[J]. IEEE Antennas and Wireless Propagation Letters, 2017, 16: 2684-2687.

[11] YU K B. Mainlobe cancellation, orthogonal nulling and product patterns[C]// 2016 IEEE International Symposium on Phased Array Systems and Technology (PAST). Piscataway: IEEE Press, 2017: 1-7.

[12] RAO C, LI R F, DAI L Y. Monopulse estimation with multipoint constrained adaptation in mainlobe jamming[C]// Proceedings of 2011 IEEE CIE International Conference on Radar. Piscataway: IEEE Press, 2012: 1054-1057.

[13] 李荣锋, 饶灿, 戴凌燕, 等. 子阵间约束自适应和差单脉冲测角算法[J]. 华中科技大学学报 (自然科学版), 2013, 41(9): 6-10.
LI R F, RAO C, DAI L Y, et al. Algorithm for constrained adaptive sum-difference monopulse among sub-arrays[J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2013, 41(9): 6-10(in Chinese).

[14] 王文涛, 周青松, 刘兴华, 等. Fast ICA盲分离算法在雷达抗主瓣干扰中的应用研究[J]. 现代雷达, 2015, 37(12): 40-44.
WANG W T, ZHOU Q S, LIU X H, et al. A study on radar mainlobe jamming suppression based on blind source separation of fast ICA[J]. Modern Radar, 2015, 37(12): 40-44(in Chinese).

[15] 王建明, 伍光新, 周伟光. 盲源分离在雷达抗主瓣干扰中的应用研究[J]. 现代雷达, 2010, 32(10): 46-49.
WANG J M, WU G X, ZHOU W G. A study on radar mainlobe jamming suppression based on blind source separation algorithm[J]. Modern Radar, 2010, 32(10): 46-49(in Chinese).

[16] CAI T T, ZHANG A R. Compressed sensing and affine rank minimization under restricted isometry[J]. IEEE Transactions on Signal Processing, 2013, 61(13): 3279-3290.

A sparse estimation method for radar target direction with sliding-window subarray configuration in mainlobe jamming

ZHOU Bilei[1,*], LI Rongfeng[2], ZENG Li[3], CHEN Hui[1], LIU Weijian[1], LI Binbin[1]

(1. Air Force Early Warning Academy, Wuhan 430000, China;
2. Aerospace Nanhu Electronic Information Technology Co., Ltd., Jingzhou 434000, China;
3. School of Electronic Information, WuHan University, Wuhan 430000, China)

Abstract: The suppression of mainlobe jamming (MLJ) is a hard task and an open problem in the radar field. This paper considers the problem of radar target direction estimation in MLJ. Hence, we propose a sparse estimation method for radar target direction with sliding-window subarray configuration in MLJ. MLJ is suppressed in each subarray with the proposed method, and the angle atom is constructed according to the phase relationship between the sliding-window subarrays. Finally, the target direction can be estimated with the orthogonal matching pursuit (OMP) algorithm. When the target input signal-to-noise ratio is 20 dB and the angle between the target and MLJ is 1/2 3 dB beam width, the target angle estimation error is smaller than 1/10 3 dB beam width. The accuracy of the angle estimation is well kept and the MLJ can be cancelled without any prior information.

Keywords: mainlobe jamming; sliding-window subarray; sparse recovery; anti-jamming; radar

Received: 2021-09-14; **Accepted**: 2021-10-29; **Published Online**: 2021-11-09 14:32
URL: kns.cnki.net/kcms/detail/11.2625.V.20211109.0904.001.html
Foundation item: National Natural Science Foundation of China (62001510)
* **Corresponding author**. E-mail: zhoubilei666888@sina.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0508

高致密弹携式蜂群布局与多体分离方案

富佳伟*，王辰
(沈阳飞机设计研究所，沈阳 110035)

摘 要：为解决蜂群突防能力弱、作战半径小的弱点，加快形成分布式协同作战的平台级能力，提出一种弹携式蜂群无人机作战系统，利用导弹载体内埋搭载多架小型低成本无人机，通过导弹-蜂群多级运载的战术制定，实现蜂群的快速有利部署。采用兼顾总体相容性的气动布局优化设计技术，综合运用布局选型、数值仿真分析及风洞试验验证，完成一种高致密嵌入式的蜂群布局方案设计，依靠对翼面折叠的小型无人机沿集束柱周向布置的内埋措施，可实现单枚导弹搭载80架小型蜂群无人机的运载能力。针对导弹-蜂群高速多体分离的设计难点，采用优化的分离策略设计，基于改进延迟分离涡模拟方法及重叠网格技术进行蜂群多体分离仿真，结果能够满足蜂群无人机的安全分离。最终的设计方案既保证了蜂群多体量饱和式攻击的作战效能，又实现了系统突防能力和续航能力的显著提升，可满足未来强对抗战场的多种作战使用需求。

关 键 词：蜂群；弹携式；多级运载；总体相容性；布局设计；多体分离
中图分类号：V221.3
文献标志码：A **文章编号**：1001-5965(2023)07-1630-09

传统的以高价值综合化平台为核心的作战体系，其平台技术复杂程度高，研发生产周期长，采购、部署和升级昂贵，为适应日趋复杂的强对抗作战环境，无人机蜂群协同作战模式应运而生。其采用分布式架构，将原有功能系统拆分为若干物理分离、结构相对简单、功能相对单一的小型无人机子系统，具备多体量的节点规模和自由灵活的作战样式，能够以信息互通、任务协同的方式实现低成本作战能力倍增，以资源分布、数据融合的形式实现体系智能涌现。小型无人机蜂群数量大，每个个体不但能够自主作战，而且个体与个体之间还具备组网和协同功能，无论是在情报监视侦察任务中还是攻潜任务中，小型无人机蜂群都有着明显的优势。其可使当今任何防空反导系统的跟踪、探测和拦截能力迅速饱和，还能根据战场态势，自行组织作战计划、自主分配作战任务，作战模式多样，适应高对抗电子战环境，且成本较低，少量小型无人机的损失不会影响整个作战计划和任务，十分适合对地、对舰等攻击任务。因此，无人机蜂群的涌现在降低研制、使用费用与风险的同时，也对未来航空技术的发展产生深远的影响[1-4]。

2018年8月，美国发布《2017—2042财年无人系统综合路线图》，将“蜂群能力”列为无人系统的15项关键技术之一，美国近年来也陆续启动了多个无人机集群作战重点项目，最具代表性的有“小精灵”(Gremlins)、“灰山鹑”(Perdix)微型无人机集群项目、“低成本无人机集群”(LOCUST)项目、拒止环境中协同作战项目等，其中绝大多数项目已推进至飞行演示验证阶段，达到整体领先的水平。除美国之外，英国、欧盟、俄罗斯等国家和地区也纷纷开展了蜂群作战关键技术研发，可见蜂群已成为各大国保持空中作战优势的热点，在未来协同作战体系网络中，无人机蜂群将成为作战力量构成最主要的角色之一[5-8]。

收稿日期：2021-09-01；录用日期：2021-10-11；网络出版时间：2022-02-18 18:30
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220218.1653.001.html
*通信作者. E-mail：fjw15210591734@163.com

引用格式：富佳伟，王辰. 高致密弹携式蜂群布局与多体分离方案[J]. 北京航空航天大学学报，2023，49(7)：1630-1638.
FU J W，WANG C. Configuration and multibody separation scheme of compact missile swarm[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1630-1638 (in Chinese).

但蜂群无人机的尺寸普遍较小, 续航能力存在较大短板, 航程问题始终是蜂群的阿喀琉斯之踵, 大大制约了其远距攻击任务的作战能力。在如何使蜂群快速安全抵达战场的关键性问题上, 各个国家的解决方案较少。美陆军曾在 2016 年寻求“导弹部署集群无人机技术”, 拟利用导弹发射“灰山鹑”这类小型集群无人机, 实现快速突防和饱和攻击的有效结合。美空军研究实验室在《2030 年的空军: 行动号召》视频中展示了集中式载体平台单次投放大量无人机的构想, 这些无人机高密度附着在无人机柱体上, C-130 运输机在货舱中通过伞牵拉投放主体, 随后无人机在空中与柱体分离进行集群作战[9]。

受上述思想启发, 本文设计了一种由载机系统、导弹布撒器系统、蜂群无人机组成的三级运载作战系统, 即弹携式蜂群作战系统。在该系统中, 大量的小型蜂群无人机布置在导弹内部由导弹布撒, 而导弹又由载机系统发射, 巧妙地为小型蜂群无人机提供了长距离初始航程和高突防速度。弹携式蜂群作战系统利用导弹超声速飞行的快速抵近和强突防能力, 实现了蜂群无人机的快速有利部署, 创新性地结合了超声速导弹载体的大作战半径、高纵深突防能力和蜂群的压倒性数量优势, 大幅提高了蜂群存活率、任务实现成功率和作战系统整体作战打击能力。

1　方案概况

1.1　弹携式蜂群系统方案

弹携式蜂群系统架构为载机+导弹+蜂群, 主要由载机系统、导弹系统、蜂群无人机、数据链系统、任务载荷和远程管控系统等组成, 导弹载体平台内埋搭载 80 架小型低成本无人机, 首先由歼击机等空基平台将导弹布撒器携带至防区外的安全区域发射离机后, 以超声速巡航状态快速飞赴目标, 在近目标区域或被拦截之前有效释放蜂群, 蜂群利用其自身装载的电磁、弹药等设备, 通过分布式协同作战, 完成电磁干扰、侦察、攻击等特定任务。弹携式蜂群系统组成如图 1 所示。

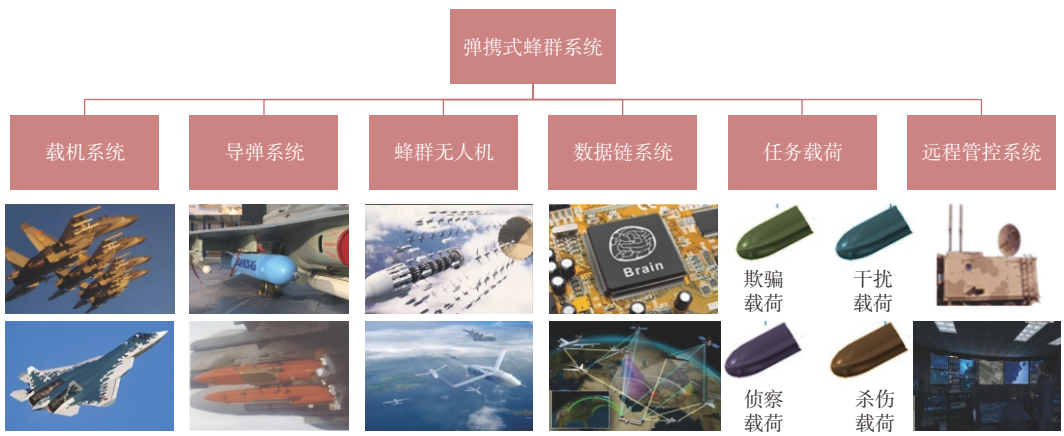


图 1　弹携式蜂群系统组成

Fig. 1　Composition of missile swarm system

1.2　导弹载体方案

导弹类别为超声速巡航导弹, 导弹载体外形与内部布置如图 2 所示, 弹头布置导航、传感、探测等固定有效装载。弹体中部外围沿周向紧密布置 80 个小型蜂群无人机, 为满足几何相容性, 嵌入式无人机翼面采用折叠收放形式, 当触发抛撒蜂群条件时, 中部弹体外罩脱落, 蜂群无人机通过电磁弹射的形式与母弹分离。弹体中部内核布置动力系统, 喷口设置在导弹最尾部。

以高马赫数、大航程、高蜂群装载能力为设计目标, 设计导弹弹重约 2.5×10^3 kg, 弹长约 7 m, 弹径约 1 m。为满足超声速巡航要求, 导弹长细比设计为 9, 弹头与弹翼前缘后掠角为 60°, 有利于减小超声速激波阻力; 弹翼与尾翼均采用周向均布的 X 型, 有利于导弹飞行过程中的稳定性。



图 2　导弹载体外形与内部布置

Fig. 2　Shape and interior layout of missile carrier

1.3　蜂群无人机方案

根据分解的续航时间、作战半径、巡航速度、

单位剩余功率、载重能力等技术指标，设计无人机总体参数[10]，考虑本平台超声速投放特殊性，兼顾机翼收放几何相容性，单个无人机总重设计为5 kg，选定巡航翼载约80 kg/m^2，设计机身长约600 mm，宽约160 mm，机翼面积约0.06 m^2。

单个无人机能根据不同作战分工装载不同的侦察、电子对抗、弹药等模块化设备载荷，微型模块化载荷位于机身头部；机身中部安装飞控、传感、天线、电源等固定装载，尾部安装电力驱动装置和风扇，通过小型无人机动力匹配设计，选用驯龙者3S 10 000 mAh锂电池2组；A2208KV2600电机，XXD6030螺桨，蜂群无人机方案示意图如图3所示。

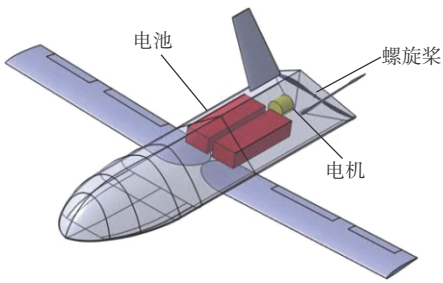
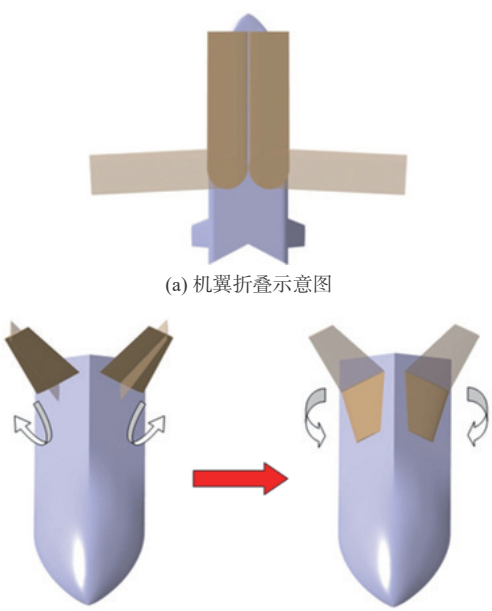


图3 蜂群无人机方案示意图

Fig. 3 Diagram of UAV swarm scheme

为满足导弹内部相容要求，确保携带装载大规模集群平台能力，充分利用导弹载体内部空间，同时应兼顾低阻、良好的工艺性等设计准则，无人机翼面应可折叠收放，采用下单翼机翼和置于机身两侧的全动V尾，折叠方式如图4所示，其中机翼折叠后宽度与机身宽度保持一致，最大化提高机翼面积，减小了翼载，V尾分2步折叠后贴附于机身两侧斜面。



(a) 机翼折叠示意图

(b) V尾折叠示意图

图4 翼面折叠示意图

Fig. 4 Diagram of wing folding

1.4 无人机多体分离方案

由于分离时速度较高，速压较大，对无人机的姿态影响较为严重，如果所有蜂群个体同时分离，强近距气动耦合会给分离安全性带来较大影响；相反，如果进行多步式分离，那么分离后的蜂群无人机较为离散，节点之间距离较远，不利于集群快速部署。综合以上因素，采用图5所示的2步分离策略，每次分离时的蜂群个体均不相邻，既保证了分离的安全性，又能实现短时间快速分离，不影响集群编队。

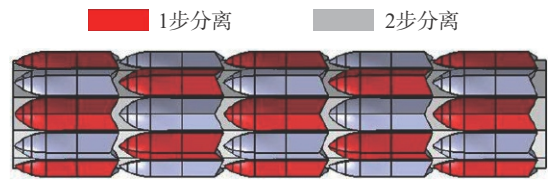


图5 蜂群无人机多体分离方式

Fig. 5 UAV swarm multi-body separation mode

1.5 方案优势

传统的以飞机为载体平台的蜂群作战概念中，飞机接近目标区域将大幅增加平台拦截击落风险，作战代价较大，若安全范围区域抛撒蜂群无人机，则要求无人机具备较高续航能力，且会降低进入目标区域的蜂群成活率；相比传统蜂群概念，超声速导弹具备更高突防能力，敌方探测和拦截难度更大，可实现蜂群无人机对目标区域的近距释放，降低无人机气动设计难度，大幅提升系统平台的任务成功率。

其次，以运输机等平台为载体的蜂群作战系统中，单架飞机造价较高，被敌方探测拦截风险较大，且具有人员伤亡概率；相比之下，导弹平台整体造价更低，具有较高纵深穿透能力，任务过程中高速飞行弹体被拦截风险小，任务失败带来的成本代价较低，且导弹发射后的全任务过程均无人，平台作战方案不会发生人员伤亡，整体风险可控。

再次，弹携式蜂群协同作战系统概念，可实现载体平台超声速飞行状态下的蜂群抛撒释放，与常规蜂群作战使用中的(低)亚声速投放情况相比，本方案无人机具备更大初始分离速度，可进行一定范围的无动力滑翔，从而获得更大航程优势，大幅提升任务打击范围和综合作战效能。图6为常规蜂群作战与弹携式蜂群作战概念对比。

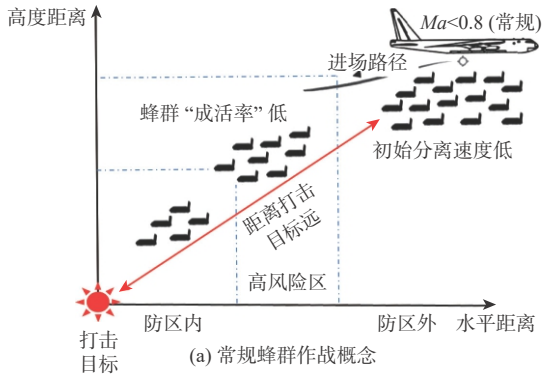


图 6　蜂群作战概念对比示意图

Fig. 6　Comparative schematic of swarm operation conception

2　蜂群无人机布局方案

2.1　气动布局设计

气动布局设计以任务指标为目标, 以总体参数为牵引, 根据平台应满足的各项任务需求, 明确了兼顾超声速无动力滑翔、低速巡航、翼面可折叠收放等能力以弹体内几何相容等约束条件下的小型无人机的关键特征。

分布式作战小型无人机平台基于低速巡航、超声速无动力滑翔、模块化装载的需要, 综合考虑导弹平台内部填充和弹射的约束, 其气动布局采用基于电力驱动的大展弦比可变后掠可折叠平直翼, 小钝头前缘升力体式机身和可折叠全动 V 尾的气动布局形式, 满足无人机的填充弹射、载荷及续航的要求。表 1 为蜂群无人机气动布局方案和设计依据。

1) 机翼设计。兼顾升阻特性、折叠要求、结构强度和重量限制, 机翼选择可折叠平直翼, 后掠角可变以适应高速滑翔和低速平飞的不同应用场景, 巡航状态下机翼后掠角为 3°。增升装置选用简单后缘襟翼, 副翼位于机翼后缘靠近翼梢处。

2) 机身设计。机身长度由导弹内部空间尺寸、蜂群数量指标、无人机内部装载、无人机总重等综合限制因素决定。考虑导弹内部空间限制, 机身为小钝头升力体式机身, 下表面为平面, 以满足嵌入式布局和机翼折叠几何相容性, 上表面采用气动外形优化设计。为保证导弹内部装载数量, 满足纵向嵌入布置, 机身尾部设计成燕尾式, 用于给螺旋桨提供布置空间, 充分提高导弹内部空间利用率。

表 1　气动布局方案和设计依据

Table 1　Aerodynamic configuration scheme and design basis

需求约束	设计依据	布局特征
巡航要求	保证50～70 m/s低速飞行续航时间1 h, 提升设计点最大升阻比, 增升减阻	大展弦比平直机翼 较薄的低速高升力翼型
超声速无动力滑翔	不损失低速巡航能力前提下, 有效抑制滑翔阶段激波阻力	小钝头前机身 中等后掠机翼
模块化装载要求	能根据不同作战分工装载不同侦察、传感、电子对抗、弹药等模块化设备载荷	机身头部模块化 侦察、打击设备 尾部推进装置
高密度填充要求	保证内部模块化装载同时, 提高填充空间利用率	折叠下单翼 折叠全动V尾
翼面折叠要求	翼面蜂巢填充时应折叠收放便于高效利用空间	

3) 机翼机身相对位置选型设计。机翼折叠方式可以有向后折叠(机翼前置)和向前折叠(机翼后置)2 种。采用二维面元法对比 2 种方案纵向力矩特性, 机翼后置静不安定度在允许范围内, 重焦匹配性更好, 因此, 选择后置机翼向前折叠方案为最终布局方案。

4) 通过迭代优化设计。最终蜂群无人机三面图如图 7 所示, 蜂群无人机采用大展弦比中等后掠可折叠平直翼, 小钝头前缘升力体式机身和可折叠全动 V 尾的气动布局形式, 控制舵面为机翼后缘内侧襟翼、外侧副翼及全动 V 尾。

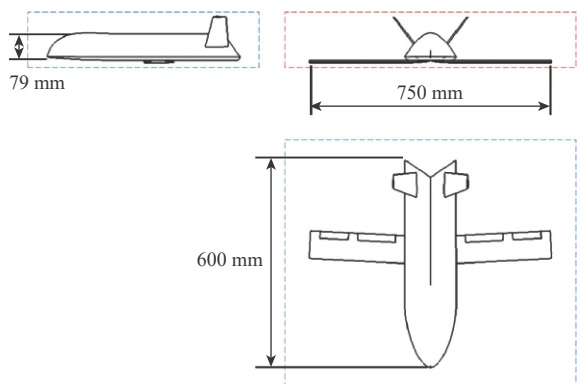


图 7　蜂群无人机三面图

Fig. 7　Three sides diagram of UAV swarm

2.2　气动特性分析验证

针对蜂群无人机低速巡航模态(3°后掠角构型), 通过气动数值仿真和低速风洞试验验证, 对蜂群无人机纵向与横航向低速基本气动特性与舵面偏转效率(内侧襟翼、外侧副翼、全动 V 尾)进行了气动特性分析。

数值仿真条件为马赫数 Ma 为 0.2, 高度 H 为 3 km。

由于低速空气压缩性可以不予考虑，空气设置为不可压恒密度，使用二阶精度 SIMPLE 算法（压力耦合方程组的半隐式方法）分别独立计算流场的压力和速度，具有较高的计算效率和精确度。

风洞试验采用 1:1 等比模型尾部支撑试验，试验风速约 35 m/s，试验模型在风洞中安装情况如图 8 所示。

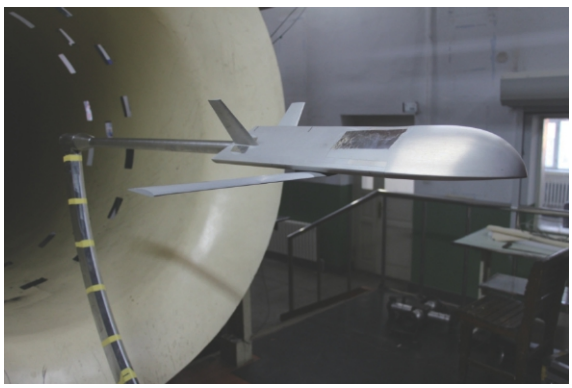
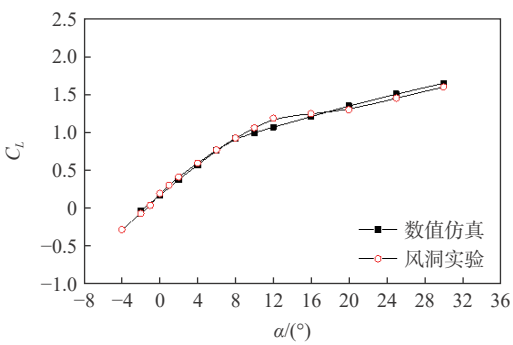


图 8　试验模型安装情况

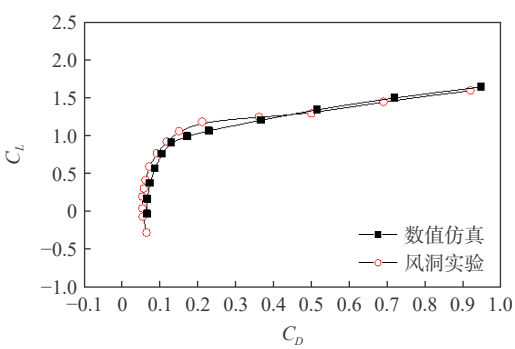
Fig. 8　Installation of test model

图 9～图 11 分别为升阻特性曲线及舵面偏转引起的纵向、横航向气动力增量曲线，α 为迎角，C_L 为升力系数、C_D 为阻力系数、C_m 为俯仰力矩系数、C_l 为滚转力矩系数、C_n 为偏航力矩系数，增量用 Δ 表示，各系数方向采用美标。

由图 9 可知，升力系数和阻力系数的仿真与试



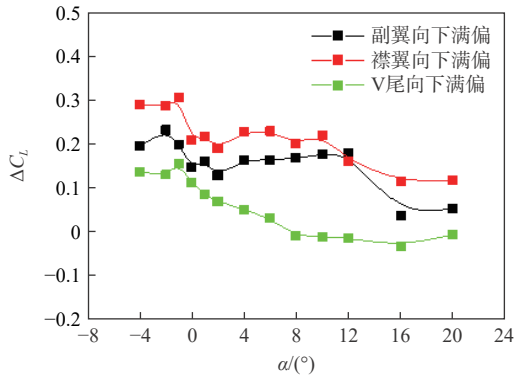
(a) 升力系数



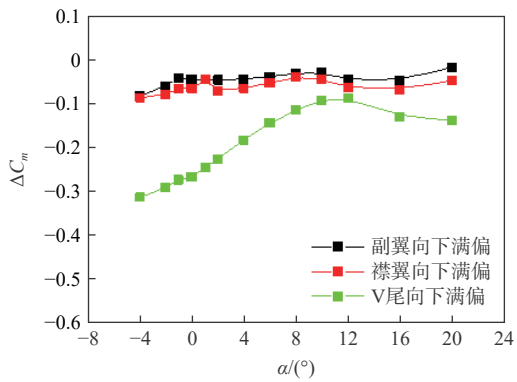
(b) 升阻比极曲线

图 9　升阻特性曲线

Fig. 9　Lift and drag characteristic curves



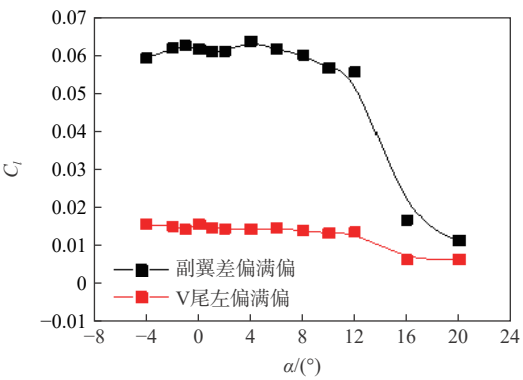
(a) 升力系数增量



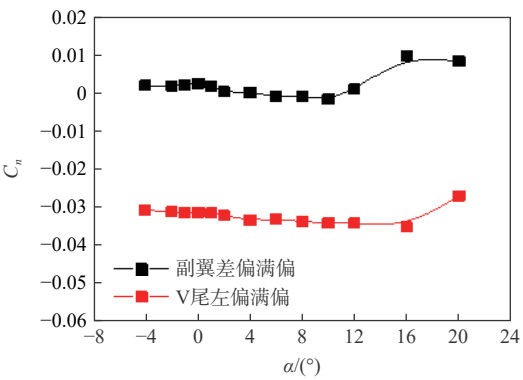
(b) 俯仰力矩系数增量

图 10　舵面偏转引起纵向气动力增量曲线

Fig. 10　Longitudinal aerodynamic increment curves by deflection



(a) 滚转力矩系数



(b) 偏航力矩系数

图 11　舵面偏转引起横航向气动力增量曲线

Fig. 11　Lateral aerodynamic increment curves by deflection

验结果基本一致, 在中、大迎角下有小幅偏差, 经分析可能是由于雷诺数条件不同而导致不同分离程度引起的, 该结果在可接受范围内, 可认为通过数值仿真结合风洞试验的气动特性分析手段合理可靠。

由图 10 可知, 襟翼下偏增升能力最大, 且随迎角增大有所减弱。V 尾上下偏转俯仰控制能力最大, 且在小迎角下, 明显强于襟翼、副翼俯仰控制能力。

由图 11 可知, 副翼差偏的滚装控制能力大于 V 尾左右差偏, 且随迎角增大逐渐减小, 但不会发生反效。V 尾差偏的偏航控制能力大于副翼差偏。

综合上述结果, 蜂群无人机平台气动特性总结如下:

1) 襟翼下偏会带来可观的升力增量, 副翼差偏滚转操纵能力较好, 但二者对飞行器的俯仰和偏航影响不大。

2) V 尾偏转的俯仰、偏航控制能力较强, V 尾上下偏转用于控制无人机俯仰, 左右差偏用于控制无人机偏航。

3) 可以根据飞行任务需求, 对操纵面进行组合使用。襟翼、副翼组合可以提供更强的增升能力; 当需要较大滚转力矩时, 副翼和 V 尾可以通过组合偏转来实现。

3　蜂群无人机多体分离设计

3.1　研究方法

无人机的多体分离问题是动边界绕流问题, 分离体间存在强非定常、非线性干扰现象, 其过程涉及空气动力学、刚体动力学、飞行力学、碰撞动力学等问题[11]。这种耦合干扰不但对分离体气动性能和飞行安全性有重要影响, 而且将决定作战任务的完成质量, 弹携式蜂群分离马赫数高、数量大、干扰强, 给分离安全和作战任务的实现带来了难题和挑战, 拟采用气动/运动耦合的空射分离仿真评估方法研究该问题[12-15]。

无人机多体分离计算模型如图 12 所示, 坐标原点位于弹头锥底部, 蜂群无人机由前到后依次编号, 由于采用 1.4 节的 2 步分离策略, 拟对奇数排进行分离动态模拟仿真。采用的坐标系为弹体坐标系, x 轴向后, y 轴向上, z 轴向左, 部件坐标系原点在蜂群无人机重心, 随蜂群无人机一起运动, 坐标轴平行于导弹弹体坐标系的各坐标轴, 蜂群无人机的姿态用欧拉角表示, 俯仰角 θ 低头为正。

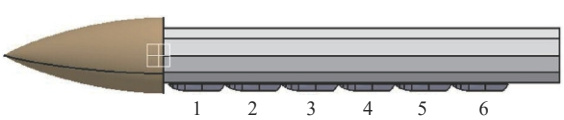


图 12　无人机多体分离计算模型示意图

Fig. 12　Diagram of UAV multibody seperation computational model

数值计算方法选用改进延迟分离涡模拟方法, 相比于最初版本的分离涡模拟 97 及延迟分离涡模拟方法, 改进延迟分离涡模拟方法加入了壁面模化的大涡模拟分支, 在保留了分离涡模拟类方法优点的同时成功克服了对数律不匹配、网格诱导分离、模型应力损耗等问题。在涡黏假设下, 涡黏系数 μ_t需要利用湍流模型进行封闭, 湍流模型选用 Menter k-ωSST 两方程模型。

在非定常计算中使用双重时间步法, 即在控制方程中引入虚拟时间项, 根据精度设定物理时间步求解真实解, 而每一物理时间步内, 通过虚拟时间内迭代达到收敛, 内迭代步应用多重网格技术加速收敛。通量项采用二阶精度 Roe 格式, 选用修正的 Venkatakrishnan 限制器保证二阶精度差值且具有 TVD 性质, 同时又有较小的数值耗散。

计算模型将弹体和固定的蜂群无人机作为背景网格区域, 3 个子区域对应 3 个分离的蜂群无人机, 图 13 为计算模型嵌套示意图, 其中黄色部分为活跃网格, 蓝色部分为接收网格, 红色部分为非活跃网格。背景区域网格单元共计 8.47×10^6 个, 3 个子区域网格单元个数分别为 29 427 个。

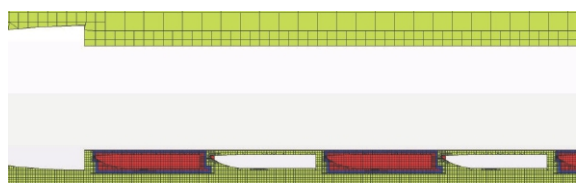


图 13　计算模型嵌套网格

Fig. 13　Computational mode nested grid

3.2　多体分离流动特性分析

给定表 2 所示的分离初始条件, 关注前排 1, 3, 5 号无人机运动轨迹和空间流场如图 14 和图 15 所示, 时间步长为 0.015 s, 即分离轨迹图和压力云图对应的时间点分别为 t=0.015 s、t=0.030 s、t=0.045 s、t=0.060 s。

可知, 1 号无人机在分离初始时, 驻点高压区位于前缘下表面而使平台产生向上的抬头力矩, 加之其自身纵向不安定导致上仰, 1 号无人机很快发生

表 2　分离初始条件

Table 2　Separation of initial conditions

Ma	H/km	初始速度/(m·s^{-1})	初始角速度/(rad·s^{-1})		
			1号	3号	5号
1.2	6	5	0	0	0

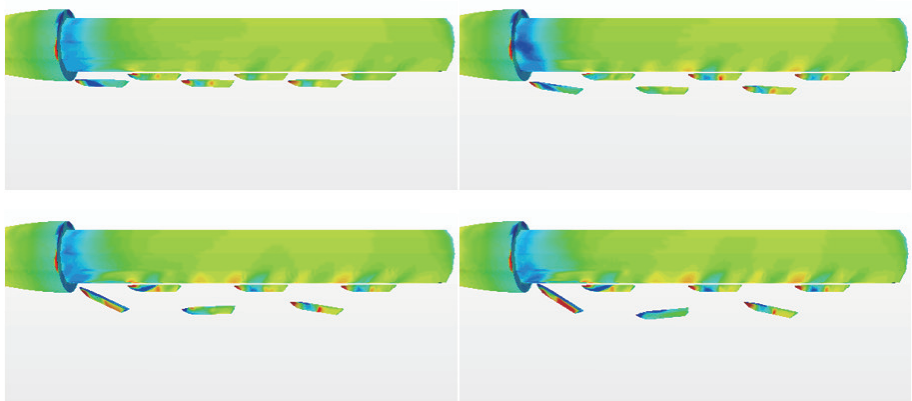


图 14 蜂群无人机分离轨迹图

Fig. 14 UAV swarm separation trajectory map

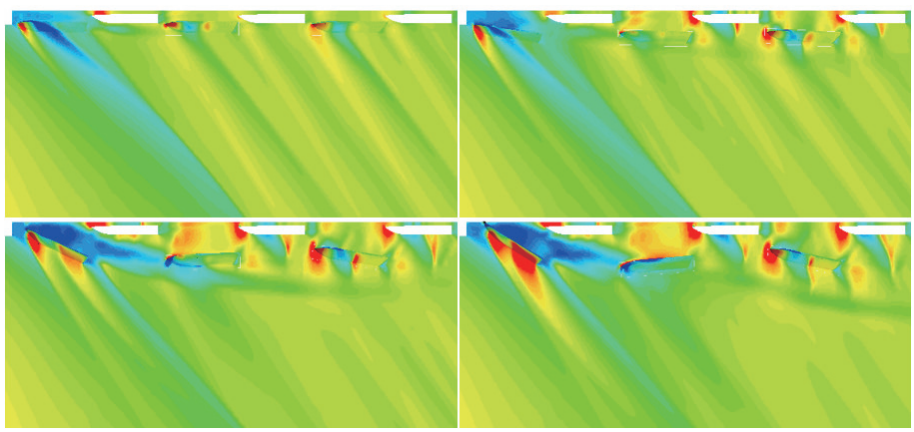


图 15 空间压力云图

Fig. 15 Space pressure contours

了头部碰撞。3 号、5 号无人机受到的尾流扰动较为复杂，其运动规律难以捕捉，但对分离初始短时间内的运动特性分析可知，5 号无人机同样有明显的力矩上仰问题，存在碰撞风险。

3.3 多体分离参数优化设计

控制分离马赫数 *Ma* 为 1.2 不变，通过参数化方法，设计最优分离策略，以满足分离安全性的要求，分离初始速度、初始角速度设计如表 3 所示。赋予了无人机初始俯仰角速度以抑制无人机抬头，

表 3 优化后初始条件设计值

Table 3 Design value of initial conditions after optimization

Ma	*H*/km	初始速度/(m·s^{-1})	初始角速度/(rad·s^{-1})		
			1号	3号	5号
1.2	6	9	12	6	6

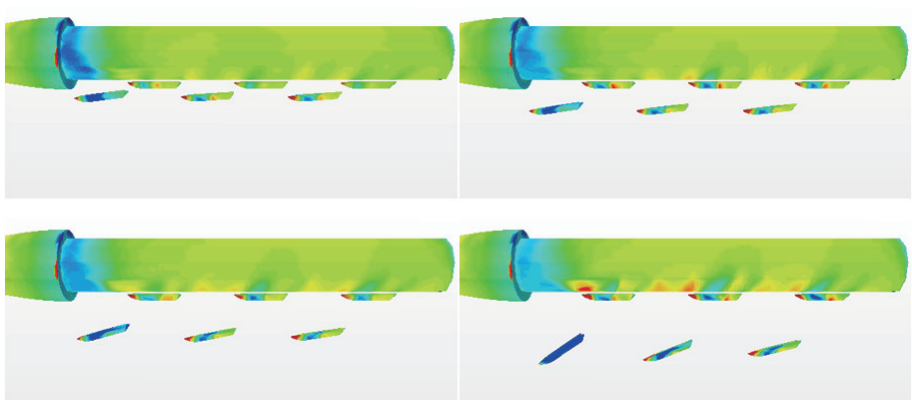


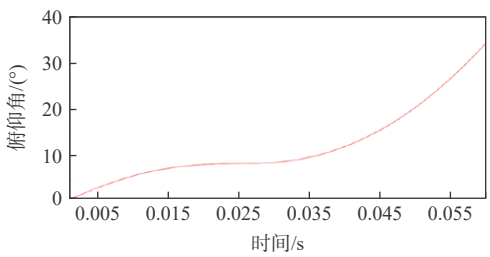
图 16 优化后无人机分离轨迹图

Fig. 16 UAV swarm separation trajectory map after optimization

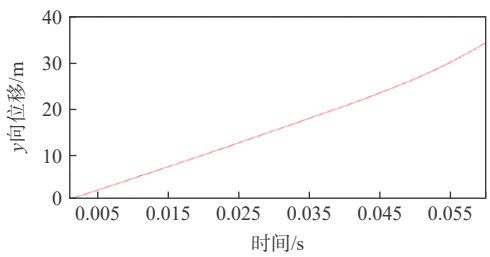
1 号无人机抬头最明显, 初始角速度设置最大为 12 rad/s, 同时增加了向下分离初始速度, 提高分离安全裕度。

图 16 为优化后蜂群无人机分离轨迹图。可知采用优化分离策略, 无人机分离后未出现上浮和抬头, 且后排无人机受前排无人机洗流影响较为稳定, 运动轨迹与姿态趋于一致, 整个系统具备较高的分离安全性。

由图 16 可知, 蜂群分离过程的主要运动为纵向运动, 而侧向偏航和滚转可忽略, 以 1 号无人机为例, 捕捉分离无人机 y 向位移和俯仰角随时间变化, 以反映系统分离运动特性规律。图 17 为优化后 1 号无人机分离俯仰角和 y 向位移变化曲线。



(a) 俯仰角



(b) y向位移

图 17　优化后 1 号无人机分离俯仰角和 y 向位移变化曲线

Fig. 17　Optimized NO.1 UAV sepration pitch angle and y-directim displacement variation

可知优化后, 位移曲线近似线性, 线性度较好, 俯仰角变化曲线均呈现双拐点三次曲线的特征, 俯仰角有发散的趋势, 因此, 需要在具备一定安全距离, 且俯仰角尚未充分发散前, 引入 V 尾单通道飞控, 对无人机俯仰姿态进行增稳控制, 进一步保证蜂群安全分离部署。

4　结　论

1) 采用高致密相容的弹携式蜂群设计, 通过紧密嵌入周向布置, 实现了单枚导弹载体可搭载 80 架小型蜂群无人机, 相比于传统蜂群作战, 弹携式蜂群具备更高突防能力, 可实现蜂群无人机对目标区域的近距释放, 提升了整体作战半径及任务成功率, 降低了成本和风险。

2) 方案兼顾了几何相容约束和性能指标要求, 形成了基于电力驱动的大展弦比可变后掠向前折叠平直翼, 小钝头前缘升力体式机身和可折叠全动 V 尾的气动布局形式, 满足无人机的填充弹射、载荷及续航的要求。内侧襟翼用于增升, 外侧副翼用于滚转控制, V 尾上下偏转可用于控制无人机俯仰, 左右差偏可用于控制无人机偏航。可以根据飞行任务需求, 对操纵面进行组合使用。襟翼、副翼组合可以提供更强的增升能力; 当需要较大滚转力矩时, 襟翼、副翼和 V 尾可以通过组合偏转来实现。

3) 通过参数优化设计方法得到的优化分离策略下, 分离无人机位移曲线近似线性, 俯仰角变化曲线呈现双拐点三次曲线的特征, 俯仰角有发散的趋势, 需要在具备一定安全距离, 且俯仰角尚未充分发散前, 引入 V 尾单通道飞控, 对无人机俯仰姿态进行增稳控制。

致谢 感谢沈阳飞机设计研究所对该创新项目的支持, 感谢浙江大学黎军教授的点拨和指导。

参考文献（References）

[1] 蒋琪, 申超, 张冬青. 认知/动态与分布式作战对导弹武器装备发展影响研究[J]. 战术导弹技术, 2016(3): 1-6.
JIANG Q, SHEN C, ZHANG D Q. Research on the influence of cognitive/dynamic and distributed operations on the development of missile weapons and equipment[J]. Tactical Missile Technology, 2016(3): 1-6(in Chinese).

[2] 时东飞, 蔡疆, 黄松华, 等. 美国空军“战斗云”作战理念及启示[J]. 指挥信息系统与技术, 2017, 8(3): 27-32.
SHI D F, CAI J, HUANG S H, et al. Operational concept and enlightenment of United States air force “combat cloud” [J]. Command Information System and Technology, 2017, 8(3): 27-32(in Chinese).

[3] 王彤, 李磊, 蒋琪. “进攻性蜂群使能战术”项目推进无人蜂群能力发展分析[J]. 战术导弹技术, 2020(1): 33-38.
WANG T, LI L, JIANG Q. Offensive swarm-enabled tactics program promotes the development of unmanned swarm capability[J]. Tactical Missile Technology, 2020(1): 33-38(in Chinese).

[4] 贾永楠, 田似营, 李擎. 无人机集群研究进展综述[J]. 航空学报. 2020, 41(S1): 723738.
JIA Y N, TIAN S Y, LI Q. Developemnt of unnamed aerial vehicle swarms[J]. Acta Aeronautica et Astronautica Sinica. 2020, 41(S1): 723738 (in Chinese).

[5] HURST J. Robotic swarms in offensive maneuver[J]. Joint Force Quarterly, 2017(87): 105-111.

[6] FISHER N , GILBERT G R. Unmanned systems in support of future medical operations in dense urban environments[J]. Journal of Article, 2016, 4(14): 48.

[7] SADRAEY M H. Manned-unmanned aircraft teaming[C]//2018 IEEE Aerospace Conference. Piscataway: IEEE Press, 2018: 1-12.

[8] United States of America Defence Science Board. Summer Research Report of Defence Science Board—Autonomous Technology: 20301-3140 [R]. Washington D. C : United States of America

Defence Science Board, 2016

[9] SATHE A, PANT R. Conceptual design studies of an unmanned combat aerial vehicle[C]// AIAA Aviation Technology, Integration, and Operations. Reston: AIAA, 2013.

[10] 宋威, 艾邦成. 多体分离动力学研究进展[J]. 航空学报, 2022, 43(9): 25590.
SONG W, AI B C. Multibody separation dynamics: Review[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(9): 25590(in Chinese).

[11] JEYAKUMAR D, RAO B N. Dynamics of satellite separation system[J]. Journal of Sound and Vibration, 2006, 297(1-2): 444-455.

[12] 唐上钦, 黄长强, 翁兴伟. 考虑气动干扰的导弹内埋式发射弹道研究[J]. 弹箭与制导学报, 2013, 33(3): 138-142.
TANG S Q, HUANG C Q, WENG X W. The study on trajectory of missile separating from cavity with aerodynamic interference considered[J]. Journal of Projectiles, Rockets, Missiles and Guidance, 2013, 33(3): 138-142(in Chinese).

[13] WANG Y J, WANG H B, QIAN F X, et al. Investigation on separation interference of typical multi-body vehicle in supersonic conditions[C]// 2013 Fourth International Conference on Digital Manufacturing & Automation. Piscataway: IEEE Press, 2013: 732-735.

[14] 张群峰, 闫盼盼, 黎军. 内埋式弹舱与弹体相互影响的精细模拟[J]. 兵工学报, 2016, 37(12): 2366-2376.
ZHANG Q F, YAN P P, LI J. Elaborate simulation of interaction effect between internal weapon bay and missile[J]. Acta Armamentarii, 2016, 37(12): 2366-2376(in Chinese).

[15] 黄蓓, 王浩, 王帅, 等. 薄片状体沉降过程中的多体干扰流场特性[J]. 弹道学报, 2012, 24(1): 41-46.
HUANG B, WANG H, WANG S, et al. Flow field characteristics of multi-plates interference in descent[J]. Journal of Ballistics, 2012, 24(1): 41-46(in Chinese).

Configuration and multibody separation scheme of compact missile swarm

FU Jiawei*, WANG Chen

(Shenyang Aircraft Design & Research Institute, Shenyang 110035, China)

Abstract: One type of missile swarm system is proposed, where numerous small UAVs are compressed in a missile carrier, to solve the weakness of swarm penetration capability and operational radius while speeding up the formation of cooperation. It uses tactics by multistage transport, making the swarm deploy rapidly. Taking advantage of the compatibility design, aerodynamic configuration optimization design technology, layout selection, numerical simulation analysis, and wind tunnel test, the research achieves a kind of compact missile swarm layout design, one missile can carry 80 small UAVs by arranging UAVs circumferentially around the cylinder. In order to overcome the challenge of multibody separation design, a swarm can safely separate by adopting an optimum separation approach that has been confirmed by improved delayed detached eddy simulation and nested grid. Finally, the design scheme can not only ensure the efficiency of saturation attack but also improve the capability of penetration and endurance, meeting the requirement of future strong confrontation battlefield.

Keywords: swarm; missile carrier; multistage transport; compatibility; configuration design; multibody separation

Received: 2021-09-01; **Accepted**: 2021-10-11; **Published Online**: 2022-02-18 18:30

URL: kns.cnki.net/kcms/detail/11.2625.V.20220218.1653.001.html

* **Corresponding author**. E-mail: fjw15210591734@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0566

高超声速再入飞行器气动布局多目标优化

樊博璇[1,2,*]，陈桂明[1]，曹逸群[2]
(1. 火箭军工程大学，西安 710025；2. 火箭军装备部驻西安地区第一军事代表室，西安 710025)

摘 要：气动布局的多目标优化是飞行器设计中的关键技术。提出一种新的高超声速再入飞行器气动外形参数的多目标优化方法，证明外形优化对高超声速流下飞行器性能的影响。通过实例仿真对飞行器所受阻力和升力对制导性能影响进行详细验证分析，将飞行器落点圆概率偏差、末速大于500 m/s的占比、最大飞行过载小于60g的占比这3个性能指标作为优化目标，将升力特性作为中间参数，将气动布局优化问题分解为2个子问题，通过基于搜索算法的升力特性优化和基于改进的模拟退火算法的外形参数优化，减少优化计算时间、提升计算效率、实现对飞行器主体和襟翼的气动布局优化、获得高超声速流下的最佳飞行器外形。仿真结果表明：在确定的约束条件下，优化算法增加了飞行器在超音速流下的气动升力，有效提高了升阻比。在不影响最大飞行过载的前提下，优化后的飞行器表现出更高的气动性能，显著提升了命中精度，同时末速也满足指标要求，制导系统性能得到有效改善。

关 键 词：高超声速再入飞行器；气动布局；升力特性；模拟退火组合优化算法；优化设计
中图分类号：V221$^+$.3; TB553
文献标志码：A **文章编号**：1001-5965(2023)07-1639-12

高超声速再入飞行器具有飞行速度快、航程远、机动能力强等特点，通过速度、射程和机动性等要素衡量其飞行性能，影响射程和速度的主要因素有环境条件、推进系统性能及空气动力学特性，具有很强的战略性和带动性。研究改变前2种因素的运营成本很高，先进空气动力学是影响高超声速飞行器设计水平的核心因素，气动布局设计的主要目标是获得大升力系数和小阻力系数，即高升阻比。这取决于机头和机身的形状和大小，以及襟翼的形状、大小和位置。随着飞行马赫数的增加，存在一定的升阻比屏障，提高升阻比较为困难[1-2]。目前针对这一问题，利用先进的计算机技术，使得在合理的时间内利用数值求解技术解决复杂几何形状的流体流动问题成为可能[3-4]。

近年来国内外有不少在理论、仿真、工程方面的研究进展[5-9]。文献[10]完成了由优化和验证步骤组成的两阶段研究，以改善导弹的气动特性。利用实验数据和导弹数据通信结果获得了最佳几何结构，并利用ANSYS-FLUENT模型验证了优化配置。文献[11]获得了具有通用形状机头的AGARD-B模型的计算流体力学(computational fluid dynamics, CFD)模拟结果，并使用从具有相同结构的模型获得的实验数据对结果进行了验证，然后引入仿真模型以获得非通用机头结构的数值预测。文献[12]研究了尾翼的流动控制，以产生不对称的压力分布，并提供改变尾翼区域流场的气动控制。文献[13]对一种导弹进行了比较研究，旨在改进导弹前部结构，以改善其空气动力学特性。结果表明，实验得到的气动载荷符合CFD模拟的气动系数。文献[14]使用自组织的间接优化(indirect optimization based

收稿日期：2021-09-23；**录用日期**：2022-01-05；**网络出版时间**：2022-02-14 16:13
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220214.1529.007.html
基金项目：国家自然科学基金(71601180)
*通信作者. E-mail：1092442646@qq.com

引用格式：樊博璇，陈桂明，曹逸群. 高超声速再入飞行器气动布局多目标优化[J]. 北京航空航天大学学报，2023，49(7)：1639-1650. FAN B X，CHEN G M，CAO Y Q. Multi-objective optimization of aerodynamic layout for hypersonic reentry vehicle[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1639-1650 (in Chinese).

upon self-organization, IOSO)算法,研究了减少旋转体在超音速下的气动阻力,阻力系数和体积分别作为目标函数和约束条件。结果表明,弹体的气动阻力减小了20%左右。文献[15]专注于N1G和AGARD-B模型在不同迎角和不同马赫数下的外部设计。为了预测升力和阻力系数,进行了CFD模型和实验研究。文献[16]提出了无人机机翼气动优化研究,预测阻力系数和升力系数,并将结果与ANSY-FLUENT进行比较。文献[17]利用遗传算法解决导弹的形状优化问题,通过AIRDOSN例行程序和导弹DATCOM软件对气动系数进行预测。文献[18]使用了非支配排序遗传算法和实数编码遗传算法研究了气动外形优化课题。类似地,文献[19]将实数编码自适应范围遗传算法与轨迹分析相结合,通过最大化射程来修改导弹的尾翼。文献[20]使用遗传算法对NREL S809翼型进行形状优化,以改善升力和阻力特性。

本文选取某型高超声速再入飞行器为研究对象,验证气动系数和进行气动外形优化。通过仿真分析,升力系数被选为优化目标,选取多个关键尺寸的可优化设计变量,利用ANSYS中的design modeler和mesh建立飞行器几何实体模型和进行网格生成,并进行多参数下CFD气动特性求解和飞行器特性计算式拟合,对高超声速马赫数下的气动系数进行了预测和验证。通过采用升力特性优化和基于一种改进模拟退火算法的外形参数向量优化2个步骤,执行优化过程,实现了飞行器整体气动布局优化。优化结果表明:与基线模型相比,升力和阻力系数得到了令人满意的改善,该优化方法能够提升高超声速飞行器的升阻比和实用价值,为未来高超声速飞行器气动布局设计提供重要参考。

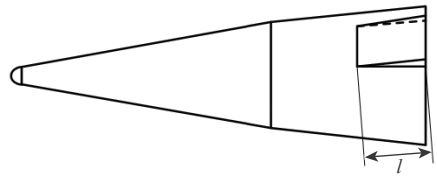
1 待优化飞行器模型

1.1 模型描述

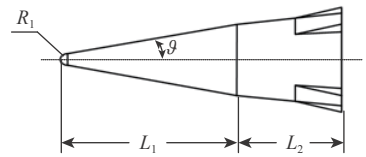
选定的飞行器模型由机体和襟翼2个主要部分组成,襟翼对飞行器的气动特性有显著影响。飞行器各特征尺寸如图1所示。图中:ϑ为第1锥角,R_1为端头半径,L_1为第1锥体长度,L_2为第2锥体长度,L_3为底部直径。体襟翼的形状包括长度l、厚度h和宽度,其中襟翼宽度可与襟翼对应的飞行器尾部圆面圆心的张角δ(襟翼张角)一一对应,即把襟翼宽度的变化转换为圆心张角的变化。

1.2 可优化设计变量确定

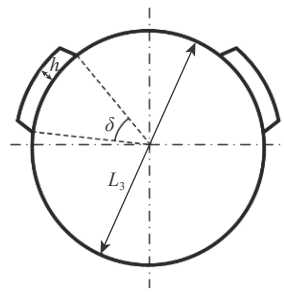
气动外形优化是飞行器的一项重要任务。就空气动力学性能而言,飞行器的所有外部部件都很重要。



(a) 主视图



(b) 俯视图



(c) 左视图

图1 高超声速再入飞行器外形三视图

Fig. 1 Three views of hypersonic reentry vehicle

实现最优解的重要任务是选择设计变量。通过理论分析发现,对于选定的基础气动外形设计方案,对气动力影响较大的为第1锥角ϑ和体襟翼的形状,后者包括体襟翼长度l、厚度h及与襟翼对应的飞行器尾部圆面圆心的张角δ。因此,拟对这4个参数进行优化,以找到满足飞行器最优气动特性的尺寸和形状。设定飞行器气动外形尺寸如表1所示。

表1 飞行器气动外形尺寸

Table 1 Aircraft aerodynamic dimensions

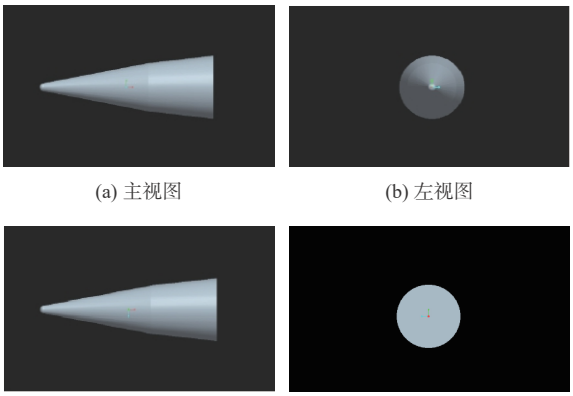
飞行器编号	ϑ/(°)	l/mm	δ/(°)	h/mm
1	7.5			
2	7.5	400	45	60
3	7.5	400	45	85
4	7.5	400	60	60
5	7.5	400	60	85
6	7.5	600	45	60
7	7.5	600	45	85
8	7.5	600	60	60
9	7.5	600	60	85
10-18	8.5			
19-27	9.5			

2 飞行器经验计算式拟合

2.1 几何建模和网格生成

通过几何建模和网格划分等步骤进行CFD前处理。分别利用ANSYS中的design modeler和mesh

建立了飞行器几何实体模型并进行网格生成。对设定的可优化参数进行排列组合, 建立飞行器无襟翼状态及飞行器在所设定不同气动外形尺寸下的几何模型, 如图 2 和图 3 所示。

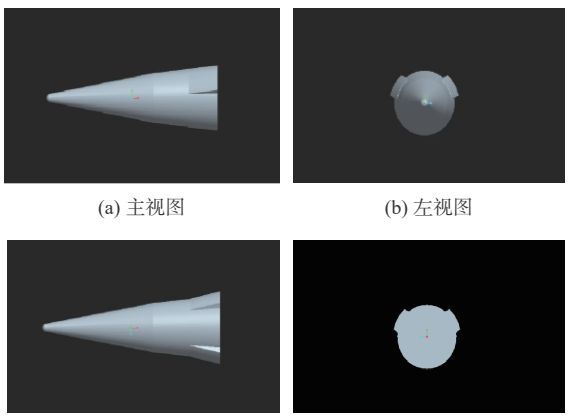


(a) 主视图　　(b) 左视图

(c) 俯视图　　(d) 右视图

图 2　飞行器无襟翼状态几何模型

Fig. 2　Geometric model of aircraft without flap



(a) 主视图　　(b) 左视图

(c) 俯视图　　(d) 右视图

图 3　飞行器几何模型

Fig. 3　Aircraft geometric model

由于双襟翼布局和迎角的影响, 不适合采用周期性网格, 需进行三维网格划分。利用 ANSYS 中的网格环境生成非结构混合网格。由于 CFD 分析是在超音速流场中进行的, 因此, 飞行器的网格生成是建立在捕捉激波的基础上的。为了解决边界层问题, 将计算域模拟成一个圆柱体, 从飞行器顶端向上游有 8 个机身长度, 从模型下游大约有 16 个机身长度, 圆柱体的半径是 8 个机身长度, 在计算区域的剩余部分生成四面体网格单元。

为保证 CFD 计算结果的准确性和网格的独立性, 以 200 万个网格数量作为划分网格时的数目标准。为保证结果的精确度, 在飞行器弹身两锥体由于斜率不同的转折处、飞行器头部半球体与弹身锥体连接处、飞行器尾部界面与弹身侧面连接处等对网格进行加密处理。网格长宽比过大会影响后续 CFD 计算的精度, 通过改变网孔的尺寸和生长速率, 保证加密后网格部分的长宽比也较小; 网格设置选择了接近度和曲率来捕捉弹体周围的流场。在对每种情况和更细的网格单元进行划分后, 观察到 2 008 827 个单元, 足以获得与实验数据很好的一致性, 得到的网格质量要优于对整个飞行器流场进行网格划分的网格质量, 同时保证了计算的气动系数和残差的收敛性。

2.2　气动特性计算

认为飞行器的气动参数是飞行器无襟翼状态下的气动参数与飞行器襟翼气动外形变化部分的叠加, 通过流场求解, 得到飞行器和飞行器无襟翼状态的气动力参数信息。

2.2.1　无襟翼状态

飞行器无襟翼状态时具有轴对称外形特征, 在选取其气动特性计算状态时, 可以只研究迎角的变化对飞行器气动特性的影响, 侧滑通道中侧滑角变化对飞行器气动特性的影响作用可以通过类比俯仰通道内迎角变化的影响规律, 故飞行器侧滑角全部取值为零。结合由飞行器制导指令得到的不同飞行高度下马赫数的大致变化范围, 设定飞行器无襟翼下气动特性状态如表 2 所示。

表 2　计算选取的飞行器无襟翼下气动特性状态

Table 2　Aerodynamic characteristic states of aircraft without flap selected for calculation

高度H/km	马赫数Ma	迎角α /(°)
80	29,28,27,26	10,9,8,7,5,4,3,2,1,0
70	28,27,26,25	10,9,8,7,5,4,3,2,1,0
60	27,26,25,24,23	10,9,8,7,5,4,3,2,1,0
50	26,25,24,23,22,21	10,9,8,7,5,4,3,2,1,0
40	26,25,24,23,22,21	10,9,8,7,5,4,3,2,1,0
30	26,25,24,23,22,21	10,9,8,7,5,4,3,2,1,0
25	23,21,20,18,16	13,12,10,9,8,7,5,4,3,2,1,0
20	20,19,18,17,16,15	13,12,10,9,8,7,5,4,3,2,1,0
15	16,14,13,12,10	13,12,10,9,8,7,5,4,3,2,1,0
10	14,13,12,11,10	13,12,10,9,8,7,5,4,3,2,1,0
5	8,7,6,5,4	16,15,13,12,10,9,8,7,5,4,3,2,1,0
0	5,4,3,2	20,18,16,15,13,12,10,9,8,7,5,4,2,0

无襟翼下飞行器需计算的飞行状态为 688 组, 3 种不同气动外形的无襟翼飞行器, 所需计算的飞行状态共 2 064 组。

2.2.2　带襟翼状态

高超声速再入飞行器的气动外形不是轴对称的, 而是面对称的, 在进行气动状态计算选取时需要同时考虑 2 个通道的情形。在俯仰通道内, 迎角绝对值相同而取值的正负不同会造成飞行器气动参数的绝对值不同; 在偏航通道内, 侧滑角绝对值相同时取值的正负则不会影响飞行器气动参数绝对值的变化。飞行器气动特性计算结果如表 3 所示。

由表3可知，每种飞行器的气动外形下需要计算的飞行状态为50组，因此，对所选择的24种不同气动外形的飞行器，所需计算的飞行状态共1 200组。

表3 带襟翼飞行器气动特性计算结果

Table 3 Calculation results aerodynamic characteristic of aircraft with flap

高度H/km	马赫数Ma	迎角α/(°)($\beta=0°$)	侧滑角β/(°)($\alpha=0°$)
80	28	−10,0,10	5,10
60	26	−10,0,10	5,10
40	25,23	−10,0,10	5,10
30	24,22	−10,0,10	5,10
20	17,15	−10,0,10	5,10
10	12,10	−10,0,10	5,10

2.3 参数特性计算式拟合

根据对气动特性变化规律数据的拟合修正，得到轴向力系数、法向力系数、俯仰力矩系数等气动参数关于可优化参数的关系。近似认为飞行器气动参数是无襟翼状态下气动参数与襟翼气动外形所造成变化的叠加，表示为

$$C_{\text{withflap}}=C_{\text{noflap}}+\Delta C \tag{1}$$

式中：C_{withflap}和C_{noflap}为飞行器和飞行器无襟翼状态的任一气动参数；ΔC为襟翼引起的气动参数变化，为飞行马赫数、迎角、侧滑角的函数，即

$$\Delta C=f(l,h,\delta,Ma,\alpha,\beta) \tag{2}$$

拟合的目的是当给定襟翼的气动外形尺寸后，可得到对应飞行器的气动特性，假设

$$\Delta C=\Delta C_l+\Delta C_h+\Delta C_\delta \tag{3}$$

式中：ΔC_l、ΔC_h、ΔC_δ分别为由襟翼长度、张角、高度所决定的部分，可表示为

$$\begin{cases}\Delta C_l=f_1(l,Ma,\alpha,\beta)\\ \Delta C_h=f_2(h,Ma,\alpha,\beta)\\ \Delta C_\delta=f_3(\delta,Ma,\alpha,\beta)\end{cases} \tag{4}$$

为减小计算量的同时保证所需数据量，对襟翼的3个可优化参数均取2个值，假设

$$\begin{cases}\Delta C_l=K_l l+b_l\\ \Delta C_h=K_h h+b_h\\ \Delta C_\delta=K_\delta\delta+b_\delta\end{cases} \tag{5}$$

式中：K_l、b_l、K_h、b_h、K_δ、b_δ为拟合参数。

将式(5)代入式(3)可得

$$\begin{aligned}\Delta C=&\Delta C_l+\Delta C_h+\Delta C_\delta=\\&(K_l l+b_l)+(K_h h+b_h)+(K_\delta\delta+b_\delta)\end{aligned} \tag{6}$$

当襟翼的可优化气动外形参数均为0时，气动参数表达式中由襟翼引起的气动变化为0，此时飞行器为飞行器无襟翼状态($\Delta C=0$)，即$b_l+b_h+b_\delta=0$，式(3)可表示为

$$\begin{aligned}\Delta C=&\Delta C_l+\Delta C_h+\Delta C_\delta=\\&K_l l+K_h h+K_\delta\delta\end{aligned} \tag{7}$$

将式(7)代入式(1)中，可得

$$\begin{aligned}C_{\text{withflap}}=&C_{\text{noflap}}+\Delta C_l+\Delta C_h+\Delta C_\delta=\\&C_{\text{noflap}}+K_l l+K_h h+K_\delta\delta\end{aligned} \tag{8}$$

式(8)为飞行器气动参数拟合经验计算式的基本形式。综合式(3)～式(7)，可知系数K_l、K_h、K_δ的影响因素为飞行马赫数、迎角、侧滑角，有

$$\begin{cases}K_l=g_1(Ma,\alpha,\beta)\\ K_h=g_2(Ma,\alpha,\beta)\\ K_\delta=g_3(Ma,\alpha,\beta)\end{cases} \tag{9}$$

选取某种气动外形的飞行器作为基准，分别选取与其仅襟翼长度不同、仅襟翼张角不同、仅襟翼高度不同的3种飞行器，按式(1)～式(9)拟合方法，可得到g_1、g_2、g_3的具体函数关系。再将式(9)代入式(8)，通过输入气动外形尺寸、飞行马赫数、迎角、侧滑角可得到飞行器气动特性拟合计算结果。

1) 轴向力系数

$$\begin{cases}C_{x,\text{withflap}}=C_{x,\text{noflap}}+K_l l+K_h h+K_\delta\delta\\ K_l=k_{l_5}Ma^5+k_{l_4}Ma^4+k_{l_3}Ma^3+k_{l_2}Ma^2+k_{l_1}Ma+k_{l_0}\\ K_h=k_{h_5}Ma^5+k_{h_4}Ma^4+k_{h_3}Ma^3+k_{h_2}Ma^2+k_{h_1}Ma+k_{h_0}\\ K_\delta=k_{\delta_5}Ma^5+k_{\delta_4}Ma^4+k_{\delta_3}Ma^3+k_{\delta_2}Ma^2+k_{\delta_1}Ma+k_{\delta_0}\end{cases} \tag{10}$$

式中：

$$\begin{cases}k_{l_i}=a_{l_i}\alpha^3+b_{l_i}\alpha^2+c_{l_i}\alpha+d_{l_i}\beta^2+e_{l_i}\beta+f_{l_i}\\ k_{h_i}=a_{h_i}\alpha^3+b_{h_i}\alpha^2+c_{h_i}\alpha+d_{h_i}\beta^2+e_{h_i}\beta+f_{h_i}\\ k_{\delta_i}=a_{\delta_i}\alpha^3+b_{\delta_i}\alpha^2+c_{\delta_i}\alpha+d_{\delta_i}\beta^2+e_{\delta_i}\beta+f_{\delta_i}\\ i=5,4,3,2,1,0\end{cases} \tag{11}$$

其中：a_Δ、b_Δ、c_Δ、d_Δ、e_Δ($\Delta=l_i,h_i,\delta_i$)为拟合参数。

2) 法向力系数

$$\begin{cases}C_{y,\text{withflap}}=C_{y,\text{noflap}}+K_l l+K_h h+K_\delta\delta\\ K_l=k_{l_4}Ma^4+k_{l_3}Ma^3+k_{l_2}Ma^2+k_{l_1}Ma+k_{l_0}\\ K_h=k_{h_4}Ma^4+k_{h_3}Ma^3+k_{h_2}Ma^2+k_{h_1}Ma+k_{h_0}\\ K_\delta=k_{\delta_4}Ma^4+k_{\delta_3}Ma^3+k_{\delta_2}Ma^2+k_{\delta_1}Ma+k_{\delta_0}\end{cases} \tag{12}$$

式中：

$$\begin{cases}k_{l_i}=a_{l_i}\alpha^3+b_{l_i}\alpha^2+c_{l_i}\alpha+d_{l_i}\beta^2+e_{l_i}\beta+f_{l_i}\\ k_{h_i}=a_{h_i}\alpha^3+b_{h_i}\alpha^2+c_{h_i}\alpha+d_{h_i}\beta^2+e_{h_i}\beta+f_{h_i}\\ k_{\delta_i}=a_{\delta_i}\alpha^3+b_{\delta_i}\alpha^2+c_{\delta_i}\alpha+d_{\delta_i}\beta^2+e_{\delta_i}\beta+f_{\delta_i}\\ i=4,3,2,1,0\end{cases} \tag{13}$$

3) 侧向力系数

$$\begin{cases} C_{z,\text{withflap}} = C_{z,\text{noflap}} + K_l l + K_h h + K_\delta \delta \\ K_l = k_{l_4} Ma^4 + k_{l_3} Ma^3 + k_{l_2} Ma^2 + k_{l_1} Ma + k_{l_0} \\ K_h = k_{h_4} Ma^4 + k_{h_3} Ma^3 + k_{h_2} Ma^2 + k_{h_1} Ma + k_{h_0} \\ K_\delta = k_{\delta_4} Ma^4 + k_{\delta_3} Ma^3 + k_{\delta_2} Ma^2 + k_{\delta_1} Ma + k_{\delta_0} \end{cases} \tag{14}$$

式中:

$$\begin{cases} k_{l_i} = a_{l_i}\alpha^2 + b_{l_i}\alpha + c_{l_i}\beta^3 + d_{l_i}\beta^2 + e_{l_i}\beta + f_{l_i} \\ k_{h_i} = a_{h_i}\alpha^2 + b_{h_i}\alpha + c_{h_i}\beta^3 + d_{h_i}\beta^2 + e_{h_i}\beta + f_{h_i} \\ k_{\delta_i} = a_{\delta_i}\alpha^2 + b_{\delta_i}\alpha + c_{\delta_i}\beta^3 + d_{\delta_i}\beta^2 + e_{\delta_i}\beta + f_{\delta_i} \\ i = 4,3,2,1,0 \end{cases} \tag{15}$$

4) 俯仰力矩系数

$$\begin{cases} C_{mz,\text{withflap}} = C_{mz,\text{noflap}} + K_l l + K_h h + K_\delta \delta \\ K_l = k_{l_5} Ma^5 + k_{l_4} Ma^4 + k_{l_3} Ma^3 + k_{l_2} Ma^2 + k_{l_1} Ma + k_{l_0} \\ K_h = k_{h_5} Ma^5 + k_{h_4} Ma^4 + k_{h_3} Ma^3 + k_{h_2} Ma^2 + k_{h_1} Ma + k_{h_0} \\ K_\delta = k_{\delta_5} Ma^5 + k_{\delta_4} Ma^4 + k_{\delta_3} Ma^3 + k_{\delta_2} Ma^2 + k_{\delta_1} Ma + k_{\delta_0} \end{cases} \tag{16}$$

式中:

$$\begin{cases} k_{l_i} = a_{l_i}\alpha^3 + b_{l_i}\alpha^2 + c_{l_i}\alpha + d_{l_i}\beta^2 + e_{l_i}\beta + f_{l_i} \\ k_{h_i} = a_{h_i}\alpha^3 + b_{h_i}\alpha^2 + c_{h_i}\alpha + d_{h_i}\beta^2 + e_{h_i}\beta + f_{h_i} \\ k_{\delta_i} = a_{\delta_i}\alpha^3 + b_{\delta_i}\alpha^2 + c_{\delta_i}\alpha + d_{\delta_i}\beta^2 + e_{\delta_i}\beta + f_{\delta_i} \\ i = 5,4,3,2,1,0 \end{cases} \tag{17}$$

5) 偏航力矩系数

$$\begin{cases} C_{my,\text{withflap}} = C_{my,\text{noflap}} + K_l l + K_h h + K_\delta \delta \\ K_l = k_{l_4} Ma^4 + k_{l_3} Ma^3 + k_{l_2} Ma^2 + k_{l_1} Ma + k_{l_0} \\ K_h = k_{h_4} Ma^4 + k_{h_3} Ma^3 + k_{h_2} Ma^2 + k_{h_1} Ma + k_{h_0} \\ K_\delta = k_{\delta_4} Ma^4 + k_{\delta_3} Ma^3 + k_{\delta_2} Ma^2 + k_{\delta_1} Ma + k_{\delta_0} \end{cases} \tag{18}$$

式中:

$$\begin{cases} k_{l_i} = a_{l_i}\alpha^2 + b_{l_i}\alpha + c_{l_i}\beta^3 + d_{l_i}\beta^2 + e_{l_i}\beta + f_{l_i} \\ k_{h_i} = a_{h_i}\alpha^2 + b_{h_i}\alpha + c_{h_i}\beta^3 + d_{h_i}\beta^2 + e_{h_i}\beta + f_{h_i} \\ k_{\delta_i} = a_{\delta_i}\alpha^2 + b_{\delta_i}\alpha + c_{\delta_i}\beta^3 + d_{\delta_i}\beta^2 + e_{\delta_i}\beta + f_{\delta_i} \\ i = 4,3,2,1,0 \end{cases} \tag{19}$$

6) 滚转力矩系数

$$\begin{cases} C_{mx,\text{withflap}} = C_{mx,\text{noflap}} + K_l l + K_h h + K_\delta \delta \\ K_l = k_{l_4} Ma^4 + k_{l_3} Ma^3 + k_{l_2} Ma^2 + k_{l_1} Ma + k_{l_0} \\ K_h = k_{h_4} Ma^4 + k_{h_3} Ma^3 + k_{h_2} Ma^2 + k_{h_1} Ma + k_{h_0} \\ K_\delta = k_{\delta_4} Ma^4 + k_{\delta_3} Ma^3 + k_{\delta_2} Ma^2 + k_{\delta_1} Ma + k_{\delta_0} \end{cases} \tag{20}$$

式中:

$$\begin{cases} k_{l_i} = a_{l_i}\alpha^2 + b_{l_i}\alpha + c_{l_i}\beta^3 + d_{l_i}\beta^2 + e_{l_i}\beta + f_{l_i} \\ k_{h_i} = a_{h_i}\alpha^2 + b_{h_i}\alpha + c_{h_i}\beta^3 + d_{h_i}\beta^2 + e_{h_i}\beta + f_{h_i} \\ k_{\delta_i} = a_{\delta_i}\alpha^2 + b_{\delta_i}\alpha + c_{\delta_i}\beta^3 + d_{\delta_i}\beta^2 + e_{\delta_i}\beta + f_{\delta_i} \\ i = 4,3,2,1,0 \end{cases} \tag{21}$$

2.4　偏差分析

用C_0表示 CFD 计算结果, C^*表示本文拟合计算结果, 则残差相对偏差e可表示为

$$e = \frac{C^* - C_0}{C_0} \tag{22}$$

以其中某迎角、侧滑角取值组合情况为例, 6 种飞行器气动特性参数拟合值与 CFD 计算结果如图 4 所示, 残差相对偏差e如图 5 所示, 可以看

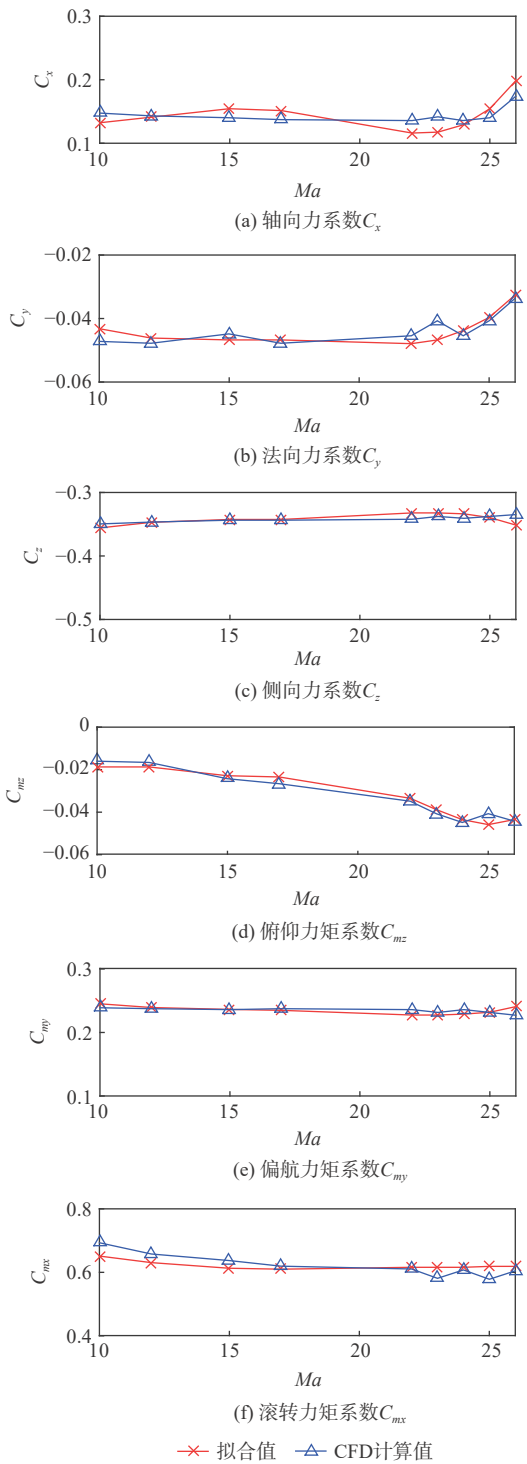


图 4　数值仿真对比

Fig. 4　Comparison of numerical simulation

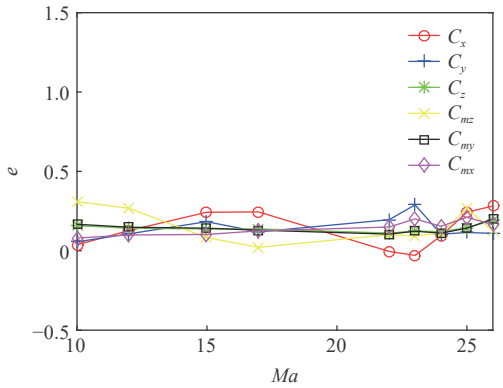


图5 残差相对偏差 e

Fig. 5 Relative error of residual e

出，拟合值与CFD计算结果基本吻合。

3 优化设计

3.1 确定目标函数

随机抽取飞行器不同气动外形下气动特性数据，采用数值仿真分析不同气动外形下，飞行器在不同迎角、侧滑角时轴向力系数C_x、法向力系数C_y、侧向力系数C_z、俯仰力矩系数C_{mz}、偏航力矩系数C_{my}随马赫数的变化规律，在飞行器的气动特性计算中，飞行器的滚转力矩系数C_{mx}始终为零。

1）在其他气动外形参数保持不变的情况下，增大第1锥角、襟翼张角和襟翼高度或减小襟翼长度会使不同马赫数下俯仰通道的配平迎角增大，同时增大相应的配平升力系数，从而能增大飞行器在整个飞行过程中受到的最大法向过载，提高飞行器的机动能力。

2）随着飞行器可优化参数的变化，飞行器在不同马赫数下俯仰通道的配平阻力系数的变化幅度远小于配平迎角或配平升力系数的变化幅度，因此，可以近似认为飞行器可优化参数的变化不影响飞行器阻力系数。

3）在飞行器外形优化过程中，主要通过调整可优化参数提升不同马赫数下的配平迎角与制导系统的匹配性，获得合适的配平升力特性，从而改善制导系统性能。

此外，利用数值仿真分析了阻力和升力对飞行器制导系统性能的影响规律，发现增大阻力系数会使圆概率偏差(circular error probable, CEP)增大，同时末速降低。当升力系数增大时，飞行器机动能力提升、弹道收敛速度较快，在中高空的飞行高度降低，法向过载增大，低空速度衰减较快。因此，目标函数是通过优化飞行器的第1锥角、襟翼长度、襟翼高度和襟翼张角来获取合适的配平升力特性，使制导性能指标最优。

3.2 约束条件

优化结果应满足飞行器性能要求，并设定一定的约束条件，即对升力和阻力系数的约束。最佳导弹几何形状的升力系数$C_{L,\mathrm{opt}}$应高于基线导弹几何形状的升力系数$C_{L,\mathrm{base}}$，而阻力系数$C_{D,\mathrm{opt}}$应小于基线导弹几何形状的阻力系数$C_{D,\mathrm{base}}$，即

$$\begin{cases} C_{L,\mathrm{opt}} \geqslant C_{L,\mathrm{base}} \\ C_{D,\mathrm{opt}} \leqslant C_{D,\mathrm{base}} \end{cases} \tag{23}$$

3.3 优化问题的数学表示

气动布局优化过程的数学表示为

1）离散马赫数序列Ma

$$Ma_1 < Ma_2 < \cdots < Ma_{N-1} < Ma_N$$

式中：N为序列中马赫数的个数。

2）飞行器初始外形参数向量$\boldsymbol{x}_\mathrm{r}$

$$\boldsymbol{x}_\mathrm{r} = [\vartheta_\mathrm{r}, l_\mathrm{r}, h_\mathrm{r}, \delta_\mathrm{r}] \tag{24}$$

式中：ϑ_r为第1锥角的原始值；l_r为襟翼长度的原始值；h_r为襟翼高度的原始值；δ_r为襟翼张角的原始值。

3）根据飞行器气动外形参数和带襟翼飞行器气动参数经验拟合计算式，计算出马赫数序列Ma中不同马赫数Ma_i下的配平升力系数$C_{L,i}(i=1,2,\cdots,N)$。

4）将仿真的统计结果作为制导系统性能的评价指标，主要包括：落点CEP所能容许的最大值$P_{d_\max}$(单位：m)；末速低于500 m/s的百分比所能容许的最大值$P_{V_\max}$(单位：%)；最大飞行过载高于60g的百分比所能容许的最大值$P_{n_\max}$(单位：%)。

设飞行器初始升力系数为$C_{L,\mathrm{r},i}(i=1,2,\cdots,N)$，通过仿真获取其落点圆概率偏差CEP为$P_{d_0}$、末速低于500 m/s的占比为$P_{V_0}$、最大飞行过载高于60$g$的占比为$P_{n_0}$。

5）设升力特性的相对变化量为Δ，配平升力系数为$(1+\Delta)C_{L,\mathrm{r},i}(i=1,2,\cdots,N)$的飞行器，通过仿真获取其落点圆概率偏差CEP为$P_{d_0}$、末速低于500 m/s的占比$P_{V_0}$、最大飞行过载高于60$g$的占比$P_{n_0}$。

6）在制导参数及干扰参数保持不变的情况下，制导系统性能指标P_{d_Δ}、P_{V_Δ}、P_{n_Δ}为飞行器初始外形参数向量$\boldsymbol{x}_\mathrm{r}$(其决定了飞行器的在不同马赫数下的配平升力系数$C_{L,\mathrm{r},i}(i=1,2,\cdots,N)$)和配平升力特性的相对变化量$\Delta$的函数，为便于描述，将其写成关于$\boldsymbol{x}_\mathrm{r}$和$\Delta$的函数表达形式，即$P_{d_\Delta}(\boldsymbol{x}_\mathrm{r},\Delta)$、$P_{V_\Delta}(\boldsymbol{x}_\mathrm{r},\Delta)$和$P_{n_\Delta}(\boldsymbol{x}_\mathrm{r},\Delta)$。则初始飞行器配平升力特性的相对变化量$\Delta$优化的目标函数为

$$\begin{aligned} \min J(\boldsymbol{x}_\mathrm{r},\Delta) = & g\left(P_{d_\Delta}, P_{d_\max}\right) + g\left(P_{V_\Delta}, P_{V_\max}\right) + \\ & g\left(P_{n_\Delta}, P_{n_\max}\right) \\ \text{s.t.} \quad & \Delta_{\min} \leqslant \Delta \leqslant \Delta_{\max} \end{aligned} \tag{25}$$

式中: 制导系统性能指标P_{d_Δ}、P_{V_Δ}、P_{n_Δ}从仿真统计结果中获取; $\Delta_{\min}$和$\Delta_{\max}$分别为相对变化量优化范围的最小值和最大值, 函数$g(x,y)$的表达式为

$$g(x,y)=\begin{cases}0 & x\leqslant y\\ x-y & x>y\end{cases} \tag{26}$$

7) 设对上述升力特性的相对变化量优化问题求解所得到的最优相对变化量为$\Delta_{\rm b}$, 飞行器任一外形参数向量$\boldsymbol{x}_{\rm r}=[\vartheta_{\rm r},l_{\rm r},h_{\rm r},\delta_{\rm r}]$在不同马赫数$Ma_i$下的配平升力系数为$C_{L,{\rm r},i}(i=1,2,\cdots,N)$。则气动外形参数优化的目标函数为

$$\begin{aligned}&\min f(\boldsymbol{x},\boldsymbol{x}_{\rm r},\Delta_{\rm b})=\sqrt{\sum_{i=1}^{N}\left[C_{L,i}-(1+\Delta_{\rm b})C_{L,{\rm r},i}\right]^2}\\ &\text{s.t.}\quad \vartheta_{\min}\leqslant\vartheta\leqslant\vartheta_{\max}\\ &\qquad l_{\min}\leqslant l\leqslant l_{\max}\\ &\qquad h_{\min}\leqslant h\leqslant h_{\max}\\ &\qquad \delta_{\min}\leqslant\delta\leqslant\delta_{\max}\end{aligned} \tag{27}$$

式中: $\vartheta_{\min}$和$\vartheta_{\max}$分别为第 1 锥角可优化范围的下限和上限; $l_{\min}$和$l_{\max}$分别为襟翼长度可优化范围的下限和上限; $h_{\min}$和$h_{\max}$分别为襟翼高度可优化范围的下限和上限; $\delta_{\min}$和$\delta_{\max}$分别为襟翼张角优化范围的最小值和最大值。

3.4　优化算法

模拟退火算法 (simulated annealing algorithm, SAA) 是一种应用广泛的随机智能优化算法, 在求解组合优化问题时, SAA 能概率性地跳出局部最优解并最终趋于全局最优, SAA 可以方便地解决复杂的问题, 如外部几何尺寸优化问题, 已被应用于工程和科学领域的许多问题。此外, SAA 在合理的时间内解决了优化问题。在工程问题中, SAA 被认为是本研究和类似应用中更有效和实用的方法,用于求解多目标优化问题。综上, 提出一种改进的模拟退火组合优化算法求解多目标优化模型。

3.5　优化问题求解

在给定的约束条件和选定的设计变量下, 利用目标函数进行优化求解。为减小优化时间, 提升计算效率, 将整个优化过程分为 2 个子优化问题。首先, 以飞行器初始升力特性为基准, 将升力特性相对变化量Δ作为自变量, 寻找使飞行器性能最优的Δ。然后, 对飞行器气动外形参数(ϑ、l、h和δ)进行优化, 使得优化后飞行器的配平升力特性尽可能在各种马赫数下都与最优的Δ对应的配平升力系数接近。

1) 优化问题 1: 升力特性优化模型求解

以飞行器升力特性的相对变化量Δ为自变量, 采用等间隔一维搜索算法在设定的范围内$\Delta\in[\Delta_{\min},\Delta_{\max}]$, 搜寻使目标函数$J(\boldsymbol{x}_{\rm r},\Delta)$取值最小的近似全局最优解$\Delta_{\rm b}$, 算法步骤如下:

步骤 1　根据飞行器初始气动布局参数计算其在不同马赫数下的升力系数, 给定升力特性相对变化量Δ的优化区间$[\Delta_{\min},\Delta_{\max}]$和搜索间隔$\varepsilon$, 设定预测制导参数、飞行器再入参数、蒙特卡罗仿真参数。

步骤 2　使用初始飞行器的升力特性$C_{L,{\rm r},i}(i=1,2,\cdots,N)$进行蒙特卡罗仿真, 并对仿真结果进行统计分析, 得到相应的制导系统性能指标P_{d_0}、P_{V_0}和P_{n_Δ}。

步骤 3　如果$P_{d_0}\leqslant P_{d_\max}$、$P_{V_0}\leqslant P_{V_\max}$、$P_{n_0}\leqslant P_{n_\max}$, 即$J(\boldsymbol{x}_{\rm r},\Delta)=0$, 初始飞行器外形满足制导系统性能指标要求, 转步骤 7; 否则, 令$\Delta=\Delta_{\min}$、$J_{\rm b}=J(\boldsymbol{x}_{\rm r},0)$、$\Delta_{\rm b}=0$, 转步骤 4。

步骤 4　使用升力特性$(1+\Delta)C_{L,{\rm r},i}(i=1,2,\cdots,N)$进行蒙特卡罗仿真, 并对仿真结果进行统计分析, 得到相应的制导系统性能指标P_{d_0}、P_{V_0}和P_{n_Δ}, 并计算$J(\boldsymbol{x}_{\rm r},\Delta)$。

步骤 5　若$J_{\rm b}>J(\boldsymbol{x}_{\rm r},\Delta)$, 则$J_{\rm b}=J(\boldsymbol{x}_{\rm r},0)$、$\Delta_{\rm b}=\Delta$; 若$J_{\rm b}=J(\boldsymbol{x}_{\rm r},\Delta)$且$|\Delta|<|\Delta_{\rm b}|$, 则$J_{\rm b}=J(\boldsymbol{x}_{\rm r},0)$、$\Delta_{\rm b}=\Delta$。

步骤 6　令$\Delta=\Delta+\varepsilon$, 若$\Delta>\Delta_{\max}$, 则转步骤 7; 否则, 转步骤 4。

步骤 7　搜索结束, 当前解Δ为近似全局最优解。

2) 优化问题 2: 外形参数向量优化模型求解

以飞行器气动外形参数向量$\boldsymbol{x}_{\rm r}$为自变量, 采用一种改进的模拟退火组合优化算法, 在设定的范围$\vartheta\in[\vartheta_{\min},\vartheta_{\max}]$、$l\in[l_{\min},l_{\max}]$、$h\in[h_{\min},h_{\max}]$、$\delta\in[\delta_{\min},\delta_{\max}]$内, 搜寻使目标函数$f(\boldsymbol{x},\boldsymbol{x}_{\rm r},\Delta_{\rm b})$取值最小的解$\boldsymbol{x}_{\rm b}$, 其算法流程(见图 6)如下。

步骤 1　根据飞行器初始气动布局参数计算其在不同马赫数下的配平升力系数。

步骤 2　给定初始温度t_0, 终止温度$t_{\rm e}$, 每一温度下的迭代步数L, 降温系数γ, 迭代初始解$\boldsymbol{x}_0=[\vartheta_0,l_0,h_0,\delta_0]$和目标函数值$f(\boldsymbol{x}_0)$。

步骤 3　当温度$t=t_k$时作L次试探搜索。

根据当前解$\boldsymbol{x}_k$产生一个随机向量$\boldsymbol{z}_k$得到$\boldsymbol{x}_k$邻域的新的试探点$\boldsymbol{x}'_k$:

$$\boldsymbol{x}'_k=\boldsymbol{x}_k+\boldsymbol{z}_k \tag{28}$$

式中: $\boldsymbol{x}$为离散变量的取值序列; k为当前解的离散位置。

步骤 4　按 Metropolis 接受准则计算出在给定当前迭代点$\boldsymbol{x}_k$和温度t_k下的转移概率P:

$$P=\begin{cases}1 & f\left(\boldsymbol{x}'_k\right)\leqslant f(\boldsymbol{x}_k)\\ \exp\left(\dfrac{f(\boldsymbol{x}_k)-f\left(\boldsymbol{x}'_k\right)}{t_k}\right) & f\left(\boldsymbol{x}'_k\right)>f(\boldsymbol{x}_k)\end{cases} \tag{29}$$

产生一个在 (0, 1) 上均匀分布的随机数η, 若

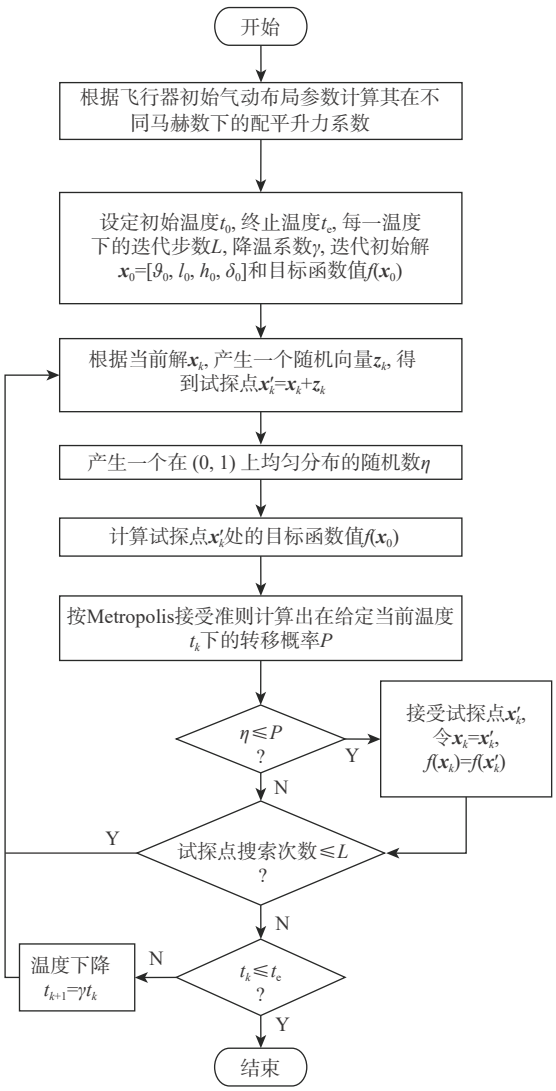


图 6 算法流程

Fig. 6 Algorithm process

$\eta\leqslant P$, 则接受新解$\boldsymbol{x}_k=\boldsymbol{x}'_k$, $f(\boldsymbol{x}_k)=f(\boldsymbol{x}'_k)$, 否则$\boldsymbol{x}_k$不变。

步骤 5 若试探点搜索小于L次, 转步骤 3, 否则转步骤 6。

步骤 6 若当前温度$t=t_k$小于终止温度t_e, 结束搜索, 当前解为近似全局最优解; 否则转步骤 7。

步骤 7 根据给定的温度衰减函数$t_{k+1}=\gamma t_k$, 产生新的温度控制参数$t=t_{k+1}$。

步骤 8 重复步骤 3～步骤 7, 直到找到最优解。

3.6 优化结果

结果表明, 在执行优化算法后, 飞行器升阻特性改善显著。表 4 为飞行器优化前设计变量取值。结果表明, 襟翼长度和襟翼张角在提升升力系数方面更为有效。

表 4 优化前后的设计变量取值

Table 4 Value of design variables before and after optimization

阶段	ϑ/(°)	l/mm	δ/(°)	h/mm
优化前	7.5	900	56	120
优化后	7.547 3	864	58.174 1	119.3

表 5 为飞行器优化前后的升阻特性, 在不同马赫数下, 优化前后的阻力系数基本保持不变, 优化后的升力系数在优化前的配平升力系数基础上相对增大了 10%。升力系数较小表明该飞行器机动能力较弱, 导致飞行器末速和命中精度的下降, 通过优化算法有效增大了飞行器在不同马赫数下的升力系数, 实现了在不影响最大飞行过载的前提下, 提高导弹末速和命中精度的目标。

表 5 优化前后的升阻特性

Table 5 Lift and drag characteristics before and after optimization

Ma	阻力系数		升力系数	
	优化前	优化后	优化前	优化后
1	0.557 3	0.558 9	0.263 4	0.288 5
2	0.430 1	0.431 1	0.217 8	0.240 0
4	0.243 0	0.243 6	0.143 6	0.160 2
6	0.187 4	0.188 1	0.126 0	0.143 0
8	0.169 4	0.170 6	0.106 5	0.123 6
10	0.151 5	0.152 9	0.094 5	0.113 1
12	0.131 3	0.132 2	0.057 5	0.073 8
14	0.158 1	0.159 3	0.099 2	0.113 1
16	0.168 3	0.169 5	0.110 4	0.124 2
18	0.153 5	0.154 7	0.087 5	0.096 8
20	0.123 3	0.124 4	0.110 4	0.121 8

3.6.1 标称条件下的仿真试验

假设飞行器采用预测制导, 弹道倾角为−5.0°, 弹道偏角为 0°, 飞行速度为 7 000 m/s, 飞行高度为 80 000 m。假设地球为圆球, 不考虑其自转角速度。飞行轨迹及各参数的变化曲线如图 7 所示。

仿真曲线表明, 标称条件下, 优化前后的飞行器命中精度基本相同, 但是末速和飞行过载曲线存在一些差异。优化后的飞行器末速比优化前的较大, 在飞行过程中优化后的法向过载大于优化前, 因此, 在满足最大飞行过载指标要求的前提下, 既提高了末速, 也增强了飞行器的机动能力, 从而提高飞行器针对不确定性的鲁棒性, 可以降低 CEP。

3.6.2 多种偏差条件下的仿真试验

采用蒙特卡罗打靶法, 设定 7 类随机偏差, 并假设随机偏差变量符合均值为零及3σ准则的正态分布, 如表 6 和表 7 所示。

根据偏差取值和分布规律对优化前和优化后的飞行器进行 500 次蒙特卡罗打靶仿真试验, 仿真统计分析结果如表 8 和表 9 所示。

仿真结果表明, 飞行器优化前的 CEP 为 516.9 m,

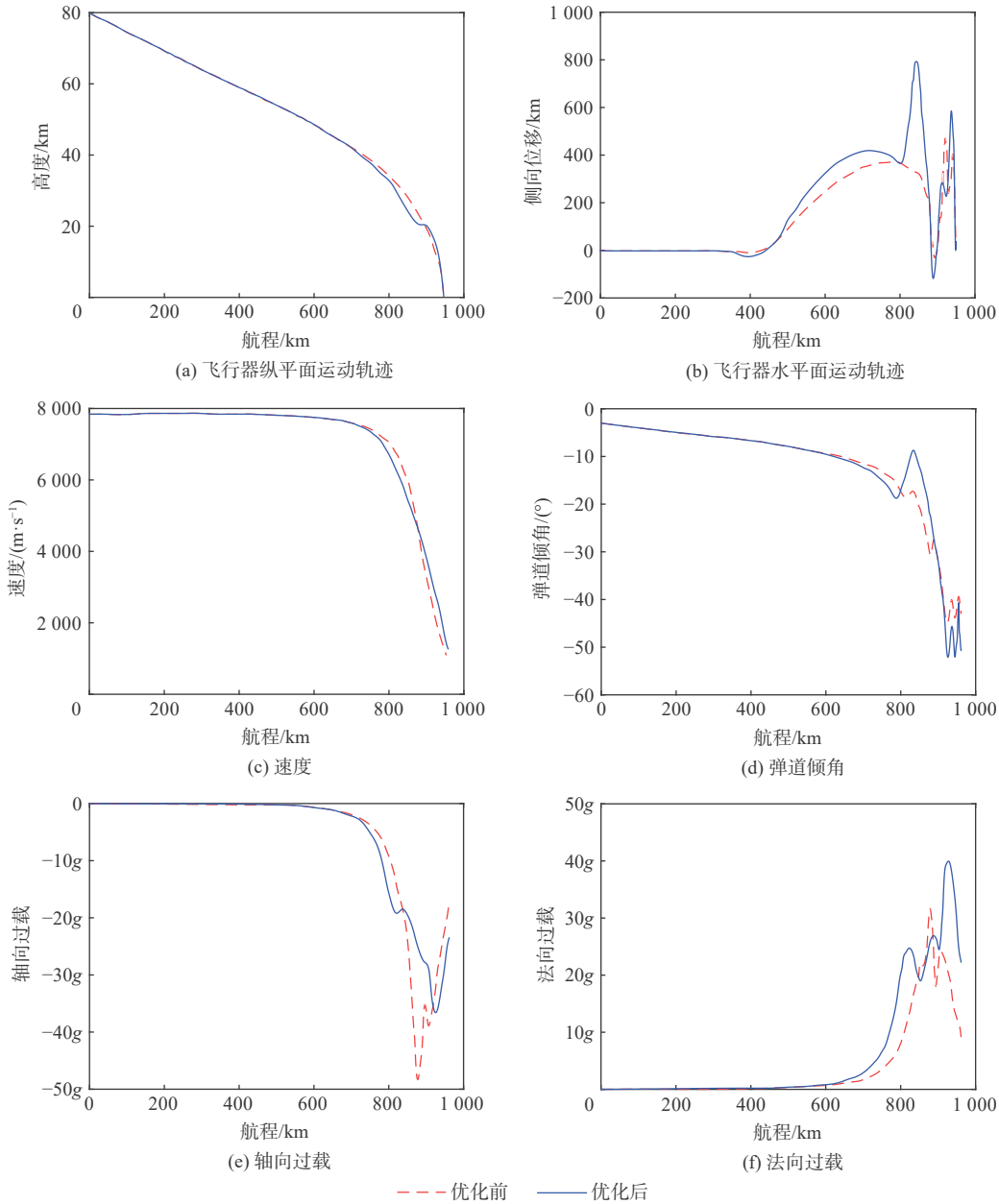


图 7　标称条件下优化前后仿真曲线

Fig. 7　Simulation curves before and after optimization under nominal conditions

表 6　随机偏差取值范围

Table 6　Range of random deviation

随机偏差变量	取值范围
初始弹道倾角偏差 $\Delta\theta_0$/(°)	[−0.4, 0.4]
大气密度偏差比例系数 ΔP_ρ/%	[−10, 10]
法向力系数偏差比例系数 ΔP_{CN}/%	[−15, 15]
轴向力系数偏差比例系数 ΔP_{CA}/%	[−15, 15]
配平迎角偏差 $\Delta\alpha_{\mathrm{T}}$/(°)	[−2.8, 2.8]
法向力系数常值偏差 $\Delta\mathrm{CN}_0$	[−0.015, 0.015]
轴向力系数常值偏差 $\Delta\mathrm{CA}_0$	[−0.015, 0.015]

表 7　随机偏差变量的正态分布规律

Table 7　Normal distribution law of random deviation variables

随机偏差变量	正态分布规律
初始弹道倾角偏差	$\Delta\theta_0 \sim N\left(0,\sigma_{\theta_0}^2\right),\ \sigma_{\theta_0}=0.4/3$
大气密度偏差比例系数	$\Delta P_\rho \sim N\left(0,\sigma_{P_\rho}^2\right),\ \sigma_{P_\rho}=0.1/3$
法向力系数偏差比例系数	$\Delta P_{\mathrm{CN}} \sim N\left(0,\sigma_{P_{\mathrm{CN}}}^2\right),\ \sigma_{P_{\mathrm{CN}}}=0.15/3$
轴向力系数偏差比例系数	$\Delta P_{\mathrm{CA}} \sim N\left(0,\sigma_{P_{\mathrm{CA}}}^2\right),\ \sigma_{P_{\mathrm{CA}}}=0.15/3$
配平迎角偏差	$\Delta\alpha_{\mathrm{T}} \sim N\left(0,\sigma_{\alpha_{\mathrm{T}}}^2\right),\ \sigma_{\alpha_{\mathrm{T}}}=2.8/3$
法向力系数常值偏差	$\Delta\mathrm{CN}_0 \sim N\left(0,\sigma_{\mathrm{CN}_0}^2\right),\ \sigma_{\mathrm{CN}_0}=0.015/3$
轴向力系数常值偏差	$\Delta\mathrm{CA}_0 \sim N\left(0,\sigma_{\mathrm{CA}_0}^2\right),\ \sigma_{\mathrm{CA}_0}=0.015/3$

表 8 优化前仿真数据统计

Table 8 Statistics of simulation data before optimization

落点偏差/km	落点偏差分布占比/%	末速/(m·s^{-1})	末速分布占比/%	末弹道倾角/(°)	末弹道倾角分布占比/%	最大过载	最大过载分布占比/%
[0,0.1)	88.11	[0,200)	0	[−90,−70)	0	[20g,40g)	53.15
[0.1,0.2)	5.59	[200,300)	0	[−70,−55)	6.29	[40g,50g)	26.57
[0.2,0.5)	0.70	[300,400)	0	[−55,−40)	28.67	[50g,55g)	9.79
[0.5,1)	0	[400,500)	0	[−40,−30)	46.85	[55g,60g)	6.99
[1,3)	2.10	[500,600)	5.59	[−30,−20)	18.18	[60g,65g)	2.10
[3,∞)	3.50	[600,∞)	94.41	[−20,0)	0	[65g,∞)	1.40

表 9 优化后仿真数据统计

Table 9 Statistics of simulation data before optimization

落点偏差/km	落点偏差分布占比/%	末速/(m·s^{-1})	末速分布占比/%	末弹道倾角/(°)	末弹道倾角分布占比/%	最大过载	最大过载分布占比/%
[0,0.1)	88.61	[0,200)	0	[−90,−70)	1.27	[20g,40g)	53.16
[0.1,0.2)	8.23	[200,300)	0	[−70,−55)	6.96	[40g,50g)	24.05
[0.2,0.5)	1.27	[300,400)	0	[−55,−40)	29.11	[50g,55g)	11.39
[0.5,1)	0.63	[400,500)	0.63	[−40,−30)	46.20	[55g,60g)	6.33
[1,3)	0.63	[500,600)	3.80	[−30,−20)	16.46	[60g,65g)	3.80
[3,∞)	0.63	[600,∞)	95.57	[−20,0)	0	[65g,∞)	1.27

落点偏差分布在 500 m 以内的占比约为 94.4%，飞行器末速大于 500 m/s 的占比约为 100%，最大飞行过载小于 60g 的占比约为 96.5%；优化后的飞行器 CEP 为 201.3 m，落点偏差分布在 500 m 以内的占比约为 98.11%，飞行器末速大于 500 m/s 的占比约为 99.37%，最大飞行过载小于 60g 的占比约为 94.93%。优化后的飞行器可以显著提升命中精度，同时末速也满足指标要求，飞行器性能得到有效提升。优化后的导弹制导性能评估结果见表 10，飞行器的制导系统性能也得到显著提升。

表 10 制导性能评估结果

Table 10 Guidance performance evaluation results

类型	CEP /m	末速低于500 m/s的占比/%	最大飞行过载
理想值	≤200	2	6g
优化前	526.8	0	3.7g
优化后	215.4	0.61	5.2g

4 结 论

1）优化结果表明，对于每个选定的设计变量都得到了改进，升阻比得到了明显提升。通过标称条件和多种偏差条件下的仿真试验，验证了包括圆概率偏差、落点偏差和末速在内的制导性能指标均得到显著提升。

2）提出的优化方法不但保持了较好的升阻特性，而且实现了高超声速飞行器的多部件多参数设计和优化，具有更好的灵活性和实际工程应用价值。

参考文献（References）

[1] DECK S, DUVEAU P, D'ESPINEY P, et al. Development and application of Spalart–Allmaras one equation turbulence model to three-dimensional supersonic complex configurations[J]. Aerospace Science and Technology, 2002, 6(3): 171-183.

[2] LÓPEZ D, DOMÍNGUEZ D, GONZALO J. Optimization of air-ejected rocket/missile geometries under validated supersonic flow field simulations[J]. AIP Conference Proceedings, 2014, 1637(1): 600-606.

[3] 乔建领, 韩忠华, 宋文萍. 基于代理模型的高效全局低音爆优化设计方法[J]. 航空学报, 2018, 39(5): 62-75.
QIAO J L, HAN Z H, SONG W P. An efficient surrogate-based global optimization for low sonic boom design[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(5): 62-75(in Chinese).

[4] 孙祥程, 韩忠华, 柳斐, 等. 高超声速飞行器宽速域翼型/机翼设计与分析[J]. 航空学报, 2018, 39(6): 26-37.
SUN X C, HAN Z H, LIU F, et al. Design and analysis of hypersonic vehicle airfoil/wing at wide-range Mach numbers[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(6): 26-37(in Chinese).

[5] 车竞, 唐硕, 何开锋. 高超声速飞行器气动布局总体性能优化设计研究[J]. 空气动力学学报, 2009, 27(2): 214-219.
CHE J, TANG S, HE K F. Research on aerodynamic configuration optimization of integral performance for hypersonic cruise vehicle[J]. Acta Aerodynamica Sinica, 2009, 27(2): 214-219(in Chinese).

[6] 刘文. 高超声速乘波体气动布局优化及稳定性研究[D]. 西安: 西

北工业大学, 2018.
LIU W. Aerodynamic layout optimization and stability study of hypersonic waverider [D]. Xi 'an: Northwestern Polytechnical University, 2018 (in Chinese).

[7] 陈立立, 郭正, 侯中喜, 等. 组合式高超声速飞行器布局设计与优化分析[J]. 气体物理, 2019, 4(6): 29-39.
CHEN L L, GUO Z, HOU Z X, et al. Layout design and optimization analysis of combined hypersonic vehicle[J]. Physics of Gases, 2019, 4(6): 29-39(in Chinese).

[8] 唐伟, 冯毅, 杨肖峰, 等. 非惯性弹道飞行器气动布局设计实践[J]. 气体物理, 2017, 2(1): 1-12.
TANG W, FENG Y, YANG X F, et al. Practices of aerodynamic configuration design for non-ballistic trajectory vehicles[J]. Physics of Gases, 2017, 2(1): 1-12(in Chinese).

[9] 徐永杰, 王志军, 吴国东, 等. 气动外形对火箭弹弹道特性影响的数值模拟研究[J]. 弹箭与制导学报, 2018, 38(4): 11-14.
XU Y J, WANG Z J, WU G D, et al. Numerical simulation research on effect of aerodynamic configuration on rocket projectile's ballistics characteristics[J]. Journal of Projectiles, Rockets, Missiles and Guidance, 2018, 38(4): 11-14(in Chinese).

[10] NGUYEN N V, TYAN M, LEE J W, et al. Investigations on missile configuration aerodynamic characteristics for design optimization[J]. Transactions of the Japan Society for Aeronautical and Space Sciences, 2014, 57(4): 210-218.

[11] VIDANOVIC N, RASUO B, DAMLJANOVIC D, et al. Validation of the CFD code used for determination of aerodynamic characteristics of nonstandard AGARD-B calibration model[J]. Thermal Science, 2014, 18(4): 1223-1233.

[12] SAHU J. CFD simulations of a finned projectile with microflaps for flow control[J]. International Journal of Aerospace Engineering, 2017, 2017: 1-15.

[13] OCOKOLJIC G, RASUO B, BENGIN A. Aerodynamic shape optimization of guided missile based on wind tunnel testing and computational fluid dynamics simulation[J]. Thermal Science, 2017, 21(3): 1543-1554.

[14] AGEEV N, PAVLENKO A. Minimization of body of revolution aerodynamic drag at supersonic speeds[J]. Aircraft Engineering and Aerospace Technology, 2016, 88(2): 246-256.

[15] VIDANOVIĆ N, RAŠUO B, KASTRATOVIĆ G, et al. Aerodynamic–structural missile fin optimization[J]. Aerospace Science and Technology, 2017, 65: 26-45.

[16] KÖRPE D S, KANAT Ö Ö. Aerodynamic optimization of a UAV wing subject to weight, geometric, root bending moment, and performance constraints[J]. International Journal of Aerospace Engineering, 2019, 2019: 1-14.

[17] RIDDLE D B, HARTFIELD R J, BURKHALTER J E, et al. Genetic-algorithm optimization of liquid-propellant missile systems[J]. Journal of Spacecraft and Rockets, 2009, 46(1): 151-159.

[18] CAO R D, ZHANG X B. Multi-objective optimization of the aerodynamic shape of a long-range guided rocket[J]. Structural and Multidisciplinary Optimization, 2018, 57(4): 1779-1792.

[19] YANG Y R, JUNG S K, CHO T H, et al. Aerodynamic shape optimization system of a canard-controlled missile using trajectory-dependent aerodynamic coefficients[J]. Journal of Spacecraft and Rockets, 2012, 49(2): 243-249.

[20] HE Y L, AGARWAL R K. Shape optimization of NREL S809 airfoil for wind turbine blades using a multi-objective genetic algorithm[C]//32nd AIAA Applied Aerodynamics Conference. Reston: AIAA, 2014.

Multi-objective optimization of aerodynamic layout for hypersonic reentry vehicle

FAN Boxuan[1, 2, *], CHEN Guiming[1], CAO Yiqun[2]

(1. Rocket Army Engineering University, Xi'an 710025, China;
2. The First Military Representative Office of Rocket Army Equipment Department in Xi'an, Xi'an 710025, China)

Abstract: Multi-objective optimization of aerodynamic layout is a key technology in aircraft design. A new multi-objective optimization method is proposed for aerodynamic shape parameters of hypersonic reentry vehicle, and then the influence of the method on the performance of hypersonic reentry vehicle is proved. Through example simulation, the influence of drag and lift on guidance performance is verified and analyzed in detail, with 3 performance indicators, including circular error probable of aircraft landing point, proportion of landing speed greater than 500 m/s, and proportion of maximum flight overload less than 60*g*, set as the optimization objective. By taking lift characteristics as intermediate parameters, the aerodynamic layout optimization problem divided into two sub problems. Through the lift characteristic optimization based on search algorithm and the shape parameter optimization based on an improved simulated annealing algorithm, the optimization calculation time is reduced, the calculation efficiency is improved, thus optimizing the overall aerodynamic layout of the main body and flap of the aircraft, and obtaining the best aircraft shape under hypersonic flow. The simulation results show that under determined constraints, the optimization algorithm increases the aerodynamic lift of the aircraft under supersonic flow and effectively improves the lift drag ratio. On the premise of not affecting the maximum flight overload, the optimized aircraft shows higher aerodynamic performance, and significantly improved hit accuracy, with the landing speed in compliance with the index requirements, and the performance of the guidance system effectively improved.

Keywords: hypersonic reentry vehicle; pneumatic layout; lift characteristic; simulated annealing combinatorial optimization algorithm; optimal design

Received: 2021-09-23; **Accepted**: 2022-01-05; **Published Online**: 2022-02-14 16:13
URL: kns.cnki.net/kcms/detail/11.2625.V.20220214.1529.007.html
Foundation item: National Natural Science Foundation of China (71601180)
* **Corresponding author**. E-mail: 1092442646@qq.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0497

基于特征运动观测的蝴蝶前飞规律及样机验证

张益鑫[1]，王兴坚[1,2,3]，王少萍[1,2,3,*]，池小楷[1]，杜韶阳[1]

(1. 北京航空航天大学 自动化科学与电气工程学院，北京 100191；
2. 北京航空航天大学 大数据精准医疗高精尖创新中心，北京 100191；
3. 北京航空航天大学 宁波创新研究院，宁波 315800)

摘 要：为了研究蝴蝶扑翼飞行的原理，研制低频扑翼的仿生器，通过蝴蝶飞行运动的生物学观测，提出蝴蝶的3种特征运动状态，分析扑翼运动、胸部俯仰运动及腹部摆动运动之间的相位关系，构建蝴蝶前飞运动学模型。基于“杆-膜”仿生翼的新工艺和定制的机载飞控系统，研制轻量化的仿生蝴蝶扑翼飞行器样机，研究蝴蝶样机的飞行控制策略。通过六维力传感器对样机做地面动力学测试，利用高速摄像机对样机飞行进行运动学跟踪，证明了基于特征运动状态的蝴蝶前飞规律和原理样机研制的有效性。

关 键 词：蝴蝶飞行规律；典型特征运动；扑翼飞行器；仿生机器人；膜翼制作
中图分类号：V221; TH113
文献标志码：A **文章编号**：1001-5965(2023)07-1651-10

蝴蝶是节肢动物门下鳞翅目的昆虫，具有高超的飞行能力。蝴蝶与其他扑翼飞行的动物相比，具有翼展弦比较小(甚至小于1)[1]、扇翅幅度较大、频率较低(平均11 Hz左右)[2]且伴随着翼和身体运动的高度耦合特点[3]。蝴蝶通过身体同侧的前翅和后翅翼面积较大且同步拍打实现飞行，其飞行轨迹具有较大的波动，但可以达到点对点的飞行目标，甚至上千公里的长途迁徙[4]，足以证明蝴蝶飞行行为的与众不同。

近30年来，随着对昆虫飞行运动观测及分析技术的不断提升，蝴蝶飞行的运动学规律得以揭示。文献[2,5-7]对蝴蝶这类扑动小展弦比翼飞行的昆虫进行了研究。文献[8]在风洞中对自由飞行的蝴蝶进行了烟流显示试验，得到扑翼过程中翼周围流场的涡结构，确定了前缘涡及“合拢剥离”机制的重要作用。文献[9]通过求解雷诺平均Navier-Stokes方程数值方法定量研究了蝴蝶前飞的空气动力学特性。文献[1]指出蝴蝶在飞行过程中翼相对身体仅有上拍和下拍动作，没有类似蜜蜂、长喙天蛾等飞行昆虫显著的翼翻转动作。蝴蝶身体存在较为明显的俯仰震荡，翼在身体的带动下产生周期性的俯仰。

由于蝴蝶的飞行具有非定常、不稳定等特点，其飞行机理较为复杂，目前世界上对仿生蝴蝶扑翼飞行器的研究较少。仿生蝴蝶扑翼飞行器飞行轨迹的杂乱使得其难以被瞄准，能够在一定程度上躲避攻击；仿生翼一般采用透明膜翅结构，在白天阳光背景下样机可被观测的体积很小，适合近距离侦查作业；由于仿生的设计特点，该类飞行器可用于研究蝴蝶飞行机理及飞行表演。未来随着微机电技术的发展，通过对帝王蝶长途飞行能力的研究，可为研制长航时微型扑翼飞行器提供仿生学原型。文献[10-11]对自然蝴蝶自由飞行及身体固定状态下的扑翼进行观测，研究了蝴蝶翼在飞行中的

收稿日期：2021-08-28；录用日期：2021-09-03；网络出版时间：2021-09-13 17:28
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20210913.1527.002.html
基金项目：国家科技重大专项(2017-V-0010-0060)；国家自然科学基金(51620105010,51805026,51675019)；国家重点基础研究发展计划(JCKY2018601C107)
*通信作者. E-mail：wangshaoping@vip.sina.com

引用格式：张益鑫，王兴坚，王少萍，等. 基于特征运动观测的蝴蝶前飞规律及样机验证[J]. 北京航空航天大学学报，2023，49(7)：1651-1660. ZHANG Y X，WANG X J，WANG S P，et al. Mechanism of butterfly forward flight and prototype verification based on characteristic motion observation[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1651-1660 (in Chinese).

动态行为，研制了2种蝴蝶型扑翼机来模拟蝴蝶翼的扑打运动，并阐明了翼变形及翅脉存在对飞行性能的影响。文献[12-13]对真实蝴蝶尺度下小型扑翼飞行器进行了持续的研究和迭代，基于曲柄摇杆急回机构设计了电机驱动的微型仿生蝴蝶扑翼飞行器，可作为危险环境下的观测系统。德国FESTO公司研制的eMotionButterfly可实现平飞、爬升和转弯飞行，飞行性能较为优异[14]。文献[15]设计了一种前后翅一体的仿生蝴蝶扑翼机，通过舵机直接驱动实现翼扑打，并完成试飞，但升力略显不足。以上研究虽然都研制了相应的仿生蝴蝶扑翼飞行器，但是均对蝴蝶扑翼及翼身耦合的机理进行简化，而这种低频、大幅度的扑翼运动对蝴蝶的飞行至关重要。

本文将蝴蝶飞行运动分解为翼扑打、胸部俯仰和腹部摆动，基于高速摄像技术对蝴蝶飞行进行了生物学观测，阐明蝴蝶扑翼的基本运动规律，以自然蝴蝶生物形态为基准等比放大设计并制作了一种新型仿生蝴蝶扑翼飞行器，实现了可控飞行，并验证了观测得到的蝴蝶飞行规律的有效性，为进一步阐明蝴蝶扑翼飞行机理提供新的思路。

1 自然蝴蝶飞行观测

1.1 坐标系统定义

自然蝴蝶可以被简化为4个部分，分别为左翼、右翼、头胸部和腹部，各部分之间存在相对运动，远比固定翼飞机复杂。双翼分别与胸部铰接，胸部和腹部之间通过胸腹关节G相铰接。图1为用于蝴蝶观测及建模研究的坐标系统定义，包括惯性坐标系、随体坐标、翼坐标系和腹部坐标系，且均为笛卡尔坐标系。

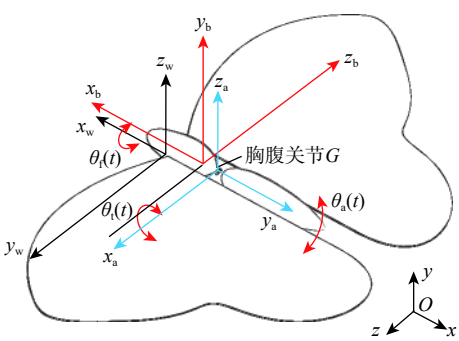


图1 蝴蝶扑翼模型坐标系统定义

Fig. 1 Definition of coordinate system of butterfly flapping wing model

为了描述蝴蝶在空间中的位置，以及相对地面的运动状态，包括绝对飞行速度和运动轨迹，引入惯性坐标系$Oxyz$。原点O为空间内确定的一点，设置在开始记录飞行时刻蝴蝶的重心处。x轴在水平面内指向某一方向，y轴竖直向上，xy平面位于竖直面内。按右手定则确定z轴，xz平面位于水平面内。为了精确描述蝴蝶飞行时自身的运动姿态，引入随体坐标系$O_bx_by_bz_b$，与蝴蝶身体固连，原点O_b设置在蝴蝶的重心，重心处在距离胸部和腹部铰链关节L_g处位置。本文为了简化处理，假设重心与铰链关节G重合($L_g=0$)，且初始时刻O_b与O重合。x_b轴与胸部中心对称轴重合，指向头部方向为正。y_b轴在机体纵向对称面内，垂直于x_b轴向上，z_b轴按右手定则确定。

为了描述蝴蝶翼的运动，引入翼坐标系$O_wx_wy_wz_w$，随着翼面运动。原点O_w设置在翼根位置，翼根定义为胸部和翼之间的连接点，处于x_bz_b面内，在x_b方向上与关节G的距离为l_w，在z_b方向上与关节G的距离定义为翼根连接位置偏移量Δl。y_w轴在左翼平面内，沿翼展向方向，由翼根指向翼尖；x_w轴沿翼弦向方向，由后缘指向前缘；z_w轴垂直于翼面。假设腹部为刚体，为了描述腹部相对胸部的灵活运动引入腹部坐标系$O_ax_ay_az_a$。原点O_a固定于G处，y_a与腹部中心对称轴重合，由G指向腹部末端，x_a轴位于x_bz_b平面内垂直于y_a，z_a轴由右手定则确定，垂于x_ay_a平面。

本文认为蝴蝶前飞的扑翼运动左右对称，并且将前后翅看做一个整体，翼的运动通过翼坐标系相对随体坐标系的旋转来描述。翼扑打轴x_w与机体纵轴x_b平行，翼面与身体水平面(x_bz_b)夹角即为扑打角θ_f。采用右手定则定义正的旋转运动。由于随体坐标系x_b轴与胸部中心对称轴重合，定义胸部俯仰角为胸部中心对称轴与惯性坐标系下水平面的夹角θ_t。由于y_a与腹部中心对称轴重合，定义腹部摆动角为腹部中心对称轴与惯性坐标系下水平面的夹角θ_a。

1.2 蝴蝶的特征运动

为了能够更准确描述蝴蝶的飞行规律，本文将蝴蝶的运动定义为3种基本特征运动：翼扑打运动、胸部俯仰运动和腹部摆动运动，如图2所示。

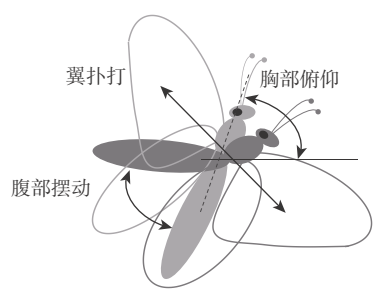


图2 蝴蝶飞行过程中的3种基本特征运动

Fig. 2 Three basic movements of butterflies during flight

1.3　生物观测实验设置

为了研究蝴蝶的飞行行为, 尤其是 3 种基本特征运动下的活体蝴蝶自由飞行规律, 详细观测并定量分析其运动规律非常必要。

本文的实验观测对象选用中国华南地区繁殖基地捕获的成年斑凤蝶。昆虫飞行观测的实验系统一般包括硬件平台和图像处理软件 2 部分。硬件平台一般包括高速摄像机、补光光源、观测箱等, 观测实验系统如图 3 所示。

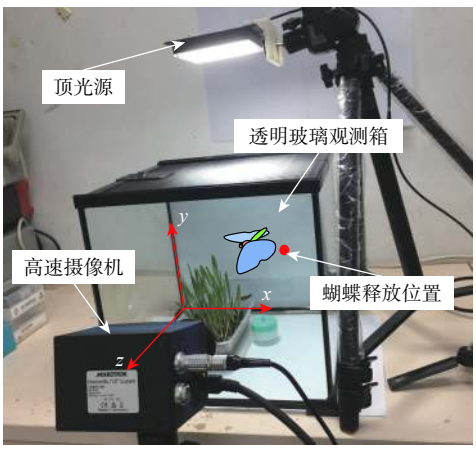


图 3　观测蝴蝶飞行的高速摄像实验系统

Fig. 3　A high-speed camera system for observing butterfly flight

与其他飞行昆虫相比, 蝴蝶具有较大的物理尺寸(翼展最大超过 8 cm), 飞行速度最高可达 2 m/s, 需要足够大的空间来保证足够的蝴蝶飞行记录时间, 因此, 本实验采用尺寸为(0.4 m×0.4 m×0.4 m)的正方体透明玻璃容器作为观测空间。由于蝴蝶扑翼运动的频率范围一般为 9～12 Hz, 需要较高的拍摄帧率, 因此, 观测空间的亮度决定了高速摄像机所拍摄频闪图像的锐度, 为此需布置足够的光源来增加相机传感器(CCD/CMOS)的进光量。红点表示蝴蝶初始的释放位置, 该位置距离地面 30 cm, 释放时使得蝴蝶身体与x轴平行, 且头部朝向x轴负方向。考虑到镜头畸变等因素, 为了保证测量精度, 需要对单摄像头进行标定, 本文采用的标定方法为直接线性变换(direct linear transform, DLT)。实验中摄像头(Motion BLITZ Cube4)参数设置如下: 帧率为 1 000 fps(帧/s), 拍摄图像的像素尺寸为 1 280×1 024。为了减少视频中特征点跟踪处理的计算量, 对采集的图像加框, 像素尺寸为 500×900, 计算量降低约 50%, 处理时间也减少一半以上。

1.4　蝴蝶前飞观测结果

本文定义前飞的状态为在一个扑打周期内蝴蝶质心在水平和竖直方向上的位移满足以下不等式$|\mathrm{d}y/\mathrm{d}x|<0.15$, 以此为标准筛选可供分析的样本。4 个观测用的特征点坐标分别为紫色的胸部特征点$(x_\mathrm{t},y_\mathrm{t})$、深蓝色的胸腹关节点$G(x_\mathrm{j},y_\mathrm{j})$、浅蓝色的腹部末端点$(x_\mathrm{a},y_\mathrm{a})$及红色的翼尖特征点$(x_\mathrm{wt},y_\mathrm{wt})$。由于实际蝴蝶左右翼翼根之间距离很小, 位置接近胸部特征点, 在观测过程中简化认为翼根与该特征点重合。跟踪结果如图 4 所示, 截取的视频片段不到 300 ms, 包含 2 个完整的扑打周期。

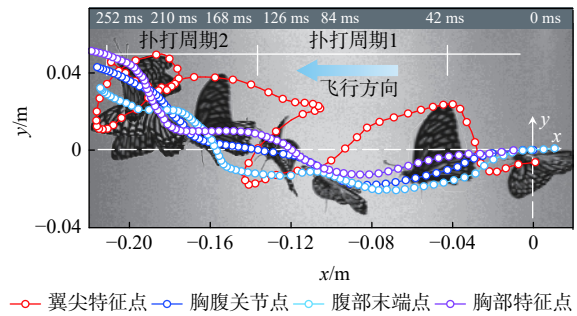


图 4　蝴蝶飞行运动中特征点跟踪结果

Fig. 4　Tracking results of feature points in butterfly flight movement

在跟踪得到蝴蝶飞行运动特征点的二维坐标时间序列的基础上, 利用有限差分法计算蝴蝶飞行时的各特征点在x(同y)方向上的运动学数值, 计算式为

$$v_{x_i}=\frac{x_{i+1}-x_{i-1}}{2\mathrm{d}t} \tag{1}$$

$$a_{x_i}=\frac{2x_{i+2}-x_{i+1}-2x_i-x_{i-1}+2x_{i-2}}{7\mathrm{d}t^2} \tag{2}$$

式中: v_{x_i}和a_{x_i}分别为特征点在图像第i帧时刻沿x轴方向的速度和加速度; x_i为特征点在第i帧时刻图像上的位置; $\mathrm{d}t$为视频相邻两帧间的时间间隔。同理可求得v_{y_i}和a_{y_i}, 根据v_{x_i}和v_{y_i}的数值可求得特征点在第i帧时刻的合速度值$v_{\mathrm{sum},i}$。

为了便于分析所有有关扑翼的时间函数曲线, 下扑阶段设置为网格背景, 上扑阶段设置为白色背景。观测到另一种特别的飞行阶段——滑翔, 用灰色背景表示, 该阶段存在于上扑结束到下一个扑打周期开始之间。蝴蝶从最初的位置释放后, 先轻轻扇动翅膀 1 到 2 个周期, 对应飞行高度略有降低, 平均飞行速度随之增加。紧接着, 胸部的平均俯仰角增加, 相当于整机的迎角增加。此后蝴蝶开始爬升, 平均飞行速度随之下降。图 5 和图 6 分别为蝴蝶翼尖特征点和胸腹关节点 G 的位移和速度曲线。蝴蝶在上扑中段附近会出现一个小的合速度峰值, 而在下扑转换为上扑时刻合速度最小。因此, 蝴蝶前飞过程中飞行速度最高可达 1.3 m/s, 且存在显著的周期波动。翼尖速度在下扑中段和上扑中段出现 2 个峰值, 最高可达 2.6 m/s。

同时, 可计算得到蝴蝶前飞过程中胸部和腹部运动的特征角θ_t和θ_a, 计算式为

$$\theta_t = \arctan\frac{y_t - y_j}{|x_t - x_j|} \tag{3}$$

$$\theta_a = \arctan\frac{y_a - y_j}{|x_a - x_j|} \tag{4}$$

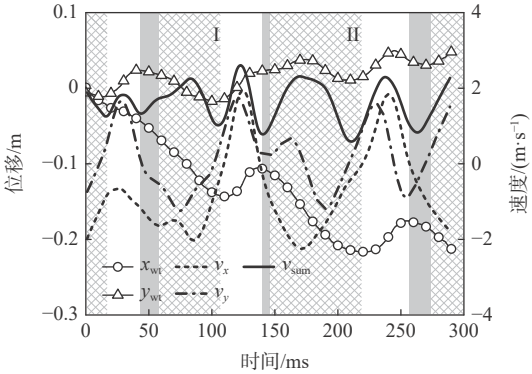


图 5 蝴蝶翼尖运动学随时间变化

Fig. 5 Time-varying kinematics of butterfly wingtip

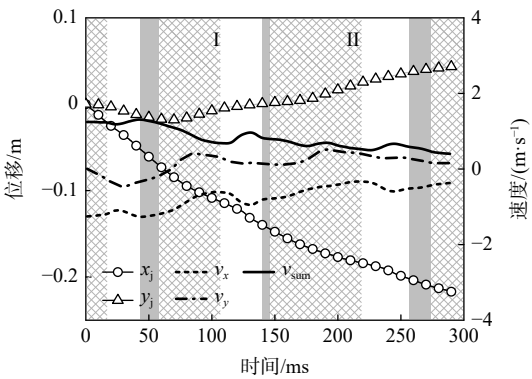


图 6 蝴蝶胸腹关节 G 运动学随时间变化

Fig. 6 Time-varying kinematics of butterfly joint G

蝴蝶通过俯仰身体来操纵扑打平面的角度，从而调整扑打运动的方向。真实蝴蝶的翼扑打运动、胸部俯仰运动和腹部摆动运动呈周期性变化。因此，蝴蝶双翼的运动可表示为扑打运动、俯仰运动及$\theta_f(t)$和$\theta_t(t)$之间的相位差γ，其中γ约为$\pi/2$或更小。胸部俯仰运动和腹部摆动之间的相位差较大约为π，如图 7 所示。

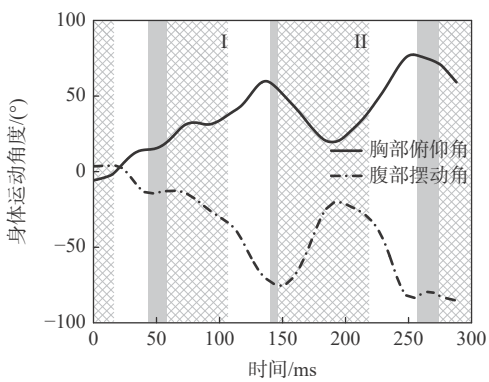


图 7 胸部俯仰角及腹部摆动角随时间变化

Fig. 7 Time-varying of thorax pitch angle and abdominal swing angle

由观测结果可知实验蝴蝶的扑打幅值约为 50°～70°，胸部俯仰角幅值约为 30°～40°，腹部摆动角幅值约为 20°～30°。由于本文中蝴蝶扫掠角在快速前飞过程变化较小，同时翼迎角变化基本与机体俯仰角相同，因此，可得到简化的蝴蝶扑翼运动方程组为

$$\begin{cases} \theta_f(t) = 60^\circ\cos(2\pi ft) + 20^\circ \\ \theta_t(t) = 30^\circ\cos\left(2\pi ft + \dfrac{\pi}{2}\right) + 34^\circ \\ \theta_a(t) = -20^\circ\cos(2\pi ft) - 35^\circ \end{cases} \tag{5}$$

式中：f为蝴蝶扑翼扑打频率。

2 仿生蝴蝶扑翼飞行器样机设计

通过生物学观测，依据生物形态学及相似原理得到仿生蝴蝶扑翼飞行器样机设计的关键参数和相应的几何模型。设置左右翼独立驱动，取消传统扑翼飞行器尾翼设计，主翼既产生升推力，又输出姿态控制力矩。仿生蝴蝶扑翼飞行器的研制目标是通过独立控制左右翼的扑打频率、相位及幅值，实现无舵面可控飞行。给定机翼拍打的起始角度和终止角度，扑打中间位置与机身水平面的夹角定义为机翼上反角，扑打中间位置又称为平衡位置。

2.1 整机设计

蝴蝶拍打翅膀是典型的节律运动，扑翼频率不高，利用伺服舵机来控制机翼拍打较为合适。将蝴蝶单侧前后翼的扑打简化为单自由度扑打模式，即伺服舵机带动前翼驱动杆摆动，扑打动作通过3D 打印结构及碳纤维杆将电机输出的旋转扭矩传递给后翼，从而使前后翼同步运动。样机的基本结构为伺服电机直驱的双翼独立扑翼机构。这种设计虽然没有了复杂的多连杆或绳驱动机构来实现扑动，但对伺服电机的控制精度和驱动力矩提出了更高的要求。对于单侧机翼采用前后翼分离设计，构成两者的碳纤维杆结构在翼根处固连，由主驱动杆同步驱动。左右翼的后翼均与“一字型”或“V 型”铰链结构相铰接，该铰链结构固连在主机体碳纤维杆上。驱动翼的动力源选择 MSK 的微型金属齿舵机，型号为 HV75K，在 7.4 V 供电条件下，输出扭矩最高可达 2.8 kg·cm，输出转速最快为每旋转 60° 需 0.10 s。该舵机设置有位置闭环反馈可实现精确的位置控制，从而确保双翼扑打幅值、频率和初始相位的精度。

图 8 为仿生蝴蝶扑翼飞行器样机的三维机械结构，主要包含左翼、右翼和机身。机身做一体化设计，包括主碳纤维杆、舵机安装支架、左右翼驱动舵机、机载飞控电子系统及微型航模电池等。样机的具体物理参数如表 1 所示。样机翼展约为 62.0 cm，最大弦长约为 38.0 cm。

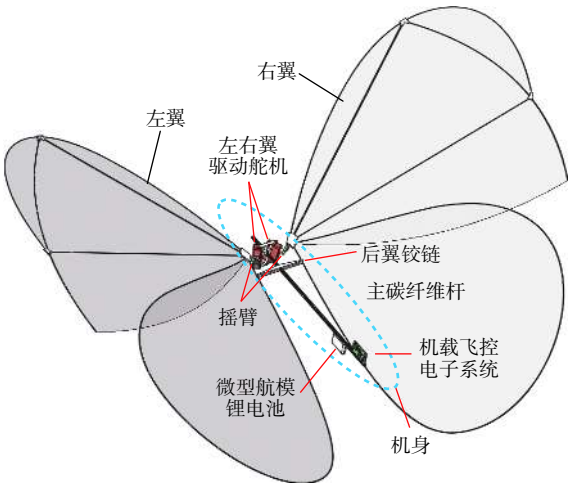


图 8　仿生蝴蝶扑翼飞行器三维机械结构

Fig. 8　3D mechanical structure of butterfly-inspired flapping wing air vehicle

表 1　样机物理参数

Table 1　Physical parameters of prototype

参数	数值
翼展/cm	62.0
最大弦长/cm	38.0
扑打频率范围/Hz	1.8～3.2
前翼前掠角/(°)	45
前翼翼面积/cm^2	575.7
后翼翼面积/cm^2	558.1
最大平飞速度/(m·s^{-1})	1.5
起飞质量/g	39.60

主碳纤维杆为截面尺寸为 2.5 mm×2.5 mm 的方形中空杆, 长度为 250 mm。机翼在扑翼机构带动下既产生升力又作为舵面对飞行姿态进行直接调整。机翼的平衡位置和机身水平面重合, 即机翼上反角设置为 0º。后翼内缘碳纤维杆与机身通过后翼连接铰链铰接, 因此, 机翼绕舵机输出轴旋转时候, 翼内缘碳纤维杆形成一个锥面。改变该锥面对称面相对机身水平面的夹角关系, 可改变样机产生的升推力矢量。经过多轮迭代, 最终制作的仿生蝴蝶扑翼飞行器样机的起飞质量约为 39.60 g, 样机各组件质量及对应占整机质量百分比如表 2 所示。

2.2　仿生翼设计及制作工艺

本文仿生蝴蝶扑翼飞行器采用碳纤维框架与热塑性聚氨酯弹性体材料 (thermoplastic polyurethanes, TPU) 薄膜组成的机翼实现飞行, 其制作工艺是整机研发过程中最为复杂和关键的环节。机翼由前翼和后翼构成, 前翼和主驱动杆相连, 为主动机构; 后翼跟随前翼运动, 为从动机构。前翼的翼展大, 且前翼后缘和后翼前缘有部分重叠。机翼基本物理特性和制作工艺要求如下: ①机翼几何外形仿生, 采用前后翼分离设计, 质量轻且具有合适的强度和弹性; ②工艺要求机翼能够成型快且制作成本低, 保证左右翼完全对称。图 9 为仿生翼制作工艺原理图。

表 2　样机各组件质量分布

Table 2　Mass distribution of each component of prototype

组件	质量/g	占比/%
左右翼驱动舵机	14.80	37.37
左翼	7.55	19.07
右翼	7.51	18.96
主碳纤维杆	0.63	1.59
机载飞控电子系统	2.66	6.72
微型航模电池	5.30	13.38
线材	0.85	2.15
其他	0.30	0.76

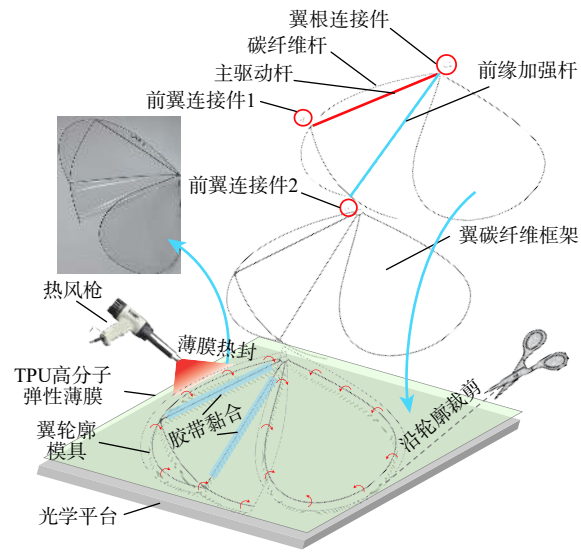


图 9　仿生翼制作工艺原理图

Fig. 9　Schematic illustration of the bionic wing fabrication

具体制作工艺流程包括以下 4 个步骤:

步骤 1　利用光固化 3D 打印技术快速完成树脂翼轮廓模具的制作, 将其固定于金属光学平台上。把尺寸与光学平台相近的矩形 TPU 翼膜展平、绷紧放置于模具上。

步骤 2　根据机翼几何参数, 截取相应长度的碳纤维杆, 沿模具上的凹槽轮廓进行弯曲, 将树脂连接件按照设计位置与碳纤维杆黏接。将整个框架嵌入模具凹槽中, 用夹具固定好。

步骤 3　根据前后翼的碳纤维外轮廓, 外延 20 mm 用剪刀进行裁剪, 将多出的薄膜沿着翼外轮廓对折, 用热风枪进行热封, 使得折叠的膜与翼轮廓内边缘的膜黏合到一起。

步骤 4　用胶带将主驱动杆、前缘加强杆与翼膜黏合到一起完成整翼的制作。

以此工艺制作的左右翼特征尺寸相差不超过 3 mm, 实际单侧仿生翼质量不超过 8 g, 左右翼质量相差不超过 0.05 g。

2.3 机载飞控系统设计

2.3.1 机载电子系统硬件

由于仿生蝴蝶扑翼飞行器载荷能力有限, 为满足轻量化要求, 本文设计了不到 3 g 的超轻机载飞控电子系统。系统由 STM32F411 芯片及相关外围电路组成飞控核心, 集成了 915 MHz 的无线通讯模块, 2.4 GHz 的无线数据传输模块, MPU-6000 的惯性测量单元, 设置的输出控制端口包括 4 路大功率伺服舵机控制端口, 如图 10 所示。由主控芯片输出脉冲亮度调制(pulse width modulation, PWM)的信号快速精准控制舵机的转动, 通过多主从串行通讯总线(inter-integrated circuit, I^2C)连接存储器 BL 24C02; 通过全双工同步串行总线(serial peripheral interface, SPI)建立与 MPU-6000 的快速数据传输; 通过串口通讯(universal synchronous/asynchronous receiver/transmitter, USART)与无线微控制器 SI1000 和 2.4 GHz 高性能射频收发器 CC2500 构建数据交互链路。

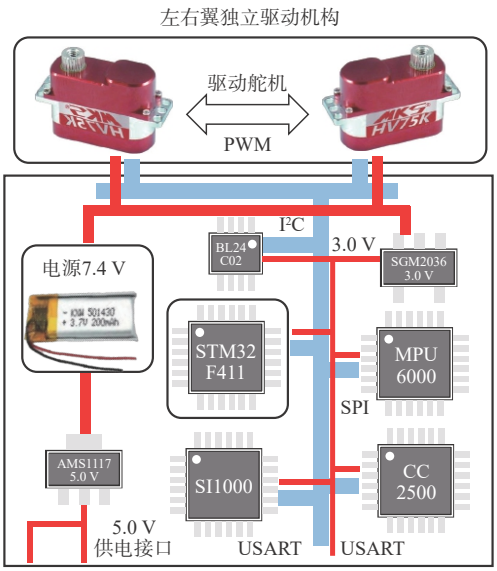


图 10 机载飞控电子系统结构示意图

Fig. 10 Schematic diagram of onboard flight control electronic system

由于本文设计的仿生蝴蝶扑翼飞行器应用于飞行距离较短的场景, 一般为距离地面基站 50 m 以内的范围, 采用基于 CC2500 为核心的 2.4 GHz 无线数据传输模块, 其最高传输速率达 250 kbps。同时, 使用 2.4 GHz 的 Futaba T6J 遥控器和飞行器建立无线数据链路。

2.3.2 PD 控制器设计

仿生扑翼飞行器的飞控系统中, 比例微分(proportional differential, PD)控制器被广泛应用, 控制效果得到验证[16-17], 对非线性被控对象的控制具有良好的鲁棒性。本文针对仿生蝴蝶扑翼飞行器设计的基于姿态反馈的 PD 控制器如图 11 所示, 姿态传感器感知当前样机的姿态角(ϕ,θ,ψ)和相应的姿态角速率(q,p,r), 其中ϕ为滚转角, θ为俯仰角, ψ为偏航角。经过滤波处理后反馈, 样机的双翼均由伺服舵机驱动, 控制器输出 PWM 波控制舵机旋转的位置、幅值及速度。

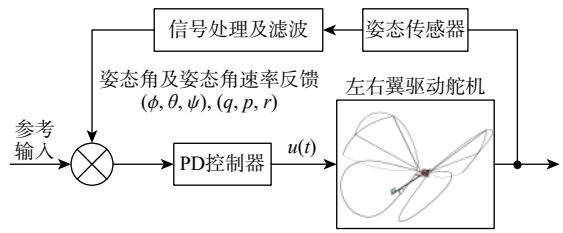


图 11 基于姿态反馈的 PD 控制器框图

Fig. 11 Block diagram of PD controller based on attitude feedback

针对扑翼飞行器的连续系统, PD 控制律方程表示为

$$u(t)=k_{\mathrm{p}}e(t)+k_{\mathrm{d}}\frac{\mathrm{d}}{\mathrm{d}t}e(t) \tag{6}$$

式中:$u(t)$为 PD 控制器输出; $e(t)$为预期的控制参量与传感器实际测量得到数值做差得到的误差变量; k_{p}为比例系数; k_{d}为微分系数。假设采样周期为 T, 在k时刻, $e(t)$对应的采样值为$e(k)$, 式 (6) 中微分计算用一阶向后差分近似替代, 则本文 PD 控制律方程的离散化可表示为

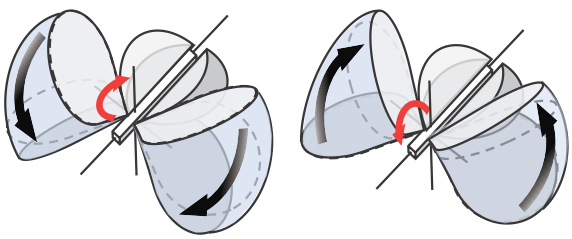
$$u(k)=k_{\mathrm{p}}e(k)+k_{\mathrm{d}}[e(k)-e(k-1)] \tag{7}$$

实验中, 通过调节k_{p}与k_{d}的数值, 实现样机的稳定操控。

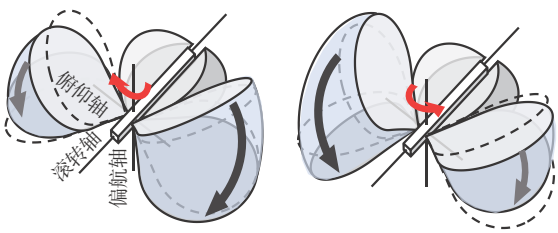
2.4 飞行策略

蝴蝶本身的无尾翼的生物形态学特征已经决定了其飞行操控的独特性。仿生蝴蝶扑翼飞行器为欠驱动系统, 是典型的独立控制变量个数小于系统自由度个数的非线性系统, 具有欠驱动系统固有的优点: 结构简单、质量轻、成本低、系统灵活度高等。飞行器左右翼由左右独立伺服电机驱动, 控制变量为左右翼扑打幅值、振翅频率、拍打的起始和末端角度。左右翼既为输出动力源, 又产生姿态调节力矩, 整机仅依靠左右翼拍打运动实现纵向及横侧向飞行操控。本文尝试将仿蝴蝶的飞行操控策略简化, 利用俯仰运动和偏航运动来实现样机在空间内的基本飞行操控。考虑到滚转运动在蝴蝶的飞行中是一种高度耦合的模态, 本文忽略滚转运动。具体样机俯仰运动控制及偏航运动控制的实现方法如图 12 所示。图 12(a)中左右翼相对机身

纵轴面对称拍打, 扑打频率保持不变, 样机的俯仰姿态变化是在平飞的基础上进行的。需要抬头时, 保持扑打的下极限位置不变, 增大扑翼幅值; 需要低头时, 保持扑打的上极限位置不变, 减小扑翼幅值。俯仰姿态改变后会产生爬升或下降的趋势, 飞行的高度也会随之改变。图 12(b)中左右翼相对机身纵轴非对称扑打, 扑打频率保持不变, 机翼上反角+15°保持不变。通过改变左右翼的扑打幅值来改变左右翼受到的气动力矩。向左偏航, 右翼扑打幅值大于左翼扑打幅值; 向右偏航, 左翼扑打幅值大于右翼扑打幅值。通过调节左右翼扑打的幅值差可以实现不同曲率的偏航运动。



(a) 俯仰运动控制



(b) 偏航运动控制

图 12　仿生蝴蝶俯仰及偏航运动控制示意图

Fig. 12　Schematic diagram of pitch and yaw motion control of prototype

3　实验验证

3.1　地面动力测试

为了得到实际样机的基本气动特性, 本文搭建了零来流地面动力实验系统。测试样机的仿生翼直接由舵机驱动, 驱动电压为 3.7～8.2 V。样机机身主碳纤维杆与地面水平, 且翼旋转轴水平设置, 在存在气动载荷的情况下进行。实验系统设置如图 13 所示, 主要硬件组成部分为: 测试台体、传感器连接结构、多通道高速数据采集卡(NI USB-6229)、六维力传感器、传感器信号调理盒、信号转接板、上位机 PC、示波器、稳压源及实验样机双翅扑打控制器。由于蝴蝶整机重量小于 0.4 N, 选用具有较高的测量精度和稳定性的 ATI 公司 nano 17 (SI-12-0.12)六维力传感器, 其三轴力的量程为: F_x和F_y为±12 N, F_z为±17 N, T_x、T_y和T_z三轴力矩的量程为±120 N·mm, 力测量精度为 1/320 N, 力矩测量精度为 1/64 N·mm, 符合实验要求。

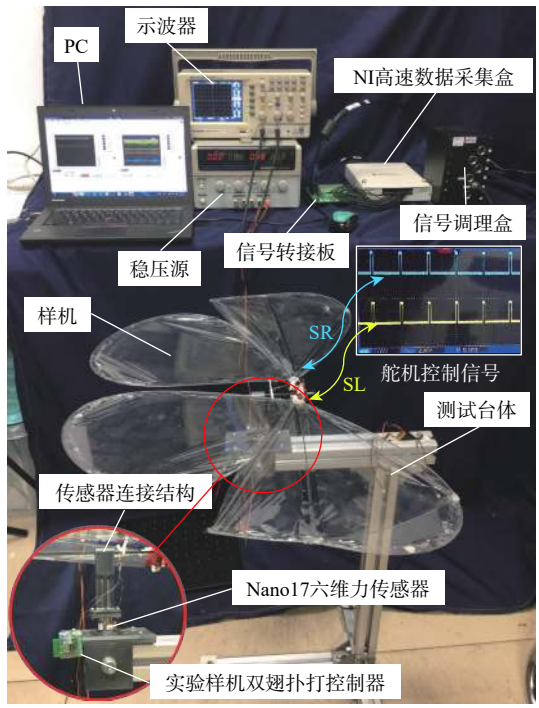


图 13　样机地面动力学实验系统

Fig. 13　Ground dynamics experiment system of prototype

实验过程中, 仿生蝴蝶扑翼飞行器的扑打频率在 2 Hz 左右, 因此, 采样频率设置为 1 kHz, 数据存储频率设置为 100 Hz, 采集的气动力和力矩数据保留 3 位有效数字。示波器可以测得实验样机控制器对左右驱动舵机输出的可调占空比 PWM 波舵机控制信号, 分别是右舵机信号 SR 和左舵机信号 SL。

在无空气来流的条件下进行试验, 输入 7.4 V 电源驱动, 样机的拍打频率暂定为 2 Hz(即拍打周期为 500 ms), 扑打幅度设置为 105°, 翼上反角设置为 0°(即扑翼平衡位置与水平面重合)。此时, 瞬时的推力和升力如图 14 所示, 选取 0～1.4 s 的数据显示, 其中推力平均值约为 0.238 N, 升力平均值约为 0.039 N, 推力峰值约为 0.789 N, 升力峰值约为 1.192 N。由此可知, 机身固定状态下测试所得升力在一个扑打周期内的平均值远小于飞行器重力, 上扑和下扑产生的气动力相互抵消, 这说明胸部俯仰运动对蝴蝶扑翼升力的产生至关重要。

由于样机的翼在扑打过程中产生被动变形, 翼沿着展向发生扭转, 出现了类似上下沉浮加俯仰的运动, 因此, 在上扑和下扑阶段均有相应的推力产生。图 14(a)中一个扑打周期内, 瞬时推力存在 4 个正峰值和一个负峰值, 其中负峰值仅出现在下扑转上扑或上扑转下扑的时刻附近, 正峰值出现在下扑和上扑中段左右。在翼扑打经过平衡位置时, 由于前翼的变形, 前翼后缘与后翼前缘存在相对运动关系, 之间产生气隙可能会影响推力的产生。图 14(b)瞬时升力仅存在一个正峰值和一个负峰值, 均出现

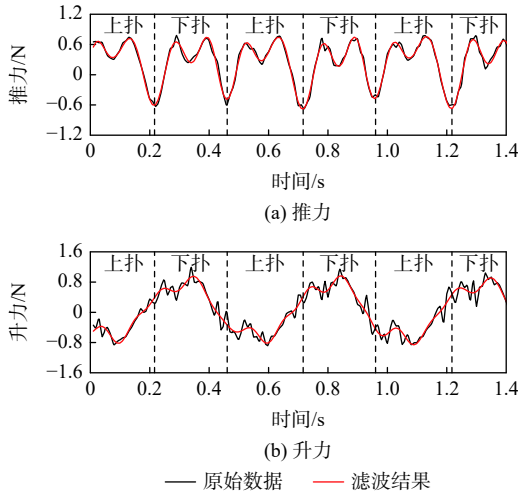


图 14 样机瞬时升力及推力曲线

Fig. 14 Instantaneous lift and thrust of the prototype

在机翼经过平衡位置时：当翼下扑经过平衡位置时，升力为正峰值；当翼上扑经过平衡位置时，升力为负峰值。

扑打频率是决定样机飞行表现的重要参数，最优的扑打频率需要通过实验来确定，同样输入 7.4 V 电源驱动，翼上反角设置为 15°（即扑翼平衡位置与水平面夹角为 15°），扑打幅度设置为 90°，改变扑打频率，得到样机所受升推力和俯仰力矩的曲线（见图 15）。推力和俯仰力矩随着扑打频率的提高逐渐减小，升力则先升高后减小，综合升推力的表现，本文选择 2.25 Hz 作为飞行器的最优扑打频率。

扑打幅度和舵机驱动电压对样机的飞行性能也存在影响。在扑打频率为 2 Hz、上反角为 0°的条件下，改变扑打幅度，分别测量驱动电压为 5.0 V 和 7.4 V 下的升推力情况，如图 16 所示。2 种电压驱动下样机的升力几乎没有变化。5 V 供电时推力随着扑打幅度增大始终维持较低值，平均约为 0.126 N，原因是这种情况下舵机输出扭矩不足且响应速度相对较慢。而 7.4 V 供电时，推力明显提高，

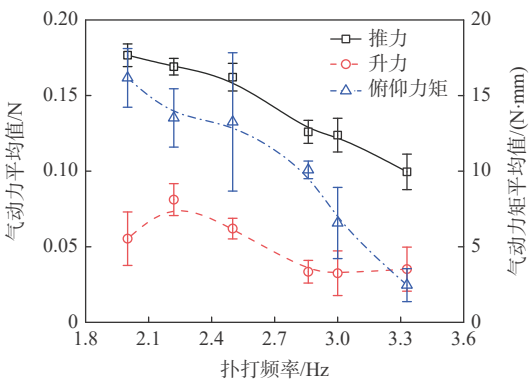


图 15 扑打频率对于扑翼气动性能的影响

Fig. 15 Influence of flapping frequency on aerodynamic performance of flapping wings

且随着扑打幅度的增大先增大后减小，减小的原因可能是扑翼幅度过大时舵机响应跟不上。因此，样机采用 2 节锂电池串联 7.4 V 电压供电。

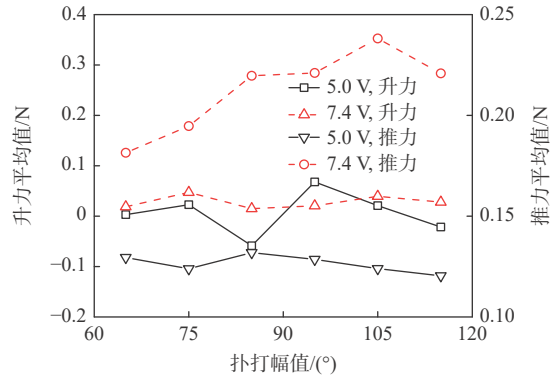


图 16 不同供电电压下样机升推力的测试结果

Fig. 16 Test results of lift and thrust of prototype under different supply voltages

在固定机体的动力测试条件下，扑打幅度变化对升力的影响甚微，而推力随幅值增大有一定的提升，结合舵机的响应速度，将样机的扑打幅度确定为 105°，其周期平均推力达到最大值，此时样机的扑打范围为−35°～70°。

3.2 自由飞测试

对样机控制器加载实验得到优化的控制参量进行样机自由飞行测试。样机以手掷方式起飞，起飞时从上扑的极限位置开始向下扑打。利用单摄像系统对样机进行了运动学跟踪分析，图 17 为放飞后 2 个扑打周期样机上特征点（分别是头部端点和翼尖）的跟踪结果。跟踪过程中存在极少数特征点被遮挡的情况，通过手动补点完善。

可见，样机实现了约 100°的大幅度扑翼运动，后翼在前翼带动下上下翻转，飞行速度接近 1.5 m/s。样机翼的扑打和胸部的俯仰也存在一定的相位差，与自然蝴蝶类似，在扑翼飞行过程中，身体存在竖直方向上的波动。

图 18 为样机遥控自由前飞的测试结果，最终实现了近 5 min 的滞空飞行。与俯仰操控相比，样

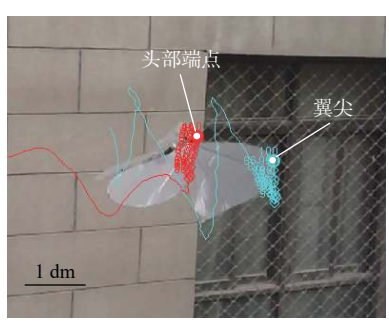


图 17 样机飞行中特征点跟踪结果

Fig. 17 Feature point tracking results during prototype flight

机对偏航指令的响应更为迅速和显著。实验中, 样机虽然飞行姿态较为稳定, 但存在轻微的掉高现象, 说明扑翼产生的升力稍显不足, 可以从整机轻量化或优化扑翼运动学参数 2 个角度出发进一步提高样机的升推力。



图 18　样机前飞逐帧图像

Fig. 18　Stroboscopic images of prototype during free forward flight

4　结　论

1) 利用高速摄像机对斑凤蝶的飞行行为进行了生物学观测, 明确了前后翼在扑翼过程中相对位置关系, 并着重分析了扑翼运动、胸部俯仰运动及腹部摆动运动之间的相位关系, 得到了简化的蝴蝶前飞运动学模型。

2) 根据仿生学原理, 样机以斑凤蝶为原型, 本着整机轻量化的原则提出具有类似蝴蝶翅脉的“杆-膜”仿生翼的新设计及新工艺, 保证了左右翼结构的对称性和质量属性的一致性。

3) 研制了具有姿态闭环反馈的超轻机载飞控电子系统, 利用六维力传感器对样机做地面动力学测试, 通过摄像机跟踪记录样机试飞的空间运动学数据, 实验验证了基于特征运动状态的前飞规律的有效性。

综上所述, 真实蝴蝶的飞行是复杂且不稳定的, 仿蝴蝶翼扑动会引起身体显著的被动俯仰。所研制的仿生蝴蝶扑翼飞行器, 在一定程度上验证了大翼面积、小展弦比及低频拍打的新扑翼模型的可行性。最终, 样机实现了前飞及横侧向可控飞行。未来, 对蝴蝶飞行机理及仿生蝴蝶扑翼飞行器气动性能优化有待更为深入的研究。腹部的主动摆动影响将被进一步考虑, 并且对样机增加腹部摆动机构, 以此验证腹部对蝴蝶独特扑翼飞行的作用。

参考文献（References）

[1] 孙茂. 昆虫飞行的空气动力学[J]. 力学进展, 2015, 45(1):201501.
SUN M. Aerodynamics of insect flight[J]. Advances in Mechanics, 2015, 45(1): 201501 (in Chinese).

[2] DUDLEY R. Biomechanics of flight in neotropical butterflies: Morphometries and kinematics[J]. Journal of Experimental Biology, 1990, 150(1): 37-53.

[3] SRIDHAR M, KANG C K, LANDRUM D B. Beneficial effect of the coupled wing-body dynamics on power consumption in butterflies[C]//AIAA Scitech 2019 Forum. Reston: AIAA, 2019.

[4] TWIGG R, SRIDHAR M, POHLY J A, et al. Aeroelastic characterization of real and artificial monarch butterfly wings[C]//AIAA Scitech 2020 Forum. Reston: AIAA, 2020.

[5] BETTS C R, WOOTTON R. Wing shape and flight behaviour in butterflies (Lepidoptera: papilionoidea and hesperioidea): A preliminary analysis[J]. The Journal of Experimental Biology, 1988, 138: 271-288.

[6] BRODSKY A K. Vortex formation in the tethered flight of the peacock butterfly Inachis io L. (Lepidoptera, Nymphalidae) and some aspects of insect flight evolution[J]. Journal of Experimental Biology, 1991, 161(1): 77.

[7] DUDLEY R. Biomechanics of flight in neotropical butterflies: Aerodynamics and mechanical power requirements[J]. The Journal of Experimental Biology, 1991, 159: 335-357.

[8] SRYGLEY R B, THOMAS A L R. Unconventional lift-generating mechanisms in free-flying butterflies[J]. Nature, 2002, 420(6916): 660-664.

[9] HUANG H, SUN M. Forward flight of a model butterfly: Simulation by equations of motion coupled with the Navier-Stokes equations[J]. Acta Mechanica Sinica, 2012, 28(6): 1590-1601.

[10] IMURA T, FUCHIWAKI M, TANAKA K. Dynamic behaviors of butterfly wing and their application to micro flight robot[C]//Proceedings of ASME 2009 Fluids Engineering Division Summer Meeting. New York: ASME, 2010: 1687-1694.

[11] TANAKA H, SHIMOYAMA I. Forward flight of swallowtail butterfly with simple flapping motion[J]. Bioinspiration & Biomimetics, 2010, 5(2): 026003.

[12] FUJIKAWA T, SATO Y, MAKATA ,et al. Development of a butterfly-style flapping robot with lead-lag mechanism using f;exible links(mechanical systems)[J]. Nippon Kikai Gakkai Ronbunshu, C Hen/Transactions of the Janpan Society of Mechanical Engineers, Part C, 2010, 76(700):2663-2670.

[13] FUJIKAWA T, HIRAKAWA K, OKUMA S, et al. Development of a small flapping robot[J]. Mechanical Systems and Signal Processing, 2008, 22(6): 1304-1315.

[14] FRONTZEK H, KNUBBEN E, MUGRAUER R, et al. eMotionbutterflies: Ultralight flying objects with collective behavior: FESTO Report 50058[R]. Esslingen: FESTO, 2015.

[15] CHI X K, WANG S P, ZHANG Y X, et al. A tailless butterfly-type ornithopter with low aspect ratio wings[C]//CSAA/IET International Conference on Aircraft Utility Systems. London: IET, 2018: 1-6.

[16] PHAN H V, AURECIANUS S, KANG T, et al. KUBeetle-S: An insect-like, tailless, hover-capable robot that can fly with a low-torque control mechanism[J]. International Journal of Micro Air Vehicles, 2019, 11: 1756829319861371.

[17] NGUYEN Q V, CHAN W L. Development and flight performance of a biologically-inspired tailless flapping-wing micro air vehicle with wing stroke plane modulation[J]. Bioinspiration & Biomimetics, 2018, 14(1): 016015.

Mechanism of butterfly forward flight and prototype verification based on characteristic motion observation

ZHANG Yixin[1], WANG Xingjian[1, 2, 3], WANG Shaoping[1, 2, 3, *], CHI Xiaokai[1], DU Shaoyang[1]

(1. School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China;
2. Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, Beihang University, Beijing 100191, China;
3. Ningbo Institute of Technology, Beihang University, Ningbo 315800, China)

Abstract: To investigate the mechanism of butterfly flapping wing flight and develop a low-frequency flapping wing bionic robot, the fligh kinematics of the butterfly are capturedand documented using a high-speed camera, three characteristic motion states of butterflies are proposed: flapping wing motion, thorax pitching motion and abdomen swinging motion. Through analysis of the phase relationship among these three states, a kinematics model of the butterfly forward flight is constructed. Then based on the new procedure of the "rod-membrane" bionic wings and a miniature onboard flight control system, a lightweight butterfly-inspired flapping wing air vehicle is developed, the flight control strategy of which is studied as well. Next, a six-dimensional force sensor is used to test the dynamics of the prototype on the ground, and a high-speed camera is used to track the flight of the prototype, which proves the effectiveness of the development of the prototype based on the characteristic motion states of the butterfly forward flight mechanism.

Keywords: butterfly flight mechanism; typical characteristic motion; flapping wing air vehicle; bionic robot; membrane wing fabrication

Received: 2021-08-28; **Accepted**: 2021-09-03; **Published Online**: 2021-09-13 17:28

URL: kns.cnki.net/kcms/detail/11.2625.V.20210913.1527.002.html

Foundation items: National Science and Technology Major Project (2017-V-0010-0060); National Natural Science Foundation of China (51620105010,51805026,51675019); National Basic Research Program of China (JCKY2018601C107)

* **Corresponding author**. E-mail: wangshaoping@vip.sina.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0478

利用北斗GEO卫星干涉信号功率反演土壤湿度

汉牟田，许志超*，常青，张波，王峰，洪学宝
(北京航空航天大学 电子信息工程学院，北京 100191)

摘 要：对利用北斗导航系统地球静止轨道（GEO）卫星干涉信号功率反演土壤湿度的方法进行研究。针对现有研究主要建立纯经验模型反演土壤湿度、反演方法单一的问题，提出一种半经验的反演方法。所提方法针对北斗GEO卫星星地之间几何构型较为稳定的特点，利用相邻两天的干涉幅度信息抵消发射信号功率的影响，实现了土壤反射系数的反演。基于交叉极化影响下的反射系数理论模型构建了一种半经验的土壤湿度反演模型，通过仿真与实验对所提方法和模型进行验证。仿真结果表明：所提方法和模型能更好地适应反演过程中出现的非线性情况。实验结果表明：土壤湿度反演的均方根误差小于0.02 cm^3/cm^3，相关系数超过0.8。

关 键 词：北斗导航系统；地球静止轨道卫星；干涉信号功率；反射系数；土壤湿度
中图分类号：V19; TP79
文献标志码：A **文章编号**：1001-5965(2023)07-1661-10

全球导航卫星系统反射信号(global navigation satellite system-remote, GNSS-R)测量土壤湿度技术是随全球导航卫星系统(global navigation satellite system, GNSS)发展而衍生出的一种新型遥感技术[1-3]，该技术的出现丰富了现有的微波遥感手段。

GNSS-R技术经过近30年的发展，目前主要有2大技术分支，①常规GNSS-R技术，该技术使用分立天线分别接收GNSS直射信号与土壤反射的GNSS信号，通过对直射信号与反射信号的协同处理以反演土壤湿度，该技术于1993年由Martín-Neira[4]提出，最初被应用于海洋遥感[4-7]；②GNSS干涉测量技术，该技术利用GNSS直射信号与反射信号间的干涉效应测量土壤湿度，该技术于2008年由美国科罗拉多大学的Larson等[8]提出，由于测量时只使用一支GNSS天线同时接收直射与反射信号，在硬件配置和应用模式上与常规GNSS接收机相同，从而可以利用现有的GNSS基础设施如连续运行参考站(continuously operating reference system, CORS)等开展土壤湿度观测[9]，因此，该技术一经提出即得到广泛关注[10-12]，是本文研究的重点。

在GNSS卫星绕地运行过程中，GNSS直射信号与反射信号传播路径不同，二者在到达天线时存在相位差，该相位差随卫星运动发生周期性变化，从而使直反射信号发生周期性干涉相长与相消，最终导致接收机记录的载噪比数据随时间变化呈现出类余弦函数状的干涉图样。研究表明，干涉图样主要出现在卫星高度角较低(小于30°)时，且土壤湿度发生变化将引起载噪比数据的震荡频率、幅度、初始相位发生相应变化，据此可反演土壤湿度[13]。目前对于干涉测量技术的研究主要利用GPS卫星信号，GPS卫星为中轨道(medium Earth orbit, MEO)卫星，绕地球做圆周运动，地球表面观测到MEO卫星最大高度角变化可达80°以上，能够引起较大的直反射信号相位差变化，干涉图样特征将十分明显，缺点是无法实现对固定区域的连续观测。

中国的北斗导航系统创新性地使用地球静止

收稿日期：2021-08-20；录用日期：2021-11-20；网络出版时间：2021-11-30 13:16
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20211129.2012.010.html
基金项目：国家自然科学基金(41774028)；博士后创新人才支持计划(BX20200039)
*通信作者. E-mail：by1602112@buaa.edu.cn

引用格式：汉牟田，许志超，常青，等. 利用北斗GEO卫星干涉信号功率反演土壤湿度[J]. 北京航空航天大学学报，2023，49(7)：1661-1670. HAN M T，XU Z C，CHANG Q，et al. Soil moisture retrieval using Beidou GEO satellite interference signal power[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1661-1670 (in Chinese).

轨道(geostationary Earth orbit, GEO)卫星播发导航信号,由于其相对地球静止,可实现对固定区域的不间断观测。实际上 GEO 卫星并非完全相对静止,而是处于一种周期性的缓慢微动状态[14],地球表面观测到 GEO 卫星 24 h 内高度角变化不超过 4°。文献 [15-16] 利用该特点开展了 GEO 卫星反射信号测量土壤湿度研究,提出一些经验的或解析的土壤湿度反演模型;文献 [14] 提出一种消除 GEO 卫星微动效应的方法用于反演土壤湿度。然而以上研究均基于常规 GNSS-R 技术,目前国内外对于利用 GEO 卫星干涉信号进行土壤湿度反演的研究尚少:文献 [17-18]建立了 GEO 卫星载噪比数据的震荡幅度/载波相位延迟与实测土壤湿度之间的一阶/二阶多项式模型用于反演土壤湿度;文献 [19] 研究了 GEO卫星载噪比数据震荡幅度的提取方法,并建立线性模型用于反演土壤湿度,但以上方法均为纯经验的反演方法,未能充分发挥理论模型对实际反演的指导作用。

针对现有研究的不足,本文根据 GEO 卫星的特点提出一种半经验的土壤湿度反演方法,首先对北斗 GEO 卫星载噪比模型进行合理近似,据此从北斗 GEO 卫星载噪比数据中反演土壤反射系数,最后根据土壤反射系数的理论模型反演土壤湿度。为了证明本文方法的有效性,同时开展了仿真与实验验证。

1 GEO 卫星干涉信号功率模型

GEO 卫星干涉信号功率模型符合通用 GNSS 干涉信号功率模型,假设土壤表面光滑,则 GNSS 干涉测量应用场景如图 1 所示。图中:H 为天线相位中心到土壤中某一等效反射面的垂直高度,G为 GNSS 天线增益,单位为 dB,θ为卫星高度角。直射信号与经过土壤反射的信号在天线处发生干涉形成一个综合信号,该综合信号被接收机跟踪处理并测量得到其瞬时功率,即干涉信号功率S,但通常情况下,商业接收机不直接输出信号功率而是输出信号的载噪比C/N_0,二者相差一个噪声功率谱密度N_0,一般认为N_0为常数,在忽略N_0的情况下,干涉信号功率可表示为

$$S = P_{\mathrm{d}} + P_{\mathrm{r}} + 2\sqrt{P_{\mathrm{d}}P_{\mathrm{r}}}\cos\left(\Delta\varphi_{\mathrm{p}} + \varphi_0\right) \tag{1}$$

$$\Delta\varphi_{\mathrm{p}} = \frac{4\pi H\sin\theta}{\lambda} \tag{2}$$

式中:$\Delta\varphi_{\mathrm{p}}$为直射信号与反射信号间的路径差导致的相位差;$\varphi_0$为干涉信号功率的初始相位,包含除$\Delta\varphi_{\mathrm{p}}$之外的其他相位差;$\lambda$为 GEO 卫星信号波长;$H$决定了干涉信号的震荡频率$f = 2H/\lambda$;$P_{\mathrm{r}}$为反射信号功率;$P_{\mathrm{d}}$为直射信号功率,可进一步表示为

$$P_{\mathrm{d}} = CG_{\mathrm{R}}(\theta) \tag{3}$$

式中:C为发射信号功率;$G_{\mathrm{R}}(\cdot)$为天线的右旋圆极化(right-handed circular polarization, RHCP)增益。

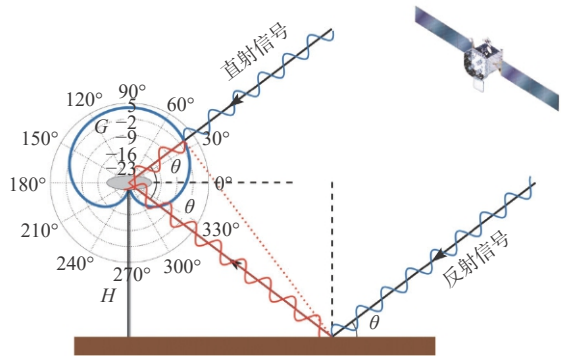


图 1 干涉测量场景

Fig. 1 Scenario of interferometric measurement.

RHCP 化直射信号经土壤反射后变为椭圆极化,可以被分解为 RHCP 分量与左旋圆极化(left-handed circular polarization, LHCP)分量。为了保证最佳的直射信号接收效果,GNSS 天线被设计为 RHCP(也称为主极化),但现实中卫星的发射天线与接收机天线并非理想的 RHCP,而是具有一定轴比的椭圆极化,因此,反射信号的 LHCP 分量(也称为交叉极化分量)也会进入天线中。文献 [20] 在建立干涉信号功率物理模型时考虑了天线交叉极化的影响,并给出了模型的推导思路,经过进一步推导得到反射信号功率P_{r}的最终表达式为

$$\begin{aligned} P_{\mathrm{r}} =& C[|\Gamma_{\mathrm{rR}}|^2 G_{\mathrm{R}}(-\theta) + |\Gamma_{\mathrm{rL}}|^2 G_{\mathrm{L}}(-\theta) + \\ & 2\sqrt{|\Gamma_{\mathrm{rR}}\Gamma_{\mathrm{rL}}|^2 G_{\mathrm{R}}(-\theta)G_{\mathrm{L}}(-\theta)}\times \\ & \cos(\varphi_{\mathrm{rR}} - \varphi_{\mathrm{rL}})] \end{aligned} \tag{4}$$

式中:Γ_{rR}与Γ_{rL}分别为 RHCP 和 LHCP 分量反射系数;φ_{rR}与φ_{rL}分别为 RHCP 和 LHCP 反射系数相位;$G_{\mathrm{L}}(\cdot)$为 GNSS 天线的 LHCP 增益;高度角θ前的负号表示反射信号从天线底部进入天线。$G_{\mathrm{L}}(\cdot)$的大小取决于天线设计,常用极化隔离度、轴比等天线指标描述[21],不同类型、不同厂商的 GNSS 天线$G_{\mathrm{L}}(\cdot)$大小不同,测绘型高精度天线通常对交叉极化抑制较好,故$G_{\mathrm{L}}(\cdot)$较小。

Γ_{rR}与Γ_{rL}是高度角与土壤复相对介电常数的函数,表达式分别为[22]

$$\Gamma_{\mathrm{rR}} = \frac{(\varepsilon_{\mathrm{r}} - 1)\cos^2\theta}{\left(\varepsilon_{\mathrm{r}}\sin\theta + \sqrt{\varepsilon_{\mathrm{r}} - \cos^2\theta}\right)\left(\sin\theta + \sqrt{\varepsilon_{\mathrm{r}} - \cos^2\theta}\right)} \tag{5}$$

$$\Gamma_{\mathrm{rL}} = \frac{(\varepsilon_{\mathrm{r}} - 1)\sin\theta\sqrt{(\varepsilon_{\mathrm{r}} - \cos^2\theta)}}{\left(\varepsilon_{\mathrm{r}}\sin\theta + \sqrt{\varepsilon_{\mathrm{r}} - \cos^2\theta}\right)\left(\sin\theta + \sqrt{\varepsilon_{\mathrm{r}} - \cos^2\theta}\right)} \tag{6}$$

式中:ε_{r}为土壤的复相对介电常数。ε_{r}主要为土壤湿度的函数,二者之间的关系可由成熟的微波波

段土壤介电模型描述如 Hallikainen 模型[23-24], 表达式为

$$\varepsilon' = 2.862 - 0.012S_s + 0.001C_s + (3.803 + 0.462S_s - 0.341C_s)m_v + (119.006 - 0.500S_s - 0.633C_s)m_v^2 \quad (7)$$

$$\varepsilon'' = 0.356 - 0.003S_s - 0.008C_s + (5.507 + 0.044S_s - 0.002C_s)m_v + (17.753 - 0.313S_s - 0.206C_s)m_v^2 \quad (8)$$

式中: ε'与ε''分别为土壤复相对介电常数的实部与虚部; S_s与C_s分别为土壤中砂土与黏土的质量百分比; m_v为体积土壤湿度(后文简称土壤湿度), cm^3/cm^3。图 2 为不同土壤湿度下, RHCP 分量与 LHCP 分量反射系数随卫星高度角的变化情况。可知 RHCP 分量反射系数反比于土壤湿度和卫星高度角, 而 LHCP 分量反射系数正比于土壤湿度和卫星高度角。

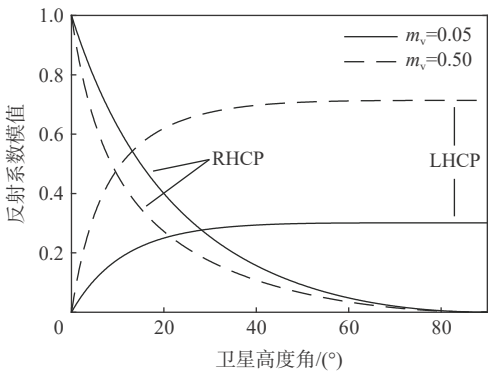


图 2　RHCP 与 LHCP 分量反射系数模值

Fig. 2　Modulus of RHCP and LHCP component reflection coefficient.

联立式 (3)～式 (8) 可得到总反射系数Γ为

$$\Gamma = \sqrt{\frac{P_r}{P_d}} \quad (9)$$

以上假设土壤表面十分光滑, 但实际上土壤表面通常具有一定粗糙度或者植被覆盖。对于轻微粗糙的土壤表面, 即土壤表面均方根高度σ满足式(10)所示的瑞利判据(Rayleigh Criterion)[25]:

$$\sigma < \frac{\lambda}{8\sin\theta} \quad (10)$$

此时粗糙度对反射信号功率的衰减可用式(11)所示的衰减因子 L_1 描述[26]:

$$L_1 = \exp\left[-\left(16\pi^2\sigma^2/\lambda^2\right)\sin^2\theta\right] \quad (11)$$

对于低矮茂密的植被覆盖如苜蓿[27]、小麦[28]等, 即植被高度h小于信号波长λ时, 可将植被视为一层均匀介质, 其对反射信号功率的衰减可用式(12)所示的衰减因子 L_2 描述[29]:

$$L_2 = \exp\left[-2\tau_p/(\sin\theta)\right] \quad (12)$$

式中: τ_p为植被光学厚度, 研究表明τ_p与植被的总含水量W_C, kg/m^2, 呈线性关系:

$$\tau_p = b_p W_C \quad (13)$$

式中: b_p为与植被类型有关的参数, 在 L 波段上, b_p的典型取值为 0.12±0.03, 对于大多数农作物均适用[29]。

因此, 在土壤粗糙度与植被的共同影响下, 反射信号功率P_r可被修正为

$$P_r^{(*)} = P_r L_1 L_2 \quad (14)$$

同时总反射系数Γ可被修正为

$$\Gamma^{(*)} = \Gamma\sqrt{L_1 L_2} \quad (15)$$

以上是通用 GNSS 干涉信号功率模型的构成。假设天线高度 H=2 m, 则对于 GPS MEO 卫星(L1 频段), 其卫星高度角从 5°变化到 30°的过程中, 由路径差变化导致的相位差变化约为 8.678×2π rad, 因此, 干涉信号功率波形将有 8.678 个震荡周期, MEO 卫星的干涉特征十分明显; 而对于北斗 GEO 卫星, 其 24 h 内高度角变化不超过 4°, 由此带来的相位差变化小于2π rad, 不满一个震荡周期, 因此, GEO 卫星的干涉特征不甚明显, 若将时间尺度扩大到数天, 则 GEO 卫星干涉图样也将呈现周期性变化, 但其周期性变化是由 GEO 卫星高度角的周期性变化引起的。

此外对于 MEO 卫星, 式 (1) 中的$P_d + P_r$一般被称为趋势项, 该项随着卫星的上升而逐渐增大或者随着卫星的下降而逐渐减小, 文献 [8] 提出使用低阶多项式拟合以剔除趋势项的影响; 对于 GEO 卫星, $P_d + P_r$的变化很小, 可视为常量, 在本文中将其称为偏置项。

2　GEO 土壤湿度干涉测量方法

2.1　土壤反射系数反演方法

根据式 (9), 反射系数是一个与发射信号功率C无关的量, 因此, 反演反射系数需设法消除发射信号功率的影响。从式 (3) 与式 (4) 可以看出, 直射信号功率与反射信号功率中均含有发射信号功率C, 常规 GNSS-R 技术使用 2 个或以上的天线分别接收直射信号与反射信号, 通过物理上的隔离实现对发射信号功率的独立测量, 从而在后续的反演过程中消除其影响。但在干涉测量技术中仅使用一支天线同时接收直射与反射信号, 两者在物理上是混叠的, 无法直接消除发射信号功率的影响, 然而考虑到 GEO 卫星的特殊性, 即 GEO 卫星高度角变化很小, 并且变化周期为 24 h, 因此, 可以认为相邻两天干涉信号功率的变化仅由土壤湿度变化引起, 而其他量如发射信号功率、卫星高度角与天线方向图的

影响均不变，于是可利用相邻两天的观测数据消除发射信号功率的影响。

对一天的GEO卫星干涉信号功率进行平均，得到偏置项并将其剔除，从而得到干涉项为

$$S^{\mathrm{d}} \approx 2\sqrt{P_{\mathrm{d}}P_{\mathrm{r}}^{(*)}}\cos\left(\Delta\varphi_{\mathrm{p}}+\varphi_0\right) \tag{16}$$

然后对式(16)进行平方积分，得到干涉幅度为

$$A=\sqrt{\int_T (\mathrm{d}S)^2\mathrm{d}t} \approx 2\sqrt{P_{\mathrm{d}}P_{\mathrm{r}}^{(*)}} \tag{17}$$

式中：积分时间T为24 h。设每天的干涉幅度测量值构成一个时间序列：A^{t_1}, A^{t_2}, …, A^{t_N}, N为总天数，则相邻两天干涉幅度的比值即为总反射系数的比值：

$$\frac{A^{t_i}}{A^{t_{i+1}}} \approx \frac{\Gamma^{t_i}}{\Gamma^{t_{i+1}}} \tag{18}$$

在此过程中，相邻两天的发射信号功率相互相抵消，同时由于土壤粗糙度与植被含水量相对于土壤湿度变化缓慢，L_1与L_2对于反演影响也在一定程度上被抵消。N个测量值可形成$N-1$个式(18)所示的方程，其矩阵形式为

$$\boldsymbol{D}\begin{bmatrix}\Gamma^{t_1}\\ \Gamma^{t_2}\\ \Gamma^{t_3}\\ \vdots\\ \Gamma^{t_{N-1}}\\ \Gamma^{t_N}\end{bmatrix}=\begin{bmatrix}0\\0\\0\\ \vdots\\0\\0\end{bmatrix} \tag{19}$$

式中：矩阵$\boldsymbol{D}$的表达式为

$$\boldsymbol{D}=\begin{bmatrix}1 & -\frac{A^{t_1}}{A^{t_2}} & 0 & 0 & \cdots & \cdots & \cdots & 0\\ 0 & 1 & -\frac{A^{t_2}}{A^{t_3}} & 0 & \cdots & \cdots & \cdots & 0\\ \vdots & 0 & 1 & -\frac{A^{t_3}}{A^{t_4}} & 0 & \cdots & \cdots & 0\\ \vdots & \vdots & 0 & 1 & \ddots & \ddots & & 0\\ \vdots & \vdots & \vdots & 0 & \ddots & \ddots & \ddots & \vdots\\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & 0\\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & -\frac{A^{t_{N-1}}}{A^{t_N}}\end{bmatrix} \tag{20}$$

由于矩阵$\boldsymbol{D}$为秩亏矩阵，理论上式(20)所示的方程有无数个解，求解该方程需要增加约束条件，从而构成一个有约束的线性方程，最终求解得到该方程在最小二乘意义下的解[30]。本文中该约束条件为待求解反射系数的变化范围，该范围可根据实际土壤湿度变化情况以及前述模型合理估算。

2.2 土壤湿度反演方法

在得到总反射系数后，可根据式(4)～式(8)进一步反演得到土壤湿度，但需要在反演之前对接收机天线增益影响进行修正，如前所述，不同的天线设计导致天线增益不同，实际上除了天线增益外，还存在通道处理增益，在无法准确获取增益信息的情况下，可用待定参数替换式(9)中的增益项，形成半经验表达式：

$$\Gamma=\Big[|\Gamma_{\mathrm{rR}}|^2 10^{a_1/10}+|\Gamma_{\mathrm{rL}}|^2 10^{a_2/10}+ 2\sqrt{|\Gamma_{\mathrm{rR}}\Gamma_{\mathrm{rL}}|^2 10^{(a_1+a_2)/10}}\times\cos(\varphi_{\mathrm{rR}}-\varphi_{\mathrm{rL}})\Big]^{\frac{1}{2}} \tag{21}$$

式中：a_1和a_2为待定增益参数。可利用若干天的实测土壤湿度与反射系数估计值作为训练数据得到a_1和a_2的估计，再将其代入式(21)中反演土壤湿度。值得注意的是由于功率增益应恒大于0，式(21)中使用待定参数的幂指数形式表示增益，此时a_1和a_2表示对数刻度下的增益，从而将约束优化问题转换为无约束优化问题，便于求解。

总体反演流程如图3所示。

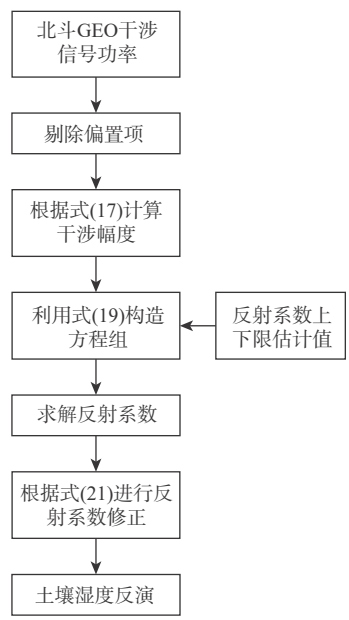


图3 总体反演流程

Fig. 3 Overall retrieval workflow

在利用式(21)进行反演时，需利用GEO卫星的高度角信息，由于其变化较小，可取一天内GEO卫星高度角的平均值参与反演运算。

3 仿真验证

3.1 GEO卫星载噪比变化仿真

假设土壤质地为典型黏土土质($S_{\mathrm{s}}=20\%$, $C_{\mathrm{s}}=40\%$)；土壤表面均方根高度$\sigma=2$ cm；植被含水量$W_{\mathrm{C}}=0.45$ kg/m^2；GEO卫星高度角$\theta\approx26.5°$；天线参数设置为$G_{\mathrm{R}}(\theta)=-1.2$ dB, $G_{\mathrm{R}}(-\theta)=-14.0$ dB, $G_{\mathrm{L}}(-\theta)=-18.3$ dB，不同土壤湿度下，24 h内北斗GEO卫星载噪比情况如图4所示。可知在交叉极化的影响下，土壤湿度增大时，GEO卫星干涉信号

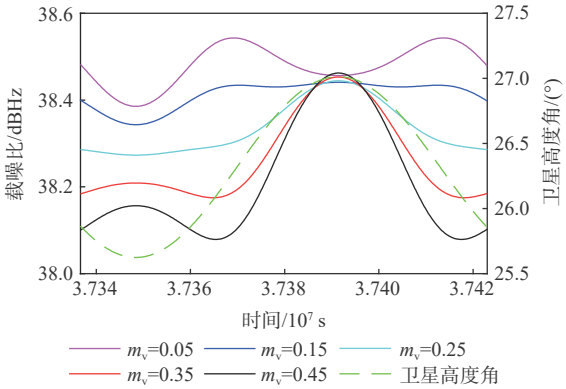


图 4　交叉极化影响下的 GEO 卫星载噪比仿真

Fig. 4　Simulation of GEO satellite C/N_0 under effect of cross-polarization

功率的震荡幅度总体上呈现出先减小后增大的趋势, 原因为当土壤湿度较小时, RHCP 分量起主导作用(根据图 2), 因此, 总反射系数随土壤湿度增加而减小, 从而导致干涉信号功率的震荡幅度减小; 当土壤湿度较大时, LHCP 分量起主导作用, 此时土壤湿度增大导致 LHCP 分量增大, 最终导致干涉信号功率震荡幅度增大。同时,当土壤湿度较高时干涉信号功率波形呈现出与土壤湿度较低时不同的形状, 这也是由于土壤湿度与卫星高度角变化时, RHCP 和 LHCP分量的变化程度不同引起的。

图 5 为将天线的交叉极化增益$G_{\mathrm{L}}(-\theta)$减小至 −48.3 dB 时(即无交叉极化影响)的载噪比变化情况。此时天线几乎完全抑制反射信号的交叉极化分量, 是一种接近理想的情况。可知干涉信号功率的震荡幅度总体上随土壤湿度的增大而减小, 同时波形形状基本相似。

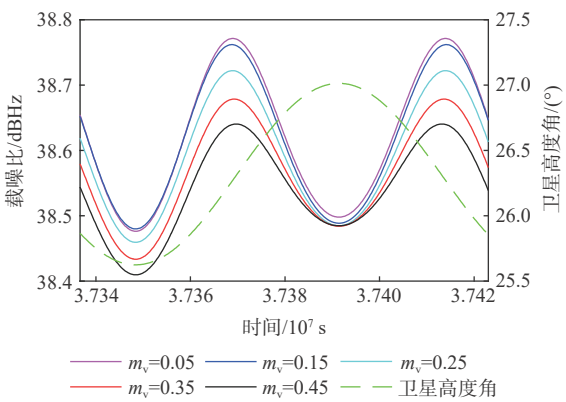


图 5　无交叉极化影响下的 GEO 卫星载噪比仿真

Fig. 5　Simulation of GEO satellite C/N_0 without effect of cross-polarization

3.2　反射系数反演仿真

图 6 为当土壤湿度在 0.02 $\mathrm{cm^3/cm^3}$ 与 0.5 $\mathrm{cm^3/cm^3}$ 之间非线性变化, 植被含水量在 0.4 $\mathrm{kg/m^2}$ 与 0.5 $\mathrm{kg/m^2}$ 之间缓慢变化时, 仿真反射系数与反演反射系数的结果对比。可知在交叉极化的影响下反射系数的反演结果与仿真设定值相比存在明显的畸变, 导致这一现象的主要原因是交叉极化影响下干涉信号功率波形随土壤湿度变化发生如图 4 所示的畸变。无交叉极化影响时, 总反射系数反演也存在畸变, 此时残余的粗糙度与植被影响是导致畸变的主要原因, 但程度相对较小。根据式 (21), 土壤湿度通过影响 RHCP 分量与 LHCP 分量反射系数以影响总反射系数, 而式 (21) 中的 2 个待定参数则分别用于修正 LHCP 分量与 RHCP 分量反射系数, 通过调整这 2 个待定参数可补偿反射系数反演误差, 因此, 图 6 所示的畸变对土壤湿度反演的影响较小。

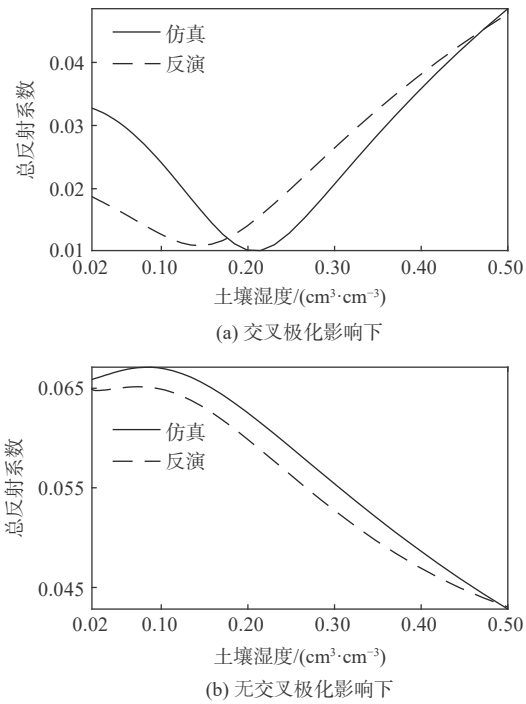


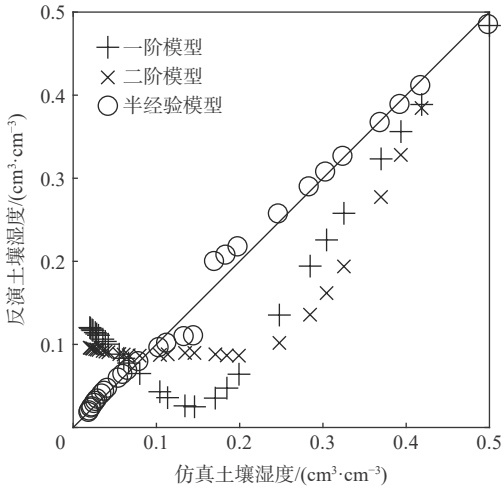
图 6　反射系数反演结果

Fig. 6　Retrieval results of the reflection coefficient

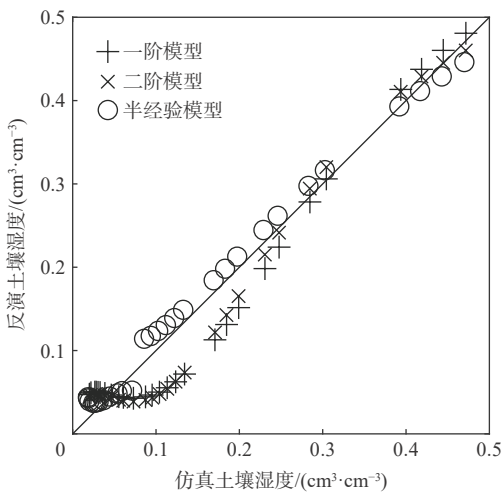
3.3　土壤湿度反演仿真

将第 3.2 节仿真产生的数据集随机平均划分为 2 组, 一组用于训练式 (21) 所示的半经验模型, 另一组用于反演土壤湿度, 图 7 为有交叉极化影响与无交叉极化影响下反演的统计结果。

图 7 中除给出本文所述的半经验模型的反演结果外, 还给出了现有研究所使用的纯经验反演模型如一阶和二阶多项模型的反演结果。通过对比可知在交叉极化的影响下, 半经验反演模型的反演误差明显小于纯经验模型的反演误差, 这是由于交叉极化影响下的反射系数与土壤湿度之间的关系呈现出图 6(a) 所示的高度非线性, 此时一阶、二阶多项式的建模能力有限, 因此, 反演存在较大误差。对于无交叉极化影响情况, 纯经验反演模型与半经验反演模型的性能相当。具体反演性能总结如表 1 所示。



(a) 交叉极化影响下



(b) 无交叉极化影响下

图 7 土壤湿度反演结果

Fig. 7 Retrieval results of soil moisture

表 1 土壤湿度反演仿真结果统计

Table 1 Statistical results of simulation of soil moisture retrieval

极化情况	反演模型	均方根误差/(cm³·cm⁻³)	相关系数
有交叉极化	半经验模型	0.012 3	0.996 2
	一阶模型	0.083 2	0.809 0
	二阶模型	0.079 3	0.829 7
无交叉极化	半经验模型	0.014 5	0.995 9
	一阶模型	0.034 6	0.974 1
	二阶模型	0.032 4	0.977 5

综合上述仿真结果可知半经验反演方法的适应性更好。

4 实验验证

4.1 实验数据采集

实验在北京市通州区的一处农田进行，实验地点卫星地图如图 8 所示。实验周期为 2019 年 11 月 10 日至 2020 年 01 月 01 日，跨度为 53 天，实验期间地表无植被覆盖，土壤表面均方根高度小于 2 cm。使用高精度 GNSS 接收机采集 GNSS 原始观测数据，数据记录频率为 1 Hz，以 Rinex 格式存储。由于 5 颗北斗 GEO 卫星均分布在南面，因此，实验过程中将 GNSS 天线的最大增益方向调整为水平朝南而非指向天顶（见图 8 左下角照片）以减少天线对反射信号的抑制，天线高度为 1.88 m，天线架设位置如图 8 中红点所示。使用探针式土壤湿度传感器测量土壤湿度，传感器填埋深度约为 5 cm，采样时间间隔为 1 min。

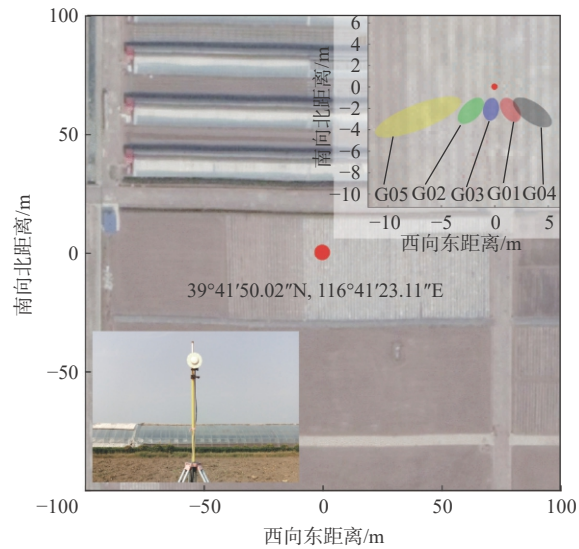


图 8 实验地点卫星地图

Fig. 8 Satellite map of experiment site

实验期间观察到 5 颗北斗 GEO 卫星的高度角变化范围如表 2 所示。

表 2 实验期间 GEO 卫星高度角变化情况

Table 2 Variation of GEO satellite elevation angle during experiment (°)

卫星号	最低高度角	最高高度角
1	36.738 9	39.510 4
2	31.821 3	34.329 8
3	42.434 4	44.878 6
4	25.624 7	27.013 6
5	14.731 7	16.990 1

结合 5 颗 GEO 卫星的方位角可绘制其第一菲涅尔反射区的位置和大小，如图 8 右上角所示。

实验期间土壤湿度变化情况如图 9 所示。

4.2 数据处理流程

步骤 1 利用 Rinex 广播星历数据与接收机天线位置，计算 GEO 卫星的高度角与方位角。

步骤 2 从每天的 Rinex 观测值文件中提取 GEO 卫星载噪比数据，并按照式 (22) 将对数刻度转换为功率线性刻度。

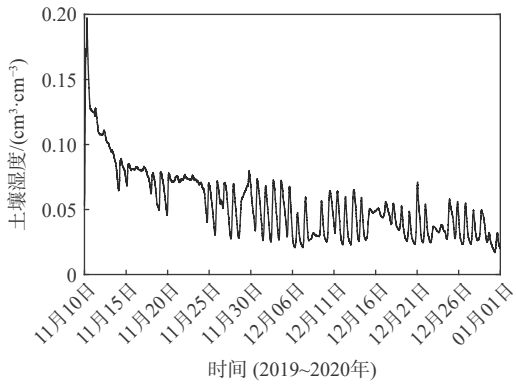


图 9　实验期间土壤湿度变化

Fig. 9　Soil moisture variation during experiment

$$\mathrm{dBHz} \xrightarrow{10^{(\cdot)/10}} \mathrm{W\cdot W^{-1}\cdot Hz} \tag{22}$$

步骤 3　由于实际的载噪比数据存在较大噪声或者毛刺, 需要先对步骤 2 中得到的载噪比数据进行中值滤波, 然后按照图 3 所示的流程图进行反演。

4.3　实验结果与分析

4.3.1　实验期间 GEO 卫星载噪比变化情况

根据表 2, 选择 5 颗 GEO 卫星中高度角小于 30°的 GEO 卫星数据进行处理, 即 4 号星与 5 号星数据, 图 10 为 4 号星和 5 号星 3 天的载噪比变化。可知 GEO 卫星的载噪比随着高度角的周期性变化而周期性变化, 变化周期为 24 h(虚线为相邻两天的分界线)。其中 5 号星数据质量较差, 其载噪比数据存在明显的毛刺现象, 并且每天数据均存在约

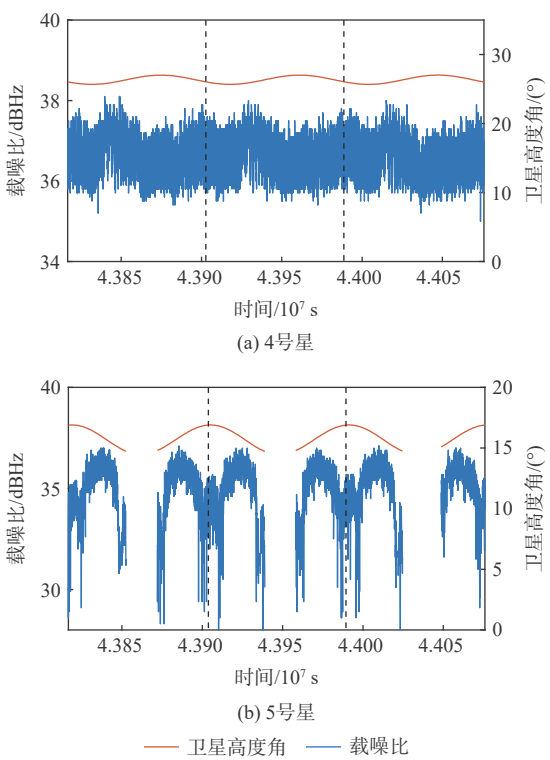


图 10　4 号星与 5 号星载噪比变化

Fig. 10　C/N_0 variation of satellite No.4 and No.5.

7 h 的缺失, 且缺失均发生在高度角最低(小于 15°)的时段内, 造成这一现象的原因可能是远处有高大遮挡物如树木的遮挡。鉴于 5 号星数据质量较差, 后续处理仅针对 4 号星。

4.3.2　反射系数反演

图 11 为 4 号星的反射系数反演结果。显示了实测土壤湿度值为 24 h 土壤湿度平均值, 实验周期为 53 天, 共计 53 组数据。可以看出反演得到的反射系数与实测土壤湿度总体上呈负相关, 这说明反射系数中右旋分量起主导作用。在利用式 (19) 求解反射系数的过程中, 需要给定反射系数上下界估计值, 本文根据实际土壤湿度变化范围, 利用式 (5) 估算出反射系数约为 0.3, 上下界范围应包含 0.3, 同时考虑噪声的影响, 上下界范围应尽可能大从而给寻优算法以较大的优化空间, 因此本文设置上下界范围为 0.1～ 0.4, 最终求解得到未经修正的反射系数的变化范围为 0.117 8～0.367 2。

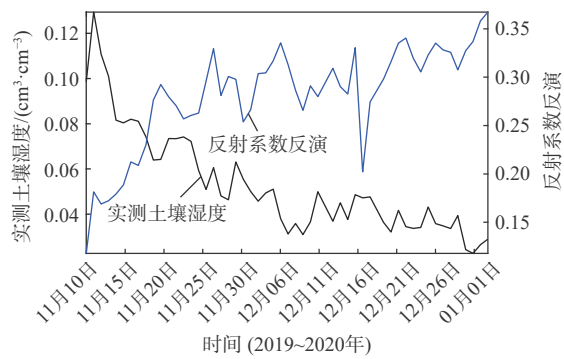


图 11　4 号星的反射系数反演结果

Fig. 11　Retrieval results of reflection coefficient for satellite No.4

4.3.3　土壤湿度反演

将上述数据随机平均划分为训练集与测试集, 训练集含 26 组数据用于训练半经验模型, 测试集含 27 组数据用于反演验证。训练结果如图 12 所示。

土壤湿度反演结果如图 13 所示。图中一阶模

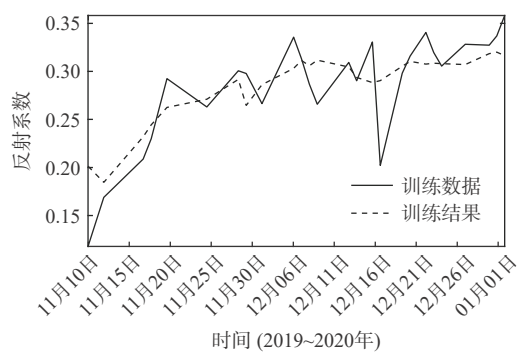


图 12　半经验模型训练结果

Fig. 12　Training results of semi-empirical model

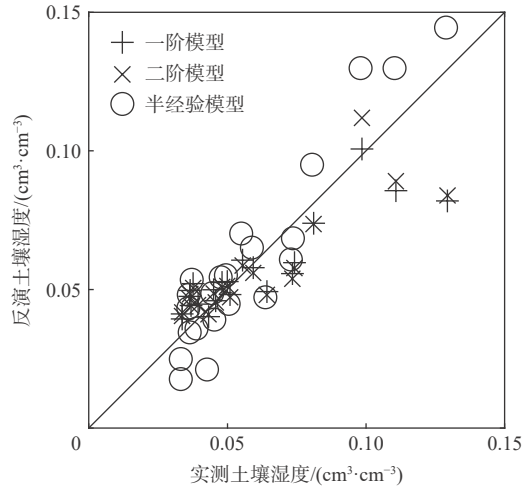


图 13 土壤湿度反演结果

Fig. 13 Soil moisture retrieval results

型、二阶模型与半经验模型反演的均方根误差分别为 0.014 2 cm^3/cm^3、0.014 1 cm^3/cm^3、0.013 3 cm^3/cm^3。

以上仅为一次随机分组、训练、反演的结果，具有偶然性，本文进一步改变用于训练半经验反演模型的数据数量，每改变一次数据量则进行 200 次随机分组、训练、反演，并计算反演的平均相关系数与平均均方根误差，结果如图 14 所示。

图 14 中训练数据量从 3 组变化到 26 组（总数据量的一半）。可知 3 个模型的反演性能均随着训练数据量的增加而收敛，收敛后一阶模型、二阶模型、半经验模型的平均均方根误差分别为：0.014 1 cm^3/cm^3、0.015 0 cm^3/cm^3、0.019 9 cm^3/cm^3。其中二阶模型的收敛速度最慢，其在训练数据量不充足的情况下存在相当大的反演误差。而一阶模型与半经验模型收敛速度相近，主要由于本实验中土壤湿度变化范围较小，此时反射系数与土壤湿度之间的关系接近线性。从图 14（b）中可知半经验模型反演结果的相关系数在训练数据量不充分的情况下依然保持较高相关性，主要源于半经验模型背后的物理规律的支撑。从反演的均方根误差来看，半经验模型稍高于纯经验的一阶模型，主要原因是在右旋分量占主导地位的情况下，反射系数对于误差与噪声的敏感性较高。图 15 为不同的反射系数估计误差下土壤湿度的反演误差。图 15 中假设土壤湿度真值为 0.2 cm^3/cm^3。图中上下两条横线表示反演误差过大以至最终的土壤湿度反演结果超过正常范围（0～1 cm^3/cm^3）而截止。从图中可以看出 RHCP 分量占主导地位的情况下，反演的误差敏感性整体较高，并且高度角越高误差敏感性越高，在 30°的高度角情况下，0.09 的反射系数估计误差或者噪声将导致 0.4 cm^3/cm^3 的土壤湿度反演误差。

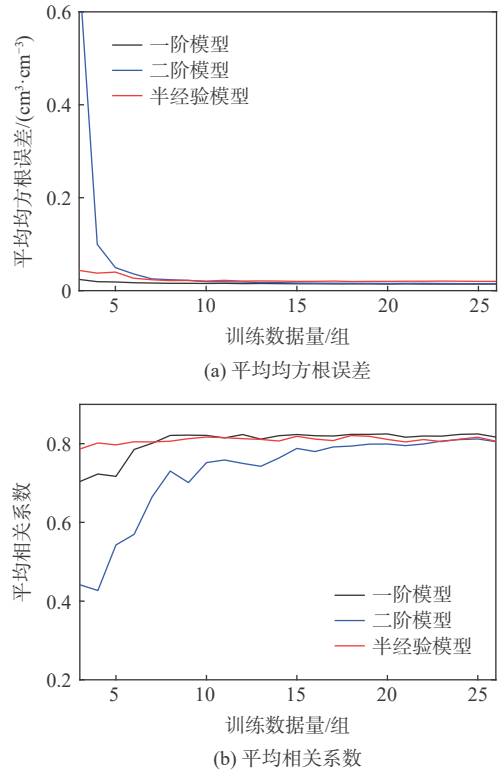


图 14 不同训练数据量下的反演性能

Fig. 14 Retrieval performance under different number of training data

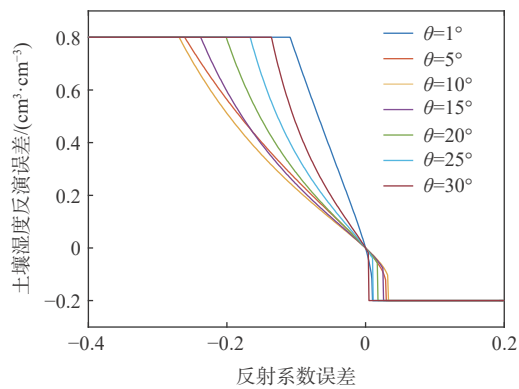


图 15 土壤湿度反演误差仿真

Fig. 15 Simulation of soil moisture retrieval error

5 结 论

1）从北斗 GEO 卫星干涉信号功率中反演土壤反射系数的关键是消除干涉幅度中发射信号功率的影响，GEO 卫星星地间稳定的几何构型为消除发射信号功率影响提供了有利条件，通过取相邻两天干涉幅度比值的方法，可抵消发射信号功率影响。同时，由于土壤表面的轻微粗糙度与低矮密集植被覆盖对反射信号功率的影响可简单地使用乘性衰减因子描述，因此，该方法在理论上也能够在一定程度上抵消二者的影响。

2）土壤反射系数在交叉极化的影响下与土壤

湿度之间的关系呈现出高度非线性, 此时纯经验的一阶与二阶多项式反演模型难以准确反演土壤湿度, 而本文模型是基于反射系数理论模型, 本文模型考虑了天线交叉极化增益的影响, 因而可以适应不同类型与不同厂商的天线, 理论上反演的均方根误差小于 0.02 cm^3/cm^3。

3) 实验研究表明, 在土壤湿度变化范围较小(小于 0.2 cm^3/cm^3)的情况, 本文模型与纯经验的一阶多项式反演模型的反演性能相当, 反演的均方根误差为 0.019 9 cm^3/cm^3, 相关系数达 0.8, 与仿真结果基本相符。

4) 本文模型对误差和噪声的敏感性较高, 尤当反射信号中 RHCP 分量占主导地位时, 这是由 RHCP 信号的电磁特性决定的。

本文模型与方法理论上适用于轻微粗糙与低矮密集植被覆盖的土壤表面, 未来仍需开展大量的实验验证以研究不同土壤条件与环境下所述模型与方法的有效性。

致谢 感谢北京市农林科学院蔬菜研究中心宋曙辉研究员为本文的实验研究提供实验场地。

参考文献 (References)

[1] 裴悦琨, 韩心新. GNSS-R探测土壤湿度综述[J]. 大地测量与地球动力学, 2021, 41(2): 140-144.
PEI Y K, HAN X X. Overview of GNSS-R technique for soil moisture detection[J]. Journal of Geodesy and Geodynamics, 2021, 41(2): 140-144(in Chinese).

[2] LI F, PENG X F, CHEN X W, et al. Analysis of key issues on GNSS-R soil moisture retrieval based on different antenna patterns[J]. Sensors, 2018, 18(8): 2498.

[3] EDOKOSSI K, CALABIA A, JIN S G, et al. GNSS-reflectometry and remote sensing of soil moisture: A review of measurement techniques, methods, and applications[J]. Remote Sensing, 2020, 12(4): 614.

[4] MARTÍN-NEIRA M. A pasive reflectometry and interferometry system (PARIS) application to ocean altimetry[J]. ESA Journal, 1993, 17(4): 331-355.

[5] ZHENG N Q, CHEN P, LI Z. Accuracy analysis of ground-based GNSS-R sea level monitoring based on multi GNSS and multi SNR[J]. Advances in Space Research, 2021, 68(4): 1789-1801.

[6] GAO F, XU T H, MENG X Y, et al. A coastal experiment for GNSS-R code-level altimetry using BDS-3 new civil signals[J]. Remote Sensing, 2021, 13(7): 1378.

[7] LIU B J, WAN W, HONG Y. Can the accuracy of sea surface salinity measurement be improved by incorporating spaceborne GNSS-reflectometry?[J]. IEEE Geoscience and Remote Sensing Letters, 2021, 18(1): 3-7.

[8] LARSON K M, SMALL E E, GUTMANN E, et al. Using GPS multipath to measure soil moisture fluctuations: Initial results[J]. GPS Solutions, 2008, 12(3): 173-177.

[9] LARSON K M, SMALL E E, GUTMANN E D, et al. Use of GPS receivers as a soil moisture network for water cycle studies[J]. Geophysical Research Letters, 2008, 35(24): L24405.

[10] JIA Y, JIN S G, SAVI P, et al. Modeling and theoretical analysis of GNSS-R soil moisture retrieval based on the random forest and support vector machine learning approach[J]. Remote Sensing, 2020, 12(22): 3679.

[11] GEREMIA-NIEVINSKI F, HOBIGER T, HAAS R, et al. SNR-based GNSS reflectometry for coastal sea-level altimetry: Results from the first IAG inter-comparison campaign[J]. Journal of Geodesy, 2020, 94(8): 1-15.

[12] MARTÍN A, LUJÁN R, ANQUELA A B. Python software tools for GNSS interferometric reflectometry (GNSS-IR)[J]. GPS Solutions, 2020, 24(4): 94.

[13] CHEW C C, SMALL E E, LARSON K M, et al. Effects of near-surface soil moisture on GPS SNR data: Development of a retrieval algorithm for soil moisture[J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 52(1): 537-543.

[14] 张楠, 严颂华, 王文伟. 北斗GEO卫星微动条件下的GNSS-R土壤湿度反演[J]. 科学技术与工程, 2019, 19(9): 154-161.
ZHANG N, YAN S H, WANG W W. Soil moisture inversion based on GNSS-R under the micro-motion condition of Beidou GEO satellites[J]. Science Technology and Engineering, 2019, 19(9): 154-161(in Chinese).

[15] 邹文博, 张波, 洪学宝, 等. 利用北斗GEO卫星反射信号反演土壤湿度[J]. 测绘学报, 2016, 45(2): 199-204.
ZOU W B, ZHANG B, HONG X B, et al. Soil moisture retrieval using reflected signals of BeiDou GEO satellites[J]. Acta Geodaetica et Cartographica Sinica, 2016, 45(2): 199-204(in Chinese).

[16] 杨磊, 吴秋兰, 张波, 等. SVRM辅助的北斗GEO卫星反射信号土壤湿度反演方法[J]. 北京航空航天大学学报, 2016, 42(6): 1134-1141.
YANG L, WU Q L, ZHANG B, et al. SVRM-assisted soil moisture retrieval method using reflected signal from BeiDou GEO satellites[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(6): 1134-1141(in Chinese).

[17] 班伟. 利用GNSS反射信号反演水位、积雪厚度和土壤湿度的方法研究[D]. 武汉: 武汉大学, 2017.
BAN W. Research on retrieving water level, snow depth and soil moisture using GNSS-reflected signal[D]. Wuhan: Wuhan University, 2017 (in Chinese).

[18] BAN W, YU K G, ZHANG X H. GEO-satellite-based reflectometry for soil moisture estimation: Signal modeling and algorithm development[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(3): 1829-1838.

[19] 李培昊. 基于北斗GEO卫星信噪比的土壤湿度反演[D]. 成都: 西南交通大学, 2019.
LI P H. Soil moisture inversion based on Beidou GEO satellite signal-to-noise ratio[D]. Chengdu: Southwest Jiaotong University, 2019 (in Chinese).

[20] ZAVOROTNY V U, LARSON K M, BRAUN J J, et al. A physical model for GPS multipath caused by land reflections: Toward bare soil moisture retrievals[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2010, 3(1): 100-110.

[21] BALANIS C A. Antenna theory analysis & design[M]. New York: John Wiley, 1996

[22] 万玮, 李黄, 洪阳. 作为外辐射源雷达的GNSS-R遥感多极化问题[J]. 雷达学报, 2014, 3(6): 641-651.
WAN W, LI H, HONG Y. Issues on multi-polarization of GNSS-R for passive radar detection[J]. Journal of Radars, 2014, 3(6): 641-651(in Chinese).

[23] 刘军, 赵少杰, 蒋玲梅, 等. 微波波段土壤的介电常数模型研究进展[J]. 遥感信息, 2015, 30(1): 5-13.
LIU J, ZHAO S J, JIANG L M, et al. Research progress on dielectric constant model of soil at microwave frequency[J]. Remote Sensing Information, 2015, 30(1): 5-13(in Chinese).

[24] HALLIKAINEN M T, ULABY F T, DOBSON M C, et al. Microwave dielectric behavior of wet soil-part 1: Empirical models and experimental observations[J]. IEEE Transactions on Geoscience and Remote Sensing, 1985, 23(1): 25-34.

[25] REES W G. Physical principles of remote sensing[M]. Cambridge: Cambridge University Press, 2013.

[26] DAVIES H. The reflection of electromagnetic waves from a rough surface[J]. Proceedings of the IEE - Part IV:Institution Monographs, 1954, 101(7): 209-214.

[27] WAN W, LARSON K M, SMALL E E, et al. Using geodetic GPS receivers to measure vegetation water content[J]. GPS Solutions, 2015, 19(2): 237-248.

[28] ZHANG S B, ROUSSEL N, BONIFACE K, et al. Use of reflected GNSS SNR data to retrieve either soil moisture or vegetation height from a wheat crop[J]. Hydrology and Earth System Sciences, 2017, 21(9): 4767-4784.

[29] WIGNERON J P, CALVET J C, PELLARIN T, et al. Retrieving near-surface soil moisture from microwave radiometric observations: Current status and future plans[J]. Remote Sensing of Environment, 2003, 85(4): 489-506.

[30] AL-KHALDI M M, JOHNSON J T, O'BRIEN A J, et al. Time-series retrieval of soil moisture using CYGNSS[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(7): 4322-4331.

Soil moisture retrieval using Beidou GEO satellite interference signal power

HAN Mutian, XU Zhichao*, CHANG Qing, ZHANG Bo, WANG Feng, HONG Xuebao

(School of Electronics and Information Engineering, Beihang University, Beijing 100191, China)

Abstract: Soil moisture retrieval using Beidou geostationary Earth orbit (GEO) interference signal power was studied. Current researches mainly establish empirical models for retrieving soil moisture. Therefore, a semi-empirical retrieval method was proposed. Taking the advantage of Beidou GEO satellite's stable geometric configuration relative to the earth, this method utilized the interference amplitude metric obtained in two consecutive days to cancel the influence of the power of the transmitted signal, therefore, achieving the retrieval of the reflection coefficient. Then, based on the reflection coefficient theoretical model under the influence of cross-polarization , a semi-empirical soil moisture retrieval model was constructed. Finally, the proposed method and model are validated through simulations and experiments. The simulation results showed that the proposed method and model can better adapt to the nonlinear situation encountered in the retrieval process. Finally, the experiment results show that the root mean square error of soil moisture retrieval is less than 0.02 $cm^3 \cdot cm^{-3}$, and the correlation coefficient exceeds 0.8.

Keywords: Beidou navigation satellite system; GEO satellite; interference signal power; reflection coefficient; soil moisture

Received: 2021-08-20; **Accepted**: 2021-11-20; **Published Online**: 2021-11-30 13:16

URL: kns.cnki.net/kcms/detail/11.2625.V.20211129.2012.010.html

Foundation items: National Natural Science Foundation of China (41774028); Post Doctoral Innovative Talent Support Program (BX20200039)

* **Corresponding author**. E-mail: by1602112@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0489

下表面射流的超临界翼型气动性能分析

王若尘[1,2]，张国鑫[1,*]，王翔宇[1]，马晓平[1]

(1. 中国科学院工程热物理研究所，北京 100190；2. 中国科学院大学，北京 100049)

摘 要：为探究下表面射流关键参数对超临界翼型气动性能的影响，采用雷诺平均Navier-Stokes（RANS）方程与Spalart-Allmaras（S-A）湍流模型进行数值模拟。通过比较基准RAE2822翼型与下表面射流翼型的流场，验证下表面射流能够在翼型后缘诱导产生逆时针分离涡，带动流线向下偏折，增加了翼型的等效弯度，同时加大前缘的吸力峰，从而提高翼型的气动性能。进一步探究射流位置、射流动量系数、射流角度、马赫数等关键参数对RAE2822翼型气动性能的影响规律。结果表明：给定状态下，下表面射流的位置越靠后，动量系数越大，翼型的气动性能越优。下表面射流在α=0°和2°时的最优射流角度为110°，在α=4°时的最优射流角度为160°，且在最优射流角度下能有效提高翼型马赫数在0.3～0.6范围内的气动性能。

关 键 词：主动流动控制；下表面射流；超临界翼型；气动性能；雷诺平均Navier-Stokes

中图分类号：V211.3

文献标志码：A **文章编号**：1001-5965(2023)07-1671-09

未来飞行器发展所面临的重大挑战之一是如何利用先进技术提高飞行器的气动效率，从而显著降低飞行油耗，增加续航时间。新型的主动流动控制方法是实现上述目标的重要途径[1-3]。此外，在飞行控制系统方面，传统的飞行控制多依赖于机械式的飞行操纵面，如升降舵、副翼、襟翼等来提高飞行器的机动性与可控性，但同时带来了维护成本高，翼面偏转过程破坏隐身性能、产生气动噪声等问题[4]。因此，采用新型的流动控制技术实现飞行器的无舵面飞行，具有十分重要的意义。

主动流动控制技术因其在增升、减阻、隐身、降噪及姿态控制等方面的优势，成为未来航空器发展的一大热点[5-6]。其中，环量控制(circulation control, CC)是主动流动控制的主要方式之一。Xu等[7]提出一种基于CC的钝后缘风力机翼型的主动流动控制方法，指出CC可以显著提高钝后缘翼型的升力，减小阻力，从而提高钝后缘翼型的气动性能。Shi等[8]开发了一套CC作动器来替代传统的舵面，并将其应用到小型无人飞行器的飞行试验中。该作动器被证明可以完全替代传统的副翼和升降舵，实现对飞机飞行姿态的控制。雷玉昌等[9-10]对超临界CC翼型进行非定常气动力建模[9]，并进一步研究了脉冲射流对CC翼型气动性能的影响[10]。除了CC之外，Zha等[11]提出的协同射流(co-flow jet, CFJ)作为一种新型的主动流动控制技术受到越来越广泛的关注。Lefebvre等[12]以CFJ6415翼型为研究对象，开展不同马赫数下翼型气动性能的分析，结果表明：CFJ翼型不仅能获得很高的最大升力系数，而且能够有效提高低迎角下的巡航性能。Xu和Zha[13]采用改进的延迟型分离涡模拟(improved delayed detached eddy simulation, IDDES)方法对基于CFJ的飞机三维控制面进行数值研究，数值模拟结果表明：采用CFJ控制的三维控制面可以在低能量消耗的情况下大幅提高飞机控制权限。刘沛清等[14]最早

收稿日期：2021-08-25；**录用日期**：2021-09-10；**网络出版时间**：2021-09-24 12:09
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20210923.1811.001.html
基金项目：国家自然科学基金(61901448)；中国科学院青年创新促进会(2020149)
*通信作者. E-mail：zhangguoxin@iet.cn

*引用格式：王若尘，张国鑫，王翔宇，等. 下表面射流的超临界翼型气动性能分析[J]. 北京航空航天大学学报，2023，49(7)：1671-1679.
WANG R C，ZHANG G X，WANG X Y，et al. Aerodynamic performance analysis of supercritical airfoil with lower surface jet[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1671-1679 (in Chinese).*

阐述了 CFJ 的增升效果和机理。许建华等[15]通过引入螺旋桨激励盘模型分析了射流动量系数、开口尺寸和位置等关键参数对 CFJ6415 翼型气动性能的影响规律。上述 2 种主动流动控制方法[7-15]虽在一定程度上提高了翼型的气动性能，但均在原有的基准翼型上进行了较大的改动，破坏了翼型本身的流线型结构。当飞行过程中射流处于关闭状态时，CC 的圆弧形后缘和 CFJ 的上表面凹槽均在一定程度上增加了飞行阻力。

翼型表面吹吸气技术在不改变翼型形状的同时，也能带来很好的流动控制效果，且该技术的发展已相对成熟[16]。Huang 等[17]采用翼型上表面定常吹吸气的方式对 NACA0012 翼型的气动性能进行系统的分析；Genc 和 Kaynak[18]研究了前缘吹吸气对 NACA2415 翼型分离流动的影响；张志勇等[16]采用大涡模拟方法，在 Ma=0.2 条件下对 NACA0012 翼型上表面吹吸气射流进行数值模拟，结果表明：上表面吹气射流对翼型的气动性能影响较差，吸气射流的效果要明显优于吹气射流。付云豪等[19]提出一种新型下表面射流的主动流动控制方式，并针对 NACA0012 翼型在 Ma=0.155 的工况下开展了相关的工作，证明下表面射流对 NACA0012 翼型的气动性能有较大的提升。目前，翼型表面吹吸气技术大量集中在翼型上表面和低马赫数工况，针对较高马赫数下的下表面射流研究相对较少。此外，现有翼型射流研究大多未考虑射流反作用力对翼型气动性能的影响，评估射流技术的气动力增益时存在不足[16-19]。

基于上述问题，在不改变翼型外形的前提下，针对 RAE2822 超临界翼型，分析下表面射流提升翼型气动性能的机理，并进一步探究射流位置、射流动量系数、射流角度及马赫数对翼型气动性能的影响，旨在为下表面射流的研究与应用提供一定的参考。

1 数值计算方法与验证

1.1 计算模型

选取的基准翼型为 RAE2822 翼型，其弦长为 1 m，下表面射流参数如图 1 所示。射流口位于翼型下表面靠近尾缘的位置，射流位置用参数 x_j 表示；射流口宽度 h 设定为弦长的 1%；射流角度 θ 定义为射流方向与射流口中心沿翼面前向切线的夹角；α 表示飞行迎角。

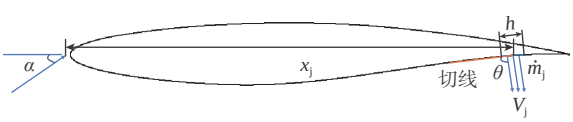


图 1 RAE2822 翼型下表面射流参数示意图

Fig. 1 Diagram of jet parameters of RAE2822 airfoil lower surface

1.2 数值计算方法

数值模拟方法采用雷诺平均 Navier-Stokes（Reynolds average Navier-Stokes, RANS）方程，其积分形式表示为

$$\frac{\partial}{\partial t}\int_{\Omega}\boldsymbol{Q}\mathrm{d}\Omega+\oint_{\Omega}(\boldsymbol{F}_{\mathrm{I}}-\boldsymbol{F}_{\mathrm{V}})\mathrm{d}S=0 \tag{1}$$

式中：$\boldsymbol{Q}$ 为流动变量；$\boldsymbol{F}_{\mathrm{I}}$ 为无黏通量；$\boldsymbol{F}_{\mathrm{V}}$ 为黏性通量；Ω 和 S 表示控制体与控制体边界。采用有限体积法对控制方程进行离散，湍流模型采用 Spalart-Allmaras（S-A）单方程模型，该模型对受到逆压力梯度作用的边界层具有良好的效果，且能够较好地模拟漩涡运动。边界条件的设定如下：远场边界为压力远场条件，翼型表面为无滑移壁面条件。此外，计算状态下射流口马赫数不超过 0.3，因此，可将射流简化为不可压缩流处理，射流口边界设为速度入口条件。

采用无量纲射流动量系数 c_μ 来表征射流的强弱，其定义为

$$c_\mu=\frac{\dot{m}_{\mathrm{j}}V_{\mathrm{j}}}{\frac{1}{2}\rho_\infty {V_\infty}^2 c} \tag{2}$$

式中：$\dot{m}_{\mathrm{j}}$ 为射流质量流量；V_{j} 为射流速度大小；V_∞ 为自由流速度；ρ_∞ 为自由流密度；c 为翼型弦长。在仿真中，设定 $V_{\mathrm{j}}/V_\infty\leqslant 0.5$。

1.3 翼型气动力计算方法

由于射流的存在，下表面射流翼型的气动力计算方法与基准翼型截然不同。对于基准翼型，翼型表面的升力和阻力可通过对表面压力分布与剪应力分布的积分求得。而对于下表面射流翼型，借助文献 [20] 的思想，翼型的气动力由 3 个部分叠加而成：压力对翼型表面及射流口的作用力、剪切力对翼型表面的作用力及射流反作用力对射流口的作用力，计算式为

$$C_L=C_{L,p}+C_{L,\tau}+C_{L,\mathrm{jet}}=C_{L,p+\tau}+C_{L,\mathrm{jet}} \tag{3}$$

$$C_D=C_{D,p}+C_{D,\tau}+C_{D,\mathrm{jet}}=C_{D,p+\tau}+C_{D,\mathrm{jet}} \tag{4}$$

式中：C_L 和 C_D 为升力系数与阻力系数，下标“p”，“τ”和“jet”分别对应压力、剪切力及射流反作用力的贡献。$C_{L,p+\tau}$ 和 $C_{D,p+\tau}$ 可通过标准的积分运算求解，而 $C_{L,\mathrm{jet}}$ 和 $C_{D,\mathrm{jet}}$ 的计算式为

$$C_{L,\mathrm{jet}}=\frac{\dot{m}\boldsymbol{V}_{\mathrm{j}}}{\frac{1}{2}\rho_\infty {V_\infty}^2 c}\cdot\boldsymbol{j} \tag{5}$$

$$C_{D,\mathrm{jet}}=\frac{\dot{m}\boldsymbol{V}_{\mathrm{j}}}{\frac{1}{2}\rho_\infty {V_\infty}^2 c}\cdot\boldsymbol{i} \tag{6}$$

式中: $\boldsymbol{V}_j$ 为射流速度的矢量; $\boldsymbol{i}$ 和 $\boldsymbol{j}$ 为阻力和升力方向的单位矢量。

1.4　计算方法验证

1.4.1　RAE2822 翼型验证

计算精度验证是仿真分析保证精确度的基础, 而网格无关性验证可以减小网格划分对计算结果的影响。因此, 为了在保证精度的同时选择合适的网格规模, 对 RAE2822 翼型进行精度验证与网格无关性验证。

采用 3 套不同规模的网格对 RAE2822 基准翼型进行数值模拟。网格划分采用 O 型拓扑结构, 远场半径为弦长的 20 倍, 并在翼型前缘与后缘处进行网格加密, 如图 2 所示。为保证 $y^+\approx1$, 第 1 层网格高度设定为 4×10^{-6} m。数值计算条件为: Ma=0.73, Re=6.5×10^6, α=2.79°。将 3 套网格的翼型压力系数分布结果与文献 [21] 的实验结果进行对比, 如图 3 所示。3 套网格参数与气动力系数的对比如表 1 所示。

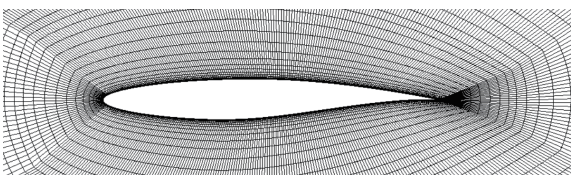


图 2　RAE2822 翼型计算网格

Fig. 2　Computational grids of RAE2822 airfoil

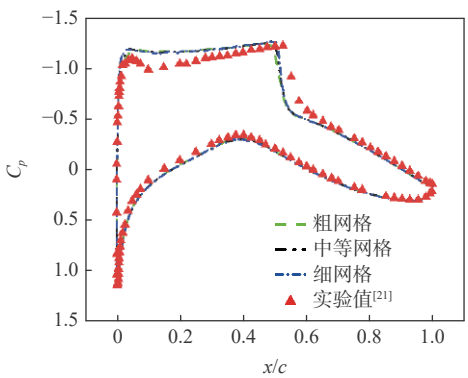


图 3　RAE2822 压力系数分布对比

Fig. 3　Comparison of pressure coefficient distribution of RAE2822

表 1　RAE2822 翼型网格参数与气动力系数的对比

Table 1　Comparison of grid parameters and aerodynamic coefficients of RAE2822 airfoil

网格	网格数量	C_L	C_D
实验值[21]		0.803	0.016 8
粗网格	23 324	0.760 31	0.016 989
中等网格	40 964	0.766 52	0.017 031
细网格	69 188	0.769 28	0.017 094

注: 3套网格与实验值相比, C_L误差5.3%, 4.5%, 4.2%, C_D误差为1.1%, 1.3%, 1.8%。

结合图 3 所示的压力系数分布对比可知, 翼型压力分布数值模拟结果与实验结果[21] 具有较好的一致性。此外, 由表 1 可观察到, 网格规模对升力系数与阻力系数的影响不大, 且与文献 [21] 的实验结果相比, 升力系数最大误差为 5.3%, 阻力系数最大误差为 1.8%, 误差均在合理的范围内, 进一步验证了计算方法的精确性。综合考虑网格量与计算精度, 将中等网格作为下表面射流翼型的网格划分依据。

1.4.2　CC020-010EJ 射流翼型验证

由于射流翼型的实验数据相对有限, 为了确保计算方法能够精确地捕捉射流的作用, 选取 CC020-010EJ 翼型对射流流动的相关数值模拟进行方法验证。计算条件为 Ma=0.1, Re=5.16×10^6, α=0°, c_μ=0.047, 并将数值仿真结果与文献 [22] 中的实验结果及数值结果进行对比, CC020-010EJ 射流翼型计算网格如图 4 所示。

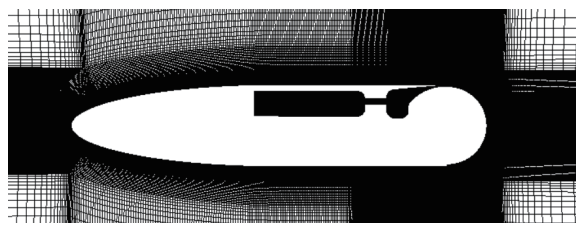


图 4　CC020-010EJ 翼型计算网格

Fig. 4　Computational grids of CC020-010EJ airfoil

图 5 为 CC020-010EJ 翼型压力系数分布的对比结果, 可知: 压力系数分布与文献 [22] 中的实验结果吻合较好。结合表 2 所示的 CC020-010EJ 翼型气动力数据对比可知: 数值模拟得到的升力和阻力均大于实验值, 除因在射流计算中使用的湍流模型及其他原因外, 还需考虑风洞试验中风洞壁的影响[6]。本文数值模拟结果虽与文献 [22]中的实验值有一定的误差, 但与文献 [22] 中的数值结果相比, 升力系数相差不大, 阻力系数误差减小了 6.1%, 从而验证了计算方法可以较准确地模拟射流流动。

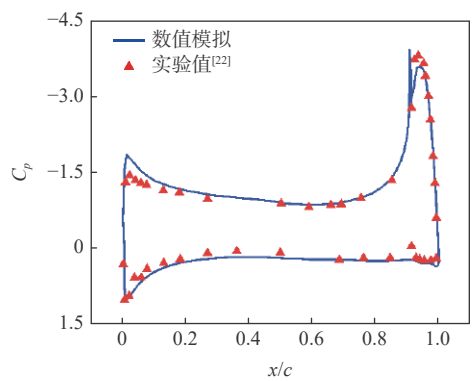


图 5　CC020-010EJ 压力系数分布对比

Fig. 5　Comparison of pressure coefficient distribution of CC020-010EJ

表 2 CC020-010EJ 翼型气动力系数的对比

Table 2 Comparison of aerodynamic coefficients of CC020-010EJ airfoil

结果	C_L	C_D
计算值	1.543	0.034 8
文献[22]数值结果	1.533	0.036 7
实验值[22]	1.363	0.031 2

注：本文，文献[22]计算值与实验值相比，C_L误差为13.2%，12.4%，C_D误差为11.5%，17.6%。

2 计算结果分析

2.1 下表面射流流动机理分析

为分析下表面射流影响翼型气动性能的机理，对 RAE2822 下表面射流翼型进行数值模拟，并与基准翼型进行对比。下表面射流翼型的计算网格如图 6 所示。为更好地捕捉射流出口的流动，在射流口处进行了网格的加密，同时保证下表面射流翼型网格规模与 RAE2822 翼型中等密度网格规模相同。计算参数如表 3 所示，所得的翼型马赫云图与流线图对比如图 7 所示。翼型压力系数分布对比如图 8 所示。

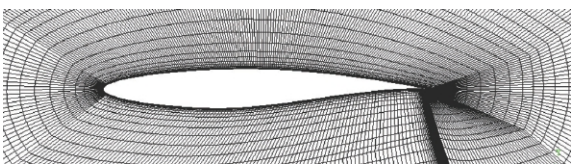


图 6 RAE2822 下表面射流翼型计算网格

Fig. 6 Computational grids of jet airfoil of RAE2822 lower surface

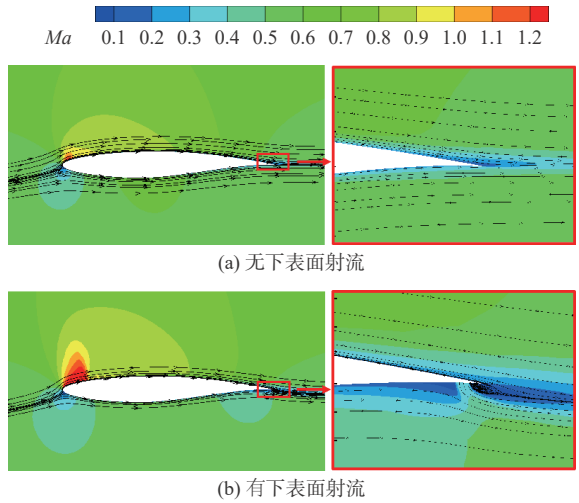
表 3 下表面射流计算参数

Table 3 Computational parameters of lower surface jet

Ma	Re	α/(°)	c_μ	x_j	θ/(°)
0.6	5.35×10^6	4	0.005	0.98c	90

结合图 7 可知，下表面射流使得前缘低压区增大，吸力峰附近马赫数增大。同时，下表面射流在翼型尾缘处形成逆时针的分离涡，使流经翼型上下表面的流线明显向下偏折，增加了翼型的等效弯度，形成了类似格尼襟翼的效果。因此，下表面射流的机理是在翼型后缘诱导产生逆时针分离涡，带动流线向下偏转，增加了翼型的等效弯度，同时增大了前缘吸力峰，从而提高翼型的升力。

结合图 8 所示的翼型压力系数分布对比图可知，下表面射流作用下压力系数曲线围成的面积有明显的增加，升力得到很大的提升。前缘吸力峰明显增大，上表面压力系数显著减小。此外，下表面射流口处的压力系数明显增大，这是由于射流阻碍了下表面的流动，增加了局部的压力。射流口之后，上下表面压力系数基本趋于一致。



(a) 无下表面射流

(b) 有下表面射流

图 7 马赫云图与流线图对比

Fig. 7 Comparison of Mach contours and streamlines

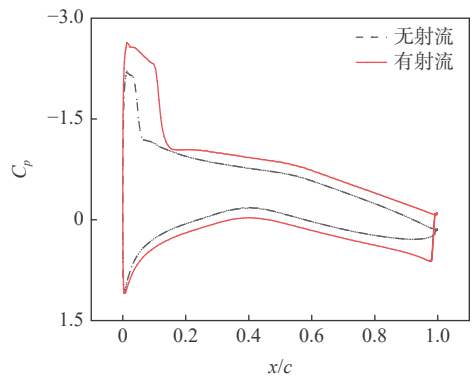


图 8 有/无射流压力系数分布对比

Fig. 8 Comparison of pressure coefficient distribution with/without jet

2.2 下表面射流关键参数对翼型气动性能的影响

2.2.1 不同射流位置的影响

为探究下表面射流位置对 RAE2822 翼型气动性能的影响，在给定 c_μ=0.005 下，针对 3 个不同射流位置 x_j 为 0.8c、0.9c、0.98c 进行对比分析，具体的计算参数设置如表 4 所示。求得的升力系数、阻力系数与极曲线如图 9 所示。气动力曲线均通过计算点之间的线性拟合得到。

表 4 不同射流位置下的计算参数

Table 4 Computational parameters of different jet positions

Ma	Re	α/(°)	c_μ	x_j	θ/(°)
0.6	5.35×10^6	−4～16	0.005	0.8c, 0.9c, 0.98c	90

结合图 9 可知，①对于升力系数，在射流动量系数一定时，射流位置越靠后，升力系数曲线越靠上，增升的效果越好。此外，当射流位置处于 0.98c 时，大迎角下的增升效果明显，而处于 0.8c 和 0.9c

时, 大迎角下的升力系数均低于无射流状态。②对于阻力系数, 在小迎角下, 射流位置越靠后, 阻力系数越小, 但整体变化不大; 大迎角下, 射流位置越靠后, 阻力系数越大。③对于极曲线, 射流位置越靠后, 极曲线整体越向上偏移, 因此气动性能越好。当射流位置处于 0.98c 时, 翼型气动性能明显优于无射流状态, 而处于 0.8c 和 0.9c 时的气动性能随着阻力的增大, 均低于无射流状态。

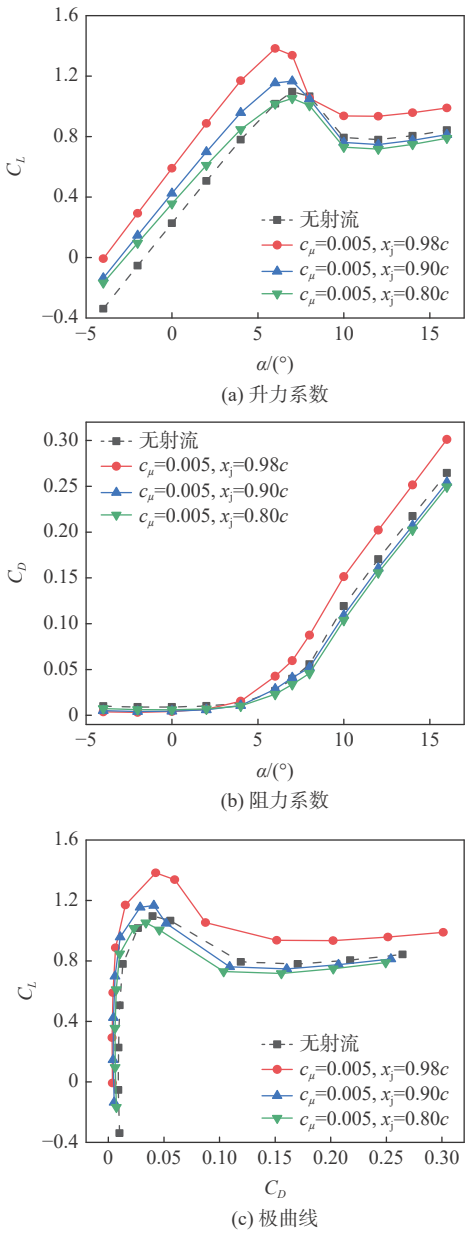


图 9　不同射流位置气动力系数对比

Fig. 9　Comparison of aerodynamic coefficients at different jet positions

为了进一步说明射流位置对翼型气动性能的影响, 选取 α=2°, c_μ=0.005 时不同射流位置马赫云图与流线图进行分析, 如图 10 所示。

由图 10 可知, 射流位置越靠后, 形成的低压分离区越靠后, 引起流线向下偏折的现象越明显。此外, 射流位置的后移增大了翼型前缘的吸力峰, 从而减小了上表面的压力, 提高了翼型的升力系数。因此, 可以得出以下结论: 下表面射流位置越靠近后缘, 翼型的气动性能越好。

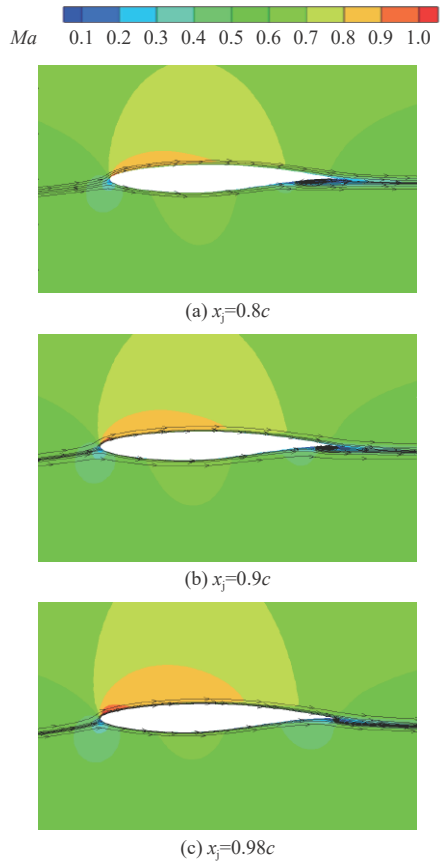


图 10　不同射流位置马赫云图与流线图对比

Fig. 10　Comparison of Mach contours and streamlines at different jet positions

2.2.2　不同射流动量系数的影响

由 2.2.1 节可知, 当下表面射流位置为 0.98c 时, 翼型具有最优的气动性能。为探究下表面射流动量系数对 RAE2822 翼型气动性能的影响, 在 x_j = 0.98c 基础上选取 c_μ为 0.000 5, 0.001, 0.005 这 3 种不同的射流动量系数进行对比分析, 具体的参数设置如表 5 所示。所得的升力系数、阻力系数与极曲线如图 11 所示。

表 5　不同射流动量系数下的计算参数

Table 5　Computational parameters of different jet momentum coefficients

Ma	Re	α/(°)	c_μ	x_j	θ/(°)
0.6	5.35×10^6	−4～16	0.000 5, 0.001, 0.005	0.98c	90

由图 11(a) 可知, 当射流位置一定时, 升力系数曲线随着射流动量系数的增大逐渐上移。因此, 高

射流动量系数可为翼型提供更大的升力。此外，当c_μ=0.005时，观察到失速迎角减小了1°，这是由于射流动量系数的增大导致流线在后缘向下偏折更明显，翼型等效弯度增大，前缘驻点下移，导致上表面流速加快，负压增大，更容易分离，因此失速迎角减小。由图11(b)可知，当α<4°时，下表面射流可减小翼型的阻力，同时，射流动量系数越大，翼型的减阻效果越好；但当$\alpha\geqslant$4°时，翼型的阻力系数随着射流动量系数的增大而增大。由图11(c)可知，射流动量系数越大，翼型的极曲线越靠上，翼型具有越优的气动性能。

为进一步量化射流动量系数对翼型气动性能的影响，选取α=4°、不同射流动量系数状态下翼型

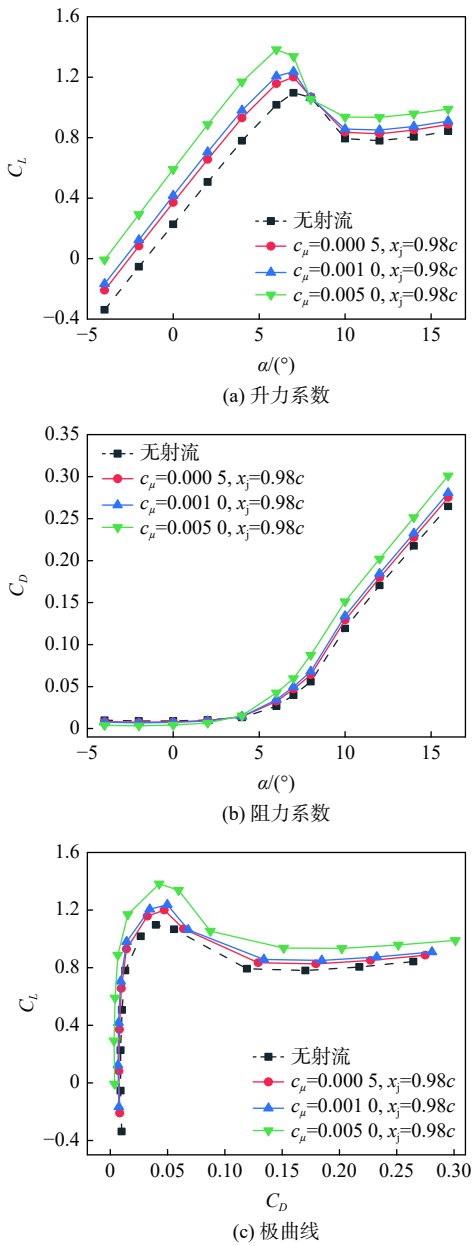


图11 不同射流动量系数下气动力系数对比图

Fig. 11 Comparison of aerodynamic coefficients at different jet momentum coefficients

的气动力系数进行分析，如表6所示。结合表6可知，当c_μ=0.005时，下表面射流翼型的升阻比提高了28.1%，说明下表面射流带来的升阻比增益非常可观。根据射流动量系数大小与翼型气动性能收益的关系，可以得出：射流动量系数越大，翼型的升阻比越高，气动性能越好。

表6 4°迎角下不同射流动量系数翼型的气动力系数

Table 6 Aerodynamic coefficients of airfoil with different jet momentum coefficients at angle of attack 4°

c_μ	C_L	C_D	C_L/C_D
0	0.779 82	0.013 095	59.55
0.000 5	0.930 05	0.014 121	65.86
0.001	0.980 81	0.014 366	68.27
0.005	1.170 1	0.015 337	76.29

2.2.3 不同射流角度的影响

为探究不同射流角度对翼型气动性能的影响，在x_j=0.98c，c_μ=0.005的基础上，对小迎角下的不同射流角度进行仿真分析，计算参数设置如表7所示。

表7 不同射流角度计算参数

Table 7 Computational parameters of different jet angles

Ma	Re	α/(°)	c_μ	x_j	θ/(°)
0.6	5.35×10^6	0, 2, 4	0.005	0.98c	10～170

图12为不同射流角度下各个气动力系数的对比，可以得出：①在小迎角下，下表面射流翼型的升力系数均优于无射流状态，且升力系数随射流角度的增大整体呈现先增大后减小的趋势，最大升力系数出现在θ=70°时；②在小迎角下，下表面射流的阻力系数整体上随着射流角度的增加呈现出减小的趋势，α分别为0°、2°、4°时翼型对应的最小阻力系数分别出现在θ为130°、160°和170°。大射流角度下阻力系数小的原因是射流反作用力在阻力方向上的投影向前，从而产生推力，减小了翼型的总阻力系数。③下表面射流翼型的升阻比随射流角度的增大呈现先增加后减小的趋势。当α=0°和2°时，最大升阻比θ=110°，相比于无射流状态，θ=110°时的下表面射流可使翼型在0°和2°时升阻比分别提高了484.5%和189.5%。当α=4°时，最大升阻比θ=160°，θ=160°时的下表面射流可使翼型升阻比提高了60.5%。综上所述，当α=0°和2°时，最优射流角度为110°，当α=4°时，最优射流角度为160°。

2.2.4 不同马赫数的影响

为进一步探究小迎角下，不同马赫数对下表面射流翼型气动性能的影响，在前文的基础上，对于

α=0°和 2°时, 选取 x_j =0.98c, c_μ=0.005, θ=110°, 对于 α=4°时, 选取 x_j =0.98c, c_μ=0.005, θ=160°, 研究 *Ma* 在 0.3～0.6 范围内下表面射流翼型的气动性能, 并进行对比分析, 图 13 为不同马赫数下翼型各个气动力系数的变化, 计算参数如表 8 所示。

可得: ①升力系数方面, 下表面射流提高了小迎角下翼型 *Ma* 在 0.3～0.6 区间内的升力系数, 且从整体上看, α=0°和 2°时下表面射流的增升效果优于 α=4°。②阻力系数方面, 下表面射流明显降低了小迎角下翼型 *Ma* 在 0.3～0.6 区间内的阻力系数, 有效减小了翼型的阻力。③升阻比方面, 可以观察到, 在小迎角状态下, 下表面射流 *Ma* 在 0.3～0.6 区间内显著提高了翼型的升阻比, 说明下表面射流不仅能提高超临界翼型在高速状态下的气动性能, 还能显著改善翼型的低速性能。综上所述, 下表面射流可以有效提高 *Ma* 在 0.3～0.6 范围内翼型的气动性能。

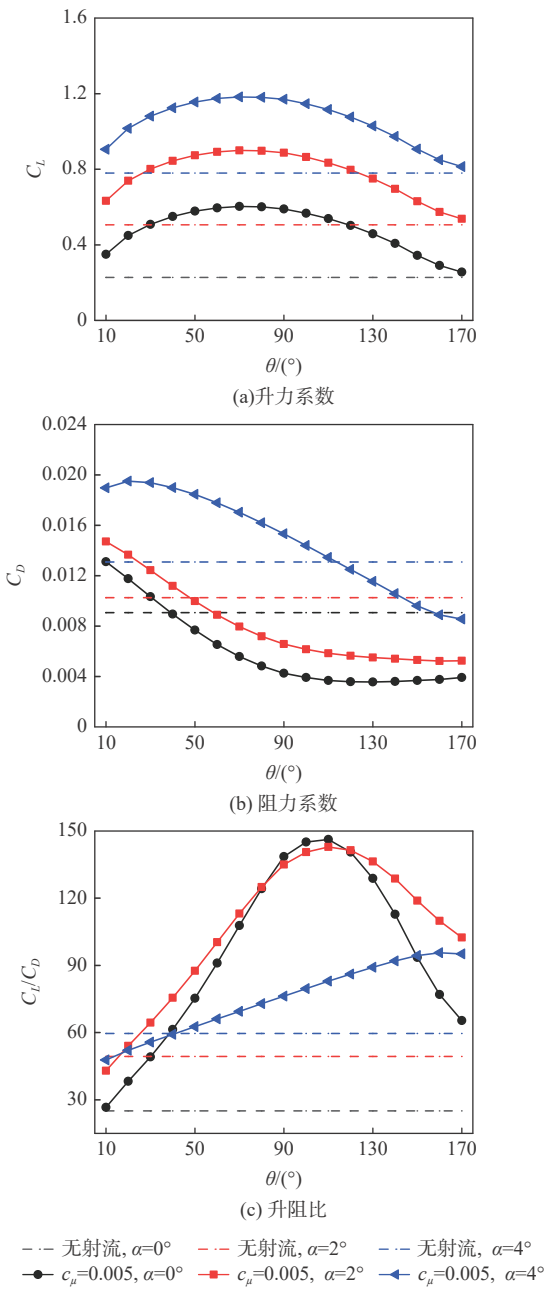


图 12　不同射流角度下各个气动力系数对比

Fig. 12　Comparison of aerodynamic coefficients at different jet angles

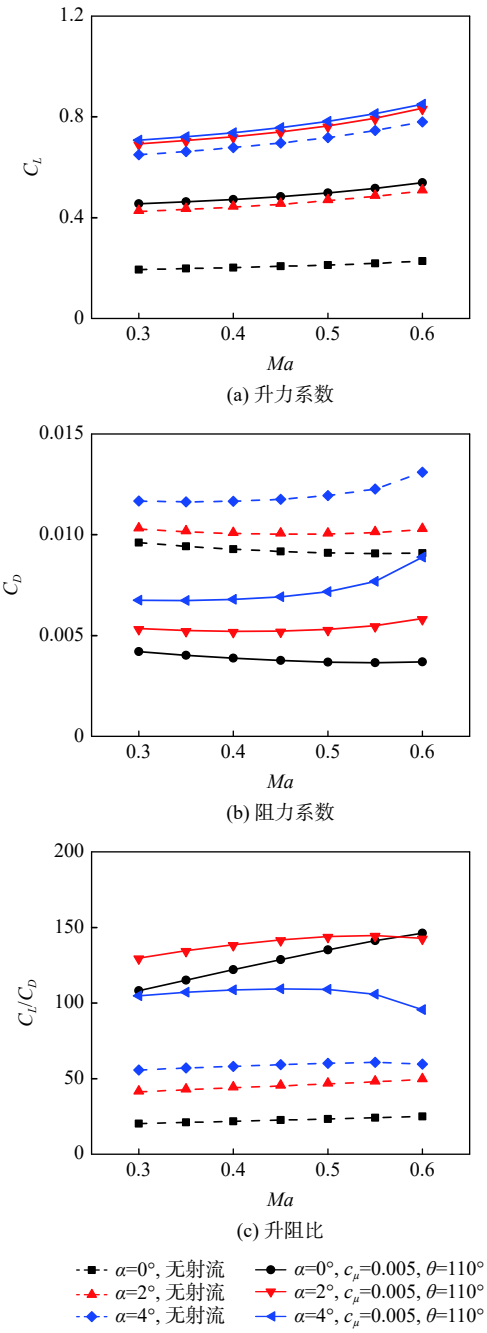


图 13　不同马赫数下各个气动力系数对比

Fig. 13　Comparison of aerodynamic coefficients at different Mach numbers

表 8　不同马赫数下的计算参数

Table 8　Computational parameters of different Mach numbers

Ma	*Re*	α/(°)	c_μ	x_j	θ/(°)
0.3～0.6	2.68×10^6～5.35×10^6	0、2、4	0.005	0.98c	110, 160

3 结 论

1) 本文所采用的基于RANS与S-A湍流模型的数值模拟方法具有足够的精确性，且采用的计算网格通过网格无关性检验。

2) 下表面射流能够在RAE2822翼型后缘诱导产生逆时针分离涡，带动流线向下偏折，增加了翼型的等效弯度，同时，加大了前缘的吸力峰，减小上翼面压力系数，从而提高翼型的气动性能。

3) 给定状态下，下表面射流位置越靠后，动量系数越大，RAE2822翼型的气动性能越优。下表面在 α=0°和2°时最优射流角度为110°，在 α=4°时最优射流角度为160°，且在最优射流角度下，下表面射流可以有效提高 Ma 在0.3～0.6范围内翼型的升阻比，从而改善翼型的气动性能。

研究结果为基于下表面射流的主动流动控制提供了一定的参考与补充，后续工作将进一步探索下表面射流对三维机翼的影响及下表面射流的控制优化等问题。

参考文献（References）

[1] SCOTT COLLIS S, JOSLIN R D, SEIFERT A, et al. Issues in active flow control: Theory, control, simulation, and experiment[J]. Progress in Aerospace Sciences, 2004, 40(4-5): 237-289.

[2] ASHILL P R, FULKER J L, HACKETT K C. A review of recent developments in flow control[J]. The Aeronautical Journal, 2005, 109(1095): 205-232.

[3] SHMILOVICH A, YADLIN Y. Active flow control for practical high-lift systems[J]. Journal of Aircraft, 2009, 46(4): 1354-1364.

[4] FU Z J, CHU Y W, CAI Y S, et al. Numerical investigation of circulation control applied to flapless aircraft[J]. Aircraft Engineering and Aerospace Technology, 2020, 92(6): 879-893.

[5] REN F, HU H, TANG H. Active flow control using machine learning: A brief review[J]. Journal of Hydrodynamics, 2020, 32(2): 247-253.

[6] 朱自强, 吴宗成. 环量控制技术研究[J]. 航空学报, 2016, 37(2): 411-428.
ZHU Z Q, WU Z C. Study of the circulation control technology[J]. Acta Aeronautica et Astronautica Sinica, 2016, 37(2): 411-428(in Chinese).

[7] XU H Y, QIAO C L, YANG H Q, et al. Active circulation control on the blunt trailing edge wind turbine airfoil[J]. AIAA Journal, 2017, 56: 554-570.

[8] SHI Z W, ZHU J C, DAI X X, et al. Aerodynamic characteristics and flight testing of a UAV without control surfaces based on circulation control[J]. Journal of Aerospace Engineering, 2019, 32(1): 04018134.

[9] 雷玉昌, 张登成, 张艳华. 环量控制翼型非定常气动力建模[J]. 北京航空航天大学学报, 2021, 47(10): 209-219.
LEI Y C, ZHANG D C, ZHANG Y H. Unsteady aerodynamic modeling of circulation control airfoil[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(10): 209-219(in Chinese).

[10] 雷玉昌, 张登成, 张艳华, 等. 脉冲射流对环量控制翼型气动性能的影响[J]. 北京航空航天大学学报, 2022, 48(3): 485-494.
LEI Y C, ZHANG D C, ZHANG Y H, et al. Effect of pulsed jet on aerodynamic performance of circulation control airfoil[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(3): 485-494(in Chinese).

[11] ZHA G C, CARROLL B F, PAXTON C D, et al. High-performance airfoil using coflow jet flow control[J]. AIAA Journal, 2007, 45(8): 2087-2090.

[12] LEFEBVRE A, DANO B, BARTOW W B, et al. Performance and energy expenditure of coflow jet airfoil with variation of mach number[J]. Journal of Aircraft, 2016, 53(6): 1757-1767.

[13] XU K W, ZHA G C. High control authority three-dimensional aircraft control surfaces using coflow jet[J]. Journal of Aircraft, 2021, 58(1): 72-84.

[14] 刘沛清, 旷建敏, 屈秋林. 联合射流控制技术的增升效果和机理[J]. 北京航空航天大学学报, 2009, 35(6): 737-740.
LIU P Q, KUANG J M, QU Q L. Effect and mechanism of lift-enhancement of the co-flow jet technology[J]. Journal of Beijing University of Aeronautics and Astronautics, 2009, 35(6): 737-740(in Chinese).

[15] 许建华, 李凯, 宋文萍, 等. 低雷诺数下协同射流关键参数对翼型气动性能的影响[J]. 航空学报, 2018, 39(8): 122018.
XU J H, LI K, SONG W P, et al. Influence of co-flow jet key parameters on airfoil aerodynamic performance at low Reynolds number[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(8): 122018(in Chinese).

[16] 张志勇, 王团团, 陈志华, 等. 低雷诺数下吹吸气射流对NACA0012翼型气动性能的影响[J]. 空气动力学学报, 2020, 38(1): 58-65.
ZHANG Z Y, WANG T T, CHEN Z H, et al. The effect of blowing/suction jet on the aerodynamic performance of airfoil NACA0012 at low Reynolds number[J]. Acta Aerodynamica Sinica, 2020, 38(1): 58-65(in Chinese).

[17] HUANG L, HUANG P G, LEBEAU R P, et al. Numerical study of blowing and suction control mechanism on NACA0012 airfoil[J]. Journal of Aircraft, 2004, 41(5): 1005-1013.

[18] GENÇ M, KAYNAK Ü. Control of flow separation and transition point over an aerofoil at low Re number using simultaneous blowing and suction[C]//19th AIAA Computational Fluid Dynamics. Reston: AIAA, 2009: 3672.

[19] 付云豪, 章卫国, 史静平, 等. 下表面射流对翼型气动性能影响的数值模拟[J]. 哈尔滨工业大学学报, 2021, 53(6): 48-53.
FU Y H, ZHANG W G, SHI J P, et al. Numerical simulation of influence of jet at lower surface on aerodynamic performance of airfoil[J]. Journal of Harbin Institute of Technology, 2021, 53(6): 48-53(in Chinese).

[20] XU H Y, QIAO C L, YANG H Q, et al. Active circulation control on the blunt trailing edge wind turbine airfoil[J]. AIAA Journal, 2018, 56(2): 554-570.

[21] COOK P, MCDONALD M A, FIRMIN M C P. Aerofoil RAE2822 pressure distributions, and boundary layer and wake measurements, experimental data base for computer program assessment: AGARD

Report AR-138[R]. Advanced Guidance for Alliance Research and Development, Part of NATO Science & Technology Organization, 1979, 138: 47.

[22] ENGLAR R, JONES G, ALLAN B, et al. 2-D circulation control airfoil benchmark experiments intended for CFD code validation[C]//47th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition. Reston: AIAA, 2009: 902.

Aerodynamic performance analysis of supercritical airfoil with lower surface jet

WANG Ruochen[1, 2], ZHANG Guoxin[1, *], WANG Xiangyu[1], MA Xiaoping[1]

(1. Institute of Engineering Thermophysics, Chinese Academy of Sciences, Beijing 100190, China;
2. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: To investigate the influence of key parameters of the lower surface jet on aerodynamic performance of supercritical airfoil, a numerical simulation is performed using Reynolds average Navier-Stokes (RANS) equation and Spalart-Allmaras (S-A) turbulence model. By comparing the flow field between the benchmark RAE2822 airfoil and the lower surface jet airfoil, the lower surface jet induces a counterclockwise separation vortex at the trailing edge of the airfoil, deflecting the streamlines downward. This deflection increases the equivalent camber of the airfoil and the suction peak of the leading edge, leading to the aerodynamic performance enhancement of the airfoil. The effects of key parameters such as the jet position, the jet momentum coefficient, the jet angle and the Mach number on the aerodynamic performance of RAE2822 airfoil are explored. The results show that under given conditions, the closer the lower surface jet is to the trailing edge and the greater the momentum coefficient is, the better the aerodynamic performance of the airfoil. The optimal angle of the lower surface jet is 110° at α=0° and 2°, and 160° at α=4°. The aerodynamic performance of the airfoil can be effectively improved at the optimal angles with the Mach number varied from 0.3 to 0.6.

Keywords: active flow control; lower surface jet; supercritical airfoil; aerodynamic performance; Reynolds average Navier-Stokes

Received: 2021-08-25; **Accepted**: 2021-09-10; **Published Online**: 2021-09-24 12:09

URL: kns.cnki.net/kcms/detail/11.2625.V.20210923.1811.001.html

Foundation items: National Natural Science Foundation of China (61901448); Youth Innovation Promotion Association of Chinese Academy of Sciences (2020149)

* **Corresponding author**. E-mail: zhangguoxin@iet.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0494

撤机对离心血泵血液相容性影响的数值研究

谢楠[1,2]，唐雨萌[1,2]，张岩[3]，柳阳威[1,2,*]

(1. 北京航空航天大学 能源与动力工程学院，北京 100191；2. 北京航空航天大学 航空发动机气动热力国家级重点实验室，北京 100191；3. 中国医学科学院 阜外医院，北京 100037)

摘　　要：采用非定常数值模拟方法系统评估某离心血泵原型与改型在4个撤机阶段的血流动力学特性和血液相容性。对速度、湍动能、叶表压力的分布规律及不同撤机阶段的血液损伤进行研究，并对各监测点的压力脉动进行分析。研究结果表明：在降低心脏泵流量和供压的第1阶段（撤机阶段Ⅰ），原型和改型均有较好的血液相容性，满足抗溶血和抗血栓的性能要求，间隙泄漏涡是造成溶血主要的流动特征；随着撤机的进行，间隙泄漏涡强度明显减弱，溶血值呈显著的下降趋势，其脉动幅度也大大降低，改型的溶血值比原型减少30%以上；各监测点压力频谱呈现典型的离散特征，相较于原型，改型的转静干涉作用更加强烈，对血泵的电磁控制技术提出更高的要求。

关 键 词：离心血泵；撤机；血液相容性；溶血；数值模拟

中图分类号：V239; R318

文献标志码：A　　**文章编号**：1001-5965(2023)07-1680-09

根据国家卫生健康委员会预测，中国心血管病患病率仍处于上升阶段，其死亡率居首位，2019年中国心血管病患者数约3.3亿[1]。心血管病给社会带来的经济负担日渐加重，已成为重大的公共卫生问题。心脏移植是诸多治疗手段中较为有效的方法。但由于自然心脏捐献数量远不能满足病人的需求，因此，开发人工心脏替代自然心脏来维持生命成为主要的治疗手段。

人工心脏从最初的第1代仿生搏动式人工心脏，到第2代有机械支撑轴承的叶片式人工心脏，再到第3代无机械轴承的人工心脏，其缺陷逐代得到改进，工作性能更加稳定，血液相容性也显著改善。常用于临床的心脏泵类型主要为离心式和轴流式。轴流血泵轴向长而轴颈小、运转稳定性好、整体体积小因而植入性好，常用于植入式心室辅助系统(ventricular assist devices, VAD)，但加工工艺要求高，并且转速高、不利于控制其抗溶血性能。离心血泵轴颈大而轴向短、整体体积大、相同的血压及流量要求下，离心泵的转速低，常用于VAD及体外膜肺氧合(extracorporeal membrane oxygenation, ECMO)，但由于体积较大，导致血细胞在血泵内的暴露时间较长，同样不利于控制其抗溶血性能。ECMO主要用于等待器官移植或心肺功能严重衰竭的患者提供体外循环，其核心器械为人工肺和人工心脏泵。在新型冠状病毒肺炎疫情期间，医护人员采用ECMO成功救治多例肺炎重症患者，临床效果显著[2]。撤机是指患者采用ECMO治疗后，其原发病可能会得到控制，心肺功能逐渐得到改善，逐步撤除ECMO对心脏和呼吸的支持，使患者自然心脏提供全部血压、恢复完全自主呼吸的过程。王军红和马青变[3]认为过早撤机可能会导致患者病情恶化，需要紧急心肺支持并再次插管，容易引发心力衰竭甚至死亡；而延长ECMO支持同样可能导致不必要的并发症或死亡。因此，在撤机过程中，医

收稿日期：2021-08-27；**录用日期**：2022-01-02；**网络出版时间**：2022-01-29 08:30
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220128.1659.004.html
基金项目：国家自然科学基金(51976006,52106039)；航空科学基金(2018ZB51013)
*通信作者. E-mail：liuyangwei@126.com

引用格式：谢楠，唐雨萌，张岩，等. 撤机对离心血泵血液相容性影响的数值研究[J]. 北京航空航天大学学报，2023，49(7)：1680-1688.
XIE N，TANG Y M，ZHANG Y，et al. Numerical study of blood pump weaning effects on hemocompatibility of centrifugal blood pump[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1680-1688 (in Chinese).

护人员会系统评估患者的心肺功能、血流动力学等因素, 并在撤机过程中采取分阶段降低血泵流量与供压的方法以提高 ECMO 的成功率。

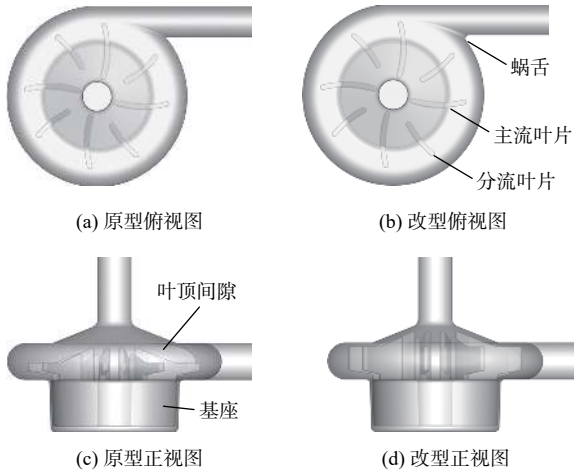
当前, 离心血泵的血液相容性设计, 与血泵配套的电机设计及相应控制技术等都是亟需解决的核心问题[4]。研究血泵的手段主要包括实验研究和数值模拟。实验研究又包括水力学特性测量实验、动物实验及临床实验。但由于血泵部件加工要求较高, 动物实验和临床实验周期长, 耗资很大。为此国内外大部分研究人员在设计人工心脏时都会采用计算流体力学(computational fluid dynamics, CFD)方法来预测血泵内部流动情况[5], 随后对可改进的部位进行优化, 最终设计出血液相容性好的血泵。CFD 方法可以大大降低研发成本, 缩短血泵的设计周期。但预测结果高度依赖于网格精度、边界条件、流动模型及溶血预测模型等。Liu 等[6-7] 采用 CFD 充分研究叶轮机械内部的复杂流动, 为改型优化提供参考。张岩等[8-10] 对轴流血泵的研制开展了一系列的工作, 并通过 CFD 技术优化血泵内部流道结构, 消除了流场中危害性较大的流动特征, 有效提升了血泵的血液相容性, 最终通过实验研究验证了数值模拟的结果。

本文基于中国患者的临床使用情况, 采用数值模拟方法, 研究撤机过程对离心血泵血液相容性的影响, 对比原型与改型的性能差异, 为血泵的临床应用及进一步优化设计提供指导。

1　研究对象几何及数值方法

1.1　离心血泵模型及运行工况

本文的离心血泵由模块化泵壳、离心叶轮及内置于离心叶轮基座的永磁体构成, 离心叶轮包含 4 个主流叶片、4 个分流叶片和离心叶轮基座, 分流叶片采取偏置设计, 与上游主流叶片距离较小, 与下游主流叶片距离较大。前期研究中, 为降低叶顶间隙泄漏涡强度并提高血液相容性对离心血泵进行了一些改进, 详细几何参数参见文献 [11]。本文通过三维建模软件 UG NX 12.0 对原型和改型进行建模, 2 种模型的俯视图及正视图如图 1 所示, 原型血泵为了追求较为宽广的工作范围, 其流量范围为 0～10 L/min[12], 为保证血泵能够在任意流量条件下提供充足的血压, 叶轮最大转速较高。因此, 为避免叶轮高速转速时在间隙内引起较大的剪切应力, 血泵原型采取大间隙的设计理念。其叶顶间隙为 5.0 mm, 蜗舌无倒圆。国内患者较优的流量工况为 3 L/min[10], 原型血泵在国内临床应用时的转速也较低, 因而在保证血液相容性的前提下允许适当减小间隙, 从而增加叶片做功面积。改型叶顶间隙为 1.0 mm, 蜗舌倒圆为 1.0 mm。



(a) 原型俯视图　(b) 改型俯视图

(c) 原型正视图　(d) 改型正视图

图 1　血泵原型与改型的俯视图及正视图

Fig. 1　Top and front views of datum and modified blood pumps

计算域划分为 3 个部分: 进口静止域、离心叶轮旋转域和出口静止域, 使用 ANSYS Meshing 在上述 3 个计算域中划分四面体网格, 在壁面附近划分 10 层以上的三棱柱网格以确保能够准确捕捉边界层, 近壁面三棱柱网格 y^+≈1。原型及改型采取同样的网格划分设置, 并进行网格无关性检验。最终, 原型网格数为 8.68×10^6, 改型网格数为 8.90×10^6, 原型叶轮表面网格如图 2 所示。

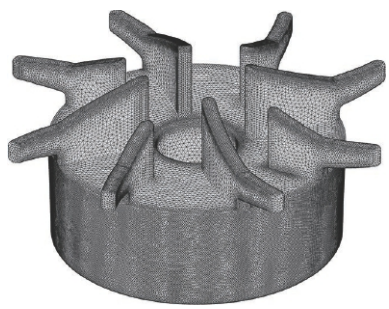


图 2　原型叶轮表面网格

Fig. 2　Surface grid of datum impeller

文献 [13-14] 研究了切应变和血液黏性之间的关系, 发现当切应变大于 100/s 时, 血液流动特性近似为牛顿流体。上述研究结论可以根据文献 [15-16] 的实验结果得到验证。人体血液黏度约为 0.003～0.004 Pa·s, 本文血液黏度为 0.003 5 Pa·s, 血液密度为 1 050 kg/m^3。基于国内患者临床应用情况确定原型血泵在不同撤机阶段的运行工况, 如表 1 所示, 并通过定常计算方法确定改型在相同流量下, 达到相同供压所需的转速。撤机阶段 Ⅰ 处于血泵正常运转工况, 随着撤机的进行, 血泵的流量和供压均逐步下降, 自然心脏的泵血机能逐渐恢复, 当撤机进行到阶段Ⅳ时, 如果患者的生命体征平稳且

心肺功能仍然稳定，则彻底关闭血泵完成撤机，使自然心脏提供全部的血压。

表 1 血泵撤机阶段运行工况

Table 1 Operating conditions in blood pump weaning stage

撤机阶段	流量/(L·min^{-1})	供压/Pa	原型转速/(r·min^{-1})	改型转速/(r·min^{-1})
I	3.0	26 707	3 200	2 622
II	2.5	19 620	2 750	2 255
III	2.0	14 892	2 400	1 971
IV	1.0	5 728	1 500	1 237

1.2 数值方法及计算设置

数值模拟通过商用软件 ANSYS Fluent 19.2 进行。相关研究表明血泵采用层流模型预测到的结果与实验值相差较大[17]，重整化群(renormalization group, RNG)k-ε和雷诺应力模型(Reynolds stress model, RSM)的预测结果与实验值误差较小。本文数值模拟采用 RNG k-ε湍流模型，在近壁区选用增强壁面处理，并通过解压力耦合方程的半隐式法(semi-implicit method for pressure-linked equations-consistent, SIMPLEC)求解压力-速度耦合方程。非定常计算以定常计算结果作为初始条件，并采用滑移网格方法，叶轮旋转 1 周分为 360 个物理时间步，每个物理时间步设置 50 个虚拟时间步，收敛判据为残差达到 10^{-6}，非定常计算 8 周以保证计算达到稳定。边界条件选择速度入口，压力出口，不考虑重力影响。

1.3 溶血预测模型

1965 年，Blackshear 等[18]基于实验结果提出预测血液损伤的幂律表示形式。Giersiepen 等[19]根据其实验中测定的游离血红蛋白浓度与总血红蛋白浓度的比值，进一步优化了 Blackshear 等[18]的血液损伤预测模型，提出以下形式：

$$D=\frac{\Delta \mathrm{Hb}}{\mathrm{Hb}}\%=3.62\times10^{-5}\tau^{2.416}t^{0.785} \tag{1}$$

式中：ΔHb为游离血红蛋白浓度；Hb为总血红蛋白浓度，本文取 140 g/L；τ为作用于红细胞的有效剪切应力；t为红细胞在血泵内的滞留时间。

本文采用的血液损伤预测模型是由 Garon 和 Farinas [20]基于双曲线型运输方程提出的一种快速预测血液损伤的模型。此模型直接从速度场和剪应力场计算标准溶血指数。双曲线型运输方程为

$$\left(\frac{\partial}{\partial t}+\nabla V\right)D_{\mathrm{I}}=\sigma \tag{2}$$

式中：V为速度；∇为梯度算子；D_{I}为引入的沿流线积分的线性损伤；σ为单位时间溶血破坏率：

$$\sigma=(3.62\times10^{-7})^{1/0.785}\tau^{2.416/0.785} \tag{3}$$

$$D_{\mathrm{I}}=D^{1/0.785}=(3.62\times10^{-7})^{1/0.785}\tau^{2.416/0.785}t \tag{4}$$

有效剪切应力的计算式[21]为

$$\tau=\left[\frac{1}{2}\left[(\sigma_1-\sigma_2)^2+(\sigma_2-\sigma_3)^2+(\sigma_3-\sigma_1)^2\right]\right]^{\frac{1}{2}} \tag{5}$$

式中：σ_1、σ_2、σ_3为作用于红细胞的主应力。血泵计算域产生的平均线性损伤计算式为

$$\overline{D}_{\mathrm{I}}=\frac{1}{Q}\int_V\sigma\mathrm{d}V \tag{6}$$

式中：Q为流量。

随后通过指数换算可得血泵溶血值为

$$D=\overline{D}_{\mathrm{I}}^{0.785} \tag{7}$$

将溶血值转化为标准溶血指数(normalized index of hemolysis, NIH)：

$$\mathrm{NIH}=\mathrm{Hb}\times D\times100 \tag{8}$$

2 结果和分析

2.1 水力性能分析

采用非定常计算方法对血泵原型与改型在不同撤机阶段进行数值模拟，血泵半叶高截面的速度分布如图 3 所示。原型与改型的叶片尾缘圆周速度随着撤机的进行而逐渐降低，显著降低了红细胞撞击壁面的速度及红细胞受到的剪切应力。但是

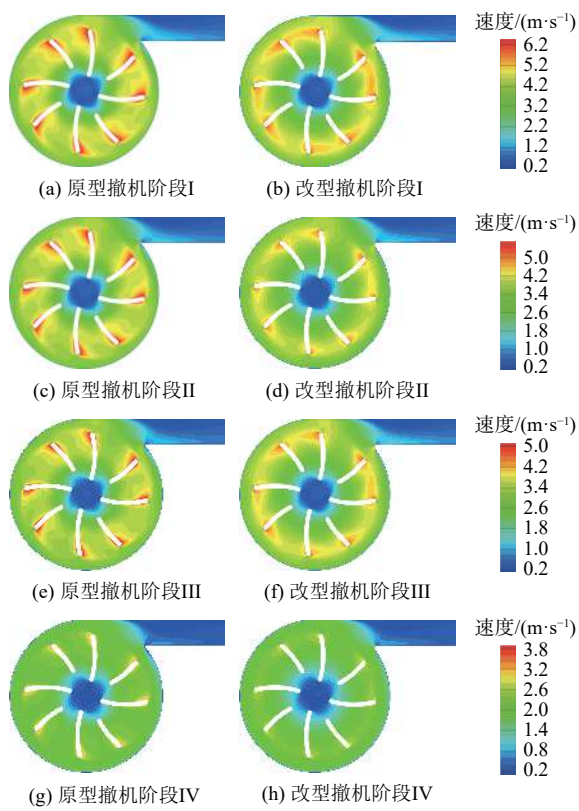


图 3 半叶高截面速度云图

Fig. 3 Velocity counters at mid span plane of impeller

随着流量下降, 泵内流速降低, 同时会增加血细胞在泵内的滞留时间。无论是原型还是改型都预测到出口管内存在大尺度的分离区, 分离区内红细胞滞留时间过长, 但由于出口管内的剪切应力较低, 不会造成严重的溶血情况, 此外, 原型蜗舌下游还存在小尺度分离区。改型在蜗舌处采用倒圆设计, 利用 Coanda 效应减小了分离区, 有效降低了形成血栓的风险。总体而言, 相较于原型, 改型的速度分布更加均匀, 体现出改型更优越的抗溶血性能。

国际上主流商业化血泵的设计指标比较符合欧美患者对心脏供血量的需求[12]。在国内临床应用中离心血泵流量较小, 导致液流角与叶片安放角相差较大, 因而叶片前缘附近存在一定的冲击损失和旋涡损失。在撤机阶段, 转速大幅降低, 有利于降低前缘冲击损失和旋涡损失。此外, 在叶轮通道及蜗室内, 流动受到压力梯度的影响, 容易产生二次流和分离区, 并形成不同尺度的旋涡结构。血泵实际运转过程中, 由于叶片与泵壳之间存在相对运动, 必须保持一定间隙。叶片压力面和吸力面存在压力差, 从而驱动叶片压力侧流体越过叶顶间隙, 形成泄漏射流。泄漏射流进入叶片吸力侧后, 与主流相互作用, 流动卷曲形成流向涡, 称为泄漏涡[22]。

离心泵叶片通道中存在的二次流还包括潜流, 潜流是指在叶片边界层内的流体由叶根流向叶尖。潜流的形成主要由于血液轴向流入离心叶轮, 而后径向流出, 当流线在叶片通道内发生弯曲时, 血液会受到离心力的影响。为平衡离心力, 叶根处的压力一般大于叶顶。叶片边界层内的血液流速低于主流区, 其离心力不足以抵抗强逆压梯度, 因此, 叶片吸力侧与压力侧边界层内的血液向叶顶流动。叶表通道涡的形成与叶表潜流息息相关, 受边界层内潜流的影响, 叶片通道中央部分的血液随之补偿到叶片表面, 在主流流动的影响下最终形成通道涡。叶顶泄漏流与叶表潜流是边界层内主要的二次流, 受叶顶间隙区域回流的影响, 叶片吸力侧边界层内泄漏流流动方向向下, 而潜流则是由叶根向上流动, 因而在吸力侧这 2 种二次流会相互抑制。

将离心血泵的叶片通道分为 2 个区, 将主流叶片吸力侧和分流叶片压力侧组成的叶片通道记为通道 I, 将主流叶片压力侧和分流叶片吸力侧组成的叶片通道记为通道 II。图 4 为原型叶轮通道内半径 r=14 mm 处湍动能云图, 叶片旋转方向为从右向左, 定义出口管与蜗壳过渡段圆周角 θ=0。泄漏涡是叶轮通道内最主要的旋涡结构, 由于叶顶间隙较大, 射流进入叶片吸力侧后, 与主流充分掺混, 卷起更多低能流体, 形成较大尺度的高湍动能区域。原

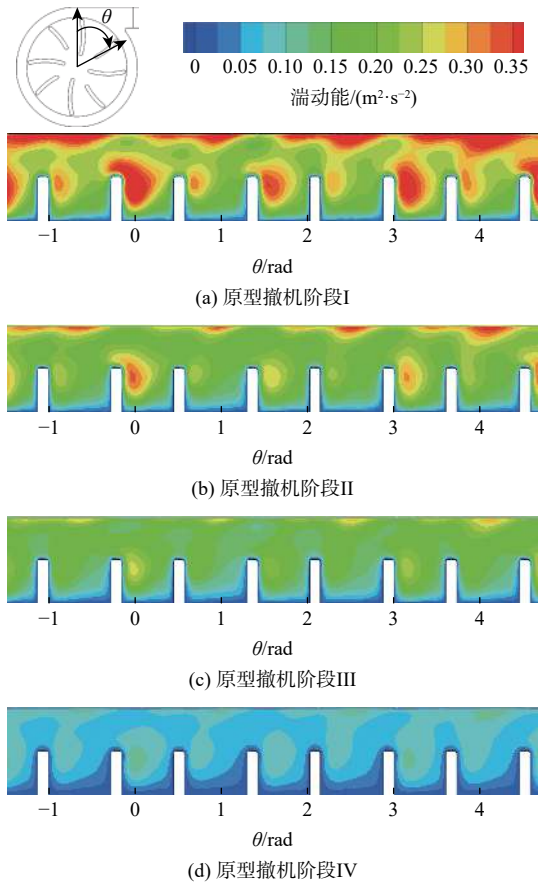


图 4　原型血泵半径 r=14 mm 湍动能云图

Fig. 4　Turbulent kinetic energy counters at r=14 mm plane of datum pump

型血泵叶顶间隙区域内的血液受叶轮带动作用较弱, 受下游强逆压梯度的影响, 在叶顶间隙内存在一股由蜗壳流向叶片前缘的血流, 这股流动与进口来流方向相反, 两股流动在叶顶间隙内掺混并产生旋涡。通道涡并不明显, 主要是由于叶片较低, 转速较高, 潜流的发展受到泄漏流的抑制作用。主流叶片吸力侧的泄漏流强度大于分流叶片, 其原因是由于主流叶片弦长较长, 泄漏涡得以充分发展。随着撤机的进行, 泄漏涡与叶顶间隙内的旋涡强度明显减弱, 造成的水力损失大大减少, 有效改善了红细胞所处的流场环境, 减轻了由于高剪切应力使得红细胞破裂导致溶血的情况。

图 5 为改型叶轮通道内半径 r=14 mm 处湍动能云图, 泄漏涡仍然是叶轮通道内最主要的旋涡结构, 但由于叶顶间隙较小, 泄漏涡能够卷起的低能流体有限, 形成的高湍动能区尺度明显小于原型。叶顶间隙内的回流现象基本消失, 因此, 上泵壳壁面附近已经没有明显的旋涡结构。叶片高度增加使得叶片做功能力增强, 在相同供压需要下, 改型所需的转速较小, 通道 I 内泄漏流强度随之下降, 不再能抑制潜流的发展, 通道涡影响范围也随之扩

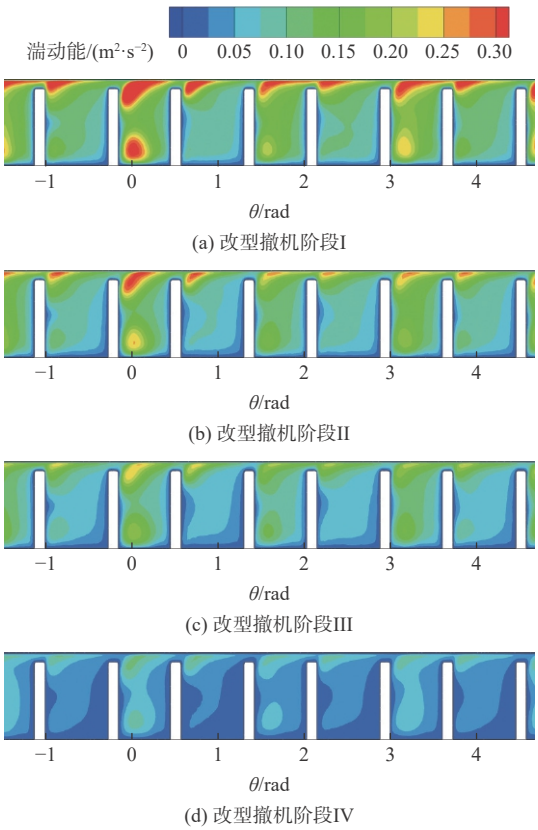


图5 改型血泵半径 r=14 mm 湍动能云图

Fig. 5 Turbulent kinetic energy counters at r=14 mm plane of modified pump

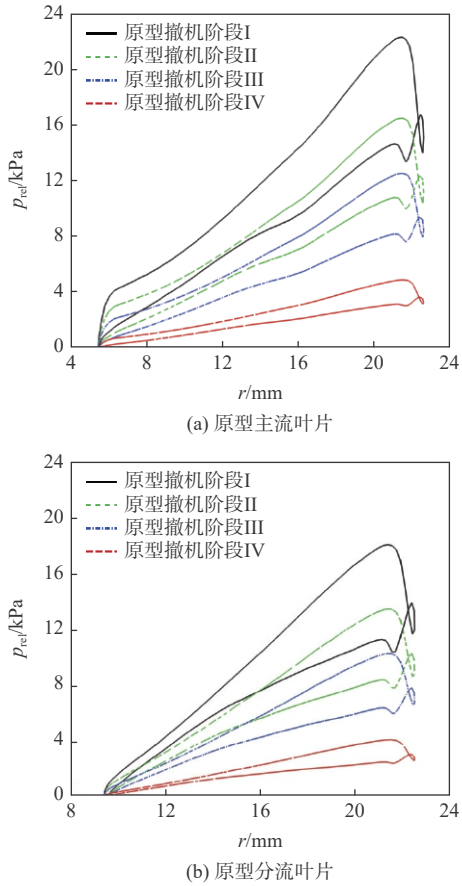


图6 原型血泵叶表相对压力分布

Fig. 6 Distribution of relative pressure on blade surface of datum pump

大。此外，分流叶片的偏置设计使得通道 I 的周向距离较小，分流叶片前缘溢流使得通道 I 的主流强度下降，抗扰动的能力较弱。无论是原型还是改型，通道涡仅存在于血液速度矢量由轴向快速转为径向的区域，主要集中在叶片前半部分，后半部分的通道涡强度已非常微弱。

原型和改型在不同撤机阶段的叶表相对压力分布如图 6 和图 7 所示。叶表相对压力计算式为

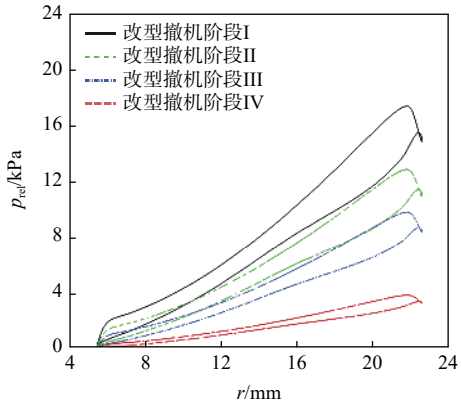
$$p_{rel} = p - p_{min} \tag{9}$$

式中：p为叶表静压；p_{min}为叶表静压最小值。在同一工况下，叶轮同一类型的叶片叶表相对压力分布相似，因此，认为任意一个叶片的相对压力分布能够代表这一类叶片。

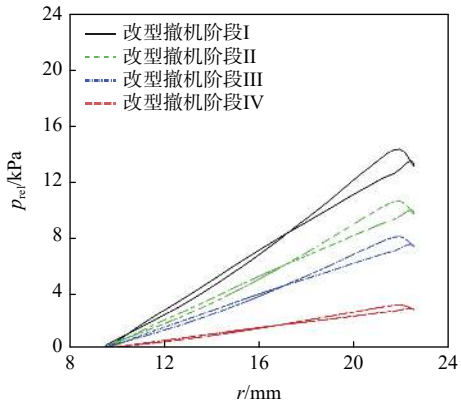
现有研究普遍认为血液在血泵内滞留时间过长、血泵中的强旋涡区、近壁面的滞止区及高剪切应力区都容易诱发血小板的激活，进而形成血栓。随着撤机的进行，原型和改型的主流叶片和分流叶片的叶表相对压力梯度和叶片负荷都明显减小，有利于降低间隙泄漏涡的强度，减少血泵的损失，提高血泵的血液相容性与效率。分离区是血栓形成的潜在区域，压力梯度和叶片负荷降低还有利于减少分离。在相同供压的条件下，改型叶片做功面积较大，单位高度上的负荷较低，表明改型比原型的抗溶血性能更优越。原型主流叶片与分流叶片吸力侧半径 22 mm 位置处存在压力突然下降的现象，这是由于原型叶片尾缘采取了大倒圆的设计，叶片压力侧和叶片吸力侧的流体在叶片尾缘汇合时，受到叶片通道面积突然扩张的影响，发生流动分离。分离区覆盖了尾缘点（最大半径处），导致原型血泵尾缘处叶表相对压力均迅速下降。改型叶片转速较低、负荷较小，尾缘分离区尺度也随之下降，压力侧分离起始点也更靠近尾缘。在半径 10.4 ～17 mm 之间，改型分流叶片压力侧的压力小于吸力侧，这是由于通道 I 内存在依附于分流叶片压力侧的通道涡。

2.2 血液相容性分析

叶轮旋转周期 T 内不同撤机阶段的瞬时溶血值如图 8 所示。随着撤机的进行，原型和改型的溶血值均大幅下降，表明撤机过程中，血泵的溶血情况会显著减弱。原型在撤机阶段 Ⅰ 的溶血值最高，瞬时溶血值波动幅度也最大，约 6%，表明当叶轮与蜗壳相对位置变化时，血泵会造成不同程度的溶血。改型在撤机阶段Ⅳ的血液损伤最低，其波动幅度小于 2%，表明叶轮旋转至不同位置时造成的溶



(a) 改型主流叶片



(b) 改型分流叶片

图 7　改型血泵叶表相对压力分布

Fig. 7　Distribution of relative pressure on blade surface of modified pump

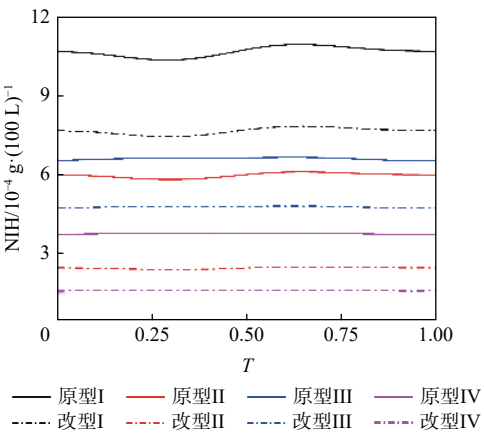


图 8　不同撤机阶段的瞬时溶血值

Fig. 8　Transient NIH of different weaning stages

血值变化已经很小。图 9 为原型和改型溶血平均值在不同撤机阶段的结果, 相较于原型, 改型在撤机阶段 Ⅰ 的溶血值降低了 38%, 在撤机阶段 Ⅱ 的溶血值降低了 37%, 在撤机阶段 Ⅲ 的溶血值降低了 36%, 在撤机阶段 Ⅳ 的溶血值降低了 34%, 溶血值减小均在 30% 以上, 说明改型的抗溶血性能显著提升。

虽然幂律模型已被广泛用于数值预测溶血, 但仍然存在一些问题可能导致预测误差较大。例如,

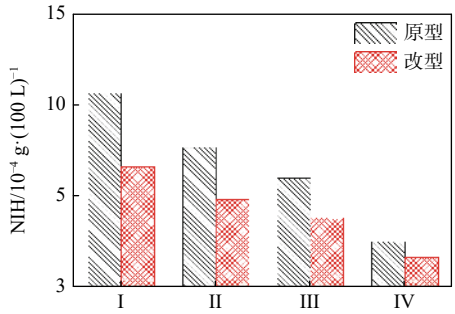


图 9　不同撤机阶段的溶血平均值

Fig. 9　Mean value of NIH in different blood pump weaning stages

幂律模型没有考虑红细胞膜的力学性能, 即细胞膜承担载荷过程会导致其力学性能下降。Marsden 等[23] 认为可以通过定义线性损伤权重[24] 和引入虚拟时间步[25] 解决 Giersiepen 溶血模型[19], 用拉格朗日积分时, 没有考虑溶血值在时间上的非线性的问题。同时, 有必要将损伤函数引入标量损伤累积模型中, 以模拟红细胞膜反复受到压力变化导致的老化问题[26]。云忠等[27-28] 采用血液流变学及应力波理论, 应用多相悬浮体 CFD 仿真技术对血液损伤机理进行研究, 给出了红细胞压差破碎的判定依据: 当红细胞内压力大于细胞外压力时, 红细胞破碎的临界压差为 0.012 Pa; 当红细胞内压力小于细胞外压力时, 红细胞破碎的临界压差为−0.002 Pa。云忠等[27-28] 研究突出了压力对血液损伤的重要性。撤机压力脉动监测点分布如图 10 所示, 8 个监测点分为 2 类, 第 1 类监测点 P_1～P_4 是位于蜗壳中心的 4 个象限点, 第 2 类监测点 P_5～P_8 是位于叶片肩部上侧, 靠近上泵壳的 4 个象限点。

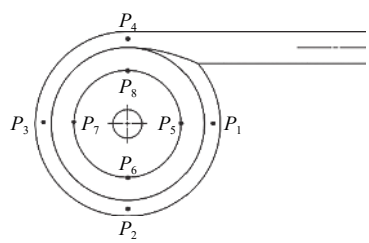
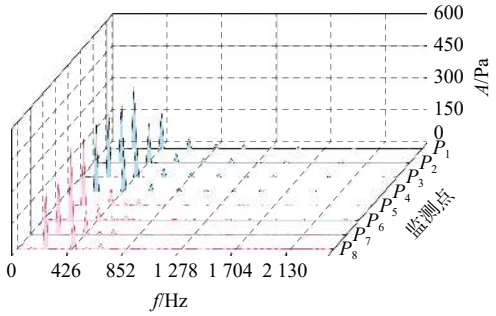


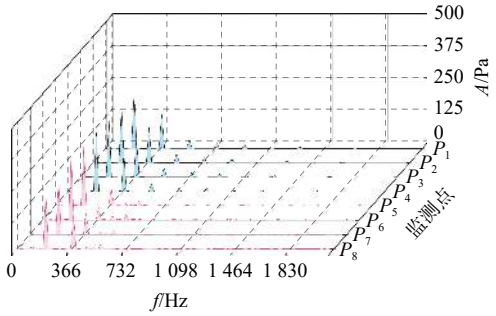
图 10　监测点分布

Fig. 10　Distribution of monitoring points

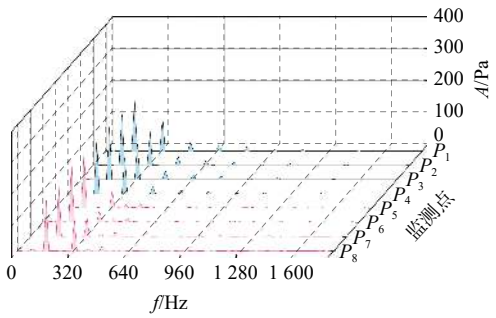
原型和改型泵内监测点压力脉动的频域结果分别如图 11 和图 12 所示。蜗壳内监测点的压力脉动与频率 f 的关系可以体现叶轮与蜗壳的转静干涉作用, 泵壳监测点的压力脉动主要体现叶轮与上泵壳的转静干涉作用。随着撤机的进行, 原型和改型的转静干涉作用均明显减弱, 血泵运转的稳定性增强, 有利于减少血泵内血液损伤, 保证患者的生命安全。改型上泵壳监测点的脉动幅值 A 大于原



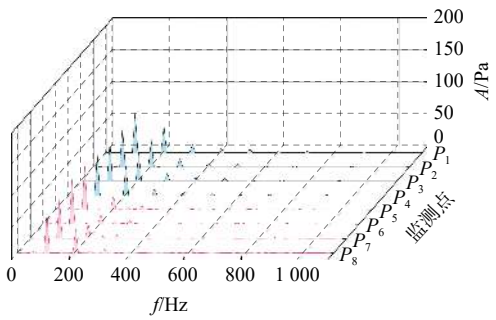
(a) 原型撤机阶段I



(b) 原型撤机阶段II



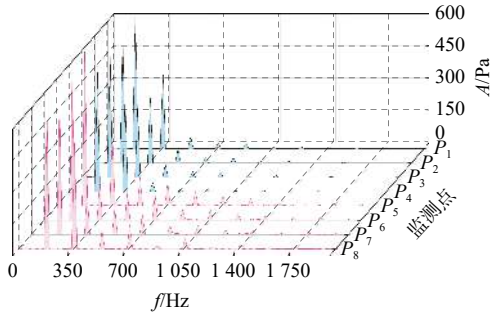
(c) 原型撤机阶段III



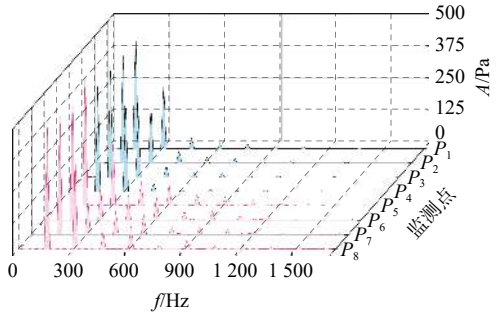
(d) 原型撤机阶段IV

图 11 原型泵内监测点压力脉动频域结果

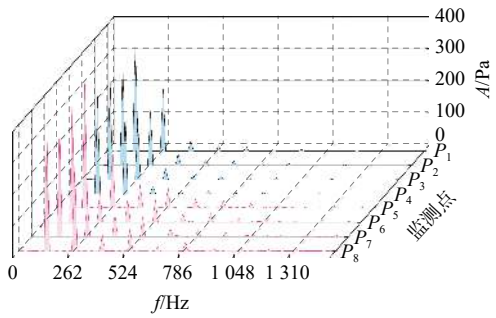
Fig. 11 Results of pressure fluctuation in frequency domain at monitoring points within the datum pump



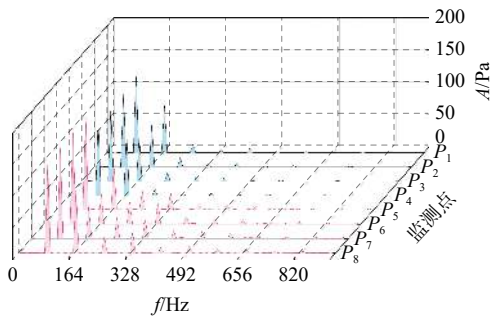
(a) 改型撤机阶段I



(b) 改型撤机阶段II



(c) 改型撤机阶段III



(d) 改型撤机阶段IV

图 12 改型泵内监测点压力脉动频域结果

Fig. 12 Results of pressure fluctuation in frequency domain at monitoring points within the modified pump

型，是由于改型的转速虽然低于同阶段的原型，但是由于改型叶片较高，更靠近监测点，叶片旋转时造成的扰动可被监测点清晰捕捉到，而原型叶顶间隙较大，虽然转速较快，但叶片旋转造成扰动会经过一定的衰减才会到达监测点，说明改型叶轮与上泵壳的转静干涉作用要强于原型。改型叶轮尾缘与原型叶轮尾缘距离蜗壳监测点的距离相等，但改型的脉动幅值较大的原因是蜗壳监测点处于蜗壳中心高度，原型叶片尾缘的高度低于监测点，导致监测点不能清晰捕捉其扰动信息。各监测点脉动主频均为叶频及其倍频，叶频时脉动幅值最大，2倍频也有明显的脉动。

原型所有监测点周向的压力脉动大致相等，说明当叶片高度较小且转速较快时，叶轮与出口管的转静干涉作用并不明显。改型流场中的脉动强度沿周向差异较大，其中蜗舌下游监测点P1和上泵壳监测点P5的压力脉动强度最大，沿顺时针方向脉动强度逐渐减弱，表明叶轮与出口管特别是蜗舌

的转静干涉作用最为明显。改型血泵内较强的转静干涉作用易诱发红细胞膜疲劳, 造成其力学性能下降, 增加破裂的风险。

3　结　论

1) 泄漏涡是血泵内最关键的流动特征, 会严重降低血泵的血液相容性, 随着撤机的进行, 血泵流量与转速逐步减小, 泄漏涡的强度也逐渐降低, 缓解了血泵内的溶血情况。改型通过减小叶顶间隙增加了叶片做功面积, 进而降低了达到标准供压所需的转速, 减小了泄漏涡的强度和尺度, 表明改型有较好的水力性能。

2) 随着撤机的进行, 2 种模型的溶血值均呈线性下降趋势, 其脉动幅度也大大降低, 表明在撤机过程中, 血泵的血液相容性显著提高。相较于原型, 改型流场中的速度分布更加均匀, 有效降低了血泵内的剪切应力, 提升了血泵的抗溶血性能和抗血栓性能。在撤机的 4 个阶段中, 改型的溶血值比相同撤机阶段的原型普遍减少 30% 以上。

3) 原型和改型内的转静干涉作用随着撤机的进行均明显减弱, 表明血泵运转的稳定性增强, 提升了血泵的安全性与可靠性。血泵各监测点脉动主频均为叶频及其倍频, 叶频时脉动幅值最大, 二倍频也有明显的脉动, 相较于原型, 处于同一撤机阶段的改型流场内转静干涉作用更加强烈, 对血泵的电磁控制技术提出了更高的要求。

参考文献（References）

[1] 《中国心血管健康与疾病报告2020》编写组.《中国心血管健康与疾病报告2020》概述[J]. 中国心血管病研究, 2021, 19(7): 582-590.
The Writing Committee of the Report on Cardiovascular Health and Diseases in China 2020. Key points of report on cardiovascular health and diseases in China 2020[J]. Chinese Journal of Cardiovascular Research, 2021, 19(7): 582-590(in Chinese).

[2] 高洁, 李建朝, 程兆云, 等. 体外膜氧合辅助9例危重型新型冠状病毒肺炎患者疗效分析[J]. 中国体外循环杂志, 2020, 18(4): 203-207.
GAO J, LI J C, CHENG Z Y, et al. Clinical analysis of COVID-19 patients supported by extracorporeal membrane oxygenation treatment[J]. Chinese Journal of Extracorporeal Circulation, 2020, 18(4): 203-207(in Chinese).

[3] 王军红, 马青变. 体外心肺复苏撤机应注意的问题[J]. 中国急救医学, 2021, 41(7): 610-612.
WANG J H, MA Q B. Problems needing attention in cardiopulmonary resuscitation[J]. Chinese Journal of Critical Care Medicine, 2021, 41(7): 610-612(in Chinese).

[4] 马纪梅. 人工心脏的经皮传能系统的研究[D]. 天津: 河北工业大学, 2011.
MA J M. Transcutaneous energy transmission system for artificial heart[D]. Tianjin: Hebei University of Technology, 2011 (in Chinese).

[5] 李永乐. 离心式心脏泵流场数值计算与分析[D]. 兰州: 兰州理工大学, 2011.
LI Y L. Numerical calculation and analysis about the flow field of centrifugal cardio pump[D]. Lanzhou: Lanzhou University of Technology, 2011 (in Chinese).

[6] LIU Y W, YU X J, LIU B J. Turbulence models assessment for large-scale tip vortices in an axial compressor rotor[J]. Journal of Propulsion and Power, 2008, 24: 15-25.

[7] LIU Y W, ZHONG L Y, LU L P. Comparison of DDES and URANS for unsteady tip leakage flow in an axial compressor rotor[J]. Journal of Fluids Engineering, 2019, 141(12): 121405.

[8] 展昭, 张岩, 桂幸民. 某轴流式血液泵的实验与改进设计[J]. 航空动力学报, 2008, 23(6): 1061-1066.
ZHAN Z, ZHANG Y, GUI X M. Experiment and design improvement of an axial flow blood pump[J]. Journal of Aerospace Power, 2008, 23(6): 1061-1066(in Chinese).

[9] 杨晓琛, 张岩, 桂幸民. 叶轮式辅助心脏泵装置优化设计[J]. 航空动力学报, 2011, 26(6): 1370-1376.
YANG X C, ZHANG Y, GUI X M. Optimization and design of an impeller assist heart pump device[J]. Journal of Aerospace Power, 2011, 26(6): 1370-1376(in Chinese).

[10] 张岩, 薛嵩, 桂幸民, 等. 运用三维数值模拟对人工心脏轴流血泵的设计和改进[J]. 中国生物医学工程学报, 2007, 26(1): 35-41.
ZHANG Y, XUE S, GUI X M, et al. Digital simulation to the development of axial blood pump for artificial heart[J]. Chinese Journal of Biomedical Engineering, 2007, 26(1): 35-41(in Chinese).

[11] 谢楠, 唐雨萌, 柳阳威, 等. 叶顶间隙对人工心脏泵血液相容性影响的数值研究[J]. 航空动力学报, 2021, 36(6): 1304-1314.
XIE N, TANG Y M, LIU Y W, et al. Numerical research on effects of blade tip clearances on the hemocompatibility of artificial heart pump[J]. Journal of Aerospace Power, 2021, 36(6): 1304-1314(in Chinese).

[12] ZHANG J T, GELLMAN B, KOERT A, et al. Computational and experimental evaluation of the fluid dynamics and hemocompatibility of the CentriMag blood pump[J]. Artificial Organs, 2006, 30(3): 168-177.

[13] FUNG Y C. Biomechanics: Mechanical properties of living tissues[M]. New York: Springer-Verlag, 1981.

[14] MULHOLLAND J W, SHELTON J C, LUO X Y. Blood flow and damage by the roller pumps during cardiopulmonary bypass[J]. Journal of Fluids and Structures, 2005, 20(1): 129-140.

[15] HUANG C R, FABISIAK W. A rheological equation characterizing both the time dependent and the steady state viscosity of whole human blood[J]. AIChE Symp Series, 1978, 74(1): 19-21.

[16] BLUDSZUWEIT C. Model for a general mechanical blood damage prediction[J]. Artificial Organs, 1995, 19(7): 583-589.

[17] ZHANG J F, ZHANG P, FRASER K H, et al. Comparison and experimental validation of fluid dynamic numerical models for a clinical ventricular assist device[J]. Artificial Organs, 2013, 37(4): 380-389.

[18] BLACKSHEAR P L, DORMAN F D, STEINBACH J H. Some mechanical effects that influence hemolysis[J]. Transactions American Society for Artificial Internal Organs, 1965, 11: 112-117.

[19] GIERSIEPEN M, WURZINGER L J, OPITZ R, et al. Estimation

of shear stress-related blood damage in heart valve prostheses: In vitro comparison of 25 aortic valves[J]. The International Journal of Artificial Organs, 1990, 13(5): 300-306.

[20] GARON A, FARINAS M I. Fast three-dimensional numerical hemolysis approximation[J]. Artificial Organs, 2004, 28(11): 1016-1025.

[21] BLUDSZUWEIT C. Three-dimensional numerical prediction of stress loading of blood particles in a centrifugal pump[J]. Artificial Organs, 1995, 19(7): 590-596.

[22] 高雁飞. 压气机转子叶尖泄漏湍流机理与数值模拟方法研究[D]. 北京: 北京航空航天大学, 2020.
GAO Y F. Study of turbulence mechanism and numerical methods in compressor tip leakage flow[D]. Beijing: Beihang University, 2020 (in Chinese).

[23] MARSDEN A L, BAZILEVS Y, LONG C C, et al. Recent advances in computational methodology for simulation of mechanical circulatory assist devices[J]. Wiley Interdisciplinary Reviews Systems Biology and Medicine, 2014, 6(2): 169-188.

[24] FARINAS M I, GARON A, LACASSE D, et al. Asymptotically consistent a numerical approximation of hemolysis[J]. Journal of Biomechanical Engineering, 2006, 128(5): 688-696.

[25] GOUBERGRITS L, AFFELD K. Numerical estimation of blood damage in artificial organs[J]. Artificial Organs, 2004, 28(5): 496-517.

[26] YELESWARAPU K K, ANTAKI J F, KAMENEVA M V, et al. A mathematical model for shear-induced hemolysis[J]. Artificial Organs, 1995, 19(7): 576-582.

[27] 云忠, 谭建平. 基于血液剪切损伤机理的高速螺旋血泵仿真分析[J]. 中山大学学报 (自然科学版), 2008, 47(1): 47-50.
YUN Z, TAN J P. Simulation analysis of the high-speed spiral blood pump based on the shear injure principle of blood[J]. Acta Scientiarum Naturalium Universitatis Sunyatseni, 2008, 47(1): 47-50(in Chinese).

[28] 云忠, 石芬, 向闯, 等. 混流血泵血液压差损伤机理分析及仿真[J]. 机械设计与研究, 2010, 26(3): 29-32.
YUN Z, SHI F, XIANG C, et al. Study on the injury principle of the blood prossure differevce in the mixed blood pump and its simulation analysis[J]. Machine Design and Research, 2010, 26(3): 29-32(in Chinese).

Numerical study of blood pump weaning effects on hemocompatibility of centrifugal blood pump

XIE Nan[1, 2], TANG Yumeng[1, 2], ZHANG Yan[3], LIU Yangwei[1, 2, *]

(1. School of Energy and Power Engineering, Beihang University, Beijing 100191, China;
2. National Key Laboratory of Science and Technology on Aero-Engine Aero-Thermodynamics, Beihang University, Beijing 100191, China;
3. Fuwai Hospital, Chinese Academy of Medical Sciences, Beijing 100037, China)

Abstract: The four stages of blood pump weaning are used to methodically assess the hemodynamic performance and hemocompatibility of the datum and modified centrifugal blood pumps. The distribution of parameters and the changes of normalized index of hemolysis in four stages of weaning are researched, the pressure pulsation of each monitoring point is also analyzed. Results show that both the datum pump and the modified pump own good hemocompatibility, and it satisfies the requirement of anti-hemolysis and anti-thrombosis in the first blood pump weaning stage. The leakage vortex formed in clearance is the main flow characteristic that induces hemolysis. The intensity of leakage vortex becomes weaken, and the index of hemolysis as well as its pulsation decrease significantly during the blood pump weaning process. Compared with the datum pump, the NIH of the modified pump is reduced by more than 30%. The pressure spectrum exhibits typical discrete characteristics, and the modified pump's rotor-stator interaction is stronger than that of the datum pump, indicating a high need for the blood pump's electromagnetic control technology.

Keywords: centrifugal blood pump; blood pump weaning; hemocompatibility; hemolysis; numerical simulation

Received: 2021-08-27; **Accepted**: 2022-01-02; **Published Online**: 2022-01-29 08:30
URL: kns.cnki.net/kcms/detail/11.2625.V.20220128.1659.004.html
Foundation items: National Natural Science Foundation of China (51976006,52106039); Aeronautical Science Foundation of China (2018ZB51013)
* **Corresponding author**. E-mail: liuyangwei@126.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0510

适用于弹性飞机飞行动力学仿真的气动力降阶方法

师妍[1]，万志强[1]，吴志刚[1,*]，巩轶男[2]

(1. 北京航空航天大学 航空科学与工程学院，北京 100191；2. 海鹰航空通用装备有限责任公司，北京 100070)

摘 要：针对低速太阳能无人机在由阵风引起的大迎角非线性气动力情况下的高精度、高效率气动弹性及飞行动力学工程仿真问题，对气动力降阶方法及其在飞行动力学仿真中的应用进行深入探索。以二维翼型为例，分别基于自回归外部输入模型 (ARX) 方法、多小波 Volterra 级数方法、反向传播 (BP)、径向基函数 (RBF) 神经网络方法和支持向量回归 (SVR) 方法建立了非定常气动力降阶模型，并比较了 5 种方法的收敛性及泛化能力。在此基础之上综合 ARX 和神经网络方法的优点提出一种计算精度高，收敛性、泛化能力较强的 ARX-BP/RBF 组合建模方法。将基于 ARX-RBF 的组合气动力降阶方法和弹性飞机刚弹耦合动力学方程相结合建立基于气动力降阶的弹性飞机飞行动力学仿真模型。以一架太阳能无人机模型为例，开展数值验证，对算例对象在阵风及舵面激励下的响应进行仿真研究，并和计算流体力学 (CFD)-计算结构动力学 (CSD) 结果及基于线性气动力的仿真结果进行对比。结果表明：所建立的弹性飞机仿真模型能有效反映气动力非线性特性，且仿真效率高于 CFD-CSD 分析方法，适用于工程实践。

关 键 词：气动力降阶；飞行仿真；刚弹耦合动力学方程；太阳能无人机；阵风响应

中图分类号：V212.1; V211; V215.3

文献标志码：A **文章编号**：1001-5965(2023)07-1689-18

近年来，太阳能无人机广泛应用于军用、民用各个领域[1]。为了满足其长航程、长航时设计指标，其气动、结构等设计要求比常规飞机更为苛刻，同时还需具有极低的能耗和极高的效率。这使太阳能无人机普遍具有大展弦比、结构质量轻、机翼柔性大、翼载低等特点。这些特点使太阳能无人机对阵风更加敏感，在阵风扰动下更容易出现大迎角情况，气动力呈现非线性特性。

为了研究阵风引起的非线性气动力对大展弦比无人机飞机结构及飞行稳定性的影响，需要建立更为精确的气动力模型来进行飞行仿真分析。传统工程应用的气动力面元分析方法基于线性气动力假设，已经不能满足仿真要求。而高精度的计算流体力学 (computational fluid dynamics, CFD)-计算结构动力学 (computational structural dynamics, CSD) 全阶有限元分析方法计算原理复杂，从网格建模到实际计算都需花费大量的时间，且仿真计算效率较低。因此，可利用气动力降阶建模方法建立气动力降阶模型来实现计算精度和计算效率方面的平衡。

目前应用于流固耦合分析系统中的非定常气动力降阶模型大致分为 2 类：①基于流场特征模态投影[2]；②基于系统辨识和数据驱动。前者建模需要全流场信息，而后者则只需要少量的输入、输出样本，目的是以简洁的数学表达式描述全阶非定常流动系统。因此，后者更适用于气动弹性和飞行仿真分析等动态响应问题。当前基于系统辨识和数

收稿日期：2021-09-02；录用日期：2021-11-22；网络出版时间：2021-12-30 07:02
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20211229.1524.003.html
*通信作者. E-mail：wuzhigang@buaa.edu.cn

引用格式：师妍，万志强，吴志刚，等. 适用于弹性飞机飞行动力学仿真的气动力降阶方法 [J]. 北京航空航天大学学报，2023，49（7）：1689-1706. SHI Y，WAN Z Q，WU Z G，et al. Aerodynamic order reduction method for elastic aircraft flight dynamics simulation [J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（7）：1689-1706（in Chinese）.

据驱动的气动力降阶模型可以大致分为4类，分别为基于自回归外部输入模型 (autoregressive model with eXogenous inputs, ARX) 的模型[3-5]、基于Volterra级数的模型[6-11]、基于神经网络的模型[12-15]及基于支持向量机 (support vector machine, SVM) 的模型[16-18]。对于这几类模型的研究和应用都有一定的基础。但是这些模型大都应用于非线性气动弹性分析，很少应用于飞行动力学仿真建模。一方面飞行动力学仿真需建立更多自由度下的气动力模型，建模难度更大，另一方面降阶模型建立大都基于有限给定来流状态下的气动力训练数据，超出训练数据范围后很难保证其适用性。且增加训练数据采集范围后还会增加训练数据获取成本，降低建模效率。因此，虽然已有很多种气动力降阶建模方法，但对这些降阶方法的适用性、收敛性、泛化能力等方面缺乏深入研究，很难直接将其应用于飞行仿真。对降阶模型特性研究的不足也制约了降阶模型的工程应用，相关研究亟待展开。

针对以上问题，本文首先基于NACA0012大迎角俯仰、沉浮运动下的CFD气动力数据建立基于ARX、基于多小波Volterra级数、基于反向传播 (back propagation, BP) 神经网络、基于径向基 (radial basis function, RBF) 神经网络和基于支持向量回归 (support vector regression, SVR) 的5种气动力降阶模型。对这5种气动力降阶模型随训练样本点增加的收敛性、来流速度变化时的泛化能力进行对比分析，并结合ARX和神经网络算法各自优点提出一种ARX-BP/RBF组合建模方法。组合模型和原神经网络的模型相比收敛性更好、泛化能力更强，和原ARX模型相比计算精度更高。最后以一架太阳能无人机模型为对象，结合气动力降阶模型和刚弹耦合飞行动力学方程对模型遭遇阵风及给定舵面偏角情况下的响应进行飞行仿真分析，并将仿真结果和CFD-CSD计算结果与基于线性气动力的飞行仿真模型仿真结果进行对比。

1 弹性飞机飞行动力学建模

1.1 刚弹耦合动力学方程建模

本文在平均体轴系下建立弹性飞机的飞行动力学方程。定义平均体轴系为 $Oxyz$，原点 O 位于飞机的瞬时质心，Ox 轴指向机头前方，Oz 轴和 Ox 正交并指向机体下方，Oy 轴按右手定则确定。根据牛顿第二定律，基于平均轴系的弹性飞机质心动力学方程为

$$\begin{cases} m(\dot{u}+qw-rv)=F_x \\ m(\dot{v}+ru-pw)=F_y \\ m(\dot{w}+pv-qu)=F_z \end{cases} \tag{1}$$

式中：m 为飞机质量；u,v,w 分别为平均体轴系下 x,y,z 这3个方向的速度值；p,q,r 分别为平均体轴系下绕 x,y,z 三轴转动的角速度值；F_x,F_y,F_z 分别为平均体轴系下机体所受 x,y,z 这3个方向的外力合力。

根据动量矩定理及飞机的面对称性，平均轴系的弹性飞机绕质心转动的动力学方程为

$$\begin{cases} \dot{p}I_{xx}-(\dot{r}+pq)I_{xz}+qr(I_{zz}-I_{yy})=M_x \\ \dot{q}I_{yy}-pr(I_{xx}-I_{zz})+(p^2-r^2)I_{xz}=M_y \\ \dot{r}I_{zz}-(qr-\dot{p})I_{xz}+pq(I_{yy}-I_{xx})=M_z \end{cases} \tag{2}$$

式中：$I_{xx},I_{yy},I_{zz},I_{xz}$ 分别为机体绕 x,y,z 轴的转动惯量和相对于 xz 平面的惯性积；M_x,M_y,M_z 分别为平均体轴系下机体所受 x,y,z 这3个方向的外力矩。

采用模态法对飞机弹性变形进行表征，可建立飞机的弹性振动动力学方程为

$$\boldsymbol{M}_e\ddot{\boldsymbol{\xi}}+\boldsymbol{C}_e\dot{\boldsymbol{\xi}}+\boldsymbol{K}_e\boldsymbol{\xi}=\boldsymbol{Q}_e \tag{3}$$

式中：$\boldsymbol{\xi}=[\xi_1,\xi_2,\cdots,\xi_n]^{\mathrm{T}}$ 为弹性模态广义坐标；$\boldsymbol{M}_e$ 为广义质量阵；$\boldsymbol{C}_e$ 为广义阻尼阵；$\boldsymbol{K}_e$ 为广义刚度阵；$\boldsymbol{Q}_e$ 为广义气动力。式(1)～式(3)组合即为弹性飞机刚弹耦合飞行动力学方程组。

1.2 运动学方程建模

本文定义大地参考系为 $OXYZ$，原点 O 为地面上的任意一点。OZ 轴沿重力加速度方向竖直向下，OX 轴方向为飞机航向在地平面上的投影方向，OY 轴通过右手定则确定。可通过飞机刚体偏航角、俯仰角、滚转角这3个欧拉角来建立平均体轴系下运动参数到大地参考系下运动参数的变换矩阵。设飞机刚体运动的偏航角为 ψ(右偏为正)，俯仰角为 θ(抬头为正)，滚转角为 ϕ(右滚为正)，飞机瞬时质心速度在 OX, OY, OZ 轴的投影分别为 $\dot{X},\dot{Y},\dot{Z}$。通过 ψ,θ,ϕ 可建立平均体轴系下运动参数到大地参考系下运动参数的关系矩阵，如式(4)，式(5)所示：

$$\begin{bmatrix}\dot{X}\\ \dot{Y}\\ \dot{Z}\end{bmatrix}=\begin{bmatrix}\cos\theta\cos\psi & \sin\phi\sin\theta\cos\psi-\cos\phi\sin\psi & \cos\phi\sin\theta\cos\psi+\sin\phi\sin\psi\\ \cos\theta\sin\psi & \sin\phi\sin\theta\sin\psi+\cos\phi\cos\psi & \cos\phi\sin\theta\sin\psi-\sin\phi\cos\psi\\ -\sin\theta & \sin\phi\cos\theta & \cos\phi\cos\theta\end{bmatrix}\cdot\begin{bmatrix}u\\ v\\ w\end{bmatrix} \tag{4}$$

$$\begin{bmatrix}\dot{\phi}\\ \dot{\theta}\\ \dot{\psi}\end{bmatrix}=\begin{bmatrix}1 & \sin\phi\tan\theta & \cos\phi\tan\theta\\ 0 & \cos\phi & -\sin\phi\\ 0 & \sin\phi\sec\theta & \cos\phi\sec\theta\end{bmatrix}\begin{bmatrix}p\\ q\\ r\end{bmatrix} \tag{5}$$

1.3　受外力项建模

将式(1)～式(3)右侧项组合, 即可得到飞机运动时受到的外力。忽略由舵面偏转产生的惯性力、推力引起的弹性模态广义力及外力距, 体轴系下的外力项为

$$\begin{bmatrix}F_x\\ F_y\\ F_z\\ M_x\\ M_y\\ M_z\\ \boldsymbol{Q}_{\mathrm{e}}\end{bmatrix}=\begin{bmatrix}-mg\sin\theta\\ mg\cos\theta\sin\phi\\ mg\cos\phi\cos\theta\\ 0\\ 0\\ 0\\ \boldsymbol{0}\end{bmatrix}+\begin{bmatrix}T_{F_x}\\ 0\\ 0\\ 0\\ 0\\ 0\\ \boldsymbol{0}\end{bmatrix}+\begin{bmatrix}f_u\\ f_v\\ f_w\\ f_p\\ f_q\\ f_r\\ \boldsymbol{Q}_{\mathrm{e}}\end{bmatrix} \tag{6}$$

式中: 右侧第 1 项表示重力; 第 2 项表示推力; 第 3 项表示气动力。本文定义飞机重力项恒定, 则外力项建模核心在于对飞机所受气动力的建模。而飞机所受的气动力不仅与飞机本身的运动及姿态有关, 还受到舵面偏转的影响, 以及外界阵风的影响, 因此, 建模参数不仅包括飞机的运动参数还包括舵面偏转参数如舵偏角$\boldsymbol{\delta}=[\delta_1,\delta_2,\cdots,\delta_n]$(下标为各舵面编号), 阵风参数如阵风速度$\boldsymbol{w}_{\mathrm{g}}=[w_{\mathrm{g}x},w_{\mathrm{g}y},w_{\mathrm{g}z}]$。

设选取前 n 阶弹性模态进行飞行动力学建模, 则$\boldsymbol{Q}_{\mathrm{e}}=[Q_1,Q_2,\cdots,Q_n]^{\mathrm{T}}$

为方便表示, 令:

$$\begin{aligned}&\boldsymbol{U}=[u,v,w,p,q,r,\xi_1,\xi_2\cdots\xi_n,\boldsymbol{\delta},\boldsymbol{w}_{\mathrm{g}}]^{\mathrm{T}}\\ &\boldsymbol{Y}=[f_u,f_v,f_w,f_p,f_q,f_r,Q_1,Q_2,\cdots,Q_n]^{\mathrm{T}}\end{aligned} \tag{7}$$

根据气动力产生部位, 飞机的外部气动力$\boldsymbol{Y}$还可写为

$$\boldsymbol{Y}=\boldsymbol{Y}_{\mathrm{w}}+\boldsymbol{Y}_{\mathrm{ht}}+\boldsymbol{Y}_{\mathrm{vt}}+\boldsymbol{Y}_{\mathrm{b}}+\boldsymbol{Y}_{\mathrm{c}} \tag{8}$$

式中: $\boldsymbol{Y}_{\mathrm{w}}$为机翼产生的气动力; $\boldsymbol{Y}_{\mathrm{ht}}$和$\boldsymbol{Y}_{\mathrm{vt}}$分别为平尾和垂尾产生的气动力; $\boldsymbol{Y}_{\mathrm{b}}$为机身产生的气动力; $\boldsymbol{Y}_{\mathrm{c}}$为舵面产生的气动力。则气动力外力项建模即建立$\boldsymbol{U}$和$\boldsymbol{Y}$的各分量$\boldsymbol{Y}_{\mathrm{w}}$, $\boldsymbol{Y}_{\mathrm{ht}}$, $\boldsymbol{Y}_{\mathrm{vt}}$, $\boldsymbol{Y}_{\mathrm{b}}$, $\boldsymbol{Y}_{\mathrm{c}}$之间的关系, 将在第 2 节详细介绍。综合本节内容可得出弹性飞机飞行动力学仿真框架, 如图 1 所示。

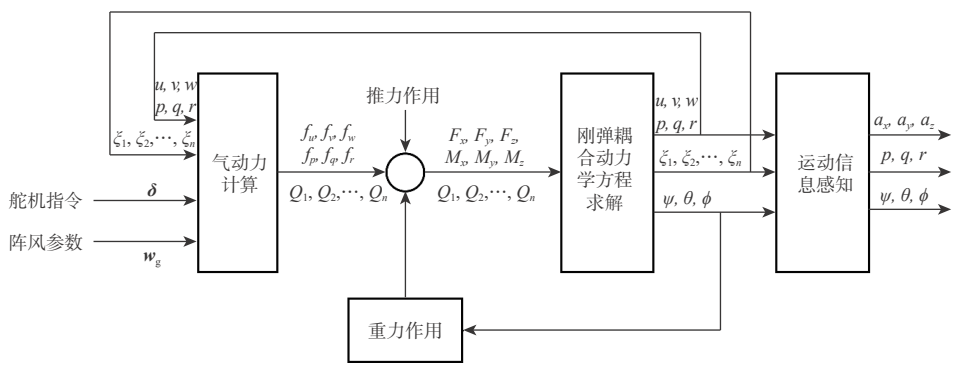


图 1　弹性飞机飞行动力学仿真框架

Fig. 1　Flight dynamics simulation framework

2　适用于飞行仿真的气动力降阶模型

本节将用基于数据驱动的气动力降阶建模方法来建立弹性飞机飞行仿真的气动力计算模块。模型的输入$\boldsymbol{u}$为给定的系统状态量变化信号, 输出$\boldsymbol{y}$为气动力或气动力系数。

2.1　气动力数据计算

本文利用 CFD 求解器对给定运动信号下的气动力进行计算和采集, 采用文献 [5] 中的气动力计算方法可获得飞机刚体运动及弹性变形产生的广义气动力, 通过设置不同的飞机气动外形边界组可实现飞机不同部件产生的气动力数据拆解。需要注意的是, 由 CFD 直接获取的气动力数据建立在风轴系下, 定义风轴系原点 O_{a} 在飞机质心, $O_{\mathrm{a}}x_{\mathrm{a}}$ 轴指向飞机速度方向; $O_{\mathrm{a}}z_{\mathrm{a}}$ 轴处于飞机对称面内, 垂直于 $O_{\mathrm{a}}x_{\mathrm{a}}$ 向下为正; $O_{\mathrm{a}}y_{\mathrm{a}}$ 轴通过右手定则建立。设风轴系下的气动力为

$$\boldsymbol{Y}_{\mathrm{a}}=\boldsymbol{Y}_{\mathrm{aw}}+\boldsymbol{Y}_{\mathrm{aht}}+\boldsymbol{Y}_{\mathrm{avt}}+\boldsymbol{Y}_{\mathrm{ab}}+\boldsymbol{Y}_{\mathrm{ac}} \tag{9}$$

2.2　状态量坐标转换

由于气动力降阶模型输入的系统状态量通常为风轴系下的状态量(如来流速度 V、迎角α、垂直于 V 向速度$\dot{h}$等), 因此, 需要先将体轴系下的状态参数转化为风轴系下的组合模型输入参数, 再进行气动力计算。

对于风轴系下的来流速度 V, 在考虑阵风干扰的情况下, 按照不同气动面可以分为 2 类, 分别为机翼与平尾的来流速度 V_1 和垂尾的来流速度 V_2:

$$\begin{cases}V_1=\sqrt{u^2+w^2}+\boldsymbol{w}_{\mathrm{b}}\cdot[-\cos\alpha_{\mathrm{r1}},0,-\sin\alpha_{\mathrm{r1}}]^{\mathrm{T}}\\ V_2=\sqrt{u^2+v^2}+\boldsymbol{w}_{\mathrm{b}}\cdot[-\cos\alpha_{\mathrm{r2}},-\sin\alpha_{\mathrm{r2}},0]^{\mathrm{T}}\end{cases} \tag{10}$$

式中：$\alpha_{r1}=\arctan(w/u)$，$\alpha_{r2}=\arctan(v/u)$为刚体运动在风轴系下形成的迎角；$\boldsymbol{w}_b=[w_{bx},w_{by},w_{bz}]$为平均体轴系下的阵风速度的3个分量，其与$\boldsymbol{w}_g$的转换关系为

$$\boldsymbol{w}_b^{\mathrm{T}}=\begin{bmatrix}\cos\theta\cos\psi & \sin\phi\sin\theta\cos\psi-\cos\psi\sin\psi & \cos\phi\sin\theta\cos\psi+\sin\psi\phi\sin\psi\\ \cos\theta\sin\psi & \sin\phi\sin\theta\sin\psi+\cos\cos\psi & \cos\phi\sin\theta\sin\psi-\sin\phi\cos\psi\\ -\sin\theta & \sin\phi\cos\theta & \cos\phi\cos\theta\end{bmatrix}\boldsymbol{w}_g^{\mathrm{T}} \tag{11}$$

对于风轴系下的迎角α，同样根据不同气动面分为机翼与平尾的迎角α_1及垂尾的迎角α_2。本文既考虑飞机的刚性运动又考虑飞机弹性模态的影响。风轴系下α_1,α_2为

$$\begin{cases}\alpha_1=\alpha_{r1}+\alpha_{e1}(\xi_1,\xi_2,\cdots,\xi_n)\\ \alpha_2=\alpha_{r2}+\alpha_{e2}(\xi_1,\xi_2,\cdots,\xi_n)\end{cases} \tag{12}$$

式中：$\alpha_{e1}(\xi_1,\xi_2,\cdots,\xi_n)$、$\alpha_{e2}(\xi_1,\xi_2,\cdots,\xi_n)$为弹性变形在风轴系下形成的等效迎角。

对于风轴系下垂直于V向的运动速度$\dot{h}$，在考虑弹性振动、飞机旋转运动及阵风干扰情况下，对应的机翼与平尾的$\dot{h}_1$和垂尾的$\dot{h}_2$分别为

$$\begin{aligned}\dot{h}_1&=(\boldsymbol{\omega}\times\boldsymbol{s}+\dot{h}_{e1}(\xi_1,\xi_2,\cdots,\xi_n))\cdot[-\sin\alpha_{r1},0,\cos\alpha_{r1}]^{\mathrm{T}}-\\&\quad\boldsymbol{w}_b\cdot[-\sin\alpha_{r1},0,\cos\alpha_{r1}]^{\mathrm{T}}\\ \dot{h}_2&=(\boldsymbol{\omega}\times\boldsymbol{s}+\dot{h}_{e2}(\xi_1,\xi_2,\cdots,\xi_n))\cdot[-\sin\alpha_{r2},\cos\alpha_{r2},0]^{\mathrm{T}}-\\&\quad\boldsymbol{w}_b\cdot[-\sin\alpha_{r2},\cos\alpha_{r2},0]^{\mathrm{T}}\end{aligned} \tag{13}$$

式中：$\boldsymbol{s}$为机翼刚轴到质心的距离矢量；$\boldsymbol{\omega}=[p,q,r]^{\mathrm{T}}$为平均体轴系相对于惯性参考系的转动角速度；$\dot{h}_{e1}(\xi_1,\xi_2,\cdots,\xi_n)$、$\dot{h}_{e2}(\xi_1,\xi_2,\cdots,\xi_n)$为弹性振动带来的垂向速度分量。

对于舵面偏转产生的气动力，由于舵偏角通常较小，因此，本文定义舵偏角产生的气动力在线性范围内。设舵偏角为$\boldsymbol{\delta}$，则舵面产生的气动力可写作$\boldsymbol{Y}_{ac}=\boldsymbol{C}_l^{\delta}\boldsymbol{\delta}^{\mathrm{T}}$的形式，建模过程也是对$\boldsymbol{C}_l^{\delta}$的辨识过程。

通过式(10)～式(13)可获得风轴系下的状态参数为

$$\boldsymbol{u}=[\boldsymbol{u}_1,\boldsymbol{u}_2,\boldsymbol{\delta}]^{\mathrm{T}} \tag{14}$$

其中：

$$\begin{cases}\boldsymbol{u}_1=[V_1,\boldsymbol{\alpha}_1,\dot{h}_1]\\ \boldsymbol{u}_2=[V_2,\boldsymbol{\alpha}_2,\dot{h}_2]\end{cases} \tag{15}$$

$\boldsymbol{\alpha}_1=[\alpha_1,\dot{\alpha}_1,\ddot{\alpha}_1]$，$\boldsymbol{\alpha}_2=[\alpha_2,\dot{\alpha}_2,\ddot{\alpha}_2]$，式(9)可写为

$$\boldsymbol{Y}_a(\boldsymbol{u})=\boldsymbol{Y}_{aw}(\boldsymbol{u}_1)+\boldsymbol{Y}_{aht}(\boldsymbol{u}_1)+\boldsymbol{Y}_{avt}(\boldsymbol{u}_2)+\boldsymbol{Y}_{ab}(\boldsymbol{u}_1)+\boldsymbol{Y}_{ac}(\boldsymbol{\delta}) \tag{16}$$

为方便后续说明，令$\boldsymbol{y}=\boldsymbol{Y}_a$。

气动力计算模块可进一步细化，如图2所示。

通过2.1节和2.2节获取$\boldsymbol{y}$和$\boldsymbol{u}$的数据后，可利用气动力降阶建模方法建立气动力降阶模型。

2.3 适用于飞行仿真的气动力降阶模型

基于系统辨识和数据驱动的气动力降阶建模方法更适用于飞行仿真。目前已有的弹性飞机气动力降阶模型大致可以分为以下4类。

2.3.1 基于ARX的气动力降阶模型

对于多输入多输出(multiple-input multiple-output, MIMO)离散时间系统，ARX模型可写为

$$\boldsymbol{y}(k)=-\sum_{i=1}^{N}\boldsymbol{A}_i\boldsymbol{y}(k-i)+\sum_{j=0}^{M}\boldsymbol{B}_j\boldsymbol{u}(k-j) \tag{17}$$

式中：$\boldsymbol{u}$为系统输入；$\boldsymbol{y}$为系统输出；k为时间标度；$\boldsymbol{A}_i$、$\boldsymbol{B}_j$为待辨识参数；N、M为模型的阶数，反映系统的非定常效应。

由式(17)直接建立的模型为线性模型，有学者对式(17)进行了非线性扩展，获得Hammerstein模型[19]。本文将这类模型也归类为基于ARX的模型。

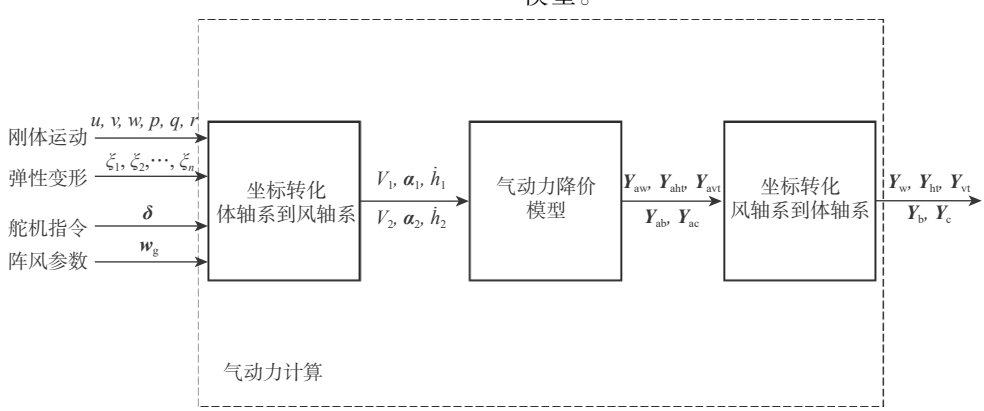


图2 气动力计算模块

Fig. 2 Aerodynamic calculation framework

2.3.2　基于 Volterra 级数的气动力降阶模型

该模型将输出表示为无穷积分求和形式:

$$y(t)=y_1(t)+y_2(t)+y_3(t)+\cdots+y_\infty(t) \tag{18}$$

式中: $y_i(t)$为第 i 阶 Volterra 级数输出。

对于具有因果关系的单输入单输出系统, 如果其因果关系性质不随时间变化而改变, 那么这个系统可写为

$$\begin{cases} y_1(t)=\int_0^t \mathrm{hh}_1(\xi)u(t-\xi)\mathrm{d}\xi \\ y_2(t)=\int_0^t\int_0^t \mathrm{hh}_2(\xi,\eta)u(t-\xi)u(t-\eta)\mathrm{d}\xi\mathrm{d}\eta \end{cases} \tag{19}$$

式中: 一阶核 hh_1 即可描述系统的线性特性, 高阶核 hh_2～ hh_n 能够描述系统的非线性。

基于 Volterra 级数的气动力降阶模型建模关键在于核函数的辨识, 本文采用基于多小波基函数展开的 Volterra 的核函数建模方法[11]。

2.3.3　基于神经网络的气动力降阶模型

神经网络本身具有强非线性、强逼近能力、强大的自学习及自适应能力, 目前已广泛应用于非线性系统辨识和建模。

1）基于 BP 神经网络的气动力降阶模型。BP 神经网络由输入层、隐藏层、输出层组成。其中输入和输出层各有一层神经元, 神经元个数分别等于输入和输出数据的维数。隐藏层神经元层数及个数需要根据实际情况, 通过人工设定。若只有一层隐藏层, 设输入层输出为 $\boldsymbol{u}=[u_1,u_2,\cdots,u_n]^{\mathrm{T}}$, 隐藏层输出$\boldsymbol{I}=[I_1,I_2,\cdots,I_{n_s}]^{\mathrm{T}}$, 则输出层输出$\boldsymbol{y}=[y_1,y_2,\cdots,y_m]^{\mathrm{T}}$ 可通过式 (20)～式 (21) 获得:

$$\begin{cases} I_j=f(\mathrm{net}_j) \\ \mathrm{net}_j=\sum_{i=0}^{n} v_{ij}u_i,\ j=1,2,\cdots,n_\mathrm{s} \end{cases} \tag{20}$$

$$\begin{cases} y_k=f(\mathrm{net}_k) \\ \mathrm{net}_k=\sum_{j=1}^{n_\mathrm{s}} w_{jk}I_j,k=1,2,\cdots,m \end{cases} \tag{21}$$

式中: v_{ij}和w_{jk}为待辨识参数。

2）基于 RBF 神经网络的气动力降阶模型。RBF 神经网络将径向基函数作为神经网络的隐含层。隐含层仅一层神经元。对于输入样本$\boldsymbol{u}_t$, RBF 网络输出y_{it}可写为

$$y_{it}=\sum_{p=0}^{P} W_{ip}\exp\left(\frac{\|\boldsymbol{u}_t-\boldsymbol{c}_p\|^2}{2\sigma^2}\right)\qquad i=1,2,\cdots,m \tag{22}$$

式中: W_{ip} 为待定系数; $\boldsymbol{c}_p$ 为中心向量; σ为宽度。

2.3.4　基于支持向量回归的气动力降阶模型

SVR 模型训练目标是找出一个超平面, 使所有数据到这个超平面的距离最小。即要找到 $\mathrm{Min}(1/2\|\boldsymbol{w}_{\mathrm{svr}}{}^2\|+C\sum_{i=1}^{n}(\xi_i^{\mathrm{svr}}+\xi_i^{\mathrm{svr}*})$, SVR 原理示意图如图 3 所示。其中, ξ_i^{svr},$\xi_i^{\mathrm{svr}*}$为训练数据点到超平面的距离, C 为可调参数。

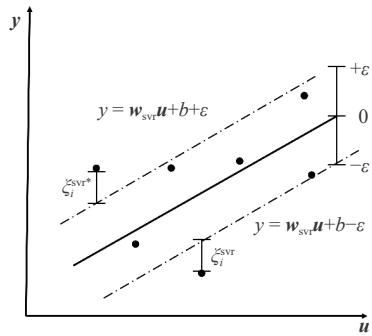


图 3　SVR 原理示意图

Fig. 3　Schematic diagram of SVR

综上, 目前已有多种气动力降阶建模方法, 但对这些气动力降阶建模方法之间的差异及适用性选择标准研究还未有定性结论。为了找到更适合用于飞行仿真的气动力降阶建模方法, 需要对各气动力降阶模型的特性进行对比研究。

3　各气动力降阶模型特性对比

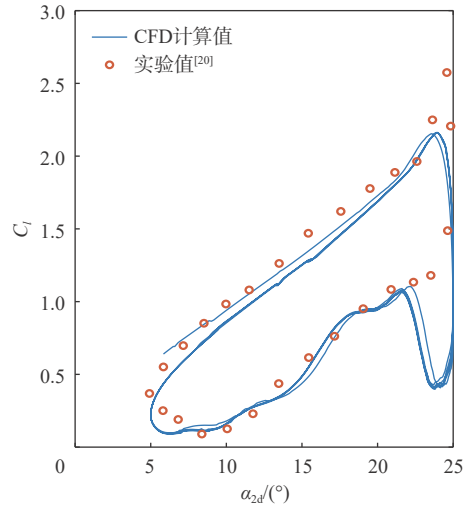
选取 ARX 方法、基于多小波的 Volterra 级数方法、BP、RBF 神经网络方法和 SVR 方法建立气动力降阶模型。其中 ARX 模型采用函数连接型神经网络对原始方法进行非线性拓展, 基函数选取幂级数。Volterra 级数模型采用二阶模型。BP 神经网络模型包含 2 个隐藏层, 每个隐藏层单元数为 5。RBF 神经网络模型激活函数选取高斯函数, 隐含层神经元的数目和输入数据数量相等。SVR 模型的核函数采用径向基函数。

3.1　算例对象选择

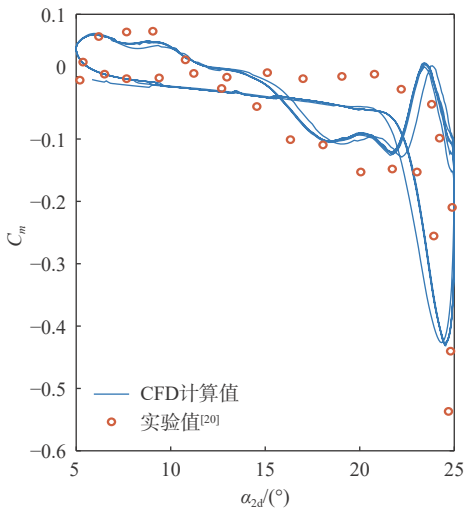
本文选取二维翼型 NACA0012 的俯仰、沉浮运动下的气动力系数作为气动力 CFD 数据采集对象。选取 fluent 压力基求解器、SST 两方程湍流模型进行 CFD 计算, 空间离散采用有限体积法, 时间离散采用隐式时间推进法。给定正弦俯仰运动下的 CFD 计算结果和文献 [20] 中的实验结果对比如图 4 所示, 所采用的验证算例各项运动参数如表 1 所示, 从表 1 可以看出翼型俯仰运动可以写作 $15+10\sin(\omega_{2\mathrm{d}}t)$的形式, 其中$\omega_{2\mathrm{d}}$为俯仰运动角频率, 可以通过表 1 数据计算得出。图 4 在一定程度验证了 CFD 计算的正确性。

3.2　模型训练、测试信号选取

对二维算例对象, 各降阶模型输出定义为



(a) 升力系数随α_{2d}变化



(b) 力矩系数随α_{2d}变化

图 4 CFD 计算结果和实验结果[20] 对比

Fig. 4 Comparison of CFD and experimental results[20]

表 1 CFD 计算验证算例翼型运动参数

Table 1 Airfoil motion parameters of CFD case

翼型弦长/m	来流速度/(m·s^{-1})	初始迎角/(°)	俯仰运动幅值/(°)	减缩频率
0.55	40.8	15	10	0.124

$\boldsymbol{y}_{2d}(k)=[C_l(k),C_m(k)]^{\mathrm{T}}$, k 为时间标度, $C_l(k),C_m(k)$为 k 时刻升力与俯仰力矩系数。训练输入$\boldsymbol{U}_{2d}(k)=[\boldsymbol{u}_{2d}(k),\boldsymbol{u}_{2d}(k-1),\cdots,\boldsymbol{u}_{2d}(k-m_{2d})]^{\mathrm{T}}$,$\boldsymbol{u}_{2d}(k)=[\alpha_{2d}(k),\dot{\alpha}_{2d}(k),\ddot{\alpha}_{2d}(k)]$, α_{2d}为有效迎角, 可通过$\alpha_{2d}=\theta_{2d}+\arctan(\dot{h}_{2d}/V)$计算获得, θ_{2d}为俯仰角, $\dot{h}_{2d}$为沉浮运动速度, V 为来流速度。m_{2d} 为模型阶数, 取值为 2。

虽然各种降阶建模方法最适用的训练数据信号形式各不相同, 但本文研究的是各降阶模型在有限信息数据输入下对整个系统特性的捕捉能力。为了方便不同模型之间的对比, 对每个降阶模型都采用相同的训练数据信号形式——变幅值 chirp 信号, 训练集、测试集部分信号形式如图 5 所示, 信号频率范围为 0～5 Hz。有效迎角变化范围为−20°～20°。

此外, 为了使模型能够应用于不同来流速度, 训练信号需在若干速度点 V_1～V_n 下进行气动力数据采集并进行降阶模型训练, 最后将各个速度点下的模型进行整合获得最终模型。

测试信号选取方面, 选取平白高斯噪声 (flat white Gaussian noise, FWGN) 信号作为测试集运动信号。测试集的有效迎角变化范围为−16°～20°。测试集部分信号形式(见图 5)。

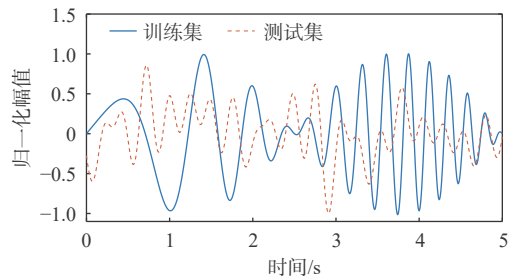


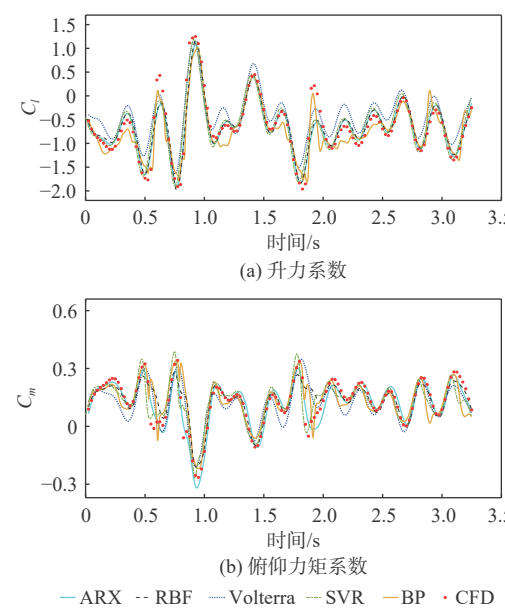
图 5 训练集和测试集部分信号形式

Fig. 5 Partial signal form of training and test set

3.3 各气动力降阶模型特性对比

3.3.1 各降阶模型有效性验证

基于不同方法的气动力降阶模型在测试集上的输出结果和 CFD 计算结果对比如图 6 所示。可以看出, 所有气动力降阶模型都能在一定程度上拟合与跟随 CFD 计算结果。从而证明了模型的有效性, 在此基础上将对各模型的收敛性及泛化能力进行研究。



(a) 升力系数

(b) 俯仰力矩系数

图 6 各模型在测试集上的输出结果和 CFD 计算结果对比

Fig. 6 Comparison between output results of different reduced order models and CFD methods ontest data set

3.3.2　各降阶模型收敛性对比

设置目标速度范围为 10～60 m/s, 训练信号形式及有效迎角变化范围与 3.2 节中相同。在这一速度范围内设置多个训练组, 训练集速度点设置如表 2 所示。测试集速度点设置如表 3 所示。计算各模型在不同训练情况下, 所有测试集输出结果与 CFD 结果相比的均方根误差 (RMSE) 值之和, 来比较各降阶模型随训练集速度点个数的收敛性, 如图 7 所示。可以看出, 各模型随着训练集速度点样本数的增加, RMSE 值逐渐减小, 模型逐渐收敛。其中基于 ARX 的气动降阶模型收敛性最好, 即需要较少的训练集速度样本点即可在目标速度范围内达到较好的模型输出结果。

表 2　训练集速度点设置

Table 2　Speed settings for training sets

训练集编号	训练集包含的速度点/(m·s^{-1})
1	10,60
2	10,35,60
3	10,20,30,40,50,60
4	10,15,20,25,30,35,40,45,50,55,60

表 3　测试集速度点设置

Table 3　Speed settings for test sets

测试集编号	测试集速度/(m·s^{-1})
1	55
2	33
3	22
4	12
5	10

3.3.3　各降阶模型泛化能力对比

采用表 2 中的训练集 3 对各降阶模型进行训练获得 10～60 m/s 速度范围内的气动力降阶模型。之后将各模型应用于 100, 150, 200, 250 m/s 来流速度下的测试集进行气动力系数计算, 将输出结果与 CFD 计算结果进行对比后, 可得各模型对来流速度的泛化能力, 如图 8 所示。可以看出, 基于 ARX 和多小波 Volterra 级数的气动力降阶模型对不同来流速度的泛化能力较强, 基于 SVR 的模型次之, 基于 BP 和 RBF 神经网络模型的泛化能力较弱。

3.3.4　各降阶模型建模及计算效率对比

各降阶模型建模效率可通过实际建模所需时间成本来衡量, 建模时间成本包括训练数据获取时间成本和模型参数辨识时间成本 2 部分。其中模型参数辨识时间成本相对于 CFD 训练数据获取时间成本只是小量, 因此, 本文只通过训练数据获取时间成本来衡量各模型建模效率。结合 3.3.2 节和 3.3.3 节的内容可知, 收敛性越好, 泛化能力越强, 模型建模所需的原始气动力数据越少, 建模训练数据获取时间成本越小。因此, 基于 ARX 的模型建模效率最高, 基于多小波 Volterra 级数的模型次之, 基于 SVR、BP 和 RBF 神经网络的模型建模效率较低。

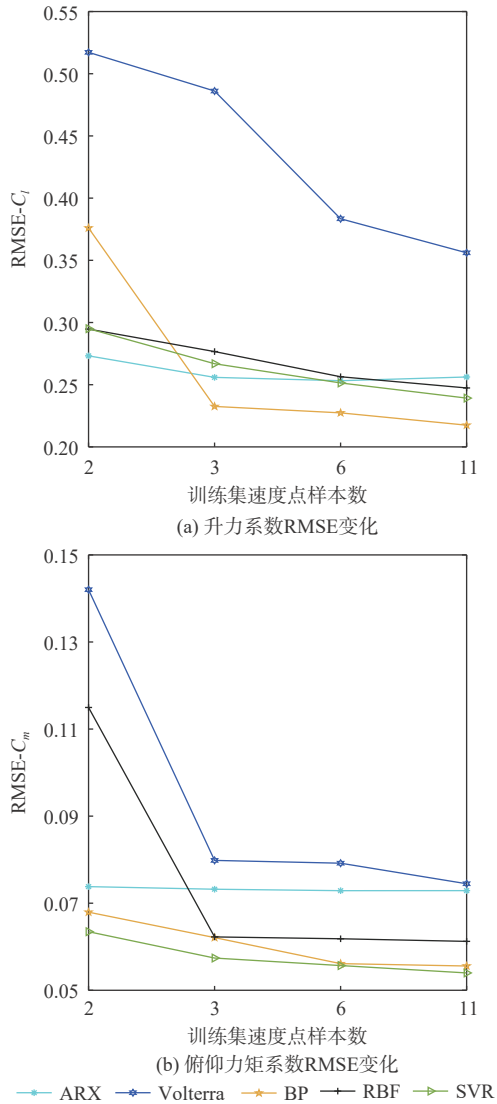


图 7　各模型随训练速度点样本数增加收敛性对比

Fig. 7　Comparison of convergence of each model with increasing sample size of training speed points

在计算效率方面, 本文对比各模型在相同测试集下的计算时间(见表 4)。测试集共有 3 500 个时间步。所有模型均利用 MATLAB 建立, 其中 BP、RBF、SVR 模型建模基于 MATLAB 工具箱和自行编写的参数优化算法, ARX 和多小波 Volterra

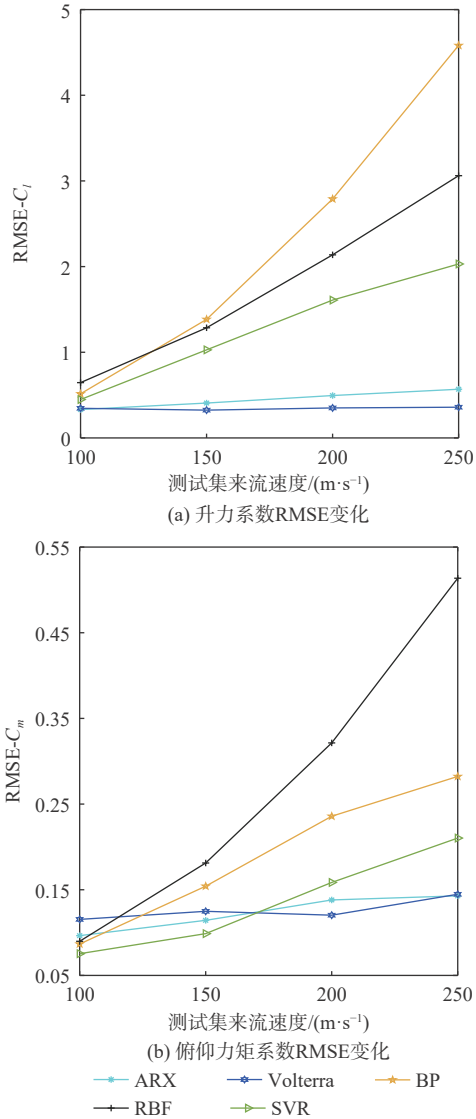


图 8 各模型对来流速度泛化能力对比

Fig. 8 Comparison of generalization ability of each model for incoming flow velocity

模型程序通过自行编写完成。各模型训练集与测试集均相同。

从表 4 中可以看出，本文建立的基于 ARX 和基于 BP 神经网络的模型计算效率最高，基于 RBF 神经网络的模型及基于 SVR 的模型次之，基于多小波 Volterra 的模型计算效率最低。需要说明的是这里的计算效率比较仅限于本文建立的模型之间的比较，如果采用不同编程语言或者模型结构可能会对本文结果产生影响。

表 4 各模型相同测试集计算时间对比

Table 4 Comparison of calculation efficiency for same test set for each model

模型	计算时间/s
ARX	0.005
Volterra	0.080
BP	0.009
RBF	0.045
SVR	0.052

3.4 组合模型建立

基于 3.3 节内容，将收敛性、泛化能力较好的 ARX 模型和非线性辨识能力强的 BP、RBF 神经网络模型组合，首先利用 ARX 模型搭建初始气动力降阶模型框架。然后基于 CFD 计算结果和初始气动力降阶模型输出之差，利用 BP/RBF 神经网络模型搭建非线性气动力细节修正模块。最后将初始模型和非线性气动力细节修正模块相组合得到最终的非线性气动力降阶模型。组合模型结构如图 9 表示。

为了验证组合模型的优越性，将 ARX-BP/RBF 组合模型的收敛性及泛化能力与基于单一方法的原始模型进行对比，如图 10 和图 11 所示。可以看出，组合模型的收敛性和泛化能力优于原始模型，且通过 RMSE 值可以看出组合模型的计算精度均优于原始 ARX 模型。

对组合模型的计算时间进行对比研究，如表 5 所示，计算时间研究测试集设置与 3.3.4 节相同。

从表 5 和表 4 可以看出，组合模型由于模型结构复杂度增加，计算时间和原始模型相比稍有降低，但在可接受范围内。

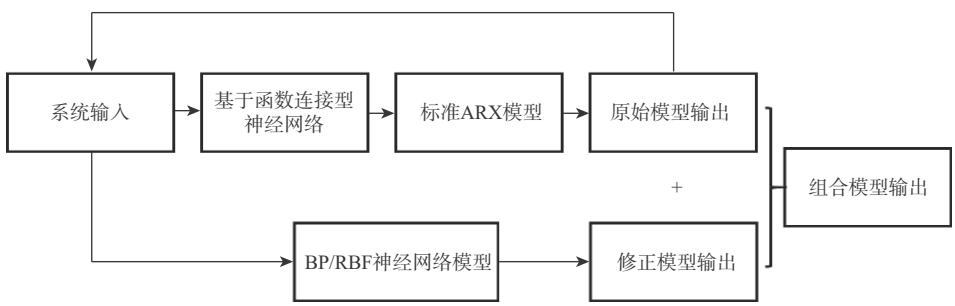


图 9 ARX-BP/RBF 组合模型结构

Fig. 9 Structure of ARX-BP/RBF combined model

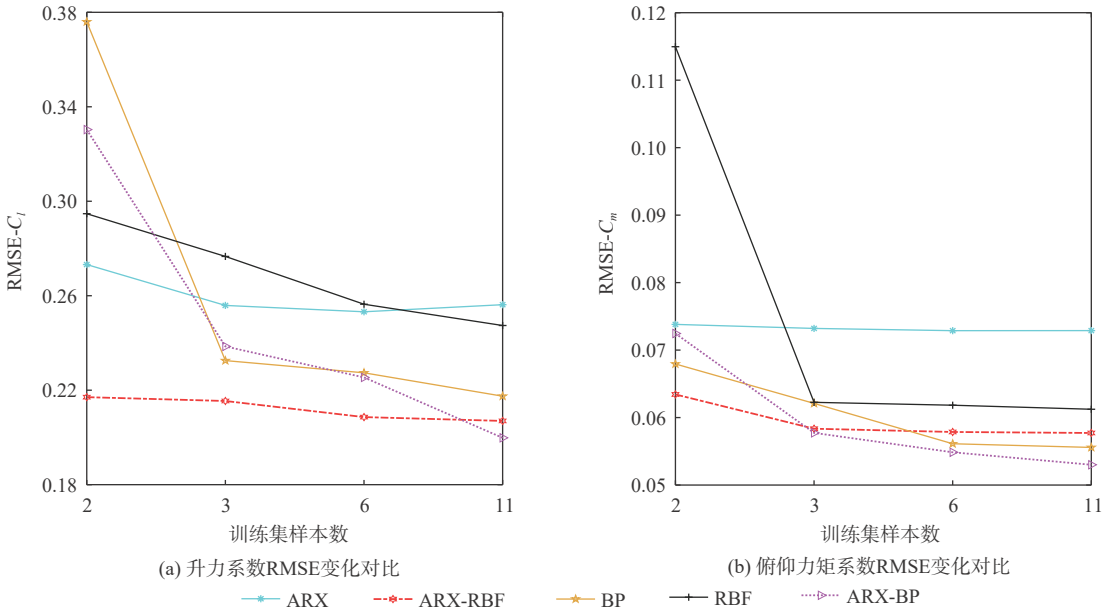


图 10　组合模型和原始模型收敛性对比

Fig. 10　Comparison of convergence between combined model and original model

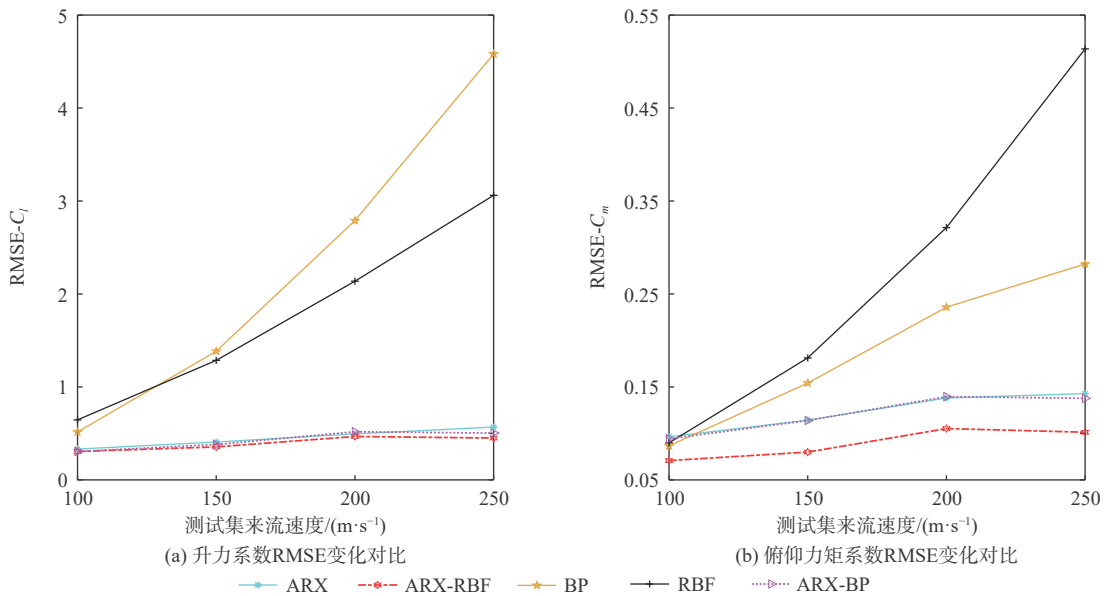


图 11　组合模型和原始模型泛化能力对比

Fig. 11　Comparison of generalization ability of combined model and original model

表 5　组合模型相同测试集计算时间对比

Table 5　Comparison of calculation efficiency for same test set for each model

模型	计算时间/s
ARX-RBF	0.050 5
ARX-BP	0.012 0

4　基于 ARX-RBF 组合模型的三维弹性飞机飞行仿真模型

4.1　算例模型选取

选取一架太阳能无人机模型作为仿真分析算例, 该模型主要参数如表 6 所示。

4.2　结构建模

图 12 为太阳能无人机算例的结构模型。取六阶刚体模态与前六阶弹性模态进行结构动力学建模, 弹性模态频率如表 7 所示, 弹性模态振型如图 13 所示。

4.3　气动建模

4.3.1　CFD 建模

太阳能无人机算例模型整机外表面网格及局部放大如图 14 所示, 气动网格为四面体网格, 模型网格节点数为 280 508, 全流场网格单元数为 3 145 935。

采用和二维算例同样的CFD设置建立CFD计算模型，并验证了网格无关性。

表6 太阳能无人机结构主要参数

Table 6 Main structural parameters of solar UAV

结构参数	数值
机翼展长/m	50.00
机翼弦长/m	1.72
平尾展长/m	10.20
平尾弦长/m	1.52
垂尾展长/m	4.77
垂尾弦长/m	1.52
质量/kg	424.70
质心到机头距离/m	10.87
转动惯量I_{xx}/(kg·m^2)	5.938 7×10^4
转动惯量/I_{yy}/(kg·m^2)	4.806 1×10^3
转动惯量/I_{zz}/(kg·m^2)	6.415 6×10^4
转动惯量/I_{xz}/(kg·m^2)	0.251 6

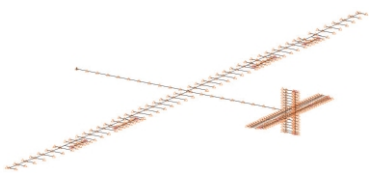


图12 太阳能无人机结构模型

Fig. 12 Structural model of solar UAV

表7 弹性模态频率

Table 7 Elastic modal frequency

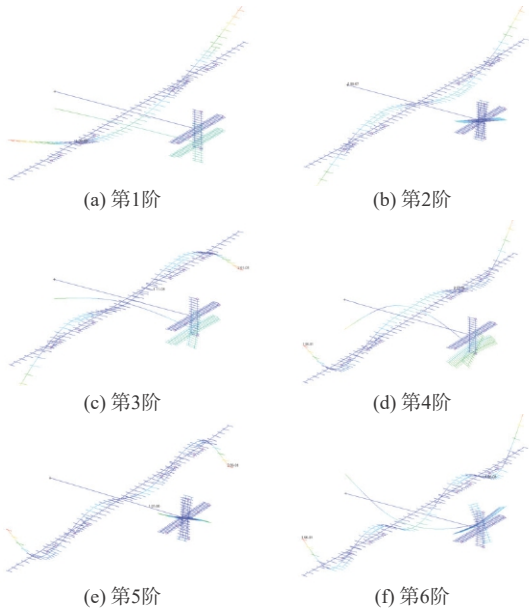
阶数	频率/Hz
1	0.48
2	1.33
3	2.01
4	3.36
5	3.55
6	4.74

4.3.2 气动力降阶建模

由于大展弦比飞机的特殊性，这里结合3.4节提出的ARX-RBF组合二维气动力模型与片条理论建立三维模型的气动力降阶初始模型，三维模型气动片条划分俯视图如图15所示。

在这种情况下，每个片条上的气动力都通过降阶模型获得。每个片条的迎角按照式(12)可写为

$$\begin{cases}\alpha_{1j}=\alpha_{\mathrm{r}1}+\sum\limits_i\varphi_{y_ij}\xi_i\\ \alpha_{2j}=\alpha_{\mathrm{r}2}-\sum\limits_i\varphi_{z_ij}\xi_i\end{cases}\tag{23}$$



(a) 第1阶 (b) 第2阶 (c) 第3阶 (d) 第4阶 (e) 第5阶 (f) 第6阶

图13 弹性模态振型

Fig. 13 Elastic mode shape

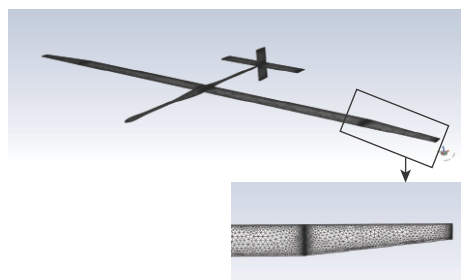


图14 整机外表面网格及局部放大

Fig. 14 Grid of external surface and local enlarged view of UAV

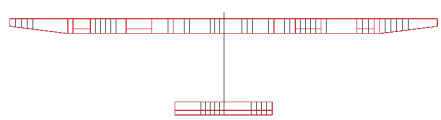


图15 气动片条划分俯视图

Fig. 15 Top view of slice model

式中：α_{1j}为机翼与平尾上第j个片条的迎角；α_{2j}为垂尾上第j个片条的迎角；i为模态编号；φ_{y_ij}为第j个片条第i阶模态沿Oy轴的转动分量；φ_{z_ij}为第j个片条第i阶模态沿Oz轴的转动分量。

同样的，机翼与平尾上第j个片条的$\dot{h}_{1j}$和垂尾上第j个片条的$\dot{h}_{2j}$为

$$\begin{cases}\dot{h}_{1j}=\left(\boldsymbol{\omega}\times\boldsymbol{s}+\mathrm{d}\left(\sum\limits_i\boldsymbol{\phi}_{z_ij}\xi_i\right)/\mathrm{d}t\right)\cdot\\ \qquad[-\sin\alpha_{\mathrm{r}1},0,\cos\alpha_{\mathrm{r}1}]^{\mathrm{T}}-\\ \qquad\boldsymbol{w}_{\mathrm{b}}{}^{\mathrm{T}}[-\sin\alpha_{\mathrm{r}1},0,\cos\alpha_{\mathrm{r}1}]^{\mathrm{T}}\\ \dot{h}_{2j}=\left(\boldsymbol{\omega}\times\boldsymbol{s}+\mathrm{d}\left(\sum\limits_i\boldsymbol{\phi}_{y_ij}\xi_i\right)/\mathrm{d}t\right)\cdot\\ \qquad[-\sin\alpha_{\mathrm{r}2},\cos\alpha_{\mathrm{r}2},0]^{\mathrm{T}}-\\ \qquad\boldsymbol{w}_{\mathrm{b}}{}^{\mathrm{T}}[-\sin\alpha_{\mathrm{r}2},\cos\alpha_{\mathrm{r}2},0]^{\mathrm{T}}\end{cases}\tag{24}$$

式中: $\boldsymbol{\phi}_{z_ij}$为第 j 个片条在第 i 阶模态作用下沿 Oz 轴的平移分量; $\boldsymbol{\phi}_{y_ij}$为第 j 个片条在第 i 阶模态作用下沿 Oy 轴的平移分量。

再将每个片条的弦长和翼段长度引入作为输入参数, 每个片条的气动力降阶模型输入参数为

$$\begin{cases}\boldsymbol{u}_{1j}=[V_1,\alpha_{1j},\dot{h}_{1j},c_{1j},l_{1j},\delta_{1j}]^{\mathrm{T}}\\ \boldsymbol{u}_{2j}=[V_2,\alpha_{2j},\dot{h}_{2j},c_{2j},l_{2j},\delta_{2j}]^{\mathrm{T}}\end{cases}\tag{25}$$

式中: $\boldsymbol{u}_{1j}$为机翼与平尾上第 j 个片条气动力降阶模型的输入参数; $\boldsymbol{u}_{2j}$为垂尾上第 j 个片条气动力降阶模型的输入参数; c_{1j}和l_{1j}分别为机翼与平尾上第 j 个片条的弦长和展长; c_{2j}和l_{2j}分别代表垂尾上第 j 个片条的弦长和展长; δ_{1j}为机翼与平尾上第 j 个片条上的控制面参数, 若第 j 个片条上有控制面, 则δ_{1j}取值为该控制面偏转角, 否则δ_{1j}取值为 0。对于垂尾上的控制面参数δ_{2j}也是同样的设置方法。$\boldsymbol{\alpha}_{1j}=[\alpha_{1j},\dot{\alpha}_{1j},\ddot{\alpha}_{1j}]$, $\boldsymbol{\alpha}_{2j}=[\alpha_{2j},\dot{\alpha}_{2j},\ddot{\alpha}_{2j}]$。

机翼与平尾上第 j 个片条气动力降阶模型的输出$\boldsymbol{y}_{1j}$和垂尾上第 j 个片条气动力降阶模型的输出$\boldsymbol{y}_{2j}$为

$$\begin{cases}\boldsymbol{y}_{1j}=[L_{1j},D_{1j},M_{1j}]^{\mathrm{T}}\\ \boldsymbol{y}_{2j}=[L_{2j},D_{2j},M_{2j}]^{\mathrm{T}}\end{cases}\tag{26}$$

式中: L_{1j}、D_{1j}和M_{1j}分别为机翼与平尾上第 j 个片条沿 $O_{\mathrm{a}}z_{\mathrm{a}}$ 轴的升力、沿 $O_{\mathrm{a}}x_{\mathrm{a}}$ 轴的阻力和绕 $O_{\mathrm{a}}y_{\mathrm{a}}$ 轴的力矩; L_{2j}、D_{2j}和M_{2j}分别代表垂尾上第 j 个片条沿 $O_{\mathrm{a}}y_{\mathrm{a}}$ 轴方向的力、沿 $O_{\mathrm{a}}x_{\mathrm{a}}$ 轴方向的力和绕 $O_{\mathrm{a}}z_{\mathrm{a}}$ 轴的力矩。

将每个片条上风轴系下气动力通过坐标轴转换即可得到体轴系的气动力。将体轴系下的气动力乘以气动模态阵即可得到弹性模态对应的广义气动力$Q_1,Q_2,\cdots,Q_n$。再将转换后的体轴系下各片条气动力进行积分求和后可获得体轴系下所有片条在 6 个自由度方向的刚体气动力合力 $f_{u1},f_{v1},f_{w1},f_{p1},f_{q1},f_{r1}$, 它们来自机翼、平尾、垂尾、控制面产生的气动力, 对应式 (8) 中的$\boldsymbol{Y}_{\mathrm{w}}+\boldsymbol{Y}_{\mathrm{ht}}+\boldsymbol{Y}_{\mathrm{vt}}+\boldsymbol{Y}_{\mathrm{c}}$。

按照以式(25)为输入, 式(26)为输出建立的气动力降阶初始模型无法考虑机身产生的气动力的影响, 本文以 CFD 计算数据为基础, 对气动力降阶初始模型引入二次修正模块, 整体构建$\boldsymbol{Y}_{\mathrm{b}}$项。该修正模块基于 RBF 神经网络模型, 模型输入 $\boldsymbol{u}_{\mathrm{b}}=[V_1,\alpha_1,\dot{h}_1]^{\mathrm{T}}$, V_1为来流速度, α_1为整机的飞行迎角参数, $\dot{h}_1$为整机质心垂直于V_1方向的运动速度。模型的目标输出为 CFD 计算结果和气动力降阶初始模型输出之差, 对应于式 (8) 中的$\boldsymbol{Y}_{\mathrm{b}}$。将气动力降阶初始模型和二次修正模块相结合即可得到最终的气动力降阶模型。三维气动力降阶模型运行流程如图 16 所示。

为验证气动力降阶模型的正确性, 将给定运动下的气动力降阶模型输出结果和 CFD 计算结果进行了对比。给定运动及变形信号如图 17 所示, 几个关键模态气动力的气动力降阶模型输出结果和 CFD 计算结果对比如图 18 所示。本文通过 Fluent 的用户自定义函数 (user-defined function, UDF) 模块将给定运动信号加载到 CFD 求解器。首先可在 UDF 模块中根据给定的各模态运动信号值计算出 t 时刻的机体表面所有网格节点位置, 之后通过动网格方法对全流场网格进行更新并进行 t 时刻 CFD 气动力计算, 计算完成后再回到 UDF 模块计算 t+1 时刻的网格节点位置, 如此循环即可实现给定运动下的 CFD 计算。从图 18 可以看出, 初始模型和二次修正模型在一定程度上都能较好地反映气动力变化趋势, 而二次修正后的气动力降阶模型的输出结果和 CFD 一致性更好, 由此验证了气动力降阶建模方法的有效性。

4.4　飞行仿真建模

将 4.2 节中的结构模型与 4.3 节中的气动模型按照式 (1)～式 (3) 进行组合即可得到弹性飞机刚弹耦合稳定性分析模型。然后将舵面运动和阵风产生的影响分别转化为各片条的运动参数和气动参数变化, 即可建立如图 1 所示的舵面及阵风扰动下的飞行力学模型。

4.4.1　与 CFD-CSD 计算结果对比

使用构建好的飞行仿真模型, 对飞机模型在放开沉浮和俯仰自由度状态下, 受到正弦阵风扰动时的响应进行仿真, 并与 CFD-CSD 耦合方法计算结果进行对比, 这里 CFD-CSD 耦合方法通过 Fluent 流体求解和 UDF 结构求解来实现。飞行仿真模型输出结果与 CFD-CSD 结果对比如图 19 所示, 来流及阵风参数如表 8 所示。可以看出, 仿真模型的输出结果与 CFD-CSD 计算结果一致性较好, 证明了模型的正确性和有效性。对于该验证算例, 基于气动力降阶模型的飞行仿真模型运行时间与 CFD-CSD 耦合方法所需时间对比如表 9 所示。可以看出, 同样计算条件下, 本文模型仿真时间只有 CFD-CSD 耦合模型的 1/33, 表明本文建立的基于气动力降阶的飞行仿真模型极大地提高了仿真效率, 适用于工程应用与分析。

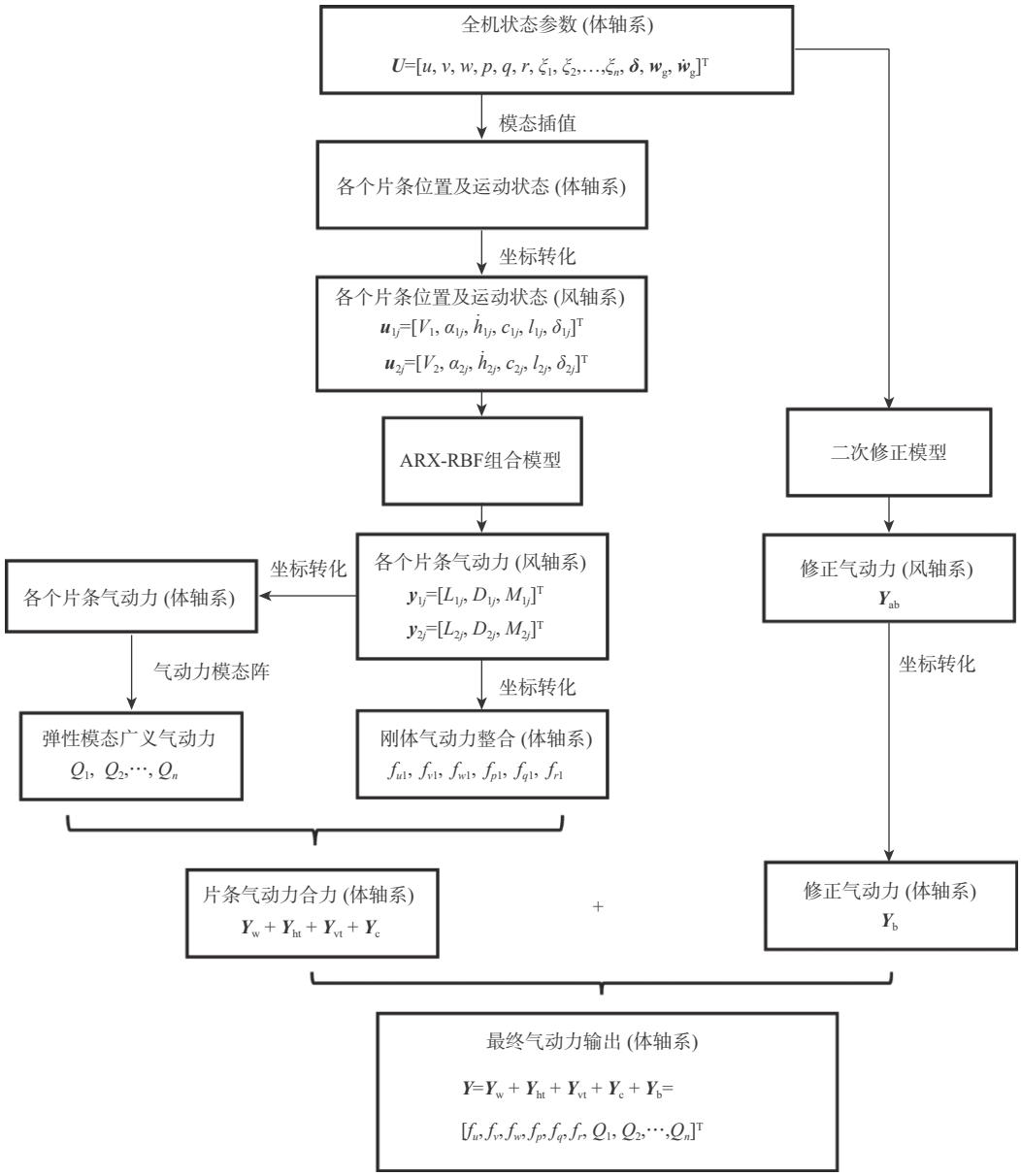


图 16 三维气动力降阶模型运行流程

Fig. 16 Operation process of three dimensional aerodynamic reduced order model

4.4.2 与基于线性气动力的仿真模型对比

对给定阵风下的飞机响应进行计算，并与基于线性气动力的飞行仿真模型结果进行对比，加速度响应对比如图 20～图 22 所示，来流及阵风参数如表 10 所示。

表 10 中的尺度 L_1、L_2 表示离散阵风风场范围位置，即在 $x \in [L_1, L_2]$ 范围内形成风场，具体数学表达式为

$$w_g(x)=\begin{cases} 0 & x < L_1 \text{或} x > L_2 \\ \dfrac{w_m}{2}\left[1-\cos\left(2\pi\dfrac{x-L_1}{L_2-L_1}\right)\right] & L_1 < x < L_2 \end{cases} \tag{27}$$

式中：w_m 为阵风幅值。

从图 20 和图 21 中可以看出，在阵风幅值较小的情况下本文模型仿真结果和基于线性气动力的仿真模型结果一致性较好，再次验证了本文方法的有效性。但随着阵风幅值的增加，气动力出现非线性特性，本文模型仿真结果和基于线性气动力的仿真模型结果的差异逐步显现，本文模型对加速度响应的仿真结果会略小于基于线性气动力的仿真结果，这是因为阵风幅值较大时，引起的附加迎角会使气流产生一定分离，气动力相应减小，如图 22 所示。表明本文模型能在一定程度上反映气动力非线性的影响。

对给定舵偏角情况下的飞机响应进行仿

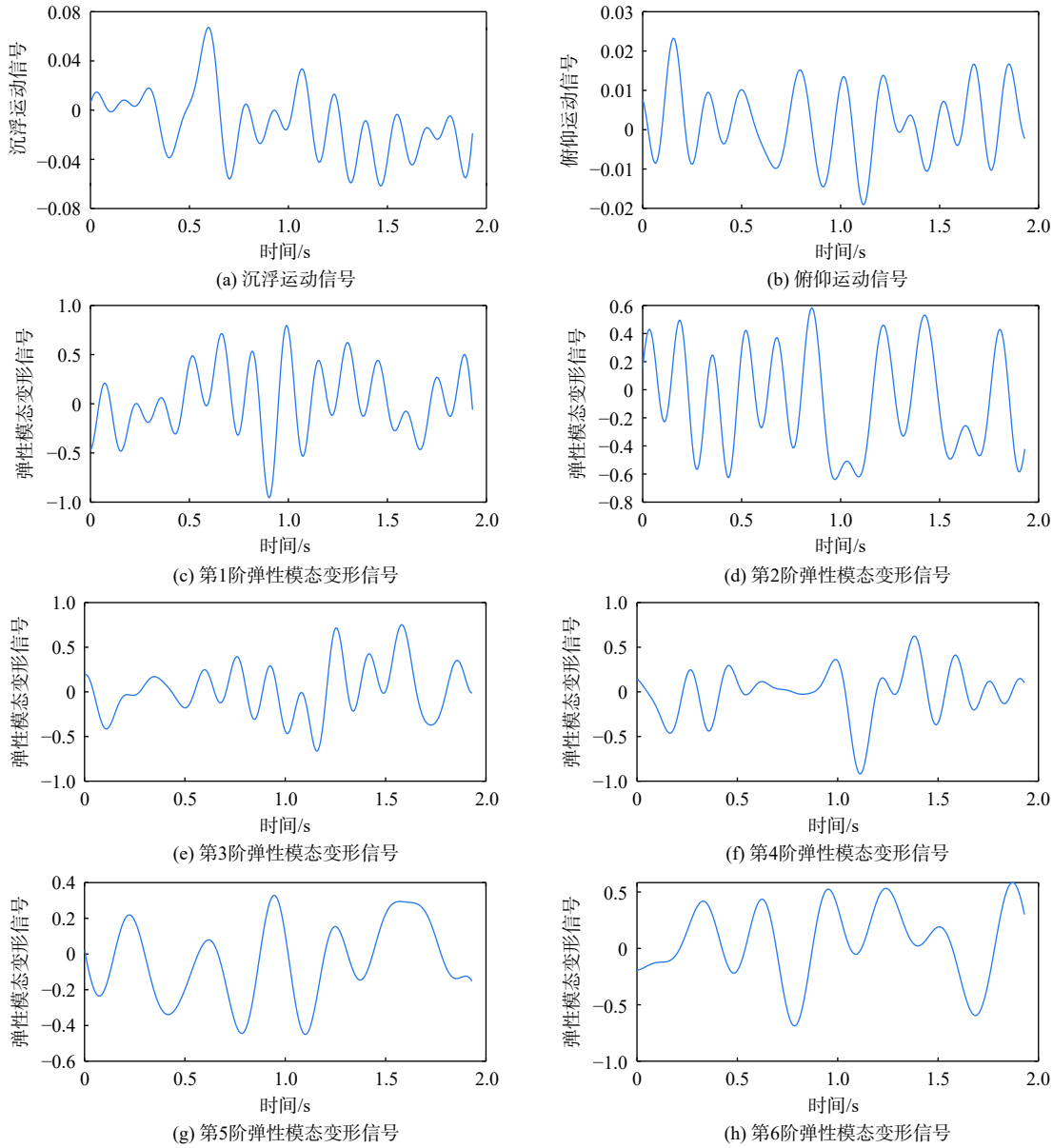


图 17　CFD 计算结果的给定运动及变形信号

Fig. 17　Motion and deformation signal for CFD calculation

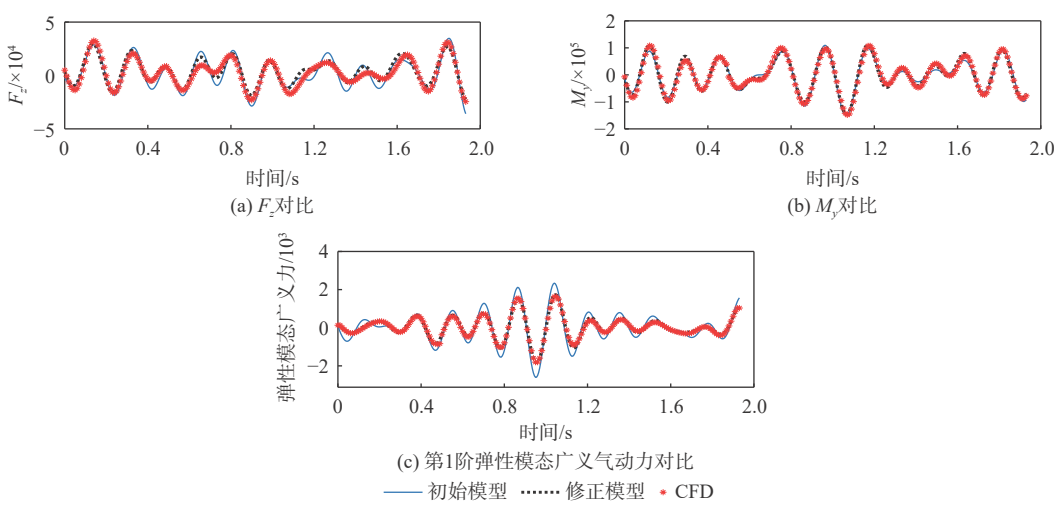


图 18　初始、修正模型输出结果与 CFD 计算结果对比

Fig. 18　Comparison of initial and modified model output results with CFD calculation results

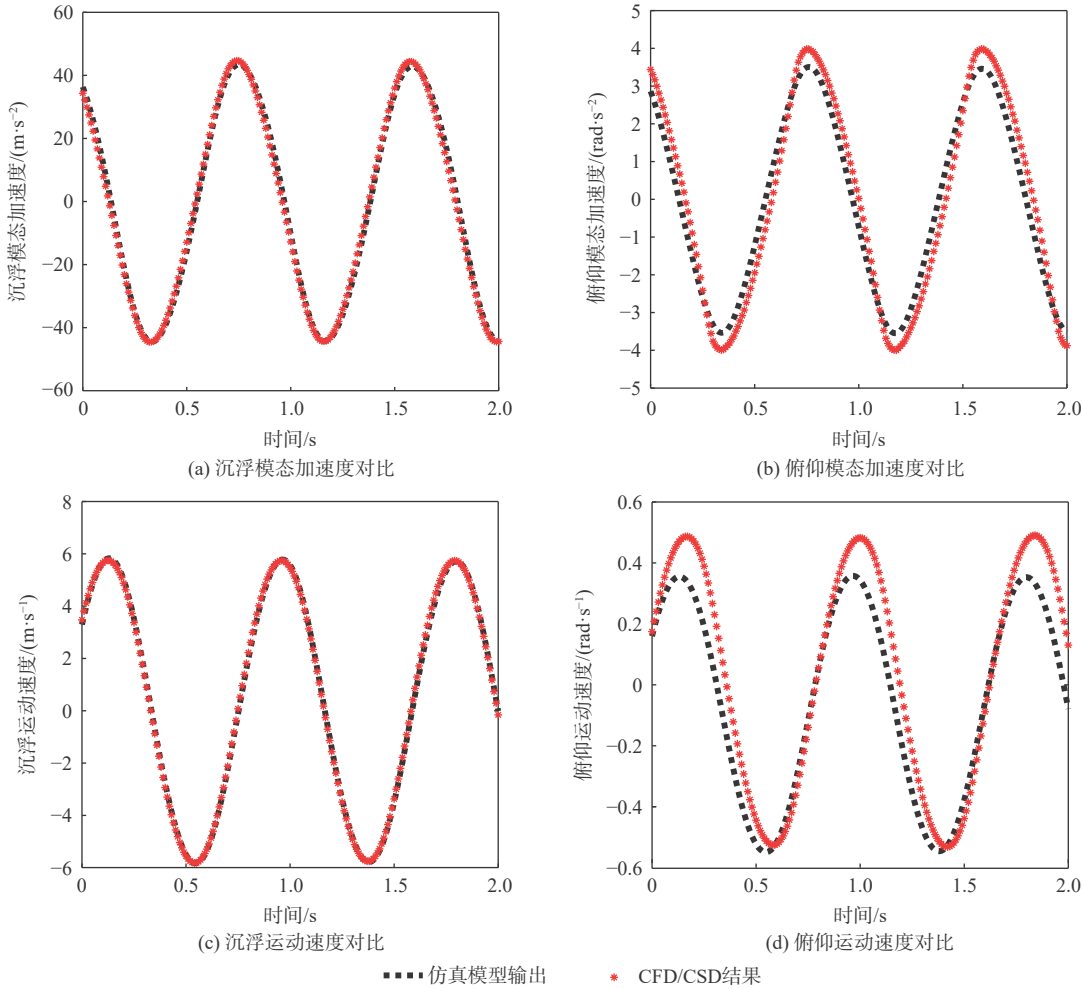


图 19 飞行仿真模型输出结果与 CFD-CSD 结果对比

Fig. 19 Comparison of flight simulation model output results with CFD-CSD results

表 8 来流及 SIN 型阵风参数

Table 8 Incoming flow and SIN type gust parameters

来流风速/(m·s⁻¹)	阵风形式	阵风幅值/(m·s⁻¹)	阵风频率/ Hz
20	SIN型	7	1.2

表 9 计算效率对比

Table 9 Comparison of calculation efficiency

模型	时间步长/s	时间步数	计算时间/min
CFD-CSD	0.001	5 000	2 398
本文模型	0.001	5 000	77

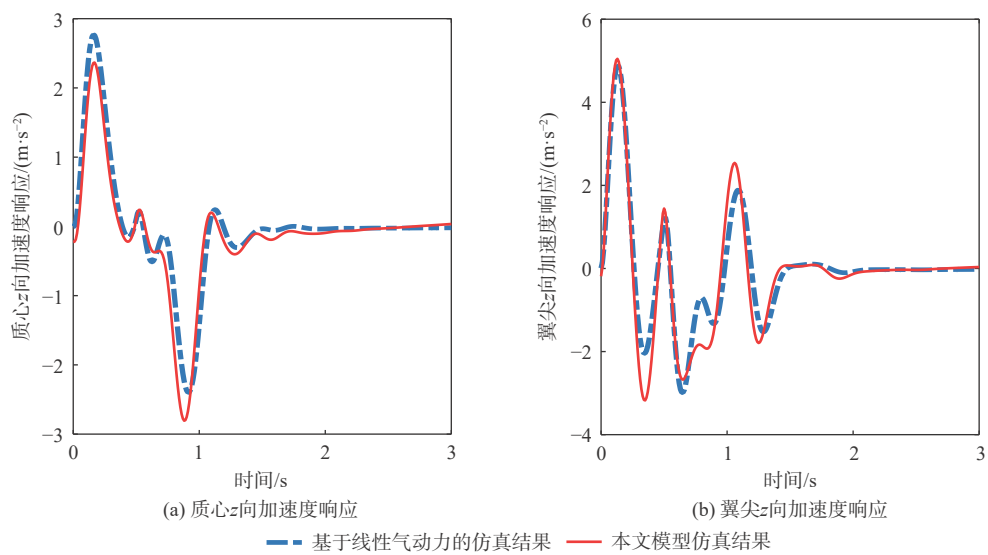


图 20 阵风算例 1 加速度响应对比

Fig. 20 Comparison of acceleration response under gust case 1

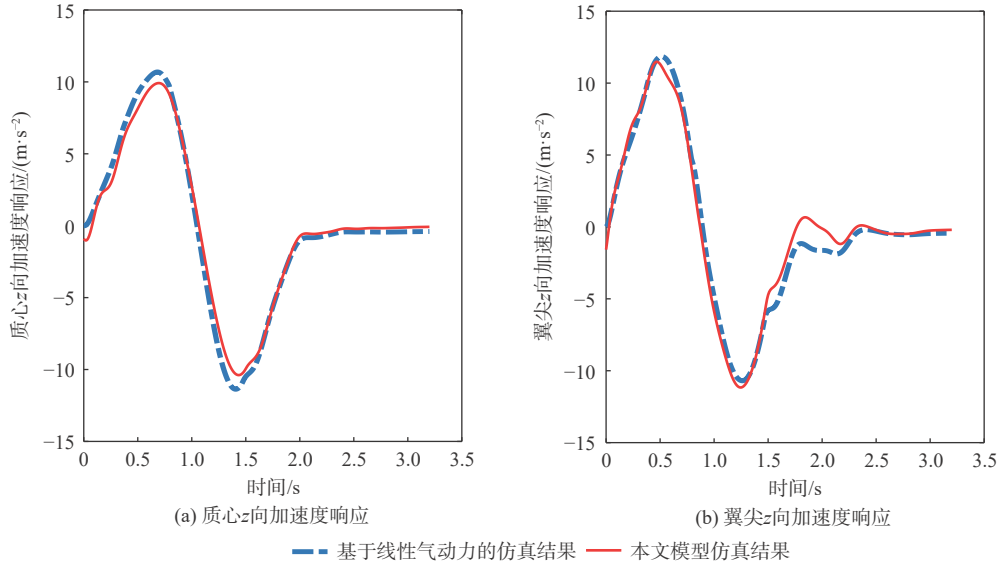


图 21　阵风算例 2 加速度响应对比

Fig. 21　Comparison of acceleration response under gust case 2

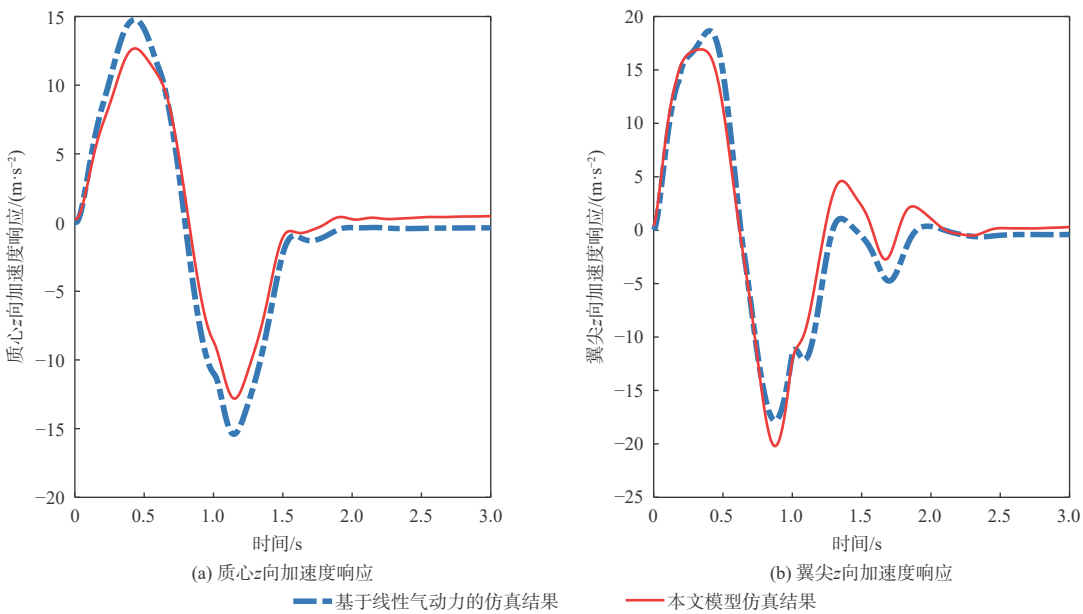


图 22　阵风算例 3 加速度响应对比

Fig. 22　Comparison of acceleration response under gust case 3

表 10　来流及 1-COS 型型阵风参数

Table 10　Incoming flow and 1-COS type gust parameters

编号	来流风速/(m·s^{-1})	阵风形式	阵风幅值/(m·s^{-1})	尺度L_1/m	尺度L_2/m
1	20	1-COS型	1	30	40
2	20	1-COS型	5	10	40
3	20	1-COS型	7	30	50

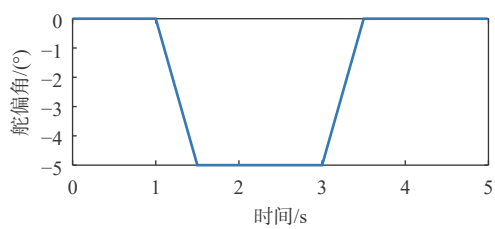
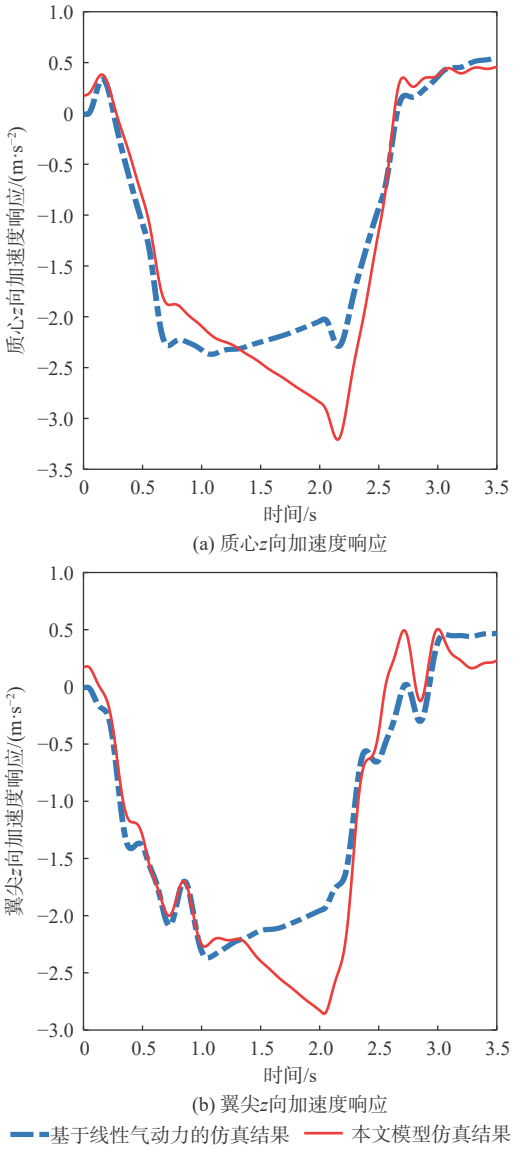


图 23　舵偏角变化曲线

Fig. 23　Variation curve of rudder deflection angle

真, 给定舵偏角变化曲线如图 23 所示。同样的也与基于线性气动力的飞行仿真结果进行对比, 如图 24 所示。从图 24 中可以看出, 本文模型仿真结果与基于线性气动力的仿真结果一致性较好。



(a) 质心z向加速度响应

(b) 翼尖z向加速度响应

图 24 给定舵偏角变化后加速度响应对比

Fig. 24 Comparison of acceleration response for a given rudder deflection angle change

5 结 论

1）本文模型中的 ARX-RBF 组合气动力降阶模型计算精度高于基于 ARX 的模型，收敛性和泛化能力强于基于 BP/RBF 神经网络的模型。

2）本文模型的弹性飞机飞行仿真模型计算结果与 CFD-CSD 耦合模型一致性较好，计算效率远高于 CFD-CSD 耦合算法。同样计算条件下，本文模型仿真时间只有 CFD-CSD 耦合模型的 1/33。

3）本文模型的弹性飞机飞行仿真模型计算结果与基于线性气动力的飞行仿真模型在线性气动力范围内仿真结果一致性较好，并且在非线性气动力范围能够有效反映出非线性气动力带来的影响。

参考文献（References）

[1] 王少奇, 马东立, 杨穆清, 等. 高空太阳能无人机三维航迹优化[J]. 北京航空航天大学学报, 2019, 45(5): 936-943.
WANG S Q, MA D L, YANG M Q, et al. Three-dimensional optimal path planning for high-altitude solar-powered UAV[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(5): 936-943(in Chinese).

[2] 刘藤, 李栋, 黄冉冉, 等. 基于降阶模型的翼型结冰冰形预测方法[J]. 北京航空航天大学学报, 2019, 45(5): 1033-1041.
LIU T, LI D, HUANG R R, et al. Ice shape prediction method of aero-icing based on reduced order model[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(5): 1033-1041(in Chinese).

[3] 杨国伟, 王济康. CFD结合降阶模型预测阵风响应[J]. 力学学报, 2008, 40(2): 145-153.
YANG G W, WANG J K. Prediction of gust response by CFD combined with reduced order model[J]. Acta Mechanica Sinica, 2008, 40(2): 145-153(in Chinese).

[4] 张伟伟, 叶正寅. 基于气动力降阶模型的跨音速气动弹性稳定性分析[J]. 振动工程学报, 2007, 24(6): 768-772.
ZHANG W W, YE Z Y. Transonic aeroelastic stability analysis based on aerodynamic reduced order model[J]. Journal of Vibration Engineering, 2007, 24(6): 768-772(in Chinese).

[5] 师妍, 万志强, 吴志刚, 等. 基于气动力降阶的弹性飞机阵风响应仿真分析及验证[J]. 航空学报, 2022, 43(1): 125474.
SHI Y, WAN Z Q, WU Z Q, et al. Gust response analysis and verification of elastic aircraft based on nonlinear aerodynamic reduced order model[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(1): 125474(in Chinese).

[6] LIND R, PRAZENICA R J, BRENNER M J, et al. Identifying parameter dependent Volterra kernels to predict aeroelastic instabilities[J]. AIAA Journal, 2005, 43(12): 2496-2502.

[7] 吴志刚, 杨超. 基于Volterra级数的跨音速非定常气动力建模[J]. 北京航空航天大学学报, 2006, 32(4): 373-376.
WU Z G, YANG C. Volterra series based transonic unsteady aerodynamics modeling[J]. Journal of Beijing University of Aeronautics and Astronautics, 2006, 32(4): 373-376(in Chinese).

[8] PRAZENICA R J, REISENTHEL P H, KURDILA A J, et al. Volterra kernel extrapolation for modeling nonlinear aeroelastic systems at novel flight conditions[J]. Journal of Aircraft, 2007, 44(1): 149-162.

[9] OMRAN A, NEWMAN B. Full envelope nonlinear parameter-varying model approach for atmospheric flight dynamics[J]. Journal of Guidance, Control, and Dynamics, 2012, 35(1): 270-283.

[10] 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. Aerodynamic second order kernel function identification method[J]. Acta Aeronautica et Astronautica Sinica, 2014, 35(11): 2949-2957(in Chinese).

[11] 陈森林, 高正红, 饶丹. 基于多小波的Volterra级数非定常气动力

建模方法[J]. 航空学报, 2018, 39(1): 121379.
CHEN S L, GAO Z H, RAO D. Modeling method of unsteady aerodynamics based on multiwavelet Volterra series[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(1): 121379(in Chinese).

[12] 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 modeling method for aerodynamic data of flight simulation[J]. Acta Aerodynamics Sinica, 2019, 37(3): 488-497(in Chinese).

[13] ZHANG W W, WANG B B, YE Z Y, et al. Efficient method for limit cycle flutter analysis based on nonlinear aerodynamic reduced-order models[J]. AIAA Journal, 2012, 50(5): 1019-1028.

[14] KOU J Q, ZHANG W W. Reduced-order modeling for nonlinear aeroelasticity with varying Mach numbers[J]. Journal of Aerospace Engineering, 2018, 31(6): 04018105.

[15] KOU J Q, ZHANG W W. Multi-kernel neural networks for nonlinear unsteady aerodynamic reduced-order modeling[J]. Aerospace Science and Technology, 2017, 67: 309-326.

[16] JIANG Y, ZHAO Q, ZHU J. Unsteady aerodynamics modeling using SVM and artificial neural network[C]//Proceedings of the 2015 Chinese Intelligent Automation Conference: Intelligent Information Processing. Berlin : Springer, 2015: 577-585.

[17] WANG Q, QIAN W Q, HE K F. Unsteady aerodynamic modeling at high angles of attack using support vector machines[J]. Chinese Journal of Aeronautics, 2015, 28(3): 659-668.

[18] CHEN S L, GAO Z H, ZHU X Q, et al. Unstable unsteady aerodynamic modeling based on least squares support vector machines with general excitation[J]. Chinese Journal of Aeronautics, 2020, 33(10): 2499-2509.

[19] NARENDRA K, GALLMAN P. An iterative method for the identification of nonlinear systems using a Hammerstein model[J]. IEEE Transactions on Automatic Control, 1966, 11(3): 634-638.

[20] SHENG W, GALBRAITH R A, COTON F N. A modified dynamic stall model for low mach numbers[J]. Journal of Solar Energy Engineering, 2007, 130(3): 653-653.

Aerodynamic order reduction method for elastic aircraft flight dynamics simulation

SHI Yan[1], WAN Zhiqiang[1], WU Zhigang[1,*], GONG Yinan[2]

(1. School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China;
2. Seahawk Aviation General Equipment Co., Ltd., Beijing 100070, China)

Abstract: We develop an aeroelasticity and flight dynamics simulation model with nonlinear aerodynamic reduced order model for low-speed solar unmanned aerial vehicles under high angle of attack caused by gust. Five aerodynamic reduced-order models are created in light of the aforementioned issues, and their use in simulating flight dynamics is further explored. The reduced order models include the model based on the autoregressive model with eXogenous inputs (ARX) method, multi-wavelet Volterra series method, back propagation (BP) neural network, radial basis function (RBF) neural network and support vector regression. Then the convergence and the generalization ability of the five models are compared, and a new model with high precision and strong generalization ability is proposed, which is a combination of ARX and neural network model. Taking a solar UAV model as an example, combining an aerodynamic reduced order model with a rigid-elastic coupling dynamic equation, the flight dynamics simulation model of elastic aircraft is established. Lastly, the computational fluid dynamics-computational structural dynamics (CFD-CSD) outputs and the outcomes of the simulation model based on linear aerodynamics are compared with the simulation results to determine how the UAV model responds to gusts. The results show that the performance of the elastic aircraft simulation model in aerodynamic prediction and gust response analysis is better than that of the model based on linear aerodynamics and is in good agreement with the CFD-CSD model, and the simulation efficiency is much higher than that of the CFD-CSD analysis.

Keywords: aerodynamic reduced order model; flight simulation; rigid elastic coupling dynamic equation; solar unmanned aerial vehicle; gust response

Received: 2021-09-02; **Accepted**: 2021-11-22; **Published Online**: 2021-12-30 07:02
URL: kns.cnki.net/kcms/detail/11.2625.V.20211229.1524.003.html
* **Corresponding author**. E-mail: wuzhigang@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0519

高速超声振动铣削钛合金实验研究

韩雄[1,2]，孙哲飞[1]，耿大喜[1,*]，张德远[1]

(1. 北京航空航天大学 机械工程及自动化学院，北京 100191；2. 成都飞机工业（集团）有限责任公司，成都 610091)

摘 要：针对钛合金在普通铣削（CM）时因切削速度低而面临的切削力大、薄壁工件变形大、加工效率低、刀具磨损严重等问题，采用高速超声振动铣削（HUVM）方法加工钛合金，实验研究其加工表面质量和切削力。从运动学角度出发对HUVM方法进行运动学分析。搭建包括超声振动系统、加工系统及测量系统在内的高速超声振动铣削实验平台。采用单因素实验对比CM和HUVM这2种方法对钛合金加工切削力和表面质量的影响规律。研究结果表明：与CM加工相比，HUVM加工可以使切削力降低32.6%～35.3%。并且HUVM加工表面粗糙度虽略有增加，但是表面结构可以更加均匀；此外，HUVM加工表面残余应力均为压应力，其绝对值随着每齿进给量和切削速度的增大而降低，而CM加工表面残余应力为拉应力。

关 键 词：超声振动；铣削；表面粗糙度；残余应力；切削力
中图分类号：V261.92
文献标志码：A **文章编号**：1001-5965(2023)07-1707-08

近年来，钛合金因具有质量轻、比强度好及耐腐蚀等良好性能，在航空、航天、兵器及船舶等工业生产中得到越来越广泛的应用[1]。然而，钛合金导热系数低、比热系数大，在机械加工过程中会产生较多的切削热，从而加剧刀具磨损，影响工件表面质量。在铣削加工钛合金时，切削速度通常低于80 m/min，这限制了钛合金的加工效率[2]。高速超声振动铣削（high-speed ultrasonic vibration milling, HUVM）方法是一种将电信号转换为振动信号并传输到刀具上，以实现高速下断续切削的新型铣削技术。这种加工技术提高了传统铣削加工时的速度限制，可以降低切削力和切削温度，减少刀具磨损。

国内外科研人员对超声振动切削钛合金进行了大量的研究[3-7]。Cakir等[8]基于有限元法进行超声振动切削Ti6Al4V模拟实验，研究结果表明：超声振动辅助切削技术可以有效降低切削力和切削温度，并且使加工工件表面有着较好的表面粗糙度。张翔宇等[9]在刀具进给方向上施加超声振动，与普通切削加工相比，切削力有所降低，切削效率得到提高。Zheng等[10]研究了超声振动铣削表面残余应力的变化规律，与普通铣削(conventional milling, CM)相比，超声振动铣削加工工件表面残余压应力较高。武民[11]研究了振幅、频率及冷却润滑方式对钛合金铣削表面粗糙度的影响规律。研究结果表明，采用较小的振幅、微量润滑可以获得较小的加工表面粗糙度。李世永[12]通过有限元仿真的方法研究了超声扭转振动铣削钛合金时振动频率、振幅、切削速度对加工表面残余应力的影响。超声振动频率增加，已加工表面的残余拉应力减小；超声振幅增加会导致表面残余拉应力减小并会产生向压应力转变的趋势；切削速度增大，表面残余拉应力增加。

尽管已经进行了较多关于超声振动铣削钛合金方面的研究，但目前为止大多数研究集中于较低

收稿日期：2021-09-16；录用日期：2021-11-05；网络出版时间：2022-01-18 20:42
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220118.1202.001.html
*通信作者. E-mail：gengdx@buaa.edu.cn

引用格式：韩雄，孙哲飞，耿大喜，等. 高速超声振动铣削钛合金实验研究[J]. 北京航空航天大学学报，2023，49（7）：1707-1714.
HAN X，SUN Z F，GENG D X，et al. Experiment research on high-speed ultrasonic vibration milling of titanium alloy[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（7）：1707-1714（in Chinese）.

切削速度范围内。因此，本文将高速超声振动技术引入钛合金铣削加工中，通过优化铣削参数来提高铣削加工质量和效率。首先从运动学角度出发分析 HUVM 断续切削原理。通过搭建 HUVM 实验平台，进行单因素实验对比研究了 CM 和 HUVM 这 2 种不同方法下每齿进给量和切削速度对加工表面质量和切削力的影响规律，从而为提高钛合金加工表面质量和加工效率提供新的工艺方法和工艺参数指导。

1 高速超声振动铣削运动学分析

HUVM 加工是在刀具上附加超声椭圆振动，使铣刀保持进给方向运动，并且附加椭圆旋转运动，从而实现铣刀对工件的分离断续切削。如图 1 所示，铣刀进给方向为 x 轴，刀具旋转轴线为 z 轴，建立直角坐标系。

刀尖运动过程可以分解为 3 部分，分别是绕主轴旋转运动、进给运动及椭圆旋转运动。在图 1 所示的直角坐标系中，建立刀尖轨迹运动方程：

$$\begin{cases} x=\dfrac{V_{\mathrm{f}}t}{60}+\dfrac{D}{2}\cos\omega_i t+A\cos\omega_z t\cos\alpha-B\sin\omega_z t\sin\alpha \\ y=\dfrac{D}{2}\cos\omega_i t+A\cos\omega_z t\sin\alpha+B\sin\omega_z t\cos\alpha \\ \mathrm{z}=0 \\ \omega_z=2\pi f \\ \alpha=\omega_i t \end{cases} \tag{1}$$

式中：V_{f} 为进给速度；D 为铣刀直径；A 和 B 分别为超声振动长轴和短轴振幅；ω_i 为主轴旋转角速度；α 为刀具转过的角度；f 为超声振动频率。

第 m 个刀尖运动轨迹方程为

$$\begin{cases} x=\dfrac{V_{\mathrm{f}}t}{60}+\dfrac{D}{2}\cos\left(\omega_i t-\dfrac{\pi(m-1)}{2}\right)+A\cos\left(\omega_z t-\dfrac{\pi(m-1)}{2}\right)\cos\alpha-B\sin\left(\omega_z t-\dfrac{\pi(m-1)}{2}\right)\sin\alpha \\ y=\dfrac{D}{2}\sin\left(\omega_i t-\dfrac{\pi(m-1)}{2}\right)+A\cos\left(\omega_z t-\dfrac{\pi(m-1)}{2}\right)\sin\alpha+B\sin\left(\omega_z t-\dfrac{\pi(m-1)}{2}\right)\cos\alpha \\ \mathrm{z}=0 \\ \omega_z=2\pi f \\ \alpha=\omega_i t \\ m=1,2,3,4 \end{cases} \tag{2}$$

当 D=12 mm，A=15 μm，B=10 μm，V_{f}=60 m/min，f=17 880 Hz，f_z=0.015 mm/齿（f_z 为每齿进给量）时，通过 MATLAB 软件获得的 HUVM 刀尖运动轨迹如图 2 所示。HUVM 刀尖运动轨迹不再是单一的等距螺旋曲线，而是“波动式”螺旋切削，切削刃与工件表面不断重复进行接触、切削和分离。在局部放大图中可以看出，铣刀 4 个刀尖运动轨迹相互交叠，钛合金材料被断续切削加工。

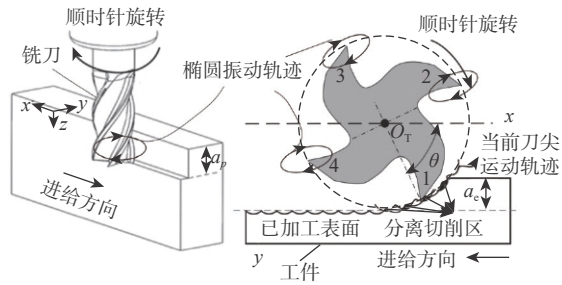


图 1 HUVM 加工过程示意图

Fig. 1 Schematic of HUVM process

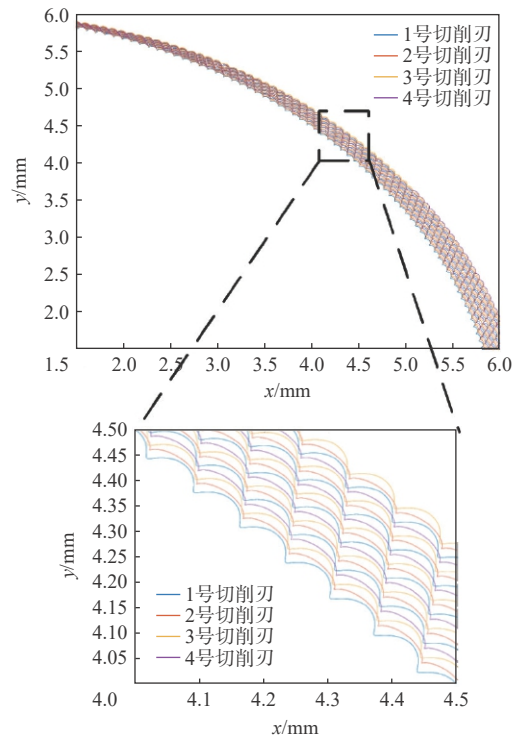


图 2 铣刀 4 个刀齿运动轨迹

Fig. 2 Motion trace of four cutter teeth

2 实验过程

为了研究 HUVM 加工方法对钛合金铣削加工的影响，搭建如图 3 所示的实验装置。实验装置主要包含 3 部分：超声振动系统、立式加工系统和测量系统。实验所用的钛合金尺寸为 200 mm×120 mm×60 mm，钛合金化学成分如表 1 所示[13]。本次实验不使用切削液。铣刀直径为 12 mm，是标准的四刃硬质合金立铣刀。

实验中，对切削力、表面粗糙度和表面残余应力等进行检测。采用 Wyko NT9300 型白光干涉仪

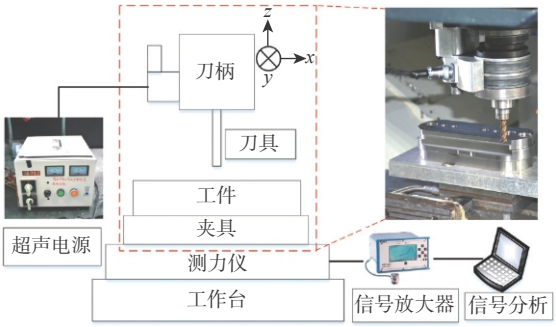


图 3　实验装置

Fig. 3　Experimental setup

表 1　钛合金化学成分[13]

Table 1　Chemical composition of titanium alloy[13]

元素	Al	V	Fe	C	N	H	O	Ti
占比/%	5.5～6.75	3.5～4.5	<0.25	<0.08	<0.05	<0.01	<0.2	余量

对表面微观形貌进行观察, 同时每个试样取 3 个不同位置测量表面粗糙度并取平均值。使用 Dino-Lite Pro 手持式显微镜对加工表面放大 200 倍后的表面形貌进行观察。采用 Xstress 3 000 型 X 射线应力分析仪测量工件表面的残余应力, 每个试样选取 3 个不同位置进行测量并取平均值, 在测量前需要进行校准来控制误差, 管电压采用 30 kV, 管电流采用 25 mA, 靶材为 Cu 靶, Kα 辐射 Ni 滤波片。切削力由 Kistler 9257B 测力仪采集, 采样频率为 5 kHz, 测量范围±15 kN, 精度为 0.01 N。

为了研究每齿进给量和切削速度对加工钛合金时表面质量与切削力的影响规律, 设置单因素实验, 实验参数设置如表 2 所示。首先, 保持其他参数不变, 切削速度为 80 m/min,研究每齿进给量对表面质量和切削力的影响规律, 设置每齿进给量分别为 0.015, 0.03, 0.045, 0.06, 0.075, 0.09, 0.105, 0.12 , 0.135 mm/齿。然后, 研究切削速度对表面质量和切削力的影响, 每齿进给量为 0.015 mm/齿, 设置切削速度分别为 100, 120, 140, 160 m/min。每组参数均进行 CM 和 HUVM 对比实验。超声振动频率为 17 880 Hz, 超声振动幅值分别为 15.8 μm 和 13.6 μm, 在实验设计调试换能器阶段, 已经对换能器在特定负载下的频率与振幅进行实时采集, 保证在实验最大载荷内, 所设计的换能器的谐振基本是稳定的。此外, 实验中所使用的非接触式感应供电超声振动系统具有实时调频功能模块, 该模块可以实时检测振动系统的频率漂移, 并根据谐振状态适时调整超声频率, 以保证振幅最大量输出。实验中轴向力与扭矩较小, 因此, 可以判断换能器谐振在加工过程中是基本稳定的。

表 2　实验参数设置

Table 2　Experimental parameter settings

参数	数值
每齿进给量f_z/(mm·齿$^{-1}$)	0.015,0.03,0.045,0.06,0.075,0.09, 0.105,0.12,0.135
切削速度v/(m·min^{-1})	80,100,120,140,160
径向切宽/mm	0.2
轴向切深/mm	8
铣削类型	CM, HUVM
振动频率/Hz	17 880
振动幅值/μm	A=15.8, B=13.6

3　结果与讨论

3.1　切削力

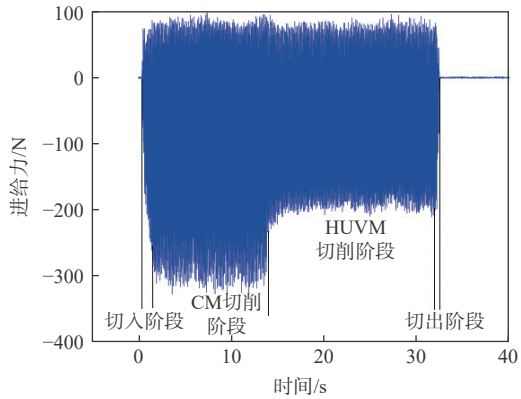
图 4 为测力仪测得的三向切削力信号, 切削速度 v_f=120 m/min、每齿进给量 f_z=0.015 mm/齿、轴向切深为 8 mm、径向切宽为 2 mm。

从图 4 可以看出, 进给力和切向力主要位于负半轴上, 这意味着测力仪获得的数值与测力仪默认正方向相反, 因此, 在随后的讨论中只考虑切削力绝对值。在切削加工过程切削力经历了 4 个阶段的变化: 切入阶段、CM 稳定阶段、HUVM 稳定阶段和切出阶段。在切入阶段, 铣刀侧刃与工件快速接触, 3 个方向的切削力数值从 0 增加到某值后, 达到 CM 稳定阶段。之后, 打开超声电源, 使切削加工进入 HUVM 过程, 这时切削力产生了明显变化。刀具沿着进给方向继续运动, 并逐渐与工件分离, 切削力迅速降至 0, 这是切出阶段。在钛合金加工过程中, 超声椭圆振动辅助加工对铣削力有明显的影响, 进给力下降约 35%, 切向力下降约 25%, 轴向力最小, 下降约 22%。

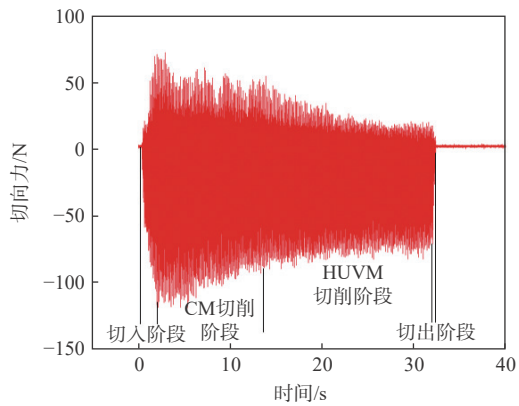
为了研究切削速度对铣削钛合金时切削力的影响, 保持每齿进给量为 0.015 mm/齿不变, 设置切削速度不同的 4 组参数, 切削速度分别为 100 , 120, 140, 160 m/min。在稳定铣削阶段平均切削力随切削速度的变化如图 5 所示。

从图 5 中可以看出, 无论是 CM 加工还是 HUVM 加工, 随着切削速度的增大, 进给力和切向力先减小后增大。在相同的切削参数下, HUVM 加工可以显著降低切削力。对于 HUVM 加工而言, 当切削速度时 120 m/min 时, 切削力下降最为显著, 在进给力和切向力分别下降了 35.3% 和 35.2%。HUVM 加工的断续切削特性使切削力转变为脉冲力, 刀具与工件是一种周期性接触、分离的加工模式, 这种分离机制有效降低了平均切削力。

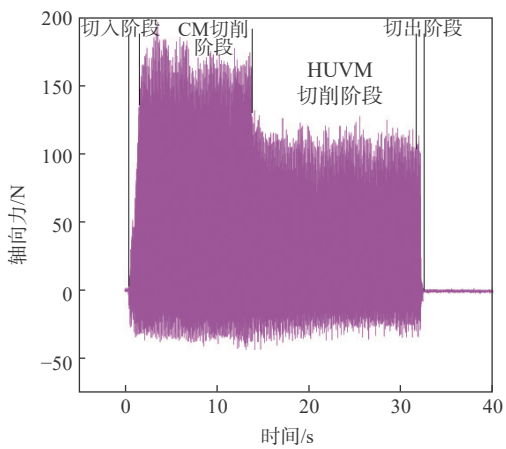
为了研究每齿进给量对切削力的影响, 设置每齿进给量分别为 0.015, 0.03, 0.045, 0.06, 0.075 , 0.09,



(a) 进给力



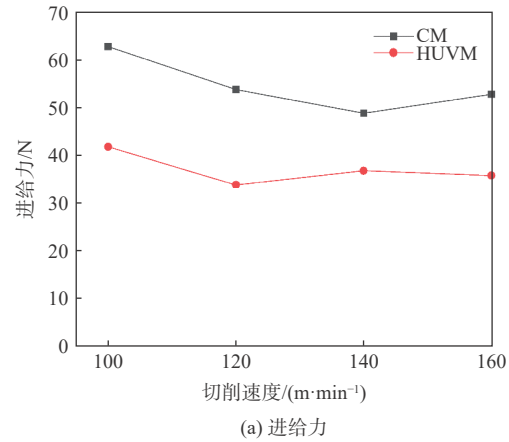
(b) 切向力



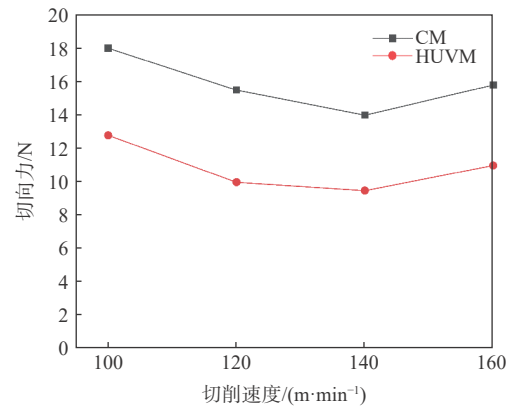
(c) 轴向力

图 4 三向切削力信号

Fig. 4 Three-way cutting force signal



(a) 进给力



(b) 切向力

图 5 切削速度与平均切削力关系

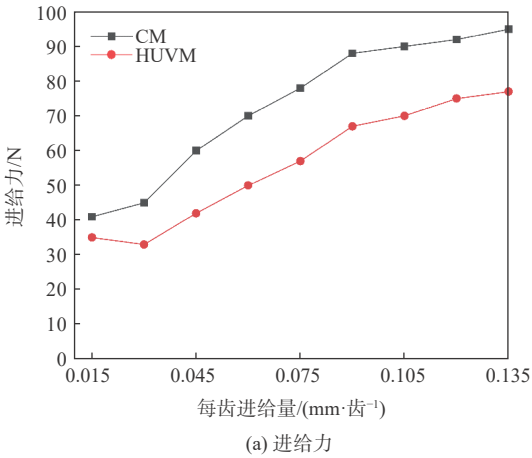
Fig. 5 Relationship between average cutting speed, feed force, and tangential force

0.105, 0.12, 0.135 mm/齿。切削速度为 80 m/min, 轴向切深为 8 mm, 径向切宽为 2 mm, 分别计算稳定切削阶段 CM 加工和 HUVM 加工的平均切削力, 切削力平均值随每齿进给量的变化如图 6 所示。

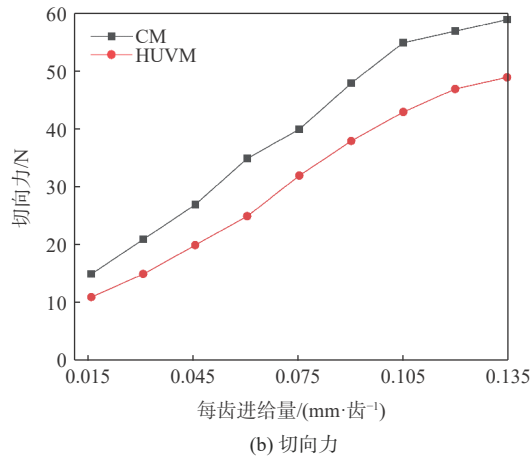
从图 6 可以看出, 在 0.015～0.135 mm/齿范围内, CM 加工和 HUVM 加工的平均切削力随着每齿进给量的增加而增加。当每齿进给量增加时, 切削量增加导致单位时间的材料去除量增大, 切削力增加。施加超声振动后, 切削力下降。当每齿进给量为 0.045 mm/齿时, 进给力下降 28.7%, 当每齿进给量为 0.045 mm/齿时切向力下降 32.6%。HUVM 加工降低切削力的主要原因是该工艺方法实现了高速切削下的切削刃-工件高频分离切削模式, 将传统铣削中长切削弧高频打断, 从微观而言切削刃受到的切削力为脉冲形式切削力, 因此, 可以降低切削过程中的平均切削力。此外, HUVM 中分离切削模式产生的切屑体积更小, 减小了切屑与前刀面的摩擦力, 因而也降低了平均切削力。

3.2 表面粗糙度

图 7 为保持切削速度为 80 m/min 不变, 研究 CM 加工和 HUVM 加工每齿进给量对工件表面粗糙度的影响规律, 设置每齿进给量分别为 0.015, 0.03, 0.045, 0.06, 0.075, 0.09, 0.105, 0.12, 0.135 mm/齿。可以看出, 无论是 CM 还是 HUVM, 表面粗糙度均随着每齿进给量的增加先增大后减小, 当每齿进给量在 0.015～0.105 mm/齿时, 表面粗糙度随着每齿进给量增加而逐渐增大。当每齿进给量在 0.105～0.135 mm/齿时, 表面粗糙度随着每齿进给量增加而逐渐减小。一方面, 随着每齿进给量的增加, 加工零件的表面粗糙度逐渐增加, 这是因为当每齿进给量增加时, 铣削厚度增加, 加工表面相邻波峰和峰



(a) 进给力



(b) 切向力

图 6　平均切削力与每齿进给量关系

Fig. 6　Relationship between average feed force, tangential force and feed rate per tooth

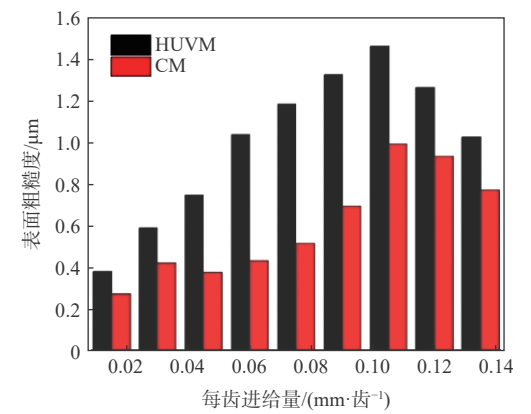


图 7　表面粗糙度与每齿进给量关系

Fig. 7　Relationship between surface roughness and feed rate per tooth

谷间距和形态差异增大, 铣刀在工件表面残留高度增加, 导致表面粗糙度增加。随着每齿进给量的进一步增大, 铣削力进一步增大, 铣削温度提高, 对金属材料表面有热软化作用, 表面粗糙度有所降低。

当每齿进给量在 0.105 mm/齿时, CM 加工表面粗糙度为 0.97 μm, HUVM 加工表面粗糙度达到最大值 1.5 μm, 远高于 CM 加工表面粗糙度。这是因为施加超声振动之后, 切削刃的运动轨迹是波动式轨迹, 在前后切削刃的交叠下导致波动形的刀痕留在加工表面上, 导致表面粗糙度增加[14]。

图 8 为使用工具显微镜观察的工件加工表面宏观形貌, 可以看出, HUVM 加工的表面形貌不同于 CM 加工表面形貌。CM 加工的表面形貌因为光滑所以更亮一些, 而 HUVM 加工的表面形貌存在微刀痕所以相对较暗。通过对表面形貌放大可以看出, CM 加工表面有明显的刀齿痕迹, 而 HUVM 加工表面更加的均匀精细, 刀痕方向与刀具螺旋角一致, 且刀痕纹理间距与每齿进给量大小大致相等。

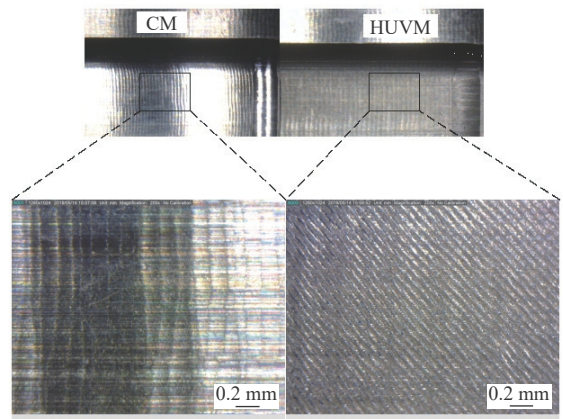


图 8　加工表面宏观形貌

Fig. 8　Macro-morphology of machined surface

图 9 为工件加工表面微观形貌, CM 加工表面相对于 HUVM 加工表面更加光滑, 而 HUVM 加工表面有更多的峰谷, 但峰谷分布相对均匀, 间距较小。这是因为 HUVM加工的分离断续切削特性使材料去除量更小, 有效削弱了加工时振动现象, 从而使表面更精细, 结构更均匀。

为了研究切削速度对表面粗糙度的影响规律, 保持每齿进给量为 0.015 mm/齿, 分别设置切削速度为 100, 120, 140, 160 m/min。图 10 为表面粗糙度和切削速度关系, HUVM 加工表面粗糙度要高于 CM 加工表面粗糙度。当切削速度达到 140 m/min 时, CM 加工表面粗糙度达到最大值。HUVM 的表面粗糙度随切削速度的变化维持在 1.4 μm 左右, 由此可见, HUVM 比 CM 加工表面粗糙度更稳定一些。

3.3　残余应力

切削过程中产生的残余应力与切削力、热载荷和材料内部微观结构有很大关系[15-16]。在进给方向上, 加工表面残余应力与每齿进给量之间的关系如图 11 所示, 加工表面残余应力与切削速度之间的关系如图 12 所示。

从图 11 和图 12 中可以看出, CM 加工表面残

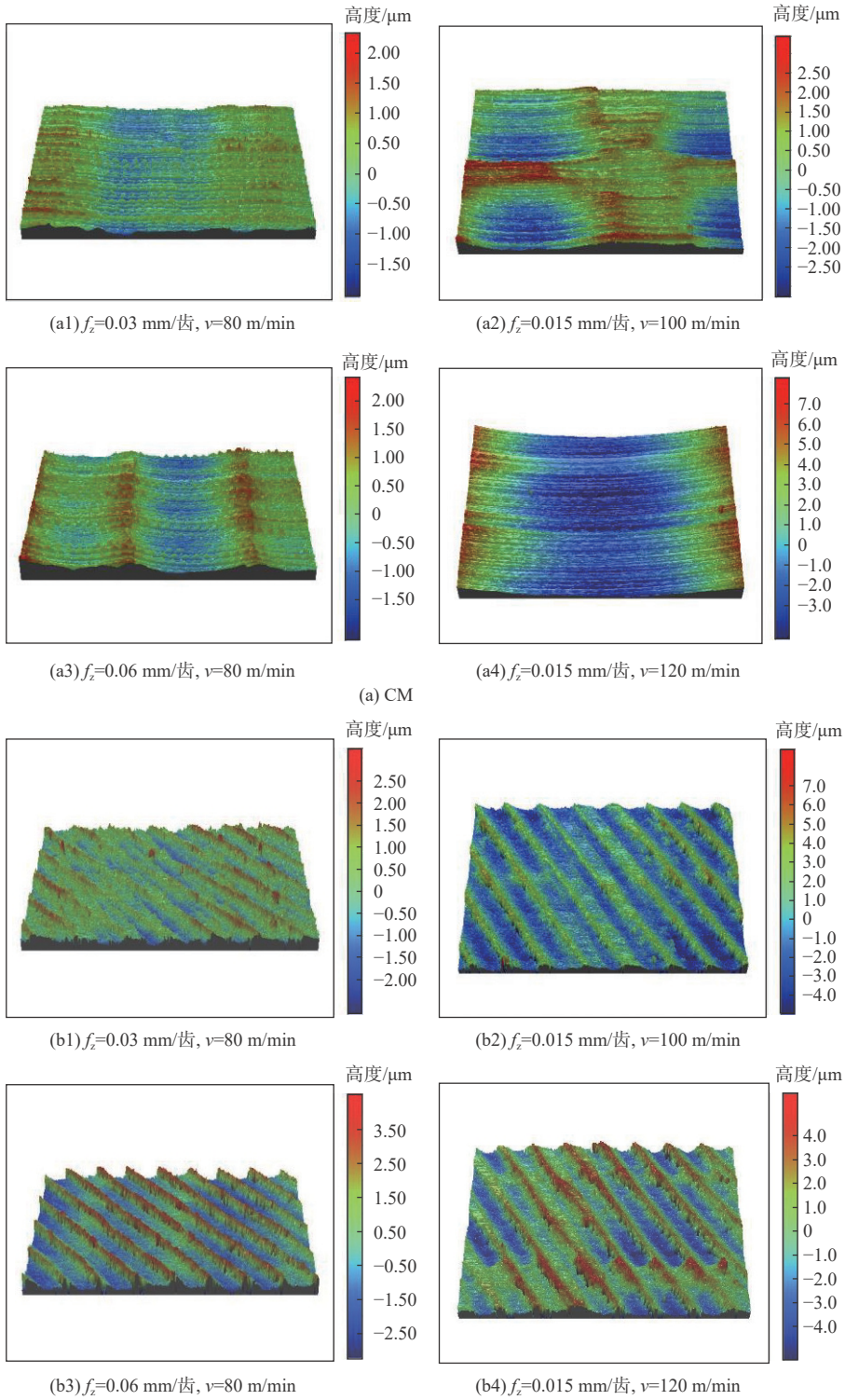


图 9 加工表面微观形貌

Fig. 9 Micro-morphology of machined surface

余应力均为拉应力，而 HUVM 加工表面残余应力均为压应力，并且 HUVM 加工表面残余压应力随着每齿进给量和切削速度的增加而减小。

加工参数的改变通常会导致工件表面机械应力载荷和热载荷分布的变化，最终导致残余应力变化。在 CM 加工中，每齿进给量和切削速度的增加导致切削区域产生热量增多，工件表面温度升高，表层材料膨胀，当工件冷却时，表面材料不能自由收缩，因此，表面产生残余拉应力[17]。

在 HUVM 加工中，刀具不仅在垂直于加工表面方向上产生铣削力分量，还对加工表面产生更强烈的机械冲击效应。由于 HUVM 加工的分离断续

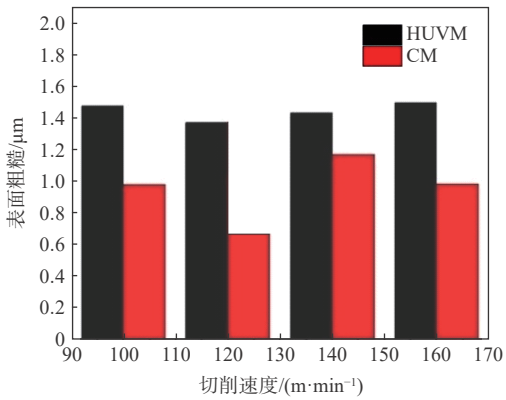


图 10　表面粗糙度和切削速度关系

Fig. 10　Relationship between surface roughness and cutting speed

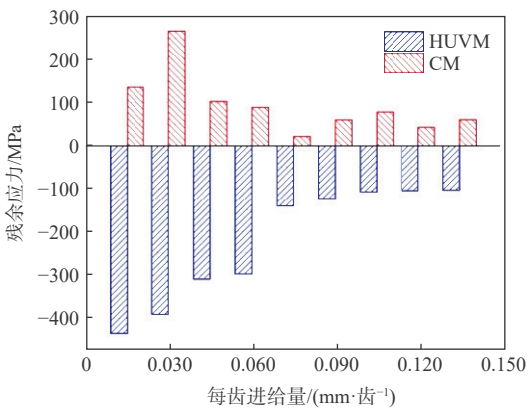


图 11　残余应力和每齿进给量关系（v_f=80 m/min）

Fig. 11　Relationship between residual stress and feed rate per tooth（v_f=80 m/min）

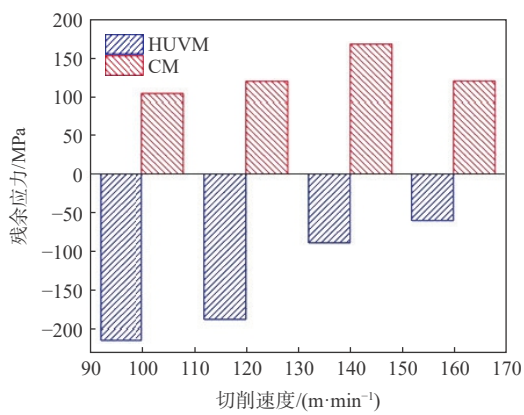


图 12　残余应力和切削速度关系（f_z=0.015 mm/齿）

Fig. 12　Relationship between residual stress and cutting speed（f_z=0.015 mm/齿）

切削特性，在切削时会产生更小的切屑，这些切屑可以快速带走产生的切削热，并延迟工件表面温度升高。此外，HUVM 加工表面上的微槽增加了工件与切削液的接触面积，有效地抑制了工件表面温度升高。因此，在这 2 个原因的共同作用下，导致 HUVM 加工表面产生残余压应力[14]。

4　结　论

1）HUVM 加工表面颜色较深，在加工表面形成更加精细均匀的波动刀痕结构，表面粗糙度略有增加。CM 加工表面较亮，表面粗糙度低，但表面存在不稳定的刀痕。

2）对 HUVM 加工而言，当切削速度增加时，表面粗糙度先增大后减小，当每齿进给量为 0.105 mm/齿时，HUVM 加工表面粗糙度最大，达到 1.5 μm。随着每齿进给量的增加，CM 加工表面粗糙度变化不大，稳定在 1.4 μm 左右。

3）在选定参数范围内，HUVM 加工表面残余应力均为压应力，残余压应力随每齿进给量和切削速度的增加而下降，而 CM 加工表面残余应力均为拉应力。

4）HUVM 加工可以显著降低切削力。随着切削速度的增加，HUVM 加工的平均切削力呈现先减小后增大的趋势。

5）超声振动系统通用性强，可以与现有数控机床直接对接。因此，可以认为 HUVM 是一种可实现钛合金高效高质低成本加工的新型加工方法。

参考文献（References）

[1] ZHANG M, ZHANG D, GUO H, et al. High-speed rotary ultrasonic elliptical milling of Ti-6Al-4V using high-pressure coolant[J]. Metals, 2020, 10(4): 500.

[2] LI A H, ZHAO J, LUO H B, et al. Progressive tool failure in high-speed dry milling of Ti-6Al-4V alloy with coated carbide tools[J]. The International Journal of Advanced Manufacturing Technology, 2012, 58(5): 465-478.

[3] WU Y, NIU J Z, FUJIMOTO M, et al. Fundamental machining characteristics of ultrasonic assisted turning of titanium alloy Ti-6Al-4V[J]. Advanced Materials Research, 2013, 797: 344-349.

[4] 路冬，蔡力钢，程强，等. 钛合金超声椭圆振动辅助车削实验研究[J]. 振动与冲击，2015, 34(6): 151-154.
LU D, CAI L G, CHENG Q, et al. Tests for ultrasonic elliptical vibration-assisted turning of titanium alloy[J]. Journal of Vibration and Shock, 2015, 34(6): 151-154(in Chinese).

[5] 张明亮，张德远，刘佳佳，等. 钛合金薄壁件高速超声椭圆振动铣削机理和试验[J]. 北京航空航天大学学报，2019, 45(8): 1606-1612.
ZHANG M L, ZHANG D Y, LIU J J, et al. Mechanism and experiment of high-speed ultrasonic elliptical vibration milling of thin-walled titanium alloy parts[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(8): 1606-1612(in Chinese).

[6] 王明海，李世永，郑耀辉. 超声铣削钛合金材料表面粗糙度研究[J]. 农业机械学报，2014, 45(6): 341-346.
WANG M H, LI S Y, ZHENG Y H. Surface roughness of titanium alloy under ultrasonic vibration milling[J]. Transactions of the Chinese Society for Agricultural Machinery, 2014, 45(6): 341-

346(in Chinese).

[7] TAO G C, MA C, SHEN X H, et al. Experimental and modeling study on cutting forces of feed direction ultrasonic vibration-assisted milling[J]. The International Journal of Advanced Manufacturing Technology, 2017, 90(1): 709-715.

[8] CAKIR F H, GURGEN S, SOFUOGLU M A, et al. Finite element modeling of ultrasonic assisted turning of Ti_6Al_4V alloy[J]. Procedia - Social and Behavioral Sciences, 2015, 195: 2839-2848.

[9] 张翔宇, 隋嵩, 张德远, 等. 高速超声振动切削钛合金可行性研究[J]. 机械工程学报, 2017, 53(19): 120-127.
ZHANG X Y, SUI H, ZHANG D Y, et al. Feasibility study of high-speed ultrasonic vibration cutting titanium alloy[J]. Journal of Mechanical Engineering, 2017, 53(19): 120-127(in Chinese).

[10] ZHENG K, LIAO W H, DONG Q, et al. Friction and wear on titanium alloy surface machined by ultrasonic vibration-assisted milling[J]. Journal of the Brazilian Society of Mechanical Sciences and Engineering, 2018, 40(9): 411.

[11] 武民. 不同振动方式下的钛合金振动辅助铣削工艺效果研究[D]. 新乡: 河南科技学院, 2018.
WU M. Study on the technological effect of vibration-assisted milling of titanium alloy under different vibration modes[D]. Xinxiang: Henan Institute of Science and Technology, 2018 (in Chinese).

[12] 李世永. 超声扭转振动辅助铣削钛合金的加工技术研究[D]. 沈阳: 沈阳航空航天大学, 2014.
LI S Y. Research on machining technology of ultrasonic torsional vibration assisted milling of titanium alloy[D]. Shenyang: Shenyang Aerospace University, 2014 (in Chinese).

[13] ZHU Z J, SUN J, LI J F, et al. Investigation on the influence of tool wear upon chip morphology in end milling titanium alloy Ti_6Al_4V[J]. The International Journal of Advanced Manufacturing Technology, 2016, 83(9): 1477-1485.

[14] LIU J J, JIANG X G, HAN X, et al. Effects of rotary ultrasonic elliptical machining for side milling on the surface integrity of Ti-6Al-4V[J]. The International Journal of Advanced Manufacturing Technology, 2019, 101(5): 1451-1465.

[15] CHEN G, REN C Z, ZOU Y H, et al. Mechanism for material removal in ultrasonic vibration helical milling of Ti_6Al_4V alloy[J]. International Journal of Machine Tools and Manufacture, 2019, 138: 1-13.

[16] 于浩, 赵军, 盖少磊, 等. 42CrMo钢车削表面完整性研究[J]. 组合机床与自动化加工技术, 2021(7): 137-140.
YU H, ZHAO J, GAI S L, et al. Research on surface integrity in turning of 42CrMo steel[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2021(7): 137-140(in Chinese).

[17] 陈建岭, 李剑峰, 孙杰, 等. 钛合金铣削加工表面残余应力研究[J]. 机械强度, 2010, 32(1): 53-57.
CHEN J L, LI J F, SUN J, et al. Surface residual stress of titanium alloy induced by milling[J]. Journal of Mechanical Strength, 2010, 32(1): 53-57(in Chinese).

Experiment research on high-speed ultrasonic vibration milling of titanium alloy

HAN Xiong[1,2], SUN Zhefei[1], GENG Daxi[1,*], ZHANG Deyuan[1]

(1. School of Mechanical Engineering and Automation, Beihang University, Beijing 100191, China;
2. Chengdu Aircraft Industrial (Group) Co., Ltd., Chengdu 610091, China)

Abstract: Conventional milling (CM) causes problems for titanium alloys, such as high cutting forces, large deformation of thin-walled workpieces, low processing efficiency, and severe tool wear due to low cutting speeds. To address these problems, a new method of high-speed ultrasonic vibration milling (HUVM) is adopted to machine titanium alloys. Their surface quality and the cutting force are also examined through experiments. The kinematic equation of HUVM is constructed. An HUVM experimental platform including ultrasonic vibration system, machining system and measurement system is built. Single factor experiments are carried out to explore the effects of CM and HUVM on the surface quality and cutting force. The experimental results show that compared with CM, HUVM could generate more uniform surface structure with increased surface roughness. Unlike CM with residual tensile stress of the machined surfaces, HUVM has residual compressive stress with its value decreasing with the increase of feed per tooth and cutting speed. Moreover, HUVM could reduce cutting force by 32.6%～35.3% compared to CM.

Keywords: ultrasonic vibration; milling; surface roughness; residual stress; cutting force

Received: 2021-09-16; **Accepted**: 2021-11-05; **Published Online**: 2022-01-18 20:42
URL: kns.cnki.net/kcms/detail/11.2625.V.20220118.1202.001.html
* **Corresponding author**. E-mail: gengdx@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0530

光纤环多极对称绕法对 Shupe 误差抑制效果仿真分析

张东波[1,2]，汪立新[1,*]，李灿[1]

(1. 火箭军工程大学 导弹工程学院，西安 710025；
2. 中国航天科技集团公司第九研究院第十六研究所 光纤传感事业部，西安 710100)

摘 要：Shupe 误差是高精度光纤陀螺（IFOG）工程化过程中的最大瓶颈问题之一。建立一个精确到匝的光纤环有限元模型，基于此模型，并在不同温度激励下分析八极、十六极和三十二极对称绕法的温度性能。仿真结果表明：相对于八极对称绕法，采用十六极对称绕法和三十二极对称绕法的光纤陀螺能够有效抑制热致陀螺漂移，且十六极对称绕法的抑制效果最好。这对高精度 IFOG 的光纤环绕制方法的选择提供了指导。

关 键 词：Shupe 误差；十六极对称绕法；光纤环；有限元模型；热致陀螺漂移
中图分类号：V279; TN253
文献标志码：A **文章编号**：1001-5965(2023)07-1715-07

从 1976 年第 1 台光纤陀螺样机诞生至今，经过几十年的发展，干涉式光纤陀螺仪原理早已成熟[1]。随着精度的提升，干涉式光纤陀螺仪（interferometric fiber optic gyro, IFOG）以其高可靠性、长寿命、大动态范围被广泛应用于导航、航海、稳定平台等方面[2]。干涉式光纤陀螺的基本原理是通过测量角运动引起的 Sagnac 相位差获得输入角速度。由于光纤环中每一点的折射率都会随外界温度变化，从而导致 2 束光经过光纤环后会产生 1 个由温度变化引起的非互异性相移，即 Shupe 误差[3]，该问题是光纤陀螺工程化过程中的最大瓶颈问题之一。对此，国内外研究人员进行了大量研究工作，Frigo[4] 发现采用特殊的光纤环绕制方法可以减小这种非互异性误差，Mohr[5] 验证了四极对称绕法 (quadrupolar, QAD) 抑制 Shupe 误差的有效性。因此，研究者们提出更多光纤环对称绕制方式，如八极对称绕法[6]、十六极对称绕法、crossover-free 绕法[7-8] 等，此外，还有学者提出随机绕法实现统计学上的对称[9]，但是由于绕制技术等限制，该方法实现难度较大，此外，在光纤环安装结构优化[10]、光纤环绕制工艺及固化胶优化[11]、光纤陀螺温度控制[12]、光纤陀螺温度补偿[13] 等方面也进行了一定的研究。

关于不同的对称绕法对 Shupe 误差的抑制效果，李绪友等[14] 通过仿真证明了相对于四极对称绕法和八极对称绕法，十六极对称绕法能更加有效地抑制 Shupe 误差，提高光纤环温度性能。由于高精度光纤环绕制所用光纤相对较长，这使三十二极对称绕法及其他更高的绕法成为可能。为研究三十二极对称绕法对光纤环温度性能的抑制效果，本文采用有限元方法建立了光纤环模型，以光纤环中的每匝光纤为单位，对采用三十二极对称绕法的光纤环施加不同的温度激励，仿真分析了其对 Shupe 误差的抑制效果，并与八极和十六极对称绕法的光纤环进行了仿真与比较。

收稿日期：2021-09-06；录用日期：2021-09-17；网络出版时间：2021-11-30 16:29
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20211129.1902.006.html
基金项目：陕西省自然科学基础研究计划 (2020JQ-491)
*通信作者. E-mail：wlxxian@sina.com

引用格式：张东波，汪立新，李灿. 光纤环多极对称绕法对 Shupe 误差抑制效果仿真分析 [J]. 北京航空航天大学学报，2023，49（7）：1715-1721. ZHANG D B，WANG L X，LI C. Simulation analysis of reduction effect of symmetrical winding method for multi-polar fiber ring on Shupe error [J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（7）：1715-1721（in Chinese）.

1 光纤环温度场数学模型

当光纤陀螺光纤环线圈中的一段光纤l存在温度扰动时，沿l传播引起的相移ϕ可表示为[4]

$$\phi(l)=\beta_0 nl+\beta_0\left(\frac{\partial n}{\partial T}\Delta T+n\alpha\Delta T\right)l \tag{1}$$

式中：$\beta_0=2\pi/\lambda_0$为真空中的传播常数，λ_0为光源中心波长；n为光纤的有效折射率；$\partial n/\partial T$为折射率温度系数；α为折射率膨胀系数；ΔT为光纤的温度变化量。把光纤长度扩展到整个光纤环，结合Sagnac效应，总的热致相移随温度的变化可表示为

$$\varphi(t)=\beta_0\int_0^L\left[\left(\frac{\partial n}{\partial T}+n\alpha\right)\dot{T}_{(l,t)}\right](L-2l)\,\mathrm{d}l \tag{2}$$

式中：L为光纤环中光纤的总长度；$\dot{T}_{(l,t)}$为位置l处的光纤温变速率，通常$\partial n/\partial T$大小为$10^{-5}/℃$，$n\alpha$比$\partial n/\partial T$至少小一个数量级，所以$n\alpha$可以忽略[15]。根据Sagnac效应，在已绕制完成的脱骨架光纤环中，Sagnac相位差与光纤环的长度L、光纤环直径D和被测光纤环绕其轴的旋转速率Ω有关，关系如下：

$$\phi(t)=\frac{2\pi LD}{\lambda_0 c}\Omega \tag{3}$$

式中：c为光在真空中的速度。

使Sagnac相移与热致相移相等，可得热瞬变过程引起的近似等效旋转误差速率：

$$\Omega_E(t)=\frac{n}{DL}\int_0^L\left[\left(\frac{\partial n}{\partial T}+n\alpha\right)\dot{T}_{(l,t)}\right](L-2l)\,\mathrm{d}l \tag{4}$$

为了计算总的热致误差速率$\Omega_E(t)$，需要对光纤环进行离散化。为了提高运算效率，假设光纤环上每一匝之内各层光纤环的温度相等，将光纤环从三维模型简化为二维模型，总长度L的光纤被分为M层，每层N匝，总匝数MN，如图1所示。则式(4)的数字化表达式可以表示为

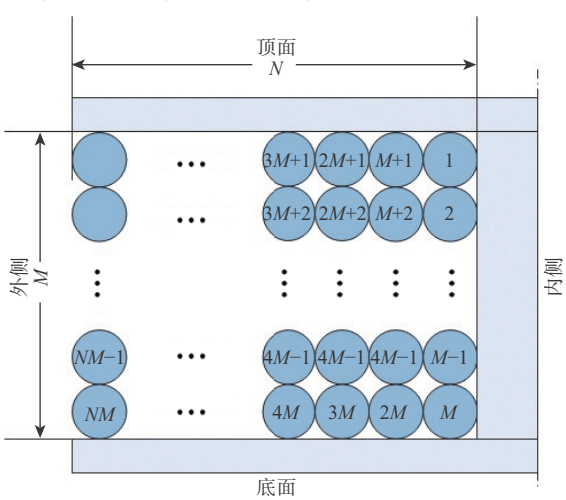


图1 光纤环绕制匝数示意图

Fig. 1 Diagram of number of fiber coil turns

$$\Omega_E(t)=\frac{n}{DL}\sum_{i=1}^{MN}\left[\left(\frac{\partial n}{\partial T}+n\alpha\right)\dot{T}(l_i,t)\right](L-2l_i)\,\mathrm{d}l_i \tag{5}$$

式中：l_i为第i匝光纤的起始点坐标长度；$\mathrm{d}l_i$为第i匝光纤的绕制长度；$\dot{T}(l_i,t)$为第i匝光纤处的温变速率。

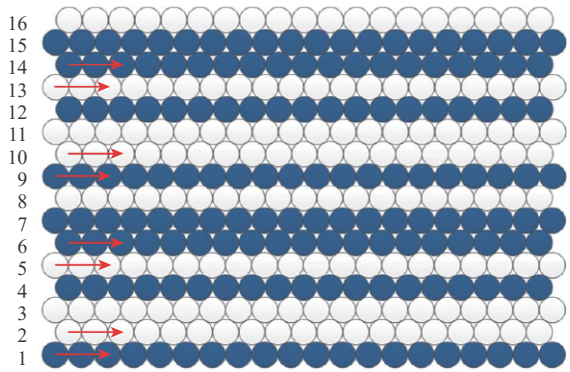
2 光纤环模型与参数设置

对称绕法是由光纤的中点开始从环内侧绕制，保证了光纤环中各点相对于光纤环中点的对称性。为了研究和比较三十二极对称绕法对光纤环温度性能的抑制效果，选择了八极对称绕法、十六极对称绕法及三十二极对称绕法3种光纤环绕制方法进行仿真分析。图2为3种对称绕法的绕制示意图[16]，对于八极对称绕法其1～4层与四极对称绕法的绕制顺序相同，5～8层与四极对称绕法的绕制顺序相反，即“正四反四”，完成八极对称绕法，十六极对称绕法1～8层与八极绕制顺序相同，9～16层与八极绕制顺序相反，即“正八反八”，同理，三十二极对称绕法为“正十六反十六”。图2中的箭头表示每匝的绕制方向，未标箭头层与箭头层绕制方向相反。

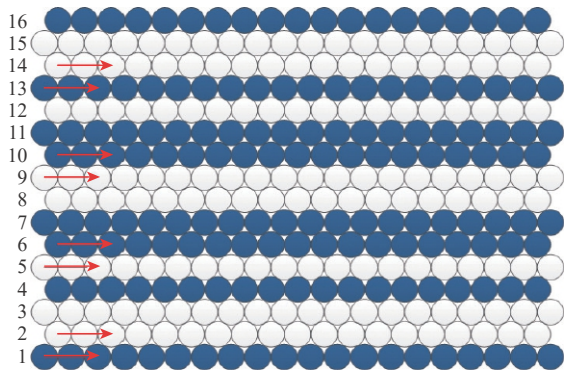
为了提高计算效率并且保证三十二极对称绕法能够完整绕制，在建模时建立32层20匝的光纤环模型进行研究分析，如图3所示，其参数如表1所示。

在实际中，光纤环的底面与光纤环骨架黏接，顶面、内侧面、外侧面与空气接触，为了便于分析和计算，仿真时简化为光纤环四周都与空气接触，表面传热系数为5 W/(K·m^2)。另外，光纤环中有光纤纤芯、包层和固化胶3种不同的材料，其各种热属性参数都会随着温度变化而变化，这会导致模型十分复杂，无法在减小网格划分的情况下保证计算精度，而且仿真模型计算量巨大，因此，需要对计算模型和材料属性参数进行简化。考虑到光纤环内部的材料在温度激励试验中部分热属性参数变化很小，因此，除了3种材料的主要热属性参数如表2所示，其余参数均设为常值。

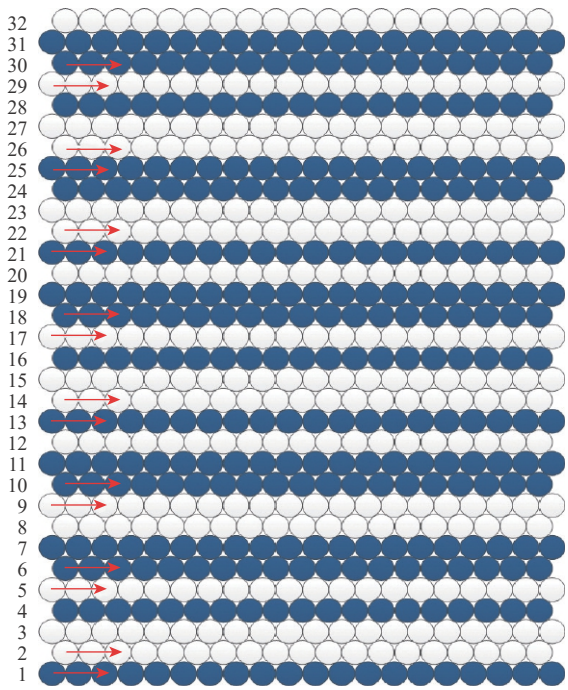
陀螺仿真选用变温条件，仿真变温曲线如图4所示。设定温箱的初始温度为20℃，保温20 min，随后以1℃/min的温变速率降温至−40℃，保温120 min，随后以1℃/min的温变速率升温至60℃，保温120 min，试验结束，试验时间总计420 min。



(a) 八极对称绕法



(b) 十六极对称绕法



(c) 三十二极对称绕法

图 2　不同对称绕法示意图

Fig. 2　Diagram of different fiber coil winding methods

仿真过程中对光纤环采用四周均匀激励，光纤环轴向单侧激励，光纤环径向单侧激励 3 种温度激励方式。

光纤环有限元模型和网格划分局部放大如图 5 所示。

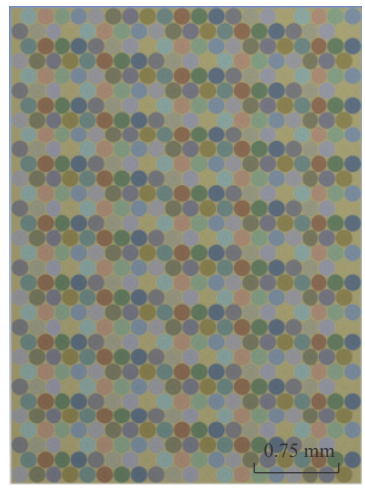


图 3　32 层 20 匝光纤环完整模型

Fig. 3　Complete model of optical fiber coil with 32 layers and 20 turns

表 1　仿真采用的光纤环结构参数

Table 1　Fiber coil structure parameter for simulation

参数	数值
层数	32
匝数	20
包层直径/m	0.000 135
纤芯直径/m	0.000 08
光纤环直径/m	0.112 8
光纤总长/m	230.72

表 2　材料参数

Table 2　Material parameters

材料	密度ρ/(kg·m^{-3})	比热c/(J·(kg·K)$^{-1}$)	导热率λ/(W·(K·m^2)$^{-1}$)
纤芯	2 203	1 990	1.38
包层	1 900	1 400	0.21
胶	970	1 600	0.21

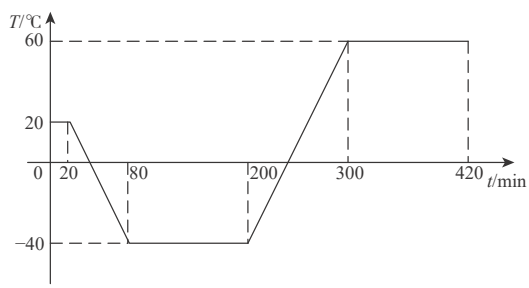
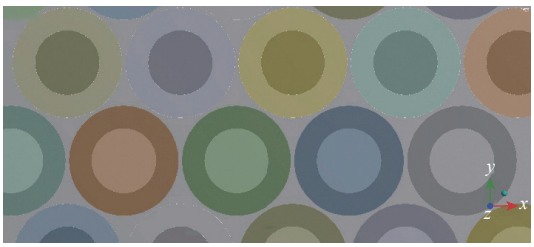
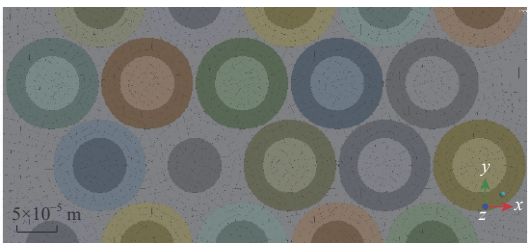


图 4　仿真变温曲线

Fig. 4　Temperature curve of simulation



(a) 光纤环有限元模型放大图



(b) 光纤环有限元模型网格划分

图 5 光纤环模型和网格划分局部放大

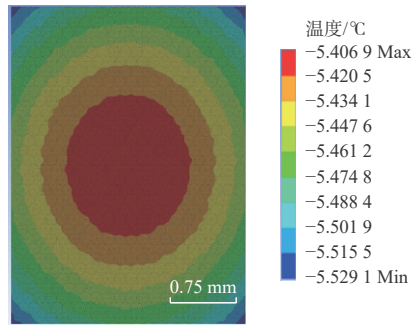
Fig. 5 Magnification and grid division of optical fiber coil model

3 仿真结果与分析

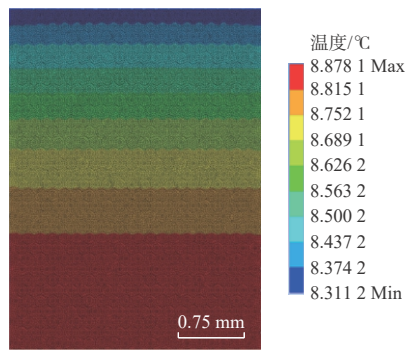
仿真算例每 30 s 计算 1 次数据，对计算数据每 10 个数据进行 1 次平滑处理。不同温度激励方式下光纤环温度分布的仿真结果如图 6 所示。将提取到的各光纤截面的温度数据和不同对称绕法中各匝光纤的位置代入离散化的热致陀螺漂移表达式 (5) 中，可以得到不同绕制方法下的热致陀螺漂移。

在四周均匀激励方式下，可以看到不同对称绕法的热值陀螺漂移曲线如图 7 所示。可以看到，当温度变化率不为 0 时，无论何种绕法，陀螺输出都会出现不同程度的漂移，其中八极对称绕法在温度变化率变化时，陀螺漂移会出现大于$\pm5\times10^{-3}$ (°)/h 的陀螺漂移峰值，对于其他 2 种绕法，则无较大变化，其中十六极对称绕法的漂移程度最小。这说明对于四周温度激励方式，十六极对称绕法能够很好地抑制陀螺热致漂移。

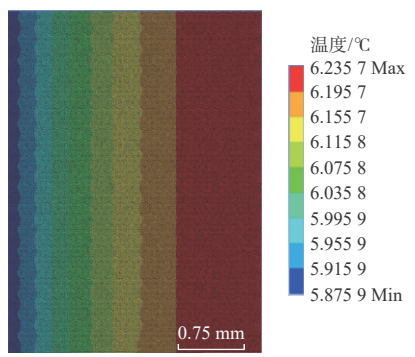
当光纤环的散热器件在图 1 所示的光纤环内侧时，会产生一个沿光纤环径向变化的温度梯度，为了研究这种情况下光纤环的不同对称绕法对热致陀螺漂移的影响，在光纤环径向单侧施加温度激励进行仿真(见图 6(b))，得到不同绕制方法下陀螺的热致漂移曲线如图 8 所示。由图 8 可以看出在温度变化率变化时，采用八极对称绕法的陀螺漂移同样出现了接近$\pm5\times10^{-3}$ (°)/h 的陀螺漂移峰值，对于其他 2 种绕法，则无较大变化。但是与四周均匀激励方式下高温及低温保持阶段陀螺漂移为 0 不同的是，径向单侧激励情况下，无论是温度梯度恒定的变温阶段还是高温或低温保持阶段，各种对称绕法的陀螺漂移都会产生波动。且在高低温保持阶段，十六极对称绕法和三十二极对称绕法的陀螺



(a) 四周均匀激励



(b) 径向单侧激励



(c) 轴向单侧激励

图 6 不同温度激励方式下光纤环温度分布仿真结果

Fig. 6 Simulation results of temperature distribution of fiber coil under different temperature excitations

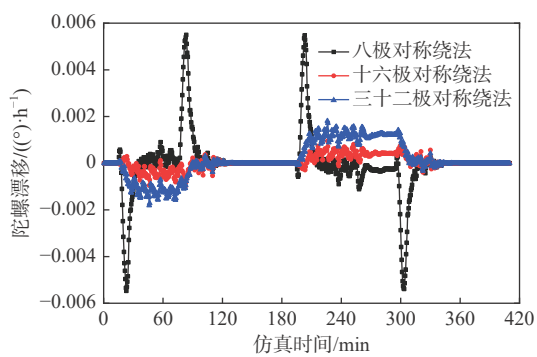


图 7 四周均匀激励下各对称绕法的陀螺热致漂移曲线

Fig. 7 Gyro drift error curve under uniform temperature excitation

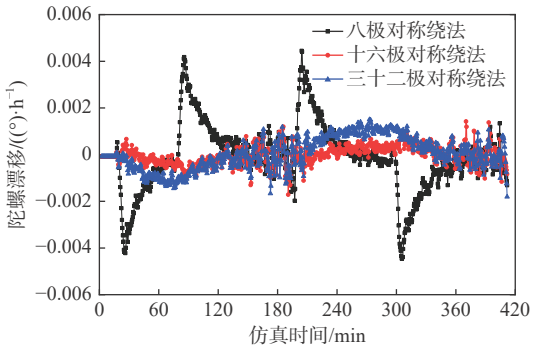


图 8　径向单侧激励下各对称绕法的陀螺热致漂移曲线

Fig. 8　Gyro drift error curve under radial temperature excitation

漂移波动有增大的趋势。对于整个激励过程而言，十六极对称绕法和三十二极对称绕法均有抑制径向热致陀螺漂移的效果，且十六极对称绕法的抑制能力更好。

此外，在光纤环轴向施加了温度激励，对各种对称绕法抑制轴向温变引起的陀螺漂移效果进行仿真分析（见图 6(c)），热致陀螺漂移曲线如图 9 所示，其中八极对称绕法的曲线与前述 2 种温度激励方式曲线变化趋势相同，十六极对称绕法和三十二极对称绕法的陀螺漂移仍存在波动。从整个过程来看，十六极对称绕法变化更小，抑制轴向温度激励下的陀螺漂移更有效。

计算并统计了上述 3 种激励方式下各对称绕法绕制的热致陀螺漂移标准差，如表 3 所示，可以看出，无论何种温度激励方式，采用十六极对称绕法的光纤热致陀螺漂移标准差均最小。综合这一

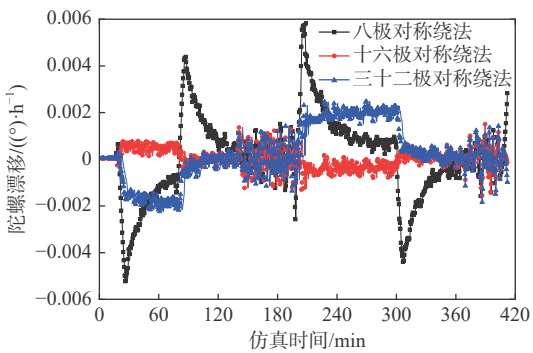


图 9　轴向单侧激励下各对称绕法的陀螺热致漂移曲线

Fig. 9　Gyro drift error curve under axial temperature excitation

表 3　各绕法不同温度激励方式热致陀螺漂移标准差

Table 3　Gyro drift SD of different temperature excitations and winding methods　(°)/h

绕制方法	四周均匀激励	径向单侧激励	轴向单侧激励
八极对称	0.001 25	0.001 49	0.001 81
十六极对称	0.000 29	0.000 43	0.000 45
三十二极对称	0.000 73	0.000 72	0.001 23

结果和文献 [10] 提供的结论，可以说明相对于采用四极、八极、三十二极对称绕法的光纤陀螺，采用十六极对称绕法的光纤陀螺对 Shupe 误差（与标准差正相关）的抑制效果最好。

从理论上讲，更多极的对称绕法能够更大程度地抑制光纤环中的 Shupe 误差，但以上仿真结果与理论出现偏差。分析其原因，可能是由于仿真采用 32 层的光纤环能够完成 2 个循环十六极对称绕法的绕制，而若使用三十二极对称绕法，则仅能完成 1 个绕制循环。为了验证这一猜想，在保持其他条件不变的情况下，采用 20 匝 64 层的光纤环做了径向单侧激励的进一步仿真，仿真结果如图 10 和表 4 所示。可以看出，当实现三十二极对称绕法 2 个循环绕制时，其性能仍然低于十六极对称绕法。

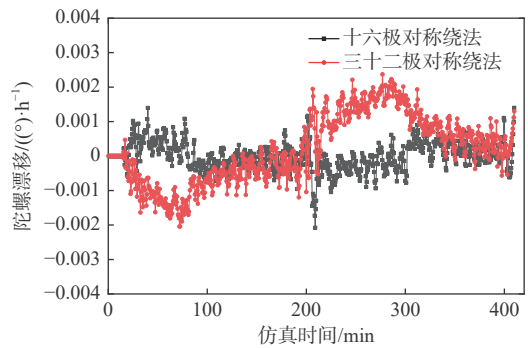


图 10　径向单侧激励下 64 层光纤环不同绕法的陀螺漂移曲线

Fig. 10　Gyro drift error curve under the radial temperature excitation with 64 layer fiber coil

表 4　64 层光纤环各绕法热致陀螺漂移标准差

Table 4　Gyro drift SD of different winding methods with 64 layer fiber coil　(°)/h

绕制方法	标准差
十六极对称	0.000 42
三十二极对称	0.000 99

本文分析是由于光纤环中各材料的导热系数差异所致，加之光纤环工作温度为−40～60℃，在该温度环境下光纤环的温度梯度产生的 Shupe 误差能够被十六极对称绕法很好地抑制，而由于材料导热系数差异，三十二极对称绕法未能发挥其效果，从而导致三十二极对称绕法性能较差。

4　结　论

1) 在本文采用的条件下，相对于八极对称绕法，采用十六极对称绕法和三十二极对称绕法的光纤陀螺能够有效抑制热致陀螺漂移；

2) 十六极和三十二极对称绕法的光纤陀螺在温度梯度改变时不会出现陀螺漂移的突变峰值，且十六极对称绕法的抑制效果最优。

上述结论对高精度光纤陀螺的光纤环绕制方法的选择上提供了指导。

参考文献（References）

[1] VALI V, SHORTHILL R W. Fiber ring interferometer[J]. Applied Optics, 1976, 15(5): 1099-1100.

[2] 霍雷, 谢良平, 谌尧周, 等. 光纤陀螺的发展与应用[J]. 电子科技, 2015, 28(8): 174-177.
HUO L, XIE L P, CHEN Y Z, et al. Development and application of fiber optic gyroscopes[J]. Electronic Science and Technology, 2015, 28(8): 174-177(in Chinese).

[3] SHUPE D. Thermally induced nonreciprocity in the fiber optic interferometer[J] Applied Optics, 1980, 19(5): 654-655.

[4] FRIGO N J. Compensation of linear source of non-reciprocity in Sagnac interferometers[J]. Fiber Optic and Laser Sensor I, 1983, 412: 268-271.

[5] MOHR F. Thermooptically induced bias drift in fiber optical Sagnac interferometers[J]. Journal of Lightwave Technology, 1995, 14(1): 27-41.

[6] GAO Z X, ZHANG Y G, WANG G C, et al. Analysis and simulation for the thermal performance of the octupolar fiber coil[J]. Optical Engineering, 2014, 53(1): 016114.

[7] WILLIAMS M . Optical fiber placement for crossover-free fiber optic gyros[C]//Conference on Photorefractive Fiber and Crystal Devices. Materials, Optical Properties, and Applications, 2006, 6314: 108-118.

[8] DU S S, GUAN Y M, JIN J, et al. Finite element model of thermal transient effect for crossover-free fiber optic gyros[J]. Optik, 2012, 123(8): 748-751.

[9] DYOTT R. Reduction of the Shupe effect in fibre optic gyros, the random-wound coil[J]. Electronics Letters, 1996, 32(23): 2177-2178.

[10] 万洵, 谢良平. 光纤陀螺温度场仿真分析与陀螺外罩结构优化设计[J]. 应用光学, 2016, 37(3): 353-358.
WAN X, XIE L P. Temperature field analysis and structure redesign of fiber optic gyroscope[J]. Journal of Applied Optics, 2016, 37(3): 353-358(in Chinese).

[11] 陈军, 王巍, 李晶, 等. 光纤线圈固化对光纤陀螺性能的影响[J]. 中国惯性技术学报, 2012, 20(6): 709-714.
CHEN J, WANG W, LI J, et al. Effect caused by potted adhesive on FOG performance[J]. Journal of Chinese Inertial Technology, 2012, 20(6): 709-714(in Chinese).

[12] 黄云柯, 李强, 胡梦纯, 等. 光纤捷联惯性测量组合热设计及温度控制技术[J]. 上海航天, 2016, 33(S1): 122-126.
HUANG Y K, LI Q, HU M C, et al. Research on temperature control and thermal design of FOG inertial measurement unit[J]. Aerospace Shanghai, 2016, 33(S1): 122-126(in Chinese).

[13] 冯卡力, 李安, 覃方君, 等. 光纤陀螺温度误差自适应神经模糊补偿方法[J]. 兵工学报, 2016, 37(4): 641-647.
FENG K L, LI A, QIN F J, et al. Temperature error compensation method based on adaptive neuro fuzzy inference for fiber-optic gyro[J]. Acta Armamentarii, 2016, 37(4): 641-647(in Chinese).

[14] 李绪友, 张春梅, 刘华兵, 等. 光纤环十六极对称绕法温度性能的仿真与分析[J]. 中国惯性技术学报, 2016, 24(6): 780-785.
LI X Y, ZHANG C M, LIU H B, et al. Simulation and analysis on temperature performance of fiber ring by 16-polar symmetrical winding method[J]. Journal of Chinese Inertial Technology, 2016, 24(6): 780-785(in Chinese).

[15] 李绪友, 凌卫伟, 许振龙, 等. 基于交叉法绕制的光纤环的槽体设计[J]. 光学学报, 2015, 36(6): 0606002.
LI X Y, LING W W, XU Z L, et al. Design of a new spool for fiber coil based on cross winding pattern[J]. Acta Optica Sinica, 2015, 36(6): 0606002(in Chinese).

[16] 王玥泽, 陈晓冬, 张桂才, 等. 八极绕法对光纤陀螺温度性能的影响[J]. 中国惯性技术学报, 2012, 20(5): 617-620.
WANG Y Z, CHEN X D, ZHANG G C, et al. Effect of octupole winding on temperature performance of fiber optic gyroscope[J]. Journal of Chinese Inertial Technology, 2012, 20(5): 617-620(in Chinese).

Simulation analysis of reduction effect of symmetrical winding method for multi-polar fiber ring on Shupe error

ZHANG Dongbo[1, 2]，WANG Lixin[1, *]，LI Can[1]

(1. College of Missile Engineering，Rocket Force University of Engineering，Xi'an 710025，China；
2. Department of Fiber Sensors，No.16 Institution of Aerospace & Technology Corporation，Xi'an 710100，China)

Abstract：The Shupe bias error is one of the biggest bottleneck problems in the engineering application of interferometric fiber optic gyroscope (IFOG). A turn-by-turn finite element model for fiber coils is established in this article. Based on this model, analyze the temperature performance of the 8 level, 16 level and 32 level symmetric winding method are analyzed under different temperature excitations. The results show that compared with octupolar symmetrical winding methods, the 16-polar and 32-polar symmetrical winding methods can effectively reduce the thermally induced gyro drift, and that the 16-polar symmetrical winding methods have the best temperature performance among the three methods. This conclusion provides guidance for the selection of fiber coil winding method for high precision IFOG.

Keywords：Shupe error；16-polar symmetrical winding method；fiber coil；finite element model；thermally induced gyro drift

Received：2021-09-06；**Accepted**：2021-09-17；**Published Online**：2021-11-30 16：29
URL：kns.cnki.net/kcms/detail/11.2625.V.20211129.1902.006.html
Foundation item：Natural Science Basic Research Program of Shaanxi (2020JQ-491)
* **Corresponding author**. E-mail：wlxxian@sina.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0548

空间大伸展并联机构的设计与性能分析

赫利涛，房海蓉*，陈宇飞，李寅
(北京交通大学 机械与电子控制工程学院，北京 100044)

摘　　要：大型工件加工时执行机构伸展空间不足是目前航天制造业一个亟待解决的问题。为增大执行机构末端的工作空间，并为其提供稳定可靠的伸展基础，提出一类具备高刚度、大伸展特性的并联可伸展机构。重点针对并联可伸展机构展开设计与分析，基于图论进行可伸展机构构型综合，利用构型演化得到PRRR、PRRR-3R和PRRR-6R（P表示移动副，R表示转动副）这3种可伸展支链单元构型，在此基础上，对3种可伸展机构支链单元的伸展性能、刚度进行对比分析，最终优选出满足要求的PRRR支链单元配置成并联可伸展机构。该并联可伸展机构可应用于需要提供大伸展及大刚度的场合，实现对其他执行机构运动空间的拓展。

关　键　词：图论；并联机构；可伸展支链；性能分析；刚度；构型综合
中图分类号：V462; TH112
文献标志码：A　　**文章编号**：1001-5965(2023)07-1722-13

随着航空航天领域的不断发展，航天结构件的尺寸越来越大，带来了一系列与结构、工艺、加工等有关的新问题和关键制造技术挑战[1]，而大型贮箱球冠面的加工就是目前面临的技术难题之一。大型贮箱的球冠直径接近10 m，球冠表面存在各种不同尺寸的法兰、支座等凸起结构[2]，整体结构复杂。而且由于大型球冠工件水平放置，整体加工呈卧式布局，所以要求加工时机构具有沿着水平方向大范围伸展的能力，而现有的机构存在伸展性能差、工作空间不足、刚度低等问题，无法同时满足要求。

为了实现加工执行机构沿着大型贮箱球冠面轴向水平伸展的目标，目前，针对卧式放置的大型工件，一种是采用串联机器人进行延伸，同时在其末端安装相应的执行装置实现加工。例如，针对大型复杂构件加工，文献[3-6]都是利用串联机器人的延伸能力，满足不同类型贮箱复杂曲面加工时的伸展要求。但现阶段的工业机器人末端负载小、运动范围有限，其刚度、精度、稳定性和运动范围均无法同时满足大型球冠面的加工要求。另一种则是利用可伸展机构作为伸展基础，为其他加工执行机构提供更大的工作空间。针对可伸展单元，文献[7-10]对空间伸展臂机构进行了设计与分析；马艳等[11]提出一种由8个转动关节依次连接组成的可折展空间连杆机构；刘仁志等[12]利用正八面体作为基本单元，对空间伸展臂结构进行构建。另外，韩博等[13]利用闭环可展机构单元的旋量约束拓扑图，对双层环形桁架可展结构进行分析；Zhao等[14]从平面连杆机构入手，提出一种基于剪叉式单元的可展开结构设计方法；史创等[15]研究了构建可展支撑机构可展单元的构型条件，提出一种曲柄滑块式可展单元。一般的可伸展机构，都是以单一支链结构为主，伸展能力强，但刚度较差，为了使可伸展机构具备一定刚度，Zhao等[16]设计了一类折叠式并联机构；房海蓉和陈江红[17]提出一种新型非完全对称的四自由度并联机构，具有结构紧凑、工作空间

收稿日期：2021-09-13；**录用日期**：2021-10-15；**网络出版时间**：2021-10-29 15:37
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20211029.1011.001.html
基金项目：中央高校基本科研业务费(2018JBZ007)
*通信作者. E-mail：hrfang@bjtu.edu.cn

引用格式：赫利涛，房海蓉，陈宇飞，等. 空间大伸展并联机构的设计与性能分析[J]. 北京航空航天大学学报，2023，49(7)：1722-1734.
HE L T，FANG H R，CHEN Y F，et al. Design and performance analysis of spatial large extension parallel mechanism[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1722-1734 (in Chinese).

大的特点。此外, 利用可重构机构的变尺度特性, 文献 [18-20] 提出不同的可重构并联机构, 将伸展机构作为可变基座进行结构配置, 以增大机构末端的运动范围, 使其具备大工作空间。

本文针对航天大型球冠结构件的轴向伸展要求, 结合可伸展机构与并联机构的优点, 提出一类具备高刚度、大伸展特点的并联可伸展机构。首先, 采用图论的方法对并联可伸展机构的支链单元进行构型综合, 利用构型演化得到 3 种可伸展支链。然后, 对 3 种可伸展支链的伸展性能进行分析与比较, 结果表明其均具有大伸展特性。接着, 采用有限元仿真分析比较 3 种支链的刚度及承载能力大小。最后, 综合机构伸展性能及刚度指标, 优选出 PRRR(P 表示移动副, R 表示转动副)支链单元配置成并联可伸展机构, 可作为单独模块为其他机构提供稳定可靠的伸展基础。

1　可伸展支链单元构型综合

在大伸展并联机构的设计中, 由于每一条支链结构的伸展性能及刚度决定了整体机构的性能, 因此, 需要对机构中的可伸展支链单元进行构型设计与分析, 通过对单一运动支链结构进行综合, 选出具备大伸展、大刚度的伸展机构。机构是具有固定构件的运动链, 故机构结构的型综合实质上为其运动链的型综合, 所以可就运动链进行机构结构的型综合, 以求得全部的机构类型。一般运动链都是由基本杆组构成, 因此, 本节以基本杆组为基础, 采用图论[21-22] 的方法, 对单自由度可伸展支链单元进行构型综合。

假设单一自由度的平面全转动副机构共有 N 个(包含机架)构件及 M 个运动副, 当运动链对应的机构自由度为 1 时, 若将机架的约束解除, 则该机构就变成了 1 个具有 4 自由度的全转动副运动链, 且该运动链应满足:

$$3N-2M=4 \tag{1}$$

在机构拓扑设计时, 根据欧拉公式可以得到拓扑结构图的内部环路与运动链之间的关系:

$$L=M-N+1 \tag{2}$$

式中: L 为拓扑图的内部区域个数, 即: 运动链的闭环数; N 为拓扑图的顶点个数, 即: 运动链的构件数; M 为拓扑图的边界个数, 即: 运动副(低副)的数量。

据此, 可以得到构件数 N 与运动副数 M 及运动链的闭环数 L 的对应关系, 如表 1 所示。

根据表 1 可以建立 5 种机构的部分拓扑简图, 其中, 空心圈表示构件, 直线表示运动副, 如图 1 所示。

表 1　构件数 *N*、运动副数 *M* 和运动链闭环数 *L* 的对应关系

Table 1　Corresponding relationship between components *N*,motion pairs *M* and closed loops *L*

构件数N	运动副数量M	运动链的闭环数L
2	1	0
4	4	1
6	7	2
8	10	3
⋮	⋮	⋮

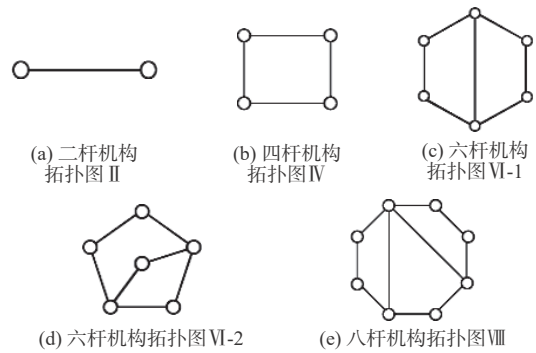


图 1　4 种基本杆组的部分拓扑图

Fig. 1　Partial topological diagrams of four basic rod groups

并联可伸展机构需要具备大伸展、大刚度、高稳定性等特点, 基于可伸展支链单元的拓扑结构设计, 同时为了能给其他机构提供更大范围工作空间, 在此, 对可伸展支链单元提出一些设计要求, 例如, 为便于拓扑变换, 可伸展支链单元均采用单自由度平面机构, 且拓扑结构的基本回路数应尽可能少; 采用平面低副来进行构件间的连接; 机构的驱动通过移动或转动的方式控制; 可伸展支链单元在伸展时需要有大承载能力等。

1.1　可伸展支链拓扑对称性分析

为了避免后续出现重复的拓扑结构, 需要利用拓扑对称性分析构件拓扑图, 去除构件重复的连接关系——同构体, 得到拓扑异构体, 进行构型综合。后续为方便对运动链拓扑对称性的分析判断, 将拓扑图中的顶点、边进行编号, 顶点用 n 表示, n=1, 2, …, 8, 边用 K_s 来表示, s=1, 2, …, 10, 下面对于不同机构进行拓扑对称性分析。

当 N=2 时, 二杆机构拓扑图Ⅱ如图 2 所示。可知, 构件 1 与构件 2 同构, 且只存在 1 种拓扑对称性关系。

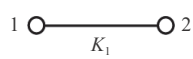


图 2　二杆机构拓扑图Ⅱ

Fig. 2　Topological diagram of two-bar mechanism Ⅱ

当 N=4 时, 四杆机构拓扑图Ⅳ如图 3 所示。根据拓扑图, 可以得到四杆机构拓扑图与邻接矩阵的

对应关系，如表2所示。分别去掉拓扑图中的构件1～构件4，则可以得到拓扑子图，如图4所示。

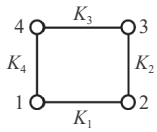


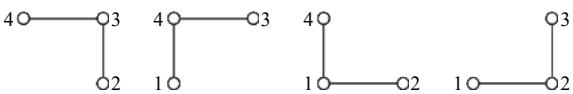
图3 四杆机构拓扑图Ⅳ

Fig. 3 Topology diagram of four-bar mechanism Ⅳ

表2 四杆机构拓扑图Ⅳ与邻接矩阵的对应关系

Table 2 Correspondence between topological diagram Ⅳ of four-bar mechanism and adjacency matrix

顶点	1	2	3	4
1	0	1	0	1
2	1	0	1	0
3	0	1	0	1
4	1	0	1	0



(a) 移除构件1 (b) 移除构件2 (c) 移除构件3 (d) 移除构件4

图4 四杆机构拓扑图Ⅳ的拓扑子图

Fig. 4 Topological subgraph of topological figure Ⅳ of four-bar mechanism

对应去掉构件1～构件4的邻接矩阵，如下：

$$\boldsymbol{X}_1=\begin{bmatrix}0&1&0\\1&0&1\\0&1&0\end{bmatrix};\quad \boldsymbol{X}_2=\begin{bmatrix}0&0&1\\0&0&1\\1&1&0\end{bmatrix};$$

$$\boldsymbol{X}_3=\begin{bmatrix}0&1&1\\1&0&0\\1&0&0\end{bmatrix};\quad \boldsymbol{X}_4=\begin{bmatrix}0&1&0\\1&0&1\\0&1&0\end{bmatrix}$$

通过初等变换，邻接矩阵$\boldsymbol{X}_2$、$\boldsymbol{X}_3$、$\boldsymbol{X}_4$均可以等价成$\boldsymbol{X}_1$的形式，表明分别移除构件1～构件4之后的4个拓扑变换子图，相互之间均是同构，因此，4个构件之间具有拓扑对称性，而且存在1种拓扑对称性关系。

当N=6时，六杆机构拓扑图Ⅵ-1如图5所示。

根据拓扑图，可以得到六杆机构拓扑图Ⅵ-1与邻接矩阵的对应关系，依次移除构件后，则可以得到6个拓扑子图，如图6所示。

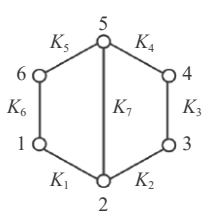
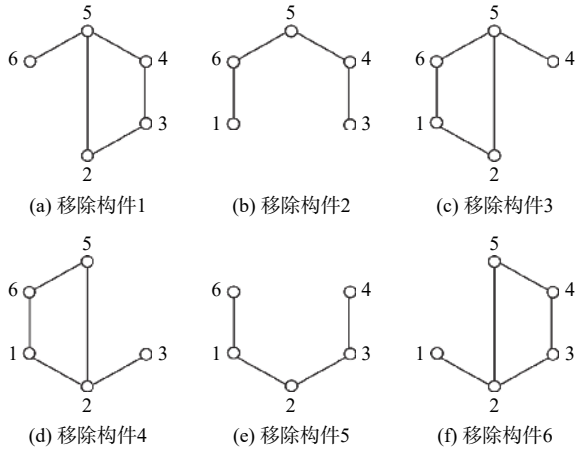


图5 六杆机构拓扑图Ⅵ-1

Fig. 5 Topology diagram of six-bar mechanism Ⅵ-1



(a) 移除构件1 (b) 移除构件2 (c) 移除构件3

(d) 移除构件4 (e) 移除构件5 (f) 移除构件6

图6 六杆机构拓扑图Ⅵ-1的拓扑子图

Fig. 6 Topological subgraph of six-bar mechanism topology Ⅵ-1

同样，由于图6的拓扑子图均为无向图，为进行同构判断，根据拓扑子图的邻接矩阵，进行初等变换后，可以得到图6(c)、图6(d)、图6(f)与图6(a)，图6(b)与图6(e)分别具有拓扑对称性。因此，在六杆机构拓扑图Ⅵ-1中存在2种拓扑对称性关系，即：构件1、构件3、构件4与构件6同构，构件2与构件5同构。

当N=6时，另一种拓扑图如图7所示。

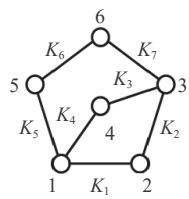


图7 六杆机构拓扑图Ⅵ-2

Fig. 7 Topology diagram of six-bar mechanism Ⅵ-2

同理也可以得到六杆机构拓扑图Ⅵ-2的拓扑对称性关系，一共存在3种拓扑对称性关系，即：构件1与构件3同构，构件2与构件4同构，构件5与构件6同构。

当N=8时，八杆机构拓扑图Ⅷ的其中一种，如图8所示。根据拓扑图，可以得到八杆机构拓扑图Ⅷ邻接矩阵的对应关系，同样，依次移除构件后，可以得到8个拓扑子图，如图9所示。

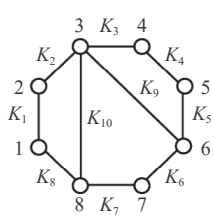
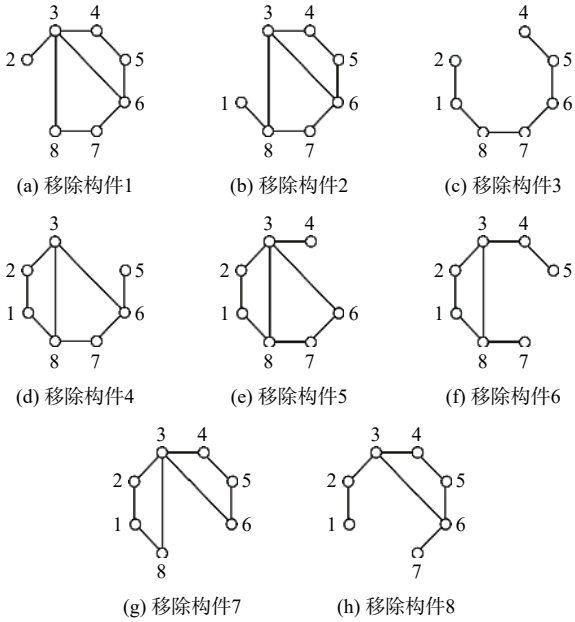


图8 八杆机构拓扑图Ⅷ

Fig. 8 Topological diagram of eight-bar mechanism Ⅷ



(a) 移除构件1　(b) 移除构件2　(c) 移除构件3

(d) 移除构件4　(e) 移除构件5　(f) 移除构件6

(g) 移除构件7　(h) 移除构件8

图 9　八杆机构拓扑图Ⅷ的拓扑子图

Fig. 9　Topological subgraph of eight-bar mechanism topology Ⅷ

进行初等变换后, 同理可得, 图 9(a)与图 9(e)、图 9(b)与图 9(d)、图 9(f)与图 9(h)具有拓扑对称性。因此, 在八杆机构拓扑图Ⅷ中存在 5 种拓扑对称性关系, 即: 构件 1 (与构件 5 同构), 构件 2 (与构件 4 同构), 构件 6(与构件 8 同构), 构件 3, 构件 7。

1.2　可伸展支链单元构型演化

根据 1.1 节的内容, 利用拓扑对称性分析构件, 得到不同的机构拓扑图, 本节针对每一类机构拓扑图, 采用双色拓扑图进行分析, 通过构型演化, 获得最终的可伸展支链单元。

在可伸展支链的基本单元中, 用 N_{II}、N_{IV}、N_{VI} 分别表示拓扑图Ⅱ、Ⅳ、Ⅵ综合出的机构总数, ${}_i^jM_N$ 表示在某种情况下综合出的机构总数, 其中, N 为拓扑图的构件数, N=2、4、6、8; i 为选定当作机架的构件编号; i=1, 2, …, 8; j 为该情况下移动副 P 的个数, j=0, 1, 2。

以四杆机构为例, 由于构件 1～构件 4 具有拓扑对称性, 且存在 1 种同构关系, 则假设构件 1 为机架, 由机架位置与移动副 P 的组合关系, 可得构型数 $N_{\mathrm{IV}}={}_1^1M_4+{}_1^2M_4+{}_1^0M_4=3+3+1=7$。在此, 仅给出构件 1 为机架且有 1 个 P 副 (移动副) 及全为 R 副(转动副)时的构型演化, 如表 3 所示。

同理可得, 当 N=6 时, 在六杆机构拓扑图Ⅵ-1 中, 利用拓扑图中存在的 2 种拓扑对称性关系, 当机架分别在构件 1、构件 2 时, 根据其与移动副 P 的组合关系, 可以得到构型数如表 4 所示。

同理, 在六杆机构的拓扑图Ⅵ-2 中, 由机架位置与移动副 P 的组合关系, 可以得到构型数如表 5 所示。

表 3　四杆机构拓扑图部分构型演化

Table 3　Partial configuration evolution of topological diagram for four-bar mechanism

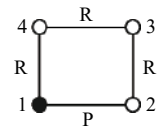
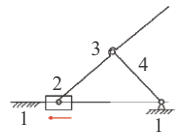
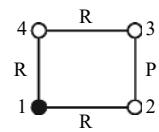
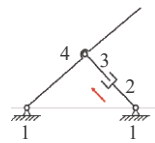
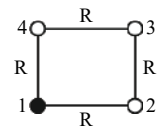
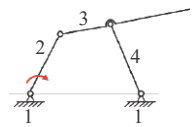
运动副情况	拓扑图	演化构型
1个P副		
1个P副		
全R副		

表 4　六杆机构拓扑 Ⅵ-1 的构型总数

Table 4　Total number of configurations of topological diagram for six-bar mechanism Ⅵ-1

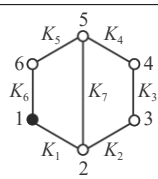
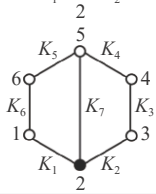
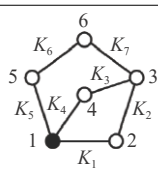
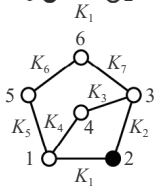
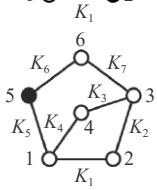
机架	拓扑图	1个P副	2个P副	全R副	构型总数
构件1		${}_1^1M_6=7$	${}_1^2M_6=21$	${}_1^0M_6=1$	29
构件2		${}_2^1M_6=4$	${}_2^2M_6=12$	${}_2^0M_6=1$	17

表 5　六杆机构拓扑 Ⅵ-2 的构型总数

Table 5　Total number of configurations of topological diagram for six-bar mechanism Ⅵ-2

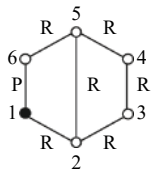
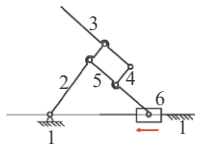
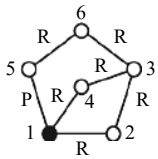
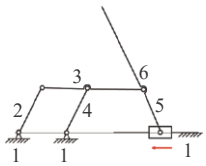
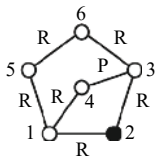
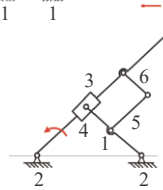
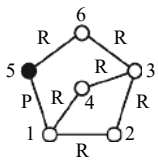
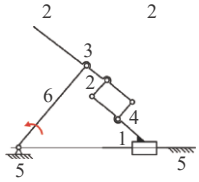
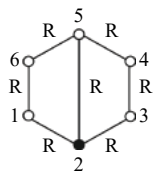
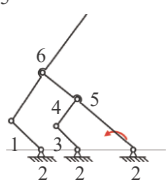
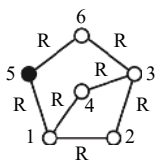
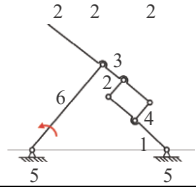
机架	拓扑图	1个P副	2个P副	全R副	构型总数
构件1		${}_1^1M_6=5$	${}_1^2M_6=13$	${}_1^0M_6=1$	19
构件2		${}_2^1M_6=4$	${}_2^2M_6=12$	${}_2^0M_6=1$	17
构件5		${}_5^1M_6=5$	${}_5^2M_6=13$	${}_5^0M_6=1$	19

因此, 当 $N=6$时, 六杆机构拓扑图的构型总数 $N_{\mathrm{VI}}=N_{\mathrm{VI-1}}+N_{\mathrm{VI-2}}=46+55=101$。

在此仅给出当构件1、构件2、构件5分别为机架时，部分拓扑图的演变构型，如表6所示。

表6 六杆机构拓扑图部分构型演化

Table 6 Partial configuration evolution of topological diagram for six-bar mechanism

序号	运动副情况	拓扑图	演化构型
1	1个P副		
2	1个P副		
3	1个P副		
4	1个P副		
5	全R副		
6	全R副		

综上所述，当构件数≤6时，平面基本单元的构型总数为 $N_{\mathrm{ALL}}=N_{\mathrm{II}}+N_{\mathrm{IV}}+N_{\mathrm{VI}}=2+7+101=110$

同样，由于八杆机构存在16种异构拓扑图，通过固定不同机架，配置不同运动副，则会有上百种构型，因篇幅有限，不作过多描述。在此只选取其中构件1为机架时，只有1个P副及全为R副的情况下的拓扑图进行构型演化，仅列出演变的2种构型，如表7所示。

本节采用图论的方法进行构型综合，当构件数≤8时，经过拓扑分析计算，可以得到基本单元的构型总数有上百种，但通过分析结构配置，当机构具有1个P副，而且特别是P副为主动副时，其伸展及折叠性能更优。依据演化构型的特点及对可伸展支链单元的设计要求的综合考虑，最终筛选出3种具备大伸展能力的平面单自由度可伸展支链构型，如表8所示。各支链单元由不同数量运动副组合形成，支链中运动副前的数字代表该运动副的数量。

表7 八杆机构拓扑图部分构型演化

Table 7 Partial configuration evolution of topological diagram of eight-bar mechanism

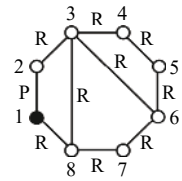
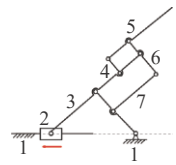
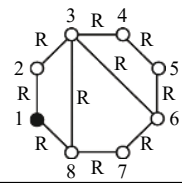
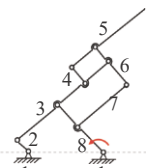
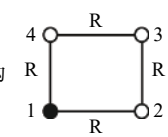
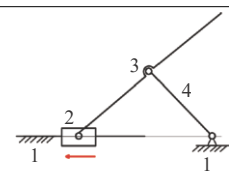
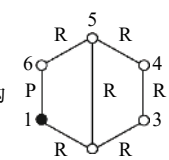
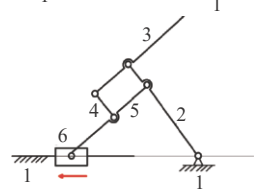
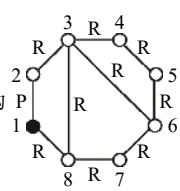
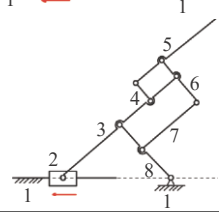
运动副情况	构型拓扑图	演化构型
1个P副		
全R副		

表8 可伸展支链构型拓扑图及其演化构型

Table 8 Topological diagram of extendable branched chain configuration and its evolutionary configuration

支链单元	机构	构型拓扑图	演化构型
PRRR支链	四杆机构		
PRRR-3R支链	六杆机构		
PRRR-6R支链	八杆机构		

2 可伸展机构伸展性能分析

对于可伸展机构伸展性能的分析，为了便于比较不同支链的伸展性能，需要定义相关指标，来表达可伸展机构伸展能力的强弱。在此，提出线伸展比这项指标，将线伸展比 Θ 定义为：机构输出端伸长距离与机构输入端移动距离的比值，即：在相同展开距离下，伸长距离越远，线伸展比就越大，表明可伸展机构的伸展能力越强。仅从可伸展支链空间伸展性上来讲，线伸展比越大越好。

$$\Theta=\frac{\Delta h}{\Delta y} \tag{3}$$

式中: Δh 为机构输出端伸长距离; Δy 为机构输入端移动距离。

完成对线伸展比的定义之后, 需要进一步对可伸展支链单元进行分析, 根据其伸展能力, 利用线伸展比 Θ 来评价不同可伸展支链单元的伸展性能。

2.1　可伸展支链伸展性能分析

2.1.1　PRRR 支链伸展性能分析

在 PRRR 支链中, 为了分析机构的伸展运动, 以 O 点为坐标原点, 建立坐标系, 水平向右为 y 轴的正方向, 竖直向上为 z 轴的正方向, E 点为机构滑块运动极限位置点, 如图 10 所示。令 $L_{AB}=l_1$, $L_{BC}=l_2$, $L_{OB}=r$, $L_{AC}=l_1+l_2=l$, 其中 l、r、a 均已知, θ的取值范围在(0°, 90°)。当滑块从 A 点经过 A_1 点, 运动到 E 点时, 达到极限位置。假设 A 点距离坐标原点的距离 y_a 为机构的输入量, 则有 $L_{OA}=y_a$。

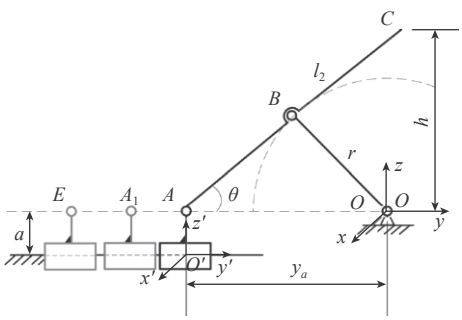


图 10　建立 PRRR 支链坐标系

Fig. 10　Establishment of PRRR branch chain coordinate system

在三角形 ΔOAB 中, 根据余弦定理有

$$\cos\theta=\frac{L_{AB}^2+L_{OA}^2-L_{OB}^2}{2L_{AB}L_{OA}}=\frac{l_1^2+{y_a}^2-r^2}{2l_1y_a} \tag{4}$$

因此, 可以得到:

$$\theta=\arccos\left(\frac{l_1^2+{y_a}^2-r^2}{2l_1y_a}\right) \tag{5}$$

其中, 夹角θ与连接杆杆长l_1、支撑杆杆长 r, 输入距离 y_a 有关。

对θ求导, 则有

$$\frac{\mathrm{d}\theta}{\mathrm{d}y_a}=-\frac{r^2-l_1^2+{y_a}^2}{2l_1{y_a}^2\sqrt{1-\dfrac{(l_1^2-r^2+{y_a}^2)^2}{(2l_1y_a)^2}}}=-\frac{r^2-l_1^2+{y_a}^2}{y_a\sqrt{(2l_1y_a)^2-(l_1^2-r^2+{y_a}^2)^2}} \tag{6}$$

在三角形 ΔOAC 中, 由几何关系可得:

$$\begin{cases}l=l_1+l_2\\ \sin\theta=\dfrac{L_{OC}}{L_{AC}}=\dfrac{h}{l_1+l_2}=\dfrac{h}{l}\\ h=l\sin\theta=(l_1+l_2)\sin\theta=(l_1+l_2)\sqrt{1-\cos^2\theta}\end{cases} \tag{7}$$

将式(5)代入式(7)中可得到伸展机构输入量 y_a 与输出量 h 之间的关系:

$$h=(l_1+l_2)\sin\theta=(l_1+l_2)\sqrt{1-\cos^2\theta}=(l_1+l_2)\sqrt{1-\left(\frac{l_1^2+{y_a}^2-r^2}{2l_1y_a}\right)^2} \tag{8}$$

在所设计的机构中, 初定 $l_1=l_2=900$ mm, $l=l_1+l_2=1\,800$ mm, r=900 mm, 则得到输入距离 y_a 与夹角θ、输入距离 y_a 与$\mathrm{d}\theta/\mathrm{d}y_a$、输入距离 y_a 与伸展高度 h 的关系曲线, 如图 11所示。可知, 随着滑块的输入距离 y_a 变大, 连接杆与水平方向夹角θ逐渐变小; θ的变化率逐渐增大; 机构伸展高度 h 逐渐变小, 与夹角θ的变化趋势一致, 即: 整个机构竖直方向的空间得到压缩。

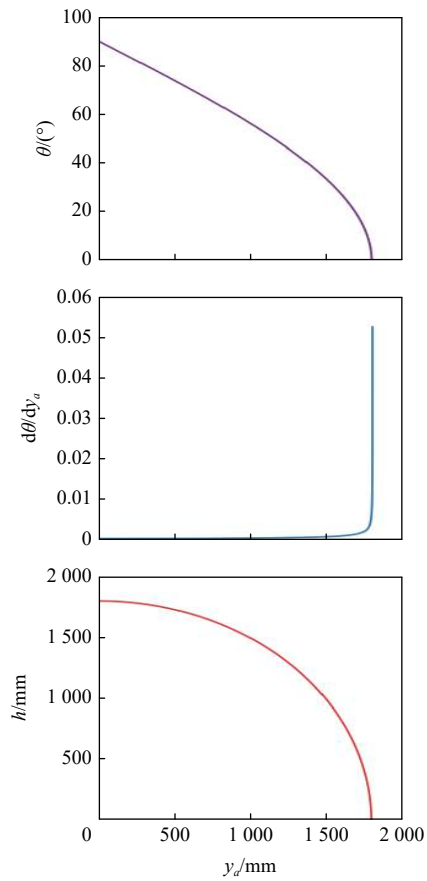
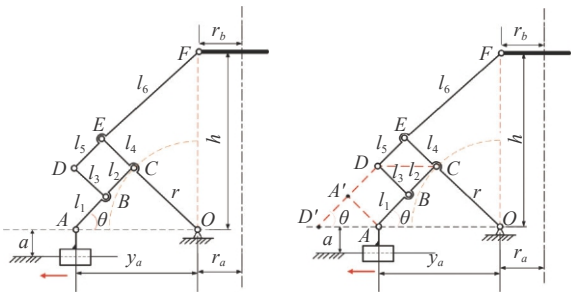


图 11　PRRR 支链伸展性能变化曲线

Fig. 11　Stretching performance change curve of PRRR branch chain

2.1.2　PRRR-3R 支链伸展性能分析

同样, 在 PRRR-3R 支链中, 以 O 点为坐标原点, 建立坐标系, 水平向右为 y 轴正方向, 竖直向上为 z 轴正方向, 如图 12 所示。令 $L_{AB}=l_1$, $L_{BC}=l_2$, $L_{OC}=r$, $L_{AC}=l_1+l_2=l$, $L_{DF}=l_5+l_6$, 其中 l、r、a 均为已知定值, θ的取值范围在(0°, 90°)。



(a) PRRR-3R支链 (b) PRRR-3R等效支链

图 12 建立 PRRR-3R 支链坐标系

Fig. 12 Establishment of PRRR-3R branch chain coordinate system

同理，在三角形 ΔOAC 中，根据余弦定理可得各变量之间的关系，在此仅给出表达式推导结果，如下：

$$\theta = \arccos\left(\frac{(l_1+l_2)^2+y_a{}^2-r^2}{2(l_1+l_2)y_a}\right) \tag{9}$$

$$\frac{\mathrm{d}\theta}{\mathrm{d}y_a} = -\frac{r^2-(l_1+l_2)^2+y_a{}^2}{2(l_1+l_2)y_a{}^2\sqrt{1-\dfrac{\left[(l_1+l_2)^2-r^2+y_a{}^2\right]^2}{\left[2(l_1+l_2)y_a\right]^2}}} \tag{10}$$

$$h = (l_1+l_2+l_5+l_6)\sqrt{1-\left(\frac{(l_1+l_2)^2+y_a{}^2-r^2}{2(l_1+l_2)y_a}\right)^2} \tag{11}$$

进一步可以分别得到机构的输入距离 y_a 与夹角 θ、$\mathrm{d}\theta/\mathrm{d}y_a$及伸展高度 h 之间的关系变化曲线，如图 13 所示。可知，随着滑块的输入距离 y_a 变大，PRRR-3R 支链伸展性能变化趋势与实际运动一致。

2.1.3 PRRR-6R 支链伸展性能分析

同样，在 PRRR-6R 支链中，以 O 点为坐标原点，建立坐标系，水平向右为 y 轴正方向，竖直向上为 z 轴正方向，如图 14 所示。令 $L_{AB}=l_1$，$L_{BC}=l_2$，$L_{OC}=r$，$L_{AC}=l_1+l_2=l$，$L_{GI}=l_8+l_9$，其中 l、r、a 均为已知定值，θ的取值范围在(0°, 90°)。

同理，在三角形 ΔOAC 中，根据余弦定理可得各变量之间的关系，由于 PRRR-6R 支链与 PRRR-3R 支链的分析过程相似，其中机构的输入距离 y_a 与夹角 θ、$\mathrm{d}\theta/\mathrm{d}y_a$的关系表达同式(9)、式(10)。因此，在此仅给出输入距离 y_a 与伸展高度 h 之间的表达式推导结果为

$$h = (l_1+l_2+l_8+l_9)\sqrt{1-\left(\frac{(l_1+l_2)^2+y_a{}^2-r^2}{2(l_1+l_2)y_a}\right)^2} \tag{12}$$

进一步可以得到机构的输入距离 y_a 与夹角 θ、$\mathrm{d}\theta/\mathrm{d}y_a$及伸展高度 h 之间的关系变化曲线，如图 15 所示。可知，随着滑块的输入距离 y_a 变大，PRRR-6R 支链伸展性能变化趋势与实际运动一致。

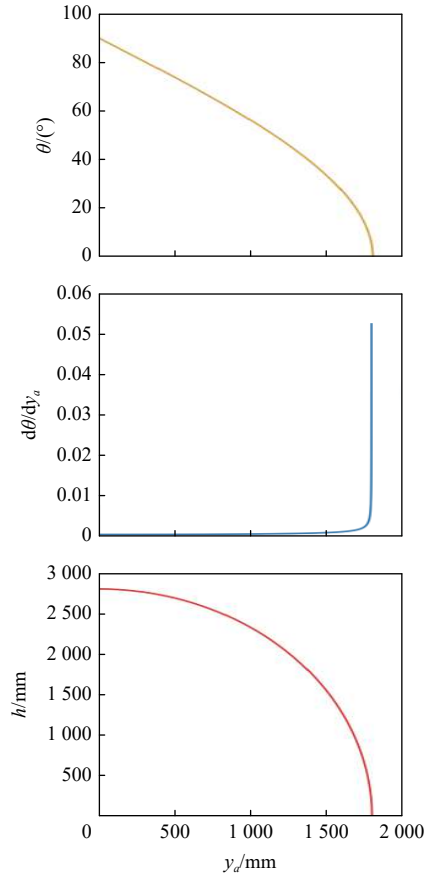
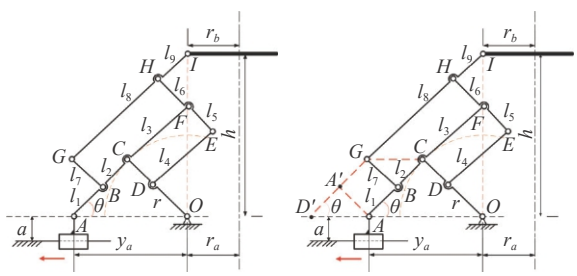


图 13 PRRR-3R 支链伸展性能变化曲线

Fig. 13 Stretching performance change curve of PRRR-3R branch chain



(a) PRRR-6R支链 (b) PRRR-6R等效支链

图 14 建立 PRRR-6R 支链坐标系

Fig. 14 Establishment of PRRR-6R branch chain coordinate system

2.2 可伸展支链伸展性能对比分析

当不同的可伸展机构，驱动滑块运动相同距离时，则可以得到其伸展高度 h 的对比、机构支链的连接杆与水平方向夹角θ变化的对比，如图 16 所示。

通过理论计算，可以得到机构末端在运动前后高度 h 的数值，将其汇总如表 9 所示。

由表 9 中的数据可以分析得出，当机构初始条件相同，驱动滑块运动相同距离时，不同支链结构的伸展高度不同，其伸展性能存在差异，机构在空间中的压缩性能也有差别。利用上述数据，通过计

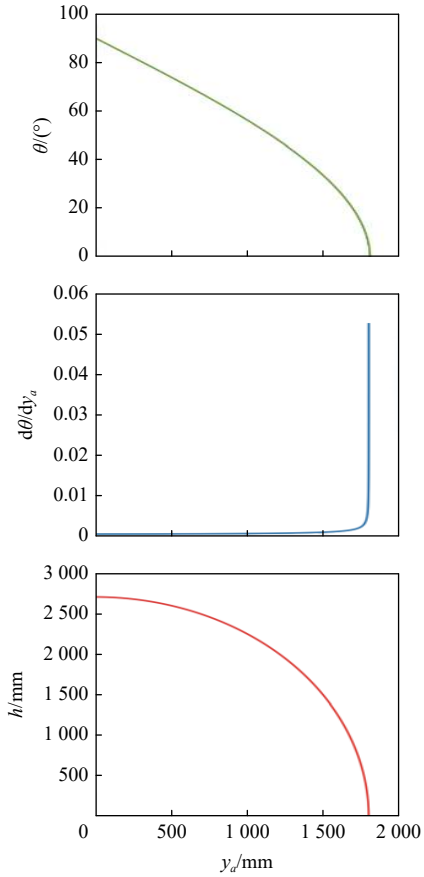


图 15　PRRR-6R 支链伸展性能变化曲线

Fig. 15　Stretching performance change curve of PRRR-6R branch chain

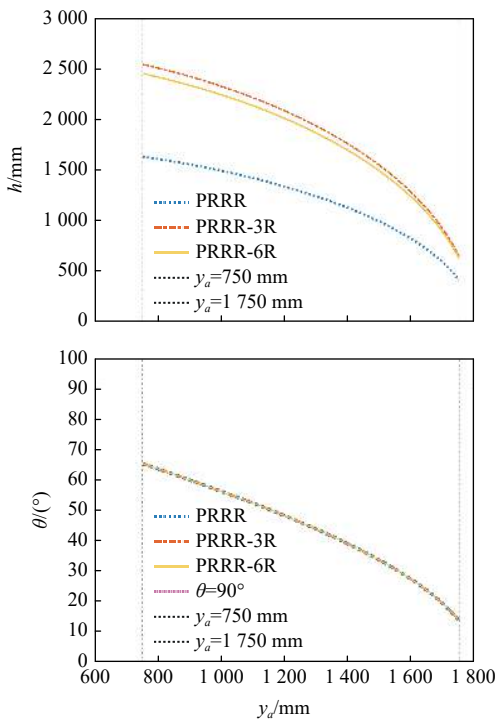


图 16　不同可伸展机构的对比

Fig. 16　Comparison of different stretchable mechanisms

算 3 种可伸展支链的线伸展比 Θ, 可以得到其伸展性能指标, 如表 10 所示。

表 9　不同可伸展机构的伸展性能比较

Table 9　Comparison of stretching performance of different extendable mechanisms

序号	基本支链单元	移动距离 Δy/mm	起始高度/mm	终止高度/mm	高度差 Δh/mm
1	PRRR 支链	1 000	1 636.306 8	421.307 5	1 214.999 3
2	PRRR-3R 支链	1 000	2 545.366 2	655.367 2	1 889.999 0
3	PRRR-6R 支链	1 000	2 454.460 2	631.961 2	1 822.499 0

表 10　3 种支链构型的线伸展比

Table 10　Linear extension ratio of three branched chain configurations

基本支链单元	支链尺寸参数	高度差 Δh/mm	线伸展比
PRRR 支链		1 214.999 3	1.215
PRRR-3R 支链		1 889.999 0	1.890
PRRR-6R 支链		1 822.499 0	1.822

3　可伸展机构刚度分析

3.1　可伸展支链单元刚度分析

仅支链的伸展性能好并不能完全满足最初的需求, 还需要各个支链结构具备大的刚度及承载能力。为了能够进一步比较 3 种支链结构的可靠性, 需要对支链的刚度及承载情况进行分析。本节采用有限元分析软件 ABAQUS 来对可伸展支链运动过程进行仿真, 通过分析连接杆在相同条件下受到的最大应力, 比较可伸展支链的变形情况, 从而找到刚度大、承载能力比较强的支链单元, 使其能够

作为并联可伸展机构的支链。本节对PRRR支链、PRRR-3R支链、PRRR-6R支链这3种支链的刚度展开分析。

对于3种支链的结构材料，选用45号钢，相关参数如表11所示。

表11 45号钢的相关性能参数

Table 11 Relevant performance parameters of 45# steel

材料名称	弹性模量/($N\cdot m^{-2}$)	泊松比	质量密度/($kg\cdot m^{-3}$)	屈服强度/($N\cdot m^{-2}$)
45号钢	2.09×10^{11}	0.269	7.89×10^{3}	3.55×10^{8}

在有限元分析软件ABAQUS中设定初始值，赋予零件材料之后，进行划分网格，如图17所示。同时，为PRRR支链、PRRR-3R支链、PRRR-6R支链的各构件间设置约束条件，并且因为整个支链是沿水平方向伸展且还要承载，故施加重力及连杆表面为0.001 MPa的压力载荷。

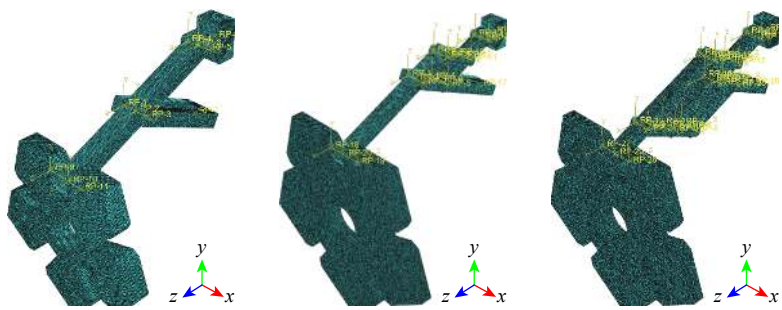
对可伸展支链进行仿真求解，得到PRRR支链单元、PRRR-3R支链单元及PRRR-6R支链单元运动过程中的应力分析结果，分别如图18～图20所示。

在运动过程中，随着整个可伸展支链单元不断压缩，主要连接部位受到的应力逐渐增大，当滑块运动到最远端时，连杆支撑区域受力最大。从应力云图中可以分析得到，可伸展支链单元的最大受力部分处于支撑杆和主连接杆铰接的部分，由于两杆连接处有开孔，故开孔周围处于应力值较高。而在驱动滑块移动1 000 mm的运动过程中，PRRR支链、PRRR-3R支链及PRRR-6R支链的连杆的最大应力强度依次为22.84，98.60，94.43 MPa，最大应力云图如图21所示，均远小于连杆的屈服强度，因此，这3种可伸展支链单元结构相对安全可靠。

综上所述，对于可伸展支链单元分析结果表明：主要承载部分位于连杆上表面，应力主要集中在铰接支撑的部分，但整体连杆的最大应力强度远低于屈服强度，结构承载可行。

3.2 可伸展支链刚度对比分析

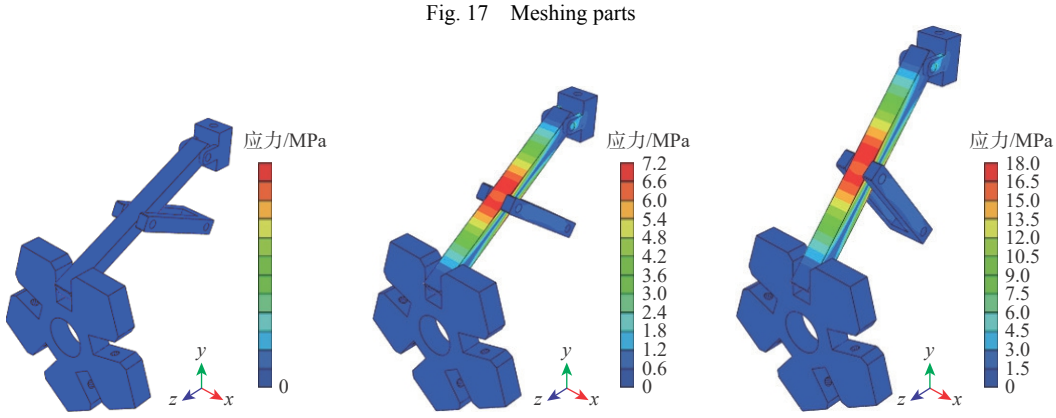
从3.1节3种不同支链结构运动过程中的刚度分析可以看出，在相同条件下，当主要受力连杆承载一定压力，且驱动滑块运动相同距离时，3种支链结构承受的最大应力值存在差异，每种支链的刚度变化不同，故其承载能力也有差异。相同承载压力下，受力杆的最大应力越大，说明受力杆越容易变形，也就表明其承载能力越弱。因此，对于可伸展支链单元的刚度及承载能力分析，可以通过相同条件下，受力杆的最大应力值来判断运动支链的承载能力强弱。



(a) PRRR支链 (b) PRRR-3R支链 (c) PRRR-6R支链

图17 划分网格

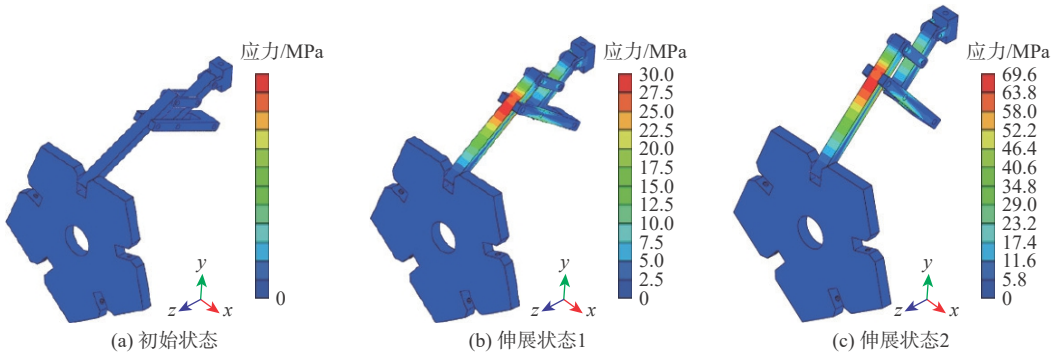
Fig. 17 Meshing parts



(a) 初始状态 (b) 伸展状态1 (c) 伸展状态2

图18 PRRR支链的应力云图

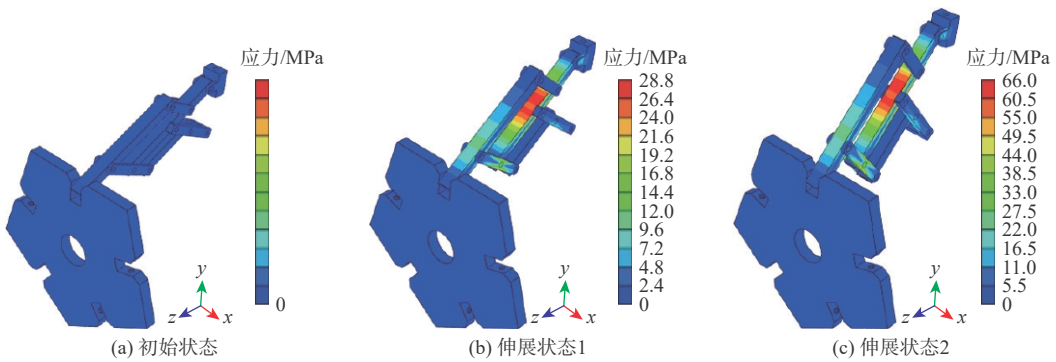
Fig. 18 Stress cloud chart of PRRR branch chain



(a) 初始状态　(b) 伸展状态1　(c) 伸展状态2

图 19　PRRR-3R 支链的应力云图

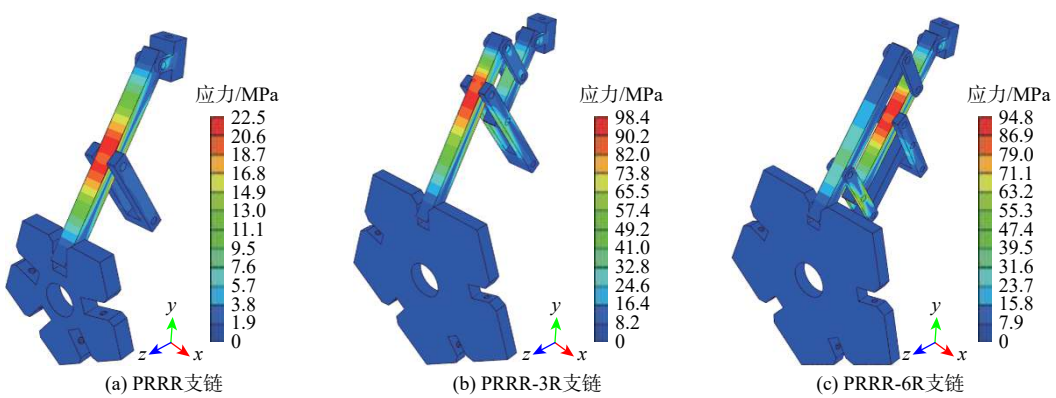
Fig. 19　Stress cloud chart of PRRR-3R branch chain



(a) 初始状态　(b) 伸展状态1　(c) 伸展状态2

图 20　PRRR-6R 支链的应力云图

Fig. 20　Stress cloud chart of PRRR-6R branch chain



(a) PRRR支链　(b) PRRR-3R支链　(c) PRRR-6R支链

图 21　最大应力云图

Fig. 21　Maximum stress cloud chart

一般情况下, 工程结构设计中用许用应力 $[\sigma]$ 来衡量允许构件承受的最大应力值。为了能够量化一种指标来比较 3 种支链结构的承载能力强弱, 用最大应力值占许用应力的百分比来比较, 由于均采用相同条件, 所以最大应力值占许用应力的百分比越小, 则表明支链结构承载能力越强。因此, 计算所设计支链中的许用应力 $[\sigma]$ 如下, 一般的钢材的安全系数 k 取 1.5～2.5。

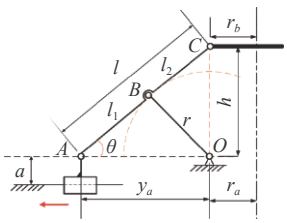
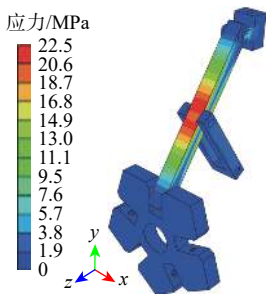
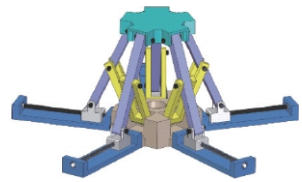
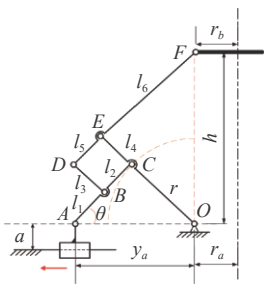
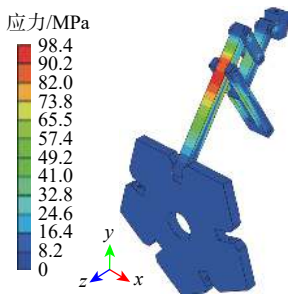
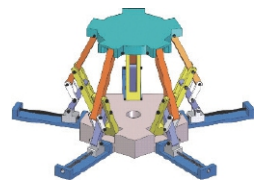
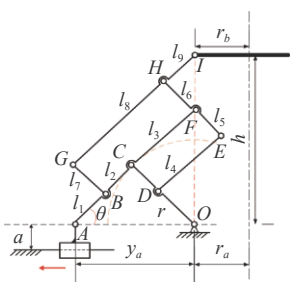
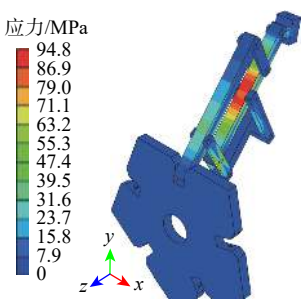
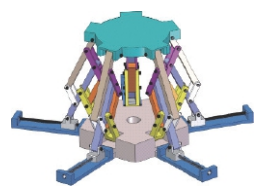
$$[\sigma]=\frac{\sigma_s}{k}=142\sim 236.67 \tag{13}$$

式中: σ_s为钢材的屈服强度, 根据表 11 取$\sigma_s=3.55\times 10^8\mathrm{N/m^2}$。安全系数取较大的值$k$=2.5, 则 $[\sigma]$=142 MPa, 进一步计算出各支链最大应力值占许用应力的百分比。由此, 可以比较 3 种可伸展支链构型的性能, 如表 12 所示。

由表 12 中的对比结果可知, 以 PRRR 支链为基本伸展支链的可伸展支链单元, 相较于其他 2 种可伸展支链, 其承载能力更强。相同条件下, 在 3 种支链结构中, PRRR 支链的主受力连杆的最大应力值占许用应力的百分比最低, 也进一步表明该支链结构的承载能力最强。

表 12 3 种可伸展支链构型的性能比较

Table 12 Stiffness comparison of three extendable branched chain configurations

基本支链单元	可伸展支链结构	可伸展支链应力云图	最大应力/ MPa	占[σ]百分比/%	伸展性能Θ	并联可伸展机构
PRRR 支链			22.5	15.8	1.215	
PRRR-3R 支链			98.4	69.3	1.890	
PRRR-6R 支链			94.8	66.8	1.822	

4 结 论

1）利用图论的方法进行可伸展支链单元的构型综合，得到多种可伸展支链，最后筛选出 PRRR 支链、PRRR-3R 支链及 PRRR-6R 支链这 3 种机构；

2）定义伸展性能指标——线伸展比，并对 3 种可伸展机构的伸展性能进行分析及比较，结果表明均具备大伸展能力；

3）对 3 种可伸展机构支链单元的刚度及承载能力进行有限元仿真分析及比较，结果表明均具备高刚度及大承载能力；

4）在综合比较 3 种可伸展机构的伸展性能及刚度之后，后续研究将以 PRRR 支链单元配置成并联可伸展机构，应用于需要提供大伸展及大刚度的场合，实现对机构运动空间的拓展。

参考文献（References）

[1] 刘欣, 王国庆, 李曙光, 等. 重型运载火箭关键制造技术发展展望[J]. 航天制造技术, 2013(1): 1-6.
LIU X, WANG G Q, LI S G, et al. Forecasts on crucial manufacturing technology development of heavy lift launch vehicle[J]. Aerospace Manufacturing Technology, 2013(1): 1-6(in Chinese).

[2] 周磊, 李新和, 俞大辉, 等. 球冠形薄壁封头在小减薄率工况下的失稳研究[J]. 锻压技术, 2016, 41(1): 25-31.
ZHOU L, LI X H, YU D H, et al. Research on instabilities of thin-walled spherical head at low thinning rate[J]. Forging & Stamping Technology, 2016, 41(1): 25-31(in Chinese).

[3] 朱云平, 张帆. 贮箱绝热层打磨机器人系统设计[J]. 上海工程技术大学学报, 2017, 31(1): 1-4.
ZHU Y P, ZHANG F. Design of grinding robot system for tank insulating layer[J]. Journal of Shanghai University of Engineering Science, 2017, 31(1): 1-4(in Chinese).

[4] 黄顺舟, 王力, 祁佩, 等. 低温贮箱隔热层打磨机器人的动力学仿真分析[J]. 环球市场信息导报, 2016, 37: 137-139.
HUANG S Z, WANG L, QI P, et al. Dynamic simulation analysis

of low temperature tank insulation polishing robot[J]. Global Market Information Guide, 2016, 37: 137-139(in Chinese).

[5] 党嘉强, 陶正瑞, 徐锦泱, 等. 聚氨酯泡沫绝热层高效精密打磨装置设计与建模[J]. 机床与液压, 2020, 48(2): 59-62.
DANG J Q, TAO Z R, XU J Y, et al. Design and modeling of high-efficiency and high-precision grinding device for polyurethane foam insulations[J]. Machine Tool & Hydraulics, 2020, 48(2): 59-62(in Chinese).

[6] 陶波, 赵兴炜, 丁汉. 大型复杂构件机器人移动加工技术研究[J]. 中国科学 (技术科学), 2018, 48(12): 1302-1312.
TAO B, ZHAO X W, DING H. Study on robotic mobile machining techniques for large complex components[J]. Scientia Sinica (Technologica), 2018, 48(12): 1302-1312(in Chinese).

[7] 李冰岩, 刘荣强, 从强, 等. 基于豆荚杆的三棱柱式可展开薄膜支撑臂设计与优化[J]. 机械工程学报, 2020, 56(7): 35-43.
LI B Y, LIU R Q, CONG Q, et al. Design and optimization of a triprism deployable membrane support arm using lenticular collapsible composite tubes[J]. Journal of Mechanical Engineering, 2020, 56(7): 35-43(in Chinese).

[8] 杨萍萍, 冯长凯. 可展开伸缩套管式伸展臂结构设计与分析[J]. 电子科技, 2008, 21(9): 8-11.
YANG P P, FENG C K. The structural design and analysis of the deployable[J]. Electronic Science and Technology, 2008, 21(9): 8-11(in Chinese).

[9] SHAN M H, GUO H W, LIU R Q, et al. Design and analysis of a triangular prism modular deployable mast[C]//2013 IEEE International Conference on Mechatronics and Automation. Piscataway: IEEE Press, 2013: 1546-1551.

[10] 高明星, 刘荣强, 李冰岩, 等. 空间可展开三棱柱伸展臂设计与优化[J]. 机械工程学报, 2020, 56(15): 129-137.
GAO M X, LIU R Q, LI B Y, et al. Design and optimization of space deployable tri-prism mast[J]. Journal of Mechanical Engineering, 2020, 56(15): 129-137(in Chinese).

[11] 马艳, 张群, 李锐明, 等. 可折展空间八转动副连杆捕获机构的设计[J]. 西安交通大学学报, 2020, 54(3): 179-187.
MA Y, ZHANG Q, LI R M, et al. Design and analysis of foldable capture mechanism based on spatial 8-rotation linkages[J]. Journal of Xi'an Jiaotong University, 2020, 54(3): 179-187(in Chinese).

[12] 刘仁志, 宋志佳, 高文杰, 等. 正八面体空间伸展臂的八面体单元结构分析[J]. 机械设计, 2013, 30(6): 72-75.
LIU R Z, SONG Z J, GAO W J, et al. Analysis of octahedron unit structure of regular octahedron space deployable arm[J]. Journal of Machine Design, 2013, 30(6): 72-75(in Chinese).

[13] 韩博, 许允斗, 姚建涛, 等. 双层环形桁架可展天线机构运动特性与动力学分析[J]. 兵工学报, 2020, 41(4): 810-821.
HAN B, XU Y D, YAO J T, et al. Kinematic characteristics and dynamics analysis of a double-ring truss deployable antenna mechanism[J]. Acta Armamentarii, 2020, 41(4): 810-821(in Chinese).

[14] ZHAO J S, CHU F L, FENG Z J. The mechanism theory and application of deployable structures based on SLE[J]. Mechanism and Machine Theory, 2009, 44(2): 324-335.

[15] 史创, 郭宏伟, 刘荣强, 等. 双层环形可展开天线机构构型优选及结构设计[J]. 宇航学报, 2016, 37(7): 869-878.
SHI C, GUO H W, LIU R Q, et al. Configuration optimization and structure design of the double-layer hoop deployable antenna mechanism[J]. Journal of Astronautics, 2016, 37(7): 869-878(in Chinese).

[16] ZHAO T S, WANG C, LIU X, et al. Stiffness and singularity analysis of foldable parallel mechanism for ship-based stabilized platform[J]. Robotica, 2016, 34(4): 913-924.

[17] 房海蓉, 陈江红. 一种新型四自由度并联机构设计与分析[J]. 北京交通大学学报, 2011, 35(4): 134-137.
FANG H R, CHEN J H. Structure synthesis and analysis of a novel four-degree-of-freedom parallel manipulator[J]. Journal of Beijing Jiaotong University, 2011, 35(4): 134-137(in Chinese).

[18] LI D, GUO S, CHEN Y. Kinematic performance and task planning analysis of a parameter reconfigurable parallel mechanism[C]//2019 IEEE International Conference on Robotics and Biomimetics (ROBIO). Piscataway: IEEE Press, 2019: 545-552.

[19] 邹琦, 曲海波, 郭盛. 一种三自由度可重构并联机构优化设计及性能分析[J]. 中国机械工程, 2018, 29(10): 1172-1178.
ZOU Q, QU H B, GUO S. Optimal design and performance analysis of a 3-DOF reconfigurable parallel mechanism[J]. China Mechanical Engineering, 2018, 29(10): 1172-1178(in Chinese).

[20] HUANG G Y, GUO S, ZHANG D, et al. Kinematic analysis and multi-objective optimization of a new reconfigurable parallel mechanism with high stiffness[J]. Robotica, 2018, 36(2): 187-203.

[21] 杨廷力. 机器人机构拓扑结构学[M]. 北京: 机械工业出版社, 2004.
YANG T L. Topology structure design of robot mechanisms[M]. Beijing: China Mechine Press, 2004 (in Chinese).

[22] 丁华锋, 黄真. 基于环路特性的运动链拓扑图及特征描述的自动生成[J]. 机械工程学报, 2007, 43(11): 40-43.
DING H F, HUANG Z. Automatic creation of topological graphs and the characteristic representations of kinematic chains based on the loop characteristics[J]. Chinese Journal of Mechanical Engineering, 2007, 43(11): 40-43(in Chinese).

Design and performance analysis of spatial large extension parallel mechanism

HE Litao，FANG Hairong*，CHEN Yufei，LI Yin

(School of Mechanical and Electronic Engineering，Beijing Jiaotong University，Beijing 100044，China)

Abstract：In the process of large-scale workpiece machining, the insufficient extension space of the executive mechanism is an urgent problem to be solved in the aerospace manufacturing industry. To increase the working space at the end of the executive mechanism, and to provide the mechanism with a stable and reliable extension foundation, a parallel extendable mechanism with high stiffness and large extension is proposed. This paper focuses on the design and analysis of parallel extensible mechanisms. Based on graph theory, the configuration of the extendable mechanism was synthesized, and the three extendable branch unit configurations, PRRR, PRRR-3R and PRRR-6R (P represents prismatic pair, R represents rotation pair) were obtained by using the configuration evolution. On this basis, the extensibility and stiffness of the branch chain unit of the three extensible mechanisms were compared and analyzed. And finally, the PRRR branch chain unit that satisfies the requirements was configured as a parallel extensible mechanism. The parallel extendable mechanism can be applied to occasions that need to provide large extension and large stiffness and realize the expansion of the movement space of other actuators.

Keywords：graph theory；parallel mechanism；extendable branched chain；performance analysis；stiffness；configuration synthesis

Received：2021-09-13；**Accepted**：2021-10-15；**Published Online**：2021-10-29 15：37

URL：kns.cnki.net/kcms/detail/11.2625.V.20211029.1011.001.html

Foundation item：Fundamental Research Funds for Central Universities (2018JBZ007)

* **Corresponding author**. E-mail：hrfang@bjtu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0554

超声波钻探器结构参数对输出特性的影响分析

张兴旺[1]，张明[2]，于登云[3,*]，曾婷[1]，殷参[1]，杨帅[1]，庞勇[1]

(1. 北京卫星制造厂有限公司，北京 100194；2. 中国空间技术研究院，北京 100194；
3. 中国航天科技集团有限公司，北京 100048)

摘 要：超声波钻探器因具有质量轻、低功耗、所需轴向力小、便于实时检测和原位分析及对象适应性强等特点，在地外天体科学钻探的新技术和新方法探索过程中备受瞩目。为进一步提高钻探器对不同钻进对象的适应性及其设计效率，通过研究超声波钻探器结构参数可知，其输出频率主要受碰撞冲击次数、弹簧钻杆的振动及钻杆本身的动力学特性等因素耦合影响。在此基础上，建立钻探器影响因素的物理模型，并结合实验研究了恢复弹簧刚度及初始预紧力和质量块质量对输出频率、位移和速度等参数的影响，得到弹簧初始预紧力对频率影响较为显著，弹簧刚度及质量块质量次之，为适应不同硬度对象高效钻进的变频超声波钻探器的优化设计提供技术参考。

关 键 词：超声波钻探器；结构；高速摄像系统；冲击碰撞；频率
中图分类号：V11
文献标志码：A **文章编号**：1001-5965(2023)07-1735-08

开展深空探测活动是人类探索宇宙奥秘、寻求长久发展的必然途径，是近地空间到更广阔太阳系空间的必然拓展[1-2]。在深空探测任务中，星体表面或表面下岩石及土壤样品的采集对研究星体环境、资源、地质构造及物质组成等具有重要意义[3]。而利用采样机构获取星体表层或表层下土壤及岩石样品，进而对其进行分析研究，是深空探测中的重要任务和关键技术[4]。

传统的旋转钻进技术因所需轴向力高、持续保持扭矩、功率消耗高、质量大、无法有效循环工作及钻头磨损严重等问题[5-7]难以在具有微弱引力环境及坚硬岩层的地外天体探测上得到更好的应用。而超声波钻探器因具有结构简单、体积小、质量轻、低功耗、所需轴向力较小、便于实时检测、不必考虑钻头磨损及强适应性等特点[8]，在探索适合深空探测的科学钻探新技术和新方法的过程中备受瞩目。为此，国内外学者和机构对超声波钻探器开展了大量研究。Bar-Cohen 等[8-9]针对火星采样探测任务设计了由压电驱动、变幅杆、自由质量块及钻杆等构成的超声波/声波钻探器(ultrasonic/sonic drilling/coring, USDC)，实现了低钻压低功耗钻进，拉开了超声波钻探器设计研究的序幕。随后，为研究钻探器驱动机制，Bao 等[10]对换能器、自由质量、钻头和岩石之间的相互作用进行建模和数值模拟分析，并通过对自由质量和变幅杆之间的相互作用进行冲击分析来推导自由质量块的最佳质量[11-12]。Sherrit 等[13]及 Zensheu 等[14]在 USDC 的研究基础上，对变幅杆构型进行设计分析，通过采用折叠变幅杆在满足钻进性能的条件下进一步减小了 USDC 的整体长度。郭俊杰等[15]针对月球及深空探测任务，率先设计了超声波/声波钻探器，并对钻探器的换能器和钻杆进行模态分析和谐响应分析。随后，陈超等[16]提出一种可用于外太空探测和取样的超声波/声波钻探采样器，建立了超声波钻探

收稿日期：2021-09-15；**录用日期**：2021-12-13；**网络出版时间**：2022-02-24 12:06
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220223.1540.003.html
基金项目：国家自然科学基金 (U1637208,U2013603)
*通信作者. E-mail：yudyun@sina.com

引用格式：张兴旺，张明，于登云，等. 超声波钻探器结构参数对输出特性的影响分析 [J]. 北京航空航天大学学报，2023，49（7）：1735-1742. ZHANG X W，ZHANG M，YU D Y，et al. Analysis of influence of ultrasonic drilling structure parameters on output characteristics [J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（7）：1735-1742（in Chinese）.

器的有限元模型，通过自由质量块设计实现了超声波-声波振动能量的耦合与转换，之后又通过高速摄影系统对自由质量块的运动规律进行分析[17]。毕亚兰[18]通过仿真和实验分析了自由质量块质量和运动范围对超声波钻探器钻进速度的影响。全齐全等[19-23]建立了自由质量块碰撞振动系统的离散模型，研究了其非线性动力学特性，在此基础上设计了冲击回转超声波钻探器。为研究变幅杆、自由质量块及钻杆间的相互作用，Vila 和 Malla[24-26]利用动量守恒原理，建立了三者的固体碰撞模型，分析了不同参数对三者间碰撞次数的影响。

目前对超声波钻探器的研究主要集中在结构构型设计、钻探器整体建模及钻探器钻进效率等方面，钻探器结构部件对输出频率等的影响研究较少。而空间采样对象等的成分不确定性及可钻性等级分布的广泛性，使得单一钻进频率很难较好地适应各种对象的高效钻进。因此，为提高钻探器适应性，需开展钻探器结构参数对频率等输出参数影响的研究。而钻杆的输出频率实际上是质量块与钻杆碰撞次数、钻杆和恢复弹簧的振动及钻杆本身的振动耦合作用的结果，很难用统一的数学模型准确描述。本文在上述超声波钻探器的基础上，通过建立系统各部分碰撞动力学模型，并结合实验研究了自由质量块质量、弹簧刚度及弹簧初始预紧力等参数对超声波钻探器输出频率等的影响，分析了频率等输出参数的影响规律，为适应空间不同对象高效钻进的频率可调的超声波钻探器的优化设计提供技术参考。

1 系统组成及工作原理

本文研究采用的超声波钻探器结构如图 1 所示。与文献 [19] 设计的钻探器类似，该钻探器主要由压电换能器、超声波发生器、自由质量块、恢复弹簧、钻杆及支撑壳等组成，其中压电换能器是系统的驱动部件，主要由预应力螺栓、后盖板、压电片叠合而成的压电叠堆及变幅杆等组成，其在超声波发生器产生的超高频正弦电压激励下，会产生 10 kHz 以上的超高频微幅振动。由于超声波产生的振动频率高、振幅小，破岩效率较低，因此，采用自由质量块和恢复弹簧系统来降低钻探器输出频率，将压电换能器产生的高频振动转化为声频或声频以下的冲击振动，同时增加振幅，并保持能量持续传递。钻杆是系统的能量输出部件，将自由质量块和恢复弹簧传递来的能量作用到钻进对象上，并通过持续的冲击使对象破碎。

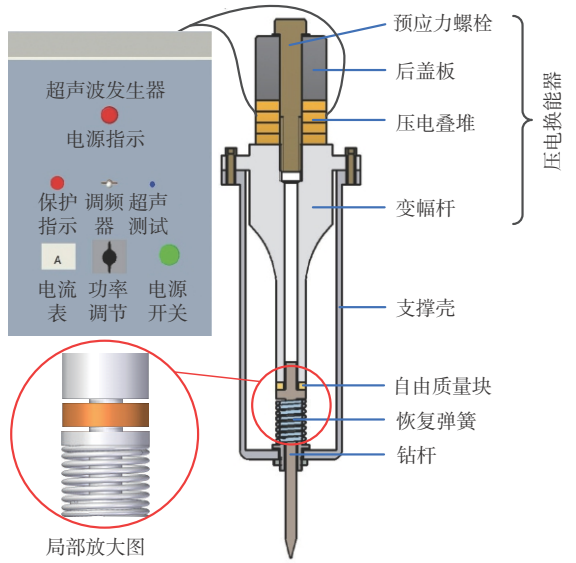


图 1 超声波钻探器结构组成

Fig. 1 Structural composition of ultrasonic drill

2 输出特性组成

变幅杆末端产生的超高频振动主要通过碰撞系统实现钻探器降频增幅输出，输出特性主要由碰撞系统中的质量块与钻杆碰撞次数、钻杆和恢复弹簧的振动及钻杆本身的振动等因素耦合而成。

碰撞系统主要包括变幅杆、自由质量块、恢复弹簧及钻杆，其碰撞系统模型如图 2 所示。

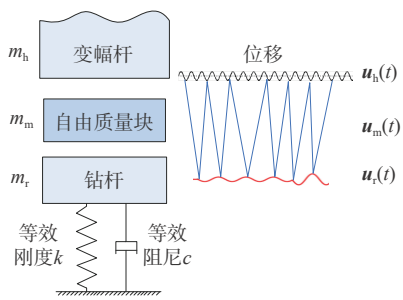


图 2 碰撞系统模型

Fig. 2 Model of collision system

变幅杆末端输出的高频振动可看作正弦运动，则有

$$\begin{cases} u_h = A_0 \sin(\omega t) \\ v_h = A_0 \omega \cos(\omega t) \end{cases} \tag{1}$$

式中：u_h为变幅杆末端位移；v_h为变幅杆末端速度；A_0为变幅杆末端振幅；ω为振动角速度。4 个参数值均可由激光测振仪测得。

对自由质量块，忽略其所受的摩擦力，对碰撞过程采用动量守恒定律可得

$$m_h v_h + m_m v_{m1} = m_h v'_h + m_m v'_{m1} \tag{2}$$

式中：m_h和m_m分别为变幅杆和自由质量块有效质

量；$\boldsymbol{v}_{\mathrm{m1}}$为自由质量块碰撞前速度；$\boldsymbol{v}'_{\mathrm{h}}$和$\boldsymbol{v}'_{\mathrm{m1}}$分别为变幅杆和自由质量块碰撞后速度，其值可由高速摄像机测得。

引入变幅杆与自由质量块碰撞恢复系数[23]e_1，即

$$e_1=\frac{\left|\boldsymbol{v}'_{\mathrm{m1}}-\boldsymbol{v}'_{\mathrm{h}}\right|}{\left|\boldsymbol{v}_{\mathrm{h}}-\boldsymbol{v}_{\mathrm{m1}}\right|} \tag{3}$$

将式(3)代入式(2)整理可得

$$\boldsymbol{v}'_{\mathrm{m1}}=\frac{m_{\mathrm{h}}\boldsymbol{v}_{\mathrm{h}}+m_{\mathrm{m}}\boldsymbol{v}_{\mathrm{m1}}+m_{\mathrm{h}}e_1(\boldsymbol{v}_{\mathrm{h}}-\boldsymbol{v}_{\mathrm{m1}})}{m_{\mathrm{h}}+m_{\mathrm{m}}} \tag{4}$$

上述碰撞完成后，自由质量块以速度$\boldsymbol{v}'_{\mathrm{m1}}$向钻杆运动，撞击钻杆，从而形成轴向冲击运动，恢复弹簧使得系统连续响应。

对自由质量块和钻杆，利用动量守恒并引入恢复系数，可得碰撞后自由质量块和钻杆的速度：

$$\begin{cases}\boldsymbol{v}'_{\mathrm{m2}}=\dfrac{m_{\mathrm{r}}\boldsymbol{v}_{\mathrm{r}}+m_{\mathrm{m}}\boldsymbol{v}_{\mathrm{m2}}+m_{\mathrm{r}}e_2(\boldsymbol{v}_{\mathrm{r}}-\boldsymbol{v}_{\mathrm{m2}})}{m_{\mathrm{r}}+m_{\mathrm{m}}}\\ \boldsymbol{v}'_{\mathrm{r}}=\dfrac{m_{\mathrm{r}}\boldsymbol{v}_{\mathrm{r}}+m_{\mathrm{m}}\boldsymbol{v}_{\mathrm{m2}}+m_{\mathrm{m}}e_2(\boldsymbol{v}_{\mathrm{m2}}-\boldsymbol{v}_{\mathrm{r}})}{m_{\mathrm{r}}+m_{\mathrm{m}}}\end{cases} \tag{5}$$

式中：m_{r}为钻杆质量；$\boldsymbol{v}_{\mathrm{m2}}$和$\boldsymbol{v}_{\mathrm{r}}$分别为自由质量块和钻杆碰撞前速度；$\boldsymbol{v}'_{\mathrm{m2}}$和$\boldsymbol{v}'_{\mathrm{r}}$分别为自由质量块和钻杆碰撞后速度；$e_2=\dfrac{\left|\boldsymbol{v}'_{\mathrm{r}}-\boldsymbol{v}'_{\mathrm{m2}}\right|}{\left|\boldsymbol{v}_{\mathrm{m2}}-\boldsymbol{v}_{\mathrm{r}}\right|}$，为自由质量块与钻杆的碰撞恢复系数。

质量块与钻杆碰撞完成后，钻杆和恢复弹簧运动可看作单自由度系统，则

$$m_{\mathrm{r}}\ddot{\boldsymbol{u}}_{\mathrm{r}}+c\dot{\boldsymbol{u}}_{\mathrm{r}}+k(\boldsymbol{u}_{\mathrm{r}}+\boldsymbol{u}_0)=\boldsymbol{0} \tag{6}$$

式中：$\boldsymbol{u}_{\mathrm{r}}$为钻杆位移；$c$为系统等效阻尼；$k$为系统等效刚度；$\boldsymbol{u}_0$为弹簧初始压缩量(常数)。

在自由质量和钻杆 2 次碰撞期间，钻杆会经历自由振动。将钻杆看作是由一种材料制成的均匀悬臂杆，则钻杆轴向自由振动的运动方程为

$$\frac{\partial^2\boldsymbol{u}_{\mathrm{r}}}{\partial\boldsymbol{x}^2}-\frac{\partial^2\boldsymbol{u}_{\mathrm{r}}}{v_{\mathrm{c}}^2\partial t^2}=0 \tag{7}$$

式中：$\boldsymbol{x}$为钻杆上任意一点到钻杆顶端的距离；$v_{\mathrm{c}}=\sqrt{E/\rho}$，为波在杆中的传播速度，$E$为钻杆材料弹性模量，$\rho$为钻杆材料密度。

式 (7) 的解可通过分离变量获得，即将其分为时间相关分量$u(t)$和位置分量$\boldsymbol{\varphi}(x)$，则解的形式可写为

$$\boldsymbol{u}_{\mathrm{r}}(\boldsymbol{x},t)=u(t)\boldsymbol{\varphi}(x) \tag{8}$$

将式 (8) 代入式 (7)，并分别对时间和位移求偏导可得通解为

$$\begin{cases}u(t)=C_1\sin(\omega_{\mathrm{n}}t)+C_2\cos(\omega_{\mathrm{n}}t)\\ \boldsymbol{\varphi}(\boldsymbol{x})=C_3\sin(\omega_{\mathrm{n}}\boldsymbol{x}/c)+C_4\cos(\omega_{\mathrm{n}}\boldsymbol{x}/c)\end{cases} \tag{9}$$

式中：C_1、C_2、C_3、C_4为常数；ω_{n}为钻杆轴向固有频率。

代入边界条件整理可得钻杆的自由振动响应为

$$\boldsymbol{u}_{\mathrm{r}}(\boldsymbol{x},t)=[D_1\sin(\omega_{\mathrm{n}}t)+D_2\cos(\omega_{\mathrm{n}}t)]\cos(\omega_{\mathrm{n}}\boldsymbol{x}/v_{\mathrm{c}}) \tag{10}$$

式中：D_1、D_2为常数，由初始条件获得。

将第 i 次碰撞后钻杆的位移和速度作为自由振动的初始条件，即$\boldsymbol{u}_{\mathrm{r}}(0,0)=\boldsymbol{u}_{ri}$，$\boldsymbol{v}_{\mathrm{r}}(0,0)=\boldsymbol{v}_{ri}$，代入式(10)可得

$$\boldsymbol{u}_{\mathrm{r}}(\boldsymbol{x},t)=\left[\frac{|\boldsymbol{v}_{ri}|}{\omega_{\mathrm{n}}}\sin(\omega_{\mathrm{n}}t)+|\boldsymbol{u}_{ri}|\cos(\omega_{\mathrm{n}}t)\right]\cos\left(\frac{\omega_{\mathrm{n}}}{v_{\mathrm{c}}}\boldsymbol{x}\right) \tag{11}$$

则由自由质量块冲击引起的考虑阻尼的钻杆的动力响应[25-26]可表示为

$$\boldsymbol{u}_{\mathrm{r}}(t)=\mathrm{e}^{-\xi\omega_{\mathrm{n}}t}\left[\frac{\boldsymbol{v}_{ri}+\boldsymbol{u}_{ri}\xi\omega_{\mathrm{n}}}{\omega_{\mathrm{d}}}\sin(\omega_{\mathrm{d}}t)+\boldsymbol{u}_{ri}\cos(\omega_{\mathrm{d}}t)\right] \tag{12}$$

式中：ξ为系统阻尼比；ω_{d}为带阻尼的系统固有频率。

则碰撞过程中钻杆的动能 T 为

$$T=\frac{\rho A}{2}\int_0^l\left(\frac{\partial\boldsymbol{u}_{\mathrm{r}}}{\partial t}\right)^2\mathrm{d}\boldsymbol{x}=\frac{\rho A}{2}\int_0^l[\dot{\boldsymbol{\psi}}^2(1-\boldsymbol{x}/l)^2]\mathrm{d}\boldsymbol{x}=\frac{m_{\mathrm{reff}}\dot{\boldsymbol{\psi}}^2}{2} \tag{13}$$

式中：A为钻杆横截面积；l为钻杆长度；$\boldsymbol{\psi}$为钻杆广义动态轴向变形；m_{reff}为钻杆有效质量。

相应的钻杆应变能 V 为

$$V=\frac{AE}{2}\int_0^l\left(\frac{\partial\boldsymbol{u}_{\mathrm{r}}}{\partial\boldsymbol{x}}\right)^2\mathrm{d}\boldsymbol{x}=\frac{EA\boldsymbol{\psi}^2}{2l} \tag{14}$$

则碰撞过程的拉格朗日方程为

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial T}{\partial\dot{\boldsymbol{\psi}}}\right)-\frac{\partial T}{\partial\boldsymbol{\psi}}+\frac{\partial V}{\partial\boldsymbol{\psi}}=\boldsymbol{F}_{ci} \tag{15}$$

式中：$\boldsymbol{F}_{ci}$为第 i 次碰撞时产生的接触力。

将式(13)和式(14)代入式(15)整理可得

$$\boldsymbol{F}_{ci}=m_{\mathrm{reff}}\ddot{\boldsymbol{\psi}}+\frac{EA}{l}\boldsymbol{\psi} \tag{16}$$

结合自由质量块的运动，整理分析可得

$$\ddot{\boldsymbol{\psi}}+\omega_{\mathrm{eff}}^2\boldsymbol{\psi}=\frac{m_{\mathrm{m}}\boldsymbol{g}}{(m_{\mathrm{m}}+m_{\mathrm{reff}})} \tag{17}$$

式中：$\omega_{\mathrm{eff}}=\sqrt{EA/[(m_{\mathrm{m}}+m_{\mathrm{reff}})l]}$为有效运动角频率。

引入边界条件，则式(17)的通解为

$$\boldsymbol{\psi}=\frac{m_{\mathrm{m}}\boldsymbol{g}}{(m_{\mathrm{m}}+m_{\mathrm{reff}})\omega_{\mathrm{eff}}^2}[1-\cos(\omega_{\mathrm{eff}}t)]+\frac{m_{\mathrm{m}}\boldsymbol{v}_{\mathrm{m0}}}{(m_{\mathrm{m}}+m_{\mathrm{reff}})\omega_{\mathrm{eff}}}\sin(\omega_{\mathrm{eff}}t) \tag{18}$$

式中：$\boldsymbol{v}_{\mathrm{m0}}$为所选取碰撞过程初始时刻自由质量块对应的速度。

对式(18)求导并赋值为 0 可得到最大碰撞接触力所对应的时间$t_{\max}$为

$$t_{\max}=\frac{|\arctan(\boldsymbol{v}_{\mathrm{m0}}\omega_{\mathrm{eff}}/\boldsymbol{g})|}{\omega_{\mathrm{eff}}} \tag{19}$$

将碰撞过程看作是弹性碰撞，则整个碰撞过程

的接触时间相应为 $2t_{max}$。

3 实验测量系统及结果分析

通过分析可知，钻杆输出频率等参数与冲击次数、系统阻尼、固有频率、碰撞速度和位移等参数有关。为进一步研究冲击钻探器结构参数对频率等的影响，本文设计了高速摄像实验系统，以测量钻杆输出参数。

3.1 实验测量系统

自由质量块及恢复弹簧的引入降低了钻杆最终的输出频率，增加了钻杆输出振幅，文献 [22] 通过建模和仿真分析了系统碰撞持续时间及碰撞过程能量和位移传递规律，但并未明确给出自由质量块、弹簧刚度及弹簧预紧力对钻杆最终输出频率和振幅的影响。为探究其影响，本文搭建了相应的实验系统，通过高速摄像原理，测得了不同条件下钻杆的输出。实验系统主要包括高速摄像系统、钻探器及其固定和控制系统等组成，如图 3 所示。

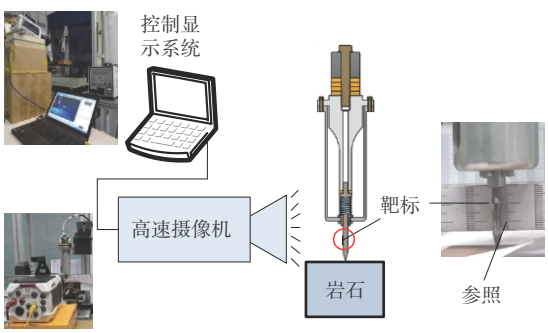


图 3 实验测量系统

Fig. 3 Experimental measurement system

高速摄像系统主要包括高速摄像机、控制显示系统、靶标及参照等。高速摄像机用来采集钻探器输出的冲击振动数据；控制显示系统用来设定高速摄像机采集参数、采集数据的显示、处理及储存等；靶标及参照用作运动分析点的选取及实际距离参照，其中靶标贴在钻杆上，用来反映钻杆运动，参照与地面保持静止，上面标有实际长度刻度，用来作为靶标运动距离的参照物。数据存储完成后利用后处理软件选取数据中靶标上的点及参照点，并对其进行运动分析，最终得到靶标上的点的运动参数，即钻杆的运动参数。

3.2 结果分析讨论

通过 3.1 节实验测量系统，采用控制单一变量法，分别测量了不同弹簧刚度、预紧力及质量块质量条件下对应的钻探器冲击频率、位移和速度等输出参数。以弹簧刚度为 35.5 N/mm，初始预紧力为 120 N，自由质量块质量为 6.06 g 为例，其位移与速度的时域和频域曲线如图 4 和图 5 所示。

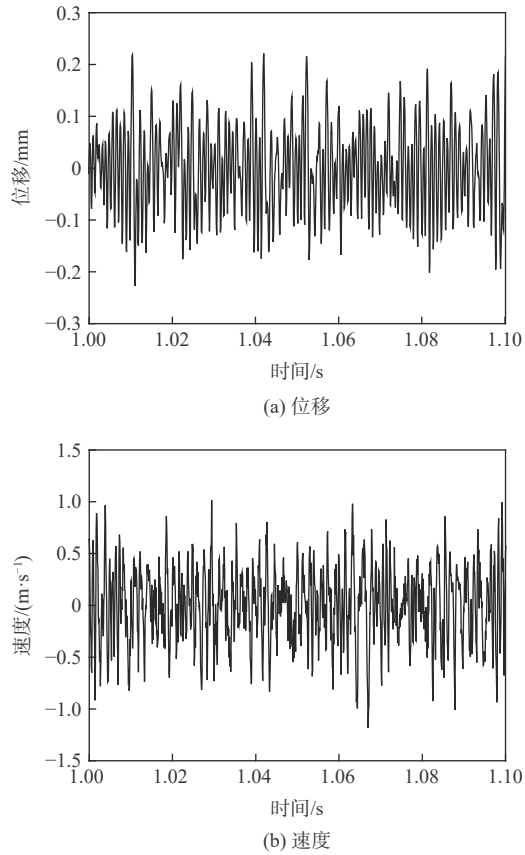


图 4 位移和速度时域曲线

Fig. 4 Time domain curves of displacement and velocity

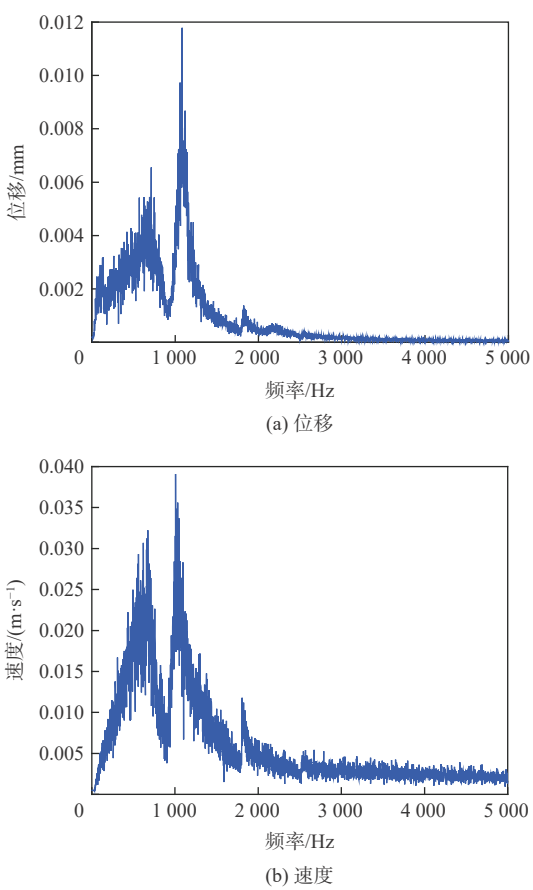


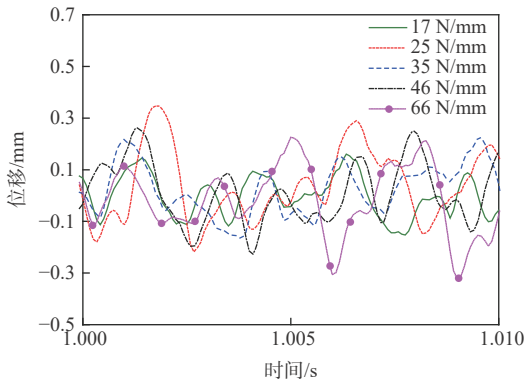
图 5 位移和速度频域曲线

Fig. 5 Frequency domain curves of displacement and velocity

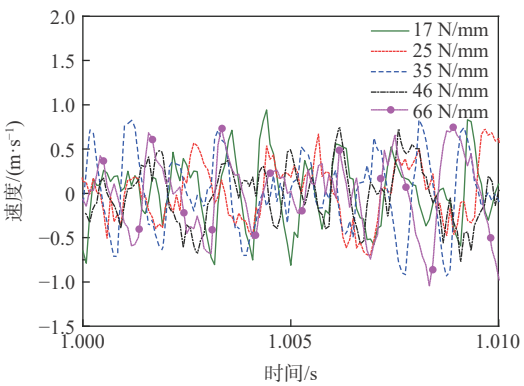
如图 4 所示，由于碰撞时间和速度等的变化，使得位移和速度的幅值具有一定的波动。由图 5 可知，位移和速度的频域变化曲线大致有 3 个峰值点，分别在 700 Hz、1000 Hz 和 1800 Hz 左右，分析其为冲击次数，钻杆与弹簧系统振动及钻杆本身的固有动力学特性作用的结果。

在此基础上，在其他变量相同的情况下，通过实验测量系统分别得到不同弹簧刚度、不同预紧力及不同质量块质量所对应的频率、位移和速度等的变化曲线，如图 6～图 8 所示。

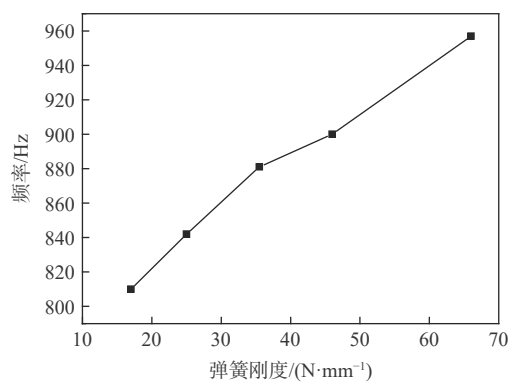
由图 6～图 8 分析可知，其他参数不变的情况下，在实验分析范围内：①随着恢复弹簧刚度的增加，钻探器钻杆输出位移和速度没有明显变化。由于刚度增加，使得碰撞后弹簧的回复力增加，碰撞次数略有增加，但增加幅度较小。此外，由于刚度增加，弹簧钻杆振动系统有效刚度增大，从而使得输出频率大致呈线性趋势增加。②随着恢复弹簧初始预紧力的增加，在输入能量相同的情况下，钻杆向下运动的阻力增大，钻杆输出位移和速度均相应减小，其中位移减小较为显著。初始预紧力的增加使得钻杆运动回复加速度增大，使得碰撞接触时间降低，钻杆运动范围降低，冲击碰撞次数增加，进而使得输出频率大致呈线性趋势增加，且增加幅度较大。③随着自由质量块质量的增加，钻探器钻杆



(a) 位移



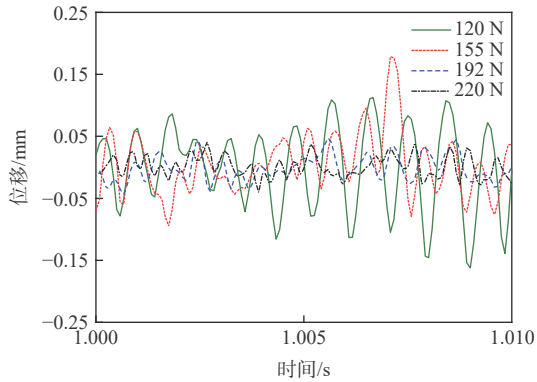
(b) 速度



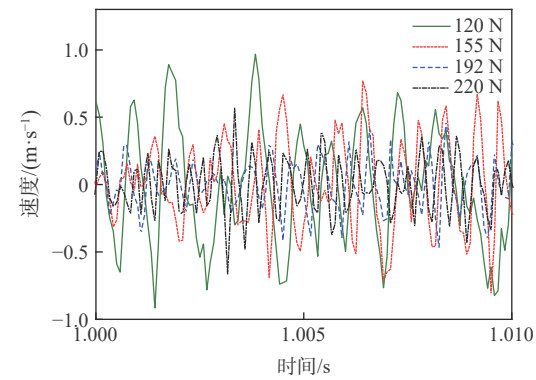
(c) 频率

图 6　不同弹簧刚度对应的钻探器参数变化曲线

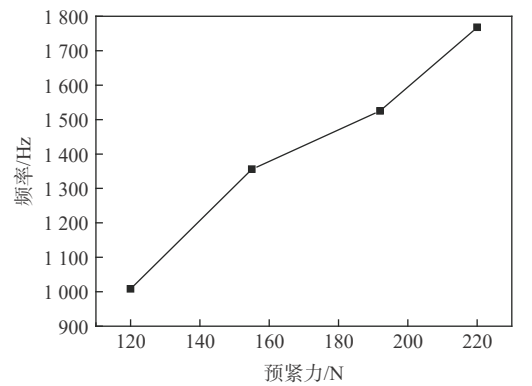
Fig. 6　Variation curves of drill parameters corresponding to different spring stiffness



(a) 位移



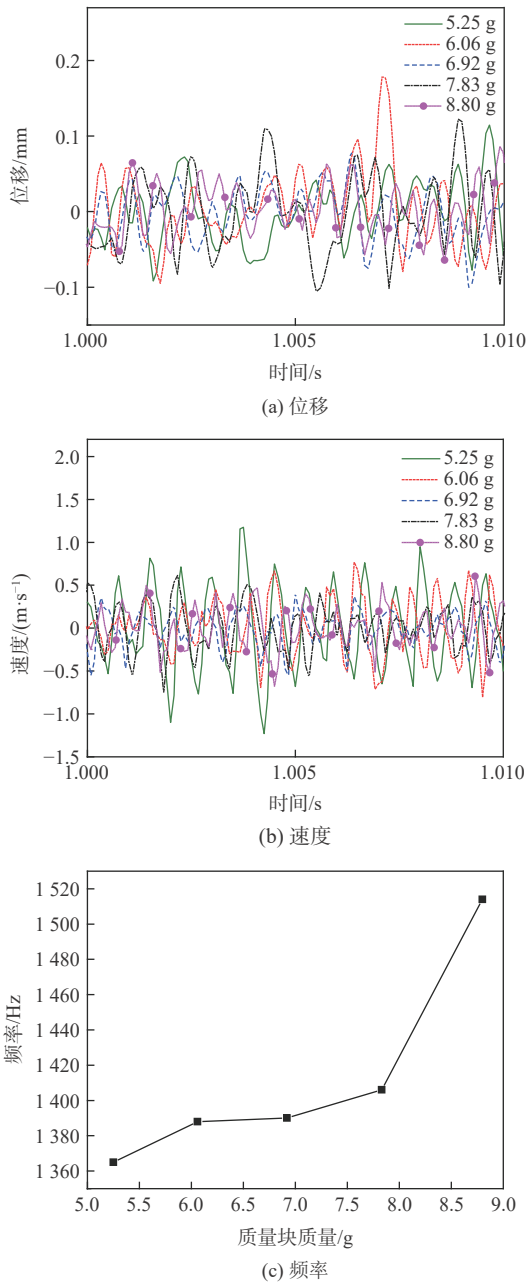
(b) 速度



(c) 频率

图 7　不同初始预紧力对应参数变化曲线

Fig. 7　Variation curves of drill parameters corresponding to different initial preloads



(a) 位移

(b) 速度

(c) 频率

图 8 不同质量块质量对应参数变化曲线

Fig. 8 Variation curves of drill parameters corresponding to different free mass block masses

的输出位移和速度略有降低，有效运动角频率降低，进而使得碰撞接触时间相应减小，同时由于位移等的降低，使得碰撞次数有所增加，且呈现逐渐加快的趋势，进而使得频率也呈现先缓慢再快速的增加趋势。

4 结 论

1）系统碰撞后的位移及速度等参数，影响着自由质量块传递到钻杆的冲击次数及钻杆输出频率等参数。

2）弹簧刚度对钻探器的位移和速度影响不明显，使频率大致呈线性增加。

3）初始预紧力使输出位移和速度均减小，频率大致呈线性增加，且影响较为显著。

4）质量块质量增加，钻探器钻杆的输出位移和速度略有降低，频率呈先缓慢再快速的增加趋势，影响幅度与刚度影响类似，均小于初始预紧力影响。

参考文献（References）

[1] 于登云, 孙泽洲, 孟林智, 等. 火星探测发展历程与未来展望[J]. 深空探测学报, 2016, 3(2): 108-113.
YU D Y, SUN Z Z, MENG L Z, et al. The development process and prospects for Mars exploration[J]. Journal of Deep Space Exploration, 2016, 3(2): 108-113(in Chinese).

[2] 于登云, 张兴旺, 张明, 等. 小天体采样探测技术发展现状及展望[J]. 航天器工程, 2020, 29(2): 1-10.
YU D Y, ZHANG X W, ZHANG M, et al. Development status and prospect of small celestial body sampling exploration techniques[J]. Spacecraft Engineering, 2020, 29(2): 1-10(in Chinese).

[3] 吴伟仁, 于登云. 深空探测发展与未来关键技术[J]. 深空探测学报, 2014, 1(1): 5-17.
WU W R, YU D Y. Development of deep space exploration and its future key technologies[J]. Journal of Deep Space Exploration, 2014, 1(1): 5-17(in Chinese).

[4] 刘志全, 庞彧, 李新立. 深空探测自动采样机构的特点及应用[J]. 航天器工程, 2011, 20(3): 120-125.
LIU Z Q, PANG Y, LI X L. Characteristics and applications of automatic sampling mechanisms for deep space exploration[J]. Spacecraft Engineering, 2011, 20(3): 120-125(in Chinese).

[5] FRANCA L F R. A bit–rock interaction model for rotary–percussive drilling[J]. International Journal of Rock Mechanics and Mining Sciences, 2011, 48(5): 827-835.

[6] TALALAY P G. Subglacial till and bedrock drilling[J]. Cold Regions Science and Technology, 2013, 86: 142-166.

[7] WANG J S, CAO P L, LIU C P, et al. Comparison and analysis of subglacial bedrock core drilling technology in Polar Regions[J]. Polar Science, 2015, 9(2): 208-220.

[8] BAR-COHEN Y, SHERRIT S, DOLGIN B P, et al. Ultrasonic/sonic drilling/coring (USDC) for planetary applications[C]//Proceedings of the SPIE—The International Society for Optical Engineering. Newport Beach: SPIE, 2001: 441-448.

[9] BAR-COHEN Y, SHERRIT S, DOLGIN B P, et al. Ultrasonic/sonic driller/corer (USDC) as a sampler for planetary exploration[C]//2001 IEEE, IEEE Aerospace Conference Proceedings. Piscataway: IEEE Press, 2001: 263-271.

[10] BAO X Q, BAR-COHEN Y, CHANG Z, et al. Modeling and computer simulation of ultrasonic/sonic driller/corer (USDC)[J]. IEEE Transactions of Ultrasonics, Ferroelectrics, and Frequency Control, 2003, 50(9): 1147-1160.

[11] BAO X Q, BAR-COHEN Y, CHANG Z, et al. Ultrasonic/sonic impacting penetrators[J]. NASA Tech Briefs, 2008, 32(4): 58-60.

[12] BAR-COHEN Y, SHERRIT S. Self mountable and extractable ultrasonic/sonic anchor: America, US7156189[P]. 2007-01-02.

[13] SHERRIT S, BADESCU M, BAO X Q, et al. Novel horn designs for power ultrasonics[C]//IEEE Ultrasonics Symposium. Piscat-

away: IEEE Press, 2004: 2263-2266.

[14] ZENSHEU C, STEWART S, MIRCEA B, et al. Design and analysis of ultrasonic actuator in consideration of length-reduction for a USDC (ultrasonic/sonic driller/corer)[C]//Proceedings of the SPIE Smart Structures Conference. Newport Beach: SPIE, 2005, 5762: 72-79.

[15] 郭俊杰, 黄卫清, 李志荣. 一种新型的超声波/声波钻探器[J]. 压电与声光, 2008, 30(5): 579-581.
GUO J J, HUANG W Q, LI Z R. A new ultrasonic/sonic drilling device[J]. Piezoelectrics&Acoustooptics, 2008, 30(5): 579-581(in Chinese).

[16] 陈超, 黄卫清, 郭俊杰, 等. 深空探测用超声波/声波钻探采样器的研究[C]//第三届全国压电和声波理论及器件技术研讨会论文集, 南京: [出版者不详], 2008: 338-341.
CHEN C, HUANG W Q, GUO J J, et al. Study on ultrasonic/sonic driller/corer for deep space exploration[C]//Proceedings of 3th Symposium on Piezoelectricity, Acoustic Waves, and Device Applications. Nanjing: [s.n.], 2008: 338-341(in Chinese).

[17] 陈超, 杨康. 超声波/声波钻探器的设计与试验[J]. 振动、测试与诊断, 2013, 33(2): 252-257.
CHEN C, YANG K. Design and text of ultrasonic/sonic drilling rig[J]. Journal of Vibration, Measurement & Diagnosis, 2013, 33(2): 252-257(in Chinese).

[18] 毕亚兰. 新型超声钻的动力特性研究[D]. 太原: 太原理工大学, 2016.
BI Y L. Study on dynamic characteristics of a new ultrasonic drill[D]. Taiyuan: Taiyuan University of Technology, 2016(in Chinese).

[19] 李贺, 全齐全, 王鑫剑, 等. 一种基于压电驱动的火星岩石钻探器的研制[J]. 深空探测学报, 2016, 3(2): 156-161.
LI H, QUAN Q Q, WANG X J, et al. Development of a piezoelectric-driven rock drilling device on Mars[J]. Journal of Deep Space Exploration, 2016, 3(2): 156-161(in Chinese).

[20] BAI D E, QUAN Q Q, WANG Y C, et al. A longitudinal & longitudinal-torsional vibration actuator for rotary-percussive ultrasonic planetary drills[J]. Advances in Space Research, 2018, 63(2): 1065-1072.

[21] WANG Y C, QUAN Q Q, YU H G, et al. Rotary-percussive ultrasonic drill: An effective subsurface penetrating tool for minor planet exploration[J]. IEEE Access, 2018, 6: 37796-37806.

[22] 王印超, 全齐全, 于红英, 等. 一种单压电叠堆驱动的回转冲击超声波钻[J]. 北京航空航天大学学报, 2018, 44(9): 1850-1859.
WANG Y C, QUAN Q Q, YU H Y, et al. A rotary-percussive ultrasonic drill driven by single piezoelectric stack[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(9): 1850-1859(in Chinese).

[23] 王童照, 全齐全, 黄江川, 等. 地外天体超声波钻探器研究与应用综述[J]. 宇航学报, 2021, 42(4): 397-408.
WANG T Z, QUAN Q Q, HUANG J C, et al. Overview of research and application of ultrasonic drills for extraterrestrial objects[J]. Journal of Astronautics, 2021, 42(4): 397-408(in Chinese).

[24] VILA L, MALLA R B. Analytical model of the contact interaction between the components of a special percussive mechanism for planetary exploration[J]. Acta Astronautica, 2016, 118: 158-167.

[25] VILA L, MALLA R B. Dynamic and contact analysis of a special percussive mechanism for planetary subsurface exploration[C]//Proceedings of the 54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. Reston: AIAA, 2013.

[26] VILA L, MALLA R B. Analysis of the hammering mechanism of a special percussive system for planetary exploration[C]//55th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. Reston: AIAA, 2014.

Analysis of influence of ultrasonic drilling structure parameters on output characteristics

ZHANG Xingwang[1], ZHANG Ming[2], YU Dengyun[3,*], ZENG Ting[1], YIN Shen[1], YANG Shuai[1], PANG Yong[1]

(1. Beijing Spacecrafts, Beijing 100194, China;
2. China Academy of Space Technology, Beijing 100194, China;
3. China Aerospace Science and Technology Corporation, Beijing 100048, China)

Abstract: The ultrasonic drill has the characteristics of lightweight, low power consumption, the small axial force required, easy real-time detection and in-situ analysis, and good adaptability to different objects. It is highly anticipated in exploring new scientific drilling technologies and new methods suitable for deep space exploration. In order to improve the adaptability of the drill to different drilling objects and its design efficiency further, the influence of the structural parameters of the ultrasonic driller is studied. The frequency is mainly considered to be affected by the number of collisions, the vibration of the spring and drilling rod, and the vibration of the drilling rod itself. On this basis, a physical model of the influencing factors was established. Experiments are used to examine how the spring stiffness, initial preload, and mass of the free masses affect the output frequency, displacement, and speed. The effect that the initial preloads of the spring have on the frequency is more significant, followed by spring stiffness and mass of the free masses. Finally, a technical reference for the optimization design of variable frequency ultrasonic drills suitable for high-efficiency drilling of objects with different hardness is provided.

Keywords: ultrasonic drill; structure; high-speed camera system; impact collision; frequency

Received: 2021-09-15; **Accepted**: 2021-12-13; **Published Online**: 2022-02-24 12:06
URL: kns.cnki.net/kcms/detail/11.2625.V.20220223.1540.003.html
Foundation items: National Natural Science Foundation of China (U1637208,U2013603)
* **Corresponding author**. E-mail: yudyun@sina.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0557

基于行人姿态的轨迹预测方法

王瑞平[1]，宋晓[2,*]，陈凯[3]，龚开奇[2]，张峻凡[4]

(1. 北京航空航天大学 自动化科学与电气工程学院，北京 100191；2. 北京航空航天大学 网络空间安全学院，北京 100191；
3. 南京航空航天大学 机电学院，南京 211106；4. 北京航空航天大学 宇航学院，北京 100191)

摘 要：在自动驾驶领域，行人轨迹预测一直是研究热点之一，行人行为的不确定性给轨迹预测带来很大的挑战。目前大部分轨迹预测方法只专注于行人之间的信息交互，忽略了行人意图和场景中其他语义信息对行人轨迹的影响。为此，提出一种基于行人姿态的卷积编码器-解码器网络(PKCEDN)来预测目标行人轨迹的方法，所提方法包含基于卷积、长短时记忆（LSTM）网络的编码器-解码器模型和能够学习当前时刻与过去时刻轨迹相关性的注意力机制。所提方法在MOT16、MOT17和MOT20公开数据集上进行了相关测试，与Linear、LSTM、Social-LSTM、Social-生成对抗网络（GAN）、SR-LSTM和M_{sgtv}等主流方法相比，在保证预测速度不降低的前提下，平均误差降低约36%。

关 键 词：行人轨迹预测；姿态提取；编码器-解码器；注意力机制；空间语义信息
中图分类号：V221^{+}.3; TB553
文献标志码：A **文章编号**：1001-5965(2023)07-1743-12

无人自动驾驶正在大力发展，行人轨迹预测问题是其中重要的组成部分，如果能够提前准确地对行人轨迹进行预测，自动驾驶系统能够使车辆事先对行人的行为做出正确的决策，促进自动驾驶、无人车配送等领域的发展，具有很好的应用价值[1-3]。因此，在计算机视觉领域，轨迹预测[3]越来越成为研究者关注的研究方向。

近些年，随着人工智能的发展，研究人员逐渐用神经网络[4-5]的方法预测行人轨迹，主要包括常规卷积网络和图卷积网络。这种预测一般是根据行人过去时刻的轨迹，结合周围环境对其的影响，预测行人未来时刻的行走轨迹。目前，该领域主要的关注点是如何充分利用丰富的视觉信息辅助行人轨迹的预测。

行人轨迹是行人与周围环境综合交互的结果。行人未来轨迹主要取决于行人过去时刻的轨迹，与周围人和车等障碍物的交互，与周围自然环境的交互[6-8]。基于此，构建行人轨迹预测模型具有较高的挑战性。目前，行人轨迹预测方法主要考虑2个方面：①行人意图。现有的行人轨迹预测方法很少考虑行人的意图。在文献[6]中，将行人相关活动分为29种，并且将行人未来活动和轨迹预测结合起来。但行人的实际意图有很多种，是行人自我的主观意识，很难去预测其真实活动目的。②时间信息。目标行人的未来轨迹首先会受到之前时刻轨迹的影响。现有的方法[6-10]利用长短时记忆(long short-term memory, LSTM[11])网络进行建模，通过其记忆单元来表征目标行人每个时刻的运动状态或与其周围空间交互特征。虽然LSTM具有强大的记忆和遗忘功能，但其局限于对前一时刻的状态记忆，对较远的过去时刻的状态记忆非常少。在实际的行人轨迹中，目标行人在过去各个时刻的轨迹状态对当前时刻的行走状态有着不同的影响[10]。

为了解决这些问题，本文从以下2个方面开展

收稿日期：2021-09-26；**录用日期**：2022-01-02；**网络出版时间**：2022-02-24 14:37
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220223.1541.004.html
基金项目：国家重点研发计划(2018YFB1702703)
*通信作者. E-mail：songxiao@buaa.edu.cn

引用格式：王瑞平，宋晓，陈凯，等. 基于行人姿态的轨迹预测方法[J]. 北京航空航天大学学报，2023，49(7)：1743-1754.
WANG R P，SONG X，CHEN K，et al. Pedestrian trajectory prediction method based on pedestrian pose[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1743-1754 (in Chinese).

创新性研究工作：①结合实际场景中最常见的13种异构交通信息，本文提出一种通过学习行人姿态及周围异构交通信息进行行人轨迹预测的方法。②提出通过注意力机制建立目标行人当前时刻轨迹与过去时刻轨迹的相关性，设计了一种基于行人姿态的卷积编码器-解码器网络（pose keypoints based convolutional encoder-decoder network, PKCEDN）来预测目标行人轨迹的方法。

PKCEDN在MOT16数据集上进行训练和测试，与现有的轨迹预测方法比较，本文方法有更好的预测精度。

1 相关工作

现有的行人轨迹预测方法主要有2类：基于机理建模的力学模型[12-15]和数据驱动的神经网络模型[6,16-18]。在力学建模方面，主流的方法为：社会力模型[12,14]、智能体模型[19-21]等。基于经验构建的力学模型需要人为设计人与周围环境交互，比如社会力模型中构建了人与人、人与物之间的力学关系。深度学习模型基于大量行人轨迹数据去训练神经网络自动构建场景中行人与其他障碍物的交互关系。由于人具有主观意识，并且所处环境错综复杂，要想全面建立行人与环境交互的力学关系是不可能的。而神经网络模型能够很好的自我学习，经过训练能够自我感知这种相互作用的关系[22]。所以在预测精度方面，基于深度学习的神经网络已经大大超越了力学建模。因此，本文主要介绍基于深度学习的模型。

基于人与人交互的轨迹预测模型。之前的部分工作都是基于人与人之间的交互进行建模，比较经典的方法包括：Social-LSTM[7]、Social-生成对抗网络[23]（generative adversarial network, GAN）。Social-LSTM以目标行人为中心建立固定大小的交互区域，通过区域内行人之间的交互预测目标行人的未来轨迹；Social-GAN同样基于LSTM，不同的是采用对抗训练的方法提升了轨迹预测精度；在实际场景中，目标行人对于不同方向的重视程度是不一样的，Social-Attention基于此设计了不同方向的行人对目标行人施加不同的影响力[24]。这类方法只关注区域内的行人，本文方法是将区域内的各类交通全部考虑在内，显然更加符合实际场景。

基于人与场景交互的预测模型。还有许多方法融入进了环境因素，提升预测的效果。文献[6]提出人们倾向于在人行道上而不是在草地上行走。Kitani等[25]采用反向强化学习预测人类的轨迹。Xie等[26]认为行人是“粒子”，其运动力学是在拉格朗日力学框架内建模。Scene-LSTM划分静态场景为曼哈顿网格，并使用LSTM预测行人的位置[9]。CAR Net方法提出一种场景语义的网络来预测行人的轨迹[27]。SoPhie方法利用场景语义分割模型的深层神经网络特征，关注行人运动轨迹[28]。还有一些方法试图用图神经网络研究行人交互模式。这些工作都很有意义，但并未考虑目标行人周围环境中不同交通类别的影响[3-4]。

基于场景视觉特征的预测模型。还有一些利用个体的视觉特征来预测目标行人的路径，而不是将其视为场景中的点[4,29-30]。Kooij等[31]通过观察行人的面部来模拟他们的意识，从而预测他们是否会过马路。Yagi等[32]提取周围行人的姿态特征，并使用卷积神经网络来预测第一人称视频中的未来轨迹。这些方法提升了轨迹预测的准确性，但并未关注行人意图的识别；Rasouli等[33]利用行人外貌特征和场景上下文预测行人意图和行人轨迹，Liang等[6]使用行人外貌特征和图像语义分割同时预测行人意图和行人轨迹。但行人外貌特征容易受到场景光线和穿着的影响，会对行人意图的识别造成干扰。与这些方法不同的是，本文通过行人的姿态特征[34-35]表征目标行人的行走意图，避免了行人外部特征变化对识别行人意图的干扰。

2 基于姿态特征提取的行人轨迹预测

行人在规划自己的路径时，会考虑到周围人和异构交通的运动情况，同时会牢记自己的目的地。目前预测行人轨迹的方法主要集中在依靠过去时刻的轨迹预测未来轨迹，忽略了对行人本身意图的关注。在本节中，首先，给出行人轨迹预测问题的定义；其次，详细阐述基于行人姿态识别行人意图的编码器-解码器网络架构。

2.1 行人轨迹预测问题的定义和分析

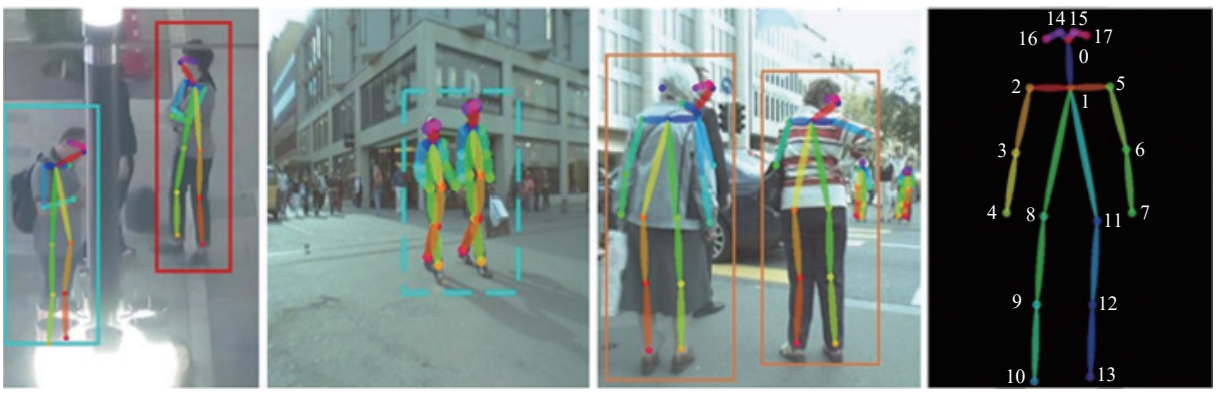
本文的目标是通过推理，预测行人在未来一段时间的轨迹。假设场景中行人 i 在时刻 t 的空间位置为 (x_t^i, y_t^i)，N 为 T_{obs} 个时刻内的行人数量，则行人轨迹预测被定义为：假设 N 个行人在 T_{obs} 个时刻内的空间位置 $\{(x_t^i, y_t^i)\}$，$t = 1,2,\cdots,T_{\text{obs}}$，$i = 1,2,\cdots,N$，通过模型预测每个行人未来 T_{pred} 时刻的空间位置。

行人轨迹预测是一个非常有挑战性的研究领域，一个重要的特点是行人行为的不确定性，因此，对行人的差异行为进行建模是非常困难的。同一场景下行人存在多种可能的行为决策，具有多模态性。例如在前方存在行人的情况下，往往存在不同的决策行为，包括跟随、绕开等。目前只能通过更多的训练数据来学习各种可能的行为，以此达到更

合理的预测。

同时, 目标行人都有自己的行走意图, 虽然在过程中可能会受到障碍物的影响而临时改变走向, 但还是会按照自己的意愿前进。因此, 目标行人意图的推断非常关键。

针对目前行人轨迹预测缺乏意图识别的问题, 本文从行人意图的角度出发, 提出一种基于行人姿态的轨迹预测方法。图 1 为姿态表征行人意图, 图 1 (a)、图 1 (b) 和图 1 (c) 中这 3 种行人姿态是不同的, 图 1 (a) 中静立等待的行人在不同时刻呈现的姿态基本保持不变; 图 1 (b) 中过马路的行人呈现的是行走时的姿态; 图 1 (c) 中等待过马路的行人在静止状态和行走状态所呈现的姿态是不同的。可以看出, 行人姿态可以反应行人的意图。



(a) 不打算过马路的行人　(b) 正在过马路的行人　(c) 打算过马路但正在等待的行人　(d) 行人姿态提取

图 1　姿态表征行人意图

Fig. 1　Pose representation of pedestrian intention

因此, 本文利用姿态估计方法获得行人姿态(见图 1(d)), 并结合空间语义信息, 提出 PKCEDN 以预测行人的未来轨迹。图 2 为 PKCEDN 的结构。其中, Conv为卷积, DeConv为反卷积, Attention Tensor 为注意力张量, $\boldsymbol{Y}_t \in \mathbf{R}^2$为模型预测的轨迹, $\boldsymbol{X}_t \in \mathbf{R}^{14\times11\times11}$为模型的输入, 表示行人在当前时刻的状态特征, 如下:

$$\boldsymbol{X}_t = \text{concatenate}(\boldsymbol{S}_t, \boldsymbol{P}_t) \tag{1}$$

式中: $\boldsymbol{S}_t \in \mathbf{R}^{13\times11\times11}$为目标行人周围的 13 类空间语义信息; $\boldsymbol{P}_t \in \mathbf{R}^{1\times11\times11}$为目标行人的姿态特征, 具体分析详见 2.3 节数据预处理模块; concatenate(·)为按通道维度进行拼接。

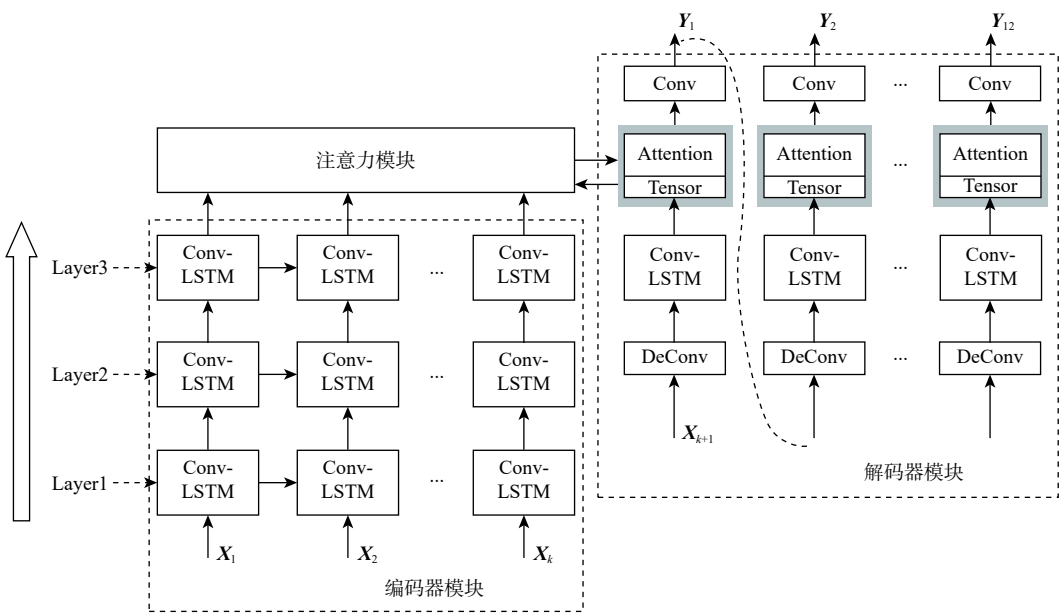


图 2　PKCEDN 的结构

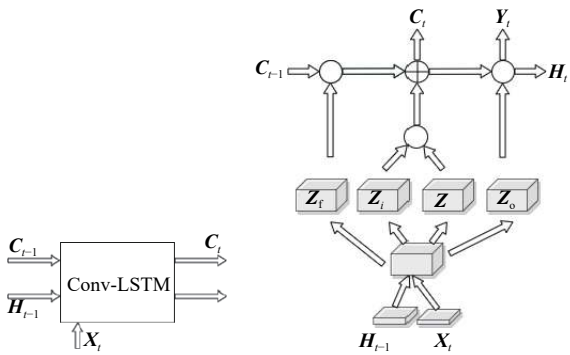
Fig. 2　Structure of PKCEDN

2.2　卷积长短期记忆网络

提取行人的空间信息对于预测目标行人轨迹至关重要。本文在数据预处理时, 用三维张量表征目标行人的状态(见 2.3 节)。但传统的 LSTM 只能

够处理一维张量，丢失了空间信息。为解决该问题，本文用 Conv-LSTM[10] 构建行人轨迹预测模型。

Conv-LSTM 用卷积运算代替 LSTM 单元中每个门的矩阵乘法。这样，其可通过在多维数据中进行卷积操作来捕获基础空间特征。图 3 为 Conv-LSTM 网络结构，其输入输出结构如图 3(a) 所示，其内部结构如图 3(b) 所示。1 个 Conv-LSTM 单元由 1 个遗忘门、1 个记忆单元、1 个输入门、1 个输出门组成。记忆单元存储和记忆过去时刻的信息（状态）。



(a) 输入输出结构 (b) 内部结构

图 3 Conv-LSTM 网络结构

Fig. 3 Structure of Conv-LSTM network

首先，遗忘单元决定记忆单元中删除哪些信息，保留哪些信息。遗忘单元的输出张量如下：

$$\boldsymbol{Z}_{\mathrm{f}}=\sigma(\boldsymbol{W}_{\mathrm{f}x}*\boldsymbol{X}_t+\boldsymbol{W}_{\mathrm{fh}}*\boldsymbol{H}_{t-1}+\boldsymbol{b}_{\mathrm{f}}) \tag{2}$$

式中：$\boldsymbol{X}_t$为 Conv-LSTM 的输入；$\boldsymbol{H}_{t-1}$为 Conv-LSTM 的隐层状态；$\sigma(\cdot)$为 sigmoid 激活函数；$\boldsymbol{b}_{\mathrm{f}}$为相应的偏置项；$*$为卷积操作；$\boldsymbol{W}_{\mathrm{fx}}$和 $\boldsymbol{W}_{\mathrm{fh}}$为 Conv-LSTM 的可学习参数。

然后，从$\boldsymbol{X}_t$中提取有用的信息得到 $\boldsymbol{Z}$，表达式为

$$\boldsymbol{Z}=\tanh(\boldsymbol{W}_x*\boldsymbol{X}_t+\boldsymbol{W}_{\mathrm{h}}*\boldsymbol{H}_{t-1}+\boldsymbol{b}) \tag{3}$$

式中：$\boldsymbol{W}_x$和 $\boldsymbol{W}_{\mathrm{h}}$为 Conv-LSTM 的可学习参数；$\boldsymbol{b}$为相应的偏置项。

结合式 (2) 和式 (3)，Conv-LSTM 通过式 (4) 和式 (5) 更新其记忆单元。

$$\boldsymbol{Z}_{\mathrm{i}}=\sigma(\boldsymbol{W}_{\mathrm{i}x}*\boldsymbol{X}_t+\boldsymbol{W}_{\mathrm{ih}}*\boldsymbol{H}_{t-1}+\boldsymbol{b}_{\mathrm{i}}) \tag{4}$$

$$\boldsymbol{C}_t=\boldsymbol{Z}_{\mathrm{f}}\circ\boldsymbol{C}_{t-1}+\boldsymbol{Z}_{\mathrm{i}}\circ\boldsymbol{Z} \tag{5}$$

式中：$\boldsymbol{b}_{\mathrm{i}}$为相应的偏置项；$\circ$为哈达玛积，$\boldsymbol{Z}_{\mathrm{i}}$为提取的输入信息；$\boldsymbol{C}_t$为当前状态张量。

最后，Conv-LSTM 利用式 (6) 和式 (7) 决定当前时刻输出的有效信息。

$$\boldsymbol{Z}_{\mathrm{o}}=\sigma(\boldsymbol{W}_{\mathrm{o}x}*\boldsymbol{X}_t+\boldsymbol{W}_{\mathrm{oh}}*\boldsymbol{H}_{t-1}+\boldsymbol{b}_{\mathrm{o}}) \tag{6}$$

$$\boldsymbol{H}_t=\boldsymbol{Z}_{\mathrm{o}}\circ\tanh(\boldsymbol{C}_t) \tag{7}$$

式中：$\boldsymbol{W}_{\mathrm{o}x}$和$\boldsymbol{W}_{\mathrm{oh}}$为 Conv-LSTM 的可学习参数；$\boldsymbol{b}_{\mathrm{o}}$为相应的偏置项。

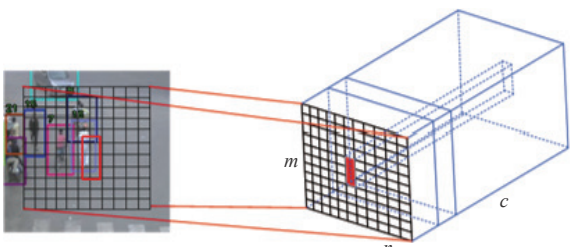
2.3 数据预处理模块

数据预处理模块主要是获取目标行人的时空序列。图 4 为空间语义信息分类，对于输入的图片，图 4(a) 中目标行人用红色方框标示，以该目标为中心，划出 11×11 的网格作为提取目标行人周围各种交通类别和障碍物的空间。每个网格的宽度为图片中行人高度的 1/3（单位为像素）。其含义为：当前时刻，在以目标行人为中心的 11×11 网格区域内的异构交通能够对目标行人的轨迹决策产生很大的影响。根据文献 [10] 对行人交互区域大小的研究，以目标行人为中心，在半径为 3 m 内的行人对目标行人的轨迹会有较大的影响；距离超过 3 m 的行人，相互影响很小，可以忽略不计。假设行人的平均身高约为 1.8 m，每个网格为行人高度的 1/3，故每个网格对应的实际物理距离约为 0.6 m。在实际中，由于成像平面和物理场景有一定的投影角度，位于图像平面不同位置同样大小的图像尺寸对应着不同的物理尺寸。本文为了估计目标行人的邻域大小，忽略了这种影响。

为了学习目标行人和周围环境的交互，本文根据异构交通和障碍物类别，将目标周围的空间语义信息分为 13 类：行人、车上的行人、汽车、自行车、摩托车、非机动车辆、站立的人、坐着的人、建筑物、地面上的路障、遮挡物 (如广告牌或电线杆)、反射 (如透明墙或窗户后的人) 和地面，并用一个 13 通道的三维张量来表示。其中，一个通道代表一种交通类别，如图 4(b) 所示，若网格内存在此目标类别，则在该通道内，其对应位置的值为 1，否则为 0。其中，c 为通道数，即交通类别，m 和 n 分别为纵向和横向网格数。由此，形成了一个三维张量 $\boldsymbol{S}_t$。

在图 4(b) 三维张量中，c 为 13，m 和 n 为 11，采用 one-hot[9] 格式编码 13 种交通类别。假设通道 1 代表行人类别，则此通道内行人所在位置网格的值为 1，其他网格的值为 0；其他通道（c 的取值范围为 [2,13]）对应位置网格的值也为 0。假设该行人所在的网格位置为 m=6、n=6，行人所在的通道为 c=1，则 [1,0,0,0,0,0,0,0,0,0,0,0,0] 代表了该行人。如图 4(b) 所示，虚线围成的三维长方体代表该行人信息。在行人意图建模方面，本文采用行人的姿态来表达目标行人的意图。图 5 为行人姿态建模与行人意图，在图 5(a) 中，有站立等待的行人，坐着休息的行人，正在滑行的行人。可以看出行人不同的运动状态会呈现不同的姿态；另外，姿态也蕴含着行人的方向信息，不同方向的行人暴露出的关键点个数不

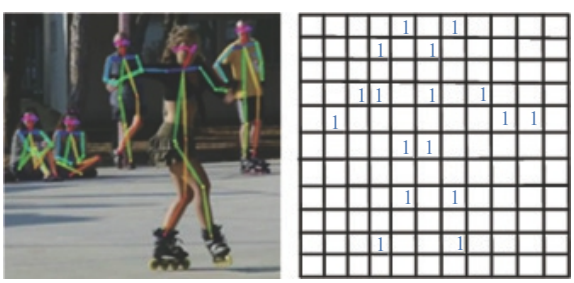
同, 位置分布也不同。图 5(b) 为图 5(a) 中正在滑行的女孩的关键点分布。这种所暗含的信息可以通过神经网络自主学习, 很好的辅助行人轨迹的预测。本文将行人的意向分为 4 类: 左、右、前、后 (所有方向都是相对于摄像头)。对于朝向不同的行人来说, 身体关键点的不同分布表明了不同的方向意图。本文采用文献 [35] 中经过预训练的模型 (Open-Pose)提取行人姿态关键点。本文把 18 个姿态关键点映射到 $m\times n$(m=n=11)个网格中, 这样可以保留姿态的原始空间信息, 如果一个单元格包含一个关键点, 将其设置为 1; 否则, 设置为 0。最后得到目标行人的姿态特征 $\boldsymbol{P}_t$。$\boldsymbol{S}_t$ 与 $\boldsymbol{P}_t$ 按通道进行拼接, 得到当前时刻目标行人的状态特征 $\boldsymbol{X}_t$。



(a) 目标行人　　(b) 三维张量

图 4　空间语义信息分类

Fig. 4　Spatial semantic information classification



(a) 姿态图　　(b) 姿态分布

图 5　行人姿态建模与行人意图

Fig. 5　Pedestrian intention modeling with pedestrian pose

2.4　编码器模块

在编码器中, 采用了 3 层 Conv-LSTM 组成的序列网络, 如图 2 所示。每个时刻都有一个输入特征 $\boldsymbol{X}$, $\boldsymbol{X}$ 是由数据预处理模块输出的, 格式如图 4(b) 所示。Conv-LSTM 的隐层尺寸随着深度增加, 分别为 32、64、128。最后一层的 Conv-LSTM 输出 128 通道 11×11 的特征。在编码器模块中, 通道数递增意味着能够提取到更多的特征。

2.5　解码器模块

解码器模块由上采样、Conv-LSTM、下采样 3 部分组成。输入为 2 通道 1×1 的张量, 经过上采样, 输出 128 通道 11×11 的张量, 之后输入到隐层为 128 通道的 Conv-LSTM, 并输入到注意力模块中, 得到 Context Tensor(128 通道 11×11), 最后, 经过下采样输出 2 通道 1×1。最后的输出表示目标行人在图像中相对于前一时刻的位移。解码器在起始时刻 Conv-LSTM 的隐层采用编码器最后时刻输出的隐层信息进行初始化的。

在训练阶段, 解码器的输入是各个时刻的真实轨迹(即 ground truth), 输出为下一时刻的预测值。在测试阶段, 解码器的输出当作下一时刻的输入。

2.6　注意力模块

虽然 2.4 节和 2.5 节所述的基于 Conv-LSTM 的编码器-解码器结构能够学习到相邻时刻的行人状态对当前时刻轨迹的影响。但对于行人轨迹来说, 当前的轨迹不仅受上一时刻行人状态的影响, 而且与前几个时刻的行人状态也有密切的关系。由图 2 可知, 编码器是用最后时刻输出的隐层初始化解码器第 1 个时刻的隐层。这种串联结构使得解码器每个时刻的输出只与上一时刻的行人状态相关。

为了让编码器-解码器结构学习每个过去时刻的行人状态对当前预测时刻轨迹的影响, 本文提出一种基于注意力机制的编码量化权重方法。注意力机制保存编码器(Encoder)输出的隐层状态 $(\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_k)$, 并通过训练, 学习解码器的隐层状态 $(\boldsymbol{h}_{k+1},\boldsymbol{h}_{k+2},\cdots,\boldsymbol{h}_{k+T})$和 $(\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_k)$的相关性。图 6 为编码器-解码器中的注意力模块详细结构, 注意力模块的具体实现包括以下 3 步:

步骤 1　在预测时刻为 t_{k+1}, 解码器的隐层状态为 $\boldsymbol{h}_{k+1}$, 计算 $(\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_k)$与 $\boldsymbol{h}_{k+1}$ 的相关性如下:

$$\text{score}(\boldsymbol{h}_i,\boldsymbol{h}_k)=\boldsymbol{h}_i * \boldsymbol{h}_{k+1}/\sqrt{d_{k+1}} \tag{8}$$

式中: *为卷积互相关; d_{k+1} 为 $\boldsymbol{h}_{k+1}$ 的大小。

步骤 2　对式(8)的结果归一化, 如式 (9) 所示。使其注意力权重之和为 1。由式 (9) 得到编码器每个时刻的对应权重。

$$\alpha_{i,k+1}=\frac{\exp(\text{score}(\boldsymbol{h}_i,\boldsymbol{h}_k))}{\sum_{i=1}^{k}\exp(\text{score}(\boldsymbol{h}_i,\boldsymbol{h}_k))} \tag{9}$$

步骤 3　将权重 $\alpha_{i,k+1}$和对应的特征 $\boldsymbol{h}_i$ 相乘, 并求和, 得到包含不同时刻特征贡献率大小的相关特征张量 $\boldsymbol{C}$ 如下:

$$\boldsymbol{C}=\alpha_{1,k+1}\boldsymbol{h}_1+\alpha_{2,k+1}\boldsymbol{h}_2+\cdots+\alpha_{k,k+1}\boldsymbol{h}_k \tag{10}$$

由式 (10) 得到相关特征张量 $\boldsymbol{C}$, 将 $\boldsymbol{C}$ 和 $\boldsymbol{h}_{k+1}$ 按通道维度进行拼接, 得到拼接张量 $\boldsymbol{T}_k$, 之后再对 $\boldsymbol{T}_k$ 进行卷积操作, 得到当前时刻的预测轨迹 $\boldsymbol{Y}_{k+1}$。

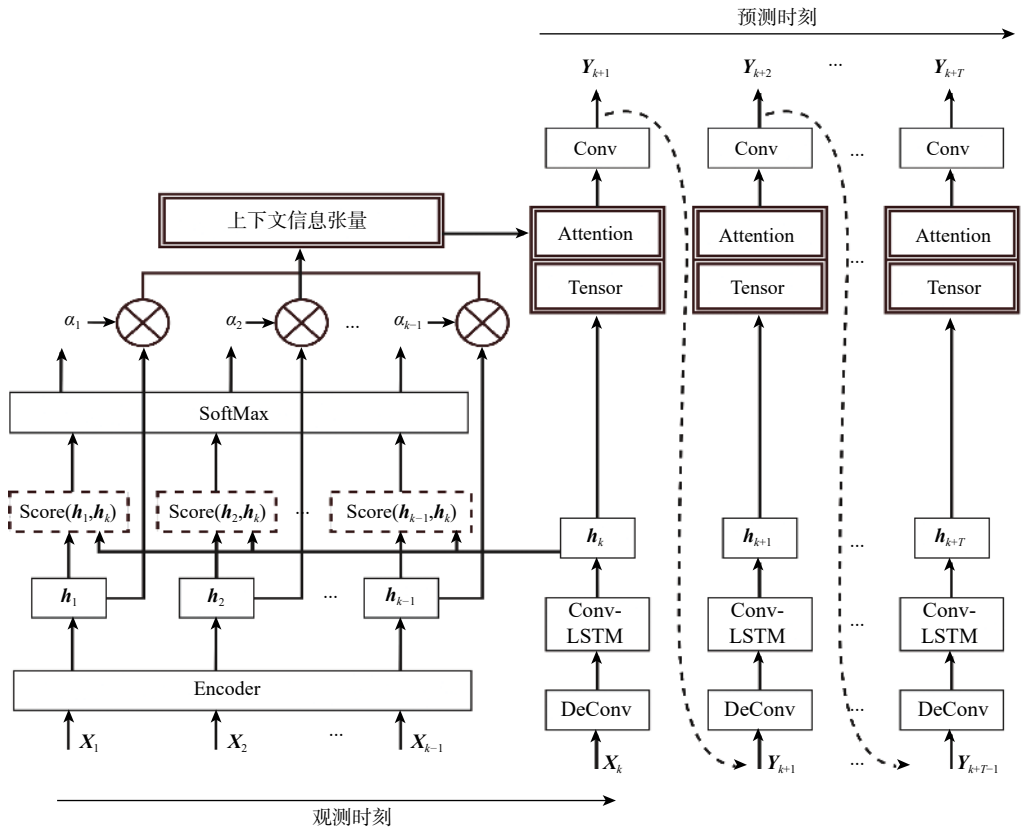


图 6 编码器-解码器中的注意力模块详细结构

Fig. 6 Detailed structure attention module for encoder-decoder architecture

3 预测网络参数设置

整个模型的实现是基于 PyTorch[36], 使用 Conv-LSTM 组成的编码器-解码器网络, 各个网络层的参数列表如表 1 所示。其中, 输入输出表示为 $c@m\times n$。编码器部分是由 3 层, 每层 7 个 Conv-LSTM 组成, 代表 7 个时间节点, 3 层 Conv-LSTM 的隐层通道数分别设置为 32、64、128; 解码器是由一层 12 个 Conv-LSTM 组成, 代表 12 个时间节点, 每个 Conv-LSTM 是由 128 通道组成。

表 1 网络层模型参数列表

Table 1 Network layer model parameters list

网络层	输入	输出
Layer1	16@11×11	32@11×11
Layer2	32@11×11	64@11×11
Layer3	64@11×11	128@11×11
Deconv	2@1×1	128@11×11
Conv-LSTM (Decoder Module)	128@11×11	128@11×11
Conv	128@11×11	2@1×1

模型用 Adam 作为优化器, 采用随机梯度下降法更新整个网络参数, 学习率设置为 0.001, dropout 设置为 0.5, 利用英伟达 V100 对模型进行训练。

网络的输出是张量, 借鉴了 2D 目标检测使用的损失函数。在本文中, 使用 Smooth-L1 损失函数, 如下:

$$L=\begin{cases}\dfrac{1}{N_t}\displaystyle\sum_{T=k+1}^{k+N}\dfrac{1}{2}\left(\left\|\boldsymbol{Y}_{T_{\text{pred}}}-\boldsymbol{Y}_T\right\|_2\right)^2 & \left\|\boldsymbol{Y}_{T_{\text{pred}}}-\boldsymbol{Y}_T\right\|_2\leqslant 1\\ \dfrac{1}{N_t}\displaystyle\sum_{T=k+1}^{k+N}\left\|\boldsymbol{Y}_{T_{\text{pred}}}-\boldsymbol{Y}_T\right\|_2-0.5 & \left\|\boldsymbol{Y}_{T_{\text{pred}}}-\boldsymbol{Y}_T\right\|_2>1\end{cases} \tag{11}$$

式中: N_t 为模型预测的输出序列的长度; k 为模型输入序列的长度; $\boldsymbol{Y}_{T\text{pred}}$ 和 $\boldsymbol{Y}_T$ 分别为预测值和真值。

4 实验与结果分析

本文方法需要通过姿态识别目标行人的意图, 适用于能够获得行人姿态的场景。为了测试本文方法在行人轨迹预测方面的表现, 在 MOT16[37]、MOT17 及 MOT20[38] 数据集上进行训练和测试。平均位移误差 (average displacement error, ADE)[23] 和最终位移误差 (final displacement error, FDE)[23] 作为评价模型性能的标准。

1)数据集和评价标准。MOT16 数据集一共提供了 14 个视频序列 (11 324 条可用行人轨迹), 其中 7 个为训练集, 7 个为测试集。这些视频序列是在无限制的环境中用静态和移动摄像机拍摄的, 主要标注目标为移动的行人与车辆。MOT16 拥有更

加丰富的画面, 不同的拍摄视角和相机运动, 也包含了不同天气状况的视频。遵从了相应的标注准则进行标注, 最后, 通过双重检测的方法来保证标注信息的高精确度。MOT16 标注的运动轨迹为二维坐标, 单位为像素。MOT17 包含 42 个具有挑战性的视频序列, 共有 43 591 条可用的行人轨迹; MOT20 包含 8 个人群密度高的视频序列, 共有 91 744 条可用的行人轨迹。以图像像素坐标进行跟踪和评估。所有序列都严格按照标准协议进行标注, 并通过多目标跟踪(multiple object tracking, MOT)基准进行高精度的注释。

本文使用来自训练集的 5 个视频序列, 每个视频每帧的分辨率都是 1 920×1 080, 并且包含 13 个对象类别。整个数据集是由移动摄像头和固定摄像头拍摄的, 并经过严格的标注, 标注的坐标是图像坐标系中生成的。

所有预测时刻的真值坐标与预测坐标之间的平均位移误差 ADE 为

$$\mathrm{ADE}=\frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\left\|\tilde{\boldsymbol{Y}}_t^i-\boldsymbol{Y}_t^i\right\|}{NT_{\mathrm{pred}}} \tag{12}$$

式中: $\tilde{\boldsymbol{Y}}_t^i$为行人$i$在时刻$t$的预测轨迹。

在最终时刻 T_{pred}, 最终预测的位置和真实位置的欧式距离误差 FDE 为

$$\mathrm{FDE}=\frac{\sum_{i=1}^{N}\left\|\tilde{\boldsymbol{Y}}_{T_{\mathrm{pred}}}^i-\boldsymbol{Y}_{T_{\mathrm{pred}}}^i\right\|_2}{N} \tag{13}$$

式中: $\tilde{\boldsymbol{Y}}_{T_{\mathrm{pred}}}^i$为行人$i$在时刻$T_{\mathrm{pred}}$的预测轨迹。

2)训练和测试。从 5 个视频序列 (MOT16-02、MOT16-04、MOT16-09、MOT16-10、MOT16-13) 中选择 4 个进行训练。在训练过程中, 使验证集 ADE 最小的模型参数是所训练的最好模型。剩下的第 5 个数据集留在测试中使用。上述过程称为“留一法”。对于 5 个视频序列, 重复上述过程 5 次, 每次在 4 个视频序列上进行训练, 可获得用于测试的 5 个参数模型, 每次训练中执行 100 个 epoch。在测试阶段: 利用第 5 个未在训练中使用的视频序列, 测试已训练好的模型。

3)方法比较。用和本文一样的方法训练和测试目前主流方法: LSTM[7]、Social-LSTM、Social-GAN、SR-LSTM[39] 和 $\mathrm{M}_{\mathrm{sgtv}}$[4], 并与本文方法进行对比, 其结果如表 2～表 4 所示。为了公平性, 本文和主流方法采用了相同的训练、测试策略及相同的超参数设置。

表 2　本文方法与主流方法的轨迹预测结果在 MOT16 数据集上的比较

Table 2　Comparison of trajectory prediction results of proposed method and state-of-the-art methods on MOT16 dataset

像素

方法	ADE					ADE均值
	MOT16-02	MOT16-04	MOT16-09	MOT16-10	MOT16-13	
LSTM	30.12	32.15	30.52	30.96	31.73	31.096
Social-LSTM	29.13	30.86	31.25	29.85	29.95	30.208
Social-GAN	27.63	29.42	28.75	29.54	30.83	29.234
SR-LSTM	28.73	29.65	30.74	29.31	29.03	29.492
$\mathrm{M}_{\mathrm{sgtv}}$	27.32	29.21	27.49	30.38	29.72	28.824
本文方法	**19.15**	**17.46**	**19.54**	**16.75**	**18.43**	**18.266**
方法	FDE					FDE均值
	MOT16-02	MOT16-04	MOT16-09	MOT16-10	MOT16-13	
LSTM	44.64	43.85	41.64	42.43	40.45	42.602
Social-LSTM	43.53	41.85	40.64	39.56	40.75	41.266
Social-GAN	34.54	35.46	32.64	37.85	38.85	35.868
SR-LSTM	38.35	37.31	38.62	34.18	32.48	36.188
$\mathrm{M}_{\mathrm{sgtv}}$	33.07	34.38	31.83	38.01	35.74	34.606
本文方法	**21.11**	**20.43**	**23.66**	**20.78**	**22.54**	**21.704**

4)定量结果分析。对于 MOT16-02、MOT16-04、MOT16-09、MOT16-10、MOT16-13 视频序列, 本文方法在 2 个度量标准 ADE 和 FDE 上的表现均优于其他方法。本文方法与其他方法最大的不同在于: 利用姿态关键点辅助预测行人的下一步动向, 表 2～表 4 中的结果表明本文方法用姿态关键点辅助判别行人意图行之有效, 最优结果在表 2～表 4 中加粗显示。另外在数据处理时保留了行人场景中的空间特征, 使得网络能够很好的学习行人所在场景的时空特性, 有效地通过行人的姿态预测其下

表3 本文方法与主流方法的轨迹预测结果在MOT17数据集上的比较

Table 3 Comparison of trajectory prediction results of PKCEDN and state-of-the-art methods on MOT17 dataset

像素

方法	ADE					ADE均值
	MOT17-02	MOT17-03	MOT17-08	MOT17-11	MOT17-14	
LSTM	35.15	36.31	33.85	35.21	39.38	35.98
Social-LSTM	30.01	31.48	32.89	35.02	33.64	32.608
Social-GAN	28.79	30.13	28.62	30.93	30.48	29.79
SR-LSTM	29.13	30.53	29.24	30.15	29.61	29.732
M_{sgtv}	26.53	30.13	28.75	29.32	27.54	28.454
本文方法	**20.04**	**17.86**	**18.42**	**15.84**	**18.05**	**18.042**
方法	FDE					FDE均值
	MOT17-02	MOT17-03	MOT17-08	MOT17-11	MOT17-14	
LSTM	50.64	55.97	61.57	60.02	63.42	58.324
Social-LSTM	44.64	52.18	54.15	43.63	42.17	47.354
Social-GAN	35.64	39.24	38.02	37.89	35.93	37.344
SR-LSTM	36.73	40.13	37.85	36.47	34.64	37.164
M_{sgtv}	34.53	37.64	35.18	37.56	33.17	35.616
本文方法	**24.75**	**23.84**	**25.03**	**22.41**	**25.75**	**24.356**

表4 本文方法与主流方法的轨迹预测结果在MOT20数据集上的比较

Table 4 Comparison of trajectory prediction results of PKCEDN and state-of-the-art methods on MOT20 dataset

像素

方法	ADE				ADE均值
	MOT20-01	MOT20-02	MOT20-03	MOT20-05	
LSTM	32.13	34.98	29.76	30.65	31.88
Social-LSTM	30.10	33.21	30.41	29.98	30.93
Social-GAN	27.02	26.98	26.81	24.06	26.22
SR-LSTM	28.31	33.01	28.73	29.53	29.895
M_{sgtv}	28.64	27.16	26.14	24.98	26.730
本文方法	**18.64**	**17.85**	**20.13**	**14.43**	**17.763**
方法	FDE				FDE均值
	MOT20-01	MOT20-02	MOT20-03	MOT20-05	
LSTM	46.54	51.83	53.54	53.41	51.33
Social-LSTM	43.83	44.12	43.84	40.75	43.14
Social-GAN	32.79	30.81	33.56	31.84	32.25
SR-LSTM	36.64	40.37	39.21	34.78	37.750
M_{sgtv}	33.69	32.42	34.15	31.19	32.863
本文方法	**24.74**	**21.63**	**24.31**	**19..46**	**22.535**

一步意图，与表2中的主流方法相比，本文方法将平均误差降低约36%。

此外，本文方法也在MOT17和MOT20数据集上进行了测试。可以看出，本文方法的预测结果同样优于其他方法，证明了本文方法具有很好的鲁棒性。

5）消融研究分析。本文通过预测时间内所有预测位置和实际位置的均方根误差 (mean average displacement error, MADE) 和预测轨迹的最终预测位置与相应的真实位置之间的均方根误差 (mean final displacement error, MADE) 这2个指标分析异构交通和姿态关键点对行人轨迹预测的影响。这2个指标常用来衡量轨迹预测方法的性能[5,7,24]。因此，通过这2个指标分别比较同时不考虑异构交通信息 Ten_h 和姿态关键点 Ten_p 的算法、只考虑异构交通信息 Ten_h 的算法、只考虑姿态关键点 Ten_p 的算法及同时融合异构交通信息 Ten_h 和姿态关键点 Ten_p 的算法之间的性能，能够间接的证明异构交通和姿态关键点的有效性。本文方法的消融实验结果如表5所示。

表 5　本文方法的消融实验

Table 5　Ablation study of proposed method

算法	参数	数值(MOT16)/像素
Ten_l	MADE	24.455
	MFDE	31.324
Ten_l + Ten_p	MADE	22.543
	MFDE	30.041
Ten_h	MADE	21.412
	MFDE	27.544
Ten_p+Ten_h	MADE	18.266
	MFDE	21.704

MADE 和 MFDE 的表达式分别为

$$\mathrm{MADE}=\frac{1}{n}\sum_{i=1}^{n}\mathrm{ADE}_i \tag{14}$$

$$\mathrm{MFDE}=\frac{1}{n}\sum_{i=1}^{n}\mathrm{FDE}_i \tag{15}$$

式中: n 为测试集中视频序列的数量。

MADE 和 MFDE 的值如表 5 所示(视频序列中每帧的尺寸为 1 080×1 920)。Ten_l 表示行人类别分布的张量元素, 是 Ten_h 中代表行人类别的一个通道。

Ten_l 为方法 Linear、LSTM、Social-LSTM、Social-GAN 的输入。在只有 Ten_l 输入的情况下, 与表 2 中其他方法的 ADE 相比较, 本文方法的 ADE 最小, 表现最好。

在加入姿态关键点 Ten_p 后, 本文对比了在 Ten_l + Ten_p 和 Ten_l 分别输入时, 本文方法的表现。可以看出, 在加入 Ten_p 后, 本文方法具有更小的 ADE, 这表明加入姿态关键点有助于网络识别行人意图, 提升网络轨迹预测的能力。

当网络输入的为异构交通类别 Ten_h 时, 相比输入 Ten_l +Ten_p 和 Ten_l, 本文方法具有更小的 ADE。这是因为目标行人轨迹不只受到周围行人的影响, 同时也会受到周围其他 12 种交通类别的影响, 比如小汽车、自行车、摩托车等。

在 Ten_p 和 Ten_h 同时作为输入时, 与其他输入情况相比, 本文方法得到了最小的 ADE。这表明, 姿态估计和异构交通类别都有助于提升网络对于行人轨迹预测的能力。

本文在数据集 MOT16 上评估了模型的处理速度。在 V100 GPU 卡上进行 6 次实验计算平均运行时间, 平台的操作系统是 CentOS 7.6。实验结果如表 6 所示。

在速度方面, LSTM 模型是最快的, 由于每个 LSTM 单元内部是由全连接层组成, 结构简单; Social-LSTM 模型针对每一个目标行人会在每一时刻都有一个池化层, 因此其处理速度明显低于 LSTM,

表 6　本文方法与主流方法在预处理和预测速度方面的比较

Table 6　Comparison with state-of-the-art methods on pre-processing speed and prediction speed

方法	数据预处理模块	预测模块速度/(次·s^{-1})	速度提升/%
LSTM	0.03	0.04	76.29
Social-LSTM	2.13	3.21	1
Social-GAN	0.20	0.26	11.60
SR-LSTM	1.24	2.01	1.64
M_{sgtv}	0.24	0.38	8.61
本文方法	0.54	0.28	6.51

Social-GAN 在处理数据时应用了多层感知机和最大池化层计算目标行人的影响区域, 并不需要在每个时刻为每个人都进行一次池化操作, 因此其处理速度明显快于 Social-LSTM; SR-LSTM 模型增加了社交信息传递层, 增大了模型的计算量, 因此需要更多的处理时间。M_{sgtv} 在每个时刻都需要进行池化操作, 因此所需时间比 Social-GAN 更多, 处理速度较慢。本文方法处理速度介于 Social-GAN 和 Social-LSTM 之间, 主要在于获取行人姿态时使用了预训练的模型 Open-Pose, 使得处理速度变慢, 但是预测的轨迹最准确。在预测速度方面, 本文方法与 Social-GAN相近。在实际的行人轨迹预测应用中, 通常是观察 3.2 s(8 帧)的行人轨迹, 预测 4.8 s(12 帧)的未来轨迹。表 6 中所有方法都满足实时性的需求。

6)定性结果分析。图 7 为 MOT16 数据集上 6 种方法预测行人轨迹的实例, 用蓝色线标注行人的历史轨迹。在图 7(a)中, 周围静止的行人不会阻碍目标行人。M_{sgtv} 和本文方法的预测效果最好。而 LSTM 和 Social-LSTM 误认为目标行人会有避障行为, 故预测的轨迹会远离旁边静止的行人。而本文方法在姿态和周围空间异构交通信息的辅助下, 能够很好地预测到目标行人的未来走向, 预测误差最小。图 7(b) 表明, 只有 Social-GAN、M_{sgtv} 和本文方法能够很好的识别结伴行为, 但本文方法更准确地预测了目标行人的未来轨迹。图 7(c) 展示了本文方法可以预测不同交通类别的轨迹, 图中目标行人骑车向左行走, 可以看到, 相比于其他方法, 本文方法的预测更加准确。

图 8 为在 MOT16、MOT20 数据集上, 预测轨迹定性结果的可视化, 在图 8(a) 中, 目标行人右侧附近的区域是静止的人群, 本文方法结合目标行人姿态和周围区域的行人空间信息能够预测更加真实的轨迹。在图 8(b) 中, 目标行人将绕过柱子, 继续前行。图 8(c) 中的行人会在人群中一直向前行走。可以看出, 本文方法的预测更接近真实轨迹。



(a) 保持直行 (b) 结伴行走 (c) 骑行行为

LSTM Social-LSTM Social-GAN SR-LSTM M_{sgtv} 本文方法

图 7 MOT16 数据集上，预测轨迹定性结果的可视化

Fig. 7 Trajectories predictions qualitative results visualization on MOT16 datasets



(a) 避让行走 (b) 避障行走 (c) 跟随行走

LSTM Social-LSTM Social-GAN SR-LSTM M_{sgtv} 本文方法

图 8 MOT20 数据集上，预测轨迹定性结果的可视化

Fig. 8 Trajectories predictions qualitative results visualization on MOT20 datasets

5 结 论

1）本文利用融合行人姿态与多种空间语义信息来预测行人轨迹，显著提升了行人轨迹预测的准确性。

2）从实验结果中可以看出，本文方法得到了较优的性能，与其他主流方法对比，本文方法将平均误差降低约 36%。

本文的工作有一些局限性。当前工作只研究了行人轨迹预测问题，下一步工作将着重研究异构交通的轨迹预测问题。另外，对于行人意图的分类问题，也需要进一步的探索研究，使其更好的辅助行人轨迹的预测。

参考文献（References）

[1] RODDENBERRY T M, GLAZE N, SEGARRA S. Principled simplicial neural networks for trajectory prediction[C]//Proceedings of Machine Learning Research. New York: Cornell Universitl, 2021: 9020-9029.

[2] CHENG H, LIAO W T, YANG M Y, et al. AMENet: Attentive maps encoder network for trajectory prediction[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 172: 253-266.

[3] CHANDRA R, GUAN T R, PANUGANTI S, et al. Forecasting trajectory and behavior of road-agents using spectral clustering in graph-LSTMs[J]. IEEE Robotics and Automation Letters, 2020, 5(3): 4882-4890.

[4] HADDAD S, LAM S K. Self-growing spatial graph networks for pedestrian trajectory prediction[C]//2020 IEEE Winter Conference on Applications of Computer Vision. Piscataway: IEEE Press, 2020: 1140-1148.

[5] WANG R, CUI Y, SONG X, et al. Multi-information-based convolutional neural network with attention mechanism for pedestrian trajectory prediction[J]. Image and Vision Computing, 2021, 107: 104110.

[6] LIANG J W, JIANG L, NIEBLES J C, et al. Peeking into the future: Predicting future person activities and locations in videos[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition . Piscataway: IEEE Press, 2020: 5718-5727.

[7] ALAHI A, GOEL K, RAMANATHAN V, et al. Social LSTM: Human trajectory prediction in crowded spaces[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition . Piscataway: IEEE Press, 2016: 961-971.

[8] FERNANDO T, DENMAN S, SRIDHARAN S, et al. GD-GAN: Generative adversarial networks for trajectory prediction and group detection in crowds[M]. Computer Vision — ACCV 2018. Berlin: Springer, 2019: 314-330.

[9] MANH H, ALAGHBAND G. ScenSe-LSTM: A model for human trajectory prediction[EB/OL]. (2019-04-15)[2021-08-20]. https://arxiv.orglabs/1808.04018.

[10] SONG X, CHEN K, LI X, et al. Pedestrian trajectory prediction based on deep convolutional LSTM network[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(6): 3285-3302.

[11] GRAVES A, GRAVES A. Long short-term memory[M]. Supervised Sequence Labelling with Recurrent Neural Networks. Berlin: Springer, 2012: 37-45.

[12] HELBING D, FARKAS I, VICSEK T. Simulating dynamical features of escape panic[J]. Nature, 2000, 407(6803): 487-490.

[13] MORRIS B T, TRIVEDI M M. Trajectory learning for activity understanding: Unsupervised, multilevel, and long-term adaptive ap-

proach[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(11): 2287-2301.

[14] HELBING D, MOLNÁR P. Social force model for pedestrian dynamics[J]. Physical Review E, 1995, 51(5): 4282-4286.

[15] KIM J, AHN C, LEE S. Modeling handicapped pedestrians considering physical characteristics using cellular automaton[J]. Physica A:Statistical Mechanics and Its Applications, 2018, 510: 507-517.

[16] QUAN R J, ZHU L C, WU Y, et al. Holistic LSTM for pedestrian trajectory prediction[J]. IEEE Transactions on Image Processing, 2021, 30: 3229-3239.

[17] ZHAO T Y, XU Y F, MONFORT M, et al. Multi-agent tensor fusion for contextual trajectory prediction[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE press, 2019: 12126-12134.

[18] XU Y Y, PIAO Z X, GAO S H. Encoding crowd interaction with deep neural network for pedestrian trajectory prediction[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 5275-5284.

[19] ZOU H S, SU H, SONG S H, et al. Understanding human behaviors in crowds by imitating the decision-making process[C]//Thirty-Second AAAI Conference on Artificial Intelligence. Piscataway: AAAI Press, 2018.

[20] TABUADA P, PAPPAS G J, LIMA P. Motion feasibility of multi-agent formations[J]. IEEE Transactions on Robotics, 2005, 21(3): 387-392.

[21] ZHANG C J, LESSER V. Multi-agent learning with policy prediction[C]//Proceedings of the AAAI Conference on Artificial Intelligence. New York: ACM, 2010, 24(1): 927-934.

[22] RUDENKO A, PALMIERI L, HERMAN M, et al. Human motion trajectory prediction: A survey[J]. The International Journal of Robotics Research, 2020, 39(8): 895-935.

[23] GUPTA A, JOHNSON J, LI F F, et al. Social GAN: Socially acceptable trajectories with generative adversarial networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 2255-2264.

[24] VEMULA A, MUELLING K, OH J. Social attention: Modeling attention in human crowds[C]//2018 IEEE international Conference on Robotics and Automation. Piscataway: IEEE Press, 2018: 4601-4607.

[25] KITANI K M, ZIEBART B D, BAGNELL J A, et al. Activity Forecasting[C]//European Conference on Computer Vision. Berlin: Springer, 2012: 201-214.

[26] XIE D, SHU T M, TODOROVIC S, et al. Learning and inferring “dark matter” and predicting human intents and trajectories in videos[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(7): 1639-1652.

[27] JAIPURIA N, HABIBI G, How J P. A transferable pedestrian motion prediction model for intersections with different geometries[EB/OL].(2018-06-25)[2021-08-06].https://arxiv.orglabs/1806.09444.

[28] SADEGHIAN A, KOSARAJU V, SADEGHIAN A, et al. SoPhie: An attentive GAN for predicting paths compliant to social and physical constraints[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition . Piscataway: IEEE Press, 2019: 1349-1358.

[29] NIKHIL N, MORRIS B T. Convolutional Neural Network for Trajectory Prediction[C]//European Conference on Computer Vision. Berlin: Springer, 2019: 186-196.

[30] SIMON T, JOO H, MATTHEWS I, et al. Hand keypoint detection in single images using multiview bootstrapping[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 4645-4653.

[31] KOOIJ J F P, SCHNEIDER N, FLOHR F, et al. Context-Based Pedestrian Path Prediction[C]//European Conference on Computer Vision. Berlin: Springer, 2014: 618-633.

[32] YAGI T, MANGALAM K, YONETANI R, et al. Future person localization in first-person videos[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscatawary: IEEE press, 2018: 7593-7602.

[33] RASOULI A, KOTSERUBA I, KUNIC T, et al. Pie: A large-scale dataset and models for pedestrian intention estimation and trajectory prediction[C]//2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2019: 6262-6271.

[34] YANG S, LUO P, LOY C C, et al. Wider face: A face detection benchmark[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscatawary: IEEE Press, 2016: 5525-5533.

[35] CAO Z, HIDALGO G, SIMON T, et al. OpenPose: Realtime multi-person 2D pose estimation using part affinity fields[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(1): 172-186.

[36] PyTorch. [EB/OL]. (2018-06-23) [2021-07-12]. https://pytorch.org/.

[37] MILAN A, LEAL-TAIXÉ L, REID I, et al. MOT16: A benchmark for multi-object tracking[EB/OL]. (2016-03-03)[2021-08-06]. https://arxiv.orglabs11603.00831.

[38] DENDORFER P, REZATOFIGHI H, MILAN A, et al. Mot20: A benchmark for multi object tracking in crowded scenes[EB/OL]. (2020-03-19)[2021-08-07]. https://arxiv.orglabs/2003.0903.

[39] ZHANG P, OUYANG W L, ZHANG P F, et al. SR-LSTM: State refinement for LSTM towards pedestrian trajectory prediction[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 12077-12086.

Pedestrian trajectory prediction method based on pedestrian pose

WANG Ruiping[1], SONG Xiao[2,*], CHEN Kai[3], GONG Kaiqi[2], ZHANG Junfan[4]

(1. School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China;
2. School of Cyber Science and Technology, Beihang University, Beijing 100191, China;
3. College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China;
4. School of Astronautics, Beihang University, Beijing 100191, China)

Abstract: In the field of autonomous driving, pedestrian trajectory prediction has been one of the research hotspots, and the uncertainty of pedestrian behavior poses a great challenge to trajectory prediction. Most of the current trajectory prediction methods only focus on the information interaction between pedestrians, ignoring the influence of pedestrian intention and other semantic information in the scene on the pedestrian trajectory. In order to achieve this, this paper suggests a method for predicting target pedestrian trajectory using pose keypoints based convolutional encoder-decoder network (PKCEDN). The method includes an attention mechanism that can learn the relationship between the current moment and past moment trajectories, as well as an encoder-decoder model based on convolutional, long and short-term memory (LSTM) networks. The proposed method has been tested on the MOT16, MOT17, and MOT20 public datasets, and the average error is reduced by about 36% compared to mainstream methods such as Linear, LSTM, Social-LSTM, Social-GAN, SR-LSTM, and M_{sgtv}, while ensuring no reduction in prediction speed.

Keywords: pedestrian trajectory prediction; pose extraction; encoder-decoder; attention mechanisms; spatial semantic information

Received: 2021-09-26; **Accepted**: 2022-01-02; **Published Online**: 2022-02-24 14:37
URL: kns.cnki.net/kcms/detail/11.2625.V.20220223.1541.004.html
Foundation item: National Key R & D Program of China (2018YFB1702703)
* **Corresponding author**. E-mail: songxiao@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0558

兔激素性股骨头坏死的造模改进及结果评价

杨丰合[1]，冯桂宇[2]，李晋玉[2]，牛旭锋[1,*]

(1. 北京航空航天大学 生物与医学工程学院 生物力学与力生物学教育部重点实验室，北京 100191；
2. 北京中医药大学 东直门医院，北京 100700)

摘　　要：股骨头坏死动物模型的建立对临床上股骨头坏死的治疗具有重要意义。对目前最常见的兔激素性股骨头坏死的造模方法进行改进，并对改进后的造模结果进行评估，希望能够在降低造模动物死亡率的同时成功诱导股骨头坏死。选择周龄介于16～20周，体重介于2.8～3.2 kg的雄性新西兰大白兔90只，随机分为空白对照组、常规造模组和造模改进组3组。其中，空白对照组仅注射生理盐水；常规造模组在特定时间点注射脂多糖和甲基泼尼松龙；造模改进组在常规造模组的基础上辅以青霉素、雷尼替丁、地克珠利、肠道润滑剂和乳酶生干预。结果表明：造模改进组动物的死亡率（13.3%）远低于常规造模组（70%）。造模完成后的微型计算机断层扫描技术（Micro-CT）分析和组织学染色结果显示：造模改进组股骨头区域出现了早期股骨头坏死的特征，其骨体积、骨体积与组织体积比、骨密度、骨小梁厚度均显著低于空白对照组，而空骨陷窝个数相较于空白对照组则出现了显著提升，证明改进后的造模方法在降低死亡率的同时成功诱导了兔股骨头坏死。

关　键　词：股骨头坏死；激素；动物模型；微型计算机断层扫描技术；组织学染色
中图分类号：R681.8; R318
文献标志码：A　　**文章编号**：1001-5965(2023)07-1755-06

股骨头坏死是一种多发于20～40岁人群的进行性疾病[1-2]。其起病隐匿，进展较快，若不经有效治疗，患者会在短时间内出现股骨头塌陷和髋关节破坏[3]。此时，就需要接受全髋关节置换术治疗[4-5]。该疗法手术创口较大，治疗费用高昂，常给患者带来生理和经济上的双重负担。据报道，中国成人非创伤性股骨头坏死的发病率为0.725%。这种疾病已经成为一个日益严重的临床问题，严重威胁中国人民健康。因此，开展相应的动物实验对股骨头坏死发病机理和治疗方式进行研究，对临床上股骨头坏死治疗具有重要指导意义。

动物实验的开展离不开动物模型的建立。目前，实验研究中最常用的动物模型是由Qin等[6]于2006年提出的通过糖皮质激素诱导的兔股骨头坏死模型。此模型造模方式简单，造模耗时较短，通过使用大量激素可以在短时间成功诱导股骨头坏死。然而，后续的研究表明，激素的过量使用会导致造模过程中出现过多数量的实验动物死亡，增加了实验经费负担，干扰了实验数据采集，同时也造成了实验动物生命不必要的浪费。因此，如何在成功造模的同时降低实验动物的死亡率是目前亟待解决的问题。

本文首先采用常规方法进行兔激素性股骨头坏死造模，通过解剖学等手段探明造模过程中动物常见的死亡原因。进而对常规造模方法进行有针对性的改进，并将造模改进组股骨头和空白对照组

收稿日期：2021-09-17；录用日期：2021-10-22；网络出版时间：2021-12-22 15:07
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20211222.0902.001.html
基金项目：国家自然科学基金(11872097,82074463)；国家重点研发计划(2020YFC0122204)；111计划(B13003)
*通信作者. E-mail：nxf@buaa.edu.cn

引用格式：杨丰合，冯桂宇，李晋玉，等. 兔激素性股骨头坏死的造模改进及结果评价[J]. 北京航空航天大学学报，2023，49(7)：1755-1760. YANG F H，FENG G Y，LI J Y，et al. Improved modeling strategy and result evaluation of steroid-induced femoral head necrosis in rabbits[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1755-1760 (in Chinese).

股骨头进行对比，验证改进后造模方法的有效性。旨在探明一种经过改进的股骨头坏死造模方法，在成功造模的同时降低实验动物的死亡率。

1 实 验

1.1 动物与材料

本文的动物实验已取得北京大学生物医学伦理委员会批准，动物饲养方式严格遵循其规定。实验选用周龄介于16～20周，体重介于2.8～3.2 kg的健康雄性新西兰大白兔90只。实验动物由中国中医科学院中药研究所提供。脂多糖购自美国Sigma-Aldrich公司。甲基泼尼松龙购自上海麦克林生化科技有限公司。青霉素、雷尼替丁、地克珠利、肠道润滑剂和乳酶生由北京中医药大学东直门医院提供。

1.2 分组与造模

经1周适应期饲养后，实验兔被随机分为空白对照组、常规造模组和造模改进组3组，每组30只。其中，空白对照组与其余2组在相同时间点注射等量的生理盐水。常规造模组于第1天经耳缘静脉注射脂多糖(10 μg/kg)，于第2～第4天经臀中肌注射糖皮质激素甲基泼尼松龙(20 mg/kg)，每次注射前均需重新称量体重以计算注射量。造模改进组在常规造模组的基础上，首先，于第2～第4天喂食地克珠利(每500 mL水加0.025 mL，摇匀后供兔子自由饮用)；其次，于第2天起喂食雷尼替丁(每500 mL水加15 mg，摇匀后供兔子自由饮用)，隔天喂服，直至造模结束；并且于第2天起肌肉注射青霉素，每日每只4万单位，直至造模结束；最后，于第2天起在饲料中添加乳酶生，每只每日3粒，直至造模结束，投喂乳酶生时间至少与青霉素注射时间间隔6 h。对于常规造模组和造模改进组，最后1针甲基泼尼松龙注射完毕14天后视为造模完成。

1.3 一般行为学及解剖学观察

观察期自注射第1针脂多糖开始，至最后1针甲基泼尼松龙注射完毕14天后为止，共18天。一般行为学观察主要是对兔子的进食状况，粪便情况和精神状态进行观察；解剖学观察主要是对观察期内死亡的兔子进行解剖，对其主要脏器进行观察。在观察的同时对各组兔子死亡率进行记录，以供后续比较。

1.4 Micro-CT分析及三维重建

最后1针甲基泼尼松龙注射完毕14天后，通过耳缘静脉注射过量戊巴比妥钠处死兔子。取出双侧股骨头，生理盐水冲洗后至于10%中性福尔马林中固定48 h。随后，使用微型计算机断层扫描技术(micro computed tomography, Micro-CT)(Skyscan 1076)对固定完毕的股骨头进行扫描，定量分析股骨头区域骨体积、骨体积与组织体积比、骨密度、骨小梁厚度和骨小梁分离度等指标，并对股骨头区域进行三维重建，以评估改进后的造模方法是否成功诱导股骨头坏死。

1.5 组织学观察

双侧股骨头在10%中性福尔马林中固定48 h。随后，置于10%EDTA脱钙液中脱钙3个月，进行苏木精−伊红(hematoxylin-eosin staining, H&E)染色，染色步骤如文献[7-8]所述。使用光学显微镜对切片的骨陷窝进行观察，以进一步评估改进后的造模方法对股骨头坏死的诱导效果。

1.6 数据处理

所有数据至少包含3个平行样，数据显示为mean ± SD(均值 ± 标准差)。数据由SPSS22.0软件进行分析，数据分析采用T检验，显著性水平用p表示，当$p<0.05$时，差异被视为具有统计学意义。其中，$p<0.05$时标记为*，$p<0.01$时标记为**，$p<0.005$时标记为***。

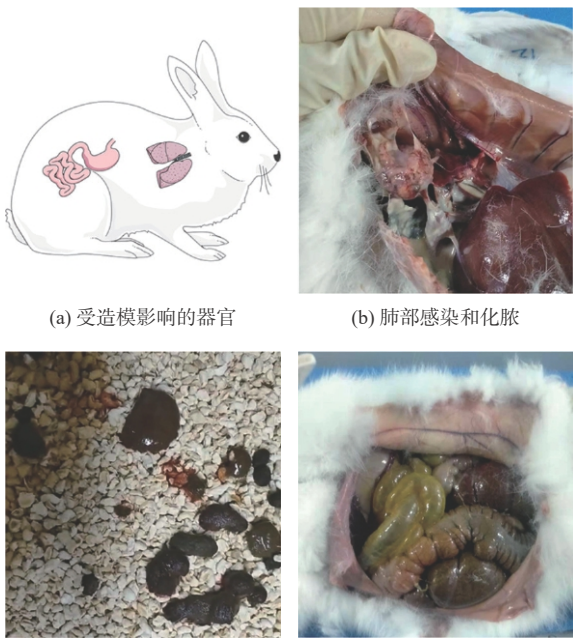
2 结果与讨论

2.1 造模后观察及死亡原因分析

在第1天注射脂多糖后，常规造模组和造模改进组兔子均出现了精神萎靡、体温升高和食欲不振等症状，极少量兔子出现了死亡。这种症状在第2天注射第1针激素后明显减轻。这是由于脂多糖是外来抗原物质，在体内会引发剧烈炎症反应[9]，而激素注入后则起到了消炎作用。在第4天最后一针激素注射完毕后，常规造模组开始出现大规模的兔子死亡，图1为激素性股骨头坏死动物模型造模过程中死亡因素。对死亡的兔子进行解剖后发现，兔子的双肺出现了严重的感染和化脓(见图1(b))。糖皮质激素是临床上公认的具有免疫抑制作用的药物。因此，可以推断，兔子肺部感染是由于过量使用激素抑制了实验动物的免疫系统导致的。对于造模改进组的兔子，由于在造模期间每日均注射青霉素，仅有2只兔子因肺部感染死亡。青霉素在杀死外源病菌的同时也会对兔子自身肠道有益菌造成损伤，造成兔子进食减少。因此，注射青霉素后，在造模改进组的饲料中添加了乳酶生以补充肠道有益菌。

除了肺部感染以外，消化道损伤也常见于死亡的兔子。图1(d)显示，许多常规造模组的死亡兔子出现了严重的肠胀气。同时，从图1(c)中可见，部

分兔子在死亡前出现了拉稀甚至便血的症状。研究表明, 在临床上大剂量使用激素会导致消化道溃疡及出血[10]。因此, 可以推断, 激素导致的消化道问题是造成兔子死亡的另一大原因。对于造模改进组的兔子, 由于使用了消化道溃疡治疗药物雷尼替丁, 仅有 1 只兔子在造模过程中出现了拉稀并死亡。除此之外, 还对本组的兔子进行了日常观察, 对于腹部涨大的兔子给予肠道润滑剂治疗, 以排出肠道气体, 进一步缓解了因消化道损伤导致的死亡现象。



(a) 受造模影响的器官　(b) 肺部感染和化脓

(c) 拉稀和便血　(d) 肠道胀气

图 1　激素性股骨头坏死动物模型造模过程中死亡因素

Fig. 1　Death factors in animal model of steroid-induced femoral head necrosis

此外, 在对常规造模组兔子的观察中发现, 兔子的腹泻常呈现小规模、集中性爆发的特点。当一只兔子出现严重腹泻并死亡后, 其周围笼位的多数兔子会于 3～5 日内出现同样的腹泻症状。结合兔子在腹泻前出现的食欲减退, 伏卧不动及精神萎靡等症状, 推测腹泻是由兔球虫病导致。兔子多为球虫携带者, 在正常情况下不会发病。然而, 当其免疫减退时, 潜伏在体内的球虫就会大量繁殖, 导致球虫病爆发。球虫病具有强传染性, 在球虫病爆发后兔子往往在几日内就会腹泻死亡。因此, 可以推断, 激素的使用降低了兔子的免疫力, 导致了球虫病的爆发, 造成了兔子的大量死亡。对于造模改进组, 在第 2～第 4 天给予兔子抗球虫药–地克珠利对球虫病进行预防, 因此, 在整个造模过程中此组并未出现像常规造模组一样的大规模腹泻死亡现象。

在注射最后一针甲基泼尼松龙 14 天后(即总天数第 18 天时), 激素诱导的股骨头坏死造模完成。由图 2 可知, 在整个造模周期内空白对照组无兔子死亡; 常规造模组共死亡 21 只, 造模死亡率为 70%; 造模改进组共死亡 4 只, 造模死亡率为 13.3%, 显著低于常规造模组。

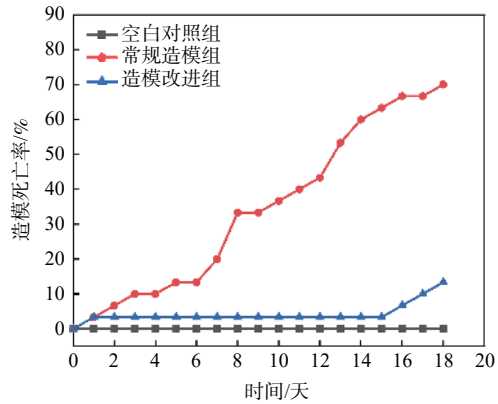


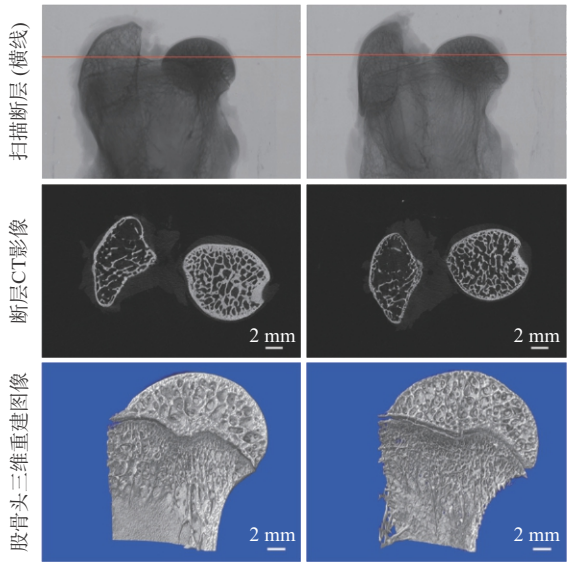
图 2　各组造模死亡率随时间变化曲线

Fig. 2　Curves of death rate with time in different groups

2.2　Micro-CT 分析造模结果

股骨头坏死发生后, 坏死部位骨小梁的形态和结构会随之发生改变[11]。Micro-CT 作为一种非侵入性检测方式, 可以在不破坏检测对象结构的情况下, 显示检测对象的内部结构, 其对股骨头坏死中的骨小梁具有很高的敏感性[12-13]。因此, 为了证明改进后的造模方式在提升存活率的同时能够保证造模的成功, Micro-CT 被用来分析空白对照组和造模改进组兔子的股骨头。通过 Micro-CT 对 2 组兔子股骨头进行断层扫描和三维重建, 可以清晰地分辨股骨头中骨小梁的情况, 进而对是否出现股骨头坏死进行初步判断。空白对照组股骨头断层扫描数据如图 3 所示, 从图中可见, 空白对照组股骨头中的骨小梁厚度适中, 排列紧密有序。和空白对照组相比, 造模改进组股骨头中的骨小梁厚度明显变薄, 同时骨小梁间隙变大, 排列也相对紊乱, 显示出早期股骨头坏死的特征。空白对照组和造模改进组股骨头部位的三维重建数据也显示出和断层扫描相似的骨小梁特征。

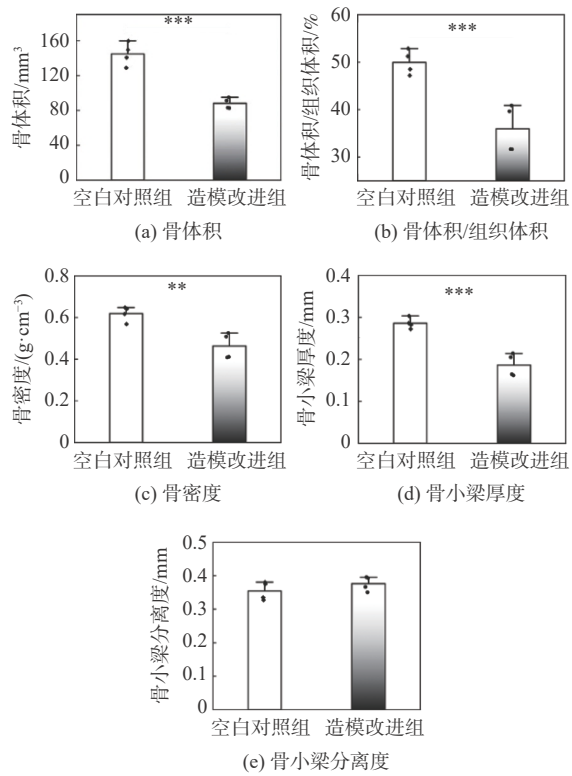
为了更为精确表征造模改进组的股骨头坏死情况, 对 2 组兔子股骨头的 Micro-CT 数据进行定量分析, 定量分析结果如图 4 所示。从图中可见, 相较于空白对照组, 造模改进组股骨头中的骨体积(p = 0.000 23)、骨体积与组织体积比(p = 0.002 5)、骨密度(p = 0.005 01)和骨小梁厚度(p = 0.000 56)均出现了显著降低, 而骨小梁分离度则出现小幅提升。以上数据表明, 造模改进组股骨头部位的骨质密度和骨小梁结构均向着使其力学性能变差的方向发展, 这是股骨头坏死的典型特征, 也是股骨头坏死后发



(a) 空白对照组 (b) 造模改进组

图 3 股骨头部位 Micro-CT 断层扫描及三维重建图像

Fig. 3 Micro-CT tomography and 3D reconstruction images of femoral head



(a) 骨体积 (b) 骨体积/组织体积

(c) 骨密度 (d) 骨小梁厚度

(e) 骨小梁分离度

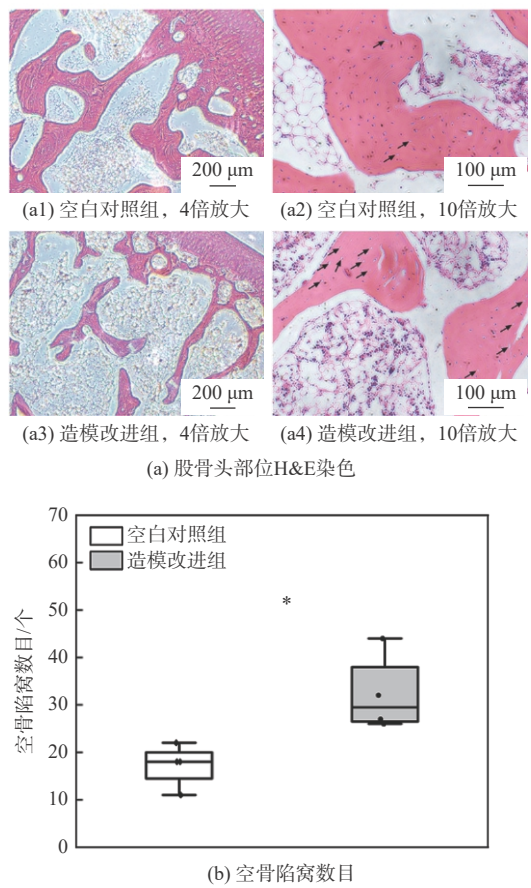
图 4 股骨头部位 Micro-CT 数据定量分析结果

Fig. 4 Quantitative analysis of Micro-CT data in femoral head region

生骨塌陷的直接原因[14-15]。

2.3 组织学染色分析造模结果

在股骨头坏死发生后，除了会在宏观上造成骨小梁结构和形态的改变之外，还会使骨小梁在微观的组织学层面发生改变。骨小梁上存在大量的骨陷窝，骨细胞就分布在这一个个骨陷窝中。研究表明，在股骨头坏死发生后，坏死部位的骨细胞会逐渐死亡，导致骨小梁上出现大量的空骨陷窝[16-17]。因此，通过对比造模改进组和空白对照组的空骨陷窝数量，即可在组织学层面对造模改进组是否出现股骨头坏死进行判断[18]。空白对照组和造模改进组不同放大倍数的 H&E 染色图像如图 5 所示。在较小的放大倍数上(4 倍)，同样可以在组织学上观察到造模改进组的骨小梁相较于空白对照组更为稀疏、紊乱。当拍摄倍数放大后(10 倍)，在造模改进组的切片上可以观察到许多空的骨陷窝(见图 5(a4)黑色箭头)；而对于空白对照组，只有极少的空骨陷窝在骨小梁上被观察到(见图 5(a2) 黑色箭头)。通过对 2 组不同切片上空骨陷窝的数量进行计数后发现，造模改进组的空骨陷窝数量显著高于空白对照组(p = 0.019 15)。因此，在 Micro-CT 的基础上，通过组织学手段进一步证明了造模改进组成功诱导了实验动物的股骨头坏死。



(a1) 空白对照组，4倍放大 (a2) 空白对照组，10倍放大

(a3) 造模改进组，4倍放大 (a4) 造模改进组，10倍放大

(a) 股骨头部位H&E染色

(b) 空骨陷窝数目

图 5 股骨头部位 H&E 染色图像及空骨陷窝数目统计

Fig. 5 H&E staining images and the number of empty bone lacunae in femoral head region

3 结 论

1) 常规的激素性股骨头坏死模型在造模过程

中会导致实验动物出现呼吸道感染和消化道损伤而死亡。通过青霉素、雷尼替丁、地克珠利、肠道润滑剂和乳酶生对常规造模方法进行改进后, 股骨头坏死的造模死亡率大幅降低。

2) 改进后的造模方法在降低了造模死亡率的同时, 保证了造模的成功。相较于空白对照组, 造模改进组的 Micro-CT 和组织学染色结果均提示该组出现早期股骨头坏死的特征。

本文对常规激素性股骨头坏死动物模型的诱导方法进行改进, 在成功造模的同时降低了动物的死亡率, 为今后股骨头坏死动物模型的顺利建立提供了参考和依据。

参考文献（References）

[1] WANG C K, HO M L, WANG G J, et al. Controlled-release of rhBMP-2 carriers in the regeneration of osteonecrotic bone[J]. Biomaterials, 2009, 30(25): 4178-4186.

[2] MARUYAMA M, NABESHIMA A, PAN C C, et al. The effects of a functionally-graded scaffold and bone marrow-derived mononuclear cells on steroid-induced femoral head osteonecrosis[J]. Biomaterials, 2018, 187: 39-46.

[3] ZHANG H X, ZHANG X P, XIAO G Y, et al. In vitro and in vivo evaluation of calcium phosphate composite scaffolds containing BMP-VEGF loaded PLGA microspheres for the treatment of avascular necrosis of the femoral head[J]. Materials Science and Engineering:C, 2016, 60: 298-307.

[4] WAEWSAWANGWONG W, RUCHIWIT P, HUDDLESTON J I, et al. Hip arthroplasty for treatment of advanced osteonecrosis: Comprehensive review of implant options, outcomes and complications[J]. Orthopedic Research and Reviews, 2016, 8: 13-29.

[5] STEVENS K, TAO C, LEE S U, et al. Subchondral fractures in osteonecrosis of the femoral head: Comparison of radiography, CT, and MR imaging[J]. American Journal of Roentgenology, 2003, 180(2): 363-368.

[6] QIN L, ZHANG G, SHENG H, et al. Multiple bioimaging modalities in evaluation of an experimental osteonecrosis induced by a combination of lipopolysaccharide and methylprednisolone[J]. Bone, 2006, 39(4): 863-871.

[7] 孙磊, 袁源, 牛睿, 等. 辐照和EO灭菌对SIS材料免疫原性的对比研究[J]. 北京航空航天大学学报, 2020, 46(12): 2245-2252.
SUN L, YUAN Y, NIU R, et al. Comparative study on immunoreactions of small intestinal submucosa by irradiation and ethylene oxide sterilization treatments[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(12): 2245-2252(in Chinese).

[8] 张林, 孙磊, 徐梦洽, 等. 可吸收胶原膜的体内免疫反应评价[J]. 北京航空航天大学学报, 2018, 44(4): 879-886.
ZHANG L, SUN L, XU M H, et al. Immunological response evaluation of absorbable collagen membrane in vivo[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(4): 879-886(in Chinese).

[9] XU J Z, GONG H P, LU S T, et al. Animal models of steroid-induced osteonecrosis of the femoral head-a comprehensive research review up to 2018[J]. International Orthopaedics, 2018, 42(7): 1729-1737.

[10] 陈冠儒, 王萧枫, 王利明, 等. 兔激素性股骨头坏死造模中的致死因素与干预分析[J]. 浙江中西医结合杂志, 2012, 22(5): 342-344.
CHEN G R, WANG X F, WANG L M, et al. Analysis on fetal factors and intervention of steriod-induced femoral head necrosis in rabbits[J]. Zhejiang Journal of Integrated Traditional Chinese and Western Medicine, 2012, 22(5): 342-344(in Chinese).

[11] 王程, 苟文隆, 徐小龙, 等. 人股骨头坏死标本不同区域骨小梁的显微结构特征及病理学表现[J]. 解放军医学院学报, 2014, 35(5): 463-465.
WANG C, GOU W L, XU X L, et al. Microstructure features and pathology of trabecula in different regions of femoral head necrosis[J]. Academic Journal of Chinese PLA Medical School, 2014, 35(5): 463-465(in Chinese).

[12] HAMBLI R. Micro-CT finite element model and experimental validation of trabecular bone damage and fracture[J]. Bone, 2013, 56(2): 363-374.

[13] ZHANG G, SHENG H, HE Y X, et al. Continuous occurrence of both insufficient neovascularization and elevated vascular permeability in rabbit proximal femur during inadequate repair of steroid-associated osteonecrotic lesions[J]. Arthritis and Rheumatism, 2009, 60(10): 2966-2977.

[14] LAI Y X, LI Y, CAO H J, et al. Osteogenic magnesium incorporated into PLGA/TCP porous scaffold by 3D printing for repairing challenging bone defect[J]. Biomaterials, 2019, 197: 207-219.

[15] HA Y C, KIM H J, KIM S Y, et al. Effects of age and body mass index on the results of transtrochanteric rotational osteotomy for femoral head osteonecrosis: Surgical technique[J]. The Journal of Bone and Joint Surgery American Volume, 2011, 93(S1): 75-84.

[16] JIN S Y, MENG C Q, HE Y, et al. Curcumin prevents osteocyte apoptosis by inhibiting M1-type macrophage polarization in mice model of glucocorticoid-associated osteonecrosis of the femoral head[J]. Journal of Orthopaedic Research, 2020, 38(9): 2020-2030.

[17] WU X, FENG X, HE Y, et al. IL-4 administration exerts preventive effects via suppression of underlying inflammation and TNF-alpha-induced apoptosis in steroid-induced osteonecrosis[J]. Osteoporosis International, 2016, 27(5): 1827-1837.

[18] WANG C Q, XU H X, LIU C, et al. CaO2/gelatin oxygen slow-releasing microspheres facilitate tissue engineering efficiency for the osteonecrosis of femoral head by enhancing the angiogenesis and survival of grafted bone marrow mesenchymal stem cells[J]. Biomaterials Science, 2021, 9(8): 3005-3018.

Improved modeling strategy and result evaluation of steroid-induced femoral head necrosis in rabbits

YANG Fenghe[1], FENG Guiyu[2], LI Jinyu[2], NIU Xufeng[1,*]

(1. Key Laboratory for Biomechanics and Mechanobiology of Ministry of Education, School of Biological Science and Medical Engineering, Beihang University, Beijing 100191, China;
2. Dongzhimen Hospital, Beijing University of Chinese Medicine, Beijing 100700, China)

Abstract: The establishment of animal model of femoral head necrosis is of great significance for the clinical treatment of this disease. To successfully induce femoral head necrosis while reducing the mortality of modeling animals, this study improves the commonly used modeling method of steroid-induced femoral head necrosis in rabbits. Ninety male New Zealand white rabbits (2.8～3.2 kg) are randomly divided into control group, conventional model group and improved model group. The control group is only injected with normal saline; the conventional model group is injected with lipopolysaccharide and methylprednisolone; the improved model group is supplemented with penicillin, ranitidine, diclazuril, intestinal lubricant and lactasin based on conventional model. The results show that the death rate of the improved model group (13.3%) is much lower than that of the conventional model group (70%). The Micro-CT and histological results exhibit that the bone volume, bone volume/tissue volume ratio, bone mineral density and trabecular thickness in improved model group are significantly lower than those in control group, while the number of empty bone lacunae is significantly higher than that in control group. These results demonstrate the successful establishment of steroid-induced femoral head necrosis in rabbits.

Keywords: femoral head necrosis; steroid; animal model; Micro-CT; histological staining

Received: 2021-09-17; **Accepted**: 2021-10-22; **Published Online**: 2021-12-22 15:07

URL: kns.cnki.net/kcms/detail/11.2625.V.20211222.0902.001.html

Foundation items: National Natural Science Foundation of China (11872097,82074463); National Key Research and Development Program of China (2020YFC0122204); 111 project (B13003)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0561

旋翼/机翼气动干扰对复合式直升机性能影响

杨克龙，韩东*

(南京航空航天大学 航空学院 直升机旋翼动力学国家级重点实验室，南京 210016)

摘　　要：为研究旋翼/机翼气动干扰对双螺旋桨推进复合式直升机的高速飞行性能的影响，建立一种可快速预测复合式直升机飞行性能的模型，以加装机翼和螺旋桨的AS365N“海豚”直升机为样例，分析了400 km/h高速飞行时，旋翼/机翼气动干扰对全机飞行性能的影响及机理，并探讨了用副翼操纵配平气动干扰引起的机翼滚转力矩时，全机功率的变化规律。研究表明：气动干扰增加了机翼的诱导速度，导致机翼的升力系数降低、阻力系数增加。在旋翼不对称涡系的作用下，位于旋翼前行侧下机翼的诱导速度和升阻力系数变化比后行侧的更显著。气动干扰导致机翼升力分配从80.00%降低至71.59%，旋翼升力分配从20.04%增加至28.48%。旋翼、螺旋桨和全机功率分别增加了16.60%、1.86%和3.76%。气动干扰使机翼滚转力矩增加、旋翼滚转力矩减小，利于全机配平，但会增加全机功率。用副翼操纵来平衡由气动干扰引起的机翼滚转力矩时，旋翼侧向周期变距和阻力减小，降低了全机功率。

关　键　词：复合式直升机；气动干扰；飞行性能；功率；旋翼；机翼
中图分类号：V211.52
文献标志码：A　　**文章编号**：1001-5965(2023)07-1761-11

21世纪以来，随着对直升机高速度、大航程、高升限等性能指标的迫切需求，复合式直升机已成为国内外直升机研究热点[1-3]。由于复合式直升机加装了辅助升力装置和推进装置(通常为机翼和螺旋桨)，高速飞行时机翼和螺旋桨可为旋翼卸载，从而极大地拓展了直升机的飞行包线和飞行性能[4-6]。但加装机翼和螺旋桨也会带来诸多问题，如旋翼/机翼气动干扰就会对全机配平和飞行性能产生不利影响[7]。因此，探讨旋翼/机翼气动干扰对复合式直升机飞行性能的影响，对复合式直升机的设计就显得尤为重要。

目前，国外在旋翼/机翼气动干扰方面已经开展了相关研究。Makino等[8]用试验方法研究了悬停时旋翼/机翼气动干扰问题。由于气动干扰的存在，旋翼悬停性能随着旋翼和机翼之间的距离减小而增加，减小旋翼和机翼之间的距离，会增大机翼下洗和旋翼拉力的比值。Hamamoto等[9]采用同样的方法探讨了高速飞行时(前进比为0.7)机翼位置对旋翼/机翼气动干扰的影响，试验研究表明：减小旋翼和机翼之间的距离，机翼升力降低，阻力和滚转力矩增加。在涡方法中，Yeo[10]则用涡方法研究了旋翼/机翼气动干扰问题，并对AH-56复合式直升机飞行测试数据进行验证，取得了较准确的预测结果。Orchard和Newman[11]采用马蹄涡来模拟机翼，用纵、横向分布的涡格来模拟旋翼，计算结果和试验值的变化趋势一致，说明用涡方法可以快速准确地预测气动干扰问题。为更加精确地模拟旋翼/机翼气动干扰，Sugawara和Tanabe[12-13]采用了“rFlow3D”计算流体力学(computational fluid dynamics，CFD)软件，对加装了矩形机翼的UH-60A直升机流场进行模拟。研究表明：气动干扰导致旋翼/机翼组合体的等效阻力增加了近20%。Frey等[14]则采用CFD方法研究了当前流行的X3构型复合式直升机旋翼/

收稿日期：2021-09-18；**录用日期**：2022-01-07；**网络出版时间**：2022-01-26 11:31
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220124.1759.008.html
基金项目：国家自然科学基金(11972181)；直升机旋翼动力学国家级重点实验室基金(2005RAL20200104)
*通信作者. E-mail：donghan@nuaa.edu.cn

引用格式：杨克龙，韩东. 旋翼/机翼气动干扰对复合式直升机性能影响[J]. 北京航空航天大学学报，2023，49(7)：1761-1771.
YANG K L，HAN D. Influence of rotor/wing aerodynamic interference on performance of compound helicopters[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1761-1771 (in Chinese).

机翼气动干扰问题，旋翼/机翼气动干扰导致位于旋翼前行侧桨叶下方的机翼升力减小，在旋翼后行侧，旋翼对机翼的影响非常小。但以上CFD方法均采用风洞配平策略，没有考虑螺旋桨和平尾等的力和力矩，与复合式直升机实际飞行时的配平策略存在差异。

国内方面，万佳和陈铭[15-16]采用自由尾迹法和涡格法，对常规构型复合式直升机的旋翼/机翼气动干扰进行初步分析。研究表明：旋翼/机翼气动干扰与飞行速度有较大关系，由于气动干扰，旋翼和机翼的效率都会降低，但该研究只针对复合式直升机构型进行分析。文献[17-18]则采用动量源方法研究双螺旋桨推进复合式直升机的气动干扰问题，悬停时机身对气流的阻塞作用降低了旋翼的升力，低速前飞时旋翼/机翼之间的气动干扰较大，主要体现在旋翼下洗造成机翼上表面压力分布增大，但没有对高速飞行时气动干扰对飞行性能的影响进行研究。

相比于高速飞行状态，在悬停和中低速飞行状态时，气动干扰更显著[19]，所以多数研究集中在中低速飞行状态。在高速飞行时，只考虑了旋翼和机翼升力的配平，对气动干扰引起的阻力和滚转力矩对全机配平和飞行性能的影响探讨较少。复合式直升机作为一种主要工作在高速飞行状态的飞行器，其高速飞行时的飞行性能很大程度上决定了复合式直升机的综合使用效能。故有必要探讨高速飞行时旋翼/机翼气动干扰对复合式直升机飞行性能的影响。为此，本文建立了一种可快速预测双螺旋推进复合式直升机飞行性能模型，以加装了螺旋桨和机翼的AS365N“海豚”直升机为样例，首先，分析了高速飞行时，旋翼/机翼气动干扰对全机飞行性能的影响及机理；然后，分析了不同升力分配时气动干扰对飞行性能影响；最后，探讨了副翼操纵在配平气动干扰引起的机翼滚转力矩时，全机各部分功率的变化规律。为复合式直升机高速飞行时的旋翼/机翼气动干扰研究提供理论支持[20]。

1 复合式直升机飞行性能预测模型

大部分复合式直升机都是在常规直升机的基础上升级发展而来，如AH-56“夏衍”、X-49A“速度鹰”和X3等[21]，故建立复合式直升机飞行性能预测模型的思路也参考此方法，即先建立常规直升机飞行性能模型，再将机翼模型和螺旋桨模型等耦合进常规直升机飞行性能模型中，从而构建复合式直升机飞行性能预测模型。常规直升机飞行性能预测模型已经在文献[22-23]中进行了详细描述和验证。本节主要建立旋翼预定尾迹模型、机翼模型、螺旋桨模型、旋翼/机翼气动干扰模型和配平模型。

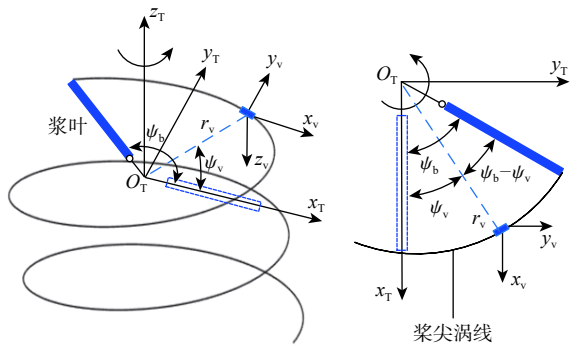
1.1 旋翼预定尾迹模型

前飞时，尾迹模型采用Beddoes[24]预定尾迹模型。图1为尾迹在旋翼桨尖平面中的示意图，其中，ψ_b为桨叶当前方位角，桨尖涡元在桨尖坐标系内的x和y方向位置分别为

$$x_v = r_v\cos\psi_v + \mu_x\Delta\psi_v \tag{1}$$

$$y_v = r_v\sin\psi_v \tag{2}$$

式中：r_v为涡元在桨叶展向的位置；ψ_v为涡元开始形成时桨叶方位角，$\Delta\psi_v$为尾迹寿命角，$\Delta\psi_v=\psi_b-\psi_v$；$\mu_x$为来流在桨尖平面内分量。



(a) 桨尖坐标系 (b) 桨尖平面俯视图

图1 尾迹在旋翼桨尖平面中的示意图

Fig. 1 Schematic of wake geometry in tip path plane of main rotor

前飞时，尾迹会后掠，在桨盘前部分区域产生上洗流，桨盘后半部分区域产生下洗流，也会出现较强的侧向入流变化[25]，故尾迹的z_v方向位移随着桨叶方位角变化而不同，可表示为非畸变尾迹和畸变尾迹的叠加：

$$z_v = -\mu_z\Delta\psi_v + \begin{cases} -\lambda_i\left(1+E\left(\cos\psi_v+\mu_x\Delta\psi_v\Big/2r_v-\left|y'^3\right|\right)\right)\Delta\psi_v & x_v<-r_v\cos\psi_v \\ -2\lambda_i\left(1-E\left|y'^3\right|\right)\Delta\psi_v & \cos\psi_v>0 \\ -2\lambda_i x_v\left(1-E\left|y'^3\right|\right)\Big/\mu_x & \text{其他} \end{cases} \tag{3}$$

式中：μ_z为来流在垂直于桨尖平面的分量；λ_i为旋翼无量钢化诱导速度；x'和y'为涡元无量纲化的纵向和侧向位置（以尾迹收缩半径尾迹为基准）；E为经验修正系数，和尾迹倾斜角有关。

尾迹形状确定后，可通过Biot-Savart定律[26]计算桨尖涡元对空间任意一点的诱导速度。图2为涡线对任意点诱导速度，由于桨尖涡线为连续曲线，在实际计算中，将其离散成微小的直线涡段，流场中所有涡线对空间任一点诱导速度可归结为直

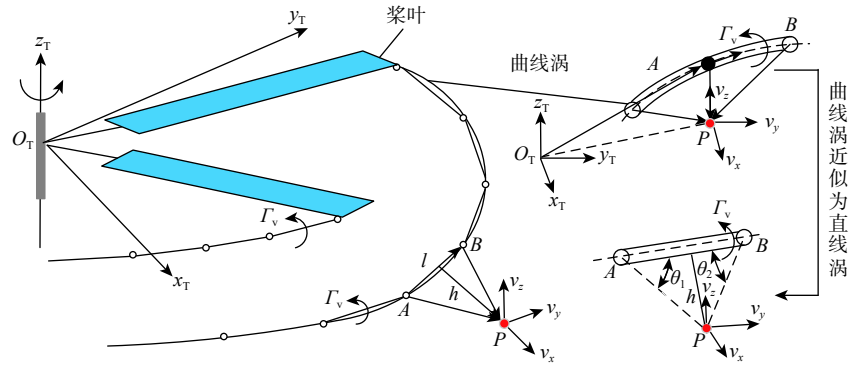


图 2　涡线对任意点诱导速度

Fig. 2　Induced velocity at an arbitrary point by vortex filament

线涡段对空间点诱导速度之和。当计算点趋近涡线时, 由 Biot-Savart 定律计算得到的诱导速度与实际值相差非常大, 为准确模拟涡元的诱导速度, 需采用涡核模型进行修正。采用 Lamb-Oseen 涡核模型[27], 得到:

$$v(r)=\frac{\Gamma_{\mathrm{v}} r_{\mathrm{s}}}{2\pi\sqrt{\left(r_{\mathrm{c}}^4+r_{\mathrm{s}}^4\right)}} \tag{4}$$

式中: $v(r)$为距离涡核r处的诱导速度; Γ_{v}为涡元环量; r_{s}为计算点至涡核的距离; r_{c}为涡核起始半径。

由图 2 可知, 假设涡段起点和终点坐标分别为点A和点B, 长度为l, 对涡核修正后, 则涡段对点P的诱导速度为

$$\begin{Bmatrix} v_x \\ v_y \\ v_z \end{Bmatrix}=\frac{\Gamma_v}{4\pi}\frac{h}{\sqrt{\left(r_{\mathrm{c}}^4+h^4\right)}}\left(\cos\theta_1-\cos\theta_2\right)\begin{Bmatrix} \boldsymbol{e}_x \\ \boldsymbol{e}_y \\ \boldsymbol{e}_z \end{Bmatrix} \tag{5}$$

式中: $\boldsymbol{e}$为单位向量; h为点P到涡段AB的距离; θ_1和θ_2分别为AP和BP连线与涡元的夹角。先对单个涡线积分, 再对桨叶片数积分, 就可得到旋翼尾迹在空间任意一点的诱导速度。

1.2　机翼模型

机翼模型基于升力面理论[28]和查翼型气动力表方法[29]建立。机翼翼型的弯度和厚度影响在查翼型气动力表方法中进行了考虑。涡格法(vortex lattice method, VLM)是升力面理论中一种比较实用的数值计算方法。在复合式直升机中, 由于气动干扰, 左、右机翼升力分布存在不对称性, 左、右机翼的涡格中的马蹄涡涡强分布也存在不对称性。为方便探讨气动干扰问题, 在计算诱导速度时, 需对左、右机翼分别进行计算, 这里以右半机翼为例进行建模, 左半机翼与右半机翼类似。

图 3 为机翼涡格示意图, 在机翼坐标系$O_{\mathrm{W}}x_{\mathrm{W}}y_{\mathrm{W}}$平面内, 将左、右半机翼分别划分为$n$个涡格, 右半翼上离散马蹄涡(有限长直线涡$AB$、无限长直线涡$A\infty$和$B\infty$)的环量分别用$\Gamma_1^{\mathrm{R}},\Gamma_2^{\mathrm{R}},\cdots,\Gamma_n^{\mathrm{R}}$来表示。同理, 左半翼上离散马蹄涡环量分别用$\Gamma_1^{\mathrm{L}},\Gamma_2^{\mathrm{L}},\cdots,\Gamma_n^{\mathrm{L}}$来表示。右半翼上涡格的马蹄涡在第$i$个涡格控制点处产生的诱导速度为

$$v_{ij}^{\mathrm{R}}(u,v,w)=\sum_{j=1}^{n}C_{ij}(u,v,w)\Gamma_j^{\mathrm{R}} \tag{6}$$

式中: $C_{ij}(u,v,w)$为诱导速度影响系数矩阵的元素, 可由 Biot-Savart 定律计算得到[28]。

对于每一个控制点, 满足物面不可穿透条件, 将翼面线化边界条件应用于第i个控制点上, 得到$v_{ij}^{\mathrm{R}}(u,v,w)$的垂向分量$w$与机翼迎角$\alpha_{\mathrm{wing}}$的关系为

$$\frac{v_{ij}^{\mathrm{R}}(w)}{V}=-\alpha_{\mathrm{wing}} \tag{7}$$

式中: V为来流速度;

联立式 (6) 和式 (7) 得到:

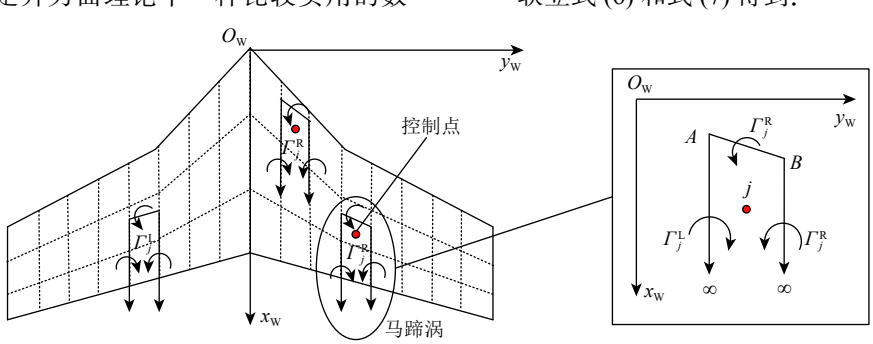


图 3　机翼涡格示意图

Fig. 3　Schematic of vortex lattice of wing

$$\sum_{j=1}^{n} C_{ij}(w)\Gamma_j^{\mathrm{R}} = -V\alpha_{\mathrm{wing}} \tag{8}$$

求解式(8),可得到右半机翼上每个涡格的环量。

右半机翼上距离机翼中轴线$O_{\mathrm{W}}x_{\mathrm{W}}$处的剖面(翼型)环量可表示为

$$\Gamma^{\mathrm{R}}(y) = \sum_{m=1}^{M} \Gamma_m^{\mathrm{R}} \tag{9}$$

式中:M为涡格行数。

根据翼型升力系数定义和Kutta-Zhukovsky定理,可得剖面(翼型)升力系数和有效迎角$\alpha_{\mathrm{eff}}(y)$的关系为[30]

$$C_L = \frac{2\Gamma^{\mathrm{R}}(y)}{Vc(y)} = C_{L_\alpha}\left[\alpha_{\mathrm{eff}}(y) - \alpha_{L=0}(y)\right] \tag{10}$$

式中:$c(y)$为展向y处的弦长;C_{L_α}为翼型升力线斜率;$\alpha_{L=0}(y)$为y处的零升迎角。

整理式(10),得到剖面处翼型的有效迎角为

$$\alpha_{\mathrm{eff}}(y) = \frac{2\Gamma^{\mathrm{R}}(y)}{Vc(y)C_{L_\alpha}} + \alpha_{L=0}(y) \tag{11}$$

根据翼型的迎角和来流马赫数,通过查翼型气动力表得到该剖面处翼型气动力。然后将剖面的气动力沿翼展积分得到右半机翼的气动力。左半机翼的气动力也可通过以上方法得到,最后将左、右半机翼的气动力相加,得到整个机翼的气动力。

副翼长度为机翼弦长的25.0%,位于机翼展向位置的55.0%~75.0%处。副翼偏转角度δ_{ail}引起的机翼升、阻力系数变化为[31]

$$\begin{cases} \Delta C_{L_{\mathrm{wing}}} = C_{L_{\mathrm{wing}}}^{\delta}\delta_{\mathrm{ail}} \\ \Delta C_{D_{\mathrm{wing}}} = 0.008 - 0.003C_{L_{\mathrm{wing}}} + 0.01C_{L_{\mathrm{wing}}}^2 \end{cases} \tag{12}$$

式中:$C_{L_{\mathrm{wing}}}^{\delta}$为副翼操纵面偏转角度$\delta_{\mathrm{ail}}$引起的机翼升力系数变化率;$C_{L_{\mathrm{wing}}}$为施加副翼操纵后机翼升力系数。

1.3 螺旋桨模型

螺旋桨模型基于组合叶素-动量理论[21]建立。相比于旋翼,螺旋桨半径小、转速和桨叶载荷高,而且工作在轴流状态。除轴向诱导速度外,还需考虑旋转平面内的周向诱导速度。图4为螺旋桨桨叶叶素所受气动力。图中:V_z和V_y分别为叶素的垂向和切向速度,r_{prop}为螺旋桨半径,θ_{prop}、ϕ_{prop}和α_{prop}分别为总距角、入流角和迎角。

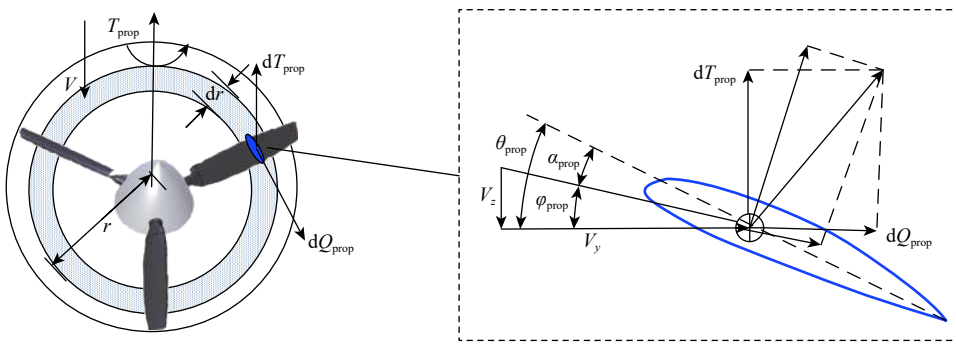


图4 螺旋桨桨叶叶素气动力

Fig. 4 Aerodynamics of blade element of propeller

根据动量和动量矩定理,距螺旋桨桨毂中心距离为r处的环形微圆上的拉力和扭矩为[32]

$$\begin{cases} \mathrm{d}T_{\mathrm{prop}} = 4\rho\pi r_{\mathrm{prop}}(V+v_{\mathrm{i}})v_{\mathrm{i}}\mathrm{d}r \\ \mathrm{d}Q_{\mathrm{prop}} = 4\rho\pi r_{\mathrm{prop}}^2 V(V+v_{\mathrm{i}})u_{\mathrm{i}}\mathrm{d}r \end{cases} \tag{13}$$

式中:v_{i}和u_{i}为轴向和周向诱导速度。

由叶素理论建立的拉力和扭矩的表达式[26]为

$$\begin{cases} \mathrm{d}T_{\mathrm{prop}} = \dfrac{1}{2}\rho V^2 c_{\mathrm{prop}}N_{\mathrm{prop}}\left(C_L\cos\phi_{\mathrm{prop}} - C_D\sin\phi_{\mathrm{prop}}\right)\mathrm{d}r \\ \mathrm{d}Q_{\mathrm{prop}} = \dfrac{1}{2}\rho V^2 c_{\mathrm{prop}}N_{\mathrm{prop}}\left(C_L\sin\phi_{\mathrm{prop}} + C_D\cos\phi_{\mathrm{prop}}\right)r\mathrm{d}r \end{cases} \tag{14}$$

式中:c_{prop}为螺旋桨桨叶弦长;N_{prop}为螺旋桨桨叶片数;C_L和C_D分别为桨叶叶素的升阻力系数。

通过对θ_{prop}进行迭代,改变螺旋桨拉力和力矩,直到式(13)和式(14)中的拉力和力矩相等,从而确定轴向和周向诱导速度,然后将所有桨叶叶素的气动力沿展向和周向积分,得到整个螺旋桨的拉力和力矩。

1.4 旋翼/机翼气动干扰模型

直升机高速前飞时,旋翼桨尖压缩性已在查翼型气动力系数表法中进行了考虑。旋翼尾迹倾斜角较大,旋翼桨尖涡系不会与机翼直接碰撞[19],故模型中没有考虑旋翼桨尖涡与机翼之间的碰撞。旋翼/机翼之间的气动干扰可通过旋翼涡系和机翼涡系的相互诱导速度来模拟。求解涡系环量时,将机翼涡与旋翼涡同时考虑,联立求解。环量与诱导速度关系为

$$\begin{bmatrix} \boldsymbol{C}_{\mathrm{RR}}\boldsymbol{C}_{\mathrm{WR}} \\ \boldsymbol{C}_{\mathrm{RW}}\boldsymbol{C}_{\mathrm{WW}} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{\Gamma}_{\mathrm{R}} \\ \boldsymbol{\Gamma}_{\mathrm{W}} \end{bmatrix} = -\begin{bmatrix} \boldsymbol{B}_{\mathrm{R}} \\ \boldsymbol{B}_{\mathrm{W}} \end{bmatrix} \tag{15}$$

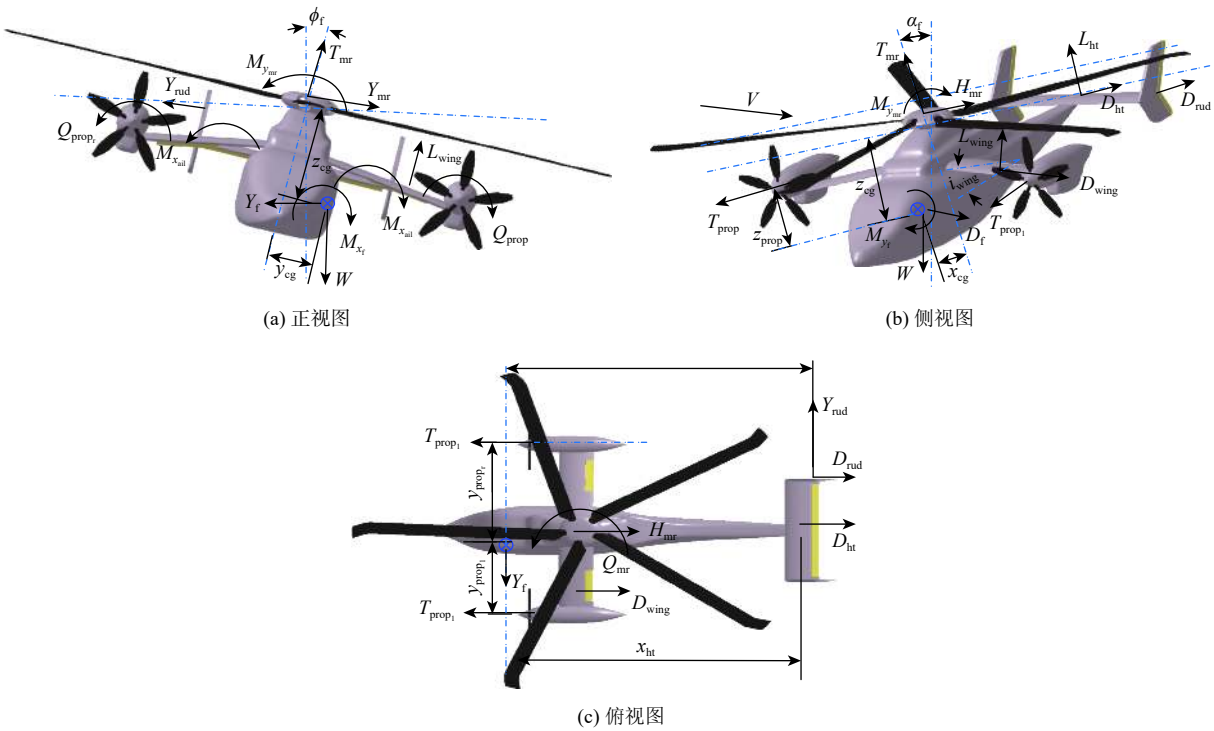
式中: $\boldsymbol{C}_{\mathrm{RR}}$和$\boldsymbol{C}_{\mathrm{WW}}$为旋翼涡系和机翼涡系分别对自身控制点的影响系数; $\boldsymbol{C}_{\mathrm{RW}}$为旋翼涡系对机翼控制点的诱导系数; $\boldsymbol{C}_{\mathrm{WR}}$为机翼涡系对旋翼控制点的诱导速度; $\boldsymbol{\Gamma}_{\mathrm{R}}$和$\boldsymbol{\Gamma}_{\mathrm{W}}$分别为待求解的旋翼涡系环量和机翼涡系环量; $\boldsymbol{B}_{\mathrm{R}}$和$\boldsymbol{B}_{\mathrm{W}}$为旋翼和机翼制点处的法向速度。

1.5　配平模型

在常规直升机的配平基础上[26], 增加了机翼和螺旋桨的力和力矩。根据六自由度平衡方程, 复合式直升机的配平方程为

$$\begin{cases}\sum F_x = D_{\mathrm{f}} + D_{\mathrm{wing}} + D_{\mathrm{ht}} + H_{\mathrm{mr}} - T_{\mathrm{mr}}\alpha_{\mathrm{f}} - L_{\mathrm{wing}} i_{\mathrm{wing}} - T_{\mathrm{prop_l}} - T_{\mathrm{prop_r}} = 0\\ \sum F_y = Y_{\mathrm{mr}} + Y_{\mathrm{f}} + T_{\mathrm{mr}}\phi_{\mathrm{f}} + Y_{\mathrm{rud}} = 0\\ \sum F_z = T_{\mathrm{mr}} + L_{\mathrm{wing}} + L_{\mathrm{ht}} - W = 0\\ \sum M_x = M_{x_{\mathrm{mr}}} + M_{x_{\mathrm{f}}} + W\left(z_{\mathrm{cg}}\phi_{\mathrm{f}} - y_{\mathrm{cg}}\right) + M_{x_{\mathrm{ail}}} + Q_{\mathrm{prop_l}} + Q_{\mathrm{prop_r}} = 0\\ \sum M_y = M_{y_{\mathrm{mr}}} + M_{y_{\mathrm{f}}} + W\left(z_{\mathrm{cg}}\alpha_{\mathrm{f}} - x_{\mathrm{cg}}\right) - D_{\mathrm{f}} z_{\mathrm{cg}} - L_{\mathrm{ht}} x_{\mathrm{ht}} - \left(T_{\mathrm{prop_l}} + T_{\mathrm{prop_r}}\right) z_{\mathrm{prop}} = 0\\ \sum M_z = Q_{\mathrm{mr}} - T_{\mathrm{prop_l}} y_{\mathrm{prop_l}} + T_{\mathrm{prop_r}} y_{\mathrm{prop_r}} = 0\end{cases} \tag{16}$$

式中: D_{f}为机身阻力; D_{wing}为机翼阻力; D_{ht}为平尾阻力; H_{mr}为旋翼阻力; T_{mr}为旋翼拉力; α_{f}为机身俯仰角; L_{wing}为机翼升力; i_{wing}为机翼安装角; $T_{\mathrm{prop_l}}$为左螺旋桨拉力; $T_{\mathrm{prop_r}}$为右螺旋桨拉力; Y_{mr}为旋翼侧向力; Y_{f}为机身侧向力; ϕ_{f}为机身侧倾角; Y_{rud}为方向舵升力; L_{ht}为平尾升力; W为全机重力; $M_{x_{\mathrm{mr}}}$为旋翼滚转力矩; $M_{x_{\mathrm{f}}}$为机身滚转力矩; z_{cg}和y_{cg}分别为旋翼桨毂中心到全机重心的垂向和侧向距离; $M_{x_{\mathrm{ail}}}$为副翼操纵引起的力矩; $Q_{\mathrm{prop_l}}$和$Q_{\mathrm{prop_r}}$分别为左、右螺旋桨的力矩; $M_{y_{\mathrm{mr}}}$为旋翼俯仰力矩; $M_{y_{\mathrm{f}}}$为机身俯仰力矩; x_{cg}为旋翼桨毂中心到全机重心的水平距离; x_{ht}为平尾压心到全机重心的水平距离; z_{prop}为螺旋桨桨毂中心到全机重心的垂向距离; Q_{mr}为旋翼反扭矩; $y_{\mathrm{prop_l}}$和$y_{\mathrm{prop_r}}$分别为左、右螺桨桨毂中心到全机重心的侧向距离。基于双螺旋桨推进的复合式直升机构型及其受力情况如图 5 所示, 其中, ⓧ为全机重心。



(a) 正视图　　(b) 侧视图

(c) 俯视图

图 5　复合式直升机受力和力矩

Fig. 5　Forces and moments acting on compound helicopter

2　模型验证

由于目前没有公开的关于双螺旋桨推进复合式直升机飞行性能的相关测试数据, 故分别对旋翼预定尾迹模型、机翼模型和螺旋桨模型进行验证。

图 6 为 HU-1D 直升机旋翼尾迹形状的预测值、试验值和自由尾迹方法的预测值[33]。图 7 为旋翼尾迹诱导速度的预测值和试验值[34](R为旋翼桨叶

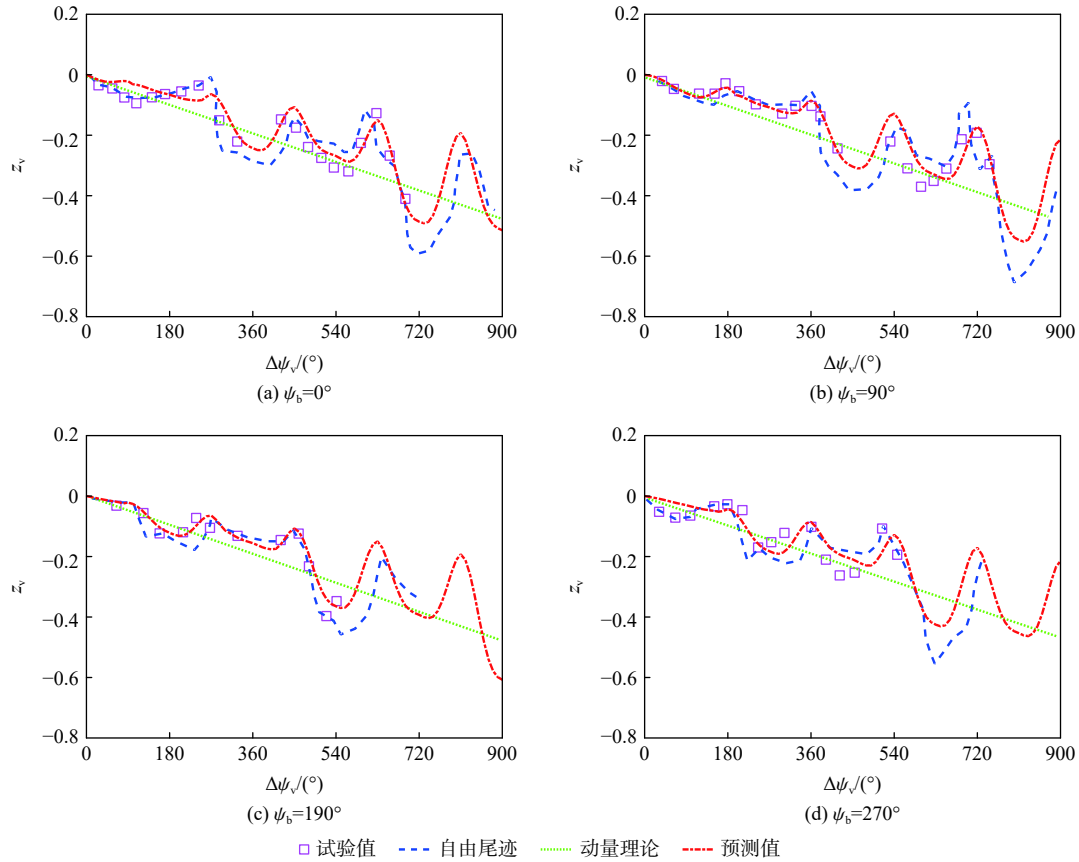


图 6 HV−ID 直升机旋翼尾迹形状预测值、试验值和自由尾迹预测值

Fig. 6 Wake geometry of HV−ID helicopter main roctor among prediction, experimental value and free wake prediction

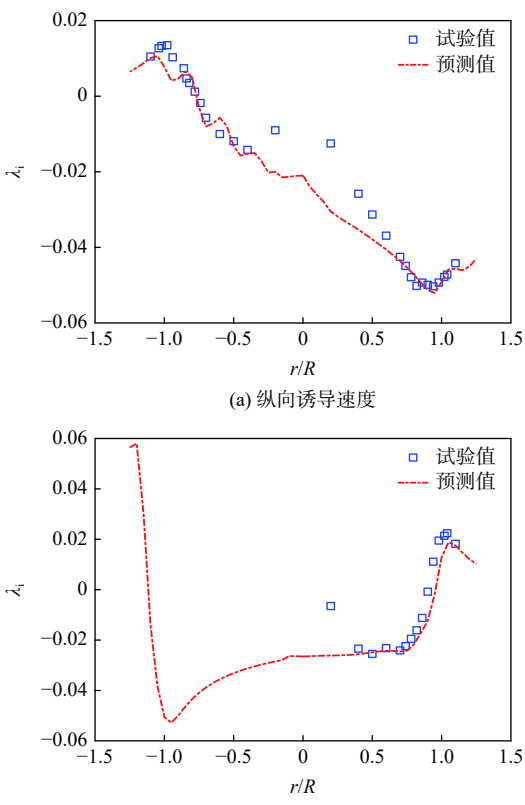


图 7 旋翼尾迹诱导速度预测值与试验值

Fig. 7 Induced velocity among of main roctor prediction, experimental value and free wake prediction

半径)。可见,本文建立的模型预测得到的旋翼尾迹形状和诱导速度与试验值吻合较好。

不同机翼迎角时,用涡格法预测到的机翼升力系数与试验值[35]和计算值[36],如图 8 所示,可见,预测值与试验值吻合较好,故机翼模型可用旋翼/机翼气动干扰研究。

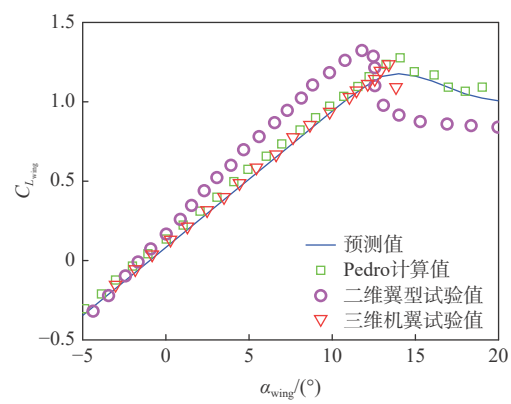


图 8 机翼升力系数的预测值、试验值和算例值

Fig. 8 Wing lift coefficient among prediction, experimental value and calculation example

图 9 为不同前进比λ_{prop}时螺旋桨拉力系数$C_{t_{prop}}$和效率η_{prop}的预测值与试验值[37]的曲线。预测值与试验值吻合较好,说明螺旋桨模型可用于复合式直升机飞行性能的研究。

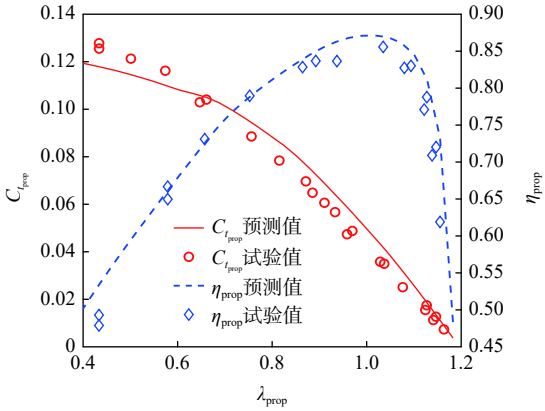


图 9　螺旋桨拉力系数和效率的预测值与试验值

Fig. 9　Wing lift coefficient among prediction, experimental value and calculation example

3　旋翼/机翼气动干扰对飞行性能的影响

样例复合式直升机参数如表 1 所示。飞行速度为 400 km/h; 起飞质量为 8 300 kg。假设无气动干扰时, 机翼承担全机 80.00% 的全机升力, 飞行高度为海平面。采用功率增加率来衡量气动干扰对复合式直升机飞行性能的影响, 功率增加率的定义为

$$\eta = \left(\frac{P}{P_{\text{base}}} - 1\right) \times 100\% \tag{17}$$

式中: P_{base}为未考虑气动干扰时各部件功率; P为考虑气动干扰时全机各部件功率。

表 1　样例复合式直升机参数

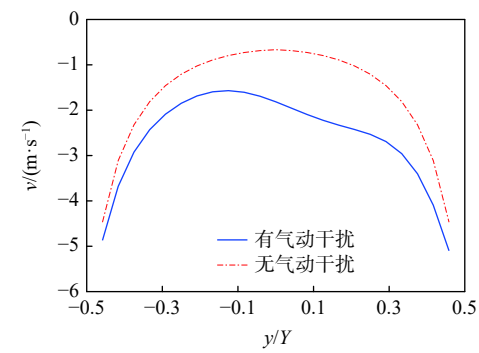
Table 1　Parameters of a compound helicopter

部件	半径/展长/m	平均弦长/m	转速/(rad·s^{-1})	负扭/(°)	安装角/(°)	翼型	桨叶片数/机翼展弦比
旋翼	5.97	0.385	37.7	−10		OA212/209	5
机翼	8.48	1.25		0	4	NACA65-21	6.78
平尾	4.0	0.90		0		NACA65-210	6.78
螺旋桨	1.35	0.25	157.1	−35	−10	Clark Y	5

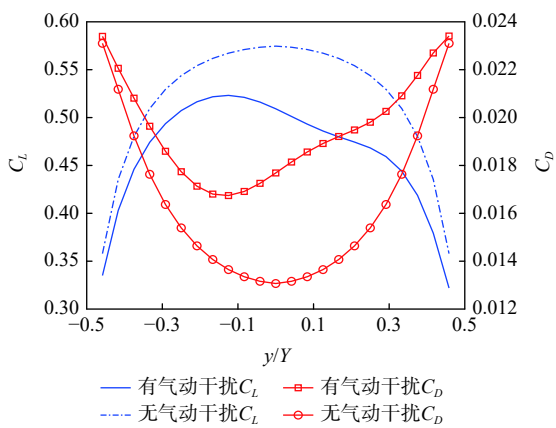
在笔记本电脑(CPU 为英特尔 i5-8250U, 1.60 GHz)上运行本文模型时, 仅需 11 min 就可计算出旋翼/机翼气动干扰对飞行性能的影响。Shi 等[38]在工作站(CPU 为英特尔 Core2 E7200, 2.60 GHz)上运行基于 CFD 建立的旋翼尾迹和流场求解模型时, 至少耗时 10 h。由此可见, 本文模型是一种实用高效的分析方法, 能快速预测复合式直升机飞行性能。

图 10 为机翼剖面(翼型)的诱导速度 v 和升阻力系数沿翼展的分布情况, 其中, Y为机翼翼展, 左、右半机翼分别位于旋翼后行侧和前行侧下方。不考虑旋翼/机翼气动干扰时, 机翼升力沿翼展均匀分布, 故机翼诱导速度分布为椭圆分布。考虑气动干扰时, 机翼诱导速度增加, 由于旋翼不对称涡系的作用, 右半机翼的诱导速度的增加比左半翼的显著[13]。如图 10(b) 所示, 机翼诱导速度的变化导致机翼剖面(翼型)的升阻力系数也相应改变: 翼型升力系数降低、阻力系数增加。由于右半机翼位于旋翼前行侧下方, 旋翼对机翼的诱导速度大, 导致右半翼的升阻力系数的变化比左半翼的更显著。

图 11 为气动干扰对旋翼和机翼的升力分配及滚转力矩的影响。气动干扰导致机翼和升力分配从 80.00% 降低至 71.59%, 旋翼升力分配从 20.04% 增加至 28.48%。这是因为旋翼尾迹对机翼的气动干扰增加了机翼的诱导速度, 导致机翼



(a) 诱导速度



(b) 升阻力系数

图 10　机翼剖面(翼型)的诱导速度和升阻力系数沿翼展分布

Fig. 10　Distribution of induced velocity, lift and drag coefficients of wing sections (airfoils) along wingspan

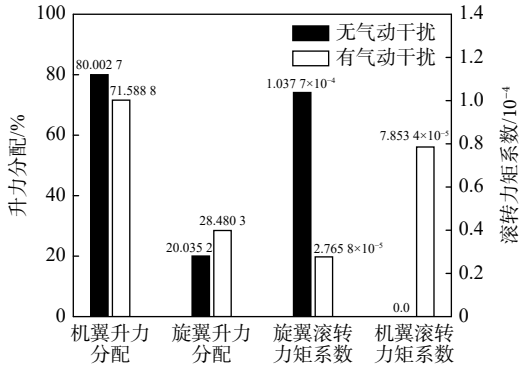


图 11 气动干扰对旋翼和机翼的升力分配及滚转力矩的影响

Fig. 11 Effect of aerodynamic interference on the lift share and rolling moment of main rotor and wing

有效迎角减小，从而导致机翼升力减小，旋翼升力增大保持垂向力配平。考虑气动干扰时，由于机翼左、右升力分布不对称，机翼产生大小为7.85×10^{-5}的滚转力矩（为方便和旋翼滚转力矩进行对比，机翼的滚转力矩系数以旋翼参数为基准进行无量纲化），旋翼不需要产生更多的滚转力矩来保持滚转力矩的配平，故旋翼的滚转力矩由1.04×10^{-4}降低至2.77×10^{-5}。

图 12 为气动干扰对旋翼、螺旋桨和全机功率的影响。P_{mr}为旋翼功率，P_{prop}为螺旋桨功率，P_{ch}为全机总功率，则$P_{ch}=P_{mr}+P_{prop}$。由于气动干扰，旋翼升力分配增加，旋翼功率也由 365.80 kW 增加至 426.54 kW（增加了 16.60%）；旋翼和机翼的等效阻力也增加[12]，螺旋桨需提供更多的推进力来平衡阻力，故螺旋桨功率增加了 1.86%。虽然气动干扰导致旋翼功率增加了 16.60%，但由于高速飞行时螺旋桨功率占全机绝大部分功率，故全机功率增加了 3.76%。综合图 11 和图 12，气动干扰导致机翼产生滚转力矩，减小了旋翼滚转力矩，利于全机配平，但气动干扰会增加全机功率，不利于飞行性能提升。

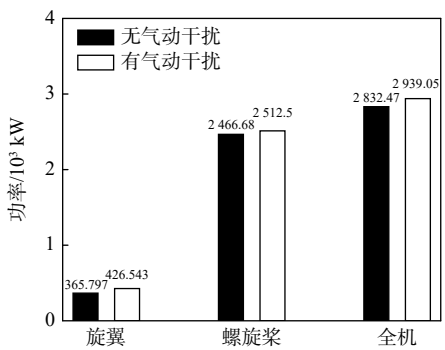
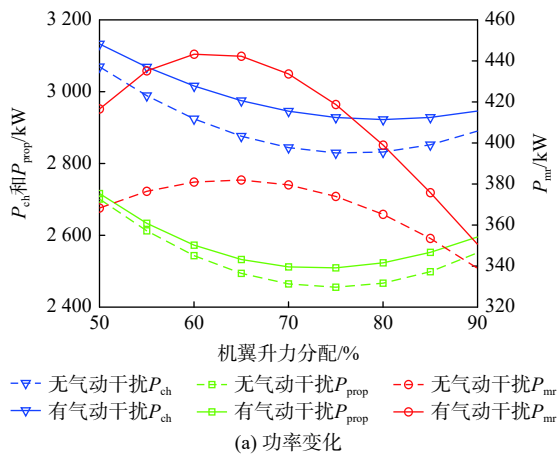


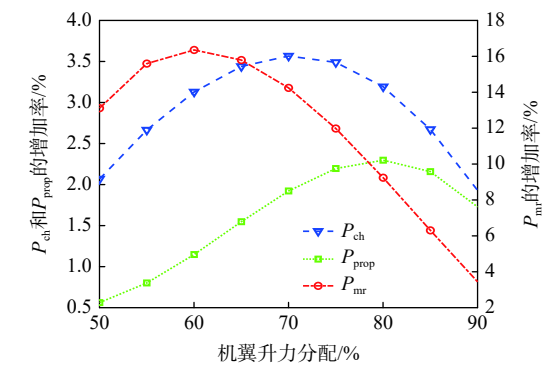
图 12 气动干扰对旋翼、螺旋桨和全机功率的影响

Fig. 12 Effect of aerodynamic interference on the power of main rotor propeller and aircraft

不同升力分配时气动干扰对旋翼、螺旋桨和全机功率的影响，如图 13 所示。气动干扰增加了旋翼、螺旋桨和全机的功率，如图 13(a) 所示。有气动干扰和无气动干扰时，全机最小功率对应的机翼升力分配均为 75.02%，气动干扰没有改变最小全机功率对应的升力分配。不同机翼升力分配时，气动干扰引起的功率增加率是不同的，如图 13(b) 所示。当机翼升力分配在 50.00%～70.00% 之间时，气动干扰引起的旋翼功率增加率较大；当升力分配为 70.00% 时，气动干扰引起的全机功率增加率最大；当升力分配大于 80.00% 时，气动干扰引起的功率增加率逐渐降低。这是因为气动干扰会改变旋翼和机翼之间的升力分配，而不同升力分配时旋翼和机翼的效率不同，旋翼螺旋桨和全机功率也不同[6]。



(a) 功率变化

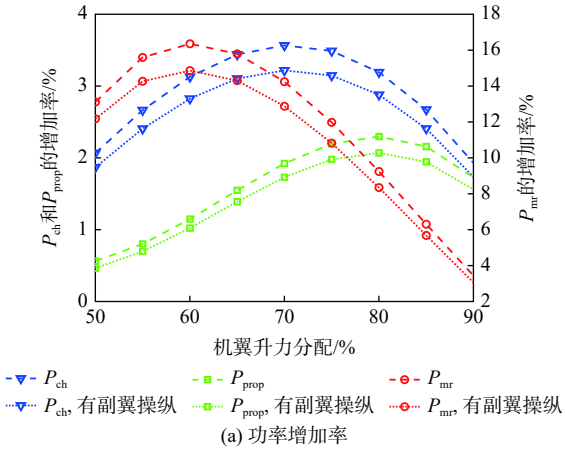


(b) 功率增加率

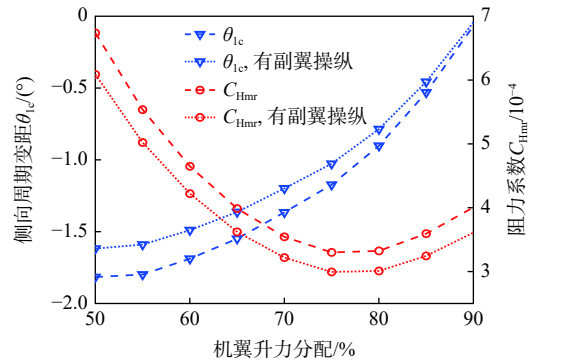
图 13 功率和功率增加率对比

Fig. 13 Comparison of power and power increase ratio

本节没有用副翼操纵来配平机翼滚转力矩。为此，考虑气动干扰时，用副翼操纵来配平机翼滚力矩时，旋翼、螺旋桨和全机功率变化情况，如图 14 所示。有副翼操纵时，旋翼、螺旋桨和全机功率均降低，如图 14(a) 所示。副翼操纵没有改变气动干扰引起的旋翼功率变化，而是降低了旋翼侧向周期变距θ_{1c}，减小了旋翼阻力系数C_{Hmr}（见图 14(b)），从而降低了螺旋桨功率和全机功率。



(a) 功率增加率



(b) 侧向周期变距和阻力系数

图 14　功率、侧向周期变距和阻力系数对比

Fig. 14　Comparison of power, cyclic pitch angle and drag coefficient

4　结　论

1) 气动干扰增加了机翼的诱导速度, 导致机翼的升力系数降低、阻力系数增加。由于旋翼不对称涡系的作用, 位于旋翼前行侧下机翼的诱导速度和升阻力系数的变化比后行侧的更显著。

2) 气动干扰导致机翼升力减小, 旋翼升力增加。机翼升力分配从 80.00% 降低至 71.59%, 旋翼升力分配从 20.04% 增加至 28.48%。旋翼、螺旋桨和全机功率分别增加了 16.60%、1.86% 和 3.76%。

3) 气动干扰导致机翼升力分布不对称, 使机翼滚转力矩增加、旋翼滚转力矩减小, 有利于全机配平, 但气动干扰会增加全机功率, 不利于飞行性能。

4) 用副翼操纵来平衡由气动干扰引起的机翼滚转力矩时, 旋翼侧向周期变距和阻力减小, 从而降低了螺旋桨功率和全机功率。

参考文献（References）

[1] 倪先平. 未来直升机技术发展展望[J]. 航空制造技术, 2008, 51(3): 32-37.
NI X P. Development prospect of helicopter technology in the future[J]. Aeronautical Manufacturing Technology, 2008, 51(3): 32-37(in Chinese).

[2] 吕春雷, 吴希明, 武庆中. 先进直升机发展现状及设计革新[J]. 航空制造技术, 2013, 56(17): 26-29.
LYU C L, WU X M, WU Q Z. Development status and design innovation of advanced helicopter[J]. Aeronautical Manufacturing Technology, 2013, 56(17): 26-29(in Chinese).

[3] 吴希明. 高速直升机发展现状、趋势与对策[J]. 南京航空航天大学学报, 2015, 47(2): 173-179.
WU X M. Current status, development trend and countermeasure for high-speed rotorcraft[J]. Journal of Nanjing University of Aeronautics and Astronautics, 2015, 47(2): 173-179(in Chinese).

[4] HIRSCHBERG M J. X2, X3, S-97 and X-49: The battle of the compounds is joined[J]. Vertiflite, 2010, 56(4): 16-22.

[5] 杨克龙, 韩东, 石启鹏. 升推力装置对复合直升机飞行性能的提升[J]. 航空动力学报, 2020, 35(11): 2429-2439.
YANG K L, HAN D, SHI Q P. Lifting and propulsion devices for flight performance improvement of a compound helicopter[J]. Journal of Aerospace Power, 2020, 35(11): 2429-2439(in Chinese).

[6] YANG K L, HAN D, SHI Q P. Study on the lift and propulsive force shares to improve the flight performance of a compound helicopter[J]. Chinese Journal of Aeronautics, 2022, 35(1): 365-375.

[7] LYNN R R. Wing-rotor interactions[J]. Journal of Aircraft, 1966, 3(4): 326-332.

[8] MAKINO K, AKASAKA T, HAMAMOTO Y, et al. Experimental investigation of the rotor-wing aerodynamic interaction on a compound helicopter in hover[C]//6th Asian/Australian Rotorcraft Forum & Heli Japan. Kanazawa: Vertical Flight Society, 2017.

[9] HAMAMOTO Y, AKASAKA T, TANABE Y. Experimental investigation of the rotor-wing aerodynamic interaction on a compound helicopter in high advance ratio[C]//6th Asian/Australian Rotorcraft Forum & Heli Japan. Kanazawa: Vertical Flight Society, 2017.

[10] YEO H. Design and aeromechanics investigation of compound helicopters[J]. Aerospace Science and Technology, 2019, 88: 158-173.

[11] ORCHARD M, NEWMAN S. Some design issues for the optimisation of the compound helicopter configuration[C]//56th American Hclicopter Society Annual Forum. Kanazawa: Vertical Flight Society, 2000.

[12] SUGAWARA H, TANABE Y. Numerical investigation of rotor/wing aerodynamic interactions at high advance ratios[J]. Journal of Aircraft, 2019, 56(6): 2285-2298.

[13] SUGAWARA H, TANABE Y. A study of rotor wing aerodynamic interaction at high speed flight on a compound helicopter[C]//6th Asian/Austrailian Rotorcraft Forum & Heli Japan. Kanazawa: Vertical Flight Society, 2017.

[14] FREY F, THIEMEIER J, ÖHRLE C, et al. Aerodynamic interactions on airbus helicopters' compound helicopter RACER in cruise flight[J]. Journal of the American Helicopter Society, 2020, 65(4): 1-14.

[15] 万佳, 陈铭. 机翼位置对复合式直升机旋翼-机翼干扰的影响[J]. 北京航空航天大学学报, 2009, 35(5): 519-522.
WAN J, CHEN M. Influence of wing location on rotor-wing interaction of compound helicopter[J]. Journal of Beijing University of Aeronautics and Astronautics, 2009, 35(5): 519-522(in Chinese).

[16] 万佳, 陈铭. 复合式直升机前飞状态旋翼-机翼干扰计算分析[J]. 航空动力学报, 2009, 24(11): 2459-2464.
WAN J, CHEN M. Rotor-wing interaction analysis of a compound helicopter in forward flight[J]. Journal of Aerospace Power, 2009, 24(11): 2459-2464(in Chinese).

[17] 赵寅宇, 黎鑫, 史勇杰, 等. 双拉力螺旋桨构型复合式高速直升机旋翼/螺旋桨干扰流场分析[J]. 南京航空航天大学学报, 2017, 49(2): 154-164.
ZHAO Y Y, LI X, SHI Y J, et al. Analysis on rotor-propellers interaction flowfield for compound double-thust-propeller high-speed helicopters[J]. Journal of Nanjing University of Aeronautics and Astronautics, 2017, 49(2): 154-164(in Chinese).

[18] 申遂愿, 朱清华, 朱振华, 等. 高速直升机旋翼/螺旋桨/机身干扰特性分析[J]. 航空工程进展, 2020, 11(1): 46-55.
SHEN S Y, ZHU Q H, ZHU Z H, et al. Analysis of rotor/propeller/fuselage interference characteristics of high-speed helicopter[J]. Advances in Aeronautical Science and Engineering, 2020, 11(1): 46-55(in Chinese).

[19] LEISHMAN J G, BI N P. Experimental investigation of rotor/lifting surface interactions[J]. Journal of Aircraft, 1994, 31(4): 846-854(in Chinese).

[20] 黄明其, 徐栋霞, 何龙, 等. 常规旋翼构型复合式高速直升机发展概况及关键技术[J]. 航空动力学报, 2021, 36(6): 1156-1168.
HUANG M Q, XU D X, HE L, et al. Development overview and key technologies of high speed hybrid helicopter with single main rotor[J]. Journal of Aerospace Power, 2021, 36(6): 1156-1168(in Chinese).

[21] ORMISTON R A. Revitalising advanced rotorcraft research and the compound helicopter[J]. The Aeronautical Journal, 2016, 120(1223): 83-129.

[22] 张勇刚, 崔钊, 韩东, 等. 加装格尼襟翼旋翼的直升机飞行性能[J]. 航空学报, 2016, 37(7): 2208-2217.
ZHANG Y G, CUI Z, HAN D, et al. Flight performance of helicopter rotors with Gurney flaps[J]. Acta Aeronautica et Astronautica Sinica, 2016, 37(7): 2208-2217(in Chinese).

[23] 董晨, 韩东, 杨克龙. 独立桨距控制对直升机飞行性能的影响[J]. 航空学报, 2018, 39(10): 222075.
DONG C, HAN D, YANG K L. Effect of individual blade pitch control on flight performance of helicopters[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(10): 222075(in Chinese).

[24] BEDDOES T S. A wake model for high resolution airloads[C]//2nd International Conference on Basic Rotorcraft Research. New York: Vertical Flight Society, 1985.

[25] HEYSON H H, KATZOFF S. Induced velocities near a lifting rotor with nonuniform disk loading: NACA-TN-3690[R]. Hampton: Hampton NACA Langley Aerorktical Laboratory, 1957: 26-34.

[26] LEISHMAN J G. Principles of helicopter aerodynamics[M]. 2nd ed. Cambridge: Cambridge University Press, 2006: 599-601.

[27] BHAGWA M J, LEISHMAN J G. Generalized viscous vortex model for applications[C]//The 58th Annual Forum and Technology Display of the American Helicopter Society. Montreal: Vertical Flight Society, 2002.

[28] CUMMINGS R M, MASON W H, MORTON S A, et al. Applied computational aerodynamics: A modern engineering approach[M]. Cambridge: Cambridge University Press, 2015: 60-630.

[29] HU H. Study of the trac airfoil table computational system: NASA-CR-1999-209323[R]. Hampton: Hampton NASA Langley Research Center, 1999: 5-15.

[30] GUDMUNDSSON S. General aviation aircraft design[M]. 2nd ed. Butterworth-Heinemann: Butterworth-Heinemann Press, 2014: 387-388.

[31] MCCORMICK B W. Aerodynamics, aeronautics, and flight mechanics[M]. 2nd ed. New York: John Wiley & Sons, Inc., 1995: 90-100.

[32] 刘沛清. 空气螺旋桨理论及其应用[M]. 北京: 北京航空航天大学出版社, 2006: 75-82.
LIU P Q. Air propeller theory and its application[M]. Beijing: Beihang University Press, 2006: 75-82 (in Chinese).

[33] LANDGREBE A J, BELLINGER E D. An investigation of the quantitative applicability of model helicopter rotor wake patterns obtained from a water tunnel: AD-739946[R]. Hampton: Hampton United Aircraft Corporation Research Laboratories, 1971: 25-36.

[34] ELLIOTT J W, ALTHOFF S L, SAILEY R H. Inflow measurements made with a laser velocimeter on a helicopter model in forward flight. Volume 1. Rectangular planform blades at an advance ratio of 0.15: NASA-TM-100541[R]. Hampton: NASA Langley Research Center, 1988: 56-64.

[35] SIVELLS J C. Experimental and calculated characteristics of three wings of NACA 64-210 and 65-210 airfoil sections with and without 2 degree washout: NACA-TN-1422[R]. Hampton: NACA Langley Memorial Aeronautical Laboratory, 1947: 68-79.

[36] DA ROCHA SANTOS P D. Aero-structural wing analysis using a vortex, lattice method coupled with an equivalent plate model[D]. Beira: Universidade da Beira Interior, 2011: 19-20.

[37] THEODORSEN T, STICKLE G W, BREVOORT M J. Characteristics of six propellers including the high-speed range: NACA-TR-594[R]. Hampton: Hampton NACA Langley Memorial Aeronautical Laboratory, 1937: 25-36.

[38] SHI Y J, XU G H, WEI P. Rotor wake and flow analysis using a coupled eulerian-lagrangian method[J]. Engineering Applications of Computational Fluid Mechanics, 2016, 10(1): 384-402.

Influence of rotor/wing aerodynamic interference on performance of compound helicopters

YANG Kelong，HAN Dong*

(National Key Laboratory of Science and Technology on Rotorcraft Aeromechanics，School of Aeronautics，Nanjing University of Aeronautics and Astronautics，Nanjing 210016，China)

Abstract: To investigate the influence of rotor/wing aerodynamic interference on the high-speed flight performance of a dual-propeller propelled compound helicopter, a flight performance prediction model of compound helicopters is established. Taking the AS365N ‘Dolphin’ helicopter equipped with a wing and propeller as an example, the influence and mechanism of rotor/wing aerodynamic interference on the flight performance are investigated at a high speed of 400 km/h. The change rule of the helicopter power is investigated when the aileron control is utilized to trim the wing rolling moment caused by the aerodynamic interference. The studies indicate that the aerodynamic interference increases the induced velocity of the wing, resulting in a decrease in the lift coefficient and an increase in the drag coefficient of the wing. Due to the asymmetrical vortex of main rotor, the induced velocity, lift and drag coefficients of the wing under the advancing side of the rotor change more significantly than those on the retreating side. Aerodynamic interference decreases the wing lift share from 80.00% to 71.59%, and increases the rotor lift share from 20.04% to 28.48%. The power of the main rotor, propeller and helicopter is increased by 16.60%, 1.86% and 3.76% respectively. Aerodynamic interference increases the wing rolling moment and decreases the main rotor rolling moment, which is conducive to the aircraft trim, but is inconducive for the power reduction of the aircraft. When the aileron is utilized to balance the wing rolling moment caused by the aerodynamic interference, the lateral cyclic pitch and drag of main rotor are reduced, and hence the aircraft power is decreased.

Keywords: compound helicopter; aerodynamic interference; flight performance; power; rotor; wing

Received：2021-09-18；**Accepted**：2022-01-07；**Published Online**：2022-01-26 11：31

URL：kns.cnki.net/kcms/detail/11.2625.V.20220124.1759.008.html

Foundation items：National Natural Science Foundation of China (11972181); Foundation of Rotorcraft Aeromechanics Laboratory (2005RAL20200104)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0568

基于度量学习的机场交通态势弱监督评估

杜婧涵[1,2]，胡明华[1,2]，张魏宁[1,2]，尹嘉男[1,2,*]

(1. 南京航空航天大学 民航学院，南京 211106；2. 国家空管飞行流量管理技术重点实验室，南京 211106)

摘 要：为准确感知机场场面运行环境，提出基于度量学习的交通态势弱监督评估方法。根据机场场面航空器的时空分布类型，从交通流量、起降队列、资源需求等视角构建交通态势指标体系；借鉴度量学习范式，利用预先定义的相似集和不相似集自动学习态势样本之间的距离度量；在此基础上，采用K均值算法实现弱监督条件下交通态势的等级划分。以上海浦东国际机场实际运行数据为例，分析并验证所提方法的有效性。实验结果表明：起始时刻离场瞬时流量、离场累计流量、离场跑道队列长度及进场累计流量的距离系数大于0.5，对场面态势影响较大；与基于欧式距离的K均值算法相比，度量学习将最优轮廓系数提升了33.3%，得到符合预期语义的聚类结果；此外，机场的平均滑行时间越长、跑道配置越复杂，其场面交通态势等级越高。

关 键 词：航空运输；机场拥堵；态势评估；度量学习；聚类分析

中图分类号：V351

文献标志码：A **文章编号**：1001-5965(2023)07-1772-07

随着航空运输业的蓬勃发展和机场新改扩建工作的逐步推进，大型繁忙机场的航班起降需求不断增长，这使得空中交通管理面临巨大压力。机场作为航空运输网络的关键节点和航班起降活动的直接场所，其场面交通态势的好坏直接影响整体运行效率[1-3]。同时，航路航线结构、扇区结构、空域灵活使用策略等不断优化，使得交通拥堵正由空域转移到机场[4-5]，进一步加大了机场场面调度和拥堵控制的难度。因此，为科学把握场面交通运行情况，准确识别交通态势关键影响因素，有力支撑场面活动的精细化管理，亟需对交通流特征进行建模，并对交通态势进行综合评估。

针对机场交通态势评估问题，李善梅等[6]从机场容量和需求方面构建拥堵态势指标，并利用神经网络分类方法对饱和度表示的拥堵等级进行预测。谢燕雯[7]应用模糊层次分析法确定以起降效率、机场资源利用率为代表地运行态势影响因素的权重，在此基础上，采用问卷调查方式将场面交通态势划分为3个等级，并用马尔可夫模型对其预测。由于缺少大量的客观运行数据支撑，此类主观层面的等级划分方法说服力不强。此外，由于飞机滑行时间能在一定程度上反映场面拥堵态势情况，Balakrishna等[8-9]基于离场交通流在跑道和滑行道上的统计量，构建强化学习模型对机场平均滑出时间进行预测。为了感知机场场面滑行态势，Yin等[10]提出从宏观视角构建指标，并利用主成分分析方法融合指标，实验验证了场面滑行时间与融合后的指标具有一致的波动性。可以看出，现有研究大多是对预先划设的态势等级或滑行时间进行直接预测，对交通态势的客观划分与综合评估方面存在不足。此外，在态势指标构建方面，大多单独针对离场航班，而忽略了进场航班和离场航班之间的相互耦合作用[11]。

为解决以上问题，本文提出基于度量学习的机

收稿日期：2021-09-26；**录用日期**：2021-12-12；**网络出版时间**：2021-12-22 15:07
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20211222.0902.001.html
基金项目：国家自然科学基金(52002178, 71731001)；国家留学基金(202106830100, 202106830077)；江苏省自然科学基金(BK20190416)
*通信作者. E-mail：j.yin@nuaa.edu.cn

引用格式：杜婧涵，胡明华，张魏宁，等. 基于度量学习的机场交通态势弱监督评估[J]. 北京航空航天大学学报，2023，49(7)：1772-1778. DU J H，HU M H，ZHANG W N，et al. Weakly supervised evaluation of airport traffic situation based on metric learning[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1772-1778 (in Chinese).

场交通态势弱监督评估方法。首先, 综合考虑离场航班流和进场航班流, 构建机场场面交通态势指标体系; 然后, 根据预先定义的约束信息, 学习态势指标的距离度量; 最后, 在此距离度量方式下, 利用 K 均值算法进行弱监督聚类, 实现对机场场面交通态势等级的客观划分与综合评估。选取上海浦东国际机场历史运行数据验证了本文方法的合理性和有效性。

1　场面交通态势指标体系构建

为进行机场场面交通态势评估, 首先需根据场面交通流特征进行建模, 构建用于描述机场场面交通态势的指标体系。图 1 为航班在机场场面滑行过程中的时空图。可以看出, 给定参考时间片, 离场航班共包含 4 种基本类型, 即在起始时刻(前/后)从停机位推出, 在终止时刻(前/后)从跑道头起飞。相应地, 进场航班也包含 4 种类型。在后续的实验分析中, 为体现各指标的区分度, 参考时间片设置为 15 min。

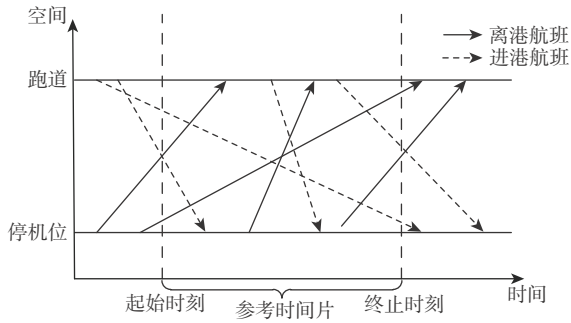


图 1　航空器滑行时空图

Fig. 1　Time-space diagram of aircraft taxiing

在此基础上, 充分考虑本节所有航班类型和进离场的相互作用, 从实际运行角度归纳并构建能够反映场面交通态势的指标体系, 如表 1 所示, 共包括 2 类 10 个指标。这些指标统计了给定参考时间片下不同运行状态的航空器架次, 涵盖了机场场面交通流量、起降队列、资源需求等方面, 能够在一定程度感知机场整体的拥堵环境。

表 1　交通态势指标体系

Table 1　Traffic situation index system

指标类型	指标名称	解释说明
离场类指标	起始时刻离场瞬时流量d_1	参考时间片起始时刻处于离场过程的航空器架次
	终止时刻离场瞬时流量d_2	参考时间片终止时刻处于离场过程的航空器架次
	离场累计流量d_3	参考时间片内所有离场航班架次
	离场跑道队列长度d_4	参考时间片内从跑道起飞的航班架次
	离场资源需求指数d_5	参考时间片内从停机位推出的航班架次
进场类指标	起始时刻进场瞬时流量a_1	参考时间片起始时刻处于进场过程的航空器架次
	终止时刻进场瞬时流量a_2	参考时间片终止时刻处于进场过程的航空器架次
	进场累计流量a_3	参考时间片内所有进场航班架次
	进场跑道队列长度a_4	参考时间片内在跑道降落的航班架次
	进场资源需求指数a_5	参考时间片内推入到停机位的航班架次

2　机场交通态势弱监督评估

机场交通态势评估流程如图 2所示, 首先利用度量学习从构建的指标集和预定义的相关约束信息中学习用于描述不同时间片交通态势相似性的距离矩阵, 然后结合距离矩阵和 K 均值算法进行有偏好的聚类分析, 最终根据聚类结果对态势等级进行划分与评估。

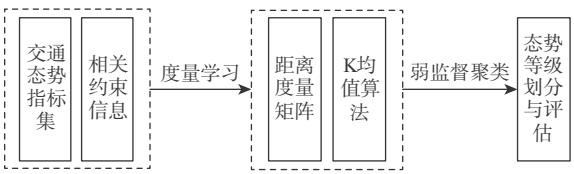


图 2　机场交通态势评估流程

Fig. 2　Flow chart of airport traffic situation assessment

2.1　度量学习

度量学习[12-14] 旨在以机器学习的方式从(弱)监督数据中自动构建特定任务的距离度量。为了使构建的指标可以更好地评估机场交通态势情况, 本文借助度量学习思想学习相应的距离矩阵, 以此来体现不同态势指标之间的差异性。具体地, 利用马氏距离来衡量 2 个态势样本的相似性:

$$\mathrm{dist}_{\boldsymbol{M}}(\boldsymbol{x}_i,\boldsymbol{x}_j)=\sqrt{(\boldsymbol{x}_i-\boldsymbol{x}_j)^{\mathrm{T}}\boldsymbol{M}(\boldsymbol{x}_i-\boldsymbol{x}_j)}=\left\|\boldsymbol{x}_i-\boldsymbol{x}_j\right\|_{\boldsymbol{M}} \quad (1)$$

式中: $\boldsymbol{x}_i\in\mathbf{R}^{10}$表示第$i$个时间片由 10 个态势指标构成的特征向量; $\boldsymbol{M}$为距离度量矩阵。一般情况下, 其是一个具有对角和非对角距离系数的全矩阵, 对角线系数是各个态势指标的权值, 非对角线系数是任意 2 个态势指标相互作用的权值, 为了保持距离度量的非负性和对称性, $\boldsymbol{M}$为半正定对称矩阵。

度量学习的关键是根据具体应用场景定义约束信息，并在约束信息的指导下学习距离矩阵。由于本文的目标是对场面交通态势等级进行划分与评估，因此，定义了2个时间片交通态势的相似集S和不相似集D。给定时间片i和j，若2个时间片的跑道配置相同且平均滑行时间之差小于s_{tt}，则$(\boldsymbol{x}_i,\boldsymbol{x}_j)\in S$；若2个时间片的跑道配置不同且平均滑行时间之差大于d_{tt}，则$(\boldsymbol{x}_i,\boldsymbol{x}_j)\in D$。$s_{\mathrm{tt}}$和$d_{\mathrm{tt}}$分别为滑行时间相似性阈值和不相似性阈值，参考欧控机场协同决策(airport-collaborative decision making, A-CDM)执行手册对可变滑行时间的短期要求[15]及机场实际管制经验，分别设置为2 min和5 min，对于不同机场运行环境可进行灵活调整。

基于本节约束信息，可以通过求解优化目标来得到距离矩阵$\boldsymbol{M}$，即

$$\begin{cases}\min\limits_{\boldsymbol{M}}\sum\limits_{(\boldsymbol{x}_i,\boldsymbol{x}_j)\in S}\|\boldsymbol{x}_i-\boldsymbol{x}_j\|_{\boldsymbol{M}}^2\\ \text{s.t.}\sum\limits_{(\boldsymbol{x}_i,\boldsymbol{x}_j)\in D}\|\boldsymbol{x}_i-\boldsymbol{x}_j\|_{\boldsymbol{M}}\geqslant 1\\ \boldsymbol{M}\geqslant\boldsymbol{0}\end{cases}\tag{2}$$

式中：$\boldsymbol{M}\geqslant\boldsymbol{0}$为半正定矩阵的约束。式(2)在不相似时间片之间距离不小于1的前提下，尽可能拉近相似时间片之间的距离。一方面，若$\boldsymbol{M}$为一个低秩矩阵，可以达到对原有态势指标降维的目的。另一方面，度量学习以相似集和不相似集作为弱监督信息，自动学习偏向预期语义的距离矩阵，有助于更好地实现后续态势等级划分与评估。

2.2 交通态势弱监督评估

为了识别机场场面潜在运行模式，对交通态势等级进行精细划分与综合评估，利用K均值算法对不同时间片态势样本进行聚类分析。作为经典的聚类算法，K均值通过最小化簇内的平方距离之和来得到簇的最优划分，其优化目标为

$$\min_{\{C_1,C_2,\cdots,C_k\}}\sum_{i=1}^{k}\sum_{\boldsymbol{x}\in C_i}\|\boldsymbol{x}-\boldsymbol{\mu}_i\|_2^2\tag{3}$$

式中：k为簇的个数；$\boldsymbol{\mu}_i=\dfrac{1}{|C_i|}\sum\limits_{\boldsymbol{x}\in C_i}\boldsymbol{x}$为簇$C_i$的聚类中心；$|C_i|$为第$i$簇的样本个数。通常情况下，使用欧氏距离来度量样本之间的相似性。然而，这种距离度量方式隐含地假设样本不同维度的重要性相同，在实际应用中具有一定的局限性。因此，本文使用度量学习得到的距离矩阵作为相似性度量方法，在此基础上改进传统K均值算法实现弱监督条件下交通态势等级的划分与评估，基本步骤如下：

步骤1 以15 min为参考时间片，构建场面交通态势的样本集$\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n\}$，其中$\boldsymbol{x}_i\in\mathbf{R}^{10}$，根据预定义的约束信息得到相似集$S$和不相似集$D$。

步骤2 根据式(2)，求解距离矩阵$\boldsymbol{M}$。

步骤3 设定聚类的类别数量k，并随机选择k个样本作为聚类中心$\{\boldsymbol{\mu}_1,\boldsymbol{\mu}_2,\cdots,\boldsymbol{\mu}_k\}$。

步骤4 基于距离矩阵$\boldsymbol{M}$，计算每个态势样本与各聚类中心的距离$d_{ji}=\|\boldsymbol{x}_j-\boldsymbol{\mu}_i\|_{\boldsymbol{M}}$，并将态势样本划分到距离最近的簇中。

步骤5 判断各聚类中心是否收敛，若收敛则进入步骤7，否则进入步骤6。

步骤6 根据每个簇中的样本，计算并更新聚类中心$\boldsymbol{\mu}_i=\dfrac{1}{|C_i|}\sum\limits_{\boldsymbol{x}\in C_i}\boldsymbol{x}$，并进入步骤4。

步骤7 输出态势样本集的划分结果，对场面交通态势进行评估分析。

3 实例分析

3.1 实验数据

本文选取上海浦东国际机场作为研究案例，其机场场面布局如图3所示，共包含16R/34L、17L/35R、16L/34R、17R/35L这4条跑道，其中16R/34L和17L/35R是离场航班的跑道配置，而16L/34R和17R/35L是进场航班的跑道配置。R和L是用来区分同一航向跑道，其中R为右侧跑道，L为左侧跑道。实验选取2019年2月实际运行数据，涉及离场航班19 159架次，进场航班19 225架次。数据项主要包括进(离)场航班降落(起飞)时间、停机位推入(推出)时间、机场场面布局及停机位和跑道配置信息。基于该运行数据计算表1中各指标的具体数值，并在表2给出指标的基本统计信息。此外，为保证各指标尺度一致，使用最小最大标准化方法将各指标映射到0～1之间。

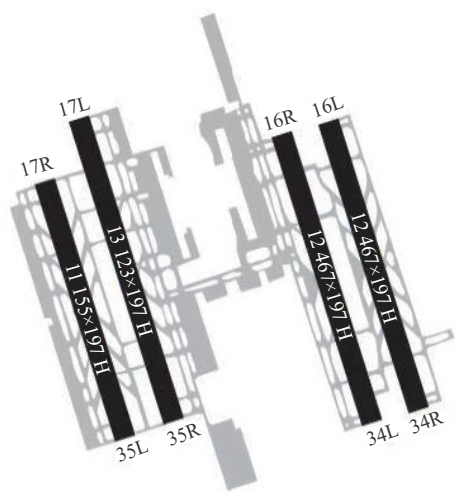


图3 上海浦东国际机场场面布局

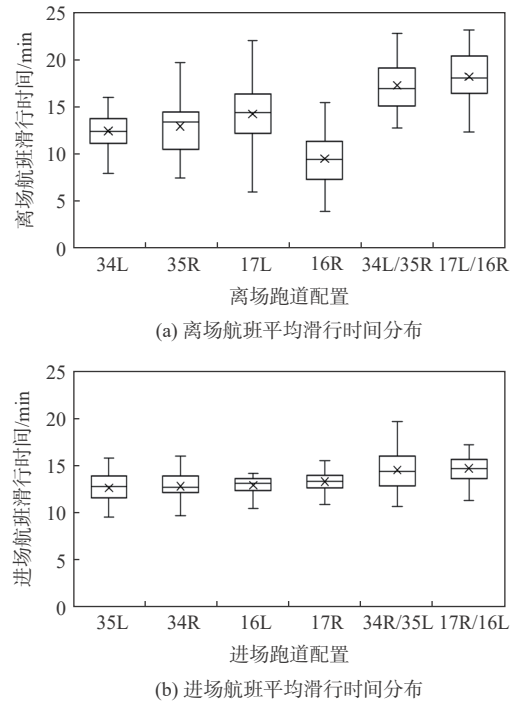
Fig. 3 Scene layout of Shanghai Pudong International Airport

表 2　指标的基本统计信息

Table 2　Basic statistics of indices

指标	统计量/架次			
	最小值	最大值	均值	中位数
d_1	0	25	8.99	9
d_2	0	23	8.92	9
d_3	1	38	16.99	18
d_4	0	17	8.07	9
d_5	0	20	8.01	8
a_1	0	18	7.53	8
a_2	0	17	7.49	8
a_3	0	31	15.67	16
a_4	0	17	8.15	8
a_5	0	18	8.18	8

一般来说, 场面平均滑行时间可以在一定程度上反映机场整体的交通态势。图 4 为离场和进场航班在不同跑道配置下以 15 min 为时间片的平均滑行时间的分布情况。可以看出, 由于机场跑道、滑行道等物理布局的原因, 不同配置的滑行时间具有一定差异。此外, 离场航班比进场航班的波动性更大, 具有更加复杂、动态的运行环境。相较单跑道运行模式而言, 多跑道运行模式的滑行时间整体偏大, 这表明在多跑道运行模式下, 机场场面更易产生拥堵现象, 导致滑行延误。



(a) 离场航班平均滑行时间分布

(b) 进场航班平均滑行时间分布

图 4　进离场航班不同跑道配置下平均滑行时间分布

Fig. 4　Distribution of average taxi time under different runway configurations of departure and arrival flights

3.2　度量学习效果分析

表 3 为度量学习算法学习到的距离矩阵的估计结果。该矩阵的对角线代表各个交通态势指标的距离系数, 非对角线代表指标间交互项的距离系数。通过观察对角线元素, 可以看出在场面交通态势指标体系中, 离场累计流量 d_3、离场跑道队列长度 d_4 及进场累计流量 a_3 的系数较大, 表明这些指标在度量态势样本之间相似性时权重更大。此外, 通过观察非对角线元素, 可以发现离场类指标和进场类指标间的交互作用较小, 指标间的交互作用更多地发生在各类指标内部。

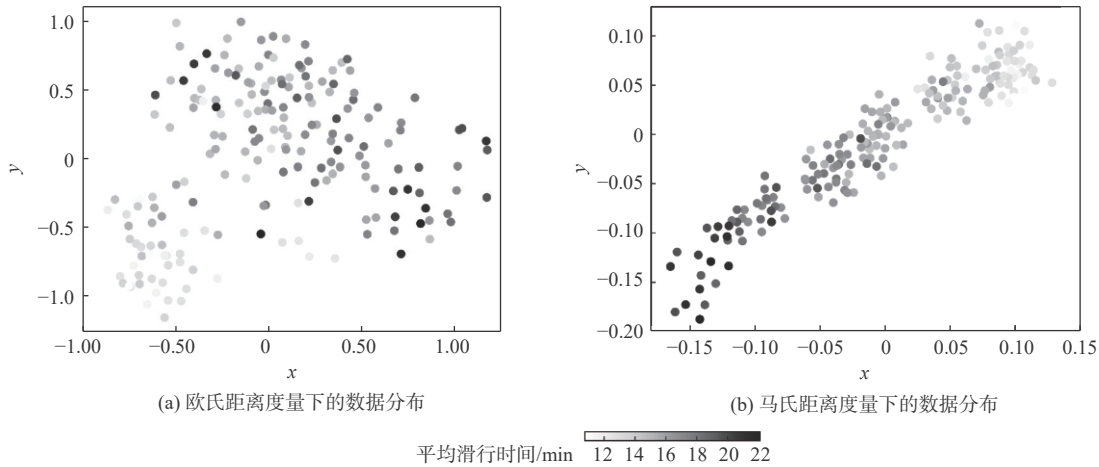
表 3　度量学习算法学习到的距离矩阵的估计结果

Table 3　Estimation results of distance matrix learned by metric learning algorithm

指标	d_1	d_2	d_3	d_4	d_5	a_1	a_2	a_3	a_4	a_5
d_1	0.50	−0.09	−0.18	−0.26	0.21	0.07	0.07	0.01	−0.07	−0.06
d_2	−0.09	0.25	−0.22	0.20	−0.29	0.06	0.06	0.06	−0.06	−0.05
d_3	−0.18	−0.22	0.77	−0.22	−0.21	0.02	0.02	0.04	−0.02	−0.02
d_4	−0.26	0.20	−0.22	0.57	−0.08	−0.04	−0.04	−0.01	0.03	0.03
d_5	0.21	−0.29	−0.21	−0.08	0.34	−0.05	−0.05	−0.02	0.05	0.04
a_1	0.07	0.06	0.02	−0.04	−0.05	0.36	−0.16	−0.22	0.28	−0.23
a_2	0.07	0.06	0.02	−0.04	−0.05	−0.16	0.32	−0.21	−0.21	0.29
a_3	0.01	0.06	0.04	−0.01	−0.02	−0.22	−0.21	0.76	−0.21	−0.22
a_4	−0.07	−0.06	−0.02	0.03	0.05	0.28	−0.21	−0.21	0.32	−0.17
a_5	−0.06	−0.05	−0.02	0.03	0.04	−0.23	0.29	−0.22	−0.17	0.35

进一步地, 利用多维缩放 (multiple dimensional scaling, MDS) 技术[12], 通过保持样本对在原始空间和二维空间距离的一致性, 分别以欧氏距离和学习的马氏距离为相似性度量方式, 可视化态势样本的分布情况。为了便于展示, 随机选择 2 天的数据进行说明。图 5 为度量学习前后数据分布, 每个点代表一个时间片的态势样本, 点的颜色是该时间片的平均滑行时间, 颜色越深代表滑行时间越长。可以看出, 以欧氏距离作为距离度量时, 数据分布较为分散, 没有自然的聚类特性, 且样本对的距离与滑

行时间关联不明显；而在利用度量学习得到的马氏距离作为距离度量时，整个数据分布集中，且平均滑行时间相近的样本距离更近。这直观地反映了度量学习的有效性，在预定义的约束信息下，拉近了相似态势样本的距离，拉远了相异态势样本的距离，从而更有利于后续态势等级划分与评估。



(a) 欧氏距离度量下的数据分布　　(b) 马氏距离度量下的数据分布

图5　度量学习前后数据分布

Fig. 5　Data distribution before and after metric learning

3.3　聚类结果分析

在度量学习的基础上，利用K均值算法对场面交通态势样本进行聚类分析。为了确定态势等级的划分数量，使用轮廓系数[16]来评估聚类的效果，轮廓系数越大表示聚类效果越好。图6为不同聚类数下的轮廓系数。可以发现，结合度量学习聚类的轮廓系数远大于直接聚类的轮廓系数，最优轮廓系数提升33.3%(从0.45提高到0.60)，能够有效增强聚类效果，进一步验证了度量学习的有效性。当聚类数为3时，轮廓系数达到最大值，故将场面态势等级划分为3级。

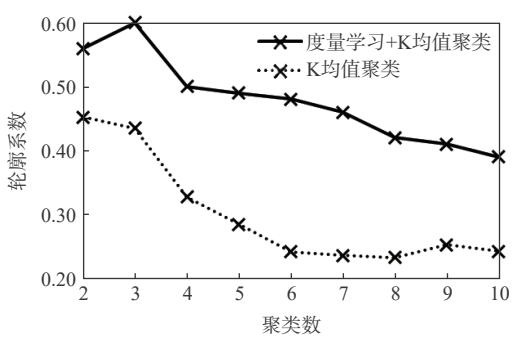


图6　不同聚类数下的轮廓系数

Fig. 6　Contour coefficients under different cluster number

此外，选取度量矩阵中距离系数大于0.5的指标作为关键指标，即起始时刻离场瞬时流量d_1、离场累计流量d_3、离场跑道队列长度d_4及进场累计流量a_3，并给出各指标在不同态势等级的聚类中心。从态势等级1到态势等级3，各指标均有明显增加，也进一步说明场面交通繁忙、拥堵的态势由低等级向高等级不断变化，如表4所示。

图7为不同态势等级下平均滑行时间的概率密度分布。可以看出，尽管各等级的概率密度分布存在一定的重叠，但均具有单一的峰值。此外，随着等级的增加，峰值对应的滑行时间增大，分布的不确定性也增大。这表明场面越繁忙，越易引发滑行延误；通过实施有效的交通流管控策略，能够在一定程度上缓解拥堵态势的恶化，降低平均滑行时间。

表4　关键指标的聚类中心

Table 4　Cluster center of key indices

指标	聚类中心		
	态势等级1	态势等级2	态势等级3
d_1	2.83	10.42	16.06
d_3	3.37	10.26	16.48
d_4	8.01	19.77	27.31
a_3	10.65	15.49	20.14

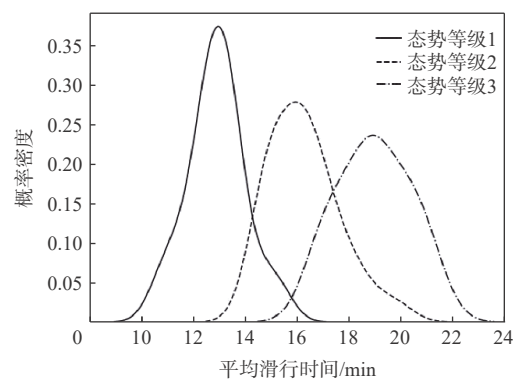


图7　不同态势等级下平均滑行时间的概率密度分布

Fig. 7　Probability density distribution of average taxi time under different situation levels

表 5 统计了不同态势等级下占比大于 5% 的跑道配置使用情况, 其中, 符号 "|" 用于区分进离场跑道配置, "|" 之前为进场跑道配置, 之后为离场跑道配置。可以看出, 态势等级 1 的跑道配置类型较多, 绝大多数是一起一落的跑道运行方式, 由于跑道物理布局原因, 该模式进离场的滑行过程相对独立, 滑行延误较小, 通常在交通流量较小的场景下使用; 态势等级 2 和态势等级 3 主要为双起双落的混合运行模式, 即 2 个离场跑道独立平行离场, 2 个进场跑道独立平行进场。根据具体的风向信息, 灵活使用 2 种跑道配置(34R, 35L|34L, 35R 和 17R, 16L|17L, 16R), 其中态势等级 2 前者占比较多, 态势等级 3 后者占比较多。在这 2 种态势下, 进离场的滑行过程耦合性更强, 场面冲突和拥堵现象频发, 整体滑行延误较大。

表 5　不同态势等级下占比大于 5% 的跑道配置情况

Table 5　Runway configuration accounting for more than 5% under different situation levels

态势等级	跑道配置(进场|离场)	占比/%
1	34R|35R	37.5
	35L|34L	26.2
	17R|16R	10.7
	16L|17L	8.0
	34R,35L|35R	5.9
2	34R,35L|34L,35R	52.1
	17R,16L|17L,16R	41.6
3	17R,16L|17L,16R	48.5
	34R,35L|34L,35R	40.7

图 8 为不同经验参数下最优轮廓系数。可以看出, 同一阈值对, 参考时间片为 15 min 时获得最优轮廓系数; 同一参考时间片, (2,5) 和 (3,6) 的阈值对获得最优轮廓系数。此外, 在所设定经验参数配置下, 绝大多数的最优轮廓系数均大于 0.5, 优于仅使用 K 均值聚类的最优轮廓系数 0.45, 也从侧面说明本方法具有一定的鲁棒性。

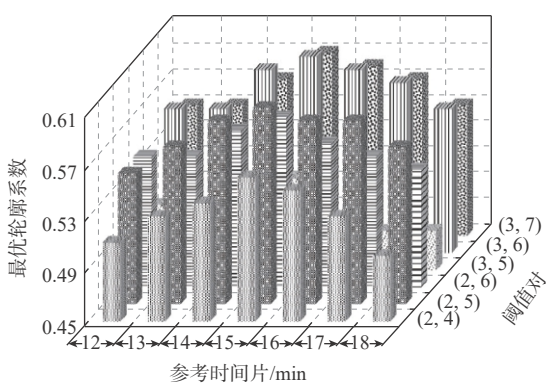


图 8　不同经验参数下最优轮廓系数

Fig. 8　Optimal contour coefficient under different empirical parameters

4　结　论

1) 度量学习模型得到的距离矩阵可以在一定程度反应各态势指标的重要性程度, 其中影响场面交通态势最重要的 4 个因素为起始时刻离场瞬时流量、离场累计流量、离场跑道队列长度及进场累计流量。

2) 结合度量学习的 K 均值算法能够有效增强聚类效果, 其最优轮廓系数相比传统 K 均值算法提升 33.3%, 且具有一定的鲁棒性。

3) 通过对聚类结果进行分析, 发现平均滑行时间和跑道配置的变化均能在一定程度上反映态势等级的变化。

在未来的工作中, 将继续研究不同态势等级下各影响因素的基本特征, 重点识别态势等级突变的关键致因及其临界特性。

参考文献 (References)

[1] SIMAIAKIS I, BALAKRISHNAN H. A queuing model of the airport departure process[J]. Transportation Science, 2016, 50(1): 94-109.

[2] MARIN A G. Airport management: Taxi planning[J]. Annals of Operations Research, 2006, 143(1): 191-202.

[3] 郑丽君, 胡荣, 张军峰, 等. 高峰时段下离港航空器绿色滑行策略设计与评价[J]. 北京航空航天大学学报, 2019, 45(11): 2320-2326.
ZHENG L J, HU R, ZHANG J F, et al. Design and evaluation of green taxiing strategy for departure aircraft during peak hours[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(11): 2320-2326(in Chinese).

[4] RODRÍGUEZ-SANZ Á, COMENDADOR F G, VALDÉS R A, et al. Characterization and prediction of the airport operational saturation[J]. Journal of Air Transport Management, 2018, 69: 147-172.

[5] SMELTINK J W, SOOMER M J, DE WAAL P R, et al. Optimisation of airport taxi planning: NLRTP-2003-475[R]. Amsterdam : National Aerospace Laboratory NLR, 2003.

[6] 李善梅, 徐肖豪, 孟令航. 基于聚类神经网络的机场拥挤等级预测[J]. 计算机工程与应用, 2013, 49(17): 254-257.
LI S M, XU X H, MENG L H. Forecasting of airport congestion level based on cluster and neural network algorithms[J]. Computer Engineering and Applications, 2013, 49(17): 254-257(in Chinese).

[7] 谢燕雯. 基于隐马尔可夫过程的机场运行态势预测方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2018: 21-32.
XIE Y W. Research on airport operation situation prediction method based on hidden Markov process[D]. Harbin: Harbin Institute of Technology, 2018: 21-32 (in Chinese).

[8] BALAKRISHNA P, GANESAN R, SHERRY L. Airport taxi-out prediction using approximate dynamic programming: Intelligence-based paradigm[J]. Transportation Research Record, 2008, 2052(1): 54-61.

[9] BALAKRISHNA P, GANESAN R, SHERRY L. Accuracy of reinforcement learning algorithms for predicting aircraft taxi-out times:

A case-study of Tampa Bay departures[J]. Transportation Research Part C:Emerging Technologies, 2010, 18(6): 950-962.

[10] YIN J N, HU M H, MA Y Y, et al. Airport taxi situation awareness with a macroscopic distribution network analysis[J]. Networks and Spatial Economics, 2019, 19(3): 669-695.

[11] CLEWLOW R, SIMAIAKIS I, BALAKRISHNAN H. Impact of arrivals on departure taxi operations at airports[C]//AIAA Guidance, Navigation, and Control Conference. Reston: AIAA, 2010: 7698.

[12] XING E P, NG A Y, JORDAN M I, et al. Distance metric learning with application to clustering with side-information[C]//Proceedings of 15th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2002, 521–528.

[13] DU J H, HU M H, ZHANG W N. Decision support for aircraft taxi time based on deep metric learning[C]//2020 IEEE 23rd International Conference on Intelligent Transportation Systems. Piscataway: IEEE Press, 2020: 1-7.

[14] LIU Y, SEELHORST M, POZDNUKHOV A, et al. Assessing terminal weather forecast similarity for strategic air traffic management [C]//6th International Conference on Research in Air Transportation. Istanbul: Istanbul Technical University, 2014: 26-30.

[15] BOOTH D, FLYNN B, SULTANA J. Airport CDM implementation[EB/OL]. (2021-08-09)[2017-03-31]. https://www.eurocontrol.int/sites/default/files/publication/files/airport-cdm-manual-2017.PDF.

[16] ROUSSEEUW P J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis[J]. Journal of Computational and Applied Mathematics, 1987, 20: 53-65.

Weakly supervised evaluation of airport traffic situation based on metric learning

DU Jinghan[1, 2], HU Minghua[1, 2], ZHANG Weining[1, 2], YIN Jianan[1, 2, *]

(1. School of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China;
2. National Key Laboratory of Air Traffic Flow Management, Nanjing 211106, China)

Abstract: In order to accurately perceive the operating environment of the airport surface, a weakly supervised evaluation method of traffic situation based on metric learning is proposed. Firstly, according to the space-time distribution types of aircraft on the airport surface, the traffic situation index system is constructed from the perspectives of traffic flow, take-off and landing queues, and resource demand; secondly, learning from the metric learning paradigm, the distance measure between situation samples is automatically learned by using pre-defined similar sets and dissimilar sets; on this basis, the K-means algorithm is used to achieve traffic situation classification under weak supervision. Taking the actual operation data of Shanghai Pudong International Airport as an example, the effectiveness of the proposed method is analyzed and verified. The experimental results show that:when the distance coefficients of the departure instantaneous flow at the beginning, the departure cumulative flow, the length of the departure runway queue, and the arrival cumulative flow are greater than 0.5, it has a greater impact on the surface situation; metric learning improves the optimal contour coefficient by 33.3% when compared to the K-means algorithm based on Euclidean distance, and obtains clustering results that meet the e-criterion; in addition, the longer the average taxi time of the airport and the more complex the runway configuration, the higher the level of the traffic situation on the surface.

Keywords: air transportation; airport congestion; situation assessment; metric learning; clustering analysis

Received: 2021-09-26; **Accepted**: 2021-12-12; **Published Online**: 2021-12-22 15:07

URL: kns.cnki.net/kcms/detail/11.2625.V.20211222.0902.001.html

Foundation items: National Natural Science Foundation of China (52002178, 71731001); China Scholarship Council (202106830100, 202106830077); Natural Science Foundation of Jiangsu Province (BK20190416)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0479

星载 GNSS-R 土壤湿度反演中开放水域的影响

杨文涛[1]，徐天河[1,2,*]，王娜子[2]，高凡[2]，荆丽丽[2]，贺匀峤[2]

(1. 长安大学 地质工程与测绘学院，西安 710054；2. 山东大学 空间科学研究院，威海 264209)

摘 要：通过星载全球导航卫星系统-反射测量（GNSS-R）技术可以获取高时空分辨率的土壤湿度（SM）信息，但开放水域会影响反演结果的精度。针对此问题，提出一种开放水域剔除的改进方法。校正其模拟功率得出反射率，结合掩膜数据精确剔除开放水域。为验证改进方法的可行性，处理了2个开放水域比较密集的试验区内为期一年的飓风全球导航卫星系统（CYGNSS）1级数据。反演结果与主被动土壤湿度（SMAP）3级土壤湿度产品进行对比分析，其均方根误差为0.052 1 cm^3/cm^3，相关性为0.654。与开放水域剔除前的结果相比分别提升6.2%和9%；与CYGNSS官方发布的产品相比分别提升32%和24%。反演结果可以作为SMAP的SM补充，提供高时间分辨率的SM值。

关 键 词：全球导航卫星系统-反射测量；土壤湿度；开放水域；飓风全球导航卫星系统；主被动土壤湿度

中图分类号：P237

文献标志码：A **文章编号**：1001-5965(2023)07-1779-08

土壤湿度(soil moisture, SM)是了解农业、水文过程及气候条件的关键地球物理参数，在陆地水循环中有着不可替代的作用。SM的准确观测对于干旱监测、洪水预报及农业灌溉等领域有着重要的意义[1]。因此，获取具有高时空和空间分辨率的SM数据非常重要。普通遥感技术进行大规模SM观测主要依靠的是光学遥感和微波遥感。利用光学遥感获取的数据具有较高的空间分辨率，但是易受云层植被的影响，而无源微波遥感获取的数据具有较高的时间分辨率，但是空间分辨率较低[2]。

全球导航卫星系统–反射测量(global navigation satellite system reflectometry, GNSS-R)技术是一种通过接收处理地球表面反射的GNSS信号，实现对地球表面物理参数提取的遥感技术[3]。而GNSS信号处于L波段，该波段电磁波具有一定的穿透地表植被、沙土、雪盖能力，对土壤水分较为敏感[4]，信号源数量充足且具有全天时全天候观测能力，在SM高时空分辨率的监测方面极具潜力。

前期，一些研究人员进行了机载和地基实验[5-6]。本文重点介绍基于星载GNSS-R的SM反演情况。在星载观测方面，英国技术演示卫星(United Kingdom technology demonstration satellite, UKTDS-1)于2014年发射[7]，为星载GNSS-R对SM的研究提供了首批数据。2016年，文献[8-9]利用技术演示卫星-1(technology demonstration satellite, TDS-1)数据进行了SM反演的研究，验证了星载SM反演的可行性。由于TDS-1卫星的设计，8天中只有2天的观测数据可用，因此，有限的数据观测量限制了星载GNSS-R的SM反演更为深入的研究[10]。直到美国国家航空航天局(National Aeronautics and Space Administration, NASA)于2016年12月发射了飓风全球导航卫星系统(cyclone global navigation satellite system,

收稿日期：2021-08-23；录用日期：2021-11-05；网络出版时间：2021-11-17 09:00
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20211116.1202.003.html
基金项目：国家自然科学基金(41604003,41704017,41704018,41874032,41931076)；国家重点研发计划(2020YFB0505800)
*通信作者. E-mail：thxu@sdu.edu.cn

引用格式：杨文涛，徐天河，王娜子，等. 星载GNSS-R土壤湿度反演中开放水域的影响[J]. 北京航空航天大学学报，2023，49(7)：1779-1786. YANG W T，XU T H，WANG N Z，et al. Influence of open water in retrieval of soil moisture by spaceborne GNSS-R[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1779-1786 (in Chinese).

CYGNSS)卫星星座，这一问题才得到有效解决。CYGNSS目的是通过估计北纬38°和南纬38°之间的海面风场来改善天气预报。虽然CYGNSS为海洋气象而设计，但是卫星也记录了大量的陆地观测数据，这为SM的反演提供了充足的数据。2018年，Chew和Small[11]发现CYGNSS信噪比的变化和主被动土壤湿度(soil moisture active and passive, SMAP)的SM值具有相关性，并进行了线性回归分析，他们将SMAP的SM平均值和基于CYGNSS信噪比得出的SM变化来估计SM值。Kim和Lakshmi[12]提出相对信噪比来进行SM反演。通过结合相对信噪比(relative signal-to-noise ratio, rSNR)和SMAP的SM值来推算每日的SM值。2019年，Clarizia等[13]提出反射-植被-粗糙度算法，该算法根据CYGNSS的反射率、植被不透明度及粗糙度系数并通过线性回归函数对SMAP的SM值进行了分析，最终得出SM的每日估计。Dong和Jin[14]利用CYGNSS相干反射数据进行了SM的反演。Yang等[15]利用机器学习的方法也进行了SM的反演。目前CYGNSS官网已经公布了SM产品[16]。

文献[8,11]说明，对于TDS-1和CYGNSS数据得出的反射率可以对地表的SM进行感知，其中水体的反射率较高，干旱地区反射率较低。目前，星载GNSS-R的SM反演大多基于反射率并将SMAP的SM作为对比值进行反演[13-16]。而开放水域的存在会降低CYGNSS反射率与SMAP的SM相关性，从而影响SM的反演结果[8,16]。而且开放水域周围常有农田等地物，快速、准确地获取开放水域密集区域的SM对农田保护也非常关键。因此，开放水域的剔除是SM反演中非常重要的步骤，但目前对此研究较少。本文改进了利用掩膜数据精确剔除开放水域的方法，之后选取2个开放水域比较密集的区域进行SM的反演，并将开放水域剔除前后的SM反演结果进行分析，并与CYGNSS官方发布的产品进行比较分析。

1 数据和实验区域

1.1 CYGNSS数据

CYGNSS是NASA的一个地球探险任务，由8颗小卫星组成，以35°的倾角围绕热带轨道运行，接收GPS卫星的直接信号和反射信号。CYGNSS的设计目的是在飓风期间观测海洋表面风速，因此，覆盖范围只有南纬38°到北纬38°，并且接收器和软件被优化用于海洋表面遥感[17]。

CYGNSS每颗卫星携带一个GNSS-R接收机，能同时提供来自4个不同GPS卫星的GNSS-R观测数据，自2017年3月在https://podaacopendap.-jpl.nasa.gov开始提供数据的免费访问和下载服务。由于CYGNSS的数据间隔为1 s，因此，每秒可以同时获得32组测量值。其中，从2019年中期开始，CYGNSS的数据间隔缩短为0.5 s，并且CYGNSS重访周期在几个小时内，因此，CYGNSS能够提供高时空间分辨率的观测数据。本文中使用的是CYGNSS Level 1 v3.0科学数据产品，数据时间跨度为2019年11月到2020年10月。较上一版本(v2.1)数据，该版本利用CYGNSS时延多普勒测量仪测量的直射信号功率和天线增益重新估计了GPS等效全向辐射功率(effective isotropic radiated power, EIRP)[18]。

1.2 掩膜数据

本文将GSWE(global surface water explorer)数据集作为掩膜数据，该数据集是欧洲委员会联合研究中心开发的数据集。该数据集绘制了过去36年中全球范围内水面的位置和时间分布，并提供了这些水面的范围和变化的统计数据。对于每一个数据集，都详细描述了创建方法，并提供目的、描述、波段和符号等信息，以便用户理解每一个数据集，并有效地和恰当地使用[19]。本文使用的是GSWE数据集提供的“seasonality”数据产品。该产品的一个像素代表了一年中开放水域存在的时间长度(0～12个月)，无数据的像素值为255。为了进行精确的开放水域的剔除，本文将大于0并且小于255的像素值标记为开放水域。图1为GSWE的季节性产品(见https://global-surface-water.appspot.com/)，大面积白色为为无水区域，左下灰色为无数据区域，其他(蓝色)为开放水域。

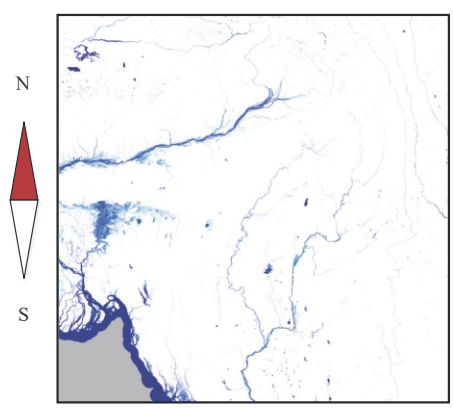


图1 GSWE的季节性产品

Fig. 1 GSWE's seasonal products

1.3 SMAP数据

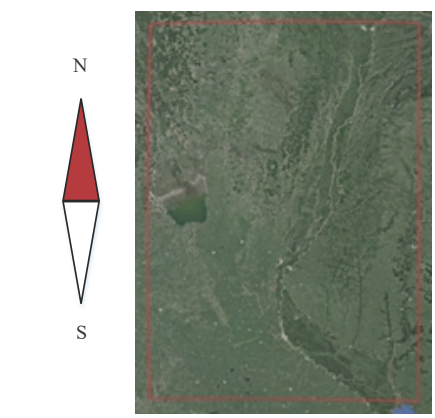
本文使用的SM对比产品来自于SMAP卫星。SMAP卫星任务于2015年1月由NASA发射，SMAP卫星搭载主动雷达和被动辐射计2种传感器，通过

测量地球亮度、温度进行地球物理反演来获取全球 SM 信息, 为用户提供全球范围内每日的 SM 估计值[20]。

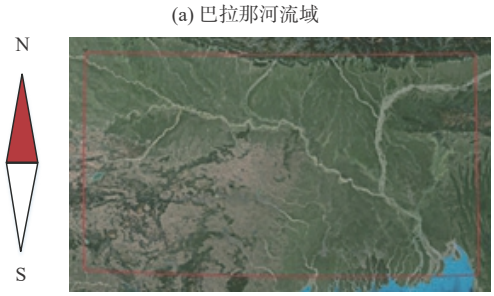
由于主动式雷达在轨运行 2 个月后失效, 本文中使用的是 SMAP L3 辐射计全球每日 36 km EASE-GRID 土壤湿度 V007 产品。该产品的空间分辨率为 36 km×36 km[21]。SMAP 每日数据包括降轨和升轨的 SM 值及相应坐标。由于 SMAP 的重访周期为 2～3 天, 为了之后数据的比较和验证, 本文使用升轨数据作为标准数据。

1.4　实验区域

本文选择的实验区域为阿根廷的巴拉那河流域(34°8′15″S～27°15′34″S, 63°29′7″W～57°53′2″W)和孟加拉国的博多河流域(21°59′17″N～26°54′0″N, 82°31′51″E～92°6′0″E)。这 2 个区域的特点是支流较多, 开放水域分布情况比较复杂, 有洪涝灾害发生, 并且周围分布大量农田。图 2 为 2 个复杂水域所对应光学影像, 详见 https://www.google.com/earth/。



(a) 巴拉那河流域



(b) 博多河流域

图 2　巴拉那河流域和博多河流域光学影像

Fig. 2　Optical images of Paraná river basin and Bodo river basin

2　理论方法

2.1　反射率的获取及质量控制

2.1.1　反射率的计算

CYGNSS 和 GPS 星座的配置功能可以看作是 L 波段的双基地雷达, 接收与散射表面特性有关的信息。其中 SM 也是该双基地雷达获取的陆地信息一部分。因此, 可以通过双基地雷达方程来获得地表反射率[10], 由此进一步获取 SM 信息。

CYGNSS 的主要产品参数是延迟多普勒图(delay Doppler map, DDM)[22]。DDM 由相干分量和非相干分量 2 部分组成[12]。假设 CYGNSS 陆地测量主要是镜面点反射, 即表面相对光滑, 反射信号是相干的, 并且来自由第 1 个菲涅耳区定义的区域, 则相干分量可表示为[23-25]

$$P_{\mathrm{RL}}^{\mathrm{coh}}=\frac{P_{\mathrm{r}}^{\mathrm{t}}G^{\mathrm{t}}}{4\pi(R_{\mathrm{ts}}+R_{\mathrm{sr}})^2}\frac{G^{\mathrm{r}}\lambda^2}{4\pi}\Gamma_{\mathrm{RL}} \tag{1}$$

式中: $P_{\mathrm{r}}^{\mathrm{t}}$为信号的发射功率; G^{t}为发射天线的增益; G^{r}为接收天线的增益; R_{ts}为镜面反射点与 GPS 发射机之间的距离; R_{sr}为镜面反射点与接收机之间的距离; λ为波长; Γ_{RL}为地表反射率, dB。

CYGNSS 在其数据集中提供了许多变量, 其中“raw_counts”是由 DDM 测量仪产生的未校准的功率值。之后, 根据量化校正, 即“L1a”校准可以得到以瓦特为单位的模拟功率; 经过由几何形状和天线参数校正, 即“L1b 校准”, 可以得到双基地雷达横截面(bistatic radar cross section, BRCS)。通过“L1b”校准的 BRCS 的表达式在文献 [26] 中的定义为

$$\sigma=\frac{P_{\mathrm{RX}}(4\pi)^2}{R_{\mathrm{ts}}^2R_{\mathrm{sr}}^2P_{\mathrm{r}}^{\mathrm{t}}\lambda^2G^{\mathrm{t}}G^{\mathrm{r}}} \tag{2}$$

式中: P_{RX}为有效反射率。

对于反射率, 利用 CYGNSS 产品有 2 种等效的方法可进行求取[27]: ①使用$P_{\mathrm{RL}}^{\mathrm{coh}}$并通过求解方程式(1)中的$\Gamma_{\mathrm{RL}}$来计算反射率; ②使$P_{\mathrm{RX}}=P_{\mathrm{RL}}^{\mathrm{coh}}$并将式(1)代入式(2), 然后将$\Gamma_{\mathrm{RL}}$写为$\sigma$的函数, 表达式为

$$\Gamma_{\mathrm{RL}}=\frac{(R_{\mathrm{ts}}+R_{\mathrm{sr}})^2}{4\pi(R_{\mathrm{ts}}R_{\mathrm{sr}})^2}\sigma \tag{3}$$

与文献 [16] 相同, 本文中使用模拟功率的峰值来计算地表反射率。通过式(1)可得

$$\begin{aligned}\Gamma_{\mathrm{RL}}=&10\lg P_{\mathrm{RL}}^{\mathrm{coh}}-10\lg P_{\mathrm{r}}^{\mathrm{t}}-10\lg G^{\mathrm{t}}-10\lg G^{\mathrm{r}}+\\&20\lg(R_{\mathrm{ts}}+R_{\mathrm{sr}})+20\lg(4\pi)-20\lg\lambda\end{aligned} \tag{4}$$

式(4)中的变量及所对应的 CYGNSS 数据集参数如表 1 所示。

表 1　式(4)中变量和所对应的 CYGNSS 数据集参数

Table 1　Variables in Eq. (4) and corresponding CYGNSS dataset parameters

变量	CYGNSS数据集参数
$P_{\mathrm{r}}^{\mathrm{t}}G^{\mathrm{t}}$	gps_eirp
G^{r}	sp_rx_gain
R_{ts}	tx_to_sp_range
R_{sr}	rx_to_sp_range

此外，λ 为 GPS 的波长(0.19 m)；$P_{\mathrm{RL}}^{\mathrm{coh}}$可用“ddm_snr”或模拟功率的峰值来代替[8,10]。

2.1.2 数据质量控制

由 2.1.1 节求得的Γ_{RL}为 CYGNSS 的地表反射率，其为入射角和介电常数ε的函数，而介电常数是和近地表的土壤湿度相关。图 3 模拟了入射角与反射率的关系。可知，相同的介电常数下，入射角在 65°左右，反射率会有明显的变化，故本文将不考虑入射角大于 65°的 CYGNSS 观测数据。另外，将信噪比小于 2 dB、接收机天线增益小于 5 dB 的观测数据也进行了剔除。

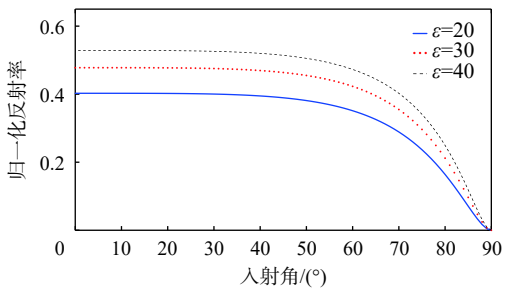


图 3 入射角与反射率的关系

Fig. 3 Relationship between incident angle and reflectivity

2.2 开放水域的剔除

笔者在文献 [28] 提出基于掩膜数据进行对开放水域精确剔除的方法。之后，为提高星载 GNSS-R 在大范围复杂水域反演 RM 的可行性，对程序进行了改进。

在文献 [28] 中将获取的 GSWE 季节性产品转化为二进制矩阵。其中，矩阵的数值范围为 0～255，其中 0 为无水，1～11 为开放水域存在的月份，12 为永久性开放水域，255 为无数据。10°×10°经纬度大小的季节性产品可以转化为 40 000×40 000 大小的二进制矩阵。然后，根据矩阵大小和经纬度信息可以精确计算出像素矩阵每个元素所对应的实际面积大小，其经纬度大小为 0.000 25°×0.000 25°。之后，将矩阵大于 0 且小于 255 每个元素标记为开放水域，求出所对应的面积作为掩膜数据，并用经纬度的形式进行记录保存。

而在本文则不记录每个大于 0 且小于 255 的元素矩阵。因为 GSWE 的季节性产品的空间分辨率为 30 m×30 m，对于 CYGNSS，这种分辨率非常高。故将连续大于 0 且小于 255 每个元素矩阵进行统计，最后，只需记录开始网格的经纬度及结束网格的经纬度，这样大大减少数据的记录量，提高大范围 SM 反演的可行性。

图 4 为在开放水域中的镜面反射点，只需将在实验区域中 CYGNSS 每个镜面点所对应的经、纬度，相对应的 CYGNSS 数据变量为“sp_lat”、“sp_lon”，与掩膜数据进行对比。若镜面点在转化后的掩膜数据中水域所覆盖的区域内，则将该镜面点进行剔除。

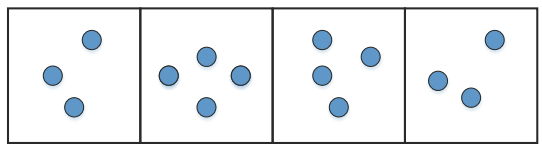


图 4 在开放水域中的镜面反射点

Fig. 4 Specular reflection point in open water

2.3 土壤湿度反演算法

在文献 [8,11] 中已经表明，SM 和反射率是线性相关，且 SM 值越大[29]，这种相关性越高。故反演算法的基础是将 SMAP 作为 SM 的真值，假设 CYGNSS 的反射率和 SMAP 的 SM 值线性相关并建立回归模型。

SMAP 的 SM 值是采用 EASE-Grid 2.0 网格进行记录，所以本文基于网格化进行 SM 反演。对于每个 EASE-Grid 网格，首先计算校正周期内 SMAP 的 SM 值和同天观测的 CYGNSS 反射率的平均值，分别记作$\overline{\mathrm{SM}}_{\mathrm{SMAP}}^{i}$、$\overline{P}_{\mathrm{CYGNSS}}^{i}$，之后计算校正周期内 SMAP 的 SM 值与$\overline{\mathrm{SM}}_{\mathrm{SMAP}}^{i}$之间的差值，记作$\Delta\mathrm{SM}_{\mathrm{SMAP}}^{t,i}$。每个 EASE-Grid 网格的大小为 36 km×36 km，而 CYGNSS 卫星的空间分辨率为 3.5 km×0.5 km，因此，在一个网格中可能会出现多个 CYGNSS 观测值，若观测值小于 3 个，则认为 CYGNSS 不足以对这个网格进行准确的采样，将不会记录此次观测。若网格区域内的地形起伏较大，也会对反射率产生影响，且如图 5 所示，进行开放水域剔除后，开放水域边缘地带由于湿润的原因也会有较高的反射率。对于地形及开放水域边缘地带的反射率，本文采用“3σ”准则进行质量控制。之后将与 SMAP 观测值同一天发生的 CYGNSS 观测值进行平均作为该网格的 CYGNSS 观测值，记作$\Delta P_{\mathrm{CYGNSS}}^{t,i}$。

然后通过线性回归，建立$\Delta P_{\mathrm{CYGNSS}}^{t,i}$与$\Delta\mathrm{SM}_{\mathrm{SMAP}}^{t,i}$的线性回归模型：

$$\Delta\mathrm{SM}_{\mathrm{SMAP}}^{t,i}=a\Delta P_{\mathrm{CYGNSS}}^{t,i}+b \tag{5}$$

式中：a 和 b 分别为线性模型参数。

最后利用线性模型参数 a、$\Delta P_{\mathrm{CYGNSS}}^{t,i}$和 $\overline{\mathrm{SM}}_{\mathrm{SMAP}}^{i}$进行 SM 的反演：

$$\mathrm{SM}_{\mathrm{CYGNSS}}^{t,i}=a\Delta P_{\mathrm{CYGNSS}}^{t,i}+\overline{\mathrm{SM}}_{\mathrm{SMAP}}^{i} \tag{6}$$

式中: i 为 EASE-Grid 网格的索引值; t 为 CYGNSS 观测值的日期。

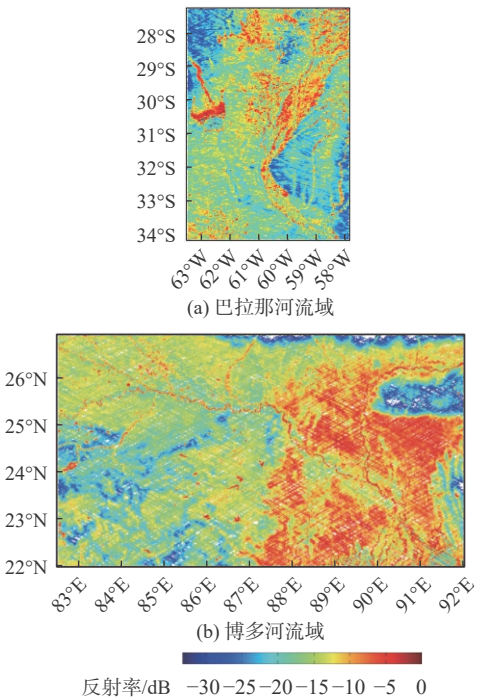


图 5　巴拉那河流域和博多河流域 2020 年 1 月～2020 年 2 月的反射率

Fig. 5　Reflectivity of Paraná river basin and bodo river basin from january to february 2020

3　结果与分析

3.1　开放水域剔除前的反射率

通过 2.1 节获取了 2019 年 11 月～2020 年 10 月之间 2 个实验水域内的 CYGNSS 反射率。图 5 为实验区域 2020 年 1 月～2020 年 2 月的反射率。可知, 实验区域内反射率在不同区域中的变化。其中开放水域中反射率较高, 其他区域的反射率相对较低, 且也与其光学影像图高度吻合。

3.2　开放水域精确剔除后的反射率

选取对应的区域的全球水资源数据集的季节性产品, 通过 2.2 节所述进行开放水域的剔除。即将开放水域中的 CYGNSS 镜面反射点进行剔除。剔除后的反射率结果如图 6 所示。

3.3　土壤湿度反演结果分析

将 SMAP 的 SM 值作为标准值, 通过 2.3 节所述, 将反演后的结果进一步分析。表 2 为开放水域剔除后的 CYGNSS 的 SM 反演结果, 其中, R 为 SM 值的相关性, RMSE为均方根误差(root mean square error)。

同时也将开放水域未剔除的 CYGNSS 反射率也进行了 SM 的反演, 反演结果也作了相同的分析。分析结果如表 3 所示。

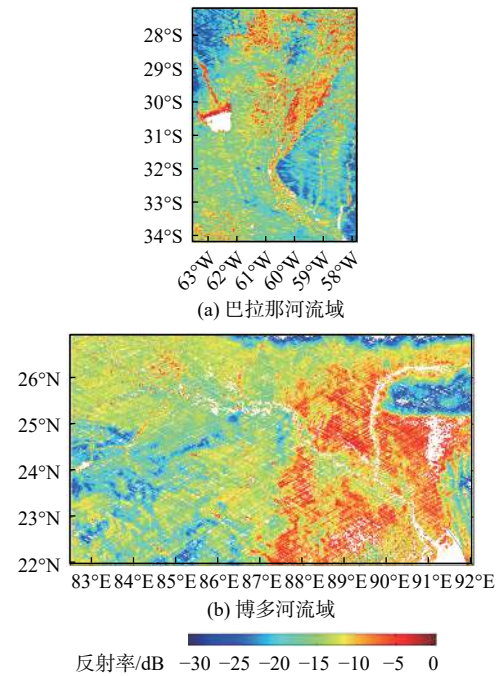


图 6　巴拉那河流域和博多河流域开放水域剔除后的反射率

Fig. 6　Reflectivity after removing open water in Paraná river basin and Bodo river basin

由表 2 和表 3 可以看出对于水域复杂的区域, 开放水域剔除后的反演结果的 RMSE 和 R 分别提升了 6.2% 和 9%, 优于不考虑开放水域的反演结果。

之后将只剔除开放水域的 CYGNSS 反演结果进行空间分析, 如图 7 和图 8 所示。从表 2、图 7 和图 8 可以得出, 基于 CYGNSS 反演的 SM 与 SMAP 有着良好的一致性。然后选取 2 个网格的 SM 反演值进行了时间序列的分析, 并且与 SMAP 的 SM 值

表 2　开放水域剔除后的 CYGNSS 的 SM 反演结果

Table 2　CYGNSS of SM retrieval results after removing open water

区域	R	RMSE/ ($cm^3 \cdot cm^{-3}$)
巴拉那河流域	0.616 7	0.045 8
博多河流域	0.673 4	0.056 2
2个区域总计	0.653 7	0.052 1

表 3　未考虑开放水域 CYGNSS 的 SM 反演结果

Table 3　CYGNSS of SM inversion results without considering open water

区域	R	RMSE/ ($cm^3 \cdot cm^{-3}$)
巴拉那河流域	0.570 0	0.047 7
博多河流域	0.631 3	0.058 9
2个区域总计	0.601 9	0.055 3

进行了比较，结果如图 9 所示。从图 9 可以得出，利用 CYGNSS 数据反演的 SM 能够作为 SMAP 的 SM 补充，也可提供更高时间分辨率的 SM 值。

由于实验区域内有大量的农田，且时常有洪涝灾害发生。基于 CYGNSS 反演得到的高时间分辨率 SM，后续可以对博多河和巴拉那河流域的洪水进行监测并及时提供预警信息以达到保护其周围农田的作用。

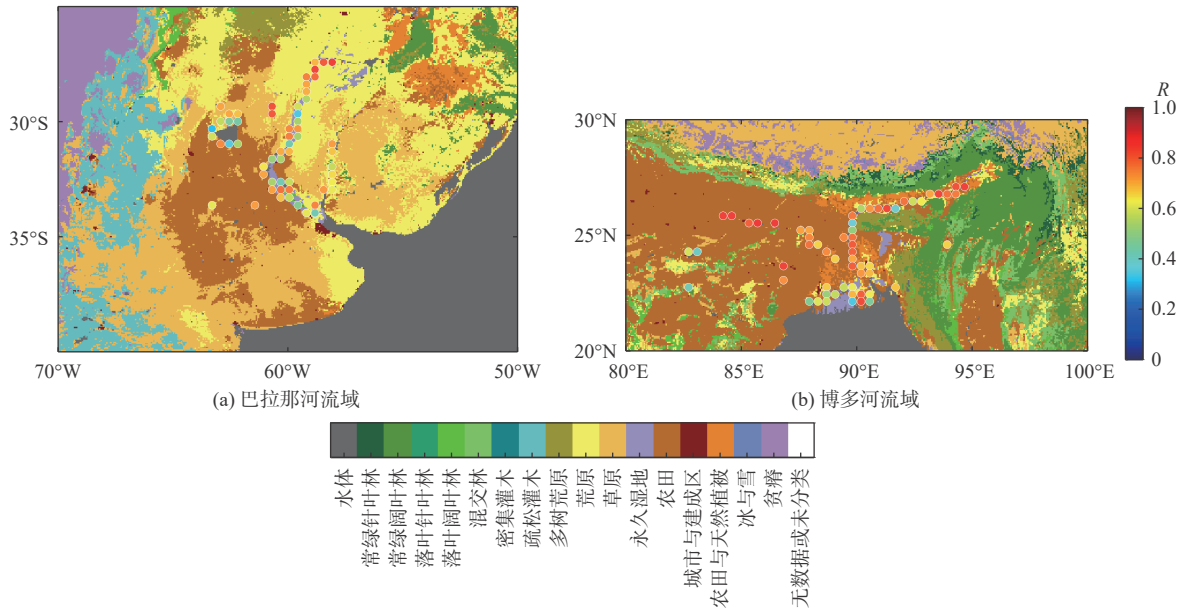


图 7 巴拉那河流域和博多河流域的反演结果的相关性分析

Fig. 7 Correlation analysis of retrieval results of Paraná river basin and Bodo river basin

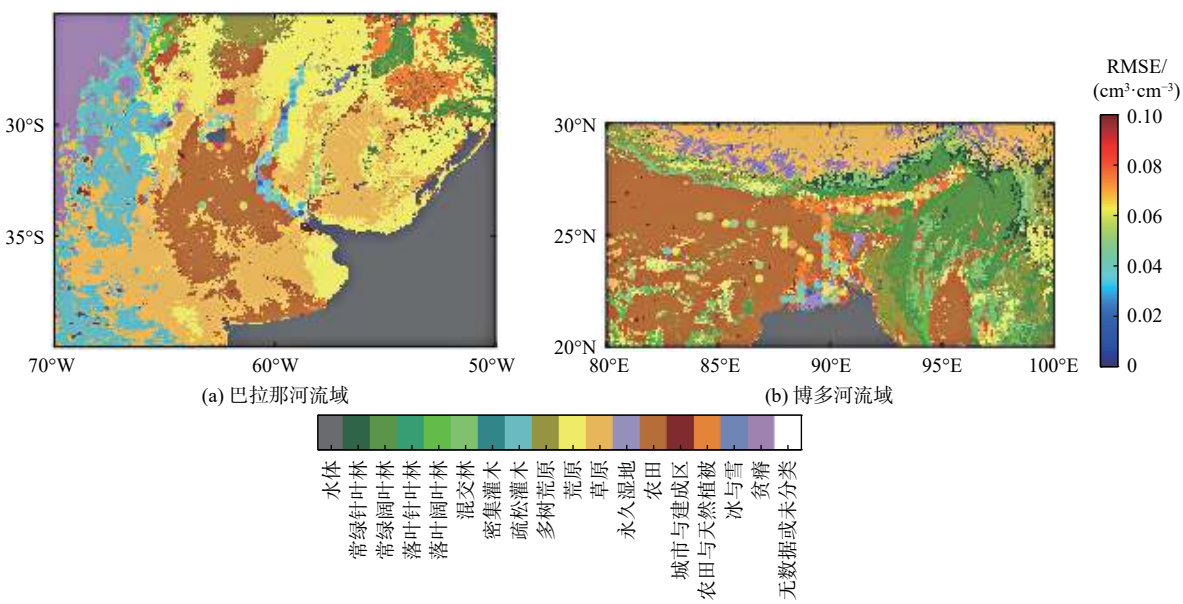


图 8 巴拉那河流域和博多河流域的反演结果的 RMSE 分析

Fig. 8 Retrieval results of Paraná river basin and Bodo river basin, respectively root mean square error analysis

3.4 与 UCAR\CU 反演结果的比较

目前，CYGNSS 也提供了相应的 SM 产品。其中，UCAR/CU 为 SM 反演产品，对于开放水域的剔除的方法是将 CYGNSS 镜面反射点作为参考点。对于每一个镜面反射点，找到围绕镜面点的 7 km×7 km 区域内的开放水域。如果 7 km×7 km 区域的开放水域的占比超过 1%，将不会对在此 CYGNSS 所有的观测量进行下一步处理。实验区域所对应的 UCAR\CU 的 SM 反演值与 SMAP 的 SM 值的比较结果如表 4 所示。从表 2 和表 4 可以得出，本文的开放水域剔除方法所反演的 SM 结果更优，相关性提高了 24%，RMSE 提高了 32%。

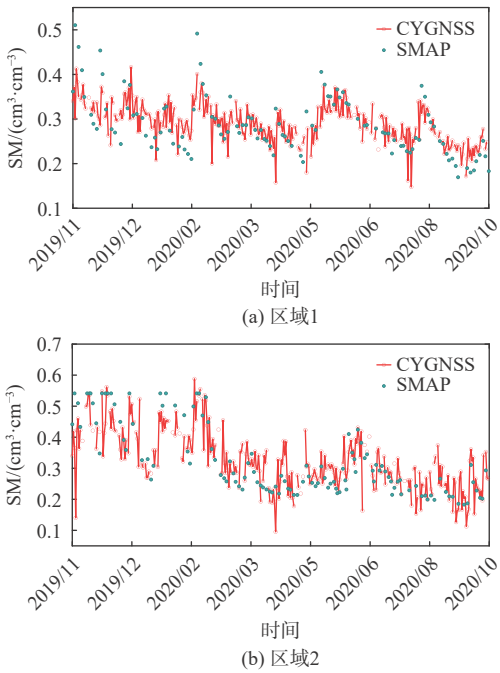


图 9　CYGNSS 反演结果与 SMAP 土壤湿度值的比较

Fig. 9　Comparison of CYGNSS retrieval results and SMAP soil moisture

表 4　UCAR\CU 的 SM 反演结果

Table 4　Retrieval results of SM from UCAR\CU

区域	R	RMSE/ ($cm^3 \cdot cm^{-3}$)
巴拉那河流域	0.522 1	0.085 0
博多河流域	0.529 6	0.073 6
2个区域总计	0.527 6	0.076 8

4　结　论

1) 将开放水域剔除后, SM 反演结果的精度和相关性得到了提升, 分别提升 6.2% 和 9%。

2) 反演结果有着较高的时间分辨率, 可以作为 SMAP 观测值的补充, 为这些水域密集的区域洪涝灾害的发生及农田保护提供预警信息。

参考文献（References）

[1] ENTEKHABI D, RODRIGUEZ-ITURBE I, CASTELLI F. Mutual interaction of soil moisture state and atmospheric processes[J]. Journal of Hydrology, 1996, 184(1-2): 3-17.

[2] KUENZER C, GUO H D, HUTH J, et al. Flood mapping and flood dynamics of the Mekong delta: ENVISAT-ASAR-WSM based time series analyses[J]. Remote Sensing, 2013, 5(2): 687-715.

[3] 杨东凯, 张其善. GNSS反射信号处理基础与实践[M]. 北京: 电子工业出版社, 2012.
YANG D K, ZHANG Q S. GNSS reflected signal processing: Fundamentals and applications[M]. Beijing: Publishing House of Electronics Industry, 2012 (in Chinese).

[4] 李黄, 夏青, 尹聪, 等. 我国GNSS-R遥感技术的研究现状与未来发展趋势[J]. 雷达学报, 2013, 2(4): 389-399.
LI H, XIA Q, YIN C, et al. The Current status of research on GNSS-R remote sensing technology in China and future development[J]. Journal of Radars, 2013, 2(4): 389-399(in Chinese).

[5] 汉牟田, 张波, 杨东凯, 等. 利用GNSS干涉信号振荡幅度反演土壤湿度[J]. 测绘学报, 2016, 45(11): 1293-1300.
HAN M T, ZHANG B, YANG D K, et al. Soil moisture retrieval utilizing GNSS interference signal amplitude[J]. Acta Geodaetica et Cartographica Sinica, 2016, 45(11): 1293-1300(in Chinese).

[6] 严颂华, 龚健雅, 张训械, 等. GNSS-R测量地表土壤湿度的地基实验[J]. 地球物理学报, 2011, 54(11): 2735-2744.
YAN S H, GONG J Y, ZHANG X X, et al. Ground based GNSS-R observations for soil moisture[J]. Chinese Journal of Geophysics, 2011, 54(11): 2735-2744(in Chinese).

[7] FOTI G, GOMMENGINGER C, JALES P, et al. Spaceborne GNSS reflectometry for ocean winds: First results from the UK TechDemoSat-1 mission[J]. Geophysical Research Letters, 2015, 42(13): 5435-5441.

[8] CHEW C, SHAH R, ZUFFADA C, et al. Demonstrating soil moisture remote sensing with observations from the UK TechDemoSat-1 satellite mission[J]. Geophysical Research Letters, 2016, 43(7): 3317-3324.

[9] CAMPS A, PARK H, PABLOS M, et al. Sensitivity of GNSS-R spaceborne observations to soil moisture and vegetation[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2016, 9(10): 4730-4742.

[10] YAN Q Y, HUANG W M, JIN S G, et al. Pan-tropical soil moisture mapping based on a three-layer model from CYGNSS GNSS-R data[J]. Remote Sensing of Environment, 2020, 247: 111944.

[11] CHEW C C, SMALL E E. Soil moisture sensing using spaceborne GNSS reflections: Comparison of CYGNSS reflectivity to SMAP soil moisture[J]. Geophysical Research Letters, 2018, 45(9): 4049-4057.

[12] KIM H, LAKSHMI V. Use of cyclone global navigation satellite system (CyGNSS) observations for estimation of soil moisture[J]. Geophysical Research Letters, 2018, 45(16): 8272-8282.

[13] CLARIZIA M P, PIERDICCA N, COSTANTINI F, et al. Analysis of CYGNSS data for soil moisture retrieval[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019, 12(7): 2227-2235.

[14] DONG Z N, JIN S G. Evaluation of the land GNSS-reflected DDM coherence on soil moisture estimation from CYGNSS data[J]. Remote Sensing, 2021, 13(4): 570.

[15] YANG T, WAN W, SUN Z G, et al. Comprehensive evaluation of using TechDemoSat-1 and CYGNSS data to estimate soil moisture over mainland China[J]. Remote Sensing, 2020, 12(11): 1699.

[16] CHEW C, SMALL E. Description of the UCAR/CU soil moisture product[J]. Remote Sensing, 2020, 12(10): 1558.

[17] RUF C S, ATLAS R, CHANG P S, et al. New Ocean winds satellite mission to probe hurricanes and tropical convection[J]. Bulletin of the American Meteorological Society, 2016, 97(3): 385-395.

[18] RUF C, ASHARAF S, BALASUBRAMANIAM R, et al. In-orbit performance of the constellation of CYGNSS hurricane satellites[J]. Bulletin of the American Meteorological Society, 2019, 100(10): 2009-2023.

[19] PEKEL J F, COTTAM A, GORELICK N, et al. High-resolution

mapping of global surface water and its long-term changes[J]. Nature, 2016, 540(7633): 418-422.

[20] ENTEKHABI D, NJOKU E G, O'NEILL P E, et al. The soil moisture active passive (SMAP) mission[J]. Proceedings of the IEEE, 2010, 98(5): 704-716.

[21] NGHIEM S V, ZUFFADA C, SHAH R, et al. Wetland monitoring with global navigation satellite system reflectometry[J]. Earth and Space Science, 2017, 4(1): 16-39.

[22] ZAVOROTNY V U, VORONOVICH A G. Scattering of GPS signals from the ocean with wind remote sensing application[J]. IEEE Transactions on Geoscience and Remote Sensing, 2000, 38(2): 951-964.

[23] VORONOVICH A G, ZAVOROTNY V U. Bistatic radar equation for signals of opportunity revisited[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 56(4): 1959-1968.

[24] DE ROO R D, ULABY F T. Bistatic specular scattering from rough dielectric surfaces[J]. IEEE Transactions on Antennas and Propagation, 1994, 42(2): 220-231.

[25] ULABY F T, LONG D G, BLACKWELL W, et al. Microwave radar and radiometric remote sensing[M]. Norwood: Artech House, 2013.

[26] WANG T L, RUF C S, BLOCK B, et al. Design and performance of a GPS constellation power monitor system for improved CYGNSS L1B calibration[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 12(1): 26-36.

[27] RODRIGUEZ-ALVAREZ N, PODEST E, JENSEN K, et al. Classifying inundation in a tropical wetlands complex with GNSS-R[J]. Remote Sensing, 2019, 11(9): 1053.

[28] YANG W T, XU T H, WANG N Z, et al. Research on the method of precisely removing open water in the retrieval of soil moisture by spaceborne GNSS-R[C]//Lecture Notes in Electrical Engineering. Singapore: Springer Singapore, 2021: 22-32.

[29] CAMPS A, PARK H, CASTELLVÍ J, et al. Single-pass soil moisture retrievals using GNSS-R: Lessons learned[J]. Remote Sensing, 2020, 12(12): 2064.

Influence of open water in retrieval of soil moisture by spaceborne GNSS-R

YANG Wentao[1], XU Tianhe[1, 2, *], WANG Nazi[2], GAO Fan[2], JING Lili[2], HE Yunqiao[2]

(1. School of Geological and Surveying Engineering, Chang'an University, Xi'an 710054, China;
2. Institute of Space Science, Shandong University, Weihai 264209, China)

Abstract: Soil moisture(SM) information can be obtained through the high spatial and temporal resolution spaceborne GNSS-R technology, however, open water will affect the accuracy of the retrieval results. To solve this problem, this paper proposes an improved method of removing open water. Prior to precisely removing the open water region based on the mask data, the simulated power must first be corrected to achieve the reflectivity.In order to verify the feasibility of the improved method, a one-year CYGNSS level 1 data was processed in two test areas with dense open water. The retrieval result was compared and analyzed with the soil moisture active passive (SMAP) level 3 soil moisture product. The root mean square error was 0.052 1 cm^3/cm^3, and the correlation was 0.654. Compared with the results before removing open water, it increased by 6.2% and 9%, respectively; Compared with the products officially released by CYGNSS, it increased by 32% and 24%. As a result, the retrieval results can be utilized to complement SMAP of SM measurements by providing highly time-resolved SM values.

Keywords: global navigation satellite system reflectometry; soil moisture; open water; cyclone global navigation satellite system; soil moisture active and passive

Received: 2021-08-23; **Accepted**: 2021-11-05; **Published Online**: 2021-11-17 09:00

URL: kns.cnki.net/kcms/detail/11.2625.V.20211116.1202.003.html

Foundation items: National Natural Science Foundation of China (41604003,41704017,41704018,41874032,41931076); National Key R & D Program of China (2020YFB0505800)

* **Corresponding author**. E-mail: thxu@sdu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0486

考虑任务分配的无人机信息交互拓扑生成

薛莹[1]，何锋[1,*]，谷晓燕[2]

(1. 北京航空航天大学 电子信息工程学院，北京 100191；2. 北京信息科技大学 信息管理学院，北京 100096)

摘 要：无人机（UAV）持久编队信息交互拓扑的优化是保证UAV编队结构稳定性和任务执行时效性的重要基础。现有的编队生成算法针对距离因素进行权重赋值和拓扑生成，由于未考虑任务分配因素，可能会引起整体任务执行时间过长甚至任务失败的问题，对UAV的能量也造成了不必要的损耗。以任务消息传输时间和能量损耗为关键优化目标，在保证UAV编队结构稳定的前提下，提出考虑任务分配因素的信息交互拓扑生成算法，优先连接承载实时性要求较高通信任务的关键汇聚链路，对剩余链路通过引入惩罚项，在权重上进一步将任务消息传输量因素考虑在内，生成最终的信息交互拓扑。使用OMNet++进行仿真验证，相比于只考虑距离因素的信息交互拓扑生成算法，所提算法在20架UAVs编队场景下，消息传输时间方面最高降低57.3%，最低降低28.1%，关键任务消息的到达时延降低了45.2%～51.6%，而任务执行过程单UAV的能量损耗总体减少了17.5%，平均每个节点减少损耗16.1%。

关 键 词：无人机编队；任务分配；信息交互拓扑；刚性矩阵；最优持久图

中图分类号：V243.1；N945.15

文献标志码：A **文章编号**：1001-5965(2023)07-1787-09

无人机(unmanned aerial vehicle, UAV)与无人机相关技术在近年来的发展越发热门和成熟，其中无人机编队进行任务执行是最值得研究的内容，例如民用领域的消防、救灾等[1]或军用领域的目标跟踪[2]、目标打击[3]等。

无人机群在保持编队飞行或进行任务执行的过程中，不管是位置的发送反馈还是任务相关消息的传递，无人机单体之间都需要保持信息交互，建立通信链路。通常，这种通信链接的集合被称为编队的信息交互拓扑[4]或通信拓扑[5]。密集的通信链接有时并不必要，保持所有链路的连接需要消耗大量能量，这将会大大影响无人机群的续航与任务的执行。因此，无人机之间信息的交互拓扑可以相对简化，成为编队通信网络拓扑中的一个子集[6]。无人机群编队作战的过程中，除了拓扑编队消息的传递，还需进行任务相关消息的传输，对于前者，使用“拓扑消息”指代，而后者则使用“任务消息”。如果在生成信息交互拓扑时能够考虑任务消息的传递，在任务消息传输时利用这些链路，那么任务消息的传输效率和机群的续航能力都能得到提升。

根据编队结构不同，编队控制方法可以分为领航跟随者编队、一致性编队、刚性编队[7]、持久编队[8]等。对于信息交互拓扑的研究也是基于这几类编队控制方法进行。其中持久编队相较于刚性编队，仅需要保持单向通信，可以大大降低能量损耗。持久编队是一种通信代价及能耗小、稳定性高的编队结构[9]，在保持飞行的过程中，与领航跟随者方法类似，通常有一个节点作为领航者按照预定路线进行运动[10]，其他节点按照生成的信息交互拓扑或与领航节点保持通信或与特定节点保持通信连接，由于

收稿日期：2021-08-25；**录用日期**：2021-11-26；**网络出版时间**：2022-01-25 08:39
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220124.1641.002.html
基金项目：国家自然科学基金 (62071023)
* **通信作者**. E-mail：robinleo@buaa.edu.cn

引用格式：薛莹，何锋，谷晓燕. 考虑任务分配的无人机信息交互拓扑生成 [J]. 北京航空航天大学学报，2023，49（7）：1787-1795.
XUE Y，HE F，GU X Y. UAV information interaction topology generation considering task allocation [J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（7）：1787-1795（in Chinese）.

相互的通信连接组成了持久图，且相对位置不变，能够很好的保持队形和编队结构的稳定性[11]。任务消息的传输与拓扑消息的传输类似，若能充分利用这些通信链接，则无需连通新链路，省掉了连接新链路的时间就相当于节省了执行该总体任务的总时间，也尽可能减少了无人机的能量损耗。

针对信息交互拓扑的研究，文献 [12] 证明，领航跟随者方法下信息交互拓扑是通信网络拓扑的一棵有向生成树，在该方法下，文献 [6] 提出使用通信距离衡量维持该通信链接的代价，文献 [4] 在此基础上考虑了位置误差对于拓扑和算法的影响，都使用了基于 Dijkstra 最短路径的领航跟随者编队信息交互拓扑优化算法。Luo 等[13] 在文献 [4] 的算法下，分别基于相对角度和相对距离，提出了 2 种适用于圆编队的算法。刚性图方法下，信息交互拓扑是通信网络拓扑的一个刚性图[12]：Priolo 等[7] 首先对文献 [13] 的2 种算法进行比较，最终在通信复杂性和最优性之间进行了折中，提出一种最优刚性网络的构造算法。

持久图方法下，信息交互拓扑是通信网络拓扑的一个持久图[14]，国内外诸多学者对此进行了探索和研究。Hendrickx 等[8] 使用一个有向图来描述节点间距离的约束，定义了持久图：如果所有相应的节点间构造的形状都保持不变，则图就是持久的。不仅推导了持久图的各种性质，给出了判定持久性的一个组合准则，而且证明了在刚性编队的基础上进行单向通信即可进一步保持编队的持久性和强壮性，为后续持久图的研究奠定了基础。罗小元等[15-16] 研究了最小加权持久图的生成方法，利用刚性矩阵证明最小加权刚性图的存在性，给出生成最小加权持久图的算法。文献 [17] 研究了基础圈为三角形、四边形及圆形圈结构下最小持久拓扑的生成。考虑到更多应用上的场景，文献 [18] 认为无人机之间的通信拓扑不仅是有向的，而且是可切换的；类似的，程潇[19] 对无人机编队组网技术进行研究，对其进行分簇，考虑在任务执行时无人机是可移动的，提出基于移动控制的机制。

现有的对于无人机群信息交互拓扑研究大多围绕在拓扑本身的构建，并未结合任务场景，生成的拓扑在编队飞行执行任务时容易造成能量浪费和因连接新链路而任务超时。本文将在已有任务分配方案情形下，使用持久图方法，以任务消息传输时间和无人机能耗为优化目标，提出考虑任务分配因素的信息交互拓扑生成算法，并通过 OMNet++ 进行了仿真验证。

1 信息交互拓扑模型

在实际应用中，无人机群编队飞行大多需要共同执行某项或某几项任务，一项任务的完成可能要靠编队中的几架无人机合作完成，这几架无人机内部需要进行感知探测、位置计算、指令下达、目标锁定等信息计算和行为实现，其之间也需要不断进行任务消息的传递。不同的拓扑结构将会影响消息的传输时效性，如若 2 节点间未事先建立链接，在需要进行消息传输时，就要进行链接的建立，而链路的距离不同，新链接建立的时长也不同，因而其成为了除任务规划因素之外，造成无人机任务执行空窗等待期的最主要原因。等待时长越久，任务的实时性就越差，执行任务的耗时越久，需要编队保持的时长越长，就会损耗越多的能量，极有可能影响后继任务[20] 的执行。因此，保证任务及时有效的完成，最小化任务执行所需耗用的能量，是无人机群编队任务执行过程中最主要的优化目标。

所有的通信链路可以认为是一个集合 U，集合 U 可以分为 3 个集合，如图 1 所示，C 越大代表该编队需要维持的链路越少，越节省能量。其中任务分配方案生成后，B 就是固定的，C 会受到信息交互拓扑集 A 的影响。A 和 B 可能会产生交集 A_2，A_2 中既需要传输拓扑消息，又需要传输任务消息的链路；因此相应的，A 可以对立的分为集合 A_1 和 A_2，A_1 表示只传输拓扑消息的链路集合；B 可以分为 2 个对立集合 A_2 和 B_2，B_2 表示需要单独连接只传输任务消息的链路。

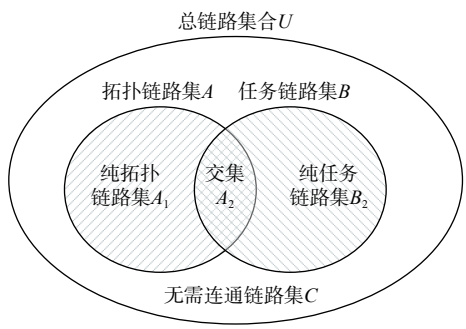


图 1 通信链路集合关系示意图

Fig. 1 Schematic diagram of relationship between communication link sets

1.1 无人机群编队的通信网络拓扑

无人机群编队的通信网络拓扑即集合 A 可以用一个赋权有向图 $D=(V,E,W)$ 来表示[6]。

1）$V=\{v_i\}(1\leqslant i\leqslant n)$ 为有向图中的节点集合，节点 v_i 表示 UAV$_i$，n 为 UAV 总数；

2）$E=\{\boldsymbol{e}_{i,j}\}$（$1\leqslant i,j\leqslant n, i\neq j$）为有向图中弧的

集合, $\boldsymbol{e}_{i,j}$表示从节点$v_i \to v_j$的有向弧, 即UAV$_i$到UAV$_j$的通信链路, 由于不考虑节点$v_i \to v_i$的环, 因此$i \neq j$;

3) $W = \{w(\boldsymbol{e}_{i,j})\}$ ($\boldsymbol{e}_{i,j} \in E$)为有向图中所有弧的权重集合, $w(\boldsymbol{e}_{i,j})$用来衡量UAV$_i$到UAV$_j$通信链路的代价。

1.2　无人机群编队的信息交互拓扑

信息交互拓扑由在维持编队过程中所需要使用的通信链路组成, 在持久编队方法下, 其是一个最优持久图。信息交互拓扑是编队通信网络拓扑的一个子集, 因此, 信息交互拓扑图可用$T = (V, E^*, W^*)$来表示, V为所有 UAV 的集合, E^*和W^*分别为E和W的子集, $T \subseteq D$。

最优持久图的相关定义如下:

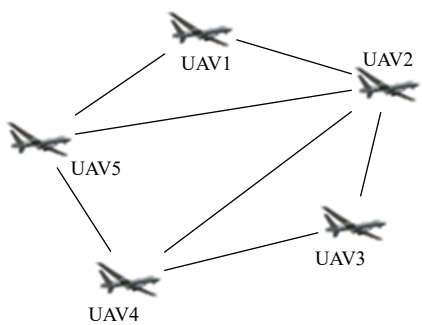
定义 1　最优持久图。如果一个有向图中各个节点的出度$d^-(i) \leqslant 2$, 且对应的无向图是最优刚性图, 则这个有向图是最优持久图。

定义 2　最优刚性图。若一个无向图是最优刚性图, 必须满足以下条件:

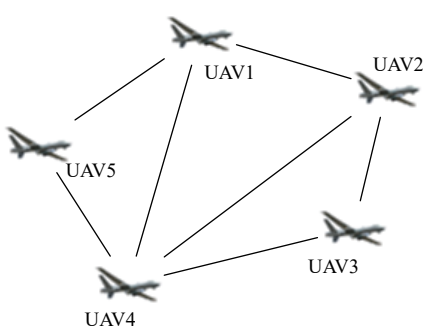
1) 对应的基础结构图是最小刚性的;

2) 在所有节点分布相同的刚性图中, 其各边边长加权和最小。

定义 3　最小刚性图。在二维空间中, 当刚性图含有$m(m > 2)$个顶点时, 当且仅当其边数$e = 2m - 3$时是最小刚性图。最小刚性图并不是唯一的。图 2 为同一组无人机不同的最小刚性图。



(a) 最小刚性图1



(b) 最小刚性图2

图 2　同一组无人机的不同的最小刚性图

Fig. 2　Different minimum rigidity diagrams of same group of UAVs

因此, 持久图方法下信息交互拓扑模型在某种程度上说是一个最优持久图, 首先, 需要满足最基础的最小刚性图的性质: 若共有 m 个节点, 则总边数$e = 2m - 3$; 其次, 在算法中, 除去关键链路之后, 剩余链路加权和最小, 满足最优刚性图; 最后, 作为最优持久图需满足所有节点的出度$d^-(i) \leqslant 2$。

由于拓扑的生成要考虑任务因素, 故链路权重不仅仅为 2 节点的距离, 本文将会详细说明权重的赋值逻辑。

2　信息交互拓扑链路权重更新方法

由第 1 节可知, 所有通信连接可以看作一个集合 U, 在 U 中若任务分配方案固定, 则集合 B 固定; 且信息交互拓扑模型属性之一为: 若共有 m 个节点, 则总边数$e = 2m - 3$, 也就相当于集合 A 的元素数量固定; 因此, 本文在生成信息交互拓扑集合 A 时, 若能将集合 A、B 的交集 A_2 扩大, 使较多的信息交互链路也能传输任务信息, 则可使连接的链路减少, 节省更多能量。

因此, 信息交互拓扑的生成需要根据任务分配方案的不同进行调整, 而信息交互拓扑生成最主要影响因素就是链路的权重, 无人机的位置、距离、信道优劣最终都反映在权重上, 故将任务因素的影响也反映在链路权重的赋值, 依据链路任务消息的负载量调整权重的大小, 使用通信代价来反映权重, 通信代价越小, 表示权重越大, 在生成时越优先考虑。在暂时忽略距离因素影响时, 任务消息传递越是频繁的有向链路, 认为在生成拓扑时该链路的优先级越高, 通信代价也将会越小。在实现上, 将会把任务分配因素和距离因素进行结合, 在距离项前提上对通信代价增加一个惩罚项, 任务消息传输越是频繁的链路, 惩罚越小, 其中传输次数最多的链路不做惩罚。链路通信代价 weight 为

$$\text{weight} = \text{normalization(distance)} + \text{punish} \tag{1}$$

式中: normalization(distance) 为归一化后的 2 节点间距离; punish 为对该链路的惩罚项。

对于一个由 k 架无人机组成的有向图$D = (V, E, W)$, 由于每条有向链路传输任务消息的负载量可能不同, 本文认为, 传输任务消息多的链路更为重要, 因此, punish 项也会根据传输次数不同而不同, 任务消息较多的链路给予较少的惩罚。具体操作过程如下: 首先对一个任务执行过程中所有弧 E 的任务消息传输次数进行统计, 传输次数最大链路的值为 max, 初始化一个 $k \times k$ 的链路通信代价矩阵, 矩阵所有元素均赋值为 max。接着对于每条有向链路, 分别用现值 max 减去链路对应的任务消息传输次数,

形成惩罚矩阵。

将归一化后的距离项矩阵与惩罚项矩阵相加后即得到了最终的链路代价矩阵，矩阵的每个元素都对应加权有向图的弧。

3 信息交互拓扑生成算法

在本节，将会引入一个“关键汇聚链路”的概念。在任务消息的传递和流转中，有一些消息是由多个无人机汇聚到一个“关键无人机”上的，这些消息对实时性的要求较高，同时到达或到达时间时延越小越有利于任务的完成，到达时间间隔较久则很可能会导致任务失败，“关键汇聚链路”就是映射这些消息流向的某几条有向链路，因此，在信息交互拓扑生成时，将会优先考虑这些链路的连接，以保证任务的完成。关键汇聚链路的指定依据消息传输的截止期限和汇聚时的节点度来完成。图3为任务分配与消息传递关系示意图，Task1～Task6为任务，依据Task3、Task4、Task5的消息截止时间，其共同汇入UAV4，形成任务执行链中的最大输入节点度，得到这些消息所对应的链路为关键汇聚链路，如虚线部分所示。

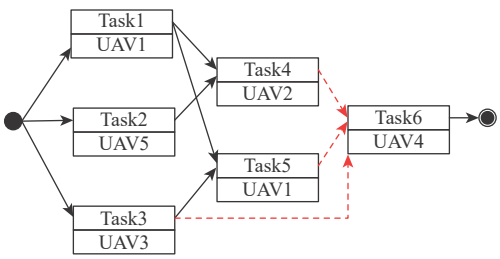


图3 任务分配与消息传递关系示意图

Fig. 3 Relationship diagram between task allocation and message passing

图4为本文算法流程，具体如下：首先，将加权有向图的弧根据链路通信代价由小到大的顺序排列，初步得到一个待处理弧集合L'；接着，将关键汇聚链路从L'中提取出来，根据权重排序后插入L'最前端，得到待处理弧集L，并初始化一个空的加权有向图D及空刚度矩阵$\boldsymbol{M}$。

然后，根据生成的待处理弧集，将每条弧依次添加进有向图D中，形成D'进行以下计算：

1）计算图D'的刚度矩阵$\boldsymbol{M}'$，判断矩阵$\boldsymbol{M}'$是否满秩；

2）统计有向图D'所有节点的出度，是否满足所有节点的出度$\leqslant 2$。

如果上述2个条件都满足，则保留该刚度矩阵$\boldsymbol{M}'$为$\boldsymbol{M}$，且更新该有向图D'为D，即保留该弧；如果上述有任意条件不满足，则将刚度矩阵中新添加的一行删除，还原为$\boldsymbol{M}$，并且不更新有向图保持原有向图D，即不添加该弧，直到L中的所有弧都执行过该操作。

刚度矩阵的计算原理如下：在二维空间$\mathbf{R}^2$中，有向图$D=(V,E,W)$所对应的任意一条弧$\boldsymbol{e}_{ij}$，若节点v_i的坐标为(x_i,y_i)，节点v_j的坐标为(x_j,y_j)，则其所对应的刚度矩阵的行元素为$[0,\cdots,x_i-x_j,y_i-y_j,0,\cdots,0,x_j-x_i,y_j-y_i,0,\cdots,0]$，其中除第$2i-1$、第$2i$、第$2j-1$、第$2j$列位置的值依次为行元素中的$x_i-x_j$、$y_i-y_j$、$x_j-x_i$、$y_j-y_i$，其他位置值均为0。

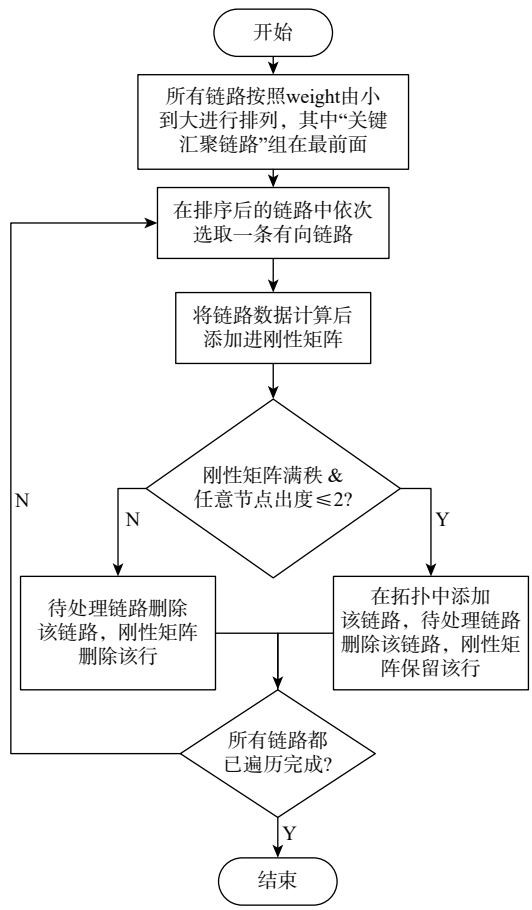


图4 本文算法流程

Fig. 4 Flow chart of proposed algorithm

4 本文算法仿真评估

4.1 算法评估指标

由于本文算法主要想解决的问题场景是任务分配下信息交互拓扑的生成，在同一任务分配方案下，不同算法生成的信息交互拓扑对任务消息传递的影响是不同的。因此，在对本文算法进行评估时主要考虑在不同拓扑下任务消息传输时间、同一任务执行时消耗的能量。

4.2 小规模实验仿真

本实验在OMNet++平台搭建数量为10的无人机编队进行作战，假设所有无人机均不会出现故障。

在仿真的场景中, 任务列表如表 1 所示。

在资源限制和无人机能力匹配的条件下, 使用遗传算法对无人机进行任务分配, 具体实现参照文献 [21]。生成的任务分配方案如表 2 所示。在表中分配方案中, 不失一般性, 假定任务消息的截止时间与其周期一致, 依据任务消息的截止时间大小和遗传算法计算出来任务分配结果, 可以计算各个节点的节点度, 并选定关键汇聚链路: 编号 11-12 消息传递关系所代表的链路, 即弧 $\boldsymbol{e}_{8,9}$、$\boldsymbol{e}_{8,10}$、$\boldsymbol{e}_{7,8}$、$\boldsymbol{e}_{7,9}$、$\boldsymbol{e}_{7,10}$, 对这几条链路应优先进行保证。

表 1　任务列表

Table 1　Task list

任务组	任务	时间/s	耗费资源	后继任务
目标侦测	TD1	20	3	TD3
	TD2	16	4	TD3
	TD3	20	8	TP1
搜索锁定	TT1	13	7	TT2
	TT2	20	10	TP1
综合导航	TN1	10	4	TN5
	TN2	17	4	TN5
	TN3	14	4	TN5
	TN4	13	5	TN5
	TN5	20	9	
信息融合	TP1	20	3	TP3, TF1, TF2
	TP2	16	1	TP3, TF1, TF2
	TP3	20	6	TN5
火力攻击	TF1	14	10	TF3
	TF2	13	10	TF3
	TF3	18	10	

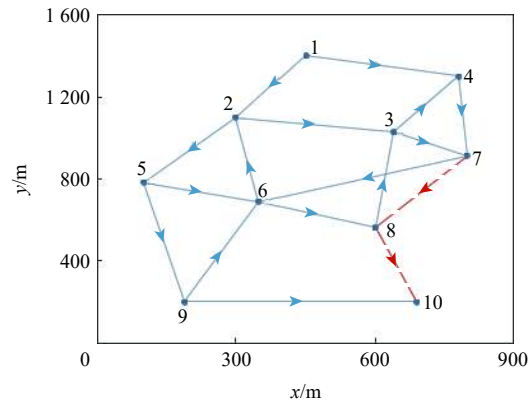
表 2　任务分配方案

Table 2　Task assignment scheme

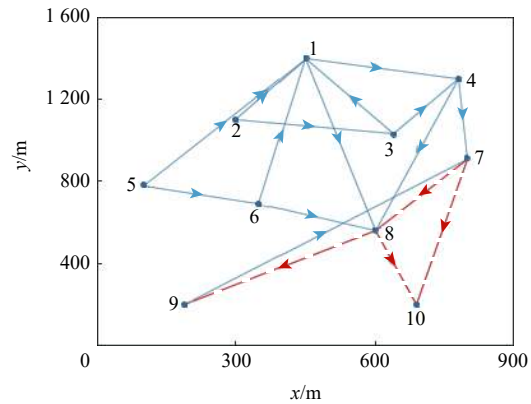
编号	任务	对应无人机	后继任务
1	TD1	UAV1	TD3
2	TD2	UAV2	TD3
3	TD3	UAV1	TP1
4	TT1	UAV3	TT2
5	TT2	UAV4	TP1
6	TN1	UAV2	TN5
7	TN2	UAV3	TN5
8	TN3	UAV6	TN5
9	TN4	UAV5	TN5
10	TN5	UAV1	
11	TP1	UAV8	TP3, TF1, TF2
12	TP2	UAV7	TP3, TF1, TF2
13	TP3	UAV8	TN5
14	TF1	UAV9	TF3
15	TF2	UAV10	TF3
16	TF3	UAV7	

1) 拓扑结构

使用距离因素算法与本文算法, 在仿真环境中随机生成无人机的位置, 进行仿真验证得到的拓扑如图 5 所示。对比图 5(a)和图 5(b), 结合第 3 节定义的关键汇聚链路与链路负载情况, 能够看出考虑任务因素的算法生成的信息交互拓扑较大程度保证了关键汇聚链路的连接, 如图中虚线(红色)链路所示, 而传统算法生成的信息交互拓扑则并不能保证。



(a) 距离因素算法



(b) 本文算法

图 5　小规模案例使用 2 种算法生成的信息交互拓扑

Fig. 5　Information interaction topology generated by two algorithms in small-scale cases

2) 消息传输时间

在本实验中, 执行任务消耗对应无人机能量为 20; 每单位距离消耗源节点能量为 5[4]; 正常发送一条任务消息消耗能量为 10, 接收一条任务消息消耗能量为 5[4]; 节点间建立新链接按照文献 [22] 所用方法, 所需耗费源节点能量为 15, 消耗目的节点能量为 15。

在本节条件约束下, 使用距离因素算法与本文算法得到的信息交互拓扑分别对表所示的任务方案进行仿真一个周期后, 消息传输时间的对比图如图 6 所示。其中, 图中未列出的 UAV 表示在该方案下无需接收任务消息。

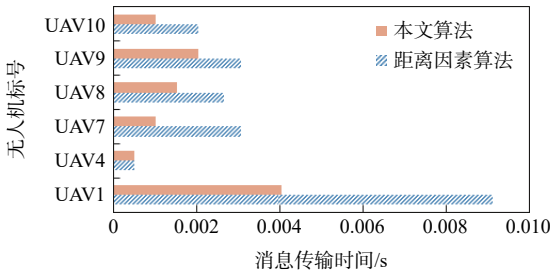


图 6 小规模实验时 2 种算法生成的拓扑下部分无人机任务执行结果对比

Fig. 6 Comparison of task execution results under two algorithms' topologies in small-scale experiment

同一无人机在 2 种算法下执行的结果不同。

1）对每架无人机来说，约半数无人机在任务因素的信息交互拓扑下进行任务执行的消息传输时间都要小于距离因素的信息交互拓扑。如 UAV1，在纯距离因素下由于需要接收其他 UAV 请求的链接，使用了约 0.009 s 的时间建立信息交互拓扑未连接的有向链路和进行任务消息的传输，而本文算法因生成的信息交互拓扑已事先考虑了任务因素的影响，建立了部分链路，仅使用了约 0.004 s 的时间，节省了约 55.4% 的时间。同理，UAV7、UAV8、UAV9 和 UAV10 分别节省 66.7%、42.3%、33.2% 和 50% 的时间。而对于其他 UAV 来说，在仿真的方案下，由于不需要作为目的节点接收其他源节点的任务消息，所以无需重新建立有向链接，在 2 种算法下传输消息所用时间相同。

2）对于关键汇聚任务消息，通过对结果观察能够得到：部分无人机如 UAV10，由于 2 种算法生成的信息交互拓扑在本案例中都连接了其作为接收无人机的关键链路，关键任务消息的接收没有什么不同。而大多无人机如 UAV8、UAV9，在距离因素算法下，接收到第 1 条关键任务消息后，由于需要接收其他任务消息才能执行第 2 条关键任务，或是由于需要建立链路才能接收到第 2 条任务消息，在第 1 条关键任务消息到达后，分别经过了约 0.003 s 和 0.001 5 s 的时间剩余任务消息才相继到达；而在本文算法下，关键任务消息的到达时延分别只有约 0.001 s 和 0.000 5s，均减少了约 66.7%，能尽可能的降低任务等待时延过长造成的任务失败概率。

3）能量损耗

图 7 为 2 种算法下执行任务的能量消耗情况，能够看出，大部分无人机在本文算法下的能量消耗情况相较于距离因素算法有明显的改善，例如 UAV9、UAV1、UAV5、UAV2、UAV6 分别能够节省能量约 45.4%、36.2%、25.7%、21.8%、17.2%。主要原因在于距离因素算法的拓扑由于没有考虑任务方案，在进行任务消息的传输时需要进行较多链路的建立，消耗较多能量；而本文算法生成的拓扑由于在拓扑生成时就较多的保证了任务消息传输链路的连接，需要重新连接的链路很少，因此能够有效减少能量消耗。

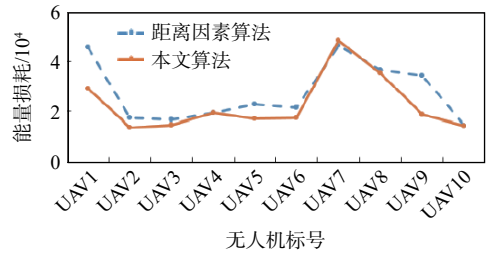


图 7 2 种算法下执行任务的能量消耗情况

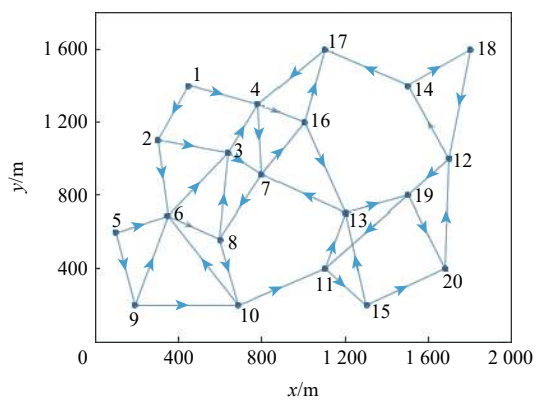
Fig. 7 Energy consumption of tasks under two algorithms

4.3 复杂场景实验仿真

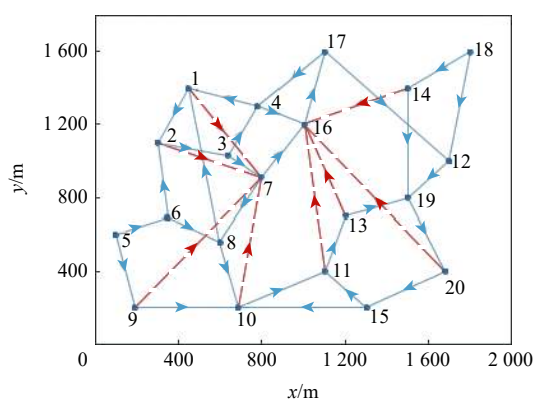
在 10 个任务组，共计 43 个任务，20 架无人机编队的场景下进行实验仿真，得到仿真结果如下：

1）拓扑结构

使用距离因素算法与本文算法，在仿真环境中随机生成无人机的位置，进行仿真验证得到的拓扑如图 8 所示。



(a) 距离因素算法



(b) 本文算法

图 8 复杂案例使用 2 种算法生成的信息交互拓扑图

Fig. 8 Information interaction topology generated by two algorithms in complex cases

根据生成的信息交互拓扑，同样能够看到，距离因素算法生成的信息交互拓扑并不能有效保证

关键汇聚链路的连接, 甚至在本算例场景下, 没有关键汇聚链路得到连接。而本文算法能够使关键汇聚链路得到连接, 如图 8(b) 中虚线(红色)链路所示。

2) 消息传输时间

在与 4.2 节同等的条件约束下, 使用距离因素算法与本文算法得到的信息交互拓扑分别对任务分配方案进行仿真一个周期后, 消息传输时间的对比如图 9 所示。

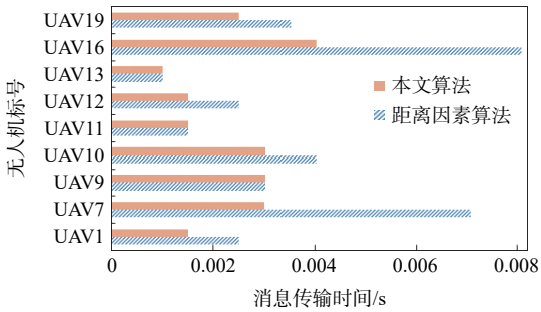


图 9　2 种算法生成的拓扑下部分无人机任务执行结果对比

Fig. 9　Comparison of mission execution results of some UAVs under two algorithms' topologies

根据本算例仿真执行结果, 可以得到: 约半数无人机在本文算法下进行任务执行的消息传输时间要小于距离因素算法的信息交互拓扑, 如 UAV1 和 UAV12 节省了 40.2% 的时间, UAV10 节省了 29.1%, UAV16 节省了 50.2%, 最低为 UAV19 节省了 28.1%, 最高为 UAV7 节省了 57.3%。对于关键汇聚任务消息的到达时延, 分别减少了约 45.2% 和 51.6%, 尽可能减少了任务等待时延过长导致的任务失败概率。

3) 能量损耗

图 10 为 2 种不同算法下执行任务能量消耗情况, 能够看出, 85% 的无人机在本文算法下的能量消耗情况相较于距离因素算法有明显的改善, 其中 UAV16 最高节省约 40% 的能量, 而最低为 UAV4 节省了 4.1%; 剩余只有 15% 的节点即 UAV2、UAV5、UAV9 这 3 架无人机损耗的能量与纯距离因素算法相同。机群总体节省了 17.5% 的能量, 平均每架无人机节省了 16.1%, 有效减少了能量损耗。

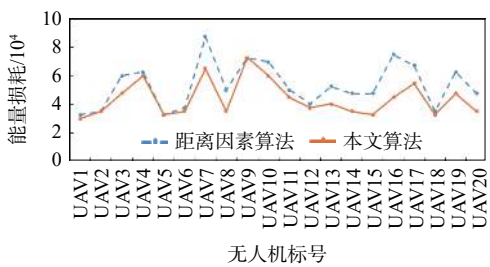


图 10　复杂场景下 2 种算法下执行任务的能量消耗情况

Fig. 10　Energy consumption of tasks under two algorithms in complex experiment

5　结　论

1) 在相同的任务分配方案和场景下, 本文算法由于考虑了任务消息传递关系, 对于那些互相有依赖关系的任务较为友好, 单无人机任务消息传输时间要短于或等于距离因素算法。

2) 对于关键汇聚链路, 由于优先保证了这些关键链路的连接, 在多任务汇聚的场景下, 关键任务消息到达的时间间隔相较距离因素算法得到了有效降低, 很大程度上可降低了任务失败的可能。

3)在能量上, 由于将任务消息传递关系考虑在了信息交互拓扑的生成算法上, 减少了新链路的建立数量, 相较距离因素算法, 对各节点都起到了降低能量损耗的作用。

参考文献 (References)

[1] LUO F, JIANG C X, DU J, et al. A distributed gateway selection algorithm for UAV networks[J]. IEEE Transactions on Emerging Topics in Computing, 2014, 3(1): 22-33.

[2] DUTTA R, SUN L, PACK D. A decentralized formation and network connectivity tracking controller for multiple unmanned systems[J]. IEEE Transactions on Control Systems Technology, 2018, 26(6): 2206-2213.

[3] HE B, ZHANG X Y, LI Q Y. Effectiveness measurement of UAV combat in uncertain environment[J]. Journal of Physics:Conference Series, 2018, 1069(1): 012048.

[4] 董文奇, 何锋. 大规模UAV编队信息交互拓扑的分级分布式生成[J]. 航空学报, 2021, 42(6): 324380.
DONG W Q, HE F. Hierarchical and distributed generation of information interaction topology for large scale UAV formation[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(6): 324380(in Chinese).

[5] LI W, WANG L, LI H B. Design of distributed time-cooperative guidance of multiple flight vehicles with un-controllable velocity[C]//2019 Chinese Control and Decision Conference. Piscataway: IEEE Press, 2019: 2419-2425.

[6] 王国强. 面向队形保持的无人机编队信息交互拓扑优化问题的研究[D]. 合肥: 合肥工业大学, 2016: 19.
WANG G Q. Research on information exchange topology optimization problem of UAV formation during formation keeping[D]. Hefei: Hefei University of Technology, 2016: 19 (in Chinese) .

[7] PRIOLO A, WILLIAMS R K, GASPARRI A, et al. Decentralized algorithms for optimally rigid network constructions[C]//2014 IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2014: 5010-5015.

[8] HENDRICKX J M, ANDERSON B D O, DELVENNE J C, et al. Directed graphs for the analysis of rigidity and persistence in autonomous agent systems[J]. International Journal of Robust and Nonlinear Control, 2007, 17(10-11): 960-981.

[9] 赵诗雪, 赵太飞, 刘昆, 等. 基于无线紫外光MIMO通信的优化持久编队算法[J]. 激光杂志, 2020, 41(10): 8-13.

ZHAO S X, ZHAO T F, LIU K, et al. Algorithm on optimal persistent formation based on wireless ultraviolet MIMO communication[J]. Laser Journal, 2020, 41(10): 8-13(in Chinese).

[10] 罗贺, 李晓多, 王国强. 能耗均衡的三维最优持久编队通信拓扑生成[J]. 航空学报, 2022, 43(1): 324922.
LUO H, LI X D, WANG G Q. Generation of three dimensional optimal persistent formation communication topology with energy consumption equilibrium[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(1): 324922(in Chinese).

[11] 王金然, 罗小元, 杨帆, 等. 三维最优持久编队拓扑生成策略[J]. 自动化学报, 2015, 41(6): 1123-1130.
WANG J R, LUO X Y, YANG F, et al. Generation strategy of optimal persistent formation topology in 3D space[J]. Acta Automatica Sinica, 2015, 41(6): 1123-1130(in Chinese).

[12] REN W. Consensus strategies for cooperative control of vehicle formations[J]. IET Control Theory & Applications, 2007, 1(2): 505-512.

[13] LUO X Y, SHAO S K, ZHANG Y Y, et al. Generation of minimally persistent circle formation for a multi-agent system[J]. Chinese Physics B, 2014, 23(2): 614-622.

[14] EREN T, ANDERSON B, MORSE A S, et al. Operations on rigid formations of autonomous agents[J]. Communications in Information and Systems, 2003, 3(4): 223-258.

[15] LUO X Y, LI S B, GUAN X P. Automatic generation of Min-weighted persistent formations[J]. Chinese Physics B, 2009, 18(8): 3104-3114.

[16] 罗小元, 杨帆, 李绍宝, 等. 多智能体系统的最优持久编队生成策略[J]. 自动化学报, 2014, 40(7): 1311-1319.
LUO X Y, YANG F, LI S B, et al. Generation of optimally persistent formation for multi-agent systems[J]. Acta Automatica Sinica, 2014, 40(7): 1311-1319(in Chinese).

[17] 邵士凯. 多智能体系统最优持久编队生成与圆形编队控制[D]. 秦皇岛: 燕山大学, 2013: 1-66.
SHAO S K. Optimally persistent formation generation and circle formation control for multi-agent syetems[D]. Qinhuangdao: Yanshan University, 2013: 1-66 (in Chinese).

[18] 刘伟, 周绍磊, 祁亚辉, 等. 有向切换通信拓扑下多无人机分布式编队控制[J]. 控制理论与应用, 2015, 32(10): 1422-1427.
LIU W, ZHOU S L, QI Y H, et al. Distributed formation control for multiple unmanned aerial vehicles with directed switching communication topologies[J]. Control Theory & Applications, 2015, 32(10): 1422-1427(in Chinese).

[19] 程潇. 无人机编队组网技术研究[D]. 南京: 南京大学, 2019: 25-38.
CHENG X. Research on networking techniques of unmanned aerial vehicle formation[D]. Nanjing: Nanjing University, 2019: 25-38 (in Chinese).

[20] 张小庆, 胡亚捷. 基于优先级与关键路径的工作流任务调度算法研究[J]. 武汉轻工大学学报, 2021, 40(2): 59-67.
ZHANG X Q, HU Y J. Research on workflow tasks scheduling algorithm based on priority and critical path[J]. Journal of Wuhan Polytechnic University, 2021, 40(2): 59-67(in Chinese).

[21] 邓启波. 多无人机协同任务规划技术研究[D]. 北京: 北京理工大学, 2014: 28-55.
DENG Q B. Cooperative task planning of multiple unmanned aerial vehicles[D]. Beijing: Beijing Institute of Technology, 2014: 28-55 (in Chinese) .

[22] 马霓, 孙礼, 贾群力, 等. 无线通信网络中点到点对等通信无线链接建立和保持的方法与装置: 中国, CN1527623A[P]. 2004-09-08.
MA N, SUN L, JIA Q L, et al. Method and apparatus for establishing and retaining point-to-point communication radio chaining in radio communication network: China, CN1527623A[P]. 2004-09-08 (in Chinese).

UAV information interaction topology generation considering task allocation

XUE Ying[1], HE Feng[1,*], GU Xiaoyan[2]

(1. School of Electronic and Information Engineering, Beihang University, Beijing 100191, China;
2. School of Information Management, Beijing Information Science and Technology University, Beijing 100096, China)

Abstract: The optimization of information interaction topology for persistent unmanned aerial vehicle (UAV) formation is an important basis for the stability of UAV formation structure and the timeliness of task execution. The existing formation generation algorithms assign weights and generate topologies based on the factor of distance rather than task allocation; therefore, the overall task execution may be too long or even fail, causing unnecessary loss of UAV energy. This study proposes an information interaction topology optimization algorithm considering the factor of task allocation, with the task message transmission time and energy loss being the key optimization objectives, and with the premise of ensuring the stability of UAV formation structure. The key aggregation links performing high real-time communication tasks are preferentially connected, and the penalty term is introduced for the remaining links. Furthermore, the weight factor of task message transmission is taken into account to generate the final information interaction topology. OMNet++ is used for simulation verification. Compared with the information interaction topology generation algorithm considering only the distance factor, the algorithm considering the task factor can reduce the message transmission time by 57.3% and 28.1% at the highest and lowest levels in the formation scenario of 20 UAVs, respectively. The arrival delay of mission-critical messages is reduced by 45.2% to 51.6%. During the mission execution, the energy loss of a single UAV is reduced by 17.5% overall, and the loss per node is reduced by 16.1% on average.

Keywords: UAV formation; task assignment; information interaction topology; rigid matrix; optimal persistence graph

Received: 2021-08-25; **Accepted**: 2021-11-26; **Published Online**: 2022-01-25 08:39

URL: kns.cnki.net/kcms/detail/11.2625.V.20220124.1641.002.html

Foundation items: National Natural Science Foundation of China (62071023)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0513

基于LPV的主动隔振平台单支腿鲁棒控制器设计

唐育聪，朱庆华*，刘付成，曹赫扬
(上海航天控制技术研究所，上海 201109)

摘 要：利用控制力矩陀螺(CMG)实现敏捷姿态机动控制时，保证机动过程控制输出力矩快速响应的同时，还需要保证机动到位后有效隔离CMG的机械振动以实现高稳定度姿态控制。针对控制CMG主动隔振平台，将控制CMG外框角速度作为变参数，提出单支腿主动隔振平台控制器的一种基于线性变参数（LPV）控制设计实现方法，通过与其他主动隔振控制方法性能比较分析，所提方法在兼顾敏捷姿态机动期间和机动到位后对主动隔振平台不同的力学传递要求方面有更好的性能。

关 键 词：控制力矩陀螺；主动隔振；姿态机动；线性变参数；鲁棒控制
中图分类号：TP273
文献标志码：A **文章编号**：1001-5965(2023)07-1796-06

敏捷姿态机动卫星的执行机构均选用控制力矩陀螺群(control moment gyros, CMGs)以保证其快速机动性能。为提高姿态机动到位后卫星的姿态稳定度，需要对CMGs配套安装隔振平台，以消除CMG的机械振动对姿态稳定度的影响。在对CMG进行隔振时，主动隔振方式虽然工程代价相对较高，但可以获取比被动隔振方式更好的振动抑制性能。主动隔振平台的控制器可以采用不同的方法设计和实现，传统的比例积分微分(proportion integral derivative, PID)控制器可以得到低转折频率且没有谐振的响应特性。当卫星根据机动任务需求可以自主调节CMGs的响应快慢时，需要CMGs根据需求调节其外框最大转速以具备不同的力矩输出能力，当CMGs的外框转速根据机动任务不断动态调整时，对其主动隔振平台的参数适应性提出了新的需求，固定转折频率的主动隔振平台已不能完全适应。

目前较为广泛应用于主动隔振平台的控制器有PID[1-2]、H∞[2]控制器等，2种控制器都能起到较好的隔振效果，但也存在参数适应性的局限性，无法兼顾不同机动需求下CMGs变带宽工况的高精度力矩传递和高性能隔振。为改善这一情况，考虑应用线性变参数(linear parameter varying, LPV)控制器，根据系统中实时可测的变参数来改变控制器的增益，并依靠H∞控制器的范数度量，使变参数能够直接影响控制器的参数。现有LPV控制器广泛用于高速飞行器等变参数范围广、速率快的背景下，对主动隔振平台的控制器设计应用较少。文献[3-5]介绍了LPV控制器的相关概念和求解过程，文献[6-7]介绍了凸分解方法，文献[8]介绍了高阶奇异值分解(higher order singular value decomposition, HOSVD)方法的LPV控制器建模过程，文献[9-10]介绍了控制器的线性矩阵不等式(linear matrix inequality, LMI)求解方法，提供了LPV控制器的设计思路和完整的求解过程。

本文主要针对CMGs的变带宽主动隔振性能和鲁棒性综合需求[11-15]，考虑星体快速机动和机动到位2种工况下控制器的自适应性。通过引入CMG外框架角速度作为系统变参数设计了LPV控制器，具备不同机动需求的主动隔振平台控制器自适应变带宽的能力。先建立了单支腿主动隔振平台单

收稿日期：2021-09-05；录用日期：2021-11-26；网络出版时间：2021-12-16 12:21
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20211215.1227.001.html
*通信作者. E-mail：zhuqh_79@sina.com

引用格式：唐育聪，朱庆华，刘付成，等. 基于LPV的主动隔振平台单支腿鲁棒控制器设计[J]. 北京航空航天大学学报，2023，49(7)：1796-1801. TANG Y C，ZHU Q H，LIU F C，et al. Design of robust controller for single outrigger of vibration active isolation platform based on LPV[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1796-1801 (in Chinese).

通道控制的动力学模型和输出加权模型, 在此基础上针对输出加权模型, 引入了 CMG 外框架角速度作为 LPV 控制器的变参数, 并通过极点配置方法设计出相应的鲁棒控制器。最后通过数学仿真比较了本文设计的 LPV 控制器与 PID、H∞控制器之间的主动隔振性能差异。

1　单支腿主动隔振平台模型描述

图 1 为单支腿隔振的简化结构模型, 弹簧隔振模块对振动干扰进行被动隔振, 音圈电机驱动模块通过控制器调节支腿隔振带宽, 提升振动抑制效果, 实现对振动的主动和被动抑制。

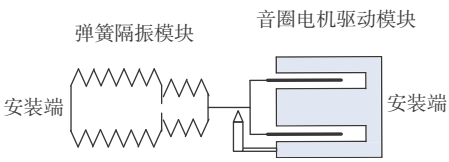


图 1　单支腿隔振的简化结构模型

Fig. 1　Simplified structure model of single vibration isolation outrigger

根据文献 [11] 中的主动隔振描述方法, 给出单支腿主动隔振平台控制器的动力学模型如图 2 所示。

考虑 CMG 端为输入端, 振源端位移大小为 s_1, 质量为 m; 支腿连接的星体端的位移大小为 s_0, 质量为 M; 弹簧等效弹性系数为 k, 音圈电机内阻为 R, 内感为 L, 传导常数 $T=nBc$, 其中 n 为电机线圈匝数, B 为电机内磁场强度, c 为切割电机内磁场的导线周长。

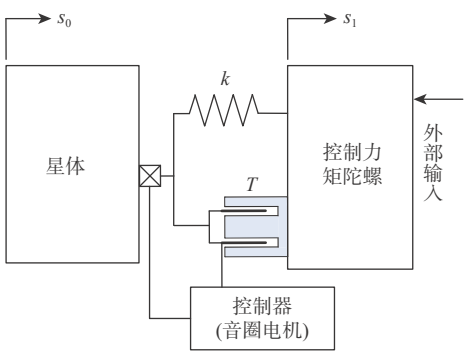


图 2　单支腿主动隔振平台控制器动力学模型

Fig. 2　Dynamic model of single leg isolation platform controller

根据牛顿第二定律和电磁学相关物理原理确定含控制器的主动隔振平台力学方程为

$$\begin{cases} m\ddot{s}_1 + k(s_1 - s_0) - Ti = 0 \\ M\ddot{s}_0 - k(s_1 - s_0) + Ti = 0 \\ L\dot{i} + T(\dot{s}_1 - \dot{s}_0) + Ri = u \end{cases} \tag{1}$$

式中: $u=Hy$ 为控制器, H 为控制器传递函数, $y=k(s_1-s_0)-Ti$ 为力传感器测量值。为得到完全可控的系统模型并用于系统综合, 选取单支腿主动隔振平台单通道相对位移大小 $l=(s_1-s_0)$、相对速度 $\dot{l}$ 和电机线圈电流 i 作为系统状态量构建状态方程, 振源端加速度 $\ddot{s}_1$ 作为系统的外部扰动输入 w, 力传感器作为测量 y 构建观测方程, 可得到系统模型为

$$\begin{cases} -M\ddot{l} - kl + Ti = -M\ddot{s}_1 \\ L\dot{i} + Tl + Ri = u \\ kl - Ti = y \end{cases} \tag{2}$$

整理得到系统状态方程为

$$\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}_1 w + \boldsymbol{B}_2 u \\ y = \boldsymbol{C}_2\boldsymbol{x} + D_{21}w + D_{22}u \end{cases} \tag{3}$$

式中:

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 \\ -k/M & 0 & T/M \\ 0 & -T/L & -R/L \end{bmatrix},\ \boldsymbol{B}_2 = \begin{bmatrix} 0 \\ 0 \\ 1/L \end{bmatrix},$$

$$\boldsymbol{C}_2^{\mathrm{T}} = \begin{bmatrix} k \\ 0 \\ -T \end{bmatrix},\ \boldsymbol{B}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},\ \boldsymbol{x} = \begin{bmatrix} l \\ \dot{l} \\ i \end{bmatrix},$$

$D_{21}=0$, $D_{22}=0$, $w=\ddot{s}_1$。

考虑到主动隔振平台支腿低频段刚性, 高频段衰减的特性, 引入 $\ddot{s}_0$ 作为系统的被调输出, 对 $\ddot{s}_0$ 添加加权函数 W_z 得到加权后的变量 z, 使 w 到 z 的传递函数的无穷范数 $\|\boldsymbol{T}_{wz}\|_\infty$ 近似等于 1, 保证单支腿在低频段实现 $\ddot{s}_1$ 与 $\ddot{s}_0$ 为 1:1 输入输出, 构建加权函数后的系统框图如图 3 所示, 其中 $H(s)$ 为控制器传递函数, $\tilde{z}$ 为被调输出。

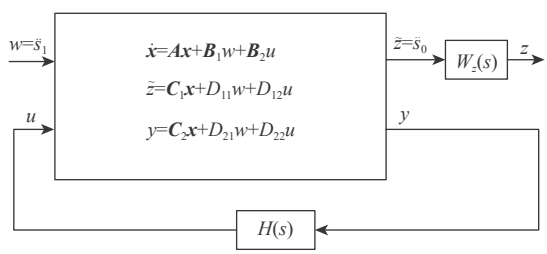


图 3　构建加权函数后的系统框图

Fig. 3　Block diagram of system with weighting function

对式(3)构建加权函数后的状态方程为

$$\begin{cases} \begin{bmatrix} \dot{\boldsymbol{x}} \\ \dot{x}_z \end{bmatrix} = \begin{bmatrix} \boldsymbol{A} & 0 \\ \boldsymbol{B}_z\boldsymbol{C}_1 & A_z \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ x_z \end{bmatrix} + \begin{bmatrix} \boldsymbol{B}_1 \\ \boldsymbol{B}_z D_{11} \end{bmatrix} w \\ \qquad + \begin{bmatrix} \boldsymbol{B}_2 \\ \boldsymbol{B}_z D_{12} \end{bmatrix} u \\ z = \begin{bmatrix} D_z\boldsymbol{C}_1, & C_z \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ x_z \end{bmatrix} + D_z D_{11} w + D_z D_{12} u \\ y = \begin{bmatrix} \boldsymbol{C}_2, & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ x_z \end{bmatrix} + D_{21} w + D_{22} u \end{cases} \tag{4}$$

式中: $\boldsymbol{C}_1 = \begin{bmatrix} k/M, & 0, & -T/M \end{bmatrix}$; $D_{11}=0$; $D_{12}=0$;

$W_z(s)$: $\begin{cases} \dot{x}_z = A_z x_z + B_z\tilde{z} \\ y_z = C_z x_z + D_z\tilde{z} \end{cases}$; $\tilde{z}=\ddot{s}_0$; $\dfrac{z(s)}{\tilde{z}(s)} = W_z(s)$; W_z 为加权函数; x_z 为矩阵增广变量; A_z、B_z、C_z、D_z 为增广后的系数矩阵一维参数。

主动隔振平台的工作特性需要高精度传递低频段的 CMG 输出力矩, 隔离高频段的 CMG 机械振动。

最终表征出的物理特性为：输入低频信号时，CMG端的输入与星体端输出保持一致，主动隔振平台体现跟踪性能；输入为高频信号时，CMG端的输入不影响星体端的输出，主动隔振平台体现干扰抑制效果。

2 单支腿主动隔振平台LPV模型设计

设计LPV模型首先需要确定变参数，当卫星姿态快速机动时，CMG外框架角速度变化较大，此时主动隔振平台应当确保CMG高精度的输出力矩传递；当卫星机动到位后的姿态稳定控制时，CMG外框架角速度较小，此时主动隔振平台应当确保CMG的各频段机械振动尽可能被隔离。为兼顾机动过程中CMG外框转速的大范围变化和机动到位后CMG外框转速的小范围波动，利用LPV变参数控制器设计方法，引入外框架角速度 ω_{out} 为变参数，根据不同的外框架角速度调整不同带宽的加权函数。以实现主动隔振平台同时具备自适应变带宽和较宽频段内对CMG振动高效隔离的能力。

本文设计方案主要考察控制器对单支腿隔振的动态性能和稳态精度的影响，控制器设计目标主要考虑阶跃输入下稳态调节时间 t_s 及对一定频率 f_c 的高频输入信号振动衰减率 σ。

引入变参数 ω_{out}（角速度变化范围 $[\omega_{\min}, \omega_{\max}]$ (°)/s）的加权函数频率特性如图4所示，设计闭环系统带宽 f_0 为开环系统特征频率的1/10，高频衰减转折频率为开环系统特征频率的10倍。因此 f_0=0.1 Hz，常数系数 α=100，f_1 为添加变参数后的系统带宽，表达式为

$$f_1 = f_0 + 5\omega_{\text{out}} \tag{5}$$

f_2 为添加变参数后的高频衰减转折频率，表达式为

$$f_2 = \alpha f_1 \tag{6}$$

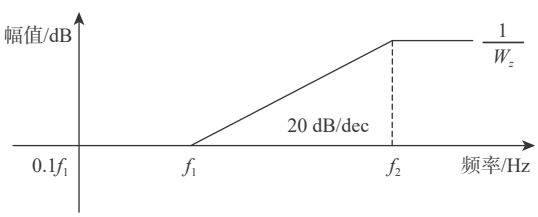


图4 ω_{out} 的加权函数频率特性

Fig. 4 Weighting function of ω_{out} frequency characteristics

由图4频率特性可以得到加权函数为

$$W_z(s) = \frac{1 + 1/2\pi\,(f_0 + 5\omega_{\text{out}})s}{1 + 1/2\pi\alpha(f_0 + 5\omega_{\text{out}})s} = \frac{\alpha s + 2\pi\alpha(f_0 + 5\omega_{\text{out}})}{s + 2\pi\alpha(f_0 + 5\omega_{\text{out}})} \tag{7}$$

式中：s 为传递函数拉氏空间下的微分算子。式(7)化简后整理可得

$$W_z(s) = \alpha + \frac{2\pi\alpha(f_0 + 5\omega_{\text{out}})(1-\alpha)}{s + 2\pi\alpha(f_0 + 5\omega_{\text{out}})} = \alpha + \frac{c_3 + c_4\omega_{\text{out}}}{s + (c_1 + c_2\omega_{\text{out}})} \tag{8}$$

式中：$c_1 = 2\pi\alpha f_0, c_2 = 10\pi\alpha, c_3 = 2\pi\alpha(1-\alpha)f_0, c_4 = 10\pi\alpha\times(1-\alpha)$，均为常数。

整理状态空间可得

$$\begin{cases} \dot{x}_z = A_z\omega_{\text{out}}x_z + B_z\tilde{z} \\ y_z = C_z\omega_{\text{out}}x_z + D_z\tilde{z} \end{cases} \Rightarrow \begin{cases} \dot{x}_z = -(c_1 + c_2\omega_{\text{out}})x_z + \tilde{z} \\ y_z = (c_3 + c_4\omega_{\text{out}})x_z + \alpha\tilde{z} \end{cases} \tag{9}$$

因此，加权函数状态空间模型中的状态转移参数 A_z 和观测参数 C_z 均为外框旋转角速度 ω_{out} 的一阶线性函数，从而得到用于控制器设计的线性变参数系统状态方程形式如式(4)所示，并且相关矩阵包含变参数 ω_{out}。

式(4)中w为外部扰动输入，本节选取的是CMG端加速度为输入；式(4)中 z 为被调输出，本节选取的是通过权函数的支腿加速度为输出。依据文献[9]和文献[6]中的LMI求解和极点配置方法，考虑设计一个状态反馈 控制器 u=$\boldsymbol{K}\boldsymbol{x}_k$，$\boldsymbol{K}$ 为控制器矩阵，$\boldsymbol{x}_k = \left[\boldsymbol{x}^{\mathrm{T}}, x_z\right]^{\mathrm{T}}$。

使 w 到 z 的闭环系统为

$$\begin{cases} \begin{bmatrix} \dot{\boldsymbol{x}} \\ \dot{x}_z \end{bmatrix} = (\overline{\boldsymbol{A}} + \overline{\boldsymbol{B}}_2\boldsymbol{K}) \begin{bmatrix} \boldsymbol{x} \\ x_z \end{bmatrix} + \overline{\boldsymbol{B}}_1 w \\ z = (\overline{\boldsymbol{C}}_1 + \overline{D}_{12}\boldsymbol{K}) \begin{bmatrix} \boldsymbol{x} \\ x_z \end{bmatrix} + \overline{D}_{11} w \end{cases} \tag{10}$$

式中：

$$\begin{cases} (\overline{\boldsymbol{A}} + \overline{\boldsymbol{B}}_2\boldsymbol{K}) = \begin{bmatrix} \boldsymbol{A} & 0 \\ B_z\boldsymbol{C}_1 & A_z\omega_{\text{out}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{B}_2 \\ B_zD_{12} \end{bmatrix} \boldsymbol{K} \\ (\overline{\boldsymbol{C}}_1 + \overline{D}_{12}\boldsymbol{K}) = \left[D_z\boldsymbol{C}_1,\ C_z\omega_{\text{out}}\right] + D_zD_{12}\boldsymbol{K} \end{cases} \tag{11}$$

$$\begin{cases} \overline{\boldsymbol{B}}_1 = \begin{bmatrix} \boldsymbol{B}_1 \\ B_zD_{11} \end{bmatrix} \\ \overline{D}_{11} = D_zD_{11} \end{cases} \tag{12}$$

为保证系统渐进稳定，并且使系统闭环传递函数 $\|\boldsymbol{T}_{wz}\|_\infty < \gamma$，其中 γ 为范数的界，通常 $\gamma > 0$。通过求解LMI将极点配置到稳定区域，保证系统的性能。图5为极点配置区域，$0 < \theta < 90°$，$\beta > 0$。θ 约束保证极点处于左半平面，r 为极点配置范围约束，β为最小衰减率，最小阻尼比 $\xi = \cos\theta$ 和最大无阻尼自然频率 $\omega_{\text{d}} = r\sin\theta$。

根据本节相关状态反馈控制器设计的内容，参考文献[11]可知，$\gamma > 0$为保证矩阵中$\gamma\boldsymbol{Z}$为负，使矩阵满足<0的负定条件，存在正定对称矩阵 $\boldsymbol{X}$ 和矩阵 $\boldsymbol{W}$=$\boldsymbol{K}\boldsymbol{X}$，使得以下LMI组成立：

$$\begin{bmatrix} \boldsymbol{M}+\boldsymbol{M}^{\mathrm{T}} & \overline{\boldsymbol{B}}_1 & \boldsymbol{N}^{\mathrm{T}} \\ \overline{\boldsymbol{B}}_1^{\mathrm{T}} & -\gamma\boldsymbol{I} & \overline{D}_{11} \\ \boldsymbol{N} & \overline{D}_{11} & -\gamma\boldsymbol{I} \end{bmatrix} < \boldsymbol{0} \tag{13}$$

$$\boldsymbol{M}+\boldsymbol{M}^{\mathrm{T}}+2\beta\boldsymbol{X}<\boldsymbol{0} \tag{14}$$

$$\begin{bmatrix} -r\boldsymbol{X} & \boldsymbol{M} \\ \boldsymbol{M}^{\mathrm{T}} & -r\boldsymbol{X} \end{bmatrix}<\boldsymbol{0} \tag{15}$$

$$\begin{bmatrix} \sin\theta(\boldsymbol{M}+\boldsymbol{M}^{\mathrm{T}}) & \cos\theta(\boldsymbol{M}-\boldsymbol{M}^{\mathrm{T}})^{\mathrm{T}} \\ \cos\theta(\boldsymbol{M}-\boldsymbol{M}^{\mathrm{T}}) & \sin\theta(\boldsymbol{M}+\boldsymbol{M}^{\mathrm{T}}) \end{bmatrix}<\boldsymbol{0} \tag{16}$$

式中:

$\boldsymbol{M}=\overline{\boldsymbol{A}}\boldsymbol{X}+\overline{\boldsymbol{B}}_2\boldsymbol{W}$, $\boldsymbol{N}=\overline{\boldsymbol{C}}_1\boldsymbol{X}+\overline{\boldsymbol{D}}_{12}\boldsymbol{W}$。

参考文献[3]和文献[6]中的凸分解方法。将CMG外框架角速度 ω_{out} 的上下限($[\omega_{\min},\omega_{\max}]$ (°)/s)分别代入式(13)～式(16)进行求解,可以分别解出 $\omega_{\mathrm{out}}=\omega_{\max}$ 的状态反馈增益矩阵 $\boldsymbol{K}_{\max}$ 和 $\omega_{\mathrm{out}}=\omega_{\min}$ 的状态反馈增益矩阵 $\boldsymbol{K}_{\min}$, 根据凸分解的几何调度策略可得状态反馈增益矩阵和CMG外框架角速度之间的关系为

$$\boldsymbol{K}(\omega_{\mathrm{out}})=\frac{(\omega_{\max}-\omega_{\mathrm{out}})}{(\omega_{\max}-\omega_{\min})}\boldsymbol{K}_{\min}+\frac{(\omega_{\mathrm{out}}-\omega_{\min})}{(\omega_{\max}-\omega_{\min})}\boldsymbol{K}_{\max} \tag{17}$$

通过输入不同的 ω_{out} 可以获得与变参数相对应的状态反馈控制增益矩阵,随后即可利用状态反馈增益阵算出对应外框架角速度下的H∞控制器,实现变参数控制器的设计。

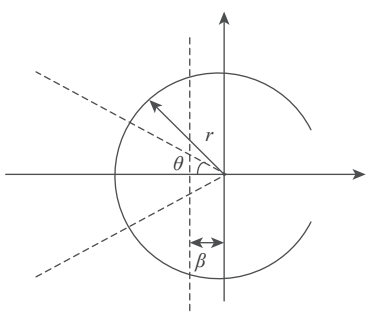


图 5　极点配置区域

Fig. 5　Pole placement area

3　仿真验证与结果分析

为验证本文LPV控制器及模型的性能及姿态快速机动的效果,同第2节中设计的PID控制器和H∞控制器与LPV相关控制器的控制效果进行比较。

根据极点配置方法[12],最终确定 β=1, r=6, $\theta=\pi/4$。利用极点配置计算方法得到的状态反馈增益矩阵为

$$\begin{cases}\boldsymbol{K}_{\max}=[-1\,190.66,\ 24.82,\ 9.97,\ 1\,776.81]\\ \boldsymbol{K}_{\min}=[-404.28,\ 24.58,\ 7.35,\ 256.79]\end{cases} \tag{18}$$

再将 $\boldsymbol{K}_{\max}$、$\boldsymbol{K}_{\min}$ 代入式(17)完成变参数控制器的设计。

第2节中得到的PID控制器为 $H(s)=k_{\mathrm{p}}+k_{\mathrm{d}}s/(as+1)+k_{\mathrm{i}}/s$, 依据Bode图中谐振峰出现在1.24 Hz处,确定系统带宽,通过多次调节控制器参数得到满足带宽和性能要求的控制器。其中PID控制器比例系数 k_{p}=30, 积分系数 k_{i}=7, 微分系数 k_{d}=0.3, 滤波系数 a=0.001。

依据第2节H∞控制器求解方法得到的H∞控制器为

$$K(s)=\frac{3.41(s+115.5)(s^2+76.32s+4\,986)}{(s+0.071\,5)(s^2+269.2s+38\,899.46)} \tag{19}$$

根据解算的控制器,分别绘制出各个控制器下闭环Bode图,如图6所示。由图可知,本文LPV控制器相较于其他控制器而言有更高的带宽,可以保证在CMG快速机动时,控制信号能够完全传递到星体端而不衰减,在该频段之后,本文LPV控制器又有一个快速的衰减频段,提升了对高频信号的隔振效率,保证了星体机动到位后能够保持高稳定度。

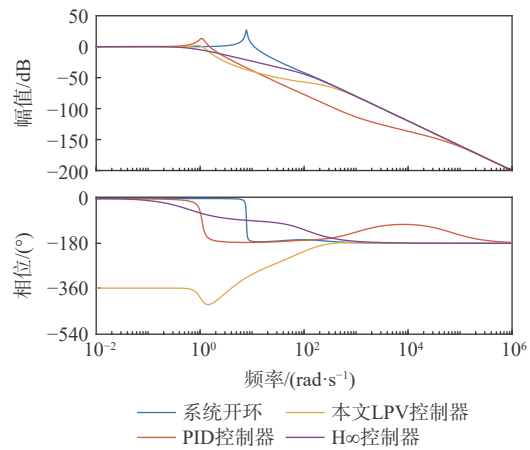


图 6　不同控制器的闭环 Bode 图

Fig. 6　Closed-loop Bode diagrams of different controllers

验证控制器动态性能,可以考虑时域响应结果。当外部输入端输入为不同信号时,PID、H∞和本文LPV控制器的控制效果图如图7～图9所示。

当输入为阶跃输入时,此时主要用来模拟快速机动后,姿态稳定状态的控制效果。

可知在3种控制器作用下,PID控制的调节时间大于H∞控制的调节时间,大于本文LPV控制的调节时间。因此,在本文LPV控制器作用下,卫星

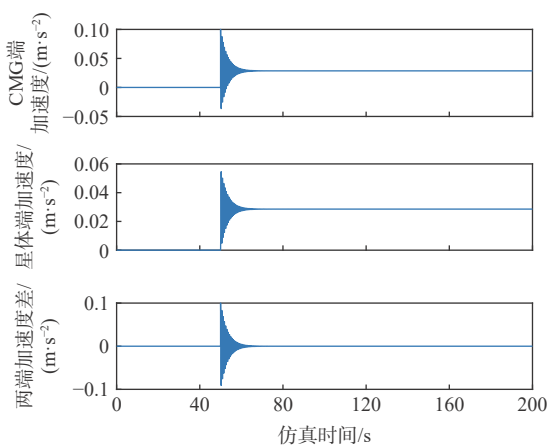


图 7　阶跃输入 PID 控制器

Fig. 7　PID controller of step input

快速机动后能够迅速进入稳态。

考虑高频干扰信号(f_c=25 Hz)输入，此时主要模拟机动到位后隔振平台的隔振效果，验证星体稳定度，隔振效果如图10所示。

表1为不同控制器的隔振性能比较。可知，3种控制器均对高频干扰能起到隔振作用，高频振动对星体端的干扰能抑制到5‰以下，且本文LPV控制器进入稳态所需要的时间比PID和H∞控制器要短，有利于执行后续快速机动的工况。

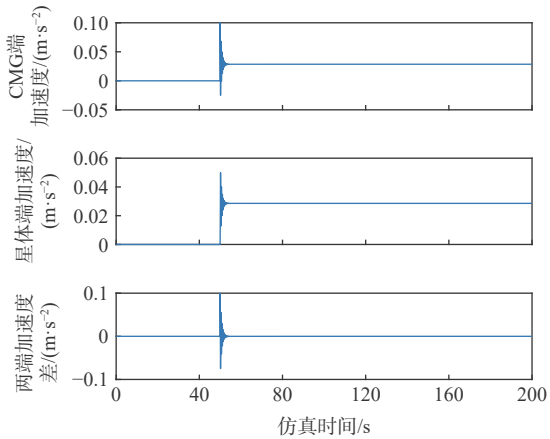


图8 阶跃输入H∞控制器

Fig. 8 H∞ controller of step input

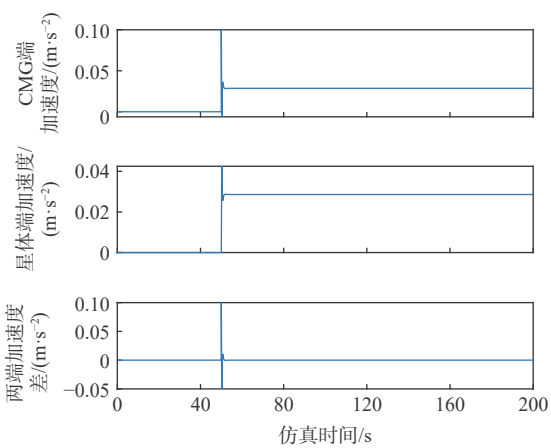


图9 阶跃输入本文LPV控制器

Fig. 9 LPV controller of step input

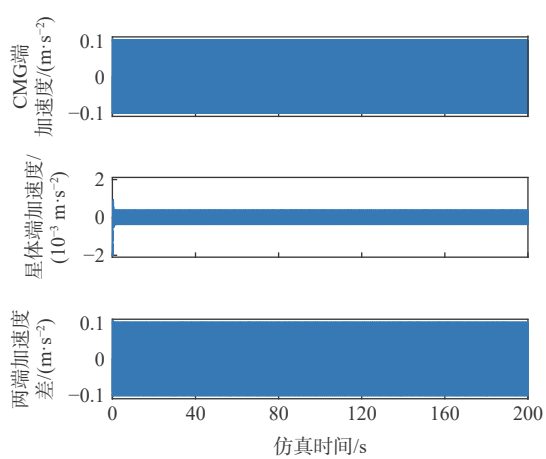


图10 高频干扰(f_c=25 Hz)输入本文LPV控制器

Fig. 10 LPV controller of high frequency interference (f_c=25 Hz) input

表1 不同控制器的隔振性能比较

Table 1 Comparison of vibration isolation performance with different controllers

控制器类型(f_c=25 Hz)	CMG端加速度/(m·s^{-2})	星体端加速度/(m·s^{-2})	稳态调节时间t_s/s	振动衰减率σ/%
PID	0.100 4	0.000 18	9.360	99.82
H∞	0.100 8	0.000 35	4.330	99.65
本文LPV	0.100 9	0.000 36	2.448	99.64

因此，本文LPV控制器相较于PID控制器和H∞控制器，具有良好的参数适应性，能够够实现预期的变带宽效果，也可以保证星体在快速机动情况下，主动隔振平台能够准确地将高频控制信号从CMG端传输到星体端；在机动到位后也能快速稳定，隔离高频信号的同时提升星体的姿态稳定度。

4 结 论

1）通过极点配置的方法调整系统性能，利用凸分解方法确定了各个顶点处控制器与变参数的线性关系，最终得到以CMG外框架角转速为变参数的LPV鲁棒控制器。

2）本文LPV控制器与PID和H∞控制器相比有着更好的动态性能和隔振效果，能够满足不同工况的系统性能。

本文采用的凸分解几何调度策略计算量相对较小，方法简单，也存在一定的保守性，如果要求更精确的方法可以采用系统广义距离调度策略或者高阶奇异值分解方法来确定变参数关系。

参考文献（References）

[1] 张尧, 张景瑞. 隔振平台对姿态控制系统影响分析及参数选择[J]. 宇航学报, 2013, 34(5): 657-664.
ZHANG Y, ZHANG J R. Analysis of influence of vibration isolation platform on attitude control system and its parameter selection[J]. Journal of Astronautics, 2013, 34(5): 657-664(in Chinese).

[2] 焦健. 柔性连接主动隔振平台的动力学建模与试验研究[D]. 哈尔滨: 哈尔滨工业大学, 2018.
JIAO J. Dynamic modeling and experimental research on active vibration isolation platform with flexible connection[D]. Harbin: Harbin Institute of Technology, 2018 (in Chinese).

[3] 李文强. LPV系统鲁棒变增益控制研究及其应用[D]. 长沙: 国防科学技术大学, 2009.
LI W Q. Research and application of robust gain-scheduling based on LPV system[D]. Changsha: National University of Defense Technology, 2009 (in Chinese).

[4] APKARIAN P, GAHINET P. A convex characterization of gain-

scheduled H∞ controllers[J]. IEEE Transactions on Automatic Control, 1995, 40(5): 853-864.

[5] APKARIAN P, ADAMS R J. Advanced gain-scheduling techniques for uncertain systems[J]. IEEE Transactions on Control Systems Technology, 1998, 6(1): 21-32.

[6] 虞忠伟, 胡东, 陈辉堂. 基于LPV方法的机器人鲁棒H∞控制[C]//第十九届中国控制会议. 北京: 中国自动化学会控制理论专业委员会, 2000: 715-719.
YU Z W, HU D, CHEN H T. Robust H∞ control of robot based on LPV method[C]//Proceedings of the 19th Chinese Control Conference. Beijing: TCCT Press, 2000: 715-719(in Chinese).

[7] 虞忠伟, 陈辉堂. 机器人多胞变增益输出反馈H∞控制[J]. 控制理论与应用, 2003, 20(6): 925-932.
YU Z W, CHEN H T. Polytopic gain scheduling output feedback H∞control for robotic manipulators[J]. Control Theory & Applications, 2003, 20(6): 925-932(in Chinese).

[8] 孙斌, 杨凌宇, 张晶. 基于高阶奇异值分解的LPV鲁棒控制器设计[J]. 北京航空航天大学学报, 2016, 42(7): 1536-1542.
SUN B, YANG L Y, ZHANG J. Robust LPV control design based on HOSVD[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1536-1542.

[9] GAHINET P, APKARIAN P. A linear matrix inequality approach to H∞ control[J]. International Journal of Robust and Nonlinear Control, 1994, 4(4): 421-448.

[10] 俞立. 鲁棒控制—线性矩阵不等式处理方法[M]. 北京: 清华大学出版社, 2002: 20-40.
YU L. Robust control linear matrix inequality processing method[M]. Beijing: Tsinghua University Press, 2002: 20-40(in Chinese).

[11] DE MARNEFFE B, AVRAAM M, DERAEMAEKER A, et al. Vibration isolation of precision payloads: A six-axis electromagnetic relaxation isolator[J]. Journal of Guidance, Control, and Dynamics, 2009, 32(2): 395-401.

[12] CHILALI M, GAHINET P. H∞ design with pole placement constraints: An LMI approach[J]. IEEE Transactions on Automatic Control, 1996, 41(3): 358-367.

[13] 胡艳梅. 线性参变系统的有限时间控制及其应用[D]. 哈尔滨: 哈尔滨工业大学, 2020.
HU Y M. Finite-time control for linear parameter-varying system and its application[D]. Harbin: Harbin Institute of Technology, 2020 (in Chinese).

[14] 乔盼盼. 一类LPV系统的最优控制数值解问题[D]. 大连: 大连理工大学, 2020.
QIAO P P. Numerical solution of optimal control for a class of LPV systems[D]. Dalian: Dalian University of Technology, 2020 (in Chinese).

[15] 温暖, 刘正华, 祝令谱, 等. 非仿射参数依赖LPV模型的变体飞行器H_∞控制[J]. 北京航空航天大学学报, 2017, 43(10): 2073-2080.
WEN N, LIU Z H, ZHU L P, et al. H∞ control for morphing aircraft via non-affine parameter dependent LPV model[J]. Journal of Beijing University of Aeronautics and Astronautics, 2017, 43(10): 2073-2080(in Chinese).

Design of robust controller for single outrigger of vibration active isolation platform based on LPV

TANG Yucong，ZHU Qinghua*，LIU Fucheng，CAO Heyang

(Shanghai Aerospace Control Technology Institute，Shanghai 201109，China)

Abstract：When the control moment gyro (CMG) is used to achieve agile attitude maneuvering control, the mechanical CMG vibration should be effectively isolated after the maneuvering is in place to achieve high-stability attitude control, while the rapid response of the control output torque is ensured during the maneuvering. This paper presents a design of outrigger controller for CMG vibration isolation platform based on linear parameter varying (LPV), setting the angular velocity of the CMG outer frame as a variable parameter. A comparison with other active vibration isolation methods demonstrates better performance of the design in different mechanical transmission requirements for the vibration isolation platform during agile attitude maneuvering and after the maneuvering in place.

Keywords：control moment gyro; active vibration isolation; attitude maneuver; linear parameter varying; robust control

Received：2021-09-05；**Accepted**：2021-11-26；**Published Online**：2021-12-16 12:21

URL：kns.cnki.net/kcms/detail/11.2625.V.20211215.1227.001.html

* **Corresponding author**. E-mail：zhuqh_79@sina.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0541

基于NARX和Kriging的时变可靠性分析双层代理模型

常泽明，李璐祎*
(西北工业大学 航空学院，西安 710072)

摘 要：针对具有动态输出性能的结构系统，传统的求解时变可靠性的代理模型方法在建模时只关注当前瞬间作用于系统的随机变量的作用，而忽视了时间累积效果，使得模型对于时变可靠性的预测效果并不理想。基于此，提出了一种基于带外生输入的非线性自回归(NARX)模型和Kriging模型的时变可靠性分析的双层代理模型方法。所提方法在内层利用NARX模型构建给定随机输入变量下输出响应随时间的变化模型，准确模拟系统的动态行为；在外层基于NARX所得极值构建系统极值与随机变量之间的Kriging模型，得到时变结构系统的可靠性。通过3个算例验证了所提方法在处理具有较强波动性输出系统的可靠性问题时的有效性和准确性。

关 键 词：时变可靠性；失效概率；代理模型；NARX模型；Kriging模型
中图分类号：TB114.3
文献标志码：A **文章编号**：1001-5965(2023)07-1802-11

不确定性广泛存在于工程结构的各个方面，如材料参数、测量误差、几何特性等。这些不确定因素对工程结构能否正常工作将会产生影响。通过分析考虑不确定性因素，可靠性分析能够量化结构系统的安全程度及其性能的有效性，从而使得工程人员能够充分掌握系统的安全性及性能，进而完成更完善的设计工作。目前，针对具有静态输出性能的结构系统已经发展出相对完备的系统可靠性分析方法[1-3]、基于可靠性的优化设计[4-6]等诸多方法。但是，实际上随着时间的推移，结构自身的材料性能、所受载荷效应、环境条件等属性并非恒定不变的，受此影响，其性能往往呈现出时变特性。因此，对不确定性结构进行时变可靠性分析具有重要的意义。

目前，常见的时变可靠性分析方法分为3类：基于跨越率的方法[7-13]、基于极值的方法[14-18]和基于代理模型的方法[19-23]。基于跨越率的方法源于20世纪40年代，Rice提出了首次超越概率计算公式[7]。随后，Crandall等[8]将数值模拟方法引入到首次超越问题中，使得基于首次超越模型的动态可靠性分析理论更加完善。在此基础上，众多研究者发展出了可微高斯过程法[9]、矩形波更新过程法[10]、拉普拉斯积分法[11]、PHI2方法[12-13]等众多基于跨越率的时变可靠性分析方法。基于极值的方法源于Hu和Du[14]提出的一种针对极值分布的抽样方法，其利用输出的极值信息来获得结构的时变可靠性，从而把时变可靠性问题转换为时不变问题。Li等[15-16]使用等效极值事件和极值分布来估计时变可靠性。通过将时变结构的性能函数线性化，Du[17]提出了时变机构可靠性分析的包络函数法。对于时空多参数时变结构，石岩等[18]提出了基于包络函数的时变可靠性分析方法。但在实际中，获得输出的极值信息往往需要反复调用功能函数才能实现，

收稿日期：2021-09-09；录用日期：2021-12-10；网络出版时间：2022-01-18 19:48
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220118.1321.003.html
基金项目：国家自然科学基金(51875464)
*通信作者. E-mail：luyili@nwpu.edu.cn

引用格式：常泽明，李璐祎. 基于NARX和Kriging的时变可靠性分析双层代理模型[J]. 北京航空航天大学学报，2023，49(7)：1802-1812. CHANG Z M，LI L Y. Double-loop surrogate model for time-dependent reliability analysis based on NARX and Kriging models[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1802-1812 (in Chinese).

尤其对于复杂工程结构, 反复调用有限元模型的成本极高。相比之下, 基于代理模型的方法显得优势更为明显。基于代理模型的方法通过构造显式功能函数来逼近真实模型, 进而通过调用显式功能函数来估计时变可靠性。Cheng 和 Lu[19] 针对高维问题提出了一种高维模型表示 (hight dimensional model representation, HDMR) 方法。Hawchar 等[20] 利用主成分分析和多项式混沌展开来逼近极限状态面。Hu 等[21] 将快速积分法和代理模型法相结合, 提出了一种有效的时变可靠性分析方法。Wang 等[22] 提出了嵌套极值响应面法, 采用双层替代法来构造极值替代模型。Hu 和 Du[23] 改进了嵌套极值响应面法, 发展了一种混合有效的全局优化技术来识别极值, 降低了计算成本。上述基于代理模型的时变可靠性分析方法通过一个双循环过程来构建替代模型, 其中, 外部循环构建输入变量与输出极值间的替代模型, 内部循环在给定输入情况下估计输出极值。需要注意的是, 这种双循环模型存在一定的缺点: ①内层极值估计的精度会影响外层模型的精度; ②较长时间范围内识别结构的极值在计算上代价昂贵。

目前, 常用的基于 Kriging 模型的双层代理模型方法同样存在上述问题。这些方法在构建代理模型的过程中, 由于只关注于某一确定时刻随机变量对于系统输出的影响, 而忽视了时间推移过程中的时间效应, 使得建立的代理模型对于时变可靠性的预测效果并不理想。尤其是对具有波动性输出的结构, 基于 Kriging 模型的双层代理模型方法甚至无法得到收敛的解。针对这一问题, 本文将能够良好描述时序行为的带有外生输入的非线性自回归 (nonlinear autoregressive with exogenous inputs, NARX) 模型引入到双层代理模型方法中, 提出了一种基于 NARX 模型和 Kriging 模型的时变可靠性分析方法。该方法在内层通过 NARX 模型来模拟给定输入变量下系统沿时间方向的动态行为, 进而得到输出的极值。在外层利用 Kriging 模型建立输出极值与输入不确定性变量之间的替代模型, 进而得到结构系统的时变可靠性。由于内层 NARX 模型考虑时间的累积效应, 将某一时刻的输出量用系统的输入输出及输入输出的延迟项来表示[24], 能够准确描述结构系统沿时间方向的动态变化, 本文方法对具有波动性输出的系统同样适用。算例结果表明, 与已有方法相比, 本文方法具有更广泛的适用范围, 且能够以小的计算成本和较高的精度得到结构系统时变可靠性。

1　NARX 模型

1.1　NARX 模型的表现形式

对于一个计算模型$M(\psi(t))$, $\psi(t)$表示与时间相关的输入, 目的是利用输入和输出的实验数据建立描述M的数学模型, 因此, 将感兴趣的时间区间$[T_1,T_2]$以时间间隔d离散为一系列时刻$t_k(k=1,2,\cdots,N_T)$。NARX 模型将一个给定时刻的输出表示为过去输入和输出在当前或之前时刻的值的函数[25], 即

$$\begin{aligned} y_{\mathrm{N}}(t_k)=&\Gamma(\boldsymbol{z}(t_k))+\varepsilon_t(t_k)=\\ &\Gamma\left(\psi(t_k),\cdots,\psi\left(t_k-n_\psi d\right),y_{\mathrm{N}}(t_k-d),\cdots,\right.\\ &\left.y_{\mathrm{N}}\left(t_k-n_y d\right)\right)+\varepsilon_t(t_k) \end{aligned} \tag{1}$$

式中: $\Gamma(\boldsymbol{z}(t_k))$为待确定的替代模型, $\boldsymbol{z}(t_k)=[\psi(t_k),\cdots,\psi\left(t_k-n_\psi d\right),y_{\mathrm{N}}(t_k-d),\cdots,y_{\mathrm{N}}\left(t_k-n_y d\right)]^{\mathrm{T}}$为包含当前和过去输入输出值的向量; n_ψ和n_y分别为最大输入和输出时间延迟; $\varepsilon_t(t_k)$为 NARX 模型的残差, 在标准 NARX 模型中, 假设残差服从均值为 0、方差为σ_ε^2的标准正态分布。

对于所要构造的模型, 其函数形式有多种选择, 常用的线性形式表示为

$$y_{\mathrm{N}}(t_k)=\sum_{i=1}^{n_g}\theta_i g_i(\boldsymbol{z}(t_k))+\varepsilon_t(t_k) \tag{2}$$

式中: $g_i(\boldsymbol{z}(t_k))$为函数模型项, 其作用是将原始数据映射到高维特征空间; n_g为模型项个数; θ_i为 NARX 模型系数。

由式(2)可以看出, NARX 模型遵循因果关系原则来描述系统的动态行为, 即系统当前状态$y_{\mathrm{N}}(t)$受其先前状态$\left[y_{\mathrm{N}}(t-d),\cdots,y_{\mathrm{N}}\left(t-n_y d\right)\right]$及当前和过去输入$\left[\psi(t),\cdots,\psi\left(t-n_\psi d\right)\right]$的影响。同时, 应注意因果效应会随着时间的推移而逐渐消失, 故为了精简模型, 只需考虑当前时间瞬间之前有限时间延迟就足够了。

建立 NARX 模型有助于为分析人员揭示系统的机制和行为, 了解一个系统是如何运行的, 从而更好地对其控制。另外, 所确定的模型可用于预测系统未来的输出。从这个角度看, 其可以被认为是原始模型$M(\psi(t))$的近似模型, 只要所划分的时间间隙足够小, 本文可以利用其来进行整个时间区间内的输出预测。

1.2　构建 NARX 模型

建立 NARX 模型包含 2 个主要步骤: ①模型结构选择, 即确定函数模型项$g_i(\boldsymbol{z}(t))$有哪些。值得注意的是, 模型项$g_i(\boldsymbol{z}(t))$可以由各种全局或局部基函数构造, 如$g_i(\boldsymbol{z}(t))$为多项式函数形式的 NARX 模型

在文献中比较常见。此外，样条函数、神经网络、小波函数等也可以作为基函数参与模型构造。②参数估计，即确定NARX模型系数。如式(2)所示对于线性形式表达的NARX模型，可将其看作一个线性回归问题，故可以使用处理线性回归分析问题的方法来解决上述问题。

在建立模型的过程中应注意结构的选择，特别是对于非线性系统，这是非常重要和困难的。在模型中包含冗余项会导致过拟合等问题，Billings建议利用最简单的模型来表示系统的原始动态过程，这样就可以通过正交最小二乘法及其导数来实现整个过程[26]。不同的结构选择方法包括试错法[27]、基于相关关系的方法[28](正交最小二乘法、正交前向回归法、最小角回归法等)等。其中，最小角回归(least angle regression, LARS)法[29]作为基于相关关系方法的一种，既保留了前向梯度算法一定程度的精确性，又简化了前向梯度算法逐步迭代的过程，并且该算法的最坏计算复杂度和最小二乘法类似，适用于线性回归问题。因此，本文采用LARS法对NARX基函数进行选择，利用最小二乘法求解系数，通过对模型的更新，得到所需NARX模型。

1.2.1 结构选择

建立NARX模型，首先需要明确NARX模型的基本形式，如函数模型项类型(多项式和神经网络等)、输入和输出的最大时间延迟、基函数的整体属性(如最大多项式阶)。建议从简单模型开始考虑，充分利用关于系统的现有信息，如系统自由度的数目、非线性行为的类型等，以使所建立的NARX模型能够充分地描述所研究的系统。本文选用多项式形式，输入输出最大时间延迟取自由度的2倍，基函数整体属性由经验或已知的功能函数确定，通过LARS法确定出合适的函数模型项用于模型构造。

1.2.2 NARX系数的求解

对于在$[T_1,T_2]$时间段内工作的动态系统，取时间间隔d，将时间区间均匀离散为N_T个时刻$t_k(k=1,2,\cdots,N_T)$。考虑到对于同一系统而言，系统输入输出关系是普遍存在且同源的，因此本文在建立NARX模型求解系数时，并不需要所有时刻的输入输出值来完成，而是以t_1时刻为起点，从离散的N_T个时刻点中均匀选取$N_{\rm F}(N_{\rm F}\ll N_T)$个时刻作为建模时刻点，建立样本时刻向量$\boldsymbol{T}_{\rm F}=\left[t_j^{({\rm F})}\right](j=1,\cdots,N_{\rm F})$。则$t_j^{({\rm F})}$时刻输出的提前一步预测可以表示为

$$\hat{y}_{\rm N}\left(t_j^{({\rm F})}\right)=\sum_{i=1}^{n_g}\theta_i g_i\left(\boldsymbol{z}\left(t_j^{({\rm F})}\right)\right) \tag{3}$$

式中：$\boldsymbol{z}\left(t_j^{({\rm F})}\right)=\left[\psi\left(t_j^{({\rm F})}\right),\cdots,\psi\left(t_j^{({\rm F})}-n_\psi d\right),y_{\rm N}\left(t_j^{({\rm F})}-d\right),\cdots,y_{\rm N}\left(t_j^{({\rm F})}-n_y d\right)\right]^{\rm T}$；模型项$g_i(\boldsymbol{z}(t))$由LARS法选择得到，$y_{\rm N}(t)$表示用于训练NARX模型的已知样本数据。

令$\boldsymbol{\phi}\left(t_j^{({\rm F})}\right)=\left[g_i\left(\boldsymbol{z}\left(t_j^{({\rm F})}\right),i=1,2,\cdots,n_g\right)\right]^{\rm T}$ $\boldsymbol{\theta}=[\theta_i,i=1,2,\cdots,n_g]^{\rm T}$，则模型残差可表示为

$$\varepsilon_t\left(t_j^{({\rm F})}\right)=y_{\rm N}\left(t_j^{({\rm F})}\right)-\hat{y}_{\rm N}\left(t_j^{({\rm F})}\right)=y_{\rm N}\left(t_j^{({\rm F})}\right)-\boldsymbol{\phi}^{\rm T}\left(t_j^{({\rm F})}\right)\boldsymbol{\theta} \tag{4}$$

因此，误差平方和为

$$\sum_{j=1}^{N_{\rm F}}\varepsilon_t^2\left(t_j^{({\rm F})}\right)=\sum_{j=1}^{N_{\rm F}}\left(y_{\rm N}\left(t_j^{({\rm F})}\right)-\boldsymbol{\phi}^{\rm T}\left(t_j^{({\rm F})}\right)\boldsymbol{\theta}\right)^2 \tag{5}$$

则所有训练时刻的输出可表示为

$$\begin{bmatrix}y_{\rm N}\left(t_1^{({\rm F})}\right)\\ \vdots\\ y_{\rm N}\left(t_{N_{\rm F}}^{({\rm F})}\right)\end{bmatrix}=\begin{bmatrix}\boldsymbol{\phi}^{\rm T}\left(t_1^{({\rm F})}\right)\\ \vdots\\ \boldsymbol{\phi}^{\rm T}\left(t_{N_{\rm F}}^{({\rm F})}\right)\end{bmatrix}\boldsymbol{\theta}+\begin{bmatrix}\varepsilon_t\left(t_1^{({\rm F})}\right)\\ \vdots\\ \varepsilon_t\left(t_{N_{\rm F}}^{({\rm F})}\right)\end{bmatrix} \tag{6}$$

方程(6)可以用矩阵符号重写为

$$\boldsymbol{y}_{\rm N}=\boldsymbol{\Phi}\boldsymbol{\theta}+\boldsymbol{\varepsilon}_t \tag{7}$$

式中：$\boldsymbol{y}_{\rm N}$为$N_{\rm F}\times 1$维向量，表示各时刻输出值；$\boldsymbol{\Phi}$为$N_{\rm F}\times n_g$维矩阵，其第j行表示NARX模型项$\boldsymbol{\phi}\left(t_j^{({\rm F})}\right)=\left[g_i\left(\boldsymbol{z}\left(t_j^{({\rm F})}\right),i=1,2,\cdots,n_g\right)\right]^{\rm T}$在瞬时$t=t_j^{({\rm F})}$的取值；$\boldsymbol{\varepsilon}_t$为$N_{\rm F}\times 1$维残差向量。

可以发现，式(7)是一个典型的线性回归问题，因此可以用最小二乘法计算其系数，即求使总误差最小的最小二乘解：

$$\boldsymbol{\theta}=\arg\min_{\boldsymbol{\theta}}\left(\boldsymbol{\varepsilon}^{\rm T}\boldsymbol{\varepsilon}\right)=\left[\boldsymbol{\Phi}^{\rm T}\boldsymbol{\Phi}\right]^{-1}\boldsymbol{\Phi}^{\rm T}\boldsymbol{y}_{\rm N} \tag{8}$$

1.2.3 NARX模型的预测和更新

初步建立NARX模型后，可将其用于预测时间间隔为d、时间区间为$[T_1,T_2]$的系统输出，即系统在离散的N_T个时刻$t_k(k=1,2,\cdots,N_T)$的所有输出。NARX模型的计算是递归重构的，即用其在某一时刻的输出估计值预测稍后时刻的输出，因此理论上只需给定初始输入$\boldsymbol{\psi}_0=\left[\psi(t_0),\psi(t_0-d),\cdots,\psi\left(t_0-n_\psi d\right)\right]$和初始输出$\boldsymbol{y}_0=\left[y_{\rm N}(t_0-d),\cdots,y_{\rm N}\left(t_0-n_y d\right)\right]$，则可递推得到时间区间内全部离散时刻的系统输出，即

$$\hat{y}_{\rm N}(t)=\sum_{i=1}^{n_g}\theta_i g_i(\hat{\boldsymbol{z}}(t)) \tag{9}$$

式中：t为离散的时刻$t_k(k=1,2,\cdots,N_T)$；向量$\hat{\boldsymbol{z}}(t)=\left[\psi(t),\cdots,\psi\left(t-n_\psi d\right),\hat{y}_{\rm N}(t-d),\cdots,\hat{y}_{\rm N}\left(t-n_y d\right)\right]^{\rm T}$中$\left[\hat{y}_{\rm N}(t-d),\cdots,\hat{y}_{\rm N}\left(t-n_y d\right)\right]$项为NARX模型估计得到的先前时刻的输出值。

系统输出的预测是通过递推得到的，因此会存在误差积累现象。为了减少累积误差，提高模型运算精度，本文充分利用NARX建模阶段已知训练样本信息，即$\left[y_{\rm N}(t-d),\cdots,y_{\rm N}\left(t-n_y d\right)\right]$，$t\in\boldsymbol{T}_{\rm F}$，在模型预测过程中，在时间区间$[T_1,T_2]$内进行分段预测，

从而实现对结果的修正。该过程在每个小时间区间$[t_j^{(\mathrm{F})}, t_{j+1}^{(\mathrm{F})}]$内, 以训练输入样本$\boldsymbol{\psi}_j=[\psi(t_j^{(\mathrm{F})}), \psi(t_j^{(\mathrm{F})}-d),\cdots,\psi(t_j^{(\mathrm{F})}-n_\psi d)]$和相应的输出$\boldsymbol{y}_j=[y_{\mathrm{N}}(t_j^{(\mathrm{F})}-d),\cdots y_{\mathrm{N}}(t_j^{(\mathrm{F})}-n_y d)]$为初始条件, 递推预测系统在该时间区间内的输出, 最终得到区间$[T_1,T_2]$内的全部输出。

由于预测结果经递推得到, 则在每一段小的时间区间$[t_j^{(\mathrm{F})}, t_{j+1}^{(\mathrm{F})}]$中, 要使末时刻预测值足够准确, 需前面所有时刻预测也足够精确。因此, 本文将该模型的误差定义为所有训练样本时刻$t_j^{(\mathrm{F})}(j=1,\cdots,N_{\mathrm{F}})$输出的原始值和预测值的差异, 即

$$\varepsilon_{\mathrm{N}}=\frac{\sum_{j=1}^{N_{\mathrm{F}}}\left(y_{\mathrm{N}}(t_j^{(\mathrm{F})})-\hat{y}_{\mathrm{N}}(t_j^{(\mathrm{F})})\right)^2}{\sum_{j=1}^{N_{\mathrm{F}}}\left(y_{\mathrm{N}}(t_j^{(\mathrm{F})})-\bar{y}_{\mathrm{N}}(t)\right)^2} \tag{10}$$

式中: $\hat{y}_{\mathrm{N}}(t_j^{(\mathrm{F})})$为 NARX 模型预测得到的训练样本时刻的输出; $y_{\mathrm{N}}(t_j^{(\mathrm{F})})$为训练样本时刻系统的真实输出; $\bar{y}_{\mathrm{N}}(t)=\sum_{j=1}^{N_{\mathrm{F}}}y_{\mathrm{N}}(t_j^{(\mathrm{F})})/N_{\mathrm{F}}$为真实输出均值。

将上述误差作为模型是否需要更新的判定标准, 若$\varepsilon_{\mathrm{N}}\leqslant 10^{-2}$, 则认为模型在各个离散时刻都能较为准确地描述系统的输出, 结束训练过程, 得到最终的 NARX 模型及系统在时间区间$[T_1,T_2]$内离散时刻点$t_k(k=1,2,\cdots,N_T)$的全部输出预测值; 若$\varepsilon_{\mathrm{N}}>10^{-2}$, 需对模型进行更新, 为了有效利用已有的样本点, 减少原始功能函数的调用, 需重新从离散的N_T个时刻点中均匀选取$2N_{\mathrm{F}}(N_{\mathrm{F}}\ll N_T)$个时刻(即保证所选择的建模时刻中包含此前的建模时刻)作为新的建模时刻点, 重复 1.2.2 节及 1.2.3 节中步骤, 更新 NARX 模型, 直到达到精度要求。

2　Kriging 模型

2.1　Kriging 代理模型的基本原理

Kriging 代理模型作为一种无偏估计模型, 通过将全局近似与局部随机误差相结合, 使其有效性不依赖于随机误差的存在, 从而能够良好地拟合具有较高非线性和局部响应突变的问题。

对于一个感兴趣的系统$Y=g(\boldsymbol{X})$, 其中, $\boldsymbol{X}=[X_1,X_2,\cdots,X_s]$为$s$维随机输入变量, 其 Kriging 模型描述成一个多项式和一个随机分布函数之和的形式[30], 表示为

$$Y\approx g_{\mathrm{K}}(\boldsymbol{X})=\sum_{j=1}^{p}f_j(\boldsymbol{X})\beta_j+z(\boldsymbol{X}) \tag{11}$$

式中: $g_{\mathrm{K}}(\boldsymbol{X})$为待定的 Kriging 模型; $\boldsymbol{f}(\boldsymbol{X})=[f_1(\boldsymbol{X}), f_2(\boldsymbol{X}),\cdots,f_p(\boldsymbol{X})]^{\mathrm{T}}$为随机变量$\boldsymbol{X}$的基函数, p为基函数的个数; $\boldsymbol{\beta}=[\beta_1,\beta_2,\cdots,\beta_p]^{\mathrm{T}}$为待定系数, 可通过训练样本进行估计; $z(\boldsymbol{X})$为表示 Kriging 模型局部偏差的一个随机过程, 其期望为 0, 方差为σ_{K}^2, 协方差矩阵分量为

$$\mathrm{Cov}\left[z(\boldsymbol{x}^{(v)}),z(\boldsymbol{x}^{(w)})\right]=\sigma_{\mathrm{K}}^2\left[R(\boldsymbol{x}^{(v)},\boldsymbol{x}^{(w)})\right] \tag{12}$$

式中: $R(\boldsymbol{x}^{(v)},\boldsymbol{x}^{(w)})$为样本点$\boldsymbol{x}^{(v)}$、$\boldsymbol{x}^{(w)}$的相关函数, $v,w=1,2,\cdots,m$, m为用于训练 Kriging 模型的样本个数。可供选择的$R(\boldsymbol{x}^{(v)},\boldsymbol{x}^{(w)})$函数类型很多, 其中, 高斯型函数表达式为

$$R(\boldsymbol{x}^{(v)},\boldsymbol{x}^{(w)})=\exp\left(-\sum_{r=1}^{s}\lambda_r\left|x_r^{(v)}-x_r^{(w)}\right|^2\right) \tag{13}$$

式中: $\lambda_r(r=1,2,,\cdots,s)$为待求的相关参数, s为变量维数。

由 Kriging 理论可知, 点$\boldsymbol{x}$处的输出估计值为

$$g_{\mathrm{K}}(\boldsymbol{x})=\boldsymbol{f}^{\mathrm{T}}(\boldsymbol{x})\hat{\boldsymbol{\beta}}+\boldsymbol{r}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{R}^{-1}(\boldsymbol{g}-\boldsymbol{F}\hat{\boldsymbol{\beta}}) \tag{14}$$

式中: $\hat{\boldsymbol{\beta}}$为$\boldsymbol{\beta}$的估计; $\boldsymbol{g}$为一个列向量, 由全部训练样本的输出值组成; $\boldsymbol{g}$为单位列向量; $\boldsymbol{F}$为p个基函数在全部m个训练样本点处的值组成的$m\times p$阶矩阵; $\boldsymbol{r}(\boldsymbol{x})$为预测点和训练样本点的相关函数向量, 可表达为

$$\boldsymbol{r}^{\mathrm{T}}(\boldsymbol{x})=\left[R(\boldsymbol{x},\boldsymbol{x}^{(1)}),R(\boldsymbol{x},\boldsymbol{x}^{(2)}),\cdots,R(\boldsymbol{x},\boldsymbol{x}^{(m)})\right]^{\mathrm{T}} \tag{15}$$

$\hat{\boldsymbol{\beta}}$表示为

$$\hat{\boldsymbol{\beta}}=\left[\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{F}\right]^{-1}\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{g} \tag{16}$$

方差估计值$\hat{\sigma}_{\mathrm{K}}^2$表示为

$$\hat{\sigma}_{\mathrm{K}}^2=(\boldsymbol{g}-\boldsymbol{F}\hat{\boldsymbol{\beta}})^{\mathrm{T}}\boldsymbol{R}^{-1}(\boldsymbol{g}-\boldsymbol{F}\hat{\boldsymbol{\beta}})/m \tag{17}$$

通过求极大似然估计的最大值可以得到相关系数$\boldsymbol{\lambda}=[\lambda_1,\lambda_2,\cdots,\lambda_s]^{\mathrm{T}}$, 即

$$\max F(\boldsymbol{\lambda})=-\frac{m\ln(\hat{\sigma}_{\mathrm{K}}^2)+\ln|\boldsymbol{R}|}{2}\qquad \lambda_r\geqslant 0(r=1,2,\cdots,s) \tag{18}$$

通过求解式 (18) 得到的$\boldsymbol{\lambda}$值用于建立 Kriging 模型, 此时可得到拟合精度最优的代理模型。

对于任意未知的$\boldsymbol{x}$, $g_{\mathrm{K}}(\boldsymbol{x})$服从高斯分布, 均值$\mu_{g_{\mathrm{K}}}(\boldsymbol{x})$及方差$\sigma_{g_{\mathrm{K}}}^2(\boldsymbol{x})$计算式为

$$\mu_{g_{\mathrm{K}}}(\boldsymbol{x})=\boldsymbol{f}^{\mathrm{T}}(\boldsymbol{x})\hat{\boldsymbol{\beta}}+\boldsymbol{r}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{R}^{-1}(\boldsymbol{g}-\boldsymbol{F}\hat{\boldsymbol{\beta}}) \tag{19}$$

$$\sigma_{g_{\mathrm{K}}}^2(\boldsymbol{x})=\sigma_{\mathrm{K}}^2\left[\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{r}(\boldsymbol{x})-\boldsymbol{f}(\boldsymbol{x})\right]^{\mathrm{T}}\left(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{F}\right)^{-1}\cdot\left[\boldsymbol{F}^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{r}(\boldsymbol{x})-\boldsymbol{f}(\boldsymbol{x})\right]+\sigma_{\mathrm{K}}^2\left(1-\boldsymbol{r}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{R}^{-1}\boldsymbol{r}(\boldsymbol{x})\right) \tag{20}$$

由 MATLAB 工具箱 DACE[31-32] 可以求得$\mu_{g_{\mathrm{K}}}(\boldsymbol{x})$和$\sigma_{g_{\mathrm{K}}}^2(\boldsymbol{x})$的值, 其中, $\sigma_{g_{\mathrm{K}}}^2(\boldsymbol{x})$表示$g_{\mathrm{K}}(\boldsymbol{x})$与$g(\boldsymbol{x})$之间的最小均方误差。Kriging 代理模型在训练样本点$\boldsymbol{x}^{(v)}$处, 有$\mu_{g_{\mathrm{K}}}(\boldsymbol{x}^{(v)})=g(\boldsymbol{x}^{(v)})$, $\sigma_{g_{\mathrm{K}}}^2(\boldsymbol{x}^{(v)})=0$, 而其他样

本的预测值的方差则一般不为0。$\sigma_{g_K}^2(\boldsymbol{x})$越大，说明$\boldsymbol{x}$处的预测值与实际值相差越悬殊，因此，可以通过$\sigma_{g_K}^2(\boldsymbol{x})$的预测值来衡量代理模型在$\boldsymbol{x}$位置处的预测精度。

2.2 基于U学习函数的自适应Kriging代理模型

对于感兴趣的系统$Y=g(\boldsymbol{X})$，在求解模型失效概率时，规定功能函数大于设定的阈值e即为失效，即$P_f=P\{g(\boldsymbol{x})>e\}$。因此，能否利用Kriging模型对系统的失效面进行准确估计对于失效概率的求解十分重要。

要建立足够准确的Kriging模型，需要对模型不断进行更新，由此衍生出了构建Kriging模型的自适应更新策略，其步骤如下：

步骤1 通过各种途径得到建立模型所需样本池$\boldsymbol{S}^x$。

步骤2 从样本池$\boldsymbol{S}^x$中随机抽取N_K^x个训练样本$\boldsymbol{x}_n$，并计算真实输出值，由此得到用于训练Kriging模型的初始训练集$\boldsymbol{K}^x=\{(\boldsymbol{x}_n,g(\boldsymbol{x}_n)),n=1,2,\cdots,N_K^x\}$，并建立初始的Kriging模型$G_K(\boldsymbol{x})$。

步骤3 利用自适应学习函数在样本池$\boldsymbol{S}^x$中挑选符合要求的样本点添加到此前训练样本集$\boldsymbol{K}^x$中，更新Kriging模型直到满足收敛条件。在这一过程中，添加到训练集用于更新Kriging模型的样本点需要满足：位于随机输入变量分布密度较大的区域；靠近极限状态面，在该点更容易误判符号。

步骤4 利用最优的Kriging代理模型$G_K(\boldsymbol{x})$实现响应预测。

本文基于U学习函数[33]来实现Kriging模型的自适应更新，U学习函数原理如下：

当系统的极限状态面为0时，即系统输出大于0时，系统失效，由Kriging模型的定义可知，$|\mu_{g_K}(\boldsymbol{x})|$较小，或当前Kriging模型预测方差$\sigma_{g_K}^2(\boldsymbol{x})$较大，或者以上2点均满足的点正负号误判风险更大。基于这一思想，Echard等[33]定义了一种U学习函数，表示为

$$U(\boldsymbol{x})=\left|\frac{\mu_{g_K}(\boldsymbol{x})}{\sigma_{g_K}(\boldsymbol{x})}\right| \tag{21}$$

U学习函数同时考虑了模型预测值距失效面的距离和标准差，靠近失效面($|\mu_{g_K}(\boldsymbol{x})|$较小)且估计值的标准差越大($\sigma_{g_K}^2(\boldsymbol{x})$较大)的点应加入到训练样本集更新Kriging模型，即样本池中U值最小的样本点应加入训练样本集中用于更新当前Kriging模型。文献[33]表明，通常情况下可以选择$\min U(\boldsymbol{x})\geqslant 2$作为基于$U$学习函数的Kriging代理模型自适应更新过程的收敛终止条件。

可以看到，U学习函数仅适用于极限状态面为0的系统，而对于一个阈值为e的系统$Y=g(\boldsymbol{X})$，当$g(\boldsymbol{X})>e$时，系统失效，其失效概率表示为

$$P_f=P\{g(\boldsymbol{x})>e\}=P\{g(\boldsymbol{x})-e>0\} \tag{22}$$

因此，在利用Kriging模型求解失效概率时，本文将不再建立$Y=g(\boldsymbol{X})$的Kriging模型，而是基于系统的功能函数及失效条件，构造一个新的函数$Y_{\text{reb}}=g_{\text{reb}}(\boldsymbol{X})=g(\boldsymbol{X})-e$，基于$U$学习函数建立描述$Y_{\text{reb}}=g_{\text{reb}}(\boldsymbol{X})$的Kriging模型，进而利用式(22)通过Kriging模型预测值计算系统的失效概率。

3 时变可靠性分析的双层代理模型方法

考虑一个时变问题，其功能函数表示为$Y(t)=G\left(\psi\left(\boldsymbol{X}^{\text{E}},t\right),\boldsymbol{X}^{\text{S}},t\right)$，其中，$\boldsymbol{X}=\left(\boldsymbol{X}^{\text{E}},\boldsymbol{X}^{\text{S}}\right)$为不确定性参数，$\boldsymbol{X}^{\text{E}}$为系统输入不确定性参数，$\boldsymbol{X}^{\text{S}}$为系统本身不确定性参数，$\psi\left(\boldsymbol{X}^{\text{E}},t\right)$为与时间相关的输入。该函数可以理解为概率空间$\boldsymbol{R}^n$中关于随机变量$\boldsymbol{X}$和时间区间$[T_1,T_2]$中的参数$t$的多元函数。当输入变量$\boldsymbol{X}\in\boldsymbol{R}^n$固定在实现值$\boldsymbol{x}=\left(\boldsymbol{x}^{\text{E}},\boldsymbol{x}^{\text{S}}\right)$时，$y(t)=G\left(\psi\left(\boldsymbol{x}^{\text{E}},t\right),\boldsymbol{x}^{\text{S}},t\right)$成为关于时间参数$t$的一元函数。假设阈值为$e$，响应超出阈值即为失效，则时变失效概率的计算表达式为

$$P_f(T_1,T_2)=P\{F\}=P\{y(t)>e,\exists t\in[T_1,T_2]\} \tag{23}$$

在时间区间$[T_1,T_2]$内，时变失效域F可等价表示为

$$F=\left\{\max_{t\in[T_1,T_2]}G\left(\psi\left(\boldsymbol{x}^{\text{E}},t\right),\boldsymbol{x}^{\text{S}},t\right)>e\right\} \tag{24}$$

式(23)可表示为

$$P_f(T_1,T_2)=P\left\{\max_{t\in[T_1,T_2]}G\left(\psi\left(\boldsymbol{x}^{\text{E}},t\right),\boldsymbol{x}^{\text{S}},t\right)>e\right\} \tag{25}$$

令$G_{\text{e}}(\boldsymbol{X})=\max\limits_{t\in[T_1,T_2]}\left(G\left(\psi\left(\boldsymbol{X}^{\text{E}},t\right),\boldsymbol{X}^{\text{S}},t\right)\right)$，通过分析$Y(t)=G\left(\psi\left(\boldsymbol{X}^{\text{E}},t\right),\boldsymbol{X}^{\text{S}},t\right)$关于时间$t$的极大值，可将时变可靠性问题转化为时不变可靠性问题，即

$$P_f(T_1,T_2)=P\{G_{\text{e}}(\boldsymbol{x})>e\}=P\{G_{\text{e}}(\boldsymbol{x})-e>0\} \tag{26}$$

基于NARX模型和Kriging模型的双层代理模型的基本思路是：在内层利用训练样本点$\boldsymbol{x}_n(n=1,\cdots,N)$的部分时刻$t_j^{(\text{F})}(j=1,\cdots,N_{\text{F}})$的真实输出建立NARX代理模型$y_{\text{N}}^{(n)}(t)=\Gamma(\boldsymbol{z}(\boldsymbol{x}_n,t))$，外层则利用$y_{\text{N}}(t)=\Gamma(\boldsymbol{z}(\boldsymbol{x},t))$预估得到所有训练样本处的动态输出的极值$G_e(\boldsymbol{x})=\max\limits_{t\in[T_1,T_2]}\Gamma(\boldsymbol{z}(\boldsymbol{x},t))$后，通过$U$学习函数建立关于$Y_{\text{reb}}=G_{\text{e}}(\boldsymbol{X})-e$与$\boldsymbol{X}$的自适应Kriging代理模型$Y_{\text{reb}}\approx G_{\text{eK}}(\boldsymbol{X})$。最终通过$G_{\text{eK}}(\boldsymbol{X})$得到$Y_{\text{reb}}$的估计值，通过式(26)计算时变失效概率$P_f(T_1,T_2)$的估计值。双层代理模型的具体步骤如下：

1) 根据输入随机变量$\boldsymbol{X}$的联合概率密度函数$f_{\boldsymbol{X}}(\boldsymbol{x})$产生容量为$N^x$的样本池$\boldsymbol{S}^x=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_{N^x}]$，作

为更新代理模型和进行时变可靠性分析的样本点。

2) 在 $\boldsymbol{S}^x$矩阵中随机抽取 N_{K}^x $\left(N_{\mathrm{K}}^x \ll N^x\right)$个样本 $\boldsymbol{x}_n\left(n=1,2,\cdots,N_{\mathrm{K}}^x\right)$, 用于构建 NARX 模型并求解极大值。

3) 在给定 $\boldsymbol{X}=\boldsymbol{x}_n$时, 输出 $y^{(n)}(t)$仅由时间 t决定, 则根据第 1 节所述, 可建立描述系统输出的 NARX 代理模型 $y_{\mathrm{N}}^{(n)}(t)=\Gamma(\boldsymbol{z}(\boldsymbol{x}_n,t))$, 从而求得给定 $\boldsymbol{X}=\boldsymbol{x}_n$ 时, 时变功能函数在 $t\in[T_1,T_2]$内关于 t的极大值 $G_{\mathrm{e}}^{(n)}(\boldsymbol{x}_n)\approx\max\limits_{t\in[T_1,T_2]} y_{\mathrm{N}}^{(n)}(t)$, 由此可以得到全部 N_{K}^x个样本对应的输出极大值 $G_{\mathrm{e}}^{(n)}(\boldsymbol{x}_n)\left(n=1,2,\cdots,N_{\mathrm{K}}^x\right)$。

4) 构建用于训练外层 Kriging 模型的初始训练集 $\boldsymbol{K}^x=\left\{(\boldsymbol{x}_n,G_{\mathrm{e}}(\boldsymbol{x}_n)-e),n=1,2,\cdots,N_{\mathrm{K}}^x\right\}$。

5) 根据第 2 节所述, 利用 $\boldsymbol{K}^x$训练得到关于函数 $Y_{\mathrm{reb}}=G_{\mathrm{e}}(\boldsymbol{X})-e$与 $\boldsymbol{X}$的自适应 Kriging 模型 $G_{\mathrm{eK}}(\boldsymbol{X})$, 在此过程中, 若收敛条件不满足, 则可由 U学习函数得到新的训练点 $\boldsymbol{x}_u$, 进而通过建立 NARX 模型得到 $\boldsymbol{X}=\boldsymbol{x}_u$时时变功能函数在 $t\in[T_1,T_2]$内的极大值 $G_{\mathrm{e}}^{(u)}(\boldsymbol{x}_u)\approx\max\limits_{t\in[T_1,T_2]} y_{\mathrm{N}}^{(u)}(t)$, 重新构造训练集 $\boldsymbol{K}^x=\{\boldsymbol{K}^x\cup\left(\boldsymbol{x}_u,G_{\mathrm{e}}^{(u)}(\boldsymbol{x}_u)-e\right)\}$对 Kriging 模型进行更新, 直到收敛条件满足。

6) 利用建立的 Kriging 模型 $G_{\mathrm{eK}}(\boldsymbol{X})$进行可靠性分析表达式为

$$\hat{P}_f(T_1,T_2)=\frac{1}{N^x}\sum_{i=1}^{N^x} I_F(G_{\mathrm{eK}}(\boldsymbol{x}_i))=\frac{N_f}{N^x} \tag{27}$$

$$\mathrm{Cov}\left(\hat{P}_f\right)=\sqrt{\left(1-\hat{P}_f\right)/(N^x-1)\hat{P}_f} \tag{28}$$

式中: $I_F(G_{\mathrm{eK}}(\boldsymbol{x}))$为失效域 $F=\{\boldsymbol{x}:G_{\mathrm{eK}}(\boldsymbol{x})\geqslant 0\}$的指示函数, 如果 $\{\boldsymbol{x}:G_{\mathrm{eK}}(\boldsymbol{x})\geqslant 0\}$, 则 $I_F(\cdot)=1$, 否则 $I_F(\cdot)=0$; N_f为 $G_{\mathrm{eK}}(\boldsymbol{x}_i)(i=1,2,\cdots,N^x)$大于 0 的个数。

4　算例分析

4.1　航空发动机涡轮盘

图 1 为航空发动机涡轮盘, 其工作载荷表示为: $F(\boldsymbol{X},t)=C\omega^2/(2\pi)+2\rho\omega^2 J$, 其中, C、ρ和 J分别为涡轮盘的转动系数、密度和截面惯性矩, ω为转动角速度, 与转动频率 n的关系为 $\omega=2n\pi$, $n=0.1n_0\sin(t/4+1.2)$, $t\in[0,7\pi]$, t为时间参数。涡轮盘横截面积为 A, 强度极限为 σ_{s}, 则时变极限状态函数可写为: $y(t)=F(\boldsymbol{X},t)-\sigma_{\mathrm{s}}A$。当工作载荷超出许用载荷时, 结构失效, 失效概率表示为: $P_f(t,t_{\mathrm{s}})=P\{y(t)>0\}$。

假设随机变量 $\boldsymbol{X}=\{\sigma_{\mathrm{s}},A,J,n_0,\rho,C\}$相互独立且均服从正态分布, 其分布参数如表 1 所示。

在此例中, 认为转动角速度 ω为系统的输入, 即 $\psi(n_0,t)=\omega$。在构建内层 NARX 模型时, 首先确定 NARX 结构, 取时间间隔 $d=0.01\pi$, 功能函数自由度为 1, 取时间延迟为 2, 模型项是输出和激励的过去值的二阶多项式函数形式: $g_i(\boldsymbol{z}(t))=\psi^l\left(n_0,t-n_\psi d\right)y^m\left(t-n_y d\right)$, 其中, $l+m\leqslant 2, 0\leqslant l\leqslant 2, 0\leqslant m\leqslant 2, n_\psi\in\{0,1,2\}, n_y\in\{1,2\}$。

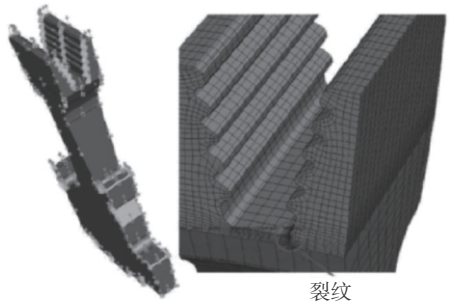


图 1　航空发动机涡轮盘

Fig. 1　Aeroengine turbine disk

表 1　随机变量分布参数

Table 1　Variables and parameters

分布类型	运算输出	$\sigma_{\mathrm{s}}/(\mathrm{N}\cdot\mathrm{m}^{-2})$	A/m^2	J/m^4	n_0/s^{-1}	$\rho/(\mathrm{kg}\cdot\mathrm{m}^{-3})$	C
正态分布	均值	1.1×10^9	0.012 4	1.22×10^{-4}	163	8 240	5.67
	标准差	1.1×10^8	1.24×10^{-3}	1.22×10^{-5}	16.3	824	0.567

根据输入随机变量 $\boldsymbol{X}$的分布参数产生容量 $N^x=10^7$的样本池 $\boldsymbol{S}^x=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_{N^x}]$, 作为更新代理模型和进行时变可靠性分析的样本点。在 $\boldsymbol{S}^x$矩阵中随机抽取 20 个样本 $\boldsymbol{x}_n(n=1,2,\cdots,20)$, 用于构建 NARX 初始模型。对于每个样本 $\boldsymbol{x}_n$, 其动态输出可以看作关于时间参数 t的一元函数, 因此可以由 NARX 模型描述这一输出。依 1.2 节所述, 从离散的时刻点中均匀抽取 3 个时刻作为建模初始时刻点, 计算对应的系统输出用于建立 NARX 模型, 经更新后得到满足精度要求的 NARX 模型同时得到输出最大值 $G_{\mathrm{e}}(\boldsymbol{x}_n)$。以随机变量 $\boldsymbol{x}_n$及其输出最大值 $G_{\mathrm{e}}(\boldsymbol{x}_n)$构建训练外层 Kriging 模型的初始训练集 $\boldsymbol{K}^x=\{(\boldsymbol{x}_n,G_{\mathrm{e}}(\boldsymbol{x}_n)-e),n=1,2,\cdots,20\}$, 利用 U学习函数建立自适应 Kriging 模型, 并进行可靠性分析。所得结果与蒙特卡罗(MC)法及双层 Kriging 法(Double-Kriging)所得结果对比如表 2 所示。其中, 样本量表示求得有效失效概率值所需要的时刻点数, 即调用功能函数的次数, 构建外层模型样本量为建立 Kriging 模型所需输出最大值的数量, 即建模所需随机变量的数量。

从表 2 可以看出, 与双层 Kriging 法相比, 建立 NARX-Kriging 模型所需的总的样本量更少, 所得失效概率的误差也更小。在构建单个内层模型时, 本文所提 NARX-Kriging 双层代理模型方法所需时刻点样本要略多于双层 Kriging 法。这是由于 NARX

表 2 不同方法的时变可靠性分析结果

Table 2 Time-dependent reliability analysis results of different methods

方法	P_f	Cov(P_f)	样本量	构建内层模型平均时刻点	构建外层模型样本量	误差/%
NARX-Kriging	8.00×10^{-5}	0.035 4	172	6	27	2.56
Double-Kriging	8.76×10^{-5}	0.033 8	196	5	39	6.70
MC	8.21×10^{-5}	0.034 9	7×10^{9}			

模型作为时序模型，其某一时刻的输出由过去时刻输出和当前或之前时刻输入值决定，使得构建NARX模型时所需的时刻点数成倍增加。然而正是由于考虑了时间的累积效应，本文所建内层模型对原始功能函数的描述更加精确。

图 2 给出了分别采用 NARX 模型和 Kriging 模型得到的原始系统的时变均值和标准差的对比。结果表明，由于 NARX 模型对原始功能函数的描述更加精确，预测的时变均值和标准差相比于 Kriging 模型与参考 MC 解吻合得更好。正是由于 NARX 内层模型更高的准确性，本文方法在构建外层模型时所需样本量，即输出最大值的数量更少，最终使得本文方法在建立完整双层模型时更加高效，预估结果也更加精确。

4.2 Bouc-Wen 模型

在工程中，在动载荷作用下很多工程结构或系统会进入非弹性状态而体现出滞回特性，如系统的反复加载、卸载时的弹塑性材料变形、铁磁元件中的磁滞特性等。人们提出很多理论模型来描述滞回系统，其中，Bouc-Wen 模型以其光滑的滞回特性和较强的通用性被广泛应用。以一个 Bouc-Wen 模型来验证本文方法的有效性和适用性，其微分控制方程为：$\ddot{y}(t)+2\varsigma\omega_{\mathrm{n}}\dot{y}(t)+\omega^2(\rho' y(t)+(1-\rho')w(t))=-A\sin(\omega_x t)$，$\dot{w}(t)=\gamma\dot{y}(t)-\alpha|\dot{y}(t)|\,|w(t)|^{n-1}w(t)-\beta\dot{y}(t)\,|w(t)|^n$，其中，$\varsigma$和$\omega_{\mathrm{n}}$分别为振荡器的阻尼比和固有频率，在本例中令$\rho'=\beta=0$，$\gamma=1$，$n=1$，振荡器的输入表示为$\psi(A,\varphi,t)=A\sin(\varphi t)$，$w(t)$称为 hysteric 位移。该系统静止时的初始条件是$y(0)=0,\dot{y}(0)=0,w(0)=0$，振荡器考虑了 5 个不确定参数，即$\boldsymbol{X}=[\varsigma,\omega_{\mathrm{n}},\alpha,A,\varphi]$，所有不确定参数均服从正态分布，分布参数如表 3 所示。

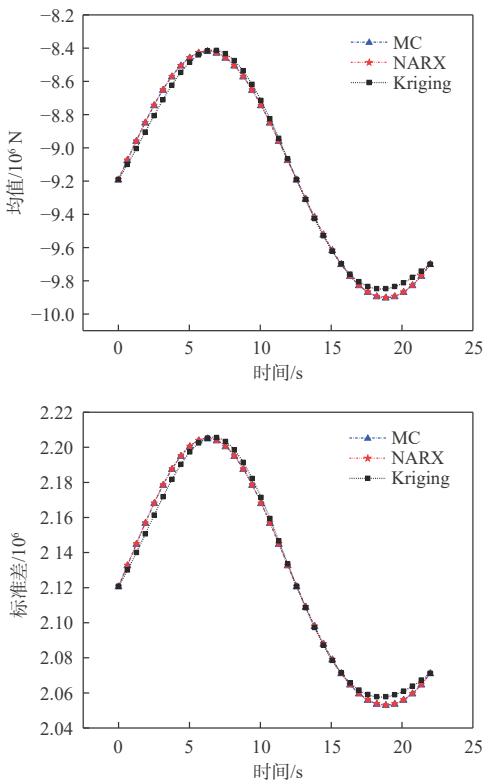


图 2 NARX 模型与 Kriging 模型预测的时变均值和标准差与参考解的对比

Fig. 2 Comparison of reference resolution with time-dependent mean and standard deviation predicted by NARX model and Kriging model

表 3 分布参数

Table 3 Variables and parameters

分布类型	运算输出	ς	ω_{n}/(rad·s^{-1})	α	A/N	φ/(rad·s^{-1})
正态分布	均值	0.02	2π	50	1	π
	标准差	0.001	0.1π	2.5	0.05	0.05π

振荡器的运行时间为 0～30 s，时间间隔为 0.01 s，当最大振幅超过 0.53 m 时，认为振荡器失效，即$P_f(t,t_{\mathrm{s}})=P\{y(t)>0.53\}=P\{y(t)-0.53>0\}$。

本例为震荡波形输出形式，取时间延迟为 2。对于 Bouc-Wen 振荡器，Spiridonakos、Chatzi 认为使用绝对项可以有效地捕捉非线性系统的滞回行为[24]，在此使用一种与前不同的多项式基函数形式：$g_i(\boldsymbol{z}(t))=\{\psi^l(A,\varphi,t-n_\psi d)\,|y(t-n_y d)|^m, y^l(t-n_y d)|y(t-n_y d)|^m\}$，其中，$l+m\leqslant3$，$l\in\{1,2,3\}$，$m\in\{0,1\}$，$n_\psi\in\{0,1,2\}$，$n_y\in\{1,2\}$。根据输入随机变量$\boldsymbol{X}$的分布参数产生容量 $N^x=2\times10^5$的样本池 $\boldsymbol{S}^x=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_{N^x}]$，在 $\boldsymbol{S}^x$ 矩阵中随机抽取 40 个样本 $\boldsymbol{x}_n(n=1,2,\cdots,40)$，构建 NARX 模型得到输出最大值$G_{\mathrm{e}}(\boldsymbol{x}_n)$。以此创建训练外层 Kriging 模型的初始训练集$\boldsymbol{K}^x=\{(\boldsymbol{x}_n,G_{\mathrm{e}}(\boldsymbol{x}_n)-e),n=1,2,\cdots,40\}$，利用 U 学习函数建立自适应 Kriging 模型，并进行可靠性分析。所得结果与 MC 法及双层 Kriging 法所得结果对比如表 4 所示，其中，MC 法的结果作为参考解。可以看到，双层 Kriging 法无法得到收敛的失效概率值。

如图 3 所示，通过 NARX 模型和 Kriging 模型

表 4　各种方法的时变可靠性分析结果

Table 4　Results of each method

方法	P_f	cov(P_f)	样本量	构建内层模型平均时刻点	构建外层模型样本量	误差/%
NARX-Kriging	7.15×10^{-4}	0.083 5	9 900	150	67	1.93
Double-Kriging						
MC	7.35×10^{-4}	0.082 4	6×10^{8}			

预测得到的原始系统输出的时变均值和标准差的对比发现, 相较与双层 Kriging 模型, 不论是在剧烈震荡还是区域平稳阶段, NARX-Kriging 模型对于系统原始状态的描述都更加准确。特别是在极值的描述上, 双层 Kriging 模型内层的 Kriging 模型对于此类问题并不能有效地识别极值, 以至于最终无法得到收敛的可靠性分析结果。而 NARX 模型因在内层对系统的时变特性描述良好, 使得本文方法最终得到的时变失效概率与 MC 法比较误差较小, 为 1.93%, 这说明本文所提 NARX-Kriging 模型对此类问题具有良好的适用性。

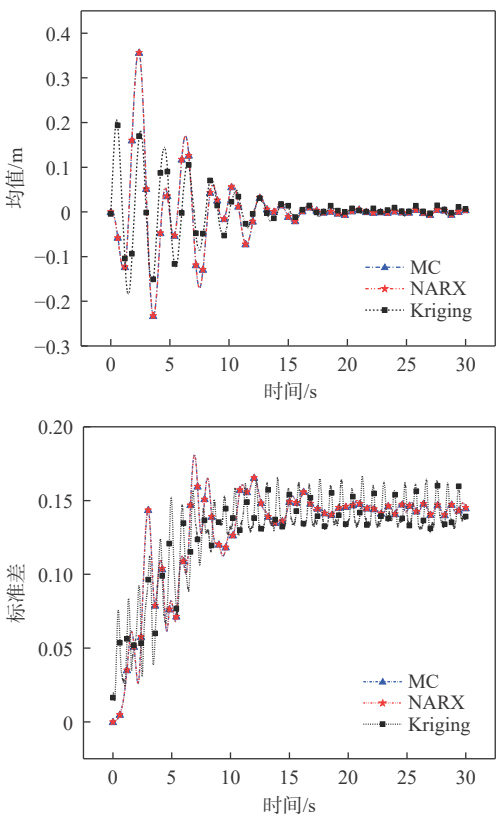


图 3　NARX 模型与 Kriging 模型预测的时变均值和标准差与参考解的对比

Fig. 3　Comparison of reference solution with time-dependent mean and standard deviation predicted by NARX model a nd Kriging model

4.3　水力涡轮叶片模型

考虑一个包含随机过程输入的非平稳动态系统, 以水轮机叶片为例, 验证本文方法的有效性。水力涡轮机利用随时间变化的河流水流荷载作用下的水动力叶片, 将水流的动能转化为机械功。水力涡轮系统的安全性和效率受涡轮叶片形状的影响, 如图 4 所示, 叶片的横截面可以由h_1、h_2、l_1三个几何参数描述。为了描述图 5 所示的随时间变化的河流流量, 河流流速$\psi(t)$由平均值$\mu_\psi(t)$、标准差$\sigma_\psi(t)$和自相关函数$\rho_\psi(t_1,t_2)$所给出的随机过程建模。

$$\begin{cases}\mu_\psi(t)=\sum_{i=1}^{4}a_i^m\sin\left(b_i^m t+c_i^m\right)\\ \sigma_\psi(t)=\sum_{j=1}^{4}a_j^s\exp\left(-\left(\left(t-b_j^s\right)/c_j^s\right)^2\right)\\ \rho_\psi(t_1,t_2)=\cos(2\pi(t_2-t_1))\end{cases}$$

式中: $a_1^m=3.815, a_2^m=2.528, a_3^m=1.176, a_4^m=-0.078\,56$, $b_1^m=0.289\,5, b_2^m=0.588\,7, b_3^m=0.761\,9, b_4^m=2.183, c_1^m=-0.266\,8, c_2^m=0.965\,1, c_3^m=3.116, c_4^m=-3.161, a_1^s=0.738\,2$, $a_2^s=1.013, a_3^s=1.875, a_4^s=1.283, b_1^s=6.456, b_2^s=4.075$, $b_3^s=9.913, b_4^s=1.035, c_1^s=0.919\,3, c_2^s=1.561, c_3^s=6.959$, $c_4^s=2.237$。

随机荷载$\psi(t)$作用下叶片根部的弯矩M_{flap}: $M_{\text{flap}}=\frac{1}{2}\rho\psi^2(t)C_m$, 其中, $\rho=1\times10^3$ kg/m^{3}为水的密度, $C_m=0.342\,2$为力矩系数。由此定义了水力涡轮叶片的极限状态函数: $y(t)=\frac{h_1M_{\text{flap}}}{EI_a}-A_l$, 其中, A_l为材料的极限应变, M_{flap}为叶片根部在转动方向的弯矩, $E=14$ GPa为弹性模量, 叶片根部的转动惯量表示为: $I_a=\frac{2}{3}l_1\left(h_1^3-h_2^3\right)$。

本例假设h_1、h_2、l_1和A_l为随机变量, 其他参数为确定性的。随机变量的分布参数如表 5 所示。在本例中, 考虑的时间间隔为 0～12 月, 时间间隔为 0.01 月。

对于随机过程输入, 通过 KL 扩展[34]将随机过

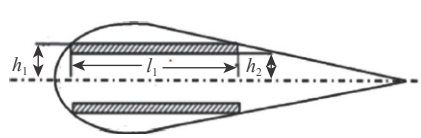


图 4　涡轮叶片根部横截面

Fig. 4　Cross section of turbine blade at the root

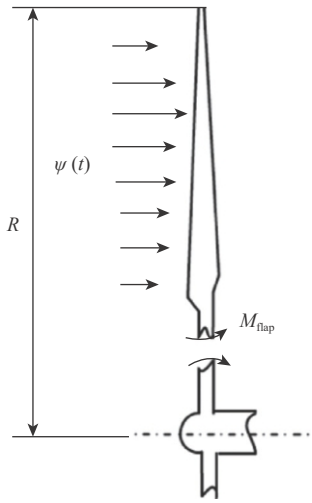


图5 涡轮叶片上的河流载荷

Fig. 5 River flow load on turbine blade

程进行分解，并保留包含随机过程输入99%方差的前4个KL项，进而将随机过程的不确定性保留在KL扩展中的4个标准正态随机变量$k_i(i=1,2,3,4)$中。因此，本例在外层构建随机变量-极值关系时

表5 分布参数

Table 5 Variables and parameters

运算输出	l_1/m	h_1/m	h_2/m	A_l	$\psi(t)/(\mathrm{m}\cdot\mathrm{s}^{-1})$
均值	0.22	0.025	0.019	0.030	$\mu_\psi(t)$
标准差	0.002 2	0.000 25	0.000 19	0.000 30	$\sigma_\psi(t)$
自相关函数					$\rho_\psi(t_1,t_2)$

注：l_1, h_1, h_2, A_l为正态分布；$\psi(t)$为随机过程。

共需要考虑对系统输出产生影响的8个随机变量，即系统输入的不确定性参数为$\boldsymbol{X}^{\mathrm{E}}=[k_1,k_2,k_3,k_4]$，系统本身的不确定性参数为$\boldsymbol{X}^{\mathrm{S}}=[l_1,h_1,h_2,A_l]$。

本例同样为震荡形输出形式，时间间隔$d=0.01$月。对于强非线性输出系统，本例采用与4.2节算例相同的多项式基函数形式：$g_i(\boldsymbol{z}(t))=\{\psi^l(A,\varphi,t-n_\psi d)\,|y(t-n_y d)|^m, y^l(t-n_y d)\,|y(t-n_y d)|^m\}$，其中，$l+m\leqslant 5, l\in\{0,1,2,3,4\}, m\in\{0,1\}, n_\psi\in\{0,1,2\}$，$n_y\in\{1,2\}$。利用不同方法计算的可靠性分析结果如表6所示。可以看到，双层Kriging法依旧无法得到收敛的失效概率值。

表6 各种方法的时变可靠性分析结果

Table 6 Results of each method

方法		P_f	$Cov(p_f)$	样本量	构建内层模型平均时刻点	构建外层模型样本量	误差/%
计算结果	NARX-Kriging	6.32×10^{-4}	0.056 3	7 640	40	191	2.27
	Double-Kriging						
	MC	6.18×10^{-4}	0.056 8	6×10^{8}			

如图6所示，对比通过NARX模型和Kriging模型得到的原始系统的时变均值和标准差可以看到，双层Kriging模型内层的Kriging模型在对系统输出的描述上仅能识别出输出的大体趋势，并不能展示系统输出的震荡细节，其对极值点的预测与真实值相比差异明显，以致最终无法得到有效的可靠性分析结果。而NARX模型则能对系统输出行为做准确的描述，最终得到的可靠性指标误差为2.27%。这再次证明了本文所提NARX-Kriging模型对长时间震荡模型的适用性和准确性。

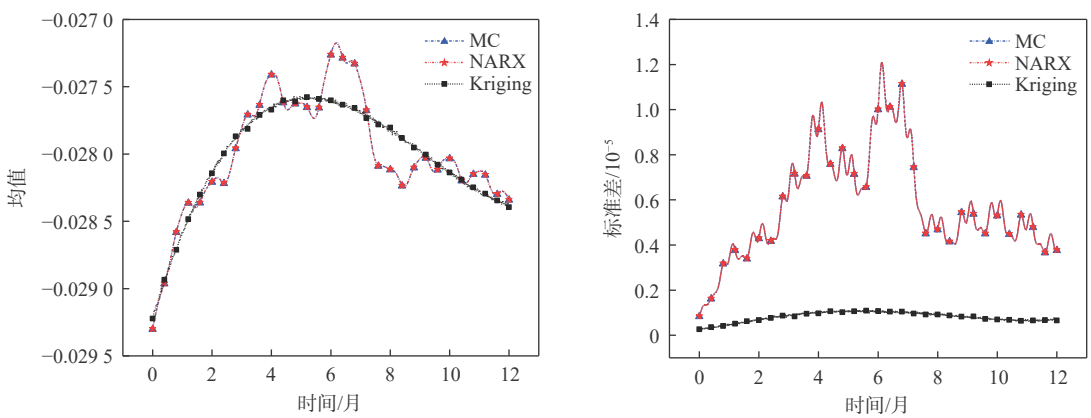


图6 NARX模型与Kriging模型预测的时变均值和标准差与参考解的对比

Fig. 6 Comparison of reference solution with time-dependent mean and standard deviation predicted by NARX model and Kriging model

5 结 论

1）在将时变可靠性问题转化为时不变问题进行求解时，将时间效应考虑在模型构造过程中，从而建立了一种精度较高的高效时变可靠性分析模型。

2）算例分析表明，与现有的双层Kriging法相比，本文方法对于输出呈较强波动性的系统具有良好的适应性，在求解系统的时变可靠性时，用于建

立模型的样本量更少, 分析结果更加准确。在实际工程中, 如震动问题、往复运动等涉及波动输出的问题非常普遍, 针对这些建立代理模型比较困难的问题, 本文方法具有一定的工程指导意义。

3) 本文方法依旧存在可供改进的空间, 如开发更为简单有效的 NARX 基本形式选择方法、更高效的 Kriging 模型的自适应准则等。此外, 基于该方法可以进一步发展全局可靠性灵敏度分析, 从而为结构的设计优化提供帮助。

参考文献（References）

[1] 黄贤振, 张义民, 吴茂昌, 等. 基于降维积分的结构系统可靠性分析研究[J]. 力学学报, 2013, 45(3): 456-460.
HUANG X Z, ZHANG Y M, WU M C, et al. Research on the method for reliability analysis of structural systems based on the dimensional-reduction integration[J]. Chinese Journal of Theoretical and Applied Mechanics, 2013, 45(3): 456-460(in Chinese).

[2] 杨绿峰, 吴文龙, 余波. 一种不依赖失效路径的结构体系可靠度分析方法[J]. 中国科学:物理学 力学 天文学, 2014, 44(11): 1220-1231.
YANG L F, WU W L, YU B. A failure-path independent method for analysis of structural system reliability[J]. Scientia Sinica: Physica, Mechanica & Astronomica, 2014, 44(11): 1220-1231(in Chinese).

[3] 亢战, 罗阳军. 基于凸模型的结构非概率可靠性优化[J]. 力学学报, 2006, 38(6): 807-815.
KANG Z, LUO Y J. On structural optimization for non-probabilistic reliability based on convex models[J]. Chinese Journal of Theoretical and Applied Mechanics, 2006, 38(6): 807-815(in Chinese).

[4] 祁武超, 邱志平. 基于区间分析的结构非概率可靠性优化设计[J]. 中国科学: 物理学 力学 天文学, 2013, 43(1): 85-93.
QI W C, QIU Z P. Non-probabilistic reliability-based structural design optimization based on the interval analysis method[J]. Scientia Sinica: Physica, Mechanica & Astronomica, 2013, 43(1): 85-93(in Chinese).

[5] LIANG J H, MOURELATOS Z P, NIKOLAIDIS E. A single-loop approach for system reliability-based design optimization[J]. Journal of Mechanical Design, 2007, 129(12): 1215-1224.

[6] DU X P, CHEN W. Sequential optimization and reliability assessment method for efficient probabilistic design[J]. Journal of Mechanical Design, 2004, 126(2): 225-233.

[7] RICE S O. Mathematical analysis of random noise[J]. The Bell System Technical Journal, 1944, 23(3): 282-332.

[8] CRANDALL S H, CHANDIRAMANI K L, COOK R G. Some first-passage problems in random vibration[J]. Journal of Applied Mechanics, 1966, 33(3): 532-538.

[9] SUNDAR V S, MANOHAR C S. Time variant reliability model updating in instrumented dynamical systems based on Girsanov's transformation[J]. International Journal of Non-Linear Mechanics, 2013, 52: 32-40.

[10] BREITUNG K, RACKWITZ R. Nonlinear combination of load processes[J]. Journal of Structural Mechanics, 1982, 10(2): 145-166.

[11] ZAYED A, GARBATOV Y, GUEDES SOARES C. Time variant reliability assessment of ship structures with fast integration techniques[J]. Probabilistic Engineering Mechanics, 2013, 32: 93-102.

[12] ANDRIEU-RENAUD C, SUDRET B, LEMAIRE M. The PHI2 method: A way to compute time-variant reliability[J]. Reliability Engineering & System Safety, 2004, 84(1): 75-86.

[13] 张德权, 韩旭, 姜潮, 等. 时变可靠性的区间PHI2分析方法[J]. 中国科学:物理学 力学 天文学, 2015, 45(5): 41-53.
ZHANG D Q, HAN X, JIANG C, et al. The interval PHI2 analysis method for time-dependent reliability[J]. Scientia Sinica: Physica, Mechanica & Astronomica, 2015, 45(5): 41-53(in Chinese).

[14] HU Z, DU X P. A sampling approach to extreme value distribution for time-dependent reliability analysis[J]. Journal of Mechanical Design, 2013, 135(7): 071003.

[15] LI J, CHEN J B, FAN W L. The equivalent extreme-value event and evaluation of the structural system reliability[J]. Structural Safety, 2007, 29(2): 112-131.

[16] CHEN J B, LI J. The extreme value distribution and dynamic reliability analysis of nonlinear structures with uncertain parameters[J]. Structural Safety, 2007, 29(2): 77-93.

[17] DU X P. Time-dependent mechanism reliability analysis with envelope functions and first-order approximation[J]. Journal of Mechanical Design, 2014, 136(8): 081010.

[18] 石岩, 吕震宙, 周易成. 随时间和空间变化可靠性及可靠性全局灵敏度分析的包络函数法[J]. 西北工业大学学报, 2017, 35(4): 591-598.
SHI Y, LU Z Z, ZHOU Y C. Temporal and spatial reliability and global sensitivity analysis with envelope functions[J]. Journal of Northwestern Polytechnical University, 2017, 35(4): 591-598(in Chinese).

[19] CHENG K, LU Z Z. Time-variant reliability analysis based on high dimensional model representation[J]. Reliability Engineering & System Safety, 2019, 188: 310-319.

[20] HAWCHAR L, EL SOUEIDY C P, SCHOEFS F. Principal component analysis and polynomial chaos expansion for time-variant reliability problems[J]. Reliability Engineering & System Safety, 2017, 167: 406-416.

[21] HU Z, MAHADEVAN S, DU X P. Uncertainty quantification of time-dependent reliability analysis in the presence of parametric uncertainty[J]. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems Part B: Mechanical Engineering, 2016, 2(3): 031005.

[22] WANG Z Q, WANG P F. A nested extreme response surface approach for time-dependent reliability-based design optimization[J]. Journal of Mechanical Design, 2012, 134(12): 121007.

[23] HU Z, DU X P. Mixed efficient global optimization for time-dependent reliability analysis[J]. Journal of Mechanical Design, 2015, 137(5): 051401.

[24] SPIRIDONAKOS M D, CHATZI E N. Metamodeling of dynamic nonlinear structural systems through polynomial chaos NARX models[J]. Computers & Structures, 2015, 157: 99-113.

[25] CHEN S, BILLINGS S A, LUO W. Orthogonal least squares methods and their application to non-linear system identification[J]. International Journal of Control, 1989, 50(5): 1873-1896.

[26] BILLINGS S A. Nonlinear system identification: NARMAX methods in the time, frequence and spatio-temporal domains[M]. New

York: John Wiley & Sons, Inc., 2013.

[27] CHEN Z H, NI Y Q. On-board identification and control performance verification of an MR damper incorporated with structure[J]. Journal of Intelligent Material Systems and Structures, 2011, 22(14): 1551-1565.

[28] WEI H L, BILLINGS S A. Model structure selection using an integrated forward orthogonal search algorithm assisted by squared correlation and mutual information[J]. International Journal of Model Identification and Control, 2008, 3(4): 341-356.

[29] EFRON B, HASTIE T, JOHNSTONE I, et al. Least angle regression[J]. The Annals of Statistics, 2004, 32(2): 407-499.

[30] LV Z Y, LU Z Z, WANG P. A new learning function for Kriging and its applications to solve reliability problems in engineering[J]. Computers & Mathematics with Applications, 2015, 70(5): 1182-1197.

[31] KLEINEN J P C. Kringing metamodeling in simulation: A review[J]. Eurpean Journal of Operational Research, 2009, 192(3): 707-716.

[32] LOPHAVEN S N, NIELSEN H B, SONDERGAARD J. Aspects of the MATLAB toolbox DACE: IMM-REP-2002-13[R]. Kongens Lyngby: Technical University of Denmark, 2002.

[33] ECHARD B, GAYTON N, LEMAIRE M. AK-MCS: An active learning reliability method combining Kriging and Monte Carlo Simulation[J]. Structural Safety, 2011, 33(2): 145-154.

[34] GHANEM R G, SPANOS P D. Stochastic finite elements: A spectral approach[M]. Berlin: Springer-Verlag, 1991.

Double-loop surrogate model for time-dependent reliability analysis based on NARX and Kriging models

CHANG Zeming，LI Luyi*

(School of Aeronautics，Northwestern Polytechnical University，Xi'an 710072，China)

Abstract：Structural systems with dynamic output performance has gained more and more attention in engineering practice. However, most existing surrogate models for estimating time-dependent reliability of such systems only consider the effect of the random variables which are acting on the system at the current moment, but ignores their effect with time-dependent accumulation. Therefore, these models cannot give an accurate prediction for time-dependent reliability of the dynamic systems. To solve this problem. this paper proposes a double-loop surrogate model method for time-dependent reliability analysis based on the nonlinear autoregressive with exogenous input (NARX) model and Kriging model. In the inner loop of the proposed method, NARX model is used to describe the variation of the output response with time under given random input variables, which can accurately simulate the dynamic behavior of the system. In the outer loop, the Kriging model of the extreme value of the dynamic systems is built based on the random input samples and the corresponding extreme values predicted by the inner NARX model. The reliability of the time-varying structure system then can be easily obtained based on the outer Kriging model. Finally, the effectiveness and accuracy of the proposed method for reliability analysis of dynamic structural systems with fluctuating outputs is verified by three examples.

Keywords：time-dependent reliability; failure probability; surrogate model; NARX model; Kriging model

Received：2021-09-09；**Accepted**：2021-12-10；**Published Online**：2022-01-18 19:48

URL：kns.cnki.net/kcms/detail/11.2625.V.20220118.1321.003.html

Foundation item：National Natural Science Foundation of China (51875464)

* **Corresponding author**. E-mail：luyili@nwpu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0542

基于可达集的无人机低空飞行冲突解脱算法

杨建航[1,2]，张福彪[1,2,*]，王江[1,2]

(1. 北京理工大学 宇航学院，北京 100081；2. 无人机自主控制技术北京重点实验室，北京 100081)

摘 要：针对无人空中交通管理（UTM）中的冲突解脱问题，提出了以可达集分析为基础的实时避撞算法。该算法可用于城市低空环境中的密集交通流空域，保证无人机（UAV）飞行过程的安全性。基于相对运动的概念，通过分析平面空域中的飞行博弈问题对避撞系统进行建模，同时利用水平集方法和最优控制理论对无人机的可达集进行分析和计算，使用机载传感器获取无人机与周围物体的信息，为每架无人机提供新的避撞策略。通过3种不同空域环境的飞行案例进行仿真实验，验证了该策略不仅可以得到平滑的飞行路径，实时安全地解决冲突解脱问题，而且针对合作/非合作目标均有效。

关 键 词：无人空中交通管理；冲突解脱；城市低空环境；可达集分析；避撞策略；平面空域

中图分类号：V221^{+}.3; TB553

文献标志码：A **文章编号**：1001-5965(2023)07-1813-15

人们对移动性日益增长的需求对传统的交通方式和交通系统提出了挑战。人们期望以更快捷、更舒适的方式到达目的地，并期望在有需要后的短时间内完成移动性任务。无人机（unmanned aircraft vehicles, UAV）体积小、灵活、可低空飞行且能够携带一套复杂的机载传感器，正在成为当今社会不可缺少的工具[1]。用于监视和侦察等军事应用的无人机系统（unmanned aircraft system, UAS）的部署早已非常成熟，最近由于其操作简单、低空空域开放等特点而被广泛应用于各种民用任务中[2]，主要包括应急搜索和救援、执法、精准农业、包裹递送、基础设施检查及灾害探测等[3]。随着民用UAS的迅速发展，解决无人机在国家空域系统运行的安全问题越来越关键，无人机必须以不干扰交通的方式行动，保持与其他飞机分离。无论在商业航空的结构化空域，还是在拥有都市型航空交通系统（urban air mobility, UAM）的低空空域，空中交通都变得越来越复杂且难以管理。在过去的50年里，有人驾驶和无人驾驶飞机数量的急剧增加给空中交通管理（air traffic management, ATM）带来了严峻的考验[4]。无人机需要一定程度的自主权，以确保在复杂环境中安全分离。为此，NASA在2015年提出了无人空中交通管理（unmanned air traffic management, UTM）系统的概念[5]。与传统的ATM相比，UTM旨在更严格、更拥挤的空域中管理更多的飞机。

UTM有许多组成部分，本文主要关注冲突解脱问题，因为其对空中交通的安全性十分重要。据统计，自2006年以来，质量小于30 kg的小型无人机系统（sUAS）占据了新型飞机的大部分市场[6]。随着现代技术的快速进步及无人机成本的大幅降低，预计未来几年低空空中交通流量将大幅增长[7]，越来越多的无人机行动将会发生在靠近建

收稿日期：2021-09-09；录用日期：2021-12-17；网络出版时间：2022-01-12 12:44
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220112.0910.002.html
基金项目：国家重点研发计划 (2021YFF0601304)；国家自然科学基金 (U1913602)；北京市科学技术委员会基金 (Z181100003218013)
* 通信作者. E-mail：fubiao.zhang@bit.edu.cn

引用格式：杨建航，张福彪，王江. 基于可达集的无人机低空飞行冲突解脱算法 [J]. 北京航空航天大学学报，2023，49（7）：1813-1827.
YANG J H，ZHANG F B，WANG J. Conflict resolution algorithms for UAV low-altitude flight based on reachable set [J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（7）：1813-1827（in Chinese）.

筑物或机场的大都市地区[8]。大都市环境的空域将由传统的载人飞行器和无人机共享，低空空域密集的操作环境大大增加了冲突的风险。在低空的城市地区，每架无人机必须避开各种障碍，如静态物体（高层建筑、树木、灯柱、禁飞区等）和动态物体（起重机、其他无人机、有人机等）。为此，需要先进的UTM系统有效处理复杂环境中的飞行冲突，及时保证安全[9]。ATM主要关注的是事先计划好的机场之间的航班，这些航班的时间表能够预先得知。相比之下，UAS具有时间和空间上的高度不确定性，无人机可以在任意时段飞行，且没有固定的路径点。此外，目前的ATM系统是“人在回路”的决策系统，人因误差无法避免，不利于飞行冲突的解决[10]，在无人驾驶的航空领域，无人机将进行多次短途飞行，需要更高水平的自动化来解决冲突[11]。因此，需要先进的冲突解脱工具和决策方法来管理UTM下涌入的大量UAS。

冲突指2架飞行器之间的距离小于特定的间隔，使飞行器安全受到威胁的一种状态。冲突探测与解脱是UTM系统的核心任务之一[12]，无人飞行器检测到可能发生的冲突后，通过安全避撞操作来解决现有的冲突威胁。目前，检测-避撞（detect-and-avoid, DAA）技术已经有了多年的研究[4]，并进行了开发和测试工作。文献[13]以复杂网络理论为基础，通过冲突DAA策略来解决局部空域冲突解脱的问题，文献[14]提供了无人机进入非隔离空域的DAA解决方案，文献[15]描述了DAA系统在复杂的城市环境中对小型无人机机动的需求，文献[16]介绍了DAA系统冲突探测的应用。本文基于可达集方法和DAA技术，提出一种新的UAS避撞策略，为UTM中的冲突解脱问题提供了一种安全有效的方法。

避撞能力是决定无人机能否与其他飞行器共享空域，以及能否在动态不确定环境中执行任务的重要因素[17]。相关文献提出了多种解决无人机避撞问题的算法。Kuchar和Yang[18]对20～21世纪的冲突解脱算法进行了全面的回顾与评估，并对无人机避撞模型包含的概念进行了分类。文献[19-20]分析和评估了各种无人机自主防撞避让算法。以主成分分析（principal component analysis, PCA）算法[21]为代表的几何算法考虑了无人机与障碍物的空间位置，计算简单，但通常无法得到平滑的飞行路径。基于概率和统计理论的蒙特卡罗算法[22]主要针对一对无人机之间的避撞问题，未考虑复杂的多机空域。人工势场法[23]的优点在于计算量小，具备防撞的实时性，但存在局部最优点的问题，不能准确到达目标。同时，解决无人机防撞问题的方法还有各类基于人工智能的避撞决策算法，包括A*搜索算法[24]和各种优化算法，如遗传算法[25]、蚁群算法[26]、粒子群优化算法[27]等。这类优化算法可以进行全局搜索，规划出最优或次优路径，但计算时需要多次迭代，较大的计算量导致规避障碍物时会出现延迟问题，而且适用范围较小，大多只能用于躲避静态障碍物，对于动态物体不能进行有效避撞。文献[28]利用可达集进行避撞检查，提高了避撞的成功率和安全性，但仍需预测轨迹，并且需要提前掌握障碍物的可达集信息，无法实现实时避障。此外，在复杂的城市环境中，障碍物密集，碰撞危险性高。因此，传统的路径规划算法在低空环境的应用具有局限性。上述针对城市环境的UAS避撞算法主要侧重于预先规划，要求UAS的飞行环境是透明的。对于陌生的任务环境，难以掌握UAS的全局飞行区域信息，无法提前通过规划算法获得安全的飞行路径[29]，执行效率会受到影响。因此，有必要探索一种实时的冲突解脱算法，应用于复杂的低空飞行空域。

本文提出的基于可达集的实时避撞算法具有简洁、高效、安全、实用等特点。飞机间的相互作用可以通过计算重要的空中交通特征（如相对位置和相对速度）来进行数学描述。为了避免碰撞，本文采用2种方式实现飞机的分离，即航向角变化（heading angle change, HAC）和速度变化（velocity change, VC）。假设无人机定高飞行，使用无人机的当前位置、速度及航向信息，基于相对运动的概念处理平面空域冲突问题。由于可达集的计算结果可以离线缓存，在规避障碍物时具有较小的计算开销，保证了实时性。提取时给定一个误差界，使可达集边界的允许误差有一个限度。采用2架飞机博弈的过程对空中避撞问题进行建模，通过可达集分析方法对无人机加以控制，从而完成空中交通避撞任务。

1 背景知识

1.1 可达集

随着空中交通系统越来越复杂，确保系统正常工作也越来越困难，因此，安全验证已应用到多种工程领域。安全验证的最简单形式是仿真，但其一次只能检查系统的一条轨迹，通过该方法检查每个可能的系统轨迹安全性的代价是十分昂贵的。因此，在控制系统的安全验证中，可达集分析方法应运而生。如果可以准确确定一组

状态集合, 系统步入此状态集合会导致危险程度剧增, 则可以通过确保系统状态不属于此状态集合来验证系统的安全性。考虑民用航空交通管制中无人机之间的避撞, 若要保证设计的控制方法不会让 2 架无人机陷入碰撞损失, 可达集分析是完成此任务的合适工具。

根据指定初始条件或终端条件, 可达集分为前向可达集(forward reachable set, FRS)和后向可达集(backward reachable set, BRS)。对于 FRS, 指定初始条件, 并确定沿开始的轨迹可以到达的所有状态的集合; 相反, 对于 BRS, 指定终端条件或目标集, 并试图确定可到达该目标集的所有轨迹的状态集合。2 种可达集如图 1 所示[30]。

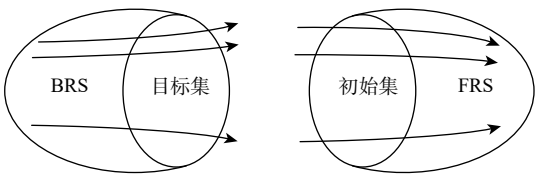


图 1　前向可达集与后向可达集[30]

Fig. 1　Forward reachable set and backward reachable set[30]

1.2　FRS 与 BRS 对比

FRS 是系统在某个时间段 t 之后可能处于的状态集合。当 2 架飞机同时飞行时, 为其中一架飞机(飞机 Ⅱ)计算 FRS, 并且另一架飞机(飞机 Ⅰ)试图避免进入此集合以确保产生无碰撞轨迹, 如图 2(a)所示。但是, 这种集合从飞机 Ⅱ 的初始状态随时间 t 向外扩张, 使得飞机 Ⅰ 的飞行路线过于保守, 导致飞机 Ⅰ 的可飞行区域大大减少。为减少这种过度保守性, 计算 FRS 时通常要使时间范围 t 保持很小, 并经常对其进行重新计算。这种方法在寻找避撞轨迹方面是有效的, 但很难应用到存在更多不确定性的低空环境中。此外, 使用 FRS 的另一缺点是安全性取决于飞机 Ⅰ 的能力和操作频率, 即使 FRS 实时计算, 飞机 Ⅰ 仍有可能无法对瞬间威胁做出反应[31]。

为研究飞机的避撞问题, 令目标集表示系统的碰撞状态集合。BRS 是一组状态集合, 这组集合在某个时间段 t 之后可能使系统状态处于目标集合中。FRS 和 BRS 之间的关键区别在于: BRS 在时间上是向后计算的, 是围绕在飞机 Ⅰ 目标集附近的封闭集合, 该集合表示在控制约束下, 能够到达碰撞状态的当前时刻所有状态的集合, 如图 2(b)所示。对于 2 架飞机的动力学系统, BRS 的计算使得飞机 Ⅰ 能在任何时间和任何状态下做出反应(如果 2 架飞机相互靠近, 则飞机 Ⅰ 可以快速转向, 远离飞机 Ⅱ 以避免碰撞)。简而言之, BRS 不像 FRS 那样过于保守。BRS 的计算结果可以离线缓存, 在运行时, 可以通过实时查找缓存可达集来计算 UAV 的最佳飞行策略。

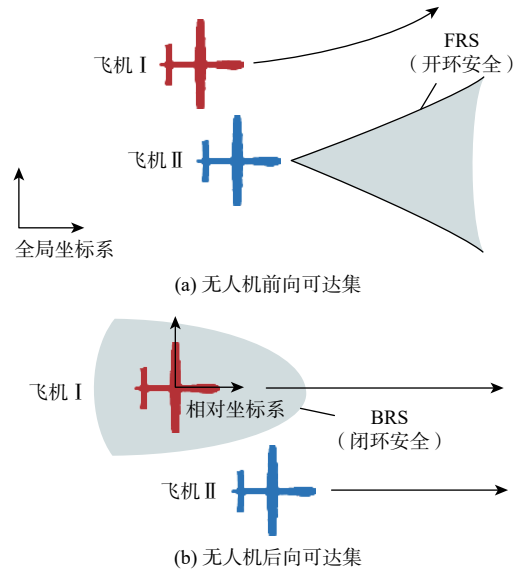


图 2　无人机前后向可达集确保安全方式

Fig. 2　FRS and BRS ensure safety measures

空中交通安全管理问题的研究中使用 BRS, 原因有如下 3 点:

1) 以目标集为导向的封闭集合的计算可确保无人机仅在必要时才使用可达集工具, 而不会过度影响无人机的预定轨迹。

2) BRS 的计算中本质上已经定义了无人机的安全性, 而 FRS 方法的安全性取决于无人机的预定轨迹是否与 FRS 相交。

3) BRS 为安全控制提供了较强的计算能力, 可以在高操作频率下对安全控制进行评估, 从而应对瞬间威胁。

2　模型与算法

2.1　BRS 计算

2.1.1　系统建模

使用常微分方程对系统进行建模:

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t}=\dot{x}(t)=f(x(t),u(t),d(t)) \qquad x(t_{\mathrm{f}})\in T_{\mathrm{f}},t\in[\tau,t_{\mathrm{f}}] \tag{1}$$

式中: $x(t)\in\mathbf{R}^n$为系统的状态变量; $u(t)\in U$为控制量; $d(t)\in D$为扰动量; τ为某一时刻; t_{f}为终端时刻; $T_{\mathrm{f}}=\{x\in\mathbf{R}^n|T(x)\leqslant 0\}$为系统的目标集。对于固定的 u和d, 假设系统动力学 $f(\cdot)$在x上是一致连续的、有界的, 且利普希茨连续的, 则后向可达集$T(\tau)$被定义为在τ时刻所有状态$x(\tau)$的集合, 该集合满足以下要

求：在给定上述系统的目标集T_f的条件下，对任意干扰输入$d(t)\in D(\tau\leqslant t\leqslant t_f)$，都存在一个控制输入$u(t)\in U(\tau\leqslant t\leqslant t_f)$，使得$x(\tau)$状态沿着满足上述动力学的轨迹可到达$x(t_f)\in T_f$的状态。

在飞机避撞问题中，$f(\cdot)$表示2架飞机之间或飞机与障碍物之间的相对动力学，$u(t)$表示飞机Ⅰ的控制输入，$d(t)$表示飞机Ⅱ的控制输入（飞机Ⅱ可用来表示障碍物，其运动输入为干扰输入），T_f表示2架飞机碰撞或飞机与障碍物碰撞时对应的状态集合。

2.1.2 水平集方法

水平集(level set)是一种用于计算动态隐式集合描述的数值方法，其基本思想是：将N维问题描述成N+1维问题的一个水平集，即将某状态x的集合看成高一维空间处于t时刻的隐函数$\phi(x,t)$的零水平集，其中，$x=\{x_1,x_2,\cdots,x_n\}\in\mathbf{R}^n$，由此将求解$x$和$t$描述的演化过程转换为求解$\phi(x,t)$的一个水平集的演化过程。$\phi(x,t)$为水平集函数，$P$为限制区域，则$t$时刻$x\in P$等价于$\phi(x,t)\leqslant 0$，如图3所示[32]。

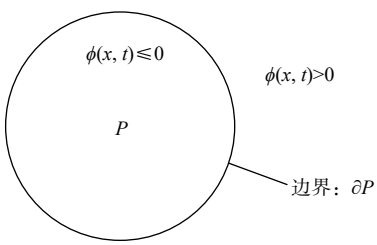


图3 水平集方法的图示[32]

Fig. 3 Illustration of level set method[32]

假设界面上每个点的速度由外部生成的速度场$f(x)$给出，若要分析集合边界∂P随时间t在速度场$f(x)$中的变化，则只需要计算使得$\phi(x,t)=0$成立的x。$\phi(x,t)$和$f(x)$存在以下关系：

$$\frac{\partial}{\partial t}\phi(x,t)+f(x)\nabla\phi(x,t)=0 \tag{2}$$

此即哈密顿-雅可比偏微分方程(HJ PDE)，也称为水平集方程。

这种方法能够计算演化中的曲线而不需对曲线参数化，Osher和Fedkiw[33]提供了有关动态隐式曲面和水平集方法的入门指导。目前，该方法已十分成熟，并得到了广泛使用。

令连续、有界的隐函数$\phi(x,t)$是HJ PDE的黏性解，则系统的HJ PDE表示如下：

$$\frac{\partial}{\partial t}\phi(x,t)+H\left(x,\frac{\partial}{\partial x}\phi(x,t)\right)=0 \tag{3}$$

令$p=\dfrac{\partial}{\partial x}\phi(x,t)$，若系统的目标是提高安全性，则哈密顿量为

$$H(x,p)=\max_{u\in U}\min_{d\in D}p^{\mathrm{T}}f(x,u,d) \tag{4}$$

文献[30,34]及相关文献中证明黏性解$\phi(x,t)$是式(3)和式(4)的正确弱解，用来生成可达集的隐式曲面。因此，计算$\phi(x,t)$的精确近似值时可以使用水平集文献中成熟的数值方法。黏性解$\phi(x,t)$可用于创建安全控制区域：如果系统状态在可达集之外，则输入u的任何控制策略都是安全的[35]。为避免系统进入危险状态，使用$\phi(x,t)$中的相对距离和相对转角信息，在系统接近边界时逐渐引入相关控制量。

2.1.3 2架飞机系统动力学

本节研究2架相同飞机在平面空域中运动的博弈问题。根据文献[36]的理论研究三维飞机的防撞示例，确定2架飞机运动学模型的可达性。为方便分析，每架飞机都被建模为具有可变的平面位置和航向、固定的线速度大小及可控制的角速度输入的简单运动对象，如图4所示。在冲突解脱问题中，通过飞行器之间的距离与碰撞半径r_{safe}相比较来确定安全程度，小于r_{safe}将发生碰撞。飞机Ⅱ试图通过合适的输入进入飞机Ⅰ的碰撞区域（见图4中实线圆），因此，需要确定飞机Ⅱ能够进入飞机Ⅰ碰撞区域的集合，即飞机Ⅰ的可达集。

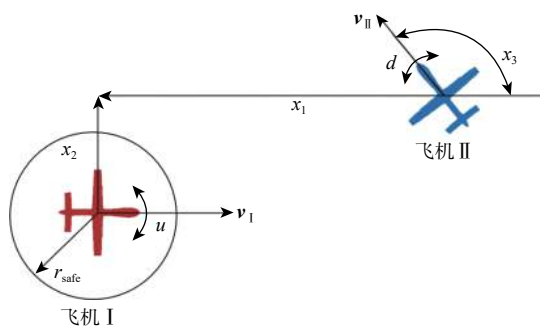


图4 2架飞机相对运动系统模型

Fig. 4 Two aircraft relative motion system model

将飞机Ⅰ固定在原点，则系统动力学描述为

$$\dot{\boldsymbol{x}}=\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}-v_{\mathrm{Ⅰ}}+v_{\mathrm{Ⅱ}}\cos x_3+ux_2\\v_2\sin x_3-ux_1\\d-u\end{bmatrix}=f(\boldsymbol{x},u,d) \tag{5}$$

式中：$\begin{bmatrix}x_1 & x_2\end{bmatrix}^{\mathrm{T}}\in\mathbf{R}^2$为2架飞机在平面内的相对位置；$x_3\in[0,2\pi]$为2架飞机航向的相对转角；$v_{\mathrm{Ⅰ}}$和$v_{\mathrm{Ⅱ}}$分别为飞机Ⅰ和飞机Ⅱ的线速度。

目标集T_f是一组危险的系统状态，系统动力学中的输入参数分为2部分：试图进入目标集的输入d和试图避免其进入目标集的输入u。其中，$u\in U=[-U^{\max},+U^{\max}]$为飞机Ⅰ的输入角速度，$d\in D=[-D^{\max},$

$+D^{\max}]$为飞机Ⅱ的输入角速度。由于 2 架飞机可以在任意相对航向上发生碰撞, 目标集T_{f}只取决于 2 架飞机的相对位置x_1和x_2, 包含以原点为中心、r_{safe}为半径的圆内任意状态:

$$T_{\mathrm{f}}=\left\{\boldsymbol{x}\in\boldsymbol{R}^2\times[0,2\pi]\,\middle|\,\sqrt{x_1^2+x_2^2}\leqslant r_{\mathrm{safe}}\right\} \tag{6}$$

2.1.4　最优控制和可达集数值解

结合式(4)和式(5), 可以得出哈密顿量为

$$H(x,p)=\max_{u\in U}\min_{d\in D}\left(\boldsymbol{p}^{\mathrm{T}}f(x,u,d)\right)=\max_{u\in U}\min_{d\in D}(-p_1v_{\mathrm{I}}+p_1v_{\mathrm{II}}\cos x_3+p_2v_{\mathrm{II}}\sin x_3+(p_1x_2-p_2x_1-p_3)u+p_3d) \tag{7}$$

式中: $p_i=\dfrac{\partial}{\partial x_i}\phi(x,t),i=1,2,3$。

通过最优控制理论解出[37]:

$$u^*=U^{\max}\mathrm{sign}(p_1x_2-p_2x_1-p_3) \tag{8}$$

$$d^*=-D^{\max}\mathrm{sign}(p_3) \tag{9}$$

将式(8)和式(9)代入式(7)中, 得到最优哈密顿算子:

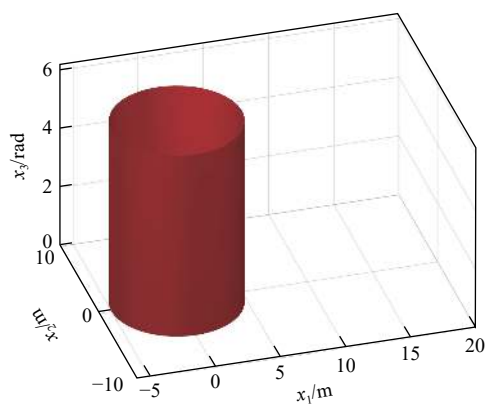
$$H^*(x,p)=-p_1v_{\mathrm{I}}+p_1v_{\mathrm{II}}\cos x_3+p_2v_{\mathrm{II}}\sin x_3+U^{\max}|p_1x_2-p_2x_1-p_3|-D^{\max}|p_3| \tag{10}$$

根据式(10)和式(3)得出 2 架飞机避撞系统的 HJPDE, 从而求得$\phi(x,t)$的近似数值解。令$\phi(x,t)=0$, 即可得零水平集下随时间变化的飞机Ⅰ可达集。

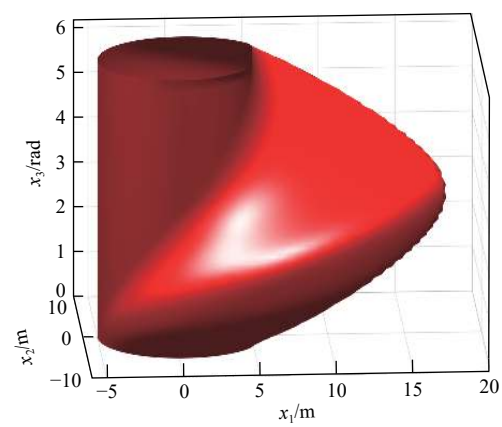
图 5(a)展示了目标集, 图 5(b)～图 5(d)展示了不同视野下 2.8 s 时的 BRS, 图 6(a)～图 6(d)为一个碰撞系统可达集随时间变化的动画演示, (MATLAB 代码及使用说明可从文献 [37] 中获得)。此示例使用以下参数: $r_{\mathrm{safe}}=5,m$, $v_{\mathrm{I}}=v_{\mathrm{II}}=5$, $U^{\max}=D^{\max}=1$。

其中, r_{safe}表示目标集的半径。考虑图 6(d)的几个水平切片, 相对航向坐标x_3是这些切片的垂直坐标。所有可达集中最大的水平切片位于$x_3=\pi$处, 对应 2 架飞机沿相反方向飞行的情况。顶部$x_3=0$和底部$x_3=2\pi$的水平切片表示 2 架飞机沿相同方向飞行的情况, 此时可达集为目标碰撞集。

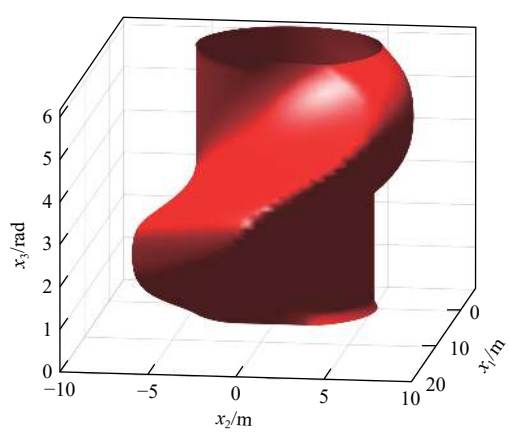
图 7 显示了来自碰撞系统动画的某一帧, 并添加了相应描述。飞机Ⅰ在实线目标集包围的左侧飞行, 飞机Ⅱ从右侧飞行, 飞机Ⅰ周围的封闭虚线是 2 架飞机当前相对航向的可达集边界。



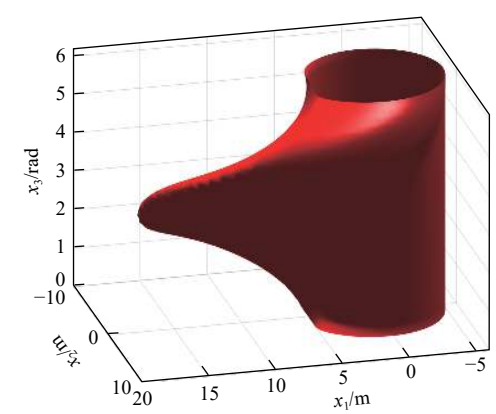
(a) 目标集(t=0 s)



(b) 可达集视野1(t=2.8 s)



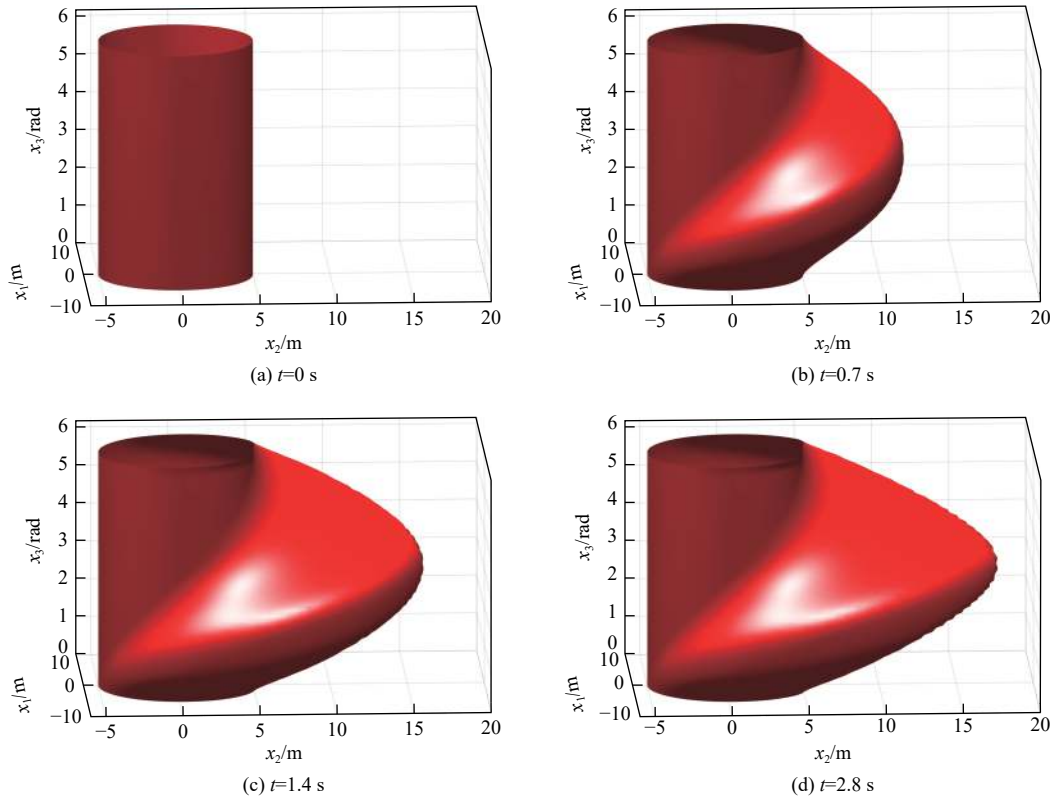
(c) 可达集视野2(t=2.8 s)



(d) 可达集视野3(t=2.8 s)

图 5　目标集与后向可达集

Fig. 5　Target set and backward reachable set



(a) t=0 s (b) t=0.7 s

(c) t=1.4 s (d) t=2.8 s

图 6 可达集随时间变化

Fig. 6 Reachable set changes over time

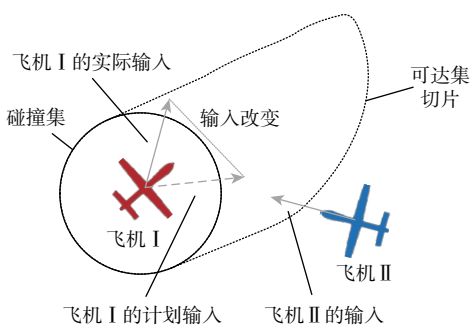
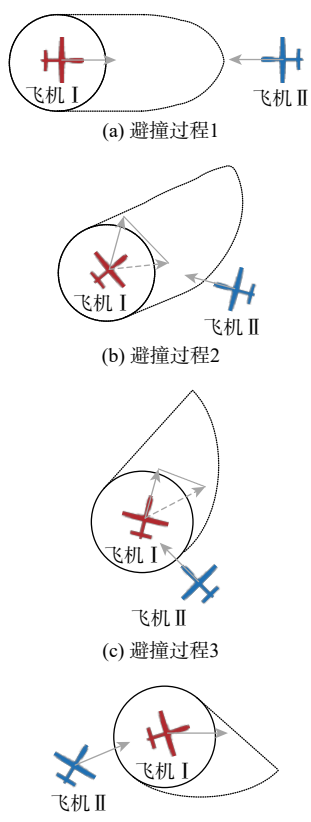


图 7 避撞动画的注释帧

Fig. 7 Comment frame of collision avoidance animation



(a) 避撞过程1

(b) 避撞过程2

(c) 避撞过程3

(d) 避撞过程4

图 8 飞机Ⅱ避撞示意图

Fig. 8 Aircraft Ⅱ cannot enter the reachable set, thus avoiding collision

图 8 显示了来自该动画的一系列帧，通过选择适当的控制输入，飞机Ⅰ可以阻止飞机Ⅱ进入可达集，进而避免飞行过程发生碰撞。

2.2 无人机与障碍物模型

当无人机在城市低空环境中飞行时，需同时考虑静态和动态障碍物。本文研究二维平面内的无人机避撞模型，即假设无人机定高飞行(单层空域)。为简化模型，所有障碍物和无人机都被建模为圆形物体，如图 9 所示。

同时，假设无人机可以通过 DAA 系统机载传感器实时测得障碍物在其运动平面内的位置和障碍物半径，并且可以实时获得无人机当前时刻的位

置、速度及航向等可测量物理量。二维无人机运动模型可以简化成一个双积分系统:

$$\begin{cases}\dot{\boldsymbol{p}}=\boldsymbol{v}\\ \dot{\boldsymbol{v}}=\boldsymbol{a}\end{cases} \tag{11}$$

式中: $\boldsymbol{p}=[x\quad y]^{\mathrm{T}}$、$\boldsymbol{v}=[\dot{x}\quad \dot{y}]^{\mathrm{T}}$、$\boldsymbol{a}=[a_x\quad a_y]^{\mathrm{T}}$分别为无人机的平面位置、速度、加速度。

图 9 中黑色实心圆表示对飞行路径造成干扰的障碍物, 半径为R的虚线圆表示无人机的冲突探测区域, r表示无人机的半径, $\boldsymbol{p}_{\mathrm{obs}}$表示障碍物中心的位置, r_{obs}表示障碍物的半径。

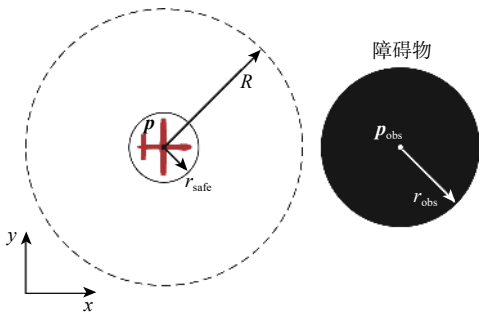


图 9　无人机与障碍物模型

Fig. 9　UAV and obstacle model

为方便表述, 定义如下区域:

1) 碰撞区域, 无人机与障碍物相撞的区域。相应的碰撞集合: $\Lambda\triangleq\{\boldsymbol{x}|\boldsymbol{x}\in\mathbf{R}^2,\|\boldsymbol{x}-\boldsymbol{p}\|\leqslant r+r_{\mathrm{obs}}\}$。

2) 冲突区域, 无人机检测到冲突的区域[38]。在此区域内, 无人机可做出避撞反应, 相应的冲突集合: $\Gamma\triangleq\{\boldsymbol{x}|\boldsymbol{x}\in\mathbf{R}^2,r+r_{\mathrm{obs}}<\|\boldsymbol{x}-\boldsymbol{p}\|\leqslant R+r_{\mathrm{obs}}\}$。

3) 安全区域, 无人机与障碍物不发生冲突的区域。在此区域内, 无人机没有碰撞风险, 相应的安全集合: $\Psi\triangleq\{\boldsymbol{x}|\boldsymbol{x}\in\mathbf{R}^2,\|\boldsymbol{x}-\boldsymbol{p}\|>R+r_{\mathrm{obs}}\}$。

无人机的可达集及 3 种区域的关系如图 10 所示。通过$\boldsymbol{p}_{\mathrm{obs}}$与$\boldsymbol{p}$之间的距离来判断障碍物进入到了无人机的哪种区域。

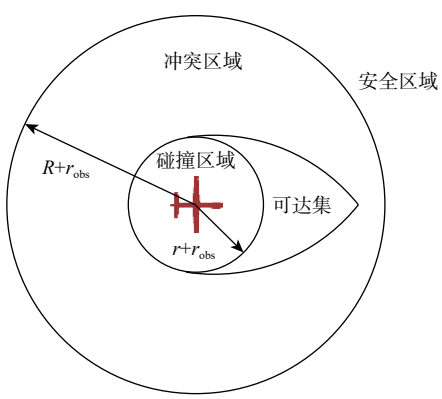


图 10　可达集与 3 种区域的关系

Fig. 10　Relation between reachable set and three areas

2.3　控制律的设计

在飞机冲突解脱问题中, 考虑了共享同一空域的有限飞机集合。当检测到任何一对飞机之间存在碰撞风险时, 即执行冲突避免机动。飞机在平面空域的初始位置、速度和航向决定了飞行的轨迹, 研究目的是为每架飞机提供一个新的避撞策略, 以解决所有可能的冲突, 且不过度影响飞机的原始轨迹。根据第 1 节介绍的可达集概念, 若使系统避免危险, 则应尽量使障碍物远离无人机的 BRS。为使无人机处于绝对安全状态, 设计如下控制律:

$$\boldsymbol{a}=\begin{cases}\boldsymbol{0} & \boldsymbol{p}_{\mathrm{obs},i}\in\Psi\\ k_a\dfrac{\|\boldsymbol{p}-\boldsymbol{p}_{\mathrm{int},i}\|+r_{\mathrm{obs},i}}{\|\boldsymbol{p}-\boldsymbol{p}_{\mathrm{obs},i}\|}\boldsymbol{n} & \boldsymbol{p}_{\mathrm{obs},i}\in\Gamma\text{且障碍物在可达集外}\\ a_{\max}\boldsymbol{n} & \text{障碍物在可达集内且}\boldsymbol{p}_{\mathrm{obs},i}\notin\Lambda\end{cases} \tag{12}$$

式中: $\boldsymbol{n}$为单位向量, 且$\boldsymbol{n}\in\{\boldsymbol{n}|\boldsymbol{n}\times\boldsymbol{v}\equiv 0,\boldsymbol{n}\cdot(\boldsymbol{p}_{\mathrm{obs},i}-\boldsymbol{p})\leqslant 0\}$, $i=1,2,\cdots,n$, n为障碍物的个数; Ψ、Γ、Λ为 2.2 节定义的 3 种区域; $a_{\max}$为最大推力对应的最大加速度; k_a为加速度比例系数; $\boldsymbol{p}_{\mathrm{int},i}$为无人机与第$i$个障碍物位置的连线和可达集交点(以下简称为交点)的位置; $r_{\mathrm{obs},i}$为第i个障碍物的半径; $\boldsymbol{p}_{\mathrm{obs},i}$为第$i$个障碍物的位置; $\boldsymbol{p}$为无人机的位置。无人机避撞模型如图 11 所示。

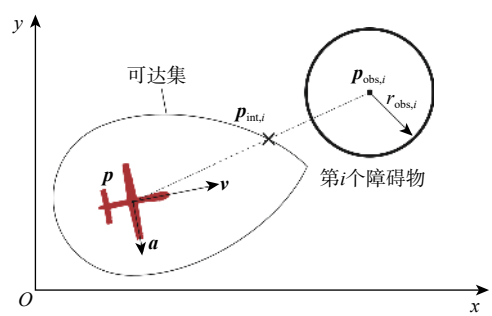


图 11　UAV 避障模型

Fig. 11　UAV collision avoidance model

当无人机未检测到冲突时, 不存在碰撞风险, 不需施加控制量; 无人机与障碍物发生冲突后, 根据无人机与障碍物之间的距离、无人机与交点之间的距离及障碍物半径大小来设置控制量; 一旦障碍物进入无人机的可达集, 则应设置最大控制量来保证无人机的安全。无人机基于可达集的平面空域实时避撞算法的具体工作流程如图 12 所示, 该算法可保证无人机以较小的转弯曲率安全地躲避路径上所有的障碍物。

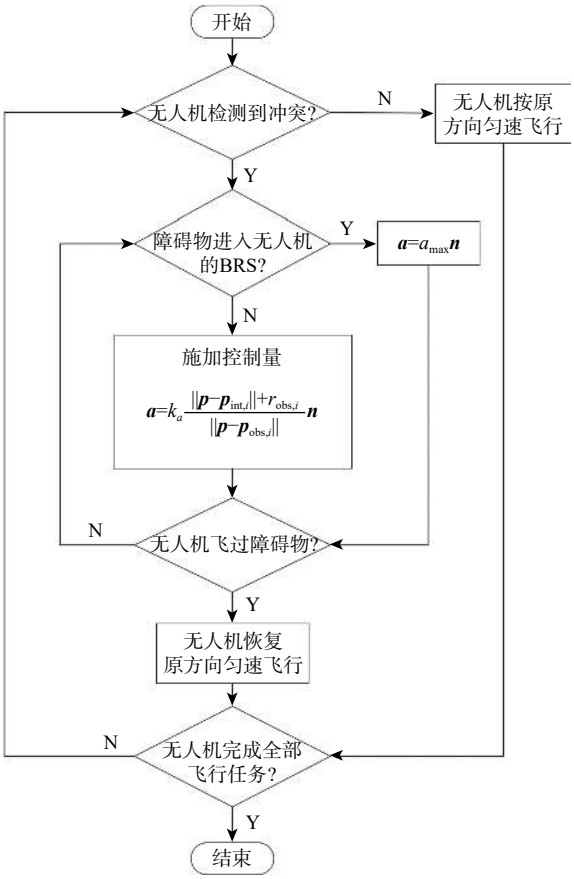


图 12 本文算法工作流程

Fig. 12 Work flow of proposed algorithm

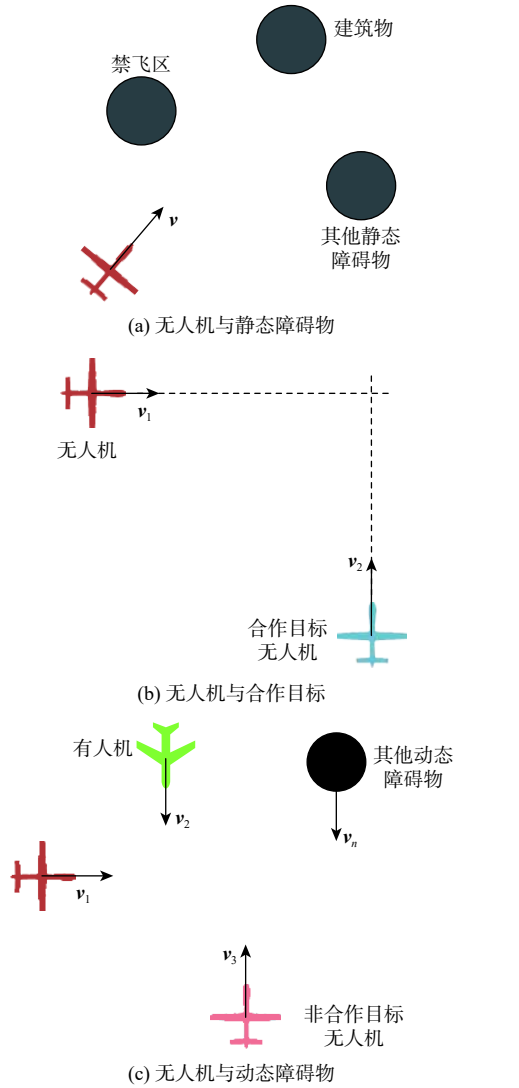


图 13 用于仿真的 3 种案例

Fig. 13 Three cases for simulation

3 仿真与结果

本节通过仿真实验来验证所提出的可达集避撞算法效果。为说明算法的可行性和有效性，模拟了以下 3 种案例：

1）无人机躲避多个静态障碍物（建筑物、禁飞区等），如图 13(a)所示。

2）2 架使用可达集避撞算法的无人机（无人机与合作目标）相互避撞，如图 13(b)所示。

3）无人机躲避多个动态障碍物（非合作目标无人机、有人机等），如图 13(c)所示。

假设无人机在经过静态障碍物（或合作目标、动态障碍物）时，若与静态障碍物（或合作目标、动态障碍物）之间的距离大于目标集的碰撞半径，并具有一定的余量，即表示无人机无法与静态障碍物（或合作目标、动态障碍物）发生碰撞，处于绝对安全状态。

3.1 仿真配置

仿真环境为 I7-8550U，主频 1.99 GHz，16 GB 内存，基于 MATLAB 2020b 平台进行仿真实验。利用 level set 工具箱计算无人机的可达集，并进行在线提取。本文以城市低空环境中的局部区域为例，将实验环境设置为城市中的小范围区域进行坐标系的建模，验证算法在局部区域中的避撞效果，无人机从原点沿 x 轴正方向匀速直线飞行，实现实时避障。为尽量真实模拟无人机避撞的过程，本文参考了文献 [39] 的实验参数，仿真实验具体参数如表 1 所示。可达集获取时间较短，保证了飞行过程中避障的实时性。

3.2 仿真结果

3.2.1 案例 1：无人机躲避多个静态障碍物

设定仿真环境条件：任务环境为城市低空的局部区域，其长为 650 m，宽为 550 m。任务区域中存

表 1 仿真参数

Table 1 Simulation parameters

参数	数值
无人机初始位置 p_0/m	[0 , 0]
反应时间/s	1
无人机质量 m/kg	20
无人机半径 r/m	1.5
无人机的初始速度 v/(m·s^{-1})	[40 , 0]
冲突探测半径 R/m	120
障碍物半径 r_{obs}/m	15
可达集获取时间/s	0.097

在多个建筑物(即静态障碍物), 通过地形建模将其转化为多个$r_{obs}=15$ m的圆形障碍, 并且会造成无人机与圆形障碍相撞。无人机平飞速度$v=40$ m/s, 传感器探测半径$R=120$ m, 可达集通过实时提取缓存 BRS 获得, 根据仿真参数, 目标集的碰撞半径$r_{safe}=r+r_{obs}=16.5$ m。由于障碍物是静态的, 在计算可达集时不考虑航向的相对转角这一维度, 此时可达集只与相对位置有关, 如图 14 所示。

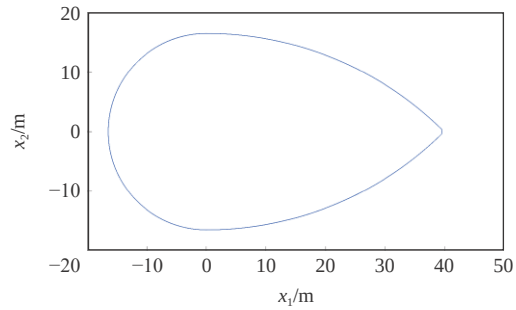


图 14　案例 1 的可达集

Fig. 14　Reachable set of case 1

在存在多个静态障碍物的情况下, 无人机完成避撞任务的飞行过程如图 15 所示。此结果是一个动态过程, 图中红色曲线表示无人机走过的路径点集, 蓝色封闭曲线为检测到障碍物及经过障碍物时实时提取的 BRS, 无人机完成避障任务后恢复原来的飞行方向。图 16 表明, 仿真过程中, 无人机与障碍物之间最小距离不小于 34 m(>16.5 m), 处于绝对安全状态。

从图 17 中可以看出, 无人机在躲避静态物体

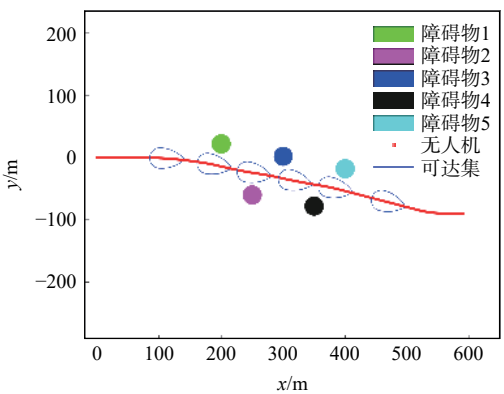


图 15　无人机躲避多个静态障碍物的过程

Fig. 15　Process of UAV avoiding multiple static obstacles

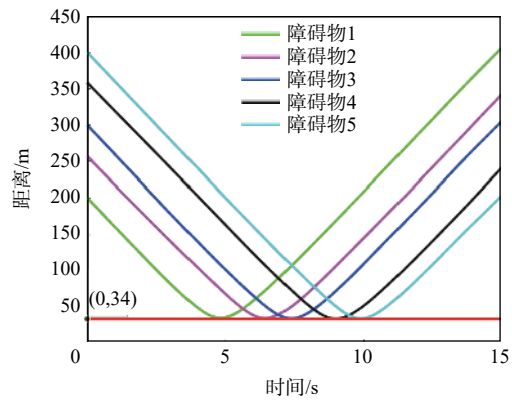


图 16　无人机与静态障碍物之间的距离

Fig. 16　Distance between UAV and static obstacles

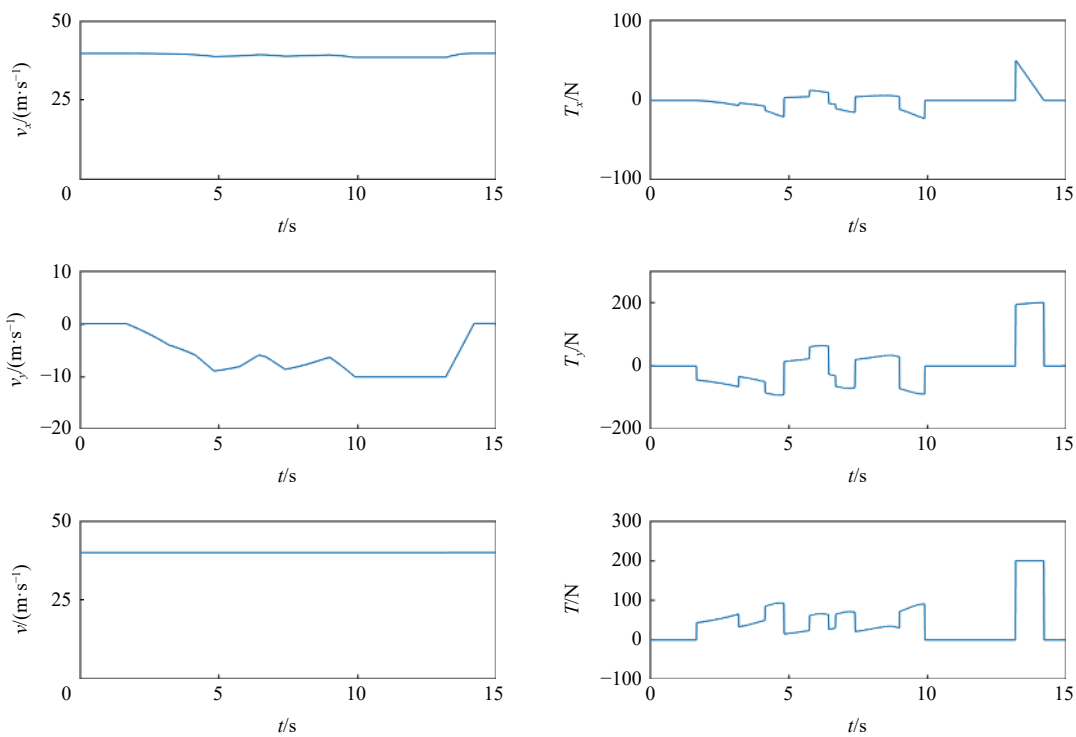


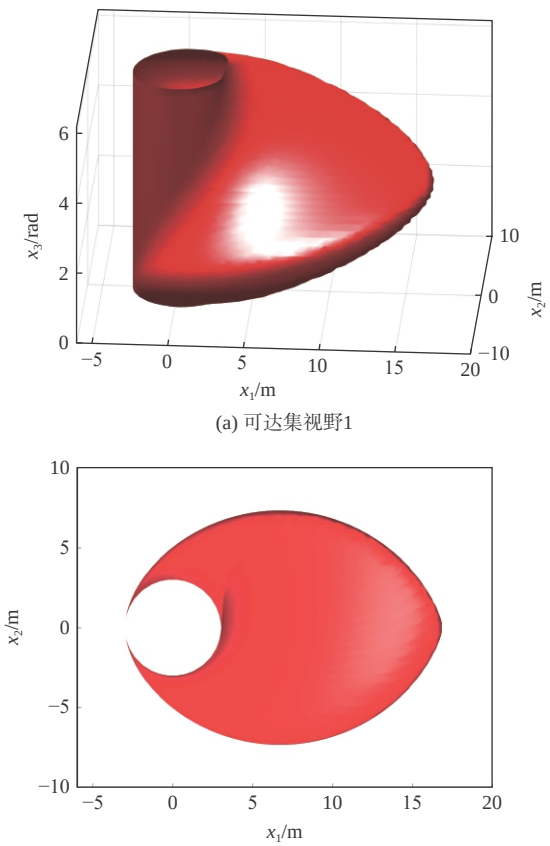
图 17　无人机速度和飞行平面内推力(案例 1)

Fig. 17　UAV speed and thrust in flight plane (case 1)

过程中没有改变自身速度大小，飞行平面内最大推力不超过 200 N。

3.2.2 案例 2：无人机与合作目标相互避撞

设定仿真环境条件：任务环境为城市低空小范围内的 2 机冲突场景，其区域范围为长 250 m、宽 180 m。任务区域中存在 2 架同样使用可达集避撞算法的无人机，且有相撞的趋势。2 架无人机平飞速度均为 $v=40$ m/s，传感器探测半径 $R=120$ m，目标集的碰撞半径 $r_{\text{safe}}=r+r=3$ m，可达集与 2 架无人机的相对位置、航向相对转角的关系如图 18 所示。



(a) 可达集视野1

(b) 可达集视野2

图 18 案例 2 的可达集

Fig. 18 Reachable set of case 2

无人机与合作目标无人机相互分离的飞行过程如图 19 所示，红线和蓝线分别表示 UAV1 和 UAV2 走过的路径点集。从图 19 可以看出，2 架无人机以平滑的路径飞行，并成功完成分离任务。在可检测范围（$R\leqslant 120$ m）内，2 架无人机均实时获取可达集（见图 19 中蓝色封闭曲线），以减小碰撞风险。图 20 表明，仿真过程中，2 架无人机之间的最小距离不小于 10 m（>3 m），处于绝对安全状态。

从图 21 和图 22 中可以看出，2 架无人机处于匀速率运动状态，且飞行平面内最大推力均不超过 500 N。

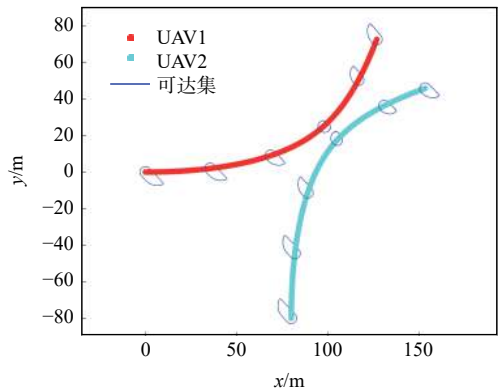


图 19 2 架无人机相互避撞的过程

Fig. 19 Process of two UAVs avoiding each other

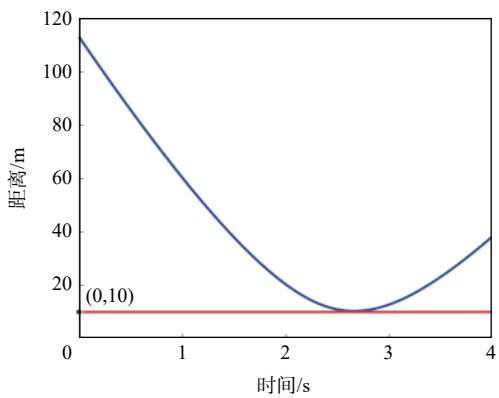


图 20 2 架无人机之间的距离

Fig. 20 Distance between two UAVs

3.2.3 案例 3：无人机躲避多个动态障碍物

设定仿真环境条件：任务环境为城市低空的局部区域，其长为 650 m，宽为 550 m。任务区域中存在未知有人机/无人机的碰撞威胁，通过地形建模将动态物体转化为多个 $r_{\text{obs}}=15$ m 的圆形障碍，并且会造成无人机与圆形障碍相撞。无人机平飞速度 $v=40$ m/s，传感器探测半径 $R=120$ m，目标集的碰撞半径 $r_{\text{safe}}=r+r_{\text{obs}}=16.5$ m，可达集与相对位置、航向相对转角的关系如图 23 所示。

图 24(a)～图 24(f)显示了在存在多个动态障碍物的飞行空域中，无人机完成避撞任务的全过程。与 3.2.1 节相同，红色曲线和蓝色封闭曲线分别表示无人机走过的路径点集和检测到障碍物及躲避障碍物时实时提取的二维可达集。

从图 25 中可看出，无人机在飞行过程中保持匀速率运动状态，且飞行平面内最大推力不超过 500 N。图 26 表明，仿真过程中，无人机与 3 个障碍物之间的最小距离不小于 33 m（>16.5 m），处于绝对安全状态。

以上 3 种案例仿真实验结果表明, 本文提出的可达集避撞算法适用于城市低空空域中的多种场景, 发挥出实时、安全、路径平滑等避撞性能优势, 适用于解决 UTM 中的冲突解脱问题。

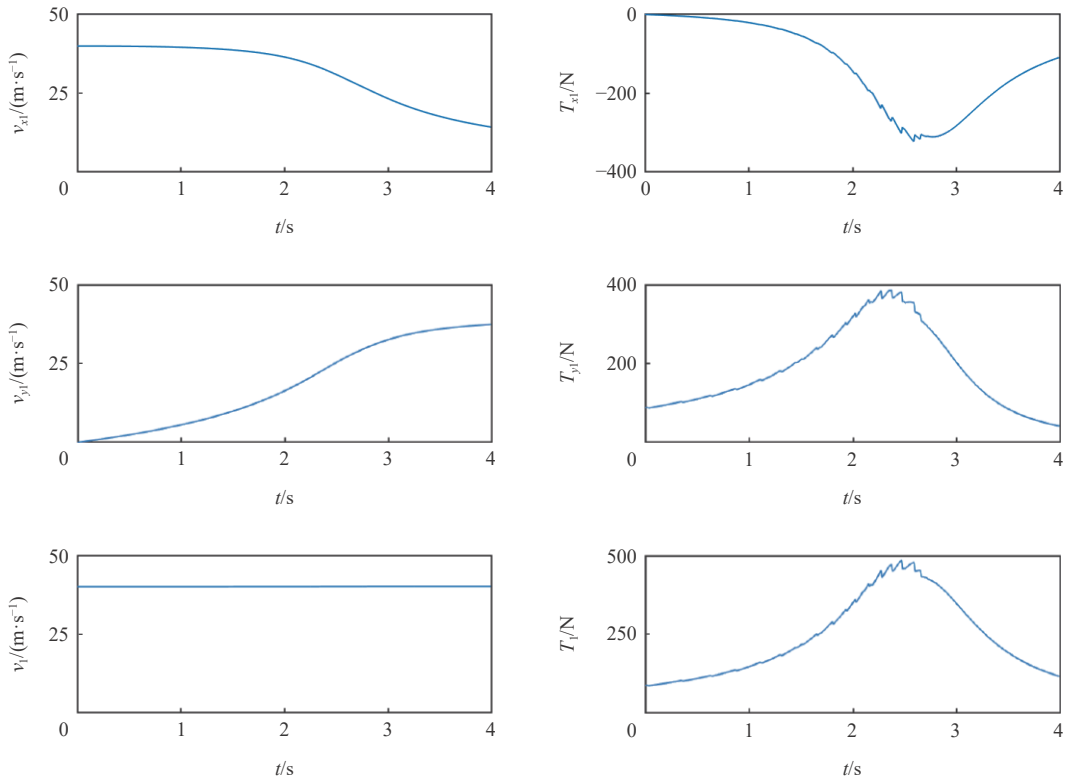


图 21　无人机 1 速度和飞行平面内推力(案例 2)

Fig. 21　UAV 1 speed and thrust in flight plane (case 2)

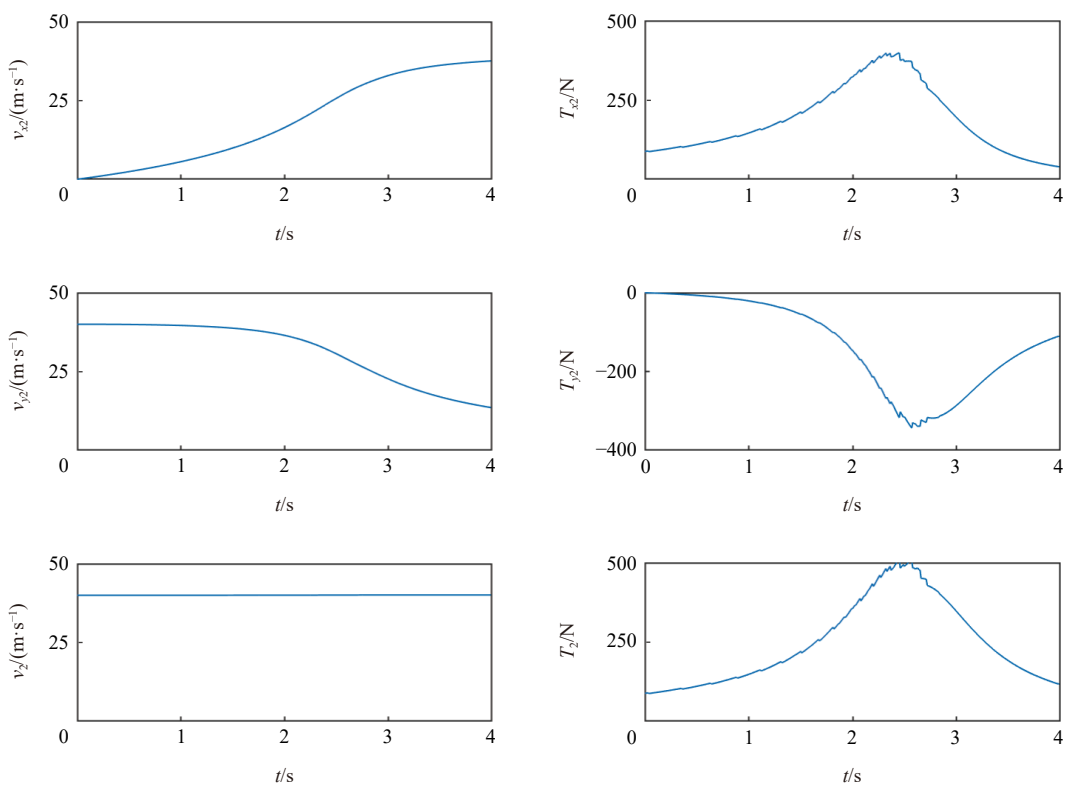
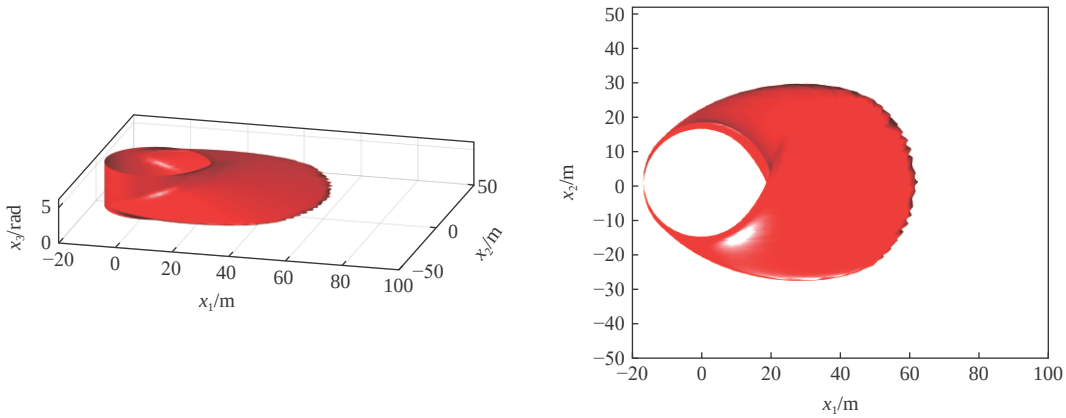


图 22　无人机 2 速度和飞行平面内推力(案例 2)

Fig. 22　UAV 2 speed and thrust in flight plane (case 2)



(a) 可达集视野1　　(b) 可达集视野2

图 23　案例 3 的可达集

Fig. 23　Reachable set of case 3

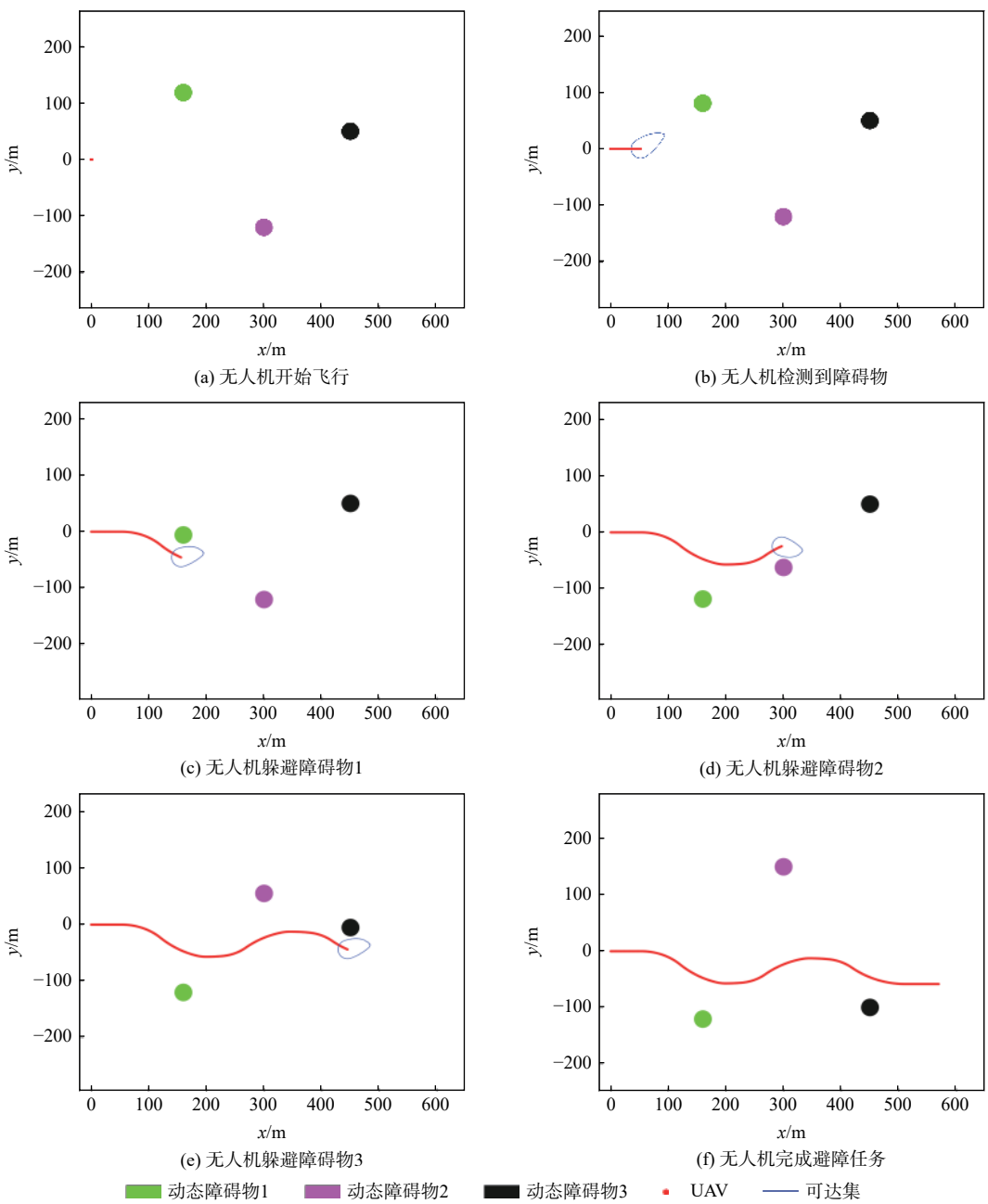


图 24　无人机躲避多个动态障碍物的过程

Fig. 24　Process of UAV avoiding multiple dynamic objects

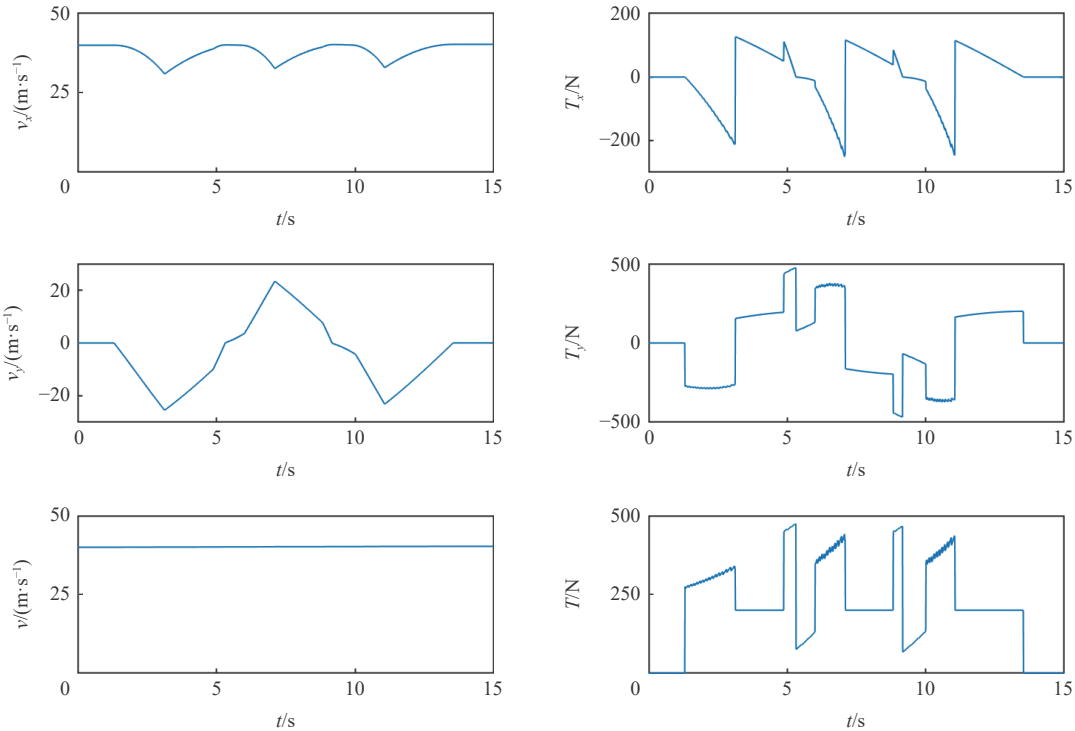


图 25　无人机速度和飞行平面内推力(案例 3)

Fig. 25　UAV speed and thrust in flight plane (case 3)

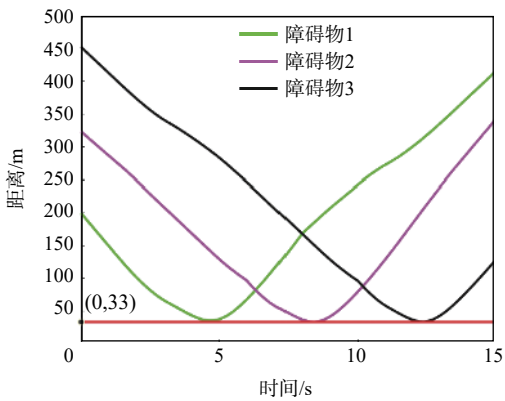


图 26　无人机与动态物体之间的距离

Fig. 26　Distance between UAV and dynamic objects

4　结　论

1) 以可达集分析为基础的实时避撞算法, 本质是对平面空域中 2 架飞机相对运动博弈问题进行分析, 通过可达集理论防止周围物体进入无人机可达范围, 从而达到避撞效果。

2) 算法使用后向可达集解决避撞问题, 发挥出后向可达集的多重优势。

3) 相比于传统的避撞算法, 该算法不需提前掌握无人机的全局飞行区域信息, 可以在陌生的任务环境中实现实时避障。

4) 可达集避撞算法保证了安全性, 针对合作/非合作目标均有效, 并且通过实验证明, 在避障过程中, 不会过度影响无人机的原始轨迹, 能够得到平滑的飞行路径。

基于以上优势, 该算法可以应用于 UTM 下交通流量密集的飞行空域。

本文实验场景主要讨论了平面空域中存在多个静态物体和动态物体时可达集避撞算法的性能表现。未来的研究会考虑将该算法扩展到多层空域, 在真实城市的三维低空环境下进行建模和仿真, 利用无人机的高度变化将其与周围物体分离, 使用可达集分析方法来保证空中交通的安全, 从而测试该算法的通用性。

参考文献 (References)

[1] BALACHANDRAN S, MUNOZ C, CONSIGLIO M C. Implicitly coordinated detect and avoid capability for safe autonomous operation of small UAS[C]//17th AIAA Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2017: 1-10.

[2] WU Y, LOW K H, HU X T. Trajectory-based flight scheduling for AirMetro in urban environments by conflict resolution[J]. Transportation Research Part C: Emerging Technologies, 2021, 131: 103355.

[3] LUNDBERG J, ARVOLA M, WESTIN C, et al. Cognitive work analysis in the conceptual design of first-of-a-kind systems — Designing urban air traffic management[J]. Behaviour & Information Technology, 2018, 37(9): 904-925.

[4] SCOTT D, RADMANESH M, SARIM M, et al. Distributed bidding-based detect-and-avoid for multiple unmanned aerial vehicles in national airspace[C]//2019 International Conference on Unmanned Aircraft Systems (ICUAS). Piscataway: IEEE Press, 2019:

930-936.

[5] ALI B S. Traffic management for drones flying in the city[J]. International Journal of Critical Infrastructure Protection, 2019, 26: 100310.

[6] BIJJAHALLI S, SABATINI R, GARDI A. Advances in intelligent and autonomous navigation systems for small UAS[J]. Progress in Aerospace Sciences, 2020, 115: 100617.

[7] HO F, GERALDES R, GONÇALVES A, et al. Pre-flight conflict detection and resolution for UAV integration in shared airspace: Sendai 2030 model case[J]. IEEE Access, 2019, 7: 170226-170237.

[8] HU J M, YANG X X, WANG W C, et al. UAS conflict resolution in continuous action space using deep reinforcement learning[C]// AIAA aviation 2020 FORUM. Reston: AIAA, 2020.

[9] ZHAO P, ERZBERGER H, LIU Y M. Multiple-aircraft-conflict resolution under uncertainties[J]. Journal of Guidance, Control, and Dynamics, 2021, 44(11): 2031-2049.

[10] 王泽坤, 吴明功, 温祥西, 等. 基于速度障碍法的飞行冲突解脱与恢复策略[J]. 北京航空航天大学学报, 2019, 45(7): 1294-1302.
WANG Z K, WU M G, WEN X X, et al. Flight collision resolution and recovery strategy based on velocity obstacle method[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(7): 1294-1302(in Chinese).

[11] DUCHAMP V, JOSEFSSON B, POLISHCHUK T, et al. Air traffic deconfliction using sum coloring[C]//2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC). Piscataway: IEEE Press, 2020: 1-6.

[12] MU L F, HAN S C. Satisficing game approach to conflict resolution for cooperative aircraft sharing airspace[J]. Big Data, 2021, 9(1): 53-62.

[13] 黄洋, 汤俊, 老松杨. 基于复杂网络的无人机飞行冲突解脱算法[J]. 航空学报, 2018, 39(12): 262-274.
HUANG Y, TANG J, LAO S Y. UAV flight conflict resolution algorithm based on complex network[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(12): 262-274(in Chinese).

[14] MIGLIACCIO G, MENGALI G, GALATOLO R, et al. A solution to detect and avoid conflicts for civil remotely piloted aircraft systems into non-segregated airspaces[J]. Proceedings of the Institution of Mechanical Engineers Part G:Journal of Aerospace Engineering, 2016, 230(9): 1655-1667.

[15] JOHNSON S C, PETZEN A, TOKOTCH D. Exploration of detect-and-avoid and well-clear requirements for small UAS maneuvering in an urban environment[C]//17th AIAA Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2017: 1-19.

[16] GREBE T, KUNZI F. Applications of conflict probes for detect and avoid systems[C]//2019 IEEE/AIAA 38th Digital Avionics Systems Conference (DASC). Piscataway: IEEE Press, 2020: 1-5.

[17] 刘畅, 王宏伦, 姚鹏, 等. 面向空中威胁的无人机动态碰撞区建模与分析[J]. 北京航空航天大学学报, 2015, 41(7): 1231-1238.
LIU C, WANG H L, YAO P, et al. Modeling and analysis of dynamic collision region for UAV avoiding aerial intruders[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(7): 1231-1238(in Chinese).

[18] KUCHAR J K, YANG L C. A review of conflict detection and resolution modeling methods[J]. IEEE Transactions on Intelligent Transportation Systems, 2000, 1(4): 179-189.

[19] 刘慧颖, 白存儒, 杨广珺. 无人机自主防撞关键技术与应用分析[J]. 航空工程进展, 2014, 5(2): 141-147.
LIU H Y, BAI C R, YANG G J. Application and analysis and discussion of autonomous collision avoidance techniques for unmanned aerial vehicle[J]. Advances in Aeronautical Science and Engineering, 2014, 5(2): 141-147(in Chinese).

[20] YU X, ZHANG Y M. Sense and avoid technologies with applications to unmanned aircraft systems: Review and prospects[J]. Progress in Aerospace Sciences, 2015, 74: 152-166.

[21] PARK J W, OH H D, TAHK M J. UAV collision avoidance based on geometric approach[C]//2008 SICE Annual Conference. Piscataway: IEEE Press, 2008: 2122-2126.

[22] WOLF T B, KOCHENDERFER M J. Aircraft collision avoidance using Monte Carlo real-time belief space search[J]. Journal of Intelligent & Robotic Systems, 2011, 64(2): 277-298.

[23] CETIN O, ZAGLI I, YILMAZ G. Establishing obstacle and collision free communication relay for UAVs with artificial potential fields[J]. Journal of Intelligent & Robotic Systems, 2013, 69(1): 361-372.

[24] YANG H, ZHAO Y J. Trajectory planning for autonomous aerospace vehicles amid known obstacles and conflicts[J]. Journal of Guidance, Control, and Dynamics, 2004, 27(6): 997-1008.

[25] FU S Y, HAN L W, TIAN Y, et al. Path planning for unmanned aerial vehicle based on genetic algorithm[C]//IEEE 11th International Conference on Cognitive Informatics and Cognitive Computing. Piscataway: IEEE Press, 2012: 140-144.

[26] CEKMEZ U, OZSIGINAN M, SAHINGOZ O K. Multi colony ant optimization for UAV path planning with obstacle avoidance[C]// 2016 International Conference on Unmanned Aircraft Systems (ICUAS). Piscataway: IEEE Press, 2016: 47-52.

[27] KARIMI J, POURTAKDOUST S H. Optimal maneuver-based motion planning over terrain and threats using a dynamic hybrid PSO algorithm[J]. Aerospace Science and Technology, 2013, 26(1): 60-71.

[28] LIN Y C, SARIPALLI S. Collision avoidance for UAVs using reachable sets[C]//2015 International Conference on Unmanned Aircraft Systems (ICUAS). Piscataway: IEEE Press, 2015: 226-235.

[29] 王亮, 魏铂淞, 熊瑜, 等. 基于边界归一化的低空无人机实时避撞路径规划[J]. 西北工业大学学报, 2017, 35(2): 213-219.
WANG L, WEI B S, XIONG Y, et al. Real-time route plan for UAV lower aerial collision avoidance based on boundary normalization[J]. Journal of Northwestern Polytechnical University, 2017, 35(2): 213-219(in Chinese).

[30] VANOORT E R. Adaptive back stepping control and safety analysis for modern fighter aircraft[D]. Delft: Delft University of Technology, 2011.

[31] LEUNG K, SCHMERLING E, ZHENG M X, et al. On infusing reachability-based safety assurance within probabilistic planning frameworks for human-robot vehicle interactions[J]. International Journal of Robotics Research, 2020, 39(10-11): 1326-1345.

[32] 陈文. 混合系统可达集计算方法研究[J]. 科技传播, 2012, 4(22): 125-126.
CHEN W. Research on computing method of reachable set of hybrid system[J]. Public Communication of Science & Technology, 2012, 4(22): 125-126(in Chinese).

[33] OSHER S, FEDKIW R. Level set methods and dynamic implicit

surfaces[M]. Berlin: Springer, 2003.

[34] MITCHELL I M, BAYEN A M, TOMLIN C J. A time-dependent Hamilton-Jacobi formulation of reachable sets for continuous dynamic games[J]. IEEE Transactions on Automatic Control, 2005, 50(7): 947-957.

[35] MITCHELL I M, TOMLIN C J. Overapproximating reachable sets by Hamilton-Jacobi projections[J]. Journal of Scientific Computing, 2003, 19(1-3): 323-346.

[36] MITCHELL I. Games of two identical vehicles[R]. Stanford: Stanford University,2001.

[37] MITCHELL I. A toolbox of level set methods[R]. Vancouver: The University of British Columbia, 2007.

[38] 管祥民, 吕人力. 基于满意博弈论的复杂低空飞行冲突解脱方法[J]. 航空学报, 2017, 38(S1): 120-128.
GUAN X M, LYU R L. Aircraft conflict resolution method based on satisfying game theory[J]. Acta Aeronautica et Astronautica Sinica, 2017, 38(S1): 120-128(in Chinese).

[39] WANG J, XIN M. Integrated optimal formation control of multiple unmanned aerial vehicles[J]. IEEE Transactions on Control Systems Technology, 2013, 21(5): 1731-1744.

Conflict resolution algorithms for UAV low-altitude flight based on reachable set

YANG Jianhang[1, 2], ZHANG Fubiao[1, 2, *], WANG Jiang[1, 2]

(1. School of Aerospace Engineering, Beijing institute of Technology, Beijing 100081, China;
2. Beijing Key Laboratory of UAV Autonomous Control Technology, Beijing 100081, China)

Abstract: Aiming at the conflict resolution problem in unmanned air traffic management (UTM), real-time collision avoidance algorithms based on reachability set analysis are proposed. To assure the secure operation of unmanned aerial vehicles (UAV) in low-altitude urban environments with dense traffic flow, these algorithms can be deployed. Based on the relative motion between the UAVs, the collision avoidance problem is modeled as a dynamic game problem in the 2D horizontal airspace, further the key concept of the reachable set of the UAV can be analyzed and calculated using the level set method and optimal control theory. Aided by airborne sensors, a new collision avoidance strategy for each drone is proposed using information about drones and surrounding objects. The technique can safely resolve the conflict resolution problem in real-time with a smooth flight route, according to simulation findings from three examples in varied airspace settings, and it is also successful against both cooperative and non-cooperative UAVs.

Keywords: unmanned air traffic management; conflict resolution; low-altitude urban environment; reachability set analysis; collision avoidance strategy; 2D horizontal airspace

Received: 2021-09-09; **Accepted**: 2021-12-17; **Published Online**: 2022-01-12 12:44

URL: kns.cnki.net/kcms/detail/11.2625.V.20220112.0910.002.html

Foundation items: National Key Research and Development Program of China (2021YFF0601304); National Natural Science Foundation of China (U1913602); Beijing Municipal Science & Technology Commission fund (Z181100003218013)

* **Corresponding author**. E-mail: fubiao.zhang@bit.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0544

遥感图像语义分割的空间增强注意力U型网络

宝音图[1,2]，刘伟[1,*]，李润生[1]，李钦[1]，胡庆[1]

(1. 中国人民解放军战略支援部队信息工程大学 数据与目标工程学院，郑州 450001；
2. 中国人民解放军31401部队，呼和浩特 010051)

摘 要：针对基于深度学习的语义分割模型在解析遥感图像时，小尺寸目标和目标边界存在分割不准确的问题，提出一种U型网络模型SGE-Unet。该模型通过优化网络结构加强模型的特征提取能力；融合空间组增强注意力，提升模型对上下文语义信息的解析能力；采用中值频率平衡交叉熵损失函数抑制类别分布不均衡的影响。在2个数据集上进行实验，SGE-Unet的整体准确率、平均交并比、$\overline{F}_1$分数和Kappa系数均高于主流模型，Vaihingen数据集中小尺寸目标车的交并比和F_1分数分别为0.719和0.901，比次优模型提升了16%和11%，实验结果表明所提模型能更精准地分割小尺寸目标及目标边界。

关 键 词：遥感图像；语义分割；深度学习；注意力；损失函数
中图分类号：TP751.1; V19
文献标志码：A **文章编号**：1001-5965(2023)07-1828-10

语义分割是图像分割的一个分支，主要任务是给图像的每个像素分配语义标签，可在像素级对图像分类标识，并细粒度地描述目标，已被广泛应用于基于遥感图像的土地覆盖利用分析、环境监测、灾害评估、城市规划和军事安全等任务[1]。近年来，随着深度学习的发展，受益于神经网络强大的表征能力，一些基于深度学习的语义分割模型在遥感图像解析中得到了较好效果[2]。根据研究对象的细粒程度，基于深度学习的遥感图像语义分割方法可分为基于区域的方法和基于像素的方法。基于区域的方法通常先将遥感图像按照一定策略划分区域，再通过深度学习模型分别辨识各区域信息并标注语义标签。划分区域主要依靠阈值、聚类、超像素和图论等方法，这些方法基于图像的颜色、纹理、光谱和位置等浅层次特征信息[3]，且需要设计人类工程特征或者专家知识支撑，因此基于区域的方法通常分割精度低、效率不高。基于像素的方法能够直接对遥感图像每个像素标注语义标签，实现端到端的像素级分割。2016年，Kampffmeyer等[4]将全卷积神经网络(fully convolutional network, FCN)[5]用于遥感图像语义分割，实现了深度学习模型在遥感图像上的像素级分割。在此基础上，Guo等[6]融合集成学习方法，提升遥感图像中小尺寸目标的分割精度。针对遥感图像目标方向尺寸多变的问题，Li[7]和Alam[8]等在U-Net[9]基础上使用不对称卷积块和索引池化提升模型对遥感图像中多方向目标的辨识能力。张小娟和汪西莉[10]利用多尺度特征融合增强模型对遥感图像目标尺寸的敏感度。Feng等[11]构建邻域像素亲和损失函数，使模型更精准地分割遥感图像的目标边界。肖春姣等[12]通过条件随机场捕获上下文信息，对遥感图像的分割结果优化处理。翟鹏博[13]和杨军[14]等采用空洞卷积增大感受野，将数字地表模型(digital surface model, DSM)信息和图像信息融合进行训练，构建双路径网络结构提升分割精度。

基于深度学习的语义分割在解析遥感图像上

收稿日期：2021-09-10；**录用日期**：2022-02-25；**网络出版时间**：2022-03-18 08:58
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220316.1023.001.html
基金项目：国家自然科学基金(41901378)
*通信作者. E-mail：greatliuliu@163.com

引用格式：宝音图，刘伟，李润生，等. 遥感图像语义分割的空间增强注意力U型网络[J]. 北京航空航天大学学报，2023，49(7)：1828-1837. BAO Y T，LIU W，LI R S，et al. Semantic segmentation of remote sensing images based on U-shaped network combined with spatial enhance attention[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1828-1837 (in Chinese).

取得了可观的成绩, 然而, 遥感图像地物信息丰富, 目标尺寸多变, 同类目标之间和不同类目标之间尺寸差别很大, 并且目标形状不规则, 这些因素影响小尺寸目标和目标边界的分割效果[11]。造成上述问题的原因可归结为: ①各类别目标的像素数量分布不均衡; ②目标边界信息在整幅图像中占比少。网络模型如果平等地计算像素类别信息, 就会导致小尺寸目标和目标边界信息被淹没。

针对上述分析, 为提升遥感图像中小尺寸目标和目标边界的分割效果, 本文提出了融合空间组增强注意力[15]的 U 型网络 (spatial group enhance and U-net, SGE-Unet) 结构用于遥感图像语义分割。为增强特征提取能力, 在编码和解码部分使用 EfficientNet[16] 做基础骨干网络, 引入空间组增强注意力机制加强对语义信息的理解。在训练阶段, 将中值频率平衡交叉熵作损失函数用于反馈训练, 抑制各类别像素数量分布不均衡的影响。在预测阶段, 用拼接策略消除大尺寸遥感图像分割拼接后预测结果的拼接痕迹。

1　原理与方法

1.1　整体网络模型设计

UNet++语义分割模型是一种 U 型网状的编码解码器结构, 其密集的跳跃连接可以获取更多语义信息[17]。在 UNet++基础上改进特征提取模块, 使用 EfficientNet[16] 作为基础结构增强特征提取能力, 在编码和解码部分添加空间组增强注意力模块, 进一步提升模型解析语义能力。SGE-Unet 网络结构如图 1 所示。图中向下的实线箭头表示 2 倍下采样的编码过程, 其中前 2 层为了保留背景信息采用了最大池化, 后 2 层为了提取特征信息采用了平均池化。向上的实线箭头表示 2 倍上采样的解码过程, 包括转置卷积和反池化; 虚线箭头表示跳跃连接, 将多张特征图进行融合, 丰富特征信息。经过最后 1 个卷积块后, 通过 Softmax 分类得到最终的分割结果。

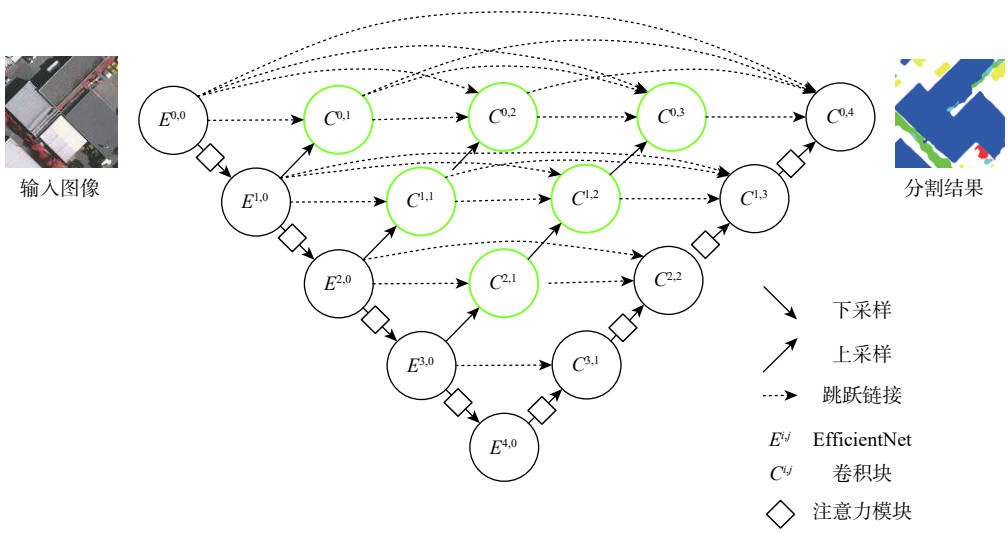


图 1　SGE-Unet 结构

Fig. 1　Structure of SGE-Unet

1.2　空间组增强注意力模块

遥感图像中不仅包含地物类别等有效信息, 也包含着云、雾、遮挡等干扰信息。对于特定的类别标签, 如果与原图中位置对应, 就能提取正确的语义特征, 如果由干扰等原因导致类别标签与原位置不能对应, 就会造成错误的学习, 混淆正确的语义特征, 导致分类困难[15]。同时, 在基于卷积神经网络的语义分割中, 下采样会导致小尺寸目标的特征信息进一步缩小甚至消失, 致使特征信息在上采样过程中较难恢复, 从而造成辨识困难[14]。注意力机制类似于人观察环境, 通过对特征图加权处理给予局部重要信息更多的关注, 有助于增强特征的表达能力, 降低噪声干扰, 加强模型对特征信息的解析[18]。针对遥感图像小尺寸目标和目标边界较难准确分割的情况, 采用空间组增强注意力[15]提升局部信息的特征表达能力。空间组增强注意力将特征图根据通道数量 C 均分为多组子特征图, 各组子特征图在空间的每个位置都有特征 $\boldsymbol{x}_i$, 则各组包含的特征为 $\boldsymbol{X}=[\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m]$($m$为特征图的宽和高的乘积), 组语义信息$\boldsymbol{g}$可通过空间平均函数衡量, 空间平均函数定义为

$$\boldsymbol{g}=F(\boldsymbol{X})=\frac{1}{m}\sum_{i=1}^{m}\boldsymbol{x}_i \quad \boldsymbol{X}=[\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m] \tag{1}$$

式中: $\boldsymbol{X}$为每组在空间中m个位置特征信息$\boldsymbol{x}_i$的合集; $F(\boldsymbol{X})$为空间平均函数。

将位置特征信息$\boldsymbol{x}_i$与该组语义信息$\boldsymbol{g}$相乘可得到位置特征信息的加权系数c_i:

$$c_i = \boldsymbol{g} \cdot \boldsymbol{x}_i = \|\boldsymbol{g}\|\|\boldsymbol{x}_i\|\cos\theta_i \tag{2}$$

式中: θ_i为位置特征信息$\boldsymbol{x}_i$与语义信息$\boldsymbol{g}$之间的夹角。显然位置特征信息$\boldsymbol{x}_i$越大, 更接近$\boldsymbol{g}$的方向(即θ_i更小), 该位置特征信息$\boldsymbol{x}_i$越重要, 因此, c_i可以有效衡量各位置特征信息的贡献情况。

为弱化加权系数c_i的数值影响, 需对c_i进行归一化处理:

$$\mu_c = \frac{1}{m}\sum_{j=1}^{m} c_j \tag{3}$$

$$\sigma_c = \frac{1}{m}\sum_{j=1}^{m} (c_j - \mu_c)^2 \tag{4}$$

$$\hat{c}_i = \frac{c_i - \mu_c}{\sigma_c} \tag{5}$$

式中: μ_c和σ_c分别为所有加权系数c_i的均值和标准差; $\hat{c}_i$为经过归一化处理的加权系数。

将$\hat{c}_i$与对应的位置特征信息$\boldsymbol{x}_i$相乘, 即可得到加权的位置特征信息$\hat{\boldsymbol{x}}_i$:

$$\hat{\boldsymbol{x}}_i = \boldsymbol{x}_i \cdot \hat{c}_i \tag{6}$$

组内所有的位置特征信息进行加权处理后, 即可获得增强后的组特征$\hat{\boldsymbol{X}} = [\hat{\boldsymbol{x}}_1, \cdots, \hat{\boldsymbol{x}}_m]$。因各组关注的区域都是独立的, 最终将增强的组特征融合, 即可更有效地关注目标区域。

空间增强注意力在充分利用全局信息的基础上, 根据语义信息的贡献程度赋予各组对应空间位置权重, 增强关注区域的语义特征学习, 更有利于在局部区域建立上下文联系, 从而更好地指导训练模型。

1.3 中值频率平衡交叉熵

损失函数可以评估模型预测值与样本真实值之间的差异度, 是指导模型学习的重要因素, 根据不同的任务需求选择合适的损失函数, 更有助于训练出性能更好的模型[19]。对于遥感图像语义分割任务, 像素级的类别标签会造成较大的类别分布不均衡。例如, 在国际摄影测量和遥感学会 (International Society for Photogrammetry and Remote Sensing, ISPRS) 发布的 Vaihingen 语义分割数据集[20]中, 不透水表面类别的像素数量超过车类别像素数量20倍。中值频率平衡交叉熵损失函数可以减弱类别分布不均衡的影响[21], 其可定义为

$$L = -\frac{1}{N}\sum_{n=1}^{N}\sum_{c=1}^{M} l_c^n \lg(\hat{p}_c^n) w_c \tag{7}$$

$$w_c = \frac{\text{median}(\{f_c | c \in M\})}{f_c} \tag{8}$$

$$l_c^n = \begin{cases} 0 & \text{样本}n\text{的真实值标签不为}c\text{类} \\ 1 & \text{样本}n\text{的真实值标签为}c\text{类} \end{cases} \tag{9}$$

式中: N为预测的样本数量; M为类别数量; w_c为类别c的权重; $\hat{p}_c^n$样本n预测为类别c的概率; f_c为类别c的像素频率; median 为取所有类别像素频率的中位数; l_c^n为符号函数。

由式 (8) 可知, 权重w_c以类别的像素频率作分子, 频率越高则权重越小, 从而平衡了类别分布不均衡的问题。

2 实验及结果分析

2.1 数据集及数据预处理

选用 ISPRS 发布的 Vaihingen 语义分割数据集[20]和 Potsdam 语义分割数据集[20]进行实验来验证本文网络模型的有效性。

Vaihingen 数据集用数字航拍方式采集自德国的 Vaihingen, 包含 33 张带有语义标签的遥感图像, 空间分辨率为 0.09 m, 图像尺寸分别为 1 000 像素×4 000 像素和 4 000 像素×1 000 像素, 遥感图像包含近红外 (IR)、红 (R) 和绿 (G)3 个波段。Potsdam 数据集用数字航拍方式采集自德国的 Potsdam, 包含 38 张带有语义标签的遥感图像, 空间分辨率为 0.05 m, 图像尺寸为 6 000 像素×6 000 像素, 遥感图像包含红 (R)、绿 (G) 和蓝 (B)3 个波段。每个数据集均被标注为不透水表面 (如道路, RGB: 255, 255, 255)、建筑 (RGB: 0, 0, 255)、树 (RGB: 0, 255, 255)、低值被 (RGB: 225, 0, 225)、车 (RGB: 255, 255, 0) 和背景 (不属于前 5 类的其他地物类别, RGB: 255, 0, 0) 共 6 种地物类别。这 2 个数据集都以城市场景为主, Vaihingen 数据集中有许多独立的建筑, 且与植被交错分布, Potsdam 数据集则有大量建筑群和狭窄的街道, 分布相对分散。

实验过程中, 参考文献 [22、23] 的数据集配置方式, 将数据集按表 1 进行分配, 表中数字为图像编号。

因选用数据集的遥感图像尺寸较大, 难以直接输入网络模型, 需先将遥感图像裁剪为多幅小尺寸子图像, 再分别输入网络模型进行预测, 而后拼接得到最终分割结果。如不采取任何措施, 有可能产生如图 2 所示的拼接痕迹。产生拼接痕迹的主要原因是: 原始遥感图像经过裁剪操作, 小尺寸子图像边缘的地物信息不完整, 导致在小尺寸子图像中丢失部分上下文信息。

表 1　数据集分配

Table 1　Allocation of dataset

类别	Vaihingen	Potsdam
训练集	1, 3, 11, 13, 15, 17, 21, 26, 28, 32, 34, 37	2_12, 3_10, 3_11, 3_12, 4_11, 4_12, 5_10, 5_12, 6_7, 6_8, 6_9, 6_10, 6_12, 7_7, 7_9, 7_10, 7_11, 7_12
验证集	5, 7, 23, 30	2_11, 4_10, 5_11, 7_8
测试集	2, 4, 6, 8, 10, 12, 14, 16, 20, 22, 24, 27, 29, 31, 33, 35, 38	2_10, 2_13, 2_14, 3_13, 3_14, 4_13, 4_14, 4_15, 5_13, 5_14, 5_15, 6_13, 6_14, 6_15, 7_13



遥感图像　　语义标签　　分割结果

图 2　拼接痕迹示意图

Fig. 2　Seam after splicing

为消除拼接痕迹, 采取一种拼接策略[11]。如图 3 所示, 对遥感图像以滑窗重叠[24]的方式裁剪成小尺寸子图像, 将小尺寸子图像经网络模型得到预测结果, 随后按顺序进行拼接。需要注意的是, 在拼接时忽略小尺寸子图像预测结果的边缘区域。图 3 中虚线区域是裁剪的尺寸, 即输入网络模型的尺寸, 实线区域是得到预测结果后保留的区域, b 为滑窗重叠的步幅。经实验对比分析, b 仅需要大于最小目标尺寸即可较好消除拼接痕迹。

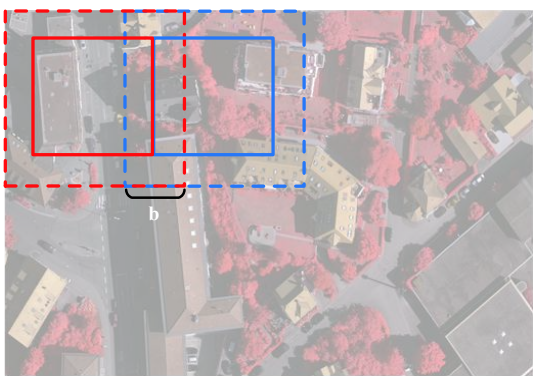


图 3　拼接策略

Fig. 3　Splicing method

实验中, 将 2 个数据集按滑窗重叠步长取 50 像素的方式裁剪为 512 像素×512 像素的子图像, 同时对训练集的图像通过旋转、缩放、翻转、颜色变换、添加噪声、随机擦除等方式扩充数据以提升模型的泛化能力。

2.2　评价标准

为了评估本文网络模型的性能, 使用整体准确率 OA、平均交并比 mIoU、$\overline{F}_1$分数和 Kappa 系数等评价指标进行分析。OA 表示预测正确像素占全部像素比例, mIoU 是各类 IoU 的平均, IoU 即预测结果与真值标签的交集与并集之比, 用于评价分割区域的完整性, $\overline{F}_1$分数则结合了精准率和召回率, Kappa 系数用于评价预测结果和标签真值的近似程度。这些评价指标侧重点不同, 综合各评价指标能更准确全面了解模型性能。相应计算公式为

$$\mathrm{OA}=\frac{\sum_{i=1}^{n}p_{ii}}{\sum_{i=1}^{n}\sum_{j=1}^{n}p_{ij}} \tag{10}$$

$$\mathrm{IoU}^{(k)}=\frac{p_{kk}}{\sum_{j=1}^{n}(p_{kj}+p_{jk})-p_{kk}} \tag{11}$$

$$\mathrm{mIoU}=\frac{1}{n}\sum_{i=1}^{n}\mathrm{IoU}^{(i)} \tag{12}$$

$$P^{(k)}=\frac{p_{kk}}{\sum_{j=1}^{n}p_{kj}} \tag{13}$$

$$R^{(k)}=\frac{p_{kk}}{\sum_{j=1}^{n}p_{jk}} \tag{14}$$

$$F_1^{(k)}=\frac{2P^{(k)}R^{(k)}}{P^{(k)}+R^{(k)}} \tag{15}$$

$$\overline{F}_1=\frac{1}{n}\sum_{i=1}^{n}F_1^{(i)} \tag{16}$$

$$\mathrm{Kappa}=\frac{\left(\sum_{i=1}^{n}\sum_{j=1}^{n}p_{ij}\right)\sum_{i=1}^{n}p_{ii}-\sum_{i=1}^{n}\left(\sum_{j=1}^{n}p_{ij}\sum_{j=1}^{n}p_{ji}\right)}{\left(\sum_{i=1}^{n}\sum_{j=1}^{n}p_{ij}\right)^2-\sum_{i=1}^{n}\left(\sum_{j=1}^{n}p_{ij}\sum_{j=1}^{n}p_{ji}\right)} \tag{17}$$

式中: n为类别数量; k表示第k类; p_{ij}表示真值为i类, 而预测值为j类的像素, 即假反例, 相应的p_{ii}为真正例, p_{ji}为假正例; P为精准率, 表示预测为正例的结果中预测对的比例; R为召回率, 表示真值为正例的结果中预测对的比例。

2.3　实验参数设置

实验平台环境为 Windows10 操作系统搭建 PyTorch 1.7 框架, CPU 型号为 12 核 Intel Xeon Gold 6226@

2.70 GHz, 配置 64 GB 内存, 搭载 2 块 NVIDIA 公司的 RTX5000 显卡, 显存为 2×16 GB, 使用 11.0 版本的 CUDA 及 8.2.0 版本的 cudnn 加速训练。

在训练过程中, 使用 Adam 优化器[25], 初始学习率为 0.00 1, 权重衰减为 0.000 5, 损失函数使用中值频率平衡交叉熵, 学习率衰减策略采用 Poly 衰减策略[26], 即用式 (18) 调整学习率:

$$\mathrm{lr}_{\mathrm{epoch}} = \mathrm{lr}_{\mathrm{eopch-1}}\left(1-\frac{\mathrm{epoch}}{\mathrm{max_epoch}}\right)^{0.9} \tag{18}$$

式中: $\mathrm{lr}_{\mathrm{epoch}}$为当前 epoch 的学习率; $\mathrm{lr}_{\mathrm{epoch-1}}$为前一次 epoch 的学习率; max_epoch为设置的最大 epoch。

实验中, max_epoch设为 200, 每个 epoch 后用验证集进行评估, 如果连续 10 个 epoch 的评价指标没有提升, 则终止训练。

2.4 结果分析

SGE-Unet 网络模型的全局分割结果如图 4 所示, 图 4(a) 遥感图像的第 1 行和第 2 行分别选自 Vaihingen 数据集和 Potsdam 数据集。观察图 4 可知, SGE-Unet 在 2 个数据集上的分割结果较好, 分割的目标边界清晰、可辨识性高, 没有明显的拼接痕迹。交错分布的树和建筑能准确分割, 对于大尺寸目标建筑和小尺寸目标车的分割效果都比较好, 易混淆的低植被和树分割结果也与标签相近, 甚至密集分布的车都能逐个区分。但也存在部分误检情况, 主要出现于目标边界分布情况复杂及阴影、遮挡等干扰因素较多区域。说明 SGE-Unet 的分割性能较好, 能准确地分割小尺寸目标和目标边界。

表 2 为 SGE-Unet 与当前遥感图像语义分割的主流模型 FCN[5]、SegNet[21]、DeepLabV3[26]、SCAttNetV2[27] 和 UNet++[17] 在 Vaihingen 数据集上的语义分割结果。如表 2 所示, SGE-Unet 在 OA、mIoU、$\overline{F}_1$分数和 Kappa系数这 4 项综合评价指标上都达到

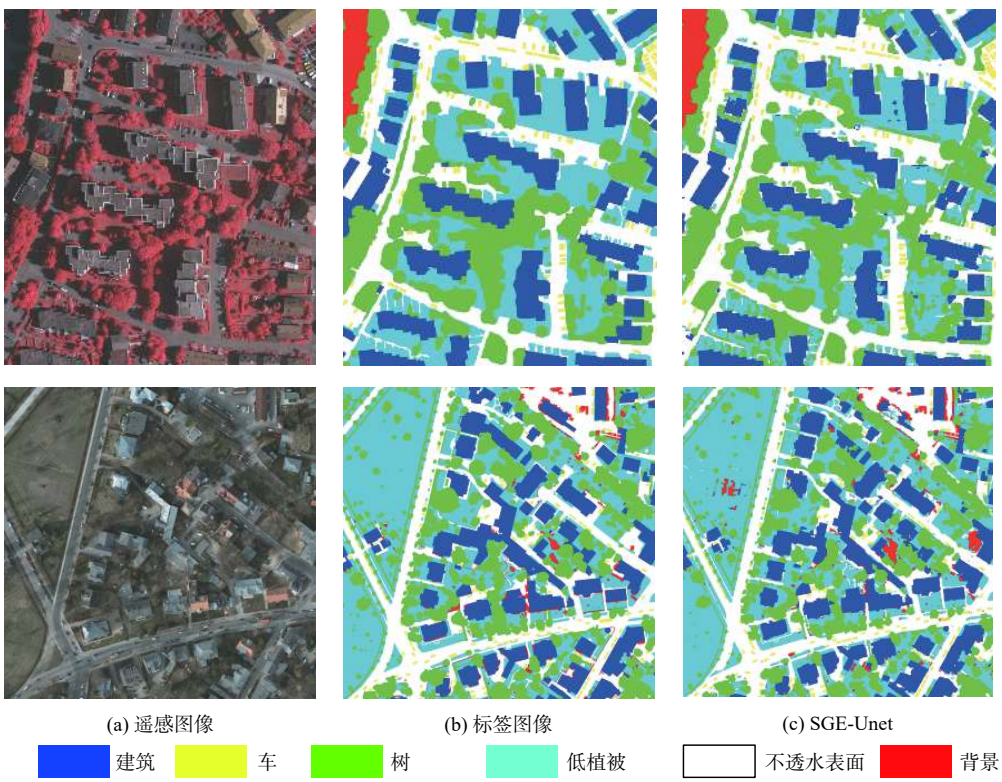


图 4 SGE-Unet 全局分割结果

Fig. 4 Global segmentation results of SGE-Unet

表 2 Vaihingen 数据集上的语义分割结果

Table 2 Semantic segmentation result on Vaihingen dataset

模型	IOU					F_1					OA	mIoU	$\overline{F}_1$	Kappa
	不透水表面	建筑	低植被	树	车	不透水表面	建筑	低植被	树	车				
FCN[5]	0.669	0.761	0.506	0.697	0.634	0.801	0.864	0.687	0.822	0.768	0.823	0.653	0.788	0.783
SegNet[21]	0.753	0.805	0.656	0.705	0.458	0.859	0.892	0.792	0.827	0.629	0.845	0.675	0.800	0.791
DeepLabV3[26]	0.822	0.912	0.711	0.770	0.568	0.902	0.954	0.831	0.870	0.724	0.858	0.757	0.856	0.809
SCAttNetV2[27]	0.804	0.823	0.667	0.671	0.544	0.891	0.903	0.801	0.803	0.705	0.855	0.702	0.821	0.788
UNet++[17]	0.783	0.878	0.614	0.716	0.619	0.952	0.952	0.817	0.931	0.809	0.854	0.722	0.892	0.808
SGE-Unet	0.803	0.890	0.706	0.830	0.719	0.935	0.966	0.934	0.941	0.901	0.866	0.790	0.935	0.824

了最高结果, 说明有较好的分割效果。在各类型目标中, SGE-Unet的 IoU 和 F_1 分数达到最高或次高结果, 优于其他模型, 尤其是小尺寸目标车的 IoU 和 F_1 分数比次高模型有近 0.1 的提升, IoU 提升了 16%, F_1 分数提升了 11%, 远高于其他模型。在树和低植被这 2 类易混淆目标上的 F_1 分数最高。同主流模型比较说明, SGE-Unet 能更精准分割目标, 特别是小尺寸目标和易混淆目标。

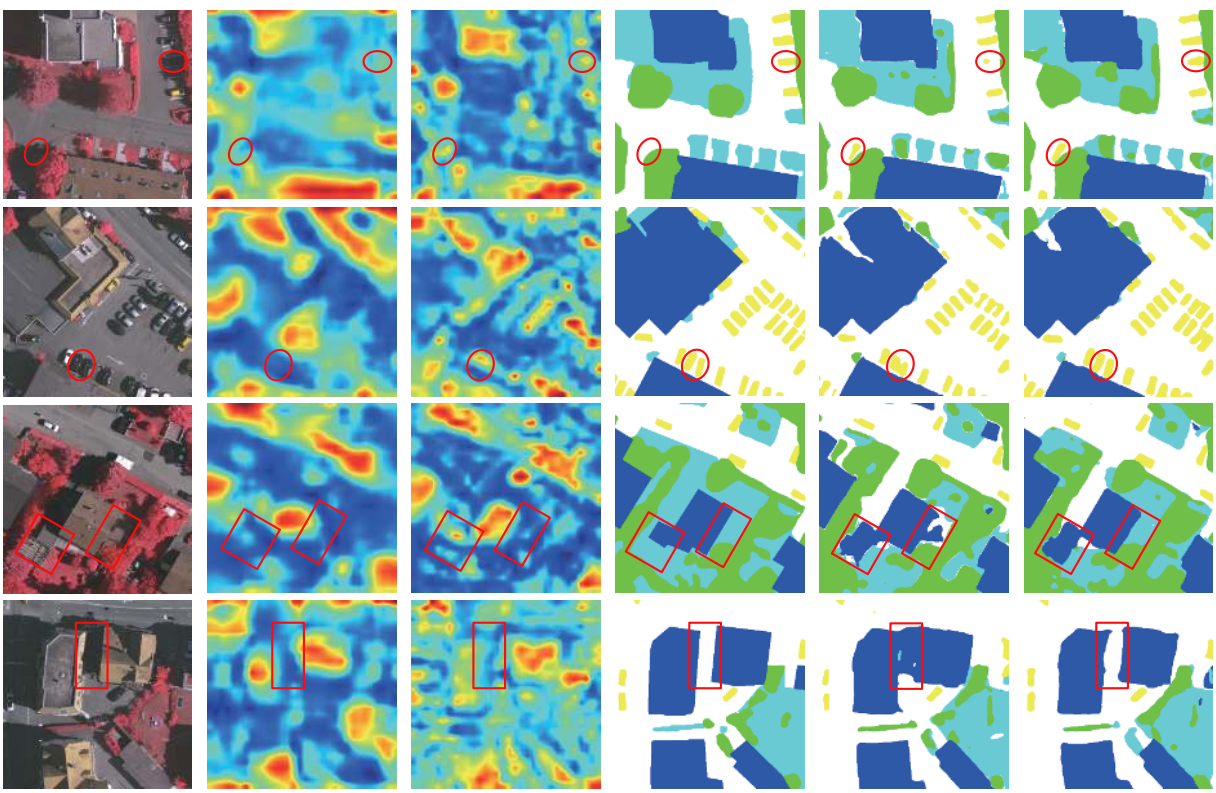
表 3 为各网络模型在 Potsdam 数据集上的语义分割结果。可以看到, 与在 Vaihingen 数据集上的结果一样, SGE-Unet 模型在综合评价指标和各类单项评价指标上都能获得较好成绩, 仅不透水表面和建筑的 IoU 得到次高结果, 但与最高结果相差不到 0.04, 基本持平。SGE-Unet 在 2 个不同的数据集上都能获得较好分割结果, 说明模型的泛化能力较强。

表 3　Potsdam 数据集上的语义分割结果

Table 3　Semantic segmentation result on Potsdam dataset

模型	IOU					F_1					OA	mIoU	$\overline{F}_1$	Kappa
	不透水表面	建筑	低植被	树	车	不透水表面	建筑	低植被	树	车				
FCN[5]	0.776	0.799	0.720	0.670	0.797	0.874	0.889	0.837	0.820	0.887	0.808	0.752	0.861	0.781
SegNet[21]	0.854	0.912	0.786	0.757	0.881	0.921	0.954	0.880	0.862	0.937	0.811	0.838	0.911	0.782
DeepLabV3[26]	0.882	0.944	0.789	0.745	0.871	0.938	0.971	0.882	0.854	0.931	0.835	0.846	0.915	0.799
SCAttNetV2[27]	0.818	0.888	0.707	0.663	0.803	0.901	0.941	0.829	0.797	0.891	0.879	0.776	0.872	0.805
UNet++[17]	0.864	0.923	0.781	0.767	0.865	0.941	0.974	0.895	0.883	0.933	0.872	0.840	0.925	0.828
SGE-Unet	0.879	0.940	0.815	0.795	0.891	0.955	0.982	0.905	0.957	0.969	0.883	0.864	0.954	0.843

图 5 为 SGE-Unet 和 UNet++在 Vaihingen 数据集上的特征热图及局部分割结果。对比 SGE-Unet 和 UNet++的特征热图, SGE-Unet 更多集中于车等小尺寸目标, 且可以辨识出部分建筑和车的轮廓, 说明 SGE-Unet 更关注小尺寸目标和目标边界。观察局部分割结果, 4 个场景中 SGE-Unet 结果与遥感图像和标签的大部分区域相近, 尤其是车、建筑和不透水表面 3 类基本完全一致。特别是在图 5 的红色标注区域, 图 5 (a) 遥感图像的第 1 行图中被阴影遮盖的车、图 5 (a) 遥感图像的第 2 行图中紧密分布的车、图 5 (a) 遥感图像的第 3 行图中建筑边界及图 5 (a) 遥感图像的第 4 行图中被阴影遮盖的道



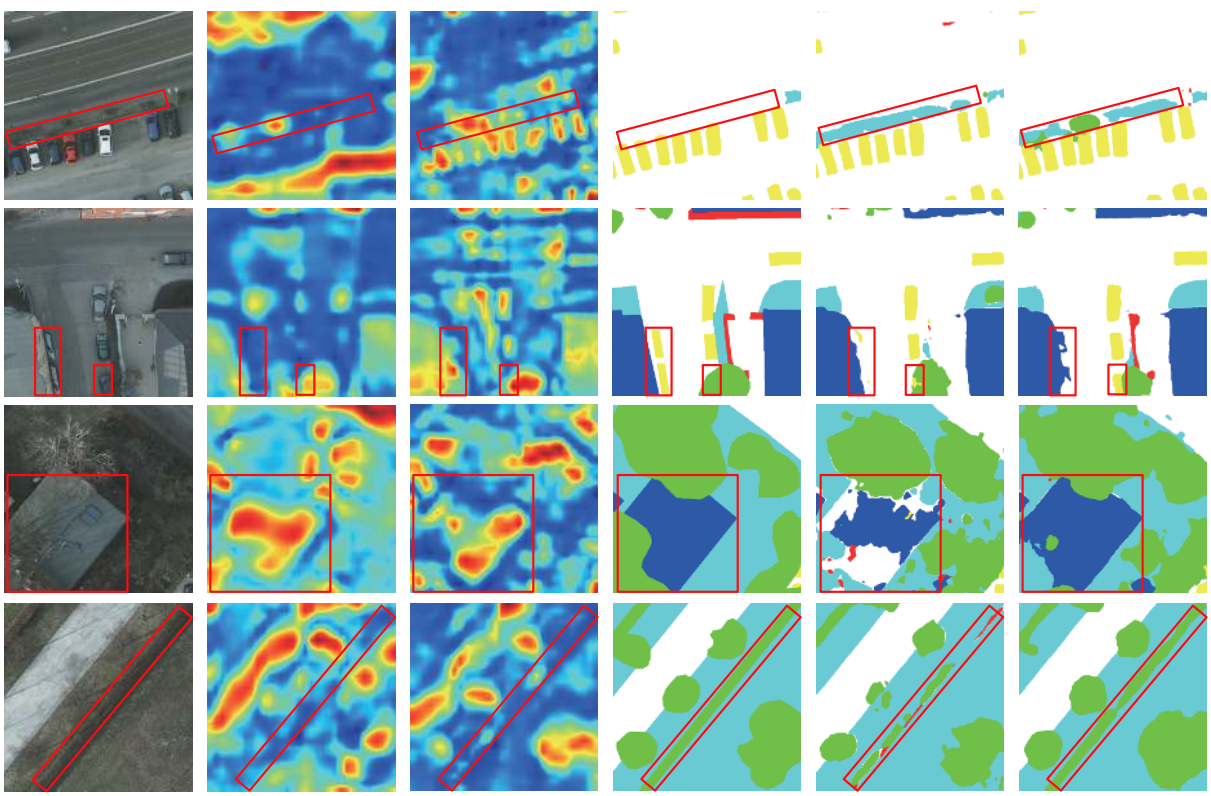
(a) 遥感图像　(b) UNet++ 特征热图　(c) SGE-Unet 特征热图　(d) 标签图像　(e) UNet++结果　(f) SGE-Unet 结果

图 5　Vaihingen 数据集上的特征热图及局部分割结果

Fig. 5　Feature heat map and local segmentation results on Vaihingen dataset

路, SGE-Unet 分割结果更好, 在小尺寸目标的完整性和目标边界的精细度方面 SGE-Unet 明显优于 UNet++。

图 6 为 SGE-Unet 和 UNet++在 Potsdam 数据集上的特征热图及局部分割结果, 为了验证模型性能, 所选场景包含目标被遮挡、目标模糊等情况。与在 Vaihingen 数据集上一样, 特征热图和分割结果表明, 相比于 UNet++, SGE-Unet 在小尺寸目标和目标边界上的分割效果更好, 分割结果更清晰、更精准。但是也有部分区域值得关注, 对于图 5 (a) 遥感图像的第 1 行图中左下红色椭圆形标注区域的不透水表面在图 5 (f)SGE-Unet 结果的第 1 行图中被识别为车, 从图 5 (a) 遥感图像的第 1 行图中可发现该处有被阴影遮盖的车, 应当为标签错误。对于图 5 (a) 遥感图像的第 3 行图, 左侧红色方框标注区域的低植被在图 5(f)SGE-Unet 结果的第 3 行图中被识别为建筑, 从图 5 (a) 遥感图像的第 3 行图可判断出该处为建筑, 应为标签错误; 右侧红色方框标注区域中图 5 (f)SGE-Unet 结果的第 3 行图中的分割结果虽然与图 5 (d) 标签图像的第 3 行图的标签一致, 但从图 5 (a) 遥感图像的第 3 行图可知该区域地物信息较为复杂且被阴影遮盖, 观察分析图 5 (e) UNet++结果的第 3 行图的分割结果应当更符合实际情况。对于图 6 (a) 遥感图像的第 1 行图, 红色斜方框标注区域的部分不透水表面在图 6 (f)SGE-Unet 结果的第 3 行图中被识别为树或者低植被, 从图 6 (a) 遥感图像的第 1 行图发现该标识区域有几株不明显的植被, 不属于不透水表面类型, 显然是标签错误, 且 SGE-Unet 的分割结果更准确。对于图 6 (a) 遥感图像的第 2 行图, 左侧红色方框标注区域的车在图 6 (f)SGE-Unet 结果的第 2 行图中被识别为建筑, 观察图 6 (a) 遥感图像的第 2 行图可知该处确有车辆, 由于成像过于扭曲, 导致模型无法辨认; 右侧红色方框标注区域的树在图 6 (f)SGE-Unet 结果的第 2 行图中被识别为车, 从图 6 (a) 遥感图像的第 2 行图中可以发现被树遮挡的车, 应属于标签错误。2 个不同数据集上的特征热图及局部分割结果进一步说明 SGE-Unet 模型的泛化性能较好。



(a) 遥感图像 (b) UNet++ 特征热图 (c) SGE-Unet 特征热图 (d) 标签图像 (e) UNet++结果 (f) SGE-Unet 结果

图 6 Potsdam 数据集上的特征热图及局部分割结果

Fig. 6 Feature heat map and local segmentation results on Potsdam dataset

表 4 和表 5 分别为在 Vaihingen 数据集和 Potsdam 数据集上的消融实验。可以看到, SGE-Unet 在大部分综合评价指标和各类单项评价指标上相较于 UNet++有明显提升, 特别是对于小尺寸目标车的 IoU 和 F_1 提升较为显著。消融实验结果说明, SGE-Unet 的改进有效提升了语义分割效果, 尤其是小尺寸目标。

表 6 列出了各网络模型的参数量和计算复杂度对比。模型参数量直接影响计算机的硬件消耗,

表 4　Vaihingen 数据集上的消融实验

Table 4　Ablation experiments result on Vaihingen dataset

模型	IOU					F_1					OA	mIoU	$\overline{F}_1$	Kappa
	不透水表面	建筑	低植被	树	车	不透水表面	建筑	低植被	树	车				
UNet++	0.783	0.877	0.613	0.716	0.619	0.973	0.952	0.817	0.930	0.809	0.853	0.722	0.896	0.808
UNet++&EfficientNet	0.815	0.911	0.627	0.721	0.655	0.976	0.967	0.823	0.937	0.832	0.866	0.746	0.907	0.825
UNet++&EfficientNet&SGE	0.813	0.901	0.626	0.722	0.714	0.971	0.963	0.826	0.938	0.894	0.867	0.755	0.918	0.824

表 5　Potsdam 数据集上的消融实验

Table 5　Ablation experiments result on Potsdam dataset

模型	IOU					F_1					OA	mIoU	$\overline{F}_1$	Kappa
	不透水表面	建筑	低植被	树	车	不透水表面	建筑	低植被	树	车				
UNet++	0.811	0.905	0.695	0.666	0.804	0.965	0.974	0.895	0.883	0.931	0.872	0776	0.930	0.828
UNet++&EfficientNet	0.829	0.918	0.711	0.691	0.810	0.960	0.986	0.892	0.909	0.943	0.879	0.792	0.938	0.839
UNet++&EfficientNet&SGE	0.826	0.915	0.709	0.694	0.823	0.955	0.982	0.898	0.913	0.970	0.883	0.793	0.944	0.843

表 6　各模型的参数量和计算复杂度对比

Table 6　Comparison of parameters and computational complexity of different modes

模型	参数量/MB	计算复杂度 GFLOPS	OA		$\overline{F}_1$	
			Vaihingen	Potsdam	Vaihingen	Potsdam
FCN[5]	21.30	19.20	0.823	0.808	0.621	0.757
SegNet[21]	18.82	117.74	0.845	0.811	0.675	0.838
DeepLabV3[26]	26.01	109.34	0.858	0.835	0.757	0.846
UNet++[17]	26.79	73.77	0.854	0.872	0.722	0.840
SGE-Unet	20.01	39.13	0.866	0.883	0.790	0.864

参数量越少, 对硬件的需求也越低。浮点运算量则影响模型的运算速度, 浮点运算量越少, 运算速度越快。SGE-Unet 的评价指标得分最高, 同时参数量较少, 计算复杂度较低, 说明模型对计算机硬件的性能要求不高。

3　结　论

通过研究基于深度学习的遥感图像语义分割方法, 提出 SEG-Unet 模型提升小尺寸目标和目标边界分割效果, 通过实验分析可以得出如下结论:

1) SGE-Unet 模型在 Vaihingen 数据集上实验结果的 OA、mIoU、$\overline{F}_1$分数和 Kappa 系数分别达到 0.866、0.790、0.935 和 0.824, 与主流模型相比均为最高值。

2) SGE-Unet 模型可以提升小尺寸目标和目标边界的分割效果, 对易混淆目标有更好区分能力, 甚至有遮挡、阴影等干扰的情况下也具备一定的辨识能力。

3) SGE-Unet 模型对计算机硬件性能要求较低。

参考文献（References）

[1] YUAN X H, SHI J F, GU L C. A review of deep learning methods for semantic segmentation of remote sensing imagery[J]. Expert Systems with Applications, 2021, 169: 114417.

[2] XING S, XIE Q, WANG M. Semantic segmentation for remote sensing images based on adaptive feature selection network[J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 8006705.

[3] 蒋晨琛, 霍宏涛, 冯琦. 一种基于PCA的面向对象多尺度分割优化算法[J]. 北京航空航天大学学报, 2020, 46(6): 1192-1203.
JIANG C C, HUO H T, FENG Q. An object-oriented multi-scale segmentation optimization algorithm based on PCA[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(6): 1192-1203(in Chinese).

[4] KAMPFFMEYER M, SALBERG A B, JENSSEN R. Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks[C]// IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE Press, 2016: 680-688.

[5] SHELHAMER E, LONG J, DARRELL T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651.

[6] GUO R, LIU J B, LI N, et al. Pixel-wise classification method for high resolution remote sensing imagery using deep neural

networks[J]. ISPRS International Journal of Geo-Information, 2018, 7(3): 110.

[7] LI R, DUAN C X, ZHENG S Y, et al. MACU-Net for semantic segmentation of fine-resolution remotely sensed images[J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 8007205.

[8] ALAM M, WANG J F, CONG G P, et al. Convolutional neural network for the semantic segmentation of remote sensing images[J]. Mobile Networks and Applications, 2021, 26(1): 200-215.

[9] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]// International Conference on Medical Image Computing and Computer-Assisted Intervention. Berlin: Springer, 2015: 234-241.

[10] 张小娟, 汪西莉. 完全残差连接与多尺度特征融合遥感图像分割[J]. 遥感学报, 2020, 24(9): 1120-1133.
ZHANG X J, WANG X L. Image segmentation models of remote sensing using full residual connection and multiscale feature fusion[J]. Journal of Remote Sensing, 2020, 24(9): 1120-1133(in Chinese).

[11] FENG Y, DIAO W, SUN X, et al. NPALoss: Neighboring pixel affinity loss for semantic segmentation in high-resolution aerial imagery[J]. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2020, V-2-2020(I-3): 475-482.

[12] 肖春姣, 李宇, 张洪群, 等. 深度融合网结合条件随机场的遥感图像语义分割[J]. 遥感学报, 2020, 24(3): 254-264.
XIAO C J, LI Y, ZHANG H Q, et al. Semantic segmentation of remote sensing image based on deep fusion networks and conditional random field[J]. Journal of Remote Sensing, 2020, 24(3): 254-264(in Chinese).

[13] 翟鹏博, 杨浩, 宋婷婷, 等. 结合注意力机制的双路径语义分割[J]. 中国图象图形学报, 2020, 25(8): 1627-1636.
ZHAI P B, YANG H, SONG T T, et al. Two-path semantic segmentation algorithm combining attention mechanism[J]. Journal of Image and Graphics, 2020, 25(8): 1627-1636(in Chinese).

[14] 杨军, 于茜子. 结合Atrous卷积的FuseNet变体网络高分遥感影响语义分割[J]. 武汉大学学报 (信息科学版), 2022, 47(7): 1071-1080.
YANG J, YU X Z. Semantic segmentation of high-resolution remote sensing images based on improved FuseNet combined with the Atrous convolution[J]. Geomatics and Information Science of Wuhan University, 2022, 47(7): 1071-1080(in Chinese).

[15] WANG X Y, CUI Z Y, CAO Z J, et al. Dense docked ship detection via spatial group-wise enhance attention in SAR images[C]// IEEE International Geoscience and Remote Sensing Symposium. Piscataway: IEEE Press, 2021: 1244-1247.

[16] TAN M X, LE Q V. EfficientNet: Rethinking model scaling for convolutional neural networks[C]// IEEE International Conference on Machine Learning (ICML).Piscataway: IEEE press, 2019: 6105-6114.

[17] ZHOU Z W, SIDDIQUEE M M R, TAJBAKHSH N, et al. UNet++: A nested U-net architecture for medical image segmentation[C]// International Workshop on Deep Learning in Medical Image Analysis, International Workshop on Multimodal Learning for Clinical Decision Support. Berlin: Springer, 2018: 3-11.

[18] 李道纪, 郭海涛, 卢俊, 等. 遥感影像地物分类多注意力融和U型网络法[J]. 测绘学报, 2020, 49(8): 1051-1064.
LI D J, GUA H T, LU J, et al. A remote sensing image classification procedure based on multilevel attention fusion U-Net[J]. Acta Geodaetica et Cartographica Sinica, 2020, 49(8): 1051-1064(in Chinese).

[19] 言有三. 深度学习之图像识别: 核心技术与案例实战[M]. 北京: 机械工业出版社, 2019: 231-232.
YAN Y S. Image recognition by deep learning: Core technologies and practices[M]. Beijing: China Machine Press, 2019: 231-232(in Chinese).

[20] ROTTENSTEINER F, SOHN G, JUNG J, et al. The ISPRS benchmark on urban object classification and 3d building reconstruction[J]. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2012(I-3): 293-298.

[21] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495.

[22] CHAI D F, NEWSAM S, HUANG J F. Aerial image semantic segmentation using DCNN predicted distance maps[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 161: 309-322.

[23] XU Z Y, SU C, ZHANG X C. A semantic segmentation method with category boundary for land use and land cover (LULC) mapping of very-high resolution (VHR) remote sensing image[J]. International Journal of Remote Sensing, 2021, 42(8): 3146-3165.

[24] 胡伟, 高博川, 黄振航, 等. 树形结构卷积神经网络优化的城区遥感图像语义分割[J]. 中国图象图形学报, 2020, 25(5): 1043-1052.
HU W, GAO B C, HUANG Z H, et al. Semantic segmentation of urban remote sensing image based on optimized tree structure convolutional neural network[J]. Journal of Image and Graphics, 2020, 25(5): 1043-1052(in Chinese).

[25] KINGMA D, BA J. Adam: A method for stochastic optimization[C]// International Conference on Learning Representations (ICLR). [S.l.]: ICLR, 2015.

[26] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834-848.

[27] LI H F, QIU K J, CHEN L, et al. SCAttNet: Semantic segmentation network with spatial and channel attention mechanism for high-resolution remote sensing images[J]. IEEE Geoscience and Remote Sensing Letters, 2021, 18(5): 905-909.

Semantic segmentation of remote sensing images based on U-shaped network combined with spatial enhance attention

BAO Yintu[1, 2]，LIU Wei[1, *]，LI Runsheng[1]，LI Qin[1]，HU Qing[1]

(1. School of Data and Target Engineering，PLA Strategic Support Force Information Engineering University，Zhengzhou 450001，China；
2. Unit 31401 of PLA，Hohhot 010051，China)

Abstract：The performance of semantic segmentation based on deep learning still need to be improved when analyzing small-sized objects and object boundaries in remote sensing images. Aiming at this problem, we propose a U-shaped network (SGE-Unet). Firstly, the structure of the model is optimized to enhance the representation of feature. Secondly, we add the attention module of spatial group enhance to extract semantic information. Finally, the median frequency balance cross-entropy loss function is used to suppress the unbalanced distribution of classes. The experiment was conducted on two datasets and shows that the overall accuracy,mean interaction over union, $\overline{F}_1$, and Kappa of SGE-Unet are better than mainstream models. In experiments of the Vaihingen dataset, the interaction over union and F_1 of the car reached 0.719 and 0.901, which were 16% and 11% higher than those of the model with the second-highest performance. The experimental results show that the proposed module greatly improves the segmentation of easily confused objects, small-sized objects, and object boundaries.

Keywords：remote sensing image; semantics segmentation; deep learning; attention; loss function

Received：2021-09-10；**Accepted**：2022-02-25；**Published Online**：2022-03-18 08：58

URL：kns.cnki.net/kcms/detail/11.2625.V.20220316.1023.001.html

Foundation item：National Natural Science Foundation of China (41901378)

* **Corresponding author**. E-mail：greatliuliu@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0553

芯片间时间触发消息堆叠调度方法

臧光界[1,2]，李峭[2,*]，王彤[2]，熊华钢[2]

(1. 中国航空研究院，北京 100029；2. 北京航空航天大学 电子信息工程学院，北京 100191)

摘 要：时间触发（TT）通信方式应用于芯片间互连网络，以保证航空电子通信任务之间消息传递的严格时间确定性。当航空电子任务具有多种操作模式，属于不同模式的芯片间的时间触发调度表会重叠占用时隙，提出芯片间时间触发消息堆叠调度方法，以提高利用网络资源的灵活性和效率，同时减小应用层消息由于等待时间触发时间窗的排队延迟。仿真实验表明：与超调度方法相比，所提方法能够减小芯片间互连网络中时间触发消息的总端到端延迟和链路平均时隙占用率，对于端到端延迟时间较长且链路平均承载消息传输较多的场景，采用所提方法减少端到端延迟的效果更显著。

关 键 词：芯片间互连；时间触发消息；操作模式；堆叠调度；端到端延迟

中图分类号：TP393

文献标志码：A **文章编号**：1001-5965(2023)07-1838-09

在综合模块化航空电子系统(integrated modular avionics, IMA)是最常见的航空电子系统体系架构[1]的基础上，分布式综合模块化航空电子系统(distributed IMA, DIMA)采用具有时间触发(time-triggered, TT)能力的交换式网络实现安全关键性消息的严格周期确定性，如SAE AS6802标准定义的时间触发以太网(time-triggered ethernet, TTE)[2]。

随着可综合模块逐渐朝着微小化方向发展，综合化互连将不仅局限于局域网，片上系统(system on chip, SoC)之间也可以进行综合化互连[3]。例如，欧洲的DREAMS(Distributed Real-time Architecture for Mixed-criticality Systems)项目提出了多核芯片构成的混合关键性流量网络模型，旨在实现网络的安全性和实时性，提高应对突发情况的能力[4]。Paukovits[5]和Kopetz[6]等提出了时间触发片上网络(time-triggered system-on-chip, TTSoC)的概念，Urbina通过TTE实现了多个多核芯片互连的跨层次体系结构[7]，文献[8−9]对片内和片间互连通信方式进行了研究，提出了实现片内片间通信的综合互连网络模型。

传统上TT消息在网络中传输依照离线生成的调度表。文献[10−11]提出了基于满足性模理论(satisfiability modulo theories, SMT)的TTE调度方法，对消息的端到端延迟、任务的优先级等进行约束，得到TT消息的调度表。文献[12]提出一种基于贪婪随机自适应搜索算法的TTE调度方法。

随着混合关键性系统的规划与设计方式的进展，需要改进调度方法提高其灵活性，使其可以在一定程度上不需要重新离线配置，就可以承载意料之外的关键流量。文献[13]提出了一种芯片间互连网络TT消息调度方法，文献[14]使用免疫算法对芯片间TT消息调度方法进行了优化，但都未提出提高其灵活性以对新加入TT消息进行调度的方法。文献[15]提出使用拓展的软件定义网络(software-defined networking, SDN)来实现动态的数据包处理和调度；文献[16]通过调整现有流量在各

收稿日期：2021-09-15；**录用日期**：2022-01-16；**网络出版时间**：2022-02-15 11:34
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220214.1419.003.html
基金项目：国家自然科学基金 (62071023)
* **通信作者**. E-mail：avionics@buaa.edu.cn

引用格式：臧光界，李峭，王彤，等. 芯片间时间触发消息堆叠调度方法[J]. 北京航空航天大学学报，2023，49(7)：1838-1846.
ZANG G J，LI Q，WANG T，et al. Stacking scheduling method for time-triggered messages in off-chip network[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1838-1846 (in Chinese).

端口的触发时间来为新加入消息预留时隙, 但是该方法改变了原有消息的实时性需求, 且在网络规模较大、流量较多时难以实现。以上方法都属于超调度方法(super-schedule approach), 即不考虑通信任务所属的模式, 只是将其所有模式中的消息依次调度。网络互连通信中的所谓“模式”, 指实现各种给定功能的不同工作阶段, 在不同模式中, 网络传输的消息不同。文献 [17] 提出了一种堆叠调度方法(stacked scheduling approach, SSA), 将不同模式下的 TT 消息堆叠在同一时隙, 该方法的目的在于为偶发消息预留带宽; 文献 [18] 给出了堆叠调度方法的具体实施方案, 但只适用于交换机控制模式切换的情况, 并只考虑了芯片间互连中交换结构均集中于各个端口的情况。

基于以上研究现状, 本文提出一种芯片间综合化互连网络 TT 消息堆叠调度方法, 使 TT 消息的调度综合考虑了 TT 消息的顺序、模式及芯片间互连网络的拓扑等因素, 并根据给定的芯片数量及消息的属性、所属模式、顺序关系等条件, 给出生成芯片间互连拓扑结构和 TT 消息调度表的方法。该方法可使链路对网络带宽的利用率达到最大, 并尽可能减小流量的端到端延迟, 提高芯片间互连网络的灵活性和适应性。

1　系统模型

1.1　芯片间互连结构

芯片间互连的拓扑结构可以是规则的(如网格形), 也可以根据应用要求设计呈现不规则的拓扑, 而前者可以被看作后者的特例。为了一般性, 本文提出并讨论的调度方法可以适用于不规则的拓扑结构。图 1 为不规则的互连拓扑结构。

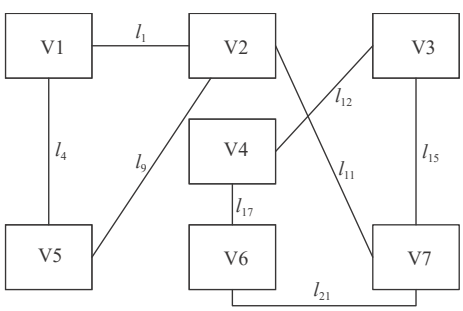


图 1　芯片间互连结构示意

Fig. 1　Sketch map of off-chip interconnection

图 1 中, V1～V7 表示芯片, 在网络拓扑中以节点形式表示, 芯片之间通过开放式的介质无关接口建立物理链路,进行全双工通信, 目前常用的芯片间互连方式有 SPI 总线协议、I^2C 总线协议、Rapid IO 架构等。图中芯片之间的连线表示物理链路, 芯片的物理链路数量和布局根据数据传输的需求确定, 在一般情况下是不对称的。为使链路序号反映链路两端节点, 不失一般性, 对链路序号 i 的标定依据式 (1) 进行:

$$i=\sum_{k=n-s}^{n-1}k-(n-d) \tag{1}$$

式中: n 为网络芯片数; s 和 d 分别为该链路两端芯片编号的较小值和较大值。

可以使用 n 阶对称矩阵 $\boldsymbol{G}$ 在算法中表示芯片间互连网络的拓扑结构, n 为网络中芯片的数量, 其主对角线上各元素为 0, 其他元素中若 $\boldsymbol{G}(i,j)=1$, 则认为芯片 Vi 与 Vj 由物理链路直接相连, 若 $\boldsymbol{G}(i,j)=0$ 则认为 2 个芯片不由物理链路直接互连。

1.2　芯片间互连 TT 消息模型

设在芯片间互连网络中传输的 n 条消息集合为 $M=\{m_1,m_2,\cdots,m_n\}$, 其中, 第 i 条 TT 消息 $m_i(i=1,2,\cdots,n)$ 的属性可表示为

$$m_i=\langle p_i,s_i,\mathrm{path}_i,o_i,\mathrm{mode}_i\rangle \tag{2}$$

式中: p_i 为消息周期; s_i 为消息长度; path_i 以 TT 消息源节点、中转节点、目的节点编号的有序数列表示 TT 消息的传输路径; o_i 为消息在经过各条物理链路上的发送起始时间偏移量; mode_i 为该 TT 消息在该网络中所处的模式, TT 消息不会在自身不处于的模式中产生。在堆叠调度方法中, 默认所有 TT 消息的周期、长度和模式已知, 计算每条 TT 消息的发送起始时间偏移量和路径, 并由此规划芯片间互连网络的拓扑。

1.3　堆叠调度原理

所谓堆叠, 即在同一时隙内, 同时安排 2 条或更多的 TT 消息传输, 被堆叠的 TT 消息不属于同一模式, 因此不会同时产生和传输引发冲突。传统的 TT 网络中, 事先已为每条周期性消息分配一段固定的传输时隙, 因此, 在一条物理链路的某段时隙, 即使该时隙上没有对应 TT 消息的实例传输, 系统也不会将其分配给其他消息。在对多模式网络进行调度时, 若采用超调度方法, 则在多模式情况下会导致很低的带宽利用率和很差的适应性。堆叠调度方法则避免了对带宽和时隙的浪费, 提高了网络的适应性。

例如, 某 TT 网络中的一条物理链路在一段时间之内的 TT 消息传输情况如图 2 和图 3 所示。设该网络具有 3 种操作模式, 分别为 mode_1、mode_2、mode_3, 在该链路上有 6 条 TT 消息传输, 分别为 m_1、m_2、m_3、m_4、m_5、m_6, 周期均为 T, 其中, m_1、m_2 属于 mode_1, m_3、m_4 属于 mode_2, m_5、m_6 属于 mode_3。使用

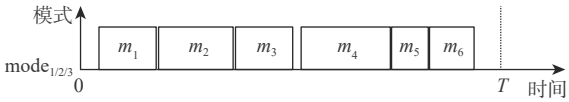


图2 物理链路在多模式情况下使用超调度方法的调度结果

Fig. 2 Scheduling results of physical link using super-schedule approach in multi-mode

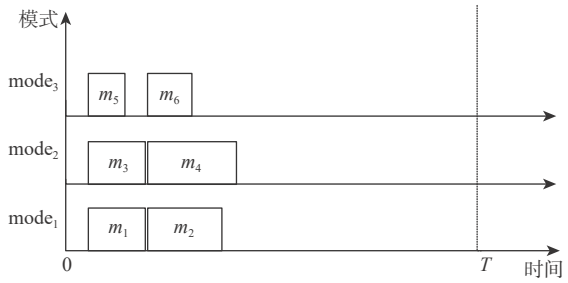


图3 物理链路在多模式情况下使用堆叠调度方法的调度结果

Fig. 3 Scheduling results of physical link using stacking-schedule approach in multi-mode

时间窗来表示TT消息在链路上的调度结果[19]，当使用超调度方法进行调度时，如图2所示，虽然由于不需要在各芯片中加入模式转换功能使得调度更加容易实现，但是在该调度方法下，时间资源要按多种模式下TT消息负载之和预留，因此留给低优先级流量的时隙更少，当该链路所在网络处于运行状态时，如果有新的消息在线加入，网络将很难为其在该链路上分配时隙，导致该网络具有较差的灵活性和适应性。

图3为该链路采用堆叠调度方法的调度结果。转换请求消息包含有要求目的节点转换为指定模式的信息，模式转换请求消息可通过修改的协议控制帧(modified protocol control frame, MPCF)传输，MPCF是具有最小长度的标准以太网帧(46Byte)，与普通协议控制帧(protocol control frame, PCF)不同，其类型字段被设置为0x06[19]，且将PCF的保留和透明时钟字段都改为数据加载字段。MPCF与PCF具有相同的优先级，高于TT和RC流量，为避免干扰时钟同步过程，MPCF在同步过程之外传输。当网络的操作模式变化时，由一个或几个节点发出模式转换请求消息，通过链路传输至网络的其他节点，收到转换请求消息的节点在该网络中所有TT消息的超周期末尾时转换模式，使用新模式下的调度表。与超调度方法相比，堆叠调度方法增加了链路时隙的利用率，能够在运行过程中为更多必要的事件触发消息分配时隙，增加了TT网络的灵活性和适应性，同时减小了被堆叠消息的端到端延迟。

1.4 芯片间网络模式切换方法

在由交换机和端系统组成的交换式网络中，由网络中的主节点发送模式更改消息，从节点接收模式更改消息后在该超周期的末尾改变操作模式[20]。但是在芯片间互连时间触发网络中，网络的拓扑及每个节点需要处理的消息数在调度完成之前都不确定，因此并不能设定一个或多个可以产生和传输模式更改消息的主节点。

在模式更改消息设计方面，若模式更改消息的传输路径与该模式下TT消息的传输路径不同，则很可能出现模式更改消息早于某条TT消息帧到达其目的或中转节点的情况，在此情况下，由于该节点的模式已被改变，将不会为该TT消息帧提供接收或转发服务，造成TT消息的丢失；另外，若某条TT消息在网络中的跳数较多，当全局时钟到达超周期的末尾时，该TT消息可能还未完成传输过程，此时更改模式同样会造成消息的丢失。

在交换式TT网络的模式切换方法基础上，提出一种以单条TT消息为单位，依节点顺序依次切换模式方法，即当前模式需要切换时，该模式下每条TT消息的起始节点在发送当前TT消息的最后一帧后紧接着发送一条与此TT消息传输路径相同的模式转换请求消息，路径上的每一个转发节点芯片在接收到此模式转换请求后即停止对应TT消息的转发和接收服务，当一个节点接收到所有该模式下此节点接收和转发的TT消息对应的模式转换请求时，改变该节点的工作模式。图4为该模式转换方法的一个实例。

图4中，操作模式1包含2条TT消息，m_1的传输路径$path_1$ = {V1,V3,V4}，TT消息m_2的传输路径$path_2$ = {V2,V3,V4}，当网络从模式1切换至模式2时，m_1的起始节点V1和m_2的起始节点V2分别沿传输路径发送模式转换请求消息cm_1和cm_2。如图4(a)所示，假设cm_1先于cm_2发送，V1发送cm_1后进入模式2，V2此时还处于模式1，在图4(b)中，V2发送cm_2后进入模式2，此时cm_1已先到达V3，由V3接收后向V4转发，此时V3停止对m_1的接收和转发服务，但仍属于模式1，仍为模式1的其他消息提供服务，在图4(c)中，当cm_2到达V3后，V3已接收到模式1的所有模式转换请求消息，V3切换为模式2，同理，V4在接收到cm_1和cm_2后也切换为模式2，整个网络模式切换完成，如图4(d)所示。

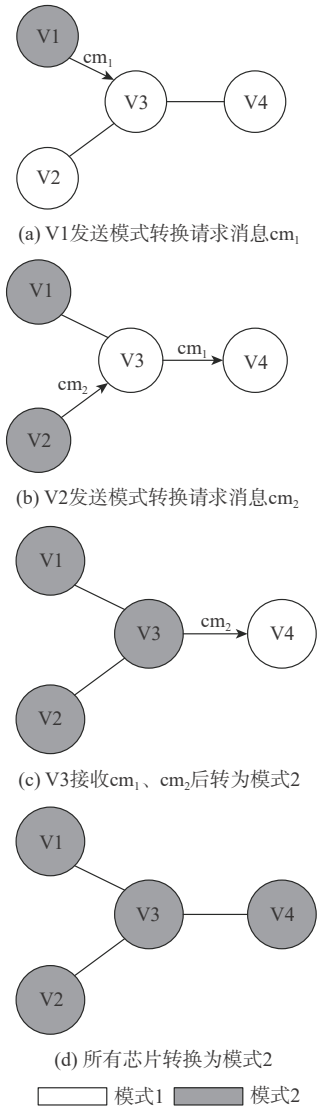


图 4　芯片间互连网络模式切换过程实例

Fig. 4　An example of mode-change in off-chip interconnection network

2　本文方法

对于一条 TT 消息 m_i, 若其发送起始时间偏移量$o_i = t$, 则其第 j 帧的发送时间o_i^j 满足:

$$o_i^j = t + jp_i \tag{3}$$

因此对于一组 TT 消息, 只要确定了其第一个超周期内的偏移量, 即可确定每条 TT 消息在所有时段上的调度结果。设 c_i 为 m_i 的执行时间, 满足:

$$c_i = \frac{s_i}{r} \tag{4}$$

式中: r 为网络数据传输速率。

在一条物理链路上, 可供调度的时间资源被记为 S, A 为一定数目的整数集合, 其中会出现几组连续的整数, 将非空集合 A 中出现的连续整数组中的数字最多的那一组的数字个数记为 LLC(A), 集合 H 表示一个超周期内所有时间, 集合 U 表示超周期内已被调度的所有时刻, 集合 R 表示一个超周期内可调度的空余时间, 则有

$$U \cup R = H \tag{5}$$

$$U \cap R = \varnothing \tag{6}$$

m_i 在该链路上可调度的条件为[20]

$$c_i \leqslant \mathrm{LLC}(R) \tag{7}$$

若 m_i 需经多个中转节点转发, 其在前一链路上传输的结束时间为 t_e, 后一链路上传输的开始时间为 t_s, 这 2 个参数应满足:

$$t_e \geqslant t_s \tag{8}$$

图 1 所示拓扑用矩阵表示为

$$G = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{9}$$

由于拓扑是不确定的, 可以先假设所有节点之间都是互相连接的, 再分别在每个模式下对消息依序在各链路上寻找时隙, 在此期间对已使用端口数等于其最大端口数的芯片周围链路进行“剪枝”, 即切断与其相连但未被使用的链路。根据以上规定和原则提出芯片间网络 TT 消息堆叠调度方法的具体算法, 在该算法中定义了几种操作函数, 包括模式提取函数 checkmode、链路检查函数 linkavailable、时隙匹配函数 LLCavailable 及剪枝函数 cutedges, 具体步骤如下。

输入: TT 消息集合 M, 芯片数量 n, 消息数量 msg, 模式数 mo, 各芯片端口数 a。

输出: 拓扑矩阵 $\boldsymbol{G}$, 调度表 $\boldsymbol{S}$。

```
function OFFCHIPSTK
S ← ∅, S_t ← ∅, G ← fillbyone(n) − E
for i ← 1 to mo
    M_i ← checkmode(M, i), msg_i ← size(M_i)
    for j ← 1 to msg_i
        route_j = kshortestpath(source_j, dest_j, G)
        rtn ← size(route_j)
        for k ← 1 to rtn
            if linkavailable(G, route_j(k)) == true
            if LLCavailable(S_t, route_j(k)) == true
                path_j ← route_j(k)
                S_t ← S_t ∪ singlesche(route_j(k), c_j, S_t)
                break
            end if
        end if
```

16: $\boldsymbol{G} \leftarrow \text{cutedges}(\text{path}_j, \boldsymbol{G}, \boldsymbol{S}_t)$
17: end for
18: end for
19: $\boldsymbol{S} \leftarrow \boldsymbol{S} \cup \boldsymbol{S}_t$, $\boldsymbol{S}_t \leftarrow \varnothing$
20: end for
21: end function

首先,对TT消息按照重要性排序,若无已给出的重要性规定,则优先调度周期小的消息。为了给通过MPCF传输的模式转换消息预留时隙,在计算每条TT消息的帧长时,都在其原长度的基础上加上MPCF的长度,这样在没有模式转换请求的情况下,该预留时隙也可供其他事件触发消息传输使用,在一定程度上减少了事件触发消息的拥塞。

在本文方法中,首先假设所有芯片之间均互相连通,函数fillbyone(n)表示创建一个所有元素均为1的n阶方阵。然后按照模式—消息—链路的顺序依次调度,对于原始消息列表M,使用函数checkmode将其中属于mode$_i$的所有消息提取出来组成新的消息列表M_i。路径确定方面,可以使用k最短路径方法(k_shortest paths)[20],依据拓扑和消息的源节点和目的节点确定k条实现传输的最短路径,依次判定该路径是否可行。函数linkavailable和LLCavailable分别用来判断该路径在拓扑中是否存在及该路径上所有链路是否满足式(7)中的TT消息可调度条件,若满足,则通过函数singlesche,依据延迟最短原则为其安排传输时隙并将其加入当前模式调度表$\boldsymbol{S}_t$,若不满足则判断下一条传输路径是否满足。当一条TT消息在其所经过的各条链路上的发送时间偏移量都确定之后,即认为其调度完成,此时使用函数cutedges完成对拓扑的“剪枝”,其具体功能为:判定网络中的每个节点已经使用的链路数是否已经达到其端口数a,若已达到则删除拓扑中所有与该节点相连的未使用的链路。当一个模式下所有的消息都调度完成后,将当前模式下调度表与其他模式调度表合并,并新建一张空的调度表进行下个模式的调度。

本文方法中,不同模式下的各组TT消息调度过程是互相独立的,即调度过程中每一模式都继承之前模式的拓扑但不继承其调度表,相当于在每个模式调度结束后都将当前调度表重置一次,这样就可以使一段时隙在多个模式中都可使用,采用延迟最短原则要求在每条链路上满足式(7)的条件下使t_s尽可能小,这2项原则共同使得消息的时间窗在链路的调度表上实现堆叠。

3 仿真实验与分析

3.1 实验参数设置

设网络中TT消息的传输速率为100 Mbit/s,已知以太网帧的帧长范围为64～1 514 Byte,时间触发流量的周期取为1～128 ms之间的幂次值,以此来随机生成TT消息,案例中将TT消息的周期设为$p = 2^m \times 3^n$ ms,其中,$0 \leqslant m \leqslant 7$、$0 \leqslant n \leqslant 4$且$m$、$n$均为整数,最终需满足$1 \leqslant p \leqslant 128$。消息的源节点、目的节点和所属模式都根据事先设定好的限制随机生成,每组实验的限制都各不相同。为进行对比,以现有的一种超调度方法——芯片间TT消息拓扑路径时刻(topology, route, schedule, TRS)调度方法的调度结果作为对照[15],以消息列表中所有TT消息的端到端延迟之和及链路的平均时隙占用率分别表示调度后网络的实时性和灵活性。其中,链路平均时隙占用率为链路中已被占用的时隙占总时间的比例,表示为

$$U = \frac{1}{f}\sum_{i=1}^{f}\frac{u_i}{h} \tag{10}$$

式中:f为网络中最终被激活的链路数;u_i为链路l_i上已被占用的时隙。U越小,表明链路的平均占用率越低,可以用来调度零星事件触发消息的时隙也就越多,网络的灵活性也就越高。

设置网络的芯片数、TT消息数和操作模式数来对比堆叠调度和TRS调度方法,通过比较评价堆叠调度方法的效果。具体参数设置如表1所示,其中对实验组内具体每条消息的属性是否相同不做限制。

表1 实验参数设置

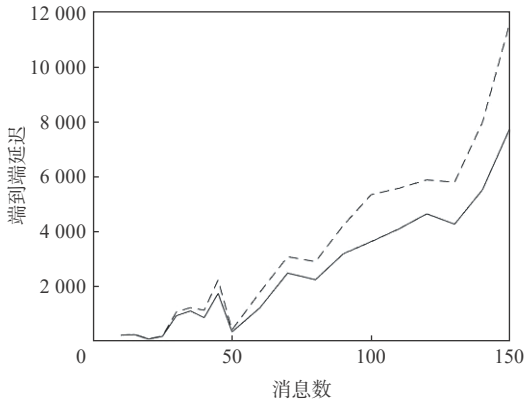
Table 1 Experimental parameter settings

实验组号	消息组数	每组消息数	网络芯片数	操作模式数
1	19	10～150	10	3
2	8	50	5～15	5
3	8	50	10	3～10

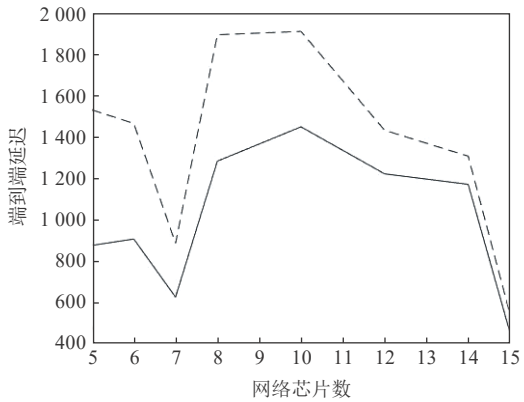
3.2 仿真结果分析

实验中,消息的端到端延迟为消息的发送起始时间偏移量与消息在其路径上各个芯片之间传输时的执行时间之和。其中,发送起始时间偏移量即为消息产生到发送所等待的时间,同时当网络的传输速度不变时,数据帧在任意2个芯片之间传输的执行时间相同,因此传输的执行时间之和可以用消息在网络中传输的跳数与在任意链路上传输时间之积表示。分别计算每组消息使用TRS调度方法和堆叠调度方法后,该组TT消息总的端到端延迟和链路平均时隙占用率,结果如图5和图6所示。

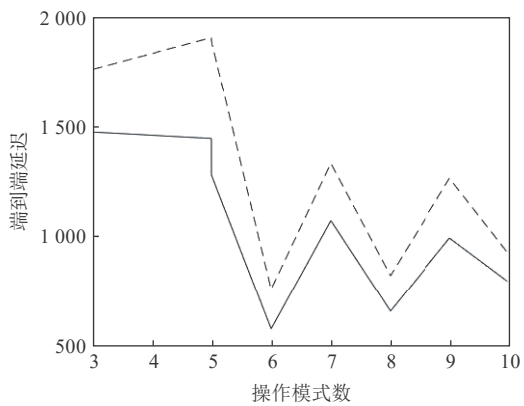
从图5可以看出,在3组实验中,由于每组TT消息生成的随机性,导致不同组消息的平均长度具有较大差异,从而使得各组TT消息的总端到端延



(a) 端到端延迟随消息数变化



(b) 端到端延迟随网络芯片数变化

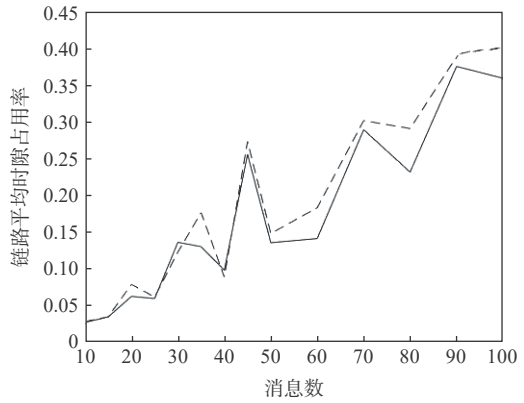


(c) 端到端延迟随操作模式数变化

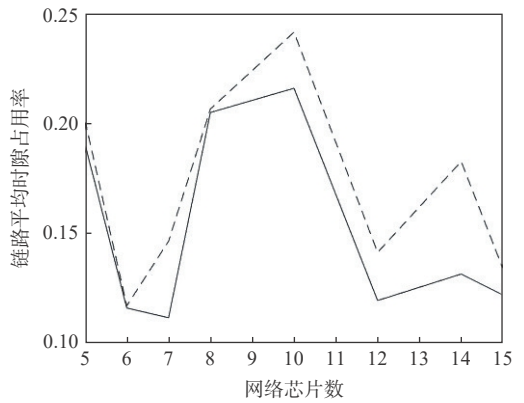
—— 堆叠调度方法　– – – TRS调度方法

图 5　不同条件下端到端延迟的变化

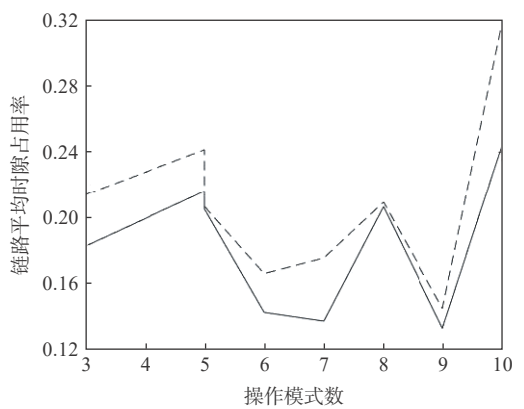
Fig. 5　Change of end-to-end delay under different conditions



(a) 链路平均时隙占用率随消息数变化



(b) 链路平均时隙占用率随网络芯片数变化



(c) 链路平均时隙占用率随操作模式数变化

—— 堆叠调度方法　– – – TRS调度方法

图 6　不同条件下链路平均时隙占用率的变化

Fig. 6　Changes of average slot utilization rate of link under different conditions

迟存在较大差异, 但是在所有组别的仿真实验中, 与 TRS 调度方法相比, 堆叠调度方法都减少了 TT 消息的端到端延迟, 在该实验所设置的仿真参数下, 减少的幅度在 0%～45% 之间。从图 6 可以看出, 与 TRS 调度方法相比, 堆叠调度方法可以减少链路平均时隙占用率, 从而增加网络运行时链路上的空余时隙, 当对事件触发消息进行在线调度时, 可以增加可供调度的事件触发消息的长度或数量, 由此可看出堆叠调度方法在提高网络的实时性和灵活性方面有着显著的效果。

实际上, 因为该实验只是从宏观上研究了消息数、网络芯片数等的影响, 实际上端到端延迟还会受到每一条消息属性(长度、源节点、目的节点等)的影响。在堆叠调度方法中, 如果能够堆叠较多的消息, 效果会更明显, 如果因为消息属性问题导致能够堆叠的消息较少(如同一链路上的 TT 消息刚好都是同一模式, 或处于不同模式下的 TT 消息不在同一链路上传输), 效果变差, 因此图 5 和图 6 中

会出现明显的转折。为了更准确地看出堆叠调度方法对端到端延迟的影响，将所有实验结果按照使用 TRS 调度方法的总端到端延迟进行排序。图 7 中给出了随 TRS 调度方法的端到端延迟增加，堆叠调度方法的端到端延迟对应变化情况的拟合曲线。可以看出，当使用 TRS 方法调度的 TT 消息组端到端延迟增加时，使用堆叠调度方法的 TT 消息组端到端延迟也呈上升趋势，但堆叠调度端到端延迟增长速率要缓于 TRS 方法，因此可推断出总端到端延迟时间越长，堆叠调度减少延迟的效果越明显。

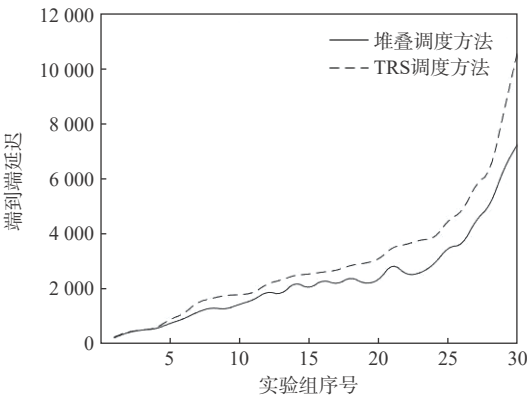


图 7 TRS 调度方法端到端延迟增加时堆叠调度方法端到端延迟增加趋势

Fig. 7 Increase trend of end-to-end delay of stack scheduling method, as end-to-end delay in TRS increases

更进一步，通过不同参数下堆叠调度方法减少端到端延迟的效果，定性研究影响使用堆叠调度方法调度 TT 消息必要性及影响堆叠调度方法效果的主要因素。图 5(a)中，在网络芯片数、操作模式数相同的条件下，随着 TT 消息数的增加，堆叠调度方法相较于 TRS 调度方法减少的端到端延迟呈现上升趋势；图 5(b) 中，在 TT 消息数、操作模式数相同的条件下，随网络中芯片数的增加，堆叠调度方法相较于 TRS 调度方法减少的端到端延迟呈现下降趋势。由于增加 TT 消息数量和减少网络中芯片数量本质上都是增加了平均每条链路需要传输的 TT 消息数，据此对使用堆叠调度方法后端到端延迟减少量随不使用堆叠调度方法时平均每条链路传输 TT 消息数变化情况进行统计并拟合曲线。如图 8 所示，随平均每条链路所传递的消息数增加，堆叠调度方法减少的端到端延迟呈明显上升趋势。因此，可得出如下结论：芯片间互连网络中，当使用常规调度方法时，平均每条链路负责传输的 TT 消息越多，转为使用堆叠调度方法时相较于平均链路传输消息少的网络减少总端到端延迟的效果越明显，即使用堆叠调度方法调度消息的必要性就越高。

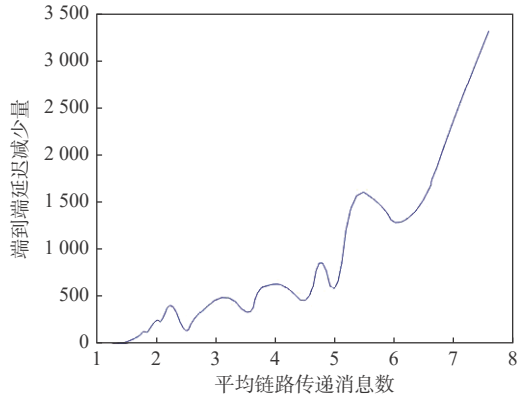


图 8 堆叠调度方法端到端延迟减小量随平均链路传递消息数变化

Fig. 8 Change of amount of end-to-end delay reduction in stack scheduling method varies with average number of messages delivered on link

此外，考虑到在使用堆叠调度方法的前提下，操作模式也会对调度效果产生影响。通过改变 3.1 节中一组 TT 消息不同的模式分布，使用堆叠调度方法对其进行调度，比较不同模式分布下总端到端延迟的变化情况。该组消息共有 150 条，当使用不考虑操作模式的 TRS 调度方法时，其总端到端延迟约为4.7×10^4 ms，在改变模式分布的过程中，所有消息的周期、数据帧长、源节点、目的节点都不改变，也不会出现不同模式下消息重组为同一模式，以此来减少其他因素干扰。实验研究了 5～20 个操作模式下的总端到端延迟的变化情况，结果如图 9 所示。从结果可看出，其他条件不变时，随着操作模式数的增加，堆叠调度方法下的端到端延迟呈现先陡峭后平缓的减少趋势，因此可得出如下结论：使用堆叠调度方法调度 TT 消息时，对于一组 TT 消息，其操作模式数越多，使用堆叠调度方法后的总端到端延迟越短。此外，通过与 TRS 调度方法端到端延迟比较，可以看出若设定除操作模式外所有属

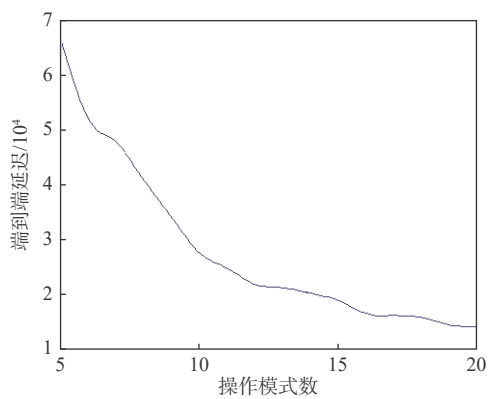


图 9 堆叠调度方法端到端延迟随操作模式数变化

Fig. 9 Change of end-to-end delay in stack scheduling method varies with number of messages' modes

性都相同, 在操作模式数较少时, 堆叠调度方法延迟减少效果不一定优于 TRS 调度方法。这是因为调度顺序改变而导致了一些“剪枝”的提前或延后进行, 从而可能使得部分消息传输的跳数与对应延迟增加, 其增加量在操作模式数较少时, 在某些情况下会大于堆叠调度方法减少的排队延迟; 同时在硬件方面, 使用堆叠调度方法需要在芯片内增加对应的模式切换消息处理模块, 这对芯片的设计提出了更高的要求, 这些都是使用堆叠调度方法应考虑到的因素和成本, 在实际工程中应加以考虑。

4　结　论

1) 建立了多操作模式情况下的芯片间互连网络与时间触发消息模型, 提出了适用于芯片间互连网络的模式转换方法和步骤。

2) 提出了芯片间时间触发消息堆叠调度方法。在不同操作模式下分别根据路径选择策略和 TT 消息严格可调度原则为 TT 消息确定传输路径和发送时间偏移量, 并通过重置不同模式下 TT 消息调度表和继承上一模式拓扑的方式保证 TT 消息的堆叠和每个芯片使用的端口数符合限制。

3) 仿真实验结果表明, 芯片间时间触发消息堆叠调度方法能够随着网络规模的变化, 不同程度减小网络中 TT 消息的端到端延迟和链路平均时隙占用率, 从而提高网络的实时性和灵活性。其中, 与 TRS 调度方法相比, 在模式较多及硬件支持的情况下, 堆叠调度方法可以减少 TT 消息的端到端延迟, 其减少效果与 TRS 调度方法调度结果的端到端延迟和平均每条链路传递消息数呈正相关。此外, 堆叠调度方法的端到端延迟随着网络中消息的操作模式数增加而减少。当链路上可堆叠的消息较多时, 可以显著减少调度结果的端到端延迟。

参考文献 (References)

[1] WOLFIG R, JAKOVLJEVIC M. Distributed IMA and DO-297: Architectural, communication and certification attributes[C]// IEEE/AIAA 27th Digital Avionics Systems Conference. Piscataway: IEEE Press, 2008: 1. E. 4-1.

[2] 何锋. 机载网络技术基础[M]. 北京: 国防工业出版社, 2018: 9-15.
HE F. Fundamentals of airborne network[M]. Beijing: National Defense Industry Press, 2018: 9-15 (in Chinese).

[3] 蒲小勃. 现代航空电子系统与综合[M]. 北京: 航空工业出版社, 2013: 70-86.
PU X B. Modern avionics system and integration[M]. Beijing: Aviation Industry Press, 2013: 70-86(in Chinese).

[4] AHMADIAN H, OBERMAISSER R, PEREZ J. Distributed real-time architecture for mixed criticality systems[M]. Boca Raton: Taylor & Francis Group, 2018: 51-57.

[5] PAUKOVITS C, KOPETZ H. Concepts of switching in the time-triggered network-on-chip[C]//2008 14th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications. Piscataway: IEEE Press, 2008: 120-129.

[6] KOPETZ H, STEINER W. Real-time systems: Design principles for distributed embedded applications[M]. 2nd ed. Berlin: Springer International Publishing, 2011: 326-337.

[7] DURRIEU G, FOHLER G, GALA G, et al. Dreams about reconfiguration and adaptation in avionics[C]//Proceedings Embedded Real Time Software and Systems. Toulouse: HAL Science Ouverte, 2016: 48-57.

[8] URBINA M, OBERMAISSER R. A gateway core between on-chip and off-chip networks for an AUTOSAR message-based multi-core platform[C]//Automotive Meets Electronics. Dortmund: VDE Verlag GmbH, 2016: 1-6.

[9] OLIVER R S, CRACIUNAS S S. Hierarchical scheduling over off- and on-chip deterministic networks[J]. ACM SIGBED Review, 2016, 13(4): 14-19.

[10] STEINER W. An evaluation of SMT-based schedule synthesis for time-triggered multi-hop networks[C]//Real-Time Systems Symposium. Piscataway: IEEE Press, 2011: 375-384.

[11] CRACIUNAS S S, OLIVIER R S. SMT-based task- and network-level static schedule generation for time-triggered networked systems[C]//22nd International Conference on Real-Time Networks and Systems (RTNS 2014). New York: ACM, 2014: 45-54.

[12] 郑重, 何锋, 李浩若, 等. 基于贪婪随机自适应搜索法的TTE通信调度算法[J]. 北京航空航天大学学报, 2021, 47(11): 2268-2276.
ZHENG Z, HE F, LI H R, et al. Scheduling algorithm of TTE network based on greedy randomized adaptive search procedure[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(11): 2268-2276(in Chinese).

[13] 孔韵雯, 李峭, 熊华钢, 等. 片间综合化互连时间触发通信调度方法[J]. 航空学报, 2018, 39(2): 258-267.
KONG Y W, LI Q, XIONG H G, et al. Time-triggered communication scheduling method for off-chip integrated interconnection[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(2): 258-267(in Chinese).

[14] 汪晶晶, 李峭, 熊华钢, 等. 芯片间时间触发通信综合规划方法及其优化[J]. 北京航空航天大学学报, 2020, 46(1): 170-180.
WANG J J, LI Q, XIONG H G, et al. Integrated planning method and optimization for off-chip time-triggered communication[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(1): 170-180(in Chinese).

[15] MOUTINHO L, PEDREIRAS P, ALMEIDA L. A real-time software defined networking framework for next-generation industrial networks[J]. IEEE Access, 2019, 7: 164468-164479.

[16] CRACIUNAS S, SERNA OLIVER R S. Online incremental scheduling method for deterministic networks: US20170070439[P]. 2017-03-09.

[17] HEILMANN F, SYED A, FOHLER G. Mode-changes in COTS time-triggered network hardware without online reconfiguration[J]. ACM SIGBED Review, 2016, 13(4): 55-60.

[18] SYED A. Model-based design and adaptive scheduling of distributed real-time systems[D]. Kaiserslautern: Technical University of Kaiserslautern, 2018: 60-63.

[19] CHEN J C, DU C L, XIE F, et al. Schedulability analysis of non-

preemptive strictly periodic tasks in multi-core real-time systems[J]. Real-Time Systems, 2016, 52(3): 239-271.

[20] TANG X Q, LI Q, ZUO Y J, et al. Online schedule of sporadic life-critical traffic in TTEthernet[C]//2019 IEEE 5th International Conference on Computer and Communications (ICCC). Piscataway: IEEE Press, 2020: 1312-1316.

Stacking scheduling method for time-triggered messages in off-chip network

ZANG Guangjie[1, 2], LI Qiao[2, *], WANG Tong[2], XIONG Huagang[2]

(1. Chinese Aeronautical Establishment, Beijing 100029, China;
2. School of Electronic Information Engineering, Beihang University, Beijing 100191, China)

Abstract: To ensure the strict time determinism of message transmission between avionics communication tasks, time-triggered (TT) communication methods are also applied to the off-chip interconnection network. When avionics tasks have multiple operation modes, the time-triggered schedules between chips belonging to different modes will overlap and occupy time slots. The stacking scheduling method for time-triggered messages in off-chip network, which can reduce the queuing delay of application layer messages due to the waiting TT time window, is proposed to improve the flexibility and efficiency of using network resources,. Simulation experiments show that compared with the super scheduling method, the stack scheduling method can reduce the total end-to-end delay of TT messages in the off-chip interconnection network and the average time slot occupancy rate of the link. For scenarios where the end-to-end delay time is long and the average link bearer message transmission is large, the effect of using the stack scheduling method to reduce the end-to-end delay is more significant.

Keywords: off-chip interconnection; time-triggered message; operation mode; stacking scheduling ; end-to-end delay

Received: 2021-09-15; **Accepted**: 2022-01-16; **Published Online**: 2022-02-15 11:34

URL: kns.cnki.net/kcms/detail/11.2625.V.20220214.1419.003.html

Foundation item: National Natural Science Foundation of China (62071023)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0555

CFRP层压板脱黏缺陷红外脉冲热波层析成像检测

钟佳岑，徐浩军，魏小龙*，韩欣珉，常怡鹏
(空军工程大学 航空工程学院，西安 710038)

摘 要：为实现基于红外脉冲热波成像检测碳纤维增强塑料（CFRP）层压板脱黏缺陷的三维层析成像，完善外场定量检测保障体系，提高层压板在服役过程中的安全性和可靠性，开展了红外脉冲热波层析成像方法与检测技术研究。制备了一种人工脱粘缺陷试样，采用红外脉冲热波层析成像检测技术对脱粘缺陷进行检测，分析了脱粘区和非脱粘区的表面热信号瞬态响应过程及红外脉冲热波层析成像对脱粘缺陷的检测能力。通过脱粘区与非脱粘区基于对数多项式拟合的重构热信号差，计算得到热信号极值时间图像，分析了脱粘区极值时间变化规律及缺陷状态；采用核函数模糊C均值聚类对相同缺陷深度对应的极值时间数组进行二分类，由此计算数组平均值作为缺陷对应极值时间；将该时间与脱粘区极值时间数组建立统计关系来构造断层图像序列，并计算其对应的缺陷深度；在此基础上，利用等值面绘制方法实现层压板脱粘缺陷三维可视化。研究表明，红外脉冲热波层析成像能够定量检测CFRP层压板脱粘缺陷，准确可靠显示层压板内部缺陷的分布和形貌，检测缺陷深度与实际缺陷深度的最大相对偏差低于15%，对工程应用具有一定指导意义。

关 键 词：碳纤维增强塑料；层压板；脱黏缺陷；红外脉冲热波成像；层析检测
中图分类号：V250.2
文献标志码：A **文章编号**：1001-5965(2023)07-1847-10

碳纤维增强塑料(carbon fiber reinforced plastic, CFRP)在航空航天飞行器制造领域的应用日益广泛[1]。采用CFRP制备的层压板由于质量轻、强度高、耐高温和耐腐蚀等特点，广泛应用于飞机主承力结构件(如机身、机翼和尾翼等)[2]。受制造工艺、使用条件等影响，CFRP层压板在制造和服役过程中极易产生脱黏、分层、裂纹等内部缺陷，严重影响飞机的安全性与可靠性[3]。在CFRP层压板发生明显内部缺陷时，无法从层压板表面通过肉眼判断其缺陷状态[4]，因此需采用合适的检测方法对缺陷进行判定和识别，以保证层压板安全服役。由于传统无损检测技术(如超声、射线和激光全息检测等)存在主观性强、检测效率低和使用条件限制等问题[5]，难以满足外场保障要求。针对CFRP层压板内部缺陷定性、定量检测问题，寻找一种高效、易实现、高灵敏度的检测方法具有重要意义[6]。由于红外脉冲热波成像检测技术具有非接触、检测速度快、检测面积大、效率高、适合在线检测等突出优势，近年来，该技术引起了广泛关注[7]。红外脉冲热波成像检测技术对CFRP层压板内部缺陷具有良好的检测能力[8]，该技术通过闪光灯产生的瞬时高能热流加热试样表面，热量传递至缺陷区时，由于缺陷区的热传导系数小于材料本身，则缺陷区对应表面温度高于非缺陷区，采用红外热像仪实时记录试样表面热异常区变化过程，从而实现对试样缺陷的判定和识别[9]。

红外脉冲热波成像检测获得的热图序列只能反映不同时刻断面上的二维信息，无法得到被检测对象的整体描述，因此需要三维可视化来帮助工程技术人员进行定量检测，即确定缺陷位置、形状、

收稿日期：2021-09-16；**录用日期**：2022-01-05；**网络出版时间**：2022-01-21 09:23
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220120.1616.003.html
*通信作者. E-mail：wei18892022001@163.com

引用格式：钟佳岑，徐浩军，魏小龙，等. CFRP层压板脱黏缺陷红外脉冲热波层析成像检测[J]. 北京航空航天大学学报，2023，49(7)：1847-1856. ZHONG J C，XU H J，WEI X L，et al. Detection of debonding defect in CFRP laminates using infrared pulse thermal wave tomography[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1847-1856 (in Chinese).

尺寸及性质等信息[10]。近年来，基于红外脉冲热波的层析成像检测技术成为了该领域的研究热点[11-12]。例如，Melnyk 等[13]提出了投影动态热层析检测方法，该方法采用多方向热加载的图像融合算法实现内部结构三维可视化。Elhassnaoui 和 Sahnoun[14]提出了基于红外脉冲热波成像的亚表层缺陷三维可视化方法，该方法通过热信号重构实现对试样厚度检测，进而得到试样三维结构。因此，基于红外脉冲热波的层析成像可以作为一种定量检测方法。

针对 CFRP 试样内部缺陷的三维可视化问题，基于红外脉冲热波的动态热层析（dynamic thermal tomography, DTT）成像是应用最广泛的方法[15]。该方法采用红外脉冲热波成像检测获得的热图序列，计算得到试样表面缺陷区与非缺陷区热信号差对应的极值时间，通过确定极值时间-深度校准曲线，将该时间转换到缺陷深度表征，进而得到深度方向的断层图像序列以实现三维可视化[16]。Vavilov 等[17-18]对 DTT 方法开展了深入的理论与试验验证工作，并将其用于 CFRP 层板和钢板缺陷检测，研究表明，该方法可以实现缺陷三维可视化，但其信噪比较差。在此基础上，Pawar 和 Vavilov[19]采用三维归一化方法抑制加热和吸收不均，试验验证了该方法在提高三维可视化精度中的有效性。Peng 等[20]提出了基于脉冲相位的 DDT 方法，通过脉冲相位法来提取特征图像以增强信噪比，从而提高三维可视化精度。然而，针对外场检测难以确定极值时间-深度校准曲线的情况，实现特定材料中某种缺陷三维可视化的研究尚十分匮乏，而三维可视化已经成为定量检测的关键技术。

基于此，本文研究了闪光灯激励下 CFRP 层压板脱粘区的表面热信号瞬态响应过程及红外脉冲热波成像对脱粘缺陷的检测能力，通过脱黏区与非脱黏区的基于对数多项式拟合重构热信号差，计算得到热信号极值时间图像，对相同缺陷深度对应的极值时间数组进行核函数模糊 C 均值聚类（kernel fuzzy C-means, KFCM）运算，由此计算缺陷对应数组平均值并与脱黏区极值时间数组建立统计关系，以确定不同深度的断层图像，在此基础上，通过等值面绘制方法实现了 CFRP 层压板内部缺陷三维可视化。

1 试验方法与过程

由于制备 CFRP 层压板自然脱黏的试样十分困难，在试验研究中采用平底孔作为人工脱黏缺陷以代替自然脱黏缺陷，在尺寸为 190 mm×100 mm×6.5 mm 的 CFRP 层压板中制备了 39 个不同尺寸/深度的平底孔。如图 1 所示，为了方便表述，将 39 个平底孔命名为 A1～A6、B1～B6、C1～C6、D1～D6、E1～E10、F1～F5。如表 1 所示，A、B、C、D、E、F 分别表示平底孔直径 d 为 20, 15, 10, 5, 3, 2 mm，其中，A1～A6、B1～B6、C1～C6 中的数字 1～6 分别表示平底孔距试样表面深度 h 为 1.5, 2, 2.5, 3, 3.5, 4 mm，D1～D6 中的数字 1～6 分别表示平底孔距试样表面深度 h 为 1, 1.5, 2, 2.5, 3, 3.5 mm，E1～E10 中的数字 1～10 分别表示平底孔距试样表面深度 h 为 1, 1.25, 1.5, 1.75, 2, 2.25, 2.5, 2.75, 3, 3.25 mm，F1～F5 中的数字 1～5 分别表示平底孔距试样表面深度 h 为 1, 1.25, 1.5, 1.75, 2 mm。

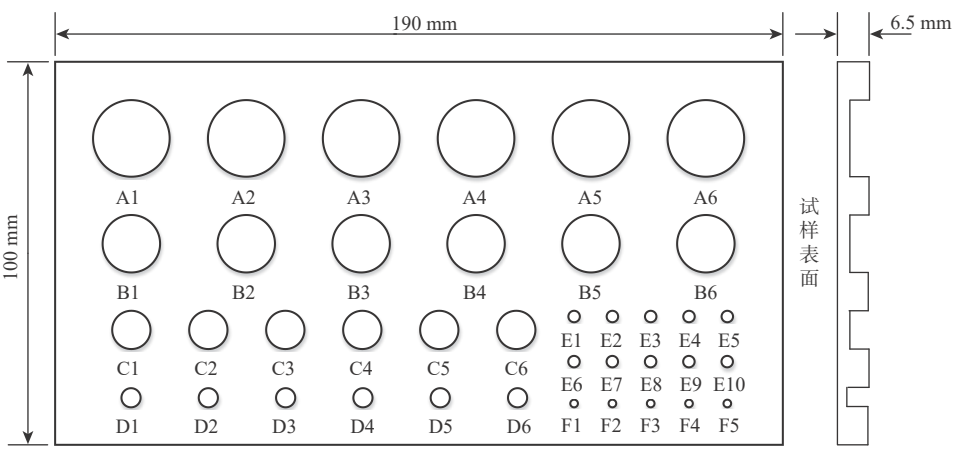


图 1 含脱黏缺陷的 CFRP 层压板示意图

Fig. 1 Sketch map of CFRP laminate with debonding defects

所采用的红外脉冲热波成像检测工作站如图 2 所示，主要包括 2 个 ThermPulse C6100 型高能闪光灯、1 个非制冷焦平面阵列红外热像仪、反光罩、计算机控制及图像采集与处理系统。采用闪光灯对试样表面进行脉冲激励，2 个闪光灯水平布置在试样表面正上方，单个闪光灯单次最大脉冲能量

表 1　试样孔深度 h

Table 1　Sample hole depth h　　mm

编号	A(d = 20 mm)	B(d = 15 mm)	C(d = 10 mm)	D(d = 5 mm)	E(d = 3 mm)	F(d = 2 mm)
1	1.5	1.5	1.5	1	1	1
2	2	2	2	1.5	1.25	1.25
3	2.5	2.5	2.5	2	1.5	1.5
4	3	3	3	2.5	1.75	1.75
5	3.5	3.5	3.5	3	2	2
6	4	4	4	3.5	2.25	
7					2.5	
8					2.75	
9					3	
10					3.25	

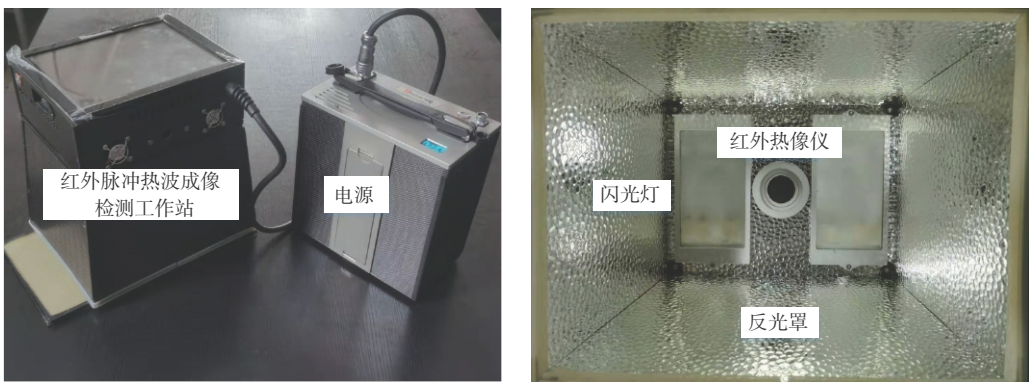


图 2　红外脉冲热波成像检测工作站

Fig. 2　Detection workstation of infrared pulse thermal wave imaging

为 3 kJ, 因此, 2 个闪光灯单次脉冲输出最大能量为 6 kJ, 闪光脉冲宽度为 1～50 ms。采用红外热像仪采集试样表面热图, 热像仪工作温度范围为−20～80 ℃, 热灵敏度小于或等于 0.045 ℃, 光谱响应范围为 8～14 μm。热像仪采集的热图为 16 位数字图像, 单帧热图大小为 300 像素×400 像素, 全画幅最高采集频率为 100 Hz。试验时, 热像仪从脉冲激励结束后采集试样表面热图, 持续采集 160 s 直至热图中缺陷信号消失。试验过程中, 通过工作站中的 FPGA 控制模块对热像仪的采集过程进行控制, 采集的试样表面热图序列经图像压缩后实时存储在工作站中。

2　层压板表面热信号瞬态响应过程及表面热图预处理

2.1　层压板表面热信号瞬态响应

试验过程中, 热像仪采集的热图序列被压缩至 256 帧并存储在工作站中, 试验结束后, 提取热图序列并对其进行处理。图 3 展示了 256 帧热图序列中的部分原始热图。在脉冲激励结束后, 缺陷并非立即显示在原始热图序列中, 而是经历了短暂过程后由弱到强逐渐显示在表面热图中, 在第 96 帧缺陷 D1 最早显示在表面热图中, 在第 192 帧可以清楚识别缺陷 A1～A4、B1～B4、C1～C3、D1～D3, 在第 224 帧能够大致识别缺陷 A1～A6、B1～B6、C1～C6, 但缺陷细节特征不明显, 随着降温过程进行, A1～A6、B1～B6、C1～C6 位置处的缺陷特征逐渐变弱直至消失。在降温过程中, 尺寸较小、深度较深的缺陷 D4～D6、E1～E10、F1～F5 均无法识别。

图 4 展示了在降温过程中, 第 32 帧之后原始热图序列中不同尺寸/深度缺陷的几何中心点与非脱黏区任意一点对应的试样表面热信号随时间变化。脉冲激励后, 脱黏区与非脱黏区热信号衰减速率不同, 相比非脱黏区, 脱黏区热信号衰减速率更慢, 且尺寸越大、深度越深的缺陷, 其热信号衰减速率越慢。最终脱黏区热信号衰减到与非脱黏区热信号接近, 但存在差异, 原因在于: 试验过程中, 受加热或表面发射率不均、热像仪灵敏度限制及噪声信号等因素的影响, 使得热信号为混叠信号, 不同时间段各混叠成分存在差异, 即降温阶段末期脱黏区与非脱黏区热信号具有差异性。

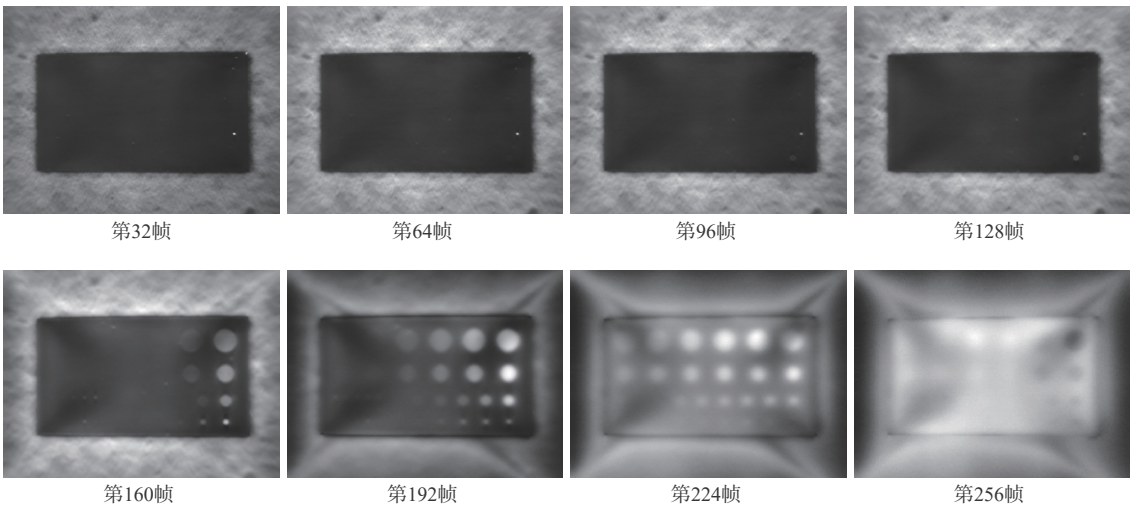


图3 脉冲激励结束后不同时刻试样表面热图

Fig. 3 Heat images of sample surface at different time after pulse excitation

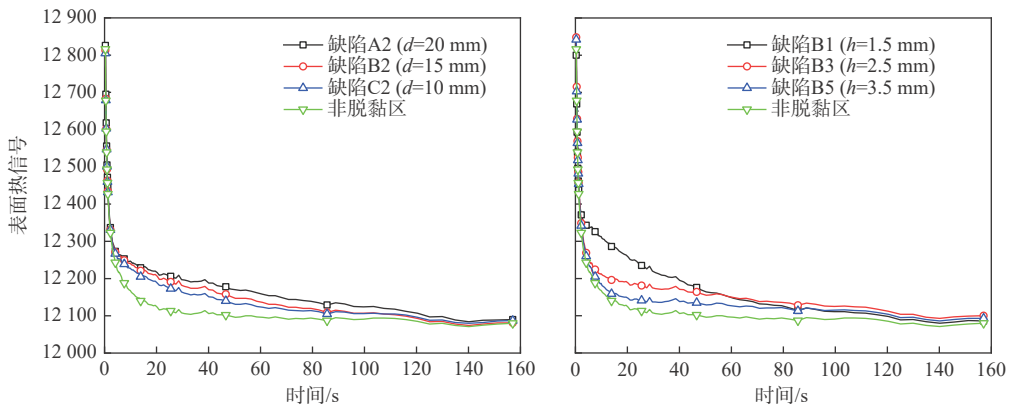


图4 表面热信号变化

Fig. 4 Surface thermal signals change

2.2 表面热图预处理

由于热像仪采集的原始热图序列中包含大量噪声，同时为了消除红外镜头反射影响及背景本身亮或暗区域影响，从提取的256帧热图序列中减去背景热图，相减的结果作为原始热图序列进入后处理过程。图5展示了减去背景热图的效果。图5(a)为脉冲激励开始前的背景热图，此时试样近表面温度与环境温度相同，即试样在自然状态下的热图；图5(b)为所截取热图序列的第192帧热图；图5(c)为图5(b)和图5(a)相减的结果。可知，减背景处理后热图整体质量更加清晰，热图整体噪声明显降低，而且热图中缺陷也更加清楚。下文中所有试验涉及的原始热图序列均指提取的256帧热图减背景后的热图序列。

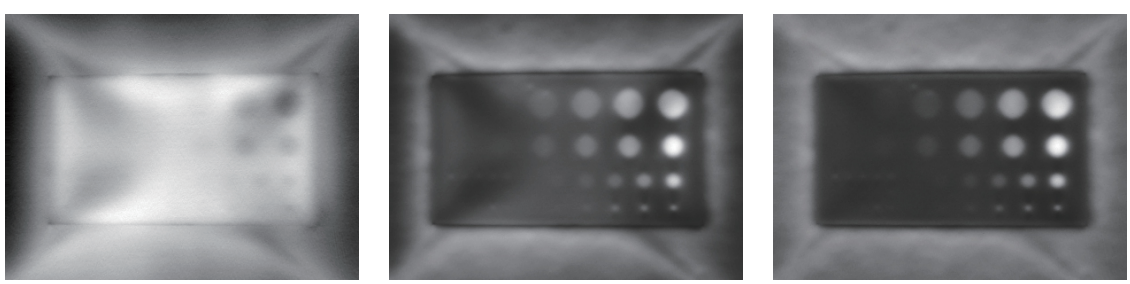


图5 原始热图减背景热图的结果

Fig. 5 Results of background heat image removed from original heat images

在试验过程中，红外脉冲热波成像对缺陷D4～D6、E1～E10、F1～F5的检测能力存在局限性，而缺陷A1～A6、B1～B6可检测性最好。由于试样放在水平桌面上，表面热图包括试样和桌面边

缘 2 部分, 为验证下文缺陷层析检测的可行性, 将桌面边缘及除 A1～A6、B1～B6 之外的缺陷部分切除, 裁剪后的目标图像如图 6 所示。

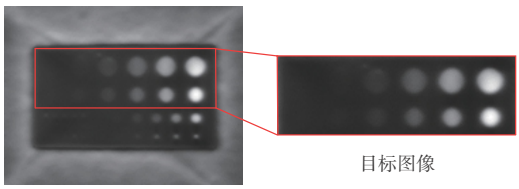


图 6　目标图像

Fig. 6　Target image

3　热信号极值时间确定及缺陷状态分析

3.1　热信号极值时间确定

通过热信号极值时间构造不同深度缺陷对应的断层图像以实现三维可视化。由于热信号极值时间受加热或表面发射率不均影响较大, 需对原始热图序列进行拟合以消除热信号极值时间的影响因素。采用基于对数多项式的拟合方法对每个像素点对应热信号响应的冷却阶段在对数域内用式 (1) 作拟合:

$$\ln[T(\tau)]=\sum_{n=1}^{N}a_n(\ln\tau)^n \tag{1}$$

每个像素点对应重构热信号可由式 (2) 表示:

$$T(\tau)=\exp\left[\sum_{n=1}^{N}a_n(\ln\tau)^n\right] \tag{2}$$

式中: T为重构热信号; τ为时间, s; N为拟合阶数。其中, N的取值依经验而定, 一般取 5 或 6 能够准确重构热信号, 经多次试算, 本文检测试验的经验值 $N=6$。

图 7 展示了不同拟合方法重构热信号对比曲线。可知, 相比其他拟合方法, 基于对数多项式拟合能够很好地逼近原始热信号。选取非脱黏区的基于对数多项式拟合的重构热信号作为参考热信号, 通过脱黏区重构热信号与参考热信号差极值得到脱黏区对应的热信号极值时间。图 8 展示了拟合前后脱黏区与非脱黏区热信号差随时间变化。可知, 拟合后热信号差值曲线更加平滑, 通过曲线极值点能够准确得到其对应的热信号极值时间 τ_m。图 9 展示了热信号极值时间图像。可知, 缺陷边缘受横向热扩散的影响较大。

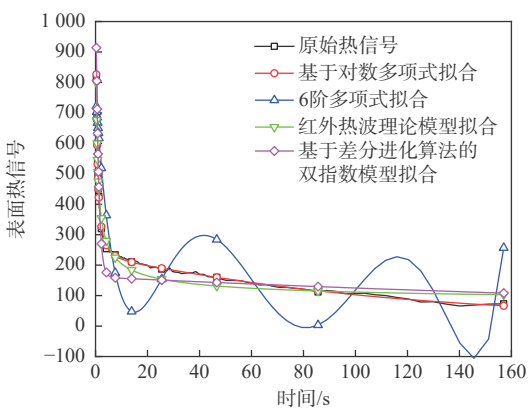


图 7　不同拟合方法重构热信号对比

Fig. 7　Comparison of reconstructed thermal signals by different fitting methods

图 10 展示了缺陷深度 2 mm, 缺陷尺寸对缺陷中心、边缘与非脱黏区热信号差的影响。可知, 对于给定缺陷深度, 缺陷尺寸减小或越靠近边缘, 热信号差极值减小, 相应极值时间也减小。图 11 展示了极值时间图像的剖面图。由图 10 和图 11 分析可知, 缺陷尺寸对热信号极值时间的影响较大, 易对缺陷深度判定造成较大误差。

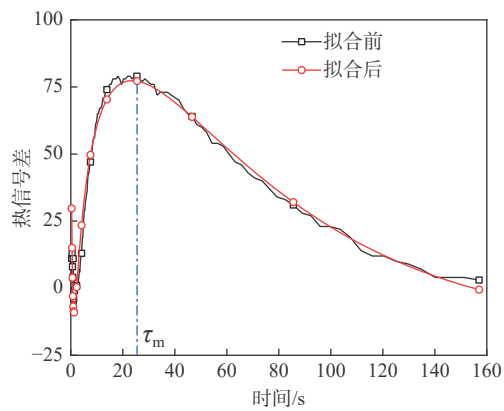


图 8　拟合前后热信号差变化

Fig. 8　Change of thermal signal difference before and after fitting

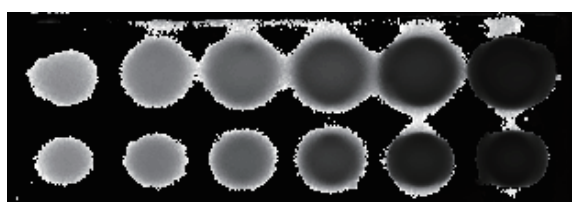


图 9　热信号极值时间图像

Fig. 9　Thermal signal extremum time image

3.2　缺陷状态分析

图 12 展示了极值时间图像的三维网格图。在实际检测过程中, 通过图 11、图 12 和式 (3) 对缺陷状态进行判定。结合 3.1 节分析可知, 忽略缺陷尺寸、边缘横向热扩散对极值时间的影响, 则可以判定缺陷上表面接近平面状态, 即缺陷 A1-B1、A2-B2、A3-B3、A4-B4、A5-B5、A6-B6 的上表面分别位于不同深度的同一水平面。

$$h=\sqrt{\frac{\alpha\tau_{\mathrm{m}}}{2}} \tag{3}$$

式中：$\alpha=\lambda/(\rho c)$为材料热扩散率，$\mathrm{m^2/s}$，λ为材料导热系数，ρ为材料密度，$\mathrm{kg/m^3}$，c为材料比热容，$\mathrm{J/(kg\cdot K)}$。

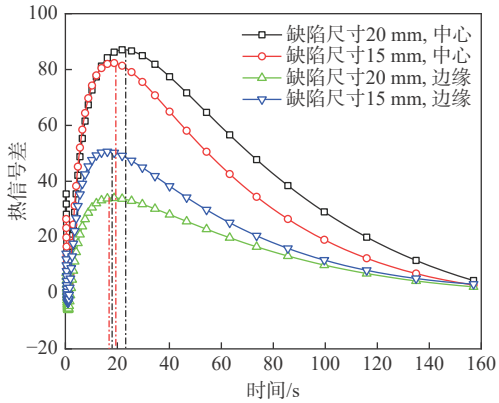


图 10 缺陷中心、边缘热信号差变化

Fig. 10 Change of thermal signal difference at defect center and edge

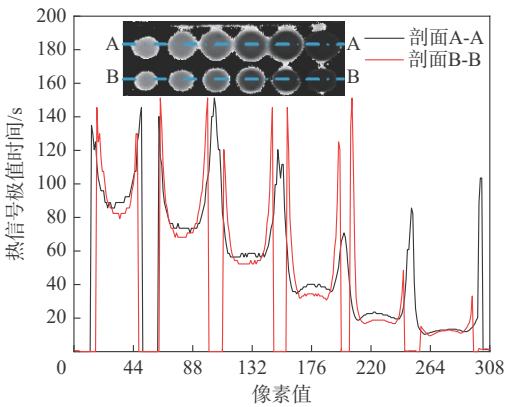


图 11 热信号校准时间剖面图

Fig. 11 Profile of extremum time image

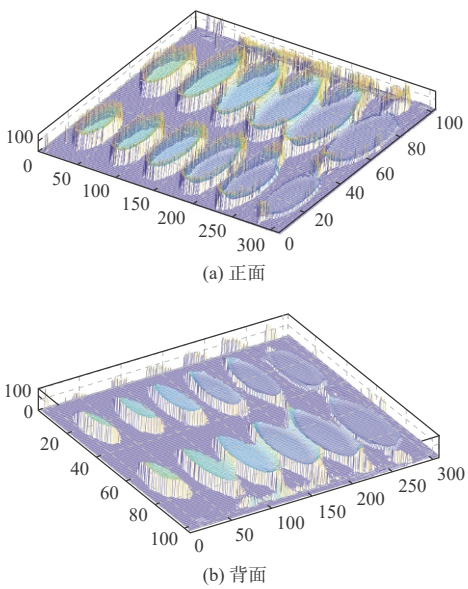


图 12 热信号极值时间三维网格图

Fig. 12 Three dimensional grid diagram of extremum time image

4 基于核函数模糊 C 均值聚类的极值时间数据处理

由图 10、图 11 可知，对于相同深度脱黏区，其对应的热信号极值时间存在差异，因此需要得到一个确定的极值时间以便表征缺陷深度。由于脱黏区对应热信号极值时间数组界限并不明确，采用模糊 C 均值聚类（fuzzy C-means, FCM）分类方法来确定相同缺陷深度对应的热信号极值时间，即将相同缺陷深度的极值时间数组进行分类，其中分布最为集中、个数最多一类的平均值作为该深度缺陷的极值时间。

采用 FCM 对极值时间图像中脱黏区对应的热信号极值时间数组进行二分类，FCM 分类结果如图 13 所示，白色部分为缺陷对应的极值时间数组，灰色部分为脱黏边缘横向热扩散对应的极值时间数组。可知，缺陷 A1-B1、A2-B2 与非脱黏区划分为一类，缺陷 A6-B6 对应极值时间数组分类效果不理想。



图 13 FCM 分类结果

Fig. 13 FCM classification results

为了提高分类精度，采用 KFCM 局部分类方法对脱黏区对应的热信号极值时间数组进行二分类，即对相同深度缺陷区域进行划分，利用核函数代替 FCM 目标函数中的欧氏距离准则，将划分区域中脱黏区对应的极值时间数组映射到高维空间进行二分类。核函数如式 (4) 所示，其对应的目标函数如式 (5) 所示：

$$k(\boldsymbol{x},\boldsymbol{y})=\exp\left(-\frac{\|\boldsymbol{x}-\boldsymbol{y}\|^2}{2\sigma^2}\right) \tag{4}$$

$$J(\boldsymbol{U},\boldsymbol{P})=\sum_{k=1}^{n}\sum_{i=1}^{c}\mu_{ik}(\boldsymbol{\Phi}(\boldsymbol{x}_k)-\boldsymbol{\Phi}(\boldsymbol{p}_i))^2 \tag{5}$$

式中：$\boldsymbol{x}$、$\boldsymbol{y}$为特征向量；σ为径向基核参数；$\boldsymbol{U}$为隶属度矩阵；$\boldsymbol{P}$为中心矩阵；μ_{ik}为隶属度；$\boldsymbol{\Phi}()$为基于核函数的特征映射。

采用 KFCM 局部分类方法对脱黏区进行二分类，试验过程中设定的加权指数为 5，最大迭代次数为 1 000，目标函数的迭代阈值为1×10^{-6}，径向基核参数$\sigma=300$。KFCM 分类结果如图 14 所示。可知，KFCM 局部分类后，缺陷对应热信号极值时间数组很好地划分为一类，分类精度有所提高。

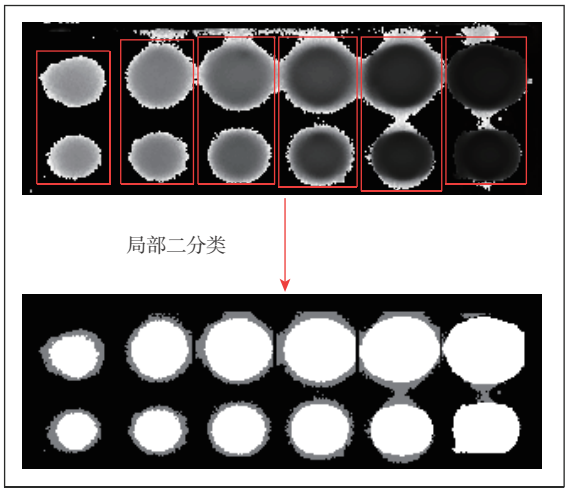


图 14　KFCM 分类结果

Fig. 14　KFCM classification results

5　断层图像及缺陷深度确定

5.1　断层图像确定

根据图 14 所示的 KFCM 分类结果，缺陷 Ai($i=1,2,\cdots,6$)对应的热信号极值时间分布最为集中、个数最多，因此计算相同缺陷深度对应极值时间数组的平均值 $\bar{\tau}_{\mathrm{mA}i}$($i=1,2,\cdots,6$)作为该缺陷深度对应的极值时间。通过如图 15 所示的流程，建立 $\bar{\tau}_{\mathrm{mA}i}$($i=1,2,\cdots,6$)与脱黏区对应极值时间 $\tau_{\mathrm{m}}(x,y)$之间的统计关系得到不同深度缺陷对应的二值化断层图像。

对断层图像分析可发现，缺陷边缘存在“毛刺”，并且图像中包含砂眼噪声。由于开运算处理后的图像只保留结构元素平移时能够填入其内部的部分，闭运算作为开运算的对偶运算相当于对原图像补集进行开运算，开闭运算可以在不明显改变物体面积和形状的情况下平滑物体边界，同时砂眼噪声相当于图像补集的孤立噪点。

因此，闭运算可以去除砂眼噪声。选取直径为 2 的方形结构元素，对断层图像进行开运算再进行闭运算，经形态学处理后的断层图像序列如图 16 所示。

5.2　缺陷深度确定

根据 5.1 节获取 $\bar{\tau}_{\mathrm{mA}i}$($i=1,2,\cdots,6$)的方法计算得到缺陷 B$i$($i=1,2,\cdots,6$)对应的极值时间 $\bar{\tau}_{\mathrm{mB}i}$($i=1,2,\cdots,6$)。为了减小缺陷尺寸对极值时间计算缺陷深度的影响，利用式 (6) 计算各断层图像对应的缺陷深度，经计算得到的缺陷深度和实际缺陷深度如表 2 所示。

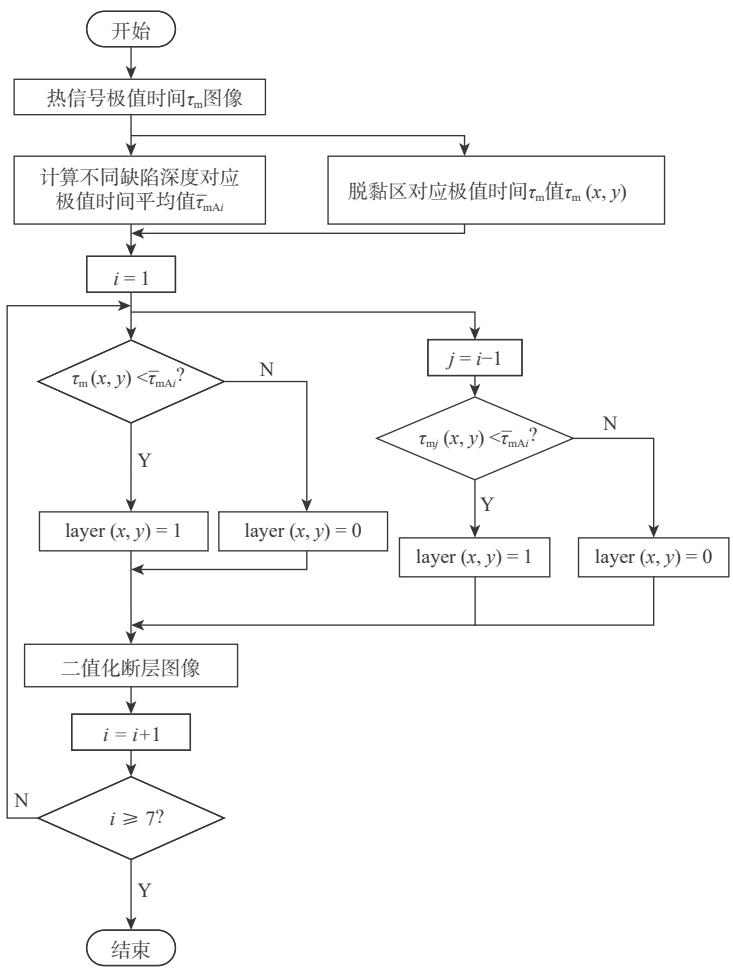


图 15　断层图像获取流程

Fig. 15　Tomographic images acquisition procedure

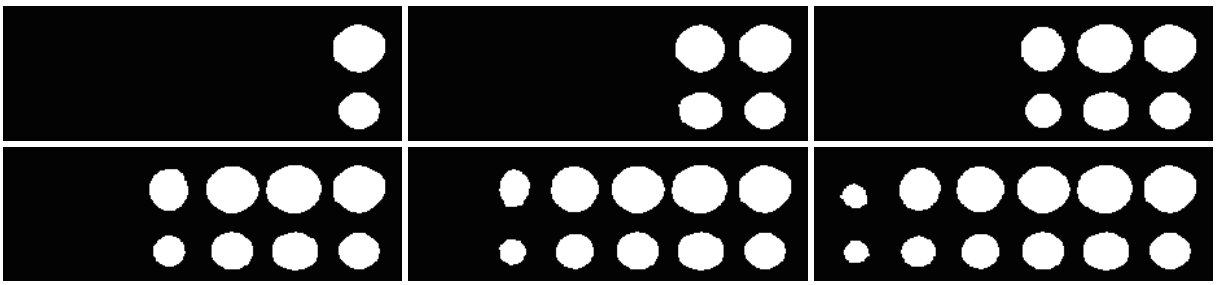


图 16 断层图像序列

Fig. 16 Tomographic images sequence

表 2 缺陷深度计算结果

Table 2 Calculation results of defect depth

编号	实际缺陷深度/mm	计算缺陷深度/mm	绝对误差/mm	相对误差/%
A6、B6	4	4.376	0.376	9.40
A5、B5	3.5	3.835	0.335	9.57
A4、B4	3	3.402	0.402	13.40
A3、B3	2.5	2.773	0.273	10.92
A2、B2	2	2.165	0.165	8.25
A1、B1	1.5	1.672	0.172	11.47

$$h=\sqrt{\frac{\alpha\bar{\tau}_{\mathrm{mA}i}}{2}}-\sqrt{\frac{\alpha(\bar{\tau}_{\mathrm{mA}i}-\bar{\tau}_{\mathrm{mB}i})}{2}},\ i=1,2,\cdots,6 \quad (6)$$

6 三维可视化

图 16 所示的二值化断层图像序列在 MATLAB 环境下通过等值面绘制方法实现三维可视化，由于断层图像只对脱黏区与非脱黏区做了区分，无法绘制试样外部轮廓，在进行等值面绘制前对断层图像边界数据赋值为 1 以便对试样轮廓进行绘制。三维可视化流程如图 17 所示。利用 isosurface 函数计算断层图像序列在显示平面的累计投影，再利用 patch 函数指定所有顶点坐标、各小面片顶点及颜色来创建面片，以图形对象的形式显示出来。

三维可视化效果如图 18 所示。可知，缺陷边缘受横向热扩散的影响较大，缺陷深度越深，缺陷上表面边缘的连续性越差。

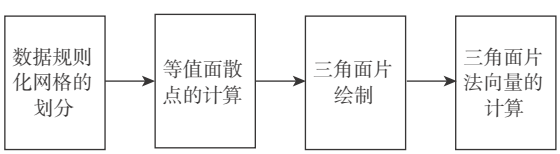


图 17 三维可视化流程

Fig. 17 3D visualization procedure

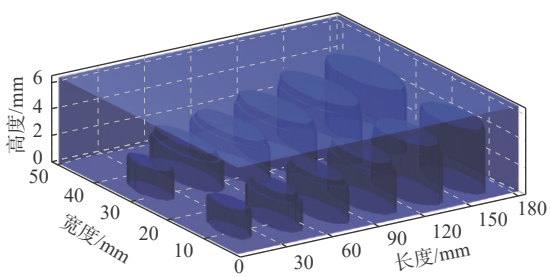


图 18 三维可视化效果

Fig. 18 3D visualization

7 结 论

1）红外脉冲热波层析成像检测技术能为提高 CFRP 层压板在服役期间的安全性、实现外场定量检测等需求提供解决方案。

2）基于对数多项式的拟合方法能很好地重构原始热信号，通过脱黏区重构热信号与非脱黏区参考热信号差，计算得到热信号极值时间及其图像。分析表明，在给定缺陷深度条件下，缺陷尺寸减小或越靠近边缘，对应极值时间减小，忽略缺陷尺寸对极值时间的影响，由此判定缺陷 A1-B1、A2-B2、A3-B3、A4-B4、A5-B5、A6-B6 的上表面接近平面状态，且分别位于不同深度的同一水平面。

3）核函数模糊 C 均值聚类局部分类方法有效提高了缺陷区极值时间数组的分类精度，通过计算数组平均值并与缺陷区极值时间数组建立统计关系，获得了不同深度的断层图像。

4）实现了基于红外脉冲热波成像检测 CFRP 层压板脱黏缺陷的三维层析重构。结果表明，红外脉冲热波层析成像能够实现缺陷分布及形貌的准确检测，层析检测缺陷深度与实际缺陷深度的最大相对偏差低于 15%。

参考文献（References）

[1] 尚金龙, 李思海, 陈贻明. 纤维增强塑料在航空航天领域中的应用[J]. 塑料工业, 2019, 47(1): 148-151.
SHANG J L, LI S H, CHEN Y M. Application of fiber reinforced plastic in aerospace field[J]. China Plastics Industry, 2019, 47(1): 148-151(in Chinese).

[2] 薄纯强. 碳纤维机翼缺陷的锁相红外无损检测研究[D]. 哈尔滨: 哈尔滨工业大学, 2019: 1.
BO C Q. Non-destructive testing research of CFRP airfoil defects

based on lock-in infrared thermography[D]. Harbin: Harbin Institute of Technology, 2019: 1(in Chinese).

[3] 梁涛. 复合材料脱粘缺陷红外热成像无损检测定量分析研究[D]. 成都: 电子科技大学, 2017: 1.
LIANG T. Quantitative analysis of infrared thermal imaging nondestructive testing for debonding defects of composite materials[D]. Chengdu: University of Electronic Science and Technology of China, 2017: 1(in Chinese).

[4] 李晓霞, 伍耐明, 段玉霞, 等. 碳纤维层合板低速冲击后的红外热波检测分析[J]. 复合材料学报, 2010, 27(6): 88-93.
LI X X, WU N M, DUAN Y X, et al. Infrared thermal wave imaging for carbon fiber laminated boards after low velocity impact[J]. Acta Materiae Compositae Sinica, 2010, 27(6): 88-93(in Chinese).

[5] PENG W, WANG F, MENG X L, et al. Dynamic thermal tomography based on continuous wavelet transform for debonding detection of the high silicon oxygen phenolic resin cladding layer[J]. Infrared Physics & Technology, 2018, 92(5): 115-121.

[6] 赵洪宝, 马丹, 马超群, 等. 航空用碳纤维复合材料典型缺陷无损检测技术研究[J]. 电子制作, 2020(24): 35-37.
ZHAO H B, MA D, MA C Q, et al. Study on nondestructive testing technology of typical defects of aviation carbon fiber composites for aviation[J]. Practical Electronics, 2020(24): 35-37(in Chinese).

[7] 南方, 肖鹏. 复合材料结构表面状态的红外热成像检测影响研究[J]. 航空制造技术, 2019, 62(14): 83-87.
NAN F, XIAO P. Research of composites surface influence on thermography detection technology[J]. Aeronautical Manufacturing Technology, 2019, 62(14): 83-87(in Chinese).

[8] 王扬, 李科, 刘俊岩. CFRP复合材料层板缺陷的红外热波成像检测方法[J]. 航空制造技术, 2016, 59(4): 36-42.
WANG Y, LI K, LIU J Y. Nondestructive testing and evaluation (NDT&E) for CFRP laminate with subsurface defects using infrared thermal wave imaging[J]. Aeronautical Manufacturing Technology, 2016, 59(4): 36-42(in Chinese).

[9] 张金玉, 杨正伟, 田干, 等. 红外热波检测及其图像序列处理技术[M]. 北京: 国防工业出版社, 2015: 16-17.
ZHANG J Y, YANG Z W, TIAN G, et al. Infrared thermal wave testing and images sequence processing technology[M]. Beijing: National Defense Industry Press, 2015: 16-17(in Chinese).

[10] 姜千辉, 姜长胜, 葛庆平, 等. 红外热波序列图像的图像分割与三维显示[J]. 无损检测, 2008, 30(2): 100-103.
JIANG Q H, JIANG C S, GE Q P, et al. Segmentation and 3D display of infrared thermal image[J]. Nondestructive Testing, 2008, 30(2): 100-103(in Chinese).

[11] TOIVANEN J M, TARVAINEN T, HUTTUNEN J M J, et al. Thermal tomography utilizing truncated Fourier series approximation of the heat diffusion equation[J]. International Journal of Heat and Mass Transfer, 2017, 108: 860-867.

[12] TAVAKOLIAN P, SIVAGURUNATHAN K, MANDELIS A. Enhanced truncated-correlation photothermal coherence tomography with application to deep subsurface defect imaging and 3-dimensional reconstructions[J]. Journal of Applied Physics, 2017, 122(2): 023103.

[13] MELNYK S I, MELNYK S S, TULUZOV I G. Method of projection dynamic thermal tomography (PDTT)[C]//International Conference on Quantitative Infrared Thermography. [S.l.]:[s.n.], 2012: 1-6.

[14] ELHASSNAOUI A, SAHNOUN S. A three-dimensional reconstruction algorithm for pulsed thermography[J]. Journal of Materials and Environmental Science, 2014, 5(4): 983-988.

[15] VAVILOV V P, NESTERUK D A, SHIRYAEV V V, et al. Thermal (infrared) tomography: Terminology, principal procedures, and application to nondestructive testing of composite materials[J]. Russian Journal of Nondestructive Testing, 2010, 46(3): 151-161.

[16] 彭伟. 药柱包覆层缺陷红外脉冲成像及动态层析检测技术研究[D]. 哈尔滨: 哈尔滨工业大学, 2020: 79-80.
PENG W. Study on infrared pulse imaging and dynamic chromatography detection technology for coating defects of grain[D]. Harbin: Harbin Institute of Technology, 2020: 79-80(in Chinese).

[17] VAVILOV V, MALDAGUE X, DUFORT B, et al. Thermal nondestructive testing of carbon epoxy composites: Detailed analysis and data processing[J]. NDT & E International, 1993, 26(2): 85-95.

[18] VAVILOV V. Dynamic thermal tomography: Recent improvements and applications[J]. NDT & E International, 2015, 71: 23-32.

[19] PAWAR S S, VAVILOV V P. Applying the heat conduction-based 3D normalization and thermal tomography to pulsed infrared thermography for defect characterization in composite materials[J]. International Journal of Heat and Mass Transfer, 2016, 94: 56-65.

[20] PENG W, WANG F, LIU J Y, et al. Pulse phase dynamic thermal tomography investigation on the defects of the solid-propellant missile engine cladding layer[J]. International Journal of Thermophysics, 2018, 39(4): 48-59.

Detection of debonding defect in CFRP laminates using infrared pulse thermal wave tomography

ZHONG Jiacen，XU Haojun，WEI Xiaolong*，HAN Xinmin，CHANG Yipeng

(School of Aeronautical Engineering，Air Force Engineering University，Xi'an 710038，China)

Abstract：To achieve the three-dimensional tomography for detecting debonding defects of carbon fiber reinforced plastic (CFRP) laminates based on infrared pulse thermal wave imaging, consummate the out-field quantitative detection guarantee system, and improve the safety and reliability of laminates in service, research on infrared pulse thermal wave tomography method and detection technology is carried out. A specimen with artificial debonding defects was prepared, infrared pulse thermal wave imaging technology is employed to detect debonding defects. The transient response process of surface thermal signal in debonding and sound region and detection ability of infrared pulse thermal wave imaging are analyzed. Through reconstructed thermal signal difference between debonding and sound region based on logarithmic polynomial fitting, extreme time of thermal signal are calculated, and the change law of debonding region extreme time and defect state are analyzed;. Kernel fuzzy C-means clustering is used to classify extreme time array corresponding to the same defect depth, and the array average value is calculated as defect extreme time. Statistical relationship between the time and debonding region extreme time array is established to construct tomographic images sequence, and the corresponding defect depth is calculated. Three-dimensional visualization of debonding defects in laminate is achieved by isosurface drawing method. Research shows that infrared pulse thermal wave tomography can detect debonding defects of CFRP laminate. This method can accurately and reliably display distribution and morphology of internal defects in laminate. Maximum relative deviation between detected and actual defect depth is less than 15%, which has certain guiding significance for engineering application.

Keywords：carbon fiber reinforced plastic；laminate；debonding defect；infrared pulsed thermal wave imaging；tomographic detection

Received：2021-09-16；**Accepted**：2022-01-05；**Published Online**：2022-01-21 09：23

URL：kns.cnki.net/kcms/detail/11.2625.V.20220120.1616.003.html

* **Corresponding author**. E-mail：wei18892022001@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0560

圆周均匀多相调制相关辐射计及其灵敏度分析

陈锐厚，罗景，胡岸勇*，苗俊刚
(北京航空航天大学 电子信息工程学院，北京 100191)

摘　　要：针对相关辐射计中由于运算放大器的零点漂移及增益起伏导致测量精度下降的问题，提出了在圆周上进行均匀多相调制的相关辐射计结构，利用最小二乘拟合方法对多相调制的测量结果进行处理，消除了运放零点漂移对相关辐射计测量偏差的影响。基于相关辐射计结构及噪声统计模型，推导了圆周均匀多相调制相关辐射计温度灵敏度的表达式，分析了系统灵敏度与调制阶数、增益起伏、调相误差的关系。分析表明：随着调制阶数的增加，增益起伏和调相误差对灵敏度的影响变小，圆周均匀多相调制相关辐射计的灵敏度趋于理想相关辐射计的灵敏度。

关　键　词：辐射测量；相关辐射计；灵敏度；多相调制；最小二乘拟合
中图分类号：TN98
文献标志码：A　　**文章编号**：1001-5965(2023)07-1857-07

辐射计是被动遥感领域用来测量物体自身辐射或反射电磁能量的设备。在微波毫米波频段，物体的电磁热辐射和其物理温度呈线性关系，因此，可通过测量物体的电磁热辐射获得物体的物理温度[1]。微波辐射计在射电天文学、卫星遥感等领域已经得到了广泛应用[2-4]。近年来，辐射测量技术在生物医疗、安防遥感领域也越来越受到重视[5-7]。灵敏度是辐射计的一个重要指标，其代表了辐射计可分辨的输入亮温的最小变化量，通常定义为输出信噪比等于1时对应的输入亮温改变量。辐射计的输出信噪比定义为：$\mathrm{SNR}=\overline{\Delta V_{\mathrm{out}}(\Delta T)}/\sigma_{\mathrm{out}}$，其中，$\overline{\Delta V_{\mathrm{out}}(\Delta T)}$表示所测量的物体温度改变$\Delta T$时引起的辐射计输出均值变化量，$\sigma_{\mathrm{out}}$表示输出波动的标准差。当$\overline{\Delta V_{\mathrm{out}}(\Delta T)}>\sigma_{\mathrm{out}}$时，即信噪比大于1时，视为该温度差能被检测出，此时的温度差ΔT即定义为该辐射计的灵敏度[8]。

相关辐射计是一种将接收天线的输出分为2路进行相关的辐射计。Faris推导了理想的相关辐射计工作在平衡状态下的灵敏度表达式[9]：$\Delta T=\sqrt{T_{\mathrm{sys}}^2\left(1+\sigma_G^2\right)/(B\tau)}$，其中，$T_{\mathrm{sys}}$为系统噪声温度，$B$为高频前端的等效带宽；$\tau$为系统的积分时间，$\sigma_G^2$为增益起伏的均方根。平衡状态即待测量物体和参考负载温度相等的状态，可以看出，工作在平衡状态的相关辐射计灵敏度受到系统增益起伏的影响很小。然而，在实际使用过程中，相关辐射计不可能一直工作在平衡状态。工作在非平衡状态的相关辐射计仍会受到增益起伏的影响而导致灵敏度恶化。此外，运算放大器的零点漂移也会导致测量出现偏差[10-12]。

针对此问题，本文提出了基于圆周均匀多相调制原理的相关辐射计结构，旨在消除运放零点漂移的影响，并降低运放增益起伏对灵敏度的影响。

1　多相调制相关辐射计的结构

本文研究的相关辐射计结构如图1所示。噪声信号X_{s}和Y_{s}分别表示辐射亮温为T_x和T_y的待测量物体和参考负载辐射的电磁信号，由魔T的2个接口输入，混合后经放大器放大得到S_1和S_2。S_1直接馈入0°功分器，S_2由调相器附加一个相位θ馈入90°功分器。z_1和z_2为放大器的热噪声，X_{s}、Y_{s}、z_1和

收稿日期：2021-09-17；录用日期：2022-01-09；网络出版时间：2022-01-12 12:44
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220112.0910.002.html
基金项目：国家自然科学基金(61731001)；中央高校基本科研业务费专项资金(FRF-BD-20-10A)
*通信作者. E-mail：hu_anyong@buaa.edu.cn

引用格式：陈锐厚，罗景，胡岸勇，等. 圆周均匀多相调制相关辐射计及其灵敏度分析[J]. 北京航空航天大学学报，2023，49(7)：1857-1863. CHEN R H，LUO J，HU A Y，et al. Uniform circular multiphase modulation correlation radiometer and its sensitivity analysis[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1857-1863 (in Chinese).

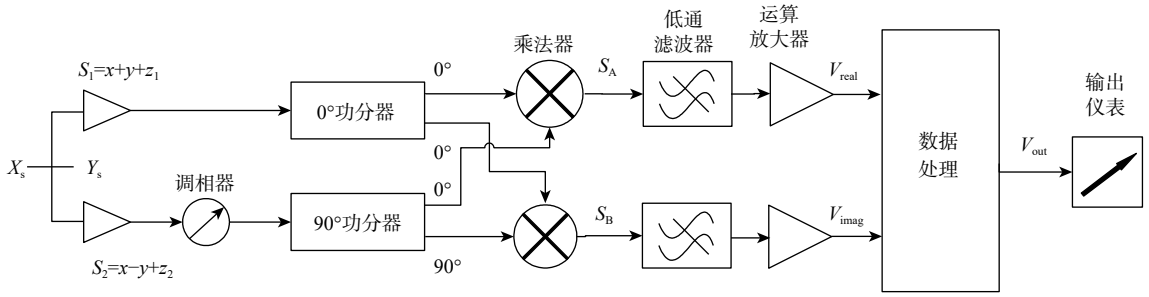


图1 多相调制相关辐射计结构

Fig. 1 Structure of multiphase modulation correlation radiometer

z_2均为平稳随机过程,且均值为0,互不相关。

经过信号分配网络和乘法器后的输出信号为

$$\begin{cases} S_{\mathrm{A}}(t)=S_1(t)S_2(t)= \\ \qquad \frac{1}{2}(x(t)+y(t)+z_1(t))(x'(t)-y'(t)+z_2'(t)) \\ S_{\mathrm{B}}(t)=S_1(t)\hat{S}_2(t)= \\ \qquad \frac{1}{2}(x(t)+y(t)+z_1(t))(\hat{x}'(t)-\hat{y}'(t)+\hat{z}_2'(t)) \end{cases} \tag{1}$$

式中:$\hat{S}_2(t)$表示$S_2(t)$的希尔伯特变换;x'表示信号x经过调相后的信号;$x(t)$与$y(t)$分别为经带宽B的带通放大器放大后的信号x和y的电压值,其功率谱密度分别为

$$\begin{cases} Q_x(f)=\frac{1}{2}akT_x \\ Q_y(f)=\frac{1}{2}akT_y \\ Q_{z_1}(f)=Q_{z_2}(f)=akT_z \end{cases} \tag{2}$$

式中:k为玻尔兹曼常数;常数a为放大器通带内的功率响应;T_z为放大器的等效噪声温度。因为魔T将源的功率平均分配给2路信号,所以有1/2的因子。

用带宽为b、通带为矩形的低通滤波器滤除信号的高频分量,即对一个信号进行平滑滤波,相当于在一段时间τ_0内对该信号取平均[13]。于是,相关值的2路输出为

$$\begin{cases} V_{\mathrm{real}}(t)=\frac{1}{2\tau_0}g_1(t)\int_{t-\tau_0}^{t}S_1(v)S_2(v)\mathrm{d}v+V_{\mathrm{r0}} \\ V_{\mathrm{imag}}(t)=\frac{1}{2\tau_0}g_2(t)\int_{t-\tau_0}^{t}S_1(v)\hat{S}_2(v)\mathrm{d}v+V_{\mathrm{i0}} \end{cases} \tag{3}$$

式中:$g_1(t)$和$g_2(t)$分别为2路运算放大器的增益;积分时间$\tau_0=1/(2b)$;V_{r0}和V_{i0}为放大器的直流偏置,来源于2路运算放大器的偏置电压,不随输入信号而改变,但随着温度和时间的变化缓慢漂移,在一段积分时间内视为常数[12]。

数据处理模块对2路输出信号进行采样、校正,并输出校正值给仪表显示温度测量值。

计算辐射计的灵敏度需要分别计算输出信号的直流分量幅值$\overline{V_{\mathrm{out}}}$和输出波动均方根$\sigma_{\mathrm{out}}^2$[14]。本文研究的相关辐射计采用复相关器,具有相关值的实部和虚部2路输出,因此分别对其进行计算[11]:

$$\overline{S_{\mathrm{A}}(t)}=\frac{1}{2}\overline{(x(t)+y(t)+z_1(t))(x'(t)-y'(t)+z_2'(t))} \tag{4}$$

式中:上划线表示取均值,如$\overline{S_{\mathrm{A}}(t)}=E[S_{\mathrm{A}}(t)]$。

式(4)中任意不相关的2项乘积期望为零,保留相关项乘积,化简得到

$$\overline{S_{\mathrm{A}}(t)}=\frac{1}{2}\overline{x(t)x'(t)}-\overline{y(t)y'(t)} \tag{5}$$

根据附录A,式(5)可化为

$$\overline{S_{\mathrm{A}}(t)}=\frac{1}{2}\int(Q_x-Q_y)\cos\theta\mathrm{d}f=\frac{1}{4}akB(T_x-T_y)\cos\theta \tag{6}$$

因此,

$$\overline{V_{\mathrm{real}}}HG_1\overline{S_{\mathrm{A}}(t)}+V_{\mathrm{r0}}=\frac{1}{4}HG_1akBT_0\cos\theta+V_{\mathrm{r0}} \tag{7}$$

式中:$T_0=T_x-T_y$;H和G_1分别为低通滤波器和放大器在零频率的放大倍数。

相同运算过程可求得

$$\overline{V_{\mathrm{imag}}}=\frac{1}{4}HG_2akBT_0\sin\theta+V_{\mathrm{i0}} \tag{8}$$

下面求输出波动均方根。根据Faris的理论[9],S_{A}的功率谱密度为

$$\begin{aligned} Q_{S_{\mathrm{A}}}(f)=&\frac{1}{4}\{[\psi_{xx}(0)-\psi_{yy}(0)]\cos\theta\cdot\delta(f)+ \\ &2(Q_x*Q_x+Q_y*Q_y)+ \\ &(Q_{z_1}+Q_{z_2})*(Q_x+Q_y)+Q_{z_1}*Q_{z_2}\}(f) \end{aligned} \tag{9}$$

式中:x、y分别为信号自相关函数在0处的取值;f为冲激函数"$*$"代表卷积。

式(9)中第1项为直流项,对输出的波动没有贡献。取式(9)中其余项的和,令

$$\begin{aligned} \tilde{Q}_{S_{\mathrm{A}}}(f)=&\frac{1}{4}\{2(Q_x*Q_x+Q_y*Q_y)+ \\ &(Q_{z_1}+Q_{z_2})*(Q_x+Q_y)+Q_{z_1}*Q_{z_2}\}(f) \end{aligned} \tag{10}$$

式(10)即代表输出波动项。假设低通滤波器是理想的,带宽为b,通带内频率响应$H(f)$等于常数。对一个信号的功率谱密度进行积分,可以得到其时域均方根[13]。

$$\sigma_{\text{real}}^2=\int \tilde{Q}_{V_{\text{real}}}(f)\mathrm{d}f=\int H^2(f)G_1^2(f)\tilde{Q}_{S_{\text{A}}}(f)\mathrm{d}f \tag{11}$$

由文献 [9] 得

$$\begin{cases} Q_x(f)*Q_x(f)\approx\dfrac{1}{8}(akT_x)^2B \\ Q_x(f)*Q_{z_1}(f)\approx\dfrac{1}{4}(ak)^2T_xT_zB \\ Q_{z_1}(f)*Q_{z_2}(f)\approx\dfrac{1}{2}(akT_z)^2B \end{cases} \tag{12}$$

根据式 (10) ～ 式 (12) 得到

$$\sigma_{\text{real}}^2=\frac{1}{8}a^2H^2G_1^2k^2BbT_{\text{sys}}^2 \tag{13}$$

式中: $T_{\text{sys}}=\sqrt{T_x^2+T_y^2+2T_xT_z+2T_yT_z+2T_z^2}$为系统的等效噪声温度。

改变信号的相位, 信号的功率谱密度不变。因此, 输出信号的实部和虚部波动均方根具有相同的表达形式:

$$\sigma_{\text{imag}}^2=\frac{1}{8}a^2H^2G_2^2k^2BbT_{\text{sys}}^2 \tag{14}$$

将 2 路输出信号写为直流分量和零均值的波动分量之和的形式:

$$\begin{cases} V_{\text{real}}(t)=G_{\text{C1}}T_0\cos\theta+n_1+V_{\text{r0}} \\ V_{\text{imag}}(t)=G_{\text{C2}}T_0\sin\theta+n_2+V_{\text{i0}} \end{cases} \tag{15}$$

式中: G_{C1}和G_{C2}分别为 2 路通道的增益因子, 即 $G_{\text{C1}}=HG_1akB/4$, $G_{\text{C2}}=HG_2akB/4$; 噪声项 n_1和n_2均值为 0, 均方根分别为σ_{real}^2和σ_{imag}^2。

2　多相调制

如第 1 节所述, 输出电压标准差的大小决定了灵敏度的高低。然而, 电路中由于器件的不理想性导致的其他误差, 如放大器的直流偏置或增益的起伏, 也会对辐射计的测量产生影响。因此, 需要采用多相调制的方法对这些误差的影响进行消减。

2.1　多相调制的基本原理

多相调制是一种通过多次改变输入信号的相位获取多组数据以对输出进行校正的方法。根据式 (15) 可以看出, 实部输出信号$V_{\text{real}}(t)$和虚部输出信号$V_{\text{imag}}(t)$是调相器的调相值θ的函数。控制调相器改变θ, 输出信号随之改变。

通常情况下, 待测物体的温度在一个积分时间τ_0内变化很小, 可以视为常数。将τ_0平均分为N段, 每段长度$\tau=\tau_0/N$, 并在每段时间τ结束时对输出信号进行采样, 得到N组测量值$\{(x_i,y_i)\}$, 即

$$\begin{cases} x_i=\dfrac{g_1\left(t-\dfrac{N-i}{N}\tau_0\right)}{\tau_0/N}\cdot\displaystyle\int_{t-\frac{N-i+1}{N}\tau_0}^{t-\frac{N-i}{N}\tau_0}S_1(v)S_2(v)\mathrm{d}v+V_{\text{r0}} \\ y_i=\dfrac{g_2\left(t-\dfrac{N-i}{N}\tau_0\right)}{\tau_0/N}\cdot\displaystyle\int_{t-\frac{N-i+1}{N}\tau_0}^{t-\frac{N-i}{N}\tau_0}S_1(v)\hat{S}_2(v)\mathrm{d}v+V_{\text{i0}} \end{cases} \tag{16}$$

式中: $i=1,2,\cdots,N$。每组测量值称为一组样本。利用这N组样本对测量物体的亮温进行估计, 从而消除一部分误差的影响。

假定输出信号$V(t)$为平稳随机过程, 在每个确定的时刻t_i所呈现的值是一个随机变量, 因此, 每组样本都是一组随机变量, 可表示为

$$\begin{cases} x_i=(1+\delta_{G_{1i}})(GT_0\cos\theta_i+n_{1i})+V_{\text{r0}} \\ y_i=(1+\delta_{G_{2i}})(GT_0\sin\theta_i+n_{2i})+V_{\text{i0}} \end{cases} \tag{17}$$

式中: θ_i为调相器的调制相位; 将G_{C1}写为$(1+\delta_{G_1})G$的形式, G_{C2}同理; δ_{G_1}和δ_{G_2}表示由于器件的不稳定性而产生的放大器增益波动, 假设均值为 0, 波动均方根都为σ_G^2, 且两者不相关; V_{r0}和V_{i0}代表运放的直流偏置; n_{1i}和n_{2i}均值为 0, 方差满足如下关系:

$$\begin{cases} D[n_{1i}]=N\sigma_{\text{real}}^2 \\ D[n_{2i}]=N\sigma_{\text{imag}}^2 \end{cases} \tag{18}$$

式 (18) 表明将积分时间平均分为N段后, 每段输出信号的方差会扩大N倍。这是因为积分时间缩短为原来的$1/N$相当于将低通滤波器的带宽b扩展N倍, 由式 (13), 输出信号的方差随之变大。从信号的功率谱角度解释, 低通滤波器的带宽变大, 经过其滤波的信号残留的高频谐波更多, 信号的波动性更强。

将这N组样本点画在 x-y 坐标系中。理想情况下, 样本点(x_i,y_i)应该分布于一个圆上, 称为复相关圆, 如图 2 所示。样本点在复相关圆上的所处位置取决于调制相位θ。若θ均匀取值, 则样本点也是等间隔均匀分布的。复相关圆的圆心由 2 路输出信号的直流偏置决定; 圆半径等于GT_0, 包含了T_0, 求得复相关圆的半径, 就可求出T_0。但是, 考虑到第 1 节叙述的各类偏差对输出结果的影响, 实际上的样本点会偏离复相关圆。因此, 需要找到一种合理的方法对复相关圆进行拟合。

2.2　最小二乘法拟合复相关圆

最小二乘法是曲线拟合常用的方法。其估计

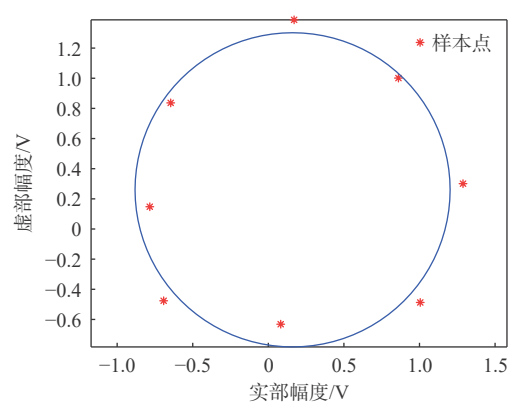


图 2　8 个样本点时的复相关圆

Fig. 2　Complex correlation circle at eight sample points

准则是选取一个残差平方和作为目标函数，将待估计的参数作为变量，将样本点作为已知量对目标函数进行优化。改变变量的取值，目标函数L的值随之变化。当变量取到某点时，L取得了值域上的最小值，此时的变量取值即作为这个参数的估计。最小二乘法用于线性问题中的求解有一个封闭的解决方案，但应用于参数非线性的问题中往往需要利用较复杂的优化方法来求解，这与待估计曲线和目标函数的选取有关[15]。

曲线拟合的第1步是选取待拟合曲线的表达形式。基于复相关圆的分析，选取带参数的曲线方程为

$$(x-p)^2+(y-q)^2=R^2 \tag{19}$$

式中：p、q分别为拟合曲线圆心的横、纵坐标；R为拟合圆的半径。

第2步是选取合适的目标函数。取目标函数为

$$L=\sum_{i=1}^{N}(x_i-R\cos\theta_i-p)^2+(y_i-R\sin\theta_i-q)^2 \tag{20}$$

当选定一个θ_i时，样本点应分布于拟合圆上某点周围，该点的坐标为$(R\cos\theta_i-p,R\sin\theta_i-q)$。将样本点与该点的距离作为残差，进而确定目标函数。

采用多元函数求极值的方法，分别对3个参数求偏导数，并求出驻点：

$$\begin{cases}\dfrac{\partial L}{\partial R}=\displaystyle\sum_{i=1}^{N}\cos\theta_i(R\cos\theta_i+p-x_i)+\\ \qquad\displaystyle\sum_{i=1}^{N}\sin\theta_i(R\sin\theta_i+q-y_i)=0\\ \dfrac{\partial L}{\partial a}=\displaystyle\sum_{i=1}^{N}(R\cos\theta_i+p-x_i)=0\\ \dfrac{\partial L}{\partial b}=\displaystyle\sum_{i=1}^{N}(R\sin\theta_i+q-y_i)=0\end{cases} \tag{21}$$

由于L是平方和的形式，一定存在一个大于或等于0的最小值，最大值为无穷大。又由于该函数仅有一个驻点，L的最小值就在这个点上取得。求得驻点，也即对参数的估计：

$$\begin{cases}\hat{p}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{N}(x_i-\hat{R}\cos\theta_i)\\ \hat{q}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{N}(y_i-\hat{R}\sin\theta_i)\\ \hat{R}=\dfrac{\overline{x_i\cos\theta_i}+\overline{y_i\sin\theta_i}-\overline{x_i}\cdot\overline{\cos\theta_i}-\overline{y_i}\cdot\overline{\sin\theta_i}}{(1-\overline{\cos\theta_i}^2-\overline{\sin\theta_i}^2)}\end{cases} \tag{22}$$

式中：N为样本数即调制阶数；上划线表示取均值，如：

$$\overline{x_i\cos\theta_i}=\frac{1}{N}\sum_{i=1}^{N}x_i\cos\theta_i \tag{23}$$

假如θ是在0～2π均匀取值，则有

$$\overline{\cos\theta_i}=\overline{\sin\theta_i}=0 \tag{24}$$

估计表达式可以进一步化简，得到圆心坐标的估计：

$$\begin{cases}\hat{p}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{N}x_i=\overline{x_i}\\ \hat{q}=\dfrac{1}{N}\displaystyle\sum_{i=1}^{N}y_i=\overline{y_i}\end{cases} \tag{25}$$

半径的估计表达式为

$$\hat{R}=\overline{x_i\cos\theta_i}+\overline{y_i\sin\theta_i}=\overline{\delta_{G_1i}\cos\theta_i(G\Delta T\cos\theta_i+n_{1i})}+\overline{\delta_{G_2i}\sin\theta_i(G\Delta T\sin\theta_i+n_{2i})}+G\Delta T+\overline{V_{\mathrm{r0}}\cos\theta_i}+\overline{V_{\mathrm{i0}}\sin\theta_i} \tag{26}$$

注意到：

$$\overline{V_{\mathrm{r0}}\cos\theta_i}=\overline{V_{\mathrm{i0}}\sin\theta_i}=0 \tag{27}$$

说明运算放大器的直流偏置影响被消除了。

3 灵敏度的推导及分析

数据处理模块的作用是在每一段调相周期τ结束时对相关器的2路输出$V_{\mathrm{real}}(t)$和$V_{\mathrm{imag}}(t)$采样、记录并进行复相关圆拟合，再将圆半径的拟合结果V_{out}输出给仪表，仪表根据V_{out}的值显示温度测量值[16]。因此，要计算灵敏度，需要计算V_{out}的波动均方根和由于待测物体温度T_x改变而引起的输出均值变化量$\Delta V_{\mathrm{out}}(\Delta T)$。

首先，计算V_{out}的输出均值，根据式(17)有

$$\begin{cases}E[x_i\cos\theta_i]=\cos^2\theta_iGT_0+V_{\mathrm{r0}}\cos\theta_i\\ E[y_i\sin\theta_i]=\sin^2\theta_iGT_0+V_{\mathrm{i0}}\sin\theta_i\end{cases} \tag{28}$$

则

$$E[V_{\mathrm{out}}]=E\left[\frac{1}{N}\sum_{i=1}^{N}(x_i\cos\theta_i+y_i\sin\theta_i)\right]=GT_0 \tag{29}$$

进而

$$E[V_{\mathrm{out}}(\Delta T)]=E[V_{\mathrm{out}}(T_x+\Delta T)-V_{\mathrm{out}}(T_x)]=G(T_0+\Delta T)-GT_0=G\Delta T \tag{30}$$

然后，计算输出波动均方根：

$$\begin{aligned}D[x_i]&=D\left[(1+\delta_{G_1})(GT_0\cos\theta_i+n_{1i})+V_{\mathrm{r0}}\right]=\\ &(GT_0\cos\theta_i)^2D\left[\delta_{G_1}\right]+D\left[(1+\delta_{G_1})n_{1i}\right]=\\ &(GT_0\cos\theta_i)^2\sigma_G^2+N\sigma_{V_{\mathrm{real}}}^2\left(1+\sigma_G^2\right)\end{aligned} \tag{31}$$

相同运算过程可得

$$D\left[y_i\right]=(GT_0\sin\theta_i)^2\sigma_G^2+N\sigma_{\mathrm{real}}^2\left(1+\sigma_G^2\right) \tag{32}$$

根据式(26)和式(27)有

$$D[V_{\mathrm{out}}]=D\left[\overline{x_i\cos\theta_i}+\overline{y_i\sin\theta_i}\right]=\frac{1}{N^2}\sum_{i=1}^{N}\left(\cos^2\theta_iD[x_i]+\sin^2\theta_iD\left[y_i\right]\right) \tag{33}$$

将式(13)、式(14)、式(31)、式(32)代入式(33)

中得到

$$D[V_{\text{out}}]=\frac{1}{N}\left(\overline{\sin^4\theta_i}+\overline{\cos^4\theta_i}\right)(GT_0)^2\sigma_G^2+\frac{2b}{B}G^2T_{\text{sys}}^2\left(1+\sigma_G^2\right) \tag{34}$$

令信噪比等于 1, 即令输出均值与输出波动标准差相等: $\overline{\Delta V_{\text{out}}(\Delta T)}=\sqrt{D[V_{\text{out}}]}$, 解得灵敏度为

$$\Delta T=\sqrt{\frac{T_{\text{sys}}^2\left(1+\sigma_G^2\right)}{B\tau}+\frac{\left(\overline{\sin^4\theta_i}+\overline{\cos^4\theta_i}\right)T_0{}^2\sigma_G^2}{N}} \tag{35}$$

以上分析是在调制相位θ_i精确的基础上进行的, 但调相器的不理想性会给调制相位带来波动。因此, θ_i也应视为一个随机变量。将其写为 2 项之和的形式, 如式 (36) 所示, 代表了在理想的调制相位上附加了一个误差项。 假设该误差项均值为 0, 方差为σ_θ^2。

$$\begin{cases}x_i=(1+\delta_{G_{1i}})(GT_0\cos(\theta_{i0}+\Delta\theta_i)+n_{1i})+V_{\text{r0}}\\ y_i=(1+\delta_{G_{2i}})(GT_0\sin(\theta_{i0}+\Delta\theta_i)+n_{2i})+V_{\text{i0}}\end{cases} \tag{36}$$

三角变换:

$$\begin{cases}\cos(\theta_{i0}+\Delta\theta_i)=\cos\theta_{i0}\cos\Delta\theta_i-\sin\theta_{i0}\sin\Delta\theta_i\\ \sin(\theta_{i0}+\Delta\theta_i)=\cos\theta_{i0}\sin\Delta\theta_i+\sin\theta_{i0}\cos\Delta\theta_i\end{cases} \tag{37}$$

对式 (37) 中包含$\Delta\theta_i$的项进行泰勒展开:

$$\begin{cases}\cos(\Delta\theta_i)=1-\dfrac{(\Delta\theta_i)^2}{2!}+\dfrac{(\Delta\theta_i)^4}{4!}+\ldots\\ \sin(\Delta\theta_i)=(\Delta\theta_i)-\dfrac{(\Delta\theta_i)^3}{3!}-\ldots\end{cases} \tag{38}$$

当调相器的误差$\Delta\theta_i$比较小时, 取展开式的第 1 项, 式 (37) 化为

$$\begin{cases}\cos(\theta_{i0}+\Delta\theta_i)=\cos\theta_{i0}-\sin\theta_{i0}\cdot\Delta\theta_i\\ \sin(\theta_{i0}+\Delta\theta_i)=\sin\theta_{i0}+\cos\theta_{i0}\cdot\Delta\theta_i\end{cases} \tag{39}$$

将式 (39) 代回式 (36) 中, 得到x_i和y_i的近似表达式为

$$\begin{cases}x_i=(1+\delta_{G_{1i}})[GT_0(\cos\theta_i-\sin\theta_{i0}\cdot\Delta\theta_i)+n_{1i}]+V_{\text{r0}}\\ y_i=(1+\delta_{G_{2i}})[GT_0(\sin\theta_i+\cos\theta_{i0}\cdot\Delta\theta_i)+n_{2i}]+V_{\text{i0}}\end{cases} \tag{40}$$

则

$$\begin{aligned}x_i\cos\theta_{i0}+y_i\sin\theta_{i0}=&GT_0[\delta_{G_{1i}}\cos^2\theta_{i0}+\delta_{G_{2i}}\sin^2\theta_{i0}+\\ &\sin\theta_{i0}\cos\theta_{i0}(\delta_{G_{1i}}-\delta_{G_{2i}})\Delta\theta_i+1]+\\ &(1+\delta_{G_{1i}})\cos\theta_{i0}\cdot n_{1i}+(1+\delta_{G_{2i}})\sin\theta_{i0}\cdot n_{2i}\end{aligned} \tag{41}$$

重新计算V_{out}的方差和$\Delta V_{\text{out}}(\Delta T)$的均值:

$$\begin{aligned}D[V_{\text{out}}]=D\left[x_i\cos\theta_i+y_i\sin\theta_i\right]=&\\ \frac{1}{N}\left(\overline{\sin^4\theta_i}+\overline{\cos^4\theta_i}\right)G^2T_0^2\sigma_G^2+&\\ \frac{2}{N}\left(\overline{\cos^2\theta_{i0}\sin^2\theta_{i0}}\right)G^2T_0^2\sigma_G^2\sigma_\theta^2+&\\ \frac{2b}{B}G^2T_{\text{sys}}^2\left(1+\sigma_G^2\right)&\end{aligned} \tag{42}$$

$$E[\Delta V_{\text{out}}(\Delta T)]=E\left[x_i\cos\theta_i+y_i\sin\theta_i\right]=G\Delta T \tag{43}$$

令信噪比等于 1, 得到灵敏度的表达式为

$$\Delta T=\sqrt{\frac{T_{\text{sys}}^2\left(1+\sigma_G^2\right)}{B\tau}+\frac{T_0^2\sigma_G^2\left(\alpha+\beta\sigma_\theta^2\right)}{N}} \tag{44}$$

式中:

$$\alpha=\frac{1}{N}\sum_{i=1}^{N}\sin^4\theta_{i0}+\cos^4\theta_{i0}$$

$$\beta=\frac{2}{N}\sum_{i=1}^{N}\cos^2\theta_{i0}\sin^2\theta_{i0}$$

当 $N=2$或4时, $\alpha=1$, $\beta=0$; $N=3$或$N>4$时, $\alpha=3/4$, $\beta=1/4$。

式 (44) 即圆周均匀多相调制相关辐射计的灵敏度表达式, 表明灵敏度受到放大器增益起伏、调相器调相误差的影响, 且与待测量物体与参考负载的温度差T_0有关。当圆周均匀多相调制相关辐射计工作在平衡状态, 即$T_0=0$时, 灵敏度表达式与传统的理想相关辐射计相同。

当辐射计工作在非平衡状态, 即$T_0\neq0$时, 放大器的增益起伏对灵敏度的影响较大, 且该影响随着调制阶数N的增大而降低。

在调相误差$\Delta\theta$比较小的情况下, 满足$\alpha+\beta\sigma_\theta^2\approx\alpha$, 调相误差对该辐射计的影响不大。

4　结　论

本文针对传统相关辐射计放大器的增益起伏和零点漂移导致的辐射亮温测量误差问题, 提出了多相调制相关辐射计的结构, 推导了在圆周上进行均匀多相调制的相关辐射计的温度灵敏度的解析表达式, 分析表明:

1) 运算放大器的零点漂移影响被消除。

2) 随着调制阶数的增加, 运算放大器的增益起伏对灵敏度的影响变小, 多相调制相关辐射计的灵敏度趋于理想相关辐射计的灵敏度。

3) 调相器的调制相位θ以预设相位θ_{i0}为中心, 且波动不大时, 对灵敏度的影响可忽略。

参考文献 (References)

[1] ORABI F T, MOORE R K. 微波遥感和辐射测量学[M]. 冯健超, 译. 北京: 科学出版社, 1988: 115-119.
ORABI F T, MOORE R K. Microwave remote sensing and radiometry[M]. FENG J C, translated. Beijing: Science Press, 1988: 115-119(in Chinese).

[2] TIURI M E. Radio telescope receivers[J]. Radio Astronomy, 1986, 7: 452-457.

[3] TIURI M E. Radio astronomy receivers[J]. IEEE Transactions on Military Electronics, 1964, 8(3): 264-272.

[4] 李延明, 姜敏, 余锐, 等. 海洋二号卫星微波辐射计系统设计[J]. 中国工程科学, 2013, 15(7): 39-43.
LI Y M, JIANG M, YU R, et al. System design of microwave radiometer of HY-2A satellite[J]. Strategic Study of CAE, 2013, 15(7): 39-43(in Chinese).

[5] CHANDRA R, ZHOW H Y, BALASINGHAM I, et al. On the opportunities and challenges in microwave medical sensing and imaging[J]. IEEE Transactions on Bio-Medical Engineering, 2015, 62(7): 1667-1682.

[6] NANZER J A. 微波毫米波安防遥感技术[M]. 苗俊刚, 胡岸勇, 孙国琳,等 译. 北京: 机械工业出版社, 2015: 156-162.
NANZER J A. Microwave and millimeter-wave remote sensing for security applications[M]. MIAO J G, HU A Y, SUN G L, et al, translated. Beijing: China Machine Press, 2015: 156-162(in Chinese).

[7] 苗俊刚, 郑成, 胡岸勇, 等. 被动毫米波实时成像技术[J]. 微波学报, 2013, 29(增刊1): 100-112.
MIAO J G, ZHENG C, HU A Y, et al. Real-time passive millimeter wave imaging technology[J]. Journal of Microwaves, 2013, 29(Sup 1): 100-112(in Chinese).

[8] HERSMAN M S, POE G A. Sensitivity of the total power radiometer with periodic absolute calibration[J]. IEEE Transactions on Microwave Theory and Techniques, 1981, 29(1): 32-40.

[9] FARIS J J. Sensitivity of a correlation radiometer[J]. Journal of Research of the National Bureau of Standards, Section C:Engineering and Instrumentation, 1967, 71C(2): 153.

[10] 桂良启, 张祖荫, 郭伟. 提高微波辐射计灵敏度研究的综述[J]. 遥测遥控, 2004, 25(6): 1-5.
GUI L Q, ZHANG Z Y, GUO W. Review of studies of microwave radiometer sensitivity improvements[J]. Telemetry & Telecontrol, 2004, 25(6): 1-5(in Chinese).

[11] 陈晨. 用于毫米波干涉成像的多芯片组件MCM复相关器研究[D]. 北京: 北京航空航天大学, 2020: 55-60.
CHEN C. Study on multi-chip module complex correlator for millimeter-wave interferometric imaging[D]. Beijing: Beihang University, 2020: 55-60(in Chinese).

[12] WANG C, XIN X, LIANG B Y, et al. Quadrature errors and DC offsets calibration of analog complex cross-correlator for interferometric passive millimeter-wave imaging applications[J]. Sensors, 2018, 18(2): 677.

[13] WAIT F D. The sensitivity of the dicke radiometer[J]. Journal of Research of the National Bureau of Standards, 1967, 71C(2): 127-152.

[14] 周荫清. 随机过程理论[M]. 3版. 北京: 北京航空航天大学出版社, 2013: 113-1116.
ZHOU Y Q. Principle of stochastic process[M]. 3rd ed. Beijing: Beihang University Press, 2013: 113-116(in Chinese).

[15] GANDER W,GOLUB G H, STREBEL R. Least-squares fitting of circles and ellipses[J]. BIT Numerical Mathematics, 1994, 34(4): 556-577.

[16] BREMER J C. Improvement of scanning radiometer performance by digital reference averaging[J]. IEEE Transactions on Instrumentation and Measurement, 1979, 28(1): 46-54.

附录 A:

令信号$x(t)$经过调相器后的信号为$y(t)$,其各自的傅里叶变换分别为$X(\omega)$、$Y(\omega)$。调相器的传输函数为

$$H(\omega)=\mathrm{e}^{\mathrm{i}\theta}$$

易知满足$Y(\omega)=X(\omega)H(\omega)=X(\omega)\mathrm{e}^{\mathrm{i}\theta}$。

根据互相关理论:

$$\begin{aligned}\overline{x(t)y(t)}&=\int \mathrm{Re}\left[\lim_{T\to\infty}\frac{1}{T}(X(\omega)Y(\omega))\right]\mathrm{d}\omega=\\&\int\lim_{T\to\infty}\frac{1}{2T}(X(\omega)Y^{*}(\omega)+X^{*}(\omega)Y(\omega))\mathrm{d}\omega=\\&\int\lim_{T\to\infty}\frac{1}{2T}\left[X^{2}(\omega)(\mathrm{e}^{-\mathrm{i}\theta}+\mathrm{e}^{\mathrm{i}\theta})\right]\mathrm{d}\omega=\\&\cos\theta\int\lim_{T\to\infty}\frac{X^{2}(\omega)}{T}\mathrm{d}\omega=\cos\theta\int Q_{x}d\omega\end{aligned}$$

Uniform circular multiphase modulation correlation radiometer and its sensitivity analysis

CHEN Ruihou，LUO Jing，HU Anyong*，MIAO Jungang

(School of Electronic Information Engineering，Beihang University，Beijing 100191，China)

Abstract：A correlation radiometer structure with uniform multiphase modulation around the circumference is proposed in order to address the issue of decreased measurement accuracy caused by the operational amplifier's offset drift and gain fluctuation in correlation radiometers. The measurement result is processed using the least square fitting method, which removes the influence of the operational amplifier's offset drift on the measurement deviation. Based on the structure of the correlation radiometer and the noise statistical model, the expression of the temperature sensitivity of the uniform circular multiphase modulation correlation radiometer is computed, and the relationship between the system sensitivity and phase modulation order, gain fluctuation and phase shift error is analyzed. The analysis shows that with the increase of the modulation order, the influence of gain fluctuation and phase modulation error on the sensitivity becomes smaller, and the sensitivity of the uniform circular multiphase modulation correlation radiometer tends to the sensitivity of the ideal correlation radiometer.

Keywords：radiometry; correlation radiometer; sensitivity; multiphase modulation; least squares fitting

Received：2021-09-17；**Accepted**：2022-01-09；**Published Online**：2022-01-12 12：44

URL：kns.cnki.net/kcms/detail/11.2625.V.20220112.0910.002.html

Foundation items：National Natural Science Foundation of China (61731001); The Fundamental Research Funds for the Central University (FRF-BD-20-10A)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0563

基于变弯度后缘的机翼阵风响应减缓数值研究

尉濡恺，戴玉婷*，杨超，于思恒
(北京航空航天大学 航空科学与工程学院，北京 100191)

摘　　要：针对带有变弯度后缘的机翼建立了阵风响应分析的数学模型，并开展了阵风响应减缓的仿真研究。采用计算流体力学（CFD）方法计算变弯度后缘给定动态偏转运动下的广义非定常气动力，并基于状态观测器法辨识CFD数据建立后缘动态偏转下的广义气动力模型，采用面元法计算模态运动、阵风引起的广义气动力，利用广义预测控制（GPC）方法设计阵风减缓控制律，在此基础上对变弯度后缘与传统铰链舵面动态气动力特性进行对比。仿真结果表明：基于变弯度后缘的GPC方法能够有效减缓由阵风引起的机翼翼尖加速度响应，翼尖加速度减缓效率为44.33%；相比传统铰链舵面，变弯度后缘偏转时弦向剖面上下表面压力分布更连续，相同舵偏下对机翼动态气动力影响更大，阵风响应减缓效率也更高，采用变弯度后缘进行阵风减缓具有更为广阔的应用前景。

关 键 词：变弯度后缘；阵风响应减缓；降阶模型；主动控制；非定常气动力
中图分类号：V211.47; TB553
文献标志码：A　　**文章编号**：1001-5965(2023)07-1864-11

飞行器在大气中飞行，不可避免地受到阵风的影响。阵风所附加的气动载荷引发飞行器飞行状态的改变，过大幅值的阵风影响飞行的性能与安全。阵风响应减缓的原理即是通过主动控制舵面产生尽可能大的气动反作用力来抑制阵风干扰，传统飞行器多采用铰链舵面偏转来减缓阵风响应。

与传统铰链舵面相比，变弯度后缘可以在保证机翼表面连续的情况下，减小气动阻力与噪声，提高气动效率[1-3]，变弯度后缘将会逐步取代传统铰链舵面成为主要的气动舵面，因此，利用变弯度后缘进行阵风减缓主动控制具有迫切的工程需求。如何快速准确预测变弯度后缘偏转引起的气动力是利用变弯度后缘进行阵风减缓主动控制的关键所在。

目前，工程上主要采用基于位势流理论的面元法来计算后缘偏转的广义气动力。Nguyen等将三维机翼分解为一系列二维机翼，利用二元Theodorson理论来构建可变弯度机翼气动弹性模型[4-7]；FlexSys公司开发的连续可变弯度机翼后缘襟翼（variable camber continuous trailing edge flap, VCCTEF）采用基于面元法的商业软件ZAERO计算了后缘偏转气动力，完成了对VCCTEF的颤振分析[8-9]；Woods等通过给基于平板面元法的Xfoil添加黏性边界层求解器来计算鱼骨变弯度翼型的升阻力特性与压力分布[3]；吴志刚等采用偶极子格网法加有理函数拟合技术计算舵面偏转广义气动力[10]。

面元法已经应用于许多成熟的商业软件，使用方便快捷，但计算精度相对较低。为了提高精度，计算流体力学(computational fluid dynamics, CFD)计算方法近年来已开始应用于变弯度后缘气动力分析与飞行器阵风响应计算中。Dhileep等采用商业软件Fluent、OpenFOAM和Xfoil对基于NACA0012翼型的单波纹板变弯度翼型和双波纹板变弯度翼型进行了气动特性分析，结果表明，基于有限体积法的Fluent、OpenFOAM计算气动力数值相近，准确度明显高于基于平板面元法的Xfoil[11]；Uliah等利用传统前缘和后缘襟翼减缓飞机阵风载荷，采用

收稿日期：2021-09-22；录用日期：2021-11-26；网络出版时间：2022-01-27 08:27
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220125.1726.001.html
基金项目：国家自然科学基金 (11672018)
*通信作者. E-mail：yutingdai@buaa.edu.cn

引用格式：尉濡恺，戴玉婷，杨超，等. 基于变弯度后缘的机翼阵风响应减缓数值研究[J]. 北京航空航天大学学报，2023，49(7)：1864-1874. WEI R K，DAI Y T，YANG C，et al. Numerical study of wing gust response alleviation based on camber morphing trailing edge[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(7)：1864-1874 (in Chinese).

欧拉方程和雷诺平均方程计算了传统前缘和后缘襟翼偏转气动力, 并在此基础上计算了通用飞机的阵风载荷[12]; 顾宁等基于动网格方法通过求解非定常 Navier-Stokes 方程计算副翼偏转时的阵风响应[13]; 许晓平等采用非定常欧拉方程对二维 NACA0006 翼型的阵风响应进行了计算分析, 对比结果较好[14]。

高精度的 CFD 计算能够准确进行阵风响应计算, 但其较长的计算周期较难应用于时域的主动控制。气动力降阶模型(reduced order model, ROM)则在处理此类问题时兼顾了精确性和计算时间。Zaide 和 Rareh 基于系统辨识技术建立了翼型在阵风激励下的气动力降阶模型, 并和 CFD 计算结果进行了对比[15]; 张伟伟等基于 CFD 技术通过系统辨识方法建立了机翼降阶非定常气动力模型, 耦合结构运动方程得到机翼阵风响应状态空间模型[16]; 杨国伟和王济康通过 CFD 计算结合自回归滑动平均(auto-regressive and moving average, ARMA)模型降阶模型准确预测了 NACA0012 翼型的阵风响应[17]; 聂雪媛和杨国伟基于 CFD 数据, 采用 ARMA 模型辨识了针对大柔性飞行器跨声速连续型阵风响应的降阶模型, 建立了有效的以翼根弯矩为减缓目标的阵风减缓控制系统[18]。

利用变弯度后缘进行阵风响应减缓控制, 阵风减缓控制律设计也是其中的重要环节, 现有的控制律研究较多, 主要包括线性的开环控制方法、最优控制理论方法、广义预测控制(generalized predictive control, GPC)理论等, 但其研究对象一般是传统的铰链舵面。杨阳等利用线性二次高斯控制方法对带有传统舵面的飞翼布局飞机设计了阵风减缓控制系统[19]; Dai 和 Yang 分别用标准的 GPC 理论和带有鲁棒性即考虑不确定性的 GPC 理论对开环阵风响应实验数据作闭环仿真, 减缓效果显著[20]。

现阶段, CFD 计算结合降阶模型主要用于机翼直接阵风响应计算与传统铰链舵面主动控制减缓, 缺乏应用于变弯度后缘主动阵风响应减缓的研究, 同时缺乏传统铰链舵面与变弯度后缘在动态气动力特性方面的对比探究。

本文较基于 CFD 计算结合状态观测器法辨识后缘偏转广义气动力模型, 建立了变弯度机翼开环阵风响应模型, 采用 GPC 理论设计闭环控制律, 在此基础上进行了变弯度机翼闭环阵风响应减缓仿真, 同时对比了传统铰链舵面与变弯度后缘动态气动力特性。

1　理论公式推导

拉氏域下典型的机翼气动弹性运动方程一般可写为[21]

$$\left[\boldsymbol{M}_{qq}s^2+\boldsymbol{C}_{qq}s+\boldsymbol{K}_{qq}-\frac{1}{2}\rho V^2\boldsymbol{Q}_{qq}(s)\right]\boldsymbol{q}(s)=\left[\frac{1}{2}\rho V^2\boldsymbol{Q}_{q\delta}(s)-\boldsymbol{M}_{q\delta}s^2\right]\delta(s)+\frac{1}{2}\rho V^2\boldsymbol{Q}_{\mathrm{g}}(s)\frac{\boldsymbol{w}_{\mathrm{g}}(s)}{V} \tag{1}$$

式中: s为拉氏变量; $\boldsymbol{q}(s)$为机翼广义模态坐标; $\delta(s)$为后缘偏转的广义坐标; V为飞行速度; ρ为大气密度; $\boldsymbol{w}_{\mathrm{g}}(s)$为阵风下洗速度向量; $\boldsymbol{M}_{qq}$和$\boldsymbol{M}_{q\delta}$为广义质量矩阵; $\boldsymbol{C}_{qq}$为广义阻尼矩阵; $\boldsymbol{K}_{qq}$为广义刚度矩阵; $\boldsymbol{Q}_{qq}(s)$、$\boldsymbol{Q}_{q\delta}(s)$和$\boldsymbol{Q}_{\mathrm{g}}(s)$为广义非定常气动力影响系数矩阵, 分别对应于机翼模态、后缘偏转、阵风产生的广义非定常气动力。其中, 机翼模态、阵风产生的广义非定常气动力采用面元法加有理函数拟合的方法计算得到, 由后缘偏转引起的广义非定常气动力通过 CFD 计算加状态观测器法降阶辨识的方法得到。

1.1　非定常气动力计算

拉氏域下的广义非定常气动力系数矩阵$\boldsymbol{Q}_{qq}(s)$和$\boldsymbol{Q}_{\mathrm{g}}(s)$由偶极子格网法计算得到, 采用最小状态(MS)法进行非定常气动力系数矩阵有理函数拟合到时域下, 拟合公式为

$$\boldsymbol{Q}(s)=\boldsymbol{A}_0+\frac{sb}{V}\boldsymbol{A}_1+\left(\frac{sb}{V}\right)^2\boldsymbol{A}_2+\boldsymbol{D}_{\mathrm{a}}\left(s\boldsymbol{I}-\frac{b}{V}\boldsymbol{R}\right)^{-1}\boldsymbol{E}s \tag{2}$$

式中: $\boldsymbol{A}_i$、$\boldsymbol{D}_{\mathrm{a}}$和$\boldsymbol{E}$为拟合多项式系数矩阵; b为参考半弦长; $\boldsymbol{R}=\mathrm{diag}(r_1,\cdots,r_{n_r})$为气动力滞后根矩阵, 文献 [21] 中给出了气动力滞后项x_{a}的具体公式。

后缘偏转的广义非定常气动力系数矩阵$\boldsymbol{Q}_{q\delta}(s)$转化到时域下可直接通过状态观测器法辨识得到的状态空间方程来替换, 离散的变弯度后缘偏转广义气动力时域状态空间模型可以表示为

$$\begin{cases}\bar{\boldsymbol{x}}_\delta(k+1)=\bar{\boldsymbol{A}}_\delta\bar{\boldsymbol{x}}_\delta(k)+\bar{\boldsymbol{B}}_\delta\bar{\delta}(k)\\ \bar{\boldsymbol{y}}_\delta(k)=\bar{\boldsymbol{C}}_\delta\bar{\boldsymbol{x}}_\delta(k)+\bar{\boldsymbol{D}}_\delta\bar{\delta}(k)\end{cases} \tag{3}$$

式中: $\bar{\boldsymbol{x}}_\delta$、$\bar{\boldsymbol{y}}_\delta$和$\bar{\delta}$分别为离散系统的状态矢量、广义气动力输出矢量和变弯度后缘偏转位移输入; $\bar{\boldsymbol{A}}_\delta$、$\bar{\boldsymbol{B}}_\delta$、$\bar{\boldsymbol{C}}_\delta$、$\bar{\boldsymbol{D}}_\delta$为离散系统状态空间方程常数矩阵。上述系统在零初始条件下, 对时间序列离散写成矩阵形式, 系统的 Markov 参数可表示为

$$\boldsymbol{Y}_\delta=\left[\bar{\boldsymbol{D}}_\delta\ \ \bar{\boldsymbol{C}}_\delta\bar{\boldsymbol{B}}_\delta\ \ \cdots\ \ \bar{\boldsymbol{C}}_\delta\bar{\boldsymbol{A}}_\delta^{l-2}\bar{\boldsymbol{B}}_\delta\right] \tag{4}$$

式中: l为数据采样个数。状态观测器法辨识的最终目的是通过辨识系统的 Markov 参数来产生系统常数矩阵$\bar{\boldsymbol{A}}_\delta$、$\bar{\boldsymbol{B}}_\delta$、$\bar{\boldsymbol{C}}_\delta$、$\bar{\boldsymbol{D}}_\delta$, 但由于原系统阶数往往很高, 需通过引入观测器, 利用带观测器系统的 Markov 参数来恢复原系统的 Markov 参数。

从系统 Markov 参数出发, 采用特征实现算法(eigensystem realization algorithm, ERA)可以得到系

统常数矩阵，通过之前求得的Markov参数，可构造广义Hankel矩阵：

$$\boldsymbol{H}(k-1)=\begin{bmatrix}\boldsymbol{Y}_\delta(k) & \cdots & \boldsymbol{Y}_\delta(k+\beta-1)\\ \vdots & & \vdots\\ \boldsymbol{Y}_\delta(k+\alpha-1) & \cdots & \boldsymbol{Y}_\delta(k+\alpha+\beta-2)\end{bmatrix} \tag{5}$$

式中：α为采样窗口移动的时间步长数；β为数据采样窗口大小。对$\boldsymbol{H}(0)$进行奇异值分解即可得到所需要的常数矩阵$\bar{\boldsymbol{A}}_\delta$、$\bar{\boldsymbol{B}}_\delta$、$\bar{\boldsymbol{C}}_\delta$、$\bar{\boldsymbol{D}}_\delta$[22]。

离散系统与连续系统常态系数矩阵对应转换关系如下：

$$\bar{\boldsymbol{A}}_\delta=\boldsymbol{A}_\delta\Delta t+\boldsymbol{I} \tag{6}$$

$$\bar{\boldsymbol{B}}_\delta=\boldsymbol{B}_\delta\Delta t \tag{7}$$

$$\bar{\boldsymbol{C}}_\delta=\boldsymbol{C}_\delta \tag{8}$$

$$\bar{\boldsymbol{D}}_\delta=\boldsymbol{D}_\delta \tag{9}$$

式中：Δt为离散系统时间步长；$\boldsymbol{I}$为单位矩阵，$\boldsymbol{A}_\delta$、$\boldsymbol{B}_\delta$、$\boldsymbol{C}_\delta$、$\boldsymbol{D}_\delta$为连续系统状态空间方程常系数矩阵。

对辨识得到的式(1)转换可得连续的变弯度后缘偏转广义气动力时域状态空间模型如下

$$\begin{cases}\dot{\boldsymbol{x}}_\delta=\boldsymbol{A}_\delta\boldsymbol{x}_\delta+\boldsymbol{B}_\delta\delta\\ \boldsymbol{y}_\delta=\boldsymbol{C}_\delta\boldsymbol{x}_\delta+\boldsymbol{D}_\delta\delta\end{cases} \tag{10}$$

1.2 开环阵风响应建模

将上述得到的气动力模型代入式(1)，转换到时域下可得

$$\begin{aligned}&\boldsymbol{M}_{qq}\ddot{\boldsymbol{q}}+\boldsymbol{M}_{q\delta}\ddot{\delta}+\boldsymbol{C}_{qq}\dot{\boldsymbol{q}}+\boldsymbol{K}_{qq}\boldsymbol{q}=\\ &\quad\frac{1}{2}\rho V^2\left[\boldsymbol{Q}_{qq0}\boldsymbol{q}+\frac{b}{V}\boldsymbol{Q}_{qq1}\dot{\boldsymbol{q}}+\left(\frac{b}{V}\right)^2\boldsymbol{Q}_{qq2}\ddot{\boldsymbol{q}}\right]+\\ &\quad\boldsymbol{D}_\delta\delta+\boldsymbol{C}_\delta\boldsymbol{x}_\delta+\frac{1}{2}\rho V^2\boldsymbol{D}_{\mathrm{a}}x_{\mathrm{a}}+\\ &\quad\frac{1}{2}\rho V^2\left[\boldsymbol{Q}_{\mathrm{g}0}\frac{\boldsymbol{w}_{\mathrm{g}}}{V}+\frac{b}{V}\boldsymbol{Q}_{\mathrm{g}1}\frac{\dot{\boldsymbol{w}}_{\mathrm{g}}}{V}+\left(\frac{b}{V}\right)^2\boldsymbol{Q}_{\mathrm{g}2}\frac{\ddot{\boldsymbol{w}}_{\mathrm{g}}}{V}\right]\end{aligned} \tag{11}$$

一般对于阵风，可令$\boldsymbol{Q}_{\mathrm{g}2}=0$，将式(11)整理成状态空间形式：

$$\begin{aligned}\begin{bmatrix}\dot{\boldsymbol{q}}\\ \ddot{\boldsymbol{q}}\\ \dot{\boldsymbol{x}}_{\mathrm{a}}\\ \dot{\boldsymbol{x}}_\delta\end{bmatrix}=&\begin{bmatrix}0 & \boldsymbol{I} & 0 & 0\\ \bar{\boldsymbol{M}}_q^{-1}\bar{\boldsymbol{K}}_q & \bar{\boldsymbol{M}}_q^{-1}\bar{\boldsymbol{C}}_q & \frac{1}{2}\rho V^2\bar{\boldsymbol{M}}_q^{-1}\boldsymbol{D}_{\mathrm{a}} & \bar{\boldsymbol{M}}_q^{-1}\boldsymbol{C}_\delta\\ 0 & \boldsymbol{E}_q & \boldsymbol{R}V/b & 0\\ 0 & 0 & 0 & 0\end{bmatrix}\cdot\\ &\begin{bmatrix}\boldsymbol{q}\\ \dot{\boldsymbol{q}}\\ \boldsymbol{x}_{\mathrm{a}}\\ \boldsymbol{x}_\delta\end{bmatrix}+\begin{bmatrix}0\\ \boldsymbol{D}_\delta\\ 0\\ \boldsymbol{B}_\delta\end{bmatrix}\delta+\\ &\begin{bmatrix}0 & 0\\ \frac{1}{2}\rho V\bar{\boldsymbol{M}}_q^{-1}\boldsymbol{Q}_{\mathrm{g}0} & \frac{1}{2}\rho b\bar{\boldsymbol{M}}_q^{-1}\boldsymbol{Q}_{\mathrm{g}1}\\ 0 & \boldsymbol{E}_q/V\\ 0 & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{w}_{\mathrm{g}}\\ \dot{\boldsymbol{w}}_{\mathrm{g}}\end{bmatrix}\end{aligned} \tag{12}$$

式中：

$$\bar{\boldsymbol{M}}_q=\boldsymbol{M}_{qq}-\frac{1}{2}\rho V^2\left(\frac{b}{V}\right)^2\boldsymbol{Q}_{qq2} \tag{13}$$

$$\bar{\boldsymbol{C}}_q=\frac{1}{2}\rho Vb\boldsymbol{Q}_{qq1}-\boldsymbol{C}_{qq} \tag{14}$$

$$\bar{\boldsymbol{K}}_q=\frac{1}{2}\rho V^2\boldsymbol{Q}_{qq0}-\boldsymbol{K}_{qq} \tag{15}$$

考虑结构输出，放置机翼翼尖处的传感器测量加速度响应为

$$\ddot{\boldsymbol{u}}=\boldsymbol{\Phi}^{\mathrm{T}}\ddot{\boldsymbol{q}} \tag{16}$$

式中：$\boldsymbol{\Phi}$为翼尖处的模态矩阵。

进一步整理可得变弯度机翼开环阵风响应状态空间方程为

$$\begin{cases}\dot{\boldsymbol{X}}_{\mathrm{ae}}=\boldsymbol{A}_{\mathrm{ae}}\boldsymbol{X}_{\mathrm{ae}}+\boldsymbol{B}_{\mathrm{ae}}\boldsymbol{u}_{\mathrm{ae}}+\boldsymbol{E}_{\mathrm{ae}}\boldsymbol{W}_{\mathrm{ae}}\\ \boldsymbol{Y}_{\mathrm{ae}}=\boldsymbol{C}_{\mathrm{ae}}\boldsymbol{X}_{\mathrm{ae}}+\boldsymbol{D}_{\mathrm{ae}}\boldsymbol{u}_{\mathrm{ae}}+\boldsymbol{F}_{\mathrm{ae}}\boldsymbol{W}_{\mathrm{ae}}\end{cases} \tag{17}$$

式中：

$$\boldsymbol{X}_{\mathrm{ae}}=[\,\boldsymbol{q}\quad\dot{\boldsymbol{q}}\quad\boldsymbol{x}_{\mathrm{a}}\quad\boldsymbol{x}_\delta\,]^{\mathrm{T}} \tag{18}$$

$$u_{\mathrm{ae}}=\delta \tag{19}$$

$$\boldsymbol{W}_{\mathrm{ae}}=\begin{bmatrix}\boldsymbol{w}_{\mathrm{g}}\quad\dot{\boldsymbol{w}}_{\mathrm{g}}\end{bmatrix}^{\mathrm{T}} \tag{20}$$

$$\boldsymbol{Y}_{\mathrm{ae}}=\ddot{\boldsymbol{u}} \tag{21}$$

$$\boldsymbol{A}_{\mathrm{ae}}=\begin{bmatrix}0 & \boldsymbol{I} & 0 & 0\\ \bar{\boldsymbol{M}}_q^{-1}\bar{\boldsymbol{K}}_q & \bar{\boldsymbol{M}}_q^{-1}\bar{\boldsymbol{C}}_q & \frac{1}{2}\rho V^2\bar{\boldsymbol{M}}_q^{-1}\boldsymbol{D}_{\mathrm{a}} & \bar{\boldsymbol{M}}_q^{-1}\boldsymbol{C}_\delta\\ 0 & \boldsymbol{E}_q & \boldsymbol{R}V/b & 0\\ 0 & 0 & 0 & 0\end{bmatrix} \tag{22}$$

$$\boldsymbol{B}_{\mathrm{ae}}=[\,0\quad\boldsymbol{D}_\delta\quad0\quad\boldsymbol{B}_\delta\,]^{\mathrm{T}} \tag{23}$$

$$\boldsymbol{E}_{\mathrm{ae}}=\begin{bmatrix}0 & 0\\ \frac{1}{2}\rho V\bar{\boldsymbol{M}}_q^{-1}\boldsymbol{Q}_{\mathrm{g}0} & \frac{1}{2}\rho b\bar{\boldsymbol{M}}_q^{-1}\boldsymbol{Q}_{\mathrm{g}1}\\ 0 & \boldsymbol{E}_q/V\\ 0 & 0\end{bmatrix} \tag{24}$$

$$\boldsymbol{C}_{\mathrm{ae}}=\begin{bmatrix}\boldsymbol{\Phi}^{\mathrm{T}}\bar{\boldsymbol{M}}_q^{-1}\bar{\boldsymbol{K}}_q & \boldsymbol{\Phi}^{\mathrm{T}}\bar{\boldsymbol{M}}_q^{-1}\bar{\boldsymbol{C}}_q & \boldsymbol{\Phi}^{\mathrm{T}}\frac{1}{2}\rho V^2\bar{\boldsymbol{M}}_q^{-1}\boldsymbol{D}_{\mathrm{a}} & \boldsymbol{\Phi}^{\mathrm{T}}\bar{\boldsymbol{M}}_q^{-1}\boldsymbol{C}_\delta\end{bmatrix} \tag{25}$$

$$\boldsymbol{D}_{\mathrm{ae}}=\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{D}_\delta \tag{26}$$

$$\boldsymbol{F}_{\mathrm{ae}}=\begin{bmatrix}\boldsymbol{\Phi}^{\mathrm{T}}\frac{1}{2}\rho V\bar{\boldsymbol{M}}_q^{-1}\boldsymbol{Q}_{\mathrm{g}0} & \boldsymbol{\Phi}^{\mathrm{T}}\frac{1}{2}\rho b\bar{\boldsymbol{M}}_q^{-1}\boldsymbol{Q}_{\mathrm{g}1}\end{bmatrix} \tag{27}$$

1.3 闭环控制律设计

利用GPC理论设计变弯度机翼闭环阵风减缓控制律，需要先依据特定舵偏输入得到的变弯度机翼开环响应输出数据辨识系统输入输出自回归模型。

在时间步长k时刻对系统参数辨识完成后，控制系统在第$k+1$个时间步长启动，给变弯度后缘舵

偏输入。在闭环系统中, 变弯度后缘舵偏输入记为 $\boldsymbol{u}_{\mathrm{ae}}$, 第$k+j$个时间步的响应还是可以表示为 3 个部分的线性组合形式: 第 1 部分是过去p个时间步的响应$\boldsymbol{Y}_{\mathrm{ae}}(k)$; 第 2 部分是未来$j$个时间步、当下的(未知)和过去$p$个时间步的变弯度后缘舵偏输入$\boldsymbol{u}_{\mathrm{ae}}$; 第 3 部分是未来$j$个时间步、当下的(未知)和过去$p$个时间步的阵风扰动$\boldsymbol{W}_{\mathrm{ae}}(k)$。未来响应的回归关系可表示为

$$\begin{aligned}\boldsymbol{Y}_{\mathrm{ae}}(k+j)=&\boldsymbol{\alpha}_1^j\boldsymbol{Y}_{\mathrm{ae}}(k-1)+\boldsymbol{\alpha}_2^j\boldsymbol{Y}_{\mathrm{ae}}(k-2)+\cdots+\\&\boldsymbol{\alpha}_p^j\boldsymbol{Y}_{\mathrm{ae}}(k-p)+\boldsymbol{\beta}_0'\boldsymbol{u}_{\mathrm{ae}}(k+j)+\boldsymbol{\beta}_0'^1\boldsymbol{u}_{\mathrm{ae}}(k+j-1)+\cdots+\\&\boldsymbol{\beta}_0'^j\boldsymbol{u}_{\mathrm{ae}}(k)+\boldsymbol{\beta}_1'^j\boldsymbol{u}_{\mathrm{ae}}(k-1)+\boldsymbol{\beta}_2'^j\boldsymbol{u}_{\mathrm{ae}}(k-2)+\cdots+\\&\boldsymbol{\beta}_p'^j\boldsymbol{u}_{\mathrm{ae}}(k-p)+\boldsymbol{\gamma}_0'\boldsymbol{W}_{\mathrm{ae}}(k+j)+\boldsymbol{\gamma}_0'^1\boldsymbol{W}_{\mathrm{ae}}(k+j-1)+\cdots+\\&\boldsymbol{\gamma}_0'^j\boldsymbol{W}_{\mathrm{ae}}(k)+\boldsymbol{\gamma}_1'^j\boldsymbol{W}_{\mathrm{ae}}(k-1)+\boldsymbol{\gamma}_2'^j\boldsymbol{W}_{\mathrm{ae}}(k-2)+\cdots+\\&\boldsymbol{\gamma}_p'^j\boldsymbol{W}_{\mathrm{ae}}(k-p)\end{aligned}\tag{28}$$

式中: α、β、γ为对应的线性组合系数。

GPC 方法是基于有限步长的系统输出的, 该有限步长称为预测步长h_{p}, 而控制输入也是在一定限度内, 称为控制步长h_{c}, 控制步长之外的控制输入假设为 0。将式 (28) 中的j分别取$1,2,\cdots,h_{\mathrm{p}}-1$, 建立多步输出预测方程:

$$\begin{aligned}\boldsymbol{Y}_{\mathrm{ae_}h_{\mathrm{p}}}(k)=&\boldsymbol{T}\boldsymbol{u}_{\mathrm{ae_}h_{\mathrm{c}}}(k)+\boldsymbol{B}_w\boldsymbol{u}_{\mathrm{ae_}p}(k-p)+\boldsymbol{A}_w\boldsymbol{Y}_{\mathrm{ae_}p}(k-p)+\\&\boldsymbol{D}_p\boldsymbol{W}_{\mathrm{ae_}p}(k-p)+\boldsymbol{D}_{h_{\mathrm{p}}}\boldsymbol{W}_{\mathrm{ae_}h_{\mathrm{p}}}(k)\end{aligned}\tag{29}$$

式中: 系数矩阵$\boldsymbol{T}$、$\boldsymbol{B}_w$、$\boldsymbol{A}_w$、$\boldsymbol{D}_p$、$\boldsymbol{D}_{h_{\mathrm{p}}}$由系统 Markov 参数组合而成; $\boldsymbol{Y}_{\mathrm{ae_}h_{\mathrm{p}}}(k)$为未来$h_{\mathrm{p}}$个时间步长内的机翼开环阵风响应系统输出; $\boldsymbol{u}_{\mathrm{ae_}h_{\mathrm{c}}}(k)$为未来$h_{\mathrm{c}}$个时间步长内的变弯度后缘偏转输入; $\boldsymbol{Y}_{\mathrm{ae_}p}(k-p)$、$\boldsymbol{W}_{\mathrm{ae_}p}(k-p)$和$\boldsymbol{u}_{\mathrm{ae_}p}(k-p)$分别为包含过去$p$个时间步长的机翼开环阵风响应系统输出、阵风输入和变弯度后缘偏转输入; $\boldsymbol{W}_{\mathrm{ae_}h_{\mathrm{p}}}(k)$为未来$h_{\mathrm{p}}$个时间步长内的阵风输入。

GPC 方法旨在减小未来h_{p}个时间步长内的翼尖加速度, 未来h_{c}个时间步长中优化控制指令的目标可以写为

$$\min\left(\boldsymbol{Y}_{\mathrm{ae_}h_{\mathrm{p}}}^{\mathrm{T}}(k)\boldsymbol{R}_w\boldsymbol{Y}_{\mathrm{ae_}h_{\mathrm{p}}}(k)+\boldsymbol{u}_{\mathrm{ae_}h_{\mathrm{c}}}^{\mathrm{T}}(k)\boldsymbol{Q}_w\boldsymbol{u}_{\mathrm{ae_}h_{\mathrm{c}}}(k)\right)\tag{30}$$

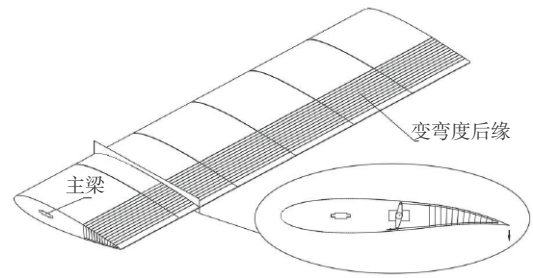
式(30)是典型的最优控制问题。方程表示需要找到一个控制命令, 使得控制能量相对较小且在未来h_{c}个时间步长中响应相对较小, 二者大小可由权重矩阵$\boldsymbol{R}_w$和$\boldsymbol{Q}_w$决定。通过求解式(30), 得到在时间步长$k+1$时变弯度后缘偏转输入为

$$\begin{aligned}\boldsymbol{u}_{\mathrm{ae}}(k+1)=\left\{-\left(\boldsymbol{T}^{\mathrm{T}}\boldsymbol{R}_w\boldsymbol{T}+\boldsymbol{Q}_w\right)\boldsymbol{T}^{\mathrm{T}}\boldsymbol{R}_w\left(\boldsymbol{B}_w\boldsymbol{u}_{\mathrm{ae_}p}(k-p)+\right.\right.\\\left.\left.\boldsymbol{A}_w\boldsymbol{Y}_{\mathrm{ae_}p}(k-p)+\boldsymbol{D}_p\boldsymbol{W}_{\mathrm{ae_}p}(k-p)+\boldsymbol{D}_{h_{\mathrm{p}}}\boldsymbol{W}_{\mathrm{ae_}h_{\mathrm{p}}}(k)\right)\right\}\end{aligned}\tag{31}$$

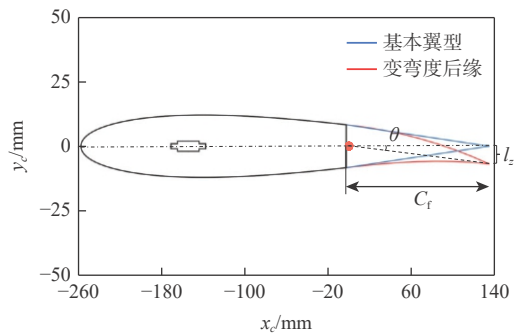
2　仿真算例

2.1　计算对象

计算模型为一个展长 1.2 m、弦长$c=0.4$ m的矩形机翼, 剖面采用 NACA0015 翼型, 共划分为 6 个翼段, 分别连接在主梁上, 每个翼段包含一个变弯度后缘, 后缘固定连接在机翼 65% 弦长处, 主梁根部采用固支边界, 机翼几何模型如图 1 所示, 图中还给出了机翼截面变形示意图及变弯度后缘偏转角定义



(a) 变弯度机翼示意图



(b) 变弯度后缘偏转角

图 1　变弯度机翼几何模型

Fig. 1　Camber morphing wing geometry model

图 1 中: x_c代表后缘弦向坐标, y_c为后缘运动后垂直翼型中线方向坐标, 以后缘起始点为原点, 定义后缘向下偏转为正。在计算变弯度后缘偏角时, 只考虑垂直翼型中线方向位移, 不考虑沿弦向变形, 偏转角定义如下:

$$\theta=\arctan\left(\frac{l_z}{C_{\mathrm{f}}}\right)\tag{32}$$

式中: θ为变弯度后缘偏转角; l_z为变弯度后缘垂直翼面方向位移变形; C_{f}为后缘变形区域长度。

2.2　开环阵风响应计算

用有限元方法对机翼进行模态分析, 保留前 4 阶模态, 采用模态叠加法近似表示结构动力学特性, 图 2 和表 1 分别展示了机翼前 4 阶模态振型及对应的模态频率。

后缘偏转引起的广义气动力采用状态观测器法参数辨识得到, 辨识用的数据需要通过 CFD 方法计算, 利用用户自定义函数(udf)给定后缘运动计算得到, 给定运动形式如下:

$$y_c(x_c,t)=(0.5x_c^2+0.033\,5x_c)\left(\sum_{n=1}^{4}a_n\sin(2\pi f_nt)\right)\tag{33}$$

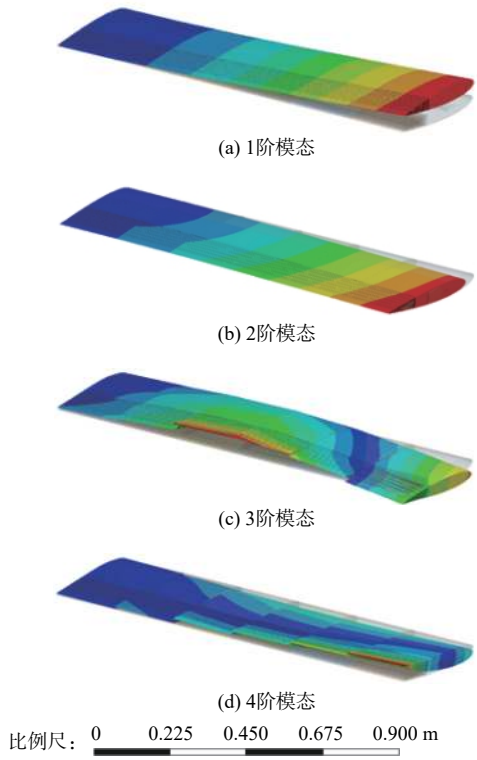


图2 前4阶模态振型

Fig. 2 The first four modes shape

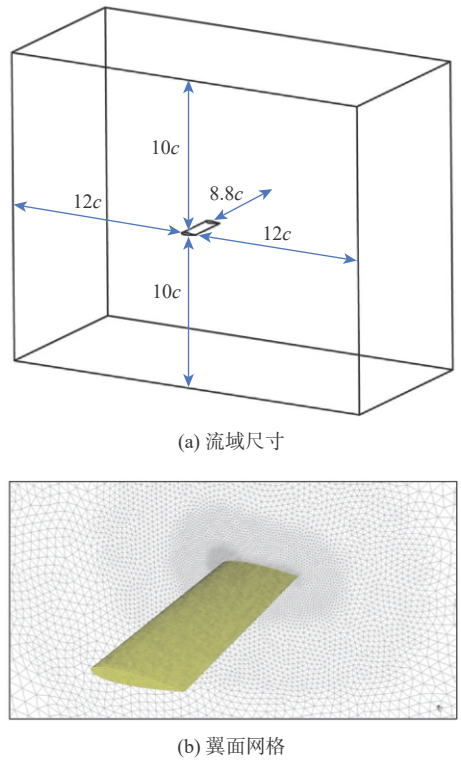


图3 流域尺寸及翼面网格示意图

Fig. 3 Schematic diagram of watershed size and airfoil mesh

表1 前4阶模态频率

Table 1 Frequency of the first four modes

模态阶数	模态描述	模态频率/Hz
1	机翼面外一弯	2.77
2	机翼面内一弯	13.45
3	机翼面外二弯	16.31
4	机翼一扭	20.25

表2 网格无关性计算

Table 2 Grid-independent computing

网格名称	网格数量	升力系数
Coarse	420万	0.091 8
Medium	538万	0.083 2
Fine	631万	0.083 4

式中：f_n为表1所给的前4阶模态频率；a_n = (0.4, 0.8, 1.2, 1.8)为对应模态的给定幅值。

计算流场网格采用非结构网格，网格总数538万，流域尺寸与翼面网格如图3所示。

流体的控制方程采用不可压缩的三维雷诺平均Navier-Stokes方程(RANS)，湍流模型采用Spalart-Allmaras(S-A)。流场的入口条件为速度入口，出口条件设置压力出口边界，机翼表面为无滑移壁面边界，计算时间步长为10^{-4} s。为排除网格对计算结果的影响，在来流速度20 m/s、攻角7°工况下对流场网格做网格无关性验证，取3套网格计算升力系数，结果如表2所示。

综合考虑计算精度和计算时间，最终选用中等数量的流场网格。

利用用户自定义函数加载式(33)给定后缘运动，在来流速度20 m/s、攻角0°工况下计算后缘变形产生的前4阶模态广义气动力。进一步利用状态观测器法辨识CFD计算的模态气动力，得到后缘偏转广义气动力模型。

通过状态观测器法辨识系统状态空间方程，适用于线性及弱非线性系统[22]。为了验证采用状态观测器法辨识的合理性，对后缘偏转流场线性度进行研究。

考虑到阵风减缓时后缘运动幅值及频率大小，给定后缘单正弦运动，运动函数如下：

$$y_c(x_c,t)=(0.5x_c^2+0.0335x_c)(A_m\sin(4\pi t)) \tag{34}$$

正弦运动频率为2 Hz，正弦运动幅值A_m为0.4、0.8、1.2、1.6，对应最大后缘偏转角分别为2.4°、4.7°、7°、9.4°，采用CFD方法计算来流速度20 m/s。攻角0°情况下变弯度机翼的升力系数C_l，后缘最大偏转角2.4°、9.4°下变弯度机翼升力系数C_l及其滞回曲线、后缘最大偏转角9.4°时刻流场压力分布云图与流线图及后缘偏转角随升力系数幅值变化关系如图4所示。

由图 4(b)、图 4(d)结果分析可知, 在后缘最大偏转角 2.4°和 9.4°下, 机翼升力系数与后缘偏转位移之间线性度关系较好, 且根据图 4(e)流场压力分布与流线图可知, 后缘最大偏转角 9.4°时刻后缘流动尚未产生分离, 由图 4(f)结果分析可知, 后缘偏转角与机翼升力系数幅值之间线性度关系较好, 故通过状态观测器法辨识后缘偏转广义气动力模型较为合理。

图 5 为状态观测器法辨识变弯度后缘偏转对应的前 4 阶模态气动力辨识结果。

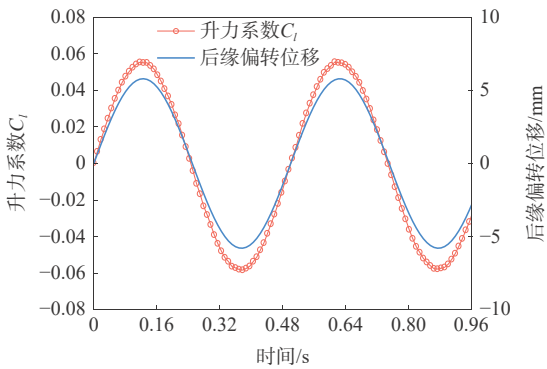
定义误差公式如下:

$$e=\sqrt{\frac{\sum(\boldsymbol{y}_1-\boldsymbol{y}_2)^2}{\sum\boldsymbol{y}_1^2}} \tag{35}$$

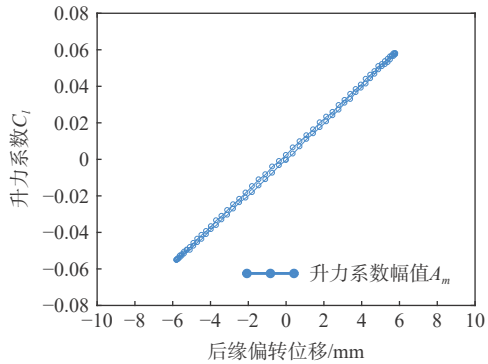
式中: $\boldsymbol{y}_1$和$\boldsymbol{y}_2$分别为比较的数据矢量和待比较的数据矢量。e为$\boldsymbol{y}_2$相比$\boldsymbol{y}_1$的数值误差, 误差越小, 则表示 2 组数据越接近, 反之远离。由此计算图 5 中的辨识效果, 辨识模型与 CFD 计算结果相比数值平均误差为 0.07, 辨识效果较好。

为进一步验证辨识得到的气动力模型的准确性, 对后缘正弦运动偏转广义气动力进行预测。考虑到阵风减缓时后缘运动幅值及频率大小, 给定后缘单正弦运动, 运动函数采用与式(34)相同的形式, 正弦运动频率 2 Hz, 正弦运动幅值A_m为 0.8、1.6, 对应后缘最大偏转角 4.7°、9.4°, 预测结果与 CFD 计算结果对比如图 6 所示。

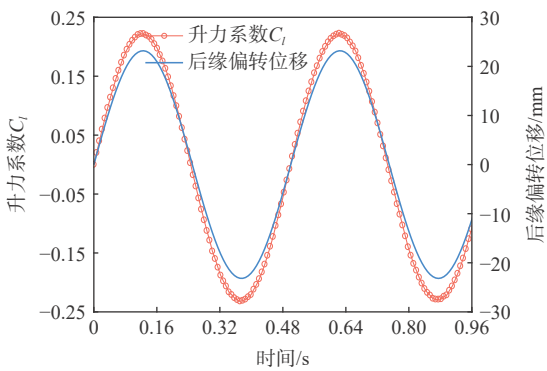
预测结果表明, 根据式(35), 在给定后缘运动幅值范围内, 后缘最大偏转角 9.4°下预测的变弯度



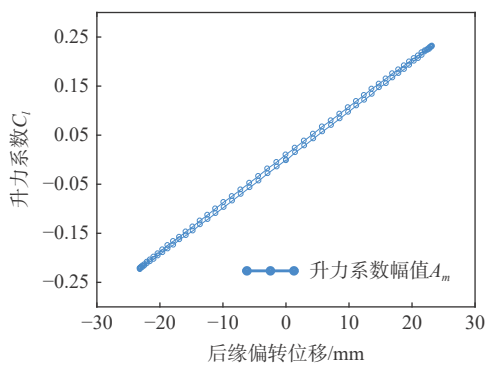
(a) 后缘最大偏转角2.4°机翼升力系数



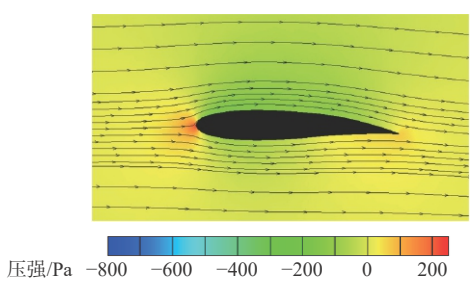
(b) 后缘最大偏转角2.4°升力系数滞回曲线



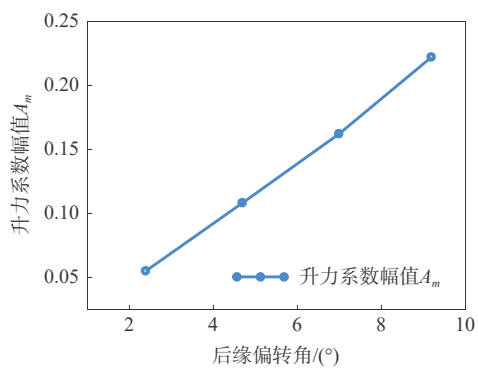
(c) 后缘最大偏转角9.4°机翼升力系数



(d) 后缘最大偏转角9.4°升力系数滞回曲线



(e) 后缘最大偏转角9.4°时刻流场压力分布



(f) 后缘偏转角随升力系数幅值变化

图 4　流场特性分析

Fig. 4　Flow field characteristic analysis

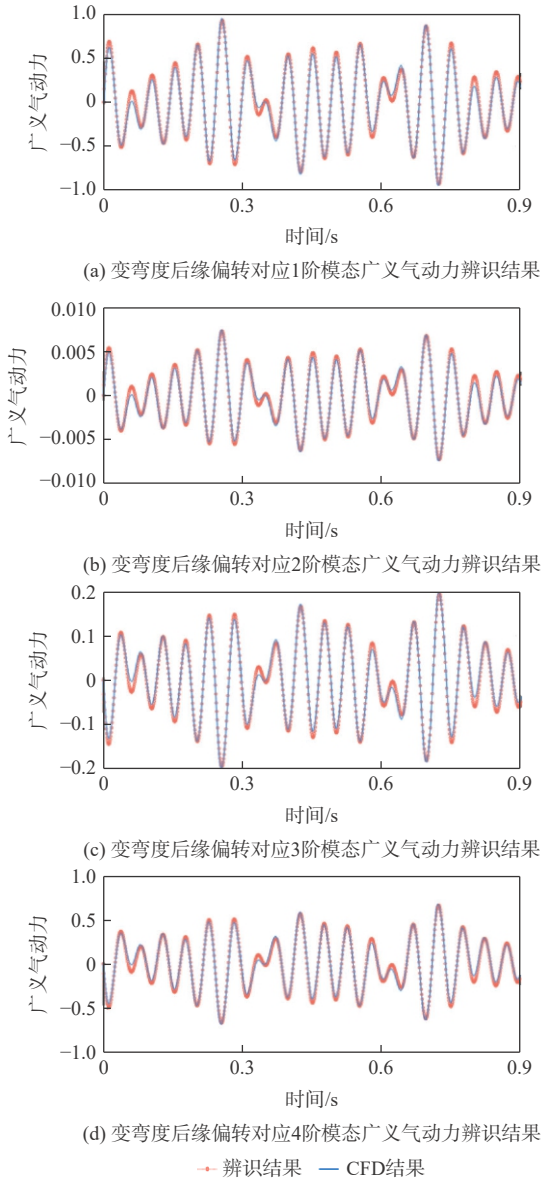


图5 变弯度后缘偏转广义模态气动力辨识结果

Fig. 5 Identification result of generalized modal aerodynamic force of trailing edge deflection

后缘偏转广义模态气动力与CFD计算结果相比数值平均误差为0.098,后缘最大偏转角4.7°下数值平均误差为0.053,满足工程实际使用要求,所辨识的后缘偏转广义气动力模型较为准确,可以用于搭建机翼开环阵风响应模型。

根据式(11),将得到的气动力与结构动力学部分耦合一起得到机翼开环阵风响应模型。

2.3 闭环阵风响应减缓仿真

对机翼进行闭环阵风响应减缓仿真,需先基于开环阵风响应结果进行控制律设计。

给开环阵风响应模型输入阵风强度4 m/s、阵风频率2 Hz的正弦阵风,并给舵面输入均值为0、方差为0.03的偏转白噪声信号,得到开环加速度响应。通过自回归模型对开环加速度响应数据进行参数辨识,从而建立多步输出预测方程,对式(30)进行最优控制求解闭环控制律,可以得到机翼闭环阵风响应,仿真结果如图7所示。

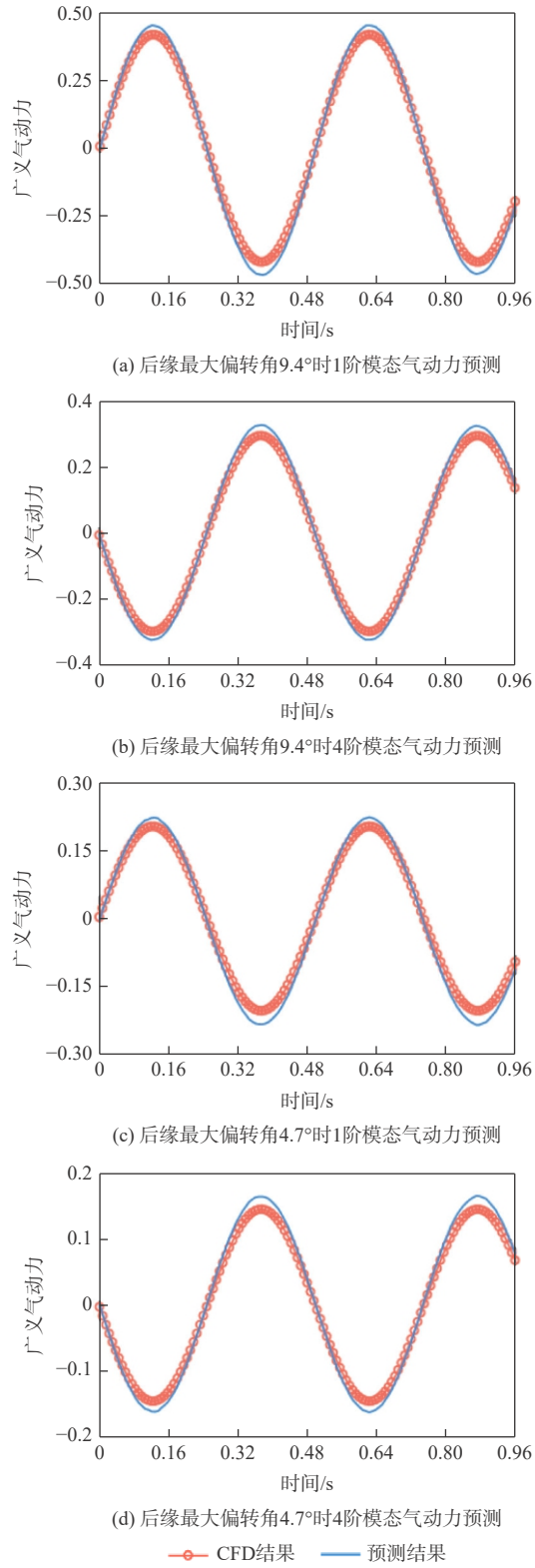
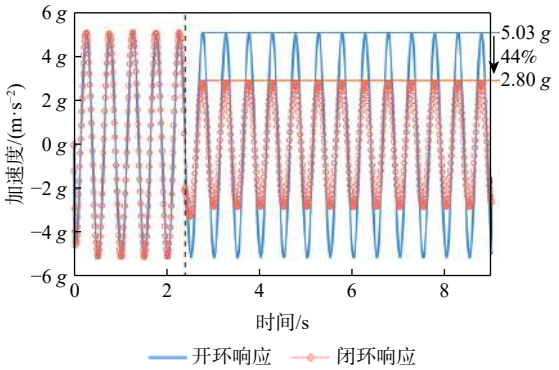


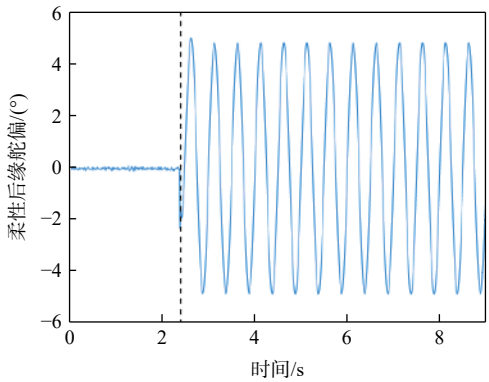
图6 不同后缘运动幅值气动力预测

Fig. 6 Aerodynamic prediction of different trailing edge motion amplitudes

由闭环仿真结果可知,前2.39 s对开环阵风响应数据进行辨识并计算得到控制律参数,辨识效果



(a) GPC开闭环仿真结果



(b) 柔性后缘舵偏

图 7　闭环阵风响应减缓仿真结果

Fig. 7　Simulation results of closed-loop gust response alleviation

较好，2.39 s 后控制器打开，对变弯度机翼加速度进行减缓主动控制，开环翼尖加速度为 5.03 g，闭环减缓后的翼尖加速度为 2.80 g，减缓效率为 44.33%，验证了针对变弯度后缘利用 GPC 理论进行阵风减缓控制设计的有效性。

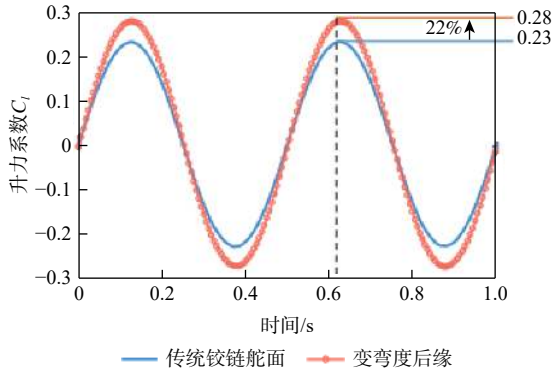
2.4　铰链舵面动态气动力与减缓效果分析

采用 CFD 方法对变弯度后缘和传统铰链舵面气动力区别进行探究，利用 udf 给定后缘刚性偏转运动计算得到，给定运动形式如下：

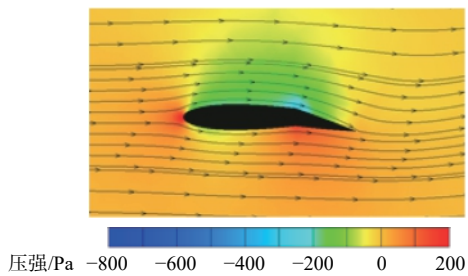
$$y_c(x_c,t)=0.103\,5x_c(2.2\sin(4\pi t)) \tag{36}$$

传统铰链舵面与变弯度后缘最大运动偏转位移相同，最大后缘偏转角 13°，式(36)中符号及所给的正弦运动频率幅值与上文相同。图 8 给出变弯度后缘与铰链舵面升力系数对比、对应最大偏转时刻二者的流场压力分布云图及沿翼型弦向压力分布。

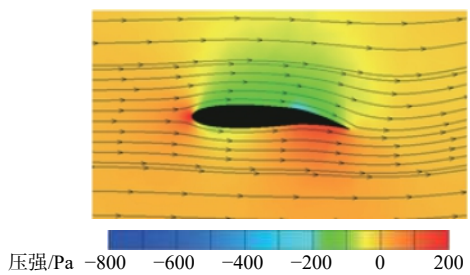
由图 8(a)结果分析可知，在相同后缘最大偏转下，变弯度后缘升力系数相比铰链舵面提高 22%，在机翼整体动态气动力方面变弯度后缘比传统铰链舵面提升效果更好。且由图 8(c)结果分析可知，相比传统铰链舵面，变弯度后缘偏转连续、无曲率突变，在沿弦向压力系数分布上无凸起，压力分布更为合理，故采用变弯度后缘进行阵风减缓控制相比统铰链舵面具有更大潜力。



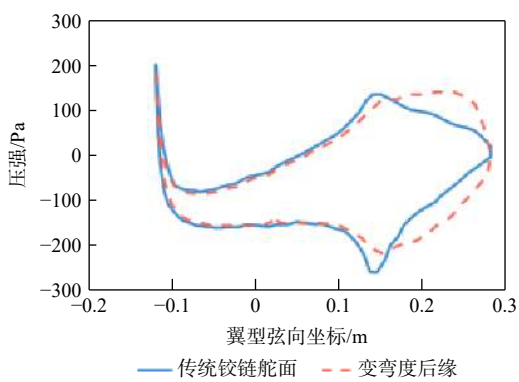
(a) 传统铰链舵面与变弯度后缘升力系数对比



(b) 传统铰链舵面偏转最大时刻流场压力分布



(c) 变弯度后缘偏转最大时刻流场压力分布



(d) 偏转最大时刻传统铰链舵面与变弯度后缘沿弦向压力分布

图 8　铰链舵面与变弯度后缘动态气动力对比

Fig. 8　Comparison of aerodynamic forces between hinged flap and camber morphing trailing edge

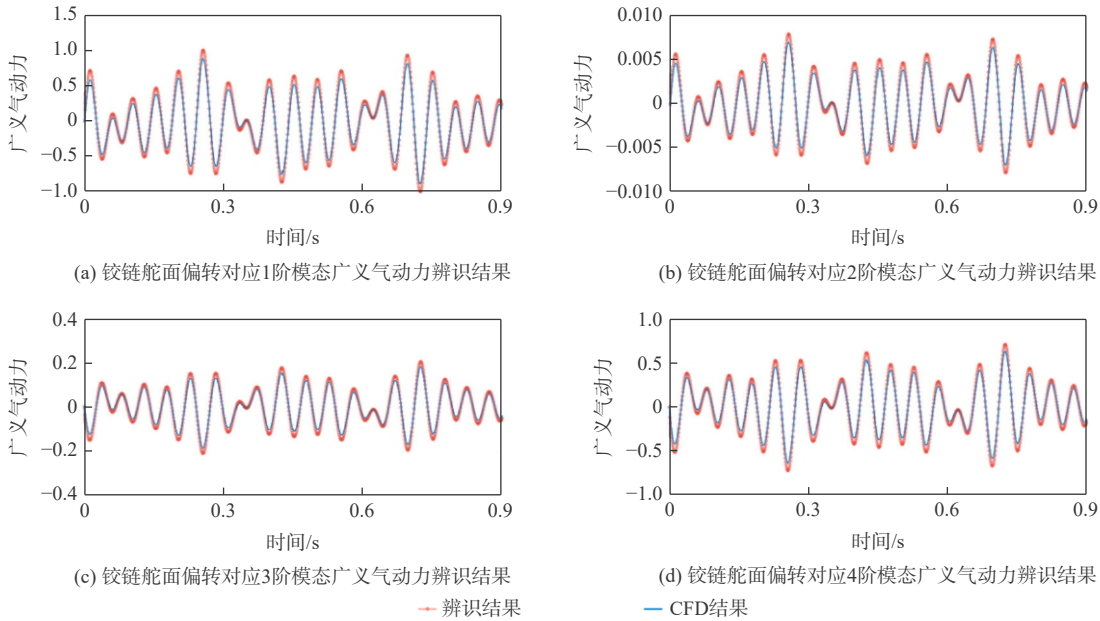
进一步对铰链舵面的阵风响应减缓效果进行探究，采用与变弯度后缘相同的方法建立传统铰链舵面开环阵风响应模型，铰链舵面给定运动形式如下：

$$y_c(x_c,t)=0.103\,5x_c\sum_{n=1}^{4}a_n\sin(2\pi f_n t) \tag{37}$$

其中，铰链舵面最大运动偏转位移与变弯度后缘相同，给出状态观测器法辨识铰链舵面偏转对应

的前4阶模态气动力辨识结果，如图9所示。

对传统铰链舵面开环阵风响应模型做GPC减缓控制仿真，输入与变弯度后缘相同的正弦阵风信号和舵面偏转白噪声，采用相同的控制参数对铰链舵面开环阵风响应数据进行参数辨识和最优控制求解，得到开闭环仿真结果如图10所示。



(a) 铰链舵面偏转对应1阶模态广义气动力辨识结果

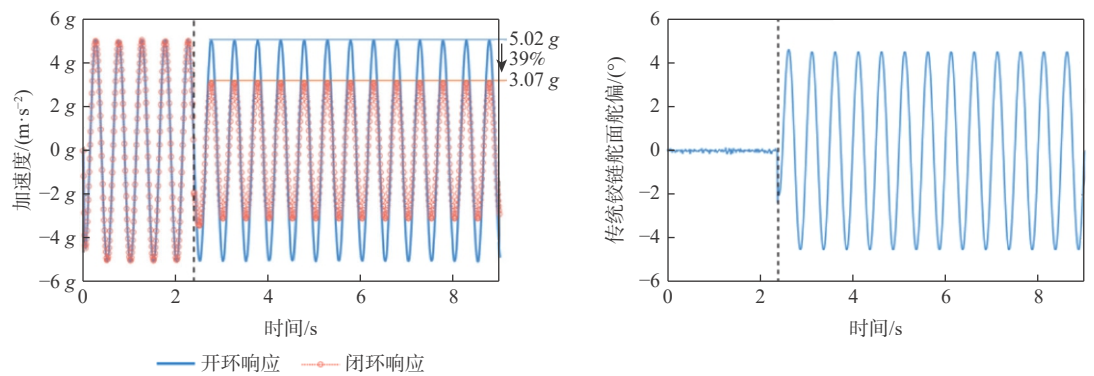
(b) 铰链舵面偏转对应2阶模态广义气动力辨识结果

(c) 铰链舵面偏转对应3阶模态广义气动力辨识结果

(d) 铰链舵面偏转对应4阶模态广义气动力辨识结果

图9 铰链舵面偏转广义模态气动力辨识结果

Fig. 9 Identification result of generalized modal aerodynamic force of hinged flap deflection



(a) GPC开闭环仿真结果

(b) 传统铰链舵面舵偏

图10 铰链舵面闭环阵风响应减缓仿真结果

Fig. 10 Simulation results of closed-loop gust response alleviation based on hinged flap

由图10结果可知，开环翼尖加速度为5.02 *g*，闭环减缓后的翼尖加速度为3.07 *g*，减缓效率为38.84%，在舵面偏转角接近的情况下，变弯度后缘的减缓效率高于传统铰链舵面，进一步验证了变弯度后缘在阵风响应减缓方面的优势。

3 结 论

采用混合气动力建模方法对所给的变弯度机翼完成了开环阵风响应建模，并在此基础上实现了变弯度机翼闭环阵风响应减缓仿真与传统铰链舵面、变弯度后缘气动力对比研究，结果表明：

1）辨识得到的变弯度后缘偏转广义气动力模型在阵风响应减缓要求幅值频率范围内准确性较好，能够有效用于变弯度机翼阵风响应减缓仿真。

2）基于变弯度后缘的GPC方法能够有效减缓由阵风引起的机翼翼尖加速度响应，减缓效率44.33%，验证了针对变弯度后缘利用GPC理论设计控制律进行阵风减缓控制设计的有效性。

3）相比传统铰链舵面，变弯度后缘偏转对机翼动态气动力提升效果更好，在弦向压力分布上无凸起，同时相比传统铰链舵面，采用变弯度后缘进行阵风减缓控制的控制效率更高，在阵风减缓控制方面具有更大的潜力，为变弯度后缘应用于阵风减缓控制提供了理论依据。

参考文献（References）

[1] RIVERO A E, FOURNIER S, MANOLESOS M, et al. Experimental aerodynamic comparison of active camber morphing and trailing-

edge flaps[J]. AIAA Journal, 2021, 59(7): 2627-2640.

[2] 祝连庆, 孙广开, 李红, 等. 智能柔性变形机翼技术的应用与发展[J]. 机械工程学报, 2018, 54(14): 28-42.
ZHU L Q, SUN G K, LI H, et al. Intelligent and flexible morphing wing technology: A review[J]. Journal of Mechanical Engineering, 2018, 54(14): 28-42(in Chinese).

[3] HUNTLEY S J, WOODS B K, ALLEN C B. Computational analysis of the aerodynamics of camber morphing: AIAA 2019-2914[R]. Reston: AIAA, 2019.

[4] NGUYEN N T, CRAMER N B, HASHEMI K E, et al. Progress on gust load alleviation wind tunnel experiment and aeroservoelastic model validation for a flexible wing with variable camber continuous trailing edge flap system: AIAA 2020-0214[R]. Reston: AIAA, 2020.

[5] TAL E A, NGUYEN N T. Unsteady aeroservoelastic modeling of flexible wing generic transport aircraft with variable camber continuous trailing edge flap: AIAA 2015-2722[R]. Reston: AIAA, 2015.

[6] NGUYEN N, TING E, LEBOFSKY S. Aeroelastic analysis of a flexible wing wind tunnel model with variable camber continuous trailing edge flap design[C]//AIAA Science and Technology Forum and Exposition. Reston: AIAA, 2015: 1405.

[7] TING E, CHAPARRO D, NGUYEN N, et al. Optimization of variable-camber continuous trailing-edge flap configuration for drag reduction[J]. Journal of Aircraft, 2018, 55(6): 2217-2239.

[8] HERRERA C Y, SPIVEY N D, LUNG S, et al. Aeroelastic airworthiness assessment of the adaptive compliant trailing edge flaps[C]//Proceedings of the 46th Society of Flight Test Engineers International Symposium, 2015: 14-17.

[9] KOTA S, FLICK P, COLLIER F S. Flight testing of FlexFloil^{TM} adaptive compliant trailing edge[C]//54th AIAA Aerospace Sciences Meeting. Reston: AIAA, 2016: 0036.

[10] 陈磊, 吴志刚, 杨超, 等. 多控制面机翼阵风减缓主动控制与风洞试验验证[J]. 航空学报, 2009, 30(12): 2250-2256.
CHEN L, WU Z G, YANG C, et al. Active control and wind tunnel test verification of multi-control surfaces wing for gust alleviation[J]. Acta Aeronautica et Astronautica Sinica, 2009, 30(12): 2250-2256(in Chinese).

[11] DHILEEP K, KUMAR D, GHOSH S, et al. Numerical study of camber morphing in NACA0012 airfoil: AIAA 2020-2781[R]. Reston: AIAA, 2020.

[12] ULLAH J, LUTZ T, KLUG L, et al. Active gust load alleviation by combined actuation of trailing edge and leading edge flap at transonic speeds: AIAA 2021-1831[R]. Reston: AIAA, 2021.

[13] 顾宁, 陆志良, 郭同庆, 等. 阵风响应及减缓的非定常数值模拟[J]. 航空计算技术, 2012, 42(3): 49-53.
GU N, LU Z L, GUO T Q, et al. Gust response and alleviation analysis of airfoil[J]. Aeronautical Computing Technique, 2012, 42(3): 49-53(in Chinese).

[14] 许晓平, 祝小平, 周洲, 等. 基于CFD方法的阵风响应与阵风减缓研究[J]. 西北工业大学学报, 2010, 28(6): 818-823.
XU X P, ZHU X P, ZHOU Z, et al. Further exploring CFD-based gust response and gust alleviation[J]. Journal of Northwestern Polytechnical University, 2010, 28(6): 818-823(in Chinese).

[15] ZAIDE A, RAVEH D. Numerical simulation and reduced-order modeling of airfoil gust response[J]. AIAA Journal, 2006, 44(8): 1826-1834.

[16] 张伟伟, 叶正寅, 杨青, 等. 基于ROM技术的阵风响应分析方法[J]. 力学学报, 2008, 40(5): 593-598.
ZHANG W W, YE Z Y, YANG Q, et al. Gust response analysis using CFD-based reduced order models[J]. Chinese Journal of Theoretical and Applied Mechanics, 2008, 40(5): 593-598(in Chinese).

[17] 杨国伟, 王济康. CFD结合降阶模型预测阵风响应[J]. 力学学报, 2008, 40(2): 145-153.
YANG G W, WANG J K. Gust response prediction with CFD-based reduced order modeling[J]. Chinese Journal of Theoretical and Applied Mechanics, 2008, 40(2): 145-153(in Chinese).

[18] 聂雪媛, 杨国伟. 基于CFD降阶模型的阵风减缓主动控制研究[J]. 航空学报, 2015, 36(4): 1103-1111.
NIE X Y, YANG G W. Gust alleviation active control based on CFD reduced-order models[J]. Acta Aeronautica et Astronautica Sinica, 2015, 36(4): 1103-1111(in Chinese).

[19] 杨阳, 杨超, 吴志刚, 等. 考虑舵机时滞的阵风减缓主动控制律设计[J]. 北京航空航天大学学报, 2020, 46(12): 2236-2244.
YANG Y, YANG C, WU Z G, et al. Design of gust alleviation active control law considering the time-delay of servo actuator[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(12): 2236-2244(in Chinese).

[20] DAI Y T, YANG C. GPC-based gust response alleviation for aircraft model adapting to various flow velocities in the wind tunnel[J]. Shock and Vibration, 2015, 2015: 348971.

[21] MOULIN B, KARPEL M. Gust loads alleviation using special control surfaces[J]. Journal of Aircraft, 2007, 44(1): 17-25.

[22] 刘晓燕. 基于非定常气动力降阶模型的气动弹性研究[D]. 北京: 北京航空航天大学, 2011: 25-33.
Liu X Y. Aeroelastic rearch based on unsteady aerodynamic reduced order model[D]. Beijing: Beihang University, 2011:25-33(in Chinese).

Numerical study of wing gust response alleviation based on camber morphing trailing edge

WEI Rukai，DAI Yuting*，YANG Chao，YU Siheng

(School of Aeronautic Science and Engineering，Beihang University，Beijing 100191，China)

Abstract：A mathematical model for gust response analysis is established for a wing with camber morphing trailing edges, and a simulation study of gust response alleviation is carried out. The computational fluid dynamics method is used to calculate the generalized unsteady aerodynamic force under the given dynamic morphing of the trailing edge, and the generalized aerodynamic force model under the dynamic deflection of the trailing edge is established based on the state observer method,. The panel method is used to calculate the generalized aerodynamic force caused by mode motion and gust, while the generalized predictive control (GPC) method is used in the design of gust alleviation control law. On this basis, the aerodynamic characteristics of the camber morphing trailing edge and the traditional hinged flap are compared. The simulation results show that the GPC method based on the camber morphing trailing edge can effectively alleviate the wing-tip acceleration response caused by gust, and the wing-tip acceleration reduction efficiency is 44.25%. The wing with morphing trailing edge has a more continuous pressure distribution on the upper and lower surfaces, a greater impact on the aerodynamics, and higher acceleration reduction efficiency. The use of camber morphing trailing edges for gust alleviation has a broader application prospect.

Keywords：camber morphing trailing edge；gust response alleviation；reduced order model；active control；unsteady aerodynamic force

Received：2021-09-22；**Accepted**：2021-11-26；**Published Online**：2022-01-27 08：27

URL：kns.cnki.net/kcms/detail/11.2625.V.20220125.1726.001.html

Foundation item：National Natural Science Foundation of China (11672018)

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

《北京航空航天大学学报》征稿简则

《北京航空航天大学学报》是北京航空航天大学主办的以航空航天科学技术为特色的综合性自然科学学术期刊(月刊)。本刊以反映航空航天领域研究成果与动态、促进学术交流、培养科技人才和推动科技成果向社会生产力转化为办刊宗旨。本刊为中国自然科学技术核心期刊,并被 Ei Compendex 等国内外权威文献检索系统收录。本刊向国内外公开发行,为进一步提高办刊质量和所刊出文章的学术水平,特制定本简则。

1 论文作者及内容

1.1 本刊面向海内外所有学者。

1.2 主要刊载与航空航天科学技术有关的材料科学及工程、飞行器设计与制造、宇航科学与工程、信息与电子技术、控制技术和自动化工程、流体力学和动力工程、计算机科学及应用技术、可靠性工程与失效分析等领域的研究文章。航空航天科学技术民用方面以及具有航空航天工程背景的应用数学、应用物理、应用力学和工程管理等方面的文章也在本刊优先考虑之列。

2 来稿要求

2.1 论文应具有创新性、科学性、学术性和可读性。

2.2 论文为原创作品,尚未公开发表过,并且不涉及泄密问题。若发生侵权或泄密问题,一切责任由作者承担。

2.3 主题明确,数据可靠,图表清晰,逻辑严谨,文字精练,公式符号标准规范,标点符号正确。

2.4 文稿撰写顺序:中文题名(一般不超过 20 个汉字),作者中文姓名、单位、所在城市、邮政编码(作者顺序始终不得更改),中文摘要(包括目的、方法、结果及结论),中文关键词(5 ~ 8 个),中图分类号,引言,正文,参考文献,英文题名,作者英文姓名、单位、所在城市、邮政编码、国别,英文摘要,英文关键词。首、末页下角注明基金项目名称及编号,作者信息。

2.5 作者请登录本刊网页进行在线投稿。

3 稿件的审核、录用与版权

3.1 来稿须经专家两审和主编、编委讨论后决定刊用与否。

3.2 若来稿经审查后认定不宜在本刊发表,将及时告知作者。如果在投稿满 3 个月后仍未收到本刊任何通知,作者有权改投其他刊。在此之前,请勿一稿多投,否则一切后果自负。

3.3 来稿一经作者签字并在本刊刊出,即表明所有作者都已经认可其版权转至本刊编辑部。本刊在与国内外文献数据库或检索系统进行交流及合作时,不再征询作者意见。

邮寄地址:100191 北京市海淀区学院路 37 号 《北京航空航天大学学报》编辑部

办公地点:北京航空航天大学办公楼 405,407 房间

电　　话:(010)82316698,82316981,82315499,82317448

E-mail: jbuaa@ buaa. edu. cn

http://bhxb. buaa. edu. cn

http://www. buaa. edu. cn

《北京航空航天大学学报》
第五届编辑委员会

主　任（主　编）：赵沁平

（以下按姓氏笔画为序）

副主任（副主编）：丁希仑　王少萍　孙志梅　李秋实　李焕喜　杨嘉陵
苗俊刚　相　艳　徐立军　钱德沛　曹晋滨

编　委：马殿富　王　琪　王　聪　邓小燕　王青云　王荣明　刘　宇
刘　红　江　洁　刘　强　闫　鹏　朱天乐　刘铁钢　齐铂金
陈万春　邹正平　苏东林　杨世春　沈成平　邱志平　宋知人
杨树斌　张晓林　杨晓奕　杨继萍　李惠峰　吴新开　张瑞丰
杨照华　宋凝芳　周　锐　林宇震　林贵平　战　强　姚仰平
胡庆雷　赵秋红　段海滨　赵巍胜　席　平　郭　宏　徐　洁
徐世杰　郭洪波　康　锐　翟　锦　熊华钢

北京航空航天大学学报

Beijing Hangkong Hangtian Daxue Xuebao

（原《北京航空学院学报》）

（月刊　1956年创刊）

第49卷 第7期　2023年7月

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS (JBUAA)

(Monthly, Started in 1956)

Vol.49 No.7　July 2023

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　　国内定价：50.00元／期

ISSN 1001-5965