北航学报
赠 阅

ISSN 1001-5965
CODEN BHHDE8

北京航空航天大學学报

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

2016-11
Vol.42 No.11

北京航空航天大学学报

第 42 卷　第 11 期　（总第 285 期）　2016 年 11 月

北航学报
赠阅

目　次

月面钻进真空环境模拟装置的设计与验证 …… 丁希仑，刘舒婷，张涛（2271）
双向旋转柱面气膜密封性能数值模拟与参数优化 …… 马纲，杨卫如（2279）
四轮驱动汽车牵引力控制系统控制策略 …… 姜立标，丘华川，吴中伟，刘巍，程铖（2289）
海基 JPALS 的多参考一致性检验算法 …… 冯涛，毛峡，宋成（2299）
冲击位置对斜搭接接头冲击后拉伸性能的影响 …… 聂恒昌，徐吉峰，关志东，黎增山，冀赵杰，谭日明（2306）
奇异谱分析在故障时间序列分析中的应用 …… 王鑫，吴际，刘超，牛文生，张华，张奎（2321）
基于稳定分簇的移动自组织网络路由协议 …… 黄金科，樊晓光，万明，禚真福，杨永建，陈少华（2332）
面向集群环境的虚拟化 GPU 计算平台 …… 杨经纬，马凯，龙翔（2340）
旋翼/平尾气动干扰对直升机配平特性的影响分析 …… 谭剑锋，吴超，王浩文，林长亮（2349）
基于场景控制特征的安全性需求分析方法 …… 朱丹江，姚淑珍，谭火彬（2358）
基于故障行为模型的产品寿命分析方法 …… 查国清，井海龙，陈云霞，康锐（2371）
未来航空燃料原料可持续性研究 …… 赵晶，郭放，阿鲁斯，杨晓奕（2378）
回转体倾斜入水空泡及弹道特性实验 …… 宋武超，王聪，魏英杰，许昊（2386）
基于子阵激励能量匹配的多子阵交错阵列设计 …… 李龙军，王布宏，夏春和，刘新波，曹帅（2395）
开放空腔壳体入水流场结构及流体动力特征研究 …… 路中磊，魏英杰，王聪，孙铁志（2403）
ARINC 653 分区实时系统的主时间框架设计 …… 谭龙华，杜承烈（2413）
基于 EHW 和 RBT 的电路故障自修复策略性能分析 …… 张峻宾，蔡金燕，孟亚峰（2423）
基于重叠因子优化的超流体陀螺量程扩展设计 …… 赵玉龙，任元，沈怀荣（2436）
基于二级声辐射模型的地面声场高效预测方法 …… 王菲，史勇杰，徐国华（2445）
俯仰操纵方式对自转旋翼机操稳特性的影响 …… 林清，蔡志浩，闫坤，王英勋（2454）
再生型壳与 ZTC4 熔体的界面稳定性 …… 柴连静，程序，吴国清，王红，南海（2466）
基于 LADRC 的无人机高精度定高控制 …… 齐鹏远，王勇，张代兵（2472）
CICQ 结构中逼近 work-conserving 的分组调度算法 …… 张元昊，熊庆旭（2481）
含泡沫镍中间层陶瓷/不锈钢接头微观组织和性能 …… 张羽，朱颖，郭伟，齐丹，祁泽武（2488）
微处理器中含噪热传感器位置分布优化方法 …… 李鑫，李鑫，杨金孝，程元乐（2495）
2×2 X 信道和 PTP 信道并存网络的自由度 …… 郭倩倩，刘锋，曾连荪，覃闻铭（2501）
基于改进 ELECTRE 法的舰船气流场方案评估方法 …… 张修远，常欣，李想，杨放青（2507）
电火花穿孔加工中管状电极振动幅值特性 …… 于志强，佟浩，李勇，柴永生，王志强，李朝将（2516）
基于优化信任评估的 Ad Hoc 安全路由协议 …… 陈少华，樊晓光，禚真福，黄金科，孙贤明（2524）
基于 CEEMDAN 的雷达信号脉内细微特征提取法 …… 王文哲，吴华，王经商，张强（2532）
升力体飞行器能量近似最优倾斜转弯飞行策略 …… 杨丁，刘明，杨凌霄，杨明，葛亚杰（2540）
惯性导航系统动基座传递对准可观测性分析 …… 黄帅，蔡洪，丁智坚（2548）

期刊基本参数：CN11-2625/V * 1956 * m * A4 * 284 * zh * P * ￥20.00 * 900 * 32 * 2016-11

（**编　辑**　娄　嘉　张　嵘　赵海容　李　晶　张欣蔚　孙　芳）

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Vol. 42 No. 11 (Sum 285) November 2016

北航学报 赠阅

CONTENTS

Design and verification of vacuum environmental simulation device for lunar surface drilling
…… DING Xilun, LIU Shuting, ZHANG Tao (2271)
Numerical simulation and parameter optimization of seal property of bidirectional rotating cylinder gas film
…… MA Gang, YANG Weiru (2279)
Control strategy of traction control system for four-wheel drive vehicle
…… JIANG Libiao, QIU Huachuan, WU Zhongwei, LIU Wei, CHENG Cheng (2289)
Multiple reference consistency check algorithm for sea-based JPALS
…… FENG Tao, MAO Xia, SONG Cheng (2299)
Tensile behavior of scarf joints after impact in different locations
…… NIE Hengchang, XU Jifeng, GUAN Zhidong, LI Zengshan, JI Zhaojie, TAN Riming (2306)
Application of singular spectrum analysis to failure time series analysis
…… WANG Xin, WU Ji, LIU Chao, NIU Wensheng, ZHANG Hua, ZHANG Kui (2321)
Stable cluster-based routing protocol for mobile Ad Hoc networks
…… HUANG Jinke, FAN Xiaoguang, WAN Ming, ZHUO Zhenfu, YANG Yongjian, CHEN Shaohua (2332)
Virtualized GPU computing platform in clustered system environment
…… YANG Jingwei, MA Kai, LONG Xiang (2340)
Analysis on influence of rotor/empennage aerodynamic interaction on helicopter trim characteristic
…… TAN Jianfeng, WU Chao, WANG Haowen, LIN Changliang (2349)
Safety requirements analysis method based on control characteristics of scenarios
…… ZHU Danjiang, YAO Shuzhen, TAN Huobin (2358)
Life analysis method for products based on failure behavior models
…… ZHA Guoqing, JING Hailong, CHEN Yunxia, KANG Rui (2371)
Research of sustainable feedstock for future alternative aviation fuels
…… ZHAO Jing, GUO Fang, A Lusi, YANG Xiaoyi (2378)
Experiment of cavity and trajectory characteristics of oblique water entry of revolution bodies
…… SONG Wuchao, WANG Cong, WEI Yingjie, XU Hao (2386)
Design of array with multiple interleaved subarrays based on subarray excitation energy-matching
…… LI Longjun, WANG Buhong, XIA Chunhe, LIU Xinbo, CAO Shuai (2395)
Flow field structure and fluid dynamic characteristics of semi-closed cylinder water-entry
…… LU Zhonglei, WEI Yingjie, WANG Cong, SUN Tiezhi (2403)
Design of major time frame for ARINC 653 partitioned real-time systems
…… TAN Longhua, DU Chenglie (2413)
Performance analysis of circuit fault self-repair strategy based on EHW and RBT
…… ZHANG Junbin, CAI Jinyan, MENG Yafeng (2423)
Range extension design on superfluid gyroscope based on overlap factor optimization
…… ZHAO Yulong, REN Yuan, SHEN Huairong (2436)
Efficient ground noise prediction method based on secondary noise radiation model
…… WANG Fei, SHI Yongjie, XU Guohua (2445)
Influence of pitch manipulation modes on controllability and stability of autogyro
…… LIN Qing, CAI Zhihao, YAN Kun, WANG Yingxun (2454)
Interfacial stability of regenerated moulds against molten ZTC4 alloy
…… CHAI Lianjing, CHENG Xu, WU Guoqing, WANG Hong, NAN Hai (2466)
Precise height control for UAV based on LADRC
…… QI Pengyuan, WANG Yong, ZHANG Daibing (2472)
Work-conserving-based packet scheduling algorithm for CICQ switches
…… ZHANG Yuanhao, XIONG Qingxu (2481)
Microstructure and properties of ceramic/stainless steel joint with Ni foam interlayer
…… ZHANG Yu, ZHU Ying, GUO Wei, QI Dan, QI Zewu (2488)
Optimization method of thermal sensor placement for microprocessor with noise
…… LI Xin, LI Xin, YANG Jinxiao, CHENG Yuanle (2495)
Degrees of freedom of networks with coexistence of 2 × 2 X channel and point-to-point channel
…… GUO Qianqian, LIU Feng, ZENG Liansun, QIN Wenming (2501)
Evaluation method of airflow field plan in warship based on improved ELECTRE method
…… ZHANG Xiuyuan, CHANG Xin, LI Xiang, YANG Fangqing (2507)
Vibration amplitude characteristics of tubular electrode in EDM drilling
…… YU Zhiqiang, TONG Hao, LI Yong, CHAI Yongsheng, WANG Zhiqiang, LI Chaojiang (2516)
Secure routing protocol based on optimized trust evaluation in Ad Hoc network
…… CHEN Shaohua, FAN Xiaoguang, ZHUO Zhenfu, HUANG Jinke, SUN Xianming (2524)
Subtle intrapulse feature extraction based on CEEMDAN for radar signals
…… WANG Wenzhe, WU Hua, WANG Jingshang, ZHANG Qiang (2532)
Energy suboptimal flight strategy of bank-to-turn maneuver for lifting vehicle
…… YANG Ding, LIU Ming, YANG Lingxiao, YANG Ming, GE Yajie (2540)
Observability analysis for transfer alignment of inertial navigation system on moving base
…… HUANG Shuai, CAI Hong, DING Zhijian (2548)

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0669

月面钻进真空环境模拟装置的设计与验证

丁希仑*, 刘舒婷, 张涛

(北京航空航天大学 机械工程及自动化学院, 北京 100083)

摘 要: 探月工程(三期)项目提出了在月面进行至少 2 m 深度的钻取采样任务,前期需要在地面进行模拟月面钻进过程热特性的研究,为此本文发明了一套月面真空环境模拟装置。根据模拟月壤钻进试验的要求,设计了一种可从底部抽气的三段式可多次拆装的真空罐结构。考虑到模拟月壤的颗粒大小和快速抽真空的要求,设计了一种筛网状具有多层过滤结构的月壤筒。根据在真空度不变的条件下可进行连续多次钻进试验的要求,设计了可从外部操作的具有缺齿结构的月壤筒转动机构。由真空度要求和可耐粉尘的特性要求,选取了三级扩散泵组。实验表明,在不放置模拟月壤的情况下,本环境模拟装置中的真空度可达 10^{-1} Pa,在放置模拟月壤的情况下,真空度可达 7 Pa,能够满足模拟月面钻进过程热特性试验的要求。此外,通过测量抽真空过程中真空度随时间的变化情况还得出了模拟月壤出气量曲线,对模拟月壤真空系统的实验应用具有重要价值。

关 键 词: 真空装置; 环境模拟; 钻进试验; 模拟月壤; 真空实现; 动密封
中图分类号: V524.3; TB771
文献标识码: A **文章编号**: 1001-5965(2016)11-2271-08

探月工程(三期)项目的任务为发射无人月面着陆器,在月面进行钻取和采样工作,并将采集到的月壤样品带回地球[1]。为了保证钻取机构的可靠性,需要在地面开展一系列模拟月面钻进过程的热特性试验研究。由于近月空间和月面环境中,真空度高达 10^{-11} Pa[2],此高真空环境中,月壤导热性差、热导率低和密实度高等特点会对钻进过程热特性试验的结果产生较大影响。因此,钻进试验需要在一个模拟月面的真空环境中进行[3],以选择合理的钻进策略和安全钻进的临界条件。

目前,国外普遍采用全比模型分析法来检验机构,苏联将整个采样器放置在月球环境模拟器中观察其工作状况,此研究方法对环境模拟器的空间和采样器的真空适应性要求较高[4]。美国国家航天局(NASA)的 Fountain 和 West[5]将少量模拟月壤置于真空度达 1.33×10^{-6} Pa 的离子泵真空系统中来研究模拟月壤的热导率。美国肯塔基大学的 Cremers 等[6]将 6 g Apollo 12 月壤样品置于直径为 300 mm、高为 400 mm 的不锈钢真空室内进行热导率相关实验,真空环境可达 10^{-4} Pa,由分子泵实现。以上 2 个真空系统均需要严格控制粉尘污染,真空室内只能放置少量模拟月壤,无法满足深层钻进试验的要求。美国喷气推进实验室 Bernett 等[7]通过实验发现真空环境对细粒橄榄玄武岩热导率有很大影响,其实验所用的真空室采用玻璃制成,直径为304.8 mm,真空环境可达 6.6×10^{-4} Pa,由口径为 100 mm 的扩散泵组实现,但是由于真空室体积较小,无法容纳过多的模拟月壤。美国 Honeybee 实验室将其研制的 LunarVader 钻取采样装置放置在真空环境中进行性能测试,真空室的长宽各1 m、高 3.5 m,内置 1 m 长的细长圆柱体模拟月壤,模拟月壤掺

收稿日期: 2015-10-16; **录用日期**: 2015-11-20; **网络出版时间**: 2015-12-17 11:33
网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151217.1133.018.html
* **通讯作者**: Tel.: 010-82338005 E-mail: xlding@buaa.edu.cn

引用格式: 丁希仑, 刘舒婷, 张涛. 月面钻进真空环境模拟装置的设计与验证[J]. 北京航空航天大学学报, 2016, 42(11): 2271-2278. DING X L, LIU S T, ZHANG T. Design and verification of vacuum environmental simulation device for lunar surface drilling [J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(11): 2271-2278 (in Chinese).

满水后压实冻结,钻进深度达750 mm,真空度可维持在853 Pa[8],该试验中模拟月壤需要为掺水状态,与真实月壤参数有较大不同,且真空室内真空度过低。中国哈尔滨工业大学崔金生等[9]在真空环境中进行钻进试验,钻进深度不超过50 mm,真空环境采用旋片式机械泵来实现,真空度变化范围为30~230 Pa,此试验环境真空度同样过低,另外允许钻进深度也较小。北京航空航天大学张涛等[10]用旋片式机械泵对模拟月壤进行抽真空实验,真空室直径为308 mm、高为200 mm,内部置满烘干压实的模拟月壤,极限真空度达19 Pa,该真空系统可允许放置大量模拟月壤,但是并未与钻进试验相结合,真空度也较低。

目前,国内外月面真空环境模拟技术均处于试验探究阶段,模拟月壤因其硬度等特性,对真空泵叶片有较大损害[11],因此真空装置内不能放置过多模拟月壤,或选择将模拟月壤冻结。此外,目前应用于钻进试验的真空室尺寸较小,不能进行深层多次钻进采样试验。

针对月面钻进过程热特性试验研究的需求,本文设计了一种用于深层多次钻进过程热特性试验研究的月面真空环境模拟装置,并验证了其不同工作条件下的真空性能。

1 真空度指标的确定

1963 年,美国喷气推进实验室 Bernett 等[7]研究了真空环境中模拟月壤的热扩散率与热导率,发现在1.33~13.3 Pa 与 1.33×10^{-4}~1.33×10^{-1} Pa 真空环境中,模拟月壤的热导率并没有明显不同,可知在10 Pa 左右与更低的真空度环境中,模拟月壤的热导率基本为常值。

1965 年,NASA 研究者 Wechsler 和 Glaser[12]对几种不同成分的模拟月壤进行了真空热导率测试实验。实验结果显示,当周围气压低于10 Pa 后,玄武岩粉末、浮石粉末和花岗岩等的热导率随气压变化很小,变化范围不超过 10^{-5} W/(cm·K)。

1975 年,美国加利福尼亚大学研究者 Tien 和 Nayak[13]对 Apollo 系列样品在真空中的热导率变化进行了对比,发现在1.33 ~133 Pa 范围内月壤热导率变化曲线与 Yovanovich 和 Kitscha[14]的理论模型趋于一致:当气压低于10 Pa 时,随着气压降低,月壤的热导率变化很小,变化范围不超过 10^{-5} W/(cm·K);当气压接近1 Pa 时,随着气压降低,月壤的热导率基本上不再变化。

由以上可知,气压低于10 Pa 后,随着气压降低,月壤的热导率变化很小。在地面环境中模拟月面极高的真空度有很大难度,若将月面真空环境模拟装置的真空度设置在10 Pa 以下,则由于月壤的热导率变化很小,钻进试验过程中的热特性与在真实月面环境中的结果近似等效。

2 月面真空环境模拟装置的研制

2.1 设计指标

根据月面钻进过程热特性试验研究对真空环境的需求,月面真空环境模拟装置的设计指标如下:

1)钻杆钻进行程:≥2 m。

2)月面真空环境模拟装置空载(不内置模拟月壤)真空度:10^{-1} Pa,抽真空时间:<24 h。

3)月面真空环境模拟装置带载(内置模拟月壤)真空度:<10 Pa,抽真空时间:<24 h。

4)每次达到真空度指标后,可连续进行3次钻进试验。

2.2 方案设计

根据第2.1节设计指标和要求,提出如下月面真空环境模拟装置的方案,装置由真空罐模块、月壤筒模块和真空获得模块组成,其具体功能模块组成如图1所示。

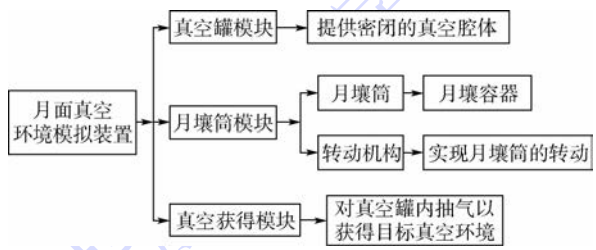


图1 装置的功能模块组成

Fig. 1 Composition of device's function modules

根据各个功能模块设计了月面真空环境模拟装置的三维模型,如图2所示。

月面真空环境模拟装置的工作流程为:首先完成月壤筒模块的安装,将盛满模拟月壤的月壤筒固定在真空罐内部转动机构上;然后完成真空罐模块的安装,将真空罐各部分之间采用橡胶圈压紧密封;最后由真空获得模块对真空罐抽气来实现目标真空环境,至此月面真空环境模拟装置内部的类月面环境模拟完成。在模拟类月面环境中进行钻进试验时,真空获得模块保持工作状态,以维持真空罐内部的真空环境,每进行一次钻进试验,可在转动月壤筒后进行下一次钻进试验,以满足在不改变真空模拟环境的前提下进行多次钻进的试验要求。

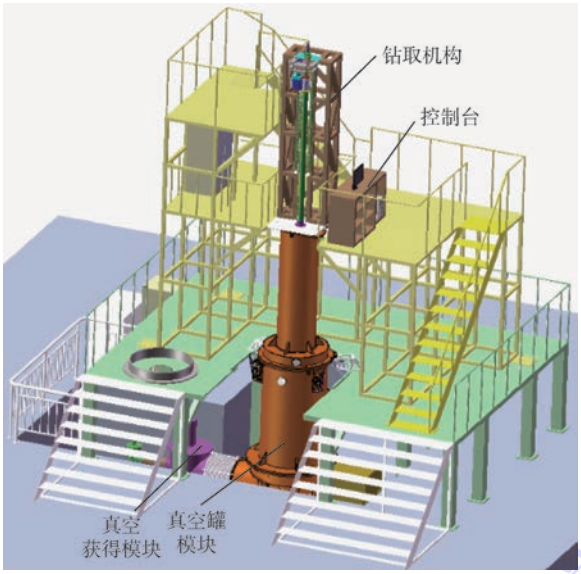


图 2　月面真空环境模拟装置的三维模型

Fig. 2　Three-dimensional model of lunar surface vacuum environmental simulation device

2.2.1　真空罐模块

根据模拟月壤钻进试验的几何尺寸、密封性要求、安装与操作的方便性及与其他模块的连接关系,将真空罐设计为上、中、下 3 段,如图 3 所示。真空罐总体高度为 5.9 m,容积为 4.148 m^3。由于从月壤筒底部抽气,可减轻月尘飞扬现象,将抽气口设计在真空罐底部。往罐内充入氮气采用皂泡检漏法对真空罐进行正压检测,通过压力表的读数变化可计算出真空罐的漏率约为 80 Pa · L/s。



图 3　真空罐

Fig. 3　Vacuum chamber

试验所采用的钻机不能在真空环境中工作,因此将钻机设计在真空罐顶端,钻杆伸入真空罐内,钻杆光轴段与真空罐顶盖之间的动密封接口采用磁流体与波纹管组合密封。钻杆在真空罐内的部分达 2.3 m,钻进深度为 2 m,在钻进过程中,极有可能出现失稳变形[15],因此需要在真空罐内设计对钻杆的导向装置,导向平台安装在真空罐上段的底端。

2.2.2　月壤筒模块

根据钻进深度 2 m 要求、钻杆尺寸等,将月壤筒设计为图 4 所示结构,内径为 500 mm,高为 2 050 mm,安装在真空罐内部。为方便更换模拟月壤,月壤筒壁和底部之间设计为螺栓连接,更换模拟月壤时,将连接螺栓拆除即可。为了实现对真空罐抽气时,可同时对模拟月壤内部进行抽气的效果,在月壤筒侧壁和底部设计有均布的直径为 30 mm 透气孔,并且根据模拟月壤颗粒大小,选择 20 目铁丝网、800 目不锈钢网、20 目铁丝网作为内胆由外至内依次紧贴月壤筒内壁放置。图 5 所示为月壤筒强度分析图(将模拟月壤假设为流体),验证得其强度满足试验要求。

钻进试验要求在不改变真空模拟环境的前提下可进行多次钻进试验,因此将钻杆设计在与月壤筒偏心的位置,转动月壤筒即可变换模拟月壤表面钻进点。如图 6 所示,月壤筒重量较大,选择

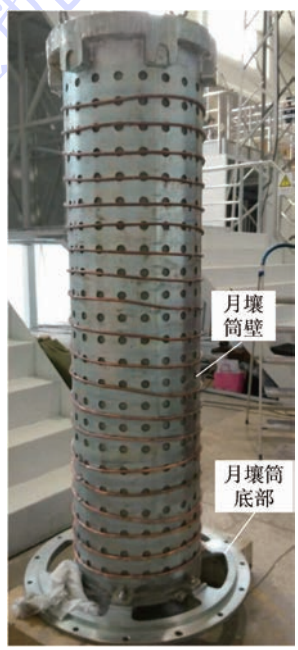


图 4　月壤筒

Fig. 4　Lunar soil container

北航学报 赠阅



图 5 月壤筒强度分析

Fig. 5 Strength analysis of lunar soil container

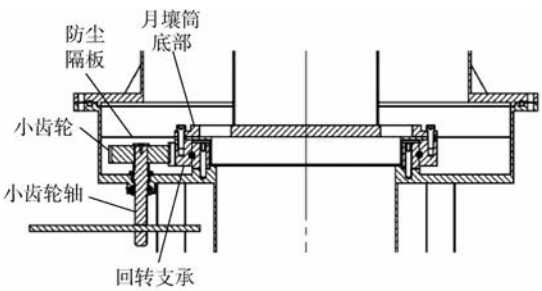


图 6 转动机构

Fig. 6 Slewing mechanism

可承受大重量的回转支承来实现月壤筒的转动，回转支承外圈带有缺齿齿轮，缺齿用于粗略定位。执行转动操作时，在真空罐外转动操作杆带动小齿轮轴转动，进而带动小齿轮转动，小齿轮通过齿轮副带动回转支承外圈转动，月壤筒跟随回转支承外圈转动，完成转动操作。小齿轮轴穿过真空罐壁，与真空罐壁之间属于低速旋转密封，选择适用于此密封状态的 J 型动密封结构。转动机构上方设置防尘隔板，可隔离部分模拟月壤粉尘。

2.2.3 真空获得模块

根据真空度要求与可耐粉尘特性要求，真空泵组选用扩散泵组，其组成如图 7 所示，包括：扩散泵 TK-630，作为维持泵的旋片式真空泵 2X-15，作为主预抽泵的罗茨泵 ZJP-600，作为辅助预抽泵的 2 台旋片式真空泵 2X-70。在扩散泵的入口与高真空阀门之间安装水冷挡板，用于阻挡扩散泵中的油蒸汽，降低泵的返油率。整个扩散泵组配有水冷系统，为各个泵散热。

为减少扩散泵组抽气口与真空罐抽气口之间管道的长度，降低抽速损失，扩散泵组的安装面比真空罐固定平台低 1.4 m。扩散泵组抽气口与真空罐抽气口之间连接有波纹管，可允许对接装配

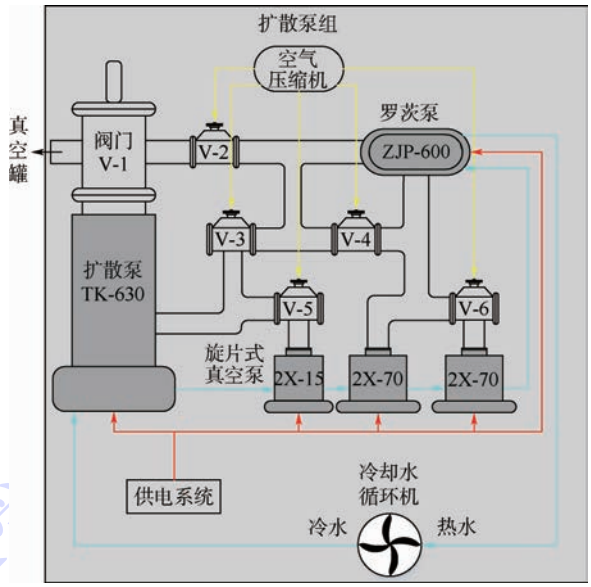


图 7 扩散泵组

Fig. 7 Diffusion pump set

中存在一定范围的误差，同时也能吸收扩散泵组产生的振动。

真空罐空载性能的理论分析如下[16]。

真空罐极限真空 P_j 为

$$P_j = P_0 + \frac{Q_1}{S} \tag{1}$$

式中：$P_0 = 5 \times 10^{-4}$ Pa 为真空泵组的极限真空；$S = 1\,800$ L/s为有效抽速，高真空状态下有效抽速受管道流导影响较大，按照泵组抽速的 30% 计算；$Q_1 = 80$ Pa · L/s 为真空罐漏率，相比于真空罐漏率，材料表面出气率忽略不计。计算得 $P_j = 4.49 \times 10^{-2}$ Pa。

真空罐的工作压力应比极限真空高一个数量级，则工作压力 $P_g = 1 \times 10^{-1}$ Pa，满足设计指标要求。

用 2 台旋片式真空泵 2X-70 从大气抽到 260 Pa，时间 t_1 为

$$t_1 = 2.3 K_q \frac{V}{S_p} \lg \frac{P_i - P_j}{P - P_j} \tag{2}$$

式中：$K_q = 1.5$ 为机械泵抽速修正系数；$V = 4.148\ \mathrm{m}^3$ 为真空罐容积；$S_p = 70$ L/s 为名义抽速，考虑抽速在泵组内管道的损失，按 2 台旋片式真空泵抽速的 50% 计算；P_i、P 分别为抽真空前、后真空罐内的压强。计算得 $t_1 = 8.8$ min。

用罗茨泵 ZJP-600 从 260 Pa 抽到 1.2 Pa，$K_q = 4$，$S_p = 600$ L/s 为罗茨泵的名义抽速，用式(2)计算得时间 $t_2 = 2.5$ min。

打开扩散泵组，从 1.2 Pa 开始抽真空，1 h 后不锈钢出气率 $q = 2.3 \times 10^{-5}$ Pa · L/(s · cm^2)，真

北航学报 赠阅

空罐内表面出气量 Q_2 为

$$Q_2 = qA \tag{3}$$

式中:$A = 22 \times 10^4$ cm^2 为真空罐内不锈钢总面积。计算得 $Q_2 = 5$ Pa · L/s。

高真空状态下,真空罐内气体总负荷 $Q_3 = Q_2 + Q_1 = 85$ Pa · L/s。

1 h 后真空罐内压强 P 为

$$P = \frac{Q_3}{S} \tag{4}$$

计算得 $P = 4.7 \times 10^{-2}$ Pa,可见 1 h 后即可达到工作压力。

综上,空载状态下总抽气时间 t 不超过 1.5 h。

3　性能验证实验

3.1　实验方法

采用 2 个压阻电阻复合真空计测量真空罐内的真空度,真空计 1、真空计 2 位置如图 3 所示。

实验过程如下:

1) 空载真空实验

目的:验证真空罐本身的焊接密封性能和一些动接口的密封性能。

步骤:

① 真空罐内不放置月壤筒,进行空载真空实验:先打开 2 台旋片式真空泵,待真空计 1 达 260 Pa时打开罗茨泵,待 1.2 Pa 时打开扩散泵。全过程中,每隔 1 s 记录一次真空计 1、真空计 2 的数值。

② 待 2 个真空计数值稳定后,分别顺时针、逆时针转动转动机构中的小齿轮轴,每隔 1 s 记录一次真空计 1 的数值,测试转动机构接口的密封性能。

③ 待 2 个真空计数值稳定后,控制钻杆回转速度为 120 r/min,进给速度为 150 mm/min,每隔 0.5 s 记录一次真空计 2 的数值,测试真空罐顶盖与钻机接口的密封性能。

2) 真空罐静态升压实验

目的:校核真空罐的漏率。

步骤:关闭泵组阀门,将真空罐与泵组阻断隔开,每隔 1 s 记录一次真空计 1、真空计 2 的数值。

3) 带载真空实验

目的:验证装置所模拟的真空环境是否满足钻进试验的真空度指标要求。

步骤:

① 在月壤筒内置满烘干后的模拟月壤(模拟月壤由玄武岩和玄武质火山渣组成),分 4 次放置,每次放置后均在三维振动台(见图 8)上振动压实30 min。三维振动台的振动频率为60 Hz,振幅为 2 mm。



图 8　三维振动台

Fig. 8　Three-dimensional vibrator

② 月壤筒顶端固定一个纸板盖,可防止抽真空时模拟月壤喷射。将月壤筒固定在真空罐内的转动机构上,然后将真空罐封闭。

③ 进行带载真空实验:先打开 2 台旋片式真空泵,待真空计 1 达 500 Pa 时打开罗茨泵,待 7 Pa 时打开扩散泵。全过程中,每隔 1 s 记录一次真空计 1、真空计 2 的数值。

3.2　空载真空实验

3.2.1　空载真空度变化曲线

图 9 为空载真空实验真空度曲线。可以看出,打开旋片式真空泵后,气压从大气降到 260 Pa 所用时间约为 11 min,打开罗茨泵后,气压从 260 Pa降到 1.2 Pa 所用时间约为 2.2 min,均与理论值接近。打开扩散泵 13 min 后,真空计 1 达到工作压力 $P_g = 1 \times 10^{-1}$ Pa,真空计 2 压力仅为 1.8×10^{-1} Pa,原因是抽速经过转动机构后有所损失。

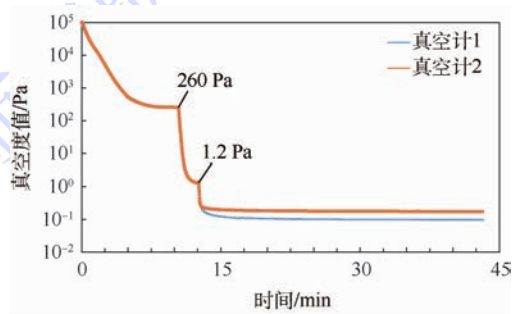


图 9　空载真空实验真空度曲线

Fig. 9　Vacuum curves of no-load experiment

3.2.2　转动机构接口密封性能测试

真空计 1 离转动机构接口较近,转动小齿轮轴驱动转动机构工作时,由图 10 可得真空计 1 示数变化范围仅为 4×10^{-4} Pa,可以得出转动机构中小齿轮轴与真空罐壁之间采用的 J 型密封结构具有良好的密封性能,可保证真空罐内的真空度在转动月壤筒的过程中仍满足钻进试验要求。

北航学报 赠阅

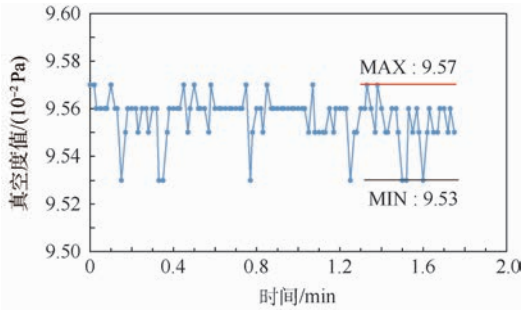


图 10 转动机构接口密封性能测试

Fig. 10 Sealing performance test of slewing mechanism's interface

3.2.3 钻机接口密封性能测试

真空计 2 离钻机接口较近，钻杆同时以 120 r/min转速回转，150 mm/min 速度进给时，由图 11 可得真空计 2 示数变化范围仅为 3×10^{-4} Pa，可以得出光轴与真空罐顶盖之间所采用的磁流体与波纹管组合密封结构具有良好的密封性能，可满足高速旋转与往复直线同步运动过程中的密封要求。

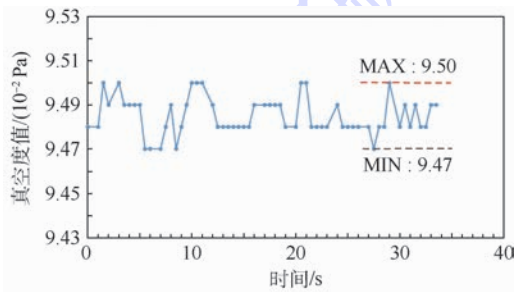


图 11 钻机接口密封性能测试

Fig. 11 Sealing performance test of drill's interface

3.3 真空罐静态升压实验

真空罐与泵组阻断隔开后，真空度-时间曲线近似为一条具有一定斜率的直线，如图 12 所示，表明容器中的漏气是压力上升的主要因素，金属放气量可以忽略。

采用静态升压法计算真空罐漏率，曲线斜率 $k=1.78\times10^{-2}$ Pa/s，漏率 Q_1' 为

$$Q_1'=Vk \tag{5}$$

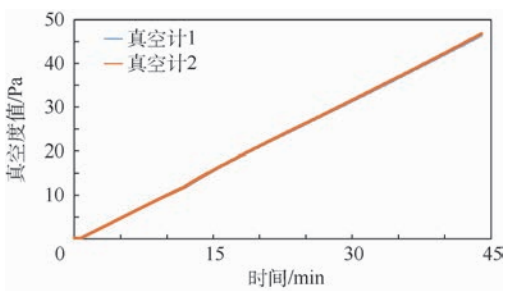


图 12 静态升压实验真空度曲线

Fig. 12 Vacuum curves of static pressurization experiment

经计算得真空罐漏率 $Q_1'=73.83$ Pa·L/s。

3.4 带载真空实验

3.4.1 带载真空度变化曲线

图 13 为第 1 次带载真空实验真空度曲线。可以看出，打开旋片式真空泵后，气压从大气降到 500 Pa 所用时间约为 13 min。打开罗茨泵后，真空计 2 示数先于真空计 1 迅速下降，这是因为放置月壤筒与防尘隔板后，到达真空计 1 测量处的抽速大大减小。4.5 min 后，真空罐内气流趋于平稳，此后模拟月壤内部的气体开始慢慢释放。真空计 2 位于月壤筒附近，受模拟月壤出气量的影响，后期其示数较大。80 min 后关闭真空泵组，此时真空计 1 达 25 Pa，真空计 2 达 34 Pa。

在第 1 次带载真空实验的基础上进行第 2 次带载真空实验，2 次实验操作过程相同，实验结果如图 14 所示。可以看出，真空计 1 从大气降到 25 Pa 仅用了 25 min，原因是在第 1 次带载真空实验时，模拟月壤内部的部分气体已被抽出，缩短了第 2 次带载真空实验的时间。最终 4.5 h 后，真空计 1 达 7 Pa，真空计 2 达 8.2 Pa，均小于 10 Pa，满足钻进试验对真空度的要求。

真空计 1 达 7 Pa 后，打开扩散泵，真空度几乎没有变化，因此带载条件下，扩散泵几乎不起作用。

实验证明，第 3 次带载真空实验仅需 2.5 h 即可达到目标真空度值 10 Pa。

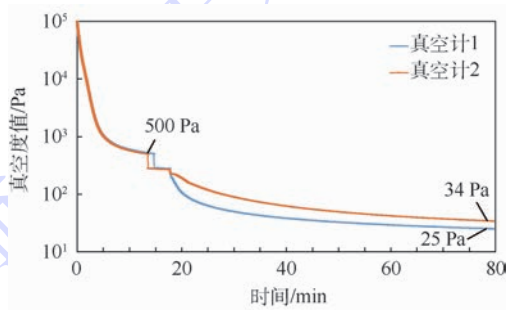


图 13 第 1 次带载真空实验真空度曲线

Fig. 13 Vacuum curves of the first load experiment

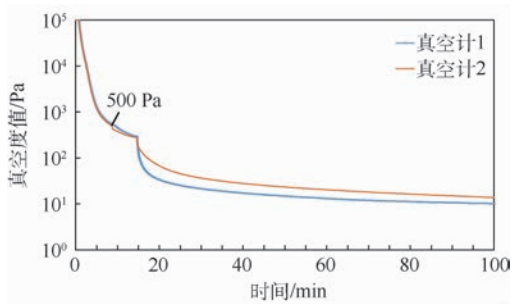


图 14 第 2 次带载真空实验真空度曲线

Fig. 14 Vacuum curves of the second load experiment

3.4.2　模拟月壤出气量曲线

真空罐内 60 Pa 以下时,相比模拟月壤出气量,不锈钢罐体出气、漏率等可忽略不计,因此真空罐内真空度主要受模拟月壤出气量的影响。真空罐内模拟月壤出气量 Q_g 为

$$Q_g = P_d S \tag{6}$$

式中:P_d 为真空罐内带载真空度;S = 300 L/s 为有效抽速,按罗茨泵抽速 50% 计算。

可得真空罐内模拟月壤出气量变化曲线,如图 15 所示。

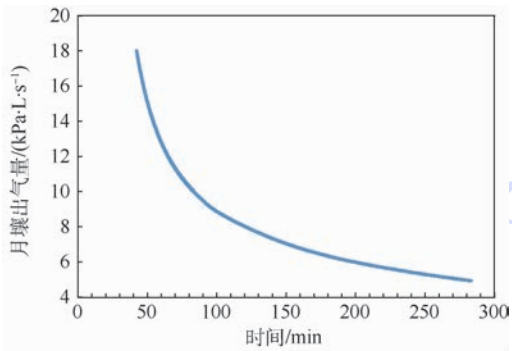


图 15　模拟月壤出气量曲线

Fig. 15　Outgassing amount curve of lunar soil simulant

3.4.3　带载真空实验后真空罐内状态

图 16 所示为带载真空实验结束后真空罐内的状态。真空罐内壁上仅附有一层很薄的模拟月尘,可见抽真空过程中,并未发生模拟月壤喷射现象,模拟月尘飞扬现象也不严重。



图 16　带载真空实验后真空罐内状态

Fig. 16　Situation in vacuum chamber after load experiment

4　结　论

本文设计并研制了一种用于深层多次钻进过程热特性试验研究的月面真空环境模拟装置,包括可从底部抽气的三段式可多次拆装的真空罐结构;筛网状具有多层过滤结构的月壤筒;可从外部操作的具有缺齿结构的月壤筒转动机构和三级扩散泵组。对装置的各个功能模块及真空度性能进行了一系列实验验证,实验结果表明:

1)在真空罐外即可对月壤筒进行转动操作,满足连续 3 次钻进试验的要求,转动机构接口所采用的 J 型密封结构密封性能良好,满足低速旋转密封要求。

2)钻机与真空罐顶盖之间接口所采用的磁流体与波纹管组合密封结构密封性能良好,满足高速旋转与往复直线同步运动过程中的密封要求。

3)带载真空实验中,真空罐内真空度可达 10 Pa 以下,满足钻进过程热特性试验对真空度的要求。

4)在不更换模拟月壤的条件下,随着带载真空实验的次数增多,模拟月壤内部越来越多的气体被抽出,所需要的抽真空时间也越来越短。

5)该装置通过多层过滤网结构和从真空罐底部抽气的设计,在实现对模拟月壤内部抽气的同时,可有效减轻模拟月尘飞扬现象,解决了创建真空环境中防止尘土污染的技术难题。

参考文献(References)

[1] 欧阳自远. 月球科学概论[M]. 北京:中国宇航出版社,2005:313-317.
OUYANG Z Y. Introduction to lunar science[M]. Beijing:China Astronautic Publishing House,2005:313-317(in Chinese).

[2] 陈金明,景甫林,黄本诚. 开展我国月球环境模拟试验技术研究的建议与设想[J]. 航天器环境工程,2003,20(2):1-4.
CHEN J M,JING F L,HUANG B C. Suggestion on the research on lunar environmental simulation test techniques in China[J]. Spacecraft Environment Engineering,2003,20(2):1-4(in Chinese).

[3] SHI X M,DENG Z Q,QUAN Q Q,et al. Development of a drilling and coring test-bed for lunar subsurface exploration and preliminary experiments[J]. Chinese Journal of Mechanical Engineering,2014,27(4):673-682.

[4] 侯佑松. 月面钻取采样过程钻具热特性测试系统研制与试验研究[D]. 哈尔滨:哈尔滨工业大学,2013:3-4.
HOU Y S. Drill thermal characteristics test system development and experiment research of the lunar surface drilling sampling process[D]. Harbin:Harbin Institute of Technology,2013:3-4(in Chinese).

[5] FOUNTAIN J A,WEST E A. Thermal conductivity of particulate basalt as a function of density in simulated lunar and Martian environments[J]. Journal of Geophysical Research,1970,75(20):4063-4069.

[6] CREMERS C J,BIRKEBAK R C,DAWSON J P. Thermal conductivity of fines from Apollo 12[C]// Proceedings of the Second Lunar Science Conference. Cambridge:MIT Press,1971:2311-2315.

[7] BERNETT E C,WOOD H L,JAFFE L D,et al. Thermal properties of a simulated lunar material in air and in vacuum[J]. AIAA Journal,1963,1(6):1402-1407.

北航学报 赠阅

[8] ZACNY K, PAULSEN G, SZCZESIAK M, et al. LunarVader: Development and testing of lunar drill in vacuum chamber and in lunar analog site of Antarctica[J]. Journal of Aerospace Engineering, 2013, 26(1): 74-86.

[9] CUI J S, HOU X Y, ZHAO D M, et al. Thermal simulation and experiment of lunar drill bit in vacuum[J]. Indonesian Journal of Electrical Engineering and Computer Science, 2014, 12(6): 4756-4763.

[10] 张涛,安迎和,丁希仑. 模拟月壤真空试验装置设计及实现方法研究[J]. 北京航空航天大学学报, 2015, 41(11): 2110-2115.
ZHANG T, AN Y H, DING X L. The design of vacuum test device and the study of pumping method about lunar soil simulants [J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(11): 2110-2115(in Chinese).

[11] 陈诗吉. 月壤热物性特征及其剖面温度垂直分布模拟[D]. 上海:华东师范大学, 2014: 12-20.
CHEN S J. Thermophysical properties of lunar regolith and simulation of its vertical temperature distribution[D]. Shanghai: East China Normal University, 2014: 12-20(in Chinese).

[12] WECHSLER A E, GLASER P E. Pressure effects on postulated lunar materials[J]. Icarus, 1965, 4(4): 335-352.

[13] TIEN C L, NAYAK A L. Analytical models for lunar soil thermal conductivity[C]// AIAA 10th Thermophysics Conference. Reston: AIAA, 1975: 27-29.

[14] YOVANOVICH M M, KITSCHA W W. Modeling the effect of air and oil upon the thermal resistance of a sphere-flat contact [C]// AIAA 8th Thermophysics Conference. Reston: AIAA, 1974: 293-319.

[15] 李京霖,丁希仑,张文明,等. 一种多功能小行星采样器的设计[J]. 机械工程学报, 2015, 51(13): 167-175.
LI J L, DING X L, ZHANG W M, et al. Design of a multi-function mintor planet soil sampler[J]. Journal of Mechanical Engineering, 2015, 51(13): 167-175(in Chinese).

[16] 达道安. 真空设计手册[M]. 北京:国防工业出版社, 2004: 769-786.
DA D A. Vacuum design manual[M]. Beijing: National Defense Industry Press, 2004: 769-786(in Chinese).

作者简介:

丁希仑 男,博士,教授,博士生导师。主要研究方向:空间机器人与空间机构学、仿生机器人、智能控制与自动化检测技术等。

Tel.: 010-82338005

E-mail: xlding@buaa.edu.cn

刘舒婷 女,硕士研究生。主要研究方向:地外环境模拟、地外钻取采样等。

Tel.: 010-82339055

E-mail: liushuting91@foxmail.com

Design and verification of vacuum environmental simulation device for lunar surface drilling

DING Xilun*, LIU Shuting, ZHANG Tao

(School of Mechanical Engineering and Automation, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: The third stage of China's lunar exploration project presented the task of drilling samples of lunar soil over 2 m deep, which motivates us to study the thermal characteristics during the drilling process. To support this research, a lunar surface vacuum environmental simulation device was designed. According to the requirements of the drilling test, a three-part and dismountable vacuum chamber was designed which was pumped from the bottom of the chamber. Considering the lunar soil simulant's size and rapid vacuum acquirement requirements, a cribriform lunar soil container with multi-layer filtrations was designed. Based on the demand of repeatable drilling tests in the same condition, the slewing mechanism with missing teeth gear was designed which allowed external operation. Finally, a three-stage diffusion pump set was selected according to the vacuum and dust tolerance requirements. Experimental results show that the vacuum of this simulation device reaches to 10^{-1} Pa without lunar soil inside, and reaches to 7 Pa with lunar soil inside. These performance indicators satisfy the vacuum requirements of the drilling test. The outgassing amount curve of simulated lunar soil is obtained by measuring the variation of vacuum with time during the pumpout process, which is valuable for the lunar soil vacuum system's application.

Key words: vacuum device; environmental simulation; drilling test; lunar soil simulant; evacuating method; dynamic sealing

Received: 2015-10-16; **Accepted**: 2015-11-20; **Published online**: 2015-12-17 11:33

URL: www.cnki.net/kcms/detail/11.2625.V.20151217.1133.018.html

* **Corresponding author**. Tel.: 010-82338005 E-mail: xlding@buaa.edu.cn

北京航空航天大学学报

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0682

双向旋转柱面气膜密封性能数值模拟与参数优化

马纲*，杨卫如

（北京航空航天大学 机械工程及自动化学院，北京 100083）

摘 要：在对气膜密封结构及性能分析的基础上，提出了双向旋转倒斜T字型与双层斜槽型柱面气膜密封界面结构；建立了2种结构密封气膜的数理模型和基于有限元方法的密封气膜数值分析模型；在VC平台上开发了基于有限元计算的密封气膜压力分布、系统密封特性及槽型结构多维参数优化的计算分析程序，实现了对密封气膜压力分布、密封特性的数值计算及结构参数的优化设计；在Gambit中建立了不同平均膜厚下倒斜T字型柱面气膜的物理模型和有限元分析模型，通过FLUENT进行了密封气膜特性数值仿真，商用软件仿真结果与自主开发程序数值计算结果对比一致性好，验证了开发程序的正确性；计算对比了2种界面结构的密封稳态特性与动态特性，在相同计算参数下，倒斜T字型的密封性能优于双层斜槽型，为密封系统设计选型提供了帮助；基于遗传思想的粒子群多维优化（GAPSO）计算程序的开发，为槽型参数的优化设计提供了有效的手段。

关 键 词：双向旋转；柱面气膜密封；数值模拟；FLUENT仿真；GAPSO
中图分类号：H117.1
文献标识码：A **文章编号**：1001-5965(2016)11-2279-10

气膜密封是利用气体动、静压原理，通过密封动、静环界面的槽型结构在密封位置形成密封气膜[1]，实现密封界面非接触状态，达到低泄漏、长寿命的密封效果[2]。气膜密封适用于高速旋转场合，广泛应用于鼓风机、压缩机等高速旋转设备中，在航空发动机中的应用也受到广泛关注[3]。工程实践中，单向旋转气膜密封应用广泛，但由于机器启动时阀门被打开或者泄漏，当压缩机及其出口阀之间存在大量气体或者驱动电机被错接时，压缩机可能发生反向旋转运动。反向旋转时，单向螺旋槽无法提供动压开启力而容易引起密封界面的磨损失效等后果[4]。

双向旋转端面气膜密封的研究已经被成功应用。John Carne公司从20世纪80年代后期开始研制双向旋转端面气膜密封，其研制的28BD已成功应用于俄罗斯等国的高速压缩机[5]。德国Burgman公司研发的V型槽结构可在多种密封结构下使用[6]。天津克兰密封有限公司研制的八字型槽气体端面密封在高速透平机上得到了应用[7]。

柱面气膜密封研究开展相对较晚，近年来逐渐引起重视。2001年，Salehi和Heshmat[8]在成功应用的径向柔顺箔轴承的基础上提出并进行了柱面形式的柔顺箔密封研究，其平滑顶箔与高速轴颈之间可形成流体动压密封气膜，并在小型气涡轮发动机模拟器上进行了柔顺箔密封系统的性能实验研究，但其密封界面为光滑圆柱，单向运转。2002年，英国帝国理工大学的Sayma等[9]提出了在大直径航空燃气轮机的转子叶尖处应用气控密封即柱面气膜密封的问题，提出了此密封应

收稿日期：2015-10-20；**录用日期**：2015-12-04；**网络出版时间**：2016-01-21 15:47
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160121.1547.011.html
基金项目：国家自然科学基金（51075018）
*通讯作者：Tel.：010-82317745 E-mail：mgbuaa@163.com

引用格式：*马纲，杨卫如. 双向旋转柱面气膜密封性能数值模拟与参数优化[J]. 北京航空航天大学学报，2016，42(11)：2279-2288. MA G，YANG W R. Numerical simulation and parameter optimization of seal property of bidirectional rotating cylinder gas film[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2279-2288 (in Chinese).*

北航学报 赠阅

有结构柔性和基于 Navier-Stokes 方程的气弹性分析模型,但未提出具体结构型式。2004 年, Shapiro 等[10-11]对光滑密封界面、螺旋槽密封界面形式下柱面气膜密封的分析方法、试验结果进行了深入的研究,但其设计及分析试验均是在单向旋转的条件下进行。国内,笔者研究团队针对柱面气膜从数值分析方法、界面结构、稳态特性及动态特性等方面开展了一系列的研究[1,12],针对的都是单向旋转条件。

柱面气膜密封易具有浮动自适应性,可使密封系统具有较大的柔性,适应系统中振动位移和变形等工况,在高速转子系统密封中具有独特的优势[13-14]。双向旋转柱面气膜密封界面结构能适应双向旋转的工况,避免了只能单向旋转的柱面气膜界面结构的局限,提高了适应范围。本文提出双向旋转的柱面气膜密封界面结构,以倒斜 T 字型和双层斜槽型柱面气膜为研究对象,采用有限元方法在 VC 平台上进行自主程序的开发,开发了气膜压力及稳态特性数值计算程序,并计算分析了动态特性。采用基于遗传思想的粒子群多维优化(GAPSO)算法对槽型结构进行多维参数优化,实现气膜密封系统优良密封特性。

1 理论分析

1.1 密封气膜数理模型

柱面气膜密封简化的密封模型如图 1 所示。图中:θ 为柱坐标 θ 向坐标维;z 为柱坐标 z 向坐标维;ω 为角速度;r 为旋转坐标系径向;t 为时间;ϕ 为偏位角;O_b 为密封环几何中心;O_j 为动环几何中心;e 为偏心距。密封系统主要包括密封动环(转子)、密封气膜、密封环(浮环)和支撑部分。密封环通过柔性支撑连接在固定结构(箱体或机匣)上,其内表面设计特殊的槽型结构,工作过程中通过转子高速旋转使得转子轴颈和密封环之间产生动压密封气膜,避免了密封副表面之间的直接接触,实现了良好的密封效果。本文提出 2 种适应双向旋转的柱面密封结构,图 2 为倒斜 T 字型和双层斜槽型界面结构的展开图。

描述柱面密封气膜压力,双柱面旋转较广义的雷诺方程为[12]

$$\frac{1}{R_j^2}\cdot\frac{\partial}{\partial\theta}\left(\frac{\rho h^3}{\eta}\cdot\frac{\partial p}{\partial\theta}\right)+\frac{\partial}{\partial z}\left(\frac{\rho h^3}{\eta}\cdot\frac{\partial p}{\partial z}\right)=6\omega\frac{\partial(\rho h)}{\partial\theta}+12\frac{\partial(\rho h)}{\partial t} \tag{1}$$

式中:h 为膜厚;p 为气膜压力;R_j 为转子半径;η 为气膜黏度;ρ 为气膜密度。

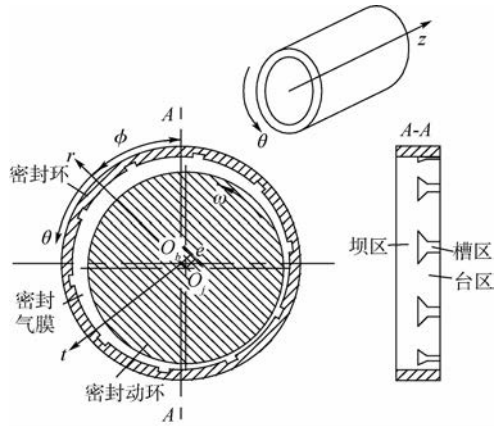


图 1 坐标系中的简化密封模型及几何关系

Fig. 1 Simplified seal model and geometric relations in coordinate system

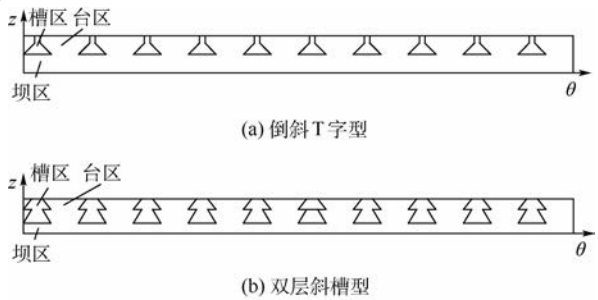


图 2 倒斜 T 字型和双层斜槽型界面展开图

Fig. 2 Interface expansion plan of inverted T type and double-groove type

1.2 柱面气膜稳态压力微分方程

在旋转坐标系中,不带节流孔的二维柱面坐标稳态压力场方程可由式(1)得到[12]:

$$\frac{1}{R_j^2}\cdot\frac{\partial}{\partial\theta}\left(\frac{\rho h^3}{\eta}\cdot\frac{\partial p}{\partial\theta}\right)+\frac{\partial}{\partial z}\left(\frac{\rho h^3}{\eta}\cdot\frac{\partial p}{\partial z}\right)=6\omega\frac{\partial(\rho h)}{\partial\theta} \tag{2}$$

采用无量纲化后的雷诺方程为

$$\frac{\partial}{\partial\bar{\theta}}\left(\bar{h}^3\frac{\partial\bar{p}^2}{\partial\bar{\theta}}\right)+\left(\frac{R_j}{L}\right)^2\frac{\partial}{\partial\bar{z}}\left(\bar{h}^3\frac{\partial\bar{p}^2}{\partial\bar{z}}\right)=2\Lambda\frac{\partial(\bar{p}\,\bar{h})}{\partial\bar{\theta}} \tag{3}$$

式中:Λ 为可压缩数,$\Lambda=\dfrac{6\omega\eta R_j^2}{p_a c^2}$表征气膜密封结构和运转状况的综合,$p_a$ 为环境压力,c 为平均膜厚;L 为密封环轴向宽度。

稳态膜厚 h 为

$$h_0(r,\theta)=c+e\cos\theta+\delta_j h_g(r,\theta) \tag{4}$$

式中:δ_j 为 Kronecker 数,在有凹凸垫或槽处为 1,无凹凸垫或槽处为 0;$h_g(r,\theta)$ 为界面凹凸垫或槽引起的稳态膜厚变化(当界面上有槽处,表示增大了稳态膜厚,相应增加量为 $h_g(r,\theta)$)。

求解气膜控制方程采用对复杂结构适应性较强的有限元方法。相关有限元分析模型的建立、数值计算过程和数值算法在文献[12]中已介绍,篇幅所限,且思路相同,本文只进行简要叙述。

北航学报 赠阅

2　有限元方法数值计算

2.1　伽辽金法建立变分方程

采用有限元方法进行压力场的数值计算分析。基于微分方程等效积分提法的加权余量思想，使用伽辽金法对柱面气膜压力微分方程改造得到其等效积分形式，将式(3)用伽辽金法建立变分方程：

$$\iint_D \left[\frac{\partial}{\partial \bar{\theta}}\left(\bar{h}^3 \frac{\partial \bar{p}^2}{\partial \bar{\theta}} \right) + \left(\frac{R_j}{L} \right)^2 \frac{\partial}{\partial \bar{z}}\left(\bar{h}^3 \frac{\partial \bar{p}^2}{\partial \bar{z}} \right) - 2\Lambda \frac{\partial (\bar{p}\bar{h})}{\partial \bar{\theta}} \right] \delta \bar{p} \mathrm{d}\theta \mathrm{d}\bar{z} = 0 \tag{5}$$

2.2　变分方程的离散求解

1）解域离散

在解域上进行有限元分割时注意：① 利用周期性将展开的平面代替圆柱解域；② 槽台界限落在单元网格边界上；③ 节点编号顺序，轴向(宽度)优先；④ 分区域划分网格，不同区域设置不同的网格参数，以利于提高计算效率。图 3 为以倒斜 T 字型界面结构为例的网格划分结果。

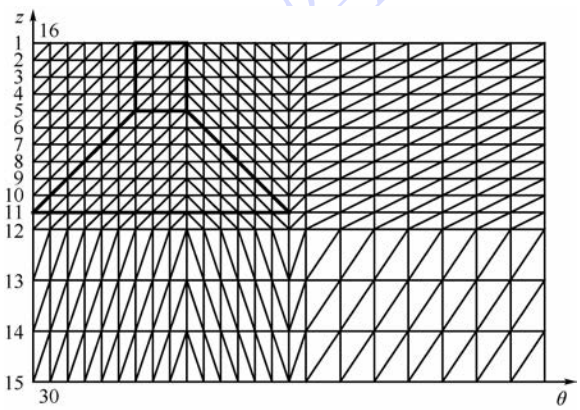


图 3　倒斜 T 字型网格划分与节点编号

Fig. 3　Mesh generation and node number of inverted T type

2）方程离散求解

使用伽辽金法对柱面气膜压力微分方程改造得到其等效积分形式，进而得到更利于使用的弱形式而建立起单元刚度矩阵，组集为整体压力场方程组，引入对应的边界条件，确定压力场分布的数值解(具体求解过程见文献[12])。

2.3　气膜的密封特性

1）气膜的稳态密封特性[12]

气膜反力：

$$\begin{cases} F_r = \int_0^l \int_0^{2\pi} (p - p_a) R_j \cos\theta \mathrm{d}\theta \mathrm{d}z \\ F_t = \int_0^l \int_0^{2\pi} (p - p_a) R_j \sin\theta \mathrm{d}\theta \mathrm{d}z \\ F = \sqrt{F_r^2 + F_t^2} \end{cases} \tag{6}$$

式中：l 为柱面密封长度。

无量纲形式：

$$\bar{F} = \frac{F}{p_a R_j L} \tag{7}$$

密封泄漏量：

$$Q = 2\int_0^{2\pi} \rho \left(-\frac{h^3}{12\eta} \cdot \frac{\partial p}{\partial z} \right) R_j \mathrm{d}\theta \mathrm{d}z \tag{8}$$

无量纲形式：

$$\bar{Q} = -24 \frac{\eta RTL}{p_a^2 c^3 R_j} Q \tag{9}$$

式中：T 为绝对温度；R 为气体常数。

摩擦转矩：

$$M = \int_0^l \int_0^{2\pi} \left(\frac{\partial p}{R_j \partial \theta} \cdot \frac{h}{2} + \eta \frac{\omega R_j}{h} \right) R_j^2 \mathrm{d}\theta \mathrm{d}z \tag{10}$$

无量纲形式：

$$\bar{M} = \frac{M}{p_a R_j L c} \tag{11}$$

气膜刚度系数：

$$K_{zz} = -\left. \frac{\partial F_z}{\partial h} \right|_{h_0} = -\frac{F_2 - F_1}{h_2 - h_1} \tag{12}$$

式中：F_2 为气膜反力 2；F_1 为气膜反力 1；h_2 为偏心距 2；h_1 为偏心距 1。

无量纲形式：

$$\bar{K}_{zz} = -\frac{\bar{F}_2 - \bar{F}_1}{\bar{H}_2 - \bar{H}_1} \tag{13}$$

2）气膜的动态密封特性

在稳态特性计算的基础上，进行动态特性分析。动态刚度系数与动态阻尼系数为[12]

$$\begin{cases} K_{rr} = -\iint_D p_r \cos\theta R_j \mathrm{d}\theta \mathrm{d}z \\ K_{rt} = -\iint_D p_t \cos\theta R_j \mathrm{d}\theta \mathrm{d}z \\ D_{rr} = -\iint_D p_r \cos\theta R_j \mathrm{d}\theta \mathrm{d}z \\ D_{rt} = -\iint_D p_t \cos\theta R_j \mathrm{d}\theta \mathrm{d}z \end{cases} \tag{14}$$

$$\begin{cases} K_{tr} = -\iint_D p_r \sin\theta R_j \mathrm{d}\theta \mathrm{d}z \\ K_{tt} = -\iint_D p_t \sin\theta R_j \mathrm{d}\theta \mathrm{d}z \\ D_{tr} = -\iint_D p_r \sin\theta R_j \mathrm{d}\theta \mathrm{d}z \\ D_{tt} = -\iint_D p_t \sin\theta R_j \mathrm{d}\theta \mathrm{d}z \end{cases} \tag{15}$$

式中：p_r、p_t 为密封气膜对应单位扰动位移的微扰压力。

以有限元分析为基础，在 VC 平台上自主开发数值计算分析程序，进行气膜稳态特性和动态特性系数的计算。

北航学报 赠阅

3 仿真计算及算例验证

3.1 仿真分析方法

选择倒斜 T 字型结构进行仿真计算。

图 4 为在 Gambit 软件中建立的倒斜 T 字型柱面密封气膜的数值模型。

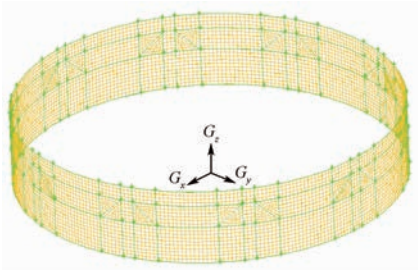


图 4 Gambit 建立模型

Fig. 4 Model in Gambit

将数值分析模型导入 FLUENT,进行气膜压力场的仿真求解计算[15]。

3.2 仿真结果

计算参数如表 1 时,仿真结果如图 5 所示。

表 1 计算参数

Table 1 Calculated parameters

参数	数值
转子半径 R_j/mm	50
密封环轴向宽度 L/mm	20
平均膜厚 c/μm	10
槽数	10
偏心率	0.5
1 槽宽比 pb_1	0.1
1 槽长比 pl_1	0.2
2 槽长比 pl_2	0.3
1 槽深比 pd_1	2.0
2 槽深比 pd_2	2.0
2 槽宽比 pb_2	0.5
上游压力 P_h/ MPa	0.3
下游压力 P_l/MPa	0.1
角速度 ω/(r · min^{-1})	20 000
绝对温度 T / ℃	20
密封气体	空气

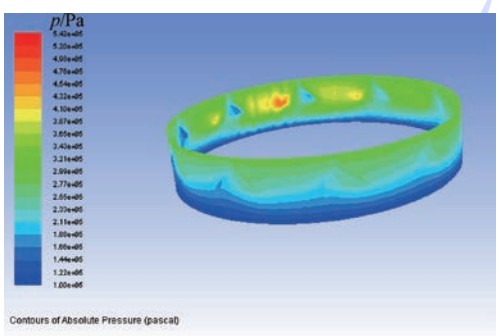


图 5 总压云图

Fig. 5 Total pressure contours

3.3 对比验证

进行不同膜厚及转速下 VC 平台上自主开发程序结果与仿真结果的对比验证。图 6 为倒斜 T 字型平均膜厚为 10 μm 和 6 μm 时,气膜反力随转速的变化情况。程序结果与仿真结果相当接近,变化趋势也一致,表明了本文程序倒斜 T 字型槽柱面气膜密封系统稳态特性的计算分析可信。

将膜厚最小处槽、台、坝交界点标记为点 1,与其处于同一轴线并且与高压端边界相邻的点标记为点 2。进行 VC 平台上自主开发程序结果与仿真结果对比与误差计算,误差在可接受范围以内,说明了程序的可行性,对比结果如表 2 所示。

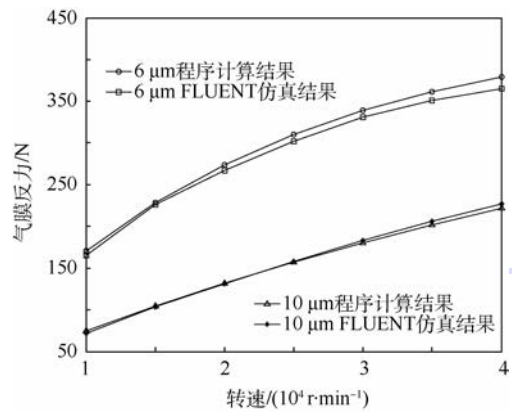


图 6 气膜反力程序结果与仿真结果对比

Fig. 6 Program results of gas film counter-force compared with simulation results

表 2 代表性节点压力值比较及误差计算

Table 2 Pressure comparison and error calculation of representative nodes

转速/(r · min^{-1})	膜厚/μm	点	节点压力/MPa		误差/%
			程序	仿真	
15 000	10	1	0.41	0.43	4.88
15 000	10	2	0.33	0.34	3.03
15 000	6	1	0.60	0.63	5.00
15 000	6	2	0.36	0.35	2.78
25 000	10	1	0.49	0.51	4.08
25 000	10	2	0.35	0.36	2.87
25 000	6	1	0.73	0.76	4.11
25 000	6	2	0.42	0.43	2.38

4 2 种界面结构气膜密封性能比较

进行倒斜 T 字型与双层斜槽型柱面气膜密封系统的稳态特性比较时,计算参数如表 3 所示。

北航学报 赠阅

表 3　稳态特性比较的计算参数

Table 3　Calculated parameters of comparison of steady state characteristics

参数	数值
转子半径 R_j/mm	50
密封环轴向宽度 L/mm	20
平均厚膜 c/μm	10
槽数	10
偏心率	0.4
1 槽宽比 pb_1	0.2
1 槽长比 pl_1	0.3
2 槽长比 pl_2	0.4
1 槽深比 pd_1	2.0
2 槽深比 pd_2	2.0
2 槽宽比 pb_2	0.5
上游压力 P_h/ MPa	0.601 325
下游压力 P_l/ MPa	0.101 325
角速度 ω/(r·min^{-1})	20 000
绝对温度 T / ℃	20

4.1　稳态特性的比较

1）气膜稳态特性随转速的变化

图 7 为倒斜 T 字型与双层斜槽型界面结构的气膜反力、泄漏量、摩擦转矩和静态刚度系数随转速变化的曲线。

由图 7 的计算结果可以看出,随着转速的增大,2 种界面结构的气膜反力逐渐增大,泄漏量先减小后增大,摩擦转矩逐渐增大,静态刚度系数逐渐增大;相同条件下,气膜反力和静态刚度系数为倒斜 T 字型大于双层斜槽型,泄漏量和摩擦转矩为倒斜 T 字型小于双层斜槽型。

2）气膜稳态特性随平均膜厚的变化

平均膜厚是一个重要的设计参数。图 8 为倒斜 T 字型与双层斜槽型界面结构的气膜反力、泄漏量、摩擦转矩和静态刚度系数随平均膜厚变化的曲线。

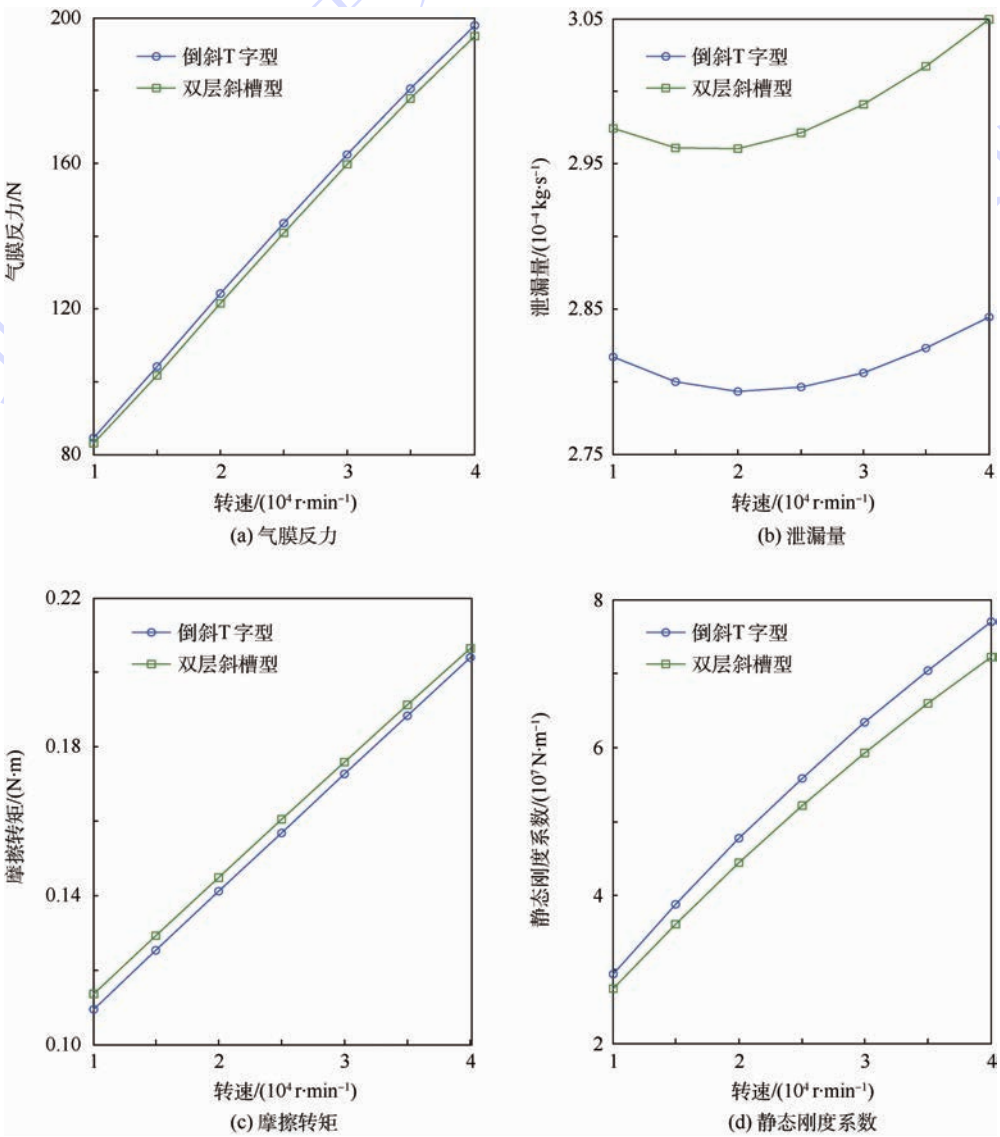


图 7　不同转速下稳态特性的比较

Fig. 7　Comparison of steady state characteristics under different rotational speeds

北航学报 赠阅

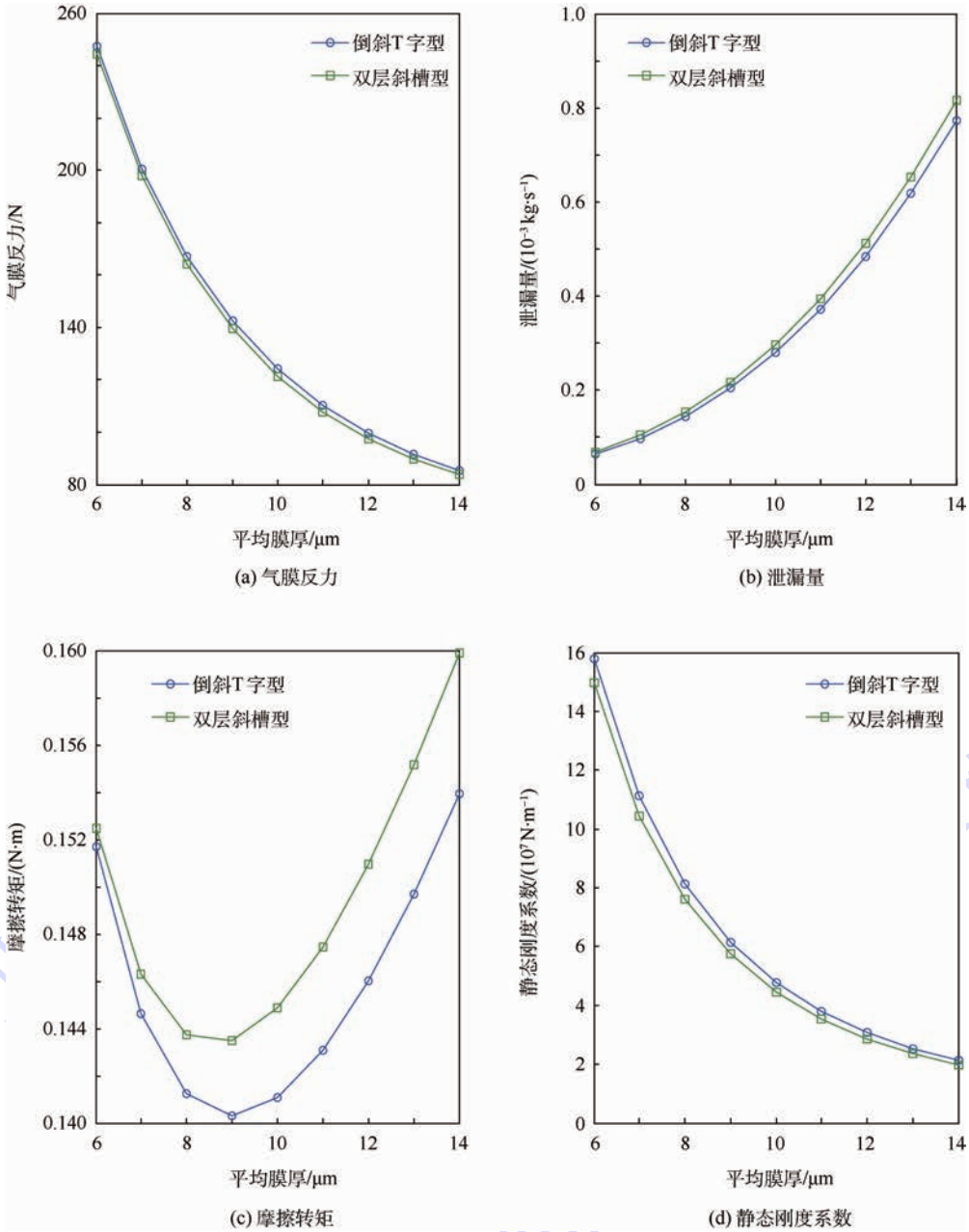


图 8 不同平均膜厚下稳态特性的比较

Fig. 8 Comparison of steady state characteristics under different average film thickness

由图 8 可以看出,随着平均膜厚增大,2 种界面结构的气膜反力逐渐减小,泄漏量逐渐增大,摩擦转矩先减小后增大,静态刚度系数逐渐增小;相同条件下,气膜反力和静态刚度系数为倒斜 T 字型大于双层斜槽型,泄漏量和摩擦转矩为倒斜 T 字型小于双层斜槽型。

4.2 动态特性的比较

1)气膜动态刚度系数随转速的变化

图 9 为 2 种界面结构动态刚度系数随转速变化的曲线。随着转速的增大,正向刚度系数 K_{rr}、K_{tt} 与交叉刚度系数 K_{rt} 逐渐增大,交叉刚度系数 K_{tr} 逐渐减小。相同参数下,K_{rr}、K_{rt} 为倒斜 T 字型大于双层斜槽型,K_{tr}、K_{tt} 为倒斜 T 字型小于双层斜槽型。

2)气膜动态阻尼系数随转速的变化

图 10 为 2 种界面结构的动态阻尼系数随转速变化的曲线。随着转速的增大,正向阻尼系数 D_{rr}、D_{tt} 与交叉阻尼系数 D_{rt} 逐渐减小,交叉阻尼系数 D_{tr} 先增大后减小。相同参数下,D_{rr}、D_{tt}、D_{tr} 为倒斜 T 字型大于双层斜槽型,D_{rt} 为倒斜 T 字型小于双层斜槽型。

北航学报 赠阅

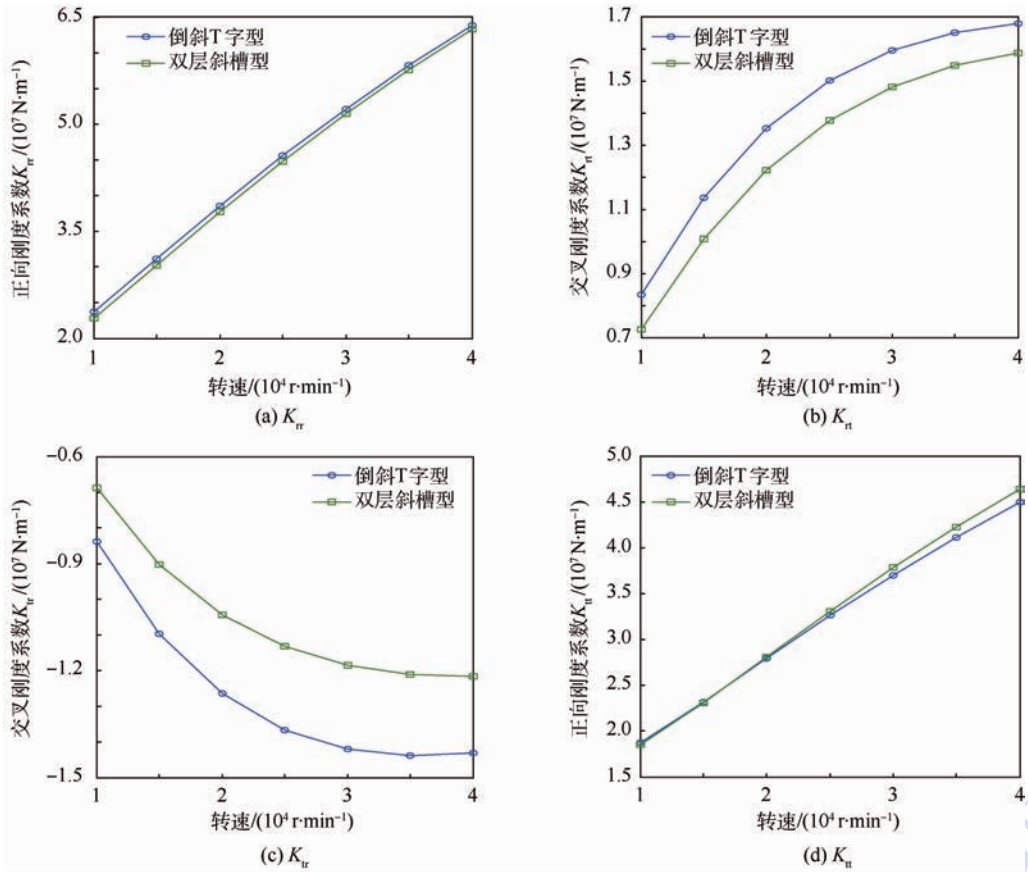


图 9　不同转速下气膜动态刚度系数比较

Fig. 9　Comparison of dynamic stiffness coefficients of gas film under different rotational speeds

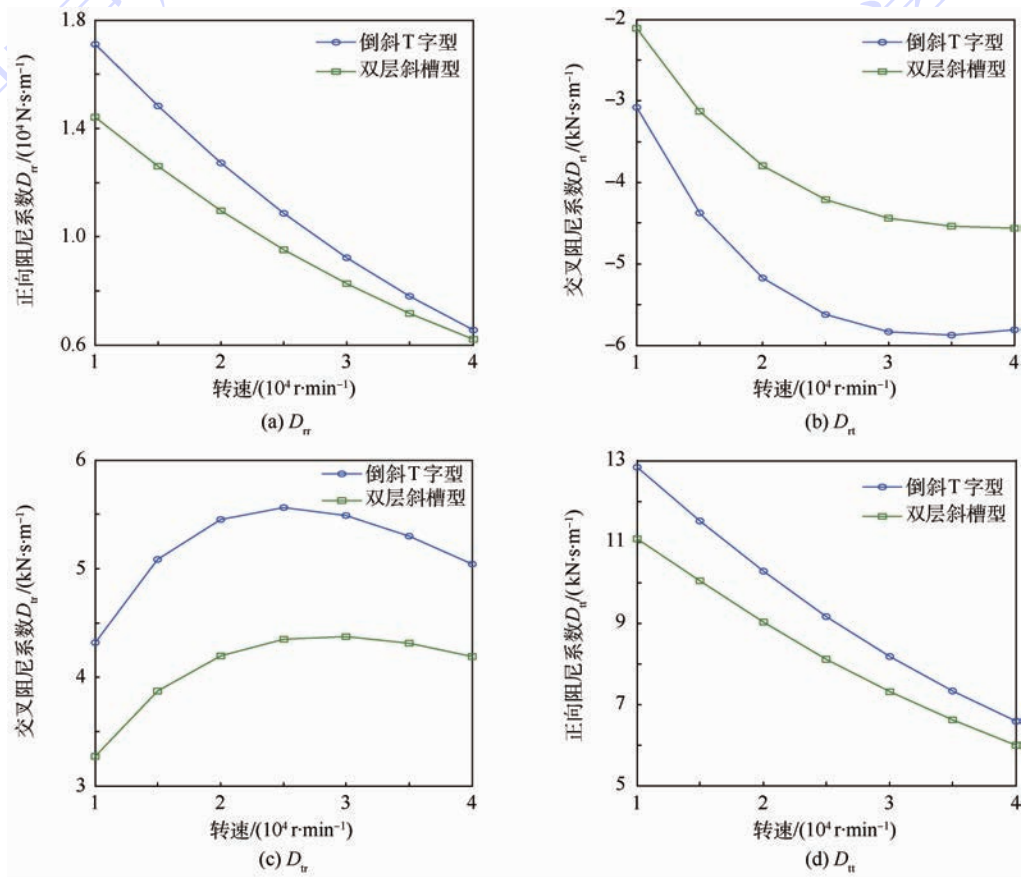


图 10　不同转速下气膜动态阻尼系数比较

Fig. 10　Comparison of dynamic damping coefficients of gas film under different rotational speeds

北航学报 赠阅

5 GAPSO 算法多维参数优化

5.1 粒子群优化算法与遗传算法结合

气膜密封系统优良密封特性的实现,首先取决于密封界面的结构,发挥特定界面结构型式优势,需要综合考虑密封界面几何参数。描述槽型几何特征的参数众多,而且具有相关性,气膜密封设计的参数优化是一个非线性、多参数和多目标优化的问题。

利用粒子群优化算法与遗传算法结合,实现柱面气膜密封界面槽形参数的多维优化。本文将各槽型参数 pb_1、pl_1、pl_2、pd_1、pb_2 作为设计变量,对密封界面结构参数进行多维优化。

建立 GAPSO 的优化模型:

$$\boldsymbol{X}_i = [pb_1, pl_1, pl_2, pd_1, pb_2]_i^{\mathrm{T}}$$

$$\max \quad F(\boldsymbol{X})$$

$$\text{s.t.} \quad \boldsymbol{X}_{\min} < \boldsymbol{X} < \boldsymbol{X}_{\max}$$

2 种目标函数 $F(\boldsymbol{X})$:① 气膜反力;② 气膜反力与泄漏量的比值。

GAPSO 算法流程图如图 11 所示(详细过程见文献[16-17])。

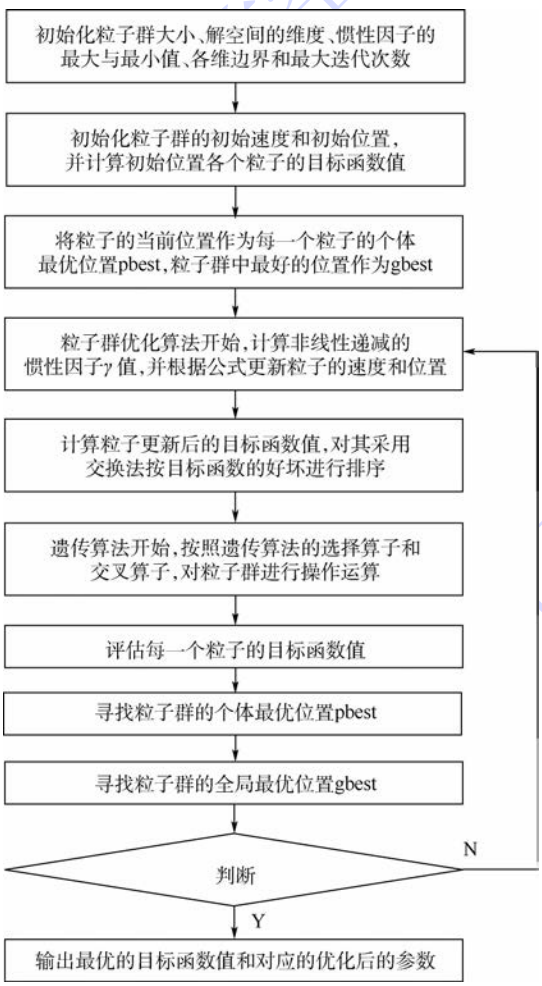


图 11 GAPSO 算法流程图

Fig. 11 Flowchart of GAPSO algorithm

5.2 界面结构参数有无优化的对比

转速不同时利用 GAPSO 算法对倒斜 T 字型进行多维优化,优化目标选为气膜反力,结果见图 12。在不同转速下,界面结构参数优化结果不同,表 4 为转速为 2×10^4 r/min 时的优化结果。

平均膜厚不同时,利用 GAPSO 算法对双层斜槽型进行多维优化,优化目标选为气膜反力与泄漏量的比值,结果见图 13。在不同平均膜厚下,界面结构参数优化结果不同,表 5 为平均膜厚为 9 μm 时的优化结果。

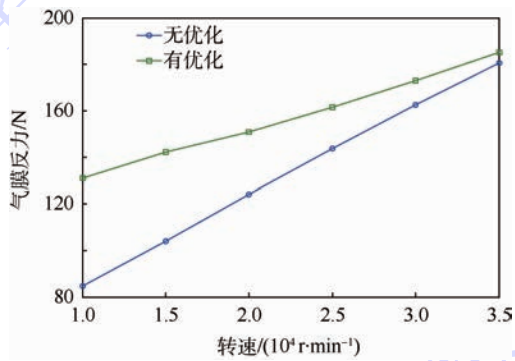


图 12 气膜反力优化对比

Fig. 12 Comparison of gas film counter-force optimization

表 4 转速为 2×10^4 r/min 时的优化结果

Table 4 Rotating speed at 2×10^4 r/min optimization results

参数	初始值	优化结果
1 槽宽比 pb_1	0.2	0.176 706
1 槽长比 pl_1	0.3	0.449 489
2 槽长比 pl_2	0.4	0.448 237
1 槽深比 pd_1	2.0	3.066 440
2 槽宽比 pb_2	0.5	0.930 887
目标函数/N	124.121	150.857

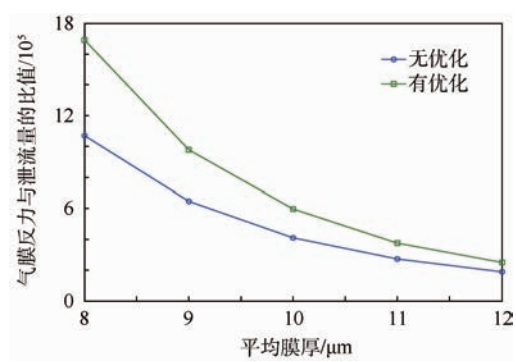


图 13 气膜反力与泄漏量的比值优化对比

Fig. 13 Comparison of ratio optimization of gas film counter-force and leakage

北航学报 赠阅

表 5　平均膜厚为 9 μm 时的优化结果

Table 5　Average film thickness at 9 μm optimization results

参数	初始值	优化结果
1 槽宽比 pb_1	0.2	0.151 854
1 槽长比 pl_1	0.3	0.419 608
2 槽长比 pl_2	0.4	0.263 190
1 槽深比 pd_1	2.0	1.150 060
2 槽宽比 pb_2	0.5	0.500 367
目标函数	645 031.54	979 997.00

6　结　论

1) 建立了双向旋转倒斜 T 字型与双层斜槽型柱面气膜密封系统密封气膜的数理模型和基于有限元方法的密封气膜数值分析模型。

2) 在 VC 平台上开发了基于有限元计算的密封气膜压力分布、密封系统稳态特性与动态特性的计算分析程序。动态特性数值计算的实现突破了目前商业软件在进行动态特性分析的局限。

3) 在 Gambit 中建立了不同平均膜厚下倒斜 T 字型柱面气膜的物理模型和有限元分析模型,通过 FLUENT 进行了密封气膜特性数值仿真,商用软件仿真结果与自主开发程序数值计算结果对比一致性好,验证了开发程序的正确性。

4) 计算分析了转速与平均膜厚对密封特性的影响,对比了相同参数下 2 种界面结构的密封特性。从稳态特性和动态特性看,倒斜 T 字型的性能均优于双层斜槽型。

5) 计算对比了 2 种界面结构有无多维优化的密封特性。通过 GAPSO 算法进行多维优化可以有效提高密封特性。

参考文献 (References)

[1] 马纲,栗秀花,沈心敏,等. 柱面气膜密封界面结构与性能分析[J]. 航空动力学报,2011,26(11):10-16.
MA G,LI X H,SHEN X M,et al. Analysis of performance and interface structure of cylinder gas film seal[J]. Journal of Aerospace Power,2011,26(11):10-16(in Chinese).

[2] KOWALSKI C A,BASU P. Reverse rotation capability of spiral-groove gas face seals[J]. Tribology Transactions,1995,38(3):549-556.

[3] YANG P,ZHU K Q,WANG X L. On the non-linear stability of self-acting gas journal bearings[J]. Tribology International,2009,42(1):71-76.

[4] NOSOWICZ J,ZEUS D. A bi-directional gas face seal experience under test conditions and in practical use[J]. Lubrication Engineering,1993,49(3):217-221.

[5] GOLD SWAIN I M. Mechanical face seals:EP,US 5375855 A[P]. 1994-12-27.

[6] 赵亚萍. 双向干气密封的研制情况及技术分析[J]. 流体机械,1996,24(4):46-47.
ZHAO Y P. Development situation of two way dry gas seal and technical analysis[J]. Journal of Fluid Mechanics,1996,24(4):46-47(in Chinese).

[7] 郝木明. 机械密封技术及应用[M]. 北京:中国石化出版社,2010:80.
HAO M M. Mechanical seal technology and its application[M]. Beijing:China Petrochemical Press,2010:80(in Chinese).

[8] SALEHI M,HESHMAT H. Performance of a complaint foil seals in a small gas turbine engine simulator employing a hybrid foil/ball bearing support system[J]. Tribology Transactions,2001,44(3):458-464.

[9] SAYMA A I,BREARD C,VAHDATI M,et al. Aerolasticity analysis of air-riding seals for aero-engine applications[J]. Journal of Tribology-Transactions of the ASME,2002,124(3):607-616.

[10] SHAPIRO W. Numerical,analytical,experimental study of fluid dynamic forces in seals:NASA/CR 2004-213199/ VOL2[R]. Latham,NY:Mechanical Technology,Inc.,2004.

[11] WALOWIT J,SHAPIRO W. Numerical,analytical,experimental study of fluid dynamic forces in seals:NASA/CR2004-213199/VOL3[R]. Latham,NY:Mechanical Technology,Inc.,2004.

[12] 马纲,席平,沈心敏,等. 柔性支承浮环柱面气膜密封准动态特性分析[J]. 航空动力学报,2010,25(5):1190-1196.
MA G,XI P,SHEN X M,et al. Analysis of quasi-dynamic characteristics of compliant floating ring gas cylinder seal[J]. Journal of Aerospace Power,2010,25(5):1190-1196(in Chinese).

[13] LIU Y C,SHEN X M,XU W F. Numerical analysis of dynamic coefficients for gas film face seals[J]. Journal of Tribology-Transaction of the ASME,2002,124(4):743-754.

[14] GREEN I,BARNSBY R M. A simultaneous numerical solution for the lubrication and dynamic stability of noncontacting gas face seals[J]. Journal of Tribology-Transactions of the ASME,2001,123(2):388-394.

[15] 马纲,孙晓军,罗先海,等. 端柱面组合密封气膜稳态特性仿真数值分析[J]. 北京航空航天大学学报,2014,40(4):439-443.
MA G,SUN X J,LUO X H,et al. Simulation numerical analysis of steady-state properties of gas face and cylinder film seal[J]. Journal of Beijing University of Aeronautics and Astronautics,2014,40(4):439-443(in Chinese).

[16] 陈贵敏,贾建援,韩琪. 粒子群优化算法的惯性权值递减策略研究研究[M]. 西安:西安交通大学出版社,2002:70.
CHEN G M,JIA J Y,HAN Q. Research on the strategy of decreasing inertia weight in particle swarm optimization algorithm[M]. Xi'an:Xi'an Jiaotong University Press,2002:70(in Chinese).

[17] 曹春红,张永坚,李文辉. 杂交粒子群算法在工程几何约束求解中的应用[J]. 仪器仪表学报,2004,25(4):397-400.
CAO C H,ZHANG Y J,LI W H. Applications of hybrid particle swarm optimization in engineering geometric constraint solving[J]. Chinese Journal of Scientific Instrument,2004,25(4):397-400(in Chinese).

北航学报 赠阅

作者简介：
马纲 女，博士，副教授，硕士生导师。主要研究方向：流体润滑密封技术及机械设计。
Tel.：010-82317745
E-mail：mgbuaa@163.com

Numerical simulation and parameter optimization of seal property of bidirectional rotating cylinder gas film

MA Gang*, YANG Weiru

(School of Mechanical Engineering and Automation, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: Based on cognizing of advantages of cylinder gas film seal system, this paper proposes a new bidirectional rotating cylinder gas film seal system with inverted T interface structure and double-groove interface structure. It establishes mathematical model and numerical analysis model of the bidirectional rotating cylinder gas film with inverted T interface structure and double-groove interface structure based on finite element method. It develops calculation program about pressure distribution, steady and dynamic property of the gas film, and multidimensional parameter optimization of groove structure on the basis of finite element method on the VC platform. It realizes calculation and analysis of pressure distribution, property of the seal system and structural parameter optimization design. It builds physical models and finite element models of the inverted T cylinder gas film with different average thickness in Gambit. And then, this paper carries out numerical simulation of the gas film under different rotational speed through FLUENT. The simulation results of commercial software are consistent with numerical results of the self-developed program, verifying the correctness of the developed program. It calculates and compares steady and dynamic properties of two kinds of interface structures, and seal performance of inverted T type is better than double-groove type under the same design conditions, which provides reference for selection model of seal system. The calculation program of genetic algorithm-particle swarm optimization (GAPSO) has guiding significance for the design of groove parameters.

Key words: bidirectional rotating; cylinder gas film seal; numerical simulation; FLUENT simulation; GAPSO

Received: 2015-10-20; **Accepted**: 2015-12-04; **Published online**: 2016-01-21 15:47
URL: www.cnki.net/kcms/detail/11.2625.V.20160121.1547.011.html
Foundation item: National Natural Science Foundation of China (51075018)
* **Corresponding author**. Tel.: 010-82317745 E-mail: mgbuaa@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0689

四轮驱动汽车牵引力控制系统控制策略

姜立标[1,*], 丘华川[1], 吴中伟[1], 刘巍[2], 程铖[1]
(1. 华南理工大学 机械与汽车工程学院, 广州 510640; 2. 浙江吉利汽车研究院有限公司, 杭州 311228)

摘 要: 针对四轮驱动汽车整车动力性和横向稳定性的问题,提出新的目标滑转率计算和修正方法的牵引力控制系统综合控制策略。首先,建立车辆系统各主要部件的数学模型;然后,针对不同工况,利用模糊 PID 控制技术建立以发动机节气门开度控制、轴间扭矩分配控制和驱动轮制动控制为一体的牵引力控制系统综合控制策略;在车轮目标滑转率的选取方面,提出了最优纵向滑转率查表法,并根据车速以及轮胎侧偏角的大小对目标滑转率进行实时修正;最后,在典型工况下使用 MATLAB/Simulink 对四轮驱动汽车牵引力控制系统进行离线仿真研究。结果表明:本文所提出的牵引力控制系统控制策略能有效地抑制驱动轮的过度滑转,充分利用路面附着条件,提高汽车纵向驱动稳定性。

关 键 词: 四轮驱动汽车; 牵引力控制; 目标滑转率; 控制策略; 动力学仿真
中图分类号: U461.1
文献标识码: A **文章编号**: 1001-5965(2016)11-2289-10

四轮驱动汽车可以根据车辆的行驶工况及路面附着条件的不同全时将发动机的输出转矩传递至前后轴,提高汽车的纵向稳定性,有利于抵抗横向滑转[1]。由于传统汽车在不同工况下行驶时,车辆驱动力矩突然增大或路面附着系数减小时,车轮发生滑转,导致车辆纵向操纵性能及转向变差,同时降低横向稳定性,影响车辆正常行驶[2]。采用全时四轮独立驱动,如何充分利用路面附着条件来保证车辆获得尽可能大的牵引力,同时保证转向平稳和侧向稳定性是控制策略[3-5]设计的关键。

文献[2]结合轴间差速器扭矩自适应原则和滑模控制理论设计了一种驱动防滑系统控制策略,但需要分模式控制。文献[6]提出基于车轮加速度和滑转率的逻辑门限值防滑控制,由于门限值参数因路面附着系数不同而有所不同,只有在工况相近时系统稳定性较好,并且门限值参数需通过反复试验获得。文献[7]采用最优扭矩分配控制策略,考虑到不同性能指标(车辆质心的力及力矩误差、执行器控制力度和轮胎使用负载)之间的干扰,提出了能够在多功能方式下工作的新目标函数,为适应不同的驱动工况对各因素进行加权调整。文献[8]提出最优扭矩矢量算法,以横摆角力矩和滑转率建立了一种优化控制分配方法。文献[9]基于角加速度和滑转率设计一种模糊逻辑控制策略,为适应电机动态转矩输出,需保证车轮滑转率控制在最优范围内。文献[10]采用增量 PI 控制方法,提出了发动机节气门开度控制和轴间扭矩分配控制的综合控制策略,其判定标准为汽车行驶 30 m 所需的时间,本文采用 5 s 直线加速所能达到的纵向车速作为牵引力控制的有效性判定依据。

本文建立了四轮驱动汽车的数学模型,采用 GIM 理论轮胎模型[11],基于轮胎法向力和轮胎侧

收稿日期: 2015-10-23; **录用日期**: 2015-12-04; **网络出版时间**: 2015-12-09 15:32
网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151209.1532.002.html
基金项目: 国家自然科学基金(50975091); 广东省自然科学基金(9451064101003049)
* **通讯作者**: E-mail: jlb620620@163.com

引用格式: 姜立标, 丘华川, 吴中伟, 等. 四轮驱动汽车牵引力控制系统控制策略[J]. 北京航空航天大学学报, 2016, 42(11): 2289-2298. JIANG L B, QIU H C, WU Z W, et al. Control strategy of traction control system for four-wheel drive vehicle[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(11): 2289-2298 (in Chinese).

北航学报 赠阅

偏角提出最优纵向滑转率查表法,建立以节气门开度控制、轴间扭矩分配控制和驱动轮制动控制[8,12-13]为一体的牵引力控制系统控制策略,结合模糊控制和 PID 控制[14-16]的优点,分别在低附着路面、高-低附着对接路面及分离路面 3 种工况下进行离线仿真分析。

1 仿真软件的整体结构

选择某汽车集团的四轮驱动汽车为研究对象。该车配置:电控汽油发动机、手动机械变速器、轴间扭矩分配器、普通主减速器、差速器及液压制动系统。建立四轮驱动汽车上各驱动系统的数学模型[17-20],如:发动机模型、传动系模型、轴间扭矩分配器模型、制动器模型、轮胎模型、整车模型、车轮动载荷计算模块、轮胎动坐标系速度计算模块、车轮中心侧偏角计算模块、发动机转速计算模块和路面附着输入模块等。四轮驱动汽车加速动力学仿真软件的主体结构如图 1 所示。

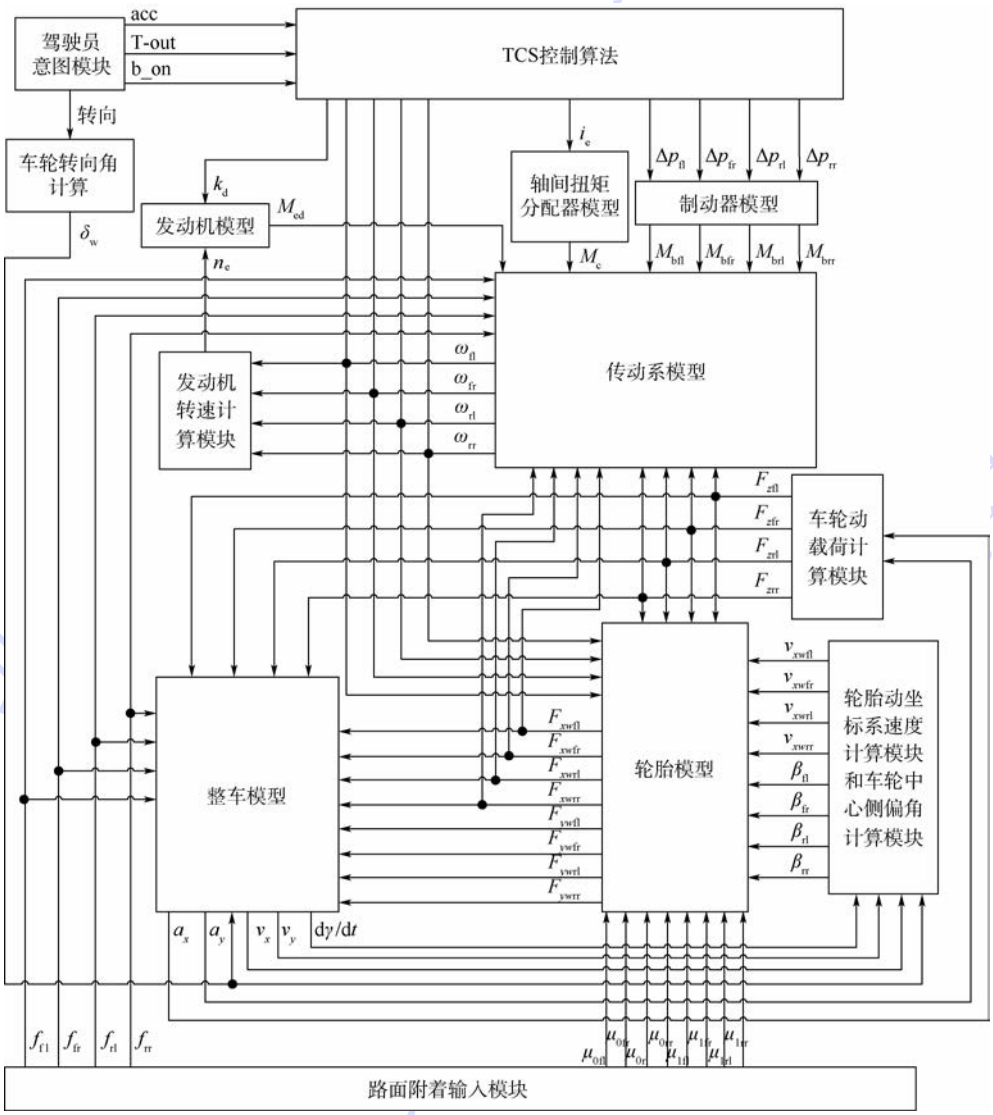


图 1 四轮驱动汽车加速动力学仿真软件的主体结构

Fig. 1 Four-wheel drive vehicle's main structure of dynamics of accelerated motion simulation software

1.1 发动机模型

发动机在汽车起动过程中以及正常行驶的工况下均属于非稳定状态,其动态响应过程由简化的一阶惯性滞后模型来表示。发动机动态输出转矩为

$$M_{ed} = M_{es}\frac{e^{-st_{e1}}}{1+st_{e2}} \tag{1}$$

式中:M_{ed}为发动机动态输出转矩;M_{es}为发动机稳态输出转矩;s 为拉氏变换变量;t_{e1}和 t_{e2}分别为系统时间常数和滞后时间常数。

1.2 整车模型

整车模型复杂程度直接反映整车的驱动性能和稳定性,影响仿真过程中模型的计算准确度和高效性。根据本文所研究的四轮驱动汽车牵引力控制系统的直接需求以及典型工况下的具体控制变量,建立图 2 所示的四轮驱动八自由度的整车模型。其中,自由度分别为整车纵向速度 v_x、侧向

速度 v_y、横摆角速度 γ、4 个车轮的转速 ω 和前轮转向角 δ_w。表 1 为仿真模型的主要参数。数学表达式如下:

$$m(\dot{v}_x - v_y\gamma) = F_x \tag{2}$$

$$m(\dot{v}_y + v_x\gamma) = F_y \tag{3}$$

$$J_z\dot{\gamma} = M_z \tag{4}$$

$$F_x = F'_{xfl} + F'_{xfr} + F_{xrl} + F_{xrr} - \frac{C_D A v_x^2}{21.15} \tag{5}$$

$$F_y = F'_{yfl} + F'_{yfr} + F_{yrl} + F_{yrr} \tag{6}$$

$$M_z = a(F'_{yfl} + F'_{yfr}) - b(F_{yrl} + F_{yrr}) + \frac{d_f}{2}(F'_{xfr} - F'_{xfl}) + \frac{d_r}{2}(F_{xrr} - F_{xrl}) \tag{7}$$

式中:m 为整车质量;F_x 和 F_y 分别为整车所受的纵向合力和侧向合力;J_z 为整车绕 z 轴的转动惯量;M_z 为整车所受的横摆合力矩;F_{xi}和 F_{yi}分别为车轮上的纵向力和侧向力(i = fl、fr、rl、rr 分别代表汽车前左轮、前右轮、后左轮、后右轮);F'_{xj}和F'_{yj}分别为前左、右轮在 x、y 轴的纵向合力和侧向合力(j = fl, fr);C_D 为风阻系数;A 为迎风面积;a 和 b 分别为整车质心至前、后轴的距离;d_f 和 d_r 分别为前、后轴的轮距。

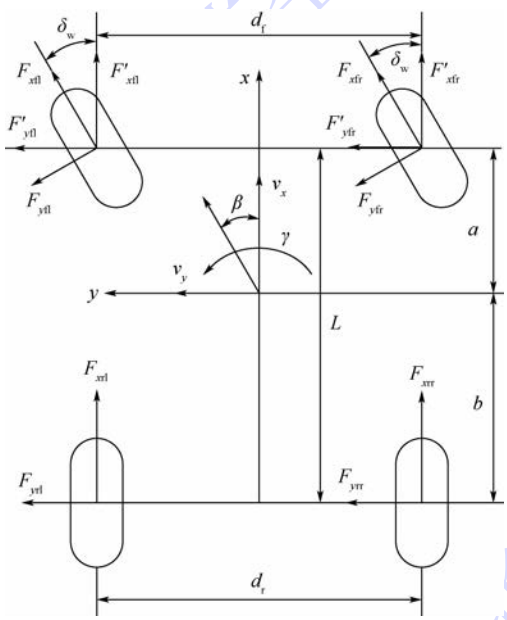


图 2　八自由度整车模型

Fig. 2　Vehicle model with 8 degrees of freedom

表 1　汽车模型主要参数

Table 1　Major parameters of vehicle model

参数	数值
整车质量 m /kg	1 730
轴距 L /m	2.8
整车质心高度 h /m	0.7
质心至前轴的距离 a /m	1.261
前轴的轮距 d_f/m	1.59
后轴的轮距 d_r/m	1.59
整车绕 z 轴的转动惯量 J_z/(kg·m²)	3 000

2　目标滑转率的确定

2.1　轮胎模型

GIM 理论轮胎模型在准确计算轮胎纵向作用力的基础上兼顾了轮胎侧向力作用,因此轮胎模型采用 GIM 理论轮胎模型。

轮胎综合附着系数为

$$\mu = \mu_0\left(\frac{1 - \dfrac{\mu_1}{\mu_0}}{1 - \dfrac{1}{S_1}\sqrt{S_s^2 + S_\alpha^2}}\right) \tag{8}$$

式中:μ_0 为路面峰值附着系数;μ_1 为路面滑转附着系数;S_s 为纵向滑转率;S_α 为侧向滑转率;S_1 = 1 为车轮的最大滑转率。

接触区滚动滑动临界点为

$$S_n = \frac{1}{3\mu F_z}\sqrt{(C_s S_s)^2 + (C_\alpha S_\alpha)^2} \tag{9}$$

式中:F_z 为轮胎法向力;C_s 为轮胎纵向刚度;C_α 为轮胎侧向刚度。

引入无量纲变量 L_n:

$$L_n = 1 - S_n \tag{10}$$

可知轮胎与路面间的纵向力为

$$F_X = \begin{cases} C_s S_s L_n^2 + \mu_x F_z(1 - 3L_n^2 + 2L_n^3) & S_s < S_{sc} \\ \mu_x F_z & S_s \geqslant S_{sc} \end{cases} \tag{11}$$

式中:F_X 为轮胎与路面间的纵向力;μ_x 为轮胎纵向附着系数;S_{sc}为滑转临界点。

2.2　目标滑转率的计算

四轮驱动汽车牵引力控制系统的理想控制目标是将 4 个车轮的滑转率实时控制在目标滑转率附近,然而目标滑转率并不是单一地采用最大纵向附着系数所对应的或者最大侧向附着系数所对应的滑转率来代替。常见的目标滑转率定义为:当纵向附着系数取得最大值时所对应滑转率的附近区域。然而,精确地计算各个车轮任一时刻的附着系数最大值需要确定其影响因素。

令纵向附着系数为

$$\mu_x = \frac{F_X}{F_z} \tag{12}$$

即

$$\mu_x = \frac{C_s S_s L_n^2}{F_z(3L_n^2 - 2L_n^3)} \tag{13}$$

消去中间变量可得

$$\mu_x = \frac{C_s S_s}{F_z + \dfrac{2\sqrt{(C_s S_s)^2 + [C_\alpha(1 - |S_s|)|\tan\beta|]^2}}{3[\mu_0 - (\mu_0 - \mu_1)\sqrt{(S_s^2 + (1 - |S_s|)^2|\tan\beta|^2}]}} \tag{14}$$

式中:β 为轮胎侧偏角。

由式(14)可知,μ_x 是关于 S_s 的函数,峰值纵向附着系数所对应滑转率可以认为是式(14)中求解方程最大值所对应的解。当轮胎数据、路面附着系数确定之后,计算最大纵向附着系数所对应滑转率的影响因素仅为轮胎法向力 F_z 和轮胎侧偏角 β。因此,任何一组二维变量(F_z,β)对应的最大纵向附着系数滑转率 $S_{\mu x\max}$ 都是唯一的。本文所提的最优纵向滑转率查表法将三维数组(F_z,β,$S_{\mu x\max}$)导入到 MATLAB/Simulink 中,利用 2-D Lookup Table 模块进行制表,其输入为轮胎法向力 F_z 和轮胎侧偏角 β,任意时刻的二维变量输入(F_z,β),可通过插值法得到最大纵向附着系数所对应的最优纵向滑转率 $S_{\mu x\max}$。

然而,计算查表得出的 $S_{\mu x\max}$ 并不能直接作为目标滑转率 S_t,本文所研究的目标滑转率修正以车辆行驶速度为临界点,原则如下:

1)车轮无侧偏角时,目标滑转率 S_t 取最大纵向附着系数所对应的 $S_{\mu x\max}$。

2)车轮有侧偏角且处于较低车速时,目标滑转率 S_t 等价于车轮能够发挥最大纵向驱动力的滑转率 $S_{\mu x\max}$,注重提高车辆的起步加速能力。

3)车轮有侧偏角且属于中等车速范围时,使目标滑转率 S_t 适当小于 $S_{\mu x\max}$,兼顾车辆的侧向稳定性和加速性能。

4)车轮有侧偏角且车辆高速行驶时,目标滑转率 S_t 应该进一步减小,侧重保证汽车的侧向稳定性。

5)由于侧向附着系数下降趋势过快,限定 S_t 不大于 0.2。

令 $S=0.9S_{\mu x\max}$,$v_1=2.8$ m/s,$v_2=8.3$ m/s,有

$\beta=0$:

$$S_t=S_{\mu x\max} \tag{15}$$

$\beta\neq0$:

$$S_t=\begin{cases}S_{\mu x\max} & v_x\leqslant v_1\\ \dfrac{(S_{\mu x\max}-S)(v_x-v_1)}{v_1-v_2}+S_{\mu x\max} & v_1<v_x\leqslant v_2\\ \dfrac{2v_2S}{v_x+v_2} & v_x>v_2\end{cases} \tag{16}$$

3 牵引力控制系统控制策略

目前,国内外的牵引力控制系统控制策略往往考虑 2 种影响因素,本文中不仅综合考虑了牵引力控制的 3 种基本影响因素,并且提高了控制策略的效率。四轮驱动汽车牵引力控制系统的有效性不仅与节气门开度、驱动轮制动、轴间扭矩分配[21-22]3 种控制方法的效率有关,控制策略的高效性同样是关键因素。好的控制策略可以将 3 种独立控制有效结合,在保证行车安全性的基础上尽可能地降低能量损耗,充分利用路面附着力,改善车辆在路面行驶的加速性能以及侧向稳定性。

3.1 节气门开度控制的开启条件

节气门开度控制的目的是使车辆的每一个车轮都在目标滑转率区域运转,此时车轮的纵向附着力和侧向稳定性达到理想的平衡状态。为了尽可能地遵循驾驶员的意图,降低发生危险的可能性,通常节气门开度控制仅仅在于在一定程度上减小驾驶员意图的油门开度。当车轮的实际滑转率小于目标滑转率时,牵引力控制系统不能起作用。只有当车轮实际滑转率大于当时时刻的目标滑转率时,节气门开度控制系统才可以通过减小油门开度从而使滑转率值减小并靠近目标滑转率区域。故节气门开度控制系统介入和工作状态的条件为

$$\begin{cases}\overline{S}\geqslant\overline{S_t} & \text{开启}\\ \overline{S}<\overline{S_t} & \text{关闭}\end{cases} \tag{17}$$

式中:$\overline{S}$和$\overline{S_t}$分别为 4 个车轮的实际滑转率平均值和目标滑转率的平均值。

3.2 驱动轮制动控制的开启条件

四轮驱动汽车牵引力控制系统的驱动轮制动控制旨在消除同轴两车轮的转速差,同时适用于消除前、后轴的左右轮转速差。驱动轮制动控制因为局限于车辆的高速稳定性和刹车总成的热衰退性,具有相对严格的开启条件:

$$\begin{cases}\omega_{fl}-\omega_{fr}\neq0\\ \omega_{rl}-\omega_{rr}\neq0\end{cases} \tag{18}$$

式中:ω_{fl}、ω_{fr}、ω_{rl}和 ω_{rr}分别为前左轮轮速、前右轮轮速、后左轮轮速和后右轮轮速。

当汽车高速行驶时,对单个车轮施加制动力矩会影响车辆的侧向稳定性;而且制动系统的频繁介入工作会加剧制动器的热衰退性,不利于行车安全。同时,车辆在进行转向行驶时内、外轮之间会有一定的转速差,考虑到本文研究的四轮驱动汽车牵引力控制系统是针对车辆直线加速工况的加速性能,故当驾驶员转动方向盘意图转向行驶时,驱动轮制动控制退出工作,驱动轮制动控制关闭条件如下:

$$\begin{cases}v_x\geqslant10\text{ m/s}\\ \delta_w\neq0\end{cases} \tag{19}$$

3.3　轴间扭矩分配控制的开启条件

四轮驱动汽车牵引力控制系统的轴间扭矩分配控制旨在消除前、后轴转速差，针对前、后轴车轮的路面附着系数差异较大、车辆急加速和急减速等前、后轴发生负荷转移等工况，轴间扭矩分配控制系统的开启、关闭条件如下：

$$\begin{cases} \dfrac{\omega_{\mathrm{fl}} + \omega_{\mathrm{fr}}}{2} = \omega_{\mathrm{of}} \\ \dfrac{\omega_{\mathrm{rl}} - \omega_{\mathrm{rr}}}{2} = \omega_{\mathrm{or}} \\ \Delta\omega \neq 0, \text{开启} \\ \Delta\omega = 0, \text{关闭} \end{cases} \tag{20}$$

式中：ω_{of} 和 ω_{or} 分别为前轴和后轴差速器壳体转速；$\Delta\omega$ 为前、后轴差速器壳体转速差。

3.4　控制策略

基于常见的路面摩擦工况，选取低附着路面、高-低附着对接路面和分离路面 3 种典型工况进行仿真分析。根据节气门开度、驱动轮制动和轴间扭矩分配 3 种控制的开启条件，得到针对低附着路面直线加速、高-低附着对接路面直线加速和分离路面直线加速 3 种典型工况的控制工作状态。表 2 为典型工况下牵引力控制系统开启状态，以仿真 5 s 之后的车辆纵向行驶车速为 3 种牵引力控制方法有效性的衡量标度，根据表 2 在每一种典型路面工况下，对比无控制条件下的汽车直线加速的纵向车速，分别对节气门开度 PID 控制、驱动轮制动 PID 控制和轴间扭矩分配 PID 控制进行有效性分析。

表 2　典型工况下牵引力控制系统开启状态

Table 2　Open state of traction control system in typical conditions

工况名称	节气门开度	驱动轮制动	轴间扭矩分配
低附着路面直线加速	开启	关闭	开启
高-低附着对接路面直线加速	开启	关闭	开启
分离路面直线加速	开启	开启	开启

图 3 为低附着路面无 PID 控制、节气门开度 PID 控制和轴间扭矩分配 PID 控制的汽车纵向车速对比。节气门开度控制可以明显提高车辆加速性能，在控制策略上作为主要控制方法。无 PID 控制、轴间扭矩分配 PID 控制的 2 条车速曲线几乎重合，车辆在低附着路面直线加速，由于附着系数很小，路面驱动力的大小主要取决于路面可以提供轮胎的附着力，通过轴间扭矩分配控制系统调节分配到前、后轴发动机转矩是徒劳的，无法改善车辆加速性能，在控制策略上标记为关闭。

图 4 为高-低附着对接路面无 PID 控制、节气门开度 PID 控制、轴间扭矩分配 PID 控制的汽车纵向车速对比。节气门开度 PID 控制对于车速的提高幅度并不明显，其主要针对车辆在低附着路面，而车辆在低附着路面加速时间较短，故可将其标记为辅助性调节。轴间扭矩分配控制主要工作在高-低附着过渡区域，因此在控制策略上属于主要控制方法。

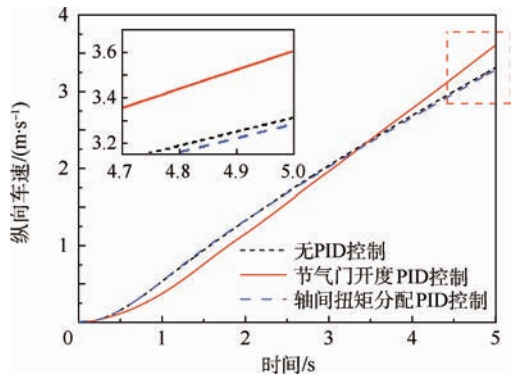


图 3　低附着路面有无控制的汽车纵向车速对比

Fig. 3　Vehicle longitudinal velocity comparison with and without control on low adhesion road

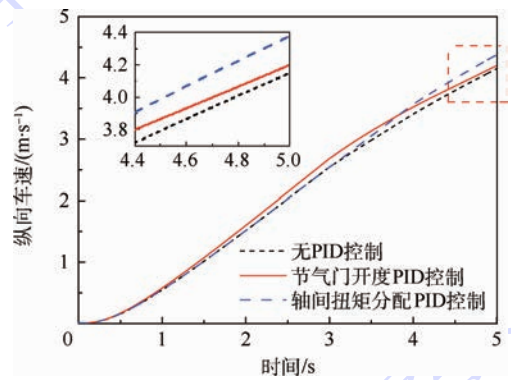


图 4　高-低附着对接路面有无控制的汽车纵向车速对比

Fig. 4　Vehicle longitudinal velocity comparison with and without control on high-low docking adhesion road

图 5 为分离路面无 PID 控制、节气门开度 PID 控制、驱动轮制动 PID 控制和轴间扭矩分配 PID 控制的汽车纵向车速对比。无 PID 控制、节气门开度 PID 控制的 2 条车速曲线基本重合，但在分离路面，节气门开度调节通常作为辅助控制手段，以减小低附着侧路面的车轮滑转程度，提高整车的侧向稳定性。驱动轮制动 PID 控制中车辆加速性能小幅度提升，为减小低附着侧路面的车轮滑转程度和保证高附着侧路面的车轮具有一定转速，所以驱动轮制动调节是主要的控制方法之一。轴间扭矩分配 PID 控制主要作用于路面附着系数变化较大的阶段，车辆加速性能明显提升，也是分离路面牵引力控制的主要方法之一。

根据上述分析结果，得出针对四轮驱动汽车的牵引力控制系统控制策略，如表 3 所示。

北航学报 赠阅

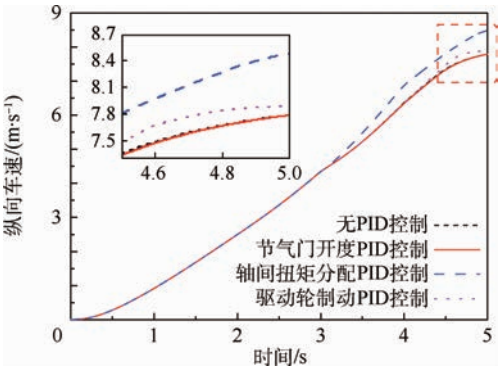


图 5 分离路面有无控制的汽车纵向车速对比

Fig. 5 Vehicle longitudinal velocity comparison with and without control on split cohesion coefficient road

表 3 牵引力控制系统的控制策略

Table 3 Control strategy of traction control system

工况名称	节气门开度	驱动轮制动	轴间扭矩分配
低附着路面直线加速	主要		
高-低附着对接路面直线加速	辅助		主要
分离路面直线加速	辅助	主要	主要

4 典型工况仿真试验

分别在低附着路面、高-低附着对接路面和分分离路面 3 种典型工况下进行直线加速离线仿真试验，并对有无牵引力控制进行对比分析。

4.1 低附着路面仿真分析

仿真初始条件为：路面环境为低附着路面（$\mu_0=0.1$，$\mu_1=0.07$），初始车速为 0 m/s，变速器档位是 3 档，节气门位置为 100%，即全油门原地直线加速起步，仿真时间为 5 s。

图 6 为四轮驱动汽车低附着路面直线加速无牵引力控制仿真结果。无牵引力控制时，由于路面附着系数过低，驾驶员意图是全油门原地加速起步。4 个车轮转速上升过快，导致车轮滑转率增大并接近 1，无法充分利用路面附着条件，车轮的路面驱动力下降过快。仿真结束时，车速为 3.31 m/s。

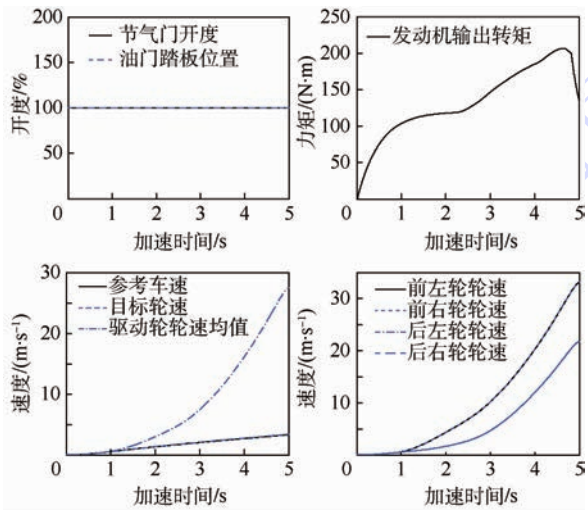


图 6 低附着路面无牵引力控制仿真结果

Fig. 6 Simulation results without traction control on low adhesion road

图 7 为四轮驱动汽车低附着路面直线加速模糊 PID 牵引力控制仿真结果。牵引力控制很好地降低车轮的过度滑转，控制结果与 PID 控制类似。仿真结束时，车速为 3.60 m/s，比无牵引力控制的车速提高 8.76%。

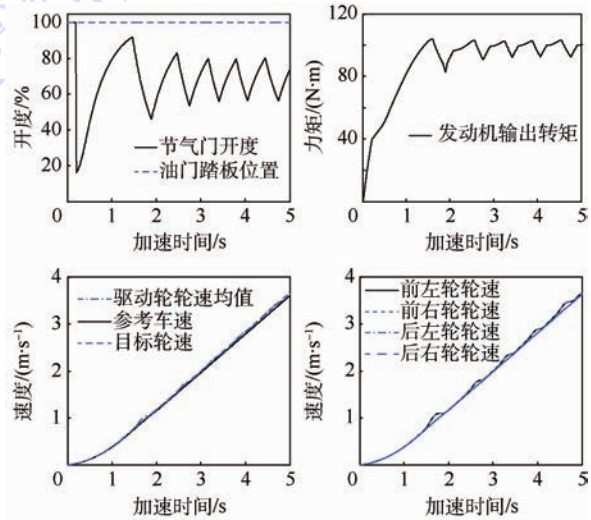


图 7 低附着路面模糊 PID 牵引力控制仿真结果

Fig. 7 Simulation results with fuzzy PID traction control on low adhesion road

4.2 高-低附着对接路面仿真分析

仿真初始条件为：路面环境为高-低附着对接路面。第 1 段（0 ~ 3 s）为高附着均一路面，$\mu_0=0.8$，$\mu_1=0.7$；第 2 段（3 ~ 5 s）为低附着均一路面，$\mu_0=0.1$，$\mu_1=0.07$。初始车速为 0 m/s，变速器档位是 3 档，节气门位置为 100%，即全油门原地直线加速起步，仿真时间为 5 s。

图 8 为四轮驱动汽车高-低附着对接路面直线加速无牵引力控制仿真结果。无牵引力控制时，完全按照驾驶员意图全油门原地加速起步。第 1 阶段是高附着均一路面，4 个车轮滑转情况良好。当车辆逐渐进入低附着路面时，前轴 2 个车轮最先处于低附着路面，出现轮速迅速加快的现象，与仍处于高附着路面的后轴 2 个车轮形成较大的转速差。当车辆完全行驶在低附着路面时，4 个车轮均出现较为严重的滑转，路面附着条件没有得到充分利用且车轮的路面驱动力下降较快。仿真结束时，车速为 4.15 m/s。

北航学报 赠阅

图 9 为四轮驱动汽车高-低附着对接路面直线加速模糊 PID 牵引力控制仿真结果。可以看出,车辆在高附着路面行驶时,4 个车轮滑转情况良好,无需牵引力控制系统介入工作。当车辆开始进入低附着路面时,前后轴的车轮分别处于不同附着系数的路面之上,轴间扭矩分配模糊 PID 控制均可迅速识别路面变化,通过改变轴间扭矩

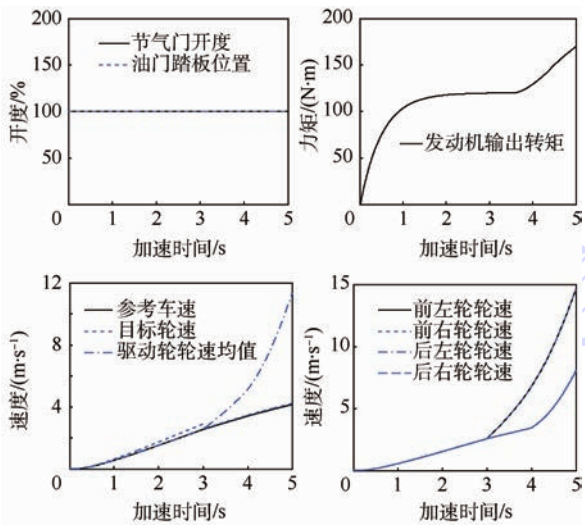


图 8　高-低附着对接路面无牵引力控制仿真结果

Fig. 8　Simulation results without traction control on high-low docking adhesion road

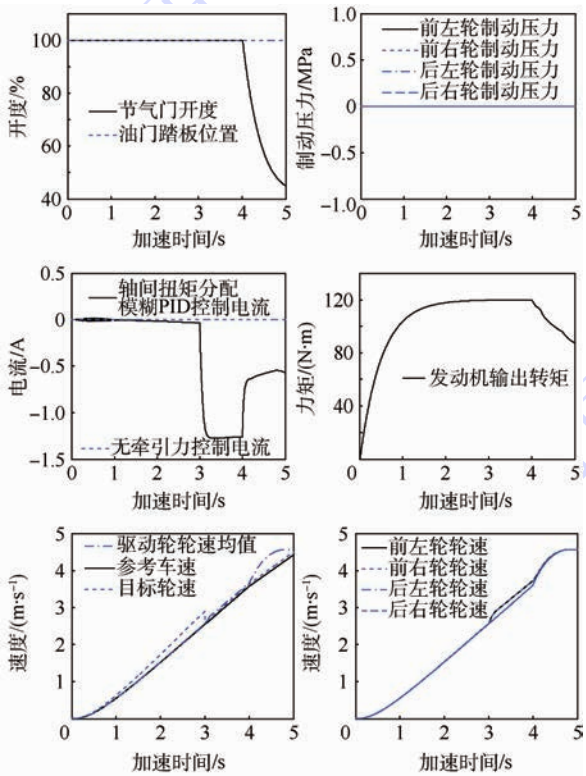


图 9　高-低附着对接路面模糊 PID 牵引力控制仿真结果

Fig. 9　Simulation results with fuzzy PID traction control on high-low docking adhesion road

分配装置中电磁离合器的通电电流值起到根据车轮所需扭矩而分配的作用。图中电流为负值,代表当前轴车轮在低附着路面、后轴车轮在高附着路面时,后轴车轮需要更大的驱动力矩,而前轴车轮的驱动力矩需要降低。将前、后轴车轮的驱动力矩作对比如图 10 所示。

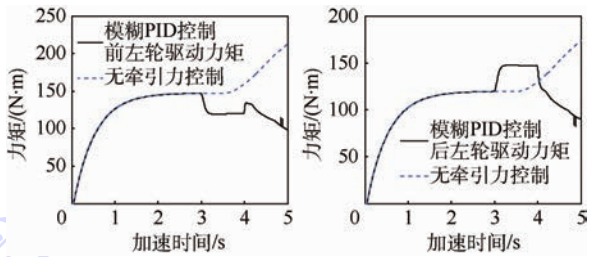


图 10　2 种控制方法的驱动力矩对比

Fig. 10　Comparison of driving moment with two control methods

由图 10 可以清晰地看出,在路面对接阶段,轴间扭矩分配控制增大后轴车轮的转矩输出,一定程度地减小前轴车轮的转矩输出,提高了车辆的起步加速性能。当车辆完全进入低附着路面时,节气门开度调节进入工作状态,及时抑制了车轮的滑转。轴间扭矩分配控制有效降低了前、后轴转速差。仿真结束时,车速为 4.56 m/s,比无牵引力控制的车速提高 9.88%。

4.3　分离路面仿真分析

仿真初始条件为:第 1 段(0 ~ 3 s)为高附着均一路面,$\mu_0 = 0.8$,$\mu_1 = 0.7$;第 2 段(3 ~ 5 s)左侧为低附着路面,$\mu_0 = 0.1$,$\mu_1 = 0.07$,右侧为高附着路面,$\mu_0 = 0.8$,$\mu_1 = 0.7$。初始车速为 0 m/s,变速器档位是 2 档,节气门位置为 100%,即全油门原地直线加速起步,仿真时间为 5 s。

对于分离路面直线加速工况,牵引力控制包括节气门开度、轴间扭矩分配和驱动轮制动 3 种方法。

图 11 为四轮驱动汽车分离路面直线加速无牵引力控制仿真结果。无牵引力控制时,同轴的左、右车轮之间存在较大的转速差,尤其是前轴。前、后轴的车轮之间同样存在一定的转速差。车轮出现滑转的现象,路面附着条件得不到充分利用。仿真结束时,车速为 7.79 m/s。

图 12 为四轮驱动汽车分离路面直线加速模糊PID 牵引力控制仿真结果。仿真结束时,车速为 8.60 m/s,比无牵引力控制的车速提高 10.40%。

北航学报 赠阅

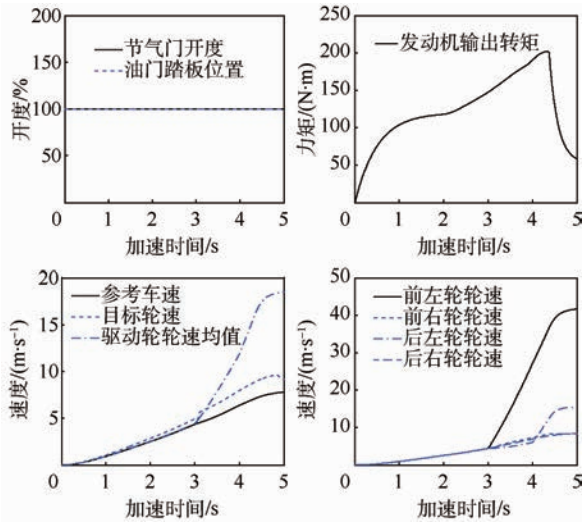


图 11 分离路面无牵引力控制仿真结果

Fig. 11 Simulation results without traction control on split cohesion coefficient road

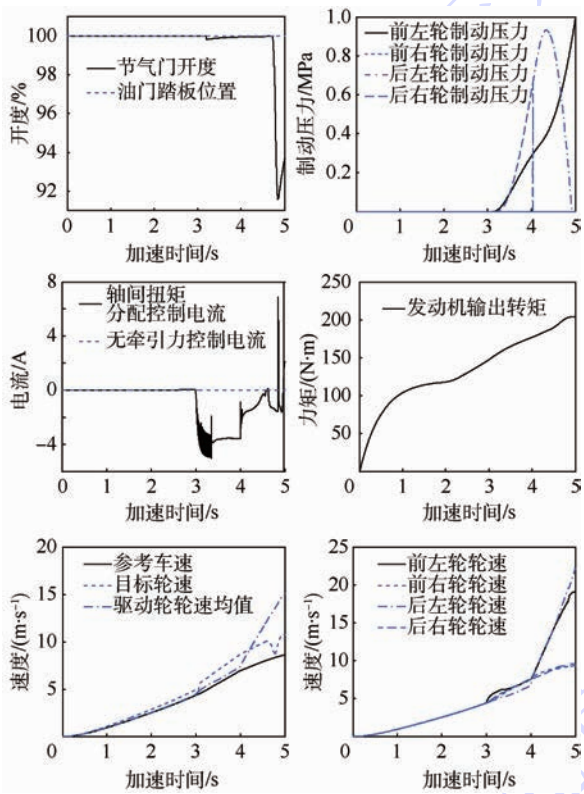


图 12 分离路面模糊 PID 牵引力控制仿真结果

Fig. 12 Simulation results with fuzzy PID traction control on split cohesion coefficient road

5 结 论

1）低附着路面工况下，无牵引力控制时，由于路面附着系数过低，车轮出现滑转，同时车轮的路面驱动力下降较快。有牵引力控制时很好地降低车轮的过度滑转，仿真结束时车速比无牵引力控制的车速提高了 8.76%。

2）高-低附着对接路面工况下，无牵引力控制时，在高附着均一路面上，4 个车轮滑转情况良好，然而当车辆逐渐进入低附着路面时，前、后轴车轮出现较大的转速差；当车辆完全行驶在低附着路面时，4 个车轮均出现较为严重的滑转。有牵引力控制时，当车辆开始进入低附着路面，前、后轴的车轮分别处于不同附着系数的路面上，轴间扭矩分配模糊 PID 控制有效地降低前、后轴的转速差；当汽车完全进入低附着路面时，节气门开度调节进入工作状态，及时抑制了车轮的滑转，仿真结束时车速比无牵引力控制的车速提高了 9.88%。

3）分离路面工况下，无牵引力控制时，同轴的左、右车轮之间存在较大的转速差，前、后轴的车轮之间同样存在一定的转速差，车轮出现滑转。有牵引力控制时，能够很好地消除同轴的左右轮的转速差，仿真结束时车速比无牵引力控制的车速提高 10.40%。

仿真结果表明，本文所提的基于最优纵向滑转率查表法的牵引力控制系统控制策略能有效抑制驱动轮的过度滑转，提高汽车的纵向动力性和横向稳定性。

参考文献（References）

[1] JALALI K, UCHIDA T, MCPHEE J, et al. Development of a fuzzy slip control system for electric vehicles with in-wheel motors[J]. SAE International Journal of Alternative Powertrains, 2012, 1(1): 46-64.

[2] HE H, PENG J, XIONG R, et al. An acceleration slip regulation strategy for four-wheel drive electric vehicles based on sliding mode control[J]. Energies, 2014, 7(6): 3748-3763.

[3] NIEWELS F, SCHUH J. Traction control system (TCS)[M]// REIF K. Brakes, brake control and driver assistance system. New York: Springer, 2014: 94-101.

[4] OSBORN R P, SHIM T. Independent control of all-wheel-drive torque distribution[J]. Vehicle System Dynamics, 2006, 44(7): 529-546.

[5] WANG J, WANG Q, JIN L, et al. Independent wheel torque control of 4WD electric vehicle for differential drive assisted steering[J]. Mechatronics, 2011, 21(1): 63-76.

[6] 赵治国，顾君，余卓平. 四轮驱动混合动力轿车驱动防滑控制研究[J]. 机械工程学报，2011，47(14)：83-98.
ZHAO Z G, GU J, YU Z P. Study of acceleration slip regulation strategy for four wheel drive hybrid electric car[J]. Journal of Mechanical Engineering, 2011, 47(14): 83-98 (in Chinese).

[7] LI B, GOODARZI A, KHAJEPOUR A, et al. An optimal torque distribution control strategy for four-independent wheel drive electric vehicles[J]. Vehicle System Dynamics, 2015, 53(8): 1172-1189.

[8] KANG J, KYONGSU Y, HEO H. Control allocation based opti-

北航学报 赠阅

mal torque vectoring for 4WD electric vehicle[J/OL]. SAE Technical Paper[2012-04-16]. doi:10.4271/2012-01-0246.

[9] YIN G,WANG S,JIN X. Optimal slip ratio based fuzzy control of acceleration slip regulation for four-wheel independent driving electric vehicles[J]. Mathematical Problems in Engineering,2013,2013:1-7.

[10] 胡建军,王银,秦大同,等. 基于轴间转矩分配的四轮驱动汽车牵引力控制[J]. 华南理工大学学报(自然科学版),2010,38(1):108-112.
HU J J,WANG Y,QIN D T,et al. Traction control of four-wheel drive vehicles based on inter-axle torque distribution[J]. Journal of South China University of Technology(Natural Science Edition),2010,38(1):108-112(in Chinese).

[11] GIM G. An analytical model of pneumatic tyres for vehicle dynamic simulations. Part 1:Pure slips[J]. International Journal of Vehicle Design,1990,11(6):589-618.

[12] SHUAI Z,ZHANG H,WANG J,et al. Lateral motion control for four-wheel-independent-drive electric vehicles using optimal torque allocation and dynamic message priority scheduling[J]. Control Engineering Practice,2014,24:55-66.

[13] WU L F,GOU J F,WANG L F,et al. Acceleration slip regulation strategy for distributed drive electric vehicles with independent front axle drive motors[J]. Energies,2015,8(5):4043-4072.

[14] KECSKES I,ODRY P. Multi-scenario optimization approach for fuzzy control of a robot-car model[C]//2015 IEEE 10th Jubilee International Symposium on Applied Computational Intelligence and Informatics (SACI). Piscataway,NJ:IEEE Press,2015:565-570.

[15] LI H Z,LI L,HE L,et al. PID plus fuzzy logic method for torque control in traction control system[J]. International Journal of Automotive Technology,2012,13(3):441-450.

[16] LI Z Y,SU D D,LI H D,et al. Research on acceleration slip regulation for vehicle based on compound control of fuzzy and PID[C]// International Conference on Materials Science and Information Technology (MSIT 2011). Zurich:Trans Tech Publications Ltd.,2012,433:4165-4172.

[17] 周云山,裘熙定,王红岩,等. 汽车无级变速传动(CVT)建模与仿真[J]. 汽车工程,1998,20(5):285-289.
ZHOU Y S,QIU X D,WANG H Y,et al. The modeling and simulation for CVT[J]. Automotive Engineering,1998,20(5):285-289(in Chinese).

[18] 程铖. 基于 dSPACE 的四轮驱动乘用车牵引力控制策略研究[D]. 广州:华南理工大学,2015:19-44.
CHENG C. The research on traction control strategy of 4 wheel drive vehicle based on dSPACE[D]. Guangzhou:South China University of Technology,2015:19-44(in Chinese).

[19] KAZEMI R,HAMEDI B,JAVADI B. A new mode controller for four-wheel sliding anti-lock braking system (ABS)[J/OL]. SAE Technical Paper[2000-05-01]. doi:10.4271/2000-01-1639.

[20] GERSTENMEIER J. Traction control (ASR)-An extension of the anti-lock braking system (ABS)[J/OL]. SAE Technical Paper[1986-10-20]. doi:10.4271/861033.

[21] PANZANI G,CORNO M,TANELLI M,et al. Designing on-demand four-wheel-drive vehicles via active control of the central transfer case[J]. IEEE Transactions on Intelligent Transportation Systems,2010,11(4):931-941.

[22] LATIF R,AZIM R,BAQAI A,et al. Vehicle modeling and performance evaluation using active torque distribution[J/OL]. SAE Technical Paper[2014-04-01]. doi:10.4271/2014-01-0103.

作者简介:

姜立标　男,博士,副教授,硕士生导师。主要研究方向:汽车系统动力学及电子控制。
E-mail:jlb620620@163.com

丘华川　男,硕士研究生。主要研究方向:汽车系统动力学及电子控制。
E-mail:qhc0327@163.com

北京航空航天大学学报

北航学报 赠阅

Control strategy of traction control system for four-wheel drive vehicle

JIANG Libiao[1,*], QIU Huachuan[1], WU Zhongwei[1], LIU Wei[2], CHENG Cheng[1]

(1. School of Mechanical & Automotive Engineering, South China University of Technology, Guangzhou 510640, China;
2. Zhejiang Geely Automobile Institute Co., Ltd., Hangzhou 311228, China)

Abstract: Aimed at the problem of the vehicle dynamics and lateral stability of four-wheel drive vehicle, the integrated traction control strategy based on a new method of target slip ratio calculation and correction was proposed. First a mathematical model describing the main components of vehicle system was established. In view of different conditions, the traction control system integrated control strategy, which is based on the engine throttle valve control, inter-axis torque distribution control and driving wheel braking-force active control, was built by using fuzzy PID control technology. In the aspect of selection in tire target slip ratio, a lookup method of the optimal longitudinal slip ratio was proposed, and real-time correction of target slip ratio was carried out according to the vehicle velocity and tire slip angle. Finally, research on off-line simulation for four-wheel drive vehicle traction control system in typical conditions was completed by using MATLAB/Simulink software. The results show that the proposed traction control strategy can effectively inhibits the driving wheel from over slipping, fully use the road adhesion conditions and improve the vehicle longitudinal driving stability.

Key words: four-wheel drive vehicle; traction control; target slip ratio; control strategy; dynamics simulation

北京航空航天大学学报

Received: 2015-10-23; **Accepted**: 2015-12-04; **Published online**: 2015-12-09 15:32

URL: www.cnki.net/kcms/detail/11.2625.V.20151209.1532.002.html

Foundation items: National Natural Science Foundation of China (50975091); Natural Science Foundation of Guangdong Province of China (9451064101003049)

* **Corresponding author**. E-mail: jlb620620@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0672

海基 JPALS 的多参考一致性检验算法

冯涛[1,2], 毛峡[1,*], 宋成[3]
(1. 北京航空航天大学 电子信息工程学院, 北京 100083; 2. 中国交通通信信息中心, 北京 100011;
3. 中国卫星导航系统管理办公室, 北京 100034)

摘 要: 针对舰载参考站的移动特性提出了海基联合精密进近着陆(舰)系统(JPALS)的多参考一致性检验(MRCC)算法。首先,利用舰船上的惯导系统输出的姿态信息和舰船尺寸测量数据,将每个参考站的原始伪距和载波相位观测量从其天线位置转换到舰船参考点上;然后,对转换后的各参考站几何无关组合观测量进行均值滤波,得到其宽巷模糊度估计值;最后,计算转换后的各原始观测量和宽巷模糊度估计值的 B 值。在无故障情况下,这些测量值的 B 值分布在一定范围(阈值)内,但在参考站故障、严重多径或无线电干扰等情形下,B 值可能超出阈值。通过模拟实验采集数据,并计算 B 值幅度和阈值,结果表明,通常情况下各观测量的 B 值幅度很小,当某个参考站的接收测量值出现异常时,可以检测到其 B 值超过阈值。研究表明,该算法满足海基 JPALS 的多参考一致性检验完好性风险指标要求。

关 键 词: 海基 JPALS; 多参考一致性检验(MRCC); 宽巷模糊度; B 值; 完好性
中图分类号: V249.3; TN967.4
文献标识码: A **文章编号**: 1001-5965(2016)11-2299-07

舰载机着舰是航空母舰的一项关键运行技术。美国正在研究一种兼容陆海空和军民两用的联合精密进近着陆(舰)系统(Joint Precision Approach and Landing System, JPALS),其分为陆基系统和海基系统。海基 JPALS 使用高精度卫星定位导航技术进行着舰过程中飞机的精确定位和进近引导,能满足全天候自动着舰要求[1]。研究海基 JPALS 对于促进中国航空母舰舰载机着舰技术的发展具有重要意义。

海基 JPALS 性能要求较高,为同时满足自动着舰要求,其垂直方向定位精度需求为 0.3 m,告警门限为 1.1 m,完好性风险不超过 $1\times10^{-7}/150\,s$[2]。海基 JPALS 采用多频接收机和特殊设计的导航算法减轻或消除了卫星星历误差、钟差、电离层延迟、对流层延迟和接收机钟差等误差项的影响,但电离层风暴、接收机故障和多径效应等导致的显著测量误差仍然难以避免[3],因此海基 JPALS 完好性监测系统的设计非常重要。目前,海基 JPALS 主要是利用舰船上多个参考站接收机的观测数据进行完好性监测,舰载机则依靠机载接收机自身定位数据和接收来自舰船广播的观测数据、完好性监测数据进行定位结果和保护门限的综合解算,以获得满足精度和完好性需求的相对定位结果。为保证舰船广播的观测数据质量,海基 JPALS 需要采用多参考一致性检验(Multiple References Consistency Check, MRCC)算法,监测并隔离存在严重多径和接收机故障的测量值。

针对静止参考站的多参考一致性检验,一般采用伪距和载波相位的差分改正数作为 B 值计算的测量值[4-6],这种算法方式在区域增强系统(Local Area Augmentation System, LAAS)和陆基 JPALS 中已被广泛研究[7-9]。但在海基 JPALS 中,

收稿日期: 2015-10-19; **录用日期**: 2015-11-06; **网络出版时间**: 2016-01-04 10:04
网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20160104.1004.011.html
* **通讯作者**: Tel.: 010-82339508 E-mail: moukyou@buaa.edu.cn

引用格式: 冯涛, 毛峡, 宋成. 海基 JPALS 的多参考一致性检验算法[J]. 北京航空航天大学学报, 2016, 42(11): 2299-2305.
FENG T, MAO X, SONG C. Multiple reference consistency check algorithm for sea-based JPALS[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(11): 2299-2305 (in Chinese).

北航学报 赠阅

参考站随舰体运动,其绝对位置无法获取,相应的伪距和载波相位差分改正数无法通过测量值减去真实距离得到[10],因此陆基 JPALS 的算法不适用于海基 JPALS。本文提出利用多个参考站与理想着舰点之间相对位置矢量已知的特性计算各参考站原始观测量和宽巷模糊度估计,进而计算其 B 值的算法,来实现海基 JPALS 的多参考一致性检验。

1 海基 JPALS 及其完好性监测系统

海基 JPALS 主要由舰载系统、机载系统、数据链系统和完好性监测系统四大部分组成,如图 1所示。舰载系统由部署在舰船上不同位置的多个全球卫星导航系统(Global Navigation Satellite System,GNSS)参考站接收机及其天线、惯导系统、数据处理系统和数据收发系统组成,其功能是计算产生宽巷模糊度、原始观测量、完好性数据、下滑道数据和舰船运动信息等,通过数据链广播给舰载机。机载系统包括 GNSS 接收机及其天线、惯导系统、数据处理系统和数据收发系统,其功能是利用自身计算产生的和来自舰载系统播发的宽巷模糊度、原始观测量和完好性数据计算出机舰相对位置矢量,并联合下滑道数据和舰船运动信息等一起传送给机载航电系统。数据链系统包括位于舰船和舰载机上的 UHF 天线和信号调制解调系统,用于实现舰船和飞机之间的数据传输。完好性监测系统位于舰船上,利用多个舰载参考站数据对信号异常和系统故障进行监测,并及时向用户发送告警信息[2],确保系统满足完好性风险指标要求。

海基 JPALS 利用多个参考站实现电离层梯度监测、天线基线监测、参考一致性检验、星历监测和无线电干扰监测等完好性监测任务。海基 JPALS 完好性监测系统由 4 个左右(本文以 4 个为例)的舰载参考站接收机和完好性数据处理系统组成,舰载参考站接收机的安装如图 2 所示。图中:1 ~4 号三角形表示参考站接收机天线的安装位置;飞机跑道上的圆点表示舰船的理想着舰点。

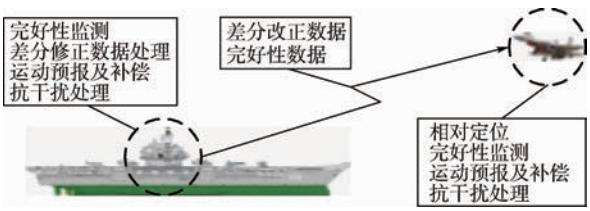


图 1 海基 JPALS 组成

Fig. 1 Structure of sea-based JPALS

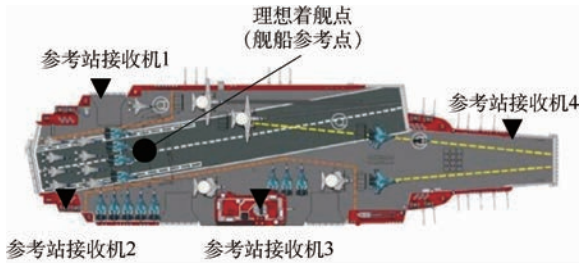


图 2 舰载多参考站安装示意图

Fig. 2 Schematic diagram of deployment of multiple reference stations on ship

2 导航算法

2.1 舰船参考点观测量

着舰过程中,飞机上的尾钩和理想着舰点(图 2 中圆点)之间的相对位置矢量对于着舰进近引导具有实质性的作用。通常,海基 JPALS 的舰载系统将理想着舰点视为舰船参考点,并将各参考站的测量值从天线位置映射转换到参考点上[11]。转换过程利用了参考站接收机到舰船参考点的相对位置矢量,该矢量可以通过舰载惯导系统和舰船自身的尺寸测量数据计算得到[12]。如图 3 所示,ρ_m 和 ϕ_m($m=1,2,3,4$)为参考站接收机的观测量,ρ_{mr}和 ϕ_{mr}($m=1,2,3,4$)为转换到舰船参考点后的测量值。

转换后每个参考站接收机通道(m,n)的 L_1 和 L_2 的伪距测量值 ρ^n_{mr,L_1}、ρ^n_{mr,L_2} 与载波相位测量值 ϕ^n_{mr,L_1}、ϕ^n_{mr,L_2} 分别为

$$\rho^n_{mr,L_1}=R+\frac{I}{f_1^2}+T+(c\delta t-c\delta t^s)+B_1+M_1+E_1+\varepsilon_T \tag{1}$$

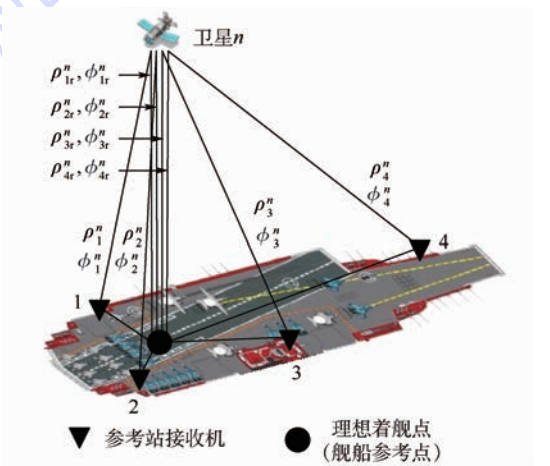


图 3 测量值映射转换示意图

Fig. 3 Schematic diagram of measurements mapping transformation

北航学报 赠阅

$$\rho^n_{mr,L_2}=R+\frac{I}{f_2^2}+T+(c\delta t-c\delta t^s)+B_2+M_2+E_2+\varepsilon_T \tag{2}$$

$$\phi^n_{mr,L_1}=R-\frac{I}{f_1^2}+T+(c\delta t-c\delta t^s)+b_1+m_1+e_1+[\lambda_1N_1+\lambda_1(\delta\varphi-\delta\varphi^s)]+\varepsilon_T \tag{3}$$

$$\phi^n_{mr,L_2}=R-\frac{I}{f_2^2}+T+(c\delta t-c\delta t^s)+b_2+m_2+e_2+[\lambda_2N_2+\lambda_2(\delta\varphi-\delta\varphi^s)]+\varepsilon_T \tag{4}$$

式中：R 为卫星到参考站的真实距离；m 为参考站编号；n 为卫星编号；I 为电离层延迟误差；T 为对流层延迟误差；δt 为参考站接收机钟差；δt^s 为卫星钟差；B_1、B_2、b_1 和 b_2 为参考站接收机通道偏差；M_1、M_2、m_1 和 m_2 为多径误差；E_1、E_2、e_1 和 e_2 为参考站接收机噪声；λ_1 和 λ_2 为信号波长；f_1 和 f_2 为信号频率；N_1 和 N_2 为载波相位模糊度；ε_T 为参考站接收机天线到舰船参考点转换引入的误差；$\delta\varphi$ 为参考站接收机天线的相位偏差；$\delta\varphi^s$ 为卫星的相位偏差；c 为光速。

2.2　两阶段导航算法

同时使用多个观测量有助于提高整周模糊度的解算效率，其适当组合可消除系统钟差、电离层延迟、对流层延迟以及 GNSS 导航引入的显著误差等。海基 JPALS 利用多频接收技术同时获取多个观测量，结合其宽巷组合值进行高精度相对定位解算，解算过程分 2 个阶段：第 1 阶段求解宽巷模糊度浮点解估计值；第 2 阶段计算固定解和相对定位矢量[3,11]。

2.2.1　第 1 阶段

第 1 阶段，通过 L_1/L_2 几何无关组合观测量和过滤算法计算宽巷模糊度。通道(m,n)的几何无关组合观测量 $Z^n_{mr,gf}$ 为 L_1、L_2 载波相位测量值 ϕ^n_{mr,L_1}、ϕ^n_{mr,L_2}和伪距测量值 ρ^n_{mr,L_1}、ρ^n_{mr,L_2}4 个测量值的线性组合。$Z^n_{mr,gf}$及其方差 σ_Z^2 计算式为

$$Z^n_{mr,gf}=\left(\frac{\phi^n_{mr,L_1}}{\lambda_1}-\frac{\phi^n_{mr,L_2}}{\lambda_2}\right)-\left(\frac{\rho^n_{mr,L_1}}{\lambda_1}+\frac{\rho^n_{mr,L_2}}{\lambda_2}\right)\left(\frac{\lambda_2-\lambda_1}{\lambda_1+\lambda_2}\right) \tag{5}$$

$$\sigma_Z^2=\left(\frac{1}{\lambda_1^2}+\frac{1}{\lambda_2^2}\right)\left[\sigma_\phi^2+\left(\frac{\lambda_2-\lambda_1}{\lambda_1+\lambda_2}\right)^2\sigma_\rho^2\right] \tag{6}$$

式中：$\sigma_\phi^2=\sigma^2_{\phi,L_1}=\sigma^2_{\phi,L_2}$为 L_1 和 L_2 的载波相位测量方差；$\sigma_\rho^2=\sigma^2_{\rho,L_1}=\sigma^2_{\rho,L_2}$为 L_1 和 L_2 的伪距测量方差。

几何无关组合观测量 $Z^n_{mr,gf}$ 是宽巷模糊度 $N^n_{mr,w}$的估计值，其包含随机误差 ε_{gf}和非测量偏差 b，因此几何无关组合观测量与宽巷模糊度的关系可以表示为

$$Z^n_{mr,gf}=N^n_{mr,w}+\varepsilon_{gf}+b \tag{7}$$

非测量偏差 b 包含频间偏差、天线偏差和故障模式偏差等，其影响可以通过第 2 阶段算法减轻，随机误差 ε_{gf}可通过移动均值滤波算法消除。宽巷模糊度的浮点解估计值 $\hat N^n_{mr,w}$及其方差$\sigma^2_{\hat N_w}$通过滤波算法求解，计算方案如下：

$$\hat N^n_{mr,w}(t)=\frac{1}{M_{fep}}\sum_{j=1}^{M_{fep}}Z^n_{mr,gf}(t-k\Delta t) \tag{8}$$

$$\sigma^2_{\hat N_w}=\frac{1}{M_{fep}}\sigma_Z^2 \tag{9}$$

式中：t 为历元时刻；Δt 为采样间隔；M_{fep}为滤波历元数；k 为滤波期间的历元编号。

2.2.2　第 2 阶段

第 2 阶段，利用宽巷模糊度和原始载波相位观测量联合解算 L_1/L_2 整周模糊度的固定解。航空母舰通过 UHF 数据链向外广播舰载接收机上参考站的原始观测量和宽巷模糊度估计值。舰载机利用接收到的舰载参考站原始观测值和宽巷模糊度估计值$(\phi_{r,L_1},\phi_{r,L_2},\hat N_{r,w})$计算舰载和机载 2 个接收机的双差观测量 $\phi_{L_1,dd}$、$\phi_{L_2,dd}$、$\hat N_{w,dd}$。

利用双差观测量求解基线矢量的方程为

$$\begin{bmatrix}\hat N_{w,dd}\\ \phi_{L_1,dd}\\ \phi_{L_2,dd}\end{bmatrix}=\begin{bmatrix}\mathbf{0}\\ -\Delta\boldsymbol{e}^T\\ -\Delta\boldsymbol{e}^T\end{bmatrix}\Delta\boldsymbol{x}+\begin{bmatrix}\boldsymbol{I}&-\boldsymbol{I}\\ \lambda_1\boldsymbol{I}&\mathbf{0}\\ \mathbf{0}&\lambda_2\boldsymbol{I}\end{bmatrix}\begin{bmatrix}\hat N_{L_1,dd}\\ \hat N_{L_2,dd}\end{bmatrix}+\boldsymbol{\varepsilon} \tag{10}$$

式中：$\Delta\boldsymbol{e}^T$ 为参考站接收机相对 2 颗卫星的单位视向矢量之差；$\boldsymbol{I}$ 为单位矩阵；$\boldsymbol{\varepsilon}$ 为残差矩阵。

首先通过最小二乘模糊度去相关平差（Least-squares Ambiguity Decorrelation Adjustment，LAMBDA）算法计算宽巷双差模糊度的固定解[13]，然后计算 L_1 和 L_2 的双差整周模糊度的浮点解 $\hat N_{L_1,dd}$、$\hat N_{L_2,dd}$及协方差矩阵，进而解算出 L_1 和 L_2 的双差整周模糊度固定解[14]，最后代入式（10）中求解得到基线矢量 $\Delta\boldsymbol{x}$。

2.3　舰船广播数据

为实现两阶段的导航算法，需要将舰船参考点观测量测量值 ρ_{r,L_1}、ρ_{r,L_2}、ϕ_{r,L_1}、ϕ_{r,L_2}和宽巷模糊度 $\hat N_{r,w}$及其标准差实时广播给舰载机。舰船参考点测量值为参考站接收机观测量转换到舰船参考点之后的测量值求平均，计算式为

$$\rho_{r,L_1}=\sum_{j=1}^{M}\rho_{jr,L_1},\rho_{r,L_2}=\sum_{j=1}^{M}\rho_{jr,L_2} \tag{11}$$

$$\phi_{r,L_1}=\sum_{j=1}^{M}\phi_{jr,L_1},\phi_{r,L_2}=\sum_{j=1}^{M}\phi_{jr,L_2} \tag{12}$$

北航学报 赠阅

$$\hat{N}_{r,w}=\sum_{j=1}^{M}\hat{N}_{jr,w} \tag{13}$$

式中：M 为参考站接收机的数量。

3 多参考一致性检验

3.1 多参考一致性检验通用理论

多参考一致性检验的方法是对多个参考站的同类型测量值进行比较，以识别测量值故障（如严重多径、接收机故障或者天线校验故障）引起的显著误差。当某个参考站上任一测量值的 B 值超过事先定义的阈值时，将参考站标记为无效状态。参考站 m 的 B 值的计算式为

$$B_{P,m}(k)=P(k)-\frac{1}{M(k)-1}\sum_{\substack{j\in S_n(k)\\ j\neq m}}^{M(k)}P_j(k) \tag{14}$$

式中：P 为参考站测量值；$M(k)$ 为 k 历元时刻有效参考站的数量；$S_n(k)$ 为 k 历元时刻有效参考站的集合；j 为 $S_n(k)$ 集合中参考站的序号。

对于海基 JPALS，需要计算并比较每个舰载参考站接收机的宽巷模糊度估计和各原始观测量的 B 值，以消除不一致的测量值[3]。

3.2 测量值

海基 JPALS 多参考一致性检验要求计算式(10)中所有测量值的 B 值，包括 L_1/L_2 的伪距观测量 ρ_{r,L_1}、ρ_{r,L_2}，载波相位观测量 ϕ_{r,L_1}、ϕ_{r,L_2} 和宽巷模糊度估计 $\hat{N}_{r,w}$ 共 5 个测量值[3]。各参考站的 L_1/L_2 的伪距和载波相位原始观测量均映射到舰船参考点表示为 ρ_{mr,L_1}、ρ_{mr,L_2}、ϕ_{mr,L_1}、ϕ_{mr,L_2}，每个参考站的宽巷模糊度估计 $\hat{N}_{mr,w}$ 及方差根据式(8)和式(9)计算得到。

3.3 B 值计算

L_1 和 L_2 的伪距和载波相位原始观测量的 B 值计算式分别为

$$B_{\rho,mr,L_1}(k)=\rho_{r,L_1}(k)-\frac{1}{M(k)-1}\sum_{\substack{j\in S_n(k)\\ j\neq m}}\rho_{jr,L_1}(k) \tag{15}$$

$$B_{\rho,mr,L_2}(k)=\rho_{r,L_2}(k)-\frac{1}{M(k)-1}\sum_{\substack{j\in S_n(k)\\ j\neq m}}\rho_{jr,L_2}(k) \tag{16}$$

$$B_{\phi,mr,L_1}(k)=\phi_{r,L_1}(k)-\frac{1}{M(k)-1}\sum_{\substack{j\in S_n(k)\\ j\neq m}}\phi_{jr,L_1}(k) \tag{17}$$

$$B_{\phi,mr,L_2}(k)=\phi_{r,L_2}(k)-\frac{1}{M(k)-1}\sum_{\substack{j\in S_n(k)\\ j\neq m}}\phi_{jr,L_2}(k) \tag{18}$$

$$B_{\hat{N}_w,mr}(k)=\hat{N}_{r,w}(k)-\frac{1}{M(k)-1}\sum_{\substack{j\in S_n(k)\\ j\neq m}}\hat{N}_{jr,w}(k) \tag{19}$$

3.4 阈值计算

为了进行故障检测，多参考一致性检验根据无故障条件下的概率设置一个检测阈值。每个通道上各观测量的 B 值服从均值为零的高斯分布，每个通道上的测量值 P（包括原始伪距观测量、载波相位观测量和宽巷模糊度估计）的 B 值检测阈值 $D_{P,\mathrm{th}}$ 则等于每个 B 值的标准差乘上一个放大因子 f，即

$$D_{P,\mathrm{th}}=f\sigma_{B,P} \tag{20}$$

式中：放大因子 f 由高斯概率密度函数的拖尾决定。

$$\sigma_{B,P}=\frac{1}{\sqrt{M(k)-1}}\sigma_P \tag{21}$$

式中：σ_P 为原始伪距观测量、载波相位观测量和宽巷模糊度估计的均方差。

如果某个参考站的 B 值超过阈值，表明参考站接收机对应的测量值误差较大，则判定该参考站发生故障，将其标记为不可用。

3.5 决策执行逻辑

海基 JPALS 多参考一致性检验故障检测流程和故障隔离程序与 LAAS 类似[15]，但两者涉及的观测量及 B 值不同。

3.5.1 故障检测流程

海基 JPALS 多参考一致性检验故障检测流程如图 4 所示。在计算出各测量值的 B 值后，将其与阈值比较。如果每个测量值的B值都没有超过阈

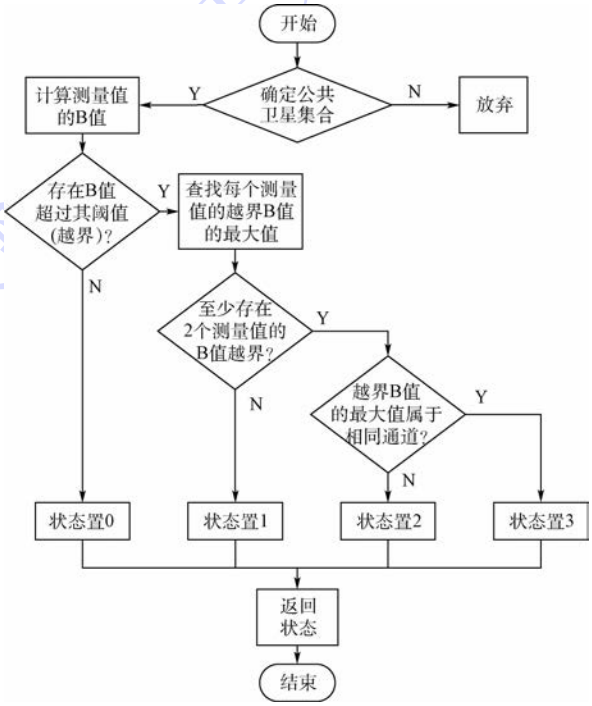


图 4 海基 JPALS 多参考一致性检验故障检测流程图

Fig. 4 Procedure diagram of MRCC fault detection for sea-based JPALS

值,则保持当前监测器状态为 0。如果发现 B 值超过阈值,则查找每个测量值超过阈值最大的 B 值。进一步判断是否有多个测量值的 B 值超过阈值,如果不是,返回状态 1,如果是,则继续判断每个超过阈值的 B 值最大值是否都处于相同通道,如果不是,返回状态 2,如果是,返回状态 3。

3.5.2　故障隔离程序

海基 JPALS 多参考一致性检验故障隔离程序流程如图 5 所示。在检测到故障状态后,进一步对存在故障的通道进行隔离处理。如果故障检测返回状态 0,表示没有 B 值越界,多参考一致性检验故障隔离程序不做任何处理。如果返回状态为 2,表明最大 B 值不在同一通道,但故障隔离程序无法处理这一类复杂情形,只能将这类处理任务传递给下一阶段更加复杂的执行监测过程[15]。对于所有其他情况,监测器将剔除包含最大 B 值的通道,然后基于剩余通道重新计算 B 值。重复该过程,直到不再出现多参考一致性检验故障标志,则故障隔离程序结束。

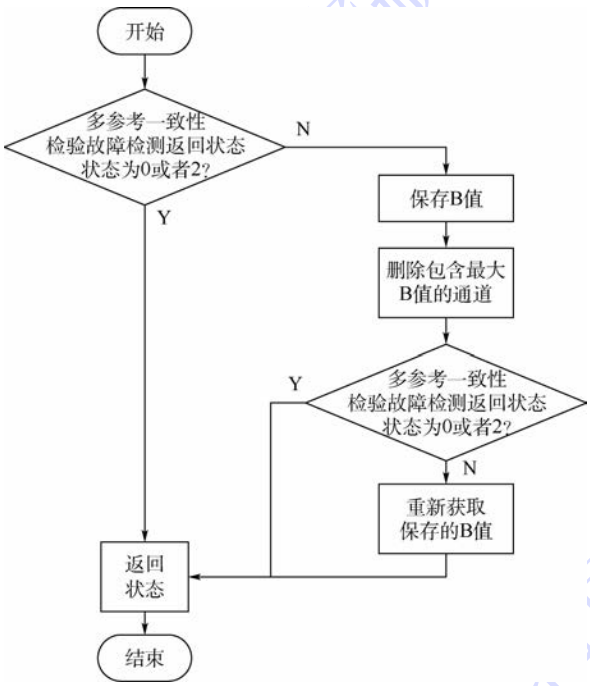


图 5　海基 JPALS 多参考一致性检验故障隔离程序流程图

Fig. 5　Procedure diagram of MRCC fault isolation for sea-based JPALS

4　实 验

4.1　实验设计

开展多参考一致性检验实验受到 2 个方面的条件限制:①找到一个足够大尺寸的移动平台架设 4 个参考站非常困难;②多参考一致性检验过程中出现 B 值超过阈值的异常是一个极小概率事件,需要数月或者更长时间的监测。考虑到多参考一致性检验的算法仅关心平台上接收机之间的相对位置[1],利用静止平台进行多参考一致性检验的初步实验是可接受的,因此,实验中将多个参考站架设在静止的房顶上,如图 6 所示。此外,为初步验证算法的理论可行性,实验仅采集了短期观测数据(1 Hz 的频率连续采集 24 h),并通过人为干扰方式模拟严重多径,而长时间的真实故障捕捉观测实验验证考虑在后续研究中进行。



图 6　多参考一致性检验数据采集实验设备

Fig. 6　MRCC data capture experimental equipment

4.2　数据分析

以实验过程中采集的 L_1 双差载波相位观测量为例进行 B 值及其阈值的计算。海基 JPALS 进行定位解算时,挑选当前几何构型最优的一批卫星进行运算。根据第 3 节的算法分别计算 4 个参考站接收机的 B 值,再根据阈值计算公式求解得到 B 值阈值。图 7 给出了各参考站接收机载波相位测量值的 B 值及其阈值。

从图 7 的实验数据处理结果可以看到,01:13 之前,各参考站无故障,其 B 值均小于阈值;01:13 之后,通过人为干扰,参考站 4 的接收机出现了 B 值超过阈值的情况,01:18 之后,解除人为干扰,参考站 4 的接收机数据又恢复正常。

实验表明,当通过本文算法实际计算得到的参考站接收机的 B 值超过阈值时,采用图 4 和图 5所描述的故障检测流程可以有效检测到参考站 4 所发生的故障(由于人为干扰导致的严重多径误差),并对其进行隔离。而根据计算式(20)得到的阈值也能够在排除所有 B 值超过阀值的接收机同时,保留未被隔离的接收机进行舰机双差定位运算,满足海基 JPALS 的多参考一致性检验完好性风险指标要求。

北航学报 赠阅

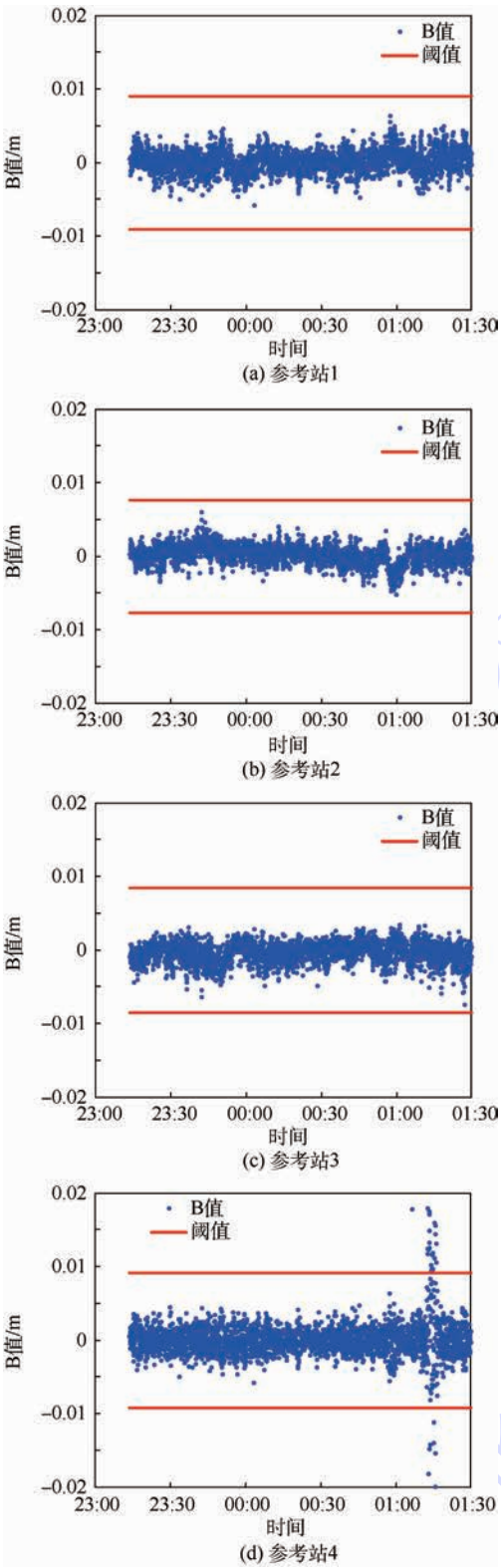


图 7 各参考站 L_1 载波相位的 B 值及其阈值

Fig. 7 B-value and threshold of L_1 carrier phase on each reference station

5 结 论

本文在分析海基 JPALS 的两阶段导航算法的基础上，提出了多参考一致性检验算法，并进行了模拟实验验证，总结如下：

1）提出计算 L_1/L_2 的伪距和载波相位观测量、宽巷模糊度估计 5 个测量值的 B 值，建立了海基 JPALS 多参考一致性检验的测量值体系。

2）利用各参考站与理想着舰点之间相对位置矢量已知的特性计算各测量值的 B 值，并通过高斯密度函数计算其阈值，理论上可以保证满足海基 JPALS 的多参考一致性检验完好性风险指标要求。

3）提出了适合海基 JPALS 的多参考一致性检验故障检测流程和故障隔离程序。

4）模拟实验结果验证了算法的可行性，当发生测量值 B 值越界时，可检测到故障通道，并及时检测隔离。

后续将进一步研究该算法在海基 JPALS 完好性监测系统其他模块中的应用，例如测量值质量监测，从而为海基 JPALS 的完好性监测体系的完善提供进一步的技术参考。

致谢 感谢陈杰博士提供的文献帮助和景少玲博士对论文提出的修改意见。

参考文献（References）

[1] PETERSON B R, PULLEN S, PERVAN B, et al. Investigation of common architectures for land- and sea-based JPALS[C]// Proceedings of the 18th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2005). Long Beach, CA: ION, 2005: 26-37.

[2] Draft system requirements document for joint precision approach and landing system (JPALS) increment 1: PMA2135-0100/R-5[R]. USA: DoN, 2007.

[3] RIFE J, KHANAFSEH S, PULLEN S, et al. Navigation, interference suppression, and fault monitoring in the sea-based joint precision approach and landing system[J]. Proceedings of the IEEE, 2008, 56(12): 1958-1975.

[4] 杨文辉，周保军. LAAS 地面系统中多基准一致性检测算法的分析[J]. 电子学报，2006，34(3)：469-471.
YANG W H, ZHOU B J. Analysis of multiple reference consistency check on LAAS ground facility[J]. Acta Electronica Sinica, 2006, 34(3): 469-471(in Chinese).

[5] 朱衍波，张晓林，薛瑞，等. 民航 GPS 地基区域完好性监视系统设计与实现[J]. 北京航空航天大学学报，2006，32(7)：797-801.
ZHU Y B, ZHANG X L, XUE R, et al. Design and implementation of GPS ground-based regional integrity monitoring system for China civil aviation[J]. Journal of Beijing University of Aeronautics and Astronautics, 2006, 32(7): 797-801(in Chinese).

[6] 王志鹏，张军，刘强. LAAS 地面站伪距纠正量误差完好性监测算法研究[J]. 电子与信息学报，2009，31(1)：196-200.
WANG Z P, ZHANG J, LIU Q. The study of integrity monitor

algorithm of the pseudorange correction error on LAAS ground facility[J]. Journal of Electronics & Information Technology, 2009,31(1):196-200(in Chinese).

[7] XIE G,PULLEN S,LUO M,et al. Integrity design and updated test results for the stanford LAAS integrity monitor testbed [C]// Proceedings of the 57th Annual Meeting of the Institute of Navigation. Albuquerque,NM:ION,2001:681-693.

[8] Performance type one local area augmentation system ground facility:FAA-E-2937[R]. USA:FAA,1999.

[9] KOENIG M,RIFE J,PULLEN S,et al. Optimizing channel selection for the JPALS' land-based integrity monitor[C]// Proceedings of the 18th International Technical Meeting of the Satellite Division of The Institute of Navigation(ION GNSS 2005). Long Beach,CA:ION,2005:62-72.

[10] MATHER C,MACAULAY A,MOLE S. Performance of integrity monitoring techniques for shipboard relative GPS landing systems[C]// Proceedings of the 18th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2005). Long Beach,CA:ION,2005:2882-2894.

[11] DOGRA S,HANSEN J W J. Sea-based JPALS relative navigation algorithm development[C]// Proceedings of the 18th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2005). Long Beach,CA:ION, 2005:2871-2881.

[12] DE LORENZO D S,ALBAN S,GAUTIER J,et al. GPS attitude determination for a JPALS testbed integer initialization and testing[C]// Position Location and Navigation Symposium,2004 (PLANS 2004). Piscataway,NJ:IEEE Press,2004:762-770.

[13] TEUNISSEN P J G,ODIJK D,JOOSTEN P. A probabilistic evaluation of correct GPS ambiguity resolution[C]// Proceedings of the 11th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS 1998). Nashville,TN:ION,1998:1315-1323.

[14] FENG S,OCHIENG W,SAMSON J,et al. Integrity monitoring for carrier phase ambiguities[J]. Journal of Navigation,2012, 65(1):41-58.

[15] XIE G. Optimal on-airport monitoring of the integrity of GPS-based landing systems[D]. Palo Alto:Stanford University,2004.

作者简介:

冯涛　男,博士研究生。主要研究方向:卫星导航。

Tel.:010-65299467

E-mail:fengtao2012@buaa.edu.cn

毛峡　女,博士,教授,博士生导师。主要研究方向:计算机视觉、模式识别和人机交互。

Tel.:010-82339508

E-mail:moukyou@buaa.edu.cn

宋成　男,博士。主要研究方向:卫星导航。

Multiple reference consistency check algorithm for sea-based JPALS

FENG Tao[1,2], MAO Xia[1,*], SONG Cheng[3]

(1. School of Electronic and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;
2. China Transport Telecommunications & Information Center, Beijing 100011, China;
3. China Satellite Navigation Office, Beijing 100034, China)

Abstract: According to the characteristics of mobile carrier reference station, a novel multiple reference consistency check (MRCC) algorithm for the sea-based joint precision approach and landing system (JPALS) is proposed. First, the original pseudorange and carrier phase measurement of each reference station are converted to the ship reference point by using the attitude information from the inertial navigation system on the ship and the ship size measurement data. Then, the wide lane ambiguity of each reference station is obtained by filtering the converted geometry-free combined observables. Finally, the B-values of these converted measurements for each reference station are calculated. Under no fault condition, the B-values of these measurements distribute in a certain range, but the B-value may exceed the threshold in the case of reference station fault, severe multipath or radio interference. Observation data were collected by simulation experiment and then the B-values and their threshold were computed. The results show that generally the B-values of these observables are very small, but when any abnormal from some measurement occurs, it can be detected that the B-value exceeds their threshold. The algorithm satisfies the MRCC integrity risk requirements of the sea-based JPALS.

Key words: sea-based JPALS; multiple reference consistency check (MRCC); wide lane ambiguity; B-value; integrity

Received: 2015-10-19; **Accepted**: 2015-11-06; **Published online**: 2016-01-04 10:04

URL: www.cnki.net/kcms/detail/11.2625.V.20160104.1004.011.html

* **Corresponding author**. Tel.: 010-82339508　E-mail: moukyou@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0694

冲击位置对斜搭接接头冲击后拉伸性能的影响

聂恒昌[1], 徐吉峰[2], 关志东[1,*], 黎增山[1], 冀赵杰[1], 谭日明[1]

(1. 北京航空航天大学 航空科学与工程学院, 北京 100083; 2. 中国商飞 北京民用飞机技术研究中心, 北京 102211)

摘 要: 低速冲击损伤对复合材料胶接结构的力学性能存在威胁。对在搭接区域3个不同位置承受低速冲击的斜搭接接头进行冲击后拉伸试验,并分别建立胶接接头冲击及冲击后拉伸有限元模型。试验结果表明:冲击位置在搭接区尖端背侧,造成的凹坑深度最大,冲击后拉伸强度最低;冲击位置在搭接区中部,凹坑深度最小,冲击后拉伸强度最大;冲击位置在搭接区尖端侧,凹坑深度及冲击后拉伸强度居中。有限元分析结果表明胶层损伤和复合材料纤维损伤的大范围扩展是导致结构失效的主要原因。

关 键 词: 复合材料结构; 斜搭接接头; 胶接连接; 低速冲击; 拉伸性能
中图分类号: V214.8; TB330.1
文献标识码: A **文章编号**: 1001-5965(2016)11-2306-15

全下文载

胶接连接具有结构轻、效率高、抗疲劳和成本低等优点,在飞机复合材料结构中得到广泛应用[1]。对于厚度较大的结构,一般选用斜搭接或阶梯搭接,可以有效降低连接区剥离应力,提高结构承载能力[2]。目前广泛应用在飞机上的胶接挖补修理结构,也可看成是斜搭接胶接接头的三维形式。连接(或修理)结构在飞行及维护过程中,经常面对冲击损伤的威胁,为满足使用要求,在受到冲击后连接结构还应当可以传递一定的载荷。因此有必要对复合材料胶接结构的损伤容限特性进行研究,为飞机结构损伤容限设计提供依据。

国内外学者对复合材料胶接结构的损伤容限性能进行了相关研究。Sayman 等[3]通过试验研究了复合材料单搭接胶接接头在4种冲击能量下的剩余强度。Goh 等[4]通过拉伸试验及有限元模型对含不同胶层缺陷的斜搭接接头的损伤容限进行考察。Campilho 等[5]对复合材料单搭接与双搭接修理结构进行拉伸试验,考察了影响胶接修理效果和破坏模式的参数。Harman 和 Wang[6]通过试验及有限元方法研究了外搭接接头在高能量冲击下的损伤容限及冲击阻抗,研究表明接头尖端对于接头强度影响很大。刘遂等[7]对不同胶层厚度的复合材料双搭接接头拉伸强度进行了试验研究,并建立解析模型对其胶层剪切破坏进行研究。郭霞等[8]针对复合材料单搭接接头的拉伸性能进行有限元分析,研究搭接长度对单搭接接头的应力分布和破坏模式的影响。乔玉等[9]通过有限元对拉伸载荷下的双阶梯形胶接接头进行渐进损伤分析,预测了接头的损伤扩展方式和连接效率。刘凯[10]基于有限元法建立对接接头受埃佐冲击时的弹塑性模型,分析了各种因素对胶层的应力和应变响应的影响。吴坤岳[11]通过试验及仿真对振动劈裂载荷下的单搭接接头进行了疲劳特性研究,研究其疲劳损伤及疲劳后的断裂机理。可见目前已有的研究主要集中在单搭接接头、对接接头等方面,对斜搭接接头的力学性能以及对胶接连接结构冲击后剩余强度的研究相对较少。

收稿日期: 2015-10-26; **录用日期**: 2016-01-15; **网络出版时间**: 2016-03-15 14:28
网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20160315.1428.004.html
* **通讯作者**: Tel.: 010-82338873 E-mail: zdguan@buaa.edu.cn

引用格式: 聂恒昌, 徐吉峰, 关志东, 等. 冲击位置对斜搭接接头冲击后拉伸性能的影响[J]. 北京航空航天大学学报, 2016, 42(11): 2306-2320. NIE H C, XU J F, GUAN Z D, et al. Tensile behavior of scarf joints after impact in different locations[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(11): 2306-2320 (in Chinese).

北航学报 赠阅

本文对复合材料斜搭接胶接接头在不同位置开展冲击试验，通过测量凹坑深度考察冲击对结构造成的损伤，并进行拉伸试验考察结构的剩余强度，在试验基础上建立三维有限元模型，分析胶接接头的冲击损伤、冲击后拉伸载荷下的应力分布和渐进损伤过程，为研究胶接接头的损伤容限特性提供参考。

1　试　验

1.1　试验对象

本文设计了如图 1 所示的斜搭接胶接接头试验件，冲击位置为搭接区接头尖端（*A*）、搭接区中部（*B*）、搭接区接头尖端背侧（*C*）。胶接接头由 T800 级碳纤维预浸料制造，采用 J116B 胶粘剂进行连接，T800/3900-2 和 J116B 胶粘剂的各项材料力学性能分别如表 1 和表 2 所示。接头的铺层顺序为$[45/0/-45/90]_{2s}$，单层厚度为 0.188 mm，胶层厚度为 0.200 mm。试验件共制备 18 件。

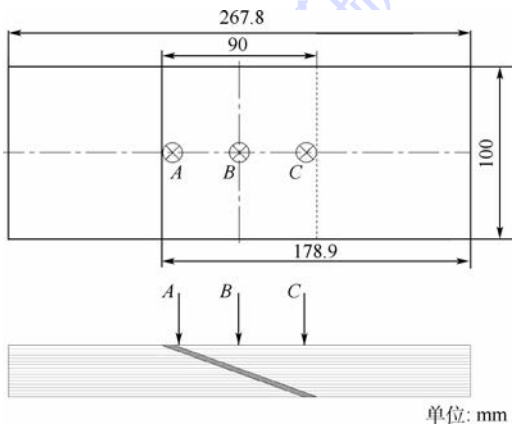


图 1　试验件基本尺寸及冲击位置图

Fig. 1　Basic dimensions and impact locations of specimens

表 1　T800/3900-2 力学性能

Table 1　Mechanical properties of T800/3900-2

参数	数值	参数	数值
E_{11}/GPa	156.50	ν_{23}	0.47
E_{22}/GPa	8.25	ν_{13}	0.32
E_{33}/GPa	8.25	X_T/MPa	2 986.30
G_{12}/GPa	4.37	X_C/MPa	1 508.00
G_{23}/GPa	2.81	Y_T/MPa	50.00
G_{13}/GPa	4.37	Y_C/MPa	220.00
ν_{12}	0.32	S/MPa	66.30

注：E_{11}，E_{22}，E_{33}—1，2，3 方向的弹性模量；G_{12}，G_{13}，G_{23}—1-2，1-3，2-3 平面内的剪切模量；ν_{12}，ν_{13}，ν_{23}—1-2，1-3，2-3 平面内的泊松比；X_T，X_C—纵向拉伸和压缩强度；Y_T，Y_C—横向拉伸和压缩强度；S—剪切强度。

表 2　J116B 胶粘剂力学性能

Table 2　Mechanical properties of J116B adhesive

参数	数值
E/GPa	1
G/GPa	0.385
ν	0.3
$t_n^0 = t_s^0 = t_t^0$/MPa	24.5
$G_n^C = G_s^C = G_t^C$/(N · mm^{-1})	7.5

注：E—弹性模量；G—剪切模量；ν—泊松比；t_n^0—法向强度；t_s^0，t_t^0—切向强度；G_n^C—法向临界断裂能量释放率；G_s^C，G_t^C—切向临界断裂能量释放率。

1.2　低速冲击试验

将试验件分为 3 组，每组 6 件，分别对 *A*、*B*、*C* 3 个冲击位置（如图 1 所示）进行低速冲击试验。试验设备为按照 ASTM D7136-07 试验标准[12]的要求制作的落锤式冲击试验机，通过调整冲头的下落高度可实现特定的冲击能量，如图 2所示，试验前利用 4 个夹钳将试验件夹持在带有 125 mm × 75 mm矩形开口的试验机底座上，实现四边固支的边界条件。

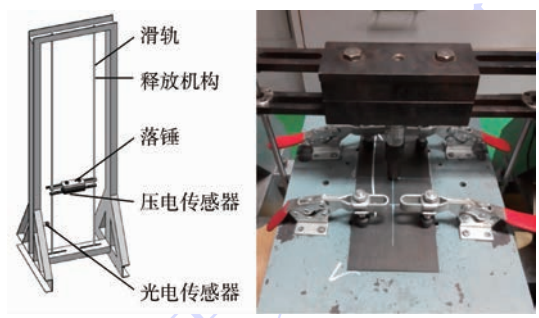


图 2　落锤式冲击试验机及冲击试验边界条件

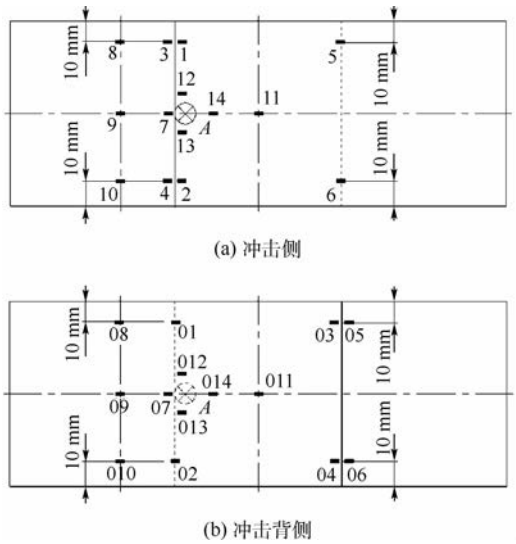
Fig. 2　Drop-weight impact testing device and boundary conditions of impact test

为确定冲击能量值，在正式试验前，按照 ASTM D7136-07 标准[12]，对尺寸为 150 mm × 100 mm、材料和铺层顺序与胶接接头相同的完好层板进行冲击试验，认为冲击后造成 1 mm 深度的凹坑为目视勉强可见冲击损伤（BVID）。经多次试验测得造成 BVID 的冲击能量值为 44 J，故正式试验中以该能量值对斜搭接胶接接头试件进行冲击。

1.3　冲击后拉伸试验

使用时代试金 WDW-200E 静力试验机，对冲击后试验件施加轴向拉伸载荷进行测试。试验温度为室温（23 ± 3℃），相对湿度为 50% ± 10%，采用 1 mm/min 的速度进行连续加载直至试验件破坏。加载过程中，通过粘贴在冲击位置附近、接头尖端附近等处的电阻应变片采集应变数据，考察试验件表面的应变分布情况。以冲击位置为 *A* 的试验件为例，应变片布置方案如图 3 所示。

北航学报 赠阅



1 ~ 14—冲击侧应变片;01 ~ 014—冲击背侧应变片。

图 3 在 *A* 位置受冲击的试验 件应变片布置方案图

Fig. 3 Layout scheme of strain gages on specimens impacted in location *A*

1.4 试验结果及分析

1.4.1 冲击后损伤

图 4 为冲击后的试验件正反面形貌及 C 扫描图像,可见冲击表面形成了永久凹坑,冲击背面则发生了明显的劈裂现象。冲击位置为 *A*、*B* 时,严重损伤面积(黑色部分)较小,背部劈裂较为明显,冲击位置为 *C* 时,严重损伤面积较大,背部劈裂较不明显。从 C 扫描图像可见,冲击位置在 *C* 处的胶接接头与完好层板的损伤形状都呈椭圆形,损伤面积也较为接近。

文献[13]的研究表明,在损伤面积、损伤宽度、凹坑深度等参数中,只有凹坑深度能较好地反映复合材料层板抵抗冲击事件的能力,因此通过测量各试验件冲击后的凹坑深度来分析冲击损伤程度。如图 5 所示,可见冲击位置在 *A*、*B* 处时平均凹坑深度略小于 1 mm,而在 *C* 处可达 1.8 mm。为进一步观测冲击对接头造成的损伤,将冲击后试验件沿通过冲击点的纵向中轴线剖开进行电镜扫描,如图 6 所示。从扫描电镜(SEM)照片可见,完好层板受到冲击时,内部出现了明显的纤维断裂、基体开裂以及分层损伤。冲击位置在 *A* 处及 *B* 处,由于胶层的缓冲作用,相当于增大了冲击接头的接触面积,因此各铺层纤维断裂不明显,特别是主要承载的 0°层损伤不大。而冲击位置在 *C* 处,由于胶层在最下方,基本没有起到缓冲作用,在冲头的直接冲击作用下,复合材料各个铺层均发生了严重的纤维断裂和明显的分层损伤。对比完好层板与冲击位置在*C*处的胶接接头的

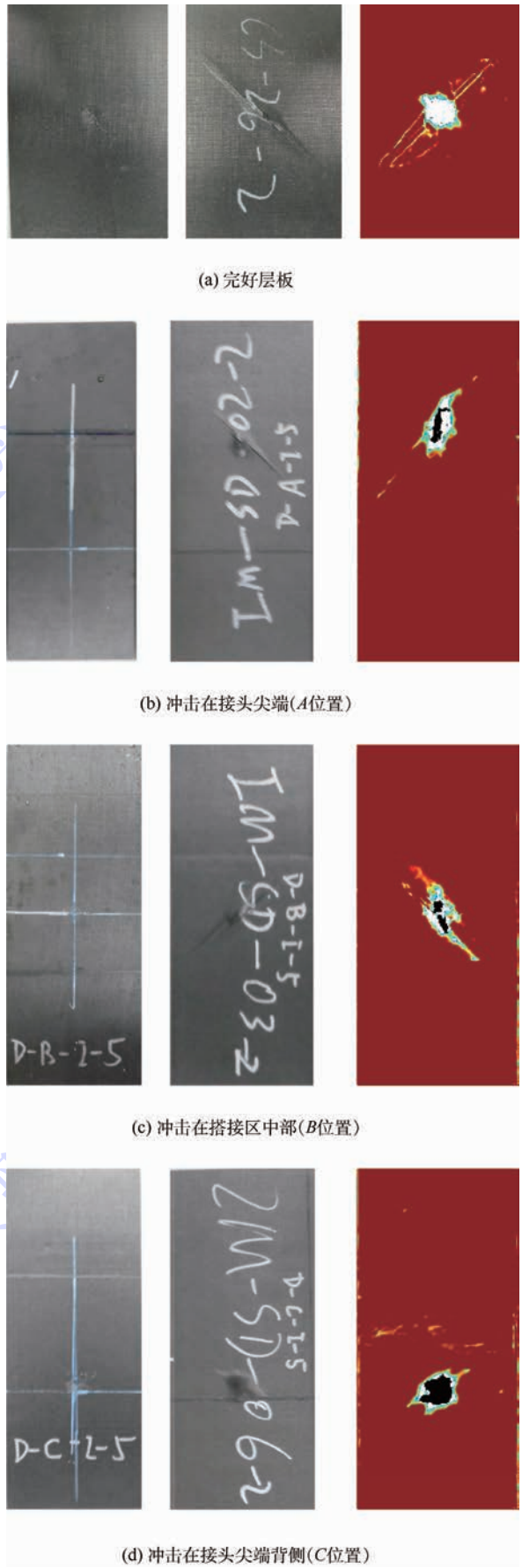


图 4 冲击后的试验件正反面形貌及 C 扫描图像

Fig. 4 Front and back photos and C-scan graphs of specimens after impact

北航学报 赠阅

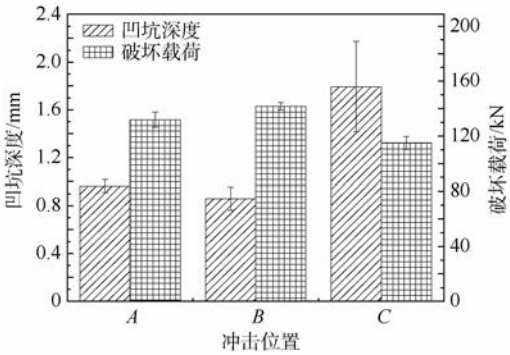


图 5　试验件的冲击凹坑深度及冲击后拉伸破坏载荷

Fig. 5　Indentation depth and tensile failure load of specimens after impact

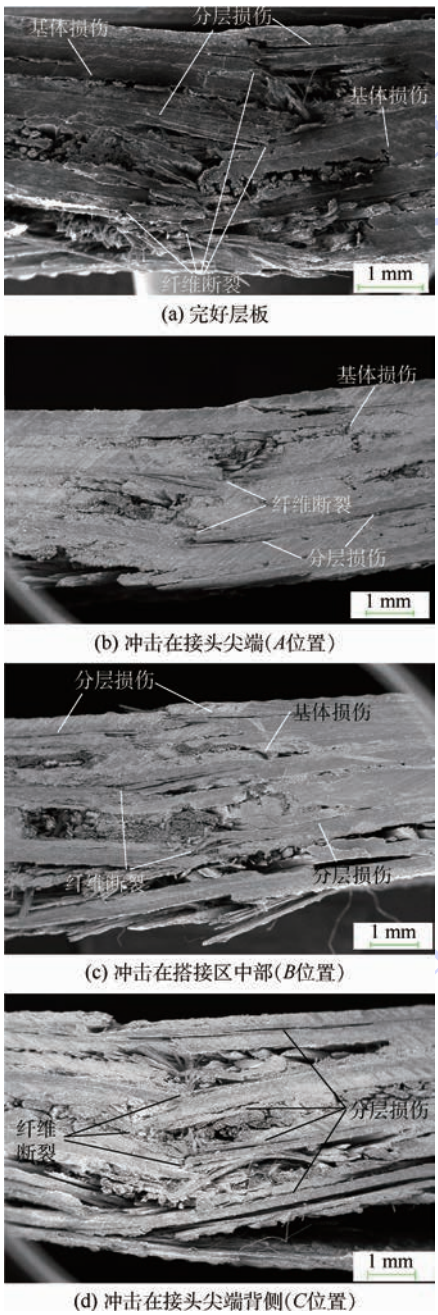


图 6　冲击后试验件纵剖面扫描电镜照片

Fig. 6　SEM photos along cut-up line of specimens after impact

SEM 照片，可见两种结构的内部均出现了较严重的纤维断裂，但后者的分层损伤更为严重，且在胶接接头尖端处发生了明显的剥离。

1.4.2　冲击后拉伸破坏模式

试验件冲击后拉伸的破坏模式主要有两种：①胶层破坏或胶层与复合材料之间的界面破坏，导致接头剥离，如图 7(a)所示；②接头没有发生剥离，试验件在搭接区域被整体拉断，如图 7(b)所示。

各组试验件的破坏模式统计如表 3 所示(其中每组各有一件没有进行冲击后拉伸试验)。

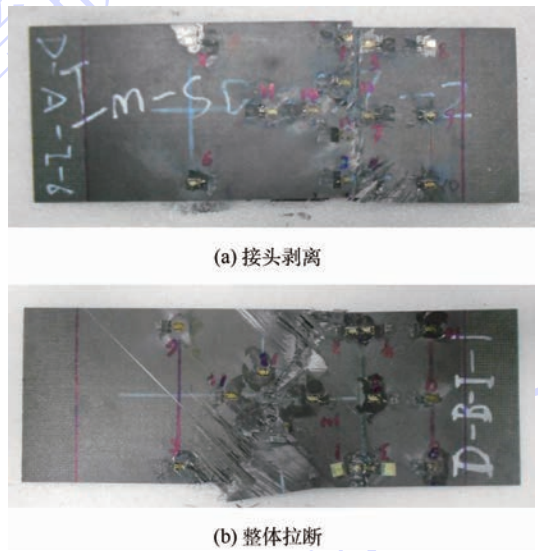


图 7　试验件两种典型的破坏模式照片

Fig. 7　Photos of two typical failure modes of specimens

表 3　各试验件的破坏模式统计

Table 3　Failure mode statistics of different specimens

试验件编号	*A*	*B*	*C*
1	整体拉断	整体拉断	接头剥离
2		整体拉断	
3	整体拉断	整体拉断	接头剥离
4	整体拉断	整体拉断	整体拉断
5	整体拉断		整体拉断
6	接头剥离	整体拉断	接头剥离

从表 3 得知，出现破坏模式①的试验件不多，主要原因是胶层强度不足、连接界面上产生较大的剥离应力导致；出现破坏模式②的试验件较多，主要是由于冲击造成的复合材料损伤引起，这种破坏模式说明试验件胶层强度足够大，结构充分发挥了复合材料的承载作用，达到了较为理想的胶接效果[2]。

1.4.3　冲击后拉伸剩余强度

各组试验件进行冲击后拉伸试验的平均破坏载荷对比如图 5 所示。

冲击位置在搭接区尖端侧(*A* 位置)时，破坏

北航学报 赠阅

模式以试验件整体拉断为主,损伤从冲击点附近开始扩展,最终试验件在接头尖端附近断裂破坏。由于冲击位置距离胶层界面很近,胶层能对冲击起一定的缓冲作用,破坏载荷介于 *B* 位置和 *C* 位置之间。

冲击位置在搭接区中部(*B* 位置)时,破坏模式均为试验件整体拉断,损伤从冲击点开始沿试验件中部短轴扩展,最终试验件直接从中部断裂。由于胶层吸收了较多的冲击能量,接头的复合材料纤维断裂较少,因此结构的破坏载荷最高。

冲击位置在搭接区尖端背侧(*C* 位置)时,破坏模式既有接头剥离,又有整体拉断。由于胶层对冲击基本不起缓冲作用,冲击侧的接头复合材料纤维损伤较为严重,导致结构承载不对称而引起冲击背侧的接头尖端处产生较大的弯曲应力,损伤从该处开始沿搭接界面扩展,使接头发生剥离破坏;而破坏模式为整体拉断的试验件,也因为纤维损伤严重导致结构承载能力较低。

1.4.4 冲击后拉伸应变分析

以冲击位置为搭接区尖端侧(*A* 位置)的试验件为例,绘制试验件拉伸过程中的应变-载荷曲线,如图 8 所示。

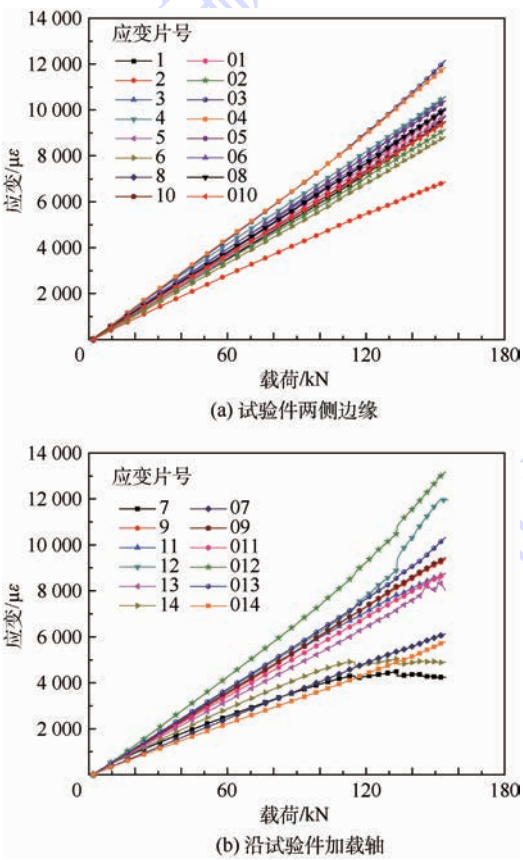


图 8 试验件在 *A* 位置受冲击后拉伸的应变-载荷曲线

Fig. 8 Tensile strain-load curves of specimens impacted in location *A*

从图 8(a)可见,分布在试验件两侧边缘的各应变片,其测得的应变值均随载荷增大而线性增长。其中,在冲击反面接头尖端内侧的 03、04 号应变片应变值较大,冲击面接头尖端内侧的 1、2 号应变片应变值较小,说明冲击造成了结构承载不对称,冲击背侧接头尖端处产生了明显的应力集中。

从图 8(b)可见,沿试验件加载中轴线附近分布的各应变片,其测得的应变值也大致随载荷增大而线性增长。其中冲击区沿加载方向两侧的 7/07、14/014 号应变片应变值最小,冲击区旁垂直于加载方向两侧的 12/012、13/013 号应变片应变值最大,原因是冲击造成的纤维断裂使通过冲击区的传力路径被截断,相应地冲击区两侧的旁路载荷增大;9/09、11/011 号应变片分别测量冲击区域两侧的远场应变,它们的应变值居中,可见离冲击区较远处没有产生明显的应力集中。

综上,试验件应力集中的主要区域为搭接区尖端位置及冲击区旁垂直于加载方向的位置。

2 有限元分析

2.1 理论基础

2.1.1 纤维断裂的引入

相关研究[14]表明,层板在低速冲击下的凹坑深度-冲击能量曲线存在拐点,如图 9 所示。拐点的出现对应于较多的纤维断裂开始在层板内发生,此后层板基本失去了继续抵抗冲击的能力,凹坑深度随即快速增长,故拐点现象表明复合材料层板对冲击事件的抵抗能力发生了突变。通常 BVID(深度为 1.0 mm)凹坑位于拐点之后,此时材料出现严重纤维断裂现象,凹坑周围的大量材料进入塑性变形区,因此,为了对进入拐点后冲击能量范围下的凹坑深度进行预测以形成较完整的曲线,引入纤维断裂建立了基于各向异性弹塑性理论的有限元模型。模型中通过纤维方向的最大应力准则判断是否发生纤维断裂,若是则对材料性能进行折减。

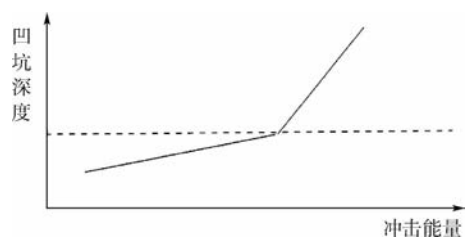


图 9 低速冲击下层板凹坑深度-冲击能量曲线示意图

Fig. 9 Schematic diagram of indentation depth-impact energy curve of laminates under low-velocity impact

北航学报 赠阅

2.1.2　层间分层的模拟

层间分层是复合材料层板在低速冲击过程中的主要损伤模式之一，也是能量耗散的重要方面，因此模型中通过在单层间插入界面单元对层间分层进行模拟，其中采用了线弹性-线性软化（也称作双线性）牵引力-相对位移本构关系建立界面单元[15]，采用基于能量释放率的幂准则（power law）判断界面单元的最终破坏[16]。

2.1.3　材料失效准则

模型中使用连续损伤准则[15]判断胶层的损伤，使用三维 Hashin 损伤准则[17]分别判断复合材料结构的纤维失效和基体失效。具体实现方法是在 ABAQUS 中编写用户自定义场变量子程序 USDFLD 进行损伤判断，在损伤出现后，对相应的材料性能参数进行折减。

2.2　有限元模型

2.2.1　冲击试验模型

复合材料单层采用 ABAQUS 中的三维 8 节点实体单元（C3D8）建立，这种单元类型可以有效地模拟复合材料的拉伸、压缩以及剪切状态，能够根据其单元应力（或应变）判断单元的纤维损伤，并且实现损伤的累积，对材料的弹性常数进行连续折减。胶层采用界面单元（COH3D8）建立，主要承受法向拉伸、压缩以及剪切载荷，通过界面单元损伤演化理论判断分层损伤。

为准确模拟冲击试验中的边界条件，分别以 A、B、C 为中心截取 125 mm × 75 mm 的方形区域，对该区域边界施加固支约束，并将冲击点附近单元网格作细化处理以保证分析精度。冲头采用解析球面刚体进行模拟，其直径为 16 mm，集中质量大小为 5.36 kg。冲头 x 方向与 y 方向的平移与3 个方向上的转动自由度均被约束，只留 z 方向平移自由度以施加初始冲击速度。如图 10 所示是冲击位置为 A 点对应的有限元模型。

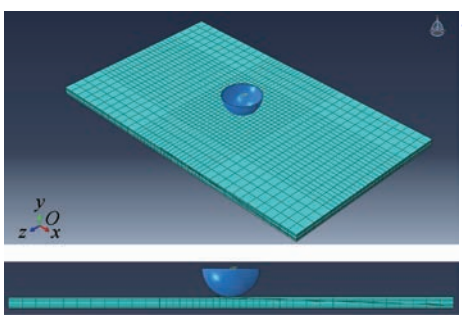


图 10　冲击位置为 A 点的斜搭接接头有限元模型

Fig. 10　Finite element models of scarf joints impacted in location A

整个冲击过程采用 ABAQUS/EXPLICIT 显式积分方式进行求解，总时间为 0.008 s（与试验实际冲击时间接近），此阶段需判断纤维及分层损伤，并对材料弹性常数进行折减。计算中记录了层板冲击点在冲击以及后续过程中的速度与变形历程，同时记录冲头与层板间的接触力，并输出层板各单层的纤维损伤、基体损伤以及各层间的分层损伤。

2.2.2　冲击后拉伸试验模型

建立有限元模型考察斜搭接接头冲击后的拉伸力学性能。为模拟冲击造成的复合材料损伤和胶层损伤，将模型中冲击位置周围的复合材料切割出与冲头大小相当的区域进行材料性能折减，并根据冲击试验后 C 扫描所得的分层面积，将冲击位置周围的胶层切割出相应大小的区域进行材料性能折减。计算表明，材料性能折减系数取0.05时计算结果与试验较为吻合。

如图 11 所示，将模型一端所有自由度限制以作为固支端，另一端沿加载轴方向施加 5 mm 的位移载荷。拉伸加载过程采用 ABAQUS/EXPLICIT 显式积分方式进行求解，输出加载端的位移和载荷，以及各部分的应力、应变、胶层损伤、复合材料纤维及基体损伤。

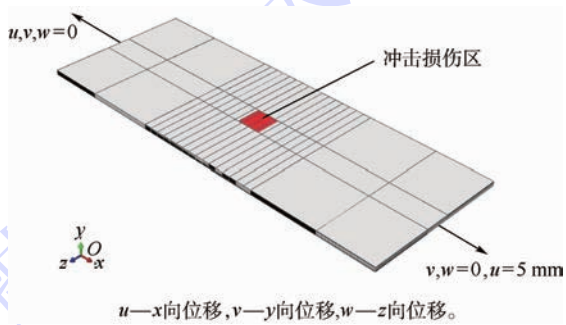


图 11　冲击后拉伸斜搭接接头有限元模型边界条件

Fig. 11　Boundary conditions of finite element models for tensile scarf joints after impact

2.3　有限元计算结果

2.3.1　冲击响应及冲击损伤分析

冲击有限元模型计算得到的 A、B、C 3 个冲击位置对应的接触力-时间、接触力-位移以及吸收能量-时间曲线分别如图 12（a）、图 12（b）和图 12（c）所示。从图 12（a）可见，A、B、C 3 个位置冲击的接触力-时间曲线大致相同，冲击位置为 C 时，最大接触力略大于 A 和 B，响应时间也

北航学报 赠阅

略短;从图 12(b)可见,冲击位置为 C 时,冲击过程中得到的最大冲头位移明显大于 A 和 B;从图 12(c)可见,冲击位置为 C 时,冲击后层板吸收的能量明显大于 A 和 B。综上,从各冲击响应曲线可得知,冲击位置在搭接区尖端背侧(C 位置)时,由于胶层的缓冲作用较小,冲击对搭接接头造成了较为严重的损伤。

有限元计算得到的冲击位置为 A、B、C 3 点对应的凹坑深度与试验结果对比如图 13 所示。由图13可以看出,计算结果与试验结果吻合良好,说明有限元模型能准确模拟冲击对胶接接头造成的损伤。

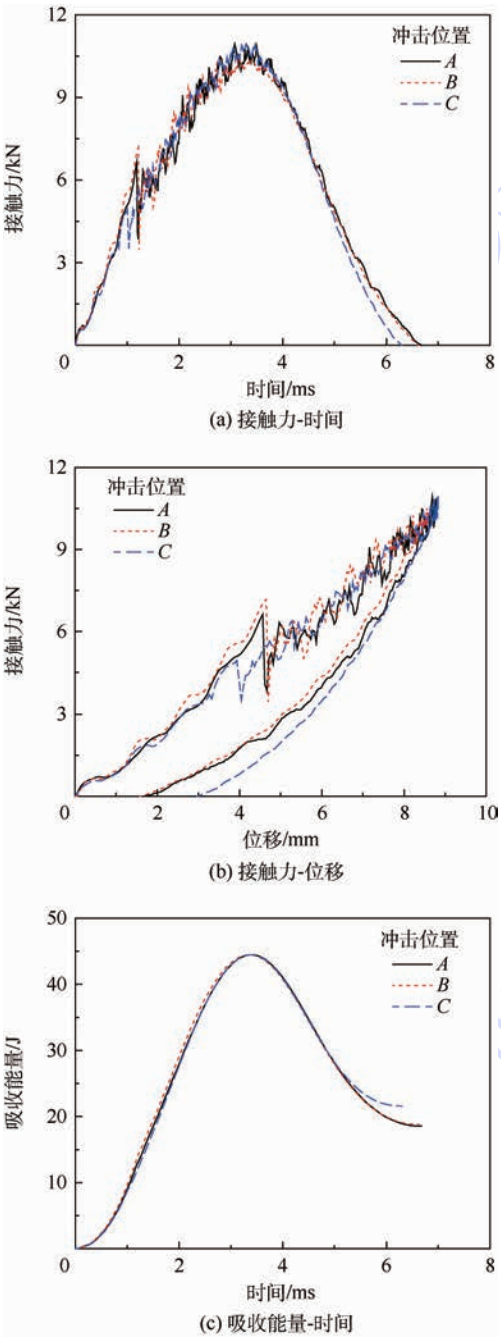


图 12 有限元模型的冲击响应曲线

Fig. 12 Impact response curves of finite element models

有限元计算得到的冲击位置为 A、B、C 3 点时纵剖面上的纤维断裂损伤区域如图 14(a)、图 14(b)和图 14(c)中红色部分所示。与 SEM 结果对比可得,胶接接头受到不同位置冲击时均出现了大范围纤维断裂。冲击在 A、B 位置时由于胶层吸收了一部分冲击能量,纤维断裂相对较少,主要出现在纵剖面靠近冲击背面的铺层处,而冲击在 C 位置时由于胶层基本不起缓冲作用,纤维断裂较为严重,在纵剖面各铺层内均有分布,分布面积也较大。

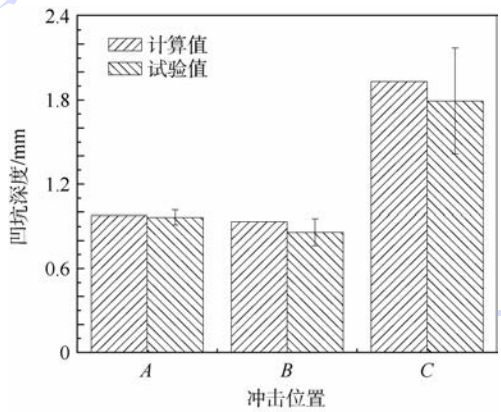


图 13 试验结果与有限元计算的凹坑深度数据对比

Fig. 13 Comparison between finite element calculation and test results of indentation depth

有限元计算得到的冲击位置为 A、B、C 3 点时纵剖面分层损伤区域如图 15(a)、图 15(b)、和图 15(c)中红色部分所示。与 SEM 结果对比可得,胶接接头受到不同位置冲击时均出现了大面积分层损伤,从有限元结果看,冲击在 3 个不同位置时,胶接接头内部的分层损伤面积大致接近。

有限元模型计算得到的冲击位置为 A、B、C 3 点时纵剖面上的基体损伤区域如图 16 中红色部分所示。可见胶接接头受到不同位置冲击时均出现了大范围基体损伤,从有限元结果看,冲击在 3 个不同位置时,胶接接头内部的基体损伤面积接近,其中冲击表面及靠近冲击背面的基体损伤面积较大,中部的基体损伤面积相对较小,而从纵剖面 SEM 结果中也可看出,基体损伤主要出现在纵剖面上部靠近冲击表面的位置。

北航学报 赠阅

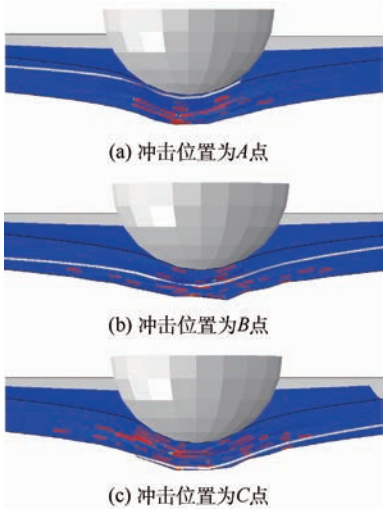


图 14　有限元模型的纤维断裂损伤区域

Fig. 14　Fiber breakage damage areas of finite element models

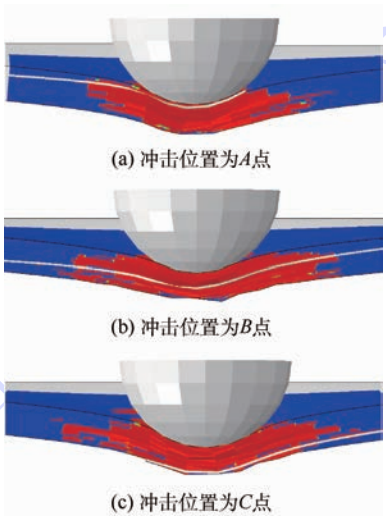


图 15　有限元模型的分层损伤区域

Fig. 15　Delamination damage areas of finite element models

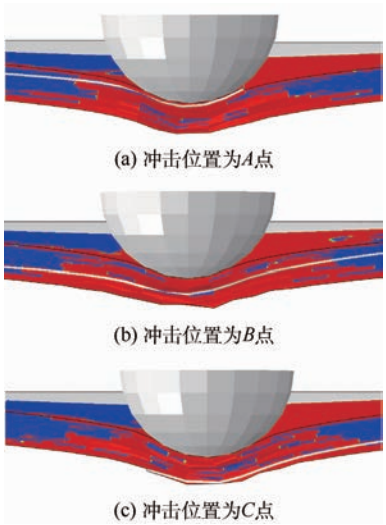


图 16　有限元模型的基体损伤区域

Fig. 16　Matrix damage areas of finite element models

2.3.2　冲击后拉伸强度分析

胶接接头在 3 个不同位置受冲击后的拉伸破坏载荷有限元计算与试验结果的对比如图 17 所示，可见胶接接头冲击后拉伸破坏载荷的计算结果与试验结果接近，当冲击位于搭接区尖端背侧时其剩余拉伸强度最低，冲击位于搭接区中部时剩余拉伸强度最高，故不同的冲击位置对胶接接头结构的强度有较大影响。

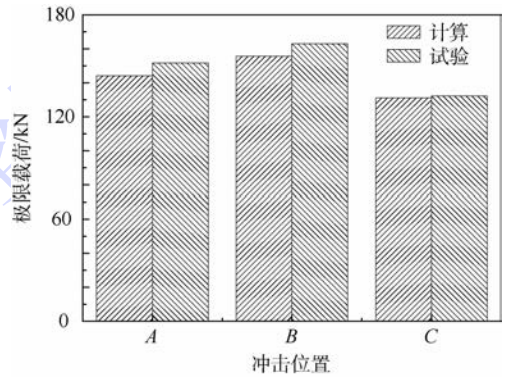


图 17　接头冲击后拉伸破坏载荷有限元计算与试验结果对比

Fig. 17　Comparison between finite element calculation and test results of tensile failure loads after impact

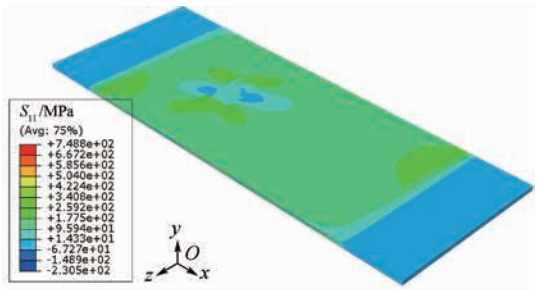
2.3.3　冲击后拉伸应力分布分析

通过应力云图分析不同冲击位置对胶接接头的应力分布和应力水平的影响差异，进而判断应力集中位置（通常也是损伤起始位置）。对冲击位置在 A、B、C 3 处时模型在承受相同拉伸载荷（约 45.5 kN）且非冲击区域未发生损伤时的应力情况进行研究，图 18、图 19 和图 20 分别为各模型整体的 S_{11}、S_{22}、S_{12} 3 个应力分量、0°铺层的 S_{11} 应力分量和胶层的 Mises 应力分量（S_{Mises}）的应力分布云图，表 4 为胶接接头整体各主要应力分量最大值。

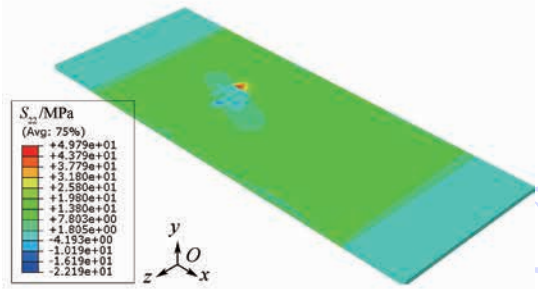
从图 18 ~ 图 20 和表 4 中可以看出，不同冲击位置对胶接接头冲击后拉伸的应力分布和应力水平的影响大致相同，即：冲击损伤区边缘均出现应力集中现象，各应力分量最大值均出现在应力集中区域中，且 S_{11} 应力远大于 S_{22} 和 S_{12} 应力；0°铺层中 S_{11} 应力水平最高，应力集中区为冲击损伤区边缘垂直于加载轴方向的位置，与试验中测得应变最大值的区域一致；胶层应力呈条带状分布，与 0°铺层胶接的区域应力水平较高，应力最大值出现在冲击损伤区边缘的应力集中区内。

不同冲击位置对应力情况的影响差异主要是：冲击位置为搭接区尖端背侧（C 位置），接头整体 S_{11} 应力水平较高，各应力分量最大值也高于

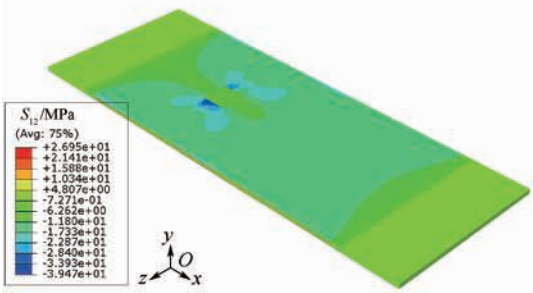
北航学报 赠阅



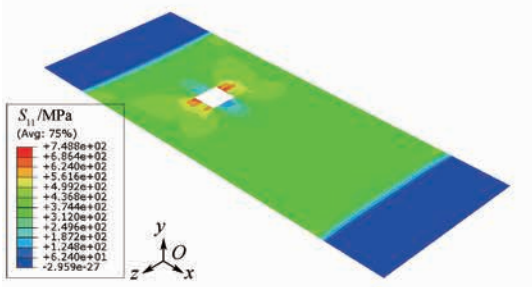
(a) S_{11} 分布云图



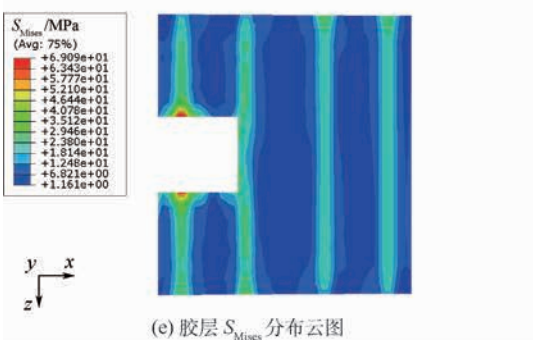
(b) S_{22} 分布云图



(c) S_{12} 分布云图



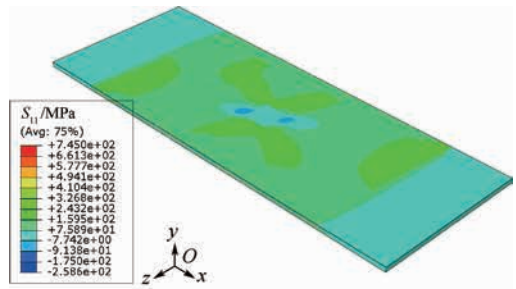
(d) 0°铺层 S_{11} 分布云图



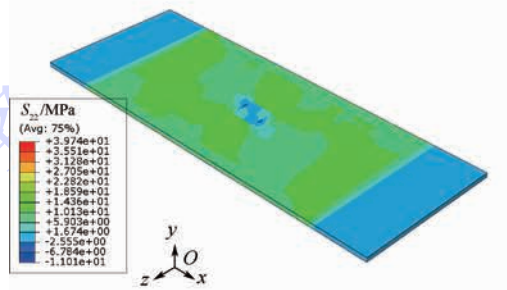
(e) 胶层 S_{Mises} 分布云图

图 18 冲击位置为 A 点的斜搭接接头的应力分布云图

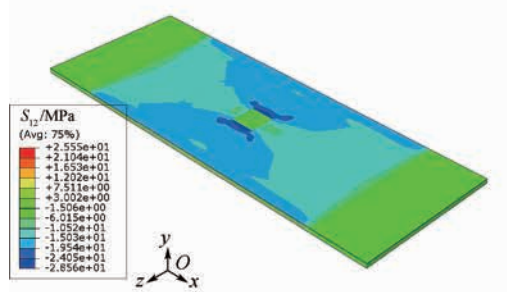
Fig. 18 Stress distribution contours of scarf joints impacted in location A



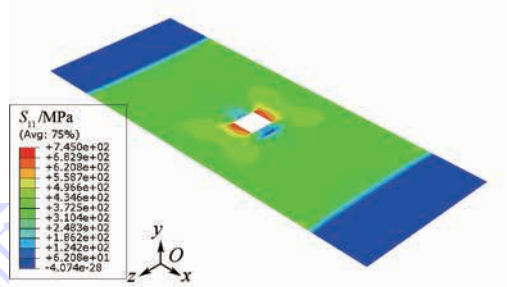
(a) S_{11} 分布云图



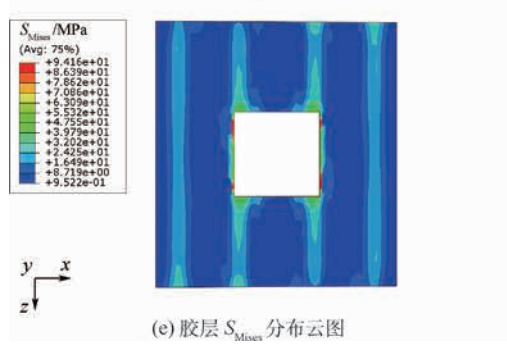
(b) S_{22} 分布云图



(c) S_{12} 分布云图



(d) 0°铺层 S_{11} 分布云图

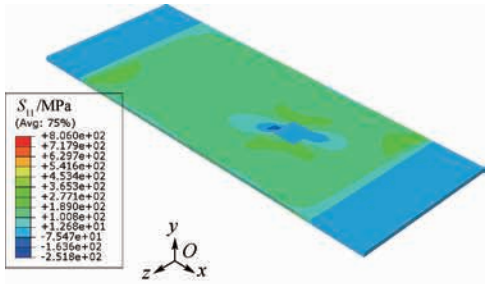


(e) 胶层 S_{Mises} 分布云图

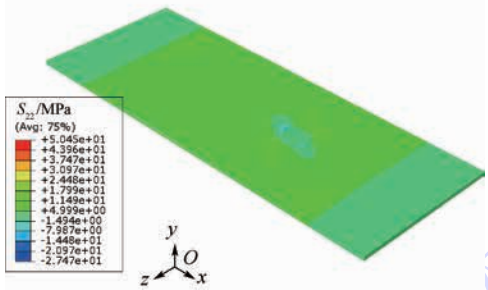
图 19 冲击位置为 B 点的斜搭接接头的应力分布云图

Fig. 19 Stress distribution contours of scarf joints impacted in location B

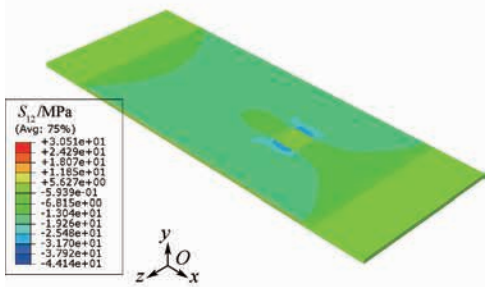
北航学报 赠阅



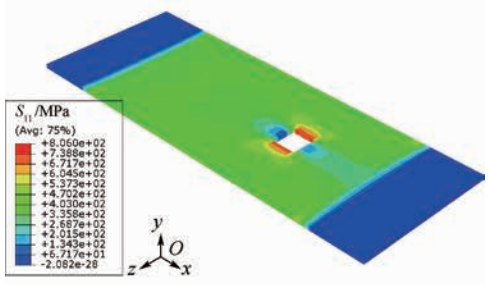
(a) S_{11} 分布云图



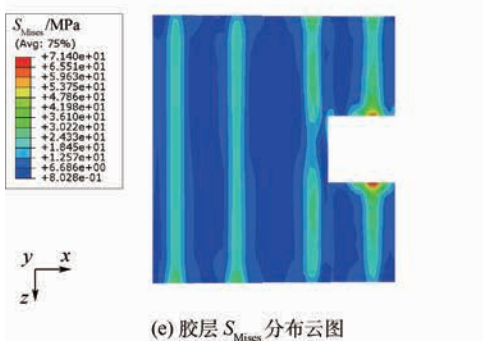
(b) S_{22} 分布云图



(c) S_{12} 分布云图



(d) 0°铺层 S_{11} 分布云图



(e) 胶层 S_{Mises} 分布云图

图 20　冲击位置为 C 点的斜搭接接头的应力分布云图

Fig. 20　Stress distribution contours of scarf joints impacted in location C

表 4　各有限元模型中应力分量最大值

Table 4　Maximum values of stress components of different finite element models

MPa

应力分量	A	B	C
S_{11}	748.81	754.05	806.03
S_{22}	49.79	39.74	50.45
S_{12}	39.47	28.56	44.12

其余二者；冲击位置为搭接区中间（B 位置），S_{22} 与 S_{12} 应力最大值较低，胶层应力集中区域较小，但胶层应力最大值明显高于其余二者。

2.3.4　冲击后拉伸渐进损伤分析

冲击发生在胶接接头的不同位置会对冲击后拉伸的损伤演化过程、进而对最终的破坏模式产生不同影响。通过复合材料和胶层在非冲击区域的损伤云图，考察胶接接头冲击后拉伸的渐进损伤过程，图 21、图 22 和图 23 中分别给出在冲击位置为 A、B、C 3 处时，模型在基体失效发生、胶层失效发生、纤维失效发生和结构整体破坏等 4 个时刻的复合材料纤维损伤、基体损伤、分层损伤以及胶层损伤（FV1、FV2、FV3 和 SDEG）云图。

冲击位置在 A 处和 C 处时，随着拉伸载荷的增加，首先出现基体损伤并逐步扩展。然后胶层的部分区域开始失效，但胶层失效扩展较慢且局限于较小的区域中，结构的承载能力没有受到太大影响。随着载荷增加，冲击区边缘开始出现纤维损伤，纤维损伤扩展较慢，而基体损伤则加快扩展至大部分区域。直到载荷继续增加至接近峰值时，纤维损伤迅速扩展，导致沿冲击区域的横截面出现大范围的纤维失效，同时胶层失效也迅速扩展，结构失去承载能力发生破坏。复合材料的分层失效与纤维失效发展规律基本一致。

冲击位置在 B 处时，模型的渐进损伤过程与 A 和 C 基本一致，不同之处在于初始纤维损伤发生的位置离复合材料冲击损伤区边缘较远。

从以上分析可知，在纤维失效大范围出现后，结构的承载能力受到严重影响，从而发生整体破坏，而胶层损伤的扩展也是导致结构发生失效的重要原因。冲击位置在 C 处时，结构整体应力水平较高，最早出现胶层失效和纤维失效，也最早发生整体破坏。而冲击位置在 B 处时，虽然纤维失效也较早出现，但其发生位置离复合材料的冲击损伤区边缘较远，减缓了冲击区边缘的纤维断裂扩展，推迟了结构的整体破坏，因此拉伸破坏载荷较高。

北航学报 赠阅

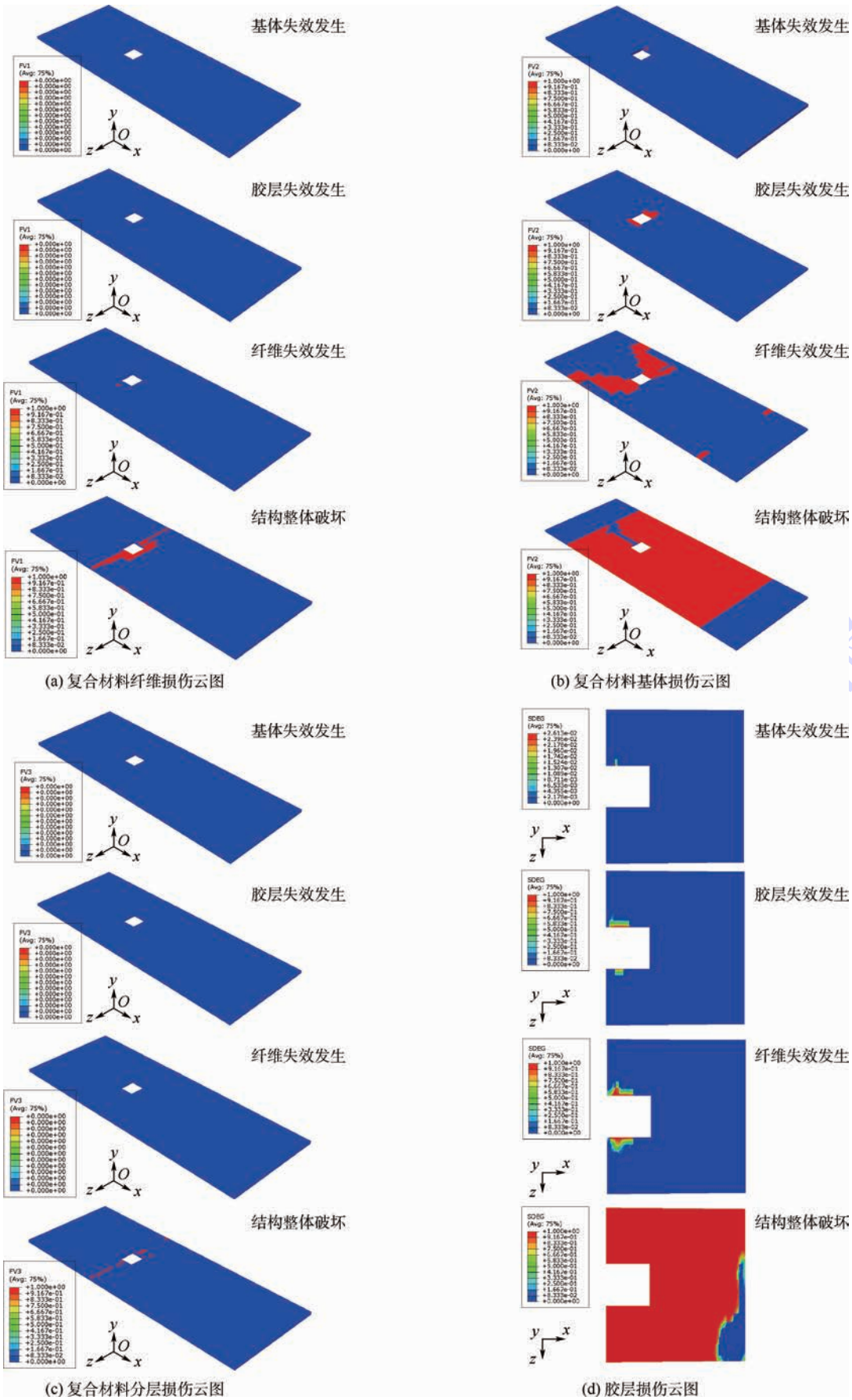


图 21 冲击位置为 *A* 处时的渐进损伤云图

Fig. 21 Damage propagation contours when impacted in location *A*

北航学报 赠阅

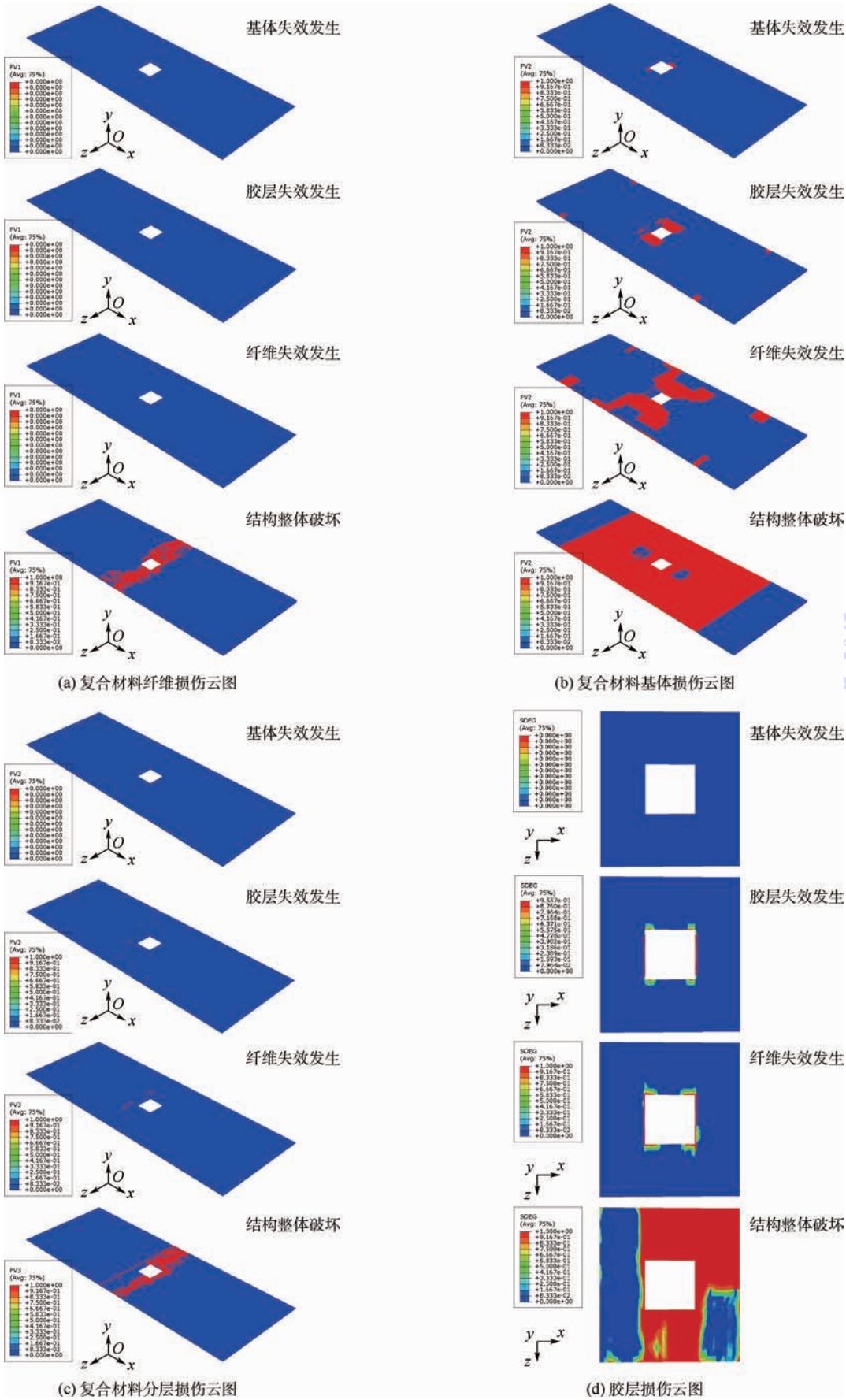
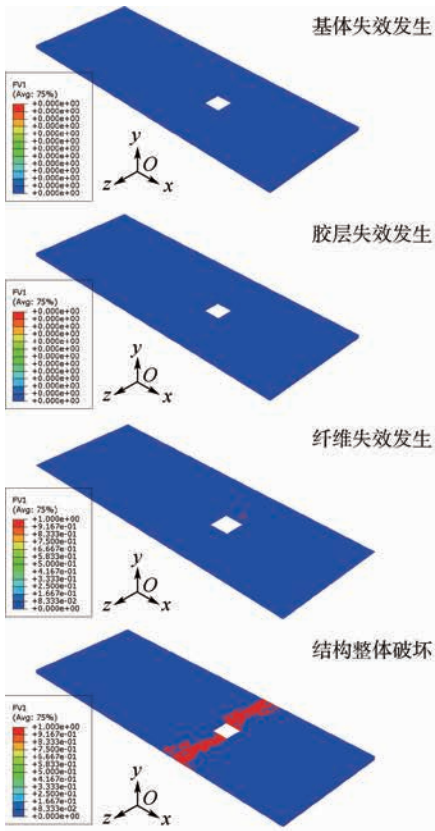


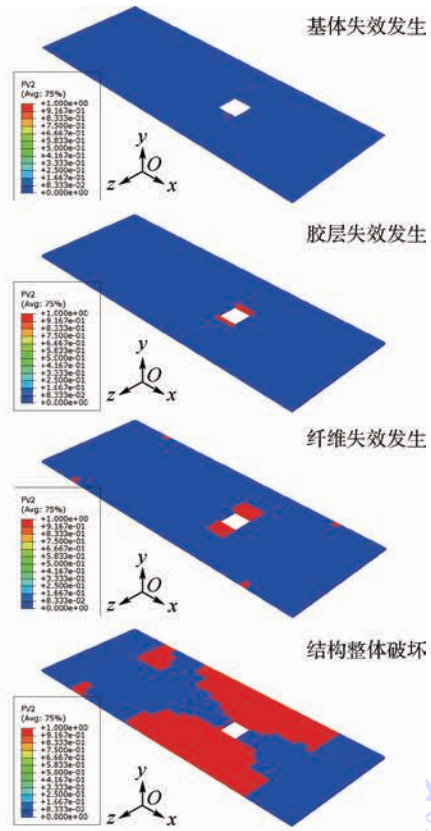
图 22　冲击位置为 B 处时的渐进损伤云图

Fig. 22　Damage propagation contours when impacted in location B

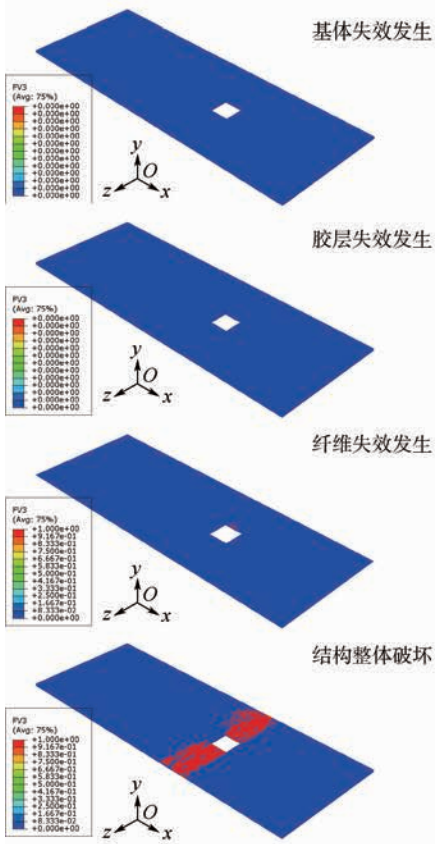
北航学报 赠阅



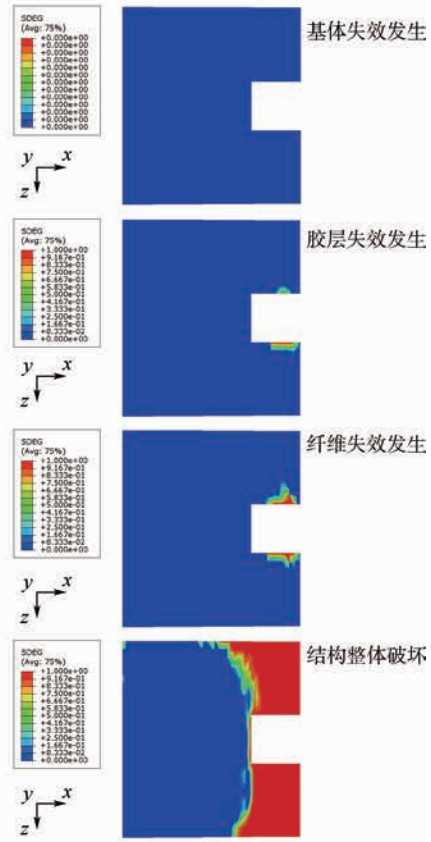
(a) 复合材料纤维损伤云图



(b) 复合材料基体损伤云图



(c) 复合材料分层损伤云图



(d) 胶层损伤云图

图 23 冲击位置为 C 处时的渐进损伤云图

Fig. 23 Damage propagation contours when impacted in location C

北航学报 赠阅

3　结　论

低速冲击损伤对斜搭接胶接接头的力学性能存在严重威胁，因为胶层的存在，不同的冲击位置对胶接接头的冲击损伤和冲击后拉伸性能会造成不同影响。对斜搭接胶接接头受到 3 个不同位置冲击后的拉伸性能进行了试验及有限元研究，得到以下结论：

1）冲击位置在搭接区尖端背侧，由于胶层缓冲作用小，造成的凹坑深度最大，冲击后拉伸强度最低；冲击位置在搭接区中部，胶层吸收了许多冲击能量，冲击造成的凹坑深度最小，冲击后拉伸强度最大；冲击位置在搭接区尖端侧，胶层也起了一定的缓冲作用，凹坑深度不大，冲击后拉伸强度居中。有限元计算表明，冲击位置在搭接区尖端背侧时层板吸收的冲击能量最大，因此造成的损伤最大。

2）由于冲击造成结构承载不对称以及冲击区域对旁路载荷传递的影响，试验件应力集中的主要区域为搭接区尖端及冲击位置旁垂直于加载方向的位置，有限元计算得到的应力集中区域与试验结果一致。

3）试验件的主要破坏模式有接头剥离和整体拉断两种，前者出现得不多，主要因胶层强度不足、连接界面存在较大的剥离应力引起，后者是大多数试验件的破坏模式，主要因冲击造成的复合材料损伤引起。

4）通过有限元方法进行渐进损伤分析可知，胶层损伤和复合材料纤维损伤的大范围扩展最终导致结构失效。冲击位置为搭接区尖端背侧时，胶层损伤和纤维损伤出现较早，结构较早发生破坏；冲击位置为搭接区中部时，冲击区附近的纤维损伤扩展得到减缓，结构破坏推迟。

参考文献（References）

[1] WU W Y, LIU Q, ZONG Z J, et al. Experimental investigation into transverse crashworthiness of CFRP adhesively bonded joints in vehicle structure[J]. Composite Structures, 2013, 106: 581-589.

[2] 谢鸣九. 复合材料连接[M]. 上海：上海交通大学出版社，2011：24-25.
XIE M J. Joints for composites materials[M]. Shanghai: Shanghai Jiao Tong University Press, 2011: 24-25 (in Chinese).

[3] SAYMAN O, ARIKAN V, DOGAN A, et al. Failure analysis of adhesively bonded composite joints under transverse impact and different temperatures[J]. Composites Part B: Engineering, 2013, 54: 409-414.

[4] GOH J Y, GEORGIADIS S, ORIFICI A C, et al. Effects of bondline flaws on the damage tolerance of composite scarf joints[J]. Composites Part A: Applied Science and Manufacturing, 2013, 55: 110-119.

[5] CAMPILHO R D S G, DE MOURA M F S F, DOMINGUES J J M S. Modelling single and double-lap repairs on composite materials[J]. Composites Science and Technology, 2005, 65(13): 1948-1958.

[6] HARMAN A B, WANG C H. Damage tolerance and impact resistance of composite scarf joints[C]// Proceedings of the 16th International Conference on Composite Materials, ICCM-16. Vancouver, BC: International Committee on Composite Materials, 2007: 1-12.

[7] 刘遂，关志东，郭霞，等. 复合材料双搭接接头拉伸强度研究[J]. 航空材料学报，2012，32(5)：86-91.
LIU S, GUAN Z D, GUO X, et al. Tensile strength of composite double-lap joint[J]. Journal of Aeronautical Materials, 2012, 32(5): 86-91 (in Chinese).

[8] 郭霞，关志东，刘遂，等. 搭接长度对复合材料单搭接胶接接头的影响[J]. 科技导报，2013，31(7)：37-41.
GUO X, GUAN Z D, LIU S, et al. Effect of lap length on the adhesive-bonded single-lap composite joints[J]. Science & Technology Review, 2013, 31(7): 37-41 (in Chinese).

[9] 乔玉，周光明，刘伟先，等. 复合材料阶梯形胶接接头渐进损伤分析[J]. 南京航空航天大学学报，2014，46(4)：632-637.
QIAO Y, ZHOU G M, LIU W X, et al. Progressive damage analysis for stepped-lap joints of adhesive-bonded composite[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2014, 46(4): 632-637 (in Chinese).

[10] 刘凯. 端部受冲击的胶接接头应力响应分析[D]. 宜昌：三峡大学，2013：62-63.
LIU K. On the stress response of adhesively bonded joint under the Izod impact test[D]. Yichang: China Three Gorges University, 2013: 62-63 (in Chinese).

[11] 吴坤岳. 基于内聚力模型的胶接结构疲劳特性研究[D]. 大连：大连理工大学，2013：60-61.
WU K Y. Research on fatigue behavior of adhesive bonded structure based on cohesive zone model[D]. Dalian: Dalian University of Technology, 2013: 60-61 (in Chinese).

[12] American Society for Testing and Materials Committiee. Standard test method for measuring the damage resistance of a fiber reinforced polymer matrix composite to a drop-weight impact event: ASTM D7136-07[S]. Philadelphia: ASTM, 2007.

[13] 沈真，张子龙，王进，等. 复合材料损伤阻抗和损伤容限的性能表征[J]. 复合材料学报，2004，21(5)：140-145.
SHEN Z, ZHANG Z L, WANG J, et al. Characterization of damage resistance and damage tolerance behaviour of composite laminates[J]. Acta Materiae Compositae Sinica, 2004, 21(5): 140-145 (in Chinese).

[14] 沈真，杨胜春，陈普会. 复合材料层压板抗冲击行为及表征方法的实验研究[J]. 复合材料学报，2008，25(5)：125-133.
SHEN Z, YANG S C, CHEN P H. Experimental study on the behavior and characterization methods of composite laminates to withstand impact[J]. Acta Materiae Compositae Sinica, 2008, 25(5): 125-133 (in Chinese).

北航学报 赠阅

[15] ABAQUS Inc. . Abaqus analysis user's manual, Volume IV: Elements[M]. Providence, RI: Dassault Systemes Simulia Corp, 2010:31.5.6-2-31.5.6-7.

[16] 杨小辉,胡坤镜,赵宁,等. 内聚力界面单元在胶接接头分层仿真中的应用[J]. 计算机仿真,2010,27(10):317-320.
YANG X H, HU K J, ZHAO N, et al. The simulation of the process of laminal damage in the adhesive bonded joint using the cohesive interface element[J]. Computer Simulation, 2010, 27(10):317-320(in Chinese).

[17] HASHIN Z. Failure criteria for unidirectional fiber composites [J]. Journal of Applied Mechanics, 1980, 47(2):329-334.

作者简介:

聂恒昌 男,博士研究生。主要研究方向:飞机复合材料结构修理。

Tel.: 010-82338873

E-mail: nhc1991@126.com

关志东 男,博士,教授,博士生导师。主要研究方向:飞机复合材料结构设计。

Tel.: 010-82338873

E-mail: zdguan@buaa.edu.cn

Tensile behavior of scarf joints after impact in different locations

NIE Hengchang[1], XU Jifeng[2], GUAN Zhidong[1,*], LI Zengshan[1], JI Zhaojie[1], TAN Riming[1]

(1. School of Aeronautic Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;
2. Beijing Aeronautical Science & Technology Research Institute, Commercial Aircraft Corporation of China, Ltd., Beijing 102211, China)

Abstract: The damage introduced by low-velocity impact is a threat to composite bonded joints. Scarf joints were impacted in three different locations of joint areas and then tensile test after impact was conducted. Finite element models of impact and tensile after impact of scarf joints were established respectively. The test results demonstrate that when impact occurs at the back face of tips, the indentation depth is the largest and the tensile strength after impact is the lowest. When impact occurs at the middle of joint area, the indentation depth is the smallest and the tensile strength after impact is the highest. When impact occurs at the front face of tips, the indentation depth and tensile strength are medium. The finite element analysis results show that the major cause of structural failure is the extensive expansion of adhesive damage and fiber cracks.

Key words: composite structure; scarf joints; adhesively bonded joints; low-velocity impact; tensile behavior

Received: 2015-10-26; **Accepted**: 2016-01-15; **Published online**: 2016-03-15 14:28

URL: www.cnki.net/kcms/detail/11.2625.V.20160315.1428.004.html

* **Corresponding author**. Tel.: 010-82338837 E-mail: zdguan@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0712

奇异谱分析在故障时间序列分析中的应用

王鑫[1], 吴际[1,*], 刘超[1], 牛文生[1,2], 张华[1], 张奎[1]
(1. 北京航空航天大学 计算机学院, 北京 100083; 2. 中航工业西安航空计算技术研究所, 西安 710068)

摘 要: 由于日益增长的飞行安全和飞机维护质量需求,飞机使用可靠性已经成为一个重要的研究领域。从某航空公司波音737飞机使用过程中现场所记录的18年的故障数据出发,应用奇异谱分析(SSA)方法,对故障时间序列进行了建模和预测,进一步以预测结果的均方根误差(RMSE)最小为优化目标对SSA模型参数进行了优选。在此基础上,提出了一种更为广泛的模型组合方法和实现算法,这种方法采用不同的时间序列模型来构造SSA分解出的趋势、周期和残差等成分。通过与三次指数平滑(Holt-Winters)、自回归移动平均(ARIMA)2种时间序列模型的实验结果对比,SSA及其参数优选和模型组合方法在故障时间序列分析中具有更好的拟合和预测精度。

关 键 词: 奇异谱分析(SSA); 故障时间序列; 预测; 参数优选; 模型组合
中图分类号: O213.2; V37; TB39
文献标识码: A **文章编号**: 1001-5965(2016)11-2321-11

数据驱动技术逐渐被证明为一种研究故障时间序列的有效方法[1],特别是对于飞机这种高可靠性需求的复杂系统。对于这种系统,人们很难从影响故障产生的众多要素入手对故障时间序列进行分析,原因是覆盖不全以及一些突发因素。因此,从历史故障数据出发,应用数学方法建立故障模型,发现其中的潜在规律并对未来做出预测,是目前故障时间序列的主要研究方法。

总的来说,故障时间序列分析覆盖众多领域和方法,如水利、软件工程等领域的人工神经网络方法[2-3],可靠性领域的奇异谱分析(Singular Spectrum Analysis, SSA)[1]、支持向量回归机[4]、人工神经网络[5-7]等方法,道路交通领域的向量自回归移动平均(Autoregressive Integrated Moving Average, ARIMA)方法[8-9]。这些方法都是以原始数据作为输入,大致可以划分为统计和人工智能2类,其中统计类又分为时域、频域2种视角。然而,在航空领域,故障时间序列预测的研究较为罕见,且研究对象也很局限,比如文献[7]中提到的人工神经网络方法对波音737飞机轮胎故障时间序列建模和预测。

SSA[10-11]是一种时域和频域相结合的非参数方法,可以用于处理非线性、非平稳性以及包含噪声的时间序列[1],能够有效提取时间序列中的主要成分,适用性广泛且操作灵活。文献[1]应用SSA对柴油机涡轮增压器和汽车发动机2个故障时间序列[4]进行了建模和预测,结果表明比其他方法更为有效。然而,该文献没有清晰地给出故障时间序列应用SSA方法的数学模型,也没有进一步分析模型参数的优选问题。文献[12]提出了一种基于SSA和ARIMA的组合模型,并用该模型预测水库中长期年径流,实验结果表明,组合模型要优于单一模型。然而,该组合方法仅通过

收稿日期: 2015-11-02; **录用日期**: 2016-02-29; **网络出版时间**: 2016-04-13 15:19
网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20160413.1519.002.html
基金项目: 中国民用航空专项研究项目(MJ-S-2013-10)
* **通讯作者**: Tel.: 010-82317624 E-mail: wuji@buaa.edu.cn

引用格式: 王鑫, 吴际, 刘超, 等. 奇异谱分析在故障时间序列分析中的应用[J]. 北京航空航天大学学报, 2016, 42(11): 2321-2331. WANG X, WU J, LIU C, et al. Application of singular spectrum analysis to failure time series analysis[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(11): 2321-2331 (in Chinese).

北航学报 赠阅

ARIMA 方法对 SSA 分解出的各个成分建立预测模型,没有针对每个成分的特点应用不同的分析方法,也没有清晰地给出建立该组合模型的具体算法。

本文首先针对故障时间序列,定义了应用 SSA 方法进行建模和预测的数学模型,在此基础上,定义了参数优选目标函数及约束条件,提出了一种基于均方根误差(Root Mean Square Error,RMSE)的参数优选算法,进一步以 SSA 为基础,定义了一种更为广泛的组合模型,给出了模型组合算法。最后,应用这些方法对某航空公司波音 737 飞机的整机故障时间序列进行实验,其中模型组合实验引入了三次指数平滑(Holt-Winters)和 ARIMA 2 种经典的时间序列分析方法,通过实验结果对比验证了 SSA 方法在故障时间序列分析中的有效性和优良性。

1 相关理论和技术

1.1 奇异谱分析

SSA 是一种广义功率谱分析[12],其主要思想是提取时间序列中的有效成分进行建模和预测。SSA 包括分解和重构 2 个过程,分解过程又分为嵌入和奇异值分解(Singular Value Decomposition,SVD),重构过程又分为分组和对角平均[11,13-14]。

1.1.1 嵌 入

定义时间序列 $X_N=(x_1,x_2,\cdots,x_N)$,序列长度为 N,设置窗口长度为 L,L 为整数且 $1<L<N/2$,$K=N-L+1$,定义延迟向量 $\boldsymbol{X}_i=(x_i,x_{i+1},\cdots,x_{i+L-1})^{\mathrm{T}}$,$X_N$ 的轨迹矩阵为

$$\boldsymbol{X}=[\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_K]=\begin{bmatrix}x_1 & \cdots & x_K\\ \vdots & & \vdots\\ x_L & \cdots & x_N\end{bmatrix} \tag{1}$$

轨迹矩阵是一个 Hankel 矩阵,其列向量和行向量均是原始序列 X_N 的子集,且对角平均后可还原为原始序列。

1.1.2 奇异值分解

对 $\boldsymbol{X}\boldsymbol{X}^{\mathrm{T}}$ 进行奇异值分解,得到 L 个降序排列的非负特征值 $\lambda_1,\lambda_2,\cdots,\lambda_L$,以及对应的正交特征向量 $\boldsymbol{U}_1,\boldsymbol{U}_2,\cdots,\boldsymbol{U}_L$,令 d 为非零特征值的个数,则 $\boldsymbol{X}$ 可表示为如式(2)所示,其中 $\sqrt{\lambda_i}$ 为 $\boldsymbol{X}$ 的奇异值,$\{\sqrt{\lambda_1},\sqrt{\lambda_2},\cdots,\sqrt{\lambda_d}\}$ 为奇异谱,$\boldsymbol{U}_i$ 通常由经验正交函数(Empirical Orthogonal Functions,EOF)表示,$\boldsymbol{V}_i$ 为主成分(Principal Components,PC),$(\sqrt{\lambda_i},\boldsymbol{U}_i,\boldsymbol{V}_i)$ 为 $\boldsymbol{X}$ 的第 i 个三重特征向量。

$$\begin{cases}\boldsymbol{X}=\boldsymbol{E}_1+\boldsymbol{E}_2+\cdots+\boldsymbol{E}_d\\ \boldsymbol{E}_i=\sqrt{\lambda_i}\boldsymbol{U}_i\boldsymbol{V}_i^{\mathrm{T}}\\ \boldsymbol{V}_i=\boldsymbol{X}^{\mathrm{T}}\boldsymbol{U}_i/\sqrt{\lambda_i}\end{cases} \tag{2}$$

1.1.3 分 组

截取式(2)中的前 r 个 $\boldsymbol{E}_i$ 近似 $\boldsymbol{X}$,并将其分为 p 组,设第 I_j 组内有 j_m 个矩阵,然后将每个组内的矩阵相加得到新的矩阵 $\boldsymbol{X}_{I_j}$,则轨迹矩阵 $\boldsymbol{X}$ 可以近似表示为

$$\begin{cases}\boldsymbol{X}\approx\boldsymbol{X}_{I_1}+\boldsymbol{X}_{I_2}+\cdots+\boldsymbol{X}_{I_p}\\ \boldsymbol{X}_{I_j}=\boldsymbol{E}_{I_{j1}}+\boldsymbol{E}_{I_{j2}}+\cdots+\boldsymbol{E}_{I_{jm}}\\ r=\sum_{j=1}^{p}\sum_{k=1}^{m}j_k\end{cases} \tag{3}$$

式中:$\boldsymbol{X}_{I_j}$的贡献率为 $\sum_{k=1}^{m}\lambda_{I_{j_k}}\Big/\sum_{i=1}^{d}\lambda_i$。

1.1.4 对角平均

对角平均是将矩阵还原为时间序列的方法。设 $\boldsymbol{X}_{I_j}$的元素为 $y_{i,j}$,$L^*=\min(L,K)$,$K^*=\max(L,K)$,则 $\boldsymbol{X}_{I_j}$ 还原为对应的时间序列 $\{g_0,g_1,\cdots,g_{N-1}\}$ 可由式(4)得出:

$$g_k=\begin{cases}\dfrac{1}{k+1}\displaystyle\sum_{m=1}^{k+1}y_{m,k-m+2} & 0<k<L^*-1\\ \dfrac{1}{L^*}\displaystyle\sum_{m=1}^{L^*}y_{m,k-m+2} & L^*-1<k<K^*\\ \dfrac{1}{N-k}\displaystyle\sum_{m=k-K^*+2}^{N-K^*+1}y_{m,k-m+2} & K^*<k<N\end{cases} \tag{4}$$

应用对角平均对式(2)中的子矩阵 $\boldsymbol{E}_i$ 还原可以得到 d 个重构序列$\{\mathrm{RC}_1,\mathrm{RC}_2,\cdots,\mathrm{RC}_d\}$。

1.1.5 预 测

SSA 的预测使用了线性递归公式(Linear Recurrent Formulae, LRF),即用前 $L-1$ 个数据点的线性组合作为当前数据点的预测值,如下:

$$\hat{y}_i=\begin{cases}\tilde{y}_i & i=1,2,\cdots,N\\ \displaystyle\sum_{j=1}^{L-1}a_j\hat{y}_{i-j} & i=N+1,N+2,\cdots,N+h\end{cases} \tag{5}$$

式中:系数 a_j 由奇异值分解中的特征向量确定[14];k 为预测点个数。

除了这种递归预测方法,文献[15]中还提到一种向量预测方法,与递归预测方法相比具有更好的稳定性,但是需要消耗更多的计算资源。

1.1.6 参数选择

SSA 的参数选择仅涉及嵌入和分组这 2 个步

骤。在嵌入过程中,窗口长度 L 的选择主要依据 2 个原则:① 不能大于时间序列长度的一半;② 可以整除时间序列的已知周期。此外,较小的 L 有利于提取趋势成分,较大的 L 有利于分析周期成分[14,16]。在分组过程中,需要确定截取的子矩阵个数 r 和分组数 p,以及子矩阵在每个分组内的分配。r 主要根据奇异值的贡献率来确定,通过比较前 r 个子矩阵的贡献率之和同预先设定的阈值 $\eta(0.85\sim0.95)$ 的大小,得到满足阈值的子矩阵个数;子矩阵的分组一般是综合考虑贡献率、谐波和噪声等因素,将子矩阵分为趋势组、周期组和噪声组,可以采用一些定量和统计的手段,如用特征向量、重构序列等的散点图来检测谐波成分等[15],用 Kendal 非参数检验的方法判断某个重构序列是否属于趋势成分[17],用加权相关系数来判断重构序列之间的可分性[18],根据周期图或者加权相关系数对重构序列自动分组[19]等。

1.2　三次指数平滑

三次指数平滑是一种处理含有趋势性和周期性成分时间序列的方法[20-22],主要思想是用历史数据的不同特征来递推当前数据点。三次指数平滑模型分为加法和乘法 2 种类型。加法预测模型可以表示为

$$\begin{cases} L_t=\alpha(y_t-S_{t-s})+(1-\alpha)(L_{t-1}+b_{t-1}) \\ b_t=\beta(L_t-L_{t-1})+(1-\beta)b_{t-1} \\ S_t=\gamma(y_t-L_t)+(1-\gamma)S_{t-s} \\ F_{t+k}=L_t+wb_t+S_{t+k-s} \end{cases} \tag{6}$$

式中:α、β 和 γ 为平滑系数,取值均为 0 ~ 1 之间,靠近 1 说明预测模型更多依赖最近观测值,其值根据均方误差最小来确定;L_t、b_t 和 S_t 分别为 t 时刻的水平、趋势的斜率和季节性成分,在实际使用中需依次计算初值;w 为预测步长;F_{t+k} 为 t 时刻的第 k 步预测值。

1.3　自回归移动平均

ARIMA 是 20 世纪 70 年代 Box 和 Jenkins[23] 提出的时间序列分析方法,其模型可以表示为 ARIMA(b,c,q),其中,b、c 和 q 分别为自回归项数、差分次数和移动平均项数,当序列平稳即 $c=0$ 时,对应的模型为

$$y_t=\sum_{i=1}^{b}\varphi_i y_{t-i}-\sum_{j=1}^{q}\theta_j\varepsilon_{t-j}+\varepsilon_t \tag{7}$$

式中:y_{t-i} 表示为 $t-i$ 时刻的观测值;ε_{t-j} 表示为 $t-j$ 时刻的随机项;φ_i 和 θ_j 为对应的系数,可以通过参数估计确定[24]。

当序列包含趋势性时,可通过 c 阶逐期差分转化为平稳序列。当序列既包含趋势性又包含季节性时,可以建立乘积季节 ARIMA 模型,记为 SARIMA(b,c,q)(B,C,Q)$^{[s]}$,其中,B、C 和 Q 分别为季节性的自回归项数、季节性差分次数和季节性移动平均项数,s 为季节长度。文献[25]给出了 SARIMA 模型的详细介绍。在实际建模中,b 和 q 可以通过观测自相关函数(Auto Correlation Function, ACF)和偏自相关函数(Partial Auto Correlation Function, PACF)确定,也可以通过计算 Akaike Information Criterion(AIC)或 Bayesian Information Criterion(BIC)值取其最小来确定。另外,文献[26]给出了 ARIMA 的自动建模方法。

1.4　故障率

设备故障率是可靠性分析的重要指标,通常故障率函数 $\lambda(t)$ 可以用 Weibull 分布来表示[27],不同参数取值对应浴盆曲线的不同阶段。在实际应用中,复杂系统常使用平均故障率来描述指定时间内的故障率[28],即

$$\lambda_a=\sum_{i=1}^{C_w}n_i\Big/\sum_{i=1}^{C_w}t_i=1/\mathrm{MTBF} \tag{8}$$

式中:n_i 为第 i 次使用周期中发生的故障数;t_i 为第 i 次使用周期的工作时长;C_w 为指定时期内的工作次数;MTBF 为平均故障间隔时间。

1.5　度量指标

为了比较不同预测模型的优劣,本文采用 RMSE 作为模型的度量指标。RMSE 与 Mean Absolute Error(MAE)、Mean Absolute Percentage Error(MAPE)等其他度量准则相比对异常点更为敏感[29],RMSE 的计算式为

$$\mathrm{RMSE}=\sqrt{\sum_{t=1}^{T}(y_t-f_t)^2/T} \tag{9}$$

式中:y_t 和 f_t 分别为时间序列在 t 时刻的观测值和模型输出值;T 为数据点的个数。

本文对训练集和测试集分别计算 RMSE 来评价模型的拟合精度和预测精度。

2　研究方法

根据故障时间序列的特点,在 SSA 方法的基础上,设计的分析框架如图 1 所示,包括数据准备和模型计算 2 个步骤。下面重点介绍 SSA 模型、参数优选和组合模型 3 个核心模块。

北航学报 赠阅

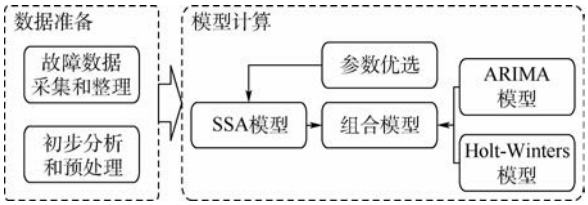


图 1 故障时间序列分析框架

Fig. 1 Analysis framework of failure time series

2.1 SSA 模型

定义故障时间序列 TFS 以及划分的训练集 TR 和测试集 TE 如下：

$$\mathrm{TFS}=\{(t_1,f_1),(t_2,f_2),\cdots,(t_n,f_n)\} \tag{10}$$

$$\mathrm{TR}=\{(t_{\mathrm{tr}_1},f_{\mathrm{tr}_1}),(t_{\mathrm{tr}_2},f_{\mathrm{tr}_2}),\cdots,(t_{\mathrm{tr}_n},f_{\mathrm{tr}_n})\} \tag{11}$$

$$\mathrm{TE}=\{(t_{\mathrm{te}_1},f_{\mathrm{te}_1}),(t_{\mathrm{te}_2},f_{\mathrm{te}_2}),\cdots,(t_{\mathrm{te}_n},f_{\mathrm{te}_n})\} \tag{12}$$

满足的约束条件为

$$\mathrm{TR}\subset\mathrm{TFS},\mathrm{TE}\subset\mathrm{TFS},\mathrm{TR}\cap\mathrm{TE}=\varnothing \tag{13}$$

$$\mathrm{tr}_n+\mathrm{te}_n\leqslant t_n,\mathrm{tr}_n<\mathrm{te}_1,\mathrm{tr}_n,\mathrm{te}_n\neq 0,\forall t\in\mathbf{N} \tag{14}$$

式中：tr_n 为训练集的大小，即样本容量；f_{tr_n} 为 tr_n 时刻对应的故障数量。

以 TR 为输入，依据第 1.1 节，构建 SSA 模型的过程可以表示为

$$\mathcal{C}_{\mathrm{ssa}}:\mathrm{TR}\xrightarrow{\text{Embedding}}\boldsymbol{F}=\begin{pmatrix}f_{\mathrm{tr}_1} & \cdots & f_{\mathrm{tr}_{\mathrm{tr}_n-L+1}}\\ \vdots & & \vdots\\ f_L & \cdots & f_{\mathrm{tr}_n}\end{pmatrix}\xrightarrow{\text{SVD}}\boldsymbol{F}=\sum_{i=1}^{d}\boldsymbol{E}_i \tag{15}$$

$$\mathcal{R}_{\mathrm{ssa}}:\xrightarrow{\text{Grouping}}\boldsymbol{F}\approx\sum_{j=1}^{m}F_j\xrightarrow{\text{Diagonal averaging}}\mathrm{TR}_{\mathrm{ssa}}=\sum_{j=1}^{m}\mathrm{TR}_{\mathrm{RC}_j}+e \tag{16}$$

$$\mathcal{F}_{\mathrm{ssa}}:\mathrm{TE}_{\mathrm{ssa}}=\mathrm{LRF}\left(\sum_{j=1}^{m}\mathrm{TR}_{\mathrm{RC}_j}\right) \tag{17}$$

式中：$\mathcal{C}_{\mathrm{ssa}}$、$\mathcal{R}_{\mathrm{ssa}}$ 和 $\mathcal{F}_{\mathrm{ssa}}$ 分别表示 SSA 分解、重构和预测。通过 $\mathcal{C}_{\mathrm{ssa}}$ 和 $\mathcal{R}_{\mathrm{ssa}}$ 2 个步骤，TR 序列被转换为一些主要成分序列（$\mathrm{TR}_{\mathrm{RC}_j}$）之和，而无效部分（$e$）被滤除，然后这些主要成分被用来建立预测模型。

2.2 参数优选

在第 2.1 节的 SSA 模型中，除了样本容量 tr_n 外，还需要确定 3 个参数：窗口长度 L、分组数 p 以及预测模型的类型 $\mathcal{F}_{\mathrm{ssa}}.\mathrm{type}$。这些参数一般通过第 1.1.6 节描述的原则人工确定。但是，由于参数取值范围较广且参数组合数目较大，需要对参数组合进行优选。优选的依据是使得最终模型的预测精度（RMSE 值）最小，目标函数可以表示为

$$\min\ \mathrm{RMSE}=\sqrt{\frac{1}{T}\sum_{\mathrm{te}=1}^{\mathrm{te}_n}(y_{\mathrm{te}}-f_{\mathrm{te}})^2}$$

$$\forall f_{\mathrm{te}}\in\mathrm{TE},\forall y_{\mathrm{te}}\in\mathcal{F}_{\mathrm{ssa}}.\mathrm{output} \tag{18}$$

约束条件为

$$2\leqslant L\leqslant \mathrm{tr}_n/2,L<50,C_{\mathrm{w}}\mid L,\forall L\in\mathbf{N} \tag{19}$$

$$1\leqslant p\leqslant 50,p<L,\forall p\in\mathbf{N} \tag{20}$$

$$\mathcal{F}_{\mathrm{ssa}}.\mathrm{type}\in\{\mathrm{recurrent},\mathrm{vector}\} \tag{21}$$

式中：y_{te} 为预测模型的输出值。根据第 1.1.6 节，L 的最大值为样本容量的一半，p 的最大值取经验值 50。当 $L=1$ 时不能进行奇异值分解，故 L 的最小值为 2。由于 SSA 递归预测方法的限制，p 的最大值不能等于 L。具体优选流程由内到外依次包括单元计算（$\mathcal{O}_{\mathrm{ssa.\,unit}}$）、分组优化（$\mathcal{O}_{\mathrm{ssa.\,groups}}$）和窗口长度优化（$\mathcal{O}_{\mathrm{ssa.\,window}}$）3 层，优化算法如算法 1 所示，其中，$\mathcal{V}_{\mathrm{RMSE}}$ 为 RMSE 值的计算函数。

算法 1 SSA 模型参数优选

输入：TFS，tr_n。

输出：不同个数预测点的最小 RMSE 值及其对应的 SSA 模型。

```
execute O_ssa.window(TFS, tr_n)
  for all L:2,3,…,min(tr_n/2,50)
execute O_ssa.groups(TFS, tr_n, L)
        for all p:1,2,…,min((L-1),50)
          execute O_ssa.unit(TFS, tr_n, L, p)
        get TR and TE
          s←C_ssa(TR, L)
          r←R_ssa(s, p)
          f←F_ssa(s, p, te_n, F_ssa.type)
            execute V_RMSE(f.output, TFS, tr_n)
            return l_G←V_RMSE.output
        end(for)
        return l_L←min(list(l_G))
  end(for)
  return l←min(list(l_L))
```

2.3 组合模型

模型组合的核心思想是将故障时间序列应用 SSA 方法分解成不同成分，包括趋势、周期和残差等，对这些成分分别建立预测模型，然后再将结果合并。模型组合流程如图 2 所示。

分解过程采用了第 2.1 节中的 $\mathcal{C}_{\mathrm{ssa}}$ 和 $\mathcal{R}_{\mathrm{ssa}}$ 2 个步骤，整合结果如式（22）所示。建模过程选用 Holt-Winters、ARIMA 和 SSA 3 种模型。由于

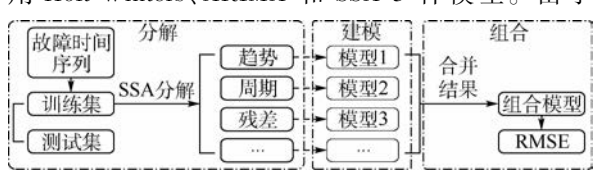


图 2 模型组合流程

Fig. 2 Process of model combination

SSA 分解的可加性,组合过程直接将每种成分的计算模型相加得到最终模型。式(23)和式(24)给出了组合模型($\mathcal{M}_{\text{com}}$)和组合预测模型($\mathcal{F}_{\text{com}}$)的表达式。

$$\text{TR} \xrightarrow{\mathcal{C}_{\text{ssa}}} \boldsymbol{F} = \sum_{i=1}^{d} \boldsymbol{E}_i \xrightarrow{\mathcal{R}_{\text{ssa}}} \text{TR}_{\text{ssa}} = \sum_{j=1}^{m} \text{TR}_{\text{RC}_j} + e \tag{22}$$

$$\mathcal{M}_{\text{com}} = \sum_{j=1}^{m,\,\text{model}\in\{\text{Holt-Winters,ARIMA,SSA}\}} \mathcal{M}_j(\text{TR}_{\text{RC}_j}, n_j) \tag{23}$$

$$\mathcal{F}_{\text{com}} = \sum_{j=1}^{m,\,\text{model}\in\{\text{Holt-Winters,ARIMA,SSA}\}} \mathcal{F}_j(\text{TR}_{\text{RC}_j}, n_j) \tag{24}$$

在分解过程中,趋势成分的提取选用故障时间序列已知的周期(定义为 A)作为窗口长度。在这种情况下,奇异值分解的第 1 个特征值一般占有较大的贡献率,并且可以重构为趋势成分。周期成分从故障时间序列去除趋势成分后的部分中进一步提取,剩余部分为残差。建立组合模型的算法如算法 2 所示,其中,$\mathcal{D}_{\text{ssa}}$为 SSA 分解函数,n_j为第 j 个成分的样本容量。

算法 2　建立组合模型

输入:TFS,n_d,n_1,n_2,…,n_m。

输出:组合模型的 RMSE 值以及模型参数。

execute $\mathcal{D}_{\text{ssa}}(\text{TFS}, n_d)$

get TR_{com} and TE by TFS, n_d

$s_1 \leftarrow \mathcal{C}_{\text{ssa}}(\text{TR}_{\text{com}}, A)$

$r_1 \leftarrow \mathcal{R}_{\text{ssa}}(s_1, \text{list}(1))$

$\text{TR}_{\text{RC}_1} \leftarrow r_1$ \$ F1

$s_2 \leftarrow \mathcal{C}_{\text{ssa}}(\text{TR}_{\text{com}} - \text{TR}_{\text{RC}_1}, L_{\max})$

get p by the features of s_2

$r_2 \leftarrow \mathcal{R}_{\text{ssa}}(s_2, p)$

for each RC in r_2

$\text{TR}_{\text{RC}_j} \leftarrow r_2$ \$ F_{j-1}

end(for)

$\text{TR}_{\text{RC}_m} \leftarrow \text{TR}_{\text{com}} - \sum_{j=1}^{m-1} \text{TR}_{\text{RC}_j}$

execute $\mathcal{M}_{\text{com}}, \mathcal{F}_{\text{com}}$, input:$\text{TR}_{\text{RC}_j}, n_1, n_2, \cdots, n_m$

for each TR_{RC_j}

get $\mathcal{M}_l$ by the features of TR_{RC_j}

execute $\mathcal{M}_j(\text{TR}_{\text{RC}_j}, n_j), \mathcal{F}_j(\text{TR}_{\text{RC}_j}, n_j)$

end(for)

$\mathcal{M}_{\text{com}} \leftarrow \sum_{j=1}^{m} \mathcal{M}_j(\text{TR}_{\text{RC}_j}), \mathcal{F}_{\text{com}} \leftarrow \sum_{j=1}^{m} \mathcal{F}_j(\text{TR}_{\text{RC}_j})$

execute $\mathcal{V}_{\text{RMSE}}(\mathcal{F}_{\text{com}}.\text{output}, \text{TFS}, \mathcal{M}_{\text{com}})$

return $\mathcal{M}_{\text{com}}, \mathcal{F}_{\text{com}}, \mathcal{V}_{\text{RMSE}}$

3　案例分析

实验原始数据来自某航空公司正在运营的 2 架波音 737 飞机的故障库。经过数据采集、整理和校验后,得到包含飞行小时数 FH、飞行起落数 FC、故障数 FA、故障率 FR 和 MTBF 在内的 5 类按月统计的数据集,截取最近 18 年 216 个数据点作为实验数据。

图 3 展示了 2 架飞机 5 类数据的散点图。可以看出,成正相关性的序列对有 FH 和 FC,FA 和 FR;成负相关性的序列对有 FR 和 MTBF;FH 和 FA 无显著相关性。根据第 1.4 节,通过 FH 和 FA 可以计算出 FR 和 MTBF,而 FH 以及 FC 可以通过飞行计划得到,预测的意义不大,故选择 FA 这个关键序列作为实验对象,并将最后一年的 12 个数据点作为测试集。

为了减少原始数据中的异常点(过大或过小值)对预测模型的不良影响,对 FA 序列进行预处理。本文采用一种简单的设置门限方法,即小于下限或者大于上限的数据点分别用下限值和上限值替换,本文实验分别以 0.1 倍和 2 倍均值为上下门限值。最终得到的 FA 序列如图 4 所示。

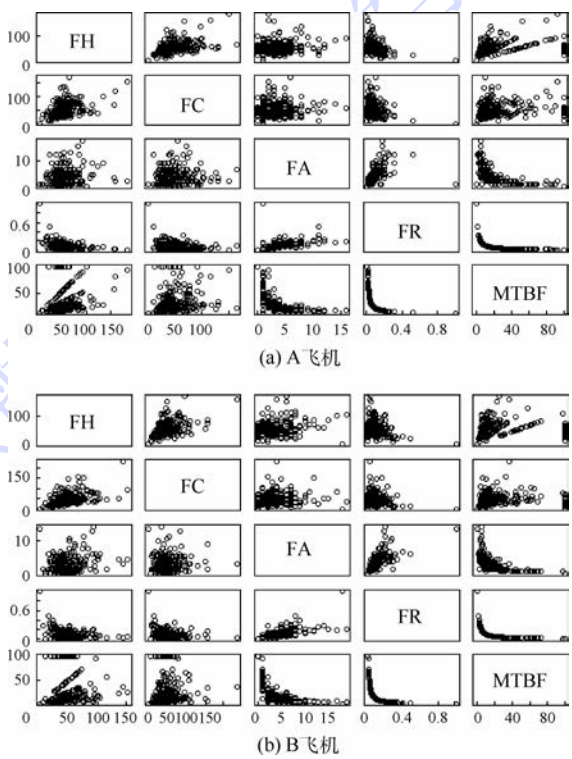


图 3　A、B 飞机故障数据集散点图

Fig. 3　Scatter diagram of failure dataset for aircraft A and B

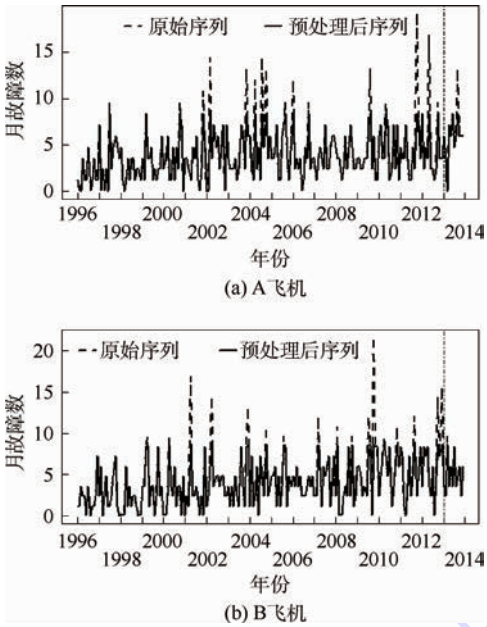


图 4 A、B 飞机预处理后的故障序列

Fig. 4 Failure series of aircraft A and B after preprocessing

3.1 基本模型实验

应用第 2.1 节的方法对 A 飞机的 FA 序列建立 SSA 模型。根据第 1.1.6 节确定参数取值范围,本例选取该范围内的最大值,即样本容量为 204,窗口长度为 96,分组数为 50。奇异值分解的结果如图 5 和图 6 所示,图 7 为对应的重构序列。

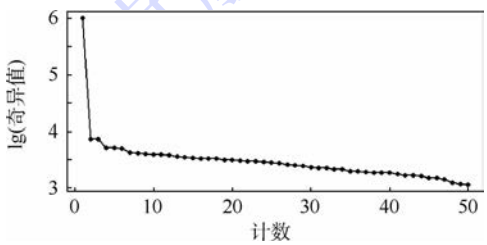


图 5 FA 序列的前 50 个奇异值

Fig. 5 First 50 singular values of FA series

可见,第 1 个奇异值贡献率较大(74.19%),蕴含了一个增长的趋势成分。另外,第 2 个和第 3 个奇异值大小相近,特征向量图和重构序列图也表现出相似的曲线,并且从图 8 中的特征对散点图中可以观察到较为明显的四边形,说明蕴含了 4 个月的周期成分,进一步计算周期估计值为4.098 372。接下来重构分解结果,选择前 50个奇异值对应的重构序列作为预测模型,模型的拟合结果和预测结果如图 9 和图 10 所示。

为了进一步比较 SSA 模型的实际效果,本文分别选取 180 和 204 作为样本容量,对 2 架飞机的 FA 序列应用 Holt-Winters、ARIMA 和 SSA 3 种方法建立预测模型,模型参数和实验结果如表 1 和表 2 所示。

可以看出,样本容量为 204 的情况下,SSA 向量预测模型12个月的预测精度最高(RMSE分

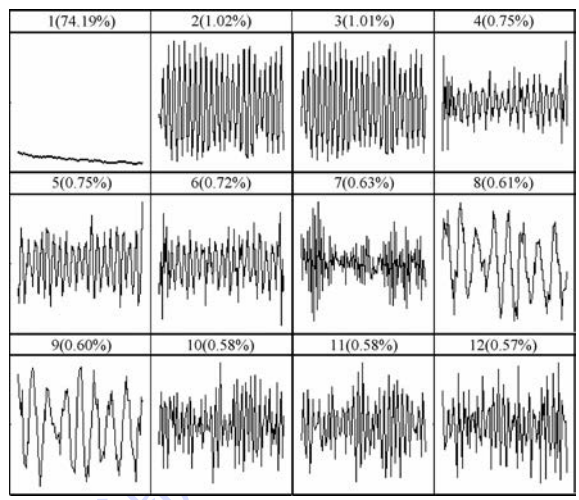


图 6 FA 序列前 12 个奇异值对应的特征向量

Fig. 6 First 12 singular values' eigenvectors of FA series

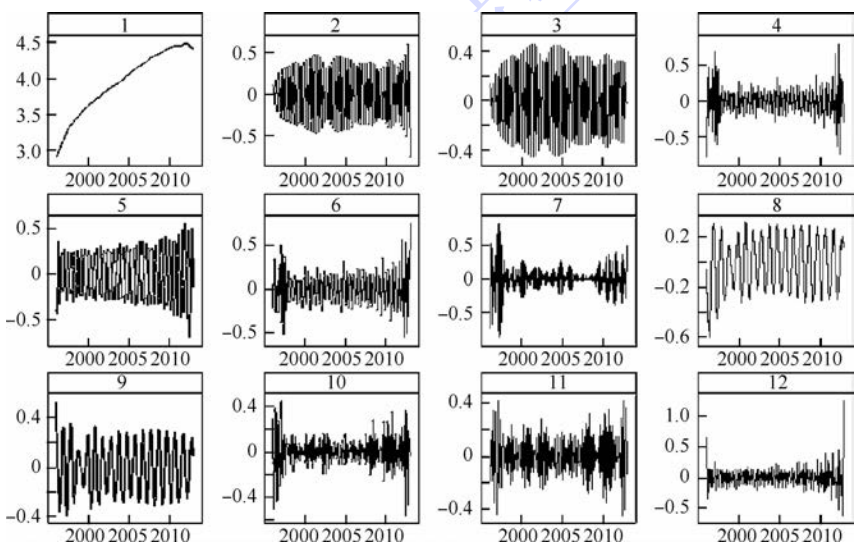


图 7 FA 序列前 12 个奇异值对应的重构序列

Fig. 7 First 12 singular values' reconstructed series of FA series

北航学报 赠阅

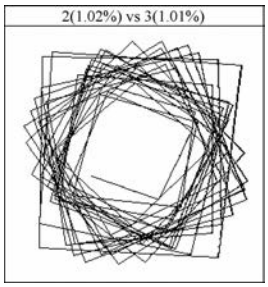


图 8　第 2 个和第 3 个奇异值对应的特征对

Fig. 8　Eigenvector pairs of 2nd and 3rd singular values

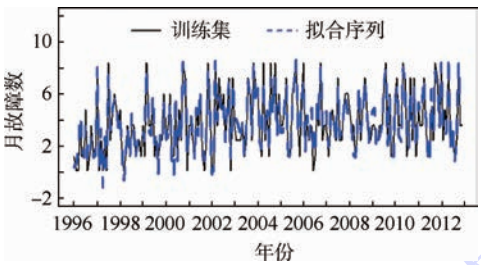


图 9　A 飞机重构序列的拟合图

Fig. 9　Fitting chart of reconstructed series for aircraft A

别为 2.294 7 和 2.361 9);SSA 模型的拟合优势明显,但是随着样本容量的增大而有所下降,这是因为随着样本容量的增大,对应的窗口长度增大,进而奇异值个数增加,分组数所占的比重减小,模型的贡献率减小,故拟合精度降低;SSA 向量预测方法比递归预测方法的结果更平稳,整体精度较高,与第 1.1.5 节的理论一致。

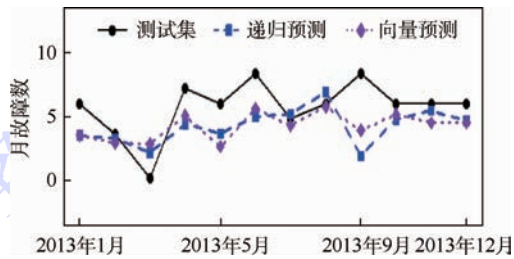


图 10　A 飞机 2 种预测方法的实验结果

Fig. 10　Experimental results of two forecast methods for aircraft A

表 1　A 飞机不同模型实验结果

Table 1　Experimental results of different models for aircraft A

Holt-Winters 样本容量	α	β	γ	拟合 RMSE	预测 RMSE				
					1	2	3	6	12
180[a]	0.007 3	0.434 8	0.223 3	2.605 9	0.118 0	0.287 4	2.591 2	2.775 2	2.415 9
180[m]	0.010 1	0.282 4	0.350 2	2.885 9	0.917 8	0.769 4	2.612 8	2.831 3	2.596 8
204[a]	0.044 0	0.072 8	0.223 3	2.617 2	0.088 4	0.068 7	2.278 3	2.882 5	2.595 1
204[m]	0.000 0	0.000 0	0.600 0	3.198 7	1.704 4	1.205 4	2.837 1	3.179 1	3.066 3
ARIMA 样本容量	b	c	q	拟合 RMSE	预测 RMSE				
					1	2	3	6	12
180	2	1	3	2.316 9	1.444 0	1.210 8	2.710 2	2.843 8	2.529 2
204	0	1	1	2.364 4	1.412 3	1.224 3	2.792 2	2.759 0	2.341 6
SSA 样本容量	L	p		拟合 RMSE	预测 RMSE				
					1	2	3	6	12
180[r]	84	list(1:50)		0.660 0	3.718 1	2.682 4	2.443 9	2.982 0	3.242 3
180[v]	84	list(1:50)		0.660 0	2.311 2	1.740 6	2.280 2	2.887 5	2.872 2
204[r]	96	list(1:50)		0.769 6	2.487 3	1.781 1	1.871 3	2.437 2	2.621 7
204[v]	96	list(1:50)		0.769 6	2.509 3	1.842 7	2.175 0	2.500 0	2.294 7

注:a 和 m 分别表示 Holt-Winters 加法和乘法季节性模型;r 和 v 分别表示 SSA 递归和向量预测模型。

表 2　B 飞机不同模型实验结果

Table 2　Experimental results of different models for aircraft B

Holt-Winters 样本容量	α	β	γ	拟合 RMSE	预测 RMSE				
					1	2	3	6	12
180[a]	0.044 7	0.091 0	0.273 7	3.045 9	1.621 2	3.540 8	3.352 4	2.585 8	2.438 6
180[m]	0.000 0	0.000 0	0.619 3	3.732 2	1.715 7	4.520 8	3.944 9	3.150 9	3.033 9
204[a]	0.010 5	0.209 6	0.191 0	2.906 7	1.748 6	3.604 5	3.352 6	2.609 1	2.474 2
204[m]	0.000 0	0.000 0	0.438 4	3.231 5	1.836 8	4.251 7	3.806 6	2.949 8	2.816 0
ARIMA 样本容量	b	c	q	拟合 RMSE	预测 RMSE				
					1	2	3	6	12
180	2	1	3	2.689 7	2.432 7	3.107 4	2.935 6	2.542 5	2.381 2
204	0	1	1	2.572 3	2.448 3	3.119 6	2.942 8	2.544 8	2.374 6
SSA 样本容量	L	p		拟合 RMSE	预测 RMSE				
					1	2	3	6	12
180[r]	84	list(1:20)		1.628 4	1.690 3	1.234 4	1.647 2	3.442 6	3.534 3
180[v]	84	list(1:20)		1.628 4	0.210 7	2.413 1	2.702 9	2.390 3	2.204 3
204[r]	96	list(1:20)		1.704 1	0.111 7	0.753 0	2.015 5	3.485 0	3.346 9
204[v]	96	list(1:20)		1.704 1	1.492 3	2.515 4	2.559 9	2.303 2	2.361 9

北航学报 赠阅

3.2 参数优选结果

以 A 飞机为例,应用第 2.2 节的方法对第 3.1 节中 SSA 模型参数进行优选。最大样本容量设为 204,根据式(19)~式(21)可知,共有 3 825 个参数组合,分 3 层计算最优参数:第 1 层,遍历窗口长度 L 的取值范围,本例为 2~102 共 101 个值,返回所有窗口长度的 RMSE 最小值及其对应的参数;第 2 层,对于每一个 L,遍历分组数 p 的取值范围,本例分 2 种情况,当 $L>50$ 时 $p\in[1,50]$,当 $L\leqslant 50$ 时 $p\in[1,L-1]$,最后返回所有分组数的 RMSE 最小值及其对应的参数;第 3 层,对于每一个分组数 p,建立 SSA 分解模型和 2 种预测模型,并计算 RMSE。最后得到的结果如表 3 所示。与表 1 对比可知,优化后的模型精度明显提高。

表 3 A 飞机 SSA 模型参数优选结果

Table 3 Parameter optimization results of SSA models for aircraft A

样本容量	优选参数组合(L,p,类型)					拟合 RMSE	预测 RMSE				
	1	2	3	6	12		1	2	3	6	12
180	50,49,r	9,3,v	43,42,v	43,42,v	31,26,v	15,3,v	0.043 7	0.003 8	0.138 5	0.632 3	1.573 8	2.052 7
204	51,50,r	44,40,v	41,39,v	49,45,v	25,23,v	91,49,v	0.094 4	0.000 9	0.045 1	0.585 4	1.624 0	2.084 0

3.3 组合模型实验

以 A 飞机为例,根据第 2.3 节的方法建立组合模型。组合模型的类型众多,除了样本容量外,还与以下几个因素相关:分解模型的参数,包括窗口长度、分组数,根据第 1.1.6 节,本例取 $L=12$,第 1 个奇异值对应的重构序列作为趋势成分,其余为残差成分;每种成分所应用的预测模型及其模型参数,本例选择 SSA(趋势)-ARIMA(残差)模型,其中参数选择取值范围内的最大值。组合过程为:首先将 FA 序列分解为趋势和残差 2 个子序列;然后分别建立趋势 SSA 模型和残差 ARIMA 模型,返回 2 个模型的拟合和预测序列;最后将拟合和预测序列相加,得到组合结果,并计算其 RMSE。

表 4 展示了趋势-残差二成分 9 种组合模型的 14 次实验结果。通过与表 1 对比可以看出,组合模型的整体拟合和预测精度要高于单一模型,特别是当趋势成分选用 SSA 模型时,其预测精度较高;趋势和残差任何一个成分采用 SSA 方法,都会提高组合模型的拟合精度,当两者均为 SSA 方法时,其拟合精度高于第 3.1 节中的基本 SSA 方法。

表 4 趋势-残差二成分组合模型实验结果

Table 4 Experimental results of trend-residual two components combination model

预测模型		参数		拟合 RMSE	预测 RMSE				
趋势	残差	趋势	残差		1	2	3	6	12
Holt-Winters[a]	Holt-Winters[a]	0.951 9, 0.553 3,1	0.003 2,1,0.228 6	2.500 1	0.284 6	0.202 0	2.403 0	2.815 1	2.553 5
	ARIMA		ARIMA(1,0,2)(1,0,0)[12]	2.025 0	0.878 2	1.450 7	3.053 6	2.799 5	2.343 3
	SSA[r]		84,50	0.544 4	3.126 2	2.556 6	2.619 5	2.431 9	2.876 5
	SSA[v]		84,50	0.544 4	2.334 5	1.819 7	2.391 0	2.359 1	2.625 7
ARIMA	Holt-Winters[a]	ARIMA(2,1,1)(2,0,2)[12]	0.003 2,1,0.228 6	2.488 9	0.318 7	0.242 0	2.230 4	3.000 0	2.874 6
	ARIMA		ARIMA(1,0,2)(1,0,0)[12]	1.957 7	0.912 2	1.367 3	2.865 2	2.937 6	2.628 2
	SSA[r]		84,50	0.534 3	3.160 2	2.631 6	2.568 3	2.590 3	3.131 0
	SSA[v]		84,50	0.534 3	2.368 5	1.887 9	2.293 0	2.541 9	2.905 8
SSA[r]	Holt-Winters[a]	84,50	0.003 2,1,0.228 6	2.488 4	0.258 0	0.208 7	2.531 5	2.717 6	2.434 5
	ARIMA		ARIMA(1,0,2)(1,0,0)[12]	1.958 9	0.851 6	1.519 6	3.193 7	2.740 0	2.274 3
	SSA[r]		84,50	0.534 4	3.099 6	2.498 9	2.662 7	2.352 3	2.830 9
SSA[v]	Holt-Winters[a]		0.003 2,1,0.228 6	2.488 4	0.267 8	0.205 4	2.519 3	2.700 5	2.404 1
	ARIMA		ARIMA(1,0,2)(1,0,0)[12]	1.958 9	0.861 4	1.502 6	3.177 3	2.722 1	2.244 2
	SSA[v]		84,50	0.534 4	2.317 7	1.783 4	2.465 6	2.263 2	2.566 9

3.4 讨 论

3.4.1 组合方法的优势

某种方法可能不适用于故障时间序列的建模和预测,但是很可能对序列分解出的某一成分有很好的建模和预测能力,这也是组合方法的优势。充分发挥每种时间序列分析方法的特性,比如 SSA 方法在残差预测中效果较好,ARIMA 方法在趋势预测中效果更优。在进行飞机故障序列的建模时,可以构造多种预测模型,针对重点关注的内容,如短期预测能力、趋势提取能力和趋势预测能

力等,选择有优势的模型进行分析,则会取得较好的效果。

3.4.2　预测结果中的分歧点

预测结果中存在一些与真实数据相悖的分歧点,这些数据点明显脱离实际水平,严重影响预测精度。如从图 10 中可以看出,2013 年 9 月份故障数的预测值与真实值形成了明显的反差:预测值为局部最小极值,真实值为局部最大极值。调整模型类型和参数后仍然得到相同的效果,这让笔者有理由认为该月的故障数据存在严重分歧。判明飞机故障序列中的分歧点有很大的应用价值,可以为故障诊断、整机可靠性分析等提供重要参考。分歧点出现的原因众多,飞机在一段时间内发生的故障数受任务强度、飞行环境和人为等多方面因素的影响,这些因素中都存在突发情况,如政策性的停飞、持续性的恶劣天气或者突发性的上级检查等。通过实地调研发现,2013 年 9 月份该架飞机接受了上级的不定期检查,导致维护人员对飞机的检查力度加大,进而发现的故障增多。

3.4.3　样本容量对预测效果的影响

为了得到更优的组合模型,可以进一步调整模型参数。如预测趋势成分时选择的建模方法及样本容量,从图 11 可以看出,以 A 飞机为例,不同方法和样本容量所得到的趋势预测结果差异很大,图中的黑线是全样本点进行 SSA 分解提取趋势后得到的曲线,有明显的增长趋势,这与单独对测试集进行简单线性回归的结果一致,而样本容量为 204 的 Holt-Winters 和 ARIMA 方法的预测结果呈现出下降的趋势,与实际不符。从图 11(Holt-Winters 为季节性加法模型;SSA 的窗口长度为 84,分组数为 50,递归预测模型)中可以观察到,2007 年开始趋势成分有明显的转变点,且波动也出现比较稳定的规律,因此选取从 2007 年开始的 72 个数据点作为趋势模型的样本更为合理,

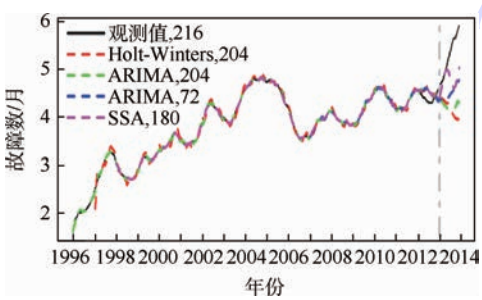


图 11　不同模型的趋势成分预测结果

Fig. 11　Trend component prediction results of different models

实验结果也证明了这一点。

4　结　论

本文在应用 SSA 技术建立飞机故障时间序列分析模型的基础上,提出了参数优选和模型组合算法,经实验验证表明:

1) SSA 方法的建模和预测性能与 Holt-Winters 和 ARIMA 相比更优,特别是拟合精度,在 A 飞机样本容量为 204 的情况下 RMSE 达到了0.769 6。

2) 参数优选算法的效果显著,在 A 飞机样本容量为 204 的情况下,模型的拟合和预测 RMSE 达到了 0.094 4 和 2.084 0。

3) 趋势-残差二成分组合模型的总体效果优于单一 SSA 方法。

4) 组合模型中趋势成分的预测结果对所选定的建模方法,特别是样本容量的大小较为敏感。

基于目前工作,后续还可以做进一步研究,如组合模型的参数优化过程、时间序列模型的扩展机制和自适应的数据预处理等内容。此外,对于故障时间序列分析,目前的研究方法是从历史数据出发逆向分析和预测,发现其内在规律。下一步可以从数据产生的领域知识入手,提取关键特征和要素,正向构建故障序列的分析模型。

致谢　感谢任健老师在实验数据方面提供的大力支持,感谢陈勇博士在语言方面给予的帮助,感谢评阅论文的各位专家。

参考文献(References)

[1] ROCCO S C M. Singular spectrum analysis and forecasting of failure time series[J]. Reliability Engineering & System Safety, 2013, 114(6): 126-136.

[2] KUTYŁOWSKA M. Neural network approach for failure rate prediction[J]. Engineering Failure Analysis, 2015, 47(1): 41-48.

[3] CAI K Y, CAI L, WANG W D, et al. On the neural network approach in software reliability modeling[J]. Journal of Systems and Software, 2001, 58(1): 47-62.

[4] MOURA M D, ZIO E, LINS I D, et al. Failure and reliability prediction by support vector machines regression of time series data[J]. Reliability Engineering & System Safety, 2011, 96(11): 1527-1534.

[5] FINK O, ZIO E, WEIDMANN U. Predicting component reliability and level of degradation with complex-valued neural networks [J]. Reliability Engineering & System Safety, 2014, 121(1): 198-206.

[6] XU K, XIE M, TANG L C, et al. Application of neural networks

in forecasting engine systems reliability[J]. Applied Soft Computing,2003,2(4):255-268.

[7] AL-GARNI A Z,JAMAL A. Artificial neural network application of modeling failure rate for Boeing 737 tires[J]. Quality and Reliability Engineering International,2011,27(2):209-219.

[8] GARCÍA F P, PEDREGAL D J, ROBERTS C. Time series methods applied to failure prediction and detection[J]. Reliability Engineering & System Safety,2010,95(6):698-703.

[9] BARBA L, RODRÍGUEZ N, MONTT C. Smoothing strategies combined with ARIMA and neural networks to improve the forecasting of traffic accidents[J]. Scientific World Journal,2014(1):152375.

[10] VATUTARD R,YIOU P,GHIL M. Singular-spectrum analysis: A toolkit for short,noisy chaotic signals[J]. Physica D:Nonlinear Phenomena,1992,58(1-4):95-126.

[11] ELSNER J B,TSONIS A A. Singular spectrum analysis-A new tool in time series analysis[M]. New York: Plenum Press, 1996:39-50.

[12] ZHANG Q,WANG B D,HE B,et al. Singular spectrum analysis and ARIMA hybrid model for annual runoff forecasting[J]. Water Resources Management,2011,25(11):2683-2703.

[13] 王解先,连丽珍,沈云中. 奇异谱分析在 GPS 站坐标监测序列分析中的应用[J]. 同济大学学报(自然科学版),2013,41(2):282-288.
WANG J X, LIAN L Z, SHEN Y Z. Application of singular spectrum analysis to GPS station coordinate monitoring series [J]. Journal of Tongji University(Natural Science),2013,41(2):282-288(in Chinese).

[14] ELSNER J B. Analysis of time series structure: SSA and related techniques[J]. Journal of the American Statistical Association, 2002,103(4):1207-1208.

[15] GOLYANDINA N,KOROBEYNIKOV A. Basic singular spectrum analysis and forecasting with R[J]. Computational Statistics & Data Analysis,2014,71(s1):934-954.

[16] GOLYANDINA N. On the choice of parameters in singular spectrum analysis and related subspace-based methods[J]. Statistics and Its Interface,2010,3(3):259-279.

[17] HIPEL K W,MCLEOD A I. Time series modelling of water resources and environmental systems[M]. Amsterdam: Elsevier, 1994:893.

[18] HASSANI H. Singular spectrum analysis:Methodology and comparison[J]. Journal of Data Science,2007,5:239-257.

[19] ALEXANDROV T,GOLYANDINA N. Automatic extraction and forecast of time series cyclic components within the framework of SSA[C]// Proceedings of the 5th Workshop on Simulation, 2005:45-50.

[20] WINTERSP R. Forecasting sales by exponentially weighted moving averages[M]// FUNKE U H. Mathematical models in marking. New York:Springer,1976:384-386.

[21] HOLT C C. Forecasting seasonals and trends by exponentially weighted moving averages[J]. International Journal of Forecasting,2004,20(1):5-10.

[22] CHATFIELD C. The Holt-Winters forecasting procedure[J]. Journal of the Royal Statistical Society,1978,27(3):264-279.

[23] BOX G E P,JENKINS G M,REINSEL G C. Time series analysis:Forecasting and control (revised edition)[J]. Journal of Marketing Research,1976,22(2):199-201.

[24] KUMAR U,JAIN V K. ARIMA forecasting of ambient air pollutants (O3, NO, NO2 and CO)[J]. Stochastic Environmental Research and Risk Assessment,2010,24(5):751-760.

[25] JEONG K,KOO C,HONG T. An estimation model for determining the annual energy cost budget in educational facilities using SARIMA (seasonal autoregressive integrated moving average) and ANN (artificial neural network)[J]. Energy,2014, 71(14):71-79.

[26] HYNDMAN R J,KHANDAKAR Y. Automatic time series forecasting: The forecast package for R[J]. Journal of Statistical Software,2008,27(3):1-22.

[27] EBELING C E. An introduction to reliability and maintainability engineering[M]. New York:McGraw-Hill,1997.

[28] SU C,JIN Q,FU Y. Correlation analysis for wind speed and failure rate of wind turbines using time series approach[J]. Journal of Renewable and Sustainable Energy,2012,4(3): 1-13.

[29] LI G,SHI J. On comparing three artificial neural networks for wind speed forecasting[J]. Applied Energy,2010,87(7): 2313-2320.

作者简介:

王鑫 男,博士研究生。主要研究方向:数据驱动技术。
Tel.: 15810128162
E-mail: wangxin_buaa@ edu. com. cn

吴际 男,博士,副教授,硕士生导师。主要研究方向:模型驱动、软件可靠性分析。
Tel.: 010-82317624
E-mail: wuji@ buaa. edu. cn

刘超 男,博士,教授,博士生导师。主要研究方向:软件工程、软件测试。
Tel.: 010-82317641
E-mail: liuchao@ buaa. edu. cn

北航学报 赠阅

Application of singular spectrum analysis to failure time series analysis

WANG Xin[1], WU Ji[1,*], LIU Chao[1], NIU Wensheng[1,2], ZHANG Hua[1], ZHANG Kui[1]

(1. School of Computer Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;
2. Xi'an Aeronautical Computing Technique Research Institute, Aviation Industry Corporation of China, Xi'an 710068, China)

Abstract: Due to significant industrial demands toward flight safety andairplane maintenance quality, improving airplane's reliability in usage stage has become an important activity and the research domain is rapidly evolving. In this paper, eighteen years' field data gathered from the maintenance phase of two Boeing 737 aircrafts are prepared as time-to-failure series. Then singular spectrum analysis (SSA) is usedto cope with this data for modeling and forecasting. Furthermore, a SSA parameter optimization algorithm is proposed by minimizing root mean square error (RMSE) of the prediction results. Based on this, a broader method of model combination is raised by utilizing different time series models to the components obtained from SSA decomposition, which represent trend, period, residuals, etc. The combination model and detailed algorithm are designed. The experimental results are compared with those of cubic exponential smoothing (Holt-Winters) and autoregressive integrated moving average (ARIMA), which verifies that the proposed models and algorithms have better fitting and prediction accuracyin failure time series analysis.

Key words: singular spectrum analysis (SSA); failure time series; prediction; parameter optimization; model combination

北京航空航天大学学报

Received: 2015-11-02; **Accepted**: 2016-02-29; **Published online**: 2016-04-13 15:19
URL: www.cnki.net/kcms/detail/11.2625.V.20160413.1519.002.html
Foundation item: China Civil Aviation Special Research Project (MJ-S-2013-10)
* **Corresponding author**. Tel.: 010-82317624　E-mail: wuji@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0720

基于稳定分簇的移动自组织网络路由协议

黄金科，樊晓光*，万明，禚真福，杨永建，陈少华

（空军工程大学 航空航天工程学院，西安 710038）

摘 要：移动自组织网络的动态拓扑特性给其分簇路由协议的设计带来了一定的挑战，而分簇算法作为分簇路由协议中重要的技术手段，如何建立一种有效稳定的分簇机制，对分簇路由协议的性能至关重要。本文在大规模网络环境下，提出一种稳定的分簇算法，并在此基础上提出一种分簇路由协议。为提高网络性能，该分簇路由协议簇间采用被动模式，簇内采用主动模式，并且各个分簇由簇首、网关、访客和普通簇节点组成。仿真结果表明：该分簇算法减少了簇的数量和簇成员在簇间的切换次数，提高了分簇的稳定性。相应的分簇路由协议较该类协议也降低了网络的平均端到端时延和平均路由开销，提高了数据包接收的成功率。

关 键 词：移动自组织网络；稳定性；分簇算法；分簇路由协议；性能分析
中图分类号：TN393.2
文献标识码：A **文章编号**：1001-5965(2016)11-2332-08

移动自组织网络(Mobile Ad Hoc Networks, MANETs)是一种高动态、多跳的无线网络，具有无中心、自组织和动态拓扑等特点[1]。由于其频繁变化的网络拓扑结构和带宽资源的有限性，给MANETs的路由协议设计带来了重大的挑战，即如何设计一种有效的路由产生、维护和选择机制，对确保网络的连通性、时效性以及网络对无线资源的利用率等方面至关重要[2]。

目前，根据更新机制的不同[3-4]，路由协议大致可以分为表驱动(主动)路由、按需(被动)路由、混合路由和地理路由4种路由策略。主动路由周期性地广播路由信息来维护路由，路由信息的周期性交换需要消耗大量的带宽资源，因此此类协议适合较小规模的网络；被动路由不需周期性地广播路由信息，只有在源、目的节点通信需要且该链路不为源节点所知时才进行路由发现过程，路由发现会造成一定的时延，因此此类协议也适合规模较小、节点移动性较大的网络；混和路由协议则是由主动路由和被动路由有机结合组成的协议，即在一定的网络区域内采用主动策略，区域间则采用被动策略，此类协议主要针对规模较大、移动性较强的网络所设计；地理路由是随着GPS、北斗等定位技术的不断发展带来的产物，在MANETs节点上装备定位装置，节点可以准确获得自己的地理位置和时标，据此，源节点可将数据“有目的”地向目的节点的方向上传输，最终到达目的节点。

根据网络的逻辑视图，路由协议可以分为平面路由协议和分簇路由协议2种。在平面路由协议中[5-6]，网络中所有的节点地位平等，功能也都相同，节点之间通常采用单一的路由策略，因此比较容易实现。但是MANETs中节点的随机移动性、加入或离开网络的不可预测性等特点引起的网络拓扑结构的频繁变化，可能会导致通信节点间路由的频繁失效，通信节点间重新建立起有效的通信路径可能会导致较长的时延和较大的网络

收稿日期：2015-11-04；**录用日期**：2016-01-28；**网络出版时间**：2016-04-01 17:07
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160401.1707.008.html
基金项目：陕西省自然科学基础研究计划资助项目（2009JM8001-4）
*通讯作者：E-mail：996970360@qq.com

引用格式：黄金科，樊晓光，万明，等. 基于稳定分簇的移动自组织网络路由协议[J]. 北京航空航天大学学报，2016，42(11)：2332-2339. HUANG J K，FAN X G，WAN M，et al. Stable cluster-based routing protocol for mobile Ad Hoc networks[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2332-2339 (in Chinese).

开销,限制了网络的规模。

分簇路由协议[7-10]的提出有效解决了平面路由中时延长和网络开销大等问题,非常适用于大规模 MANETs。在分簇路由协议中,网络按照某种特定的规则将网络节点划分成不同的簇,每个簇按特定规则选举产生簇首,相邻簇间通信时通过网关节点进行转发,而对转发后数据如何到达目的节点的细节并不关心。分簇路由协议不仅可以减少参与路由维护的节点数目,还可以降低因路由信息交换带来的网络开销,而且采用特定的分簇策略将网络划分成不同的稳定子网络,可以降低网络拓扑的动态变化对路由协议性能的影响,极大地提高了网络的通信效率。

ZRP[11]是第一个混合使用主动驱动和被动策略的自组织网络分簇路由协议,其规定在区域内采用 DVA 主动路由策略,区域间则采用类似 DSR 协议中的被动路由机制。ZHLS 和 SLURP 等协议[12]也采用类似 ZRP 协议的策略,区域内采用主动路由机制,区域间采用被动路由机制。它们有一个共同缺点,就是数据流量较大(特别是参与通信的节点较多)时,路由发现将消耗大量的网络带宽,严重时将会导致网络瘫痪。同时,频繁的路由发现过程也会造成一定的传输时延,影响数据的时效性。CBRP[13]等协议不论是在簇内还是簇间均采用被动路由策略,因此路由发现的时间较长,且当网络流量较小时,消耗的网络带宽也小,但随着流量的逐渐增大也会占用大量的网络带宽,影响网络的整体性能。HSR 和 CGSR[14-15]等协议则在簇内和簇间都使用主动路由机制,虽然降低了路由发现所需的时延,但大大增加了网络的控制开销,因此这类协议适用于节点移动速度较小的网络。

在分簇路由协议中,分簇机制作为网络拓扑管理的一种有效模式,不仅可以降低网络管理的复杂度,而且还很好地提高了 MANETs 的扩展性。因此,合理有效的分簇机制对于分簇类路由协议来说至关重要。文献[16-19]详细介绍了现有的一些分簇算法,但主要是针对一些特定的需求,例如最小 ID 簇和最大连通度算法等。文献[20]提出了一种动态多普勒分簇(Dynamic Doppler Velocity Clustering,DDVC)算法和动态链路持续时间分簇(Dynamic Link Duration Clustering,DLDC)算法,二者都只适合于网络节点具有伪线性运动轨迹的高速 MANETs。文献[21]提出了一种利用连续 2 次发送 Hello 消息的多普勒频移计算节点移动速度的 MPBC 算法,该算法将与所有邻居节点的移动特性具有最大相似度的节点选举为簇首。

基于此,本文提出了一种基于联合度量指标 T_m 的分簇算法。仿真结果表明:T_m 分簇算法通过提高簇内节点的数目,减少了簇成员在簇间的切换次数,提高了分簇算法的稳定性。同时还在该分簇算法的基础上提出了一种新的分簇路由协议 T_m-R。仿真结果表明:T_m-R 分簇路由协议较其他 4 种典型协议降低了网络的平均端到端时延和平均路由开销,并增加了数据包接收的成功率。

1　分簇算法

1.1　分簇度量指标

1.1.1　链路保持时间

假定 MANETs 中的每个网络节点都按照三维 Gauss-Markov 模型移动[22],节点的最大通信半径均为 r,任意 2 个节点之间可实现双向通信,而且每个网络节点都能通过 GPS 或北斗卫星实时获得自身的地理位置和时间信息,并周期性地向其相邻节点发送自己的最新坐标和时间信息,据此可大致预测出每条链路的保持时间。

设节点 i 前后 2 次向其邻居节点 j 发送自己的时间和坐标信息,分别为 t_1,(x_{i1},y_{i1}) 和 t_2,(x_{i2},y_{i2}),且在 t_1 和 t_2 时刻,节点 j 的坐标信息分别为 (x_{j1},y_{j1}) 和 (x_{j2},y_{j2})。在实际网络中,节点 i 和节点 j 都是移动的,为简化分析,在此假定节点 i 固定,节点 j 移动。图 1(a)为节点 i 和节点 j 相互"远离"的情形,图 1(b)为节点 i 和节点 j 相互"接近"的情形。

在图 1 中,节点 i 固定在坐标原点 O,而节点 j 相对节点 i 由 A 点沿向量$\overrightarrow{AB}$的方向移动。假定在 $\Delta t=t_2-t_1$ 时间内,节点 i 和节点 j 运动速度和方向都保持不变,经过时段 Δt,节点 j 由 A 点运动到 B 点。

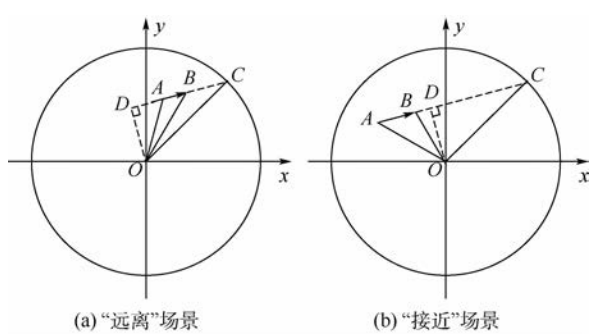


图 1　链路保持时间示意图

Fig. 1　Sketch map of link holding time

北航学报 赠阅

图 1(a)中有如下关系成立:

$$d_{OA}=\sqrt{(x_{n1}-x_{m1})^2+(y_{n1}-y_{m1})^2} \tag{1}$$

$$d_{OB}=\sqrt{(x_{n2}-x_{m2})^2+(y_{n2}-y_{m2})^2} \tag{2}$$

由余弦定理可得

$$d_{AB}=\sqrt{d_{OA}^2+d_{OB}^2-2d_{OA}d_{OB}\cos\angle AOB} \tag{3}$$

式中:$\angle AOB=\arcsin\left(\frac{y_{n1}}{d_{OA}}\right)-\arcsin\left(\frac{y_{n2}}{d_{OB}}\right)$。

则节点 i 和节点 n 之间的相对移动速度为

$$v=d_{AB}/\Delta t \tag{4}$$

由于 Δt 时间内,节点 i 和节点 j 运动速度和方向都假定保持不变,因此,图 1 中节点 j 脱离节点 i(即飞出节点 i 的通信范围)所需的运动轨迹为线段 BC。即从时刻 t_2 开始算起,节点 i 和节点 j 之间的链路还可以继续保持的时间为

$$T=d_{BC}/v=(d_{CD}-d_{BD})/v \tag{5}$$

式中:

$$d_{CD}=\sqrt{r^2-(d_{OB}\sin\angle OBA)^2}$$

$$d_{BD}=d_{OB}\cos\angle OBA$$

$$\angle OBA=\arccos[(d_{AB}^2+d_{OB}^2-d_{OA}^2)/(2d_{OB}d_{AB})]$$

图 1(b)的分析方法与图 1(a)类似,同理可得

$$T=d_{BC}/v=(d_{BD}+d_{CD})/v=(d_{BD}+d_{CD})/(d_{AB}/\Delta t) \tag{6}$$

式中:d_{OA} 和 d_{OB} 定义与图 1(a)相同。

$$d_{AB}=\sqrt{d_{OA}^2+d_{OB}^2-2d_{OA}d_{OB}\cos\angle AOB}$$

$$\angle AOB=\arcsin\left(\frac{y_{n2}}{d_{OB}}\right)-\arcsin\left(\frac{y_{n1}}{d_{OA}}\right)$$

$$d_{BD}=d_{AD}-d_{AB}=d_{OA}\cos\angle OAB-d_{AB}$$

$$d_{CD}=\sqrt{r^2-(d_{OA}\sin\angle OAB)^2}$$

$$\angle OAB=\arccos[(d_{AB}^2+d_{OA}^2-d_{OB}^2)/(2d_{OA}d_{AB})]$$

1.1.2 移动性因子

为度量节点的移动性因子,引入远隔性函数[23-24]的概念。远隔性是节点移动性的一种标准度量。

假定网络中共有 N 个节点,$\boldsymbol{n}_i(t)$ 表示节点 i 在 t 时刻的位置向量,$i=1,2,\cdots,N$,则在 t 时刻节点 i 和节点 j 之间的距离可以定义为

$$d_{ij}(t)=|\boldsymbol{n}_i(t)-\boldsymbol{n}_j(t)| \tag{7}$$

因此,t 时刻节点 i 到节点 j 的远隔性可定义为

$$R_{ij}(t)=F(d_{ij}(t)) \tag{8}$$

式中:$F(\cdot)$ 为关于距离的函数。通常,最简单的 $F(\cdot)$ 函数是恒等函数,即

$$R_{ij}(t)=d_{ij}(t) \tag{9}$$

远隔性是随着节点的动态移动而实时改变的,因此移动性因子应该根据远隔性的时间倒数来计算,即移动性因子 $M_i(t)$ 为

$$M_i(t)=\frac{1}{N}\sum_{j=1}^{N}\left|\frac{\mathrm{d}}{\mathrm{d}t}F(d_{ij}(t))\right| \tag{10}$$

将式(9)代入式(10)得

$$M_i(t)=\frac{1}{N}\sum_{j=1}^{N}|d'_{ij}(t)|=\frac{1}{N}\sum_{j=1}^{N}\left|\frac{d_{ij}(t)}{\mathrm{d}t}\right|\approx\frac{1}{N}\sum_{j=1}^{N}\left|\frac{d_{ij}(t_2)-d_{ij}(t_1)}{t_2-t_1}\right| \tag{11}$$

式中:$t_2-t_1\to 0$。

继续沿用图 1 的假设,因此无论是节点 i 和节点 j 相互“远离”还是“靠近”,式(11)都可简化为

$$M_i(t)=\frac{1}{N}\sum_{j=1}^{N}\left|\frac{d_{OB}(t_2)-d_{OA}(t_1)}{t_2-t_1}\right|=\frac{1}{N}\sum_{j=1}^{N}\left|\frac{d_{OB}(t_2)-d_{OA}(t_1)}{\Delta t}\right| \tag{12}$$

1.1.3 联合度量指标

在设计 MANETs 分簇算法时,本文将链路的保持时间和节点的移动特性有机结合起来作为网络分簇时选择簇首的联合度量指标,即

$$T_m=T_{ij}/M_i \tag{13}$$

式中:T_{ij} 为选择节点 i 作为簇首时,节点 i 与邻居节点 j 之间链路的保持时间;M_i 为节点 i 的移动性因子。

1.2 分簇策略

1.2.1 簇的初始化

当 MANETs 刚建立时,所有的网络节点都处于孤立状态。每个节点首先按照式(13)计算其成为簇首节点的联合度量指标,然后向其一跳邻居节点发送包含该联合度量指标的 Hello 消息,当邻居节点收到 Hello 消息后,将其值与自己的联合度量指标相比较,将联合度量指标较大的节点选举为簇首,并将其簇首声明信息广播给它的所有一跳邻居节点。当邻居节点收到簇首声明信息之后,则向其发送加入请求信息,当簇首收到加入请求消息并“同意”该邻居节点成为其簇成员时,将向该邻居节点发送一个请求接受消息,当邻居节点收到该请求接受消息之后即意味着其成功加入到该簇中,该邻居节点就会周期性地广播一个成员信息来宣告自己的存在状态,同时簇首也会周期性地发送一个簇首信息来维护其“统治地位”。

1.2.2 网关的选择

如果网络节点在发送加入请求消息之后,收到 2 个或 2 个以上的请求接受消息,则表明该网络节点处于 2 个或多个簇的相交地带,因此可将

其选举为网关节点。

1.2.3　簇的维护

簇完成初始化之后，簇首和簇成员之间必须周期性地交换包含各自 ID 号的信息来维护“友好”关系。当 MANETs 的拓扑结构发生变化时，本文分以下 3 种情形来确保网络划分成合适的簇。

1）删除或增加节点

当簇成员收不到簇首周期性广播的包含簇首 ID 的信息或簇首收不到簇成员周期性广播的包含簇成员 ID 的信息，表明该节点已不在该簇中，则将其信息从该簇中删除。

当新节点（包括新加入网络的节点或被其他簇剔除的节点）想要加入一个簇中时，就必须周期性地广播一个加入请求信息，当某簇簇首收到该消息后，若同意其加入，则向该节点发送一个加入接受消息，新节点收到该加入接受消息后正式成为该簇中的成员。

2）簇首的更换

如果某簇的簇首离开该簇或失效时，与其相关联的邻居节点就会重新变成孤立节点，因此这些孤立节点必须重新周期性地广播加入请求消息，如果某些孤立节点收到过多的加入请求消息，则表明这些孤立节点附近不存在簇首，因此必须重启簇的初始化过程。

3）簇的合并

如果 2 个簇进入到彼此的通信范围，并且 2 个簇首之间的联合度量指标相差不大，则表明 2 个簇会在相互的范围内停留较长的时间，这意味着二者合并的可能较大，因此值得合并这 2 个分簇。原来联合度量指标较小的簇首则成为合成新簇的簇首，原来联合度量指标较大的簇首则变为普通簇成员。如 2 个簇首之间的联合度量指标相差很大，表明这 2 个簇只是“擦肩而过”，不需要合并。

2　分簇路由协议

通常，分簇路由协议可以降低大规模网络的路由开销，即通常簇内采用源路由策略，簇间采用被动路由策略，其伪代码如下：

```
if source and destination in the same cluster
{ if d_sd is small than twice of transmission range
    Node source and destination communicate directly
  else
    Node source and destination communicate via relay
}
else
{ L1: Node source send RREQ to neighbor clusterhead via gateway
  if destination is in the neighbor cluster
    Node destination RREP to source along the coming way
  else
    goto L1;
}
```

2.1　簇内通信

当源、目的节点处于同一分簇时，两节点之间可以直接通信或者通过该分簇的簇首中继转发。即当目的节点处于源节点的通信覆盖范围时，两节点可直接或通过簇首中继转发来完成通信任务；否则，只有通过簇首的转发才能完成数据传输。

2.2　簇间通信

当源、目的节点不在同一簇中时，源节点就会发起簇间路由策略。首先，源节点向其所在簇的簇首发送请求消息，该簇首收到请求消息后，通过网关节点将其转发给相邻簇的簇首。当目的节点所在分簇的簇首收到该消息后，就会沿着发送请求消息相反的路径发送应答消息。处于该反向路径上的簇首收到该应答消息后，会根据当前链路状态选择局部最短路径向下继续传递该应答消息。因此，源节点最终会获得通往目的节点的最短路径。

2.3　路由维护

由于 MANETs 中节点的高速移动性，网络的拓扑结构也在快速的变化当中，已建立链路极可能中断。因此，高效的路由维护机制对保障协议的时效性等方面至关重要。当网络中某条链路中断后，网络就会进行局部修复，即用中断节点间的一条最短路径替换中断路径，然后进行路径更新并将修复后的结果告知给目的节点。

由于簇首周期性进行路由维护，因此为确保簇间路由的时效性，只有当簇间链路发生中断时才发起局部链路修复工作。这种局部链路修复技术对减少路径重建开销和端到端时延非常有效。

3　性能分析

本文利用 NS3 网络仿真软件对动态分簇算法和分簇路由协议的性能进行仿真分析。网络中各移动节点随机分布在 300 km × 300 km 空域范围内，且按照 Gauss-Markov 移动模型移动，节点

北航学报 赠阅

的传播范围为 15 km,速度为 50 ~ 100 m/s 之间的随机值,仿真时间设为 60 min。

仿真中,每个节点的仿真模型均相同。物理层采用自由空间模型,并考虑实际航空通信的物理参数,如发射功率和接收机灵敏度等;MAC 层采用 TDMA 协议;网络层采用分簇路由协议(簇内采用 OLSR,簇间采用 AOVD);应用层采用均匀分布的恒定比特业务。网络成员以相同概率产生新的业务且持续时间服从负指数分布,数据包长度服从泊松分布。参数设置如表 1 所示。

表 1 仿真参数设置

Table 1 Simulation parameter setting

参数	数值
信道速率/(Mb · s^{-1})	5
数据包平均大小/KB	512
发射功率/W	125
接收机灵敏度/dBm	-24.67
物理层特征	直接序列扩频(DSSS)

3.1 分簇算法仿真对比与分析

在分簇路由协议中,网络控制开销主要来源于分簇的过程和簇间的信息交换,因此分簇时产生的簇越少,相应地网络开销也会减少。并且,由于 MANETs 拓扑结构不断发生变化,原来属于本簇的簇成员很可能离开本簇而加入到另外的一个簇中去,若这种情形频繁发生,则会严重影响网络的性能。

综上,本文将用簇的数量和簇间切换次数这 2 个指标来衡量分簇算法的优劣,并将本文提出的分簇算法与现有 DLDC、DDVC、MPBC 和MOBIC 算法进行比较分析。

图 2 反映了网络节点数对簇数量的影响。可以看出,当网络节点的数量增加时,T_m、DLDC、DDVC、MPBC 和 MOBIC 这 5 种算法所形成的簇的数量都在减少,但 T_m 分簇算法充分考虑了节点间链路的保持时间和移动性因子,因而形成的簇的数量相对较少且稳定。

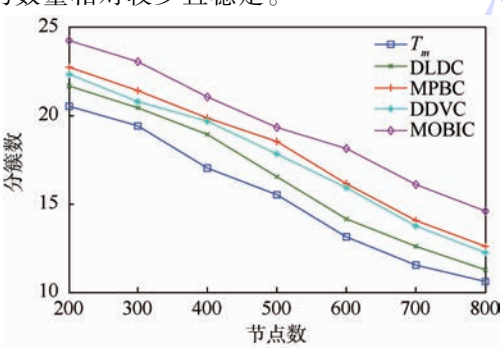


图 2 节点数量对簇数量的影响

Fig. 2 Influence of nodes' number on cluster's number

图 3 反映了网络节点数对簇间切换次数的影响。当网络节点的数量增加时,节点之间的平均距离在逐渐缩短,而节点的通信范围固定不变,因此节点通信覆盖范围内的节点数必将增加,这就会带来频繁的拓扑结构变化和通信链路的断开与连接。如图 3 所示,由于当节点数量增加时,T_m 分簇算法形成的簇较少,因此在相同的网络规模下,簇内节点的数量就会相对较多,簇内成员的簇间切换的次数也会相继减少。因此相比其他4 种分簇算法,T_m 分簇算法具有更好的稳定性。

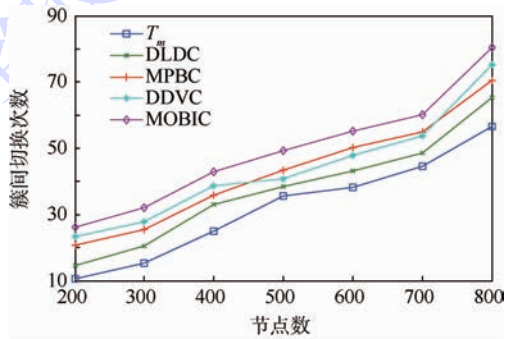


图 3 节点数量对簇间切换次数的影响

Fig. 3 Influence of nodes' number on nodes' switch times between clusters

3.2 分簇路由协议仿真对比与分析

为评估 T_m-R 分簇路由协议的性能,本文将其与现有的 MPBC-R、DLDC-R、DDVC-R 和 MOBIC-R 协议从平均端到端时延、平均路由开销和丢包率 3 个方面进行仿真对比。

显然,当节点的数量增加时,平均端到端时延必然会增加。这是因为当节点数量增加后,网络中参与通信的节点的数量也随之增加,这样就会消耗更多的带宽,也势必会增加多跳节点对之间通信的可能,因此平均端到端时延会迅速增大。图 4 反映了节点数量对网络平均端到端时延的影响。可以看出,T_m-R 分簇路由协议的平均端到端时延低于其他 4 种分簇路由协议。这是因为 T_m-R 分簇路由协议在分簇时更为精确地考虑了节点间链路的保持时间,因此形成的簇最为稳定,进而避免了不必要的时间浪费。

图 5 反映了节点数量对平均路由开销的影响。网络的路由开销主要包括路由重建和路由维护所产生的控制消息数据包。显然,当网络中的节点数量增多时,网络的平均路由开销必然会增加,这与图 5 所示相符。但由于本文提出的 T_m-R 分簇路由协议具有稳定的分簇机制,簇的重建和簇间切换的概率相对其他 4 种协议较低,不需要

过多的路由建立和维护控制数据包的传输，因此路由开销较小。

图 6 反映了节点的最大移动速度对数据包接收成功率的影响。可以看出，当节点的最大移动速度增加时，数据包的接收成功率随之降低。这是因为节点移动速度的增加势必会造成网络拓扑结构的频繁变化，这就会增加节点间原有链路“失联”的概率，因此造成了更多的丢包。由于 T_m-R 分簇路由协议的分簇机制较其他协议最为稳定，因此其数据接收成功率最高。

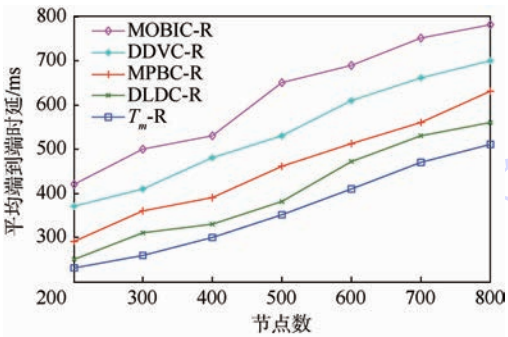


图 4　节点数量对平均端到端时延的影响

Fig. 4　Influence of nodes’ number on average end-to-end delay

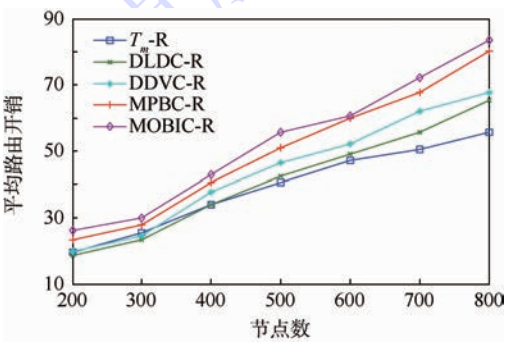


图 5　节点数量对平均路由开销的影响

Fig. 5　Influence of nodes’ number on normalized routing overhead

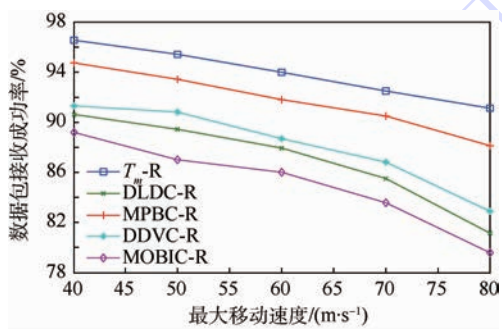


图 6　最大移动速度对数据包接收成功率的影响

Fig. 6　Influence of maximum movement speed on packet acceptance ratio

4　结　论

在分析现有分簇算法和分簇路由协议特点的基础上，提出了一种新的分簇算法和分簇路由协议，仿真结果表明：

1）该分簇算法具有良好的稳定性。

2）基于此分簇算法的分簇路由协议较其他 4 种典型分簇路由协议，降低了整个网络的平均端到端时延和平均路由开销，并提高了数据包接收的成功率。

参考文献（References）

[1] JAWHAR I, WU J. QoS support in TDMA-based mobile Ad Hoc networks[J]. Journal of Computer Science and Technology, 2005, 20(6): 797-810.

[2] KUMAR S, RAGHAVAN V S, DENG J. Medium access control protocols for Ad Hoc wireless networks[J]. Ad Hoc Networks, 2006, 4(3): 326-358.

[3] GAVRILOVSKA L, ATANASOVSKI V. Ad Hoc networking towards 4G: Challenges and QoS solutions[C]// 7th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services. Piscataway, NJ: IEEE Press, 2005: 71-80.

[4] BEN MAHMOUD M S, LARRIEU N. An ADS-B based secure geographical routing protocol for aeronautical Ad Hoc networks [C]// 2013 IEEE 37th Annual Computer Software and Applications Conference Workshops. Piscataway, NJ: IEEE Press, 2013: 556-562.

[5] SANTHI K, PARVATHAVARTHINI B. Randomized routing techniques for Ad-Hoc on-demand distance vector of wireless networks[C]// 2013 International Conference on Human Computer Interactions (ICHCI). Piscataway, NJ: IEEE Press, 2013: 1-6.

[6] ISTIKMAL. Analysis and evaluation optimization dynamic source routing (DSR) protocol in mobile Ad Hoc network based on ant algorithm[C]// 2013 International Conference of Information and Communication Technology (ICoICT). Piscataway, NJ: IEEE Press, 2013: 400-404.

[7] NIU X, TAO Z, WU G, et al. Hybrid cluster routing: An efficient routing protocol for mobile Ad Hoc networks[C]// 2006 IEEE International Conference on Communications. Piscataway, NJ: IEEE Press, 2006: 3554-3559.

[8] WOO S C, SINGH S. Scalable routing protocol for Ad Hoc networks[J]. Wireless Networks, 2001, 7(5): 513-529.

[9] ZHOU J H, LEI L, LIU W K, et al. A simulation analysis of nodes mobility and traffic load aware routing strategy in aeronautical Ad Hoc networks[C]// Proceedings of 2012 9th International Bhurban Conference on Applied Sciences and Technology. Piscataway, NJ: IEEE Press, 2012: 423-426.

[10] HYEON S, KIM K, YANG S. A new geographic routing protocol for aircraft Ad Hoc networks[C]// 2010 IEEE/AIAA 29th Digital Avionics Systems Conference. Piscataway, NJ: IEEE Press,

北航学报 赠阅

2010:255-260.

[11] DUBE R, RAIS C D, WANG K Y, et al. Signal stability based adaptive routing for Ad Hoc mobile networks[J]. IEEE Personal Communications Magazine, 1997, 4(1):36-45.

[12] NG J M, LU I. A peer-to-peer zone-based two-level link state routing for mobile Ad Hoc networks[J]. IEEE Journal on Selected Areas in Communications, 1999, 17(8):1415-1425.

[13] WANG Y, DONG L, LIANG T T, et al. Cluster-based location aided routing algorithm for mobile Ad Hoc networks[J]. Journal of Software, 2009, 11(20):3086-3100.

[14] PEI G, GERLA M, HONG X, et al. A wireless hierarchical routing protocol with group mobility [C]// 1999 IEEE Wireless Communications and Networking Conference. Piscataway, NJ: IEEE Press, 1999:1538-1542.

[15] CHIANG C C, WU H K, LIU W, et al. Routing in clustered multihop, mobile wireless networks with fading channel[C]// 1997 IEEE Singapore International Conference on Networks-Networks: The Next Millennium. Piscataway, NJ: IEEE Press, 1997:197-211.

[16] YU J Y, CHONG P H J. A survey of clustering schemes for mobile Ad Hoc networks[J]. IEEE Communications Surveys and Tutorials, 2005, 7(1):32-48.

[17] ABOLHASAN M, WYSOCKI T, DUTKIEWICZ E. A review of routing protocols for mobile Ad Hoc networks[J]. Ad Hoc Networks, 2004, 2(1):1-22.

[18] MEHTA S, SHARMA P, KOTECHA K. A survey on various cluster head election algorithms for MANET[C]// 2011 Nirma University International Conference on Engineering. Piscataway, NJ: IEEE Press, 2011:1-6.

[19] VENKATESWARAN A, SARANGAN V, GAUTAM N, et al. Impact of mobility prediction on the temporal stability of MANET clustering algorithms[C]// International Workshop on Modeling Analysis and Simulation of Wireless and Mobile System. New York: ACM, 2005:144-151.

[20] SAKHAEE E, JAMALIPOUR A. Stable clustering and communications in pseudolinear highly mobile Ad Hoc networks[J]. IEEE Transactions on Vehicular Technology, 2008, 57(6): 3769-3777.

[21] NI M M, ZHONG Z D, ZHAO D M. MPBC: A mobility prediction-based clustering scheme for Ad Hoc networks[J]. IEEE Transactions on Vehicular Technology, 2011, 60(9):4549-4559.

[22] YU J Y, CHONG P H J. An efficient clustering scheme for large and dense mobile Ad Hoc networks (MANETs)[J]. Computer Communications, 2006, 30(1):5-6.

[23] NG J M, ZHANG Y. A mobility model with group partitioning for wireless Ad Hoc networks[C]// 3rd International Conference on Information Technology and Applications. Piscataway, NJ: IEEE Press, 2005:289-294.

[24] XU K, HONG X, GERLA M. An Ad Hoc network with mobile backbones[C]// IEEE International Conference on Communications. Piscataway, NJ: IEEE Press, 2002:3138-3143.

作者简介:

黄金科 男,博士研究生。主要研究方向:无线通信、综合航电技术。

Tel.: 15829717886

E-mail: 86297609@qq.com

樊晓光 男,博士,教授,博士生导师。主要研究方向:无线通信、综合航电技术。

E-mail: 996970360@qq.com

万明 男,博士,讲师。主要研究方向:综合航电技术。

E-mail: afwanming@163.com

禚真福 男,硕士,讲师。主要研究方向:综合航电技术。

E-mail: zzf_tiger@126.com

杨永建 男,博士研究生。主要研究方向:综合航电技术。

E-mail: 645970360@qq.com

陈少华 男,博士研究生。主要研究方向:综合航电技术。

E-mail: chensh10@126.com

北京航空航天大学学报

北航学报 赠阅

Stable cluster-based routing protocol for mobile Ad Hoc networks

HUANG Jinke, FAN Xiaoguang*, WAN Ming, ZHUO Zhenfu, YANG Yongjian, CHEN Shaohua

(Aeronautics and Astronautics Engineering College, Air Force Engineering University, Xi'an 710038, China)

Abstract: The dynamic topology of a mobile Ad Hoc network poses a real challenge in the design of clustering routing protocol. And as an essential technique of clustering routing protocol, clustering of nodes provides an efficient method for establishing a clustering structure in mobile Ad Hoc networks. In this paper, we designed a stable clustering algorithm and a corresponding clustering routing protocol for large-scale mobile Ad Hoc networks. Each cluster is composed of a clusterhead, several cluster gateway nodes, several cluster guest nodes and other cluster members. The proposed routing protocol uses proactive protocol between nodes within individual clusters and reactive protocol between clusters. Simulation results show that the proposed clustering algorithm improves the stability by decreasing the clusters' number and nodes' switch times between clusters. And the clustering routing protocol provides low average end-to-end delay, normalized routing overhead and high packet acceptance ratio compared with some existing routing protocols.

Key words: mobile Ad Hoc network; stability; clustering algorithm; clustering routing protocol; performance analysis

北京航空航天大学学报

Received: 2015-11-04; **Accepted**: 2016-01-28; **Published online**: 2016-04-01 17:07

URL: www.cnki.net/kcms/detail/11.2625.V.20160401.1707.008.html

Foundation item: the Project Supported by Natural Science Basic Research Plan in Shannxi Province of China (2009JM8001-4)

* **Corresponding author**. E-mail: 996970360@qq.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0731

面向集群环境的虚拟化 GPU 计算平台

杨经纬，马凯，龙翔*
（北京航空航天大学 计算机学院，北京 100083）

摘 要：针对集群系统的多节点多 GPU 环境，提出一种新型虚拟化 GPU 计算平台。该平台实现对集群系统所有节点上 GPU 资源的统一抽象与管理，构建公共 GPU 资源池。原有 GPU 应用程序可以不经任何修改而迁移到虚拟化 GPU 计算平台，并具备访问资源池内任何 GPU 的能力，编程人员无需显式针对多节点多 GPU 应用展开 MPI 编程。应用程序摆脱了单个节点上 GPU 资源的限制，并具备无差别地访问集群系统中任何可用 GPU 资源的能力，能有效提高系统总体资源利用率以及吞吐量。采用流水化通信技术，实现对虚拟化 GPU 计算平台的运行时开销以及节点间数据传输延迟的隐藏。实验表明：与非流水化通信相比，系统总体数据传输延迟降低了 50%~70%，具备与节点机本地数据传输等同的通信性能。

关 键 词：GPU；MPI；CUDA；集群系统；硬件加速；并行计算；高性能计算
中图分类号：TP315
文献标识码：A **文章编号**：1001-5965(2016)11-2340-09

近年来，GPU 的应用场景从传统的图形图像处理领域扩展到通用计算领域，如航空航天装备研制、卫星遥感数据处理、天气预报、石油勘探和流体动力学等。鉴于 GPU 的高效能、高性价比特性，越来越多的编程人员开始关注 GPU 编程，并将其用于相关领域的高性能计算。

GPU 通用计算一般在集群系统中的多个节点上配备数量不等的 GPU 卡，现有的 GPU 编程技术（如 CUDA、OpenCL）只支持单节点应用，GPU 应用程序只能直接访问（访存、计算）节点本地的 GPU 资源。为了实现多节点多 GPU 应用，开发人员需要展开 MPI 编程或者借助于第三方运行时，使得单个应用程序能够访问更多的 GPU 资源，提高计算性能。MPI 一定程度上增加了程序的复杂度与编程人员的负担，给系统的开发、调试以及维护带来困难。

为了充分利用集群系统的 GPU 计算能力，并降低应用程序的开发难度，本文提出一种基于 CUDA 架构的 GPU 虚拟化技术，实现对集群系统所有节点上 GPU 资源的统一抽象与管理。集群系统中的每一个 GPU 都可以看作节点本地设备，应用程序按照访问本地设备的方式访问集群系统的任何一个 GPU。

基于虚拟化 GPU 计算平台，应用程序可以同时请求节点本地和其他节点上的 GPU 计算服务，每个节点也可以同时为本地和其他节点上的应用程序提供 GPU 计算服务。系统根据应用需求在全局范围静态或者动态优化 GPU 分配，从而提升计算资源的整体利用率，减少计算任务的响应时间。

1 GPU 虚拟化技术

虚拟化技术被广泛应用于现代计算机系统，如 VMware、KVM[1] 和 Xen[2]，能有效提高系统整体效能，并降低运行与维护成本。鉴于 GPU 在高性能计算领域的重要性，虚拟化技术被应用于 GPU。GPU 虚拟化技术可分为 2 类：

收稿日期：2015-11-09；**录用日期**：2016-01-15；**网络出版时间**：2016-03-15 14:40
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160315.1440.008.html
* **通讯作者**：Tel.：010-82339685 E-mail：long@buaa.edu.cn

引用格式：杨经纬，马凯，龙翔. 面向集群环境的虚拟化 GPU 计算平台[J]. 北京航空航天大学学报，2016，42(11)：2340-2348.
YANG J W，MA K，LONG X. Virtualized GPU computing platform in clustered system environment[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2340-2348 (in Chinese).

1) 硬件虚拟化:基于虚拟化 PCIE 总线(如 NextIO N2800-ICA[3]),允许某个节点访问其他节点上的 GPU,实现资源的复用。此类方法灵活性较差,且底层接口难以标准化,因此实施成本较高,一般鲜为工业界和学术界所用。

2) 软件虚拟化:基于函数库或者运行时系统,屏蔽底层硬件细节,采用标准化的高层接口实现对 GPU 资源的分时复用与透明共享(如 OpenGL[4]、Direct 3D[5]、CUDA[6]和 OpenCL[7])。

软件虚拟化技术又可分为 2 种应用场景:

1) 面向虚拟机监控器(Virtual Machine Monitor,VMM),为虚拟机(VM)提供图形加速及通用计算支持。VMware Virtual GPU[8]提供了 Windows 虚拟机环境下对 Direct 3D 的支持,而 VMGL[9]则提供了类 Unix 虚拟机环境下对 OpenGL的支持,但二者并不提供对 GPU 通用计算功能的支持。vCUDA[10]基于前后端模式实现虚拟机环境下的 GPU 通用计算,前端 CUDA 接口以封装库的形式运行于客户机,而后端 GPU 计算服务则放置于宿主机。GViM[11]、gVirtuS[12]采用类似的架构方案,并提供 CUDA 支持。

2) 面向集群环境,为客户端提供访问其他节点 GPU 的 API 接口,实现分布式 GPU 资源的虚拟化。dOpenCL[13]通过对底层分布式细节的完全隐藏,实现对 CPU、GPU 以及其他加速部件的本地化抽象,并支持 OpenCL 编程接口。VCL[14]同样实现了 OpenCL 规范,并扩展 OpenMP,允许集群系统某个节点上的 GPU 应用程序调用其他节点上 GPU 资源。除此之外,VCL 还提供基于任务依赖关系的调度以及缓冲区管理功能。rCUDA[15-16]基于分布式系统环境,实现了虚拟化 CUDA 编程接口,支持对远程 GPU 的并发使用,使得没有 GPU 设备的节点也具备 GPU 计算能力。

rCUDA 基于设备底层接口实现节点间通信,系统软硬件耦合紧密,可移植性较差。此外,rCUDA为闭源软件,其具体实现细节并未公开。鉴于此,本文提出一种面向集群系统环境的新型虚拟化 GPU 计算平台。

2　面向集群环境的 GPU 虚拟化

2.1　传统 GPU 计算架构

一个典型的集群系统架构[16]如图 1 所示。系统中存在多个节点,每个节点上部署数量不等的 GPU 资源,节点之间基于互联网络方式实现通信。根据对计算资源需求的不同,GPU 应用程序可分为 3 类:① 单节点单 GPU 应用;② 单节点多 GPU 应用;③ 多节点多 GPU 应用。由于 GPU 卡计算、存储资源的限制,单个节点上所部署的 GPU 卡无法满足某些 GPU 应用程序的需求,甚至某些节点由于没有部署 GPU 卡而无法提供 GPU 计算服务。多节点多 GPU 计算架构为这种需求提供了合理的解决方案,通过将 GPU 应用程序划分为多个分布在不同节点上的进程的方式实现对其他节点上 GPU 资源的调用,满足其对超出本节点负载能力的 GPU 资源的需求。

多节点多 GPU 计算架构如图 2 所示。该架构下的 GPU 应用程序开发一般基于 MPI、CUDA 相结合的方式实现,要求开发人员显式展开分布式 MPI 编程,并根据 GPU 资源的分布情况,优化进程在节点上配置。每个部署 GPU 资源的节点需要启动一个或者多个 MPI 进程,各 MPI 进程基于 CUDA 接口访问本地 GPU 资源,MPI 进程之间基于 MPI 消息接口通信,实现计算任务的分发以及计算结果的回收。应用程序并行性的开发需要在 2 个层次上展开:① MPI 层:MPI 进程间并行性的开发;② CUDA 层:GPU 线程间并行性的开发。

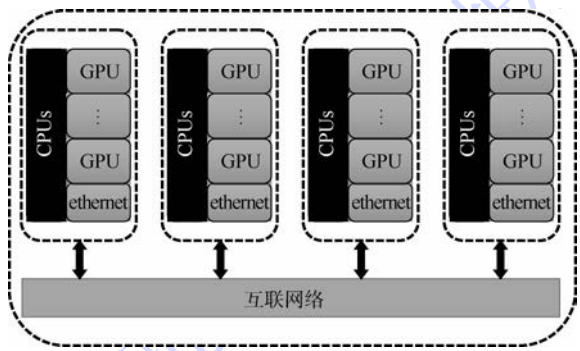


图 1　集群系统架构[16]

Fig. 1　Architecture of clustered system[16]

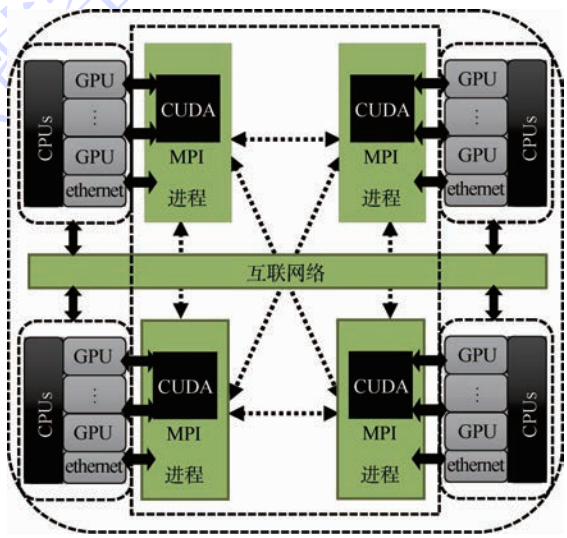


图 2　多节点多 GPU 计算架构

Fig. 2　Computing framework of multiple GPUs on multiple nodes

北航学报 赠阅

由于采用双层并行架构,程序的开发过程变得复杂,编程人员不仅要考虑层次内的并行性设计、优化问题,还需要考虑层次间的性能优化问题。此外,该架构灵活性较差,系统难以根据GPU资源的变化而动态改变程序在集群中的部署。为了提升应用程序的灵活性以及可移植性,并降低其开发难度,本文提出一种面向集群系统的新型虚拟化GPU计算架构。

2.2 虚拟化 GPU 计算平台框架

在本系统中,GPU的虚拟化基于CUDA运行时的虚拟化而实现,该设计为CUDA运行时引入了抽象层和通信层。通信层位于抽象层之下,基于MPI接口实现节点间通信。相对于rCUDA直接基于底层接口实现节点间通信,MPI能够适配各种硬件平台,具备更好的可移植性。通信层的引入为GPU应用程序屏蔽了MPI接口调用。

图3为虚拟化GPU计算平台逻辑视图[16]。虚拟化之后的CUDA运行时基于通信中间件实现对集群GPU资源的统一抽象与管理,GPU资源的物理分布对于应用程序变得透明。基于虚拟化CUDA运行时,集群系统中的所有GPU设备构成一个设备库,任意节点上用户程序均以统一的逻辑视图访问设备库的GPU资源。各应用程序可以像访问本地GPU一样调用集群系统中的所有GPU,开发人员只需专注于CUDA编程,而关于GPU资源的分布、访问以及MPI接口等问题,则交由虚拟化CUDA运行时处理。由于抽象层实现对原有CUDA API的兼容性封装,原有GPU应用程序则可以不经任何修改而无缝迁移到虚拟化GPU计算平台运行。

虚拟化GPU计算平台的软件栈如图4所示。虚拟化CUDA运行时由抽象层、通信层以及节点本地CUDA运行时三部分组成。通信层由分布于各节点的GPU访问代理服务器进程(以下简称

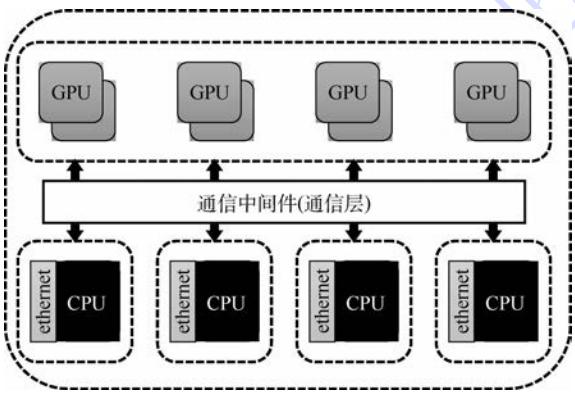


图3 虚拟化GPU计算平台逻辑视图[16]

Fig. 3 Logic view of virtualized GPU computing platform[16]

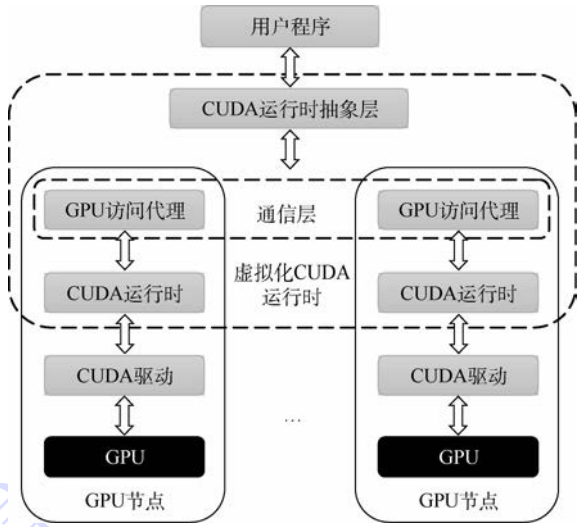


图4 虚拟化GPU计算平台软件栈

Fig. 4 Software stack of virtualized GPU computing platform

GPU访问代理)以及MPI通信接口构成,是实现对各节点GPU资源统一抽象与管理的关键。GPU访问代理负责处理所有对其所属节点GPU的访问请求,并将请求转发给节点本地CUDA运行时,每个GPU拥有独立的访问代理。抽象层同样具备MPI通信能力,基于MPI与各GPU访问代理进行交互。

当运行于某个节点上的应用程序发起GPU访问请求时,会首先调用CUDA运行时的抽象层,抽象层根据API参数以及集群系统当前GPU资源的配置以及状态,将访问请求传递到一个或多个GPU访问代理。访问代理根据参数调用本地CUDA运行时执行相应的操作,待操作完成之后再经由通信层回传相应的结果。

除了具备开发难度较低、可移植性好等优点之外,虚拟化GPU计算平台还具有以下特性:

1)充分利用系统资源,对于没有配备GPU的节点,也能运行GPU应用程序。

2)提升GPU应用程序的加速比,在单个节点上GPU资源有限的情况下,能够利用其他节点上的GPU资源,提升系统并行计算能力。

由于增加了额外的软件层次以及网络通信开销,系统性能的提升受到了一定程度的影响。

3 虚拟化 GPU 计算平台设计

如图4所示,虚拟化CUDA运行时采用分布式架构,通过在抽象层与节点CUDA运行时之间插入通信层实现对分布式GPU资源的虚拟化。

3.1 CUDA 运行时抽象层

为了实现对GPU资源的统一抽象与管理,抽

北航学报 赠阅

象层除了对集群系统中分布式 CUDA 运行时 API 进行兼容性封装之外,还对各节点上的 GPU 资源进行全局管理。在 GPU 应用程序启动阶段,抽象层主要负责:① 完成 MPI 运行环境以及 GPU 资源信息的初始化,并根据配置要求启动 GPU 访问代理,记录其所属的应用句柄;② 维护 GPU 代码段(FatBinary)的注册信息,如 FatBinary 在节点主机内存偏移位置、空间大小以及对应的 GPU 设备编号等;③ 记录核函数以及 GPU 全局变量信息及其所属的 FatBinary 编号;④ 给目标 GPU 访问代理传递参数,启动核函数,开始计算过程;⑤ 为目标节点 GPU 和本地主机申请、释放内存空间,传输数据。此外,在 GPU 应用程序结束阶段,抽象层还需要注销 FatBinary 信息,并释放其所占用的内存空间。以上各项操作需由各 GPU 访问代理以及节点本地 CUDA 运行时配合完成。

3.2　GPU 访问代理

GPU 访问代理是衔接抽象层与节点本地 CUDA运行时的重要部件。GPU 访问代理与其所服务的 GPU 应用程序具有相同的生命周期,随着 GPU 应用程序的启动而启动,并在完成 GPU 计算任务之后退出。每个 GPU 访问代理同时只为一个特定 GPU 应用程序提供服务,当新的 GPU 应用程序启动时,系统根据参数配置,在指定的一个或者多个节点上启动数量不等的 GPU 访问代理。应用程序不同,其所拥有的 GPU 访问代理在集群系统中的数量与分布也不同。对于具体的 GPU 应用程序,GPU 访问代理的数量与分布由用户或者上层作业调度系统指定。

GPU 访问代理是核函数调用以及数据传输的最终执行者,GPU 访问代理基于 MPI 方式在抽象层与节点本地 CUDA 运行时之间执行核函数调用、参数传递以及数值拷贝等操作。在系统初始化阶段,GPU 访问代理还需要完成 FatBinary、GPU 核函数以及 GPU 全局变量在其所属节点环境的注册工作,为 GPU 计算任务创建执行环境。鉴于同一节点上的不同 GPU 可能会服务于不同的 GPU 应用程序,该节点上会出现不同 GPU 访问代理隶属于不同的 GPU 应用程序的情况。

3.3　节点本地 CUDA 运行时

节点本地 CUDA 运行时直接继承原有 CUDA 运行时,服务于所有对本地 GPU 资源的访问请求。抽象层维护集群系统可用 GPU 资源的全局状态信息,而节点本地 GPU 运行时则只负责维护节点本地可用 GPU 资源的状态信息,并保持与抽象层资源信息的一致性。

3.4　数据传输

在虚拟化 GPU 计算平台下,一个典型的 GPU 程序运行流程如图 5 所示。主机与 GPU 的交互主要包括核函数调用以及数据传输 2 类操作。对于数据密集型的 GPU 应用程序,每次计算过程可能需要一到多次数据传输,这些操作均通过CUDA API 实现。CUDA API 调用需要由抽象层、通信层以及节点本地 CUDA 运行时 3 部分共同完成,涉及数据在不同组件、不同节点之间的传输,因此,数据传输性能对应用程序的性能会产生重要的影响,是系统设计要考虑的重点问题。拟采用流水化通信技术,实现节点间数据的高效传输。

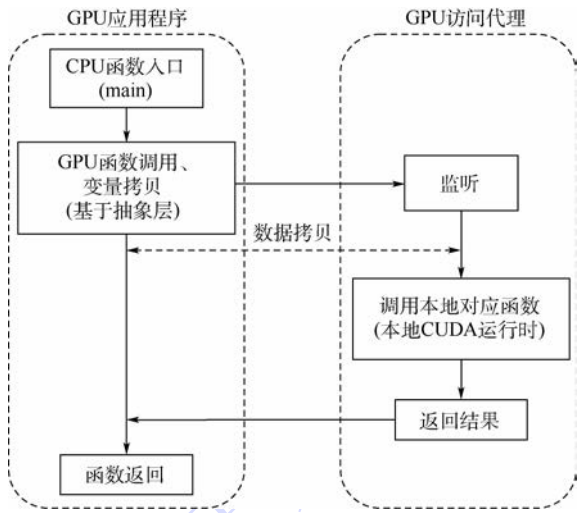


图 5　GPU 应用程序运行流程简图

Fig. 5　Simplified execution flowchart of GPU program

4　流水化通信技术

一次 GPU 计算过程可分为 3 个执行步骤:

1) 将输入数据由主机内存拷贝到 GPU 内存。

2) 启动 GPU 核函数,开始 GPU 计算过程。

3) 将计算结果由 GPU 内存拷贝到主机内存。

大多计算任务往往需要多次循环执行以上 3 个步骤。最新的 GPU 架构配备 2 个复制引擎,支持数据双向并行传输,因此理论上 3 个步骤可以并行执行,在整个计算任务的循环执行过程中,考虑采用多执行流模型,将 3 个执行步骤流水化,充分发挥系统性能。

对于虚拟化 CUDA 运行时环境,数据传输又可以分为 2 个步骤,即 GPU 应用程序与目标 GPU 访问代理之间的数据传输以及目标 GPU 访

北航学报 赠阅

问代理与目标 GPU 之间的数据传输。因此,一次完整的计算过程可以划分 5 个执行步骤,其流水线时间框图如图 6 所示,各步骤对应的操作如下:

步骤 1 GPU 应用程序将数据传输到目标 GPU 访问代理。

步骤 2 目标 GPU 访问代理将数据拷贝到目标 GPU 内存。

步骤 3 启动核函数,目标 GPU 开始计算过程。

步骤 4 将计算结果由目标 GPU 内存拷贝到目标 GPU 访问代理。

步骤 5 将计算结果由目标 GPU 访问代理传输到 GPU 应用程序。

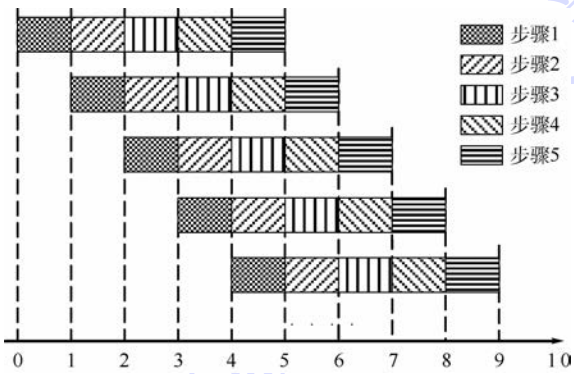


图 6 GPU 计算任务流水线

Fig. 6 Pipeline of GPU computing task

理想情况下,各步骤执行时间相同,而实际运行过程中,各步骤执行时间因应用程序以及系统状态的不同而发生变化。以上计算过程的流水化由虚拟化 CUDA 运行时的抽象层与节点本地 CUDA 运行时自动实现,不需要用户程序作出任何特殊设计或修改。其中步骤 1 ~ 步骤 2、步骤 4 ~ 步骤 5 的流水化由抽象层实施,而步骤 2 ~ 步骤 4 的流水化则由目标节点本地 CUDA 运行时实施。如图 7 所示,抽象层采用异步数据传输方式,对于每次数据传输,会首先创建专用数据传输线程并立即返回执行其他操作或者下一次数据传输。数据传输线程与目标 GPU 访问代理建立连接以执行数据传输。GPU 访问代理端监听到数据传输请求之后也会创建专用数据传输线程,与抽象层所创建的线程共同完成数据传输工作。数据传输完毕之后,各数据传输线程退出。为了实现步骤 1 ~ 步骤 2、步骤 4 ~ 步骤 5 的并行流水化,GPU 访问代理在数据传输的同时开启新的线程执行与本地 GPU 内存之间的数据拷贝工作,并采用多缓冲区机制,实现二者的并行执行。

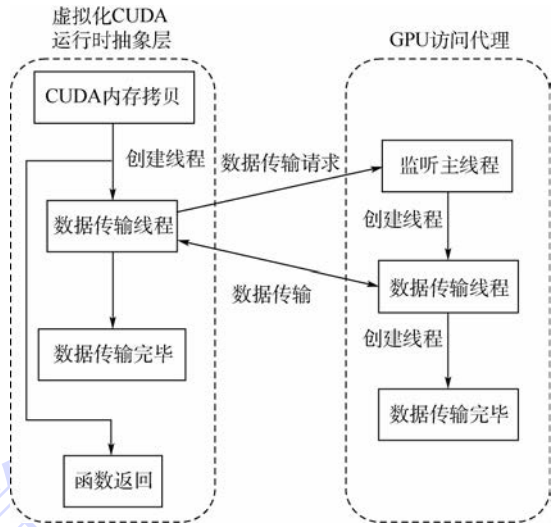


图 7 虚拟化 CUDA 运行时数据传输流程

Fig. 7 Data transmitting flow of virtualized CUDA run-time system

5 虚拟化 GPU 计算平台的性能实验

5.1 实验原理与方法

虚拟化 GPU 计算平台的实现,一方面使得 GPU 应用程序可以访问其他节点上 GPU 资源,提升计算性能。另一方面,由于虚拟化技术引入了额外的软件栈以及节点通信开销,一定程度上影响了系统整体性能的提升。本实验拟基于典型的 GPU 应用程序,测试虚拟化 GPU 计算平台的计算与通信性能。

计算性能的对比在计算平台与应用规模 2 个维度展开。

计算平台主要分为 2 类:

1) 节点本地 GPU 计算平台:应用程序的核函数只在节点本地 GPU 上运行。

2) 虚拟化 GPU 计算平台:应用程序的核函数在集群系统全局范围内的 GPU 上运行。

应用程序的规模的变化主要体现在其所处理的数据规模的变化。有关通信性能及开销的实验对比分析也在计算平台、数据规模 2 个维度展开。

5.2 实验环境

本实验集群环境配备 4 个节点,节点之间基于 InfiniBand(IB)网络互联,IB 网卡控制器型号为 Mellanox ConnectX-3 Double-Port 4X QDR,实测最大带宽约为 3 600 MB/s。每个节点配置 2 颗 Intel(R) Xeon(R) E5-2690 处理器,标称主频为 2.90 GHz,每颗处理器为 8 核 32 线程。单个节点内存容量为 160 GB,频率为 1 066 Hz。此外,每个节点配备 8 块 Tesla K20M GPU 卡,每个 GPU 拥

北航学报 赠阅

有 2496 颗流处理器以及 5 GB 内存。节点操作系统环境为 Ubuntu 14.04 LTS,CUDA 以及 MPI 环境分别为 CUDA 6.4、MVAPICH2 2.1。

5.3　实验内容与分析

5.3.1　通信性能

由于虚拟化 GPU 计算环境引入了节点通信开销,因此通信性能是考量计算平台性能的重要指标。流水化通信技术实现对节点通信开销的隐藏,并充分利用 IB 带宽。实验测量本地主机到目标 GPU 内存的异步传输带宽。按不同的传输缓冲区设置,实验被分为 6 组,其传输缓冲区大小依次设置为 64 KB,128 KB,…,2 MB 不等。每组实验的传输带宽如图 8 所示。

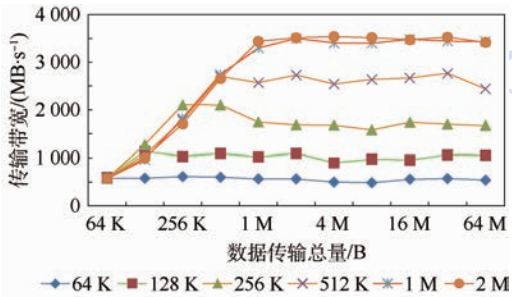


图 8　虚拟化 CUDA 运行时数据传输带宽

Fig. 8　Data transmitting bandwidth of virtualized CUDA run-time system

在每组实验中,数据传输带宽受限于传输缓冲区大小、数据传输量以及 IB 的带宽上限。在传输缓冲区大小固定的情况下,系统的数据传输带宽随着数据传输量的增加而呈线性增长趋势,直至数据传输量达到缓冲区容量的上限或者传输带宽达到 IB 的带宽上限。在达到 IB 带宽上限之前,传输缓冲区越大,传输带宽所能达到的上限值越大。实验表明,当传输缓冲区达到 1 MB 时,数据传输带宽达到 IB 带宽上限值。在到达传输带宽上限之前,增加缓冲区大小或者数据传输量能有效提升 IB 带宽的利用率,并隐藏虚拟化 CUDA 运行时所带来的开销。当传输缓冲区以及数据传输量均超过 1 MB 时,系统对虚拟化 CUDA 运行系统开销的隐藏能力达到上限,IB 带宽的利用率趋于饱和,传输速度为 3 550 MB/s,基本达到 IB 的传输带宽的实测上限值 3 650 MB/s。

图 9 与图 10 分别针对矩阵乘法以及 FFT 计算 2 种典型 GPU 应用程序展开数据传输延迟实验。矩阵乘法为计算密集型应用,而 FFT 则为数据密集型应用。在不影响性能评估的前提下,本实验仅仅分析由 GPU 应用程序节点到目标节点 GPU 的单向数据传输延迟,实验对比在如

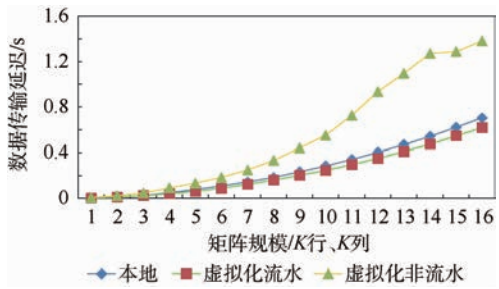


图 9　矩阵乘法数据传输延迟

Fig. 9　Data transmitting latency of matrix multiplication

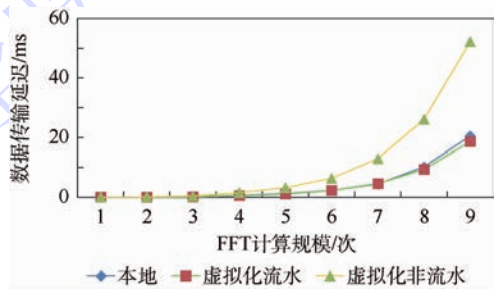


图 10　FFT 数据传输延迟

Fig. 10　Data transmitting latency of FFT

第 5.1 节所述的 2 类平台之间展开。为了突出通信流水化所带来的性能提升效果,虚拟化 GPU 计算平台上的实验又分为流水化传输和非流水化传输 2 类,因此最终实验数据分为 3 组。

在虚拟化 GPU 计算平台中,数据传输过程一般由节点间传输、目标节点主机内存到 GPU 内存传输两部分组成,而传统 GPU 计算平台下的数据传输仅仅需要完成主机内存到本地 GPU 内存传输过程。因此,如果采用非流水化数据传输,由于引入了额外的控制开销以及节点间传输,虚拟化 GPU 计算平台下的传输延迟大于传统 GPU 计算平台下的本地传输延迟。而采用流水化数据传输机制之后,目标节点主机内存到 GPU 内存的传输隐藏了节点间传输延迟以及相关控制开销,传输过程的总体延迟显著降低。实验结果表明,采用流水化通信技术之后,虚拟化 GPU 计算平台的数据传输延迟降低 50% ~70%。

在本实验环境中,由于 IB 传输带宽低于 PCIE 带宽,因此采用流水化通信技术的虚拟化 GPU 计算平台的数据传输相对于传统 GPU 计算平台的本地数据传输具有更大的延迟,系统的数据传输性能受限于节点间的 IB 互联带宽。而在图 9 与图 10 中,虚拟化 GPU 计算平台的流水化数据传输反而具有较低的延迟,这主要得益于 DMA 技术在 IB 带宽传输中的应用,这在计算性能实验部分也有所体现。在实际应用中,为了提

北航学报 赠阅

升节点间的数据传输性能,需要采用具有更高带宽的 IB 互联网络。

5.3.2 计算性能

计算性能实验环节评测 GPU 应用程序的单次计算过程从启动到结束的总体延迟,该延迟由主机内存到目标 GPU 内存的传输延迟、目标 GPU 代码执行时间以及目标 GPU 内存到主机内存的传输延迟 3 部分构成。数据传输是构成整体计算过程的不可获取的重要组成部分,随着数据传输所占比例的不同,节点间数据传输以及其开销对 GPU 应用程序的性能影响也各不相同。表 1 与表 2 分别为矩阵乘法与 FFT 的总体计算延迟,与通信性能实验类似,计算性能的实验也在 3 类平台分别展开,即节点本地 GPU 计算平台、基于流水化通信技术的虚拟化 GPU 计算平台以及非流水化的虚拟化 GPU 计算平台。

表 1 矩阵乘法计算延迟

Table 1 Total computing latency of matrix multiplication

矩阵规模/行、列	计算延迟/s		
	本地平台	虚拟化流水平台	虚拟化非流水平台
1024	0.0135	0.0131	0.0175
2048	0.0867	0.0846	0.1019
3072	0.2620	0.2565	0.3029
4096	0.5960	0.5849	0.6806
5120	1.1333	1.1192	1.2874
6144	1.9270	1.9087	2.1512
7168	3.0309	3.0047	3.2794
8192	4.4760	4.4430	4.7193
9216	6.3196	6.2937	6.6417
10240	8.6168	8.5922	9.0155
11264	11.4213	11.3951	11.9786
12288	14.7529	14.7424	15.6143
13312	18.6737	18.6943	19.6530
14336	23.2603	23.3639	24.2306
15360	28.5631	28.8291	29.6420
16384	34.5847	34.9069	35.8820

表 2 FFT 计算延迟

Table 2 Total computing latency of FFT

FFT 计算规模/行、列	计算延迟/ms		
	本地平台	虚拟化流水平台	虚拟化非流水平台
8	0.1014	0.2759	0.2167
16	0.1396	0.3135	0.3191
64	0.4128	0.6905	1.0005
256	1.7061	2.5051	3.0628
512	2.7434	3.4342	6.0090
1024	4.9513	5.6387	12.4230
2048	9.8512	10.6143	26.1070
4096	20.6565	20.5740	49.3056
8192	41.1518	39.8459	102.3965

在虚拟化 GPU 计算平台下,由于矩阵乘法计算为计算密集型应用,因此流水化数据传输技术对计算性能的提升有限,而 FFT 计算为数据密集型应用,流水化数据传输技术对计算性能的提升作用显著。实验数据显示,采用流水化数据传输技术的虚拟化 GPU 计算平台能够提供与本地 GPU 计算平台等同的计算性能。在实际应用中,在算法允许的情况下,可以将 GPU 应用程序需要处理的输入数据划分为多个片段,依次传输到 GPU 内存并展开计算,将计算与数据传输流水化,从而提高系统性能。

6 结 论

本文面向集群环境,设计并实现了一种新型的虚拟化 GPU 计算平台,将多节点多 GPU 资源抽象,形成统一的编程视图,使得 GPU 应用程序可以无差别地访问集群系统上的任一 GPU 资源。

1)与基于 MPI 编程的多节点多 GPU 应用程序不同的是,原有单节点多 GPU 应用程序在无需修改的情况下可以直接运行于虚拟化 GPU 计算平台,实现对多节点多 GPU 的访问,增加了应用程序的可移植性。同时,由于通信层基于 MPI 接口实现,屏蔽了底层硬件接口的差异性,系统具有更好的硬件平台普适性。

2)虚拟化 GPU 计算平台使得 GPU 应用程序摆脱了单个节点上 GPU 资源的限制,可以充分利用集群系统多个节点上的 GPU 资源,从而提升 GPU 代码的并行度,对于 GPU 应用程序具有良好的加速潜力。

3)基于流水化通信技术,本系统隐藏了节点间数据传输所带来的延迟以及控制开销。实验表明,虚拟化 GPU 计算平台可以提供与原有 GPU 计算平台等同的通信性能,因此在多数情况下 GPU 应用程序的整体性能不会因为系统引入了额外的软件栈以及通信环节而下降。

系统性能达到了预期目标,具有良好的应用前景。

参考文献(References)

[1] KIVITY A,KAMAY Y,LAOR D,et al. KVM:The Linux virtual machine monitor[EB/OL]. Proceedings of the Linux Symposium,Ottawa[2015-11-01]. https://www.kernel.org/doc/ols/2007/ols2007v1-pages-225-230.pdf.

[2] BARHAM P,DRAGOVIC B,FRASER K,et al. Xen and the art of virtualization[C]// Proceedings of the 19th ACM Symposium on Operating Systems Principles. New York: ACM,2003:164-177.

北航学报 赠阅

[3] NextION 2800-ICA-Flexible and manageable I/O expansion and virtualization[EB/OL]. Austin: NEXTIO[2015-11-01]. http://www.nextio.com/docs/NextIO20N2800-ICAIOConsolidationApplianceProduct Briefv0.18.pdf.

[4] SHREINER D. OpenGL programming guide[M]. 7th ed. Boston: Addison-Wesley Professional, 2009: 1-28.

[5] BLYTHE D. The direct 3D 10 system[J]. ACM Transactions on Graphics, 2006, 25(3): 724-734.

[6] NVIDIA. CUDA C programming guide[EB/OL]. Santa Clara: NVIDIA [2015-11-01]. https://docs.nvidia.com/cuda/cuda-c-programming-guide/index.html.

[7] KHRONOS OpenCL Working Group. OpenCL 2.0 specification [EB/OL]. OR: Khronos [2015-11-01]. http://www.khronos.org/registry/cl/specs/opencl-1.2.pdf.

[8] DOWTY M, SUGERMAN J. GPU virtualization on VMware's hosted I/O architecture[J]. ACM SIGOPS Operating Systems Review, 2009, 43(3): 73-82.

[9] LAGAR-CAVILLA H A, TOLIA N, SATYANARAYANAN M, et al. VMM-independent graphics acceleration[C]// VEE 07: Proceedings of the 3rd International Conference on Virtual Execution Environments. New York: ACM, 2007: 33-43.

[10] SHI L, CHEN H, SUN J. vCUDA: GPU accelerated high performance computing in virtual machines[C]// 23rd IEEE International Symposium on Parallel and Distributed Processing (IPDPS'09). Piscataway, NJ: IEEE Press, 2009: 418-428.

[11] GUPTA V, GAVRILOVSKA A, SCHWAN K, et al. GViM: GPU-accelerated virtual machines[C]// Proceedings of the 3rd ACM Workshop on System-level Virtualization for High Performance Computing. New York: ACM, 2009: 17-24.

[12] GIUNTA G, MONTELLA R, AGRILLO G, et al. A GPGPU transparent virtualization component for high performance computing clouds[C]// 16th International Euro-Par-Conference on Parallel Processing. Berlin: Springer, 2010: 379-391.

[13] KEGEL P, STEUWER M, GORLATCH S. dOpenCL: Towards a uniform programming approach for distributed heterogeneous multi-/many-core systems[C]// 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum (IPDPSW). Piscataway, NJ: IEEE Press, 2012: 174-186.

[14] BARAK A, BEN-NUN T, LEVY E, et al. A package for OpenCL based heterogeneous computing on clusters with many GPU devices[C]// 2010 IEEE International Conference on Cluster Computing Workshops and Posters (CLUSTER WORKSHOPS). Piscataway, NJ: IEEE Press, 2010: 1-7.

[15] DUATO J, PEÑA A J, SILLA F, et al. rCUDA: Reducing the number of GPU-based accelerators in high performance clusters[C]// 2010 International Conference on High Performance Computing and Simulation (HPCS). Piscataway, NJ: IEEE Press, 2010, 6: 224-231.

[16] PEÑA A J, REAÑO C, SILLA F, et al. A complete and efficient CUDA-sharing solution for HPC clusters[J]. Parallel Computing, 2014, 40(10): 574-588.

作者简介:

杨经纬　男,博士研究生。主要研究方向:计算系统结构、嵌入式系统、多核实时操作系统与实时调度。

E-mail: yaungjw@buaa.edu.cn

马凯　男,硕士研究生。主要研究方向:计算机系统结构、并行与分布式计算。

E-mail: makai@buaa.edu.cn

龙翔　男,博士,教授,博士生导师。主要研究方向:计算系统结构、并行与分布式系统、实时系统。

Tel.: 010-82339685

E-mail: long@buaa.edu.cn

北京航空航天大学学报

北航学报 赠阅

Virtualized GPU computing platform in clustered system environment

YANG Jingwei, MA Kai, LONG Xiang*

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

Abstract: A virtualized GPU computing platform is proposed for clustered systems, which are often equippedwith GPUs in some nodes. All GPUs in system are uniformly abstracted as virtualized ones in a commonly accessed resource pool. Legacy GPU programs can execute on the virtualized GPU computing platform without any modification and any free virtualized GPU in the common resource pool is available to it, which relieves the burden of MPI programming. The platform frees programs with the limit of GPUs in local node and makes it possible for them to access any available GPU in distributed nodes, leading to higher system utilization and throughput. Based on pipelined communication, the run-time overhead and inter-node transmitting latency in virtualized GPU computing platform are hidden by intra-node memory copying and GPU computing. Compared with the non-pipelined communication, the total transmission latency is decreased by approximately 50% – 70%. It results in a comparable performance with intra-node local data transmission.

Key words: GPU; MPI; CUDA; clustered systems; hardware acceleration; parallel computing; high performance computing

北京航空航天大学学报

Received: 2015-11-09; **Accepted**: 2016-01-15; **Published online**: 2016-03-15 14:40

URL: www.cnki.net/kcms/detail/11.2625.V.20160315.1440.008.html

* **Corresponding author**. Tel.: 010-82339685 E-mail: long@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0770

旋翼/平尾气动干扰对直升机配平特性的影响分析

谭剑锋[1,*]，吴超[2]，王浩文[3]，林长亮[4]

（1. 南京工业大学 机械与动力工程学院，南京 211816； 2. 南京航空航天大学 航空宇航学院，南京 210016；
3. 清华大学 航天航空学院，北京 100084； 4. 中航工业哈尔滨飞机工业集团有限责任公司 飞机设计研究所，哈尔滨 150066）

摘 要：旋翼/平尾气动干扰建模是直升机全机配平模型和飞行动力学模型的核心内容。常规全机配平模型主要通过旋翼下洗流或诱导速度等方式间接考虑旋翼对平尾的气动干扰作用，但未充分考虑旋翼对平尾非定常气动干扰产生的非线性气动载荷，因而仍难于准确体现旋翼/平尾气动干扰对全机配平特性的影响。为此本文基于非定常面元/黏性涡粒子法，通过在平尾面元中增加由旋翼和平尾尾迹时变干扰产生的非定常压力项，建立旋翼/平尾气动干扰模型，直接计算平尾非线性气动载荷，并耦合基于 GA/LM 混合优化的直升机全机配平方法，构建旋翼/平尾气动干扰作用下的直升机全机配平特性分析方法。通过计算UH-60A直升机的旋翼操纵量和机体姿态，并与试验测试值对比验证本文方法的准确性。通过与基于诱导速度考虑旋翼/平尾气动干扰的直升机配平结果比较表明，后者难于体现直升机低速纵向操纵量和机体俯仰角突增现象，而本文方法能较好地体现直升机低速纵向操纵量和机体俯仰角突增，且与 OH-6A、EH-101 等试验测量的特性一致。研究不同平尾构型对旋翼/平尾气动干扰下直升机全机配平特性的影响，分析表明低平尾产生较大的低速纵向操纵量突增，而高平尾则增加高速纵向操纵量；前置平尾产生较大的低速纵向操纵量突增，右旋直升机的右置平尾有利于减小低速纵向操纵量突增和机体俯仰角。

关 键 词：平尾气动；黏性涡粒子法；直升机配平；旋翼操纵；机体姿态
中图分类号：V211.52
文献标识码：A **文章编号**：1001-5965(2016)11-2349-09

直升机低速飞行特性对于武装直升机和运输直升机具有重要作用，同时也是直升机飞行动力学和空气动力学的研究重点和难点之一[1]。由于全尺寸直升机风洞试验难度较大，而理论分析仍然难于准确模拟低速前飞状态下的直升机非定常气动干扰特性，从而降低飞行动力学模拟分析精度。过去 40 多年间，众多技术先进的直升机（UH-60A[2]、AH-64[3]、EH-101[4] 和 NH-90[5]）均存在纵向操纵问题，国内于 2012 年对某型号直升机试飞表明低速前飞状态下存在显著的机身抬头现象[6]，而这些问题均与旋翼/平尾气动干扰对直升机全机配平和操纵特性影响直接相关，由此准确分析旋翼尾迹与平尾气动干扰对直升机全机配平特性的影响就显得至关重要。

直升机飞行动力学建模是分析全机配平特性的基础，且气动干扰建模是直升机飞行动力学模型和全机配平特性分析的核心内容之一。基于准定常假设，早期飞行动力学模型中的旋翼气动力忽略尾迹对诱导入流和非定常气动载荷的影响[7]，因而主要用于直升机全机配平和稳定性分

收稿日期：2015-11-23；**录用日期**：2016-01-27；**网络出版时间**：2016-04-13 15:26
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160413.1526.004.html
基金项目：国家自然科学基金（11502105）；江苏省自然科学基金（BK20161537）；江苏省普通高校自然科学研究计划（15KJB130004）

* **通讯作者**：Tel.：025-58139357 E-mail：Jianfengtan@njtech.edu.cn

引用格式：谭剑锋，吴超，王浩文，等. 旋翼/平尾气动干扰对直升机配平特性的影响分析[J]. 北京航空航天大学学报，2016，42(11)：2349-2357. TAN J F，WU C，WANG H W，et al. Analysis on influence of rotor/empennage aerodynamic interaction on helicopter trim characteristic[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2349-2357(in Chinese).

北航学报 赠阅

析。随着直升机对飞行品质的要求提高,建立包括旋翼入流动态特性的高阶直升机飞行动力学模型得到逐步发展[8-10],旋翼气动力模型普遍采用线性非均匀非定常入流模型、高阶动态入流模型以及二维翼型非定常气动力模型,并根据静态吹风数据或经验公式体现旋翼对机身和平尾的气动干扰作用。相比早期直升机飞行动力学模型,计算精度得到改善,但仍然未能准确考虑旋翼畸变尾迹对平尾气动和操纵特性的影响。为此,通过自由尾迹模型模拟旋翼畸变尾迹,得到更真实的桨盘诱导速度分布,并计算机身和平尾处诱导速度,而后通过经验模型或风洞试验数据间接考虑旋翼/机身/平尾气动干扰,建立精度更高的直升机飞行动力学模型[11-13],但仍未很好体现旋翼尾迹与平尾的非定常气动干扰对全机配平特性和操纵特性的影响,因而难于准确捕捉直升机低速前飞状态下的操纵量突增和显著的机体"抬头"现象[11]。

为更准确分析旋翼尾迹与平尾气动干扰对直升机全机配平特性的影响,直升机飞行动力学模型需嵌入旋翼/平尾气动干扰模型。旋翼/平尾气动干扰模型主要采用相对简单的旋翼尾迹模型。通过采用势流理论和升力面方法预测靠近平尾的旋翼桨尖涡强度和平尾非定常载荷[14]。由于预定尾迹难于准确得到不同飞行状态下的旋翼尾迹,因此分析精度较低。通过采用平面尾迹模型分析高速前飞状态的平尾载荷[15],然而平面尾迹难于准确得到低速状态的畸变尾迹,因此难于准确分析旋翼与平尾的气动干扰特性。基于 Eulerian 体系的涡量输运法(Vorticity Transport Method, VTM)可用于分析旋翼流场和旋翼/机身气动干扰[16],但基于 Eulerian 体系而存在非物理耗散,因而该方法仍然存在旋翼尾迹耗散问题,且基于升力面模型而难于体现旋翼尾迹对平面表面的非线性影响。基于 Eulerian 体系的计算流体力学(Computational Fluid Dynamics, CFD)能较准确分析旋翼流场[17],但数值耗散较大,计算资源需求仍然较多,且计算得到的旋翼尾迹耗散较快[18],因此在分析旋翼尾迹与平尾的气动干扰方面仍需发展。

基于 Lagrangian 体系定义涡量场,并求解 Navier-Stokes 方程,黏性涡粒子法可较好地分析旋翼尾迹的非定常特性,并于近年成功用于分析旋翼流场[19-20]。耦合非定常面元,黏性涡粒子法[21]能较准确捕捉旋翼尾迹的非定常特性,并能较准确体现旋翼干扰下的平尾和尾桨气动载荷[22-23]。为更准确分析旋翼尾迹与平尾的非定常气动干扰对直升机全机配平特性的影响,本文将基于非定常面元/黏性涡粒子法,通过在平尾面元中增加由旋翼和平尾尾迹时变干扰产生的非定常压力项,建立旋翼/平尾气动干扰模型,直接计算平尾非线性气动载荷,并嵌入直升机飞行动力学模型,耦合基于 GA/LM 混合优化的直升机全机配平方法,构建基于旋翼/平尾气动干扰的直升机全机配平特性分析方法。通过计算 UH-60A 直升机的旋翼操纵量和机体姿态,并与 UH-60A、OH-6A、EH-101 试验结果对比,分析旋翼/平尾气动干扰对全机配平特性的影响,并分析平尾构型对直升机全机配平特性的影响规律。

1 计算方法

1.1 旋翼/平尾气动干扰模型

1) 非定常面元/黏性涡粒子法

在桨叶和平尾的表面附近和尾流区之外,可假设流场无黏、无旋、不可压,由此旋翼流场势函数 ϕ_v 由旋翼桨叶、平尾和尾迹涡面上的汇 σ_s 和偶极子 μ_d 组成[21]。

$$\phi_v(x,y,z,t)=\frac{1}{4\pi}\int_{S_B}\mu_d\boldsymbol{n}\cdot\nabla\left(\frac{1}{\boldsymbol{r}_t}\right)\mathrm{d}S-\frac{1}{4\pi}\int_{S_B}\sigma_s\left(\frac{1}{\boldsymbol{r}_t}\right)\mathrm{d}S+\frac{1}{4\pi}\int_{S_W}\mu_d\boldsymbol{n}\cdot\nabla\left(\frac{1}{\boldsymbol{r}_t}\right)\mathrm{d}S \quad (1)$$

式中:S_B 和 S_W 分别为旋翼或平尾等物面和尾迹涡面;$\boldsymbol{n}$ 为旋翼或平尾的外法线单位矢量;$\boldsymbol{r}_t=(x,y,z)$ 为时变位置矢量。

由相对于旋翼或平尾表面的法向速度为零和旋翼或平尾对流体扰动在无限远处为零条件可得

$$\begin{cases}\dfrac{\partial\phi_v}{\partial n}-\boldsymbol{v}_B\cdot\boldsymbol{n}=0 & \text{物体表面}\\ \lim\nabla\phi_{r_t\to\infty}=0 & \text{远场边界}\end{cases} \quad (2)$$

式中:$\boldsymbol{v}_B$ 为表面速度矢量。

由物面 Neumann 边界条件、桨叶和平尾后缘 Kutta 条件,可将式(2)的边界条件转为代数方程,由此求得桨叶和平尾的汇 σ_s 和偶极子 μ_d 分布。

旋翼和平尾的涡量场可由 Navier-Stokes 方程描述[19-22]。基于 Lagrangian 体系定义涡量场,得到黏性涡粒子控制方程:

$$\frac{\mathrm{d}\boldsymbol{x}_p}{\mathrm{d}t}=\boldsymbol{u}_p(\boldsymbol{x}_p,t) \quad (3a)$$

$$\frac{\mathrm{d}\boldsymbol{\alpha}_p}{\mathrm{d}t}=\nu_p\nabla^2\boldsymbol{\alpha}_p+\nabla\boldsymbol{u}_p\cdot\boldsymbol{\alpha}_p \quad (3b)$$

式中:$\boldsymbol{x}_p$ 为涡粒子的位置矢量;$\boldsymbol{\alpha}_p$ 为涡粒子的涡

北航学报 赠阅

矢量；ν_p 为流体运动黏性系数；$\boldsymbol{u}_p$ 为涡粒子速度矢量。式(3a)体现涡量的输运效应；式(3b)体现涡量的耗散和拉伸效应，可采用粒子强度交换法计算旋翼尾迹涡量耗散效应 $\nu_p \nabla^2 \boldsymbol{\alpha}_p$；通过采用直接求解法计算由速度梯度产生的旋翼尾迹涡量拉伸和旋转效应 $\nabla \boldsymbol{u}_p \cdot \boldsymbol{\alpha}_p$。

基于涡量等效原则，将桨叶和平尾尾随偶极子面元转为由等效面涡 γ 和线涡 μ_l 积分的涡粒子[21]，并基于 Neumann 条件计算得到面元/黏性涡粒子法。基于此方法，通过非定常面元模拟旋翼和平尾气动载荷，黏性涡粒子法模拟旋翼和平尾的畸变尾迹特性，得到随前飞速度变化的平尾气动载荷[22]。

2）平尾气动模型

前飞状态下，由于旋翼桨尖涡贴近并穿过平尾表面，因此对平尾气动载荷产生较显著的影响。在旋翼尾迹黏性涡粒子法的黏性效应求解中，对于无物面边界的旋翼涡量场，对称粒子强度交换法的涡量守恒，即 $\sum \mathrm{d}\boldsymbol{\alpha}_p/\mathrm{d}t = 0$，但低速飞行时，平尾将浸润在旋翼涡量场中，旋翼桨尖涡贴近平尾表面，受到平尾边界的剪截作用，从而导致部分旋翼涡量穿过平尾上下表面边界而丢失，由此改变旋翼尾迹干扰下的平尾非定常气动载荷。为此本文基于镜面映射方式确保旋翼尾迹涡量守恒(见图 1)。图 1 中涡粒子 i 在平尾表面附近的部分涡量进入平尾内部。假设涡粒子 i 的涡强为均匀分布，则进入平尾内部的涡强可由涡粒子涡量与体积 V_{ACB} 之积求得。通过平尾表面涡量镜像映射方式，该涡强也等于镜像涡量与体积 V_{ADB} 之积，因此平尾镜面映射方式可确保旋翼涡量守恒，则镜面映射后的涡粒子涡量可表示为

$$\alpha_x^i = \alpha_x, \quad \alpha_y^i = \alpha_y, \quad \alpha_z^i = -\alpha_z \tag{4}$$

式中：α_x 和 α_y 分别为涡粒子 i 的涡量在平尾表面 x 和 y 方向的分量；α_z 为涡粒子 i 的涡量在平尾表面的法线分量；α_x^i 和 α_y^i 分别为镜面涡粒子 i' 的涡量在平尾表面 x 和 y 方向的分量；α_z^i 为镜面涡粒子 i' 的涡量在平尾表面的法线分量。

由于旋翼尾迹在靠近平尾表面时存在速度变化，且受到旋翼尾迹几何的时变影响，平尾表面压力发生变化，由此产生时变非线性的平尾气动载荷。基于非定常伯努利方程，采用流场速度势和平尾表面速度计算平尾表面的压力分布。为更准确体现由旋翼尾迹速度和几何的时变特性对平尾表面非定常压力的影响，本文在计算平尾气动载荷过程中增加由旋翼和平尾尾迹产生的非定常压力项以及由旋翼桨叶时变气动载荷诱发的非定常压力项，由此平尾表面无量纲压力系数为

$$C_p = \frac{p_p - p_{ref}}{(1/2)\rho_d v_{ref}^2} = 1 - \frac{v_s^2}{v_{ref}^2} - \frac{2}{v_{ref}^2}\left(\frac{\partial \phi_{blade}}{\partial t} + \frac{\partial \phi_{wake}}{\partial t}\right) \tag{5}$$

式中：$\boldsymbol{v}_s$、p_p、p_{ref}、$\boldsymbol{v}_{ref}$、ρ_d、ϕ_{blade} 和 ϕ_{wake} 分别为平尾表面的空气速度、当地压力、参考压力、参考速度、空气密度、桨叶产生的速度势和尾迹产生的速度势。

旋翼桨叶的时变诱导影响可基于速度势的微分直接求解得到

$$\frac{\partial \phi_{blade}}{\partial t} = \frac{\phi_{blade}^{t} - \phi_{blade}^{t-\Delta t}}{\Delta t} \tag{6}$$

旋翼和平尾尾迹产生的非定常压力项可通过计算尾迹对平尾表面面元的诱导速度和尾迹自身速度之积得到，即

$$\frac{\partial \phi}{\partial t}_{wake} = -\sum \frac{\partial \phi}{\partial \boldsymbol{x}} \cdot \frac{\partial \boldsymbol{x}'}{\partial t} = -\sum \boldsymbol{v}_{ind}(\boldsymbol{x}) \cdot \boldsymbol{v}_{wake}(\boldsymbol{x}') \tag{7}$$

式中：$\boldsymbol{v}_{ind}$ 为尾迹对平尾表面面元的诱导速度；$\boldsymbol{v}_{wake}$ 为尾迹自身速度。

基于此，旋翼尾迹干扰下的平尾非定常气动载荷可表示为

$$\Delta \boldsymbol{F}_k = -C_{pk}(\rho_d v_{ref}^2/2)_k \Delta S_k \boldsymbol{n}_k \tag{8}$$

式中：$\Delta \boldsymbol{F}_k$ 为表面面元气动力；ΔS_k 为面元面积；C_{pk} 为面元压力系数；$\boldsymbol{n}_k$ 为面元外法线矢量。

1.2　直升机全机配平模型

直升机全机配平建模是直升机空气动力学、飞行力学及旋翼结构动力学的基础，在机体姿态响应计算和稳定性分析方面均发挥重要作用。直升机全机配平的基础是飞行动力学模型，目前国内外已对全机飞行动力学建模做了较深入的研究[7-13]。基于直升机旋翼、机身、平尾及尾桨等典型气动力部件的动力学方程和运动方程，建立直升机全机动力学方程[24]为

$$\begin{cases} \dot{\boldsymbol{V}} = \boldsymbol{F}/m - \boldsymbol{\Omega V} \\ \dot{\boldsymbol{W}} = \boldsymbol{I}^{-1}\boldsymbol{M} - \boldsymbol{I}^{-1}\boldsymbol{\Omega I W} \\ \dot{\boldsymbol{\alpha}} = \boldsymbol{E W} \end{cases} \tag{9}$$

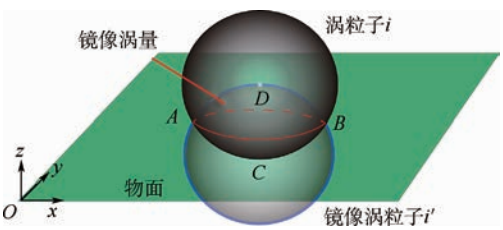


图 1　涡粒子涡量镜像映射

Fig. 1　Mirror-mapping of vortex particle vorticity

式中：m 为全机质量；$\boldsymbol{V}=(u,v,w)^{\mathrm{T}}$ 为线速度；$\boldsymbol{W}=(p,q,r)^{\mathrm{T}}$ 为角速度；$\boldsymbol{F}$ 为各部件合力；$\boldsymbol{M}$ 为外力矩；$\boldsymbol{I}$ 为全机惯性矩阵；$\boldsymbol{\Omega}$ 为叉乘向量算子；$\boldsymbol{\alpha}=(\phi,\theta,\psi)^{\mathrm{T}}$ 为欧拉角；$\boldsymbol{E}$ 为角速度至欧拉角速度的转换矩阵。

全机配平点定义为状态量导数为常数或零，同时操作输入为固定值。对于直升机稳定飞行状态，$\dot{\boldsymbol{V}}$、$\dot{\boldsymbol{W}}$、$\dot{\boldsymbol{a}}$ 和 $\ddot{\boldsymbol{a}}$ 为零（$\boldsymbol{a}$ 为一阶谐波形式的挥舞运动向量），$\dot{\boldsymbol{\alpha}}$ 及 $\boldsymbol{V}$、$\boldsymbol{W}$、$\boldsymbol{a}$ 均为定值，由此直升机全机飞行动力学方程可表示为状态方程，即

$$\dot{\boldsymbol{X}}_{\mathrm{e}}=g(\boldsymbol{X}_{\mathrm{e}},\boldsymbol{U}_{\mathrm{e}}) \tag{10}$$

式中：e 表示平衡状态；$\boldsymbol{X}_{\mathrm{e}}=[\boldsymbol{V}_{\mathrm{e}},\boldsymbol{W}_{\mathrm{e}},\boldsymbol{\alpha}_{\mathrm{e}},\boldsymbol{a}_{\mathrm{e}}]^{\mathrm{T}}$ 为配平状态量；$\boldsymbol{U}_{\mathrm{e}}=(\theta_0,\theta_{1\mathrm{c}},\theta_{1\mathrm{s}},\theta_{\mathrm{T0}})_{\mathrm{e}}^{\mathrm{T}}$ 为配平操纵量。

基于状态方程式(10)，直升机非线性方程组可表示为

$$g_i(\boldsymbol{X}_I)=0 \qquad i=1,2,\cdots,9 \tag{11}$$

为此，构造目标函数：

$$J=h(\boldsymbol{X}_I)=\min\sum_{i=1}^{9}g_i^2(\boldsymbol{X}_I)=g(\boldsymbol{X}_I)^{\mathrm{T}}g(\boldsymbol{X}_I) \tag{12}$$

当 $J=0$ 时，$\boldsymbol{X}_I$ 为式(11)的精确解，从而将直升机全机配平转变为优化问题，也即寻找 $\boldsymbol{X}_I$ 使 J 最小。针对直升机稳定飞行状态，以 UH-60A 直升机为例，旋翼总距范围为[9.9°,25.9°]，纵向周期变距范围为[-8.0°,8.0°]，横向周期变距范围为[-12.5°,16.3°]，尾桨总距范围为[4.5°,36.5°]，旋翼挥舞角范围为[-6°,25°]，俯仰角和滚转角范围为[-90°,90°]，从而构建得到配平解空间的约束条件。根据笔者建立的基于 GA/LM 混合优化方法[24]实现全局直升机全机配平解搜索，采用遗传算法(Genetic Algorithm，GA)搜索全局较优配平解，而后将较优解定义为 Levenberg-Marquardt(LM)优化方法的初值，利用此方法的快速收敛特性搜索物理最优配平真解，从而得到直升机全机配平操纵量。

1.3 耦合方法

常规直升机全机配平模型主要通过计算旋翼下洗流或尾迹诱导速度，而后基于二维翼型气动特性间接考虑旋翼对平尾气动干扰，但未充分体现旋翼尾迹对平尾非定常气动干扰产生的非线性气动载荷影响。为此，本文将第 1.1 节的非定常面元/黏性涡粒子法嵌入第 1.2 节的直升机飞行动力学模型中，根据式(8)计算旋翼尾迹干扰下的平尾非定常气动载荷，代入直升机动力学模型式(11)，以体现直升机低速前飞状态下，旋翼/平尾气动干扰对直升机全机配平特性的影响。首先基于旋翼诱导入流和二维翼型气动特性的 GA/LM 混合优化全机配平方法计算直升机初始操纵量，而后采用非定常面元/黏性涡粒子法计算旋翼非定常气动载荷和平尾非定常气动载荷，随后将旋翼和平尾非定常气动载荷代入直升机飞行动力学方程，通过基于 GA/LM 混合优化的直升机全机配平方法重新计算直升机的操纵量和机体姿态，并根据解空间约束和优化条件得到最终操纵量和机体姿态。计算流程如图 2 所示。

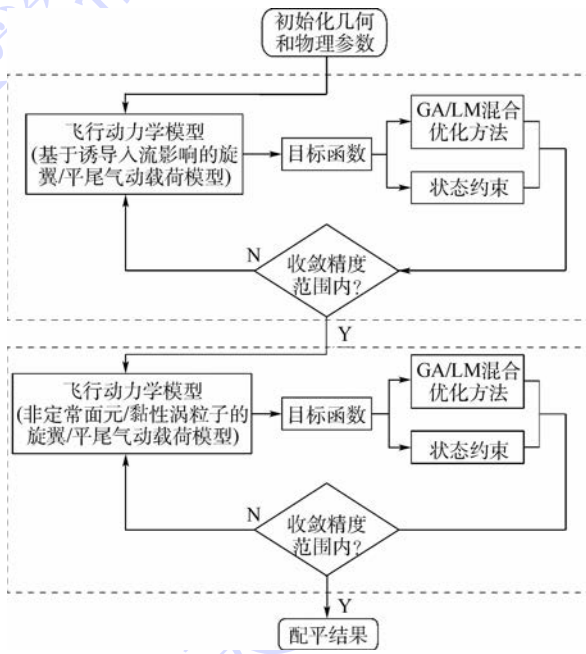


图 2 直升机全机配平流程图

Fig. 2 Flowchart of whole helicopter trim

1.4 方法验证

为验证本文方法的准确性，计算 UH-60A 各飞行速度操纵量，并与飞行测试数据[25]对比验证。计算得到 UH-60A 旋翼操纵量和机体姿态变化特性如图 3 所示，其中旋翼总距 θ_0、纵向周期变距 $\theta_{1\mathrm{c}}$、横向周期变距 $\theta_{1\mathrm{s}}$ 以及尾桨总距 θ_{T0} 均为桨根值，θ_{pitch} 为俯仰角，θ_{roll} 为滚转角。从图 3 中可以看出，本文计算得到的旋翼操纵量和机体俯仰角与 UH-60A 试飞测量值总体吻合较好，由此表明本文方法的可行性。各前飞状态，旋翼纵向操纵量与试飞测量值吻合较好，机体姿态与试飞测试略有误差，由于暂未直接计算旋翼与机身、旋翼与尾桨之间干扰产生的非定常气动载荷，因而机体俯仰角与试飞测量值略有误差。

北航学报 赠阅

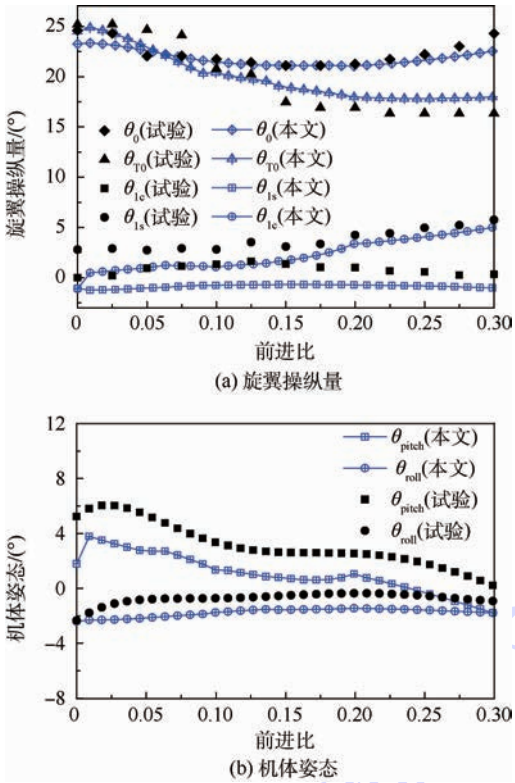


图 3　旋翼操纵量和机体姿态的变化

Fig. 3　Variation of rotor control and fuselage attitude

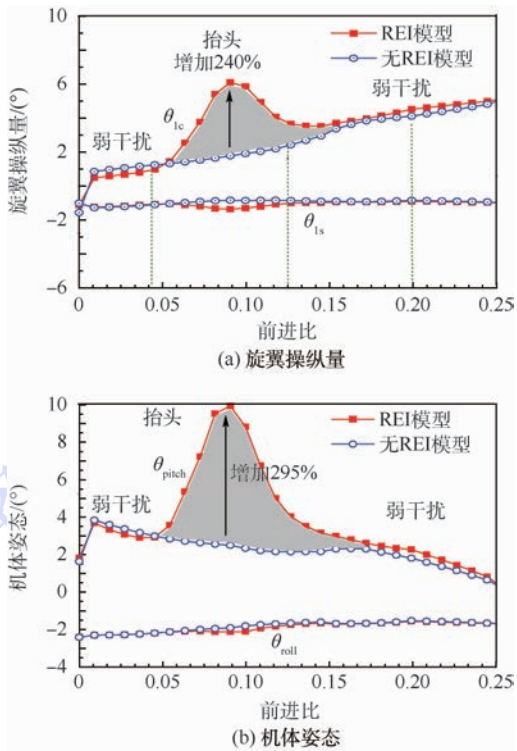


图 4　旋翼/平尾气动干扰下的旋翼操纵量和机体姿态

Fig. 4　Rotor control and fuselage attitude under rotor/empennage aerodynamic interaction

2　旋翼/平尾气动干扰对全机配平特性的影响

为分析旋翼/平尾气动干扰对直升机全机配平特性的影响，采用本文耦合旋翼/平尾气动干扰(Rotor/Empennage aerodynamic Interaction，REI)模型的全机配平方法和通过诱导速度考虑旋翼/平尾气动干扰的常规全机配平方法[24]分别计算UH-60A固定低平尾下的旋翼操纵量和机体姿态(见图4)，并与OH-6A和EH-101试验结果对比。

由于难于获得OH-6A和EH-101试验的详细测试数据，因此本节将主要对比计算结果与试验结果的特性趋势。

在悬停到小速度前飞的过渡阶段，由于平尾气动效率较低，旋翼尾迹/平尾干扰较弱，即使直接考虑旋翼对平尾的非定常影响，平尾气动载荷也较小，因此本文方法计算的旋翼纵向操纵量和机体俯仰角与基于诱导速度的常规全机配平方法计算结果基本相同。低速阶段，由于旋翼尾迹/平尾气动干扰显著，旋翼尾迹贴近平尾表面移动，使得平尾气动载荷迅速增加，并随旋翼尾迹的远离又快速减小，因此采用本文方法计算得到的旋翼纵向操纵量和机体俯仰角显著增加，形成典型的“抬头”现象，并与UH-60A[2]、OH-6A试验[3]中所出现的纵向操纵量突增(见图5(a))、EH-101飞行试验[4]中所出现的机体俯仰角快速增加相吻合(见图5(b))。受旋翼/平尾非定常气动干扰的影响，旋翼纵向操纵量和机体俯仰角分别增加240%和295%，纵向操纵量大于中等速前飞(200 km/h)的纵向操纵量，增量约为21.2%，且机体俯仰角高达10°。而基于诱导速度的常规全机配平方法计算的旋翼纵向操纵量和机体俯仰角未明显增加，并未出现UH-60A、OH-6A、EH-101飞行测试中出现的纵向操纵量和机体俯仰角突增现象[2-4]，因此通过诱导速度间接考虑旋翼/平尾气动干扰的全机配平难于准确反映低速状态旋翼/平尾气动干扰对旋翼纵向操纵量和机体姿态特性的影响。中等速度和高速前飞状态下，旋翼尾迹离开平尾上表面，旋翼与平尾气动干扰强度较弱，平尾气动载荷变化较小，因此直接计算旋翼/平尾非定常气动干扰后得到的旋翼纵向操纵量和机体俯仰角与基于诱导速度的计算结果相似。因此，通过直接计算旋翼尾迹干扰下的平尾非定常气动载荷，本文方法更准确体现旋翼/平尾气动干扰对直升机全机配平特性影响，并更准确分析直升机低速纵向操纵量和机体俯仰突增特性。

北航学报 赠阅

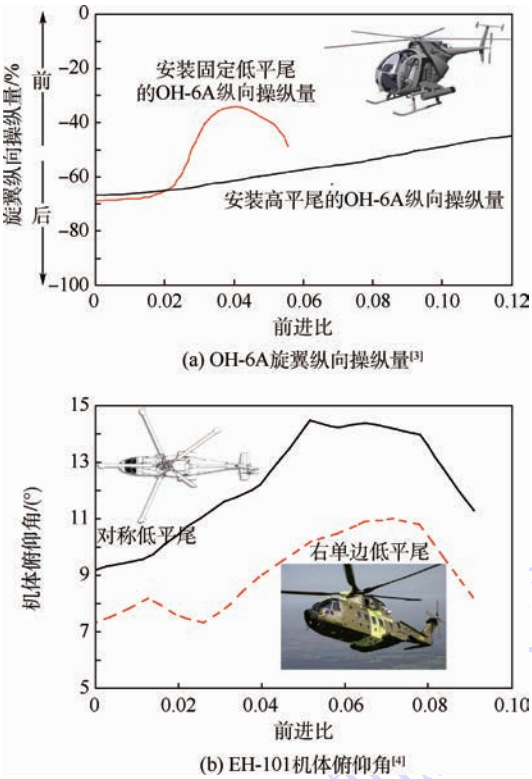


图 5 OH-6A 旋翼纵向操纵量和 EH-101 机体俯仰角

Fig. 5 Rotor longitudinal control of OH-6A and fuselage pitch angle of EH-101

3 平尾构型对直升机全机配平特性的影响

直升机[26]平尾构型定义如图 6 所示。

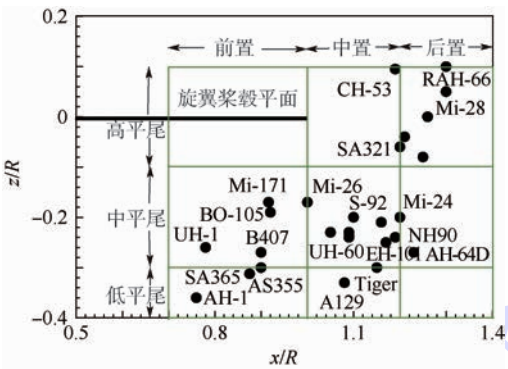


图 6 直升机平尾位置

Fig. 6 Empennage's position of helicopters

3.1 平尾垂向位置

根据图 6 的平尾构型定义,保持平尾的纵向位置不变,将平尾高度分别调整为 $-0.328R$、$-0.24R$、$0.05R$,得到低、中、高平尾。采用本文方法计算得到低、中、高平尾的旋翼操纵量和机体姿态随前飞速度变化,如图 7 所示。

过渡阶段,由于旋翼尾迹与平尾干扰较弱,3 种平尾向下载荷基本相同,因此旋翼纵向操纵量和机体俯仰角均相似。低速阶段,由于旋翼尾迹与低、中置平尾干扰显著,旋翼尾迹贴近低、中平尾表面,低、中置平尾向下气动载荷较大,因而旋翼纵向操纵量和机体俯仰角突增量较大。相比于低平尾,旋翼尾迹的干扰作用减弱,气动载荷减小,中平尾旋翼纵向操纵量和机体俯仰角明显减小,约为低平尾的 58.6% 和 56.5%。由于旋翼尾迹远离高平尾下表面,旋翼尾迹/高平尾干扰较弱,因此高平尾旋翼纵向操纵量和机体俯仰角缓慢增加。中高速阶段,旋翼尾迹逐步远离低、中置平尾上表面,旋翼尾迹/平尾干扰较弱,气动载荷较小,因此中、低平尾旋翼纵向操纵量和机体俯仰角突增量较小,而旋翼尾迹逐步靠近高平尾,并贴近高平尾表面,旋翼尾迹/高平尾气动干扰仍然明显,气动载荷较大,因此高平尾旋翼纵向操纵量和机体俯仰角均出现一定突增,但幅度较弱。

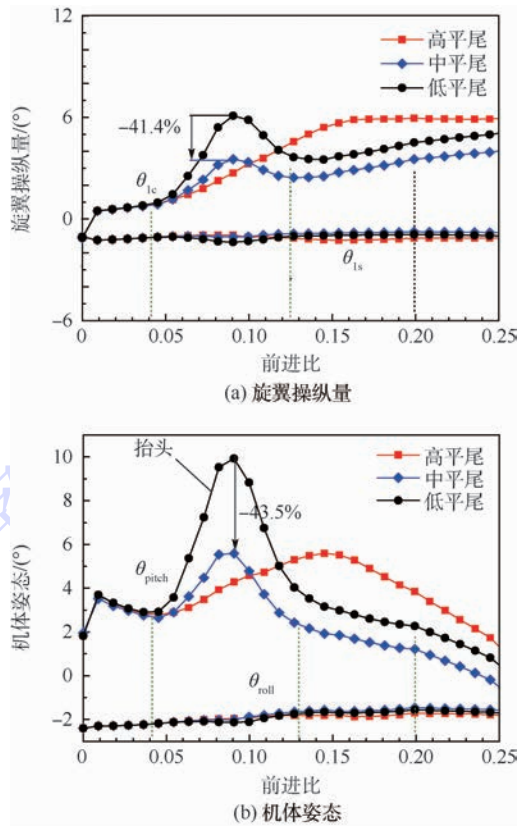


图 7 低、中、高平尾的旋翼操纵量和机体姿态

Fig. 7 Rotor control and fuselage attitude for low-, middle- and high-set empennage

3.2 平尾纵向位置

保持高度不变,将平尾纵向位置分别调整为 $0.9R$、$1.18R$、$1.3R$,得到前、中、后置平尾,对应的旋翼操纵量和机体俯仰角变化如图 8 所示。

前、中、后置平尾的旋翼操纵量和机体俯仰角随前飞速度的变化特性基本相似。过渡阶段到低速阶段,旋翼尾迹均靠近平尾表面,形成较明显的

北航学报 赠阅

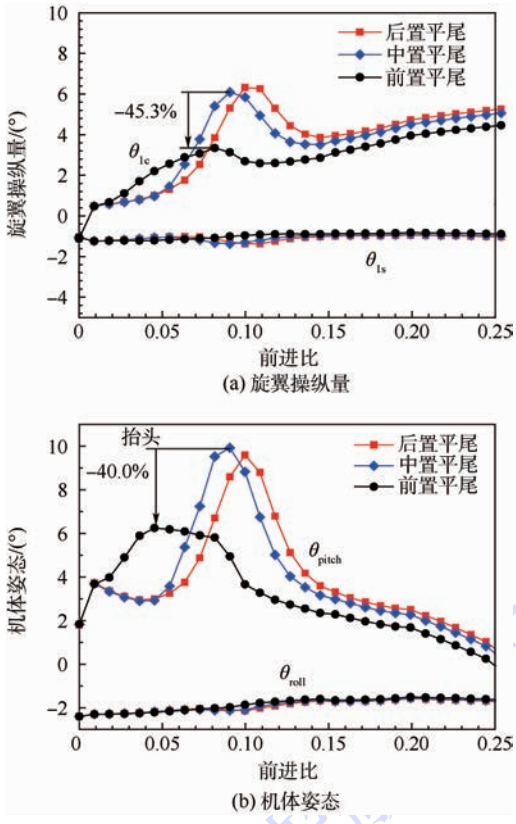


图 8　前、中、后置平尾的旋翼操纵量和机体姿态

Fig. 8　Rotor control and fuselage attitude for forward-, middle- and backward-set empennage

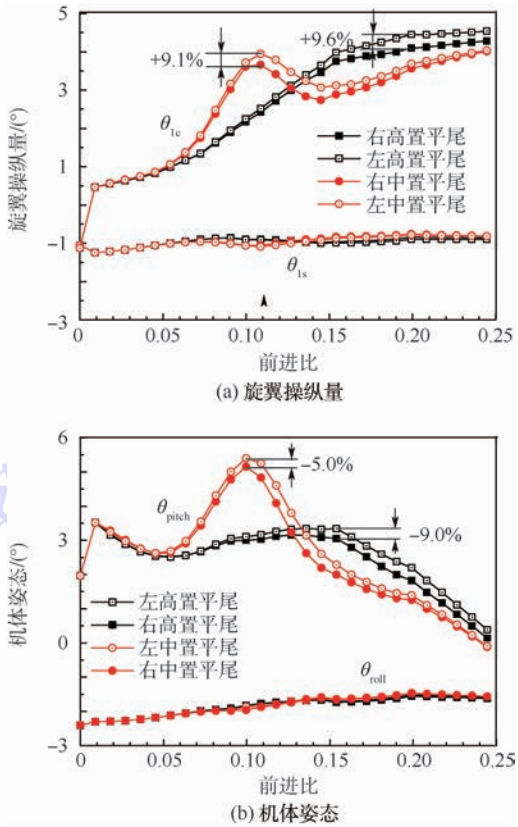


图 9　单边平尾的旋翼操纵量和机体姿态

Fig. 9　Rotor control and fuselage attitude of half-set empennage

旋翼/平尾干扰，平尾气动载荷突增较大，因此 3 种平尾构型的旋翼纵向操纵量和机体俯仰角均出现突增。但旋翼/前置平尾干扰对应的前飞速度较小，气动载荷较小，在平尾非线性气动载荷和力臂均减小的双重作用下，前置平尾的旋翼纵向操纵量和机体俯仰角突增量较小，约为中置平尾的 54.7% 和 60.0%，机体俯仰角得到抑制。相比中置平尾，后置平尾的旋翼纵向操纵量和机体俯仰角突增峰值对应的前进比后移，推迟机体抬头的出现，但突增减小量较小，机体俯仰角并未得到有效抑制。

3.3　左右单边平尾

将平尾展长减半，得到左、右单边平尾。右旋直升机采用左、右单边平尾的旋翼操纵量和机体姿态角随前飞速度变化如图 9 所示。

单边平尾的旋翼纵向操纵量和机体俯仰角随前飞速度的变化特性与对称平尾相似。过渡阶段，旋翼尾迹远离平尾下表面，旋翼/平尾干扰较弱，左、右单边平尾向下载荷基本相同，因此左、右单边平尾旋翼纵向操纵量和机体俯仰角基本相同。低速阶段，受到后行桨距角增加，尾迹涡量强度增加，旋翼尾迹/左平尾干扰较强，左平尾向下载荷大于右平尾，因此左平尾旋翼纵向操纵量和机体俯仰角大于右平尾。中高速阶段，为保证旋翼横向力矩平衡，后行边桨距增加，前行边桨距减小，后行边桨尖涡强度更大，旋翼尾迹/左平尾干扰较强，高置左单边平尾向下载荷大于右单边平尾，因此左单边平尾旋翼纵向操纵量和机体俯仰角大于右单边平尾，右平尾有利于抑制旋翼/平尾气动干扰产生的旋翼纵向操纵量和机体俯仰角突增。

4　结　论

1）耦合旋翼/平尾气动干扰模型和 GA/LM 混合优化全机配平方法，建立旋翼/平尾气动干扰作用下的全机配平特性分析方法，相比于由诱导速度考虑旋翼/平尾干扰的常规全机配平方法，本文方法更准确体现旋翼/平尾气动干扰对直升机全机配平特性的影响，更准确分析直升机低速纵向操纵量和机体俯仰角突增特性。

2）直升机低速阶段的纵向操纵量和机体俯仰角突增主要由旋翼/平尾气动干扰产生的平尾载荷变化所致，旋翼/平尾气动干扰对旋翼纵向操纵量和机体俯仰姿态特性影响显著。

3）受旋翼尾迹较强的干扰，低平尾直升机低速阶段纵向操纵量和机体抬头显著，高平尾直升

机则较弱,但对应速度转移至中等速度。

4)相比于后置平尾,前置平尾直升机低速纵向操纵量和机体俯仰角显著减弱,但速度区间增加。

5)受旋翼后行桨叶桨尖涡强度较大的影响,右旋直升机右单边平尾有利于减小纵向操纵量和机体俯仰角。

参考文献(References)

[1] LEISHMAN J G. Principles of helicopter aerodynamics[M]. 2nd ed. New York:Cambridge University Press,2006:658-682.

[2] LEONI R D. Black hawk:The story of a world class helicopter[M]. Reston:AIAA,2007:88-102.

[3] PROUTY R W,AMER K B. The YAH-64 empennage and tail rotor-a technical history[C]//38th American Helicopter Society Annual Forum. Fairfax,VA:AHS,1982:247-261.

[4] MAIN B J,MUSSI F. EH101-development status report[C]// Proceedings of the 16th European Rotorcraft Forum. Glasgow:ERF,1990:III.2.1.1-12.

[5] EGLIN P. Aerodynamic design of the NH90 helicopter stabilizer[C]// Proceedings of the 23rd European Rotorcraft Forum. Glasgow:ERF,1997:68.1-68.10.

[6] 朱宇,张宏林,刘莉. 直升机低速飞行特性试验分析[J]. 航空工程进展,2012,3(4):433-437.
ZHU Y,ZHANG H L,LIU L. Flight test research on helicopter low speed flight characteristics[J]. Advances in Aeronautical Science and Engineering,2012,3(4):433-437(in Chinese).

[7] TALBOT P D,TINLING B E,DECKER W A,et al. A mathematical model of a single main rotor helicopter for piloted simulation:NASA TM-84281[R]. Moffett Field:NASA,1982.

[8] HE C,LEWIS W D. A parametric study of real time mathematical modeling incorporating dynamic wake and elastic blades[C]//Proceedings of the 48th Annual Forum of the American Helicopter Society. Fairfax,VA:AHS,1992:1181-1196.

[9] 陈仁良. 直升机飞行动力学数学建模及机动性研究[D]. 南京:南京航空航天大学,1998:25-56.
CHEN R L. Helicopter flight dynamics mathematical modeling and maneuvering research[D]. Nanjing:Nanjing University of Aeronautics and Astronautics,1998:25-56(in Chinese).

[10] 曹义华. 直升机机动飞行旋翼的气动力模拟[J]. 航空学报,1999,20(1):39-42.
CAO Y H. Modeling the aerodynamic force of a maneuvering helicopter rotor[J]. Acta Aeronautica et Astronautica Sinica,1999,20(1):39-42(in Chinese).

[11] THEODORE C R. Helicopter flight dynamics simulation with refined aerodynamic modeling[D]. College Park, Maryland:University of Maryland,2000:35-105.

[12] SPOLDI S,RUCKEL P. High fidelity helicopter simulation using free wake,lifting line tail,and blade element tail rotor models[C]// Proceedings of the 59th Annual Forum of the American Helicopter Society. Fairfax,VA:AHS,2003:92-99.

[13] 李攀. 旋翼非定常自由尾迹及高置信度直升机飞行力学建模研究[D]. 南京:南京航空航天大学,2010:14-58.
LI P. Rotor unsteady free-vortex wake model and investigation on high-fidelity modeling of helicopter flight dynamics[D]. Nanjing:Nanjing University of Aeronautics and Astronautics,2010:14-58(in Chinese).

[14] GANGWANI S T. Calculation of rotor wake induced empennage airloads[J]. Journal of the American Helicopter Society,1983,28(2):37-46.

[15] CURTISS H C,QUACKENBUSH T R. The influence of the rotor wake on rotorcraft stability and control[C]//15th European Rotorcraft Forum. Glasgow:ERF,1989:70.1-70.24.

[16] FLETCHER T M,BROWN R E. Main rotor-empennage interaction and its effects on helicopter flight dynamics[C]// Proceedings of the American Helicopter Society 63rd Annual Forum. Fairfax,VA:AHS,2007:1-11.

[17] WANG B,ZHAO Q J,XU G H,et al. Numerical analysis on noise of rotor with unconventional blade tips based on CFD/Kirchhoff method[J]. Chinese Journal of Aeronautics,2013,26(3):572-582.

[18] KOMERATH N M,SMITH M J,TUNG C. A review of rotor wake physics and modeling[J]. Journal of the American Helicopter Society,2011,56(2):022006-1-21.

[19] HE C J,ZHAO J G. Modeling rotor wake dynamics with viscous vortex particle method[J]. AIAA Journal,2009,47(4):902-915.

[20] 魏鹏,史勇杰,徐国华,等. 基于粘性涡模型的旋翼流场数值方法研究[J]. 航空学报,2011,33(5):771-780.
WEI P,SHI Y J,XU G H,et al. Numerical method for simulating rotor flow field based upon viscous vortex model[J]. Acta Aeronautica et Astronautica Sinica,2011,33(5):771-780(in Chinese).

[21] TAN J F,WANG H W. Simulating unsteady aerodynamics of helicopter rotor with panel/viscous vortex particle method[J]. Aerospace Science and Technology,2013,30(1):255-268.

[22] 谭剑锋,王浩文,吴超,等. 基于非定常面元/黏性涡粒子混合法的旋翼/平尾非定常气动干扰研究[J]. 航空学报,2014,35(3):643-656.
TAN J F,WANG H W,WU C,et al. Rotor/empennage unsteady aerodynamic interaction with unsteady panel/viscous vortex particle hybrid method[J]. Acta Aeronautica et Astronautica Sinica,2014,35(3):643-656(in Chinese).

[23] 谭剑锋. 直升机旋翼对尾桨非定常气动载荷的影响[J]. 航空学报,2015,36(10):3228-3240.
TAN J F. Influence of helicopter rotor on tail rotor unsteady aerodynamic loads[J]. Acta Aeronauica et Astronautica Sinica,2015,36(10):3228-3240(in Chinese).

[24] 吴超,谭剑锋,王浩文,等. 基于GA/LM混合优化的直升机全机配平算法[J]. 飞行力学,2014,32(1):5-19.
WU C,TAN J F,WANG H W,et al. Optimal trim for helicopter based on GA and LM hybrid algorithm[J]. Flight Dynamics,2014,32(1):5-19(in Chinese).

[25] CROSS J,BRILLA J,KUFELD R,et al. The modern rotor aerodynamic limits survey:A report and data survey:NASA TM-4446[R]. Moffett Field:NASA,1993.

[26] 倪先平. 直升机手册[M]. 北京:航空工业出版社,2003:65-68.

北航学报 赠阅

NI X P. Manual of helicopter[M]. Beijing: Aviation Industry Press, 2003: 65-68(in Chinese).

作者简介:

谭剑锋　男,博士,讲师。主要研究方向:旋翼空气动力学与结构动力学、风力机空气动力学。

Tel.: 025-58139357

E-mail: Jianfengtan@njtech.edu.cn

吴超　男,博士研究生。主要研究方向:直升机飞行力学、直升机仿真与控制。

E-mail: wuchao_nuaa@163.com

王浩文　男,博士,教授,博士生导师。主要研究方向:旋翼动力学、结构强度及振动载荷分析。

Tel.: 010-62792661

E-mail: bobwang@mail.tsinghua.edu.cn

林长亮　男,博士,工程师。主要研究方向:旋翼结构动力学。

E-mail: lclnuaa@nuaa.edu.cn

Analysis on influence of rotor/empennage aerodynamic interaction on helicopter trim characteristic

TAN Jianfeng[1,*], WU Chao[2], WANG Haowen[3], LIN Changliang[4]

(1. School of Mechanical and Power Engineering, Nanjing Tech University, Nanjing 211816, China;
2. College of Aerospace Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;
3. School of Aerospace, Tsinghua University, Beijing 100084, China;
4. Institute of Aircraft Design, AVIC Harbin Aircraft Industry Group Corporation, Harbin 150066, China)

Abstract: Unsteady aerodynamic interaction of rotor and empennage is the core content of helicopter trim model and flight dynamic model. The rotor/empennage aerodynamic interaction effect was considered indirectly through rotor inflow or induced velocity of rotor wake in classical trim method. However, the nonlinear airload induced by the unsteady aerodynamic interaction was not fully taken into account, and it is then difficult to reflect the influence of the rotor/empennage aerodynamic interaction on the trim characteristic. Therefore, based on unsteady panel/viscous vortex particle method, a rotor/empennage unsteady aerodynamic interactional model is established by adding an unsteady pressure term into empennage panel which is induced by unsteady interaction of rotor and empennage wake, and the nonlinear airload on empennage is simulated. The helicopter trim characteristic analysis method under rotor/empennage aerodynamic interaction is then established by coupling with helicopter trim method based on GA/LM hybrid optimization. The control of rotor and attitude of fuselage of UH-60A are simulated and compared with the flight test data to validate the present method. Compared with the result of the classic helicopter trim method in which rotor/empennage aerodynamic interaction is indirectly accounted through induced velocity, it is shown that the increase of longitudinal control and fuselage attitude is not simulated by the classic helicopter trim method; however, the phenomenon is captured by the present method, and is consistent with the tested characteristics of OH-6A, EH-101, etc. The influence of empennage's type on helicopter trim characteristic is then analyzed, and it is shown that the increase of rotor longitudinal control is great with low-set empennage, forward-set empennage in low-speed forward flight, and high-set empennage in high-speed forward flight. Start-set empennage for right revel rotor is useful to decrease the increasement of longitudinal control and fuselage attitude in low-speed forward.

Key words: empennage aerodynamics; viscous vortex particle method; helicopter trim; rotor control; fuselage attitude

Received: 2015-11-23; **Accepted**: 2016-01-27; **Published online**: 2016-04-13 15:26

URL: www.cnki.net/kcms/detail/11.2625.V.20160413.1526.004.html

Foundation items: National Natural Science Foundation of China (11502105); Natural Science Foundation of Jiangsu Province (BK20161537); the Project of Natural Science Research of Higher Education Institutions of Jiangsu Province (15KJB130004)

* **Corresponding author**. Tel.: 025-58139357　E-mail: Jianfengtan@njtech.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0757

基于场景控制特征的安全性需求分析方法

朱丹江，姚淑珍*，谭火彬

（北京航空航天大学 计算机学院，北京 100083）

摘 要：安全性需求是系统安全性保证的关键。随着系统复杂度和耦合度的剧增，安全性需求的分析提取日益困难。通过对系统需求场景的控制结构和过程分析建模，提出描述控制过程中系统变量间关系的变量影响图模型，进一步给出了安全性需求分析方法。通过该方法，使用变量影响图等对控制过程进行分析，生成基于系统理论事故模型和过程（STAMP）的危险性控制活动，并以此获得系统安全性需求。经实验验证，所提出的安全性需求分析方法在正确性和一致性方面具有较好的效果。

关 键 词：安全性需求；需求分析；控制分析；场景分析；控制建模
中图分类号：TP311
文献标识码：A **文章编号**：1001-5965(2016)11-2358-13

随着软件密集型系统和信息物理系统的兴起，系统的规模剧增，系统组件间的交互、人机交互和软硬件交互越来越复杂，从而导致系统事故频发[1]。其中几乎所有涉及软件的严重事故都可以追溯到需求缺陷[2]，安全性需求在保证系统安全性中的作用日益关键。

在需求工程领域，通常将安全性需求作为非功能性需求/目标[3-4]或质量因素[5-6]进行处理，通过对安全性目标的精化和可操作化[7]、对安全性质量因素的分解和分类[5,8]以及使用安全性需求模式[9]，由安全工程师人工分析得到安全性需求。但这种对安全性需求和功能性需求分别独立定义的方法，难以得到较完备的安全性需求，导致系统安全性问题没有被很好的解决[10]。并且该方法依赖安全工程师的人工分析和定义，费时耗力，带有一定的盲目性和倾向性。因此，结合系统功能性需求自动生成安全性需求的相关研究很有必要。

安全性需求生成的一种重要方法是安全性分析[11]，通过对系统功能模型进行安全性分析，得到安全性需求，并据此完善系统功能模型，从而得到较完备的系统功能模型和安全性需求。目前，一般用基于链式/树式事故因果模型的 FTA/SFTA、SFMEA 等安全性分析方法，对系统场景[12-13]或系统结构[14-15]进行分析，得到系统失效模式并生成相应的安全性需求。然而，SFTA 等安全性分析方法主要针对简单的低耦合系统，不适用于软件密集型系统或信息物理系统等高耦合复杂系统的安全性分析[11]。并且其所用的链式/树式事故因果模型将事故的原因局限于系统组件失效，因此不能生成较完备的安全性需求。

为了解决传统链式/树式事故因果模型及安全性分析方法的局限性，Leveson[2]提出了基于系统理论事故模型和过程（STAMP），将系统看作层次化的控制结构，认为事故是由系统间的控制过程没有成功施加安全性约束导致的，并在此基础上提出了系统理论过程分析（System-Theoretic Process Analysis，STPA）方法[2]。该方法分 2 步：①确定导致系统进入不安全状态的危险性控制活

收稿日期：2015-11-17；**录用日期**：2015-12-18；**网络出版时间**：2016-01-20 17:07
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160120.1707.007.html
基金项目：航空科学基金（2013ZC51023）
* **通讯作者**：Tel.：010-82339434 E-mail：szyao@buaa.edu.cn

引用格式：朱丹江，姚淑珍，谭火彬. 基于场景控制特征的安全性需求分析方法[J]. 北京航空航天大学学报，2016，42(11)：2358-2370. ZHU D J，YAO S Z，TAN H B. Safety requirements analysis method based on control characteristics of scenarios[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2358-2370 (in Chinese).

动(Hazardous Control Action,HCA);②在控制结构中确定 HCA 的原因。由于 STPA 将系统的不完备控制作为事故原因,涵盖了系统组件失效、组件间交互不一致等原因,因此可用于生成更加完备的安全性需求。但在使用 STPA 生成安全性需求时面临着以下挑战:①STPA 的分析对象是系统设计阶段的系统控制结构,不能直接用于对常用的系统功能性需求模型进行分析;②STPA 只给出了 HCA 的分类,分析过程依赖于人工,缺乏详细的可用于自动化分析的方法。

本文在 HCA 的结构模型[16]基础上,提出一种安全性需求分析方法 Auto-STPA。该方法在系统功能场景及其控制特征描述的基础上,进行系统控制结构和过程分析建模,根据控制过程中系统变量(System Variables,SV)对控制的影响构建变量影响图,然后在此基础上自动生成 HCA 并转化为相应的安全性需求。

1　Auto-STPA 方法总体架构

为了结合系统功能性需求得到较完备的安全性需求,并且借助于自动化方法保证安全性需求的生成质量和效率,在提高系统安全性的同时降低成本,本文在 STPA 的 HCA 分类[2]及结构模型[16]基础上,针对系统功能性需求场景,提出了一种安全性需求分析方法 Auto-STPA,如图 1 所示,该方法包括场景建模、控制结构分析、控制过程分析,以及在此基础上的安全性需求分析。

图 1 中所描述的 Auto-STPA 方法的主体处理步骤如下:

1) 场景建模。为了使用 STAMP 因果模型对场景进行安全性分析,本文结合控制特征的要求对时序图进行扩展,并进行规格化表示,从而用于场景建模,获得用于表示场景的控制特征时序图。

2) 控制结构分析。由于控制活动来源于系统控制结构,本文从场景控制特征时序图所描述的交互中,分析得到系统的控制过程,然后根据这些控制过程间的关系,建立系统的控制结构。

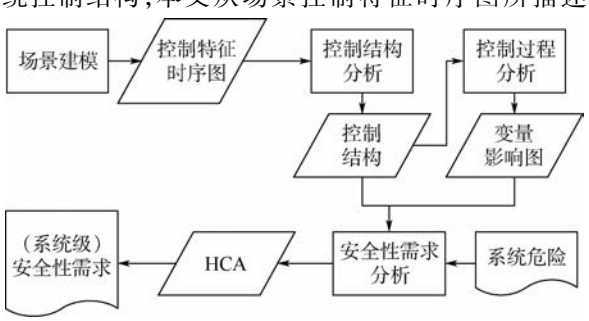


图 1　Auto-STPA 方法总体架构

Fig. 1　Framework of Auto-STPA method

3) 控制过程分析。因为 HCA 是在系统控制过程中分析得到的,所以需要对系统控制过程进行分析建模。该步骤对场景控制特征时序图中的消息等进行系统变量提取,并根据其在控制过程的作用分析其关系,用变量影响图建模。

4) 安全性需求分析。根据控制结构和过程分析模型,自动构建 HCA 树(HCAT)进行安全性分析,得到 HCA,并将其转化为安全性需求。在分析过程中,进行错误和矛盾自动检测,保证安全性需求的正确性和一致性。

2　需求场景建模与分析

对需求场景进行控制结构和过程分析是进行安全性需求分析的基础。

2.1　场景及其控制特征描述

场景一般包含在系统用例中,目前常用的场景建模方法有自然语言描述、规格化的自然语言描述和活动图等。控制作为一种系统不同层次间的交互,难以从这些场景模型中直接分析出来。而时序图由于能够表示场景中的交互,可以较容易地从中得到控制结构。

为了从场景的控制特征时序图模型中分析得到控制结构,时序图应能对控制的特征进行描述。根据进行控制所需的 4 个基本条件[2,17]可知,有效的控制过程(见图 2)具有以下特征:①控制路径和反馈路径形成一个控制闭环;②控制活动(Control Action,CA)是不同对象间的交互,且该交互的结果应达到一定的控制目的;③通过过程模型中的系统变量,对被控制过程进行控制并反馈控制结果。标准的 UML 时序图对此缺乏相应的描述机制。例如,时序图中的消息缺乏对控制活动的明确标识;由于异步消息没有返回消息,时序图的交互并没有形成显式的控制闭环;时序图交互的消息只有返回值,不能完全表示控制反馈结果等。

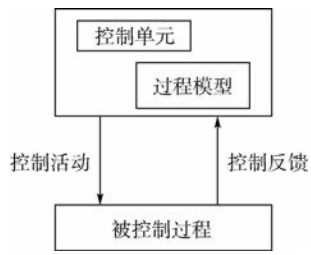


图 2　基本控制过程示例

Fig. 2　Example of basic control process

北航学报 赠阅

针对上述问题,本文对标准的 UML 时序图进行了扩展,使其能够满足对场景及其控制特征描述的要求。

1)在时序图的消息上引入构造型 << control action >> 标识出控制单元发出的控制活动。如果不同对象间的某个交互,其后置条件包含于场景所在用例的后置条件,则该交互的发起消息为控制活动。每个用例的"触发事件"通常对应一个控制活动。

2)为了描述控制结果并识别控制活动,对消息添加可选的[#control results]$_{opt}$:{SV_i = value$_i$}部分,value$_i$ 表示 SV_i 的结果。对于 << control action >> 消息,[#control results]表示预期的控制结果,其响应消息的[#control results]表示实际控制结果。

3)每条 << control action >> 消息必须有控制反馈消息与其对应,用于反馈控制的结果。消息的处理过程用以消息接收为起始,以发送响应消息为结束的"执行说明"表示,从而能够对控制过程的控制闭环特征进行描述。在 << control action >> 消息处理的执行说明中,最后发送的响应消息通常为其控制反馈消息,可以用构造型 << feedback >> 对其进行标识。

以电梯系统 Request Elevator 用例的主要成功场景[13]为例,该场景触发事件为用户按动电梯按钮,电梯接收到相应命令后提取目的楼层信息,并计算电梯运行方向和运行距离,生成电梯调动命令。在电梯运行之前,系统先发送命令确保电梯门关闭。在电梯停靠在目的楼层后,发送命令打开电梯门。根据上述描述机制,该场景及其控制特征描述如图 3 所示。其中,press_elevator_button、CloseDoor 等消息作为控制活动,用 << control action >> 进行标识,并且与其反馈消息组成控制闭环。由于消息 8、11、12 反馈的[#control results]包含于系统整体预期控制结果中,即消息 1 的[#control results],因此该实例包含3个直接

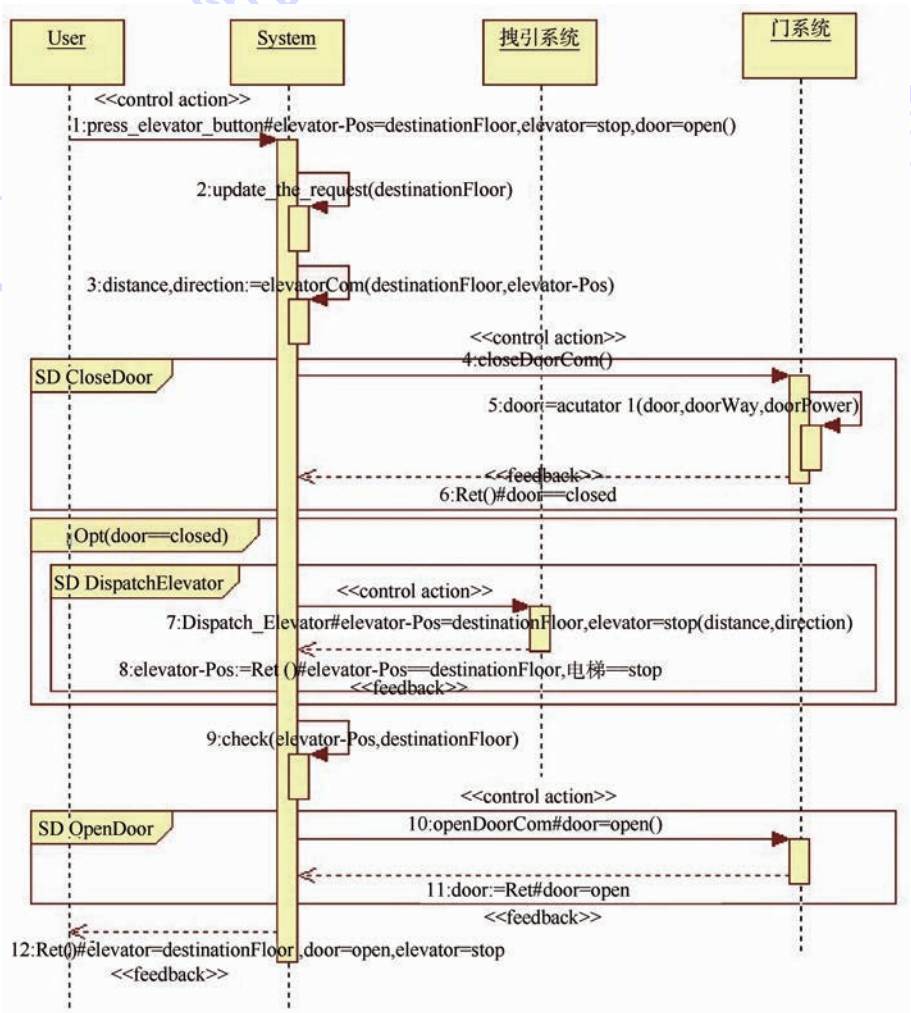


图 3 电梯系统控制特征时序图

Fig. 3 Control characteristic sequence diagram of elevator system

控制过程:User 控制 System、System 控制门系统和拽引系统;消息 1 的[#control results]不包含消息 6 的[#control results],因此 System 控制门系统的过程为间接控制过程。

为了便于自动化分析,本文将上述针对控制特征扩展的时序图定义如下。

定义 1　控制特征时序图

控制特征时序图 C-SD 可以表示为 5 元组 $<O, M, E, \rightarrow, ES>$。

1) $O=\{O_1, O_2, \cdots, O_m\}$ 为对象集合,$O_i(0<i<m)$ 为时序图中第 i 个对象。

2) $M=\{m_i \mid m_i \subseteq [\text{attribute}=]_{opt}[<\text{type}>.]_{opt}\text{name}[\text{argument}_{list}]_{opt}[:\text{return-value}]_{opt}[\#\text{control results}]_{opt}\}$ 为消息集合。<type>表示消息类型:<CA>说明是 CA 消息;<CR>说明是控制反馈消息;<COMM>说明消息是普通的通信消息,一般可省略。

3) $E=M\times\{s,r\}$ 为事件集合。s 表示消息的发送,r 表示消息的接收。对于消息 m_i,发送事件用 $<m_i,s>$ 表示,接收事件用 $<m_i,r>$ 表示。

4) $\rightarrow=\{<m_1,m_2,\cdots m_i,\cdots,m_n> \mid m_i \in M \cup \text{COMB}\}$ 为消息的偏序关系,表示时序图的消息轨迹。COMB 为表示交互片段的特殊消息集合,每个交互片段由 2 个特殊消息组成:①片段起始消息,包含交互片段操作符构成的消息名和监护条件构成的参数列表;②片段结束消息,一般以"END_交互片段操作符"命名。

5) $ES=O\times Es\times Ee$ 为执行说明集合,表示交互的执行过程。$Es\subseteq E$ 和 $Ee\subseteq E$ 分别为起始事件和终止事件集合。

根据定义 1,图 3 时序图形式化表示为

$O=\{\text{User},\text{System},\text{DoorSys},\text{MoveSys}\}$

$M=\{1<\text{CA}>.\text{press_elevator_button}()\#\{\text{elevator-Pos}=\text{destinationFloor},\text{elevator}=\text{stop},\text{door}=\text{open}\},2\text{update_the_reguest}(\text{destinationFloor}),3,4,5,6,7,8,9,10,11,12\}$

$\rightarrow=<1,2,3,\text{REF}(\text{SD-CloseDoor})(4,5,6)\text{ENDREF},\text{Opt}(\text{door}==\text{closed})(\text{REF}(\text{SD-Dispatch Elevator})(7,8),\text{ENDREF})\text{ENDOpt},9,\text{REF}(10,11)\text{ENDREF},12>$

$E=\{<m_i,s>,<m_i,r> \mid m_i=1,2,3,4,5,6,7,8,9,10,11,12\}$

$ES=\{<\text{User},<1,s>,<12,r>>,<\text{System},<2,s>,<2,r>>,<\text{System},<3,s>,<3,r>>,<\text{DoorSys},<4,s>,<6,r>>,<\text{DoorSys},<5,s>,<5,r>>,<\text{MoveSys},<7,s>,<8,r>>,<\text{System},<9,s>,<9,r>>,<\text{DoorSys},<10,s>,<11,r>>\}$

式中:M 中的消息除部分给出具体形式外,其他的用消息序号表示。根据消息的格式可以从中提取出系统变量以及控制结果。ES 中的每个元素表示一个过程,并且满足控制过程的闭环结构特征,便于后文根据 << control action >> 消息及其反馈消息分析控制过程。另外,可以发现 ES 中的执行说明存在嵌套情况,形成了过程和子过程的关系,这种关系在后文被用来进行控制结构分析。

与传统的自然语言等描述方法相比,本文控制特征时序图对场景的描述更加直观,其形式化定义便于进行自动化分析;与标准的 UML 时序图相比,本文控制特征时序图在交互描述中突出了场景中的控制特征,并且所用的消息结构可以更好地进行系统变量和控制结果等信息的提取,便于后文控制结构和过程的分析。

2.2　控制结构和过程分析

在系统控制理论中,系统被看作是层次化的控制结构,每层系统对其下层系统的行为进行约束,即高层通过约束对低层的行为进行控制[2]。这种系统控制结构层次间的控制及其反馈过程形成了如图 4 所示的控制环,使得系统保持动态平衡状态。控制环作为基本控制结构和过程的描述,定义如下。

定义 2　控制环

控制环 CL 表示为 5 元组 $<\text{controller}, CP, CR, RCA, \{<CA, CEDP>\}>$。

1) controller 表示 CL 所描述的控制过程中的控制单元。

2) $CP\in ES$ 表示 CL 所描述的"控制过程"。ES:$<obj,es,ee>$ 为 CP 对应的执行说明。

3) CR:{SV = value}表示 CP 的"控制结果"。

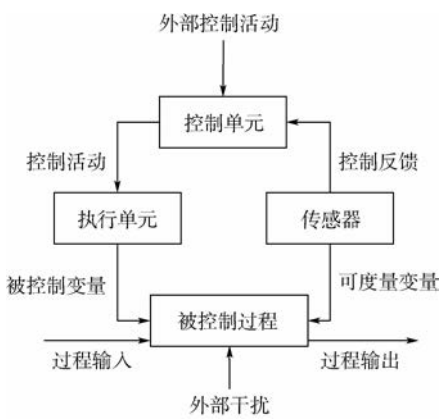


图 4　基本控制环

Fig. 4　Basic control loop

北航学报 赠阅

value 为系统变量 SV 对应的值。

4) RCA 表示 controller 所接收到的控制活动,即该 CP 所处理的控制活动。对于顶层 CL,该字段为 Null。

5) < CA, CEDP > 表示 CP 中的“被控制过程”。CA 为 controller 发出的控制活动,CEDP ∈ ES 为 CA 对应的“子控制过程”。

由定义 2 可知,在场景的控制特征时序图中,ES 集合的每个执行说明 es_k,满足 $es_k \in ES$ 且 $\bigcup_{k=1}^{n} es_k = ES$, $n = card(ES)$。假设 es_k 的起始事件 Es 对应消息 m_i,其终止事件 Ee 对应消息 m_j,若 m_i 为 CA 消息,m_j 为其控制反馈消息,则 es_k 对应一个控制环 cl,且 cl. CP = es_k, cl. controller = es_k. obj, cl. CR 为 m_j 的[#control results], cl. RCA = m_i。若 es_k 存在子执行说明 $es' \in ES$,使得 es'对应消息序列是 $< m_i, \cdots, m_j >$ 的子序列,且 es'为控制环,则 cl. < CA, CEDP > 为 < es'. obj, es' >,从而从场景的控制特征时序图中自动提取出控制环。

在一个控制环 CL 中,如果存在 < CA, CEDP >,则说明 controller 在对 RCA 的处理过程 CP 中生成了新的控制过程,即发生控制过程嵌套,并且 CEDP 是 CP 的子过程,CEDP 对应的 CL_{CEDP} 是其子控制环。

对于存在嵌套的控制过程,其嵌套的子控制过程存在 3 种作用:①控制过程将要进行的控制传递到子控制过程处理;②控制过程分解为多个较简单的控制在子控制过程中处理;③子控制过程通过生成父控制过程执行所需的某些条件,实现间接控制。其中,①和②是完整的控制过程,其控制结果(部分)达到整体控制目的;③不是完整的控制过程,其控制结果是父控制过程中的某些条件,没有直接(部分)达到父过程的控制目的。本文现阶段只考虑完整的控制过程,即控制环 CL 满足条件:$\forall CL_i$,如果 $\exists CL_j$ 使得 CL_j. CP $\subseteq$ CL_i. CEDP,即 CL_j 嵌套于 CL_i,则 CL_j. CR $\subseteq$ CL_i. CR。

根据上述控制环的定义及条件,图 3 所示场景中存在 4 个完整的控制环。

1) ControlLoop1

(User, <<1, s >, <12, r >>, {elevator-St == stop, elevator-Pos == destinationFloor, door == open}, NULL, { < System, <1, r >, <12, s >> })

2) ControlLoop2

(System, <<1, r >, <12, s >>, {elevator-St == stop, elevator-Pos == destinationFloor, door == open}, press_elevator_button( ), { < MoveSys, <7, r >, <8, s >>, < DoorSys, <10, r >, <11, s >> })

3) ControlLoop3

(MoveSys, <<7, r >, <8, s >>, {elevator-St == stop, elevator-Pos == destinationFloor}, DispatchElevator(direction, distance), NULL)

4) ControlLoop4

(DoorSys, <<10, r >, <11, s >>, {door == open }, openDoorCom( ), NULL)

控制结构作为一种通过系统/组件间的控制/反馈关系进行组织的系统结构,期望通过系统间的控制施加安全性约束,从而达到一定的控制目的。控制结构中的分层是通过层次间的控制过程来确定的[18],过程之间只具有嵌套关系或不相交关系[19],因此,系统的控制结构可以由若干个描述控制过程的控制环间的嵌套或不相交关系组成,其定义如下。

定义 3 控制结构

控制结构 $CS = \{CL_{ij} \mid CL_{i+1\,j'} \subseteq CL_{ij}, 0 < i, 0 < j, 0 < j'\}$。$CL_{ij}$ 处于 CS 中第 i 层且为该层第 j 个 CL,$CL_{i+1\,j'}$ 处于 CS 中第 $i+1$ 层且为该层第 j' 个 CL,使得 $CL_{i+1\,j'}$. CP $\not\subseteq CL_{ij}$. { < CA, CEDP > }. CEDP 成立。

根据定义 3,可以得到如图 5 所示的控制结构图。从图 5 可知,$CL_{1,1}$ = ControlLoop1, $CL_{2,1}$ = ControlLoop2, $CL_{3,1}$ = ControlLoop3, $CL_{3,2}$ = ControlLoop4,且 $CL_{1,1}$的 controller 将控制传递到 $CL_{2,1}$ 处理,$CL_{2,1}$的 controller 将控制过程分解为两个子控制,分别通过 $CL_{3,1}$ 和 $CL_{3,2}$ 进行处理,达到最终的控制目的。与目前只表示系统/组件间的控制关系的控制结构定义[20]相比,本文的控制结构定义除了用控制环表示系统/组件间的控制关系外,还用控制环间的关系表示控制过程的分解和传递,从而更加便于对系统控制结构和过程进行安全性分析。

通过控制结构完整性检测,能够发现控制过程缺陷导致的危险,如缺少控制反馈等,但是由于

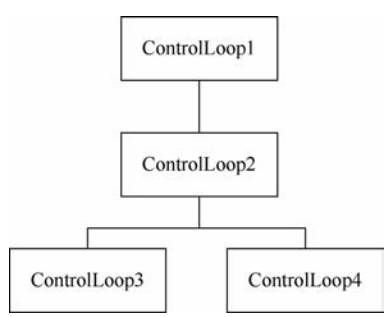


图 5 电梯系统控制结构

Fig. 5 Control structure of elevator system

控制结构不能表示控制过程中所施加的约束以及约束的作用过程,因此不能从中分析发现 STAMP 因果模型的约束没有成功施加等 6 种危险因素[2]。为了分析得到由于控制施加的安全性约束不完备,或约束施加过程缺陷而导致系统进入危险状态的 6 种 HCA,本文对系统功能性需求场景进行分析,从中提取出控制过程所涉及的系统变量及其之间的关系,并用变量影响图进行建模。通过系统变量及其之间的关系对控制活动所施加的约束进行控制结果预判,发现导致系统进入不安全状态的 HCA,并在控制施加约束的过程中,通过系统变量间的相互作用分析出没有达到控制目标的原因。另外,系统变量之间的关系还可以用来发现错误的 HCA,以及将 HCA 精化到更低级别。

在控制过程中,系统变量之间的关系是控制规则的来源[2],安全性约束往往通过对系统变量的限制来体现,并且通过系统变量间的关系进行施加,从而达到控制目的。本文根据系统变量在控制过程中的作用,将系统变量间关系分为横向关系和纵向关系,并定义如下。

定义 4　系统变量间的关系

系统变量间关系 SVR = $\{<x,y>, x$ 和 y 为 SV$\}$,表示 x 在控制过程中影响 y。SVR = $H \cup V$。

1) H = SV × K_SV, SV = $\{sv \mid sv \in \rightarrow_{CL.(CP\text{-}CEDP).m}\}$, K_SV = $\{sv \mid sv \in CL.CR\}$,表示系统变量间的横向关系,即:如果 SV 通过作用于控制活动的生成和传递过程影响 K_SV,从而达到控制目的,则称 SV 和 K_SV 的笛卡儿乘积组成的关系是横向关系。

2) V = SV′ × K_SV, SV′ = $\{sv \mid sv \in CR\}$,表示系统变量间的纵向关系,即:如果 SV′通过作用于控制活动的执行或处理过程影响 K_SV,从而达到控制目的,则称 SV′和 K_SV 的笛卡儿乘积组成的关系是纵向关系。

根据定义 4,可以用以下规则判断控制环 CL 中系统变量间的关系:

1) CL. CP 中新的系统变量与 CL. CR 中的系统变量是纵向关系。

2) CL.CEDP 之前子过程中的系统变量与 CL_{CEDP}. CR(即该被控制过程的控制结果)中的系统变量是横向关系。

3) CL. RCA 中的系统变量与 CL. CR 中的系统变量是横向关系。

4) CL. CP 中如果存在组合片段生成的特殊消息,则组合片段条件中的系统变量与其交互片段中的系统变量是横向关系。

通过上述关系定义和判断规则,可以从控制环组成的控制结构中分析得到如下定义的变量影响图,对系统变量及其之间的关系进行建模。

定义 5　变量影响图

变量影响图 VEG = < SVs, H, V >。SVs 为 CS 中系统变量的非空集合,H 为 SVR 的横向关系,V 为 SVR 的纵向关系。

图 6 所示电梯系统的变量影响图中,节点表示系统变量对应的状态机,H 边和 V 边分别描述了系统变量间的横向和纵向关系。控制活动直接作用于某些系统变量(如 direction 和 distance),利用变量影响图中系统变量间的作用关系,将约束传递到被控制对象(如 elevator-Pos),达到控制目的(如电梯对齐停靠在某楼层)。这样通过变量影响图,在一定程度上刻画了控制活动施加约束的过程。因此,变量影响图不但能够用于发现控制过程中的错误,还可以用来分析导致系统危险的控制偏差产生的原因,解决了现有的被控制系统模型 SED[21] 和层次过程模型[16] 所表示系统变量间的函数组成关系只能用于系统危险或目标分解细化的局限性。

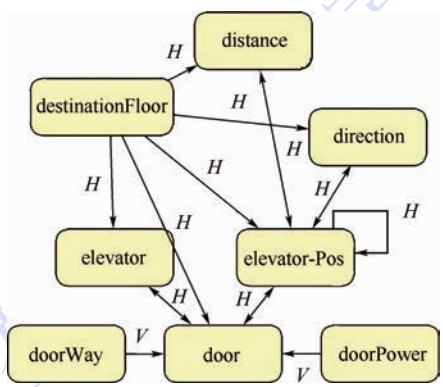


图 6　电梯系统局部变量影响图

Fig. 6　Part of variables effect graph for elevator system

另外,由于目前的控制结构和过程分析模型一般由设计人员直接给出[22],控制结构和过程分析模型与需求间的一致性得不到保证,并且使得目前 STPA 分析局限于系统设计阶段。本文从描述功能性需求场景的控制特征时序图中自动提取控制结构和过程分析模型,不但保证了控制结构和过程分析模型与需求描述的一致性,而且将安全性分析提前到需求阶段,降低了系统安全性保证的成本。

3　安全性需求分析

安全性需求分析一般以系统安全性分析为基础,本文在 STPA 方法[2] 的基础上,根据系统控制

北航学报 赠阅

结构和变量影响图,对施加约束的控制活动进行安全性分析,生成可能导致系统进入不安全状态的 HCA,并且根据其结构模型[16]将其自动转化为安全性需求。

3.1 HCAT 和时间因素分析

安全性需求一般由安全性分析得到的危险原因生成。根据 STAMP 因果模型,STPA 将导致系统危险的 HCA 分为 4 类[2],但 STPA 并没有给出 HCA 的分析方法[23]。本文安全性需求分析方法在 STPA 的 HCA 分类基础上,对系统进行安全性分析,生成 HCA 并转化为安全性需求。HCA 作为 STAMP 因果模型的事故原因,表示因违背安全性约束而导致系统危险的控制行为。Thomas[16]将 HCA 定义为 4 元组(SC,T, CA, Co),SC 为控制单元,CA 为控制单元提供的控制活动,T 为 CA 的类型:提供(provided)或不提供(not provided),Co 为控制活动的上下文。由该定义可知,HCA 主要由 CA 及其作用/应该作用的 Co 组成,其中 CA 可以用其作用的系统变量(称为关键 SV)表示,Co 也可用系统变量集合表示。因此,本文安全性需求分析方法根据控制结构和变量影响图,生成如下定义的 HCAT,对控制过程中导致系统进入不安全状态的 CA 及其上下文进行分析。

定义 6 危险性控制活动树

HCAT 是满足如下性质的树结构:

1) 树的根节点为控制活动 CA,有 2 个 T 节点作为其子节点,分别表示 CA 的不同类型 T 的取值。

2) T 节点的子节点为关键 SV 节点,表示 CA 是否作用造成关键 SV 的不同取值情况。

3) 关键 SV 节点的每个子节点为 sv 的不同取值情况,sv 为组成 Co 的系统变量,即关键 SV 节点的子树为 Co 子树,表示 CA 的上下文。

4) Co 子树中,其根节点 sv 的子树是一个由其他系统变量组成的 Co 子树。

由于危险被定义为导致事故的系统状态或条件[2],也可用系统变量集合表示,因此可以在 HCAT 中,用危险对其中路径的合法性进行检测,从而保证树中每条从根节点到叶子节点的路径都表示一个 HCA。

STPA 方法认为除控制活动提供与否导致的 2 种 HCA 外,在控制活动提供并执行的情况下,还存在如下与时间因素相关的 4 种 HCA 需要进行消除或控制[1]:

1) PTE(Provided Too Early):由于提供的太早,或提供顺序错误,而导致系统进入不安全状态的 HCA。

2) PTL(Provided Too Late):由于提供的太晚,或提供顺序错误,而导致系统进入不安全状态的 HCA。

3) STS(Stopped Too Soon):由于结束太早而导致系统进入不安全状态的 HCA。

4) ATL(Applied Too Long):由于应用时间太长而导致系统进入不安全状态的 HCA。

因为 HCA 的 4 元组定义[16]和定义 6 中缺乏明确的时间约束,所以 HCAT 无法直接生成上述与时间因素相关的 HCA。本文对 HCA 中的时间因素进行了分析,得到如表 1 所示的规则。通过这些规则,可以根据系统变量的状态变迁,用 HCAT 生成所有类型的 HCA。

HCA 可用于生成系统/组件行为的安全性需求和约束[1]。从表 2 所示的安全性需求和约束可知,安全性需求和约束一般具有“系统-对象-系统行为-行为发生上下文”的结构(系统行为中可以包含行为的作用对象,行为发生上下文中可以包含时间等约束),因此可以将安全性需求表示为与 HCA 对应的 3 元组(system, behavior, context),system = HCA. SC, behavior = HCA. T + HCA. CA, context = HCA. Co,这样本文安全性需求分析可以在自动生成 HCA 后将其转化为安全性需求。

表 1 HCA 时间因素对应规则

Table 1 Correspondence rules of HCA time factor

HCA 种类	时间因素对应规则
PTE	1)“关键 SV”变迁后的状态 +“上下文”的前一个状态; 2)“关键 SV”变迁后的状态 +“上下文”当前状态
PTL	1)“关键 SV”不发生变迁的状态 +“上下文”的当前状态; 2)“关键 SV”变迁后的状态 +“上下文”下一个状态
STS	1)“关键 SV”变迁后状态的下一个状态 +“上下文”的当前状态; 2)“关键 SV”变迁后的状态 +“上下文”当前状态
ATL	1)“关键 SV”不发生变迁的状态 +“上下文”的当前状态; 2)“关键 SV”变迁后的状态 +“上下文”下一个状态; 3)“关键 SV”不发生变迁的状态 +“上下文”下一个状态

表 2 安全性需求和约束示例

Table 2 Examples of safety requirements and constraints

编号	安全性需求和约束
1)	Door must designed to facilitate immediate reversal of direction 门系统必须能够快速进行反向动作
2)	Door must be designed and located to withstand emergency conditions 门系统必须在紧急情况下进行开关门操作
3)	Sensor must be designed to operate in the presence of X, Y, Z materials 传感器必须在 X、Y、Z 物质存在时仍能正常工作

3.2　安全性需求分析方法

在第 3.1 节分析的基础上,本文提出了根据变量影响图等自动生成安全性需求的方法。首先,根据变量影响图中变量间的横向关系构造 HCAT;然后,根据 HCAT 和表 1 所示规则生成 HCA;最后,根据安全性需求的表示和转换规则,将 HCA 转换为安全性需求。另外,在用变量影响图构建 HCAT 的过程中,通过路径检测保证 HCAT 中每条路径都能够转变为一个 HCA,并且通过错误及冲突检测和处理保证生成结果的正确性和一致性。

输入:危险 H,变量影响图 VEG,控制结构 CS。

输出:安全性需求。

1) 初始化 HCAT。$\forall \mathrm{CA} \in \mathrm{CS}, K_\mathrm{SV} \subseteq \mathrm{CA}\text{-}K_\mathrm{SV}$,树根节点 root = CA,root. lchild = T,root. rchild = T',K_SV_i. parent = T & K_SV_i. parent = T',其中,CA-K_SV 为 CA 作用的"关键 SV"集合,T 和 T'分别表示 CA 是否提供的 T 节点,K_SV_i 为根据 K_SV 取值生成的节点。

2) $\forall \mathrm{sv} \in \mathrm{VEG.SVs}$,如果 sv 与 K_SV 中系统变量具有横向关系,且 sv 不存在于 HCAT 中,根据系统变量间关系 SVR,对其每种可能取值生成节点 $\mathrm{CO_SV}_i$ 作为当前 HCAT 叶子节点的子节点。对于当前 HCAT 进行危险检测,如果 $\forall h_i(v_{i1}, v_{i2}, \cdots, v_{in}) \in H$, $\exists p_j(\mathrm{sv}_{j1}, \mathrm{sv}_{j2}, \cdots, \mathrm{sv}_{jm}) \in \mathrm{PATH}: \{v_{i1}, v_{i2}, \cdots, v_{in}\} \subset \{\mathrm{sv}_{j1}, \mathrm{sv}_{j2}, \cdots, \mathrm{sv}_{jm}\}$,其中 PATH 为 HCAT 的路径集合,说明当前路径能导致危险,标注出所导致的危险。

3) 根据 HCA 时间因素对应规则,在 T 子树上进行危险检测。之后,删除 HCAT 中没有标注危险的路径。

4) 如果 HCAT 的左右子树中具有相同 Co 路径,则说明这两条路径标注的危险相互冲突,根据危险优先级,删除优先级低的路径。

5) 如果 HCAT 中存在路径包含 CL_k.CP 中"关键 SV"及其控制路径上 SV 的取值,其中 $\mathrm{CL}_k \in \mathrm{CS}$,则说明系统目标和系统危险间存在矛盾,删除该路径。

6) 遍历 HCAT 中从根节点到叶子节点的所有路径,并删除路径上的 K_SV,生成 HCA 集合。

7) 将 HCA 转化为相应的安全性需求。

该方法中,步骤 1) 和步骤 2) 是根据定义 6 构造 HCAT,并且在步骤 2) 中对导致的危险进行了标注,另外用系统变量间的关系进行 Co 子树构造,避免了生成错误的 HCA,保证了所生成需求的正确性。步骤 3) ~ 步骤 5) 是 HCAT 剪枝过程,其中,步骤 3) 是进一步做时间因素分析后剪去安全路径;步骤 4) 和步骤 5) 是进行冲突检测和解决,从而保证了所生成需求的一致性。在生成 HCA 后,可以根据文献[16]的方法,用变量影响图中的纵向关系对 HCA 进行精化。步骤 6)、步骤 7) 是 HCA 及其对应安全性需求的生成。本文安全性需求分析方法不但在 HCAT 构造和剪枝过程中对 CA 所有的上下文情况进行了处理,保证了能够使用 VEG 和 CS 生成完备的 HCA 集合,而且根据 HCA 时间因素对应规则,生成了 STAMP 因果模型所有类型的 HCA,从而保证了最终生成的安全性需求在数量和类型上的完备性。

下面以电梯系统为例,对图 6 所示的变量影响图中所示的系统变量,做出如下假设:

1) SV:destinationFloor 由于在场景的一次执行过程中恒定不变,因此作为常量处理。

2) K_SV:elevator-Pos 的值为 right 和 wrong,分别表示其是否在 destinationFloor。

3) SV:door 的值为 open 和 closed,分别表示调度时门开和关。

4) SV:distance 具有值 aligned 和 not aligned,分别表示是否与楼层对齐。

5) SV:direction 具有值 right 和 wrong,分别表示调度方向对错。

为了表明该方法的矛盾发现和处理能力,引入 2 个危险。H1:门开时电梯进行调度;H2:电梯没有对齐时停止调度,且优先级 H1 > H2。根据上述方法,对于电梯系统的控制活动 DispatchElevator,所生成的 HCAT 如图 7(a)、图 7(b)所示。

图 7 中:用 H 标识的路径表示 HCA 路径;用"×"标识的路径是在 HCAT 构建过程中被剪枝的安全路径;方框表示存在危险冲突;标识有"×"的方框表示经冲突处理所删除的节点。所生成的安全性需求有:

1) 电梯调度活动不能在电梯门打开且电梯与楼层对齐时进行。

2) 电梯调度活动不能在电梯门打开且电梯没有与楼层对齐时进行。

3) 电梯门关闭且电梯没有对齐楼层时必须进行电梯调度活动。

4) 电梯不能在电梯门没有完全关闭时就太早进行调度。

5) 电梯不能调度太晚,以至在电梯门再次打开时进行。

北航学报 赠阅

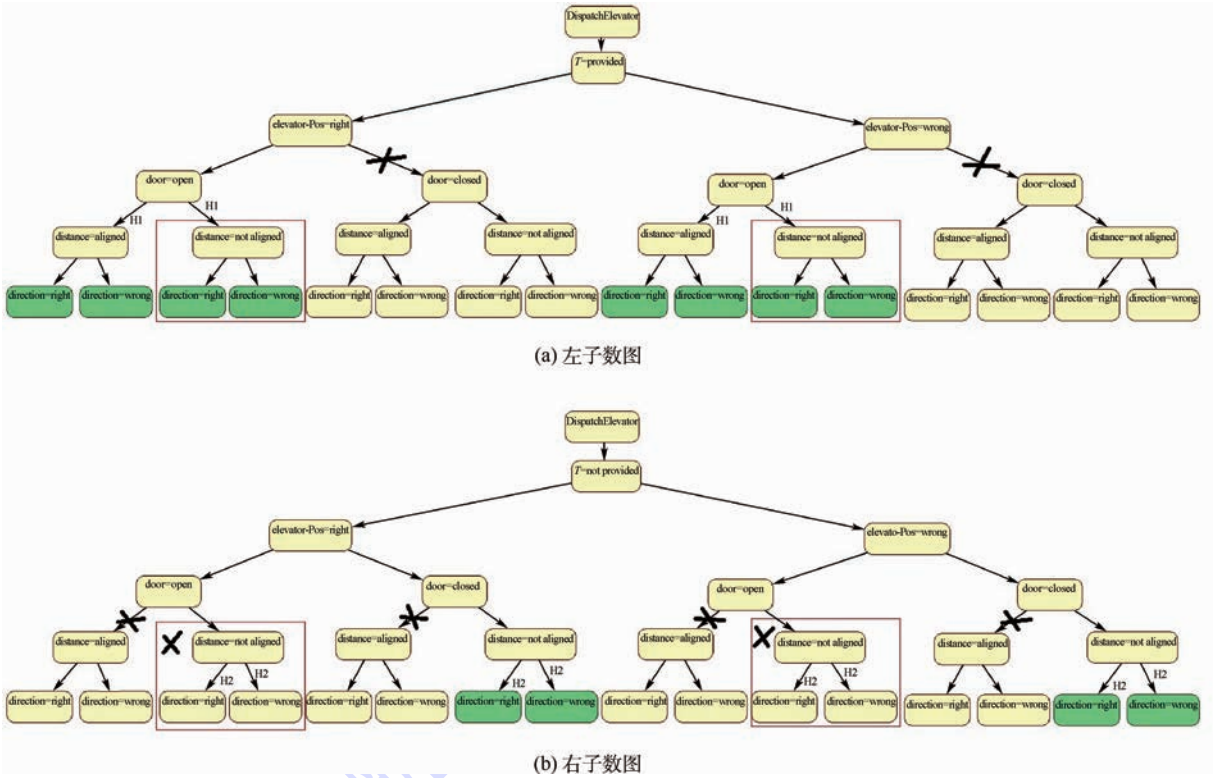


图 7 电梯系统 HCAT

Fig. 7 HCAT of elevator system

6）电梯调度活动不能结束太早，以至在电梯没有对齐楼层时就停止。

7）电梯调度活动应用时间太长，以至在电梯没有对齐楼层或电梯门再次打开时进行。

本文的安全性需求分析方法除了生成文献[16]定义的 2 种基本 HCA 外，还生成了文献[1]所述的 4 种与时间因素相关的 HCA。并且在 HCA 的生成过程中，通过危险检测和利用系统变量间横向关系进行剪枝处理，不但提高了方法效率，而且保证了所生成 HCA 和安全性需求的正确性。另外，方法中的矛盾检测处理保证了 HCA 和安全性需求的一致性。

4 实验分析

本节用人工构建的系统功能场景控制特征时序图作为输入，通过与现有的分析方法进行对比，对本文控制结构和过程分析的有效性进行评估。另外，由于本文 HCA 和安全性需求一一对应，本节通过对本文安全性需求分析方法生成的文本形式的 HCA 与文献[16]方法得到的 HCA 进行对比，对本文方法的有效性进行评估。

4.1 实验场景介绍

为了对基于场景控制特征的安全性需求分析方法进行检验，本文设计了安全性需求提取的原型工具 AutoSR 用于实验分析。

利用该原型工具，本文选择 Train Door Controller（TDC）[16]、ITP[1] 和 Nuclear Power Plant（NPP）[16] 系统实例作为实验对象，对本文方法进行评估验证。其中，TDC 是用于火车门开启和关闭的控制系统，要求 TDC 在正常或紧急情况下都能控制车门动力系统进行车门打开或关闭操作，并能够通过传感器信息采集系统获取相关信息，用于车门控制；ITP 是简化的管理飞机飞行高度并保证安全间隔的空中交通管理系统，要求飞机的 ITP 系统能够检测飞行空域内其他相关参考飞机的信息，判断能否变更飞行高度，并能够经由所在空域控制台向空中交通管理系统发送飞行高度变更请求，接收其相关指令；NPP 是简化的核电站反应堆冷却系统，用于人工或通过各种自动控制系统对冷却系统的冷却液阀门进行开关控制，并用优先级管理模块对接收到的各种控制命令进行协调，对冷却液阀门进行控制。这 3 个实例系统/组件间的控制关系如图 8 所示。图中：每个方框代表需求中涉及的系统/组件，连接线表示系统间的关系，每个控制关系对应需求中的一个表示控制功能的需求。

下面分 3 部分对本文安全性需求分析方法进行评估。首先，对原型工具生成的控制结构进行评估，分析验证本文控制结构模型的有效性；然

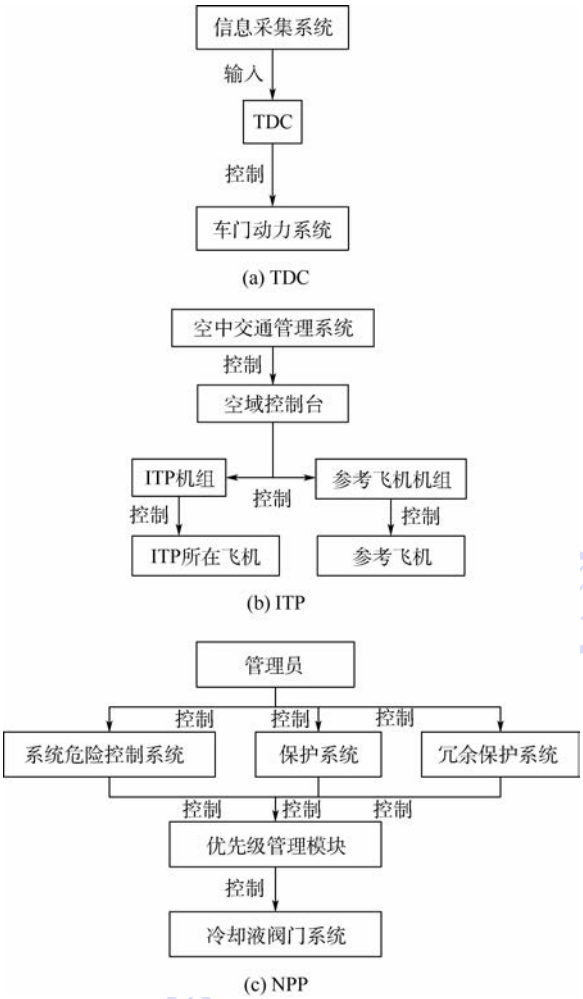


图 8　TDC、ITP 和 NPP 控制关系

Fig. 8　Control relationship of TDC, ITP and NPP

后,通过与安全分析常用的过程分析模型对比,分析验证变量影响图的有效性;最后,验证本文基于控制结构和过程的安全性需求分析方法的有效性。

4.2　控制结构有效性验证与分析

为了验证控制结构模型的有效性,本文分别对 TDC、ITP 和 NPP 进行控制特征时序图建模,用控制结构和过程分析模块提取控制结构模型,并将其与人工构建的对应(Hierarchical Control Structure, HCS)[20]模型进行对比分析。根据实验对象的需求与图 8,可以得到实验对象相关的控制信息,如表 3 所示。

假设模型满足所有控制需求,即每个控制需求都有控制环描述的控制过程与之对应。本文用 $CLM = \sum CL_{ij}$ 表示控制环的总数,评估模型描述控制关系和控制过程的能力;控制强度 CC = CLM/控制需求数量,表示平均每条控制需求对应的控制环数量,评估模型满足控制需求的程度,由于控制的传递和分解会导致一条控制需求对应多个控制过程,CC 可能大于 1;控制密度 CD = CLM/控制模型层次数,表示控制结构中平均每层的控制环数量,评估模型平均每个控制层次所含的控制信息量。所得的各项评估结果如表 4 所示。

表 3　实验对象控制信息

Table 3　Control information of subjects for experiments

实例	系统数量	被控制系统数	系统控制层次	控制需求数量
TDC	3	1	2	1
ITP	6	5	4	5
NPP	6	5	4	8

表 4　控制结构模型评估结果

Table 4　Assessment results of control structure model

实例	HCS 模型			本文模型		
	CLM_{HCS}	CC_{HCS}	CD_{HCS}	CLM	CC	CD
TDC	1	1	0.50	1	1.00	0.50
ITP	5	1	1.25	11	2.20	2.75
NPP	8	1	2.00	12	1.38	3.00

由表 4 可知,HCS 模型的 CLM 评估普遍小于本文模型,这是由于 HCS 模型只描述系统间的直接控制关系,而本文模型还包含由控制分解、传递形成的非直接控制关系,表明本文模型克服了 HCS 模型无法描述非直接控制关系的缺陷。本文模型 CC 值一般大于 HCS 模型,这是因为一条控制需求可能需要通过控制传递和分解而被多个控制过程实现,表明本文模型相对于 HCS 模型包含控制需求对应的更多控制信息。因为 HCS 形式化定义中每个控制层次只有一个控制/被控制系统,而本文模型的每个控制层次可以存在多个控制环,本文模型 CD 值不小于 HCS 模型,表明本文模型解决了 HCS 模型对 ITP 的一对多、NNP 的多对多等复杂控制结构建模的局限性。总之,对于同一系统的控制需求,本文用控制环定义的控制结构描述了系统间的直接/非直接控制关系,相对于只描述直接控制关系的传统 HCS 模型,其对控制信息的描述更加完备;此外,本文模型通过描述控制环间关系并形成层次结构,不但方便了后文对系统变量间关系的分析,还使得对复杂控制结构的建模更加简洁,避免了 HCS 模型每层只要一个控制/被控制系统可能造成大量重复的情况。

4.3　变量影响图模型有效性验证与分析

在生成系统控制结构之后,用控制结构和过程分析模块对控制环所对应的控制过程进行系统变量提取及其关系分析,并进行变量影响图建模,以便进行安全性分析。为了验证变量影响图的有效性,本文将每条控制需求作为一个系统目标,并且根据文献[1,16]所示的 TDC、ITP、NPP 系统危

北航学报 赠阅

险,对这 3 个系统实例的全局控制过程分别用基于 objectives[21]、hazards[16]、场景的方法提取系统变量及其关系,并进行变量合并、层次化等处理,构建相应的被控制系统模型 SED[21]、层次过程模型 HPM[16] 以及变量影响图 VEG,所生成的 3 种过程模型情况如表 5 所示。

表 5 过程模型情况

Table 5 Information of process models

实例	基于 objectives 的 SED		基于 hazards 的 HPM		基于场景的 VEG	
	SV 数量	关系数量	SV 数量	关系数量	SV 数量	横向/纵向关系数量
TDC	5	0	5	0	5	2/0
ITP	15	7	8	4	19	4/6
NPP	16	6	10	3	20	3/12

对于 3 种过程模型,分别用 SVC 和 SVRC 对其进行评估。

SVC = 模型中 SV 数量/SVM,SVM = 所有模型 SV 数量和 - SV 重复数,表示模型构造方法提取的系统变量数量在全部系统变量中的比例。

SVRC = 2 × 模型中关系数量/[SVM(SVM - 1)],表示模型构造方法分析出的关系数量在所有可能关系中的比例。

评估结果如图 9 所示。

由图 9(a)可知,对于 3 个实例,VEG 的 SVC 评估优于或等于 HPM 和 SED,表明 VEG 通过对场景进行分析提取系统变量的方法,解决了 HPM 和 SED 建模方法受限于系统 hazards 和 objectives 描述简单且数量较少的问题,可以得到相对比较完备的系统变量集合。

图 9(b)中,对于简单的 TDC 系统,所得系统变量属于同一层次,不存在组成关系或纵向关系,其 HPM、SED 模型的 SVRC 值为 0,而 VEG 模型中由于包含了这些系统变量间的 2 个横向关系,所以其 SVRC 值为 0.2;对于较复杂的 ITP 和 NPP 系统,VEG 模型的横向关系和纵向关系之和大于其 HPM 模型和 SED 模型中的关系数,因此本文 VEG 模型的 SVRC 评估结果较优。由上可知,对于同一系统,本文的 VEG 模型所得的系统变量间关系较完备。

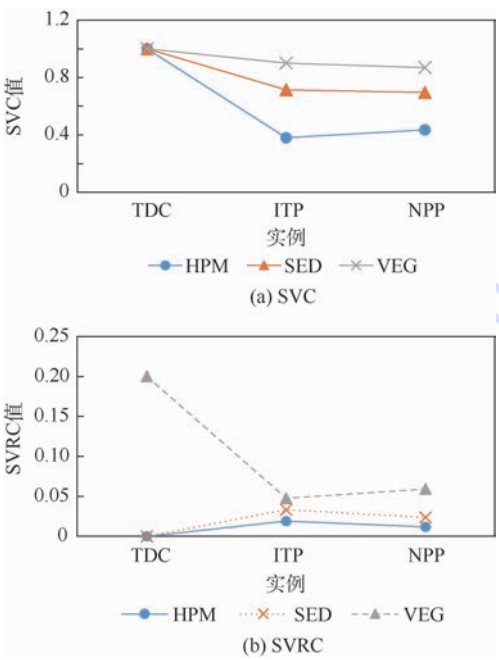


图 9 SVC 和 SVRC 评估结果

Fig. 9 Results of SVC and SVRC assessment

4.4 安全性需求分析方法有效性验证与分析

根据控制结构和变量影响图等,安全性需求分析模块自动构建 HCAT 并生成 HCA。由于在相同系统级别下,HCA 和安全性需求一一对应,本文通过 HCA 对安全性需求提取方法进行验证。该节对 TDC、ITP、NPP 相同顶层控制过程的过程模型进行简化后,分别用文献[16]方法和本文方法生成 HCA,并用生成的 HCA 数量、HCA 错误数和 HCA 矛盾数作为指标,在安全性需求的生成能力和生成质量方面,验证本文安全性需求分析方法的有效性,用平均空间使用度 ASD 验证方法的效率,ASD = 当前层次存储所需空间数/(HCA 数量 - HCA 错误数)。实验结果如图 10 和表 6 所示。

从表 6 可知,本文方法生成的 HCA 总数少于或等于文献[16],但去除其中错误的 HCA 和不一致的 HCA 后,本文方法所得的有效 HCA 数大于或等于文献[16]方法。这表明本文的安全性需求分析方法得到的安全性需求比较完备。本文

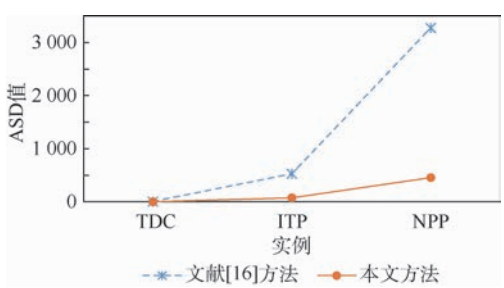


图 10 ASD 评估结果

Fig. 10 Results of ASD assessment

北航学报 赠阅

表 6　安全性需求分析方法评估结果

Table 6　Assessment results of safety requirements analysis methods

实例	CA 名	文献[16]方法				本文方法			
		HCA 总数	HCA 错误数	存在的矛盾数	有效 HCA 数	HCA 总数	HCA 错误数	解决的矛盾数	有效 HCA 数
TDC	Open Door	12	0	3	9	9	0	3	9
ITP	ITP Request	23	4	0	19	21	0	0	21
NPP	Close MSIV	52	9	14	36	39	0	14	39

在分析过程中对矛盾和错误进行自动检测和处理的方法,不但解决了文献[16]方法将矛盾检测放在后期由人工处理的不足,而且使得 HCA 错误数为 0,克服了文献[16]方法由于缺乏错误检测机制所造成的 HCA 可能存在错误的缺陷。另外,由于 2 种方法矛盾检测原理相同,并且 3 个实例中不存在功能性需求和危险间的矛盾,所以检测到的 HCA 矛盾相同。从图 10 可知,本文方法的空间使用度随模型复杂度的增加,具有明显的优势,这是因为本文在安全性需求分析过程中,使用变量影响图中的横向关系做了优化和剪枝。对同一系统,相对于文献[16]方法,本文方法不但能用较少的存储资源生成相对较完备的安全性需求,而且通过矛盾和错误检测处理,保证了所生成的安全性需求的正确性和一致性。

5　结　论

本文使用 STAMP 因果模型,对基于需求场景的控制特征进行安全性需求分析的方法进行了研究。结果表明:

1)通过对系统场景的控制特征时序图进行分析,在一定程度上将基于 STAMP 的安全性分析从设计阶段提前到需求阶段,扩展了 STAMP 的应用范围。

2)使用变量影响图模型自动构造 HCAT,分析得到 HCA 和安全性需求,弥补了 STPA 方法缺乏危险性活动生成方法的缺陷。

3)通过在安全性需求分析过程中进行矛盾和错误检测处理,保证了所得到的安全性需求的正确性和一致性,并通过剪枝优化,提高了方法的分析效率。

参考文献(References)

[1] FLEMING C H, SPENCER M, THOMAS J, et al. Safety assurance in NextGen and complex transportation systems[J]. Safety Science, 2013, 55: 173-187.

[2] LEVESON N G. Engineering a safer world: Systems thinking applied to safety[M]. Boston, Massachusetts: MIT Press, 2012: 212-227.

[3] BLACK J, KOOPMAN P. Indirect control path analysis and goal coverage strategies for elaborating system safety goals in composite systems[C]// Proceedings of the 2008 14th IEEE Pacific Rim International Symposium on Dependable Computing. Piscataway, NJ: IEEE Press, 2008: 184-191.

[4] SUPAKKUL S, CHUNG L. Applying a goal-oriented method for hazard analysis: A case study[C]// Proceedings of the 4th International Conference on Software Engineering Research, Management and Applications. Piscataway, NJ: IEEE Press, 2006: 22-30.

[5] FIRESMITH D. Engineering safety-related requirements for software-intensive systems[C]// Proceedings of the 27th International Conference on Software Engineering. Piscataway, NJ: IEEE Press, 2005: 720-721.

[6] FIRESMITH D. Engineering safety requirements, safety constraints, and safety-critical requirements[J]. Journal of Object Technology, 2004, 3(3): 27-42.

[7] CHUNG L, SUPAKKUL S. Capturing and reusing functional and non-functional requirements knowledge: A goal-object pattern approach[C]// 2006 IEEE International Conference on Information Reuse and Integration. Piscataway, NJ: IEEE Press, 2006: 539-544.

[8] FIRESMITH D. A taxonomy of safety-related requirements[C]// 2004 Proceedings of the Workshop on Requirements for High Assurance Systems (RHAS'04). Piscataway, NJ: IEEE Press, 2004: 11.

[9] CHANG W, BAO X, LI X. A study on airborne software safety requirements patterns[C]// 2013 IEEE 7th International Conference on Software Security and Reliability. Piscataway, NJ: IEEE Press, 2013: 131-136.

[10] CHEN Z, MOTET G. System safety requirements as control structures[C]// 2009 33rd Annual IEEE International Computer Software and Applications Conference. Piscataway, NJ: IEEE Press, 2009: 324-331.

[11] YANG C. Software safety testing based on STPA[C]// 3rd International Symposium on Aircraft Airworthiness (ISAA). Amsterdam: Elsevier, 2014, 80: 399-406.

[12] ALLENBY K, KELLY T. Deriving safety requirements using scenarios[C]// 5th IEEE International Symposium on Requirements Engineering. Piscataway, NJ: IEEE Press, 2001: 228-235.

[13] VYAS P, MITTAL R K. Eliciting additional safety requirements from use cases using SFTA[C]// 2012 1st International Conference on Recent Advances in Information Technology. Piscataway, NJ: IEEE Press, 2012: 163-169.

[14] GUILLERM R, DEMMOU H, SADOU N. A safety requirement engineering method and tool[C]// 2013 21st IEEE Internation-

北航学报 赠阅

al Conference on Requirements Engineering. Piscataway, NJ: IEEE Press, 2013: 328-329.

[15] MENON C, KELLY T. Eliciting software safety requirements in complex systems[C]// 2010 4th Annual IEEE Systems Conference. Piscataway, NJ: IEEE Press, 2010: 616-621.

[16] THOMAS J P. Extending and automating a systems-theoretic hazard analysis for requirements generation and analysis[D]. Boston, Massachusetts: MIT, 2013: 71-87.

[17] ASHBY W R. An introduction to cybernetics[M]. London: Chapman & Hall, 1957: 202-216.

[18] CHECKLAND P. Systems thinking, systems practice[M]. New York: John Wiley & Sons, 1981: 271-284.

[19] VANHATALO J, VÖLZER H, LEYMANN F. Faster and more focused control-flow analysis for business process models through SESE decomposition[C]//5th International Conference on Service-Oriented Computing (ICSOC 2007). Berlin: Springer, 2007, 4749: 43-55.

[20] FLEMING C H. Safety-driven early concept analysis and development[D]. Boston, Massachusetts: MIT, 2015: 53-92.

[21] INGHAM M D, RASMUSSEN R D, BENNETT M B, et al. Generating requirements for complex embedded systems using state analysis[J]. Acta Astronautica, 2006, 58(12): 648-661.

[22] LEVESON N G. A systems-theoretic approach to safety in software-intensive systems[J]. IEEE Transactions on Dependable and Secure Computing, 2004, 1(1): 66-86.

[23] ABDULKHALEQ A, WAGNER S. A-STPA: An open tool support for system-theoretic process analysis[C]//2014 STAMP Conference. Boston, Massachusetts: MIT Press, 2014.

作者简介：

朱丹江 男，博士研究生。主要研究方向：软件工程、软件测试。

E-mail: danjiang@buaa.edu.cn

姚淑珍 女，博士，教授，博士生导师。主要研究方向：先进软件工程技术、形式化方法、Petri 网理论。

Tel.：010-82339434

E-mail: szyao@buaa.edu.cn

谭火彬 男，博士，讲师。主要研究方向：软件工程、软件建模。

Safety requirements analysis method based on control characteristics of scenarios

ZHU Danjiang, YAO Shuzhen*, TAN Huobin

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

Abstract: Safety requirements are critical to ensure the system safety. With the increase of system complexity and coupling degree, it becomes more difficult to analyze and extract the safety requirements. We construct the variables effect graph which describes the relationships among system variables in the control process through control structure and process analysis modeling for the system scenario. Then we present a safety requirements analysis method. By using the method, the control process is analyzed with variables effect graph etc., and the hazardous control actions based on the systems-theoretic accident model and process (STAMP) are created, and thereby the system safety requirements are generated. The experimental results show that our method is effective on correctness and consistency.

Key words: safety requirements; requirements analysis; control analysis; scenario analysis; control modeling

Received: 2015-11-17; **Accepted**: 2015-12-18; **Published online**: 2016-01-20 17:07

URL: www.cnki.net/kcms/detail/11.2625.V.20160120.1707.007.html

Foundation item: Aeronautical Science Foundation of China (2013ZC51023)

* **Corresponding author**. Tel.: 010-82339434 E-mail: szyao@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0739

基于故障行为模型的产品寿命分析方法

查国清，井海龙，陈云霞*，康锐

（北京航空航天大学 可靠性与系统工程学院，北京 100083）

摘 要：针对传统可靠性工程方法无法满足高可靠长寿命需求产品设计与分析要求的现状。本文充分考虑引起产品寿命分散性的内因和外因的不确定性，在产品功能模型或数字样机模型的基础上，结合产品的故障行为模型，建立了一种基于产品故障行为模型的产品寿命分析方法。在该方法中，内因参数与外因参数的分散性分别可以用相应的概率密度函数来表示，都是通过对实际使用情况和生产条件的估计得到的，因此，按照这种方法评估得到的产品寿命特征，能够弥补基于故障物理的可靠性预计方法无法体现产品正常使用条件的不足。

关 键 词：故障行为模型；机理分析；耗损性失效机理；寿命分析；故障物理
中图分类号：V120.50
文献标识码：A **文章编号**：1001-5965(2016)11-2371-07

可靠性工程是研究产品全寿命过程中与故障作斗争的技术科学，它从系统的整体性及其同外界环境的辩证关系出发，研究产品发生故障的机理与规律，预防、预测和控制产品故障的理论与方法，并运用这些机理与规律、理论与方法开展一系列相关的技术与管理活动[1-2]。

从20世纪80年代后期，我国在装备研制中开始全面推广可靠性工程技术，这些技术来源于20世纪五六十年代国外（以美国为主）形成的定性分析与定量评估相结合的工程方法[3-5]，在设计上主要依赖设计准则（基于工程经验的定性设计）、裕度设计（预留安全系数）[6]和余度设计（增加冗余备份）等提高产品可靠性，然后基于可靠性统计试验定量评估产品可靠性水平。这种方法的认识基础是产品故障具有不可归因的随机特性，理论基础是概率论与数理统计，工程实施中更多关注产品故障的结果数据，不能关注产品的故障过程。这就导致在设计分析方面，只能采取“事后”改进的方法，而对故障原因缺乏认知，针对高可靠长寿命需求的产品，面临改进无方向的弊端[7-9]。综上所述，面对高可靠长寿命的技术要求，需要研究建立一种基于故障过程认知的可靠性技术来设计和分析产品的可靠性水平。

产品的故障行为模型能够描述系统或单元在复杂应力条件下，多种故障机理综合作用的故障规律[10-11]。从本质上看，故障行为模型属于确定性模型，即在一组给定的输入条件下，系统或单元的故障行为是确定的，故障时间可以由故障行为模型进行预测。但是，实际使用中的产品，所经历的使用条件和环境条件等外因条件的差异是很大的。另一方面，即使采用同样的设计方法，同一批次生产出的同一批产品，由于材料性质、工艺条件等内因方面的不可能完全相同，产品的性能也存在着差异性[12]。

本文在前人研究成果的基础上，充分考虑引

收稿日期：2015-11-11；**录用日期**：2015-12-06；**网络出版时间**：2016-01-21 15:47
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160121.1547.004.html
*通讯作者：Tel.：010-82338909 E-mail：chenyunxia@buaa.edu.cn

引用格式：查国清，井海龙，陈云霞，等. 基于故障行为模型的产品寿命分析方法[J]. 北京航空航天大学学报，2016，42(11)：2371-2377. ZHA G Q，JING H L，CHEN Y X，et al. Life analysis method for products based on failure behavior models[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2371-2377 (in Chinese).

北航学报 赠阅

起产品寿命分散性的内因和外因的不确定性,在产品功能模型或数字样机模型的基础上,结合产品的故障行为模型,建立一种基于产品故障行为模型的产品寿命分析方法。在该方法中,内因参数与外因参数的分散性分别可以用相应的概率密度函数 $p(\boldsymbol{I})$ 和 $p(\boldsymbol{E})$ 来表示,其中 $\boldsymbol{I}$ 和 $\boldsymbol{E}$ 分别为内因参数向量和外因参数向量。在故障行为模型中,将内因参数、外因参数分别视为来自这两个分布的随机变量,就可以通过数理统计的方法估计产品在实际使用条件下的平均无故障间隔时间(Mean Time to Failure,MTTF)。由于概率密度函数 $p(\boldsymbol{I})$ 和 $p(\boldsymbol{E})$ 是通过对实际使用情况和生产条件的估计得到的,因此,按照这种方法评估得到的 MTTF,能够弥补基于故障物理的可靠性预计方法无法体现产品正常使用条件的不足。

1 产品寿命分析方法

1.1 原理与流程

基于故障行为模型的产品寿命分析方法首先在进行内因和外因的特征参数提取及分布规律确定的基础上,利用相应的概率密度函数来表示故障行为模型中的内因参数和外因参数的分散性;其次,在故障行为模型中,将内因参数、外因参数分别视为来自这两个分布的随机变量,通过数理统计的方法估计产品在实际使用条件下的故障行为,确定产品的敏感故障机理和理论寿命,为产品的设计改进提供支撑。具体分析原理与流程如图 1所示。

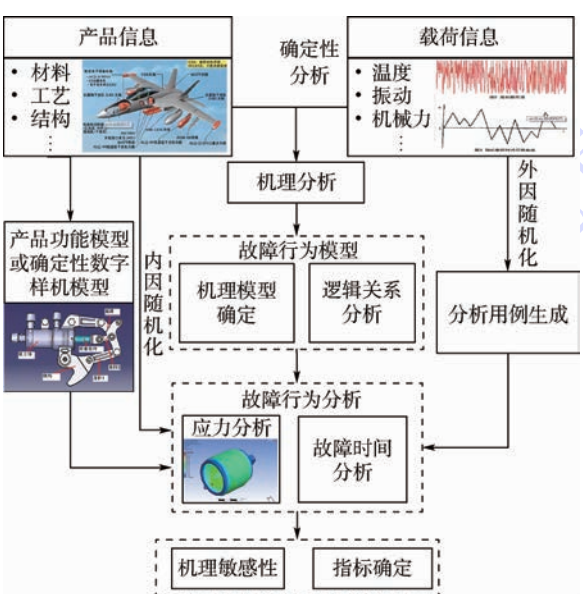


图 1 产品寿命分析原理与流程

Fig. 1 Principle and process of life analysis for products

1.2 机理分析流程

机械类产品在寿命周期内主要表现出耗损型失效机理,机理分析根据给定的载荷谱或任务剖面,结合产品的组成、结构、原理,定性分析,确定产品每个单元对应的耗损型失效机理以及对应的敏感应力。机理分析的主要工具是故障模式、机理和影响分析(Failure Mode, Mechanisms and Effects Analysis,FMMEA)[13-14],机理分析流程如图 2 所示。

1.3 故障行为分析方法

故障行为分析是指在产品性能模型或数字样机模型的基础上,考虑产品内外因参数分散性,利用故障行为模型进行产品应力分析和寿命分析的过程。以下以产品的 MTTF 的计算为例介绍产品的故障行为分析方法。

MTTF 的计算是利用故障行为模型,综合考虑产品内因参数和外因参数的分散性,对产品的寿命指标进行预计。指标计算的原理如式(1)所示:

$$\mathrm{MTTF}=\int_{I}\int_{E}B_t(\boldsymbol{I},\boldsymbol{E})p(\boldsymbol{I})p(\boldsymbol{E})\mathrm{d}\boldsymbol{I}\mathrm{d}\boldsymbol{E} \tag{1}$$

式中:$B_t(\boldsymbol{I},\boldsymbol{E})$ 表示以故障时间表征的故障行为模型。

对于内因参数与外因参数取值为离散的情况,式(1)可以表示为

$$\mathrm{MTTF}=\sum_{i=1}^{m}\sum_{j=1}^{n}B_t(\boldsymbol{I},\boldsymbol{E}\,|\,\boldsymbol{I}=I_i,\boldsymbol{E}=E_j)p(\boldsymbol{I}=I_i)p(\boldsymbol{E}=E_j) \tag{2}$$

式中:$p(\boldsymbol{I}=I_i)$ 和 $p(\boldsymbol{E}=E_j)$ 分别为内因参数、外因参数的分布律。

故障行为分析的具体实施步骤如图 3 所示。

1) 建立故障行为模型

采用故障时间表述方式,建立产品的故障行为模型:

$$T=B_t(\boldsymbol{I},\boldsymbol{E}) \tag{3}$$

式中:T 为产品的寿命。

2) 内因特征参数提取

分析确定式(3)中的内因集合中的参数可能产生的分散性,用分布律或者密度函数的形式表示,即 $p(\boldsymbol{I})$。内因参数中的分散性主要来源于生产、制造等过程。

3) 环境特征参数提取

分析确定式(3)中的外因集合中的参数可能产生的分散性,用分布律或者密度函数的形式表示,即 $p(\boldsymbol{E})$。外因参数中的分散性主要产品使用环境的不同,可以由产品寿命周期中的环境

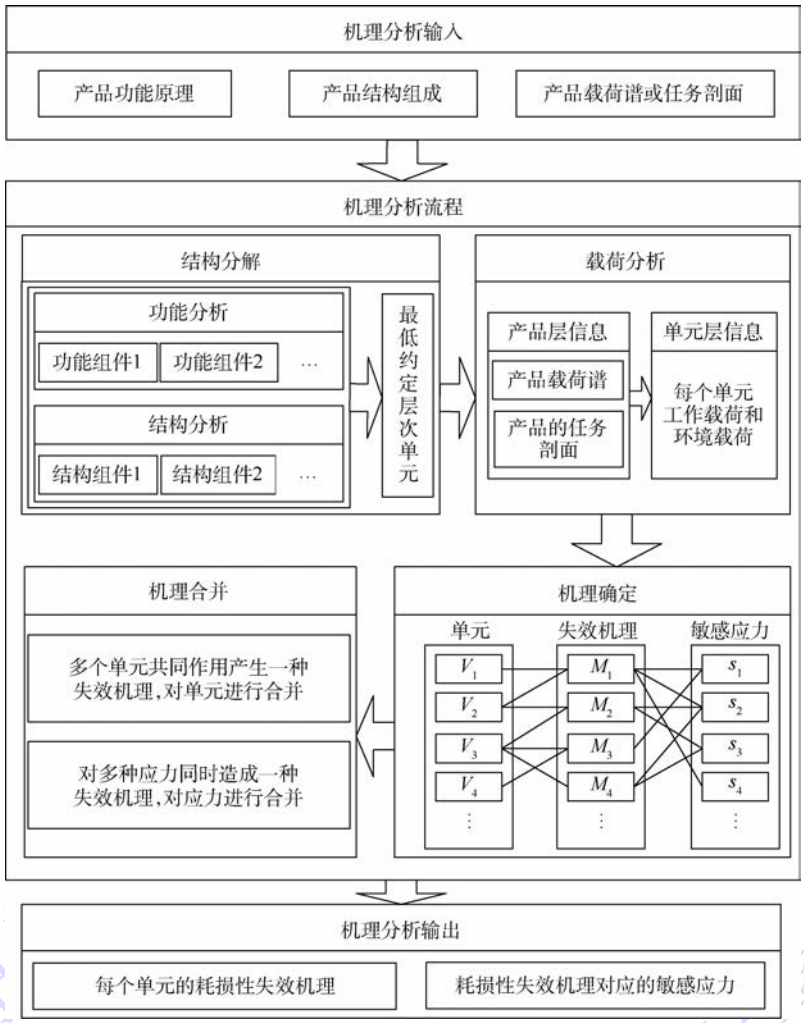


图 2　机理分析流程

Fig. 2　Process of mechanism analysis

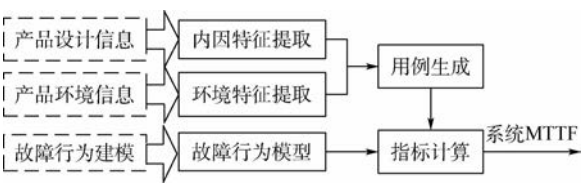


图 3　MTTF 分析步骤

Fig. 3　Analysis procedure of MTTF

剖面和使用剖面转化获得。

4) 按照式(1)或式(2)计算产品的 MTTF

当故障行为模型较为复杂时,则可以通过仿真的方法计算式(1)或式(2)的值:首先,通过环境特征提取确定能够表征实际使用环境的内因参数、外因参数分布 $p(\boldsymbol{I})$ 和 $p(\boldsymbol{E})$;然后,进行分析用例生成,即按照 $p(\boldsymbol{I})$ 和 $p(\boldsymbol{E})$ 仿真生成 n 组测试条件,分别按照式(1)或式(2)计算故障时间。对 n 个故障时间求取算术平均值,即得到所需评估的 MTTF 指标。

1.4　分析用例生成方法

对于形式复杂的故障行为模型,式(1)或式(2)的解析解往往是无法求得的,此时需要通过数值模拟的方法进行 MTTF 的评估。其中,一个重要的环节即为分析用例生成。

分析用例生成是指从给定的内因、外因概率密度函数 $p(\boldsymbol{I})$ 和 $p(\boldsymbol{E})$ 中,生成一组样本量为 N 的随机样本。利用这些样本,可以通过蒙特卡罗方法,进行 MTTF 的估计。生成随机样本的方法有很多种,例如,直接抽样法(反函数法)、挑选抽样法、复合抽样法等。以下以直接抽样法为例,示意测试用例生成的方法[15]。

对于任意给定的分布函数 $F(x)$,直接抽样法的流程如下:首先,生成一组样本量为 N 的服从区间[0,1]均匀分布的伪随机数,记为 $\xi_1,\xi_2,\cdots,\xi_N$。然后,按照式(4)确定服从 $F(x)$ 的随机序列。

$$X_n=\inf_x(F(x)\geqslant\xi_n)\qquad n=1,2,\cdots,N \tag{4}$$

对于离散型分布 $F(x)=\sum_{x_i<x}P_i$,其中,$x_1,x_2,\cdots$ 为相应的分布函数间断点,$P_1,P_2,\cdots$ 为相应

北航学报 赠阅

的概率,则直接抽样法可以表示为

$$X_F = x_I \qquad \sum_{i=1}^{I-1} P_i < \xi \leqslant \sum_{i=1}^{I} P_i \tag{5}$$

对于连续型分布,如果分布函数 $F(x)$ 的反函数 $F^{-1}(y)$ 存在,则直接抽样法可以表示为

$$X_F = F^{-1}(\xi) \tag{6}$$

1.5 机理敏感性分析方法

机理敏感分析的目的是通过分析确定系统的薄弱环节或主机理,从而为系统的设计改进提供方向。薄弱环节,是针对系统而言的,指的是系统中最容易发生故障的单元;主机理,是针对单元而言的,指的是单元最容易出现的故障机理。由于二者从分析方法上是相同的,故以下仅以主机理分析为例进行说明。

假定某一单元的故障行为模型的时间表示为式(3)所示,该单元共有 n 种潜在的故障机理,其机理模型为式(7),给定环境条件 E_g,则满足式(8)的第 j 种机理称为环境条件 E_g 下的敏感机理。

$$T_i = M_{ti}(I_i, E_i) \tag{7}$$

$$j = \arg\inf_{i} \{ M_{ti}(\boldsymbol{I}_i, E_g) \geqslant B_t(\boldsymbol{I}, E_g) \} \tag{8}$$

式中:T_i 为第 i 个机理对应的寿命值。

通常,机理之间的关系表现为时间竞争关系,即

$$T = \bigwedge_{i=1}^{n} T_i \tag{9}$$

此时,式(8)简化为

$$j = \arg_i \{ B_{ti}(I_i, E_g) = \bigwedge_{i=1}^{n} B_{ti}(I_i, E_g) \} \tag{10}$$

假设产品实际使用的环境条件可以用密度函数 $p(\boldsymbol{E})$ 来表示,按照 $p(\boldsymbol{E})$ 随机生成 m 种环境条件,对每一种环境条件分别求取敏感机理,则成为敏感机理次数最多的故障机理即为该单元的主机理。主机理分析的实施步骤如下:

1) 建立故障行为模型,如式(3)或式(7)所示。

2) 通过环境特征提取的方法,获得表征产品实际工作条件的环境应力分布,记为 $p(\boldsymbol{E})$。

3) 按照 $p(\boldsymbol{E})$ 随机生成 m 组环境条件,分别按照式(8)确定每一种环境条件下的敏感机理。称敏感机理次数最多的故障机理即为主机理。

4) 在通过分析确定主机理或薄弱环节之后,应该针对主机理或者薄弱环境,进行相应的设计改进,从而提高产品的可靠性水平。

2 案例分析

本节将以某航空液压柱塞泵为应用案例,综合运用上述方法,开展基于故障行为模型的寿命分析。该柱塞泵是利用发动机提供的动力作为输入,为液压系统各部件提供一定压力、流量的液压能。某柱塞泵机理分析结果如表 1 所示。

表 1 某柱塞泵机理分析结果

Table 1 Results of mechanism analysis for piston pump

序号	薄弱环节单元	主机理
1	滑履、卡盘	磨损
2	转子	疲劳
3	内轴、传动轴	接触疲劳
4	弹簧	应力松弛
5	四氟塑料圈	老化
⋮	⋮	⋮

根据机理分析结果,选择相应的寿命模型,计算各耗损性失效机理的失效时间,式(11)~式(13)为建立的滑履-卡盘磨损寿命模型[16]:

$$T = \frac{2\pi H h_{\max}(R_2 - R_3)}{KF\omega_0} \tag{11}$$

$$\lg K = 5\lg f - 2.27 \tag{12}$$

$$\omega_0 = \sqrt{\frac{\pi^3 (R_2^2 - R_1^2)\mu n^2 R^2 t_0}{1\,800 I h_0}} \tag{13}$$

式中:R_1 为滑履下表面内圈直径,mm;R_2 为滑履上表面外圈直径,mm;R_3 为滑履上表面内圈直径,mm;H 为材料硬度(布氏硬度),MPa;h_0 为油膜厚度,mm;$h_{\max}$ 为最大允许磨损深度,mm;f 为摩擦系数;F 为法向载荷,N;I 为滑履绕自身中心轴线转动惯量,kg·m^2;μ 为油液的动力黏度,Pa·s;R 为柱塞分布圆半径,mm;n 为转子转速,r/min;t_0 为从启动至额定转速所需时间,s;ω_0 为滑履自转角速度,rad/s。

利用式(11)~式(13)的寿命计算模型计算滑履-卡盘磨损机理的理论寿命值,模型参数表如表 2 所示,可得滑履-卡盘的磨损寿命为 1 413 h。

表 2 滑履-卡盘磨损寿命模型参数

Table 2 Parameters of wear life model for slipper and chuck

序号	参数	数值
1	F/N	34.1
2	R_1/mm	5.15
3	R_2/mm	7.2
4	R_3/mm	5.5
5	$h_{\max}$/mm	1
6	f	0.075
7	H/MPa	500
8	I/(10^{-7}kg·m^2)	1.927
9	μ/(Pa·s)	0.006 63
10	R/mm	21.5
11	n/(r·min^{-1})	4 000
12	t_0/s	20
13	h_0/mm	0.085

通过上述方法可以计算出所有最低约定层次单元的理论寿命计算结果如表 3 所示。

根据确定性寿命模型,考虑寿命模型参数的不确定特征,进行内因参数提取并确定其分布特性;同时根据全寿命周期载荷谱,确定其载荷的特征参数及分布特性。

对每个内因参数和外因参数进行蒙特卡罗抽样,将每次抽样值代入寿命模型计算得到每个失效机理的若干个寿命值。然后对每个失效机理的所有寿命值进行分布曲线拟合,得到每个失效机理对应的寿命分布,确定每个失效机理在给定置信度下的寿命区间。

下面以滑履-卡盘磨损寿命区间计算为例说明求解过程。

为了方便计算,本案例采取内因参数、外因参数和变异系数皆取为 0.01。根据上述步骤,确定抽样次数为 1 000 次,将每次抽样值代入寿命模型计算得到 1 000 个寿命值,通过分布曲线拟合得到寿命服从的正态分布,最后通过分析寿命分布特点确定置信度为 95% 的滑履-卡盘磨损寿命区间为[1 236,1 606]h。

同理,可对所有最低约定层次单元寿命区间进行计算,计算结果汇总如表 4 所示。

考虑柱塞泵的各机理竞争失效机制,采取短板原则,给出柱塞泵的整寿命区间的计算方法。在表 4 的基础上,分别对每个失效机理对应的理论寿命分布进行蒙特卡罗抽样,通过取小原则确定的每次抽样的理论寿命最小值作为当次抽样的柱塞泵整体寿命,抽样确定的理论寿命最小值对应的失效机理为柱塞泵的敏感机理,敏感机理对应的单元为柱塞泵的薄弱环节。

对表 4 确定的理论寿命分别进行 1 000 次仿真抽样则有 1 000 个取小值,则得到柱塞泵整体寿命值的 1 000 个计算结果,然后通过分布曲线拟合可认为柱塞泵整体寿命服从正态分布,最后基于寿命分布计算得到柱塞泵的敏感机理为滑履-卡盘磨损机理,薄弱环节为滑履和卡盘,最终确定的柱塞泵置信度为 95% 的寿命区间为[1 219,1 540]h。

表 3　最低约定层次单元理论寿命计算结果

Table 3　Calculation results of theoretical life for the lowest level units

序号	薄弱环节单元	主机理	模型名称	寿命/h
1	滑履、卡盘	磨损	滑履-卡盘磨损寿命计算模型	1 413
2	转子	疲劳	名义应力(时域法)寿命计算模型	1 742
3	内轴、传动轴	接触疲劳	渐开线花键接触疲劳寿命计算模型	5 512
4	弹簧	应力松弛	弹簧应力松弛寿命计算模型	3 200
5	四氟塑料圈	老化	塑料老化寿命计算模型	4 786
⋮	⋮	⋮	⋮	⋮

表 4　最低约定层次单元寿命区间计算结果

Table 4　Calculation results of life interval for the lowest level units

序号	单元名称	机理	模型名称	寿命均值/h	寿命区间/h
1	滑履、卡盘	磨损	滑履-卡盘磨损寿命计算模型	1413	[1236,1606]
2	转子	疲劳	名义应力(时域法)寿命计算模型	1742	[1533,2054]
3	内轴、传动轴	接触疲劳	渐开线花键接触疲劳寿命计算模型	5512	[5021,6102]
4	弹簧	应力松弛	弹簧应力松弛寿命计算模型	3200	[2713,3518]
5	四氟塑料圈	老化	塑料老化寿命计算模型	4786	[4205,5432]
⋮	⋮	⋮	⋮	⋮	⋮

3　结　论

本文充分考虑了引起产品寿命分散性的内因和外因的不确定性,在产品功能模型或数字样机模型的基础上,结合产品的故障行为模型,提出了一种基于产品故障行为模型的产品寿命分析方法。具体创新点如下:

1) 在产品功能模型或数字样机模型的基础上,系统地提出了一种基于故障行为模型的产品寿命分析方法,为具有长寿命指标产品的设计改进提供了技术支撑。

2) 给出了综合考虑内因和外因分散性的故障行为分析方法,内因参数与外因参数的分散性分别用相应的概率密度函数来表示,并且从实际使用情况和生产条件的估计得到,为准确评估产品在正常使用工况下的故障行为提供了方法支撑。

3) 提出了机理敏感性的分析方法,确定了产

北航学报 赠阅

品的寿命薄弱环节和敏感机理,为具有长寿命指标产品的设计改进指明了方向。

参考文献(References)

[1] SALEH J H, MARAIS K. Highlights from the early (and pre-) history of reliability engineering[J]. Reliability Engineering and System Safety, 2006, 91(2): 249-256.

[2] MCLINN J. A short history of reliability[J]. Journal of Reliability Data Analysis Center, 2011, 54(1): 8-15.

[3] MEEKER W Q, ESCOBAR L A. Statistical methods for reliability data[M]. Hoboken: John Wiley & Sons, 2014: 350-420.

[4] ZACKS S. Introduction to reliability analysis: Probability models and statistical methods[M]. Berlin: Springer Science & Business Media, 2012: 100-150.

[5] GNEDENKO B V, BELYAYEV Y K, SOLOVYEV A D. Mathematical methods of reliability theory[M]. Manhattan: Academic Press, 2014: 300-500.

[6] ZHOU T, DONG J B, PENG L X, et al. Designing method research of generalized reliability margin for complex system[C]// IEEE Conference on Prognostics and System Health Management (PHM). Piscataway, NJ: IEEE Press, 2012: 1-4.

[7] MEEKER W Q, HAMADA M. Statistical tools for the rapid development and evaluation of high-reliability products[J]. IEEE Transactions on Reliability, 1995, 44(2): 187-198.

[8] 邓爱民. 高可靠长寿命产品可靠性技术研究[D]. 长沙: 国防科学技术大学, 2006: 1-7.
DENG A M. Research on reliability technology of high-reliability and long-lifetime products[D]. Changsha: National University of Defense Technology, 2006: 1-7.

[9] 陈云霞, 赵安, 康锐. 非电产品寿命分析和试验方法研究[C]// 中国运筹学会可靠性分会第九届可靠性学术会议论文集. 长沙: 中国运筹学会, 2013: 205-211.
CHEN Y X, ZHAO A, KANG R. Research on life analysis and test method for non-electric products[C]// The Reliability of the Operational Research Society of China, the Ninth Session of the Reliability Academic Conference. Changsha: Operations Research Society of China, 2013: 205-211 (in Chinese).

[10] ZENG Z G, CHEN Y X, KANG R. Failure behavior modeling: Towards a better characterization of product failures[J]. Chemical Engineering Transactions, 2013, 33(2): 571-576.

[11] ZENG Z G, KANG R, CHEN Y X. A physics-of-failure-based approach for failure behavior modeling: With a focus on failure collaborations[C]// Proceedings of Annual European Safety and Reliability Conference (ESREL2014). Piscataway, NJ: IEEE Press, 2014: 1-9.

[12] 陈云霞. 性能与可靠性一体化建模和分析方法研究[D]. 北京: 北京航空航天大学, 2004: 22-24.
CHEN Y X. Study on method of modeling and analysis for integrating performance and reliability[D]. Beijing: Beihang University, 2004: 22-24 (in Chinese).

[13] MATHEW S, ALAM M, PECHT M. Identification of failure mechanisms to enhance prognostic outcomes[J]. Journal of Failure Analysis and Prevention, 2012, 12(1): 66-73.

[14] CHEN Y X, JING H L, LIAO X, et al. Methods to determine stress profile in ALT based on theoretical life models[J]. Chinese Journal of Aeronautics, 2015, 28(3): 729-736.

[15] 杨振海, 程维虎. 非均匀随机数产生[J]. 数理统计与管理, 2006, 25(6): 750-756.
YANG Z H, CHENG W H. The common method of generating random number for the non-uniform distribution[J]. Jouranl of Applied Statistics and Management, 2006, 25(6): 750-756 (in Chinese).

[16] LIAO X, CHEN Y X. Research on the wear life of slipper in aero-hydraulic pump based on Archard wear model and liquid friction work[C]// 2013 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (QR2MSE). Piscataway, NJ: IEEE Press, 2013: 930-934.

作者简介:

查国清 男,博士研究生。主要研究方向:故障控制论、公共安全。

E-mail: zgq@buaa.edu.cn

井海龙 男,博士研究生。主要研究方向:产品耐久性分析与试验技术。

Tel.: 010-82338236

E-mail: jinghailong@buaa.edu.cn

陈云霞 女,博士,教授,博士生导师。主要研究方向:产品故障学、复杂系统可靠性建模与仿真、高可靠长寿命产品设计与试验技术等。

Tel.: 010-82338909

E-mail: chenyunxia@buaa.edu.cn

康锐 男,教授,博士生导师。主要研究方向:确信可靠性理论与方法、信息物理系统可靠性建模与仿真、可靠性系统工程方法论、故障控制论。

E-mail: kangrui@buaa.edu.cn

北航学报 赠阅

Life analysis method for products based on failure behavior models

ZHA Guoqing, JING Hailong, CHEN Yunxia*, KANG Rui

(School of Reliability and Systems Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: With the situation that traditional reliability engineering methods cannot meet the demand of design and analysis for the products with high reliability and long life index, this paper considers the uncertainty of inner and outside factors which cause the life dispersion, and on the basis of product functional models or digital prototype models, with the failure behavior model of products, a method is proposed to conduct the life analysis for products with long life index. The dispersion of inner and outside parameters can be expressed by the probability density function and obtained by analyzing the actual use and production conditions. Thus the life characteristics of products obtained with this method can remedy the disadvantage that the reliability prediction method based on physics of failure cannot reflect the actual normal use conditions of the products.

Key words: failure behavior model; mechanism analysis; wear-out failure mechanism; life analysis; physics of failure

北京航空航天大学学报

Received: 2015-11-11; **Accepted**: 2015-12-06; **Published online**: 2016-01-21 15:47

URL: www.cnki.net/kcms/detail/11.2625.V.20160121.1547.004.html

* **Corresponding author**. Tel.: 010-82338909　E-mail: chenyunxia@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0676

未来航空燃料原料可持续性研究

赵晶，郭放，阿鲁斯，杨晓奕*

（北京航空航天大学 能源与动力工程学院 能源与环境国际中心，北京 100083）

摘 要：运用全生命周期评价方法，评价未来航空燃料原料可持续性。以传统石油基航空燃料为基准，分析未来8种航空燃料原料（煤、天然气、能源藻、麻疯树、大豆、棕榈、油菜籽及亚麻荠）的可持续性。评价路径包括原料阶段、燃料阶段和应用阶段。原料阶段考虑基础设施消耗，燃料阶段考虑因电能消耗嵌套引起的碳排放。评价指标不仅包括能量、化石能源、水资源消耗和温室气体（GHG）排放，而且包括了环境污染雾霾中的主要成分PM10和PM2.5排放。本文为选择航空替代燃料原料来源提供理论和技术基础。结果表明，在原料阶段，生物基温室气体排放相比石油基均降低，其中大豆油碳减排最明显，麻疯树基航空燃料PM10和PM2.5排放量大，因在种植过程中加入大量化肥。在燃料阶段，煤基费托过程能量消耗大，碳排放高。通过全生命周期分析，煤基费托温室气体排放高，大豆温室气体排放较低，其次是能源藻。在不与人争粮，不与粮争地的前提下、不占用耕地的能源藻具有可持续大规模制备航空替代燃料的应用前景。

关 键 词：可持续；原料；全生命周期；航空燃料；替代燃料
中图分类号：V312+.3
文献标识码：A **文章编号**：1001-5965(2016)11-2378-08

随着航空快速发展，我国对航空燃料的需求越来越大，原油进口数量逐年提高，对进口原油的依赖性日益加大。据国家统计局数据显示，“十一五”期间航空煤油需求增长率约为10%，“十二五”期间航空煤油需求增长达到11%，消费量由2000年的460万t上升到2008年的1 276万t，“十三五”期间我国航空业将进一步发展，国内航空煤油需求预计将在9%左右增长。

全球气候变化也给航空运输业提出了新的要求，需要寻求更加环保的航空替代燃料。2009年，国际航空运输协会（IATA）代表整个航空业向国际民航组织（ICAO）提出三大承诺目标：2020年实现碳排放零增长；从2009年到2020年，平均每年燃油效率提高1.5%；2050年碳排放量比2005年减少50%。IATA预测此目标的实现，发动机、飞机轻量化和效率提升可贡献20%碳减排，航空基础设施可实现10%，其他50%碳减排主要依靠生物质航空替代燃料及相关技术。

摆脱化石能源依赖带来的能源危机，降低温室气体（GHG）排放是航空替代燃料的研究动力。无论从军事还是民用的角度看，我国都应该在航空替代燃料的研究方面做出更大的努力，尽力达到航空替代燃料来源的多元化，并完成不同原料来源航空替代燃料的可持续性评价。

燃料从原材料开采到生产，最后到燃烧，整个过程都存在碳排放，都要损耗能源，因此衡量一种燃料可持续与否需要对燃料的整个生命周期进行

收稿日期：2015-10-19；**录用日期**：2015-11-13；**网络出版时间**：2015-12-21 17:53
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20151221.1753.002.html
基金项目：国家国际科技合作专项项目（2013DFA61590）
* **通讯作者**：Tel.：010-82317346 E-mail：yangxiaoyi@buaa.edu.cn

引用格式：赵晶，郭放，阿鲁斯，等. 未来航空燃料原料可持续性研究[J]. 北京航空航天大学学报，2016，42(11)：2378-2385.
ZHAO J，GUO F，A L S，et al. Research of sustainable feedstock for future alternative aviation fuels[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2378-2385 (in Chinese).

北航学报 赠阅

评价,即对燃料摇篮到坟墓(WTW)的全过程进行评价[1]。

本文进行评价使用的计算模型为美国阿贡国家实验室的 2014 年发布的 GREET1_2014 版(Green house gases Regulated Emissions and Energy of Transportation)。GREET 是美国阿贡国家实验室从 1995 年开始开发用于评估替代车用燃料以及先进车辆技术的全生命周期燃料循环的能源效率和温室气体排放的模型。GREET 模型已经被广泛应用,美国能源部、西北太平洋实验室和通用汽车公司都曾使用 GREET 模型进行评价,而且 GREET 模型从 2013 年开始包含了航空替代燃料的评价体系[1]。柴志勇[1]分析了 GREET 模型机理,进行 GREET 模型能耗及排放计算方法及仿真能力研究。GREET 模型计算结果包括能量消耗、化石能源消耗以及温室气体排放等数据,能够满足全生命周期评价需要的数据。

目前,航空替代燃料研究涉及的原料涵盖化石能源和生物质,本文选取煤、天然气、麻疯树、能源藻、大豆、棕榈、油菜籽及亚麻荠 8 种不同原料,运用 GREET1_2014 进行全生命周期的能量消耗、化石能源消耗、水资源消耗、PM10 和 PM2.5 排放和温室气体排放进行计算,并将计算结果与传统石油基航空燃料进行对比分析,为未来可持续航空替代燃料的选取提供依据。

1　航空燃料矿井到油泵过程

1.1　原料来源

我国能源结构具有“富煤、贫油、少气”的特点,在化石能源总量中煤炭占比高达 94.4%,而石油只有 2.4%。2010 年我国石油产量为20 301.4 万 t,消费量为 43 245 万 t;天然气产量为 12.5 亿 m^3,消费量为 130.7 亿 m^3;煤炭产量为 323 500 万 t,消费量为 312 236.5 万 t[2]。在生物质航空替代燃料商业化之前,煤和天然气航空燃料可作为过渡产品。

生物质的选取依据是“不与人争粮,不与粮争地”,应可持续。能源藻的培养不占用耕地,麻疯树可以在沙漠化土地里生长。目前以能源藻和麻疯树为原料的加氢处理制备生物基航空替代燃料最具发展潜力,同时其对淡水的消耗也很少。本文选取了几种已经大规模培养的生物质大豆、棕榈、油菜籽、亚麻荠作为对比。

本文计算中将全生命周期过程分为原料阶段(原料成长(开采)和原料运输)、燃料阶段(粗油提取、粗油精制、燃料运输、储存)和应用阶段(见图 1),关键参数见表 1。原料阶段和燃料阶段属于“矿井到油泵”过程,计算过程中考虑了基础设施建设消耗以及系统内产生电能的循环利用。应用阶段为“油泵到车轮”过程,整个全生命周期过程包括这两个过程,为“矿井到车轮”过程。以传统石油基航空燃料为基准,本文对 8 种不同原料分别计算其原料阶段能量消耗、化石能源消耗及温室气体排放,燃料阶段能量消耗、化石能源消耗及温室气体排放,最后对全生命周期过程能量消耗、物质消耗、水资源消耗、PM10 和 PM2.5 排放及温室气体排放进行分析。计算过程中部分基础数据来源为文献[3-14]。

基础数据主要参考中国统计年鉴 2014 中 2012 年数据,其中:

原油开采能效为 93%[15],进口油占比 58.5%,国产油占比 41.5%。原油开采过程中各基础能源占比:原油 39.39%,燃料油 0.50%,柴油 2.43%,汽油 0.55%,天然气 39.18%,煤 5.16%,电 12.8%[16]。运输方式占比和运输距离:远洋游轮 58.5%,11 000 km;驳船 10%,400 km;管道 78%,1 030 km;铁路 46.8%,950 km[17]。

煤炭开采能效为 97%[15],井下开采占比 86%,地表开采占比 14%[18]。煤炭开采过程中各基础能源占比:燃料油 0.01%,柴油 2.57%,汽油 0.19%,天然气 0.74%,煤 87.73%,电 0.51%[16]。运输方式占比和运输距离:驳船 17%,650 km;铁路 50%,1 000 km;公路 33%,310 km[17]。

电力行业各基础能源占比:燃煤 76.34%,燃油 0.11%,燃气 2.27%,水电 17.16%,核电 1.97%,风电 2.07%,其他 0.08%[19]。电力行业燃煤烟煤占比 90%,褐煤占比 10%;

目前,国内外制备航空燃料的技术主要有油脂加氢技术(加氢法(HRJ)),生物质气化费托合成技术(气化-费托合成法(FT)),生物质水相催化合成技术(水热液化法(HTL))等[20]。9 种原料全生命周期路径:①石油直接精制(refining);②原料煤间接液化费托合成油(FT);③天然气气化费托合成油(FT);④能源藻加氢精制(HRJ);⑤麻疯树加氢精制(HRJ);⑥大豆加氢精制(HRJ);⑦ 棕榈加氢精制(HRJ);⑧油菜籽加氢精制(HRJ);⑨亚麻荠加氢精制(HRJ)。

北航学报 赠阅

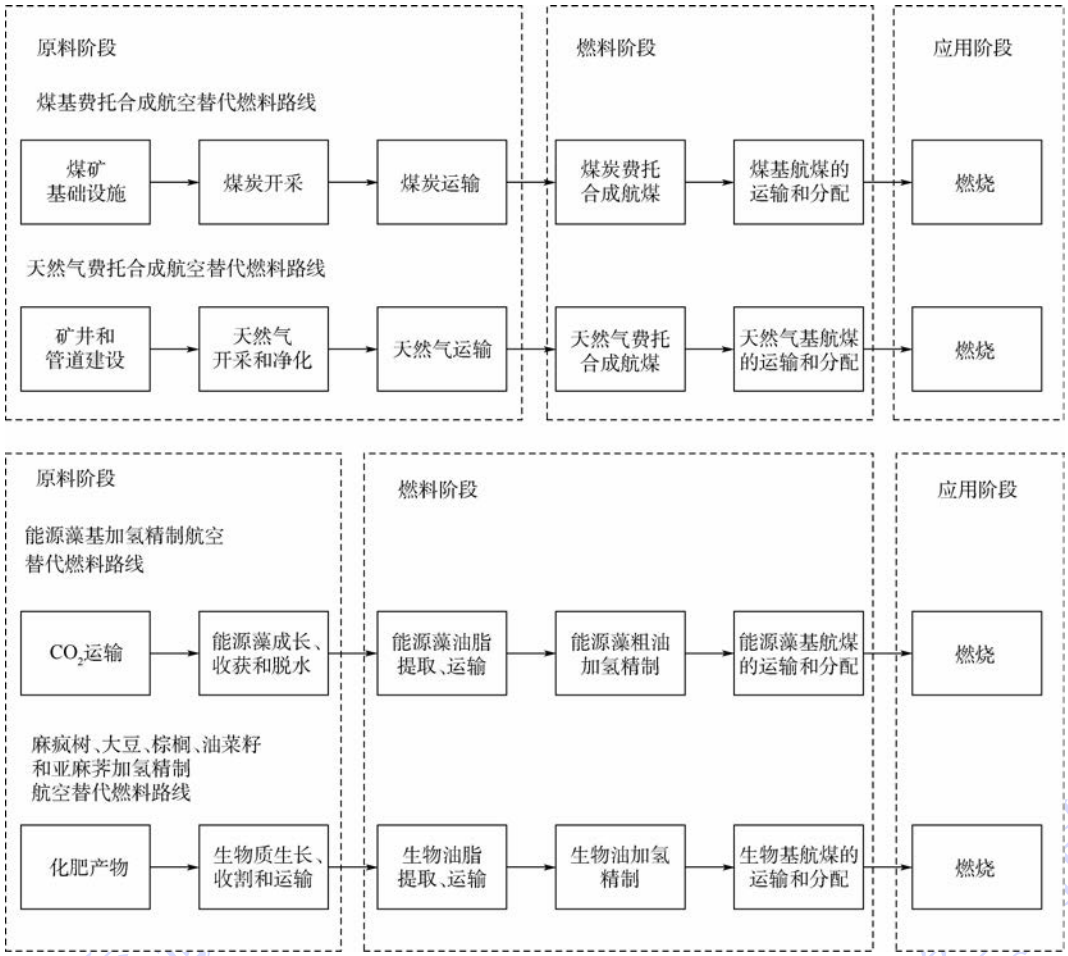


图 1 全生命周期分析路径

Fig. 1 Whole life cycle analysis pathway

表 1 GREET1_2014 中关键参数

Table 1 Key parameters in GREET1_2014

项目	内容
1. 模拟年份	2015
2. 系统中分配方法	按能量含量分配
3. 蒸馏情景	最大化航空燃料
4. 石油	
4.1 是否考虑石油开采过程基础设施	是
5. 煤炭	
5.1 是否考虑煤炭开采过程基础设施	是
6. 天然气	
6.1 是否考虑天然气开采过程基础设施	是
7. 能源藻输入	
7.1 能源藻提油工艺	水热液化
7.2 CO_2 是否回收	是
7.3 是否考虑基础设施建设	是
7.4 共产物计算方法	按能量含量分配
8. 麻疯树、大豆、棕榈、油菜籽、亚麻荠	22.3%
8.1 油脂提取过程分配方法	按能量含量分配
8.2 共产物分配方法	按能量含量分配
9. 生物质油加氢精制过程	
9.1 加氢精制过程分配方法	按能量含量分配

1.2 原料阶段能量消耗、物质消耗及温室气体排放

石油、煤和天然气的原料阶段包括开采阶段和原料运输。开采过程包括基础设施建设的能耗和排放。生物质原料阶段包括原料成长阶段和运输阶段，其中原料成长阶段包括基础设施投入和营养盐需求能耗和排放。原料阶段基于中国能耗基础数据的 GREET 模型计算结果见表 2。结果给出了石油、煤、天然气、能源藻（油脂含量 25%）[3]、麻疯树（油脂含量 37.7%）[5-6]、大豆（油脂含量 21.3%）[4]、棕榈（油脂含量 32.3%）[5-6]、油菜籽（油脂含量 48.4%）[5-6]、亚麻荠（油脂含量 41.9%）[7] 不同原料在原料阶段化石能源消耗和温室气体排放数据。图 2 为原料阶段能量消耗与化石能源消耗。原料阶段包括原料的成长（开采）阶段、原料运输阶段。

从图 2 中可以看出，原料阶段能量消耗（MJ/MJ）从多到少依次为：麻疯树 0.29 > 能源藻 0.27 > 油菜籽 0.23 > 亚麻荠 0.22 > 石油 0.15 > 大豆 0.11 > 天然气 0.10 > 棕榈 0.083 > 煤 0.053。以石油为参照，煤在原料阶段能量消耗最小，仅为石油开采阶段能量消耗 36%。煤、天然气原料阶段能量消耗低，说明其开采成本低于石油开采能耗，大豆、棕榈的能量消耗低主要原因是相较于其他几种生物质，其氮肥和磷肥的输入量较少。

北航学报 赠阅

表 2　原料阶段化石能源消耗、温室气体排放

Table 2　Fossil fuel consumption and GHG emissions in feedstock process

原料	化石能源消耗/(MJ·MJ^{-1})			油含量/%	GHG 排放量/(g CO_2e·MJ^{-1})
	煤	天然气	石油		
石油	0.047 9	0.046 3	0.050 6		14.41
煤	0.040 2	0.001 7	0.010 1		13.56
天然气	0.019 5	0.076 8	0.004 2		10.51
能源藻	0.023 7	0.100 4	0.005 6	25.0	-48.15
麻疯树	0.031 0	0.131 2	0.128 4	35.0	-37.64
大豆	0.017 1	0.026 7	0.061 8	18.5	-59.07
棕榈	0.005 9	0.046 8	0.030 0	22.3	-58.87
油菜籽	0.023 2	0.129 7	0.074 6	44.0	-38.38
亚麻荠	0.019 9	0.114 7	0.080 0	36.0	-45.16

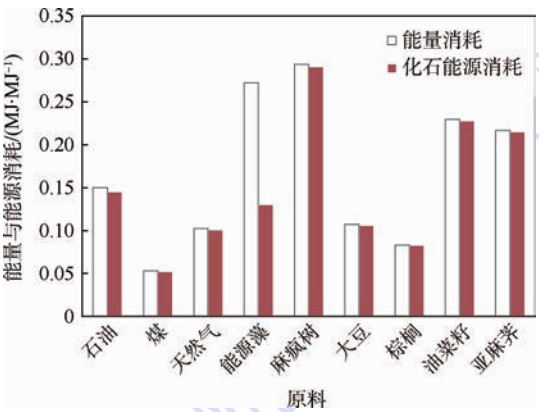


图 2　原料阶段能量消耗与化石能源消耗

Fig. 2　Energy consumption and fossil fuel consumption in feedstock process

对比 9 种原料化石能源消耗与能量消耗,除能源藻外,其他原料其化石能源消耗与其能量消耗大体相同。能源藻生长速率是其他植物的 30 ~ 100倍[21],油脂提取后,藻渣量较大,本研究将其进行发电利用,节约化石能源消耗。同样以石油原料为参照,煤原料阶段化石能源消耗最少,消耗量为 0.052 MJ/MJ,仅为石油原料阶段化石能源消耗的 35.9%;而麻疯树的种植过程中较高的能量消耗和化肥输入量导致其化石能源消耗最多,消耗量为 0.29 MJ/MJ,为石油原料阶段的 2 倍。

对比 9 种原料在原料阶段的温室气体排放,石油、煤、天然气的排放为正值,表明在整个原料阶段这 3 种原料的获取是向大气排放温室气体的。而能源藻、麻疯树、大豆、棕榈、油菜籽和亚麻荠的排放为负值,主要原因是在原料生长阶段生物质吸收空气中的二氧化碳进行成长,达到了碳减排的效果。在整个原料阶段大豆的温室气体排放最低,具有最优减排效果,相对于石油碳减排达到了 510%。碳减排效果最差的生物质为麻疯树,相对于石油碳减排 361%。

1.3　燃料阶段能量消耗、物质消耗及温室气体排放

原料阶段基于中国能耗基础数据的 GREET 模型计算结果见表 3。表 3 列出了石油、煤、天然气、能源藻、麻疯树、大豆、棕榈、油菜籽、亚麻荠燃料阶段化石能源消耗及温室气体排放。石油采用精制工艺制取航空燃料,煤和天然气采用费托合成工艺制取航空燃料。石油基精制过程的能量效率为 95.3%,天然气费托合成过程的能量效率为 64%,煤费托合成过程的能量效率为 50%。能源藻、麻疯树、大豆、棕榈、油菜籽和亚麻荠采用加氢精制制备工艺制取航空燃料。能源藻油的航空燃料质量产油率为 72.4%,能量产油率 54.2%;麻疯树的质量产油率为 71.4%,能量产油率为 61.9%;大豆油的航空燃料质量产油率为71.9%,能量产油率为 35.8%;棕榈油的航空燃料质量产油率为 71.4%,能量产油率为 79.5%;油菜籽油的航空燃料产油率为 73.5%,能量产油率为58.6%;亚麻荠油的航空燃料质量产油率为 72.4%,能量产油率为 54.2%。其中质量产油率为 1 kg 原料油产出航空燃料重量。数据来自文献[3-14]。能量产油率为航空燃料产物的能量与所有产物能量总和的比值。按照能量产油率的数值计算过程中的能量消耗和碳排放。

表 3　燃料阶段化石能源消耗、温室气体排放

Table 3　Fossil fuel consumption and GHG emissions in fuel process

燃料阶段	化石能源消耗/(MJ·MJ^{-1})			GHG 排放量/(g CO_2e·MJ^{-1})
	煤	天然气	石油	
石油精制	0.006 7	0.034 5	0.024 2	4.13
煤费托合成	1.046 6	0.002 7	0.015 6	114.71
天然气费托合成	0.022 5	0.606 6	0.006 3	24.84
能源藻加氢精制	0.023 6	0.339 7	0.013 5	11.63
麻疯树加氢精制	0.018 0	0.283 3	0.017 4	22.66
大豆加氢精制	0.067 5	0.212 4	0.014 8	20.76
棕榈加氢精制	0.027 2	0.303 7	0.052 9	25.54
油菜籽加氢精制	0.035 3	0.288 2	0.016 0	22.34
亚麻荠加氢精制	0.020 9	0.261 2	0.024 7	19.96

图 3 中对比了不同原料在燃料阶段能量消耗与化石能源消耗。石油因其炼制工艺成熟,在制备阶段石油能量消耗和化石能源消耗最少,消耗量仅为 0.066 MJ/MJ 和 0.065 MJ/MJ。亚麻荠加氢精制航空替代燃料制备阶段能量消耗最高,达到 1.11 MJ/MJ,是石油燃料阶段 16.8 倍。煤基费托合成工艺因其耗能,能量消耗仅低于亚麻荠为 1.07 MJ/MJ。燃料阶段能量消耗(MJ/MJ)从多到少依次为:亚麻荠 1.11 > 煤 1.07 > 大豆 0.83 > 能源藻 0.75 > 天然气 0.64 > 油菜籽

北航学报 赠阅

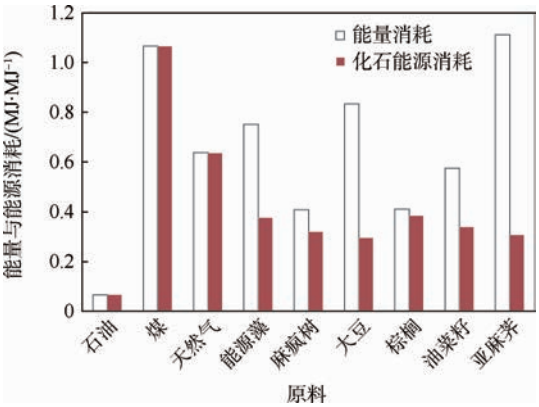


图 3 不同原料燃料阶段能量消耗与化石能源消耗

Fig. 3 Energy consumption and fossil fuel consumption of different feedstocks in fuel process

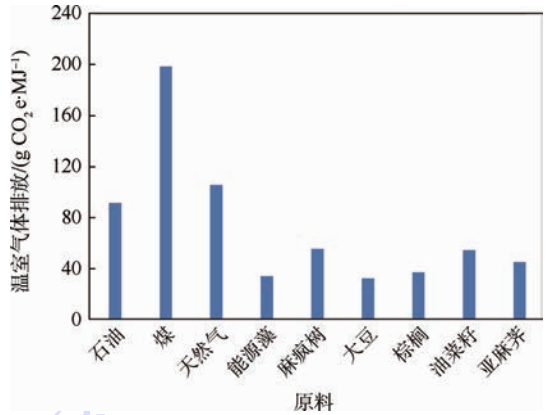


图 4 航空燃料全生命周期过程温室气体排放

Fig. 4 GHG emissions of aviation fuels during their whole life-cycles

0.57 > 棕榈 0.41 > 麻疯树 0.408 > 石油 0.066。

同样以石油基航空燃料为参照,对比 8 种原料化石能源消耗与能量消耗,假设加氢工艺中的氢主要来源为天然气。其中,煤费托合成航空替代燃料阶段化石能源消耗为石油基航空燃料制备阶段的 16.3 倍。燃料阶段化石能源消耗从多到少依次为:煤 > 天然气 > 棕榈 > 能源藻 > 油菜籽 > 麻疯树 > 亚麻荠 > 大豆 > 石油。

对比 9 种原料在燃料阶段的温室气体排放,石油基排放最低,因为在燃料过程中石油精制工艺能耗最低,其他原料燃料阶段的碳排放相对于石油基航空燃料均有不同程度的增加,其中煤基费托合成因为其较低的能量效率,碳排放最高为 114.71 g CO_2e/MJ,是石油基的 27.8 倍。大豆加氢精制工艺碳排放是石油基的 5 倍。

2 航空替代燃料全生命周期评价

2.1 温室气体排放

全生命周期分析假设航空燃料最终用于单通道飞机。图 4 为不同原料制备航空替代燃料并最终应用到飞机过程的全生命周期过程中温室气体排放的计算结果。以石油基航空燃料为基准,其中煤费托合成航空替代燃料全生命周期过程排放最高,排放量为 198.75 g CO_2e/MJ,是石油基航空燃料排放量的 2.17 倍。大豆加氢精制航空替代燃料全生命周期过程排放最少,排放量为 32.17 g CO_2e/MJ,仅为石油基航空燃料的 35%,具有最优碳减排效果,能源藻加氢精制排放为 34.0 g CO_2e/MJ,仅为石油基航空燃料的 37%。不同原料制备航空替代燃料全生命周期过程温室气体排放从高到低依次为:煤 > 天然气 > 石油 > 麻疯树 > 油菜籽 > 亚麻荠 > 棕榈 > 能源藻 > 大豆。

2.2 能量消耗和化石能源消耗

图 5 对比了不同原料制备航空替代燃料全生命周期过程能量消耗与化石能源消耗的计算结果。能量消耗主要包含全生命周期过程中能量输入和非化石能源的物质输入(化肥、除草剂、杀虫剂),非化石能源物质输入将物质全生命周期的能量消耗和碳排放嵌套到燃料的全生命周期计算过程中进行加和。以石油基航空燃料为参照进行对比,其他 8 种航空替代燃料能量消耗均高于传统石油,其中亚麻荠加氢精制航空替代燃料全生命周期过程中能量总消耗为最多,消耗量为 2.3 MJ/MJ,是石油基航空燃料 1.9 倍。不同原料制备航空替代燃料全生命周期过程中能量总消耗从多到少依次为:亚麻荠 > 煤 > 能源藻 > 大豆 > 油菜籽 > 天然气 > 麻疯树 > 棕榈 > 石油。在全生命周期总能量消耗方面,由于石油基航空燃料制造工艺技术已经非常成熟,具有较低能量消耗。因此,除石油外的 8 种原料在制造工艺技术上具有较大改进空间。

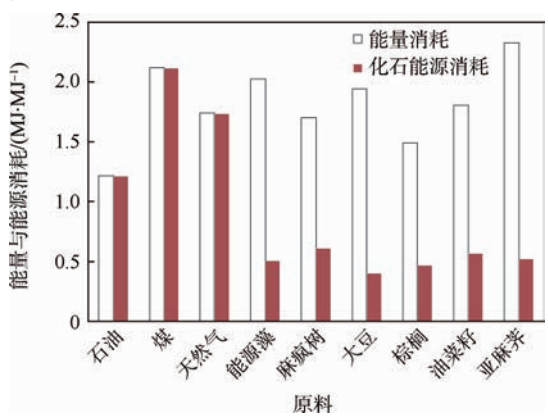


图 5 航空燃料全生命周期过程能量消耗与化石能源消耗

Fig. 5 Energy consumption and fossil fuel consumption of aviation fuels during their whole life cycles

北航学报 赠阅

以石油基航空燃料为参照,煤基和天然气基航空替代燃料全生命周期过程化石能源消耗均高于石油基,其中煤费托合成航空替代燃料化石能源消耗最多,消耗量为 2.12 MJ/MJ,是石油基的 1.75 倍;大豆加氢精制航空替代燃料过程中化石能源消耗最少,消耗量为 0.40 MJ/MJ,是石油基的 33.1%。不同原料制备航空替代燃料全生命周期过程化石能源消耗从多到少依次为:煤 > 天然气 > 石油 > 麻疯树 > 油菜籽 > 亚麻荠 > 能源藻 > 棕榈 > 大豆。其中原料煤、天然气、石油化石能源消耗量高,主要原因源于原料来源于化石能源。在化石能源消耗方面,生物质航空替代燃料均低于传统航空燃料,具有明显优势。

2.3　水资源消耗

基于中国能耗基础数据的 GREET 模型航空燃料全生命周期过程水资源消耗计算结果见表 4。表中对不同原料制备航空替代燃料全生命周期过程中水资源消耗进行了对比。石油基航空燃料制备过程耗水量为 81.04 gal/MJ (1 gal = 3.79 L),而其他 8 种原料制备航空替代燃料耗水量远远低于传统石油,仅为传统石油基航空燃料耗水量的 0.02% ~ 1.50%。因此在节约水资源方面,其他 8 种原料相比传统石油具有突出的优势。

表 4　航空燃料全生命周期过程水资源消耗

Table 4　Water consumption of aviation fuels during their whole life cycles

全生命周期阶段	耗水量/(gal · MJ^{-1})
石油精制	81.04
煤费托合成	0.02
天然气费托合成	0.01
能源藻加氢精制	0.03
麻疯树加氢精制	0.05
大豆加氢精制	1.22
棕榈加氢精制	0.03
油菜籽加氢精制	0.05
亚麻荠加氢精制	0.04

2.4　PM10 和 PM2.5 排放

PM10 和 PM2.5 是当今中国环境污染雾霾中的主要成分,分别代表可吸入颗粒物指空气动力学当量直径小于等于 10 μm 的颗粒物和小于等于 2.5 μm 的颗粒物。定量研究其排放对改善雾霾污染具有重要意义。图 6 为基于中国能耗基础数据的 GREET 模型不同原料来源全生命周期过程 PM10 和 PM2.5 排放情况计算结果。图中 PM10 排放量均大于 PM2.5 是因为 PM10 排放包含 PM2.5 的排放。

以石油基航空燃料为参照进行对比,其全生命周期过程中 PM10 排放量为 0.008 8 g/MJ,PM2.5 排放量为 0.007 6 g/MJ。麻疯树加氢精制航空替代燃料全生命周期过程中 PM10 和 PM2.5 排放量最大,因在种植过程中加入大量化肥,其分别为石油基的 17 倍和 6.8 倍。煤基费托合成航空替代燃料全生命周期过程中 PM2.5 排放量最低,为石油基航空燃料排放量的 35.7%,PM10 排放量为石油基航空燃料排放量的 75.3%。能源藻基制备航空替代燃料全生命周期过程中 PM10 和 PM2.5 排放量分别为石油基航空燃料的 71.4% 和 74.7%,大豆基为 70% 和 67.5%。因此利用生物原料制备航空替代燃料在大气污染防治上具有优势。

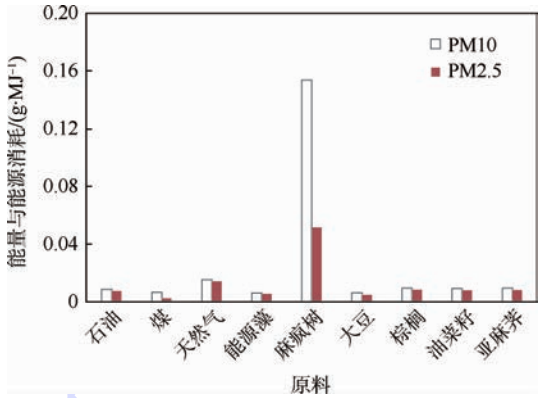


图 6　航空燃料全生命周期过程 PM10 和 PM2.5 排放

Fig. 6　PM10 and PM2.5 emissions of aviation fuels during their whole life cycles

3　结　论

本文运用生命周期评价方法,以传统石油基航空燃料为基准,分析未来 8 种航空燃料原料(煤、天然气、能源藻、麻疯树、大豆、棕榈、油菜籽、亚麻荠)的可持续性。生长(开采)、制备阶段的评价指标为能量、化石能源消耗,温室气体排放。全生命周期阶段的评价指标为能量、化石能源、水资源消耗、PM10 和 PM2.5 排放以及温室气体排放。为选择未来航空替代燃料原料来源提供理论和技术基础。

1)原料阶段能量生物基航空替代燃料的温室气体排放均为负值,主要是基于生物质在生长过程中的二氧化碳吸收。

2)燃料阶段,生物质的能量消耗远高于化石能源消耗,消耗的部分能量来自于生物质生长过程中光合作用吸收的太阳能,这一部分能量转化为电能在系统内循环后降低了化石能源的消耗,使得生物质航空替代燃料的化石能源消耗和温室

北航学报 赠阅

气体排放远低于化石原料制备的航空替代燃料。

3）全生命周期阶段，相对化石来源制备的航空燃料，生物质航空替代燃料的化石能源消耗和温室气体排放均有不同程度的降低。其中能源藻和大豆的化石能源消耗和温室气体排放最少。

现阶段航空燃料面对的是能源安全和环境问题，从能源安全出发，我国有着煤多油少的储量现状，需要采用煤作为航空替代燃料的原料来源保证能源安全。从保护环境出发，大豆油和能源藻全生命周期温室气体排放最少，可以做为航空替代燃料的原料来源。作为食品油的大豆油大规模生产航空替代燃料时，必然会造成航空替代燃料与粮争地、与人争粮的尴尬状态，不占用耕地的能源藻具有可持续大规模制备航空替代燃料的应用前景。

参考文献（References）

[1] 柴志勇. 基于 GREET 模型的车用生物质燃料能耗及排放研究[D]. 长春：吉林大学，2007：23-30.
CHAI Z Y. Research on vehicular biofuels energy consumption and emission based on GREET model[D]. Changchun: Jilin University, 2007: 23-30(in Chinese).

[2] 赵光辉，姜伟，牛欣宇，等. 航空生物燃料制备技术及应用前景[J]. 中外能源，2014，19(8)：30-34.
ZHAO G H, JIANG W, NIU X Y, et al. Jet fuel preparation technology and application prospects[J]. Sino-Global Energy, 2014, 19(8): 30-34(in Chinese).

[3] FRANK E, HAN J, PALOU-RIVERA I, et al. Life-cycle analysis of algal lipid fuels with the GREET model: ANL/ESD/11-5[R]. Chicago: Center for Transportation Research, Argonne National Laboratory, 2011: 16-39.

[4] HUO H, WANG M, BLOYD C, et al. Life cycle assessment of energy and greenhouse gas effects of soybean-derived biodiesel and renewable fuels: ANL/ESD/08-2[R]. Chicago: Center for Transportation Research, Argonne National Laboratory, 2008: 13-33.

[5] STRATTON R W, WONG H M, HILEMAN J I. Life cycle greenhouse gas emissions from alternative jet fuels: LAE-2010-007-R/PARTNER-COE-2010-001[R]. Cambridge: Massachusetts Institute of Technology, 2010: 4-107.

[6] STRATTON R W, WONG H M, HILEMAN J I. Quantifying variability in life cycle greenhouse gas inventories of alternative middle distillate transportation fuels[J]. Environmental Science and Technology, 2011, 45(10): 4637-4644.

[7] SHONNARD D R, WILLIAMS L, KALNES T N, et al. Camelina-derived jet fuel and diesel: Sustainable advanced biofuels[J]. Environmental Progress and Sustainable Energy, 2010, 29(3): 382-392.

[8] XIE X, WANG M, HAN J. Assessment of fuel-cycle energy use and greenhouse gas emissions for Fischer-Tropsch diesel from coal and cellulosic biomass[J]. Environmental Science and Technology, 2011, 45(7): 3047-3053.

[9] BURNHAM A, HAN J, CLARK C E, et al. Life-cycle greenhouse gas emissions of shale gas, natural gas, coal and petroleum[J]. Environmental Science and Technology, 2012, 46(2): 619-627.

[10] PRADHAN A, SHRESTHA D S, MCALOON A, et al. Energy life-cycle assessment of soybean biodiesel revisited[J]. Transactions of the ASABE, 2011, 54(3): 1031-1039.

[11] FRANK E, HAN J, PALOU-RIVERA I, et al. Methane and nitrous oxide emissions affect the life-cycle analysis of algal biofuels[J]. Environmental Research Letters, 2012, 7(1): 1-10.

[12] VASUDEVAN V, STRATTON R W, PEARLSON M N, et al. The environmental performance of algal biofuel technology options[J]. Environmental Science and Technology, 2012, 46(4): 2451-2459.

[13] PEARLSON M. A techno-economic and environmental assessment of hydroprocessed renewable distillate fuels[D]. Cambridge: Massachusetts Institute of Technology, 2011: 17-40.

[14] WANG M H, HUANG A. Full-fuel-cycle analysis of energy and emissions impacts of transportation fuels produced from natural gas: ANL/ESD/-40[R]. Chicago: Center for Transportation Research, Argonne National Laboratory, 1999: 8-37.

[15] OU X M, ZHANG X L, CHANG S Y. Alternative fuel buses currently in use in China: Life-cycle fossil energy use, GHG emissions and policy recommendations[J]. Energy Policy, 2010, 38(1): 406-418.

[16] 中华人民共和国国家统计局. 中国统计年鉴 2014[M]. 北京：中国统计出版社，2012：263-267.
National Bureau of Statistic of the People's Republic of China. China statistical yearbook 2014[M]. Beijing: China Statistics Press, 2012: 263-267(in Chinese).

[17] 张庭婷，谢晓敏，黄震. 微藻生物柴油全生命周期分析[J]. 上海交通大学学报，2014，48(6)：750-755.
ZHANG T T, XIE X M, HUANG Z. Life cycle analysis of microalgal biodiesel in China[J]. Journal of Shanghai Jiao Tong University, 2014, 48(6): 750-755(in Chinese).

[18] 田会，才庆祥，甄选. 中国露天采煤事业的发展展望[J]. 煤炭工程，2014，46(10)：11-14.
TIAN H, CAI Q X, ZHEN X. Development prospects of surface coal mining industry in China[J]. Coal Engineering, 2014, 46(10): 11-14(in Chinese).

[19] 中国电力企业联合会. 2012 年电力统计基本数据一览表[EB/OL]. 北京：中国电力企业联合会，2013-11-07[2015-07-19]. http://www.cec.org.cn/guihuayutongji/tongjxinxi/niandushuju/2013-11-07/111737.html.
China Electricity Council. Power statistics of basic data list 2012[EB/OL]. Beijing: China Electricity Council, 2013-11-07[2015-07-19]. http://www.cec.org.cn/guihuayutongji/tongjxinxi/niandushuju/2013-11-07/111737.html(in Chinese).

[20] PACHAURI R K, REISINGER A. IPCC fourth assessment report: Climate Change2007[R]. Geneva: IPCC, 2007.

[21] YANG X Y, WANG X, ZHAO B W, et al. Simulation model of pyrolysis biofuel yield based on algal components and pyrolysis kinetics[J]. Bioenergy Research, 2014, 7(4): 1293-1304.

北航学报 赠阅

作者简介:

赵晶　女,博士,助理研究员。主要研究方向:航空替代燃料。
Tel.:010-82316682
E-mail:zhao_jing@buaa.edu.cn

杨晓奕　女,博士,教授。主要研究方向:航空替代燃料。
Tel.:010-82317346
E-mail:yangxiaoyi@buaa.edu.cn

Research of sustainable feedstock for future alternative aviation fuels

ZHAO Jing, GUO Fang, A Lusi, YANG Xiaoyi*

(Energy and Environment International Centre, School of Energy and Power Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: Based on the whole life-cycle analysis, the feedstock sustainability has been evaluated for alternative aviation fuels. Eight sources, including coal, natural gas, algae, jatropha, soybean, palm, rapeseed and camelina, have been compared with petroleum-based aviation fuel in energy consumption, fossil fuel consumption, water consumption, greenhouse gas (GHG) emissions, and PM10 and PM2.5 emissions. The whole life cycle concluded feedstock production, fuel refining and combustion application. In the first stage of feedstock production, energy and emissions comprise the influence of infrastructure construction. The fuel refining process considered the emissions caused by nested electricity utilization. This paper provides theoretical and data support to assess the sustainable feedstock for alternative aviation fuel. The results indicate that compared with petroleum-based one, the GHG emissions of biomass all decrease in feedstock production stage, and soybean conducts the minimum GHG emission. Jatropha-based fuel gives the highest PM10 and PM2.5 emissions as the result of high fertilizer input. During the fuel refining stage, coal-based Fischer-Tropsch aviation fuel performs the highest energy consumption and GHG emission. Among eight feedstocks, coal-based Fischer-Tropsch aviation fuel shows the highest GHG emission in the whole life cycle, while algae-based aviation fuel shows the lowest GHG emission and followed by soybean. Owing to non-arable land and non-competition with food, algae are very promising as the sustainable feedstock for alternative aviation fuel in the future.

Key words: sustainable; sources; whole life cycle; aviation fuel; alternative fuel

Received: 2015-10-19; **Accepted**: 2015-11-13; **Published online**: 2015-12-21 17:53
URL: www.cnki.net/kcms/detail/11.2625.V.20151221.1753.002.html
Foundation item: Program of International Science and Technology Cooperation Program of China (2013DFA61590)
* **Corresponding author**. Tel.:010-82317346　E-mail:yangxiaoyi@buaa.edu.cn

北京航空航天大学学报

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0690

回转体倾斜入水空泡及弹道特性实验

宋武超，王聪*，魏英杰，许昊

（哈尔滨工业大学 航天学院，哈尔滨 150001）

摘 要：针对回转体低速倾斜入水过程空泡的生成机理和演化特性开展研究。通过不同入水速度和入水角度回转体倾斜入水对比实验，研究了入水速度和入水角度对入水空泡、回转体速度、俯仰角及阻力系数的影响规律。研究结果表明：空泡深闭合时刻的空泡长度、最大空泡直径和空泡闭合点深度随着回转体入水速度的增加而增大；空泡深闭合发生时间随着入水速度的增加而小幅增大；同一入水时刻回转体俯仰角和俯仰角速度均随入水速度的增加而减小；随着回转体入水角度的增加，同一时刻入水空泡直径和回转体速度逐渐减小，阻力系数逐渐增大。

关 键 词：回转体；运动参数；倾斜入水；入水空泡；实验研究
中图分类号：O351
文献标识码：A **文章编号**：1001-5965(2016)11-2386-09

运动体入水问题广泛存在于水上飞机和宇宙飞船的水上着落、船体在风浪中的砰击、空投鱼雷入水、跳水运动等领域，具有较强的工程应用背景[1]。从空气介质穿过气水交界面侵入水体的入水过程，涉及到气体、液体和固体3种介质的耦合作用，具有高度非线性和非定常性。

Worthington 和 Cole[2-3] 利用闪光摄影技术对球体入水进行了大量实验，分析了入水速度和球体表面条件对入水喷溅的影响。May[4] 研究了环境压力和入水速度对球体垂直入水空泡形成和发展的影响规律。Tabuteau 等[5] 针对球体垂直入水过程中入水速度对空泡直径及入水喷溅高度的影响规律进行了研究。Charters[6] 针对钢质球体开展了高速垂直入水实验，得到了球体在不同入水速度下的阻力系数。Abelson[7] 针对锥头圆柱体倾斜入水过程中入水角度对空泡内压降的影响规律进行了研究。施红辉等[8-10] 研究了细长体入水角度和入水速度等因素对入水后自由液面的稳定性及空泡形态的影响规律。何春涛等[11] 研究了圆柱体倾斜入水条件下空泡闭合方式与入水速度的关系。杨衡等[12] 开展了弹体低速入水空泡实验研究，分析了入水速度和入水角度对空泡形态及入水喷溅的影响规律。胡青青[13] 针对锥头圆柱体开展入水实验研究，分析了入水速度和入水角度等对圆柱体入水后姿态稳定性及空泡形态的影响规律。梅哲力[14] 研究了入水角度对细长圆柱体入水过程中空泡形态及阻力系数的影响规律。马庆鹏[15] 和陈晨[16] 等针对圆柱体高速垂直入水问题进行数值模拟研究，分析了圆柱体入水速度及空气域压力对空泡形态发展及流场压力分布的影响。

目前国内对入水问题的实验研究较多关注于垂直入水问题，针对入水空泡、入水弹道等发展变化过程往往做定性分析，较少涉及到倾斜入水过

收稿日期：2015-10-23；**录用日期**：2016-05-13；**网络出版时间**：2016-05-17 11:11
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160517.1111.003.html
基金项目：黑龙江省自然科学基金（A201409）；中央高校基本科研业务费专项资金（HIT.NSRIF.201159）；哈尔滨市科技创新人才研究专项资金（2013RFLXJ007）
* **通讯作者**：Tel.：0451-86413510 E-mail：alanwang@hit.edu.cn

引用格式：宋武超，王聪，魏英杰，等. 回转体倾斜入水空泡及弹道特性试验[J]. 北京航空航天大学学报，2016，42(11)：2386-2394. SONG W C，WANG C，WEI Y J，et al. Experiment of cavity and trajectory characteristics of oblique water entry of revolution bodies[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2386-2394 (in Chinese).

程中入水速度、入水角度对空泡形态、运动体运动特性及流体动力特性的影响规律。本文通过回转体低速倾斜入水实验,研究在倾斜入水情况下,入水空泡的生成机理和演化特性;开展对比实验定量分析入水速度和入水角度对入水空泡形成、发展特性的影响,以及回转体运动特性和阻力系数的变化规律。

1　实验系统与模型参数

图 1 为倾斜入水实验系统示意图。实验水箱四周壁和底部材质为 9 mm 钢化玻璃,结构尺寸为 1.5 m × 0.8 m × 0.9 m,并于水箱底部放置一层聚氨酯减振板作为防护层。滑轨采用直角断面铝合金,倾斜角度 β 调整范围为 30° ~ 70°。

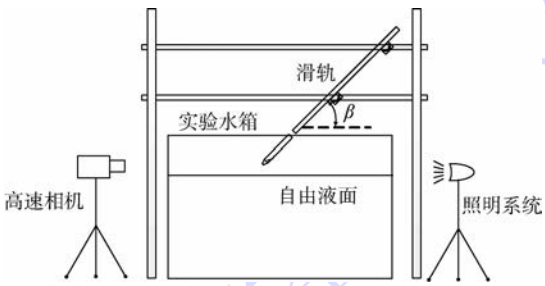


图 1　倾斜入水实验系统示意图

Fig. 1　Schematic of oblique water entry experiment system

本文实验过程采用 FASTCAMSA5 高速照相机进行拍摄,拍摄帧率为 1 000 帧/s,曝光时间设定为 1/100 000 s。照明系统采用 8 × 1 000 W 光源点阵组成的平行光源,并采用磨砂玻璃对其进行处理,以形成良好的成像效果。实验模型材料采用普通钢,密度 $\rho = 7.85\ \mathrm{g/cm^3}$。如图2所示,直径 $D_0 = 9$ mm,柱段长度 $L = 40$ mm,头型为 $\theta = 140°$锥角头型。

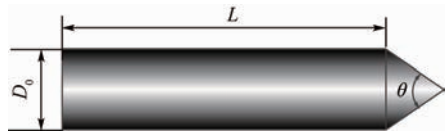


图 2　入水实验模型

Fig. 2　Water entry experiment model

2　实验结果分析

2.1　回转体倾斜入水空泡形态分析

回转体入水过程中经历入水撞击、流动形成、开空泡和空泡闭合 4 个阶段,且各阶段历时非常短暂,具有极强的瞬时性和非定常性[11]。在入水撞击阶段,回转体在极短的时间内经历了由空气介质到水介质的突变,导致了一系列复杂的流动现象,回转体在该阶段受到了巨大的冲击力。在流动形成阶段(图 3 中入水时间 $t = 1$ ms),处于被流体质点包裹状态的回转体,其自身的动能通过阻力的形式传递给附近流域的流体质点,使原来处于静止状态的流体质点开始运动;当回转体壁面附近的流体质点获得足够大的速度后,则与其分离,形成入水空泡,同时回转体在该阶段将经受入水期间最大的加速度。在开空泡阶段($t = 10$ ~ 50 ms),回转体在空泡的包裹下继续运动,且空泡径向和轴向尺寸逐渐增大,空泡尾部与大气处于连通状态。在空泡闭合阶段($t = 50$ ~ 60 ms),空泡壁面向空泡轴线处收缩,最终在自由液面或者自由液面以下某一深度收缩至一点,空泡完全闭合。在自由液面附近发生的闭合称之为表面闭

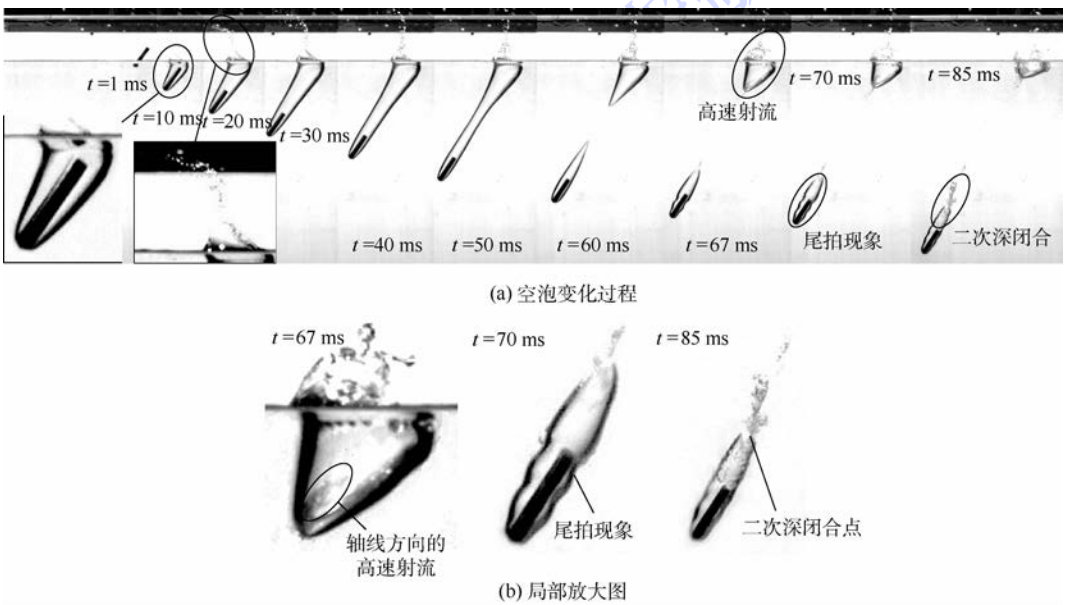


图 3　倾斜入水空泡变化过程

Fig. 3　Variation process of oblique water entry cavity

北航学报 赠阅

合,在液面以下某点发生的闭合称之为深闭合。一般来说,当入水速度较高时,表面闭合现象会发生;当入水速度较低时,则出现深闭合现象。实验研究表明,回转体入水速度越高,其空泡表面闭合发生的时间就越快[17]。

如图 3 所示,回转体以一定的角度倾斜入水后,在原来静止的水域形成一个非对称入水空泡,且空泡背水面一侧半径明显小于迎水面一侧半径。随着入水深度的增加,空泡逐渐被拉长;当回转体到达一定入水深度时,环境压力处于主导地位,空泡中部逐渐收缩至一点,被隔断成两个独立的空泡,即发生深闭合现象(t = 60 ms);深闭合发生后,闭合点附近的入水空泡在重力的作用下发生一定的上漂;从图 3(b)的局部放大图(t = 67 ms)中可以看出,由于空泡壁面在闭合点的瞬时撞击作用,闭合点处产生了沿空泡轴线方向的高速射流;随着入水时间增加,下部空泡被继续压缩,从图 3(b)(t = 70 ms)中可以看出,回转体尾部与空泡壁面发生碰撞,出现尾拍现象,入水空泡产生一定的扰动。在深闭合发生之前,空泡的非对称性随着入水深度的增加而逐渐减弱,且同一入水过程中,空泡一般会发生多次深闭合现象[18]。

回转体入水过程中,自由液面处入水点附近出现明显的不对称喷溅现象(t = 10 ms),该喷溅主要沿回转体水平速度分量方向运动,且背水面一侧处的喷溅较为强烈。这是因为在倾斜入水过程中,回转体下部的近水壁面首先撞击水面,且由于存在显著的水平方向速度分量,撞击点附近的流体质点获得较大的水平方向的动量,因而入水过程中产生沿水平速度分量方向的喷溅。随着时间的推移,喷溅先是向回转体运动方向的斜上方迅速运动,当向外扩张速度减小到零时开始向中间收缩,最后在回转体头部撞击自由液面处的斜上方汇聚,形成一个不规则的锥形体(t = 30 ms)。从图 3 中可以看出,在整个入水过程中,喷溅现象始终存在;根据高速摄像机所拍摄的喷溅流动过程发现,入水过程中喷溅速度高于回转体入水速度[11]。

2.2 入水速度对倾斜入水过程影响

为研究倾斜入水情况下,入水速度对空泡形态和闭合时间及回转体运动特性的影响规律,开展不同入水速度回转体低速倾斜入水实验。实验中滑轨与水平面之间夹角 β = 55°。

2.2.1 入水速度对回转体入水空泡形态的影响

图4 给出了3 种不同入水速度 V_0 回转体低

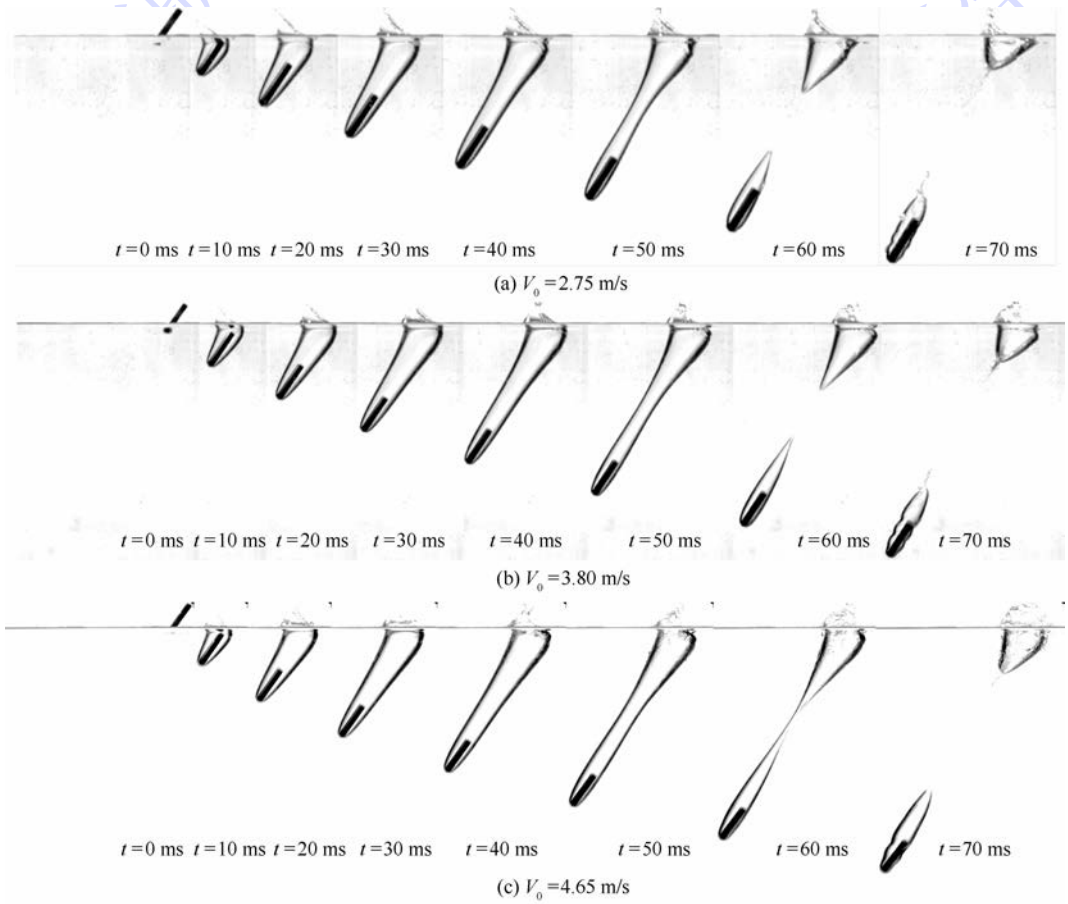


图 4 不同入水速度回转体倾斜入水空泡形态变化特性

Fig. 4 Morphologic variation characteristics of oblique water entry cavity of revolution bodies with various water entry velocity

速倾斜入水过程空泡形态变化特性。取回转体头部触水时刻为零时刻。

从图 4 可以看出,回转体以不同速度入水,所产生的入水空泡有较大差异。在入水前期,同一入水时刻最大空泡半径随着回转体入水速度的增加而增大;空泡发生深闭合之后,回转体尾部与空泡壁面碰撞导致空泡产生波动现象,其波长随着入水速度的增加而增大,波动频率随着入水速度的增加而减小。

图 5 给出了 3 种入水速度回转体倾斜入水过程中,空泡发生第一次深闭合瞬间的空泡形态。图 6 和图 7 给出了 6 种入水速度回转体入水空泡发生深闭合瞬时的空泡长度、最大空泡直径及闭合点深度和深闭合发生时间。从图中可以看出,深闭合时刻入水空泡长度、最大空泡直径和闭合点深度均随着入水速度的增加而增大;空泡深闭合发生时间随着入水速度的增加而小幅增大,在本文实验速度范围内,最大差值仅为2.6 ms。

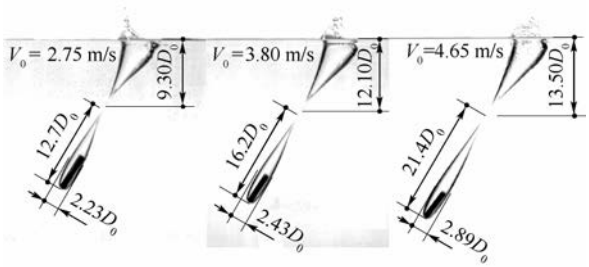


图 5　不同入水速度空泡深闭合

Fig. 5　Deep closure of cavity with various water entry velocity

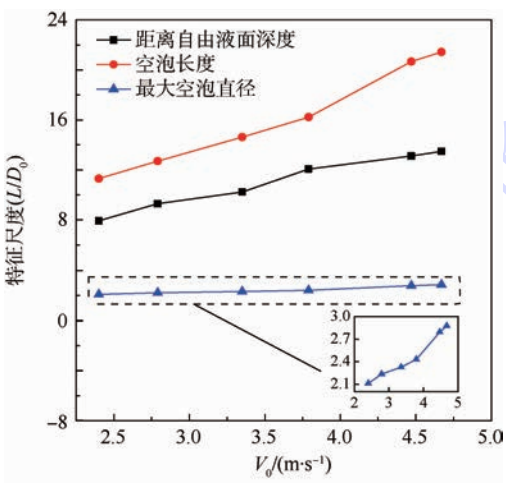


图 6　不同入水速度空泡深闭合瞬间特征尺寸

Fig. 6　Instant feature size of cavity at deep closure with various water entry velocity

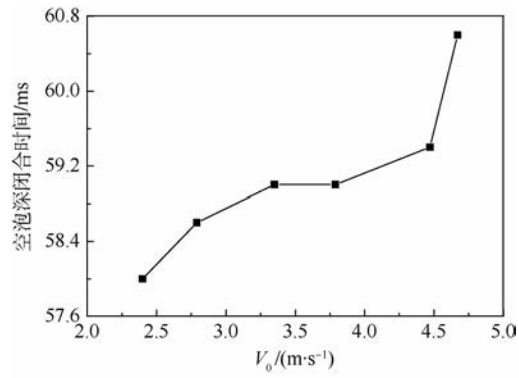


图 7　空泡深闭合时间与入水速度的关系

Fig. 7　Relationship between cavity deep closure time and water entry velocity

2.2.2　入水速度对回转体运动特性的影响

回转体入水过程中姿态的描述采用导弹飞行力学中的地面坐标系,其定义如图 8 所示。坐标原点定义在回转体撞击自由液面处;俯仰角 α 定义为回转体轴线与地面坐标系 x 轴的夹角;V 为回转体速度。

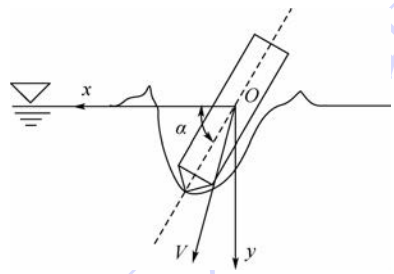


图 8　坐标系定义

Fig. 8　Definition of coordinate system

根据高速摄像机拍摄帧率和回转体相对于自由液面的位置,可以得到回转体入水过程中位移和俯仰角变化曲线。对该曲线进行五阶多项式拟合[19],并对拟合后的曲线取一阶和二阶导数,即可得到回转体入水过程中速度、俯仰角速度和加速度曲线,进而根据阻力系数的定义可推导得到阻力系数曲线。

图 9 为不同入水速度条件下回转体倾斜入水后速度与加速度变化规律对比。本文中入水时间取到空泡发生深闭合后为止,不同入水条件下入水时间不尽相同。从图中可以看出,当入水速度较大时,回转体在入水过程中速度呈减小规律;当入水速度较小时,其速度呈先增大后减小规律。对比图 9(b)中加速度变化过程可以发现,当入水速度较大时,回转体撞击自由液面后受到的流体阻力占据主要地位,使得回转体产生与速度方向相反的加速度,运动速度呈减小趋势。当入水速

北航学报 赠阅

度较小时,在速度方向上回转体所受到的阻力小于其本身的重力分量,因而在入水初期做加速运动;随着回转体速度的增加,其所受的阻力亦随之增大,使得回转体加速度减小,但仍为正值,速度增大趋势减缓;当阻力与速度方向上的重力分量相平衡时,回转体加速度减小到零,速度增加到最大;随着入水时间的继续增加,运动过程中速度方向上阻力大于重力分量,加速度变为负值,回转体做减速运动。此外,由于相机拍摄帧率的限制,致使未能准确捕捉回转体撞击水面阶段的数据信息,导致加速度曲线在入水瞬间有一定的不连续性。

图 10 为倾斜入水过程中不同入水速度回转体俯仰角和角速度随时间变化规律对比。从图中可以看出,当入水速度较大时,回转体俯仰角呈减小趋势;当入水速度较小时,俯仰角呈增大趋势。对比图 10(b)可以看出,在撞击自由液面时刻附近,不同入水速度回转体的角速度差别较小。随着时间的推移,入水速度较大的回转体,角速度呈先增大后减小的规律,且其角速度也较大;入水速度较小的回转体,角速度呈先减小后增大的规律;同一入水时刻,回转体俯仰角和角速度均随着入水速度的增加而减小。

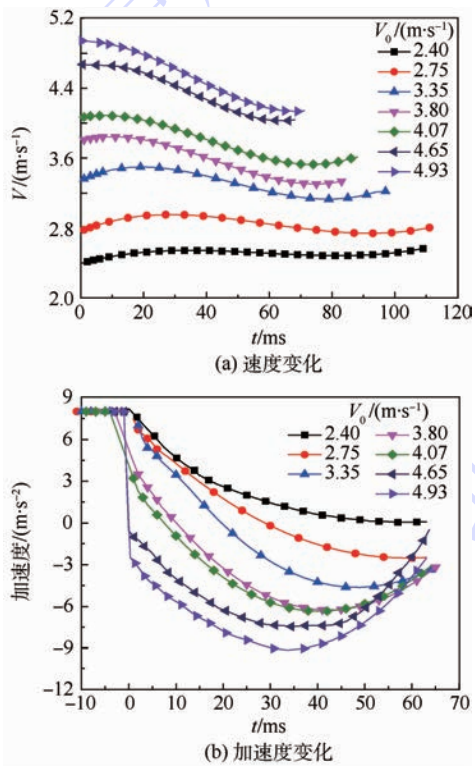


图 9 不同入水速度回转体倾斜入水速度和加速度变化规律对比

Fig. 9 Change rule comparison of oblique water entry velocities and acceleration of revolution bodies with various water entry velocities

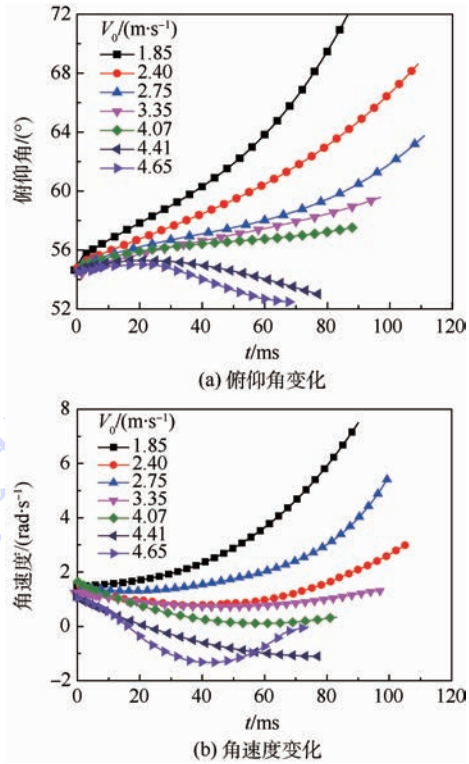


图 10 不同入水速度回转体倾斜入水俯仰角和角速度变化规律对比

Fig. 10 Change rule comparison of oblique water entry pitch angle and angle velocities of revolution bodies with various water entry velocities

2.3 入水角度对倾斜入水过程影响

为进一步研究回转体入水角度对入水空泡形态、回转体运动特性及流体动力特性的影响规律,开展不同入水角度回转体低速倾斜入水实验。实验过程中滑轨与水平面之间夹角分别为 $\beta=45°$, 55°,65°,入水速度 $V_0=4.08$ m/s。

2.3.1 入水角度对回转体入水空泡形态的影响

图 11 给出了回转体以不同角度入水过程空泡形态变化特性。从图中可以看出,随着入水角度的减小,入水空泡的非对称性逐渐加强,且深闭合发生后闭合点附近空泡的上漂现象更为显著。入水角度 $\beta=45°$ 时,入水过程中产生的喷溅,较另外两种入水角度更为强烈。与垂直入水产生的近似轴对称喷溅不同,倾斜入水过程中产生的喷溅大部分出现在水平速度分量方向,这主要是由于倾斜入水过程中回转体头部撞击水面的不对称性造成的。在相同的入水速度情况下,入水角度越小,其水平方向的速度分量就越大,引起的喷溅也更为强烈。

图 12 为不同入水角度回转体入水过程中空泡无量纲直径随时间的变化曲线。图 12(a)为空泡扩张过程中入水深度 $H=2D_0$ 处空泡无量纲直径变化过程,图12(b)为整个入水过程中入水

北航学报 赠阅

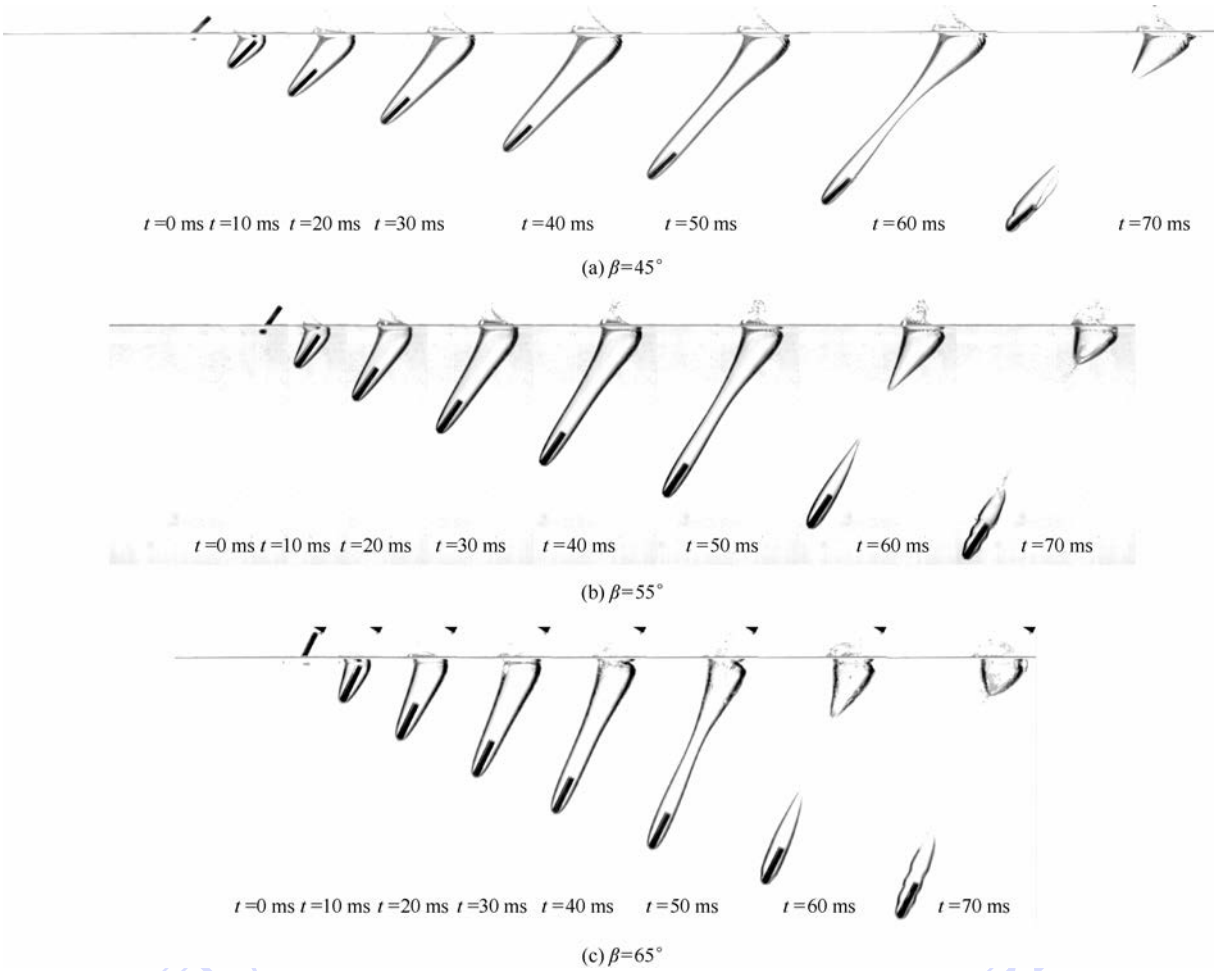


图 11　不同入水角度回转体倾斜入水空泡形态变化特性

Fig. 11　Morphologic variation characteristics of oblique water entry cavity of revolution bodies at various water entry angles

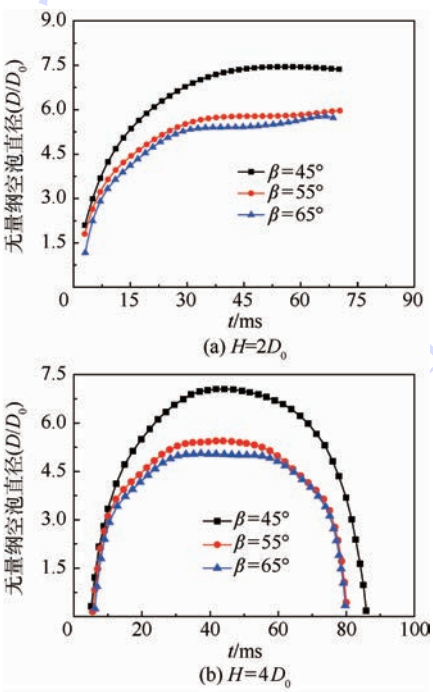


图 12　不同入水角度入水无量纲空泡直径随时间的变化

Fig. 12　Temporal change of dimensionless diameter of water entry cavity at various water entry angles

深度 $H=4D_0$ 处空泡无量纲直径变化过程。从图中可以看出,随着入水时间的推移,相同入水深度处不同入水角度下入水空泡的无量纲直径逐渐出现差异,但均呈相似性较强的非线性变化趋势;整个入水过程中,同一入水时刻相同入水深度处空泡的无量纲直径随着入水角度的减小而增加,且增长幅度越来越大。

图 13 为不同入水深度处空泡壁面无量纲速度随时间变化规律对比(正值为空泡扩张速度,负值为空泡收缩速度)。从图中可以看出,在撞击自由液面时刻附近,不同入水角度回转体的入水空泡具有相近的无量纲扩张速度;入水后随着时间的推移,同一入水时刻不同入水角度下的空泡壁面无量纲扩张速度随着入水角度的减小而增加,且增长幅度越来越大;对于入水深度 $H=4D_0$ 处入水空泡,在收缩过程中其壁面无量纲收缩速度随着时间的推移缓慢增加;在空泡发生深闭合后,由于空泡溃灭产生的扰动,无量纲速度发生不规则变化,而后其值在短时间内迅速增大,且同一时刻无量纲速度随着入水角度的减小而减小。

北航学报 赠阅

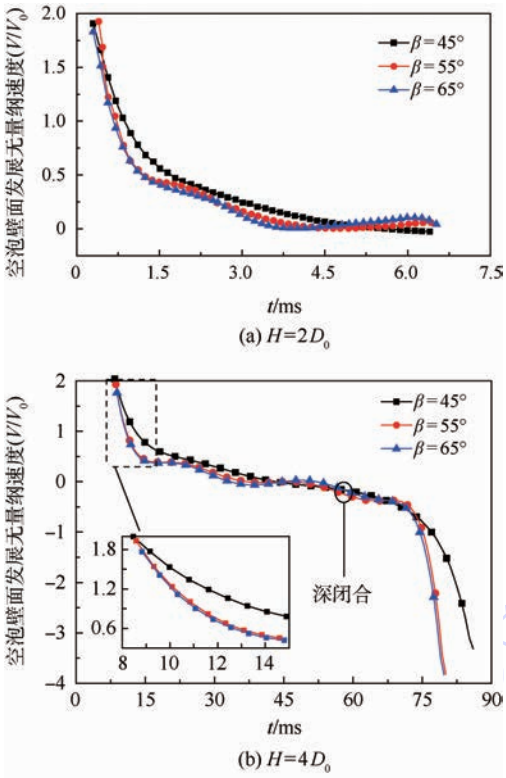


图 13 空泡壁面发展无量纲速度

Fig. 13 Dimensionless velocities of cavity wall development

2.3.2 入水角度对回转体运动特性的影响

图 14 为不同入水角度回转体速度随时间变化规律对比。从图中可以看出,入水初期 3 种入水角度回转体速度值均小幅增加,而后随着入水时间的推移,开始逐渐减小;同一入水时刻,回转体速度值随着入水角度的增加而减小。在入水初期,由于速度方向上回转体的重力分量大于其受到的流体阻力,回转体的加速度为正值,速度缓慢增加;入水过程中物体受到的阻力与其速度的平方成正比,随着入水速度的增加,回转体受到的流体阻力越来越大,回转体加速度逐渐由正值变为负值,继而开始做减速运动。

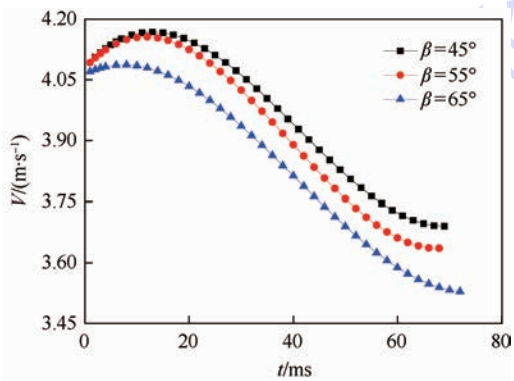


图 14 不同入水角度回转体倾斜入水速度变化规律对比

Fig. 14 Change rule comparison of oblique water entry velocities of revolution bodies at various water entry angles

图 15 为不同入水角度回转体以相同速度入水后阻力系数随时间变化规律对比。从图中可以看出,入水过程中不同入水角度回转体阻力系数呈先增大后减小的趋势,且具有较强的非线性特性;在同一入水时刻,回转体阻力系数随着入水角度的增加而增大。

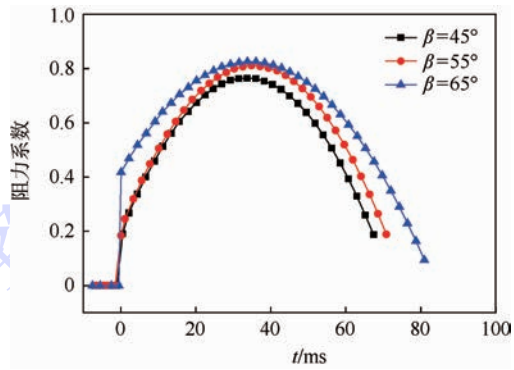


图 15 不同入水角度回转体阻力系数变化规律对比

Fig. 15 Change rule comparison of drag coefficients of revolution bodies at various water entry angles

3 结 论

本文针对回转体低速倾斜入水过程空泡的生成和演化特性进行了研究,开展对比实验量化分析了回转体入水速度和入水角度对回转体入水空泡、速度、俯仰角和阻力系数的影响规律,得到结论如下:

1) 入水空泡深闭合时刻空泡长度、最大空泡直径和空泡闭合点深度随回转体入水速度的增加而增大;空泡深闭合发生时间随着入水速度的增加而小幅增大,在本文实验速度范围内,最大差值仅为 2.6 ms。

2) 入水过程中速度和俯仰角变化规律随入水初始速度的不同而出现差异;同一入水时刻回转体俯仰角和俯仰角速度均随着回转体入水速度的增加而减小。

3) 入水角度对入水空泡直径、回转体速度和阻力系数变化规律有较大影响;同一入水时刻,随着入水角度的增加,回转体入水空泡直径和速度随之减小,阻力系数随之增大。

由于入水问题具有高度非线性和非定常性,且影响因素众多,尚有许多工作需要进一步开展。入水过程中,入水物体的流体动力特性取决于其结构参数及入水参数的共同作用,其结构参数包括:头部形状、长细比、尾部形状质量及惯性矩等,入水参数包括:入水速度、入水角度、俯仰角及偏航角等。近年来随着水下武器及宇宙飞船和火箭

发动机等在水面回收等技术的应用,研究运动体在多参数共同作用下入水过程中的流体动力特性显得更加迫切和重要。

参考文献(References)

[1] 王永虎,石秀华. 入水冲击问题研究的现状与进展[J]. 爆炸与冲击,2008,28(3):276-282.
WANG Y H,SHI X H. Review on research and development of water-entry impact problem[J]. Explosion and Shock Waves,2008,28(3):276-282(in Chinese).

[2] WORTHINGTON A M,COLE R S. Impact with a liquid surface studied by the aid of instantaneous photography [J]. Philosophical Transactions of the Royal Society,1900,194(A):175-200.

[3] WORTHINGTON A M,COLE R S. A study of splashes [M]. New York:Longmans Green and Company,1908:25-76.

[4] MAY A. Vertical entry of missiles into water[J]. Journal of Applied Physics,1953,23(12):1362-1372.

[5] TABUTEAU H,SIKORSKI D,SIMON J,et al. Impact of spherical projectiles into a viscoplastic fluid[J]. Physical Review E. Statistical,Nonlinear & Soft Matter Physics,2011,84(3):1971-1982.

[6] CHARTERS A C. The aerodynamic performance of small spheres from subsonic to high supersonic velocities[J]. Journal of the Aeronautical Sciences,1945,12(4):468-476.

[7] ABELSON H I. Pressure measurements in the water-entry cavity [J]. Journal of Fluid Mechanics,1970,44(part 1):129-144.

[8] SHI H H,ITOH M,TAKAMI T. Optical observation of the supercavitation induced by high-speed water entry[J]. Journal of Fluids Engineering-Transactions of the ASME,2000,122(4):806-810.

[9] 施红辉,周浩磊,吴岩,等. 伴随超空泡产生的高速细长体入水实验研究[J]. 力学学报,2012,44(1):49-55.
SHI H H,ZHOU H L,WU Y,et al. Experiment on water entry of high-speed slender body and the resulting supercavitation [J]. Chinese Journal of Theoretical and Applied Mechanics,2012,44(1):49-55(in Chinese).

[10] SHI H H,MAKOTO K. Underwater acoustics and cavitating flow of water entry[J]. Acta Mechanica Sinica,2004,20(4):374-382.

[11] 何春涛,王聪,何乾坤,等. 圆柱体低速入水空泡试验研究[J]. 物理学报,2012,61(13):134701-134708.
HE C T,WANG C,HE Q K,et al. Low speed water-entry of cylindrical projectile[J]. Acta Physica Sinica,2012,61(13):134701-134708(in Chinese).

[12] 杨衡,张阿漫,龚小超,等. 不同头型弹体低速入水空泡试验研究[J]. 哈尔滨工程大学学报,2014,35(9):1060-1066.
YANG H,ZHANG A M,GONG X C,et al. Experimental study of the cavity of low speed water entry of different head shape projectiles[J]. Journal of Harbin Engineering University,2014,35(9):1060-1066(in Chinese).

[13] 胡青青. 不同倾角下钝体入水后的超空泡流动的实验观察及数值计算[D]. 杭州:浙江理工大学,2014:8-9.
HU Q Q. Experimental observation and numerical calculation of supercavity flow of blunt body under different inclination angle into the water[D]. Hangzhou:Zhejiang Sci-Tech University,2014:8-9(in Chinese).

[14] 梅哲力. 细长圆柱体入水的特性研究[D]. 南京:南京理工大学,2014:32-49.
MEI Z L. Characteristic research on the water entry of slender cylinder[D]. Nanjing:Nanjing University of Science and Technology,2014:32-49(in Chinese).

[15] 马庆鹏,魏英杰,王聪,等. 锥头圆柱体高速入水空泡数值模拟[J]. 北京航空航天大学学报,2014,40(2):204-209.
MA Q P,WEI Y J,WANG C,et al. Numerical simulation of high-speed water-entry cavity of cone cylinder[J]. Journal of Beijing University of Aeronautics and Astronautics,2014,40(2):204-209(in Chinese).

[16] 陈晨,马庆鹏,魏英杰,等. 空气域压力对高速射弹入水流场影响[J]. 北京航空航天大学学报,2015,41(8):1443-1450.
CHEN C,MA Q P,WEI Y J,et al. Effects of operating pressure on high-speed projectile's water-entry flow[J]. Journal of Beijing University of Aeronautics and Astronautics,2015,41(8):1443-1450(in Chinese).

[17] BLACKWELL B D. Preliminary remarks on the cavity made by a projectile entering water:S. R. E./U. 1/4[R]. New York:Department of Science Rescue,1944.

[18] MAY A. Water entry and the cavity-running behavior of missiles:SEAHAC/TR-75-2 [R]. New York:US Department of Commerce,1975.

[19] EPPS B P. An impulse framework for hydrodynamic force analysis:Fish propulsion,water entry of spheres,and marine propellers[D]. Cambridge:Massachusetts Institute of Technology,2010:35-68.

作者简介:

宋武超　男,博士研究生。主要研究方向:多相流计算及实验。
Tel.:0451-86413510
E-mail:songwuchao2015@163.com

王聪　男,博士,教授,博士生导师。主要研究方向:运动体多相流动力学、水下发射技术。
Tel.:0451-86413510
E-mail:alanwang@hit.edu.cn

北京航空航天大学学报

北航学报 赠阅

Experiment of cavity and trajectory characteristics of oblique water entry of revolution bodies

SONG Wuchao, WANG Cong*, WEI Yingjie, XU Hao

(School of Astronautics, Harbin Institute of Technology, Harbin 150001, China)

Abstract: Studies of the oblique water entry of revolution bodies are conducted. By the experiment of revolution body velocity water entry with different velocities and angles, cavity evolution on revolution bodies is discussed, including formation, development and pinch-off of oblique entry. The influence of water entry velocity and angle on the cavities, revolution body velocity, pitching angle and drag coefficients was experimentally obtained. The experiments show that at the moment when the cavity is pinched off, the length and largest diameters of the cavity, and the depth of the closure point are larger as the velocity of revolution bodies increases. The cavity closure time increases slightly with the increase of revolution body velocity. When the velocity of revolution bodies is larger, the pitching angle and angle velocity are smaller. With the increase of water-entry angle of revolution bodies, the diameter of the cavity and the velocity of revolution bodies gradually become smaller, but the drag coefficients become larger.

Key words: revolution body; kinematical parameters; oblique water entry; water entry cavity; experiment study

北京航空航天大学学报

Received: 2015-10-23; **Accepted**: 2016-05-13; **Published online**: 2016-05-17 11:11

URL: www.cnki.net/kcms/detail/11.2625.V.20160517.1111.003.html

Foundation items: Natural Science Foundation of Heilongjiang Province of China (A201409); the Fundamental Research Funds for the Central Universities (HIT.NSRIF.201159); Special Foundation for Harbin Science and Technology Innovation Talents of China (2013RFLXJ007)

* **Corresponding author**. Tel.: 0451-86413510 E-mail: alanwang@hit.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0680

基于子阵激励能量匹配的多子阵交错阵列设计

李龙军[1,2], 王布宏[1,*], 夏春和[2], 刘新波[1], 曹帅[1]

(1. 空军工程大学 信息与导航学院, 西安 710077; 2. 北京航空航天大学 计算机学院, 北京 100083)

摘 要: 多子阵稀疏交错共享孔径阵列天线是实现多功能阵列天线设计的有效途径。提出了一种基于子阵激励能量匹配的多子阵稀疏交错优化方法。根据均匀线阵激励与其方向图之间存在的傅里叶变换关系,该方法首先通过快速傅里叶变换(FFT)获得特定方向图的激励频谱能量分布,然后分析均匀线阵目标方向图频谱能量的分布特征,采用交叉选取子阵激励的方法,确定各子阵单元位置,使得阵列天线单元激励能量均匀分配,从而确保各子阵方向图近似一致。在此基础上通过迭代 FFT 的方法降低各子阵天线方向图旁瓣峰值(PSL),实现低旁瓣,同性能的多子阵交错共享孔径阵列天线设计。仿真研究与实验表明,通过本文提出多子阵交错方法设计的共享孔径多稀疏交错子阵具有运算量小、孔径利用率高、各子阵峰值旁瓣电平低且拥有相似方向图的优点。利用子阵交错的方法,通过控制各子阵主波瓣指向,能轻松实现多波束指向的多功能阵列天线设计。

关 键 词: 多子阵交错; 共享孔径; 迭代快速傅里叶变换 (FFT); 多功能阵列天线; 旁瓣电平

中图分类号: TN820

文献标识码: A **文章编号**: 1001-5965(2016)11-2395-08

随着定位导航和通信技术的不断发展,航电雷达系统的高度集成化和功能多样化使得机载天线的数量和复杂程度大幅增加,这些都对阵列天线设计提出了新的挑战[1-2]。共享孔径阵列天线能够通过孔径的空分复用,使用交错布置于同一个天线孔径的不同子阵完成不同功能。这种方法不但可以减少机载天线的数量,节省平台空间,而且可以有效避免天线单元间的电磁兼容问题。但目前为止,对多子阵交错稀疏布阵有效方法的研究还比较少。如何使阵列天线的稀疏优化方法和交错方法有机结合,设计出旁瓣峰值(PSL)电平低和方向图性能近似一致的交错子阵是当前迫切需要解决的问题。

多子阵交错稀疏优化布阵的关键问题在于如何在保证子阵方向图性能近似一致的情况下减小子阵方向图的旁瓣峰值,目前已有一些学者对此类问题提出了探索性的解决方案,并取得了不错的效果。研究主要集中在利用差集与互补差集及其延拓实现多子阵的交错稀疏布阵。文献[3-5]利用循环差集及其“补集”实现了多子阵的稀疏交错布阵,并将其成功地应用于宽带天线、极化捷变天线和雷达收发双置天线的综合设计。文献[6-8]利用几乎差集对两子阵的稀疏交错优化设计进行了深入研究,并将其应用于极化捷变天线的综合和超声成像收发双置天线的设计。但现有的差集还比较少,基于差集的交错稀疏优化方法还只能对一些阵元数量特定的部分阵列进行优化,而且产生的交错子阵方向图旁瓣峰值还有进一步优化的空间。文献[9-10]利用遗传算法设计了两子阵交错布阵的共享孔径线阵天线,并将其

收稿日期: 2015-10-20; **录用日期**: 2015-12-25; **网络出版时间**: 2016-01-21 15:47
网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20160121.1547.010.html
基金项目: 国家自然科学基金 (61172148)
* **通讯作者**: Tel.: 029-84791762 E-mail: wbhyl@aliyun.com

引用格式: 李龙军, 王布宏, 夏春和, 等. 基于子阵激励能量匹配的多子阵交错阵列设计[J]. 北京航空航天大学学报, 2016, 42(11): 2395-2402. LI L J, WANG B H, XIA C H, et al. Design of array with multiple interleaved subarrays based on subarray excitation energy-matching[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(11): 2395-2402 (in Chinese).

北航学报 赠阅

成功应用于单脉冲和差波束阵列天线的实现。但对于大孔径多子阵的稀疏交错共享孔径天线,遗传算法存在优化目标函数过多,收敛速度慢,运算效率低,容易陷入局部最优等缺陷。

迭代快速傅里叶变换(FFT)算法利用均匀阵列天线方向图与阵元激励间存在的傅里叶变换的关系,通过对阵元激励的优化选择,实现稀疏、低副瓣的天线设计[11-13]。本文通过对线阵天线方向图的频谱激励能量分布的分析,根据密度加权阵的相关原理[14],通过将天线方向图的频谱能量等额均分,实现了多子阵的稀疏交错布阵。本文的主要工作如下:

1)基于均匀阵列天线方向图与阵元激励间存在的傅里叶变换的关系,详细分析并提出了基于方向图频谱能量平均分配的均匀线阵多子阵交错方法,从理论和仿真上验证了该方法的有效性和可行性。

2)提出了改进迭代 FFT 算法,实现了多子阵交错稀疏布阵。

3)通过实验仿真证明,利用新提出的均匀线阵交错方法能够有效对多子阵交错的共享孔径天线进行优化设计。并通过分析孔径利用率,子阵间旁瓣峰值的差值及子阵旁瓣峰值等参数对交错方法的性能进行了详细的评估。通过控制不同子阵主瓣波束指向,实现了多波束指向的多功能稀疏交错共享孔径天线的设计。

1 多子阵交错阵列天线优化模型

由于线阵天线方向图满足乘法定理,为了方便研究阵元位置与天线阵列方向图之间的关系,可以将天线单元看成全向性点元。当主瓣指向 θ_0 方向,来波信号波长为 λ,一个阵元数为 N、阵元栅格间距为 $\lambda/2$、入射方位角为 θ、阵元激励为 I_n 的均匀线阵,其方向图可表示为

$$F(\theta)=\sum_{n=0}^{N-1}I_n\mathrm{e}^{\mathrm{j}n\pi\cos(\theta-\theta_0)} \tag{1}$$

对式(1)进行傅里叶逆变换可得

$$f(t)=\frac{1}{M}\sum_{n=0}^{M-1}F(n)\mathrm{e}^{\mathrm{j}nt\omega_0} \tag{2}$$

式中:M 为周期;t 为傅里叶变换的采样点数,$1<t<M$;$\omega_0=2\pi/M$ 为采样间隔。

由式(1)和式(2)可知,阵元的激励 I_n 与阵因子之间存在傅里叶变换的关系[11]。多子阵的稀疏交错共享孔径阵列天线是指利用孔径空分复用技术,将多个子阵交错布置在同一个孔径上,使各个子阵的方向图性能近似一致,旁瓣电平都能处于一个较低的水平。根据交错阵列天线的设计要求可知,其数学优化模型可以表示为

$$\begin{cases}\text{子阵 } i\ \min\left\{P_{\mathrm{PSL}}^{(i)}(d_1,d_2,\cdots,d_{N/t})=\max\left(\left|\dfrac{F_i(\theta_i)}{P_{\max}^{(i)}}\right|\right)\right\}\\ \text{子阵 } j\ \min\left\{P_{\mathrm{PSL}}^{(j)}(d'_1,d'_2,\cdots,d'_{N/t})=\max\left(\left|\dfrac{F_j(\theta_j)}{P_{\max}^{(j)}}\right|\right)\right\}\\ \min\{\Delta=\mathrm{PSL}_i-\mathrm{PSL}_j\}\\ \text{同时 } d_i\neq d'_j\end{cases} \tag{3}$$

式中 $P_{\mathrm{PSL}}^{(i)}$ 和 $P_{\mathrm{PSL}}^{(j)}$ 分别为第 i、j 个子阵的旁瓣峰值;$F_i(\theta_i)$ 和 $F_j(\theta_j)$ 分别为第 i、j 个子阵旁瓣区域的值;$P_{\max}^{(i)}$ 和 $P_{\max}^{(j)}$ 分别为第 i、j 个子阵的主瓣增益值;θ_i 和 θ_j 分别为子阵 i、j 旁瓣区域的方位角。

2 直线阵列多子阵交错方法

由于直线阵列天线阵元激励与方向图函数之间存在傅里叶变换的关系,将阵元激励看成离散的“频谱信息”,则其对应的方向图曲线可以看成是阵列天线的“时域信号”。不失一般性,以 400 元直线阵列天线为例,对本文提出的方法进行详细的描述。假设子阵数目为 4($T=4$),阵列天线的稀疏率为 75%,生成初始的随机 0、1 序列 $\boldsymbol{I}$,该序列含 100 个 1 值,300 个 0 值。对 $\boldsymbol{I}$ 做 2 048 点的 FFT 逆变换生成时域信号 P,将 P 旁瓣区域中大于约束旁瓣的值置为 -25 dB,其余值保持不变,从而生成新的时域信号,如图 1(a)所示。对约束后的方向图做 FFT 变换生成激励矢量,由文献[13]可知,可以截取其前 400 个点的激励频谱值作为迭代约束后的阵元满阵激励 $\boldsymbol{I}_N$,对其进行由大到小的排序,生成 $\boldsymbol{I}_{\mathrm{r}}$ 如图 1(b)所示。

为了均分阵列天线的频谱能量,以交错的方式对 I_1 进行选择,被选择激励相对于原天线孔径所在的位置即为子阵阵元所在的位置,其数学表述如式(4)所示。4 个子阵阵元的位置及其激励值如图 2 所示。从图 2 中可知,4 个子阵的阵元位置是交错稀疏分布在同一个天线孔径上的,子阵的激励分布趋势及相应的值基本相同。

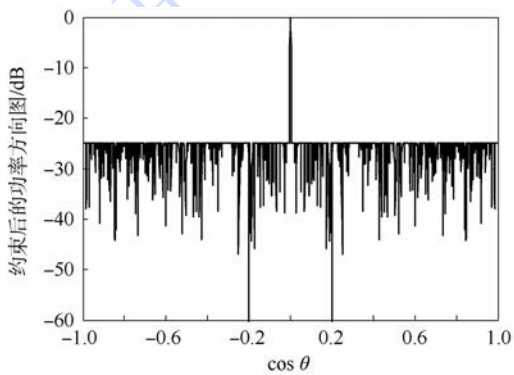
$$\begin{cases}\boldsymbol{I}_{\mathrm{s}1}=\boldsymbol{I}_{\mathrm{r}}(4n+1)\\ \boldsymbol{I}_{\mathrm{s}2}=\boldsymbol{I}_{\mathrm{r}}(4n+2)\\ \boldsymbol{I}_{\mathrm{s}3}=\boldsymbol{I}_{\mathrm{r}}(4n+3)\\ \boldsymbol{I}_{\mathrm{s}4}=\boldsymbol{I}_{\mathrm{r}}(4n+4)\\ n=(0,1,\cdots,99)\end{cases} \tag{4}$$

式中:$\boldsymbol{I}_{si}$ 为第 i 个子阵的激励。根据密度加权阵的相关原理,线阵栅格点上阵元存在的概率取决于线阵栅格点上的激励权值分布,权值幅度大的栅格点上阵元存在的概率就大。因此可以设置存

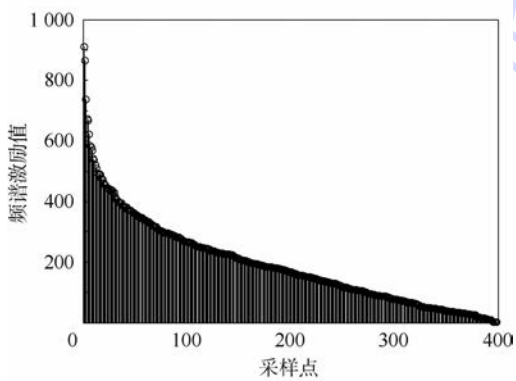
北航学报 赠阅

在阵元位置上的值为“1”,表示阵元存在,无阵元位置的栅格处的值为“0”。经过一次迭代后的子阵天线方向图如图 3 所示。从图 3 中可知,通过该交错方法形成的 4 个子阵方向图旁瓣值近似相同,子阵 1 的旁瓣峰值为 -12.01 dB,子阵 2 的旁瓣峰值为 -15.15 dB,子阵 3 的旁瓣峰值为 -13.11 dB,子阵 4 的旁瓣峰值为 -15.14 dB,子阵间的旁瓣峰值最大差值为 3.14 dB,最小差值仅为 0.01 dB。为了进一步减小子阵的旁瓣峰值,可以循环迭代使用 FFT 算法,最终实现多子阵交错的共享孔径直线阵列天线的设计。

考虑一个阵元数为 N,交错子阵数为 T,稀疏率为 R 的等栅格分布的均匀直线阵列天线,首先随机生成 N/T 个“1”值的 0、1 序列作为阵列天线的初始激励值,对其作 Q 点的 FFT 变换,得到阵列天线的方向图值,令阵列天线方向图旁瓣区域中大于约束旁瓣值的点上的值等于约束旁瓣,然后对约束后的阵列天线方向图进行 Q 点的 FFT 变换,获得 Q 个点激励幅值 $\boldsymbol{I}$。截取前 N 个值作为满阵激励值 $\boldsymbol{I}_N$,并对其进行由大到小排序,形成新的序列 $\boldsymbol{I}_r$,则可以选取向量($\boldsymbol{I}_r(1)$,$\boldsymbol{I}_r(1+T)$,…,$\boldsymbol{I}_r(1+T\times K)$)对应向量 $\boldsymbol{A}$ 所在的位置作为稀疏子阵 1 的单元位置,($\boldsymbol{I}_r(2)$,$\boldsymbol{I}_r(2+$



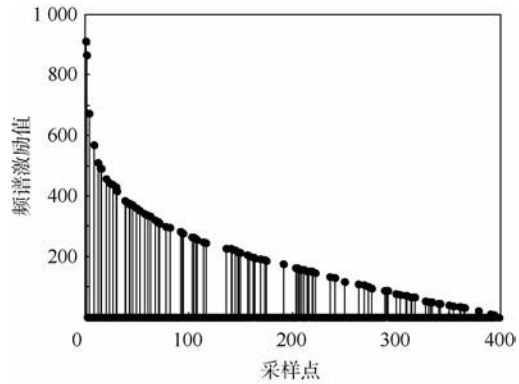
(a) 约束后的阵列天线方向图



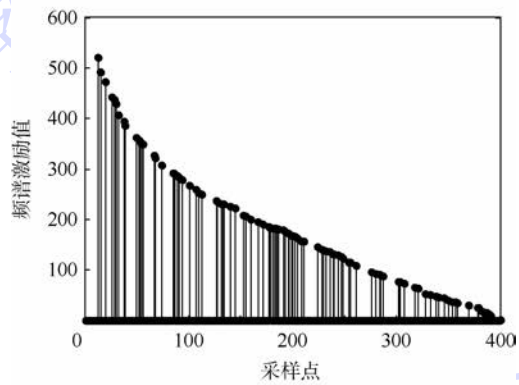
(b) 排序后激励的分布

图 1　约束后的阵列天线方向图及其阵元激励

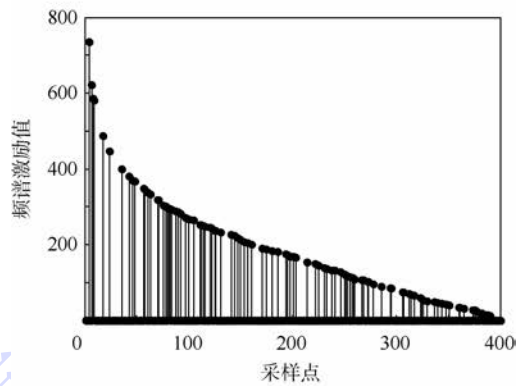
Fig. 1　Restraint pattern and array element excitations



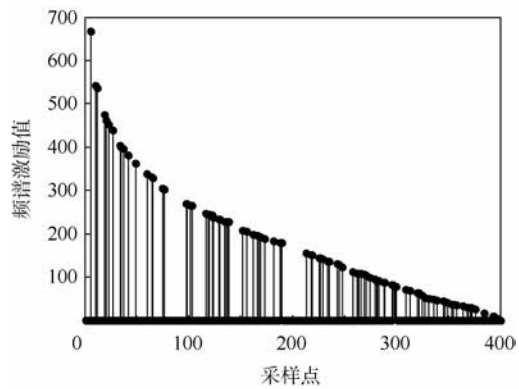
(a) 子阵1



(b) 子阵2



(c) 子阵3



(d) 子阵4

图 2　交错子阵单元的激励与位置

Fig. 2　Excitations and positions of interleaved subarray elements

$T),\cdots,\boldsymbol{I}_r(2+T\times K))$对应向量$\boldsymbol{I}_N$所在的位置作为子阵 2 的阵元位置，……，$(\boldsymbol{I}_r(T),\boldsymbol{I}_r(2T),\cdots,\boldsymbol{I}_r(T\times(K+1))$对应向量$\boldsymbol{I}_N$所在的位置作为子阵 T 的阵元位置($K=N/T-1$)。所有子阵均为密度加权阵，激励值置为 1。以第一个子阵天线单元激励值作为输入值，进入下一次的迭代。采用这样的迭代方式原因在于，根据文献[11]可知，通过少量的迭代进行 FFT 变换，天线单元的激励能量值会更集中于采样点的前端，这样能让子阵天线的旁瓣峰值更接近于约束旁瓣值，从而达到降低子阵旁瓣的目的。图 4 为改进型迭代 FFT 算法的流程图，其具体步骤可概括如下。

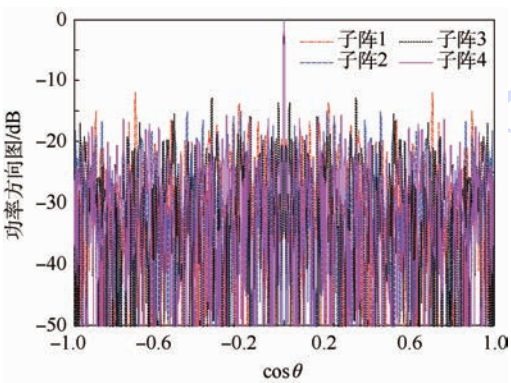


图 3 四子阵交错的天线方向图

Fig. 3 Patterns of four interleaved subarrays

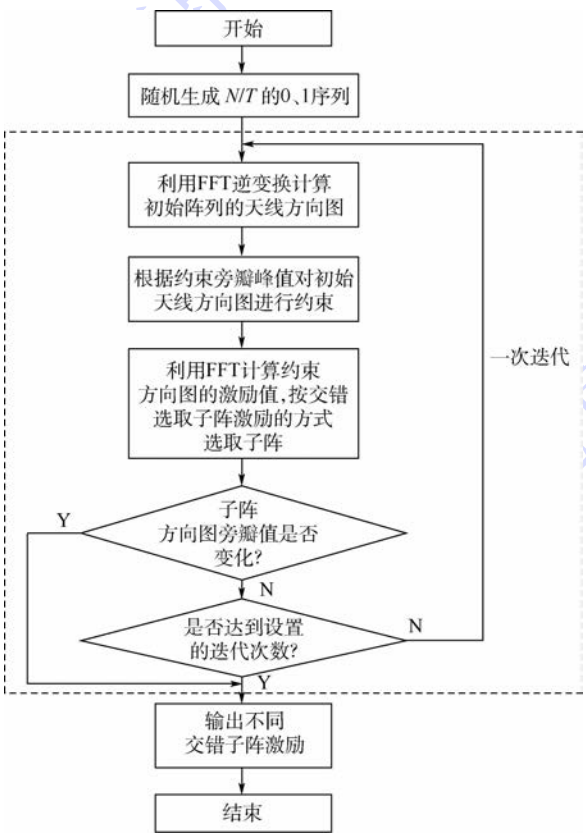


图 4 改进型迭代 FFT 算法流程图

Fig. 4 Flowchart of modified iterative FFT algorithm

1）按稀疏率 $R=1/T$ 随机稀疏阵元数为 N 的均匀直线阵列，设置阵列激励值 $\boldsymbol{I}$ 为 1。

2）对 $\boldsymbol{I}$ 做 Q 点的 FFT 逆变换，得到 P。

3）对 P 旁瓣区域的值进行判定，令幅值大于约束旁瓣值区域上的点值等于约束旁瓣值，其他点上的值保持不变。

4）对校正后的 P 进行 Q 点的 FFT 变换，得到新的激励值 $\boldsymbol{I}$。

5）截取 $\boldsymbol{I}$ 中的前 N 个值，生成新的激励矩阵，对其进行由大到小排列，生成新的激励向量 $\boldsymbol{I}_N$。

6）以$(\boldsymbol{I}_r(i),\boldsymbol{I}_r(i+T),\cdots,\boldsymbol{I}_r(i+T\times K))$位置处的激励对应在 $\boldsymbol{A}$ 上的位置作为第 i 个子阵单元所在的位置，同时将该位置上的激励值置为 1。

7）对新生成的子阵 1 的激励向量进行 FFT 逆变换，判断更新后的天线方向图旁瓣峰值相对于上一次迭代时的旁瓣峰值是否有变化，是则继续下一步的计算，否则输出结果。

8）将子阵 1 的单元激励序列作为新的输入值进入下一轮的迭代，重复步骤 2）~步骤 8），直到新生成的天线方向图旁瓣峰值不再变化，输出优化后的结果。

3 实验仿真与分析

3.1 四子阵交错的共享孔径直线阵列天线

为了验证基于迭代 FFT 算法的直线阵列多子阵交错的有效性和可行性，首先对 100 元等间隔分布的均匀直线阵列进行四子阵交错的优化设计($T=4$)。阵列的稀疏率为 75%，单个子阵阵元数量为 25，此时，在利用迭代 FFT 算法时，应以以下方式选取子阵单元的激励：

$$\begin{cases}\boldsymbol{I}_{s1}=\boldsymbol{I}_r(4n+1)\\ \boldsymbol{I}_{s2}=\boldsymbol{I}_r(4n+2)\\ \boldsymbol{I}_{s3}=\boldsymbol{I}_r(4n+3)\\ \boldsymbol{I}_{s4}=\boldsymbol{I}_r(4n+4)\\ n=(0,1,\cdots,24)\end{cases}\tag{5}$$

设置旁瓣约束值为 -20 dB，对其做 2 048 点的 FFT 变换。得到图 5 所示的四子阵交错的单元结构。其中，“1”代表该点处有阵元，“0”代表该点处的阵元被稀疏。从图 5 中可知，4 个子阵交错分布在同一个天线孔径，各个子阵孔径大小与原均匀阵列天线孔径基本一致，且避免了不同子阵天线单元的位置重叠。若以式(6)表示阵列的孔径利用率

$$\eta=\frac{L_i}{L_{\text{full}}}\times 100\%\tag{6}$$

式中:L_i 为稀疏子阵 i 的阵列天线孔径电尺寸大小;L_{full} 为均匀阵列天线孔径电尺寸大小。可以得出子阵 1 的孔径利用率为 97%,子阵 2 的孔径利用率为 92%,子阵 3 的孔径利用率为 88%,子阵 4 的孔径利用率为 96%。交错布阵的 4 个子阵使得 100 元的均匀阵列天线的孔径利用率大幅提高,单个子阵的阵列天线分辨率得到大幅提高。

为了对本文提出的交错方法的性能进行评估,首先对 25 元均匀布阵的直线阵列和 100 个阵元以稀疏率为 25% 随机稀疏的直线阵列天线方向图进行实验仿真,仿真结果如图 6 所示。从图 6(a)中可知,当阵列以 25 单元均匀布阵时,阵列天线方向图旁瓣值为 -13.21 dB,但由于该阵列天线的天线孔径为 10λ,导致阵列天线的分辨率较低,此时方向图零点主瓣宽度为 10°。若采用 100 个等间隔栅格以随机稀疏的方式布置 25 个阵元,由于此时阵列天线的孔径为 45.5λ,阵列天线的分辨率大幅增加,方向图零点主瓣宽度为 2.5°,但由此导致的问题就是方向图旁瓣峰值较大,此时的阵列天线副瓣电平为 -6.448 dB。交错稀疏共享孔径阵列天线就是将交错方法与阵列稀疏优化方法有机协同,设计出子阵方向图旁瓣低且孔径利用率高的阵列天线。对图 5 结构的阵列天线进行仿真计算,得到的仿真结果如图 7 所示。

从图 7 中可知,通过交错稀疏的共享孔径阵列天线的旁瓣峰值相对于随机稀疏的子阵有了相应的减小,且各个子阵充分利用了天线孔径,4 个子阵方向图的零点主瓣宽度同为 3.6°,且可以得出子阵 1 的方向图旁瓣峰值为 -9.911 dB,子阵 2 的旁瓣峰值电平为 -9.556 dB,子阵3的旁瓣峰

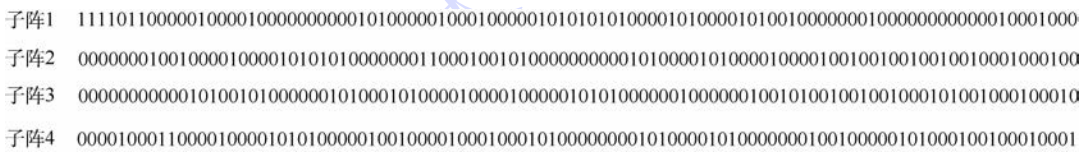


图 5　四子阵交错的天线孔径结构

Fig. 5　Aperture architecture of four interleaved subarrays

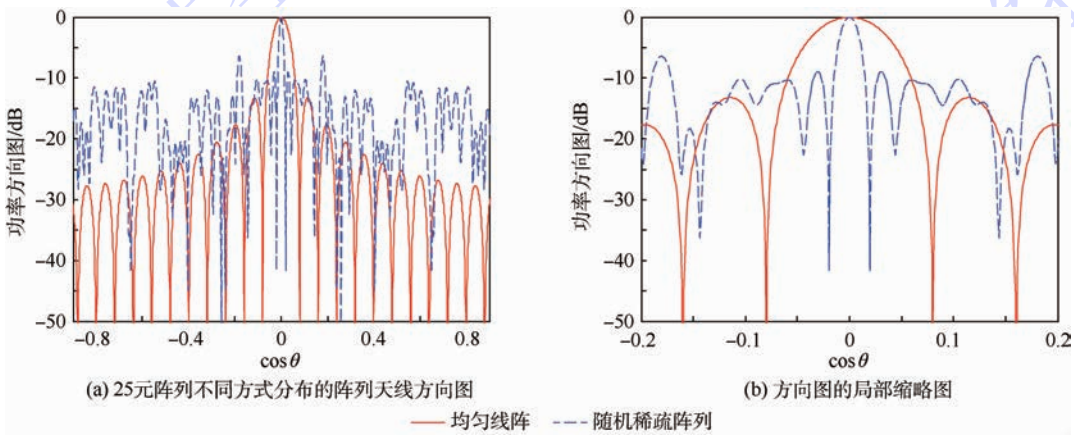


图 6　不同方式分布的直线阵列天线方向图

Fig. 6　Patterns of linear array antenna with different layouts

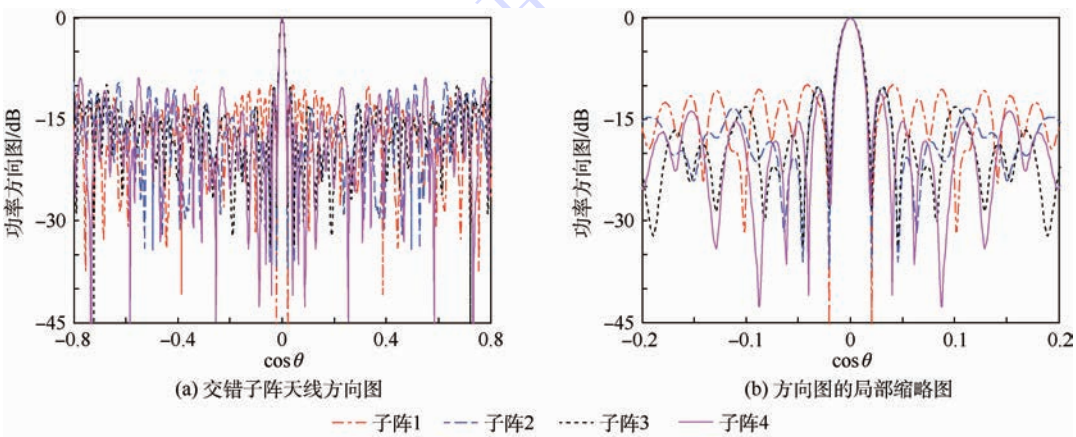


图 7　四子阵交错的直线阵列天线方向图

Fig. 7　Patterns of linear array antenna for four interleaved subarrays

值电平为 -9.965 dB,子阵 4 的旁瓣峰值电平为 -8.901 dB。子阵间的旁瓣峰值电平最大差值为 1.064 dB,最小差值仅为 0.054 dB,这就保证了 4 个子阵方向图性能基本一致,证明基于迭代 FFT 算法的直线阵列多子阵交错方法是有效的。

3.1.1 激励权值对基于迭代 FFT 交错子阵方向图的影响

本文主要研究的是阵列单元位置对天线方向图的影响,寻求多子阵交错稀疏分布在同一个天线孔径的优化方法,通过优化单元位置来降低子阵天线方向图旁瓣峰值,实现阵列天线稀疏优化布阵。需要特别说明的是,阵列天线设计在工程应用中是一个需要系统解决的优化问题。虽然利用本文方法稀疏交错优化后的阵列天线旁瓣峰值相对于工程应用要求来说较高,但阵列天线单元位置优化是阵列天线其他优化研究工作的基础和前提(包括阵元激励优化、阵元扰动、互耦和通道不一致误差的估计与消除),在优化子阵单元位置后,可以通过优化算法(如遗传算法、萤火虫算法[15-16]等)对交错稀疏子阵单元上的激励进行进一步优化,以进一步降低子阵天线方向图旁瓣峰值,本文不再详述。

3.1.2 改进型迭代 FFT 算法旁瓣约束值的讨论

旁瓣约束值主要是针对满阵天线,它的选取直接影响了阵列采样点激励在前段的集中程度,其目的主要是为了降低子阵天线旁瓣峰值,当阵列天线为满阵时,天线的旁瓣值能够满足约束要求,但由于稀疏子阵只选取了阵列天线的部分激励点,因此稀疏后的子阵天线方向图(部分激励点做傅里叶变换获得的值)旁瓣峰值是无法满足约束要求的,选取一个合理的旁瓣约束值能够使目标方向图对应的激励点值集中分布在激励的前段部分,使得在只选取激励前段部分点值做傅里叶逆变换得到的方向图旁瓣峰值更好地接近约束旁瓣值,达到降低子阵天线方向图旁瓣值与约束旁瓣值之间差值的目的,因此旁瓣约束值并不是越大越好或者越小越好,前期通过大量的实验仿真结果得出,在选取约束旁瓣峰值时一般选大于优化后稀疏阵列天线旁瓣峰值 10 ~ 15 dB 的值优化效果最好。

3.2 多子阵交错的多波束指向共享孔径阵列天线设计

现代导航定位系统往往要求人们对多个空中和地面目标进行精确定位,这就要求雷达天线能够同时接收不同方向的来波信号。而传统数字多波束天线的设计相对固定,收发使用不灵活。设计多子阵交错的共享孔径阵列天线能有效解决这些问题。为了实现多波束指向阵列天线的设计,本文对阵元数为 100、稀疏率为 80% 的直线阵列进行 5 子阵交错的优化布阵。为了使子阵方向图频谱能量近似相同,在确定子阵天线结构,选取子阵单元方向图频谱能量时,应以式(7)的方式进行选取,优化后的阵列天线单元位置结构图如图 8所示。从图中可以得出,子阵 1 的孔径利用率为 96%,子阵 2 的孔径利用率为 96%,子阵 3 的孔径利用率为 96%,子阵 4 的孔径利用率为 95%,子阵 5 的孔径利用率为 94%,5 个子阵交错共享于同一个天线孔径,孔径利用率得到了大幅提高。

$$\begin{cases} \boldsymbol{I}_{s1} = \boldsymbol{I}_r(5n+1) \\ \boldsymbol{I}_{s2} = \boldsymbol{I}_r(5n+2) \\ \boldsymbol{I}_{s3} = \boldsymbol{I}_r(5n+3) \\ \boldsymbol{I}_{s4} = \boldsymbol{I}_r(5n+4) \\ \boldsymbol{I}_{s5} = \boldsymbol{I}_r(5n+5) \\ n = (0,1,\cdots,19) \end{cases} \tag{7}$$

对图 8 结构的阵列天线方向图进行仿真,仿真结果如图 9 所示。图 9(a)和图 9(b)分别为 5 个子阵主瓣都指向 90°时天线方向图及其缩略图。从图中可看出,5 个子阵方向图形状大致相同,由于单个子阵阵元数量的减少,相对于四子阵交错的共享孔径阵列天线,五子阵交错的阵列天线旁瓣峰值略有增大。其中子阵 1 的旁瓣峰值电平为 -7.49 dB,子阵 2 的旁瓣峰值为 -7.09 dB,子阵 3 的旁瓣峰值为 -6.98 dB,子阵 4 的旁瓣峰值电平为 -7.17 dB,子阵 5 的旁瓣峰值电平为 -8.66 dB,子阵间最大峰值电平差为1.68 dB,最小峰值旁瓣电平差为 0.017 dB。证明基于迭代 FFT 算法的多子阵交错方法是稳健的。使子阵主瓣波束分别指向不同的方向,能形成不同波束指向的共享孔径多功能阵列天线,如图 10 所示。

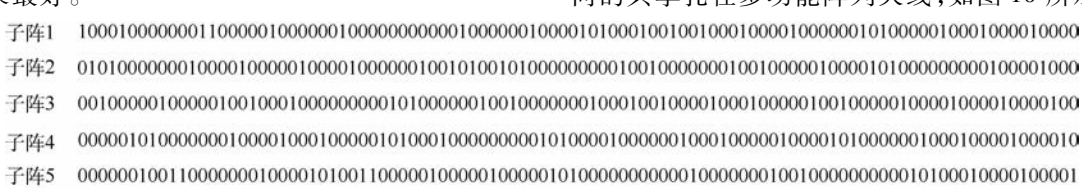


图 8 五子阵交错的天线孔径结构

Fig. 8 Aperture architecture of five interleaved subarrays

北航学报 赠阅

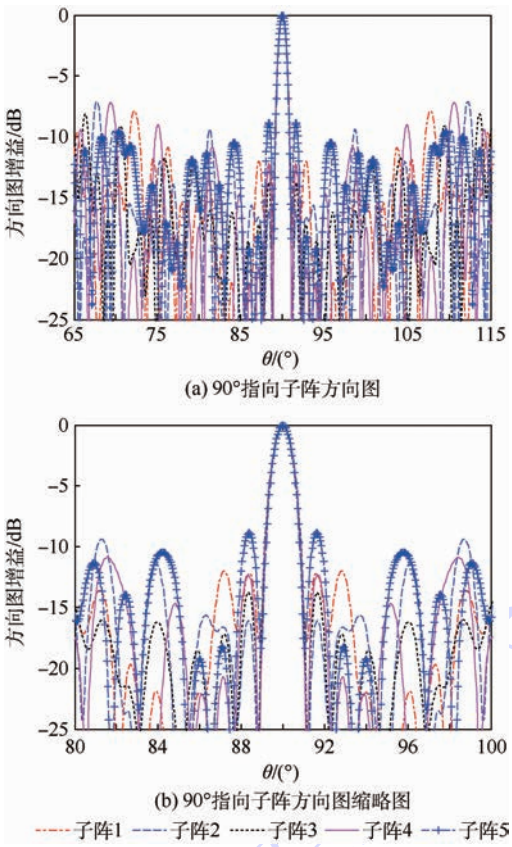


图 9　同一波束指向的五子阵交错阵列方向图

Fig. 9　Patterns of five interleaved subarrays array in the same beam direction

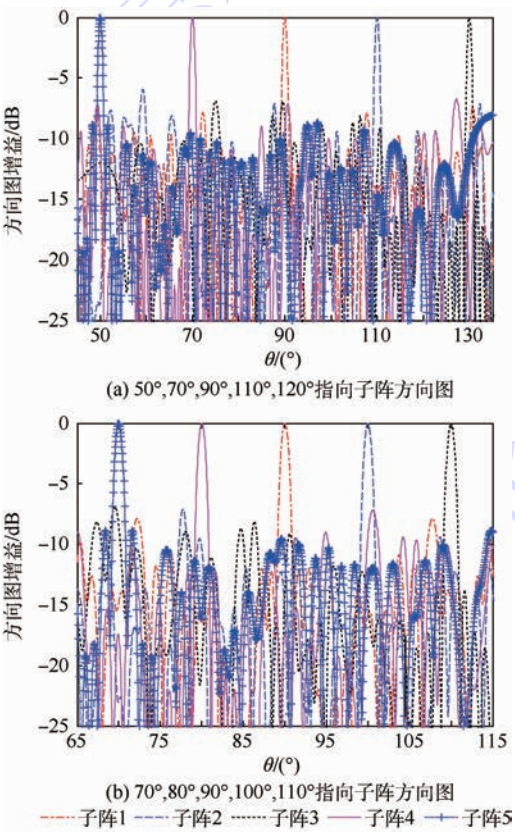


图 10　不同波束指向的五子阵交错阵列方向图

Fig. 10　Patterns of five interleaved subarrays array in different beam directions

从图 10 中可知，当子阵分别指向 50°、70°、90°、110°、120°和 70°、80°、90°、100°、110°时阵列天线拥有近似一致的方向图性能，因此可以通过交错稀疏多个子阵，灵活控制子阵主瓣波束指向，实现多波束指向的多功能阵列天线设计。

4　结　论

利用改进型迭代 FFT 算法，通过将均匀线阵方向图频谱能量等额分配给不同的交错子阵单元，得出以下结论：

1）该方法可以实现多子阵的交错稀疏设计，且每个子阵都有较高的孔径利用率和性能近似一致的功率方向图，其方向图旁瓣峰值能够接近同等孔径下的均匀直线阵列天线旁瓣峰值。

2）通过控制不同子阵方向图主瓣波束指向，可以实现空域多功能阵列天线的设计。

3）需要说明的是阵列天线优化设计是一个需要系统解决的问题，本文主要研究的是阵元位置对阵列天线方向图的影响，虽然稀疏阵列由于阵元之间的距离较大，使得子阵内和子阵间的耦合效应较小，但互耦仍然是影响阵列天线方向图性能的重要因素，因此下一步工作需要着重解决子阵间和子阵内部单元互耦的联合校正问题。

参考文献（References）

[1] JOSEFSSON L P. Conformal array antenna theory and design [M]. Toronto: Wiley-IEEE Press, 2006: 230-246.

[2] ASIM A K, ANTHONY K B. Null steering in irregularly spaced sparse antenna arrays using aperture distributed subarrays and hybrid optimizer [J]. IET Microwaves, Antennas & Propagation, 2014, 8(2): 86-92.

[3] COMAN C I, LAGER I E, LIGTHART L P. A deterministic solution to the problem of interleaving multiple sparse array antennas[C]// Proceedings of 2nd European Radar Conference-EuRAD. Piscataway, NJ: IEEE Press, 2005: 243-246.

[4] QIN Y, GAO S, SAMBELL A, et, al. Design of a broadband square-ring-slot coupled patch antenna[J]. Microwave Optical Technology Letter, 2005, 47(5): 454-457.

[5] SIMEONI M, LAGER I E, COMAN C I. Implementation of polarization agility in planar phased-array antennas by means of interleaved subarray[J]. Radio Science, 2009, 44(44): 538-549.

[6] OLIVERI G, MASSA A. Fully interleaved linear arrays with predictable sidelobes based on almost difference sets[J]. IET Radar, Sonar and Navigation, 2010, 4(5): 649-661.

[7] OLIVERI G, LIZZI L, MASSA A. ADS interleaved arrays with reconfigurable polarization[C]// 2010 IEEE International Symposium on Antennas and Propagation and CNC-USNC/URSI Radio Science Meeting-Leading the Wave, AP-S/URSI 2010. Piscataway, NJ: IEEE Press, 2010: 1-4.

[8] OLIVERI G, MASSA A. ADS-based array design for 2-D and 3-D

ultrasound imaging[J]. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 2010, 57(7): 1568-1582.

[9] HAUPT R L. Interleaved thinned linear arrays [J]. IEEE Transactions on Antennas and Propagation, 2005, 53(9): 2858-2864.

[10] HAUPT R L. Interleaving thinned sum and difference linear arrays [C]// Proceedings of 2006 IEEE Antennas and Propagation Society International Symposium, Albuquerque. Piscataway, NJ: IEEE Press, 2006: 4773-4776.

[11] HOU Q S, GUO Y, WANG B H, et al. A novel method for analyzing sidelobe structure of nonuniform sparse linear array[C]// IET International Radar Conference 2009. Stevenage: IET Press, 2009: 1-4.

[12] DAVID B D. Low-sidelobe pattern synthesis using iterative fourier techniques coded in MATLAB[J]. IEEE Antennas and Propagation Magazine, 2009, 51(2): 137-150.

[13] WARREN P D P. Weighted thinned linear array design with the iterative FFT technique [J]. IEEE Transactions on Antennas and Propagation, 2011, 59(9): 3473-3477.

[14] WILL P M N K. Synthesis of thinned planar circular and square arrays using density tapering [J]. IEEE Transactions on Antennas and Propagation, 2014, 62(4): 1555-1563.

[15] 严韬, 陈建文, 鲍拯. 基于改进遗传算法的天波超视距雷达二维阵列稀疏优化设计[J]. 电子与信息学报, 2014, 36(12): 3014-3020.
YAN T, CHEN J W, BAO Z. Optimization design of sparse 2-D arrays for over-the-horizon radar (OTHR) based on improved genetic algorithm [J]. Journal of Electronics & Information Technology, 2014, 36(12): 3014-3020 (in Chinese).

[16] WANG G G, GUO L, DUAN H, et al. A new improved firefly algorithm for global numerical optimization[J]. Journal of Computational and Theoretical Nanoscience, 2014, 11(2): 477-485.

作者简介:

李龙军 男, 博士研究生。主要研究方向: 阵列信号处理。
Tel.: 15829920225
E-mail: lilongjunboy@buaa.edu.cn

王布宏 男, 博士, 教授, 博士生导师。主要研究方向: 阵列信号处理、网络与信息安全。
Tel.: 029-84791762
E-mail: wbhyl@aliyun.com

夏春和 男, 博士, 教授, 博士生导师。主要研究方向: 网络安全与防护、网络测量等。
Tel.: 010-82317656
E-mail: XCH@buaa.edu.cn

Design of array with multiple interleaved subarrays based on subarray excitation energy-matching

LI Longjun[1,2], WANG Buhong[1,*], XIA Chunhe[2], LIU Xinbo[1], CAO Shuai[1]

(1. School of Information and Navigation, Air Force Engineering University, Xi'an 710077, China;
2. School of Computer Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: The shared aperture array antennas with multiple interleaved subarrays is an effective solution of the design of multi-functional array antennas. A new method for interleaving multiple subarrays based on excitation energy-matching is proposed in this paper. In light of the Fourier transform mapping between the element excitations and array factor of uniform linear array, the spectrum of the array factor is first obtained by fast Fourier transform (FFT) and its excitation energy distributions are analyzed thoroughly. The element positions are then selected carefully in a staggered way which ensures that the uniform distribution of the excitation energy of the array antenna elements and the similar patterns can be obtained. Then the peak sidelobe levels (PSL) of the subarrays are reduced with the iterative FFT algorithm. Futhermore, the design of shared aperture antenna array with low sidelobe levels and same performance is achieved. Some numerical simulations and experiments demonstrate that the interleaved subarrays generated from the proposed method possess many favorable characteristics, such as, reduced computational burden, high aperture efficiency, lower PSL than the random thinned array, and perfect matched patterns. By controlling the mainbeam direction, the multi-functional array antenna can be easily designed with multiple beam lobes.

Key words: multiple interleaved subarrays; shared aperture; iterative fast Fourier transform (FFT); multi-functional array antennas; sidelobe levels

Received: 2015-10-20; **Accepted**: 2015-12-25; **Published online**: 2016-01-21 15:47
URL: www.cnki.net/kcms/detail/11.2625.V.20160121.1547.010.html
Foundation item: National Natural Science Foundation of China (61172148)
* **Corresponding author**. Tel.: 029-84791762 E-mail: wbhyl@aliyun.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0711

开放空腔壳体入水流场结构及流体动力特征研究

路中磊，魏英杰*，王聪，孙铁志

（哈尔滨工业大学 航天学院，哈尔滨 150001）

摘 要：基于混合介质雷诺平均 Navier-Stokes 方程，对开放空腔壳体垂直入水运动过程开展了数值研究，得到了压力场、速度场分布，空泡波动、闭合特征，空腔气体涨缩规律，以及流体动力变化规律，并分析了空腔结构在入水运动过程中对流场结构和流体动力的影响。结果表明：液体随气体涨缩同步进出开放端；开放端局部形成波动的压力源和周期性的压力场、速度场分布；入水空泡呈现波动形态，其扩展程度与开放端液体流速相关；空泡内形成气体漩涡，随空腔涨缩往返进出空泡，对空泡闭合具有抑制作用；流体动力呈波动变化规律，频率与气体涨缩频率一致，幅值与气体涨缩程度成正比。开放空腔结构在入水过程中空腔内气体发生涨缩运动，对流场结构和流体动力产生周期性扰动作用，在一定程度上可以减缓冲击、维持空泡及运动的稳定性。

关 键 词：入水；开放空腔壳体；流场结构；流体动力；数值计算
中图分类号：TB126；V19
文献标识码：A **文章编号**：1001-5965(2016)11-2403-10

运动体由空气域穿越自由液面进入水域完成入水运动，入水过程运动体将改变静态流场结构，产生湍动涡、空泡流等现象，同时在流体动力作用下，改变入水弹道，甚至引发结构变形，流场运动与结构体运动两者相互耦合，是一个非常复杂的过程，一直以来，入水问题都是国内外学者重点关注的问题之一。对于开放空腔壳体入水而言，其开放端位于下方首先撞击水面，流入的液体将空腔内气体封闭形成一个相对独立的流域空间。腔体内气液两相流体相互作用、对外部流场的扰动以及附加质量作用，均对流场结构和流动动力等特征产生影响，相对于传统实心体入水过程[1]，其结构的特殊性增加了入水问题的复杂性。相关应用也较为广泛，例如航天飞机固体火箭助推器(Solid Rocket Boosters，SRB)的回收[2]、机载声呐浮标的投放[3]等，此类特殊结构体入水问题已经成为入水问题研究的重要分支。

对入水问题的早期研究多采用试验方法，最早可以追溯到 19 世纪初，Worthington[4] 首次借助于高速摄像技术观察记录了液滴和球型结构体入水全过程，记录并描述了入水喷溅、空泡演化等现象；近年来，Duclaux[5]、Truscott[6-7] 以及何春涛[8]等众多国内外学者依然延续摄影试验方法对结构体入水问题开展了研究，对入水运动规律、入水空泡演化、流体动力特征以及影响因素进行了分析。随着计算机技术和计算流体力学的发展，数值方法在入水问题研究中得到了广泛应用，可以获得更加丰富、更加精细的流场结构信息。Lee 等[9]针对射弹入水空泡的闭合方式和闭合点特征开展了理论和数值计算，其基于能量守恒原理建立了

收稿日期：2015-11-02；**录用日期**：2015-12-18；**网络出版时间**：2016-01-14 17:42
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160114.1742.001.html
基金项目：国家自然科学基金（11672094）；黑龙江省自然科学基金（A201409）；中央高校基本科研业务费专项资金（HIT.NSRIF.201159）；哈尔滨市科技创新人才研究专项资金（2013RFLXJ007）
* **通讯作者**：Tel.：0451-86413510 E-mail：weiyingjie@gmail.com

引用格式：*路中磊，魏英杰，王聪，等. 开放空腔壳体入水流场结构及流体动力特征研究[J]. 北京航空航天大学学报，2016，42(11)：2403-2412. LU Z L，WEI Y J，WANG C，et al. Flow field structure and fluid dynamic characteristics of semi-closed cylinder water-entry[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2403-2412 (in Chinese).*

空泡演化的理论模型,分析了影响闭合方式的因素,以及深闭合方式下空泡的几何、运动特征参数;何春涛等[10]对锥头圆柱体入水流场特征进行数值计算,获得了不同入水速度条件下泡内压力变化规律,以及空泡内压力场、速度场分布情况;陈晨等[11]对高速射弹入水过程开展了数值计算研究,分析了空气域压力对入水喷溅形态、空泡闭合时间及内部空化效应的影响;赵蛟龙等[12]对底部带空腔的圆柱体不同俯仰角入水载荷进行计算,分析了空腔结构在入水冲击载荷作用下的结构响应,以及与俯仰角间的关系;潘龙等[13]针对圆柱体头部喷气入水缓冲特征开展研究,建立了顶部喷气平头圆柱体入水模型;杨衡等[14]开展了横流作用下圆柱体入水数值仿真,分析了在波、流作用下入水弹道和姿态响应。目前,对开放空腔结构入水问题的研究尚处于探索阶段,公开发表的文献资料相对较少,空腔内部流场结构以及对外部流场的扰动均是十分复杂的流动过程,具有较广泛的研究空间和应用前景。

针对含开放空腔的壳体结构在入水过程中内外流场耦合作用特点,忽略壳体结构变形,本文对下方完全敞开的开放空腔壳体垂直入水过程开展了数值模拟,结合腔内气体运动情况,分析了气体涨缩对流场结构和空泡形态的周期性扰动规律;空泡内气体漩涡的运动特征和其对空泡闭合的影响;以及扰动情况下壳体流体动力特征及分布情况。

1 数值计算理论

1.1 基本控制方程

入水运动是一个非定常问题,流动参数均随空间和时间变化,在宏观上,表现为质量的扩散、动量的黏滞耗散、能量的传导等现象,流体流动将受到物理基本定律的支配,包括连续性方程、动量方程和能量方程。

连续性方程:

$$\frac{\partial \rho}{\partial t}+\frac{\partial(\rho \boldsymbol{u}_i)}{\partial x_i}=0 \tag{1}$$

式中:ρ 为流体密度;t 为时间;$\boldsymbol{u}_i$ 为 i 方向速度;x_i 为 i 方向位置。

动量方程:

$$\frac{\partial}{\partial t}(\rho \boldsymbol{u}_i)+\frac{\partial(\rho \boldsymbol{u}_i \boldsymbol{u}_j)}{\partial x_j}=\rho \boldsymbol{g}_i-\frac{\partial P}{\partial x_i}+\frac{\partial}{\partial x_i}\left[(\mu+\mu_t)\left(\frac{\partial \boldsymbol{u}_i}{\partial x_j}+\frac{\partial \boldsymbol{u}_j}{\partial x_i}\right)\right] \tag{2}$$

式中:$\boldsymbol{g}_i$ 为 i 方向重力加速度分量;P 为压强;μ 为动力黏度;μ_t 为湍动黏度。

能量方程:

$$\frac{\partial}{\partial t}(\rho E)+\frac{\partial}{\partial x_i}(\rho \boldsymbol{u}_i E)=-\frac{\partial}{\partial x_i}(\boldsymbol{u}_i P)+\frac{\partial}{\partial x_i}\left(k_{\text{eff}}\frac{\partial T}{\partial x_i}\right) \tag{3}$$

式中:E 为流体能量;k_{eff} 为有效热导率系数;T 为流体温度。

1.2 多相流模型

低速入水问题是一种典型的多相流问题,通常形成具有明确气液界面的泡状流动,气液两相介质互不掺混,并且空泡壁面处气液两相流动速度基本一致,因此,对于此问题可应用均相流(homogeneous flow)模型,以一等效可压缩流体代替气液两相流。

本文选用流体体积(Volume of Fluid,VOF)均相流模型对入水过程中气液两相流体的流动进行描述,且不考虑气液相变影响。VOF 模型将气液两相流看作单一介质的混合流体,各相共享同一压力场和速度场,共用一套动量方程,通过求解单独的动量方程和处理穿过区域的每一流体的容积比来描述各相流体。混合流体的混合密度 ρ_m、混合动力黏度系数 μ_m 和混合能量 E_m 分别如下:

$$\rho_m=\alpha_l\rho_l+\alpha_g\rho_g \tag{4}$$

$$\mu_m=\alpha_l\mu_l+\alpha_g\mu_g \tag{5}$$

$$E_m=\frac{1}{\rho_m}(\alpha_l\rho_l E_l+\alpha_g\rho_g E_g) \tag{6}$$

式中:α_l 和 α_g 分别为液体和气体所占体积分数,$\alpha_l+\alpha_g=1$,下标 m 表示气液混合物,l 表示液体,g 表示气体。

1.3 湍流模型

在壳体入水过程中,开放空腔内气体将对外部流场产生扰动,引起空泡壁面波动流动,是开放腔体结构体入水过程的主要特征之一,波动强度具有周期性,存在黏性流动和湍动流动交替出现,导致空泡壁面流动复杂多变。另一方面,对于低速入水条件,空腔气体对流场的扰动微弱,雷诺数相对较小,因此对此问题的模拟需要考虑低雷诺数模型。

本文选用 SST k-ω 低雷诺数湍流模型,其通过湍动能(turbulence kinetic energy)k 和比耗散率(specific dissipation rate)ω 来描述湍动黏度,并以湍动黏度修正系数 γ 实现在黏性流动区域和湍动流动区域均具有适合的湍动黏度,表达式如下:

$$\mu_t=\rho\frac{k}{\omega}\gamma \tag{7}$$

式中:$k=\frac{1}{2}(u'^2+v'^2+w'^2)$ 为湍动能;$\omega=\frac{\varepsilon}{C_\mu k}$ 为

北航学报 赠阅

比耗散率,ε 为湍动耗散率;常数 $C_\mu = 0.09$;$\gamma = \min\left[a^*, \frac{a_1\omega}{SF}\right]$ 为湍动黏度修正系数,$a^* = a_\infty^*\left(\frac{a_0^* R_k + Re_t}{R_k + Re_t}\right)$ 为低雷诺数修正系数,$Re_t = \frac{\rho k}{\mu\omega}$,$R_k = 6$、$a_\infty^* = 1$、$a_0^* = 0.024$ 及 $a_1 = 0.31$ 为经验系数,S 为应变率不变测度,F 为调和函数。

湍动能输运方程和比耗散率输运方程:

$$\frac{\partial(\rho k)}{\partial t} + \frac{\partial(\rho k \boldsymbol{u}_i)}{\partial x_i} = \frac{\partial}{\partial x_j}\left[\left(\mu + \frac{\mu_t}{\sigma_k}\right)\frac{\partial k}{\partial x_j}\right] + \widetilde{G}_k - Y_k + S_k \tag{8}$$

$$\frac{\partial(\rho\omega)}{\partial t} + \frac{\partial(\rho\omega \boldsymbol{u}_j)}{\partial x_j} = \frac{\partial}{\partial x_j}\left[\left(\mu + \frac{\mu_t}{\sigma_\omega}\right)\frac{\partial\omega}{\partial x_j}\right] + G_\omega - Y_\omega + D_\omega + S_\omega \tag{9}$$

式中:σ_k、σ_ω 为普朗特数;$\widetilde{G}_k = \min(G_k, 10\rho\beta^* k\omega)$ 为湍动能生成项,G_k 为标准 k-ω 模型湍动能生成项,β^* 为系数;$Y_k = \rho\beta^* k\omega$ 为湍动能耗散项;S_k 为湍动能源项;$G_\omega = \frac{a}{v_t}\widetilde{G}_k$ 为比耗散率产生项;$Y_\omega = \rho\beta\omega^2$ 为比耗散率耗散项,β 为系数;S_ω 为比耗散率源项;$D_\omega = \frac{2}{\sigma_{\omega,2}}(1 - F_1)\frac{\rho}{\omega}\cdot\frac{\partial k}{\partial x_j}\cdot\frac{\partial\omega}{\partial x_j}$ 为交叉扩散项,F_1 为调和函数,$\sigma_{\omega,2} = 1.168$。

1.4　数值方法

基于 Fluent 软件平台,采用有限体积法(Finite Volume Method,FVM)[15] 对控制方程离散化求解,时间离散为一阶精度,空间离散均为二阶精度,瞬态项采用时间显式积分离散方式;对流项采用二阶迎风格式;扩散项中梯度离散采用基于单元体的最小二乘法插值(least-squares cell based)格式;压力插值采用基于体积力(body force weighted)离散格式;体积率插值采用 Geo-Reconstruct 离散格式;其他均采用二阶迎风格式(second order upwind)。选用基于压力的分离式求解器(segregation solver),压力场、速度场采用 Coupled 耦合算法,联立求解离散化连续性方程和离散化动量方程。

2　计算模型

2.1　几何模型

开放空腔壳体模型选用一端具有完全开放的全腔体刚性结构,图 1 给出模型示意图,为方便后续分析,对模型各面进行定义。模型几何参数及属性见表 1。

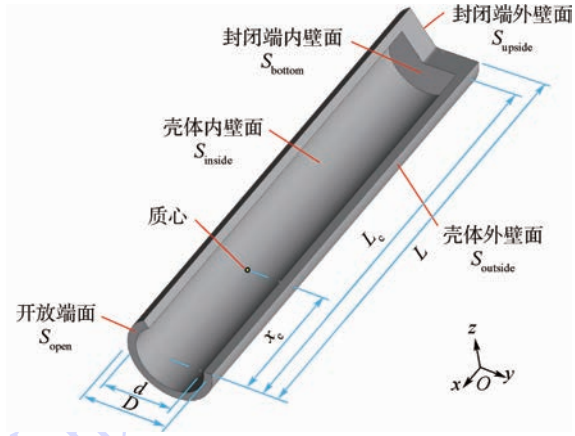


图 1　开放空腔壳体模型示意图

Fig. 1　Schematic of model of semi-closed cylinder

表 1　模型几何参数及属性

Table 1　Geometric parameters and properties of model

参数	L/m	L_c/m	D/m	d/m	M/kg	x_c/m
数值	0.20	0.19	0.04	0.032	0.35	0.08

2.2　计算流域及边界条件

图 2 给出了模型入水过程的计算流域和边界条件设置,选用圆柱型计算域,流域尺度参照试算结果,考虑计算效率和计算精度,设定边界区域为压力、速度和温度等流动参数输运现象较为微弱的最小计算流域尺度。计算应用动网格技术,采用动计算域方式,即计算域整体随壳体同步运动,实现沿重力方向单自由度运动。另一方面,动计算域方式导致流域边界不固定,要求在边界上给出时变的压力、体积分数参数分布,通过设定边界压力、液面位置与惯性坐标间的关系,依据壳体

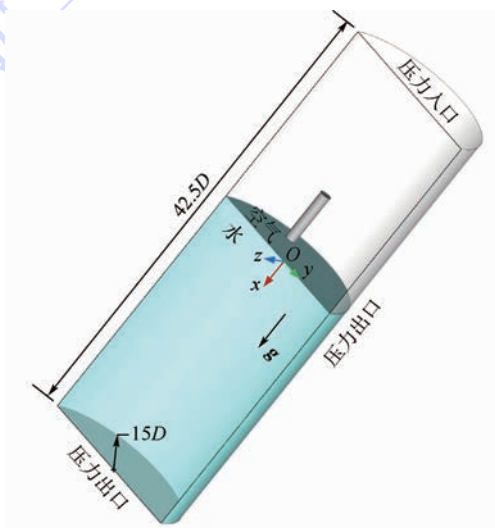


图 2　计算流域及边界条件

Fig. 2　Computational domain and boundary conditions

北航学报 赠阅

运动反馈的位置参数,确定动态边界条件,使边界条件与运动参数匹配,实现网格和数据的更新。通过给定的初始条件,流体动力与弹道参数耦合计算,得到壳体入水的运动规律和流体动力情况。

为方便对计算结果分析,设定坐标系如下:以开放端撞击自由液面时刻为时间零点,自由液面位置为坐标原点,重力方向为运动正方向,如图 2 所示。计算初始条件:壳体开放端面距自由液面垂向距离为 1.25D,即 -0.05 m;初始运动速度 2.50 m/s;对应计算开始时间为 -0.02 s。气相流体采用理想气体,液相流体采用常温下不可压缩的水。具体计算条件和流体属性参数如表 2 所示。

表 2 计算条件及流体属性

Table 2 Computational conditions and fluid properties

入水速度 v_0/(m·s^{-1})	重力加速度 g/(m·s^{-2})	环境压力 P_∞/Pa	液体属性 ρ_l/(kg·m^{-3})	液体属性 μ_l/(kg·ms^{-1})	气体属性 比热 c_p/(J·(kg·K)$^{-1}$)
2.55	9.8	101 325	998.2	0.001 003	1 006.43

2.3 计算网格划分

图 3 给出了壳体模型表面网格和流域空间网格的划分方式,均采用了正六面体结构化网格,可以较好控制网格质量和数量。为了保证计算结果的有效性和提高计算精度,更好地捕捉开放空腔壳体入水过程中的流动特征,对壳体开放端附近扰动区域的网格、开放端下方和腔体内部流场参数畸变区域的网格以及尾缘涡区域的网格进行了加密处理。

根据 SST k-ω 湍流模型对近壁面网格距壁面无量纲长度 y^+ 的要求,同时避免网格出现较大的横纵比,对近壁面网格细化,图 4 给出空泡闭合时刻的开放空腔壳体壁面 y^+ 分布,高速回射流使壁面黏性层厚度达到极小值,其值均小于 11。全域网格最大横纵比网格均出现在远场,比例

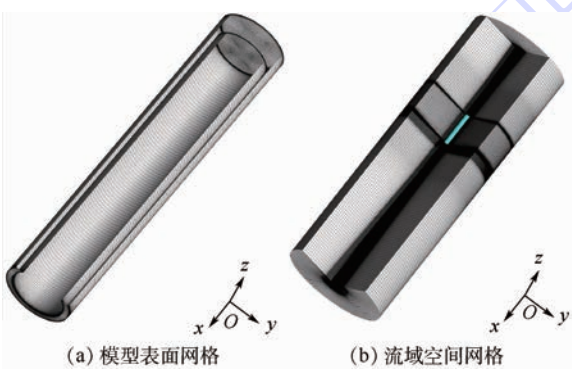


图 3 计算网格

Fig. 3 Computational grid

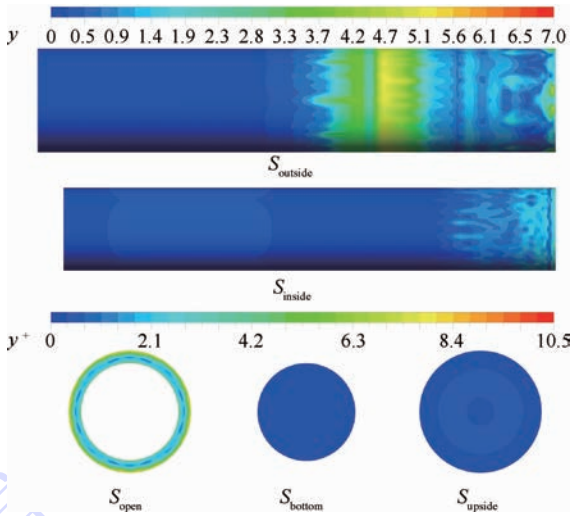


图 4 圆柱壳体外表面 y^+ 分布

Fig. 4 Distribution of y^+ on external surface of cylinder shell

大于 0.006;最小网格出现在壳体开放端,尺度为 0.00 002 m,最大网格出现在远场,尺度为 0.015 m;网格总量为 648 万。

3 结果分析

入水过程中,开放空腔内气体在流体动力作用下将发生周期涨缩现象,是开放空腔壳体入水问题的主要特点之一,一方面其运动对液体流域产生扰动,改变流场结构,形成特殊流动现象;另一方面气体涨缩作用于壳体,改变流体动力特征。下面依次对流场结构、空泡特征和流体动力 3 个问题进行分析。

3.1 流场结构特征分析

入水后空腔内气体涨缩将引起液体流域受到周期性扰动,流场结构特征也表现出周期性变化,图 5 分别给出了开放端液体排出阶段和吸入阶段流线图和等压线图。

对于液体排出阶段,空腔内气体压缩到极大值后,腔内压力到达极值,开放端流域形成局部高压,由于作用范围较小,使压力梯度较大,液面开始下降,但初始流速较小,外部流域未产生明显流动现象。排出的液体逐渐加速,压力作用范围扩展,在动压的作用下,开放端流域继续处于高压状态,但腔内气体静压减弱,压力梯度略有减小,随流速增加空泡区域环流形成,促进了局部空泡扩展。空腔持续性排出液体导致压力急剧下降,开放端逐渐形成相对低压区域,出现逆向压力梯度,液体流动速度减缓,特别是横向流动开始衰减,环流减弱,空泡扩展阶段结束。

北航学报 赠阅

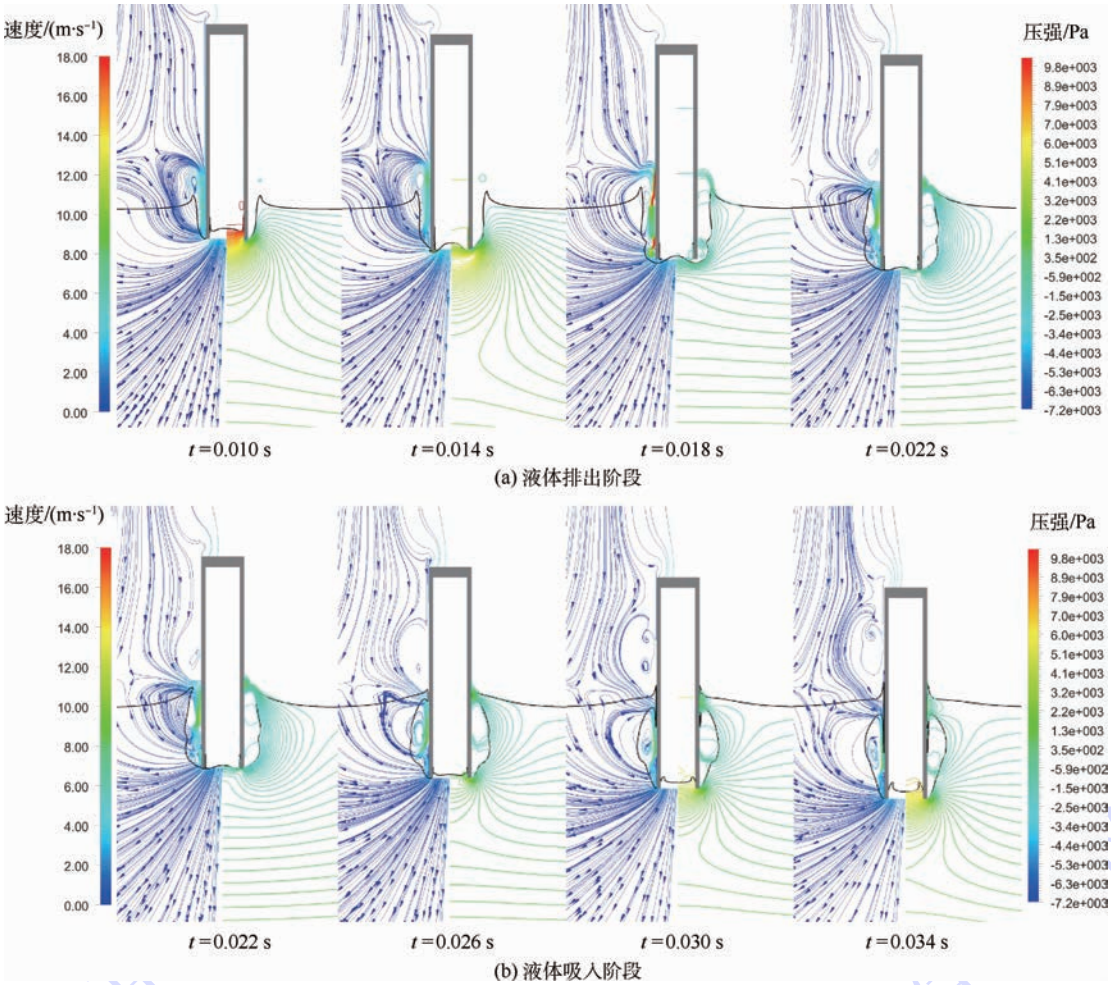


图 5　流线及等压线

Fig. 5　Streamlines and pressure isobars

对于液体吸入阶段,空腔内气体过度膨胀导致压力小于环境压力,开放端附近流域形成指向空腔内部的压力梯度,外部液体回流,腔内液面上升。随着流入液体的加速,空腔内气体压缩、开放端下方向外侧排开的液体流量减小,空泡扩展程度受到抑制,开始进入局部收缩阶段。空腔压缩程度进一步增加,流入的液体开始减速,开放端再次形成高压区,并形成以开放端向外辐射的压力梯度,直至腔内压缩至最大值,然后再一次进入到液体排出阶段,同时空泡收缩阶段结束,完成一次周期运动。

考虑入水空泡结构特征,开放端的这种扰动使空泡区域的局部流场产生压力梯度,导致沿轴线不同位置处空泡的扩展程度出现差异,由于扰动是周期性的,因此空泡在整体上表现出波动现象。可见,这种空泡的波动流动是开放空腔对流场扰动的一种传播形式,其可以反映出空腔内气体涨缩的频率及强度,同时又将反作用于流场,特别对于空泡内易压缩的气体介质,空泡的波动影响到形成气体环流。闭合前,在空泡与大气连接的开口处形成明显的气体漩涡,并受空泡波动影响而出现移动;闭合后,形成附着于壳体运动的独立空泡,其完全处于液面以下,此阶段空腔内气体涨缩极度衰减,空泡波动程度减弱。同时由于液体密度、黏度均较大,液体漩涡耗散十分迅速,难以形成大尺度涡,对于液体流场主要为势流流动,空腔形成的谐振腔在开放端产生周期性压力源,形成沿径向扩展的速度势,沿深度出现周期性的压力场和速度场分布。

3.2　空腔气体扰动对入水空泡的影响

入水空泡是入水运动的重要特征之一,是气液两相流动相互作用的结果,对近壁面流场结构和流体动力具有较大影响。对于开放空腔壳体而言,开放空腔结构对入水空泡具有一定的扰动影响,为得到较为直观的空泡流动特征,针对开放腔体和封闭腔体 2 种壳体模型入水空泡形态开展了入水试验,图 6 给出具有封闭腔体和开放腔体的 2 种模型入水空泡试验照片,为方便对入水空泡扰动分析,试验采用与文中数值计算 1∶1 模型和一致的运动初始条件,见表 1、表 2,可以发现 2 种

北航学报 赠阅

模型的入水空泡特征及形态具有明显差异，在开放空腔气体扰动作用下，入水空泡出现波动，并抑制、推迟空泡的闭合，空泡闭合时刻照片在图中用方框标记。

通过 3.1 节对入水流场结构特征的分析，知道空腔内气体涨缩将引起空泡形态周期性变化，在空泡内形成气体漩涡流动，漩涡运动又将对空泡波动、闭合特征和闭合时间产生影响，下面针对空泡内的气体漩涡流动情况开展分析。



图 6 入水空泡演化试验照片

Fig. 6 Test photographs of development of water-entry cavities

漩涡强度通常可以采用速度梯度张量第二不变量（the second invariant of the velocity gradient tensor）[16] 来描述，其表达式为

$$Q=\frac{1}{2}(1-\delta_{ij})\left(\frac{\partial \boldsymbol{u}_i}{\partial x_i}\cdot\frac{\partial \boldsymbol{u}_j}{\partial x_j}-\frac{\partial \boldsymbol{u}_i}{\partial x_j}\cdot\frac{\partial \boldsymbol{u}_j}{\partial x_i}\right) \tag{10}$$

式中：δ_{ij} 为 Kronecker 符号；Q 为正值表示流体微元旋转运动强于轴向变形，为负值表示流体微元旋转运动弱于轴向变形，可以有效地对漩涡强度进行描述。

图 7 给出了开放端液体排出阶段和吸入阶段空泡内气体涡旋分布及演化规律。在液体排出阶段，由于空泡未闭合，泡内空间与外界大气形成气体环流，在空泡开口处产生气体漩涡，随着排出液体流速的增加，涡旋强度也逐渐增强，并通过空泡与大气连接的开口处向空泡外部移动。另一方面对于空腔内部，液体沿空腔内壁面流速相对较高，特别在开放端首次射流阶段，部分气体将由开放端外缘泻出，在空泡内部形成一个反向旋转的气

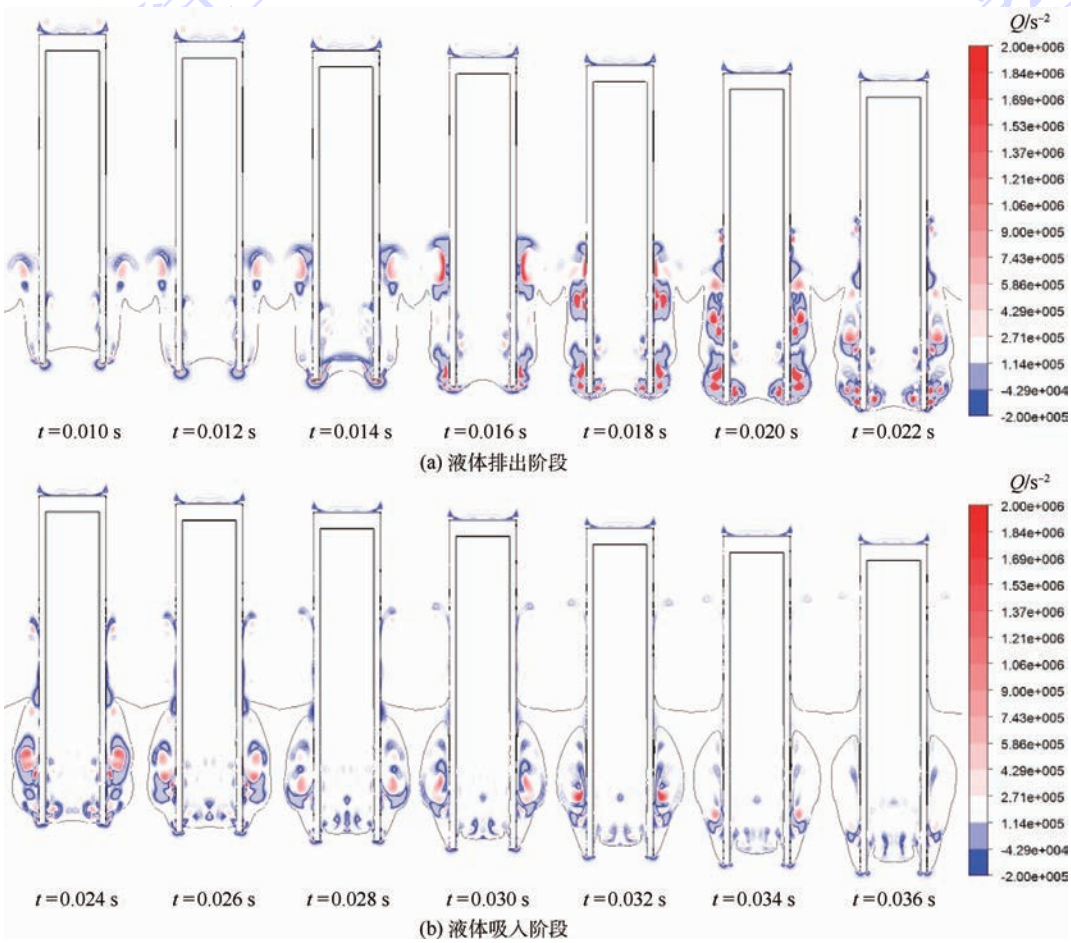


图 7 入水空泡内 Q 值分布

Fig. 7 Distribution of Q value in water-entry cavity

北航学报 赠阅

体漩涡。随着开放端排出液体速率开始减小,大气中气旋运动方向改变,通过空泡与大气连接的开口处再次进入空泡内部,并与空泡内部的气体漩涡融合、耗散,涡旋强度不断减弱。在液体吸入阶段,开放端射流形成的气体漩涡迅速耗散消失,空泡内气体漩涡再次向上方运动,此时空泡出现闭合,形成独立的空泡,其内部气体运动相对平稳,最终完全耗散。

空腔周期性涨缩引起的流场扰动使空泡处流场出现扰动,导致泡内空间局部气体出现压力梯度,根据开尔文定理(Kelwin's theorem):

$$\frac{D\boldsymbol{\Gamma}}{Dt}=-\oint_c\frac{1}{\rho}\nabla P\mathrm{d}l+\oint_c\boldsymbol{f}\mathrm{d}l+\oint_c\upsilon\nabla^2\boldsymbol{u}\mathrm{d}l \tag{11}$$

式中:$\boldsymbol{\Gamma}$ 为速度环量,$\boldsymbol{\Gamma}=\nabla\times\boldsymbol{v}$,表征涡旋流动强度,$\boldsymbol{v}$ 为流体速度;$\boldsymbol{f}$ 为体积力;υ 为运动黏度。压力梯度导致泡内出现速度环量,形成气体涡旋。涡旋沿压力梯度作用方向移动,导致气体漩涡通过空泡与大气连通的开口位置往复进出空泡,再次影响空泡界面上的压力分布,阻碍空泡开口收缩,推迟空泡闭合时间。

3.3　流体动力特征分析

入水后壳体受到重力和流体动力的共同作用,其中流体动力主要由压差阻力和黏滞阻力构成,前者表现为流体动压和静压,垂向作用于运动体表面,后者是由流体黏性产生,切向作用于运动体表面,并与相对流速反向。本文中开放空腔壳体垂直入水过程仅存在沿重力方向的单自由度运动,且运动无攻角,具有轴对称特征,因此,轴向作用力是问题研究的重点。

空腔内气体涨缩引起开放端液体往复吸入、排出,对流场结构产生扰动,在一定程度上会影响流体动力特征,图 8 给出了流入空腔的液体质量流量(q)规律,根据腔内液体的运动规律,可将壳体运动过程分为 A 空气域运动阶段,B 液体吸入阶段,C 液体排出阶段。开放端撞水后,液体周期性往返通过,使B、C阶段交替出现,在开放端液体排出末期,即 C-B 过程,质量流量曲线出现波动拐点,特别是首次液体射流末期,由于气体部分外泄,导致出现流入空腔内的液体出现一次波动。与射流过程不同,在液体吸入过程,即 B-C 过程,曲线则相对光滑。

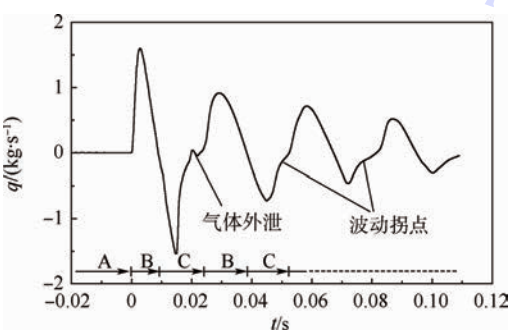


图 8　流入空腔的液体质量流量时程曲线

Fig. 8　Time history of water mass flow rate into cell

图 9 给出了开放端面 S_{open} 和封闭端内壁面 S_{bottom} 上的压力(F)变化规律,具有以下 3 个典型特征:压力出现明显波动规律,2 个面上压力波动具有一定相位差,以及在液体排放阶段出现压力波动反常现象。

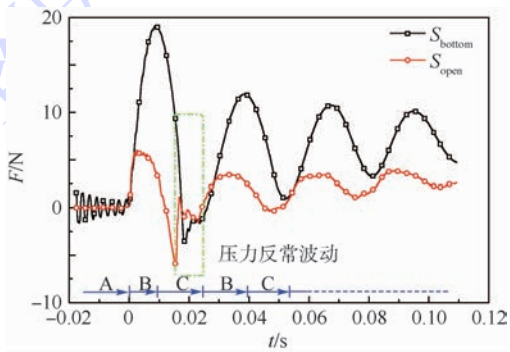


图 9　S_{bottom} 面和 S_{open} 面上压力变化时程曲线

Fig. 9　Time history of pressure on the surface of S_{bottom} and S_{open}

压力波动主要是由于气液两相流体相互作用,引起腔体内气体周期性涨缩和开放端周期性流动,改变了流场结构,使流体动力出现相应的周期性变化。在入水前($t<0$ s),空腔内气体同样伴随涨缩现象,S_{bottom} 上压差阻力出现波动,由于气体惯性较小,压差阻力表现为高频率、低幅值特征,对后续入水阶段的影响微弱;S_{open} 则仅受到恒定的气体动压作用,压差阻力表现平稳。入水后($t>0$ s),由于气体的易压缩性,液体流入空腔内部,相对于气体,液体惯性明显,导致封闭气体涨缩现象更为明显,幅值增加、频率降低。由于 S_{open} 的面积相对于 S_{bottom} 较小,因此作用在 S_{open} 上的力在量值上较小。

波动不同步是由于作用流体介质属性导致,S_{bottom} 仅受气体作用,气体惯性较小,压力主要表现为气体静压,其与气体压缩程度呈正比关系,即 S_{bottom} 压力由空腔压缩量决定;S_{open} 则主要受液体作用,液体惯性较大,压力主要表现为液体动压,其数值大小与流入腔体的液体流速有关,即 S_{open} 压力由空腔压缩速度决定。根据开放端流入液体质量流量的波动变化特征,可知空腔内气体涨缩形成谐振腔,气体压缩速度为气体压缩量的时间导数,压缩速度与压缩量之间具有 $\pi/2$ 相位差。因此,S_{open} 上的压力波动提前于 S_{bottom} 1/4 周期。

北航学报 赠阅

在开放端首次射流末期(C 阶段),壁面压力出现反常波动,如图 9 中虚线框所示,参照图 8 中流入空腔液体质量流量,此阶段部分气体外泄,一方面引起空腔内液体涌动,出现压力梯度(图 5(a)),使 S_{bottom} 上的压力突增,且伴随高频波动;另一方面,溢出的气体改变开放端流场结构,开放端再次封闭后,S_{open} 上将受到液体拌击作用,导致压力突增,表现出压力的反常波动。

图 10 给出了壳体内壁面 S_{inside} 和外壁面 S_{outside} 上的黏滞阻力变化规律。入水前($t<0$ s),近壁面附近的气体流动将引起黏滞作用,腔内气体周期性涨缩形成 S_{inside} 上的波动黏滞阻力。入水后($t>0$ s),液体流入空腔,出现沾湿区域,在液体高黏度作用下黏滞阻力明显增大,由于腔内液面波动、气体环流、外泄等作用,在前 2 个涨缩周期内 S_{inside} 上的黏滞阻力波动规律性较差,此后才表现出明显的周期性,阻力值与开放端流入液体的质量流量成正比关系。S_{outside} 的黏滞阻力受到入水空泡的影响,特别是入水初始阶段(0 ~ 0.016 s),空泡将液体与壳体外壁面隔离,泡内形成低压区域,气体流入产生负向的黏滞力;当腔内气体外泄后,在空泡内形成较为强烈的气体环流,黏滞阻力负向增长,并随气体漩涡的分离、融合运动而出现同步波动;空泡闭合后(0.026 s),外壁面被水沾湿,由于液体黏度远远大于气体,液体的黏滞作用占据主导,但在空泡回流作用下黏滞阻力仍处于负值;空泡持续性增长,并出现多次脱落,导致空泡覆盖面积出现周期性变化,S_{outside} 上黏滞阻力将随其沾湿面积的变化出现波动;当空泡逐渐稳定,覆盖面积相对较小,回射流强度减弱,黏滞阻力出现正值。

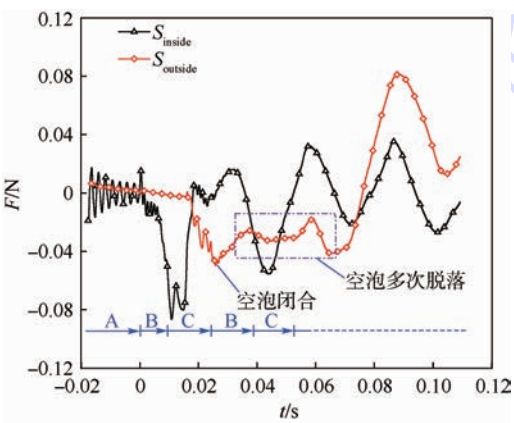


图 10 黏滞阻力变化规律时程图

Fig. 10 Time history of viscous resistance

通常采用无量纲化阻力系数来描述流体动力情况,分别定义压差阻力系数 C_p 和黏滞阻力系数 C_ν 如下:

$$C_p=\frac{\boldsymbol{F}_p}{\frac{1}{2}\rho v_0^2 A} \tag{12}$$

$$C_\nu=\frac{\boldsymbol{F}_\nu}{\frac{1}{2}\rho v_0^2 A} \tag{13}$$

式中:$\boldsymbol{F}_p$ 和 $\boldsymbol{F}_\nu$ 分别为压差阻力和黏滞阻力;本文 $\rho=998.2\ \text{kg/m}^3$;$v_0$ 为特征速度,选取入水速度作为特征速度;A 为特征面积,选取壳体最大截面积 $A=0.25\pi D^2$。

图 11 和图 12 分别给出开放空腔壳体入水过程中压差阻力系数和黏滞阻力系数的时程曲线。压差阻力作为流体动力的主导部分受到空腔气体扰动影响,呈现出周期性变化趋势,特别是气体首次涨缩过程的压力波动幅值尤为明显,并在首次射流末段压差阻力系数出现瞬时负值。黏滞阻力在入水初始阶段主要受到空泡形态、运动等特征的影响,扰动的流场导致空泡产生特殊的演化过程,黏滞阻力系数出现图中较为复杂的规律。在空泡闭合后,黏滞阻力出现波动增加趋势。对比压差阻力系数和黏滞阻力系数,发现在入水过程中压差阻力始终占据主导,黏滞阻力几乎不决定壳体运动及姿态特征。

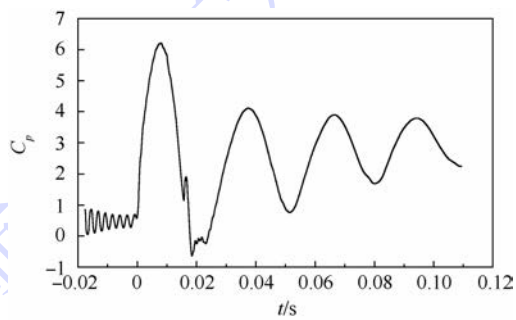


图 11 入水过程压差阻力系数

Fig. 11 Pressure drag coefficients during water-entry

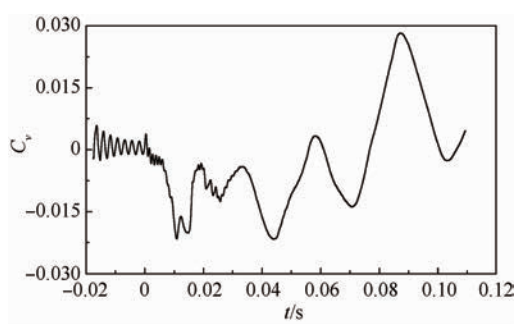


图 12 入水过程黏滞阻力系数

Fig. 12 Viscous drag coefficients during water-entry

北航学报 赠阅

4　结　论

本文针对撞水端敞开的开放空腔壳体垂直入水过程开展了数值仿真,分析了开放空腔壳体入水对流场结构的扰动和壳体流体动力变化规律,获得了压力场、流线和涡量场分布及变化情况,以及壳体各部分受力及整体的阻力特征等规律。具体结论如下:

1)入水前空腔气体便出现了涨缩现象,但扰动极其微弱;入水后液体流入空腔与气体同步运动,在液体惯性作用下空腔涨缩幅度明显增强,在开放端形成压力源,对压力场、速度场产生周期性扰动。

2)空泡受到流场扰动影响沿轴线出现扩展尺度差异,形成空泡波动现象;空泡波动使泡内气体流场结构变化显著,产生速度环量,形成气体漩涡,并随空腔涨缩反复通过空泡与大气连接处进出空泡内部,阻碍、推迟了空泡闭合的发生。

3)扰动的流场产生附加流体动力,使流体动力均呈现波动变化,波动频率与空腔涨缩频率一致,压差阻力远大于黏滞阻力,占据主导作用。封闭端静压波动幅值与压缩量成正比,撞水面动压波动幅值与液体流入速度成正比,两者波动相差 1/4 周期;受空泡和空腔液体流动影响,入水初期内壁面黏滞阻力表现明显,空泡溃灭后外壁面表现明显。

参考文献(References)

[1] DE BACKER G, VANTORRE M, BEELS C, et al. Experimental investigation of water impact on axisymmetric bodies[J]. Applied Ocean Research, 2009, 31(3): 143-156.

[2] MASON D, FOLKMAN S, BEHRING M. Thrust oscillations of the space shuttle solid rocket booster motor during static tests [C]// 15th Joint Propulsion Conference. Reston: AIAA, 1979.

[3] 刘若辰,王英民. 声呐浮标空投入水受力特性仿真研究[J]. 水声工程, 2011, 35(10): 59-62.
LIU R C, WANG Y M. Reserch of water-entry mechanical property of Sonobuoy[J]. Underwater Acoustic Engineering, 2011, 35(10): 59-62(in Chinese).

[4] WORTHINGTON A M. Impact with a liquid surface studied with aid of instantaneous photography[J]. Philosophical Transactions of the Royal Society of London, 1900, 194A: 175-199.

[5] DUCLAUX V, CAILLÉ F, DUEZ C, et al. Dynamics of transient cavities[J]. Journal of Fluid Mechanics, 2007, 591: 1-19.

[6] TRUSCOTT T T, TECHET A H. Water-entry of spinning spheres[J]. Journal of Fluid Mechanics, 2009, 625: 135-165.

[7] ARISTOFF J M, TRUSCOTT T T, TECHET A H, et al. The water entry of decelerating spheres[J]. Physics of Fluids, 2010, 22(3): 417-422.

[8] 何春涛,王聪,何乾坤,等. 圆柱体低速入水空泡试验研究[J]. 物理学报, 2012, 61(13): 134701-134701.
HE C T, WANG C, HE Q K, et al. Low speed water-entry of cylindrical projectile[J]. Acta physica Sinica, 2012, 61(13): 134701-8(in Chiness).

[9] LEE M, LONGORIA R G, WILSON D E. Cavity dynamics in high-speed water entry[J]. Physics of Fluids, 1997, 9(3): 540-550.

[10] HE C T, WANG C, WEI Y J, et al. Numerical simulation of pressure distribution in vertical water-entry cavity[J]. Journal of Ship Mechanics, 2011, 15(9): 960-968.

[11] 陈晨,马庆鹏,魏英杰,等. 空气域压力对高速射弹入水流场影响[J]. 北京航空航天大学学报, 2015, 41(8): 1443-1450.
CHEN C, MA Q P, WEI Y J, et al. Effects of operating pressure on high-speed projectile's water-entry flow[J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(8): 1443-1450(in Chinese).

[12] 赵蛟龙,孙龙泉,张忠宇,等. 柱形空腔结构落水载荷及冲击响应研究[J]. 振动与冲击, 2013, 32(20): 113-118.
ZHAO J L, SUN L Q, ZHANG Z Y, et al. Hydrodynamic loads and impact response for a water entry of a cylindrical cavitary structure[J]. Journal of Vibration and Shock, 2013, 32(20): 113-118(in Chinese).

[13] 潘龙,王焕然,姚尔人,等. 头部喷气平头圆柱体入水缓冲机制研究[J]. 工程热物理学报, 2015, 36(8): 1691-1695.
PAN L, WANG H R, YAO E R, et al. Mechanism reserch on the water-entry impact of the head-jetting flat cylinder[J]. Journal of Engineering Thermophysics, 2015, 36(8): 1691-1695(in Chinese).

[14] 杨衡,孙龙泉,刘莹,等. 波流作用下圆柱体入水特性的三维数值模拟研究[J]. 船舶力学, 2015, 19(10): 1186-1196.
YANG H, SUN L Q, LIU Y, et al. 3D numerical simulation on water entry of cylindrical under wave and stream action[J]. Journal of Ship Mechanics, 2015, 19(10): 1186-1196(in Chinese).

[15] HIRT C W, NICHOLS B D. Volume of fluid (VOF) method for the dynamics of free boundaries[J]. Journal of Computational Physics, 1981, 39(1): 201-225.

[16] JEONG J, HUSSAIN F. On the identification of a vortex[J]. Journal of Fluid Mechanics, 1995, 285(4): 69-94.

作者简介:

路中磊　男,博士研究生。主要研究方向:多相流体动力学数值仿真及试验研究。

E-mail: lu_zhonglei@163.com

魏英杰　男,博士,教授,博士生导师。主要研究方向:多相流体动力学、空气动力学研究。

Tel.: 0451-86413510

E-mail: weiyingjie@gmail.com

北航学报 赠阅

Flow field structure and fluid dynamic characteristics of semi-closed cylinder water-entry

LU Zhonglei, WEI Yingjie*, WANG Cong, SUN Tiezhi

(School of Astronautics, Harbin Institute of Technology, Harbin 150001, China)

Abstract: Numerical simulation for the vertical water-entry process of an end-closed cylinder shell was performed based on reynolds-averaged Navier-Stokes equations. The results of pressure and velocity contours, features of cavity wave and closure, rule of mass flow rate into the cavum, and hydrodynamic change rule were achieved. The influence of air compression-expansion on flow field structure and fluid dynamics was analyzed. The results show that water flow passing the open end in and out of cavum is synchronized with air movement; fluctuant pressure source appears at the open end of shell, leading to periodic pressure and velocity distribution; the expansion diameters of different axisymmetrical cavity sections are inversely proportional to speed of water into cavum; an air vortex generates in cavity, periodically enters and discharges from cavity, and prevents cavity closing; fluid dynamics present fluctuation characteristics whose frequency is equal to that of air, and the amplitude is in proportional relationship with that of air compression-expansion. The air compression-expansion motion in the progress of water-entry by the open cavum structure produces periodic perturbation of flow field and fluid dynamics, reduces impact of water-entry, and enhances the stability of cavity and motion of the an end-closed cylinder shell.

Key words: water-entry; semi-closed cylinder; flow field structure; fluid dynamics; numerical calculation

北京航空航天大学学报

Received: 2015-11-02; **Accepted**: 2015-12-18; **Published online**: 2016-01-14 17:42

URL: www.cnki.net/kcms/detail/11.2625.V.20160114.1742.001.html

Foundation items: National Natural Science Foundation of China (11672094); Natural Science Foundation of Heilongjiang Province of China (A201409); the Fundamental Research Funds for the Central Universities (HIT.NSRIF.201159); Special Foundation for Harbin Science and Technology Innovation Talents of China (2013RFLXJ007)

* **Corresponding author**. Tel.: 0451-86413510 E-mail: weiyingjie@gmail.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0726

ARINC 653 分区实时系统的主时间框架设计

谭龙华，杜承烈*
（西北工业大学 计算机学院，西安 710072）

摘 要：综合模块化航空电子系统中的 ARINC 653 标准规定系统采用分区内调度和分区间调度双层调度机制。根据标准，分区内的调度由分区设计者指定，分区之间则按照静态的主时间框架进行调度。如何基于多个分区应用设计用于分区间调度的主时间框架是系统集成阶段需要解决的问题。首先利用可调度分析导出了分区的有界延迟模型参数，进一步将该参数转化为分区的调度参数用于分区间调度。然后进行分区间调度生成主时间框架，提出了最少窗口数目匹配-最佳匹配(MFBF)算法用于减少分区窗口的切换次数。提出的从分区参数推导到分区间调度流程能够基于若干分区应用生成 ARINC 653 系统的主时间框架。实验结果表明：时间窗口优化算法能有效减少分区窗口切换次数。

关 键 词：实时系统；资源分区；综合模块化航空电子；ARINC 653；主时间框架
中图分类号：V247.1；TP302.1
文献标识码：A **文章编号**：1001-5965(2016)11-2413-10

ARINC 653 标准是美国航空电子工程委员会为新一代综合模块化航电软件制定的操作系统标准规范，包括实时操作系统的行为逻辑规范以及对应用程序的接口规范[1]。规范利用分区概念将应用进行时空隔离，这使得某个分区的错误被控制在分区之内，同时提升了应用的可移植性。

在综合模块化航电软件中，传统的位于分布式节点的应用以分区应用方式集成到综合模块化计算平台上[2]。针对于此，ARINC 653 规范规定系统采用分区内调度和分区间调度双层调度机制。分区内调度一般采用传统嵌入式的实时调度策略，由分区应用的开发者决定；分区之间按照静态主时间框架进行确定性的周期调度，在系统集成层次确定。航电系统作为一类强实时系统，需要研究双层调度机制下的 2 个方面问题：①给定主时间框架分析系统的可调度性；②生成主时间框架对系统进行调度。关于 ARINC 653 系统的可调度性问题已经存在一些研究结果[3-5]。本文研究 ARINC 653 实时操作系统背景下可调度性分析的对立问题：给定若干分区应用，如何生成主时间框架在系统层进行分区间调度。

从分区应用生成主时间框架面临如何从分区应用设计分区以及进一步分区间调度这 2 个阶段问题。何锋等[6]以分区任务执行系数和轮转调度周期作为参数考察分区能力和系统集成问题，给出了关键参数的设计方法。文献[7]推导出了分区内任务集合在单调速率[8](Rate Monotonic，RM)调度下的分区利用率最小上界。

Shin 和 Lee[9]提出了分区的周期资源模型，在该模型中，分区参数包含分区周期和分区执行时间。类似于周期资源模型，Lee 等[10]给出了分区周期和分区利用率满足的关系。借助将其他分区视作当前分区的堵塞，文献[11]同样给出了分区周期和利用率的可行域，并提出了系统低耗能

收稿日期：2015-11-06；**录用日期**：2016-01-08；**网络出版时间**：2016-03-15 14:39
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160315.1439.007.html

* **通讯作者**：Tel.：029-88431545 E-mail：ducl@nwpu.edu.cn

引用格式：*谭龙华，杜承烈. ARINC 653 分区实时系统的主时间框架设计[J]. 北京航空航天大学学报，2016，42(11)：2413-2422.*
TAN L H，DU C L. Design of major time frame for ARINC 653 partitioned real-time systems[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2413-2422 (in Chinese).

北航学报 赠阅

优化算法。Lipari 和 Bini[12]对分区内部采用固定优先级(Fixed Priority,FP)调度的分区应用,从分区内任务集合导出了分区的有界延迟模型[13]参数。对于文献[12]中的问题,文献[14]利用最坏情况分析方法导出了分区参数,该方法较文献[12]中的算法较大幅度提升了运算速度,但仅能给出近似解。

不同于以上介绍的关于分区内任务对分区能力需求的研究,Mok 和 Austin[15]给出了有界延迟分区之间集成的相关结论并且考虑了原子时间片的问题。文献[15]属于分区间集成问题的工作,但是没有涉及分区间调度生成静态主时间窗口框架的问题。

在已有分区参数选择工作的基础上,本文的工作在于考虑分区之间影响的因素下将分区的有界延迟模型参数转化为分区周期和分区执行时间参数,提出了基于这 2 个参数生成主时间框架的算法,从而给出了从若干分区应用生成主时间框架的完整流程。

1 系统模型

1.1 ARINC 653 系统调度模型

ARINC 653 分区实时操作系统规范采用层次调度机制,分为分区间调度和分区内调度。其中分区内调度由分区应用的设计者指定,分区之间按照主时间框架周期地静态调度。主时间窗口框架详细规定了每个分区拥有的时间窗口的起始时间和持续时间。图 1 所示的即是一个主时间窗口框架[1],该图显示了一个共存在 4 个分区的计算模块,主时间框架总长度为 200 ms。在 200 ms 的主时间框架内,P_1 分区获得了 2 个时间窗口,P_2 分区获得了 2 个时间窗口,P_3 分区共获得了 4 个时间窗口,P_4 分区仅有 1 个时间窗口。

分区	P_1	P_3	P_2	P_3	P_1	P_3	P_2	P_3	P_4
偏移/s	0.000	0.020	0.030	0.070	0.100	0.120	0.130	0.170	0.180
持续时间/s	0.020	0.010	0.010	0.010	0.020	0.010	0.010	0.010	0.020

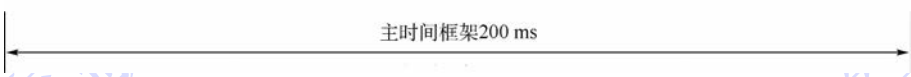


图 1 主时间窗口框架中的分区窗口[1]

Fig. 1 Partition windows within a major time frame[1]

1.2 相关概念定义

ARINC 653 规范限定的适应对象是单核处理器,所以本文下面的探讨范围适用于多个分区在单核处理器上集成的情况。对于多核处理器,在限制应用不可核间迁移情况下,可将其视作单核情况[16]。

一个计算任务表征为一个三元组$\langle C,D,T\rangle$,其中:C 表示周期任务在一个周期内的最坏情况下执行时间(Worst Case Execute Time,WCET),D 表示相对截止时间,如果周期任务在就绪后 D 时间内没有执行结束意味着超时,T 表示周期任务的周期。

一个应用程序 Γ 是一个计算任务集合$\{\tau_1,\tau_2,\cdots,\tau_n\}$,其中:$\tau_i$ 为三元组模型$\langle C,D,T\rangle$描述的任务,该集合内的计算任务被加载于某一 ARINC 653 分区上。一个 ARINC 653 资源分区 Φ 是一个二元组$\langle W,P\rangle$[13]。其中:W 表示一个 n 个时间区间(窗口)集合$\{(s_1,e_1),(s_2,e_2),\cdots,(s_n,e_n)\}$,$P$ 表示分区的周期,并且满足:$0\leqslant s_1<e_1<s_2<e_2<\cdots<s_n<e_n\leqslant P$。分区在时间区间$(s_i,e_i)$内获得资源用于执行分区内任务($e_i$ 和 s_i 为第 i 个时间区间的起点和终点),因此对于一个周期分区,其获得资源的时间区间为$(s_i+j\cdot P,e_i+j\cdot P)$,$1\leqslant i\leqslant n$,$j$ 为自然数。

定义 1 分区利用率[13]

分区 Φ 的分区利用率为

$$\alpha(\Phi)=\sum_{i=1}^{n}(e_i-s_i)/P \tag{1}$$

定义 2 匀速等效执行[13]

假设分区 Φ 所在的处理器的执行速度是 r,它的分区利用率为 α,则它的匀速等效执行是指以独占非打断方式在执行速度 $\alpha\cdot r$ 的处理器上的执行。

定义 3 服务时间函数[13]

服务时间函数 $S(t)$ 是 0 ~ t 时间间隔内分区获得的计算时间总和。

定义 4 分区延迟[15]

分区延迟 λ 是满足不等式(2)的 Δ 最小值。

$\forall t_0,t_1(t_1\geqslant t_0)$

$$\alpha(t_1 - t_0 - \Delta) \leqslant S(t_1) - S(t_0) \leqslant \alpha(t_1 - t_0 + \Delta) \tag{2}$$

定义 5　有界延迟分区[13,15]

有界延迟分区 Φ 是一个二元组$\langle\alpha,\lambda\rangle$。

对于周期为 P、分区利用率为 α 的分区容易得出分区执行时间 $O = P \cdot \alpha$,最坏情况下分区执行窗口延迟 $\lambda_{\max} = 2P \cdot (1-\alpha)$[12]。

定义 6　Harmonic 集合

一个整数集合是 Harmonic 集合当且仅当集合中的任意 2 个元素 e_1 和 e_2 满足式(3)约束关系:

$$e_1 \% e_2 = 0 \ \vee \ e_2 \% e_1 = 0 \tag{3}$$

式中:$e_1 \% e_2$ 表示 e_1 对 e_2 进行取余运算。

在综合模块化航电软件系统中,被调度元素的周期通常被建议配置成具有 Harmonic 关系[4,10,17],以减少分析复杂度。

2　分区参数设计

本节将从分区内的计算任务集合导出分区参数,保证分区内实时任务的可调度性。分区参数的设计过程如图 2 所示,首先利用可调度分析从分区内任务集合参数导出分区有界延迟模型参数:分区利用率 α、最大延迟 λ;然后将参数转化为可实际用于分区配置的分区周期 P、周期内执行时间参数 O,用于分区间调度。需要指出的是,2.1 节中从分区任务集合导出参数 $<\alpha,\lambda>$ 的方法是已有工作[12-13],本文仍将其简要归纳、扩展讨论从而完整给出基于若干分区任务集合设计系统主时间框架的全部过程。本文基于已有工作,在 2.2 节分析讨论了当分区周期满足 Harmonic 关系、分区遵循严格周期调度情况下,模型参数转化的问题。

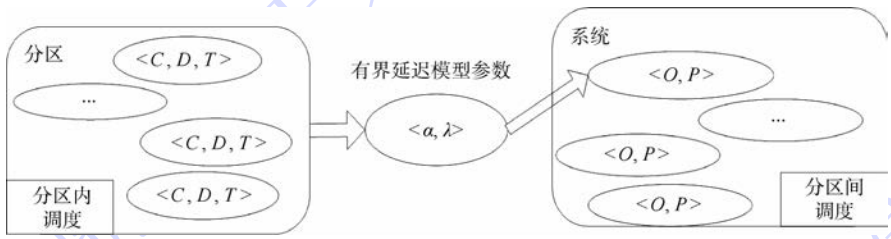


图 2　分区参数设计过程

Fig. 2　Process of partition parameter design

2.1　有界延迟模型参数的可行域

定理 1[13]　给定一个任务集合 $\Gamma = \{\tau_1, \tau_2, \cdots, \tau_n\}$和一个有界延迟分区 $\Phi = (\alpha, \lambda_n)$,设 S_n 是任务集合 Γ 在有界延迟分区 Φ 的等价匀速执行处理器上的一个调度,S_p 是 Γ 在有界延迟分区 $\Phi = (\alpha, \lambda_n)$上和 S_n 的调度次序和资源分配量都相同的一个调度。设 λ 是保证 S_n 中所有任务都能够在截止期前 λ 时间完成计算时取得的最大值,那么 S_p 可调度当且仅当 $\lambda > \lambda_n$。

该定理指出了有界延迟分区参数的可调度条件,以此可以推导出相应的参数计算公式:

$$\min_{i=1,2,\cdots,n} \{D_i - R_i\} > \lambda \tag{4}$$

式中:D_i 为任务 τ_i 的截止期;R_i 为响应时间。

对于采用最早截止时间优先(Earliest Deadline First,EDF)调度的分区,计算某一任务的最大响应时间比较复杂[18]。文献[13]指出可以采用将所有计算任务的截止期提前 λ 后验证可调度性的方法确定延迟分区的最大延迟参数是否能够取值为 λ。例如,借助文献[19]中提出的快速收敛分算法(Quick convergence Processor-demans Analysis,QPA),可以二分搜索得出最大允许的延迟 $\lambda_{\max}$。

对于采用 FP 调度的分区,Lipari 和 Bini[12]给出了分区延迟的计算公式(5)。给定分区利用率,借助式(5)可计算出能够保证分区内任务集合可调度时分区服务器允许的最大延迟。

$$\lambda \leqslant \min_{i=1,2,\cdots,n} \{\lambda_i\}$$

$$\begin{cases} \lambda_i = \max\limits_{t \in P_{i-1}(D_i)} \left\{ t - \dfrac{1}{\alpha}\left(C_i + \sum\limits_{j=1}^{i-1} \left\lceil \dfrac{t}{T_j} \right\rceil C_j \right) \right\} \\ P_0(t) = \{t\} \\ P_i(t) = P_{i-1}\left(\left\lfloor \dfrac{t}{T_i} \right\rfloor T_i \right) \cup P_{i-1}(t) \end{cases} \tag{5}$$

例 1　设分区内的计算任务参数如表 1 所示,分区内部采用 RM 调度算法调度。图 3 显示出当分区利用率取不同值时,各个任务可调度时允许的最大分区延迟。例如,当分区利用率为

表 1　任务参数

Table 1　Parameters of tasks

任务编号	最坏执行时间/ms	截止期/ms	周期/ms
T_1	1	6	6
T_2	1	10	10
T_3	3	20	20

北航学报 赠阅

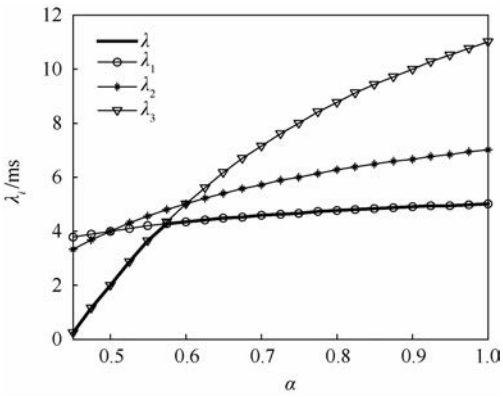


图 3 不同分区利用率下允许的最大延迟

Fig. 3 Maximum feasible delay with different partition utilizations

0.6 时,任务 T_1 具有最高的优先级,任务 T_3 具有最低的优先级。按照优先级从高至低顺序,对于任务 T_1,首先将 $D_1=6$ 代入计算得出集合 $P_0(6)=\{6\}$,进一步可以根据 λ_i 计算公式计算出分区允许延迟时间 $\lambda_1=4.33$;类似地,对于任务 T_2 和任务 T_3,分别可以计算出其允许延迟时间 $\lambda_2=5$,$\lambda_3=5$。最终,根据计算公式(5)可以计算该分区服务器在分区利用率为 0.6 时允许的最大延迟:

$$\lambda_{\max}=\min_{i=1,2,3}\{\lambda_i\}=\min\{4.33,5,5\}=4.33$$

图 3 显示出当选择分区利用率 α 为自变量时,在能够保证分区内任务集合可调度条件下,分区允许的最大延迟可以表征为分区利用率 α 的函数[20],即 $\lambda_{\max}=f(\alpha)$。因此,有界延迟分区参数可行域可以表征为区域集合 $\{(\alpha,\lambda)\mid b_l\leqslant\alpha\leqslant b_r,0\leqslant\lambda\leqslant f(\alpha)\}$,其中 b_r、b_l 表征分区利用率上下界。下面的推论说明函数 $\lambda_{\max}=f(\alpha)$ 具有单调递增性。

推论 1 设 $\lambda_{\max}=f(\alpha)$ 是分区允许的最大延迟关于分区利用率 α 的函数,则该函数在可行区间内单调递增。

证明 设原分区利用率为 α_0,现将分区利用率提升至 $\alpha_1>\alpha_0$,其匀速等效执行速度增大 α_1/α_0 倍。考察分区的匀速等效执行系统。

假设分区内采用 EDF 调度,所有任务周期的最小公倍数为 L,考察 $[0,L]$ 区间内各个任务的各次周期执行的结束时刻组成的有穷集合 $\{f_1,f_2,\cdots,f_i,\cdots,f_n\}$。对于第一个时刻 f_1,其对应某个任务 τ_i 的第一次执行结束,执行区间为 $[s_1,f_1]$,且 $C_i=f_1-s_1$,其中 s_1 是执行开始时间。当处理器速度增大 α_1/α_0 倍时,在 EDF 调度下,因为所有任务的周期和截止期不变,所以任务 τ_i 在 $[s_1,f_1]$ 内仍是最高优先级任务,容易得出其执行结束时刻

$$f_1'=s_1+C_i/(\alpha_1/\alpha_0)<f_1=s_1+C_i$$

假设对于前 k 个结束时刻结束执行的任务,当处理器速度变大 α_1/α_0 倍时响应时间缩短,现考察第 $k+1$ 个执行结束时刻。

设 f_{k+1} 对应任务 τ_x 的第 j 次执行,设其释放至截止期区间 $(r_{x,j},d_{x,j}]$ 内的执行时间区间为

$$[s_1,e_1]\cup[s_2,e_2]\cup\cdots\cup[s_n,e_n]$$

并且满足:

$$C_x=\sum_{l=1}^{n}e_l-s_l$$

其中:C_x 为 τ_x 的执行时间,出现多个执行区间是因为可能被更高优先级任务抢占。处理器速度增大 α_1/α_0 倍后,因为所有任务周期的周期和截止期不变,所以任务 τ_x 在原执行时间区间 $[s_1,e_1]\cup[s_2,e_2]\cup\cdots\cup[s_n,e_n]$ 内仍是最高优先级任务。此时,在 e_n 时刻,处理器至少已经分配了 C_x 时间用于 τ_x 的第 j 次执行,由于 $C_x/(\alpha_1/\alpha_0)<C_x$,所以 τ_x 的第 j 次执行在 e_n 时刻前已经结束,即响应时间缩短。结论对于第 $k+1$ 个时刻结束执行的任务仍然成立。

由于所有任务的响应时间缩短,由定理 1,可以得出保证分区内部可调度的最大分区延迟增大。采用 FP 调度策略的系统类似可证。

该性质的证明过程表明,在单核情况下,当任务集合被限制在分区服务器提供的时间片内执行时,增大分区服务器的分区利用率有助于缩短分区内计算任务的响应时间,进而能够容许更大的分区服务器延迟。

2.2 参数变换与参数选择

有界延迟模型参数能够描述分区内任务对分区的能力需求,但无法用于分区间调度。本节基于分析参数的可行域性质,将有界延迟模型参数〈α,λ〉转化为分区服务器的调度参数〈O,P〉。

2.2.1 固定偏移执行的分区服务器

引理 1 设分区服务器集合的调度参数为 $\{\langle O_1,P_1\rangle,\langle O_2,P_2\rangle,\cdots,\langle O_n,P_n\rangle\}$,其中周期参数 $\{P_1\leqslant P_2\leqslant\cdots\leqslant P_n\}$ 满足 Harmonic 关系。若采用 RM 调度算法进行分区间调度,每个分区将固定偏移执行,分区窗口具有周期性。

证明 由于不同的分区周期满足 Harmonic 关系,$P_i\%P_j=0\vee P_j\%P_i=0$,当 $P_i\leqslant P_j$ 时,易得 $P_j\%P_i=0$。即第 $i+1$ 个分区的周期是前 i 个分区周期的公倍数。

根据 RM 调度算法,周期更小的分区有更高

北航学报 赠阅

的优先级。第 1 个分区被调度时,处理器空闲,容易得知其将固定偏移执行,分区窗口具有周期性。

假设前 i 个分区的执行时间具有周期性。对于第 $i+1$ 个分区,由于 P_{i+1} 是前 i 个分区周期的公倍数,且前 i 个分区周期执行,所以它也将在前 i 个分区执行后形成的空闲时间区间内被固定偏移执行,分区窗口具有周期性执行特征。所以,前 $i+1$ 个分区组合成以 P_{i+1} 为周期的调度。根据数学归纳法,易得所有分区的将固定偏移执行,分区窗口具有周期性。

由于每个分区将被固定偏移执行,分区窗口具有周期性,记 $S(t)$ 为分区资源服务时间函数,由该引理容易得出如下关系:

$$S(t+kP)=S(t)+kO$$

其中:$k\in\mathbf{N}$。

2.2.2　分区周期的可行区间

文献[12]指出,若干周期分区服务器被调度时,因调度引起的延迟最大为 $\lambda_{\text{server}}=2P(1-\alpha)$。实际上,对于严格周期执行的分区服务器,由于分区间调度引起的延迟可以通过更精确的分析限制为 $\lambda_{\text{server}}=P(1-\alpha)$。通过更小的分区延迟,分区服务器可以在更小的分区利用率情况下使得分区内的任务可调度。该结论优化自文献[12],因其对于后续的分区周期参数选择有重要作用,将其作为定理 2 列出,并给出证明。

定理 2　分区周期满足 Harmonic 关系的若干分区在采用 RM 调度算法调度时,分区由于调度引起的延迟 $\lambda_{\text{server}}\leqslant P(1-\alpha)$。

证明　假设某个分区在 t 时刻获得的执行时间为 $S(t)$,当分区在某种调度下的分布窗口使得区间 $[t_0,t_1]$ $(t_1>t_0)$ 内获得的资源 $S(t_1)-S(t_0)$ 最少,不等式 $\alpha(t_1-t_0-\lambda)\leqslant S(t_1)-S(t_0)$ 两侧取等号时,资源分区延迟 λ 取最大值,即

$$\alpha(t_1-t_0-\lambda)=S(t_1)-S(t_0)$$

设 $t_1-t_0=d+kP$,$d\in\mathbf{N}\wedge d\in[0,P)$,$k\in\mathbf{N}$,由 $S(t+kP)=S(t)+kO$ 得

$$\alpha(t_1-t_0-\lambda)=S(t_1)-S(t_0)\Leftrightarrow$$
$$\alpha(d+kP-\lambda)=S(t_0+d+kP)-S(t_0)\Leftrightarrow$$
$$\alpha(d-\lambda)+\alpha kP=S(t_0+d)+\alpha kP-S(t_0)\Leftrightarrow$$
$$\alpha(d-\lambda)=S(t_0+d)-S(t_0)\tag{6}$$

假设 $\lambda>P(1-\alpha)$ 成立,则

$$\alpha(d-\lambda)=S(t_0+d)-S(t_0)\geqslant 0\Rightarrow$$
$$d-\lambda\geqslant 0\Rightarrow$$
$$d\geqslant\lambda>P(1-\alpha)\tag{7}$$

即 $d>P(1-\alpha)$。

又由 $P>d>P(1-\alpha)$,$S(t+P)=P\alpha+S(t)$,利用鸽巢原理得

$$S(t_0+d)-S(t_0)\geqslant$$
$$t_0+d-t_0-P(1-\alpha)=$$
$$d-P(1-\alpha)\tag{8}$$

进一步联合式(6)的结论 $\alpha(d-\lambda)=S(t_0+d)-S(t_0)$ 容易得出

$$\alpha(d-\lambda)\geqslant d-P(1-\alpha)\Leftrightarrow(P-d)(1-\alpha)\geqslant\alpha\lambda\tag{9}$$

由假设 $\lambda>P(1-\alpha)$,从式(9)容易得出

$$(P-d)(1-\alpha)\geqslant\alpha\lambda\geqslant\alpha P(1-\alpha)\Leftrightarrow$$
$$P-d\geqslant\alpha P\Leftrightarrow$$
$$P(1-\alpha)\geqslant d\tag{10}$$

式(10)结论 $P(1-\alpha)\geqslant d$ 与式(7)已推出的结论 $d>P(1-\alpha)$ 矛盾,所以假设错误,即分区延迟参数 λ 不可能超过 $P(1-\alpha)$。

对于优先级最高的分区,当 $t_1=P$,$t_0=P\alpha$ 时,$\alpha(P-P\alpha-\lambda)=S(P)-S(P\alpha)=P\alpha-P\alpha=0$。此时分区延迟参数 λ 取得最大值 $P(1-\alpha)$。

该定理指出在分区服务器被固定偏移执行、分区窗口具有周期性时,分区服务器能够保证提供的资源延迟 λ_{server} 不超过 $P(1-\alpha)$。

由推论 1,当分区的利用率为 α 时,能够保证分区内任务集合可调度的分区服务器最大延迟为 $\lambda_{\max}=f(\alpha)$。为使分区服务器满足分区内任务集合的可调度要求,应保证分区服务器的供给能力不小于需求能力,即

$$\lambda_{\text{server}}\leqslant P(1-\alpha)\leqslant f(\alpha)=\lambda_{\max}$$

变换可得

$$P\leqslant\frac{f(\alpha)}{1-\alpha}\tag{11}$$

式(11)给出了分区周期 P 的上界。推论 1 指出 $\lambda_{\max}=f(\alpha)$ 具有单调递增性,$1/(1-\alpha)$ 在 $(0,1)$ 区间内也单调递增,所以式(11)右侧在区间 $(0,1)$ 内单调递增。因此,分区周期的最大值与分区的利用率最大值相关,当分区利用率取最大值 $\alpha_{\max}$ 时,分区周期取得最大值 $P_{\max}$。

对于存在多个分区的系统,单个分区利用率的上界是能够保证其他分区内的任务可调度,因此,对于某个分区 i,它的利用率取值范围至多为

$$\alpha_i\in\left[\sum_{l=1}^{m}\alpha_{i,l},1-\sum_{j=1,i\neq j}^{n}\alpha_j\right]\tag{12}$$

式中:m 为分区 i 内的任务个数;$\alpha_{i,l}$ 为分区内任务 l 的利用率;n 为系统内分区个数;α_j 为分区 j 的分区利用率。此时,分区周期最大值 $P_{\max}$ 可由式(11)和式(12)计算:

北航学报 赠阅

$$P_{\max}=\frac{f(\alpha_{\max})}{1-\alpha_{\max}}$$

分区周期与分区的切换次数正相关,在文献[21]中,通过增大分区利用率的方式获得更大的分区周期参数,以实现减少分区切换次数的目的。将该结论应用于本文时,可以限制分区周期不小于某一最小值 $P_{\min}$,从而限制分区的切换次数。此时可以确定单个分区周期的可能取值区间至多为$[P_{\min},P_{\max}]$。

2.2.3 分区调度参数选择

已知分区周期 P,若分区执行时间 O 能够使得分区内的任务集合可调度,则对于相同的周期和执行时间 $O'>O$,同样能够使得任务集合可调度。这一性质使得在分区周期固定时,可以利用二分搜索算法搜索出使得分区内部任务集合可调度的最小分区执行时间参数 $O\in[0,P]$。借助式(12)给出的分区利用率约束,可以进一步缩小参数 O 的搜索范围。

在周期资源模型参数选择时,可依据不同的系统设计目标进行相应的周期参数、执行时间参数选择。如果选择分区利用率之和最小使得系统具有更强的可扩展能力为优化目标,可通过遍历方法搜索出最优的分区周期和执行时间参数,分区参数选择伪代码如下:

```
Pseudocode 1 Parameter Selection
1: O = [ ]
2: P = [ ] // O and P are target parameters
3: U_min = +∞
4: for each v ∈ P_s do
5:    U = 0
6:    for each P_i ∈ v do
7:       O_l = O_min, O_r = O_max
8:       while O_l < O_r do
9:          O_mid = (O_l + O_r)/2
10:            if < O_mid, P_i > is schedulable then
11:               O_r = O_mid
12:            else
13:               O_l = O_mid
14:            end if
15:      end while
16:      U = U + O_l/P_i
17:   end for
18:   if U < U_min ∧ U < 1 then
19:         update (O, P, U_min)
20:   end if
21: end for
```

根据2.2.2节得出的分区周期取值范围,利用搜索算法可以得出所有分区周期满足Harmonic关系的分区周期向量集合。记所有可能的Harmonic分区向量组成的集合(代码第4行)为

$$P_s=\{\langle P_1,P_2,\cdots,P_i,\cdots,P_n\rangle \mid P_i\%P_j=0\vee P_j\%P_i=0\}$$

算法第4~21行的循环用于搜索所有可能的分区周期取值,其中第6~17行的代码对一个分区进行二分搜索最小执行时间。第18~20行更新搜索出的最新结果。

对于某个分区服务器的周期 P,需二分搜索一次最小执行时间,对应的时间复杂度为 $O(P\text{lb}\ P)\cdot O(\text{SchTest})$,其中 $O(\text{SchTest})$ 为用于检测分区是否可调度的某种算法的复杂度(如QPA[19])。对于所有可能的周期取值,共需要搜索 $|P_s|$ 次,因此对于 n 个分区的系统,该算法的最终复杂度为:$O(n|P_s|P\text{lb}P)\cdot O(\text{SchTest})$。

3 主时间窗口框架设计

确定每个分区的周期和执行时间后,分区间调度问题可以转化为传统实时任务调度问题。引理1表明,当分区周期满足Harmonic关系且采用RM调度算法进行分区间调度时,每个分区将被固定偏移执行,分区窗口具有严格周期性。为了减少分区切换次数,本文提出最少窗口数目匹配-最佳匹配(Minimum number of windows Fit-Best Fit, MFBF)算法分配调度时间片。

MFBF算法按照分区优先级顺序进行分区调度,首先根据RM算法指定的分区优先级对所有分区进行由高到低排序。其中优先级相同的分区按照执行时间由小到大排列优先级,对于此时优先级仍相同的分区,按照出现顺序赋值优先级。

对于第 i 个分区,在一个分区周期内,分区所需要的执行时间是 O,分区所能获得的空闲处理器时间是一个区间或者多个区间的并集。现需要将 O 分配在这若干个区间上。如果不存在一个区间能够满足执行时间 O,依次从这多个区间中按照区间长度由大到小顺序依次将 O 分配在这多个区间中(对应伪代码9~15行)。在选择最后一个的区间时,选择能够满足剩余执行时间且区间长度最小的区间(对应伪代码16~18行),即最匹配的区间。MFBF匹配算法伪代码如下所示:

Pseudocode 2 MFBF

1: $S_p = \{S_{p1}, S_{p2}, \cdots, S_{pn}\}$
　// partition set, priority desc
2: for each $i = 1: n$ do
3:　P_i = period of S_{pi}
4:　O_i = capacity of S_{pi}
　　// scheduling parameters of partition S_{pi}
5:　$S_f = \{[l_1, r_1], [l_2, r_2], \cdots. [l_m, r_m]\}$
6:　$L = [r_1 + 1 - l_1, r_2 + 1 - l_2, \cdots, r_m + 1 - l_m]$
　　// free intervals in $[0, P_i)$
7:　sort S_f, L as length of intervals desc
8:　index = 1
9:　while $O_i > L$ (index) do
10:　　$O_i = O_i - L$ (index)
11:　　allocate S_f(index) to partition S_{pi}
12:　　delete L (index) from L
13:　　delete S_f(index) from S_f
14:　　index = index + 1
15:　end while
16:　[l, r] = available interval with minimum length in S_f
17:　allocate$[r + 1 - O_i, r]$ to partition S_{pi}
18:　delete$[r + 1 - O_i, r]$ from S_f
19: end for

因为分配时间片时按照优先级由高到低进行,所以该算法能够保证分区周期执行。除此以外,由于每次分配时间片时使用尽可能少的分区窗口完成一次分区服务,该算法能在一定程度上减少分区切换次数。下文 4.2 节中的实验结果表明,算法能够减少分区的切换次数,提升系统效率。

对于一组周期升序排列且满足 Harmonic 关系的周期资源模型分区$\{\langle O_1, P_1\rangle, \langle O_2, P_2\rangle, \cdots, \langle O_n, P_n\rangle\}$,可利用 MFBF 调度算法进行调度形成一个以 P_n 为周期长度的离线调度结果。该离线调度结果可用于生成静态的主时间框架。

离线调度结果保存了处理器的时间片分配信息。一个分区获得的一个连续时间片对应该分区服务器的一个时间窗口。遍历长度为 P_n 的时间片分配方案,可以得出每个时间窗口所属的分区、开始时间和持续时间信息。这些信息可以组成类似于图 1 所示的 ARINC 653 分区实时系统的主时间框架。进一步地,可将主时间窗口框架按照规定格式记录在 XML 文件中[1],用于平台部署,最终完成 ARINC 653 分区实时系统主时间窗口配置的全部过程。

4　案例与实验分析

4.1　案例分析

本节通过一个案例分析演示从分区应用集合出发生成系统的主时间框架的完整过程。设系统中的分区以及分区内的任务参数如表 2 所示。任务的调度参数表示为$\langle C, D, T\rangle$三元组,且每个分区内部按照 RM 调度算法调度。

表 2　系统中的分区及分区内任务参数

Table 2　Partitions within system and task parameters within partitions

ms

分区 P_1 中任务 $\langle C, D, T\rangle$	分区 P_2 中任务 $\langle C, D, T\rangle$	分区 P_3 中任务 $\langle C, D, T\rangle$
$\langle 2,20,20\rangle$	$\langle 4,40,40\rangle$	$\langle 5,40,40\rangle$
$\langle 4,25,25\rangle$	$\langle 4,50,50\rangle$	$\langle 6,100,100\rangle$
$\langle 1,50,50\rangle$		$\langle 5,200,200\rangle$

首先,根据式(12),可以计算得出分区 P_1 的利用率最小值为分区内任务利用率之和 $\alpha_{min} = 2/20 + 4/25 + 1/50 = 0.28$,分区 P_2 和 P_3 分区利用率最小值为 0.18 和 0.21。为保证分区 P_2 和 P_3 可调度,分区 P_1 的分区利用率至多为 $1 - 0.18 - 0.21 = 0.61$。同理,最终可计算得 3 个分区的利用率取值范围分别是[0.28, 0.61], [0.18, 0.51], [0.21, 0.54]。

对于每个分区,当分区利用率取最大值时,分区内任务集合可以承受的最大分区服务器延迟 $\lambda_{max} = f(\alpha_{max})$。类似于例 1 计算过程,当分区利用率取最大值[0.61, 0.51, 0.54]时,可计算得出分区内任务集合能够承受的最大分区服务器延迟:[11.89, 26.47, 30.74] ms。进一步将分区利用率取值上界代入式(11)得

$$P_{max}(\alpha) = \frac{f(\alpha_{max})}{1 - \alpha_{max}}$$

计算得出的分区周期的上界是[30, 54, 66] ms。如果为控制分区切换次数,依次选择 3 个分区的周期不小于[10, 10, 20] ms,此时可以得出 3 个分区服务器的周期取值范围 I = [[10, 34], [10, 54], [20, 66]]。

在 3 个分区允许的周期范围内,存在[10, 10, 20] ~ [30, 30, 60]共 113 个满足 Harmonic 关系的周期选择方案。当选择系统利用率最小作为优化目标时,利用 2.2.3 节中搜索算法且结果保留 1 位小数精度,最终得出的分区执行时间-分区周期参数为:$\langle 4.2, 10\rangle$, $\langle 2.5, 10\rangle$和$\langle 5.0, 20\rangle$,此时对应的系统利用率为 0.92。将分区调

北航学报 赠阅

度参数代入第 3 节给出的 MFBF 调度算法后，得出如图 4 所示的主周期长度为 20 ms 的分区调度结果。

将调度结果统一偏移调度开始的空闲分区后得出主时间框架如表 3 所示。

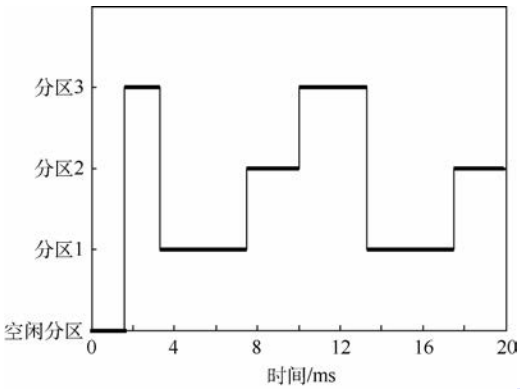


图 4 案例系统的分区调度结果

Fig. 4 Scheduling result of partitions within example system

表 3 案例系统的主时间框架

Table 3 Major time frame of example system

窗口 id	窗口起始时间/ms	窗口持续时间/ms
3.1	0	1.7
1.1	1.7	4.2
2.1	5.9	2.5
3.2	8.4	3.3
1.2	11.7	4.2
2.2	15.9	2.5

4.2 窗口微调算法的性能试验

为了对第 3 节中提出的分区窗口匹配算法 MFBF 进行实验测试，按照 UUFast[22] 算法随机产生分区的利用率，并随机选择分区周期进行实验。为避免分区利用率过小导致在分区最小执行时间为 1(原子时间)时周期特别大的情况，对于分区利用率小于 0.005 的分区重新随机生成分区利用率。

由于分区的周期需要满足 Harmonic 关系，实验采用如下算法随机生成一组满足 Harmonic 关系的分区周期。

1) 从[2, 10]区间内的正整数中，等概率随机选择一个整数 B 作为基数。

2) 对于每个待生成的分区周期，首先从[10, 10 000]区间内的正整数中，等概率随机选择一个正整数 P 作为分区周期。

3) 按照文献[23]中 Harmonic 关系调整计算公式进行分区周期 Harmonic 调整，作为最后的分区周期，同样记作 P。计算公式为

$$P = B \cdot 2^{\lfloor \mathrm{lb}(P/B) \rfloor} \tag{13}$$

式中：等号右侧的 B 和 P 参数分别为步骤 1)和步骤 2)随机生成的基数与周期。

4) 将调整后的周期 P 乘以 UUFast 算法生成的该分区的分区利用率能够生成该分区的执行时间 O，如果 O 小于 1(原子时间)，即出现了最小执行时间不足 1 的情况，转步骤 2)，重新随机生成周期。

由于 UUFast 算法能够保证分区的利用率等概率均匀分布，上述的分区周期生成算法用到的基数 B 和初始周期 P 也都是等概率随机选择生成，所以经过式(13)进行 Harmonic 关系调整后最终的分区执行时间-周期参数能够保持一定的随机特征。

产生的执行时间-周期参数作为 MFBF 算法输入，将其调度结果与 RM 算法调度结果进行比较。为考察算法在不同系统负载情况下的运行效果，实验测试了分区数为 3、6 和 9，分区利用率之和为 0.3 ~ 1.0 之间的系统。在每次实验中，对于给定的分区数目和系统所有分区利用率之和，随机生成调度参数，实验 100 次取平均结果。MFBF 算法的分区切换次数优化效果如图 5 所示，其中横坐标 U 为所有分区利用率之和，纵坐标 R 为 MFBF 算法的分区切换次数与 RM 算法分区切换次数的比值。如图 5 中所示，MFBF 算法在系统不同负载的情况下均能减少分区切换次数，提升系统效率。

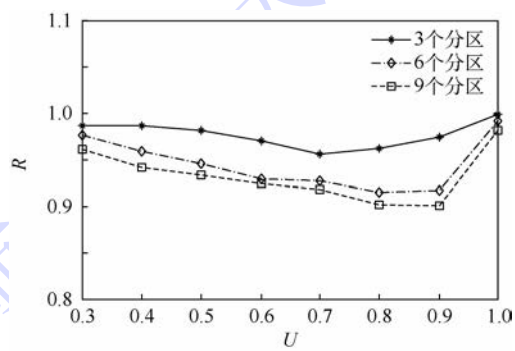


图 5 MFBF 算法和 RM 算法的分区切换次数比率

Fig. 5 Ratio of partition switch times between MFBF algorithm and RM algorithm

5 结 论

1) 本文中描述的主时间框架设计流程能够从分区任务集合的调度参数出发生成系统主时间框架。在最终生成的主时间框架中，各分区服务器周期执行，分区的周期满足 Harmonic 关系。分区服务器提供的计算时间资源能够保证分区内的任务集合可调度。

北航学报 赠阅

2) MFBF 算法的调度结果能够使得分区服务器周期执行。算法通过对分区时间窗口进行优化配置减少时间窗口切换次数,实验结果表明,MFBF 算法能够在不同的系统负载情况下减少分区时间窗口切换次数,提升系统效率。

参考文献(References)

[1] Airlines Electronic Engineering Committee. ARINC specification, avionics application software standard interface:ARINC 653-2[S]. Annapolis, MA:Aeronautical Radio, Inc., 2006:1-205.

[2] 褚文奎,张凤鸣,樊晓光. 综合模块化航空电子系统软件体系结构综述[J]. 航空学报,2009,30(10):1912-1917.
CHU W K, ZHANG F M, FAN X G. Overview on software architecture of integrated modular avionic systems[J]. Acta Aeronautica et Astronautica Sinica, 2009, 30(10):1912-1917(in Chinese).

[3] 高晓光,薛亚勇,温增葵. IMA 双层调度算法中的任务可调度性分析方法[J]. 航空学报,2015,36(2):585-595.
GAO X G, XUE Y Y, WEN Z K. Task schedulability analyzing method of two-level hierarchical scheduling algorithm in integrated modular avionics[J]. Acta Aeronautica et Astronautica Sinica, 2015, 36(2):585-595(in Chinese).

[4] ARVIND E, INSUP L, OLEG S, et al. A compositional framework for avionics (ARINC-653) systems: MS-CIS-09-04[R]. Philadelphia: Department of Computer & Information Science, University of Pennsylvania, 2009.

[5] CHANG Y, DAVIS R, WELLINGS A. Schedulability analysis for a real-time multiprocessor system based on service contracts and resource partitioning: YCS-432[R]. York: Department of Computer Science, University of York, 2008.

[6] 何锋,宋丽茹,熊华钢. 航空电子双层任务分区调度设计[J]. 北京航空航天大学学报,2008,34(11):1364-1368.
HE F, SONG L R, XIONG H G. Two-level task partition scheduling design in integrated modular avionics[J]. Journal of Beijing University of Aeronautics and Astronautics, 2008, 34(11):1364-1368(in Chinese).

[7] 李昕颖,顾健,何锋,等. 硬实时系统在强分区约束下的双层分区调度[J]. 计算机学报,2010,33(6):1032-1039.
LI X Y, GU J, HE F, et al. Two-level partition scheduling in hard real time system under strong partition constraints[J]. Chinese Journal of Computers, 2010, 33(6):1032-1039(in Chinese).

[8] LIU C L, LAYLAND J W. Scheduling algorithms for multiprogramming in a hard-real-time environment[J]. Journal of the ACM, 1973, 20(1):46-61.

[9] SHIN I, LEE I. Periodic resource model for compositional real-time guarantees[C]// Proceedings of 24th IEEE Real-Time Systems Symposium. Piscataway, NJ: IEEE Press, 2003:2-13.

[10] LEE Y H, KIM D, YOUNIS M, et al. Partition scheduling in APEX runtime environment for embedded avionics software [C]// Proceedings of 5th International Conference on Real-Time Computing Systems and Applications. Piscataway, NJ: IEEE Press, 1998:103-109.

[11] ZHOU T, XIONG H. Design of energy-efficient hierarchical scheduling for integrated modular avionics systems[J]. Chinese Journal of Aeronautics, 2012, 25(1):109-114.

[12] LIPARI G, BINI E. Resource partitioning among real-time applications[C]// Proceedings of 15th Euromicro Conference on Real-Time Systems. Piscataway, NJ: IEEE Press, 2003:151-158.

[13] MOK A K, FENG X, CHEN D. Resource partition for real-time systems[C]// Proceedings of Real-Time Technology and Applications Symposium. Piscataway, NJ: IEEE Press, 2001:75-84.

[14] ALMEIDA L, PEDREIRAS P. Scheduling within temporal partitions: Response-time analysis and server design[C]// Proceedings of 4th ACM International Conference on Embedded Software. New York: ACM Press, 2004:95-103.

[15] MOK A K, AUSTIN U O T A. A model of hierarchical real-time virtual resources[C]// Proceedings of 23rd Real-Time Systems Symposium. Piscataway, NJ: IEEE Press, 2002:26-35.

[16] DAVIS R I, BURNS A. A survey of hard real-time scheduling for multiprocessor systems[J]. ACM Computing Surveys(CSUR), 2011, 43(4):1-35.

[17] LEE Y H, KIM D, YOUNIS M, et al. Resource scheduling in dependable integrated modular avionics[C]// Proceedings of International Conference on Dependable Systems and Networks. Piscataway, NJ: IEEE Press, 2000:14-23.

[18] MARCO S. Analysis of deadline scheduled real-time systems: Technical report RR-2772[R]. Pairs: INRIA, 1996.

[19] ZHANG F, BURNS A. Schedulability analysis for real-time systems with EDF scheduling[J]. IEEE Transactions on Computers, 2009, 58(9):1250-1258.

[20] BUTTAZZO G, BINI E, WU Y F. Partitioning real-time applications over multicore reservations[J]. IEEE Transactions on Industrial Informatics, 2011, 7(2):302-315.

[21] PATHAN R M, STENSTROM P, GREEN L G, et al. Overhead-aware temporal partitioning on multicore processors[C]// Proceedings of the IEEE 20th Real-Time and Embedded Technology and Applications Symposium. Piscataway, NJ: IEEE Press, 2014:251-262.

[22] BINI E, BUTTAZZO G. Biasing effects in schedulability measures[C]// Proceedings of 16th Euromicro Conference on Real-Time Systems. Piscataway, NJ: IEEE Press, 2004:196-203.

[23] CHAN M Y, CHIN F. Schedulers for larger classes of pinwheel instances[J]. Algorithmica, 1993, 9(5):425-462.

作者简介:

谭龙华　男,博士研究生。主要研究方向:综合模块化航空电子系统、实时控制系统。

E-mail: tanlonghua@mail.nwpu.edu.cn

杜承烈　男,博士,教授,博士生导师。主要研究方向:实时分布计算、信息物理融合系统的设计与验证、安全关键领域软件工程。

Tel.: 029-88431545

E-mail: ducl@nwpu.edu.cnm

北航学报 赠阅

Design of major time frame for ARINC 653 partitioned real-time systems

TAN Longhua, DU Chenglie*

(School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: The ARINC 653 standard for integrated modular avionics defines that the system adopts intra-partition scheduling and inter-partition scheduling. As specified by the standard, local scheduling is specified by partition application designers and partitions are scheduled statically by the major time frame. How to design the major time frame to schedule partitions is a problem in the phase of partition integrating. The bounded delay parameters of a partition are derived based on schedulability analysis and converted to parameters which can be used for partition scheduling afterwards. An algorithm referred to as the minimum number of windows fit-best fit (MFBF) is proposed to reduce the number of partition switches when partitions are scheduled to generate the major time frame. The proposed process from partition parameters deriving to inter-partition scheduling completes the design of the major time frame from several partition applications. Experimental results show that the MFBF algorithm reduces the number of partition switches effectively.

Key words: real-time systems; resource partition; integrated modular avionics; ARINC 653; major time frame

北京航空航天大学学报

Received: 2015-11-06; **Accepted**: 2016-01-08; **Published online**: 2016-03-15 14:39

URL: www.cnki.net/kcms/detail/11.2625.V.20160315.1439.007.html

Foundation item: Defense Industrial Technology Development Program (1000-GEAC0001)

* **Corresponding author**. Tel.: 029-88431545 E-mail: ducl@nwpu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0742

基于 EHW 和 RBT 的电路故障自修复策略性能分析

张峻宾，蔡金燕*，孟亚峰

（军械工程学院 电子与光学工程系，石家庄 050003）

摘 要：在诸如深空、深海等特殊环境中的电子电路，传统的基于冗余容错技术提高电子系统可靠性的方法受到了很大程度的限制，进而基于硬件演化（EHW）的故障自修复策略开始被广泛研究。但是后者也存在一些弊端，如存在电路演化规模大、电路演化速度慢、故障修复能力有限等问题。因此，在前期工作中，提出了基于 EHW 和补偿平衡技术（RBT）的电路故障自修复策略。通过从故障自修复能力、故障修复速度、硬件资源消耗等角度对比分析，相比于常规的基于 EHW 的故障自修复策略，基于 EHW 和 RBT 的电路故障自修复策略的故障修复方法灵活、故障修复类型多，且能缩减电路演化规模、缩短电路演化时间、提高电路修复速度、硬件资源消耗可控，其可行性和有效性得到了论证，具有重要的工程应用价值。

关 键 词：硬件演化（EHW）；故障自修复；补偿平衡技术（RBT）；矫正电路（RTC）；故障在线修复

中图分类号：TP302.8

文献标识码：A **文章编号**：1001-5965(2016)11-2423-13

常规的电子系统主要采用冗余容错技术提高系统的可靠性[1-2]。尤其在深空、深海等特殊复杂环境中工作的电子系统，基于冗余容错技术提高电子系统可靠性受到了很大程度的限制，使得探索新技术提高电子系统可靠性和及时现场修复故障显得格外重要[3-7]。

在 20 世纪末，硬件演化（Evolvable Hardware，EHW）技术的提出为提高电子系统的可靠性和实现故障修复提供了一种全新的途径。其具有自组织、自适应等特点，采用 EHW 技术实现故障电路自组织，从而保证电子系统正常工作。以上称为常规的基于 EHW 的电子电路故障自修复策略，但仍存在一些弊端。

1）故障修复模式较为简单，故障修复类型有限，故障现场及时修复的能力受到了一定约束。如深空、深海等环境中的电子系统出现故障时，若不能实现故障现场及时修复，将直接影响系统正常运行，甚至造成整个系统瘫痪。当被测单元（Unit Under Test，UUT）出现故障时（UUT 是可编程器件），需要将其隔离，再进行电路演化配置。在配置过程中，可能会因 UUT 进行电路配置使得整个电子系统停止运行，此时电子系统是否具有故障在线修复能力将显得格外重要。

2）电路演化耗时较长。常规的基于 EHW 的故障修复策略需要对整个出现故障的电路系统进行演化，电路演化规模伴随着演化电路的输入输出端数量的增加而增长。当演化电路规模较大时，由于搜索空间大，成功搜索到预期目标电路的时间较长，甚至有在规定的时间内根本搜索不到预期目标电路的可能性，造成故障很难及时修复。

针对以上问题，前期工作中提出了基于 EHW 和补偿平衡技术（Reparation Balance Technology，

收稿日期：2015-11-11；**录用日期**：2016-03-04；**网络出版时间**：2016-04-15 16:56
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160415.1656.010.html
基金项目：国家自然科学基金（61271153，61372039）
* **通讯作者**：Tel.：0311-87994289 E-mail：cjyrad@163.com

引用格式：张峻宾，蔡金燕，孟亚峰. 基于 EHW 和 RBT 的电路故障自修复策略性能分析[J]. 北京航空航天大学学报，2016，42(11)：2423-2435. ZHANG J B，CAI J Y，MENG Y F. Performance analysis of circuit fault self-repair strategy based on EHW and RBT[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2423-2435 (in Chinese).

北航学报 赠阅

RBT)的电子电路故障自修复策略[8]。采用了RBT技术,针对正在运行的电子系统出现的故障,能够快速演化出矫正电路,通过多路选择器(Multiplexer, MUX)的切换,实现故障现场、在线及时修复,从而克服了常规基于EHW的电子电路故障自修复策略中遇到的各种问题。

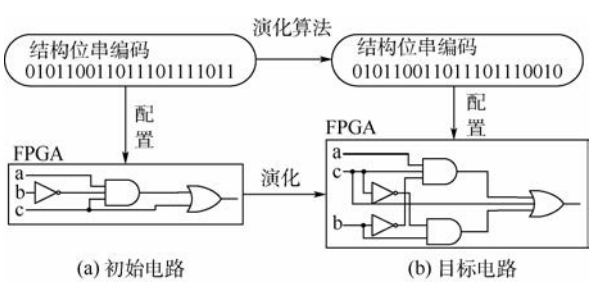
本文从故障自修复能力、故障修复时间、硬件资源消耗3个角度,对提出的基于EHW和RBT的电路故障自修复策略和常规的基于EHW的故障自修复策略的性能进行了对比分析,论证了所提出的基于EHW和RBT的电路故障自修复策略的有效性、优越性和可行性。

1 硬件演化基本原理

EHW技术是以演化算法(Evolutionary Algorithm, EA)为组合优化和全局搜索工具,通过模拟进化来获得具备预期功能的电路和系统结构[8-12]。将需要设计的电路(结构位串)、参数等作为演化算法的染色体编码,通过不断交叉、变异等演化操作,找到与电路预期输入输出特性一致的电路编码。利用EHW技术具有的良好的鲁棒性、自组织、自适应、自修复的优点,为故障自修复提供了技术保障[13-17]。

EHW公式可表示为:演化算法 + 可编程逻辑器件 = 硬件演化[8,13-14,17]。即:EAs + PLDs = EHW。

EHW演化过程如图1所示。首先演化算法会初始化一段既定长度的结构位串编码01011001101110111011,其对应的电路如图1(a)所示;在一定的约束条件和适应度评判函数的作用下,通过演化算法,结构位串编码进化成01011001101110111 0010,对应的电路是目标电路,如图1(b)所示。从电路演变过程可以明显地看出,演化算法能从混沌无规则的结构位串编码按一定的规则找到符合目标电路的结构位串编码[8]。



FPGA—现场可编程门阵列。

图1 EHW基本原理

Fig. 1 Basic theory scheme of EHW

2 新颖故障自修复策略模型建立

EHW技术是近年来新兴的技术,已被众多的研究机构和学者所重视,设计了许多基于EHW的硬件系统。

国外的Lohn[15]、Gregory[16]、Pauline[17]、Chang[18]和James[19]等的研究成果阐述了整个电路演化系统包括软件和硬件两部分,软件主要实现电路编码、演化及译码,硬件主要实现对电路演化的具体结果进行执行和故障检测。国内武汉大学的朱继祥和李元香[20-21]、南京航空航天大学的姚睿和王友仁[22-24]、张强[25]等提出了多种硬件演化、容错、重构系统,这些系统都具有相同模块,它们包括用于电路演化的PC、UUT、故障检测模块等。

综合以上分析,可将常规的基于EHW的电子电路故障自修复通用模型用图2表示。通用模型主要包括UUT、上位机、下位机、MUX和串口通信等模块。除串口通信模块外,其他模块都是自修复通用模型必要的组成部分。

MUX指多路开关,即当UUT出现故障的时候,切断UUT与整个电路系统的连接,直至电路成功演化重新配置之后,恢复UUT与整个电路系统的连接。即形成三通的通路,便于信号(包括故障信号、无故障信号、补偿平衡信号)在FPGA、UUT、输入输出电路之间的互连。

此时,UUT需要是可编程器件,可为复杂可编程逻辑器件(Complex Programmable Logic Device, CPLD)/FPGA。上位机可以是个人计算机(Personal Computer, PC),也可以是最小系统,其能够运行演化算法,并能够配置演化出的数据。下位机需要和UUT相结合,主要实现对UUT电路的故障检测、分析,并根据故障信息实现对MUX开关控制。此时的下位机可以和UUT合并

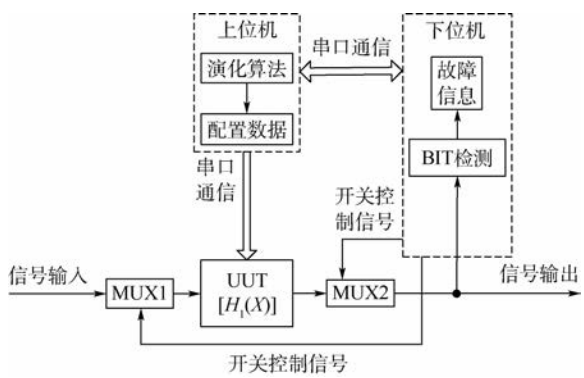


图2 常规的基于EHW的故障自修复通用模型

Fig. 2 Universal model of general fault self-repair circuit system based on EHW

北航学报 赠阅

在一起,也可以与 UUT 相分离。现有的电子系统都具备 BIT 检测的功能,BIT 检测模块可以融合在插件/电路系统(UUT)之中。

串口通信模块主要实现信号在模块之间的传递,串口通信只是信号传递的一种方式,也可以进行有线传输。

而在前期的研究中,提出了基于 EHW 和 RBT 的电子电路故障自修复策略[8],其通用模型如图 3 所示。

基于 EHW 和 RBT 的故障自修复策略,主要实现对 UUT 出现的故障信号进行矫正修复,不需要进行精确的故障定位分析,只需要对 UUT 的输入输出信号进行检测,进行故障分析,对不符合预期输入的输出组合进行矫正修复,详细的矫正修复原理见文献[8]。

从图 3 可以看出,在常规的基于 EHW 的故障自修复模型的基础上,基于EHW和RBT的故障自修复通用模型只额外增加了用于产生矫正信号的 FPGA,一些因额外增加 FPGA 而增加的布线,可以忽略。

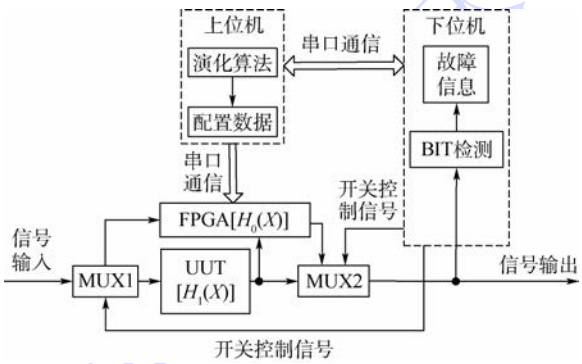


图 3　基于 EHW 和 RBT 的故障自修复通用模型

Fig. 3　Universal model of fault self-repair circuit system based on EHW and RBT

3　故障自修复能力分析

分析故障自修复能力,主要包含故障修复机制和故障修复类型 2 个指标。

3.1　故障修复机制

在故障修复机制方面,常规的基于 EHW 的故障自修复策略的故障修复示意图如图 4 所示。而基于 EHW 和 RBT 的故障自修复策略的故障修复示意图如图 5 所示。

在图 4 中,在采用 CPLD/FPGA 作为硬件实现平台时,采用常规的 EHW 技术,如果采用片外演化,需要检测出存在故障的开关节点或逻辑块,以便于修复时绕开故障节点。图 4(a1)中出现故障的节点对应的第 2 行未被使用,此时交换第 1 行 AND 线和第 2 行 AND 线,可实现故障修复(图 4(a2));图 4(b1)中编号 12 的逻辑块出现故障,将第 3 行中使用的逻辑块顺序后移,可绕开故障单元,实现故障修复(图 4(b2))。如果采用片内演化,需要演化一次,下载配置一次,检测一次。如果检测后还存在故障,还需要再进行演化、配置下载、检测,依次循环,直到所有故障消失为止。无论是片外演化还是片内演化,新配置的电路均绕开了故障节点。

在图 5 中,采用 FPGA 作为硬件的实现平台,均不需要进行故障定位,只需要检测出存在故障

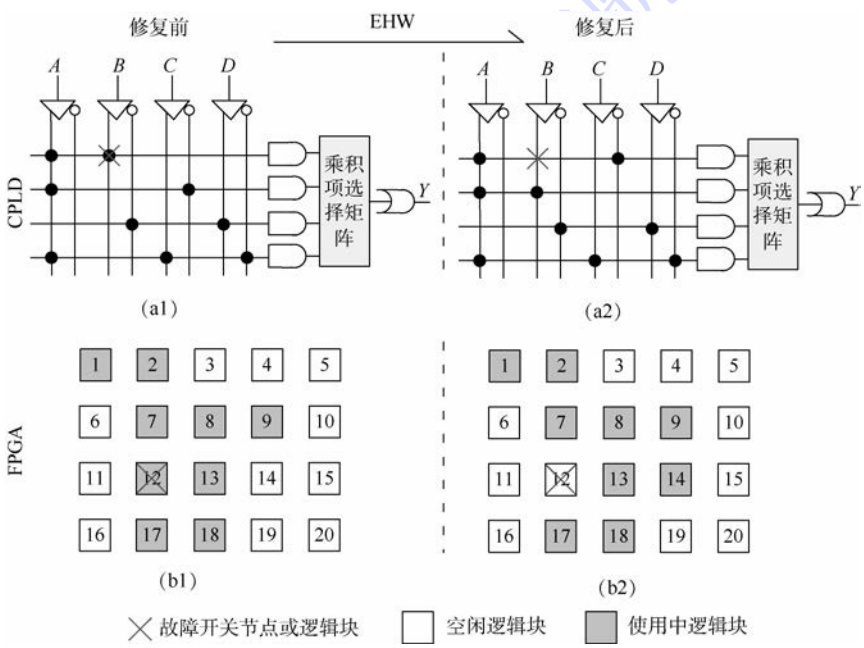


图 4　常规的基于 EHW 的故障自修复机制

Fig. 4　Mechanism of general fault self-repair circuit system based on EHW

北航学报 赠阅

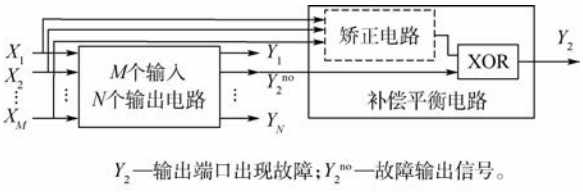


图 5 基于 EHW 和 RBT 的故障自修复机制

Fig. 5 Mechanism of fault self-repair circuit system based on EHW and RBT

的输入输出组合即可。修复是对输出状态进行矫正，详细原理可参考文献[8]。

以上可以看出，常规的基于 EHW 的故障自修复策略，在片外演化中会因故障定位的难度大于直接对输出状态判决的难度，或在片内演化中因检测不合格而循环演化，这些都会消耗较长的时间，造成故障不能及时修复的现象。而基于 EHW 和 RBT 的故障自修复策略，没有前者存在的问题，在故障信息判断方面能够节约时间，这是后者的主要优点。

3.2 故障修复类型

数字电路出现的故障类型较多，根据不同故障类型具有的特点，紧密结合常规的基于 EHW 的故障自修复策略和基于 EHW 和 RBT 的故障自修复策略的故障修复机制，不同故障类型下的故障修复能力见表 1。表 1 中 t 表示故障修复时间，其包括以下几部分：检测故障(t_1)、故障信息分析(t_2)、电路演化(t_3)、串口通信(t_4)、数据配置(t_5)、开关切换(t_6)，且 $t=t_1+t_2+t_3+t_4+t_5+t_6$；$t_s$ 表示间歇性故障时间间隔。

针对固定型故障和时滞故障，如表 1 所示，由于常规的基于 EHW 的故障自修复策略的故障修复机制是针对出现的故障，直接演化出新的电路重新配置，绕开故障单元，所以这 2 种故障能够被修复，只有在冗余单元不足以替换故障单元时，故障修复才会失败。而基于 EHW 和 RBT 的故障自修复策略能够修复这 2 种故障。

表 1 不同故障类型下的 2 种策略的故障自修复能力

Table 1 Fault self-repair ability of two strategies with different fault types

故障类型		能否修复故障	
		常规的基于 EHW 的故障自修复策略	基于 EHW 和 RBT 的故障自修复策略
固定型故障		能	能
时滞故障		能	能
瞬态故障		能	能
桥接故障		不能	能
间歇性故障	$t \geqslant t_s$	不能	能
	$t < t_s$	不一定	能

针对桥接故障，由于信号已经不能形成通路，这种故障和 UUT 内部电路没有任何关系，因此常规的基于 EHW 的故障自修复策略不能修复桥接故障。而基于 EHW 和 RBT 的故障自修复策略能够修复桥接故障，在旁路 FPGA 内部，配置需要的电路，通过 MUX 转接，对存在故障的输出状态进行矫正修复。

由于瞬态故障不是由电路或系统中硬件引起的，多是由外部原因造成，通过重启等可以复原。即使不重启，两种演化策略均可修复。

重点在于分析 2 种故障修复策略在间歇性故障时的故障修复能力。目前，MUX 多由互补金属氧化物半导体(Complementary Metal Oxide Semiconductor, CMOS)构成，MUX 切换速度很快，耗时在几纳秒至几百纳秒之间[26-28]，而纳秒级的时间间隔相对于间歇性故障时间间隔(t_s)是可以忽略不计的。此时基于 EHW 和 RBT 的故障自修复策略可根据间歇性故障时间间隔、自适应地判断 MUX 接通和断开时机，实现故障修复。

而常规的基于 EHW 的故障自修复策略，在对间歇性故障进行修复时，和 t_s 有直接关系。当 $t \geqslant t_s$ 时，在电路演化好的时候又会出现新的故障，此时故障不能修复；当 $t < t_s$ 时，此时只在 $t_s - t$ 的时间内实现了故障修复，而当出现新故障后，电路又需要重新演化，此时常规的基于 EHW 的故障自修复策略表现出了其局限性。

综上分析可得，基于 EHW 和 RBT 的故障自修复策略在对桥接故障、间歇性故障进行修复的时候具有明显的优势，而常规的基于 EHW 的故障自修复策略此时故障修复能力明显不足。

4 故障修复时间

故障修复时间是衡量故障修复速度的重要指标。在常规的基于 EHW 的故障自修复过程中，且在 3.2 节中所示，故障修复时间存在以下关系：$t=t_1+t_2+t_3+t_4+t_5+t_6$。而在基于 EHW 和 RBT 的电子电路故障自修复策略过程中，时间消耗的位置和常规的基于 EHW 的故障自修复策略一致，只是电路演化时间 t_3 消耗量不一样，具有明显的区别，这也是所提出的 EHW 和 RBT 故障自修复策略的优势所在。

串口通信 RS232 标准接口的传输速率可达到 115 200 b/s，RS485 最高则可达到 10 Mb/s，常规情况下的数据能够在几秒内传输完毕。基于 CMOS 组成的 MUX 开关切换时间可忽略不计。由于所提出的基于 EHW 和 RBT 的故障自修复策略，只需检测出哪些端口的哪些组合存在故障，不

北航学报 赠阅

需关心电路内部到底出现了什么类型的故障。因此,检测故障就显得容易,最简单的方法就是将测量值和预期值进行比对。

而信号测量与信号的时钟频率有直接关系。信号的频率越高,其周期越短,根据奈奎斯特采样定理,只要能采样一个周期的信号,就能对信号进行评判。绝大多数数字信号的周期都在 1 s 之内,也就是信号检测和比较能够在 1 s 内完成。

数据配置耗时与配置文件大小和 DCLK 频率有直接关系,通常情况下的数据配置耗时可控制在几秒内[29]。

而电路演化时间受演化算法、硬件配置、电路演化规模影响。不同硬件配置的 PC,运行速度有所不同。当在同一配置的 PC 对不同规模的电路进行演化的时候,所消耗的时间将不同。在基于处理器为 AMD Athlon X2 250 双核,内存为 2 GB,显卡为 ATI Radeon 3000 的 PC 上,采用遗传离散粒子群算法作为演化算法,使用 MATLAB 软件,对 4 输入 1 输出和 4 输入 2 输出的电路进行演化,进化次数从 0 ~ 120 000 次,对每增加 6 000 次的进化时间消耗进行了统计,统计后的数据经过最小二乘支持向量机(Least Square Support Vector Machine, LS-SVM)拟合,得到如图 6 所示的曲线。

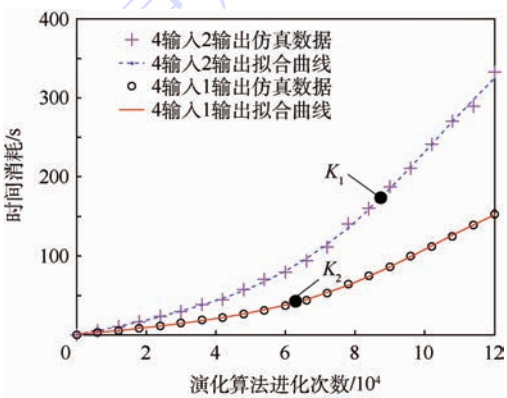


图 6　演化算法进化次数与时间消耗量之间的关系

Fig. 6　Relationship between iterative number of EA and time consumption

深究其原因:当演化电路规模越大,演化过程中电路编码长度越长,解空间就会很大,寻找到目标解的耗时将直接增加。假设编码长度为 H,其二进制编码的解空间为 2^H;当编码长度为 $H+1$ 时的解空间为 2^{H+1},新增量为 2^H。当电路规模越大时,H 越大,解空间的新增量呈 2^H 指数增长。目前国内外的很多研究者[30-32]都在致力于解决大规模电路快速演化问题。当电路规模增加时,电路成功演化的难度也相应增加,电路演化成功率会逐渐降低,且电路演化时间消耗也增加。

如图 6 所示,论编码长度,4 输入 2 输出大于 4 输入 1 输出。就相同的进化次数而言,前者所消耗的时间明显长于后者。随着进化次数的增加,时间消耗的差距会越来越大。随着电路规模增大,因搜索空间增大而成功搜索到目标电路的进化次数也将增加,对应的时间也将增加。从图 6可以看出,K_1 点的演化迭代次数大于 K_2 点,且前者的时间消耗明显大于后者。

基于此,前期提出的基于 EHW 和 RBT 的故障自修复策略相比于常规的基于 EHW 的故障自修复策略,在电路演化时间上具有明显优势。无论是组合逻辑电路,还是时序逻辑电路,前者只针对出现故障的输出端进行电路演化,演化电路的规模得到了降低,演化电路规模如图 7 所示。

针对规模为 $M \times N$ 的电路($X_1, X_2, \cdots, X_M$ 为输入端编号,$Y_1, Y_2, \cdots, Y_N$ 为输出端编号),当 1 个输出端出现故障时,演化电路的规模为 $M \times 1$;当 2 个输出端出现故障时,演化电路的规模为 $M \times 2$;当 i($i \in [1, N]$,且 i 为整数)个输出端出现故障时,演化电路的规模为 $M \times i$;只有当 N 个输出端都出现故障时,演化电路的规模才是故障 UUT 电路的规模(即演化电路规模为 $M \times N$)。

由于故障自修复的时间主要消耗在电路演化上,因此缩减电路演化规模,在尽量短时间内成功演化出目标电路,将显得格外重要。缩减演化电

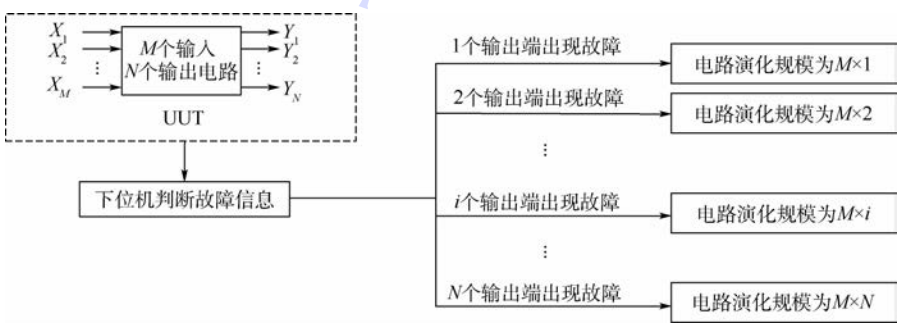


图 7　基于 EHW 和 RBT 的故障自修复策略中电路演化规模

Fig. 7　Circuit evolutionary scales of fault self-repair strategy based on EHW and RBT

北航学报 赠阅

路规模、缩短演化时间、提高电路演化成功率和保证故障自修复的速度四者具有递进关系，而基于EHW 和 RBT 的电子电路故障自修复策略在此方面具有明显的优势和重要的工程应用价值。

5 硬件资源消耗

硬件资源消耗主要从定量/微观、定性/宏观2 个层面进行分析。对比分析图 2 和图 3，排除相同的单元/模块，基于 EHW 和 RBT 的故障自修复策略新增硬件资源包括 3 个部分：MUX、与输出端口对应的 XOR 门和配置矫正电路的 FPGA。

5.1 定量/微观的分析角度

在分析硬件资源消耗时，可以将 CMOS 或基本门等变化量作为分析指标。CMOS 和基本门之间存在固定的数值关系：1 输入的非门（NOT）需要2 个 CMOS，2 输入的与门（AND）需要 6 个 CMOS，2输入的或门（OR）需要6个CMOS，而一个典型的单刀单掷开关需要 4 个 CMOS[33-34]。基本门和 FPGA 逻辑资源之间也存在一定的换算关系，采用等效门数原则，即实现相同的功能电路，在标准门阵列中需要的门数就是 FPGA 该资源的等效门数。以 Xilinx 的 XC4000 系列的 FPGA 为例，一个 4 输入的查找表（Look Up Table，LUT）等价于 1 ~ 9 个基本门，一个触发器等价于 6 ~ 12 个基本门，一个基本可编程逻辑单元（Configurable Logic Block，CLB）等价于 15 ~ 48 个基本门[35]。当使用一个 CLB、LUT 或者触发器实现不同功能的时候，对应的等价基本门数量会发生变化。由于 XOR 和 MUX 可独立于 FPGA 存在，因此可采用 CMOS 指标进行衡量。

根据 RBT 技术基本原理，每个输出端需要使用 1 个 2 输入的异或门（XOR），1 个 2 输入的 XOR 总的 CMOS 消耗量为 22。针对一个具有 M 个输入端 N 个输出端的电路，假设每个输出端都受到所有输入端的影响。由于需将 UUT 输入信号引入到 FPGA，此时新增 M 个单刀单掷开关（等价于 $4M$ 个 CMOS）；由于 UUT 的故障信号需引入 FPGA，且 UUT 的矫正信号需引出，此时 1 个输出端口新增 2 个单刀单掷开关（等价于 8 个 CMOS）。用 S_{mxr} 表示为修复故障时 XOR 和 MUX 的 CMOS 使用量，修复 i 个输出端口新增 CMOS 数量用 S^i_{mxr} 表示：

$$S^i_{mxr} = 4M + 8i + 22i = 4M + 30i \quad (1)$$

式中：$i \in [1:N]$，i 为整数。从式（1）可以看出，在故障修复率为 100% 时，当 M 和 i 值越小，对应越少；反之则越多。

以 ISCAS85 电路库中的 C17 和 C6288 电路为例，对故障修复中 S_{mxr} 进行分析。此时仍假设每个输出端都受到所有输入端的影响，一个 LUT 等价于 8 个基本门，且一个基本门等价于 6 个 CMOS，此时得到了如图 8 和图 9 所示数据。

从图 8 和图 9 可以看出，针对同一个 UUT，随着故障输出端数量的增加，采用 RBT 技术后，S_{mxr} 消耗量呈线性增长，但增长较为缓慢。由于 C17 电路规模小，且输入/输出端数量少，所以 S_{mxr} 消耗量占 C17 电路的比值偏大；而 C6288 电路规模较大，且新增输入/输出端数量较少，所以 S_{mxr} 消耗量占 C6288 电路的比值远远小于 C17。即使当 C6288 的故障修复率达到了 100%，S_{mxr} 消耗量占 C6288 电路的比值还不到 4%。

因此可以得到以下结论：当 UUT 规模越大，且其输入端/输出端数目越少时，采用 RBT 技术，修复故障时 S_{mxr} 和 UUT 本身的 CMOS 使用数量的比值将越小，甚至可忽略；反之则越大。

最后重点比较用于修复 UUT 的 RTC 的硬件资源消耗。由于现阶段 EHW 主要实现平台是FPGA，因此选取FPGA的基本组成单元（LUT、

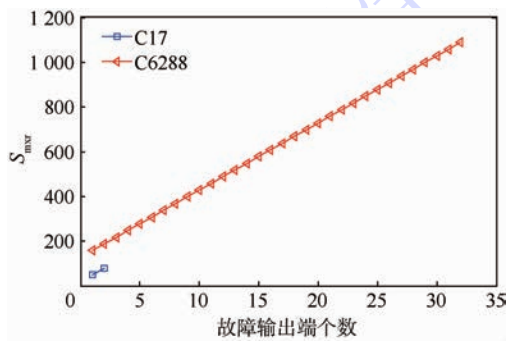


图 8 故障输出端个数与 S_{mxr} 的关系

Fig. 8 Relationship between number of fault output ports and S_{mxr}

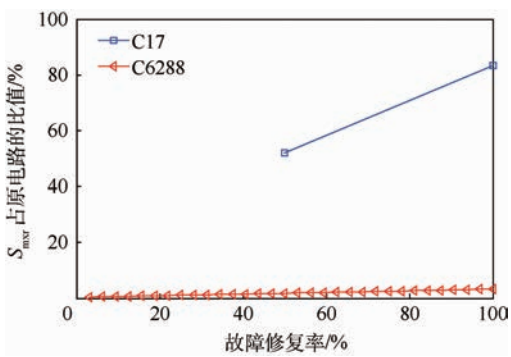


图 9 故障修复率与 S_{mxr} 增量百分比的关系

Fig. 9 Relationship between fault repair rate and increment percentage of S_{mxr}

北航学报 赠阅

Slice 和可编程输入输出单元(Input Output Block, IOB))作为衡量电路硬件资源消耗的评价标准。目前主流的 FPGA 整合了许多常用功能,大多是基于 4 输入 LUT 技术。

数字电路的单个输出端存在只使用电路系统一部分输入端的可能性,针对多个故障输出端演化出的 RTC 电路可能存在共用 LUT 的可能性,共用的 LUT 数量用 L 表示。输出端个数的增加也势必影响 LUT 数量的增加。针对同一个输出端,因电路复杂度增加可能额外增加一些 LUT,用 Q 表示。为了便于统计分析,需 3 个约束条件:

1) 针对每个故障输出端,根据补偿修复技术演化出的 RTC 电路,均受所有输入端的影响。

2) 同一电路的不同输出端不存在共用 LUT 的情形,即 $L=0$。

3) 针对同一个输出端因电路复杂度而新增的 LUT 数量为 0,即 $Q=0$。

为了定量分析硬件资源消耗,首先选取了 ISCAS85 中的 2 个典型电路进行分析比较,分别是 C17 电路和 C432 电路。根据表 2 所示的电路规模划分原则,C17 属于小规模电路,C432 属于大规模电路。

C17 电路具有 5 个输入端和 2 个输出端,每个输出端只与其中 4 个输入端有关系。不同数量输出端出现故障时采用 RBT 技术后,得到的 RTC 的硬件资源消耗见表 3。表中的 OC (Original Circuit) 为原始电路,CRTC(Corresponding RTC)为与之对应的矫正电路。

从表 3 可以看出,针对中规模电路,由于其本身的硬件资源消耗较少,在进行补偿平衡修复时,当只有 1 个输出端出现故障时,设计的 RTC 电路硬件资源消耗 Slices 增加了 100%,LUT 增加了 50%,IOB 增加了 71.429%。当 2 个输出端出现故障时,Slices、LUT 及 IOB 都增加了 100%。

而 C432 电路具有 36 个输入端,7 个输出端。输出端 N223 和 N329 受共同的 18 个输入端的影响,剩余的 5 个输出端都受全部 32 个输入端的影响。对应的 C432 电路单个输出端出现故障之后,对应的 RTC 硬件资源消耗见表 4。

表 2　集成电路的集成度分类

Table 2　Integration level classification of integrated ciruit (IC)

名称	SSI	MSI	LSI	VLSI
门数量	1 ~ 10	10 ~ 100	100 ~ 100 000	≥100 000

表 3　C17 电路不同数量输出端故障时 RTC 硬件资源消耗

Table 3　Consumption of hardware resource of RTC when C17 circuit has different numbers of fault output ports

故障输出端口数量	使用的逻辑单元名称	OC	CRTC	(CRTC/OC)/%
1	Slices	1	1	100.000
	4 输入 LUT	2	1	50.000
	IOB	7	5	71.429
2	Slices	1	1	100.000
	4 输入 LUT	2	2	100.000
	IOB	7	7	100.000

表 4　C432 电路单个输出端故障时 RTC 硬件资源消耗

Table 4　Consumption of hardware resource of RTC when C432 circuit has one fault output port

故障输出端口名称	使用的逻辑单元名称	OC	CRTC	(CRTC/OC)/%
N223 N329	Slices	56	3	5.357
	4 输入 LUT	109	5	4.587
	IOB	43	19	44.186
N370 N421 N430 N431 N432	Slices	56	5	8.929
	4 输入 LUT	109	9	8.257
	IOB	43	37	86.047

从表 4 可以看出,当输出端口 N223 和 N329 出现故障时,对应的硬件资源消耗 Slices 增加了 5.357%,LUT 增加了 4.587%,IOB 增加了 44.186%。当输出端口剩余的 5 个输出端出现故障时,对应的硬件资源消耗 Slices 增加了 8.929%,LUT 增加了 8.257%,IOB 增加了 86.047%。

相比于 C17 电路,在单个输出端出现故障时,C432 电路的 Slices、LUT 和 IOB 的增加量占 OC 硬件资源消耗量的百分比在逐渐下降。就 C432 电路而言,与输出端有关系的输入端的数量也直接影响 RTC 电路的硬件资源消耗。

为了更加直观定量地分析输入端个数、电路规模等指标对 RTC 电路硬件消耗的影响,对 ISCAS85 中的 10 个不同规模电路(C499、C880、C880a、C1355、C1908、C1908a、C2670、C3540、C5315、C6288)进行了分析。此时假设 $L=0$,单输出端出现故障对应的 Slices、LUT 和 IOB 硬件资源消耗如表 5 ~ 表 7 所示。

北航学报 赠阅

表 5 Slice 消耗对比表

Table 5 Contrastive table of Slice consumption

电路名称	C499	C880	C880a	C1355	C1908	C1908a	C2670	C3540	C5315	C6288
OC	45	59	63	45	57	43	84	187	210	387
CRTC	6	8	8	6	5	5	30	7	23	4
(CRTC/OC)/%	13.33	13.56	12.70	13.33	8.77	11.63	35.71	3.74	10.95	1.03

表 6 LUT 消耗对比表

Table 6 Contrastive table of LUT consumption

电路名称	C499	C880	C880a	C1355	C1908	C1908a	C2670	C3540	C5315	C6288
OC	78	112	108	78	101	75	160	331	399	690
CRTC	11	15	15	11	9	9	59	13	45	8
(CRTC/OC)/%	14.10	13.39	13.89	14.10	8.91	12.00	36.88	3.93	11.28	1.16

表 7 IOB 消耗对比表

Table 7 Contrastive table of IOB consumption

电路名称	C499	C880	C880a	C1355	C1908	C1908a	C2670	C3540	C5315	C6288
OC	73	86	86	73	58	58	373	72	301	64
CRTC	42	61	61	42	34	34	234	51	179	33
(CRTC/OC)/%	57.53	70.93	70.93	57.53	58.62	58.62	62.73	70.83	59.47	51.56

由于不易清晰从表 5 ~ 表 7 中看出变化规律,得到了如图 10 所示的曲线图。在图 10 中,横坐标按照 10 个典型电路硬件资源消耗从少到多的顺序排序,与之对应括号中的数值表示所消耗的 Slice 数量;左边的纵坐标表示的是矫正电路与之对应的原始电路的 IOB 和 LUT 消耗的比值;右边纵坐标表示的 10 个电路对应的输入端的数量。由于 1 个 Slice 包含 2 个 LUT,所以图 10 中 Slice 和 LUT 的曲线走势一致,因此只需要分析其中一个数据即可。

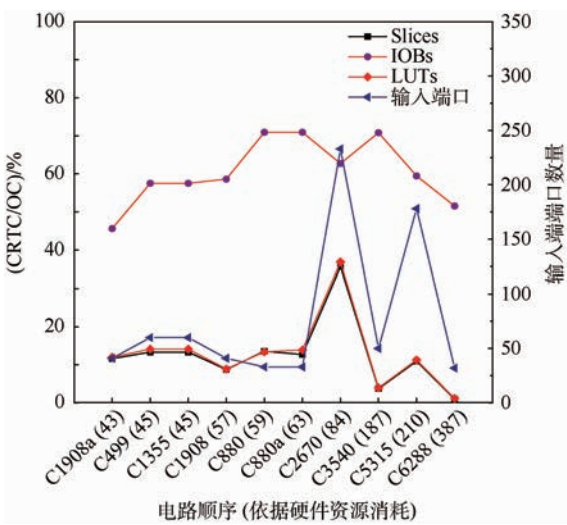


图 10 不同规模电路单输出端故障时 RTC 硬件资源消耗对比曲线

Fig. 10 Contrastive curves of consumption of hardware resource of corresponding RTC when single fault output port of different scale circuits occurs

在电路单个输出端出现故障时,从图 10 中能够清晰地看出,10 个典型电路的输入端数量的变化趋势和 Slice 的 ORTC/OC 的值走势一致。即当电路的输入端数量增加(减少)的时候,对应的 Slice 和 LUT 的 ORTC/OC 的值都随着增加(减少)。就具体的电路而言,当 10 个电路均无故障时,C1908a 电路消耗的 Slice 数量最少,为 43 个;C6288 消耗的 Slice 数量最多,为 387 个。就 Slice 的 ORTC/OC 而言,C6288 电路的矫正电路 Slice 增加百分比最少,为 1.03%;C2670 电路的矫正电路 Slice 增加百分比最大,为 35.71%。

相反地,C499 电路本身需要消耗 45 个 Slice,其对应的 ORTC/OC 的值达到了 13.33%,对应的输入端为 41 个。如果是表 3 中的 C17 电路,只消耗 1 个 Slice,而对应的 ORTC/OC 的值达到了 100%,对应的输入端个数为 5 个。

从表 5 和表 6 还可以看出,C880 和 C880a 电路、C1908 和 C1908a 电路,这 2 组电路的输入端和输出端数量相等,但由于其内部电路结构复杂度不一样,即 L 和 Q 不一样, Slice 和 LUT 的 ORTC/OC 的值均不相等,当 $Q-L$ 越大时,ORTC/OC 值越小,反之当 $Q-L$ 越小时,ORTC/OC 值越大。

从以上分析可以得到以下结论:无故障 UUT 电路的总硬件资源消耗越大,且输入端个数越少时,则 UUT 故障后对应的矫正电路所消耗的硬件资源(Slice、LUT、IOB)越少;反之,UUT 总的硬件资源消耗越小,或输入端个数越多,则故障 UUT

对应的矫正电路所消耗的硬件资源(Slice、LUT、IOB)越多。

以上是在单输出端存在故障的情况下进行的定量分析,根据前面的约束条件,现对多个输出端出现故障进行分析。对单输出端故障对应的表 5、表 6 和表 7 中的数据进行反推,由于同一个电路不同输出端可使用 1 个 Slice 中的不同 LUT,因此对 LUT 的使用量进行分析,得到了表 8 所示的数据。

选取表 8 中的几个典型电路进行对比分析。当 $L=0$ 且 $Q=0$ 时,在硬件资源消耗不大于无故障电路的情况下,并采用补偿平衡技术,C2670 无故障时的 LUT 消耗量为 160,且存在 140 个输出端口,其故障修复率只有 1.43%,主要是因为其输出端口和输入端口太多,而其无故障电路的 $Q-L$ 远远小于 0,即 $Q-L=-8\,100$;C6288 无故障时的 LUT 消耗量为 690,且存在 32 个输出端口,其故障修复率能够达到 65.63%,其无故障电路的 $Q-L$ 大于 0;C3540 无故障时的 LUT 消耗量为 331,其故障修复率能够达到 100.00%,主要因为其无故障电路的 $Q-L=45$,远远大于 0。

因此可以得到以下结论:当 $L=0$ 且 $Q=0$ 时,在硬件资源消耗不大于无故障电路的情况下,输出端数量越少,且无故障电路的硬件资源消耗越多,采用 RBT 技术后,对应的故障修复率将越高。

表 8　硬件资源消耗不大于无故障电路情况下 RBT 的故障修复率

Table 8　Fault repair rate based on RBT when consumption of LUT is less than that of original circuit

电路名称	C499	C880	C880a	C1355	C1908	C1908a	C2670	C3540	C5315	C6288
原电路 LUT 数量	78	112	108	78	101	75	160	331	399	690
对应矫正电路 LUT 数量	11	15	15	11	9	9	59	13	45	8
总端口数量	73	86	86	73	58	58	373	72	301	64
输出端口数量	32	26	26	32	25	25	140	22	123	32
最大能够修复故障端口数量	7	7	7	7	11	8	2	22	8	21
故障修复率/%	21.88	26.92	26.92	21.88	44.00	32.00	1.43	100.00	6.50	65.63

注:当 C3540 故障修复率达到 100% 时,只使用了 286 个 LUT,且 CRTC/OC 的比值是 86.40%($L=0$ 且 $Q=0$)。

以上是对 10 个典型电路的对比分析。针对一个具有 M 个输入端 N 个输出端的电路,UUT 电路本身 LUT 的消耗量为 S_{OC},在 RBT 技术下修复单个输出端的 LUT 消耗量为 S_{RO}。则得到

$$S_{RO}=M/4 \tag{2}$$

当修复多个输出端出现故障时,存在因复杂度增加的 LUT,因共用 LUT 而减少使用 LUT。存在修正参数 R_{lq}:

$$R_{lq}=Q-L \tag{3}$$

在硬件资源消耗不大于无故障电路的情况下,采用 RBT 技术后的故障修复率用 P_{re} 表示:

$$P_{re}=\frac{(S_{OC}-R_{lq})/S_{RO}}{N}\times 100\%=\frac{\frac{S_{OC}-Q+L}{M/4}}{N}\times 100\%=\frac{4(S_{OC}-Q+L)}{M\times N}\times 100\% \tag{4}$$

从式(4)中同样可以看出,在硬件资源消耗不大于无故障电路的情况下,$M\times N$ 越小,即对应的输入端和输出端数量乘积越小,修正参数 R_{lq} 越小,S_{OC} 越大,在采用 RBT 技术下的故障修复率越高,即验证了表 8 中各个电路的故障修复率。当 $M\times N$、R_{lq} 更小,修复故障的矫正电路的硬件消耗相对于无故障电路的硬件消耗的比值将更小,尤其是在少量输出端口出现故障的情况下,新增的硬件资源消耗基本可忽略不计。

另外,通过演化算法实现电路进化,修复故障所使用的硬件资源还能够进一步降低,其原因如下:

1)以上分析的过程中,均假设 $Q=L=0$,而在实际电路演化过程中,通过演化算法,能够消除一些冗余单元,使得 $L>0$。

2)以上分析过程均假设每个输出端均受所有输入端的影响,这是一个最大值。而通过演化算法,能够消除与输出端无关的输入端,MUX 的使用量和 LUT 的使用量能够进一步缩减。

从以上分析综合来看,UUT 硬件资源消耗越大,输入/输出端个数越少,且出现故障的输出端个数较少,修复故障所使用的硬件资源(RTC)占 UUT 本身硬件消耗的比值将越小,甚至可能显得微不足道。此时基于 EHW 和 RBT 的故障自修复策略优势将得到体现。设计者还能够采用式(1)和式(4),对故障修复的硬件资源消耗和故障修复率进行预先评估。

5.2　定性/宏观的分析角度

在定性/宏观的分析角度,主要基于芯片的重量、体积、面积等成本进行评估。在详细分析常规的基于 EHW 的故障自修复电路系统和 EHW 与

北航学报 赠阅

RBT 相结合的故障自修复电路系统之间的区别后,排除两者必备的模块,实质上后者只增加了一片 FPGA 和相应的布线。布线可以忽略,下面主要评估新增 FPGA 对整个硬件系统的影响。

目前,CPLD/FPGA 制造工艺已经达到了 16 nm,系统门数已经达到近千万门。以 Xilinx 的产品为例,早期的中低端产品 FPGA Spartan-3E 系列的 XC3S500E 芯片[36],包含了 1 164 个 CLB、73 K 的 Distributed RAM、360 K 的 Block RAM、用户能使用的最大 I/O 数量为 232。其封装面积为 0.64 ~ 9 cm^2,厚度为 1.1 ~ 4.1 mm,重量为 0.1 ~ 5.3 g,价格几十到几百元人民币不等[36]。Xilinx 的最新高端产品 FPGA XCVU13P 芯片,具有 1 636 个CLB、432 Mb 的 UltraRAM、94.5 Mb 的 Total Block RAM、51.6 Mb 的 Max. Distributed RAM,I/O 总数达到了 832,其封装面积约为 5 × 5 cm^2。功耗低,成本更低,整体性能大幅提升,可利用资源大幅扩大。而重量、面积、体积保持不变,价格在 10 000 美元左右[37]。CPLD/FPGA 的工作环境温度范围为 -40 ~ 100℃,能够满足工作要求。其价格随着芯片资源、功能的丰富呈增加之势[38]。

针对基于 EHW 和 RBT 的电子电路故障自修复系统,一方面,系统中的一片 CPLD/FPGA 可以配置多个 RTC 电路,通过 MUX 转接,实现对多个 UUT 进行修复,修复原理示意图如图 11 所示。

图 11 为单个 FPGA 实现对多个 UUT 进行故障修复电路系统的简化模型。从模型中可以清晰地看出,每个 UUT 模块均能通过各自两侧的 MUX,实现和 FPGA 间信号互通。这样可以减少硬件系统中 CPLD/FPGA 的使用量,降低硬件成本。

另一方面,根据设计的自修复电路系统的功能、用途不同(主要指使用环境),可以有选择地增加 CPLD/FPGA 的数量,控制硬件成本。针对一个较大的电路系统,采用 RBT 技术,图 3 中的 UUT 不局限于一个芯片,可为一个局部电路或者一个小/中/大规模电路系统。此时,相对与整个电路系统,所增加的硬件资源将显得微不足道。

设计在不同环境中工作的电子系统的主导因素不一样,有时是价格,有时是重量、面积、体积,使用者可根据自己需求进行权衡。

MUX 能实现信号转接,因此也需要关注 MUX 的硬件资源开销。90% 以上的集成电路生产技术都是属于 CMOS 生产工艺,MUX 也不例外。

现今,集成电路实现的超高速数字电路速率已超过 10 Gb/s,射频电路的最高工作频率已超过 6 GHz。工艺从 1999 年的 0.18 μm,到 2009 年已经达到了 0.07 μm。CMOS 集成电路的温度稳定性好,抗辐射能力强,因此适合于在特殊环境下工作。基于 CMOS 的存储单元 ROM 的面积只有几平方微米,基于 CMOS 的电容大小为 1 ~ 4 fF/μm^2[26]。CMOS 模拟电路的工作频率也越来越高,现在它已经能处理 5 GHz 以上的信号[27]。例如,ADG715、ADG5207、ADG918 等多路开关芯片[28],ADG715 的封装 RU-24 TSSOP 的表面积只有 7.8 × 4.4 mm^2,厚度才 1.2 mm[28]。ADG5207 芯片的温度范围为 -40 ~ 125℃,面积约 42 mm^2[27]。ADG5209 芯片的面积只有 30 mm^2[28],ADG918 功耗低,小于 1 μA,工作频率高,达到了 1 GHz,封装面积不到 10 mm^2[28]。现在有很多基于 BiMOS 的集成电路,能够满足射频和模拟电路的需求。而且 BiMOS 门电路的功耗极低,约为 0.0003 ~ 7.5 mW,信号传输时间极小,约为 3 ns[39]。另外,一片 FPGA 里面有大量的开关盒,具有很小的体积和面积,这些开关盒完全可以用来替代 MUX。

综上可以看出,在基于 EHW 和 RBT 的故障自修复电路系统中,所增加的硬件资消耗,无论是面积、重量、体积,亦或是数量、成本等都能够容忍,即使在对体积、面积和重量极为苛刻的宇航、深海等领域的电子系统,所增加的硬件消耗都在可控范围内,都不存在技术层面的障碍。

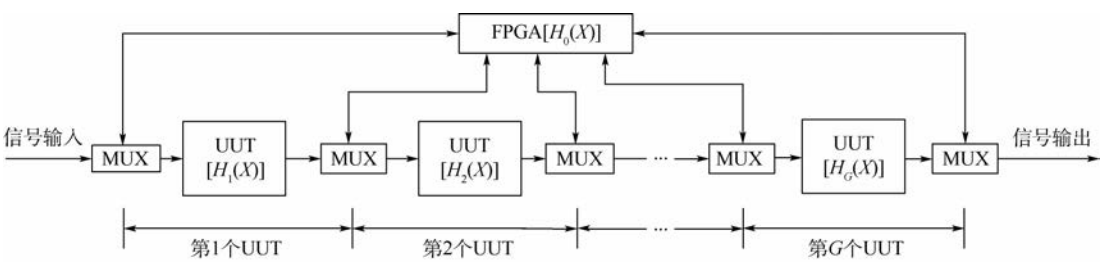


图 11 单个 FPGA 对多个 UUT 实现故障自修复的电路系统简化模型

Fig. 11 Simplified circuit system model for fault self-repair multiple UUTs by single FPGA

北航学报 赠阅

6 结　论

为解决常规的基于 EHW 的故障自修复策略存在故障定位难度大、演化电路规模大、故障在线修复不及时和电路演化速度慢等问题,首先建立了常规的基于 EHW 的故障自修复策略模型,提出了基于 EHW 和 RBT 的故障自修复策略模型,并对模型中各模块的功能进行了简要分析。本文主要对后者的性能进行了详细分析。

1) 对 2 种故障修复策略的故障自修复能力进行了详细的分析,主要包含故障修复机制和故障修复时间两个指标。相比于常规的基于 EHW 的故障自修复策略,基于 EHW 和 RBT 的故障自修复策略不需关注故障电路内部的故障信息,只需关心电路外部哪个输出口出现了故障,根据故障信息和补偿平衡修复原理,实现对故障信号矫正修复。在故障修复类型方面,尤其是在对桥接故障和间歇性故障进行修复时,前者的故障修复能力受到了一定的限制,而基于 EHW 和 RBT 的故障自修复策略的优势比较明显。

2) 对比分析了 2 种故障修复策略的故障修复时间。通过分析得到故障修复时间的长短主要取决于电路演化时长的结论。而相比于常规的基于 EHW 的自故障修复策略,基于 EHW 和 RBT 的故障自修复策略能够直接缩减电路演化规模,从而缩短了电路演化时间,保证电路演化的时效性,具有明显的优势。

3) 在硬件资源消耗方面,从定量/定性、微观/宏观等角度分析了新增硬件的成本、数量、体积、面积、重量等指标。在定量/微观方面,设计者能够依据式(1)和式(4),对故障修复的硬件资源消耗和故障修复率进行预先评估。当无故障 UUT 的硬件消耗越大,且输入端/输出端个数越少时,S_{mxr}增量将越小,CRTC/OC 的值也越小。在定性/宏观方面,新增硬件资源都是能够容忍可控。

综合以上多个角度的分析,基于 EHW 和 RBT 的电路故障自修复策略的有效性、可行性得到充分的论证。实现了故障在线修复,缩减了电路演化规模,缩短了电路演化时间,保证了故障修复速度,缩短故障修复时间,且新增硬件资源消耗具有可控性。

参考文献 (References)

[1] JESSICA D, ROGER D, ROGER M G, et al. Multi-agent control of high redundancy actuation[J]. International Journal of Automation and Computing, 2014, 11(1): 1-9.

[2] NORDSTROM S G, SHETTY S S, NEEMA S K, et al. Modeling reflex-healing autonomy for large-scale embedded systems[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2006, 36(3): 292-303.

[3] EBRAHIMI M, MIREMADI S G, ASADI H, et al. Low-cost scan-chain-based technique to recover multiple errors in TMR systems[J]. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2013, 21(8): 1454-1468.

[4] ALMUKHAIZIM S, SINANOGLU O. Novel hazard-free majority voter for *n*-modular redundancy-based fault tolerance in asynchronous circuits[J]. IET Computers & Digital Techniques, 2011, 5(4): 306-315.

[5] 姜昱光,韩建伟,朱翔,等. SRAM 型 FPGA 单粒子翻转效应加固方法[J]. 北京航空航天大学学报, 2014, 40(8): 1073-1077.
JIANG Y G, HAN J W, ZHU X, et al. Single event upset mitigation testing of SRAM-based FPGAs [J]. Journal of Beijing University of Aeronautics and Astronautics, 2014, 40(8): 1073-1077(in Chinese).

[6] 赵磊,王祖林,周丽娜,等. 星载 SRAM 型 FPGA 可靠性快速评估技术[J]. 北京航空航天大学学报, 2013, 39(7): 863-868.
ZHAO L, WANG Z L, ZHOU L N, et al. Fast reliability evaluation for SRAM-based space borne FPGAs[J]. Journal of Beijing University of Aeronautics and Astronautics, 2013, 39(7): 863-868(in Chinese).

[7] PENG J J, LIU Y P, CHEN Y Y. A dependability model for TMR system[J]. International Journal of Automation & Computing, 2012, 9(3): 315-324.

[8] ZHANG J B, CAI J Y, MENG Y F, et al. The fault self-repair strategy based on evolvable hardware and reparation balance technology[J]. Chinese Journal of Aeronautics, 2014, 27(5): 1211-1222.

[9] YAO X, TETSUYA H. Promises and challenges of evolvable hardware[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 1999, 29(1): 87-97.

[10] VASSILEV V K, MILLER J F. Scalability problems of digital circuit evolution: Evolvability and efficient designs[C] // The 2nd NASA/DoD Workshop on Evolvable Hardware. Piscataway, NJ: IEEE Press, 2000: 55-64.

[11] LIANG H J, LUO W J, WANG X F. A three-step decomposition method for the evolutionary design of sequential logic circuits [J]. Genetic Programming and Evolvable Machines, 2009, 10(3): 231-262.

[12] 何国良,李元香,史忠植. 基于精英池演化算法的数字电路在片演化方法[J]. 计算机学报, 2010, 33(2): 365-372.
HE G L, LI Y X, SHI Z Z. Elitist pool evolutionary algorithm for on-line evolution of digital circuits[J]. Chinese Journal of Computers, 2010, 33(2): 365-372(in Chinese).

[13] VASICEK Z, SEKANINA L. Formal verification of candidate solutions for post-synthesis evolutionary optimization in evolvable hardware[J]. Genetic Programming and Evolvable Machines, 2011, 12(3): 305-327.

[14] WANG J, CHEN Q S, LEE C. Design and implementation of a

北航学报 赠阅

virtual reconfigurable architecture for different applications of intrinsic evolvable hardware [J]. IET Computers & Digital Techniques, 2008, 22(5): 386-400.

[15] LOHN J D, HORNBY G S, LINDEN D S. Evolutionary antenna design for a NASA spacecraft [M]// Genetic Programming Theory and Practice II. Berlin: Springer, 2004: 301-315.

[16] GREGORY S H, JASON D L, DEREK S L. Computer-automated evolution of an X-band antenna for NASA's space technology 5 mission [J]. Evolutionary Computation, 2011, 19(1): 1-23.

[17] PAULINE C H, ANDY M T. Challenges of evolvable hardware: Past, present and the path to a promising future [J]. Genetic Programming and Evolvable Machines, 2011, 12(3): 183-215.

[18] CHANG P C, WANG Y W, TSAI C Y. Evolving neural network for printed circuit board sales forecasting [J]. Expert Systems with Applications, 2005, 29(1): 83-92.

[19] JAMES M H. Fault-tolerant sensor systems using evolvable hardware [J]. IEEE Transactions on Instrumentation and Measurement, 2006, 55(3): 846-853.

[20] 朱继祥, 李元香, 邢建国. 可重构系统的演化修复机制 [J]. 计算机学报, 2014, 37(7): 1599-1606.
ZHU J X, LI Y X, XING J G. The evolvable recovery of reconfigurable system [J]. Chinese Journal of Computers, 2014, 37(7): 1599-1606 (in Chinese).

[21] 朱继祥, 李元香, 夏学文. 演化硬件的容错模式研究 [J]. 小型微型计算机系统, 2010, 31(12): 2472-2475.
ZHU J X, LI Y X, XIA X W. Research on fault tolerance of evolvable hardware [J]. Journal of Chinese Computer Systems, 2010, 31(12): 2472-2475 (in Chinese).

[22] 姚睿, 王友仁, 于盛林, 等. 具有在线修复能力的强容错三模冗余系统设计及实验研究 [J]. 电子学报, 2010, 38(1): 177-183.
YAO R, WANG Y R, YU S L, et al. Design and experiments of enhanced fault-tolerant triple-module redundancy systems capable of online self-repairing [J]. Acta Electronica Sinica, 2010, 38(1): 177-183 (in Chinese).

[23] YAO R, CHEN Q Q, LI Z W, et al. Multi-objective evolutionary design of selective triple modular redundancy systems against SEUs [J]. Chinese Journal of Aeronautics, 2015, 28(3): 804-813.

[24] 姚睿, 于盛林, 王友仁, 等. 采用主流 FPGA 的数字电路在线生长进化方法 [J]. 南京航空航天大学学报, 2007, 39(5): 582-587.
YAO R, YU S L, WANG Y R, et al. Online growing evolution and evaluation approach based on mainstream FPGA [J]. Journal of Nanjing University of Aeronautics and Astronautics, 2007, 39(5): 582-587 (in Chinese).

[25] 张强, 周军, 于晓洲. 一种可局部动态实时重构的演化硬件平台 [J]. 西北工业大学学报, 2011, 29(5): 761-765.
ZHANG Q, ZHOU J, YU X Z. An efficient high-speed real-time partially reconfigurable platform for evolvable hardware [J]. Journal of Northwestern Polytechnic University, 2011, 29(5): 761-765 (in Chinese).

[26] 李本俊, 刘丽华, 辛德禄. CMOS 集成电路原理与设计 [M]. 北京: 北京邮电大学出版社, 1997: 88-94.
LI B J, LIU L H, XIN D L. CMOS integrated circuit theory and design [M]. Beijing: Beijing University of Posts and Telecommunications Press, 1997: 88-94 (in Chinese).

[27] 赵巍. 数字电子技术 [M]. 北京: 北京邮电大学出版社, 2010: 81-93.
ZHAO W. Digital electronic technology [M]. Beijing: Beijing University of Posts and Telecommunications Press, 2010: 81-93 (in Chinese).

[28] REV C. High voltage, latch-up proof, 4-/8-channel multiplexers data sheet adg5208/adg5209 [R]. Norword MA: Analog Devices Inc., 2011: 1.

[29] 李春雨, 张丽霞. 利用 CPLD 提高 FPGA 加载速度 [J]. 电子器件, 2013, 36(4): 550-553.
LI C Y, ZHANG L X. Using CPLD improve FPGA loading speed [J]. Chinese Journal of Electron Devices, 2013, 36(4): 550-553 (in Chinese).

[30] LI X D, YAO X. Cooperatively coevolving particle swarms for large scale optimization [J]. IEEE Transactions on Evolutionary Computation, 2012, 16(2): 210-224.

[31] ASAFUDDOULA M, RAY T, SARKER R. A decomposition-based evolutionary algorithm for many objective optimization [J]. IEEE Transactions on Evolutionary Computation, 2015, 19(3): 445-160.

[32] 姚睿, 陈芹芹, 孙艳梅, 等. 采用输入输出分解的分区分段演化机制 [J]. 哈尔滨工程大学学报, 2015, 36(4): 522-527.
YAO R, CHEN Q Q, SUN Y M, et al. In-partition in-stage evolutionary mechanism by applying input and output decomposition [J]. Journal of Harbin Engineering University, 2015, 36(4): 522-527 (in Chinese).

[33] 宋吉江, 牛轶霞, 于春战, 等. CMOS 模拟开关及其选择问题 [J]. 微电子技术, 2011, 29(3): 58-60.
SONG J J, NIU Y X, YU C Z, et al. CMOS analog switch and its choices [J]. Microelectronic Technology, 2011, 29(3): 58-60 (in Chinese).

[34] 王迎栋, 孙明杰. 基于 CMOS 模拟开关实现平衡混频器 [J]. 计算机与网络, 2012(15): 70-72.
WANG Y D, SUN M J. Design of balanced mixer based on CMOS analog switch [J]. Computer & Network, 2012(15): 70-72 (in Chinese).

[35] Xilinx Inc.. Application note: Gate account capacity metrics for FPGAs XAPP 059 (Version 1.1) [R]. San Jose: Xilinx Inc., 1997: 1-2.

[36] Xilinx Inc.. Product specification: Spartan-3 FPGA family data sheet (v4.1) [R]. San Jose: Xilinx Inc., 2013: 1-2.

[37] Xilinx Inc.. Ultrascale + FPGAs product tables and product selection guide [R]. San Jose: Xilinx Inc., 2013: 2-3.

[38] Xilinx Inc.. Spartan-3 generation FPGAs [R]. San Jose: Xilinx Inc., 2008: 8.

[39] 杨之廉, 许军. 集成电路导论 [M]. 3 版. 北京: 清华大学出版社, 2012: 123-127, 141.
YANG Z L, XU J. Integrate circuit introduction [M]. 3rd ed. Beijing: Tsinghua University Press, 2012: 123-127, 141 (in Chinese).

北航学报 赠阅

作者简介:

张峻宾　男,博士研究生。主要研究方向:电子电路故障自修复与硬件演化技术。

E-mail: kmemee@163.com

蔡金燕　女,博士,教授,博士生导师。主要研究方向:检测理论、故障诊断与自修复,硬件演化技术及可靠性共性技术。

Tel.: 0311-87994289

E-mail: cjyrad@163.com

孟亚峰　男,博士,副教授,硕士生导师。主要研究方向:电子系统性能检测与故障诊断。

E-mail: myfrad@163.com

Performance analysis of circuit fault self-repair strategy based on EHW and RBT

ZHANG Junbin, CAI Jinyan*, MENG Yafeng

(Department of Electronic and Optical Engineering, Ordnance Engineering College, Shijiazhuang 050003, China)

Abstract: The method of improving the reliability of the electronic systems based on redundant fault-tolerant technology is limited in aerospace, deep sea and some specific fields. Then the fault self-repair strategy based on evolvable hardware (EHW) begins to be researched. However, there are some kinds of disadvantages, such as large evolvable circuit scales, slow evolvable speed, finite ability of fault self-repair and so on, for which a novel fault self-repair strategy based on EHW and reparation balance technology (RBT) had been proposed in previous research. The contrast analysis is carried out in three aspects, which are the ability of fault self-repair, the speed of fault self-repair and the consumption of hardware resource. Compared to the conventional fault self-repair strategy based on EHW, the fault self-repair strategy based on EHW and RBT, which has flexible methods of troubleshooting, repairs many fault types, reduces the scales of evolvable circuit, decreases the evolvable time, and increases the speed of evolvable circuit. The consumption of hardware resource can be controlled. Its feasibility and validity have been demonstrated. The proposed fault self-repair strategy has important engineering application significance.

Key words: evolvable hardware (EHW); fault self-repair; reparation balance technology (RBT); rectification circuit (RTC); online fault repair

Received: 2015-11-11; **Accepted**: 2016-03-04; **Published online**: 2016-04-15 16:56

URL: www.cnki.net/kcms/detail/11.2625.V.20160415.1656.010.html

Foundation items: National Natural Science Foundation of China (61271153, 61372039)

* **Corresponding author**. Tel.: 0311-87994289　E-mail: cjyrad@163.com

北京航空航天大学学报

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0747

基于重叠因子优化的超流体陀螺量程扩展设计

赵玉龙[1]，任元[2,*]，沈怀荣[2]

（1. 装备学院 研究生管理大队，北京 101416； 2. 装备学院 航天装备系，北京 101416）

摘 要：针对热驱动型双弱连接超流体陀螺的幅值-角速度单值求解区间因工作曲线呈周期性而制约超流体陀螺量程的突出问题，提出了一种基于重叠因子优化的超流体陀螺量程扩展方法。首先在明晰陀螺量程受限机理的基础上，设计了幅值锁定量程扩展方案。其次分析了热相位与温升延时的关系，引入重叠因子的概念，提出了一种基于重叠因子的遗传优化方法来实现幅值锁定参数优化：通过间接编码三角形隶属函数左半宽度，在避免出现相邻隶属函数无重叠或重叠太多的前提下，对模糊线性隶属函数进行遗传优化，寻找幅值锁定系统最优参数。通过与传统比例(P)控制下幅值锁定的仿真对比表明，该方法能够使超流体陀螺的薄膜幅值快速且稳定地锁定在期望值，在保证测量精度的同时提高了陀螺系统的量程范围。

关 键 词：超流体陀螺；量程扩展；幅值锁定；热相位；重叠因子
中图分类号：V448.2
文献标识码：A **文章编号**：1001-5965(2016)11-2436-09

惯性导航技术是空间站、卫星、导弹、舰船、机器人等高性能装备的关键技术之一[1]。惯导系统的核心部件是实现检测功能的陀螺仪，自19世纪机械陀螺问世至今，其先后经历了传统的机电式陀螺（如液浮陀螺、气浮陀螺和磁浮陀螺）、光学陀螺到量子陀螺的发展阶段[2]。随着高分辨率对地观测系统、载人航天和探月工程等重大专项任务的推进，对未来惯导系统在体积、精度、灵敏度等方面提出了越来越高的指标要求。为了突破传统的思维模式，实现惯导系统的跨越式发展，许多学者开始致力于寻求新的物理效应，研究基于新的测量原理或测量方式的陀螺仪。

近年来随着低温物理学科的发展，超流体陀螺开始进入人们的视线[3-4]，由美国加州大学 Packard 提出的基于交流约瑟夫森效应的双弱连接超流体陀螺[5]，在精度、体积、温度要求和受噪声影响等方面具有巨大的优势。理论上基于物质波干涉的超流体干涉仪的角速度检测灵敏度比基于光子干涉的光纤陀螺灵敏度高将近10个数量级[6]。

目前超流体陀螺的研究尚停留在机理推导和初步实验阶段[7]，特别是在超流体陀螺高精度大量程设计和控制系统设计方面，还存在很多问题需要深入研究。

文献[8]对双弱连接超流体陀螺干涉仪进行了原理推导和相关实验，从薄膜幅值中提取出地球自转角速度分量，从而提出将其发展为新型陀螺的构想；文献[9-10]的研究表明，超流体陀螺可高灵敏、高精度感知角速度，其可检测角速度范围有限，但未提出可行的量程扩展方案；文献[11-12]提出热像移辅助方法来扩大超流体陀螺的量程，但未考虑系统惯性延迟的问题，从而不可避免地制约超流体陀螺的敏感精度和动态特性。文献[13]虽然考虑了一阶惯性滞后，为热像移辅助回路设计了简单可行的比例(P)控制器，但系统的

收稿日期：2015-11-12；**录用日期**：2015-11-28；**网络出版时间**：2015-12-09 14:33
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20151209.1433.001.html
基金项目：国家自然科学基金（51475472）；国家“863”计划（2015AA8018038C）
* **通讯作者**：Tel.：010-66364384 E-mail：renyuan823@aliyun.com

引用格式：赵玉龙，任元，沈怀荣. 基于重叠因子优化的超流体陀螺量程扩展设计[J]. 北京航空航天大学学报，2016，42(11)：2436-2444. ZHAO Y L，REN Y，SHEN H R. Range extension design on superfluid gyroscope based on overlap factor optimization [J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2436-2444 (in Chinese).

北航学报 赠阅

动态性能和控制精度有待进一步提高。此外，关于模糊控制系统优化设计方面，文献[14]提出采用粒子群算法来优化模糊控制规则，算法简单且计算方便，但容易陷入局部最优，导致结果不理想。文献[15]提出使用 BP 神经网络优化模糊隶属函数，该优化算法具有自适应自学习的能力，但其优化结果对训练样本依赖大且学习过程较复杂，工程应用较难实现。

为了克服以上问题，本文在分析超流体陀螺幅值-角速度工作曲线特点的基础上，从工程应用角度出发，旨在实现陀螺量程范围的扩展。考虑了热相位与温升延时关系的基础上，设计并优化了基于热相位补偿实现陀螺量程范围的扩展方案。在幅值锁定模糊控制系统中，引入重叠因子的概念，实现隶属函数左半宽度的间接编码，在提高遗传优化效率的同时确保了相邻隶属不会没有重叠或者重叠部分太多的情况，实现了超流体陀螺薄膜幅值快速且稳定地锁定在期望值，在保证测量精度的同时提高了陀螺系统的量程范围。

1　超流体陀螺的敏感原理

超流体陀螺如图 1 所示，由加热电阻、干涉腔、弱连接和薄膜振动幅值输出模块组成。超流体陀螺工作时，通过加热铜镍合金电阻丝使弱连接两侧形成一定的化学势差，干涉腔内的液氦Ⅱ在该化学势差的作用下，在布满由孔径 70 nm，间隔 3 μm 的若干微孔组成的弱连接处发生约瑟夫森效应，形成超流体物质波。当外界有微小角速度输入时，超流体陀螺能以极高的灵敏度敏感并通过物质波干涉效应引起超流体陀螺的薄膜产生振动位移，使得薄膜的幅值被输入角速度调制。通过高精度的超导量子干涉仪对薄膜振动检测并记录下来，结合幅值-角速度解算关系便可求解出相应的角速度，实现角速度的敏感。

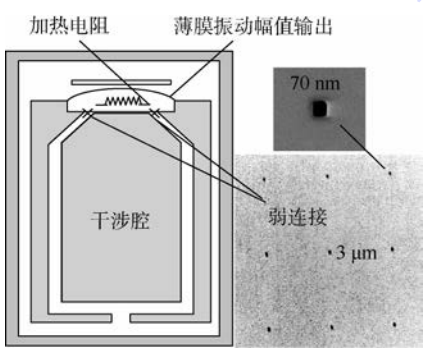


图 1　超流体陀螺示意图

Fig. 1　Schematic diagram of superfluid gyroscope

热驱动装置通过加热铜镍合金的电阻丝来改变干涉腔内的温度和压力，使弱连接两端形成一定的驱动势能[7]：

$$\Delta u = m(\Delta p/\rho - l\Delta T) \tag{1}$$

式中：Δu 为化学势差；m 和 ρ 分别为超流体液氦Ⅱ的原子质量和密度；l 为液氦Ⅱ超流体熵密度；ΔT 和 Δp 分别为热驱动装置控制弱连接两端温差和压差。

在驱动势能的作用下，超流体流经 2 个弱连接，并在双弱连接的微孔处分别形成临界流量幅值为 I_c、相位为 $\Delta\phi_0$ 和 $\Delta\phi_1$ 的物质波。考虑双弱连接的结构完全对称相同，则流经 2 个弱连接的对称微孔的流量也相同，管道内的总流量为[7]

$$I(t) = 2I_c\cos\left(\frac{\Delta\phi_0 - \Delta\phi_1}{2}\right)\sin\left(\frac{\Delta\phi_0 + \Delta\phi_1}{2}\right) \tag{2}$$

由于超流体物质波的 Sagnac 效应，当外部以角速度 $\boldsymbol{\omega}$ 旋转时，在超流体陀螺干涉回路中会形成与角速度量相对应的 Sagnac 相位差：

$$\Delta\phi_{\omega} = \frac{4\pi m}{h}\boldsymbol{\omega}\cdot\boldsymbol{A} \tag{3}$$

式中：h 为普朗克常数；$\boldsymbol{A}$ 为感应面积的矢量。在超流体陀螺工作回路中，两弱连接的相位差和外部转动引起的相位差满足

$$\Delta\phi_{-} = \Delta\phi_0 - \Delta\phi_1 = \frac{4\pi m}{h}\boldsymbol{\omega}\cdot\boldsymbol{A} \tag{4}$$

在热驱动势能作用下超流体陀螺开始工作，超流体陀螺环路中相位变化和驱动势差的关系满足约瑟夫森效应原理[7]

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = -\frac{2\pi\Delta u}{h} \tag{5}$$

进一步对式(5)沿着双弱连接干涉腔积分得到环路内总的超流体相位为

$$\Delta\phi_{+} = \Delta\phi_0 + \Delta\phi_1 = \frac{-4\pi t\Delta u}{h} \tag{6}$$

考虑弱连接所含孔数为 N，将式(4)、式(6)代入式(2)，可得干涉腔内总的超流体流量为[13]

$$I(t) = 2NI_c\left|\cos\left(\frac{2\pi m}{h}\boldsymbol{\omega}\cdot\boldsymbol{A}\right)\right|\sin\left(-2\pi t\frac{\Delta u}{h}\right) \tag{7}$$

在超流体流量作用下，薄膜内腔的体积发生变化，驱动薄膜发生位移变化，弹性薄膜的位移和超流体流量的关系为

$$\rho A_{\mathrm{f}}\dot{x}(t) = I(t) \tag{8}$$

式中：A_{f} 为弹性薄膜的面积。对超流体陀螺质量流量进行积分得到薄膜的位移：

$$x(t) = \frac{NI_c}{\pi\rho A_{\mathrm{f}}f_{\mathrm{J}}}\left|\cos\left(\frac{2\pi m_4}{h}\boldsymbol{\omega}\cdot\boldsymbol{A}\right)\right|(\cos(2\pi f_{\mathrm{J}}t) - 1) \tag{9}$$

北航学报 赠阅

式中:f_J 为约瑟夫森振荡频率;m_4 为氦 4 原子质量。

通过位移传感器实时检测薄膜的位移,结合式(3)、式(7)~式(9)便可解算出外部转动的角速度。

2 幅值锁定量程扩展方案

外部转动引起的相位,对弱连接处产生的超流体流量振动信号的振幅进行了调制。由式(9)得到薄膜振动幅值为

$$X=\frac{NI_c}{\pi\rho A_1 f_J}\left|\cos\left(\frac{2\pi m_4}{h}\boldsymbol{\omega}\cdot\boldsymbol{A}\right)\right| \tag{10}$$

定义 $F=2m_4\boldsymbol{\omega}\cdot\boldsymbol{A}/h$ 为角速度引起的旋转通量,得到角速度引起的相位变化量 πF 和薄膜振动幅值的超流体陀螺工作曲线如图 2 所示。

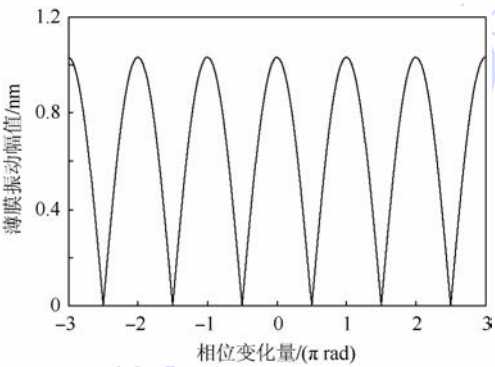


图 2 超流体陀螺工作曲线

Fig. 2 Superfluid gyroscope working curve

从图 2 可以看出角速度引起的相位变化和薄膜振动幅值之间呈周期变化关系。通过幅值检测得到薄膜的振动幅值,当角速度输入引起的相位变化超过区间(0,π/2)时,由幅值解算得到相位变化量只是 πF 的余量 $\pi F-n\pi/2$,并非真实的相位变化量 πF。只有当 πF 在(0,π/2)的区间内,幅值和角速度才是单值对应关系,可直接解算角速度 $\boldsymbol{\omega}=F\cdot h/2m_4\boldsymbol{A}$。综上分析,基于双弱连接的超流体陀螺实现角速度检测的工作区间仅为(0,π/2),直接限制了陀螺的量程范围。

为此,设计如图 3 所示的幅值锁定量程扩展方案,通过加热电阻,使得热相位注入模块开始启动。当外界角速度引起相位变化 πF 时,通过热相位注入 $\Delta\phi_h$ 来抵消这种相位变化,将陀螺继续维持在某一固定相位处,则对应的薄膜振动幅值也将维持在恒定位置。

理论上陀螺的量程可以通过注入补偿实现无限的扩展。工作点相位可在区间(0,π/2)内取值,此处选取 π/4 作为系统的工作锁定相位。热相位补偿量为 $\Delta\phi_h$,则工作中满足

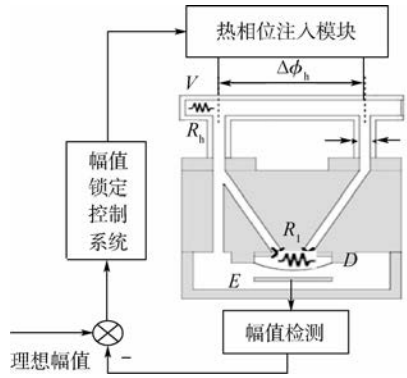


图 3 幅值锁定量程扩展方案图

Fig. 3 Amplitude-locking range extension scheme

$$\pi F+\Delta\phi_h=\pi/4 \tag{11}$$

此时可直接反解得到外部转动角速度为

$$\boldsymbol{\omega}=\frac{h}{2\pi m\boldsymbol{A}}(\pi/4-\Delta\phi_h) \tag{12}$$

考虑在长度为 L、半径为 R 的管道内,流体的流速 v 与管路压力的变化 Δp 的关系[7]

$$v=\frac{R^2\Delta p}{8\delta L} \tag{13}$$

式中:δ 为普通流体的黏度。

又 Δp 与管内温度的变化 ΔT 的关系为

$$\Delta p=k\Delta T \tag{14}$$

式中:k 为温度差与压力差的相关系数。

由质量守恒定律对环内的流体有

$$v\rho_1=v\rho_2 \tag{15}$$

式中:v 为超流体的流速;ρ_1 为超流部分的密度;ρ_2 为常流部分的密度。

将式(13)和式(14)代入式(15)得超流体流速:

$$v=\frac{k\rho_2R^2}{8\rho_1\delta L}\Delta T \tag{16}$$

超流体的相位差 $\Delta\phi$ 与超流体的流速 v 满足

$$\Delta\phi=mv/h \tag{17}$$

将式(16)代入式(17),最终得到注入的热相位与环内的温度关系为

$$\Delta\phi_h=\frac{km\rho_2R^2}{8h\rho_1\delta L}\Delta T \tag{18}$$

3 超流体陀螺幅值锁定优化设计

3.1 幅值锁定系统设计

由式(18)可知,注入的热相位与腔内的温度变化成正比例关系,而腔内温度变化主要是通过加热电阻实现,由于加热电阻的温升过程一般近似为一阶惯性环节 $1/(1+Ts)$,其中 T 为与结构相关的时间常数,s 为复变量,所以热相位注入必

北航学报 赠阅

然存在延迟影响。又由式(12)可知，角速度解算依赖于注入的热相位，热相位注入不及时将导致超流体的幅值不能精准地锁定在目标位置，直接影响超流体陀螺的角速度解算精度。

考虑陀螺包含位移检测的积分环节，整体陀螺系统是一个非线性时变系统，为此设计具有高动态的实时补偿的幅值锁定控制系统，使得热相位能实时补偿角速度输入引起的相位变化，从而维持薄膜的幅值始终锁定在固定值处。设计如图 4所示的模糊自适应 PID 幅值锁定控制系统。

图 4 中 e、e_c 分别表示期望薄膜幅值与实际薄膜的幅值差和幅值差的变化率。通过实时检测 e 和 e_c，将检测量通过量化处理得到 E、E_C，经过模糊推理得出不同时刻 E、E_C 对应的 PID 输出比例、积分、微分的模糊量参数 K_p、K_i 和 K_d。经过解模糊得到对应的模糊自适应 PID 参数实际修正 $\Delta K(k_p, k_i, k_d)$；结合 PID 初始参数 $K(k_{p0}, k_{i0}, k_{d0})$，最终模糊自适应 PID 控制器的控制参数为 $K'(k_{p'}, k_{i'}, k_{d'})$：

$$K' = K + \Delta K \tag{19}$$

式中：k_{p0}、k_{i0}和 k_{d0}分别为 PID 控制器的初始比例、积分、微分参数；k_p、k_i 和 k_d 分别为比例、积分、微分参数修正值；$k_{p'}$、$k_{i'}$和 $k_{d'}$分别为模糊自适应 PID 控制器的比例、积分、微分参数。

则模糊自适应 PID 输出的实际控制量为

$$u(n) = k_p e(n) + k_i \sum e(n) + k_d e_c(n) \tag{20}$$

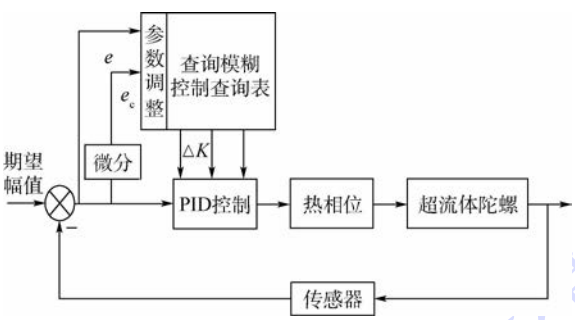


图 4　幅值锁定控制系统示意图

Fig. 4　Schematic diagram of amplitude-locking control system

3.2　引入重叠因子的优化设计

模糊控制器虽然具有鲁棒性强、无需被控对象精确建模等优点[16]，但在参数选取方面大多还是依赖工作经验，人为性和主观性强[17]。遗传算法对所求解的优化问题没有太多的数学要求，有良好的进化特性，寻优过程不依赖问题的内在性质，可处理任意形式的目标函数[18]。

输入输出在论域的隶属函数分布通常为如图 5所示的典型等腰三角形隶属函数[19]。图 5 中：O_1、O_2、O_3、O_4、O_5、O_6、O_7 分别为隶属函数 NL(负大)、NM(负中)、NS(负小)、ZO(零)、PS(正小)、PM(正中)、PL(正大)的中心点；d_1、d_2、d_3、d_4、d_5、d_6 分别为各隶属函数的中心点到右边界点的宽度，称为隶属函数的右半宽度；a_1、a_2、a_3、a_4、a_5、a_6 为相邻隶属函数交点的纵坐标，即重叠因子。

对 PS 和 PM 隶属函数分析，其分布如图 6 所示。

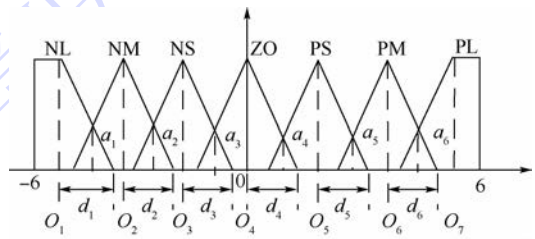


图 5　输入输出隶属函数分布

Fig. 5　Input-output membership function distribution

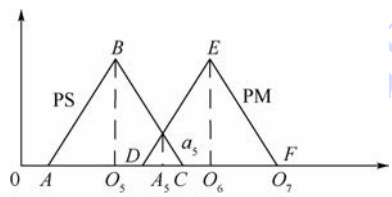


图 6　PS 和 PM 的隶属函数分布

Fig. 6　PS and PM membership function distribution

三角形隶属函数由中心、左半宽度和右半宽度共同决定。考虑遗传优化中，可能会出现相邻模糊隶属函数没有交集或者重叠部分太多导致优化求解速率下降或者出现无用解集。在此引入重叠因子的概念，即相邻隶属函数的交点的纵坐标定义为重叠因子。直接对三角形的右半宽度进行编码，而对于其左半宽度则是通过引进重叠因子来间接编码，同时设定重叠因子在合理的区间[0.2　0.6]。

由三角形相似定理有在$\triangle BCO_5$ 中

$$|A_5O_5|/|CO_5| = (1 - a_5)/|BO_5| \tag{21}$$

则 $|A_5O_5| = d_5(1 - a_5)$，其中 d_5 表示 PS 隶属函数的右半宽度。在$\triangle DEO_7$ 中，$|A_5O_6| = d_1(1 - a_5)$，其中 d_1 表示 PS 隶属函数的左半宽度。

又由于

$$|A_5O_6| = (O_6 - O_5) - |A_5O_5| \tag{22}$$

综合式(21)、式(22)得到隶属函数的左半宽度

$$d_1 = (O_6 - O_5)/(1 - a_5) - d_5 \tag{23}$$

北航学报 赠阅

这样通过引入重叠因子,结合隶属函数右半半宽度便实现了隶属函数左半宽度的间接表达。对于本文设计的模糊推理系统,包含 2 个输入 E、E_C,3 个输出 K_p、K_i、K_d,需要编码的有 6 个相应的右半宽度、6 个重叠因子和 7 个中心,根据每个参数的范围(U_{min},U_{max}),依据式(24)再对 M 取整,便实现了将区间上的实数 m 转变为 10 位二进制码。

$$M=(2^{10}-1)(m-U_{min})/(U_{max}-U_{min}) \tag{24}$$

优化过程中首先将每个参数优化成二进制串,再串联成为一个染色体,通过对该染色体进行遗传操作直到求得最优解 G(G 为使代价函数值取最小时对应的解值)。随后通过拆分染色体得到对应最优参数,按照式(25)进行解码,将得到的数 G 转化为区间(U_{min},U_{max})上的实数

$$m=U_{min}+G(U_{max}-U_{min})/(2^{10}-1) \tag{25}$$

适应度函数 f 的设计需避免寻优开始较早出现过大的适应度值,而湮灭其他个体[18]。防止控制能量过大,目标函数中加入控制输入项,同时综合考虑薄膜幅值锁定系统的动态性能和超调,将惩罚函数融于优化指标,设计参数代价函数为

$$J=\begin{cases}\int_0^\infty(\omega_1+\omega_4)|e(t)|+\omega_2u^2(t)\mathrm{d}t+\omega_3t_u & e<0\\ \int_0^\infty(\omega_1|e(t)|+\omega_2u^2(t))\mathrm{d}t+\omega_3t_u & e\geqslant 0\end{cases} \tag{26}$$

式中:ω_1、ω_2、ω_3 和 ω_4 为权值,$\omega_4\gg\omega_1$;$e(t)$ 和 $u(t)$ 分别为误差和控制器输出;t_u 为上升时间。

根据适应度函数 $f=1/J$,个体 i 被选中的概率正比于个体的适应度值。个体的复制概率和其适应度值成比例,即

$$p_i=f_i\Big/\sum_{i=1}^{n}f_i \tag{27}$$

式中:n 为群体大小;f_i 为第 i 个个体的适应度。

间接编码隶属函数左半宽度后,通过遗传寻优的实现过程是通过模糊推理系统计算出在参数范围内取不同值时,超流体陀螺薄膜幅值的实际输出幅值和此时对应的指标值,从种群中选择复制群体,同时保留每一代的最优个体,完成交叉和变异操作,形成下一代新个体,重复操作,在进化代数范围内,综合考虑系统动态性和超调性的代价函数,得到模糊自适应 PID 控制器最优参数。

4 仿真验证

采用间接编码三角形隶属函数左半宽度的遗传优化算法,首先初始化输入输出隶属函数为等腰三角形隶属函数,且相邻两三角形隶属函数的交点垂直坐标均为 0.5,即重叠因子为 0.5;接着初始化遗传算法的参数;最后利用模糊推理系统计算出此时系统的实际输出,计算对应的最优指标。样本个数为 30 个,交叉、变异概率分别为0.93、0.05,取 $\omega_1=0.92$,$\omega_2=0.98$,$\omega_3=2.5$,$\omega_4=100$。超流体陀螺弱连接的微孔数目为4 225 个,工作温度为 2.175 K 时,根据二流体模型,此时超流体密度为 3 kg/m^3,普通液氦Ⅱ的密度为 143 kg/m^3,超流体流过弱连接的临界流量为 5.2×10^{-12} kg/s,弹性薄膜的面积为 0.5×10^{-4} m^2,发生约瑟夫森振荡的频率为 1 000 Hz,相位锁定值设为 $\pi/4$,注入热相位上限为 500π。

为了验证本文对超流体陀螺系统优化设计的有效性和可行性,在 MATLAB/Simulink 环境中,基于重叠因子编码后,经遗传寻优得到优化前后输入隶属函数参数表(表 1)及优化前后输出隶属函数参数表(表 2);图 7 为经过间接编码隶属函数左半宽度后,优化得到的最终输入-输出 K_p、K_i、K_d 模糊推理特性曲面。

表 1 优化前后输入隶属函数参数

Table 1 Original and optimized input membership function parameters

隶属函数	优化前 $E(E_C)$	优化后	
		E	E_C
NL	[-6 -6 -4]	[-6 -6 -3.922]	[-6 -6 -4.048]
NM	[-6 -4 -2]	[-5.46 -3.48 -1.857]	[-5.8 -3.8 -2.016]
NS	[-4 -2 0]	[-3.41 -1.476 -0.146]	[-3.44 -1.698 0]
ZO	[-2 0 2]	[-0.9365 -0.126 0.832]	[-0.968 -0.0952 1.349]
PS	[0 2 4]	[0.254 1.127 3.41]	[0.3492 1.603 3.946]
PM	[2 4 6]	[1.73 3.476 5.32]	[1.921 3.4 6]
PL	[4 6 6]	[3.81 6 6]	[4.27 6 6]

北航学报 赠阅

表 2　优化前后输出隶属函数参数

Table 2　Original and optimized output membership function parameters

隶属函数	优化前 K_p、K_i、K_d	优化后		
		K_p	K_i	K_d
NL	[-0.6　-0.6　-0.4]	[-0.6　-0.6　-0.395]	[-0.6　-0.6　-0.4402]	[-0.6　-0.6　-0.3952]
NM	[-0.6　-0.4　-0.2]	[-0.6　-0.45　-0.2]	[-0.547　-0.3571　-0.183]	[-0.5444　-0.375　-0.176]
NS	[-0.4　-0.2　0]	[-0.36　-0.2　0]	[-0.403　-0.1698　-0.00317]	[-0.36　-0.17　-0.01746]
ZO	[-0.2　0　0.2]	[-0.08　0　0.12]	[-0.0746　0　0.1444]	[-0.1063　0.00317　0.132]
PS	[0　0.2　0.4]	[0　0.1286　0.29]	[0.0302　0.1635　0.38]	[0.03333　0.16　0.367]
PM	[0.2　0.4　0.6]	[0.144　0.3063　0.6]	[0.22　0.354　0.6]	[0.22　0.3921　0.6]
PL	[0.4　0.6　0.6]	[0.3698　0.6　0.6]	[0.4333　0.6　0.6]	[0.3952　0.6　0.6]

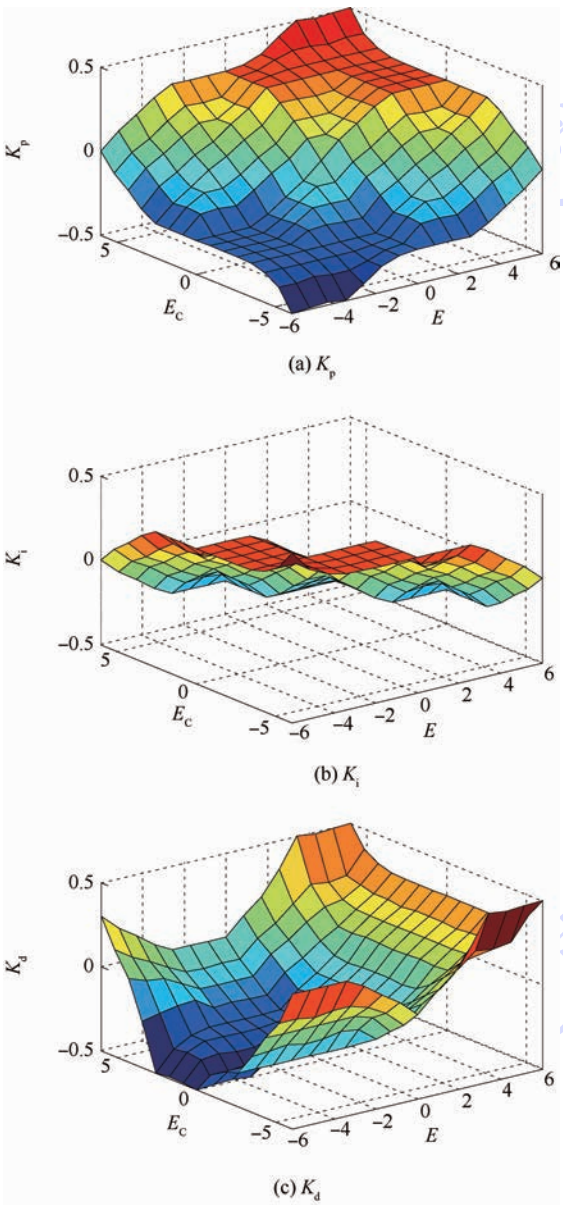


图 7　输入-输出 K_p、K_i、K_d 推理特性曲面

Fig. 7　Input-output reasoning feature surface of K_p, K_i and K_d

从优化前后的输入输出隶属函数参数可以看出，优化后的隶属函数分布不再是简单的均匀三角形分布。隶属函数的形状在误差较小的论域，对应的坡度较陡，此时系统具有较好的灵敏度；在误差较大的论域，对应的隶属函数形状较缓，此时系统的隶属函数分辨率较低，控制特性较平缓，利于系统的稳定性。通过引入重叠因子间接编码三角形隶属函数，克服了优化结果可能存在相邻隶属函数无交点或重叠太多等无效解的情况，提高了优化效率。

通过模糊推理特性曲面可以直观分析出控制系统输入输出特性，曲面上某一点的三轴坐标表征着相应的控制参数。其中水平坐标对应不同的 E、E_C 值；竖直坐标对应不同 E、E_C 取值时，模糊推理系统的输出控制量 K_p、K_i、K_d，推理特性曲面较为平滑，表征控制较为理想。

代价函数 J 的优化过程如图 8 所示，随着优化代数逐渐增大，代价函数值逐渐减小，最终趋于稳定的较小值，此时对应参数解是最优的。

将优化后的参数作为幅值锁定系统控制参数，图 9 为传统 P 控制和本文方法下超流体陀螺系统的薄膜幅值锁定对比情况。传统 P 控制下薄膜幅值在期望幅值 73 nm 附近波动，波动幅值约为 100 nm，由于热相位注入延迟的影响，导致薄膜幅值始终不能很好地锁定在期望值；经过遗传优化的模糊控制下，由局部放大图可知，陀螺系统在0.17s时，薄膜幅值已经能很好地锁定在理想

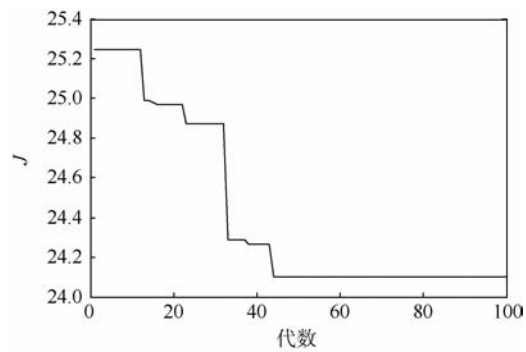


图 8　代价函数 J 的优化过程

Fig. 8　Optimization process of cost function J

北航学报 赠阅

值,且能长时间地工作在锁定幅值处,没有明显的幅值波动。

图 10 为传统 P 控制和本文方法下超流体陀螺系统的幅值锁定误差对比情况。从局部放大图可知,传统 P 控制下系统始终有较大的锁定误差,陀螺系统在 0.17 s 时的锁定误差比本文方法下的误差高将近 23 倍。在 0.17 s 之后,系统幅值波动始终在 13.7% 左右,这是由于温升延迟和控制欠佳共同造成的;而经过遗传优化模糊控制下的陀螺系统在 0.17 s 后,相对期望值的幅值波动约为 0.3%,这个误差对角速度解算的影响已经很小了,可实现较高精度的角速度测量。

综上所述,通过引入重叠因子间接编码的隶属函数,经过遗传寻优,使得隶属函数的形状根据误差论域的大小而自适应地进行调整,避免了单一均匀隶属形状选取盲目性。同时遗传优化模糊控制改善了超流体陀螺系统的动态响应性能,使系统达到稳定的时间大大提高。克服了由于加热电阻升温延迟对陀螺系统的影响,使得薄膜的幅值尽快且良好地锁定在期望薄膜幅值,提高了陀螺系统的敏感精度。

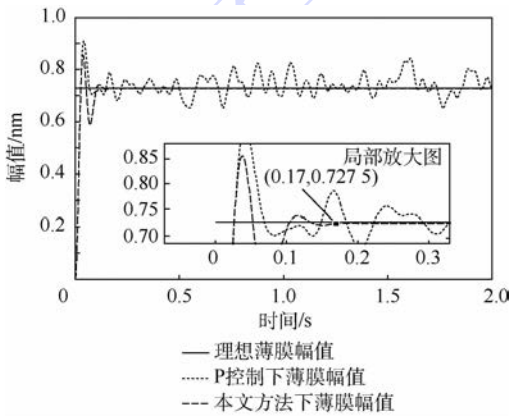


图 9 薄膜幅值锁定对比图

Fig. 9 Film amplitude-locking contrast diagram

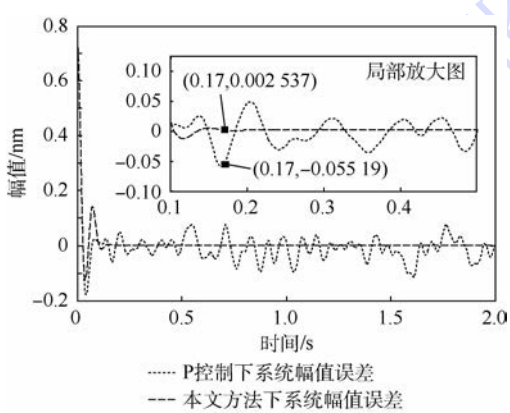


图 10 薄膜幅值锁定误差对比图

Fig. 10 Film amplitude-locking error contrast diagram

图 11 为未施加锁定控制和施加锁定控制后超流体陀螺的角速度测量对比情况。

从图 11 可以看出,幅值锁定控制前后的陀螺最小检测角速度曲线重合,可见幅值锁定控制方法并不改变陀螺的最小检测角速度。

外界角速度输入时,薄膜幅值将被调制,当调制量恰与测量元件能检测最小薄膜位移的精度相等时,对应的角速度值为陀螺的最小检测角速度。受限于目前检测薄膜幅值的超导量子干涉仪检测精度仅为 10^{-15} m,在敏感面积10 cm^2 时超流体陀螺的最小检测角速度为 2.195×10^{-8} rad/s。此外随着敏感面积的增大,陀螺的最大检测角速度变小,限制超流体陀螺使用的关键是如何扩展其角速度测量量程上限,陀螺做的越小,其敏感面积就越小,对应可实现的检测角速度越大。在敏感面积为 10 cm^2 的情况下,幅值锁定前后的最大检测角速度由 2.476×10^{-5} rad/s 增大到 2.476×10^{-2} rad/s,可见幅值锁定控制较好地实现了陀螺的量程扩展。

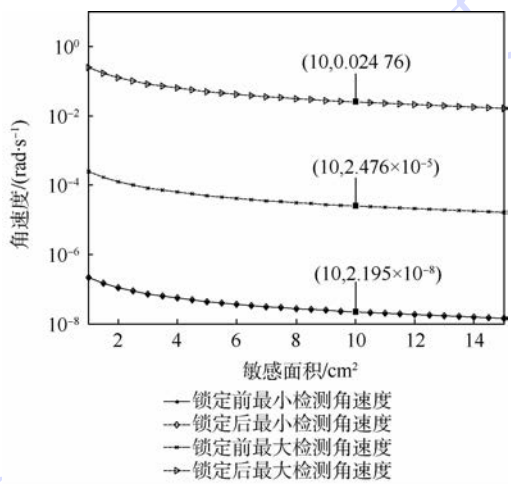


图 11 角速度测量对比图

Fig. 11 Angular velocity measurement contrast diagram

5 结 论

热驱动型双弱连接超流体陀螺利于物质波的干涉效应可实现极高灵敏度的角速度检测,然而目前其角速度检测范围仅为 $10^{-8}\sim10^{-5}$ rad/s,本文对此引入幅值锁定概念进行了量程扩展方案的研究,得到以下结论:

1) 基于重叠因子优化的幅值锁定控制系统可实现较好的薄膜幅值锁定效果,在较短调节时间内将幅值波动控制在 0.3% 左右,可实现较高精度的角速度测量。

2) 幅值锁定控制方案实现了超流体陀螺角速度检测范围的扩展。使用超导量子干涉仪检测

北航学报 赠阅

薄膜位移时,在相同敏感面积下,最大检测角速度提高了 3 个数量级,最小检测角速度不变。为实现陀螺的更大量程范围,在施加幅值锁定控制的同时,可使用更高分辨率的检测器件,且陀螺面积也应该做得越小越好。

参考文献(References)

[1] FANG J C, SUN J J, XU Y L, et al. A new structure of permanent magnet biased axial hybrid magnetic bearing [J]. IEEE Transactions on Magnetics, 2009, 45(12): 5319-5325.

[2] HAN B C, HU G, FANG J C. Optimization design of magnetic suspended gyroscope rotor[J]. Optics and Precision Engineering, 2006, 14(4): 662-666.

[3] KAPALE K T, DIDOMENICO L D, LEE H. Quantum interferometric sensors[C]// Proceedings of SPIE -The International Society for Optical Engineering, V 6603, 2007, Noise and Fluctuations in Photonics, Quantum Optics, and Communications. Bellingham, WA: SPIE, 2007.

[4] HODBY E, HOPKINS S A. Experimental observation of a superfluid gyroscope in dilute Bose-Einstein condensate [J]. Physics Review Letters, 2003, 91(9): 090403.

[5] PACKARD R. The role of the Josephson-Anderson equation in superfluid helium [J]. Reviews of Modern Physics, 1998, 70(2): 641-651.

[6] HOSKINSON E. Superfluid 4He weak links [D]. Berkeley: University of California, 2005: 5-9.

[7] GOLOVASHKIN A I, ZHERIKHINA L N, TSKHOVREBOV A M, et al. Ordinary SQUID interferometers and superfluid helium matter wave interferometers: The role of quantum fluctuations [J]. Journal of Experimental and Theoretical Physics, 2010, 111(2): 332-339.

[8] HAKONEN A P, VAROQUAUX E. Detection of the rotation of the earth with a superfluid gyrometer[J]. Physical Review Letters, 1997, 78(19): 3602-3605.

[9] PACKARD R E. Principles of superfluid helium gyroscopes [J]. Physical Review B, 1992, 46(6): 3540-3549.

[10] HOSKINSON E, PACKARD R. Thermally driven Josephson oscillations in superfluid 4He[J]. Physical Review Letters, 2005, 94(15): 155303.

[11] SATO Y, ADITYA J, PACKARD R. Flux locking a superfluid interferometer[J]. Applied Physics Letters, 2007, 91(7): 1-3.

[12] SATO Y. Experiments using 4He weak link [D]. Berkeley: University of California, 2007: 51-58.

[13] 谢征. 基于物质波干涉效应的新型低温超流体陀螺关键技术研究[D]. 南京:南京航空航天大学, 2012: 121-125.
XIE Z. Research on the key techniques for new cryogenic superfluid gyroscope based on matter wave interference effect[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2012: 121-125(in Chinese).

[14] STIMAC G, BRAUT S, ZIGULI R. Comparative analysis of PSO algorithms for PID controller tuning[J]. Chinese Journal of Mechanical Engineering, 2014, 27(5): 928-936.

[15] YANG Y Q, WANG G J. Parameters optimization of polygonal fuzzy neural networks based on GA-BP hybrid algorithm[J]. International Journal of Machine Learning and Cybernetics, 2014, 5(5): 815-822.

[16] SONG S J, MOON Y J, LEE D H, et al. Comparative study of fuzzy PID control algorithms for enhanced position control in laparoscopic surgery robot[J]. Journal of Medical and Biological Engineering, 2015, 35(1): 33-44.

[17] ESFANDYARI M, FANAEI M A, ZOHREIE A. Adaptive fuzzy tuning of PID controllers[J]. Neural Computing and Applications, 2013, 23(1): 19-28.

[18] ZHANG N N, GUO Y B. Control strategy for hybrid electric vehicle based on fuzzy logic [J]. International Journal of Automotive Technology, 2012(112): 563-571.

[19] HAN J Q. From PID to active disturbance rejection control[J]. IEEE Transactions on Industrial Electronics, 2009, 56(3): 900-906.

作者简介:

赵玉龙　男,博士研究生。主要研究方向:先进量子陀螺技术。
Tel.: 15801412805
E-mail: 460151954@qq.com

任元　男,博士,讲师。主要研究方向:先进量子陀螺技术。
Tel.: 010-66364384
E-mail: renyuan823@aliyun.com

北航学报 赠阅

Range extension design on superfluid gyroscope based on overlap factor optimization

ZHAO Yulong[1], REN Yuan[2,*], SHEN Huairong[2]

(1. Department of Graduate Management, Equipment Academy, Beijing 101416, China;
2. Department of Space Equipment, Equipment Academy, Beijing 101416, China)

Abstract: Aimed at the limited range caused by single value interval of amplitude-angular velocity of heat driven double weak link superfluid gyroscope whose working curve is periodic, a superfluid gyroscope range extension method based on overlap factor optimization is proposed. First, on the basis of range limiting mechanism of gyroscope, the amplitude-locking range extension scheme is designed. Second, considering the relationship between hot phase and heating delay, with the concept of overlap factor, a genetic optimization method based on overlap factor is proposed to optimize the gyroscope amplitude-locking control parameter. Through indirect coding the left width of triangle membership function, the fuzzy linear membership function is optimized to find the best parameters of amplitude-locking system under the premise of avoiding adjacent membership function with no overlap or too much overlap. Finally, comparison results between this method and traditional P method show that the film amplitude of superfluid gyroscope can be locked to the expectations quickly and stably by this method, which extend the range of the superfluid gyroscope with measurement accuracy guaranteed.

Key words: superfluid gyroscope; range extension; amplitude-locking; hot phase; overlap factor

北京航空航天大学学报

Received: 2015-11-12; **Accepted**: 2015-11-28; **Published online**: 2015-12-09 14:33

URL: www.cnki.net/kcms/detail/11.2625.V.20151209.1433.001.html

Foundation items: National Natural Science Foundation of China (51475472); National High-tech Research and Development Program of China (2015AA8018038C)

* **Corresponding author**. Tel.: 010-66364384 E-mail: renyuan823@aliyun.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0756

基于二级声辐射模型的地面声场高效预测方法

王菲，史勇杰*，徐国华

（南京航空航天大学 直升机旋翼动力学国家级重点实验室，南京 210016）

摘 要：建立了可用于直升机运动状态下噪声传播特性分析的高效地面声场预测方法。该方法包括计算旋翼气动、噪声的自由尾迹方法和时域 FW-H 方程；计入大气、地面等环境因素的声传播模型及地面声场计算模型；为提高计算效率，在声源计算与噪声传播之间引入基于“紧致球声源”的二级声辐射模型；在此基础上，还提出了通过建立“特征参数声辐射球库”以实现直升机运动状态下噪声实时预测的方法。以 AH-1 旋翼为算例，通过与地面声场直接计算法对比，说明了方法的有效性及高效性；此外，文中还分析了大气、地面等环境因素对噪声传播以及地面声场特性的影响。

关 键 词：噪声；旋翼；二级声辐射球；大气吸声；地面吸声、反射
中图分类号：V211.52
文献标识码：A **文章编号**：1001-5965(2016)11-2445-09

随着直升机的广泛应用，直升机噪声大的缺点越来越引起人们的重视。对于军用直升机，噪声大使其过早地暴露目标，与其隐身要求相悖；对于民用直升机，由于常在人口稠密的地区起飞着陆以及低空慢速飞行，噪声污染已成为公害之一。旋翼气动噪声是直升机的主要噪声源，国内外在此方面已开展了很多研究工作。然而这些研究主要集中在旋翼气动噪声分析方法建模[1-2]、发声机理[3-5]、声学特性以及设计降噪[6-7]等方面，对于直升机飞行过程噪声的预测及其地面声场特性的研究则开展较少。

一般而言，地面噪声预测分为噪声源计算、噪声传播及地面噪声合成等部分[8]。若采用直接计算法，需要计算旋翼旋转一周内桨叶上所有辐射点对地面观测点的噪声。当观测点数目较多时，计算将非常耗时。这对于实时分析直升机运动过程中的动态声场特性显然是不现实的。20 世纪 80 年代，Boxwell 等[6]开展了飞行参数对旋翼气动噪声的影响研究，并对旋翼噪声进行了参数化表达。基于此，Fleming[9] 和 Conner[10] 等开展了基于试验测量的直升机噪声实时预测方法研究。首先由外场测量不同飞行状态下的地面声场数据；通过噪声逆向传播技术，将地面声场映射形成声辐射球；建立关于飞行参数的声辐射球库，当计算直升机某一飞行过程中地面声场时，根据状态参数插值获取相应的声辐射球，再将其传播至地面。根据上述研究，目前 NASA 已经发展形成了直升机噪声实时分析系统——Rotor Noise Model(RNM)[11]。

虽然实飞测量可以获得更丰富直升机噪声数据，但是也存在技术难度大、代价高等缺点。为此，本文借鉴 RNM 思想，建立基于数值计算的直升机噪声地面声场实时预测方法：其中声辐射球采用自由尾迹/FW-H 方程进行求解；引入大气、地面等环境因素对噪声传播特性的影响噪声传播模型。以 AH-1/OLS 试验旋翼为算例，分析验证

收稿日期：2015-11-17；**录用日期**：2016-02-29；**网络出版时间**：2016-04-13 15:23
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160413.1523.003.html
基金项目：国家自然科学基金（11302103）
* **通讯作者**：Tel.：025-84892117 E-mail：shiyongjie@nuaa.edu.cn

引用格式：王菲，史勇杰，徐国华. 基于二级声辐射模型的地面声场高效预测方法[J]. 北京航空航天大学学报，2016，42(11)：2445-2453. WANG F，SHI Y J，XU G H. Efficient ground noise prediction method based on secondary noise radiation model[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2445-2453 (in Chinese).

了方法的有效性。

1 计算方法

图 1 是本文建立的方法与直接地面噪声预测方法的对比。地面声场的计算包括旋翼气动/噪声预测、声传播(包含环境因素对噪声传播特性影响)、地面声场合成等。本文在①、②过程间加入了声辐射球,由二级声辐射球来传递旋翼的噪声信息,由此建立了高效的地面声场预测方法。

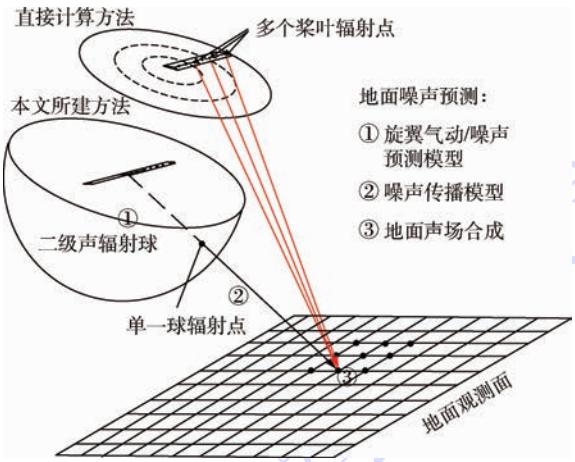


图 1 2 种地面噪声预测方法对比

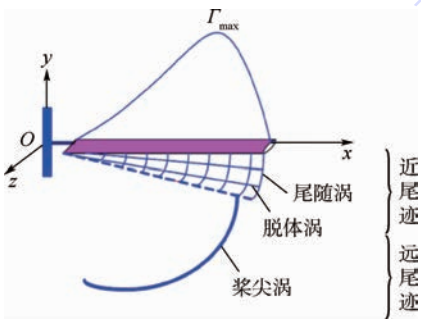
Fig. 1 Comparison between two ground noise prediction methods

1.1 旋翼气动/噪声预测模型

旋翼噪声预测包括气动力和噪声计算两部分。本文中旋翼气动力采用课题组发展的自由尾迹方法进行求解[12]。旋翼尾迹模型由近尾迹和远尾迹两部分构成。近尾迹由桨叶环量沿径向变化引起的尾随涡和沿方位角方向变化引起的脱体涡组成。远尾迹仅包括单根桨尖涡线,由沿桨叶后缘脱出的尾迹在转过 30°~60°方位角后卷起形成,如图 2 所示。桨尖涡线的控制方程为

$$\frac{\mathrm{d}\boldsymbol{r}(\psi,\zeta)}{\mathrm{d}t}=\boldsymbol{V}(\boldsymbol{r}(\psi,\zeta)) \tag{1}$$

式中:$\boldsymbol{r}(\psi,\zeta)$为控制点的位置矢量,$\psi$为桨叶方位角,$\zeta$为尾迹寿命角;$\boldsymbol{V}(\boldsymbol{r}(\psi,\zeta))$为控制点空间运动速度,包括自由流速度、附着涡及其他涡线引起的诱导速度。诱导速度采用毕奥-萨伐尔定理进行求解,涡核模型采用了 Scully 代数涡核模型[13],初始涡核半径为 $0.2c$(c 为桨叶弦长)。



Γ_{max}—桨叶展向最大环量。

图 2 旋翼自由尾迹模型示意图

Fig. 2 Schematic of rotor free-wake model

为更好地模拟桨叶三维效应的影响,采用了 Weissinger-L 升力面理论进行了新生涡的求解,该方法的具体实现可参见文献[14]。

在获得桨叶气动力基础上,可以进行旋翼气动噪声的计算。噪声计算采用基于 FW-H 方程的 Farassat 1A 公式[15]。Farassat 1A 公式将噪声分为两部分,即

$$p'(X,t)=p'_{\mathrm{t}}(X,t)+p'_{\mathrm{l}}(X,t) \tag{2}$$

式中:p'、p'_{t}和 p'_{l}分别为总噪声、厚度噪声和载荷噪声;X 为观测位置;t 为观测时间。

1.2 二级声辐射模型

当噪声传播距离较远时,可将发声体看作点声源[16],从而简化计算。然而先前研究表明,旋翼噪声辐射具有明显的方向性,即不同方位观测点接收到的旋翼噪声大小差别较大[17]。因此,将旋翼噪声看作单纯的点声源将失去传播的方向性特点。借鉴点声源思想,本文建立了基于球声源的二级声辐射模型,如图 3 所示。在该模型中,以

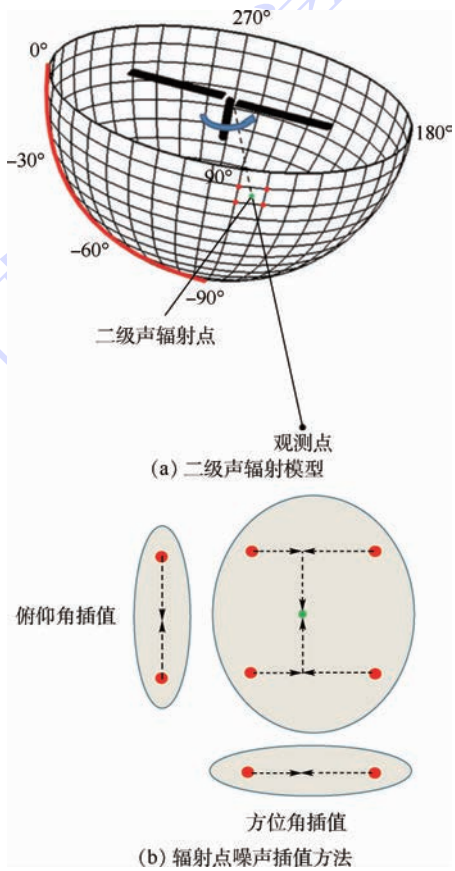


图 3 二级声辐射模型示意图

Fig. 3 Schematic of secondary noise radiation model

位于桨毂处的声辐射球球心代替实际桨叶声源作为噪声辐射中心，声辐射球上等角度间隔的观测点噪声代表了不同方位旋翼噪声辐射的实际声源。旋翼噪声沿从辐射中心到观测点的直线传播，声辐射点为该直线与声辐射球面的交点，噪声信息可通过二维插值方法在相邻 4 个观测点上插值得到。由此将随桨叶旋转的多个运动声源转换为声辐射球面上单一静态辐射点，提高了计算效率。

1.3　噪声传播模型

噪声传播过程中不仅有随着传播距离增加造成的扩散损失，还有大气吸声效应对其造成的衰减和地面对其的反射与阻抗作用[18]。对于距地面一定高度的观测点，其接受到的噪声有直达声波和地面反射声波两部分（见图 4）。地面观测点接收到的噪声声压可表示为

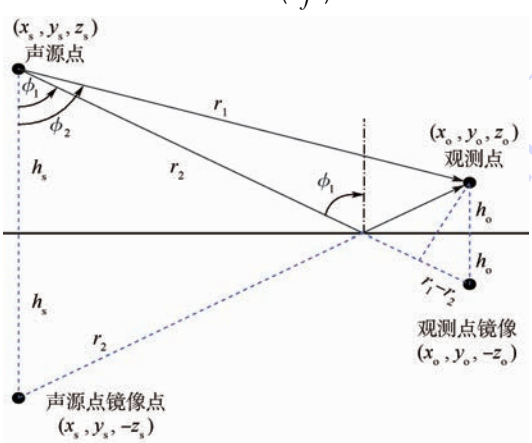
$$p(x,t)=\frac{1}{r_1}p_{\mathrm{ff}}(x,t)\mathrm{e}^{-\alpha r_1}+\frac{1}{r_2}p_{\mathrm{gr}}(x,t)\mathrm{e}^{-\alpha r_2} \tag{3}$$

$$p_{\mathrm{gr}}(x,t)=\Pi(\phi_1,\omega)p_{\mathrm{ff}}(x,t) \tag{4}$$

式中：$p(x,t)$ 为观测点接收到的噪声；$p_{\mathrm{ff}}(x,t)$ 为直达声波；$p_{\mathrm{gr}}(x,t)$ 为地面反射声波；r_1 为直达声波传播距离；r_2 为反射声波传播距离；$\frac{1}{r_{1(2)}}$ 为扩散损失；x 为观测点位置；α 和 $\Pi(\phi_1,\omega)$ 分别为大气吸声系数和地面吸声衰减系数，其表达式如下：

$$\alpha=-8.686f^2\left[1.83\times10^{-11}+\left(\frac{0.012\,75\mathrm{e}^{-2\,239.1/T}}{f^2/f_{\mathrm{rl,O}}+f_{\mathrm{rl,O}}}+\frac{0.11\mathrm{e}^{-3\,352/T}}{f^2/f_{\mathrm{rl,N}}+f_{\mathrm{rl,N}}}\right)\right] \tag{5}$$

$$\Pi(\phi_1,\omega)=\frac{0.218(1+\mathrm{i})\left(\frac{\sigma_{\mathrm{e}}}{f}\right)^{\frac{1}{2}}\cos\phi_1-1}{0.218(1+\mathrm{i})\left(\frac{\sigma_{\mathrm{e}}}{f}\right)^{\frac{1}{2}}\cos\phi_1+1} \tag{6}$$



h_s—声源距地面高度；h_o—观测点距地面高度；

ϕ_1—反射声波入射角；ϕ_2—直达声波入射角。

图 4　观测点噪声接收示意图

Fig. 4　Schematic of noise receiving at observation point

式中：σ_{e} 为地面等效流阻。其他参数意义可参见文献[17]。

需要指出的是，现有的声传播理论一般是建立在频域内，而 2.1 节中旋翼气动噪声的计算则是在时域内开展的。为此，本文采用了傅里叶变换和反变换进行噪声的时、频变换。

1.4　声场计算流程

实际飞行过程中，直升机的飞行姿态（飞行速度、加速度、航迹角等）是不断变化的，每个飞行姿态对应着一个声辐射球。研究表明，定常或小机动飞行状态下旋翼噪声可归结为参数前进比和桨盘倾角的函数[6]。因此，本文基于“参数化建模”思想，建立了以前进比和桨盘倾角为特征参数的声辐射球库。这样在地面噪声预测中，可根据所需状态的特征参数，直接选取相应的声辐射球用于后续噪声传播计算，而不用实时计算相应的声辐射球，从而进一步提高计算效率。

图 5 给出了采用本文所建方法进行地面噪声预测的流程图。计算流程主要分为 3 个部分。首先，建立了以前进比（μ）和桨盘倾角（α_{TPP}）为特征参数的声辐射球库；其次，根据所需计算状态的特征参数，选取相应的声辐射球，确定相应的声辐射点并插值得到相应的噪声信息；最后，进行噪声传播计算并进行相应的噪声后处理。

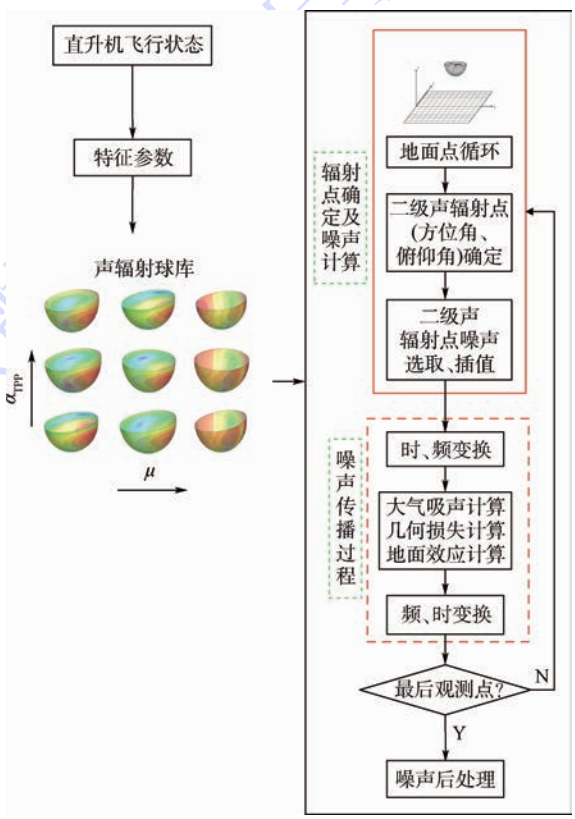


图 5　快速地面声场计算流程图

Fig. 5　Flowchart of quick ground noise computing

北航学报 赠阅

2 结果讨论与分析

2.1 计算细节

本节以 AH-1/OLS 试验旋翼[6]作为算例，对所建地面声场预测方法的有效性进行了分析。OLS 试验旋翼是 AH-1 直升机旋翼的 1/7 缩比模型，共有 2 片桨叶，采用 BHT-540 对称翼型，直径为 1.916 m，展弦比为 9.22，具有 −10°线性扭转，桨叶根切比为 18.2%。算例状态参数为拉力系数 0.005 4，前进比 0.164，总距 6.2°，纵横向周期变距分别为 −2.4°、1.0°。图 6 给出了算例分析中旋翼空间位置及地面观测点分布示意图，其中 O 点为坐标系原点，桨毂中心位于地面中心正上方 50 m 处。参考直升机适航标准中测量点分布，文中设置了 B、C、D 3 个地面观测点，坐标分别为：(100,40,0)，(100,0,0)和(100，−40,0)。此外，图中还给出了半径分别为 25R、50R(R 为旋翼半径)的声辐射球；B_1、C_1、D_1，B_2、C_2、D_2 分别为 B、C、D 观测点噪声传播路径与声辐射球的交点，即噪声传播至 25R、50R 时，3 个地面观测点所在位置。

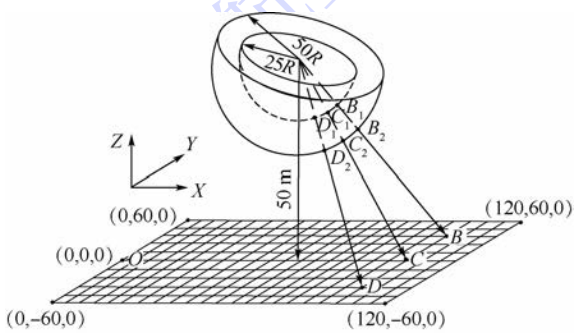


图 6 旋翼空间位置及地面观测点分布示意图

Fig. 6 Schematic of rotor spatial position and distribution of ground observation point

2.2 方法验证

图 7 给出了该状态下旋翼的自由尾迹几何形状。从图 7 中看出，本文方法较好地模拟了旋翼尾迹畸变及旋翼尾迹的运动。

图 8 给出了采用 2 种计算方法得到的半径为 50R 的噪声声压级 SPL(单位为 dB)等高图，其中快速计算法采用了半径为 25R 的二级声辐射球。该噪声等高图是将球面空间投影到二维平面图得出的。投影方法采用了具有良好保角性(即投影图形与原图形相比没有角度变形)的兰勃特投影[19]，这样有利于准确反映出噪声辐射的方向性。图中周向角度表示旋翼旋转的方位角，径向角度表示观测点仰角(仰角为负表示观测点位于旋翼下方)。

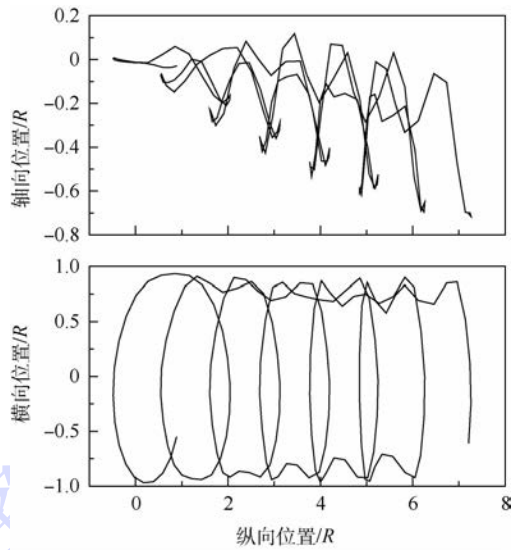


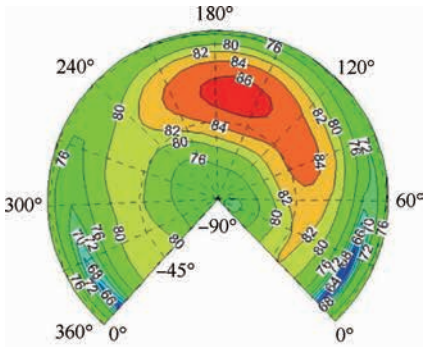
图 7 旋翼桨尖涡尾迹几何形状

Fig. 7 Wake geometry of rotor blade tip vortex

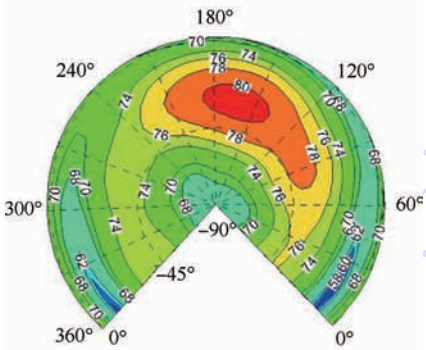
对比图 8(a)～图 8(c)可以看出，相比于 25R 处的噪声声压级 50R 处的噪声声压级降低了约 6 dB，这与传播距离增加一倍声压级降低 6 dB 的噪声衰减结果[20]一致。这说明半径为 25R 的二级声辐射球较好地模拟了旋翼远场噪声及噪声传播模型较好地模拟了噪声的衰减。图 8 中还可以看出，该状态下旋翼最大噪声位于前行侧桨盘平面下方 30°附近，体现出较强的载荷噪声辐射特性，这与文献[12]分析基本一致。此外，对比同一传播距离下由 2 种计算方法计算得出的声压级等高图可以看出，2 种计算方法在噪声幅值及方向性两方面基本一致。

旋翼噪声包含近场噪声及远场噪声两部分。近场噪声随传播距离增加衰减较快而不会传播较远的距离，计算中应尽可能选择较大半径的声辐射球，使近场噪声充分衰减掉。然而过大的声辐射球半径将引起较大的大气吸声计算误差，因此合理选择二级声辐射球是非常重要的。图 8(d)给出了由半径为 20R 的声辐射球计算得到的 50R 处的噪声声压级等高图。与图 8(b)对比可知，虽然该图较好地体现出旋翼噪声传播的方向性，但是最大声压级高出约 2 dB，反映出该半径的声辐射球并没有充分衰减掉近场噪声。综合上述分析可以看出本文所采用二级声辐射球是合理的。

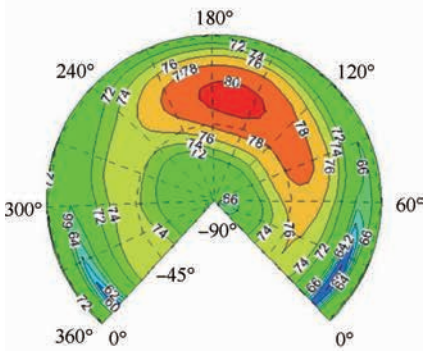
图 9 给出了由 2 种计算方法计算得到 B_2、C_2 和 D_2 3 个球面观测点的声压时间历程。可以看出，二者在幅值及相位两方面基本一致。综合图 8和图 9 的计算分析说明了本文快速地面声场预测方法的有效性。



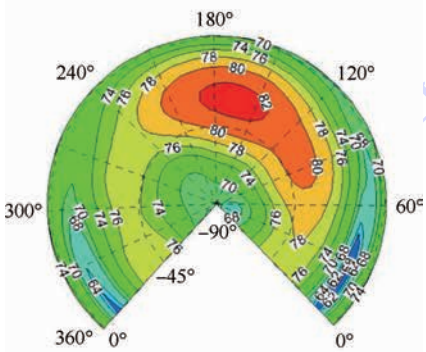
(a) 二级声辐射球(半径25R)



(b) 传播距离50R、辐射球半径25R(快速计算法)



(c) 传播距离50R(直接计算法)



(d) 传播距离50R、辐射球半径20R(快速计算法)

图 8　2 种计算方法噪声声压级计算结果对比

Fig. 8　Comparison of computed noise SPL results of two different methods

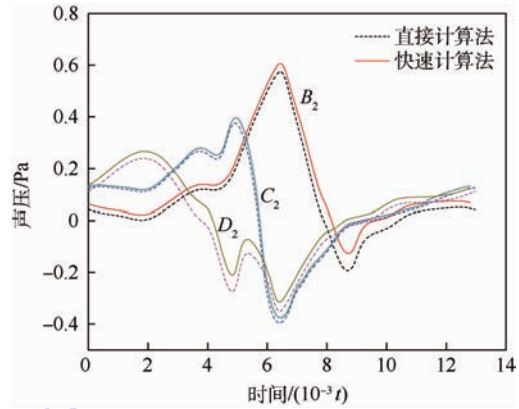


图 9　地面观测点声压时间历程

Fig. 9　Sound pressure time history of ground observation points

表 1 给出了 2 种方法的计算时间对比。算例中桨叶声源辐射点共有 18 × 28 × 201(桨叶分段数 × 弦向分段数 × 观测时间)。在直接计算法中，需要计算所有声源点对地面观测点的噪声辐射；而快速地面声场预测模型中建立了旋翼各方位角噪声辐射特性的二级声辐射球，地面声场预测时，仅需将相应的二级声辐射点噪声映射至地面观测点即可。当地面观测点数目均为 10 万时，前者需要耗时 3.2 h，而后者仅为 305 s。需要指出的是，快速地面声场计算模型虽然可以大幅度提高地面声场计算效率，但是当传播距离较近时，噪声中含有近场项及远场项，噪声衰减并不随传播距离的增加线性衰减，此时应采用直接计算法进行地面声场预测。

表 1　2 种方法的声辐射点数及计算时间对比

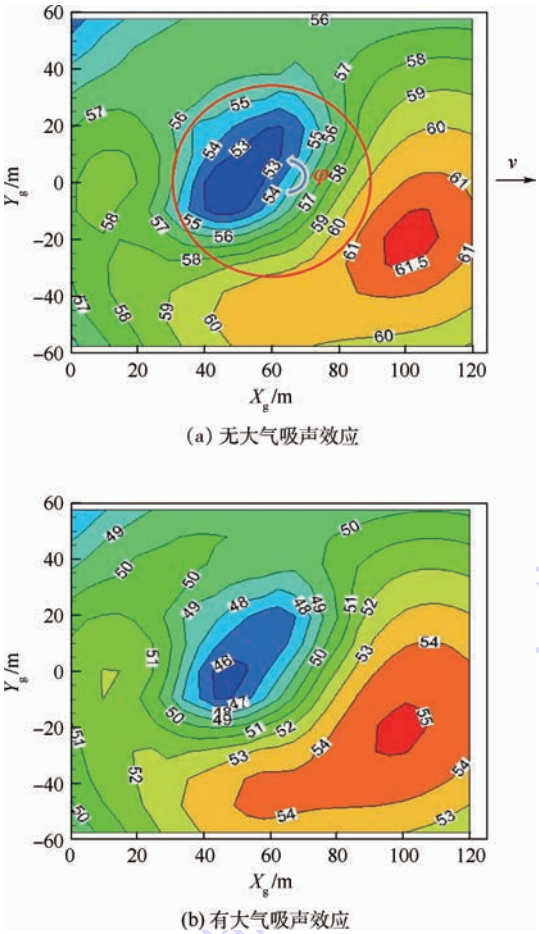
Table 1　Comparisons of radiation points and computation time between two methods

计算方法	声辐射点数目	观测点数目	计算用时
二级声源法	1	10 万	305 s
直接计算法	101 304	10 万	3.2 h

2.3　大气吸声影响分析

本节进行了大气吸声效应对噪声传播特性的影响分析。图 10 对比了有、无大气吸声效应情况下地面观测面的噪声声压级(单位为 dB)等高图。算例中直升机前飞方向(v)沿 X_g 轴正向，旋翼投影位于地面中心，如图 10(a)所示。可以看出，该状态下最大噪声位于地面右前方、最小噪声出现在地面中心附近。此外，由于大气吸声作用地面旋翼噪声降低了约 6.5 dB，但大气吸声效应并没有改变噪声辐射的方向性。图 11 为地面观测点 B、C、D 的声压时间历程。可以看出大气吸声对不同地面观测点作用处处相似，仅改变了噪声的幅值大小而没有改变噪声的相位特征。

北航学报 赠阅



(a) 无大气吸声效应

(b) 有大气吸声效应

图 10 大气吸声对噪声传播特性影响

Fig. 10 Effects of atmosphere absorption on noise radiation characteristics

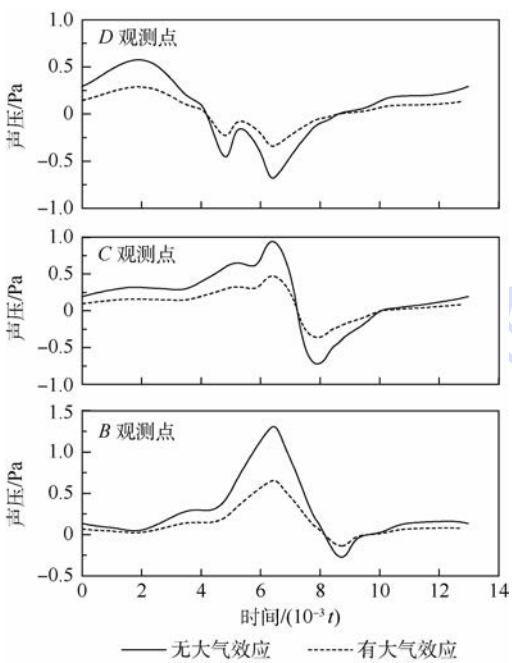
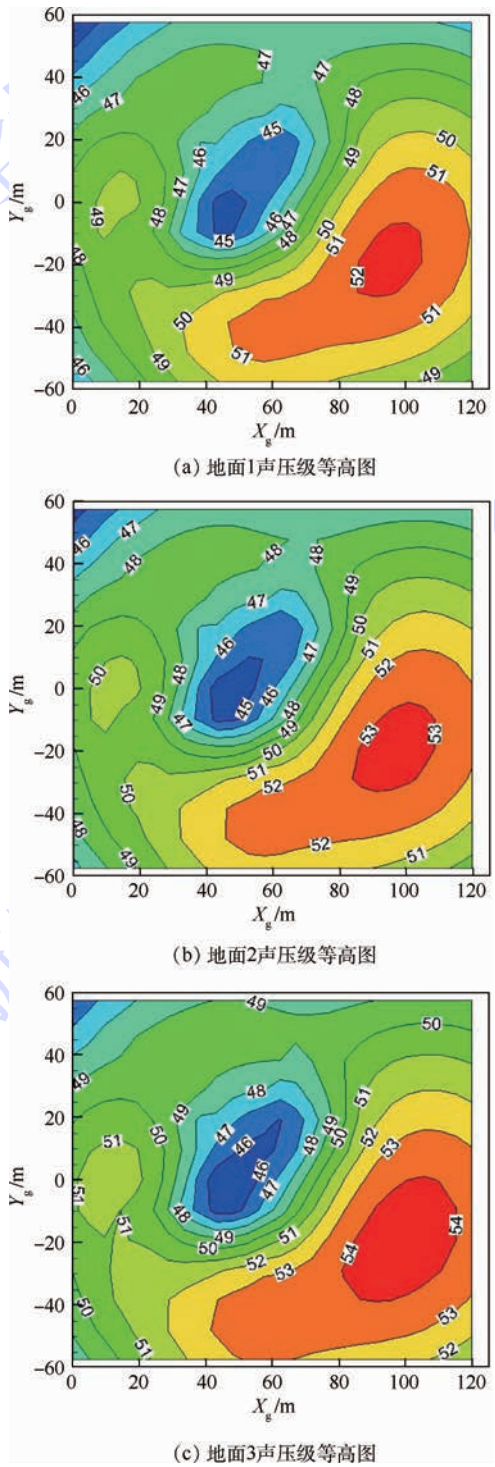


图 11 大气吸声对噪声声压时间历程的影响

Fig. 11 Effects of atmosphere absorption on noise sound pressure time history

2.4 地面效应影响分析

图 12 进一步分析了地面吸声效应对旋翼噪声传播特性的影响，给出了 3 种常见地面（地面类型及地面声学特性描述参数如表 2[21] 所示）的噪声声压级（单位为 dB）等高图。对比图 12 与图 10可以看出，地面吸声效应除进一步降低噪声强度外，对噪声的传播方向几乎没有影响。此外对比不同地面声压级等高图可以看出，流阻对地



(a) 地面1声压级等高图

(b) 地面2声压级等高图

(c) 地面3声压级等高图

图 12 不同地面类型对噪声传播特性影响

Fig. 12 Effects of different grounds on noise radiation characteristics

北航学报 赠阅

表 2　不同地面声学特性参数[21]

Table 2　Acoustic characteristic parameters of different grounds[21]

编号	地面类型	空隙率	流阻/(kg·s^{-1}·m^{-3})
地面 1	间隙填以碎石的旧土路	0.400	7 500
地面 2	裸露并被雨水填实的土地	0.350	17 100
地面 3	被尘封并很少使用的沥青路面	0.250	120 000

面吸声效应具有较大的影响，流阻较大的地面噪声声压级较低。因而，对直升机而言，应尽量选择旧土路等流阻较大的地面飞行以降低直升机的可探测性。图 13 以 C 观测点为例，给出了不同地面类型对旋翼噪声的阻抗作用。可以看出，不同地面下观测点接受到噪声的声压历程变化曲线相位一致，而幅值不同；这说明地面的阻抗作用对噪声幅值影响较大，几乎不会影响噪声的相位特征。

实际的地面噪声接收点（如人耳等）距地面有一定的高度，因此其接受到的噪声不仅含有旋翼直接辐射的直达声波，还有经地面反射后接收到的反射声波（如图 4 所示）。图 14 给出了噪声观测点位于地面观测点 C 点上方 1.0R、2.1R、3.0R时，观测点接收到的声波时域信号。算例中由于观测点距噪声源较远，其高度变化对反射声波入射角影响较小，因此实际计算中采用了直达声波作为反射声波进行计算分析。由图 14 可以看出，不同高度观测点接收到的反射声波与总声波均发生了较明显的变化。反射声波，由于传播距离较大，扩散衰减和大气吸声效应进一步增强，因而声压幅值较小；此外传播距离的改变引起观测点接收到的反射声波相位发生改变。总声波则由于直达声波与反射声波的叠加作用，发生了较明显改变，且当观测点高度为 2.1R 时，直达声波与反射声波相位变化一致，发生“声共振”现象，使得总声波声压幅值明显增大。图 15 给出了观测点接收到的总噪声频域图，可以看出不同高度观测点接收到的总噪声频域信号发生了较为明显的变化。当发生“声共振”时，总声波低频段噪声明显增强，而低频噪声具有较强的穿透性，非常不利于直升机的声隐身。

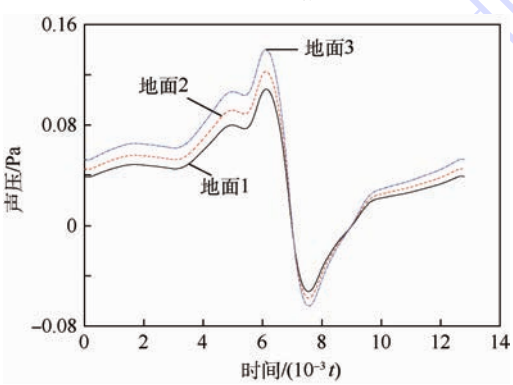


图 13　观测点噪声声压时间历程

Fig. 13　Noise sound pressure time history of observation points

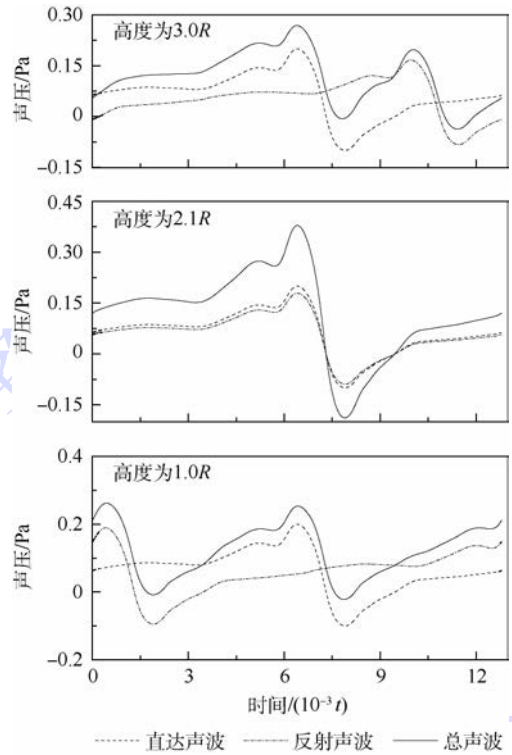


图 14　不同高度观测点接收到的声压时间历程

Fig. 14　Accepted sound pressure time history of observation point at different heights

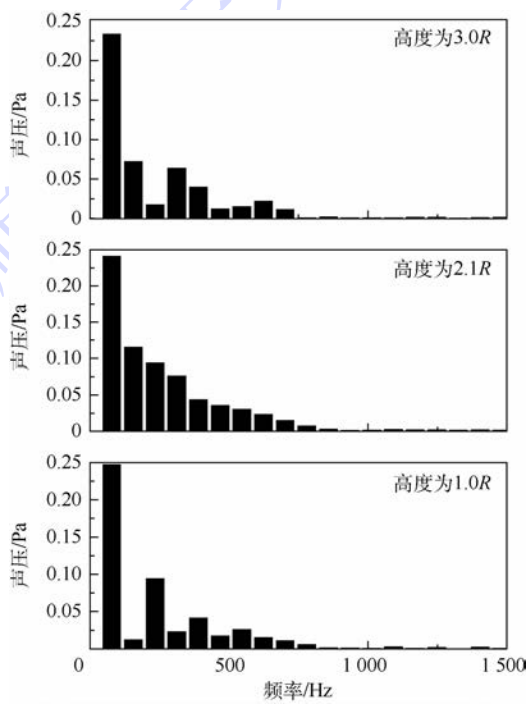


图 15　不同高度观测点总声波频域特性

Fig. 15　Frequency-domain characteristics of total acoustic wave of observation point at different heights

北航学报 赠阅

3 结 论

本文开展了直升机飞行过程中高效地面声场预测方法的研究并建立了相应计算模型,应用所建方法计算分析了大气、地面等环境因素对噪声传播特性的影响。

1)在旋翼噪声预测及噪声传播之间,结合点声源假设和旋翼噪声辐射的方向性特点建立了二级声辐射球,从而将随桨叶旋转的多个运动声源转换为声辐射球面上单一静态辐射点,显著提高了计算效率。当地面观测点数量为 10 万时,传统计算方法耗时约为 3.2 h,本文所建方法仅为 305 s。

2)为实时预测直升机运动过程中旋翼噪声对地面声场影响,文中提出了“特征参数声辐射球库”。在地面声场预测中,可根据旋翼所处状态的特征参数,直接选取相应的二级声辐射球进行地面噪声辐射计算。

3)大气、地面吸声效应对旋翼噪声具有较强的衰减作用,但几乎不会改变旋翼噪声传播的方向性;地面流阻对地面吸声效应具有较大的影响。

4)对于距地面具有一定的高度的观测点,其接收到的噪声有直达声波与反射声波;地面反射对观测点时、频域噪声均具有较大的影响,且会发生“声共振”现象,此时低频噪声将会增大,对地面人员产生较大的影响。

5)应当指出,本文所建快速地面声场预测方法对二级声辐射球半径具有一定的要求,当噪声传播距离较近时,应采用直接计算法进行地面声场预测。

参考文献(References)

[1] BRENTNER K S, LYRINTZIS A S, KOUTSAVDIS E K. Comparison of computational aeroacoustic prediction methods for transonic rotor noise prediction[J]. Journal of Aircraft, 1997, 34(4):531-538.

[2] BOYD D D JR. HART-Ⅱ acoustic predictions using a coupled CFD/CSD method[C]// Proceedings of 65th Annual Forum of the AHS, 2009.

[3] DESOPPER A, LAFON P, PHILIPPE J J, et al. Effect of an anhedral sweptback tip on the performance of a helicopter rotor [J]. Vertica, 1988, 12(4):345-355.

[4] YU Y H. Rotor blade-vortex interaction noise[J]. Progress in Aerospace Sciences, 2000, 36(2):97-115.

[5] EDWARDS B, COX C, BOOTH E R JR. Revolutionary concepts for helicopter noise reduction-S. I. L. E. N. T program: NASA-CR-2002-211650[R]. Washington, D. C. BOOTH E R J R.: NASA, 2002.

[6] BOXWELL D A, SCHMITZ F H, SPLETTSTOESSER W R, et al. Helicopter model rotor-blade vortex interaction impulsive noise: Scability and parameteric variations[J]. Journal of the AmericanHelicopter Society, 1987, 32(1):3-12.

[7] SAKOSWKY P C, CHARLES B D. Noise measurement test results of AH-1G operational loads survey: Bell Helicopter Co. Report 299-099-831[R]. Fort Worth, TX: Bell Helicopter Co. 1976.

[8] GOPALAN G, SCHMITZ F H, SIM B W. Flight path management and control methodology to reduce helicopter blade-vortex (BVI) noise[J]. Journal of Aircraft, 2002, 39(2):193-205.

[9] FLEMING G G, RICKLEY E J. HNM heliport noise model, version 2.2, user's guide: Dot/FAA/EE/94-01, DOT-VNTSC-FAA-94-3[R]. Washington, D. C.: Office of Environment and Energy, 1994.

[10] CONNER D A, PAGE J A. A tool for low noise procedures design and community noise impact assessment: The rotorcraft noise model(RNM)[C]// Heli Japan 2002, 2002: T213-6.

[11] CONNER D A, BURLEY C L. Flight acoustic testing and data acquisition for the rotor noise model(RNM)[C]// Proceedings of 62nd Annual Forum of AHS, 2006.

[12] 吕维梁,招启军,徐国华. 计入畸变修正的旋翼尾迹前飞状态稳定性分析[J]. 航空学报,2012,33(11):1958-1966.
LU W L, ZHAO Q J, XU G H. Analysis of rotor wake s stability in forward flight associated with distortion revise[J]. Acta Aeronautica et Astronautica Sinica, 2012, 33(11):1958-1966 (in Chinese).

[13] LEISHMAN J G. Principles of helicopter aerodynamics[M]. 2nd ed. Cambridge: Cambridge University Press, 2006: 580-605.

[14] WEISSINGER J. The lift distribution of swept-back wings: NASA-TM-1120[R]. Washington, D. C.: NASA, 1947.

[15] FARASSAT F, SUCCI G P. The prediction of helicopter rotor discrete frequency noise[J]. Vertica, 1983, 7(4):309-320.

[16] DOWLING A P, FFOWCS WILLIAMS J E, WRIGHT W M. Sound and sources of sound[J]. American Journal of Physics, 1985, 53(6):320.

[17] SIM B W, SCHMITZ F H. Acoustic phasing and amplification effects of single-rotor helicopter blade-vortex interactions[C]// Proceedings of 55th Annua Forum of AHS, 1998.

[18] WEIR D, JUMPER S, BURLEY C L. Aircraft noise prediction program theoretical manual: Rotorcraft system noise prediction system: NASA-TM-83199[R]. Washington, D. C.: NASA, 1995.

[19] SNYDER J P. Map projections: A working manual: USGS Paper 1395[R]. Washington, D. C.: U. S. Government Printing Office, 1987.

[20] 马大猷,沈豫. 声学手册[M]. 北京:科学出版社,2001:31-38.
MA D Y, SHEN H. Manual of acoustics[M]. Beijing: Science Press, 2001:31-38(in Chinese).

[21] ATTENBOROUGH K. Acoustical impendance models for outdoor ground surfaces[J]. Journal of Sound and Vibration, 1985, 99(4):521-544.

作者简介:

王菲 男,博士研究生。主要研究方向:直升机旋翼计算流体力学和直升机旋翼气动噪声。

Tel.: 025-84892117

北航学报 赠阅

E-mail: wangfeinan@ nuaa. edu. cn

史勇杰　男,博士,副教授。主要研究方向:直升机旋翼 CFD 和直升机旋翼气动噪声。
Tel.: 025-84892117
E-mail: shiyongjie@ nuaa. edu. cn

徐国华　男,博士,教授,博士生导师。主要研究方向:直升机空气动力学、旋翼 CFD 和气动声学。
Tel.: 025-84892117
E-mail: ghxu@ nuaa. edu. cn

Efficient ground noise prediction method based on secondary noise radiation model

WANG Fei, SHI Yongjie*, XU Guohua

(National Key Laboratory of Science and Technology on Rotorcraft Aeromechanics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: An efficient ground noise prediction method for the analysis of helicopter noise in flight based on the secondary noise radiation is developed. The method to predict the aerodynamic force and aeroacoustic characteristics based on the free-wake/FW-H, the noise propagation model that can take the effects of atmosphere absorption, ground absorption and reflection on noise radiation characteristics into account, as well as ground noise computation model are included in this method. In order to improve computational efficiency, the secondary sound radiation model built on the hypothesis of compact source is introduced between the noise source computation and noise radiation. To achieve the real time simulation of ground noise, the noise radiation model database is built according to the idea of parametric modeling. The efficiency and effectiveness are illustrated through the comparison of computed results and compute time with the direct computing method for AH-1 rotor. Furthermore, the effects of atmosphere and ground on noise radiation as well as ground noise characteristics are analyzed with this method.

Key words: noise; rotor; secondary noise radiation; atmosphere absorption; ground absorption and reflection

Received: 2015-11-17; **Accepted**: 2016-02-29; **Published online**: 2016-04-13 15:23
URL: www. cnki. net/kcms/detail/11. 2625. V. 20160413. 1523. 003. html
Foundation item: National Natural Science Foundation of China (11302103)
* **Corresponding author**. Tel.: 025-84892117　E-mail: shiyongjie@ nuaa. edu. cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0776

俯仰操纵方式对自转旋翼机操稳特性的影响

林清[1,2], 蔡志浩[1,2,*], 闫坤[1,2], 王英勋[1,2]
(1. 北京航空航天大学 自动化科学与电气工程学院, 北京 100083;
2. 北京航空航天大学 飞行器控制一体化国防重点实验室, 北京 100083)

摘 要: 为了研究不同的俯仰操纵方式对于旋翼机飞行动力学特性的影响,首先基于解析形式叶素法给出自转旋翼的建模方法,并建立了对象无人旋翼机的数学模型;然后分析了2种操纵方式在配平、稳定性及操纵性等方面的差异。研究表明,2种操纵方式各有优缺点:旋翼操纵方式的配平俯仰姿态变化更小且长周期稳定性更好,但螺旋模态不稳定;升降舵操纵方式的螺旋稳定性更好,且俯仰可达力矩较大,但高速配平迎角为负且存在速度静不稳定的问题。针对2种操纵方式,分别设计集成了2架样例无人旋翼机并进行了飞行试验。基于试验数据分析了无人旋翼机飞行过程中自转旋翼的转速变化特性;分别对2架无人旋翼机进行了姿态控制律设计与试验,较好地实现了姿态跟踪控制,并基于试飞数据验证了无人旋翼机数学模型。

关 键 词: 自转旋翼机; 操纵方式; 叶素法; 姿态控制; 模型验证
中图分类号: V249.12
文献标识码: A **文章编号**: 1001-5965(2016)11-2454-12

自转旋翼机(简称旋翼机)是以自转旋翼作为主要升力面和操纵面,发动机牵引/推动作为前进动力的旋翼类飞行器。旋翼机早于直升机出现,其在旋翼原理与技术等方面的突破为后来直升机的诞生与发展奠定了基础[1]。

旋翼机具有独特的技术优势:与飞机相比,旋翼机可以短距起飞,采用预旋后可以实现超短距起飞或跳飞;降落时,通过旋翼后倒可进行有效的减速,实现点式着陆,且低速飞行性能好,不易失速;与直升机相比,旋翼机结构简单,且气动效率更高[2]。近年来,旋翼机在商业和军事上的应用潜力逐渐受到重视,再次成为航空领域关注的热点,随着无人机(Unmanned Aerial Vehicle, UAV)的快速发展与大量应用,无人自转旋翼机也成为其重要的发展方向。

英国 Glasgow 大学在深化旋翼机认识和促进旋翼机发展上作出了巨大的贡献。Houston 等[3-5]以轻型有人旋翼机为研究对象,先后开展了旋翼机空气动力学[3]和飞行动力学[4]、适航性[5]及飞行品质[6]等方面的研究。南京航空航天大学的李建波研究员先后在自转旋翼空气动力学[2,7]、大中型旋翼机总体设计[8]、旋翼机飞行动力学及飞行性能[9]、复合式旋翼机飞行动力学[10-11]等方面进行了深入的研究。此外,陈淼[12]对无人自转旋翼机的飞行控制问题进行了有益的探索。

早期的旋翼机是在固定翼飞机的基础上加装自转旋翼改装而来,为了提高低速时的控制效率,引入了旋翼轴倒角控制方式[1,13],并在随后的发展过程中取消了原有的机翼、副翼、升降舵甚至平尾,形成了目前主流旋翼机的布局形式与操纵方式,即“旋翼 + 方向舵”的操纵方式。1934 年 Hafner 将桨叶变距控制系统引入旋翼机[1],这样,

收稿日期: 2015-11-25; **录用日期**: 2016-01-27; **网络出版时间**: 2016-03-15 14:47
网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20160315.1447.009.html
* **通讯作者**: Tel.: 010-82338792 E-mail: czh@buaa.edu.cn

引用格式: 林清, 蔡志浩, 闫坤, 等. 俯仰操纵方式对自转旋翼机操稳特性的影响[J]. 北京航空航天大学学报, 2016, 42(11): 2454-2465. LIN Q, CAI Z H, YAN K, et al. Influence of pitch manipulation modes on controllability and stability of autogyro[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(11): 2454-2465 (in Chinese).

北航学报 赠阅

旋翼的操纵方式可分为 2 种:一种是直扳式,即直接控制旋翼轴倒角,另一种是通过周期变距实现桨盘的控制。2 种方式各具特点[14]:直扳式结构简单,且成本低,所以在小型旋翼机中广泛使用,但控制精度较差,且随着起飞重量增加,所需操纵力矩会变得非常大,目前为大部分轻型旋翼机所采用;周期变距控制方式所需控制力矩较小,操纵载荷小,但机构较复杂,目前只有较大型旋翼机采用,例如美国 GAC 公司的 Hawk4 旋翼机、南京航空航天大学的 ZX1 旋翼机。从飞行动力学研究的角度,2 种操纵方式可以认为是等效的,故本文不区分 2 种方式,统称为"旋翼周期变距"。

以上对于旋翼机的研究都是针对"旋翼 + 方向舵"布局开展。然而在旋翼机发展初期,局限于当时旋翼空气动力学、飞行力学等学科的发展水平,机翼、副翼及升降舵等的取消并没有进行深入的考量。近年来,学术界和工业界基于现代直升机理论、飞行动力学理论等,重新审视机翼、升降舵等对于旋翼机的影响。Matthew 等[15-16]通过风洞试验研究了升降舵对于自转旋翼机俯仰控制的辅助作用,研究表明升降舵有助于旋翼机从低空速高桨盘迎角状态解出并避免掉高,可以提高低速飞行时旋翼机的安全性。Lopez 和 Wells[17]提出了一种自转旋翼/机翼转换式无人机方案,该无人机使用了自转旋翼与升降副翼来实现俯仰和滚转控制,研究表明自转旋翼操纵模式时,该无人机的纵向周期变距俯仰控制效率较低,需要使用升降副翼来辅助控制。王俊超和李建波[10]给出了一种复合式自转旋翼机方案,并采用"旋翼 + 方向舵"的操纵方式进行了配平和操稳特性分析,但主要研究了机翼对旋翼拉力的卸载作用,配平和操纵性分析中都未使用升降舵和副翼。

为了研究旋翼与升降舵 2 种俯仰操纵方式对自转旋翼机操稳特性的影响,本文首先基于解析叶素法建立了自转旋翼模型,并通过与风洞数据及数值叶素法计算结果的对比说明了其合理性;然后建立了升降舵增强后的对象旋翼机数学模型,对比研究了 2 种操纵方式在配平特性、模态特性、稳定性和操纵性上的差异;最后针对 2 种操纵方式,分别设计和集成了样例无人旋翼机及相应的控制与测量系统,通过飞行试验,分析了全飞行过程中旋翼机的转速变化过程,对设计的比例积分微分(Proportion Integration Differentiation,PID)姿态控制律进行了试验,并基于试飞数据验证了旋翼机的数学模型。

1　无人自转旋翼机

本文选择德国 AUTOGYRO 公司设计制造的 Calidus 有人旋翼机作为全尺寸样机,设计制造了 9/42(旋翼直径比)缩比的基础布局无人旋翼机,三视图如图 1 所示,主要参数如表 1 所示。

在基础布局无人旋翼机的基础上,我们通过增加升降舵,给出了本文建模与对比研究的对象无人旋翼机,如图 2 所示。表 2 给出了两种操纵方式控制通道的分配情况。

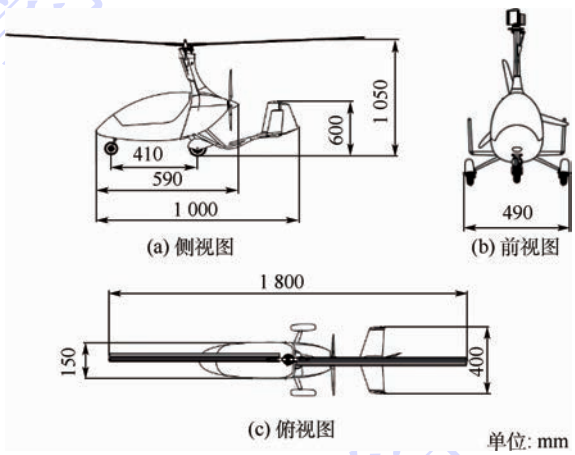


图 1　基础布局无人旋翼机三视图

Fig. 1　Three views of basic autogyro UAV

表 1　基础布局无人旋翼机主要参数

Table 1　Key parameters of basic autogyro UAV

参数	数值
旋翼半径/m	0.85
旋翼实度	0.049
旋翼弦长/m	0.07
桨盘初始后倒角/(°)	7.8
旋翼翼型	HD45
旋翼翼型升力线斜率	5.443
旋翼翼型平均阻力系数	0.012
速度范围/(m·s^{-1})	10 ~ 25

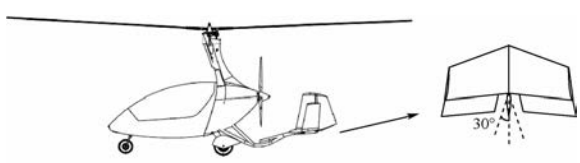


图 2　对比研究使用的升降舵增强无人旋翼机

Fig. 2　Autogyro UAV augmented with elevator used for comparative study

北航学报 赠阅

表 2 2 种操纵方式的控制通道分配情况

Table 2 Control channels allocation of two manipulation modes

操纵方式	俯仰操纵量	滚转操纵量	航向操纵量
A	纵向周期变距	横向周期变距	方向舵
B	升降舵	横向周期变距	方向舵

2 旋翼机建模

2.1 基于解析形式叶素法的自转旋翼建模

对于自转旋翼机而言,除了预旋时旋翼处于驱转状态,全飞行过程旋翼始终处于自转状态,与直升机自转飞行及自转着陆相似,旋翼转速为飞行中的关键参数之一。因此,建立合理的自转旋翼模型是开展自转旋翼机飞行动力学与飞行控制研究的基础。

本文根据文献[13,17]建立自转旋翼模型,为了提高解算速度,便于模型动力学特性分析与后续的控制律设计,采用了解析形式叶素法来建立自转旋翼模型。该方法在低前进比情况下具有较高的计算精度,同时避免了数值叶素法[18]求解非线性多重积分带来的高复杂度与计算缓慢的问题;相对于系统辨识方法[4,19]只能确定某些操作点或小范围包线内有效的线性模型,该方法可建立较大包线内的非线性自转旋翼模型。

首先,根据动量理论迭代计算诱导速度。

$$\begin{cases}\mu=\dfrac{V_k}{\Omega R}\\ \lambda=\mu\alpha_{\mathrm{TPP}}-\dfrac{v_{\mathrm{i}}}{\Omega R}\\ C_T=\dfrac{\sigma a}{4(1+1.5\mu^2)}\left[\left(\dfrac{2}{3}-\dfrac{2}{3}\mu^2+\dfrac{3}{2}\mu^4\right)\theta_0+\left(1-\dfrac{\mu^2}{2}\right)\lambda\right]\\ v_{\mathrm{i}}=\dfrac{C_T\Omega R}{2\mu}\end{cases}\tag{1}$$

式中:μ 为旋翼前进比;V_k 为空速;Ω 为旋翼转速;R 为旋翼半径;λ 为旋翼流入比;α_{TPP} 为桨盘气流迎角;v_{i} 为诱导速度;C_T 为拉力系数;σ 为桨盘实度;a 为旋翼翼型升力线斜率;θ_0 为旋翼总距。

得到稳定的诱导速度 v_{i} 后,根据式(2)计算自转旋翼的拉力 T_{M}、旋翼后向力 H_{M} 和反扭矩 Q_{M}。

$$\begin{cases}T_{\mathrm{M}}=C_T\rho\pi R^2(\Omega R)^2\\ H_{\mathrm{M}}=C_H\rho\pi R^2(\Omega R)^2\\ Q_{\mathrm{M}}=C_Q\rho\pi R^3(\Omega R)^2\end{cases}\tag{2}$$

式中:反扭矩系数 C_Q 和侧力系数 C_H 定义为

$$C_Q=\frac{\sigma c_d(1+\mu^2)}{8}-\sigma\frac{a\lambda}{4(1+1.5\mu^2)}\left[\frac{\theta_0(2-\mu^2)}{3}+\left(1+\frac{\mu^2}{2}\right)\lambda\right]\tag{3}$$

$$C_H=\frac{\sigma}{\mu}\left[\frac{c_d}{8}(1+3\mu^2)-\lambda\frac{C_T}{\sigma}-\frac{C_Q}{\sigma}\right]\tag{4}$$

式中:c_d 为旋翼翼型平均阻力系数。

旋翼的挥舞运动包含了吹风挥舞、操纵挥舞和随动挥舞,可由式(5)和式(6)计算得到。

$$a_{1s}=\underbrace{\left(\frac{\mu}{1-0.5\mu^2}\right)\left[\frac{8\theta_0}{3}+2\left(\mu\alpha_{\mathrm{TPP}}-\frac{v_{\mathrm{i}}}{\Omega R}\right)\right]}_{\text{吹风挥舞}}-\underbrace{\left(\frac{1+1.5\mu^2}{1-0.5\mu^2}\right)B_1}_{\text{操纵挥舞}}+\underbrace{\frac{1}{1-0.5\mu^2}\left[\frac{-16(q/\Omega)}{\gamma_b}+\frac{p}{\Omega}\right]}_{\text{随动挥舞}}\tag{5}$$

$$b_{1s}=\underbrace{A_1}_{\text{操纵挥舞}}+\underbrace{\frac{v_{\mathrm{i}}/(\Omega R)}{1+0.5\mu^2}}_{\text{吹风挥舞}}+\underbrace{\frac{1}{1+0.5\mu^2}\left[\frac{-16(p/\Omega)}{\gamma_b}-\frac{q}{\Omega}\right]}_{\text{随动挥舞}}\tag{6}$$

式中:a_{1s} 和 b_{1s} 分别为纵向挥舞角和横向挥舞角;B_1 和 A_1 分别为纵向周期变距和横向周期变距;γ_b 为桨叶洛克数;p 和 q 分别为旋翼机的滚转角速率和俯仰角速率。

根据旋翼机受力分析可得到机体坐标系中旋翼的气动力和气动力矩:

$$\begin{cases}F_x^{\mathrm{rotor}}=-H_{\mathrm{M}}\cos i_{\mathrm{M}}-T_{\mathrm{M}}\sin(a_{1s}+i_{\mathrm{M}})\\ F_y^{\mathrm{rotor}}=-H_{\mathrm{M}}\sin\beta+T_{\mathrm{M}}\sin b_{1s}\\ F_z^{\mathrm{rotor}}=-T_{\mathrm{M}}\cos(a_{1s}+i_{\mathrm{M}})\\ M_x^{\mathrm{rotor}}=F_y^{\mathrm{rotor}}(-z_{\mathrm{M}})+F_z^{\mathrm{rotor}}y_{\mathrm{M}}+Q_{\mathrm{M}}\sin(a_{1s}+i_{\mathrm{M}})\\ M_y^{\mathrm{rotor}}=F_x^{\mathrm{rotor}}z_{\mathrm{M}}+F_z^{\mathrm{rotor}}(-x_{\mathrm{M}})-Q_{\mathrm{M}}\sin b_{1s}\\ M_z^{\mathrm{rotor}}=F_y^{\mathrm{rotor}}x_{\mathrm{M}}+F_x^{\mathrm{rotor}}(-y_{\mathrm{M}})+Q_{\mathrm{M}}\cos(a_{1s}+i_{\mathrm{M}})\cos b_{1s}\end{cases}\tag{7}$$

式中:i_{M} 为旋翼轴初始后倒角;$(x_{\mathrm{M}},y_{\mathrm{M}},z_{\mathrm{M}})$ 为机体坐标系中桨毂中心的坐标;β 为侧滑角;$*^{\mathrm{rotor}}$ 为自转旋翼气动力/力矩在体轴系的 3 轴分量。

2.2 自转旋翼建模方法验证

为了验证本文中自转旋翼建模方法的有效性,将计算结果与文献[8]中的数值叶素法及风洞试验结果进行了对比。对比自转旋翼半径为 1.25 m,桨叶片数为 2、弦长为 0.095 m、翼型为 OA212,单片桨叶质量为 0.772 kg。

试验分为 2 组:第 1 组试验中自转旋翼后倒角为 4°,旋翼总距分别取 0°、2°、4°,空速范围为 15~30 m/s;第 2 组试验中自转旋翼总距固定为

4°，自转旋翼后倒角分别取 2°、4°、6°，空速范围为 20 ~ 30 m/s。

图 3 和图 4 分别对比了 2 种方法计算结果与风洞试验数据在自转旋翼稳定转速、升力（L）与风速（V）关系上的差异。

关于对比试验中自转旋翼稳定转速的计算，数值叶素法按照式（10）根据反扭矩和桨叶转动惯量进行迭代计算。由于此处关注的是给定空速、自转旋翼后倒角和旋翼总距下的自转旋翼稳定转速、气动力和力矩计算结果而不是其动态调整过程，所以解析叶素法计算时采用了与式（10）不同的迭代计算方法：根据计算所得的反扭力矩迭代调整转速，且转速调整步长依迭代次数而定，而不采用实际的转动加速度，最终可以实现反扭矩的快速收敛与转速的快速稳定。

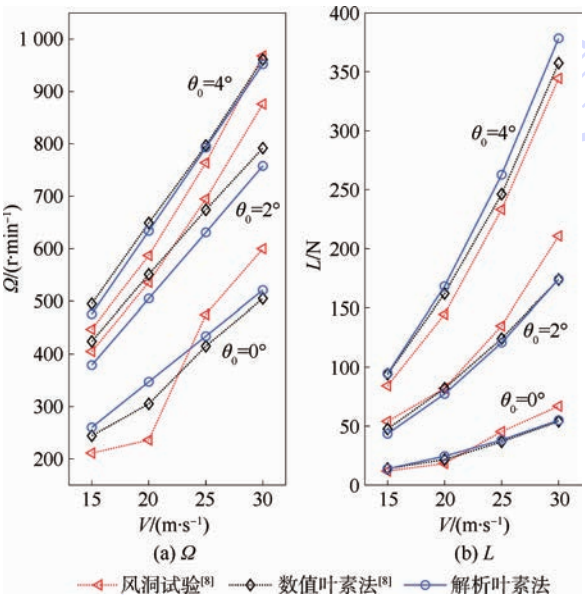


图 3　对比试验 1 中自转旋翼稳定转速、升力与风速关系

Fig. 3　Stable speed and lift of autorotating rotor versus wind speed in comparison test 1

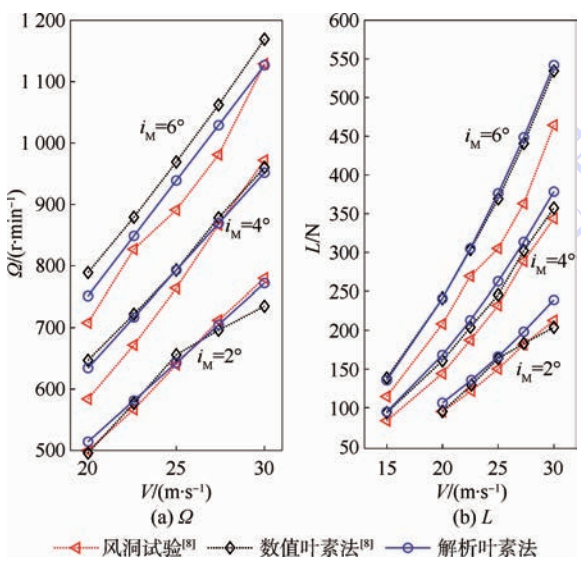


图 4　对比试验 2 中自转旋翼稳定转速、升力与风速关系

Fig. 4　Stable speed and lift of autorotating rotor versus wind speed in comparison test 2

从图 3 和图 4 中可以看出，本文方法计算结果与文献[8]中的风洞试验结果及基于数值叶素法的结果吻合程度较高，模型的精度能够满足小型无人旋翼机动力学分析的需求。

2.3　旋翼机数学模型

旋翼机区别于直升机，其转速在较大范围内变化，并与其他模态相互耦合，所以在建模时需要增加旋翼转速自由度，这样就构成了旋翼机的 7 自由度动力学模型，由式（8）~ 式（10）给出。

$$\begin{cases}\dot{u}=vr-wq+F_x/m\\ \dot{v}=-ur-wq+F_y/m\\ \dot{w}=uq-vp+F_z/m\end{cases}\tag{8}$$

$$\begin{cases}\dot{p}=(c_1r+c_2p)q+c_3\bar{L}+c_4N\\ \dot{q}=c_5pr-c_6(p^2-r^2)+c_7M\\ \dot{r}=(c_8p-c_2r)q+c_4\bar{L}+c_9N\end{cases}\tag{9}$$

$$\dot{\Omega}=-Q_M/J_M\tag{10}$$

旋翼机的运动学方程由式（11）与式（12）给出

$$\begin{cases}\dot{x}=u\cos\theta\cos\psi+v(\sin\theta\cos\psi\sin\phi-\sin\psi\cos\phi)+\\ \qquad w(\sin\theta\cos\psi\cos\phi+\sin\psi\sin\phi)\\ \dot{y}=u\cos\theta\sin\psi+v(\sin\theta\sin\psi\sin\phi+\cos\psi\cos\phi)+\\ \qquad w(\sin\theta\sin\psi\cos\phi-\cos\psi\sin\phi)\\ \dot{z}=-u\sin\theta+v\cos\theta\sin\phi+w\cos\theta\cos\phi\end{cases}\tag{11}$$

$$\begin{cases}\dot{\phi}=p+(r\cos\phi+q\sin\phi)\tan\theta\\ \dot{\theta}=q\cos\phi-r\sin\phi\\ \dot{\psi}=\dfrac{r\cos\phi+q\sin\phi}{\cos\theta}\end{cases}\tag{12}$$

式中：m 为飞机质量；x、y 和 z 分别为地轴系 3 轴位置；u、v 和 w 分别为体轴系 3 轴地速；ϕ、θ 和 ψ 分别为滚转角、俯仰角和偏航角；r 为偏航角速率；常数 c_i 的定义可参考文献[20]；J_M 为旋翼相对于桨毂中心的转动惯量；$\bar{L}$、M 和 N 分别为旋翼机的滚转力矩、俯仰力矩和偏航力矩；F_x、F_y 和 F_z 为机体坐标系下旋翼机的 3 轴力分量，具体组成如下：

$$\begin{cases}\bar{L}=M_x^{\text{rotor}}+M_x^{\text{plane}}+M_x^{\text{P}}\\ M=M_y^{\text{rotor}}+M_y^{\text{plane}}+M_y^{\text{P}}\\ N=M_z^{\text{rotor}}+M_z^{\text{plane}}+M_z^{\text{P}}\end{cases}\tag{13}$$

$$\begin{cases}F_x=F_x^{\text{rotor}}+F_x^{\text{plane}}+F_x^{\text{P}}+G_x\\ F_y=F_y^{\text{rotor}}+F_y^{\text{plane}}+F_y^{\text{P}}+G_y\\ F_z=F_z^{\text{rotor}}+F_z^{\text{plane}}+F_y^{\text{P}}+G_z\end{cases}\tag{14}$$

北航学报 赠阅

式中：$*^{plane}$ 为包括机身、平尾、垂尾、方向舵和升降舵在内的固定翼气动部件产生的气动力/力矩在机体系下的 3 轴分量；$*^{P}$ 为推力/力矩在机体轴上的分量；G_i 为重力在机体系下的 3 轴分量。

3 对比分析

3.1 配平特性对比

常用的配平方法有周期打靶法、Gauss-Seidel 迭代法、Newton 法[8] 及飞行中配平法(fly-to-trim)[21] 等，不同的方法各有优劣。为了更好地分析旋翼机自转旋翼转速、挥舞角等关键物理量在飞行中的变化过程，本文选择飞行中配平法进行旋翼机的配平，通过设计配平控制律，动态调节操纵量和姿态角，在飞行过程中实现力与力矩的平衡。本文中旋翼机的配平状态为定速、定高、无侧偏及无侧滑飞行。

图 5 给出了 2 种操纵方式下旋翼机主要状态量的配平结果，α 为迎角。可以看到 2 种操纵方式下旋翼机的横向配平状态基本一致，主要差异体现在配平俯仰角与配平迎角。从图 5(a)可以看到，不同操纵方式下配平迎角和俯仰角都随着飞行速度的提高而减小，但由于操纵方式 A 可以直接控制桨盘的气流迎角，所以其配平迎角和俯仰角虽然持续减小但一直保持为正，且全速度范围内变化很小；而操纵方式 B 只能通过控制整机俯仰角来间接控制桨盘气流迎角，所以其配平俯仰角和迎角随着空速的提高持续减小，且当速度高于 15 m/s 后甚至出现负的迎角和俯仰角。

旋翼机相对于直升机的重要优势在于旋翼机的飞行姿态非常平稳，主要原因在于直升机需通过较大的桨盘前倾获得前进的动力，而旋翼机的主旋翼只需产生升力，前进动力通过后推螺旋桨提供。操纵方式 A 充分体现了旋翼机的优势，而操纵方式 B 则存在俯仰姿态变化较大的问题，且高速飞行时低头比较严重，一方面增大了阻力和负升力，气动效率较低，同时也会影响到对俯仰姿态要求较高的任务设备的使用。

图 6 和图 7 分别给出了速度包线内纵横向配平控制量的对比，δ_e 为操纵方式 B 的升降舵控制量，δ_p 为油门开度，δ_r 为方向舵控制量。从图 6(a)和图 6(b)中可以看到两种操纵方式都是从低速时的抬头操纵到高速时的低头操纵：对于操纵方式 A 而言，旋翼纵向周期变距从后倒 0.5°到前倒3.9°；对于操纵方式 A 而言，升降舵从上偏 1.3°

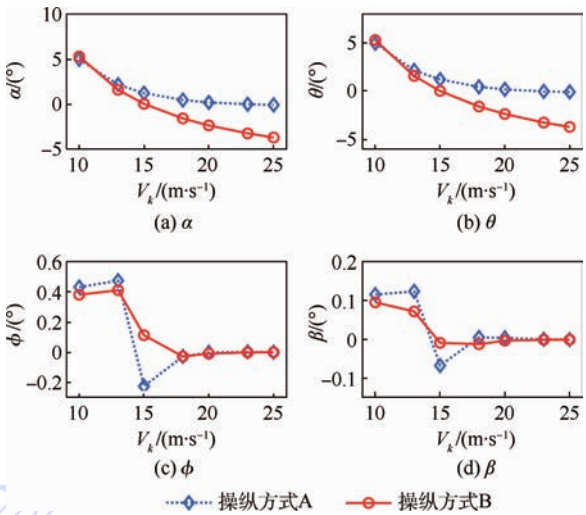


图 5 速度包线内主要状态配平对比

Fig. 5 Comparison of main trimmed states in speed envelope

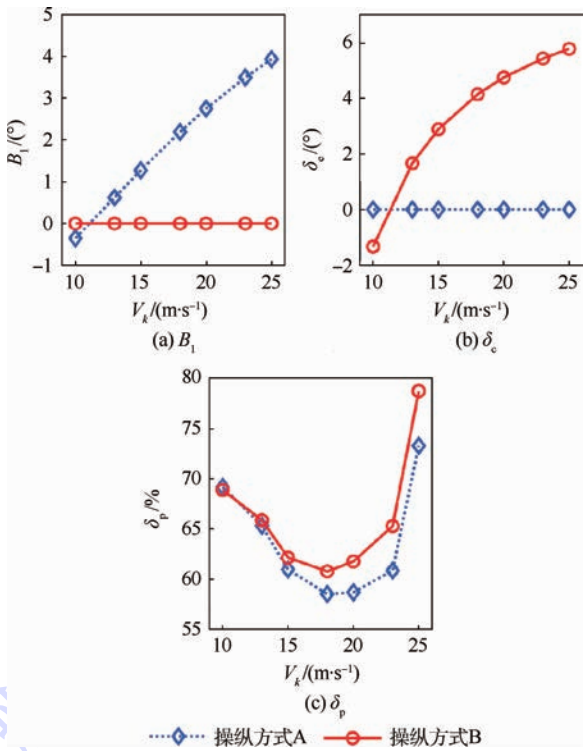


图 6 纵向控制量配平对比

Fig. 6 Comparison of longitudinal trimmed control inputs

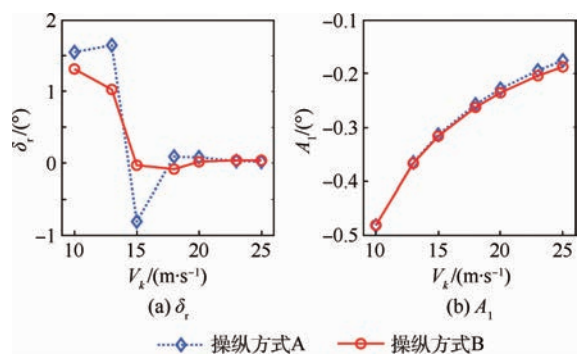


图 7 横向控制量配平对比

Fig. 7 Comparison of lateral trimmed control inputs

北航学报 赠阅

到下偏 5.7°。主要原因是随着空速的提高，桨盘的空气流入量增加，为了保持平飞，需要减小桨盘迎角。操纵方式 A 可以直接控制桨盘后倒角，而操纵方式 B 只能通过控制俯仰角的方式来减小桨盘迎角。从图 6(c)中可以看到，在低速飞行时由于要使用推力的垂直分量来平衡部分重力，所以低速时油门开度较大，随着空速的提高油门开度先减小，而后随着速度的提高，废阻变大，油门开度变大。由于操纵方式 B 的飞行姿态多为负迎角，废阻较操纵方式 A 更大，所以其油门开度更大一些。

从图 7 中可以看到，2 种操纵方式下旋翼机的横向配平控制量基本一致，方向舵低速时变化比较大，高速时变化幅度很小；随着空速的提高，旋翼的横向控制量(左倒)逐渐减小，体现了右旋主旋翼横向吹风挥舞与正桨螺旋桨反扭矩的综合作用。

图 8 给出了旋翼挥舞角及转速的配平结果对比，从图 8(a)中可以看到操纵方式 A 的纵向挥舞角在操纵挥舞与吹风挥舞的叠加作用下，随着速度的提高从后倒变为前倒角；操纵方式 B 由于不直接控制旋翼，所以旋翼纵向挥舞角只包含了吹风挥舞和随动挥舞，随着前飞速度的提高，挥舞角后倒增大。2 种旋翼的横向挥舞角差异很小。就桨盘迎角而言，桨盘迎角随着前飞速度的增大逐渐减小，但即使在俯仰角为负的情况下，由于旋翼轴初始后倒角的存在，桨盘迎角仍为正值，保证了旋翼的正常自转。操纵方式 B 由于高速时主要处于负迎角状态，旋翼需要提供较操纵方式 A 更大的拉力来消除负升力的影响，所以配平旋翼气流迎角更大且配平转速更高，如图 8(c)和图 8(d)所示，α_s 为桨盘气流迎角。

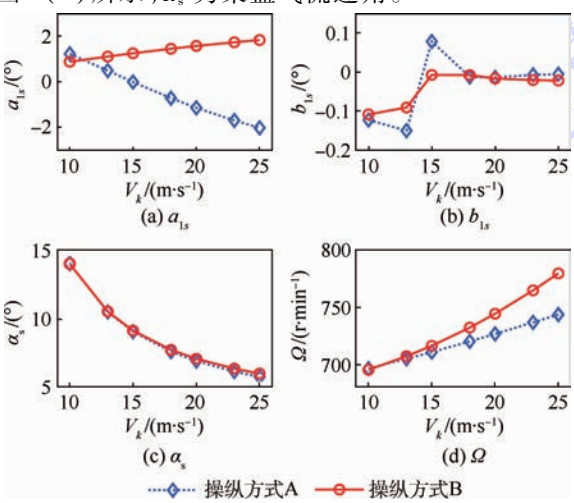


图 8　主要旋翼状态配平对比

Fig. 8　Comparison of trimmed main rotor states

3.2　模态特性对比

在配平点处，针对采用 2 种操纵方式的无人旋翼机非线性模型进行小扰动线性化。根据线性化结果，分析各运动模态的特性[14]，分析结果如图 9～图 12 所示，ξ 为阻尼比，ω_n 为无阻尼自然频率，T_2 为倍幅时间，$T_{1/2}$ 为半幅时间，T 为模态周期。

从图 9 中可以看到，长周期的稳定性受操纵方式影响明显，操纵方式 A 在速度为 10 m/s 时为临界稳定，速度高于 13 m/s 之后，随速度的提高，半幅时间从 99 s 缩短至 22 s，稳定性增强；而操纵方式 B 的长周期模态在 15 m/s 之前都是不稳定。随速度提高，长周期的倍幅时间从 63 s 逐渐增大至 347 s，稳定性逐渐提高，并在 18 m/s 之后变为稳定，此后，随着速度的提高，半幅时间从 257 s 逐渐减小至48 s，稳定性逐渐提高。总的来说，随着

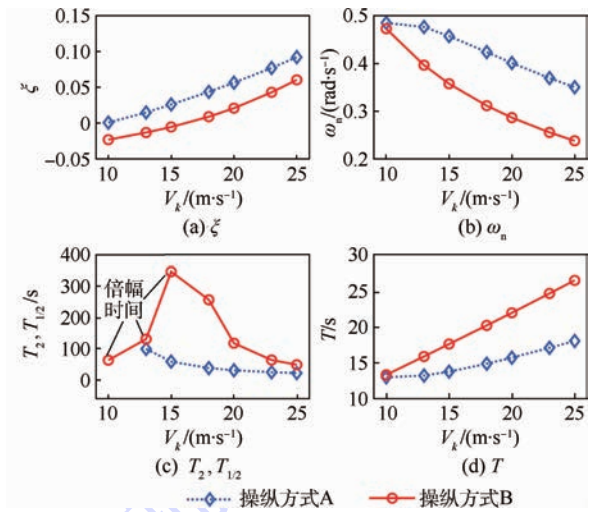


图 9　长周期模态特性对比

Fig. 9　Comparison of phugoid mode characteristics

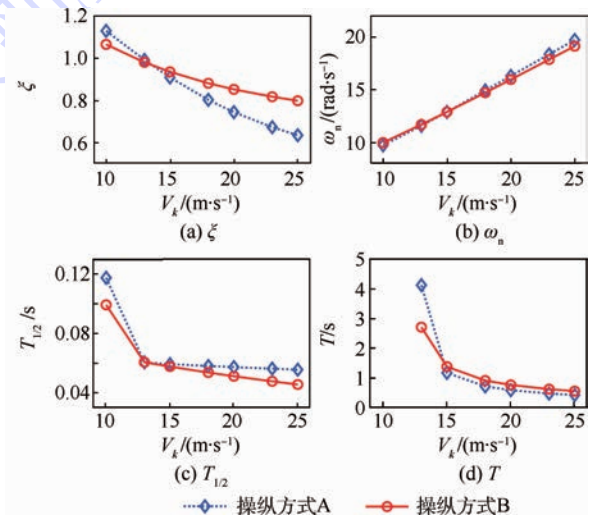


图 10　短周期模态特性对比

Fig. 10　Comparison of short-period mode characteristics

北航学报 赠阅

速度的提高,采用 2 种操纵方式的旋翼机长周期运动模态的稳定性都有所提高,但采用操纵方式 A 的旋翼机长周期稳定性更好。

图 10 和图 11 分别给出了 2 种操纵方式下,旋翼机短周期模态和荷兰滚模态的特征参数,可以看到,2 种操纵方式的短周期和荷兰滚运动特性较为一致:随着速度的提高,短周期阻尼比降低,无阻尼自然频率增大,半幅时间逐渐减小,模态稳定性随速度增大而提高。

图 12 给出了采用 2 种操纵方式时,旋翼机单调运动模态的特征参数,从图 12(a)中可以看到,2 种操纵方式的转速收敛模态半幅时间均小于 1.5 s,并随速度的提高而减小,转速模态的收敛速度加快,且操纵方式 A 下转速的收敛速度高于操纵方式 B。从图 12(b)中可以看到,操纵方式 A 下,旋翼机的螺旋模态一直处于不稳定状态,但随空速的提高,倍幅时间从 5 s 增加至 59 s,发散速度变慢;操纵方式 B 下,旋翼机在速度低于 18 m/s时螺旋模态不稳定,但随速度提高,倍幅时间从 5 s增加至 108 s,并在 18 m/s 后变为稳定。从图 12(c)中可以看到,2 种操纵方式的滚转阻尼模态特性基本一致,半幅时间都小于 0.05 s,能够快速收敛。

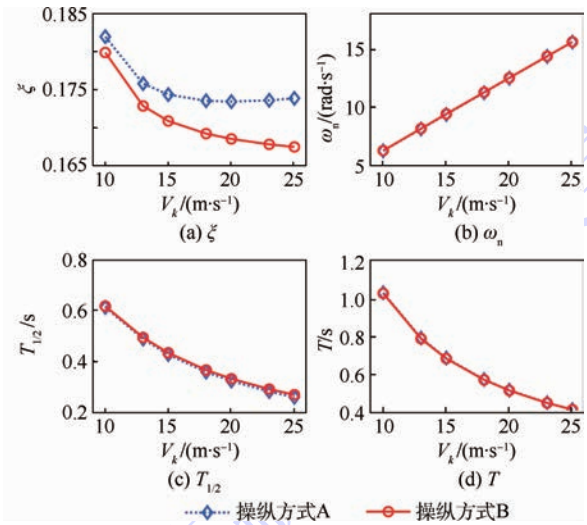


图 11 荷兰滚模态特性对比

Fig. 11 Comparison of Dutch-roll mode characteristics

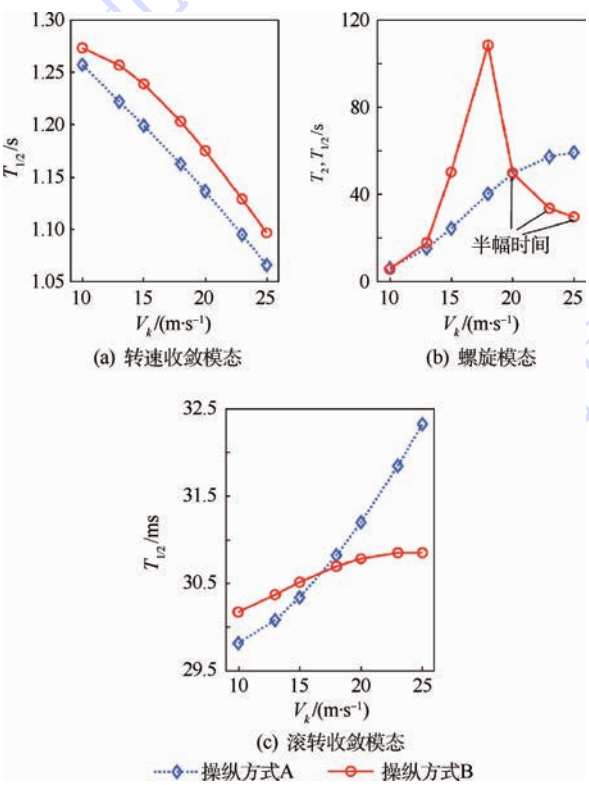


图 12 单调运动模态特性对比

Fig. 12 Comparison of monotone mode characteristics

3.3 稳定性对比

图 13 给出了采用 2 种操纵方式时旋翼机的主要纵向稳定性导数,M_u、M_w 和 M_Ω 分别为空速稳定性导数、迎角静稳定性导数和转速静稳定性导数。

空速稳定性导数 $M_u>0$ 时满足速度静稳定条件,从图 13(a)中可以看到操纵方式 A 下旋翼机的速度静稳定性随空速提高而增大,采用操纵方式 B 时,旋翼机为速度静不稳定,且随空速的提高其速度静不稳定度增大。

图 13(b)给出了 2 种操纵方式下旋翼机的迎角静稳定性导数 M_w 的对比结果,迎角静稳定性

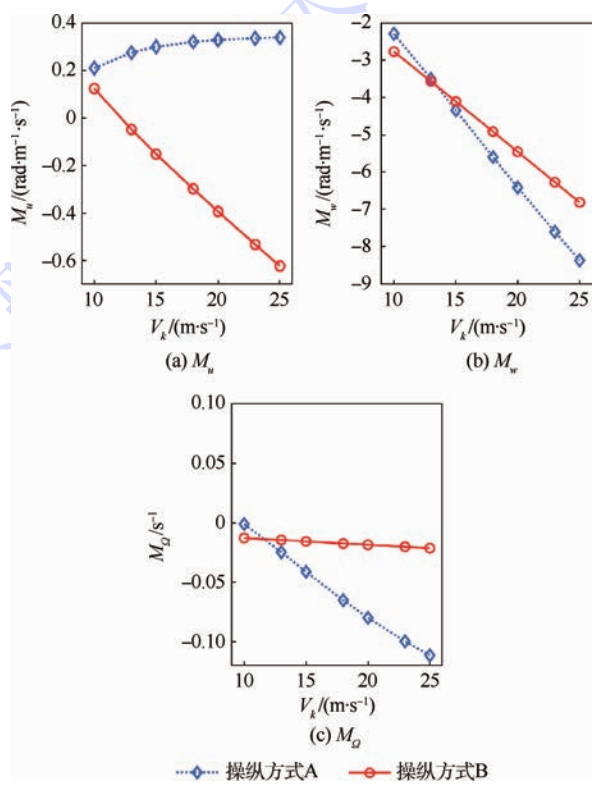


图 13 主要纵向稳定性导数曲线

Fig. 13 Curves of key longitudinal stable moment derivatives

北航学报 赠阅

导数 $M_w<0$ 时满足迎角静稳定条件,从中可以看到 2 种操纵方式都满足迎角静稳定性条件,并随空速提高而增大,且高速时操纵方式 A 的迎角静稳定性高于操纵方式 B。

图 13(c)给出了 2 种操纵方式下旋翼机的转速静稳定性导数 M_Ω 的对比结果,转速静稳定性导数 $M_\Omega<0$ 时满足转速静稳定条件。从中可以看到 2 种操纵方式都满足转速静稳定性条件,并随空速提高而增大,且高速时操纵方式 A 的转速静稳定性高于操纵方式 B。

可见,采用操纵方式 A 时旋翼机的纵向稳定特性要好于采用操纵方式 B。

3.4　操纵性对比

为了对比 2 种操纵方式导致的旋翼机操纵性上的差异,分析了 20 m/s 时 2 种操纵方式的虚拟可达力矩集,各舵面的偏转范围为 $|A_1|\leqslant 15°$、$|B_1|\leqslant 15°$、$|\delta_e|\leqslant 25°$及$|\delta_r|\leqslant 30°$。

图 14 给出了 2 种操纵方式对应的可达力矩集,C_l、C_m 和 C_n 分别为滚转力矩系数、俯仰力矩系数和偏航力矩系数,其中操纵方式 A 的可达力矩集凸包体积为 0.61,操纵方式 B 的可达力矩集凸包体积为 1.18。从图中可以看到操纵方式 B 在俯仰通道的可达力矩稍大于操纵方式 A,这一方面是由于升降舵与重心的距离较远,同时螺旋桨滑流也起到了增强升降舵舵效的作用。

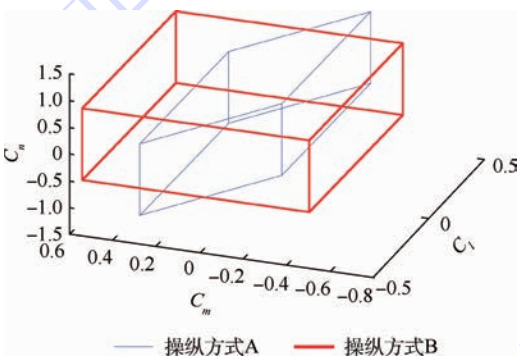


图 14　可达力矩集对比

Fig. 14　Comparison of attainable moment sets

4　试验研究

为了更好地熟悉无人旋翼机的飞行特性,并验证所设计的 PID 姿态控制律,本文分别针对不同的操纵方式设计集成了 2 架样例无人旋翼机并设计了相应的旋翼转速测量装置[22-23]和自动驾驶仪[24]。图 15(a)给出的样例无人旋翼机 A 采用了操纵方式 A,为图 1 中缩比无人旋翼机实物;图 15(b)给出了样例无人旋翼机 B,该无人机采用操纵方式 B,图 15(c)给出了样例无人旋翼机 B 的试飞截图。



图 15　样例无人旋翼机

Fig. 15　Example autogyro UAVs

4.1　无人旋翼机自转特性分析

旋翼转速随前飞速度与桨盘迎角动态变化是旋翼机的关键特性,且转速模态与长短周期模态发生耦合,对于旋翼机稳定性产生重要影响,所以清晰地认识转速变化是分析和控制旋翼机的基础,但现有文献很少给出全飞行过程中旋翼机的自转旋翼转速变化过程。

本文基于飞行试验给出全飞行过程中无人旋翼机自转旋翼的转速变化过程。飞行试验可以分为 7 个阶段,如图 16 所示,各阶段分别为:Ⅰ为检查预旋阶段,Ⅱ为起飞预旋阶段,Ⅲ为滑跑起飞阶段,Ⅳ为巡航阶段,Ⅴ为下滑着陆阶段,Ⅵ为着陆滑跑及停机降转阶段,Ⅶ为手动停转阶段。

检查预旋时(Ⅰ)时先短时预旋旋翼,旋翼转速最高达到 518 r/min;检查旋翼状态正常后进行起飞预旋(Ⅱ),在预旋电机的驱转下,预旋转速最高达到 828 r/min;切断预旋电机后增大推力开始滑跑,进入到滑跑起飞阶段(Ⅲ),滑跑阶段由于旋翼失去驱动力,且旋翼逐渐承载,故转速逐渐降低,离地爬升后随着空速的提高,转速逐渐提高,爬升结束后转速升高至 719 r/min;之后进入巡航阶段(Ⅳ),巡航过程中旋翼转速在 549 ~ 951 r/min 之间波动,平均值约 681 r/min;下滑着陆过程(Ⅴ)中,油门处于怠速状态,转速随空速的降低而降低,着陆时降至约 626 r/min;在着陆滑跑及停机降转阶段(Ⅵ)中,转速降至 49 r/min;最终由操作手手动停转(Ⅶ)。

图 17 给出了样例无人旋翼机 B 在一个架次中转速的变化曲线。虽然由于无人旋翼机 B 续航能力较差,且所采用的转速测量装置采样频率较低(仅1Hz),但从图中仍可以看到其旋翼转

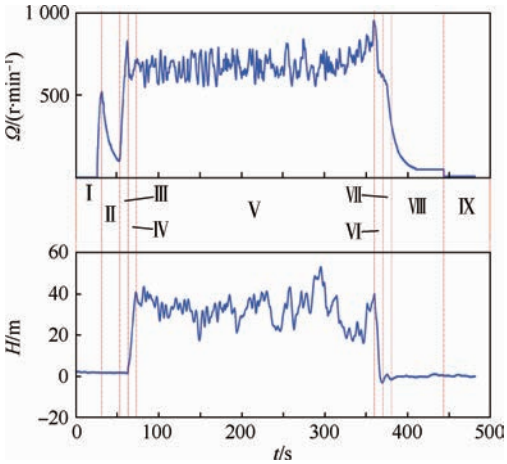


图 16 样例无人旋翼机 A 的转速与高度变化曲线

Fig. 16 Variation curves of rotor speed and height of example autogyro UAV-A

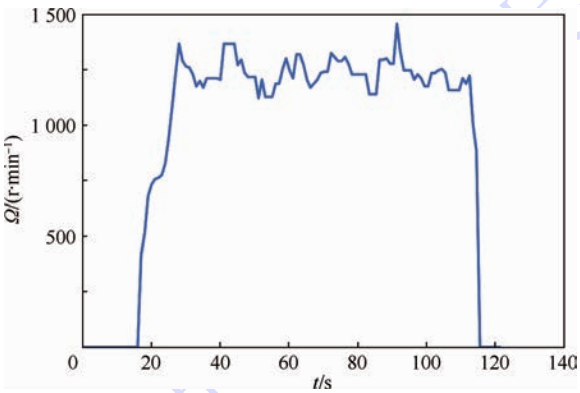


图 17 样例无人旋翼机 B 转速变化

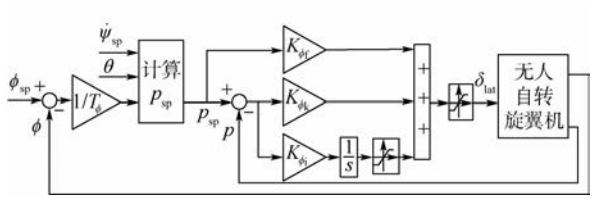
Fig. 17 Rotor speed variation of example autogyro UAV-B

速的变化过程。起飞滑跑前,旋翼先通过一级预旋转速升高到约 780 r/min,然后通过二级预旋升高到约 1 380 r/min,切断预旋后转速快速下降至约1 200 r/min,随着滑跑加速,旋翼逐渐承载,转速持续降低;起飞后,随着空速的提高,旋翼转速提高;巡航飞行过程中,旋翼转速随空速、风速和姿态的变化在 1 100 ~ 1 450 r/min 之间波动。

4.2 基于 PID 的姿态控制器

经典 PID 控制方法面向工程实现,具有非线性控制方法无法比拟的易用性,所以本文采用 PID 控制方法设计了 2 架样例无人旋翼机的姿态控制律并完成了飞行试验,并将其作为后续的旋翼机飞行控制研究的参考。

图 18 以滚转通道为例给出了基于 PID 方法的自转旋翼机姿态控制结构。图中的控制结构与经典姿态控制[20]中的 Propotion Differentiation (PD)或 Propotion Integration(PI)控制结构稍有区别,通过将姿态控制分为姿态环与速率环,可方便地实现姿态跟踪模式、角速率跟踪模式和舵面控



δ_{lat}—横向操纵量。

图 18 滚转通道控制结构

Fig. 18 Control structure of roll channel

制模式等多种控制模式。

各通道的控制律如式(15) ~式(21)所示。

1) 滚转通道控制律为

$$A_1 = K_{\phi f}p_{sp} + K_{\phi k}(p_{sp} - p) + K_{\phi i}\int(p_{sp} - p)\mathrm{d}t \quad (15)$$

式中:$K_{\phi f}$、$K_{\phi k}$和 $K_{\phi i}$分别为滚转通道前馈系数、比例系数和微分系数;期望滚转角速率 p_{sp}的计算公式为

$$p_{sp} = \dot{\phi}_{sp} - \dot{\psi}_{sp}\sin\theta = \frac{\phi_{sp} - \phi}{T_\phi} - \dot{\psi}_{sp}\sin\theta \quad (16)$$

式中:T_ϕ 为滚转通道的期望响应时间常数;$\dot{\phi}_{sp}$为期望滚转角变化率;$\dot{\psi}_{sp}$为期望的航向角变化率,定义为

$$\dot{\psi}_{sp} = \frac{w\dot{\phi}_{sp} + g\sin\phi\cos\theta + u\dot{\theta}_{sp}\sin\phi}{u\cos\phi\cos\theta + w\sin\theta} \quad (17)$$

2) 俯仰通道的控制律为

$$\delta_{lon} = K_{\theta f}q_{sp} + K_{\theta k}(q_{sp} - q) + K_{\theta i}\int(q_{sp} - q)\mathrm{d}t \quad (18)$$

式中:δ_{lon}为俯仰操纵量,对于操纵方式 A 和操纵方式 B 而言分别为纵向周期变距 B_1 和升降舵 δ_e;$K_{\theta f}$、$K_{\theta k}$和 $K_{\theta i}$分别为俯仰通道前馈系数、比例系数和微分系数;期望俯仰角速率 q_{sp}的计算公式为

$$q_{sp} = \dot{\theta}_{sp}\cos\phi + \dot{\psi}_{sp}\cos\theta\sin\phi = \frac{\theta_{sp} - \theta}{T_\theta}\cos\phi + \dot{\psi}_{sp}\cos\theta\sin\phi \quad (19)$$

式中:T_θ 为俯仰通道的期望响应时间常数;$\dot{\theta}_{sp}$为期望俯仰角变化率。

3) 偏航通道控制律为

$$\delta_r = K_{\psi f}r_{sp} + K_{\psi k}(r_{sp} - r) + K_{\psi i}\int(r_{sp} - r)\mathrm{d}t \quad (20)$$

式中:$K_{\psi f}$、$K_{\psi k}$和 $K_{\psi i}$分别为偏航通道前馈系数、比例系数和微分系数;期望偏航角速率 r_{sp}的计算公式为

$$r_{sp} = -\dot{\theta}_{sp}\sin\phi + \dot{\psi}_{sp}\cos\theta\cos\phi = -\frac{\theta_{sp} - \theta}{T_r}\sin\phi + \dot{\psi}_{sp}\cos\theta\cos\phi \quad (21)$$

式中:T_r 为偏航通道的期望响应时间常数。

4.3　基于 PID 控制器的姿态控制试验

基于 4.2 节 PID 控制结构，经过初始设计和飞行调参，实现了样例无人自转旋翼机的姿态跟踪控制，姿态跟踪试验结果如图 19 和图 20 所示。

图 19 为样例无人旋翼机 A 的俯仰角和滚转角跟踪试验结果。可以看到，本文控制律可以较好地实现无人旋翼机的姿态控制，能够快速地、以较小的静差跟踪给定姿态指令，其中滚转角平均跟踪误差为 1.1°，俯仰角平均跟踪误差为 -0.9°。

图 20 为样例无人旋翼机 B 的俯仰角和滚转角跟踪试验结果。可以看出，基于 PID 的姿态控制律同样可以实现样例无人旋翼机 B 的姿态跟踪控制。但由于无人旋翼机 B 的机型太小，速度较低，受风干扰影响较大，所以其姿态控制效果并不理想。

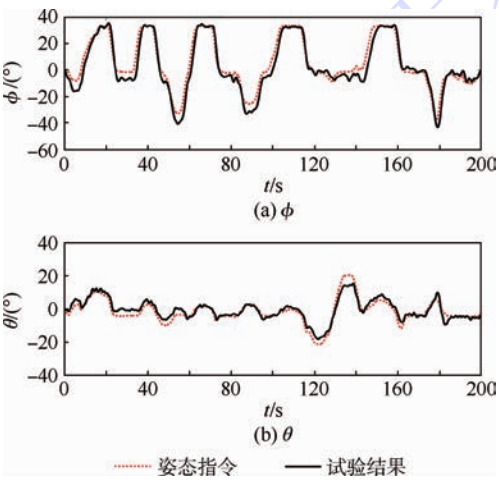


图 19　样例无人旋翼机 A 姿态跟踪试验结果

Fig. 19　Attitude tracking test results of example autogyro UAV-A

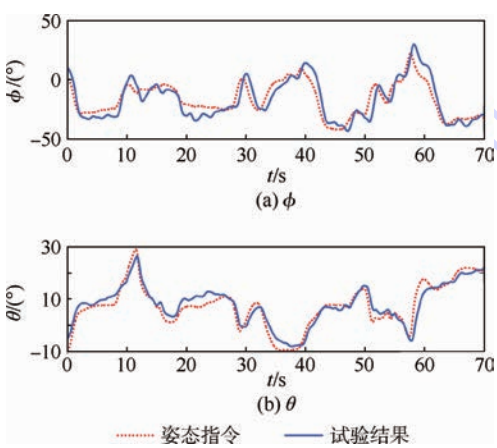


图 20　样例无人旋翼机 B 姿态跟踪试验结果

Fig. 20　Attitude tracking test results of example autogyro UAV-B

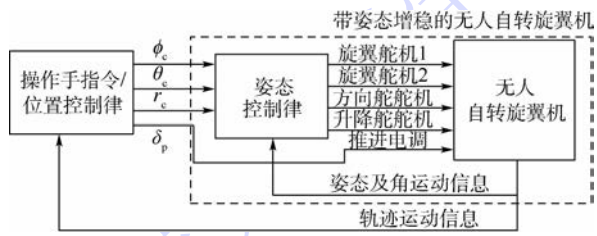
4.4　基于试飞数据的模型验证

文献[17]中针对提出的无人自转旋翼机方案，建立了无人旋翼机的七自由度模型，但囿于试验条件，只通过分析模型对于典型控制输入的响应定性地验证了模型的合理性。基于 4.3 节的飞行试验数据，本文开展了无人旋翼机数学模型的定量验证。

图 21 给出了样例无人旋翼机的飞行控制结构，由于旋翼机的操纵特性与常规的固定翼及直升机差异较大，且小型无人旋翼机更难以操控，为了提高安全性与试验效率，本文在手控飞行与自动飞行时都基于 4.3 节中设计的姿态增稳控制律进行。

基于图 21 给出的控制结构，本文的模型验证思路如下：将姿态控制律与无人旋翼机作为一个整体即带姿态增稳的无人自转旋翼机，通过采集飞行时的姿态给定指令与姿态跟踪响应，对比采用同样的姿态控制律时，样例无人旋翼机与无人旋翼机数学模型对于同样的姿态给定指令的跟踪效果。

对比结果如图 22 所示，整体看来，带姿态增稳的无人自转旋翼机数学模型与无人机实际跟踪



ϕ_c—滚转角指令；θ_c—俯仰角指令；r_c—偏航角速率指令。

图 21　样例无人旋翼机飞行控制结构

Fig. 21　Flight control structure of example autogyro UAV

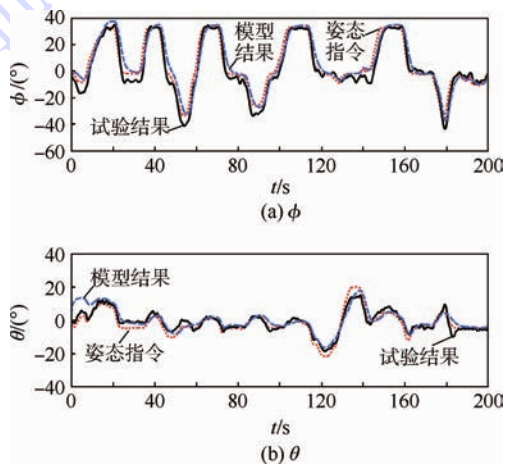


图 22　姿态跟踪试验与仿真对比

Fig. 22　Comparison of attitude tracking between flight test and simulation

北航学报 赠阅

效果吻合度较高,由于建模中对于旋翼挥舞动态、非均匀入流等的简化处理、对传感器模型及风干扰的省略,二者在对比的初始时刻与大的姿态指令变化时,存在一定的跟踪误差。虽然本文的模型验证方法较文献[18]简化,但在一定程度上验证了本文所建立的小型无人旋翼机数学模型的预测能力。

5 结 论

1) 基于解析形式叶素法建立的自转旋翼气动模型与风洞试验数据及数值求解叶素法吻合程度较高,基于该方法建立的旋翼机模型可用于旋翼机飞行动力学特性分析。

2) 对比分析了 2 种操纵方式在配平、稳定性及操纵性等方面的差异,研究表明,旋翼操纵方式的配平俯仰姿态变化更小、纵向稳定性更好,而升降舵操纵方式具有更大的俯仰可达力矩,并能改善螺旋模态的稳定性。

3) 基于经典 PID 理论设计了样例无人旋翼机姿态控制律,实现了姿态跟踪控制飞行试验,分析了全飞行过程中旋翼转速变化特性,并基于试飞数据对旋翼机模型进行了验证。

致谢 感谢北京航空航天大学交通科学与工程学院吴江浩教授、王浩在旋翼机建模方面提供的帮助。感谢胡亮在样例无人旋翼机 A 的设计制造与试验方面做出的贡献。

参考文献 (References)

[1] LEISHMAN J G. Development of the autogiro: A technical perspective[J]. Journal of Aircraft, 2004, 41(4): 765-781.

[2] 王焕瑾,高正. 自转旋翼的气动优势和稳定转速[J]. 航空学报,2001,22(4):337-339.
WANG H J, GAO Z. Aerodynamic virtue and steady rotary speed of autorotating rotor[J]. Acta Aeronautica et Astronautica Sinica, 2001, 22(4): 337-339(in Chinese).

[3] HOUSTON S S, THOMSON D G. The aerodynamics of gyroplanes: CAA Paper 2009/02[R]. West Sussex: Civil Aviation Authority, 2010.

[4] HOUSTON S S. Identification of autogyro longitudinal stability and control characteristics[J]. Journal of Guidance, Control, and Dynamics, 1998, 21(3): 391-399.

[5] THOMSON D G, HOUSTON S S, SPATHOPOULOS V M. Experiments in autogiro airworthiness for improved handling qualities[J]. Journal of the American Helicopter Society, 2005, 50(4): 295-301.

[6] BAGIEV M, THOMSON D G. Handling qualities evaluation of an autogiro against the existing rotorcraft criteria[J]. Journal of Aircraft, 2009, 46(1): 168-174.

[7] 崔钊,韩东,李建波,等. 加装格尼襟翼的自转旋翼气动特性研究[J]. 航空学报,2012,33(10):1791-1799.
CUI Z, HAN D, LI J B, et al. Study on aerodynamic characteristics of auto-rotating rotors with Gurney flaps[J]. Acta Aeronautica et Astronautica Sinica, 2012, 33(10): 1791-1799(in Chinese).

[8] 朱清华. 自转旋翼飞行器总体设计关键技术研究[D]. 南京:南京航空航天大学,2007:37-48.
ZHU Q H. Research on key technologies of gyroplane preliminary design[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2007: 37-48(in Chinese).

[9] 王俊超,李建波,韩东. 自转旋翼机飞行性能理论建模技术[J]. 航空学报,2014,35(12):3244-3253.
WANG J C, LI J B, HAN D. Theoretical modeling technology for gyroplane flight performance[J]. Acta Aeronautica et Astronautica Sinica, 2014, 35(12): 3244-3253(in Chinese).

[10] 王俊超,李建波. 自转旋翼/机翼组合构型飞行器飞行动力学特性[J]. 南京航空航天大学学报,2011,43(3):399-405.
WANG J C, LI J B. Flight dynamics characteristics of autorotating rotor/wing combination aircraft[J]. Journal of Nanjing University of Aeronautics and Astronautics, 2011, 43(3): 399-405(in Chinese).

[11] 王俊超,李建波. 机翼对自转旋翼机纵向稳定性的影响[J]. 航空学报,2014,35(1):151-160.
WANG J C, LI J B. Effects of wing on autogyro longitudinal stability[J]. Acta Aeronautica et Astronautica Sinica, 2014, 35(1): 151-160(in Chinese).

[12] 陈淼. 自转式无人旋翼机飞行控制技术研究[D]. 南京:南京航空航天大学,2012:16-88.
CHEN M. Research on flight control technologies for unmanned gyroplane[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2012: 16-88(in Chinese).

[13] PROUTY R W. Helicopter performance, stability, and control[M]. Boston: PWS Engineering, 1986: 163-187.

[14] 陈仁良,高正. 直升机飞行动力学[M]. 北京:科学出版社,2003:18-19.
CHEN R L, GAO Z. Helicopter flight dynamics[M]. Beijing: Science Press, 2003: 18-19(in Chinese).

[15] MATTHEW J T. Elevators in autogyro propeller wake enable low-speed pitch control[J]. Aircraft Engineering and Aerospace Technology, 2011, 83(3): 154-159.

[16] MATTHEW J T, CARTER R G. Pitch control benefits of elevators for autogyros in low-speed forward flight[C]// Proceedings of 43rd AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2005: 11953-11961.

[17] LOPEZ C A, WELLS V L. Dynamics and stability of an autorotating rotor/wing unmanned aircraft[J]. Journal of Guidance, Control, and Dynamics, 2004, 27(2): 258-270.

[18] HOUSTON S S. Validation of a rotorcraft mathematical model for autogyro simulation[J]. Journal of Aircraft, 2000, 37(3): 403-409.

[19] HOUSTON S S. Identification of gyroplane lateral/directional stability and control characteristics from flight test[J]. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 1998, 212(4): 271-285.

[20] STEVENS B L, LEWIS F L. Aircraft control and simulation [M]. Hoboken, NJ: John Wiley & Sons, 2003: 101-138.

[21] DREIER M E. 直升机和倾转旋翼飞行器飞行仿真引论 [M]. 孙传伟，译. 北京：航空工业出版社，2014：323-346.
DREIER M E. Introduction to helicopter and tiltrotor simulation [M]. Translated by SUN C W. Beijing: Aviation Industry Press, 2014: 323-346 (in Chinese).

[22] 闫坤，蔡志浩，林清，等. 带预旋自转旋翼机起降阶段关键飞行参数测量方法[C]// IEEE Chinese Guidance, Navigation and Control Conference 2014, 2014: 1643-1648.
YAN K, CAI Z H, LIN Q, et al. Key parameters measurement method in takeoff and landing of autogyro with prerotating[C]// IEEE Chinese Guidance, Navigation and Control Conference 2014, 2014: 1643-1648 (in Chinese).

[23] FUTABA. Futaba SBS-01RM Magnetic RPM Sensor[EB/OL]. Japan: Futaba radio control (RC-R/C) systems and accessories, 2015 [2015-11-20]. http://www.futabarc.com/accessories/futm0850.html.

[24] 杨金鹏，蔡志浩，林清，等. 一种复合式短距起降无人机自动飞行控制系统设计[C]// IEEE Chinese Guidance, Navigation and Control Conference 2014, 2014: 1044-1049.
YANG J P, CAI Z H, LIN Q, et al. Autonomous flight control system design for STOL flight of a compound autogyro UAV[C]// IEEE Chinese Guidance, Navigation and Control Conference 2014, 2014: 1044-1049 (in Chinese).

作者简介：

林清　男，博士，工程师。主要研究方向：自转旋翼无人机飞行控制、旋翼-机翼复合式无人机多模态飞行控制。
Tel.：15101522879
E-mail：linqingbh@163.com

蔡志浩　男，博士，副教授。主要研究方向：无人机自主导航与控制、多机协同与任务规划、机器视觉。
Tel.：010-82338792
E-mail：czh@buaa.edu.cn

闫坤　男，硕士研究生。主要研究方向：复合式无人机飞行控制。
Tel.：010-82338792
E-mail：1182668247@qq.com

王英勋　男，研究员，博士生导师。主要研究方向：无人机自主控制、无人机系统工程。
Tel.：010-82338792
E-mail：wangyx@buaa.edu.cn

Influence of pitch manipulation modes on controllability and stability of autogyro

LIN Qing[1,2], CAI Zhihao[1,2,*], YAN Kun[1,2], WANG Yingxun[1,2]

(1. School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;
2. Science and Technology on Aircraft Control Laboratory, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: To study the influence of two manipulation modes on flight dynamics characteristics of autogyro, an example autogyro UAV was chosen and modeled based on closed-form blade element method. Then the differences of the two modes in trim results, stability and handling characteristics were analyzed in the whole speed envelope. The results indicate that each manipulation mode has its advantages and disadvantages: for the rotor manipulation mode, the trim pitch angle has little change and the phugoid mode is stable, but the spiral mode is unstable in the whole envelope. For the elevator manipulation mode, the spiral mode is stable at relatively high speed and the attainable pitch moment is larger, but the trim angle of attack is negative at high speed and the speed static stability is negative. Then two prototype autogyro UAVs were assembled and flight-tested for the two manipulation modes respectively. The autorotation of the main rotor during the whole flight were analyzed. Attitude control laws were designed and flight-tested for the two prototype UAVs, which both achieveattitude tracking relatively well. Finally, the model of the autogyro UAV was validated with the flight test data.

Key words: autogyro; manipulation mode; blade element method; attitude control; model validation

Received: 2015-11-25; **Accepted**: 2016-01-27; **Published online**: 2016-03-15 14:47
URL: www.cnki.net/kcms/detail/11.2625.V.20160315.1447.009.html
* **Corresponding author**. Tel.: 010-82338792　E-mail: czh@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0683

再生型壳与 ZTC4 熔体的界面稳定性

柴连静[1], 程序[1], 吴国清[1,*], 王红[2], 南海[2,3]

(1. 北京航空航天大学 材料科学与工程学院, 北京 100083; 2. 北京百慕航材高科技股份有限公司, 北京 100080;
3. 北京航空材料研究院, 北京 100096)

摘 要: 从研究再生型壳中几种典型金属离子对铸件表面的污染影响规律入手,通过在型壳材料中添加 Al、Nb 和 Mn 等金属元素实际浇铸 ZTC4 合金(Ti-6Al-4V),探讨不同残留金属在铸造过程中对铸件表面质量、陶瓷/钛合金界面成分、组织形貌及显微硬度等的影响,揭示废弃型壳材料中铸造残留金属对钛合金新铸坯陶瓷/合金界面稳定性的影响规律。结果表明:含 Al、Nb 和 Mn 等金属残留的型壳浇铸 ZTC4 合金,铸件表面外观平整无明显铸造缺陷;相比无金属残留型壳,含残留金属型壳的铸件表面粗糙度明显增加;Al 金属残留对型壳与熔融钛合金界面稳定性的影响较小,而 Nb 和 Mn 金属残留能明显降低型壳与熔融钛合金的界面稳定性,铸件表面粗糙度可以在一定程度上反映界面稳定性;当型壳中含有 Al 金属残留时,陶瓷/合金界面处析出 Ti_5Si_3 相。

关 键 词: 钛合金; 精密铸造; 陶瓷型壳; 界面反应; α 层; 回收
中图分类号: V252.2; TG146.23
文献标识码: A **文章编号**: 1001-5965(2016)11-2466-06

熔模铸造钛合金目前已被广泛应用于航空航天、船舶汽车等多个领域[1-2]。但由于钛合金在其熔融状态下化学活性较高[3],易与型壳材料发生界面反应,在合金表面形成氧含量较高的 α 反应层[4-6]。由于 α 层硬度高、脆性大,降低了铸件的塑性和疲劳性能[7],因此,研究型壳材料在铸造过程中与熔融钛合金的界面稳定性,对准确评估及彻底清除铸件反应层十分关键。

钛合金精密铸造的快速发展导致了铸造用废弃模具材料大量增加。据统计,全世界每年钛合金铸造业重达 10 万 t 的废弃陶瓷铸型中,精密铸造用废弃陶瓷型壳大约占 60% 以上[8]。目前为止,这些废弃铸型主要是通过填埋处理,不仅增加了成本,更会造成环境污染。Stephen 和 Kunes[9] 在 1981 年就明确分析了工业陶瓷固体废弃物的回收及工业生产成本。1999 年,Fang 和 Smith[10-11] 等在前人研究的基础上进一步深入评估,指出回收利用模具材料可大幅降低精密铸造生产成本,具有非常高的研究价值和应用意义。2010 年,Holtzer等[8] 发现,用废弃模具多次破碎后制得型壳的强度,比氧化铝型壳约高 60%。

目前钛合金精密铸造用型壳主要由化学稳定性较高的面层材料[12-13](直接接触熔融合金)和提供强度及透气性的背层[14]两部分组成,而背层一般认为对铸造过程中陶瓷/熔融金属界面稳定性影响较小,因此可将回收材料在用于型壳背层材料。然而,在浇铸过程中面层材料可能会由于受到冲击和侵蚀而被破坏,导致背层材料暴露,造成活性较高的钛合金与背层材料接触,引起型壳与熔融合金的剧烈反应[15],如 Yuan 等[16] 的研究指出,在铸造 TiAl 合金的反应层发现含 Si 相 $(Ti,Zr)_5(Al,Si)_3$,表明铸造过程中有背层材料的 Si

收稿日期: 2015-10-20; **录用日期**: 2016-01-22; **网络出版时间**: 2016-02-29 17:59
网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20160229.1759.001.html
基金项目: 中国博士后科学基金(2014M550587)
* **通讯作者**: Tel.: 010-82313240 E-mail: guoqingwu@buaa.edu.cn

引用格式: 柴连静, 程序, 吴国清, 等. 再生型壳与 ZTC4 熔体的界面稳定性[J]. 北京航空航天大学学报, 2016, 42(11): 2466-2471. CHAI L J, CHENG X, WU G Q, et al. Interfacial stability of regenerated moulds against molten ZTC4 alloy[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(11): 2466-2471 (in Chinese).

元素渗入到合金表面。除此之外，由于铸造废弃型壳材料中含有铸造金属残留，直接重复使用废弃型壳材料将带来型壳界面稳定性降低及增加新铸坯表面污染的风险。因此掌握金属残留在铸造过程与熔融钛合金的界面稳定性，成为实现废弃型壳循环利用的关键问题。

本文通过在精密铸造型壳背层材料中添加几种典型金属元素，研究 Al、Nb 和 Mn 等残留金属对 ZTC4 合金铸件表面质量、型壳/铸件界面成分、组织形貌及显微硬度的影响，分析残留金属对 ZTC4 合金铸坯界面稳定性影响规律，为实现废弃型壳回收利用及其铸件 α 层清除提供科学依据。

1　试验材料及方法

1.1　型壳制备

图 1 所示为试验中含金属残留的型壳结构示意图及实际型壳结构。其中型壳面层浆料由氧化钇粉末（200 目）、醋酸锆溶胶（自制）及一些添加剂按一定固/液重量比制备，散沙选用氧化钇（50/80 目）；背层浆料分别由莫来石粉末（200 目）及均匀混有 5wt% Al、Nb、Mn 金属粉末（200 目）的莫来石粉末按一定固/液比和硅溶胶溶液（30wt%）制备而成，散沙选用莫来石（30/50 目），且含金属残留的背层涂有 3 层。试验中型壳制备材料为北京百慕航材高科技股份有限公司提供，金属粉末为江西章贡金镱镀膜材料公司提供，纯度均达到 99% 以上。制备完成的型壳面层在温度为25℃，相对湿度为 50% 的环境中干燥 24 h，背层在相同环境中干燥 8 h。在 300℃下用电脱蜡斧对型壳脱蜡，将清洗后的试块在1 050℃预烧结 1 h。

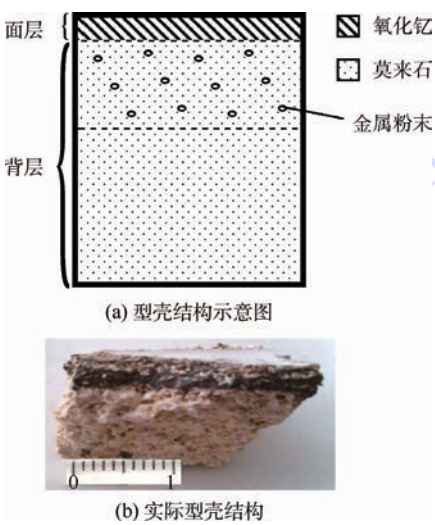


图 1　含金属残留型壳的结构示意图及实际型壳结构

Fig. 1　Sketch map of structure of mould with metal residue and actual mould structure

1.2　ZTC4 合金精密铸造

ZTC4 合金化学成分见表 1。试验在北京航空材料研究院完成，采用真空重力浇注法，型壳预热温度为 450℃，铸块尺寸 70 mm × 70 mm × 100 mm（中空铸件，壁薄 20 mm），铸造后采用机械法破碎型壳。铸件取样位置如图 2 所示。

表 1　ZTC4 合金化学成分

Table 1　Chemical composition of ZTC4 alloy

wt%

元素	Al	V	Fe	Si	C
含量	6.120 0	4.100 0	0.062 0	0.011 0	0.020 0
元素	N	H	O	其他	Ti
含量	0.014 0	0.002 4	0.170 0	<0.400 0	余量

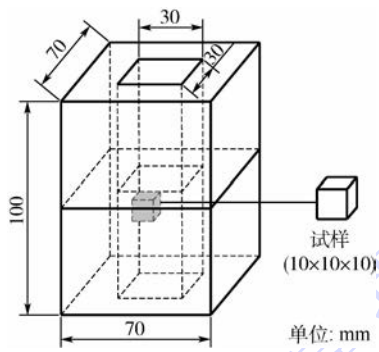


图 2　铸件取样位置示意图

Fig. 2　Sketch map of sampling position of cast bars

1.3　测试方法

ZTC4 合金截面试样经抛光后，首先用 96 mL H_2O + 3 mL HNO_3 + 1 mL HF 的 Kroll 腐蚀液腐蚀约 20 s，腐蚀出铸造钛合金组织，再用质量分数 1% 的 NH_4F 水溶液腐蚀 15 ~ 30 s，腐蚀出 α 层组织。采用 LEICA DM4000 M 光学金相显微镜对合金的显微组织进行观察。利用 FEI QUANTA 200F 扫描电镜、JXA-8100 电子探针对型壳/ZTC4 铸件界面的形貌及成分进行分析。采用 D/max 2200PC XRD 分析试样物相成分，扫描角度为 10° ~ 90°。采用 Buehler micromet 6040 显微硬度仪测量铸件的显微硬度。硬度测试沿试样横截面表面开始进行，从铸件表面每隔 20 μm 测量 3 个点取其平均，测试载荷为 25 gf，1 000 gf ≈ 9.8 N，保压时间为 15 s，获得铸件显微硬度沿厚度方向的分布。

2　试验结果及讨论

2.1　残留金属对铸件表面形貌及成分的影响

如图 3 为 ZTC4 铸件表面形貌（表面未经任何喷砂、清洗处理），可见铸件表面外观平整，无铸瘤等明显铸造缺陷。由表 2 中不同型壳铸件表

北京航空航天大学学报

面粗糙度可看出，Al、Nb 和 Mn 等残留金属会使铸件表面粗糙度（轮廓算术平均偏差 R_a 和轮廓最大高度 R_z）明显增加，且表面粗糙度按照 Mn 残留 < Al 残留 < Nb 残留依次递增。经 XRD 分析，铸件表面粗糙凸起物成分主要为面层残留 Y_2O_3，Ti 及其氧化物 TiO、Ti_2O，且不同型壳铸件凸起物成分基本相同（见图 4），未见残留金属扩散至界面处对铸件表面成分产生明显影响。

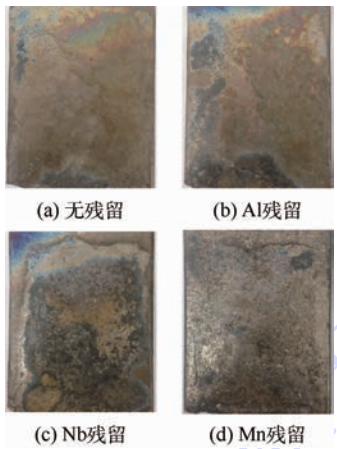


图 3 含不同金属残留型壳的 ZTC4 铸件表面形貌

Fig. 3 Surface appearance of ZTC4 casts using moulds with different metal residues

表 2 含不同金属残留型壳 ZTC4 铸件的表面粗糙度

Table 2 Surface roughness of ZTC4 casts using moulds with different metal residues

型壳中的金属残留 / 铸件种类	粗糙度/μm	
	R_a	R_z
无残留	10.6 ± 2.32	51.5 ± 11.06
Al 残留	22.5 ± 7.76	101.6 ± 49.72
Nb 残留	32.6 ± 2.98	139.5 ± 16.13
Mn 残留	21.1 ± 5.97	95.3 ± 23.90

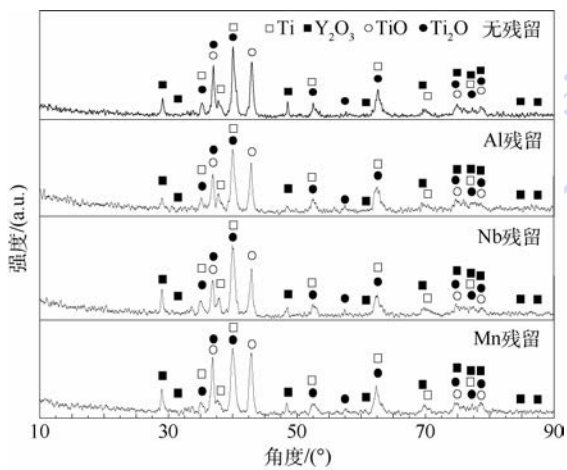


图 4 含不同金属残留型壳的 ZTC4 铸件表面凸起物的 XRD 结果

Fig. 4 XRD results of protrusions at surface of ZTC4 casts using moulds with different metal residues

2.2 残留金属对型壳/铸件反应层厚度的影响

铸造过程中，型壳中的氧元素扩散进入熔融钛合金，使得钛合金表面形成粗大的 α 相片层组织，同时导致合金表面硬度增加。因此，铸件表面 α 层和硬化层的厚度反映了型壳/铸件的界面稳定程度。如图 5(a) ~ 图 5(d)所示，ZTC4 合金铸造后金相组织为典型的 α + β 片层结构，本文所指 α 层厚度为铸件表面粗大片层区厚度（图中箭头所示）。在铸件表面随机取 10 个位置测量其 α 层厚度，去掉一个最高值和一个最低值，取其余8 个值的平均值，获得平均 α 层厚度。各铸件表面平均 α 层厚度如图 5(e)所示，含 Al 金属残留的型壳铸件表面的 α 层厚度与无残留型壳铸件差异不大，残留 Nb 和 Mn 金属型壳的铸件表面 α 脆性层厚度明显增加。如图 6 所示为合金表面显微硬度分布情况，利用指数函数进行拟合（图中实线），可以看出，ZTC4 合金的显微硬度约为 380 HV，使用无金属残留型壳的铸件表面硬化层厚度约为1 250 μm，采用含 Al、Nb 和 Mn 残留的型壳铸件表面硬化层厚度分别为1 250、1 450 和 1 450 μm 左右。可见，Al 金属残留对铸件表面 α 层和硬化层厚度的影响不明显，对陶瓷/钛合金界面稳定性的影响较小，而 Nb 和 Mn 会明显降低陶瓷/钛合金界面稳定性，钛合金表面的 α 层及硬化层厚度增加了 15% 左右。

由表 2、表 3 中不同型壳铸件的表面粗糙度、α 层厚度及硬化层厚度，铸件表面粗糙度按 Mn 残留 < Al 残留 < Nb 残留递增，硬度层厚度较 α 层略小，α 层厚度按 Al 残留 < Mn 残留 < Nb 残留递增，硬化层厚度按 Al 残留 < Mn 残留 ≈ Nb 残留递增，可见铸件表面粗糙度可以在一定程度上反映陶瓷/ZTC4 合金界面稳定性。

表 3 含不同金属残留型壳 ZTC4 铸件的平均 α 层厚度及硬化层厚度

Table 3 Average alpha-case layer thickness and hardened layer thickness of ZTC4 casts using moulds with different metal residues

型壳中的金属残留 / 铸件种类	厚度/μm	
	平均 α 层	平均硬化层
无残留	1 280 ± 189.5	1 250
Al 残留	1 285 ± 174.1	1 250
Nb 残留	1 479 ± 85.4	1 450
Mn 残留	1 419 ± 142.2	1 450

2.3 残留金属对型壳/铸件界面微观结构及成分的影响

如图 7 为铸件显微组织的背散射电子像，可见 ZTC4 合金的显微组织主要由 α（深灰色相）+ β（浅灰色相）的双相魏氏体组织组成。铸造

北航学报 赠阅

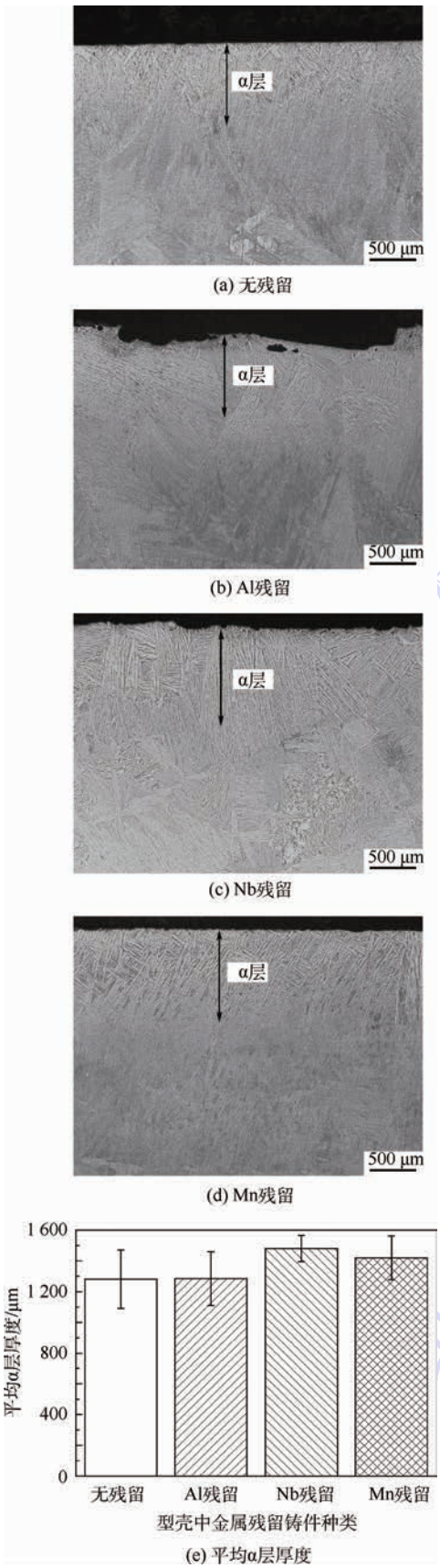


图 5　含不同金属残留的型壳 ZTC4 铸件的金相组织及平均 α 层厚度

Fig. 5　Metallographic structures and average alpha-case layer thickness of ZTC4 casts using moulds with different metal residues

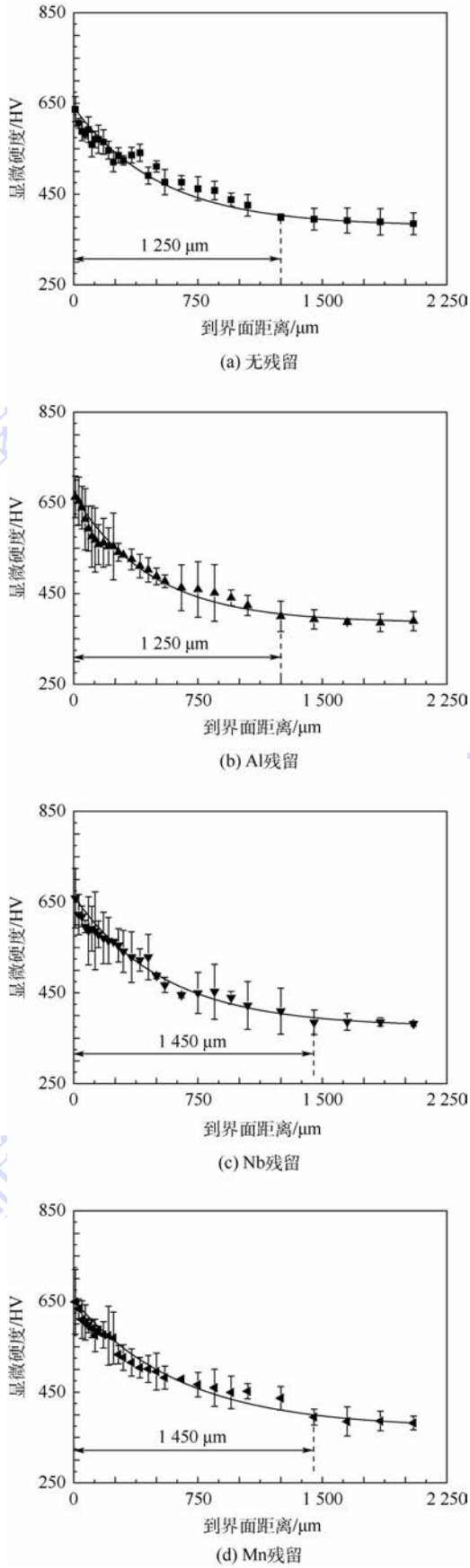


图 6　含不同金属残留型壳的 ZTC4 铸件的表面显微硬度

Fig. 6　Surface micro-hardness of ZTC4 casts using moulds with different metal residues

北航学报 赠阅

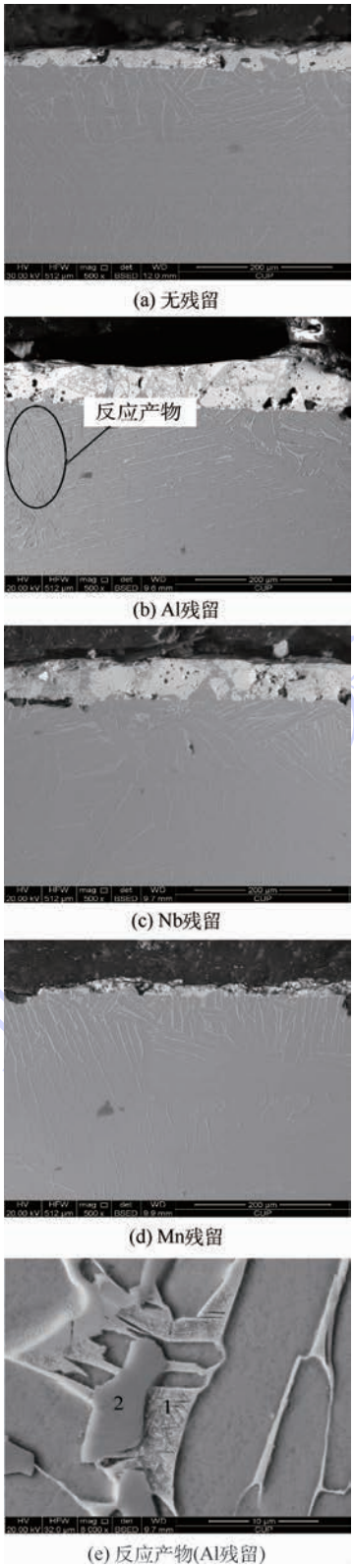


图 7 含不同金属残留的型壳与 ZTC4 铸件反应界面的背散射电子像

Fig. 7 Backscattered images of interface of ZTC4 casts against moulds with different metal residues

过程中型壳中的氧元素扩散进入 ZTC4 合金,使铸件表面形成粗大的 α 组织,且采用不同的型壳材料,试样表面粗化 α 相片层厚度不同:当使用不含金属残留的型壳时,粗化 α 相片层厚度约为 10.1 μm;采用含有金属残留的型壳后,粗化层厚度均有增加,且按 Al(12.9 μm) < Mn(13.8 μm) < Nb(14.3 μm)依次增加。此外,在各合金表面的 β 相中均观察到了细小的深灰色针状析出物,经 EDX 分析可知(图 7(e)位置 1),其中含有少量的 Zr 元素,可推断是面层浆料中的 Zr 元素扩散进入熔融合金所致。当使用含 Al 残留的型壳时,在陶瓷/合金界面观察到了大量随机分布的析出物(颜色与 α 相近),EDX 分析可知(图 7(e)位置 2),这些析出物中 Si、Ti 含量较高,Al 含量较低,可判定为 Ti_5Si_3 相。

3 结 论

1) 背层含有 Al、Nb、Mn 等金属残留的型壳浇铸 ZTC4 合金,铸件表面外观平整,无铸瘤等明显铸造缺陷。相较无金属残留型壳,含残留金属型壳的铸件表面粗糙度按 Mn 残留 < Al 残留 < Nb 残留明显增加。

2) 含 Al 金属残留的型壳铸件表面的 α 脆性层及硬化层厚度与无金属残留型壳的铸件差异不大,而残留 Mn 和 Nb 的型壳铸件表面 α 层及硬化层厚度明显增加。对比铸件表面粗糙度及陶瓷/合金界面反应层厚度发现,表面粗糙度可以在一定程度上反映界面反应稳定性。

3) 铸造过程中,型壳面层浆料中的 Zr 元素少量扩散进入合金,在 β 相内析出含 Zr 的细小针状析出物;使用残留 Al 的型壳时,背层材料中 Si 元素扩散进入熔融钛合金,生成 Ti_5Si_3 相随机分布于界面处。

参考文献(References)

[1] GURRAPPA I. Characterization of titanium alloy Ti-6Al-4V for chemical, marine and industrial applications [J]. Materials Characterization, 2003, 51(2): 131-139.

[2] BOYER R R. An overview on the use of titanium in the aerospace industry [J]. Materials Science and Engineering: A, 1996, 213(1-2): 103-114.

[3] KOSTOV A, FRIEDRICH B. Predicting thermodynamic stability of crucible oxides in molten titanium and titanium alloys [J]. Computational Materials Science, 2006, 38(2): 374-385.

[4] SUNG S Y, KIM Y J. Alpha-case formation mechanism on titanium investment castings [J]. Materials Science and Engineering: A, 2005, 405(1): 173-177.

[5] CHAN K S, KOIKE M, JOHNSON B W, et al. Modeling of alpha-case formation and its effects on the mechanical properties of titanium alloy castings [J]. Metallurgical and Materials Transactions A, 2008, 39(1): 171-180.

北航学报 赠阅

[6] BOETTINGER W, WILLIAMS M, CORIELL S, et al. Alpha case thickness modeling in investment castings[J]. Metallurgical and Materials Transactions B, 2000, 31(6): 1419-1427.

[7] YUE L Y, WANG A B, LI L. Material morphological characteristics in laser ablation of alpha case fromtitanium alloy[J]. Applied Surface Science, 2012, 258(20): 8065-8071.

[8] HOLTZER M, ZYCH J, DANKO R, et al. Reclamation of material from used ceramic moulds applied in the investment casting technology[J]. Journal of Achives of Foundry Engineering, 2010, 10(3): 199-104.

[9] STEPHEN W A, KUNES T P. Cutting the cost of disposal through innovative and constructive uses of foundry wastes[J]. AFS Transactions, 1981, 89: 697-708.

[10] FANG H, SMITH J D, PESSLEE K D. Study of spent refractory waste recycling from metal manufacturers in Missouri[J]. Resource Conservation and Recycling, 1999, 25(2): 111-124.

[11] SMITH J D, FANG H, PEASLEE K D. Characterization and recycling of spent refractory wastes fron metal manufacturers in Missouri[J]. Resource Conservation and Rcycling, 1999, 25(2): 151-169.

[12] SAHA R L, NANDY T K, MISRA R D K, et al. On the evaluation of stability of rare earth oxides as face coats for investment casting of titanium[J]. Metallurgical Transactions B, 1990, 21(3): 559-566.

[13] SAHA R L, MISRA R D K. Formation of low-melting eutectic at the metal-mould interface during titanium casting in zircon sand moulds[J]. Journal of Materials Science Letters, 1991, 10(22): 1318-1319.

[14] JOVANOVIÓ M T, DIMŎIÓ B, BOBIÓ I, et al. Microstructure and mechanical properties of precision cast TiAl turbocharger wheel[J]. Journal of Materials Processing Technology, 2005, 167(1): 14-21.

[15] YUAN C, COMPTON D, CHENG X, et al. The influence of polymer content and sintering temperature on yttria face-coat moulds for TiAl casting[J]. Journal of the European Ceramic Society, 2012, 32(16): 4041-4049.

[16] YUAN C, CHENG X, HOLT G S, et al. Investment casting of Ti-46Al-8Nb-1B alloy using moulds with CaO-stabilized zirconia face coat at various mould pre-heat temperatures[J]. Ceramics International, 2014, 41(3): 4129-4139.

作者简介：

柴连静　女，硕士研究生。主要研究方向：钛合金精密铸造。
E-mail：chailianjing@163.com

吴国清　男，博士，副教授。主要研究方向：轻质合金及复合材料设计与成型技术。
Tel.：010-82313240
E-mail：guoqingwu@buaa.edu.cn

Interfacial stability of regenerated moulds against molten ZTC4 alloy

CHAI Lianjing[1], CHENG Xu[1], WU Guoqing[1,*], WANG Hong[2], NAN Hai[2,3]

(1. School of Materials Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;
2. Baimtec Materials Co., Ltd., Beijing 100080, China;
3. Beijing Institute of Aeronautical Materials, Beijing 100096, China)

Abstract: Aimed at the effect rules of several typical metal ions from reclaimed shells on contaminating casting surface, with the method of adding metallic elements such as Al, Nb and Mn to shell materials used to actually cast ZTC4 alloy (Ti-6Al-4V), this study analyzed the influence of different residual metals in casting process on casting surface quality, ceramic/titanium alloy interaction compositions, microstructure morphology and micro-hardness, and discussed the impact of residual casting metal in recycled mould materials on the interface stability between ceramic mould and new titanium alloy cast. The results show that the cast surface of ZTC4 alloy using moulds containing Al, Nb or Mn ions looks smooth with no obvious casting defects; compared with none metal containing mould, the roughness of ZTC4 cast using shells with residual metal increases markedly; Al vestigital had little effect on the interface inertness of ceramic against molten titanium alloy, while Nb and Mn can obviously reduce the interface inertness, and to some degree, casting surface roughness can reflect the interface stability; when the mould contains Al, Ti_5Si_3 phase is found at ceramic/metal interface.

Key words: titanium alloy; investment casting; ceramic mould; interface reaction; alpha-case layer; recycling

Received: 2015-10-20; **Accepted**: 2016-01-22; **Published online**: 2016-02-29 17:59
URL: www.cnki.net/kcms/detail/11.2625.V.20160229.1759.001.html
Foundation item: China Postdoctoral Science Foundation (2014M550587)
* **Corresponding author**. Tel.: 010-82313240　E-mail: guoqingwu@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0725

基于 LADRC 的无人机高精度定高控制

齐鹏远[1], 王勇[1,*], 张代兵[2]

(1. 北京航空航天大学 自动化科学与电气工程学院, 北京 100083; 2. 国防科技大学 机电工程与自动化学院, 长沙 410073)

摘 要: 针对复杂气流扰动对无人机(UAV)航迹高度控制的影响,对存在复杂气流扰动下的定高控制策略、控制结构和控制器参数优化展开研究,实现高精度高度控制。基于线性自抗扰控制(LADRC)确定总体控制架构,设计扩张状态观测器(ESO)观测估计纵向高度通道和速度通道中存在的总扰动,在控制中引入扰动补偿,减小扰动对系统输出造成的影响。对UAV在飞行过程中存在的大气紊流扰动或离散突风等风干扰分析其功率谱密度,构造考虑风扰动对高度影响、时域响应特性和稳定裕度的综合目标函数,通过粒子群优化算法得到具有高精度、高抗干扰性能的控制器参数,优化中考虑风干扰的功率谱密度分布,减小了控制器参数设计的保守性。通过与常规比例-积分-微分(PID)控制器控制效果进行对比,说明基于线性自抗扰控制器的纵向高度控制的优异性能。

关 键 词: "天钩"回收; 高精度高度控制; 线性自抗扰控制(LADRC); 风干扰分析; 粒子群优化算法

中图分类号: V249

文献标识码: A **文章编号**: 1001-5965(2016)11-2472-09

"天钩"回收方式是一种新型的无人机(UAV)零长回收方式,具有对回收场地要求低、回收方式简单且安全可靠等优点,非常适合小型固定翼无人机。当无人机"天钩"回收时,首先下降至回收高度,跟踪期望的回收航线保持定高直线飞行,撞向"天钩"回收机构的回收索,回收索沿着具有后掠角的机翼向翼尖滑动,最终滑入安装在翼尖的回收钩,实现"天钩"回收[1]。回收飞行过程中存在的复杂气流扰动将对飞行轨迹产生影响,可能导致回收失败。因此追求高精度的纵向高度控制和横侧向侧偏距控制在此类问题中十分必要,本文仅对其中的高精度高度控制进行研究。

工程中常使用比例-积分-微分(PID)控制器进行高度控制,在无扰动情况下可以实现无静差控制,但在扰动情况下,控制精度往往难以达到要求,为此国内外学者在高精度高度控制上展开了大量研究。文献[2]针对无人机的高精度着陆控制采用了L1自适应控制,仿真结果表明误差满足精度要求,但如何选取最优的反馈回路低通滤波器难度较大。文献[3]采用模糊增益调度控制在紊流等扰动存在的情况下,提升了轨迹跟踪性能,但规则库的设计对经验要求较高。文献[4]构造非线性干扰观测器对建模误差和外在干扰进行观测,并在此基础上设计滑模控制器,虽可实现较高的控制精度,但存在控制抖振问题。文献[5]对俯仰角控制应用自抗扰控制器,取得了较好的控制效果,但并未给出有效的参数设计方法。

韩京清[6-7]提出的自抗扰控制方法具有鲁棒性强,可控范围大,算法简单,工程易实现等优点。本文基于线性自抗扰控制(LADRC),以无人机高度通道为研究对象,对定高控制策略、控制结构和控制器参数优化设计展开研究,实现高精度高

收稿日期: 2015-11-05; **录用日期**: 2016-01-22; **网络出版时间**: 2016-03-15 14:34
网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20160315.1434.006.html
* **通讯作者**: Tel.: 010-82317544 E-mail: wy_buaa@sina.com

引用格式: 齐鹏远, 王勇, 张代兵. 基于LADRC的无人机高精度定高控制[J]. 北京航空航天大学学报, 2016, 42(11): 2472-2480.
QI P Y, WANG Y, ZHANG D B. Precise height control for UAV based on LADRC[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(11): 2472-2480 (in Chinese).

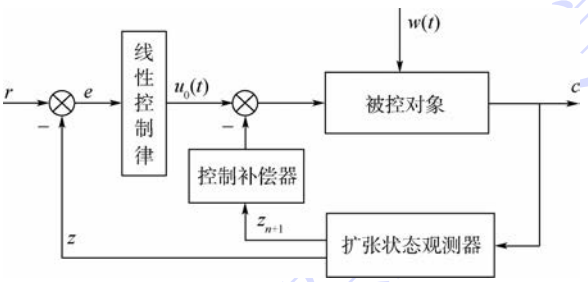
北航学报 赠阅

度控制。

1　线性自抗扰控制器理论

自抗扰控制是针对观测系统的“干扰总和”，并加以补偿的一种控制方法。其利用扩张状态观测器(ESO)观测出系统的未建模动态、非线性动态特性以及外部扰动来进行闭环系统的控制补偿。

虽然传统自抗扰方法具有高精度、高反馈效率等方面的优势，但其选取的参数过多，参数整定比较繁琐，且在工程上难以进行稳定性等指标分析。为此，高志强等[8-9]提出了 LADRC 方法，控制器参数较少，且具有明确的物理含义，十分便于理论分析和工程应用。LADRC 结构如图 1 所示。



r—参考输入；c—被控输出；e—偏差量；$u_0(t)$—初始控制量；z—状态观测量；z_{n+1}—扩张状态观测量；n—状态个数；$w(t)$—作用在系统中的扰动总和。

图 1　LADRC 结构

Fig. 1　Structure of LADRC

设带有未知扰动的不确定对象为

$$y^{(n)}=f(y,y^{(1)},\cdots,y^{(n-1)})+w(t)+b_0u \tag{1}$$

式中：u 和 y 分别为系统输入和输出；b_0 为已知常数。取状态变量 $[x_1,x_2,\cdots,x_n]^{\mathrm{T}}=[y,y^{(1)},\cdots,y^{(n-1)}]^{\mathrm{T}}$，令 $a(t)=f+w(t)$，它包含了系统模型和外扰的总和，称之为系统的扩张状态变量。令 $b(t)=\dot{a}(t)$，可得到系统的扩张状态方程为

$$\begin{cases}\dot{x}_1=x_2\\ \dot{x}_2=x_3\\ \quad\vdots\\ \dot{x}_{n-1}=x_n\\ \dot{x}_n=a(t)+b_0u\\ \dot{x}_{n+1}=b(t)\\ y=x_1\end{cases} \tag{2}$$

构造式(1)系统的线性 ESO 为

$$\begin{cases}\dot{z}_1=z_2-b_1(z_1-x_1)\\ \dot{z}_2=z_3-b_2(z_1-x_1)\\ \quad\vdots\\ \dot{z}_n=z_{n+1}-b_n(z_1-x_1)+b_0u\\ \dot{z}_{n+1}=-b_{n+1}(z_1-x_1)\end{cases} \tag{3}$$

式中：$b_1,b_2,\cdots,b_{n+1}$ 为 $(s+\omega)^{n+1}$ 展开后从第 2 项开始的各项系数，ω 的选取可根据系统带宽的要求确定或在线整定。ESO 可以以一定速度跟踪系统的各个状态，且观测出的扩张状态值即为作用在系统上的扰动总和，即

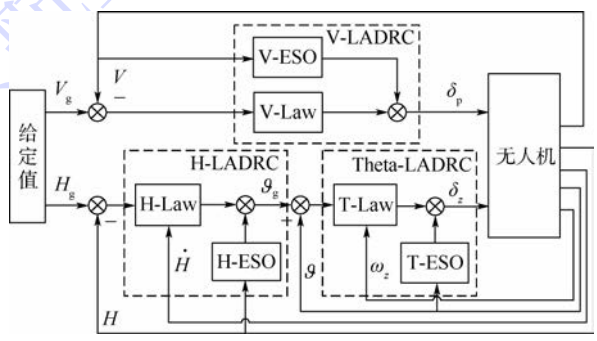
$$\begin{cases}z_1(t)\rightarrow x_1(t)\\ z_2(t)\rightarrow x_2(t)\\ \quad\vdots\\ z_n(t)\rightarrow x_n(t)\\ z_{n+1}(t)\rightarrow x_{n+1}(t)=a(t)\end{cases} \tag{4}$$

自抗扰控制器的本质就是利用 ESO 观测系统的内部和外部扰动。将观测出的扰动总和引入到控制量中进行补偿，即可抑制扰动对系统输出带来的影响，从而提高控制器的控制性能。

2　无人机纵向控制器结构

无人机纵向高度控制通道的常规操纵面主要有升降舵和油门。本文对纵向的高度和速度控制采用俯仰高度/油门空速控制策略，即通过升降舵控制内环俯仰角实现外环高度控制，通过油门控制空速。

将纵向控制通道划分为高度控制通道和速度控制通道，其中高度通道采用内外环控制结构，分为内环俯仰角控制通道和外环高度控制通道。反馈高度 H，俯仰角 ϑ 和空速 V，分别在 3 个通道构造 LADRC 控制器，LADRC 利用 ESO 观测出纵向内环俯仰角通道、外环高度通道和空速通道中存在的扰动，将扰动量引入控制量中进行补偿，达到提高控制精度的效果。无人机纵向通道控制结构如图 2 所示，本文为了便于分析设计，反馈控制律均取为线性形式。



V_g—空速给定值；H_g—高度给定值；$\dot{H}$—高度变化率；ϑ_g—俯仰角给定值；ω_z—俯仰角速度；δ_p—油门开度；δ_z—升降舵舵偏角；H-Law—高度通道控制律；H-ESO—高度通道 ESO；H-LADRC—高度通道 LADRC；T-Law—俯仰角通道控制律；T-ESO—俯仰角通道 ESO；Theta-LADRC—俯仰角通道 LADRC；V-Law—速度通道控制律；V-ESO—速度通道 ESO；V-LADRC—速度通道 LADRC。

图 2　无人机纵向通道控制结构

Fig. 2　Control structure of UAV longitudinal channel

北航学报 赠阅

2.1 内环俯仰角控制器

无人机的姿态运动可用关于俯仰角的二阶微分方程来描述,从纵向运动方程中提取出 ϑ 和 ω_z 的相关项,引入二阶 LADRC 控制器,实现对纵向俯仰角的姿态控制。

首先引入三阶线性 ESO,观测纵向内环可能存在的扰动总和,ESO 表达式为

$$\begin{cases} e_\vartheta = z_{1\vartheta} - \vartheta \\ \dot{z}_{1\vartheta} = z_{2\vartheta} - \beta_{1\vartheta} e_\vartheta \\ \dot{z}_{2\vartheta} = z_{3\vartheta} - \beta_{2\vartheta} e_\vartheta + b_\vartheta u_\vartheta \\ \dot{z}_{3\vartheta} = -\beta_{3\vartheta} e_\vartheta \\ \beta_{1\vartheta} = 3\omega_\vartheta \\ \beta_{2\vartheta} = 3\omega_\vartheta^2 \\ \beta_{3\vartheta} = \omega_\vartheta^3 \end{cases} \tag{5}$$

式中:e_ϑ 为俯仰角跟踪误差;$z_{1\vartheta}$ 和 $z_{2\vartheta}$ 分别为观测得到的俯仰角和俯仰角速度;$z_{3\vartheta}$ 为俯仰角通道观测到的总扰动;b_ϑ 为升降舵舵偏角对俯仰角的控制比例系数;u_ϑ 为实际升降舵舵偏角;$\beta_{1\vartheta}$、$\beta_{2\vartheta}$ 和 $\beta_{3\vartheta}$ 为俯仰角通道 ESO 系数,取决于待整定观测器参数 ω_ϑ。

根据俯仰角给定值与实际值的偏差,以及俯仰角速度,设计线性误差反馈律,得到初始控制量 $u_{0\vartheta}$,再将 ESO 观测的扰动 $z_{3\vartheta}$ 引入到控制中进行补偿,最终得到实际控制量 u_ϑ。

$$\begin{cases} u_{0\vartheta} = K_\vartheta(\vartheta_g - \vartheta) - K_{\omega_z}\omega_z \\ u_\vartheta = u_{0\vartheta} - \dfrac{z_{3\vartheta}}{b_\vartheta} \end{cases} \tag{6}$$

b_ϑ 可根据无人机气动参数和飞行状态确定具体数值。优化控制器参数 K_ϑ、K_{ω_z} 和 ω_ϑ 就可以获得满意的俯仰角观测效果和控制效果。

2.2 外环高度控制器

无人机纵向外环高度运动可用以高度 H 为输出,俯仰角 ϑ 为输入的一阶微分方程描述,将迎角 α 看作扰动量,引入一阶 LADRC 控制器,实现高度控制。

$$\dot{H} = -V\alpha + V\vartheta \tag{7}$$

对此一阶系统引入二阶线性 ESO 观测纵向外环可能存在的扰动总和,引入线性误差反馈控制律实现外环高度控制,具体形式为

$$\begin{cases} e_H = z_{1H} - H \\ \dot{z}_{1H} = z_{2H} - \beta_{1H} e_H + b_H u_H \\ \dot{z}_{2H} = -\beta_{2H} e_H \\ \beta_{1H} = 2\omega_H \\ \beta_{2H} = \omega_H^2 \\ u_{0H} = K_{H_p}(H_g - H) + K_{Hd}\dot{H} \\ u_H = u_{0H} - \dfrac{z_{2H}}{b_H} \end{cases} \tag{8}$$

式中:e_H 为高度跟踪误差;u_{0H} 为高度通道初始控制量;z_{1H} 为观测得到的航迹高度;z_{2H} 为高度通道观测出的总扰动;u_H 为实际俯仰角给定值;β_{1H} 和 β_{2H} 为高度通道 ESO 系数,取决于待整定观测器参数 ω_H;$b_H = V$ 为无人机当前飞行空速。优化控制器参数 K_{H_p}、K_{Hd} 和 ω_H 可以获得满意的高度观测效果和控制效果。

2.3 空速控制器

无人机的空速通道同外环高度通道类似,可用以空速 V 为输出,油门 δ_p 为输入的一阶微分方程来描述,将其余项均看作扰动量 f,引入一阶 LADRC 控制器,实现空速控制。

$$\dot{V} = f + b_V\delta_p \tag{9}$$

式中:$b_V = \cos(\alpha_* + \varphi_p)/m$ 可根据无人机气动参数和飞行状态确定具体数值,φ_p 为发动机安装角,m 为无人机质量,α_* 为配平迎角。对此一阶系统引入二阶线性 ESO 观测速度通道可能存在的扰动总和,引入线性误差反馈控制律实现速度闭环控制,具体形式为

$$\begin{cases} e_V = z_{1V} - V \\ \dot{z}_{1V} = z_{2V} - \beta_{1V} e_V + b_V u_V \\ \dot{z}_{2V} = -\beta_{2V} e_V \\ \beta_{1V} = 2\omega_V \\ \beta_{2V} = \omega_V^2 \\ u_{0V} = K_{V_p}(V_g - V) \\ u_V = u_{0V} - \dfrac{z_{2V}}{b_V} \end{cases} \tag{10}$$

式中:e_V 为速度跟踪误差;b_V 为油门对空速的控制比例系数;z_{1V} 为观测得到的空速;z_{2V} 为速度通道观测出的总扰动;u_V 为实际油门控制量;β_{1V} 和 β_{2V} 为高度通道 ESO 系数,取决于待整定观测器参数 ω_V。优化控制器参数 K_{V_p} 和 ω_V 可以获得满意的速度观测效果和控制效果。

3 复杂气流扰动分析

飞行过程中存在的复杂气流扰动主要有大气紊流和离散突风 2 种[10],控制律能否有效地减小复杂气流扰动带来的影响,是实现高精度控制的关键,因此对这 2 种扰动特性进行研究分析,使控制参数设计具有针对性,减小控制器参数设计的保守性。

对大气紊流(德莱顿模型)和离散突风(1 - cos 模型)[11]进行频谱分析,大气紊流的功率谱密度 $X(\omega)$、$Y(\omega)$ 和 $Z(\omega)$ 如图 3 所示,离散突风的功率谱密度如图 4 所示,ω 为频率。

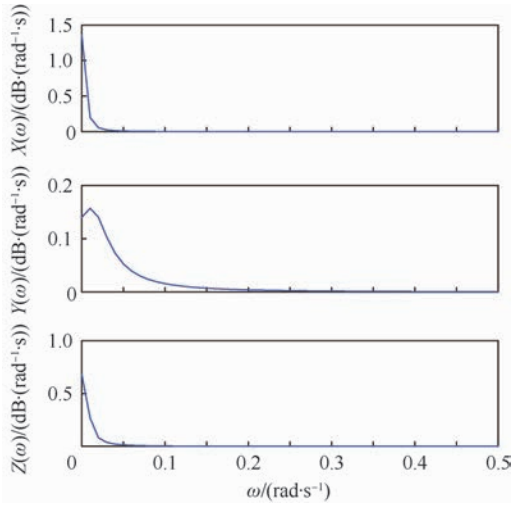


图 3　大气紊流(德莱顿模型)功率谱密度

Fig. 3　Power spectrum density of wind turbulence (Dryden model)

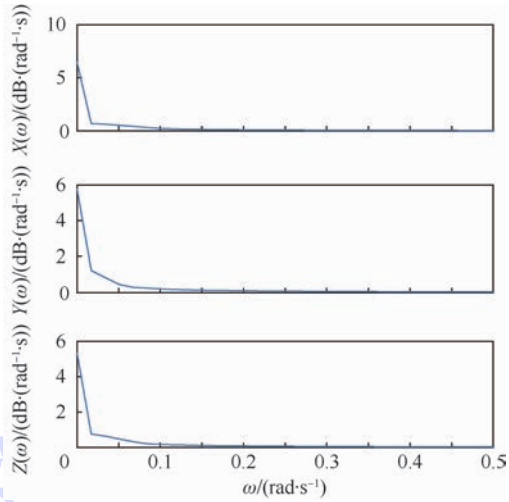


图 4　离散突风(1 - cos 模型)功率谱密度

Fig. 4　Power spectrum density of discrete wind gust (1 - cos model)

从图 3 和图 4 中可以看出，无论是大气紊流还是离散突风，其主要成分频率均在区间[0 ~ 0.2] rad/s 之内。设计时应重点考虑频率在此区间内的大气扰动带来的影响，从而减小设计的保守性。

4　复杂气流扰动下无人机线性模型

文献[12]中对在非平静大气中固定翼飞机的运动方程进行了比较详细的分析。大气扰动的作用实际上就是改变了飞行器上的相对气流，从而改变了飞行器所受到的气动力。根据分析与计算，可以得到在复杂气流扰动下的纵向运动方程为

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\Delta V\\ \Delta\alpha\\ \Delta\omega_z\\ \Delta\vartheta\\ \Delta H\end{bmatrix}=\boldsymbol{A}\begin{bmatrix}\Delta V\\ \Delta\alpha\\ \Delta\omega_z\\ \Delta\vartheta\\ \Delta H\end{bmatrix}+\boldsymbol{B}\begin{bmatrix}\Delta\delta_z\\ \Delta P\\ w_x\\ w_y\\ w_{yx}\end{bmatrix} \tag{11}$$

式中：$\boldsymbol{A}$ 和 $\boldsymbol{B}$ 的定义如下，其余各变量具体意义参见文献[12]。由于小型无人机在飞行过程中可以看作质点，因此方程的输入量紊流梯度量 w_{yx} 可认为是零。在此线性化模型的基础上，可以对复杂扰动对无人机纵向飞行状态的影响进行定量分析，从而为控制器参数优化设计提供抗扰性方面的优化准则。

$$\boldsymbol{A}=\begin{bmatrix}
x^{v} & x^{\alpha}+g\cos\theta_* & 0 & -g\cos\theta_* & x^{h}\\
-\dfrac{y^{v}}{1+y^{\dot\alpha}} & -\dfrac{y^{\alpha}-g\sin\theta_*/v_*}{1+y^{\dot\alpha}} & \dfrac{1-y^{\omega_z}}{1+y^{\dot\alpha}} & -\dfrac{g\sin\theta_*/v_*}{1+y^{\dot\alpha}} & -\dfrac{y^{h}}{1+y^{\dot\alpha}}\\
\mu_z^{v}-\dfrac{\mu_z^{\dot\alpha}y^{v}}{1+y^{\dot\alpha}} & \mu_z^{\alpha}-\dfrac{\mu_z^{\dot\alpha}(y^{\alpha}-g\sin\theta_*/v_*)}{1+y^{\dot\alpha}} & \mu_z^{\omega_z}+\dfrac{\mu_z^{\dot\alpha}(1-y^{\omega_z})}{1+y^{\dot\alpha}} & -\dfrac{\mu_z^{\dot\alpha}g\sin\theta_*/v_*}{1+y^{\dot\alpha}} & \mu_z^{h}-\dfrac{\mu_z^{\dot\alpha}y^{h}}{1+y^{\dot\alpha}}\\
0 & 0 & 1 & 0 & 0\\
0 & -v_* & 0 & v_* & 0
\end{bmatrix}$$

$$\boldsymbol{B}=\begin{bmatrix}
x^{\delta_z} & x^{P} & -x^{v} & x_a^{\alpha}/v_* & 0\\
-\dfrac{y^{\delta_z}}{1+y^{\dot\alpha}} & -\dfrac{y^{P}}{1+y^{\dot\alpha}} & y^{v} & -y_a^{\alpha}/v_* & 0\\
\mu_z^{\delta_z}-\dfrac{\mu_z^{\dot\alpha}y^{\delta_z}}{1+y^{\dot\alpha}} & \mu_z^{P}-\dfrac{\mu_z^{\dot\alpha}y^{P}}{1+y^{\dot\alpha}} & -\mu_z^{v}-\dfrac{\mu_z^{\dot\alpha}y^{v}}{1+y^{\dot\alpha}} & \mu_z^{\alpha}-\dfrac{\mu_z^{\dot\alpha}y^{\alpha}}{(1+y^{\dot\alpha})v_*} & -\mu_z^{\omega_z}+\dfrac{\mu_z^{\dot\alpha}y^{\omega_z}}{1+y^{\dot\alpha}}\\
0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

北航学报 赠阅

5 控制器参数优化设计

在本文中,基于 LADRC 的无人机纵向高度控制器共有 8 个待整定参数,如果通过人工试凑的方法来确定,不仅工作量比较大,而且并不一定能保证找到最优的参数。利用粒子群优化算法(PSO)进行控制器参数的寻优设计,可以提高设计效率和控制性能。粒子群优化算法[13-15]是由 Eberhart 和 Kennedy 在对鸟群捕食行为进行研究的基础上,提出的一种新型进化算法,其具有实现容易、精度高和收敛快等优点,在解决实际问题中展示出了优越性。

采用粒子群优化算法进行参数寻优设计,其关键是给出目标适应度函数。根据经典控制理论,控制律以大气扰动为输入、航迹高度为输出的闭环放大倍数在区间[0 ~ 0.2] rad/s 上越小,则其对相应频率区间扰动的抗扰能力越强,控制器的控制精度也就越高。为了兼顾控制器的快速性和稳态性能,以高度阶跃响应的上升时间 $t_r = 7$ s,无稳态误差等指标安排过渡过程,要求幅值裕度 GM ≥ 10 dB,相角裕度 PM ≥ 45°,构建粒子群优化算法适应度函数。

$$\begin{cases} J_{\min} = \int_0^t X_1 \mathrm{d}\tau + \int_0^t X_2 \mathrm{d}\tau + \mathrm{PM} + \mathrm{GM} \\ X_1 = k_1 \tau \mid e_H(\tau) \mid \\ X_2 = k_2 \tau \mid A_{wx} \mid + k_2 \tau \mid A_{wy} \mid \\ \text{若 PM} \geqslant 60°, \text{则 PM} = 0 \\ \text{若 GM} \geqslant 10\ \text{dB}, \text{则 GM} = 0 \end{cases} \tag{12}$$

式中:$J_{\min}$ 为目标函数;t 为时间;τ 为积分步长;A_{wx} 和 A_{wy} 分别为区间[0 ~ 0.2] rad/s 上的误差闭环放大倍数;k_1、k_2 根据各项的重要程度进行选取,在本文中取 $k_1 = 1$,$k_2 = 2$。

在该适应度函数中,第 1 项 $\int_0^t X_1 \mathrm{d}\tau$ 的物理意义在于保证高度响应与给定过渡过程间具有最小的跟踪误差;第 2 项 $\int_0^t X_2 \mathrm{d}\tau$ 的物理意义在于使特定频率范围的风干扰到高度输出的闭环放大倍数最小,提高对风干扰的抑制能力;最后两项 PM + GM 用以兼顾控制器的稳定裕度,相当于设立“惩罚”机制,若稳定裕度不满足幅值裕度 GM ≥ 10 dB,相角裕度 PM ≥ 60°,则在适应度计算结果中会直接叠加上响应的稳定裕度数值,从而排除不满足所要求稳定裕度的控制器参数。

设定粒子群优化算法的最大迭代次数为 100,粒子群种群大小为 40,变量维数为 8,惯性因子 $w = 0.9$,学习因子 $c_1 = 2$,社会因子 $c_2 = 2$,约束因子 $r_1 = r_2 = 1$。适应度变化曲线如图 5 所示。

从图 5 中可以看出,算法在第 25 次迭代时收敛到最优值,最终得到的控制器参数为 $\omega_H = 1.12$,$\omega_\vartheta = 1.27$,$\omega_V = 3.8$,$K_{H\mathrm{p}} = 3.34$,$K_{H\mathrm{d}} = 6.76$,$K_{V\mathrm{p}} = 0.32$,$K_\vartheta = 1.39$,$K_{\omega_z} = 0.82$,LADRC 的高度阶跃响应曲线和伯德图分别如图 6 和图 7 所示。

从图 7 中可以看出高度阶跃响应无超调,上升时间 $t_r \approx 7$ s,幅值裕度 GM ≈ 10.7 dB,相角裕度 PM ≈ 60.4°,完全满足指标要求。

采用 C 语言与 MATLAB 语言混合编程,用 C 语言实现MATLAB中耗时较长的插值函数,降低

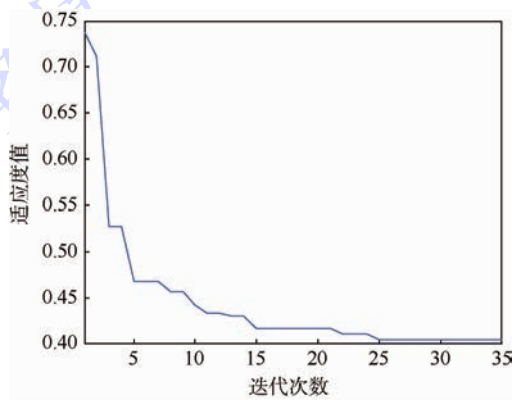


图 5 适应度变化曲线

Fig. 5 Changing curve of fitness

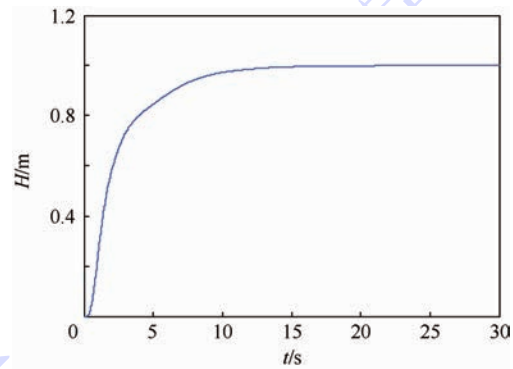


图 6 LADRC 高度阶跃响应曲线

Fig. 6 Height step response curve of LADRC

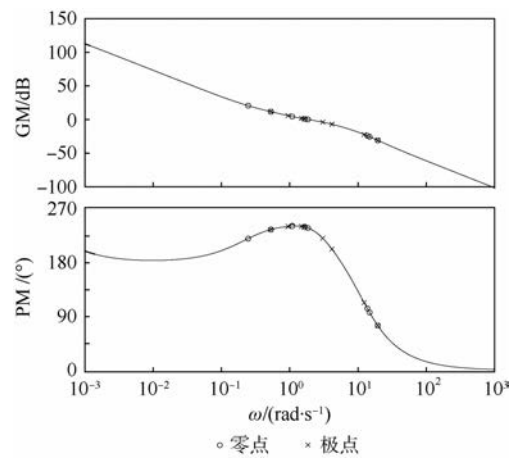


图 7 LADRC 伯德图

Fig. 7 Bode diagram of LADRC

仿真过程中无人机全量模型的解算时间，从而缩短粒子群优化算法参数优化设计的仿真时间。经过多次仿真，结果表明该算法均可快速收敛，且在配置为Core i7 处理器，3G 内存，32 位 Windows 7 操作系统的电脑下，控制器参数优化设计的平均时间消耗约为 40 min，完全可以满足参数离线设计的要求。

6　仿真分析

在第 2 节控制器的基础上，分别从 ESO 观测扰动效果、在复杂气流扰动作用下的 LADRC 效果和气动参数摄动下的 LADRC 效果 3 个方面，验证 LADRC 的控制性能。

6.1　ESO 观测扰动效果仿真结果分析

在俯仰角内回路、高度外回路和速度回路中分别对俯仰角加速度、高度和空速上叠加扰动量 $w(t)=0.5\sin(0.5t)$，LADRC 中相应 ESO 的观测效果如图 8 所示，ϑ''为俯仰角加速度。

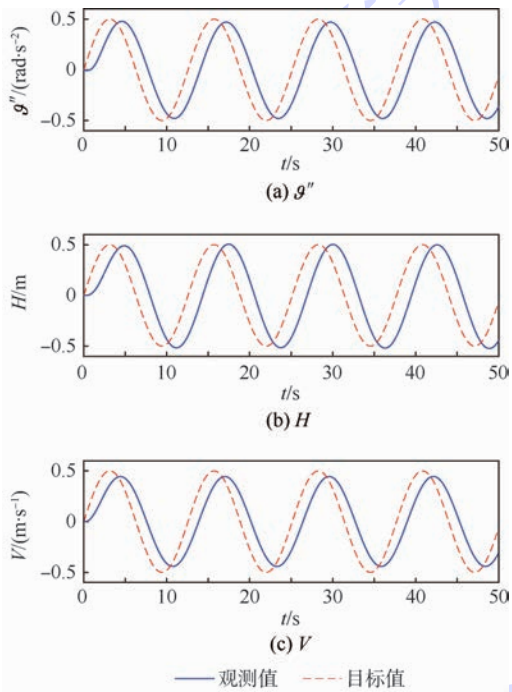


图 8　ESO 观测结果

Fig. 8　Observed results of ESO

由图 8 可知，3 个回路的 ESO 扰动跟踪整体效果较好，其中扰动幅值观测值和实际值基本相同，相位观测值与实际值间的偏差均在可接受的范围内。良好的 ESO 对干扰的观测跟踪效果，为 LADRC 抵抗扰动实现高精度控制奠定了基础。

6.2　复杂气流扰动仿真结果分析

为了说明在复杂气流扰动作用下 LADRC 的控制性能，分别引入离散突风和大气紊流 2 种风扰动，在 $H=8$ m，$V=25$ m/s 状态下，将 LADRC 的高度控制效果与 PID 的高度控制效果进行对比，其中 PID 控制器在无风干扰情况下的单位阶跃高度控制响应与 LADRC 基本相同。

在离散突风（半波长 1 - cos 模型）干扰仿真中，在 0 s 给定 8 ~ 9 m 的高度阶跃指令，2 种控制以相似的动态过程将无人机控制在 9 m 高度。在 20 s 后分别引入幅值为 4 m/s 的离散垂直突风扰动和离散顺向突风扰动，无人机相关运动情况如图 9 和图 10 所示，升降舵和油门响应以及各 LADRC 的控制补偿量如图 11 和图 12 所示。图中：$\Delta\vartheta_g$ 为高度通道补偿量，其实际为对俯仰角给定值的补偿；$\Delta\delta_z$ 为俯仰角通道补偿量，其实际为对升降舵舵偏角的补偿量；$\Delta\delta_p$ 为速度通道补偿量，其实际为对油门的补偿量。

从图 12 仿真结果中可看出 LADRC 的控制性能无论在受扰幅度或恢复基准值的快速性上，都明显优于 PID 控制器。以顺向突风为例，对仿真结果进行分析。

若无人机在飞行过程中受到顺向离散突风扰动，无人机空速减小导致升力减小，航迹高度降低，为了保持给定空速和高度，升降舵将上偏使飞机抬头，油门增加使空速增大。对于 LADRC 控制器，由于各通道 ESO 对扰动的观测作用，俯仰角给定值、升降舵舵偏角和油门会在控制中叠加补偿量，PID 控制器相比，相应控制量均较大，因此高度和空速均可以更快恢复到基准值，从而提高控制律对风扰动的抑制作用，增强了抗干扰能力。垂直突风结果分析与此类似，此处不再赘述。

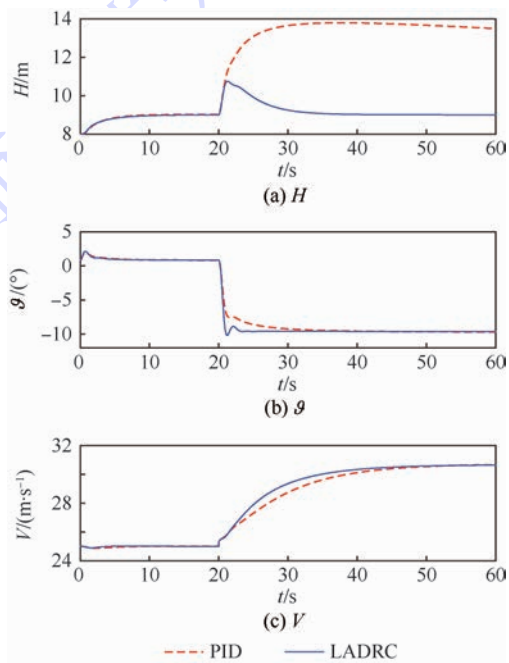


图 9　离散垂直突风下高度控制仿真结果

Fig. 9　Simulation results of height control under discrete vertical wind gust condition

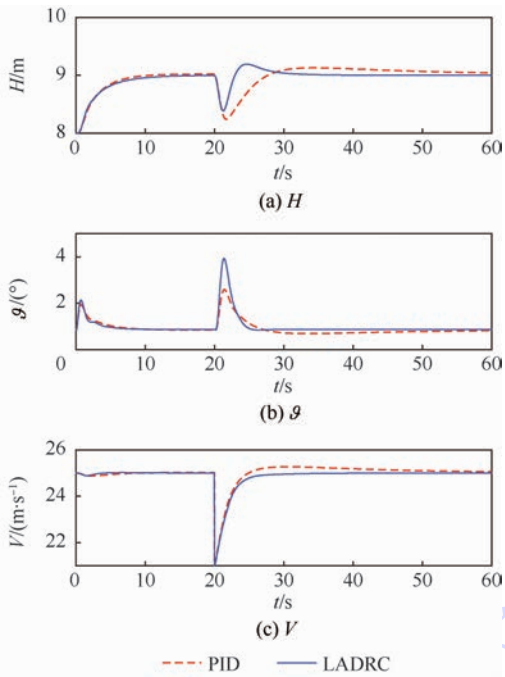


图 10 离散顺向突风下高度控制仿真结果

Fig. 10 Simulation results of height control under discrete forward wind gust condition

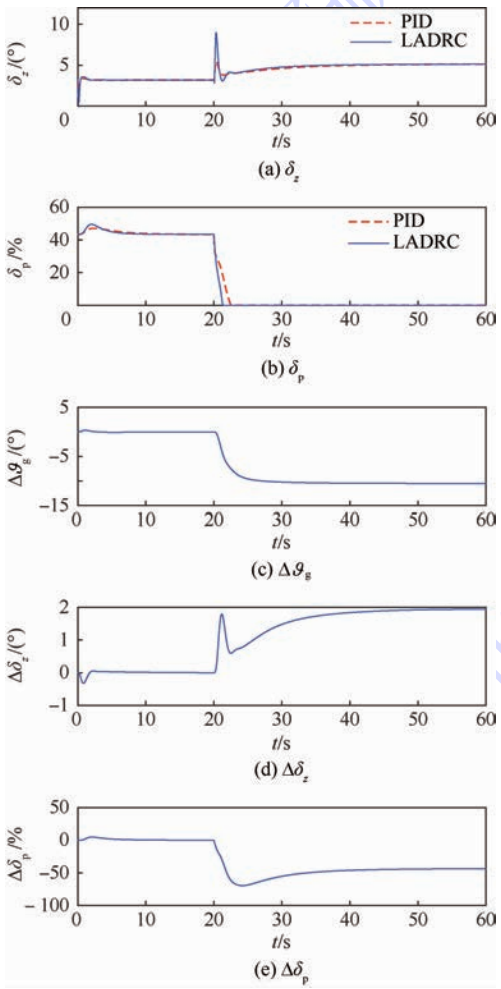


图 11 离散垂直突风下控制量与补偿量仿真结果

Fig. 11 Simulation results of control and compensating quantities under discrete vertical wind gust condition

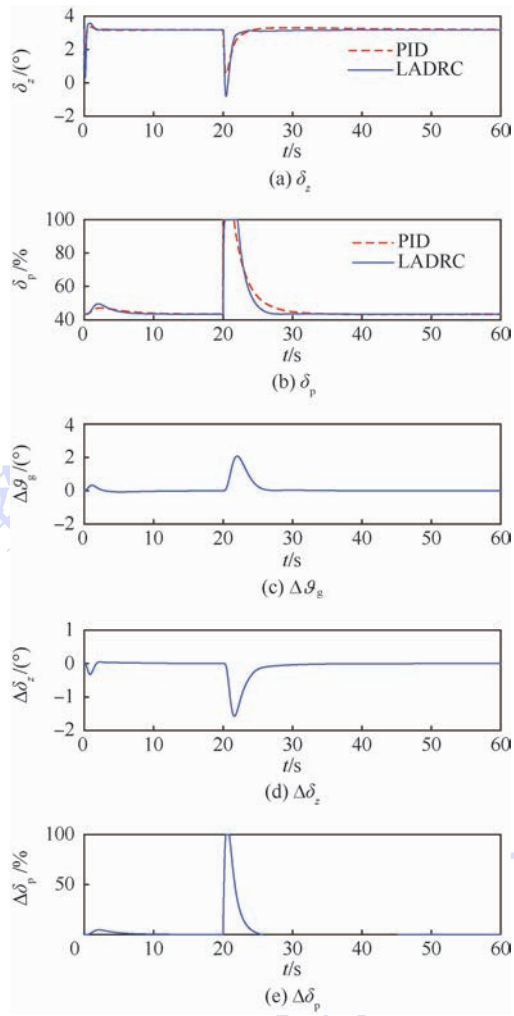


图 12 离散顺向突风下控制量与补偿量仿真结果

Fig. 12 Simulation results of control and compensating quantities under discrete forward wind gust condition

在大气紊流(德莱顿模型)干扰仿真中,假定无人机在 $H=8$ m, $V=25$ m/s 状态下作定高平飞,引入大气紊流扰动,升降舵和油门控制量以及各控制通道补偿量如图 13 所示,PID 控制器和 LADRC 的高度、俯仰角和空速控制效果对比如图 14 所示。

从图 13 中可以看出 LADRC 下航迹高度和空速在大气紊流扰动作用下的波动幅度明显小于 PID 控制器,实现了高精度高度控制,其本质同上述分析相同,由 ESO 观测风扰动,并在控制中引入控制补偿以减小风扰动的影响。

6.3 气动参数摄动仿真结果分析

对于低速小型无人机而言,气动参数的不确定性对控制器的控制性能也有着较大的影响。为了说明气动参数摄动下 LADRC 的控制性能,引入气动参数摄动,将升力对迎角的导数 C_y^{α} 和升降舵舵效 $m_z^{\delta_z}$ 减小 40%,俯仰力矩对迎角的导数 m_z^{α} 增大 40%,PID 控制与 LADRC 下的高度阶跃响应结果对比如图 15 所示。

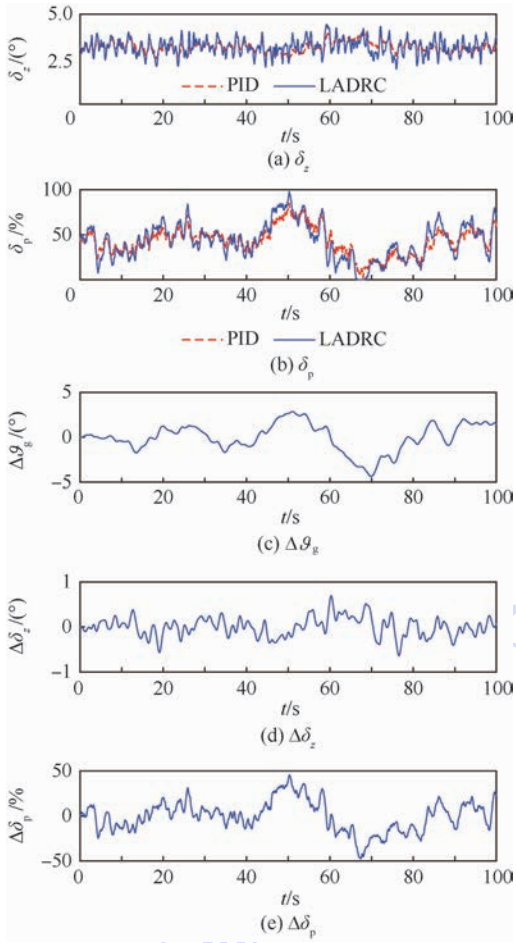


图 13　大气紊流作用下控制量与补偿量仿真结果

Fig. 13　Simulation results of control and compensating quantities under wind turbulence condition

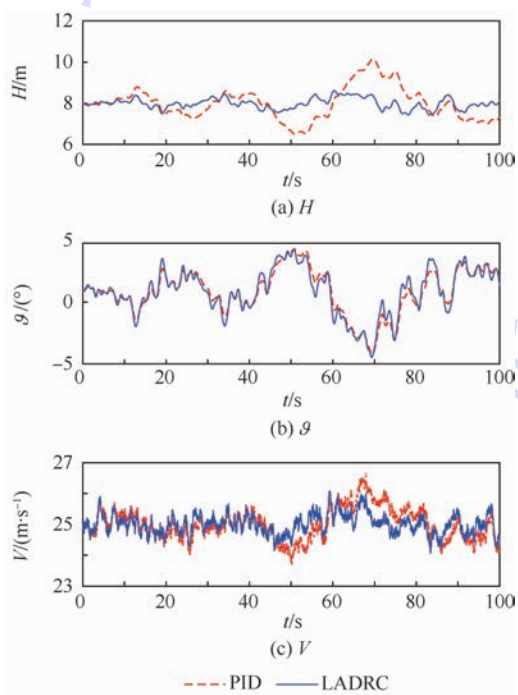


图 14　大气紊流作用下高度控制仿真结果

Fig. 14　Simulation results of height control under wind turbulence condition

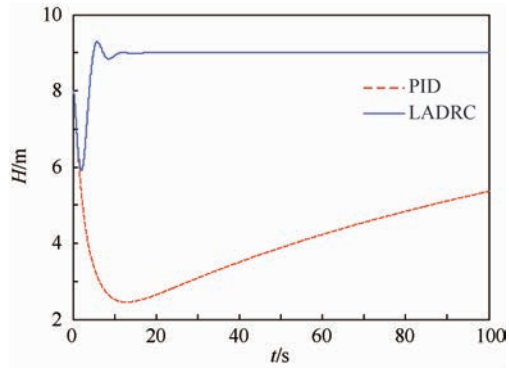


图 15　气动参数摄动仿真结果

Fig. 15　Simulation results of aerodynamic parameter perturbation

从图 15 仿真结果可以看出，PID 控制器高度响应需要很长的时间才能达到给定值，而 LADRC 则可以很快达到给定值，LADRC 较 PID 控制在存在气动参数摄动的情况下具有更好的控制效果。其本质是将气动参数摄动归结为系统的内部扰动，通过 ESO 估计出这一部分扰动，在控制量中叠加补偿，减小气动参数摄动的影响。

7　结　论

本文基于 LADRC 理论，实现了存在复杂气流扰动和气动参数摄动情况下的高精度高度控制，主要结论如下：

1）在控制律设计过程中考虑风干扰的功率谱分布，可以减小控制器设计的保守性。

2）仿真结果表明在复杂气流扰动情况下，与 PID 控制器相比较，LADRC 具有更好的高度控制精度和抗干扰性能。

本文以“天钩”方式回收为背景，对其他应用场合的相关控制也具有参考价值。

参考文献（References）

[1] 孙林峰，马晓平，吴佳凯. 无人机绳钩回收仿真研究[J]. 科学技术与工程，2012，12(7)：1572-1575.
SUN L F，MA X P，WU J K. Simulation research of rope-hook recovery of unmanned aerial vehicle[J]. Science Technology and Engineering，2012，12(7)：1572-1575(in Chinese).

[2] DONG I Y，JUNG Y D，SUNG W C. A guidance and control law design for precision automatic take-off and landing of fixed-wing UAVs：AIAA-2012-4674 [R]. Reston：AIAA，2012.

[3] CHEOLKEUN H. Gain-scheduled directional guidance controller design using a genetic algorithm for automatic precision landing[J]. International Journal of Control Automation and Systems，2010，8(1)：107-117.

[4] 张建宏，张平. 无人机自主精确着陆控制律设计及仿真研究[J]. 系统仿真学报，2009，21(3)：743-748.
ZHANG J H，ZHANG P. Autonomous precise landing control

北航学报 赠阅

law for UAV[J]. Journal of System Simulation, 2009, 21(3): 743-748(in Chinese).

[5] 熊治国,孙秀霞,胡孟权,等. 自抗扰控制器的飞机纵向运动控制律设计[J]. 电机与控制学报,2006,10(4):420-430.
XIONG Z G, SUN X X, HU M Q, et al. Control law design of aircraft longitudinal movement based on active disturbances rejection controller[J]. Electric Machines and Control, 2006, 10(4):420-430(in Chinese).

[6] 韩京清. 控制理论-模型论还是控制论[J]. 系统科学与数学,1998,9(4):328-335.
HAN J Q. Control theory-model or control[J]. Journal of System Science and Mathmetic, 1998, 9(4):328-335(in Chinese).

[7] 韩京清. 自抗扰控制技术-估计补偿不确定因素的控制技术[M]. 北京:国防工业出版社,2009:255-261.
HAN J Q. Active disturbance rejection control-control technique for estimating and compensating uncertain factor[M]. Beijing: National Defense Industry Press, 2009:255-261(in Chinese).

[8] GAO Z Q. Scaling and bandwidth-parameterization based controller tuning[C]// Proceedings of the American Control Conference. Piscataway, NJ: IEEE Press, 2003:4989-4996.

[9] GAO Z Q, HU S H, JIANG F J. A novel motion control design approach based on active disturbance rejection[C]// Proceedings of the 40th IEEE Conference on Decision and Control. Piscataway, NJ: IEEE Press, 2001:1547-1552.

[10] 肖业伦,金长江. 大气扰动中的飞行原理[M]. 北京:国防工业出版社,1993:11-14.
XIAO Y L, JIN C J. Flight principles in wind disturbances[M]. Beijing: National Defense Industry Press, 1993:11-14(in Chinese).

[11] ROBINSON P A, REID L D. Modeling of turbulence and downbursts for flight simulators[J]. Journal of Aircraft, 1990, 27(8):700-707.

[12] 肖业伦. 飞行器运动方程[M]. 北京:航空工业出版社,1987:93-94.
XIAO Y L. Motion equations of aerial vehicles[M]. Beijing: Aeronautics Industry Press, 1987:93-94(in Chinese).

[13] SHI Y H. Particle swarm optimization[J]. IEEE Connections, 2004, 2(1):8-13.

[14] JACA F S, ALBERT A G. A study of global optimizaition using particle swarms[J]. Journal of Global Optimization, 2005, 31(1):93-108.

[15] CHATTERJEE S P. Nonlinear inertia weight variation for dynamic adaptation in particle swarm optimization[J]. Computers & Operations Research, 2006, 33(3):859-871.

作者简介:

齐鹏远 男,硕士研究生。主要研究方向:无人机飞行控制。
E-mail: bttre@sina.com

王勇 男,博士,研究员,博士生导师。主要研究方向:先进控制理论及飞行控制工程应用。
Tel.: 010-82317544
E-mail: wy_buaa@sina.com

Precise height control for UAV based on LADRC

QI Pengyuan[1], WANG Yong[1,*], ZHANG Daibing[2]

(1. School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;
2. School of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha 410073, China)

Abstract: Aimed at solving the impact of sophisticated wind disturbance, research on the control strategy, control structure and parameter optimization design is developed in this paper to realize height control with high precision performance in the longitudinal channel of unmanned aerial vehicle (UAV). First, based on the features of linear active disturbance rejection control (LADRC), the whole control structure is determined, and the probable disturbance from both inside and outside of the system can be estimated using the extended state observer (ESO). As a result, influence of disturbance on the output of the system can be eliminated by introducing this estimated value into the control law. Second, the power spectrum density of the wind turbulence and discrete gust are also analyzed, and a comprehensive fitness function considering impact of wind disturbance on height, characteristics of time domain response and stability margin is constructed, so that controller parameters with high precision and high disturbance-rejection performance can be produced through the particle swarm optimization algorithm to reduce the conservatism of controller parameter design. Finally, the outstanding performance of linear active disturbance rejection controller in the longitude channel of UAV is illustrated through being compared with the regular proportinal-integral-derivative (PID) height controller.

Key words: "sky-hook" recovery; precise height control; linear active disturbance rejection control (LADRC); wind turbulence analysis; particle swarm optimization algorithm

Received: 2015-11-05; **Accepted**: 2016-01-22; **Published online**: 2016-03-15 14:34
URL: www.cnki.net/kcms/detail/11.2625.V.20160315.1434.006.html

* **Corresponding author**. Tel.: 010-82317544 E-mail: wy_buaa@sina.com

北京航空航天大学学报

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0688

CICQ 结构中逼近 work-conserving 的分组调度算法

张元昊，熊庆旭*
（北京航空航天大学 电子信息工程学院，北京 100083）

摘 要：联合输入交叉点排队(CICQ)结构的分组调度算法是一个得到了充分研究的领域，但已有算法在吞吐率和分组平均时延方面与输出排队(OQ)的结果相比，依然不够令人满意，其关键在于 OQ 交换机可以工作于 work-conserving 状态。不同于已有的研究，本文提出了以使得交换机最大程度工作于 work-conserving 状态为目标的新的研究思路，给出并证明了 CICQ 交换机实现 work-conserving 状态的充分且必要条件。以此为基础，提出了一种新的 CICQ 输入调度的算法即交叉缓存队列均衡(CQB)算法，并将经典的最大队长优先(LQF)算法用于输出调度，结合得到 CQB-LQF 算法。仿真结果显示，与经典的及最新的 CICQ 分组调度算法相比，CQB-LQF 算法显著提高了吞吐率及分组平均时延性能。

关 键 词：分组交换；联合输入交叉点排队(CICQ)；work-conserving；调度算法；时延性能

中图分类号：TP393

文献标识码：A **文章编号**：1001-5965(2016)11-2481-07

互联网技术的发展使得以 Internet Protocol (IP)为核心的分组交换网络对传输速度和效率的需求随之加大。而交换机的吞吐率与时延性能作为 Best effort 服务的重要指标，直接影响着整个网络的效率和速度。为高性能交换机设计良好的交换结构和调度算法，以提高吞吐率与时延性能，成为一个重要的问题。

输出排队(Output Queueing, OQ)结构由于其 work-conserving 的特性，具有良好的吞吐率和时延性能，但由于其需要与端口数 N 等量的加速比，可扩展性不足。而输入排队(Input Queueing, IQ)结构的加速比为 1，具有良好的可扩展性[1]，但存在信头 (Head-of-Line, HoL)堵塞问题[2]。虚拟输出排队(Virtual Output Queuing, VOQ)结构[3]克服了 HoL 堵塞问题[4]，但 VOQ 在调度控制过程中较为复杂，根本原因是输入输出竞争的密切耦合。联合输入交叉点排队(Combined Input-Crosspoint-Queued, CICQ)结构[5]通过在所有交叉节点(crosspoint)上配置一定容量(单个或多个信元)的缓存(crossbuffer)来解耦输入竞争和输出竞争，使得分组调度更有效地进行。本文讨论交叉缓存容量为单个信元(cell)的情况。

到目前为止，人们对 CICQ 结构的调度进行了大量研究，经典算法有输入端采用 Round-Robin (RR)或其改进算法的 RR-RR[6]、Differential Round-Robin (DRR)[7]、Tracking Fair Quota Allocation (TFQA)[8]及 Round-Robin with Longest Queue Detecting (RR-LQD)[9]等；以及以队长、交叉缓存占用率、阻塞时间为权重的最大权重匹配法例如 Longest Queue First and RR (LQF-RR)[10]、Most Critical Buffer First (MCBF)[11]、Shortest Crosspoint Buffer First (SCBF)[12]、the Shortest

收稿日期：2015-10-23；**录用日期**：2016-01-22；**网络出版时间**：2016-04-01 17:07
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160401.1707.007.html
基金项目：国家自然科学基金 (61271196)
* **通讯作者**：Tel.：010-82338616 E-mail：qxxiong@buaa.edu.cn

引用格式：张元昊，熊庆旭. CICQ 结构中逼近 work-conserving 的分组调度算法[J]. 北京航空航天大学学报，2016，42(11)：2481-2487. ZHANG Y H, XIONG Q X. Work-conserving-based packet scheduling algorithm for CICQ switches [J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(11): 2481-2487 (in Chinese).

北航学报 赠阅

Buffer First and the Greatest Weigh buffer First (SBF-GWF)[13]、Hybrid Optimization Packet Scheduling(HOPS)[14]和 Most Critical Queue First (MCQF)[15]等。RR 及其改进算法实现简单,复杂度一般为$O(1)$,且大多数算法对可接入业务的吞吐率可接近 100%,但分组平均时延较大。最大权重匹配法以提高复杂度的代价来获取分组时延的下降。例如代表性的 MCBF 算法以特定输出和输入端口对应的 crossbar 中缓存的分组数目为权重,充分利用 crossbar 资源,但没有考虑 VOQ 的队列长度。而近来提出的 SBF-GWF 算法,在 MCBF 的基础上,进一步地,输出调度采用 VOQ 队长为权重,相比 MCBF 及其他算法,时延性能有显著的提高。

在已有的 CICQ 调度算法的研究中,即使是性能突出的 SBF-GWF 算法,相比采用最简单的 First In First Out(FIFO)算法的 OQ 结构,其时延性能依然有较大差距。其根本原因是 OQ 结构能够使交换机工作于 work-conserving 状态,实现 100% 的吞吐率,从而使得时延下降。基于这个思路,不同于已有的算法,本文以实现或最大程度使得交换机工作于 work-conserving 状态为核心,研究提出了 CICQ 结构中的一种新算法,即交叉缓存队列均衡(Crossbuffer Queue Balance, CQB)算法,尽量均衡所有输出端口的 crossbuffer 分组占用。仿真结果显示其时延性能优于 SBF-GWF 算法。

首先,本文通过证明给出了 CICQ 结构 crossbuffer 容量为单个 cell 条件下可能实现 work-conserving 的前提,以及充分必要条件,并以此为基础,具体给出了 CQB 调度算法;其次,给出了 CQB 算法性能的数值仿真结果,并和现有的代表性算法进行了分析比较;最后,进行了小结并给出了今后研究工作。

1 Work-conserving 的充要条件

图 1 给出了 CICQ 交换结构。在本文讨论中,假设在同一个时隙,输入调度和输出调度同时进行,那么当前时隙通过输入调度进入 crossbar 的分组,最早只能在下个时隙进行输出调度。本文将一次的输入调度和其下个时隙的输出调度合称为一次调度。在 CICQ 结构中 work-conserving 的含义为:在任意一次调度中,若输入端有去往某个输出端口的分组,那么输出调度中该端口必有分组被调度离开交换机。本节将通过提出一些定理并给出证明,讨论CICQ结构中,在输入调度前后

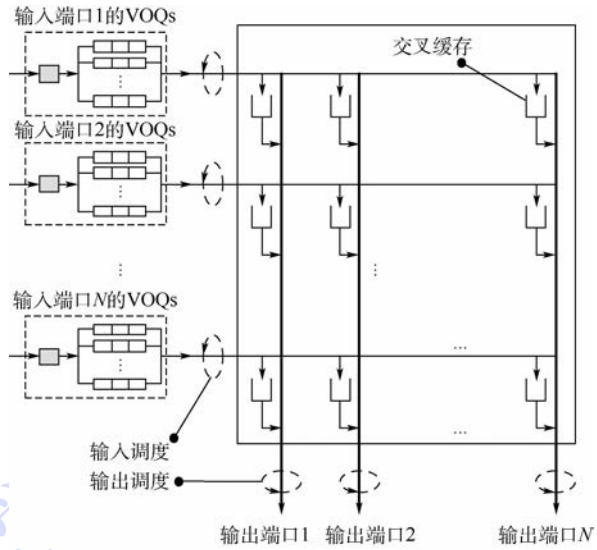


图 1 CICQ 交换结构

Fig. 1 CICQ switch architecture

VOQ 与 crossbar 的状态与 work-conserving 的关系。

符号说明:V_{ij}为输入端口 i 和输出端口 j 对应的 VOQ;N 为交换机端口数;C_{ij}为输入端口 i 和输出端口 j 对应的交叉缓存(crossbuffer);f_{ij}为 C_{ij}的状态标识,若 C_{ij}为空,则 $f_{ij}=0$,否则,$f_{ij}=1$;B_j 为所有去往输出端口 j 的 crossbuffer 中排队分组的队列长度之和,即 $B_j=\sum_{i=1}^{N} f_{ij}$;矩阵 $\boldsymbol{P}_{N\times N}=\{p_{ij}\}_{N\times N}$,若 V_{ij}不为空且 C_{ij}为空,则 $p_{ij}=1$,否则 $p_{ij}=0$;E 为输出端口 j 的集合,$E=\{1\leqslant j\leqslant N\mid$一次调度的开始时刻,VOQ 中有去往输出端口 j 的分组,且 $B_j=0\}$;M 为集合 E 中的端口数目;$\boldsymbol{Q}_{N\times M}$为在矩阵 $\boldsymbol{P}$ 中取集合 E 中的全部输出端口对应的 M 列,组成新的矩阵 $\boldsymbol{Q}_{N\times M}$。

定理 1 在一次调度中,可以实现 work-conserving 的充要条件是,在输入调度后、输出调度前,任一 $j\in E$,都有 $B_j>0$。

证明 充分性。对于该次调度没有分组传出的任一输出端口 k,由 work-conserving 的定义可知,则在输入调度后、输出调度前必有 $B_k=0$。因为在输入调度后、输出调度前,对所有的 $j\in E$,都有 $B_j>0$,而 $B_k=0$,则 $k\notin E$。由集合 E 的条件可知,在该次调度的开始时刻(即输入调度前),由 $k\notin E$可以说明输出端口 k 至少不满足以下 2 个条件之一:①输入端 VOQ 中有去往输出端口 k 的分组。②$B_k=0$。

由于输入调度后 $B_k=0$,而输入调度不会使 B_k 减小,又 B_k 恒非负,则该次调度开始时刻也有 $B_k=0$。即输出端口 k 满足上述条件②,则必然不满足条件①。

即在该次调度中对于没有分组传出的任一输

北航学报 赠阅

出端口 k,证得在输入端没有去往该端口的分组。此为 work-conserving 的逆否命题,则该次调度可以实现 work-conserving,充分性得证。

必要性。若在输入调度后、输出调度前,$\exists j \in E$,有 $B_j = 0$,那么本次调度输出端口 j 将没有分组传出。

而又因为 $j \in E$,可知 VOQ 中有去往输出端口 j 的分组,但本次调度中输出端口 j 却无分组传出,则不满足 work-conserving 定义,即本次调度是非 work-conserving 的。必要性得证。　**证毕**

定理 1 给出了实现 work-conserving,输入调度所要实现的目标,即对于有分组传输需求的输出端口,需要在输入调度后,其所对应的交叉节点缓存队列非空。例如在一次调度开始时刻,输入端 VOQ 和 crossbar 的状态分别为

$$\begin{cases} [V_{ij}]_{4\times4} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ [C_{ij}]_{4\times4} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{cases} \tag{1}$$

式中:0 代表 V_{ij} 或 C_{ij} 为空;1 代表非空。

此时只有 $B_4 = 1$ 非零,$E = \{2,3\}$。输出端口 1 由于无传输需求,故不会影响 work-conserving 的实现与否;而输出端口 4 由于在输入调度前就有 $B_4 > 0$,输入调度不会使 B_j 减小,则在输出调度前也必有 $B_4 > 0$,可以保证本时隙有分组传出,也不会对实现 work-conserving 造成影响。故影响 work-conserving 实现的关键就在于集合 E 中的输出端口 2 和 3,所以输入调度的目标是使得 B_2 和 B_3 在输出调度前均大于零,即定理 1 所阐述的 work-conserving 的充要条件。

而是否存在一种输入调度满足以上条件,这取决于输入调度前 VOQ 与 crossbar 的状态。

定理 2　在一次调度中,时隙开始时刻存在一种输入调度可以完成 work-conserving 的充要条件是,矩阵 $\boldsymbol{Q}$ 对应的二分图最大匹配数可以达到 M。

证明　充分性。若矩阵 $\boldsymbol{Q}$ 的列数为 M,其对应二分图最大匹配数也可达到 M,那么任取一种最大匹配,则该匹配矩阵 $\boldsymbol{U}_{N\times M}$ 每一列均有一个元素为 1。

由于矩阵 $\boldsymbol{Q}$ 中的 M 列与集合 E 中的 M 个输出端口一一对应,则按照上述矩阵 $\boldsymbol{Q}$ 的最大匹配矩阵 $\boldsymbol{U}$ 进行输入调度,即调度匹配矩阵中对应位置元素为 1 的 VOQ 中的分组进入相应的 crossbuffer,完成之后,对所有的 $j \in E$,都有 $B_j = 1 > 0$。

则根据定理 1,该次调度可以实现 work-conserving。

必要性。若存在一种输入调度可以完成 work-conserving,取其调度矩阵中集合 E 中的输出端口对应的 M 列组成新的矩阵 $\boldsymbol{S}_{N\times M}$。

对于所有 $j \in E$,有输入调度前 $B_j = 0$,由定理 1,有输入调度后 $B_j > 0$,则 $\boldsymbol{S}$ 中任一列均至少有一个元素非零为 1,将这些每列中的一个非零元素取出组成新的矩阵 $\boldsymbol{U}_{N\times M}$,则其即为矩阵 $\boldsymbol{Q}$ 的匹配数为 M 的最大匹配矩阵。　**证毕**

定理 2 给出了通过合适的输入调度实现 work-conserving 的前提条件。这说明在理论上 work-conserving 并不能仅通过输入调度算法 100% 实现。例如图 2 所示状态,有 $M = 2$,$\boldsymbol{P}_{4\times4} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$,$\boldsymbol{Q}_{4\times2} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$,矩阵 $\boldsymbol{Q}$ 的最大匹配数为 1,小于 M,则任意输入调度都无法完成 work-conserving。

现就一次调度考虑,在输出调度完成后,对输出端口给出如下定义:

non-work-conserving 端口为 VOQ 队列中有去往该输出端口的分组,但本次调度无分组传出的输出端口;work-conserving 端口为 non-work-conserving 端口的补集。

推论 1　在一次调度中,令通过调度能达到的最大 work-conserving 端口数为 L,矩阵 $\boldsymbol{Q}$ 对应的二分图最大匹配数为 K,则有

$$L = N - M + K \tag{2}$$

证明　最大 work-conserving 端口数为 L,则最小 non-work-conserving 端口数为 $N - L$,由定理 2 可得 $N - L = M - K$,则有 $L = N - M + K$。　**证毕**

由定理 2 可以看出,存在类似于图 2 所示的某些业务状态,使得无论通过何种输入调度,均无法在该时隙实现 work-conserving。

因此要尽可能实现交换机的 work-conserving 状态,调度算法除了要保证当前时隙实现 work-conserving 外,还需要尽量保证后续时隙不出现上述的不可能实现 work-conserving 的情况。

北航学报 赠阅

2 CQB 算法

根据定理 1,可以看出在当前时隙内,是否可以实现 work-conserving 的关键在于集合 E 中的输出端口所对应的 crossbuffer 队列,这些队列在输入调度前为空,而实现 work-conserving 的充要条件就是在输入调度后将其均变为非空。故在输入调度时,优先处理这些空的输出 crossbuffer 队列,为其寻找合适的输入端口使其建立匹配,并将相应分组调度进来。之所以称其为 CQB 算法,就是因为该算法优先处理 B_j 最小(一般为 0)的输出 crossbuffer 队列,通过调度使其 B_j 提高,然后继续这个过程直至调度结束。至于输入端口的选择,根据传输需求来进行,传输需求说明如式(3)所示。当为输出端口 1 的 crossbuffer 队列选择匹配时,输入端口 1 和输入端口 2 均可供选择,但由于输入端口 2 只有去往输出端口 1 的分组,而输入端口 1 除此之外还有去往输出端口 2 的分组,则认为对于输出端口 1,输入端口 2 的传输需求比输入端口 1 大,故选择输入端口 2 与输出端口 1 进行匹配。

$$\begin{cases} [V_{ij}]_{2\times2}=\begin{bmatrix}1 & 1\\ 1 & 0\end{bmatrix} \\ [C_{ij}]_{2\times2}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix} \end{cases} \tag{3}$$

需要说明的是,CQB 算法是一种输入调度的算法。在输出端,本文采用权重匹配的方式,如最大队长优先(Longest Queue First,LQF)、Oldest Cell First(OCF)等,并根据输出调度的算法不同,将整个算法称为 CQB-LQF 算法、CQB-OCF 算法等。

本文具体讨论 CQB-LQF 算法。方便起见,算法中的符号含义如下:t_i 为输入端口 i 包含的非空 VOQ 队列数目;R_j 为满足 V_{ij} 不为空且 C_{ij} 为空的所有输入端口的集合;O 为输出端口的集合;I 为输入端口的集合;W_j 为 R_j 与 I 的交集;p_i 为输入端口的优先级指针;p_o 为输出端口的优先级指针。

调度开始时令 p_i 指向输入端口 1,p_o 指向输出端口 1。CQB-LQF 算法的具体步骤如下。

输入调度算法:

1) 每个时隙开始时,令 O 包含所有输出端口,I 包含所有输入端口。

2) 如果 O 为空,则该时隙输入竞争裁决结束。

3) 从 p_o 指向的输出端口开始,在 O 中选择第一个 B_j 最小的输出端口 j,并将 p_o 指向其下一个输出端口的位置。

4) 如果 W_j 为空,从 O 中剔除输出端口 j,回到第 2)步。

5) 从 p_i 指向的输入端口开始,在 W_j 中选择第一个 t_i 最小的输入端口 i,并将 p_i 指向其下一个输入端口的位置,将 V_{ij} 的头信元发送到 C_{ij} 中。

6) 将 B_j 加 1,从 I 中剔除输入端口 i,更新 R_j 和所有的 W,回到第 3)步。

输出调度算法:

每个时隙中,对于每个输出端口 j,选择满足 C_{ij} 不为空且 V_{ij} 队长最大的 i,将 C_{ij} 中的分组调度输出。同样仿照输入调度使用优先级指针裁决 VOQ 队长相等的情况。

3 性能仿真与比较

文献[13]将 SBF-GWF 算法与一些经典的算法(*i*SLIP、RR-RR、LQF-RR、OCF-OCF、MCBF 等)做了平均时延的仿真比较,可以看出在各种不同的业务以及不同的负载下,SBF-GWF 算法的时延性能均优于其他经典算法。在本节,将用本文提出的 CQB-LQF 算法,与 SBF-GWF 算法同样在各种业务各个负载下进行仿真比较,并以 OQ 结构 FIFO 算法的结果作为参照。比较的性能除了吞吐率和平均时延外,还涉及到一个新的性能参数:non-work-conserving 比率:

$$R_{\mathrm{nwc}}=\frac{T_{\mathrm{nwc}}}{T_{\mathrm{S}}} \tag{4}$$

式中:T_{nwc} 为 non-work-conserving 的时隙数;T_{S} 为总仿真时隙数。R_{nwc} 可以表征逼近 work-conserving 的程度,以此来观察 work-conserving 对交换机性能的影响。

本文仿真时间为 50 万个时隙,交换机端口数为 32 × 32。其中 ON-OFF 业务的平均突发长度(mean burst length)为 120。仿真结果如下。

3.1 去向分布均匀业务的结果

由于在低负载下各算法的吞吐率基本为 1,故观察归一化负载(下文简称“负载”)较高(0.90、0.95、0.99)时的吞吐率。

为区别吞吐率的细小差别,吞吐率数据以表格的形式呈现。均匀业务的吞吐率数据如表 1、表 2 所示,可以看出在均匀业务下,2 种算法的吞吐率几乎一致,并极接近于 OQ。

图 2 和图 3 分别给出了 2 种均匀业务下的分组平均时延的对比结果。其中 l 为归一化负载,d 为平均时延/时隙。可以看出在均匀业务下,2 种算法的性能几乎相同,并且均与 OQ 基本一致,进一步改进幅度相当有限。

北航学报 赠阅

表 1　均匀 Bernoulli 业务吞吐率

Table 1　Throughput under uniform Bernoulli traffic

算法	业务吞吐率		
	$l=0.90$	$l=0.95$	$l=0.99$
SBF-GWF	0.999 989 6	0.999 980 8	0.999 782 9
CQB-LQF	0.999 989 6	0.999 980 8	0.999 782 9
OQ	0.999 991 7	0.999 982 8	0.999 785 0

表 2　均匀 ON-OFF 业务吞吐率

Table 2　Throughput under uniform ON-OFF traffic

算法	业务吞吐率		
	$l=0.90$	$l=0.95$	$l=0.99$
SBF-GWF	0.997 505 7	0.986 224 8	0.975 294 1
CQB-LQF	0.997 505 5	0.986 229 5	0.975 294 4
OQ	0.997 508 2	0.986 232 2	0.975 298 2

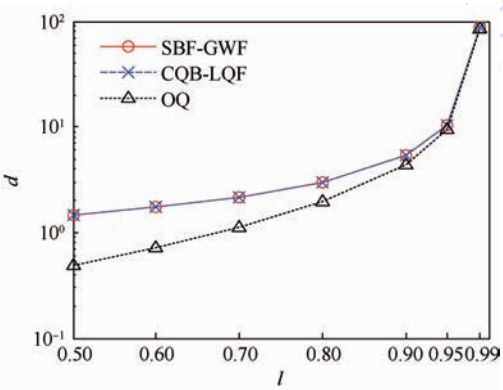


图 2　均匀 Bernoulli 业务平均时延

Fig. 2　Average delay under uniform Bernoulli traffic

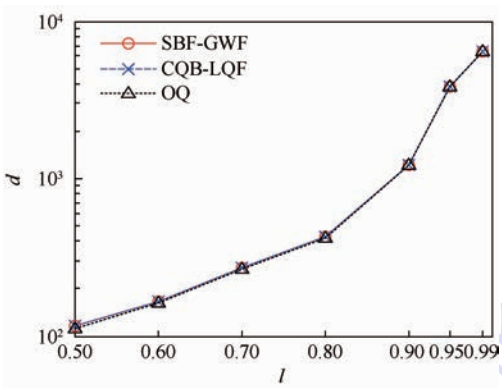


图 3　均匀 ON-OFF 业务平均时延

Fig. 3　Average delay under uniform ON-OFF traffic

3.2　非均匀去向分布业务的结果

非均匀业务采用经典的 Diagonal 模型,其表达式为

$$\rho_{ij}=\begin{cases}\frac{2}{3}\rho & i=j\\ \frac{1}{3}\rho & i=(j-1)\bmod N\\ 0 & \text{其他}\end{cases} \tag{5}$$

式中:ρ_{ij}为 V_{ij}的负载;ρ 为交换机输入负载。

非均匀业务的吞吐率如表 3、表 4 所示,可以看出在与 OQ 的接近程度上,CQB-LQF 算法明显优于 SBF-GWF。

图 4、图 5 给出了在非均匀业务高负载条件下分组平均时延的结果对比。从中可以看出,CQB-LQF 算法的平均时延也明显优于 SBF-GWF 算法,并且与 OQ 的接近程度几乎是 SBF-GWF 算法的 2 倍。

CQB-LQF 在非均匀业务下吞吐率和平均时延性能的优势,体现了追求尽可能逼近 work-conserving 这一算法设计思路的作用。这一点可以由图 6、图 7 看出,在大部分情况下,尤其是负载较高时,CQB-LQF 算法的 R_{nwc} 值显著低于 SBF-GWF 算法,即 CQB-LQF 算法更逼近 work-conserving。

对于图 7 所示的非均匀 ON-OFF 业务负载较低的情况下,CQB-LQF 算法与 SBF-GWF 相比,致力于实现 work-conserving,但 R_{nwc} 反而较高的现象,其原因可以由定理2进行解释。定理2表明

表 3　非均匀 Bernoulli 业务吞吐率

Table 3　Throughput under unbalanced Bernoulli traffic

算法	业务吞吐率		
	$l=0.90$	$l=0.95$	$l=0.99$
SBF-GWF	0.999 993 3	0.999 986 6	0.999 920 3
CQB-LQF	0.999 993 5	0.999 989 8	0.999 942 7
OQ	0.999 996 0	0.999 992 8	0.999 958 8

表 4　非均匀 ON-OFF 业务吞吐率

Table 4　Throughput under unbalanced ON-OFF traffic

算法	业务吞吐率		
	$l=0.90$	$l=0.95$	$l=0.99$
SBF-GWF	0.998 737 2	0.996 882 0	0.991 328 3
CQB-LQF	0.998 817 8	0.997 234 8	0.992 578 5
OQ	0.999 052 4	0.997 835 2	0.995 108 1

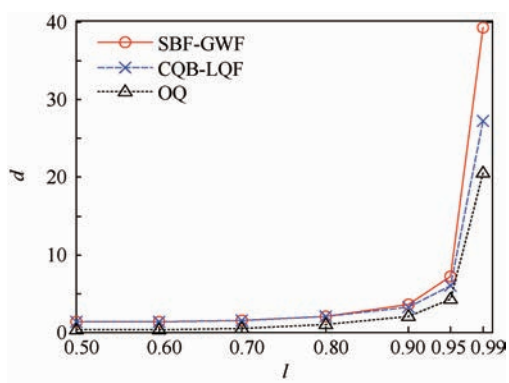


图 4　非均匀 Bernoulli 业务平均时延

Fig. 4　Average delay under unbalanced Bernoulli traffic

北航学报 赠阅

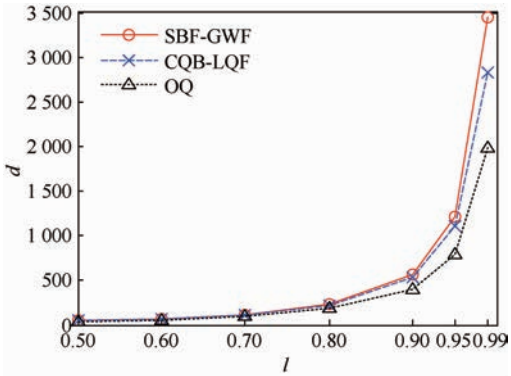


图 5 非均匀 ON-OFF 业务平均时延

Fig. 5 Average delay under unbalanced ON-OFF traffic

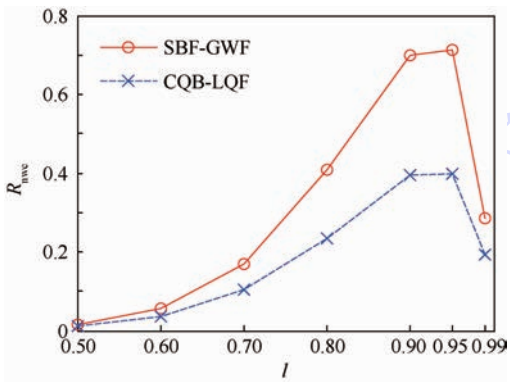


图 6 非均匀 Bernoulli 业务 R_{nwc}

Fig. 6 R_{nwc} under unbalanced Bernoulli traffic

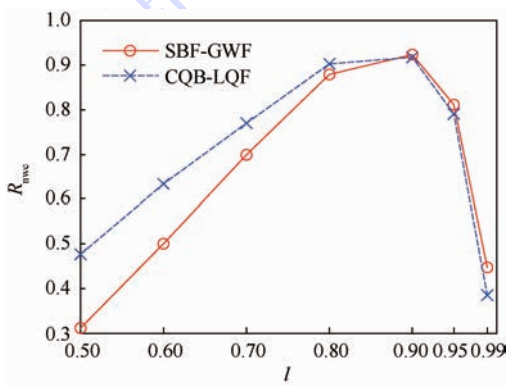


图 7 非均匀 ON-OFF 业务 R_{nwc}

Fig. 7 R_{nwc} under unbalanced ON-OFF traffic

了输入调度算法实现 work-conserving 有前提条件,这个前提条件的满足与否受 VOQ 的状态影响,更进一步受业务到达的影响。而 CQB-LQF 算法着重考虑当前时隙的 work-conserving 实现,没有更多地考虑业务到达情况,这样做是考虑了算法的简单。而 ON-OFF 业务和 Bernoulli 业务相比,业务到达的因素对调度的影响前者更明显,对调度算法的要求更大,使得出现了上述现象。这一点由图 5 中 ON-OFF 业务和图 6 中 Bernoulli 业务两种算法对比结果的差异可以说明。

上述现象为今后对 CQB 算法进一步的改进提出了方向,即可以加入一些关于业务特性和状态的考虑,以进一步提升 work-conserving 实现比率和时延性能。但在高速网络中算法的简单也是同样重要的,同时 CQB-LQF 算法的时延性能比 SBF-GWF 更好,因此可以说 CQB-LQF 算法是一种性能更优的调度算法。而 work-conserving 实现比率与负载大小、业务状态等因素更详细具体的关系,有待于进一步地深入研究。

4 结 论

1) 不同于已有基于轮询以及以队长或时延等为权重实现最大权重匹配的研究,本文以交换机最大程度工作于 work-conserving 状态为目标,讨论并证明交叉节点容量为单个 cell 的 CICQ 结构实现 work-conserving 的充分且必要条件。

2) 以上述条件为基础,研究提出了一种新的输入调度算法——CQB 算法。通过仿真与现有研究中性能突出的 SBF-GWF 算法进行对比,结果显示在均匀业务中,CQB-LQF 算法与 SBF-GWF 性能相近,而在非均匀业务中,CQB-LQF 算法性能更优。

由于 CQB 算法是一种输入调度算法,在未来的研究中也可进一步研究相应的输出调度算法,使之更为逼近 work-conserving。

参考文献 (References)

[1] 熊庆旭. 输入排队结构交换机分组调度研究[J]. 通信学报,2005,26(6):118-129.
XIONG Q X. Research on packet scheduling in input-queued switches[J]. Journal on Communications, 2005, 26(6): 118-129(in Chinese).

[2] KAROL M J, HLUCHYJ M G, MORGAN S P. Input versus output queueing on a space-division packet switch[J]. IEEE Transactions on Communications, 1987, 35(12): 1347-1356.

[3] NONG G, HAMDI M. On the provision of quality-of-service guarantees for input queued switches[J]. IEEE Communications Magazine, 2000, 38(12): 62-69.

[4] MCKEOWN N, MEKKITTIKUL A, ANANTHARAM V, et al. Achieving 100% throughput in an input-queued switch[J]. IEEE Transactions on Communications, 1999, 47(8): 1260-1267.

[5] NABESHIMA M. Performance evaluation of a combined input- and crosspoint-queued switch[J]. IEICE Transactions on Communications, 2000, 83(3): 737-741.

[6] ROJAS-CESSA R, OKI E, JING Z, et al. CIXB-1: Combined input-one-cell-crosspoint buffered switch[C]// 2001 IEEE Workshop on High Performance Switching and Routing. Piscataway, NJ: IEEE Press, 2001: 324-329.

[7] LUO J, LEE Y, WU J. DRR a fast high-throughput scheduling

北航学报 赠阅

algorithm for combined input crosspoint-queued cicq switches [C]// IEEE 20th IEEE International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems. Piscataway, NJ: IEEE Press, 2005: 329-332.

[8] HUA N, WANG P, JIN D, et al. Simple and fair scheduling algorithm for combined input-crosspoint-queued switch [C]// IEEE International Conference on Communications, ICC' 2007. Piscataway, NJ: IEEE Press, 2007: 6305-6310.

[9] YUN Z, PENG L, ZHAO W, et al. RR-LQD: A novel scheduling algorithm for CICQ switching fabrics [C]// Proceedings of the 15th Asia-Pacific Conference on Communications, APCC 09. Piscataway, NJ: IEEE Press, 2009: 846-849.

[10] JAVIDI T, MAGILL R, HRABIK T. A high-throughput scheduling algorithm for a buffered crossbar switch fabric [C]// IEEE International Conference on Connmunications, ICC' 2001. Piscataway, NJ: IEEE Press, 2001: 1586-1591.

[11] MHAMDI L, HAMDI M. CBF: A high-performance scheduling algorithm for buffered crossbar switches [C]// Workshop on High Performance Switching and Routing, 2003, HPSR. Piscataway, NJ: IEEE Press, 2003: 67-72.

[12] ZHANG X, BHUYAN L N. An efficient scheduling algorithm for combined input-crosspoint-queued (CICQ) switches [C]// IEEE Global Telecommunications Conference, 2004. GLOBECOM'04. Piscataway, NJ: IEEE Press, 2004, 2: 1168-1173.

[13] GAO Z, ZENG H, XIA Y, et al. SBF-GWF scheduling for combined input-crosspoint-queued (CICQ) switches [C]// 2011 6th International Conference on Computer Sciences and Convergence Information Technoloy (ICCIT). Piscataway, NJ: IEEE Press, 2011: 404-408.

[14] 高志江,曾华燊,申志军. 混合优化的 CICQ 交换结构调度算法[J]. 计算机应用,2012,32(7):1791-1795.
GAO Z J, ZENG H Y, SHEN Z J. Hybrid optimization packet scheduling algorithm for CICQ switches [J]. Journal of Computer Applications, 2012, 32(7): 1791-1795 (in Chinese).

[15] WANG X T, WANG Y W, LI S C, et al. A novel high performance scheduling algorithm for crosspoint buffered crossbar switches [C]// International Conference on Computer Information Systems and Industrial Applications. Paris: Atlantis Press, 2015: 59-62.

作者简介:

张元昊　男,硕士研究生。主要研究方向:卫星交换机。
E-mail: zyh_demon@163.com

熊庆旭　男,博士,教授,博士生导师。主要研究方向:通信网络、无线通信等。
Tel.: 010-82338616
E-mail: qxxiong@buaa.edu.cn

Work-conserving-based packet scheduling algorithm for CICQ switches

ZHANG Yuanhao, XIONG Qingxu*

(School of Electronic and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: Although packet scheduling for combined input-crosspoint-queued (CICQ) switches has been extensively investigated, the performance in terms of throughput and average delay of packets is not satisfied, especially compared to the results obtained by employing output queued architecture. The key reason is that output queneing (OQ) switch can operate in work-conserving state. Different from the popular approaches appeared in the literatures, our work focuses on scheduling packet so that the switch works approximately in work-conserving state to the largest extent. First, the sufficient and necessary conditions for the switch with CICQ architecture to work in work-conserving state are investigated and proved. Then, a new CICQ input scheduling algorithm called crossbuffer queue balance (CQB) is put forward, which is combined with the classical longest queue first (LQF) scheduling used for output arbiter to obtain CQB-LQF algorithm. Simulation results show that the performance in terms of throughput and average packet delay obtained by CQB-LQF is much better than the classical and popular algorithms.

Key words: packet switching; combined input-crosspoint-queued (CICQ); work-conserving; scheduling algorithm; delay performance

Received: 2015-10-23; **Accepted**: 2016-01-22; **Published online**: 2016-04-01 17:07
URL: www.cnki.net/kcms/detail/11.2625.V.20160401.1707.007.html
Foundation item: National Natural Science Foundation of China (61271196)
* **Corresponding author**. Tel.: 010-82338616　E-mail: qxxiong@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0760

含泡沫镍中间层陶瓷/不锈钢接头微观组织和性能

张羽，朱颖，郭伟*，齐丹，祁泽武

（北京航空航天大学 机械工程及自动化学院，北京 100083）

摘 要：连接温度850℃，保温时间60 min，用泡沫镍金属作为中间层真空钎焊Al_2O_3陶瓷与1Cr18Ni9Ti不锈钢，利用扫描电子显微镜（SEM）、能谱（EDS）分析钎焊接头的微观组织，利用剪切实验检测接头的力学性能。实验结果表明，不添加中间层时，接头平均剪切强度只有7.7 MPa，断裂位置发生在陶瓷侧；添加泡沫镍中间层时，接头平均剪切强度达到101.7 MPa，断裂位置发生在陶瓷与泡沫镍金属连接界面处。不添加中间层时，Ti元素主要分布在钎料与陶瓷以及钎料与不锈钢反应界面处；添加中间层后钎焊接头中Ti元素主要分布在中间层，与Ni元素形成$TiNi_3$，Ag、Cu和Ti元素分布更加均匀。

关 键 词：泡沫镍；剪切强度；真空钎焊；Al_2O_3陶瓷；1Cr18Ni9Ti不锈钢
中图分类号：TG454
文献标识码：A **文章编号**：1001-5965(2016)11-2488-07

陶瓷材料具有强度高、硬度高、弹性模量大、热胀系数小、密度低、抗腐蚀及耐磨等优良的力学性能和物理化学性能，广泛应用于机械、化工、电子、宇航、能源和光学等领域[1-2]。在其应用过程中，经常遇到金属与陶瓷连接问题，从而实现二者性能上的互补。由于陶瓷和金属之间化学及物理性能差异及冶金不相容性，使得一般焊接方法很难连接2种材料。钎焊是常用的金属与陶瓷的连接方法，目前陶瓷与金属的钎焊主要存在2个问题：一是陶瓷与金属的润湿性问题，主要通过陶瓷金属化或利用活性金属元素加以解决；另一个是陶瓷与金属连接接头中的残余应力问题，在接头连接的加热和冷却过程中，陶瓷、金属各自产生差异较大的膨胀和收缩，在接头界面附近产生较大的热应力，因此，如何降低陶瓷与金属焊接接头的残余应力以及对于残余应力分布的分析俨然成为当前一个非常重要的问题。这一问题可以利用热膨胀系数较低的材料作为中间层来缓解。

为缓解焊接接头残余应力，提高连接强度，一般采用添加中间层的方法[3]。中间层一般选用热膨胀系数相对较低的金属，如Cu、Mo和SiC等材料的添加对于缓解陶瓷金属焊接接头残余应力已经取得了一定的成效[4-6]。泡沫金属[7]也可作为中间层缓解陶瓷与金属之间残余应力。Shirzadi等[8]用不锈钢泡沫作为中间层焊接陶瓷与金属，最大剪切强度达到了33 MPa，热循环性能可以达到60次，相对于不含不锈钢泡沫中间层得到了很大的提升。Park和Eagar[9]选用带有孔洞的镍金属作为中间层焊接Si_3N_4和Inconel 718，在镍金属上添加孔隙，结果显示尽管反应面积降低了，但是剪切强度得到了提升，孔洞的添加对缓解残余应力起到了很大的作用。

泡沫金属由于其独特结构而具有许多优异性能，如具有超轻型、高强度比、吸声、隔声、隔热、散热、减振和吸收冲击能等性能，在航空航天、汽车工业、建筑等领域展现了广泛的应用前景[10-11]。

收稿日期：2015-11-19；**录用日期**：2016-01-28；**网络出版时间**：2016-04-15 16:56
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160415.1656.009.html
基金项目：国家自然科学基金（11372019）
* **通讯作者**：E-mail：gwei@buaa.edu.cn

引用格式：*张羽，朱颖，郭伟，等. 含泡沫镍中间层陶瓷/不锈钢接头微观组织和性能[J]. 北京航空航天大学学报，2016，42(11)：2488-2494. ZHANG Y，ZHU Y，GUO W，et al. Microstructure and properties of ceramic/stainless steel joint with Ni foam interlayer[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2488-2494 (in Chinese).*

北航学报 赠阅

金属泡沫具有的独特网状、无方向性多孔结构使它具有特殊的压缩应力-应变特性,具备了优良的吸能减震性能,可用作能量吸收材料。

本实验选用 0.2 mm 厚的泡沫镍金属作为中间层,采用 Ag-Cu-Ti 钎料来焊接 Al_2O_3 陶瓷与 1Cr18Ni9Ti 不锈钢,同时作为对比,同样参数下不添加中间层焊接 Al_2O_3 陶瓷与 1Cr18Ni9Ti 不锈钢。然后对焊接接头微观组织、元素分布、断面成分以及强度进行研究。

1　试验方法和方案

钎焊实验过程中所采用的母材分别为 95wt% Al_2O_3 陶瓷和 1Cr18Ni9Ti 不锈钢,钎料选用 Ag-26.77Cu-4.4Ti(wt%),Ag-Cu-Ti 钎料对 Al_2O_3 陶瓷与 1Cr18Ni9Ti 不锈钢有很好的润湿性[12-13]。试验所采用的泡沫镍金属材料孔径为 0.1 mm、孔隙率为 90%,其宏观形貌如图 1 所示。

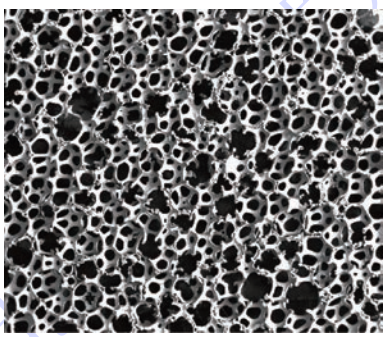


图 1　泡沫镍金属宏观形貌

Fig. 1　Macroscopic morphology of Ni foam

Al_2O_3 陶瓷和 1Cr18Ni9Ti 不锈钢加工成 ϕ10 mm × 5 mm 的试样,实验前用 400#、600#、800# 和 1 200# 水砂纸逐级打磨去除试件表面的氧化膜,并用丙酮溶剂对不锈钢和泡沫镍金属材料进行去油、去污染处理,再用无水乙醇进行脱水处理。试样装配如图 2 所示。Ag-Cu-Ti 钎料厚度为 0.1 mm,泡沫镍金属厚度为 0.2 mm。

钎焊试验采用沈阳真空技术研究所制造的 VBF-113 型真空钎焊炉。试样装配图如图 2 所示。钎焊温度为 850℃,保温时间 60 min,真空度为 1.0×10^{-3} Pa,钎焊温度曲线如图 3 所示。

采用 GLEEBLE-1500 热模拟试验机对钎焊接头进行剪切强度试验测试,剪切速度为0.5 mm/min,每一组做 3 次剪切强度测试,求其平均值,钎焊接头剪切试验示意图如图 4 所示。利用扫描电子显微镜(SEM)(CamScan3400)观察接头微观组织以及通过能谱(EDS)测定其元素组成。

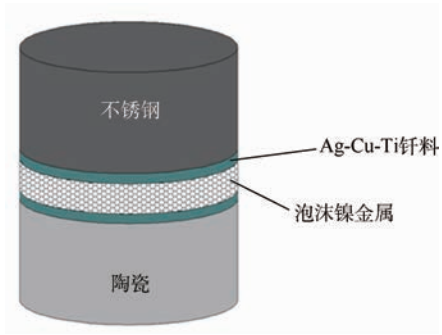


图 2　试样装配图

Fig. 2　Assembly diagram of brazing specimen

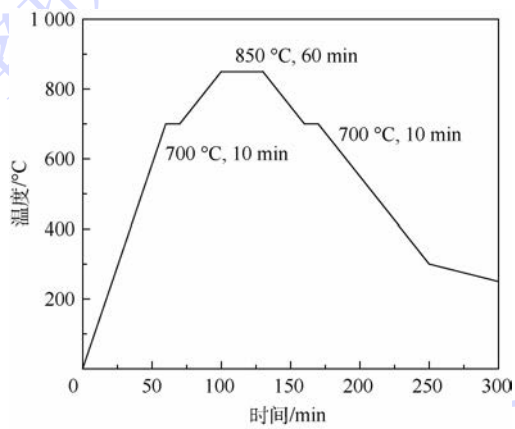


图 3　钎焊温度曲线

Fig. 3　Temperature curve of brazing

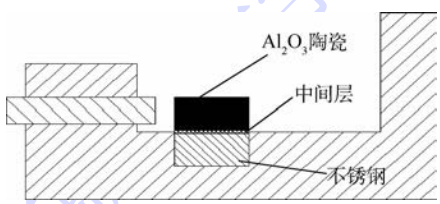


图 4　钎焊接头剪切试验示意图

Fig. 4　Shear test schematic of brazed joint

2　结果及分析

2.1　剪切强度

表 1 为钎焊温度 850℃,保温时间 60 min 条件下,添加和不添加泡沫镍中间层真空钎焊接头剪切强度。

表 1　钎焊接头剪切强度

Table 1　Shear strength of brazed joints

MPa

中间层	剪切强度			平均剪切强度
	组 1	组 2	组 3	
含泡沫镍金属	95.5	120.1	89.5	101.7
不含泡沫镍金属	8.1	6.1	8.8	7.7

北航学报 赠阅

不含泡沫镍中间层平均剪切强度只有 7.7 MPa;当添加 0.2 mm 泡沫镍金属中间层时,接头平均剪切应力达到 101.7 MPa。可以看出,添加泡沫镍金属中间层后,接头平均剪切强度提高了将近 13 倍。

2.2 接头微观组织及结构分析

图 5 为陶瓷/不锈钢焊接接头微观组织。图 5(a)和图 5(c)为钎焊温度 850℃、保温时间 60 min、不含泡沫镍中间层陶瓷/不锈钢焊接接头微观组织。图 5(b)和图 5(d)为钎焊温度 850℃,保温时间 60 min、添加 0.2 mm 泡沫镍金属中间层,陶瓷/不锈钢焊接接头微观组织。图 5(a)可以看出,在 Al_2O_3 陶瓷母材上出现了裂纹,这直接影响了其力学性能,裂纹产生原因主要是在焊接降温过程中,Al_2O_3 陶瓷和 1Cr18Ni9Ti 不锈钢热膨胀系数相差很大,2 种材料膨胀与收缩过程不一致而产生较大的残余应力。图 5(b)可见,当添加泡沫镍金属作为中间层时,陶瓷与不锈钢均未出现裂纹。图 5(c)可见,不添加镍中间层接头界面连接较好,并未出现裂纹等缺陷,中间层包括银白色相和灰色相,相间分布,有助于连接界面的强度提高;图 5(d)因为中间层泡沫镍金属的加入,白色相和灰色相分布更加细化。

为了确定接头各反应区的相组成,对不含泡沫镍中间层及添加泡沫镍中间层的钎焊接头进行了 EDS 分析,表 2 为图 5(c)不含中间层接头各区域的 EDS 结果。图 5(c)中灰色区域 *D* 为 Cu 基固溶体,白色区域 *C* 为 Ag 基固溶体;区域 *A* 以及区域 *E*、区域 *F* 分别为两侧反应界面,含有大量 Ti 元素。Ti 元素主要分布在两侧反应界面处,Ti 元素对于陶瓷与金属的连接起到了很大的作用[12]。

表 3 为图 5(d)含 0.2 mm 泡沫镍接头各区域的 EDS 结果。图 5(d)中区域 *H*、区域 *I* 靠近陶瓷侧主要为 Cu 基固溶体和 Ag 基固溶体;区域 *M*、区域 *N* 靠近金属侧主要富含 Ag、Cu 元素,Ti 元素在两侧反应界面处含量很少;区域 *J*、区域 *K* 主要由 Ti、Ni 元素组成,其比例约为 1∶3,该 2 个区域生产的可能相为 $TiNi_3$[14]。通过以上分析可以看出,不添加泡沫镍中间层时,Ti 元素主要富集在钎料与陶瓷连接界面以及钎料与不锈钢连接界面处;添加中间层后,Ti 主要分布在泡沫镍中间层,与 Ni 元素形成 $TiNi_3$ 化合物。

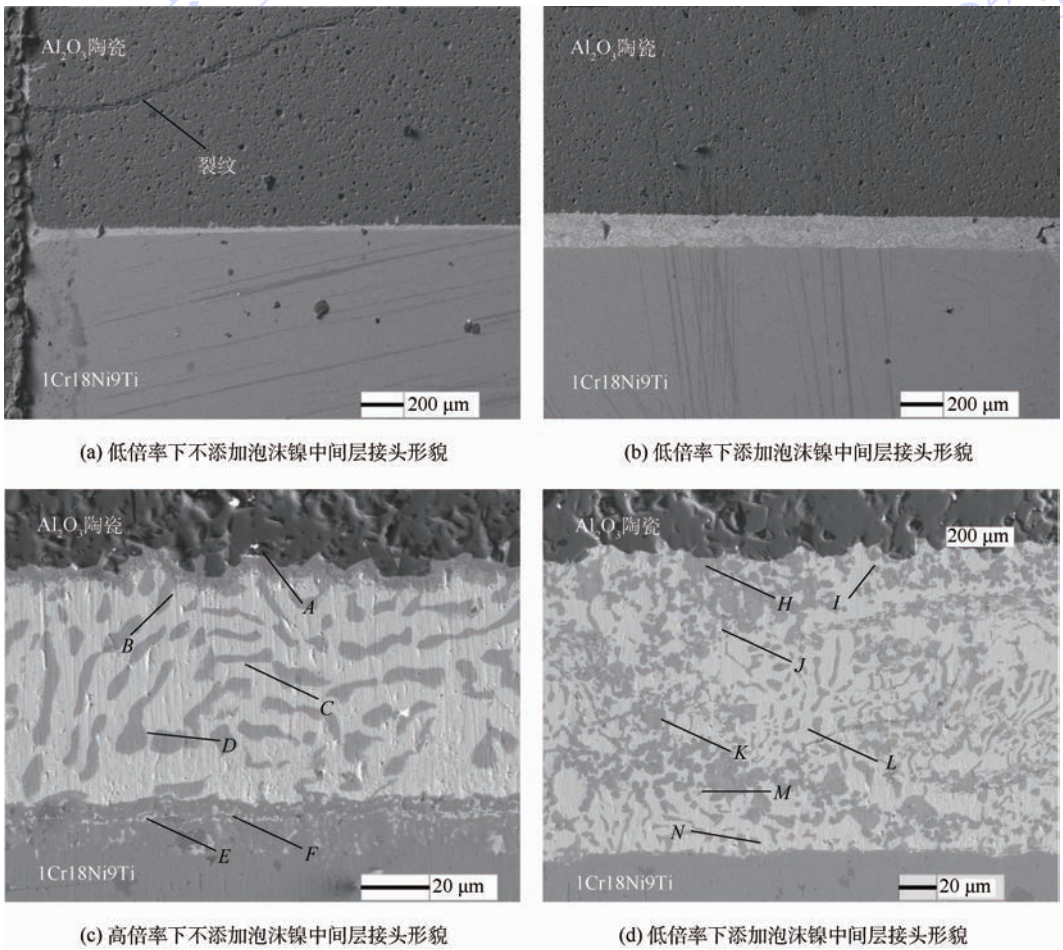


图 5 陶瓷/不锈钢焊接接头微观组织

Fig. 5 Microstructures of brazed ceramic/stainless steel joint

北航学报 赠阅

表 2　不含泡沫镍中间层焊接接头各区域的能谱结果

Table 2　Regional EDS results of brazed joint without Ni foam interlayer

区域	含量/wt%						可能相
	Ag	Cu	Ti	Fe	Al	O	
A	6.12	20.79	47.44	1.72	10.57	10.48	富 Ti
B	11.95	86.25	0.24	0.32	0.13	1.12	富 Cu
C	87.80	11.68	0	0.07	0.58	0	Ag(s, s)
D	2.87	94.50	0.03	0.30	1.24	1.05	Cu(s, s)
E	19.25	1.95	40.36	27.55	0.38	10.51	Ti-Fe-O,Ag
F	9.16	1.27	28.76	47.55	0.43	12.82	Ti-Fe-O

表 3　含泡沫镍中间层焊接接头各区域的能谱结果

Table 3　Regional EDS results of brazed joint with Ni foam interlayer

区域	含量/wt%							可能相
	Ag	Cu	Ti	Fe	Al	Ni	O	
H	1.93	89.11	1.06	0.51	0.20	4.65	2.38	Cu(s, s)
I	86.60	12.20	0	0.31	0.55	0.67	0	Ag(s, s)
J	20.65	6.5	17.65	2.05	0.68	52.53	0	$TiNi_3$
K	0.64	20.60	19.08	2.19	0.45	57.34	0	$TiNi_3$
L	88.51	11.02	0	0.33	0.19	0	0	富 Ag
M	65.83	32.27	0.29	0.21	0.46	0.94	0	富 Ag, Cu
N	77.41	21.89	0	0.87	0.02	0.22	0	富 Ag, Cu

2.3　断裂模式

图 6 为 Al_2O_3 陶瓷与 1Cr18Ni9Ti 不锈钢接头断口宏观形貌和断裂模式。图 6(b)的断口上可以清晰看到钎焊断口表面存在泡沫所特有的网格结构。不含泡沫中间层钎焊接头可以看出断裂位置发生在陶瓷母材，向陶瓷内部凹进，其断裂模式如图 6(c)所示；含泡沫镍中间层，断裂位置发生在泡沫 Ni 中间层与陶瓷结合面处，且不锈钢侧还有一定陶瓷残留，断裂模式如图 6(d)所示。

图 7 为含泡沫镍金属钎焊接头断口微观形貌与 Cu、Ni、Ag 和 Ti 元素分布，区域 *O* 为孔洞，Ag-Cu-Ti 钎料流入并与 Ni 发生反应从而融合。从 EDS 分析观察可得，断面上残余 Ag、Cu 和 Ti 元素，Ti 与 Ni 元素分布相似，形成了化合物 Ti_xNi_y。网格孔洞结构能够大大缓解由于 Al_2O_3 陶瓷和 1Cr18Ni9Ti 不锈钢热膨胀系数的不同而导致的内应力。在钎料熔化阶段，泡沫镍中间层起到了骨架的作用，液态 Ag-Cu-Ti 钎料流入泡沫镍，并与之发生一定反应，使得 Ag-Cu-Ti 三元素都更加均匀的分布在中间层与母材之间，对于提高界面连接强度起到了很大的作用。

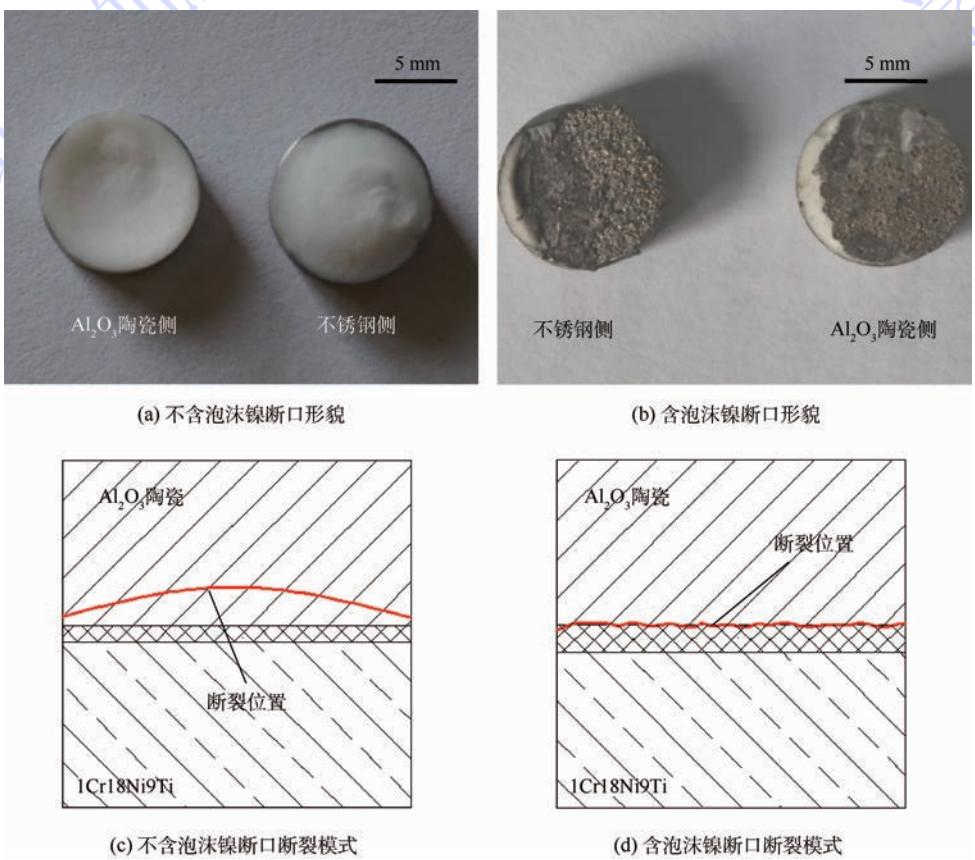


图 6　Al_2O_3 陶瓷与 1Cr18Ni9Ti 不锈钢焊接接头断口形貌和断裂模式

Fig. 6　Fracture morphology and mode of brazed joint of Al_2O_3 ceramic and 1Cr18Ni9Ti stainless steel

北航学报 赠阅

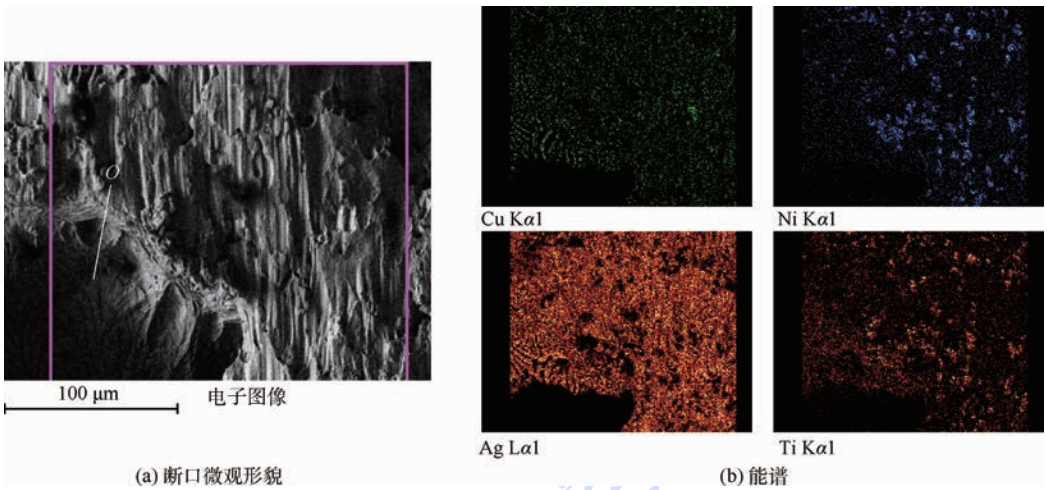


图 7 含泡沫镍金属 Al_2O_3 陶瓷与 1Cr18Ni9Ti 不锈钢焊接接头断口微观形貌和能谱结果

Fig. 7 Fracture microscopic morphology and EDS results of brazed joint of Al_2O_3 ceramic and 1Cr18Ni9Ti stainless steel with Ni foam

2.4 接头残余应力缓解机理

在降温过程中,由于 1Cr18Ni9Ti 不锈钢热膨胀系数较大,Al_2O_3 陶瓷热膨胀系数较小,金属体积收缩较大,而陶瓷体积变化很小,这时在 Al_2O_3 陶瓷靠近焊缝处的表面上容易产生较大的残余应力,这也是不含泡沫镍中间层时,断裂发生在陶瓷母材的原因。当加入金属泡沫镍中间层后,泡沫材料具有良好的塑性变形可以吸收残余应变能,同时镍的热膨胀系数处于 Al_2O_3 陶瓷与 1Cr18Ni9Ti 不锈钢之间,所以在降温冷却过程中,陶瓷与不锈钢的直接连接转变为陶瓷与泡沫镍的连接以及不锈钢与泡沫镍的连接,2 个连接界面处的应力幅值从而降低,因此对于整个接头残余应力随之降低。

陶瓷金属连接弹性简单模型,通常界面应力计算公式[3,15]为

$$\sigma_m = \sigma_c = \frac{E_m E_c}{E_m + E_c}(\alpha_m - \alpha_c)\Delta T \tag{1}$$

式中:σ_m 为不锈钢侧残余应力;σ_c 为陶瓷侧残余应力;E_m 为金属弹性模量;E_c 为陶瓷弹性模量;α_m 为金属热膨胀系数;α_c 为陶瓷热膨胀系数;ΔT 为降温区间。

表 4 为不同温度下材料的弹性模量与热膨胀系数。从式(1)中可以得出残余应力与温度差 ΔT 以及热膨胀系数之差成正比,实验过程保温温度为 850℃,降温至 20℃,ΔT = 830℃,Ag-Cu-Ti 钎料在 850℃ 保温阶段会进入母材以及泡沫镍中间层,暂且不考虑其缓解残余应力的作用,将模型简单化。所以不含泡沫镍中间层时 $\sigma_m = \sigma_c$ = 761 MPa。加泡沫镍金属作为中间层时,可以计算陶瓷/泡沫镍金属连接界面残余应力大概为494 MPa,不锈钢/泡沫镍金属连接界面残余应力大概为 325 MPa。由此可见,当添加泡沫镍金属作为中间层时,连接界面上的残余应力得到了一定的缓解,同时陶瓷界面处的残余应力要大于不锈钢侧的残余应力,这也验证了添加泡沫镍中间层后断裂位置发生在陶瓷连接界面处。

表 4 不同温度下材料的弹性模量与热膨胀系数

Table 4 Coefficients of materials' thermal expansion and elastic modulus under different temperature

材料	温度/℃	弹性模量/GPa	热膨胀系数/($10^{-6}K^{-1}$)
陶瓷	20	380	5.96
	200	380	6.85
	400	370	7.94
	600	360	8.99
	800	350	10.03
不锈钢	20	184	16.6
	200	169	17.0
	400	149	17.5
	600	137	18.2
	800	88	18.6
镍	20	207	13.0

同时,由于泡沫金属本身结构具有减振、吸收冲击能等性能,可以进一步降低连接界面的残余应力。添加泡沫镍金属中间层,钎焊后降温过程中,泡沫镍金属起到骨架的作用,钎料凝固过程中,流入泡沫镍金属,变成了多个熔区独立凝固收缩,相邻熔区之间会产生相反的收缩力,从而相邻小熔区之间的应力得以相消,所以金属镍泡沫的加入,对钎料起到了细化、分流的作用,每个小熔区的应力幅值也相应降低,当所有的小熔区应力相互叠加之后,钎焊接头的残余应力得到了极大的缓解;不含泡沫镍金属时,钎料在凝固时收缩变形,会产生向钎料内部的应力,由于熔区面积较大,所有会产生较大的残余应力。

北航学报 赠阅

3　结　论

1）连接温度 850℃，保温时间 60 min，真空钎焊 Al_2O_3 陶瓷与 1Cr18Ni9Ti 不锈钢，不添加中间层时，剪切强度只有 7.7 MPa；添加泡沫镍中间层时，剪切强度达到 101.7 MPa。接头剪切强度提高了将近 13 倍。

2）泡沫镍金属中间层的添加，使 Ag-Cu-Ti 钎料的元素分布发生很大的变化，不添加泡沫镍中间层时，Ti 元素主要富集在钎料与陶瓷连接界面以及钎料与不锈钢连接界面处；添加中间层后，Ti 主要分布在泡沫镍中间层，与 Ni 元素形成 $TiNi_3$ 化合物，而且 Ag-Cu-Ti 三元素的分布更加均匀。

3）不含泡沫镍中间层时，断裂位置发生在陶瓷母材侧，断口宏观形貌为椭圆形，向陶瓷内部凹进；添加泡沫镍中间层后，断裂位置发生在泡沫镍中间层与陶瓷结合面。

参考文献（References）

[1] 路学成，任莹. 先进结构陶瓷材料的研究进展[J]. 佛山陶瓷，2009(1)：37-43.
LU X C，REN Y. Research progress of advanced structural ceramic[J]. Foshan Ceramic，2009(1)：37-43(in Chinese).

[2] SONG S H，YOON S E，CHOI J，et al. A high-performance ceramic composite anode for protonic ceramic fuel cells based on lanthanum strontium vanadate[J]. International Journal of Hydrogen Energy，2014，39(29)：16534-16540.

[3] 欧孝玺，张福勤，夏莉红，等. 以 Nb 为中间层 AgCuTi 为钎料连接炭/炭复合材料与不锈钢[J]. 粉末冶金材料科学与工程，2011，16(3)：442-447.
OU X X，ZHANG F Q，XIA L H，et al. Joining of carbon/carbon composite with stainless steel using AgCuTi as brazing alloy and Nb interlayer[J]. State Key Laboratory of Powder Metallurgy，2011，16(3)：442-447(in Chinese).

[4] JADOON A K，RALPH B，HORNSBY P R. Metal to ceramic joining via a metallic interlayer bonding technique[J]. Journal of Materials Processing Technology，2004，152(3)：257-265.

[5] HE Y M，ZHANG J，SUN Y，et al. Microstructure and mechanical properties of the Si_3N_4/42CrMo steel joints brazed with Ag-Cu-Ti + Mo composite filler[J]. Journal of the European Ceramic Society，2010，30(15)：3245-3251.

[6] HE Y M，ZHANG J，LIU C F，et al. Microstructure and mechanical properties of Si_3N_4/Si_3N_4 joint brazed with Ag-Cu-Ti + SiCp composite filler[J]. Materials Science and Engineering：A，2010，527(12)：2819-2825.

[7] ZHU Y，QI D，GUO W，et al. The braze joint between Al_2O_3 to 1Cr18Ni9Ti using a nickel foam[J]. Welding in the World，2015，59(4)：491-496.

[8] SHIRZADI A A，ZHU Y，BHADESHIA H K D H. Joining ceramics to metals using metallic foam[J]. Materials Science and Engineering：A，2008，496(1-2)：501-506.

[9] PARK J W，EAGAR T W. Strain energy release in ceramic-to-metal joints with patterned interlayers[J]. Scripta Materialia，2004，50(4)：555-559.

[10] 曹立宏，马颖. 多孔泡沫金属材料的性能及其应用[J]. 甘肃科技，2006，22(6)：117-119.
CAO L H，MA Y. Properties and application of porous metal foam material[J]. Gansu Science and Technology，2006，22(6)：117-119(in Chinese).

[11] 曹国英，王芳，王录才. 泡沫金属的力学性能及研究进展[J]. 铸造设备研究，2008(2)：51-54.
CAO G Y，WANG F，WANG L C. The mechanical property and progress of metal foams[J]. Foundry Equipment Research，2008(2)：51-54(in Chinese).

[12] YUE X，HE P，FENG J C，et al. Microstructure and interfacial reactions of vacuum brazing titanium alloy to stainless steel using an AgCuTi filler metal[J]. Materials Characterization，2008，59(12)：1721-1727.

[13] NOGI K. The role of wettability in metal-ceramic joining[J]. Scripta Materialia，2010，62(12)：945-948.

[14] KUO D H，YEH K T，SHIUE R K，et al. Joining of Al_2O_3 to 316SS using braze-infiltrated Ni net[J]. ISIJ International，2011，1(6)：1017-1019.

[15] 张小勇，吕宏，王林山，等. SiC 陶瓷与金属 Ta 连接的残余应力[J]. 中国稀土学报，2003，21(12)：98-101.
ZHANG X Y，LU H，WANG L S，et al. Calculation of welding residual stress between SiC ceramic and Ta[J]. Journal of the Chinese Rare Earth Society，2003，21(12)：98-101(in Chinese).

作者简介：

张羽　男，硕士。主要研究方向：异种材料连接与数值模拟。
E-mail：13126737586@163.com

郭伟　男，博士，副教授。主要研究方向：特种材料的钎焊扩散连接技术。
Email：gwei@buaa.edu.cn

朱颖　女，博士，副教授。主要研究方向：焊接、钎焊和扩散焊、微连接。
Email：y.june@163.com

北京航空航天大学学报

北航学报 赠阅

Microstructure and properties of ceramic/stainless steel joint with Ni foam interlayer

ZHANG Yu, ZHU Ying, GUO Wei*, QI Dan, QI Zewu

(School of Mechanical Engineering and Automation, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: Al_2O_3 ceramic and 1Cr18Ni9Ti stainless steel were brazed with or without Ni foam interlayer at the temperature of 850°C for 60 min at vacuum. The microstructure of the joint was investigated by scanning electron microscopy (SEM) and energy dispersive spectroscopy (EDS), besides shearing tests were used to test the mechanical properties of the joints. The experimental results show that the average shear strength of the joint without Ni foam interlayer is 7.7 MPa, and the fracture occurs in the ceramic. While Ni foam interlayer is used, the average shear strength of the joint is up to 101.7 MPa, and the fracture is located at the interface of ceramic and Ni foam interlayer. Without Ni foam, Ti is mainly distributed at the interface of ceramic and Ag-Cu-Ti and the interface of steel and Ag-Cu-Ti. With Ni foam, element Ti is distributed at the interlayer zone where $TiNi_3$ is formed during the brazing process, besides, the distribution of Ag, Cu, Ti element is more uniform.

Key words: Ni foam; shear strength; vacuum brazing; Al_2O_3 ceramic; 1Cr18Ni9Ti stainless steel

北京航空航天大学学报

Received: 2015-11-19; **Accepted**: 2016-01-28; **Published online**: 2016-04-15 16:56

URL: www.cnki.net/kcms/detail/11.2625.V.20160415.1656.009.html

Foundation item: National Natural Science Foundation of China (11372019)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2016.0204

微处理器中含噪热传感器位置分布优化方法

李鑫, 李鑫*, 杨金孝, 程元乐

(西北工业大学 电子信息学院, 西安 710072)

摘 要: 高性能处理器普遍集成热传感器,采用动态热管理技术对芯片实施连续热监控。然而,由于实际芯片中的模拟或者数字热传感器不可避免伴随噪声,使动态热管理的可靠性受到很大影响。因此,为了提高热监控的精确性,本文运用主成分分析(PCA)技术对原始热图像样本矩阵进行降维近似处理,并结合矩阵扰动分析提出基于模拟退火算法的热传感器位置分布优化方法。实验结果表明:该方法比现有的贪婪算法在热重构误差、信噪比(SNR)和误警率等性能方面有了一定提高,能够有效运用在动态热管理中实现精确的热监控。

关 键 词: 热传感器; 动态热管理; 热重构; 噪声; 位置分布
中图分类号: V221+.92; TP212
文献标识码: A **文章编号**: 1001-5965(2016)11-2495-06

随着高性能处理器特征尺寸的缩小和性能需求的增加,其功率密度呈指数级增长,导致芯片温度不断升高。近年来,高性能处理器普遍集成热传感器,采用动态热管理(Dynamic Thermal Management, DTM)[1]技术对芯片实施连续热监控。然而,实际芯片中所植入的模拟或者数字热传感器不可避免地伴随有多种噪声,在一定程度上会加剧错误的预警和不必要的响应,使动态热管理的可靠性受到影响[2],给系统性能带来不必要的损失。因此,在有噪声存在的情况下,如何获得精确的芯片热点温度信息对动态热管理技术至关重要。

有效的热传感器位置分布方法可以在使用有限数量热传感器的情况下,估计芯片的所有热点温度。Long 等[3]提出了一种基于网格的插值技术,用于确定最优的热传感器位置。Memik 等[4]提出了一种优化的热传感器位置分布算法,可以在热点温度估计方面获得较高的精度。对于均匀间隔放置的热传感器,采用虚拟插值方法确定最优的传感器位置;对于非均匀间隔放置的热传感器,在二维 k-均值聚类(k-means clustering)算法的基础上,加入热点的温度信息,构造出三维"距离",并进一步引入感知因子,使得优化的传感器位置向温度较高的方向移动。文献[4]的目的是最大化热点温度估计精度,并不考虑全局监控问题,因而根据该方法获得热传感器的位置分布后,进行热分布重构得到的平均温度误差较大。Reda等[5]使用启发式迭代方法将热传感器数量分配问题近似为一个非确定性多项式困难问题(Non-deterministic Polynomial hard, NP-hard),并提出了一种新颖的软测量技术,该技术结合实际热传感器温度读数对没有放置传感器的位置温度进行估计,进一步提高了热点温度估计精度。Li 等[6]针对热传感器位置分布优化问题,提出了3 种简单、有效的热传感器位置分布策略;针对热传感器数量分配问题,提出了一种静态热传感器数量分配技术,能够使用相对较少数量的热传感器监控所有热点温度。

收稿日期: 2016-03-15; **录用日期**: 2016-06-12; **网络出版时间**: 2016-07-01 19:15
网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20160701.1915.001.html
基金项目: 国家自然科学基金 (61501377)
* **通讯作者**: E-mail: xinli@nwpu.edu.cn

引用格式: *李鑫, 李鑫, 杨金孝, 等. 微处理器中含噪热传感器位置分布优化方法[J]. 北京航空航天大学学报, 2016, 42(11): 2495-2500. LI X, LI X, YANG J X, et al. Optimization method of thermal sensor placement for microprocessor with noise[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(11): 2495-2500 (in Chinese).*

北航学报 赠阅

总体而言,目前已有的局部监控相关研究成果尚未有考虑热传感器噪声因素的影响,使其实用性受到了一定的限制。本文在考虑热传感器噪声因素影响下,提出了一种基于模拟退火算法(simulated annealing)的热传感器优化位置分布方法,能够在很大程度上减小噪声的影响,降低热点的误警率和提高热分布重构的性能,对动态热管理的温度监控具有很大的实用意义。

1 热图像降维

主成分分析(Principal Component Analysis, PCA)[7]是一种常用的数据分析方法,其通过线性变换将原始数据表示为一组各维度线性无关的组合,可用于提取数据的主要特征分量,已在信号处理、数据压缩、人脸识别等领域相对成熟。

由于在计算机处理过程中任何连续的变量必须离散化,因此,将整个芯片区域中连续的温度信号离散化为 $L\times W$ 的网格表述,L 和 W 分别为离散热图像的长和宽。假设温度 $t[x,y]$ 为定义在(x,y)坐标上的离散温度,$0\leqslant x\leqslant L-1,0\leqslant y\leqslant W-1$。通过提取热图像的列来获得向量化的热图像$w[i]$,其中 $0\leqslant i\leqslant N-1,N=LW$。提取公式为

$$w[i]=t\left[i \bmod L,\left\lfloor\frac{i}{L}\right\rfloor\right] \tag{1}$$

式中:$\lfloor\ \rfloor$表示取下整函数;mod 为求余计算。

假定有 M 幅热图像,得到向量化热图像集合 $\{\boldsymbol{w}_j\}_{j=0}^{M-1}$,第 j 幅热图像的均值为 $\boldsymbol{\mu}_j=\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{w}_j$。因此,零均值向量集合为 $\boldsymbol{\varphi}=\{\boldsymbol{w}_j-\boldsymbol{\mu}_j\}_{j=0}^{M-1}$,计算协方差矩阵 $\boldsymbol{C}=E[\boldsymbol{\varphi}\ \ \boldsymbol{\varphi}^{\mathrm{T}}]$,$E$ 为数学期望;$\boldsymbol{\varphi}^{\mathrm{T}}$ 为 $\boldsymbol{\varphi}$ 的转置矩阵,利用公式 $\boldsymbol{C}\boldsymbol{\Phi}_i=\lambda_i\boldsymbol{\Phi}_i$,求矩阵 $\boldsymbol{C}$ 的特征值 λ_i 和对应的特征向量 $\boldsymbol{\Phi}_i$。通过上述步骤后可得到 PCA 处理的结果 $\boldsymbol{\Omega}$,其定义为

$$\boldsymbol{\Omega}=\begin{bmatrix}\Phi[0,0] & \cdots & \Phi[N-1,0]\\ \vdots & & \vdots\\ \Phi[0,N-1] & \cdots & \Phi[N-1,N-1]\end{bmatrix}\times\begin{bmatrix}\varphi[0,0] & \cdots & \varphi[0,M]\\ \vdots & & \vdots\\ \varphi[N-1,0] & \cdots & \varphi[N-1,M]\end{bmatrix} \tag{2}$$

将协方差矩阵 $\boldsymbol{C}$ 的特征值从小到大排列:$\lambda_0\geqslant\lambda_1\geqslant\cdots\geqslant\lambda_{N-1}\geqslant 0$,按照特征值所占的能量比例来选取前面 $K(K\ll N)$ 个特征向量:

$$\sum_{i=0}^{K-1}\lambda_i\Big/\sum_{i=0}^{N-1}\lambda_i\geqslant\eta \tag{3}$$

通常取 $\eta=90\%\sim99\%$。在满足均方差最小的条件下,选取所占能量比最大的 K 个特征值所对应的$\{\boldsymbol{\Phi}_i\}_{i=0}^{K-1}$为最佳近似正交向量子空间。降维后近似的热图像集合$\{\hat{\boldsymbol{w}}_j\}_{j=0}^{M-1}$可表示为

$$\{\hat{\boldsymbol{w}}_j\}_{j=0}^{M-1}=\begin{bmatrix}\Phi[0,0] & \cdots & \Phi[0,K-1]\\ \vdots & & \vdots\\ \Phi[N-1,0] & \cdots & \Phi[N-1,K-1]\end{bmatrix}\times\begin{bmatrix}\Omega[0,0] & \cdots & \Omega[0,M]\\ \vdots & & \vdots\\ \Omega[K-1,0] & \cdots & \Omega[K-1,M]\end{bmatrix}+\{\boldsymbol{\mu}_j\}_{j=0}^{M-1}=\boldsymbol{\Phi}_K\boldsymbol{\Omega}_K+\{\boldsymbol{\mu}_j\}_{j=0}^{M-1} \tag{4}$$

这样,每幅热图像都可以投影到 $\boldsymbol{\Phi}_K$ 的子空间中,用来表示热图像的主要特征信息。

2 优化热传感器放置技术

2.1 传感器位置分布理论

假定在坐标$\{(x_n,y_n)\}_{n=1}^{P}$上放 P 个传感器,根据式(1)可得向量化的传感器位置$\{d_i\}_{i=1}^{P}$,由式(4)得传感器测量矩阵$\{\boldsymbol{w}_j^{\mathrm{S}}\}_{j=0}^{M-1}$为

$$\{\boldsymbol{w}_j^{\mathrm{S}}\}_{j=0}^{M-1}=\begin{bmatrix}\Phi[d_1,0] & \cdots & \Phi[d_1,K-1]\\ \vdots & & \vdots\\ \Phi[d_P,0] & \cdots & \Phi[d_P,K-1]\end{bmatrix}\times\begin{bmatrix}\Omega[0,0] & \cdots & \Omega[0,M]\\ \vdots & & \vdots\\ \Omega[K-1,0] & \cdots & \Omega[K-1,M]\end{bmatrix}+\{\boldsymbol{\mu}_j^{\mathrm{S}}\}_{j=0}^{M-1}=\widetilde{\boldsymbol{\Phi}}_K\boldsymbol{\Omega}_K+\{\boldsymbol{\mu}_j^{\mathrm{S}}\}_{j=0}^{M-1} \tag{5}$$

从 $\boldsymbol{\Phi}_K$ 中取出对应的 P 行得到 $P\times K$ 维传感矩阵 $\widetilde{\boldsymbol{\Phi}}_K$,由于$\{\boldsymbol{\mu}_j\}_{j=0}^{M-1}$每行具有相同值,则可以取$\{\boldsymbol{\mu}_j\}_{j=0}^{M-1}$中任意 P 行组成 $P\times M$ 维均值矩阵$\{\boldsymbol{\mu}_j^{\mathrm{S}}\}_{j=0}^{M-1}$。

通常传感器的读数易于受到多种噪声的影响,因此,加入噪声因素,式(5)可以修正为

$$\{\boldsymbol{w}_j^{\mathrm{S}}\}_{j=0}^{M-1}+\boldsymbol{z}=\widetilde{\boldsymbol{\Phi}}_K(\boldsymbol{\Omega}_K+\boldsymbol{\varepsilon})+\{\boldsymbol{\mu}_j^{\mathrm{S}}\}_{j=0}^{M-1} \tag{6}$$

式中:$\boldsymbol{z}$ 为噪声项;$\boldsymbol{\varepsilon}$ 为在降维过程中产生近似误差项 $\boldsymbol{\varepsilon}=\{\hat{\boldsymbol{w}}_j\}_{j=0}^{M-1}-\{\boldsymbol{w}_j\}_{j=0}^{M-1}$。

令 $\boldsymbol{\lambda}^{\mathrm{S}}=\{\boldsymbol{w}_j^{\mathrm{S}}\}_{j=0}^{M-1}-\{\boldsymbol{\mu}_j^{\mathrm{S}}\}_{j=0}^{M-1}$,可以将式(6)变为

$$\boldsymbol{\lambda}^{\mathrm{S}}+\boldsymbol{z}=\widetilde{\boldsymbol{\Phi}}_K(\boldsymbol{\Omega}_K+\boldsymbol{\varepsilon}) \tag{7}$$

分析式(6),可以推出:

$$\frac{\|\boldsymbol{\varepsilon}\|_2}{\|\boldsymbol{\Omega}_K\|_2}\leqslant\|\widetilde{\boldsymbol{\Phi}}_K^{-1}\|_2\|\widetilde{\boldsymbol{\Phi}}_K\|_2\cdot\frac{\|\boldsymbol{z}\|_2}{\|\boldsymbol{\lambda}^{\mathrm{S}}\|_2}=\mathrm{Cond}_2(\widetilde{\boldsymbol{\Phi}}K)\cdot\frac{\|\boldsymbol{z}\|_2}{\|\boldsymbol{\lambda}^{\mathrm{S}}\|_2} \tag{8}$$

式中:$\|\ \ \|_2$ 为二范数;$\mathrm{Cond}_2(\ )$为矩阵二范数条件数。

北航学报 赠阅

根据式(8)可知,条件数表示矩阵计算对于误差的敏感性。因此,在给定噪声的基础上,提高热重构的稳定性和精度必须使 $\widetilde{\boldsymbol{\Phi}}_K$ 的条件数最小。这个结论可以用来确定传感器的最优位置:在矩阵 $\boldsymbol{\Phi}_K$ 中挑选 P 行得到具有最小条件数的传感矩阵 $\widetilde{\boldsymbol{\Phi}}_K$,其相应的行所在位置即为传感器的最优位置,这是一个 NP-hard 问题。

2.2　模拟退火算法放置传感器方法

现代进化算法主要包括模拟退火算法、遗传算法和蚁群算法等,它们有一个共同的目标——求解 NP-hard 组合优化问题的全局最优解。为此,本文采用基于模拟退火算法进行启发式搜索获得最小条件数的传感矩阵 $\widetilde{\boldsymbol{\Phi}}_K$。

模拟退火算法[8]是一种通用概率算法,相对于其他最优化算法,大量实践表明模拟退火算法非常适合求解组合优化问题,在参数合适的情况下基本上可以 100% 得到全局最优解,在一定程度上可以解决 NP-hard 问题。在本节中,提出一种基于模拟退火算法的热传感器优化位置分布方法。在 $\boldsymbol{\Phi}_K$ 中选择 P 行得到有最小的条件数的满秩矩阵 $\widetilde{\boldsymbol{\Phi}}_K$。算法可以描述为:

1) 设置退火时间表,包括初始温度 T、终止温度 T_{stop}、衰减参数 α 和马尔科夫链长度 L_{M}。

2) 从 $\boldsymbol{\Phi}_K$ 矩阵 N 行中随机挑选 2 组 $P(P\geqslant K)$ 行得到传感器位置向量 $\{d_i\}_{i=1}^{P}$ 和 $\{d_i^{\text{c}}\}_{i=1}^{P}$,形成传感矩阵 $\widetilde{\boldsymbol{\Phi}}_K$ 和 $\widetilde{\boldsymbol{\Phi}}_K^{\text{c}}$。如果 $\text{rank}(\widetilde{\boldsymbol{\Phi}}_K)=K$ 和 $\text{rank}(\widetilde{\boldsymbol{\Phi}}_K^{\text{c}})=K$,分别计算 $\widetilde{\boldsymbol{\Phi}}_K$ 和 $\widetilde{\boldsymbol{\Phi}}_K^{\text{c}}$ 的条件数,条件数小的为最佳解 X 而另一个为当前解 X^{c}。

3) 初步搜索步骤如下:

① 生成一个随机整数补偿量 $\{m_i\}_{i=1}^{P}$,并且令 $\{\overrightarrow{d_i^{\text{c}}}\}_{i=1}^{P}=\{d_i^{\text{c}}\}_{i=1}^{P}+\{m_i\}_{i=1}^{P}$ 作为新的位置向量来形成下一个传感矩阵 $\overrightarrow{\widetilde{\boldsymbol{\Phi}}_K^{\text{c}}}$。

② 计算 $\overrightarrow{\widetilde{\boldsymbol{\Phi}}_K^{\text{c}}}$ 的条件数作为新解 $\overrightarrow{X^{\text{c}}}$。

③ 如果 $\overrightarrow{X^{\text{c}}}<X$,令 $X=\overrightarrow{X^{\text{c}}}$,$\widetilde{\boldsymbol{\Phi}}_K=\overrightarrow{\widetilde{\boldsymbol{\Phi}}_K^{\text{c}}}$ 并且 $\{d_i\}_{i=1}^{P}=\{\overrightarrow{d_i^{\text{c}}}\}_{i=1}^{P}$。

④ 在[0, 1]内产生一个均匀分布的随机数 θ,并根据 Metropolis 准则,计算可接受概率为

$$\rho=\begin{cases}1 & \overrightarrow{X^{\text{c}}}<X^{\text{c}}\\ \exp\left(\dfrac{X^{\text{c}}-\overrightarrow{X^{\text{c}}}}{T}\right) & \text{其他}\end{cases}\tag{9}$$

如果 $\theta<\rho$,接受一个新解。设定 $X^{\text{c}}=\overrightarrow{X^{\text{c}}}$,令 $\{d_i^{\text{c}}\}_{i=1}^{P}=\{\overrightarrow{d_i^{\text{c}}}\}_{i=1}^{P}$。

⑤ 重复执行步骤① ~ ④,直到搜索步骤的数目等于 L_{M}。

4) 如果 $T>T_{\text{stop}}$,设定 $T=\alpha T$ 并跳至步骤 3)。

5) 如果 $\text{rank}(\widetilde{\boldsymbol{\Phi}}_K)=K$,储存 $\widetilde{\boldsymbol{\Phi}}_K$ 作为最佳传感矩阵,并将 $\{d_i\}_{i=1}^{P}$ 作为最优传感器位置向量,否则跳转至步骤 2)。

值得注意的是,在确定传感器最优位置分布时,需要计算 $\boldsymbol{\Phi}_K$ 中所有 P 个可能子矩阵的条件数,计算量大概是 $N!$。模拟退火算法在搜索过程中需要精确调整接受退化结果的条件才可能取得较优解,在无法得到实际问题精确最优解的情况下,无论从存储空间还是计算效率考虑,其都不失为一种优越的近似方法,特别是对解大规模 NP-hard 问题更是如此。遗传算法的主要缺点在于,一方面其局部搜索能力相对较差,尤其是在优化后期搜索效率较低,增加了优化时间消耗;另一方面,该算法需要对问题进行编码,在求解结束时还需要对问题解码,同时交叉率与变异率以及全体规模等参数的选择对算法性能影响很大,而合适值的选择通常对经验要求较高。对于结构复杂的组合优化问题,搜索空间大,不合适参数的选择会让优化陷入“早熟”。而蚁群算法是一个建模方案复杂的算法,具有非常高的时间复杂度。因此,本文选择模拟退火算法用于求解传感器的最优位置问题。

3　算法性能验证

3.1　含噪热图像重构技术

给定的 M 幅热图像的向量化集合 $\{\boldsymbol{w}_j\}_{j=0}^{M-1}$,根据式(5)可以得到的传感矩阵 $\widetilde{\boldsymbol{\Phi}}_K$,在此基础上,采用模拟退火算法计算出 P 个含噪传感器的优化位置分布 $\{d_i\}_{i=1}^{P}$。如果 $P\geqslant K$ 并且 $\text{rank}(\widetilde{\boldsymbol{\Phi}}_K)=K$,重构图像向量集合 $\{\widetilde{\boldsymbol{w}}_j\}_{j=0}^{M-1}$ 可以表示为[9]

$$\{\widetilde{\boldsymbol{w}}_j\}_{j=0}^{M-1}=\widetilde{\boldsymbol{\Phi}}_K(\widetilde{\boldsymbol{\Phi}}_K^{*}\widetilde{\boldsymbol{\Phi}}_K)^{-1}\widetilde{\boldsymbol{\Phi}}_K^{*}(\boldsymbol{\lambda}^{\text{S}})+\{\boldsymbol{\mu}_j\}_{j=0}^{M-1}\tag{10}$$

3.2　仿真平台的搭建

为了验证上述算法的有效性,本文采用仿真工具链[10]来模拟一个基于 Alpha 21264 架构的 16 核处理器[11]的热特性分布。通过微处理器性能仿真软件 SimpleScalar[12]在该芯片架构上仿真 SPEC2006[13]标准测试程序,并通过 Wattch infrastructure[14]来计算模块的功耗,最后由热特性仿真软件 Hotspot[15]计算出处理器的温度分布。

3.3　实验结果

实验中所有程序均在物理内存为 8 GB、主频为 4 GHz 的 Intel Skylake 四核处理器上由 MAT-

北航学报 赠阅

LAB 编写完成。这里给出以下性能指标：

1）平均重构误差。考虑设定 M 幅原始温度图像 $\boldsymbol{T}=\{\boldsymbol{T}_j\}_{j=0}^{M-1}$，平均重构误差定义为

$$E_{\text{avg}}=\frac{1}{MLW}\sum_{j=0}^{M-1}\sum_{x=0}^{L-1}\sum_{y=0}^{W-1}\left|\boldsymbol{T}_j(x,y)-\widetilde{\boldsymbol{T}}_j(x,y)\right| \tag{11}$$

2）最大重构误差。其定义为

$$E_{\max}=\max_{x,y,j}\left|\boldsymbol{T}_j(x,y)-\widetilde{\boldsymbol{T}}_j(x,y)\right| \tag{12}$$

3）均方误差。其表达式为

$$M_{\text{SE}}=\frac{1}{MLW}\sum_{j=0}^{M-1}\sum_{x=0}^{L-1}\sum_{y=0}^{W-1}\left|\boldsymbol{T}_j(x,y)-\widetilde{\boldsymbol{T}}_j(x,y)\right|^2 \tag{13}$$

4）误警率。误警热点的定义包含 2 种情况：①热点的实际温度达到或超过门限温度，而传感器对相应热点的温度估计值低于门限温度；②热点的实际温度没有达到门限温度，而传感器对相应热点的温度估计值达到或超过门限温度。其公式为

$$F_{\text{AR}}=\left(\frac{1}{MLW}\sum_{j=0}^{M-1}F_j\right)\times 100\% \tag{14}$$

式中：F_j 为 $\boldsymbol{T}_j$ 的误警热点数量。

3.4 传感器分布性能比较

为了验证本文方法的有效性，分别作了 3 组实验进行比较。随机从 SPEC2006 所有 29 个应用程序产生的热图像中选择 $M=22$ 个样本图像，每幅热图像 $N=1\,024$，离散为 $L=W=32$。第 1 组实验中，在放置不同数量传感器的条件下，分别对 4 种传感器放置的方法（本文算法、贪婪算法[9]、随机放置方法和均匀放置方法）所得到传感矩阵 $\widetilde{\boldsymbol{\Phi}}_K$ 的条件数进行比较。其结果如图 1 所示，可以观察到，相对于其他方法，本文算法能够获得传感矩阵的最小条件数。

热重构技术可以近似复原整个芯片温度，主要应用于动态热管理中实现全局监控。为了验证本文算法的热重构性能，第 2 组实验中在噪声标准差为 5% 的情况下，采用 EigenMaps[9] 热重构方法进行重构实验。根据式（10）执行 10 000 次模拟获得性能指标数据，如表 1 所示。对于上述 4 种不同传感器分布方法的热重构性能，分别从平均重构误差、最大重构误差、均方误差和信噪比等方面进行比较。图 2 ~ 图 5 的实验结果证明了本文算法在热重构方面的优越性能。

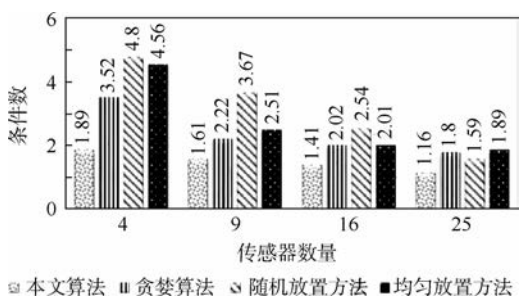


图 1 传感矩阵条件数对比

Fig. 1 Comparison of sensing matrix condition number

表 1 噪声标准差为 5%时，热重构性能比较

Table 1 Comparison of thermal reconstruction performance with noise standard deviation of 5%

方法	传感器数量	平均重构误差/℃	最大重构误差/℃	均方误差/℃	信噪比/dB	误警率/%
EigenMaps + 本文算法	4	2.39	12.74	9.73	0.65	18.85
	9	1.21	6.29	2.48	6.58	9.79
	16	1.13	5.96	2.15	7.18	9.06
	25	0.84	4.50	1.19	9.76	6.52
EigenMaps + 贪婪算法	4	2.77	17.27	14.31	-0.94	20.55
	9	1.81	11.36	5.97	2.83	14.19
	16	1.40	7.96	3.39	5.24	11.46
	25	1.07	6.11	1.99	7.53	8.77
EigenMaps + 随机放置方法	4	3.76	26.31	29.07	-3.95	24.29
	9	2.54	13.28	11.70	-0.05	19.66
	16	1.43	8.46	3.72	4.83	11.47
	25	0.95	5.76	1.59	8.52	7.32
EigenMaps + 均匀放置方法	4	3.74	23.33	26.19	-3.57	24.93
	9	2.06	11.79	7.47	1.83	16.38
	16	1.22	7.62	2.63	6.36	9.94
	25	1.15	6.67	2.31	6.90	9.05

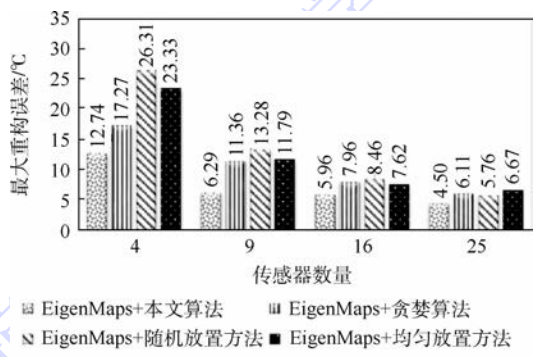


图 2 最大重构误差对比

Fig. 2 Comparison of maximum reconstruction error

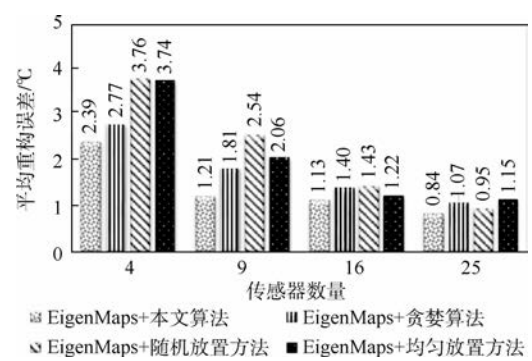


图 3 平均重构误差对比

Fig. 3 Comparison of average reconstruction error

由于实际芯片中传感器不可避免地伴随有多种噪声,这无论给局部监控还是全局监控都带来了热点误警率的增加,在一定程度上会加剧错误的预警和不必要的响应,使动态热管理的可靠性受到影响。第 3 组实验中,在噪声标准差为 5% 的情况下,比较 4 种传感器分布方法在误警率方面的性能,如图 6 所示。实验结果证实,当传感器读数受到噪声影响时,本文算法可以显著减少误警率,能够有效运用在动态热管理中实现精确的热监控。

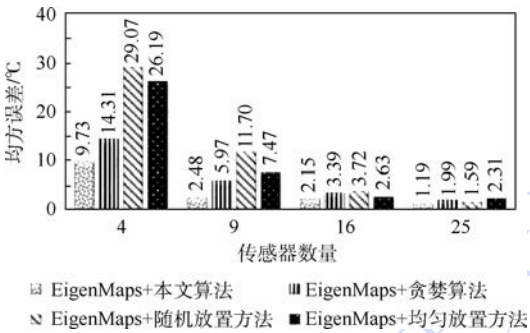


图 4　均方误差对比

Fig. 4　Comparison of mean square error

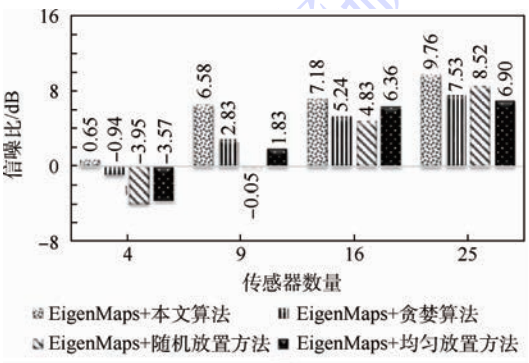


图 5　信噪比对比

Fig. 5　Comparison of SNR

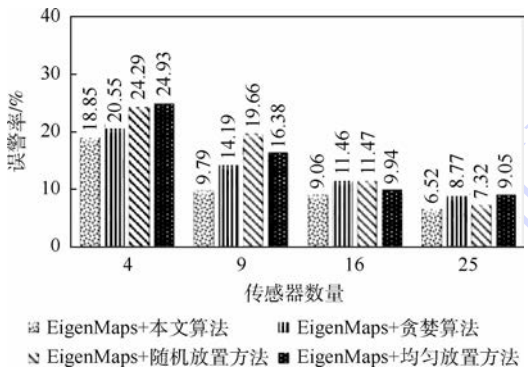


图 6　误警率对比

Fig. 6　Comparison of false alarm rate

4　结　论

本文提出微处理器中考虑噪声因素的热传感器位置分布优化方法。

1) 在均方误差最小的原则下,对原始热图像样本矩阵进行降维近似处理,并结合矩阵扰动分析建立了含噪热传感器位置分布理论。

2) 使用基于模拟退火算法寻找最优传感器位置。

3) 实验结果证明,本文算法比现有的贪婪算法在热重构误差、信噪比和误警率等性能方面有了一定提高,能够有效运用在动态热管理中实现精确的热监控。

在今后的工作中,会对遗传算法和蚁群算法的性能进行详细比较,并结合其他现代优化算法做进一步的优化和完善。

参考文献 (References)

[1] SHI B, ZHANG Y, SRIVA A, et al. Dynamic thermal management under soft thermal constraints[J]. IEEE Transactions on VLSI Systems, 2013, 21(11): 2045-2054.

[2] ZHANG Y, SRIVA A. Accurate temperature estimation using noisy thermal sensors for Gaussian and non-Gaussian cases[J]. IEEE Transactions on Very Large Scale Integration Systems, 2011, 19(9): 1617-1626.

[3] LONG J, MEMIK S O, MEMIK G, et al. Thermal monitoring mechanisms for chip multiprocessors[J]. ACM Transactions on Architecture and Code Optimization, 2008, 5(2): 9: 1-9: 33.

[4] MEMIK S O, MUKHERJ R, LONG J. et al. Optimizing thermal sensor allocation for microprocessors[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits, 2008, 27(3): 516-527.

[5] REDA S, COCHRAN R, NOWROZ A N. Improved thermal tracking for processors using hard and soft sensor allocation techniques[J]. IEEE Transactions on Computers, 2011, 60(6): 841-851.

[6] LI X, RONG M, LIU T, et al. Inverse distance weighting method based on a dynamic voronoi diagram for thermal reconstruction with limited sensor data on multiprocessors[J]. IEICE Transactions on Electronics, 2011, 94(8): 1295-1301.

[7] KOTTAIMALAI R, RAJASEKARAN M P, SEVAM V, et al. EEG signal classification using principal component analysis with neural network in brain computer interface applications [C] // 2013 IEEE International Conference on Emerging Trends in Computing, Communication and Nanotechnology (ICECCN). Piscataway, NJ: IEEE Press, 2013: 227-231.

[8] KEIKHA M M. Improved simulated annealing using momentum terms[C] // 2011 Second International Conference on Intelligent Systems, Modelling and Simulation (ISMS). Piscataway, NJ: IEEE Press, 2011: 44-48.

[9] RANIERI J, VINCENZI A, CHEBIRA A, et al. 'EigenMaps': Algorithms for optimal thermal maps extraction and sensor placement on multicore processors[C] // Proceedings of the 49th Annual Design Automation Conference. Piscataway, NJ: IEEE Press, 2012: 636-641.

[10] LI X, RONG M, WANG R, et al. Reducing the number of sen-

北航学报 赠阅

sors under hotspot temperature error bound for microprocessors based on dual clustering[J]. IET Circuits, Devices & Systems, 2013, 7(4): 211-220.

[11] MANIATAKOS M, MICHAEL M K, MAKRIS Y, et al. Multiple-bit upset protection in microprocessor memory arrays using vulnerability-based parity optimization and interleaving[J]. IEEE Transactions on VLSI Systems, 2015, 23(11): 2447-2460.

[12] BAZZAZ M, SALEHI M, EJLAI A, et al. An accurate instruction-level energy estimation model and tool for embedded systems[J]. IEEE Transactions on Instrumentation and Measurement, 2013, 62(7): 1927-1934.

[13] ZOU Q, YUE J, SEGEE B, et al. Temporal characterization of SPEC CPU2006 workloads analysis and synthesis[C]// 2012 IEEE 31st International Performance Computing and Communications Conference (IPCCC). Piscataway, NJ: IEEE Press, 2012: 11-20.

[14] BROOKS D, TIWARI V, MARTONOSI M. Wattch: A framework for architectural-level power analysis and optimizations [C]// Proceedings of the 27th International Symposium on Computer Architecture. Piscataway, NJ: IEEE Press, 2000: 83-94.

[15] HUANG W, GHOSH S, VEKYSAMY S, et al. HotSpot: A compact thermal modeling methodology for early-stage VLSI design [J]. IEEE Transactions on VLSI Systems, 2006, 14(5): 501-513.

作者简介：

李鑫 男，硕士研究生。主要研究方向：高性能处理器中热传感器噪声误差分析。

E-mail: louis_li@mail.nwpu.edu.cn

李鑫 男，博士，讲师。主要研究方向：高性能处理器热特性分析。

E-mail: xinli@nwpu.edu.cn

杨金孝 男，博士，副教授。主要研究方向：热传感器数量分配和位置分布方法。

E-mail: yjx@nwpu.edu.cn

程元乐 男，硕士研究生。主要研究方向：热特性仿真与验证平台搭建。

Optimization method of thermal sensor placement for microprocessor with noise

LI Xin, LI Xin*, YANG Jinxiao, CHENG Yuanle

(School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: High-performance processors employ dynamic thermal management (DTM) techniques, which use a set of on-chip thermal sensors to continuously monitor temperatures during the runtime. However, on-chip sensors are inevitably accompanied by a range of noise sources due to fabrication randomness and fluctuations, which greatly affects the reliability of DTM. Therefore, in order to improve the accuracy of thermal monitoring, this paper utilizes principal component analysis (PCA) to approximately reduce dimensionality of the original sample matrix of the thermal image, and combines it with matrix perturbation analysis to propose the optimization method of thermal sensor placement which is based on simulated annealing algorithm. Experimental results indicate that the proposed method significantly outperforms the existing greedy algorithm in the aspects of thermal reconstruction error, Signal to noise ratio (SNR) and false alarm, which can be better applied in dynamic thermal management techniques to achieve accurate thermal monitoring.

Key words: thermal sensor; dynamic thermal management; thermal reconstruction; noise; placement

Received: 2016-03-15; **Accepted**: 2016-06-12; **Published online**: 2016-07-01 19:15

URL: www.cnki.net/kcms/detail/11.2625.V.20160701.1915.001.html

Foundation item: National Natural Science Foundation of China (61501377)

* **Corresponding author**. E-mail: xinli@nwpu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0696

2×2 X 信道和 PTP 信道并存网络的自由度

郭倩倩，刘锋*，曾连荪，覃闻铭

（上海海事大学 信息工程学院，上海 201306）

摘 要：自由度(DOF)是一个空间复用概念，是在多天线系统中引入的，体现了无线信道对空间资源的利用效率。在认知无线电中，由于主、次网络会相互干扰，从而会降低DOF。本文主要研究的是2×2 X信道(XC)和点对点(PTP)信道并存时网络的DOF。研究结论也适合于XC优先于PTP信道的情况。在已知发送端信道状态(CSI)信息的前提下，基于符号扩展模型，利用渐进干扰对齐方法给出了对应的信号处理方案，并证明当 n 趋于无穷大时，自由度内界可达5/3。利用传播时延模型示范了其可达性。此外，分析了DOF上界。

关 键 词：X信道(XC)；点对点(PTP)信道；信道并存；渐进干扰对齐；自由度(DOF)
中图分类号：TN929.5
文献标识码：A **文章编号**：1001-5965(2016)11-2501-06

多输入多输出(Multiple-Input Multiple-Output, MIMO)系统可以在不增加系统带宽和传输功率的前提下，利用空间自由度(Degrees of Freedom，DOF)来成倍地提高无线信道的容量。然而，即使将MIMO技术应用到多网络系统中，由于网络之间消息相互干扰，系统性能受到干扰的影响还是很严重。因为研究无线网络容量一般较为困难，可以用空间自由度进行近似。关于MIMO无线网络现有研究中，已经有对点对点(Point to Point，PTP)信道、广播信道(Broadcast Channel，BC)、多址接入信道(Multiple Access Channel，MAC)、干扰信道(Interference Channel，IC)和X信道(X Channel，XC)等单个信道的自由度研究。多用户信道并存网络模型的自由度研究却很少。在无线网络中，干扰是限制信道自由度的一个主要因素。尤其是多个网络并存时，网络内部和网络之间都会有干扰，那么对干扰进行处理在研究信道的自由度中就显得尤为重要。干扰处理的方法主要有干扰回避、干扰正交化、干扰对齐和干扰中和等，本文主要用到干扰对齐。干扰对齐[1]的基本思想是：通过收发端的联合设计，使得在每个接收端其收到的干扰可以尽量地对齐到信号空间的某些维度，达到压缩干扰占用空间、增大期望信号可用空间的目的；同时还要保持期望信号与干扰信号相互独立，期望信号处于零干扰的空间中，这样期望信号才能够被解析出来。由于干扰对齐的重要性，干扰对齐技术在很多模型中被得到广泛应用。干扰对齐的具体实现方法有很多，比如线性干扰对齐、渐近干扰对齐和盲干扰对齐等。其中，线性干扰对齐的方案是最容易实现也是最常用的一种方法，主要基于线性预编码(波束成形)机制的信号空间对齐方法。波束成型方案在现存的MIMOPTP信道、BC和MAC中应用非常普遍，例如K用户的IC中，通过线性干扰对齐，每个用户可以得到1/2的自由度[1]；在 $M\times N$ 用户的XC[2]中，通过线性干扰对齐方案[3]能达到

收稿日期：2015-10-27；**录用日期**：2015-12-18；**网络出版时间**：2016-01-21 15:47
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160121.1547.012.html
基金项目：国家自然科学基金(61271283，61472237，61401270)；交通运输部应用基础研究项目(2015329810030)；上海市教委科研创新基金(14YZ113)
* **通讯作者**：Tel.：021-38282863 E-mail：liufeng@shmtu.edu.cn

引用格式：郭倩倩，刘锋，曾连荪，等. 2×2 X 信道和 PTP 信道并存网络的自由度[J]. 北京航空航天大学学报，2016，42(11)：2501-2506. GUO Q Q，LIU F，ZENG L S，et al. Degrees of freedom of networks with coexistence of 2×2 X channel and point-to-point channel [J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2501-2506 (in Chinese).

北航学报 赠阅

$MN/(M+N-1)$ 的自由度。

过去的研究中已经通过干扰对齐和干扰中和方法获知了 PTP 信道、BC、MAC、IC 和 XC 等简单信道的自由度[4-9]。对于更加复杂的信道,如 2 个或多个简单信道共存构成的复合信道,这些复合信道的自由度研究仅有少量的进展。文献[6]推导了 2 个 PTP 信道并存时的网络模型的自由度。在 2 个 BC 并存时的自由度研究[7-8]中,比较发现,使用干扰对齐的方法比使用迫零方法得到的自由度要大。同样,使用干扰对齐方法也研究得出了 2 个 MAC 共存的网络模型的自由度[9]。所以,干扰对齐在复杂网络中的应用越来越广泛。

本文中研究的是 2×2 XC 与 PTP 信道 2 种信道共存时网络整体的自由度,这是一种新的并存模型。不同于认知无线电[10]中主次网络次序不同的假设,本文模型中 2 个信道的次序可以无关;尽管如此,依然借用了主网与次网的概念来描述与区分所涉及的 2 个信道。使用渐近干扰对齐方法,证明其可达自由度为 5/3;另一方面,目前找到的自由度上界为 2。

符号说明:本文中的 $\boldsymbol{m}_{ij}$ 为发送端 j 发送到接收端 i 相应的消息;$\hat{\boldsymbol{m}}_{ij}$ 为接收端 i 的期望信号;$\mathrm{diag}(a_1,a_2,\cdots,a_N)$ 为对角元素为 $a_1,a_2,\cdots,a_N$ 的对角阵;$(\cdot)^{\mathrm{T}}$ 为向量或矩阵的转置;C_{ij} 为矩阵第 i 行第 j 列元素的代数余子式。

1 系统模型

本文研究的 2×2 XC 与 PTP 信道并存系统模型如图 1 所示,上方虚线框图内的 2×2 XC 代表主网络,下方虚线框图内的 PTP 信道代表次网络(图 2 和图 3 中的图形均适用)。

主网络是 2×2 XC,次网络是 PTP 信道,考虑每个节点仅有一根天线的基本情况。由于单天线

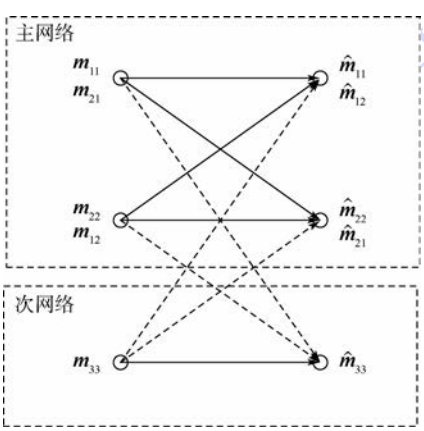


图 1 2×2 XC 和 PTP 信道共存模型

Fig. 1 Coexistence model of 2×2 XC and PTP channel

导致信号空间不足以进行信号处理,采用符号扩展[11-14]的方法来进行分析,这样收发端才有足够的空间可以信号检测与干扰处理。注意:得到的自由度结论需要除以所用的符号扩展数目,即进行归一化处理[15-16]才能与原模型相对应。

消息与干扰情况说明如下:$\boldsymbol{m}_{11}$、$\boldsymbol{m}_{21}$ 为发送端 1 要发送的 2 个消息,$\boldsymbol{m}_{22}$、$\boldsymbol{m}_{12}$ 为发送端 2 要发送的 2 个消息,$\boldsymbol{m}_{33}$ 为发送端 3 要发送的消息,则发送端总共发送 5 个消息。为了不失一般性,假设各发送端信号满足功率约束关系:$E\{|\boldsymbol{X}_i|^2\}\leqslant P$,其中:$\boldsymbol{X}_i$ 为发送信号;E 为数学期望,P 为功率。图 1 中实线代表的是发送端向对应接收端发送的期望信号,虚线代表的是对此接收端的干扰信号。接收端 1 的期望消息是 $\boldsymbol{m}_{11}$、$\boldsymbol{m}_{12}$,其他 3 个消息是干扰消息;接收端 2 的期望消息是 $\boldsymbol{m}_{21}$、$\boldsymbol{m}_{22}$,其他 3 个消息是干扰消息;接收端 3 的期望消息是 $\boldsymbol{m}_{33}$,其余 4 个消息是干扰消息。设所用符号扩展数目为 M_n。输入输出关系为

$$\boldsymbol{Y}_j=\boldsymbol{H}_{j1}\boldsymbol{v}_1^1\boldsymbol{m}_{11}+\boldsymbol{H}_{j2}\boldsymbol{v}_2^2\boldsymbol{m}_{12}+\boldsymbol{H}_{j3}\boldsymbol{v}_3^1\boldsymbol{m}_{33}+\boldsymbol{H}_{j1}\boldsymbol{v}_1^2\boldsymbol{m}_{21}+\boldsymbol{H}_{j2}\boldsymbol{v}_2^1\boldsymbol{m}_{22}+\boldsymbol{z}_j \tag{1}$$

式中:$\boldsymbol{Y}_j$ 为维度为 $M_n\times 1$、第 $j(j=1,2,3)$ 个接收端的输出信号;$\boldsymbol{H}_{ji}$ 为发送端 i 到接收端 j、维度为 $M_n\times M_n$ 的信道矩阵;$\boldsymbol{v}_j^i(j=1,2,3;i=1,2)$ 为发送端 j 的第 i 个波束成形向量,其中 $\boldsymbol{v}_1^1$ 和 $\boldsymbol{v}_2^1$ 是 $(n+1)^N\times 1$ 维,其他波束成形向量是 $n^N\times 1$ 维,n 为变量,自由度随 n 变化而变化,当 n 趋于无穷大时,内界可达 5/3,N 为 2.1 节中 6 个约束关系式分成的两部分中,每个部分的 T_k 个数,可得 $N=2$;$\boldsymbol{z}_j(j=1,2,3)$ 代表的是 $M_n\times 1$ 维信道的加性高斯白噪声。

2 可达自由度

2.1 自由度下界

证明思路主要为让每个消息均获得 1/3 的自由度。为此,需要的符号扩展个数可以确定为 $M_n=2(n+1)^N+n^N$,其中消息 $\boldsymbol{m}_{11}$、$\boldsymbol{m}_{22}$ 各包含 $(n+1)^N$ 个符号,消息 $\boldsymbol{m}_{12}$、$\boldsymbol{m}_{21}$ 和 $\boldsymbol{m}_{33}$ 各包含 n^N 个符号。

定理 1 2×2 XC 和 PTP 共存的网络,使用渐近干扰对齐方法,网络自由度为 5/3 是可达的,其中 2×2 XC 自由度为 4/3,PTP 自由度为 1/3。

证明 采用的干扰处理方案如下。

在接收端 1,期望信号为 $\boldsymbol{m}_{11}$、$\boldsymbol{m}_{12}$,干扰信号为 $\boldsymbol{m}_{21}$、$\boldsymbol{m}_{22}$ 和 $\boldsymbol{m}_{33}$,输入输出关系是

$$\boldsymbol{Y}_1=\boldsymbol{H}_{11}\boldsymbol{v}_1^1\boldsymbol{m}_{11}+\boldsymbol{H}_{12}\boldsymbol{v}_2^2\boldsymbol{m}_{12}+\boldsymbol{H}_{13}\boldsymbol{v}_3^1\boldsymbol{m}_{33}+\boldsymbol{H}_{11}\boldsymbol{v}_1^2\boldsymbol{m}_{21}+\boldsymbol{H}_{12}\boldsymbol{v}_2^1\boldsymbol{m}_{22}+\boldsymbol{z}_1$$

北航学报 赠阅

因为每个消息自由度为 1/3,两个期望消息已经占接收端 1 空间的 2/3,所以只剩 1/3 空间。将干扰 $\boldsymbol{m}_{21}$ 和 $\boldsymbol{m}_{33}$ 对齐到 $\boldsymbol{m}_{22}$ 的空间中,约束条件为

$$\begin{cases}\boldsymbol{H}_{13}\boldsymbol{v}_3^1 \prec \boldsymbol{H}_{12}\boldsymbol{v}_2^1\\ \boldsymbol{H}_{11}\boldsymbol{v}_1^2 \prec \boldsymbol{H}_{12}\boldsymbol{v}_2^1\end{cases} \tag{2}$$

式中:$\prec$ 表示归属于某个集合,例如 $\boldsymbol{a}\prec\boldsymbol{b}$ 就是矩阵 $\boldsymbol{a}$ 的列向量组成的集合是矩阵 $\boldsymbol{b}$ 的列向量组成集合的子集。在接收端 2,使用渐进干扰对齐,期望信号为 $\boldsymbol{m}_{21}$、$\boldsymbol{m}_{22}$,干扰信号为 $\boldsymbol{m}_{11}$、$\boldsymbol{m}_{12}$ 和 $\boldsymbol{m}_{33}$,输入输出关系为

$$\boldsymbol{Y}_2=\boldsymbol{H}_{21}\boldsymbol{v}_1^1\boldsymbol{m}_{11}+\boldsymbol{H}_{22}\boldsymbol{v}_2^2\boldsymbol{m}_{12}+\boldsymbol{H}_{23}\boldsymbol{v}_3^1\boldsymbol{m}_{33}+\boldsymbol{H}_{21}\boldsymbol{v}_1^2\boldsymbol{m}_{21}+\boldsymbol{H}_{22}\boldsymbol{v}_2^1\boldsymbol{m}_{22}+\boldsymbol{z}_2$$

将干扰 $\boldsymbol{m}_{12}$、$\boldsymbol{m}_{33}$ 对齐到 $\boldsymbol{m}_{11}$ 的空间中,则约束条件为

$$\begin{cases}\boldsymbol{H}_{22}\boldsymbol{v}_2^2 \prec \boldsymbol{H}_{21}\boldsymbol{v}_1^1\\ \boldsymbol{H}_{23}\boldsymbol{v}_3^1 \prec \boldsymbol{H}_{21}\boldsymbol{v}_1^1\end{cases} \tag{3}$$

在接收端 3,期望信号为 $\boldsymbol{m}_{33}$,干扰信号为 $\boldsymbol{m}_{11}$、$\boldsymbol{m}_{12}$、$\boldsymbol{m}_{21}$ 和 $\boldsymbol{m}_{22}$,输入输出关系为

$$\boldsymbol{Y}_3=\boldsymbol{H}_{31}\boldsymbol{v}_1^1\boldsymbol{m}_{11}+\boldsymbol{H}_{32}\boldsymbol{v}_2^2\boldsymbol{m}_{12}+\boldsymbol{H}_{33}\boldsymbol{v}_3^1\boldsymbol{m}_{33}+\boldsymbol{H}_{31}\boldsymbol{v}_1^2\boldsymbol{m}_{21}+\boldsymbol{H}_{32}\boldsymbol{v}_2^1\boldsymbol{m}_{22}+\boldsymbol{z}_3$$

因为期望信号仅占接收端空间的 1/3,剩余 2/3 的空间放干扰消息。将干扰消息 $\boldsymbol{m}_{12}$ 对齐到 $\boldsymbol{m}_{11}$ 的空间中,干扰消息 $\boldsymbol{m}_{21}$ 对齐到 $\boldsymbol{m}_{22}$ 的空间中,则约束关系为

$$\begin{cases}\boldsymbol{H}_{32}\boldsymbol{v}_2^2 \prec \boldsymbol{H}_{31}\boldsymbol{v}_1^1\\ \boldsymbol{H}_{31}\boldsymbol{v}_1^2 \prec \boldsymbol{H}_{32}\boldsymbol{v}_2^1\end{cases} \tag{4}$$

由接收端 1、接收端 2、接收端 3 干扰对齐方案的约束条件式(2)、式(3)和式(4)可以得到

$$\begin{cases}\boldsymbol{H}_{12}^{-1}\boldsymbol{H}_{13}\boldsymbol{v}_3^1 \prec \boldsymbol{v}_2^1\\ \boldsymbol{H}_{21}^{-1}\boldsymbol{H}_{22}\boldsymbol{v}_2^2 \prec \boldsymbol{v}_1^1\\ \boldsymbol{H}_{12}^{-1}\boldsymbol{H}_{11}\boldsymbol{v}_1^2 \prec \boldsymbol{v}_2^1\end{cases}\quad\begin{cases}\boldsymbol{H}_{21}^{-1}\boldsymbol{H}_{23}\boldsymbol{v}_3^1 \prec \boldsymbol{v}_1^1\\ \boldsymbol{H}_{32}^{-1}\boldsymbol{H}_{31}\boldsymbol{v}_1^2 \prec \boldsymbol{v}_2^1\\ \boldsymbol{H}_{31}^{-1}\boldsymbol{H}_{32}\boldsymbol{v}_2^2 \prec \boldsymbol{v}_1^1\end{cases} \tag{5}$$

令 $\boldsymbol{B}_1=\boldsymbol{H}_{12}^{-1}\boldsymbol{H}_{13}\boldsymbol{v}_3^1$,对应波束成形向量 $\boldsymbol{v}_3^1$;令 $\boldsymbol{B}_2=\boldsymbol{H}_{21}^{-1}\boldsymbol{H}_{22}\boldsymbol{v}_2^2$,对应波束成形向量 $\boldsymbol{v}_2^2$。对角阵 $\boldsymbol{g}_i=\mathrm{diag}(g_{11},g_{22},\cdots,g_{n^Nn^N})$ $(j=1,2,3)$ 来自一个连续分布,以保证矩阵的随机性。令 $\boldsymbol{g}_1\boldsymbol{v}_3^1=\boldsymbol{g}_2\boldsymbol{v}_1^2=\boldsymbol{g}_3\boldsymbol{v}_2^2$,目的是让 $\boldsymbol{v}_3^1$ 分别能用 $\boldsymbol{v}_1^2$、$\boldsymbol{v}_2^2$ 表示。记:

$$\boldsymbol{T}_1=\boldsymbol{H}_{12}^{-1}\boldsymbol{H}_{11}\boldsymbol{g}_2^{-1}\boldsymbol{g}_1\boldsymbol{H}_{13}^{-1}\boldsymbol{H}_{12} \tag{6}$$

$$\boldsymbol{T}_2=\boldsymbol{H}_{32}^{-1}\boldsymbol{H}_{31}\boldsymbol{g}_2^{-1}\boldsymbol{g}_1\boldsymbol{H}_{13}^{-1}\boldsymbol{H}_{12} \tag{7}$$

$$\boldsymbol{T}_3=\boldsymbol{H}_{21}^{-1}\boldsymbol{H}_{23}\boldsymbol{g}_1^{-1}\boldsymbol{g}_3\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21} \tag{8}$$

$$\boldsymbol{T}_4=\boldsymbol{H}_{31}^{-1}\boldsymbol{H}_{32}\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21} \tag{9}$$

通过式(6)~式(9)可以将式(5)改写为

$$\begin{cases}\boldsymbol{B}_1 \prec \boldsymbol{v}_2^1\\ \boldsymbol{T}_1\boldsymbol{B}_1 \prec \boldsymbol{v}_2^1\\ \boldsymbol{T}_2\boldsymbol{B}_1 \prec \boldsymbol{v}_2^1\end{cases}\quad\begin{cases}\boldsymbol{B}_2 \prec \boldsymbol{v}_1^1\\ \boldsymbol{T}_3\boldsymbol{B}_2 \prec \boldsymbol{v}_1^1\\ \boldsymbol{T}_4\boldsymbol{B}_2 \prec \boldsymbol{v}_1^1\end{cases} \tag{10}$$

令 $\boldsymbol{w}_1$ 和 $\boldsymbol{w}_2$ 为 $M_n\times 1$ 的列向量,具体定义为

$$\boldsymbol{w}_1=[1\quad 1\quad 1\quad\cdots\quad 1]^{\mathrm{T}}$$

$$\boldsymbol{w}_2=[1\quad 2\quad 1\quad\cdots\quad 1]^{\mathrm{T}}$$

这样可保证 $\boldsymbol{w}_1$ 和 $\boldsymbol{w}_2$ 线性独立。要满足条件式(10),矩阵 $\boldsymbol{B}_1$ 和 $\boldsymbol{B}_2$、$\boldsymbol{v}_2^1$ 和 $\boldsymbol{v}_1^1$ 分别从式(11)~式(14)对应的集合中选取。

$$B_1=\left\{\left(\prod_{k\in\{1,2\}}(\boldsymbol{T}_k)^{\alpha_k}\right)\boldsymbol{w}_1:\forall\alpha_k\in\{0,1,\cdots,n-1\}\right\} \tag{11}$$

$$v_2^1=\left\{\left(\prod_{k\in\{1,2\}}(\boldsymbol{T}_k)^{\beta_k}\right)\boldsymbol{w}_1:\forall\beta_k\in\{0,1,\cdots,n\}\right\} \tag{12}$$

$$B_2=\left\{\left(\prod_{k\in\{3,4\}}(\boldsymbol{T}_k)^{\alpha_k}\right)\boldsymbol{w}_2:\forall\alpha_k\in\{0,1,\cdots,n-1\}\right\} \tag{13}$$

$$v_1^1=\left\{\left(\prod_{k\in\{3,4\}}(\boldsymbol{T}_k)^{\beta_k}\right)\boldsymbol{w}_2:\forall\beta_k\in\{0,1,\cdots,n\}\right\} \tag{14}$$

要使接收端的期望信号能被解析出来,需要期望信号与干扰信号是相互独立的,即期望信号与干扰信号构成的空间是满秩的。即证明矩阵接收端 1 的信号空间$[\boldsymbol{H}_{11}\boldsymbol{v}_1^1\quad\boldsymbol{H}_{12}\boldsymbol{v}_2^2\quad\boldsymbol{H}_{12}\boldsymbol{v}_2^1]$是满秩的。矩阵$[\boldsymbol{H}_{11}\boldsymbol{v}_1^1\quad\boldsymbol{H}_{12}\boldsymbol{v}_2^2\quad\boldsymbol{H}_{12}\boldsymbol{v}_2^1]$左乘 $\boldsymbol{H}_{11}^{-1}$ 得到:$\boldsymbol{S}=[\boldsymbol{v}_1^1\quad\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{v}_2^2\quad\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{v}_2^1]$。

$\boldsymbol{w}_1\prec\boldsymbol{S}$、$\boldsymbol{w}_2\prec\boldsymbol{S}$,显然,如果 $\boldsymbol{w}_1$ 和 $\boldsymbol{w}_2$ 不是线性独立的,则 $\boldsymbol{S}$ 不是满秩的。矩阵 $\boldsymbol{S}$ 中第 l 行为 $\left[\prod_{k=3,4}(\lambda_{lk})^{\alpha_k}\quad d_l\prod_{k=3,4}(\gamma_{lk})^{\alpha_k}\quad d_l\prod_{k=1,2}(\lambda_{lk})^{\beta_k}\right]$,要证明接收端干扰信号和期望信号是满秩的,即证明 $\boldsymbol{S}$ 是满秩的,这里采用反证法证明。其中 $\alpha_k\in\{0,1,\cdots,n-1\}$,$\beta_k\in\{0,1,\cdots,n\}$,$d_l$ 取自一个连续的分布函数的随机变量。用反证法证明矩阵 $\boldsymbol{S}$ 是满秩的,即只要推出概率 $P(|\boldsymbol{S}|=0)>0$($|\boldsymbol{S}|$代表 $\boldsymbol{S}$ 的行列式)矛盾。

证明:

1)将矩阵 $\boldsymbol{S}$ 按第 1 行展开得到其行列式为

$$\begin{aligned}|\boldsymbol{S}|=&C_{1,1}+\lambda_{13}C_{1,2}+\cdots+\lambda_{13}^n\lambda_{14}^nC_{1,(n+1)^2}+\\&d_1(C_{1,(n+1)^2+1}+\gamma_{13}C_{1,(n+1)^5+2}+\cdots+\\&\gamma_{13}^n\gamma_{14}^nC_{1,2(n+1)^2})+d_1(C_{1,2(n+1)^2+1}+\\&\lambda_{11}C_{1,2(n+1)^2+2}+\cdots+\lambda_{11}^{n-1}\lambda_{12}^{n-1}C_{1,2(n+1)^2+n^2})\end{aligned} \tag{15}$$

如果 λ_{1k}、γ_{1k} $(k=1,2,3,4)$ 的值给定,那么 $|\boldsymbol{S}|=0$ 只需满足下列条件之一:①d_1 是此方程的根;②λ_{1k}、γ_{1k} 的各次项系数都等于 0。

由于 d_1 来自于一个连续的分布,所以满足上

北航学报 赠阅

述条件①的概率几乎为零,所以下面需要证明满足条件②的概率大于 0 即可,即:$P(|\boldsymbol{S}|=0)>0$ 可推出式(16)成立的概率大于 0。

$$C_{1,(n+1)^2+1}+\gamma_{13}C_{1,(n+1)^2+2}+\cdots+\gamma_{13}^{n}\gamma_{14}^{n}C_{1,2(n+1)^2}+C_{1,2(n+1)^2+1}+\lambda_{11}C_{1,2(n+1)^2+2}+\cdots+\lambda_{11}^{n-1}\lambda_{12}^{n-1}C_{1,2(n+1)^2+n^2}=0 \tag{16}$$

2) 接下来考虑式(16)。如果 λ_{1k}、γ_{1k} 给定,同理要保证等式(16)成立的概率大于 0,只需满足下列 2 个条件之一即可:①λ_{11}、λ_{12}、γ_{13} 和 γ_{14} 是此方程的根;②λ_{11}、λ_{12}、γ_{13} 和 γ_{14} 的各次项系数都等于 0。

满足条件①的概率几乎为零,所以只需证明满足条件②的概率大于零即可。

3) 由式(16)可以推出一个充分条件,即 $C_{1,2(n+1)^2+n^2}=0$ 的概率大于 0。以此类推,按第 l 层展开后,得到的充分条件是 $C_{l,2(n+1)^2+n^2-l}=0$ 成立的概率大于 0。

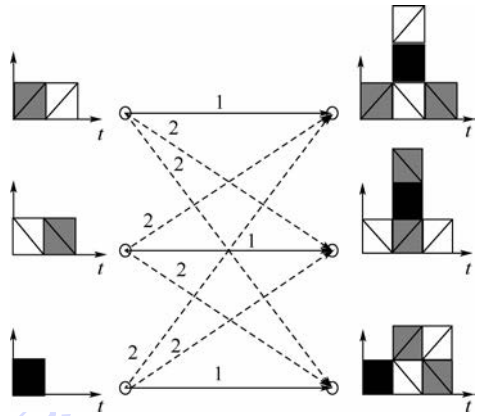
重复之前的步骤,这样一步步将第 1 行最后一列的代数余子式取出来,取 M_n-1 次之后会推出矛盾。不失一般性,将 $\boldsymbol{w}_1$ 放在第 1 列,那么会推出“1 = 0”的矛盾。因此原命题得证,即期望消息和干扰消息空间独立。同理,接收端 2 和接收端 3 期望消息和干扰消息空间也是独立的。

由以上证明可得,总共发送消息符号数 $2(n+1)^2+3n^2$ 个,符号扩展 $2(n+1)^2+n^2$,自由度为 $\dfrac{2(n+1)^2+3n^2}{2(n+1)^2+n^2}$。当 n 趋于无穷的时候,自由度达到 5/3。其中,2 × 2 XC 占用 4/5,所获自由度为 4/3;PTP 信道占用 1/5,所获自由度为 1/3。

由文献[2]可知 2 × 2 XC 本身信道自由度为 4/3,点对点信道与之共存时,只要经过良好设计,可以避免 2 × 2 XC 的自由度受到影响,同时点对点信道可以获得 1/3 的自由度。如果以点对点信道作为主网,其自由度为 1,此时如果不降低 PTP 信道自由度的话,由于没有多余信号空间容纳干扰,故次网 XC 将不能获得任何自由度。但是如果 PTP 信道可以降低自由度的话,XC 可以获得部分自由度,此时问题回到第 1 节的模型。故本文研究模型适用于 2 个网络无优先级并存,也适用于 2 × 2 XC 作为主网优先满足其自由度的情况。

2.2 基于传播时延的例证

下面举一个特殊例子来证明 2.1 节自由度下界是可以达到的。基于传播时延假设,图 2 中给出了 5/3 自由度如何达到的具体实例。图中实线代表传播时延为1,虚线代表传播时延为2,可以



t—时隙。

图 2 基于传播时延对自由度的分析图

Fig. 2 Analysis diagram of DOF based on propagation delay

看出,发送端发送 5 个消息,灰色方块代表接收端 1 要接收的消息,方块中有左对角线表示发送端 1 发送的消息;白色方块代表接收端 2 要接收的消息,方块中有右对角线表示发送端 2 发送的消息;黑色方块代表发送端 3 发给接收端 3 的消息。

基于上述传播时延假设及消息配置,如图 2 所示,发送端 1 在第 1 时隙发送灰色左对角线方块,到达接收端 1 存在 1 个传播时延;发送端 2 在第 1 时隙发送白色右对角线方块,发送端 3 在第 1时隙发送黑色方块,它们到达接收端 1 有2 个传播时延。在第 2 时隙,发送端 1 发送白色左对角方块消息,发送端 2 发送灰色右对角方块消息。接收端 1 在第 1 时隙和第 3 时隙得到期望消息(灰色方块)占 2/3,干扰消息占 1/3。同样地,在接收端 2,期望消息(白色方块)占 2/3,干扰消息占 1/3。发送端 3 在第 1 时隙得到期望消息,第 2、第 3 时隙得到干扰消息,从而期望消息占1/3。由系统模型传播时延的分析可得,模型自由度为 2/3 + 2/3 + 1/3 = 5/3 是可实现的。

2.3 自由度上界

根据文献[1],对于两用户的多输入多输出高斯干扰信道,发送端的天线数分别为 M_1、M_2 时,接收端的天线数分别为 N_1、N_2 时,其自由度为 $\min(M_1+M_2, N_1+N_2, \max(M_1, N_2), \max(M_2, N_1))$。

如图 3 所示,将 2 × 2 XC 合并为每节点天线数为 2 的点对点,与次网络每节点单天线点对点组成两用户 IC,即 $M_1=2, M_2=1, N_1=2, N_2=1$,由文献[1]得到此模型自由度为 2。由于节点合并不会降低原系统自由度,故其自由度上界为 2。与前面推导出的自由度下界相比,该上界较为宽松。

北航学报 赠阅

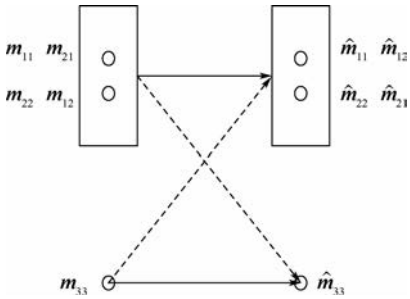


图 3　自由度上界分析图

Fig. 3　Analysis of upper bound on DOF

3　结　论

本文研究了 2×2 XC 与点对点信道并存时的网络模型,使用渐近干扰对齐方法分析了系统的自由度,提出了基于符号扩展的信号处理方案,证明可以达到的自由度为 5/3,并给出了参考自由度上界为 2。该自由度下界的证明可以为具体网络收发设计提供理论上的参考。

参考文献(References)

[1] CADAMBE V R, JAFAR S A. Interference alignment and degrees of freedom of the *K*-user interference channel[J]. IEEE Transactions on Information Theory, 2008, 54(8): 3425-3441.

[2] JAFAR S A, SHAMAI S. Degrees of freedom region of the MIMO X channel[J]. IEEE Transactions on Information Theory, 2008, 54(1): 151-170.

[3] YETIS C M, GOU T, JAFAR S A, et al. Feasibility conditions for interference alignment[C]// Global Telecommunications Conference, 2009 (GLOBECOM 2009). Piscataway, NJ: IEEE Press, 2009: 1-6.

[4] HUANG C, JAFAR S A. Degrees of freedom of the MIMO interference channel with cooperation and cognition[J]. IEEE Transactions on Information Theory, 2009, 55(9): 4211-4220.

[5] CADAMBE V R, JAFAR S A. Interference alignment and the degrees of freedom of wireless X networks[J]. IEEE Transactions on Information Theory, 2009, 55(9): 3893-3908.

[6] 周璐娜,刘锋,曾连荪. 两个点对点(PTP)网络并存时的自由度研究[J]. 计算机科学, 2015, 42(1): 110-112.
ZHOU L N, LIU F, ZENG L S. Research on degrees of freedom of two PTP networks coexistence[J]. Computer Science, 2015, 42(1): 110-112(in Chinese).

[7] KIM J, PARK S H, SUNG H, et al. Sum rate analysis of two-cell MIMO broadcast channels: Spatial multiplexing gain[C]// IEEE International Conference on Communications (ICC), 2010. Piscataway, NJ: IEEE Press, 2010: 1-5.

[8] SHIN W, LEE N, LIM J B, et al. On the design of interference alignment scheme for two-cell MIMO interfering broadcast channels[J]. IEEE Transactions on Wireless Communications, 2011, 10(2): 437-442.

[9] LEE N, SHIN W, HEATH R W, et al. Interference alignment with limited feedback for two-cell interfering MIMO-MAC[C]// International Symposium on Wireless Communication Systems (ISWCS). Piscataway, NJ: IEEE Press, 2012: 566-570.

[10] 周璐娜,刘锋,曾连荪. CSIT 对无线 MIMO 网络自由度的影响[J]. 计算机应用研究, 2014, 31(12): 3813-3817.
ZHOU L N, LIU F, ZENG L S. Influence of CSIT on degrees of freedom of wireless MIMO networks[J]. Application Research of Computer, 2014, 31(12): 3813-3817(in Chinese).

[11] TRESCH R, GUILLAUD M, RIEGLER E. On the achievability of interference alignment in the *K*-user constant MIMO interference channel[C]// IEEE/SP 15th Workshop on Statistical Signal Processing, 2009 (SSP'09). Piscataway, NJ: IEEE Press, 2009: 277-280.

[12] JAFAR S A, FAKHEREDDIN M J. Degrees of freedom for the MIMO interference channel[J]. IEEE Transactions on Information Theory, 2007, 53(7): 2637-2642.

[13] ABDOLI M J, GHASEMI A, KHANDANI A K. On the degrees of freedom of *K*-user SISO interference and X channels with delayed CSIT[J]. IEEE Transactions on Information Theory, 2013, 59(10): 6542-6561.

[14] LU Y, WEI Z. On degrees of freedom region of three-user MIMO interference channels[J]. IEEE Transactions on Signal Processing, 2015, 63(3): 590-603.

[15] GOU T G, JAFAR S A. Degrees of freedom of the *K* user $M \times N$ MIMO interference channel[J]. IEEE Transactions on Information Theory, 2010, 56(12): 6040-6057.

[16] LEE N, LIM J B, CHUN J. Degree of freedom of the MIMO Y channel: Signal space alignment for network coding[J]. IEEE Transactions on Information Theory, 2010, 56(7): 3332-3342.

作者简介:

郭倩倩　女,硕士研究生。主要研究方向:无线通信与网络。
E-mail: 1491538684@qq.com

刘锋　男,博士,讲师。主要研究方向:无线通信与网络。
Tel.: 021-38282863
E-mail: liufeng@shmtu.edu.cn

曾连荪　男,教授,硕士生导师。主要研究方向:定位导航系统
Tel.: 021-38282863
E-mail: lszeng@shmtu.edu.cn

北航学报 赠阅

Degrees of freedom of networks with coexistence of 2×2 X channel and point-to-point channel

GUO Qianqian, LIU Feng*, ZENG Liansun, QIN Wenming

(Department of Information Engineering, Shanghai Maritime University, Shanghai 201306, China)

Abstract: Degrees of freedom (DOF) is a spatial multiplexing conception, which was introduced in multiple antenna systems. And it characterizes the utilization efficiency of spatial resource for wireless channels. In the area of cognitive radio, since the primary and secondary networks interfere with each other, the DOF will be reduced. This paper mainly studies the DOF of the network on 2×2 X channel (XC) and point-to-point (PTP) channel which coexist. The result of this paper is also applied to the scenario when the XC has priority over PTP channel. If channel state information (CSI) is available at all transmitters, we propose the signal processing scheme using asymptotic interference alignment approach based on symbol extension model. It is shown that the lower bound of DOF tends to 5/3 when n tends to infinity. A propagation delay model is used to show its achievability. In addition, we analyze the upper bound on the DOF.

Key words: X channel (XC); point-to-point (PTP) channel; channel coexistence; asymptotic interference alignment; degrees of freedom (DOF)

北京航空航天大学学报

Received: 2015-10-27; **Accepted**: 2015-12-18; **Published online**: 2016-01-21 15:47

URL: www.cnki.net/kcms/detail/11.2625.V.20160121.1547.012.html

Foundation items: National Natural Science Foundation of China (61271283, 61472237, 61401270); Application Basic Research Project of Ministry of Transportation of China (2015329810030); Innovation Program of Shanghai Municipal Education Commission (14YZ113)

* **Corresponding author**. Tel.: 021-38282863 E-mail: liufeng@shmtu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0703

基于改进 ELECTRE 法的舰船气流场方案评估方法

张修远，常欣*，李想，杨放青

（哈尔滨工程大学 船舶工程学院，哈尔滨 150001）

摘 要：从气流场对舰载机作业的影响出发，构建了气流场评估指标体系。针对层次分析法一致性检验的缺陷，结合群决策特征根法（GEM）和决策熵理论，提出了结果可靠性可量化的气流场评估的赋权方法。结合模糊集理论知识，将经典 ELECTRE 法拓展到模糊环境下；排序计算中涉及到的指标评价值以三角模糊数的形式给出；使用改进 ELECTRE 法进行气流场方案的排序，模糊数计算保证了对大量模糊数据的完整处理。通过对 3 种不同气流场方案进行评估应用实例结果表明，所提出的舰船总体方案群决策方法具有较好的可行性，为舰船气流场方案评价提供了一种新的方法。

关 键 词：舰船气流场；指标体系；群决策特征根法（GEM）；决策熵；ELECTRE
中图分类号：U662.3
文献标识码：A **文章编号**：1001-5965(2016)11-2507-09

随着舰载机被广泛地应用于大型水面舰船上，舰载机在舰船上的作业能力越来越受到人们的关注，其中气流场的研究是一个重要的方面。大型水面舰船的气流场是指空气运动物理量在水面舰船周围的分布情况，其产生是由于海上的空气流过水面舰船的水上部分而改变原有的运动状态，在水面舰船周围形成了有其运动特征的流动状态。舰载机在舰船上作业，特别是起飞和降落作业时会穿过气流场的范围，而由于舰载机的运动依赖于其空气动力学特性，所以气流场会对其运动产生较大影响，会对其作业产生有利或不利因素。所以说，气流场是评估舰载机作业安全、有效的重要要素[1]。

近些年，随着计算流体动力学（Computational Fluid Dynamics，CFD）的进步，气流场的研究也愈加地全面。2002 年，Polsky[2] 对 LHA 两栖攻击舰空气尾流进行了计算，证明了 CFD 方法对于计算气流场的可行性。2007 年，赵维义和王占勇[3] 应用粒子图像测速技术获得了舰船空气尾流场特征，验证了上层建筑对飞行甲板上方区域的流场影响较大。2013 年，郜冶和刘长猛[4] 研究了风向变化产生的航母甲板涡结构特征。

而与此同时，对于气流场计算结果的分析评估还处于初级阶段。在舰船评估方法的方面，目前应用较多的是针对总体方案评估方法，其中的赋权方法都基本使用了层次分析法（Analytic Hierarchy Process，AHP），而排序方法有文献[5]提出的基于改进优劣解距离法（Technique for Order Preference by Similarity to an Ideal Solution，TOPSIS）的舰船总体方案群决策方法、文献[6]提出的基于面积和斜率灰色关联度的舰船方案群决策方法、文献[7]提出的基于相对熵排序的舰船总体方案评价方法等。这些方法都适用于舰船总体方案的评估，而对于气流场这种专家主观成分多、赋权结果一致性检验困难、打分结果模糊性大的评估对象并不适合。

本文提出一种能够全面、综合评价大型水面舰船气流场的方法，从而对不同大型水面舰船方

收稿日期：2015-10-29；**录用日期**：2015-12-18；**网络出版时间**：2016-01-21 15:47
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160121.1547.003.html
基金项目：国家自然科学基金（51309061，51379040）；中央高校基本科研业务费专项资金（HEUCFR150113）
* **通讯作者**：Tel.：0451-82568592 E-mail：changxin@hrbeu.edu.cn

引用格式：张修远，常欣，李想，等. 基于改进ELECTRE 法的舰船气流场方案评估方法[J]. 北京航空航天大学学报，2016，42(11)：2507-2515. ZHANG X Y，CHANG X，LI X，et al. Evaluation method of airflow field plan in warship based on improved ELECTRE method[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2507-2515 (in Chinese).

北航学报 赠阅

案在气流场方面做出优劣判断。本文在对不同舰船方案气流场进行 CFD 计算的基础上,针对气流场特征以及其影响对象舰载机的作业特点,建立评价气流场的指标体系,基于群决策特征根法(Group Eigenvalue Method,GEM)和决策熵理论给出各级指标的权重,使用改进模糊(ELECTRE,Elimination et Choise Translating Reality)法对不同舰船方案在气流场方面进行排序,从而建立多专家群决策的大型水面舰船气流场综合评估模型,得到对不同方案气流场方面的优劣判断,为今后气流场的总体研究与评价提供指导信息和理论依据。

1 气流场评估指标体系

舰船的气流场评估研究是一个大信息量、多目标的复杂的系统问题,几何空间巨大,涉及的物理量众多,如果考察所有涉及的物理量,不仅会耗费大量成本,而且也不尽科学。对于气流场的研究,应该紧紧围绕气流场的主要服务对象——固定翼舰载机来展开,以气流场对舰载机作业的影响为中心,建立气流场评估指标体系。

大型水面舰船在进行舰载机起降作业时,分为起飞和降落 2 种状态。在起飞状态时,由于舰载机在开始启动的一段时间里处于加速阶段,其与空气的相对速度较小,气流场对其的影响较小,所以在气流场评估时不考察这一阶段,而应考察舰载机在甲板上加速起飞的后半程以及舰载机离舰后的这一范围的气流场数据。其中,空气与舰载机的相对风速直接影响舰载机的升力和操纵性;侧风风速会使舰载机产生侧滑、横滚运动,对舰载机的安全起飞有负面影响;流场压力的变化以及涡量分布对舰载机的飞行稳定性有一定影响。所以,在评估起飞状态时的气流场采用相对风速、侧风风速、压力分布、涡量分布这 4 项底层指标进行评估[8-9]。需要注意的是,一般的大型水面舰船有多条起飞跑道,本文考察的是舰艏的 2 条起飞跑道,从右舷到左舷依次标注为 1 号、2 号起飞跑道。

在降落状态时,舰载机沿最后进场方向,即五边(final)进近,越过最低决断高度(minimum decision attitude)后保持飞行状态下降直至尾钩勾住阻拦索制动后完成降落,在这个过程中舰船气流场主要在舰载机越过最低决断高度后、尾钩勾住阻拦索前对舰载机有较大影响,所以在气流场评估时应考察这一范围的气流场数据。其中,空气与舰载机的相对风速直接影响舰载机的升力和操纵性,过低的相对风速容易使低速飞行的舰载机陷入失速状态;侧风风速会对舰载机产生侧滑、横滚运动,对舰载机的安全降落有负面影响;下洗速度会使舰载机在最后进近过程中掉高,对舰载机的安全降落有负面影响;涡量分布对舰载机的飞行稳定性有一定影响;流线的变化对舰载机的飞行稳定性有直接影响,如果流线变化剧烈,可能使舰载机的飞行状态发生突变,对降落产生不利影响。所以,在评估降落状态时的气流场采用相对风速、侧风风速、下洗速度、涡量分布和流线变化这 5 项底层指标进行评估[8]。

在 5 项的指标中,相对风速、侧风风速、下洗速度是定量指标,即可以由 CFD 计算直接得出客观数值,直接进行比较评估;压力分布、涡量分布、流线变化是定性指标,即不能直接由 CFD 计算得出客观数值,需要专家根据 CFD 计算得出的压力云图、涡量云图、流线图进行主观评判打分后再进行比较评估。气流场的评估范围如图 1 所示。综上,气流场评估指标体系如图 2 所示。同时对指标进行编号(如 B1、C1),便于之后的说明与计算。

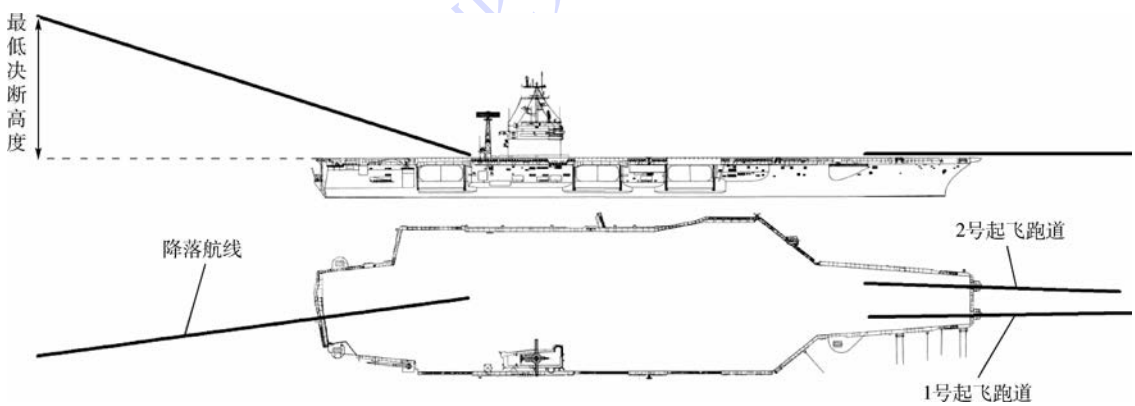


图 1 气流场评估范围示意图

Fig. 1 Schematic diagram of evaluation area of airflow field

北航学报 赠阅

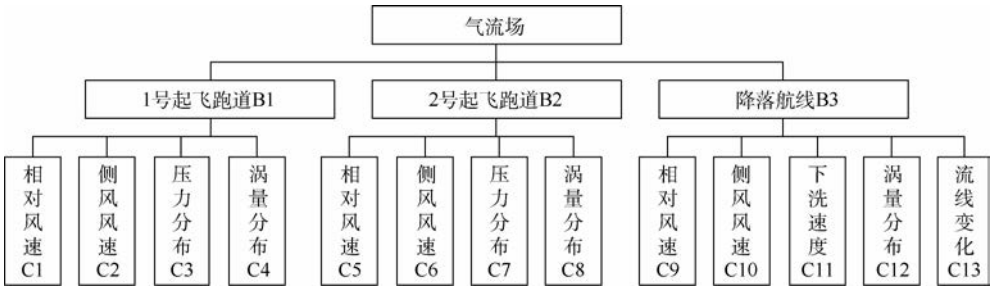


图 2　气流场评估指标体系

Fig. 2　Evaluation index system of airflow field

2　基于 GEM 与决策熵的赋权方法

由于指标体系中各个分量对于总体的重要性各异,所以必须对各个分量赋予权重,以计量各个分量的重要性,并用于之后的排序计算。

目前在舰船总体评估计算中,赋权方法一般采用层次分析法,并以 C. R. 为判据进行一致性检验[5,9]。在应用实践中,由于专家的意见各异,特别是像本文针对的大型水面舰船气流场评估这种研究处于起步阶段的赋权,专家很难达成趋于统一的意见,群赋权的一致性很难保证。并且一般用 0.1 作为一致性检验 C. R. 的临界值是相当粗略的,很难说有什么客观标准和意义,在理论上讲,把 C. R. 作为一致性检验是有缺陷的[10]。

针对这一问题,本文采用基于 GEM 和决策熵理论的赋权方法,将一致性与可靠性进行科学量化,让决策者可以直观地控制赋权结果的可靠性,特别针对大型水面舰船气流场评估的赋权,可以在量化可靠性的基础上得到可以接受的赋权结果。

2.1　基于 GEM 的群赋权计算法

假定有由 $S_1,S_2,\cdots,S_m$ 组成的 m 个专家群组决策系统 G,对 n 个指标 $B_1,B_2,\cdots,B_n$ 进行赋权,第 i 个专家 S_i 对第 j 个被赋权对象 B_j 的权重值记为 $x_{ij}\in[0,1]\ (i=1,2,\cdots,m;j=1,2,\cdots,n)$。$x_{ij}$ 的值越大,目标 B_j 的权重越高,在所有指标中的重要性越高。S_i 及其群组 G 的赋权结果组成 n 维列向量 $\boldsymbol{x}_i$ 和 $m\times n$ 阶矩阵 $\boldsymbol{x}$ 为

$$\begin{cases}\boldsymbol{x}_i=[x_{i1},x_{i2},\cdots,x_{in}]^{\mathrm{T}}\\ \boldsymbol{x}=[x_{ij}]_{m\times n}=\begin{bmatrix}x_{11}&\cdots&x_{1n}\\ \vdots& &\vdots\\ x_{m1}&\cdots&x_{mn}\end{bmatrix}\end{cases}\tag{1}$$

$\sum_{a=1}^{n}x_{ia}=1$ 为专家群组在一次赋权过程中所做的结论,代表各自对被赋权项的赋权值。

专家的赋权水平不仅取决于他的专业水平、经验、知识面和综合能力,而且与赋权时的精神状态、情绪和偏好密切相关。所以,现实中赋权可靠性达到最大值 1 的专家是不存在的。因此,假设一个赋权最准、最公正,即赋权水平最高的专家叫做理想专家 S_*,他的评分向量为

$$\boldsymbol{x}_*=[x_{*1},x_{*2},\cdots,x_{*n}]^{\mathrm{T}}$$

由于人们总是聘请水平较高的专家参与,所以人们现实地定义理想专家为对被赋权项的认识与专家群体 G 有最高一致性的专家,即 S_* 的赋权结论与 G 的完全一致,与专家个体间的差异最小。因此,理想专家 S_* 的赋权结果 $\boldsymbol{x}_*$ 就是要求出的这个群赋权事件中专家群体的赋权结果。而求在群赋权事件中的 $\boldsymbol{x}_*$ 可以通过求解方程 $\max\limits_{\|\boldsymbol{b}\|^2=1}\sum\limits_{i=1}^{m}(\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x}_i)^2=\sum\limits_{i=1}^{m}(\boldsymbol{x}_*^{\mathrm{T}}\boldsymbol{x}_i)^2=\rho_{\max}$ 得到。其中:$\rho_{\max}$ 为矩阵 $\boldsymbol{F}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$ 的最大正特征根;$\boldsymbol{x}_*$ 为 $\rho_{\max}$ 对应于 $\boldsymbol{F}$ 的正特征向量,且 $\|\boldsymbol{x}_*\|^2=1$。

因此,求得矩阵 $\boldsymbol{F}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$ 的特征向量 $\boldsymbol{x}_*$,之后将其中的各元素进行归一化处理后,就可以得到在这个群赋权事件中各个指标所得出的对应的权重[11]。

2.2　基于决策熵的群赋权可靠性分析方法

在群赋权事件中,由于专家及其群组在进行赋权时存在赋权水平差异,所以需要对于整体的赋权结果进行可靠性检验,而使用基于决策熵理论的可靠性模型可以很好地验证整体赋权的可靠性[12]。

将各已知因素单位化

$$\boldsymbol{D}_i=[d_{i1},d_{i2},\cdots,d_{in}]^{\mathrm{T}}\qquad i=1,2,\cdots,m\tag{2}$$

式中:$d_{ij}=x_{ij}/\sqrt{x_{i1}^2+x_{i2}^2+\cdots+x_{in}^2}$,$j=1,2,\cdots,n$。

再以 $\boldsymbol{N}_i=[N_{i1},N_{i2},\cdots,N_{in}]$ 表示按专家 $S_i(i=1,2,\cdots,m)$ 的赋权大小排序的被赋权目标 $B_1,B_2,\cdots,B_n$ 的大小名次。

则专家 S_i 的赋权水平向量为

$$\boldsymbol{E}_i=[e_{i1},e_{i2},\cdots,e_{in}]\tag{3}$$

式中:$e_{ij}=1-|N_{*j}-N_{ij}|-|d_{*j}-d_{ij}|$,$i=1,2,\cdots,m$,$j=1,2,\cdots,n$。

由此可以计算出专家 S_i 赋权结果的不准确性的度量——决策熵 H_i。H_i 等于 S_i 的赋权水平

北航学报 赠阅

向量各分量的广义熵之和,即

$$H_i = \sum_{j=1}^{n} h_{ij} \qquad i = 1,2,\cdots,m$$

$$h_{ij} = \begin{cases} -e_{ij}\ln e_{ij} & 1/\mathrm{e} \leqslant e_{ij} \leqslant 1 \\ \dfrac{2}{\mathrm{e}} - e_{ij}\left|\ln e_{ij}\right| & e_{ij} < 1/\mathrm{e} \end{cases} \tag{4}$$

而群组 G 的整体赋权水平测度函数为

$$H_G = \frac{1}{m}\sum_{i=1}^{m} H_i \tag{5}$$

式中:H_G 为群组决策熵,简称群组熵。

而通过决策熵值和可靠性的关系,就可以得出个体和群组在这个群赋权事件中的可靠性。

2.3 基于 GEM 和决策熵理论的赋权方法流程

在群赋权事件中,由于个体专家的赋权水平问题,仅仅通过使用 GEM 计算有时得不到令人满意的整体赋权结果,这时就需要进行修正。通过对个体和群组在这个群赋权事件中的可靠性分析,可以对群赋权结果进行修正和优化,从而得到令人满意的结果。基于 GEM 和决策熵理论的赋权方法流程图如图 3 所示。

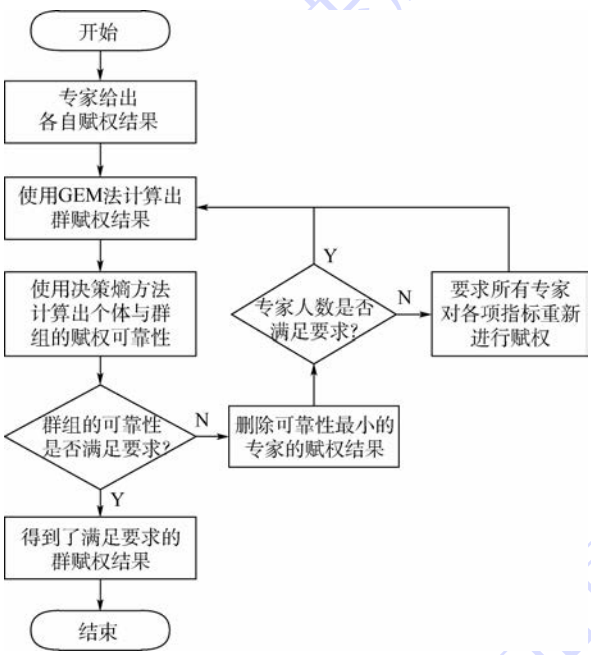


图 3 基于 GEM 和决策熵理论的赋权方法流程图

Fig. 3 Flowchart of weighting method based on GEM and theory of decision entropy

3 基于改进 ELECTRE 法的排序方法

3.1 基本原理

ELECTRE 是由 Roy 等在 20 世纪 60 年代提出的求解多属性决策问题的常见方法之一,这种方法使用简便,广泛应用于决策问题[13]。

ELECTRE 法实质上是一种消除与选择转化算法,即先淘汰部分备选方案,使决策者可把全部备选方案排列成序,从而选出最合理的方案。

传统 ELECTRE 法缺点在于不能直接处理不确定信息,在解决复杂的现实决策问题方面还有局限性。目前采用的模糊 ELECTRE 法用来计算全定性指标,并不能很好处理定性、定量混合指标体系。在本文中,对已有 ELECTRE 法进行改进,将其拓展到模糊环境下,同时将定性、定量指标一同处理,其中不确定信息以三角模糊数形式给出,确定信息以特殊三角模糊数形式给出,这样可在专家评价定性指标时给予更多的自由空间而不仅局限于传统的优良评价,同时结合定量指标计算,可更好地解决航母气流场评价这种有较多定性指标的模糊多属性群决策问题[14-15]。

3.2 计算步骤

第 1 步 每位专家给出对所有方案与指标的模糊评价矩阵,即

$$\boldsymbol{X}^k = (\tilde{x}_{ij}^k)_{m\times n} = \begin{array}{c} \\ C_1 \\ \vdots \\ C_m \end{array}\begin{array}{c} \begin{array}{ccc} A_1 & \cdots & A_n \end{array} \\ \begin{bmatrix} \tilde{x}_{11}^k & \cdots & \tilde{x}_{1n}^k \\ \vdots & & \vdots \\ \tilde{x}_{m1}^k & \cdots & \tilde{x}_{mn}^k \end{bmatrix} \end{array} \quad k = 1,2,\cdots,K$$

其中:$A_1, A_2, \cdots, A_n$ 为 n 个被排序的方案;$C_1, C_2, \cdots, C_m$ 为方案中用于分析比较的 m 个独立指标;k 为专家的编号,总共有 K 个专家参与评价。对于定量指标,评价结果直接使用实际数值,用特殊三角模糊数,即 $q = (a,a,a)$ 表示;对于定性指标,需要专家进行主观评判,专家对指标的评价 $\tilde{x}_{ij}^k$ 使用三角模糊数表示,即 $\tilde{x}_{ij}^k = (x_{ijl}^k, x_{ijm}^k, x_{ijr}^k)$,$x_{ijl}^k$、$x_{ijr}^k$ 和 x_{ijm}^k 分别为专家认为的最低值、最高值和最有可能的值,3 个值都在[0,1]上表示。

第 2 步 构造模糊群决策矩阵,令

$$\tilde{x}_{ij} = \frac{\tilde{x}_{ij}^1 + \tilde{x}_{ij}^2 + \cdots + \tilde{x}_{ij}^k}{K} = \left(\frac{1}{K}\sum_{k=1}^{K} x_{ijl}^k, \frac{1}{K}\sum_{k=1}^{K} x_{ijm}^k, \frac{1}{K}\sum_{k=1}^{K} x_{ijr}^k\right) \tag{6}$$

这样可以得到模糊群决策矩阵 $\tilde{\boldsymbol{X}}$,即

$$\tilde{\boldsymbol{X}} = (\tilde{x}_{ij})_{m\times n} = \begin{array}{c} \\ C_1 \\ \vdots \\ C_m \end{array}\begin{array}{c} \begin{array}{ccc} A_1 & \cdots & A_n \end{array} \\ \begin{bmatrix} \tilde{x}_{11} & \cdots & \tilde{x}_{1n} \\ \vdots & & \vdots \\ \tilde{x}_{m1} & \cdots & \tilde{x}_{mn} \end{bmatrix} \end{array} \tag{7}$$

第 3 步 将矩阵 $\tilde{\boldsymbol{X}}$ 规范化为 $\tilde{\boldsymbol{R}} = (\tilde{r}_{ij})_{m\times n}$。这里,将所有的指标值进行横向比较,将原有

北航学报 赠阅

的有量纲数都化为数值在 0 到 1 之间的无量纲数,这个无量纲数即为对该指标的评价值,以便式(8)对不同指标之间的比较计算。

$$\begin{cases}\tilde{r}_{ij}=\left(\dfrac{x_{ijl}}{x_{ir}^{+}},\dfrac{x_{ijm}}{x_{ir}^{+}},\dfrac{x_{ijr}}{x_{ir}^{+}}\right),x_{ir}^{+}=\max\limits_{i}(x_{ijr}) & C_i\ \text{为效益型属性}\\ \tilde{r}_{ij}=\left(\dfrac{x_{il}^{-}}{x_{ijr}},\dfrac{x_{il}^{-}}{x_{ijm}},\dfrac{x_{il}^{-}}{x_{ijl}}\right),x_{il}^{-}=\min\limits_{i}(x_{ijl}) & C_i\ \text{为成本型属性}\end{cases} \tag{8}$$

第 4 步　进行模糊和谐性检验,令

$$h_{jk}=\sum_{i\in I^{+}}w_i \tag{9}$$

式中:h_{jk}为和谐性指数;$I^{+}=\{i\mid\tilde{r}_{ji}>\tilde{r}_{ki}\}$,$i=1,2,\cdots,n$;$w_i$ 为各指标的权重值,三角模糊数之间使用 R_S 准则进行比较。由此构造出方案之间的和谐性指数矩阵:

$$\boldsymbol{H}=\begin{bmatrix}0 & h_{12} & \cdots & h_{1n}\\ h_{21} & 0 & \cdots & h_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ h_{n1} & h_{n2} & \cdots & 0\end{bmatrix}$$

第 5 步　进行模糊非和谐性检验,令

$$d_{jk}=\max_{i\in I^{-}}\{\mid R_S(\tilde{r}_{ji})-R_S(\tilde{r}_{ki})\mid w_i\} \tag{10}$$

式中:$I^{-}=I\backslash I^{+}$,$I=\{1,2,\cdots,n\}$。

第 6 步　建立一致性比较矩阵,令

$$\boldsymbol{C}=\begin{bmatrix}0 & c_{12} & \cdots & c_{1n}\\ c_{21} & 0 & \cdots & c_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ c_{n1} & c_{n2} & \cdots & 0\end{bmatrix} \tag{11}$$

式中:

$$c_{jk}=\begin{cases}0 & d_{jk}\geqslant q\\ h_{jk} & \text{其他}\end{cases}\qquad \forall j\neq k$$

这里 q 为否决门槛值,c_{jk}体现了在总体上 A_j 级别高于 A_k 的程度。以此为判据,可以衡量个方案的优劣,进而确定各方案的次序关系。

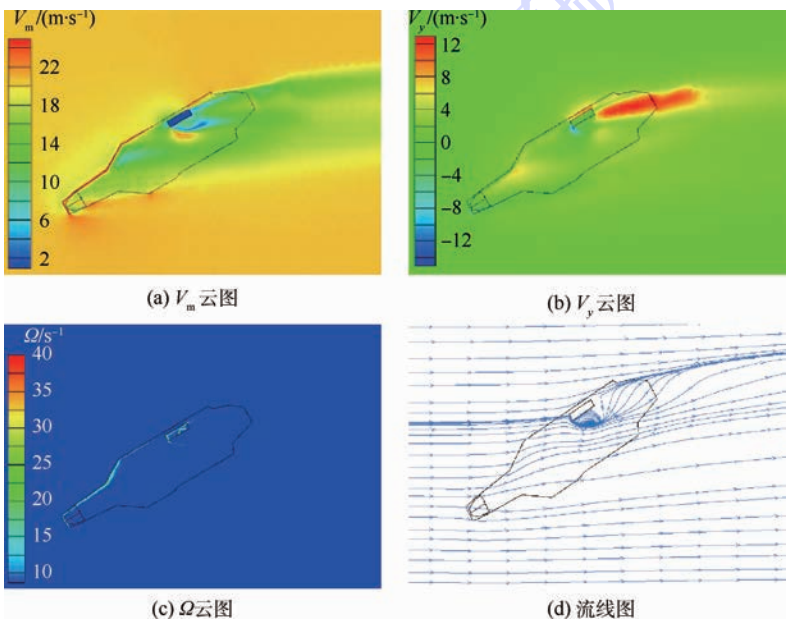
4　应用实例

现通过赋权与排序方法对 3 个大型水面舰船气流场方案进行评估。3 个设计方案包括了核动力弹射起飞方案、常规动力弹射起飞方案以及常规动力滑跃起飞方案,以验证评估模型的广泛适用性。在指标体系中涉及的定量指标数值如表 1 所示。将 3 个方案的设置为典型工况,即迎风 30°、40 kn 风速,按照已有的 CFD 方法对 3 种方案进行数值计算[16-17],计算结果如图 4 ~ 图 6 所示,分别给出了 3 种方案的合速度 V_m 云图、侧风风速 V_y 云图、涡量 Ω 云图和流线图。

表 1　各方案定量指标计算结果

Table 1　Calculation results of quantitative index in each plan

指标	方案 1	方案 2	方案 3
C1	20.6	16.5	17.0
C2	2.92	4.27	2.29
C5	19.3	14.8	15.2
C6	0.68	1.52	0.49
C9	13.9	15.1	14.4
C10	5.33	5.82	2.51
C11	0.14	1.39	0.06

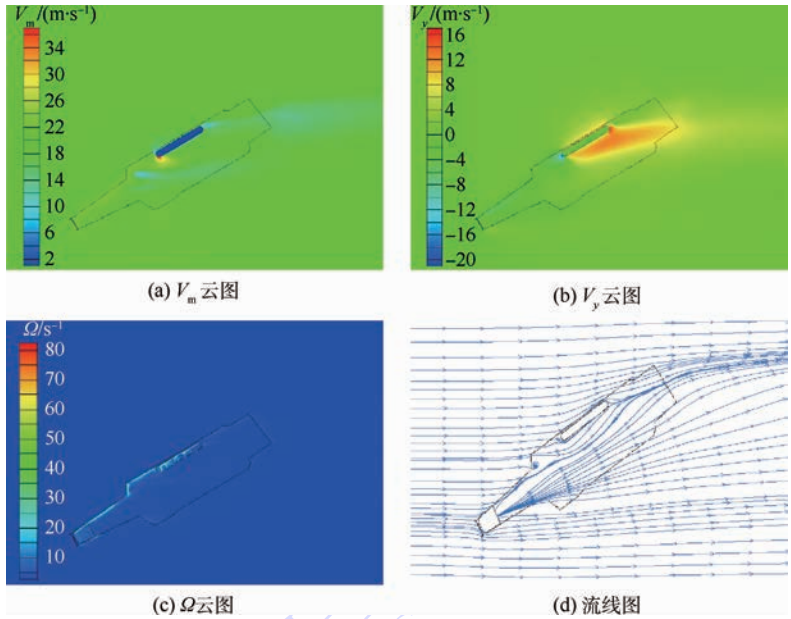


(a) V_m云图　(b) V_y云图

(c) Ω云图　(d) 流线图

图 4　方案 1 的 V_m、V_y 和 Ω 云图及流线图

Fig. 4　Contours of V_m, V_y and Ω and streamline diagram in plan 1

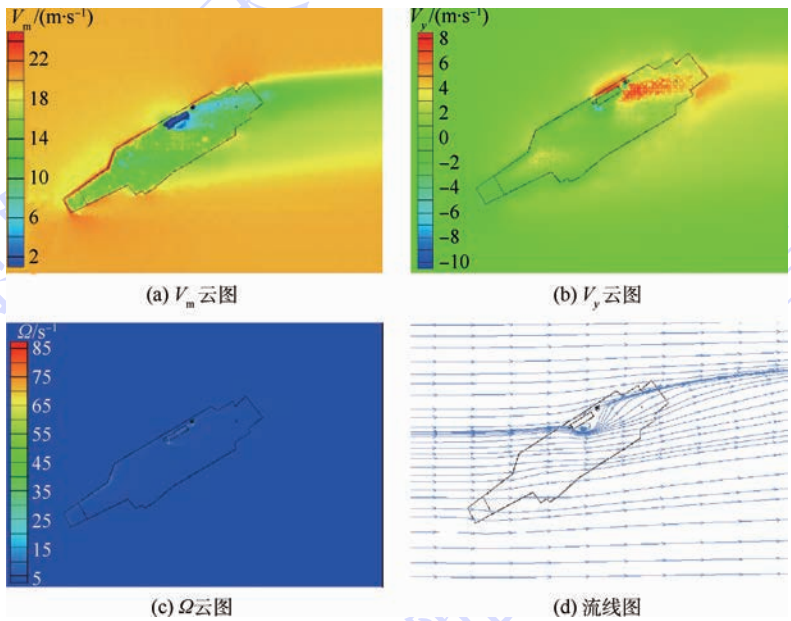


(a) V_m 云图　(b) V_y 云图

(c) Ω云图　(d) 流线图

图 5　方案 2 的 V_m、V_y 和 Ω 云图及流线图

Fig. 5　Contours of V_m, V_y and Ω and streamline diagram in plan 2



(a) V_m 云图　(b) V_y 云图

(c) Ω云图　(d) 流线图

图 6　方案 3 的 V_m、V_y 和 Ω 云图及流线图

Fig. 6　Contours of V_m, V_y and Ω and streamline diagram in plan 3

首先邀请 5 位专家对指标体系的各层、各部分进行赋权，同时设定可以接受的群决策可靠性为 60%，需要有半数以上的有效赋权。以 3 个一级指标 B1、B2、B3 为例，专家的赋权结果见表 2。

表 2　一级指标的专家赋权结果

Table 2　Results of weighting for indexes in the first class by experts

专家编号	B1	B2	B3
专家 1	0.30	0.30	0.40
专家 2	0.33	0.33	0.34
专家 3	0.35	0.35	0.30
专家 4	0.40	0.30	0.30
专家 5	0.33	0.33	0.34

由此得

$$\boldsymbol{x}=\begin{bmatrix}0.30 & 0.33 & 0.35 & 0.40 & 0.33\\ 0.30 & 0.33 & 0.35 & 0.30 & 0.33\\ 0.40 & 0.34 & 0.30 & 0.30 & 0.34\end{bmatrix}^{\mathrm{T}}$$

$$\boldsymbol{F}=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}=\begin{bmatrix}0.5903 & 0.5503 & 0.5694\\ 0.5503 & 0.5203 & 0.5394\\ 0.5694 & 0.5394 & 0.5712\end{bmatrix}$$

求得 $\boldsymbol{F}$ 的特征向量并经过标准化后得到群赋权结果 $\boldsymbol{x}^*=(0.3345, 0.3345, 0.3111)$，再经过式(2)～式(5)计算得群组决策熵 $\boldsymbol{H}_C=0.8061$，查表得群赋权可靠性为 65%，符合要求。所以 3 个一级指标 B1、B2、B3 的权重为

北航学报 赠阅

0.334 5、0.334 5、0.311 1。

同理,计算底层权重,经过标准化处理后得到的底层权重如表 3 所示。

表 3　底层权重结果

Table 3　Results of weighting for indexes in the second class

指标	权重	指标	权重
C1	0.118 1	C8	0.049 1
C2	0.093 8	C9	0.056 2
C3	0.073 4	C10	0.091 5
C4	0.049 1	C11	0.068 3
C5	0.118 1	C12	0.035 6
C6	0.093 8	C13	0.059 5
C7	0.073 4		

5 位专家对定性指标主观评判在按照式(6)处理后得到群体三角模糊数评判值,见表 4。

表 4　群体的三角模糊数评判值

Table 4　Group evaluation value of triangular fuzzy number

指标	方案 1	方案 2	方案 3
C3	(0.68,0.80,0.88)	(0.56,0.68,0.74)	(0.66,0.76,0.86)
C4	(0.58,0.68,0.78)	(0.46,0.56,0.66)	(0.62,0.74,0.80)
C7	(0.76,0.84,0.88)	(0.62,0.74,0.82)	(0.72,0.80,0.86)
C8	(0.76,0.86,0.96)	(0.54,0.66,0.70)	(0.74,0.82,0.90)
C12	(0.60,0.72,0.82)	(0.50,0.62,0.72)	(0.60,0.70,0.82)
C13	(0.44,0.58,0.64)	(0.66,0.74,0.84)	(0.38,0.54,0.62)

加上定量指标的数值,按照式(7)共同构成群模糊决策矩阵 $\tilde{\boldsymbol{X}}$:

$$\tilde{\boldsymbol{X}}=\begin{bmatrix}(20.6,20.6,20.6) & (16.5,16.5,16.5) & (17.0,17.0,17.0)\\ (2.92,2.92,2.92) & (4.27,4.27,4.27) & (2.29,2.29,2.29)\\ (0.68,0.80,0.88) & (0.56,0.68,0.74) & (0.66,0.76,0.86)\\ (0.58,0.68,0.78) & (0.46,0.56,0.66) & (0.62,0.74,0.8)\\ (19.3,19.3,19.3) & (14.8,14.8,14.8) & (15.2,15.2,15.2)\\ (0.68,0.68,0.68) & (1.52,1.52,1.52) & (0.49,0.49,0.49)\\ (0.76,0.84,0.88) & (0.62,0.74,0.82) & (0.72,0.80,0.86)\\ (0.76,0.86,0.96) & (0.54,0.66,0.7) & (0.74,0.82,0.90)\\ (13.9,13.9,13.9) & (15.1,15.1,15.1) & (14.4,14.4,14.4)\\ (5.33,5.33,5.33) & (5.82,5.82,5.82) & (2.51,2.51,2.51)\\ (0.14,0.14,0.14) & (1.39,1.39,1.39) & (0.06,0.06,0.06)\\ (0.60,0.72,0.82) & (0.50,0.62,0.72) & (0.60,0.70,0.82)\\ (0.44,0.58,0.64) & (0.66,0.74,0.84) & (0.38,0.54,0.62)\end{bmatrix}$$

接着由式(8)得到规范化的群决策矩阵,再代入各指标的权重值,根据式(9)和式(10)分别进行和谐性检验和非和谐性检验,取否决门槛值 $q=0.5$,最终由式(11)可以得到一致性比较矩阵 $\boldsymbol{C}$:

$$\boldsymbol{C}=\begin{bmatrix}0 & 0.280\,6 & 0.291\\ 0 & 0 & 0\\ 0 & 0.280\,6 & 0\end{bmatrix}$$

根据矩阵 $\boldsymbol{C}$ 得到的方案的优先次序为:方案 1 优于方案 3 优于方案 2,可见方案 1 在气流场方面为最优方案。

5　结　论

本文针对大型水面舰船气流场评估问题:

1) 以气流场对舰载机作业的影响为中心,建立了气流场评估指标体系。

2) 分析之前方案评估所用方法对气流场评估的适用性的基础上,提出了基于 GEM 和决策熵理论的赋权方法。

3) 基于改进模糊 ELECTRE 法的方案排序方法,提出了舰船气流场方案排序方法。

4) 建立了一套完善的大型水面舰船气流场评估体系,并通过对 3 种不同的大型水面舰船设计方案的气流场进行 CFD 计算并经过完整的评估计算,验证了所提出的评估体系的可行性,证明了所提出的评估模型具有广泛的适用性。

在实际应用中,提出的评估模型可进行一定程度上的简化和调整,如对群决策可靠性的取值范围、否决门槛值参数等,使其满足实际应用要求。

参考文献 (References)

[1] 贺少华,刘东岳,谭大力,等. 载机舰船气流场相关研究综述[J]. 舰船科学技术,2014,36(2):1-7.
HE S H,LIU D Y,TAN D L,et al. A review of researches on ship airwakes[J]. Ship Science and Tchonology,2014,36(2):1-7(in Chinese).

[2] POLSKY S A. A computational study of unsteady ship airwake:AIAA-2002-1022[R]. Reston:AIAA,2002.

[3] 赵维义,王占勇. 舰船空气尾流场对直升机着舰的影响研究[J]. 海军航空工程学院学报,2007,22(7):435-438.
ZHAO W Y,WANG Z Y. Research on the effect of warship air-

北航学报 赠阅

wake on helicopter landing on shipboard [J]. Journal of Naval Aeronautical Engineering Institute, 2007, 22(7): 435-438 (in Chinese).

[4] 郜冶, 刘长猛. 护卫舰气流场数值计算研究[J]. 哈尔滨工程大学学报, 2013, 34(5): 599-603.
GAO Y, LIU C M. Numerical calculation of frigate ship airwake [J]. Journal of Harbin Engineering University, 2013, 34(5): 599-603 (in Chinese).

[5] 黄胜, 郭海鹏, 侯远杭, 等. 基于改进 TOPSIS 法的舰船总体方案群决策方法[J]. 上海交通大学学报, 2014, 48(1): 92-97.
HUANG S, GUO H P, HOU Y H, et al. Group decision-making method of warship based on improved TOPSIS [J]. Journal of Shanghai Jiao Tong University, 2014, 48(1): 92-97 (in Chinese).

[6] 黄胜, 郭海鹏, 王超. 基于面积和斜率灰色关联度的舰船方案群决策方法[J]. 船舶工程, 2014, 36(1): 100-103.
HUANG S, GUO H P, WANG C. Group decision-making of warship scheme based on area and slope grey degree[J]. Ship Engineering, 2014, 36(1): 100-103 (in Chinese).

[7] 侯远杭, 黄胜, 胡玉龙, 等. 基于相对熵排序的舰船总体方案评价方法[J]. 上海交通大学学报, 2012, 46(8): 1218-1222.
HOU Y H, HUANG S, HU Y L, et al. Evaluation of warship based on relative entropy method[J]. Journal of Shanghai Jiao Tong University, 2012, 46(8): 1218-1222 (in Chinese).

[8] POLSKY S A, NAYLOR S. CVN airwake modeling and Integration: Initial steps in the creation and implementation of a virtual burble for F-18 carrier landing simulations: AIAA-2005-6298 [R]. Reston: AIAA, 2005.

[9] 廖全蜜, 黄胜, 王宇, 等. 基于改进 TOPSIS 法的舰船通道人机效能研究[J]. 西南交通大学学报, 2015, 50(3): 536-542.
LIAO Q M, HUANG S, WANG Y, et al. Research of human performance in warship passageway based on improved TOPSIS method[J]. Journal of Southwest Jiaotong University, 2015, 50(3): 536-542.

[10] 李学全, 李松仁, 韩旭里. AHP 理论与方法研究: 一致性检验与权重计算[J]. 系统工程学报, 1997, 12(2): 111-117.
LI X Q, LI S R, HAN X L. Study of AHP theory and method: Test of consistency and calculation of weight vector[J]. Journal of Systems Engineering, 1997, 12(2): 111-117 (in Chinese).

[11] 邱菀华, 刘美芳. 管理决策及其应用[M]. 北京: 机械工业出版社, 2012: 131-137.
QIU W H, LIU M F. Managerial decision making and its application[M]. Beijing: China Machine Press, 2012: 131-137 (in Chinese).

[12] 马俊, 邱菀华, 张浩. 招投标决策模型及其应用[J]. 北京航空航天大学学报, 2000, 26(4): 470-472.
MA J, QIU W H, ZHANG H. The bidding decision model and its application[J]. Journal of Beijing University of Aeronautics and Astronautics, 2000, 26(4): 470-472 (in Chinese).

[13] 张吉军. 模糊层次分析法(FAHP)[J]. 模糊系统与数学, 2002, 14(2): 80-88.
ZHANG J J. Fuzzy analytical hierarchy process (FAHP) [J]. Fuzzy Systems and Mathematics, 2002, 14(2): 80-88 (in Chinese).

[14] 高璟, 张强. 基于 ELECTRE 方法的模糊多属性群决策[J]. 数学实践与认识, 2009, 39(9): 21-27.
GAO J, ZHANG Q. Extensions of the ELECTRE for group decision-making under fuzzy environment[J]. Mathematics in Practice and Theory, 2009, 39(9): 21-27 (in Chinese).

[15] XU Z, CHEN J. An interactive method for fuzzy multiple attribute group decision making[J]. Journal of Information Science, 2007, 177(1): 248-263.

[16] 陆超, 姜治芳, 王涛. 利用缩比模型 CFD 数值模拟计算舰面空气流场相似准则数的影响探讨[J]. 中国舰船研究, 2008, 3(6): 45-48.
LU C, JIANG Z F, WANG T. Discussion on comparability of scaled models for CFD numerical simulation for ship airwake [J]. Chinese Journal of Ship Research, 2008, 3(6): 45-48 (in Chinese).

[17] REDDY K, TOFFOLETTO R, JONES K. Numerical simulation of ship airwake [J]. Computer & Fluids, 2000, 29(4): 451-465.

作者简介:

张修远 男, 硕士研究生。主要研究方向: 舰船总体。
Tel.: 0451-82519824
E-mail: lyyg-zxy@163.com

常欣 男, 博士, 副教授, 硕士生导师。主要研究方向: 舰船总体与船舶推进。
Tel.: 0451-82568592
E-mail: changxin@hrbeu.edu.cn

北航学报 赠阅

Evaluation method of airflow field plan in warship based on improved ELECTRE method

ZHANG Xiuyuan, CHANG Xin*, LI Xiang, YANG Fangqing

(College of Shipbuilding Engineering, Harbin Engineering University, Harbin 150001, China)

Abstract: Based on the influence of airflow field on deck-landing aeroplane operations, an evaluation index system for airflow field of warship is constructed. In view of the defect of consistency check by analytic hierarchy process (AHP), a weighting method which is combined with group eigenvalue method (GEM) and the theory of decision entropy is formed; the method can obtain outcomes, whose reliability is quantifiable. Combining the concepts of fuzzy theory, the classical elimination et choice translating reality (ELECTRE) method is extended for fuzzy multiattribute group decision making; all attribute values in sorting are expressed in triangular fuzzy numbers; when airflow field plans of warship are sorted with improved fuzzy ELECTRE, the use of fuzzy number ensures the complete processing of fuzzy data which is from the experts. Evaluation-computing is run for three airflow field plans, which verifies the feasibility of this improved evaluation method that can be a new method for general airflow field plan evaluation.

Key words: airflow field in warship; index system; group eigenvalue method (GEM); decision entropy; elimination et choice translating reality (ELECTRE)

北京航空航天大学学报

Received: 2015-10-29; **Accepted**: 2015-12-18; **Published online**: 2016-01-21 15:47

URL: www.cnki.net/kcms/detail/11.2625.V.20160121.1547.003.html

Foundation items: National Natural Science Foundation of China (51309061, 51379040); the Fundamental Research Funds for the Central Universities (HEUCFR150113)

* **Corresponding author**. Tel.: 0451-82568592　E-mail: changxin@hrbeu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0736

电火花穿孔加工中管状电极振动幅值特性

于志强[1], 佟浩[2,3,*], 李勇[2,3], 柴永生[1], 王志强[3], 李朝将[3]
(1. 烟台大学 机电汽车工程学院, 烟台 264005; 2. 清华大学 摩擦学国家重点实验室, 北京 100084;
3. 清华大学 机械工程系 精密超精密制造装备及控制北京市重点实验室, 北京 100084)

摘 要: 在采用管状电极高压冲液的电火花穿孔加工中,较大幅值的电极随机振动将导致加工中侧壁短路和二次放电,这对穿孔加工精度和效率造成不利影响。本文以高压冲液管状电极为研究对象,采用仿真、实验验证及对比分析的方法研究了电极微观振动幅值特性。基于 ANSYS 有限元软件平台,建立了描述电极振动特性的有限元模型,通过数值模拟分析了冲液压力、导向器间隙、穿孔深度、放电间隙、主轴转速、电极悬长多因素对电极振动幅值的影响,分析了多因素耦合情况下各因素对电极振动幅值的影响程度,并进一步进行了实验验证。研究表明导向器间隙、穿孔深度、冲液压力与电极振动幅值成正相关性,减小导向器间隙由 10 μm 至 1 μm,可有效降低孔径误差 20 μm。多因素耦合分析得到电极振动幅值影响因素主次顺序为:穿孔深度、冲液压力、放电间隙、主轴转速、电极悬长。此项研究为优化电火花穿孔加工工艺与主轴设计提供依据。

关 键 词: 电火花加工 (EDM); 穿孔成形; 管状电极; 振动特性; 流固耦合
中图分类号: V261.6+1; G661
文献标识码: A **文章编号**: 1001-5965(2016)11-2516-08

利用管状电极高压冲液的电火花穿孔加工,可及时更新工作液并快速排出电蚀产物,通过细长电极旋转和轴向伺服进给,可实现深小孔的高效批量化加工。此高速穿孔工艺已广泛应用于航空、航天、模具和液压等领域难加工材料的小孔加工。比如,已应用于航空发动机涡轮叶片上大量的气膜冷却孔加工[1]。

传统的高速穿孔加工采用长电极后推式进给,细长电极的低刚度会产生加工中弯曲变形问题,从而对孔批量化加工精度和效率造成不利影响。针对此问题,清华大学已研发出一种采用夹头夹在电极前端的双夹子蠕动式进给、高精度旋转及高压内冲液的复合功能主轴[2]。但前期实验表明:在加工较大深径比的小孔时,孔深和冲液压力等因素变化对孔径精度和加工效率有明显影响。分析其原因认为:由于电极具有一定的悬伸长度,在较高压力中空冲液扰动的作用下,加工中电极易产生不规则振动现象。这不仅直接影响孔加工尺寸形状精度,而且造成的侧壁放电或短路将使电极频繁回退,也降低孔的加工效率。

目前,高速穿孔加工研究还主要集中在工艺参数优化选择、电极和工件材料特性影响、电极损耗规律等方面[3-5],缺乏对高压中空冲液引起的电极扰动规律研究,国内外研究仅对这种扰动现象进行了一些报道。比如,2013 年郭维城等试验研究发现随着孔深增加,电极振动现象加剧[6]。然而,为优化高速穿孔工艺过程,特别是对于加工精度和效率有较高要求的较大深径比小孔批量化

收稿日期: 2015-11-09; **录用日期**: 2015-12-18; **网络出版时间**: 2016-01-19 10:00
网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20160119.1000.001.html
基金项目: 国家自然科学基金 (51305230); 北京高等学校青年英才计划项目 (YETP0084); 清华大学自主科研计划课题 (20131089209)
* **通讯作者**: Tel.: 010-62796337 E-mail: tonghao@mail.tsinghua.edu.cn

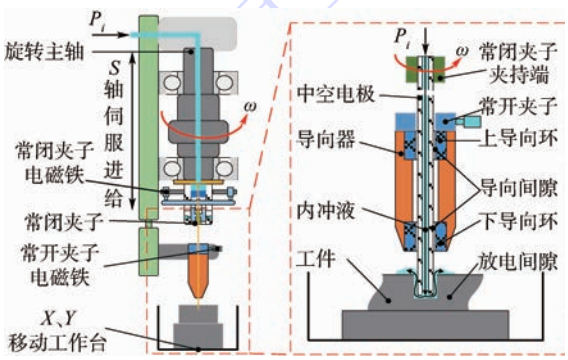
引用格式: 于志强, 佟浩, 李勇, 等. 电火花穿孔加工中管状电极振动幅值特性[J]. 北京航空航天大学学报, 2016, 42(11): 2516-2523. YU Z Q, TONG H, LI Y, et al. Vibration amplitude characteristics of tubular electrode in EDM drilling[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(11): 2516-2523 (in Chinese).

加工（比如，航空发动机涡轮叶片气膜冷却孔加工），研究电极扰动规律及其影响因素具有较重要意义和实用价值。

本研究通过建立中空管状电极内冲液物理模型及流固耦合数值仿真模型，并借助 ANSYS 有限元仿真分析软件平台，综合考虑中空冲液压力、穿孔深度、电极悬长、主轴转速、放电间隙及导向间隙等多参数影响，研究高速穿孔过程中电极振动幅值的影响规律，并通过实验验证仿真结果及规律的正确性。

1　穿孔加工中空管状电极物理及仿真建模

如图 1 所示，物理及数值仿真建模过程是基于最新开发的双夹子蠕动式进给主轴。该主轴可实现电极旋转、高压内冲液及蠕动进给功能，通过协调控制常开夹子、常闭夹子及 S 轴进给运动，实现电极蠕动进给补偿电极损耗，并采用精密导向器约束加工中电极的旋转径跳误差，导向器设计为具有上下导向环的结构来保证导向长度并减少摩擦面积[2]。



P_i—冲液压力；ω—主轴转速。

图 1　高速电火花穿孔加工主轴结构

Fig. 1　Spindle structure used in hole machining by high-speed EDM

根据工作液的流动状态，流场模型求解器设置为耦合、三维、定常流及绝对压力求解算法。放电间隙中工作液流动速度较快，通过雷诺数计算得出工作液流动状态属于湍流。因此数值仿真采用标准的 k-ε 模型[7-9]，湍动能 k 和耗散率 ε 的方程为

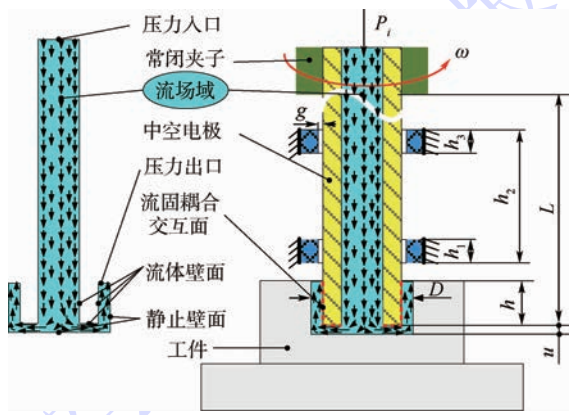
$$\frac{\partial(\rho k)}{\partial t}+\frac{\partial(\rho k u_i)}{\partial x_i}=\frac{\partial}{\partial x_j}\left[\left(\mu+\frac{\mu_t}{\sigma_k}\right)\frac{\partial k}{\partial x_j}\right]+G_k-\rho\varepsilon \tag{1}$$

$$\frac{\partial(\rho\varepsilon)}{\partial t}+\frac{\partial(\rho\varepsilon u_i)}{\partial x_i}=\frac{\partial}{\partial x_j}\left[\left(\mu+\frac{\mu_t}{\sigma_\varepsilon}\right)\frac{\partial\varepsilon}{\partial x_j}\right]+\frac{C_{1\varepsilon}\varepsilon}{k}G_k-C_{2\varepsilon}\rho\frac{\varepsilon^2}{k} \tag{2}$$

式中：$\rho=1.0\ \mathrm{g/cm^3}$ 为工作液密度；$\sigma_k=1.0$、$\sigma_\varepsilon=1.3$ 为湍流动能和耗散率的普朗特常数；$C_{1\varepsilon}=1.44$、$C_{2\varepsilon}=1.92$ 为经验常数；μ_t 为湍流黏度；u_i 为时均速度；G_k 为由于平均速度梯度引起的湍动能量；t 为时间；x_i 和 x_j 为坐标张量；μ 为流体动力黏度。

鉴于单向流固耦合中的弱耦合方式易于收敛并满足物理模型（图 1）仿真要求，因此采用单向流固耦合求解方式，依次进行流场计算和结构力学计算。

根据图 1 主轴结构参数及 k 和 ε 的方程，建立电极振动特性分析的数值仿真物理模型如图 2 所示。模型中包含了导向间隙 g、穿孔深度 h、P_i、ω、放电间隙 u 和电极悬长 L 等重要参数。设置边界条件为：定义大气压力作用在放电间隙出口；流场计算中运动边界为流体壁面；结构力学计算中运动边界为旋转电极。为实现流固耦合面数据完全传递，设置流体壁面与中空电极旋转速度大小、方向相同。模型相关参数为：中空冲液为纯净水、电极材料为黄铜、电极中空内径为 ϕ0.24 mm、电极外径为 ϕ0.5 mm。



D—小孔直径；$h_1=2$ mm—下导向环厚度；$h_2=30$ mm—导向器长度；$h_3=2$ mm—上导向环厚度。

图 2　数值仿真物理模型及边界条件

Fig. 2　Numerical simulation of physical model and boundary conditions

2　中空冲液电极振动特性数值仿真分析

数值仿真借助 Pro/E 和 ANSYS Workbench 软件平台中 Fluid Structure Interaction（FSI）单向流固耦合模块，开展中空冲液电极振动幅值特性的仿真分析研究。划分网格时使流场计算与结构力学计算网格交互面节点 100% 重合，避免因节点不重合造成的数据传递误差[10]。基于考虑工作液高速高压流动、并且及时的冷却作用，忽略非接触式火花放电瞬间较小作用力、瞬时放电温度

北航学报 赠阅

对电极的影响。

仿真流程如图 3 所示，首先通过 Pro/E 建立几何模型并导入 FSI 模块中，将结构力学计算模型和流体模型网格化且保证网格节点重合。然后在 Fluent 中设定模型求解器、能量方程、湍流模型、流体几何属性、边界条件及耦合参数，在 Mechanical中设定边界条件和映射激励。最后设置残差系数并初始化求解环境，进行流固耦合求解得到流场压力和中空电极最大变形量。

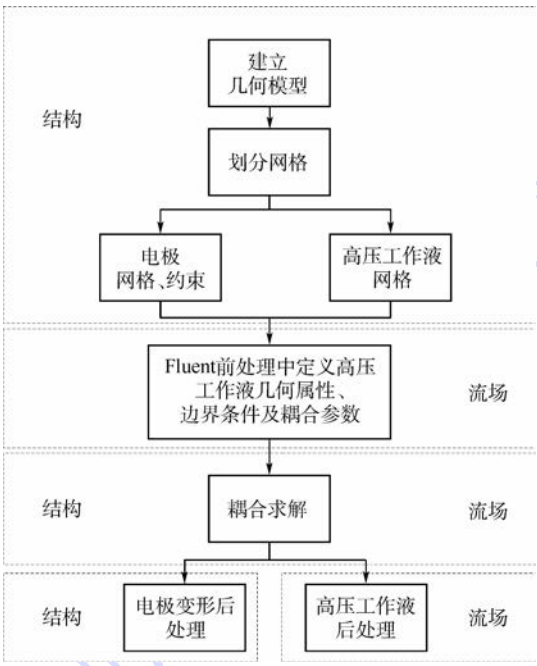


图 3 仿真分析流程图

Fig. 3 Flowchart of simulation analysis

数值仿真预期得到各个关键单因素、以及多因素耦合对电极振动特性的影响规律，因此首先仿真关键单因素影响，然后正交仿真研究多因素耦合影响规律。采用极差分析、方差分析的统计方法，表征各因素优化规律的可信度。

2.1 关键因素对电极振动特性影响规律

在给定条件下研究 P_i、g、h、ω 和 u 等关键因素对电极振动特性的影响规律，利用电极最大变形量表征其影响程度。关键因素仿真参数如表 1 所示，对电极振动幅值 A 的影响规律曲线如图 4 所示。

表 1 关键因素仿真参数

Table 1 Simulation parameters of key factors

关键因素	仿真参数
P_i/MPa	1,3,5,7,9
g/μm	1,5,10
h/mm	1,2,3,6
ω/(r·min^{-1})	0,400,800,1 600,3 200
u/μm	20,40

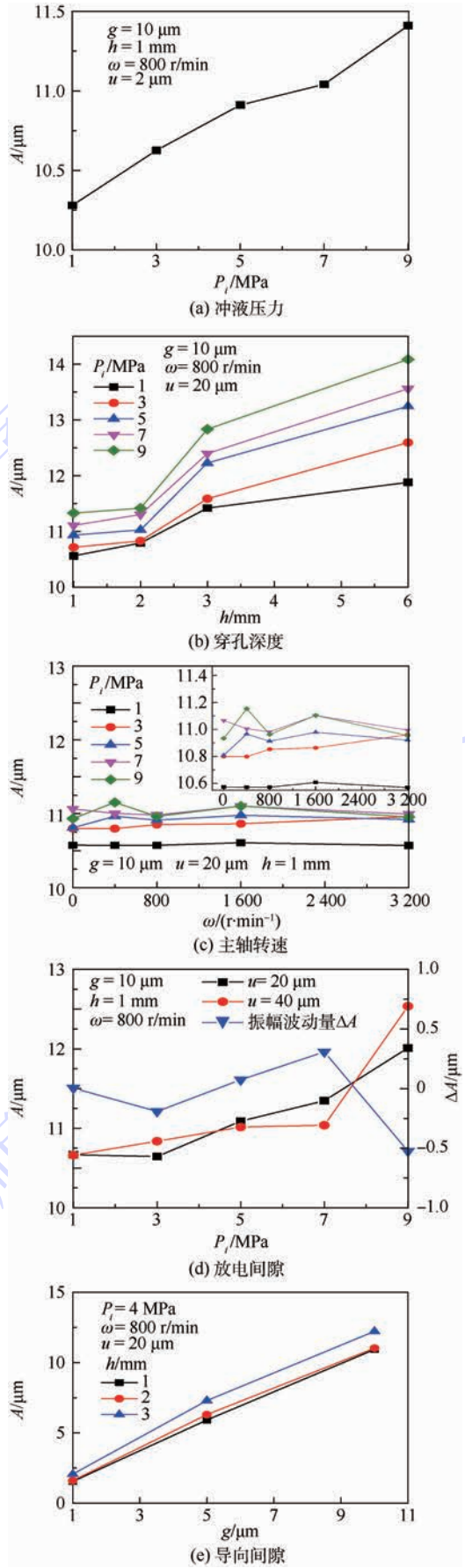


图 4 关键因素与电极振动幅值关系曲线

Fig. 4 Relation curves of key factors and electrode vibration amplitude

北航学报 赠阅

2.1.1　冲液压力对电极振动幅值的影响

电极振动幅值随 P_i 的变化规律如图 4(a)所示,电极振动幅值随冲液压力 P_i 的增加逐渐增大,且大于导向器间隙;当 P_i = 1 MPa 时,电极振动幅值即达到 10.5 μm(导向间隙 10 μm 时),但 P_i 增加到 9 MPa 时,其振动幅值为 11.42 μm,即电极振动幅值波动量小于 1.0 μm。这表明若要实现较好的排屑效果(P_i > 1 MPa),必然会产生一定的振动幅值(> 10 μm),而 P_i 变化(继续增大时)对电极振动的影响较小。分析认为,放电间隙中工作液流经电极前端面时受小孔底面阻挡,工作液速度的重新分布加剧了工作液主流中的内部摩擦(主流湍动性增加),进而加大了对电极扰动作用,但导向器的约束作用使振幅在较小范围内变化。

2.1.2　穿孔深度对电极振动幅值的影响

如图 4(b)所示,电极振动幅值随 h 增加而增加,且增长速率随 P_i 增加而加快;,当 P_i 保持1 MPa 不变时,h 从 1 mm 增加到 6 mm 时,振动幅值从 10.5 mm 增大到 11.5 μm。但当 P_i 达到 9 MPa 时,穿孔深度 6 mm 时振动幅值甚至达到 14 μm,即由此产生的孔径误差可达 28 μm。在较深孔加工中为实现有效排屑,孔越深冲液压力要求越高,这时随着孔深增加而产生的电极振动将对孔径误差影响较大。分析认为,这是由于电极悬伸长度增加降低了电极挠曲强度,使电极抵抗弯曲变形能力下降造成的。

2.1.3　主轴转速对电极振动幅值的影响

图 4(c)结果显示,ω 在 0 ~ 3 200 r/min 范围内变化时,即使 P_i = 9 MPa 时电极振动幅值变化量仍小于 0.3 μm,即主轴转速变化对电极振动的影响甚至可忽略。分析认为,由于电极壁面较光滑,转动时对工作液湍动性影响微弱。

2.1.4　放电间隙对电极振动幅值的影响

图 4(d)结果显示,随着 P_i 增加,不同放电间隙(u = 20,40 μm)曲线交替上升,即表明 10 μm 量级的加工间隙变化对电极振动幅值影响无明显规律。可以认为在较高中空冲液压力条件下,10 μm量级放电间隙对工作液流动状态影响较小,进而对电极振动幅值影响微弱。

2.1.5　导向间隙对电极振动幅值的影响

根据图 4(e)结果,导向间隙对电极振动幅值产生显著影响。分析 h = 3 mm 的情况:g = 1 μm 时振动幅值 2.4 μm,而 g = 10 μm 时振动幅值达到 12.5 μm,即由此产生的孔径误差可达 25 μm。这是由于工作液湍流扰动使电极产生了较大的变形,导向器约束振动幅值限制在一定的范围内。为证实分析的正确性,进一步对电极与导向器之间的接触压力进行分析如图 5 所示,当 P_i = 4 MPa,g = 20 μm, h = 1 mm 时存在接触压力 0.046 MPa,即导向器对电极具有约束作用。因此,实际加工中应尽可能减小导向器间隙,以降低电极振动幅值,提高孔径尺寸精度。

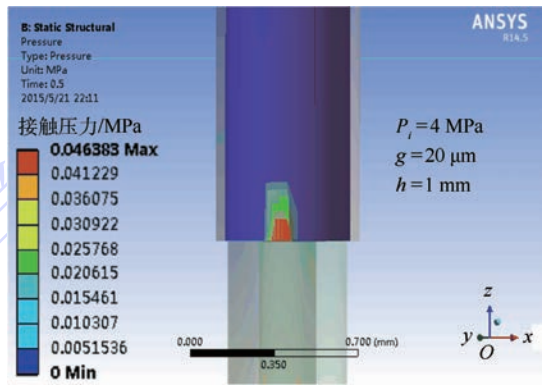


图 5　电极与导向器接触压力分布

Fig. 5　Contact pressure distribution of electrode and guide apparatus

2.2　多因素耦合对电极振动特性影响规律

进而通过正交仿真研究多因素耦合作用效果,为指导加工工艺提供依据。根据实际加工经验选取加工参数范围,如表 2 所示采用 5 因素 5 水平进行正交数值仿真研究,导向间隙根据实际情况取值为 10 μm。

表 2　多因素耦合水平正交表

Table 2　Level orthogonal array of multi-factor coupling

水平	仿真因素				
	P_i/MPa	h/mm	L/mm	ω/(r·min^{-1})	u/μm
1	1	1	83	200	20
2	2	2	84	400	30
3	3	3	85	600	40
4	4	4	86	800	50
5	5	5	87	1 000	60

通过数据统计分析方法评价多因素影响权重,均值与极差分析结果如图 6 所示。得到优化水平依次为 P_i(1)、h(1)、L(5)、ω(4)、u(3),即得到一组优化组合为:P_i = 1 MPa,h = 1 mm,L = 87 mm、ω = 800 r/min、u = 40 μm。极差分析中,较大数值说明该因素对电极振动幅值影响能力较大,由极差分析可得对电极振动幅值影响的主次顺序为:穿孔深度、冲液压力、放电间隙、主轴转速、电极悬长。

为验证数值仿真结果的可信度,对其进行方差分析,引入误差列 e 进行结果有效性评估[11-12]。

北航学报 赠阅

方差分析结果如图 7 所示,因素的偏差平方和均大于 e 的偏差平方和,可认为数值仿真结果有效。为得到因素的显著性水平,查询 F 分布表知$F_{0.05}$ = 6.390,由 F_h = 11.730(F_h 为 h 的 F 统计量)且满足 $F_h > F_{0.05}$,则有 95% 的把握认为穿孔深度 h 对电极振动有显著的影响。同理,由 F_{P_i} = 5.134(F_{P_i}为 P_i 的 F 统计量)且满足 $F_{0.05} > F_{P_i} > F_{0.1}$,则有 90% 的把握认为冲液压力 P_i 对电极振动影响显著。同时 L、ω、u 的 F 统计量均小于 $F_{0.1}$,说明这些因素对电极振动的影响不显著。通过方差分析 F 统计量大小得到因素主次排列顺序为:穿孔深度、冲液压力、放电间隙、主轴转速、电极悬长。与极差分析结果一致,表明了数值仿真结果的准确性。

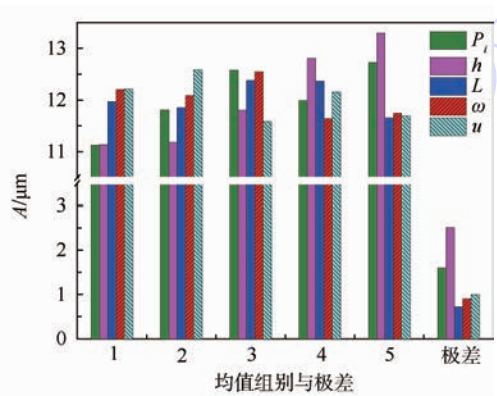


图 6 正交仿真均值与极差分析

Fig. 6 Mean and range analysis of orthogonal simulation

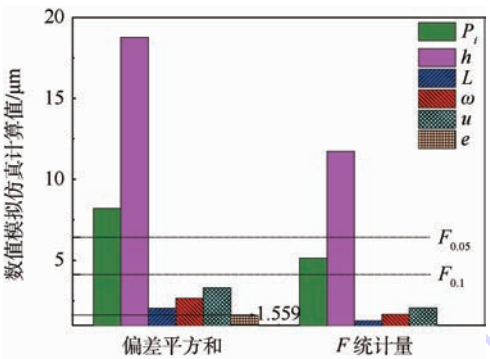


图 7 正交仿真方差分析

Fig. 7 Variance analysis of orthogonal simulation

3 电火花穿孔加工过程实验验证

受到实验系统限制,实验验证无法达到仿真研究所覆盖的参数范围。但为验证仿真结果的正确性,根据实验机床能达到的系统参数,开展内冲液压力、穿孔深度和主轴转速的趋势性验证实验。基本实验条件如表 3 所示,不同因素的参数变化范围如表 4 所示。采用中空电极(内径 240 μm、外径 500 μm、长度 400 mm)、高压内冲液方式,每个验证实验参数下分别加工出 3 个通孔。根据通孔直径与电极振动幅值具有正相关性,通过测量多个小孔直径取平均值的方法,实验得到各因素与电极振动幅值的关系。

在 h = 6 mm、ω = 140 r/min 参数条件下,依次进行不同内冲液压力下对比加工实验。加工出的孔出、入口及典型通孔脱模显微镜测量结果如图 8所示,入口直径为 ϕ553 ~ 571 μm,对应出口直径为 ϕ552 ~ 568 μm。

表 3 电火花穿孔加工实验基本参数

Table 3 Basic parameters of EDM drilling processing experiment

参数	数值
电压/ V	90
峰值电流/ A	10
脉宽/ μs	24
脉间/ μs	40

注:工件材料为 Inconel 738;电极材料为黄铜;电极类型为中空管状;工作液为纯净水。

表 4 实验验证参数变化范围

Table 4 Parameters range of verification experiments

参数	范围
P_i/ MPa	1,1.5,2,2.5,3,3.5,4
h/ mm	2,3,4,6
ω/ (r · min^{-1})	100,120,140,160,180

图 8(a)为其他加工参数相同、内冲液压力分别为 P_i = 4 MPa 与 P_i = 1 MPa 所对应的孔入口图,图 8(b)为与之对应的出口图。通过图 8 的孔脱模测量可知:与 1 MPa 冲液压力下加工孔径相比,4 MPa 情况下通孔直径较大且最大相差 19.7 μm。通过此测量方法,得到孔径平均值 $\bar{D}$ 随 P_i 变化曲线如图 9 所示,此变化趋势与仿真结果一致,即随着内冲液压力增加而电极振动幅值增大(孔径变大),但孔径变化数值大于仿真振动幅值。从穿孔加工过程来看这样结果也是合理的,因为孔径误差不仅来源于电极振动幅值几何因素,而且振动还会造成更多的二次放电,进一步扩大了加工的孔径。此研究表明在保证有效排屑前提下,适当地减小内冲液的压力有利于提高孔径尺寸精度。

在 P_i = 4 MPa、ω = 140 r/min 参数条件下,依次进行不同穿孔深度的对比加工实验。穿孔深度为 2 mm 的阵列孔及典型通孔脱模测量图如图 10 所示。对比图 8(c)与图 10(b)可得,通孔平均值由穿孔深度 2 mm 的 ϕ557 μm 增加至穿孔深度为 6 mm 的 ϕ569 μm,孔径平均值最大相差 12 μm。

孔径平均值随穿孔深度变化曲线如图 11 所示，此实验结果与仿真结果趋势一致，即随着加工深度增加而电极振动幅值增大（孔径变大），但孔径变化数值仍大于仿真振动幅值。

在 $P_i=4$ MPa、$h=6$ mm 参数条件下，依次进行不同主轴转速参数下对比实验结果如图 12 所示，孔径随主轴转速增加呈无规律的震荡状态，且孔径平均值最大差值小于4 μm。这样结果与

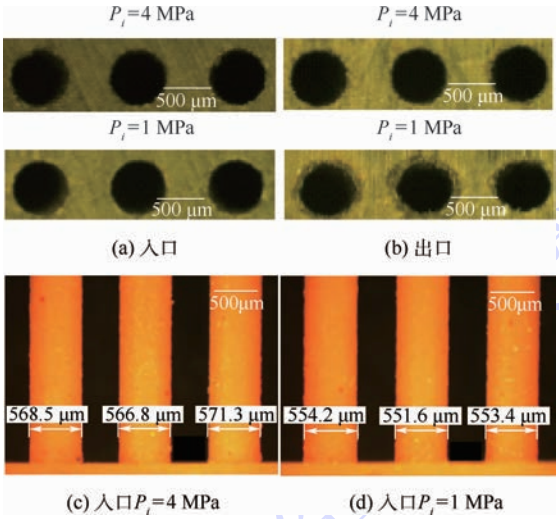


图 8　$h=6$ mm 时阵列孔出、入口及典型通孔脱模尺寸图

Fig. 8　Inlets, outlets and injection models of array holes at $h=6$ mm

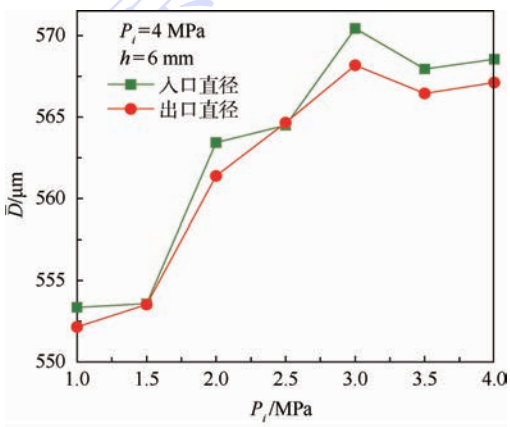


图 9　冲液压力与孔径平均值曲线

Fig. 9　Curve of hydraulic pressure of dielectric fluid and mean of aperture

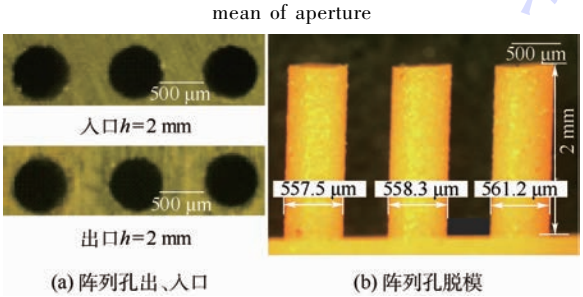


图 10　$h=2$ mm 时阵列孔出、入口及典型通孔脱模尺寸图

Fig. 10　Inlets, outlets and injection models of array holes at $h=2$ mm

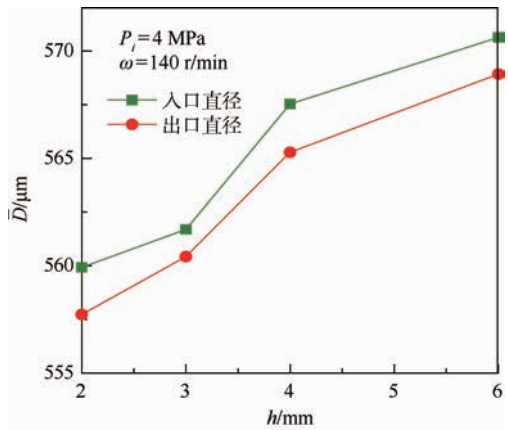


图 11　穿孔深度与孔径平均值曲线

Fig. 11　Curve of drilling depth and mean of aperture

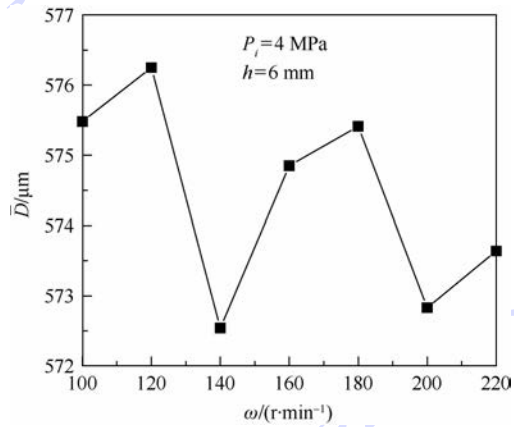


图 12　主轴转速与孔径平均值曲线

Fig. 12　Curve of spindle speed and mean of aperture

仿真趋势一致得出：在采用中空冲液的高速电火花穿孔加工中，无需采用过高的主轴转速，这可降低主轴的设计难度，有利于加工系统的稳定运行。

4　结　论

通过流固耦合建模及数值仿真表明，降低冲液压力、减小穿孔深度、减小导向器间隙有利于降低电极振动幅值，提高孔径尺寸复制精度。且在导向器间隙一定的实际情况下，多因素耦合数值仿真分析得到影响电极振动幅值因素显著性排序为：穿孔深度、冲液压力、放电间隙、主轴转速、电极悬长。这从有效控制电极扰动振动的角度，得到改善加工过程和提高孔径复制精度的依据。得出如下主要结论：

1）通过对比中空冲液电极振动仿真分析和多孔加工实验结果可得，穿孔过程加工实验结果与电极振动数值仿真结果趋势相吻合，但由于孔径误差不仅来源于电极振动幅值几何因素，实验结果中孔径变化数值大于数值仿真振动幅值。

2）在采用中空冲液的高速电火花穿孔加工

北航学报 赠阅

中,导向器间隙对电极振动幅值影响最为显著,通过提高导向器导向精度,可有效降低电极振动幅值。且由于主轴转速与电极悬长对电极振动幅值较小的影响作用,因此无需采用过高的主轴转速,这可降低主轴的设计难度,有利于加工系统的稳定运行。

3) 采用脱模观测孔径尺寸方法,实验得到冲液压力从 4 MPa 降至 1 MPa 时孔径尺寸可减小 19.7 μm,即穿孔深度一定时,在保证有效排屑的前提下,适当减小内冲液压力有利于提高孔径尺寸精度。

参考文献(References)

[1] 叶树林,刘晋春. 高速电火花小孔加工机理的研究[J]. 电加工,1997(4):15-18.
YE S L, LIU J C. The research on machining mechanism of high speed small hole EDM drilling[J]. Electromachining & Mould, 1997(4):15-18(in Chinese).

[2] 王志强,佟浩,李勇,等. 气膜冷却孔电火花加工用复合功能主轴[J]. 清华大学学报(自然科学版),2014,54(9):1131-1137.
WANG Z Q, TONG H, LI Y, et al. Multifunctional spindle for machining film cooling holes in EDM[J]. Journal of Tsinghua University (Science and Technology), 2014, 54(9): 1131-1137(in Chinese).

[3] 刘正勋,陈祥康. 深小孔的高速电火花加工工艺[J]. 机械科学与技术,1989(5):5-8.
LIU Z X, CHEN X K. The machining technology of EDM on deep hole[J]. Mechanical Science and Technology, 1989(5): 5-8(in Chinese).

[4] YILMAZ O, OKKA M A. Effect of single and multi-channel electrodes application on EDM fast hole drilling performance [J]. The International Journal of Advanced Manufacturing Technology, 2010, 51(1-4): 185-194.

[5] PHAM D T, IVANOV A, BIGOT S, et al. An investigation of tube and rod electrode wear in micro EDM drilling [J]. The International Journal of Advanced Manufacturing Technology, 2007, 33(1-2): 103-109.

[6] 郭维城,郭永丰,侯朋举,等. 电火花加工阵列穿丝孔的试验研究[J]. 机械设计与制造,2013 (2):226-228.
GUO W C, GUO Y F, HOU P J, et al. Study on the array wire threading hole test in electrical-discharge process[J]. Machinery Design and Manufacture, 2013(2): 226-228(in Chinese).

[7] 贾振元,郑新毅,王福吉,等. 微孔电火花加工极间工作液流动状态研究[J]. 大连理工大学学报,2010,50(2):188-193.
JIA Z Y, ZHEN X Y, WANG F J, et al. Research on flow state of interelectrode working fluid in micro-EDN[J]. Journal of Dalian University of Technology, 2010, 50(2): 188-193(in Chinese).

[8] 张淑奎,王燕青,黄河,等. 旋转电极内冲液电火花铣削加工流场仿真及实验研究[J]. 电加工与模具,2011(1):1-5.
ZHANG S K, WANG Y Q, HUANG H, et al. Research on rotation electrode and iner jetted dielectric EDM milling flow simulation and experiment[J]. Electromachining & Mould, 2011(1): 1-5(in Chinese).

[9] 唐家鹏. Fluent14.0 超级学习手册[M]. 北京:北京人民邮电出版社,2013:30-38.
TANG J P. The study guide of Fluent14.0[M]. Beijing: Posts & Telecom Press, 2013: 30-38(in Chinese).

[10] 黄志新. ANSYS Workbench 14.0 超级学习手册[M]. 北京:人民邮电出版社,2013:505-520.
HUANG Z X, The study guide of ANSYS Workbench 14.0[M]. Beijing: Posts & Telecom Press, 2013: 505-520(in Chinese).

[11] 茆诗松. 实验设计[M]. 2 版. 北京:中国统计出版社,2012:111-123.
MAO S S. Design of experiments[M]. 2nd ed. Beijing: China Statistics Press, 2012: 111-123(in Chinese).

[12] 何为. 优化实验设计方法及数据分析[M]. 北京:化学工业出版社,2012:29-57.
HE W. Optimization design of experiment method and data analysis[M]. Beijing: Chemical Industry Press, 2012: 29-57(in Chinese).

作者简介:

于志强 男,硕士研究生。主要研究方向:特种加工工艺与装备。

Tel.: 010-62796339

E-mail: yzq_09@163.com

佟浩 男,博士,副研究员,硕士生导师。主要研究方向:特种加工工艺与装备。

Tel.: 010-62796337

E-mail: tonghao@mail.tsinghua.edu.cn

北航学报 赠阅

Vibration amplitude characteristics of tubular electrode in EDM drilling

YU Zhiqiang [1], TONG Hao[2,3,*], LI Yong[2,3], CHAI Yongsheng[1], WANG Zhiqiang [3], LI Chaojiang[3]

(1. School of Electromechanical & Automobile Engineering, Yantai University, Yantai 264005, China;
2. State Key Laboratory of Tribology, Tsinghua University, Beijing 100084, China;
3. Beijing Key Lab of Precision/Ultra-precision Manufacturing Equipments and Control, Department of Mechanical Engineering, Tsinghua University, Beijing 100084, China)

Abstract: In the drilling of electric discharge machining (EDM) by use of a tubular electrode with high-pressure dielectric fluid, the random electrode vibration causes short circuit and second discharge which have a bad effect on machining accuracy and efficiency. In this paper, the microscopic amplitude characteristics of the electrode vibration are researched by adopting methods of simulations, experiments and comparative analysis. Based on ANSYS simulation platform of finite-element software, a finite element model was established to simulate the vibration characteristics of the tubular electrode, and then the effects of multiple factors including hydraulic pressure of dielectric fluid, electrode-guide gap, drilling depth, discharge gap, spindle speed and electrode overhanging length on the electrode vibration amplitude were analyzed. In addition, the significant level of every factor was analyzed by numerical simulation under the condition of coupling these factors. The EDM drilling experiments were carried out to verify the simulation results. Our research shows that the guide gap, drilling depth and hydraulic pressure have positive correlation with the electrode vibration amplitude, and reducing the guide gap from 10 μm to 1 μm can effectively reduce 20 μm of the error of the aperture. By analyzing these coupling factors, the order of significance affecting the electrode vibration is drilling depth, hydraulic pressure of dielectric fluid, discharge gap, spindle speed and electrode overhanging length. This research provides a basis for optimizing machining process and spindle design of EDM drilling using a tubular electrode.

Key words: electric discharge machining (EDM); drilling forming; tubular electrode; vibration characteristic; fluid-structure interaction

Received: 2015-11-09; **Accepted**: 2015-12-18; **Published online**: 2016-01-19 10:00
URL: www.cnki.net/kcms/detail/11.2625.V.20160119.1000.001.html
Foundation items: National Natural Science Foundation of China (51305230); Beijing Higher Education Young Elite Teacher Project (YETP0084); Independent Scientific Research Project of Tsinghua University (20131089209)

* **Corresponding author**. Tel.: 010-62796337　E-mail: tonghao@mail.tsinghua.edu.cn

北京航空航天大学学报

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2016.0184

基于优化信任评估的 Ad Hoc 安全路由协议

陈少华，樊晓光*，禚真福，黄金科，孙贤明

（空军工程大学 航空航天工程学院，西安 710038）

摘 要：针对 Ad Hoc 网络中自私节点造成的攻击问题，提出了一种基于信任评估的 Ad Hoc 安全路由协议 TEAR。该协议在节点行为信任评估方案的基础上，通过使用链路信任值和路径信任值表示节点和路径的可信度，同时，在链路信任值计算时，引入了基于公共邻居节点的间接信任评估方法和可变时间窗机制，保证了信任值计算的准确性和实效性；协议中每个节点维护一张信任表，并重新定义了路由请求和路由回复消息格式，实现了安全路径的快速选择。使用 NS2 进行模拟验证，又与 AODV 以及基于固定时间窗信任评估的安全路由协议 TA-AODV 进行性能对比。结果表明，改进的信任评估方案具有更好的性能，采用该方案的安全路由协议能较快地发现自私节点，并准确选择安全路径，提高了 Ad Hoc 网络性能。

关 键 词：Ad Hoc 网络；自私节点；安全路由；信任评估；可变时间窗；节点行为
中图分类号：TP393.08
文献标识码：A **文章编号**：1001-5965(2016)11-2524-08

Ad Hoc 网络是一种不依赖任何固定基础设施的无线移动网络，具有无中心、自组织、高动态和能量受限等特点。在战场环境、紧急救援和移动会议中具有广泛的应用前景，同时其安全问题也更加突出，网络抗攻击能力尤为脆弱，因此需要设计有效的安全路由算法[1-2]。

Ad Hoc 网络中，由于能量和带宽受限，会产生自私节点，这些节点会部分或者全部地丢弃接收到的分组，使网络安全性能降低[3]。现有的安全路由协议，如 SAODV[4]、Ariadne[5] 和 ARAN[6] 等通过完整性校验、签名认证和 RSA 密码体制来满足安全需求，但是存在计算开销大、无法发现不协作的自私节点的缺陷。针对自私节点，近年来，国内外研究者提出了很多基于信任评估方法的安全路由协议。信任评估方法的优点是计算开销小，模型运用灵活，能在整个网络范围内检测节点的信任值。文献[7]基于信任保留的认证协议提高了节点认证的可靠性，但是增加了信任评估方法的复杂度。文献[8]构造了一个新的信任模型，但在路由过程中没有确切的实现。文献[9-10]基于信任评估的安全路由协议都假设节点行为保持不变或者忽略节点行为的变化，无法检测如通断攻击[11]节点行为的改变，从而会出现自私节点的误判和安全路径的错误选择。文献[12]在判断节点是否自私时使用了固定时间窗机制，对节点行为的判断存在延迟。文献[13]中的协议虽然能够检测节点行为的变化，但是在计算信任值时存在证据不足的问题。

针对上述问题，本文提出了一种基于节点行为检测的 Ad Hoc 安全路由协议信任评估方案 TESBB，并将其整合到 AODV 协议中。在信任评估方案中，引入间接信任评估来计算链路信任值(link trust)，每个节点通过比较前后链路信任值的变化来判断节点的行为改变。当发现某个邻居

收稿日期：2016-03-09；**录用日期**：2016-04-22；**网络出版时间**：2016-05-13 11:36
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160513.1136.001.html
基金项目：国家自然科学基金（61402517）；中航工业集团预研基金（619010601）
* **通讯作者**：Tel.：029-84787655 E-mail：1203574700@qq.com

引用格式：陈少华，樊晓光，禚真福，等. 基于优化信任评估的 Ad Hoc 安全路由协议[J]. 北京航空航天大学学报，2016，42(11)：2524-2531. CHEN S H，FAN X G，ZHUO Z F，et al. Secure routing protocol based on optimized trust evaluation in Ad Hoc network[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2524-2531 (in Chinese).

节点的行为发生改变时,通过可变时间窗来重新评估链路信任值;在路由过程中,每个节点的路由表中维护一个路径信任值,并通过控制分组的计算和路由表的查询来选择安全的分组传送路径。该方案满足了 Ad Hoc 安全路由协议快速发现自私节点和准确选择安全路径的需求。

1　基本理论

1.1　公共邻居

如果一个节点同时处在另外两个不同节点的传输范围内,那么这个节点就称为公共邻居节点。如图 1 所示,在 Ad Hoc 网络中,节点 A 和节点 H 既是节点 E 的邻居,又是节点 F 的邻居,则节点 A 和节点 H 就被称为公共邻居。

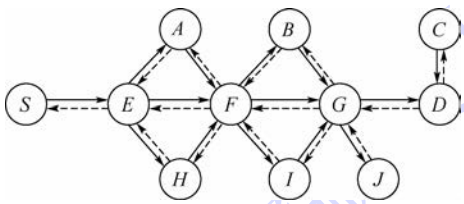


图 1　公共邻居

Fig. 1　Public neighbors

1.2　节点行为

Ad Hoc 网络节点通常无法在一跳范围内完成分组的传送任务,需要中间节点的转发,自私节点为了节省自身资源会出现不转发的行为。这些节点的行为会引起如黑洞攻击[14]、灰洞攻击[15]等威胁。

1) 黑洞攻击:在路由请求过程中,恶意节点收到 RREQ 后,在并没有到目的节点的路由的情况下,立即进行路由应答 RREP,谎称自己有到达目的节点的最近路由,使源节点和该恶意节点建立路由。此后源节点就沿该路径发送分组,恶意节点在此过程中丢弃经过它的分组,形成丢弃报文的黑洞。

2) 灰洞攻击:灰洞攻击是针对分组的 Dos 攻击。攻击节点在网络状况良好的情况下,首先以正常方式参与路由发现过程,然后根据分组的类型或者数量选择性地丢弃数据。有时该类节点只丢弃数据分组,使得路由信息正确,但是数据无法传送。灰洞攻击节点的存在会降低系统的性能,扰乱路由的建立过程,从而造成安全隐患。

2　信任模型

2.1　信任模型及其改进

在安全路由协议中有两种信任值:链路信任值和路径信任值。

1) 链路信任值:其定义类似于节点信任值[16-17],但是为了更加确切描述不同节点对公共邻居的信任值,本文使用链路信任值。当一个节点要求其邻居节点转发分组时,会立即使用监视模块对邻居节点的行为进行监视,同时,若邻居节点正确地转发了分组,则计算链路信任值。本文中链路信任值引入间接信任值评估方法,在计算直接信任值 $T^{\mathrm{D}}_{i\to j}$ 的基础上,通过和公共邻居之间进行信息交换,获得推荐信任值 $T^{\mathrm{I}}_{i\to j}$,最终进行综合评估得出链路信任值 T_{l}。

在 t 时刻节点 i 对节点 j 的直接链路信任值表示为 $T^{\mathrm{D}}_{i\to j}(t)$,定义如下:

$$T^{\mathrm{D}}_{i\to j}(t)=\begin{cases}\dfrac{N_{\mathrm{f}}(t)-N_{\mathrm{f}}(t-W)+1}{N_{\mathrm{r}}(t)-N_{\mathrm{r}}(t-W)+2} & t>W\\ \dfrac{N_{\mathrm{f}}(t)+1}{N_{\mathrm{r}}(t)+2} & t\leqslant W\end{cases}\tag{1}$$

式中:$N_{\mathrm{f}}(t)$ 为节点 j 在 t 时刻之前正确转发的分组;$N_{\mathrm{r}}(t)$ 为在 t 时刻之前节点 i 发送到节点 j 的全部分组;W 为时间窗。$T^{\mathrm{D}}_{i\to j}(t)$ 表示在时间窗内分组的转发率,初始化值为 0.5。

在 t 时刻节点 i 对节点 j 的间接链路信任值表示为 $T^{\mathrm{I}}_{i\to j}(t)$,如图 2 所示,节点 i 计算 a 提供给它的间接信任值为

$$T^{\mathrm{I}}_{(i\to j)a}(t)=\lambda T^{\mathrm{D}}_{i\to a}(t)(1-\lambda)T^{\mathrm{D}}_{a\to j}\quad \lambda>0.5\tag{2}$$

节点 i 将所有公共邻居节点的推荐信任值按式(2)进行处理,求得节点 i 对节点 j 的间接链路信任值。

$$T^{\mathrm{I}}_{i\to j}=\sum_{k=1}^{n}T^{\mathrm{I}}_{(i\to j)L_k}(t)/n$$

式中:L_k 为节点 i 与节点 j 的第 k 个公共节点。

综合评估节点 i 对节点 j 的链路信任值为

$$T_{i\to j}(t)=\mu T^{\mathrm{D}}_{i\to j}(t)+(1-\mu)T^{\mathrm{I}}_{i\to j}(t)$$

当 $T_{i\to j}(t)$ 接近 0 时,就认为该节点不可信。因此,若 $T_{i\to j}(t)$ 小于门限值 $T_{\mathrm{Threshold}}$,每个节点就将该邻居节点视为自私节点,并列入黑名单。当该自私节点存在于所有邻居节点的黑名单时,就会被移除网络。$T_{\mathrm{Threshold}}=(T_t^{\min}+T_t^{\max})/2.0$,$T_t^{\min}$ 和 $T_t^{\max}$ 分别为网络在 t 时刻的最小链路信任值和

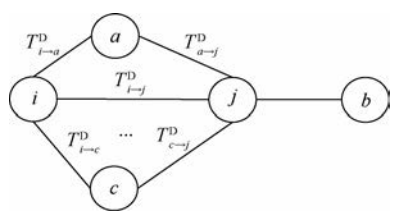


图 2　间接链路信任值评估方法

Fig. 2　Indirect link trust evaluation method

北航学报 赠阅

最大链路信任值,其初始化值为 0.5,随着网络的变化及时更新。

2)路径信任值:分组从源节点到目的节点需要一条路径,因此,考虑一条路径的可靠性是很有必要的。路径信任值 T_p 用来表示一个分组从源节点传送到目的节点的可能性,其定义如下:

$$T_p = \prod (T_{i\to i}(t) \mid n_i, n_j \in \text{path}, n_i \to n_j, n_j \neq N_d) \quad (3)$$

式中:$n_i \to n_j$ 表示节点 j 是节点 i 的下一跳节点;N_d 为目的节点。

在安全路由协议中,T_p 由转发控制分组来计算,并在每个节点的路由表中进行维护,其计算过程如图 3 所示。

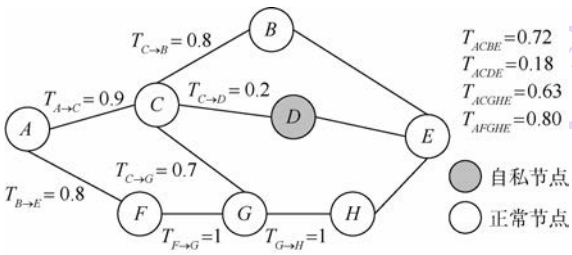


图 3 路径信任值计算示例

Fig. 3 Example of path trust value calculation

图 3 中,从节点 A 到节点 E 有 4 条路径,但是路径 L_{ACDE} 存在自私节点 D,若节点 A 只考虑链路信任值,则其有可能会选择让节点 C 转发分组,不能选择最优路径。本方案在路由表中维护路径信任值 T_p,节点 A 能够查询路由表,选择最安全的路径 L_{AFGHE}。

2.2 安全路由缺陷

现有安全路由协议中,所有节点在计算链路信任值时都使用固定时间窗,使得计算得到的信任值可以长时间使用。这种方法能很好地检测出如黑洞攻击节点这种行为保持不变的节点,但是对于其行为随时间变化的节点,该方法不能及时检测出节点的行为变化,因此就不能评估出准确的链路信任值,在数值上表现为其收敛速度变慢,而链路信任值的收敛速度较慢会引起两个问题:一是自私节点的误判;二是不能选择安全的路径。自私节点的检测会因为信任值收敛变慢而延缓,进一步由于不能及时计算出路径信任值 T_p,节点维护的路由表也不能及时更新。因此,使用可变时间窗来及时计算链路信任值 T_l 对提高安全路由协议性能至关重要。

3 TESBB 描述

TESBB 是一种基于节点行为的 Ad Hoc 网络安全路由信任评估方案,在该方案中,每个节点都可以通过计算邻居节点的链路信任值来发现自私节点。在实际 Ad Hoc 网络中,节点的行为会随着地理位置和环境的变化而变化,这就要求节点能够快速准确地评估邻居节点的行为,及时发现自私节点,并且正确选择安全路径。

TESBB 方案包括两个部分:检测节点行为变化和确定评估时间窗。

3.1 检测节点行为变化

每个节点在网络里都有能力侦听到其周围每个邻居节点的分组收发情况,并且具备监控机制使得一个节点能够检测和收集其邻居节点的行为。

定义链路信任值的变化为 ΔT,节点 i 对节点 j 在 t 时刻的链路信任值变化为 $\Delta T_{i\to j}(t) = T_{i\to j}(t) - T_{i\to j}(t')$,$t$ 和 t' 分别为评估链路信任值的当前时间和上一次时间。

$\Delta T_{i\to j}(t) > 0$:节点 j 转发了节点 i 要求转发的分组。

$\Delta T_{i\to j}(t) < 0$:节点 j 丢弃了节点 i 要求转发的分组。

因此,当 $\Delta T_{i\to j}(t)$ 为非 0 值时,说明节点 j 的行为发生了变化。为了加快节点信任值的收敛速度,定义 $\text{Con}_{i\to j}(t)$ 来反映节点行为的变化,$\text{Con}_{i\to j}(t) = \Delta T_{i\to j}(t) \times \Delta T_{i\to j}(t')$。

$\text{Con}_{i\to j}(t) > 0$:节点 j 的行为保持不变。

$\text{Con}_{i\to j}(t) < 0$:节点 j 的行为有较大变化。

因此,当 $\text{Con}_{i\to j}(t) < 0$ 时,节点 i 发现节点 j 的行为有明显改变,确定其为自私节点。

3.2 确定评估时间窗

在现有安全路由协议中,大多都采用固定时间窗来判断邻居节点自私与否,但是在实际网络中这种方法的假设条件很难实现,无法用节点以前的信任值来反映节点当前的行为状态。在 TESBB 中,自私节点的链路信任值被明显减小。引入适应性时间窗,根据邻居节点的行为灵活决定评估时间窗,从而精确地表示评估链路信任值所需的时间。当一个节点检测到其邻居节点行为发生改变时,该节点在链路信任值计算中使用可变时间窗来反映邻居节点的新行为。若节点从协作节点变为自私节点,链路信任值通过评估会降低。

在 t 时刻节点 i 对节点 j 的评估时间窗为 $W_{i\to j}(t)$,计算公式如下:

$$W_{i\to j}(t) = \begin{cases} W_{\max} & t > W_{\max} + t_{i\to j} \\ t - t_{i\to j} & W_{\min} + t_{i\to j} < t \leqslant W_{\max} + t_{i\to j} \\ W_{\min} & t \leqslant W_{\min} + t_{i\to j} \end{cases} \quad (4)$$

式中:$t_{i\to j}$为节点 i 检测到节点 j 行为发生变化的时刻;$W_{\max}$为固定时间窗;$W_{\min}$为设置的最短检测时间间隔。

某一节点在 $t_{i\to j}$时刻发现其邻居节点行为改变时,其时间窗 W 变化如图 4 所示。当邻居节点行为不变时,时间窗为 $W_{\max}$。

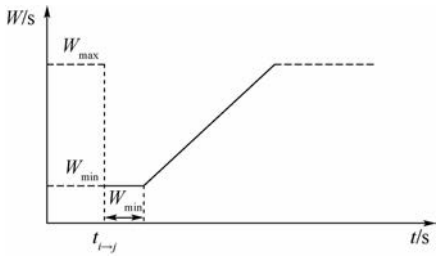


图 4　可变时间窗

Fig. 4　Variable time window

4　安全路由协议

以 AODV 为基础路由协议进行扩展,设计出了一种安全路由协议 TEAR,去除依据跳数值得最优路径选择方法,实现了多路径功能,考虑了路径信任值的传递和计算,最终选择出安全的传输路径。本方案中,Ad Hoc 网路的每个节点要维护一张路由表和一张信任表。信任表的格式如表 1 所示,表中信任值初始值为 0.5。

表 1　信任表结构

Table 1　Structure of a trust table

字段	符号
节点	IP
链路信任值	T_l
转发的分组	N_f
接收的分组	N_r
时间窗	W
检测到行为变化的时刻	$t_{i\to j}$

图 5 给出了信任表的更新过程。当节点 i 要求节点 j 转发分组时,节点 i 就开始更新对节点 j 的信任记录。首先节点 i 在发送缓存中存储要发送的分组,N_r 进行加一操作,若节点 j 能够正确地转发分组,N_f 进行加一操作,然后节点 i 使用更新消息计算 $T_{i\to j}(t)$、$\Delta T_{i\to j}(t)$、$\mathrm{Con}_{i\to j}(t)$,如果 $\mathrm{Con}_{i\to j}(t)<0$,节点 i 更新 $t_{i\to j}$为当前时间,最后根据 $t_{i\to j}$来更新 $W_{i\to j}(t)$。

TEAR 的过程描述如下:

在没有目的节点活动路由的情况下,源节点

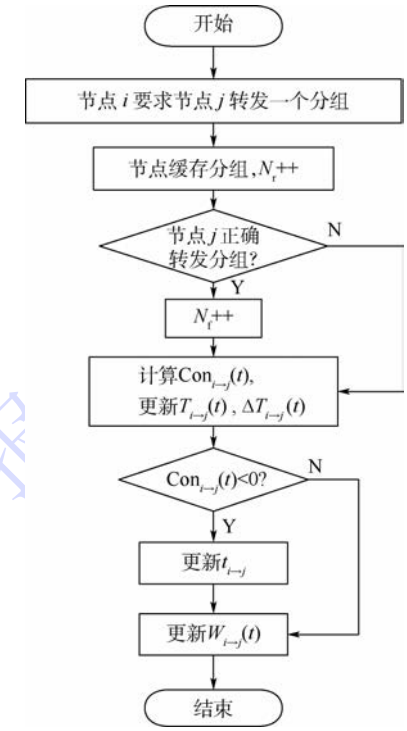


图 5　信任表更新流程图

Fig. 5　Flowchart of trust table updating

需要广播 RREQ 分组,其关键消息格式如图 6 所示。RREQ 从源节点转发到不同的目的节点时,沿途所经过的节点都要建立到源节点的反向路由 Reroute。当中间节点接收到 RREQ 分组时,查询路由表,根据路由表中的 T_p 和自己节点的信任表记录更新路径信任值,并将该节点加入到 Reroute 中,进行上述过程直到满足下述条件:① 当前节点是目的节点;② 当前节点的路由表中有到目的节点的表项,并且表项中 DestSeq 大于 RREQ 中的 DestSeq。

当 RREQ 分组到达目的节点或有到达目的节点的路由时,能够根据到达的分组内容来做出路由回复,将 RREP 分组沿着信任值最大的反向路径传输到源节点,建立路由,进行数据分组发送。

安全路由协议 TEAR 的伪代码描述如下所示,其中信任表的更新过程在图 5 中。

```
if(N_s! →N_d)
  send RREQ; //choice node i to analyse
while(t < TTL){
  if( RREQ. SourIP == Node_i SourIP &&
    RREQ. SourSeq == Node_i. SourSeq)
      drop RREQ;
  else{ //update node record
```

PktType	SourIP	SourSeq	BrdcastID	DestIP	DestSeq	HopCont	TTL	ReRoute

图 6　RREQ 和 RREP 消息格式

Fig. 6　Format of RREQ Packet and RREP Packet

```
        DestSeq = RREQ. SourSeq;
        precursor = TESBB( nodes) ;
        // nodes represent the neighbours of node i
        HopCount = RREQ. HopCont;
        Reset TTL;
    }
    if( Node_i == N_d ||
      Node_i → N_d ) // node i keep the route to dest node
      Node_end = Node_i ;
      reply REEP;
      goto L1 ;
    else{ // update RREQ
        RREQ. DestSeq = MAX( DestSeq) ;
        // the maximum DestSeq related to dest node
in record
        RREQ. HopCont + + ;
        RREQ. ReRoute ← Node i; // add node i to
reversse route
    }
} // end while
L1 ;
// initialize RREP
if( RREQ. DestSeq == Nodeend. DestSeq) {
    Nodeend. DestSeq + + ;
    RREP. DestSeq = Nodeend. DestSeq;
    RREP. ReRoute = RREQ. ReRoute;
    // copy RREQ's reversse route to RREP
if( Node_eng == N_d )
    RREP. HopCont = 0;
else
    RREP. HopCont = Node_end. HopCont;
}
if( Node_j. DestSeq == RREP. DestSeq ||
    RREP. HopCont < Node_j. HopCont)
    forward pointer = RREP. ReRoute;
    RREP. HopCont + + ;
if( Node_j == N_s )
    Set up routing connection;
```

5 模拟实验与分析

5.1 模拟环境设置

本文使用 NS2 来建立网络模拟环境，对 TEAR 协议进行仿真和分析，并与 AODV 和固定时间窗安全路由协议 TA-AODV[12] 进行比较，其中 TA_AODV 的时间窗设置为 50 s。在 AODV 源代码中添加自私节点类，加入转发分组计数函数与链路信任值和路径信任值的计算函数，修改路由代码后重新编译。仿真比较，分析 trace 文件，选择相应的字段，编写 awk 脚本生成最终分析所需文件。

模拟中在 1 000 m × 1 000 m 的二维环境中设置 100 个节点，节点之间的通信使用点到点的 CBR 数据流，发送速度为每秒 4 个报文，报文大小为 512 B，节点通信半径为 250 m，无线链路的宽度为 2 Mb/s，使用 Random Waypoint 移动模型，PauseTime 设置为 0，最大速度为 2 m/s，仿真时间为 2 000 s，用 setdest 程序产生相应代码。模拟中的节点分为正常节点和自私节点，自私节点转发额定数量的分组就开始把接收到的分组全部丢弃，自私节点的数量不超过 30。为了减小误差，取 10 次模拟结果的平均值进行比较。仿真参数的设定如表 2 所示。

表 2 模拟参数设定

Table 2 Setting of simulation parameter

参数类型	数值
模拟时间/s	2 000
模拟区域/m^2	1 000 × 1 000
节点数量	100
自私节点数量	0 ~ 30
MAC 类型	IEEE 802.11b
数据传输类型	CBR(UDP)
发包频率/(packets · s^{-1})	4
数据包长/B	512
初始化链路信任值	0.5
$T_{Threshold}$	0.5
最大时间窗 W_{max}/s	100

选取以下标准来进行评价：

1）分组交付率：目的节点接收到的分组与发送的总的分组的比率。

2）网络吞吐量：单位时间内所有目的节点的平均接收数据速率。

3）评估时间：检测出自私节点行为变化所用的平均时间。

5.2 模拟结果

5.2.1 分组交付率

分析比较自私节点数量对网络平均分组交付率的影响，仿真结果如图 7 所示。可以看出，随着自私节点数量的增加，所有协议的分组交付率都在下降。AODV 协议受自私节点的影响最大。网络中不存在自私节点时，由于 TA-AODV 在 50 s 的固定时间窗内不能对节点行为做出合理评估，有时会把正常节点当作自私节点，致使其分组交付率较低。TEAR 在可变时间窗内监测节点的行为，能及时发现自私节点，所以在自私节点数达到

北航学报 赠阅

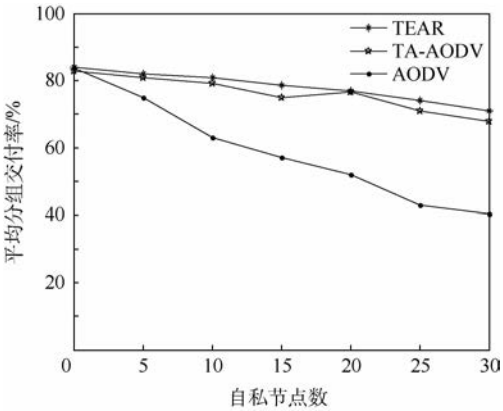


图 7　平均分组交付率比较

Fig. 7　Comparison of average packet delivery rate

30 时仍能保持 70% 以上的分组交付率。

5.2.2　网络吞吐量

分析比较自私节点数量对平均网络吞吐量的影响,仿真结果如图 8 所示。当不存在自私节点时,TEAR 和 TA-AODV 的平均网络吞吐量低于 AODV,这是因为 AODV 协议中没有添加安全相关的字段以及函数,在绝对安全条件下,其具有更高的吞吐量。随着自私节点数量的增多,AODV 吞吐量下降明显,转发的分组数量变少,最终到达目的节点的分组数减少,而 TEAR 能够更加有效地避开自私节点,一次性选择更加安全的路径,其路径重新路由的概率较小。TEAR 的平均网络吞吐量高于 AODV 和 TA-AODV 的最大值分别为 46.2% 和 10.5% 。

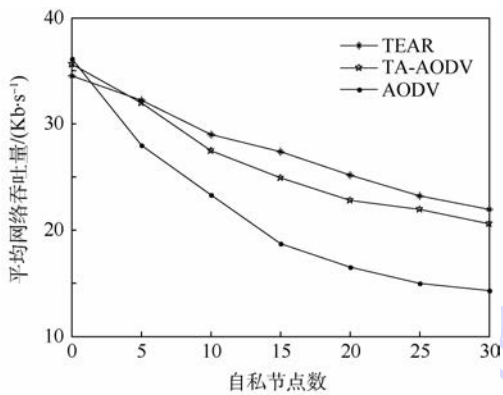


图 8　平均网络吞吐量比较

Fig. 8　Comparison of average network throughput

5.2.3　评估时间

1) 检测相同数量的自私节点所用的时间能够反映节点信任评估模型的性能。记录仿真结果如图 9 所示。可以看出,当自私节点数目较少时,TEAR 能较快地发现自私节点,而当自私节点数目达到 14 个时,两种方法所用时间基本相同,这是因为 TEAR 评估模型引入了公共邻居节点推荐

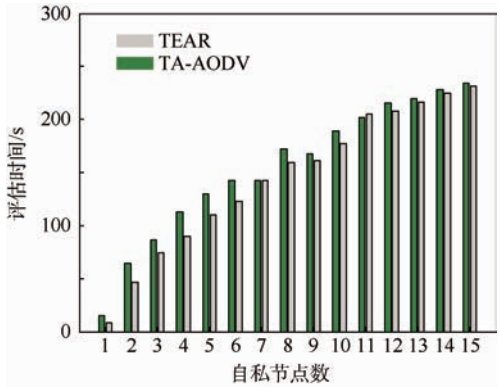


图 9　自私节点评估时间比较

Fig. 9　Comparison of selfish node evaluating time

的间接链路信任值,能更加精确地描述节点信任值的变化。

2) 自私节点数达到 15 时,分析得出结果如图 10 所示。评估时间表示发现一个自私节点所用时间,节点数目表示有发现该自私节点的平均正常节点数目。从图 10 中可以看出,TEAR 能够较快地发现节点行为的改变,要使 TA-AODV 中的 5 个节点发现一个自私节点行为的改变,需要 600 s,而在 TEAR 中只需要 400 s,这是因为固定时间窗的信任评估方案信任值收敛速度较慢,无法及时反映节点的行为变化,当其行为已经发生改变时,仍使用以前的评估值来判断节点的状态。

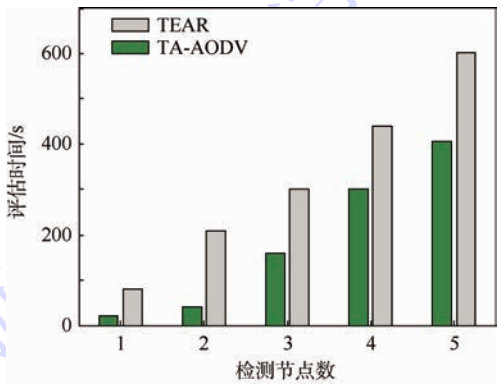


图 10　评估节点数目比较

Fig. 10　Comparison of evaluating node number

6　结　论

本文提出了一种基于节点行为的信任评估方案,并且以 AODV 协议为基础进行扩展得到安全路由协议 TEAR。TEAR 使用直接信任评估和间接信任评估结合的方案,保证了链路信任值的可靠性;使用可变时间窗机制,当节点行为改变比较频繁时,不必等待固定的时间后再去评估,能够及时发现自私节点,确保了评估的时效性。通过对分组交付率和网络吞吐量的比较发现,TEAR 能

北航学报 赠阅

够有效地避开自私节点,选择仅由合作节点构成的安全路径进行分组传递,减少重复路由概率,提高了网络性能。本文提出的协议在信任值门限的更新以及自私节点丢包行为方面没有详细说明,如何及时根据当前网络状态更新信任门限值并采取多种机制模拟节点丢包行为,以确保 Ad Hoc 路由协议更加安全高效,是未来的一个研究方向。

参考文献(References)

[1] YANG H,LUO H Y,YE F,et al. Security in mobile Ad Hoc networks:Challenges and solutions[J]. IEEE Wireless Communications,2004,11(1):38-47.

[2] GUPTE S,SINGHAL M. Secure routing in mobile wireless Ad Hoc networks[J]. Ad Hoc Networks,2003,1(1):151-174.

[3] KANNHAVONG B,NAKAYAMA H,NEMOTO Y,et al. A survey of routing attacks in mobile Ad Hoc networks[J]. IEEE Wireless Communications,2007,14(5):85-91.

[4] ZAPATA M G. Secure Ad Hoc on-demand distance vector routing[J]. ACM SIGMOBILE Mobile Computing and Communications Review,2002,6(3):106-107.

[5] HU Y C,PERRIG A,JOHNSON D B. Ariadne:A secure on-demand routing protocol for Ad Hoc networks[J]. Wireless Networks,2005,11(1):21-38.

[6] HU Y C,JOHNSON D B,PERRIG A. SEAD:Secure efficient distance vector routing for mobile wireless Ad Hoc networks[J]. Ad Hoc Networks,2003,1(1):175-192.

[7] 付才,洪帆,洪亮,等. 基于信任保留的移动 Ad Hoc 网络安全路由协议 TPSRP[J]. 计算机学报,2007,30(10):1854-1864.
FU C,HONG F,HONG L,et al. Mobile Ad Hoc secure routing protocol based on trust preserving[J]. Chinese Journal of Computers,2007,30(10):1854-1864(in Chinese).

[8] BALAKRISHNAN V,VARADHARAJAN V,LUCS P,et al. Trust enhanced secure mobile Ad-Hoc network routing[C]// Proceedings of the 21st International Conference on Advanced Networking and Applications. Piscataway,NJ:IEEE Press,2007:27-33.

[9] ABUSALAH L,KHOKHAR A,GUIZANI M. A survey of secure mobile Ad Hoc routing protocols[J]. IEEE Communications Surveys and Tutorials,2008,10(4):78-93.

[10] LI X,JIA Z,ZHANG P,et al. Trust-based on-demand multipath routing in mobile Ad Hoc networks[J]. IET Information Security,2010,4(4):212-232.

[11] PERRONE L F,NELSON S C. A study of on-off attack models for wireless Ad Hoc networks[C]// Proceedings of 1st Workshop on Operator-Assisted(Wireless Mesh)Community Networks. Piscataway,NJ:IEEE Press,2006:58-67.

[12] USHIKUBO H,TAKEDA S,SHIGENO H. An effective secure routing protocol considering trust in mobile Ad Hoc networks[J]. IPSJ Journal,2014,55(2):649-658.

[13] UMEDA S,TAKEDA S,SHIGENO H. Trust evaluation method adapted to node behavior for secure routing in mobile Ad Hoc networks[C]// Proceedings of 2015 8th International Conference on Mobile Computing and Ubiquitous Networking(ICMU). Berlin:Springer,2015:143-148.

[14] KUROSAWA S,NAKAYAMA H,KATO N,et al. Detecting blackhole attack on AODV-based mobile ad hoc networks by dynamic learning method[J]. International Journal of Network Security,2007,5(3):338-345.

[15] SEN J,CHANDRA M,HARIHARA S,et al. A mechanism for detection of grayhole attack in mobile Ad Hoc networks[C]// Proceedings of 2007 6th International Conference on Information,Communications and Signal Processing. Piscataway,NJ:IEEE Press,2007:1-5.

[16] GHOSH T,PISSINOU N,MAKKI K. Collaborative trust-based secure routing against colluding malicious nodes in multi-hop Ad Hoc networks[C]// Proceedings of the 29th Annual IEEE International Conference on Local Computer Networks(LCN'04). Piscataway,NJ:IEEE Press,2004:224-231.

[17] CHENG Y,HUANG C,SHI W. Trusted dynamic source routing protocol[C]// Proceedings of the 3rd International Conference on Wireless Communications,Networking and Mobile Computing. Piscataway,NJ:IEEE Press,2007:1632-1636.

作者简介:

陈少华 男,硕士研究生。主要研究方向:移动自组织网络、机载计算机网络。
Tel.:17792021962
E-mail:876915107@qq.com

樊晓光 男,博士,教授,博士生导师。主要研究方向:无线通信、综合航电技术。
Tel.:029-84787655
E-mail:1203574700@qq.com

禚真福 男,博士,讲师。主要研究方向:嵌入式系统开发。
E-mail:zzf_tiger@126.com

黄金科 男,博士研究生。主要研究方向:无线通信、综合航电技术。
Tel.:15829717886
E-mail:86297609@qq.com

孙贤明 男,硕士研究生。主要研究方向:机载计算机网络及检测技术。
E-mail:1553552609@qq.com

北航学报 赠阅

Secure routing protocol based on optimized trust evaluation in Ad Hoc network

CHEN Shaohua, FAN Xiaoguang*, ZHUO Zhenfu, HUANG Jinke, SUN Xianming

(College of Aeronautics and Astronautics Engineering, Air Force Engineering University, Xi'an 710038, China)

Abstract: Aimed at the attack of selfish nodes in mobile Ad Hoc networks, this paper proposes a secure routing protocol TEAR, which is based on trust evaluation. In order to realize the accuracy and timeliness of trust values, the proposed protocol, which is based on the trust evaluation scheme of node behavior, uses the link trust and path trust to represent the reliability of a node and a path respectively, and introduces the indirect trust evaluation scheme with public neighbors' trust values and variable time window. The trust value table of each node and the new structure of RREQ and RREP provide the quick selection of secure routing. Finally, TEAR protocol was verified by NS2 simulation and compared with AODV and TA-AODV which is based on fixed time window trust evaluation. The simulation results show that the improved trust evaluation scheme has a superior performance, and TEAR can not only detect the selfish nodes correctly but also select secure path, which improves the function of Ad Hoc networks.

Key words: Ad Hoc network; selfish nodes; secure routing; trust evaluation; variable time window; node behavior

北京航空航天大学学报

Received: 2016-03-09; **Accepted**: 2016-04-22; **Published online**: 2015-05-13 11:36

URL: www.cnki.net/kcms/detail/11.2625.V.20160513.1136.001.html

Foundation items: National Natural Science Foundation of China (61402517); AVIC Pre-Research Foundation of China (619010601)

* **Corresponding author**. Tel.: 029-84787655　E-mail: 1203574700@qq.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2016.0410

基于 CEEMDAN 的雷达信号脉内细微特征提取法

王文哲，吴华*，王经商，张强

（空军工程大学 航空航天工程学院，西安 710038）

摘 要：有效的信号特征提取是高精度雷达辐射源识别的基础，以脉冲描述字为代表的传统特征已无法满足复杂电磁环境的需要。本文提出一种基于自适应噪声完备集合经验模态分解(CEEMDAN)的有效雷达辐射源脉内细微特征提取算法。雷达信号由对非平稳、非线性信号尤为有效的 CEEMDAN 分解产生的个别分量重构，抑噪效果通过 1 000 次蒙特卡罗实验得到验证，同时设计基于该重构的一种脉内特征空间。本文方法与主流特征提取方法的识别精度在 6 部雷达辐射源产生的 3000 个不同脉内调制的加噪信号样本上进行了实验对比，结果表明不同种类信号样本在本文特征空间中清晰可分，本文方法较之主流方法更加精确，尤其在 0 dB 信噪比(SNR)下仍保持 90% 以上的高精度。

关 键 词：特征提取；自适应噪声完备集合经验模态分解(CEEMDAN)；经验模态分解；噪声协助；雷达辐射源识别；信号重构

中图分类号：TN97

文献标识码：A **文章编号**：1001-5965(2016)11-2532-08

辐射源识别是现代雷达侦察的核心功能。识别包括两个过程，即信号特征提取和信号分类。表面上看，后者是辐射源识别的最终目的，但是由于雷达信号取样后的高维性，没有合适的分类器能够直接对其进行处理，因此信号分类必须建立在特征提取的基础上。以脉冲描述字(PDW)[1]为代表的传统特征集在电子战实践中长期占据中心地位，然而脉冲压缩雷达、相控阵雷达等复杂体制雷达已对其性能构成严峻挑战。尤其亟需解决的是，在处理复杂体制辐射源或者辐射源数目较多时，由于 PDW 参数基本上基于脉间特性，而脉间参数完全无法充分反映辐射源信号调制特性，直接导致 PDW 特征空间的脉冲重叠概率过高[2]。为解决这一问题，一种思路是设计更佳的分类器来补偿分类特征的性能下降。Matuszewski 和 Paradowski[3]提出一种专家系统来实现辐射源识别。刘海军等[4]利用基于联合参数建模的方法对雷达辐射源进行识别。关一夫和张国毅[5]则提出一种基于隐马尔可夫模型的雷达辐射源识别方法。Xiao 等[6]设计一种基于仿射传播聚类(AP 聚类)的支持向量机(SVM)雷达辐射源分类识别方法。尽管这些方法在一定程度上提升了识别效能，但是没有对特征提取过程进行改进，因此仍然没有从本质上解决问题。另一种思路即通过利用信号脉内特征改进特征提取[7-8]。Pu 等[9]用非模糊相位扩展和滑动平均改进瞬时自相关，并提取雷达信号的主瞬时频率。这种方法仅适用于平稳线性信号。Zhu 和 Jin[10]提出一种基于连续小波变换(CWT)的雷达信号时频特征提取方法。然而小波变换仍然具有无法克服的缺陷[11]，例如无法处理非线性信号，以及无法在时频域同时取得高分辨率等。一般来说，侦察接收机接收到的基本均为非平稳非线性信号，因此找到能够处理此类信号的分析工具尤为关键。自适应噪声完备

收稿日期：2016-05-17；**录用日期**：2016-07-16；**网络出版时间**：2016-08-30 15:52
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160830.1552.003.html
基金项目：国家自然科学基金（61379104）；航空科学基金（20152096019）
* **通讯作者**：Tel.：029-84787681 E-mail：15129291728@163.com

引用格式：王文哲，吴华，王经商，等. 基于 CEEMDAN 的雷达信号脉内细微特征提取法[J]. 北京航空航天大学学报，2016，42(11)：2532-2539. WANG W Z，WU H，WANG J S，et al. Subtle intrapulse feature extraction based on CEEMDAN for radar signals[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2532-2539 (in Chinese).

北航学报 赠阅

集合经验模态分解(CEEMDAN)由 Torres 等[12]在 2011 年提出,正是尤为适合非线性非平稳过程分析的一类时频分析工具。利用希尔伯特变换,经验模态分解(EMD)[13]可以得到信号完整的能量-频率-时间分布,称为希尔伯特-黄谱。这一类信号分析方法被 NASA 定义为希尔伯特-黄变换(HHT),尽管问世不久,EMD 被实践证明是真实信号的高效分析方法,广泛应用于诸多工程领域,然而却仍具有所谓的"模式混淆"问题[14]。该问题可以由噪声协助方法解决,CEEMDAN 正是此类方法中具有重要影响的一种。本文提出一种基于 CEEMDAN 的雷达辐射源脉内细微特征提取方法,该方法非常适于处理非线性非平稳雷达信号,与传统方法相比识别精度更高,并具有良好的抑噪性能。

1　CEEMDAN 核心理论

本节介绍 EMD 和 CEEMDAN 的基本概念。但出于逻辑上的清晰,首先简要介绍希尔伯特变换[15],以说明 CEEMDAN 算法的必要性。

1.1　希尔伯特变换

对于任意实信号 $x(t)$,其希尔伯特变换为

$$y(t)=\frac{1}{\pi}P\int_{-\infty}^{\infty}\frac{x(\tau)}{t-\tau}\mathrm{d}\tau \tag{1}$$

式中:P 为柯西主分量。结合 $y(t)$,可获得实信号的解析形式:

$$z(t)=x(t)+\mathrm{i}y(t)=a(t)\mathrm{e}^{\mathrm{i}\theta(t)} \tag{2}$$

式中:

$$\begin{cases}a(t)=(x^2(t)+y^2(t))^{1/2}\\ \theta(t)=\arctan\dfrac{y(t)}{x(t)}\end{cases} \tag{3}$$

$a(t)$为瞬时幅度,而 $\theta(t)$为瞬时相位。则瞬时频率为

$$\omega(t)=\frac{\mathrm{d}\theta(t)}{\mathrm{d}t} \tag{4}$$

希尔伯特变换是分析单成分信号的高效工具。然而由于真实信号往往是多成分的,希尔伯特变换获得的瞬时频率对其来说往往没有物理意义[16],这一点极大地限制了希尔伯特谱分析的应用范围。为解决这一问题,可采用类 EMD 算法,如 CEEMDAN。

1.2　EMD 和 CEEMDAN

EMD 的主要思想是将一信号 $x(t)$分解为一系列本征模态函数(IMF),再使用希尔伯特变换分析 IMF。IMF 必须满足两个条件:①在全局特性上看,极值点数和过零点数相差不能超过 1;②在局部特性上看,各处由局部极大值点定义的包络线和局部极小值定义的包络线的均值都为 0。

EMD 的局部特性可能导致单个 IMF 产生不同幅度的震荡,或者多个 IMF 产生相似幅度的震荡。当不需要这种现象出现或者不同 IMF 需要相似幅度的时候,这就成为所谓的"模式混淆"问题。CEEMDAN 使用噪声协助类方法来解决该问题。算法具体表述如下:令 $E_k(\cdot)$为产生第 k 个 IMF 的算子,且令 $w(i)$为零均值单位方差白噪声的一个实现,接着:

步骤 1　对 $i=1,2,\cdots,I$ 使用 EMD 分解 $x^{(i)}=x+\beta_0 w^{(i)}$,并定义第一 CEEMDAN IMF 为 $\tilde{d}_1=\dfrac{1}{I}\sum_{i=1}^{k}d_1^{(i)}=\bar{d}_1$。

步骤 2　计算第一残差:$r_1=x-\tilde{d}_1$。

步骤 3　使用 EMD 获得 $r_1+\beta_1E_1(w^{(i)})$,$i=1,2,\cdots,I$ 的第一 IMF,定义第二 CEEMDAN IMF 为:$\tilde{d}_2=\dfrac{1}{I}\sum_{i=1}^{I}E_1(r_1+\beta_1E_1(w^{(i)}))$。

步骤 4　对 $k=2,3,\cdots,K$ 计算第 k 残差:$r_k=r_{(k-1)}-\tilde{d}_k$。

步骤 5　使用 EMD 获得 $r_k+\beta_kE_k(w^{(i)})$,$i=1,2,\cdots,I$ 的第一 IMF,定义第$(k+1)$CEEMDAN IMF 为 $\tilde{d}_{(k+1)}=\dfrac{1}{I}\sum_{i=1}^{I}E_1(r_k+\beta_kE_k(w^{(i)}))$。

步骤 6　对下一个 k 回到步骤 4。

重复步骤 4 ~ 步骤 6 直至获取的残差不能被 EMD 继续分解。系数 $\beta_k=\varepsilon_k\mathrm{std}(r_k)$允许在每一次迭代中选择信噪比(SNR),其中:std(·)为标准偏差算子。

2　基于 CEEMDAN 的雷达信号重构和脉内细微特征向量提取

下面提出本文的特征提取方法。本文首先验证 CEEMDAN 对雷达信号数据的挖掘能力,并利用雷达信号的 IMF 成分重构信号,同时测试该信号重构算法的抑噪性能,以考查其在真实噪声环境下的重构效果。最后,基于该重构算法提取了 3 种脉内特征构成新的特征空间。该特征空间的识别效能将在下一节进行实验验证与分析。

假设 $x(t)$为接收到的某部雷达辐射的待识别雷达信号,且 $\tilde{d}_k$,$k=1,2,\cdots,K$ 为 CEEMDAN 算法计算出的 $x(t)$的 IMF。根据 1.2 节,$x(t)$可以被表示为

$$x(t)=\sum_{k=1}^{K}\tilde{d}_k+r_K \tag{5}$$

将希尔伯特变换施于每个 IMF,而该信号可表示为

$$x(t)=\sum_{k=1}^{K}a_k(t)\exp\left(\mathrm{i}\int\omega_k(t)\mathrm{d}t\right) \tag{6}$$

北航学报 赠阅

等式(6)使得幅度(或者能量)可以被表达为时间和频率的函数。

雷达信号为了抗干扰和功能需要,通常包含复杂的调制,并且为典型的多成分非线性非平稳随机过程。为此,这里采用两种现代雷达系统常用的发射调制样式进行分析,即线性调频(LFM)和二相频移键控(BPSK)。

2.1 基于 CEEMDAN 的信号重构

首先对以上提到的两种调制波形施以CEEMDAN,析取的IMF在图1中给出。结果表

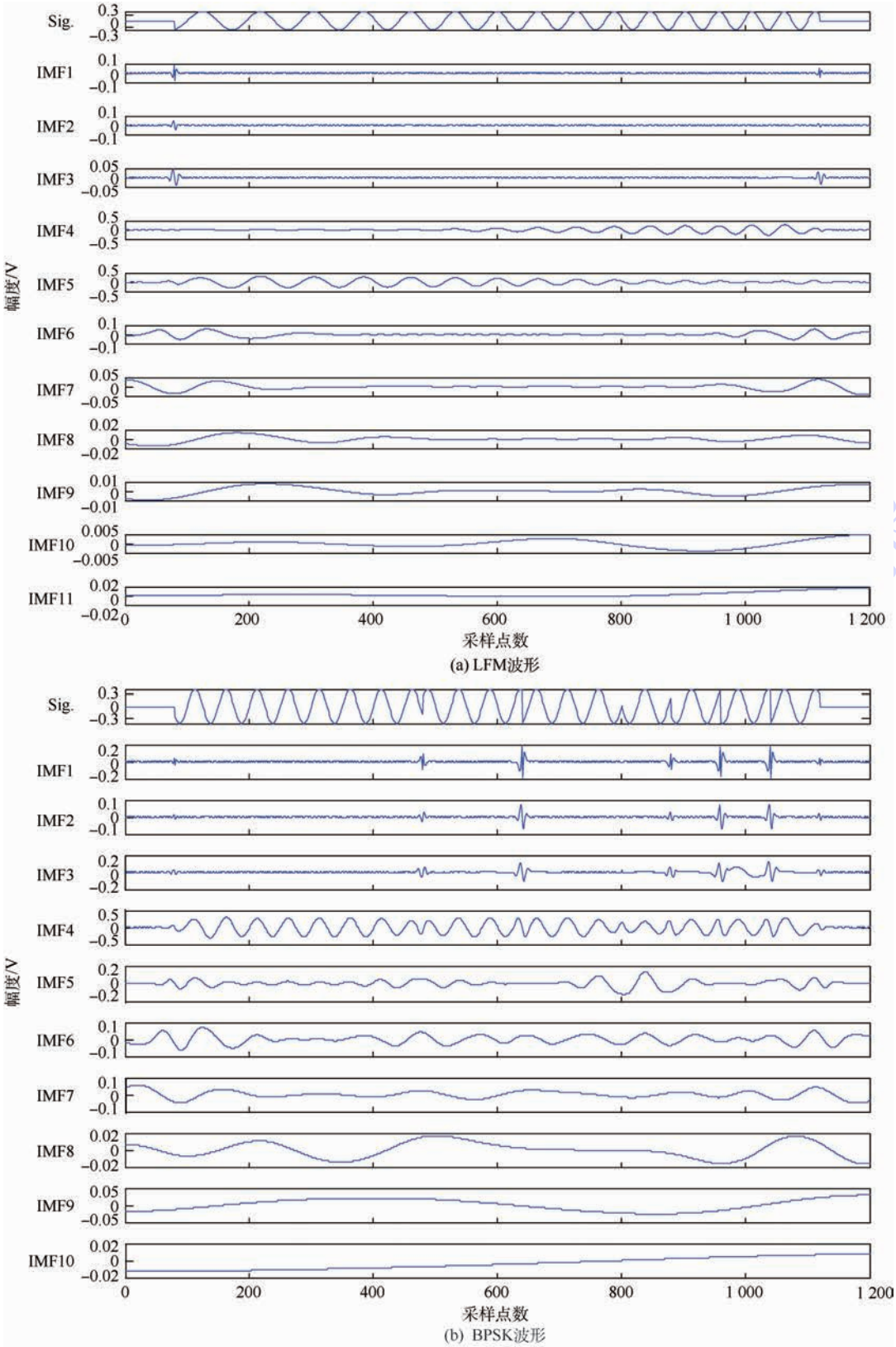


图 1 两种波形的 CEEMDAN 分解

Fig. 1 CEEMDAN decomposition of two waveforms

北航学报 赠阅

明,可以在一个雷达信号中分解出多达 10 个成分,这充分说明雷达信号中存在过多的信号成分,亦即是说在每一瞬间存在过多的瞬时频率值;同时可以看出,信号的主要成分集中在 IMF4 ~ IMF7 之间,因此本文使用这些成分重构被噪声污染的信号,亦即取其和信号为重构信号。

事实上,ECM 接收机收到的真实雷达信号不可能是上文讨论的纯净信号,而往往是噪声污染过的。因为 CEEMDAN 是基于噪声协助的类 EMD 方法,算法中使用的独立同分布噪声集在集合平均计算中彼此抵消,在本节实验中表现出了对信号加性噪声较好的压制作用。本文采用了信噪比分别为 10 dB 和 20 dB 的信号,对每种信号施加 CEEMDAN(协助噪声标准差为 0.05 V)并提取 IMF4 ~ IMF7 对信号进行重构,进行 1 000 次蒙特卡罗实验。结果表明对于 10 dB信噪比 LFM 和 BPSK 雷达信号,污染信号样本和纯净信号之间均方误差平均值分别为 0.316 3 V和 0.315 3 V,而重构信号和纯净信号之间均方误差平均值为0.100 7 V和 0.108 7 V;而对于 20 dB 信噪比信号,该均方误差平均值在重构之前为 0.098 9 V 和0.097 1 V,重构之后则为 0.042 5 V 和 0.051 8 V。从图 2 中可以看到,本文信号重构方法的去噪性能非常好,说明该重构方法有效地捕捉了带噪信号空间中的信号主要成分。

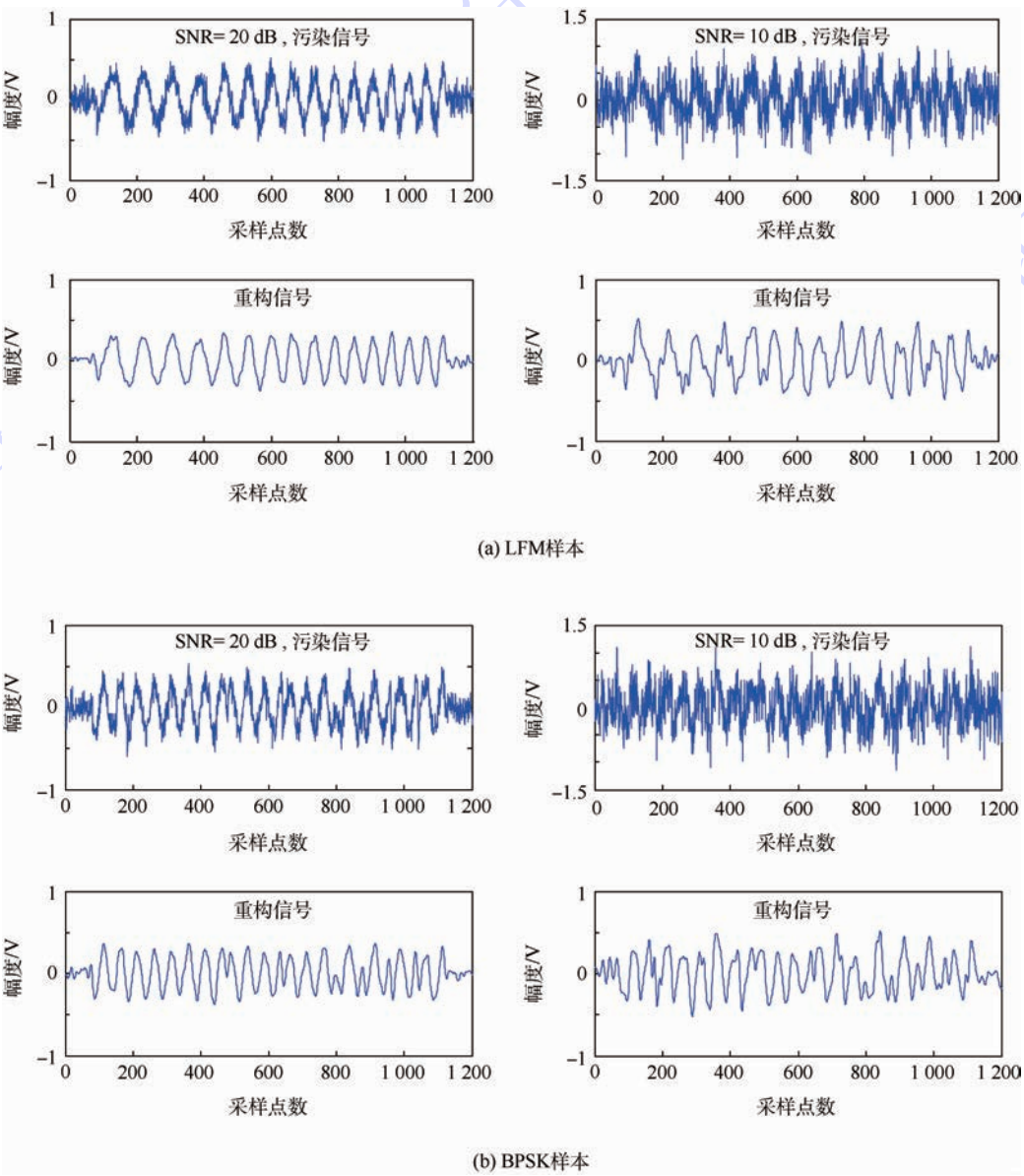


图 2　对两种波形样本的抑噪效果

Fig. 2　De-noising effect on two kinds of waveform samples

北航学报 赠阅

2.2 特征向量和特征空间构建

据 2.1 节,雷达信号的主要信息可以通过其重构信号表达,由于 IMF 是单一成分信号,可以对其施加希尔伯特变换提取其瞬时特征,如式(1)~式(4),因此与原始雷达信号不同,重构信号在保留了其主要信息的同时,是可以进行希尔伯特变换的,因此就可提取出基于其时频分布的信号特征。基于重构信号,本文提出 3 种信号特征,并构成特征向量和特征空间。

1) 首要成分相似比

$$F_1=\frac{\sum_{n=0}^{N}(c_{1n}-\bar{c}_1)(x_n-\bar{x})}{\sqrt{\sum_{n=0}^{N}(c_{1n}-\bar{c}_1)^2}\sqrt{\sum_{n=0}^{N}(x_n-\bar{x})^2}} \tag{7}$$

式中:N 为采样点数;x 为雷达信号;c_1 为信号首要 IMF 成分,本文中为 IMF4;x_n 和 c_{1n} 分别为 x 和 c_1 的采样后形式;F_1 描述 c_1 与 x 的相似程度。

2) 首要成分带宽

$$F_2=\frac{|\max(f_{1n})-\min(f_{1n})|}{f_{1c}} \tag{8}$$

$$f_{1c}=\sum_{n=0}^{N}f_{1n}/N \tag{9}$$

式中:f_{1n} 为首要 IMF 成分的瞬时频率;f_{1c} 为其平均值,即首要 IMF 成分的中心频率;F_2 描述 c_1 的频域跨度。

3) 成分瞬时频率方差

$$F_3=\frac{1}{K-1}\sum_{k=1}^{K}(f_{ck}-\bar{f}_c)^2 \tag{10}$$

$$\bar{f}_c=\frac{1}{K}\sum_{k=1}^{K}f_{ck} \tag{11}$$

式中:K 为用于重构信号的 IMF 成分个数;f_{ck} 为第 k 成分中心频率;F_3 描述用于重构信号的各 IMF 之间的频域差异。

事实上,基于重构信号可以提取的特征数目是无穷的,然而从计算成本角度考虑,特别是航空应用条件下的空间限制和计算实时性要求,本文只提出3种特征以刻画雷达脉内细微特征。其中,由于用于重构信号的各 IMF 之中,首要成分 c_1 又居其主干,因此 F_1 及 F_2 的计算只涉及信号首要成分,以进一步降低算法时间和空间复杂度;而 F_3 则从全局角度考察重构信号,以提供更丰富的信号细节。总的来讲,本文提出特征的考量在于在保留特征向量充足信息量和降低算法时空复杂度之间达到均衡。

则本文提出的特征向量可表示为 $\boldsymbol{F}=[F_1,F_2,F_3]$,选用合适的分类器,雷达信号就可以被有效识别。本文特征提取方法的计算流程详见图 3伪代码,并将在第 3 节进行验证。

```
Procedure  GenerateFeature
  c = SignalReconstruction(x(t));          ● 信号重构
  F1 = ComputeFeature1(x(t), c);           ● 计算特征F1
  for k = 1 … K do                         ● 计算重构瞬时频率
                                             和中心频率
    fk = Hilbert( ck );
    fck = Mean( fk );
  F2 = ComputeFeature2( f );               ● 计算特征F2
  f̄c = Mean( fc );                         ● 计算中心频率均值
  F3 = ComputeFeature3( fc, f̄c );          ● 计算特征F3
  Return F;
- - - - - - - - - -
Procedure  SignalReconstruction
  IMF = Ceemdan(x(t));                     ● CEEMDAN分解
  for i = 1 … 4 do                         ● 重构成分赋值
    ci = IMF i+3;
  Return@GenerateFeature;                  ● 返回主流程
```

图 3 本文雷达辐射源识别特征提取方法计算流程伪代码

Fig. 3 Computational pseudocode of proposed feature extraction method for radar emitter identification

3 实验与结果分析

为验证本文算法,本节进行基于 MATLAB (R2014a)的仿真。本节模拟了位于同一方位的 6 部雷达,即脉冲到达方位(DOA)认为全部相同。发射信号的调制形式与彼此略有不同,见表 1。

表 1 仿真辐射源信号的调制参数

Table 1 Signal modulation parameters of simulated emitters

辐射源	调制参数				
	载频 f_0/GHz	脉冲宽度 T/μs	脉冲幅度 A/V	调频斜率 α/(GHz·μs^{-1})	相位编码函数 φ_i
E1	2.1	10.4	0.32	0.10	
E2	1.9	10.3	0.33	0.15	
E3	2.0	10.2	0.31	0.20	
E4	1.8	10.3	0.30		[1111100110101]
E5	1.9	10.5	0.32		[11100010010]
E6	2.2	10.4	0.31		[1110010]

北航学报 赠阅

由表 1 可见,前 3 种信号为 LFM 信号,后3 种为 BPSK 信号。6 种信号在载频 f_0、脉冲宽度 T 和脉冲幅度 A 上基本一致,但 3 种 LFM 信号在调频斜率 α 相互区分,而 BPSK 信号在相位编码函数 φ_i 上相互区分。这样设计实验调制参数的原因在于充分模拟待识别信号 PDW 高度重叠的情况。为验证本文算法的实用性,将不同能量的白噪声叠加在表 1 所述信号上,并附加大小不同的相移,产生信噪比 0 ~ 20 dB、相位在 $-\pi/2 \sim \pi/2$ 的样本信号。表 1 中每种信号产生 500 个样本,也即共有 3 000 个样本。所有样本在图 4(a)所示的 PDW 特征空间中以不同颜色的点表示,每种信号对应一种颜色。从中可以看出,任何一种信号都无法在 PDW 特征空间中与其他辐射源信号区分开。实际上 PDW 中尚有其他参数,但出于图形表示的明晰性,不失一般性地予以略去。接下来将全部样本在归一化的本文特征空间中绘出,如图 4(b)所示。可以看出,除了部分信号类别的边缘略有接触,各个辐射源信号样本点都是可以清晰区分的。这证明本文特征空间对于雷达脉内细微特征能够良好地捕捉,对复杂体制雷达信号识别是可用的。

尔后随机选取 1 000 个样本训练 SVM 分类器,余下的 2 000 个样本用来测试该分类器。作为与本文算法的对比,同时用 Pu 的方法和 CWT 方法提取雷达信号特征,如文献[9-10],采用的分类器同样为 SVM。3 种算法的识别精度见表 2。

从表 2 中可见,在信噪比高达 20 dB 时,3 种方法的识别精度都相当高,但 2 种对照算法的精度随着信噪比下降均有大幅下滑,而本文算法在各种信噪比条件下仍保持很高的识别精度。0 dB 信噪比时,CWT方法的识别精度下降到 80.1%,Pu 的方法精度为 68.9%,这对于电子战来说已经相当低,但本文方法仍有 90% 以上的识别精度,这是由于本文方法的高效抑噪性能。此外,应注意到 CWT 和 Pu 的方法的识别精度上限分别在 98.5% 和 97.6% 左右,说明该两种算法仍未完全利用雷达信号的细微脉内特性,在特征空间中的各类样本对于 SVM 是线性不可分的,也即 SVM 算法无法给出没有错误的分离超平面;而本文方法的精度在信噪比为 20 dB 时已经达到 100%,说明本文方法给出的特征在提取脉内细微特征方面较之两种对照方法性能更优。

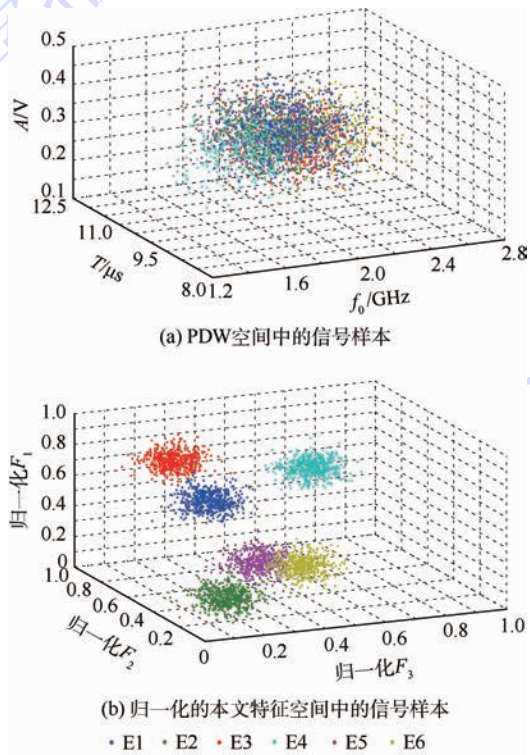


图 4　两种特征空间中的信号样本对比

Fig. 4　Comparison of signal samples in two feature space

表 2　在变化信噪比下的 3 种方法仿真结果

Table 2　Simulation results of three methods under changing SNR

方法	精度/%					
	SNR = 0 dB	SNR = 4 dB	SNR = 8 dB	SNR = 12 dB	SNR = 16 dB	SNR = 20 dB
Pu 的方法	68.9	77.2	85.1	91.2	97.3	97.6
CWT 方法	80.1	85.4	90.5	94.1	98.2	98.5
本文方法	90.7	92.8	94.1	96.3	99.6	100

4　结　论

脉内特征是雷达信号特征提取的重要研究方向,然而目前大部分脉内特征提取方法仍不能在非平稳非线性信号上取得足够好的效果,而此类信号恰恰构成了实际环境中雷达信号的主体。本文提出一种基于 CEEMDAN 的雷达辐射源识别脉内细微特征提取算法,仿真实验表明:

1) CEEMDAN 可以准确提取出构成常见雷达调制信号的 IMF 成分,且部分成分可用于重构雷达信号。

2) 由于 CEEMDAN 在非平稳非线性信号处

北航学报 赠阅

理上的优势,本文信号重构方法具有足够好的抑噪性能以在低信噪比条件下工作。

3) 本文特征空间可以充分区分 PDW 高度重叠的雷达信号样本,同时本文方法较之该领域主流方法具有更好的雷达辐射源识别精度。

4) 类 EMD 方法在雷达辐射源识别信号特征提取领域有较大的应用空间,为进一步探索类 EMD 方法在该领域的应用打下了基础。

由于类 EMD 方法属于经验方法,目前缺少相应的数学支撑,使得本文无法在数学层面上进行论证,这也是类 EMD 雷达信号处理仍需解决的课题。

参考文献(References)

[1] MONTAZER G, KHOSHNIAT H, FATHI V. Improvement of RBF neural networks using fuzzy-OSD algorithm in an online radar pulse classification system[J]. Applied Soft Computing, 2013,13(9):3831-3838.

[2] JIANG H, PANG Z, TANG P, et al. Intrapulse modulation recognition based on pulse description words[C]// Proceedings of the 2013 6th International Congress on Image and Signal Processing. Piscataway, NJ: IEEE Press, 2013, 3: 1367-1371.

[3] MATUSZEWSKI J, PARADOWSKI L. The knowledge based approach for emitter identification[C]// Proceedings of the 12th International Conference on Microware & Radar, MIKON '98. Piscataway, NJ: IEEE Press, 1998, 3: 810-814.

[4] 刘海军,柳征,姜文利,等. 基于联合参数建模的雷达辐射源识别方法[J]. 宇航学报,2011,32(1):142-149.
LIU H J, LIU Z, JIANG W L, et al. A joint-parameter modeling based radar emitter identification method[J]. Journal of Astronautics, 2011, 32(1): 142-149(in Chinese).

[5] 关一夫,张国毅. 一种基于隐马尔科夫模型的雷达辐射源识别方法[J]. 火力与指挥控制,2015,40(10):98-103.
GUAN Y F, ZHANG G Y. A radar emitter recognition algorithm based on hidden markov models[J]. Fire Control & Command Control, 2015, 40(10): 98-103(in Chinese).

[6] XIAO W, WU H, YANG C. Support vector machine radar emitter identification algorithm based on AP clustering[C]// Proceedings of the 2013 International Conference on Quality, Reliability, Risk maintenance, and Safety Engineering (QR2MSE). Piscataway, NJ: IEEE Press, 2013: 2062-2064.

[7] SINGH A, RAO K. Digital receiver-based electronic intelligence system configuration for the detection and identification of intra-pulse modulated radar signals[J]. Defence Science Journal, 2014, 64(2): 152-158.

[8] KAWALEC A, OWCZAREK R. Radar emitter recognition using intrapulse data[C]// Proceedings of the 15th International Conference on Microwaves, Radar and Wireless Communications, 2004, MIKON-2004. Piscataway, NJ: IEEE Press, 2004, 2: 435-438.

[9] PU Y W, JIN W D, ZHU M, et al. Classification of radar emitter signal based on cascade feature extraction and hierarchical decision technique[C]// Proceedings of the 8th International Conference on Signal Processing. Piscataway, NJ: IEEE Press, 2006.

[10] ZHU B, JIN W D. Feature analysis of advanced radar emitter signals based on continuous wavelet transform[J]. Applied Mechanics and Materials, 2013, 246-247: 1125-1129.

[11] PENG Z, TSE P, CHU F. A comparison study of improved Hilbert-Huang transform and wavelet transform: Application to fault diagnosis for rolling bearing[J]. Mechanical Systems and Signal Processing, 2005, 19(5): 974-988.

[12] TORRES M, COLOMINAS M, SCHOLOTTHAUER G, et al. A complete ensemble empirical mode decomposition with adaptive noise[C]// Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway, NJ: IEEE Press, 2011: 4144-4147.

[13] HUANG N E, SHEN Z, LONG S, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[C]// Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences. London: The Royal Society, 1998, 454(1971): 903-995.

[14] WU Z H, HUANG N E. Ensemble empirical mode decomposition: A noise-assisted data analysis method[J]. Advances in Adaptive Data Analysis, 2009, 1(1): 1-41.

[15] HAHN S L. Hilbert transforms in signal processing[M]. Norwood: Artech House on Demand, 1996: 1-55.

[16] HUANG N E. Computing instantaneous frequency by normalizing Hilbert transform: US, 6901353[P]. 2005-05-31.

作者简介:

王文哲 男,硕士研究生。主要研究方向:雷达信号处理、电子对抗理论与技术。

Tel.: 029-84787681

E-mail: 524940802@qq.com

吴华 女,硕士,副教授。主要研究方向:雷达辐射源识别、电子对抗理论与技术。

Tel.: 029-84787681

E-mail: 15129291728@163.com

北京航空航天大学学报

北航学报 赠阅

Subtle intrapulse feature extraction based on CEEMDAN for radar signals

WANG Wenzhe, WU Hua*, WANG Jingshang, ZHANG Qiang

(Aeronautics and Astronautics Engineering College, Air Force Engineering University, Xi'an 710038, China)

Abstract: Effective signal feature extraction builds the foundation of highly accurate radar emitter identification, a key function of the electronic warfare. Conventional features used in practice such as the pulse description words cannot fulfill the task in complex electromagnetic environments. An effective subtle intrapulse radar feature extraction method based on the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) was proposed. Radar signals were reconstructed by components provided by the CEEMDAN decomposition process, which was highly effective for non-stationary and nonlinear signals; the de-noising effect of the reconstruction on radar signals was validated through 1 000 Monte Carlo experiments, and an intrapulse feature space based on the reconstruction was designed. The identification accuracy of the proposed feature space was compared to that of the mainstream methods in the area, on 3 000 noise-contaminated signal samples supposed to be generated by 6 radar emitters, with different intrapulse modulation. Experimental results show that the samples are totally distinguishable in the proposed feature space, and the proposed method is more accurate in the comparison, especially in highly noisy environment, with accuracy above 90% at 0 dB signal to noise ratio (SNR).

Key words: feature extraction; complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN); empirical mode decomposition; noise assisted; radar emitter identification; signal reconstruction

北京航空航天大学学报

Received: 2016-05-17; **Accepted**: 2016-07-16; **Published online**: 2016-08-30 15:52

URL: www.cnki.net/kcms/detail/11.2625.V.20160830.1552.003.html

Foundation items: National Natural Science Foundation of China (61379104); Aeronautical Science Foundation of China (20152096019)

* **Corresponding author**. Tel.: 029-84787681　E-mail: 15129291728@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2016.0613

升力体飞行器能量近似最优倾斜转弯飞行策略

杨丁[1,2], 刘明[1,2,*], 杨凌霄[2], 杨明[1,2], 葛亚杰[2]
(1. 国防科学技术大学 航天科学与工程学院, 长沙 410073; 2. 空间物理重点实验室, 北京 100076)

摘 要: 升力体飞行器返回地球大气层内时受到热流、动压及过载等约束条件,为使得飞行器在倾斜转弯飞行过程中能量损失最小,需要研究一种能量最优的倾斜转弯机动飞行策略。本文的主要研究内容包括:从再入动力学模型出发,分析了升力体飞行器倾斜转弯的弹道特性,推导了终端速度与倾斜转弯幅度相关的解析解,提出了一种多约束条件下能量近似最优的倾斜转弯飞行策略;为进一步验证飞行策略的能量近似最优性,建立了能量最优的非线性轨迹优化模型,通过高斯伪谱法进行求解,获得能量最优的飞行轨迹。仿真结果表明,该飞行策略与优化方法获得的结果高度一致,并且该方法求解效率更高,工程应用性更强。

关 键 词: 升力体飞行器; 近似最优飞行策略; 倾斜转弯机动; 最优控制; 高斯伪谱法
中图分类号: V412.4
文献标识码: A **文章编号**: 1001-5965(2016)11-2540-08

不同于传统轴对称再入飞行器,升力体飞行器通常为面对称外形,具备较高的升阻比,其中,美国通用空天飞行器(CAV)[1-2]就是典型的高升阻比升力体飞行器。升力体飞行器在返回大气层内时,相比传统惯性轨迹,其利用高升阻比气动特性可进行大范围的倾斜转弯机动飞行,从而适应飞行器再入后返场点的灵活选择。飞行器在大气层内飞行过程中通过气动升力的倾斜实现转弯机动,受到气动阻力的作用持续减速,而当返场点确定的情况下,需调整飞行器的速度方向指向返场点,随后直线飞行完成返场任务,因此,飞行器通过倾斜机动使得速度方向指向返场点时对应的速度最大即倾斜转弯过程中能量损失最小是判断能否完成返场任务的关键点,规划满足驻点热流、过载及动压等多约束条件下能量最优的倾斜转弯飞行轨迹非常有意义。

针对倾斜转弯过程中能量损失最小问题,文献[3]在考虑动压、过载及热流密度等过程约束下,采用常值升阻比假设,获得了倾斜转弯转(BTT)弯半径等运动参数的解析关系,文献[4]分析了倾斜转弯的基本原理和控制特性;文献[5-11]分别采用遗传算法、直接法加序列二次规划法、高斯伪谱法对升力式飞行器再入机动轨迹进行了优化,为本文选择高斯伪谱法作为最优控制问题的优化求解方法提供了参考;文献[12]提出几何控制方案来解决再入飞行器的倾斜机动飞行问题。基于以上研究基础,从倾斜转弯的基本原理出发,推导给出能量最优的转弯机动策略,进一步采用优化算法进行仿真验证。

本文选取 CAV 为研究对象,考虑力、热等飞行过程约束,建立了飞行器动力学模型和高度-速度飞行走廊模型;基于升力体飞行器倾斜转弯基本原理和准平衡滑翔假设,推导获得了倾斜转弯终端速度与升阻比、倾侧角及终端弹道偏角的解析关系,提出了能量近似最优的倾斜转弯机动飞行策略;最后,建立了能量最优的非线性轨迹优化问题模型,采用高斯伪谱法获得了终端能量最优轨迹。仿真结果表明本文所提飞行策略与优化方法获得的结果高度一致,并且具有更高的计算效率。

收稿日期: 2016-07-20; **录用日期**: 2016-08-10; **网络出版时间**: 2016-08-31 15:28
网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20160831.1528.003.html
* **通讯作者**: Tel.: 010-88523261 E-mail: liumingbuaa@163.com

引用格式: 杨丁, 刘明, 杨凌霄, 等. 升力体飞行器能量近似最优倾斜转弯飞行策略[J]. 北京航空航天大学学报, 2016, 42(11): 2540-2547. YANG D, LIU M, YANG L X, et al. Energy suboptimal flight strategy of bank-to-turn maneuver for lifting vehicle [J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(11): 2540-2547 (in Chinese).

北航学报 赠阅

1　数学模型

1.1　动力学模型

为研究升力体飞行器大气层内再入倾斜转弯的弹道特性和规律,假设地球为均质无旋圆球,飞行器采用倾斜转弯飞行控制方式,因此,飞行器在空间内的再入运动方程可表述为

$$\begin{cases}\dot{r}=V\sin\gamma\\ \dot{\theta}=\dfrac{V\cos\gamma\sin\psi}{r\cos\phi}\\ \dot{\phi}=\dfrac{V\cos\gamma\cos\psi}{r}\\ \dot{V}=-D-g\sin\gamma\\ \dot{\gamma}=\dfrac{L\cos\sigma}{V}-\left(\dfrac{g}{V}-\dfrac{V}{r}\right)\cos\gamma\\ \dot{\psi}=\dfrac{L\sin\sigma}{V\cos\gamma}+\dfrac{V}{r}\cos\gamma\sin\psi\tan\phi\end{cases}\tag{1}$$

式中:r、θ、ϕ、V、γ、ψ 分别为飞行器质心到地心的距离、经度、纬度、飞行器相对地球速度、弹道倾角及航向角;D、L 分别为飞行器阻力加速度和升力加速度;σ 为飞行倾侧角;g 为引力加速度。

1.2　气动模型

本文选择洛马公司的通用航空飞行器 CAV-H[13] 升力飞行器作为研究对象,该飞行器最大升阻比可达 3.2。基于马赫数无关原理,当飞行马赫数大于 8 时,飞行器气动升力系数 C_L、气动阻力系数 C_D 可拟合为攻角 α 相关的表达式[14]:

$$\begin{cases}C_L=0.046\,75\alpha-0.105\,68\\ C_D=0.000\,508\alpha^2+0.004\,228\alpha+0.016\,1\end{cases}\tag{2}$$

1.3　飞行约束模型

飞行器在大气层内飞行约束主要包括过程约束和终端约束。其中,过程约束包括热流峰值约束、过载约束、动压约束以及控制量约束等[15],终端约束包括位置、速度以及航向角约束等。

1)热流峰值约束

$$\dot{Q}=\frac{C_1}{\sqrt{R_{\rm d}}}\rho^{0.5}V^{3.15}\leqslant\dot{Q}_{\max}\tag{3}$$

式中:$R_{\rm d}$ 为飞行器端头半径;C_1 为与飞行器特性相关的常数;ρ 为飞行高度的大气密度;$\dot{Q}_{\max}$ 为驻点热流密度约束值。取 $C_1/\sqrt{R_{\rm d}}=5.188\times10^{-8}$。

大气密度可采用如下指数函数近似计算:

$$\rho(H)=\rho_0{\rm e}^{-\beta H}\tag{4}$$

式中:H 为飞行高度;$\beta=\dfrac{1}{H_{\rm MCP}}$,$H_{\rm MCP}=7.11$ km 为大气刻度系数;$\rho_0=1.225$ kg/m^3 为海拔 0 km 密度。

由于热流峰值仅与高度、速度相关,可进一步转化为飞行高度的下边界约束:

$$H>H_{\dot{Q}_{\max}}(V)=\frac{2}{\beta}\ln\left(\frac{C_1V^{3.15}}{\dot{Q}_{\max}\sqrt{R_{\rm d}}}\sqrt{\rho_0}\right)\tag{5}$$

由式(5)可知,热流约束仅与飞行高度及速度相关,与飞行攻角无关。

2)动压约束

将大气密度公式代入动压计算公式并转换为高度和速度的关系式,得

$$q=\frac{1}{2}\rho_0{\rm e}^{-\beta H}V^2\geqslant q_{\min}\tag{6}$$

进一步转化为飞行高度的上边界约束:

$$H<H_{q_{\min}}(V)=\frac{1}{\beta}\ln\left(\frac{\rho_0V^2}{2q_{\min}}\right)\tag{7}$$

由式(7)可知,动压约束仅与飞行高度及速度相关,与飞行攻角无关。

3)过载约束

飞行器升力加速度及阻力加速度计算公式为

$$\begin{cases}L=\rho(H)V^2S_{\rm ref}C_L/(2m)\\ D=\rho(H)V^2S_{\rm ref}C_D/(2m)\end{cases}\tag{8}$$

过载约束可表述为

$$\frac{L\cos\alpha+D\sin\alpha}{g}<n_{\max}\tag{9}$$

进一步转化为飞行高度的下边界约束:

$$H>H_{n_{\max}}$$

$$H_{n_{\max}}=\frac{1}{\beta}\ln\left[\frac{\rho_0V^2S_{\rm ref}(C_L\cos\alpha+C_D\sin\alpha)}{2n_{\max}mg}\right]\tag{10}$$

由式(10)可知,过载的约束与飞行高度、速度及飞行攻角相关。

4)控制量约束

飞行器在大气层内飞行时,主要控制量为飞行攻角 α 和倾侧角 σ。对应控制量约束如下:

$$\begin{cases}7.5^\circ\leqslant\alpha\leqslant20^\circ\\ |\dot{\alpha}|\leqslant5(^\circ)/{\rm s}\\ |\sigma|\leqslant90^\circ\\ |\dot{\sigma}|\leqslant20(^\circ)/{\rm s}\end{cases}\tag{11}$$

5)终端约束

飞行器在大气层内倾斜转弯的幅度可通过航向角 ψ 的变化量来描述。为研究能量最优的倾斜转弯策略,考虑终端航向角约束,可表述为

$$\psi(t_{\rm f})=\psi_{\rm f}\tag{12}$$

式中:$\psi_{\rm f}$ 为终端航向角约束值。

2　飞行器倾斜转弯基本原理

飞行器在大气层内飞行的受力主要包括重力、气动升力及气动阻力,气动升力和气动阻力的改变实现纵向平面内飞行轨迹的控制,气动升力

北航学报 赠阅

大小及方向的改变实现倾斜转弯机动。飞行器在倾斜转弯机动过程中,其对应的运动学方程为

$$R_s = V^2/(L\sin\sigma) \tag{13}$$

式中:R_s 对应为飞行器倾斜转弯机动过程中的转弯半径;$L=\rho(H)V^2S_{ref}C_L/(2m)$,当飞行器飞行速度 $V\neq 0$,则可得

$$R_s = \frac{2m}{\rho(H)S_{ref}C_L\sin\sigma} \tag{14}$$

由式(14)可知,在飞行器质量一定的情况下,飞行转弯半径 R_s 与飞行高度 H、升力系数 C_L 及倾侧角 σ 相关。

以 CAV-H 飞行器为例,对应的飞行器倾斜转弯机动的转弯半径特性曲线如图 1 所示。

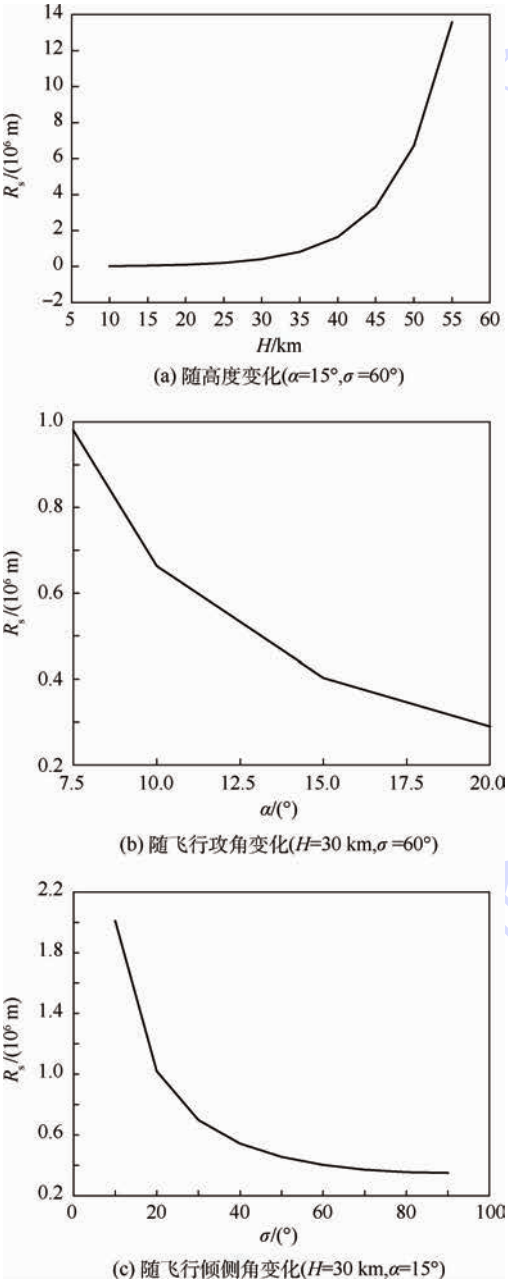


图 1 转弯半径特性曲线

Fig. 1 Characteristic curves of bank-to-turn radius

由图 1 可知:

1)飞行高度越低,对应飞行器倾斜转弯半径越小。

2)飞行倾侧角不为 0 时,飞行攻角越大,对应升力系数越大,则倾斜转弯半径越小。

3)倾斜转弯机动过程中,飞行倾侧角对应的侧向升力分量越大,则倾斜转弯半径越小。

飞行高度、攻角及倾侧角的匹配关系决定了飞行器倾斜转弯机动过程中的转弯半径,同时,一定飞行速度的条件下,飞行器实际飞行过程中还需考虑驻点热流、动压及法向过载等约束条件。在飞行攻角确定的条件下,可将飞行约束转化为飞行高度约束条件,如表述为式(5)、式(7)及式(10)。

3 能量近似最优倾斜转弯飞行策略

如图 2 所示,飞行器倾斜转弯机动飞行过程中,飞行速度由初始速度 V_0 变化为终端速度 V_f,在速度方向偏转角度 $\Delta\psi$ 一定的条件下,如何使得终端速度 V_f 最大是以下分析的重点。

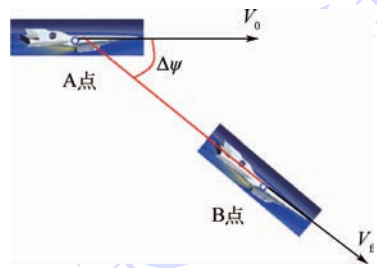


图 2 倾斜转弯示意图

Fig. 2 Schematic diagram of bank-to-turn

飞行器在倾斜转弯过程中弹道倾角较小,则 $\gamma\approx 0$,则动力学模型式(1)中速度对时间的导数表示为

$$\dot{V} = -D \tag{15}$$

式(1)中航向角表达式右边第一项 $L\sin\sigma/(V\cos\gamma)$ 为气动升力在当地水平面的分量对航向角变化率的影响,第二项 $V\cos\gamma\sin\psi\tan\phi/r$ 为地球曲率对航向角变化率的影响。可以发现,纬度 ϕ 越低,第二项对航向角变化率影响越小,当 $\phi\to 0$,则 $\tan\phi\to 0$。为研究气动力对航向角改变能力的客观规律,定义以下变量

$$\dot{\psi}_1 = \frac{L\sin\sigma}{V} \tag{16}$$

式(15)反映了飞行器速度大小变化与受力关系;式(16)反映了飞行器在水平面速度方向变化与受力关系。由式(15)和式(16)可得

北航学报 赠阅

$$-L\sin\sigma/D = V\frac{\mathrm{d}\psi_1}{\mathrm{d}V} \tag{17}$$

假设飞行器飞行过程中升阻比为 K_{LD},其中 $K_{LD}=L/D$。则可得

$$-\frac{\mathrm{d}V}{V}=\frac{\mathrm{d}\psi_1}{K_{LD}\sin\sigma} \tag{18}$$

对式(18)进行积分,则表达式可得

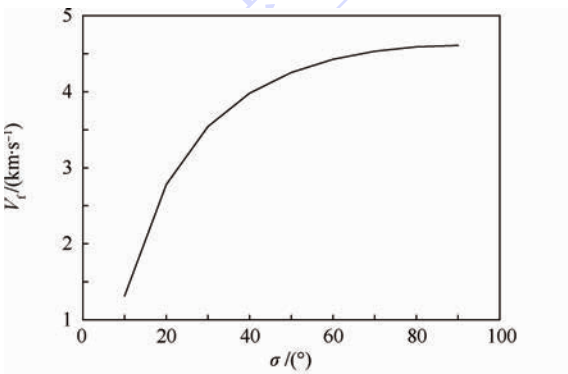
$$-\int_{V_0}^{V_f}\frac{1}{V}\mathrm{d}V=\int_{\psi_0}^{\psi_f}\frac{\mathrm{d}\psi_1}{K_{LD}\sin\sigma} \tag{19}$$

假设飞行过程中 K_{LD}、倾侧角 σ 为常值,由式(19)可得出

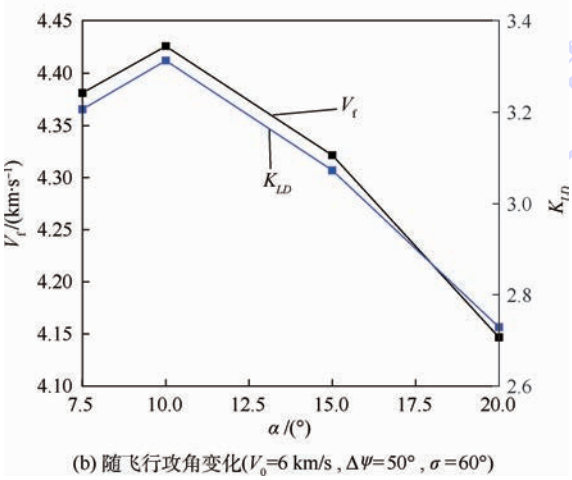
$$V_f=V_0\mathrm{e}^{-\frac{\psi_f-\psi_0}{K_{LD}\sin\sigma}} \tag{20}$$

由式(20)可以看出,在 V_0 及 $\Delta\psi$($\Delta\psi=\psi_f-\psi_0$)一定的情况下,飞行器转弯结束点的速度仅与飞行过程中的 K_{LD} 和 σ 相关。以 CAV-H 飞行器为例,对应飞行器倾斜转弯终端速度的特性曲线如图 3 所示。

为使得倾斜转弯的终端速度最大,飞行器在飞行过程中的升阻比 K_{LD} 及倾侧角 σ 的选取需尽可能大。



(a) 随飞行倾侧角变化(V_0=6 km/s, $\Delta\Psi$=50°, α=10°)



(b) 随飞行攻角变化(V_0=6 km/s, $\Delta\Psi$=50°, σ=60°)

图 3　倾斜转弯终端速度的特性曲线

Fig. 3　Characteristic curves of bank-to-turn terminal velocity

考虑飞行器飞行过程中的热流、动压及法向过载约束,对应的飞行高度-速度(H-V)约束边界如图 4 所示。

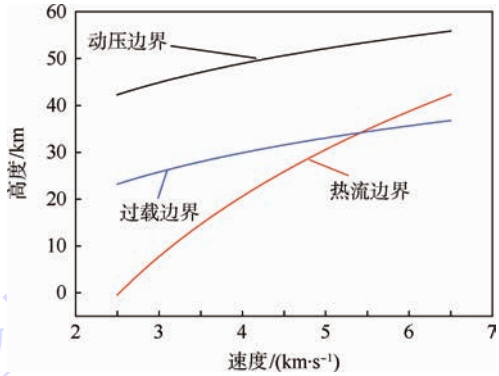


图 4　高度-速度走廊边界示意图

Fig. 4　Schematic diagram of H-V corridor boundary

由图 4 可知,在飞行速度较大时,飞行边界由热流约束决定,随着飞行速度降低,飞行边界由过载约束决定。

由以上理论推导及弹道特性分析结果可知飞行器能量近似最优的倾斜转弯飞行策略如下:

1)基于最优升阻比攻角-速度飞行剖面及飞行过程约束,规划高度-速度飞行走廊。

2)基于高度-速度飞行走廊下边界获得倾斜机动能量最优高度-速度飞行剖面。

3)纵向基于高度-速度及攻角-速度剖面、横向通过攻角及高度实时求解倾侧角,获得能量近似最优倾斜转弯轨迹。

能量近似最优倾斜转弯机动策略的流程图如图 5 所示。

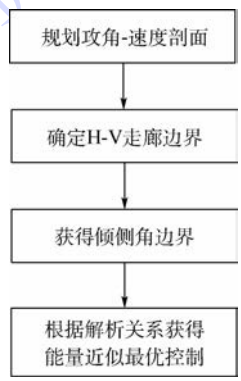


图 5　能量近似最优飞行策略流程图

Fig. 5　Flowchart of energy suboptimal flight strategy

4　最优转弯转弯轨迹优化

为获取倾斜转弯能量最优的轨迹,可以通过间接法和直接法求解建立的相应最优控制问题。考虑到间接法求解多点边值问题时对初值估计精

北航学报 赠阅

度要求高且对于过程约束问题求解困难,本文采用直接法对性能指标进行寻优,其基本思路是,通过对控制变量和状态变量进行参数化离散,从而将轨迹优化这一连续的最优控制问题转换成有限维的非线性规划(NLP)问题,并采用成熟的非线性规划求解工具进行求解。直接法应用时具有不需要推导一阶最优性条件、对初值估计精度要求不高以及通用性好等优点。

离散后的轨迹优化问题如下。

状态变量:

$$\boldsymbol{x}(\tau)\approx\sum_{i=0}^{N}\boldsymbol{X}(\tau_i)L_i(\tau)\qquad L_i(\tau)=\prod_{j=0,j\neq i}^{N}\frac{\tau-\tau_j}{\tau_i-\tau_j} \tag{21}$$

式中:τ_i 为配点,$\tau_i\in(-1,1)$,取自 N 阶 Legendre 多项式 $P_K(\tau)$ 的根。

控制变量:

$$\boldsymbol{u}(\tau)\approx\sum_{i=1}^{N}\boldsymbol{U}(\tau_i)L_i'(\tau)\qquad L_i'(\tau)=\prod_{j=1,j\neq i}^{N}\frac{\tau-\tau_j}{\tau_i-\tau_j} \tag{22}$$

运动方程约束:

$$\sum_{i=0}^{N}\dot{L}_i(\tau_k)\boldsymbol{X}_i=\frac{t_f-t_0}{2}f(\boldsymbol{X}_k,\boldsymbol{U}_k,\tau_k;t_0,t_f)$$

$$\dot{L}_i(\tau_k)=\sum_{l=0,l\neq i}^{N}\frac{\prod_{j=0,j\neq i,l}^{N}(\tau_k-\tau_j)}{\prod_{j=0,j\neq i}^{N}(\tau_i-\tau_j)}\qquad k=1,2,\cdots,N \tag{23}$$

终端状态约束:

$$\boldsymbol{X}_f=\boldsymbol{X}_0+\frac{t_f-t_0}{2}\sum_{k=1}^{N}\omega_k f(\boldsymbol{X}_k,\boldsymbol{U}_k,\tau_k;t_0,t_f) \tag{24}$$

式中:ω_k 为高斯积分权重系数。

过程约束和端点约束:

$$\begin{cases}\boldsymbol{\Phi}(\boldsymbol{X}_0,t_0,\boldsymbol{X}_f,t_f)=\boldsymbol{0}\\ \boldsymbol{C}(\boldsymbol{X}_k,\boldsymbol{U}_k,\tau_k;t_0,t_f)\leqslant\boldsymbol{0}\end{cases}\qquad k=1,2,\cdots,N \tag{25}$$

目标函数:

$$\boldsymbol{J}=\phi(\boldsymbol{X}_0,t_0,\boldsymbol{X}_f,t_f)+\frac{t_f-t_0}{2}\sum_{k=1}^{N}\omega_k G(\boldsymbol{X}_k,\boldsymbol{U}_k,\tau_k;t_0,t_f) \tag{26}$$

经过上述转化后,将连续最优控制问题转化为 NLP 问题,然后可采用 NLP 方法进行求解。

为获取能量最优的倾斜转弯轨迹,建立能量最优的轨迹优化问题。

1)性能指标选取

为了验证最优飞行策略的最优性,选择终端速度最大作为性能指标。

$$\max J=V(t_f) \tag{27}$$

2)控制方程

满足非线性动力学方程式(1)。

3)过程约束

满足热流峰值、动压及过载等过程约束方程式(3)、式(6)、式(9),满足控制量约束方程式(11)。

4)终端约束

满足终端约束方程式(12)。

通过求解该最优控制问题,即可获得能量最优的轨迹。

5 仿真算例

本文选择 CAV-H 作为研究对象,飞行器质量 907 kg,飞行过程中最大热流峰值、最小动压及最大过载约束分别为 3 000 kW/m^2、10 kPa 及 3g。计算机环境为 Windows XP 2.5 GHz,仿真软件基于 VC++6.0。基于能量最优的倾斜转弯飞行策略及高斯伪谱法分别求解飞行轨迹,仿真计算初始条件如表 1 所示,计算算例如表 2 所示。

仿真结果如图 6 ~ 图 11 所示。

表 1 仿真初始条件

Table 1 Initial conditions of simulation

初始条件	高度/km	速度/(m·s^{-1})	弹道倾角/(°)	经度/(°)	纬度/(°)	弹道偏角/(°)
数值	50	6 000	0	0	0	0

表 2 典型计算算例

Table 2 Typical calculation example

状态	高斯伪谱法算例 1	高斯伪谱法算例 2	高斯伪谱法算例 3
$\Delta\psi$/(°)	45	90	135

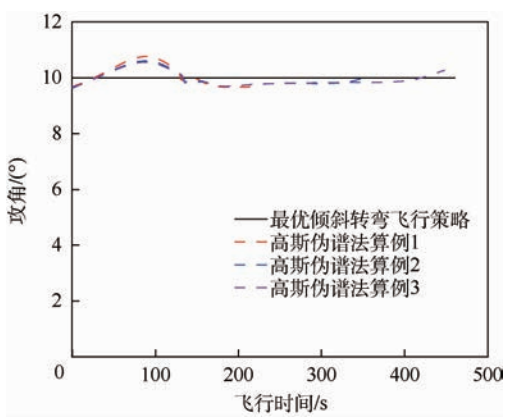


图 6 攻角随飞行时间变化

Fig. 6 Angle of attack changes with flight time

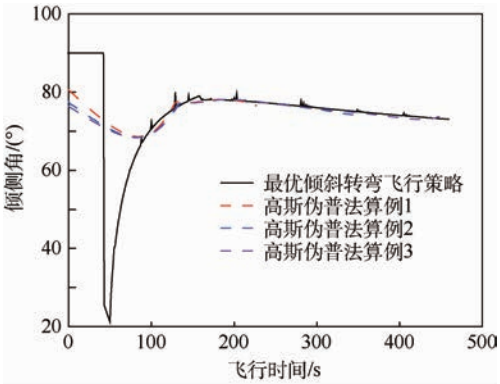


图 7　倾侧角随飞行时间变化

Fig. 7　Bank angle changes with flight time

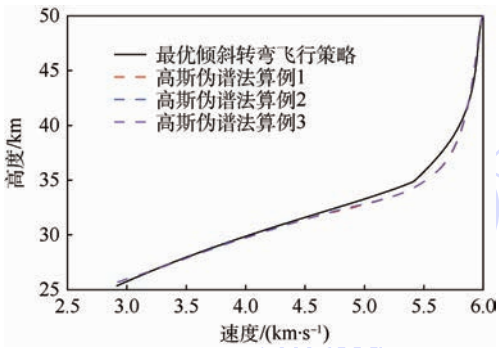


图 8　高度随速度变化关系

Fig. 8　Altitude changes with velocity

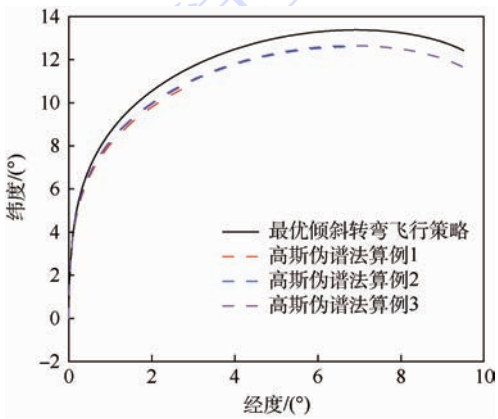


图 9　地面轨迹

Fig. 9　Ground track

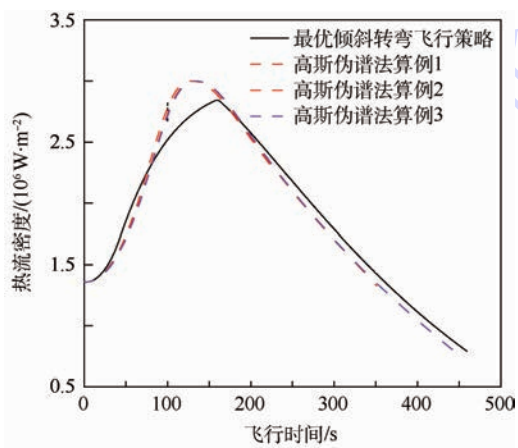


图 10　热流密度随飞行时间变化

Fig. 10　Heating rate changes with flight time

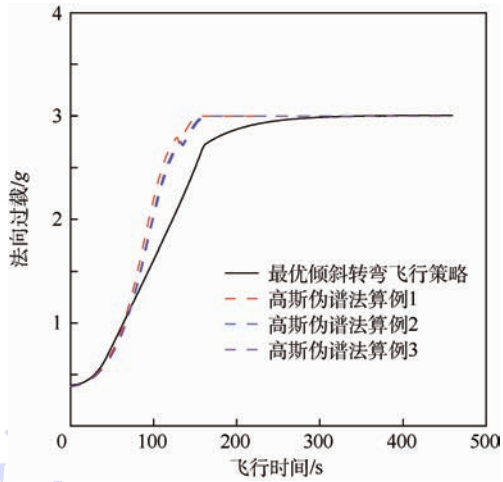


图 11　法向过载随飞行时间变化

Fig. 11　Normal load factor changes with flight time

针对 3 个典型算例进行了仿真,轨迹优化结果如表 3 所示。由表 3 可以看出,基于能量近似最优倾斜转弯飞行策略对应的终端速度与高斯伪谱法的优化结果相比,飞行热流密度、法向过载等均在过程约束范围内,终端速度偏差小于 20 m/s,高度小于 0.4 km。同时,基于高斯伪谱法直接寻优的攻角、倾侧角控制量与能量近似最优转弯策略对应的攻角、倾侧角飞行策略是一致的。

轨迹优化所需时间如表 4 所示,可以看出,基于能量最优倾斜转弯飞行策略计算最优轨迹所需时间均小于 1 s,高斯伪谱法优化轨迹所需时间均大于 600 s。在优化结果性能一致的情况下,本文提出的能量最优倾斜转弯机动飞行策略具有更高的计算效率。

表 3　轨迹优化结果

Table 3　Trajectory optimization results

状态	终端弹道偏角/(°)	终端速度 V_f/(m·s^{-1})		终端高度 H_f/km	
		近似最优倾斜转弯飞行策略	高斯伪谱法	近似最优转弯飞行策略	高斯伪谱法
算例 1	45	4 702	4 720	32.3	32.1
算例 2	90	3 712	3 726	28.8	28.9
算例 3	135	2 917	2 926	25.3	25.7

表 4　轨迹优化所需时间

Table 4　Time of trajectory optimization

s

状态	近似最优倾斜转弯飞行策略计算时间	高斯伪谱法计算时间
算例 1	<1	630
算例 2	<1	972
算例 3	<1	1 316

6　结　论

升力体外形飞行器大气层内能量最优的倾斜转弯轨迹优化问题是一个较为复杂的最优控制问题。本文基于飞行器倾斜转弯的基本原理和准平

北航学报 赠阅

衡滑翔假设,从动力学模型出发,推导获得了倾斜转弯终端速度与升阻比、倾侧角及终端弹道偏角相关的解析解,提出了能量最优的倾斜转弯机动飞行策略,并与高斯伪谱法的优化轨迹进行了对比,结果表明最优轨迹高度一致,验证了该方法的最优性。并且,该方法的计算效率可大幅度提高,1 s 之内可完成多约束条件下的最优倾侧转弯轨迹规划,具有较强的工程实际应用价值。

参考文献(References)

[1] 童雄辉. 美国未来全球快速精确打击体系预测分析[J]. 导弹与航天运载技术,2008(5):57-61.
TONG X H. Forecast and analysis of USA's future conventional prompt global strike system[J]. Missile and Space Vehicle, 2008(5):57-61(in Chinese).

[2] 康开华,才满瑞,伍赣湘,等. 美国高超声速技术飞行器[J]. 航天制造技术,2010(1):7-11.
KANG K H,CAI M R,WU G X,et al. U. S. hypersonic technology vehicle[J]. Aerospace Manufacturing Technology, 2010(1):7-11(in Chinese).

[3] 李广华,张洪波,汤国建. 高超声速滑翔飞行器典型弹道特性分析[J]. 宇航学报,2015,36(4):397-403.
LI G H,ZHANG H B,TANG G J. Typical trajectory characteristics of hypersonic glide vehicle[J]. Journal of Astronautics, 2015,36(4):397-403(in Chinese).

[4] 陈小庆,侯中喜,刘建霞. 高超声速滑翔飞行器倾斜转弯分析及控制系统设计[J]. 国防科技大学学报,2012,34(3):17-23.
CHEN X Q,HOU Z X,LIU J X. Analysis and controller design of bank-to-turn system for hypersonic gliding vehicle[J]. Journal of National University of Defense Technology,2012,34(3):17-23(in Chinese).

[5] 张忠峰,高云峰,宝音贺西. 高超声速飞行器机动航迹优化[J]. 弹箭与制导学报,2009,29(3):164-172.
ZHANG Z F,GAO Y F,BAOYIN H X. Optimization of maneuvering flight path of hypersonic vehicle[J]. Journal of Projectiles,Rockets, Missiles and Guidance, 2009, 29(3):164-172(in Chinese).

[6] 黄育秋,何麟书. 升力式再入飞行器多约束多阶段弹道优化设计[J]. 哈尔滨工业大学学报,2011,43(7):144-148.
HUANG Y Q,HE L S. Multi-constraints and multi-phases trajectory optimization of lift reentry vehicle[J]. Journal of Harbin Institute of Technology,2011,43(7):144-148(in Chinese).

[7] 周文雅,杨涤,李顺利. 利用高斯伪谱法求解升力航天器最优再入轨迹[J]. 南京理工大学学报,2010,34(1):85-90.
ZHOU W Y, YANG D, LI S L. Using Gauss pseudospectral method to solve optimal reentry trajectory for lifting vehicle[J]. Journal of Nanjing University of Science and Technology,2010, 34(1):85-90(in Chinese).

[8] 周浩,陈万春. 基于拟平衡滑翔的横程最大轨迹研[J]. 飞行力学,2010,28(3):64-68.
ZHOU H,CHEN W C. Maximum cross range trajectory optimization based on quasi-equilibrium glide conditions[J]. Flight Dynamics,2010,28(3):64-68(in Chinese).

[9] 张红文,张科南,陈万春. 带静态参数的高超声速飞行器轨迹优化算法[J]. 北京航空航天大学学报,2014,40(2):141-147.
ZHANG H W,ZHANG K N,CHEN W C. Indirect method for trajectory optimization of hypersonic vehicle with static parameters[J]. Journal of Beijing University of Aeronautics and Astronautics,2014,40(2):141-147(in Chinese).

[10] 李瑜,杨志红,崔乃刚. 助推-滑翔导弹最大射程优化[J]. 弹道学报,2008,20(4):53-56.
LI Y,YANG Z H,CUI N G. Optimization of maximum range for boost-glide missile[J]. Journal of Ballistics,2008,20(4):53-56(in Chinese).

[11] ZHANG K N,CHEN W C. Reentry vehicle constrained trajectory optimization[C]// 17th AIAA International Space Planes and Hypersonic Systems and Technologies Conference 2011. Reston:AIAA,2011.

[12] ZHAO Z G. A rapid reentry trajectory planning method for CAV[C]//2010 3rd International Symposium on Systems and Control in Aeronautics and Astronautics. Piscataway NJ: IEEE Press,2010:1255-1260.

[13] VINH N X. Flight mechanics of high-performance aircraft[M]. Cambridge: Cambridge University Press,1993:374.

[14] TERRY H P. A common aero vehicle model, description, and employment guidance[R]. Arlington: Schafer Corporation for AFRL and AFSPC,2003.

[15] 汪朝群. 航天飞机的再入飞行走廊及再入飞行轨迹[J]. 航天控制,1990(1):1-10.
WANG C Q. On entry flight corridor and entry trajectory for space shuttle reentry[J]. Aerospace Control,1990(1):1-10(in Chinese).

作者简介:

杨丁 男,博士,高级工程师。主要研究方向:飞行器动力学与控制。
Tel.:010-88523743
E-mail:50732554@qq.com

刘明 男,博士,高级工程师。主要研究方向:飞行器动力学与控制。
Tel.:010-88523261
E-mail:liumingbuaa@163.com

杨凌霄 女,硕士。主要研究方向:飞行器设计、导航制导与控制等。
Tel.:010-88520815
E-mail:799390175@qq.com

杨明 男,博士,工程师。主要研究方向:飞行器动力学与控制。
Tel.:010-88523743
E-mail:ym20081210163@163.com

葛亚杰 女,硕士。主要研究方向:飞行器设计、导航制导与控制等。
Tel.:010-88520815
E-mail:geyajie87@sina.com

北航学报 赠阅

Energy suboptimal flight strategy of bank-to-turn maneuver for lifting vehicle

YANG Ding[1,2], LIU Ming[1,2,*], YANG Lingxiao[2], YANG Ming[1,2], GE Yajie[2]

(1. College of Aerospace Science and Engineering, National University of Defense Technology, Changsha 410073, China;
2. Science and Technology on Space Physics Laboratory, Beijing 100076, China)

Abstract: When lifting vehicle is returning into earth's atmosphere, some path constraints including heating rate, dynamic pressure and overload should be enforced strictly. In order to reduce the energy losing during the performance of bank-to-turn maneuver, it is necessary to develop a suboptimal flight strategy. The main research contents of this paper are as follows. The trajectory characteristic of bank-to-turn maneuver is analyzed via studying the entry dynamics. The analytical solution to terminal velocity at the special terminal heading angle is derived. Furthermore, a suboptimal flight strategy, which considers the path constraints, is proposed based on this analytical solution. In order to verify the optimality of this flight strategy, a trajectory optimization problem is formulated so as to provide optimal energy trajectory. Then, Gauss pseudospectral method is used to solve this problem. The results show that both trajectories are highly consistent, and the flight strategy is more efficient and has greater engineering applicable values.

Key words: lifting vehicle; suboptimal flight strategy; bank-to-turn maneuver; optimal control; Gauss pseudospectral method

北京航空航天大学学报

Received: 2016-07-20; **Accepted**: 2016-08-10; **Published online**: 2016-08-31 15:28
URL: www.cnki.net/kcms/detail/11.2625.V.20160831.1528.003.html
* **Corresponding author**. Tel.: 010-88523261　E-mail: liumingbuaa@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2015.0765

惯性导航系统动基座传递对准可观测性分析

黄帅，蔡洪*，丁智坚

（国防科学技术大学 航天科学与工程学院，长沙 410073）

摘 要：针对机载武器惯性导航系统动基座传递对准问题，以系统原始的非线性模型为分析对象，根据可观测性的基本定义，将系统可观测性问题转化为判断系统状态量是否存在唯一解。通过分析系统各状态量的解同载体运动方式之间的关系，得出了系统各状态量可观测的充分条件，并设计了相应的机动方案。结果表明，载体在不同加速度下的平移运动可增强惯性器件误差、安装误差角和姿态的可观测性，而载体的旋转运动则有助于提高杆臂的可观测性。与传统方法不同，该方法分析过程简洁明了，且物理意义明确，为合理规划载体的机动方案提供了理论指导，同时，也为其他非线性系统的分析与设计提供了新的思路。最后，采用扩展 Kalman 滤波（EKF）对系统进行了仿真，仿真结果验证了理论分析结果的正确性和有效性。

关 键 词：可观测性分析；传递对准；机动方案；动基座；惯性导航系统

中图分类号：V448.12

文献标识码：A **文章编号**：1001-5965(2016)11-2548-07

对于机载武器，其惯性导航系统（简称惯导系统）的初始化采用传递对准[1]完成，以克服由于基座运动产生的各种干扰影响以及惯性传感器的精度限制，在短时间内实现高精度的初始对准。

惯导系统动基座传递对准的过程就是在高精度主惯导的辅助条件下，利用主、子惯导系统导航参数或测量信息的差值作为观测量，对主、子惯导系统间的失准角、速度误差、位置误差等导航参数误差及惯性元件误差进行估计，并根据估计结果对子惯导系统进行高精度对准，从而达到提高子惯导导航精度的目的[2]。

目前，针对惯导系统动基座传递对准技术的研究主要集中在匹配算法的应用和改进[3]、运载体弹性形变及杆臂效应的影响分析与补偿[4-6]、时间延时效应的分析与补偿[6-8]、对准精度的评估[9]与非线性滤波算法研究[10]等方面。而系统可观测性分析作为对系统状态可估计能力的描述，是 Kalman 滤波收敛的前提条件[11-12]，同样也引起了广泛研究[13-14]。

然而，由于动基座条件下传递对准系统是非线性的、时变的，虽然理论上可通过检查 Grammian 矩阵的秩来判断其可观测性，但计算量较大，且只能通过数值分析的方法对系统的性质进行研究，在实际应用中十分不便[11]。传统的可观测性分析方法通常用线性化模型代替原系统非线性模型，利用分段定常系统（Piece-Wise Constant System，PWCS）理论[15-16]，通过求解系统可观测性矩阵的秩或奇异值判断系统的可观测性。这种方法虽然大大简化了非线性时变对准模型的可观测性分析，但存在以下问题[17]：①当系统的维数较高或分段过多时，工作量较大；②经过线性化与分段定常两次近似后得到的系统模型与原始的系统模型之间存在一定的误差；③不能充分揭示系统可观测性与载体运动状态之间的联系和内在机理。

针对上述问题，本文从惯导系统动基座传递对准原始的非线性模型出发，根据可观测性的基

收稿日期：2015-11-20；**录用日期**：2016-04-15；**网络出版时间**：2016-04-19 17:33
网络出版地址：www.cnki.net/kcms/detail/11.2625.V.20160419.1733.001.html
* **通讯作者**：Tel.：0731-87001191 E-mail：hcai@nudt.edu.cn

引用格式：黄帅，蔡洪，丁智坚. 惯性导航系统动基座传递对准可观测性分析[J]. 北京航空航天大学学报，2016，42(11)：2548-2554. HUANG S，CAI H，DING Z J. Observability analysis for transfer alignment of inertial navigation system on moving base[J]. Journal of Beijing University of Aeronautics and Astronautics，2016，42(11)：2548-2554 (in Chinese).

本定义对系统状态变量进行了分析,给出了系统可观测性的充分条件,深刻地揭示出系统可观测性同载体运动状态的内在联系。该方法避免了由于线性化过程而引入的模型误差,且无需经过复杂的数值计算过程,极大地减小了计算量,并为合理规划载体的机动方案提供了理论依据。

1　系统模型建立

在惯导系统动基座传递对准过程中,存在主、子两套惯导系统,为便于说明问题,同时避免发生混淆,对坐标系及部分相关符号进行了定义。定义主惯导体坐标系为 a 系;子惯导体坐标系为 b 系;采用当地地理坐标系(即北-天-东坐标系)作为导航坐标系,记为 n 系,并记子惯导计算机建立的数字平台坐标系为 n′系;定义惯性坐标系 i 系,并与初始时刻的地面坐标系 e 系和导航坐标系 n 系重合。考虑惯性器件的测量误差,子惯导系统的运动学方程可表示为[18]

$$\dot{\boldsymbol{r}}^{\mathrm{n}}=\boldsymbol{v}^{\mathrm{n}}$$

$$\text{或}\begin{cases}\dot{\lambda}=v_{\mathrm{E}}\sec L/(R+h)\\ \dot{L}=v_{\mathrm{N}}/(R+h)\\ \dot{h}=v_{\mathrm{U}}\end{cases}\tag{1}$$

$$\dot{\boldsymbol{v}}^{\mathrm{n}}=\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}(\boldsymbol{f}^{\mathrm{b}}-\boldsymbol{b}_{\mathrm{a}})-(2\boldsymbol{\omega}_{\mathrm{ie}}^{\mathrm{n}}+\boldsymbol{\omega}_{\mathrm{en}}^{\mathrm{n}})\times\boldsymbol{v}^{\mathrm{n}}+\boldsymbol{g}^{\mathrm{n}}\tag{2}$$

$$\dot{\boldsymbol{C}}_{\mathrm{b}}^{\mathrm{n}}=\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}[\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}\times]\tag{3}$$

$$\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}=(\boldsymbol{\omega}^{\mathrm{b}}-\boldsymbol{b}_{\mathrm{g}})-\boldsymbol{C}_{\mathrm{n}}^{\mathrm{b}}(\boldsymbol{\omega}_{\mathrm{ie}}^{\mathrm{n}}+\boldsymbol{\omega}_{\mathrm{en}}^{\mathrm{n}})\tag{4}$$

式中:$\boldsymbol{r}^{\mathrm{n}}$ 为惯性测量单元(Inertial Measurement Unit,IMU)在 n 系下的位置;λ、L 和 h 分别为经度、纬度和高程;R 为地球平均半径;$\boldsymbol{v}^{\mathrm{n}}=[v_{\mathrm{N}}\quad v_{\mathrm{U}}\quad v_{\mathrm{E}}]^{\mathrm{T}}$ 为 IMU 在 n 系下的地速(下标 N、U、E 分别代表速度的北向、天向和东向分量);$\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}$ 为子惯导体坐标系 b 系到其导航坐标系 n 系的坐标转换矩阵,也即子惯导系统输出的姿态矩阵;$\boldsymbol{\omega}_{\mathrm{ie}}^{\mathrm{n}}=[\omega_{\mathrm{ie}}\cos L\quad\omega_{\mathrm{ie}}\sin L\quad 0]^{\mathrm{T}}$ 为在 n 系下表示的地球自转角速度;$\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}$ 为弹体相对导航系的转动角速度在子惯导体坐标系下的投影;$\boldsymbol{\omega}_{\mathrm{en}}^{\mathrm{n}}=[v_{\mathrm{E}}\quad v_{\mathrm{E}}\tan L\quad -v_{\mathrm{N}}]^{\mathrm{T}}/(R+h)$ 为导航坐标系 n 系相对于地面坐标系 e 系的转动角速度在 n 系下的投影;$\boldsymbol{f}^{\mathrm{b}}$ 为加速度计测量的比力;$\boldsymbol{\omega}^{\mathrm{b}}$ 为陀螺仪测量的弹体相对于惯性空间的转动角速度;$\boldsymbol{g}^{\mathrm{n}}$ 为在 n 系表示的当地重力加速度;× 表示矢量积运算;$[\boldsymbol{a}\times]$ 表示为向量 $\boldsymbol{a}$ 的反对称矩阵;$\boldsymbol{b}_{\mathrm{a}}$、$\boldsymbol{b}_{\mathrm{g}}$ 分别为加速度计零偏和陀螺漂移,本文将其视为随机常数,即

$$\dot{\boldsymbol{b}}_{\mathrm{a}}=0\tag{5}$$

$$\dot{\boldsymbol{b}}_{\mathrm{g}}=0\tag{6}$$

由于主、子惯导安装位置不同,杆臂的影响导致主、子惯导输出的位置、速度存在如下关系:

$$\boldsymbol{r}_{\mathrm{m}}^{\mathrm{n}}=\boldsymbol{r}^{\mathrm{n}}+\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}\boldsymbol{l}^{\mathrm{b}}\tag{7}$$

$$\boldsymbol{v}_{\mathrm{m}}^{\mathrm{n}}=\boldsymbol{v}^{n}+\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}(\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}\times\boldsymbol{l}^{\mathrm{b}})\tag{8}$$

式中:$\boldsymbol{r}_{\mathrm{m}}^{\mathrm{n}}$ 和 $\boldsymbol{v}_{\mathrm{m}}^{\mathrm{n}}$ 分别为主惯导系统输出的载体在导航系下的位置和地速。

在不考虑载体变形及弹体弹性形变的前提下,主、子惯导间的安装关系不会发生改变,杆臂矢量可视为常值,即

$$\dot{\boldsymbol{l}}^{\mathrm{b}}=0\tag{9}$$

式(1)~式(9)构成了惯导系统动基座传递对准模型,其状态量包括子惯导所在弹体的位置、速度、姿态以及子惯导的惯性器件误差(加速度计零偏和陀螺漂移)和杆臂矢量,系统观测量则包括主惯导实时输出的精确的位置、速度和姿态信息以及子惯导加速度计和陀螺仪的测量输出。

2　可观测性分析

根据文献[19]中关于系统可观测性定义可知,若系统可观,则意味着根据系统的输出和输入信息即可唯一确定系统的初始状态量。因此,系统的可观测性分析问题可转化为判断系统状态量是否存在唯一解的问题。与此同时,通过对状态量解的分析,可以看出载体机动方式对状态量可观测性的影响,进而可获得满足系统可观测的充分条件。

利用改进方法,本节将对系统的状态量逐一进行分析,并给出系统可观测的充分条件,为进一步设计载体的运动轨迹提供理论指导。

1)加速度计零偏的可观测性

假设载体作直线运动,且姿态不发生变化,即 $\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}=0$,$\dot{\boldsymbol{C}}_{\mathrm{b}}^{\mathrm{n}}=0$,则式(8)可写为

$$\boldsymbol{v}_{\mathrm{m}}^{\mathrm{n}}=\boldsymbol{v}^{\mathrm{n}}\tag{10}$$

即主、子惯导速度相等,不受杆臂影响。

对式(10)两端求导,并代入式(2),有

$$\dot{\boldsymbol{v}}_{\mathrm{m}}^{\mathrm{n}}=\dot{\boldsymbol{v}}^{\mathrm{n}}=\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}(\boldsymbol{f}^{\mathrm{b}}-\boldsymbol{b}_{\mathrm{a}})-(2\boldsymbol{\omega}_{\mathrm{ie}}^{\mathrm{n}}+\boldsymbol{\omega}_{\mathrm{en}}^{\mathrm{n}})\times\boldsymbol{v}_{\mathrm{m}}^{\mathrm{n}}+\boldsymbol{g}^{\mathrm{n}}\tag{11}$$

由式(11)可解得

$$\boldsymbol{b}_{\mathrm{a}}=\boldsymbol{f}^{\mathrm{b}}-\boldsymbol{C}_{\mathrm{n}}^{\mathrm{b}}[\dot{\boldsymbol{v}}_{\mathrm{m}}^{\mathrm{n}}+(2\boldsymbol{\omega}_{\mathrm{ie}}^{\mathrm{n}}+\boldsymbol{\omega}_{\mathrm{en}}^{\mathrm{n}})\times\boldsymbol{v}_{\mathrm{m}}^{\mathrm{n}}-\boldsymbol{g}^{\mathrm{n}}]\tag{12}$$

式中:$\boldsymbol{f}^{\mathrm{b}}$ 由子惯导加速度计测量获得,$\boldsymbol{\omega}_{\mathrm{ie}}^{\mathrm{n}}$ 和 $\boldsymbol{\omega}_{\mathrm{en}}^{\mathrm{n}}$ 可通过主惯导输出的位置及速度信息计算获得,而 $\boldsymbol{g}^{\mathrm{n}}$ 在位置矢量已知后,即为已知(由杆臂造成的位置误差较小,可忽略不计)。故式(11)等号右端除 $\boldsymbol{C}_{\mathrm{n}}^{\mathrm{b}}$ 外,各项均为已知。因此,令

北航学报 赠阅

$$\boldsymbol{f}_{\mathrm{m}}^{\mathrm{a}}=\dot{\boldsymbol{v}}_{\mathrm{m}}^{\mathrm{n}}+(2\boldsymbol{\omega}_{\mathrm{ie}}^{\mathrm{n}}+\boldsymbol{\omega}_{\mathrm{en}}^{\mathrm{n}})\times\boldsymbol{v}_{\mathrm{m}}^{\mathrm{n}}-\boldsymbol{g}^{\mathrm{n}}=-\boldsymbol{g}^{\mathrm{n}} \tag{13}$$

即载体加速度为0,便可唯一确定 $\boldsymbol{b}_{\mathrm{a}}$,故加速度计零偏是可观测的。

2) 陀螺漂移的可观测性

已知在直线运动状态下,$\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}=0$,则式(4)可改写为

$$\boldsymbol{b}_{\mathrm{g}}=\boldsymbol{\omega}^{\mathrm{b}}-\boldsymbol{C}_{\mathrm{n}}^{\mathrm{b}}(\boldsymbol{\omega}_{\mathrm{ie}}^{\mathrm{n}}+\boldsymbol{\omega}_{\mathrm{en}}^{\mathrm{n}}) \tag{14}$$

对式(11)两端求导,并根据式(5),可得

$$\ddot{\boldsymbol{v}}_{\mathrm{m}}^{\mathrm{n}}=\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}\dot{\boldsymbol{f}}^{\mathrm{b}}-(2\dot{\boldsymbol{\omega}}_{\mathrm{ie}}^{\mathrm{n}}+\dot{\boldsymbol{\omega}}_{\mathrm{en}}^{\mathrm{n}})\times\boldsymbol{v}_{\mathrm{m}}^{\mathrm{n}}-(2\boldsymbol{\omega}_{\mathrm{ie}}^{\mathrm{n}}+\boldsymbol{\omega}_{\mathrm{en}}^{\mathrm{n}})\times\dot{\boldsymbol{v}}_{\mathrm{m}}^{\mathrm{n}}+\dot{\boldsymbol{g}}^{\mathrm{n}} \tag{15}$$

整理得

$$\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}\dot{\boldsymbol{f}}^{\mathrm{b}}=\ddot{\boldsymbol{v}}_{\mathrm{m}}^{\mathrm{n}}+(2\dot{\boldsymbol{\omega}}_{\mathrm{ie}}^{\mathrm{n}}+\dot{\boldsymbol{\omega}}_{\mathrm{en}}^{\mathrm{n}})\times\boldsymbol{v}_{\mathrm{m}}^{\mathrm{n}}+(2\boldsymbol{\omega}_{\mathrm{ie}}^{\mathrm{n}}+\boldsymbol{\omega}_{\mathrm{en}}^{\mathrm{n}})\times\dot{\boldsymbol{v}}_{\mathrm{m}}^{\mathrm{n}}-\dot{\boldsymbol{g}}^{\mathrm{n}} \tag{16}$$

式中:$\boldsymbol{f}^{\mathrm{b}}$、$\boldsymbol{\omega}_{\mathrm{ie}}^{\mathrm{n}}$、$\boldsymbol{\omega}_{\mathrm{en}}^{\mathrm{n}}$和$\boldsymbol{g}^{\mathrm{n}}$各项已知,故其导数也均为已知量。

由文献[20-21]可知,若存在两个不相等的时刻 t_1、t_2,满足 $\dot{\boldsymbol{f}}^{\mathrm{b}}(t_1)$和$\dot{\boldsymbol{f}}^{\mathrm{b}}(t_2)$线性不相关,则可确定该运动状态下的姿态矩阵 $\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}$。

将已知的 $\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}$ 代入式(14),即可确定 $\boldsymbol{b}_{\mathrm{g}}$,即陀螺漂移是可观测的。注意,此条件同样可满足加速度计零偏可观测。

3) 姿态的可观测性

直线运动状态下,式(11)成立,其中 $\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}$ 项满足

$$\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}=\boldsymbol{\phi}_{\mathrm{n}}(t_0,t)\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}(t_0)\boldsymbol{\phi}_{\mathrm{b}}^{\mathrm{T}}(t_0,t) \tag{17}$$

式中:$\boldsymbol{\phi}_{\mathrm{n}}(t_0,t)$和$\boldsymbol{\phi}_{\mathrm{b}}(t_0,t)$分别为n系和b系从 t_0到 t 时刻的姿态转移矩阵,可由主惯导输出信息及子惯导补偿后的陀螺测量数据计算得到,为已知量。将式(17)代入式(11),进一步整理可得

$$\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}(t_0)\boldsymbol{\phi}_{\mathrm{b}}^{\mathrm{T}}(t_0,t)(\boldsymbol{f}^{\mathrm{b}}-\boldsymbol{b}_{\mathrm{a}})=\boldsymbol{\phi}_{\mathrm{n}}^{\mathrm{T}}(t_0,t)[\dot{\boldsymbol{v}}_{\mathrm{m}}^{\mathrm{n}}+(2\boldsymbol{\omega}_{\mathrm{ie}}^{\mathrm{n}}+\boldsymbol{\omega}_{\mathrm{en}}^{\mathrm{n}})\times\boldsymbol{v}_{\mathrm{m}}^{\mathrm{n}}-\boldsymbol{g}^{\mathrm{n}}] \tag{18}$$

式中:除 $\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}(t_0)$外,各项皆为已知。注意到$\boldsymbol{\phi}_{\mathrm{b}}^{\mathrm{T}}(t_0,t)(\boldsymbol{f}^{\mathrm{b}}-\boldsymbol{b}_{\mathrm{a}})$为子惯导真实比力在其初始时刻体坐标系下的投影,记

$$\boldsymbol{f}^{\mathrm{b}(0)}(t)=\boldsymbol{\phi}_{\mathrm{b}}^{\mathrm{T}}(t_0,t)(\boldsymbol{f}^{\mathrm{b}}-\boldsymbol{b}_{\mathrm{a}}) \tag{19}$$

因此,若存在$\boldsymbol{f}^{\mathrm{b}(0)}(t_3)$和$\boldsymbol{f}^{\mathrm{b}(0)}(t_4)$且 $t_3\neq t_4$,使这两个时刻的$\boldsymbol{f}^{\mathrm{b}(0)}(t)$线性不相关,则 $\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}(t_0)$可被唯一地确定[20-21],进而任意时刻的姿态都是已知的,即姿态是可观测的。

4) 安装误差角的可观测性

除子惯导惯性器件误差及杆臂影响外,安装误差同样是导致主、子惯导输出不一致的因素之一。定义 $\boldsymbol{\psi}_{\mathrm{a}}$ 为弹体的安装误差角,即

$$\boldsymbol{C}_{\mathrm{b}}^{\mathrm{a}}=\boldsymbol{I}+[\boldsymbol{\psi}_{\mathrm{a}}\times] \tag{20}$$

已知载体姿态矩阵 $\boldsymbol{C}_{\mathrm{a}}^{\mathrm{n}}$ 可由主惯导精确提供,故 $\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}$ 已知后,安装误差矩阵 $\boldsymbol{C}_{\mathrm{b}}^{\mathrm{a}}=(\boldsymbol{C}_{\mathrm{a}}^{\mathrm{n}})^{\mathrm{T}}\cdot\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}$即为已知。由定义式(20),可解得

$$[\boldsymbol{\psi}_{\mathrm{a}}\times]=(\boldsymbol{C}_{\mathrm{a}}^{\mathrm{n}})^{\mathrm{T}}\cdot\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}-\boldsymbol{I} \tag{21}$$

由此可唯一确定 $\boldsymbol{\psi}_{\mathrm{a}}$,即安装误差角是可观测的。

5) 速度的可观测性

在直线运动假设下,对式(11)进行积分,得

$$\boldsymbol{v}_{\mathrm{m}}^{\mathrm{n}}=\boldsymbol{v}^{\mathrm{n}}=\boldsymbol{v}^{\mathrm{n}}(t_0)+\int_{t_0}^{t}[\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}(\boldsymbol{f}^{\mathrm{b}}-\boldsymbol{b}_{\mathrm{a}})-(2\boldsymbol{\omega}_{\mathrm{ie}}^{\mathrm{n}}+\boldsymbol{\omega}_{\mathrm{en}}^{\mathrm{n}})\times\boldsymbol{v}_{\mathrm{m}}^{\mathrm{n}}+\boldsymbol{g}^{\mathrm{n}}]\mathrm{d}t \tag{22}$$

由式(22)可解得速度初值

$$\boldsymbol{v}^{\mathrm{n}}(t_0)=\boldsymbol{v}_{\mathrm{m}}^{\mathrm{n}}-\int_{t_0}^{t}[\boldsymbol{C}_{\mathrm{b}}^{\mathrm{n}}(\boldsymbol{f}^{\mathrm{b}}-\boldsymbol{b}_{\mathrm{a}})-(2\boldsymbol{\omega}_{\mathrm{ie}}^{\mathrm{n}}+\boldsymbol{\omega}_{\mathrm{en}}^{\mathrm{n}})\times\boldsymbol{v}_{\mathrm{m}}^{\mathrm{n}}+\boldsymbol{g}^{\mathrm{n}}]\mathrm{d}t \tag{23}$$

则任意时刻速度都可确定,速度可观。

6) 杆臂的可观测性

将式(8)改写为

$$[\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}\times]\boldsymbol{l}^{\mathrm{b}}=\boldsymbol{C}_{\mathrm{n}}^{\mathrm{b}}(\boldsymbol{v}_{\mathrm{m}}^{\mathrm{n}}-\boldsymbol{v}^{\mathrm{n}})=\delta\boldsymbol{v}^{\mathrm{b}} \tag{24}$$

式中:除 $\boldsymbol{l}^{\mathrm{b}}$ 外各项均为已知。注意到矩阵$[\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}\times]$的秩为2,因此若存在时刻 $t_5\neq t_6$,满足 $\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}(t_5)$与$\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}(t_6)$线性不相关,则有

$$\begin{bmatrix}\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}(t_5)\times\\ \boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}(t_6)\times\end{bmatrix}\boldsymbol{l}^{\mathrm{b}}=\begin{bmatrix}\delta\boldsymbol{v}^{\mathrm{b}}(t_5)\\ \delta\boldsymbol{v}^{\mathrm{b}}(t_6)\end{bmatrix} \tag{25}$$

由于

$$\mathrm{rank}\left(\begin{bmatrix}\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}(t_5)\times\\ \boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}(t_6)\times\end{bmatrix}\right)=3 \tag{26}$$

$\boldsymbol{l}^{\mathrm{b}}$ 有唯一解。

因此,需载体作曲线运动,且存在两个不同的时刻满足 $\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}$线性不相关,则可满足杆臂可观测。

综上所述,对于惯性导航动基座传递对准系统,有如下结论。

结论1 载体作直线运动,姿态不发生变化,即 $\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}=0$,$\dot{\boldsymbol{C}}_{\mathrm{b}}^{\mathrm{n}}=0$,且载体加速度为0,则加速度计的零偏是可观测的。

结论2 载体作直线运动,姿态不发生变化,即 $\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}=0$,$\dot{\boldsymbol{C}}_{\mathrm{b}}^{\mathrm{n}}=0$,且存在两个时刻 $t_1\neq t_2$,满足比力导数$\dot{\boldsymbol{f}}^{\mathrm{b}}(t_1)$和$\dot{\boldsymbol{f}}^{\mathrm{b}}(t_2)$线性不相关,则陀螺的常值漂移是可观测的。

结论3 载体作直线运动,姿态不发生变化,即 $\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}=0$,$\dot{\boldsymbol{C}}_{\mathrm{b}}^{\mathrm{n}}=0$,且存在两个时刻 $t_3\neq t_4$,满足$\boldsymbol{f}^{\mathrm{b}(0)}(t_3)$和$\boldsymbol{f}^{\mathrm{b}(0)}(t_4)$线性不相关,则子惯导姿态、速度和安装误差角是可观测的。

结论4 载体作曲线运动,姿态发生变化,且存在两个时刻 $t_5\neq t_6$,满足 $\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}(t_5)$与 $\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}(t_6)$线性不相关,则杆臂矢量是可观测的。

北航学报 赠阅

3　仿真验证

为了对可观测性分析的结论进行验证,本节用 MATLAB 进行了数值仿真。以导航参数误差为状态变量,同时将惯性器件(加速度计及陀螺仪)测量误差的常值项和子惯导的安装误差增广为状态变量,构成传递对准的状态方程组,采用“速度 + 姿态”的匹配模式[2],利用扩展 Kalman 滤波器(EKF)对系统的状态变量进行估计,并通过滤波效果判断各状态变量的可观测性。

3.1　仿真参数设计

仿真的基本参数设置为:惯导系统的数据更新频率为 100 Hz;杆臂向量设置为 $\boldsymbol{l}^{\mathrm{b}} = [1 \quad 4 \quad 2]^{\mathrm{T}}$ (m),仿真时间为 60 s。

子惯导系统的基本参数设置为:安装误差为 $\boldsymbol{\psi}_{\mathrm{a}} = [1 \quad 3 \quad 2]^{\mathrm{T}}$ (°),加速度计零偏为 $\boldsymbol{b}_{\mathrm{a}} = [0.01 \quad 0.01 \quad 0.01]^{\mathrm{T}}$ (m/s^2),陀螺漂移为 $\boldsymbol{b}_{\mathrm{g}} = [0.005 \quad 0.005 \quad 0.005]^{\mathrm{T}}$((°)/s)。为达到较好的仿真效果,惯性器件的测量噪声设置为 0。此外,在对准开始时,子惯导系统状态量的初值由主惯导赋予。

滤波器误差状态的初值均取为 0,滤波周期为 0.01 s,其余参数根据系统进行设置,采用 4 阶龙格-库塔积分算法。

3.2　载体机动方案规划

由第 2 节推导得出的结论可知,系统各状态变量可观测性的确定具有一定的先后顺序:①加速度计零偏;②陀螺漂移;③子惯导姿态;④子惯导的速度和安装误差角;⑤杆臂。因此,载体的机动方案也应依据前述的 4 个条件分段设计。

由结论 1,载体静止或作匀速直线运动,即可满足加速度计零偏可观测。

根据结论 2,若载体作直线运动,则比力的变化将始终沿着载体前进的方向,无法满足线性不相关的要求。因此,在直线行驶的末端,载体的运动轨迹必须发生变化,依靠转弯瞬间的侧向加速度来实现不同时刻比力导数 $\dot{\boldsymbol{f}}^{\mathrm{b}}(t)$ 的线性不相关。但这样的机动只能使结论 2 中的条件在很短的时间段内得到满足,难以达到满意的估计效果。为得到较好的仿真结果,以验证理论的正确性,不妨假设载体在变加速直线运动的同时进行上升、下降或侧向平移等机动,且过程中保持姿态不发生变化,则载体在机动前后的比力导数是线性不相关的,即满足结论 2 中陀螺漂移可观测的要求。

在结论 3 中,只要求存在两个不同的时刻使 $\boldsymbol{f}^{\mathrm{b}(0)}(t)$ 线性不相关。此处需要注意的是,比力矢量等于惯性加速度与重力加速度之差,因此只要 $|\boldsymbol{a}(t)|$ 不恒定,即载体作变加速直线运动,便可满足要求,使子惯导的姿态、速度及安装误差角可观测。由此可见,载体第 3 段的机动方案包含于第 2 段中,因此无需再额外进行特定的机动。

依据结论 4 的要求,载体作曲线运动,若行进过程中存在俯仰、偏航或滚转 3 种姿态运动中的任意两种或全部,即可满足不同时刻 $\boldsymbol{\omega}_{\mathrm{nb}}^{\mathrm{b}}(t)$ 线性不相关的要求,使杆臂可观测。

由上述分析,对载体的机动方案进行如下设计。

机动段 1:0 ~ 10 s,载体静止不动。

机动段 2:10 ~ 40 s,载体沿纵轴作变加速直线运动,25 ~ 40 s,载体作变加速的升降运动。

机动段 3:40 ~ 60 s,载体同时进行滚转运动和俯仰运动。

载体比力和角速度的具体变化情况分别如图 1和图 2 所示。

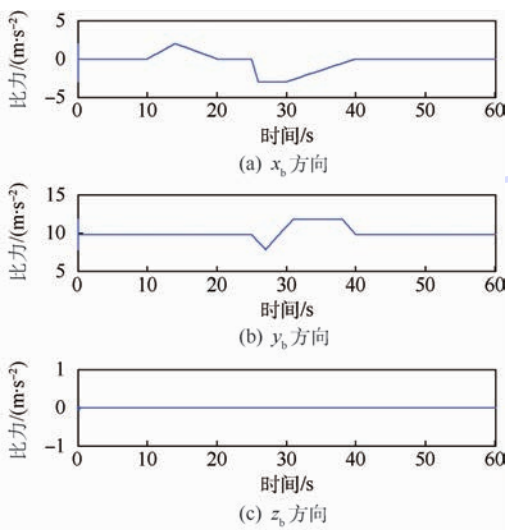


图 1　载体的比力

Fig. 1　Specific force of vehicle

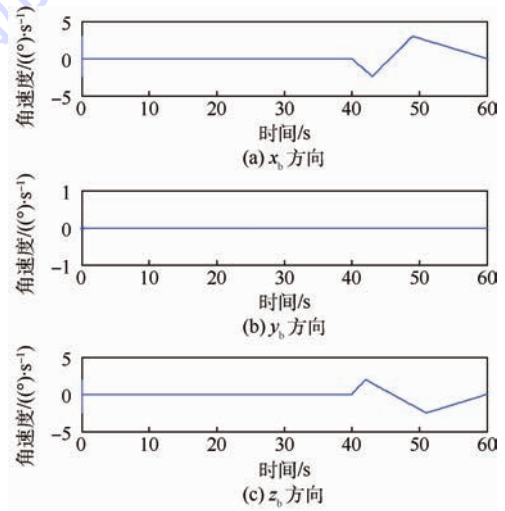


图 2　载体的角速度

Fig. 2　Angular velocity of vehicle

北航学报 赠阅

3.3 仿真结果

图 3 给出了子惯导加速度计零偏的估值曲线。在 0 ~ 10 s 之间，载体保持静止，满足结论 1 中的条件，加速度计零偏自仿真开始即迅速收敛，趋于真值。

图 4 给出了子惯导陀螺漂移的估值曲线。由图 1 可见，在 25 s 之前，比力仅沿 x_b 轴方向变化，不能唯一确定陀螺漂移，z_b 轴方向上的陀螺漂移始终未收敛至真值。而在 25 s 之后，y_b 轴方向上的比力也开始发生变化，满足 $\dot{\boldsymbol{f}}^b$ 线性不相关的条件，根据结论 2，此时陀螺完全可观。因此，曲线开始向真值收敛。

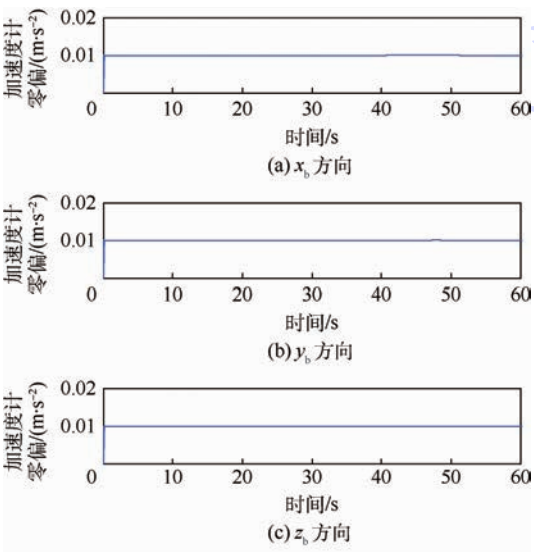


图 3 加速度计零偏估计结果

Fig. 3 Estimation results of accelerometer bias

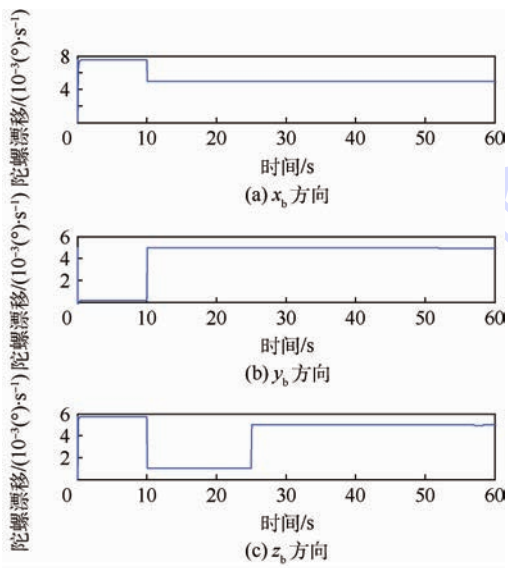


图 4 陀螺漂移估计结果

Fig. 4 Estimation results of gyro drift

图 5 给出了子惯导安装误差角的估值曲线。由于使安装误差角可观测的载体机动条件同陀螺漂移相同，同样，其估值曲线在 25 s 后开始向真值收敛，仿真结果与结论 3 一致。

图 6 给出了杆臂的估值曲线。对比图 2 可以看出，在 40 s 前，载体未进行任何姿态运动，杆臂在各方向上的估值基本保持在初值附近。而在 40 s 后，载体在滚转和俯仰方向的姿态同时发生变化，满足 $\boldsymbol{\omega}_{nb}^b$ 线性不相关的条件，此时杆臂可观，估值曲线开始向真值收敛，仿真结果与结论 4 相符。

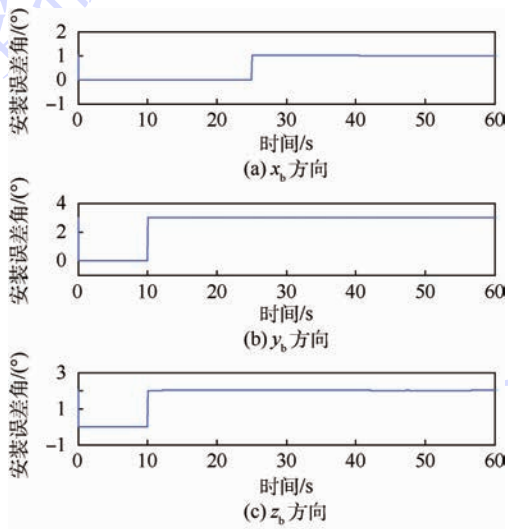


图 5 安装误差角估计结果

Fig. 5 Estimation results of fix error angle

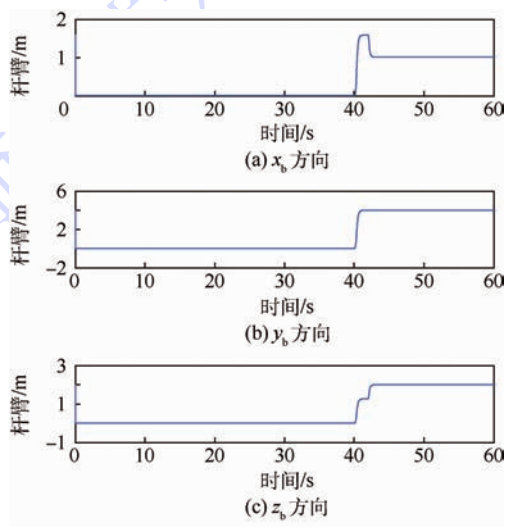


图 6 杆臂估计结果

Fig. 6 Estimation results of lever arm

上述仿真结果与第 2 节中的理论分析一致，充分验证了理论分析的正确性，也直观地反映了载体机动状态对系统状态可观测性的影响。

4　结　论

本文针对惯性导航系统传递对准过程中的可观测性问题，提出了从基本定义出发的非线性可观测性分析方法，解决了传统方法存在的模型误差不可避免、计算量较大、物理意义不明确等问题。

1）根据理论分析的结果对载体的机动策略进行了规划，通过 EKF 仿真验证了理论分析的正确性，分析过程直观、简洁，且结论更加准确、清晰。

2）系统可观测性的条件较易满足，只需载体先后进行不同方向上的变加速直线运动及姿态运动即可，且变加速直线运动可以增强系统惯性器件零偏、安装误差角及姿态的可观测性，而姿态运动则有助于对杆臂矢量的估计。

3）全面、深刻地揭示了系统状态估计同载体机动之间的联系，为载体运动轨迹的规划和滤波器的设计提供了理论指导，同时也为惯性导航系统高精度传递对准奠定了基础。

参考文献（References）

[1] KAIN J E, CLOUTIER J R. Rapid transfer alignment for tactical weapon application: AIAA-1989-3581 [R]. Reston: AIAA, 1989: 1290-1300.

[2] 高伟，奔粤阳，李倩. 捷联惯性导航系统初始对准技术[M]. 北京：国防工业出版社，2014：124.
GAO W, BEN Y Y, LI Q. Initial alignment for strapdown inertial navigation system [M]. Beijing: National Defense Industry Press, 2014: 124 (in Chinese).

[3] 韩英宏，陈万春. 新型快速传递对准方法[J]. 北京航空航天大学学报，2012，38(2)：149-152.
HAN Y H, CHEN W C. New method for rapid transfer alignment[J]. Journal of Beijing University of Aeronautics and Astronautics, 2012, 38(2): 149-152 (in Chinses).

[4] 解春明，赵剡，王纪南. 传递对准中机翼弹性变形建模与滤波处理[J]. 北京航空航天大学学报，2010，36(8)：931-935.
XIE C M, ZHAO Y, WANG J N. Modeling of wing flexure and filtering solution in transfer alignment [J]. Journal of Beijing University of Aeronautics and Astronautics, 2010, 36(8): 931-935 (in Chinses).

[5] 高伟，张亚，孙骞，等. 传递对准中杆臂效应的误差分析与补偿[J]. 仪器仪表学报，2013，34(3)：559-565.
GAO W, ZHANG Y, SUN Q, et al. Error analysis and compensation for lever-arm effect in transfer alignment [J]. Chinese Journal of Scientific Instrument, 2013, 34(3): 559-565 (in Chinese).

[6] LYOU J, LIM Y C. Transfer alignment considering measurement time delay and ship body flexure [J]. Journal of Mechanical Science and Technology, 2009, 23: 195-203.

[7] 刘红光，陈志刚，陈刚. 基准信息时间延迟对速度匹配传递对准性能的影响分析[J]. 中国惯性技术学报，2012，20(5)：544-551.
LIU H G, CHEN Z G, CHEN G. Time-delay's effect on velocity matching transfer alignment [J]. Journal of Chinese Inertial Technology, 2012, 20(5): 544-551 (in Chinese).

[8] 陈雨，赵剡，李群生，等. 快速传递对准中主惯导信息滞后补偿方法[J]. 中国惯性技术学报，2013，21(5)：576-580.
CHEN Y, ZHAO Y, LI Q S, et al. Compensation method for master inertial navigation system information delay in rapid transfer alignment[J]. Journal of Chinese Inertial Technology, 2013, 21(5): 576-580 (in Chinese).

[9] 江云，鲁浩，位晓峰. 弹载捷联惯导系统空中传递对准精度评估[J]. 弹箭与制导学报，2012，32(1)：21-26.
JIANG Y, LU H, WEI X F. Accuracy estimation for missile-borne strapdown inertial navigation systems during in-flight transfer alignment [J]. Journal of Projectiles, Rockets, Missiles and Guidance, 2012, 32(1): 21-26 (in Chinese).

[10] 解春明，赵剡，邓俊云. 一种改进的自适应平方根传递对准滤波算法[J]. 系统工程与电子技术，2011，33(3)：622-626.
XIE C M, ZHAO Y, DENG J Y. Improved adaptive square root filtering algorithm for transfer alignment[J]. Systems Engineering and Electronics, 2011, 33(3): 622-626 (in Chinese).

[11] HAM F M, BAR-ITZHACK I Y. Observability, eigenvalues, and Kalman filtering [J]. IEEE Transaction on Aerospace and Electronic Systems, 1983, 19(2): 269-273.

[12] KRELNDLER E, SARACHIK B E. On the concepts of controllability and observability of linear systems [J]. IEEE Transactions on Automatic Control, 1964, AC-9(2): 129-136.

[13] 房建成，周锐，祝世平. 捷联惯导系统动基座对准的可观测性分析[J]. 北京航空航天大学学报，1999，25(6)：714-719.
FANG J C, ZHOU R, ZHU S P. Observability analysis of strapdown inertial navigation system on moving base [J]. Journal of Beijing University of Aeronautics and Astronautics, 1999, 25(6): 714-719 (in Chinses).

[14] 秦峰，战兴群，湛雷，等. 机载武器传递对准可观测性分析[J]. 北京航空航天大学学报，2012，38(9)：1134-1138.
QIN F, ZHAN X Q, ZHAN L, et al. Observability analysis of transfer alignment of airborne weapons [J]. Journal of Beijing University of Aeronautics and Astronautics, 2012, 38(9): 1134-1138 (in Chinese).

[15] GOSHEN-MESKIN D, BAR-ITZHACK I Y. Observability analysis of piece-wise constant systems-part I: Theory [J]. IEEE Transaction on Aerospace and Electronic Systems, 1992, 28(4): 1056-1067.

[16] GOSHEN-MESKIN D, BAR-ITZHACK I Y. Observability analysis of piece-wise constant systems-part II: Application to inertial navigation in-flight alignment [J]. IEEE Transaction on Aerospace and Electronic Systems, 1992, 28(4): 1068-1075.

[17] 汤勇刚. 载波相位时间差分/捷联惯导组合导航方法研究[D]. 长沙：国防科技大学，2007：9-11.
TANG Y G. Research on time-differenced carrier phase/SINS integrated navigation approach[D]. Changsha: National University of Defense Technology, 2007: 9-11 (in Chinese).

[18] WEI M, SCHWARZ K P. A strapdown inertial algorithm using an earth-fixed cartesian frame [J]. Journal of Institute of Navi-

北航学报 赠阅

gation, 1990, 37(2): 153-167.
[19] CHEN C T. Linear system theory and design[M]. 3rd ed. New York: Oxford University Press, 1999.
[20] BLACK H D. A passive system for determining the attitude of a satellite[J]. AIAA Journal of Guidance, Control, and Dynamics, 1964, 2(7): 1350-1351.
[21] SHUSTER M D, OH S D. Three-axis attitude determination from vector observations[J]. AIAA Journal of Guidance, Control, and Dynamics, 1981, 4(1): 70-77.

作者简介:

黄帅 女,硕士研究生。主要研究方向:飞行器导航制导与控制。
Tel.: 0731-87001191
E-mail: ase_huang@hotmail.com

蔡洪 男,博士,教授,博士生导师。主要研究方向:飞行器导航制导与控制。
Tel.: 0731-87001191
E-mail: hcai@nudt.edu.cn

丁智坚 男,博士研究生。主要研究方向:惯性导航技术。
Tel.: 0731-87001191
E-mail: dzjqe@126.com

Observability analysis for transfer alignment of inertial navigation system on moving base

HUANG Shuai, CAI Hong*, DING Zhijian

(College of Aerospace Science and Engineering, National University of Defense Technology, Changsha 410073, China)

Abstract: A novel approach that starts from the basic observability definition is used to investigate the observability of the transfer alignment system on moving base. The original nonlinear state equations are analyzed directly to eliminate errors of the model. Sufficient conditions to make alignment fully observable are analytically derived. And maneuver scheme of vehicle is designed according to it. The results show that translatory motions with varying acceleration can enhance the observability of sensor bias, fix error angle and attitude, whereas angular motions are in favor of the observability of the lever arm. Different from traditional linearization-based observability studies, this kind of analysis approach is straightforward and comprehensive, with no need of nonanalytical support from numerical simulations. Besides, it has explicit physical meanings, revealing the inherent connecting link between the observability of transfer alignment system and vehicle maneuvers, and providing new insights into the analysis and design of nonlinear systems. Finally, covariance simulations with an extended Kalman filter (EKF) are carried out to estimate the states, and the numerical results accord with the theoretical analysis.

Key words: observability analysis; transfer alignment; maneuver scheme; moving base; inertial navigation system

Received: 2015-11-20; **Accepted**: 2016-04-15; **Published online**: 2016-04-19 17:33
URL: www.cnki.net/kcms/detail/11.2625.V.20160419.1733.001.html
* **Corresponding author**. Tel.: 0731-87001191 E-mail: hcai@nudt.edu.cn

北京航空航天大学学报

北航学报
赠 阅

《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,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,作者有权改投它刊。在此之前,请勿一稿多投,否则一切后果自负。

3.3 来稿一经刊登,即赠送单行本。

3.4 来稿一经作者签字并在本刊刊出,即表明所有作者都已经认可其版权转至本刊编辑部。本刊在与国内外文献数据库或检索系统进行交流及合作时,不再征询作者意见。

邮寄地址:100083 北京市海淀区学院路 37 号 北京航空航天大学学报编辑部
办公地点:北京航空航天大学办公楼 405,407,409 房间
电 话:(010)82315594,82338922,82314839,82315426
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

北航学报
赠 阅

《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

（以下按姓氏笔划为序）

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 42 卷 第 11 期　2016 年 11 月

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS (JBUAA)
(Monthly, Started in 1956)
Vol.42 No.11　November 2016

主管单位　中华人民共和国工业和信息化部
主办单位　北京航空航天大学
主　　编　赵沁平
编辑出版　《北京航空航天大学学报》编辑部
邮　　编　100083
地　　址　北京市海淀区学院路 37 号
印　　刷　北京科信印刷有限公司
发　　行　北航文化传媒集团
发行范围　国内外发行
联系电话　(010)　82315594　82338922　82314839
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　100083,　P.　R.　China)
Chief Editor　Zhao Qinping
Edited and Published by　Editorial Board of JBUAA
Printed by　Beijing Kexin Printing Co., Ltd.
Distributed by　BUAA Culture Media Group Limited
Telephone　(010)　82315594　82338922　82314839
E-mail　jbuaa@buaa.edu.cn
http://bhxb.buaa.edu.cn

刊　　号　ISSN 1001-5965　CN 11-2625/V
国内定价　20.00 元／期

ISSN 1001-5965
9 771001 596168 11>