



北航学报
赠阅

ISSN 1001-5965
CODEN BHHDE8

北京航空航天大学 学报

JOURNAL OF BEIJING UNIVERSITY OF
AERONAUTICS AND ASTRONAUTICS



2016-07

Vol.42 No.7

目 次

- 二维层状材料过渡金属硫化物 张千帆, 高磊, 田洪镇, 徐忠菲, 王亚鹏 (1311)
- 航空高动态网络链路感知 OLSR 路由算法 马林华, 张嵩, 茹乐, 徐扬, 田雨, 于云龙 (1326)
- 基于网络编码的无线 NoC 多播映射 陈亦欧, 谭芳, 凌翔 (1335)
- 高速切削高温合金 GH4169 数值模拟与实验 范孝良, 吴学华, 王进峰, 康文利 (1344)
- 动态总方差设计及其快速算法 汪立新, 李灿, 姜周, 朱战辉, 田颖 (1352)
- 基于 CCR 带并行腔双集束型设备调度方法 周炳海, 黎明 (1361)
- 电动伺服舵系统动力学建模及颤振分析 张仁嘉, 吴志刚, 杨超 (1368)
- 磁悬浮反作用飞轮无刷直流电机转矩脉动抑制策略 汤继强, 王营许, 周新秀 (1377)
- 空基伪卫星组网部署的航路规划算法 曾凌川, 李大朋, 曲艺, 任爱爱, 巩应奎 (1388)
- EHA 反馈线性化最优滑模面双模糊滑模控制 张振, 李海军, 诸德放 (1398)
- 基于 EEG 的脑力疲劳特征研究 范晓丽, 牛海燕, 周前祥, 柳忠起 (1406)
- 考虑系统不确定性的高超声速飞行器容错控制 彭程, 王新民, 谢蓉, 穆凌霞 (1414)
- 动态火力接入下要地防空作战目标威胁评估 刘敬蜀, 姜文志, 雷宇曜, 代进进 (1422)
- 可重复使用飞行器进场着陆拉平纵向控制 郝现伟, 王勇, 杨业, 郭涛, 张代兵 (1432)
- 基于母线电流的无刷直流电机换相位置优化策略 张前, 冯明 (1441)
- 不完备信息条件下的并发故障诊断方法 孙伟超, 许爱强, 李文海 (1449)
- 泡沫子弹撞击刚性靶板的理论模型 刘虎, 刘华, 杨嘉陵 (1461)
- 横向加强构件作用下的开口薄壁梁等效建模方法 邓昊, 程伟 (1469)
- 利于减少配平损失的太阳能飞机构型设计 阙建锋, 王维军, 吴宇 (1479)
- 图模型与学习算法结合的贝叶斯网络自动建模 沈琳, 于劲松, 唐荻音, 刘浩 (1486)
- 多约束下备件多层多级库存优化模型 蔡芝明, 金家善, 陈砚桥 (1494)
- 加速度场中环路热管温度波动现象分析 薛好, 谢永奇, 戴华, 余建祖, 马伟 (1502)
- 含区间参数不确定结构的损伤识别方法 杨君坦, 邱志平, 李琦 (1509)
- 基于诱导航线的多无人机编队飞行控制方法 吴俊成, 周锐, 董卓宁, 车军 (1518)
- 基于在线轨迹迭代的自适应再入制导 赵頔, 沈作军 (1526)
- 基于高阶奇异值分解的 LPV 鲁棒控制器设计 孙斌, 杨凌宇, 张晶 (1536)
- 基于拟态电势能的飞行器航路规划方法 何仁珂, 魏瑞轩, 张启瑞, 许卓凡 (1543)
- 重型直升机-吊挂耦合系统闭环飞行品质分析 朱笑宇, 曹义华, 曹龙 (1550)

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Vol. 42 No. 7 (Sum 281) July 2016

CONTENTS

- Two-dimensional layered materials of transition metal dichalcogenides
..... ZHANG Qianfan, GAO Lei, TIAN Hongzhen, XU Zhongfei, WANG Yapeng (1311)
- Link awareness-based OLSR routing algorithm for airborne highly dynamic networks
..... MA Linhua, ZHANG Song, RU Le, XU Yang, TIAN Yu, YU Yunlong (1326)
- Multicast wireless NoC mapping based on network coding
..... CHEN Yiou, TAN Fang, LING Xiang (1335)
- Numerical simulation and experiment in high-speed cutting superalloy GH4169
..... FAN Xiaoliang, WU Xuehua, WANG Jinfeng, KANG Wenli (1344)
- Design of dynamic total variance and its fast algorithm
..... WANG Lixin, LI Can, JIANG Zhou, ZHU Zhanhui, TIAN Ying (1352)
- Scheduling method for double-cluster tools with parallel chambers based on capacity constraint resource
..... ZHOU Binghai, LI Ming (1361)
- Dynamic modeling and flutter analysis of a fin-actuator system
..... ZHANG Renjia, WU Zhigang, YANG Chao (1368)
- Torque ripple reduction strategy for magnet suspended reactive flywheel brushless DC motor
..... TANG Jiqiang, WANG Yingxu, ZHOU Xinxiu (1377)
- Route planning algorithm for configuring airborne pseudolites
..... ZENG Lingchuan, LI Dapeng, QU Yi, REN Aiai, GONG Yingkui (1388)
- Double fuzzy sliding mode control for EHA based on feedback linearization optimal sliding surface
..... ZHANG Zhen, LI Haijun, ZHU Defang (1398)
- Mental fatigue characteristics based on EEG analysis
..... FAN Xiaoli, NIU Haiyan, ZHOU Qianxiang, LIU Zhongqi (1406)
- Fault-tolerant control for hypersonic vehicle with system uncertainty
..... PENG Cheng, WANG Xinmin, XIE Rong, MU Lingxia (1414)
- Threat evaluation of air-targets for key positions air-defense under dynamic fire access
..... LIU Jingshu, JIANG Wenzhi, LEI Yuyao, DAI Jinjin (1422)
- Flare longitudinal control for approach and landing of reusable launch vehicle
..... HAO Xianwei, WANG Yong, YANG Ye, GUO Tao, ZHANG Daibing (1432)
- BLDC motor commutation position optimization strategy based on bus current
..... ZHANG Qian, FENG Ming (1441)
- Diagnosis method of simultaneous fault with incomplete information
..... SUN Weichao, XU Aiqiang, LI Wenhai (1449)
- Theoretical model for a porous projectile striking on flat rigid anvil
..... LIU Hu, LIU Hua, YANG Jialing (1461)
- Equivalent modeling method of open thin-walled beam under action of transverse stiffening member
..... DENG Hao, CHENG Wei (1469)
- Design of solar-powered aircraft configuration for reducing trim loss
..... QUE Jianfeng, WANG Weijun, WU Yu (1479)
- Automatic learning of Bayesian network structure using graph model and learning algorithm
..... SHEN Lin, YU Jinsong, TANG Diyin, LIU Hao (1486)
- Optimal inventory modeling of spare parts multi-indenture multi-echelon under multi-constraints
..... CAI Zhiming, JIN Jiashan, CHEN Yanqiao (1494)
- Investigation of temperature fluctuation in loop heat pipe under acceleration fields
..... XUE Hao, XIE Yongqi, DAI Hua, YU Jianzu, MA Wei (1502)
- Damage detection of structures with interval uncertain parameters
..... YANG Juntan, QIU Zhiping, LI Qi (1509)
- Formation flight control method of multiple UAVs based on guidance route
..... WU Juncheng, ZHOU Rui, DONG Zhuoning, CHE Jun (1518)
- Adaptive reentry guidance based on on-board trajectory iterations
..... ZHAO Di, SHEN Zuojun (1526)
- Robust LPV control design based on HOSVD
..... SUN Bin, YANG Lingyu, ZHANG Jing (1536)
- Mimetism electric potential energy motion planning algorithm for aircraft
..... HE Renke, WEI Ruixuan, ZHANG Qirui, XU Zhuofan (1543)
- Heavy helicopter-slung-load coupling system flying qualities in closed-loop state
..... ZHU Xiaoyu, CAO Yihua, CAO Long (1550)

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2016.0092

二维层状材料过渡金属硫化物



张千帆*, 高磊, 田洪镇, 徐忠菲, 王亚鹏

(北京航空航天大学 材料科学与工程学院, 北京 100083)

摘 要: 过渡金属硫化物(TMDCs)纳米材料优异的性能有利于它在纳米电子学、光电子学和自旋电子学等领域广泛地应用。主要介绍了TMDCs材料最新的进展,概述了TMDCs的制备方法、性能以及纳米结构形态。纳米TMDCs层状材料边缘态对材料性能会产生重要的影响,着重介绍了低维纳米材料边缘态对性能(化学性能、电子性能和磁性能)的影响以及边缘态最新的研究进展,特别是TMDCs材料边缘态的磁性能为其在磁性和自旋领域潜在性的应用提供了可能,并且引起了广大研究者的兴趣,从实验和理论2个方面总结了该领域的研究成果,为二维层状TMDCs材料的研究提供了支持。

关键词: 过渡金属硫化物(TMDCs); 层状材料; 纳米结构; 边缘态; 磁性

中图分类号: TQ584; O464

文献标识码: A **文章编号:** 1001-5965(2016)07-1311-15

纳米材料由于具有较大的表面积/体积比和量子效应而具有不同于体材料的优异性能。石墨烯作为一种较成功的层状材料,其优异的电子结构、电导性和较大强度吸引了广大研究者的注意。但是石墨烯的零带隙这一缺陷严重限制了它在某些领域中的应用。近年来,由于过渡金属硫化物材料(如二硫化钼、二硒化钼、二硫化钨、二硒化钨、二硫化铌等)具有半导体或超导性质,且在纳米级光电子学等领域广泛地应用,因而引起了广大研究者的兴趣^[1-3]。过渡金属硫化物层状材料不仅可应用于电子学领域(晶体管、逻辑元件、传感器和柔性电子学)、能源技术(太阳能电池、氢存储和电池)、先进工程材料(纳米复合材料),而且还可广泛地应用于自旋电子学领域。

当材料的尺寸变成纳米级时,低维材料表面的缺陷如边缘、角落将对材料的性能起主导作用,低维的过渡金属硫化物层状材料由于其具有低维

性和量子尺寸效应而具有优异的电子和物理性能。由于过渡金属硫化物材料层与层之间是由弱的范德华力相连接的,因此可以通过实验手段从体材料中制备出低维的过渡金属硫化物材料。过渡金属硫化物体材料的性质是不活泼的,但当它们变成低维结构后显示出较活泼的化学反应性。这是因为其存在的边缘态具有活性位点,对材料的性能会产生显著的影响。

目前,也有许多文献报道了过渡金属硫化物层状材料的一些性能,但是目前仍没有文章系统全面地介绍过渡金属硫化物材料的制备方法、结构和性能。本文主要就过渡金属硫化物层状材料结构、制备及性能进行总结,综述了低维纳米过渡金属硫化物材料性能的特点,并提出边缘态的存在对材料性能的意义,介绍了边缘态磁性能的研究进展,本课题组也研究了二维过渡金属硫化物层状材料的磁性能。

收稿日期: 2016-01-26; 录用日期: 2016-03-04; 网络出版时间: 2016-03-10 15:31

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20160310.1531.001.html

基金项目: 国家自然科学基金(11404017); 博士点基金新教师类(20131102120001); 人力资源和社会保障部留学人员科技活动项目择优资助; 新世纪优秀人才支持计划(NCET-12-0033)

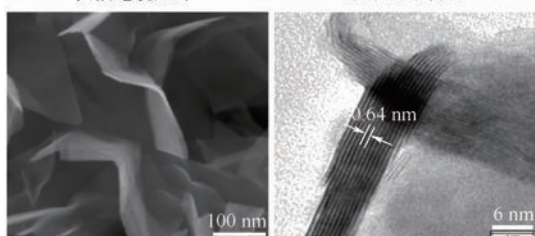
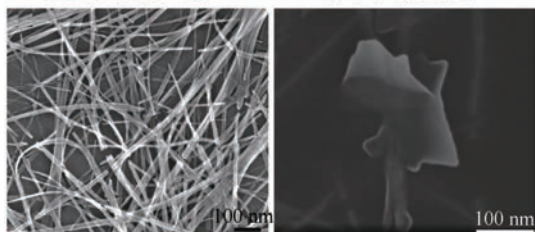
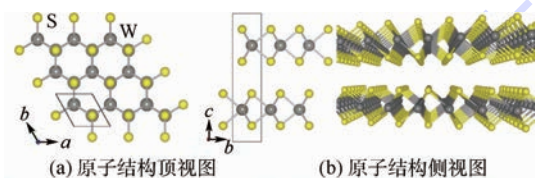
* 通讯作者: E-mail: qianfan@buaa.edu.cn

引用格式: 张千帆, 高磊, 田洪镇, 等. 二维层状材料过渡金属硫化物[J]. 北京航空航天大学学报, 2016, 42(7): 1311-1325.
ZHANG Q F, GAO L, TIAN H Z, et al. Two-dimensional layered materials of transition metal dichalcogenides[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1311-1325 (in Chinese).

1 过渡金属硫化物层状材料的制备与性能

本节主要介绍过渡金属硫化物层状材料的结构、形态、制备方法以及性能。

二硫化钨的结构如图 1^[4]所示。由图 1(a)和图 1(b)可见,过渡金属硫化物层状材料包含 3 个原子层,过渡金属层镶嵌在 2 个硫原子层中间。每一个过渡金属硫化物层都是六方晶格结构。相应地,通过堆叠单层的过渡金属硫化物材料构成体材料。层内是化学键相连接,层间是范德华力相连接。过渡金属硫化物通过硫族元素和过渡金属元素的组合可以形成 40 多种不同类型的材料。过渡金属硫化物层状材料的性能依赖于过渡金属原子 M 和硫族元素 $X(X = S, Se, Te)$ 。



a, b, c —3 个不同的方向。

图 1 二硫化钨的结构^[4]

Fig. 1 Structure of WS_2 ^[4]

过渡金属硫化物层状材料具有半导体性、半金属磁性^[5]和超导电性^[6],因而广泛地应用于光电子学、自旋电子学、催化剂^[7]、润滑剂^[8]、化学和生物传感器、超级电容器^[9]、太阳能电池和锂电池中。过渡金属硫化物层状材料可以衍生出不同形态的材料(纳米管、纳米条带、纳米团簇和纳米颗粒),不同形态具有不同的性能,因而过渡金属硫化物层状材料具有广泛的应用领域。其中,二硫化钨和二硒化钨纳米条带材料具有较高的电

荷载流子迁移率^[10-11]和较高的开/关比率^[12-13]。目前,基于二硫化钨和二硒化钨的半导体器件已经成功制备出来^[14-15]。较薄的二硫化钨载体运输对于化学吸附物非常敏感,这种特性使二硫化钨可以应用在化学传感器上^[16-17]。

1.1 制备与表征

1.1.1 制备方法

由于过渡金属硫化物材料层间是较弱的范德华力,可以沿着层间方向将层剥离。目前已经通过实验手段制备出单层的过渡金属硫化物层状材料,比如微机械剥离技术^[18]、大块样品的化学剥离法^[18]、物理气相沉积(PVD)^[19]、化学气相沉积(CVD)^[20],另外,利用水热合成方法^[21]在基板上制备出超薄的二硫化钨薄膜^[22],并使它们自由悬浮^[23]。除此之外,一维的二硫化钨纳米条带可以通过两步电化学/化学方法制备得到^[24]。

Novoselov 等^[25]证实了可以通过微机械剥离的方法制备出较薄的二硫化钨晶体,这是由于层与层之间存在较弱的相互作用和层间较强的相互作用。通过机械剥离法分离出来的二维过渡金属硫化物层状材料是没有点缺陷和扩展缺陷的,这为过渡金属硫化物层状材料大规模的应用提供了技术手段。Coleman 等^[26]提出了液相剥离的方法,这种方法能够制备出 3 ~ 12 nm 较薄的二硫化钨薄层,对应 5 ~ 20 个单分子层,但是单层的二硫化钨不能用此方法制备出。

将过渡金属硫化物材料分离成单层材料是一个关键的步骤,这关系到分离出来的材料是否具有优异的光学活性并且可否成功应用到器件中。Morrison 等^[27]报道了通过迫使水合作用可以使锂插入二硫化钨(Li_xMoS_2)中并且分离出单层的二硫化钨,形成一种稳定的胶体悬浮液。虽然 Li_xMoS_2 的剥离是一种比较简单的制备单层二硫化钨的方法,但是锂插入过程中伴随着相变,发生一些性能的改变。特别是,当锂插入二硫化钨,由于金属性的出现导致半导体性能的丧失。接着,可以克服这个缺点的化学剥离的方法^[18]被提出,锂插入能够改变被剥离材料的性能,但是温和的热处理过程能够促使半导体相的恢复。当热处理温度达到 300 °C 时,化学剥离的二硫化钨显示出显著的带隙光致发光性,半导体的性能能够被恢复。如图 1 所示有研究者直接通过水热合成法来制备单层的二硫化钨薄膜^[4,28]。另外,声波降解法也大规模地应用在制备过渡金属硫化物层状材料中^[26]。这些制备方法的成功应用使单层过渡金属硫化物材料的制备成为可能,因而可以扩

大过渡金属硫化物材料的应用领域。

1.1.2 表征方法

目前,实验中运用一些测量仪器来研究过渡金属硫化物层状材料的结构和性能。透射电子显微镜和扫描探针显微镜可以用来观察单层或者多层材料的形貌特征。Qiu 等^[29]利用高分辨透射显微镜对单层二硫化钼的结构进行观测,观察到由单晶二硫化钼剥离得到单层二硫化钼表面的空位缺陷。一些过渡金属硫化物材料比如单层的二硫化钼和二硫化钨是直接带隙的材料,具有荧光性,可以用荧光显微镜^[30]来观察。扫描隧道显微镜表征主要倾向于研究过渡金属硫化物材料上面吸附的原子。原子力显微镜已经成为测量过渡金属材料层状材料厚度的一种有效的测量仪器。在一些研究中,拉曼光谱可以用来研究过渡金属硫化物层状材料性能和层数之间的关系。在 1.2 节、3.3 节和 3.4 节内容的介绍中会涉及到利用测量仪器来研究过渡金属硫化物材料的性能,本节中就不过多介绍。

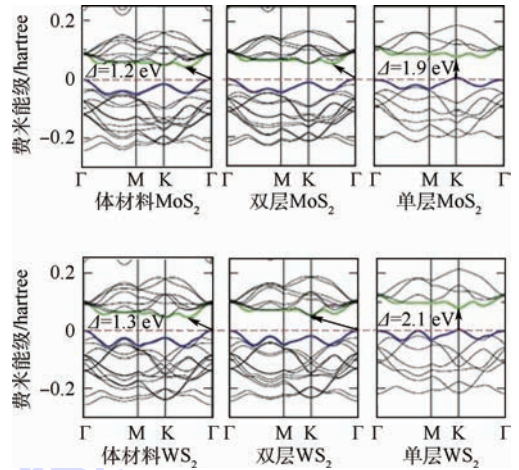
1.2 材料性能

1.2.1 电子性能

第一性原理计算的过渡金属硫化物材料中有绝缘体、半导体、半金属和金属^[31-35]。典型的例子是二硫化钼,它是一种半导体材料。体相中,它是一种间接带隙的材料,而它的单层材料却是一种直接带隙的材料^[34,36]。

过渡金属硫化物层状材料的能带结构依赖于层数,当体材料厚度减少为多层或者单层时,层状材料的电子性能发生显著的变化。由于过渡金属硫化物材料具有量子效应,因而带隙会随着层数的变化而发生变化^[34-37]。一些研究^[30,38]已经证实当二硫化钼材料厚度减少为单层时,间接带隙转变为直接带隙。Kuc 等^[34]研究了量子效应对单层和多层 MS_2 ($M = W, Nb, Re$) 电子结构的影响。研究发现 WS_2 和 MoS_2 相似,当层数减少时,显示出间接带隙(体材料)到直接带隙的转变。图 2^[34]是 MoS_2 和 WS_2 的能带结构,图中箭头由价带顶指向导带底。而 NbS_2 和 ReS_2 是金属性的,其电子性能与层数无关。单层 $MoSe_2$ 和 $MoTe_2$ 也显示出间接带隙到直接带隙的转变,带隙(Δ)大小分别是 1.44 eV 和 1.07 eV^[39] (1 hartree = 27.2114 eV)。

也有第一性原理研究发现,当材料层数减少时,会引起带隙能量的变化,当从体材料变成单层材料时, MoS_2 、 WS_2 、 $MoSe_2$ 、 WSe_2 、 $MoTe_2$ 和 WTe_2 带隙能量分别变化了 1.14、1.16、0.78、0.64、0.57



Γ, M, K —布里渊区的高对称性点。

图 2 MoS_2 和 WS_2 的能带结构^[34]

Fig. 2 Band structures of MoS_2 and WS_2 ^[34]

和 0.37 eV,说明过渡金属硫化物层状材料能带带隙是可调节的,这个发现使过渡金属硫化物层状材料可以更好地应用于光电子学器件中^[40]。

在过渡金属硫化物材料的电子应用领域,需要通过调整带隙的大小来增加半导体器件的载流子迁移率或者发光二极管器件中的发射效率。这些都可以通过外加电场、化学功能化、纳米图案成形或者外加应变等手段实现,而外加应变是一种改变带隙的有效的手段。第一性原理计算证实二维过渡金属硫化物层状材料的带隙可以通过拉伸、压缩和纯剪切应变来调整^[41-45]。理论计算表明,外加应力增大, MoX_2 和 WX_2 ($X = S, Se, Te$) 单层和双层的材料能够由直接带隙转化为间接带隙^[43,46]。Yun 等^[46]运用线性增广平面波法发现拉伸应变能够减少隙能,而压缩应变能够增大隙能。Shi 等^[45]运用第一性原理计算方法发现二硫化钼的部分能态密度(Density of States, DOS)图可以通过双轴拉伸来改变。如图 3^[45]所示,钼原子 d 轨道的能态密度图可以通过外加应变显著地改变。此研究通过外加应变调节单层二硫化钼的电子性能,为制备高性能的电子器件提供了思路。另一种手段就是通过外加电场来改变带隙,根据第一性原理计算^[47],在外加电场的作用下,带隙能够连续不断地降低到零(由半导体到金属的转变)。由于过渡金属硫化物层状材料是由表面构成的,吸附原子和表面缺陷也能够改变过渡金属硫化物材料的性能。通过对过渡金属硫化物层状材料的功能化,可以使过渡金属硫化物材料表面吸附原子或者产生表面缺陷,因而可以改善材料的性能且扩大其应用领域^[48-50]。Ataca 和 Ciraci^[48]研究了单层二硫化钼表面吸附原子的过程,

发现吸附不同的原子对单层二硫化钼电子性能会产生不同的影响。进一步地,He 等^[49]发现吸附 H、B、C、N 和 F 原子能够使单层的二硫化钼产生磁矩。Yue 等^[42]则发现二硫化钼中的硫原子被非金属原子(H、B、N 和 F)和过渡金属原子(V、Cr、Mn、Fe 和 Co)取代后能够产生磁矩。通过外加应力、电场及材料功能化等手段,可以使过渡金属硫化物材料的性能发生显著的变化,从而应用到特定的领域。

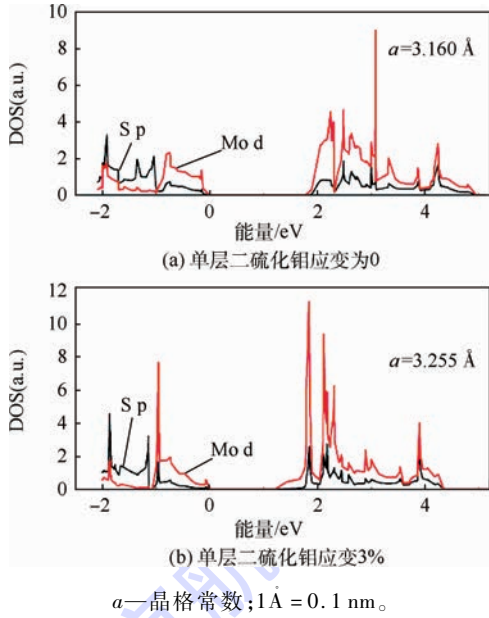


图3 钼原子 d 轨道和硫原子 p 轨道的部分能态密度^[45]
Fig.3 Projected density of states of d-orbitals for Mo-atoms and p-orbitals for S-atoms^[45]

1.2.2 谷电子学性能

基于电荷的输运和电子的自旋来制备出自旋电子和谷电子器件。“谷”状结构是指导带底和价带顶处的简并。过渡金属硫化物如二硫化钼由于具有自旋轨道耦合以及对称结构破缺,导致了 K 谷和 K' 谷对左旋和右旋圆偏振光的吸收具有选择性,因此在自旋电子学和谷电子学等领域具有广泛的应用。图 4^[51]表示价带顶和导带底。在二硫化钼的第一布里渊区内,不同自旋(+和-)会和非等值点 K_+ 、 K_- 发生耦合。

要制备出谷电子器件,关键是实现所谓的谷极化。迄今为止,通过产生特定的能谷来产生能谷极化是比较困难的。谷极化是通过圆偏振光激发,由于光学选择性,给定的圆偏振光只能激发出自旋向上或者自旋向下的光子。Zeng 等^[52]对 2 层的二硫化钼进行实验,由于存在反转对称,在布里渊区 K 和 K' 处发生简并,所以没有谷极化出现,但在单层的二硫化钼中发现了谷的旋光选择性。Feng 等^[51]从实验和理论上证实了单层二硫

化钼的谷选择圆偏振光吸收性质,这项研究表明单层二硫化钼的能带在六边形布里渊区的顶点附近拥有“谷”状结构(见图 4^[51]),而相邻点的谷并不等价,它们分别吸收左旋光和右旋光,其选择性近乎完美。这项研究首次发现了材料中谷的旋光选择性,单层二硫化钼实现了谷极化,说明过渡金属硫化物材料在自旋电子学和谷电子学领域具有较好的发展前景。二硫化钼没有空间反演中心,第一性原理计算表明,自旋轨道耦合会使价带发生劈裂^[53-54]。由于自旋-谷耦合,根据对称性,劈裂后的能带分为自旋向上和自旋向下,时间反演性决定了不同谷中的劈裂必然相反,价带顶的自旋劈裂决定了谷的旋光选择性依赖于自旋自由度^[55]。自旋自由度和谷自由度的耦合图如图 5^[55]所示。图 5 中实线代表沿着面外自旋向下的能带,虚线代表沿着面外自旋向上的能带。

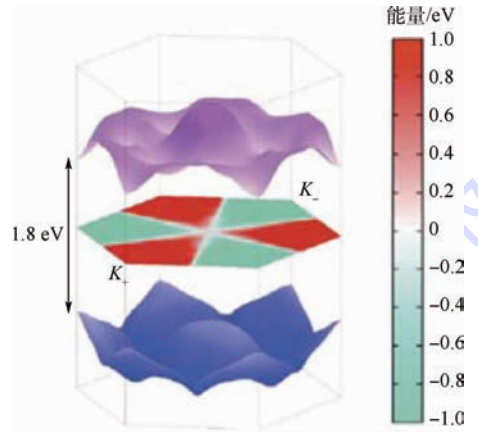
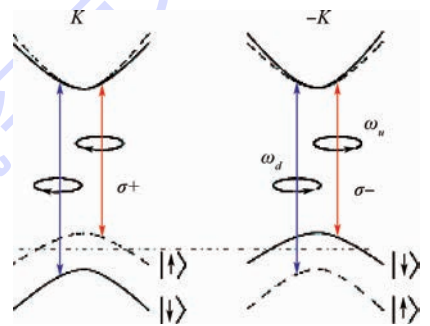


图4 价带顶和导带底^[51]

Fig.4 Top valence band and bottom conduction band^[51]



$K, -K$ —不同的谷; $\sigma+, \sigma-$ —左旋圆偏振光和右旋圆偏振光;
 ω_u, ω_d —从劈裂的价带顶到导带底的过渡频率。

图5 自旋自由度和谷自由度的耦合图^[55]

Fig.5 Schematics of coupled spin drgree of freedom and valley drgree of freedom^[55]

1.2.3 光学性能

过渡金属硫化物材料具有优异的光学性能,在光探测和光致发光等领域广泛地应用。研究发现体材料的二硫化钼光致发光性不明显,而单层的二硫化钼具有较强的光致发光性^[18],因而单层

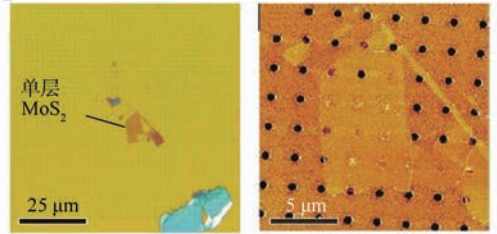
二硫化钨可以应用于太阳光电板、光电探测器和光电发射器中^[38]。过渡金属硫化物材料的光致发光性与层厚度有密切的联系。有研究^[30]发现随着过渡金属硫化物层状材料层数的降低,光致发光性增强,而单层的光致发光性是最强的。Zhao等^[56]研究了不同层数($n=1,2,\dots,5$)机械剥离的二硫化钨和二硒化钨的光吸收和光致发光性。图6^[56]中1L表示机械剥离成一层的二硫化钨,2L表示机械剥离成2层的二硫化钨,以此类推。二硫化钨在625和550 nm出现电子吸收峰(A和B),二硒化钨在760和600 nm处出现电子吸收峰(A和B),当层厚度降低时,所有的吸收峰逐渐蓝移。图6(c)和图6(d)表示机械剥离二硫化钨的荧光图像。从图6(c)和图6(d)中可以看出二硫化钨和二硒化钨从双层变为单层时,光致发光性增强。利用拉曼光谱研究单层二硫化钨的光致发光性,单层二硫化钨在1.9 eV和2.0 eV之间出现光致发光性峰,这与直接带隙转变产生的激发有关^[57]。对于二硫化钨,有研究发现在单层的三角形边缘处光致发光性增强,强度比内部原子区域增强了25~40倍,这个发现促进了单层过渡金属硫化物材料光电性质调节的研究^[28]。另外,过渡金属硫化物也广泛地应用在光探测器上,

而对二硒化钨进行光探测的研究,发现响应比二硫化钨有3个数量级的提高,显示出巨大的优异性^[58]。过渡金属硫化物优异的光学性质为其在光电子学领域的应用奠定了基础。

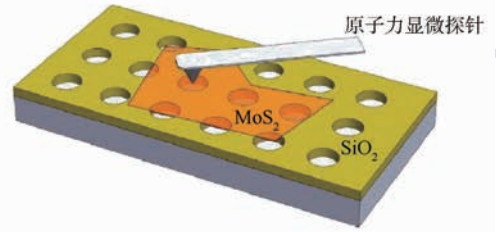
1.2.4 机械性能

与石墨烯相似,二维过渡金属硫化物材料也是一种较薄的具有内在柔韧性的力学材料,从而在力学领域具有良好的前景。

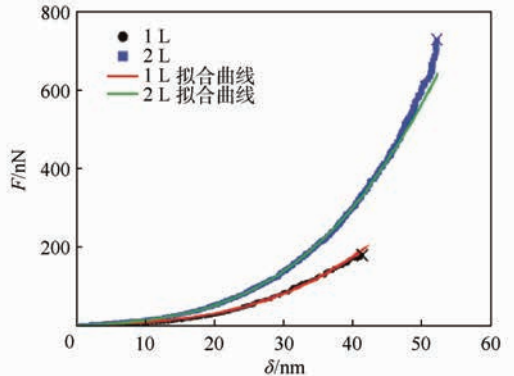
Bertolazzi等^[59]通过原子力显微镜对二硫化钨薄膜进行纳米压痕实验来测量单层和多层二硫化钨的强度和断裂强度。图7^[59]为悬浮二硫化钨板的纳米压痕。如图7(a)和图7(b)所示,在二氧化硅/硅基板上面通过转移较薄的二硫化钨



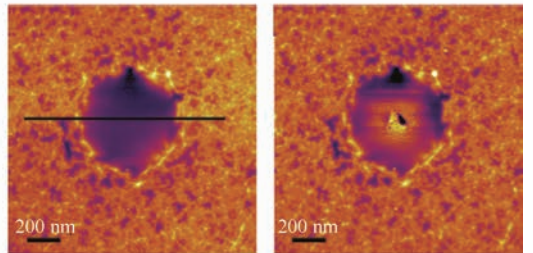
(a) 沉积在二氧化硅/硅基板上的单层二硫化钨的光学图像



(b) 原子力显微镜纳米压痕实验



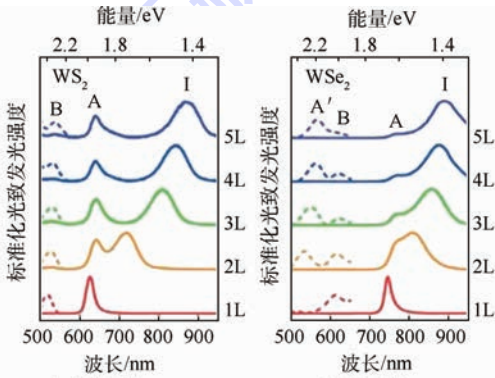
(c) 单层和双层二硫化钨力和薄膜挠度曲线



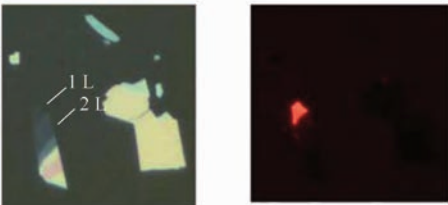
(d) 悬浮单层二硫化钨在纳米压痕之前的原子力显微照片 (e) 断裂发生之后的原子力显微照片

图7 悬浮二硫化钨板的纳米压痕^[59]

Fig. 7 Nanoindentation of suspended MoS₂ sheets^[59]



(a) 二硫化钨层数 $n=1,2,\dots,5$ 时的标准化光致发光性光谱 (b) 二硒化钨层数 $n=1,2,\dots,5$ 时的标准化光致发光性光谱



(c) 机械剥离二硫化钨的明场光学图像 (d) 机械剥离二硫化钨的荧光图像

A、A'、B和I—电子吸收峰。

图6 二硫化钨、二硒化钨的标准光致发光性光谱和机械剥离二硫化钨的荧光图像^[56]

Fig. 6 Normalized photoluminescence spectra of mechanically exfoliated WS₂, WSe₂ and fluorescence images of mechanically exfoliated WS₂^[56]

片到一排微制备出来的圆形孔上面来制备出二硫化钼薄膜。如图 7(c) 所示,在压痕实验中,原子力显微镜的尖端压在二硫化钼薄膜的中央,检测悬臂挠度,得到了力 F 和薄膜挠度 δ 的曲线。薄膜进一步被弯曲达到断裂点,检测到断裂发生时的力(见图 7(c))。在断裂发生时,在薄膜的中间会产生孔洞,也就是在原子力显微镜尖端刺破二硫化钼薄膜的位置。图 7(d) 和图 7(e) 是在拉伸和断裂实验前和后二硫化钼间歇接触的原子力显微照片。实验发现单层二硫化钼的弹性模量约为 70 GPa,高于钢铁的弹性模量(约为 205 GPa),该研究结果说明单层二硫化钼可以应用于复合材料的增强体和柔性电子器件等领域。 $\text{Li}^{[41]}$ 通过第一性原理研究发现超薄的二硫化钼片弹性模量约为 250 GPa,进一步说明过渡金属硫化物材料具有优异的机械性能。

1.2.5 摩擦性能

由于过渡金属硫化物层状材料层内由较强的化学键相连接,层间由较弱的范德华力相连接,因而具有较低的剪切阻力和摩擦系数,像二硫化钼和二硫化钨纳米结构可以作为润滑剂。据报道二硫化钼的摩擦系数在超高真空条件下量级是 10^{-3} 。在一般环境条件下,测量出的摩擦系数范围为 0.01 ~ 0.1^[60]。因此,过渡金属硫化物材料作为润滑材料可以广泛应用于超真空和车的动力系统中。

1.2.6 热性能

对过渡金属硫化物材料热性能的研究有利于其在热学器件中的应用。过渡金属硫化物材料具有较优异的热学稳定性,在高温下依然能保持较好的热稳定性,不发生分解,因而可以应用于高温器件中。另外,过渡金属硫化物材料具有优异的热力学传导性,Varshney 等^[61]发现二硫化钼具有热力学传导性质,但是 Yan 等^[62]发现二硫化钼的热力学传导性比石墨烯的要低。为了解决这个问题,Zhang 等^[63]构造二硫化钼/石墨烯纳米片板,发现热传导性比单独的二硫化钼材料要好,此研究成果说明可以通过构造异质结构来调节功能器件的热传导性。2014 年 Huang 等^[64]运用第一性原理研究了二硫化钼和二硫化钨的热电性能,其研究结果说明过渡金属硫化物材料是一种具有前景的热电材料。

1.2.7 储能性能

储能器件像锂硫电池或者超级电容器,对于便携电子设备、电动汽车和智能电网具有重要的意义。而过渡金属硫化物材料像二硫化钼在能量储存方面已经显示出巨大的优越性。最近,锂硫

电池由于具有低成本、较高的比容量($1673 \text{ mA} \cdot \text{h/g}$)和较高的理论比能量($2600 \text{ W} \cdot \text{h/kg}$)^[65-66]可以用于便携电子设备和电子运输中^[67]。但是, Li_2S 具有一些缺点,比如 Li_2S 绝缘性质以及锂聚硫物($\text{Li}_2\text{S}_m, m = 4 \sim 8$)会溶解于电解液中,导致能量的损失和低的库伦效率,为了解决这个问题,需要寻找一种传导性的材料将 Li_2S 阴极材料封装起来。Cui 课题组^[68]首次证实了二维过渡金属硫化物由于与 Li_2S 之间具有较强的相互作用,对锂硫电池的阴极材料来说是一种有效的封装材料,克服了以往碳基材料存在的缺陷,为过渡金属硫化物材料应用于高容量电极的封装材料提供了可能。但是位点与硫原子结合能之间的关系还不明确,理解晶体位点与探究晶体表面原子位点的活性以及暴露程度对于可循环能源储存具有重要的意义,而研究聚硫和不同晶体位点的结合对于控制聚硫沉积十分关键,为容量损失的控制提供了思路。图 8^[69]是 Li_2S 沉积在二硫化钼不同位点的示意图。结合能是指 Li_2S 与二硫化钼不同位点结合成键所产生的键能。Cui 课题组^[69]运用二维二硫化钼研究在锂硫电池和析氢反应中边缘的电化学选择性能。如图 8 所示,二硫化钼钼边缘和硫边缘具有悬键而与 Li_2S 相互作用比较强。该研究从理论和实验角度说明二维材料不同位点化学反应性的不同,为更好地提高理论电池的性能提供了途径。

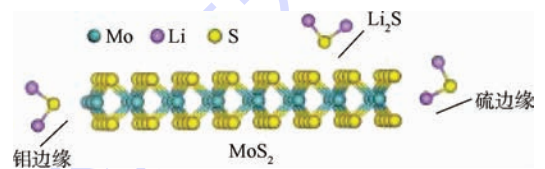


图 8 Li_2S 分子和 MoS_2 原子位点之间相互作用示意图^[69]

Fig. 8 Schematic of interaction between Li_2S molecules and different MoS_2 atomic sites^[69]

在锂硫电池里引入粘结材料来改善电池的性能,而有研究发现过渡金属硫化物材料可以作为粘结材料^[70]。但是目前还不能对粘结材料和 Li_2S_m 间的相互作用进行微观观察,现阶段粘结材料的选择还是主要靠经验。虽然也有人利用简化的锂硫电池做过一些计算模拟,但是还不能对其作用机理做清晰的描述^[71-72]。图 9^[73]是 Li-S 化合物在 6 个不同的锂化阶段(S_8 、 Li_2S_8 、 Li_2S_6 、 Li_2S_4 、 Li_2S_2 和 Li_2S)于不同的锚定材料上的结合能(E_b)示意图。在吸附未锂化的 S_8 时,相比于石墨烯上的吸附情况,这些锚定材料的结合能基本相同(0.75 ~ 0.85 eV)。但只要锂化过程一开始,

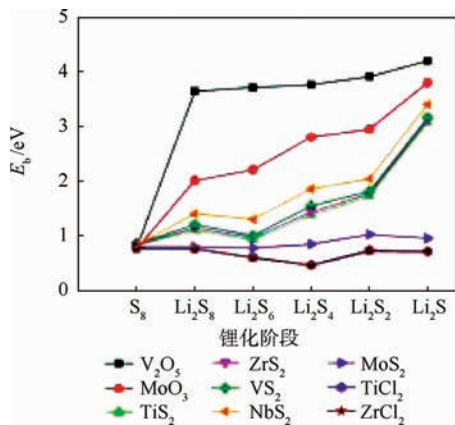


图9 Li-S化合物在6个不同的锂化阶段(S_8 、 Li_2S_8 、 Li_2S_6 、 Li_2S_4 、 Li_2S_2 和 Li_2S)于不同的锚定材料上的结合能示意图^[73]

Fig. 9 Schematic of binding energies for Li-S composites at six different lithiation stages (S_8 , Li_2S_8 , Li_2S_6 , Li_2S_4 , Li_2S_2 and Li_2S) on different anchoring materials we select^[73]

不同的材料就会产生差别明显的锚定效果。根据结合能大小的不同,这些材料可被归类为3种:氧化物(V_2O_5 、 MoO_3),硫化物(TiS_2 、 ZrS_2 、 VS_2 、 NbS_2 和 MoS_2)和氯化物($TiCl_2$ 、 $ZrCl_2$),并可被分别视作具有强的、中等强度的和弱的相互作用的锚定材料,而它们的结合能分别处在 $2.0 \sim 4.2$ eV、 $0.8 \sim 2.0$ eV和 $0.4 \sim 0.8$ eV的范围内。对于强的和中等强度的锚定材料,它们锂化过程中结合能的变化具有相同的趋势。对于强的锚定材料,结合能会随着锂化过程的进行而增大;对于中等强度的锚定材料,在 Li_2S_6 处结合能达到最小值,并在之后逐渐增大。强的和中等强度的锚定材料都显示了具有与碳基石墨烯锂化过程结合能逐渐降低相反的趋势。另一方面,效果弱的锚定材料的锚定效果与石墨烯类似。Zhang课题组^[73]研究了二维过渡金属硫化物层状材料对锂硫电池的粘结效应,对粘结材料与 Li_2S_m 分子在不同锂化阶段的结合能进行了系统的计算模拟。从图9^[73]中可以看出,过渡金属硫化物材料结合能处于 $0.8 \sim 2.0$ eV,为中等强度的粘结材料,其结合能在 Li_2S_6 处达到最低,之后逐渐增大。锂硫基团的分解意味着锂原子倾向于独立存在,而不是与硫原子成键,从而导致这2种元素的分离。这种结构上的破坏削弱了硫元素的作用,而硫元素是正极材料中的活性物质。锂硫键的削弱将会引起 Li_2S_m 分子中锂原子和硫原子的分离,使它们分别形成 Li^+ 和 S^{2-} ,并最终使硫元素溶解在电解液中。因此,诸如 V_2O_5 和 MoO_3 结合能大于 2.0 eV,过强的粘结材料并不是较好的选择,相比较而言,过渡金属硫化物材料能够在结合强度和

保持 Li_2S_m 的完整性间取得平衡,是比较优异的粘结材料。该研究为之后锂硫电池性能的改善提供了理论支持。另一方面,在过渡金属硫化物材料层间插入碱离子(Li^+ 、 Na^+ 和 K^+)可以改善过渡金属硫化物材料的性能,这种方法已经广泛应用于能量储存领域^[74]。Cui课题组^[75]研究了在二硫化钼层间插入锂离子来调节二硫化钼化学和物理性能,重点介绍了通过调节层间锂原子的浓度来调节二硫化钼的性能,并解释锂化过程二硫化钼光学性能和电导性能增强的原因。该研究有利于更好地理解锂插入过程和反复锂插入和脱嵌引起二硫化钼结构的改变,有力地解释了二硫化钼薄膜的容量损失,并阐明了基于二硫化钼能量储存器件性能提高的机理,反向研究提高过渡金属硫化物材料性能,为其在能量储存领域更好的应用提供了途径。

2 过渡金属硫化物纳米结构

低维的纳米条带结构由于具有优异的性能和广泛的应用,引起了广大研究者的注意。过渡金属硫化物层状材料具有不同的形态,比如纳米管、纳米团簇、纳米条带、纳米片和类富勒烯的纳米颗粒等。由于过渡金属硫化物材料衍生出来的结构十分丰富,因而具有许多优异的性能,可以广泛应用于各个领域。下面简要介绍一下过渡金属硫化物材料的几种形态。

2.1 过渡金属硫化物纳米管

过渡金属硫化物纳米管和碳纳米管^[74]的形态相似,在矿物质的沉积过程中^[76]能够形成过渡金属硫化物纳米管的结构。第一个合成出来的过渡金属硫化物纳米管是Tenne等^[77]于1992年合成出来的二硫化钨纳米管。有研究发现,过渡金属硫化物层具有较大的倾向形成纳米管的结构^[78]。

2.2 过渡金属硫化物纳米团簇

有限尺寸的纳米团簇可以合成或者从过渡金属硫化物材料中切割出来。有研究^[18]发现在 $Au(111)$ 上可以合成二硫化钼纳米团簇,成三角形形状。过渡金属硫化物材料不同于石墨烯,不是由同一种元素构成,所以不可能切割成矩形边缘以相同原子终止的纳米团簇,而三角形的纳米团簇结构是最简单的,而且能够具有同种边缘。实验上发现三角形纳米团簇以相同的边缘终止是最稳定的结构。体材料切割的六边形的二硫化钼团簇具有钼边缘和硫边缘的模型,如图10^[18]所示。三角形纳米团簇可以以钼边缘或者硫边缘终止,裸露的钼边缘具有悬键,能量上不稳定,而且

从实验角度制备不出来具有裸露钼边缘的纳米团簇,在硫原子过量的情况下,硫原子可以饱和这些悬键,这样钼边缘可以具有 50% 的硫覆盖率或者 100% 的硫覆盖率。纳米团簇另一个特性就是特定尺寸的纳米团簇是比较稳定的。实验中,边缘钼原子的个数为偶数时的纳米团簇可以制备出来 ($l=6,8,10,12$)。这种特性是由于边缘上过多的硫原子饱和钼原子的悬键,过多的硫原子就会形成二聚体,导致边缘发生重构,而二聚体是以倍数增长的形式进行重构,所以只有当二聚体的个数为偶数时才能发生重构,每一个二聚体连接一个钼原子,所以钼原子的个数也为偶数^[19]。

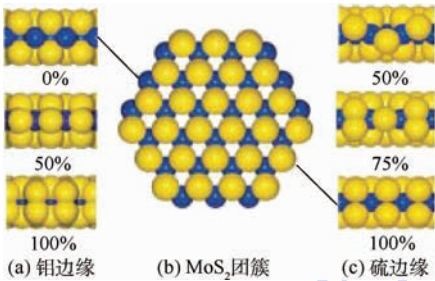


图 10 体材料切割的六边形的二硫化钼团簇具有钼边缘和硫边缘的模型^[18]

Fig. 10 Model of a bulk-truncated hexagonal MoS_2 nanocluster exposing Mo-edges and S-edges^[18]

2.3 过渡金属硫化物纳米条带

过渡金属硫化物纳米条带是宽度小于 100 nm 的条带结构。纳米条带的结构可以用于纳米器件的制造。近年来,有一些研究者研究了超薄纳米条带的制备方法。Liu 等^[79]通过电子束辐射制备出较窄的二硫化钼纳米条带。Helveg 等^[80]通过扫描电子显微镜发现 Au(111) 基体上可以生长出单层的二硫化钼。大量的研究报道发现纳米条带的性能依赖于边缘态,所以在第 3 节中,主要介绍边缘态对过渡金属硫化物纳米团簇或者纳米条带性能的影响。

3 过渡金属硫化物层状材料边缘态

当过渡金属硫化物层状材料变成纳米条带或者三角形的纳米团簇的结构时,由于边缘态的存在,材料的性能发生显著的变化。

3.1 边缘态的稳定性

图 11 为二硫化钼纳米条带计算的边缘态能量^[81]。图中:ac-u 为不对称的扶手椅形二硫化钼边缘,ac-s 为对称的扶手椅形二硫化钼边缘,zz-u 为不对称的之字形二硫化钼边缘,zz-s 为对称的之字形二硫化钼边缘,zz-Mo-u 为不对称的以钼原

子终止的之字形二硫化钼边缘,zz-Mo-s 为对称的以钼原子终止的之字形二硫化钼边缘,zz-S-u 为不对称的以硫原子终止的之字形二硫化钼边缘,zz-S-s 为对称的以硫原子终止的之字形二硫化钼边缘。从图 11 中可以看出纳米条带边缘态的稳定性依赖于手征性和对称性。根据第一性原理的计算,在不考虑氢饱和度的情况下,硫终止的之字形纳米条带是最稳定的,可以看出硫终止的具有对称性的之字形纳米条带的边缘态能量比其他的纳米条带的能量都要低,因而可以推测出硫原子终止方式可以稳定边缘态^[81]。另外一种稳定边缘态的方式就是通过氢饱和和边缘的悬键,来稳定纳米条带的边缘态。从图中可以看出氢饱和的纳米条带比没有氢饱和的条带能量更低,也更稳定。为了研究纳米条带稳定性和纳米条带宽度之间的关系,Li 等^[82]计算了纳米条带的键能,进而得到纳米条带的稳定性。图 12^[82]给出了纳米条带键能和纳米条带宽度的曲线。图中: p 为 MoS_2 三角形条带边缘 Mo 原子的个数。之字形和扶手椅形纳米条带随着宽度的上升键能单调上升,而之字形纳米条带的结合能比扶手椅形纳米条带的高,进一步证实了 Pan 和 Zhang^[81]之字形二硫化钼纳米条带更稳定的结论。

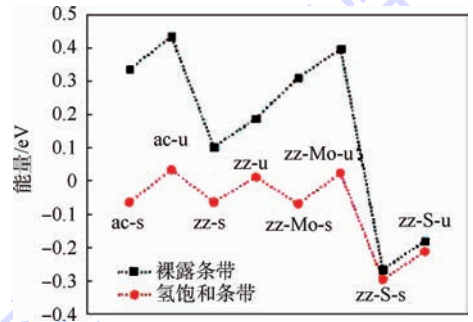


图 11 二硫化钼纳米条带计算的边缘态能量^[81]

Fig. 11 Calculated edge energies of MoS_2 nanoribbons^[81]

3.2 边缘态的电子性能

过渡金属硫化物纳米条带的电子态依赖于它的边缘结构。从实验的角度,在金基板上三角形的二硫化钼纳米团簇边缘处在扫描电镜下呈现金属电子态^[19,80]。Bollinger 等^[83]通过扫描隧道显微镜 (STM) 发现单层的二硫化钼三角形纳米微粒 (约为 10 nm^2) 在边界呈现局域态。如图 13^[83]所示,浅色的边缘具有金属性,在图中,对于比较连续的边缘态可以观察到纳米尺寸的金属性的封闭的线路。根据计算的能量^[84],在 H_2S 过量的情况下,Mo 边缘吸附 S 原子二聚体最可能存在的均衡边缘结构。通过第一性原理对单层二硫化钼进行第一性原理的计算,发现是由于 Mo 边缘存在过

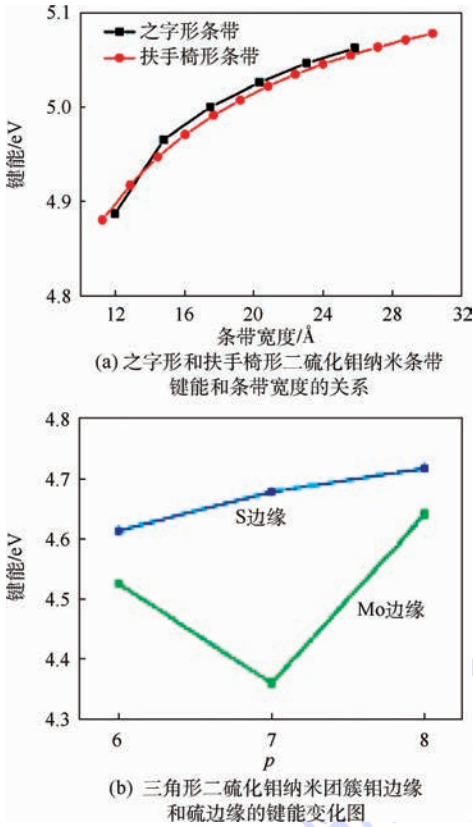


图 12 二硫化钼纳米条带键能和条带宽度的关系及三角形二硫化钼纳米团簇钼边缘和硫边缘的键能变化^[82]

Fig. 12 Relationship between binding energy and width of MoS₂ nanoribbons and variation in binding energies of triangular MoS₂ nanoclusters terminated by a Mo-edge or S-edge^[82]

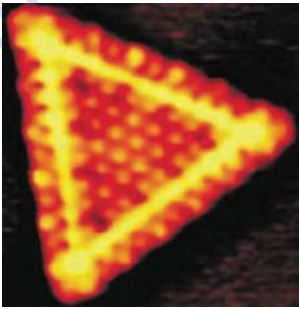


图 13 单个二硫化钼纳米团簇在重建金(111)上的 STM 照片^[83]

Fig. 13 STM photograph showing single triangular MoS₂ nanocluster dispersed on a reconstructed Au(111) surface^[83]

多的 S 原子,所以单层二硫化钼呈现金属性。

另外, Li 和 Galli^[85] 计算得到堆叠起来多层的二硫化钼纳米线的边缘态具有金属性。而 Ataca 等^[86] 得到裸露的二硫化钼扶手椅形纳米条带具有半导体性质,且在氢气饱和的情况下带隙变大。Humberto 等^[28] 计算了二硫化钼的部分能态密度图,结果发现边缘态是金属性的,而内部区域是半导体性。

需要特别强调的是,边缘态的性质对边缘态的结构非常敏感,根据第一性原理计算,钼边缘吸收硫单体呈现的是半导体性,而吸收的是硫二聚体则呈现为金属性。如图 14^[37] 所示,有研究已经发现电子结构受边缘的构型和覆盖率的影响,这为电子结构的调节又提供了一条思路^[87]。

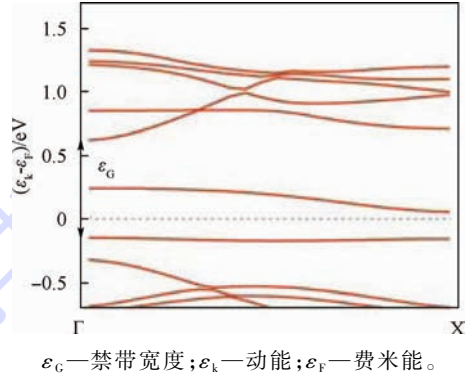


图 14 二硫化钼条带钼边缘吸收硫单体的费米能附近电子能带结构^[37]

Fig. 14 Electronic band structure around Fermi level of MoS₂ stripe with S-monomers adsorbed at Mo-edge^[37]

一些研究发现也证实了过渡金属硫化物层状材料的边缘态对性能具有重要的作用,且边缘态影响费米能附近的能带。由于边缘态的存在,过渡金属硫化物材料电子性能在外界刺激下会发生相应的变化。Dolui 等^[88] 的研究表明了二硫化钼纳米条带边缘态的存在,其扶手椅形的带隙在外加横向电场作用下显著地下降,并且在特定的临界场值发生了金属-绝缘体的转变。

3.3 边缘态的催化活性

加氢脱硫反应多年来一直是催化剂领域的一个研究热点。加氢脱硫反应是将精炼石油产物(汽油、喷气燃料和柴油)脱去硫的化学催化过程。过渡金属硫化物材料比如二硫化钼和二硫化钨可以广泛地应用于脱氢加硫反应^[89]。而加氢脱硫反应研究的重点是确定催化剂结构和加氢脱硫活性之间的关系。一般认为活性位点位于边缘态并且和边缘上不饱和原子有关。通过原子力扫描隧道显微镜和第一性原理相结合的方法证实了二硫化钼存在化学反应活性是由于边缘态的存在^[90]。运用一些原位技术,例如扩展 X 射线吸收精细结构来研究二硫化钼小的团簇,也发现了边缘态存在活性位点^[91-92]。也有研究用氧气、一氧化碳和一氧化氮作为探针分子来探测到活性位点,也证实了活性位点位于二硫化钼结构的边缘^[93-94]。

电化学析氢反应被看作是可持续生产氢气的手段^[95]。在电催化反应中,发现二硫化钼纳米结构的边缘态在析氢反应中具有活性^[96]。目前,二

硫化钼被看作是一种可用于电催化析氢反应中具有前景性的材料^[97]。虽然,应用比较广泛的电催化析氢反应的催化剂是铂,但是二硫化钼价廉、丰富,可以代替铂应用于电催化析氢反应中。二硫化钼材料是一种催化性能较差,而低维的二硫化钼团簇在基体(如金^[98]、石墨^[99]、氧化石墨烯^[100])是一种比较优异的电催化析氢反应的催化剂。二硫化钼已经被证实其边缘态在电催化析氢反应中具有反应活性。对于其他的过渡金属硫化物材料,像二硫化钨、二硒化钼和二碲化钨也有研究发现它们在类似的电催化析氢反应过程中具有活性^[101-103]。这些研究成果的发现为过渡金属硫化物材料在催化剂领域的应用提供了可能。

3.4 边缘态的磁性能

一些研究发现磁矩产生在二硫化钼纳米团簇、纳米颗粒和纳米条带中。一般来说,磁性产生主要有2种原因,一种是材料表面存在大量的点缺陷,另一种是材料具有边缘态。金属空位附近存在高密度的自旋态^[39]。纳米团簇的磁性能分析得出磁性主要来源于边缘态的不饱和原子。

之前对非磁性过渡金属硫化物材料的研究主要集中在通过化学功能化^[48]、对二维过渡金属硫化物层状材料施加应力^[104]和外加电场^[3]等手段来引发磁矩。另外也有研究发现过渡金属硫化物层状材料中产生空位也会导致磁矩的产生^[105]。实验中运用一些外加手段可以在过渡金属硫化物层状材料中引起缺陷,像离子辐射能够在材料中引入大量的缺陷,进而改变材料的结构和磁学性能。研究发现^[106],室温条件下二硫化钼在2 MeV 质子束作用下能够显示出磁序,这个发现说明可以利用活性离子调节过渡金属硫化物材料的性能。

除了外界条件引发生缺陷和空位外,边缘态的存在对磁性的影响也得到了广泛地研究。理论计算^[107]表明三维的 $(\text{MoS}_2)_q$ 中 Mo_6S_8 是具有磁性的。图15^[108]是不同厚度和形态的磁滞回线, H 为磁场强度。单晶二硫化钼在相似测试条件下显示出无磁性的状态,较厚的样本I具有较厚的边缘。而较薄的样本II是较薄的纳米片具有高密度边缘态。结果表明样本II比样本I具有较强的磁性强度。Zhang等^[108]通过不规则的纳米晶体合成了二硫化钼薄膜,在室温条件下呈现出铁磁响应。作者认为滞后曲线非空矫顽磁场是由于纳米微粒边缘存在大量的自旋态。而这个结果说明边缘态的存在能够影响材料的电子、磁学和化学性能,因而纳米材料可以应用于更广泛的领域。Zhang等^[108]还对三角形团簇不同覆盖率

的模型进行了自旋计算,结果发现钼边缘0 硫覆盖率率和硫边缘100% 硫覆盖率率的情况具有磁性,而磁性的存在主要与边缘自旋态有关。Zhang等^[108]的研究说明边缘态的种类和覆盖率都会对磁性产生影响。

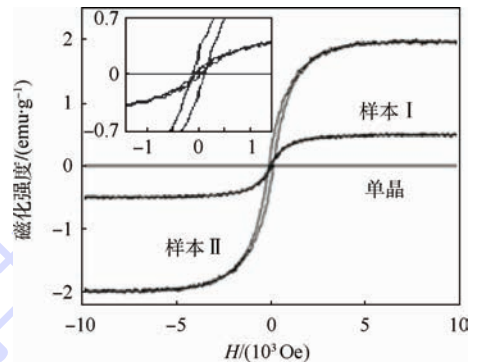


图15 室温下2个二硫化钼样本(I和II)记录的磁滞回线^[108]

Fig. 15 Hysteresis curves recorded at room temperature from two MoS_2 samples (I and II)^[108]

实验中,由于磁力显微镜(MFM)具有和原子力显微镜类似的纳米尺度的分辨率,可以观察到纳米级的磁畴,所以运用磁力显微镜可以观察到微米级和纳米级的磁性和非磁性的响应。有研究者用磁力显微镜来表征机械剥离的单层和多层的二硫化钼和石墨烯纳米条带。通过分析相和振幅的改变发现单层或者多层的二硫化钼和石墨烯纳米条带具有磁信号响应,当二硫化钼和石墨烯的厚度超过一定值后,就会变成非磁性状态^[109]。也有实验进一步指出是由于在二硫化钼材料中存在边缘态才会出现磁信号响应^[110]。图16^[109]是二硫化钼和石墨烯纳米条带的磁力显微镜的相图。通过分析,二硫化钼和石墨烯的磁性响应依赖于层数。

除此之外,第一性原理计算^[81-82]表明二硫化钼纳米条带之字形边缘呈现铁磁性,扶手椅形呈现非磁性,而有研究发现^[111]扶手椅形纳米条带在氢饱和情况下呈金属性,且具有磁性。并且它

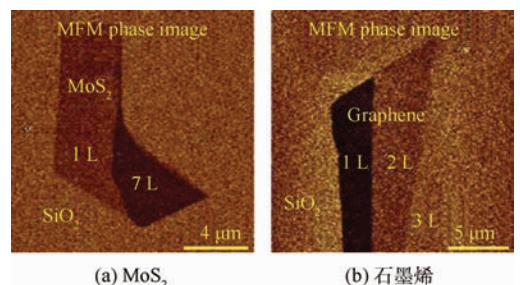


图16 二硫化钼和石墨烯纳米条带磁力显微镜的相图^[109]

Fig. 16 Magnetic force microscopy phase image of MoS_2 and graphene nanosheets^[109]

们的磁性可以通过一些手段进行调整,比如外加电场、弹性力或者引入一些缺陷等^[56,88,112-114]。进一步,有研究者对纳米条带的边缘用氢气进行钝化处理,发现硫边缘在加氢脱硫反应的条件下仍然显示磁性^[114]。Pan等^[56]研究了应变效应对二硫化钼纳米条带磁性能的影响,在外加应变作用下,之字形边缘的磁性能提高,这为边缘态磁性能的调节提供了思路。随后,有研究者通过调整应力和电场的组合效应来调整二硫化钼之字形边缘的磁性能^[113]。基于之前的研究,应力和电场的组合效应对扶手椅形二硫化钼纳米条带磁性的影响也被提出来^[115]。

文献[105-108]的研究对象主要集中在二硫化钼材料上,文献[116]发现对化学气相沉积的方法制备出的具有边缘态的二硫化钼纳米片进行磁性能的研究,测试结果显示测试样品在室温条件下磁化强度为 $1.4 \text{ emu} \cdot \text{g}^{-1}$,磁滞行为达到了700 K说明其具有较强的铁磁性,且具有较高的居里温度。另外,作者也运用第一性原理自旋极化计算得到二硫化钼纳米片铁磁性磁矩是由于之字形边缘的存在。这个结果说明具有高密度边缘自旋的二硫化钼纳米片也可以用于制作自旋电子器件^[116]。也有研究者^[18]对制备出来的二硫化钨纳米层状材料进行研究,发现二硫化钨纳米层状材料具有铁磁性,室温下饱和磁化强度(M_s)达到了 $0.0058 \text{ emu} \cdot \text{g}^{-1}$,矫顽磁场为 92 Oe。

文献[4,116]已经比较全面地探究了磁性产生的原因,边缘态原子不饱和,配位不同于内层原子,因而边缘态原子呈现磁性。二硫化钼之字形边缘的自旋密度图 17^[106]所示,从自旋密度分布图中可以看出磁矩主要产生在边缘的钼原子和硒原子上。图 18^[28]是长方形二硫化钨纳米片之字形边缘的能态密度。对具有磁性的二硫化钨进行总态密度分析,发现在费米能附近自旋向上和自旋向下不对称(如图 18(a)所示),因此说明二硫化钨纳米层状材料具有磁性。对边缘硫原子和

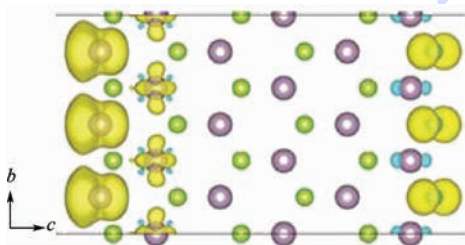


图 17 二硫化钼纳米条带之字形边缘自旋密度的等值面图^[106]

Fig. 17 Isosurface plot of spin density of MoSe_2 nanoribbons with zigzag edges^[106]

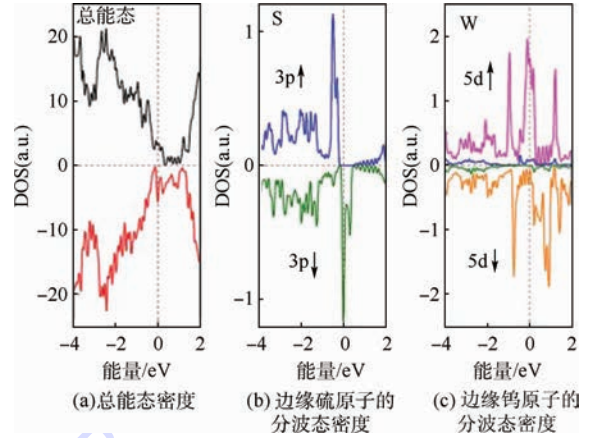


图 18 长方形二硫化钨纳米片之字形边缘的能态密度^[28]

Fig. 18 Density of state of rectangular zigzag-edge WS_2 nanosheets^[28]

钨原子进行分波态密度分析,如图 18(b)和图 18(c)所示,发现在费米能附件硫原子的 3p 轨道和钨原子的 5d 轨道出现劈裂,进一步证实边缘原子对纳米层状材料的磁性有贡献^[4]。另外,也有第一性原理计算^[43]对边缘磁矩分布进行分析,发现磁性主要产生在边缘钨原子的 4d 轨道和硫原子的 3p 轨道。不饱和的自旋态主要集中在边缘的钨原子和硫原子上。

4 结论

- 1) 以二硫化钼为代表的过渡金属硫化物层状材料由于具有电子性能、光学性能和机械性能等而成为应用于纳米电子学和光电子学领域的新一代材料。
- 2) 过渡金属硫化物层状材料纳米条带可以通过实验的方法得到,这使得纳米条带应用于更广的领域。
- 3) 过渡金属硫化物材料可以衍生出不同形态的材料,因而具有更加广泛的应用。
- 4) 边缘态的存在能够影响低维层状材料的性能,使其与体材料的性能具有明显的差异,因而引起广大研究者的注意。

本文介绍了过渡金属硫化物层状材料近年来的研究热点,重点放在低维纳米过渡金属硫化物层状材料边缘态对性能的影响,其中,边缘态磁性能的研究已经成为当今的研究热点,边缘态磁性能的研究为过渡金属硫化物材料应用于自旋电子学领域提供了支持。

参考文献 (References)

- [1] WOLF S A, AWSCHALOM D D, BUHRMAN R A, et al. Spintronics: A spin-based electronics vision for the future [J] Sci-

- ence, 2001, 294 (5546) : 1488-1495.
- [2] HUANG J, DONG Z, LI Y, et al. MoS₂ nanosheet functionalized with Cu nanoparticles and its application for glucose detection [J]. *Materials Research Bulletin*, 2013, 48 (11) : 4544-4547.
- [3] NOVOSELOV K S, GEIM A K, MOROZOV S V, et al. Electric field effect in atomically thin carbon films [J]. *Science*, 2004, 306 (5696) : 666-669.
- [4] HUO N, LI Y, KANG J, et al. Edge-states ferromagnetism of WS₂ nanosheets [J]. *Applied Physics Letters*, 2014, 104 : 202406.
- [5] SHISHIDOU T, FREEMAN A J, ASAHI R. Effect of GGA on the half-metallicity of the itinerant ferromagnet CoS₂ [J]. *Physical Review B*, 2001, 64 (18) : 180401.
- [6] TAKADA K, SAKURAI H, TAKAYAMA-MUROMACHI E, et al. Superconductivity in two-dimensional CoO₂ layers [J]. *Nature*, 2003, 422 (6927) : 53-55.
- [7] MENG F, LI J, CUSHING S K, et al. Solar hydrogen generation by nanoscale p-n junction of p-type molybdenum disulfide/n-type nitrogen-doped reduced graphene oxide [J]. *Journal of the American Chemical Society*, 2013, 135 (28) : 10286-10289.
- [8] LEE C, LI Q, KALB W, et al. Frictional characteristics of atomically thin sheets [J]. *Science*, 2010, 328 (5974) : 76-80.
- [9] FENG J, SUN X, WU C, et al. Metallic few-layered VS₂ ultra-thin nanosheets: High two-dimensional conductivity for in-plane supercapacitors [J]. *Journal of the American Chemical Society*, 2011, 133 (44) : 17832-17838.
- [10] RADISAVLJEVIC B, RADENOVIC A, BRIVIO J, et al. Single-layer MoS₂ transistors [J]. *Nature Nanotechnology*, 2011, 6 (3) : 147-150.
- [11] PERERA M M, LIN M W, CHUANG H J, et al. Improved carrier mobility in few-layer MoS₂ field-effect transistors with ionic-liquid gating [J]. *ACS Nano*, 2013, 7 (5) : 4449-4458.
- [12] CHANG M H, CHEN C J, LAI M S, et al. Universal hepatitis B vaccination in Taiwan and the incidence of hepatocellular carcinoma in children [J]. *New England Journal of Medicine*, 1997, 336 (26) : 1855-1859.
- [13] LARENTIS S, FALLAHAZAD B, TUTUC E. Field-effect transistors and intrinsic mobility in ultra-thin MoSe₂ layers [J]. *Applied Physics Letters*, 2012, 101 (22) : 223104.
- [14] CHAMLAGAIN B, LI Q, GHIMIRE N J, et al. Mobility improvement and temperature dependence in MoSe₂ field-effect transistors on parylene-C substrate [J]. *ACS Nano*, 2014, 8 (5) : 5079-5088.
- [15] WABG H, YU L, LEE Y H, et al. Integrated circuits based on bilayer MoS₂ transistors [J]. *Nano Letters*, 2012, 12 (9) : 4674-4680.
- [16] LI H, YIN Z, HE Q, et al. Fabrication of single-and multilayer MoS₂ film-based field-effect transistors for sensing NO at room temperature [J]. *Small*, 2012, 8 (1) : 63-67.
- [17] PERKINS F K, FRIEDMAN A L, COBAS E, et al. Chemical vapor sensing with monolayer MoS₂ [J]. *Nano Letters*, 2013, 13 (2) : 668-673.
- [18] EDA G, YAMAGUCHI H, VOIRY D, et al. Photoluminescence from chemically exfoliated MoS₂ [J]. *Nano Letters*, 2011, 11 (12) : 5111-5116.
- [19] LAURITSEN J V, KIBSGAARD J, HELVEG S, et al. Size-dependent structure of MoS₂ nanocrystals [J]. *Nature Nanotechnology*, 2007, 2 (1) : 53-58.
- [20] ZHAN Y, LIU Z, NAJMAEI S, et al. Large-area vapor-phase growth and characterization of MoS₂ atomic layers on a SiO₂ substrate [J]. *Small*, 2012, 8 (7) : 966-971.
- [21] PENG Y, MENG Z, ZHONG C, et al. Hydrothermal synthesis and characterization of single-molecular-layer MoS₂ and MoSe₂ [J]. *Chemistry Letters*, 2001 (8) : 772-773.
- [22] SALVATORE G A, MÜNZENRIEDER N, BARRAUD C, et al. Fabrication and transfer of flexible few-layers MoS₂ thin film transistors to any arbitrary substrate [J]. *ACS Nano*, 2013, 7 (10) : 8809-8815.
- [23] WU J, LI H, YIN Z, et al. Layer thinning and etching of mechanically exfoliated MoS₂ nanosheets by thermal annealing in air [J]. *Small*, 2013, 9 (19) : 3314-3319.
- [24] LI Q, NEWBERG J T, WALTE E C, et al. Polycrystalline molybdenum disulfide (2H-MoS₂) nano-and microribbons by electrochemical/chemical synthesis [J]. *Nano Letters*, 2004, 4 (2) : 277-281.
- [25] NOVOSELOV K S, JIANG D, SCHEIDIN F, et al. Two-dimensional atomic crystals [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2005, 102 (30) : 10451-10453.
- [26] COLEMAN J N, LOYTA M, O' NEILL A, et al. Two-dimensional nanosheets produced by liquid exfoliation of layered materials [J]. *Science*, 2011, 331 (6017) : 568-571.
- [27] JOENSEN P, FRINDT R F, MORRISON S R. Single-layer MoS₂ [J]. *Materials Research Bulletin*, 1986, 21 (4) : 457-461.
- [28] GUTIÉRREZ H R, PEREA-LÓPEZ N, ELÍAS A L, et al. Extraordinary room-temperature photoluminescence in triangular WS₂ monolayers [J]. *Nano Letters*, 2013, 13 (8) : 3447-3454.
- [29] QIU H, XU T, WANG Z, et al. Hopping transport through defect-induced localized states in molybdenum disulphide [J]. *Nature Communications*, 2013, 4 : 2642-2647.
- [30] SPLENDIANI A, SUN L, ZHANG Y, et al. Emerging photoluminescence in monolayer MoS₂ [J]. *Nano Letters*, 2010, 10 (4) : 1271-1275.
- [31] ATACA C, SAHIN H, CIRAC S. Stable, single-layer MX₂ transition-metal oxides and dichalcogenides in a honeycomb-like structure [J]. *The Journal of Physical Chemistry C*, 2012, 116 (16) : 8983-8999.
- [32] DING Y, WANG Y, NI J, et al. First principles study of structural, vibrational and electronic properties of graphene-like MX₂ (M = Mo, Nb, W, Ta; X = S, Se, Te) monolayers [J]. *Physica B: Condensed Matter*, 2011, 406 (11) : 2254-2260.
- [33] KOBAYASHI K, YAMAUCHI J. Electronic structure and scanning-tunneling-microscopy image of molybdenum dichalcogenide surfaces [J]. *Physical Review B*, 1995, 51 (23) : 17085-17095.
- [34] KUC A, ZIBOUCHE N, HEINE T. Influence of quantum confinement on the electronic structure of the transition metal sulfide TS₂ [J]. *Physical Review B*, 2011, 83 (24) : 2237-2249.
- [35] LIU L, KUMAR S B, OUYANG Y, et al. Performance limits of monolayer transition metal dichalcogenide transistors [J]. *IEEE Transactions on Electron Devices*, 2011, 58 (9) : 3042-3047.
- [36] LEBEGUE S, ERIKSSON O. Electronic structure of two-dimen-

- sional crystals from ab initio theory [J]. *Physical Review B*, 2009, 79(11):115409-115409.
- [37] BOLLINGER M V, JACOBSEN K W, NØRSKOV J K. Atomic and electronic structure of MoS₂ nanoparticles [J]. *Physical Review B*, 2003, 67(8):283-287.
- [38] KOM T, HEYDRICH S, HIRMER M, et al. Low-temperature photocarrier dynamics in monolayer MoS₂ [J]. *Applied Physics Letters*, 2011, 99(10):102109.
- [39] MA Y, DAI Y, GUO M, et al. Electronic and magnetic properties of perfect, vacancy-doped, and nonmetal adsorbed MoSe₂, MoTe₂ and WS₂ monolayers [J]. *Physical Chemistry Chemical Physics*, 2011, 13(34):15546-15553.
- [40] KUMAR A, AHLUWALIA P K. Electronic structure of transition metal dichalcogenides monolayers 1H-MX₂ (M = Mo, W; X = S, Se, Te) from ab-initio theory: New direct band gap semiconductors [J]. *The European Physical Journal B*, 2012, 85(6):1-7.
- [41] LI T. Ideal strength and phonon instability in single-layer MoS₂ [J]. *Physical Review B*, 2012, 85(23):235407.
- [42] YUE Q, KANG J, SHAO Z, et al. Mechanical and electronic properties of monolayer MoS₂ under elastic strain [J]. *Physics Letters A*, 2012, 376(12):1166-1170.
- [43] JOHARI P, SHENOY V B. Tuning the electronic properties of semiconducting transition metal dichalcogenides by applying mechanical strains [J]. *ACS Nano*, 2012, 6(6):5449-5456.
- [44] SCALISE E, HOUSSA M, POURTOIS G, et al. Strain-induced semiconductor to metal transition in the two-dimensional honeycomb structure of MoS₂ [J]. *Nano Research*, 2012, 5(1):43-48.
- [45] SHI H, PAN H, ZHANG Y W, et al. Quasiparticle band structures and optical properties of strained monolayer MoS₂ and WS₂ [J]. *Physical Review B*, 2013, 87(15):155304.
- [46] YUN W S, HAN S W, HONG S C, et al. Thickness and strain effects on electronic structures of transition metal dichalcogenides: 2H-MX₂ semiconductors (M = Mo, W; X = S, Se, Te) [J]. *Physical Review B*, 2012, 85(3):033305.
- [47] RAMASUBRAMANIAM A, NAVEH D, TOWE E. Tunable band gaps in bilayer transition-metal dichalcogenides [J]. *Physical Review B*, 2011, 84(20):205325.
- [48] ATACA C, CIRACI S. Functionalization of single-layer MoS₂ honeycomb structures [J]. *The Journal of Physical Chemistry C*, 2011, 115(27):13303-13311.
- [49] HE J, WU K, SA R, et al. Magnetic properties of nonmetal atoms adsorbed MoS₂ monolayers [J]. *Applied Physics Letters*, 2010, 96(8):082504.
- [50] KOH E W K, CHIU C H, LIM Y K, et al. Hydrogen adsorption on and diffusion through MoS₂ monolayer: First-principles study [J]. *International Journal of Hydrogen Energy*, 2012, 37(19):14323-14328.
- [51] CAO T, WANG G, HAN W, et al. Valley-selective circular dichroism of monolayer molybdenum disulphide [J]. *Nature Communications*, 2012, 3(2):177-180.
- [52] ZENG H, DAI J, YAO W, et al. Valley polarization in MoS₂ monolayers by optical pumping [J]. *Nature Nanotechnology*, 2012, 7(8):490-493.
- [53] RAMASUBRAMANIAM A. Large excitonic effects in monolayers of molybdenum and tungsten dichalcogenides [J]. *Physical Review B*, 2012, 86(11):115409.
- [54] SALLEN G, BOUET L, MARIE X, et al. Robust optical emission polarization in MoS₂ monolayers through selective valley excitation [J]. *Physical Review B*, 2012, 86(8):081301.
- [55] XIAO D, LIU G B, FENG W, et al. Coupled spin and valley physics in monolayers of MoS₂ and other group-VI dichalcogenides [J]. *Physical Review Letters*, 2012, 108(19):196802.
- [56] ZHAO W, GHORANNEVIS Z, CHU L, et al. Evolution of electronic structure in atomically thin sheets of WS₂ and WSe₂ [J]. *ACS Nano*, 2013, 7(1):791-797.
- [57] PAN H, ZHANG Y W. Tuning the electronic and magnetic properties of MoS₂ nanoribbons by strain engineering [J]. *The Journal of Physical Chemistry C*, 2012, 116(21):11752-11757.
- [58] ZHU Z Y, CHENG Y C, SCHWINGENSHLÖGL U. Giant spin-orbit-induced spin splitting in two-dimensional transition-metal dichalcogenide semiconductors [J]. *Physical Review B*, 2011, 84(15):153402.
- [59] BERTOLAZZI S, BRIVIO J, KIS A. Stretching and breaking of ultrathin MoS₂ [J]. *ACS Nano*, 2011, 5(12):9703-9709.
- [60] DALLAYALLE M, SÄNDIG N, ZERBETTO F. Stability, dynamics, and lubrication of MoS₂ platelets and nanotubes [J]. *Langmuir*, 2012, 28(19):7393-7400.
- [61] VARSHNEY V, PATNAIK S S, MURATORE C, et al. MD simulations of molybdenum disulphide (MoS₂): Force-field parameterization and thermal transport behavior [J]. *Computational Materials Science*, 2010, 48(1):101-108.
- [62] YAN R, SIMPSON J R, BERTOLAZZI S, et al. Thermal conductivity of monolayer molybdenum disulfide obtained from temperature-dependent Raman spectroscopy [J]. *ACS Nano*, 2014, 8(1):986-993.
- [63] ZHANG Z, XIE Y, PENG Q, et al. Thermal transport in MoS₂/graphene hybrid nanosheets [J]. *Nanotechnology*, 2015, 26(37):375402.
- [64] HUANG W, LUO X, GAN C K, et al. Theoretical study of thermoelectric properties of few-layer MoS₂ and WSe₂ [J]. *Physical Chemistry Chemical Physics*, 2014, 16(22):10866-10874.
- [65] JI X, LEE K T, NAZAR L F. A highly ordered nanostructured carbon-sulphur cathode for lithium-sulphur batteries [J]. *Nature Materials*, 2009, 8(6):500-506.
- [66] BRUCE P G, FREUNBERGER S A, HARDWICK L J, et al. Li-O₂ and Li-S batteries with high energy storage [J]. *Nature Materials*, 2012, 11(1):19-29.
- [67] YANG Y, ZHENG G, CUI Y. Nanostructured sulfur cathodes [J]. *Chemical Society Reviews*, 2013, 42(7):3018-3032.
- [68] SEH Z W, YU J H, LI W, et al. Two-dimensional layered transition metal disulphides for effective encapsulation of high-capacity lithium sulphide cathodes [J]. *Nature Communications*, 2014, 5:5017-5017.
- [69] WANG H T, ZHANG Q F, YAO H B, et al. High electrochemical selectivity of edge versus terrace sites in two-dimensional layered MoS₂ materials [J]. *Nano Letters*, 2014, 14(12):7138-7144.
- [70] SEH Z W, ZHANG Q, LI W, et al. Stable cycling of lithium sulfide cathodes through strong affinity with a bifunctional binder [J]. *Chemical Science*, 2013, 4(9):3673-3677.
- [71] LI W, ZHANG Q, ZHENG G, et al. Understanding the role of

- different conductive polymers in improving the nanostructured sulfur cathode performance [J]. *Nano Letters*, 2013, 13 (11) : 5534-5540.
- [72] DAVID L, BHANAVAT R, SINGH G. MoS₂/graphene composite paper for sodium-ion battery electrodes [J]. *ACS Nano*, 2014, 8 (2) : 1759-1770.
- [73] ZHANG Q, WANG Y, SEH Z W, et al. Understanding the anchoring effect of two-dimensional layered materials for lithium-sulfur batteries [J]. *Nano Letters*, 2015, 15 (6) : 3780-3786.
- [74] HJIMAIJ S. Helical microtubules of graphitic carbon [J]. *Nature*, 1991, 354 (6348) : 56-58.
- [75] XIONG F, WANG H, LIU X, et al. Li intercalation in MoS₂: In situ observation of its dynamics and tuning optical and electrical properties [J]. *Nano Letters*, 2015, 15 (10) : 6777-6784.
- [76] HARRIS P J F. Carbon nanotubes and related structures: New materials for the twenty-first century [J]. *American Journal of Physics*, 2001, 72 (3) : 415-415.
- [77] TENNE R, MARGULIS L, GENUT M, et al. Polyhedral and cylindrical structures of tungsten disulphide [J]. *Nature*, 1992, 360 (6403) : 444-446.
- [78] TENNE R, RAO C N R. Inorganic nanotubes [J]. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 2004, 362 (1823) : 2099-2125.
- [79] LIU X, XU T, WU X, et al. Top-down fabrication of sub-nanometre semiconducting nanoribbons derived from molybdenum disulfide sheets [J]. *Nature Communications*, 2013, 4 (2) : 216-219.
- [80] HELVEG S, LAURITSEN J V, LÆGSGAARD E, et al. Atomic-scale structure of single-layer MoS₂ nanoclusters [J]. *Physical Review Letters*, 2000, 84 (5) : 951-954.
- [81] PAN H, ZHANG Y W. Edge-dependent structural, electronic and magnetic properties of MoS₂ nanoribbons [J]. *Journal of Materials Chemistry*, 2012, 22 (15) : 7280-7290.
- [82] LI Y, ZHOU Z, ZHANG S, et al. MoS₂ nanoribbons: High stability and unusual electronic and magnetic properties [J]. *Journal of the American Chemical Society*, 2008, 130 (49) : 16739-16744.
- [83] BOLLINGER M V, LAURITSEN J V, JACOBSEN K W, et al. One-dimensional metallic edge states in MoS₂ [J]. *Physical Review Letters*, 2001, 87 (19) : 196803-196804.
- [84] BYSKOV L S, NØRSKOV J K, CLAUSEN B S, et al. DFT calculations of unpromoted and promoted MoS₂-based hydrodesulfurization catalysts [J]. *Journal of Catalysis*, 1999, 187 (1) : 109-122.
- [85] LI T, GALLI G. Electronic properties of MoS₂ nanoparticles [J]. *The Journal of Physical Chemistry C*, 2007, 111 (44) : 16192-16196.
- [86] ATACA C, SAHIN H, AKTURK E, et al. Mechanical and electronic properties of MoS₂ nanoribbons and their defects [J]. *The Journal of Physical Chemistry C*, 2011, 115 (10) : 3934-3941.
- [87] KARTHIKEYAN J, KUMAR V, MURUGAN P. Atomic structure and edge magnetism in MoS_{2+x} parallelogram shaped platelets [J]. *Physical Chemistry Chemical Physics*, 2013, 15 (31) : 13077-13082.
- [88] DOLUI K, PEMMARAJU C D, SANVITO S. Electric field effects on armchair MoS₂ nanoribbons [J]. *ACS Nano*, 2012, 6 (6) : 4823-4834.
- [89] PRINS R, DE BEER V H J, SOMORJAI G A. Structure and function of the catalyst and the promoter in Co-Mo hydrodesulfurization catalysts [J]. *Catalysis Reviews—Science and Engineering*, 1989, 31 (1-2) : 1-41.
- [90] LAURITSEN J V, NYBERG M, VANG R T, et al. Chemistry of one-dimensional metallic edge states in MoS₂ nanoclusters [J]. *Nanotechnology*, 2003, 14 (3) : 385-389.
- [91] CLAUSEN B S, LENGELER B, CANDIA R, et al. Exafs studies of calcined and sulfided CO-Mo HDS catalysts [J]. *Bulletin des Sociétés Chimiques Belges*, 1981, 90 (12) : 1249-1259.
- [92] PARHAM T G, MERRILL R P. An EXAFS study of the structure of supported cobalt molybdate catalysts as a function of sulfiding temperature [J]. *Journal of Catalysis*, 1984, 85 (2) : 295-310.
- [93] TAUSTER S J, PECORARO T A, CHIANELLI R R. Structure and properties of molybdenum sulfide: Correlation of O₂ chemisorption with hydrodesulfurization activity [J]. *Journal of Catalysis*, 1980, 63 (2) : 515-519.
- [94] CANDIA R, SØRENSEN O, VILLADSEN J Ø R, et al. Effect of sulfiding temperature on activity and structures of CO-MO/Al₂O₃ catalysts. II [J]. *Bulletin des Sociétés Chimiques Belges*, 1984, 93 (8-9) : 763-774.
- [95] DRESSELHAUS M S, THOMAS I L. Alternative energy technologies [J]. *Nature*, 2001, 414 (6861) : 332-337.
- [96] BONDE J, MOSES P G, JARAMILLO T F, et al. Hydrogen evolution on nano-particulate transition metal sulfides [J]. *Faraday Discussions*, 2009, 140 : 219-231.
- [97] VIELSTICH W, GASTEIGER H A, YOKOKAWA H. Handbook of fuel cells: Advances in electrocatalysis, materials, diagnostics and durability [M]. [S. l.] : Wiley-Blackwell, 2009 : 403-419.
- [98] JARAMILLO T F, JØRGENSEN K P, BONDE J, et al. Identification of active edge sites for electrochemical H₂ evolution from MoS₂ nanocatalysts [J]. *Science*, 2007, 317 (5834) : 100-102.
- [99] JARAMILLO T F, BONDE J, ZHANG J, et al. Hydrogen evolution on supported incomplete cubane-type [Mo₃S₄]₄₊ electrocatalysts [J]. *The Journal of Physical Chemistry C*, 2008, 112 (45) : 17492-17498.
- [100] LI Y, WANG H, XIE L, et al. MoS₂ nanoparticles grown on graphene: An advanced catalyst for the hydrogen evolution reaction [J]. *Journal of the American Chemical Society*, 2011, 133 (19) : 7296-7299.
- [101] TSAI C, CHAN K, ABILD-PEDERSEN F, et al. Active edge sites in MoSe₂ and WSe₂ catalysts for the hydrogen evolution reaction: A density functional study [J]. *Physical Chemistry Chemical Physics*, 2014, 16 (26) : 13156-13164.
- [102] XU C, PENG S, TAN C, et al. Ultrathin S-doped MoSe₂ nanosheets for efficient hydrogen evolution [J]. *Journal of Materials Chemistry A*, 2014, 2 (16) : 5597-5601.
- [103] VOIRY D, YAMAGUCHI H, LI J, et al. Enhanced catalytic activity in strained chemically exfoliated WS₂ nanosheets for hydrogen evolution [J]. *Nature Materials*, 2013, 12 (9) : 850-855.
- [104] MA Y, DAI Y, GUO M, et al. Evidence of the existence of magnetism in pristine VX₂ monolayers (X = S, Se) and their strain-induced tunable magnetic properties [J]. *ACS Nano*, 2012, 6 (2) : 1695-1701.

- [105] ZHOU Y, YANG P, ZU H, et al. Electronic structures and magnetic properties of MoS₂ nanostructures: Atomic defects, nanoholes, nanodots and antidots [J]. *Physical Chemistry Chemical Physics*, 2013, 15(25):10385-10394.
- [106] MATHEW S, GOPINADHAN K, CHAN T K, et al. Magnetism in MoS₂ induced by proton irradiation [J]. *Applied Physics Letters*, 2012, 101(10):102103.
- [107] MURUGAN P, KUMAR V, KAWAZO Y, et al. Atomic structures and magnetism in small MoS₂ and WS₂ clusters [J]. *Physical Review A*, 2005, 71(6):362-368.
- [108] ZHANG J, SOON J M, LOH K P, et al. Magnetic molybdenum disulfide nanosheet films [J]. *Nano Letters*, 2007, 7(8):2370-2376.
- [109] LI H, QI X, WU J, et al. Investigation of MoS₂ and graphene nanosheets by magnetic force microscopy [J]. *ACS Nano*, 2013, 7(3):2842-2849.
- [110] TONGAY S, VAMOOSFADERANI S S, APPLETON B R, et al. Magnetic properties of MoS₂: Existence of ferromagnetism [J]. *Applied Physics Letters*, 2012, 101(12):123105.
- [111] BOTELLO-MÉNDEZ A R, LOPEZ-URIAS F, TERRONES M, et al. Metallic and ferromagnetic edges in molybdenum disulfide nanoribbons [J]. *Nanotechnology*, 2009, 20(32):14582-14582.
- [112] LU P, WU X, GUO W, et al. Strain-dependent electronic and magnetic properties of MoS₂ monolayer, bilayer, nanoribbons and nanotubes [J]. *Physical Chemistry Chemical Physics*, 2012, 14(37):13035-13040.
- [113] KOU L, TANG C, ZHANG Y, et al. Tuning magnetism and electronic phase transitions by strain and electric field in zigzag MoS₂ nanoribbons [J]. *The Journal of Physical Chemistry Letters*, 2012, 3(20):2934-2941.
- [114] VOJVODIC A, HINNEMANN B, NØRSKOV J K. Magnetic edge states in MoS₂ characterized using density-functional theory [J]. *Physical Review B*, 2009, 80(12):125416.
- [115] HU T, ZHOU J, DONG J, et al. Electronic and magnetic properties of armchair MoS₂ nanoribbons under both external strain and electric field, studied by first principles calculations [J]. *Journal of Applied Physics*, 2014, 116(6):064301.
- [116] XIA B, GAO D, LIU P, et al. Zigzag-edge related ferromagnetism in MoSe₂ nanoflakes [J]. *Physical Chemistry Chemical Physics*, 2015, 17(48):32505-32510.

作者简介:

张千帆 男,博士,副教授。主要研究方向:第一性原理材料计算及能源材料、拓扑绝缘体和层状结构材料。

E-mail: qianfan@buaa.edu.cn

Two-dimensional layered materials of transition metal dichalcogenides

ZHANG Qianfan*, GAO Lei, TIAN Hongzhen, XU Zhongfei, WANG Yapeng

(School of Materials Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: Transition metal dichalcogenides (TMDCs) nanomaterials have become a focus of research for their unique properties. Various remarkable properties of TMDCs materials provide possibility for their applications in nanoelectronics, optoelectronics and spintronics devices. This paper mainly introduces the recent advances of TMDCs in different fields with various excellent properties, and summarizes the preparation method, properties and nanostructure morphology of layered TMDCs materials. The edge states of layered TMDCs materials have important effect on the properties, which leads to superior properties different from bulk materials. In this paper, we emphasize the effects of edge states in low dimensional nanomaterials on the properties (chemical activity, electronic property and magnetic property) and the newest progress in edge states. Especially, the magnetic property in edges of TMDCs materials provides the opportunity for their potential usage in the field of magnetism and spintronics, which have aroused great interest of researchers. This paper summarizes the research achievements in TMDCs materials from theoretical and experimental perspectives, which would be supportive for the development of two-dimensional layered TMDCs materials.

Key words: transition metal dichalcogenides (TMDCs); layered materials; nanostructure; edge states; magnetic property

Received: 2016-01-26; **Accepted:** 2016-03-04; **Published online:** 2016-03-10 15:31

URL: www.cnki.net/kcms/detail/11.2625.V.20160310.1531.001.html

Foundation items: National Natural Science Foundation of China (11404017); Doctoral Fund of New Teachers (20131102120001); Preferential Subsidy of Overseas Students' Science and Technology Activities from Ministry of Human and Social Resources; Program for New Century Excellent Talents in University (NCET-12-0033)

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

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0470

航空高动态网络链路感知 OLSR 路由算法



马林华^{1,2,*}, 张嵩¹, 茹乐¹, 徐扬², 田雨³, 于云龙¹

(1. 空军工程大学 航空航天工程学院, 西安 710038;

2. 西安电子科技大学 综合业务网理论与关键技术国家重点实验室, 西安 710071;

3. 中国人民解放军 95876 部队, 张掖 734100)

摘 要: 针对航空高动态无人机(UAV)网络环境中节点移动速度快、网络拓扑变化快, 导致网络链路稳定性差、数据到达率低和信息拥塞度高等问题, 提出了一种航空高动态网络链路感知 OLSR(OLSR-LA)路由算法, 该算法利用接收的 2 个连续 Hello 消息的多普勒频移、能量等信号特征, 计算出航空高动态无人机网络中 2 个相邻节点的相对速度和移动趋势, 从而得出这 2 个节点之间链路的保持时间。根据节点 MAC 层接口队列长度衡量网络局部的负载程度, 并利用 ARIMA-WNN 组合预测模型预测下一时刻节点负载的预测值, 并通过 Hello 消息传递给邻居节点。根据链路感知情况, 采用基于局部路由负载均衡(RRLB)算法避免拥塞的发生。仿真结果表明, 与传统 OLSR 算法相比, 本文提出的算法有效提高了分组交付率, 降低了端到端的传输延时, 增加了网络吞吐量, 从而提高了整个无人机网络传输的有效性和实时性。

关键词: 航空高动态网络; 链路感知; 链路的保持时间; 负载均衡; 路由算法

中图分类号: TP393

文献标识码: A

文章编号: 1001-5965(2016)07-1326-09

航空高动态无人机网络(Airborne Highly Dynamic UAV Network, AHDUAVN)是移动 Ad Hoc 网络在无人机航空通信领域^[1]的应用。在执行任务中, 航空高动态无人机网络中的多架无人机之间存在较高的相对移动, 这会造成网络拓扑快速变化, 传统的路由算法一般不能很好地适应无人机网络的高动态性^[2]; 并且在网络承载较多数据流量^[3]的情况下, 某些节点链路的负载较高, 甚至发生拥塞, 这样都会造成丢包的现象, 端到端的延时也会随之增加, 进而影响整个无人机网络的通信效能。因此, 如何设计高可靠、低时延的航空高动态无人机路由算法是该领域的重要研究课题之一。

为了克服节点高速移动对路由的影响, 研究人员已提出了不少方法来对链路进行预测^[4]。

文献[5]提出了一种简单高效的链路稳定性预测模型, 其核心思想是计算所接收数据分组信号强度的方差来预测所关注链路的稳定性强弱, 但该算法并不能很好地适应航空高动态无人机快速拓扑变化的场景。文献[6]提出一种基于控制分组多路径多普勒路由(Multipath Doppler Routing, MUDOR)算法, 将链路预测的思想应用到航空通信网络中, 以分组多普勒频移^[7]作为衡量飞机间链路质量的指标, 通过选择多普勒值最小的链路构建路由。文献[8]提出了基于移动预测的 Ad Hoc 网络分簇算法, 利用多普勒频移计算相对速度, 并利用接收信号强度估计链路保持时间, 然后选择与所有邻居节点的移动特性相似度最大的节点作为簇首, 但该算法并没有考虑负载过重时信

收稿日期: 2015-07-14; 录用日期: 2015-09-06; 网络出版时间: 2015-11-16 15:00

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151116.1500.004.html

基金项目: 国家自然科学基金(61372167); 航空科学基金(20130596008); 综合业务网理论与关键技术国家重点实验室开放研究课题(ISNH15-13)

* 通讯作者: Tel.: 029-84787700 E-mail: land_max@126.com

引用格式: 马林华, 张嵩, 茹乐, 等. 航空高动态网络链路感知 OLSR 路由算法[J]. 北京航空航天大学学报, 2016, 42(7): 1326-1334. MA L H, ZHANG S, RU L, et al. Link awareness-based OLSR routing algorithm for airborne highly dynamic networks[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1326-1334 (in Chinese).

息拥塞对航空通信网络的影响。

针对高速数据传输带来的拥塞问题,文献[9]将 MAC 层接口队列长度作为路径的负载程度值,通过搜集路径所经历的所有中间节点的接口队列长度信息来选择负载最小的路径。针对航空高动态网络,文献[10]提出了一种具有负载感知特性的路由算法,该算法同时考虑了节点的相对移动速度和数据拥塞情况,给出了新的动态路由因子度量来适应拓扑结构的变化。但以上 2 种算法都是以上一个时刻的负载情况作为参考,没有考虑到网络负载的动态变化问题。

针对航空高动态无人机网络环境中链路稳定性差、数据到达率低和信息拥塞度高等问题,本文提出了航空高动态网络链路感知 OLSR (OLSR-LA) 路由算法。在传统的 OLSR 算法^[11]的基础上,引入了链路稳定性预测、负载预测机制,利用接收到的 2 个连续 Hello 消息的多普勒频移、能量等信号特征,计算出航空高动态无人机网络中 2 个相邻节点的相对速度和移动趋势,从而得出该 2 个节点间链路的保持时间,并且在链路处于不可靠状态时及时进行局部路径修复策略,有效减少了不可靠链路对数据传输的影响。通过交互邻居节点 MAC 层接口队列长度来衡量网络局部的负载程度,并通过 ARIMA-WNN 组合预测模型^[12]来预测邻居节点下一时刻负载的预测值。当网络中业务量增大时,根据感知路径的负载情况,采用基于局部路由负载均衡 (Regional Routing Load-Balancing, RRLB) 算法来避免拥塞的发生,有效提高了网络吞吐量,降低了网络传输的端到端的延时。

1 网络模型及链路感知机制

1.1 网络模型

航空高动态无人机网络可以表示为无向图 $G(V, E)$, V 为网络拓扑 G 中的移动节点集, E 为节点间可双向通信的链路集。假设网络中所有飞行器节点都处于同一高度上,从而将网络仿真场景从三维近似为二维;每个无人机节点配置相同的全向天线,且具有相同的传输半径 R 。

1.2 链路稳定性预测

根据自由空间衰减模型^[8]可知,接收功率与收发天线之间的距离关系为

$$P_R = \frac{P_T G_T G_R \lambda^2}{(4\pi d)^2} \quad (1)$$

式中: P_T 为发射功率; P_R 为接收功率; d 为收发天线间距离; G_T 为发射天线增益; G_R 为接收天线增

益; λ 为载波波长。

图 1 所示为无人机节点 n_i 和 n_j 的相对移动模型。当 n_j 从 B 移动 C 时,称其为“远离”场景;当 n_j 从 C 移动 B 时,称其为“接近”场景。 n_j 在 B 和 C 处分别发送一次 Hello 消息,时间间隔为 Δt ,设 Δt 时间内节点之间相对运动的速度大小和方向保持不变。 f 为载波的原频率, f' 为 n_i 接收到的载波频率。根据多普勒效应,可得在“远离”场景中,节点 n_i 接收到 C 处发送的 Hello 消息载波频率可以计算为

$$f' = \frac{V_c - v_p \cos \angle ACB}{V_c} f \quad (2)$$

式中: V_c 为电磁波的传输速度; v_p 为“远离”场景中的相对移动速度,即

$$v_p = \frac{V_c (f - f')}{f \cos \angle ACB} \quad (3)$$

由余弦定理可得

$$\cos \angle ACB = \frac{d_{AC}^2 + d_{BC}^2 - d_{AB}^2}{2d_{AC}d_{BC}} \quad (4)$$

式中: d_{AB} 、 d_{AC} 可由式(1)得到。

设 Δt 时间内节点之间相对运动的速度大小和方向保持不变,则 B 和 C 之间的距离可以表示为

$$d_{BC} = v_p \Delta t \quad (5)$$

根据式(1)、式(3)~式(5),可得“远离”场景中的平均相对移动速度为

$$v_p = \sqrt{\frac{\lambda V_c (f - f')}{2\pi f \Delta t} \sqrt{\frac{P_T G_T G_R}{P_R^C} + \frac{\lambda^2 P_T G_T G_R}{16\pi^2 (\Delta t)^2} \left(\frac{1}{P_R^B} - \frac{1}{P_R^C} \right)}} \quad (6)$$

式中: P_R^B 、 P_R^C 分别为发送节点 n_j 位于 B 、 C 时,节点 n_i 接收信号的接收功率。

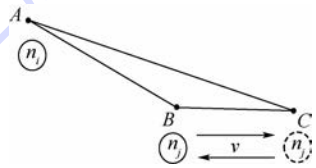


图 1 相对移动示意图

Fig. 1 Sketch map of relative mobility

同理可得,“接近”场景中的平均相对移动速度为

$$v_q = \sqrt{\frac{\lambda V_c (f - f')}{2\pi f \Delta t} \sqrt{\frac{P_T G_T G_R}{P_R^B} + \frac{\lambda^2 P_T G_T G_R}{16\pi^2 (\Delta t)^2} \left(\frac{1}{P_R^C} - \frac{1}{P_R^B} \right)}} \quad (7)$$

如图 2(a) 所示,在“远离”场景中, n_i 和 n_j 之间链路最长保持时间为

$$t_p = \frac{d_{CD}}{v_p} = \frac{d_{BD} - d_{BC}}{v_p} \quad (8)$$

在△ABC中,通过余弦定理可得

$$\cos \angle ABC = \frac{d_{AB}^2 + d_{BC}^2 - d_{AC}^2}{2d_{AB}d_{BC}} \quad (9)$$

根据式(1)、式(5)和式(6),可得

$$\cos \angle ABC = \frac{\frac{4\pi V_c(f-f')\Delta t}{f\lambda\sqrt{P_R^c P_T G_T G_R}} + \frac{1}{P_R^b} - \frac{1}{P_R^c}}{\sqrt{\frac{1}{P_R^b}} \sqrt{\frac{8\pi V_c(f-f')\Delta t}{f\lambda\sqrt{P_R^c P_T G_T G_R}} + \frac{1}{P_R^b} - \frac{1}{P_R^c}}} = \frac{M_1 + N_1}{\sqrt{\frac{1}{P_R^b}} \sqrt{2M_1 + N_1}} \quad (10)$$

式中: $M_1 = \frac{4\pi V_c(f-f')\Delta t}{f\lambda\sqrt{P_R^c P_T G_T G_R}}$; $N_1 = \frac{1}{P_R^b} - \frac{1}{P_R^c}$ 。

在△ABD中,通过余弦定理可得

$$\cos \angle ABD = \frac{d_{AB}^2 + d_{BD}^2 - d_{AD}^2}{2d_{AB}d_{BD}} = \frac{d_{AB}^2 + d_{BD}^2 - R^2}{2d_{AB}d_{BD}} \quad (11)$$

将式(11)求解可得

$$d_{BD} = d_{AB} \cos \angle ABD + \sqrt{R^2 - d_{AB}^2 + d_{AB}^2 (\cos \angle ABD)^2} \quad (12)$$

因为∠ABD和∠ABC是相等的,所以根据式(6)、式(10)和式(12)可得,“远离”场景中, n_i 和 n_j 链路的最长保持时间为

$$t_p = \frac{d_{CD}}{v_p} = \frac{d_{BD}}{v_p} - \Delta t = \left[\frac{M_1 + N_1}{\sqrt{2M_1 + N_1}} + \sqrt{\frac{16\pi^2 R^2}{\lambda^2 P_T G_T G_R} + \frac{(M_1 + N_1)^2}{2M_1 + N_1} - \frac{1}{P_R^b}} - 1 \right] \Delta t \quad (13)$$

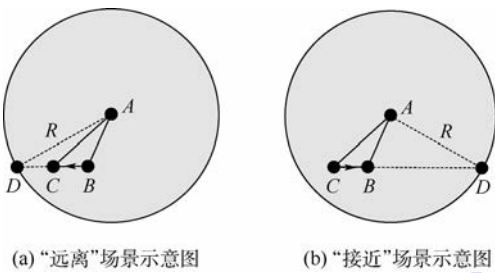


图2 2种相对运动趋势示意图

Fig.2 Sketch map of two relative mobility directions

同理可得,“接近”场景中, n_i 和 n_j 链路的最长保持时间为

$$t_q = \frac{d_{BD}}{v_q} = \frac{d_{CD}}{v_q} - \Delta t = \left[\frac{M_2 - N_1}{\sqrt{2M_2 - N_1}} + \sqrt{\frac{16\pi^2 R^2}{\lambda^2 P_T G_T G_R} + \frac{(M_2 - N_1)^2}{2M_2 - N_1} - \frac{1}{P_R^c}} - 1 \right] \Delta t \quad (14)$$

$$\text{式中: } M_2 = \frac{4\pi V_c(f-f')\Delta t}{f\lambda\sqrt{P_R^b P_T G_T G_R}}。$$

1.3 链路负载预测

由于航空高动态无人机网络业务具有突发性、随机性和规律性等特点,网络中链路负载的时间序列是非线性、非平稳和多时间尺度动态变化的,所以网络中每个因素与链路的负载函数关系很难用准确的数学模型进行逼近。本文采用基于ARIMA-WNN组合预测模型^[12]预测链路负载的动态变化情况,该组合预测模型不仅可以通过ARIMA模型有效捕捉到负载时间序列间的相关性,还能通过WNN模型来弥补ARIMA模型无法刻画的非线性关系,从而对负载时间序列做出精确的预测。

1.3.1 负载时间序列模型

节点 n_i 在每次发送Hello消息时,对其MAC层接口的队列长度进行采样,经过相同间隔的多次采样,按时间顺序进行排列就可得到一组负载观测值的时间序列集 $q_i(k)$ 。设 $q_i(l)$ 表示节点 n_i 第 l 次发送Hello消息时MAC层接口的队列长度的观测值,节点 n_i 的MAC层接口的最大队列长度为 $q_{\max}(i)$,则定义节点 n_i 第 l 次发送Hello消息时的负载拥塞程度函数为

$$f_i(l) = \frac{q_i(l)}{q_{\max}(i)} \quad (15)$$

1.3.2 ARIMA模型

ARIMA模型是由Box和Jenkins^[13]提出的一种时间序列建模方法,其基本思想是:对非平稳的时间序列通过 z 阶逐级差分使其成为平稳序列,再利用ARMA模型对该平稳序列建模,经逆变换得到预测序列。理论上,ARIMA模型的数学描述为

$$\begin{cases} \varphi(B)\nabla^r X_t = \theta(B)\varepsilon_t \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma_\varepsilon^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ E(X_s \varepsilon_t) = 0, \forall s < t \end{cases} \quad (16)$$

式中: $\varphi(B) = 1 - \sum_{i=1}^p \varphi_i B_i$; $\theta(B) = 1 - \sum_{i=1}^q \theta_i B_i$; $\nabla = 1 - B$ 为差分算子; $|B| \leq 1$ 为后移算子,即 $B_i X_t = X_{t-i}$,且 $\varphi(B)$ 与 $\theta(B)$ 互质, $\varphi_p \theta_q \neq 0$; X_t 为样本值; $\{\varepsilon_t\}$ 为白噪声序列,服从均值为0、方差为常量 σ_ε^2 的正态分布; p 为自回归阶数; q 为滑动平均阶数; φ_i 和 θ_i 为待估计参数。具体建模参数求解过程参考文献[12]。

1.3.3 WNN模型

网络负载具有非线性、多时间尺度变化的特

点。小波变换具有多分辨分析能力,而神经网络具有很强的自学习、自适应和非线性逼近能力,因此 WNN 模型就是将两者有效地结合起来进行预测的一种预测模型。WNN 模型的隐节点函数是小波函数,而输入层到隐含层的权值和阈值分别对应小波伸缩和平移参数。本文采用的 WNN 模型包括输入层、隐含层和输出层 3 层网络,输出层采用线性输出。

图 3 为 WNN 模型的拓扑结构。输入层有 m 个神经元,隐含层有 n 个神经元,输出层有 1 个神经元。 $X_k, X_{k+1}, \dots, X_{k-1+m}$ 为 WNN 模型的输入参数, \hat{X}_{k+l} 为输出第 $k+l$ 个时间序列的预测值, w_{ij} 为输入层与隐含层的连接权值, w_{jl} 为隐含层与输出层的连接权值。

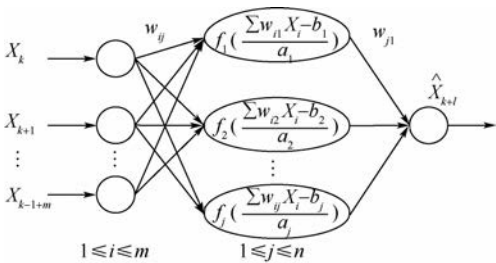


图 3 WNN 模型的拓扑结构

Fig. 3 Topology structure of WNN model

当输入为 $X_k, X_{k+1}, \dots, X_{k-1+m}$ 时,隐含层输出为

$$f(j) = f_j \left(\frac{\sum_{i=k}^{k-1+m} w_{ij} X_i - b_j}{a_j} \right) \quad (17)$$

$j = k, k + 1, \dots, k - 1 + m$

式中: $f(j)$ 为隐含层第 j 个节点的输出值; b_j 为小波基函数的平移参数; a_j 为小波基函数的伸缩参数; f_j 为小波基函数。

本文所采用的小波基函数为 Morlet 母小波基函数,其表达式为

$$y = \cos(1.75x) e^{-\frac{x^2}{2}} \quad (18)$$

则输出为

$$\hat{X}_{k+l} = \sum_{j=1}^n w_{jl} f(j) \quad (19)$$

1.3.4 ARIMA-WNN 组合预测模型

本文采用的 ARIMA-WNN 组合预测模型将一组负载时间序列数据看成是由线性自相关部分 L_t 与非线性部分 U_t 组成。因此,该组合预测模型预测过程分为三步:①利用 ARIMA 模型对负载时间序列的线性部分建模,得到预测结果 \hat{L}_t ,则该预测结果与原负载时间序列的残差隐含了该负载时间序列的非线性特征。②将这部分残差代入 WNN 模型,通过不断修正神经元之间的连接权值

和阈值来逼近负载时间序列的非线性特性,并进行预测,得到预测结果 \hat{U}_t 。③将两部分预测结果相加,即得到下一时刻节点的负载预测值。图 4 为网络中某个节点负载的相同时间间隔预测值,仿真显示其预测值与真实值的相似度可达到 89%。

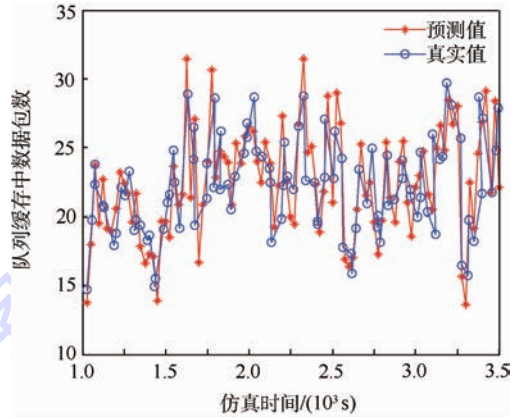


图 4 ARIMA-WNN 组合预测模型的负载预测值

Fig. 4 Load predicted values of ARIMA-WNN combination forecasting model

2 OLSR-LA 路由算法

OLSR-LA 路由算法在传统 OLSR 算法基础上,引入了链路稳定性预测和负载预测机制。

图 5 为 OLSR-LA 路由算法流程。可知,OLSR-LA 路由算法主要由三部分组成:链路稳定性预测模块、负载预测交互模块和路由建立与维护模块。其中,路由建立与维护模块是将链路稳定性预测和负载预测机制有效加入路由策略的核心。

2.1 链路稳定性预测机制

OLSR-LA 路由算法与传统 OLSR 算法一样,网络中每个节点仍然采用定期发送 Hello 消息、TC 消息和 MID 消息来广播自己的状态,并获取邻居节点的状态。OLSR-LA 路由算法中,任意节点收到来自邻居节点的 Hello 消息后,首先提取 Hello 消息中邻居节点的负载情况和信号的特征信息,然后通过接收该节点连续发送的 2 个 Hello 消息的多普勒频移、接收信号强度值等信号特征,利用式(13)和式(14)推算出该 2 个节点间链路有效保持时间,保存到节点的邻居表和拓扑表中,以此作为 MPR 选择、MPR 信息分发^[11]和及时删除失效链路节点的依据。每个节点在传统 OLSR 算法的 MPR 选择原则基础上加入链路稳定性预测机制来得到自己的链路状态,保持良好的相邻节点集、二跳邻居节点集、MPR 集和 MPR 选择集后,再通过 TC 消息将自己的携带有链路状态保持良好的 MPR 集泛洪到整个网络。

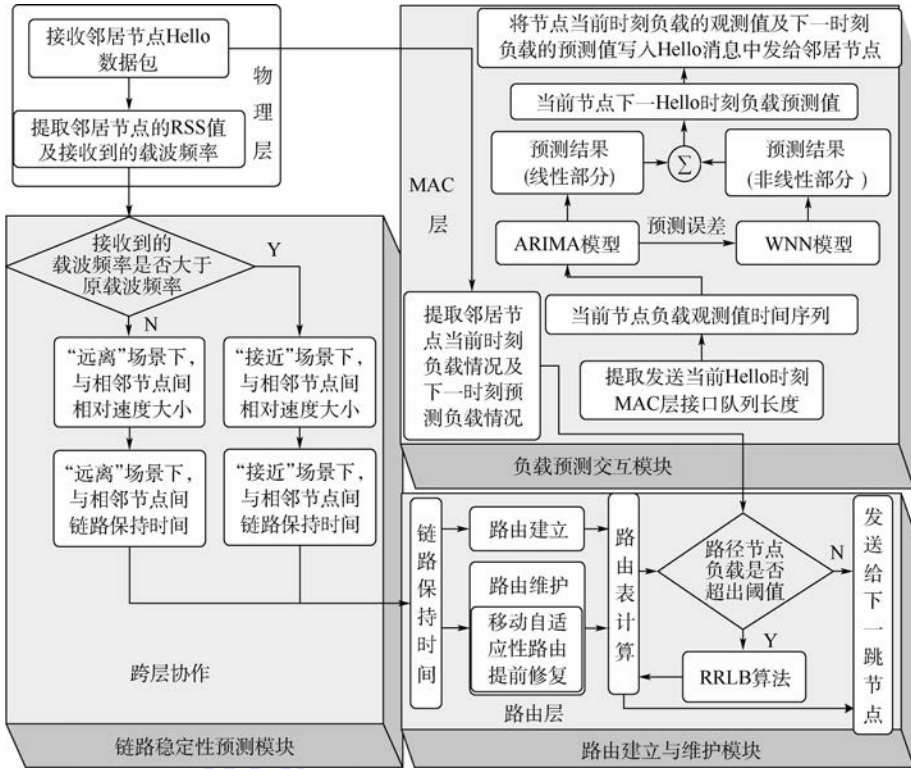


图5 OLSR-LA 路由算法流程

Fig.5 Flowchart of OLSR-LA routing algorithm

2.2 链路负载预测机制

OLSR-LA 路由算法运行时,每个节点以 Hello 消息的间隔不断地对自己 MAC 层接口队列长度进行采样,提取自身节点负载情况,从而构成负载观测值的一个时间序列集,然后通过第 1.3.4 节的 ARIMA-WNN 组合预测模型对本节点下一时刻的 MAC 层接口队列长度进行预测,并将当前时刻负载的观测值和下一时刻预测负载的预测值写入 Hello 消息中发送给邻居节点,因此邻居节点间通过 Hello 消息交换,每个节点都可获得两跳邻居节点相应的负载情况。

2.3 路由建立与维护优化

节点根据接收到 TC 消息中的拓扑表的全局拓扑信息和自己邻居表中的局部拓扑信息,采用 Dijkstra 算法计算出链路状态保持路由的路径,将计算出的路由存储在自己的路由表中。当节点检测到路由表中任意一条链路保持时间低于危险阈值 T 时,则认为该链路进入危险期,删除该路由路径上的失效节点,再根据更新的拓扑表和邻居表来提前重新计算路由的路径,从而减轻拓扑变化对路径稳定性的影响,其中,危险阈值 T 设为 $30\text{ms}^{[14]}$ 。

在传统 OLSR 算法基础上引入链路稳定性预

测机制,可以提前预测出即将失效的路径,从而有效减轻了航空高动态无人机网络中拓扑快速变化对链路稳定性的影响,然而随着网络业务量的增加,这种机制建立的链路状态良好的最短路径并不是最优路径。当一条链路稳定性很好的路径上的某个节点即将或者已经发生拥塞时,只能选择丢弃发送给它的数据包,即使网络中其他路径还是空闲的,因此,OLSR-LA 路由算法在建立链路状态良好的最短路径后进一步引入负载预测机制。当一条链路稳定性良好的路径上的某个节点要给该路径上的下一跳节点发送数据时,先探测到下一跳节点负载拥塞程度函数是否达到危险阈值,若有,则此节点启动 RRLB 算法;若没有,则直接沿着该条链路稳定性良好的路径传输,直到发送到目的节点。

2.4 RRLB 算法

为了更好地避免局部拥塞的发生^[15],平衡网络业务的分布,尽可能降低拥塞带来的丢包率,本文提出了一种 RRLB 算法。该算法是分布式进行的,能使业务在传输过程中主动绕过潜在的拥塞链路而到达目的节点。根据算法设计需要,RRLB 算法扩展了 OLSR 协议原有的路由表结构。表 1 为扩展后的路由表,保存了两跳的路由节点。

表 1 OLSR-LA 路由算法路由表结构
Table 1 Routing table structure of OLSR-LA routing algorithm

序号	地址	地址	地址	跳数	本地接口地址
(1)	R_dest_addr	R_next_1_addr	R_next_2_addr	R_dist	R_iface_addr
(2)	R_dest_addr	R_next_1_addr	R_next_2_addr	R_dist	R_iface_addr
(3)

注:R_next_1_addr 为下一跳节点的接口地址;R_next_2_addr 为下第二跳节点的接口地址;其他项目与 OLSR 原协议相同。

不失一般性,考虑网络中一条链路稳定性很好的路径上的某个节点为节点 i 。节点 i 路由表中保存的下一跳节点为节点 j ,节点 j 下一跳节点为节点 k 。节点 i 通过 Hello 消息获得节点 j 下一时刻负载拥塞程度函数大于 $2/3$,则认为其可能拥塞时,启动 RRLB 算法,具体如下:

- 1) while (节点 i 预测到下一跳节点 j 可能发生拥塞)
- 2) 搜索所有邻居节点(一跳和二跳)中是否存在空闲或者低负载的链路稳定的邻居节点(其负载拥塞程度函数小于 $1/3$)。
- 3) If (存在这样的邻居节点)
- 4) If (这样的邻居节点中存在节点 i 和节点 k 共同的邻居节点)
- 5) 执行策略 1:节点 i 选择链路稳定最优的节点 h 替代节点 j 作为下一跳节点。
- 6) Else if (这样的邻居节点中存在一个节点是节点 i 的一跳邻居节点和一个节点是节点 k 的一跳邻居节点,而这 2 个节点也互为邻居节点)
- 7) 执行策略 2:节点 i 选择链路稳定最优的节点 p 和节点 q 替代节点 j 进行两跳转发(设节点 p 为节点 i 的一跳邻居节点,节点 q 为节点 k 的一跳邻居节点)。
- 8) Else 按原路径发送给下一跳节点 j
- 9) Else 按原路径发送给下一跳节点 j

图 6 为 RRLB 算法执行示意图。从源节点 S 到目的节点 D 的链路稳定性良好的最优路径为 $\{S, a, b, c, d, e, D\}$ (为了表述清晰,省去了网络中其他节点)。当节点 b 通过组合预测模型预测到其下一时刻可能发生拥塞时,通过 Hello 消息发送给节点 a 。此时,节点 a 启动 RRLB 算法,将过多的业务负载分担给低负载的节点。如图 6 所示,节点 a 搜索发现其邻居节点中存在节点 a 和节点 c 共同的合适邻居节点 i ,节点 a 执行策略 1,将其局部路径更改为 $\{a, i, c\}$ 。当节点 c 通过来自节点 d 的 Hello 消息感知到节点 d 可能发生拥

塞,节点 c 启动 RRLB 算法搜索发现其邻居节点中不存在节点 c 和节点 e 共同的合适邻居节点,但存在一个低负载节点 j 是节点 c 的一跳邻居节点和一个低负载节点 k 是节点 e 的一跳邻居节点,而这 2 个节点也互为邻居节点,节点 a 执行策略 2,将其局部路径更改为 $\{c, j, k, e\}$ 。

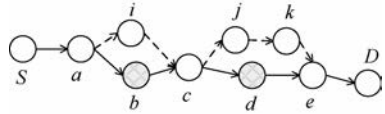


图 6 RRLB 算法执行示意图

Fig.6 Sketch map of RRLB algorithm

值得注意的是,在 RRLB 算法运行过程中,OLSR-LA 路由算法将过多的业务调整到空闲或者低负载的节点中,有效减少了网络拥塞的可能性;又因为 RRLB 算法是分布式运行的,不需要任何中心控制机制,也不会带来额外的网络开销;并尽可能减小为负载均衡需求所带来的路径增加。因此,RRLB 算法对航空高动态无人机网络具有很好的适应性和扩展性。

3 仿真分析

本文利用 NS-3.21 仿真软件对 OLSR-LA 路由算法与传统 OLSR 算法及采用基于多普勒预测策略^[8]改进的 OLSR(OLSR-DOP)算法的性能进行了仿真比较。仿真中,80 架无人机随机分布在 $100 \text{ km} \times 100 \text{ km}$ 的区域内,采用 Friis 传输模型^[1,7]和 Gauss-Markov 移动模型^[16]。每架无人机的发射功率为 $31 \text{ dB} \cdot \text{m}$,接收机灵敏度设为 $-90 \text{ dB} \cdot \text{m}$,通信半径均为 25 km ,载波频率为 800 MHz 。MAC 层采用 IEEE 802.11 DCF,传输速率为 2 Mb/s ,通信模型采用 CBR 数据源,共有 15 个通信对,每个节点的数据包产生速率分别为 2、4、6、8、10、12 packet/s,数据包分组大小为 512 B ,节点 MAC 层缓存中最大队列长度为 40 个,仿真时间为 60 min 。根据不同的场景种子,随机生成 10 个运动场景,最终结果取 10 次仿真结果的平均值。

比较不同路由算法在不同移动速度对分组交付率与延时的影响。设每个发送节点的数据包产生速率恒定为 4 packet/s 。

图 7 为不同移动速度区间下分组交付率性能。可知,3 种路由算法的分组交付率随移动速度的增加而下降,其中,传统 OLSR 算法下降速度最快,OLSR-LA 路由算法下降速度最慢,OLSR-DOP 算法介于两者之间,这种差异性随着速度的增加而更加明显。这是由于高速移动的无人机网

络环境中,随着速度的增加,网络中链路的中断情况加剧,从而导致大量的数据包发送失败。而 OLSR-DOP 算法与 OLSR-LA 路由算法引入了链路稳定性预测机制,可以有效预测不可靠链路的影响,及时修复路由,但 OLSR-DOP 算法计算精度不如 OLSR-LA 路由算法。OLSR-LA 路由算法还考虑了负载对链路的影响,从而有效避免了拥塞节点的发生,进一步减少丢包,有效提高了数据传输的成功率。

图 8 为网络传输延时随节点移动速度变化的性能。可以看出,相比于传统 OLSR 算法和 OLSR-DOP 算法,OLSR-LA 路由算法能有效减少网络传输时端到端的延时,这主要是由于 OLSR-LA 路由算法采用路由提前修复机制,尽可能减少不可靠链路对网络传输的影响,同时引入了分布式的 RRLB 算法,减少网络传输时拥塞发生的概率。

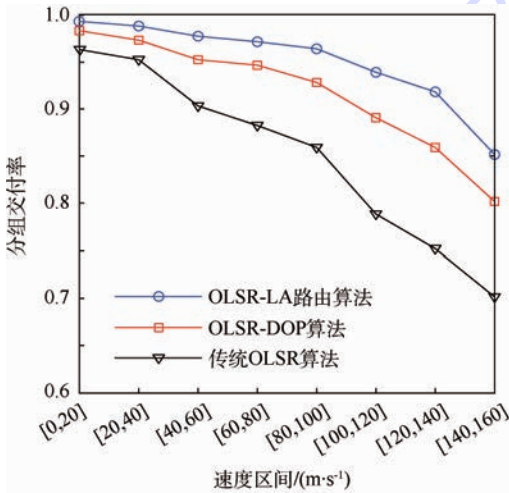


图 7 不同移动速度区间下分组交付率性能

Fig. 7 Performance of packet delivery rate with different mobility velocities

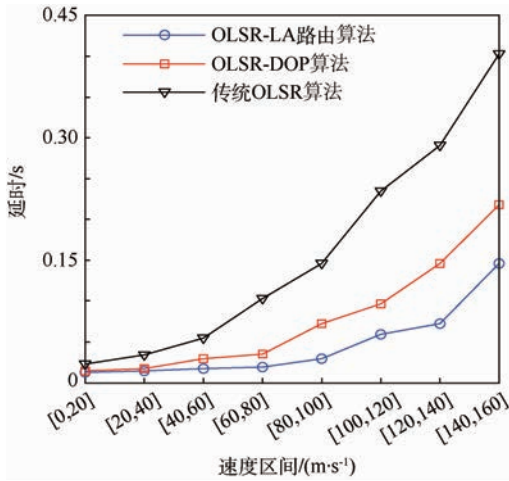


图 8 不同移动速度区间下端到端传输延时性能

Fig. 8 Performance of end-to-end transmission latency with different mobility velocities

图 9 为每个节点移动速度区间为 [40, 60] m/s 时,不同业务量下网络的吞吐量性能。仿真结果显示,OLSR-LA 路由算法相比于传统 OLSR 算法和 OLSR-DOP 算法获得了更好的网络吞吐量。随着业务量的增大,OLSR-LA 路由算法相比于其他 2 种算法网络吞吐量的性能优势是先增大后减小。这是因为当业务负载较小时,网络中很少有节点出现缓存溢出的情况,所以 RRLB 算法基本未启用,OLSR-LA 路由算法和 OLSR-DOP 算法相比于传统 OLSR 算法网络的增加主要是由链路稳定性预测机制带来的,而 OLSR-DOP 算法的预测精度不如 OLSR-LA 路由算法,所以在低负载情况时网络吞吐量稍逊 OLSR-LA 路由算法。随着业务量的增加,网络中拥塞出现的概率增大,RRLB 算法的优势就体现得越来越明显。当业务量增加到一定程度时,OLSR-LA 路由算法相比于其他 2 种算法的差距开始缩小,这是因为网络中所有节点负载都几乎饱和,所以 OLSR-LA 路由算法不能有效地发现空闲或者低负载的节点来有效地均衡负载了。

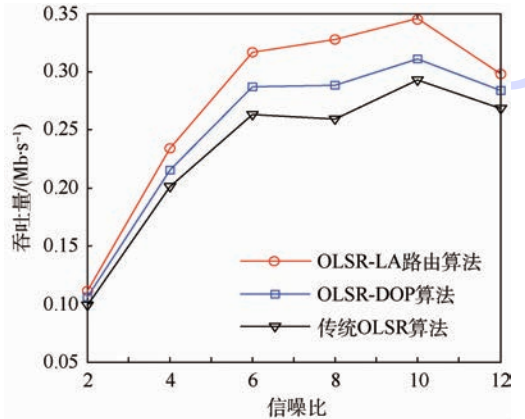


图 9 不同业务量下网络吞吐量性能

Fig. 9 Performance of network throughput with different traffic values

图 10 为每个节点移动速度区间为 [40, 60] m/s 时,不同业务量下网络的端到端传输延时性能。仿真结果显示,随着业务量的增加,OLSR-LA 路由算法和 OLSR-DOP 算法相比于传统 OLSR 算法能有效减少网络传输时端到端的延时,这主要是因为 2 种算法都尽可能减少了不可靠链路对数据传输的影响。在业务负载较小时,OLSR-LA 路由算法相比于其他 2 种算法优势并不明显;随着业务量的增加,OLSR-LA 路由算法可以通过 RRLB 算法有效降低拥塞的发生,减小网络传输的延时。当业务量增加到一定程度时,OLSR-LA 路由算法相比于 OLSR-DOP 算法的差距开始缩小。当业务产生速率恒定为 8 packet/s 时,OLSR-LA 路由算法在延时方面的优势最为明显。

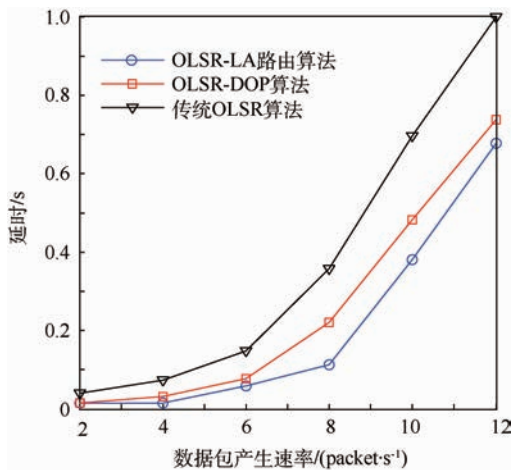


图10 不同业务量下端到端传输延时性能

Fig. 10 Performance of end-to-end transmission latency with different traffic values

4 结论

针对航空高动态无人机网络环境中链路稳定性差、数据到达率低和信息拥塞度高等问题,在传统 OLSR 算法基础上,提出航空高动态网络链路感知 OLSR (OLSR-LA) 路由算法,仿真结果表明:

1) 提出的 OLSR-LA 路由算法与传统 OLSR 算法相比,可以有效提高分组交付率,降低端到端的传输延时,增加网络的吞吐量,从而增强了整个无人机网络传输的有效性和实时性。

2) OLSR-LA 路由算法引入了链路稳定性预测机制,在链路处于不可靠状态时及时进行局部路径修复策略,有效减少不可靠链路对数据传输的影响。

3) 当网络业务量增大时,根据 ARIMA-WNN 组合预测模型预测邻居节点下一时刻负载的预测情况,提出的分布式 RRLB 算法能有效地避免拥塞的发生。

参考文献 (References)

- [1] ROHRER J P, JABBAR A, CETINKAYA E K, et al. Highly-dynamic cross-layered aeronautical network architecture [J]. IEEE Transactions on Aerospace and Electronic Systems, 2011, 47(4): 2742-2765.
- [2] ERTURK M, HAQUE J, ARSLAN H. Challenges of aeronautical data networks [C] // Proceedings of IEEE Aerospace Conference. Piscataway, NJ: IEEE Press, 2010: 1-7.
- [3] 赵璇, 蒋铃鹤, 何晨. 一种基于最小负载的 Ad Hoc 移动网络路由协议 [J]. 上海交通大学学报, 2004, 38(5): 684-687. ZHAO X, JINAG L G, HE C. A least-loaded dynamic routing protocol Ad Hoc networks [J]. Journal of Shanghai Jiao Tong University, 2004, 38(5): 684-687 (in Chinese).
- [4] BIRADAR R, MANVI S, REDDY M. Link atability based multicast routing scheme in MANET [J]. Computer Networks, 2010, 54(7): 1183-1196.
- [5] 夏辉, 贾智平, 张志勇, 等. 移动 Ad Hoc 网络中基于链路稳定性预测的组播路由协议 [J]. 计算机学报, 2013, 36(5): 926-936. XIA H, JIA Z P, ZHANG Z Y, et al. A link stability prediction-based multicast routing protocol in mobile Ad Hoc networks [J]. Chinese Journal of Computer, 2013, 36(5): 926-936 (in Chinese).
- [6] SAKHAE E, JAMALIPOUR A, KATO N. Aeronautical Ad Hoc networks [C] // IEEE Wireless Communications and Networking Conference. Piscataway, NJ: IEEE Press, 2006: 246-251.
- [7] SAKHAE E, TALEB T, JAMALIPOUR A. A novel scheme to reduce control overhead and increase link duraion in highly mobility Ad Hoc networks [C] // IEEE Wireless Communications and Networking Conference. Piscataway, NJ: IEEE Press, 2007: 3972-3977.
- [8] 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.
- [9] LEE S J, GERLA M. Dynamic load-aware routing in Ad Hoc networks [C] // Proceedings of IEEE International Conference on Communications. Piscataway, NJ: IEEE Press, 2001: 3206-3210.
- [10] 刘智, 徐桢. 航空高动态网络负载感知路由算法 [J]. 北京航空航天大学学报, 2014, 40(12): 1697-1701. LIU Z, XU Z. Geographic load aware routing algorithm for highly dynamic airborne networks [J]. Journal of Beijing University of Aeronautics and Astronautics, 2014, 40(12): 1697-1701 (in Chinese).
- [11] WANG A, ZHU B. Improving MPR selection algorithm in OLSR protocol based on node localization technology [J]. Journal of Networks, 2014, 9(7): 1674-1681.
- [12] 谭满春, 冯萃斌, 徐建闽. 基于 ARIMA 与人工神经网络组合模型的交通流预测 [J]. 中国公路学报, 2007, 20(4): 118-121. TAN M C, FENG L B, XU J M. Traffic flow prediction based on hybrid ARIMA and ANN model [J]. China Journal of Highway and Transport, 2007, 20(4): 118-121 (in Chinese).
- [13] BOX G E P, JENKINS G M. 时间序列分析: 预测与控制 [M]. 顾岚, 译. 北京: 中国统计出版社, 1997: 101-149. BOX G E P, JENKINS G M. Time series analysis: Forecasting and control [M]. GU L, translated. Beijing: China Statistics Press, 1997: 101-149 (in Chinese).
- [14] 王炫, 李建东, 张文柱. 支持多速率传输的动态 Ad Hoc 路由协议 [J]. 电子与信息学报, 2006, 28(10): 1907-1911. WANG X, LI J D, ZHANG W Z. A novel routing protocol for multi-rate moblie Ad Hoc networks [J]. Journal of Electronics and Information Technology, 2006, 28(10): 1907-1911 (in Chinese).
- [15] 王传臣, 张润彤, 李丹丹, 等. 认知网络中的拥塞规避多径路由算法 [J]. 电子学报, 2011, 39(8): 1858-1863. WANG C C, ZHANG R T, LI D D, et al. Multiple-path routing algorithm with congestion avoidance in cognitive networks [J]. Acta Electronica Sinica, 2011, 39(8): 1858-1863 (in Chinese).

Chinese).

- [16] BROYLES D, JABBAR A, STERBENZ J P, et al. Design and analysis of a 3-D Gauss-Markov mobility model for highly-dynamic airborne networks[C]//Proceedings of International Telemetering Conference. Las Vegas; ISA, 2009:1-10.

作者简介:

马林华 男,博士,教授,博士生导师。主要研究方向:移动自

组织网络协同通信、信道编码。

Tel.: 029-84787700

E-mail: land_max@126.com

张嵩 男,博士研究生。主要研究方向:移动自组织网络、协同通信。

Tel.: 029-84787700

E-mail: zhangsong1949@163.com

Link awareness-based OLSR routing algorithm for airborne highly dynamic networks

MA Linhua^{1,2,*}, ZHANG Song¹, RU Le¹, XU Yang², TIAN Yu³, YU Yunlong¹

(1. School of Aeronautics and Astronautics Engineering, Air Force Engineering University, Xi'an 710038, China;

2. The State Key Laboratory of Intergrated Service Networks, Xidian University, Xi'an 710071, China;

3. Unit 95876 of the People's Liberation Army, Zhangye 734100, China)

Abstract: Due to the high mobility of the unmanned aerial vehicle (UAV) node, quick changes of the network topology structure, the airborne highly dynamic UAV network suffers some problems such as poor stability of the network link, low data delivery ratio and high data congestion information. In order to overcome these problems, a link awareness-based OLSR (OLSR-LA) routing algorithm for airborne highly dynamic networks is proposed. The characteristics of two received consecutive Hello messages, such as the Doppler shift and the power strength of received Hello messages, can be used to obtain the relative speed and direction of motion between the two adjacent nodes in the airborne highly dynamic UAV network. Then the link connection lifetime is estimated by the relative speed. The OLSR-LA routing algorithm uses the queue length in the buffer of the MAC layer to indicate the local load level. Then the predicted value of the node's load level in the next time can be predicted by the ARIMA-WNN combination forecasting model and passed to neighbor by Hello message. Finally, according to the conditions of link awareness, regional routing load-balancing (RRLB) algorithm is employed to avoid network congestion. The simulation results show that compared to the traditional OLSR algorithm, the proposed OLSR-LA routing algorithm can effectively improve the packet delivery rate, increase the traffic of network, reduce the end-to-end transmission latency and enhance the real-time and effectiveness of the data transmission in the whole UAV networks.

Key words: airborne highly dynamic network; link awareness; link connection lifetime; load-balancing; routing algorithm

Received: 2015-07-14; Accepted: 2015-09-06; Published online: 2015-11-16 15:00

URL: www.cnki.net/kcms/detail/11.2625.V.20151116.1500.004.html

Foundation items: National Natural Science Foundation of China (61372167); Aeronautical Science Foundation of China (20130596008); The State Key Laboratory Open Foundation of Intergrated Service Networks (ISNI15-13)

* Corresponding author. Tel.: 029-84787700 E-mail: land_max@126.com

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0482

基于网络编码的无线 NoC 多播映射



陈亦欧*, 谭芳, 凌翔

(电子科技大学 通信抗干扰技术国家级重点实验室, 成都 611731)

摘 要: 具有高带宽、QoS 保障和广播等优点的无线片上网络 (NoC) 能为通信信号系统的多核并行处理提供平台, 而网络编码能够以较小功耗开销为代价显著提高 NoC 的传输效率。本文从通信信号系统的多播特性与需求出发, 为无线 NoC 建立合适的架构与映射模型, 在多约束条件下设计满足网络吞吐率性能最大化及核间通信延时、数据传输和编码功耗等开销最小化的任务映射与网络编码技术联合优化方法, 实现了任务的高效并行与多处理核的协同工作。实验表明, 本文提出的网络架构及映射算法与传统方式相比能节约 6% 的功耗, 得到 16% 的吞吐率增益, 并且路由节点硬件实现复杂度低。

关键词: 无线片上网络; 网络编码; 映射; 多播; 低功耗

中图分类号: TP393

文献标识码: A

文章编号: 1001-5965(2016)07-1335-09

随着并行计算与并行处理的发展, 片上网络 (Network on Chip, NoC) 上承载的应用包含大量的多播业务, 传统的数据传输方式已不能满足需求。另外, 随着片上系统 (System on Chip, SoC) 规模的增大, 采用传统的金属连线的网络传输效率低, 且固定网络拓扑很难对各种不同的应用算法及多播业务进行网络传输的弹性适应和调整。为了解决这些问题, 业界开发了毫米波天线和射频互联波导 (RF-I)^[1-4], 为芯片内或芯片间 (板级) 的高速数据交互提供了技术支撑。在此基础上提出的无线 NoC 既具有高速传输的特点, 又因节点互联采用虚拟拓扑, 可灵活地进行改变以适应不同应用系统的需求。然而, 无线信道的数据传输仍不可避免地受到信道带宽和容量的约束。由于无线 NoC 具有天然的广播特性, 为了提高网络的传输效率, 本文引入无线通信中热门的网络编码技术, 一方面显著提高网络吞吐率, 降低由网络拥塞带来的额外通信开销; 另一方面, 通过数据包的

编码合并减少传输数据的总量, 进而降低网络传输功耗。在多核阵列平台采用网络编码技术还有一个显著优点, 即线性网络编码在芯片上的实现逻辑非常简单。

在 NoC 上实现高效多播最首要的任务是进行多播应用的任务映射。现已有很多科研人员对有线 NoC 的任务映射进行了研究, 提出了利用蚁群算法、遗传算法、离散粒子群算法和 NAGA-II 等对延时或功耗进行优化的映射策略。传统的多播数据通过多播树 (Multicast Tree, MT) 传播, 已有的多播树构造算法包括最短路径多播树的 Dijkstra 算法^[5]、Bellman-Ford 算法^[6]、最小代价多播树的 KMB 算法^[7] 和 MPH 算法^[8] 等。然而由于多播树自身结构的局限性, 无法实现多播传输的最大容量。Ahlswede 等^[9] 用网络编码技术对网络中间节点接收到的信息进行“融合”处理, 从而使网络的数据传输速率逼近多播容量的极限。张思为^[10] 提出了一种基于局部总线的 NoC 架构,

收稿日期: 2015-07-17; 录用日期: 2015-09-30; 网络出版时间: 2016-01-04 10:03

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20160104.1003.002.html

基金项目: 中央高校基本科研业务费专项资金 (ZYGX2014J011); 国家重大科技专项 (2013ZX03003015); 国家“863”计划 (2015AA01A701)

* 通讯作者: Tel.: 028-61830326 E-mail: chenyou@uestc.edu.cn

引用格式: 陈亦欧, 谭芳, 凌翔. 基于网络编码的无线 NoC 多播映射[J]. 北京航空航天大学学报, 2016, 42(7): 1335-1343. CHEN Y O, TAN F, LING X. Multicast wireless NoC mapping based on network coding[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1335-1343 (in Chinese).

首次将网络编码引入 NoC。Shalaby^[11] 和 Vonbun^[12] 等验证了在 NoC 上采用蝶形网络编码的设计用于路由跳数和数据包存活数上的性能优势。然而,蝶形结构是网络编码中最简单的一种结构,只适用于多播业务信源和信宿都为 2 的情况,应用范围受限,故缺乏通用性。为了解决此问题,本文从网络编码的原理出发,提出了一种适用于多播业务的无线 NoC 平台,利用网络编码构造与任务映射的相对独立性,设计出采用网络编码的无线 NoC 的两步映射算法,以降低延时、减小功耗与提高网络传输效率为目标,建立了基于网络编码的无线 NoC 资源映射模型与网络编码映射模型。仿真结果表明,在多种多播业务模式下,与普通无线 NoC 相比,本文提出的网络架构及映射算法能够至少节约 6% 的功耗,并得到 16% 的吞吐率增益,同时具有较小的复杂度开销。

1 网络编码原理

假定多播网络 $G(V, E)$ 的信源产生的信息用符号向量 $\mathbf{x} = (x_1, x_2, \dots, x_h)$ 表示, h 为多播的最大理论传输容量,信宿接收的信息由符号向量 $\mathbf{y} = (y_1, y_2, \dots, y_h)$ 表示。对于线性网络编码,其链路传输的信息、网络节点接收与发送的信息均为信源信息 \mathbf{x} 中各元素的线性组合,即 $y_i = q_{i1}x_1 + q_{i2}x_2 + \dots + q_{ih}x_h = \mathbf{f}_{e_i} \cdot \mathbf{x}^T, i = 1, 2, \dots, h$, \mathbf{f}_{e_i} 为信宿的第 i 条接收边 e_i 上的全局编码向量, q_{ij} 为转移变量,用矩阵表示,可得

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_h \end{bmatrix} = \begin{bmatrix} q_{11} & q_{12} & \dots & q_{1h} \\ q_{21} & q_{22} & \dots & q_{2h} \\ \vdots & \vdots & \ddots & \vdots \\ q_{h1} & q_{h2} & \dots & q_{hh} \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_h \end{bmatrix} = \mathbf{Q}\mathbf{x}^T \quad (1)$$

式中: \mathbf{Q} 为转移矩阵,当其可逆,信宿可译码得出信源信息 \mathbf{x} , 即 $\mathbf{x} = (\mathbf{Q}^{-1}\mathbf{y}^T)^T$ 。

为了实现简单,采用线性网络编码,在中间节点与信宿节点进行的编码操作被简化为对接收到的数据进行模二加运算。模二加运算是数字电路中最基本的异或运算,在集成电路里实现非常简单,而信宿节点收到多条入边送来的数据包后,交给其处理节点进行译码操作,还原为信源信息,不会增加路由节点的复杂度,因此在无线 NoC 上应用网络编码的实现复杂度非常低。

2 无线 NoC 平台

2.1 互联模型

无线 NoC 中,每个节点上放置一个低功耗

UWB 天线,通过无线射频接口将数据高速广播到相邻节点。天线的广播范围决定了无线 NoC 的拓扑。当节点以网格形式放置,并且天线的传输距离为网格中相邻 2 个节点之间的距离时,该网络的拓扑为有线 NoC 中最常用的 Mesh 拓扑。本文以 Mesh 拓扑为例进行研究。不失一般性,该平台设计与映射算法也可以应用在其他 NoC 拓扑上。图 1 为无线 Mesh NoC 互联结构示意图。每个无线路由节点 (R) 与一个处理节点 (PE) 相连,路由节点间采用双向数据传递。

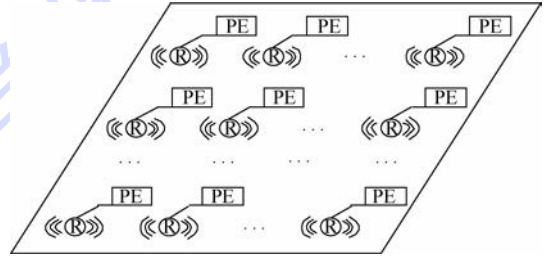


图 1 无线 Mesh NoC 互联结构示意图

Fig. 1 Schematic of wireless Mesh NoC interconnection architecture

2.2 路由机制

与有线 NoC 类似,数据路由与网络拓扑有关,例如 Mesh 拓扑可采用 xy 路由,数据包从源到目的先沿着 x 轴方向传输,再沿着 y 轴方向传输。具有网络编码功能的节点必然需要额外的逻辑资源进行编码,面积和功耗相比也会较大。出于复杂度考虑,在 NoC 上会采用较简单的网络编码结构,其对应的编码节点个数较少,因此仅需要为网络中部分路由节点增加编码操作逻辑。

综上所述,网络中的无线路由节点分为 2 类:编码节点(进行网络编码的节点)和非编码节点(不进行网络编码的节点)。编码节点对接收到的来自多条链路的数据包进行网络编码操作,然后广播转发;而非编码节点直接将接收到的数据进行广播转发。为了保证数据快速地从源扩散到目的,并且避免死锁,采用最短路径算法,即数据包在 x 轴或 y 轴方向上只能沿着从源节点到目的节点的方向进行传输,而不能进行反向传输。

2.3 信道复用

无线 NoC 上,数据以局部广播的方式进行传播,某节点发送数据,其所有相邻节点都会收到。为了能让广播范围重叠的相邻节点同时传输数据,以达到在节点对不同路径到来的数据进行编

码的目的,片上无线链路采用信道复用的方式,即用 TDMA、FDMA、CDMA 或它们的组合方式对冲突链路进行区分。这里以 FDMA 为例,为了避免冲突,传输域交叠的节点采用不同的载波发射频率,即与任一个节点相邻的各节点必须采用不同载波频率。

节点载波频率的分配问题是一个典型的边(点)着色问题。根据文献[13]的着色算法可确定载波个数 χ 的取值范围为

$$\Delta(G) \leq \chi \leq \Delta(G) + 1 \quad (2)$$

式中: $\Delta(G)$ 为图 G 的顶点的最大度。

本文主要研究基于网络编码的映射问题,故暂不考虑信道仲裁和分配过程,即假设信道是逻辑独立的。

2.4 网络编码策略

数据在网络的传输过程中,不同节点送来的数据包由于经过的路径不同,以及网络拥塞等原因,到达同一节点的时间可能会不一致。而进行网络编码的路由节点必须在收到所有路径送来的数据包后才能进行编码操作。因此,当路由节点收到一条路径送来的数据后,需要将数据保存在缓存中,等待其他路径的数据到来后一起编码。若各条路径的数据到达时间差较大,则路由节点需要较大的缓存空间,对 NoC 有限的存储空间提出挑战。为了解决此问题,本文采用缓存感知的网络编码策略。网络的最优缓存大小为 B_0 ^[14],如果多条路径的数据没有同时到达,则将先到的数据保存在缓存里,等待其他路径的数据都到达后进行编码发送。若等待的时间超过门限值,则先到达的数据无需进行网络编码,直接排队等待传输。这样既节约存储资源,又不会因为数据缓存等待造成额外的传输延时。

3 无线 NoC 映射过程

基于网络编码的无线 NoC 映射过程可分为 2 个步骤:①任务映射,即将各项任务分配到 NoC 的处理节点上;②编码映射,即确定网络编码系数以及编码节点在无线 NoC 拓扑上的位置。本节介绍任务流图(Task Graph, TG)、资源流图(Resource Graph, RG)和多播图等基本概念以及 NoC 映射过程。

3.1 任务流图

定义 1 任务流图是由节点 $u_i (\forall u_i \in U, i = 1, 2, \dots, |U|)$ 和有向边 $c_k = c_{i,j} (\forall c_k \in C, k = 1, 2, \dots, |C|)$ 构成的表征实时信号处理系统工作流

程的赋权图, $G_1 = G(U, C)$, 节点 U 表示系统的一个任务,有向边 C 表示任务间的通信数据流,分单播流和多播流 2 类。任务数为 $|U|$, 数据流总数为 $|C|$ 。节点的权值 $p_i (i = 1, 2, \dots, |U|)$ 表征节点 u_i 的处理负载(即运算时间),有向边的权值 $f_k = f_{i,j} (k = 1, 2, \dots, |C|)$ 表征节点 i 到节点 j 的数据流上的通信量。

3.2 资源流图

定义 2 资源流图是表示无线 NoC 互联信息的有向图, $G_2 = G(T, R)$, 节点数为 $|T|$, 边数为 $|R|$ 。节点 $t_i \in T (i = 1, 2, \dots, |T|)$ 表示无线 NoC 中的瓦片,包含路由节点和处理节点。有向边 $r_{i,j} \in R (i, j = 1, 2, \dots, |R|)$ 表示从瓦片 t_i 到瓦片 t_j 的单向数据链路。

3.3 多播图

定义 3 多播树是由任务流图中具有相同信宿的一组多播流构成的图。节点表示多播应用的信源或信宿,每棵多播树包含一个信源与多个信宿,边表示从多播信源到多系信宿的一条多播流,边的权值即多播流的通信流量。由任务流图所表征的系统中所有的多播树构成多播图。

3.4 任务映射和编码映射

任务映射是按一定的优化规则将任务流图中的各个任务分配到资源流图中的节点上,采用细颗粒度的任务划分,多个任务可同时分配到一个资源节点上。

编码映射是确定了多播信源和信宿在资源流图上的位置的基础上,按一定的优化规则,确定网络编码的结构以及编码节点在资源流图中的位置。每个路由节点最多分配一个编码节点,原因如下:若多个编码节点分配到同一个路由节点上,由于不同编码节点在网络编码结构上的位置不同,其链接链路与相邻节点也不同,因此流经这多个编码节点的数据流在经过不同的编码向量编码后,需要不同的链路发到不同的相邻编码节点去。此时,为了区分这几组不同的编码数据,需要在数据包里加入额外的信息进行标识,这会带来额外的逻辑开销。另外,会使得编码链路的映射变得复杂,不利于映射时的最优化选择。

映射过程中,如果选择的编码路径或编码节点不同,网络各条链路上传递的数据也不同,相应得到的网络通信性能与代价成本也会不同。编码节点与无线路由节点的映射是一一对应的关系。图 2 为基于网络编码的无线 NoC 映射示意图。

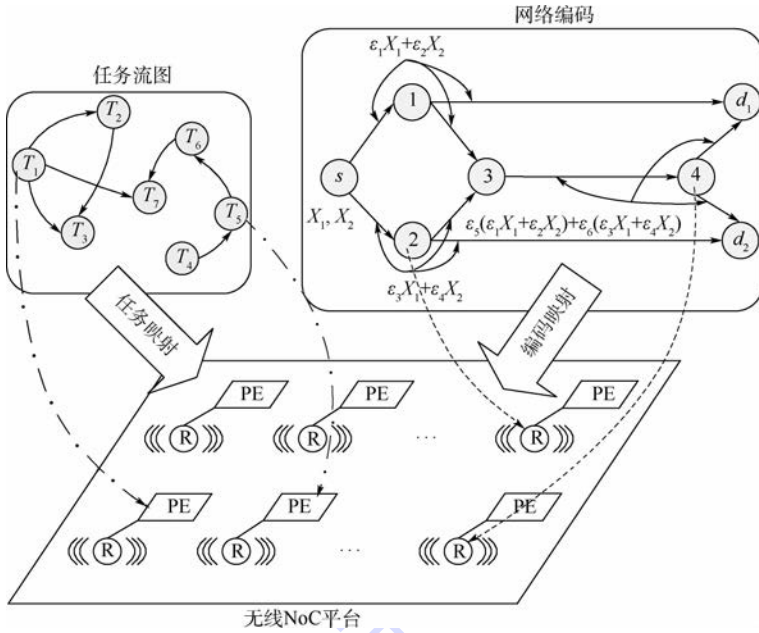


图2 基于网络编码的无线 NoC 映射示意图

Fig.2 Schematic of wireless NoC mapping based on network coding

4 问题描述和数学模型

4.1 优化目标

映射问题是一个多目标优化问题,涉及的优化目标包括低功耗、实时性与高网络吞吐率,下面分别对其进行分析。

1) 芯片功耗

低功耗设计是芯片设计中需要着重考虑的问题。芯片功耗过大,将造成芯片过热、工作不稳定以及供电等问题。无线 NoC 的功耗分为 2 个部分:处理功耗和网络功耗。处理功耗是所有处理节点运行的功耗之和,即 $P_{PE} = \sum_{k \in T} P_k$, $P_k (k=1, 2, \dots, |T|)$ 为处理节点 k 的功耗。处理节点的工作状态决定功耗的大小。网络功耗指网络中无线路由节点进行数据传输带来的功耗,包括节点的交换、仲裁与控制功耗、编码功耗和无线信道传输功耗。不同映射阶段考虑的功耗不同。

2) 实时性

数字信号处理系统有实时性要求,本文用响应时间来描述系统的实时性。响应时间指从输入到输出的最长路径延时,其数学表达式为

$$D_R = \max_{l_k \in L} \left(\sum_{u_i \in U_k} p_i + \sum_{e_m \in E_k} t_m \right) \quad (3)$$

式中:集合 $L = \{l_k | k=1, 2, \dots, |L|\}$ 包含任务流图中从输入节点到输出节点之间所有路径; U_k 为路径 l_k 上的任务节点集; $E_k = \{e_m\}$ 为 l_k 上的直连边集; $t_m = t_{i,j}$ 为任务 i 到任务 j 的通信时间。

3) 网络吞吐率

网络传输性能用网络吞吐率来衡量。为评估网络编码的效果,定义网络吞吐率为单位时间内网络中传递的多播信息量,其计算式为

$$O = \sum_{s \in S} \frac{f_s n_s}{T_M} \quad (4)$$

式中: $f_s (s \in S)$ 为多播图中源节点 s 到单个目的节点的通信量, S 为源节点集合; n_s 为源节点 s 对应的目的节点个数; T_M 为多播任务的传输时间。

4.2 任务映射

任务映射采用多目标优化遗传算法 NSGA-II 求解最优方案^[15-16],优化目标是功耗和响应时间同时最小化,即

$$\begin{cases} \min \left(\sum_{u_i \in U} P_i + \sum_{u_i \in U} \sum_{u_j \in U} f_{i,j} w_{i,j} P_0 \right) \\ \min \left[\max_{l_k \in L} \left(\sum_{u_i \in U_k} p_i + \sum_{u_i, u_j \in U_k} t_{i,j} \right) \right] \end{cases} \quad (5)$$

式中: P_0 为传输单比特数据一跳的功耗; $w_{i,j}$ 为从节点 i 到节点 j 的跳数; $f_{i,j}$ 为从节点 i 传到节点 j 的数据流量。

因为这一阶段还没有进行网络编码的设计,没有获得编码的信息,所以功耗和响应时间的计算里不考虑网络编码带来的影响。

从实际情况出发,对任意的任务映射 $TM: DFG \rightarrow RG$,本映射平台的约束条件如下:

$$\forall i, \sum_j X_{i,j} = 1, \quad u_i \in U, t_j \in T \quad (6)$$

$$\forall j, \sum_i X_{i,j} p_i \leq \frac{1}{A}, \quad u_i \in U, t_j \in T \quad (7)$$

$$\forall r, \sum_{i,j \in U} \sum_{m,n \in T} b_{i,j} h_r^{m,n} \leq B \quad (8)$$

$$T_{\infty} \leq \frac{1}{A} \quad (9)$$

式中: $X_{i,j}$ 为任务 i 与处理节点 j 的映射关系; $b_{i,j}$ 为节点 i 到节点 j 通信所需的最小带宽; $h_r^{m,n}$ 为链路 r 是否在从节点 i 到节点 j 的路由路径上; B 为链路带宽; T_{∞} 为有环网络的迭代边界; A 为数据注入率。

式(6)约束每个任务只能映射到一个处理节点上,式(7)保证分配到同一处理节点的任务的总运算时间满足数据注入率要求,式(8)为带宽约束,式(9)为迭代边界约束,有环网络的迭代边界 T_{∞} 不能超过数据注入率 A 的倒数。

4.3 编码映射

由第3.4节可知,编码映射是在已知多播图和任务映射结果的基础上,在资源流图上构造网络编码向量的过程。编码映射包含以下步骤:

1) 根据任务映射方案,将任务流图的多播图 G_1 映射到资源流图上得到资源流图的多播图 G_2 ,如图3(a)所示。

2) 去掉 G_2 中没有交叉关系的多播流,仅保留具有胶连关系的多播流,构成简化的多播图 G_3 ,如图3(b)所示。

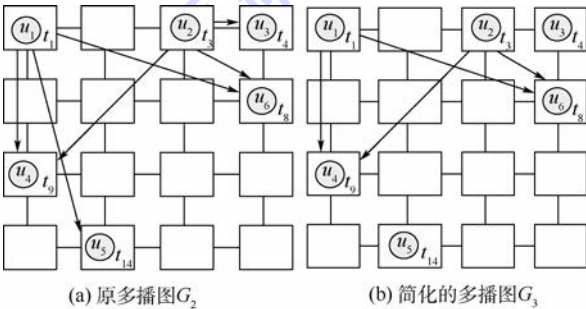


图3 简化多播图的实例

Fig.3 An example of simplified multicast graph

具有胶连关系的多播任务指具有相同信宿,并且相互之间有交叉的多播流。由此,多播流在传输路径上可能发生冲突。网络编码用来提高冲突链路的数据传输效率,如果数据传输没有冲突,则无法发挥效果,因此本文仅对有冲突的多播任务进行网络编码设计。图3给出了本步骤的一个简单实例。 G_2 中原本有2组多播流,一组的信源是瓦片 t_1 ,信宿是瓦片 t_8 、 t_9 和 t_{14} ;另一组的信源是 t_3 ,信宿是 t_4 、 t_8 和 t_9 。从 G_2 中可以看出,多播流 $t_3 \rightarrow t_4$ 是孤立的,与其他多播流都不相交;而多播流 $t_1 \rightarrow t_{14}$ 虽与 $t_3 \rightarrow t_9$ 有交叉关系,但没有其他流和它共享信宿,故剔除这2条多播流,得到简化

的多播图 G_3 。

3) 用 NSGA-II 选取最优多播路由路径,并对最优路径的各条边构造全局网络编码向量,确定编码节点位置。本文中网络拓扑是已知的,故采用多项式时间算法快速构造多播路径上各条链路的网络编码向量。该算法包括2个步骤:①为每个目的节点寻找网络中 h 条边不相交的路的集合 (h 为源节点个数);②为步骤①中所找的路构造全局编码向量。在最短路由算法下,步骤①的寻路过程若采用遍历搜索,搜索空间巨大,因此采用 NSGA-II 进行优化选路,同时将边不相交作为约束条件,使得优化结果满足多项式时间算法的要求。步骤②的全局网络编码向量构造算法的详细步骤可参见文献[17-18]。

编码映射的优化目标为最大化通信性能增益(网络吞吐率)和最小化编码代价(功耗),即

$$\begin{cases} \min(P_c + P_{tr}) \\ \max\left(\sum_{s \in S} \frac{f_s n_s}{T_M}\right) \end{cases} \quad (10)$$

式中: P_c 为网络编码的功耗; P_{tr} 为多播数据传输功耗。

由于编码操作是在无线路由节点上进行的,编码方案不会对处理节点的功耗造成影响,因此本文优化功耗只考虑网络编码的功耗 P_c 和多播数据传输功耗 P_{tr} 。

由于给定任务流图时,多播的通量和目的节点数是已知的,因此网络吞吐率的大小与多播任务的传输时间 T_M ,即各条多播流传输时间 $T_{s,d}$ 中的最大值相关。考虑网络编码的影响,多播流传输时满足网络编码的数据会编码后传输,其余未编码数据在规定的路由算法下依次传输。因而多播流传输时间为这两部分数据传输时间之和。编码数据和未编码数据的网络传输时间的数学表达式分别为

$$T_{NC_{s,d}} = T_{NC_{Tr_{s,d}}} + T_{C_{s,d}} + T_{NC_{W_{s,d}}} \quad (11)$$

$$T_{NonNC_{s,d}} = T_{NonNC_{Tr_{s,d}}} + T_{NonNC_{W_{s,d}}} \quad (12)$$

式中: $T_{NC_{Tr_{s,d}}}$ 和 $T_{NC_{W_{s,d}}}$ 分别为编码数据在无线信道上的传输时间和由信道冲突造成的排队等待时间; $T_{NonNC_{Tr_{s,d}}}$ 和 $T_{NonNC_{W_{s,d}}}$ 分别为未编码数据在无线信道上的传输时间和冲突等待时间; $T_{C_{s,d}}$ 为数据在路由节点进行网络编码的时间。

在无线 NoC 平台上,各参数的计算式如下:

$$T_{NC_{Tr_{s,d}}} = \frac{f_s g_{s,d}}{B} w_{s,d} \quad (13)$$

$$T_{NC_{W_{s,d}}} = \sum_{m \in M_{s,d}} \frac{f_{NC_m}}{B} \quad (14)$$

$$T_{\text{NonNC_Tr}_{s,d}} = \frac{f_s(1 - g_{s,d})}{B} w_{s,d} \quad (15)$$

$$T_{\text{NonNC_W}_{s,d}} = \sum_{m \in M_{s,d}} \frac{f_{\text{Non_NC}_m}}{B} \quad (16)$$

式中： $g_{s,d}$ 为多播流中编码数据的比例； $M_{s,d}$ 为路径上路由节点集合； f_{NC_m} 和 $f_{\text{Non_NC}_m}$ 分别为在路径上节点 m 处与本路径编码数据和未编码数据相冲突的其他路径数据量之和。

5 仿真性能评估

为了验证本文提出算法的性能,本节首先对具有多播业务的通信系统进行建模及性能分析;同时,评估了本文算法在获得性能提升时付出的

开销情况,包括路由器的实现复杂度问题与映射算法的执行复杂度问题。

5.1 性能分析

4×4 MIMO-OFDM 通信接收系统可抽象为图4所示的任务流图形式,包含32个任务,根据其功能不同,分为4类节点:A类节点(同步与解调)、B类节点(MIMO信道估计)、C类节点(MIMO迭代检测)和D类节点(LDPC译码)。其中,MIMO迭代检测器向LDPC译码器发送的是多播数据。任务节点的处理负载由各节点算法的乘加运算量等价得到,节点的处理负载及节点间的通信流数据^[19]详见表1和表2。

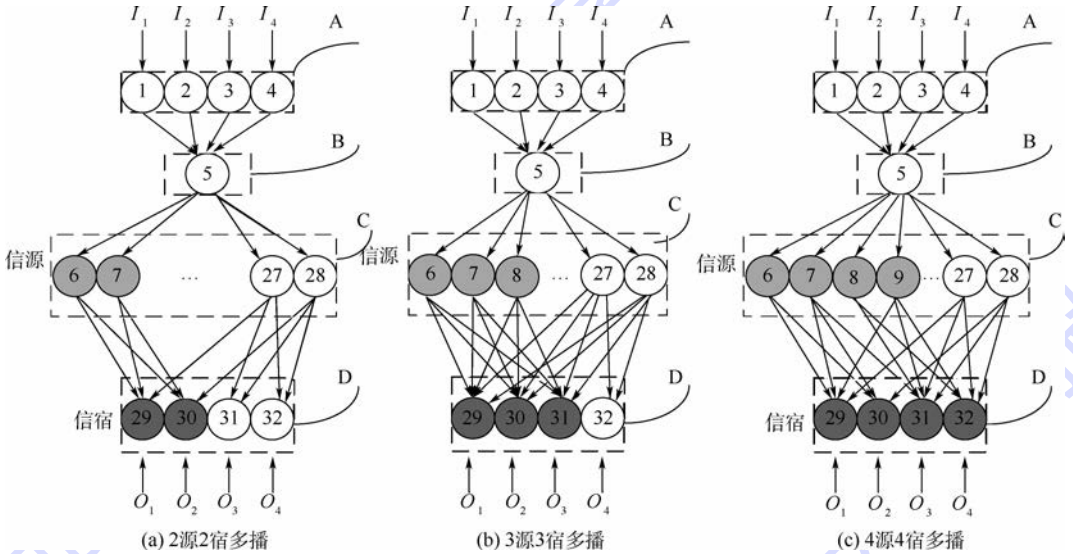


图4 MIMO-OFDM接收端任务流图

Fig.4 Task graph of MIMO-OFDM receiver

表1 任务节点的负载量

Table 1 Work Loads of each task

节点类型(节点标号)	处理负载/ms
A (1 ~ 4)	10.1
B (5)	21.4
C (6 ~ 28)	6.3
D (29 ~ 32)	13.2

表2 任务间的通信流量

Table 2 Traffic flows among tasks

通信流类型	链路	通信流量/(Mb · s ⁻¹)
单播	A→B	3.253 12
单播	B→C	1.689 60
单播	C→D	0.084 48
多播	C→D	0.084 48

由网络编码的定义可知,信宿能译码的充要条件是其输入链路上有 h 条边的全局编码向量线性无关。因此,在确定的网络拓扑下,信源个数的上界是确定的,即: $h \leq \max(\alpha_i)$, α_i 为信宿节点 i 的度, h 为信源个数。Mesh拓扑中节点的最大

出度和入度为4,在这个约束条件下,考察2源2宿(多播1)、3源3宿(多播2)和4源4宿(多播3)这3种多播应用,如图4所示。

由于缓存空间限制,当不同多播流到来时间差较大时,无法进行网络编码。定义如下参数:编码比例 $\alpha = f_c/f_0$ 为衡量编码效率的指标, f_c 为经过网络编码操作的多播流量, f_0 为多播总通信量。目标SoC平台采用32nm工艺的无线Mesh NoC结构,取 $B_{\text{max}} = 20 \text{ Gb/s}$,处理核满负载功率 $p_0 = 50 \text{ mW}$,线性负载功耗模型,片上无线传输每比特能耗 $e_0 = 2.7 \text{ pJ}$ ^[4]。遗传算法的主要参数设置如下:种群规模为100,代沟为0.9,交叉与变异概率分别为0.5和0.1,迭代100次。算法采用MATLAB编程,外加系统的遗传算法工具箱gatbx,仿真结果如下:

1) 将采用了网络编码的本文算法与未采用网络编码的现有的最短路径多播树算法进行比

较,图 5(a)给出了 5×5 Mesh(拓扑 1)下 2 种算法在网络吞吐率和功耗上的性能对比。图 5(b)给出了 10×10 Mesh(拓扑 2)下 2 种算法在网络吞吐率和功耗上的性能对比。为便于量化网络吞吐率和功耗的优化程度,图 5 中对数据进行了如下归一化处理:

$$O_{\text{normalized}} = \frac{O}{O_{\text{NC}}} \quad (17)$$

$$P_{\text{normalized}} = \frac{P_{\text{NC}}}{P} \quad (18)$$

式中: $O_{\text{normalized}}$ 为归一化吞吐率; O_{NC} 为采用网络编码 NoC 吞吐率; $P_{\text{normalized}}$ 为归一化功耗; P 为 NoC 功耗; P_{NC} 为采用网络编码 NoC 功耗。

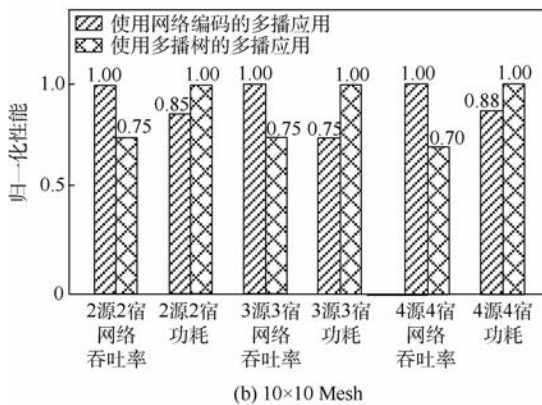
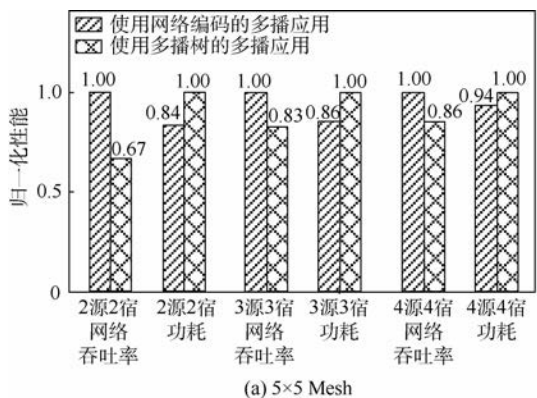


图 5 多播应用在 5×5 Mesh 与 10×10 Mesh 的性能对比
Fig. 5 Performance comparisons of multicast application on 5×5 Mesh and 10×10 Mesh

在 2 源 2 宿条件下,2 类拓扑下本文算法比现有的最短路径多播树算法的网络吞吐率分别提高了 49% 和 33%,功耗分别降低了 16% 和 15%;在 3 源 3 宿条件下,2 类拓扑下本文算法比现有的最短路径多播树算法的网络吞吐率分别提高了 20% 和 33%,功耗分别降低了 14% 和 25%;在 4 源 4 宿条件下,2 类拓扑下本文算法比现有的最短路径多播树算法的网络吞吐率分别提高了 16% 和 43%,功耗分别降低了 6% 和 12%。证明采用了网络编码的 NoC 可以获得性能的大幅提高。

在 3 源 3 宿,4 源 4 宿条件下,拓扑 2 比拓扑 1 的网络吞吐率分别提高了 13% 和 27%,功耗降低了 11% 和 6%。在 2 源 2 宿条件下,拓扑 2 比拓扑 1 的网络吞吐率降低了 16%,功耗提高了 1%。说明网络中存在多个多播业务应用时,网络规模的增大可进一步提高 NoC 的性能。原因在于当任务流中包含较多多播时,网络规模越大,编码映射路由路径的选择越多,更容易实现多个多播的网络编码,从而提高网络吞吐率并降低网络功耗。但当网络中多播均已实现网络编码时,网络规模的增大反而会增大数据的传输时间和传递次数,从而降低网络吞吐率并增加功耗。因此,网络规模必须和多播应用的规模适应,才能使 NoC 的性能达到最优。

2) 比较不同编码比例下的性能增益,结果如图 6 所示。可以看出,随着编码比例的增大,吞吐率增益与功耗节省呈准线性规律增长。因此,可采用并行调度技术使得到达网络编码节点的不同路径的数据尽可能同时到达,提高网络编码的比例,以获得最大的性能增益。其中,吞吐率增益和功耗节约计算公式如下:

$$O_{\text{gain}} = \frac{O_{\text{NC}} - O_{\text{NonNC}}}{O_{\text{NC}}} \quad (19)$$

$$P_{\text{save}} = \frac{P_{\text{NonNC}} - P_{\text{NC}}}{P_{\text{NonNC}}} \quad (20)$$

式中: O_{gain} 为吞吐率增益; O_{NonNC} 为未采用网络编码 NoC 吞吐率; P_{save} 为功耗的节约; P_{NonNC} 为未采用网络编码 NoC 功耗。

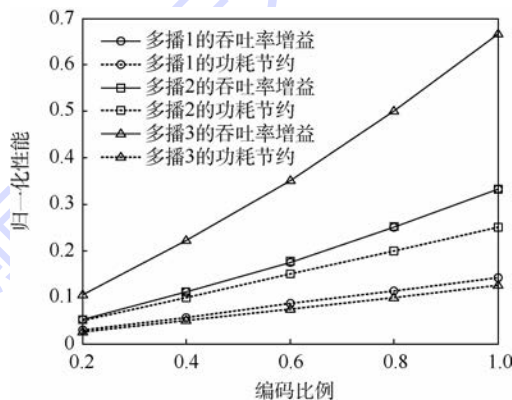


图 6 多播应用在不同编码比例下的性能增益
Fig. 6 Performance enhancement of multicast application under different encoding ratios

5.2 路由节点的实现复杂度分析

在对路由节点进行建模实现时,考虑到是否进行网络编码对射频与天线接入部分没有影响,故在建模时忽略这部分,仅对路由转发这部分功能进行建模与比较。基于网络编码的无线节点的交换模块的结构示意如图 7 所示,其中阴影部分

是编码功能模块,根据网络编码系数控制开关 $k_1 \sim k_5$ 的通断,将编码后的数据送到复用器,根据仲裁结果进行数据发送。

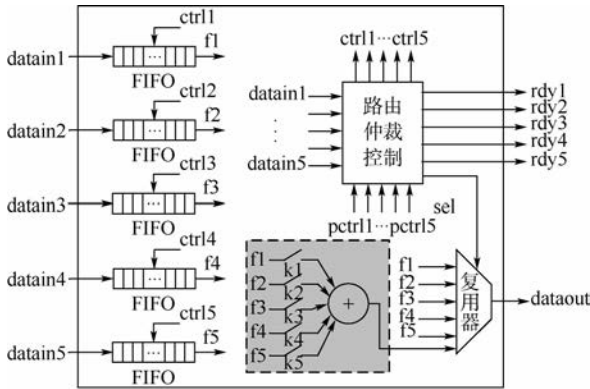


图7 无线节点的交换模块示意图

Fig. 7 Schematic of switch architecture of wireless router

用 Verilog HDL 建立路由节点的功能模型,在 Synopsys 的 Design Compiler 下面用 90 nm 工艺进行综合,从综合结果的面积和功耗分析可知,与普通无线 NoC 的路由节点相比,基于网络编码的无线 NoC 路由节点的功耗和面积略有增加,分别为 0.06% 和 0.8%。相对于第 5.1 节讨论的传输延时和功耗性能提高而言,这样的附加开销是较小的、可接受的。

5.3 映射算法的执行时间分析

在 MATLAB 平台上,采用本文提出的算法在网络拓扑为 5×5 Mesh 和 10×10 Mesh 下映射 MIMO-OFDM 接收端,得到一个映射结果的时间分别为 28.4846 s 和 62.8402 s;相比普通映射算法的 8.1626 s 和 10.0816 s,分别多耗时 3 倍和 6 倍。考虑到映射算法是离线执行的,只有在 NoC 承载应用的开始时刻执行一次,其后应用持续处理时无需执行映射算法,因此算法执行时间的略微增加是可接受的,不会降低 NoC 的性能。

6 结论

本文提出了一种网络编码在无线 NoC 上的应用机制,并且给出了将具有多播流的任务流图映射到无线 NoC 平台上的实现算法,经实验验证表明:

1) 与普通无线 NoC 相比,本文提出的网络编码方法及映射算法具有很好的网络性能,功耗方面,至少节约 6%,网络吞吐率方面,能提高至少 16%。

2) 本文算法带来的性能开销非常小,相比普通无线 NoC,路由节点的面积和功耗分别仅增加了 0.8% 和 0.06%。

参考文献 (References)

- [1] LIN J, WU H T, SU Y, et al. Communication using antennas fabricated in silicon integrated circuits [J]. IEEE Journal of Solid-State Circuits, 2007, 42(8): 1678-1687.
- [2] SASAKI N, FUKUDA M, KIMOTO K, et al. CMOS UWB transmitter and receiver with silicon integrated antennas for inter-chip wireless interconnection [C] // Proceedings of 2008 IEEE on Radio and Wireless Symposium. Piscataway, NJ: IEEE Press, 2008: 795-798.
- [3] ZHAO D, WANG Y. SD-MAC: Design and synthesis of a hardware-efficient collision-free QoS-aware MAC protocol for wireless network-on-chip [J]. IEEE Transactions on Computers, 2008, 57(9): 1230-1245.
- [4] ZHAO D, WANG Y, LI J, et al. Design of multi-channel wireless NoC to improve on-chip communication capacity [C] // Proceedings of 5th IEEE/ACM International Symposium on Networks on Chip. Piscataway, NJ: IEEE Press, 2011: 177-184.
- [5] DIJKSTRA E W. A note on two problems in connexion with graphs [J]. Numerische Mathematik, 1959, 1(1): 269-271.
- [6] BERTSEKAS D P, GALLAGER R G. Data networks [M]. 2nd ed. Englewood Cliffs: Prentice-Hall, 1992: 396-398.
- [7] KOU L, MARKOWSKY G, BERMAN L. A fast algorithm for Steiner trees [J]. Acta Informatica, 1981, 15(2): 141-145.
- [8] TAKAHASHI H, MATSUYAMA A. An approximate solution for the Steiner problem in graphs [J]. Math Japonica, 1980, 24(6): 573-577.
- [9] AHLWEDE R, CAI N, YEUNG R W, et al. Network information flow theory [C] // Proceedings of 1998 IEEE International Symposium on Information Theory. Piscataway, NJ: IEEE Press, 1998: 186.
- [10] 张思为. 网络编码在基于局部总线的 NoC 的应用研究 [D]. 成都: 电子科技大学, 2010: 24-34.
ZHANG S W. Application research on network coding based on NoC with local bus [D]. Chengdu: University of Electronic Science and Technology of China, 2010: 24-34 (in Chinese).
- [11] SHALABY A, GOULART V, RAGAB M E S. Study of application of network coding on NoCs for multicast communications [C] // Proceedings of IEEE 7th International Symposium on Embedded Multicore SoCs. Piscataway, NJ: IEEE Press, 2013: 37-42.
- [12] VONBUN M, WALLENTOWITZ S, FEILEN M, et al. Evaluation of hop count advantages of network-coded 2D-mesh NoCs [C] // Proceedings of 23rd International Workshop on Power and Timing Modeling, Optimization and Simulation. Piscataway, NJ: IEEE Press, 2013: 134-141.
- [13] REINHARD D. Graph theory III [M]. New York: Springer-Verlag, 2005: 111-126.
- [14] CHEN W, LETAIEF K B, CAO Z. Buffer-aware network coding for wireless networks [J]. IEEE/ACM Transactions on Networking, 2012, 20(5): 1389-1401.
- [15] CHEN Y, HU J, LING X. Topology and mapping co-design for complex communication systems on wireless NoC platforms [C] // Proceedings of 8th IEEE Conference on Industrial Elec-

- tronics and Applications. Piscataway, NJ: IEEE Press, 2013: 1442-1447.
- [16] 陈亦欧,胡剑浩. 面向无线 NoC 平台的拓扑与映射联合设计[J]. 微电子学,2012,42(6):846-849.
CHEN Y O, HU J H. Co-design of topology and mapping for wireless NoC platform[J]. Journal of Microelectronics,2012,42(6):846-849(in Chinese).
- [17] LAUFER R, DUBOIS-FERRIÈRE H, KLEINROCK L. Polynomial-time algorithms for multirate anypath routing in wireless multihop networks[J]. IEEE/ACM Transactions on Networking,2012,20(3):742-755.
- [18] YAZDI S M S, SAVARI S A. A deterministic polynomial time algorithm for constructing a multicast coding scheme for linear deterministic relay networks[C]// Proceedings of 45th Annual Conference on Information Sciences and Systems. Piscataway, NJ: IEEE Press,2011:1-6.
- [19] 陈亦欧,胡剑浩,陈庚生. 面向实时复杂系统的基于片上网络多核平台的映射技术研究[J]. 计算机应用研究,2012,29(7):2589-2592.
CHEN Y O, HU J H, CHEN G S. Energy and performance-aware mapping for real-time complex system based on NoC platform[J]. Journal of Application Research of Computers,2012,29(7):2589-2592(in Chinese).
- 作者简介:**
陈亦欧 女,博士,副教授。主要研究方向:通信集成电路、信号处理与多核并行技术。
Tel.: 028-61830326
E-mail: chenyou@uestc.edu.cn
- 凌翔 男,博士,教授,博士生导师。主要研究方向:无线通信与 VLSI 设计。
Tel.: 028-61831270
E-mail: xiangling@uestc.edu.cn

Multicast wireless NoC mapping based on network coding

CHEN Yiou*, TAN Fang, LING Xiang

(National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract: Featured by high bandwidth, QoS guarantee, and broadcasting, wireless network on chip (NoC) can provide an efficient platform for multicore and parallel processing for communication and signal processing systems. In addition, network coding can greatly increase the transmission efficiency at the cost of small power consumption in wireless NoC. On this basis, we proceeded from the features and requirements of communication and signal processing systems, and built proper architectures and mapping models for wireless NoCs. We also proposed a joint optimization method including task mapping and network coding mapping, to meet the specialized constraints and requirements of wireless NoC multicast services, such as maximizing network throughput, data transmission, and minimizing communication latency and power consumption, so as to achieve efficient task paralleling and multi-processor cooperation. Experimental results prove that the proposed architecture and mapping algorithm can achieve at least 6% power consumption saving and 16% throughput gain, with acceptable physical complexity increasement than the traditional methods.

Key words: wireless network on chip; network coding; mapping; multicast; low power

Received: 2015-07-17; **Accepted:** 2015-09-30; **Published online:** 2016-01-04 10:03

URL: www.cnki.net/kcms/detail/11.2625.V.20160104.1003.002.html

Foundation items: the Fundamental Research Funds for the Central Universities (ZYGX2014J011); National Science and Technology Major Project (2013ZX03003015); National High-tech Research and Development Program of China (2015AA01A701)

* **Corresponding author.** Tel.: 028-61830326 E-mail: chenyou@uestc.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0436

高速切削高温合金 GH4169 数值模拟与实验



范孝良, 吴学华, 王进峰*, 康文利

(华北电力大学 能源动力与机械工程学院, 保定 071003)

摘 要: 为了准确模拟高温合金 GH4169 高速切削过程, 深入研究了高速切削 GH4169 的有限元建模技术, 包括有限元模型的建立、材料本构模型、切屑分离准则以及接触摩擦模型等关键技术。为了模拟高速切削 GH4169 的切屑分离过程, 研究切屑形态及其形成机理, 分别采用 Johnson-Cook 和各项同性硬化本构关系模型对 GH4169 的高速加工过程进行二维正交切削有限元模拟, 2 种模型都获得了相类似的锯齿状切屑。在此基础上, 模拟了基于上述 2 种模型的应力场、温度场和切削力曲线。为了验证有限元模型的有效性和正确性, 在 CA6140 机床进行了 GH4169 高速车削实验, 实验获得的锯齿形切屑验证了 2 种有限元模型的正确性, 实验结果表明: 随着切削速度的增大, 锯齿状切屑的锯齿化程度增大; 绝热剪切是导致高速切削 GH4169 生成锯齿状切屑的主要原因。实验测量的切削力曲线和切削温度场, 与有限元模型 A 输出结果更好地吻合, 进一步表明模型 A 比模型 B 更能反映 GH4169 的实际高速加工特性。

关键词: GH4169; 锯齿切屑; 有限元模拟; Johnson-Cook 本构; 各项同性硬化本构; 绝热剪切

中图分类号: TG506.1

文献标识码: A **文章编号:** 1001-5965(2016)07-1344-08

镍基高温合金 GH4169 因其具有高屈服应力和抗疲劳等优良性能, 常用于制造燃气轮机的轮盘、航空发动机与叶片等耐热零部件, 但由于它同时具有导热系数低、弹性模量小等特点, 是典型的难加工材料。对于 GH4169 这种难加工材料, 采用传统的切削加工, 效率极低, 刀具寿命短, 必须建立适合高效切削技术的新理论和研究方法。高速切削高温合金切削加工的主要特征是产生锯齿状切屑, 锯齿状切屑的产生往往会导致切削温度高以及切削力周期性振荡, 从而加重刀具磨损, 影响工件表面加工质量。关于锯齿形切屑的形成机理的解释主要有 2 种: 绝热剪切理论和周期性脆断理论^[1]。国内外学者对锯齿形切屑形成原因进行了实验研究, Komanduri 和 Brown^[1] 提出用

绝热剪切理论来解释锯齿形切屑形成原理, 他们采用高速摄影和 SEM 内原位低速切削的方法, 研究钛合金和高强度钢的锯齿形切屑。Poulachon 和 Moisan^[2] 研究了在硬态切削 100Cr 6 (相当于 GCr15) 时的切削变形机理, 认为锯齿形切屑的形成机理是脆性断裂。吴申峰和张雪萍^[3] 研究了 AISI52100 轴承钢在低、高速切削条件下锯齿状切屑的形成过程及绝热剪切机理。哈尔滨工业大学郝兆朋^[4] 也对高温合金锯齿形切屑的形成机理进行实验研究。Wang 等^[5] 研究了高速铣削淬硬钢锯齿形切削的形成机理, 并建立了锯齿形切屑的几何模型以预测绝热剪切变形区的剪应变和应变率。大连理工大学谷丽瑶等^[6] 认为高速切削过程的绝热剪切演化随切削速度的提高经历了

收稿日期: 2015-06-30; 录用日期: 2015-09-06; 网络出版时间: 2015-09-16 19:37

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150916.1937.001.html

基金项目: 中央高校基本科研业务费专项资金(14ZD37); 国家自然科学基金(51301068); 河北省自然科学基金(E2014502003)

* 通讯作者: Tel.: 0312-7525045 E-mail: wj266@163.com

引用格式: 范孝良, 吴学华, 王进峰, 等. 高速切削高温合金 GH4169 数值模拟与实验[J]. 北京航空航天大学学报, 2016, 42(7): 1344-1351. FAN X L, WU X H, WANG J F, et al. Numerical simulation and experiment in high-speed cutting superalloy GH4169 [J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1344-1351 (in Chinese).

绝热剪切的发生、形变带、转变带和绝热剪切局部化断裂。目前,对高温合金锯齿状切屑形成原因的研究不多,且始终没有得到统一的理论。

近几年来,有限元数值模拟技术被广泛应用于金属切削仿真中,为研究难加工材料切削加工机理提供方便。国内外学者对锯齿形切屑仿真的研究主要集中在类似于难加工材料高强度钢和钛合金,对高速切削高温镍基合金的切削变形,研究的不多。Mabrouki 和 Rigal^[7]对 AISI 4340 钢硬锯齿状切屑形成过程进行了仿真研究,发现锯齿形切屑是由于材料热软化现象而引起的。Rhim 和 Oh^[8]运用修正的流变应力模型对钛合金等难加工材料进行仿真,模拟出了锯齿形切屑中绝热剪切带的存在。Arrazola等^[9]建立二维有限元模型研究了锯齿形切屑形成过程中材料参数对其形成的影响,其研究表明材料的热参数特性,尤其是热导率和比热,对锯齿形切屑的形成有重要的影响。山东大学杨奇彪和刘战强^[10]建立了 Ti6Al4V 绝热剪切带变形的解析模型,以绝热剪切带带内剪切应变和剪切应变率来表征绝热剪切带的变形程度,并进行了相关的切削实验。对于高温合金有限元仿真的研究,还存在一些问题,比如,仿真过程中采取哪类材料模型和何种失效准则,到目前为止,尚无统一的观点。

针对高温合金高速切削加工中存在的这些问题,本文使用有限元软件 ABAQUS 建立 2 种仿真模型,模型 A 选用 Johnson-Cook(J-C)材料模型和 J-C 材料失效准则,模型 B 选用各项同性硬化本构关系模型和 Shear Damage 切屑分离准则,模拟高温合金的高速切削过程,对 2 种模型获得的相关参数进行分析,并与实验进行对比。

1 高速切削有限元建模

金属高速切削是一个复杂的动力效应过程,在建模时要考虑动力学、材料断裂准则以及接触摩擦原理等因素。同时金属切削过程又是热力耦合过程,切削过程中会产生大量的切削热,切削热主要分布在 3 个区域:主变形区(剪切区)、刀具-工件摩擦区以及刀具-切屑摩擦区,切削区分布如图 1 所示^[11]。能否获得理想的仿真结果主要取决于有限元算法、材料本构模型、刀具-切屑的接触设置、摩擦模型以及正确划分网格等方面。本文采用拉格朗日有限元算法,选取 ABAQUS 中的动态-显示温度位移耦合分析步,对工件采用稳定性较好

的四节点平面应变缩减积分单元 CPE4RT。刀具采用刚度相对较大的三节点平面应力单元 CPE3T,在刀具与工件接触的部位划分较密的网格,以提高计算精度。

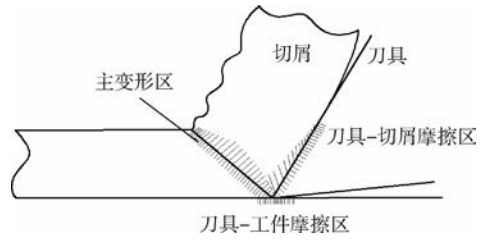


图 1 正交切削模型的 3 个变形区^[11]

Fig.1 Three deformation zones of orthogonal cutting model^[11]

1.1 材料本构关系及切屑分离准则

材料本构关系和切屑分离准则设置是否合理,直接影响到仿真结果的好坏。本文选取不同材料本构关系及分离准则建立 2 种仿真模型。

1.1.1 模型 A

1) 本构关系模型。选择能够较好表达金属材料粘塑性本构关系特性——J-C 本构关系模型^[12]。本构关系模型的数学表达式为

$$\sigma = [A + B(\bar{\epsilon})^n] \cdot \left[1 + C \ln \left(\frac{\dot{\epsilon}}{\dot{\epsilon}_0} \right) \right] \cdot \left[1 - \left(\frac{Q - Q_0}{Q_{\text{melt}} - Q_0} \right)^m \right] \quad (1)$$

式中: A 、 B 、 C 、 m 、 n 均为材料常数, A 为材料的屈服应力, B 为应变强化参数, C 为材料应变速率敏感系数, n 为材料应变强化项指数, m 为材料热软化系数; σ 为材料的屈服极限; Q_{melt} 、 Q_0 和 Q 分别为材料的熔点、转变温度以及变形温度; $\bar{\epsilon}$ 和 $\dot{\epsilon}$ 分别为等效塑性应变和应变率; $\dot{\epsilon}_0$ 为参考应变率。本构方程所取参数如表 1 所示^[13]。

表 1 高温合金 GH4169 本构方程参数^[13]

Table 1 Parameters of superalloy GH4169 constitutive equation^[13]

参数	A/MPa	B/MPa	C	m	n
数值	980	1370	0.164	1.03	0.02

2) 模型的失效准则。Johnson 和 Cook 建立了 J-C 破坏准则,它是以前述等效塑性应变来衡量的。

$$\bar{\epsilon}^{\text{fp}} = \left[D_1 + D_2 \exp \left(D_3 \frac{p}{S} \right) \right] \left(1 + D_4 \ln \frac{\dot{\epsilon}}{\dot{\epsilon}_0} \right) \cdot \left[1 - D_5 \left(\frac{Q - Q_0}{Q_{\text{melt}} - Q_0} \right)^m \right] \quad (2)$$

式中: $\bar{\epsilon}^{\text{fp}}$ 为失效的等效塑性应变; p/S 为无量纲的偏应力比值, p 为压应力, S 为 Mises 应力; D_1 、

D_2 、 D_3 、 D_4 和 D_5 为材料的失效参数。

衡量破坏准则标准的参数定义如式(3)所示,当 V 值累加到 1 时,说明材料失效。

$$V = \sum \left(\frac{\Delta \varepsilon}{\varepsilon_{rp}} \right) \quad (3)$$

1.1.2 模型 B

1) 本构关系模型。选择各向同性硬化本构关系模型^[14],此模型能够体现工件材料热弹塑性行为,且当温度不同时能反映材料的屈服应力和塑性应变之间关系。在 ABAQUS 软件中设置材料的 Hardening 为 Isotropic,输入如表 2 所示的数值^[15]。

表 2 屈服应力、塑性应变、温度数据表^[15]

Table 2 Data sheet of yield stress, plastic strain and temperature^[15]

序号	屈服应力/MPa	塑性应变	温度/K
1	989.0	0	293
2	1119.0	0.042	293
3	1190.0	0.092	293
4	1236.4	0.142	293
5	1282.7	0.192	293
6	1298.0	0.272	293
7	896.0	0	673
8	989.0	0.042	673
9	1051.0	0.092	673
10	1112.7	0.142	673
11	1128.2	0.192	673
12	1174.5	0.242	673
13	1190.0	0.292	673
14	803.6	0	823
15	927.3	0.042	823
16	989.0	0.092	823
17	1004.5	0.142	823
18	1020.0	0.192	823
19	1035.0	0.242	823
20	772.7	0	973
21	881.0	0.042	973
22	942.7	0.092	973
23	989.0	0.142	973
24	1020.0	0.212	973
25	432.7	0	1123
26	541.0	0.046	1123
27	587.3	0.096	1123
28	649.0	0.146	1123
29	680.0	0.196	1123
30	695.0	0.246	1123

2) 切屑分离准则。材料的分离是一个复杂的现象,涉及在微观层面上发生的物理变化过程。此模型采用 Shear Damage 准则,实现切屑与工件的分离。

1.2 接触-摩擦模型

本文模拟选取修正的库伦摩擦接触模型,实验研究表明,切屑与前刀面的接触部分可以划分为粘结摩擦区和滑动摩擦区^[16]。划分 2 个区域

的模型表达式为

$$\tau = \min(\mu\sigma_0, \sigma_1) \quad (4)$$

式中: τ 为接触区域剪切应力; μ 为前刀面的摩擦因数,此处取值 0.9; σ_0 为法向应力; σ_1 为工件材料临界屈服强度。

2 有限元模拟结果与分析

为了更好地研究高速切削高温合金 GH4169 的切削加工性能,模拟实验采取的切削参数如下:切削速度 V_c 分别取 90 m/min、144 m/min 和 200 m/min,切削厚度 $a_p = 0.15$ mm、刀具前角 γ_0 为 -5° 、刀具后角 α_0 为 0° 。

图 2 为模型 A 和模型 B 锯齿形切屑的形成过程。 a 时刻刀具刚切入时,工件在刀具作用下发生挤压变形,开始形成切屑;随着刀具的前进,切削温度逐渐升高,主变形区开始发生变形,在 b 时刻产生锯齿状切屑的第 1 个节块,主变形区的变形继续进行, c 时刻锯齿状切屑的第 2 个节块开始出现,且 2 种模型在第 1 个节块和第 2 个节块之间都能明显地观察到绝热剪切带的存在;刀具的继续前进,同时伴随着应力和温度不断增加,导致材料发生热软化效应,致使主变形区剪切变形和剪切滑移现象更加明显,连续形成锯齿状切屑如 d 、 e 及 f 时刻所示,且形成的锯齿状切屑形状规则,呈现周期性变化,锯齿状切屑相邻节块之间都能明显地观察到绝热剪切带的存在。

图 3 和图 4 是在不同切削速度,2 种模型 Mises 应力云图分布,同时也可以观察 2 种模型不同切削速度下的切屑形态。由图所示,任何切削速度下,模型 A 和模型 B 最大 Mises 应力都主要分布在切屑的主变形区,刀具和工件的接触使材料应变增加,出现加工硬化效应,此时切屑在主变形区受到剪切和挤压作用,产生较大的塑性变形,从而引起较大的应力。从仿真结果还可以发现,高速切削高温合金的过程中,切屑都呈现锯齿状,且随着切削速度增大,锯齿化程度更加明显,模型 A 锯齿化程度比模型 B 更显著。

伴随着锯齿状切屑的周期性形成,切削力发生周期性波动,图 5 和图 6 分别是模型 A 和模型 B 模拟得到的主切削力随时间变化图。由图 5 和图 6 可以发现:高速切削高温合金过程中,切削力随着切削速度的增大而减小;模型 A 模拟出的切削力曲线比模型 B 波动更明显,更符合高速切削高温合金 GH4169 的加工特性。图 7 和图 8 是 2 种模型的工件和刀具温度 T 分布云图。

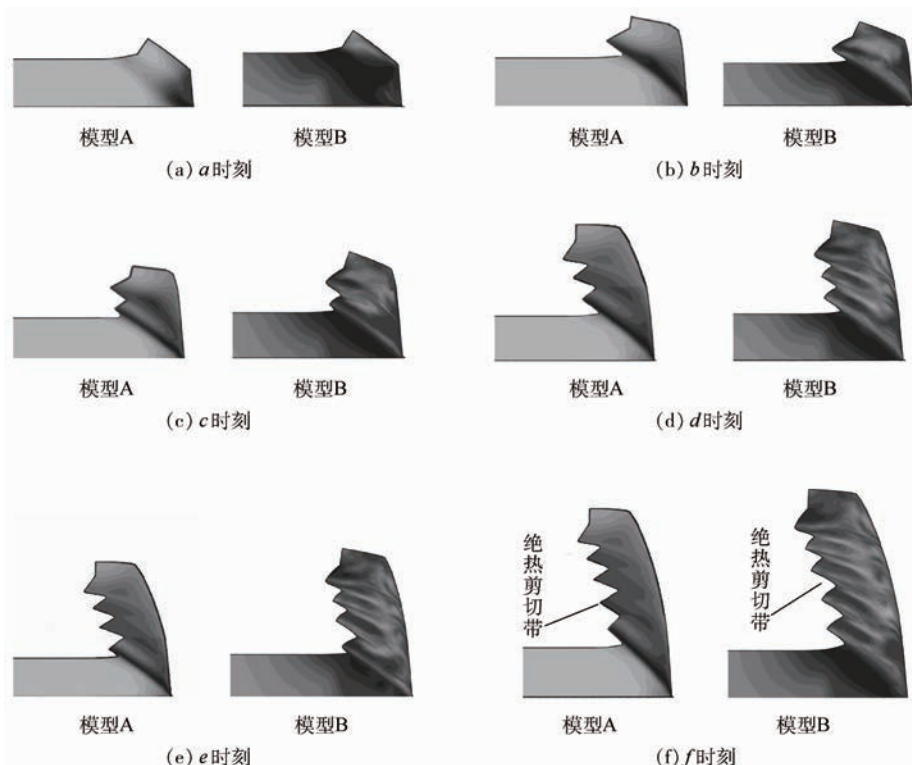


图 2 锯齿形切屑的形成过程

Fig.2 Serrated chip formation process

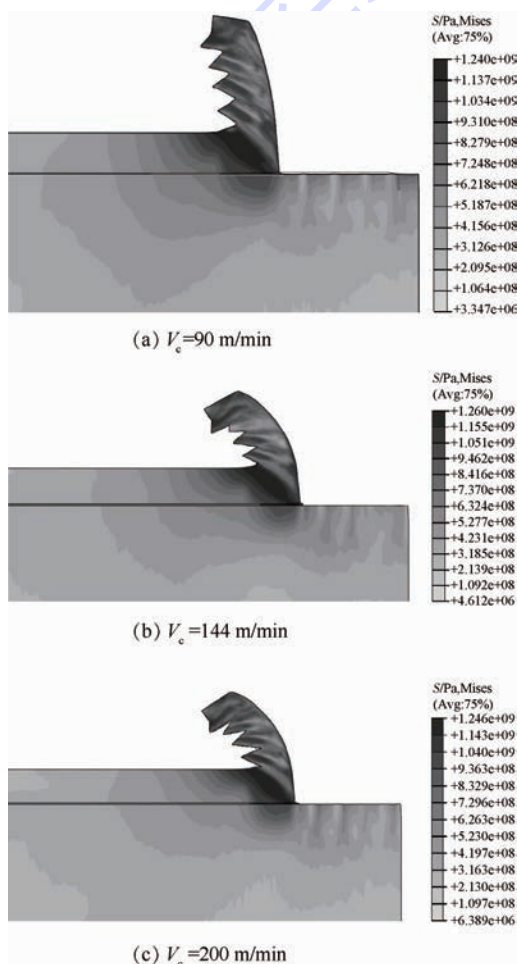


图 3 模型 A Mises 应力分布云图

Fig.3 Model A Mises stress distribution contours

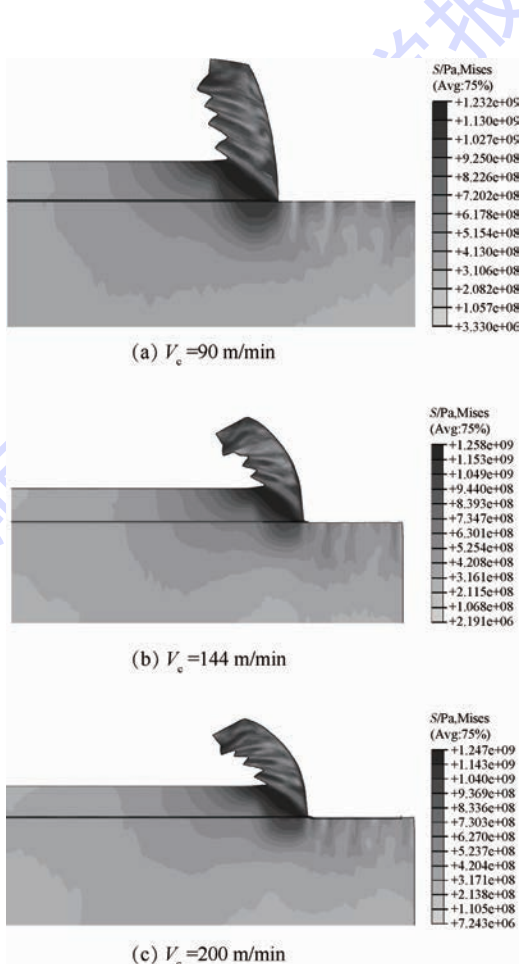


图 4 模型 B Mises 应力分布云图

Fig.4 Model B Mises stress distribution contours

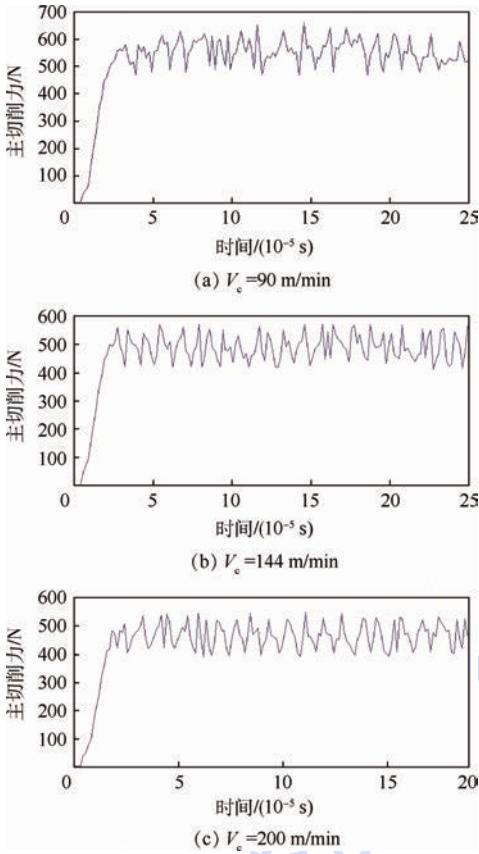


图5 模型 A 主切削力变化曲线

Fig.5 Model A main cutting force curve

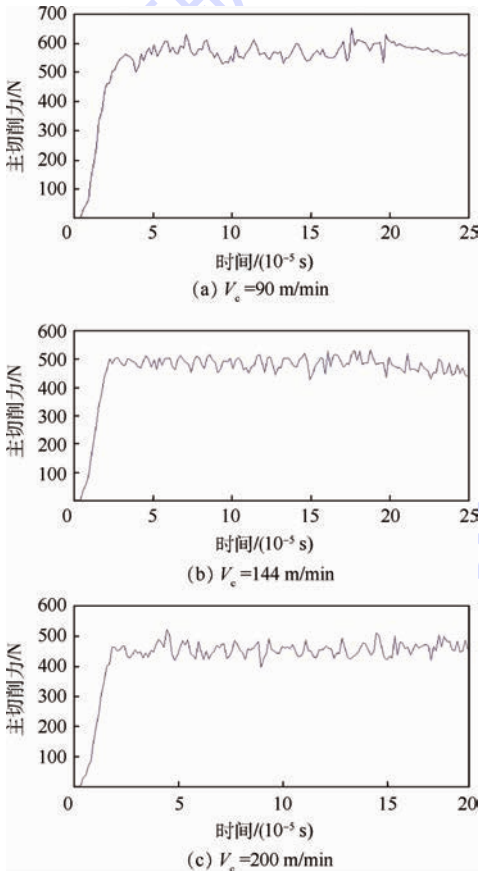


图6 模型 B 主切削力变化曲线

Fig.6 Model B main cutting force curves

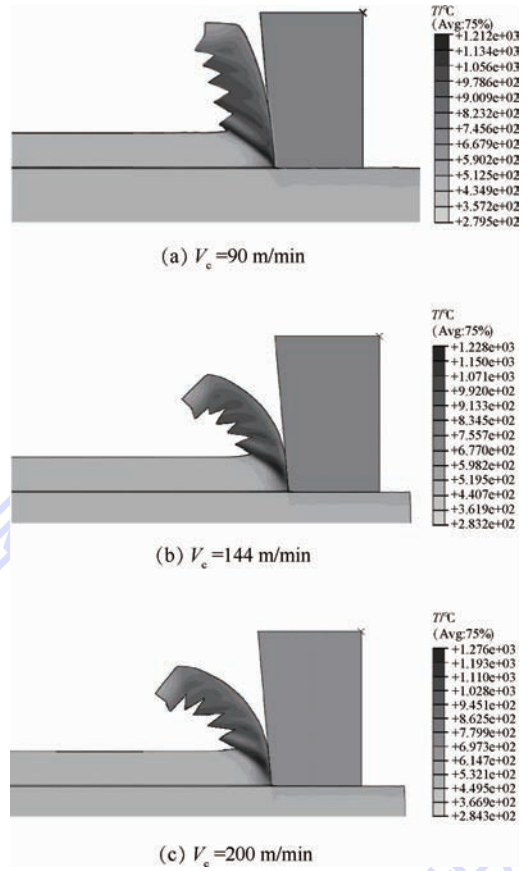


图7 模型 A 温度分布云图

Fig.7 Model A temperature contours

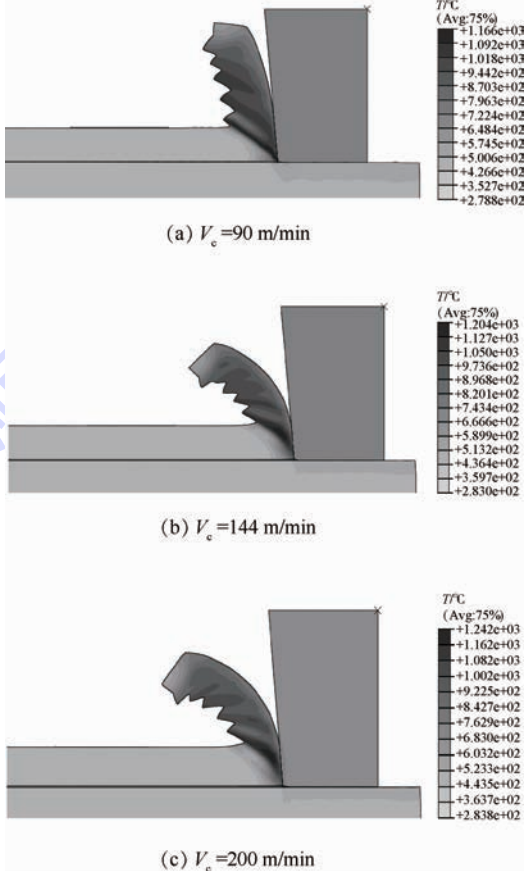


图8 模型 B 温度分布云图

Fig.8 Model B temperature contours

由图 7 和图 8 可以看出,在高速切削过程中,最高温度分布在切屑部分,切屑带走大量的切削热,高速切削过程相当于绝热的过程,刀具和工件接触时间短,剪切变形区产生的热量来不及更多的传递,致使切屑温度很高,且产生的大量热量大部分分布在绝热剪切带附近形成绝热剪切^[5]。其中会有一部分热量传递给刀具,工件内部温度几乎不变,只有表层发生温度的变化。由此可知,高速切削时,工件温升较小,能够获得较好的表面加工质量,提高加工精度。因此,在实际加工高温合金 GH4169 时,可以根据实际加工需求,适当提高切削速度。

3 有限元模拟的实验验证

为了验证所建立有限元模型和模拟结果的正确性,在 CA6140 车床上,对高温合金 GH4169 进行高速正交车削实验,工件为棒料 GH4169,尺寸为 $\phi 80\text{mm} \times 100\text{mm}$,其组成元素及含量百分比如表 3 所示,刀具选用 YG8 硬质合金刀具,切削环境为干切削。采用 LEICA MEF4A 金相显微镜观察切屑形态及金相组织并进行拍照,采用 YDC-III 89 三向压电车削测力仪来测量实际切削力,采用自然热电偶法测量切削温度。

表 3 高温合金 GH4169 化学成分

元素	Ni	Cr	Mo	Al	Ti	Nb
含量/ vol%	50.0 ~ 55.0	17.0 ~ 21.0	2.8 ~ 3.3	0.3 ~ 0.7	0.75 ~ 1.15	4.75 ~ 5.50

切削实验过程中,选择与仿真一致的刀具参数和切削参数,进一步探究高速切削高温合金锯齿状切屑的形成原因,研究切削速度对切削力和切削温度的影响趋势,并与模拟结果对比分析。

图 9 为实验所得的切屑形态,将其与仿真结果对比,可以发现:相同切削参数下,实验所得切屑形态与有限元模拟产生的切屑形态基本一致,所产生的切屑都呈锯齿状,切屑上相邻的节块之间都能明显地观察到绝热剪切带的存在,因此本文认为高速切削高温合金所形成的锯齿形切屑由绝热剪切现象引起的,应该用绝热剪切理论来解释;观察实验所得锯齿状切屑,随着切削速度的增大,锯齿化程度更加明显^[17],这与仿真结果基本一致;模型 A 所得切屑形态更接近实验值。

图 10 为实验所得主切削力值与仿真值比较,由图可知:高速切削高温合金 GH4169 过程中,切削力随着切削速度的增大而减小;模型 A 和模型 B 仿真出切削力变化趋势与实验所得基

本一致,模型 A 所得结果更接近实验值,且模型 A 和模型 B 得到的切削力值与实验值误差在 10% 以内,因此 2 种模型得到的切削力的模拟结果都是可以接受的。

本实验,用自然热电偶法测量的切削温度是指刀具前刀面与切屑接触区域的平均温度,而仿真中很难精确地得到切削过程中的平均温度,但能够准确得出某个单元或节点在某个位置某个时刻的温度,故选取仿真中刀具与工件接触部分的最高温度与实验获得的平均温度对比^[18],得到切削温度的仿真值与实验值如图 11 所示。由图可知,模型 A 和模型 B 模拟得到的最高温度和实验得到的平均温度变化趋势相似。

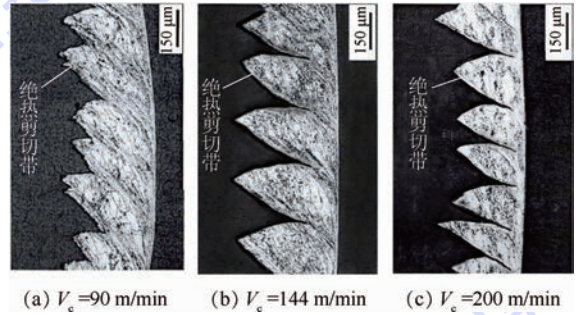


图 9 实验切屑形态

Fig. 9 Experimental chip morphology

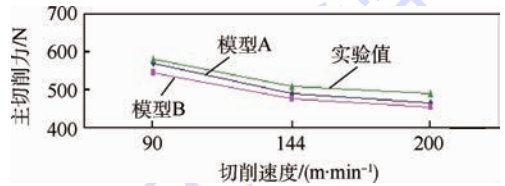


图 10 主切削力的实验值与仿真值比较

Fig. 10 Comparison between experimental and simulated main cutting force

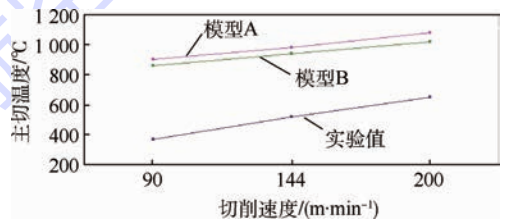


图 11 实验平均温度与仿真最高温度比较

Fig. 11 Comparison between experimental average temperature and simulated maximum temperature

4 结论

1) 本文利用 ABAQUS 有限元软件,选用了 2 种不同的有限元模型模拟高速切削高温合金 GH4169 的切削过程。2 种模型成功地模拟出高速切削 GH4169 过程中所产生的锯齿状切屑,且

切屑相邻节块之间都有明显的绝热剪切带。同时,研究了锯齿状切屑的形成过程,分析了2种模型在相同模拟参数下的应力和温度分布状况以及主切削力随时间的变化规律。

2) 通过高速正交车削实验,对2种模型的有效性进行验证。实验结果表明:2种有限元模型所得切屑形态与实验所得基本一致,模型A所得切屑形态与实验值更接近。由实验所得切屑形态可以发现,随着切削速度的增加,锯齿化程度增大;实验获得的锯齿状切屑上相邻节块之间能观察到明显的绝热剪切带,表明高速切削高温合金所形成的锯齿形切屑由绝热剪切现象引起的。实验获得的切削力曲线和切削温度与有限元模拟具有较好的一致性,再次证明有限元模型的正确性,且无论是切屑形态还是切削力曲线都表明模型A比模型B更能反映GH4169的实际高速加工特性。

参考文献 (References)

- [1] KOMANDURI R, BROWN R H. On the mechanics of chip segmentation in machining[J]. *Journal of Engineering for Industry*, 1981, 103(1): 33-51.
- [2] POULACHON G, MOISAN A L. Hard turning: Chip formation mechanisms and metallurgical aspects[J]. *Journal of Manufacturing Science and Engineering*, 2000, 122(8): 406-412.
- [3] 吴申峰, 张雪萍. 淬硬轴承钢锯齿形切屑形成机理[J]. *上海交通大学学报*, 2011, 45(1): 71-77.
WU S F, ZHANG X P. Saw-tooth chip formation process and mechanism in hard turning AISI 52100 bearing steel [J]. *Journal of Shanghai Jiao Tong University*, 2011, 45(1): 71-77 (in Chinese).
- [4] 郝兆朋. 切削GH4169的相关机理及高效切削技术的基础研究[D]. 哈尔滨: 哈尔滨工业大学, 2013: 19-24.
HAO Z P. Research on related mechanism and high efficiency cutting technology in machining nickel-based superalloy GH4169 [D]. Harbin: Harbin Institute of Technology, 2013: 19-24 (in Chinese).
- [5] WANG C Y, XIE Y Y, ZHENG L J, et al. Research on the chip formation mechanism during the high-speed milling of hardened steel [J]. *International Journal of Machine Tools and Manufacture*, 2014, 79: 31-48.
- [6] 谷丽瑶, 王敏杰, 孙传俊. 高速切削过程绝热剪切局部化断裂的特性实验[J]. *机械工程学报*, 2015, 50(15): 166-171.
GU L Y, WANG M J, SUN C J. Experimental study on characteristics of adiabatic shear localization fracture in high speed machining [J]. *Journal of Mechanical Engineering*, 2015, 50(15): 166-171 (in Chinese).
- [7] MABROUKI T, RIGAL J F. A contribution to a qualitative understanding of thermo-mechanical effects during chip formation in hard turning [J]. *Journal of Materials Processing Technology*, 2006, 176(1-3): 214-221.
- [8] RHIM S H, OH S I. Prediction of serrated chip formation in metal cutting process with new flow stress model for AISI 1045 steel [J]. *Journal of Materials Processing Technology*, 2006, 171(3): 417-422.
- [9] ARRAZOLA P J, BARBERO O, URRESTI I. Influence of material parameters on serrated chip prediction in finite element modeling of chip formation process [J]. *International Journal of Material Forming*, 2010, 3(Suppl. 1): 519-522.
- [10] 杨奇彪, 刘战强. Ti6Al4V 切削锯齿形切屑绝热剪切带变形的表征[J]. *沈阳工业大学学报*, 2013, 35(2): 181-186.
YANG Q B, LIU Z Q. Deformation characterization of adiabatic shear band for Ti6Al4V serrated chips [J]. *Journal of Shenyang University of Technology*, 2013, 35(2): 181-186 (in Chinese).
- [11] 肖茂华, 何宁, 李亮, 等. 镍基高温合金高速洗削的切削热研究[J]. *哈尔滨工程大学学报*, 2011, 43(11): 105-109.
XIAO M H, HE N, LI L, et al. Cutting heat of Ni-base superalloy high speed milling [J]. *Journal of Harbin Engineering University*, 2011, 43(11): 105-109 (in Chinese).
- [12] ZHU D Z, CHEN W P, LI Y Y. Strain-rate relationship of aluminum matrix composites predicted by Johnson-Cook model [C]// *Proceedings of the 6th International Conference on Physical and Numerical Simulation of Materials Processing*. Zurich-Dumten: Trans Tech Publications Ltd, 2012: 935-940.
- [13] DING H T, SHEN N G. Thermal and mechanical modeling analysis of laser-assisted micro-milling of difficult to machine alloys [J]. *Journal of Materials Processing Technology*, 2012, 212(3): 601-603.
- [14] 段春争, 王肇喜, 李红华. 高速切削锯齿形切屑形成过程的有限元模拟[J]. *哈尔滨工程大学学报*, 2014, 35(2): 226-232.
DUAN C Z, WANG Z X, LI H H. Finite element simulation of the formation process of a serrated chip in high-speed cutting [J]. *Journal of Harbin Engineering University*, 2014, 35(2): 226-232 (in Chinese).
- [15] 黄美霞. 虚拟高速切削加工仿真技术研究[D]. 太原: 中北大学, 2010.
HUANG M X. Virtual simulation technology for high speed cutting [D]. Taiyuan: North University of China, 2010 (in Chinese).
- [16] 魏彦杰. 金属切削过程刀-屑接触区摩擦状态实验研究[D]. 秦皇岛: 燕山大学, 2008: 7-9.
WEI Y J. Experimental study of friction on the tool-chip contact field in metal cutting process [D]. Qinghuangdao: Yanshan University, 2008: 7-9 (in Chinese).
- [17] SCHULZ H, ABELE E. 高速加工理论与应用[M]. 何宁, 译. 北京: 科学出版社, 2010: 18-21.
SCHULZ H, ABELE E. The high speed machining fundamentals and applications [M]. HE N, translated. Beijing: Science Press, 2010: 18-21 (in Chinese).
- [18] 曹自洋, 何宁, 李亮, 等. 高速切削钛合金Ti6Al4V切屑的形成及其数值模拟[J]. *中国机械工程*, 2008, 19(20): 2450-2454.
CAO Z Y, HE N, LI L, et al. High-speed cutting of titanium alloy Ti6Al4V chip formation and its numerical simulation [J]. *Chinese Mechanical Engineering*, 2008, 19(20): 2450-2454 (in Chinese).

作者简介:

范孝良 男,硕士,教授。主要研究方向:数字化设计与制造。

Tel.: 0321-7525029

E-mail: wx803@163.com

王进峰 男,博士,副教授。主要研究方向:智能制造、金属切削等。

Tel.: 0312-7525045

E-mail: wjf266@163.com

Numerical simulation and experiment in high-speed cutting superalloy GH4169

FAN Xiaoliang, WU Xuehua, WANG Jinfeng*, KANG Wenli

(School of Energy, Power and Mechanical Engineering, North China Electric Power University, Baoding 071003, China)

Abstract: The finite element modeling technologies of the high-speed cutting superalloy GH4169 were investigated to simulate the cutting process correctly, which include establishing the finite element model, selecting the material constitutive model, and determining the chip separation criteria and contacting friction model, etc. To simulate the chip separation process and research the chip formation mechanism in high-speed cutting GH4169, the two-dimensional orthogonal cutting finite element models were constructed using Johnson-Cook model and isotropic hardening material constitutive model respectively. As a result, two models obtained the similar serrated chip. On this basis, equivalent plastic strain, cutting temperature and cutting force curve were simulated further. To verify the validity of the finite element model, an experiment was carried out by high-speed turning GH4169 on the machine of CA6140, and the serrated chip generated from the turning experiment verifies the validity of two finite element models. The results show that the serrated chip sawtooth degree increases with the increase of cutting speed and the adiabatic shear is the main reason to cause the serrated chip on high-speed cutting superalloy GH4169. The cutting force curve and cutting temperature generated from the experiment accord better with model A, which demonstrates that model A exceeds model B in reflecting the machining characteristics of high-speed cutting GH4169.

Key words: GH4169; serrated chip; finite element simulation; Johnson-Cook constitutive; isotropic hardening material constitutive; adiabatic shear

Received: 2015-06-30; **Accepted:** 2015-09-06; **Published online:** 2015-09-16 19:37

URL: www.cnki.net/kcms/detail/11.2625.V.20150916.1937.001.html

Foundation items: the Fundamental Research Funds for the Central Universities (14ZD37); National Natural Science Foundation of China (51301068); Natural Science Foundation of Hebei Province (E2014502003)

* **Corresponding author.** Tel.: 0312-7525045 E-mail: wjf266@163.com

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0488



动态总方差设计及其快速算法

汪立新^{1,*}, 李灿¹, 姜周², 朱战辉¹, 田颖¹

(1. 第二炮兵工程大学 控制科学与工程系, 西安 710025; 2. 北京航空航天大学 能源与动力工程学院, 北京 100083)

摘 要: Allan 方差法分析随机误差时存在 2 点缺陷, 一是长相关时间下估计值震荡较大, 二是无法跟踪信号的动态变化。本文融合总方差法和动态 Allan 方差法的思想和优势, 提出了动态总方差法。首先使用窗函数截取原始数据, 然后对窗内数据进行延拓, 对延拓后的数据进行总方差分析, 得到原始信号的局部随机特性, 随着窗函数的滑动可以得到原始信号的随机变化特性。经验证, 动态总方差法同时解决了 Allan 方差法存在的 2 个缺陷。最后设计了半球谐振陀螺 (HRG) 的线振动试验验证此算法的有效性, 分析结果表明动态总方差法在分析精度和使用的数据量上具有优势。但是动态总方差法又存在计算量大、分析速度慢的问题, 因此本文又推导出了动态总方差法的递推公式, 从而给出了动态总方差的快速算法。

关 键 词: Allan 方差法; 总方差法; 动态总方差法; 递推公式; 快速算法

中图分类号: V241.5

文献标识码: A

文章编号: 1001-5965(2016)07-1352-09

陀螺仪在平台稳定系统中起到传感器的作用, 是平台稳定系统控制信号的起始端, 因此陀螺的测量精度会影响整个平台系统的稳定精度。陀螺仪的测量结果中不可避免地含有随机误差, 目前 Allan 方差法是分析陀螺随机误差的标准工具^[1]。但是 Allan 方差估计值在长相关时间时误差大、震荡剧烈, 而且 Allan 方差法无法分析陀螺随机误差系数随时间的变化特性。

目前解决 Allan 方差法估计误差大、震荡剧烈的理论有很多, 包括 #1 理论方差^[2]、Théoh 方差^[3]、总方差^[4]和重叠 Allan 方差法^[5]等。其中总方差法是通过镜像映射的方法对数据进行延拓, 新数据序列约为原始序列的 3 倍, 增加了数据自由度和估计值稳定性, 很好地解决了 Allan 方差法估计值震荡剧烈的问题。

动态 Allan 方差法^[6-8]使用窗函数截取数据, 再对窗内数据进行 Allan 方差分析, 得到原始信

号局部的随机误差特性, 随着窗函数的移动就可以观测到整个数据段的随机误差特性。这就相当于在原始数据上使用了显微镜来观察其局部特性, 然后在时间轴上滑动“显微镜”得到随机误差特性随时间的动态变化。

本文融合总方差法延拓数据和动态 Allan 方差法加窗的思想, 提出动态总方差法。首先使用窗函数对原始随机序列进行分割, 然后对窗内数据进行镜像映射来延拓数据, 再对延拓后数据进行总方差分析, 得到此时总方差值和各随机误差组分系数。随着窗函数的移动, 就可得到总方差的二维分布图和各随机误差系数随时间的变化规律。最后针对动态总方差法计算量大、分析速度慢的问题, 又推导了动态总方差法的递推公式和初值计算式, 从而给出了动态总方差法的快速算法。

收稿日期: 2015-07-20; 录用日期: 2015-10-30; 网络出版时间: 2015-12-17 10:41

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151217.1041.008.html

基金项目: 二炮装备技术基础项目 (EP114054)

* 通讯作者: E-mail: wlxxian@sina.com

引用格式: 汪立新, 李灿, 姜周, 等. 动态总方差设计及其快速算法 [J]. 北京航空航天大学学报, 2016, 42(7): 1352-1360.
WANG L X, LI C, JIANG Z, et al. Design of dynamic total variance and its fast algorithm [J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1352-1360 (in Chinese).

1 Allan 方差及陀螺随机误差分析

1.1 Allan 方差法原理

本节主要从 Allan 方差的定义、误差和功率谱关系等 3 个方面对 Allan 方差法进行分析。Allan 方差法就是不断改变数据采集频率的一个过程,假设原始数据的采样周期为 T_s ,共采集了 N 个数据,则采集时间为 $N \cdot T_s$,记原始数据序列为 $\{\Omega_i\}, i=1, 2, \dots, N$ 。将采集频率增大为 $m \cdot T_s$,那么就需要对相邻的 m 个数据取平均值作为更改采样频率后的输出,而且求平均值时数据不允许重叠,那么改变采样频率后的数据长度为 $K = \frac{N}{m}$,记此时的输出为 $\bar{\Omega}_k$,则

$$\bar{\Omega}_k(m) = \frac{1}{m} \sum_{i=1}^m \Omega_{(k-1)m+i} \quad (1)$$

式中: k 为数据重新分组后 $\bar{\Omega}$ 的序列。

则 Allan 方差的计算式为

$$\sigma_A^2(\tau) = \frac{1}{2(K-1)} \sum_{j=1}^{K-1} [\bar{\Omega}_{j+1}(m) - \bar{\Omega}_j(m)]^2 \quad (2)$$

式中: $\tau = m \cdot T_s$ 为相关时间, m 为平均因子; j 为 $\bar{\Omega}(m)$ 的序列号。需要强调的是式(2)给出的是 Allan 方差计算式,而不是定义式。那么计算值相对于真实值的相对误差为

$$\delta_A = \frac{1}{\sqrt{2(P-1)}} \times 100\% \quad (3)$$

式中: P 为重新采样后的输出数据数量,也可以理解为原始数据进行分组得到的组数;而 $P-1$ 为 Allan 方差估计的独立数,由式(3)可以看出,Allan 方差计算值的相对误差取决于 P 值,当采样频率过小(即输出数据量 P 很小)时,估计值相对误差 δ_A 就会很大,这就是长相关时间下误差大、震荡剧烈的原因。

参考文献[9]可知,Allan 方差与功率谱密度的关系式为

$$\sigma_A^2(\tau) = 4 \int_0^{+\infty} S_\Omega(f) \frac{\sin^4(\pi f \tau)}{(\pi f \tau)^2} df \quad (4)$$

式中: $S_\Omega(f)$ 为功率谱密度, f 为信号频率。可以将 $\frac{\sin^4(\pi f \tau)}{(\pi f \tau)^2}$ 看成一个带通滤波器,滤波器的带通取决于相关时间 τ 。由式(4)可以看出,将原始数据通过带通滤波器后,Allan 方差值与信号的能量成正比,因此可以根据 Allan 方差值分析不同频段信号的含量。

1.2 陀螺随机误差

一般来讲,将陀螺随机误差^[10-11]分为角度随

机游走(Angular Random Walk, ARW)、零偏不稳定性(Bias Instability, BI)、速率随机游走(Rate Random Walk, RRW)、速率斜坡(Rate Ramp, RR)和量化噪声(Quantization Noise, QN)。为了避免重复,只简要介绍其功率谱密度。

1) 角度随机游走

角度随机游走的功率谱密度表达式为

$$S_\Omega^{\text{ARW}}(f) = V^2 \quad (5)$$

式中: V 为角度随机游走系数,式(5)代入式(4)可得其 Allen 方差为

$$\sigma_{\text{ARW}}^2 = \frac{V^2}{\tau} \Rightarrow \lg \sigma_{\text{ARW}} = -\frac{1}{2} \lg \tau + \lg V \quad (6)$$

容易看出在 Allan 方差的双对数曲线中,斜率为 $-\frac{1}{2}$ 的噪声为角度随机游走。

2) 零偏不稳定性

零偏不稳定性的功率谱密度表达式为

$$S_\Omega^{\text{BI}}(f) = \begin{cases} \frac{B^2}{2\pi f} & f \leq f_0 \\ 0 & f > f_0 \end{cases} \quad (7)$$

式中: B 为零偏不稳定性系数,式(7)代入式(4)可得其 Allen 方差为

$$\sigma_{\text{BI}}^2 \approx (0.6643B)^2 \quad f \leq f_0 \quad (8)$$

容易看出此时 Allan 方差为一常值,所以在 Allan 方差的双对数曲线中,斜率为 0 的噪声为零偏不稳定性。

3) 速率随机游走

速率随机游走的功率谱密度表达式为

$$S_\Omega^{\text{RRW}}(f) = \left(\frac{U}{2\pi f}\right)^2 \quad (9)$$

式中: U 为速率随机游走系数,式(9)代入式(4)可得其 Allen 方差为

$$\sigma_{\text{RRW}}^2 = \frac{U^2 \tau}{3} \quad (10)$$

容易看出在 Allan 方差的双对数曲线中,斜率为 $\frac{1}{2}$ 的噪声为速率随机游走。

4) 速率斜坡

速率斜坡的功率谱密度表达式为

$$S_\Omega^{\text{RR}}(f) = \left(\frac{R}{2\pi f}\right)^2 \quad (11)$$

式中: R 为速率斜坡系数,式(11)代入式(4)可得其 Allen 方差为

$$\sigma_{\text{RR}}^2 = \frac{R^2 \tau^2}{2} \quad (12)$$

容易看出在 Allan 方差的双对数曲线中,斜率为 1 的噪声为速率斜坡。

5) 量化噪声

量化噪声的功率谱密度表达式为

$$S_n^{QN}(f) = \begin{cases} \tau_0 Q^2 \left[\frac{\sin^2(\pi f \tau_0)}{(\pi f \tau_0)^2} \right] & f \geq \frac{1}{2\tau_0} \\ \tau_0 Q^2 & f < \frac{1}{2\tau_0} \end{cases} \quad (13)$$

式中: Q 为量化噪声系数, 式(13)代入到式(4)可得其 Allen 方差为

$$\sigma_{QN}^2 = \frac{3Q^2}{\tau^2} \quad (14)$$

容易看出在 Allen 方差的双对数曲线中, 斜率为 -1 的噪声为速率斜坡。

1.3 Allan 方差法的缺陷

为了找出 Allan 方差的缺陷, 分别仿真出平稳白噪声和方差渐大的白噪声如图 1 所示, 然后使用 Allan 方差法进行分析, 分析结果如图 2 所示。

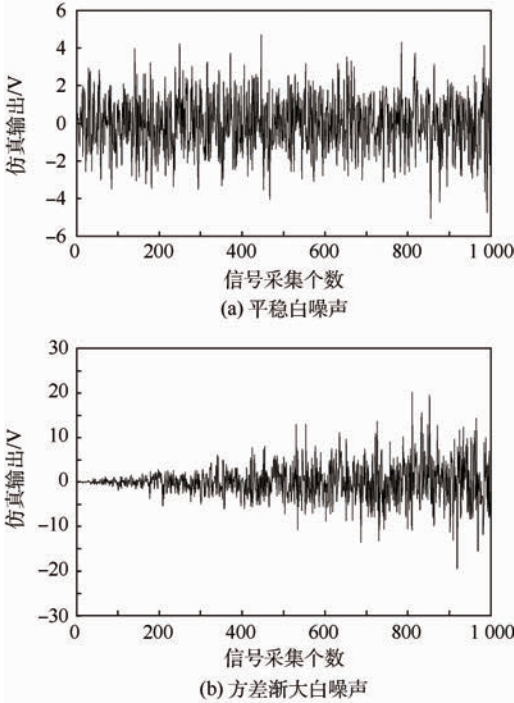


图 1 平稳白噪声和方差渐大白噪声

Fig. 1 Stationary white noise and white noise with gradually bigger variance

对比图 2 的 Allan 方差分析结果, 可以分析出 Allan 方差的 2 点缺陷:

1) 在长相关时间下, 由于 Allan 方差估计的自由度比较小, 导致 Allan 方差估计值置信度比较低, 相对误差比较大, 震荡非常严重。

2) 平稳白噪声和方差渐大白噪声分析结果相同, 说明 Allan 方差无法跟踪随机序列的动态变化(即无法跟踪信号的方差变化)。

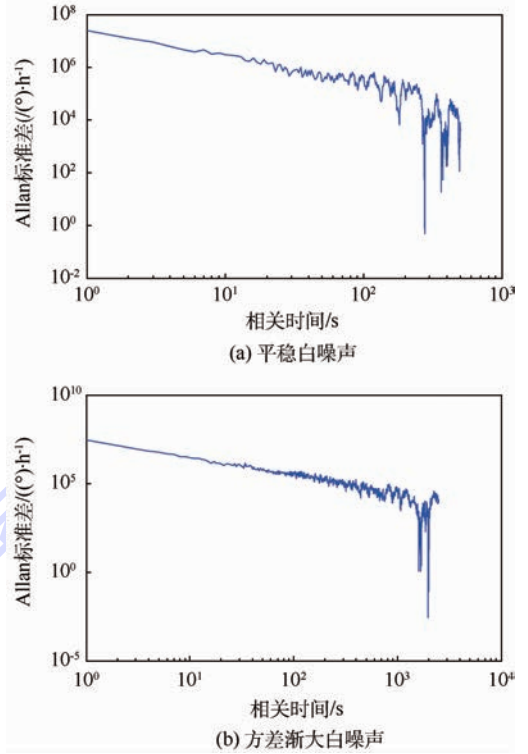


图 2 平稳白噪声和方差渐大白噪声 Allan 方差分析
Fig. 2 Allan variance analysis of stationary white noise and white noise with gradually bigger variance

2 总方差法原理分析

Allan 方差法存在的第 1 个缺陷是由于估计值的自由度比较小, 那么很自然的一个想法就是通过数据延拓的方法增加数据的自由度。总方差(total variance)法就是通过镜像映射将原始数据进行延长, 达到增加估计值置信度、稳定估计值的效果。

假设原始数据序列为 $\{\Omega_i\}$, 序列长度为 N , 通过镜像映射后的序列为 $\{\Omega_i^*\}$, 其映射规则^[12]为

$$\begin{cases} \Omega_i^* = \Omega_{1-i} & i = 2 - N, 3 - N, \dots, 0 \\ \Omega_i^* = \Omega_i & i = 1, 2, \dots, N \\ \Omega_i^* = \Omega_{2N-i} & i = N + 1, N + 2, \dots, 2N - 1 \end{cases} \quad (15)$$

延拓后的序列长度为 $3N - 2$, 约为原始数据的 3 倍。

总方差的计算式^[13]为

$$\sigma_{tot}^2(\tau) = \frac{1}{2(N-1)} \sum_{n=2}^N (\Omega_n^* - \Omega_{n-m}^*)^2 \quad (16)$$

对比式(16)与式(1)和式(2)可以看出, 总方差法没有重新采样的过程, 而是直接使用原始数据, 使数据的利用效率增加了。在总方差中, τ 的最大值为 $N - 1$, 在 Allan 方差中, τ 的最大值为

$\frac{N}{2}$,说明总方差法的分析范围约为 Allan 方差法的 2 倍。

总方差的估计值相对误差^[13]为

$$\delta_{\text{tol}} = \frac{1}{\sqrt{2F(\tau)}} \times 100\% \quad (17)$$

式中: $F(\tau)$ 为相关时间为 τ 时的等效自由度,其计算式为

$$F(\tau) = b \cdot \frac{T}{\tau} - c \quad (18)$$

其中: T 为数据采集的总时间, $\frac{T}{\tau}$ 可以理解为独立组数; b 、 c 为偏差系数,其值与随机噪声的组分有关,表 1 给出了 3 种随机误差组分的偏差系数。

表 1 3 种随机误差组分的偏差系数

Table 1 Deviation coefficients of three random error components

随机误差组分	b	c	$F(\tau)$
零偏不稳定性	3/2	0	3.000
角度随机游走	$24(\ln 2)^2/\pi^2$	0.222	2.097
量化噪声	140/151	0.358	1.514

使用式(3)和式(17)计算出 Allan 方差法和总方差法在不同相关时间下的相对误差。为了便于统一和比较,将相关时间进行归一化处理 τ/T ,计算结果如图 3 所示。

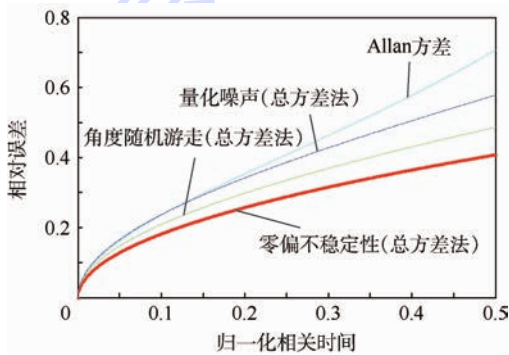


图 3 Allan 方差与总方差相对误差

Fig. 3 Relative error of Allan variance and total variance

从图 3 中可以看出,在相关时间较小的情况下,Allan 方差和总方差法的相对误差相差不大;但是在相关时间较大的情况下,总方差法的相对误差比 Allan 方差法的相对误差小很多;当归一化相关时间为 0.5 时(即在 Allan 方差法中将所有数据分为 2 组),Allan 方差估计值的相对误差为 70.71%,量化噪声的总方差估计值相对误差为 57.81%,角度随机游走的总方差估计值相对误差为 40.82%,零偏不稳定性的总方差估计值相对误差为 48.63%,说明了当归一化相关时间为 0.5 时,总方差法的估计值依然有效,其精度远

远高于 Allan 方差法。

同时使用 Allan 方差法和总方差法对图 1(b)所示的随机序列进行分析,分析结果如图 4 所示,从图中可以看出,总方差法在长相关时间下依然保持了较好的稳定性,解决了 Allan 方差法在长相关时间下震荡较大的问题。

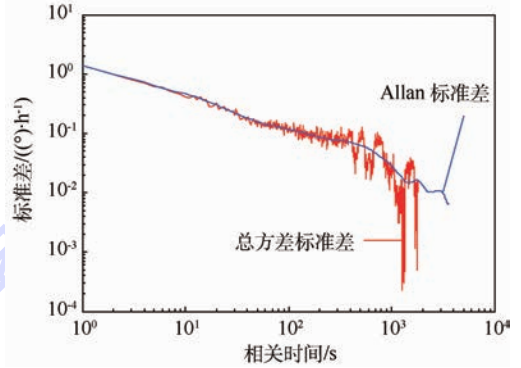


图 4 随机序列的 Allan 方差和总方差分析

Fig. 4 Allan variance and total variance analysis of random sequence

3 动态 Allan 方差法分析

Allan 方差法的第 2 个缺陷是无法跟踪信号的非平稳变化,那么可以用窗函数的方法,逐段截取数据对其进行 Allan 方差分析,窗函数就像显微镜一般,显微出原始数据的局部特性,然后将这些特性在时间轴上展开就可以得到随机特性随时间的变化,这就是动态 Allan 方差法的理论基础。动态 Allan 方差法的具体实现步骤^[14-15]为

- 1) 确定原始数据的分析时间起点 t_1 和窗函数宽度 W ,要求 $t - \frac{W}{2} \leq t_1 \leq t + \frac{W}{2}$ 。
- 2) 用中心点为 t_1 、宽度为 W 的窗函数截断原始信号 $\Omega(t)$ 。
- 3) 计算窗口内数据的 Allan 方差值 $\sigma_A^2(t_1, \tau)$,并使用最小二乘拟合的方法得到此时的随机误差系数。
- 4) 将窗口滑动到下一个时刻 t_2 ,要求以 t_2 为中心的窗要与以 t_1 为中心的窗互相重叠。重复步骤 2) 和步骤 3),得到此时窗口内的 Allan 方差值 $\sigma_A^2(t_2, \tau)$ 和随机误差系数。
- 5) 以此类推,重复上述过程,得到 Allan 方差值的集合 $\{\sigma_A^2(t_i, \tau)\} (i=1, 2, \dots, n)$ 和随机误差系数序列。以 t 和 τ 为变量,将 Allan 方差值的变化展现在三维分布图上,就得到了动态 Allan 分析的结果;将各时刻的随机误差系数在时间轴上展开,得到随机误差特性的动态变化过程。

使用动态 Allan 方差法分别分析图 1(a)和图 1(b)中的 2 种信号,分析结果如图 5 所示。图中图形灰度的变化表示方差大小的变化,从图 5(b)中可以看出,随着原始信号方差的增大,动态 Allan 方差的灰度越来越深,可以很好地跟踪原始数据的动态变化。说明动态 Allan 方差法解决了 Allan 方差的第 2 个缺陷。

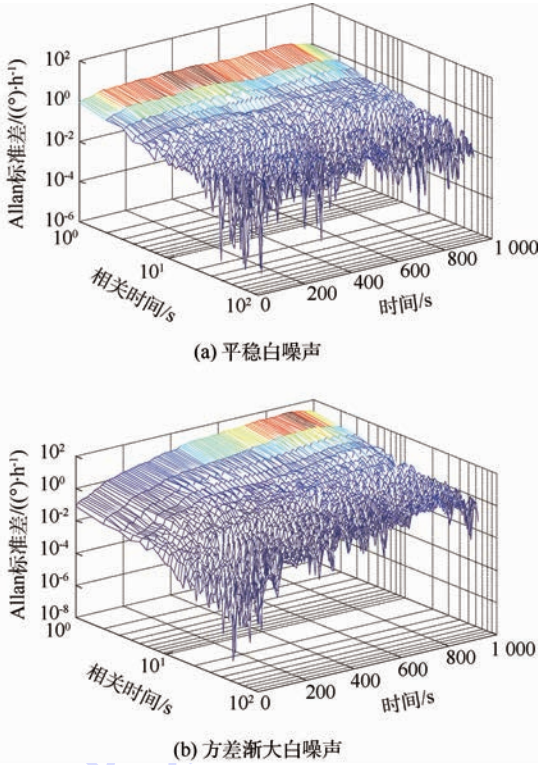


图 5 平稳白噪声和方差渐大白噪声动态 Allan 方差分析
Fig. 5 Dynamic Allan variance analysis of stationary white noise and white noise with gradually bigger variance

4 动态总方差法设计

总方差法解决了 Allan 方差法长相关时间下震荡较大的问题;动态 Allan 方差法解决了 Allan 方差法无法跟踪随机信号动态变化的缺陷。结合 2 种算法的优势,本文提出了动态总方差法,其基本思想是:首先利用窗函数截取原始数据,然后对窗口内的数据进行镜像映射来延拓数据,再对延拓后的数据计算其总方差值,最终得到原始数据的动态总方差图和随机误差系数随时间的变化规律。

动态总方差法的具体实现步骤包括:

- 1) 确定原始数据的分析起点 t_1 和窗口宽度 L 。
- 2) 用中心点为 t_1 、宽度为 L 窗函数截取原始信号 $\Omega(t)$,得到窗口截断信号 $\Omega_T(t_1)$,支撑变量描述窗内渐逝的时间,记为 t' ,则

$$t - L/2 \leq t' \leq t + L/2 \quad (19)$$

记窗函数为 $P_w(t')$,则窗内信号为

$$\Omega_T(t_1, t') = \Omega(t')P_w(t_1 - t') \quad (20)$$

3) 将窗内信号 $\Omega_T(t_1, t')$ 按式(15)给出的方式进行镜像映射,得到延拓序列 $\Omega_T^*(t_1, t')$ 。

4) 按照式(16)给出的总方差法计算式计算得到 t_1 时刻的总方差值 $\sigma_{tol}^2(t_1, \tau)$,同时使用最小二乘拟合得到 t_1 时刻的随机误差系数。

5) 将窗口滑动到 t_2 ,重复步骤 2) ~ 步骤 4) 得到 t_2 时刻的动态总方差值 $\sigma_{tol}^2(t_2, \tau)$ 和随机误差系数。以此类推,得到动态总方差法序列 $\{\sigma_{tol}^2(t_i, \tau) | (i = 1, 2, \dots, N)$ 和各时刻的随机误差系数。

6) 将动态总方差值展示在 $\sigma_{tol}-\tau-t$ 三维图中来分析信号的非平稳特征;将随机误差系数在时间轴上展开得到随机误差随时间的变化规律。

离散情况下,第 n 个窗截得的数据序列动态总方差的计算公式为

$$\sigma_{DT}^2[n, M] = \frac{1}{2(W-1)} \cdot \sum_{j=n-\frac{W}{2}+1}^{n+\frac{W}{2}} (\bar{\Omega}_j^*[M] - \bar{\Omega}_{j-M}^*[M])^2 \quad (21)$$

式中: $\bar{\Omega}_j^*[M]$ 为第 j 个 $\bar{\Omega}^*[M]$; $\bar{\Omega}_{j-M}^*[M]$ 为第 $j-M$ 个 $\bar{\Omega}^*[M]$ 。

5 动态试验验证

为了验证所设计的动态总方差法的有效性,本文设计了半球谐振陀螺的线振动试验。线振动的初始加速度为 $5g$,采用弦函数方式施加在陀螺上。陀螺数据采集的频率为 10 kHz ,共采集了 4200 个数据。首先去除输出信号中的趋势项而得到随机信号,如图 6 所示,其动态总方差法分析如图 7 所示。

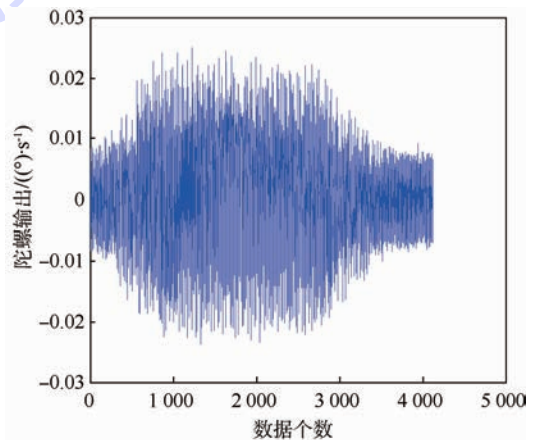


图 6 原始随机信号

Fig. 6 Original random signal

从图 7 中可以看出灰度的变化,这是由于原始数据方差的变化引起的,说明动态总方差法可以跟踪信号的非平稳性变化。为了观察信号中随机误差组分的变化,分别使用动态 Allan 方差法和动态总方差法分析原始数据中的量化噪声系数 Q 、角度随机游走系数 N 、零偏不稳定性 B ,结果如图 8 和图 9 所示,图中 t 为时间。

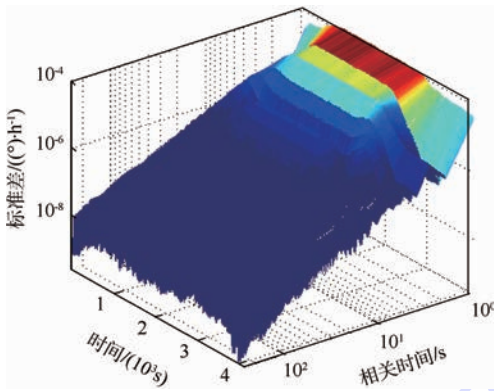


图 7 陀螺输出的动态总方差分析

Fig. 7 Dynamic total variance analysis of gyro output

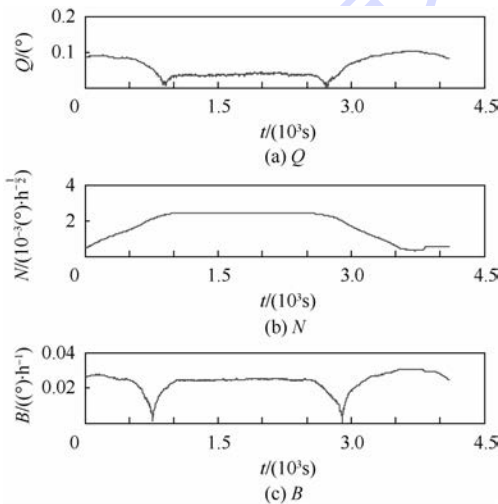


图 8 动态 Allan 方差分析结果

Fig. 8 Results of dynamic Allan variance analysis

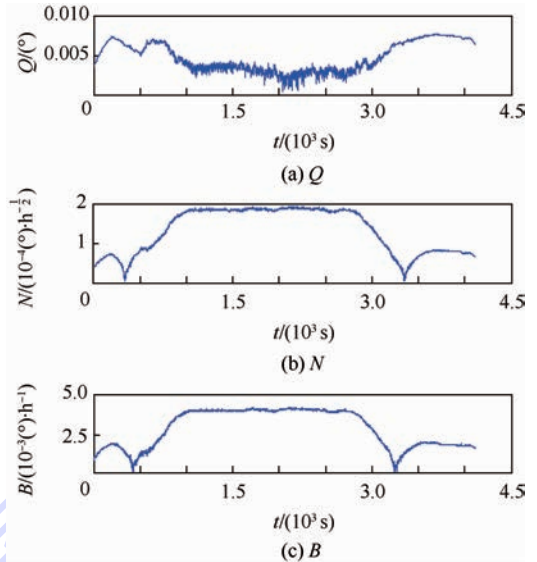


图 9 动态总方差法分析结果

Fig. 9 Results of dynamic total variance analysis

对比图 8 和图 9 的处理结果可以得到:

- 1) 从数量级上来讲,动态总方差法分析精度比动态 Allan 方差法至少高一个数量级。
- 2) 角度随机游走和零偏不稳定性系数在振动过程中增大,而量化噪声系数在振动过程中减小,说明周期性振动弥补了量化噪声的周期性误差。

为了对比动态总方差法和动态 Allan 方差法的数据处理精度,选取 1 000 ~ 2 600 段数据,共 1 600 个。对这 1 600 个数据进行 Allan 方差分析,然后使用长度为 401 的窗口分割数据,分别进行动态 Allan 方差和动态总方差分析。由于 Allan 方差分析使用了 1 600 个点,具有数据量上的绝对优势,所以可以将其作为评价动态 Allan 方差法和动态总方差法的一个标准值。处理结果如表 2 所示。

表 2 振动条件下陀螺噪声系数分析结果

Table 2 Noise coefficient analysis result of gyro in vibration condition

处理算法	Q	N	B	V	R
Allan 方差(1600)	3.8×10^{-3}	1.8×10^{-4}	3.8×10^{-3}	2.8×10^{-2}	6.7×10^{-2}
动态总方差(401)	3.2×10^{-3}	1.9×10^{-4}	3.9×10^{-3}	2.9×10^{-2}	7.5×10^{-2}
动态 Allan(401)	4.8×10^{-2}	2.5×10^{-3}	3.7×10^{-2}	0.4	1.2

以表 2 中 Allan 方差法得到的结果作为标准,可以看出使用 401 点的动态总方差法处理结果与标准结果精度一致,说明在同等精度下,动态总方差法使用的数据约为 Allan 方差法的 1/4;比较动态 Allan 方差法和动态总方差法处理结果可以看出,使用同等数据量时,动态总方差法达到的精度比动态 Allan 方差法高一个数量级。

6 动态总方差快速算法

从上面的分析可以看出,动态总方差法同时解决了 Allan 方差法存在的 2 个缺陷,而且在分析精度和使用数据量上也有了很大提高。但是动态总方差法使用窗函数截取数据,需要对每个窗函数内的数据进行总方差分析,计算量非常大,算

法运行需要的时间特别长。针对这个问题,本文推导了基于递推公式的快速算法。

从第 n 个窗内数据的动态总方差计算式开始推导,其计算式为式(21),则第 $n+1$ 个窗内数据的动态总方差计算式为

$$\sigma_{\text{DT}}^2[n+1, M] = \frac{1}{2(W-1)} \cdot \sum_{j=n-\frac{W}{2}+2}^{n+1+\frac{W}{2}} (\bar{\Omega}_j^*[M] - \bar{\Omega}_{j-M}^*[M])^2 \quad (22)$$

将式(22)变形为

$$\begin{aligned} \sigma_{\text{DT}}^2[n+1, M] &= \frac{1}{2(W-1)} \cdot \sum_{j=n-\frac{W}{2}+2}^{n+1+\frac{W}{2}} (\bar{\Omega}_j^*[M] - \bar{\Omega}_{j-M}^*[M])^2 = \\ &= \frac{1}{2(W-1)} \sum_{j=n-\frac{W}{2}+1}^{n+\frac{W}{2}} (\bar{\Omega}_j^*[M] - \bar{\Omega}_{j-M}^*[M])^2 + \\ &= \frac{1}{2(W-1)} \cdot [(\bar{\Omega}_{n+1+\frac{W}{2}}^*[M] - \bar{\Omega}_{n+1+\frac{W}{2}-M}^*[M])^2 - \\ &(\bar{\Omega}_{n-\frac{W}{2}+1}^*[M] - \bar{\Omega}_{n-\frac{W}{2}+1-M}^*[M])^2] \quad (23) \end{aligned}$$

式中:第2个等号后的第1项即为 $\sigma_{\text{DT}}^2[n, M]$,所以动态总方差法的递推公式为

$$\begin{aligned} \sigma_{\text{DT}}^2[n+1, M] &= \frac{1}{2(W-1)} \sum_{j=n-\frac{W}{2}+2}^{n+1+\frac{W}{2}} (\bar{\Omega}_j^*[M] - \bar{\Omega}_{j-M}^*[M])^2 = \\ &= \sigma_{\text{DT}}^2[n, M] + \frac{1}{2(W-1)} \cdot \\ &= [(\bar{\Omega}_{n+1+\frac{W}{2}}^*[M] - \bar{\Omega}_{n+1+\frac{W}{2}-M}^*[M])^2 - \\ &(\bar{\Omega}_{n-\frac{W}{2}+1}^*[M] - \bar{\Omega}_{n-\frac{W}{2}+1-M}^*[M])^2] \quad (24) \end{aligned}$$

此递推公式的初始项即为第一个窗口的动态总方差值,设其为第 n_0 个窗函数,其值需要按照定义式进行计算,即为

$$\sigma_{\text{DT}}^2[n_0, M] = \frac{1}{2(W-1)} \cdot \sum_{j=n_0-\frac{W}{2}+1}^{n_0+\frac{W}{2}} (\bar{\Omega}_j^*[M] - \bar{\Omega}_{j-M}^*[M])^2 \quad (25)$$

式(25)和式(24)分别给出了动态总方差快速算法的初始值和递推公式。式(24)和式(25)组合在一起就是动态总方差法的快速算法。

通过对比总方差法的定义式和递推公式可以看出,按照定义式进行计算时,每个窗口需要计算 $W+1$ 次乘法、 $2W-1$ 次加法;而按照递推公式进行计算时,每个窗口需要计算 3 次乘法、4 次加

法。而整个分析过程共有 $N-W+1$ 个窗口,那么快速算法就会大大节约算法的运行时间。

为了对上述分析结果进行验证,将窗函数固定为 $W = \frac{N}{10}$,使用同一台计算机分别处理数据长度为 10^2 、 10^3 和 10^4 的数据序列,统计定义式算法和快速算法的运行时间,结果如表3所示。

表3 传统算法与快速算法运行时间

Table 3 Runtime of traditional and fast algorithms

数据长度	运行时间/s	
	传统算法	快速算法
10^2	3.0×10^{-3}	3.1×10^{-4}
10^3	0.30	4.5×10^{-3}
10^4	118	0.25

从表3给出的结果可以看出,快速算法相对于定义式算法大大节约了运行时间;而且随着数据量的增加,2种算法运行的时间相差也越来越大,快速算法的优势也越明显。

7 结论

1) 本文设计了动态总方差法,当使用相同的数据量进行分析时,动态总方差法相对于动态Allan方差法的分析精度至少提高了一个数量级;当分析精度一致时,动态总方差法需要的数据量是Allan方差法的1/4。

2) 本文设计了动态总方差法的快速算法,对比快速算法和定义式算法的运行时间可以看出,快速算法极大地减少了算法的运行时间,而且原始数据量越大,则快速算法的优势越明显。

参考文献 (References)

- [1] Gyro and Accelerometer Panel of the IEEE Aerospace and Electronic Systems Society. IEEE standard specification format guide and test procedure for single-axis interferometric fiber optic gyros: IEEE STD 952TM-1997 [S]. Piscataway, NJ: IEEE Standards Board, 1998.
- [2] 程旭维,汤霞清,黄湘远.基于#1理论方差的光学陀螺长期随机误差分析[J].中国激光,2014,41(10):146-153. CHENG X W, TANG X Q, HUANG X Y. Investigation on random error properties of optic gyroscope based on theoretical variance #1 [J]. Chinese Journal of Lasers, 2014, 41(10): 146-153 (in Chinese).
- [3] HOWE D A, Théo H: A hybrid, high-confidence statistic that improves on the Allan deviation [J]. Metrologia, 2006, 43: 322-331.
- [4] HOWE D A. The total deviation approach to long-term characterization of frequency stability [J]. IEEE Transactions on Ultrasonic, Ferroelectrics, and Frequency Control, 2000, 47(5): 1102-1110.
- [5] RILEY W J. Techniques for frequency stability analysis [C] //

- IEEE International Frequency Control Symposium. Piscataway, NJ: IEEE Press, 2003.
- [6] GALLEANI L. The dynamic Allan variance III: Confidence and detection surfaces [J]. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 2011, 58(8) : 1550-1558.
- [7] DOBROGOWSKI A, KASZANIA M. Real-time assessment of dynamic Allan deviation and dynamic time deviation [C] // Proceeding of the 2012 European Frequency and Time Forum. Piscataway, NJ: IEEE Press, 2012: 247-252.
- [8] GALLEANI L, TAVELLA P. The dynamic Allan variance [J]. IEEE Transaction on Ultrasonics, Ferroelectrics, and Frequency Control, 2009, 56(3) : 450-460.
- [9] 李颖. 慢光光纤陀螺关键技术研究 [D]. 哈尔滨: 哈尔滨工业大学, 2009.
- LI Y. Key techniques for the slow light fiber optic gyroscope [D]. Harbin: Harbin Institute of Technology, 2009 (in Chinese).
- [10] 吕琳, 全伟. 基于 GP + GA 的陀螺随机误差建模分析 [J]. 北京航空航天大学学报, 2015, 41(6) : 1135-1140.
- LYU L, QUAN W. Modeling and analysis of gyroscope's random drift based on GP + GA method [J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(6) : 1135-1140 (in Chinese).
- [11] 祁家毅, 任顺清, 冯士伟, 等. 半球谐振陀螺随机误差分析 [J]. 中国惯性技术学报, 2009, 17(1) : 98-101.
- QI J Y, REN S Q, FENG S W, et al. Random error analysis of hemispherical resonator gyro [J]. Journal of Chinese Inertial Technology, 2009, 17(1) : 98-101 (in Chinese).
- [12] 韩军良, 葛升民, 沈毅. 基于总方差方法的光纤陀螺随机误差特性研究 [J]. 哈尔滨工业大学学报, 2007, 39(5) : 708-711.
- HAN J L, GE S M, SHEN Y. Research on the random error properties of FOG based on total variance [J]. Journal of Harbin Institute of Technology, 2007, 39(5) : 708-711 (in Chinese).
- [13] 石国祥, 陈坚, 叶军, 等. 总方差方法在光纤陀螺随机误差分析中的应用 [J]. 光电工程, 2012, 39(1) : 63-67.
- SHI G X, CHEN J, YE J, et al. Applications of total variance method in random error analysis of the fiber optic gyro signal [J]. Opto-Electronic Engineering, 2012, 39(1) : 63-67 (in Chinese).
- [14] 李冀辰, 高凤岐, 王广龙, 等. 光纤陀螺振动和变温条件下的 DAVAR 分析 [J]. 中国激光, 2013, 40(9) : 184-190.
- LI J C, GAO F Q, WANG G L, et al. Analysis of dynamic allan variance for fiber optic gyro under vibration and variable temperature conditions [J]. Chinese Journal of Laser, 2013, 40(9) : 184-190 (in Chinese).
- [15] ZHANG C X, WANG L, GAO S, et al. Dynamic Allan variance analysis for stochastic errors of fiber optic gyroscope [J]. Infrared and Laser Engineering, 2014, 43(9) : 3081-3088.

作者简介:

汪立新 男, 教授。主要研究方向: 惯性技术、组合导航以及微小信号测试与处理。

E-mail: wlxxian@sina.com

李灿 男, 硕士研究生。主要研究方向: 惯性系统及测试。

E-mail: 15339020629@163.com

姜周 男, 硕士研究生。主要研究方向: 能源与动力。

E-mail: sy1404110@buaa.edu.cn

朱战辉 男, 博士研究生。主要研究方向: 惯性系统及测试。

E-mail: 18049637941@163.com

杨浩天 男, 硕士研究生。主要研究方向: 惯性系统及测试。

E-mail: 15191918796@163.com

Design of dynamic total variance and its fast algorithm

WANG Lixin^{1,*}, LI Can¹, JIANG Zhou², ZHU Zhanhui¹, TIAN Ying¹

(1. Department of Control Science and Engineering, The Second Artillery Engineering University, Xi'an 710025, China;

2. School of Energy and Power Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: There exist two shortcomings in Allan variance approach when it is used to analyze random error. First, the vibration of estimated value is large in long-term correlation. Second it can't follow the dynamic change of signal. This paper integrates the thought and advantage of total variance and dynamic allan variance, and proposes the dynamic total variance approach. Firstly, original data is truncated by window function. Secondly, we continue the data in the window. Then the data after continuation is analyzed by total variance, so that local random characteristic is acquired. With the window sliding, the change of random characteristic of original signal with time is obtained. Through verification, the dynamic total variance can solve the two shortcomings of Allan variance approach. Eventually, linear vibration trial of hemispherical resonator gyro (HRG) is designed to clarify this new algorithm, which indicates that dynamic total variance possesses advantages in analysis accuracy and data quantity that it used. However, this new algorithm has the problem of excessive calculation and low speed of analysis, so we derive recursive formula of it, based on which fast algorithm is provided.

Key words: Allan variance approach; total variance algorithm; dynamic total variance algorithm; recursive formula; fast algorithm

Received: 2015-07-20; **Accepted:** 2015-10-30; **Published online:** 2015-12-17 10:41

URL: www.cnki.net/kcms/detail/11.2625.V.20151217.1041.008.html

Foundation item: Equipment Technical Infrastructure Projects of the Second Artillery(EP114054)

* **Corresponding author.** E-mail: wlxxian@sina.com

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0538

基于 CCR 带并行腔双集束型设备调度方法



周炳海*, 黎明

(同济大学 机械与能源工程学院, 上海 201804)

摘 要: 为有效解决 450 mm 晶圆制造中带有并行处理腔的集束型设备群调度问题, 提出了基于产能约束资源 (CCR) 的调度方法。首先, 综合考虑多品种加工、晶圆驻留和资源约束等特征, 以系统总完工时间最小作为目标, 建立带并行腔双集束型设备调度数学模型; 其次, 为优化机械手作业顺序, 对 CCR 实施锁定-收紧-松弛 (LTL) 策略, 构建了一种以 CCR 为界的分段调度算法; 最后, 进行了仿真实验分析。结果表明本文提出的算法是有效和有竞争性的。

关键词: 集束型设备群; 并行腔; 驻留约束; 产能约束资源; 调度算法

中图分类号: F270; TP391

文献标识码: A

文章编号: 1001-5965(2016)07-1361-07

集束型设备群是半导体制造系统的核心和关键设备, 主要用于晶圆的蚀刻、光刻以及离子注射等处理过程。由于其调度问题的重要性以及复杂性^[1], 近年来受到了业界和学术界的广泛关注。

目前关于集束型设备群调度研究尚处于初始阶段。大多数有关其调度文献集中于同种晶圆在稳态模式下的周期性调度的研究。Yi 等^[2]将集束型设备群调度问题分解为多个单集束型设备调度问题建模分析, 并求解出晶圆无驻留约束条件下调度的基本周期下界 (lower-bound fundamental period)。Chan 等^[3-4]将机械手搬运时间假定为固定常数, 对 2 种不同结构的集束型设备群进行研究, 以最小循环周期为目标, 提出了一种基于资源的集束型设备群分析建模方法。Ding 等^[5]利用关键路径松弛法对具有固定动作顺序的单品晶圆集束型设备群调度进行研究, 并给出了具有固定动作顺序机械手调度方案, 但该方法不能保证其松弛调度结果最优。Liu 和 Zhou^[6]在单臂集束型设备群中同时考虑晶圆驻留约束和多晶圆流, 提

出了基于时间约束集的逐级回溯算法。文献 [1-6] 均未考虑集束型设备群实际工程应用中普遍存在的带并行处理腔的调度。

并行腔广泛存在于晶圆制造过程加工时间相对较长的处理模块, 通过平衡不同工序之间负荷, 可提高系统有效产出。已有学者已对于带并行腔的单集束型设备调度问题进行了研究^[7-12]。但对于具有多机械手同时作业的集束型设备群, 由于系统中各单集束型设备间的耦合和相互间的依赖增加其调度复杂性, 使得传统单集束型设备的调度方法无法直接应用^[13], 同时目前对此类调度问题的研究鲜有报道, 仅有 Kumar 等^[14]研究了与集束型设备群结构类似但缓冲连接不同的多机械手制造单元的调度, 但并未考虑驻留约束。因此, 对于问题规模更加复杂的此类调度问题亟需探索有效的解决方法。

本文对带并行腔双集束型设备调度问题进行研究, 在问题域中考虑晶圆多品种加工和驻留约束, 建立调度模型, 提出了基于产能约束资源

收稿日期: 2015-08-24; 录用日期: 2015-10-30; 网络出版时间: 2016-01-05 10:36

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20160105.1036.002.html

基金项目: 国家自然科学基金(71471135, 61273035)

* 通讯作者: Tel.: 021-54810409 E-mail: bhzhou@tongji.edu.cn

引用格式: 周炳海, 黎明. 基于 CCR 带并行腔双集束型设备调度方法[J]. 北京航空航天大学学报, 2016, 42(7): 1361-1367. ZHOU B H, LI M. Scheduling method for double-cluster tools with parallel chambers based on capacity constraint resource [J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1361-1367 (in Chinese).

(Capacity Constraint Resource, CCR) 的启发式算法。

1 问题描述

典型带并行腔的双集束型设备如图 1 所示。图中： C_1, C_2 为在双集束型设备群的 2 个单集束型设备； R_1, R_2 为对应的搬运模块； $M_{h,j}^p$ 为第 h 个集束型设备中的第 j 个处理模块； $M_{h,j(k)}^p$ 为模块 $M_{h,j}^p$ 的第 k 个加工腔（若 $k = 1$, 可省略），且 $k \leq K_{h,j}$, 处理模块 $M_{h,j}^p$ 共有 $K_{h,j}$ 个并行腔； $M_{h,j}^b$ 为用于连接 C_h 和 C_{h+1} 的缓冲模块。模型中用坐标集 X 表示晶圆加工路径。 $X = \{x_1, x_2, \dots, x_i, \dots, x_\varphi\}$, $x_i = (h, j)$ 为晶圆第 i 步工序处理模块序号坐标, φ 为晶圆加工工序总数；系统并行腔数集合 $P_n = \{K_1, K_2, \dots, K_i, \dots, K_\varphi\}$, K_i 为 $M_{x_i}^p$ 含有的并行腔数目。

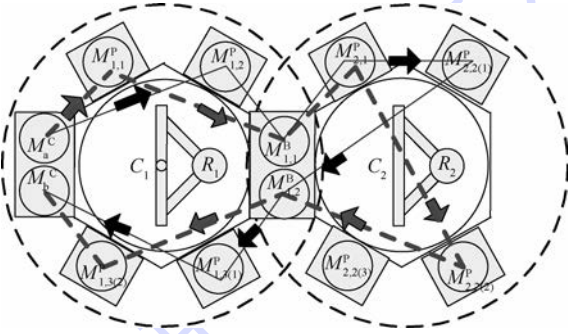


图 1 带并行腔的双集束型设备群示意图
Fig.1 Schematic of double-cluster tools with parallel chambers

晶圆按照其工艺路径从输入卡闸模块 M_a^c 进入系统, 依次经过位于各个集束型设备中的处理模块 (processing module), 连接相邻集束型设备的缓冲模块 (buffer module), 最终加工完成后存放在输出卡闸模块 M_b^c , 晶圆在不同模块间的转移通过搬运模块 (robotic module) 完成。

晶圆调度有关符号与变量定义如下： $S_{x_i}(w)$ 为晶圆 w 进入 $M_{x_i}^p$ 的时间； $L_{x_i}(w)$ 为晶圆 w 离开 $M_{x_i}^p$ 的时间； $B_{h,j}^s(w)$ 为晶圆 w 进入 $M_{h,j}^b$ 的时间； $B_{h,j}^l(w)$ 为晶圆 w 离开 $M_{h,j}^b$ 的时间； t_M 为机械手单次搬运时间, $t_M = t_1 + t_u + t_m$, t_1, t_u 和 t_m 分别为机械手装载、卸载以及移动时间； $P(w)$ 为晶圆 w 在所有处理模块的加工时间集合 $P(w) = \{P_{x_1}(w), P_{x_2}(w), \dots, P_{x_i}(w), \dots, P_{x_\varphi}(w)\}$, 其中 $P_{x_i}(w)$ 为晶圆 w 在 $M_{x_i}^p$ 的加工时间 (若 $P_{x_i}(w) = 0$, 表示晶圆不经过该模块加工, 反之则需经过, 不同种晶圆具有不同加工路径)； $U(w)$ 为晶圆 w 驻留时间集

合 $U(w) = \{U_{x_1}(w), U_{x_2}(w), \dots, U_{x_i}(w), \dots, U_{x_\varphi}(w)\}$, $U_{x_i}(w)$ 为晶圆 w 在 $M_{x_i}^p$ 加工完成后允许的最大驻留时间； P_{M^b} 为晶圆 w 在 $M_{h,j}^b$ 的加工时间； U_{M^b} 为晶圆 w 在 $M_{h,j}^b$ 的驻留时间； $f_{x_i}(w)$ 为晶圆 w 在 $M_{x_i}^p$ 中的实际停留时间； $d_{x_i}(w)$ 为晶圆 w 在 $M_{x_i}^p$ 加工完成后实际驻留时间； $P_{x_i}^c(t)$ 为 t 时刻在 $M_{x_i}^p$ 进行加工的晶圆数量； $T_{R,h}(w)$ 为对晶圆 w 调度时 R_h 的可用时间区间集, $T_{R,h}(w) = \{t_{R,h,1}(w), t_{R,h,2}(w), \dots, t_{R,h,\varphi t(w,h)}(w)\}$, 包含 $\varphi t(w, h)$ 个区间, q 为 $T_{R,h}(w)$ 的第 q 个可用时间区间。

为有效描述带并行腔的双集束型设备群调度问题, 作如下假设：① 晶圆从 M_a^c 进入系统, 处理完后存放在 M_b^c 中, 且只进入每个模块一次, 不考虑重入, 即晶圆路径坐标集为 $X = \{(1, 1), (1, 2), \dots, (h, j), \dots, (1, n)\}$ (包含所有加工模块)；② 晶圆搬运模块采用单臂机械手, 每次只能搬运一片晶圆, 且晶圆在不同模块间的搬运时间固定；③ 相邻的集束型设备用 2 个缓冲模块连接, 其中一个用于暂存进入集束型设备的晶圆, 另一个用于暂存离开集束型设备的晶圆。 $M_{h,j}^b$ 仅对晶圆起到缓冲作用, 不作任何处理, 驻留时间无限；④ 晶圆在处理模块内驻留时间有上限限制, 即驻留约束, 超过该时间上限, 晶圆变成废品；⑤ 同一处理模块的并行腔在同一个集束型设备中, 不跨集束型设备存在；⑥ 同类型晶圆在同一 $M_{h,j}^p$ 并行腔的处理时间和驻留时间相同；⑦ 同一批 (lot) 中含有不同种类的晶圆, 不同种晶圆在各处理腔的加工时间不同。

根据假设前提, 建立问题模型：

由假设②可知：晶圆搬运模块采用单臂机械手, 则

$$S_{x_{i+1}}(w) - L_{x_i}(w) > t_M \quad (1)$$

连续晶圆进入同一模块的时间差大于 t_M 。

$$S_{x_i}(w + 1) - S_{x_i}(w) > t_M \quad (2)$$

同理, 缓冲模块：

$$B_{h,j}^s(w + 1) - B_{h,j}^s(w) > t_M \quad (3)$$

由假设②可知：搬运必须发生在机械手可用时间集内。

$$\exists q \in \{1, 2, \dots, \varphi t(w, h)\} : [S_{x_i}(w) - t_M, S_{x_i}(w)] \in t_{R,h,q}(w) \quad (4)$$

由假设③可知：

$$P_{M^b} = 0 \quad (5)$$

$$U_{M^b} = +\infty \quad (6)$$

由假设④可知, 驻留约束：

$$f_{x_i}(w) = P_{x_i}(w) + d_{x_i}(w) \quad (7)$$

$$d_{x_i}(w) \leq U_{x_i}(w) \quad (8)$$

由假设⑤可知: $M_{x_i}^p$ 共有 K_i 个并行腔。则

$$C_{x_i}^p(t) \leq K_{x_i} \quad t \in [0, +\infty] \quad (9)$$

晶圆离开 $M_{x_i}^p$ 的必要条件是下一工序存在未被占用的加工腔:

$$C_{x_{i+1}}^p(t) < K_{x_{i+1}} \quad t = L_{x_i}(w) \quad (10)$$

调度目标是晶圆的总加工完成时间(makespan)最小,即最小化最后一块晶圆 W 进入 M_b^c 的时间:

$$\min S_{M_b^c}(W) \quad (11)$$

综上所述,本文对应的调度问题是以式(11)为目标函数,以式(1)~式(10)为约束条件的非线性规划问题。

2 算法构建

为了解决综合考虑多品种加工、晶圆驻留和资源约束等条件下的带并行腔集束型设备群调度问题,结合上述数学模型,构建基于CCR的调度算法,采用逐次计算设备平均活跃时间的方式进行产能约束资源识别,并在产能约束资源前后采用拉动和递推分段调度的优化策略,以达到提高资源利用率的目的。同时为了优化机械手作业顺序,对产能约束资源采取锁定-收紧-松弛(Locking-Tightening-Loosening, LTL)策略。

根据Roser等^[15]的研究内容,对带并行腔的集束型设备群CCR作如下定义。

定义1 集束型设备群中,CCR是平均活跃时间 $\overline{a_{x_i}}$ 最长的处理模块或设备。其计算公式为

$$\overline{a_{x_i}} = \frac{\sum_{w=1}^W [P_{x_i}(w) + 2(t_l + t_u + t_m)]}{K_i W} \quad (12)$$

$$a_l = \max \{ \overline{a_{x_1}}, \overline{a_{x_2}}, \dots, \overline{a_{x_i}}, \dots, \overline{a_{x_\varphi}} \} \quad (13)$$

式中: a_l 为CCR的设备的活跃时间, l 为CCR的设备编号, $l \pm i$ 为CCR前(后)第 i 个处理模块。

定义2 模块饱和状态,在带有 K_i 个并行腔的 $M_{x_i}^p$,满足 $C_{x_i}^p(t) = K_i$ 时,该状态称为饱和状态。

定理1 若含有 $K_i (K_i \geq 2)$ 个并行腔的 $M_{x_i}^p$ 可达到其饱和状态,则须满足 $P_{x_i} \geq (K_i - 2) \cdot t_m$ 。

证明 (反证法)假设 $M_{x_i}^p$ 处于饱和状态,且 $P_{x_i} < (K_i - 2) \cdot t_m$ 。根据递推,易得第 j 片晶圆进入设备时间满足:

$$S_{x_i}(j) = \begin{cases} j \cdot t_m & j \leq K_i \\ [(j-1)/K_i] \cdot [(K_i+1)t_m + P_{x_i}] + \eta \cdot t_m & j > K_i \end{cases} \quad (14)$$

式中:

$$\eta = \begin{cases} \text{mod}(j, K_i) & \text{mod}(j, K_i) \neq 0 \\ K_i & \text{mod}(j, K_i) = 0 \end{cases}$$

因为 $P_{x_i} < (K_i - 2) \cdot t_m$,则易推出 $\forall t \in [0, +\infty]$, $P_{x_i}^c(t) < K_i$,则假设不成立,定理1得证。

引理1 根据定理1和约束理论可得,当 $M_{x_i}^p$ 为CCR,且 $P_{x_i} < (K_i - 2) \cdot t_m$ 时,增加并行腔数量对throughput提升无影响(3.3节结果可说明)。

定义3 对于在 $M_{x_i}^p$ 进行加工的晶圆 w ,定义3种基本操作:①锁定(Locking),令 $L_{x_i}^\xi(w) = S_{x_i}^\xi(w) + P_{x_i}(w)$,即 $d_{x_i}^\xi(w) = 0$;②收紧(Tightening),令 $S_{x_i}^{\xi+1}(w) = S_{x_i}^\xi(w) + 1$ 或 $L_{x_i}^{\xi+1}(w) = L_{x_i}^\xi(w) - 1$,即 $d_{x_i}^{\xi+1}(w) = d_{x_i}^\xi(w) - 1$;③松弛(Loosening),令 $S_{x_i}^{\xi+1}(w) = S_{x_i}^\xi(w) - 1$ 或 $L_{x_i}^{\xi+1}(w) = L_{x_i}^\xi(w) + 1$,即 $d_{x_i}^{\xi+1}(w) = d_{x_i}^\xi(w) + 1$ (“1”为一单位时间, ξ 为变量的更新次数)。

本文算法的核心思想包括4个阶段:①CCR识别;②CCR调度;③CCR前模块调度;④CCR后模块调度。首先,根据定义1确定系统的CCR。然后对CCR实施Locking操作,最大限度使其处于活跃状态。再以CCR为界,在CCR之前模块实施拉动调度,在驻留时间范围内,优化机械手作业时间,若出现不可行时间区间,则返回上一阶段对CCR实施Loosening操作,重新搜索机械手可行作业时间,直至完成CCR之前所有模块调度。然后对CCR之后模块,采用基于Tightening策略下的推式调度,使完成加工的晶圆尽早离开处理模块,减少晶圆驻留时间,逐级推动,直至完成晶圆在所有模块的调度。

具体算法步骤如下。

步骤1 CCR识别。

步骤1.1 录入加工参数数据。

步骤1.2 根据式(12)计算加工腔平均活跃时间。

步骤1.3 根据式(13)识别CCR处理模块 l 。

步骤2 CCR调度。

步骤2.1 初始化各参数,当前调度的晶圆为 w ,令 $\xi = 1$ 。

步骤2.2 计算晶圆 w 在CCR第 k 个并行腔的开始加工时间 $S_{l(k)}(w)$, $k \in [1, K_i]$ 。

若式(15)成立,令 $\xi = \xi + 1$,按式(16)逐次叠加,直至式(17)成立;若式(17)成立,按式(18)计算 $S_{l(k)}(w)$ 。

$$L_{I-1}^{\xi}(w) < \sum_{i=1}^{I-1} P_i(w) + (I-1)(t_1 + t_u + t_m) \quad (15)$$

$$L_{I-1}^{\xi}(w) = L_{I-1}^{\xi-1}(w) + 1 \quad (16)$$

$$L_{I-1}^{\xi}(w) \geq \sum_{i=1}^{I-1} P_i(w) + (I-1)(t_1 + t_u + t_m) \quad (17)$$

$$S_{I(k)}(w) = L_{I-1}^{\xi}(w) + t_1 + t_u + t_m \quad (18)$$

步骤 2.3 判断机械手是否空闲。

判断式(19)是否成立,若不成立,令 $\xi = \xi + 1$,按式(16)更新,直至成立;否则,读取当前各参数:

$$[L_{I-1}^{\xi}(w), L_{I-1}^{\xi}(w) + t_m] \in T_{R,h} \quad (19)$$

步骤 2.4 对 CCR 实施 Locking 操作,根据式(20)计算 $L_{I(k)}(w)$:

$$L_{I(k)}(w) = S_{I(k)}(w) + P_I(w) \quad (20)$$

步骤 2.5 判断式(21)是否成立,若不成立,令 $\xi = \xi + 1$,按式(22)更新,直至式(21)成立。

$$[L_{I(k)}^{\xi}(w), L_{I(k)}^{\xi}(w) + t_m] \in T_{R,h} \quad (21)$$

$$L_{I(k)}^{\xi}(w) = L_{I(k)}^{\xi-1}(w) + 1 \quad (22)$$

步骤 2.6 根据 Locking 策略计算晶圆 $w + 1$ 的 $S_{I(w+1)}, L_{I(w+1)}$ 。

若 $C_I^p(t) < K_I$ 且 $w < K_I$,则

$$S_{I(k+1)}(w+1) = S_{I(k)}(w) + \max\{P_{I+1}, P_{I+2}, \dots\} + 2(t_1 + t_u + t_m) \quad (23)$$

$$L_{I(k+1)}(w+1) = S_{I(k+1)}(w+1) + P_I(w+1) \quad (24)$$

否则

$$S_{I(k)}(w+1) = L_{I(k)}(w - K_I + 1) + 2(t_1 + t_u + t_m) \quad (25)$$

$$L_{I(k)}(w+1) = S_{I(k)}(w+1) + P_I(w+1) \quad (26)$$

步骤 2.7 根据上述完成 CCR 调度。

步骤 3 CCR 前模块调度。

步骤 3.1 初始化参数 $d_c(w)$ (实际驻留时间)、 c (当前被调度的处理模块编号)。

步骤 3.2 判断是否满足驻留约束,如果满足,继续下一步;否则,对 CCR 执行一次 Loosening 操作,返回步骤 2。

步骤 3.3 计算。

$$S_{c(k)}(w) = L_{c(k)}(w) - P_c(w) - d_c(w) \quad (27)$$

$$L_{c-1(g)}(w) = S_{c(k)}(w) - (t_1 + t_u + t_m) \quad (28)$$

步骤 3.4 判断机械手是否空闲,若机械手空闲,则继续下一步;若机械手忙碌,则执行一次 Loosening 操作,返回步骤 2。

步骤 3.5 判断当前处理模块是否空闲

$$L_{c(g)}(w) - L_{c-1(g)}(w) \geq t_1 + t_u + t_m \quad (29)$$

步骤 3.6 更新参数,判断是否完成所有 CCR 前处理模块的调度,如果已完成即进入下一阶段,否则返回步骤 3.2。

步骤 4 CCR 后模块调度。

步骤 4.1 初始化参数 $d_c(w)$ 、 c 。

步骤 4.2 判断是否满足驻留约束,若满足,则执行一次 Tightening 操作,继续下一步;否则:若当前处理模块为 CCR,则返回前一处理模块,按式(16)更新,返回步骤 2.2。若当前模块为非 CCR,则增加一单位驻留时间,返回步骤 4.2。

步骤 4.3 计算。

$$S_{c(k)}(w) = L_{c-1(g)}(w) + (t_1 + t_u + t_m) \quad (30)$$

$$L_{c(g)}(w) = S_{c(k)}(w) + P_c(w) + d_c(w) \quad (31)$$

步骤 4.4 判断处理模块是否繁忙,如式(32),若繁忙,增加一单位驻留时间,并返回步骤 4.2;否则,记录各参数。

$$L_{c(g)}(w) - L_{c-1(g)}(w) \geq t_1 + t_u + t_m \quad (32)$$

步骤 4.5 更新参数,判断是否完成 CCR 后的所有模块调度,若未完成,返回步骤 4.2,否则计算结束,完成晶圆 w 的调度。

步骤 5 重复上述步骤 1 ~ 步骤 4,直至完成所有晶圆调度,算法结束。

以图 1 所示集束型设备群为例,运用基于 CCR 的 TLT 算法对其中 6 个不同种类的晶圆 ($w_1 \sim w_6$) 进行调度,调度结果的甘特图如图 2 所示。

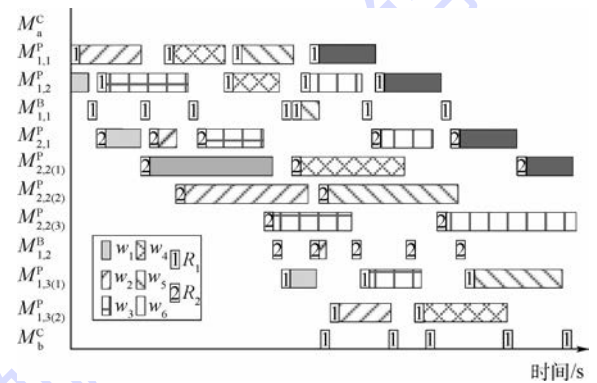


图 2 计算实例甘特图

Fig. 2 Gantt chart of calculative example

3 仿真实验与分析

为有效评价本文算法,选取无驻留约束条件下带并行腔单集束型设备调度最优策略——LCM 策略进行比较,并以 Kumar 等^[14]提出的算法作为 Benchmark。

现将相关符号与变量定义如下: $R = \frac{C_{LCM} - C_{TLT}}{C_{LCM}} \times 100\%$, 该值越大,表明调度得到的制造期相对 C_{LCM} 的提升越大,即算法的性能越好, C_{TLT} 为本文提出的 LTL 策略下算法的制造周

期, C_{LCM} 为基于 LCM 策略算法的制造周期; $F^C = \frac{(M^H + 1)t_M}{\max P + t_M}$ 为设备因子 (cluster factor), 表征集束型设备群中机械手繁忙程度的函数, 其值越大, 机械手完成一次搬运所需时间与最大处理时间的比例越大, 即机械手越繁忙, 反之越空闲, $\max P$ 为晶圆在集束型设备中各个处理模块的最大处理时间, M^H 为每个集束型设备中的处理模块数目。

实验用 C++ 编程实现, 仿真的环境为主频为 2.13 GHz、内存为 8 GB 的笔记本电脑。实验中的数据, 均为 10 次相同实验的平均结果。

3.1 运算时间

假设晶圆在各处理模块的处理时间服从 $P \sim N(40, 10)$; 在各处理模块的驻留约束时间服从 $U \sim N(20, 5)$; 搬运时间 $t_M = 2$ s; 每个集束型设备所含的处理模块的数目 $M^H = 4$; 每个处理模块并行腔数集合 $P_n = \{1, 2, 1, 3, 1, 1, 1, 1\}$ 。仿真运行结果如图 3 所示。

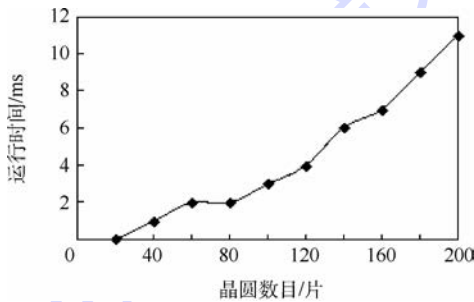


图 3 算法运行时间分析

Fig. 3 Running time analysis of algorithm

图 3 显示, 算法在晶圆数目较少时 (小于 20 片), 运行时间几乎为 0, 算法的运行时间随着晶圆数目的增加大体上呈正相关, 当晶圆数目为 200 片, 算法运行时间仍未超过 0.2 s, 由此可见, 算法响应速度极快, 适合于晶圆实时调度。

3.2 加工时间对调度的影响

令 $M^H = 4; t_M = 2; P_n = \{1, 2, 1, 3, 1, 1, 1, 1\}$; 晶圆在各个处理模块的处理时间服从期望 $\mu = 40$, 方差 $\sigma = 0, 1/8, 1/4, 1/2$ 均值的正态分布; 驻留约束时间分别服从 $U \sim N(20, 5)$ 。仿真运行的结果如图 4 所示。由图 4 可见, 随着加工晶圆的数量增加, R 值逐渐增加, 说明在加工批量越大的情况下本文算法的调度结果相对于 Benchmark 的提升越大。同时, 图 4 中的 4 条不同曲线的走向趋势基本一致, 且在 σ 的较大的情况下, 本文算法的优化结果更为明显, 说明了算法对不同处理时间波动的适应性。

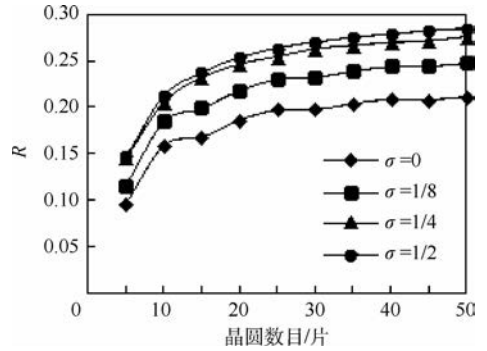


图 4 不同方差下 R 值与晶圆数目的关系

Fig. 4 Relationship between R value and number of wafer with different variance

3.3 CCR 并行腔对调度的影响

令 $M^H = 4, t_M = 2$ s; 晶圆在各个处理模块的处理时间服从 $P \sim N(40, 10)$; 驻留约束时间分别服从 $U \sim N(20, 5)$; CCR 模块的并行腔数目 $K_i = 1, 2, \dots, 7$ 。仿真运行的结果如图 5 所示。

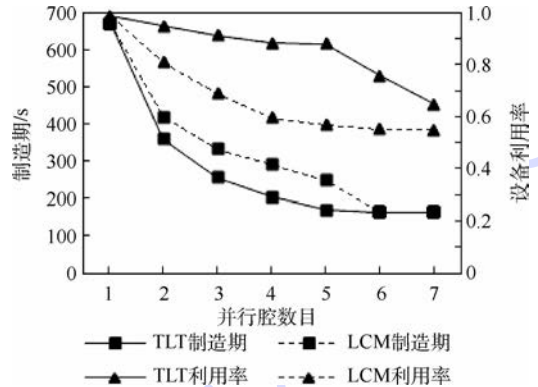


图 5 CCR 并行腔对调度的影响

Fig. 5 Effect of parallel chambers at CCR on scheduling

图 5 显示了在 CCR 并行腔数目不同的情况下, 本文算法的调度 makespan 均小于基于 LCM 的调度结果, 并且 CCR 设备的利用率也相对较高。同时, 随着 CCR 并行腔数目的增多, 可以看出 makespan 不断减小, 设备利用也相对降低, 这是 TOC 理论下对 CCR 不断改善的结果, 当 CCR 并行腔数超过 6 时, 增加并行腔数目, makespan 不再下降, 说明此时 CCR 发生转移。

3.4 机械手对调度的影响

令 $M^H = 4, t_M = 2$ s, $P_n = \{1, 2, 1, 3, 1, 1, 1, 1\}$; 晶圆在各个处理模块的处理时间服从 $P \sim N(40, 10)$; 驻留约束时间分别服从 $U \sim N(20, 5)$; 分别令 $F^C = 1, 2, \dots, 6$ 。仿真运行的结果如图 6 所示。

图 6 显示了不同机械手繁忙程度下 2 种算法累计驻留时间和 makespan 的关系。随着 F^C 的增加, makespan 趋向于线性增加, 且本文算法调度结果均优于 Benchmark。在一般情况下, 累计驻

留时间越短,则调度可得到的 makespan 越小,图中的时间比例也说明这种关系。同时,算法在获取较短驻留时间方面也存在优势。

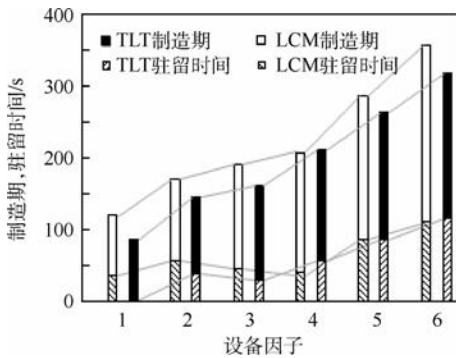


图6 机械手对调度的影响

Fig.6 Effect of manipulator on scheduling

4 结论

1) 算法可以有效解决在集束型设备群的调度过程中普遍存在的带并行腔多品种加工的问题,特别在减少驻留提升 CCR 设备利用率方面有明显优势。

2) 在晶圆批量较小(每批少于 20 片)的情况下,得出算法的仿真运行时间几乎为零,可有效地进行实时调度,即使批量较大(每批达到 200 片)也能满足调度的要求,证明了算法的高效性。

3) 随着晶圆数目和 σ 增大, R 值不断增加,表明本文算法相对于 LCM 算法的提升效果明显,且在并行腔数目和设备因子变化的情况下,也能获得较优的调度结果。

4) 算法对于晶圆加工时间波动性、并行腔数目和机械手繁忙程度的仿真证明算法对不同的加工数据和设备具有非常好的适应性。

针对此类问题的优化,今后可不仅考虑 makespan 最小这一个目标,还可以引入拖期时间最少、设备利用率最高等其他因素,采用多目标优化技术。此外,通过甘特图看到调度结果还有进一步优化的空间,因此今后可以考虑以启发式算法作为初始解,采用迭代搜索算法进行进一步的优化。

参考文献 (References)

[1] PERKINSON T L, MCLARTY P K, GYURCSIK R S, et al. Single-wafer cluster tool performance: An analysis of throughput [J]. IEEE Transactions on Semiconductor Manufacturing, 1994, 7(3):369-373.

[2] YI J, DING S, SONG D, et al. Steady-state throughput and scheduling analysis of multi-cluster tools: A decomposition ap-

proach [J]. IEEE Transactions on Automation Science and Engineering, 2008, 5(2):321-336.

- [3] CHAN W K, YI J, DING S. Optimal scheduling of multi-cluster tools with constant robot moving times, Part I: Two-cluster analysis [J]. IEEE Transactions on Automation Science and Engineering, 2011, 8(1):5-16.
- [4] CHAN W K, DING S, YI J, et al. Optimal scheduling of multi-cluster tools with constant robot moving times, Part II: Tree-like topology configurations [J]. IEEE Transactions on Automation Science and Engineering, 2011, 8(1):17-28.
- [5] DING S, YI J, ZHANG M T, et al. Performance evaluation and schedule optimization of multi-cluster tools with process times uncertainty [C] // Proceeding of the 2006 IEEE International Conference on Automation Science and Engineering, Piscataway, NJ: IEEE Press, 2006:7-10.
- [6] LIU M X, ZHOU B H. Modeling and scheduling analysis of multi-cluster tools with residency constraints based on time constraint sets [J]. International Journal of Production Research, 2013, 51(16):4835-4852.
- [7] PERKINSON T L, GYURCSIK R S, MCLARTY P K. Single-wafer cluster tool performance: An analysis of the effects of redundant chambers and revisitation sequences on throughput [J]. IEEE Transactions on Semiconductor Manufacturing, 1996, 9(3):384-400.
- [8] GEISMAR H N, DAWANDE M W, SRISKANDARAJAH C. Robotic cells with parallel machines: Throughput maximization in constant travel-time cells [J]. Journal of Scheduling, 2004, 7(5):375-395.
- [9] ZHENG X H, YU H B, HU J T. A general throughput model for parallel cluster tools [C] // International Conference on ICCE2011, AISC 110. Berlin: Springer Press, 2011:215-222.
- [10] 卢睿, 李林琰. 有晶滞留时间约束的集束型装备调度问题研究 [J]. 控制仿真学报, 2014, 26(8):1775-1780.
- LU R, LI L Y. Research on scheduling problem of cluster tools with residency time constraints [J]. Journal of System Simulation, 2014, 26(8):1775-1780 (in Chinese).
- [11] WIKBORG U, LEE T E. Noncyclic scheduling for timed discrete-event systems with application to single-armed cluster tools using pareto-optimal optimization [J]. IEEE Transactions on Automation Science and Engineering, 2013, 10(3):699-710.
- [12] ZHANG J, FANG X, QI L. LCM cycle based optimal scheduling in robotic cell with parallel workstations [C] // 2014 IEEE International Conference on Robotics and Automation (ICRA). Piscataway, NJ: IEEE Press, 2014:1367-1373.
- [13] DING S, YI J. An event graph based simulation and scheduling analysis of multi-cluster tools [C] // Proceedings of 2004 Winter Simulation Conference. Piscataway, NJ: IEEE Press, 2004:1915-1924.
- [14] KUMAR S, RAMANAN N, SRISKANDARAJAH C. Minimizing cycle time in large robotic cells [J]. IIE Transactions, 2005, 37(2):123-136.
- [15] ROSER C, NAKANO M, TANAKA M. A practical bottleneck detection method [C] // Proceedings of Winter Simulation Conference, 2001. Piscataway, NJ: IEEE Press, 2001, 2:949-953.

作者简介:

周炳海 男,博士,教授,博士生导师。主要研究方向:制造系统调度、建模与仿真评价,设备可靠性与预防性维护建模。

黎明 男,硕士研究生。主要研究方向:半导体制造系统生产调度、建模与仿真评价。

Tel: 021-54810409

E-mail: bhzhou@tongji.edu.cn

Scheduling method for double-cluster tools with parallel chambers based on capacity constraint resource

ZHOU Binghai*, LI Ming

(School of Mechanical Engineering, Tongji University, Shanghai 201804, China)

Abstract: To effectively solve scheduling problems of multiple cluster tools of 450 mm wafer fabrication systems with parallel processing chambers, a scheduling method based on capacity constraint resource (CCR) was proposed. Firstly, with comprehensive consideration of the characteristics of different types of wafers, resources and residency constraints, a mathematical programming model of double-cluster tools with parallel chambers was established to minimize the makespan of the system. Then, to optimize manipulator movements by adopting locking-tightening-loosening (LTL) strategy to the CCR, a piecewise scheduling algorithm was constructed with the CCR. Finally, through the analysis of simulation experiments, the results indicate that the proposed algorithm is valid and competitive.

Key words: multiple cluster tools; parallel chambers; residency constraint; capacity constraint resource; scheduling algorithm

Received: 2015-08-24; **Accepted:** 2015-10-30; **Published online:** 2016-01-05 10:36

URL: www.cnki.net/kcms/detail/11.2625.V.20160105.1036.002.html

Foundation items: National Natural Science Foundation of China (71471135,61273035)

* **Corresponding author.** Tel.: 021-54810409 E-mail: bhzhou@tongji.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0448

电动伺服舵系统动力学建模及颤振分析



张仁嘉, 吴志刚*, 杨超

(北京航空航天大学 航空科学与工程学院, 北京 100083)

摘 要: 舵机动力学对舵系统颤振有最直接影响,其动力学模型的准确建立是颤振分析的关键。以某电动伺服舵系统为对象,分别建立了电机、减速器及控制回路模型。基于其地面振动试验(GVT)数据特征,考虑了减速器中零部件间隙及滚珠丝杠间接接触刚度2类非线性因素。动力学仿真结果复现了试验现象,证明了接触刚度和间隙对该伺服舵系统动力学特征的主导作用。在颤振分析中,对比了接触刚度和常数刚度2种假设下,舵机在不同阶跃角指令下的舵偏角响应。发现常数刚度下舵系统仅存在极限环(LCO)和失稳2个域,临界速度不会随间隙大小而变化;而接触刚度下存在稳定、LCO和失稳3个域,且相同风速下,其临界指令角较常数刚度假设更高,但其临界速度却会随间隙增大而减小,甚至小于常数刚度假设,应引起足够重视。

关键词: 气动弹性; 颤振; 非线性; 接触刚度; 间隙; 舵机

中图分类号: V211.4

文献标识码: A **文章编号:** 1001-5965(2016)07-1368-09

舵系统是舵机和舵面耦合系统的统称,承担着实时改变飞行器姿态的任务。作为重要构成部件,舵系统的稳定性一直都是设计部门关注的焦点。其中,舵机本身的动态稳定性是考核重点之一,舵系统伺服颤振稳定性是另一重点。过去,工程中主要关注舵面结构本身的颤振特性,很少考虑舵机伺服的影响。舵机伺服常以弹簧代替,其刚度为一常数。而研究表明舵机伺服刚度其实是一个随激励频率变化而变化的复数^[1-4]。基于这一结论会得到和常规方法不同的颤振边界。

国外研究人员在20世纪90年代已将舵机伺服纳入颤振分析框架^[1],主要由韩国学者Lee等^[2]主导,典型工作是考虑舵机动刚度和间隙非线性的颤振分析。国内的相关研究也一直在跟进,如张新替等^[3]发展了一种可考虑舵机动刚度的颤振分析方法,并提出一种通过改变舵机控制

律来调节舵机动刚度,从而实现颤振抑制的方法^[4];杨宁等^[5]分析了结构非线性和舵机动刚度对舵系统颤振的综合影响。典型结论是动刚度改变了系统阻尼,从而影响到颤振边界;间隙会削弱舵面的支持刚度,还会带来极限环振荡或混沌效应,对系统颤振有不利影响。

在文献[1-5]中提到的动刚度及非线性因素之外,还有一种因素一直没有引起关注,那就是接触刚度。接触刚度广泛存在于发生力接触的两物体之间,它是作用力的函数,表现出强非线性特征。无论是经典的Hertz接触理论,还是考虑了接触表面粗糙度的GW模型^[6],或是更精细的分型模型^[7],都证明了这一结论。在舵机减速器内部,接触无处不在,如键与槽的接触,齿轮轮齿间的接触和滚珠和丝杠的接触等。其中,尤以滚珠丝杠副这类点接触最“软”,加上它位置靠近输出

收稿日期: 2015-07-03; 录用日期: 2015-08-20; 网络出版时间: 2015-09-17 10:41

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150917.1041.004.html

基金项目: 国家自然科学基金(11372023)

* 通讯作者: Tel.: 010-82317510 E-mail: wuzhigang@buaa.edu.cn

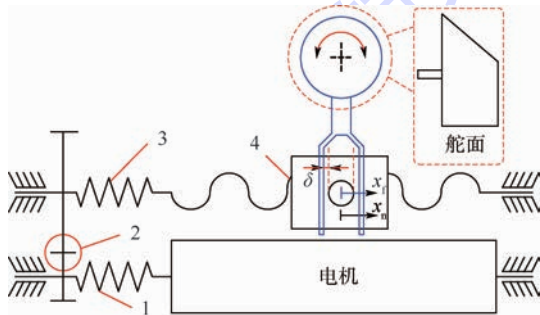
引用格式: 张仁嘉, 吴志刚, 杨超. 电动伺服舵系统动力学建模及颤振分析[J]. 北京航空航天大学学报, 2016, 42(7): 1368-1376.
ZHANG R J, WU Z G, YANG C. Dynamic modeling and flutter analysis of a fin-actuator system [J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1368-1376 (in Chinese).

级这一因素,足以对舵面旋转频率造成极大影响,直接影响舵系统整体颤振特性。

现有文献有关接触刚度的报道多集中于机械加工过程^[8-12],其中有关滚珠丝杠的研究也多是数控机床中的长程丝杠^[8-10],其结构特征、承载形式和工作环境,都与舵机减速器中的丝杠有很大差别。本文将以某电动伺服舵系统为对象,探索接触刚度对舵机动力学的影响。从地面振动试验现象出发,分析舵机内部可能起主导作用的非线性因素,并以此建立舵机动力学模型;将仿真和试验结果进行对比,验证建模方法的可行性;最后建立舵系统颤振时域分析模型,对接触刚度和间隙耦合作用下舵系统的响应行为展开研究。

1 研究对象

研究对象为某电动伺服舵系统,其结构如图1所示。舵系统由电动舵机和全动梯形舵组成。舵机内含两级减速器,第1级为齿轮副,第2级为滚珠丝杠副。舵轴下方设有可调阻尼器。



δ —拨叉和螺母连接处的一个“综合间隙”半宽度,为减速器中各零部件配合处间隙统一向拨叉螺母处归集的一个等效间隙; x_r —螺母位移; x_n —拨叉位移;1,3—扭转刚度;2—啮合刚度;4—接触+拉压刚度。

图1 舵系统结构

Fig.1 Structure of fin-actuator system

对于这类舵系统,舵机对舵面的支持作用十分关键。当转角指令被给定时,舵机处于位置保持状态。此时,舵面若受外力作用迫使舵机转角偏移,电机会产生回复力矩进行纠偏。这种作用类似于扭转弹簧,不同点在于扭转弹簧的刚度是一个定值,而舵机因为有能量源存在,其刚度会随激励频率变化而变化,是带有相位的复值。

2 试验

从地面振动试验中获取了舵面模式、频率以及舵机频响(FRF)特性。考虑到舵机的分散性,选择了多组舵机进行试验。试验方案如图2所示。

进行带有力反馈的步进正弦扫频,以激振力

(F_t)为输入,激振点处加速度为输出,得到舵系统频响曲线。将激振力在40~120 N之间的一组典型频响曲线绘于图3。发现随着激振力的提高,舵面旋转频率从70 Hz向84 Hz迁移,表明舵机刚度随激励力的变化发生变化。

变换多组舵机进行相同测试,将其中4台(1#~4#)的舵面旋转频率随激振力变化趋势绘于图4。4条曲线走势相同,说明舵机一致性较好。舵面旋转频率随着激振力的增大而提高,且增速渐缓,显示出典型的多项式刚度特性。

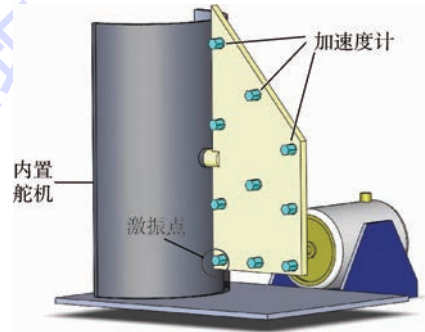


图2 地面振动试验方案

Fig.2 Scheme of ground vibration test

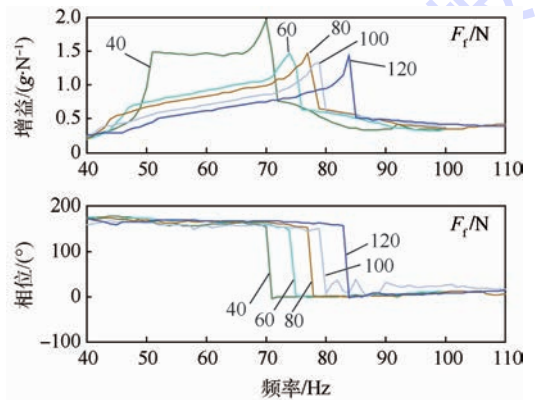


图3 不同激振力下舵面频响曲线

Fig.3 FRF curves of fin under different excitation forces

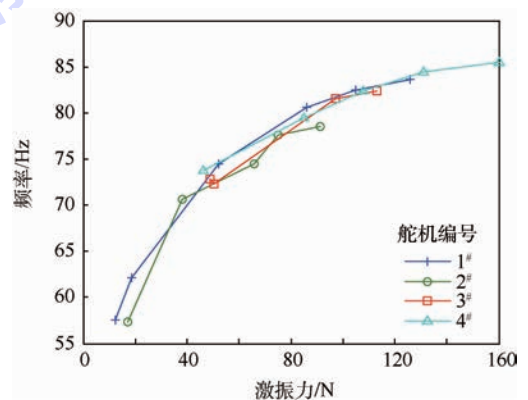


图4 舵面旋转频率随激振力的变化趋势

Fig.4 Change trend of rotational frequency of fin according to excitation force

根据图 1 所示舵系统结构,分析造成这一现象的可能原因。电机输出轴以键连接方式与小齿轮相连,存在轴扭转刚度;小齿轮与大齿轮间存在啮合刚度;丝杠存在轴扭转刚度和拉压刚度;滚珠丝杠副存在接触刚度;其余部位刚度或量级较高,或位置远离舵轴,对舵面旋转频率不构成主要影响,可不予考虑;各零部件配合处有间隙存在,尤以拨叉和滑动螺母相接处最大。所有这些因素中,造成非线性刚度的关键因素可能是滚珠丝杠接触刚度及间隙。

3 理论建模

3.1 舵机动力学及运动学建模

将减速器结构等效为多自由度质量-阻尼-弹簧系统,逐级建立其动力学方程为

$$\begin{cases} J_0 \ddot{\theta}_0 + b_0 \dot{\theta}_0 + T_1 = T_0 \\ J_1 \ddot{\theta}_1 + b_1 \dot{\theta}_1 + T_2 = T_1 \\ J_2 \ddot{\theta}_2 + b_2 \dot{\theta}_2 + T_3 = T_2 \\ J_3 \ddot{\theta}_3 + b_3 \dot{\theta}_3 + T_4 = T_3 \\ J_4 \ddot{\theta}_4 + b_4 \dot{\theta}_4 + T_{load} = T_4 \end{cases} \quad (1)$$

式中: $J_0 \sim J_4$ 分别为电机转子惯量、小齿轮惯量、大齿轮惯量、丝杠旋转惯量和轴套旋转惯量; $\theta_0 \sim \theta_4$ 和 $b_0 \sim b_4$ 分别为电机转子、小齿轮、大齿轮、丝杠和轴套的旋转角与旋转方向的摩擦阻尼系数; T_0 为作用于电机转子上的洛伦兹力; T_{load} 为舵机受到的外部作用力矩。其余参数定义为

$$\begin{cases} T_1 = k_0 (\theta_0 - \theta_1) \\ T_2 = k_1 r_2 (\theta_1 r_1 - \theta_2 r_2) \\ T_3 = k_2 (\theta_2 - \theta_3) \\ T_4 = \frac{T_5 r_3 \tan \lambda}{d} \\ T_5 = \begin{cases} k_3 x_e \operatorname{sgn}(\Delta x) \cdot d & |\Delta x| > \delta \\ 0 & |\Delta x| \leq \delta \end{cases} \end{cases} \quad (2)$$

式中: $r_1 \sim r_3$ 分别为小齿轮、大齿轮和丝杠节圆半径; $k_0 \sim k_2$ 分别为电机转子轴扭转刚度、大小齿轮啮合刚度和丝杠扭转刚度; $k_3 = \frac{1}{\left(\frac{1}{k_{tc}} + \frac{1}{k_c}\right)}$ 为一个

综合刚度, k_{tc} 为丝杠拉压刚度, k_c 为滚珠丝杠接触刚度; d 为舵轴与丝杠轴之间垂直距离; $\Delta x = x_n - x_f, x_n = \frac{\theta_3 s}{2\pi}, x_f = d \tan \theta_4$; λ 和 s 分别为丝杠螺旋角和导程; $x_e = |\Delta x| - \delta$ 为丝杠有效变形。

需要说明的是,啮合刚度 k_1 虽是随旋转运

动周期变化的刚度,但舵机处于位置保持状态时的啮合刚度变化幅度很小,所以 k_1 可按常数处理。

关于接触刚度 k_c ,其来源及受力关系如图 5 所示。Hertz 接触理论认为,滚珠和曲面之间的接触是点接触,有负荷 Q 作用时,接触点将扩展成为一个椭圆接触面(见图 5(a))。对于丝杠-滚珠-螺母副,滚珠沿螺旋滚道运动,受到来自螺母和丝杠滚道的双向挤压(见图 5(b)),滚珠和滚道在接触点处同时发生弹性变形。

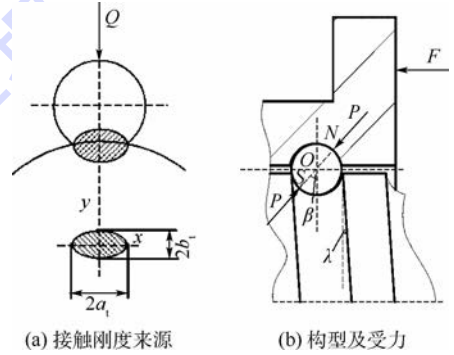


图 5 丝杠-滚珠-螺母副受力与接触模型
 a_1 —椭圆长轴半径; b_1 —椭圆短轴半径; N —滚珠与螺母接触点; S —滚珠与丝杠接触点; P —法向接触力; F —轴向载荷。

图 5 Force and contact model of screw-ball-nut pair

假设螺母受到的轴向载荷为 F ,丝杠-滚珠-螺母副因接触刚度导致的轴向变形量为 δ_a ,由 Hertz 接触理论知理想接触情况下轴向变形 δ_a 与轴向工作载荷 F 间的关系为

$$\delta_a = \left(\frac{\cos \lambda}{z^2 \sin^5 \beta} \right)^{\frac{1}{3}} \left[\frac{2K(e_{np})}{\pi m_{asp}} \sqrt{\frac{1}{8} \left(\frac{3}{E'} \right)^2 \sum \rho_{np}} + \frac{2K(e_{sp})}{\pi m_{asp}} \sqrt{\frac{1}{8} \left(\frac{3}{E'} \right)^2 \sum \rho_{sp}} \right] F^{\frac{2}{3}} \quad (3)$$

式中: z 为工作滚珠数; β 为滚珠丝杠副接触角; $K(e)$ 为与椭圆偏心率有关的第一类完全椭圆积分; e_{np} 为滚珠与螺母滚道椭圆接触面的偏心率; m_{asp} 为该接触面的长半轴系数; ρ_{np} 为该接触面的主曲率; E' 为当量弹性模量; e_{sp} 为滚珠与丝杠滚道椭圆接触面的偏心率; m_{asp} 为该接触面的长半轴系数; ρ_{sp} 为该接触面的主曲率。

将式(3)简写为

$$\delta_a = C_d F^{\frac{2}{3}} \quad (4)$$

当接触副的结构形式和制造材料确定后,轴向接触变形系数 C_d 就是一个确定值。可见接触副的轴向接触变形与轴向载荷的 $\frac{2}{3}$ 次方成正比。

由式(4)进一步得到接触副轴向接触刚度 k_c

的表达式为

$$k_c = \frac{dF}{d\delta_a} = \frac{F^{\frac{1}{3}}}{K_d} \quad (5)$$

式中: $K_d = \frac{2C_d}{3}$ 。

可知丝杠-滚珠-螺母副的轴向接触刚度随着轴向载荷的增大而增大,呈非线性关系。实际当中,因为几乎不存在赫兹接触理论中的理想接触,所以 k_c 和 F 之间的关系通常表示为

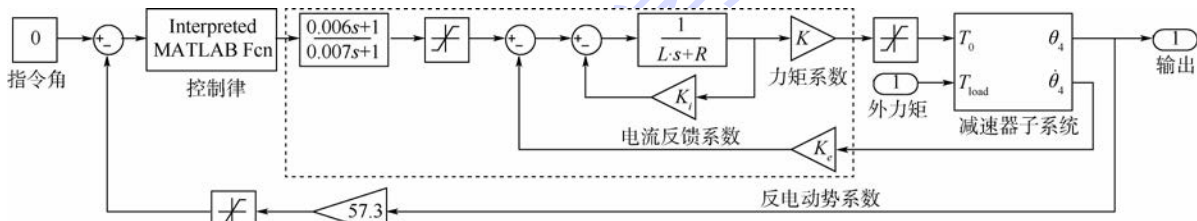
$$k_c = \frac{F^\gamma}{K_d} \quad (6)$$

式中:次方数 γ 不会严格等于 $\frac{1}{3}$,而是有一定

范围^[10]。

电机电路部分建模可参见文献[13],图6直接给出基于前述理论建立的舵机综合模型 Simulink 框图。其中虚线框部分为电机电路模型,减速器子系统根据式(1)的动力学方程建立。

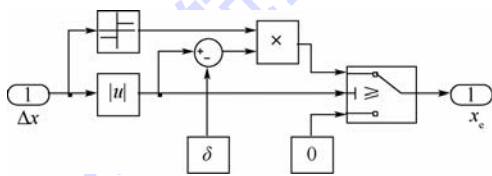
对于减速器内部结构,仅给出丝杠-滚珠-螺母副这一关键部位的详细模型,其余部分比较常规,不再赘述。图7为螺母和拨叉结合处的间隙模型。图8为由丝杠拉压刚度和接触刚度合成的综合刚度模型。图9为丝杠-滚珠-螺母副力学模型,其中间隙模型如图7所示,刚度模型如图8所示。



s—Laplace 变量;L—电感;R—电阻; K_i —电流反馈系数; K_e —反电动势系数; K —力矩系数。

图6 舵机 Simulink 框图

Fig. 6 Simulink block diagram of synthesized fin-actuator system



u—输入值。

图7 间隙模型

Fig. 7 Freeplay model

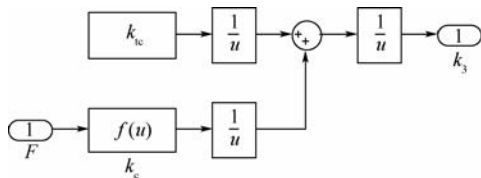
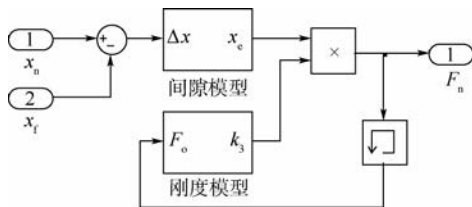


图8 丝杠-滚珠-螺母副刚度模型

Fig. 8 Stiffness model of screw-ball-nut pair



F_n — F 的新值; F_o — F 的旧值。

图9 丝杠-滚珠-螺母副力学模型

Fig. 9 Force model of screw-ball-nut pair

3.2 颤振时域建模

为了舵面颤振方程中引入舵机的“复杂”支持,基于分支模态法^[14]对方程进行改写。取舵轴根部固支情况下的舵面弹性模态作为弹性分支,舵轴根部放开旋转自由度得到的舵面刚体旋转模态作为刚体分支,得

$$M\ddot{x} + C\dot{x} + Kx = F_a \quad (7)$$

式中:

$$x = \begin{bmatrix} q_e \\ \theta_r \end{bmatrix}$$

$$M = \begin{bmatrix} M_{ee}^{n \times n} & M_{er}^{n \times 1} \\ M_{re}^{1 \times n} & M_{rr} \end{bmatrix}$$

$$C = \begin{bmatrix} C_{ee}^{n \times n} \\ 0 \end{bmatrix}$$

$$K = \begin{bmatrix} K_{ee}^{n \times n} \\ k_{rr} \end{bmatrix}$$

其中: $M_{ee}^{n \times n}$ 、 $C_{ee}^{n \times n}$ 和 $K_{ee}^{n \times n}$ 分别为舵面弹性分支 n 阶模态对应的广义质量、阻尼和刚度阵,均为对角阵; q_e 为广义坐标; M_{rr} 、 k_{rr} 和 θ_r 分别为刚体分支旋转模态对应的物理惯量、刚度以及坐标; $M_{er}^{n \times 1}$ 为 $M_{re}^{1 \times n}$ 的转置,代表弹性和刚体分支的耦合质量; F_a 为气动力。

在计算 F_a 时,输入模态为舵面弹性模态和刚体模态分支的集合,即 1 阶刚体旋转模态和 n 阶弹性模态的组合。

频域下舵面广义非定常气动力表达式为

$$F_{ta} = q_{\infty} A_r \mathbf{x} \quad (8)$$

式中: $q_{\infty} = \rho V^2/2$ 为飞行动压, ρ 为大气密度, V 为来流速度; A_r 为气动力影响系数矩阵, 其为减缩频率 $k = \omega b/V$ 的函数, ω 为圆频率, b 为参考弦长。

时域下的气动力影响系数矩阵可借助最小状态 (Minimum-State, MS) [15] 法进行拟合。先给出拉氏域下的气动力影响系数矩阵:

$$A_L = A_0 + pA_1 + p^2A_2 + p\tilde{D}(pI - \tilde{R})^{-1}\tilde{E} \quad (9)$$

式中: $p = bs/V$; $A_0 \sim A_2$ 、 \tilde{D} 和 \tilde{E} 均为系数矩阵; $\tilde{R} = \text{diag}(r_1, r_2, \dots, r_m)$ 为气动力滞后根对角阵。

略去滞后根的影响, 则时域下非定常气动力可近似为

$$F_{ta} = q_{\infty} \left(A_0 \mathbf{x} + \frac{b}{V} A_1 \dot{\mathbf{x}} + \frac{b^2}{V^2} A_2 \ddot{\mathbf{x}} \right) \quad (10)$$

式中: $F_{ta} = [f_{q_e} \ f_{\theta_r}]^T$ 为频域广义气动力向量, f_{θ_r} 为舵面刚体旋转方向的气动力分量, f_{q_e} 为舵面弹性坐标下的气动力分量。

将式(10)代入式(7), 并将式(7)写成状态空间的形式:

$$\begin{cases} \dot{\bar{\mathbf{x}}} = \bar{A}\bar{\mathbf{x}} + \bar{B}\bar{u} \\ \bar{\mathbf{y}} = \bar{C}\bar{\mathbf{x}} + \bar{D}\bar{u} \end{cases} \quad (11)$$

式中: $\bar{\mathbf{x}} = [\mathbf{x} \ \dot{\mathbf{x}}]^T$, 输入 \bar{u} 为舵机转角, 输出 $\bar{\mathbf{y}}$ 为舵面施加给舵机的外力矩。其他变量定义如下:

$$\bar{A} = \begin{bmatrix} \mathbf{O}^{(n+1) \times (n+1)} & \mathbf{I}^{(n+1) \times (n+1)} \\ -\bar{M}^{-1}\bar{K} & -\bar{M}^{-1}\bar{C} \end{bmatrix}$$

$$\bar{B} = \begin{bmatrix} \mathbf{O}^{(n+1) \times 1} \\ \bar{M}^{-1}\bar{K}' \end{bmatrix}$$

$$\bar{C} = [\mathbf{O}^{1 \times n} \ k_{rr} \ \mathbf{O}^{1 \times (n+1)}]$$

$$\bar{D} = [-k_{rr}]$$

$$\bar{M} = M - \frac{1}{2}\rho b^2 A_2$$

$$\bar{C} = C - \frac{1}{2}\rho V b A_1$$

$$\bar{K} = K - \frac{1}{2}\rho V^2 b A_0$$

$$\bar{K} = \begin{bmatrix} \mathbf{O}^{n \times 1} \\ k_{rr} \end{bmatrix}$$

最终建立如图 10 所示的舵系统颤振时域综合模型。舵系统对应式(11)的舵面 + 气动力子系统, 舵机对应图 6 所示的舵机子系统。

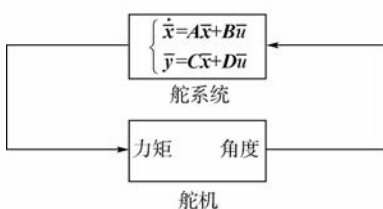


图 10 舵系统颤振模型 Simulink 框图

Fig. 10 Simulink block diagram of fin-actuator flutter model

4 结果与讨论

舵系统虽涉及诸多参数, 但其中大部分参数是确定的或可测的, 如控制系统和电机电路参数、舵机零部件质量 (或惯量)、舵轴阻尼器阻尼系数等。表 1 列出的是仿真计算过程中用到的主要参数, 这些参数都是确定的。

表 1 仿真参数

Table 1 Simulation parameters

类目	参数	数值
	电阻 R/Ω	1.1
	电感 L/H	6.53×10^{-4}
电机	力矩系数 $k_t/(N \cdot m \cdot A^{-1})$	0.034
	反电动势系数 $k_{emf}/(V \cdot \text{rad}^{-1} \cdot \text{s}^{-1})$	0.034
	转子惯量 $J_0/(\text{kg} \cdot \text{m}^2)$	1.09×10^{-6}
	转子阻尼系数 $b_0/(N \cdot m \cdot \text{rad}^{-1} \cdot \text{s}^{-1})$	0.003
	小齿轮惯量 $J_1/(\text{kg} \cdot \text{m}^2)$	2.4×10^{-8}
	小齿轮节圆半径 r_1/mm	5.0
	大齿轮惯量 $J_2/(\text{kg} \cdot \text{m}^2)$	4.0×10^{-6}
	大齿轮节圆半径 r_2/mm	22.5
	丝杠惯量 $J_3/(\text{kg} \cdot \text{m}^2)$	3.18×10^{-5}
减速器	丝杠螺旋角 $\lambda/(\circ)$	6.06
	丝杠导程 s_s/mm	4.0
	丝杠节圆半径 r_3/mm	6.0
	舵轴与丝杠距离 d/mm	28.5
	轴套惯量 $J_4/(\text{kg} \cdot \text{m}^2)$	4.1×10^{-5}
	阻尼器阻尼系数 $b_4/(N \cdot m \cdot \text{rad}^{-1} \cdot \text{s}^{-1})$	0.25
舵面	旋转惯量 $M_{rr}/(\text{kg} \cdot \text{m}^2)$	9.95×10^{-3}

其他非确定参数主要根据理论或试验方法进行估算。如 $k_0 \sim k_3$ 这 4 个刚度参数先用材料力学方法进行计算, 然后根据试验结果进行修正。除舵轴阻尼器外的其余各处阻尼系数是根据工程经验进行设置的。

图 11 ~ 图 14 中的仿真曲线, 是用快速傅里叶变换 (Fast Fourier Transform, FFT) 的方法对 Simulink 模型的“仿真扫频力”输入和“舵轴转角”输出求频响所得。

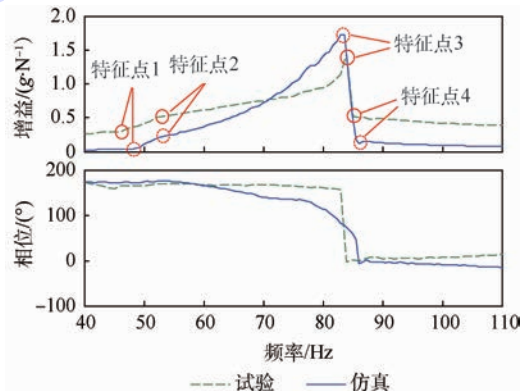


图 11 试验和仿真频响对比 (带接触刚度和间隙)

Fig. 11 Comparison of FRFs from simulation and test (with contact stiffness and freeplay)

图 11 是激振力为 120 N 时,得到的试验和仿真频响对比。仿真频响准确地捕捉到了试验频响中的 4 处特征点(响应突变点)。

作为对照,图 12 给出了考虑接触刚度和不考虑接触刚度时的 2 组仿真曲线。这里不考虑接触刚度指用舵面旋转频率为 84 Hz 对应的常数刚度代替接触刚度。可以看出常数刚度下,频响曲线体现出典型的二阶系统特征,完全失去接触刚度下的几处特征。

图 13 考察的是间隙对舵面旋转频率的影响。随着间隙半宽度 δ 值的增大,频率呈下降趋势。 δ 为 1×10^{-4} 时,换算得到的舵轴间隙大概为 0.1° 。

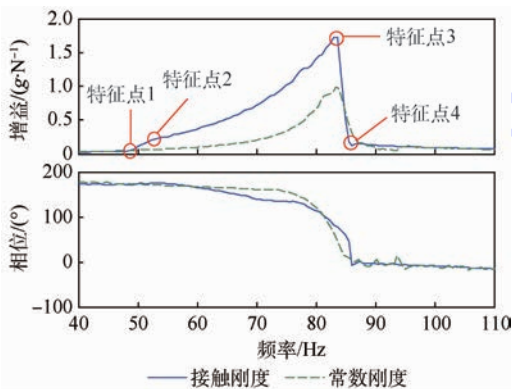


图 12 接触刚度和常数刚度下仿真频响对比(带间隙)
Fig. 12 Comparison of FRFs from simulations under contact stiffness and constant stiffness (with freeplay)

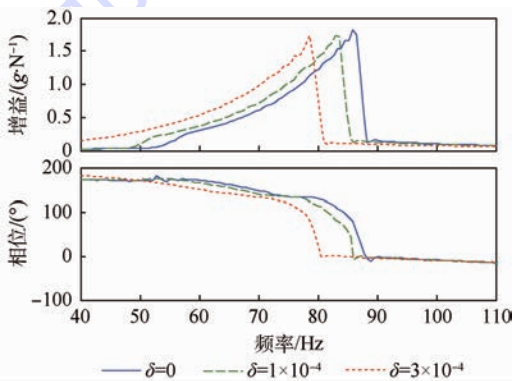


图 13 不同 δ 下频响对比(带接触刚度)
Fig. 13 Comparison of FRFs under different δ (with contact stiffness)

图 14 是不同激振力 F_i 下系统仿真频响曲线的对比。对照图 3, 可以看到仿真结果和试验结果规律一致。

图 15 将不同激振力下舵面旋转频率的仿真结果和图 4 中 4 组舵机(1[#]~4[#])的试验结果进行对比, 同样能够看到一致规律。

在颤振分析环节, 主要对比舵系统在接触刚度和常数刚度 2 种假设下, 气动弹性响应的不同。

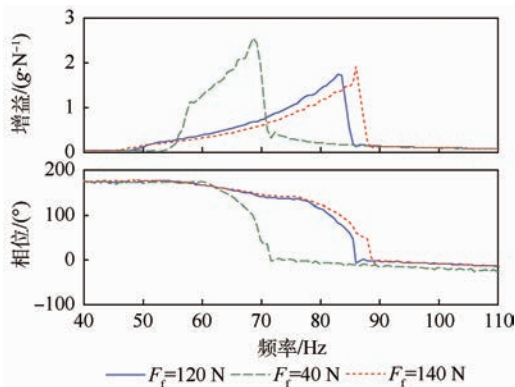


图 14 不同激振力下仿真频响对比
Fig. 14 Comparison of FRFs from simulations under different excitation forces

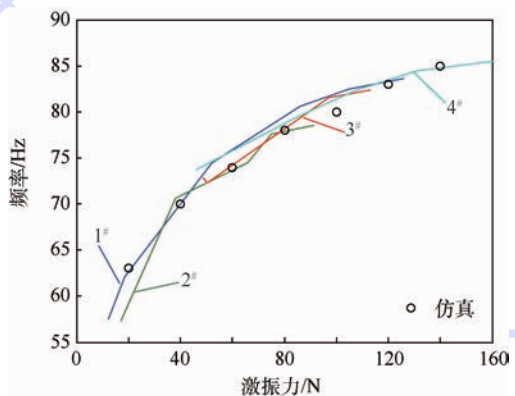


图 15 仿真和试验旋转频率随激振力变化的对比
Fig. 15 Comparison of simulation and experimental data for change trend of fin rotational frequency according to excitation force

舵机模型取经试验数据修正过的模型。常数刚度假设下, 舵面刚体旋转频率依然保持 84 Hz(对应接触刚度假设下激振力为 120 N 时的频率)。舵面弹性分支仅保留一阶弯曲模态, 固有频率 72 Hz。非常气动算法取超音速偶极子网格法。

对于非线性系统, 变量的初始状态对系统稳定性影响很大。飞行器快速机动时, 经常会令舵面快速打出偏角, 以下点考察这类阶跃指令角和来流速度联合作用下舵系统的稳定性。

讨论之前, 首先定义 2 个概念:

临界角: 当前风速下, 系统处于发散和极限环(Limit Cycle Oscillation, LCO)(或收敛)交界的舵机指令角。

临界速度: 无论舵机指令角多小, 系统都会发散的最小速度。

图 16 所示是舵轴旋转角间隙为 0.1° (对应 $\delta = 1 \times 10^{-4}$) 时, 接触刚度和常数刚度假设下的稳定边界对比。横轴为无量纲速度来流速度, 纵轴为舵机指令角大小。

从图 16 中看到,常数刚度下系统响应存在 LCO 和失稳 2 个域,且随着风速提高,临界角渐小;接触刚度下,临界角同样随风速提高而渐小,但与前者相比系统响应多了一个稳定域,这样系统在 LCO 域上下就存在 2 个边界。当风速大于 0.7 时,两边界重合,LCO 域消失。对比两种假设的失稳边界发现,相同风速下,接触刚度的临界角更大。在 0.1° 间隙下,常数刚度的临界速度比接触刚度小 2.5%。

图 17、图 18 所示是不同工况下的舵偏角响应曲线。图中标记 S、E 分别代表振动起始点和终止点(发散情况下为过程点)。

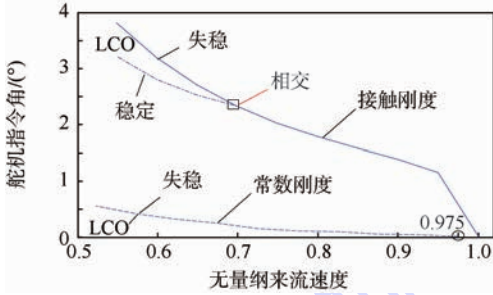


图 16 接触刚度和常数刚度假设下舵系统稳定边界对比(间隙: 0.1°)

Fig. 16 Comparison of stable boundaries of fin-actuator system under contact stiffness and constant stiffness (freeplay: 0.1°)

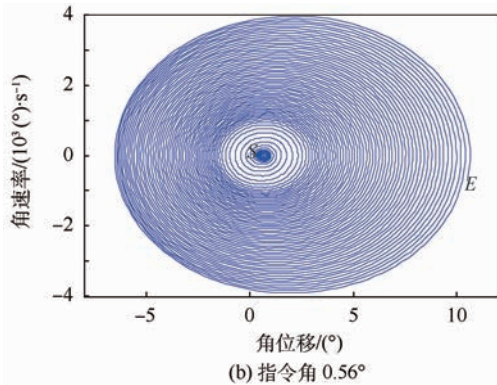
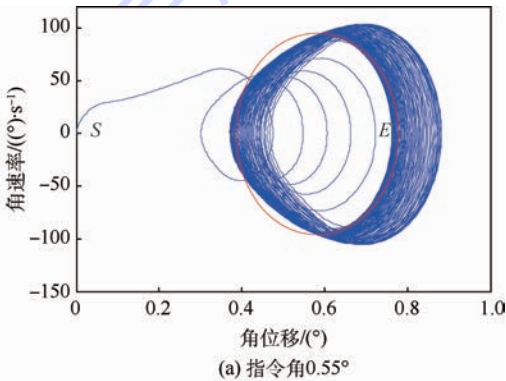


图 17 舵偏角相图(常数刚度,速度: 0.52,间隙: 0.1°)
Fig. 17 Phase diagram of fin shaft angle(constant stiffness, velocity: 0.52, freeplay: 0.1°)

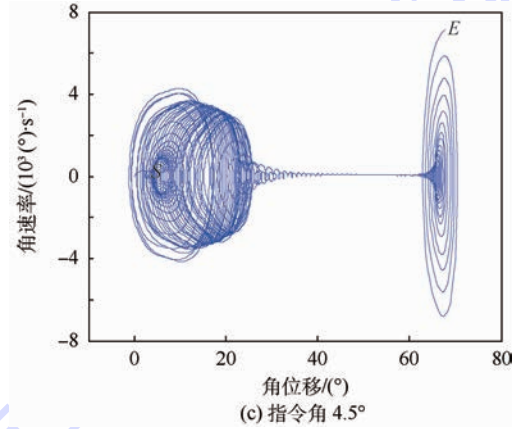
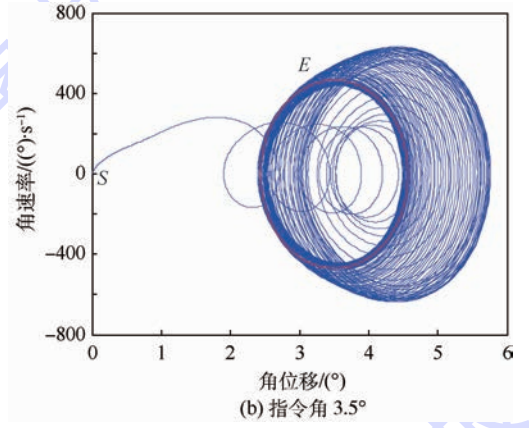
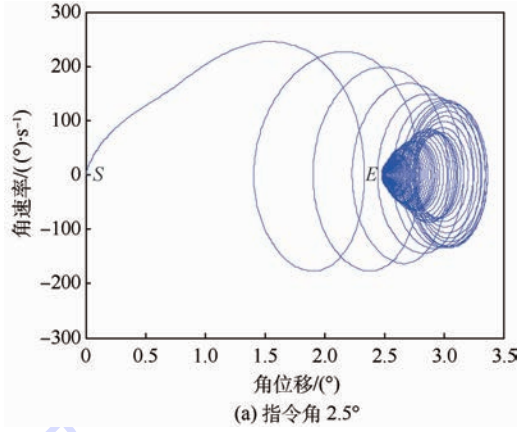


图 18 舵偏角相图(接触刚度,速度 0.55,间隙: 0.1°)
Fig. 18 Phase diagram of fin shaft angle(contact stiffness velocity: 0.55, freeplay: 0.1°)

图 17 为常数刚度,风速 0.52,间隙 0.1° 时,不同指令角下的舵机转角响应相图。阶跃指令角为 0.55° 时,系统维持稳定的 LCO 运动。当指令角增加到 0.56° 时,系统响应发散。同时也看到常数刚度假设下没有收敛域出现。图 18 为接触刚度,风速 0.55,间隙 0.1° ,指令角分别为 2.5° 、 3.5° 和 4.5° 时的舵机转角响应相图。系统分别处于收敛、LCO 和发散状态。其中发散状态特征不同于常数刚度假设,舵偏角没有逐渐打向 90° ,而是在到达 68° 之后,振荡发散。

图 19 所示是接触刚度和常数刚度下舵系统稳定边界在不同间隙下的变化趋势对比。图中箭头所示为趋势走向。可以看到,相同风速下,随着间隙的增大,常数刚度假设的临界角增大。其临界速度在 3 种间隙下均位于 0.975 处,保持不变;对于接触刚度假设,当风速小于 0.75 时,临界角随间隙增大而增大。但当风速大于 0.75 时,临界角却是减小的。系统的临界速度也会随间隙的增大而减小。此外,还发现系统的 LCO 域下边界不会受间隙的影响。

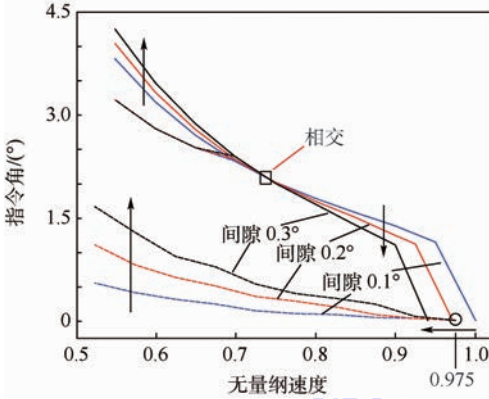


图 19 接触刚度和常数刚度假设下舵系统稳定边界与间隙大小的关系

Fig. 19 Relations between freeplays and stable boundaries of fin-actuator system under contact stiffness and constant stiffness

5 结论

1) 考虑接触刚度和间隙的舵系统动力学模型,准确捕捉到了地面振动试验中的非线性现象:随着激振力的增大,舵面旋转频率逐渐升高,舵机提供的支持刚度在提高。说明以接触刚度为主,间隙为辅的非线性因素组合,共同主导了舵系统整体的动力学行为。

2) 接触刚度下的舵系统气动弹性响应存在稳定、LCO 和失稳 3 个域。当风速大于 0.7 时,LCO 域上下边界重合,LCO 域消失。而常数刚度下,系统仅存在 LCO 和失稳 2 个域。

3) 2 种刚度假设下,随着风速的提高,系统边界角均减小。相同风速下,接触刚度下的边界角一般大于常数刚度下的边界角,更加稳定。

4) 常数刚度下,边界角随间隙增大而增大;接触刚度下,风速低于 0.75 时,边界角与间隙正相关。风速高于 0.75 时,反之。

5) 常数刚度下,边界速度不随间隙变化;而接触刚度下,间隙越大,边界速度越低,甚至会低

于常数刚度下的边界速度。

6) 本文舵面模态频率为弯低旋高。随风速提高,舵机受力增大,此时接触刚度会提高舵机整体刚度,从而拉大舵面弯旋频率差,改善系统稳定性。倘若是弯高旋低的情形,接触刚度的作用很可能是相反的,需谨慎对待。

综上,鉴于 2 种刚度假设下,舵系统动力学行为及气动弹性响应的多处不同,认为工程分析中,应当对舵机的支持特性引起足够重视,尤其要关注接触刚度和其他非线性因素联合作用下可能带来的不利影响。

参考文献 (References)

- [1] YEHEZKELY E, KARPEL M. Nonlinear flutter analysis of missiles with pneumatic fin actuators[J]. Journal of Guidance, Control, and Dynamics, 1996, 19(3): 664-670.
- [2] SHIN W H, LEE I, SHIN Y S, et al. Nonlinear aeroelastic analysis for a control fin with an actuator[J]. Journal of Aircraft, 2007, 44(2): 597-605.
- [3] 张新替, 吴志刚, 杨超. 考虑舵机动力学的舵系统颤振分析[J]. 北京航空航天大学学报, 2011, 37(8): 927-932.
ZHANG X T, WU Z G, YANG C. Flutter analysis for rudder considering actuator's dynamics[J]. Journal of Beijing University of Aeronautics and Astronautics, 2011, 37(8): 927-932 (in Chinese).
- [4] ZHANG X T, WU Z G, YANG C. New flutter-suppression method for a missile fin with an actuator[J]. Journal of Aircraft, 2013, 50(3): 989-994.
- [5] YANG N, WU Z G, YANG C. Structural nonlinear flutter characteristics analysis for an actuator-fin system with dynamic stiffness[J]. Chinese Journal of Aeronautics, 2011, 24(5): 590-599.
- [6] VALENTIN L P. Contact mechanics and frictions[M]. Berlin: Springer Berlin Heidelberg, 1987: 64-68.
- [7] 张学良, 黄玉美, 傅卫平, 等. 粗糙表面法向接触刚度的分形模型[J]. 应用力学学报, 2000, 17(2): 31-35.
ZHANG X L, HUANG Y M, FU W P, et al. Fractal model of normal contact stiffness between rough surfaces[J]. Chinese Journal of Applied Mechanics, 2000, 17(2): 31-35 (in Chinese).
- [8] KATUHIRO N, KAZUKI T. Stiffness of a pre-loaded ball screw[J]. Transactions of the Japan Society of Mechanical Engineers, 1987, 53(492): 1898-1904.
- [9] KRIPA K V, SAMIR A N. The dynamics of lead-screw drives: Low-order modeling and experiments[J]. Journal of Dynamic Systems, Measurement, and Control, 2004, 126(2): 388-396.
- [10] XIAO H, BRENNAN M J, SHAO Y M. On the undamped free vibration of a mass interacting with a Hertzian contact stiffness[J]. Mechanics Research Communications, 2011, 38(8): 560-564.
- [11] KONO D, INAGAKI T, MATSUBARA A, et al. Stiffness model of machine tool supports using contact stiffness[J]. Precision Engineering, 2013, 37(3): 650-657.

- [12] KONO D, NISHIO S, YAMAJI I, et al. A method for stiffness tuning of machine tool supports considering contact stiffness [J]. *International Journal of Machine Tools and Manufacture*, 2015, 90:50-59.
- [13] ZHANG R J, WU Z G, YANG C. Fin-actuator system modeling and experimental validating for aeroelastic research [C] // 56th AIAA/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. Reston: AIAA, 2015.
- [14] GLADWELL G M L. Branch mode analysis of vibrating systems [J]. *Journal of Sound and Vibration*, 1964, 1(1):41-59.
- [15] KARPEL M. Extensions to the minimum-state aeroelastic modeling method [J]. *AIAA Journal*, 1991, 29(11):2007-2009.
- 作者简介:**
张仁嘉 男, 博士研究生。主要研究方向: 气动伺服弹性力学。
Tel.: 010-82313376
E-mail: zhrj559@163.com
- 吴志刚 男, 博士, 副教授。主要研究方向: 气动伺服弹性力学。
Tel.: 010-82317510
E-mail: wuzhigang@buaa.edu.cn

Dynamic modeling and flutter analysis of a fin-actuator system

ZHANG Renjia, WU Zhigang*, YANG Chao

(School of Aeronautic Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: Dynamics of an actuator has a direct impact on the flutter characteristics of a fin structure. Thus it is important to build the dynamic model of the fin-actuator system precisely before a flutter analysis. A typical fin-actuator system is investigated by constructing its motor model, reducer model and controller model, respectively. Based on the characteristics of the ground vibration test (GVT) data, two nonlinear factors are considered including the contact stiffness between the screw and the balls, as well as the freeplay. Simulation results reproduce what were seen in the test. It is found that the contact stiffness and freeplay dominate the dynamic characteristics of the fin-actuator system. In the flutter analysis, the angular responses of the actuator stimulated by different angular step commands under both contact stiffness assumption and constant stiffness assumption are compared. The results show that a limit cycle oscillation (LCO) domain and an unstable domain exist under the constant stiffness. Its critical velocities do not vary with the freeplays. However, besides the LCO and unstable domains, a stable domain appears under the contact stiffness. At a specific flow velocity, its critical command angle is higher than the one under constant stiffness. But its critical velocity decreases as the freeplay increases, and even reaches a lower level than the critical velocity under constant stiffness. Enough attention should be paid to this issue.

Key words: aeroelasticity; flutter; nonlinear; contact stiffness; freeplay; actuator

Received: 2015-07-03; **Accepted:** 2015-08-20; **Published online:** 2015-09-17 10:41

URL: www.cnki.net/kcms/detail/11.2625.V.20150917.1041.004.html

Foundation item: National Natural Science Foundation of China (11372023)

* **Corresponding author.** Tel.: 010-82317510 E-mail: wuzhigang@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0451

磁悬浮反作用飞轮无刷直流电机转矩脉动抑制策略



汤继强*, 王菅许, 周新秀

(北京航空航天大学 仪器科学与光电工程学院, 北京 100083)

摘 要: 无磁链反馈直接转矩控制(DTC)因转矩响应速度快,为小电感磁悬浮反作用飞轮(MSRF)无刷直流电机(BLDCM)的转矩脉动抑制提供了可能,但其 bang-bang 控制器使电机转矩脉动较大。为解决此问题,对影响转矩脉动的换相和非导通相续流过程进行了数学建模,并得出电机转矩与绕组电流的关系,提出一种基于转矩预测的转矩控制方法,能够有效减小转矩脉动,并证明了所提策略的稳定性和鲁棒性。在采用滑模观测器(SMO)进行反电动势估计时,提出一种新的带有参数的光滑连续函数替换符号函数,有效地抑制了滑模观测器的抖振现象。仿真和实验结果表明,所提出的基于改进滑模观测器和转矩预测的改进转矩控制方法相对于传统直接转矩控制能够更好地抑制转矩脉动,而且转矩响应速度基本不变。

关键词: 磁悬浮反作用飞轮(MSRF); 无刷直流电机(BLDCM); 转矩脉动; 转矩预测; 滑模观测器(SMO); 抖振

中图分类号: TM351

文献标识码: A **文章编号:** 1001-5965(2016)07-1377-11

磁悬浮反作用飞轮(Magnet Suspended Reactive Flywheel, MSRF)是航天器姿态控制的常用惯性执行机构,其输出力矩的稳定性直接影响航天器姿态控制的精度。无刷直流电机(Brushless DC Motor, BLDCM)因体积小、能量密度高和易于实现智能控制等优点被广泛用于驱动MSRF,但其换相、非导通相续流等会引起无刷直流电机的转矩脉动,使MSRF输出力矩不稳定。

针对无刷直流电机转矩脉动的研究主要集中在电机结构优化和改进控制策略2个方面^[1]。MSRF无刷直流电机一般采用无齿槽结构,基本消除了齿槽转矩脉动^[2],已经实现了优化。在改进控制策略方面,可进一步细分为分散控制策略和整体控制策略。前者主要针对引起转矩脉动的具体因素,比如换相、非导通相续流、非理想梯形波和电枢反应等;后者主要采用直接转矩控制

(Direct Torque Control, DTC)、自抗扰控制、模糊控制和神经网络控制等方法抑制整体转矩脉动。分散控制策略中,重叠换相法^[3]、PWM斩波法、滞环电流法^[4-5]和电流预测法^[6]等都是针对无刷直流电机换相转矩脉动进行抑制的,并取得了一定效果,但是方法作用区域都是在无刷直流电机的换相阶段或者换相阶段前后,受到换相时间的限制。除换相过程外,非理想反电动势(Electromotive Force, EMF)、非导通相续流等问题均会引起转矩脉动,针对这些因素,也有很多学者进行了研究^[7-10]。但分散控制策略都是针对具体因素,如果对引起无刷直流电机转矩脉动的因素分别进行处理,算法会很复杂,而且有些造成转矩脉动的因素无法进行数学建模和对应的转矩脉动抑制,较好的处理方法是对电机转矩进行估计和控制。

在整体抑制转矩脉动的控制策略中,自抗扰

收稿日期: 2015-07-06; 录用日期: 2015-08-26; 网络出版时间: 2015-10-19 15:20

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151019.1520.005.html

基金项目: 国家自然科学基金(61174003, 61473018, 61403015)

* 通讯作者: Tel.: 010-82339273 E-mail: tjq_72@163.com

引用格式: 汤继强, 王菅许, 周新秀. 磁悬浮反作用飞轮无刷直流电机转矩脉动抑制策略[J]. 北京航空航天大学学报, 2016, 42(7): 1377-1387. TANG J Q, WANG Y X, ZHOU X X. Torque ripple reduction strategy for magnet suspended reactive flywheel brushless DC motor[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1377-1387 (in Chinese).

控制、模糊控制和神经网络控制等方法存在参数整定困难、计算复杂和实时性差等缺点,不利于实现。直接转矩控制以电磁转矩作为直接反馈控制量,通过选择适当的定子电压矢量把转矩和定子磁链误差限制在滞环内^[1]。该方法因受不确定的转矩脉动影响小,得到致力于抑制无刷直流电机转矩脉动工作者的青睐,也得到了广泛深入的研究。

文献[11]指出,120°导通工作方式下无刷直流电机控制特性使其磁链形状为带有6个尖角的不规则圆形,导致传统直接转矩控制对磁链幅值控制精度不高,基于此,文献[12]结合无刷直流电机120°导通工作方式的特点,提出了一种无磁链反馈直接转矩控制方案,因为控制结构简单,为无刷直流电机直接转矩控制普遍采用,本文也采用无磁链反馈直接转矩控制对电机转矩进行控制。

直接转矩控制首先需要对转矩进行估计,滑模观测器(Sliding Mode Observer, SMO)因受电机参数影响小、易于实现,得到了广泛研究和应用。文献[13]在无磁链反馈直接转矩控制的基础上,采用滑模观测器对反电动势进行观测,提高了转矩的观测精度和控制精度,并对存在抖振的反电动势观测结果进行滤波处理,但是没有考虑滤波器导致的相移问题,使反电动势观测不准确。文献[14]针对滑模观测器出现的抖振问题,提出一种分段式滑模变结构直接转矩控制,减小了系统抖振,提高了对反电动势的估计精度,但是该方法难以确定参数,而且影响了稳定精度。文献[15-16]将电阻和电感进行自适应辨识,提高了滑模变观测器对反电动势的估计精度,转矩的控制精度有所提高,但是控制算法复杂。

为了提高直接转矩控制对转矩的控制精度,文献[17]将6个电压矢量增加到12个,得到了平稳的转矩,但控制复杂度增加,而且只是单纯增加电压矢量个数,没有对施加的电压矢量进行限制,对转矩脉动抑制效果的提高改善不大。文献[18]将传统按照数学模型调制开关管的PWM方法与直接转矩控制方法相结合,实现了转矩脉动的最小化,有效抑制了转矩脉动,但是调制开关管的PWM占空比需要提前计算,在运行时进行查表操作,实现复杂。文献[19-20]分别利用模糊控制方法、神经网络控制方法改善直接转矩控制的效果,但其主要问题是算法复杂,不易实现。

传统直接转矩控制的 bang-bang 控制器对大电感无刷直流电机控制效果较好,但对小电感无

刷直流电机作用较差,而 MSRF 无刷直流电机典型特点就是电感较小,因此必须对传统直接转矩控制的 bang-bang 控制器进行改进。基于此,本文首先针对引起转矩脉动的换相和非导通相续流建立了数学模型,分析得到电机转矩与绕组电流关系函数,提出了一种判断电机状态,按数学模型进行转矩预测的转矩控制方法,能够在保持转矩响应速度的情况下减小转矩脉动。同时为减弱传统滑模观测器对反电动势观测过程中出现的抖振,提出一种参数可调的光滑连续函数替换符号函数,大幅削弱了抖振,实现了绕组反电动势以及转矩的精确估计。这2种方法结合形成了一种改进转矩控制方法,仿真和实验结果表明,该方法相对于传统直接转矩控制能大幅度减小无刷直流电机在非换相阶段转矩脉动,对换相阶段转矩脉动也有一定的抑制效果。

1 无刷直流电机的数学模型

1.1 基本数学模型

图1为无刷直流电机三相绕组等效模型及驱动模型。假设电机三相绕组完全对称。图中: R 为每相绕组的电阻; L 为每相绕组等效电感; U_0 为中性点电压; U_d 为直流母线电压; e_A 、 e_B 和 e_C 分别为对应A、B和C相的反电动势; i_A 、 i_B 和 i_C 分别为A、B和C相的电流; u_A 、 u_B 和 u_C 分别为A、B和C相绕组端电压。正方向如图1所示。

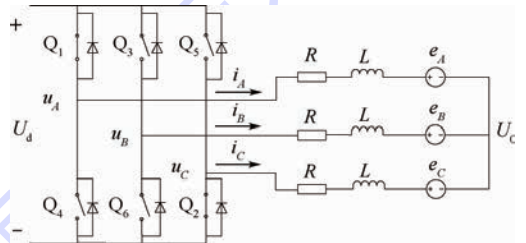


图1 无刷直流电机等效模型及驱动模型

Fig.1 Equivalent model of BLDCM and inverter model

根据图1得无刷直流电机电压平衡方程为

$$\begin{cases} u_A - U_0 = u_{AP} = Ri_A + L \frac{di_A}{dt} + e_A \\ u_B - U_0 = u_{BP} = Ri_B + L \frac{di_B}{dt} + e_B \\ u_C - U_0 = u_{CP} = Ri_C + L \frac{di_C}{dt} + e_C \end{cases} \quad (1)$$

式中: u_{AP} 、 u_{BP} 和 u_{CP} 为三相绕组相电压。

1.2 绕组不同导通状态下数学模型

MSRF 无刷直流电机在运行过程中,造成电机输出力矩脉动的因素主要有:①换相;②非

导通相续流;③ 三相绕组反电动势不是标准梯形波;④ 其他因素。前 2 种因素能够引起转矩脉动,是因为电机导通状态发生变化。由于 MSRF 无刷直流电机结构经过设计,其磁密均匀而且转速一般不高,可在建立数学模型时不考虑非理想梯形波的影响。下面对无刷直流电机正常运行状态、换相状态和非导通相续流状态进行数学建模。

1.2.1 正常运行状态

无刷直流电机正常运行时,电机绕组只有两相导通。根据图 1 所示,电机状态为 AC 相导通,有

$$U_d = 2Ri_A + 2L \frac{di_A}{dt} + 2e$$

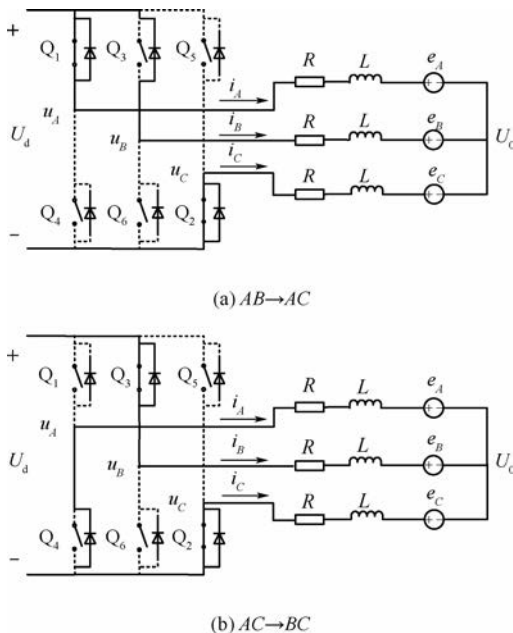
同时有

$$T_e = \frac{e_A i_A + e_B i_B + e_C i_C}{\omega} = \frac{2e i_A}{\omega}$$

式中: $e = e_A$; ω 为转子机械角速度; T_e 为电机电磁转矩。

1.2.2 换相状态

对于实际的无刷直流电机来说,较常见的情况往往是换相转矩减小的情况^[6],因此本文对换相过程中关断相电流下降速率小于开通相电流上升速率的情况不予研究。电机换相有开关管上管导通和下管导通 2 种,分别以 $AB \rightarrow AC$ 、 $AC \rightarrow BC$ 换相过程为例进行分析。这 2 种情况下电机导通状态如图 2 所示。



(a) $AB \rightarrow AC$

(b) $AC \rightarrow BC$

图 2 $AB \rightarrow AC$ 、 $AC \rightarrow BC$ 换相过程无刷直流电机等效电路

Fig. 2 Equivalent circuits of BLDCM in commutation period of $AB \rightarrow AC$, $AC \rightarrow BC$

根据图 2(a),有

$$\begin{cases} U_d = Ri_A + L \frac{di_A}{dt} + e_A + U_0 \\ U_d = Ri_B + L \frac{di_B}{dt} + e_B + U_0 \\ 0 = Ri_C + L \frac{di_C}{dt} + e_C + U_0 \\ i_A + i_B + i_C = 0 \end{cases}$$

则

$$U_d = 3Ri_A + 3L \frac{di_A}{dt} + 4e$$

式中: $e = e_A$ 。

同时有

$$T_e = \frac{(e_A - e_C)i_A}{\omega} + \frac{(e_B - e_C)i_B}{\omega} \approx \frac{2e i_A}{\omega} \quad (2)$$

考虑到 MSRF 无刷直流电机电感很小,换相时间很短,可忽略换相期间反电动势变化。分析式(2),前一项 $(e_A - e_C)i_A/\omega$ 与非换相相电流成正比,后一项 $(e_B - e_C)i_B/\omega$ 与关断相电流成正比,而关断相电流下降速率非常快,同时换相期间, e_B 、 e_C 相差不大,可将后一项忽略。

同理,得到 $AC \rightarrow BC$ 换相期间有

$$U_d = -3Ri_C - 3L \frac{di_C}{dt} + 4e$$

$$T_e = \frac{(e_C - e_B)i_C}{\omega} + \frac{(e_A - e_B)i_A}{\omega} \approx \frac{2e i_C}{\omega}$$

式中: $e = e_C$ 。

其他换相情况以此类推。

电机换相状态的识别,自霍尔信号发生变化到关断相电流降为零为换相阶段。

1.2.3 非导通相续流状态

根据文献[10],非导通相续流是因换相延迟引起。非导通相续流有 2 种形式:上管旁路导通和下管旁路导通。图 3 分别为 CB 相导通换相至 AB 相导通时换相延迟和 BC 相导通换相至 BA 相导通时换相延迟引起的非导通相续流时三相绕组等效电路。

从图 3 可以看出,上管旁路导通时,电机绕组导通状态同上管换相,下管旁路导通时,电机绕组导通状态同下管换相。因此,非导通相续流时,可以认为电机“换相”,非导通相为开通相。例如, CB 导通换相至 AB 导通换相延迟时,可认为 A 为开通相, C 为关断相。在这种假设下,非导通相续流与正常换相情况完全相同。非导通相续流判断条件为非导通相端电压小于零电压(下管旁路导通)或者大于直流母线电压

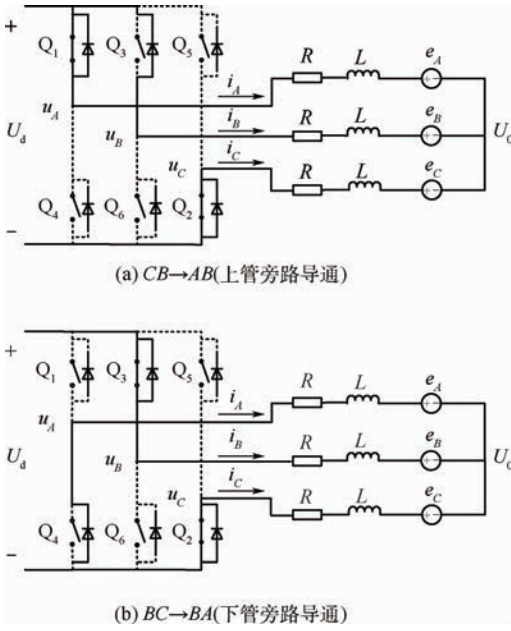


图3 不同非导通相续流情况下无刷直流电机等效电路
Fig.3 Equivalent circuits of BLDCM with diode freewheeling of different inactive phases

(上管旁路导通)。

综上,不同电机绕组导通状态下数学模型可合并如下:

$$U_d = \begin{cases} 2Ri + 2L \frac{di}{dt} + 2e & \text{com} = 0 \\ 3Ri + 3L \frac{di}{dt} + 4e & \text{com} = 1 \end{cases} \quad (3)$$

$$T_e = \frac{2ei}{\omega} \quad (4)$$

式中:com = 0 为电机处于两相导通状态, i 为上管导通相电流, e 为上管导通相反电动势; com = 1 为电机处于换相状态或者非导通相续流状态, i 为非换相相电流, e 为非换相相反电动势。

2 改进型滑模观测器

考虑到三相绕组反电动势有可能不是标准梯形波,而且滑模观测器具有对电机参数不敏感、鲁棒性强的优点,选用滑模观测器对绕组反电动势进行估计。

将式(3)经过 Clark 变换,得到静止坐标系下电压方程如下:

$$\begin{cases} u_\alpha = Ri_\alpha + L \frac{di_\alpha}{dt} + e_\alpha \\ u_\beta = Ri_\beta + L \frac{di_\beta}{dt} + e_\beta \end{cases} \quad (5)$$

式中: u_α 、 u_β 、 i_α 、 i_β 、 e_α 和 e_β 分别为 $\alpha\beta$ 坐标系下的绕组定子电压、定子电流和绕组反电动势。

于是,在两相静止坐标系下无刷直流电机电磁转矩为^[14]

$$T_e = \frac{3}{2} \cdot \frac{p}{2} \left(\frac{e_\alpha i_\alpha}{\omega} + \frac{e_\beta i_\beta}{\omega} \right) \quad (6)$$

式中: p 为电机极数。

考虑到 MSRF 转动惯量大,认为采样过程中,转子转速不变,即 $\alpha\beta$ 坐标系下绕组反电动势不变。以 $x = [i \ e]^T = [i_\alpha \ i_\beta \ e_\alpha \ e_\beta]^T$ 为状态变量,得到系统状态方程如下:

$$\dot{x} = Ax + Bu \quad (7)$$

式中: $A = \begin{bmatrix} A_1 & A_2 \\ 0 & 0 \end{bmatrix}$, $A_1 = -\frac{R}{L}I$, $A_2 = -\frac{1}{L}I$, I 为

单位阵; $B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}$, $B_1 = \frac{1}{L}I$; $u = \begin{bmatrix} u_\alpha \\ u_\beta \end{bmatrix}$ 。

所提出的改进型滑模观测器假设电机绕组电阻和等效绕组电感不变,根据式(7)建立反电动势滑模观测器如下:

$$\dot{\hat{x}} = A\hat{x} + Bu + Kh(\hat{i} - i) \quad (8)$$

式中: $\hat{x} = [\hat{i}_\alpha \ \hat{i}_\beta \ \hat{e}_\alpha \ \hat{e}_\beta]^T$ 为观测状态变量; $K =$

$$\begin{bmatrix} k_1 & 0 & -k_1 h_1 & 0 \\ 0 & k_2 & 0 & -k_2 h_2 \end{bmatrix}^T, k_1, k_2, h_1 \text{ 和 } h_2 \text{ 为滑模}$$

增益; $h(t)$ 为提出的参数可调的光滑连续函数,即

$$h(t) = \begin{cases} \text{sgn}(t) & |t| \geq \frac{\pi}{2\lambda} \\ \sin(\lambda t) & |t| < \frac{\pi}{2\lambda} \end{cases}$$

其中: λ 为函数参数; $\text{sgn}(\cdot)$ 为符号函数。

$h(t)$ 波形如图4所示。可知,函数 $h(t)$ 为奇函数,且 $t > 0$ 时, $h(t) > 0$, $t < 0$ 时, $h(t) < 0$ 。此特点与符号函数类似,将符号函数替换为 $h(t)$ 函数后依旧会使滑模观测器具有状态切换功能。函数 $h(t)$ 有 1 个参数,可以根据系统需求调节 λ 的大小。 $h(t)$ 函数在 $t = 0$ 处收敛至 0。在系统状态进入滑模区域后,相当于减小了滑模观测器的滑模增益,减小抖振现象。

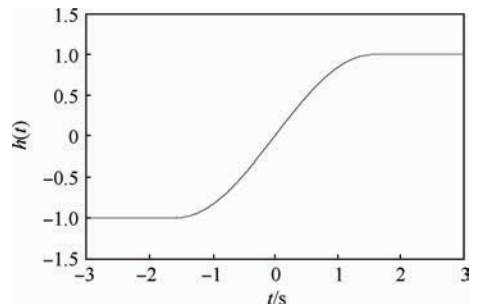


图4 $h(t)$ 波形 ($\lambda = 1$)

Fig.4 Waveform of $h(t)$ ($\lambda = 1$)

式(8)减去式(7)得到

$$\dot{\hat{x}} - \dot{x} = A(\hat{x} - x) + Kh(\hat{i} - i) \quad (9)$$

选取滑模面

$$s = \hat{i} - i = \mathbf{0}$$

于是滑模观测器如下:

$$\dot{\hat{x}} = A\hat{x} + Bu \quad (10)$$

根据观测器进入滑模条件:

$$s^T \dot{s} < 0$$

即

$$\begin{aligned} s^T \dot{s} = & -\frac{R}{L}(\hat{i}_\alpha - i_\alpha)^2 - \frac{R}{L}(\hat{i}_\beta - i_\beta)^2 - \\ & \frac{1}{L}(\hat{i}_\alpha - i_\alpha)(\hat{e}_\alpha - e_\alpha) - \\ & \frac{1}{L}(\hat{i}_\beta - i_\beta)(\hat{e}_\beta - e_\beta) + \\ & k_1 h(\hat{i}_\alpha - i_\alpha)(\hat{i}_\alpha - i_\alpha) + \\ & k_2 h(\hat{i}_\beta - i_\beta)(\hat{i}_\beta - i_\beta) < 0 \end{aligned} \quad (11)$$

又 $h(t)$ 与 t 符号相同,且 $h(t) < 1$ 。只需

保证:

$$\begin{aligned} \frac{1}{L}|\hat{i}_\alpha - i_\alpha||\hat{e}_\alpha - e_\alpha| + \frac{1}{L}|\hat{i}_\beta - i_\beta||\hat{e}_\beta - e_\beta| + \\ k_1|\hat{i}_\alpha - i_\alpha| + k_2|\hat{i}_\beta - i_\beta| < 0 \end{aligned} \quad (12)$$

即可保证式(11)成立,可取:

$$\begin{cases} k_1 < -\frac{\max(|\hat{e}_\alpha - e_\alpha|)}{L} \\ k_2 < -\frac{\max(|\hat{e}_\beta - e_\beta|)}{L} \end{cases} \quad (13)$$

按式(13)选取 k_1 、 k_2 ,可保证所设计的观测器进入滑模状态。根据等效输入理论,观测器进入滑模区域后,有

$$(\hat{i} - i)' = \hat{i} - i = \mathbf{0}$$

结合式(9)得到

$$A_2(\hat{e} - e) + \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix} h(\hat{i} - i) = \mathbf{0}$$

$$(\hat{e} - e)' = \begin{bmatrix} -h_1 & 0 \\ 0 & -h_2 \end{bmatrix} \begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix} h(\hat{i} - i)$$

则

$$(\hat{e} - e)' = -\frac{1}{L} \begin{bmatrix} h_1 & 0 \\ 0 & h_2 \end{bmatrix} (\hat{e} - e) \quad (14)$$

于是,当 $h_1 > 0, h_2 > 0$ 时,反电动势观测误差会以指数形式收敛到零,实现对绕组反电动势估计。

参数 λ 主要影响抖振的收敛速度。对无刷直流电机而言,滑模观测器跟踪的 i_α 、 i_β 信号具有连续阶跃的特点,设 i_α 、 i_β 信号的最大阶跃值为 I ,

则 λ 值的经验取值范围为: $\frac{\pi}{2I} < \lambda < \frac{5\pi}{2I}$ 。

3 转矩控制器

3.1 基于转矩预测的转矩控制器设计

由于采样周期很短,将式(3)、式(4)在 k 时刻进行离散化,有

$$U_d(k) = \begin{cases} 2Ri(k) + \frac{2L(i(k+1) - i(k))}{\Delta t} + \\ 2e(k) & \text{com} = 0 \\ 3Ri(k) + \frac{3L(i(k+1) - i(k))}{\Delta t} + \\ 4e(k) & \text{com} = 1 \end{cases} \quad (15)$$

$$T_e(k+1) = T_e(k) + \frac{2e(k)}{\omega(k)}(i(k+1) - i(k)) \quad (16)$$

式(15)表明,电机在不同运行状态下,可以通过控制 k 时刻直流母线电压来控制 $k+1$ 时刻电机上管导通相电流 $i(k+1)$ 。根据式(16)可知,通过控制 k 时刻直流母线电压可进而间接控制 $k+1$ 时刻电机转矩。

同时, $i(k)$ 、 $\omega(k)$ 可以通过传感器得到, $e(k)$ 、 $T_e(k)$ 可以通过改进滑模观测器得到。电机状态判断方法在第 1.2.2 节和第 1.2.3 节已说明。

令 $T_e(k+1) = T_e^*$, T_e^* 为参考转矩。即使每个下一时刻期望转矩均为参考力矩,有

$$D(k)U = \begin{cases} 2Ri(k) + \frac{\omega(k)L(T_e^* - T_e(k))}{e(k)\Delta t} + \\ 2e(k) & \text{com} = 0 \\ 3Ri(k) + \frac{3\omega(k)L(T_e^* - T_e(k))}{2e(k)\Delta t} + \\ 4e(k) & \text{com} = 1 \end{cases} \quad (17)$$

式中: $D(k)$ 为电机不同状态下,使 $k+1$ 时刻电机转矩理论上为期望转矩时,需要对电机供电电压的调制占空比。

式(17)说明,根据电机状态,按照不同数学模型得出的对电机的供电电压 U 进行占空比为 $D(k)$ 的占空比调制,可以在理论上使得 $k+1$ 时刻电机转矩为期望转矩。

图 5 为实际应用的一种无刷直流电机驱动等效电路。图中: U 为额定电压;二极管 D_0 、电感 L_r 、电容 C_0 和开关管 Q_7 构成 Buck 电路。可以看出,通过 Q_7 可以调制电机供电电压,调制后通过电感 L_r 、电容 C_0 构成的低通滤波器滤波,使实际作用于三相绕组的 U_d 近似于直流电压。如果按式(17)对 Q_7 输入占空比为 $D(k)$ 的数字信号,就可以实现对转矩的控制。

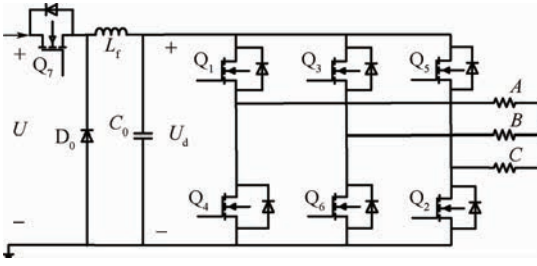


图5 实际应用的无刷直流电机驱动等效电路

Fig.5 Actual used equivalent circuit of inverter of BLDCM

3.2 控制器稳定性和鲁棒性

由式(3)可知,无刷直流电机在换相阶段与非换相阶段数学模型类似,只证明非换相阶段基于转矩补偿策略的转矩控制器的收敛性。

无刷直流电机数学模型是对实际系统的近似,考虑电机参数估计误差,根据式(3)可假设实际系统方程为

$$i = -\frac{R'}{L'}i - \frac{1}{L'}e + \frac{1}{2L'}u \quad (18)$$

式中: u 为系统输入电压; R' 和 L' 分别为无刷直流电机实际绕组电阻和绕组等效电感。

忽略滑模观测器对电机反电动势的估计误差,则转矩补偿策略对应的系统方程为

$$\left(\frac{\omega}{2e}T_c^* - i\right)/\Delta t = -\frac{R}{L}i - \frac{1}{L}e + \frac{1}{2L}u \quad (19)$$

综合式(18)、式(19),可得

$$L'i + (R - R' + L/\Delta t)i - \frac{L}{\Delta t} \cdot \frac{\omega T_c^*}{2e} = 0$$

假设收敛过程中电机转速不变,且设:

$$\Delta i = i - \frac{\omega}{2e}T_c^*$$

则有

$$(\Delta i)' = -\left(\frac{R' - R}{L'} + \frac{L}{L'\Delta t}\right)\Delta i - \frac{R' - R}{L'} \cdot \frac{\omega T_c^*}{2e} \quad (20)$$

由式(20)可知,系统在 $R' = R$ 时稳定,系统在稳定的前提下,控制器对电机电感参数具有鲁棒性;在电机参数估计准确的情况下,有

$$(\Delta i)' = -\frac{1}{\Delta t}\Delta i$$

即电流误差以指数形式收敛到零,系统稳定,且收敛速度与系统控制周期有关。

4 总体控制方案

图6为本文提出的无刷直流电机改进转矩控制系统框图。无刷直流电机利用本身的位置传感器确认转子位置,控制逆变器开关管。采样电机任意两相的电流和绕组端电压以及中性点电压,经过Clark变换得到 $\alpha\beta$ 坐标系下定子电流与定

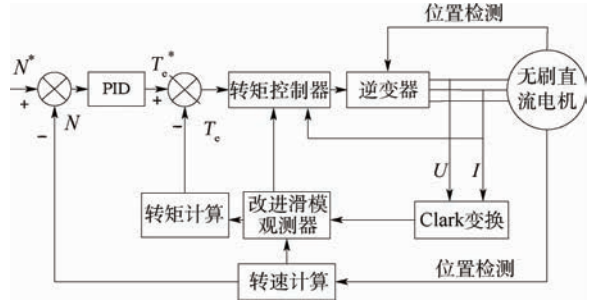


图6 改进转矩控制方法的结构框图

Fig.6 Control framework of advanced torque control method

子电压。利用滑模观测器对绕组反电动势进行估计,进而估计电机转矩。根据此时电机绕组导通状态对应的数学模型,输出电源额定电压的调制占空比,使下一时刻理论电磁转矩为期望电磁转矩。为了保证系统稳定,采用离线绕组电阻测量方案,即利用绕组中热敏电阻测量绕组温度,将不同温度下绕组电阻值存于程序表格中,程序运行时进行查询。图中: N^* 和 N 分别为无刷直流电机的参考转速和实际转速。

5 仿真和实验验证

5.1 仿真验证

根据实际应用的MSRF无刷直流电机参数,利用MATLAB建立无刷直流电机模型。参数为:额定电压 $U = 28\text{V}$,相绕组电阻为 $0.17\ \Omega$,相绕组自感为 0.78mH ,相绕组互感为 0.15mH ,极对数为4,黏滞系数为 $0.0002\text{N}\cdot\text{m}/(\text{rad}\cdot\text{s}^{-1})$,额定负载转矩为 $0.02\text{N}\cdot\text{m}$,反电动势系数为 $0.0079\text{V}/(\text{r}\cdot\text{min}^{-1})$,转子转动惯量为 $0.0273\text{kg}\cdot\text{m}^2$,电机加速阶段最大参考力矩为 $0.302\text{N}\cdot\text{m}$ 。

为验证改进滑模观测器能否削弱传统滑模观测器存在的抖振,在指令转速为 $400\text{r}/\text{min}$ 、仿真(采样、控制)频率为 50kHz 条件下进行仿真验证, e_β 的观测误差 $\Delta e_\beta = \hat{e}_\beta - e_\beta$ 仿真结果如图7(a)所示。改进型滑模增益参数为: $k_1 = k_2 = -10000$, $h_1 = h_2 = 2.5$, $\lambda = 3$;传统滑模增益参数为: $k_1 = k_2 = -10000$, $h_1 = h_2 = 2.5$ 。

更改指令转速为 $800\text{r}/\text{min}$,再次仿真,结果见图7(b)。滑模增益参数保持不变。

从图7可知,利用所提出的改进滑模观测器相对传统滑模观测器对定子绕组反电动势的估计值更平滑,基本消除了抖振现象。由于滑模观测器跟踪的实际电流基本上是阶段性阶跃电流,造成了滑模观测器对 e_β 的观测存在阶跃性的观测误差。

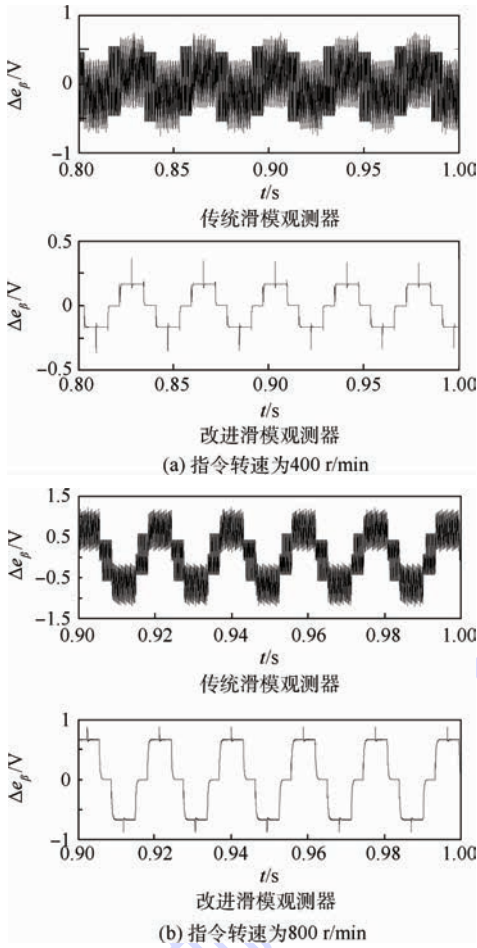


图7 400 r/min与800 r/min指令转速下绕组反电动势估计误差仿真结果

Fig.7 Simulation results of back-EMF estimation error under command speeds of 400 r/min and 800 r/min

再验证基于转矩预测的转矩控制方法(称改进方法)是否相对于传统无磁链反馈直接转矩控制方法(称传统DTC方法)能够更好地抑制转矩脉动。2种方法都采用改进滑模观测器,滑模增益参数不变,设传统DTC方法的转矩滞环宽度为 $0.005\text{ N}\cdot\text{m}$,仿真频率为 50 kHz 。图8为指令转速为 400 r/min (初始转速 350 r/min)时转矩仿真结果。可以得到,电机转速为 $350\sim 400\text{ r/min}$ 情况下,传统DTC方法转矩脉动 36.7% (按加速阶段计算),改进方法转矩脉动在换相时为 16.7% (按加速阶段计算)。

将图8一个周期内细节展开如图9所示,且加入电机状态信号、2种方法的控制器输出信号进行对比。图9(a)中纵坐标表示电机运行状态,1为换相阶段,0为非换相阶段,可以看出,传统DTC控制器输出电压在 0 和 28 V 间跳变,转矩波形明显超过了设定的转矩滞环宽度。改进方法在换相阶段对电机转矩进行补偿后,在换相阶段完成后 0.0005 s 内收敛到参考转矩 $0.302\text{ N}\cdot\text{m}$,

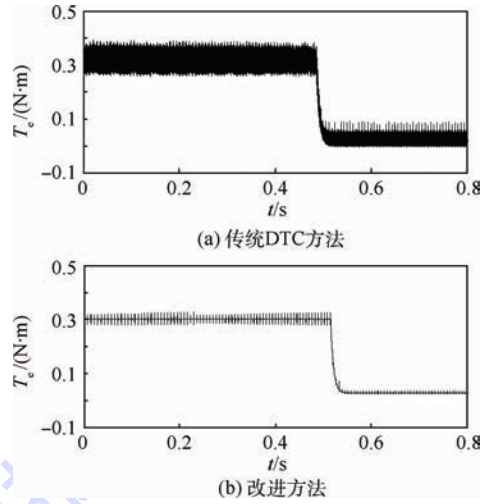


图8 指令转速为 400 r/min 时转矩仿真结果

Fig.8 Simulation results of torque under command speed of 400 r/min

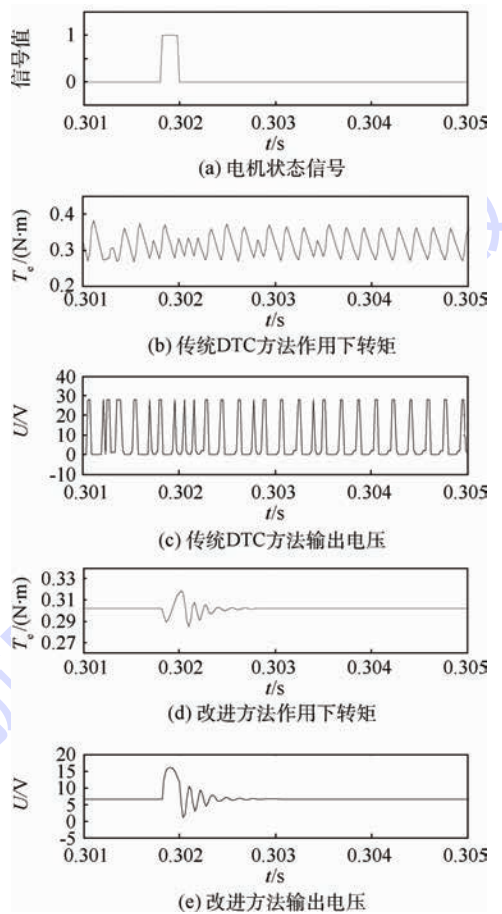


图9 单位周期内2种方法转矩结果与输出电压对比
Fig.9 Comparison of torque and output voltage in one unit cycle between two methods

在换相阶段由于转矩估计误差的存在,使转矩控制精度较差,超过了参考转矩,形成了新的脉动。但是总体来说,改进方法能够减小整个运行阶段的转矩脉动,尤其在非换相阶段相对于传统DTC方法得到平稳的转矩。

图 10 为指令转速为 800 r/min(初始转速为 750 r/min)时 2 种转矩控制方法的转矩仿真结果。可以看出,改进方法在加速阶段和稳速阶段均比传统 DTC 方法转矩控制精度高。

图 10 细节以及对应的控制器输出电压如图 11 所示。可以看出,0.200~0.210 s 之间,电机转矩上限边界很整齐,这是由于电机转速提高,改进方法为补偿跌落转矩,而输出的电压相等,在基本相同转速下,对转矩的补偿效果相同。同样,改进方法作用下非换相阶段转矩很平稳。

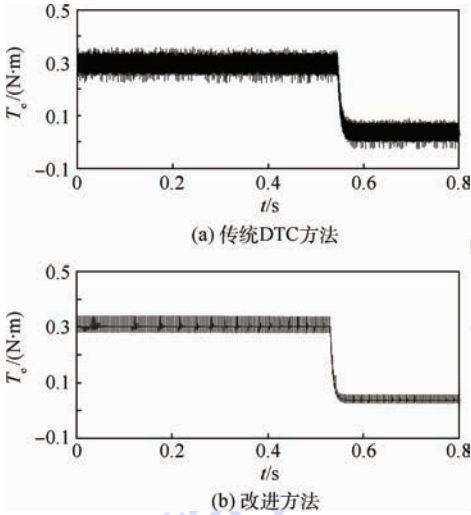


图 10 指令转速为 800 r/min 时转矩仿真结果
Fig. 10 Simulation results of torque under command speed of 800 r/min

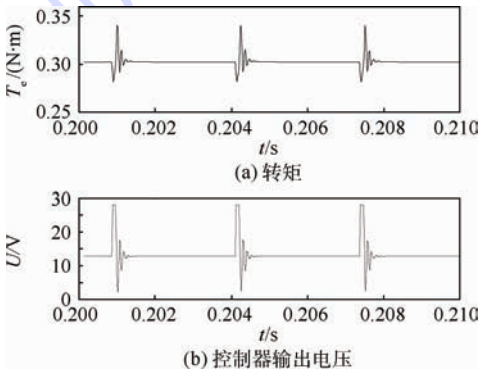


图 11 改进方法转矩和控制器输出电压波形
Fig. 11 Torque and output voltage waveform of controller by advanced method

为了验证相对于传统 DTC 方法,改进方法是否会大幅度降低转矩响应速度,在电机转速稳定为 600 r/min 时,负载转矩从 0.02 N·m 突变为 0.10 N·m,2 种方法作用下转矩响应结果见图 12。可以得到,传统 DTC 方法和改进方法在负载力矩突变时,响应时间均在 0.02 s 左右,说明改进方法相对于传统 DTC 方法没有大幅度降低转矩响应速度。

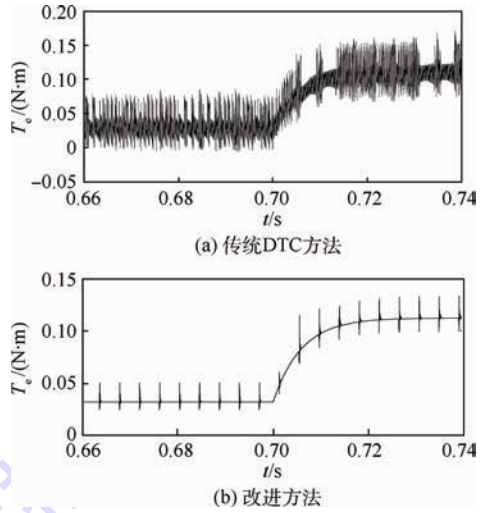


图 12 负载突变时转矩响应仿真结果
Fig. 12 Simulation results of torque response with step load

5.2 实验验证

为了进一步验证所提基于改进滑模观测器和转矩预测的转矩控制方法的有效性,在 MSRF 无刷直流电机上进行实验。实验用 MSRF 无刷直流电机参数与所建立的仿真电机模型相同。由于电机的绕组反电动势和转矩均无法直接测量,因此在实验中,选择将电机的估计定子电流、估计绕组反电动势和估计转矩通过串口上传至上位机,然后利用 MATLAB 绘制图像。

首先验证所提出的改进滑模观测器能否减小传统滑模观测器中的系统抖振。在电机加速阶段(转速接近 600 r/min)进行实验,得到传统滑模观测器与改进滑模观测器的 i_β 跟踪结果见图 13,2 种方法对 e_β 的估计波形见图 14。

从图 13 可以看出,实验条件下,传统滑模观测器对 i_β 的跟踪效果中存在很大的抖振,抖振幅度达到 1.5 A,改进滑模观测器作用下, i_β 的抖振

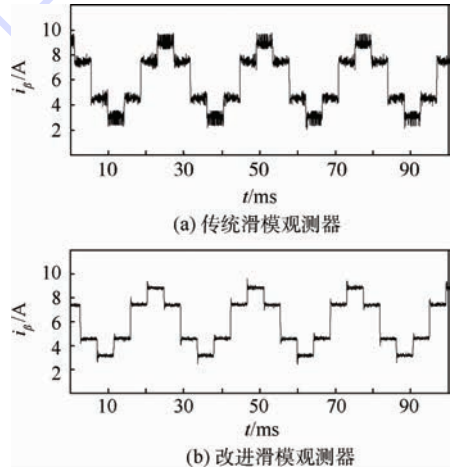
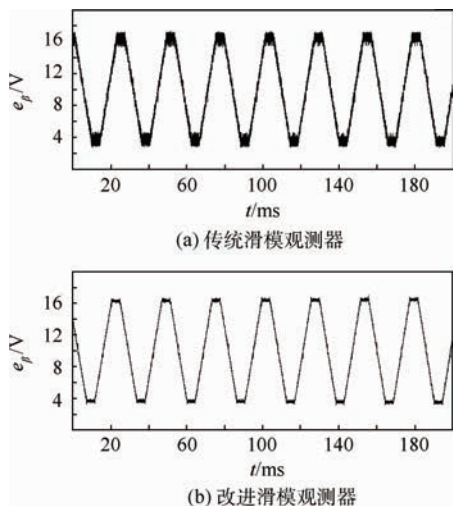


图 13 传统和改进滑模观测器的 i_β 跟踪结果
Fig. 13 Tracking results of i_β by traditional and advanced SMOs

图 14 传统和改进滑模观测器对 e_{β} 的估计结果Fig. 14 Estimation results of e_{β} by traditional and advanced SMOs

幅度降到了 0.4 A 左右。这种情况下,改进滑模观测器相对于传统滑模观测器能够更准确地估计绕组反电动势,图 14 的实验结果也说明了此问题。明显,利用新提出的函数替换符号函数的改进滑模观测器能够有效抑制系统抖振。同时由于三相电流和电压测量误差、滑模观测器的参数整定不准确,以及滑模观测器原理上无法消除跟踪误差、系统外来干扰等问题的存在,改进滑模观测器对 i_{β} 的跟踪仍旧存在一定的抖振。

为了验证改进方法相对于传统 DTC 方法能否更好地抑制转矩脉动,在电机处于加速阶段(转速接近 600 r/min)时,进行实验验证。传统 DTC 方法也采用改进滑模观测器估计绕组反电动势,进而估计电机转矩,转矩滞环宽度为 $0.005 \text{ N} \cdot \text{m}$ 。图 15 为 2 种方法作用下转矩实验结果。

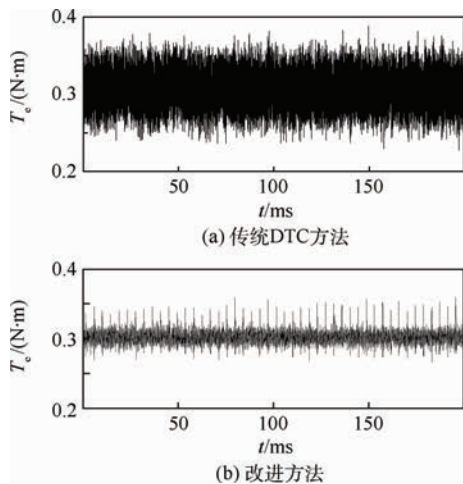


图 15 传统 DTC 方法和改进方法转矩实验波形

Fig. 15 Two experimental curves of torque by traditional DTC method and advanced method

从图 15 可以看出,实验条件下,传统 DTC 方法作用下,电机转矩脉动为 40% 左右,改进方法作用下,在换相阶段转矩脉动近似为 26% ,非换相阶段转矩脉动近似为 12% 。由于实验条件下传统 DTC 方法也采用改进滑模观测器对转矩进行估计,实验结果说明,改进方法能够有效地改进传统 DTC 方法的 bang-bang 观测器的缺陷,进而更好地抑制转矩脉动。

在实验条件下,电机转矩三相电流以及电压的测量误差和测量时间差的存在造成对转矩的估计是不准确的,而且改进滑模观测器不可能完全消除系统抖振,进而消除绕组反电动势的估计误差,造成电机的转矩估计存在误差。另外,实验对象电机绕组的电阻参数不对称、逆变器开关器件正向导通电阻和二极续流导通电阻的存在也影响所提方法的精度。本实验系统中采用的集成逆变器芯片 MSK4300 在使用过程中的发热问题也会影响系统的参数。这些原因导致实验条件下所提出的改进方法转矩脉动抑制效果相对于仿真条件有所下降。

为了验证所提出的改进方法的转矩响应速度是否下降,在实验条件下,电机处于加速状态时,在电机转速超过 600 r/min 时,将给定最大参考转矩从 $0.07 \text{ N} \cdot \text{m}$ 增大至 $0.32 \text{ N} \cdot \text{m}$,实验结果如图 16 所示。

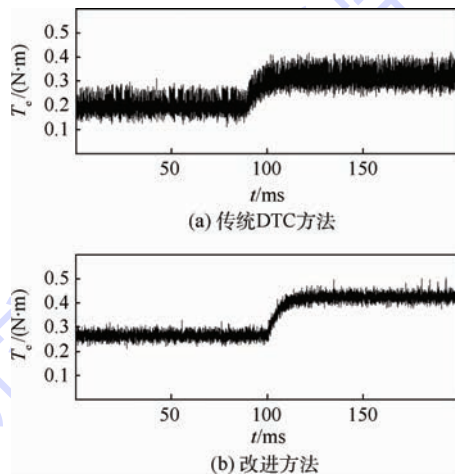


图 16 实验条件下负载突变时转矩响应曲线

Fig. 16 Response curves of torque under mutation loads in the condition of experiment

从图 16 可以看出,在实验条件下,传统 DTC 方法和改进方法基本都在 20 ms 左右完成转矩的重新稳定,2 种方法具有基本相同的转矩响应速度,这与仿真结果相符。

6 结 论

本文针对造成转矩脉动的换相过程和非导通续流建立了数学模型,提出了一种基于数学模

型的转矩预测的转矩控制方法。针对传统滑模观测器的抖振问题,提出了一种含参数的光滑连续函数代替符号函数,形成了一种改进滑模观测器。仿真和实验结果说明:

1) 改进滑模观测器能够有效抑制系统抖振,提高了绕组反电动势和转矩的估计精度。

2) 基于改进型滑模观测器和转矩预测相结合的改进转矩控制方法能够有效抑制 MSRF 无刷直流电机运行过程中的转矩脉动,提高了输出转矩的稳定性。

参考文献 (References)

- [1] 巫庆辉,邵诚,徐占国. 直接转矩控制技术的发展现状与发展趋势[J]. 信息与控制, 2005, 34(4): 444-450.
WU Q H, SHAO C, XU Z G. Survey of research status quo and development trends about direct torque control[J]. Information and Control, 2005, 34(4): 444-450 (in Chinese).
- [2] 王志强,赵波,周新秀. 反作用飞轮电机换相转矩脉动分析[J]. 微特电机, 2010, 38(10): 1-4.
WANG Z Q, ZHAO B, ZHOU X X. Commutation torque ripple analysis of brushless DC motor for reaction flywheels[J]. Small & Spacial Electrical Machines, 2010, 38(10): 1-4 (in Chinese).
- [3] YOSHIDA M, MURAI Y, TAKADA M. Noise reduction by torque ripple suppression in brushless DC motor[C] // 29th Annual IEEE Power Eletronics Specialist Conference. Piscataway, NJ: IEEE Press, 1998, 2: 1397-1401.
- [4] KIM G H, KANG S J, WON J S. Analysis of the commutation torque ripple effect for BLDCM fed by HCRPWMVSI[C] // 7th Annual Applied Power Electronics Conference and Exposition. Piscataway, NJ: IEEE Press, 1992: 277-284.
- [5] 杨进,杨向宇. 一种减小无刷直流电机纹波转矩的新方法[J]. 微电机, 2005, 38(1): 9-11.
YANG J, YANG X Y. The application of tysteresis control in reducing ripple torque of brushless DC motor[J]. Micromotors, 2005, 38(1): 9-11 (in Chinese).
- [6] 林平,韦鲲,张仲超. 新型无刷直流电机换相转矩脉动抑制控制方法[J]. 中国电机工程学报, 2006, 26(3): 153-158.
LIN P, WEI K, ZHANG Z C. A novel control scheme to suppress the commutation torque ripple in BLDCM[J]. Proceedings of the CSEE, 2006, 26(3): 153-158 (in Chinese).
- [7] SONG J H, CHOY I. Commutation torque ripple reduction in brushless DC motor drives using a single DC current sensor[J]. IEEE Transactions on Power Electronics, 2004, 19(2): 312-319.
- [8] 任军军. 永磁无刷直流电机的转矩脉动抑制的控制策略研究[D]. 杭州: 浙江大学, 2004.
REN J J. Research on torque ripple minimization techniques for PM brushless DC motor[D]. Hangzhou: Zhejiang University, 2004 (in Chinese).
- [9] 陈冬,房建成. 非理想梯形波反电动势永磁无刷直流电机换相转矩抑制方法[J]. 中国电机工程学报, 2008, 28(30): 79-83.
CHEN D, FANG J C. Commutation torque ripple reduction in PM brushless DC motor with nonideal trapezoidal back EMF[J]. Proceedings of the CSEE, 2008, 28(30): 79-83 (in Chinese).
- [10] 郭方正,韩邦成,刘刚. PAM 调制方式下高速无刷直流电机非导通相续流抑制方法研究[J]. 微电机, 2009, 42(8): 42-46.
GUO F Z, HAN B C, LIU G. Research on suppression for diode freewheeling of inactive phase in high-speed BLDCM using PAM strategy[J]. Micromotors, 2009, 42(8): 42-46 (in Chinese).
- [11] OZTURK S B, TOLIYAT H A. Direct torque control of brushless DC motor with non-sinusoidal back-EMF[C] // IEEE IEMDC 2007: Proceedings of the International Electric Machines and Drives Conference. Piscataway, NJ: IEEE Press, 2007.
- [12] 林海,梁中,闫茂德,等. 无刷直流电机改进型直接转矩控制研究[J]. 电气传动, 2014, 44(8): 11-14.
LIN H, LIANG Z, YAN M D, et al. Improved direct torque control in brushless DC motor drives[J]. Electric Drive, 2014, 44(8): 11-14 (in Chinese).
- [13] 安群涛,孙立志,刘超,等. 无刷直流电机的磁链自控直接转矩控制[J]. 中国电机工程学报, 2010, 30(12): 86-92.
AN Q T, SUN L Z, LIU C, et al. Flux linkage self-control based direct torque control of brushless DC motor[J]. Proceedings of the CSEE, 2010, 30(12): 86-92 (in Chinese).
- [14] 朱俊杰,粟梅,王湘中,等. 分段式滑模变结构无刷直流电机直接转矩控制[J]. 仪器仪表学报, 2013, 34(11): 2634-2640.
ZHU J J, SU M, WANG X Z, et al. Direct-torque-control of brushless DC motors based on segmented sliding-mode-variable-structure[J]. Chinese Journal of Scientific Instrument, 2013, 34(11): 2634-2640 (in Chinese).
- [15] 郭鸿浩,周波,左广杰,等. 无刷直流电机反电势自适应滑模观测[J]. 中国电机工程学报, 2011, 31(21): 142-149.
GUO H H, ZHOU B, ZUO G J, et al. Adaptive sliding-mode observer for back electromotive force estimation of brushless DC motor[J]. Proceedings of the CSEE, 2011, 31(21): 142-149 (in Chinese).
- [16] 郭鸿浩,周波,左广杰,等. 无刷直流电机转矩观测与电感自适应辨识[J]. 中国电机工程学报, 2011, 31(33): 151-158.
GUO H H, ZHOU B, ZUO G J, et al. Torque estimation and adaptive inductance identification for brushless DC motor[J]. Proceedings of the CSEE, 2011, 31(33): 151-158 (in Chinese).
- [17] 周衍,张兴华. 一种简易的无刷直流电动机直接转矩控制[J]. 微特电机, 2014, 42(7): 54-60.
ZHOU Y, ZHANG X H. A kind of simplified direct torque control for brushless DC motor[J]. Small & Spacial Electrical Machines, 2014, 42(7): 54-60 (in Chinese).
- [18] 李珍国,章松发,周生海,等. 考虑转矩脉动最小化的无刷直流电机直接转矩控制系统[J]. 电工技术学报, 2014, 29(1): 139-146.
LI Z G, ZHANG S F, ZHOU S H, et al. Direct torque control of brushless DC motor considering torque ripple minimization[J].

Transaction of China Electrotechnical Society, 2014, 29(1): 139-146 (in Chinese).

- [19] 王晓远,田亮,冯华. 无刷直流电机直接转矩模糊控制研究[J]. 中国电机工程学报, 2006, 26(15): 134-138.

WANG X Y, TIAN L, FENG H. Study of the fuzzy control on direct torque control for brushless BLDCM [J]. Proceedings of the CSEE, 2006, 26(15): 134-138 (in Chinese).

- [20] 许鹏,郭桂芳,曹军义,等. 直流无刷电机神经网络直接转矩控制[J]. 中国电机工程学报, 2009, 29(Suppl.): 192-196.

XU P, GUO G F, CAO J Y, et al. Neural network control on direct torque control for brushless DC motor [J]. Proceedings of

the CSEE, 2009, 29(Suppl.): 192-196 (in Chinese).

作者简介:

汤继强 男,博士,副教授。主要研究方向:磁悬浮飞轮及磁悬浮控制力陀螺技术。

Tel.: 010-82339273

E-mail: tjq_72@163.com

王营许 男,硕士研究生。主要研究方向:磁悬浮飞轮无刷直流电机控制。

E-mail: wangyingxu@buaa.edu.cn

Torque ripple reduction strategy for magnet suspended reactive flywheel brushless DC motor

TANG Jiqiang*, WANG Yingxu, ZHOU Xinxiu

(School of Instrumentation Science & Opto-electronics Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: Due to its high torque response speed, direct torque control (DTC) without flux feedback makes the torque ripple reduction of magnet suspended reactive flywheel (MSRF) brushless DC motor (BLDCM) possible, which has a small inductance. However, the bang-bang controller of the DTC causes high torque ripples. To solve this problem, the mathematic models of commutation and diode freewheeling of the inactive phase were established and the relationship between torque and phase current was obtained. Then a torque prediction method based on the established models was proposed to reduce the torque ripple effectively, and the stability and robustness were proved. While estimating the back electromotive force of the BLDCM with sliding mode observer (SMO), a smooth and continuous function with one parameter was employed to replace the sign function which leads to a higher back electromotive force estimation accuracy. Simulation and experimental results show that compared to the traditional DTC, the advanced torque control method with torque prediction and advanced SMO can significantly suppress the torque ripples and has almost the same torque response speed.

Key words: magnet suspended reactive flywheel (MSRF); brushless DC motor (BLDCM); torque ripple; torque prediction; sliding mode observer (SMO); chattering

Received: 2015-07-06; Accepted: 2015-08-26; Published online: 2015-10-19 15:20

URL: www.cnki.net/kcms/detail/11.2625.V.20151019.1520.005.html

Foundation items: National Natural Science Foundation of China(61174003,61473018,61403015)

* Corresponding author. Tel.: 010-82339273 E-mail: tjq_72@163.com

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0477

空基伪卫星组网部署的航路规划算法



曾凌川, 李大朋, 曲艺, 任爱爱, 巩应奎*

(中国科学院光电研究院, 北京 100094)

摘 要: 针对以临近空间飞艇为平台的空基伪卫星组网问题, 提出了一种空基伪卫星从初始位置到目标位置进行组网部署的航路规划算法。首先, 通过梳理空基伪卫星组网的性能指标, 设计了基于距离和几何精度因子(GDOP)的航路规划代价函数; 然后, 对航路规划的约束条件进行了分析, 在稀疏A*算法的基础上, 设计了距离代价和GDOP代价权重可自适应调整的空基伪卫星航路规划算法, 并对规划出的航路采用Dubins曲线进行平滑; 最后, 通过仿真实验, 对提出的航路规划算法进行了验证。仿真实验结果表明, 航路规划算法在GDOP较恶劣的情况下, 能够有效地在后续的航路规划中降低GDOP代价, 提高用户定位精度; 而在GDOP较为良好时, 则可以增加距离代价的权重, 在航路规划时使伪卫星尽快向目标位置收敛, 减少时间消耗。

关键词: 导航增强; 空基伪卫星; 航路规划; 稀疏A*搜索算法; Dubins曲线

中图分类号: V249.31

文献标识码: A **文章编号:** 1001-5965(2016)07-1388-10

伪卫星是一种成本低廉、易于部署的导航增强平台。在一些导航卫星可视性降低、卫星导航信号被干扰的情况下, 通过部署伪卫星, 可在卫星导航服务性能下降的情况下为特定区域的用户提供导航增强服务。根据部署空间区域的不同, 伪卫星可分为空基和地基(海基)两类。与地基伪卫星相比, 空基伪卫星无论是在覆盖范围还是部署灵活性方面都具有前者不可比拟的优势。在各类空基平台中, 临近空间飞艇飞行动态性较小、工作高度较高, 且具有一定的悬停能力, 不仅有利于平台本身的定位, 而且具有更大的导航增强服务范围。因此, 基于临近空间飞艇的空基伪卫星技术已经成为当前区域导航增强领域的研究热点。

目前, 澳大利亚、美国、印度和日本等国已经对基于临近空间飞艇的伪卫星导航增强系统进行了建模仿真和半实物外场试验。文献[1]将伪卫

星作为差分基准站, 对伪卫星位置误差、信号的多径效应对流层延迟进行了建模仿真; 文献[2]模拟了城市峡谷环境, 对部署伪卫星来改善导航源的几何精度因子(Geometric Dilution Of Precision, GDOP)和提高用户可见星数量进行了仿真分析; 文献[3]分析了空基和地基伪卫星的协同定位问题, 通过采用载波相位二次差分的办法来提高定位精度, 对伪卫星GDOP恶劣导致的用户定位精度下降进行补偿; 文献[4]模拟了由4颗基于平流层飞艇的空基伪卫星、6个地面监测站和1个控制中心组成的导航增强系统, 对GPS不可用情况下的用户伪距测量误差进行了仿真分析; 文献[5-6]重点研究了利用地面监测站对基于平流层飞艇的空基伪卫星进行“倒GPS”定位的方法, 并进行了半实物外场试验, 通过差分来提高对伪卫星的定位精度。

收稿日期: 2015-07-16; 录用日期: 2015-10-23; 网络出版时间: 2015-12-17 10:41

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151217.1041.005.html

基金项目: 国家自然科学基金(91438207)

* 通讯作者: Tel.: 010-82178658 E-mail: ykgong@aoe.ac.cn

引用格式: 曾凌川, 李大朋, 曲艺, 等. 空基伪卫星组网部署的航路规划算法[J]. 北京航空航天大学学报, 2016, 42(7): 1388-1397.
ZENG L C, LI D P, QU Y, et al. Route planning algorithm for configuring airborne pseudolites[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1388-1397 (in Chinese).

国内对基于临近空间飞艇的相关研究重点集中于伪卫星平台的组网方案和编队控制上。文献[7]提出了一种利用平流层飞艇提高北斗区域导航系统 GDOP 性能的方法;文献[8-10]重点分析了空基伪卫星高度角和方位角等因素对 GDOP 的影响,并以此为基础提出了空基伪卫星的布局方案;文献[11]则采用遗传算法来计算使 GDOP 最优的伪卫星几何布局;文献[12]考虑了基于无人机的空基伪卫星编队的协同控制问题,采用“长机-僚机”方式的控制策略,设计了一种基于神经网络的编队控制器。

从国内外的研究现状可知,空基伪卫星组网的主要任务之一是设计伪卫星平台的最优几何构型来降低用户定位误差,具体表现为以 GDOP 作为优化目标,通过采用经验法或最优化算法来计算伪卫星的最优部署位置。这些研究通常仅给出伪卫星最终的静态组网结果,而未考虑伪卫星以什么样的方式到达部署位置。事实上,伪卫星在部署时总是存在从初始位置向目标位置的运动过程,此时伪卫星编队的几何构型将会随着伪卫星本身运动状态的变化而变化,导致用户所在区域的 GDOP 也在不断改变。

临近空间飞艇运动速度较慢,组网部署通常需要一定时间。一般用户普遍对导航服务有着连续性和完好性的要求,在伪卫星部署过程中也可能会有定位和定时方面的需求。因此,针对国内外相关研究的不足之处,有必要研究飞艇编队在组网部署过程中的航路规划问题,通过设计合适的航路,使飞艇编队在运动过程中尽量保持良好的几何构型,降低用户在伪卫星组网过程中的定位误差。针对此问题,本文首先梳理了空基伪卫星平台组网性能指标;然后指出了伪卫星航路规划问题中不同于一般无人飞行器航路规划的特殊性,设计了一种基于稀疏 A* 算法的伪卫星航路规划算法;最后对提出的空基伪卫星航路规划算法进行了仿真验证。

1 空基伪卫星组网性能指标

从用户定位和定时的角度出发,伪卫星组网性能的优劣可用 GDOP 来衡量。GDOP 定义为伪卫星伪距测量误差到定位定时误差的放大倍数,其反映了导航源几何位置对用户定位误差的影响。GDOP 的计算方法如下:

设用户接收机的三维位置和钟差分别为 (x_u, y_u, z_u) 和 t_u ,在校正对流层传播延迟等误差项后,用户接收机到伪卫星的伪距测量值为

$$\rho_j = \sqrt{(x_j - x_u)^2 + (y_j - y_u)^2 + (z_j - z_u)^2} + ct_u \quad (1)$$

式中: (x_j, y_j, z_j) 为第 j 颗伪卫星的三维位置坐标; c 为光速。

设用户近似位置和近似钟差分别为 $(\tilde{x}_u, \tilde{y}_u, \tilde{z}_u)$ 和 \tilde{t}_u ,将式(1)在 $(\tilde{x}_u, \tilde{y}_u, \tilde{z}_u, \tilde{t}_u)$ 处进行泰勒展开并保留 1 阶项,得到

$$\Delta\rho_j = a_{xj}\Delta x_u + a_{yj}\Delta y_u + a_{zj}\Delta z_u - c\Delta t_u \quad (2)$$

式中: $(\Delta x_u, \Delta y_u, \Delta z_u)$ 为用户近似位置与真实位置之差; (a_{xj}, a_{yj}, a_{zj}) 为用户位置到第 j 颗伪卫星的方向余弦。

假设伪卫星数量为 n ,则式(2)可扩展为矩阵形式:

$$\Delta\rho = H\Delta x \quad (3)$$

式中:

$$\left\{ \begin{array}{l} \Delta\rho = \begin{bmatrix} \Delta\rho_1 \\ \Delta\rho_2 \\ \vdots \\ \Delta\rho_n \end{bmatrix} \quad \Delta x = \begin{bmatrix} \Delta x_u \\ \Delta y_u \\ \Delta z_u \\ -c\Delta t_u \end{bmatrix} \\ H = \begin{bmatrix} a_{x1} & a_{y1} & a_{z1} & 1 \\ a_{x2} & a_{y2} & a_{z2} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ a_{xn} & a_{yn} & a_{zn} & 1 \end{bmatrix} \end{array} \right. \quad (4)$$

由式(3),采用最小二乘法,得到用户位置和时间偏差为

$$\Delta x = (H^T H)^{-1} H^T \Delta\rho \quad (5)$$

式(5)中,若 Δx 存在扰动,则可得 Δx 的条件数为

$$\text{cov}(\delta\Delta x) = [(H^T H)^{-1} H^T] \text{cov}(\delta\Delta\rho) [(H^T H)^{-1} H^T]^T \quad (6)$$

式中: $\delta\Delta x$ 为用户接收机的定位和定时误差; $\delta\Delta\rho$ 为用户接收机伪距测距标准差。

若各伪卫星的测距误差统计独立,均值为 0,方差为 δ^2 ,则由式(6)有

$$\text{cov}(\delta\Delta x) = (H^T H)^{-1} \delta^2 \quad (7)$$

由式(7)可见, $(H^T H)^{-1}$ 体现了用户接收机的定位定时误差与伪距测量误差的关系,称为精度衰减矩阵。GDOP 定义为 $(H^T H)^{-1}$ 迹运算的平方根:

$$\text{GDOP} \triangleq \sqrt{\text{tr}(H^T H)^{-1}} \quad (8)$$

通常而言,将伪卫星所覆盖的服务区域内进行均匀采样,将采样点的 GDOP 求和后进行平均,就可得到伪卫星网络对服务区域的定位增强性能优劣程度:

$$\bar{G} = \sum_{i=1}^n G_i / m \quad (9)$$

式中： \bar{G} 为增强区域的GDOP均值； G_i 为采样点*i*处的GDOP值； m 为覆盖区域的有效采样点数。所谓有效指的是覆盖区域某个采样点到各伪卫星的方向余弦满足 $\mathbf{H}^T \mathbf{H}$ 可逆的条件，此时该采样点对应的GDOP可被求解。

2 空基伪卫星航路规划算法设计

空基伪卫星航路规划指的是在规划区域内，在给定的约束条件下寻找一条从起始点到目标点的最优或可行的飞行轨迹。根据空基伪卫星的业务性质，动态组网过程中的航路代价可分为伪卫星距离代价和覆盖区域GDOP代价2个方面。出于导航定位的需要，伪卫星数量通常在4个以上。这样，空基伪卫星的航路规划是一个编队协同航路规划问题。

由GDOP的定义可知，GDOP代价始终随着各伪卫星的运动而变化。这种航路代价与运动状态的耦合关系，使空基伪卫星的航路规划问题存在不同于一般飞行器航路规划的特殊性：在对一般飞行器进行协同规划时，规划区域中的威胁源/障碍往往不随飞行器的运动而发生变化，这就可以先对每个飞行器单独进行航路规划，然后再考虑时间上的协同；然而，对空基伪卫星而言，GDOP代价是由多颗伪卫星的运动状态所共同决定的，因此在协同航路规划中，无法为每个伪卫星单独规划相应的航路，而必须将多颗伪卫星作为一个整体来进行规划。

2.1 航路规划代价函数设计

A^* 算法是一种标准启发式搜索算法，在路径规划和图搜索领域中应用非常广泛。在利用 A^* 算法进行航路搜索时，通常将规划环境表示为网格的形式，通过预先设定的代价函数寻找最小代价航路。算法对当前位置的每一个可能到达网格单元计算代价函数，然后选择代价函数最低的网格单元加入搜索空间。加入搜索空间的这一新网格单元又被用来进行下一次的搜索，产生更多的可能路径^[13]。

A^* 算法的代价函数定义为

$$f(x) = g(x) + h(x) \quad (10)$$

式中： $g(x)$ 为实际代价函数； $h(x)$ 为启发代价函数。

本文定义 $g(t)$ 为从起始时刻运动到当前时刻*t*的实际代价。设空基伪卫星数量为*n*，则本文将 $g(t)$ 设计为

$$g(t) = \omega_1 \left(\sum_{j=1}^n S_j(t) \right) + \omega_2 G(t) \quad (11)$$

式中： $S_j(x)$ 为第*j*颗伪卫星从起始时刻到当前时刻*t*的运动距离； $G(t)$ 为在当前时刻*t*条件下，增强服务区域各采样点对应的GDOP值的总和； ω_1 和 ω_2 分别为伪卫星距离代价和GDOP代价的权重，调节 ω_1 、 ω_2 的大小可决定伪卫星是选择优先保障航路上GDOP最优而增加距离代价，还是减少距离代价而牺牲GDOP性能。

定义 $h(t)$ 为从当前时刻*t*运动到最终目标位置的估计代价， $h(t)$ 可设计为

$$h(t) = \omega_1 \left(\sum_{j=1}^n D_j(t) \right) + \omega_2 G_{\text{best}} \quad (12)$$

式中： $D_j(t)$ 为第*j*颗伪卫星从当前时刻*t*所在位置与目标部署位置的欧氏距离； G_{best} 为伪卫星全部位于目标部署位置时对应的增强区域各采样点的GDOP总和，也是GDOP代价的最优值；权重 ω_1 、 ω_2 取值与式(11)相同。

文献[13]的相关研究已经证明，只要启发式函数满足一致性($h(t)$ 组成结构与 $g(t)$ 相同)和可接纳性($h(t)$ 小于或等于从当前时刻*t*所处位置到目标位置的真实代价)，并且搜索空间中存在可行解，那么 A^* 算法就一定能找到其中的最优解。由式(11)可知， $h(t)$ 满足一致性和可接纳性。因此，采用式(11)和式(12)的 A^* 算法可以找到全局最优解。

2.2 代价函数权重自适应调整

空基伪卫星在从初始位置向目标位置的组网部署过程中，距离代价和GDOP代价始终随伪卫星编队运动状态的改变而不断变化。若伪卫星编队初始位置对应的GDOP代价远高于最终目标位置对应的最优GDOP，则伪卫星在运动的初期就有必要为提高用户定位精度而增加GDOP代价的权重，对应的航路规划就需要尽量使伪卫星编队保持良好的几何构型；而随着伪卫星逐渐接近目标位置，GDOP代价也逐渐收敛于最优值，当GDOP代价与最优值的偏差小到一定程度时，则可以考虑增大距离代价的权重，使伪卫星能够沿着较短的路径尽快到达目标位置，以减少燃料和时间的消耗。

由此可见，为使伪卫星在动态组网时能够根据当前的运动状态调节代价函数 $f(t)$ ，需要在航路规划时对距离代价和GDOP代价的权重进行自适应调整。针对式(11)中GDOP代价的权重 ω_2 ，本文将 ω_2 设计为关于当前时刻*t*的函数：

$$\omega_2 = \frac{G(t) - G_{\text{best}}}{G(t)} \quad (13)$$

相应地,距离代价权重 ω_1 表示为

$$\omega_1 = 1 - \omega_2 = G_{\text{best}} / G(t) \quad (14)$$

由式(13)和式(14)可见,在当前时刻 t ,若伪卫星编队的 GDOP 代价 $G(t)$ 接近最优值 G_{best} ,那么在下一时刻的航路规划中就相应地减小 GDOP 代价权重 ω_2 ,而增加距离代价的权重 ω_1 ,使航路尽量接近飞艇当前位置与目标位置的最短距离,这样在飞艇匀速运动的情况下,将会缩短飞艇到达目标所需时间;反之,若伪卫星编队几何结构在当前时刻 t 比较恶劣,则 GDOP 代价权重 ω_2 也相应较高,下一时刻航路规划将朝着降低 GDOP 代价的方向进行。

2.3 航路规划约束条件

采用 A* 算法进行航路搜索,在扩展节点时需要搜索当前节点邻域中的所有网格单元。这样,算法收敛到最优解的时间会随着规划区域的增大而呈指数增长,并且生成的航路不一定满足伪卫星运动的物理约束条件。因此,结合第 2.1 节中航路规划代价函数设计,本文考虑基于稀疏 A* 算法来进行空基伪卫星的三维航路搜索。稀疏 A* 算法与普通 A* 算法的不同之处在于:在扩展节点时,稀疏 A* 算法通过把约束条件结合到搜索算法中去,减少节点需要搜索的邻域网格单元,有效缩短搜索时间。

在 Szczerba 等^[14]提出的稀疏 A* 算法中,主要考虑的约束条件为飞行器运动时的几何约束,包括最小航迹段步长 L_{min} 、最大水平拐弯角 ϕ_{max} 等。由于稀疏 A* 算法是在离散的网格单元中搜索,规划出的路径由一系列的折线段所组成,并不满足飞行器的运动特性。因此,为规划出可飞行的航路,还需考虑飞行器动力学方面的约束。

飞行器的动力学特性主要包括在机体表面产生力和力矩的空气动力,这些力和力矩产生了侧向、纵向和滚转加速度^[15]。为生成可飞行航路,在规划航路时需要考虑飞艇转弯速率的约束,在三维空间中可转化为航路中的曲率和挠率约束。曲率和挠率分别与加速度满足如下关系:

$$\begin{cases} |a_x| = |\mathbf{v}|^2 \kappa \\ |a_z| = |\mathbf{v}|^2 \tau \end{cases} \quad (15)$$

式中: \mathbf{v} 为飞行器的速度; a_x 为侧向加速度; a_z 为俯仰加速度; κ 为曲率; τ 为挠率。

由式(15)可知,曲率(挠率)与加速度成正比,因此规划路径上任意一点的曲率(挠率)应小于飞行器可达到的最大曲率(挠率)。

Dubins 曲线由一系列圆弧及其切线组成,已

被证明是在满足一定曲率条件下,连接给定方向向量的两点间的最短轨迹曲线^[15]。这样,在给定飞艇最大曲率(挠率)的情况下,可以采用 Dubins 曲线对由传统稀疏 A* 算法规划出的路径进行平滑处理,得到可飞行路径。

下面以二维平面情况为例说明 Dubins 曲线的曲率与稀疏 A* 算法中最大水平拐弯角的约束关系。如图 1 所示,A、B 和 C 为稀疏 A* 算法规划出的航路上的 3 个位置点,航路在 B 点转弯; r 为曲率半径; ϕ 为航路水平拐弯角; L 为航路步长。由曲率与曲率半径的关系 $\kappa = 1/r$ 可知, κ 越大,则 r 越小;当 κ 达到最大值 κ_{max} , r 对应为最小值 r_{min} 。此外,由图 1 不难得到以下关系:

$$\tan \frac{\phi}{2} = \frac{L}{2r} \quad (16)$$

式(16)表明,当 L 一定时, r 越小,则 ϕ 越大;当 $L = L_{\text{max}}$, $r = r_{\text{min}}$ 时, ϕ 达到最大值 ϕ_{max} , ϕ_{max} 即为稀疏 A* 算法中的最大水平拐弯角。此外,为能够对航路中可能出现连续的折线段进行平滑,取线段 AB 和 BC 的长度均为 $L/2$ 。

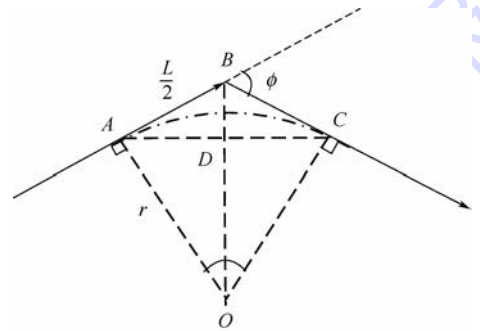


图 1 Dubins 曲线平滑处理原理

Fig.1 Principle of smoothing with Dubins curve

因此,在二维情况下,若在航路规划时能给出约束条件最大航路步长 L_{max} 和飞艇最大曲率 κ_{max} ,就可以计算出航路最大水平拐弯角 ϕ_{max} ,从而通过稀疏 A* 算法进行航路规划,规划完成后采用 Dubins 曲线对航路进行平滑处理。推广到三维空间,约束条件除 L_{max} 和 κ_{max} 外,若再给定飞艇最大挠率 τ_{max} ,可计算出最大水平拐弯角 ϕ_{max} 和最大俯仰角 θ_{max} ,再结合高程约束,就可以对三维空间中的航路进行规划和平滑。

航路规划约束也包括航迹段长度的约束。由式(16)可知,在曲率一定的条件下,如果航迹段步长 L 值较大,则水平拐弯角 ϕ 也将增大,这样稀疏 A* 算法在每次扩展节点时就要搜索更多的邻域网格,增加了时间和空间代价;结合图 1 还可见,B 点是从 A 点扩展而来的航路代价最小点,而

飞艇实际却是沿 AC 运动,若 L 过大,则 B 与 AC 的距离也更长,航路代价也相差更大。当航路代价相差到一定程度时, B 点实际上就不能作为 A 点在约束条件下扩展出来的航路代价最小点,成为无效点。若 L 取得过小,在规划空间较大时,不仅航路规划时间将会指数增长,而且飞艇将会增加更多的转弯机动。然而,以飞艇为平台的空基伪卫星作为导航源,只有在本身位置能够精确测量的前提下才能为用户提供导航增强服务,如果飞艇机动较多,将不利于飞艇本身位置的确定,从而影响用户定位。因此,在确定航迹段步长 L 时,应结合曲率等约束条件,在可接受的上限以内合理选择其大小。

对飞艇编队进行航路规划时,除了单个飞艇的动力学约束外,还需考虑各飞艇之间相对状态的约束,主要包括碰撞条件约束和能量消耗均衡性约束。对于碰撞条件约束,可通过设置飞艇间距阈值进行判断,在航路规划过程中,若飞艇间距小于阈值,就认为存在碰撞危险;对于能量消耗均衡性约束,可在航路规划之前预先进行任务分配,避免个别飞艇距离目标位置过长或过短。具体方法是:以各伪卫星到目标部署位置的欧式距离均方差最小为准则,遍历初始位置与目标位置的所有排列组合情况,选出最优组合。

2.4 基于稀疏 A^* 算法的航路规划算法

稀疏 A^* 算法是在隐式图中进行搜索,图中的节点在图搜索过程中可分为3类:已被扩展的节点;当前已经产生但尚未扩展的节点;尚未产生的节点。其中,第1类节点称为封闭节点,在搜索过程中存放在CLOSED表中;第2类节点是当前已经产生、正在等待扩展的节点,称为开放节点,在搜索过程中存放在OPEN表中^[16]。

基于权重自适应调整的改进稀疏 A^* 算法的空基伪卫星航路规划算法流程如下:

步骤1 对各个飞艇进行任务分配,确定各飞艇所对应的目标部署位置。同时,根据输入曲率和挠率的最大值约束,计算曲率半径和挠率半径。设置航迹段步长 L ,由式(15)计算最大水平拐弯角 ϕ_{\max} 和最大俯仰角 θ_{\max} 。

步骤2 将各伪卫星初始位置点插入OPEN表中,将CLOSED表置空。

步骤3 如果OPEN表为空,算法以搜索失败结束。调整算法参数(相邻网格单元),然后重新运行规划算法。

步骤4 按式(10)计算节点代价 $f(t)$,从OPEN表中移出代价最小的元素作为各伪卫星的

开放节点,将它放入CLOSED表中,同时更新GDOP代价和距离代价的权重值。

步骤5 如果伪卫星当前节点与目标点之间的距离小于航迹段步长,则将目标点的父节点指针指向当前节点,航路搜索过程结束。从目标点开始向上回溯至起始位置,得到从起始点到目标的最小代价路径。

步骤6 扩展各伪卫星的开放节点。扩展过程如下:

1) 对每颗伪卫星构造当前节点待扩展区。待扩展区的水平剖面是夹角大小为 ϕ_{\max} 的2倍的扇面,扇面的半径在数值上等于航迹段步长 L 。并以进入当前节点速度在水平面上投影的方向为对称轴。垂直剖面的夹角大小为 θ_{\max} 的2倍,关于水平方向对称。

2) 分割各颗伪卫星当前位置点的待扩展区。如图2所示,首先在垂直剖面内均匀取 M 个方向,得到 M 个2倍于 ϕ_{\max} 大小的扇面;然后把每一个扇面均匀分为 N 个扇区,这样可得到 $M \times N$ 个小扇区,共形成 $M \times (N+1)$ 个扩展节点。判断每个扩展节点位置坐标是否满足飞行高度约束,若不满足,则将该节点位置高度限制为高程约束的边界值。

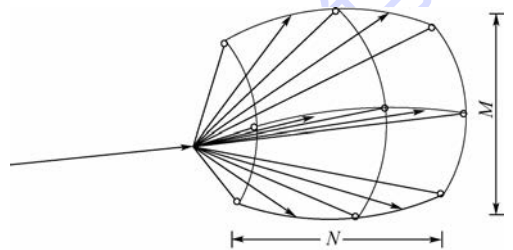


图2 开放节点扩展示意图

Fig.2 Schematic diagram of expanding open nodes

3) 将 n 个伪卫星对应的待扩展区组合成一个搜索空间,各个伪卫星分别遍历各自对应的待扩展区,经排列组合后共形成 $[M \times (N+1)]^n$ 种组网方式。

4) 设置伪卫星之间碰撞危险的距离阈值 d_{\min} ,遍历检测搜索空间,检测各个伪卫星的间距。若在某种组网方式中,存在2个及以上的伪卫星的间距距离小于距离阈值,则将该组网方式从搜索空间中剔除。若剔除以后搜索空间为空集,则OPEN表置空,返回步骤3。

5) 将第3)步计算得到的每种组网情况对应的各伪卫星位置坐标集合,插入到OPEN表中。

6) 返回步骤3。

步骤7 对规划出的路径采用Dubins曲线进

行平滑。结合图1,平滑方法如下:

1) 沿航路前行方向,每隔 $L/2$ 的距离选取一个航路上的位置点(大地坐标系),组成位置点序列,设第一个位置点为当前位置点。

2) 从当前位置点开始,依次往后连续取3个点,分别设为 $A(x_A, y_A, z_A)$ 、 $B(x_B, y_B, z_B)$ 和 $C(x_C, y_C, z_C)$ 。计算 \overrightarrow{AB} 和 \overrightarrow{BC} 夹角(拐弯角)的余弦值,若为0则转到第7)步;否则执行第3)步。

3) 计算水平拐弯角 ϕ ,若 ϕ 为钝角,则 ϕ 的值用 $\pi - \phi$ 替代。根据式(16)计算曲率半径 r ;同时,根据 $\overrightarrow{OA} \cdot \overrightarrow{AB} = 0, \overrightarrow{OB} \cdot \overrightarrow{AC} = 0, \overrightarrow{OC} \cdot \overrightarrow{BC} = 0$,计算出曲率圆的圆心 O 在大地坐标系下的坐标 $O(x_o, y_o, z_o)$ 。

4) 设 D 为 AC 与 OB 的交点,根据图1中几何关系,计算得到如下向量长度:

$$\begin{cases} |\overrightarrow{OB}| = \frac{L}{2\sin(\phi/2)} \\ |\overrightarrow{AD}| = |\overrightarrow{DC}| = r\sin(\phi/2) \end{cases} \quad (17)$$

构造本地坐标系:以 O 为本地坐标系原点,以 \overrightarrow{OB} 方向为 y 轴, \overrightarrow{AC} 方向为 x 轴,令 $\overrightarrow{BE} = \overrightarrow{BC} \times \overrightarrow{AB}$,以 \overrightarrow{BE} 方向为 z 轴。根据式(17),可分别求得 E, A 和 C 在本地坐标系中的坐标,分别设为 \hat{p}_E 、 \hat{p}_A 和 \hat{p}_C ;同时,设 E, A 和 C 3 点在大地坐标系中的坐标分别为 p_E, p_A 和 p_C 。

令 $\hat{p}_E, \hat{p}_A, \hat{p}_C, p_E, p_A$ 和 p_C 均为 3×1 列向量, T 为本地坐标系到大地坐标系的转换矩阵(3×3 方阵),于是可得线性方程组:

$$\begin{cases} T\hat{p}_E = p_E \\ T\hat{p}_A = p_A \\ T\hat{p}_C = p_C \end{cases} \quad (18)$$

通过求解式(18)的方程组,可求解出转换矩阵 T 中各个元素。由于 E, A 和 C 在本地坐标系中的坐标并不在同一个坐标轴平面上,因此,式(18)中的9个方程互不相关, T 中的9个未知数均可被求解。

5) 设置旋转角度步长 $\Delta\omega$,在本地坐标系中将 \overrightarrow{OA} 以 O 为圆心向 C 点旋转,每转过 $\Delta\omega$ 即在圆弧上采样一个位置点,计算该点在本地坐标系下的坐标 \hat{p}_k ;再通过计算 $T\hat{p}_k$ 得到采样点在大地坐标系下的坐标 p_k 。

6) 连接 A, C 及采样点,得到 Dubins 曲线 \widehat{AC} 。

7) 若 C 为位置点序列中的最后一个位置点,航路规划算法结束;否则,将 B 设为当前位置点,

返回第2)步。

图3显示了基于 Dubins 曲线平滑前后的航路对比。可见,平滑后的航路更加符合飞艇的实际运动轨迹。

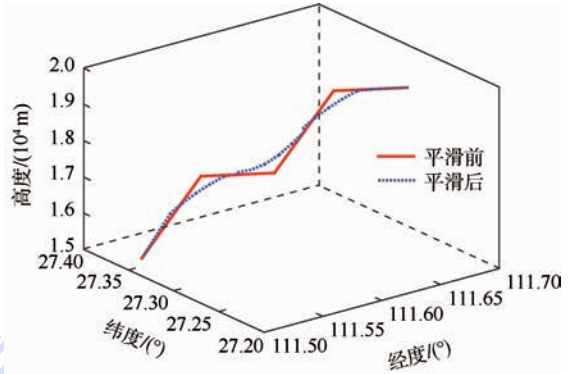


图3 Dubins 曲线平滑前后的航路对比

Fig. 3 Comparison of route with and without smoothing by Dubins curve

3 仿真结果分析

本文仿真实验采用临近空间飞艇作为空基伪卫星平台。由于飞艇的数量、初始位置和初始速度方向等因素由飞艇放飞地点、飞艇前期任务规划等诸多因素决定,存在很大的随机性。因此,本文仿真实验对于飞艇数量、初始位置和初始速度方向参数不作特别要求。

仿真实验的相关参数见表1。飞艇的初始速度方向假设为从正南至正北方向,速度大小恒定。仿真平台采用 MATLAB R2015a。

表1 仿真实验相关参数

Table 1 Related parameters of simulated experiments

参数	数值
飞艇数量	5
飞艇编队初始位置 (经度/(°), 纬度/(°), 高度/m)	1号艇:(122.0, 24.5, 10000) 2号艇:(122.5, 25.0, 10000) 3号艇:(123.0, 25.5, 10000) 4号艇:(123.5, 25.0, 10000) 5号艇:(124.0, 24.5, 10000)
飞艇高程范围/m	[2000, 30000]
航迹段长度(步长)/m	5000
最小曲率半径/m	4330
最小挠率半径/m	4330
最小间距阈值/m	5000
服务区域经度范围/(°)	[112.0, 114.0]
服务区域纬度范围/(°)	[25.5, 27.5]
垂直方向搜索扇面数	3
水平方向搜索扇面数	4

3.1 离线航路规划仿真

离线航路规划指在飞艇编队组网部署之前规划出各飞艇的预定航路。本文分别在4种不同权

重条件下对飞艇航路规划进行仿真,这4组不同权重条件分别如下:

- 1) 距离代价权重与 GDOP 代价权重按式(13)和式(14)进行自适应调节。
- 2) 距离代价权重优先, GDOP 代价权重为 0。
- 3) GDOP 代价权重优先, 距离代价权重为 0。
- 4) 距离代价与 GDOP 代价权重为固定值 1/2。

在以上4组不同权重条件下,飞艇航路规划仿真结果如图4所示。由于飞艇具有一定的悬停能力,因此对于先到达目标部署位置的飞艇,本文假定其到达目标部署位置后即悬停不动,相应的路径不再继续规划。飞艇编队的目标部署位置在航路规划前采用粒子群搜索算法^[17-18]求出。

由图4可见,对于权重自适应可调、距离代价权重优先以及权重固定3种情况,航路规划算法均可以收敛,飞艇最终运动到目标部署位置;只有在 GDOP 代价权重优先的情况下,航路规划算法不能使飞艇收敛到目标部署位置。

基于上述4种权重条件下的航路规划结果,可分析各飞艇在沿相应航路运动时,飞艇与目标部署位置的距离以及导航增强服务区域的 GDOP 变化情况。各飞艇与目标部署位置的平均距离变化如图5所示。

由图5可见,对于 GDOP 代价权重优先的情况,由于算法中没有目标部署位置的信息,导致航路规划过程中飞艇与目标部署位置的偏离越来越大,而对于其他3种情况,算法在每一次搜索中都包含了目标部署位置信息,因而飞艇总是会最终收敛于目标部署位置。此外,在权重自适应可调情况下进行航路规划时,飞艇距离代价与距离权重优先条件下的距离代价比较接近。这表明采用基于权重自适应可调情况下的航路规划算法,不会明显地增加飞艇组网部署的时间开销。

在上述4种权重条件下,航路规划中 GDOP 代价权重变化情况如图6所示。可见,对于距离代价权重优先的情况,由于算法在搜索过程中没有考虑 GDOP 因素,因此航路规划过程中导航增强服务区域的 GDOP 均值没有像其他3种权重条件下那样快速下降。而对于权重自适应可调和固定权重2种情况,前者对应的航路规划中的 GDOP 均值下降更快,更加接近于 GDOP 代价权重优先条件下的 GDOP 均值。

结合图5和图6可知,与其他相比,采用权重自适应可调的航路规划算法,能够在不明显增加距离代价的情况下使飞艇编队保持良好的几何构

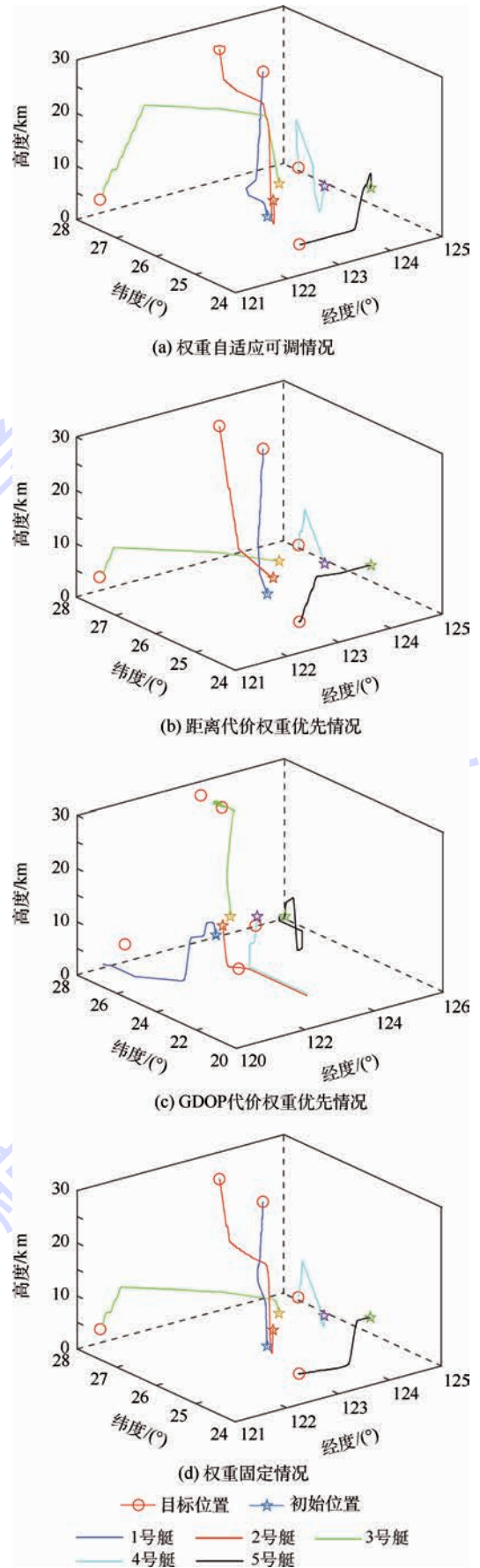


图4 不同权重条件下的三维航路规划结果
Fig. 4 3D route planning of airships in different conditions of weight

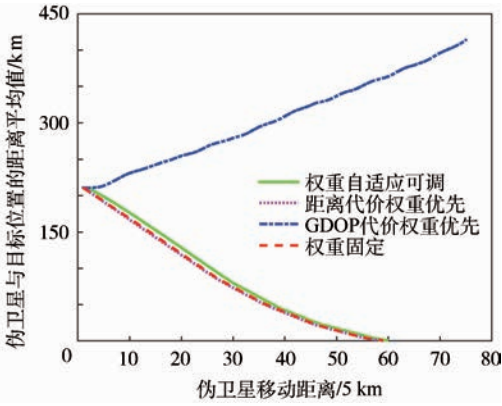


图5 不同权重条件下航路规划中各飞艇与目标部署位置平均距离变化曲线

Fig. 5 Curves of average distance from airships' position to deployed target position in different conditions of weight during route planning

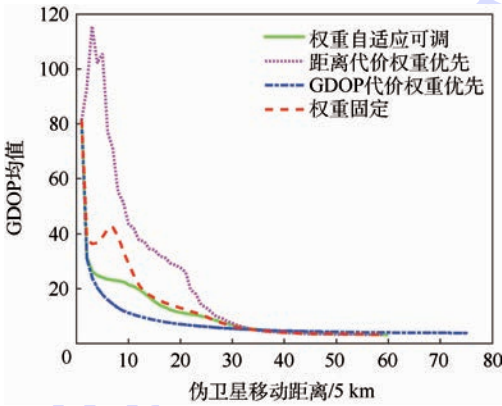


图6 不同代价权重条件下航路规划中的GDOP代价变化曲线

Fig. 6 Curves of GDOP cost in different conditions of weight during route planning

型,有效降低 GDOP 代价,体现了在距离代价和 GDOP 代价之间的调节能力。

3.2 在线航路重规划仿真

离线航路规划为每个飞艇确定了预定的飞行路径。然而,飞艇在实际飞行时,难免会受到风场等外界因素的影响而偏离预定航路。在这种情况下,需要对飞艇编队的航路进行在线重规划。本文考虑在单个飞艇偏离预定航路情况下,飞艇编队在线航路重规划的2种方法:基于单艇调整的在线航路重规划和基于编队调整的在线航路重规划。

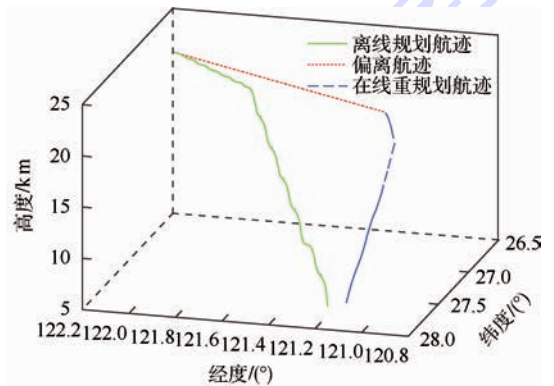
基于单艇调整的在线航路重规划仅对偏离预定航路的飞艇进行航路重规划,其他飞艇则继续按原定航路运动。这样,航路重规划在搜索时,搜索空间仅为单个飞艇的待扩展区,与飞艇编队整

体的航路规划相比,搜索空间大大减小,具有较好的实时性。

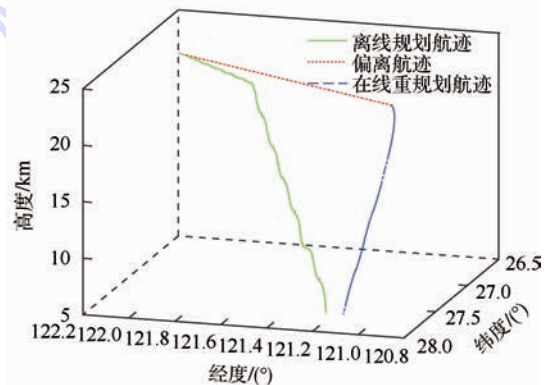
由于单个飞艇偏离原定航路破坏了飞艇编队的整体几何构型,因此基于编队调整的在线航路重规划将从修正编队几何构型的角度考虑,采用的方法与离线航路规划方法相同。由于基于编队调整的航路重规划重新优化了编队的几何结构,因而与基于单艇调整的航路重规划而言,具有更低的 GDOP 代价。

在本节仿真中,假设图4(a)仿真中的3号艇(绿色)在沿预定航路运动到经度为 122.1252° 、纬度为 26.7657° 、高度为 22500 m 的位置后开始发生偏离,继续运动至经度为 120.9994° 、纬度为 27.4151° 、高度为 22500 m 的位置后,飞艇检测到自身已偏离预定航路,并开始在线航路重规划。3号艇的在线航路重规划结果如图7所示。图7(a)为基于3号艇自身调整的在线航路重规划结果,图7(b)所示为基于编队调整的在线航路重规划结果。可见,2种方法规划出的航路都可收敛于原定航路。

图8显示了3号艇在线航路规划前后,增强服务区域 GDOP 均值的变化情况。可见,虽然基



(a) 基于单艇调整的在线航路重规划



(b) 基于编队调整的在线航路重规划

图7 3号艇在线航路重规划的仿真

Fig. 7 Simulation of online route re-planning of No. 3 airship

于编队调整的在线航路重规划与基于单艇调整的在线航路重规划相比, GDOP 代价有所降低, 但降幅仅约为 0.006, 相应的航路代价却增加了一个步长(5 km)。这表明在线航路重规划若对整个编队的几何构型进行调整, 不仅算法搜索时间增多, 而且飞艇沿规划航路运动的时间也 longer。考虑到在线航路重规划对实时性的要求, 实际中可能更适合采用基于单艇调整的在线航路重规划方法。

编队的几何构型, 从而对 GDOP 代价产生扰动。后续工作将进一步重点研究变速情况下的在线航路规划。

参考文献 (References)

- [1] DAI L W, WANG J L, TSUJII T, et al. Pseudolite applications in positioning and navigation: Modeling and geometric analysis [C] // International Symposium on Kinematic System in Geodesy, Geomatics & Navigation. Berlin: Springer, 2001: 482-489.
- [2] OKTAY H, STEPANIAK M. Airborne pseudolite in a global positioning system degraded environment [C] // 2011 5th International Conference on Recent Advances in Space Technologies (RAST). Piscataway, NJ: IEEE Press, 2011: 280-285.
- [3] TIWARY K, BEHERA S, SHARADA G, et al. Modelling and simulation of pseudolite-based navigation: A GPS-independent radio navigation system [J]. Defence Science Journal, 2010, 60 (5): 541-550.
- [4] CHANDU B, PANT R, MOUDGALYA K. Modeling and simulation of a precision navigation system using pseudolites mounted on airships: AIAA-2007-7746 [R]. Reston: AIAA, 2007.
- [5] WANG J. Pseudolite applications in positioning and navigation: Progress and problems [J]. Journal of Global Positioning Systems, 2002, 1(1): 48-56.
- [6] TSUJII T. Experiments of inverted pseudolite positioning for airship-based GPS augmentation system [C] // Proceedings of the 15th International Technical Meeting of the Satellite Division of the U. S. Institute of Navigation. Manassas: Institute of Navigation, 2002: 1689-1695.
- [7] MA W H, YUAN J P, LUO J J. Airborne pseudolite aiding BeiDou system to improve positioning precision in low latitude areas [C] // International Conference on Space Information Technology. Bellingham: SPIE Press, 2006, 5985: 955-959.
- [8] 杨一, 高社生, 阎海峰. 临近空间伪卫星几何布局方案设计 [J]. 系统工程与电子技术, 2014, 36(3): 532-538.
YANG Y, GAO S S, YAN H F. Design on geometric configuration schemes of pseudolite in near space [J]. Systems Engineering & Electronics, 2014, 36(3): 532-538 (in Chinese).
- [9] 呼玮, 杨建军, 何苹. 基于临近空间飞艇定位的伪卫星布局研究 [J]. 无线电工程, 2009, 39(10): 24-27.
HU W, YANG J J, HE P. Study on pseudolite configuration scheme based on near space airships [J]. Radio Engineering, 2009, 39(10): 24-27 (in Chinese).
- [10] 孟键, 孙付平, 丛佃伟. 伪卫星增强区域卫星导航系统组网仿真 [J]. 测绘科学技术学报, 2008, 25(3): 213-215.
MENG J, SUN F P, CONG D W. Plans of augmenting regional satellite navigation using PLs [J]. Journal of Geomatics Science and Technology, 2008, 25(3): 213-215 (in Chinese).
- [11] 杨龙, 周建华, 陈金平. 基于遗传算法的空基伪卫星导航信号增强平台组网优化方法研究 [J]. 全球定位系统, 2008, 33(3): 9-13.
YANG L, ZHOU J H, CHEN J P. The study of optimization of formation flying navigation augmentation platforms based on genetic algorithm [J]. GNSS World of China, 2008, 33(3): 9-13 (in Chinese).

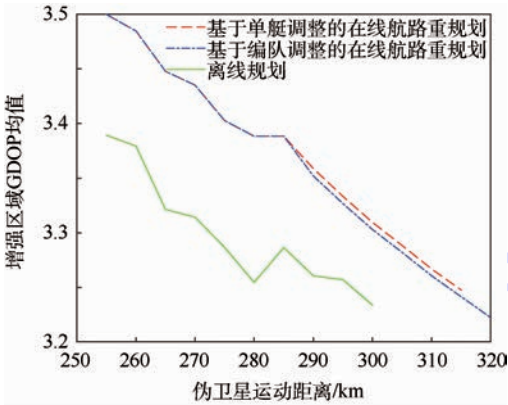


图8 在线航路重规划的 GDOP 代价对比
Fig. 8 Comparison of GDOP cost of online route re-planning

4 结 论

本文在相关研究的基础上, 提出了空基伪卫星从初始位置到目标位置如何进行组网部署的问题。在稀疏 A* 算法的基础上, 结合相关约束, 提出了距离代价和 GDOP 代价权重自适应可调节的空基伪卫星航路规划算法。通过仿真实验, 得出如下结论:

1) 采用权重自适应可调的航路规划算法, 规划的航路一方面能够有效降低导航增强服务区域的 GDOP 值, 提高用户定位精度; 另一方面又能够使飞艇较快到达目标部署位置, 不会明显增加飞艇运动的时间和燃料消耗, 表明航路规划算法具有良好的经济性。

2) 当编队中的个别飞艇在运动过程中偏离预定航路时, 可采用基于单艇调整和基于编队调整 2 种在线航路重规划方法, 前者的算法搜索速度更快, 规划路径更短, 而 GDOP 代价与后者相比相差并不大, 因此更适合在线实时的航路重规划。

本文的仿真实验中, 均假定飞艇在运动过程中保持匀速。实际上, 飞艇在运动过程中受风场、空气密度等环境因素的影响, 很难始终保持匀速运动。飞艇运动过程中的速度变化将会实时改变

- [12] 党进伟,高社生,张极. 伪卫星协同编队控制与仿真[J]. 测控技术,2013,32(11):76-79.
DANG J W, GAO S S, ZHANG J. Control and simulation of collaborative formation for pseudolites[J]. Measurement & Control Technology, 2013, 32(11):76-79 (in Chinese).
- [13] HART P E, NILSSON N J, RAPHAEL B. A formal basis for the heuristic determination of minimum cost routes [J]. IEEE Transactions on Systems Science and Cybernetics, 1968, 4(2): 100-107.
- [14] SZCZERBA R J, GALKOWSKI P, GLICKSTEIN I S, et al. Robust algorithm for real-time route planning [J]. IEEE Transactions on Aerospace and Electronic Systems, 2000, 36(3): 869-878.
- [15] TSOURDOS A, WHITE B, SHANMUGAVEL M. 无人机协同路径规划[M]. 祝小平,周洲,王恽,译. 北京:国防工业出版社,2013:27-56.
TSOURDOS A, WHITE B, SHANMUGAVEL M. Cooperative path planning of unmanned aerial vehicles [M]. ZHU X P, ZHOU Z, WANG Y, translated. Beijing: National Defense Industry Press, 2013:27-56 (in Chinese).
- [16] 丁明跃,郑昌文,周成平,等. 无人飞行器航迹规划[M]. 北京:电子工业出版社,2009:46-57.
DING M Y, ZHENG C W, ZHOU C P, et al. Route planning for unmanned aerial vehicles [M]. Beijing: Publishing House of Electronics Industry, 2009:46-57 (in Chinese).
- [17] 朱志良,叶宁,刘军,等. 基于临近空间飞行器的区域自组网优化部署算法[J]. 电子与信息学报, 2011, 33(4): 915-921.
ZHU Z L, YE N, LIU J, et al. Deployment optimization algorithm for regional MANET containing near space vehicles as a part [J]. Journal of Electronics & Information Technology, 2011, 33(4):915-921 (in Chinese).
- [18] LI D P, DENG P K, LIU B, et al. Research on the dynamic configuration of air-based pseudolite network [C] // China Satellite Navigation Conference (CSNC) 2015 Proceedings: Volume II. Berlin: Springer, 2015, 341:357-367.

作者简介:

曾凌川 男, 硕士, 助理工程师。主要研究方向: 导航增强技术。

Tel.: 010-82178014

E-mail: zenglingchuan@aoe.ac.cn

巩应奎 男, 博士, 副研究员。主要研究方向: 卫星导航仿真技术、导航增强技术。

Tel.: 010-82178658

E-mail: ykgong@aoe.ac.cn

Route planning algorithm for configuring airborne pseudolites

ZENG Lingchuan, LI Dapeng, QU Yi, REN Aiai, GONG Yingkui*

(Academy of Opto-Electronics, Chinese Academy of Sciences, Beijing 100094, China)

Abstract: This paper focuses on the airborne pseudolites based on airships and proposes a route planning method to configure airborne pseudolites from the initial position to the target position. First, by analyzing the geometric dilution of precision (GDOP) as the performance of navigation enhancement, a route planning cost function was designed based on the distance and GDOP. Then, on the basis of sparse A* algorithm and analysis of constraint conditions, a route planning algorithm was proposed which can adjust the weights of distance and GDOP adaptively, and the route was smoothed by the Dubins curve. Finally, the route planning algorithm was validated through the simulated experiments. The simulated experimental results show that the route planning algorithm can reduce the GDOP cost effectively and improve user positional accuracy in the subsequent route planning if the GDOP value is severe at current time; on the other hand, while the GDOP approaches its optimal value, the algorithm will increase the weight of distance to make pseudolites move to their destinations as quickly as possible and reduce the time consumption during the subsequent route planning.

Key words: navigation enhancement; airborne pseudolites; route planning; sparse A* search algorithm; Dubins curve

Received: 2015-07-16; Accepted: 2015-10-23; Published online: 2015-12-17 10:41

URL: www.cnki.net/kcms/detail/11.2625.V.20151217.1041.005.html

Foundation item: National Natural Science Foundation of China (91438207)

* Corresponding author. Tel.: 010-82178658 E-mail: ykgong@aoe.ac.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0454

EHA 反馈线性化最优滑模面双模糊滑模控制

张振¹, 李海军^{1,*}, 诸德放²

(1. 海军航空工程学院 兵器科学与技术系, 烟台 264001; 2. 空军勤务学院 航空弹药系, 徐州 221000)



摘 要: 为提高电静液作动器(EHA)控制性能,提出了一种反馈线性化最优滑模面的双模糊滑模控制方法。该方法在EHA的非线性模型上,利用反馈线性化方法将其线性化,在此基础上建立线性切换函数,并采用最优控制理论对切换函数进行设计。为削弱抖振将模糊控制算法引入滑模控制中,采用一个模糊控制器,根据最优滑模函数运动特性的数值对切换控制增益进行估计;采用另一个模糊控制器,根据滑模控制原理对切换控制项进行调整。仿真结果证明了该方法的有效性。

关键词: 电静液作动器(EHA); 反馈线性化; 滑模控制; 最优控制; 模糊控制

中图分类号: TH137

文献标识码: A

文章编号: 1001-5965(2016)07-1398-08

多/全电飞机是现代飞机的一个新兴发展方向,其作动系统将采取新型功率电传作动器,而电静液作动器(Electro-Hydrostatic Actuator, EHA)是其实现的一种重要形式^[1-4]。与传统作动器相比, EHA 具有占用空间小、效率高等优点,但同时其是一种强非线性的电液位置伺服系统,存在参数时变性及外界干扰的不确定性,这些都严重影响系统的动态性能。

为提高 EHA 动态性能,许多学者基于非线性控制理论来设计其控制器,如滑模变结构控制^[5-7]、自适应控制^[8-10]和模糊控制^[11]等,并取得了一定的成果。在这些非线性控制策略中,滑模变结构控制由于设计简单且对系统参数摄动和外界扰动强鲁棒性而深受重视^[12]。对非线性系统而言,在设计滑模变结构控制器时,应取非线性的切换函数,但目前大多采用线性的切换函数,这将降低系统的控制性能,且滑模变结构控制由于其本质上的不连续特性将引起系统的抖振,这种抖振不仅增加了控制器的负担,而且有可能使系统跟踪误差在零点附近产生抖

动,从而增大稳态误差,有时甚至会造成系统的不稳定^[13]。反馈线性化是一种非线性控制设计方法,其基本思想是:通过适当的非线性状态变换和反馈变换,使非线性系统实现全部或部分的精确线性化,以方便应用熟知的线性控制方法进行设计,但其需要精确的模型信息,无法克服干扰,鲁棒性较差。

基于以上原因,本文以 EHA 为研究对象,将反馈线性化与滑模变结构控制相结合:一方面利用滑模变结构对外界干扰强鲁棒性特点,减弱反馈线性化对精确模型的依赖,以提高 EHA 控制系统的鲁棒性;另一方面将 EHA 的非线性模型反馈线性化后,可建立线性滑模切换函数,并应用线性理论来改善其动态性能;同时引入了模糊控制算法来削弱系统的抖振。最后,对 EHA 的跟踪控制进行了仿真,通过仿真验证了本文所提方法的有效性。

1 EHA 非线性建模

1.1 系统描述

EHA 系统组成原理如图 1 所示,该系统执

收稿日期: 2015-07-07; 录用日期: 2015-09-11; 网络出版时间: 2015-10-14 16:22

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151014.1622.011.html

* 通讯作者: Tel.: 13081617309 E-mail: li_haijun@sohu.com

引用格式: 张振, 李海军, 诸德放. EHA 反馈线性化最优滑模面双模糊滑模控制[J]. 北京航空航天大学学报, 2016, 42(7): 1398-1405. ZHANG Z, LI H J, ZHU D F. Double fuzzy sliding mode control for EHA based on feedback linearization optimal sliding surface[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1398-1405 (in Chinese).

行元件为对称双作用液压缸,液压泵为双向定量泵,动力源为永磁同步电机。EHA控制器通过比较指令位移与实际位移的大小,来改变电机的转速和转向,从而达到控制作动器位移的目的。

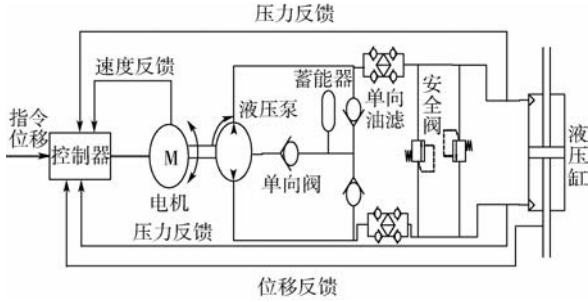


图1 EHA系统组成原理

Fig.1 Composition principle of EHA system

1.2 系统的非线性模型

1.2.1 永磁同步电机模型

永磁同步电机分梯形波和正弦波2种,后者比前者调速范围大、力矩脉动小,且考虑现有飞机主电源为交流电源,故本文中选取正弦波永磁同步电机作为动力源。为建模方便,假设铁心不饱和,且忽略气隙中高次谐波和磁损耗,电机采用 $i_d=0$ 矢量控制,其在 dq 坐标系有以下方程:

电机 q 轴定子电压方程为

$$u_q = L_q \frac{di_q}{dt} + Ri_q + n_p w_{rm} \psi_r \quad (1)$$

电机运动方程为

$$\frac{dw_{rm}}{dt} = [T_e - D_p(P_1 - P_2) - B_m w_{rm}] / J_m \quad (2)$$

电磁转矩方程为

$$T_e = \frac{3}{2} n_p \psi_r i_q \quad (3)$$

式中: u_q 、 L_q 和 i_q 分别为 q 轴的定子电压、电感和电流; R 为定子电阻; n_p 为电机极对数; w_{rm} 为电机转速; ψ_r 为永磁体产生的磁链; D_p 为定量泵排量; P_1 、 P_2 分别为液压泵高、低压腔压力; B_m 和 J_m 分别为电机阻尼系数和转动惯量; T_e 为电机电磁转矩。

1.2.2 泵控缸模型

根据文献[14],泵控缸流量连续方程为

$$D_p w_{rm} = A \frac{dx_l}{dt} + \frac{V_0}{2\beta_e} \left(\frac{dP_1}{dt} - \frac{dP_2}{dt} \right) + (\xi + L_{ext}/2)(P_1 - P_2) + 2\xi P_{pipe} \quad (4)$$

式中: A 为液压缸液压杆活塞面积; x_l 为液压缸液压杆的位移; V_0 为液压管路和液压缸平均体

积; β_e 为液压油等效容积弹性模数; ξ 、 L_{ext} 分别为液压泵内、外泄漏系数; P_{pipe} 为液压管路压降。

在式(4)中,液压管路压降 P_{pipe} 大小与流量大小有关,根据Darcy管流动方程,其与流量的关系近似为

$$P_{pipe} \approx K_{pipe} D_p^2 w_{rm}^2 \quad (5)$$

式中: K_{pipe} 为黏性系数。

液压缸力平衡方程为

$$(P_1 - P_2)A = M_l \frac{d^2 x_l}{dt^2} + F_f + F_L \quad (6)$$

式中: M_l 为液压缸液压杆及负载等效到其的总质量; F_f 为摩擦力; F_L 为作用在液压杆的外负载力。

在式(6)中,摩擦力 F_f 通常认为由静摩擦力、库仑摩擦力和黏性摩擦力组成^[15],可表示为

$$F_f = \left[a_1 \left(\frac{dx_l}{dt} \right)^2 + a_3 \right] \text{sgn} \left(\frac{dx_l}{dt} \right) + a_2 \frac{dx_l}{dt} \quad (7)$$

式中: a_1 、 a_2 和 a_3 为摩擦系数。

1.3 系统的非线性模型状态空间描述

取 q 轴定子电压 u_q 作为模型输入量。此时令 $u = u_q$, $x_1 = i_q$, $x_2 = w_{rm}$, $x_3 = P_1 - P_2$, $x_4 = x_l$, $x_5 = dx_l/dt$,通过式(1)~式(7),可得EHA非线性模型状态空间描述为

$$\begin{cases} \dot{x}_1 = -\frac{R}{L_q} x_1 - \frac{n_p \psi_r}{L_q} x_2 + \frac{u_q}{L_q} \\ \dot{x}_2 = \frac{3n_p \psi_r}{2J_m} x_1 - \frac{B_m}{J_m} x_2 - \frac{D_p}{J_m} x_3 \\ \dot{x}_3 = -\frac{4\beta_e \xi K_{pipe} D_p^2}{V_0} x_2^2 + \frac{2\beta_e D_p}{V_0} x_2 - \frac{2\beta_e}{V_0} \left(\xi + \frac{L_{ext}}{2} \right) x_3 - \frac{2\beta_e A}{V_0} x_5 \\ \dot{x}_4 = x_5 \\ \dot{x}_5 = \frac{A}{M_l} x_3 - \frac{a_1}{M_l} x_5^2 \text{sgn}(x_5) - \frac{a_2}{M_l} x_5 - \frac{a_3}{M_l} \text{sgn}(x_5) - \frac{F_L}{M_l} \end{cases} \quad (8)$$

由式(8)可以看出,EHA模型呈强非线性且特征参数多,计算量大,直接求解比较困难,需要对其进行线性化。

2 非线性模型反馈线性化

EHA非线性模型状态空间描述形式所对应的仿射非线性系统的标准形式为

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x})u + \mathbf{d} \\ \mathbf{y} = \mathbf{h}(\mathbf{x}) \end{cases} \quad (9)$$

式中: $f(\mathbf{x})$ 和 $g(\mathbf{x})$ 为充分光滑的向量场; \mathbf{d} 为系统所受干扰量; y 为系统输出量; $h(\mathbf{x})$ 为系统输出函数。

因此, EHA 非线性模型可改写为

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \end{pmatrix} = \begin{pmatrix} f(x_1, x_2) \\ f(x_1, x_2, x_3) \\ f(x_2, x_3, x_5) \\ f(x_5) \\ f(x_3, x_5) \end{pmatrix} + g(\mathbf{x})u + \mathbf{d} \quad (10)$$

式中:

$$f(x_1, x_2) = -\frac{R}{L_q}x_1 - \frac{n_p \psi_r}{L_q}x_2$$

$$f(x_1, x_2, x_3) = \frac{3n_p \psi_r}{2J_m}x_1 - \frac{B_m}{J_m}x_2 - \frac{D_p}{J_m}x_3$$

$$f(x_2, x_3, x_5) = -\frac{4\beta_e \xi K_{\text{pipe}} D_p^2}{V_0}x_2^2 + \frac{2\beta_e D_p}{V_0}x_2 -$$

$$\frac{2\beta_e}{V_0}(\xi + \frac{L_{\text{ext}}}{2})x_3 - \frac{2\beta_e A}{V_0}x_5$$

$$f(x_5) = x_5$$

$$f(x_3, x_5) = \frac{A}{M_t}x_3 - \frac{a_1}{M_t}x_5^2 \text{sgn}(x_5) - \frac{a_2}{M_t}x_5$$

$$g(\mathbf{x}) = [1/L_q \quad 0 \quad 0 \quad 0 \quad 0]^T$$

$$\mathbf{d} = [0 \quad 0 \quad 0 \quad 0 \quad -\frac{a_3}{M_t} \text{sgn}(x_5) - \frac{F_L}{M_t}]^T$$

$$h(\mathbf{x}) = x_4$$

2.1 反馈线性化条件

在对 EHA 非线性模型反馈线性化之前, 要验证其是否满足反馈线性化的条件。按照相关定义^[16-17]: 对原点附近的所有 \mathbf{x} , 矩阵 $[g \quad ad_f g \cdots ad_f^{n-1} g]$ 的秩等于系统维数 n 且集合 $\{g, ad_f g, \cdots, ad_f^{n-2} g\}$ 是非奇异对合分布 (f, g 为向量函数 $f(\mathbf{x}), g(\mathbf{x})$ 的缩写), 则其满足非线性模型状态反馈线性化条件。当矩阵的秩小于 n 时, 可以实现部分线性化, 但此时需保证零动态子系统的稳定性。

2.2 反馈线性化求解

设 EHA 系统相对阶为 r , 则根据式(11)计算其相对阶:

$$\begin{cases} L_g L_f^i h(\mathbf{x}) = 0 & 0 \leq i < r-1 \\ L_g L_f^{r-1} h(\mathbf{x}) \neq 0 \end{cases} \quad (11)$$

根据式(11)计算得到: $L_g h(\mathbf{x}) = 0, L_g L_f h(\mathbf{x}) = 0, L_g L_f^2 h(\mathbf{x}) = 0, L_g L_f^3 h(\mathbf{x}) = 0, L_g L_f^4 h(\mathbf{x}) = \frac{3A\beta_e D_p n_p \psi_r}{V_0 M_t J_m L_q} \neq 0$ 。可知 EHA 系统相对阶为 5, 与系

统维数相等, 因而可进行状态反馈线性化。取反馈控制律和非线性变化映射 $\mathbf{z} = \Phi(\mathbf{x})$ 分别为

$$u = \frac{z_6 - L_g L_f^4 h(\mathbf{x}) - L_f^5 h(\mathbf{x})}{L_g L_f^4 h(\mathbf{x})} \quad (12)$$

$$z_i = \phi_i(\mathbf{x}) = L_f^{i-1} h(\mathbf{x}) \quad 1 \leq i \leq 5 \quad (13)$$

式(12)中, z_6 为新取的控制量, 通过式(12)和式(13), EHA 非线性系统可转换为如下线性系统:

$$\begin{cases} \dot{z}_i = \frac{\partial \phi_i}{\partial \mathbf{x}^T} (f(\mathbf{x}) + g(\mathbf{x})u + \mathbf{d}) = z_{i+1} & 1 \leq i \leq 5 \\ y = z_1 \end{cases} \quad (14)$$

3 模糊滑模控制器设计

3.1 最优滑模面的滑模控制

3.1.1 滑模面函数及其参数确定

EHA 的控制目标是使作动器状态向量 \mathbf{z} 跟踪期望状态向量 \mathbf{z}_r , 因此定义跟踪误差向量为

$$\mathbf{e} = \mathbf{z}_r - \mathbf{z} = [z_{ir} - z_i]^T = [e_i]^T \quad 1 \leq i \leq 5 \quad (15)$$

式中: e_i 为跟踪误差向量 \mathbf{e} 的各个分量形式。

对 EHA 模型进行微分流形变换后得

$$\begin{cases} \dot{e}_i = e_{i+1} & 1 \leq i \leq 4 \\ \dot{e}_5 = z_{5r} - z_6 = \sigma \end{cases} \quad (16)$$

为在下文中方便表示和书写, 定义 $z_{5r} - z_6 = \sigma$ 。将式(16)表示为误差状态方程为

$$[\dot{e}_i]^T = \mathbf{A}[e_i]^T + \mathbf{B}\sigma \quad 1 \leq i \leq 5 \quad (17)$$

式中:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

选取线性切换面:

$$s = \mathbf{C}[e_i]^T = c_1 e_1 + c_2 e_2 + c_3 e_3 + c_4 e_4 + e_5 \quad 1 \leq i \leq 5 \quad (18)$$

式中: $\mathbf{C} = [c_1 \quad c_2 \quad c_3 \quad c_4 \quad 1]$, 其数值大小影响滑动模态的动态品质。

为使滑模具有良好的动态品质和鲁棒性, 本文采用最优控制方法确定滑模切换面的参数值。由于 (\mathbf{A}, \mathbf{B}) 完全能控, 可将式(17)和式(18)化为标准型, 即

$$\begin{cases} \begin{bmatrix} \dot{e}'_1 \\ \dot{e}'_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} e'_1 \\ e'_2 \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{B}_2 \end{bmatrix} \sigma \\ s = [\mathbf{C}_1 \quad \mathbf{C}_2] \begin{bmatrix} e'_1 \\ e'_2 \end{bmatrix} \end{cases} \quad (19)$$

式中:

$$[\dot{e}'_1 \ \dot{e}'_2]^T = [e_1 \ e_2 \ e_3 \ e_4 \ | \ e_5]^T$$

$$[e'_1 \ e'_2]^T = [e_1 \ e_2 \ e_3 \ e_4 \ | \ e_5]^T$$

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & | & 0 \\ 0 & 0 & 1 & 0 & | & 0 \\ 0 & 0 & 0 & 1 & | & 0 \\ 0 & 0 & 0 & 0 & | & 1 \\ \hline 0 & 0 & 0 & 0 & | & 0 \end{bmatrix}$$

$$B = [0 \ B_2]^T = [0 \ 0 \ 0 \ 0 \ | \ 1]^T$$

$$C = [C_1 \ C_2] = [c_1 \ c_2 \ c_3 \ c_4 \ | \ 1]$$

当系统在滑模面上运动时:

$$s = c_1 e_1 + c_2 e_2 + c_3 e_3 + c_4 e_4 + e_5 = 0 \quad (20)$$

根据线性系统最优跟踪问题性能指标定义^[18],同时考虑到系统在滑模面运动时与控制量无关,此时定义系统性能指标函数为

$$J = \int_{t_0}^{\infty} (e^T Q e) dt \quad (21)$$

式中: Q 为给定的正定矩阵,取 $Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$,

且 $Q_{12}^T = Q_{21}$ 。

故二次型展开后可表示为

$$e^T Q e = [e_1'^T \ e_2'^T] \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \begin{bmatrix} e_1' \\ e_2' \end{bmatrix} = e_1'^T Q_{11} e_1' + e_2'^T Q_{21} e_1' + e_1'^T Q_{12} e_2' + e_2'^T Q_{22} e_2' \quad (22)$$

令 $H = Q_{22}^{-1} Q_{21} e_1' + e_2'$,则式(22)可表示为

$$e^T Q e = e_1'^T Q_{11} e_1' + (H - Q_{22}^{-1} Q_{21} e_1')^T Q_{21} e_1' + e_1'^T Q_{12} (H - Q_{22}^{-1} Q_{21} e_1') + (H - Q_{22}^{-1} Q_{21} e_1')^T \cdot Q_{22} (H - Q_{22}^{-1} Q_{21} e_1') = e_1'^T Q_{11}^* e_1' + H^T Q_{22} H \quad (23)$$

则式(19)可表示为

$$\dot{e}'_1 = A_{11}^* e_1' + A_{12} H \quad (24)$$

联立式(23)和式(24),得到一个等价的典型二次型性能指标最优控制问题,其中, $Q_{11}^* = Q_{11} - Q_{12} Q_{22}^{-1} Q_{21}$; $A_{11}^* = A_{11} - A_{12} Q_{22}^{-1} Q_{21}$ 。

由 (A, B) 可控和 Q 正定,可得出 (A_{11}, A_{12}) 可控和 Q_{22} 及 Q_{11}^* 正定,故上面所描述的最优问题有解。根据二次型性能指标最优控制解法可得

$$H = -Q_{22}^{-1} A_{12}^T P e_1' \quad (25)$$

式中: P 可根据黎卡提方程 $PA_{11}^* + A_{11}^{*T} P - PA_{12} Q_{22}^{-1} A_{12}^T P + Q_{11}^* = 0$ 解得。

根据式(25)得

$$(A_{12}^T P + Q_{21}) e_1' + Q_{22} e_2' = 0 \quad (26)$$

从而得到

$$C = [A_{12}^T P + Q_{21} \quad Q_{22}] = [A_{12}^T P + Q_{21} \quad I] \quad (27)$$

3.1.2 滑模控制器设计

为了实现作动器对不确定性和外界干扰的鲁棒控制,设计如下滑模控制:

$$u = u_{eq} + u_{sw} \quad (28)$$

式中: u_{eq} 为滑模控制中等效控制部分; u_{sw} 为滑模控制中切换部分。

将式(12)中的 $L_d L_f^4 h(x)$ 看作外部干扰,先不考虑其影响,令 $\dot{s} = 0$, 则

$$\dot{s} = \sum_{i=1}^4 c_i \dot{e}_i + z_{5r} - z_6 \quad (29)$$

将式(12)代入式(29)得

$$\dot{s} = \sum_{i=1}^4 c_i \dot{e}_i + z_{5r} - L_f^5 h(x) - u L_g L_f^4 h(x) \quad (30)$$

从而可以得到

$$u_{eq} = \frac{1}{L_g L_f^4 h(x)} \left(\sum_{i=1}^4 c_i \dot{e}_i + z_{5r} - L_f^5 h(x) \right) \quad (31)$$

切换控制表示为

$$u_{sw} = \frac{1}{L_g L_f^4 h(x)} K \text{sgn}(s) \quad (32)$$

式中: K 为切换控制中的控制增益, $K = L_d L_f^4 h(x) + \eta, \eta > 0$ 。

构造 Lyapunov 函数 $V(s) = s^2/2$, 显然 $V(0) = 0$, 当 $s \neq 0$ 时, $V(s) > 0$, 故 $V(s)$ 为正定函数,对其求导可得

$$\begin{aligned} \dot{V}(s) = s \dot{s} &= \left\{ s \sum_{i=1}^4 c_i \dot{e}_i + z_{5r} - L_f^5 h(x) - \right. \\ &L_g L_f^4 h(x) \left[\frac{1}{L_g L_f^4 h(x)} \left(\sum_{i=1}^4 c_i \dot{e}_i + z_{5r} - \right. \right. \\ &L_f^5 h(x) \left. \left. \right) + \frac{1}{L_g L_f^4 h(x)} K \text{sgn}(s) \right] - L_d L_f^4 h(x) \left. \right\} = \\ &s(-K \text{sgn}(s) - L_d L_f^4 h(x)) = -\eta |s| \leq 0 \end{aligned} \quad (33)$$

由 Lyapunov 稳定判别定理可知系统是稳定的。

3.2 模糊控制器设计

3.2.1 切换控制增益的模糊控制

在滑模控制中,切换控制中的控制增益用于补偿系统的不确定性及外界干扰,取较大的控制增益可使系统有较好的鲁棒性,但会造成较大抖振。为了保证系统鲁棒性同时削弱抖振,控制增益应随着系统不确定性及外界干扰变化而变化。鉴于此,本文引入模糊控制,通过调节切换控制增益大小的方法,从而减弱抖振。

由指定参数滑模(即第3节所设计最优滑模面)发生条件得

$$\dot{s}/s \leq -\gamma \quad (34)$$

式中： γ 为正常数。

定义 $r = \dot{s}/s$ ，通过比较 r 与 $-\gamma$ 的大小从而调整控制增益的大小。具体分析如下：

1) 当 $r \leq -\gamma$ 时，系统满足指定滑模发生条件，无需改变 K 的大小，即 $\Delta K = 0$ 。

2) 当 $-\gamma < r < 0$ 时，系统满足滑模发生条件，但不满足指定滑模发生条件，需要 K 有较小的改变。 ΔK 的正负与 $L_g L_f^4 h(\mathbf{x})$ 和 s 有关，不失一般性，设 $L_g L_f^4 h(\mathbf{x}) > 0$ ，此时系统已趋向滑模面，则 $s > 0$ 时，应有 $\Delta K > 0$ ，以使 \dot{s} 向 $-\infty$ 移动，从而 r 趋向于更小；则 $s < 0$ 时，应有 $\Delta K < 0$ ，以使 \dot{s} 向 $+\infty$ 移动，从而 r 趋向于更小。

3) 当 $r = 0$ 时，系统不满足滑模发生条件，需要 K 有较大的改变。 ΔK 的正负分析同上。

4) 当 $r > 0$ 时，系统不满足滑模发生条件，且远离滑模面运动，需要 K 有很大的改变。 ΔK 的正负分析同上。

根据上述分析，设计的模糊控制器以 r 为输入， $|\Delta K|$ 为输出，将 $|\Delta K|$ 乘以 $\text{sgn}(s)$ 即得到其真实变化量。因此定义模糊集：

$$r = \{NB, NM, NS, ZO, P\}$$

$$|\Delta K| = \{ZO, VS, S, M, B\}$$

式中：NB 为负大；NM 为负中；NS 为负小；ZO 为零；P 为正；VS 为很小；S 为小；M 为中；B 为大。隶属度函数采用 Sigmoid 型与三角形函数结合获得。

具体模糊规则如下：

- 1) 规则 1。若 r 为 NB，则 $|\Delta K|$ 为 ZO。
- 2) 规则 2。若 r 为 NM，则 $|\Delta K|$ 为 VS。
- 3) 规则 3。若 r 为 NS，则 $|\Delta K|$ 为 S。
- 4) 规则 4。若 r 为 ZO，则 $|\Delta K|$ 为 M。
- 5) 规则 5。若 r 为 P，则 $|\Delta K|$ 为 B。

采用积分的方法对 \hat{K} 的上界进行估计，则

$$\hat{K} = G \text{sgn}(s) \int_0^t |\Delta K| dt \quad (35)$$

式中： G 为比例系数，且 $G > 0$ 。

从而切换控制器为

$$\hat{u}_{sw} = \frac{1}{L_g L_f^4 h(\mathbf{x})} \hat{K} \text{sgn}(s) \quad (36)$$

则模糊滑模控制器为

$$u = u_{eq} + \hat{u}_{sw} \quad (37)$$

3.2.2 切换项的模糊控制

为进一步削弱抖振，可对切换项模糊化。根据滑模控制特点可知，当 s 为零时，控制器为等效

控制；当 s 非零时，控制器为等效控制加切换控制。

根据上述分析，取 $u = u_{eq} + \mu \hat{u}_{sw}$ ，通过模糊系数 μ 将切换项模糊化，定义模糊集：

$$s = \{N, ZO, P\}$$

$$\mu = \{N, ZO, P\}$$

式中： N 为负。隶属度函数采用三角函数。

具体模糊规则如下：

1) 规则 1。若 s 为 N，则 μ 为 P。

2) 规则 2。若 s 为 ZO，则 μ 为 ZO。

3) 规则 3。若 s 为 P，则 μ 为 P。

因此控制律为

$$u = u_{eq} + \mu \hat{u}_{sw} \quad (38)$$

式中：当 $\mu = 1$ 时， $u = u_{eq} + \hat{u}_{sw}$ ，此时控制律与式(37)一样；当 $\mu \neq 1$ 时，此时控制律为切换项模糊化的等效控制。

4 仿真实验

为验证本文所设计的控制器的有效性，在 MATLAB/Simulink 环境下将 EHA 系统模型式(8)和控制算法实现，对 EHA 位移跟踪控制进行仿真。EHA 相关参数如下： $L_q = 2 \text{ mH}$ ， $R = 0.2 \Omega$ ， $n_p = 2$ ， $\psi_r = 0.186 \text{ Wb}$ ， $D_p = 1.0 \times 10^{-6} \text{ m}^3/\text{rad}$ ， $B_m = 3 \times 10^{-4} \text{ N} \cdot \text{m}/(\text{rad} \cdot \text{s}^{-1})$ ， $J_m = 1.5 \times 10^{-3} \text{ kg} \cdot \text{m}^2$ ， $A = 1.8 \times 10^{-3} \text{ m}^2$ ， $V_0 = 1.4 \times 10^{-4} \text{ m}^3$ ， $\beta_e = 6.8 \times 10^8 \text{ N}/\text{m}^2$ ， $\xi = L_{ext} = 1.02 \times 10^{-12} \text{ m}^3/(\text{s} \cdot \text{Pa})$ ， $M_t = 1.4 \text{ kg}$ 。

EHA 理想跟踪轨迹设置为 $y_d = \sin(2\pi t)$ ，初始位置设置为 0.8，外界干扰主要考虑外负载力，其采用高斯函数表示为 $F_L = a \exp[-(t-c)^2/2b^2]$ ， a 决定负载力大小， b 表示负载力作用时间范围， c 表示负载力的中心， $a = 10 \text{ kN}$ ， $b = 0.5$ ， $c = 1.5$ 。同时由于作动器工作频宽为 0~3 Hz，故本文所设计控制器对跟踪轨迹为 $y_d = \sin(4\pi t)$ 和 $y_d = \sin(6\pi t)$ 信号进行了仿真。

仿真结果如图 2~图 6 所示。图 2 为反馈线性化控制仿真结果，图 3 为反馈线性化滑模控制仿真结果，图 4 为反馈线性化模糊滑模控制仿真结果，图 5 为反馈线性化最优滑模面双模糊滑模控制仿真结果，图 6 为 4 种控制方法跟踪误差的对比，图 7 为本文方法在跟踪信号频率为 2 Hz 和 3 Hz 下的跟踪误差对比。

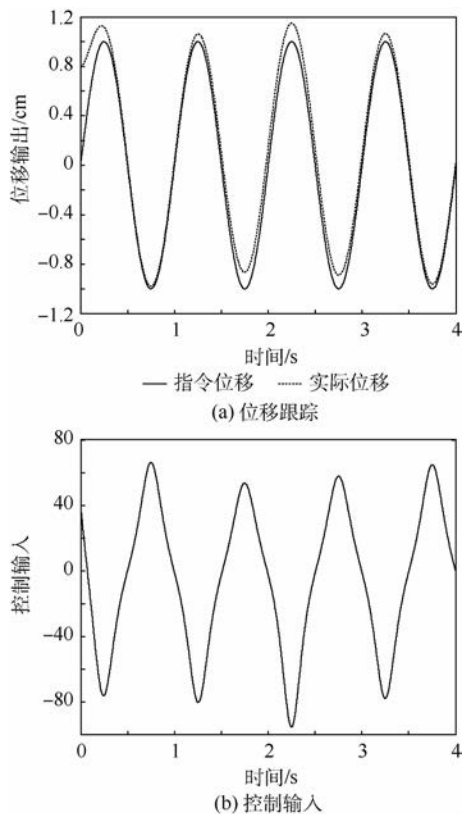


图 2 反馈线性化控制仿真结果

Fig. 2 Simulation results of feedback linearization control

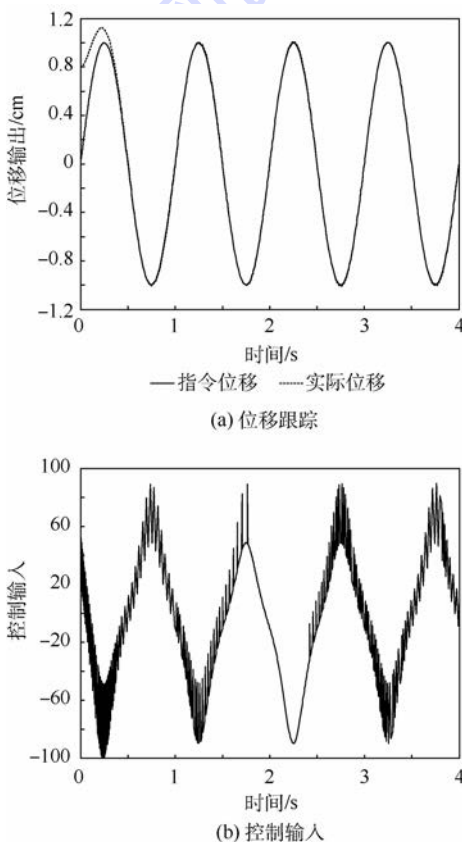


图 3 反馈线性化滑模控制仿真结果

Fig. 3 Simulation results of sliding mode control based on feedback linearization

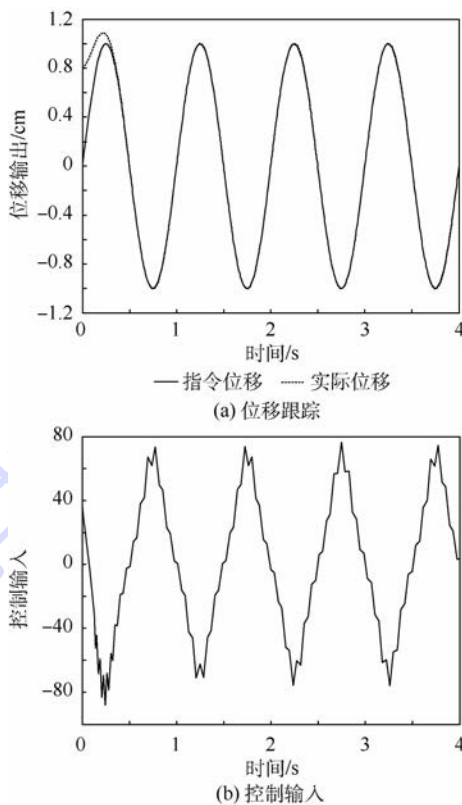


图 4 反馈线性化模糊滑模控制仿真结果

Fig. 4 Simulation results of fuzzy sliding mode control based on feedback linearization

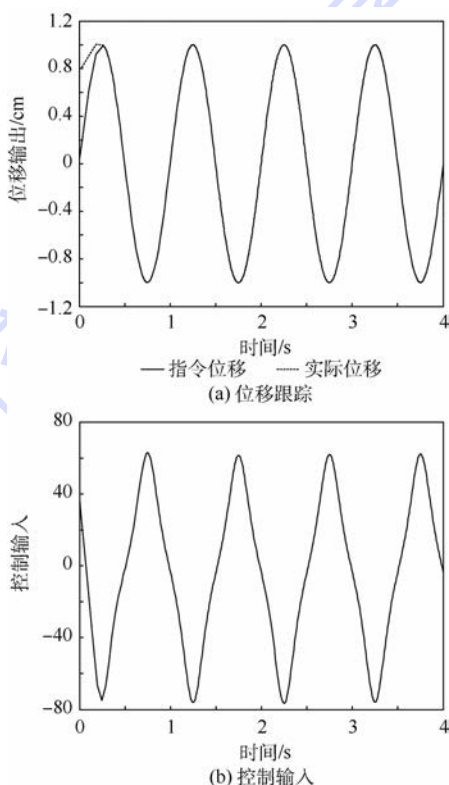


图 5 反馈线性化最优滑模面双模糊滑模控制仿真结果

Fig. 5 Simulation results of double fuzzy sliding mode control with optimal sliding surface based on feedback linearization

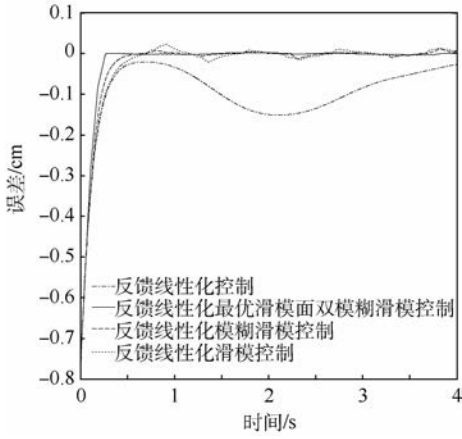


图6 4种控制方法跟踪误差的对比

Fig.6 Tracking errors of four control schemes

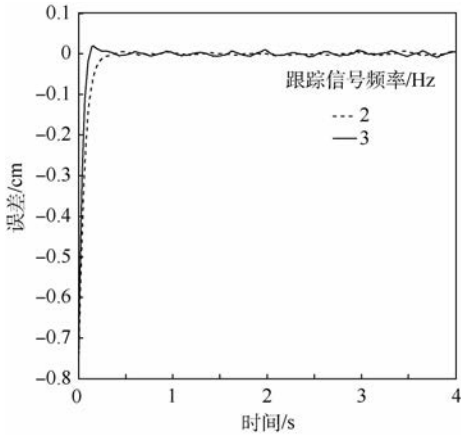


图7 不同频率跟踪信号下的跟踪误差的对比

Fig.7 Tracking errors under different frequency tracking signals

从图2和图6中可以看出,在有外界干扰的情况下,反馈线性化控制会有较大的误差,而有滑模控制的则不受外界干扰的影响,说明其有较好的鲁棒性,从而实现系统控制的鲁棒性目标。

结合图3~图5可知,滑模控制虽具有较好的鲁棒性,但存在抖振,造成控制输入及跟踪误差曲线在零点附近的抖动。模糊滑模控制削弱了抖振,而本文所提出的方法更进一步地削弱了抖振;同时可以看出本文所提方法的调整时间大概在0.3s左右,而另外3种方法的调整时间大概在0.5s左右,因此本文所提出方法还减少了调整时间,提高了响应速度。

从图6和图7中可以看出,在高频率下跟踪误差会变大,这是因为在高频率下更易引起滑模控制的抖振,从而造成跟踪误差曲线的抖动;同时,可以看出其跟踪误差幅值基本在0.05mm之内,可以满足EHA控制要求。

通过仿真结果的对比分析可知,本文所提方法不仅提高了系统的鲁棒性,而且通过反馈线性化建立最优滑模切换函数提高了系统的响应速

度,同时通过双模糊控制器的设计削弱了系统的抖振,可以满足EHA控制精度要求,说明本文所提出方法是有效的。

5 结论

本文针对EHA存在的强非线性和外界干扰,设计了一种新型控制器,仿真结果证明:

1) 本文所提控制算法与反馈线性化控制相比具有较好鲁棒性,与传统的滑模和模糊滑模控制相比进一步削弱了抖振且提高了响应速度。

2) 在EHA工作频宽范围内,可以满足其控制精度要求。

本文方法同样适用于类似的非线性控制系统,尤其是在航空、航天等对控制精度要求较高的领域。

参考文献 (References)

- [1] QI H T, FU Y L, QI X Y, et al. Architecture optimization of more electric aircraft actuation system [J]. Chinese Journal of Aeronautics, 2011, 24(4): 506-513.
- [2] DE LUCENA S E, SUZUKI E Y. Electro-hydraulic actuator tester for fly-by-wire aircrafts [C] // 2007 IEEE Instrumentation and Measurement Technology Conference. Piscataway, NJ: IEEE Press, 2007: 59-64.
- [3] YU S, GANEV E. Next generation power and thermal management system [J]. SAE International Journal of Aerospace, 2008, 1(1): 1107-1121.
- [4] TAKAHASHI N, KONDO T, TAKADA M, et al. Development of prototype electro-hydrostatic actuator for landing gear extension and retraction system [C] // Proceedings of the 7th JFPS International Symposium on Fluid Power. Piscataway, NJ: IEEE Press, 2008: 165-168.
- [5] WANG S, BURTON R, HABIBI S. Sliding mode controller and filter applied to a model of an electro-hydrostatic actuator system [J]. Journal of Dynamic Systems, Measurement, and Control, 2011, 133(3): 024504-1-024504-7.
- [6] GHAZALI R, SAM Y M, RAHMAT M F, et al. Two-degree-of-freedom robust control for a non-minimum phase electro-hydraulic system [C] // Proceedings of 24th Chinese Control and Decision Conference. Piscataway, NJ: IEEE Press, 2012: 1184-1190.
- [7] YANG L, YANG S, BURTON R. Modeling and robust discrete time sliding mode control design for a fluid power electro-hydraulic actuator system [J]. IEEE/ASME Transactions on Mechatronics, 2013, 18(1): 1-10.
- [8] LEE J M, KIM H M, PARK S H, et al. A position control of electro-hydraulic actuator systems using the adaptive control scheme [C] // Proceedings of the 7th Asian Control Conference. Piscataway, NJ: IEEE Press, 2009: 21-26.
- [9] SEO I S, SHIN J M, HAN M K, et al. Adaptive position control scheme with neural networks for electro-hydrostatic actuator systems [C] // Proceedings of the 2010 IEEE International Conference on Mechatronics and Automation. Piscataway, NJ: IEEE

Press,2010:501-506.

- [10] LING T G, RAHMAT M F, HUSAIN A R. ANFIS modeling and direct ANFIS inverse control of an electro-hydraulic actuator system [C]//Proceedings of 8th Conference on Industrial Electronics and Applications. Piscataway, NJ:IEEE Press,2013:370-375.
- [11] AHN K K, TRUONG D Q, SOO Y H. Self-tuning fuzzy PID control for hydraulic load simulator[C]// Proceedings of IEEE International Conference on Control, Automation and Systems. Piscataway, NJ:IEEE Press,2007:2098-2102.
- [12] XU Z, RAHMAN M F. Direct torque and flux control of an IPM synchronous motor using variable structure control approach [C]//The 30th Annual Conference of the IEEE Industrial Electronic Society. Piscataway, NJ:IEEE Press,2004:2733-2738.
- [13] HA Q P, NGUYEN Q H, RYE D C. Fuzzy sliding-mode controllers with applications [J]. IEEE Transactions on Industrial Electronics,2001,48(1):38-46.
- [14] 康荣杰,焦宗夏,JEAN C M,等.电动静液作动器非线性框图建模与鲁棒控制方法[J].航空学报,2009,30(3):518-525.
KANG R J, JIAO Z X, JEAN C M, et al. Nonlinear block diagram model and robust control of electro-hydrostatic actuator [J]. Acta Aeronautica et Astronautica Sinica, 2009, 30(3): 518-525 (in Chinese).
- [15] CHINNIAH Y A. Fault detection in the electrohydraulic actuator using extended Kalman filter[D]. Saskatoon: University of Saskatchewan,2004:170-183.
- [16] 贺昱曜,闫茂德.非线性控制理论及应用[M].西安:电子科技大学出版社,2007:115-119.
HE Y Y, YAN M D. Nonlinear control theory and application

[M]. Xi'an: Electronic Science and Technology University Press,2007:115-119 (in Chinese).

- [17] 李殿璞.非线性控制系统理论基础[M].北京:清华大学出版社,2014:188-215.
LI D P. Theoretical basis of nonlinear control system [M]. Beijing: Tsinghua University Press,2014:188-215 (in Chinese).
- [18] 于长官.现代控制理论及应用[M].哈尔滨:哈尔滨工业大学出版社,2005:170-176.
YU C G. Modern control theory and application [M]. Harbin: Harbin Institute of Technology Press,2005:170-176 (in Chinese).

作者简介:

张振 男,博士研究生。主要研究方向:机载及弹载作动器一体化集成设计与控制。

Tel.: 15653527266

E-mail: 755310683@qq.com

李海军 男,博士,教授,博士生导师。主要研究方向:机载及弹载作动器一体化集成设计与控制,航空导弹测试与故障诊断。

Tel.: 13081617309

E-mail: li_haijun@sohu.com

诸德放 男,硕士,副教授,硕士生导师。主要研究方向:武器系统与运用工程。

Tel.: 13505205558

E-mail: zdf5558@163.com

Double fuzzy sliding mode control for EHA based on feedback linearization optimal sliding surface

ZHANG Zhen¹, LI Haijun^{1,*}, ZHU Defang²

(1. Department of Ordnance Science and Technology, Naval Aeronautical and Astronautical University, Yantai 264001, China;

2. Department of Aerial Ammunition, Air Force Logistics College, Xuzhou 221000, China)

Abstract: In order to improve the control performance for electro-hydrostatic actuator (EHA), double fuzzy sliding mode control based on feedback linearization optimal sliding surface was proposed. The nonlinear model of EHA was linearized based on the feedback linearization by establishing linear switching function and adopting optimal control theory. Fuzzy control was introduced into sliding mode control to reduce the chattering, and a fuzzy controller was adopted to estimate the switching control gain based on the characteristics of optimal sliding mode switching function; another fuzzy controller was adopted to adjust the corrective controller based on the principle of the sliding mode control. The simulation results show that the proposed control scheme is effective.

Key words: electro-hydrostatic actuator (EHA); feedback linearization; sliding mode control; optimal control; fuzzy control

Received: 2015-07-07; Accepted: 2015-09-11; Published online: 2015-10-14 16:22

URL: www.cnki.net/kcms/detail/11.2625.V.20151014.1622.011.html

* Corresponding author. Tel.: 13081617309 E-mail: li_haijun@sohu.com

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0428

基于 EEG 的脑力疲劳特征研究



范晓丽¹, 牛海燕², 周前祥^{1,*}, 柳忠起¹

(1. 北京航空航天大学 生物与医学工程学院, 北京 100083; 2. 北京特种车辆研究所, 北京 100072)

摘 要: 模拟飞行员在飞行过程中监视仪表信息的过程, 分析脑电 (EEG) 随脑力疲劳变化的特点及规律, 从而为后期对抗脑力疲劳提供科学根据。通过设计 2 级不同难度的视觉监控任务分别诱发脑力疲劳, 采用多种方法相结合进行研究, 比较 EEG 参数 (δ 、 θ 、 α 、 β 、 $(\alpha + \theta)/\beta$ 、 α/β 、 $(\alpha + \theta)/(\alpha + \beta)$ 和 θ/β) 在任务前后的变化情况。结果表明: 从正常到疲劳状态, 额区、中央区、顶区和枕区的 α 波相对能量显著增加 ($P < 0.05$); 前额区、侧额区、后额区以及枕区的 β 波相对能量显著降低 ($P < 0.05$); δ 波和 θ 波相对能量变化未达到显著性差异 (均有 $P > 0.05$); 参数 $(\alpha + \theta)/\beta$ 、 α/β 、 $(\alpha + \theta)/(\alpha + \beta)$ 和 θ/β 在除额区外的各脑区都显著增大 ($P < 0.05$); 在额区, 只有 α/β 在疲劳前后增加明显 ($P < 0.05$); 与较高难度的任务比较, 低难度任务中的各 EEG 参数变化较为明显。因此, 除 δ 波和 θ 波以外的其他特征参数被证实可在特定的脑区域可以作为衡量脑力疲劳的潜在指标, 同时可以验证适当地增加任务难度可以在某种程度上对抗脑力疲劳的产生。

关键词: 脑力疲劳; 任务难度; 视觉监控; 脑电 (EEG); 小波分析

中图分类号: TB18; B842. 1

文献标识码: A **文章编号:** 1001-5965(2016)07-1406-08

随着航空技术的进步, 现代飞机系统也在不断地更新与发展, 并逐渐趋于机械化以及智能化, 使任务集中在少数几个人身上。虽然某种程度上减少了体力劳动, 但飞行员的生理和心理承受的负荷越来越大, 再加上夜航、跨时区飞行, 导致飞行员生物钟紊乱, 睡眠缺乏, 因而引发严重的脑力疲劳^[1]。脑力疲劳会导致人的作业机能衰退, 主要症状为头脑昏沉, 注意力不集中, 思考困难, 健忘, 工作绩效下降且易出差错, 因而会导致严重的飞行故障^[2-3]。在这样的背景下, 如何评价和预测脑力疲劳已成为航空医学中的重要研究课题之一。

目前, 关于脑力疲劳的评定方法有主观评定法和客观评定法。主观评定法主要通过问卷调查

的形式进行, 这种评定方式可提供脑力疲劳的多种信息, 如脑力疲劳是否产生, 产生时间、产生的原因以及疲劳程度, 常用的有 Piper 疲乏量表、Epworth 嗜睡量表和 Stanford 疲乏量表等^[4-5]。但主观评定法具有评分主观、评分标准不统一及受个体差异影响严重等缺点, 不能准确地对人的脑力疲劳情况进行客观反映, 因此在实际的脑力疲劳评定工作中仅用主观评定法是不够的, 还需要有客观评定法予以支持。客观评定法包括心理学、行为学指标评定法, 如通过心理运动测验取得人的知觉、认知解释和运动反应等, 前提是假设作业绩效降低是疲劳的标志。其次还包括生理学指标评定法, 主要是一些电生理指标, 如脑电 (Electroencephalogram, EEG)、眼电 (Electrooculogram,

收稿日期: 2015-06-26; 录用日期: 2015-10-10; 网络出版时间: 2015-12-17 10:41

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151217.1041.003.html

基金项目: 国家自然科学基金 (31170895); 国防预研基金 (A0920132003); 中国航天员中心人因工程重点实验室开放课题 (HF2013-K-06)

* 通讯作者: Tel.: 010-82338696 E-mail: zqxg@buaa.edu.cn

引用格式: 范晓丽, 牛海燕, 周前祥, 等. 基于 EEG 的脑力疲劳特征研究[J]. 北京航空航天大学学报, 2016, 42(7): 1406-1413.
FAN X L, NIU H Y, ZHOU Q X, et al. Mental fatigue characteristics based on EEG analysis[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1406-1413 (in Chinese).

EOG)、肌电 (Electromyogram, EMG) 以及心电图 (Electrocardiogram, ECG) 等^[6]。在这些生理参数中,由于 EEG 直接反应大脑活动,因而被认为是最可靠有效的指标之一,广泛应用于脑力疲劳研究中。一般,通过分析入睡过程或者长时间脑力作业过程中的 EEG 变化来评价疲劳和困倦状态^[7-8]。

关于疲劳引起 EEG 变化的研究很多,一般认为从正常状态向疲劳状态转变的过程中,慢波增加(δ, θ),快波降低(α, β)^[9-12]。有研究直接将 EEG 各节律波的比值($\alpha + \theta$)/ β 、 α/β 、($\alpha + \theta$)/($\alpha + \beta$)和 θ/β 作为脑力疲劳的检测指标^[13-14]。但是由于不同研究中实验任务设计不同,导致这些指标在疲劳状态时表现出来的特性也会有一些差异,目前还未达成一致的共识。

在飞机飞行过程中,飞行员的一个重要任务就是通过监视仪表盘来获取各种参数信息,从而实施对系统的操纵和控制。脑力疲劳又是一种涉及生理、精神和行为的复杂过程^[12],它的评定不能依据单一指标进行,而应在整合多学科对脑力疲劳研究的成果基础上,从多角度出发,并且通过主观和客观评定相结合的方法对脑力疲劳进行定性定量研究。本文通过设计 2 级不同难度的视觉监控任务来分别诱发疲劳,并将多种方法相结合,其中,主观问卷和反应时间/正确率法作为辅助评定方法来检验疲劳是否诱发产生并验证 EEG 评定的可靠性,并采用小波包分析 EEG 信号,通过对被试在执行每种难度任务前后 EEG 信号各节律($\delta, \theta, \alpha, \beta$)相对能量以及比值参数($\alpha + \theta$)/ β 、 α/β 、($\alpha + \theta$)/($\alpha + \beta$)和 θ/β 进行分析,从而探讨在视觉监控任务中利用 EEG 特性客观评估脑力疲劳的可行性,以及任务难度对这些 EEG 参数的影响。

1 研究方法

1.1 被试

北京航空航天大学研究生,共 10 人,年龄 23~30 岁,经临床检查无中枢神经系统疾病,临床 EEG 图检查均正常,视力正常或矫正正在 1.0 以上,都具有规律的休息习惯,要求在实验前将头发洗干净、饱餐,保证正常睡眠时间,禁止实验前 24 小时饮酒、喝茶或者咖啡以及吃药物。所有被试自愿参加实验,并填写书面知情同意书,实验结束后均向其支付一定的酬劳。本实验符合相关伦理学的规定和准则,并且取得了北京航空航天大学生物医学与工程学院相关负责部门的同意与认可。

1.2 监控作业设计

在 VC 平台上编制程序“视觉监控实验软件”,其主要是模拟仪表盘来进行设计,界面被分为 $n \times n$ 个小区域,每个小区域上分布一个刺激。实验通过设定 n 值来控制任务难度,本实验选取的 2 级难度的任务分别为 $n = 2$ 与 $n = 3$, 并且都在人的正常认知范围内。图 1(a)与图 1(b)分别为对应的低难度和高难度任务界面,其中,⊕表示靶子刺激,⊖、⊙、⊗、⊘、⊙、⊗为干扰刺激。画面重复切换,间隔时间都固定为 3 s,且每切换一幅画面,刺激重新随机分布,每幅画面中可能“有靶子”,也可能“没有靶子”,有靶子的概率是 50%。实验开始后,需被试观察每幅画面中是否有靶子刺激⊖:①若发现靶子,则迅速按下鼠标左键。②若通过搜索之后确定没有靶子则按鼠标右键。实验程序会自动记录每次操作的正确性和反应时间。

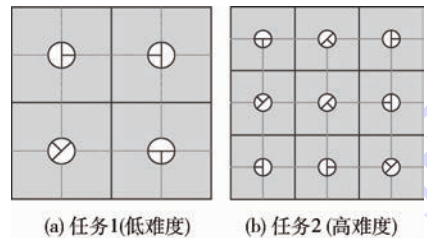


图 1 视觉监控任务界面

Fig. 1 Visual detection task interfaces

1.3 实验步骤

实验安排在一间隔音效果良好的电磁屏蔽室里,根据生物钟时间表以及主观问卷调查初步确立实验时间为上午 10:00-11:00 和下午 3:00-4:00, 10 名被试均参与高、低任务难度下的视觉监控任务,每级任务持续时间为 60 min。为获得被试在实验过程中的操作绩效及主观疲劳感,每级任务过程被分为 12 个 5 min 任务片段,如图 2 所示,被试每完成 5 min 任务,软件自动记录这期间的平均反应时间和正确率,同时要求被试参照疲劳量表(见表 1)以书面的形式给出此刻的疲劳分值,然后迅速进入下一个 5 min 任务。同时记录整个实验过程的 EEG 信号。

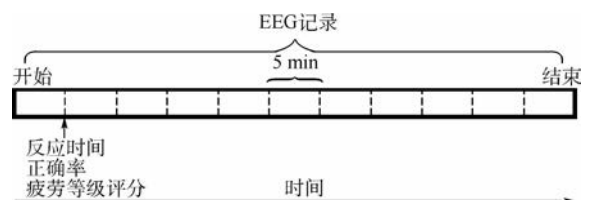


图 2 实验过程设计

Fig. 2 Experiment process design

表1 疲劳量表

Table 1 Subjective fatigue scale

疲劳等级	症状描述	打分
警觉	注意集中, 高度兴奋	0
清醒	清醒, 没有明显感觉	1~2
轻度疲劳	轻微眼乏, 眼涩	3~4
中度疲劳	中度眼乏, 眼涩, 刺痛, 眩晕, 并伴有轻度的头痛与恶心	5~7
严重疲劳	严重眼乏, 眩晕, 头痛厉害, 恶心呕吐	8~10

1.4 信号采集及处理

采用国际脑电图学会标定的 10~20 电极导联定位标准, 并使用 Neuroscan64 导脑电仪采集 EEG 信号。取前额区 FP(Fp1, Fp2)、侧额区 IF(F7, F8)、额区 F(F3, F4, Fz)、颞区 T(T7, T8)、中央区 C(C3, C4, Cz)、后颞区 PT(P7, P8)、顶区 P(P3, P4)和枕区 O(O1, O2, Oz)放置头皮电极, 以乳突电极 M1 和 M2 为参考电极, 垂直 EOG 信号同时被采集。控制头皮阻抗小于 5 kΩ, EEG 波采集率 500 Hz, AC 采集。

原始 EEG 信号非常微弱, 一般在几十微伏量级, 并且在检测过程中不可避免地混入一些干扰成分, 包括 ECG、EOG 和 EMG 等非神经活动的生理信号伪迹, 以及其他仪器设备电磁干扰, 如工频干扰、电极接触不良以及身体移动等原因产生的各种伪差。在本文的研究中, 幅度在 50~70 μV 的波段视为伪迹并直接去除^[15-17]; 采用独立分量分析 (Independent Component Analysis, ICA) 法去除眨眼、眼动等导致的伪迹; 身体移动产生的噪声较明显, 故直接手动去除。最后, EEG 经 0.5~40 Hz 无相移带通数字滤波器滤波。

由于 EEG 波是时变非平稳信号, 选择一种合适的方法来更好地获取反映大脑活动和状态的有效信息是进行 EEG 分析的一个重要前提。近年来, 小波变换不断发展, 相比传统的傅里叶变换, 其具有良好的局部化特性, 非常适于分析非平稳信号的瞬态特性和时变特性。本文利用离散小波包变换检测各 EEG 节律 δ (0.5~4 Hz)、 θ (4~8 Hz)、 α (8~16 Hz) 和 β (16~32 Hz), 考虑到每个节律的频率范围, 采用 Daubechies (dB4) 小波函数对 EEG 信号进行 6 层分解, 分为 64 个频率成分, 最小的频率分辨率用式 (1) 估计:

$$\Delta f = \frac{1}{2^6} \times \frac{f_s}{2} = 3.906 \text{ Hz} \quad (1)$$

式中: f_s 为采集频率, 这里为 500 Hz。

重排后小波包系数依次与各节律 EEG 波 (δ , θ , α , β) 的对应关系如图 3 所示。重建小波分解系数并分别提取 4 个 EEG 节律信号。各频段波

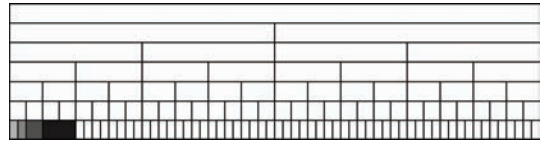


图3 小波包的空间分解

Fig. 3 Spatial decomposition of wavelet packet

的能量 E_j 通过式 (2) 得到:

$$E_j = \sum_{k=1}^L C_j^2(k) \quad j = 1, 2, 3, 4 \quad (2)$$

式中: j 为不同节律的 EEG 波; C_j 为每频段波对应的小波系数; L 为一个频段波中分解系数的数目。

各频段波的相对能量 R_{E_j} 表示为

$$R_{E_j} = \frac{E_j}{\sum_{k=1}^4 E_k} \quad j = 1, 2, 3, 4 \quad (3)$$

比值参数 $(\alpha + \theta)/\beta$ 、 α/β 、 $(\alpha + \theta)/(\alpha + \theta)$ 以及 θ/β 都可以根据式 (2) 和式 (3) 计算得到。

2 结果

2.1 主观分析

以被试 A 为例, 发现其主观脑疲劳分值随着时间的推移逐渐上升, 如图 4 所示。结果说明, 随着监控任务的持续, 被试 A 的疲劳程度不断加深。

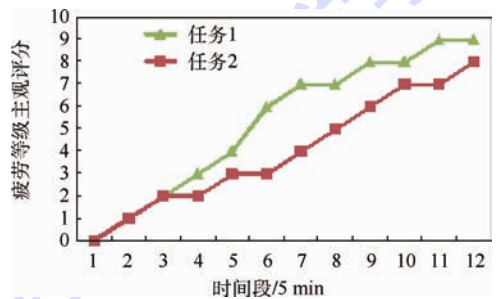


图4 被试 A 主观疲劳量化分值

Fig. 4 Subjective fatigue evaluation scale scores of subject A

从图 4 中还可以看出, 相比高难度的任务 2, 任务 1 更容易导致人疲劳; 并且大多数被试 (7:3) 存在相同的变化趋势, 他们普遍认为在任务 1 持续到 25 min 左右, 而任务 2 到 30 min 左右时, 开始感觉较难集中注意力, 到任务结束时都会伴随心慌、易怒和恶心想吐等症状, 由此可推断持续了 60 min 的 2 级难度任务都已成功诱发了脑力疲劳。

2.2 行为数据分析

高、低任务难度下被试执行监控任务的操作正确率和反应时间随时间变化的曲线如图 5 所示。单因素重复测量方差分析表明: 低难度任务状态下 (任务 1), 被试对刺激信息反应时间显著

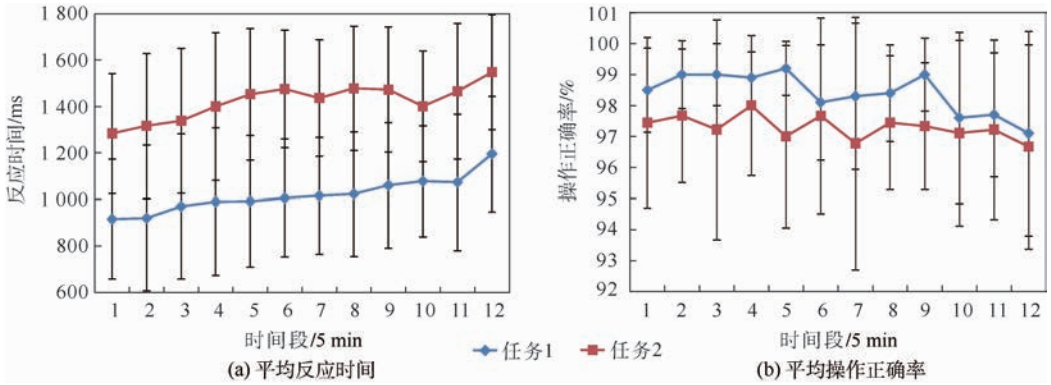


图 5 行为数据随时间变化曲线

Fig. 5 Curves of changes over time in behavior data

延长 ($df = 11, F = 11.413, P = 0$), 并随时间呈线性增长趋势 (拟合系数 $R^2 = 0.88$), df 为自由度, F 为 F 检验的结果, P 为 F 分布的概率, 一般 $P \leq 0.05$ 表示差异显著; 操作正确率降低, 但未达到显著水平。高难度任务状态下 (任务 2), 被试对刺激信息反应时间显著延长 ($df = 11, F = 3.101, P = 0.046$), 并随时间呈线性增长趋势 ($R^2 = 0.69$); 操作正确率降低, 也未达到显著水平。说明随着任务的持续, 被试疲劳的加深, 导致工作绩效下降。

2.3 EEG 分析

为了获取疲劳前后各 EEG 参数的变化特性, 选取每级任务前 5 min 和后 5 min 的 EEG 数据进行比较分析。分别对 2 级任务中 8 个 EEG 参数 (δ 、 θ 、 α 、 β 、 $(\alpha + \theta)/\beta$ 、 α/β 、 $(\alpha + \theta)/(\alpha + \beta)$ 和 θ/β) 在各脑区 (FP、IF、F、C、T、PT、P 和 O) 进

行配对 T 检验分析, 结果如表 2 和表 3 所示。选 O 区并作出疲劳前后各 EEG 参数对比柱状图 (见图 6)。

综合表 2、表 3 和图 6 可以看出, 疲劳前后 EEG 参数在 2 种不同任务难度条件下的变化规律基本是一致的, 即从正常状态到疲劳阶段, F、C、P 和 O 区的 α 波相对能量显著增加 ($P < 0.05$); FP、IF、PT 和 O 区 β 波相对能量显著降低 ($P < 0.05$); δ 和 θ 波疲劳前后变化表现的较为微弱, 没有达到显著性差异 ($P > 0.05$); 4 个比例参数 ($\alpha + \theta$)/ β 、 α/β 、 $(\alpha + \theta)/(\alpha + \beta)$ 和 θ/β 在除 T 区外的各脑区都显著增大 ($P < 0.05$), 在 T 区只有参数 θ/β 显著增大。

进一步比较 2 种任务难度下各 EEG 参数疲劳前后的变化程度, 同样选取 O 区, 通过式 (4) 计

表 2 任务 1 疲劳前后各 EEG 参数在各脑区的配对 T 检验结果

Table 2 Paired T-test results of EEG parameters of task 1 before and after fatigue in all brain regions

参数	FP	IF	F	T	C	PT	P	O
δ								
θ								0.045
α			0.001		0.001		0.001	<0.001
β	<0.001	0.002			0.047	0.001	0.002	<0.001
$(\alpha + \theta)/\beta$	<0.001	<0.001	<0.001		<0.001	<0.001	<0.001	<0.001
α/β	<0.001	<0.001	<0.001	0.018	<0.001	<0.001	<0.001	<0.001
$(\alpha + \theta)/(\alpha + \beta)$	<0.001	<0.001	<0.001		<0.001	<0.001	<0.001	<0.001
θ/β	<0.001	<0.001	0.001		0.001	<0.001	<0.001	<0.001

表 3 任务 2 疲劳前后各 EEG 参数在各脑区的配对 T 检验结果

Table 3 Paired T-test results of EEG parameters of task 2 before and after fatigue in all brain regions

参数	FP	IF	F	T	C	PT	P	O
δ								
θ								
α			0.001		<0.001	<0.001	<0.001	<0.001
β	0.030	0.048				0.040		0.049
$(\alpha + \theta)/\beta$	<0.001	<0.001	<0.001		<0.001	<0.001	<0.001	<0.001
α/β	<0.001	<0.001	<0.001	0.045	<0.001	<0.001	<0.001	<0.001
$(\alpha + \theta)/(\alpha + \beta)$	<0.001	<0.001	<0.001		<0.001	<0.001	<0.001	<0.001
θ/β	<0.001	0.011	<0.001		<0.001	<0.001	<0.001	<0.001

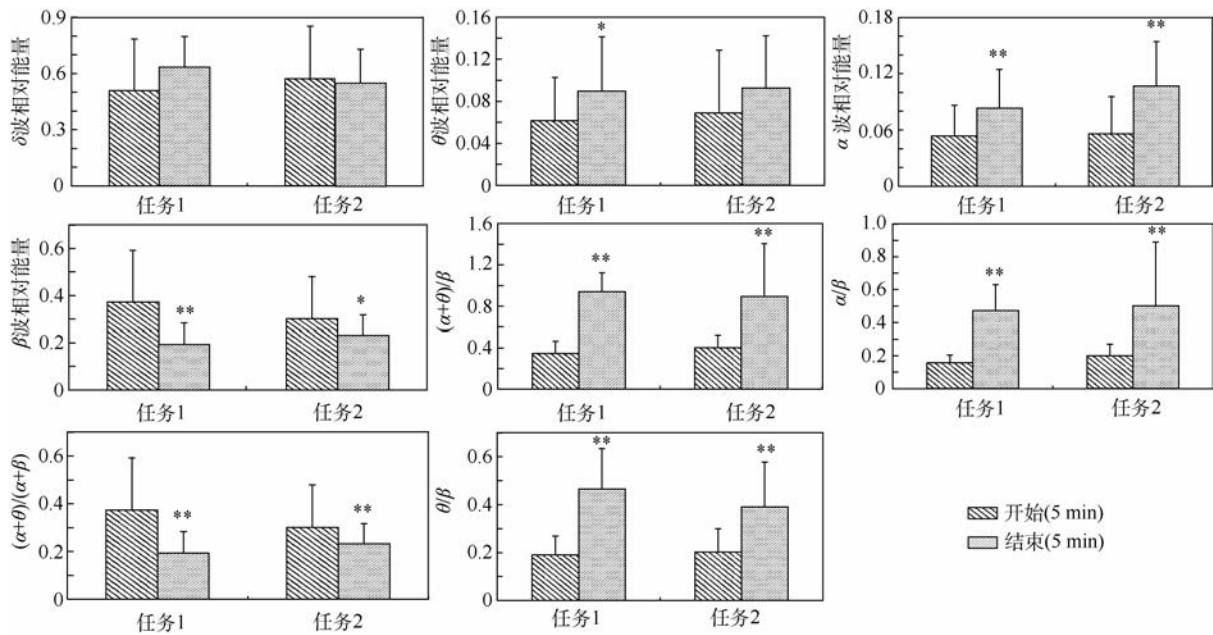


图6 O区疲劳前后各EEG参数对比 (* $P < 0.05$, ** $P < 0.001$)

Fig.6 Comparison of EEG parameters in occipital O between before and after fatigue (* $P < 0.05$, ** $P < 0.001$)

算每个参数疲劳前后的相对变化量:

$$\Delta = \frac{|y_2 - y_1|}{y_1} \quad (4)$$

式中: y_1, y_2 分别为任务前5 min、后5 min对应的参数值。

分别对8个指标在2级难度任务中的 Δ 值进行独立样本T检验,结果表明:疲劳前后, δ ($t = -1.656, df = 39, P = 0.106$), θ ($t = -0.605, df = 57, P = 0.548$), α ($t = -0.770, df = 57, P = 0.445$), β ($t = -0.818, df = 43, P = 0.417$) 波的能量变化以及 $(\alpha + \theta)/\beta$ ($t = -1.906, df = 58, P = 0.062$), α/β ($t = -0.024, df = 58, P = 0.981$), θ/β ($t = -1.916, df = 40, P = 0.062$) 的能量变化在2种难度的任务中表现均无显著性差异;而 $(\alpha + \theta)/(\alpha + \beta)$ ($t = -2.692, df = 58, P = 0.009$) 在2种难度的任务中表现有显著性差异。图7为不同任务难度级别状态下 EEG 参数疲劳前后变化对比。可以看出,在较低难度的任务1中,疲劳后 EEG 参数的变化相比任务2中较为明显,从某种程度上可反映出任务1较易使人产生疲劳,这与主观评价的结果是相吻合的。而通过独立T检验得到大部分指标未达到显著性差异可能是由2级任务的难度级别差距太小而导致的。从图7中还可看出,在 δ, θ, α 和 β 4个节律的波中, α 波在疲劳后的变化程度最大。

以被试B为例,图8显示了被试B在任务1前后 EEG 功率密度。对比可知,在任务后5 min, EEG 功率在8 Hz 和 12 Hz 左右明显增大,正好在 α 波频率范围内,验证了图7的结果,即 α 波变化最明显,能敏锐地反映脑力疲劳的变化。

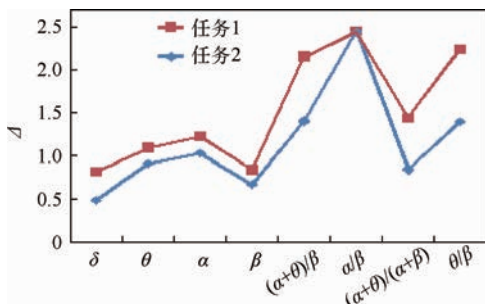


图7 不同任务难度级别状态下 EEG 参数疲劳前后变化对比

Fig.7 Comparison of EEG parameters with different task complexity between before and after fatigue

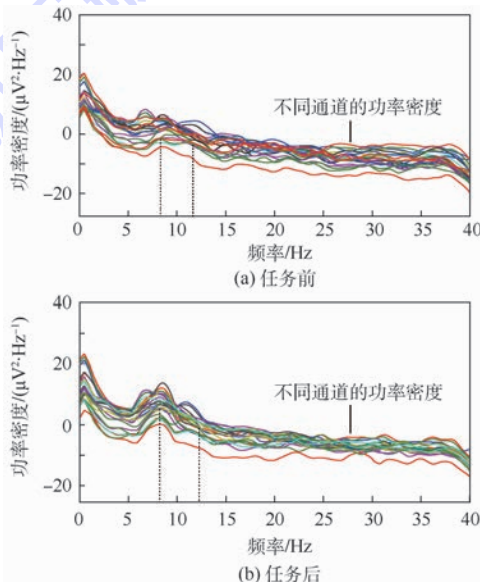


图8 被试B的EEG功率密度

Fig.8 EEG Power spectral density of subject B

3 讨论

脑力疲劳从表象上解释,主要是指心理学、行为学涉及的注意力不集中、反应迟缓和工作能力降低,主观疲劳感增加。本文设计了2种不同难度级别的视觉监控任务来诱发疲劳,使被试在执行过程中视觉器官和中枢神经处于高度紧张状态而导致疲劳。通过主观调查问卷获取被试对自身疲劳程度的评价,根据得分情况定性地推断随着时间的推移,被试脑力疲劳逐渐产生并不断加深。被试在执行任务过程中的行为绩效数据也表明随着任务的持续,工作绩效降低。从主客观2个方面验证了脑力疲劳的产生。

关于脑力疲劳的 EEG 特性的研究,本文采用能量谱估计分析了不同频段波的变化规律,发现 α 波在额区、中央区、顶区和枕区活动增强,而 β 波在前额区、侧额区、后颞区以及枕区活动减弱,与以往的研究基本保持一致。Belyavin 和 Wright^[18]在其研究中指出, β 波活动会在警觉度降低的情况下明显减弱。Torsvall 和 Akerstedt^[19]认为 α 波能灵敏地反映脑力疲劳的变化。同时,Lal 等^[9]认为 δ 波和 θ 波在疲劳期间活动会明显增强,但本文中 δ 波和 θ 波没有发生显著差异,只有轻微的增长趋势,分析原因可能是由于监控任务诱发产生的疲劳较微弱,不足以引起 δ 波和 θ 波发生较明显的波动。而当诱发产生的疲劳不足以引起各频率成分发生明显变化时,选择合理的比例特征参数能更好区分疲劳等级。本文选取反映 EEG 快波与慢波之间比例关系的特征参数 $(\alpha + \theta)/\beta$ 、 α/β 、 $(\alpha + \theta)/(\alpha + \beta)$ 和 θ/β 作为评价指标进行研究,结果表明这些参数在除颞区的整个头皮区域都有显著的增加。Eoh 等^[20]在其研究中得到了类似的结果,即 $(\alpha + \theta)/\beta$ 和 α/β 随着疲劳的增加逐渐增大。由于 δ 波和 θ 波处于一个轻微的上升趋势,因而 $(\alpha + \theta)/(\alpha + \beta)$ 和 θ/β 均处于上升状态。这些 EEG 特征参数在颞区表现迟钝,可能是由于颞区的功能与视觉认知无关所致。

脑力负荷与任务难度之间是一种正比例的关系^[21],本文选定的2级不同难度的视觉监控任务都在人的正常脑力负荷承载范围内。相对难度较高的任务2,EEG 参数在任务1中变化更加明显,这是因为任务1更简单、单调和无挑战性,长时间的重复操作更容易使被试感到疲劳、乏味和厌倦,很难集中精力来完成实验。因此,适当地增加任务难度在某种程度上可以对抗疲劳的产生。

4 结论

本文在小波变换的基础上分析了疲劳时各 EEG 特征参数变化情况,同时将视觉监控任务的反应时间和主观问卷作为疲劳产生及程度的界定指标,结果表明:

1) 除 δ 波和 θ 波相对能量参数,其余特征参数都被证实可以作为衡量脑力疲劳的指标。

2) 任务难度太低,长时间的简单重复操作更易产生乏味和厌倦心理,从而引发脑力疲劳。因此,在飞机座舱人机界面设计时,针对飞行员适当增加任务难度可以在某种程度上对抗脑力疲劳的产生。

本文只是脑力疲劳的初步研究,更深入全面的研究有待于进一步展开:

1) 在任务难度等级划分上,本文只选择了2个难度等级的任务,而且低难度任务选取得过于简单,因而具有一定的局限性,后期将考虑设定更多的难度等级,以便更加深入地探索难度等级对脑力疲劳的影响规律。

2) 被试数量太少,个体差异很容易影响到最后结果。

3) 由于 EEG 信号处理和分析的工作量较大,只选取了任务开始5 min 和任务结束5 min 进行研究,中间过程未进行研究。

4) 除了主观评价和时间反应法,其他测量方法(例如 EOG、EMG 和 ECG 等)都可以与 EEG 结合起来对脑疲劳程度进行综合分析。

在后续的脑力疲劳研究中,将综合考虑这些因素,以期得到更加全面准确的结论。

参考文献 (References)

- [1] 张佳丽,李靖,蒙果,等. 飞行员在模拟飞行训练中脑力疲劳的脑电图研究[J]. 中国应用生理学杂志,2013,29(3): 267-270.
ZHANG J L, LI J, MENG G, et al. Study on fatigue of pilots during simulated flight training based on electroencephalogram [J]. Chinese Journal of Applied Physiology, 2013, 29(3): 267-270 (in Chinese).
- [2] HOPSTAKEN J F, LINDEN D, BAKKER A B, et al. A multi-faceted investigation of the link between mental fatigue and task disengagement[J]. Psychophysiology, 2015, 52(3): 305-315.
- [3] MURATA A, UETAKE A, TAKASAWA Y. Evaluation of mental fatigue using feature parameter extracted from event-related potential [J]. International Journal of Industrial Ergonomics, 2005, 35(8): 761-770.
- [4] LI H, SEO J, KHAM K, et al. Method of measuring subjective 3D visual fatigue: A five-factor model[C] // Digital Holography

- and Three-Dimensional Imaging. [S. l. : s. n.], 2008:44-50.
- [5] HODDES E, ZARCONI V, SMYTHE H, et al. Quantification of sleepiness: A new approach [J]. *Psychophysiology*, 1973, 10 (4):431-436.
- [6] KAR S, BHAGAT M, ROUTHAY A. EEG signal analysis for the assessment and quantification of driver's fatigue[J]. *Transportation Research Part F: Traffic Psychology and Behavior*, 2010, 13(5):297-306.
- [7] SCHMIDT E A, SCHRAUF M, SIMON M, et al. Drivers' misjudgment of vigilance state during prolonged monotonous daytime driving [J]. *Accident Analysis and Prevention*, 2009, 41 (5):1087-1093.
- [8] HOWARD M E, JACKSON M L, BERLOWITZ D, et al. Specific sleepiness symptoms are indicators of performance impairment during sleep deprivation [J]. *Accident Analysis and Prevention*, 2014, 62:1-8.
- [9] LAL S, CRAG A, BOORD P, et al. Development of an algorithm for EEG-based drive fatigue countermeasure [J]. *Journal of Safety Research*, 2003, 34(3):321-328.
- [10] GALE A, DAVIES R, SMALLBONE A. EEG correlates of signal rate, time in task and individual differences in reaction time during a five-stage sustained attention task [J]. *Ergonomics*, 1977, 20(4):363-376.
- [11] PHIPPS-NELSON J, REDMAN J, RAJARATNAM S. Temporal profile of prolonged, night-time driving performance: Breaks from driving temporarily reduce time-on-task fatigue but not sleepiness [J]. *Journal of Sleep Research*, 2011, 20(3):404-415.
- [12] ZHAO C, ZHAO M, LIU J, et al. Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator [J]. *Accident Analysis and Prevention*, 2012, 45:83-90.
- [13] DE WAARD D, BROOKHUIS K. Assessing driver status: A demonstration experiment on the road [J]. *Accident Analysis and Prevention*, 1991, 23(4):297-307.
- [14] JAP B, LAL S, FISCHER P, et al. Using EEG spectral components to assess algorithms for detecting fatigue [J]. *Expert Systems with Applications*, 2009, 36(2):2352-2359.
- [15] LAL S, CRAIG A. Driver fatigue: Electroencephalography and psychological assessment [J]. *Psychophysiology*, 2002, 39(3):313-321.
- [16] SCHIER M A. Changes in EEG alpha power during simulated driving: A demonstration [J]. *International Journal of Psychophysiological*, 2000, 37(2):155-162.
- [17] WILSON G, SWAIN C, ULLSPERGER P. EEG power changes during a multiple level memory retention task [J]. *International Journal of Psychophysiology*, 1999, 32(2):107-118.
- [18] BELYAVIN A, WRIGHT N. Changes in electrical activity of the brain with vigilance [J]. *Electroencephalography and Clinical Neurophysiology*, 1987, 66(2):137-144.
- [19] TORSVALL L, AKERSTEDT T. Sleepiness on the job: Continuously measured EEG changes in train drivers [J]. *Electroencephalography and Clinical Neurophysiology*, 1987, 66(6):502-511.
- [20] EOH H, CHUNG M, KIM S. Electroencephalographic study of drowsiness in simulated driving with sleep deprivation [J]. *International Journal of Industrial Ergonomics*, 2005, 35(4):307-320.
- [21] 曾庆新, 庄达民, 马银香. 脑力负荷与目标辨认 [J]. *航空学报*, 2007, 28(Supl.):76-80.
ZENG Q X, ZHUANG D M, MA Y X. Mental workload and target identification [J]. *Acta Aeronautica et Astronautica Sinica*, 2007, 28(Supl.):76-80 (in Chinese).

作者简介:

范晓丽 女, 博士研究生。主要研究方向: 航空航天人因工程。
Tel.: 18810152142
E-mail: fanfan19851414@163.com

周前祥 男, 博士, 教授, 博士生导师。主要研究方向: 航空航天人因工程。
Tel.: 010-82338696
E-mail: zqx@buaa.edu.cn

Mental fatigue characteristics based on EEG analysis

FAN Xiaoli¹, NIU Haiyan², ZHOU Qianxiang^{1,*}, LIU Zhongqi¹

(1. School of Biology and Medical Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;

2. Beijing Special Vehicle Research Institute, Beijing 100072, China)

Abstract: By simulating the process of pilots monitoring instrument information during flight, the change characteristics of electroencephalogram (EEG) waves along with mental fatigue were analyzed, which will provide scientific bases for future development of countermeasure to fatigue. Two visual detection tasks of different difficulties were designed to induce fatigue respectively, and many measurements were combined to study the EEG characteristics of fatigue. The EEG parameters [$\delta, \theta, \alpha, \beta, (\alpha + \theta)/\beta, \alpha/\beta, (\alpha + \theta)/(\alpha + \beta), \theta/\beta$] at the beginning and end sections of the task were compared and analyzed. The results show that there is a significant increase in α activity in the frontal, central, parietal and occipital lobes ($P < 0.05$), and a decrease in the β activity in the pre-frontal, inferior frontal, posterior temporal and occipital lobes ($P < 0.05$); there is no significant difference in δ rhythm and θ rhythm in any brain region (all in $P > 0.05$); The four formulas increase significantly in all brain regions except the temporal ($P < 0.05$), where only α/β changes clearly ($P < 0.05$); compared with the task with higher difficulty, the parameters in the task of lower difficulty change more obviously. Therefore, all these characteristic parameters in specific brain regions except for δ and θ can be considered as possible indicators for mental fatigue. It was also verified that adding right amount of task difficulty could counter mental fatigue.

Key words: mental fatigue; task difficulty; visual detection; electroencephalogram (EEG); wavelet analysis

Received: 2015-06-26; **Accepted:** 2015-10-10; **Published online:** 2015-12-17 10:41

URL: www.cnki.net/kcms/detail/11.2625.V.20151217.1041.003.html

Foundation items: National Natural Science Foundation of China (31170895); National Defense Pre-research Foundation of China (A0920132003); the Open Project of Human Factors Engineering Key Laboratory (HF2013-K-06)

* **Corresponding author.** Tel. : 010-82338696 E-mail: zqyg@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0462

考虑系统不确定性的高超声速飞行器容错控制



彭程, 王新民*, 谢蓉, 穆凌霞

(西北工业大学 自动化学院, 西安 710072)

摘 要: 针对一类含有执行器失效故障和范数有界不确定性的系统, 提出了一种鲁棒容错跟踪控制设计方法。该方法利用有界实引理, 给出了保证闭环系统稳定且满足鲁棒性能要求的线性矩阵不等式(LMI)的表达式; 为了削弱控制器的保守性, 使用不同的 Lyapunov 变量对应不同的系统状态; 由此带来的非凸优化问题, 发展了一种迭代线性矩阵不等式算法。考虑到该迭代算法的收敛性取决于初值的选择, 经过推导给出了一种求解合适初值的算法。将该算法应用于高超声速飞行器 X-33 跟踪控制器的设计, 仿真结果表明本文算法是可行和有效的。

关键词: 高超声速飞行器; 线性矩阵不等式; 容错控制; 跟踪控制; 鲁棒控制

中图分类号: TP13

文献标识码: A **文章编号:** 1001-5965(2016)07-1414-08

随着科学技术的快速发展与进步, 人们对控制系统的安全性、可靠性及可维护性要求越来越高, 因此控制系统越来越趋向于大型化和复杂化。众所周知, 故障对于任何一个系统来说都是不可避免的, 尤其对于复杂系统而言, 一旦故障发生, 就可能造成重大损失。而容错控制的目的就是当发生故障时, 虽然某些性能出现了降级, 但整个闭环系统仍然可以稳定运行, 并且具有可以接受的性能指标。作为一门新兴的交叉学科, 容错控制为提高复杂动态系统的可靠性和安全性开辟了一条新的途径^[1-2]。

由于实际工程中被控对象的动态特性一般难以用精确的数学模型来描述, 存在不确定性和外部干扰, 因此在容错控制系统设计中同时考虑系统的容错性和鲁棒性具有重要意义。鲁棒容错控制能够在被控对象具有不确定性的情况下, 依然保证闭环系统的联立镇定和完整性, 并且系统的性能指标能够达到设计要求。文献[3]针对航空

发动机, 设计了基于特征结构配置方法的容错控制系统, 但没有考虑输入矩阵不确定性的影响。文献[4]采用基于滑模观测器的控制方法对航天器的飞轮故障进行了鲁棒容错控制。文献[5-6]利用自适应理论设计了高超声速飞行器的鲁棒容错控制器, 但无论是自适应理论还是基于观测器的方法实质上都属于主动容错控制, 需要对故障进行实时准确的检测, 因此其应用具有一定的局限性。由于 H_{∞} 指标既考虑了自身参数摄动的影响, 又刻画了抗外界扰动的能力, 因此基于 H_{∞} 指标的容错控制方法也得到了广泛的研究。文献[7]利用 Riccati 方程来求解鲁棒容错控制器, 求得的控制器不仅能够保证闭环系统的渐近稳定性, 还保证了 H_{∞} 性能, 但在 Riccati 方程的求解过程中缺乏有效的方法来确定待定参数的最佳值, 这给控制器设计带来了很大的保守性。文献[8]利用结构 H_{∞} 和旋翼状态反馈 (Rotor State Feedback, RSF) 控制理论对旋翼的传感器故障进行了

收稿日期: 2015-07-09; 录用日期: 2015-09-25; 网络出版时间: 2015-10-30 17:09

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151030.1709.012.html

基金项目: 陕西省自然科学基金(2014JQ8342)

* 通讯作者: Tel.: 029-88431356 E-mail: wxmin@nwpu.edu.cn

引用格式: 彭程, 王新民, 谢蓉, 等. 考虑系统不确定性的高超声速飞行器容错控制[J]. 北京航空航天大学学报, 2016, 42(7): 1414-1421. PENG C, WANG X M, XIE R, et al. Fault-tolerant control for hypersonic vehicle with system uncertainty[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1414-1421 (in Chinese).

鲁棒容错控制。随着线性矩阵不等式(Linear Matrix Inequality, LMI)技术的发展,利用LMI来解决鲁棒 H_∞ 容错控制问题成为发展趋势^[9-15],其核心是寻找一个闭环系统的二次型Lyapunov函数,从而保证系统具有渐近稳定性,为了方便计算需要对所有的系统状态使用不同的Lyapunov变量,这会导致控制器的求解变为非凸优化。针对该问题,文献[9-11]对所有的故障情况采用同一种Lyapunov变量,显然由此得到的控制器具有较大的保守性;文献[12-15]利用扩展LMI技术,引入松弛变量,使得Lyapunov变量和系统矩阵解耦,但松弛变量的引入在增加设计自由度的同时也会增加LMI维数。

本文针对一类考虑执行器失效故障和范数有界参数不确定的系统,研究了一种用来求解鲁棒 H_∞ 容错跟踪控制问题的迭代LMI算法。该算法允许所有的系统状态使用不同的Lyapunov变量,从而降低了控制系统的保守性;相比于扩展LMI技术,该算法无需额外增加LMI的阶数,就可以解决Lyapunov变量和系统矩阵的耦合问题。通过该算法求得的控制器的鲁棒性小于给定的 H_∞ 范数上界;无论在正常和故障情况下都可以实现跟踪控制,并且正常情况下的系统性能最优,从而实现了系统的鲁棒容错控制。

1 问题描述

考虑如下线性系统:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}'\mathbf{x}(t) + \mathbf{B}'\mathbf{u}(t) + \mathbf{G}\mathbf{w}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \end{cases} \quad (1)$$

式中: $\mathbf{x}(t) \in \mathbf{R}^n, \mathbf{y}(t) \in \mathbf{R}^p, \mathbf{u}(t) \in \mathbf{R}^m, \mathbf{w}(t) \in \mathbf{R}^h$ 分别为系统状态、输出、输入和外部干扰; \mathbf{A} 和 \mathbf{B} 分别为合适维数的系统矩阵和输入矩阵, $\mathbf{A}' = \mathbf{A} + \Delta\mathbf{A}(t), \mathbf{B}' = \mathbf{B} + \Delta\mathbf{B}(t), \Delta\mathbf{A}(t) = \mathbf{E}_a\boldsymbol{\Sigma}(t)\mathbf{F}_a$ 和 $\Delta\mathbf{B}(t) = \mathbf{E}_b\boldsymbol{\Sigma}(t)\mathbf{F}_b$ 为范数有界不确定参数, $\mathbf{E}_a, \mathbf{F}_a, \mathbf{E}_b$ 和 \mathbf{F}_b 为维数合适矩阵, $\boldsymbol{\Sigma}^T(t)\boldsymbol{\Sigma}(t) \leq \mathbf{I}$ 为时变不确定变量。

在系统式(1)中加入执行器故障,可得系统故障模型为

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}'\mathbf{x}(t) + \mathbf{B}'\boldsymbol{\Pi}_L\mathbf{u}(t) + \mathbf{G}\mathbf{w}(t) \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) \end{cases} \quad (2)$$

式中: $\boldsymbol{\Pi}_L = \text{diag}(\pi_{l1}, \pi_{l2}, \dots, \pi_{lm})$ 为执行器效率矩阵, $L = 0, 1, \dots, 2^m - 1$ 。

在控制系统中往往需要对输入指令 $\mathbf{r}(t)$ 进行跟踪,即使系统式(2)的跟踪误差 $\mathbf{e}(t) = \mathbf{r}(t) -$

$\mathbf{C}_s\mathbf{y}(t)$ 满足 $\lim_{t \rightarrow \infty} \mathbf{e}(t) = \mathbf{0}$,矩阵 \mathbf{C}_s 为用来选择与 $\mathbf{r}(t)$ 一致的输出信号。为了保证在执行器故障情况下系统跟踪没有静态误差,引入跟踪误差的积分信号 $\int_0^t \mathbf{e}(t) dt$,得到如下所示的增广系统:

$$\begin{cases} \begin{bmatrix} \dot{\mathbf{e}}(t) \\ \dot{\mathbf{x}}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{0} & -\mathbf{C}_s\mathbf{C} \\ \mathbf{0} & \mathbf{A}' \end{bmatrix} \begin{bmatrix} \int_0^t \mathbf{e}(t) dt \\ \mathbf{x}(t) \end{bmatrix} + \\ \begin{bmatrix} \mathbf{0} \\ \mathbf{B}' \end{bmatrix} \boldsymbol{\Pi}_L \mathbf{u}(t) + \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{bmatrix} \begin{bmatrix} \mathbf{r}(t) \\ \mathbf{w}(t) \end{bmatrix} \end{cases} \quad (3)$$

不妨记增广系统状态和系统外部扰动为

$$\bar{\mathbf{x}}(t) = \begin{bmatrix} \int_0^t \mathbf{e}(t) dt \\ \mathbf{x}(t) \end{bmatrix}, \quad \mathbf{v}(t) = \begin{bmatrix} \mathbf{r}(t) \\ \mathbf{w}(t) \end{bmatrix} \quad (4)$$

于是可得

$$\dot{\bar{\mathbf{x}}}(t) = (\bar{\mathbf{A}} + \bar{\mathbf{E}}_a\bar{\boldsymbol{\Sigma}}(t)\bar{\mathbf{F}}_a)\bar{\mathbf{x}}(t) + (\bar{\mathbf{B}} + \bar{\mathbf{E}}_b\bar{\boldsymbol{\Sigma}}(t)\bar{\mathbf{F}}_b)\boldsymbol{\Pi}_L\mathbf{u}(t) + \bar{\mathbf{G}}\mathbf{v}(t) \quad (5)$$

式中:

$$\bar{\mathbf{A}} = \begin{bmatrix} \mathbf{0} & -\mathbf{C}_s\mathbf{C} \\ \mathbf{0} & \mathbf{A}' \end{bmatrix}, \quad \bar{\mathbf{B}} = \begin{bmatrix} \mathbf{0} \\ \mathbf{B}' \end{bmatrix}, \quad \bar{\mathbf{G}} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{bmatrix} \quad (6)$$

$$\begin{cases} \bar{\mathbf{E}}_a = \begin{bmatrix} \mathbf{0} \\ \mathbf{E}_a \end{bmatrix}, \quad \bar{\mathbf{F}}_a = [\mathbf{0} \quad \mathbf{F}_a], \quad \bar{\mathbf{E}}_b = \begin{bmatrix} \mathbf{0} \\ \mathbf{E}_b \end{bmatrix} \\ \bar{\mathbf{F}}_b = \mathbf{F}_b, \quad \bar{\boldsymbol{\Sigma}}(t) = \boldsymbol{\Sigma}(t) \end{cases} \quad (7)$$

本文控制的目的是:针对增广系统式(6)设计一种控制器使得闭环系统能够渐近稳定。如果只考虑系统的稳定性,可能会导致控制器输出信号 $\mathbf{u}(t)$ 变得很大以致无法实现。为此,需要在设计过程中考虑对控制信号的度量。根据以上设计思想,在增广系统中引入包含系统状态和控制信号的关注输出 $\mathbf{z}(t)$,于是可得

$$\mathbf{z}(t) = \mathbf{C}_z\bar{\mathbf{x}}(t) + \mathbf{D}_z\boldsymbol{\Pi}_L\mathbf{u}(t) \quad (8)$$

式中: \mathbf{C}_z 和 \mathbf{D}_z 为加权矩阵,用来平衡跟踪误差和控制能量。

考虑如下状态反馈控制器:

$$\mathbf{u}(t) = \mathbf{K}\bar{\mathbf{x}}(t) = \mathbf{K}_i \int_0^t \mathbf{e}(t) dt + \mathbf{K}_p\mathbf{x}(t) \quad (9)$$

式中: $\mathbf{K} = [\mathbf{K}_i \quad \mathbf{K}_p] \in \mathbf{R}^{m \times (l+n)}$ 为需求解的控制增益。

将式(9)代入式(6)、式(8),可得带有执行器故障的闭环系统为

$$\begin{cases} \dot{\bar{\mathbf{x}}}(t) = (\bar{\mathbf{A}}' + \bar{\mathbf{B}}'\boldsymbol{\Pi}_L\mathbf{K})\bar{\mathbf{x}}(t) + \bar{\mathbf{G}}\mathbf{v}(t) \\ \mathbf{z}(t) = (\mathbf{C}_z + \mathbf{D}_z\boldsymbol{\Pi}_L\mathbf{K})\bar{\mathbf{x}}(t) \end{cases} \quad (10)$$

式中:

$$\bar{\mathbf{A}}' = \bar{\mathbf{A}} + \Delta\bar{\mathbf{A}}, \quad \Delta\bar{\mathbf{A}} = \bar{\mathbf{E}}_a\bar{\boldsymbol{\Sigma}}(t)\bar{\mathbf{F}}_a \quad (11)$$

$$\bar{\mathbf{B}}' = \bar{\mathbf{B}} + \Delta\bar{\mathbf{B}}, \quad \Delta\bar{\mathbf{B}} = \bar{\mathbf{E}}_b\bar{\boldsymbol{\Sigma}}(t)\bar{\mathbf{F}}_b \quad (12)$$

本文设计的控制器需要使得闭环系统式(10)满足以下要求:

1) 无论在正常或故障情况下,控制器均能够保证闭环系统渐近稳定。

2) 对于带有范数有界不确定性系统式(10),所设计的控制器满足设定的性能指标,同时保证无故障时系统性能最优。

3) 无论在正常或故障情况下,输出信号 $C_y(t)$ 与参考信号 $r(t)$ 之间无稳态误差。

为了方便下文的描述,给出如下引理。

引理 1 Schur 补引理。对给定的对称矩阵

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}, S_{11} \in \mathbf{R}^{n \times n}, \text{以下 3 个条件是等价的:}$$

$$\begin{cases} S < 0 \\ S_{11} < 0, \quad S_{22} - S_{12}^T S_{11}^{-1} S_{12} < 0 \\ S_{22} < 0, \quad S_{11} - S_{12} S_{22}^{-1} S_{12}^T < 0 \end{cases} \quad (13)$$

引理 2^[12,16-17] 设 E, F 和 Ψ 为给定矩阵, Ψ 为对称矩阵,且时变参数矩阵 $\Sigma(t)$ 满足 $\Sigma^T(t) \cdot \Sigma(t) \leq I$ 。当

$$\Psi + E\Sigma(t)F + F^T\Sigma^T(t)E^T < 0 \quad (14)$$

成立时,那么必定存在一个正常数 ε 满足:

$$\Psi + \varepsilon EE^T + \varepsilon^{-1} F^T F < 0 \quad (15)$$

推论 对于实矩阵 X, Y ,其维数符合相乘法则,存在 $\varepsilon > 0$ 使得如下不等式成立:

$$X^T Y + Y^T X \leq \varepsilon X^T X + \varepsilon^{-1} Y^T Y \quad (16)$$

证明 只需令 $\Psi = 0, \Sigma(t) = I$ 且 $E = X^T, F = Y$,即可得结论式(16)。 **证毕**

H_∞ 范数指标是最优控制理论中最有效的性能指标之一,其主要用来刻画系统抗外来干扰的能力,同时也考虑了对自身参数摄动的抑制能力。 H_∞ 范数指标数学描述如下。

考虑如下系统:

$$\begin{cases} \dot{x}(t) = Ax(t) + B_1 u(t) + B_2 w(t) \\ z(t) = C_1 x(t) + D_1 u(t) + D_2 w(t) \end{cases} \quad (17)$$

式(17)中,外界干扰 $w(t)$ 是能量有界信号。 H_∞ 控制问题可以表述为:设计一个控制器使得闭环系统渐近稳定,同时关注输出 $z(t)$ 能量要小于 γ 倍外部干扰信号的能量,即 $\|z(t)\|_2 < \gamma \|w(t)\|_2, \gamma$ 为给定正数。由控制理论可知,这实质上就是要求 $z(t)$ 到 $w(t)$ 传递函数的无穷范数小于给定的 γ 值,即 $\|G(j\omega)\|_\infty < \gamma$ 。

下面将给出一个将 H_∞ 控制问题转化为 LMI 表达式的重要引理——有界实引理。

引理 3 对于可控系统式(17),如果存在一个正定对称矩阵 P 满足如下不等式:

$$\begin{bmatrix} (A')^T P + PA' & PG & C^T \\ G^T P & -\gamma I & 0 \\ C & 0 & -\gamma I \end{bmatrix} < 0 \quad (18)$$

那么该系统是渐近稳定的,且满足 $\|G_{zw}(j\omega)\|_\infty < \gamma$ 。

2 鲁棒容错跟踪控制器设计

为了将本文所要求的鲁棒控制器设计转换为 LMI 描述,将闭环系统式(10)代入引理 3 可得

$$\begin{bmatrix} (\bar{A}' + \bar{B}' \Pi_L K)^T P_L + P_L (\bar{A}' + \bar{B}' \Pi_L K) \\ \bar{G}^T P_L & -\gamma_L I \\ C_z + D_z \Pi_L K & 0 & -\gamma_L I \end{bmatrix} < 0 \quad (19)$$

式中: P_L 和 γ_L 分别表示不同故障状态下的 Lyapunov 变量和鲁棒性能指标。

为了分离出式(19)中不确定参数 $\bar{\Sigma}(t)$,对式(19)进行分解可得

$$\begin{bmatrix} (\bar{A} + \bar{B} \Pi_L K)^T P_L + P_L (\bar{A} + \bar{B} \Pi_L K) \\ \bar{G}^T P_L & -\gamma_L I \\ C_z + D_z \Pi_L K & 0 & -\gamma_L I \end{bmatrix} + \begin{bmatrix} P_L \bar{E}_a \\ 0 \\ 0 \end{bmatrix} \bar{\Sigma}(t) \begin{bmatrix} \bar{F}_a^T \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \bar{F}_a^T \\ 0 \\ 0 \end{bmatrix} \bar{\Sigma}^T(t) \begin{bmatrix} P_L \bar{E}_a \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} P_L \bar{E}_b \\ 0 \\ 0 \end{bmatrix} \bar{\Sigma}(t) \begin{bmatrix} (\bar{F}_b \Pi_L K)^T \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} (\bar{F}_b \Pi_L K)^T \\ 0 \\ 0 \end{bmatrix} \bar{\Sigma}^T(t) \begin{bmatrix} P_L \bar{E}_b \\ 0 \\ 0 \end{bmatrix} < 0 \quad (20)$$

易得式(20)中不等式右边第 1 项为对称矩阵,则由引理 2 可得必存在正标量 $\varepsilon_1, \varepsilon_2$, 满足:

$$\begin{bmatrix} (\bar{A} + \bar{B} \Pi_L K)^T P_L + P_L (\bar{A} + \bar{B} \Pi_L K) \\ \bar{G}^T P_L & -\gamma_L I \\ C_z + D_z \Pi_L K & 0 & -\gamma_L I \end{bmatrix} + \varepsilon_1 \begin{bmatrix} P_L \bar{E}_a \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} P_L \bar{E}_a \\ 0 \\ 0 \end{bmatrix}^T + \varepsilon_1^{-1} \begin{bmatrix} \bar{F}_a^T \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} \bar{F}_a^T \\ 0 \\ 0 \end{bmatrix}^T + \varepsilon_2 \begin{bmatrix} P_L \bar{E}_b \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} P_L \bar{E}_b \\ 0 \\ 0 \end{bmatrix}^T + \varepsilon_2^{-1} \begin{bmatrix} (\bar{F}_b \Pi_L K)^T \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} (\bar{F}_b \Pi_L K)^T \\ 0 \\ 0 \end{bmatrix}^T < 0 \quad (21)$$

由引理 1 进一步可得

$$\begin{bmatrix} (\bar{A} + \bar{B}\Pi_L K)^T P_L + P_L (\bar{A} + \bar{B}\Pi_L K) + \varepsilon_1^{-1} \bar{F}_a^T \bar{F}_a & & & & & \\ \bar{G}^T P_L & & & & & -\gamma_L I \\ C_z + D_z \Pi_L K & & & & & 0 & -\gamma_L I \\ \bar{E}_a^T P_L & & & & & 0 & 0 & -\varepsilon_1^{-1} I \\ \bar{E}_b^T P_L & & & & & 0 & 0 & 0 & -\varepsilon_2^{-1} I \\ \bar{F}_b \Pi_L K & & & & & 0 & 0 & 0 & 0 & -\varepsilon_2 I \end{bmatrix} < 0$$

$$L = 0, 1, \dots, 2^{m-1} \quad (22)$$

注意式(22)中没有包含不确定参数 $\bar{\Sigma}(t)$, 使得不等式求解成为可能。式(22)针对所有故障状态使用不同的 Lyapunov 变量 P_L , 减小了控制器的保守性。为了使得正常情况时系统性能最优, 需要最小化 $\text{tr}(P_0)$ ^[18], tr 表示矩阵的迹, P_0 为无故障状态下的 Lyapunov 变量。如此便将鲁棒控制器设计问题转化为: 在 LMI 式(22)约束下目标函数为 $\min(\text{tr}(P_0))$ 的优化问题。

易知式(22)仍属于非凸优化, 无法利用现有的 LMI 工具求解。因此, 本文发展了一种迭代 LMI 算法来解决此问题, 为保证该迭代算法的有效运行, 需要给出一个合适的初值。文献[19]利用互补线性化方法求解初始控制器, 但在多参数线性化过程中, 该方法往往无解。因此, 本文给出了一种初始控制器求解方法。为了方便叙述, 将执行器故障模型进行改写。记:

$$\begin{cases} \Pi_0 = \text{diag}(\pi_{01}, \pi_{02}, \dots, \pi_{0m}) \\ V = \text{diag}(v_1, v_2, \dots, v_m) \\ W = \text{diag}(w_1, w_2, \dots, w_m) \\ |V| = \text{diag}(|v_1|, |v_2|, \dots, |v_m|) \end{cases} \quad (23)$$

$$\begin{cases} \pi_{0j} = 0.5(\pi_{lj} + \pi_{uj}) \\ w_j = \frac{\pi_{uj} - \pi_{lj}}{\pi_{lj} + \pi_{uj}} \\ v_j = \frac{\pi_{lj} - \pi_{0j}}{\pi_{0j}} \end{cases} \quad j = 1, 2, \dots, m \quad (24)$$

式中: π_{lj} 、 π_{uj} 分别为控制效率矩阵的上、下界。

将式(24)代入式(23), 即可得执行器效率矩阵为

$$\Pi_L = \Pi_0(I + V) \quad (25)$$

并且易得 $|V| \leq W \leq I$ 。

定理 1 考虑闭环扩维系统式(10), 对于给定的 H_∞ 范数上界 γ , 如果存在对称正定矩阵 X , 矩阵 Y 、 Z 和正标量 $\varepsilon_1 \sim \varepsilon_5$, 满足:

$$\min(\text{tr}(Y))$$

$$\text{s. t. } \begin{bmatrix} N_1 & & & & \\ \bar{G}^T & & & & -\gamma I \\ C_z X + D_z Z & & & & 0 & -\gamma I + \varepsilon_5 D_z W D_z^T \\ N_2^T & & & & 0 & 0 & N_3 \end{bmatrix} < 0 \quad (26)$$

$$\begin{bmatrix} Y & I \\ I & X \end{bmatrix} > 0 \quad (27)$$

$$N_1 = \bar{A}X + \bar{B}Z + (\varepsilon_3 + \varepsilon_4)\lambda_b I + \varepsilon_1 \lambda_a I + (\bar{A}X + \bar{B}Z)^T + \varepsilon_2 \bar{B}W\bar{B}^T$$

$$N_2 = [Z^T, X, Z^T, Z^T, Z^T]$$

$$N_3 = -\text{diag}(\varepsilon_2 W^{-1}, \varepsilon_1, \varepsilon_3, \varepsilon_4 W^{-1}, \varepsilon_5 W^{-1})$$

则由此获得的状态反馈增益矩阵 $K = \Pi_0^{-1} Z X^{-1}$ 必能够保证闭环系统式(10)是渐近稳定的, 闭环系统的鲁棒性能小于给定的 H_∞ 范数上界。

证明 由引理 3 可知, 闭环系统是渐近稳定的, 且满足 $\|G_{zw}(j\omega)\|_\infty < \gamma$ 的充要条件是式(19)成立。不妨对所有的故障情况均使用同一个 Lyapunov 函数变量 P , 且对式(19)左右两边同乘以 $\text{diag}(P^{-1}, I, I)$, 并记 $X = P^{-1}$, 于是可得

$$\begin{bmatrix} N_4 & & & & \\ \bar{G}^T & & & & -\gamma I \\ (C_z + D_z \Pi_L K)X & & & & 0 & -\gamma I \end{bmatrix} < 0 \quad (28)$$

$$N_4 = [\bar{A} + \Delta\bar{A} + (\bar{B} + \Delta\bar{B})\Pi_L K]X + X[\bar{A} + \Delta\bar{A} + (\bar{B} + \Delta\bar{B})\Pi_L K]^T$$

将式(28)进行分解可得

$$\begin{bmatrix} N_4 & & & & \\ \bar{G}^T & & & & -\gamma I \\ (C_z + D_z \Pi_0 K)X & & & & 0 & -\gamma I \end{bmatrix} + \begin{bmatrix} 0 & 0 & X(D_z \Pi_0 V K)^T \\ 0 & 0 & 0 \\ (D_z \Pi_0 V K)X & 0 & 0 \end{bmatrix} < 0 \quad (29)$$

令 $Z = \Pi_0 K X$, 同时由引理 2 的推论可得

$$N_4 \leq \bar{A}X + \bar{B}Z + (\bar{A}X + \bar{B}Z)^T + \varsigma_1 \Delta \bar{A}^T \Delta \bar{A} + \varsigma_1^{-1} XX + \varsigma_2 \bar{B} |V| \bar{B}^T + \varsigma_2^{-1} Z^T |V| Z + \varsigma_3 \Delta \bar{B} \Delta \bar{B}^T + \varsigma_3^{-1} Z^T Z + \varsigma_4 \Delta \bar{B} |V| \Delta \bar{B}^T + \varsigma_4^{-1} Z^T |V| Z \quad (30)$$

将式(25)代入式(30)可得

$$N_4 \leq \bar{A}X + \bar{B}Z + (\bar{A}X + \bar{B}Z)^T + (\varsigma_1 \lambda_a + \varsigma_3 \lambda_b + \varsigma_4 \lambda_b) I + \varsigma_1^{-1} XX + \varsigma_2 \bar{B} \bar{W} \bar{B}^T + \varsigma_2^{-1} Z^T W Z + \varsigma_3^{-1} Z^T Z + \varsigma_4^{-1} Z^T W Z \quad (31)$$

其中: $\|\Delta \bar{A}\|_2 \leq \sqrt{\lambda_a}$; $\|\Delta \bar{B}\|_2 \leq \sqrt{\lambda_b}$ 。

同理可得

$$\begin{bmatrix} 0 & 0 & X(D_z \Pi_0 V K)^T \\ 0 & 0 & 0 \\ (D_z \Pi_0 V K) X & 0 & 0 \end{bmatrix} \leq \begin{bmatrix} \varsigma_5^{-1} Z^T W Z & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \varsigma_5 D_z W D_z^T \end{bmatrix} \quad (32)$$

于是可得

$$\begin{bmatrix} N_4 \\ \bar{G}^T & -\gamma I \\ (C_z + D_z \Pi_L K) X & 0 & -\gamma I \end{bmatrix} \leq \begin{bmatrix} N_5 \\ \bar{G}^T & -\gamma I \\ C_z X + D_z Z & 0 & -\gamma I + \varsigma_5 D_z W D_z^T \end{bmatrix} \quad (33)$$

$$N_5 = \bar{A}X + \bar{B}Z + (\bar{A}X + \bar{B}Z)^T + (\varsigma_1^{-1} + \varsigma_3) X^2 + \varsigma_2 \bar{B} \bar{W} \bar{B}^T + (\varsigma_1 \lambda_a + \varsigma_3 \lambda_b + \varsigma_4 \lambda_b) I + \varsigma_2^{-1} Z^T W Z + \varsigma_3^{-1} Z^T Z + \varsigma_4^{-1} Z^T W Z$$

由此可知,不等式(33)右边是式(28)的一个强约束条件,因此不等式

$$\begin{bmatrix} N_5 \\ \bar{G}^T & -\gamma I \\ C_z X + D_z Z & 0 & -\gamma I + \varsigma_5 D_z W D_z^T \end{bmatrix} < 0 \quad (34)$$

的解也是不等式(28)的解,进一步由引理1可知,式(34)和式(26)等价,即若有矩阵变量 (X, Z) 使得不等式(26)成立,则 $P = X^{-1}$ 必然是引理3所要求的LMI的解,因而由此得到的控制器就能保证闭环系统是渐近稳定的,且鲁棒性能小于给定的 H_∞ 范数上界,这就是式(26)的物理意义。

对式(27)应用引理1,可得 $X^{-1} < Y$ 等价于 $P < Y$,因此要求 $\text{tr}(Y)$ 极小就是要求 $\text{tr}(P)$ 极小,由保性能控制理论可知,这样做可以保证系统的性能最优,这也就是式(27)的物理意义。证毕

根据得到的初始控制器 $K = \Pi_0^{-1} Z X^{-1}$,利用

下面给出的迭代LMI算法用来解决式(22)中的非凸求解问题。

Step 1 选择合适的 H_∞ 范数上界 γ ,根据定理1计算初始控制器 $K^0 = \Pi_0^{-1} Z^0 (X^0)^{-1}$ 。

Step 2 令 $K = K^0$,由LMI式(22)所约束的优化问题,计算得到对应不同故障情况的初始Lyapunov变量 P_L^0 。

Step 3 在第 j 步迭代($j > 0$)。

1) 将 $P_L = P_L^{j-1}$ 代入不等式(22),求解其中的优化问题,得到第 j 步的控制器增益 K_{opt}^j 。

2) 将 $K = K_{opt}^j$ 代入不等式(22),求解其中的优化问题,得到第 j 步的Lyapunov变量 P_L^j 。

Step 4 给定一个足够小的阈值 σ ,若 $|\text{tr}(P_L^j - P_L^{j-1})| < \sigma$,输出的 $K_{opt} = K_{opt}^j$ 作为最优 H_∞ 鲁棒容错控制器增益,算法结束。否则,令 $j = j + 1$,返回Step 3。

算法的收敛性证明可以参看文献[20]。从迭代算法的步骤中可以看出,本文算法实质是通过两步优化方法实现了Lyapunov变量和控制增益的解耦,从而将非凸优化转化为凸优化。本文算法相比于扩展LMI算法,由于没有引入松弛变量,因此大大降低了LMI表达式的阶数,降低了LMI的求解难度,提高了运算效率;付出的代价是需要对低阶LMI(相对于扩展LMI算法)进行循环求解,但考虑到控制器的计算是离线的,因此这样的代价是可以接受的。

3 仿真实例

为了验证本文算法的有效性,以多操纵面高超声速飞行器X-33为对象,设计了飞行跟踪控制器。关于X-33模型的详细内容可以参考文献[21]。本文选取再入段做仿真研究,飞行器的系统状态 $x(t) = [\beta, \phi, \alpha, p, q, r]^T$,分别为侧滑角、滚转角、迎角、滚转角速率、俯仰角速率和偏航角速率;飞行器的输出 $y(t) = x(t)$;舵面输入 $u(t) = [\delta_1^r, \delta_1^l, \delta_2^r, \delta_2^l, \delta_3^r, \delta_3^l, \delta_4^r, \delta_4^l]^T$,分别为右侧内升降副翼、左侧内升降副翼、右侧襟翼、左侧襟翼、右侧方向舵、左侧方向舵、右侧外升降副翼和左侧外升降副翼。在控制器设计过程中,考虑2种故障,分别为右侧内升降副翼完全失效和右侧方向舵完全失效,即令相应的 $\pi_{Li} = 0$ 。考虑系统的不确定性,系统矩阵扰动40%。飞行器需要跟踪的指令信号 $r(t) = [\beta_{cmd}, \phi_{cmd}, \alpha_{cmd}]^T$;选择加权矩阵:

$$C_z = \begin{bmatrix} \text{diag}(4, 4, 4, 1, 1, 1) & 0_{6 \times 3} \\ 0_{8 \times 9} \end{bmatrix} \quad D_z = \begin{bmatrix} 0_{6 \times 8} \\ 2I_8 \end{bmatrix}$$

根据本文算法求出的鲁棒容错跟踪控制器如式(35)所示。为了验证本文算法的优越性,将其与所有故障状态均使用同一个 Lyapunov 变量的方法进行对比(也即本文算法的初值)。图1给出了不存在故障时系统动态跟踪响应,此时2种控制器均能够很好地跟踪参考信号。图2和图3分别给出了右侧内升降副翼和右侧方向舵失效故障时仿真结果,此时2种控制器仍然能够跟踪控

制指令,但很明显无迭代优化情况的超调量比较大,控制效果显然不如本文算法。图4给出了右侧内升降副翼和右侧方向舵同时失效故障时的仿真结果。可以看出,此时使用相同 Lyapunov 变量方法得到的控制器有较为明显的振荡,几乎无法跟踪指令,而本文算法的控制器虽然也有一些波动,但依然能够得到较好的稳定系统,这充分说明了本文算法可以有效削弱控制器的保守性。

$$K = \begin{bmatrix} -1.1025 & 1.3231 & 0.3518 & 1.2369 & 1.0544 & 1.5356 & 1.4994 & -0.7889 & -0.7287 \\ 6.4476 & -0.3740 & 0.1224 & -0.6272 & 0.7627 & -5.1900 & -2.9108 & 0.4121 & -0.4906 \\ 7.7972 & 2.1710 & 0.9338 & 1.1990 & 2.0122 & -4.8100 & -2.6162 & -1.2661 & -1.6125 \\ -7.9956 & -2.2693 & 1.0663 & -1.2243 & 0.6530 & 4.9235 & 2.7348 & 1.3680 & -1.1091 \\ -0.6207 & -0.0773 & -0.0248 & -0.0375 & -0.1429 & 0.4422 & 0.1989 & 0.0277 & 0.0878 \\ -0.1688 & 0.0011 & 0.0166 & 0.0141 & 0.0213 & 0.1298 & 0.0855 & -0.0019 & -0.0217 \\ -1.1025 & 1.3231 & 0.3518 & 1.2369 & 1.0544 & 1.5356 & 1.4994 & -0.7889 & -0.7287 \\ 6.4476 & -0.3740 & 0.1224 & -0.6272 & 0.7627 & -5.1900 & -2.9108 & 0.4121 & -0.4906 \end{bmatrix} \quad (35)$$

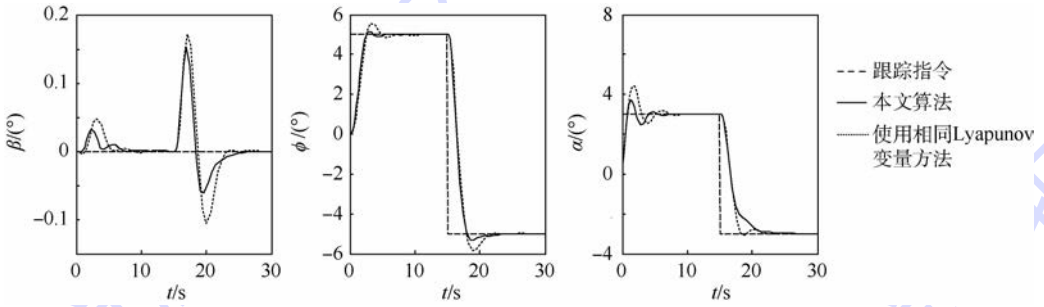


图1 无故障时飞行器指令响应曲线

Fig.1 Command response curves of vehicle without fault

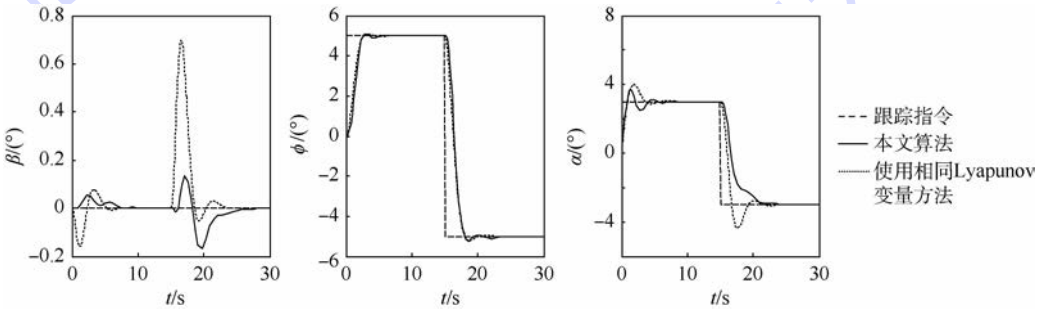


图2 右侧内升降副翼失效时飞行器指令响应曲线

Fig.2 Command response curves of vehicle under actuator outage of right inboard elevon

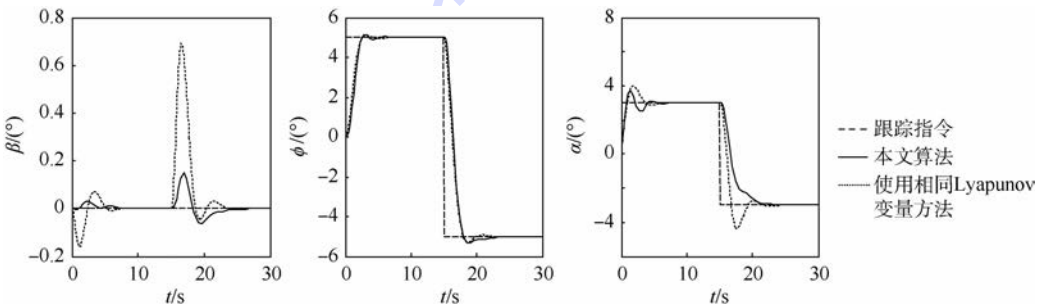


图3 右侧方向舵失效时飞行器指令响应曲线

Fig.3 Command response curves of vehicle under actuator outage of right rudder

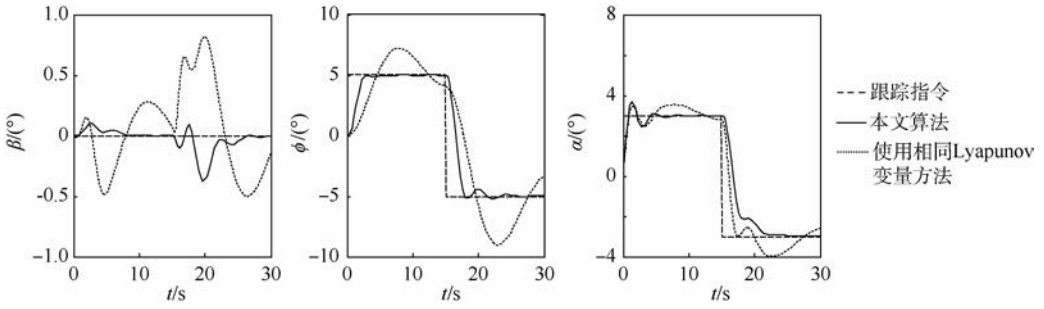


图4 右侧内升降副翼和右侧方向舵同时失效时飞行器指令响应曲线

Fig.4 Command response curves of vehicle under actuators outage of both right inboard elevon and right rudder

4 结论

本文研究了一种迭代 LMI 算法,用来解决鲁棒容错控制器设计过程中存在的非凸优化问题。通过在高超声速飞行器 X-33 上的仿真实验,得到了以下结论:

1) 所设计的控制器不仅可以保证对指令的跟踪,而且能够使得正常情况下的跟踪性能最优,在故障情况下也能够保证对指令的跟踪。

2) 与所有故障状态均使用同一个 Lyapunov 变量方法相比,本文算法可以降低控制器设计的保守性,提高系统的动态性能,并且闭环系统系统具有较好的鲁棒性,甚至在多故障情况下仍能够保持较好的跟踪性能,从而验证了本文算法的正确性。

参考文献 (References)

- [1] JI Y,ZHOU H L,ZONG Q. Adaptive active fault-tolerant control of generic hypersonic flight vehicles[J]. Journal of Systems and Control Engineering,2015,229(2):130-138.
- [2] ZHANG P,LIU H H T,LI X B,et al. Fault tolerance of cooperative interception using multiple flight vehicles[J]. Journal of the Franklin Institute,2013,350(9):2373-2395.
- [3] 傅强. 航空发动机被动容错控制系统鲁棒性设计[J]. 测控技术,2013,32(5):32-34.
FU Q. Robustness design of passive fault tolerant control system for aero engine[J]. Measurement & Control Technology, 2013,32(5):32-34(in Chinese).
- [4] 张爱华,胡庆雷,霍星,等. 过驱动航天器飞轮故障重构与姿态容错控制[J]. 宇航学报,2013,34(3):369-376.
ZHANG A H,HU Q L,HUO X,et al. Fault reconstruction and fault tolerant attitude control for over-activated spacecraft under reaction wheel failure[J]. Journal of Astronautics,2013,34(3):369-376(in Chinese).
- [5] CHEN F Y,WANG Z,TAO G,et al. Robust adaptive fault-tolerant control for hypersonic flight vehicles with multiple faults[J]. Journal of Aerospace Engineering,2015,28(4):04014111.
- [6] HE J J,QI R Y,JIANG B. Adaptive fault-tolerant control design for hypersonic flight vehicles based on feedback linearization[C]//33rd Chinese Control Conference. Piscataway, NJ: IEEE Press, 2014:3197-3202.
- [7] YANG G H,ZHANG S Y,LAM J,et al. Reliable control using redundant controllers[J]. IEEE Transactions on Automatic Control,1998,43(11):1588-1593.
- [8] PANZA S,LOVERA M. Rotor state feedback in helicopter flight control: Robustness and fault tolerance[C]//2014 IEEE Conference on Control Applications. Piscataway, NJ: IEEE Press, 2014:451-456.
- [9] 陈雪芹,耿云海,张迎春,等. 基于 LMI 的鲁棒容错控制及其在卫星姿态控制中的应用[J]. 控制理论与应用,2008,25(1):95-99.
CHEN X Q,GENG Y H,ZHANG Y C,et al. Robust fault-tolerant control based on LMI approach and application in satellite attitude control system[J]. Control Theory & Applications, 2008,25(1):95-99(in Chinese).
- [10] 陈明,童朝南. 不确定系统鲁棒容错控制的 LMI 设计方法[J]. 控制与决策,2009,24(4):526-531.
CHEN M,TONG C N. LMI approach to robust fault-tolerant H-infinity control for uncertain systems[J]. Control and Decision,2009,24(4):526-531(in Chinese).
- [11] 杨冬梅,孙俊娜. 不确定时滞线性离散系统的鲁棒容错控制[J]. 东北大学学报(自然科学版),2012,33(2):161-164.
YANG D M,SUN J N. Robust fault tolerant control for uncertain linear discrete systems with time-delay[J]. Journal of Northeastern University(Natural Science),2012,33(2):161-164(in Chinese).
- [12] YE S J,ZHANG Y M,WANG X M,et al. An improved LMI approach for static output feedback fault-tolerant control with application to flight tracking control[C]//4th IEEE Conference on Industrial Electronics and Applications. Piscataway, NJ: IEEE Press,2009:35-40.
- [13] 欧阳高翔,倪茂林,孙承启. 视故障为结构不确定项的鲁棒可靠跟踪控制器设计[J]. 控制理论与应用,2009,26(1):80-84.
OUYANG G X,NI M L,SUN C Q. Robust reliable tracking controller design when the fault is viewed as a structural uncertainty[J]. Control Theory & Applications,2009,26(1):80-84(in Chinese).
- [14] ZHANG Q J,YE S J,LI Y,et al. An enhanced LMI approach

- for mixed H_2/H_∞ flight tracking control[J]. Chinese Journal of Aeronautics, 2011, 24(3):324-328.
- [15] 王明昊,刘刚,杨述华.高超声速飞行器的多胞LPV系统控制器设计[J].空间控制技术与应用,2013,39(1):15-22.
WANG M H, LIU G, YANG S H. Polytopic-LPV-system-based control design for hypersonic vehicle[J]. Aerospace Control and Application, 2013, 39(1):15-22 (in Chinese).
- [16] XIE L. Output feedback H_∞ control of systems with parameter uncertainty[J]. International Journal of Control, 1996, 63(4):741-750.
- [17] 俞立,陈国定,杨马英.不确定系统具有圆盘区域极点约束的鲁棒控制[J].自动化学报,2000,26(1):116-120.
YU L, CHEN G D, YANG M Y. Robust control of uncertain linear system with disk pole constraints[J]. Acta Automatica Sinica, 2000, 26(1):116-120 (in Chinese).
- [18] HAN X D, LIU J G, XIE D X, et al. Robust H_∞ guaranteed cost satisfactory fault-tolerant control for discrete-time systems with quadratic D stabilizability[J]. Journal of Systems Engineering and Electronics, 2010, 21(3):496-502.
- [19] HE Y, WANG Q G. An improved ILMI method for static output feedback control with application to multivariable PID control[J]. IEEE Transactions on Automatic Control, 2006, 51(10):1678-1683.
- [20] CAO Y Y, LAM J, SUN Y X. Static output feedback stabilization: An ILMI approach[J]. Automatica, 1998, 34(12):1641-1645.
- [21] HOLLIS B R, THOMPSON R A, MURPHY K J, et al. X-33 aerodynamic computations and comparisons with wind-tunnel data[J]. Journal of Spacecraft and Rockets, 2001, 38(5):684-691.

作者简介:**彭程** 男,博士研究生。主要研究方向:容错飞行控制。

Tel.: 029-88431356

E-mail: wanli1818@163.com

王新民 男,硕士,教授,博士生导师。主要研究方向:飞行控制。

Tel.: 029-88431356

E-mail: wxmin@nwpu.edu.cn

谢蓉 女,博士,副教授。主要研究方向:飞行控制与控制理论。

E-mail: xierong@nwpu.edu.cn

穆凌霞 女,博士研究生。主要研究方向:高超声速控制。

E-mail: mulx@nwpu.edu.cn

Fault-tolerant control for hypersonic vehicle with system uncertainty

PENG Cheng, WANG Xinmin*, XIE Rong, MU Lingxia

(College of Automation, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: For the systems with norm-bounded parameter uncertainty and actuator failure faults, a robust fault-tolerant tracking control method is presented. In this method, the linear matrix inequalities (LMI) which guarantee the stability of the closed-loop system and meet the demand of the robust performance were derived utilizing the bounded real lemma. In order to reduce the conservativeness of the designed controller, the different Lyapunov variables were adopted under different system fault states. An iterative LMI algorithm is developed to solve the resulting non-convex optimization problem caused by the use of different Lyapunov variables. Since the convergence of the iterative algorithm lies on the choice of the initial values, an efficient way to find a good initial value was given after derivation. The method was applied to the design of the hypersonic vehicle X-33 tracking controller. Simulation results show that the proposed method is available and effective.

Key words: hypersonic vehicle; linear matrix inequality (LMI); fault-tolerant control; tracking control; robust control

Received: 2015-07-09; Accepted: 2015-09-25; Published online: 2015-10-30 17:09

URL: www.cnki.net/kcms/detail/11.2625.V.20151030.1709.012.html

Foundation item: Natural Science Foundation of Shaanxi Province (2014JQ8342)

* Corresponding author. Tel.: 029-88431356 E-mail: wxmin@nwpu.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0475

动态火力接入下要地防空作战目标威胁评估



刘敬蜀, 姜文志*, 雷宇曜, 代进进

(海军航空工程学院 兵器科学与技术系, 烟台 264001)

摘 要: 以动态火力接入下海军要地防空作战为背景, 综合考虑要地相对价值和目标攻击能力、攻击意图以及到达时间, 构建空袭目标威胁评估模型。首先, 阐述了动态火力接入下要地防空作战系统及其动态接入与退出策略; 然后, 结合要地相对价值和空袭目标数量以及毁伤能力构建了空中目标威胁力量量化模型; 其次, 基于目标相对要地的航路角构建了目标攻击意图量化模型, 并根据目标运动特性分别给出到达要地时间的计算方法; 最后, 根据目标距离防空体系拦截区的关系进行威胁等级排序, 并提出威胁度计算方法, 据此给出目标威胁等级排序。仿真结果表明, 本文方法符合实际作战中的一般战术推理, 能够客观地给出空袭目标威胁等级排序, 可为要地防空作战决策研究提供一定的参考。

关键词: 动态火力接入; 要地防空; 威胁评估; 要地相对价值; 威胁紧急程度; 威胁等级排序

中图分类号: V247; E273.4; TJ761.1+3

文献标识码: A **文章编号:** 1001-5965(2016)07-1422-10

海军要地大多地处沿海前沿地区, 要地内目标多、分布广, 容易遭到远程精确打击与隐身攻击、火力压制与电磁压制的综合攻击。有别于空军的国土防空和陆军的野战防空, 海军要地防空属于预设阵地防空, 在未来信息化作战样式下, 其面临的威胁将更加突出^[1]。为了适应未来海军要地防空作战的需求, 必须改变传统的防空作战样式。因此, 提出动态火力接入下的要地防空作战系统, 其是基于战术单位内, 利用计算机、通信与网络技术, 将地理上分布的各防空火力单元联网成统一高效的防空体系, 各防空火力单元可以随时接入, 也可以随时退出^[2]。

威胁评估就是综合考虑目标威胁的各影响因素, 合理选取威胁评估指标并对其进行量化, 进而建立威胁评估模型对目标威胁程度进行估计, 并给出目标威胁度排序^[3]。其作为要地防空指挥决策的一个重要环节, 可以利用战场态势信息推

断敌方各类空袭目标对我方的威胁程度, 是防空作战指挥过程中进行目标分配的前提条件^[4]。近年来, 国内外关于威胁评估的研究已经取得了许多成果, 主要集中在以下 2 个方面:

1) 威胁指标量化。威胁指标有定性指标和定量指标两大类。定性指标量化方法有 Miller 量化法^[5]、两级比例标度法^[6]和专家评分法^[7]等; 定量指标量化方法有区间量化法^[8]、模糊数表示法^[9]和区间数表示法^[10]等。威胁指标量化方法的选取应以充分体现其几何和物理意义为原则。

2) 威胁评估方法。主要包括定量计算和定性推理两大类: ①定量计算方法包括多属性决策理论^[11]、案例推理^[12]、时空推理^[13]和工作域方法^[14]等, 此类方法的特点是计算快速简单; ②定性推理方法包括模糊理论方法^[15]、贝叶斯网络^[16]、神经网络^[17]、专家系统^[18]和支持向量机^[19]等, 此类方法多需要引入专家知识构建规则

收稿日期: 2015-07-15; 录用日期: 2015-09-06; 网络出版时间: 2015-10-19 15:20

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151019.1520.006.html

* 通讯作者: Tel.: 0535-6635609 E-mail: ytjwz@sina.cn

引用格式: 刘敬蜀, 姜文志, 雷宇曜, 等. 动态火力接入下要地防空作战目标威胁评估[J]. 北京航空航天大学学报, 2016, 42(7): 1422-1431. LIU J S, JIANG W Z, LEI Y Y, et al. Threat evaluation of air-targets for key positions air-defense under dynamic fire access[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1422-1431 (in Chinese).

或推理网络,其主要特点是推理过程与人类思维相似。

总结上述方法,存在两方面突出的不足之处:

①威胁指标量化的过程中未能充分体现其几何和物理意义,并且未充分考虑到各威胁指标之间的耦合性;②多数威胁评估方法都需要确定指标权重系数,这必然要结合专家打分,从而引入一定的主观性和经验性。本文基于动态火力接入下海军要地防空目标威胁评估的特点,分别从威胁等级、威胁能力和威胁紧急程度 3 个方面进行量化分析,旨在寻找一种符合战术推理、尽量避免引入主观因素并易于工程实践的威胁评估方法。

1 动态火力接入下要地防空作战系统

动态火力接入下要地防空作战系统分为两级指挥系统:战术单位级和火力单元级。战术单位级可以同时动态接入多个火力单元,并对其进行协调指挥。海军要地防空作战的火力单元是指用于海军要地防空的可以独立作战的一组设备,其直接作战设备主要包括通信系统、指挥控制系统、探测跟踪系统、制导系统、发射装置、弹(丸)和干扰系统等^[20]。根据动态火力接入下要地防空作战系统的描述,基于战术单位集中指挥的防空作战系统组织结构如图 1 所示。

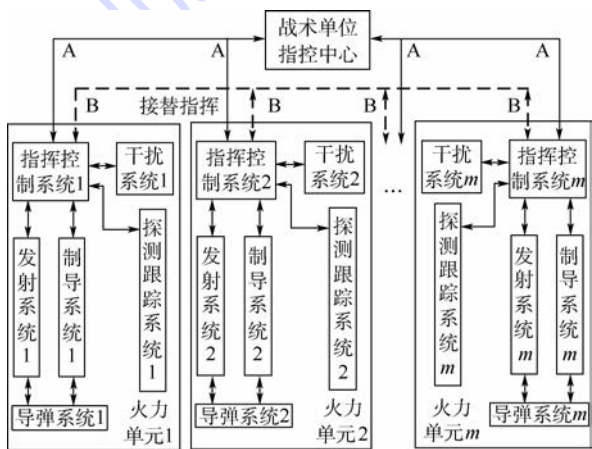


图 1 动态火力接入下战术单位防空作战系统组成结构
Fig.1 Composition structure of tactical unit air-defense operation system under dynamic fire access

当战术单位指挥控制系统完好时,接入的 m 个火力单元通过链路 A 在战术单位指挥控制中心组织下进行协同作战;当战术单位指挥控制系统战损时,指定某火力单元(如火力单元 1)进行接替指挥,接入的火力单元通过链路 B(图 1 中虚线所示)直接连通进行协同作战,战损的火力单元动

态退出作战系统,整个系统的运行不受任何影响,而修复好的火力单元可根据战术单位指控中心设定的网络通信协议动态接入作战系统,这里的动态接入与退出策略主要体现在整个作战系统在运行过程中,如果系统的功能需要重组,无需系统先停止运行,按照新的功能要求对系统进行重新配置,系统重新启动运行才能完成系统功能的重组。另外,动态火力接入下要地防空作战实现了组网作战方式,从而使得各火力单元的制导系统、探测跟踪系统和指挥控制系统也可以根据作战需求进行动态重组^[20]。

不难发现,在动态火力接入作战条件下,威胁评估过程也将发生变化,其结果会根据火力单元动态接入与退出的变化状态而实时发生变化。

2 要地防空威胁评估模型

由于威胁评估是在特定战场背景下进行的,因此,选取评估指标是应具有针对性的。对海军要地防空作战而言,其与舰艇防空以及空空作战最大区别在于:防空导弹阵地与被保护的要地对象在地理位置上是分离开的,并且个别要地前出于防御阵地,这均使得威胁评估研究内容发生了很大变化。本文将从以下 3 个方面进行威胁评估建模:威胁等级,即威胁到的要地的重要程度;空中目标对要地的威胁能力;威胁紧急程度。影响目标威胁度的因素如图 2 所示^[21]。

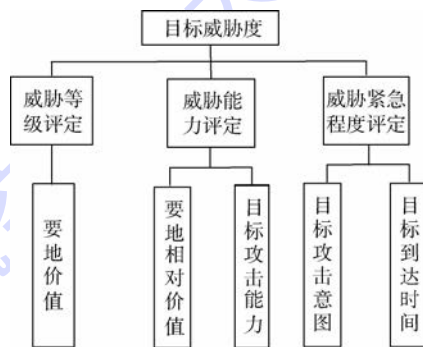


图 2 目标威胁度影响因素^[21]

Fig.2 Influencing factor of target threat extent^[21]

2.1 坐标系建立

假设被保护要地序列为 $K = \{K_1, K_2, \dots, K_m\}$, 来袭空中目标序列为 $T = \{T_1, T_2, \dots, T_n\}$, 我方动态接入 L 套防空导弹武器系统 F_1, F_2, \dots, F_L 对要地进行保护,接入的火力单元均接受战术单位 TU 的指挥,来袭目标威胁评估均由战术单位指控中心完成。以战术单位 TU 所在地理位置为原点 O , O_x 和 O_y 在水平面上,分别指向正北(N)和正东(E), O_z 由右手定则确定建立“北-天-东”坐标

系,如图3所示,本文约定角度值均以北偏东为正。图中: T_j 为目标,其坐标为 $(x_{T_j}, y_{T_j}, z_{T_j})$; T'_j 为目标在 Oxz 面的投影; v_{T_j} 为目标速度;探测系统能够提供的目标位置和速度信息分别为 $(d_{T_j}, \varepsilon_{T_j}, q_{T_j})$ 和 $(\dot{x}_{T_j}, \dot{y}_{T_j}, \dot{z}_{T_j})$; K_i 为要地,其位置信息为 $(x_{K_i}, 0, z_{K_i})$ 或 (d_{K_i}, q_{K_i}) 。战术单位指控中心通过融合各火力单元探测信息获得以上位置和速度信息,并由通信网传输给个火力单元指控中心。

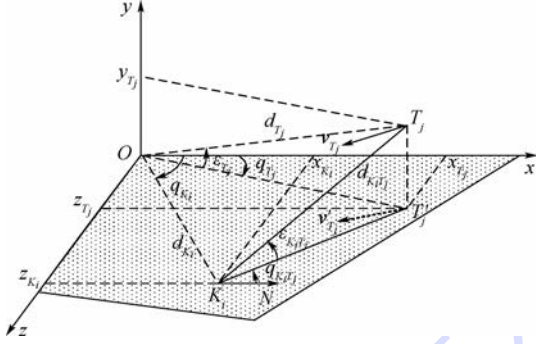


图3 统一坐标系

Fig. 3 Uniform coordinate system

要地防空作战中,防空导弹武器系统与受保护要地一般是分离的,而威胁估计是目标相对要地,因此,可以根据以上信息得到目标 T_j 相对要地 K_i 的距离 $d_{K_i T_j}$ 、俯仰角 $\varepsilon_{K_i T_j}$ 和方位角 $q_{K_i T_j}$ 分别为

$$d_{K_i T_j} = \sqrt{(x_{K_i} - x_{T_j})^2 + y_{T_j}^2 + (z_{K_i} - z_{T_j})^2} \quad (1)$$

$$\varepsilon_{K_i T_j} = \arcsin(y_{T_j}/d_{K_i T_j}) \quad (2)$$

$$q_{K_i T_j} = \arcsin \left[\frac{|z_{K_i} - z_{T_j}|}{\sqrt{(x_{K_i} - x_{T_j})^2 + (z_{K_i} - z_{T_j})^2}} \right] \quad (3)$$

2.2 威胁等级评定

海军要地是水面舰艇、潜艇和航空兵等海军主战兵力的驻屯地域,主要包括各级作战指挥中心、舰艇(潜艇)驻泊码头、防空导弹阵地、雷达站、航空兵机场和通信枢纽等,并且这些要地目标分布相近且比较集中,需要进行统一防空,从而构成了分布型区域要地目标。

由于构成分布型区域目标的各要地重要程度不同,威胁到各个不同要地的空中目标就形成了不同威胁等级,因此,需要对各个要地进行价值评定。选取要地的重要性值来评价要地的价值,其是指要地被击毁后在政治、经济和军事上可能造成影响的大小,通常由军事性、经济性、社会性、危险性、易损性和可恢复性等因素综合决定^[22],计算公式为

$$C_i = \ln(c_{i1} c_{i2} \cdots c_{is}) + 1 \quad (4)$$

式中: $c_{i'}$ 为评价第 i 个要地考虑的第 i' 个因素值,

其取值范围为 $c_{i'}$ $\in [1, 10]$; s 为评价因素的个数。

将 C_i 进行归一化处理得

$$C_i = \frac{\ln(c_{i1} c_{i2} \cdots c_{is}) + 1}{\ln(10^s) + 1} \quad (5)$$

则 C_i 的取值范围为 $C_i \in [0, 1]$ 。

显然,要地目标的重要性值越大,其重要程度就越高,则威胁到该要地目标的空中目标威胁等级也就越高。在实际作战中,各要地的价值量化一般由上级直接指定,以要地价值参数表的形式具体给出。

2.3 空中目标威胁能力评定

空中目标的威胁能力是指其突防后对防空导弹武器系统所保卫的要地可能造成的损失程度。对于分布型区域要地目标,要评定目标威胁能力主要考虑2个方面:一方面是要地相对价值,即各个要地价值占整个分布型区域总价值的比例;另一方面是目标攻击能力,即不同种类的空中目标对要地的作战能力。

2.3.1 要地相对价值

根据第2.1节计算各要地价值为 C_1, C_2, \dots, C_m , 则可得要地总价值为

$$C_S = \sum_{i=1}^m C_i \quad (6)$$

从而得到第 k 个要地的相对价值为

$$\delta(C_i) = C_i / C_S \quad (7)$$

显然, $\delta(C_i)$ 的值越大,该要地的价值越高,则空中目标对其威胁值也就越大。

2.3.2 目标攻击能力

1) 威胁目标类型

对于海军要地防空,其面临的主要空中威胁目标类型包括以下2类:

(1) 精确制导武器类。包括战术弹道导弹(Tactical Ballistic Missile, TBM)、巡航导弹(Cruise Missile, CM)、空地导弹(Air to Ground Missile, AGM)、反辐射导弹(Anti-Radiation Missile, ARM)和制导炸弹(Precision Guided Bomb, PGB)等。精确制导武器凭借精度高、射程远和速度大等诸多优点大大提高了空袭作战的效率,提高了飞机的生存能力,更好地达到了作战目的。

(2) 飞机类。又可分为两类:一类是硬杀伤为主的飞机,主要包括轰炸机(Bomber Aircraft, BA)、歼击轰炸机(Fighter-Bomber, FB)和武装直升机(Armed Helicopter, AH)等,用来对要地实施直接攻击;另一类是软杀伤为主的飞机,如预警指挥飞机、各种支援干扰飞机及战术与战略侦察机等,它们通常处于防空体系防区外,虽然不对要地

构成直接威胁,但是在它们的支援下会使得空袭效果倍增,但是由于在防区外,它们对要地的威胁远小于硬杀伤飞机,因此,本文只讨论硬杀伤飞机的威胁度。作为空袭作战任务最重要的武器载体和发射平台,不同类别飞机有不同的作战目的,这也决定了其机载武器也各不相同。一般的,轰炸机载武器包括巡航导弹、空地导弹和制导炸弹;歼击轰炸机载武器包括空地导弹、反辐射导弹和制导炸弹;武装直升机载武器包括空地导弹和制导炸弹。

2) 空中目标威胁值

空中目标威胁值是指在不受防空体系对抗条件下,对所攻击要地目标可能造成的毁伤程度^[23]。对于精确制导武器类,其主要影响因素包括武器数量和对所攻击要地的毁伤能力;对于飞机类,主要包括机载武器种类、数量以及各类武器对所攻击要地的毁伤能力。同时,对于分布式区

域要地防空,空中目标威胁值还必须计入要地的相对价值。为了计算空中目标威胁值,将要地能够承受的弹药数和各类空中目标对各要地的毁伤概率列于表1。

表1 要地承受弹药数和空中目标毁伤概率

Table 1 Key position acceptability ammo number and air-target killing probability

要地	承受弹药数/毁伤概率				
	TBM	CM	AGM	ARM	PGB
K_1	n_{TBM}^1/p_{TBM}^1	n_{CM}^1/p_{CM}^1	n_{AGM}^1/p_{AGM}^1	n_{ARM}^1/p_{ARM}^1	n_{PGB}^1/p_{PGB}^1
K_2	n_{TBM}^2/p_{TBM}^2	n_{CM}^2/p_{CM}^2	n_{AGM}^2/p_{AGM}^2	n_{ARM}^2/p_{ARM}^2	n_{PGB}^2/p_{PGB}^2
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
K_m	n_{TBM}^m/p_{TBM}^m	n_{CM}^m/p_{CM}^m	n_{AGM}^m/p_{AGM}^m	n_{ARM}^m/p_{ARM}^m	n_{PGB}^m/p_{PGB}^m

注:要地承受弹药数是综合考虑要地的面积、功能、抗毁程度以及目标的毁伤能力等因素获得的;而空中目标毁伤概率可通过情报资料查询或统计方法求得,二者不会对威胁评估有直接影响。

(1) 当来袭目标 T_j 为飞机类,其对要地 K_i 的威胁值为

$$b_{T_j K_i} = \begin{cases} \frac{1 - (1 - p_{CM}^i)^{M_{CM}^{BA}}}{n_{CM}^i} \cdot \frac{1 - (1 - p_{AGM}^i)^{M_{AGM}^{BA}}}{n_{AGM}^i} \cdot \frac{1 - (1 - p_{PGB}^i)^{M_{PGB}^{BA}}}{n_{PGB}^i} \delta(C_i) & T_j \in BA \\ \frac{1 - (1 - p_{AGM}^i)^{M_{AGM}^{FB}}}{n_{AGM}^i} \cdot \frac{1 - (1 - p_{ARM}^i)^{M_{ARM}^{FB}}}{n_{ARM}^i} \cdot \frac{1 - (1 - p_{PGB}^i)^{M_{PGB}^{FB}}}{n_{PGB}^i} \delta(C_i) & T_j \in FB \\ \frac{1 - (1 - p_{AGM}^i)^{M_{AGM}^{AH}}}{n_{AGM}^i} \cdot \frac{1 - (1 - p_{PGB}^i)^{M_{PGB}^{AH}}}{n_{PGB}^i} \delta(C_i) & T_j \in AH \end{cases} \quad (8)$$

式中: M 为一架飞机类目标一次攻击对要地投放的弹药数,其上标代表飞机类型,下标代表弹药类型。

(2) 当来袭目标 T_j 为精确制导武器类时,其对要地 K_i 的威胁值为

$$b_{T_j K_i} = \begin{cases} (p_{TBM}^i/n_{TBM}^i) \delta(C_i) & T_j \in TBM \\ (p_{CM}^i/n_{CM}^i) \delta(C_i) & T_j \in CM \\ (p_{AGM}^i/n_{AGM}^i) \delta(C_i) & T_j \in AGM \\ (p_{ARM}^i/n_{ARM}^i) \delta(C_i) & T_j \in ARM \\ (p_{PGB}^i/n_{PGB}^i) \delta(C_i) & T_j \in PGB \end{cases} \quad (9)$$

通过式(8)和式(9)不难发现,不管来袭空中目标是精确制导武器类还是飞机类,其对要地的毁伤概率越大,则其威胁值越大,并且对被攻击要地而言,其相对价值越大、可承受弹药数越小,则来袭武器对其威胁值也越大,这与实际作战情况是一致的。

2.4 威胁紧急程度评定

2.4.1 目标攻击意图

目标攻击意图是指目标对要地进行攻击的可能性大小的量度。影响空袭目标攻击意图的因素很多,本文考虑采用目标航路角作为量化攻击意图

的指标^[24]。目标的航路角是指目标速度在水平面的投影与敌我连线的夹角,如图4所示。图中: θ_{κ, T_j} 为目标相对要地的航路角; P_{CPA}^{κ, T_j} 为目标相对要地的航路捷径; v'_{T_j} 为目标速度向量在 Oxz 的投影。

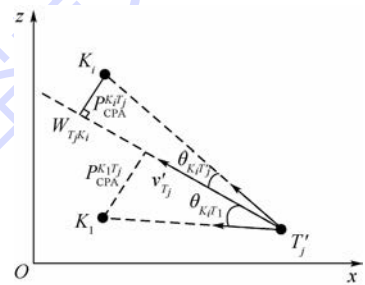


图4 目标与要地的位置关系

Fig. 4 Position relation between target and key position

$T_j K_i$ 为目标与要地连线向量,则可得二者夹角余弦为

$$\cos \theta_{\kappa, T_j} = \frac{v'_{T_j} \cdot T_j K_i}{|v'_{T_j}| \cdot |T_j K_i|} \quad (10)$$

且有

$$\begin{cases} |v'_{T_j}| = \sqrt{x_{T_j}^2 + z_{T_j}^2} \\ |T_j K_i| = \sqrt{(x_{T_j} - x_{K_i})^2 + (z_{T_j} - z_{K_i})^2} \end{cases} \quad (11)$$

所以可得

$$\cos \theta_{K_i T_j} = \frac{\dot{x}_{T_j}(x_{T_j} - x_{K_i}) + \dot{z}_{T_j}(z_{T_j} - z_{K_i})}{\sqrt{\dot{x}_{T_j}^2 + \dot{z}_{T_j}^2} \sqrt{(x_{T_j} - x_{K_i})^2 + (z_{T_j} - z_{K_i})^2}} \quad (12)$$

从而得到航路捷径为

$$P_{CPA}^{K_i T_j} = r_{K_i T_j} \sin \theta_{K_i T_j} \quad (13)$$

式中: $r_{K_i T_j}$ 为目标 T_j 在水平面投影点 T_j' 到要地 K_i 的距离。

目标接近要地的速度为

$$v_{T_j}'' = v_{T_j}' \cdot \cos \theta_{K_i T_j} \quad (14)$$

通常航路角 $|\theta_{K_i T_j}|$ 的取值范围为 $[0, \pi]$, 但是一般认为航路角在 $\pi/2$ 以内的目标对要地具有明显攻击意图。不难发现, 当 $|\theta_{K_i T_j}| \in [0, \pi/2]$ 时, 航路捷径与航路角成正比, 目标接近速度与航路角成反比, 即航路角越小时, 航路捷径越小, 目标接近速度越大, 目标攻击意图也就越大。因此, 航路角可综合速度和航路捷径的威胁效应作为量化攻击意图的指标。

1) 精确制导武器类目标

考虑到攻击意图随航路角的增大而递减的趋势, 选取负指数函数来量化目标 T_j 对要地 K_i 攻击意图系数为

$$a_{T_j K_i} = \begin{cases} e^{-|\theta_{T_j K_i}|} & |\theta_{T_j K_i}| \in \left[\arcsin \frac{r_{T_j K_i}'}{2R_{\min}^{T_j}}, \frac{\pi}{2} \right] \\ 0 & \text{其他} \end{cases} \quad (15)$$

式中: $r_{T_j K_i}'$ 为目标与要地连线水平距离; $R_{\min}^{T_j}$ 为目标 T_j 的最小水平转弯半径。

特别地, 对于反辐射导弹目标, 由于其具有向目标作连续径向运动特征, 探测跟踪网可以据此判断其攻击意图。因此, 当 T_j 为反辐射导弹时, 记探测跟踪网识别到 T_j 向要地序列 $\{K_{m_1}, K_{m_2}, \dots, K_{m_\lambda}\}$ ($m_i \in \{1, 2, \dots, m\}$) 作连续径向运动, 则 T_j 对 K_{m_i} 的攻击意图系数为

$$a_{T_j K_i} = \begin{cases} \delta(C_i) / \lambda & i \in \{m_1, m_2, \dots, m_\lambda\} \\ 0 & i \notin \{m_1, m_2, \dots, m_\lambda\} \end{cases} \quad (16)$$

2) 飞机类目标

对于飞机类目标, 其真正的威胁来自于机载武器, 而当前时刻探测跟踪网只能测得载机航路角, 因此, 其攻击意图系数实际上是根据载机航路角信息确定的机载武器的攻击意图。根据不同机型, 可以分别用行向量表示为

$$a_{T_j K_i} = \begin{cases} (a_{T_j K_i}^{CM} & a_{T_j K_i}^{AGM} & a_{T_j K_i}^{PCB}) & T_j \in BA \\ (a_{T_j K_i}^{AGM} & a_{T_j K_i}^{ARM} & a_{T_j K_i}^{PCB}) & T_j \in FB \\ (a_{T_j K_i}^{AGM} & a_{T_j K_i}^{PCB}) & & T_j \in AH \end{cases} \quad (17)$$

式中: 各个机载武器的攻击意图系数由式(15)确定。对于反辐射导弹, 由于此时无法确定其攻击对象, 因此也根据式(15)确定。

记目标 T_j 对各要地的意图为 $A_{T_j} = (a_{T_j K_1}, a_{T_j K_2}, \dots, a_{T_j K_m})$, 若 $\forall a_{T_j K_i} \in A_{T_j}$, 都有 $a_{T_j K_i} = 0$, 则有 $A_{T_j} = \mathbf{0}$, 即可判定目标 T_j 对各要地无攻击意图; 否则, $A_{T_j} \neq \mathbf{0}$ 时, 可判定目标 T_j 对各要地有攻击意图。

2.4.2 目标到达时间

1) 精确制导武器类目标

(A) 战术弹道导弹。对于战术弹道导弹目标, 一般认为其弹道为椭圆曲线, 因此, 到达时间可采用椭圆弹道理论计算。记 t_k 时刻, 探测跟踪网测得目标 T_j 的位置和速度分别为 $(x_{T_j}^k, y_{T_j}^k, z_{T_j}^k)$ 和 $(\dot{x}_{T_j}^k, \dot{y}_{T_j}^k, \dot{z}_{T_j}^k)$, 为了方便计算, 将目标参数由统一坐标系转换到地球直角坐标系 $O_d x_d y_d z_d$ 。

易知统一坐标系原点 O 在地球直角坐标系中的值为

$$\begin{cases} x_{Od} = (R_e + h_o) \cos \varphi_o \cos \lambda_o \\ y_{Od} = (R_e + h_o) \cos \varphi_o \sin \lambda_o \\ z_{Od} = (R_e + h_o) \sin \varphi_o \end{cases} \quad (18)$$

式中: R_e 为地球半径(本文取 6 371 km); φ_o, λ_o 和 h_o 分别为战术单位的经度、纬度和海拔高度。

目标相对地球直角坐标系的位置和速度分别为

$$\begin{cases} \begin{bmatrix} x_{T_j d}^k \\ y_{T_j d}^k \\ z_{T_j d}^k \end{bmatrix} = C_d \begin{bmatrix} x_{T_j}^k \\ y_{T_j}^k \\ z_{T_j}^k \end{bmatrix} + \begin{bmatrix} x_{Od} \\ y_{Od} \\ z_{Od} \end{bmatrix} \\ \begin{bmatrix} \dot{x}_{T_j d}^k \\ \dot{y}_{T_j d}^k \\ \dot{z}_{T_j d}^k \end{bmatrix} = C_d \begin{bmatrix} \dot{x}_{T_j}^k \\ \dot{y}_{T_j}^k \\ \dot{z}_{T_j}^k \end{bmatrix} \end{cases} \quad (19)$$

式中: C_d 为统一坐标系与地球直角坐标系的转换矩阵, 具体可参看文献[25]。

目标到地球球心的距离、目标高度和速度大小分别为

$$|r_{T_j d}^k| = \sqrt{(x_{T_j d}^k)^2 + (y_{T_j d}^k)^2 + (z_{T_j d}^k)^2} \quad (20)$$

$$h_{T_j}^k = |r_{T_j d}^k| - R_e \quad (21)$$

$$|v_{T_j d}^k| = \sqrt{(\dot{x}_{T_j d}^k)^2 + (\dot{y}_{T_j d}^k)^2 + (\dot{z}_{T_j d}^k)^2} \quad (22)$$

式中: $r_{T_j d}^k$ 为目标到地球球心构成的向量。则可得目标到达时间为^[25]

$$t_{T_i}^k = \mu^{-\frac{1}{2}} (\alpha_{T_j}^k)^{\frac{3}{2}} (E_l - e_{T_j}^k \sin E_l - E + e_{T_j}^k \sin E) \quad (23)$$

式中: μ 为与引力有关的常数, 取值为 $3.986 \times 10^{14} \text{ m}^3/\text{s}^2$ 。

$$\alpha_{T_j}^k = \frac{\mu |r_{T_{jd}}^k|}{2\mu - |r_{T_{jd}}^k| |v_{T_{jd}}^k|}$$

$$e_{T_j}^k = \left(1 - \frac{|r_{T_{jd}}^k \times v_{T_{jd}}^k|^2}{\mu \alpha_{T_j}^k} \right)^{\frac{1}{2}}$$

$$|r_{T_{jd}}^k \times v_{T_{jd}}^k|^2 = (x_{T_{jd}}^k \dot{y}_{T_{jd}}^k - y_{T_{jd}}^k \dot{x}_{T_{jd}}^k)^2 + (x_{T_{jd}}^k \dot{z}_{T_{jd}}^k - z_{T_{jd}}^k \dot{x}_{T_{jd}}^k)^2 + (y_{T_{jd}}^k \dot{z}_{T_{jd}}^k - z_{T_{jd}}^k \dot{y}_{T_{jd}}^k)^2$$

$$E_l = \arccos\left(\frac{\alpha_{T_j}^k - R_e}{\alpha_{T_j}^k e_{T_j}^k}\right)$$

$$E = \arccos\left(\frac{\alpha_{T_j}^k - |r_{T_{jd}}^k|}{\alpha_{T_j}^k e_{T_j}^k}\right)$$

(B) 反辐射导弹。由于反辐射导弹一般向目标作连续径向运动,因此,目标 T_j 到达要地 K_i 的时间为

$$t_{T_j K_i}^k = r_{T_j K_i}^k / |v_{T_j}^k| \quad (24)$$

式中: $r_{T_j K_i}^k$ 为 t_k 时刻目标到要地的距离; $|v_{T_j}^k|$ 为 t_k 时刻目标的速度值。

(C) 其他制导武器。对于巡航导弹和空地导弹,凡是向要地接近的导弹目标,都有航路捷径点(见图 4),记为 W_{ij} 点,对于不同航向的目标,其航路捷径点就不同,因此,可以选取目标到达航路捷径点 W_{ij} 的时间作为目标临近要地上空的快慢程度的指标,得到

$$t_{T_j K_i}^k = \frac{r_{T_j K_i}^k \cdot \cos \theta_{K_i T_j}^k}{|v_{T_j}^k|} \quad (25)$$

式中: $r_{T_j K_i}^k$ 为 t_k 时刻目标距离要地的水平距离; $|v_{T_j}^k|$ 为 t_k 时刻目标的水平速度。

特别地,对于制导炸弹,一般认为其在水平面运动为减速运动,因此,引入速度变化率 $\Delta v_{T_j}^k$,从而得到

$$t_{T_j K_i}^k = \frac{\sqrt{|v_{T_j}^k|^2 + 2\Delta v_{T_j}^k \cdot r_{T_j K_i}^k} - |v_{T_j}^k|}{\Delta v_{T_j}^k} \quad (26)$$

通过式(23)~式(26)可以发现,目标到达时间综合体现了目标到要地的距离和速度的威胁效应,这一指标体现了三者的相关性。

2) 飞机类目标

对于飞机类目标,目标到达时间分为两方面讨论:一方面从当前时刻 t_k 至目标到达投弹圈边界的时间。所谓投弹圈,是指空中目标向防空体系保卫的要地投放空袭武器的最大距离为半径的圆(见图 5)。显然,防空导弹体系必须在飞机类目标进入投弹圈以前将其杀伤,因为一旦其进入投弹圈再进行拦截,则防空效果很低,特别是对要地目标的保卫已不起太大作用;另一方面是机载武器从投放点到达要地的时间。

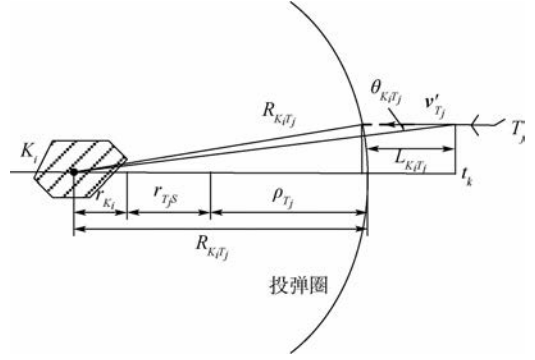


图 5 飞机类目标投弹圈示意图

Fig. 5 Aircraft target bombing circle sketch

图 5 中: r_{K_i} 为要地的平均半径; $r_{T_{jS}}$ 为飞机目标投放武器的杀伤半径; ρ_{T_j} 为飞机目标空袭武器的投放距离,其主要取决于飞机目标的飞行高度、速度以及投放的空袭武器的射程;则投弹圈半径为 $R_{K_i T_j} = r_{K_i} + r_{T_{jS}} + \rho_{T_j}$,从而 t_k 时刻目标 T_j 距离投弹圈的投放距离为

$$L_{K_i T_j}^k = r_{K_i T_j}^k \cos \theta_{K_i T_j}^k - \sqrt{(R_{K_i T_j})^2 - (r_{K_i T_j}^k \cdot \sin \theta_{K_i T_j}^k)^2} \quad (27)$$

目标到达投弹圈的时间为

$$t_{T_j C_j}^k = L_{T_j K_i}^k / |v_{T_j}^k| \quad (28)$$

对于同一个飞机类目标上的不同类型的机载武器,不同的射程决定了其投放距离的不同,从而决定了其不同的投放点,以行向量形式表示不同机载武器决定的飞机类目标到达投弹圈的时间,即

$$t_{T_j C_j}^k = \begin{cases} (t_{T_j C_j}^{CM}, t_{T_j C_j}^{AGM}, t_{T_j C_j}^{PGB}) & T_j \in BA \\ (t_{T_j C_j}^{ARM}, t_{T_j C_j}^{AGM}, t_{T_j C_j}^{PGB}) & T_j \in FB \\ (t_{T_j C_j}^{AGM}, t_{T_j C_j}^{PGB}) & T_j \in AH \end{cases} \quad (29)$$

对于不同机载武器,其从投放点到达要地时间可分别按照式(24)~式(26)计算,分别记为 $t_{C_j K_i}^{CM}$ 、 $t_{C_j K_i}^{AGM}$ 、 $t_{C_j K_i}^{ARM}$ 、 $t_{C_j K_i}^{PGB}$,从而得到目标到达时间为

$$t_{T_j K_i}^k = \begin{cases} (t_{T_j C_j}^{CM} + t_{C_j K_i}^{CM}, t_{T_j C_j}^{AGM} + t_{C_j K_i}^{AGM}, t_{T_j C_j}^{PGB} + t_{C_j K_i}^{PGB}) & T_j \in BA \\ (t_{T_j C_j}^{ARM} + t_{C_j K_i}^{ARM}, t_{T_j C_j}^{AGM} + t_{C_j K_i}^{AGM}, t_{T_j C_j}^{PGB} + t_{C_j K_i}^{PGB}) & T_j \in FB \\ (t_{T_j C_j}^{AGM} + t_{C_j K_i}^{AGM}, t_{T_j C_j}^{PGB} + t_{C_j K_i}^{PGB}) & T_j \in AH \end{cases} \quad (30)$$

为了方便比较,下面将目标到达时间进行归一化处理,记 $T_j \in BA$ 时, $t_{T_j K_i}^k = (t_{T_j K_i}^{\tau k})$, $\tau = 1, 2, 3$; $T_j \in FB$ 时, $t_{T_j K_i}^k = (t_{T_j K_i}^{\zeta k})$, $\zeta = 1, 2, 3$; $T_j \in AH$ 时, $t_{T_j K_i}^k = (t_{T_j K_i}^{\chi k})$, $\chi = 1, 2$, 则有

$$\begin{cases}
 t_{\max T_j K_i}^k = \max \{ \max(t_{T_j K_i}^k), \max(t_{T_j K_i}^{\tau k}), \max(t_{T_j K_i}^{\xi k}), \\
 \max(t_{T_j K_i}^{\chi k}) \} \\
 T_{T_j K_i}^k = 1 - \frac{t_{T_j K_i}^k}{t_{\max T_j K_i}^k} + \max \left\{ 1 - \max \left(1 - \frac{t_{T_j K_i}^k}{t_{\max T_j K_i}^k} \right) \right\} \\
 T_{T_j K_i}^{\tau k} = 1 - \frac{t_{T_j K_i}^{\tau k}}{t_{\max T_j K_i}^{\tau k}} + \max \left\{ 1 - \max \left(1 - \frac{t_{T_j K_i}^{\tau k}}{t_{\max T_j K_i}^{\tau k}} \right) \right\} \\
 T_{T_j K_i}^{\xi k} = 1 - \frac{t_{T_j K_i}^{\xi k}}{t_{\max T_j K_i}^{\xi k}} + \max \left\{ 1 - \max \left(1 - \frac{t_{T_j K_i}^{\xi k}}{t_{\max T_j K_i}^{\xi k}} \right) \right\} \\
 T_{T_j K_i}^{\chi k} = 1 - \frac{t_{T_j K_i}^{\chi k}}{t_{\max T_j K_i}^{\chi k}} + \max \left\{ 1 - \max \left(1 - \frac{t_{T_j K_i}^{\chi k}}{t_{\max T_j K_i}^{\chi k}} \right) \right\}
 \end{cases} \quad (31)$$

3 要地防空目标威胁等级排序

3.1 威胁等级划分

对于来袭目标的等级划分,一个关键性因素是目标的最大威判距离,其是指确保防空导弹与目标在杀伤区远界相遇的威胁判断决策距离^[26],则目标 T_j 相对火力单元 F_k 的最大威判距离为

$$D_{\max F_k}^{TE} = R_{\max F_k}^{\text{kill}} + t_{\Sigma} |v_{T_j}| \quad (32)$$

式中: $R_{\max F_k}^{\text{kill}}$ 为动态火力接入下火力单元 F_k 的杀伤区远界; $t_{\Sigma} = t_D + t_R + t_F$, t_D 为战术单位决策时

$$\text{Th}_{T_j}^k = \begin{cases}
 M_{\text{CM}_j}^{\text{BA}} \sum_{i=1}^m b_{T_j K_i}^{\text{CM}} a_{T_j K_i}^{\text{CM}} T_{T_j K_i}^{1k} + M_{\text{AGM}_j}^{\text{BA}} \sum_{i=1}^m b_{T_j K_i}^{\text{AGM}} a_{T_j K_i}^{\text{AGM}} T_{T_j K_i}^{2k} + M_{\text{PGB}_j}^{\text{BA}} \sum_{i=1}^m b_{T_j K_i}^{\text{PGB}} a_{T_j K_i}^{\text{PGB}} T_{T_j K_i}^{3k} & T_j \in \text{BA} \\
 M_{\text{AGM}_j}^{\text{FB}} \sum_{i=1}^m b_{T_j K_i}^{\text{AGM}} a_{T_j K_i}^{\text{AGM}} T_{T_j K_i}^{1k} + M_{\text{AGM}_j}^{\text{FB}} \sum_{i=1}^m b_{T_j K_i}^{\text{ARM}} a_{T_j K_i}^{\text{ARM}} T_{T_j K_i}^{2k} + M_{\text{PGB}_j}^{\text{FB}} \sum_{i=1}^m b_{T_j K_i}^{\text{PGB}} a_{T_j K_i}^{\text{PGB}} T_{T_j K_i}^{3k} & T_j \in \text{FB} \\
 M_{\text{AGM}_j}^{\text{AH}} \sum_{i=1}^m b_{T_j K_i}^{\text{AGM}} a_{T_j K_i}^{\text{AGM}} T_{T_j K_i}^{1k} + M_{\text{PGB}_j}^{\text{AH}} \sum_{i=1}^m b_{T_j K_i}^{\text{PGB}} a_{T_j K_i}^{\text{PGB}} T_{T_j K_i}^{3k} & T_j \in \text{AH} \\
 \sum_{i=1}^m b_{T_j K_i} a_{T_j K_i} T_{T_j K_i}^k & \text{其他}
 \end{cases} \quad (33)$$

根据 $\text{Th}_{T_j}^k$ 值的大小即可获得空中目标威胁度由高到低的排序。

4 实例仿真

为保护某海军要地 K_1 和 K_2 , 沿海岸线部署 3 个防空导弹火力单元 K_4 、 K_5 、 K_6 以及战术单位 K_3 , 构成序列 $K = \{K_1, K_2, K_3, K_4, K_5, K_6\}$, 如图 6 所示。

各要地的距离和方位信息分别为 $K_1(30, 60^\circ)$, $K_2(40, 135^\circ)$, $K_3(0, 0^\circ)$, $K_4(25, 20^\circ)$, $K_5(22, 110^\circ)$, $K_6(49, 150^\circ)$ (距离单位为 km), 战术单位指控中心雷达天线海拔高度为 45 m, 跟踪网最大跟踪距离 1.8ψ km, 战术单位决策时间

间, t_R 为火力单元系统反应时间, t_F 为防空导弹与目标遭遇的飞行时间。

对于已经进入防空体系杀伤区并对要地有攻击意图的空中目标, 应判定其较高的威胁等级, 从而有如下威胁等级划分:

1) I 级威胁。我方要地处于目标武器的攻击范围内, 目标具有明显攻击意图, 能够并且必须立即采取对抗措施的目标, 即满足条件 $\{ \exists k, \text{s. t.}, 0 < r_{T_j F_k} \leq D_{\max T_j F_k}^{\text{TE}}, \mathbf{A}_{T_j} \neq \mathbf{0} \}$, $r_{T_j F_k}$ 为目标相对于火力单元 F_k 的距离。

2) II 级威胁。我方要地处于目标武器的攻击范围内, 目标没有攻击意图, 能够采取对抗措施的目标, 即满足 $\{ \exists k, \text{s. t.}, 0 < r_{T_j F_k} \leq D_{\max T_j F_k}^{\text{TE}}, \mathbf{A}_{T_j} = \mathbf{0} \}$ 。

3) III 级威胁。对我方要地有近期威胁, 但是处于防区外, 我方无法采取对抗措施, 应对其进行监视的目标, 即满足条件 $\{ \forall k, r_{T_j F_k} > D_{\max T_j F_k}^{\text{TE}} \}$ 。

3.2 目标威胁度

由于只有 I 级威胁目标需要立即进行拦截, 因此只需对这类空中目标进行具体的威胁度排序即可。综合本文考虑的要地相对价值、目标攻击能力、目标攻击意图和目标到达时间几个因素之间的关联性, 以及它们对威胁度的影响, 它们之间适宜处理为相乘的关系。因此, T_j 对要地的威胁度为

$t_D = 2 \text{ s}$, 火力单元系统反应时间 $t_R = 10 \text{ s}$, 防空导弹的高空最大射程为 $12\varphi \text{ km}$, 低空最大射程为

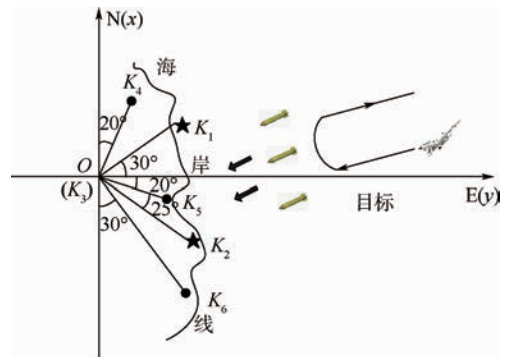


图6 要地分布示意图

Fig. 6 Key positions distribution sketch

5φ km,平均速度为 10ψ km/s (φ 、 ψ 为无量纲数量单位)。空袭方首先采用战术弹道导弹和巡航导弹在攻击机群到达前直接攻击各类指挥中心,然后攻击机群在离要地一定距离和高度上发射空地导弹、制导炸弹攻击保卫要地和防御阵地,并发射反辐射导弹探测跟踪设备进行攻击,而后低空返航, t_k 时刻探测跟踪网获得的空中目标信息如表 2 所示(相对战术单位指控中心 O)。

表 2 空袭目标信息
Table 2 Air attack target information

类型	斜距/ km	方位/ (°)	高度/ km	速度/ (m·s ⁻¹)	速度倾角/ (°)	武器 数量
TBM	70	90	50	1400	-120	
CM	40	20	0.02	340	-90	
AGM	60	50	5	340	-110	
ARM	100	135	10	680	-70	
PGB	80	150	15	340	-30	
BA	180	45	12	600	-80	[2,2,1]
FB	100	50	15	680	10	[2,1,1]

t_k 时刻各要地相对价值、可承受弹药数和空中目标毁伤概率如表 3 所示。

表 3 t_k 时刻要地相对价值、承受弹药数和空中目标毁伤概率

Table 3 Key positions relative value, acceptability ammo number and air-target killing probability at t_k

要地 $\delta(C_i)$	承受弹药数/毁伤概率					
	TBM	CM	AGM	ARM	PGB	
K_1	0.25	1/0.70	2/0.85	4/0.87	1/0.78	2/0.75
K_2	0.17	1/0.70	1/0.85	2/0.87	1/0.78	2/0.75
K_3	0.17	1/0.70	1/0.85	2/0.87	1/0.78	2/0.75
K_4	0.18	1/0.70	3/0.85	3/0.87	3/0.78	3/0.75
K_5	0.20	1/0.70	3/0.85	3/0.87	3/0.78	2/0.75

t_k 时刻目标对要地的威胁等级为

$$J = (I \quad I \quad I \quad I \quad I \quad III \quad II)$$

因此,只需计算 I 级目标的威胁度进行排序。根据目标态势信息,战术单位指控中心协调控制火力单元(要地) K_4 和 K_5 对来袭目标进行拦截,则 t_k 时刻目标对序列 $K = \{K_1, K_2, K_3, K_4, K_5\}$ 攻击能力为

$$B = \begin{bmatrix} 0.140 & 0.126 & 0.175 & 0.119 & 0.119 \\ 0.057 & 0.051 & 0.106 & 0.145 & 0.145 \\ 0.058 & 0.052 & 0.054 & 0.074 & 0.074 \\ 0.052 & 0.047 & 0.195 & 0.133 & 0.133 \\ 0.075 & 0.045 & 0.094 & 0.064 & 0.064 \end{bmatrix}$$

探测跟踪系统识别到反辐射导弹攻击火力单元(要地) K_5 , 则 t_k 时刻制导武器类目标对要地序列的攻击意图为

$$A_1 = \begin{bmatrix} 0.3391 & 0.8399 & 0.5925 & 0.3635 & 0.7971 \\ 0.0730 & 0.0582 & 0.2477 & 0.2477 & 0.0613 \\ 0.5924 & 0.3509 & 0.7053 & 0.9326 & 0.5152 \\ 0 & 0 & 0 & 0 & 0.1700 \\ 0.8850 & 0.7697 & 0.7697 & 0.9490 & 0.8697 \end{bmatrix}$$

t_k 时刻目标对要地的威胁度为

$Th_{T_j}^k = (0.3951 \quad 0.0659 \quad 0.0918 \quad 0.0133 \quad 0.0517)$
则目标威胁排序结果为 I 级威胁目标 $\{TBM > AGM > CM > PGB > ARM\}$, II 级威胁目标 $\{FB\}$, III 级威胁目标 $\{BA\}$, 即必须立即对 I 级制导武器类目标进行拦截,对歼击轰炸机目标和轰炸机目标继续进行监视。

火力单元(要地) K_5 受到反辐射导弹攻击,其探测制导设备静默,诱偏系统实现自防护,由于目标通道数有限,战术单位指控中心根据作战需要向火力单元(要地) K_6 发出动态接入命令,火力单元(要地) K_6 与战术单位指控中心动态建链,同时战术单位指控中心计算目标相对火力单元(要地) K_6 的攻击能力、攻击意图以及目标到达时间(归一化)为

$$b_{T,K_6} = (0.119 \quad 0.145 \quad 0.074 \quad 0.133 \quad 0.064)^T$$

$$a_{T,K_6} = (0.7053 \quad 0.1746 \quad 0.3897 \quad 0 \quad 0.5243)^T$$

$$t_{T,K_6} = (0.9478 \quad 0.2761 \quad 0.4180 \quad 1 \quad 1)^T$$

从而得到 t_{k+1} 时刻目标对要地序列 $K = \{K_1, K_2, K_3, K_4, K_5, K_6\}$ 的威胁度为

$$Th_{T_j}^{k+1} = (0.4647 \quad 0.1995 \quad 0.0983 \quad 0.1738 \quad 0.0739)$$

此时,巡航导弹和反辐射导弹的威胁度进一步增大,因此,需要尽快组织目标通道对其进行拦截。一方面,由于巡航多采用低空、超低空飞行,待探测跟踪系统发现巡航导弹时,留给防空系统的剩余拦截时间就非常有限;而反辐射导弹的径向打击,使得防空阵地雷达静默,大大降低了整个防空体系的防御能力。因而,应优先对其进行拦截,可以使被攻击火力单元再次动态接入防空体系进行拦截作战。

5 结 论

1) 本文基于要地相对价值和目标攻击能力、攻击意图以及到达时间,提出了一种空袭目标威胁评估方法,所得的目标威胁度能较好地体现目标对要地威胁的大小和紧急程度。

2) 目标威胁等级排序排除了防空火力区以外的目标,为后续目标分配环节减少了决策时间,增加了防空体系拦截时间。

3) 在量化威胁指标时,本文尽量从威胁因素对威胁度的影响机理出发,避免引入主观因素,评估结果更贴近实际作战情况,具有一定的工程应用价值。

参考文献 (References)

[1] 吴志飞,马曲立,饶世钧.空海一体战对海军要地防空作战的

- 影响及对策[J]. 海军大连舰艇学院学报, 2013, 36(4): 6-8.
- WU Z F, MA Q L, RAO S J. Influence of air-sea battle for protecting naval important place in air defense combat and countermeasures[J]. Journal of Dalian Naval Academy, 2013, 36(4): 6-8 (in Chinese).
- [2] 刘敬蜀, 姜文志, 代进进, 等. 动态火力接入下要地防空制导决策方法[J]. 海军航空工程学院学报, 2015, 30(5): 537-542.
- LIU J S, JIANG W Z, DAI J J, et al. Guidance-making method of essential air-defense under dynamic fire access[J]. Journal of Navel Aeronautical and Astronautical University, 2015, 30(5): 537-542 (in Chinese).
- [3] 代进进, 李相民. 舰艇编队网络化防空作战空袭目标威胁评估[J]. 系统工程与电子技术, 2013, 35(3): 532-538.
- DAI J J, LI X M. Threat evaluation of air strike targets for warship formation networked air-defense operation[J]. Journal of Systems Engineering and Electronics, 2013, 35(3): 532-538 (in Chinese).
- [4] 何友, 王国宏, 关欣, 等. 信息融合理论及应用[M]. 北京: 电子工业出版社, 2010: 393-395.
- HE Y, WANG G H, GUAN X, et al. Information fusion theory with applications[M]. Beijing: Publishing House of Electronics Industry, 2010: 393-395 (in Chinese).
- [5] MILLER G A. The magical number seven plus or minus two: Some limits on our capacity for processing information[J]. Psychological Review, 1956, 63(2): 81-97.
- [6] MACCRIMMON K R. Decision making among multiple-attribute alternatives: A survey and consolidated approach; RM-4823-ARPA[R]. Santa Monica: RAND Corporation, 1968.
- [7] 张鑫, 万新敏, 李争, 等. 运用 AHP 和云重心评价法的弹道导弹威胁评估[J]. 空军雷达学院学报, 2010, 24(5): 340-343.
- ZHANG X, WAN X M, LI Z, et al. Threat evaluation of ballistic missile by using AHP and cloud centroid method[J]. Journal of Air Force Radar Academy, 2010, 24(5): 340-343 (in Chinese).
- [8] 羊彦, 张继光, 景占荣, 等. 战术弹道导弹防御中的威胁评估算法[J]. 空军工程大学学报(自然科学版), 2008, 9(2): 31-35.
- YANG Y, ZHANG J G, JING Z R, et al. Research on an algorithm of threat assessment in tactics ballistic missile defense[J]. Journal of Air Force Engineer University (Natural Science Edition), 2008, 9(2): 31-35 (in Chinese).
- [9] 张磊, 童幼堂, 徐奕航. 舰艇编队空中目标威胁排序模型研究[J]. 舰船电子工程, 2009, 29(6): 136-138.
- ZHANG L, TONG Y T, XU Y H. Research on air target threatening order model of warship formation[J]. Ship Electronic Engineering, 2009, 29(6): 136-138 (in Chinese).
- [10] 郭辉, 徐浩军, 周莉. 基于区间数 TOPSIS 法的空袭目标威胁评估[J]. 空军工程大学学报(自然科学版), 2011, 12(1): 40-45.
- GUO H, XU H J, ZHOU L. Evaluation of air attack threat against target based on interval numbers TOPSIS[J]. Journal of Air Force Engineer University (Natural Science Edition), 2011, 12(1): 40-45 (in Chinese).
- [11] 程明, 周德云, 张堃. 基于混合型多属性决策方法的目标威胁评估[J]. 光电与控制, 2010, 17(1): 11-13.
- CHENG M, ZHOU D Y, ZHANG K. Threat assessment of target based on a hybrid multi-attribute decision-making method[J]. Electronics Optics & Control, 2010, 17(1): 11-13 (in Chinese).
- [12] 万开方, 高晓光, 刘宇, 等. 结合离差最大化的多属性群体决策 TOPSIS 威胁评估[J]. 火力与指挥控制, 2012, 37(8): 66-70.
- WAN K F, GAO X G, LIU Y, et al. TOPSIS threat assessment model based on the principle of maximum deviation and multi-attribute-group-decision-making[J]. Fire Control & Command Control, 2012, 37(8): 66-70 (in Chinese).
- [13] LOONEY C G, LIANG L R. Cognitive situation and threat assessment of ground battle spaces[J]. Information Fusion, 2003, 4(4): 297-308.
- [14] PETERSON G, AXELSSON L, JENSEN T, et al. Multi-source integration and temporal situation assessment in air combat [C]//Proceeding of Information, Decision and Control. Piscataway, NJ: IEEE Press, 1999: 371-375.
- [15] FENG L Y, XUE Q, LIU M X. Threat evaluation model of targets based on information entropy and fuzzy optimization theory [C]//2011 IEEE International Conference on Industrial Engineering and Engineering Management. Piscataway, NJ: IEEE Press, 2011: 1789-1793.
- [16] JOHANSSON F, FALKMAN G. A Bayesian network approach to threat evaluation with application to an air defense scenario [C]//2008 11th International Conference on Information Fusion. Piscataway, NJ: IEEE Press, 2008: 1-7.
- [17] FU L, WANG Q, XU J, et al. Target assignment and sorting for multi-target attack in multi-aircraft coordinated based on RBF [C]//2012 24th Chinese Control and Decision Conference. Xi'an: Northwestern Polytechnical University Press, 2012: 1935-1938.
- [18] 黄大荣, 郭安学, 李云生, 等. 基于专家知识属性重要度的集群目标威胁评估方法[J]. 兵工学报, 2009, 30(10): 1357-1362.
- HUANG D R, GUO A X, LI Y S, et al. An object group threat assessment method based on attribute significance of multi-field expert systems[J]. Acta Armamentarii, 2009, 30(10): 1357-1362 (in Chinese).
- [19] 郭辉, 徐浩军, 刘凌. 基于回归型支持向量机的空战目标威胁评估[J]. 北京航空航天大学学报, 2010, 36(1): 123-126.
- GUO H, XU H J, LIU L. Target threat assessment of air combat based on support vector machines for regression[J]. Journal of Beijing University of Aeronautics and Astronautics, 2010, 36(1): 123-126 (in Chinese).
- [20] 刘敬蜀, 姜文志, 刘涛, 等. 动态火力介入下要地防空作战系统研究[J]. 指挥控制与仿真, 2015, 37(5): 24-29.
- LIU J S, JIANG W Z, LIU T, et al. Research on operation system of key position air-defense under dynamic fire access[J]. Command Conetol & Simulation, 2015, 37(5): 24-29 (in Chinese).
- [21] 陈怀瑾. 防空导弹武器系统总体设计和试验[M]. 北京: 中国宇航出版社, 2009: 522-530.
- CHEN H J. Overall design and trial of air-defense missile weapon system[M]. Beijing: China Astronautic Publishing House,

2009;522-530(in Chinese).

- [22] 杨进佩,王俊,梁维泰.反导作战中的目标威胁排序方法研究[J].中国电子科学研究院学报,2012,7(4):432-436.

YANG J P, WANG J, LIANG W T. Research on method of the threaten queuing based on an-missile[J]. Journal of CAEIT, 2012,7(4):432-436(in Chinese).

- [23] 徐品高.空中目标威胁值的评定[J].战术导弹技术,2000(2):1-8.

XU P G. Threat value assessment for air target[J]. Tactical Missile Technology,2000(2):1-8(in Chinese).

- [24] 刘博.基于ANFIS的空袭目标攻击意图量化方法[J].指挥控制与仿真,2012,34(5):14-17.

LIU B. Quantitative method of targets attack intention based on ANFIS[J]. Command Control & Simulation,2012,34(5):14-17(in Chinese).

- [25] 温羨峤,刘谭军.反战术弹道导弹(ATBM)攻防对抗过程有关参数计算模型研究[J].现代防御技术,1996,(4):11-23.

WEN X Q, LIU T J. Computational model of parameters of attack-defense confrontation process for anti-tactical ballistic missile[J]. Modern Defense Technology, 1996,24(4):11-23(in Chinese).

- [26] 吴金平,陆铭华,黄文斌,等.水面舰艇编队对反舰导弹的威胁判断研究[J].系统仿真学报,2004,16(5):879-882.

WU J P, LU M H, HUANG W B, et al. Threat estimation re-

search of surface warship formation against anti-ship missiles

[J]. Journal of System Simulation,2004,16(5):879-882(in Chinese).

作者简介:

刘敬蜀 女,博士研究生。主要研究方向:作战指挥一体化技术。

Tel.: 18510992857

Email: liujingshux@163.com

姜文志 男,博士,教授,博士生导师。主要研究方向:作战指挥一体化技术。

Tel.: 0535-6635609

E-mail: ytjwz@sina.cn

雷宇曜 男,博士研究生。主要研究方向:作战指挥一体化技术。

Tel.: 15953599270

E-mail: 494511426@qq.com

代进进 男,博士,讲师。主要研究方向:作战指挥一体化技术。

Tel.: 13780953801

E-mail: dj8213@163.com

Threat evaluation of air-targets for key positions air-defense under dynamic fire access

LIU Jingshu, JIANG Wenzhi*, LEI Yuyao, DAI Jinjin

(Department of Ordnance Science and Technology, Naval Aeronautical and Astronautical University, Yantai 264001, China)

Abstract: Under the background of naval key position air-defense combat, air-targets threat evaluation model is established based on key positions relative value, target attack capability, target attack intent and target arrival time. Firstly, the key positions air-defense combat system under dynamic access is introduced, and the dynamic access and quit strategy are presented. Secondly, the key positions relative value, air-target number and target killing probabilities are used to quantize the air-target threat capability. Then, air-target attack quantization model is proposed with the target-to-key positions course angle, and the time that targets arrived at key positions is calculated based on different target types. Finally, the threat level order is carried out by comparing the distance between the targets and air-defense system intercept area, and the air-target threat value is obtained. The simulation results show that the quantized method conforms with the ordinary tactics reasoning, and it can develop the air-target threat order results objectively. This paper can provide a decision reference for the key positions defense operation.

Key words: dynamic fire access; key position air-defense; threat evaluation; key position relative value; threat urgency degree; threat level order

Received: 2015-07-15; Accepted: 2015-09-06; Published online: 2015-10-19 15:20

URL: www.cnki.net/kcms/detail/11.2625.V.20151019.1520.006.html

* Corresponding author. Tel.: 0535-6635609 E-mail: ytjwz@sina.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0441



可重复使用飞行器进场着陆拉平纵向控制

郝现伟¹, 王勇^{1,*}, 杨业², 郭涛², 张代兵³

(1. 北京航空航天大学 自动化科学与电气工程学院, 北京 100083; 2. 北京航天自动控制研究所, 北京 100854;
3. 国防科技大学 机电工程与自动化学院, 长沙 410073)

摘 要: 针对可重复使用飞行器(RLV)进场着陆拉平段的纵向控制问题,提出了反馈控制与前馈控制相结合的复合控制策略。设计反馈控制律参数,在此基础上基于时间加权高度跟踪误差/误差变化率平方积分指标优化设计前馈控制律参数。按输入补偿的前馈控制在不影响系统稳定性的前提下,提高了RLV对拉平着陆轨迹的跟踪精度,减小了RLV的接地散布。提出了基于积分器初值的控制律平滑切换方法,实现了RLV起落架释放前后不同控制律之间的平滑切换。仿真验证了拉平纵向复合控制和拉平过程中不同控制律之间平滑切换方法的有效性。

关键词: 可重复使用飞行器(RLV); 进场着陆; 前馈校正; 模型降阶; 控制律切换
中图分类号: V417+.6; V249.32

文献标识码: A **文章编号:** 1001-5965(2016)07-1432-09

可重复使用飞行器(Reusable Launch Vehicle, RLV)的无动力特性表明,RLV在进场着陆过程中不能复飞,仅能通过操纵相关的气动舵面来控制高度和动压,使RLV沿着期望的着陆轨迹实现安全着陆^[1]。整个进场着陆过程分为深下滑段、拉平段和地面滑跑段,其中拉平段是决定飞行器安全着陆的关键阶段。在拉平段,RLV跟踪轨迹由深下滑段的斜直线轨迹变为圆弧/指数形式的轨迹,这对轨迹跟踪控制性能提出了更高的要求,同时RLV的构型也将发生变化(为增升需要,襟翼偏角将下偏至某个位置,同时随着动压的减小,RLV起落架满足释放条件,起落架状态将发生变化),构型变化又会使RLV的气动不确定性增大,因此为保证着陆控制系统既具有较好的轨迹跟踪性能,又具有较强的鲁棒性,如何设计拉平纵向控制律成为决定RLV安全着陆的关键技术。

对于RLV飞行控制系统的设计,当前国内外

采用了非线性跟踪控制^[2]、神经网络自适应控制^[3]、滑模干扰观测器控制^[4]、模型参考变结构控制^[5]及轨迹线性化^[6]等非线性控制方法。但由于采用经典控制理论设计的控制器具有物理意义明确、结构简单及鲁棒性强等优点,因此目前工程上仍常采用经典控制理论设计各类飞行器控制律。文献[7]在拉平段采用PID控制,通过控制升降舵实现升降速度控制,进而实现了飞行器安全着陆,该控制方案常用于有动力飞行器着陆。无动力RLV在采用文献[7]的控制方案时,飞行器接地时刻的升降速度可满足接地要求,但缺点是接地点散布较大,易超出安全范围。为保证在复杂的风干扰或参数不确定情况下RLV仍具有较好的接地性能,在拉平段通过升降舵进行高度控制也是一种控制方案,但该控制方案的难点是RLV如何精确跟踪由圆弧拉起/指数拉起/浅下滑轨迹组成的拉平着陆轨迹^[8]。RLV拉平着陆

收稿日期: 2015-07-01; 录用日期: 2015-09-25; 网络出版时间: 2015-10-30 16:58

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151030.1658.009.html

基金项目: 总装预研重点基金(9140A25030109HK0103)

* 通讯作者: Tel.: 010-82317544 E-mail: wy_buaa@sina.com

引用格式: 郝现伟, 王勇, 杨业, 等. 可重复使用飞行器进场着陆拉平纵向控制[J]. 北京航空航天大学学报, 2016, 42(7): 1432-1440. HAO X W, WANG Y, YANG Y, et al. Flare longitudinal control for approach and landing of reusable launch vehicle[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1432-1440 (in Chinese).

轨迹的高度给定指令可视为加速度输入函数,由于采用PID控制的高度跟踪控制系统为II型系统,因而系统存在加速度静差,为此可通过提高系统的型别来消除系统加速度稳态静差。文献[9-10]研究了弹射座椅和导弹的飞行控制系统,采用鲁棒伺服线性二次型调节器(Linear Quadratic Regulator, LQR)理论在外回路增加积分环节提高了系统的型别,实现了控制系统对输入指令的精确跟踪,该方法的难点是加权矩阵 Q 、 R 的选取依赖于设计者的经验,缺点是易导致系统稳定性恶化。除在反馈回路增加积分环节提高系统的型别外,还可通过引入前馈校正装置来提高系统的跟踪精度。文献[11]采用经典控制理论在反馈系统中引入前馈校正装置实现了火炮系统对输入指令的精确跟踪,具有较强的工程实用性。

针对无动力RLV的拉平纵向控制问题,本文基于经典控制理论中前馈控制思想,提出一种按输入补偿的高度跟踪复合控制律设计方法。首先基于线性模型设计反馈回路控制律参数,然后采用平衡截断法^[12]对由姿态回路构成的等效飞行器传递函数进行模型降阶,最后基于时间加权高度跟踪误差/误差变化率平方积分性能指标设计前馈校正环节参数。对于起落架释放前后2组控制律的切换问题,提出一种基于积分器初值的控制律平滑切换方法。通过算例验证控制策略与控制结构的合理性,以及控制律设计和控制律平滑切换方法的有效性。

1 RLV纵向数学模型

RLV进场着陆为无动力着陆,仅通过操纵相关的气动舵面跟踪期望的着陆轨迹,在规定的接地状态范围内实现安全着陆。根据理论力学中刚体质心运动及转动运动定理,RLV的纵向动力学模型^[13]为

$$\begin{cases} \dot{V} = -Q/m - g \sin \theta \\ \dot{\theta} = \frac{Y}{mV} - \frac{g \cos \theta}{V} \\ \dot{\omega}_z = M_z/I_z \\ \dot{\vartheta} = \omega_z \\ \dot{\alpha} = \dot{\vartheta} - \dot{\theta} \\ \dot{h} = V \sin \theta \\ ma_{y_{cg}} = Q \sin \alpha + Y \cos \alpha - mg \cos \vartheta \\ a_y = a_{y_{cg}} + x_a \dot{\omega}_z \end{cases} \quad (1)$$

式中: h 、 V 、 θ 、 α 、 ϑ 和 ω_z 分别为高度、速度、轨迹倾角、迎角、俯仰角和俯仰角速率; x_a 为加速度计安装在飞行器对称面内距重心的距离,置于重心前

时 x_a 定义为正,反之为负; $a_{y_{cg}}$ 和 a_y 分别为飞行器重心处的法向加速度和加速度计安装处的法向加速度; m 和 g 分别为飞行器质量和重力加速度; I_z 为俯仰转动惯量; Q 、 Y 和 M_z 分别为阻力、升力和俯仰力矩,其值通常取决于飞行器的速度、迎角及各气动舵面偏角,可用以下形式表述:

$$\begin{cases} Q \approx q S C_x(Ma, \alpha, \delta) \\ Y \approx q S C_y(Ma, \alpha, \delta) \\ M_z \approx q S b_A m_z(Ma, \alpha, \omega_z, \delta) \\ C_x(Ma, \alpha, \delta) = C_{x0} + C_x^{Ma} Ma + C_x^\alpha \alpha + C_x^\delta \delta \\ C_y(Ma, \alpha, \delta) = C_{y0} + C_y^{Ma} Ma + C_y^\alpha \alpha + C_y^\delta \delta \\ m_z(Ma, \alpha, \omega_z, \delta) = m_{z0} + m_z^{Ma} Ma + m_z^\alpha \alpha + m_z^{\omega_z} \omega_z + m_z^\delta \delta \end{cases} \quad (2)$$

其中: $\delta = [\delta_e, \delta_{ca}, \delta_{sb}]^T$, δ_e 、 δ_{ca} 和 δ_{sb} 分别为升降舵、襟翼和阻力板; $q = \rho V^2/2$ 为动压, ρ 为空气密度; Ma 、 S 和 b_A 分别为马赫数、参考面积和平均气动弦长; C_x 、 C_y 和 m_z 分别为阻力系数、升力系数和俯仰力矩系数; C_{x0} 、 C_{y0} 和 m_{z0} 分别为基准点阻力系数、升力系数和俯仰力矩系数; C_x^* 、 C_y^* 和 m_z^* 为气动导数,*为式(2)中相关参数。

为便于分析问题,在基准点为铅垂面内固定轨迹倾角恒定动压飞行的条件下,对非线性系统式(1)进行小扰动线性化,考虑 θ 、 ϑ 和 α 之间的线性关系 $\theta = \vartheta - \alpha$,得到RLV的线性化模型:

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{A} \mathbf{x} + \mathbf{B} \mathbf{u} \\ \mathbf{y} = \mathbf{C} \mathbf{x} + \mathbf{D} \mathbf{u} \end{cases} \quad (3)$$

式中: $\mathbf{x} = [V, \alpha, \omega_z, \vartheta, h]^T$ 为自由变量; $\mathbf{u} = [\delta_e, \delta_{ca}, \delta_{sb}]^T$ 为输入变量; $\mathbf{y} = [V, \alpha, \omega_z, \vartheta, h, a_y]^T$ 为输出变量。

$$\mathbf{A} = \begin{bmatrix} x^v & x^\alpha + g \cos \theta_* & 0 & -g \cos \theta_* & 0 \\ -y^v & -y^\alpha + \frac{g \sin \theta_*}{V_*} & 1 - y^{\omega_z} & -\frac{g \sin \theta_*}{V_*} & 0 \\ \mu_z^v & \mu_z^\alpha & \mu_z^{\omega_z} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \sin \theta_* & -V_* \cos \theta_* & 0 & V_* \cos \theta_* & 0 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} x^{\delta_e} & x^{\delta_{ca}} & x^{\delta_{sb}} \\ -y^{\delta_e} & -y^{\delta_{ca}} & -y^{\delta_{sb}} \\ \mu_z^{\delta_e} & \mu_z^{\delta_{ca}} & \mu_z^{\delta_{sb}} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{C} = \begin{bmatrix} I_{5 \times 5} \\ C_{1 \times 5} \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} \mathbf{0}_{5 \times 3} \\ D_{1 \times 3} \end{bmatrix}$$

其中: θ_* 、 V_* 为基准运动对应的轨迹倾角、空速;

$I_{5 \times 5}$ 为单位矩阵; $0_{5 \times 3}$ 为零矩阵。

$$C_{1 \times 5} = [c_{61} \quad c_{62} \quad 0 \quad g \sin \alpha_* \quad 0]$$

$$D_{1 \times 3} = [d_{61} \quad d_{62} \quad d_{63}]$$

式中:

$$c_{61} = x_a (\mu_z^v - y^v) + \frac{Q^v \sin \alpha_* + Y^v \cos \alpha_*}{m}$$

$$c_{62} = x_a \mu_z^\alpha + \frac{(Q + Y^\alpha) \cos \alpha_* + (Q^\alpha - Y) \sin \alpha_*}{m}$$

$$d_{61} = x_a (\mu_z^{\delta_c} - y^{\delta_c}) + \frac{Q^{\delta_c} \sin \alpha_* + Y^{\delta_c} \cos \alpha_*}{m}$$

$$d_{62} = x_a (\mu_z^{\delta_{ca}} - y^{\delta_{ca}}) + \frac{Q^{\delta_{ca}} \sin \alpha_* + Y^{\delta_{ca}} \cos \alpha_*}{m}$$

$$d_{63} = x_a (\mu_z^{\delta_{sb}} - y^{\delta_{sb}}) + \frac{Q^{\delta_{sb}} \sin \alpha_* + Y^{\delta_{sb}} \cos \alpha_*}{m}$$

其中: α_* 为基准运动对应的迎角; 参数 $x^v, x^\alpha, x^{\delta_*}, y^v, y^\alpha, y^{\omega_z}, y^{\delta_*}, \mu_z^v, \mu_z^\alpha, \mu_z^{\omega_z}, \mu_z^{\delta_*}, Q^v, Q^\alpha, Q^{\delta_*}, Y^v, Y^\alpha$ 和 Y^{δ_*} 的具体表述形式见文献[13], δ_* 为相应的气动舵面。

2 拉平纵向控制策略与控制结构

RLV 进入拉平段后, 以飞行器高度与期望着陆轨迹高度之差为引导信号, 将高度跟踪误差输入给着陆控制系统, 通过控制升降舵偏转实现高度跟踪控制。在常规控制方案中, 升降舵通道由姿态控制内回路和高度跟踪控制外回路组成。其中, 姿态控制内回路由俯仰角反馈、俯仰角速度反馈和法向加速度反馈构成, 用于得到升降舵偏角指令, 通过控制升降舵偏转使飞行器产生俯仰运动, 改变迎角, 从而改变升力使飞行器产生升降运动, 实现高度跟踪控制; 高度跟踪控制外回路由高度反馈构成 PID 控制, 用于得到内回路所需要的俯仰角给定指令。由于常规控制方案中升降舵通道的系统为 II 型系统, 对于深下滑段, 高度给定指令为斜坡输入信号的情况下, 高度跟踪静差为零, 但对于拉平段, 此时高度给定指令变为拉平段的圆弧和指数信号, 将致使飞行器无法及时跟踪上期望拉平轨迹, 因此将产生较大的高度跟踪误差。为此, RLV 进入拉平段后, 在常规控制方案中高度跟踪控制外回路的基础上引入前馈校正装置, 相当于引入输入量的微分信号, 根据轨迹变化对俯仰角给定指令进行超前修正, 然后经过内回路控制律改变飞行器升降舵偏角指令, 使得飞行器获得抬头力矩具有一定的超前量, 从而有利于消除高次轨迹指令下的高度跟踪误差。

另外, 襟翼和阻力板的舵面偏转对飞行器安全着陆也起到关键作用。襟翼偏转用于增加飞行

器的升力, 有利于减小拉平高度跟踪误差, 同时还便于使飞行器的接地状态满足性能指标。阻力板偏转用于控制飞行器动压, 为轨迹跟踪控制提供动压保障。

RLV 拉平纵向控制律结构如图 1 所示。图中: H_g, V_{bg} 和 ϑ_g 分别为高度给定指令、表速给定指令和俯仰角给定指令; V_b 为表速; K_{ph}, K_{ih} 和 K_{dh} 为高度跟踪控制外回路 PID 控制比例系数; $K_\vartheta, K_{\omega_z}$ 和 K_{a_y} 为姿态控制内回路的比例系数。图 1 中, 由于 RLV 具有纵向静不稳定性特性, 引入法向加速度 a_y 反馈用以增加 RLV 的纵向静稳定性; ϑ 反馈和 ω_z 反馈用于改善 RLV 的长周期和短周期模态动态性能; 对高度进行 PID 控制和输入补偿前馈校正, 实现飞行器在加速度输入信号情况下高度跟踪静差为零。对于襟翼舵面, 为增升需要, 襟翼偏角将由拉平起始点偏角 δ_{ca0} 随时间 t 匀速下偏至规定偏角 δ_{ca1} 。阻力板通道为表速控制, 采用 PI 控制形式实现表速无静差跟踪控制。

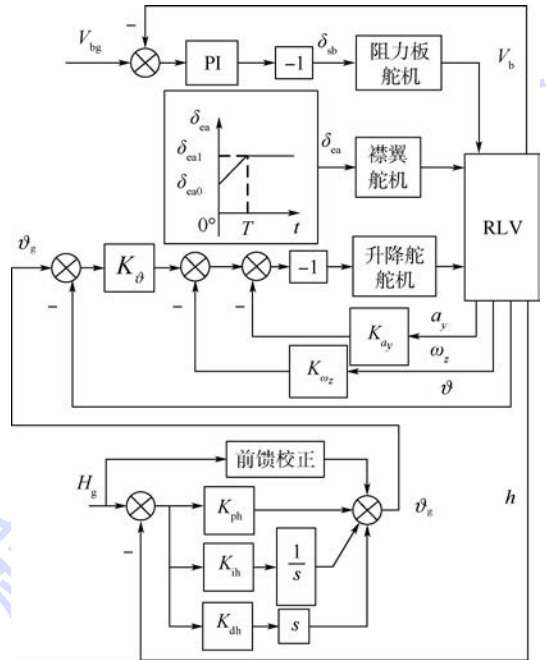


图 1 RLV 拉平纵向控制律结构

Fig. 1 Flare longitudinal control law structure of RLV

拉平纵向各通道的控制律表达式为

$$\vartheta_g = (K_{ph} + K_{ih}/s + K_{dh}s)(H_g - h) + G_f(s)H_g \quad (4)$$

$$\delta_c = - [K_\vartheta (\vartheta_g - \vartheta) - K_{\omega_z} \omega_z - K_{a_y} a_y] \quad (5)$$

$$\delta_{ca} = \begin{cases} \delta_{ca0} & t \leq T_{lp0} \\ \frac{\delta_{ca1} - \delta_{ca0}}{T_{rh}}(t - T_{lp0}) + \delta_{ca0} & T_{lp0} < t \leq T_{lp0} + T_{rh} \\ \delta_{ca1} & T_{lp0} + T_{rh} < t \end{cases} \quad (6)$$

$$\delta_{sb} = (K_{pv} + K_{iv}/s)(V_{bg} - V_b) \quad (7)$$

式中: s 为拉普拉斯算子; $G_f(s)$ 为前馈校正环节; K_{pv} 和 K_{iv} 为表速回路PI控制比例系数; T_{lp0} 为RLV进入拉平段的时刻; T_{th} 为襟翼偏角软化时间。

3 拉平纵向控制律参数设计

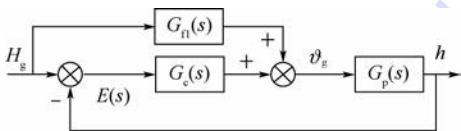
RLV拉平纵向控制律参数可归纳为2类参数:一类是反馈控制律参数(如 K_{ph} 、 K_{ih} 、 K_{dh} 、 K_{θ} 、 K_{ω_z} 、 K_{α_y} 、 K_{pv} 和 K_{iv});另一类是前馈校正环节参数。在控制律设计过程中,先将反馈控制律参数设计问题转换为多参数多约束优化设计问题,再利用MATLAB提供的SISO Design Tool工具箱^[14]对反馈控制律参数在可能的取值范围内寻优。根据这种设计方法,首先优化设计升降舵通道内回路反馈控制律参数,然后优化设计外回路反馈控制律参数,其次基于时间加权高度跟踪误差/误差变化率平方积分指标优化设计前馈校正环节参数,最后在升降舵通道控制律参数基础上再优化设计阻力板通道控制律参数。本文重点讨论如何优化设计升降舵通道前馈校正环节参数,使复合系统既能消除加速度静差,又能改善系统的动态跟踪性能。

将图1中升降舵回路控制结构简化为图2所示的2种形式。图中: $G_p(s)$ 为输入是俯仰角给定指令、输出是飞行高度的传递函数,由姿态控制内回路传递函数与飞行器传递函数组成; $G_c(s)$ 为校正装置传递函数; $G_{f1}(s)$ 和 $G_{f2}(s)$ 为前馈校正环节; $E(s)$ 为误差传递函数。

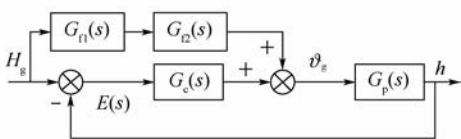
等效飞行器传递函数 $G_p(s)$ 的一般表达式如式(8)所示,传递函数 $G_c(s)$ 、 $G_{f1}(s)$ 和 $G_{f2}(s)$ 的表达式如式(9)和式(10)所示。

$$G_p(s) = \frac{b_{n-1}s^{n-1} + b_{n-2}s^{n-2} + \dots + b_1s + b_0}{s(a_n s^n + a_{n-1}s^{n-1} + \dots + a_1s + a_0)} \quad (8)$$

$$G_c(s) = (K_{dh}s^2 + K_{ph}s + K_{ih})/s \quad (9)$$



(a) 前馈控制中无动态校正环节



(b) 前馈控制中有动态校正环节

图2 升降舵回路控制结构简化图

Fig. 2 Simplified graph of elevator loop control structure

$$G_{f1}(s) = K_{qk}s \quad G_{f2}(s) = \frac{T_2s + 1}{T_1s + 1} \quad (10)$$

式中: a_n, a_{n-1}, \dots, a_0 和 $b_{n-1}, b_{n-2}, \dots, b_0$ 为 $G_p(s)$ 的系数; K_{qk} 为静态前馈系数; T_1 和 T_2 为动态前馈控制器时间常数。 $G_{f1}(s)$ 的作用是使复合系统等效为更高型别的系统,使系统加速度误差为零; $G_{f2}(s)$ 的作用是提高复合系统的动态响应性能。

对于升阻比大、飞行速度较小和着陆轨迹较缓的常规无动力飞行器,采用图2(a)所示前馈控制中无动态校正环节的控制结构,即可实现对拉平着陆轨迹的跟踪控制。但对于升阻比小、飞行速度较大和着陆轨迹较陡的RLV,若仍采用图2(a)所示的控制结构,可能无法快速地跟踪上期望轨迹,进而致使RLV无法安全着陆;若采用图2(b)所示的前馈控制中有动态校正环节的控制结构,则能改善系统的动态特性,在能消除系统加速度误差的同时,还能快速减小RLV的轨迹跟踪误差。

设计参数 K_{qk} 、 T_1 和 T_2 时,首先设计参数 K_{qk} ,然后在此基础上再设计参数 T_1 和 T_2 。

设计参数 K_{qk} 时,令 $G_{f2}(s) = 1$,则前馈校正环节为 $G_f(s) = K_{qk}s$ 。

复合系统的闭环传递函数 $\Phi(s)$ 为

$$\Phi(s) = \frac{(G_{f1}(s) + G_c(s))G_p(s)}{1 + G_c(s)G_p(s)} = \frac{[(K_{dh}s^2 + K_{ph}s + K_{ih}) + K_{qk}s^2](b_{n-1}s^{n-1} + b_{n-2}s^{n-2} + \dots + b_0)}{\Omega_1(s)} \quad (11)$$

式中:

$$\Omega_1(s) = s^2(a_n s^n + a_{n-1}s^{n-1} + \dots + a_0) + (K_{dh}s^2 + K_{ph}s + K_{ih})(b_{n-1}s^{n-1} + b_{n-2}s^{n-2} + \dots + b_0)$$

复合系统的误差传递函数为

$$\Phi_e(s) = \frac{1 - G_{f1}(s)G_p(s)}{1 + G_c(s)G_p(s)} = \frac{\{s^2[a_n s^n + \dots + (a_1 - K_{qk}b_1)s + (a_0 - K_{qk}b_0)]\}}{\Omega_1(s)} \quad (12)$$

取 $K_{qk} = a_0/b_0$ 。当输入信号为 $r(t) = t^2/2$ 时,利用终值定理推导可得系统的稳态误差为

$$e_{ss} = \lim_{s \rightarrow 0} E(s) = \lim_{s \rightarrow 0} \Phi_e(s)R(s) = 0 \quad (13)$$

式中: $R(s)$ 为输入传递函数。

可见,复合系统的加速度稳态误差为零,复合系统等效为Ⅲ型系统。

设计完参数 K_{qk} 后,再设计参数 T_1 和 T_2 。前馈校正环节变为

$$G_f(s) = K_{qk}s \frac{T_2s + 1}{T_1s + 1} \quad (14)$$

复合系统的闭环传递函数变为

$$\Phi(s) = \{ [(K_{ih}s^2 + K_{ph}s + K_{ih})(T_1s + 1) + K_{qk}s^2(T_2s + 1)] / (T_1s + 1) \} \cdot [(b_{n-1}s^{n-1} + b_{n-2}s^{n-2} + \dots + b_0) / \Omega_1(s)] \quad (15)$$

可见,引入前馈校正环节 $G_r(s)$ 后的闭环系统特征方程增加了 1 个闭环极点 $-1/T_1$ 。 T_1 的值直接影响系统的控制性能,为尽量减小闭环极点 $-1/T_1$ 对系统稳定性的影响,根据以下原则选取 T_1 的值:

- 1) T_1 大于零,以保证闭环极点 $-1/T_1$ 位于左半 s 平面,从而对系统的稳定性没有影响。
- 2) T_1 尽量小,使闭环极点 $-1/T_1$ 远离虚轴,从而降低 $-1/T_1$ 对系统动态性能的影响。
- 3) T_1 不宜过小,需结合飞控计算机的运行周期确定。

设计完参数 K_{qk} 和 T_1 后,再基于时间加权高度跟踪误差 $e(t)$ /误差变化率 $\dot{e}(t)$ 平方积分指标设计参数 T_2 。

时间加权高度跟踪误差 $e(t)$ /误差变化率 $\dot{e}(t)$ 平方积分指标形式如下:

$$J = \lambda_1 \int_0^\infty te^2(t) dt + \lambda_2 \int_0^\infty t\dot{e}^2(t) dt \quad (16)$$

式中: λ_1 和 λ_2 为加权系数。

性能指标 J 总为正值, J 越小表明系统的动态跟踪性能越好。性能指标 J 中引入误差变化率 $\dot{e}(t)$ 有利于限制过渡过程中误差的起伏变化,提高系统的动态跟踪性能。时间 t 的加权,有利于降低动态误差的持续时间。

复合系统的误差传递函数为

$$\overline{\Phi}_e(s) = \frac{1 - G_{fl}(s)G_{l2}(s)G_p(s)}{1 + G_c(s)G_p(s)} \quad (17)$$

将式(8)~式(10)代入式(17),整理为一般形式,即

$$\overline{\Phi}_e(s) = \frac{s^3(\overline{b_{n+1}}s^{n+1} + \overline{b_n}s^n + \dots + \overline{b_1}s + \overline{b_0})}{a_{n+3}s^{n+3} + a_{n+2}s^{n+2} + \dots + a_1s + a_0}$$

$$E(s) (\overline{a_{n+3}}s^{n+3} + \overline{a_{n+2}}s^{n+2} + \dots + \overline{a_1}s + \overline{a_0}) = s^3(\overline{b_{n+1}}s^{n+1} + \overline{b_n}s^n + \dots + \overline{b_1}s + \overline{b_0})R(s) \quad (18)$$

式中: $\overline{a_{n+3}}, \overline{a_{n+2}}, \dots, \overline{a_0}$ 和 $\overline{b_{n+1}}, \overline{b_n}, \dots, \overline{b_0}$ 为 $\overline{\Phi}_e(s)$ 的系数。

系统误差函数 $e(t)$ 的微分方程为

$$\overline{a_{n+3}}e^{(n+3)} + \overline{a_{n+2}}e^{(n+2)} + \dots + \overline{a_1}\dot{e} + \overline{a_0}e = \overline{b_{n+1}}r^{(n+4)} + \overline{b_{n-1}}r^{(n+3)} + \dots + \overline{b_1}r^{(4)} + \overline{b_0}\ddot{r} \quad (19)$$

将加速度输入信号 $r(t) = t^2/2$ 代入式(19),得

$$\overline{a_{n+3}}e^{(n+3)} + \overline{a_{n+2}}e^{(n+2)} + \dots + \overline{a_1}\dot{e} + \overline{a_0}e = 0 \quad (20)$$

依次将 $e, \dot{e}, \dots, e^{(n+2)}$ 在式(20)左右两边相乘,得

$$\begin{cases} \overline{a_{n+3}}e^{(n+3)}e + \overline{a_{n+2}}e^{(n+2)}e + \dots + \overline{a_1}\dot{e}e + \overline{a_0}ee = 0 \\ \overline{a_{n+3}}e^{(n+3)}\dot{e} + \overline{a_{n+2}}e^{(n+2)}\dot{e} + \dots + \overline{a_1}\dot{e}\dot{e} + \overline{a_0}e\dot{e} = 0 \\ \vdots \\ \overline{a_{n+3}}e^{(n+3)}e^{(n+2)} + \overline{a_{n+2}}e^{(n+2)}e^{(n+2)} + \dots + \overline{a_1}\dot{e}e^{(n+2)} + \overline{a_0}ee^{(n+2)} = 0 \end{cases} \quad (21)$$

在区间 $[0, \infty)$ 上分别对等式(21)两边积分,为书写方便,令

$$I_{i,j} = \int_0^\infty e^{(i)}e^{(j)} dt \quad i, j = 0, 1, \dots, n+3 \quad (22)$$

$$\begin{cases} \overline{a_{n+3}}I_{n+3,0} + \overline{a_{n+2}}I_{n+2,0} + \dots + \overline{a_1}I_{1,0} + \overline{a_0}I_{0,0} = 0 \\ \overline{a_{n+3}}I_{n+3,1} + \overline{a_{n+2}}I_{n+2,1} + \dots + \overline{a_1}I_{1,1} + \overline{a_0}I_{0,1} = 0 \\ \vdots \\ \overline{a_{n+3}}I_{n+3,n+2} + \overline{a_{n+2}}I_{n+2,n+2} + \dots + \overline{a_1}I_{1,n+2} + \overline{a_0}I_{0,n+2} = 0 \end{cases} \quad (23)$$

对于 $I_{i,j}$,当 $i \neq j$ 时, $I_{i,j}$ 通过分部积分法化简,可用 $I_{l,l}$ 初值 $e^{(l)}(0)$ 和终值 $e^{(l)}(\infty)$ 表示,其中, $l = 0, 1, \dots, n+2$ 。

由于系统为稳定系统,则终值 $e^{(l)}(\infty) = 0$ 。

初值 $e^{(l)}(0)$ 可由拉普拉斯变换微分公式(24)和初值定理式(25)求得:

$$E^{(l)}(s) = s^l E(s) - \sum_{n=0}^{l-1} s^{(l-1-n)} e^{(n)}(0) \quad (24)$$

$$e^{(l)}(0) = \lim_{s \rightarrow \infty} s E^{(l)}(s) \quad (25)$$

将式(23)整理为矩阵和向量的形式,即

$$\mathbf{H}(T_2)\mathbf{I} = \mathbf{b}(T_2) \quad (26)$$

式中: $\mathbf{I} = [I_{n+2,n+2}, I_{n+1,n+1}, \dots, I_{1,1}, I_{0,0}]^T$; \mathbf{H} 和 \mathbf{b} 分别为关于参数 T_2 的矩阵和向量。

整理式(26),可得到 $I_{l,l} = \int_0^\infty (e^{(l)})^2 dt$ 关于参数 T_2 的表达式。

同理,依次将 $te, t\dot{e}, \dots, te^{(n+2)}$ 在式(20)左右两边相乘,得

$$\begin{cases} \overline{a_{n+3}}te^{(n+3)}e + \overline{a_{n+2}}te^{(n+2)}e + \dots + \overline{a_1}t\dot{e}e + \overline{a_0}tee = 0 \\ \overline{a_{n+3}}te^{(n+3)}\dot{e} + \overline{a_{n+2}}te^{(n+2)}\dot{e} + \dots + \overline{a_1}t\dot{e}\dot{e} + \overline{a_0}te\dot{e} = 0 \\ \vdots \\ \overline{a_{n+3}}te^{(n+3)}e^{(n+2)} + \overline{a_{n+2}}te^{(n+2)}e^{(n+2)} + \dots + \overline{a_1}t\dot{e}e^{(n+2)} + \overline{a_0}tee^{(n+2)} = 0 \end{cases} \quad (27)$$

在区间 $[0, \infty)$ 上分别对等式(27)左右两边

积分,为书写方便,令

$$P_{i,j} = \int_0^{\infty} t e^{(i)} e^{(j)} dt \quad i, j = 0, 1, \dots, n+3 \quad (28)$$

则

$$\begin{cases} \overline{a_{n+3}} P_{n+3,0} + \overline{a_{n+2}} P_{n+2,0} + \dots + \overline{a_1} P_{1,0} + \overline{a_0} P_{0,0} = 0 \\ \overline{a_{n+3}} P_{n+3,1} + \overline{a_{n+2}} P_{n+2,1} + \dots + \overline{a_1} P_{1,1} + \overline{a_0} P_{0,1} = 0 \\ \vdots \\ \overline{a_{n+3}} P_{n+3,n+2} + \overline{a_{n+2}} P_{n+2,n+2} + \dots + \overline{a_1} P_{1,n+2} + \overline{a_0} P_{0,n+2} = 0 \end{cases} \quad (29)$$

对于 $P_{i,j}$, 当 $i \neq j$ 时, $P_{i,j}$ 通过分部积分法化简, 可用 $I_{l,l}$ 、 $P_{l,l}$ 和初值 $e^{(l)}(0)$ 表示, 其中, $l = 0, 1, \dots, n+2$ 。

将式(29)整理成矩阵和向量的形式:

$$\mathbf{K}(T_2)\mathbf{P} = \mathbf{d}(T_2) \quad (30)$$

式中: $\mathbf{P} = [P_{n+2,n+2}, P_{n+1,n+1}, \dots, P_{1,1}, P_{0,0}]^T$; \mathbf{K} 和 \mathbf{d} 分别为关于参数 T_2 的矩阵和向量。

通过整理式(30), 得到 $P_{l,l} = \int_0^{\infty} t(e^{(l)})^2 dt$ 关于参数 T_2 的表达式。

联立 $P_{0,0} = \int_0^{\infty} t e^2(t) dt$ 和 $P_{1,1} = \int_0^{\infty} t e^2(t) dt$, 可得性能指标 J 关于 T_2 的函数。使性能指标 J 关于 T_2 的导数 $dJ/dT_2 = 0$, 求得 T_2 。此时, 系统的动态性能达到最优。

综上所述, 等效系统 $G_p(s)$ 的阶数越高, 则误差传递函数 $\overline{\Phi}_e(s)$ 的阶数越高, 参数 T_2 的设计过程越复杂。为降低参数 T_2 的设计难度, 采用平衡截断法先对传递函数 $G_p(s)$ 进行模型降阶, 然后再基于新的等效系统设计参数 T_2 。

4 控制律平滑切换

RLV 进入拉平段后, 襟翼偏角开始随时间线性偏转至某规定位置, 当 RLV 满足起落架释放条件后, 起落架开始放下。襟翼偏角变化和起落架释放称为构型变化, 构型变化对飞行器气动有较大影响。为保证构型变化后的着陆控制系统仍具有较好的控制性能, 在控制结构不变的情况下, 需重新设计控制律参数。

起落架释放前后升降舵通道控制律参数的不同, 必然面临着 2 组控制律之间的切换问题。控制律之间的切换可能会引起升降舵舵面瞬间跳变, 从而产生较大的俯仰力矩, 致使飞行器姿态剧烈变化, 因此需采用抑制措施, 避免舵面瞬变。目前, 常采用双模态同步运算瞬变抑制法或单模态

运算瞬变抑制法^[15]在一定时间内对 2 组控制律采用淡入/淡出的方式抑制舵面瞬变, 但这 2 种方法均依赖设计者经验, 通过调节特定的参数达到较好的控制效果。为避免上述方法经验试凑, 提高切换效率, 以下基于积分器初值提出一种新的控制律平滑切换方法。

根据控制律中积分器的工作原理, 飞控计算机在控制律切换后的第 1 个运算周期, 积分器输出值为其初值 $x_{\text{Ini}0}$, 在之后的运算周期, 积分器输出值为初值 $x_{\text{Ini}0}$ 累加积分器输入值得到的值。根据积分器的该特性, 在控制律切换后的第 1 个运算周期, 可通过对新控制律中积分器赋初值, 从而使 2 组控制律得到的升降舵偏角指令相等。

对于图 1 所示的控制结构, 通过对高度跟踪控制外回路 PID 控制中积分器进行赋初值 $x_{\text{Ini}0}$, 以实现控制律平滑切换。由于控制律切换前一个运行周期的舵面偏角 $\delta_{\text{ec}1}$ 已知, 飞行器的飞行状态已知, 新的控制律参数已知, 采用从后向前推导的原则, 推导出积分器初值 $x_{\text{Ini}0}$ 的表达式为

$$x_{\text{Ini}0} = \vartheta_g - K_{\text{ph}}(H_g - h) - K_{\text{dh}}(\dot{H}_g - \dot{h}) \quad (31)$$

式中: $\vartheta_g = (-\delta_{\text{ec}1} + K_{a_y} a_y + K_{\omega_z} \omega_z) / K_{\vartheta} + \vartheta$ 。

赋入积分器初值后, 便可实现 2 组控制律之间的平滑切换。

5 仿真算例

以某 RLV 为研究对象, 按本文方法进行拉平纵向控制律设计。对 RLV 进行气动特性分析, 确定在拉平起始点襟翼舵面偏角 $\delta_{\text{ea}0}$ 为 0° , 进入拉平后襟翼偏角 $\delta_{\text{ea}1}$ 保持在 15° , 襟翼偏角软化时间 T_{th} 为 5 s。要求 RLV 的接地状态满足: 飞行器接地地点距理想接地点纵向偏差不得超过 ± 200 m; 升降速度小于 3 m/s; 俯仰角在 $3^\circ \sim 12^\circ$ 范围内。

根据 RLV 起落架状态变化过程, 选取拉平起始点为起落架收起状态时的控制律设计点, 选取起落架释放起始点距地面的中间高度点为起落架放下状态时的控制律设计点。起落架收起和放下状态控制律设计点对应的 RLV 飞行状态分别为: $q = 10\ 142$ Pa, $h = 423$ m, $\alpha = 3.1^\circ$, $\theta = -21^\circ$, $\omega_z = 0^\circ/\text{s}$, $\delta_e = -4.3^\circ$, $\delta_{\text{ea}} = 0^\circ$, $\delta_{\text{sb}} = 35^\circ$; $q = 6\ 724$ Pa, $h = 216$ m, $\alpha = 8.6^\circ$, $\theta = -9.2^\circ$, $\omega_z = 6.2^\circ/\text{s}$, $\delta_e = -6.5^\circ$, $\delta_{\text{ea}} = 15^\circ$, $\delta_{\text{sb}} = 17^\circ$ 。

在设计点处建立飞行器线性化模型, 设计控制律参数, 使 RLV 线性控制系统的时域/频域性能满足控制目标, 并使 RLV 的接地状态接近理想接地状态。其中, RLV 理想接地状态为: 接地纵向偏差

为 0 m, 升降速度为 1.5 m/s, 俯仰角为 7.5°。

根据第 3 节中方法设计拉平纵向控制律, 性能指标式 (16) 中的加权系数 λ_1 和 λ_2 分别取 0.6 和 0.4。控制律参数如下所示:

1) 起落架释放前。 $K_{ph} = 1.14, K_{ih} = 0.33, K_{dh} = 1.25, K_{\theta} = 1.81, K_{\omega_z} = 1.21, K_{a_y} = 0.71, K_{pv} = 8.2, K_{iv} = 0.75, K_{qk} = 0.42, T_1 = 0.5, T_2 = 6$ 。

2) 起落架释放后。 $K_{ph} = 1.43, K_{ih} = 0.26, K_{dh} = 1.88, K_{\theta} = 2.74, K_{\omega_z} = 1.71, K_{a_y} = 0.7, K_{pv} = 8.2, K_{iv} = 0.75, K_{qk} = 0.53, T_1 = 0.5, T_2 = 0.3$ 。

起落架开始释放时, 根据第 3 节中方法实现起落架释放前后不同控制律之间的平滑切换。

为检验本文提出的控制律设计方法的合理性和有效性, 以下分别采用图 2(a) 所示前馈控制中无动态校正环节和图 2(b) 所示前馈控制中有动态校正环节 2 种控制结构进行非线性全量仿真, 记录并画出 RLV 的飞行高度、高度跟踪误差 (正误差表示期望高度大于实际飞行高度)、升降速度、俯仰角和升降舵响应曲线, 见图 3 ~ 图 7, 图中横坐标 x_d 为 RLV 距理想接地点的距离。

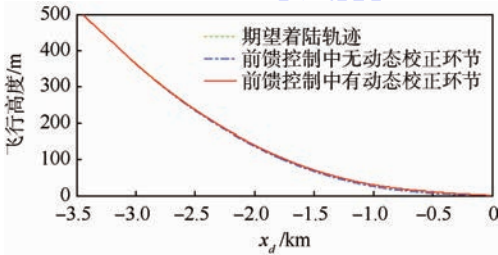


图 3 飞行高度响应曲线

Fig. 3 Flight altitude response curves

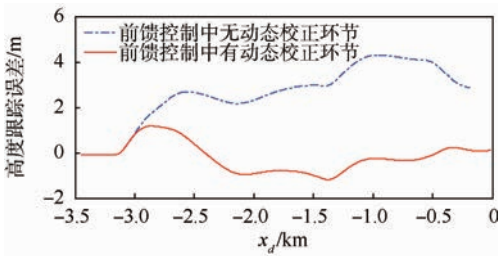


图 4 高度跟踪误差响应曲线

Fig. 4 Altitude tracking error response curves

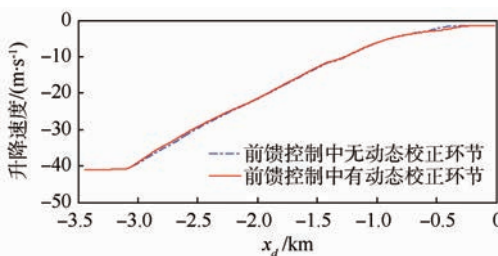


图 5 升降速度响应曲线

Fig. 5 Vertical speed response curves

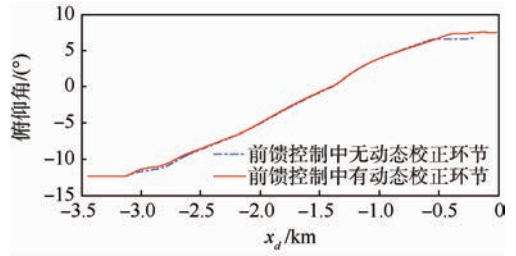


图 6 俯仰角响应曲线

Fig. 6 Pitch angle response curves

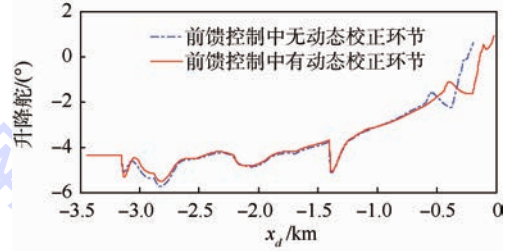


图 7 升降舵响应曲线

Fig. 7 Elevator response curves

由图 3 ~ 图 7 可见, 采用图 2(a) 所示的控制结构时, 虽然 RLV 接地时刻的升降速度 (1.3 m/s) 和俯仰角 (6.7°) 均落在要求的范围内, 但 RLV 对期望着陆轨迹的跟踪精度较差 (最大高度跟踪误差为 4.3 m), 致使 RLV 的实际接地点距理想接地点的纵向偏差较大 (-193.8 m); 采用图 2(b) 所示的控制结构时, RLV 对期望着陆轨迹的跟踪精度较高 (最大高度跟踪误差仅为 1.18 m), RLV 各接地状态 (升降速度为 1.37 m/s, 俯仰角为 7.46°, 接地纵向偏差为 -19.1 m) 均接近理想接地状态。另外, 采用本文提出的控制律平滑切换方法, 也实现了起落架释放前后 2 组控制律之间的平滑切换。

6 结论

本文根据 RLV 拉平着陆的轨迹跟踪纵向控制需求, 提出了一种常规反馈控制与按输入补偿的前馈控制相结合的复合控制方案。

1) 提高了 RLV 的轨迹跟踪精度, 减小了飞行器的接地散布。

2) 基于时间加权高度跟踪误差/误差变化率平方积分指标的前馈控制律参数设计方法是有效的。

3) 通过对新的控制律中积分器赋初值, 实现了不同控制律之间的平滑切换。

参考文献 (References)

[1] KLUEVER C A, NEAL D A. Approach and landing range guidance for an unpowered reusable launch vehicle [J]. Journal of

- Guidance, Control, and Dynamics, 2015, 38(11):2057-2066.
- [2] SUN H F, YANG Z L, ZENG J P. New tracking-control strategy for airbreathing hypersonic vehicles [J]. Journal of Guidance, Control, and Dynamics, 2013, 36(3):846-859.
- [3] XU H J, MIRMIRANI M, IOANNOU P A. Robust neural adaptive control of a hypersonic aircraft [C] // AIAA Guidance, Navigation, and Control Conference and Exhibit. Reston: AIAA, 2003:1-8.
- [4] HALL C E, SHTESSEL Y B. Sliding mode disturbance observer-based control for a reusable launch vehicle [J]. Journal of Guidance, Control, and Dynamics, 2006, 29(6):1315-1328.
- [5] 潘彦鹏, 周军, 呼卫军. 临近空间飞行器再入段复合控制律设计 [J]. 系统工程与电子技术, 2013, 35(11):2364-2369.
- PAN Y P, ZHOU J, HU W J. Design of compound control law for near space vehicle reentry [J]. Systems Engineering and Electronics, 2013, 35(11):2364-2369 (in Chinese).
- [6] BEVACQUA T, BEST E, HUIZERGA A, et al. Improved trajectory linearization flight controller for reusable launch vehicles [C] // 42nd AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2004:875-887.
- [7] LIU Z, WANG Y, HAO X W. Coordinated landing control of unmanned aerial vehicle [C] // 2011 International Conference on Electronics, Communications and Control. Piscataway, NJ: IEEE Press, 2011:1965-1970.
- [8] 郝现伟, 王勇, 杨业, 等. 变构型多操纵面 RLV 进场着陆轨迹优化设计 [J]. 北京航空航天大学学报, 2015, 41(12):2232-2239.
- HAO X W, WANG Y, YANG Y, et al. Optimization design of approach and landing trajectory for variable configuration RLV with multi-control surfaces [J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(12):2232-2239 (in Chinese).
- [9] WISE K A, BRINKER J S. Linear quadratic flight control for ejection seats [J]. Journal of Guidance, Control, and Dynamics, 1996, 19(1):15-22.
- [10] WISE K A. A trade study on missile autopilot design using optimal control theory: AIAA-2007-6673 [R]. Reston: AIAA, 2007.
- [11] 姜文, 贺显曜, 贺恭性. 改进火炮拖动随动系统跟随特性的方法 [J]. 火力与指挥控制, 2009, 34(3):92-95.
- JIANG W, HE Y Y, HE G X. An improved method for the following character of gun servo system [J]. Fire Control and Command Control, 2009, 34(3):92-95 (in Chinese).
- [12] GUGERCIN S, ANTOULAS A C. A survey of model reduction by balanced truncation and some new results [J]. International Journal of Control, 2004, 77(8):748-766.
- [13] 肖业伦. 飞行器运动方程 [M]. 北京: 航空工业出版社, 1987:64-81.
- XIAO Y L. Model of aircraft [M]. Beijing: Aviation Industry Press, 1987:64-81 (in Chinese).
- [14] 薛定宇. 控制系统计算机辅助设计——Matlab 语言与应用 [M]. 3 版. 北京: 清华大学出版社, 2012:284-290.
- XUE D Y. Computer aided control systems design using Matlab language [M]. 3rd ed. Beijing: Tsinghua University Press, 2012:284-290 (in Chinese).
- [15] 杨一栋, 牛佩翼, 江浩. 飞行控制系统模式转换瞬变抑制技术的研究 [J]. 航空学报, 1990, 11(1):88-92.
- YANG Y D, NIU P Y, JIANG H. The study of transient suppression techniques for multimode flight control system [J]. Acta Aeronautica et Astronautica Sinica, 1990, 11(1):88-92 (in Chinese).

作者简介:

郝现伟 男, 博士研究生。主要研究方向: 无人驾驶飞行器飞行控制。

E-mail: haoxw_1220@163.com

王勇 男, 博士, 研究员, 博士生导师。主要研究方向: 无人驾驶飞行器飞行控制。

Tel.: 010-82317544

E-mail: wy_buaa@sina.com

Flare longitudinal control for approach and landing of reusable launch vehicle

HAO Xianwei¹, WANG Yong^{1,*}, YANG Ye², GUO Tao², ZHANG Daibing³

(1. School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;

2. Beijing Aerospace Automatic Control Institute, Beijing 100854, China;

3. School of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha 410073, China)

Abstract: To solve the flare longitudinal control problem for approach and landing of reusable launch vehicle (RLV), a compound control strategy based on combined feedback control and feedforward control is proposed. The feedback control law parameters were designed. The feedforward control law parameters were designed and optimized based on the index of time-weighted height tracking error and the integral for square of error change rate. Under the premise of not affecting the stability of the system, according to feedforward control with the input compensation, the tracking precision of RLV for the flare trajectory was improved and the ground distribution of RLV was reduced. The method of control law smooth switching was proposed by initial value of integrator, which realizes the smooth switching between the different control laws of the RLV landing gear before and after release. Simulation verifies that flare longitudinal compound control law and the method of control law smooth switching are effective.

Key words: reusable launch vehicle (RLV); approach and landing; feedforward correction; model reduction; control laws switching

Received: 2015-07-01; **Accepted:** 2015-09-25; **Published online:** 2015-10-30 16:58

URL: www.cnki.net/kcms/detail/11.2625.V.20151030.1658.009.html

Foundation item: Key Pre-research Foundation of General Armament Department (9140A25030109HK0103)

* **Corresponding author.** Tel.: 010-82317544 E-mail: wy_buaa@sina.com

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0429

基于母线电流的无刷直流电机换相位置优化策略



张前, 冯明*

(北京科技大学 机械工程学院, 北京 100083)

摘 要: 通过分析无刷直流(BLDC)电机换相位置与母线电流的关系,建立了母线电流与换相位置偏差的数学关系,在此基础上提出了一种无位置传感器无刷直流电机换相位置优化策略。该策略以控制绕组电流与反电动势(BEMF)同相位为目标,以换相位置补偿角为被控制量,以母线电流随换相角度的变化率作为偏差量,通过比例积分(PI)调节器调节电机换相位置;并且可以同时校正位置信号检测误差和绕组电感引起的相位偏差。通过实验验证了该策略可以准确、快速地找到最佳换相位置,有效提高电机运行效率。

关键词: 无刷直流(BLDC)电机; 无位置传感器; 母线电流; 换相位置优化; 反电动势(BEMF)法

中图分类号: TM301.2

文献标识码: A

文章编号: 1001-5965(2016)07-1441-08

无刷直流(BLDC)电机因其高可靠性、高功率密度和高效率等优点,在众多领域得到广泛应用。其中,无位置传感器无刷直流电机省去了位置传感器,可以使电机结构更紧凑,可靠性更高,有效降低电机系统成本。因此,对于无位置传感器无刷直流电机控制策略的研究具有重要意义。

常见的无位置传感器无刷直流电机控制方法主要有反电动势(BEMF)法^[1-4]、续流二极管法^[5]、磁链法^[6]、滑模观测器法^[7]和电流注入法^[8]等。以上各种无位置传感器无刷直流电机控制方法均需要通过电机参数及运行状态得到转子的位置信息,该位置信息的准确性是影响无位置传感器控制方式能否高效运行的关键。但是任意一种无位置传感器无刷直流电机控制方法都无法保证其得到的位置信息绝对准确,电机参数的误差、温度、外界环境的变化、滤波及干扰因素的影响等都会导致提取到的转子位置信息出现偏差。文献[9]提出了一种以非导通相续流电流偏

差为反馈量,以控制非导通相续流电流对称为目标的位置信号闭环控制策略,但是该控制策略只能在特定斩波方式下使用,且无法考虑绕组电感导致的电流相位滞后问题。文献[10]提出了一种通过端电压对称校正换相位置的方法,以端电压差值作为偏差量,以换相位置补偿角作为被控制量,建立比例积分(PI)调节器,校正换相位置,该方法可以有效校正位置信号检测误差,但是无法补偿由绕组电感引起的电流相位滞后误差。文献[11]通过实验检测不同转速下测得的位置信号和实际位置之间的相位误差,建立表格,然后通过查表法补偿误差,该方法可以在一定程度上减小换相误差,但是该补偿过程是开环控制的,无法根据电机运行状态实时调节换相位置补偿角,当电机运行参数与建立表格时运行参数不同时,可能会导致换相偏差增大,甚至导致运行失败。文献[12-13]分析了超前滞后换相对电流的影响,指出了超前或滞后换相都会引起导通相电流增大甚

收稿日期: 2015-06-26; 录用日期: 2015-08-26; 网络出版时间: 2015-11-16 15:00

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151116.1500.003.html

基金项目: 国家“863”计划(SS2012AA110501)

* 通讯作者: Tel.: 010-82376928 E-mail: mingfeng@me.ustb.edu.cn

引用格式: 张前, 冯明. 基于母线电流的无刷直流电机换相位置优化策略[J]. 北京航空航天大学学报, 2016, 42(7): 1441-1448. ZHANG Q, FENG M. BLDC motor commutation position optimization strategy based on bus current[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1441-1448 (in Chinese).

至非导通相导通,而电流的增加必然导致损耗的增加。

无位置传感器无刷直流电机换相位置的准确性,不仅包含提取到的转子位置信息的准确性,还须考虑电机绕组电感引起的电流滞后^[14-16]。因此,换相位置准确与否包括2个方面:①检测到的转子位置是否准确;②能否准确计算绕组电感引起的电流滞后角度。一般的转子位置补偿策略通常只考虑检测到的转子位置信号的准确性,而未考虑绕组电感引起的电流相位滞后。

本文通过推导无刷直流电机换相位置与母线电流的数学关系,构造了以换相位置补偿角为控制对象,以母线电流随换相角度的变化率为偏差量的PI调节器,并在此基础上提出了一种无刷直流电机换相位置优化策略,该策略可以同时校正位置信号检测和绕组电感引起的换相误差,无需将两者分开计算。经实验验证,该策略能够以简单的优化过程实现良好的优化效果,有效提高了无刷直流电机的运行效率。

1 换相位置对母线电流的影响

无刷直流电机输出的电磁功率等于反电动势与电枢电流的数量积,换相位置不准确会引起电枢电流与反电动势相位不一致,造成功率因数降低,进而导致无刷直流电机效率下降,而对换相位置优化的最终目的是使电枢电流与反电动势相位保持一致。为了简化计算,设梯形波反电动势的平顶宽度为120°,将电枢电流等效为标准的方波电流,不考虑弱磁增磁作用的影响。

1.1 电枢电流与反电动势同相时的功率

图1所示为a相电枢电流与反电动势同相位时的波形。此时无刷直流电机处于最佳运行状态,其力矩波动最小,效率最高。

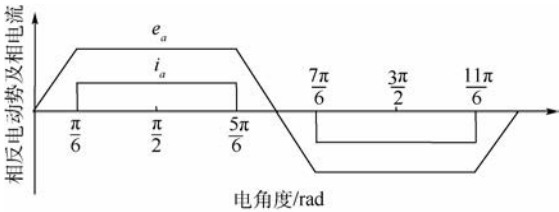


图1 无相位差时a相电枢电流与反电动势波形

Fig. 1 Armature current and BEMF waveform of phase a with no commutation error

选取图1中一个换相周期 $\pi/6 \sim \pi/2$ 为分析对象,在 $\pi/6 \sim \pi/2$ 期间,无刷直流电机输出的电磁功率 P_0 为

$$P_0 = e_a \cdot i_a + e_b \cdot i_b + e_c \cdot i_c \quad (1)$$

式中: e_a 、 e_b 和 e_c 为相反电动势; i_a 、 i_b 和 i_c 为相电

流。在此换相周期内,有 $e_a = -e_b = E$, $i_a = -i_b = I_0$, E 为相反电动势最大值, I_0 为相电流和相反电动势同相位时的直流母线电流, $i_c = 0$ 。

由式(1)可知:

$$P_0 = 2EI_0 \quad (2)$$

1.2 电枢电流与反电动势不同相时的功率

图2所示为电枢电流与反电动势存在相位差 θ 时的三相反电动势和电流波形。 θ 为负时,表示电流超前, θ 为正时,表示电流滞后。取一个换相周期 $(\pi/6 + \theta) \sim (\pi/2 + \theta)$ 作为分析对象,该换相周期内输出的电磁功率 P_θ 为

$$P_\theta = e'_a \cdot i'_a + e'_b \cdot i'_b + e'_c \cdot i'_c \quad (3)$$

式中: e'_a 、 e'_b 和 e'_c 为存在相位差 θ 时的相反电动势; i'_a 、 i'_b 和 i'_c 为存在相位差 θ 时的相电流。在此换相周期内,有 $i'_a = -i'_b = I_\theta$, $i'_c = 0$, I_θ 为存在相位差 θ 时的母线电流。

因此,可由式(3)得

$$P_\theta = (e'_a - e'_b)I_\theta \quad (4)$$

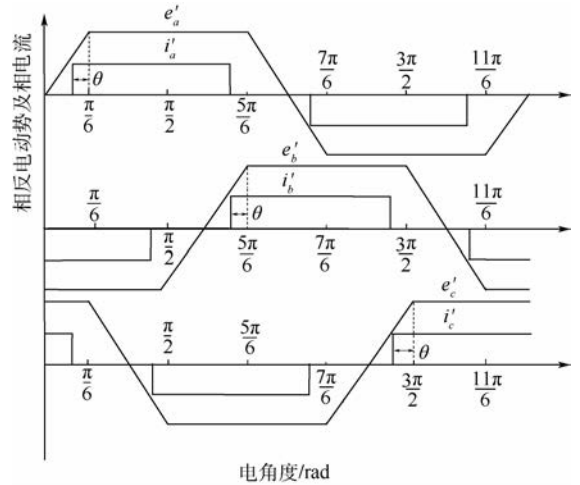


图2 不同相时三相电枢电流与反电动势波形

Fig. 2 Armature current and BEMF waveform of triphase with commutation error

1.3 换相角度偏差对母线电流的影响

无刷直流电机在固定转速、固定功率运行时,在每一个换相周期内输出的电磁功相同,因此有

$$\int_0^{3\pi/\omega} P_0 dt = \int_0^{3\pi/\omega} P_\theta dt \quad (5)$$

式中: ω 为电角速度。

由图2可知, $\theta < 0$ 时,换相周期 $(\pi/6 + \theta) \sim (\pi/2 + \theta)$ 内各相反电动势的值为

$$e'_a = \begin{cases} E + \frac{6E\theta}{\pi} + \frac{6Et\omega}{\pi} & 0 < t \leq -\frac{\theta}{\omega} \\ E & -\frac{\theta}{\omega} < t \leq \frac{\pi}{3\omega} \end{cases} \quad (6a)$$

$$e'_b = -E \quad 0 < t \leq \frac{\pi}{3\omega} \quad (6b)$$

$$e'_c = \begin{cases} E & 0 < t \leq -\frac{\theta}{\omega} \\ E - \frac{6E\theta}{\pi} - \frac{6Et\omega}{\pi} & -\frac{\theta}{\omega} < t \leq \frac{\pi}{3\omega} \end{cases} \quad (6c)$$

$\theta > 0$ 时,该换相周期内各相反电动势的值为

$$e'_a = E \quad 0 < t \leq \frac{\pi}{3\omega} \quad (7a)$$

$$e'_b = \begin{cases} -E & 0 < t \leq \frac{\pi}{3\omega} - \frac{\theta}{\omega} \\ -3E + \frac{6E\theta}{\pi} + \frac{6Et\omega}{\pi} & \frac{\pi}{3\omega} - \frac{\theta}{\omega} < t \leq \frac{\pi}{3\omega} \end{cases} \quad (7b)$$

$$e'_c = \begin{cases} E - \frac{6E\theta}{\pi} - \frac{6Et\omega}{\pi} & 0 < t \leq \frac{\pi}{3\omega} - \frac{\theta}{\omega} \\ -E & \frac{\pi}{3\omega} - \frac{\theta}{\omega} < t \leq \frac{\pi}{3\omega} \end{cases} \quad (7c)$$

以 $\pi/6 + \theta$ 对应的时刻作为式(5)中积分的零时刻,由式(2)、式(4)~式(6)或式(2)、式(4)、式(5)、式(7)均可得,换相超前和滞后时的母线电流 I_θ 均满足:

$$\frac{I_\theta}{I_0} = \frac{2\pi^2}{2\pi^2 - 9\theta^2} \quad (8)$$

随着电流超前滞后角的增大,得到的母线电流 I_θ 会增大,其最小值为 I_0 ,为相电流和相反电动势同相位时的直流母线电流,其值不受换相角度偏差的影响, I_0 的具体值对后续换相位置优化策略影响较小,因此无需计算其精确值,用该转速下的额定电流近似代替即可。

式(8)是在 $(\pi/6 + \theta) \sim (\pi/2 + \theta)$ 换相周期计算得到的,但由于三相对称的无刷直流电机 6 个换相周期是等价的,因此在无刷直流电机运行期间,换相角度偏差和母线电流之间恒满足式(8)。

母线电流与换相角度偏差的关系如图 3 所示。随着换相角度偏差增大,母线电流增加越来越明显, θ 为 0 时,表示无刷直流电机相电流和相反电势同相位。由于换相位置偏差超过 $\pi/3$ 时,

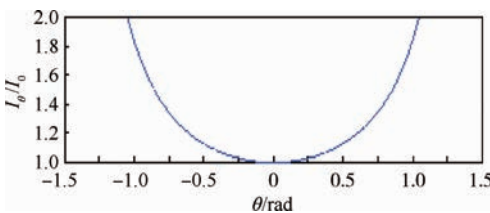


图3 母线电流与换相角度偏差的关系

Fig.3 Relationship between bus current and angular commutation deviation

无刷直流电机已经失步,此时其已无法正常运行,因此上述分析均在 $-\pi/3 < \theta < \pi/3$ 区间进行。

2 换相位置优化策略在驱动系统中的实现

2.1 构建 PI 调节器

由第 1.3 节分析可知, I_θ 是随换相角度偏差 θ 变化的值,取 I_θ 对 θ 的变化率记为

$$A = dI_\theta/d\theta \quad (9)$$

母线电流变化率与换相角度偏差的关系如图 4 所示。 θ 越接近 0, A 的绝对值越小;反之, A 的绝对值越大,且电流超前时 A 为负值,电流滞后时 A 为正值。

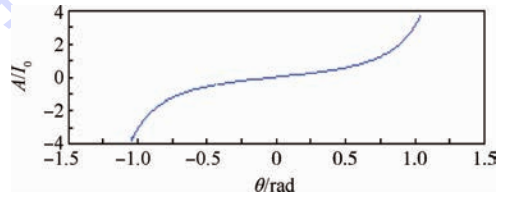


图4 母线电流变化率与换相角度偏差的关系

Fig.4 Relationship between bus current gradient and angular commutation deviation

通过上述数学分析,可以构造以 A 为偏差量,以换相位置补偿角 α 作为被控制量的 PI 调节器。构建的 A 和 α 的关系为

$$\alpha = K_p A + K_i \int A dt \quad (10)$$

式中: K_p 为比例系数; K_i 为积分系数。

以无位置传感器反电动势法为例,实际工作中,通过反电动势法得到的反电动势过零点信号与实际换相位置并非直接对应(采用 30° 或 90° 相位延时),通常需要外加一定延时角度 τ ,最终延时角度 φ 可以表示为

$$\varphi = \tau + \alpha \quad (11)$$

式中: τ 为综合考虑信号滤波电路引起的相位延时、电感引起的电流滞后、电压比较器和光耦等芯片引起的延时及计算过程引起的延时等得到的相位延时角。

由图 4 可知,取 $K_p < 0, K_i < 0$,当电流超前于反电动势时, $\theta < 0, A < 0$,根据式(10)得 $\alpha > 0$,由式(11)可得延时角度 φ 增大,从而减小电流与反电动势的相位差;反之,延时角度 φ 减小。

将式(11)按照增量式 PI 调节器离散化后可以得到

$$\Delta\alpha_k = K_p(A_k - A_{k-1}) + K_i A_k \quad (12)$$

式中: A_k 为第 k 次进入 PI 调节时的 A 值; $\Delta\alpha_k$ 为第 k 次换相位置补偿角的变化量; k 为正整数。

离散化以后的软件延时角度可以表示为

$$\varphi = \tau + \Delta\alpha_1 + \Delta\alpha_2 + \dots + \Delta\alpha_k \quad (13)$$

2.2 调速系统逻辑框图

本文通过分析母线电流与换相角度偏差之间的关系提出了以母线电流变化率为基础的换相位置优化策略。图5为采用该换相位置优化策略的调速系统控制逻辑框图。图中： ω_0 为给定的参考速度；PID为速度环的比例积分微分调节器；BUCK为用于调节转速的降压电路。与传统的无刷直流电机调速系统相比，采用换相位置优化策略的调速系统增加了母线电流检测、计算A值和PI调节三部分。该控制系统中，首先通过三相端电压分压、滤波和比较等环节得到反电动势过零点；然后根据转速、绕组电感、滤波电路和计算耗时等估算

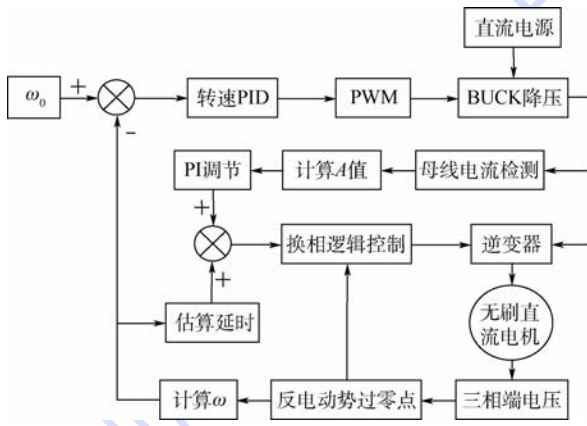


图5 采用换相位置优化策略的调速系统逻辑框图

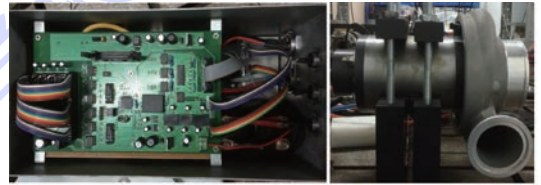
Fig. 5 Logic diagram of speed control system using commutation position optimization strategy

延时时间；最后在估算延时的基础上通过本文提出的换相位置优化策略进行PI调节，得到换相位置补偿角，进一步优化换相位置。

3 实验验证

3.1 实验装置

实验对象为氢燃料电池汽车高速离心式空气压缩机。空气压缩机属于风机类型负载，其输出转矩及功率随转速升高而增加，叶轮克服空气阻力做功即其有效输出功。实验装置如图6所示。



(a) 驱动器 (b) 空气压缩机

图6 实验装置

Fig. 6 Experimental devices

驱动器原理框图如图7所示。图中： V_i 为直流电源的输入电压； V_a 、 V_b 和 V_c 为三端的端电压；JTAG为调试接口；CAN总线为控制器局域网总线。通过BUCK变换器输出的直流电压 V_o 调节无刷直流电机转速，通过反电动势法提取转子位置信号，母线上装有电流传感器，通过单片机的AD转换模块采集电流传感器输出的信号用于A的计算。

氢燃料电池汽车高速电动离心式空气压缩机的电机参数如表1所示。

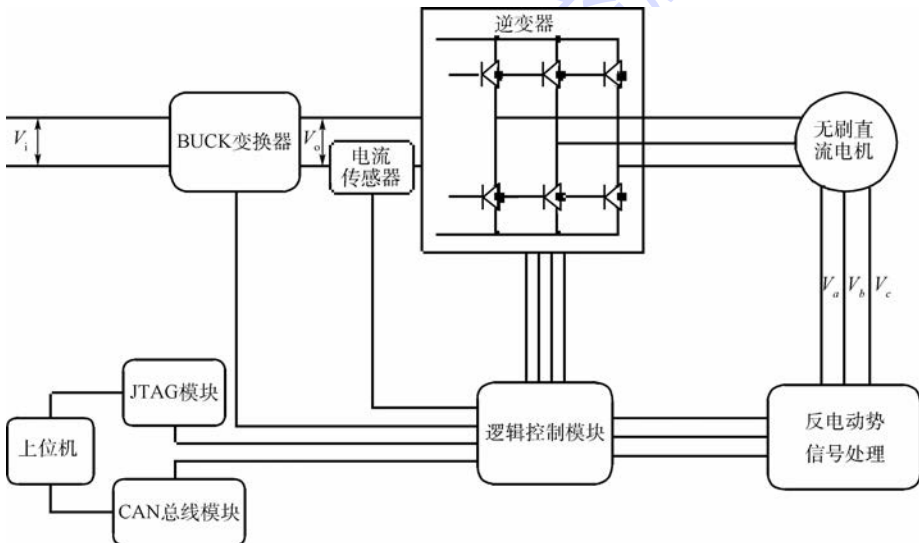


图7 实验用驱动器原理框图

Fig. 7 Principle diagram of experimental drive

表1 空气压缩机电机参数

Table 1 Parameters of electric machine as air compressor

电机参数	数值
额定功率/kW	10
额定转速/(r·min ⁻¹)	100 000
待机转速/(r·min ⁻¹)	15 000
极对数	1
线电感/ μ H	80
线电阻/m Ω	40
额定电流/A	50
额定电压/V	250

3.2 稳态实验分析

图8为30 000 r/min时a相电流及端电压波形。上半屏为a相电流波形,下半屏为a相对地的端电压波形。由图8(a)可知,当换相超前时,a相导通前,S1处的反电动势明显小于a相关断后S2处的反电动势;由图8(b)可知,当换相滞后时,S1处的反电动势大于S2处的反电动势;由图8(c)可

知,经过换相位置优化以后,S1与S2处的反电动势大致相同。通过端电压波形对比可以看出,该换相位置优化策略能有效补偿换相位置偏差。

为了进一步验证该换相位置优化策略的优化效果,实验中将至式(11)所示换相延时公式改为

$$\varphi = \tau + \tau_e + \alpha \quad (14)$$

式中: τ_e 为验证该换相位置优化策略的优化效果而额外增加的延时。实验中, τ_e 从 -10° 到 20° 每隔 2° 做一次实验,记录换相位置优化前后的母线电流值,优化前 α 取值为0,优化后其值为通过该换相位置优化策略得到的值。在转速15 000、20 000、25 000、30 000 r/min得到的4组实验结果如图9所示。通过实验结果可以看出,在不同转速下,优化前的母线电流随换相位置的变化规律符合式(8),因而该实验可以进一步证明本文建立的母线电流与换相位置的数学关系的正确性。

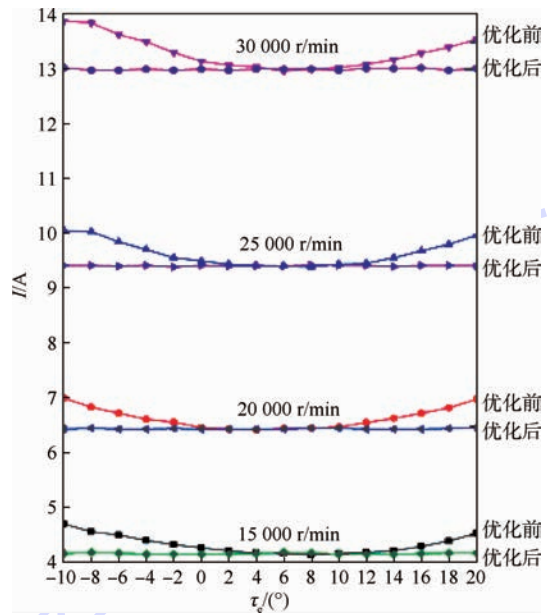


图9 换相位置优化前后母线电流

Fig. 9 Bus current before and after commutation position optimization

从图9中25 000 r/min和30 000 r/min换相位置优化前的电流曲线可以发现,在 $\tau_e = -10^\circ$ 时得到的母线电流值与 $\tau_e = -8^\circ$ 时基本相同,这是因为换相位置优化前,在 $\tau_e = -8^\circ$ 时得到的 φ 已经接近于0,当 $\tau_e = -10^\circ$ 时,通过式(14)得到 φ 为负值,而 φ 为负值在实际换相位置调整中是无法实现的,在程序中,当得到 φ 为负值时,将其调整为0。

通过不同转速下换相位置优化前后的数据对比,可以确定出各个转速下最佳换相点的位置和换相位置优化策略可以达到的换相延时区间,如表2所示。

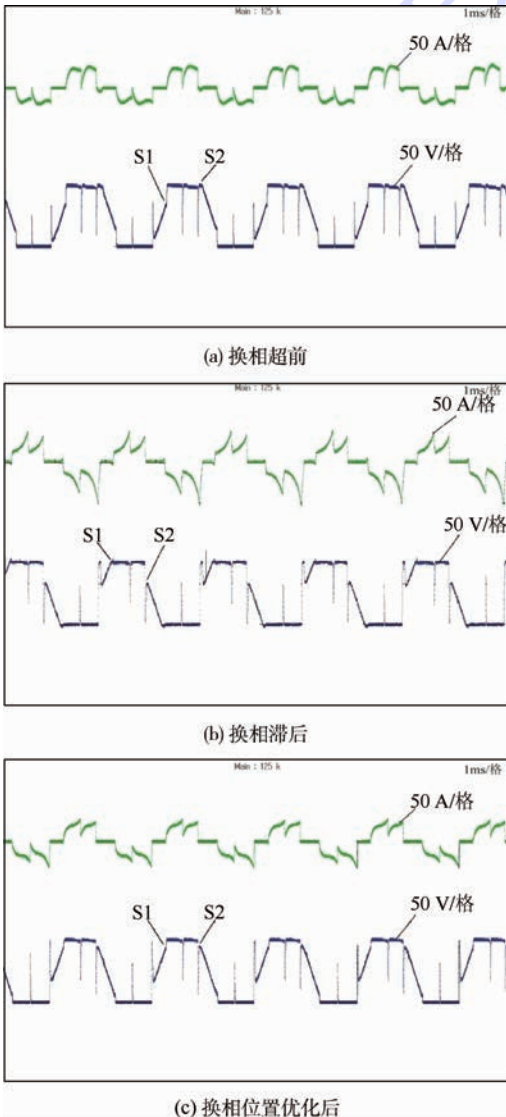


图8 a相电流及端电压波形

Fig. 8 Current and terminal voltage waveforms of phase a

表2 不同转速下换相位置优化结果

Table 2 Commutation position optimization results at different rotating speeds

转速/(r·min ⁻¹)	最佳换相延时/(°)	换相延时区间/(°)
15 000	$\tau + 8$	$(\tau + 4) \sim (\tau + 12)$
20 000	$\tau + 4$	$(\tau + 2) \sim (\tau + 8)$
25 000	$\tau + 6$	$(\tau + 4) \sim (\tau + 10)$
30 000	$\tau + 6$	$(\tau + 4) \sim (\tau + 8)$

表2中,最佳换相延时指通过换相位置优化前的曲线得到的母线电流最小时的换相延时角度。换相延时区间指通过换相位置优化策略可以达到的换相角度范围。以图9中15 000 r/min为例,在-10°~20°范围内优化后的电流值均小于优化前($\tau + 4^\circ$)和($\tau + 12^\circ$)时的电流值,这就意味着换相位置优化策略可以将换相位置由任意角度调节到($\tau + 4^\circ$)~($\tau + 12^\circ$)之间,因此,15 000 r/min时的换相延时区间即为($\tau + 4^\circ$)~($\tau + 12^\circ$)。

从表2可以看出,15 000 r/min时优化后的换相延时区间为8°,随着转速的升高,母线电流值的增加,在30 000 r/min时,优化后的换相延时区间减小到4°,换相位置优化策略的优化效果随母线电流值的增加逐渐提高。

3.3 动态实验分析

为验证动态条件下的换相位置优化效果,分别测量换相位置优化前后15 000~20 000 r/min,20 000~25 000 r/min,25 000~30 000 r/min加速过程的母线电流有效值及加速时间,加速过程采用同样的控制参数,是否进行位置优化通过CAN总线进行设置,转速达到设定值时立刻停止输出,以便于区分加速阶段与稳态运行阶段,实验结果如表3所示。

表3 优化前后母线电流及加速时间

Table3 Bus current and accelerating time before and after optimization

加速过程转速/ (r·min ⁻¹)	加速时间/s		加速过程电流/A	
	优化前	优化后	优化前	优化后
15 000~20 000	3.6	3.4	6.37	6.29
20 000~25 000	3.2	3.0	9.27	9.19
25 000~30 000	3.4	3.0	12.87	12.69

通过表3可以看出,换相位置优化后的加速时间小于优化前,同时其母线电流的有效值也小于换相位置优化前,说明在加速过程中通过换相位置优化策略提高了无刷直流电机运行的效率,因此可以从侧面证明该换相位置优化策略对于动态过程同样具有良好效果。25 000~30 000 r/min优化前后加速过程如图10所示。

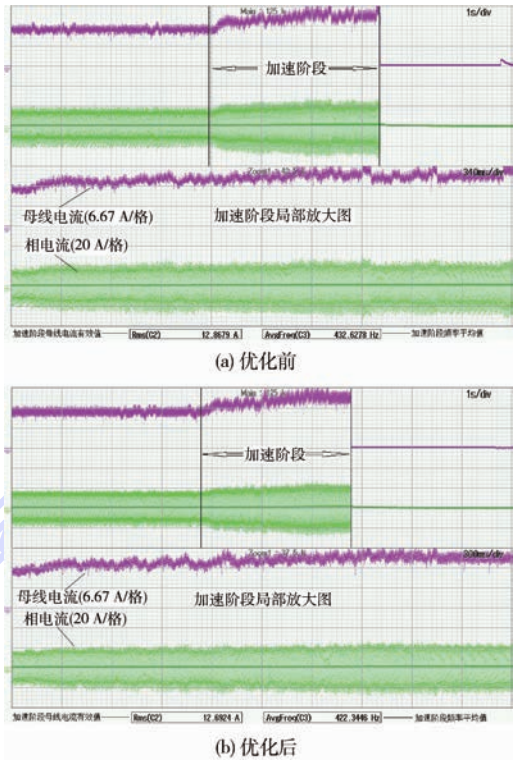


图10 换相位置优化前后实验波形

Fig. 10 Experimental waveform before and after commutation position optimization

4 结论

本文通过综合分析换相位置偏差与母线电流的关系,提出了以母线电流对换相角度变化率为基础的换相位置优化策略,通过理论分析及实验表明:

1) 该策略可以准确调节换相位置到最优换相点,有效减小母线电流,改善相电流及端电压波形,提高无刷直流电机运行效率。

2) 可以同时校正位置信号检测误差和绕组电感引起的电流相位滞后,避免了采用换相位置优化策略时无法进行相位超前角控制的问题。

3) 无需计算偏差的具体值,而是通过PI调节使换相位置在最优换相点附近小幅波动,因此无需复杂的计算,简化了优化过程。

4) 通过采样母线电流作为调节依据,无需检测每次换相时的电流或电压,对运算速度及采样频率要求较低,可以节约CPU资源,适用于高速驱动场合。

本文提出的换相位置优化策略在稳态时优化效果良好,在动态时对换相位置也可以进行相应的优化,但是当负载波动频率与该换相位置优化策略调节的频率接近时,优化效果并不理想,因此,仍需进一步改进控制策略以提高其在各种工况下的优化效果。

致谢 感谢同济大学汽车学院对本文研究工作的支持与帮助。

参考文献 (References)

- [1] ACARNLEY P P, WATSON J F. Review of position-sensorless operation of brushless permanent-magnet machines [J]. IEEE Transactions on Industrial Electronics, 2006, 53 (2) : 352-362.
- [2] DAMODHARAN P, VASUDEVAN K. Sensorless brushless DC motor drive based on the zero-crossing detection of back electromotive force (EMF) from the line voltage difference [J]. IEEE Transactions on Energy Conversion, 2010, 25 (3) : 661-668.
- [3] BI C, HLA N P, JIANG Q, et al. Back-EMF ZCP error induced by electromagnetic structure of spindle motor [J]. IEEE Transactions on Magnetics, 2011, 47 (7) : 1899-1905.
- [4] ZWYSSIG C, KOLAR J W, ROUND S D. Megaspeed drive systems: Pushing beyond 1 million r/min [J]. IEEE/ASME Transactions on Mechatronics, 2009, 14 (5) : 564-574.
- [5] OGASAWARA S, AKAGI H. An approach to position sensorless drive for brushless DC motors [J]. IEEE Transactions on Industry Applications, 1991, 27 (5) : 928-933.
- [6] ERTUGRUL N, ACARNLEY P. A new algorithm for sensorless operation of permanent magnet motors [J]. IEEE Transactions on Industry Applications, 1994, 30 (1) : 126-133.
- [7] FAKHAM H, DJEMAI M, BUSAWON K. Design and practical implementation of a back-EMF sliding-mode observer for a brushless DC motor [J]. IET Electric Power Applications, 2008, 2 (6) : 353-361.
- [8] KASA N, WATANABE H. A mechanical sensorless control system for salient-pole brushless DC motor with autocalibration of estimated position angles [J]. IEEE Transactions on Power Electronics, 2000, 47 (2) : 389-395.
- [9] 宋飞,周波,吴小婧.校正无位置传感器无刷直流电机位置信号相位的闭环控制策略 [J].中国电机工程学报,2009,29(12):52-57.
SONG F, ZHOU B, WU X J. Closed loop control method to correct position phase for sensorless brushless DC motor [J]. Proceedings of the CSEE, 2009, 29 (12) : 52-57 (in Chinese).
- [10] 吴小婧,周波,宋飞.基于端电压对称的无位置传感器无刷直流电机位置信号相位校正 [J].电工技术学报,2009,24(4):54-60.
WU X J, ZHOU B, SONG F. A new control method to correct position phase for sensorless brushless DC motor [J]. Transactions of China Electrotechnical Society, 2009, 24 (4) : 54-60 (in Chinese).
- [11] 韦鲲,任建军,张仲超.三次谐波检测无刷直流电机转子位置的研究 [J].中国电机工程学报,2004,24(5):163-167.
WEI K, REN J J, ZHANG Z C. Research on the scheme of sensing rotor position of BLDCM based on the third harmonic component [J]. Proceedings of the CSEE, 2004, 24 (5) : 163-167 (in Chinese).
- [12] LIN M Y, LI Q, GU W G. Effect of rotor position error on commutation in sensorless BLDC motor drives [C] // 8th International Conference on Electrical Machines and Systems (ICEMS 2005). Piscataway, NJ: IEEE Press, 2005.
- [13] SAMOYLENKO N, HAN Q, JATSKEVICH J. Balancing hall-effect signals in low-precision brushless DC motors [C] // Applied Power Electronics Conference (APEC 2007). Piscataway, NJ: IEEE Press, 2007.
- [14] JIANG Q, BI C, HUANG R Y. A new phase-delay-free method to detect back EMF zero-crossing points for sensorless control of spindle motors [J]. IEEE Transactions on Magnetics, 2005, 41 (7) : 2287-2294.
- [15] SHEN J X, TSENG K J. Analysis and compensation of rotor position detection error in sensorless PM brushless DC motor drives [J]. IEEE Transactions on Energy Conversion, 2003, 18 (1) : 87-93.
- [16] HAN Q, SAMOYLENKO N, JATSKEVICH J. Average-value modeling of brushless DC motors with 120° voltage source inverter [J]. IEEE Transactions on Energy Conversion, 2008, 23 (2) : 423-432.

作者简介:

张前 男,博士研究生。主要研究方向:无刷直流电机及永磁同步电机控制系统。

Tel. : 18600003063

E-mail: jz_ustb@163.com

冯明 男,博士,教授,博士生导师。主要研究方向:高速电机设计及其控制。

Tel. : 010-82376928

E-mail: mingfeng@me.ustb.edu.cn

BLDC motor commutation position optimization strategy based on bus current

ZHANG Qian, FENG Ming*

(School of Mechanical Engineering, University of Science and Technology Beijing, Beijing 100083, China)

Abstract: This paper analyzes the relation between commutation position of brushless direct current (BLDC) motors and bus current, builds up their mathematical relation, and proposes a sensorless BLDC motors commutation position optimization strategy. This strategy regulates the commutation position by a proportional integral (PI) adjuster, aiming at regulating the current and the back electromotive force (BEMF) in-phase. The controlled variable of the PI adjuster is the commutation position compensating angle and the deviation value is the ratio of current change over angle. This optimization strategy can revise the detected position signal error and the phase deviation caused by the winding inductance at the same time. The experimental results verify that this optimization strategy can get the best commutation position accurately and fast, and improve the operating efficiency effectively.

Key words: brushless direct current (BLDC) motor; sensorless; bus current; commutation position optimization; back electromotive force (BEMF) method

Received: 2015-06-26; **Accepted:** 2015-08-26; **Published online:** 2015-11-16 15:00

URL: www.cnki.net/kcms/detail/11.2625.V.20151116.1500.003.html

Foundation item: National High-tech Research and Development Program of China (SS2012AA110501)

* **Corresponding author.** Tel.: 010-82376928 E-mail: mingfeng@me.ustb.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0433

不完备信息条件下的并发故障诊断方法



孙伟超¹, 许爱强^{2,*}, 李文海²

(1. 海军航空工程学院 研究生管理大队, 烟台 264001; 2. 海军航空工程学院 科研部, 烟台 264001)

摘 要: 在实际的装备故障诊断过程中,经常存在测试信息不完整的情况,而此时对并发故障进行诊断,则更为困难。针对这种情况,首先,对目前应用于不完备信息条件下的故障诊断方法以及并发故障诊断方法进行了分析。定义了不完备并发故障诊断决策系统对存在缺失的测试信息进行表述。提出了不完备边界粗糙熵对决策系统的不确定性进行度量,给出每条测试属性的重要度,同时给出了不完备信息条件下属性值频率的计算方法。然后,为了对并发故障进行诊断,在 DSmT 框架下构建了并发故障诊断模型,在此模型下提出一种融合证据特征的区间信度合成规则。最后,通过故障诊断实例,验证了方法的有效性和适用性。

关键词: 并发故障诊断; DSmT 理论; 区间信度合成; 粗糙集理论; 边界粗糙熵

中图分类号: TP181; V24

文献标识码: A

文章编号: 1001-5965(2016)07-1449-12

随着技术的发展,现代设备结构日趋复杂,逐渐向大型化、自动化等方向发展,各部件之间联系更加紧密,这就为复杂装备的故障诊断带来了新的挑战。在诊断知识获取过程中,由于被测对象结构复杂、诊断系统采集的信息缺失等原因,致使所得到的故障信息存在不完备的情况;同时,在设备实际运行过程中,常常伴随有并发故障的发生。因此,为了保证设备运行安全稳定,对不完备信息条件下的并发故障诊断方法进行研究具有一定的现实意义。

在故障诊断领域,信息不完备是指诊断系统中某一数据或多个数据丢失、不完全或无法确定,即数据信息不完整。Kryszkiewicz^[1]提出了不完备信息系统的容差关系模型及其知识约简方法,使得粗糙集理论成为处理不完备信息的有效手段。目前,针对不完备信息条件下的故障诊断方法主要也是围绕粗糙集理论展开,通过构建不完备决策表对故障信息进行表示。在实际诊断应用

中,文献[2]提出了基于信息熵的约简算法;文献[3]通过定义最大特征相似集完成约简与规则提取;文献[4]则根据机械故障的特征定义了一种新的相容关系;文献[5]基于约简的不唯一性,提出了一种最优测试信息的选取方法;文献[6]则根据不完备信息的特点设计了新的属性约简方法与基于可信度的诊断方法;文献[7]通过建立中间变量直接将不完备信息系统分解为简单规则的集合,提取诊断规则,同时,将粗糙集理论与其他人工智能方法结合进行不完备信息条件下的故障诊断也取得了较好的效果;文献[8]引入模糊决策系统与神经网络,设计了转台的粗糙神经网络故障诊断系统;文献[9]提出了一种基于粗糙集与灰色理论的不完备信息条件下的变电站故障诊断方法,利用粗糙集进行约简,通过属性的灰色相关梯度进行决策;文献[10]在对锅炉水位控制系统的故障诊断中,面对不完备信息,首先用粗糙集方法构建决策表,然后使用 BP 神经网络

收稿日期: 2015-06-30; 录用日期: 2015-09-18; 网络出版时间: 2016-01-08 15:17

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20160108.1517.001.html

* 通讯作者: Tel.: 0535-6635836 E-mail: xuaq6342@yahoo.com.cn

引用格式: 孙伟超, 许爱强, 李文海. 不完备信息条件下的并发故障诊断方法[J]. 北京航空航天大学学报, 2016, 42(7): 1449-1460. SUN W C, XU A Q, LI W H. Diagnosis method of simultaneous fault with incomplete information[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1449-1460 (in Chinese).

络对冗余属性进行约简;文献[11]针对直升机传动系统诊断决策表,在概念上对缺失属性值进行分析,提出粒计算模型的定义,构建算法进行属性约简。

由于多重故障并发时,其表现形式多样,故障和征兆之间并非简单的一一对应关系,造成了对并发故障诊断的困难。针对这一问题,一些新的方法不断地被应用于其中。例如,文献[12]针对核动力装置的并发故障利用复合神经网络实现故障的分离;文献[13]提出了一种基于先验知识和相关向量机的并发故障诊断方法,在新的框架下,只需要使用单发故障模式对相关向量机进行训练,就可以完成对并发故障的判断;文献[14]通过构建2层贝叶斯网络对多源信息进行融合,完成了对地热泵系统的并发故障诊断,有效地提升了诊断准确度。而更为普遍的并发故障诊断研究是在DSmT理论框架中展开的。自DSmT提出以来^[15],因其允许辨识框架内元素不互斥,为并发故障模型的建立提供了新的思路。目前,已有学者开展了基于DSmT的并发故障诊断方法的研究。李植良等^[16]基于DSmT中的广义辨识框架构建了并发故障诊断模型,通过DSmT组合与决策规则得到最终诊断结果;文献[17]利用随机集方法获得诊断证据,并以此构建了新的适用于单发与并发故障的组合规则;文献[18-20]则给出了DSmT理论在实际并发故障诊断中的应用实例。

目前,针对不完备信息条件下的并发故障诊断还少有研究。为了有效地解决这一问题,本文在已有研究的基础上,将粗糙集与DSmT理论相结合进行并发故障诊断。利用粗糙集构建不完备诊断信息决策表,对存在缺失的测试信息进行表示,鉴于不同的测试属性在决策中的重要程度并不相同,提出不完备边界粗糙熵的概念对测试属性的重要度进行度量;同时给出了不完备情况下属性值频率的计算方法,以构建证据的基本信度赋值函数。在DSmT框架下的并发故障诊断模型中进行证据合成。同时,为了在合成过程中体现出不同测试属性对决策的不同重要度,获得更为精确的诊断结果,在DSmT框架下提出一种基于先验知识的区间信度合成规则。将此方法应用于实际诊断过程中,对效果进行验证。

1 不完备故障信息表示与特征提取

1.1 相容关系下不完备并发故障诊断决策系统

定义1 有序数组, $*$ $\in V$ 表示不完备并发

故障诊断决策系统 $I_{DS} = (U, C \cup \{d\}, V, f)$, U 为关于测试样本的非空有限集合, C 为测试属性的非空有限集合, $\{d\}$ 为故障决策属性, V 为测试属性的值域, f 为测试信息函数,它为每一次测试样本的每个测试属性赋予一个测试值,同时,系统中至少存在一个 $c \in C, x \in U$,使得 V 为缺省值,用符号“ $*$ ”表示。

定义2 给定一个不完备并发故障诊断决策系统 $I_{DS} = (U, C \cup \{d\}, V, f)$, $*$ $\in V$,对于任意测试属性的子集 $B \subseteq C$, x, y 分别为测试样本集中的一个测试样本,对于 $\forall x, y \in U$,定义测试样本在测试属性集合 B 下的相容关系为

$$S_{IM}(B) = \{(x, y) \in U \times U \mid \forall b \in B, f_b(x) = f_b(y) \vee f_b(x) = * \vee f_b(y) = *\}$$

令 $S_B(x)$ 表示对象集 $\{y \in U \mid (x, y) \in S_{IM}(B)\}$,对于集合 B 而言, $S_B(x)$ 是与 x 不可区分的对象的最大集合, $f_b(x)$ 表示在属性 b 下与 x 不可区分的对象的最大集合。

令 $S_{IM}(B) = \{S_B(x) \mid x \in U\}$ 表示对 U 的一种分类,则 $U/S_{IM}(B)$ 中的元素称为相容类。 $U/S_{IM}(B)$ 中的相容类一般不构成 U 的划分,其构成 U 的覆盖, $\cup U/S_{IM}(B) = U$ 。

定义3 给定一个不完备并发故障诊断决策系统 $I_{DS} = (U, C \cup \{d\}, V, f)$, $*$ $\in V$ 。对于集合 $X \subseteq U$,在相容关系 $S_{IM}(B)$ 下, X 的上下近似集分别定义为

$$\overline{S_{IM}(B)}(X) = \{x \in U \mid S_B(x) \cap X \neq \emptyset\} = \cup \{S_B(x) \mid x \in X\}$$

$$\underline{S_{IM}(B)}(X) = \{x \in U \mid S_B(x) \subseteq X\} = \{x \in X \mid S_B(x) \subseteq X\}$$

1.2 不完备并发故障诊断决策系统中属性的重要度计算

在不完备并发故障诊断决策系统中,每一条测试属性相对于决策所起的作用均不相同,为此引进测试属性重要度的概念对其进行度量,该重要度随属性的不确定性大小而变化;并且提出不完备并发故障诊断决策系统中的不完备边界粗糙熵的概念,将属性的不确定性予以量化,并依据量化后的值给出测试属性重要度的计算方法。

定义4 给定 $I_{DS} = (U, C \cup \{d\}, V, f)$, $*$ $\in V$,集合 $X \subseteq U$,对于任意测试属性的子集 $B \subseteq C$, $S_{IM}(B)$ 为 U 上的相容关系,则定义集合 $B_{nb}(X) = \overline{S_{IM}(B)}(X) - \underline{S_{IM}(B)}(X)$ 为集合 X 在相容关系 $S_{IM}(B)$ 下的边界域。

定义5 给定 $I_{DS} = (U, C \cup \{d\}, V, f)$, $*$ $\in V$,集合 $X \subseteq U, Y \subseteq U, B \subseteq C$ 。相容关系 $S_{IM}(B)$ 对论

域 U 的分类为 $U/S_{\text{IM}}(B) = \{S_B(x_1), S_B(x_2), \dots, S_B(x_m)\}$, 决策 $\{d\}$ 对论域 U 的划分记为 $U/\{d\} = \{Y_1, Y_2, \dots, Y_n\}$ 。如果记 $Y_i = d^{-1}(\{v_d\}) = \{k \in U: d(k) = v_d\}$, v_d 为决策属性值, 则定义论域 U 在相容关系 $S_{\text{IM}}(B)$ 下得到的分类对于决策分类 Y_i 的决策边界域为

$$BN_B(Y_i) = \overline{S_{\text{IM}}(B)(Y_i)} - S_{\text{IM}}(B)(Y_i)$$

定义6 在 $I_{\text{DS}} = (U, C \cup \{d\}, V, f)$, $* \in V$ 中, $B \subseteq C$, 相容关系 $S_{\text{IM}}(B)$ 对论域 U 的分类为 $U/S_{\text{IM}}(B) = \{S_B(x_1), S_B(x_2), \dots, S_B(x_m)\}$, 决策 $\{d\}$ 对论域 U 的划分记为 $U/\{d\} = \{Y_1, Y_2, \dots, Y_n\}$ 。定义知识 B 的不完备边界粗糙熵为

$$E_{\text{IB}}(B) = \sum_{i=1}^n r_i \left(\frac{|B_{nB}(Y_i)|}{|U|} \text{lb} |U| \right) - \sum_{j=1}^m \frac{|S_B(x_j)|}{|U|} \text{lb} \frac{1}{|S_B(x_j)|} \quad (1)$$

式中: $i=1, 2, \dots, n$; $j=1, 2, \dots, m$; $r_i = |Y_i|/|U|$ 为等价类 Y_i 的基数与 U 的基数之比。

I_{DS} 中的不完备边界粗糙熵的定义式(1)同时考虑了由相容关系 $S_{\text{IM}}(B)$ 在对 U 进行分类时, 导致不确定的2个方面因素。式(1)等号右端第1项 $\sum_{i=1}^n r_i \left(\frac{|B_{nB}(Y_i)|}{|U|} \text{lb} |U| \right)$ 表示决策表中的对象集合在相容关系 $S_{\text{IM}}(B)$ 下相对于决策类产生的粗集边界所引起的知识不确定; 式(1)等号右端第2项 $-\sum_{j=1}^m \frac{|S_B(x_j)|}{|U|} \text{lb} \frac{1}{|S_B(x_j)|}$ 刻画了在相容关系 $S_{\text{IM}}(B)$ 下产生的知识块大小引起的不确定。

在最优约简结果中, 如果去掉某一属性, 则随着分类能力的下降, 会导致不完备边界粗糙熵的增加, 因此, 用去掉某一条测试属性后不完备边界粗糙熵的增加量定义该属性的决策重要度。

定义7 给定 $I_{\text{DS}} = (U, C \cup \{d\}, V, f)$, $* \in V$, $B \subseteq C$, 对于 $\forall a \in B$, 属性 a 的决策重要度为 $s(a) = E_{\text{IB}}(B \setminus \{a\}) - E_{\text{IB}}(B)$ (2)

对式(2)定义的决策重要度经过归一化处理, 就可以得到该测试属性的权重:

$$\omega_a = s(a) / \sum_{x \in B} s(x) \quad (3)$$

1.3 不完备并发故障诊断决策系统中的属性值频率计算

进行并发故障诊断时, 合成的证据群来源于不完备决策表的相对约简, 其中每一条测试属性作为证据时的信度赋值函数同样由决策表给出, 通过对属性值频率的计算获得。

定义8 在不完备并发故障诊断决策系统 $I_{\text{DS}} = (U, C \cup \{d\}, V, f)$, $* \in V$ 中, 对于 $\forall a \in C$, 在决策属性 $d_j \in \{d\}$ 下, 在属性 a 上的属性值 $f_a(x_i)$ 频率区间为

$$p[f_a^{d_j}(x_i)] = \left[\frac{|f_a^{d_j}(x_i)|}{|V_a| - (|f_a^{d_j'}(x_i)| - |f_a^{d_j}(x_i)|)}, \frac{|f_a^{d_j'}(x_i)|}{|V_a| + (|f_a^{d_j'}(x_i)| - |f_a^{d_j}(x_i)|)} \right]$$

式中: $|f_a^{d_j}(x_i)|$ 为决策表中某一决策属性 d_j 下与属性 a 中属性值为 $f_a(x_i)$ 相对应的对象个数; $|f_a^{d_j'}(x_i)|$ 为属性值为 $f_a(x_i)$ 所对应的对象个数与属性值为“*”所对应的对象个数之和; $|V_a|$ 为决策表中所有决策属性 d 下与属性 a 列上属性值为 $f_a(x_i)$ 相对应的对象个数; $|V_a|$ 为决策表中所有决策属性下与属性 a 中属性值为 $f_a(x_i)$ 相对应的对象个数与属性值为“*”所对应的对象个数之和。

如果在决策属性中有 n 种取值, 即系统中存在的故障种类为 d_1, d_2, \dots, d_n 。当属性 a 取值为 $f_a(x_i)$ 时, 其作为证据的基本信度赋值为 $m_a = \{p[f_a^{d_1}(x_i)], p[f_a^{d_2}(x_i)], \dots, p[f_a^{d_n}(x_i)]\}$

2 并发故障诊断模型与信度合成

2.1 DSMT 框架下并发故障诊断模型

DSMT 作为信息融合算法中一个新的分支, 有效地对 D-S 证据理论进行了扩展。在 DSMT 中, $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ 作为所考虑融合问题的辨识框架, 其是包含 n 个完备命题的有限集合。在 DSMT 框架下, 定义超幂集 D^θ , 其是由 Θ 中的命题, 以及其通过并、交运算组成的所有符合命题的集合。可表示为: ① $\emptyset, \theta_1, \theta_2, \dots, \theta_n \in D^\theta$; ② 如果 $A, B \in D^\theta$, 那么 $A \cap B \in D^\theta, A \cup B \in D^\theta$; ③ 除了①和②中包含的命题, 再无其他命题属于 D^θ 。

框架 Θ 中的命题 $\theta_i (i=1, 2, \dots, n)$ 组成了表述融合问题所有特征的有限命题集合, 其超幂集空间 D^θ 组成了自由 DSMT 模型, 允许表述连续、模糊和相对概念的命题包含其中。在自由 DSMT 模型中, 不存在其他约束。但由于框架 Θ 中的某些子集可能含有相互排斥的命题, 或某些命题并不存在, 需要将约束条件添加到自由 DSMT 模型中, 以便更好地表述真实情况, 这样就构造了混合 DSMT 模型。

而 DSMT 中的 DSMT 模型为并发故障的表述提供了依据。在进行并发故障诊断时, $\Theta = \{\theta_1,$

$\theta_2, \dots, \theta_i, \dots, \theta_n$ 是一个包含 n 个单故障模式的有限集合, θ_i 表示其中的一种故障模式。

通过超幂集 D^θ 组成的自由 DSm 模型表示整个系统所有可能的故障模式。相比较 D-S 证据理论框架下必须要求元素间互斥, 由于 D^θ 中包含了 \cap 的运算, 所以很好地解决了原有理论无法表示并发故障的问题。例如, 当系统有 2 种单故障模式, $\Theta = \{\theta_1, \theta_2\}$, 则系统所有可能的故障模式, 即此时的自由 DSm 模型可以表示为: $\{\emptyset, \theta_1, \theta_2, \theta_1 \cap \theta_2, \theta_1 \cup \theta_2\}$, \emptyset 表示系统无故障, θ_1, θ_2 表示 2 种单发故障, $\theta_1 \cap \theta_2$ 表示 2 种故障并发, $\theta_1 \cup \theta_2$ 表示故障存在, 但无法进行判断。

当系统的单故障模式种类为 n , 即 $|\Theta| = n$ 时, 系统所有可能的故障模式, 即 D^θ 的势可达到 2^{2^n} 的大多数。这是由于 D^θ 的产生与列举单调布尔函数集合的 Dedekind 问题密切相关, 当 $n \geq 1$ 时, $|D^\theta|$ 满足 Dedekind 序列: 1, 2, 5, 19, 167, 7580, ...。

但在实际故障诊断过程中, 并非所有可能的故障模式都会发生, 在构建诊断模型时应将相应约束条件考虑进来, 减少 $|D^\theta|$ 中故障模式的种类, 降低计算量。由于约束条件的引入, 自由 DSm 模型此时变为混合 DSm 模型。例如, 当系统存在单故障为 3 种时, $\Theta = \{\theta_1, \theta_2, \theta_3\}$, 则自由 DSm 模型下系统全部可能的故障模式为 19 种, 即 $\emptyset, \theta_1, \theta_2, \theta_3, \theta_1 \cap \theta_2, \theta_1 \cap \theta_3, \theta_2 \cap \theta_3, \theta_1 \cap \theta_2 \cap \theta_3, \theta_1 \cup (\theta_2 \cap \theta_3), \theta_2 \cup (\theta_1 \cap \theta_3), \theta_3 \cup (\theta_1 \cap \theta_2), \theta_1 \cap (\theta_2 \cup \theta_3), \theta_2 \cap (\theta_1 \cup \theta_3), \theta_3 \cap (\theta_1 \cup \theta_2), \theta_1 \cup \theta_2, \theta_1 \cup \theta_3, \theta_2 \cup \theta_3, \theta_1 \cup \theta_2 \cup \theta_3, (\theta_1 \cap \theta_2) \cup (\theta_1 \cap \theta_3) \cup (\theta_2 \cap \theta_3)$ 。由于实际故障中不存在并集情况, 故引入非存在性约束条件, 此时, 模型为混合 DSm 模型, 系统可能故障模式为 8 种, 即 $\emptyset, \theta_1, \theta_2, \theta_3, \theta_1 \cap \theta_2, \theta_1 \cap \theta_3, \theta_2 \cap \theta_3, \theta_1 \cap \theta_2 \cap \theta_3$ 。又知故障 θ_1, θ_3 并不会同时发生, 再引入排斥性约束条件 $\theta_1 \cap \theta_3 = \emptyset$, 此时混合 DSm 模型变为: $\{\emptyset, \theta_1, \theta_2, \theta_3, \theta_1 \cap \theta_2, \theta_2 \cap \theta_3\}$, 表示系统目前可能存在无故障 (\emptyset)、3 种单发故障 ($\theta_1, \theta_2, \theta_3$) 及 2 种并发故障 ($\theta_1 \cap \theta_2, \theta_2 \cap \theta_3$)。

2.2 基于先验知识的区间信度合成规则

测试过程中存在的不确定信息导致得到的信度赋值函数为区间信度赋值函数。并发故障诊断模型作为一种混合 DSm 模型, 目前在 DSmT 框架下, Smarandache 和 Dezert^[21] 通过引入区间值运算方法, 将精确信度结构上定义的经典 DSm 组合规则和混合 DSm 组合规则推广到区间值信度结构上, 给出了 DSmT 框架下的区间证据合成规则。

用 $m_{\mu(\Theta)}^l(A)$ 和 $m_{\mu(\Theta)}^r(A)$ 表示区间的混合 DSm 组合规则与经典 DSm 组合规则得到的结果。在进行区间证据合成时, 对于所有的 $A \in D^\theta$, 混合 DSm 组合规则表示如下:

$$m_{\mu(\Theta)}^l(A) \triangleq \phi(A) (S_1^{\text{inf}}(A) + S_2^{\text{inf}}(A) + S_3^{\text{inf}}(A)) \quad (4)$$

式中:

$$S_1^{\text{inf}}(A) = \sum_{\substack{X_1, X_2, \dots, X_k \in D^\theta \\ (X_1 \cap X_2 \cap \dots \cap X_k) = A}} \prod_{i=1}^k \text{inf}(m_i^l(X_i))$$

$$S_2^{\text{inf}}(A) = \sum_{\substack{X_1, X_2, \dots, X_k \in \emptyset \\ [u=A], \forall [(u \in \emptyset) \wedge (A=I_j)]}} \prod_{i=1}^k \text{inf}(m_i^l(X_i))$$

$$S_3^{\text{inf}}(A) = \sum_{\substack{X_1, X_2, \dots, X_k \in D^\theta \\ (X_1 \cap X_2 \cap \dots \cap X_k) = \emptyset \\ (X_1 \cup X_2 \cup \dots \cup X_k) = A}} \prod_{i=1}^k \text{inf}(m_i^l(X_i))$$

以及

$$m_{\mu(\Theta)}^r(A) \triangleq \phi(A) (S_1^{\text{sup}}(A) + S_2^{\text{sup}}(A) + S_3^{\text{sup}}(A)) \quad (5)$$

式中:

$$S_1^{\text{sup}}(A) = \sum_{\substack{X_1, X_2, \dots, X_k \in D^\theta \\ (X_1 \cap X_2 \cap \dots \cap X_k) = A}} \prod_{i=1}^k \text{sup}(m_i^r(X_i))$$

$$S_2^{\text{sup}}(A) = \sum_{\substack{X_1, X_2, \dots, X_k \in \emptyset \\ [u=A] \vee [(u \in \emptyset) \wedge (A=I_j)]}} \prod_{i=1}^k \text{sup}(m_i^r(X_i))$$

$$S_3^{\text{sup}}(A) = \sum_{\substack{X_1, X_2, \dots, X_k \in D^\theta \\ (X_1 \cap X_2 \cap \dots \cap X_k) = \emptyset \\ (X_1 \cup X_2 \cup \dots \cup X_k) = A}} \prod_{i=1}^k \text{sup}(m_i^r(X_i))$$

根据式(4)和式(5), 可以得到最后区间合成结果为: $m_{\mu(\Theta)}^l(A) = [m_{\mu(\Theta)}^l(A), m_{\mu(\Theta)}^r(A)]$ 。当在自由 DSm 模型下使用经典 DSm 组合规则, 则最后合成区间的边界直接由 $S_1^{\text{inf}}(A)$ 和 $S_1^{\text{sup}}(A)$ 给出, 即 $m_{\mu(\Theta)}^l(A) = [S_1^{\text{inf}}(A), S_1^{\text{sup}}(A)]$ 。

可以看出, 整个融合过程等价于将区间基本信度赋值的上界和下界分别应用精确信度结构下的 DSm 组合规则进行合成。此时得到的结果为次优的, 因区间算术在进行区间证据合成时其本身存在的缺陷会导致结果区间过于宽泛。而且当使用混合 DSm 模型时, 可以看出在式(4)中, 第 1 个和式对应信度赋值由合取规则获得; 第 2 个和式将所有对绝对和相对空集的信度赋值, 传递到完全或相对未知集上; 第 3 个和式将相对空集的信度赋值传递到非空集合上。这表明如果故障 θ_1 与 θ_2 没有并发故障, 即 $\theta_1 \cap \theta_2 = \emptyset$ 时, 冲突信度质量 $m(\theta_1 \cap \theta_2)$ 被传递到焦元 $\theta_1 \cup \theta_2$ 上, 作

为未知信息处理。对于故障诊断问题, $\theta_1 \cup \theta_2$ 意味着诊断结果为“故障 θ_1 或故障 θ_2 发生”, 此时诊断就失去了意义, 同时会造成有用信息的丢失。

由于对区间基本信度赋值 2 个端点进行运算会割裂其内在联系, 造成结果误差, 故本文使用最优化方法替代区间运算; 同时本文认为应将冲突质量重新进行分配, 而非如混合 DSm 规则将冲突质量直接分配到并集上做未知处理, 因此, 对 DSmT 下区间证据的组合规则和证据源进行修改, 给出新的 DSmT 下区间信度合成规则, 将其称为 CDIP 规则。

2 个区间证据合成时, CDIP 方法定义如下。

定义 9 $m_1(\cdot)$ 和 $m_2(\cdot)$ 是广义幂集 G 下的区间基本信度赋值, 满足: $m_1^-(X_i) \leq m_1(X_i) \leq m_1^+(X_i)$, $1 \leq i \leq g$, $m_2^-(X_j) \leq m_2(X_j) \leq m_2^+(X_j)$, $1 \leq j \leq g$, 则通过 CDIP 方法合成后的区间基本信度赋值表示为

$m_{CDIS}(X) = [(m_1 \oplus m_2)_{CDIP}^-(X), (m_1 \oplus m_2)_{CDIP}^+(X)]$
式中: $(m_1 \oplus m_2)_{CDIP}^-(X)$, $(m_1 \oplus m_2)_{CDIP}^+(X)$ 由如下优化模型给出:

$$\begin{aligned} \max/\min [(m_1 \oplus m_2)_{CDIP}(X)] &= m_{12}(X) + \\ &\sum_{\substack{Y \in G \setminus \{X\} \\ Y \cap X = \emptyset}} \left[\frac{\omega_1 \cdot m_1(X)^2 m_2(Y)}{\omega_1 \cdot m_1(X) + \omega_2 \cdot m_2(Y)} + \right. \\ &\left. \frac{\omega_2 \cdot m_2(X)^2 m_1(Y)}{\omega_2 \cdot m_2(X) + \omega_1 \cdot m_1(Y)} \right] \\ \text{s. t. } \sum_{i=1}^g m_1(X_i) &= 1 \quad \sum_{j=1}^g m_2(X_j) = 1 \\ m_1^-(X_i) &\leq m_1(X_i) \leq m_1^+(X_i) \quad i = 1, 2, \dots, g \\ m_2^-(X_j) &\leq m_2(X_j) \leq m_2^+(X_j) \quad j = 1, 2, \dots, g \end{aligned} \quad (6)$$

其中: $m_{12}(X) = \sum_{\substack{X_i, X_j \in G \\ X_i \cap X_j = X}} m_1(X_i) m_2(X_j)$, $i, j = 1, \dots, g$

$\max/\min [(m_1 \oplus m_2 \oplus \dots \oplus m_n)_{CDIP}(X)] =$

$$\begin{aligned} &m_{12\dots n}(X) + \sum_{\substack{X_{j_1}, X_{j_2}, \dots, X_{j_n} \in G \setminus \{X\} \\ X \cap X_{j_1} \cap X_{j_2} \cap \dots \cap X_{j_n} = \emptyset \\ \{j_1, j_2, \dots, j_n\} \in P(\{1, 2, \dots, n-1\})}} \sum_{s=1}^{n-1} \frac{\left[\sum_{k_1=1}^s \omega_{i_{k_1}} \cdot m_{i_{k_1}}(X) \right] \left[\prod_{k_1=1}^s m_{i_{k_1}}(X) \prod_{k_2=s+1}^n m_{i_{k_2}}(X_{j_t}) \right]}{\sum_{k_1=1}^s \omega_{i_{k_1}} \cdot m_{i_{k_1}}(X) + \sum_{k_2=s+1}^n \omega_{i_{k_2}} \cdot m_{i_{k_2}}(X_{j_t})} \\ \text{s. t. } \sum_{j=1}^g m_i(X_j) &= 1 \\ m_i^-(X_j) &\leq m_i(X_j) \leq m_i^+(X_j) \\ j &= 1, 2, \dots, g, \forall i = 1, 2, \dots, n \end{aligned}$$

$2, \dots, g$; ω_1 和 ω_2 分别为证据 S_1 和 S_2 的权重系数, 由第 1.2 节中测试属性重要度给出。

此时, 当 $X \cap Y = \emptyset$ 时, 总冲突为 $m(X \cap Y) = m_1(X)m_2(Y) + m_2(X)m_1(Y)$ 。CDIP 方法将其视作 2 个部分, 即由 $m_1(X)$ 和 $m_2(Y)$ 引起的冲突质量 $m_1(X)m_2(Y)$ 以及由 $m_2(X)$ 和 $m_1(Y)$ 引起的冲突质量 $m_2(X)m_1(Y)$, 分别在焦元 X 与 Y 间进行分配。而分配的比例则由先验知识 ω_1, ω_2 以及 2 条证据在其上的信度赋值 $m_1(X), m_2(X), m_1(Y)$ 和 $m_2(Y)$ 给出。

在计算 $(m_1 \oplus m_2)_{CDIP}(X)$ 时, 首先通过合取规则获得 $m_{12}(X)$, 然后计算涉及焦元 X 的冲突质量分到 X 上的部分。对于 $\forall Y \in G \setminus \{X\}, Y \cap X = \emptyset$, 冲突质量 $m_1(X)m_2(Y)$ 中按照比例 $\frac{\omega_1 \cdot m_1(X)}{\omega_1 \cdot m_1(X) + \omega_2 \cdot m_2(Y)}$ 传递到焦元 X 上; 冲突质量 $m_2(X)m_1(Y)$ 中按照比例 $\frac{\omega_2 \cdot m_2(X)}{\omega_2 \cdot m_2(X) + \omega_1 \cdot m_1(Y)}$ 传递到焦元 X 上。需注意的是, 如果分配比例中分母为零, 即 $m_1(X) = m_2(Y) = 0$ 或者 $m_2(X) = m_1(Y) = 0$, 说明并无冲突存在, 则需要在式 (6) 中去掉该部分质量。

在进行多个区间证据合成时, CDIP 方法定义如下。

定义 10 $m_1(\cdot), m_2(\cdot), \dots, m_n(\cdot)$ 为 n 个广义幂集 G 下的区间基本信度赋值, 满足: $m_i^-(X_j) \leq m_i(X_j) \leq m_i^+(X_j)$, $1 \leq i \leq n, 1 \leq j \leq g$, 则通过 CDIP 方法合成后的区间基本信度赋值表示为

$m_{CDIS}(X) = [(m_1 \oplus m_2 \oplus \dots \oplus m_n)_{CDIP}^-(X), (m_1 \oplus m_2 \oplus \dots \oplus m_n)_{CDIP}^+(X)]$
式中: $(m_1 \oplus m_2 \oplus \dots \oplus m_n)_{CDIP}^-(X)$, $(m_1 \oplus m_2 \oplus \dots \oplus m_n)_{CDIP}^+(X)$ 由如下优化模型给出:

其中: $m_{12\dots n}(X) = \sum_{\substack{X_{j_1}, X_{j_2}, \dots, X_{j_n} \in G \\ X_{j_1} \cap X_{j_2} \cap \dots \cap X_{j_n} = X}} \prod_{l=1}^n m_i(X_{j_l})$, $\{i_1, i_2, \dots, i_n\} \in P^n(\{1, 2, \dots, n\})$, $P^k(\{1, 2, \dots, n\})$ 表示由 $\{1, 2, \dots, n\}$ 中的 k 个元素所组成的集合;

ω_i 为对应的证据 S_i 的权重系数。

3 基于 CDIP 规则的并发故障诊断方法

在并发故障诊断过程中,各种传感器采集现场数据,产生测试信息,反映出装备的状态,在此过程中,有些测试信息遗漏或无法获取,造成了测试信息的不完备。此时,进行故障诊断时,需要将测试信息进行处理,然后结合当前传感器信息,使用 CDIP 方法进行信息融合,给出测试结果。

1) 选取装备故障信息,构建不完备并发故障诊断决策系统 $I_{DS} = (U, C \cup \{d\}, V, f)$, $* \in V$ 。

2) 对 I_{DS} 中数据进行预处理。根据测试指标正常区间,按照“高于”、“低于”和“属于”该区间将测试数据进行离散化。

3) 使用粗糙集方法对决策表进行约简,获得约简结果 B 。

4) 若得到的约简结果不唯一,计算不完备边界粗糙熵 $E_{IB}(B_i)$,选取值最小的约简结果 B_i 。

5) 分别计算每一条测试属性的重要度 $s(a_i)$,通过计算 ω_{a_i} 得到每条证据的权重。

6) 计算 B_i 中每条证据的属性值频率 $p[f_{a_i}^{d_i}(x_i)]$,构建证据的区间信度赋值函数 m_i 。

7) 构建 DS_m 模型,在 DS_m 框架下通过 CDIP 规则对区间证据进行组合。通过合成后的区间信度赋值对故障进行判断,给出诊断结果。

4 数值实验

本节选取模拟放大电路和旋转变压器激励电路的不完备诊断数据,采用本文所提方法对其中存在的并发故障进行诊断;并采用其他 DS_mT 框架下的区间证据合成方法对获得的区间信度赋值函数进行组合,与 CDIP 方法所得到的结果进行比较,从而对该方法的有效性与适用性进行验证。

本文对比使用的 DS_mT 区间证据合成方法包括区间 DS_m 组合规则、CDI2 方法和 CDI5 方法。区间 DS_m 组合规则通过引入区间运算,将经典 DS_m 规则及混合 DS_m 规则推广到区间值信度结构上;而 CDI2 方法、CDI5 方法与 CDIP 方法同为基于最优化思想的区间证据合成方法,区别在于冲突证据在各焦元间的分配比例存在差别。以两证据合成为例,CDI2 方法给出的优化模型为

$$\max/\min [(m_1 \oplus m_2)_{CDI2}(X)] =$$

$$\sum_{\substack{X_i, X_j \in G \\ X_i \cap X_j = X}} m_1(X_i)m_2(X_j) + C(X) \frac{c_{12}(X)}{e_{12}} k_{12}$$

式中: $C(X)$ 为 X 的特征函数,当焦元 X 与冲突有关时, $C(X) = 1$, 否则, $C(X) = 0$; $c_{12}(X)$ 为区间信度质量矩阵 M 中焦元 X 所对应的列和,即 $c_{12}(X) = m_1(X) + m_2(X)$; e_{12} 为 M 中与冲突有关的非零焦元所对应的列和; k_{12} 为总的冲突质量,

$$k_{12} = \sum_{\substack{X_i, X_j \in G \\ X_i \cap X_j = \emptyset}} m_1(X_i)m_2(X_j)$$

优化模型为

$$\max/\min [(m_1 \oplus m_2)_{CDI5}(X)] = m_{12}(X) + \sum_{\substack{Y \in G \setminus \{X\} \\ Y \cap X = \emptyset}} \left[\frac{m_1(X)^2 m_2(Y)}{m_1(X) + m_2(Y)} + \frac{m_2(X)^2 m_1(Y)}{m_2(X) + m_1(Y)} \right]$$

4.1 运放电路并发故障诊断实例

本节选用三运放构成的仪表放大器电路作为并发故障诊断实例。电路主要由运放 OP07 组成,电路原理图如图 1 所示。

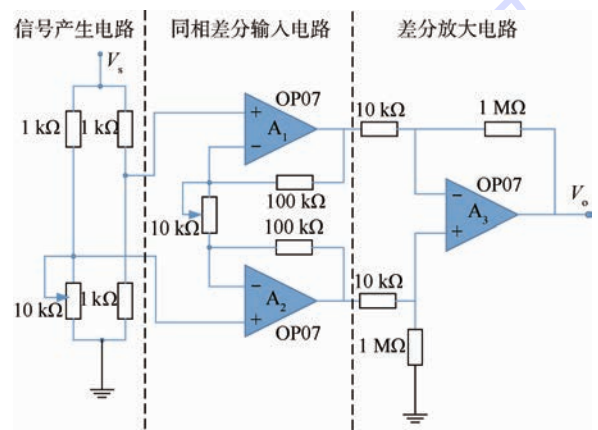


图 1 仪表放大器电路原理

Fig. 1 Principle of instrumentation amplifier circuit

在过往的故障诊断过程中,发现仪表放大器电路存在 3 种故障形式,即差分输入电路故障,记为 F_1 ;信号产生电路故障,记为 F_2 ;差分输入电路与放大电路并发故障,记为 $F_1 \cap F_3$ 。

在 3 种故障形式下,将对应的信号源 V_s 的最大、最小输入,电路的最大增益,共模抑制比及电路工作温度进行记录。其中,信号 V_s 的最大、最小输入是在信号频率 $f = 50 \text{ Hz}$,增益 $G_{AIN} = 72 \text{ dB}$, $R_p = 5 \text{ k}\Omega$ 时,即电路输出不失真时的信号源的最大、最小输入,在输入端进行采集;最大增益是在信号 $V_s = 100 \mu\text{V}$, $R_p = 5 \text{ k}\Omega$ 时,即输出不失真时可实现的电路最大增益值,在输出端进行采集。选取典型的原始测试数据在表 1 中给出。

表1 运放电路各故障下原始测试数据

Table 1 Raw test data of instrumentation amplifier circuit fault

编号	V_s 最大 输入/mV	V_s 最小 输入/mV	共模抑 制比/dB	电路工 作温度/°C	最大增 益/dB	故障 模式
1	19.5	0.61	89	50.5	119	F_1
2	10.1	1.03	20	65.1	108	F_1
3	13.9	*	102	105.9	140	F_1
4	17.8	0.90	117	90.4	106	F_1
5	*	1.12	73	93.6	117	F_1
6	14.4	0.48	120	77.4	103	F_2
7	18.2	0.77	51	81.2	84	F_2
8	18.5	0.01	93	*	40	F_2
9	13.3	0.53	105	110.0	76	F_2
10	*	0.49	81	66.3	81	F_2
11	17.6	1.07	68	72.1	93	$F_1 \cap F_3$
12	18.0	0.96	*	58.6	101	$F_1 \cap F_3$
13	12.7	1.50	44	55.9	88	$F_1 \cap F_3$
14	11.9	1.91	62	61.7	79	$F_1 \cap F_3$
15	16.5	1.00	110	88.3	96	$F_1 \cap F_3$

注: * 一缺失的测试数据。

以测试属性 $C_1 \sim C_5$ 分别表示与表1对应的5个项目,通过仿真实验确定各指标的正常区间,以此对原始数据进行离散化。因此, C_1 表示 V_s 的最大输入: 12.0 ~ 18.0 mV; C_2 表示 V_s 的最小输入: 0.50 ~ 1.00 mV; C_3 表示共模抑制比: ≥ 70 dB; C_4 表示电路工作温度: ≤ 85.0 °C; C_5 表示最大增益: 85 ~ 105 dB。其中,测试属性 C_3 、 C_4 赋值过程中,满足正常条件为“1”,不满足为“0”。由此构建诊断决策表如表2所示。

针对不完备并发故障诊断决策表,采用文献[22]的方法进行属性约简,得到的结果为 $B = \{C_2, C_5\}$ 。通过计算属性 C_2 与 C_5 的边界粗糙

表2 运放电路不完备并发故障诊断决策表

Table 2 Simultaneous fault diagnosis decision table with incomplete information of instrumentation amplifier circuit

编号	C_1	C_2	C_3	C_4	C_5	故障模式
1	2	1	1	1	2	F_1
2	0	2	0	1	2	F_1
3	1	*	1	0	2	F_1
4	1	1	1	0	2	F_1
5	*	2	1	0	2	F_1
6	1	0	1	1	1	F_2
7	2	1	0	1	0	F_2
8	2	0	1	*	0	F_2
9	1	1	1	0	0	F_2
10	*	0	1	1	0	F_2
11	1	2	0	1	1	$F_1 \cap F_3$
12	1	1	*	1	1	$F_1 \cap F_3$
13	1	2	0	1	1	$F_1 \cap F_3$
14	0	2	0	1	0	$F_1 \cap F_3$
15	1	1	1	0	1	$F_1 \cap F_3$

熵,得到其作为证据的权重为: $\omega(a_2) = 0.364$, $\omega(a_5) = 0.636$ 。然后通过计算属性 C_2 与 C_5 的属性频率区间,得到信度赋值函数(见表3)。

表3 运放电路区间信度赋值

Table 3 Interval-value belief assignment of instrumentation amplifier circuit

i	C_i	$m_i(F_1)$	$m_i(F_2)$	$m_i(F_1 \cap F_3)$
	0	[0, 0.250]	[0.750, 1.000]	0
2	1	[0.333, 0.429]	[0.286, 0.333]	[0.286, 0.333]
	2	[0.400, 0.500]	0	[0.500, 0.600]
5	0	0	0.800	0.200
	1	0	0.200	0.800
	2	1.000	0	0

选取实测的4组故障数据,对诊断结果进行验证。对数据离散化后,按照对应约简取得实例为: $T_1 = \{0, 2\}$, $T_2 = \{1, 2\}$, $T_3 = \{1, 1\}$, $T_4 = \{2, 2\}$ 。

在进行区间证据合成前,需要根据系统存在的故障模式构建 DS_m 模型。由于在电路中存在单发故障 F_1 、 F_2 以及并发故障 $F_1 \cap F_3$,故电路的诊断框架为 $\Theta = \{F_1, F_2, F_3\}$,由于不存在其他并发故障,因此需引入排斥性约束 $F_1 \cap F_2 = \emptyset$, $F_2 \cap F_3 = \emptyset$, $F_1 \cap F_2 \cap F_3 = \emptyset$ 。在此框架下,用4种方法分别对区间证据进行组合,得到的结果如表4所示。

通过表4中可以看出,样本 T_1 中2个指标的信息存在严重冲突。在面对高冲突区间证据合成时,区间 DS_m 组合规则会将其冲突部分分配到相应的并集上作为未知信息进行处理。例如,其将 $m(F_1 \cap F_2)$ 的质量转移到了焦元 $F_1 \cup F_2$ 上,而 $F_1 \cup F_2$ 由于包含故障模式 F_1 与 F_2 ,因此无法做出诊断。而将冲突在焦元间进行分配的 CDI2 方法、CDI5 方法与 CDIP 方法都得到了正确的结果。但由于 CDIP 方法赋予 C_5 更多的权重,使得结果的准确度明显高于 CDI2 方法与 CDI5 方法。而在样本 T_4 中,由于决策 F_1 实际包含2种故障模式,单发故障 F_1 和并发故障 $F_1 \cap F_3$,即 $F_1 = F_1 \cup (F_1 \cap F_3)$,说明证据间完全不存在冲突,在这种不涉及冲突质量分配的情况下,所有4种方法基于合取原则得到了同样的结果。在样本 T_2 中,2条证据相一致,都为焦元 F_1 赋最大值,从合成结果来看,4种方法都得到了正确的结论。由于 C_5 显著支持 F_1 ,而 CDI2 方法、CDI5 方法认为2条证据重要度相同,导致结果准确度低于 CDIP 方法。这点在样本 T_3 上也有所体现。

表4 运放电路合成后诊断结果

Table 4 Results of fault diagnosis through combination of instrumentation amplifier circuit

方法	$T_1 = \{0, 2\}$			
	$m(F_1)$	$m(F_2)$	$m(F_1 \cap F_3)$	$m(F_1 \cup F_2 \cup F_3)$
区间 DSm 规则	[0,0.250]	[0,0]	[0,0]	[0.750,1.000]
CDI2 方法	[0.500,0.719]	[0.281,0.500]	[0,0]	
CDI5 方法	[0.500,0.679]	[0.321,0.500]	[0,0]	
本文方法	[0.652,0.786]	[0.214,0.348]	[0,0]	
方法	$T_2 = \{1, 2\}$			
	$m(F_1)$	$m(F_2)$	$m(F_1 \cap F_3)$	$m(F_1 \cup F_2 \cup F_3)$
区间 DSm 规则	[0.333,0.429]	[0,0]	[0.286,0.333]	[0.286,0.333]
CDI2 方法	[0.556,0.632]	[0.041,0.055]	[0.327,0.388]	
CDI5 方法	[0.584,0.650]	[0.064,0.083]	[0.286,0.333]	
本文方法	[0.617,0.676]	[0.038,0.050]	[0.286,0.333]	
方法	$T_3 = \{1, 1\}$			
	$m(F_1)$	$m(F_2)$	$m(F_1 \cap F_3)$	$m(F_1 \cup F_2 \cup F_3)$
区间 DSm 规则	[0,0]	[0.057,0.067]	[0.495,0.610]	[0.352,0.419]
CDI2 方法	[0.067,0.080]	[0.147,0.173]	[0.751,0.781]	
CDI5 方法	[0.042,0.058]	[0.168,0.195]	[0.755,0.781]	
本文方法	[0.031,0.046]	[0.166,0.186]	[0.776,0.795]	
方法	$T_4 = \{2, 2\}$			
	$m(F_1)$	$m(F_2)$	$m(F_1 \cap F_3)$	$m(F_1 \cup F_2 \cup F_3)$
区间 DSm 规则	[0.400,0.500]	[0,0]	[0.500,0.600]	[0,0]
CDI2 方法	[0.400,0.500]	[0,0]	[0.500,0.600]	
CDI5 方法	[0.400,0.500]	[0,0]	[0.500,0.600]	
本文方法	[0.400,0.500]	[0,0]	[0.500,0.600]	

4.2 激励发生电路并发故障诊断实例

本节选用某航电系统的旋转变压器激励发生电路作为诊断验证对象,该电路用于产生旋转变压器励磁绕组所需要的正弦信号,为旋转变压器运行提供激励。电路原理图如图2所示。

由图2可以看出,该电路主要由电源模块、频率控制模块、正弦信号产生模块以及幅值调理及驱动能力调节模块组成。在之前的测试记录中共有9个测试项目,包括:信号频率、信号幅值、正弦模块输入电压值、信号频率稳定度、

+15V电压值、-15V电压值、电路工作温度、+5V电压值和+10V电压值。其中,信号频率、幅值及频率稳定度数据在整体电路的输出端进行采集;频率控制模块输出电压在ICL8038的输入端进行采集;而各电压值在电源模块相应输出端进行采集。各故障模式下的原始测试数据如表5所示。

以测试属性 $C_1 \sim C_9$ 分别表示9个测试项目,根据对产生的正弦信号的要求以及各芯片的正常范围,对原始测试数据进行离散化。

离散化区间依次为: C_1 表示输出信号频率: 1610~1650 Hz; C_2 表示输出信号幅值: 13.1~14.5 V; C_3 表示正弦模块的输入电压值: 2.2~2.8 V; C_4 表示输出信号频率稳定度: $\leq 10\%$; C_5 表示+15V电压输出: +13.5~+16.5 V; C_6 表示-15V电压输出: -16.5~-13.5 V; C_7 表示电路工作温度: $\leq 85^\circ\text{C}$; C_8 表示+5V电压输出: +4.5~+5.5 V; C_9 表示+10V电压输出: +9.5~+10.5 V。按照“高于”、“低于”和“属于”该区间为测试属性赋值“2”、“1”和“0”;在测试属性 C_4 、 C_7 赋值过程中,满足正常条件为“1”,不满足为“0”。以此构建诊断决策表如表6所示。

对该决策表约简后可得到2个约简结果, $B_1 = \{C_1, C_2, C_3, C_5\}$, $B_2 = \{C_1, C_2, C_5, C_6\}$, 通过计算这2个约简的不完备边界粗糙熵可得:

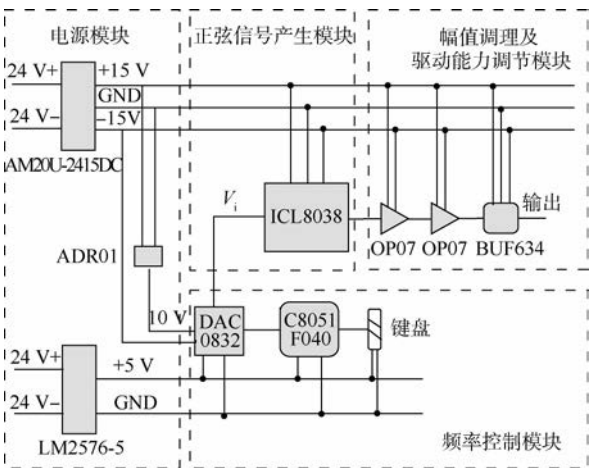


图2 旋转变压器激励发生电路原理

Fig. 2 Principle of rotating transformer excitation generating circuit

$E_{IB}(B_1) = 3.242, E_{IB}(B_2) = 3.576$, 故选择约简 $B_1 = \{C_1, C_2, C_3, C_5\}$ 作为最后的结果。分别对约简中每条属性的重要度进行计算, 从而得到权重 $w(a_1) = 0.312, w(a_2) = 0.218, w(a_3) = 0.220, w(a_5) = 0.250$; 计算属性值频率区间获得约简中每条属性的信度赋值(见表7)。

由于在被诊断电路中共存在6种故障模式, 其中4种为单发故障, 2种并发故障 $F_1 \cap F_2$ 和 $F_3 \cap F_4$ 。故诊断系统辨识框架为 $\Theta = \{F_1, F_2, F_3, F_4\}$, 同时引入排斥性约束 $F_1 \cap F_3, F_1 \cap F_4, F_2 \cap F_3, F_2 \cap F_4$, 这同时意味着系统存在排斥性约束 $F_1 \cap F_2 \cap F_3 = \emptyset, F_1 \cap F_2 \cap F_4 = \emptyset, F_1 \cap F_3 \cap F_4 = \emptyset, F_2 \cap F_3 \cap F_4 = \emptyset, F_1 \cap F_2 \cap F_3 \cap F_4 = \emptyset$ 。选取实测的4组故障数据作为验证实例, 分别为: $T_1 = \{2, 1, 2, 1\}, T_2 = \{1, 0, 1, 0\}, T_3 = \{2, 1, 0, 0\}, T_4 = \{1, 1, 2, 2\}$ 。采用4种方法对区间信度进行合成, 得到的结果如表8所示。

表5 激励发生电路各故障下原始测试数据

Table 5 Raw test data of excitation generating circuit fault

编号	信号频率/幅值/Hz	信号输入电压/V	频率/稳定度/%	+15V -15V 电压/V		电路工作温度/°C	+5V +10V 电压/V		故障模式	
				V	V		V	V		
1	1620	13.8	3.0	5.0	*	-13.6	*	5.8	10.1	F_1
2	*	9.0	3.3	9.3	18.9	-12.7	105	5.5	*	F_1
3	1630	13.2	*	7.9	15.1	-15.9	70	*	9.0	F_1
4	1600	13.8	3.0	3.9	16.8	-16.5	72	6.6	11.9	F_1
5	1640	14.0	2.1	*	14.0	*	*	6.0	*	F_1
6	1630	*	4.1	7.2	13.9	-16.2	80	5.0	9.3	F_1
7	1440	13.1	*	14.0	17.2	-13.1	89	5.6	10.6	F_1
8	1260	13.6	2.6	2.9	15.5	*	91	6.1	10.3	F_2
9	1020	12.9	2.6	4.4	17.0	-14.5	69	4.9	*	F_2
10	*	13.0	2.8	*	13.8	-13.4	88	5.0	10.0	F_2
11	1580	14.2	2.5	12.9	14.2	-15.9	102	6.1	10.8	F_2
12	1530	13.8	*	8.0	16.2	-16.9	85	5.0	10.1	F_2
13	1690	10.9	1.1	10.5	*	-13.6	86	6.9	9.4	F_2
14	1600	14.5	2.1	4.1	15.0	-19.0	72	*	*	F_2
15	*	11.7	2.6	8.9	12.2	-11.3	91	5.6	8.4	F_3
16	1610	14.6	2.8	17.2	*	-10.9	*	4.3	9.5	F_3
17	1660	9.8	2.5	11.5	10.0	-17.9	*	4.6	*	F_3
18	1650	12.9	1.9	5.4	*	-16.7	93	5.7	9.8	F_3
19	1650	13.0	*	3.3	13.2	*	70	*	9.1	F_3
20	1620	*	2.2	8.3	18.4	-17.7	74	6.0	*	F_3
21	2640	12.1	2.0	8.2	13.5	-16.9	77	6.1	9.9	F_3
22	1730	13.9	2.1	14.4	*	-15.0	*	5.8	10.3	F_4
23	*	14.2	2.4	6.5	13.2	-18.2	80	4.1	10.6	F_4
24	1690	13.1	2.8	3.1	*	-17.8	*	*	*	F_4
25	1810	*	3.3	*	16.7	-16.9	75	5.6	11.0	F_4
26	1670	12.5	3.0	9.4	15.5	-17.3	*	3.0	*	F_4
27	1700	14.0	*	*	14.3	*	70	5.1	10.1	F_4
28	*	13.1	1.8	2.5	12.0	-19.0	88	6.0	9.8	F_4

注: * 一缺失的测试数据; $F_1 \sim F_4$ —频率控制模块故障、幅值调制及驱动能力调节模块故障、电源模块故障及正弦信号产生模块故障。

在本例中, 根据系统存在的并发故障模式可知, 决策表中的决策属性实际应为: $F_1 = F_1 \cup (F_1 \cap F_2), F_2 = F_2 \cup (F_1 \cap F_2), F_3 = F_3 \cup (F_3 \cap F_4), F_4 = F_4 \cup (F_3 \cap F_4)$ 。由于模型中无非存在性约束, 在使用区间 DSm 规则进行证据合成时, 其将冲突部分质量作为未知转移为并集质量 $m(\cup)$, 由于此部分对于诊断无意义, 故不在表8中列出。可以看出, 区间 DSm 规则在验证样本上得到的诊断结果大多无效, 这主要因为其结果中焦元过多且质量分布过于松散, 而大量的冲突质量作为无意义信息处理, 导致在给定焦元上的信度赋值过低, 从而无法做出判断。而 CD12 方法、CD15 方法和 CDIP 方法最终都得到了正确的结果, 但与第4.1节中的结果相比, 得到的结果区间偏大, 这是由不完备并发故障诊断决策表中缺失数据过多, 使得通过计算属性值频率获得的区间信度范围较大所导致。

表6 激励发生电路不完备并发故障诊断决策表

Table 6 Simultaneous fault diagnosis decision table with incomplete information of excitation generating circuit

编号	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	故障模式
1	1	1	2	1	*	1	*	2	1	F_1
2	*	0	2	1	2	2	0	1	*	F_1
3	1	1	*	1	1	1	1	*	0	F_1
4	0	1	2	1	2	1	1	2	2	F_1
5	1	1	0	*	1	*	*	2	*	F_1
6	1	*	2	1	1	1	1	1	0	F_1
7	0	1	*	0	2	2	0	2	2	F_1
8	0	1	1	1	1	*	0	2	1	F_2
9	0	0	1	1	2	1	1	1	*	F_2
10	*	0	1	*	1	2	0	1	1	F_2
11	0	1	1	0	1	1	0	2	2	F_2
12	0	1	*	1	1	0	1	1	1	F_2
13	2	0	0	0	*	1	0	2	0	F_2
14	0	1	0	1	1	0	1	*	*	F_2
15	*	0	1	1	0	2	0	2	0	F_3
16	1	2	1	0	*	2	*	0	1	F_3
17	2	0	1	0	0	0	*	1	*	F_3
18	1	0	0	1	*	0	0	2	1	F_3
19	1	0	*	1	0	*	1	*	0	F_3
20	1	*	1	1	2	0	1	2	*	F_3
21	1	0	0	1	1	0	1	2	1	F_3
22	2	1	0	0	*	1	*	2	1	F_4
23	*	1	1	1	0	0	1	0	2	F_4
24	2	1	1	1	*	0	*	*	*	F_4
25	2	*	2	*	2	0	1	2	2	F_4
26	2	0	2	1	1	0	*	0	*	F_4
27	2	1	*	*	1	*	1	1	1	F_4
28	*	1	0	1	0	0	0	2	1	F_4

表7 激励发生电路区间信度赋值表

Table 7 Interval-value belief assignment of excitation generating circuit

i	C_i	$m_i(F_1)$	$m_i(F_2)$	$m_i(F_3)$	$m_i(F_4)$
1	0	[0.182, 0.375]	[0.455, 0.750]	[0, 0.125]	[0, 0.222]
	1	[0.308, 0.500]	[0, 0.100]	[0.385, 0.600]	[0, 0.182]
	2	[0, 0.125]	[0.091, 0.250]	[0.091, 0.250]	[0.500, 0.778]
2	0	[0.083, 0.182]	[0.231, 0.300]	[0.417, 0.545]	[0.083, 0.182]
	1	[0.313, 0.400]	[0.235, 0.286]	[0, 0.067]	[0.313, 0.400]
	2	[0, 0.500]	[0, 0]	[0.333, 1.000]	[0, 0.500]
3	0	[0.100, 0.333]	[0.182, 0.375]	[0.182, 0.375]	[0.182, 0.375]
	1	[0, 0.167]	[0.286, 0.455]	[0.286, 0.455]	[0.143, 0.273]
	2	[0.444, 0.750]	[0, 0.143]	[0, 0.143]	[0.200, 0.429]
5	0	[0, 0.167]	[0, 0.167]	[0.333, 0.714]	[0.222, 0.571]
	1	[0.176, 0.308]	[0.294, 0.462]	[0.063, 0.214]	[0.188, 0.357]
	2	[0.300, 0.666]	[0.100, 0.333]	[0.111, 0.429]	[0, 0.286]

表8 激励发生电路合成后诊断结果

Table 8 Results of fault diagnosis through combination of excitation generating circuit

方法	$T_1 = \{2, 1, 2, 1\}$					
	$m(F_1)$	$m(F_2)$	$m(F_3)$	$m(F_4)$	$m(F_1 \cap F_2)$	$m(F_3 \cap F_4)$
区间 DSm 规则	[0, 0.012]	[0, 0.005]	[0, 0.001]	[0.006, 0.048]	[0.010, 0.161]	[0.003, 0.109]
CDI2 方法	[0.143, 0.370]	[0.062, 0.180]	[0.002, 0.063]	[0.144, 0.475]	[0.091, 0.483]	[0.019, 0.293]
CDI5 方法	[0.111, 0.463]	[0.028, 0.177]	[0, 0.028]	[0.152, 0.604]	[0.074, 0.481]	[0.007, 0.268]
本文方法	[0.105, 0.451]	[0.028, 0.173]	[0, 0.028]	[0.157, 0.630]	[0.069, 0.474]	[0.008, 0.283]
方法	$T_2 = \{1, 0, 1, 0\}$					
	$m(F_1)$	$m(F_2)$	$m(F_3)$	$m(F_4)$	$m(F_1 \cap F_2)$	$m(F_3 \cap F_4)$
区间 DSm 规则	[0, 0.003]	[0, 0.002]	[0.015, 0.106]	[0, 0.005]	[0, 0.055]	[0.030, 0.420]
CDI12	[0.016, 0.118]	[0.057, 0.185]	[0.218, 0.553]	[0.029, 0.132]	[0.012, 0.243]	[0.141, 0.590]
CDI15	[0.012, 0.141]	[0.079, 0.274]	[0.117, 0.618]	[0.008, 0.134]	[0.021, 0.276]	[0.094, 0.618]
本文方法	[0.012, 0.150]	[0.074, 0.251]	[0.120, 0.621]	[0.008, 0.132]	[0.023, 0.286]	[0.095, 0.620]
方法	$T_3 = \{2, 1, 0, 0\}$					
	$m(F_1)$	$m(F_2)$	$m(F_3)$	$m(F_4)$	$m(F_1 \cap F_2)$	$m(F_3 \cap F_4)$
区间 DSm 规则	[0, 0.003]	[0, 0.004]	[0, 0.004]	[0.006, 0.067]	[0, 0.053]	[0.031, 0.391]
CDI12	[0.089, 0.235]	[0.074, 0.207]	[0.018, 0.125]	[0.150, 0.442]	[0.012, 0.249]	[0.150, 0.552]
CDI15	[0.089, 0.294]	[0.069, 0.243]	[0.001, 0.123]	[0.046, 0.442]	[0.021, 0.260]	[0.132, 0.660]
本文方法	[0.084, 0.279]	[0.064, 0.235]	[0.001, 0.130]	[0.046, 0.454]	[0.019, 0.254]	[0.143, 0.679]
方法	$T_4 = \{1, 1, 2, 2\}$					
	$m(F_1)$	$m(F_2)$	$m(F_3)$	$m(F_4)$	$m(F_1 \cap F_2)$	$m(F_3 \cap F_4)$
区间 DSm 规则	[0.013, 0.100]	[0, 0.001]	[0, 0.002]	[0, 0.009]	[0.017, 0.265]	[0.003, 0.138]
CDI12	[0.213, 0.543]	[0.047, 0.135]	[0.004, 0.109]	[0.067, 0.267]	[0.090, 0.492]	[0.021, 0.379]
CDI15	[0.175, 0.610]	[0.017, 0.140]	[0.001, 0.147]	[0.057, 0.358]	[0.062, 0.478]	[0.020, 0.432]
本文方法	[0.174, 0.613]	[0.018, 0.133]	[0.002, 0.167]	[0.051, 0.343]	[0.060, 0.478]	[0.023, 0.450]

确的合成结果。

随着需要合成的测试属性数量增多,故障模式的增多,方法会出现计算量增大的问题,需要对方法进行进一步的优化。

参考文献 (References)

- [1] KRYSZKIEWICZ M. Rough set approach to incomplete information systems [J]. Information Sciences, 1998, 12 (1-4): 39-49.
- [2] 陈利安,肖明清,赵鑫.基于粗糙集与信息熵的不完备测试信息条件下故障诊断[J].振动与冲击, 2012, 31 (22): 24-28.

5 结论

本文在综合分析粗糙集及 DSmT 理论上,提出了可以在诊断信息不完备条件下进行的并发故障诊断的方法。经验证表明:

1) 可以有效地对不完备信息,以及在 DSmT 框架下对并发故障进行表示。

2) 通过对测试信息中的特征进行提取,度量其对决策的重要程度。

3) 在对区间证据进行合成时,有效地融入了证据特征,可以获得较之前区间合成方法更为精

- CHEN L A, XIAO M Q, ZHAO X. Fault diagnosis under condition of incomplete test information based on rough set and information entropy[J]. Journal of Vibration and Shock, 2012, 31(22): 24-28 (in Chinese).
- [3] 王珉,胡莺庆,秦国军.基于不完备信息的直升机传动系统故障诊断规则提取方法[J].振动与冲击,2011,30(12): 185-190.
- WANG M, HU N Q, QIN G J. Method of rule extraction for fault diagnosis of a helicopter power train based on incomplete information[J]. Journal of Vibration and Shock, 2011, 30(12): 185-190 (in Chinese).
- [4] LU X, HUANG W, WANG W, et al. Multi-interpretations valued tolerance relation model of incomplete information in fault diagnosis system[C]//2014 11th World Congress on Intelligent Control and Automation (WCICA). Piscataway, NJ: IEEE Press, 2014:6035-6040.
- [5] 杨召,肖明清,胡斌,等.测试信息不完备条件下航空电子装备故障诊断[J].计算机应用,2013,33(A01):46-47.
- YANG Z, XIAO M Q, HU B, et al. Fault diagnosis under condition of incomplete test information based on rough set and information entropy[J]. Journal of Computer Applications, 2013, 33(A01):46-47 (in Chinese).
- [6] 胡雷刚,肖明清,禹航,等.不完备信息条件下的航空发动机故障诊断方法[J].振动、测试与诊断,2012,32(6): 903-908.
- HU L G, XIAO M Q, YU H, et al. Aero engine fault diagnosis method under incomplete information[J]. Journal of Vibration, Measurement & Diagnosis, 2012, 32(6): 903-908 (in Chinese).
- [7] RONG Y, ZHAO J, WANG J, et al. Fault diagnosis of transformer based on incomplete information system rule extraction and Petri nets[J]. Power System Protection and Control, 2009, 37(18):1-4.
- [8] 赵佰亭.混合决策系统的粗集模型及在转台故障诊断中的应用[D].哈尔滨:哈尔滨工业大学,2010:101-110.
- ZHAO B T. Rough set model of hybrid decision system and application in turntable fault diagnosis[D]. Harbin: Harbin Institute of Technology, 2010:101-110 (in Chinese).
- [9] DONG H, LI X, WEI Z. Substation fault diagnosis based on rough sets and grey relational analysis[J]. Telkomnika Indonesian Journal of Electrical Engineering, 2014, 12(2):1162-1168.
- [10] GAO Q, YIN C, DONG G. A novel fault diagnosis method for boiler drum water level based on rough sets and evidence theory [C]// 2012 3rd International Conference on Intelligent Control and Information Processing (ICICIP). Piscataway, NJ: IEEE Press, 2012:488-492.
- [11] WANG M, HU N, QIN G. A method for rule extraction based on granular computing: Application in the fault diagnosis of a helicopter transmission system[J]. Journal of Intelligent & Robotic Systems, 2013, 71(3-4):445-455.
- [12] 彭俏,郭立峰,马杰.基于混合神经网络的核动力装置并发故障诊断研究[J].核动力工程,2013,33(6):30-34.
- PENG Q, GUO L F, MA J. Intercurrent fault diagnosis of nuclear power plants based on hybrid artificial neural network[J]. Nuclear Power Engineering, 2013, 33(6):30-34 (in Chinese).
- [13] VONG C M, WONG P K, IP W F, et al. Simultaneous-fault diagnosis of automotive engine ignition systems using prior domain knowledge and relevance vector machine [J]. Mathematical Problems in Engineering, 2013, 2013:1-19.
- [14] CAI B, LIU Y, FAN Q, et al. Multi-source information fusion based fault diagnosis of ground-source heat pump using Bayesian network[J]. Applied Energy, 2014, 114:1-9.
- [15] DEZERT J. Foundations for a new theory of plausible and paradoxical reasoning[J]. Information and Security Journal, 2002, 9:13-57.
- [16] 李植良,徐晓滨,文成林.基于扩展证据理论的并发故障诊断方法[J].杭州电子科技大学学报,2009,28(6):107-110.
- LI Z L, XU X B, WEN C L. Simultaneous fault diagnosis method based on extended evidence theory[J]. Journal of Hangzhou Dianzi University, 2009, 28(6):107-110 (in Chinese).
- [17] 徐晓滨,文成林,蒋海娜,等.基于随机集理论的并发故障诊断信息融合方法[J].仪器仪表学报,2010,31(2): 334-340.
- XU X B, WEN C L, JIANG H N, et al. Information fusion method of simultaneous fault diagnosis based on random set theory [J]. Chinese Journal of Scientific Instrument, 2010, 31(2): 334-340 (in Chinese).
- [18] 郭清,夏虹,韩文伟. DSMT 的主冷却剂泵并发故障融合方法分析[J].哈尔滨工业大学学报,2014,46(9):111-115.
- GUO Q, XIA H, HAN W W. Concurrent fault information fusion methods of main coolant pump based on DSMT[J]. Journal of Harbin Institute of Technology, 2014, 46(9):111-115 (in Chinese).
- [19] WU X, LIU Y, LIU Z, et al. Research on simultaneous faults diagnosis for nuclear power plant based on dezert-smarandache theory [C] // 2013 21st International Conference on Nuclear Engineering. New York: ASME, 2013:23-29.
- [20] REN J, JIANG B, ZHAO F, et al. Simultaneous fault diagnosis of the reactor coolant system based on the DSM evidence theory [C] // 2013 21st International Conference on Nuclear Engineering. New York: ASME, 2013:31-34.
- [21] SMARANDACHE F, DEZERT J. An introduction to DSMT theory of plausible, paradoxical, uncertain, and imprecise reasoning for information fusion [J]. Octagon Mathematical Magazine, 2007, 15(2):681-722.
- [22] DAI J, WANG W, TIAN H, et al. Attribute selection based on a new conditional entropy for incomplete decision systems [J]. Knowledge-Based Systems, 2013, 39:207-213.

作者简介:

孙伟超 男,博士,助理研究员。主要研究方向:智能故障诊断。

Tel.: 18653579273

E-mail: ben_phoenix@163.com

许爱强 男,博士,教授,博士生导师。主要研究方向:复杂电子装备故障诊断。

Tel.: 0535-6635836

E-mail: xuaq6342@yahoo.com.cn

Diagnosis method of simultaneous fault with incomplete information

SUN Weichao¹, XU Aiqiang^{2,*}, LI Wenhai²

(1. Graduate Management Brigade, Naval Aeronautical Engineering Institute, Yantai 264001, China;

2. Department of Scientific Research, Naval Aeronautical Engineering Institute, Yantai 264001, China)

Abstract: Test information is often incomplete in the fault diagnosis process of equipment. And simultaneous fault diagnosis process is more difficult at the same time. In response to this situation, the current methods of fault diagnosis with incomplete test information and of simultaneous fault diagnosis are studied firstly. Then we define incomplete fault diagnosis decision system to describe incomplete test information. And it defines incomplete boundary rough entropy to measure the level of uncertainty in the system and assign the importance of each attribute. Meanwhile, the method to calculate frequency of attributes' value under incomplete condition is proposed. In order to diagnose simultaneous fault, the paper constructs the diagnosis model under DS_mT framework, and proposes a new combination rule of interval-value belief to overcome the shortcomings of previous methods. Finally, the validity and applicability of the method are proved by two equipment fault diagnosis examples.

Key words: simultaneous fault diagnosis; DS_mT theory; interval-value belief combination; rough set theory; boundary rough entropy

Received: 2015-06-30; **Accepted:** 2015-09-18; **Published online:** 2016-01-08 15:17

URL: www.cnki.net/kcms/detail/11.2625.V.20160108.1517.001.html

Foundation item: Pre Research Fund for Weapons and Equipment of PLA (9140A27020214JB14436)

* **Corresponding author.** Tel. : 0535-6635836 E-mail: xuaq6342@yahoo.com.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0471

泡沫子弹撞击刚性靶板的理论模型



刘虎, 刘华*, 杨嘉陵

(北京航空航天大学 航空科学与工程学院, 北京 100083)

摘 要: Taylor 撞击常用于测试材料的动态屈服强度。泡沫压缩密度与压缩应变的关系对泡沫子弹 Taylor 撞击的理论分析起着重要的作用。本文通过引入塑性泊松比, 得到了泡沫压缩密度与压缩应变的准确关系, 进一步建立了可压缩泡沫子弹撞击刚性靶板的理论模型。当塑性泊松比为常数时, 文中建立的泡沫受压后密度比的一阶泰勒展开式可以退化到已有模型。当塑性泊松比为塑性应变和子弹相对密度的函数时, 相对密度会影响泡沫子弹的冲击响应历程及最终变形, 但对撞击持续时间影响较小。同时, 初始速度也会影响子弹最终变形和撞击持续时间。本文的理论工作可以为分析泡沫材料的动态力学行为提供有益指导。

关键词: Taylor 撞击; 泡沫材料; 塑性; 塑性泊松比; 刚性

中图分类号: O344.2

文献标识码: A **文章编号:** 1001-5965(2016)07-1461-08

泡沫材料由于具有轻质、高比强度、高比刚度和良好的吸能特性而被广泛用于航空、航天、汽车、人员和结构防护领域^[1], 特别是用于爆炸防护、汽车耐撞性设计和飞机跑道阻拦等高速防护领域, 所以在高应变率条件下泡沫材料的力学性能备受关注。其中, 高应变率条件下材料参数测试的常用方法包括 Taylor 测试和 Hopkinson 杆测试, 而 Taylor 测试由于操作简单广受欢迎^[2]。

1948年, Taylor^[3]率先提出了 Taylor 测试的基本原理, 他利用圆柱形子弹撞击刚性靶板, 通过子弹的最终变形获得了材料高应变率状态下的力学特性。同年, Whiffin^[4]利用该方法对多种金属材料进行了试验分析, 通过子弹的最终变形得到了材料动态屈服强度, 此后多名学者对 Taylor 模型进行了改进。例如, Hawkyard^[5]利用能量守恒的办法代替动量定理对塑性波前进行了分析, 得到的子弹变形模式比 Taylor 模型更加接近于试验结果。Jones 等^[6-7]同时也分析了应变率和应变强

化效应对 Taylor 撞击结果的影响。另外还有不少学者对 Taylor 测试进行了有限元仿真分析^[8-10]。

除了金属材料, Taylor 撞击还广泛用于测试聚合物^[11-12]和泡沫材料^[13-14]的动态力学性能, Lu 等^[13-14]是较早对多孔材料的 Taylor 撞击进行理论和试验研究的, 他们通过多孔材料的压缩试验, 指出压缩过程中密度与压缩应变呈线性关系, 但这种线性关系只适用于压缩应变较小的情形。在本文中, 通过引入多孔材料的塑性泊松比, 对多孔泡沫材料压缩过程中的密度变化进行了详细推导, 当塑性泊松比为常数时, 密度表达式的一阶泰勒展开式可以退化到 Lu 等^[13]的分析结果, 当塑性泊松比为塑性应变和相对密度的函数时, 本文进一步分析了泡沫材料相对密度对 Taylor 撞击响应的影响。

1 问题描述及基本方程

假设一个圆柱体泡沫子弹以初速度 U 撞击

收稿日期: 2015-07-14; 录用日期: 2015-08-06; 网络出版时间: 2015-09-17 10:43

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20150917.1043.005.html

基金项目: 国家自然科学基金(11472035, 11472034)

* 通讯作者: Tel.: 010-82317507 E-mail: liuhuarui@buaa.edu.cn

引用格式: 刘虎, 刘华, 杨嘉陵. 泡沫子弹撞击刚性靶板的理论模型[J]. 北京航空航天大学学报, 2016, 42(7): 1461-1468. LIU H, LIU H, YANG J L. Theoretical model for a porous projectile striking on flat rigid anvil[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1461-1468 (in Chinese).

刚性靶板, 撞击过程可分为3个阶段, 图1给出了撞击前、撞击中间阶段及子弹的最终变形。如图1(a)所示, 假设子弹的初始长度为 L_0 , 截面面积为 A_0 , 子弹为理想刚塑性材料, 屈服应力 σ_Y , 同时假设泡沫子弹的直径远大于孔隙直径, 子弹可看作连续介质。当子弹撞击刚性靶板时, 在撞击面上形成的塑性波立刻以速度 v 向子弹的自由端传播, 如图1(b)所示, 塑性波前到达的区域立刻进入屈服, 截面面积由 A_0 变为 A , 子弹的密度由 ρ_0 变为 ρ , u 为未变形区域的速度, z 和 h 分别为未变形和变形区域的长度。如图1(c)所示, 撞击结束时子弹未变形和变形区域的长度分别为 Z 和 H , L_f 为子弹的总长度。

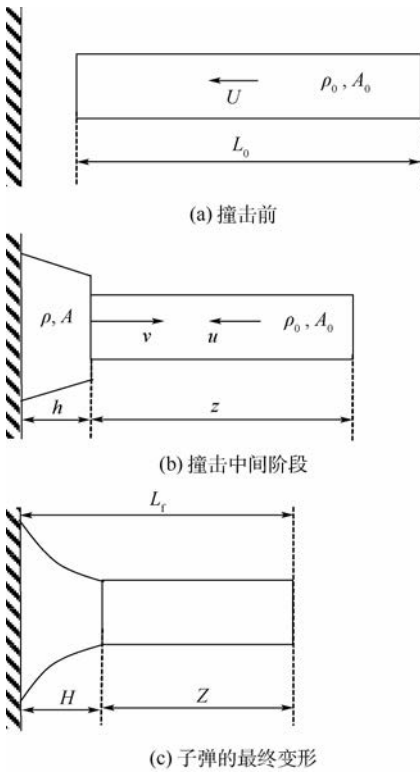


图1 圆柱体泡沫子弹撞击刚性靶板前、撞击中间阶段及子弹的最终变形

Fig. 1 A cylindrical porous projectile striking on a rigid anvil before impact, in intermediate stage of deformation, and in final stage of deformation

与 Taylor 模型一致, 可得

$$\frac{dh}{dt} = v \tag{1}$$

$$\frac{dz}{dt} = -(u + v) \tag{2}$$

$$\frac{du}{dt} = -\frac{\sigma_Y}{\rho_0 z} \tag{3}$$

式中: t 为时间。

由质量守恒可得连续方程:

$$\rho_0 A_0 (u + v) = \rho A v \tag{4}$$

在塑性波前截面上, 由冲量定理可得

$$\rho_0 A_0 (u + v) u = \sigma_Y (A - A_0) \tag{5}$$

假设沿撞击方向有限长度 dl_0 发生塑性变形后长度变为 dl , 截面面积由 A_0 变为 A , 则变形前后的体积分别为

$$dV_0 = A_0 dl_0 \tag{6}$$

$$dV = A dl = dl_0 (1 - e) A_0 (1 + \mu_p e)^2 = (1 - e) (1 + \mu_p e)^2 dV_0 \tag{7}$$

子弹压缩前后密度比可以写为

$$\frac{\rho}{\rho_0} = \frac{dV_0}{dV} = \frac{1}{(1 - e) (1 + \mu_p e)^2} \tag{8}$$

式中: V_0 为变形前的体积; V 为变形后的体积; μ_p 为塑性泊松比; e 为塑性应变, 其定义为

$$e = \frac{dl_0 - dl}{dl_0} = \frac{dV_0/A_0 - dV/A}{dV_0/A_0} = 1 - \frac{\rho_0 A_0}{\rho A} \tag{9}$$

引入无量纲量:

$$\left\{ \begin{aligned} \lambda &= \frac{u}{\sqrt{\sigma_Y/\rho_0}} \\ \lambda_0 &= \frac{U}{\sqrt{\sigma_Y/\rho_0}} \\ \eta &= \frac{v}{\sqrt{\sigma_Y/\rho_0}} \\ \tau &= \frac{t}{L_0/\sqrt{\sigma_Y/\rho_0}} \end{aligned} \right. \tag{10}$$

式中: λ 和 λ_0 分别为子弹未变形区域的无量纲瞬时速度及初始速度; η 为塑性波前的无量纲速度, 其初始值为 η_0 ; τ 为无量纲时间。

1.1 子弹未变形区域的速度

联立式(4)和式(5)并将式(9)代入可得

$$\frac{\rho_0 u^2}{\sigma_Y} = e \frac{A - A_0}{A_0} = e \left[\frac{\rho_0}{\rho(1 - e)} - 1 \right] \tag{11}$$

将式(8)和式(10)代入式(11)可得

$$\lambda = \sqrt{e \left[(1 + \mu_p e)^2 - 1 \right]} \tag{12}$$

1.2 子弹未变形区域的长度

式(2)与式(3)相除得

$$\frac{dz}{z} = \frac{\rho_0}{\sigma_Y} (u + v) du = \frac{\rho_0 u}{\sigma_Y e} du \tag{13}$$

对式(11)两边求导, 并将式(13)代入得

$$\frac{dz}{z} = \frac{1}{2e} d \left\{ e \left[\frac{\rho_0}{\rho(1 - e)} - 1 \right] \right\} \tag{14}$$

在子弹撞击刚性平面瞬时, 子弹具有最大应变 e_0 , 将式(8)代入式(14), 两端积分可得未变形区域长度 z 与应变 e 的关系:

$$\ln \left(\frac{z}{L_0} \right) = \frac{1}{2} \int_{e_0}^e \frac{1}{e} \left[3 \mu_p^2 e^2 + 4 \mu_p e + 2e^2 (1 + \mu_p e) \frac{d\mu_p}{de} \right] de \tag{15}$$

当撞击结束时,应变 $e = 0$,可以得到未变形

区域的最终长度 Z 为

$$\ln\left(\frac{Z}{L_0}\right) = \frac{1}{2} \int_{e_0}^0 \frac{1}{e} \left[3\mu_p^2 e^2 + 4\mu_p e + 2e^2(1 + \mu_p e) \frac{d\mu_p}{de} \right] de \quad (16)$$

1.3 子弹塑性变形区域的长度

式(1)与式(2)相除可得

$$dh = -\frac{v}{v+u} dz = -(1-e) dz \quad (17)$$

式(17)两边同除以 L_0 可得

$$d\left(\frac{h}{L_0}\right) = -(1-e) d\left(\frac{z}{L_0}\right) \quad (18)$$

令式(15)中 $\ln\left(\frac{z}{L_0}\right) = \Theta(e)$, 代入到式(18),

两端积分可得

$$\frac{h}{L_0} = \frac{1}{2} \int_e^{e_0} \frac{1-e}{e} \exp(\Theta(e)) \cdot \left[3\mu_p^2 e^2 + 4\mu_p e + 2e^2(1 + \mu_p e) \frac{d\mu_p}{de} \right] de \quad (19)$$

撞击结束时,应变 $e = 0$,此时可得塑性变形区的最终长度 H 为

$$\frac{H}{L_0} = \frac{1}{2} \int_0^{e_0} \frac{1-e}{e} \exp(\Theta(e)) \cdot \left[3\mu_p^2 e^2 + 4\mu_p e + 2e^2(1 + \mu_p e) \frac{d\mu_p}{de} \right] de \quad (20)$$

撞击结束时, L_f 等于未变形区域和塑性变形区域的长度之和:

$$L_f = Z + H \quad (21)$$

1.4 撞击持续时间

由式(3)和式(10)可得

$$d\tau = -\frac{z}{L_0} \sqrt{\frac{\rho_0}{\sigma_Y}} du \quad (22)$$

将式(12)和式(15)代入式(22),并结合初始条件 $\tau = 0$ 时 $e = e_0$,两边积分得

$$\tau = \int_{e_0}^e \exp(\Theta(e)) \cdot \frac{3\mu_p^2 e^2 + 4\mu_p e + 2e^2(1 + \mu_p e) \frac{d\mu_p}{de}}{2\sqrt{\mu_p^2 e^3 + 2\mu_p e^2}} de \quad (23)$$

在撞击结束时 $e = 0$,代入到式(23)可以得到撞击持续时间 τ_f :

$$\tau_f = \int_0^{e_0} \exp(\Theta(e)) \cdot \frac{3\mu_p^2 e^2 + 4\mu_p e + 2e^2(1 + \mu_p e) \frac{d\mu_p}{de}}{2\sqrt{\mu_p^2 e^3 + 2\mu_p e^2}} de \quad (24)$$

2 压缩密度比的退化

2.1 压缩密度比的一阶泰勒展开

由式(8),可得到压缩密度比的一阶泰勒展开式:

$$\frac{\rho}{\rho_0} = \frac{1}{(1-e)(1+\mu_p e)^2} \approx 1 + (1-2\mu_p)e \quad (25)$$

Lu 等^[14]通过多孔材料试验,认为子弹的压缩密度比与应变 e 满足 $\frac{\rho}{\rho_0} = 1 + \alpha e$,其中 α 为常数。

显然这里的 α 与式(25)中的 $1-2\mu_p$ 相对应,这也说明在 Lu 等^[13]的分析模型中,子弹的塑性泊松比 μ_p 为常数。

将式(25)代替式(8),类似推导可得到子弹未变形区域的无量纲瞬时速度 λ 、子弹未变形区域和变形区域的最终长度 Z 和 H ,以及无量纲撞击持续时间 τ_f 如下:

$$\lambda = \sqrt{e \frac{1}{(1-e)[1+(1-2\mu_p)e]} - 1} \quad (26)$$

$$\ln\left(\frac{Z}{L_0}\right) = \frac{1}{2} \int_{e_0}^0 \frac{1}{e} \left\{ \frac{1+(1-2\mu_p)e^2+2e^2(1-e) \frac{d\mu_p}{de}}{(1-e)^2[1+(1-2\mu_p)e]^2} - 1 \right\} de \quad (27)$$

$$\frac{H}{L_0} = \frac{1}{2} \int_0^{e_0} \frac{1-e}{e} \exp(\Theta_1(e)) \cdot \left\{ \frac{1+(1-2\mu_p)e^2+2e^2(1-e) \frac{d\mu_p}{de}}{(1-e)^2[1+(1-2\mu_p)e]^2} - 1 \right\} de \quad (28)$$

$$\tau_f = \frac{1}{2} \int_0^{e_0} \exp(\Theta_1(e)) \sqrt{\frac{(1-e)[1+(1-2\mu_p)e]}{2\mu_p e^2 + (1-2\mu_p)e^3}} \cdot \left\{ \frac{1+(1-2\mu_p)e^2+2e^2(1-e) \frac{d\mu_p}{de}}{(1-e)^2[1+(1-2\mu_p)e]^2} - 1 \right\} de \quad (29)$$

式中:

$$\Theta_1(e) = \frac{1}{2} \int_{e_0}^e \frac{1}{e} \left\{ \frac{1+(1-2\mu_p)e^2+2e^2(1-e) \frac{d\mu_p}{de}}{(1-e)^2[1+(1-2\mu_p)e]^2} - 1 \right\} de \quad (30)$$

当 μ_p 为常数时, $\frac{d\mu_p}{de} = 0$,令 $1-2\mu_p = \alpha$,式(26)~式(30)可以退化到 Lu 等^[13]的分析模型。

2.2 压缩密度比的二阶泰勒展开

由式(8),也可得到压缩密度比的二阶泰勒展开式:

$$\frac{\rho}{\rho_0} = \frac{1}{(1-e)(1+\mu_p e)^2} \approx 1 + (1-2\mu_p)e + (3\mu_p^2 - 2\mu_p + 1)e^2 \quad (31)$$

与2.1节类似,用式(31)代替式(8),得到的冲击响应结果为

$$\lambda = \sqrt{e \left\{ \frac{1}{(1-e)[1+(1-2\mu_p)e+(3\mu_p^2-2\mu_p+1)e^2]} - 1 \right\}} \quad (32)$$

$$\ln\left(\frac{Z}{L_0}\right) = \frac{1}{2} \int_{e_0}^e \left\{ \frac{1+2(3\mu_p^2-2\mu_p+1)e^3-3e^2\mu_p^2}{(1-e)^2[1+(1-2\mu_p)e+(3\mu_p^2-2\mu_p+1)e^2]^2} + \frac{(2e^2-6\mu_p e^3+2e^3)(1-e)\frac{d\mu_p}{de}}{(1-e)^2[1+(1-2\mu_p)e+(3\mu_p^2-2\mu_p+1)e^2]^2} - 1 \right\} \frac{1}{e} de \quad (33)$$

$$\frac{H}{L_0} = \frac{1}{2} \int_0^{e_0} \frac{1-e}{e} \exp(\Theta_2(e)) \cdot \left\{ \frac{1+2(3\mu_p^2-2\mu_p+1)e^3-3e^2\mu_p^2}{(1-e)^2[1+(1-2\mu_p)e+(3\mu_p^2-2\mu_p+1)e^2]^2} + \frac{(2e^2-6\mu_p e^3+2e^3)(1-e)\frac{d\mu_p}{de}}{(1-e)^2[1+(1-2\mu_p)e+(3\mu_p^2-2\mu_p+1)e^2]^2} - 1 \right\} de \quad (34)$$

$$\tau_t = \frac{1}{2} \int_0^{e_0} \exp[\Theta_2(e)] \cdot \left\{ \frac{\sqrt{(1-e)[1+(1-2\mu_p)e+(3\mu_p^2-2\mu_p+1)e^2]}}{e[2\mu_p e-3\mu_p^2 e^2+(3\mu_p^2-2\mu_p+1)e^3]} \cdot \frac{1+2(3\mu_p^2-2\mu_p+1)e^3-3e^2\mu_p^2}{(1-e)^2[1+(1-2\mu_p)e+(3\mu_p^2-2\mu_p+1)e^2]^2} - 1 + \frac{(2e^2-6\mu_p e^3+2e^3)(1-e)\frac{d\mu_p}{de}}{(1-e)^2[1+(1-2\mu_p)e+(3\mu_p^2-2\mu_p+1)e^2]^2} \right\} de \quad (35)$$

其中:

$$\Theta_2(e) = \frac{1}{2} \int_{e_0}^e \left\{ \frac{1+2(3\mu_p^2-2\mu_p+1)e^3-3e^2\mu_p^2}{(1-e)^2[1+(1-2\mu_p)e+(3\mu_p^2-2\mu_p+1)e^2]^2} + \frac{(2e^2-6\mu_p e^3+2e^3)(1-e)\frac{d\mu_p}{de}}{(1-e)^2[1+(1-2\mu_p)e+(3\mu_p^2-2\mu_p+1)e^2]^2} - 1 \right\} \frac{1}{e} de \quad (36)$$

3 结果与讨论

3.1 塑性泊松比为常数

图2给出了压缩密度比的一阶泰勒展开式与Lu等^[13]模型的计算结果对比,这里取初始应

变 $e_0 = 0.3$,同时取 $\mu_p = 0.25$,对应于Lu等^[13]模型中 $\alpha = 0.5$ 。可以看出两者计算得到的冲击响应结果一致,这也说明一阶泰勒展开模型可以退化到Lu等^[13]模型。

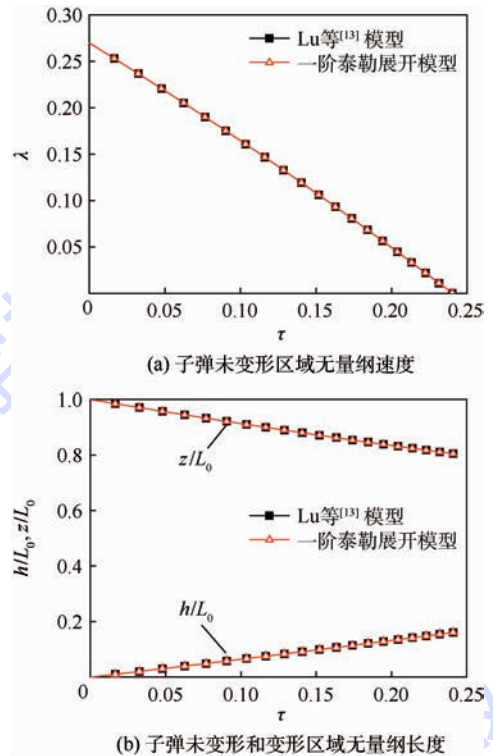


图2 $\mu_p = 0.25, e_0 = 0.3$ 时的冲击响应
Fig. 2 Impact responses when $\mu_p = 0.25$ and $e_0 = 0.3$

3.2 塑性泊松比为压缩应变的函数

多孔材料的塑性泊松比并不为常数,与冲击速度和压缩应变有关,对于某闭孔金属泡沫材料,文献[15]通过数值模拟给出了如下低速冲击情况下的塑性泊松比与压缩应变的关系:

$$\mu_p = \alpha_0 - \frac{\alpha_1 e_1}{e_1 + \alpha_2} \quad (37)$$

式中: e_1 为工程应变,满足 $e = \ln(1 + e_1)$; α_0, α_1 和 α_2 为材料常数并满足:

$$\begin{cases} \alpha_0 = 0.3 \\ \alpha_1 = 0.25 - 0.06\zeta \\ \alpha_2 = 0.47\zeta \end{cases} \quad (38)$$

式中: ζ 为泡沫子弹的相对密度,即金属泡沫的表现密度与相应金属材料的密度之比。

虽然式(37)是低速冲击条件下得到的数值模拟结果,但由文献[15]可知,冲击速度仅对塑性泊松比的峰值有较大影响,而该峰值对应的压缩应变很小,在Taylor撞击实验中,泡沫子弹的初始应变通常较大,此时冲击速度对塑性泊松比的影响较小,而在撞击后期,泡沫子弹的速度已较低,因此式(37)可以近似分析Taylor撞击问题。

由式(37),可以得到

$$\frac{d\mu_p}{de} = -\frac{\alpha_1 \alpha_2 \exp(e)}{(\exp(e) + \alpha_2 - 1)^2} \quad (39)$$

将式(37)分别代入到压缩密度比的原始表达式(8)及其一阶、二阶泰勒展开式(25)和式(31),可以分别得到对应的压缩密度比与压缩应变 e 的关系式。取泡沫材料的相对密度 $\zeta = 0.4$,图3给出了它们之间的关系,可以看出在压缩应变 e 较小时,压缩密度比与应变近似呈线性关系,这与 Lu 等^[14]的试验结果一致,但当压缩应变 e 较大时,这种线性关系不再成立。

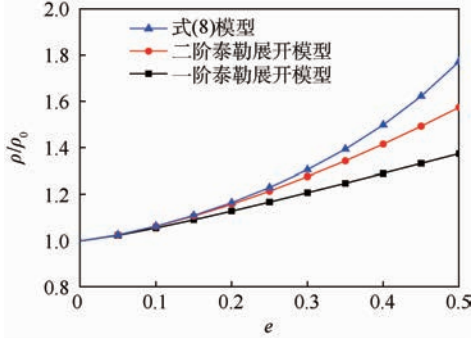


图 3 压缩密度比与压缩应变关系

Fig. 3 Relations between compressive density ratio and compressive strain

图4给出了由压缩密度比原始表达式(8)及其泰勒展开式(25)和式(31)得到的子弹的冲击响应历程,这里取泡沫材料的相对密度 $\zeta = 0.2$,初始无量纲速度 $\lambda_0 = 0.20435$,将 λ_0 和式(37)分别代入到无量纲速度表达式(12)、式(26)和式(32)可以计算得到对应的初始应变 e_0 ,从而得到各参数随时间的变化规律。由图4中可以看出它们具有相似的变化规律,而且随着展开阶次的增加,计算结果逐渐向压缩密度比原始表达式的结果靠近。

采用压缩密度比的原始表达式(8),并取初始无量纲速度 $\lambda_0 = 0.20435$,可以计算出相对密度不同时泡沫子弹的冲击响应历程,如图5所示。图中可以看出相对密度 ζ 不变时,子弹的塑性应变 e 、未变形区域长度 z/L_0 和未变形区域的速度 λ 在冲击历程中逐渐减小,而变形区域的长度 h/L_0 和塑性波前的速度 η 则随着时间的推移逐渐增大。同时可以看出,在同一时刻,随着相对密度 ζ 的增加,塑性应变 e 和子弹未变形区域长度 z/L_0 减小,变形区域长度 h/L_0 和塑性波前速度 η 则会增大,而未变形区域的速度 λ 则基本不变。

图6给出了泡沫材料相对密度不同时子弹的冲击响应随无量纲初始速度 λ_0 的变化规律,

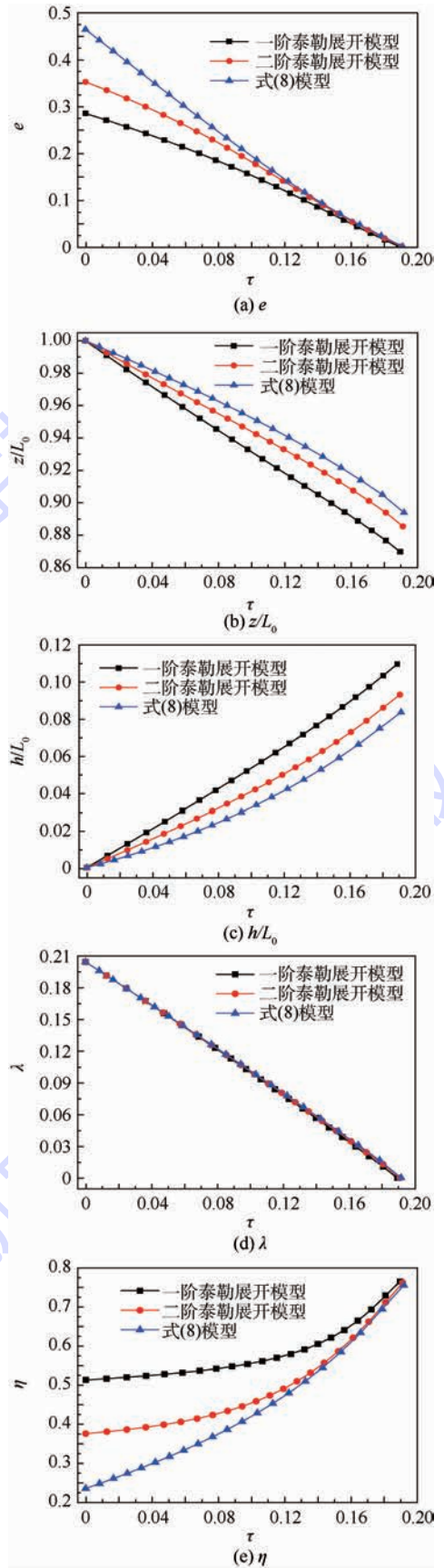


图 4 不同压缩密度比模型的冲击响应比较 ($\zeta = 0.2, \lambda_0 = 0.20435$)

Fig. 4 Comparison of impact responses predicted by different compressive density ratio models ($\zeta = 0.2, \lambda_0 = 0.20435$)

这里的压实应变取 $e_0 = 0.8$ 。可以看出 ζ 一定时, 最大塑性应变 e_0 、子弹最终的变形区域长度 H/L_0 和冲击持续时间 τ_f 随着 λ_0 的增加而增加, 子弹最终的未变形区域长度 Z/L_0 和子弹总剩余长度 L_t/L_0 及塑性波前无量纲初始速度 η_0 则随着 λ_0 的增加而相应减小。在 λ_0 一定时, 随着 ζ 的增大, 子弹最终的变形区域长度 H/L_0 和初始塑性

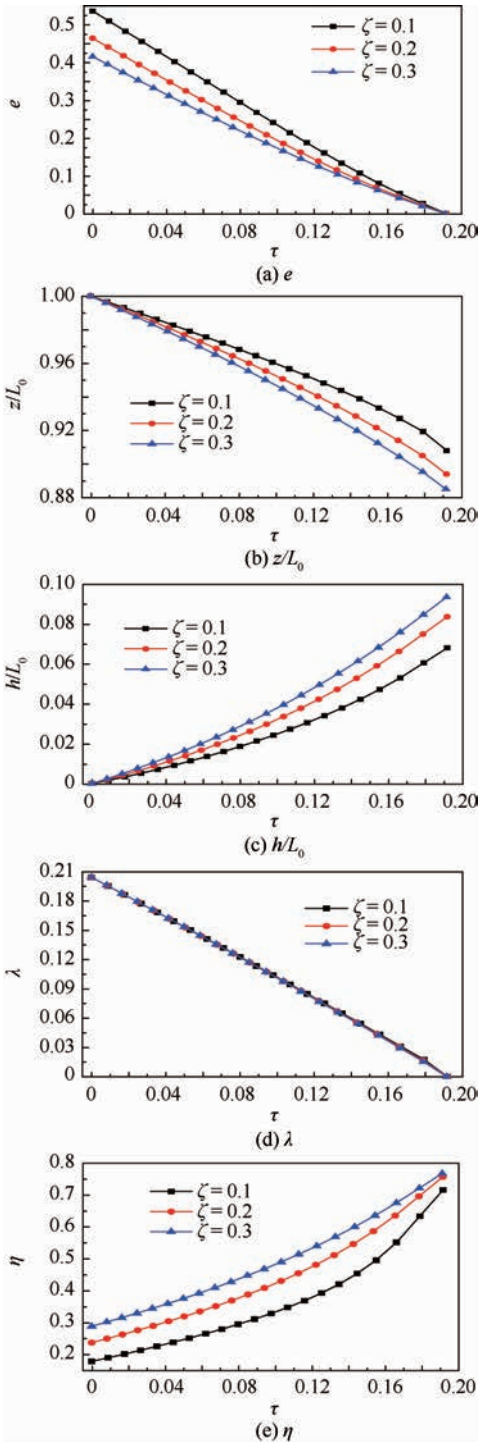


图5 ζ 不同时的冲击响应-时间曲线 ($\lambda_0 = 0.20435$)

Fig. 5 Impact responses versus time for different ζ ($\lambda_0 = 0.20435$)

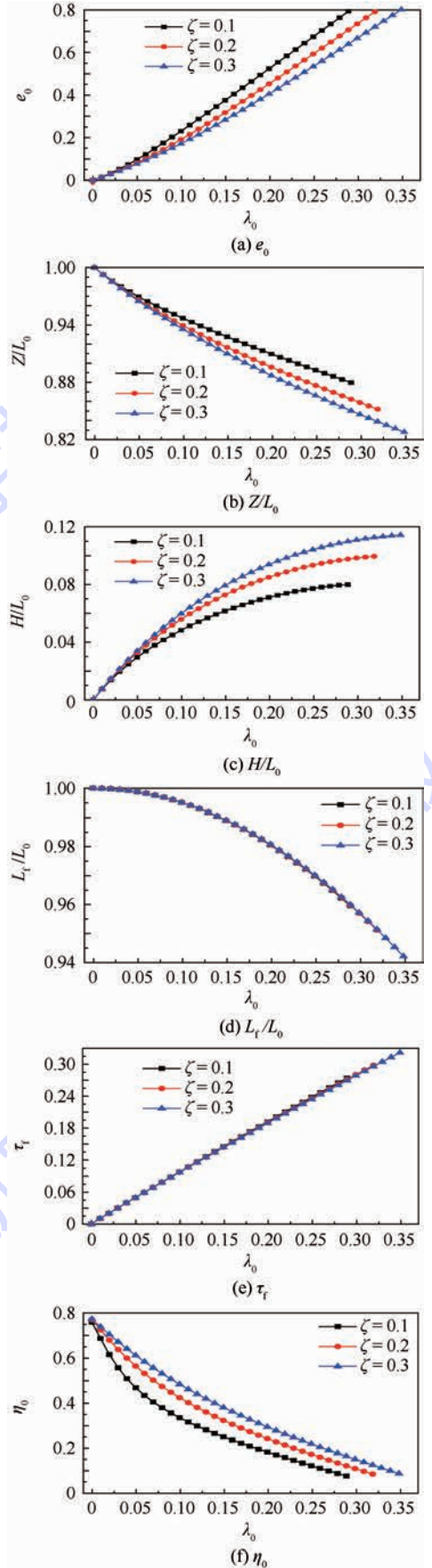


图6 ζ 不同时的最终冲击响应- λ_0 曲线

Fig. 6 Final impact responses versus λ_0 for different ζ

波前无量纲速度 η_0 相应的增大,而最大塑性应变 e_0 、子弹最终的未变形区域长度 Z/L_0 则减小,冲击持续时间 τ_i 和子弹总剩余长度 L_i/L_0 则基本保持不变。

4 结 论

本文通过引入塑性泊松比,从理论上推导了泡沫材料 Taylor 撞击中压缩密度比与压缩应变之间的关系,进而对泡沫子弹的 Taylor 撞击响应进行了分析,主要可以得到:

1) 当塑性泊松比为常数时,压缩密度比的一阶泰勒展开式可以退化到 Lu 的分析模型。同时,通过分析发现,当压缩应变较小时,撞击过程中的压缩密度比与压缩应变近似呈线性关系,这与 Lu 的试验分析结果一致,但压缩应变较大时,这种线性关系将不再成立。而后又计算了压缩密度比二阶泰勒展开的结果,发现它们的冲击响应历程具有相似的变化规律,同时随着展开阶次的提高,计算结果将逐渐向着原始表达式的结果靠近。

2) 当塑性泊松比为压缩应变的函数时,通过塑性泊松比与压缩应变和相对密度的关系式(37),分析了子弹相对密度对冲击响应的影响,发现相对密度会对泡沫子弹的冲击响应历程及最终变形产生影响,但对冲击响应持续时间影响较小。

3) 子弹的初始速度会对子弹最终变形和冲击响应持续时间产生重要影响。

本文的分析结果对利用 Taylor 模型分析泡沫材料的动力学行为具有重要的指导意义。

参考文献 (References)

- [1] LU G X, YU T X. Energy absorption of structures and materials [M]. Cambridge: Woodhead Publishing Ltd., 2003: 1-10.
- [2] 余同希, 邱信明. 冲击动力学 [M]. 北京: 清华大学出版社, 2011: 27-36.
- YU T X, QIU X M. Impact dynamics [M]. Beijing: Tsinghua University Press, 2011: 27-36 (in Chinese).
- [3] TAYLOR G. The use of flat-ended projectiles for determining dynamic yield stress I: Theoretical considerations [J]. Proceedings of the Royal Society of London A, 1948, 194 (1038): 289-299.
- [4] WHIFFIN A C. The use of flat-ended projectiles for determining dynamic yield stress II: Tests on various metallic materials [J]. Proceedings of the Royal Society of London A, 1948, 194 (1038): 300-322.
- [5] HAWKYARD J B. A theory for the mushrooming of flat-ended projectiles impinging on a flat rigid anvil, using energy consideration [J]. International Journal of Mechanical Sciences, 1969, 11 (3): 313-333.
- [6] JONES S E, GILLIS P P, FOSTER JR J C, et al. A one-dimensional, two-phase flow model for Taylor impact specimens [J]. Journal of Engineering Materials and Technology, 1991, 113 (2): 228-235.
- [7] JONES S E, DRINKARD J A, RULE W K, et al. An elementary theory for the Taylor impact test [J]. International Journal of Impact Engineering, 1998, 21 (1-2): 1-13.
- [8] BRÜIG M, DRIEMEIER L. Numerical simulation of Taylor impact tests [J]. International Journal of Plasticity, 2007, 23 (12): 1979-2003.
- [9] TENG X, WIERZBICKI T, HIERMAIER S, et al. Numerical prediction of fracture in the Taylor test [J]. International Journal of Solids and Structures, 2005, 42 (9): 2929-2948.
- [10] MA X, ZHANG D Z, GIGUERE P T, et al. Axisymmetric computation of Taylor cylinder impacts of ductile and brittle materials using original and dual domain material point methods [J]. International Journal of Impact Engineering, 2013, 54: 96-104.
- [11] HUTCHINGS I M. Estimation of yield stress in polymers at high strain-rates using Taylor's impact technique [J]. Journal of the Mechanics and Physics of Solids, 1978, 26 (5): 289-301.
- [12] 胡文军, 张方举, 陈裕泽, 等. 聚碳酸酯的 Taylor 撞击实验及光塑性分析 [J]. 高分子材料科学与工程, 2010, 26 (10): 115-117.
- HU W J, ZHANG F J, CHEN Y Z, et al. Taylor impact test of polycarbonate and photoplastic analysis [J]. Polymer Materials Science and Engineering, 2010, 26 (10): 115-117 (in Chinese).
- [13] LU G X, WANG B, ZHANG T G. Taylor impact test for ductile porous materials part I: Theory [J]. International Journal of Impact Engineering, 2001, 25 (10): 981-991.
- [14] WANG B, ZHANG J, LU G X. Taylor impact test for ductile porous materials part II: Experiments [J]. International Journal of Impact Engineering, 2003, 28 (5): 499-511.
- [15] 王长峰. 泡沫金属的动态压溃模型和率敏感性分析 [D]. 合肥: 中国科技大学, 2013.
- WANG C F. Dynamic crushing models and rate-sensitivity analysis of metallic foams [D]. Hefei: University of Science and Technology of China, 2013 (in Chinese).

作者简介:

刘虎 男, 博士研究生。主要研究方向: 冲击动力学。

Tel.: 010-82317507

E-mail: liuhu5260@126.com

刘华 女, 博士, 教授。主要研究方向: 纳米和复合材料力学、冲击动力学及结构耐撞性设计等。

Tel.: 010-82317507

E-mail: liuhuarui@buaa.edu.cn

Theoretical model for a porous projectile striking on flat rigid anvil

LIU Hu, LIU Hua*, YANG Jialing

(School of Aeronautic Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: Taylor impact is often applied to the determination of the dynamic yield stress of materials. For theoretical analysis of the Taylor impact of porous projectiles, the relationship between the density of a compressed porous projectile and the compressive plastic strain is very important. This paper proposes an exact density model for the compressible porous projectile by inducing the plastic Poisson's ratio, and further, an analytical model is established for the compressible porous projectile striking on a flat rigid anvil. As the plastic Poisson's ratio is a constant, the first order Taylor series expansion of the compression density ratio model can be reduced to the existing model. As the plastic Poisson's ratio is a function of the compressive plastic strain and the relative density, the relative density has a major influence on the impact response and the final deformation of the projectile, but the duration of impact-contact process is almost unaffected. The initial velocity of the projectile has considerable effects on both the final deformation of the projectile and the duration of impact-contact process. The present theoretical model is useful in analyzing the dynamic behavior of porous materials.

Key words: Taylor impact; porous material; plastic; plastic Poisson's ratio; rigid

Received: 2015-07-14; **Accepted:** 2015-08-06; **Published online:** 2015-09-17 10:43

URL: www.cnki.net/kcms/detail/11.2625.V.20150917.1043.005.html

Foundation items: National Natural Science Foundation of China (11472035, 11472034)

* **Corresponding author.** Tel.: 010-82317507 E-mail: liuhuarui@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0456

横向加强构件作用下的开口薄壁梁等效建模方法



邓昊, 程伟*

(北京航空航天大学 航空科学与工程学院, 北京 100083)

摘 要: 卫星结构中通常有大量横向加强构件作用下的开口薄壁杆件, 横向加强构件一般沿杆件轴向均匀分布, 通过对这类结构研究, 理论上证明了这种结构的振动微分方程与普通的开口薄壁梁振动微分方程有相同的形式, 因此可以用开口薄壁梁单元进行等效计算。本文建立了开口薄壁梁的3种数学模型: 有限元模型、传递矩阵模型和解析模型。采用优化理论中序列二次规划对等效截面参数进行辨识, 同时分析了不同目标函数对辨识结果的影响, 并且提出了一种对截面初始参数进行估计的方法。对于有限元模型, 提出采用 MATLAB 与 ABAQUS 交互式参数优化方法, 充分结合二者优点可以快速高效地对截面参数进行优化辨识, 具有较强的通用性。通过实验验证了这种等效建模方法的正确性与精确性。所提出的等效建模方法可以减少 90% 以上的单元数量。通过建立这种简化模型, 可以大幅度提高整星结构模型修正与结构重分析的效率。

关键词: 横向加强构件; 开口薄壁梁; 振动微分方程; 等效建模; 参数优化辨识

中图分类号: TV312

文献标识码: A **文章编号:** 1001-5965(2016)07-1469-10

在航空航天的结构中存在着大量的开口薄壁结构, 由于开口薄壁结构的抗扭性能较差, 而横向加强构件的存在可以大大改善开口薄壁结构受扭后的变形应力状态。因此, 为了提高其抗扭性能, 通常在结构中使用有横向加强构件的开口薄壁梁。由于横向加强构件的存在, 使得结构的建模通常采用壳单元或实体单元。实体单元虽然可以很好地模拟真实结构, 但通常所需单元的数量较大, 计算代价高昂。而使用壳单元虽能有效地减小单元数量, 但是薄壁构件通常存在大量的圆角, 壳单元对圆角进行建模通常很困难且步骤繁琐^[1]。因此, 为了更有效地对有横向加强构件的薄壁杆件建模, 本文提出运用开口薄壁梁单元对其进行等效建模, 单元的截面参数可以通过参数辨识的方法得到。这种方法可以大幅度地减小单元数量与计算规模并且有很高的精度。

Arpaci 等^[2-3]研究了考虑扭转效应的开口薄壁梁的耦合振动特性。Prokić^[4]研究了任意截面形状的开口薄壁梁的动力学特性。Ambrosini 等^[5-6]运用 Vlasov 理论计算了任意截面开口薄壁梁的振动频率。Prokić^[7]考虑了复杂薄壁梁振动的耦合效应。本文在此基础上建立了有横向加强构件的开口薄壁梁的弯扭耦合振动方程。通过实验测得的固有频率, 分别采用有限元法、传递矩阵法和解析法对模型进行参数优化辨识, 从而建立了等效梁单元模型, 通过与实测频率进行对比, 验证了方法的正确性与精确性。

1 开口薄壁杆件弯扭耦合振动方程

图 1 为开口薄壁梁横截面。考虑图 1 所示的弯扭耦合非对称薄壁梁单元截面, X 和 Y 轴通过截面剪切中心(剪心) S 且平行于过形心的坐标轴

收稿日期: 2015-07-07; 录用日期: 2015-10-10; 网络出版时间: 2015-11-19 10:01

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151119.1001.002.html

* 通讯作者: E-mail: cheng_wei@buaa.edu.cn

引用格式: 邓昊, 程伟. 横向加强构件作用下的开口薄壁梁等效建模方法[J]. 北京航空航天大学学报, 2016, 42(7): 1469-1478.
DENG H, CHENG W. Equivalent modeling method of open thin-walled beam under action of transverse stiffening member [J].
Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1469-1478 (in Chinese).

x 和 y , S 点与 G_M 点为横截面的剪心与质心, 剪心与质心之间的 2 个坐标轴之间的距离分别为 e_1 和 e_2 。一般细长的开口薄壁杆件弯扭耦合振动方程^[8]为

$$\begin{cases} EI_x \frac{\partial^4 v}{\partial z^4} - \frac{\partial^2}{\partial t^2} \left[\rho I_x \frac{\partial^2 v}{\partial z^2} - \rho A(v + \theta e_1) \right] = 0 \\ EI_y \frac{\partial^4 u}{\partial z^4} + \frac{\partial^2}{\partial t^2} \left[\rho I_y \frac{\partial^2 u}{\partial z^2} + \rho A(u - \theta e_2) \right] = 0 \\ EJ_w \frac{\partial^4 \theta}{\partial z^4} - GJ_\varphi \frac{\partial^2 \theta}{\partial z^2} + \frac{\partial^2}{\partial t^2} (\rho A e_1 v - \rho A e_2 u + \rho I_s \theta) = 0 \end{cases} \quad (1)$$

式中: v 为横截面剪心沿 y 方向的位移; u 为剪心沿 x 轴方向的位移; θ 为横截面扭转角; ρ 为材料密度; I_x 和 I_y 为截面关于 x 轴与 y 轴的惯性矩; J_w 和 J_φ 为截面翘曲常数和扭转常数; E 为材料的弹性模量; I_s 为截面关于剪心的极惯性矩; t 为时间; A 为隔板区域面积; G 为剪切模量。

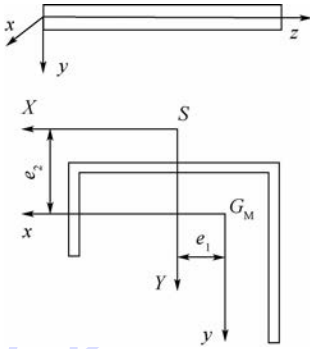


图1 开口薄壁梁横截面

Fig. 1 Cross-section of open thin-walled beam

2 横向加强构件作用下的开口薄壁杆件

横向隔板构件作用下的开口杆件如图2所示, 由于隔板对于杆件的翘曲有一定的抑制作用, 因此杆件的弯扭耦合运动方程与开口杆件将是有所差别的^[9], 容易证明, 对于纯弯曲变形状态, 这种杆件可以等效为普通的开口薄壁梁。而对于纯扭转变形, 本节将会证明有隔板加强杆件的运动方程与普通开口薄壁杆件运动方程在形式上也是一致的。当有横向隔板构件作用下的薄壁杆件发生整体扭转变形时, 由于隔板与杆件连接处横向弯矩与剪力的作用, 隔板除了产生扭曲以外还会发生弯曲。从杆件中取出含有隔板的一个微段 dz , 如图3所示。隔板的厚度为 d , 长宽分别为 b 和 h , 隔板平面示意图如图4所示。

若不考虑连接处的弯矩作用, 隔板相当于铰接在槽型杆件上, 由虚位移原理可以得到微段的

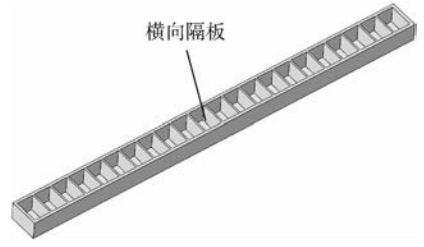


图2 横向隔板构件作用下的开口杆件
Fig. 2 Open bar under action of transverse partition member

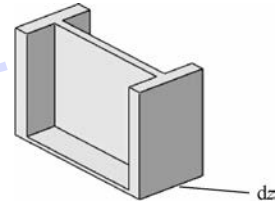


图3 含有横向隔板的一个微段
Fig. 3 A micro-segment containing transverse partition member

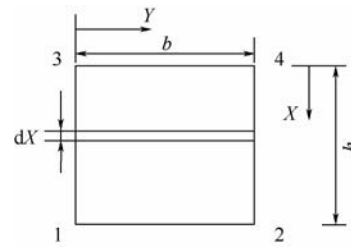


图4 横向隔板示意图
Fig. 4 Schematic diagram of transverse partition member

平衡微分方程^[9]:

$$B(z_0) \theta'(z_0) = 2(1 - \nu) D \iint_A \left(\frac{\partial^2 \omega}{\partial x \partial y} \right)^2 dx dy \quad (2)$$

式中: D 为隔板抗弯刚度; ν 为材料泊松比; $B(z_0)$ 为隔板对槽型截面作用的双力矩, z_0 为轴向坐标; ω 为隔板的面外位移, 与杆件轴向平行。根据薄板理论, 由于不考虑连接处横向弯矩作用, 隔板只承受扭矩, 板的面外位移为

$$\omega(x, y) = Kxy \quad (3)$$

式中: K 为任意常数, 可以通过边界条件确定。当杆件只承受扭矩作用时, 槽型杆件周边纵向位移为

$$u(z, s) = -\theta'(z) w(s) \quad (4)$$

式中: $w(s)$ 为以隔板中心为极点的扇形坐标, s 为周向曲线坐标。联立式(2) ~ 式(4)可以得到

$$B(z_0) = \frac{Ed^3 A}{12(1 + \nu)} \theta'(z_0) \quad (5)$$

由式(5)推导可知横向隔板对该槽型截面处的作用相当于一双力矩。

若考虑隔板与槽型截面连接处的弯矩时,将隔板可以看作一系列宽为 dx 、长为 b 的梁所组成,如图 4 所示,由于 1 和 2 边不受外力,因此可以忽略梁与梁之间的内力,考虑到槽型截面外形轮廓线上各点的翘曲不同,隔板将会发生面外弯曲。将宽为 dx 且两端固支的梁由中间位置截开。在切口处作用一对方向沿杆轴线的单位剪力。

可以由图乘法计算得到切口两端相对位移为

$$\Delta = \frac{b^3}{12EI} \quad (6)$$

式中: I 为梁的截面惯性矩:

$$I = \frac{(dx)d^3}{12} \quad (7)$$

开口截面薄壁杆件受扭矩作用时,梁切口两端的位移 δ 可以由单位力法求解得到

$$\delta = -\theta'(z_0)A_x \quad (8)$$

$$A_x = 2xb \quad (9)$$

式中: A_x 为隔板面积。由于切口两端相对位移为 0,将式(6)~式(9)代入力法方程,可以得到

$$F_x\Delta + \delta = 0 \quad (10)$$

式中: F_x 为切口两端作用力:

$$F_x = \frac{2Exd^3\theta'(z_0)}{b^2}dx \quad (11)$$

因此力 F_x 在槽型截面上引起的双力矩为

$$B_x = F_x A_x = \frac{4x^2 E d^3 \theta'(z_0)}{b} dx \quad (12)$$

式中: B_x 为长为 b 的梁在槽型截面 z_0 处的双力矩。现在对 B_x 从 0 到 h 进行积分,可以得到隔板整体对槽型截面的双力矩 $B(z_0)$ 为

$$B(z_0) = \int_0^h B_x dx = \frac{4Eh^3 d^3}{3b} \theta'(z_0) \quad (13)$$

考虑到隔板同时发生弯曲与扭转变形,因此总的双力矩为二者的叠加,即

$$B(z_0) = \left[\frac{4Eh^3 d^3}{3b} + \frac{Ed^3 A}{12(1+\nu)} \right] \theta'(z_0) \quad (14)$$

由于杆件为周期薄壁结构,且隔板沿杆纵向布置较密,可以采用连续化处理,将多层隔板等效成为连续的弹性介质,弹性介质对槽型截面的双力矩集度可由式(15)等效计算:

$$b(z) = \frac{B(z)}{L_c} \quad (15)$$

式中: $B(z)$ 为双力矩; L_c 为隔板之间的间距。槽型截面剪力流为

$$Q(z) = \frac{b(z)}{A} \quad (16)$$

根据有加强构件的开口薄壁梁的连续化解法,可以得到一般意义下的有加强构件薄壁杆件

受静态扭转载荷作用的平衡微分方程^[9]为

$$EJ_w \frac{d^4\theta}{dz^4} - GJ \frac{d^2\theta}{dz^2} = m_t + Q'(z)A \quad (17)$$

$$Q'(z) = \frac{dQ(z)}{dz} \quad (18)$$

式中: m_t 为沿杆长分布的外扭矩; J 为扭转惯性矩。

将式(14)~式(16)代入式(17)和式(18)中,可以得到微段受扭矩作用的平衡微分方程:

$$EJ_w \theta'''' - \left(\frac{\bar{K}}{L_c} + GJ \right) \theta'' = m_t \quad (19)$$

$$\bar{K} = \frac{4Eh^3 d^3}{3b} + \frac{Ed^3 A}{12(1+\nu)} \quad (20)$$

式中: \bar{K} 为修正刚度。

一般非对称开口薄壁杆件受静态扭转载荷作用的平衡微分方程为

$$EJ_w \theta'''' - GJ \theta'' = m_t \quad (21)$$

对比式(19)与式(21)可知,有隔板作用下的开口薄壁杆件受扭静态平衡方程与一般开口薄壁杆件方程形式相同,只是在扭转刚度 GJ 上增加了一个修正项 $\frac{\bar{K}}{L_c}$;而对于横向构件作用下的薄壁构件弯扭耦合动态方程可以通过达朗贝尔原理将静态力用惯性力替换而得到。类似地可以得到相同的结论。根据本节分析可知,有隔板作用的开口薄壁梁与一般开口薄壁梁的运动微分方程形式完全相同,因此这种结构可以用开口薄壁梁单元进行等效计算。

3 复杂薄壁开口杆件等效截面参数优化识别

3.1 有限元模型

正确的辨识等效截面参数首先需要建立正确的数学模型。本节通过有限元法建立考虑剪切变形的开口薄壁梁单元的数学模型。王晓峰^[10]、Mohri^[11]和 Kim^[12]等分别建立了考虑剪切变形的空间薄壁梁单元。这种单元基于 Timoshenko 梁理论和 Vlasov 薄壁杆件约束扭转理论,通过设置单元内部结点,对弯曲转角和翘曲角采用独立插值的方法,薄壁杆件横截面正应力对应的余能为

$$\Pi_{c1} = \int_V \frac{\sigma_x^2}{2E} dV \quad (22)$$

式中: V 为体积; σ_x 为正应力。

剪应力对应的余能为

$$\Pi_{c2} = \int_V \frac{\tau^2}{2G} dV \quad (23)$$

式中: τ 为剪应力。

圣维南剪力所对应的余能为

$$\Pi_{c3} = \int_0^l \frac{1}{2G} \cdot \frac{M_T^2}{J} dx \quad (24)$$

式中: M_T 为转矩; l 为梁的长度。

因此单元的总余能为三者的叠加,即

$$\Pi_c = \Pi_{c1} + \Pi_{c2} + \Pi_{c3} \quad (25)$$

根据 Hellinger-Reissner 广义变分原理, 推得可同时考虑剪切变形及弯扭耦合等因素的空间薄壁梁单元刚度矩阵。单元每个节点有 7 个自由度: $\{u, v, w, \theta_x, \theta_y, \theta_z, \theta\}$, u, v 和 w 为沿 3 个坐标轴方向的横向位移, θ_x, θ_y 和 θ_z 为绕 3 个坐标轴的转角。通过引入单元内部节点, 由插值函数建立单元位移场, 通过虚功原理建立单元平衡微分方程, 同时将单元内部自由度凝聚后可以得到单元刚度矩阵 K , 同理用类似的方法可以得到单元的质量矩阵, 以上所描述的有限元模型在一般的有限元通用程序中均有相应的单元, 例如 ABAQUS 中开口薄壁梁单元。

3.2 传递矩阵模型

Ambrosini^[13]建立了考虑剪切效应与转动惯量的开口薄壁梁单元的传递矩阵模型, 即

$$\frac{\partial \mathbf{v}}{\partial z} = \mathbf{A} \mathbf{v} \quad (26)$$

$$\mathbf{v}(z, w) = [u, \varphi, Q_y, M_x, v, \varphi_x, Q_x, M_y, \theta, \theta', M_T, B]^T$$

式中: $\mathbf{v}(z, \omega_f)$ 为状态矢量, ω_f 为圆频率; φ_x 和 φ_y 为绕 x 轴和 y 轴的截面转角; Q_x 和 Q_y 为截面剪力; M_x 和 M_y 为弯曲力矩; θ' 为 θ 沿薄壁梁轴向坐标的一阶微分; B 为双力矩; 矩阵 \mathbf{A} 的具体形式可以参见文献 [12]。式 (26) 适用于非对称开口薄壁杆件, 并且由于考虑了剪切效应, 因此也适用于深梁。因此状态矢量可以表示为

$$\mathbf{v} = e^{Az} \cdot \mathbf{v}_0 \quad (27)$$

式中: \mathbf{v}_0 为初始状态矢量。将 e^{Az} 展开成级数的形式, 即

$$e^{Az} = \frac{1}{k!} \sum_{k=0}^{\infty} A^k z^k \quad (28)$$

式中: \mathbf{A} 为 12 阶 \times 12 阶函数传递矩阵。对于固支边界条件、简支边界条件和自由边界条件依次有

$$\begin{cases} u = v = 0 \\ \varphi_x = \varphi_y = \theta = 0 \\ \theta' = 0 \end{cases} \quad (29)$$

$$\begin{cases} Q_y = Q_x = 0 \\ M_x = M_y = M_T = 0 \\ B = 0 \end{cases} \quad (30)$$

$$\begin{cases} u = v = 0 \\ M_x = M_y = 0 \\ \theta = 0 \\ B = 0 \end{cases} \quad (31)$$

根据开口薄壁梁两端的边界条件可以求出固有频率。

3.3 解析模型

开口薄壁杆件振动微分方程式 (1) 并未考虑梁的剪切效应, 因此仅对于细长梁有较高的精度。令

$$\begin{cases} u = U(z) e^{i\omega_f t} \\ v = V(z) e^{i\omega_f t} \\ \theta = \Phi(z) e^{i\omega_f t} \\ \lambda_x = \frac{\omega_f^2 \rho A}{EI_x} \\ \lambda_y = \frac{\omega_f^2 \rho A}{EI_y} \\ \lambda_{\omega_f} = \frac{\omega_f^2 \rho A}{EI_w} \\ \lambda_b = \frac{GJ}{EI_w} \\ \lambda_p = \frac{\omega_f^2 \rho I_p}{EI_w} \\ \mu = \frac{\rho \omega_f^2}{E} \end{cases} \quad (32)$$

式中: I_w 为翘曲系数; I_p 为极惯性矩; $U(z)$ 为横向位移幅值; $V(z)$ 为纵向位移幅值; $\Phi(z)$ 为转角位移幅值, 假设:

$$\begin{cases} U(z) = C_U e^{rz} \\ V(z) = C_V e^{rz} \\ \Phi(z) = C_\Phi e^{rz} \end{cases} \quad (33)$$

式中: C_U 为横向位移常数; C_V 为纵向位移常数; C_Φ 为转角位移常数; r 为常数。

将式 (32) 和式 (33) 代入振动微分方程式 (1) 并化简可以得到其矩阵形式为

$$\begin{bmatrix} r^4 - \mu r^2 - \lambda_y & 0 & \lambda_y e_2 \\ 0 & r^4 + \mu r^2 - \lambda_x & -\lambda_x e_1 \\ \lambda_w e_2 & -\lambda_w e_1 & r^4 - \lambda_b r^2 - \lambda_0 \end{bmatrix} \cdot \begin{bmatrix} C_U \\ C_V \\ C_\Phi \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (34)$$

若方程式 (34) 存在非零解, 则其系数行列式应为 0, 于是有

$$r^{12} + a_1 r^{10} + a_2 r^8 + a_3 r^6 + a_4 r^4 + a_5 r^2 + a_6 = 0 \quad (35)$$

式中:

$$\begin{cases} a_1 = -\lambda_b \\ a_2 = -(\lambda_p + \mu^2 + \lambda_y + \lambda_x) \\ a_3 = \lambda_b(\mu^2 + \lambda_y + \lambda_x) + \mu(-\lambda_y + \lambda_x) \\ a_4 = \lambda_p\mu^2 + \mu\lambda_b(\lambda_y - \lambda_x) + \lambda_x(\lambda_p - e_1^2\lambda_w^2) + \lambda_y(\lambda_p + \lambda_x - e_2^2\lambda_w^2) \\ a_5 = -\lambda_b\lambda_x\lambda_y + \mu[\lambda_x(-\lambda_p + e_1^2\lambda_w) + \lambda_y(\lambda_p - e_2^2\lambda_w)] \\ a_6 = \lambda_y[e_2^2\lambda_x\lambda_y + \lambda_x(-\lambda_p + e_1^2\lambda_w)] \end{cases}$$

可以通过数值计算求解方程式(35)的根,因此可以得到 $U(z)$ 、 $V(z)$ 和 $\Phi(z)$ 的一般表达式为

$$\begin{cases} U(z) = \sum_{i=1}^{12} C_{U_i} e^{r_i z} \\ V(z) = \sum_{i=1}^{12} C_{V_i} e^{r_i z} \\ \Phi(z) = \sum_{i=1}^{12} C_{\phi_i} e^{r_i z} \end{cases} \quad (36)$$

式中: r_i 为常数,未知量 C_{U_i} 、 C_{V_i} 和 C_{ϕ_i} 可以由边界条件得到。边界条件与式(29)~式(31)相同。

3.4 截面参数优化识别流程

3.4.1 目标函数建立

进行参数优化识别首先要建立正确的目标函数,目标函数的好坏将会直接影响到辨识参数的精度,本文通过采用3种不同的目标函数对参数进行辨识,并同时比较3种辨识结果的精度。目标函数为

$$f_1(p) = \sum_{i=1}^N \left(1 - \frac{f_{iA}}{f_{iE}} \right)^2 \quad (37)$$

$$f_2(p) = \sum_{i=1}^N \left| 1 - \frac{f_{iA}}{f_{iE}} \right| \quad (38)$$

$$f_3(p) = \min_{1 \leq i \leq n} \max \left(\left| 1 - \frac{f_{iA}}{f_{iE}} \right| \right) \quad (39)$$

式中: p 为需要辨识的参数; f_{iA} 为计算固有频率; f_{iE} 为实验频率或目标频率。后2种目标函数是非光滑的函数,可以通过变换将其转化为光滑优化问题。

对于目标函数式(38),可以将其转化为

$$\begin{aligned} \min \sum_{i=1}^n t_i \\ \text{s. t. } -t_i \leq 1 - \frac{f_{iA}}{f_{iE}} \leq t_i \quad i = 1, 2, \dots, n \end{aligned} \quad (40)$$

对于目标函数式(39),可以将其转化为

$$\begin{aligned} \min t \\ \text{s. t. } -t \leq 1 - \frac{f_{iA}}{f_{iE}} \leq t \quad i = 1, 2, \dots, n \end{aligned} \quad (41)$$

通过以上变量替换将截面参数辨识问题转化为带约束的非线性优化问题。序列二次规划(SQP)^[14-16]算法是求解中小规模约束最优化问题的一类有效算法。SQP的基本思想是用一系列二次规划问题来逼近带约束的非线性优化问题的解。本文采用全局SQP方法对截面参数进行优化,截面参数优化识别流程如图5和图6所示。

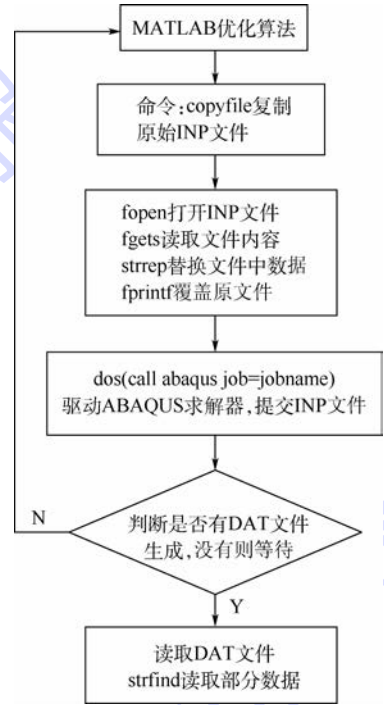


图5 MATLAB与ABAQUS交互优化流程
Fig.5 MATLAB and ABAQUS interactive optimization process

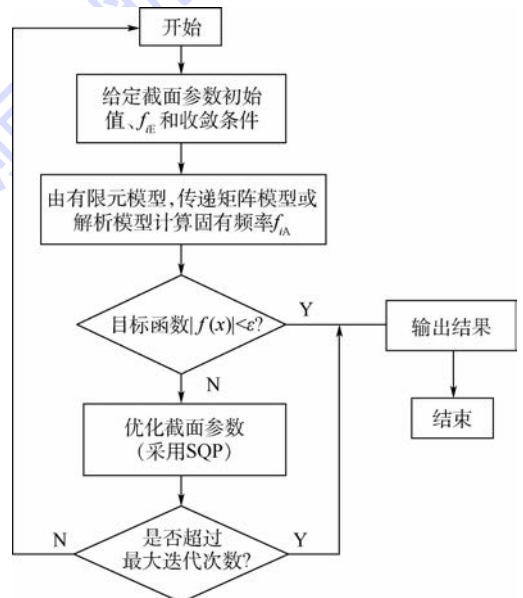


图6 截面参数识别流程

Fig.6 Cross-section parameter identification procedure

3.4.2 MATLAB 与 ABAQUS 交互式优化

MATLAB 是一种影响大、流行广的科学计算语言,其中拥有大量优秀的数值算法,而 ABAQUS 是一款非常强大的有限元分析软件,可以模拟非常复杂的模型。在运用有限元模型进行参数辨识的过程中,本文将结合二者的优点,同时使用 MATLAB 和 ABAQUS 进行交互式的优化参数辨识。INP 文件是 ABAQUS 的输入文件,它包含了对整个模型的完整描述,而 INP 文件本质上是一种文本文件,因此 MATLAB 可以对其进行读写从而控制 ABAQUS 的分析过程。ABAQUS 通常可以将其计算结果输出至 DAT 文件, MATLAB 可以直接读取 DAT 文件中的数据,由此可知只需要通过文件之间的交换即可实现 MATLAB 与 ABAQUS 联合优化。

在优化过程中,每一步迭代, MATLAB 中可以通过修改 INP 文件中的部分数据然后运用 dos 命令驱动 ABAQUS 程序进行计算,同时 MATLAB 可以判断是否有 DAT 结果文件生成从而读取其中的结果数据,由结果数据通过优化算法进行一步迭代得到一组新的数据作为 ABAQUS 输入数据,从而完成一次循环。MATLAB 与 ABAQUS 交互优化流程如图 5 所示。

3.4.3 截面参数初始值估算

截面参数初始值对参数优化将会产生较大的影响。合理的设置初始值才能得到正确的结果。初始值的估计可以通过对结构施加静态载荷,运用有限元程序计算得到的静态位移估算截面的等效参数。截面等效惯性矩 I_x 和 I_y 的估计可以通过对一端固支的实体在另一端分别施加弯矩 M_x 和 M_y ,通过实体单元计算截面端部的位移 w_x 和 w_y ,由材料力学可知

$$I = \frac{ML^2}{2EW} \quad (42)$$

式中: M 为端部所受的力矩; W 为端部位移; L ($0 \leq L \leq 1$) 为隔板沿轴向厚度之和与结构的总长度之比。

截面的扭转常数 J_φ 与翘曲常数 J_w 对结构的固有频率的影响很大,因此合理的设置初值可以提高截面参数的辨识精度。

工况 1 考虑将结构的一端固支,在另一端施加扭矩 M_T 。采用里兹法求解截面扭转角,开口截面薄壁杆件的约束扭转势能可以表示为

$$\Pi = \frac{1}{2} \int_0^l (EJ_w \theta'^2 + GJ_\varphi \theta^2) dz - \int_0^l M_T \theta dz \quad (43)$$

由于薄壁杆件的一端完全固支,因此在固定

端处的扭转角 θ 和翘曲 $\frac{d\theta}{dz}$ 应为 0。根据位移边界条件,可以假设扭转角 θ 为多项式的形式:

$$\theta = a_1 z^2 + a_2 z^3 + \dots + a_n z^{n+1} \quad (44)$$

易知式(44)满足位移边界条件。将式(44)代入式(43),得

$$\begin{aligned} \Pi = & \frac{1}{2} \int_0^l \{ EJ_w [\sum_{i=1}^n (i^2 + i) a_i z^{i-1}]^2 + \\ & GJ_\varphi [\sum_{i=1}^n (i + 1) a_i z^i]^2 \} dz - \int_0^l M_T (\sum_{i=1}^n a_i z^i) dz \end{aligned} \quad (45)$$

由势能驻值原理可以得到

$$\frac{\partial \Pi}{\partial a_1} = 0, \frac{\partial \Pi}{\partial a_2} = 0, \dots, \frac{\partial \Pi}{\partial a_n} = 0 \quad (46)$$

由此可以解出 n 个参数。当令 $n \rightarrow \infty$, 可以解得精确解为

$$\theta(z) = -\frac{ml}{KGJ_\varphi} \left[\text{thk} \left(1 - \text{ch} \frac{k}{l} z \right) - \left(\frac{k}{l} z - \text{sh} \frac{k}{l} \right) \right] \quad (47)$$

式中: $k = \sqrt{\frac{GJ_\varphi l^2}{EJ_w}}$ 。令 $z = l$ 可以得到扭矩作用点处的扭转角。

工况 2 考虑将结构的两端均固支,在杆件的 $1/2$ 截面处施加扭矩 M_T ,由于在固定端处要满足位移边界条件,可以假设扭转角 θ 为

$$\theta(z) = \sum_{n=1}^{\infty} a_n \left(1 - \cos \frac{2n\pi z}{l} \right) \quad (48)$$

易知式(48)满足两端固支的边界条件,因此开口薄壁杆件的应变能为

$$\begin{aligned} U_E = & \frac{EJ_w}{2} \int_0^l \theta'^2 dz + \frac{GJ_\varphi}{2} \int_0^l \theta^2 dz = \\ & \frac{EJ_w}{2} \int_0^l \left(\sum_{n=1}^{\infty} a_n \frac{4n^2 \pi^2}{l^2} \cos \frac{2n\pi z}{l} \right)^2 dz + \\ & \frac{GJ_\varphi}{2} \int_0^l \left(\sum_{n=1}^{\infty} a_n \frac{2n\pi}{l} \sin \frac{2n\pi z}{l} \right)^2 dz \end{aligned} \quad (49)$$

利用积分公式,当 $n \neq m$ 时有

$$\begin{cases} \int_0^l \sin \left(\frac{m\pi z}{l} \right) \sin \left(\frac{n\pi z}{l} \right) dz = 0 \\ \int_0^l \cos \left(\frac{m\pi z}{l} \right) \cos \left(\frac{n\pi z}{l} \right) dz = 0 \end{cases} \quad (50)$$

当 $n = m$ 时有

$$\begin{cases} \int_0^l \sin \left(\frac{m\pi z}{l} \right) \sin \left(\frac{n\pi z}{l} \right) dz = \frac{l}{2} \\ \int_0^l \cos \left(\frac{m\pi z}{l} \right) \cos \left(\frac{n\pi z}{l} \right) dz = \frac{l}{2} \end{cases} \quad (51)$$

因此式(49)可以化简为

$$U_E = \frac{\pi^2 GJ_\varphi}{l} \sum_{n=1}^{\infty} n^2 a_n^2 + \frac{4\pi^4 EJ_w}{l^3} \sum_{n=1}^{\infty} n^4 a_n^2 \quad (52)$$

载荷势能为

$$m = -M_T \sum_{n=1}^{\infty} a_n [1 - (-1)^n] \quad (53)$$

因此薄壁杆件的总势能为

$$\Pi = U + m \quad (54)$$

根据势能驻值条件可以得到

$$\frac{\partial \Pi}{\partial a_n} = \left(\frac{2\pi^2 GJ_{\varphi} n^2}{l} + \frac{8\pi^4 EJ_w n^4}{l^3} \right) a_n + [\cos(n\pi) - 1] = 0 \quad (55)$$

可以得到

$$a_n = \frac{l[1 - \cos(n\pi)]}{2\pi^2 n^2 GJ_{\varphi} \left(\frac{4\pi^2 n^2}{h^2} + 1 \right)} \quad (56)$$

由此可以得到扭转角 $\theta(z)$ 的解析式,令 $z = \frac{l}{2}$

可以得到扭矩作用点处的扭转角。实际使用中,通常取解析式的前3项来近似扭转角 $\theta(z)$ 。本节推导了2种不同工况下扭转角 $\theta(z)$ 的解析式。对于扭转常数 J_{φ} 与翘曲常数 J_w 初值的估算,可以在有限元程序中运用实体单元计算2种工况下结构的静态位移,然后联立式(47)和式(48)可以得到关于 J_{φ} 与 J_w 的非线性方程组。非线性方程组可以采用牛顿法进行求解进而得到 J_{φ} 与 J_w 的初始估计值。

4 仿真算例

运用有横向隔板构件加强的薄壁结构进行仿真验证,槽型横截面宽0.1 m,高0.1 m,壁厚0.008 m,隔板厚度为0.005 m,隔板的间距为0.071 m,薄壁构件的总长为1.5 m,材料弹性模量为71 GPa,材料的密度为 $2.78 \times 10^3 \text{ kg/m}^3$ 。在ABAQUS中运用实体单元计算结构两端自由时的前10阶固有频率,建立目标函数,设置辨识参数的初始值然后进行参数优化,x轴与y轴为过截

面形心的主惯性轴,且y轴为截面对称轴,如图1所示。易知剪心在此对称轴上,由此可知 $e_1 = 0$,因此需要辨识的参数有7个,辨识结果如表1所示。表2为简化模型计算所得的各阶频率与真实模型频率的误差绝对值。

表1中, K_x 和 K_y 为x方向和y方向的横向剪切刚度系数。由表1可以看出,目标函数为二范数与无穷范数辨识所得的结果相接近,并且由表2反映上述2种模型计算的结果相对于其他模型更加接近真实值。当目标函数为二范数时,等效模型与真实模型的模态应变能及其误差如表3所示,由于模态应变能不仅包含各阶模态的刚度信息,同时还包含振型的相关信息,因此表3的结果可以说明等效梁单元计算精度可以逼近真实解且误差在工程要求的范围之内。

有横向隔板的开口结构的扭转刚度与翘曲刚度介于普通的开口结构与闭口结构之间,随着横向隔板数量的增加,其结构的力学性能在理论上应接近于闭口结构,反之,应接近于无横向加强的薄壁开口结构,为了验证此结论,定义式(42)中的L,对于开口截面与闭口截面均可以得到其等效扭转常数的精确解。对于开口截面,截面的扭转常数精确解为^[17]

$$J_{\varphi} = a\delta_i^3 - 0.07\delta_i^4 \quad (57)$$

对于闭口截面,截面扭转常数的精确解为

$$J_{\varphi} = 2.24923a^4 \quad (58)$$

式中:a为槽型截面尺寸; δ_i 为槽型截面的壁厚。图7反映了随着隔板数量与厚度的增加,等效截面扭转常数可以趋近于闭口截面的精确解 $2.2492 \times 10^{-4} \text{ m}^4$,而随着隔板数量与厚度的减少,等效截面扭转常数可以趋近于槽形开口截面扭转常数精确解 $5.0913 \times 10^{-8} \text{ m}^4$ 。由此可以说明等效梁模型的正确性与合理性。

表1 截面参数辨识结果

Table 1 Cross-section parameter identification results

截面参数	梁单元模型			传递矩阵模型	解析模型
	二范数	无穷范数	一范数		
$I_x/(10^{-6} \text{ m}^4)$	2.26	2.34	2.24	2.16	2.30
$I_y/(10^{-6} \text{ m}^4)$	3.88	2.50	3.76	3.76	3.71
$J_{\varphi}/(10^{-7} \text{ m}^4)$	1.47	1.44	1.42	1.39	1.44
$J_w/(10^{-9} \text{ m}^6)$	2.08	2.19	2.06	2.17	2.13
e_1/m	0	0	0	0	0
e_2/mm	66.4	68.4	65.8	68.2	67.2
K_x	0.837	0.905	0.824	0.788	
K_y	0.411	0.443	0.424	0.352	

表2 简化模型与真实模型频率的误差绝对值

Table 2 Frequency absolute error between simple model and real model %

模态阶次	梁单元模型			传递矩阵模型	解析模型
	二范数	无穷范数	一范数		
1	0.37	2.12	0.29	0.01	0.27
2	1.08	0.98	1.06	0.08	0.80
3	0.09	1.91	0.15	1.40	0.88
4	0.57	2.11	1.05	1.71	1.36
5	1.68	2.02	2.18	1.22	1.77
6	0.07	1.99	0.06	0.18	0.57
7	0.45	0.45	1.01	0.20	0.53
8	0.67	0.92	0.72	0.22	0.63
9	0.02	2.13	0.09	2.35	1.15
10	2.07	2.18	2.06	2.19	2.12
11	2.05	2.08	1.11	2.18	1.85

表3 等效模型与真实模型的模态应变能及其误差

Table 3 Model strain energy of equivalent model and real model and their error

模态阶次	模态应变能/(N·m)		模态应变能误差/%
	真实模型	等效模型	
1	530070	538680	1.62
2	595130	604227	1.51
3	923810	922350	-0.16
4	2421200	2362411	-2.41
5	5509182	5574967	1.18
6	6.456×10^6	6.460×10^6	0.06
7	7.208×10^6	7040179	-2.33
8	1.419×10^7	1.396×10^7	-1.62
9	2.215×10^7	2.226×10^7	0.49
10	2.733×10^7	2.790×10^7	2.01
11	3.067×10^7	3.129×10^7	1.96

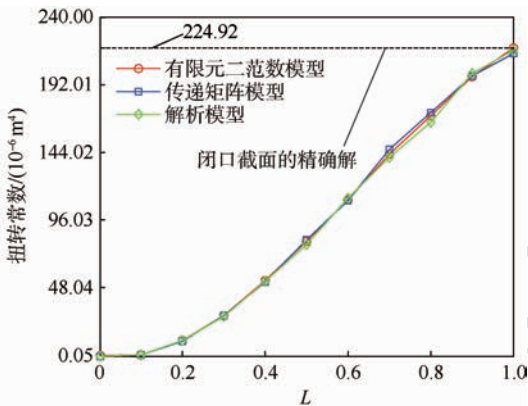


图7 等效截面扭转常数与L的关系曲线

Fig.7 Relationship curves between equivalent cross-section torsion constant and L

5 实验验证

某测试结构长 1.1 m,槽型截面宽 0.1 m,高 0.06 m,壁厚 5 mm,隔板厚度 3 mm,隔板间距 0.1 m,材料弹性模量为 71GPa,材料的密度为

2780 kg/m³。测试结构如图 2 所示。通过锤击法模态测试测量结构的固有频率。将结构用软弹簧悬挂,设置 LMS 振动噪声测试系统的采样时间为 8 s,分析带宽 1024 Hz,触发电压为 65.399 mV,测量的平均次数为 5 次。由于所需测量频带较宽,因此使用硬锤头,测试方法采用单输入多输出识别法。沿薄壁结构轴向布置多个加速度传感器,加速度传感器的响应信号由 LMS 系统采集。结构测试的示意图如图 8 所示,薄壁结构部分测量点的加速度频响函数 Bode 图如图 9 所示。

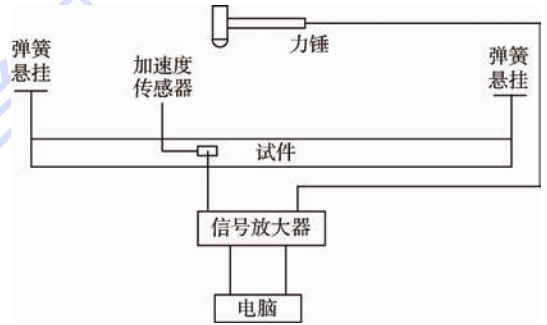


图8 实验模态测试示意图

Fig.8 Schematic diagram of experimental modal testing

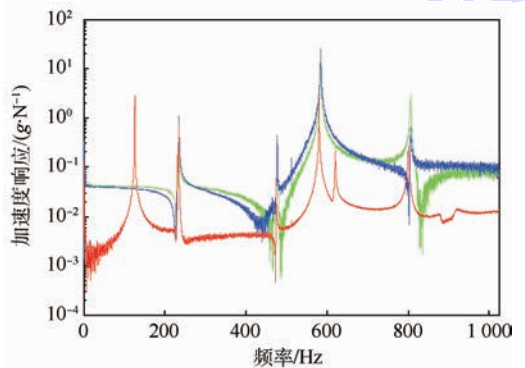


图9 不同测点的加速度频响函数 Bode 图

Fig.9 Bode diagram of acceleration frequency response function at different points

根据频响函数可以得到结构的各阶频率,等效截面参数的辨识结果如表 4 所示,表 5 为等效梁单元计算频率与实验频率对比。图 10 为各阶模态频率误差。由图 10 计算结果可以看出,由梁单元计算所得的频率与实验频率误差基本在 2% 以内,计算与实验误差主要来源于 3 个方面:①结构的几何模型误差,由于薄壁结构中往往存在着大量的圆角,而圆角对结构的扭转频率影响较大,因此圆角的加工误差通常会造造成结构频率的改变从而导致理论模型与真实结构的差异;②理论模型或者有限元模型本身的计算精度所引入的误差,由于有限元模型通常只能在低频段有较高的

表 4 等效截面参数辨识结果

Table 4 Equivalent cross-section parameter identification results

截面参数	梁单元模型			传递矩阵模型	解析模型
	二范数	无穷范数	一范数		
$I_x/(10^{-7}m^4)$	3.07	3.07	3.07	3.07	3.07
$I_y/(10^{-6}m^4)$	1.20	1.23	1.19	1.17	1.22
$J_\varphi/(10^{-8}m^4)$	1.09	1.05	1.07	1.06	1.07
$J_w/(10^{-10}m^6)$	2.11	2.06	2.16	2.15	2.10
e_1/m	0	0	0	0	0
e_2/m	0.0645	0.0631	0.0656	0.0663	0.0642

表 5 等效梁实验频率与计算频率比较

Table 5 Comparison between experimental frequency and calculated frequency of equivalent beam Hz

模态阶次	实验频率	计算频率				
		二范数	无穷范数	一范数	传递矩阵模型	解析模型
1	126.8	125.6	128.1	125.9	126.5	125.1
2	231.1	231.8	233.3	227.9	232.5	227.5
3	234.9	233.4	233.9	232.1	236.8	234.9
4	477.4	478.4	480.4	482.0	482.3	486.2
5	585.2	586.9	577.3	588.4	579.7	581.3
6	621.6	625.6	626.7	622.9	626.1	623.7
7	806.4	815.6	791.3	814.8	811.4	797.5

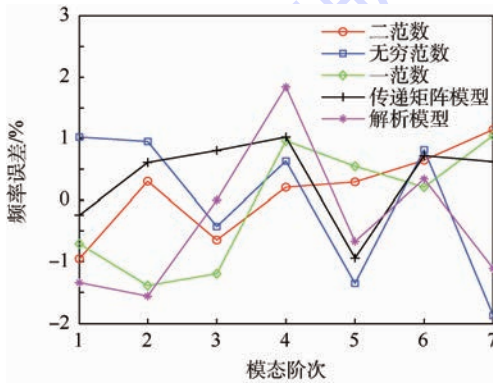


图 10 模态频率误差

Fig. 10 Modal frequency error

计算精度,而等效梁模型本身就是对真实结构的近似,这是计算与实验之间误差的主要来源;③优化程序与目标函数所带来的误差,由表 4 的结果可以看出不同的目标函数对计算结果的影响较大,同时优化算法本身的性能也会对结果产生一定的影响。

6 结 论

1) 本文提出用开口薄壁梁来等效近似计算有均匀分布横向加强构件的薄壁杆件,这种等效方法不仅适用于有隔板的构件,对于缀板或缀条加强的薄壁杆件也有同样的适用性,具有一定的通用性。

2) 本文建立了 3 种开口薄壁梁模型,通过对

比不同模型辨识结果说明目标函数为二范数与一范数的有限元模型的精度普遍高于其他模型。

3) 本文提出了通过实体有限元在静态载荷作用下产生的静态位移来估算等效截面的扭转常数与翘曲常数,并且这种估计方法有较强的通用性。

4) 本文所使用的 ABAQUS 与 MATLAB 交互式的参数优化辨识方法,可以有效地将 MATLAB 中各种优化算法与有限元程序相结合,实现了对 ABAQUS 的二次开发。

5) 用梁来等效替代横向构件作用的薄壁结构可以大幅度缩减单元的数量,同时在一定的频率范围内有较高的精度。

6) 本文所提出的等效模型物理意义明确,不仅适用于结构模型修正,同时也适用于结构动态载荷识别与损伤识别,模型应用范围广。

参考文献 (References)

[1] BROWN A M, SEUGLING R M. Using plate finite elements for modeling fillets in global response analysis[J]. Finite Elements in Analysis and Design, 2004, 40(13-14): 1963-1975.

[2] ARPACI A, BOZDAG E. On free vibration analysis of thin-walled beams with nonsymmetrical open cross-sections [J]. Computers and Structures, 2002, 80(7-8): 691-695.

[3] ARPACI A, BOZDAG S E, SUNBULOGLU E. Triply coupled vibrations of thin-walled open cross-section beams including rotary inertia effects[J]. Journal of Sound and Vibration, 2003, 260(5): 889-900.

[4] PROKIĆ A. On triply coupled vibrations of thin-walled beams with arbitrary cross-section [J]. Journal of Sound and Vibration, 2005, 279(3-5): 723-737.

[5] AMBROSINI R D. A modified Vlasov theory for dynamic analysis of thin-walled and variable open section beams [J]. Engineering Structures, 2000, 22(8): 890-900.

[6] AMBROSINI R D, RIERA J D, DANESI R F. Dynamic analysis of thin-walled and variable open section beams with shear flexibility [J]. International Journal for Numerical Methods in Engineering, 1995, 38(17): 2867-2885.

[7] PROKIĆ A. On fivefold coupled vibrations of Timoshenko thin-walled beams [J]. Engineering Structures, 2006(28): 54-62.

- [8] YAMAN Y. Vibrations of open-section channels: A coupled flexural and torsional wave analysis [J]. *Journal of Sound and Vibration*, 1997, 204(1) : 131-158.
- [9] 包世华,周坚. 薄壁杆件结构力学 [M]. 北京: 中国建筑工业出版社, 2006: 190-193.
BAO S H, ZHOU J. Structural mechanics of thin-walled rod member [M]. Beijing: China Building Industry Press, 2006: 190-193 (in Chinese).
- [10] 王晓峰,张其林,杨庆山. 新型空间薄壁梁单元 [J]. *应用数学和力学*, 2010, 31(9) : 1089-1100.
WANG X F, ZHANG Q L, YANG Q S. New finite element of spatial thin-walled beams [J]. *Applied Mathematics and Mechanics*, 2010, 31(9) : 1089-1100 (in Chinese).
- [11] MOHRI F, EDDINARI A, DAMIL N, et al. A beam finite element for non-linear analyses of thin walled elements [J]. *Thin-Walled Structures*, 2008, 46(7-9) : 981-990.
- [12] KIM N, KIM M Y. Exact dynamic/static stiffness matrices of non-symmetric thin-walled beams considering coupled shear deformation effects [J]. *Thin-Walled Structures*, 2005, 43(5) : 701-734.
- [13] AMBROSINI D. On free vibration of nonsymmetrical thin-walled beams [J]. *Thin-Walled Structures*, 2009, 47(6-7) : 629-636.
- [14] HAN S P. Super linearly convergent variable metric algorithms for general nonlinear programming problems [J]. *Mathematical Programming*, 1976, 11(1) : 263-282.
- [15] HAN S P. A globally convergent method for nonlinear programming [J]. *Journal of Optimization Theory and Applications*, 1977, 22(3) : 297-309.
- [16] POWELL M J D. A fast algorithms for nonlinearly constrained optimization calculations [M] // WATSON G. *Numerical analysis*. Heidelberg: Springer, 1978: 144-157.
- [17] 钱伟长,林鸿荪,胡海昌. 弹性柱体扭转理论 [M]. 北京: 科学出版社, 1956: 178-179.
QIAN W C, LIN H S, HU H C. Torsion bar elastic theory [M]. Beijing: Science Press, 1956: 178-179 (in Chinese).

作者简介:

邓昊 男, 硕士研究生。主要研究方向: 计算结构动力学。

E-mail: 15313257835@163.com

程伟 男, 博士, 教授, 博士生导师。主要研究方向: 结构动力学。

E-mail: cheng_wei@buaa.edu.cn

Equivalent modeling method of open thin-walled beam under action of transverse stiffening member

DENG Hao, CHENG Wei*

(School of Aeronautic Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: Satellite structures usually have a lot of open thin-walled beams under the action of transverse stiffening member. Transverse member is generally evenly distributed along the axial direction of the beam. Through the study of such structures, it is theoretically proved that the differential equation of vibration of this structure has the same form as the equation of ordinary open thin-walled beam. Therefore, equivalent calculation was carried out using the open thin-walled beam element. Three kinds of mathematical models of open thin-walled beams, finite element model, transfer matrix model and analytical model, were established. The equivalent cross-section parameters were identified using sequential quadratic programming. At the same time, the influence of different objective functions on the identification results was analyzed. And a method for estimating the initial parameters of the cross-section was presented. For the finite element model, the MATLAB and ABAQUS interactive parameters optimization method was proposed. A combination of both full advantages can quickly and efficiently optimize the cross-section parameters and this method has strong versatility. Finally, the correctness and accuracy of the equivalent modeling method are verified by experiments. The proposed equivalent modeling method can reduce more than 90% of the number of elements. By establishing the simplified model, the efficiency of the structural model updating and structure reanalysis can be greatly improved.

Key words: transverse stiffening member; open thin-walled beam; differential equation of vibration; equivalent modeling; parameter optimization identification

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0464

利于减少配平损失的太阳能飞机构型设计



阙建锋, 王维军*, 吴宇

(北京航空航天大学 航空科学与工程学院, 北京 100083)

摘 要: 针对具有静稳定性的太阳能飞机一般采用尾翼配平造成气动配平损失的问题,提出了一种“T”构型太阳能飞机。概述该太阳能飞机的总体构型,分析该构型飞机巡航状态下降低重心的自配平原理,建立能量平衡和质量分析模型,给出适合于该构型太阳能飞机的概念设计方法,并优化构型设计参数。结果表明,“T”构型太阳能飞机不仅具有静稳定性,还降低了配平损失。在巡航状态下,“T”构型太阳能飞机单位面积平飞需用功率比常规构型太阳能飞机减少了6.2%,具有明显的应用效果。

关键词: 太阳能飞机; 气动配平损失; 概念设计; 构型设计参数; 平飞需用功率

中图分类号: V221+.2

文献标识码: A

文章编号: 1001-5965(2016)07-1479-07

太阳能飞机通过太阳能电池接受太阳能辐射,将太阳能转化为电能。白天将一部分能量提供给飞行所需的能耗并将剩余的能量储存在蓄电池中,提供飞机夜间飞行。太阳能飞机的能量来源优势,让它可以巡航飞行几个月甚至更久^[1-5]。因此太阳能飞机在中继通信,情报、监视和侦察(ISR),野火警告系统,农业协助,边境巡逻,污染和核能的观测等领域具有广阔的应用前景,成为未来军事和民用发展的热点^[6-7]。

但是光伏电池的能源密度比汽油等能源低许多,限制了太阳能飞机的设计空间。于是考虑到能源短缺问题,太阳能飞机的设计和传统飞机不同,太阳能飞机的翼载荷和飞行功率较小。因此太阳能飞机的结构密度,能量采集和传递效率,能源的储存和管理^[8],推进系统以及飞机构型需要进行详细设计进而达到能源循环的要求,以期实现长航时飞行^[9]。

现有的太阳能飞机构型中,常规构型太阳能飞机占多数。对于非常规构型的太阳能飞机研究

较少,例如鸭式构型的 Solair I 号^[10]、飞翼构型的太阳神号^[11],以及“Z”型构型。其中“Z”型构型可以是美国秃鹰计划候选机型的构型,采用多机身尾翼配平^[9],也可以是飞翼构型^[12]。上述太阳能飞机都采用舵面偏转配平。考虑到太阳能飞机翼载荷较轻,结合新构型结构设计的可行性,提出“T”构型太阳能飞机。其原理是在巡航状态下采用降低重心的自配平方法,减少了气动配平损失。然后建立太阳能飞机数学模型,研究适合该构型的优化设计方法,分析其巡航状态的配平原理和静稳定性,并与常规构型太阳能飞机进行对比,分析该构型的应用效果,以及定量分析不同配平迎角下对应的操纵量。结果表明,“T”构型太阳能飞机具有较好的纵向静稳定性,降低了单位面积平飞需用功率,减少了舵面偏转配平带来的配平损失。

1 “T”构型的提出

1.1 现有太阳能飞机构型

为了实现长航时,获取更多的太阳能和减少

收稿日期: 2015-07-13; 录用日期: 2015-08-20; 网络出版时间: 2015-10-09 15:46

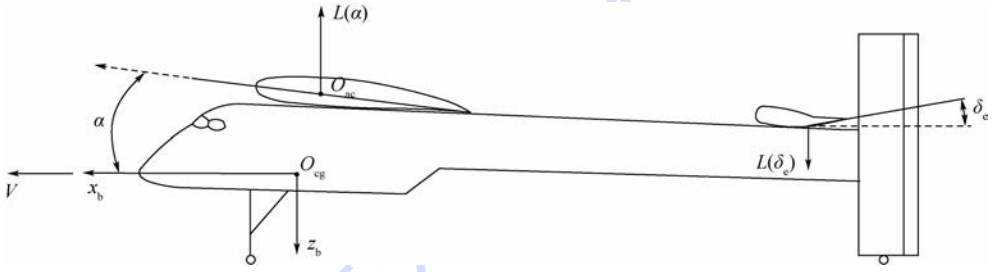
网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151009.1546.005.html

* 通讯作者: Tel.: 010-82317519 E-mail: wangwj@buaa.edu.cn

引用格式: 阙建锋, 王维军, 吴宇. 利于减少配平损失的太阳能飞机构型设计[J]. 北京航空航天大学学报, 2016, 42(7): 1479-1485. QUE J F, WANG W J, WU Y. Design of solar-powered aircraft configuration for reducing trim loss[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1479-1485 (in Chinese).

飞行能耗,成为现有太阳能飞机构型设计研究的重点。波音提出了一种非常规构型的太阳能飞机“太阳鹰”,采用的是多机身布局^[13]。每个机身上安装的主动光伏组件可根据太阳辐射角情况绕机身轴线转动来获得更多的太阳能。但是这种方式会带来阻力和结构质量的代价。而飞翼构型的太阳能飞机增大了飞机的面积,减轻了结构质量,在此基础上文献[12]提出的折叠机翼设计可获得更多的太阳能。但是舵面配平时力臂减小,则需要的舵面升力增加,造成全机升力减小。放宽

静稳定性对低速太阳能飞机控制系统也有要求。对于具有纵向静稳定性的飞机,其需要舵面偏转操纵产生配平所需的升力,而且这部分升力与迎角产生升力方向相反,造成全机升力损失一部分,如图1所示。而太阳能飞机飞行速度较低,并且不追求高机动性,具有静稳定性是有意义的。现有的太阳能飞机大多采用的是舵面偏转配平。上述舵面偏转配平方式,对于具有静稳定性的太阳能飞机而言,伴随着气动配平损失带来的飞行能耗。



V—飞行速度; α —迎角; x_b, z_b —机体轴; O_{cg} —飞机的重心; O_{ac} —飞机的气动焦点;
 $L(\alpha)$ —升力; $L(\delta_e)$ —升降舵配平升力; δ_e —操纵量。

图1 常规构型定直飞行时升力情况

Fig. 1 Lift of conventional configuration in direct flight

1.2 “T”构型的配平原理

针对具有静稳定性太阳能飞机的气动配平损失带来的持续能耗的问题,提出了一种“T”构型布局方式,如图2所示。机翼采用正弯度翼型,在机翼中段后缘处布置了操纵舵面。其下方有一个垂直的支撑面,一端连接机翼,另一端连接机身。由于太阳能飞机翼载荷较低,对T构型结构设计有利。机身可以放置蓄电池和负载。太阳能飞机的动力装置布置于机身前端和机翼前缘处。太阳能电池布置在机翼上表面。

分析飞机的纵向受力情况,如图3所示,并建立飞机纵向的平衡方程。图中:来流速度与机翼平均气动弦夹角为 α ,平均气动弦相对机体轴的安装角为 φ 。在纵向对称面内,飞机受到升

力 L 、阻力 D 、推力 T 、重力 mg (m 为飞机质量, g 为重力加速度) 以及正弯度翼型零升俯仰力矩 M_0 。

飞机在不同飞行迎角下定直飞行的平衡方程为

$$\begin{cases} T = D \cos(\alpha - \varphi) - L \sin(\alpha - \varphi) \\ mg = L \cos(\alpha - \varphi) + D \sin(\alpha - \varphi) \\ M_0 + M_T - (L \cos(\alpha - \varphi) + D \sin(\alpha - \varphi)) \Delta x_{ac} - (L \sin(\alpha - \varphi) - D \cos(\alpha - \varphi)) \Delta z_{ac} = 0 \end{cases} \quad (1)$$

式中: Δz_{ac} 为焦点与重心距离沿垂直支撑面方向的投影; Δx_{ac} 为焦点与重心距离沿机体轴方向的投影; M_T 为推力对重心的力矩。飞机在不同飞行迎角下定直飞行时,气动力是 α 和 δ_e 的函数,因

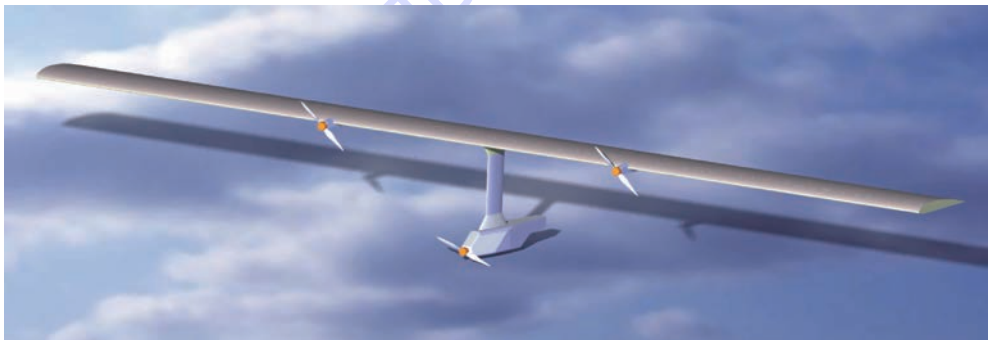


图2 “T”构型示意图

Fig. 2 Sketch map of “T” configuration

此全机俯仰力矩特性不仅和迎角、重心位置有关,还和操纵量有关。

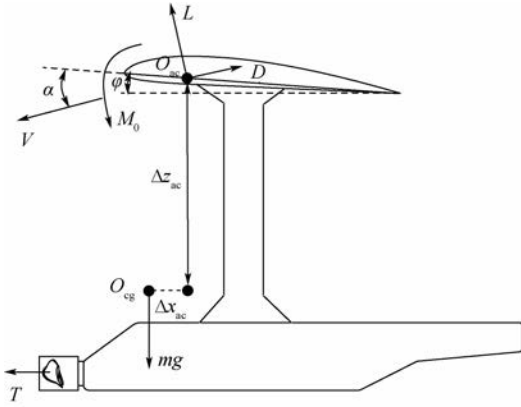


图3 “T”构型飞机定直飞行时受力情况

Fig.3 Force conditions of “T” configuration in direct flight

以机体坐标系为参照,飞机巡航飞行时, $\alpha = \varphi$, 给出飞机巡航飞行时的平衡方程:

$$\begin{cases} T = D \\ L = mg \\ M_0 + M_T + D\Delta z_{ac} - L\Delta x_{ac} = 0 \end{cases} \quad (2)$$

可以看出当飞机巡航飞行时,截面为正弯度翼型的机翼产生的低头力矩和机翼气动力产生的抬头力矩平衡时,飞机实现自配平。这种配平方式利用大面积机翼产生的气动力,提供抬头力矩,在巡航状态下采用降低重心的自配平方法,减少了气动配平损失。

2 “T”构型参数设计的数学模型

2.1 能量平衡模型

太阳能飞机飞行时,升力系数为 C_L , 阻力系数为 C_D , 大气密度为 ρ , 机翼面积为 S 。全机升力可以表示为

$$L = \frac{1}{2}\rho V^2 C_L \quad (3)$$

全机阻力可以表示为

$$D = \frac{1}{2}\rho V^2 C_D \quad (4)$$

由于升力等于重力,结合式(3),可得

$$V = \sqrt{\frac{2mg}{C_L \rho S}} \quad (5)$$

由于推力等于阻力,结合式(4)和式(5)可得推进系统平飞需用功率与机翼面积比为

$$\frac{P_p}{S} = \frac{TV}{S} = \sqrt{\frac{2g^3}{\rho}} \cdot \frac{C_D}{C_L^{3/2}} \left(\frac{m}{S}\right)^{3/2} \quad (6)$$

式中: P_p 为推进系统平飞需用功率; $C_D/C_L^{3/2}$ 为功率因子。

太阳能飞机用电设备包括推进装置和机载设备,飞机平飞需用总功率为

$$P_{tot} = P_p / (\eta_{plr} \eta_{mot}) + P_{eq} \quad (7)$$

式中: η_{plr} 为螺旋桨功率; η_{mot} 为电机功率; P_{eq} 为机载设备功率。

太阳能飞机的能源动力系统通过太阳能电池板获取太阳能,传递、存储,最后转化为自身飞行所需的动力。在这个过程中必然伴随着能量损失。能源动力系统及能源利用率如图4^[14]所示,MPPT为最大功率跟踪点。

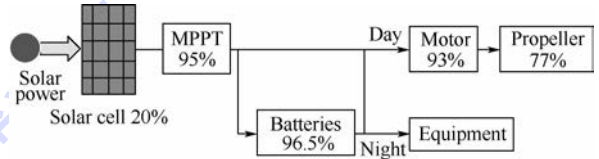


图4 能源动力系统及能源利用率^[14]

Fig.4 Energy power system and energy utilization efficiency^[14]

太阳能飞机通过光伏电池将太阳能转化为电能,太阳能电池输出功率^[15]为

$$P = I_0 \tau \eta_{sc} S_{sc} \sin h \quad (8)$$

式中: I_0 为太阳垂直照射强度; τ 为太阳能辐射的衰变系数; η_{sc} 为太阳能电池转化效率; S_{sc} 为太阳能电池的面积; h 为太阳高度角。

白天太阳能电池产生的能量供给飞机正常运转。多余的能量则存储于蓄电池中,在太阳辐射量不足或夜间维持飞机正常飞行。设白天的时间为 T_d , 夜晚的时间为 T_n 。为了让太阳能飞机能连续24 h不间断飞行,获得的太阳能总能量应该满足飞机飞行所需能量。假设飞机在某高度巡航飞行,则能量平衡方程为

$$\int_{t_r}^{t_s} P dt \cdot \eta_{mppt} = P_{tot} (T_d + T_n \frac{1}{\eta_{bat}}) \quad (9)$$

式中: t_r 为日出时刻; t_s 为日落时刻; η_{mppt} 为最大功率跟踪器效率; η_{bat} 为蓄电池效率; t 为时间。根据能量平衡方程可推出单位时间单位面积太阳能电池获得的能量和单位面积的平飞需用功率关系为

$$\int_{t_r}^{t_s} \frac{P}{S_{sc}} dt = \frac{S}{\eta_{mppt} S_{sc}} \left(T_d + T_n \frac{1}{\eta_{bat}} \right) \frac{P_{tot}}{S} \quad (10)$$

2.2 质量分析模型

太阳能飞机的质量由机体结构、太阳能电池、蓄电池、推进系统和负载的质量组成,即

$$m = m_{af} + m_{sc} + m_{bat} + m_p + m_{eq} \quad (11)$$

式中: m_{af} 、 m_{sc} 、 m_{bat} 、 m_p 和 m_{eq} 分别为机体结构、太阳能电池、蓄电池、推进系统和负载的质量。各个分系统的质量为

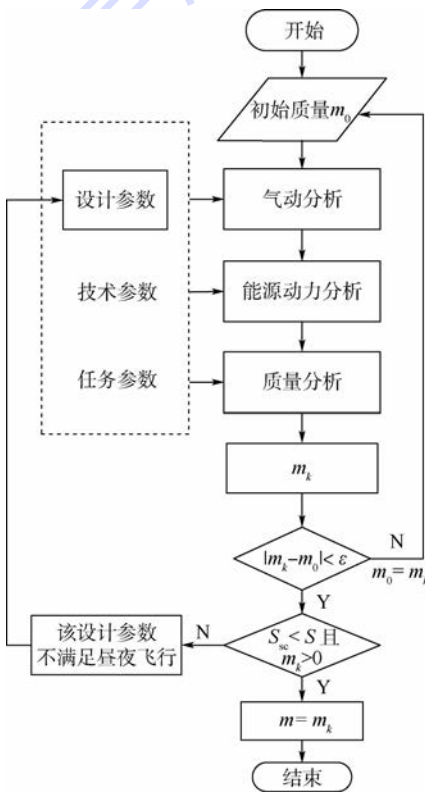
$$\begin{cases} m_{\text{aft}} = \rho_{\text{aft}} AR^{x_2} b^{x_1} + \rho_{\text{vs}} S_{\text{vs}} \\ m_{\text{sc}} = \rho_{\text{sc}} S_{\text{sc}} \\ m_{\text{bat}} = (P_{\text{tot}} T_n / \eta_{\text{bat}}) \rho_{\text{bat}} \\ m_p = P_p / \rho_p \end{cases} \quad (12)$$

式中： ρ_{aft} 为机身结构的密度；AR为展弦比； b 为展长； x_1 和 x_2 分别为与展长和展弦比对应的指数因子^[16]； ρ_{vs} 为垂直支撑面的密度； S_{vs} 为垂直支撑面的面积； ρ_{sc} 为太阳能电池密度； ρ_{bat} 为蓄电池能量密度； ρ_p 为推进装置功重比。

2.3 总体参数设计方法

太阳能飞机的所有参数可分为3类：一类是技术参数，和太阳能飞机的各子系统的技术水平有关；另一类是任务参数，和太阳能飞机需要完成的设计任务有关；还有一类是设计参数，和太阳能飞机的构型以及飞行状态有关，在这里选择机翼展弦比、展长、重心位置以及垂直支撑面长度作为设计参数。

太阳能飞机的总体参数众多，而且可以互相推导，根据飞机的质量组成，可以很好地得到其质量的迭代方程，所以选取质量作为主要的迭代参数，求解出质量可以推导出其他的总体参数。程序的算法框图见图5。根据质量分析模型方程式(11)和式(12)，联立式(6)和式(7)得到质量迭代方程 $m_{k+1} = f(m_k)$ ， $f(m_k)$ 为各子系统质量函数之



ϵ —误差精度。

图5 程序的算法框图

Fig. 5 Block diagram of program algorithm

和 k 为迭代次数。选取初值 m_0 可迭代求解飞机的质量，若为正实根，则方程有解，同时代入方程式(9)判断飞机是否可以实现昼夜飞行。如果满足昼夜能量平衡方程，则设计参数满足要求；如果不满足昼夜平衡方程，则改变设计参数，继续进行质量迭代。通过判断是否满足昼夜飞行条件，可以将符合条件的构型设计参数可行域确定下来。以巡航飞行功率和静稳定性为设计指标，从可行构型中确定最优构型。

3 结果分析与讨论

“T”构型与常规构型太阳能飞机最大区别是降低重心，增加了垂直支撑面连接机身和机翼。因此主要分析重心位置和垂直支撑面的设计参数。对于静稳定的太阳能飞机，其静稳定裕度有一定要求，假设静稳定性边界为 $\partial C_m / \partial C_L = -0.15$ (C_m 为俯仰力矩系数)。任务指标如表1所示。

表1 太阳能飞机的任务指标

Table 1 Mission index of solar-powered aircraft

任务指标	纬度/(°)	日期	h/km	P_{eq}/W	m_{eq}/kg
数值	40	春分日	8.0	100	150

3.1 静稳定性分析

设飞机重心坐标为 $(x_{\text{cg}}, 0, z_{\text{cg}})$ ，坐标原点取平均气动弦前缘点在机身轴线上的投影点，其重心坐标进行无量纲化， x_{cg} 为真实位置除以平均气动弦长， z_{cg} 为真实位置除以垂直支撑面长度。 z_{cg} 和飞机机翼尺寸有关， z_{cg} 和展弦比、展长的关系如图6所示。因此主要研究太阳能飞机垂直支撑面的长度和重心坐标位置 x_{cg} 对飞机静稳定性的影响。图7(a)所示为垂直支撑面的长度 l 对俯仰力矩曲线的影响。可知，随着垂直支撑面长度的增加，俯仰力矩系数对迎角的导数 $C_{m\alpha}$ 增大，飞机的静稳定性增强。

在一定移动范围内，飞机重心坐标位置 x_{cg} 可

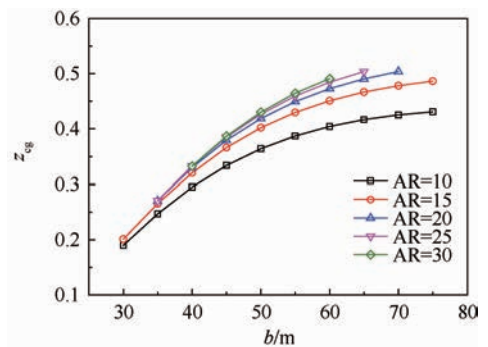


图6 z_{cg} 和展弦比、展长的关系

Fig. 6 Relationship between z_{cg} and aspect ratio and span

以通过质量布置调整,并不影响飞机的总质量大小和升阻特性。图 7(b)为重心坐标位置 x_{cg} 对俯仰力矩曲线的影响。可见和舵面偏转相似,改变重心坐标沿机身方向的位置可以改变配平点,因此可将配平迎角设计在最佳的巡航迎角处。但从俯仰力矩系数对迎角的导数来看,改变重心坐标位置 x_{cg} 对静稳定性也有一定影响。

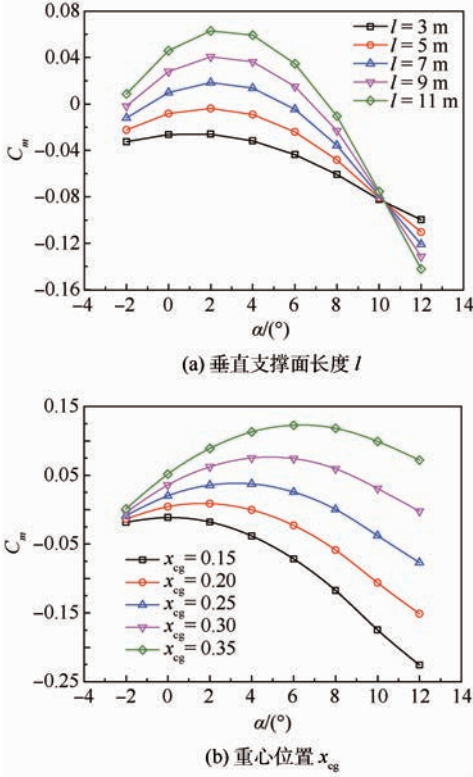


图 7 垂直支撑面长度 l 与重心位置 x_{cg} 对俯仰力矩曲线的影响

Fig. 7 Influences of length of vertical supporting surface l and centre of gravity position x_{cg} on pitch moment curves

3.2 平飞需用功率分析

由式(10)可知,对于太阳能飞机,不仅希望单位时间单位面积太阳能电池获得的太阳能越多,而且希望平飞需用功率和面积的比值越小。前者是由太阳能电池的效率和入射角决定,后者是由飞机构型决定。故研究机翼展弦比和展长变化对单位面积平飞需用功率的影响,如图 8 所示。从图中可以看出,仅考虑机翼展弦比和展长变化情况下,单位面积平飞需用功率存在最小值。

垂直支撑面长度对单位面积平飞需用功率和静稳定性的影响如图 9 所示。可见,单位面积平飞需用功率随垂直支撑面长度减小而降低,但 $\partial C_m / \partial C_L$ 增加,即静稳定性降低, $\partial C_m / \partial C_L$ 需满足最小稳定裕度要求。

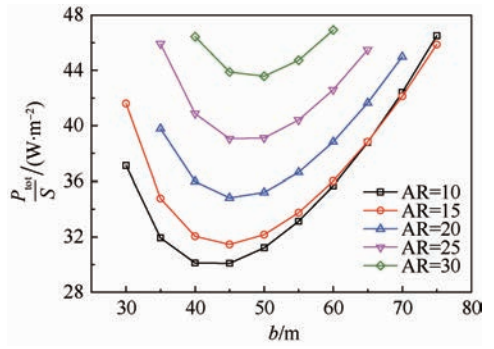


图 8 P_{tot}/S 和 AR、 b 关系

Fig. 8 Relationship between P_{tot}/S and AR, b

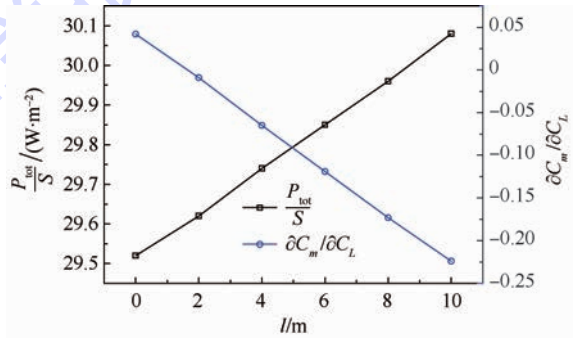


图 9 垂直支撑面长度对 P_{tot}/S 和静稳定性的影响

Fig. 9 Influences of length of vertical supporting surface on P_{tot}/S and static stability

3.3 对比研究

在相同的技术任务参数下,满足静稳定裕度 $\partial C_m / \partial C_L = -0.15$,当展长为 42.7 m,展弦比为 10.5,重心坐标为(0.22, 0, 0.32),垂直支撑面长度为 7.1 m 时,“T”构型太阳能飞机最小单位面积平飞需用功率为 29.91 W/m²。而满足相同静稳定裕度情况下,常规构型太阳能飞机的最小单位面积平飞需用功率为 31.77 W/m²。与之相比,该构型太阳能飞机的单位面积平飞需用功率减小了 6.2%,应用效果明显。

3.4 操纵特性分析

为了实现太阳能飞机以不同速度定直平飞,在机翼中段后缘处布置可上下偏转的操纵舵面。其操纵力矩对俯仰力矩系数的影响如图 10(a)所示。不同迎角下定直飞行采用舵面偏转进行配平,其静稳定性是变化的。由图 7(b)可知, x_{cg} 变化对配平点是有影响的,同时占全机质量接近 1/4 的锂电池布置于机身内,因此考虑 x_{cg} 和舵面偏转的组合操纵。其俯仰力矩曲线如图 10(b)所示。组合操纵不仅满足配平需求还改善了配平点的静稳定性。

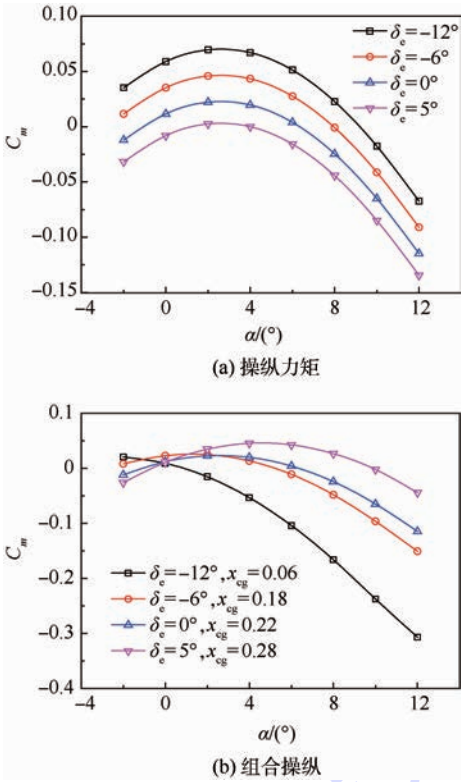


图 11 操纵力矩和组合操纵对俯仰力矩系数的影响

Fig. 11 Influences of operating moment and combined control on pitch moment coefficient

4 结 论

本文针对气动配平损失,提出了“T”构型太阳能飞机,经建模仿真表明:

1) 该构型太阳能飞机采用降低重心的自配平原理,降低了巡航状态下气动配平损失。增加垂直支撑面的长度有利于提高飞机的静稳定性,但也增加了飞机的阻力和质量。根据静稳定裕度边界要求,设计垂直支撑面的长度。改变重心坐标沿机身方向的位置可以改变配平点,可将巡航配平迎角设计在最小平飞功率对应迎角处。

2) 满足任务要求,当静稳定裕度 $\partial C_m / \partial C_L = -0.15$ 时,以单位面积平飞需用功率最小为目标,“T”构型太阳能飞机存在最佳设计参数:展长为 42.7 m,展弦比为 10.5,重心坐标为 (0.22, 0, 0.32),垂直支撑面长度为 7.1 m。

3) 与常规构型太阳能飞机相比,该构型太阳能飞机的单位面积平飞需用功率减小了 6.2%,降低了配平损失,应用效果明显。

4) x_{cg} 变化和舵面偏转的组合操纵不仅满足以不同迎角飞行时的配平需求,还改善了配平点的静稳定性。

综上所述,定量分析了“T”构型太阳能飞机的配平原理和设计方法,该构型不仅能减少配平损失,而且满足静稳定性设计要求。今后还将对飞机结构设计、纵向动稳定性以及横侧向稳定性进行进一步研究,优化与完善构型设计。

参考文献 (References)

- [1] IRVING F G, MORGAN D. The feasibility of an aircraft propelled by solar energy [C] // Proceedings of the AIAA/MIT/SSA 2nd International Symposium on the Technology and Science of Low Speed and Motorless Flight. Reston: AIAA, 1974: 1-9.
- [2] YOUNGBLOOD J W, TALAY T A. Solar-powered airplane design for long-endurance, high-altitude flight [C] // AIAA 2nd International Very Large Vehicles Conference. Reston: AIAA, 1982: 1-8.
- [3] BRANDT S A, GILLIAM F T. Design analysis methodology for solar-powered aircraft [J]. Journal of Aircraft, 1995, 32 (4): 703-709.
- [4] ENRICO C. Design of solar high altitude long endurance aircraft for multi payload and operations [J]. Aerospace Science and Technology, 2006, 10 (6): 541-550.
- [5] MEYER J, DU PLESSIS J A F, ELLIS P, et al. Design considerations for a low altitude long endurance solar powered unmanned aerial vehicle [C] // AFRICON, 2007. Piscataway, NJ: IEEE Press, 2007: 1-7.
- [6] LEUTENEGGER S, JABAS M, SIEGWART R Y. Solar airplane conceptual design and performance estimation [J]. Journal of Intelligent & Robotic Systems, 2011, 61 (1-4): 545-561.
- [7] ZHU X F, GUO Z, FAN R F, et al. How high can solar-powered airplanes fly [J]. Journal of Aircraft, 2014, 51 (5): 1653-1659.
- [8] GAO X Z, HOU Z X, GUO Z, et al. Energy management strategy for solar-powered high-altitude long-endurance aircraft [J]. Energy Conversion and Management, 2013, 70: 20-30.
- [9] ZHU X F, GUO Z, HOU Z X, et al. Solar-powered airplanes: A historical perspective and future challenges [J]. Progress in Aerospace Sciences, 2014, 71: 36-53.
- [10] BOUCHER R J. History of solar flight [C] // AIAA/SAE/ASME. 20th Joint Propulsion Conference. Reston: AIAA, 1984: 14-29.
- [11] NOLL T E, BROWN J M, PEREZ-DAVIS M E, et al. Investigation of the helios prototype aircraft mishap: NASA Report 64317 [R]. Washington, D. C.: NASA, 2004.
- [12] MARDANPOUR P, HODGES D H. Passive morphing of flying wing aircraft: Z-shaped configuration [J]. Journal of Fluids and Structures, 2014, 44: 17-30.
- [13] 马东立, 包文卓, 乔宇航. 利于冬季飞行的太阳能飞机构型研究 [J]. 航空学报, 2014, 35 (6): 1581-1591.
MA D L, BAO W Z, QIAO Y H. Study of solar-powered aircraft configuration beneficial to winter flight [J]. Acta Aeronautica et Astronautica Sinica, 2014, 35 (6): 1581-1591 (in Chinese).
- [14] ROSS H. Fly around the world with a solar-powered airplane

- [C] // Proceedings of the 26th Congress of International Council of the Aeronautical Science. Reston: AIAA, 2008: 1-11.
- [15] 沈辉, 曾祖勤. 太阳能光伏发电技术[M]. 北京: 化学工业出版社, 2005: 6-14.
- SHEN H, ZENG Z Q. Solar photovoltaic power generation technology[M]. Beijing: Chemical Industry Press, 2005: 6-14 (in Chinese).
- [16] NOTH A, SIEGWART R, ENGEL W. Design of solar powered airplanes for continuous flight[D]. Zürich: ETH Zürich, 2008.

作者简介:

阙建锋 男 硕士研究生。主要研究方向: 太阳能飞机概念设计和建模仿真。

E-mail: jfq7708@126.com

王维军 男, 博士, 副教授。主要研究方向: 飞行动力学仿真与控制、微小型飞行器设计试验。

Tel.: 010-82317519

E-mail: wangwj@buaa.edu.cn

Design of solar-powered aircraft configuration for reducing trim loss

QUE Jianfeng, WANG Weijun*, WU Yu

(School of Aeronautic Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: The solar-powered aircraft with the static stability generally uses tail trim, which causes the trim loss. To address the problem of the aerodynamic trim loss, a new solar-powered aircraft configuration called "T" configuration was put forward. The general solar-powered aircraft configuration was summarized. The trimming principle by lowering center of gravity in the state of cruise flight was analyzed. The models of energy balance and mass analysis were built. A concept design method which is suitable for the new solar-powered aircraft configuration was introduced, and then the design parameters of the aircraft configuration were optimized. The results show that the solar-powered aircraft of "T" configuration not only has static stability, but also reduces the trim loss. In the state of cruise flight, the solar-powered aircraft of "T" configuration can decrease the flying power required in unit area by 6.2% compared with the normal solar-powered aircraft, which shows obvious application effect.

Key words: solar-powered aircraft; aerodynamic trim loss; concept design; configuration design parameters; flying power required

Received: 2015-07-13; Accepted: 2015-08-20; Published online: 2015-10-09 15:46

URL: www.cnki.net/kcms/detail/11.2625.V.20151009.1546.005.html

* Corresponding author. Tel.: 010-82317519 E-mail: wangwj@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0445

图模型与学习算法结合的贝叶斯网络自动建模

沈琳¹, 于劲松^{1,2,*}, 唐荻音¹, 刘浩^{1,3}

(1. 北京航空航天大学 自动化科学与电气工程学院, 北京 100083; 2. 北京航空航天大学 先进航空发动机协同创新中心, 北京 100083;
3. 解放军 95809 部队, 93 分队, 沧州 061736)

摘 要: 针对纯数据驱动的贝叶斯网络结构学习算法的准确度和效率较低的问题, 提出了一种融合多信号流图模型与 K2 学习算法的贝叶斯网络自动建模方法。该方法利用多信号流图模型能够描述信号之间传递与依赖关系的能力, 结合 K2 学习算法在结构学习中的优势, 实现了专家知识与数据驱动方法有效融合的贝叶斯网络结构自动学习算法。通过与常用网络结构学习算法的对比实验证明, 该融合算法显著降低了结构学习对学习范围和训练数据规模的要求, 具有更高的学习准确度和运算效率。采用真实系统实例阐述了该融合算法的应用过程, 验证了算法的实用性。

关键词: 贝叶斯网络; 结构学习; 多信号流图; K2 算法; 故障诊断

中图分类号: TP277

文献标识码: A **文章编号:** 1001-5965(2016)07-1486-08



贝叶斯网络是基于概率分析、图论的一种表达和推理不确定性知识的模型^[1], 因其在概率推理和故障诊断方面具有的显著优势, 被广泛应用于机械、医疗等各个行业。贝叶斯网络的建立通常需要经过结构学习和参数学习 2 个步骤, 其中结构学习难度较大, 成为了制约贝叶斯网络建立的关键问题^[2]。

贝叶斯网络结构学习的主要方法有基于数据驱动的方法和基于专家知识的方法 2 大类。数据驱动方法学习效率低, 建模过程受到数据样本和算法自身缺陷的制约, 难以保证精度^[3]; 基于专家知识的方法完全依赖于咨询专家后的手工构建, 对大型复杂系统建模困难, 且不利于方法的集成化与自动化。但是专家知识含有的先验信息可以提高建模的效率和准确性。因此, 开展数据驱动与专家知识结合的贝叶斯网络结构学习, 具有理论研究和实际工程应用价值^[4-5]。

多信号流图是一种基于专家知识构建的图模

型, 其描述信号之间传递关系的能力能够为贝叶斯网络结构的学习提供重要参考。文献[6]提出了一种将多信号流图翻译为诊断贝叶斯网络的手动建模方法, 该方法虽然能够充分运用多信号流图中的专家知识, 但因其没有引入学习算法, 导致建模开发需要大量经验, 当面对多层次复杂系统时, 手动翻译工作量大、效率低, 且无法判断模型是否最优。文献[7]利用多信号流图的故障-测试相关性矩阵(D -矩阵)实现了自动化结构学习, 但是该方法并没有考虑故障节点之间的影响关系, 因此生成的朴素贝叶斯模型并未充分利用多信号流图蕴含的结构因果关系信息。

本文在上述研究的基础上, 充分利用多信号流图的先验知识, 同时结合基于数据驱动的 K2 学习算法, 实现了诊断贝叶斯网络结构的自动建模。通过对比实验证明了该建模方法的性能和优势, 并通过在某观瞄系统中的应用证明了该算法的工程应用价值。

收稿日期: 2015-07-02; 录用日期: 2015-09-30; 网络出版时间: 2015-10-30 17:00

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151030.1700.010.html

* 通讯作者: Tel.: 010-82338693 E-mail: yujs@buaa.edu.cn

引用格式: 沈琳, 于劲松, 唐荻音, 等. 图模型与学习算法结合的贝叶斯网络自动建模[J]. 北京航空航天大学学报, 2016, 42(7): 1486-1493. SHEN L, YU J S, TANG D Y, et al. Automatic learning of Bayesian network structure using graph model and learning algorithm [J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1486-1493 (in Chinese).

冗余边的情况^[9]。

多信号流图模型与系统物理模型相契合,可以充分利用系统设计方案所包含的专家知识进行建模,具有建模过程简单、包含系统知识完整等优势,并且由于在诊断系统设计阶段多信号流图模型能够提供可测试性分析和故障模式、影响及危害度分析(FMECA),该模型受到大量复杂系统的青睐,应用广泛。多信号流图模型所包含的专家知识和其模型特点为它与贝叶斯模型间系统知识的共享提供了便利。基于多信号流图模型获取先验知识克服了以往专家知识的不足,并且可以实现知识的自动获取,极大地提高了诊断贝叶斯网络结构学习的效率和准确性。

1 多信号流图模型

1.1 多信号流图概述

多信号流图模型采用有向图描述故障传播依赖关系,可用于实现复杂系统诊断建模、算法设计等功能,目前已广泛使用于航空航天、工业控制和医疗设备等高新技术系统的可测性设计、维护管理等领域。多信号流图的主要元件包括模块、子系统、测试、开关、与节点和连接线^[8]。其中,模块元件与故障模式对应;子系统模块则用于实现系统的多层次建模;开关元件可用于表示系统的多模式特性;与节点元件则用于表示系统的冗余特性;连接线表示了信号的流向。

1.2 多信号流图与诊断贝叶斯网络对比

多信号流图与诊断贝叶斯网络均属于图模型,二者在知识表达方面有诸多相似之处,例如^[6]:①2种图模型结构都由节点和有向边2个部分元素组成;②节点代表的均为对象的状态或特征,有向边则表示节点间的关联关系;③节点均可以划分为故障节点和测试节点,前者代表组件或者模块,后者通常是传感器信息;④2种推理模型都支持从原因到结果的正向推理和从结果到原因的反向推理。

但是二者之间也存在显著区别,包括:①多信号流图中的有向连接线反映2个节点之间信号的传递关系,而诊断贝叶斯网络中代表了故障的因果联系,因此二者并不一一对应。②多信号流图模型允许信号流的传递成环形连接,而贝叶斯网络则是严格的有向无环图,这一属性也反映了二者有向含义的不同;③多信号流图模型节点涵盖信息包括该节点的失效率和虚警率,而贝叶斯网络中节点所具有的信息不仅包含节点本身的故障先验概率,还反映父节点状态对该节点的影响关系;④基于多信号流图的诊断过程是定性分析,推理结果往往是一组可疑组件,而基于贝叶斯网络的诊断是定量分析,不仅可以给出每个怀疑组件的故障概率,还可以处理证据冲突的问题。

以往基于专家知识的建模方法均是人工给出先验知识,因此该知识可能是零散的、片面性的,并且通常只是故障关系存在或发生的可能性。基于系统设计方案的专家知识可以完全满足多信号流图的建模需求,但是对于贝叶斯网络而言,就形成了不确定的先验知识,直接采用该知识进行贝叶斯网络建模而忽略先验知识的不确定性,往往会导致网络结构出现丢失边或

2 结构学习算法

2.1 数据驱动算法的选取

基于数据驱动的贝叶斯网络结构学习算法已存在大量研究,受到使用条件的限制或算法准确度的制约,通过某一种算法单独完成高质量贝叶斯网络的建立仍然十分困难^[10]。K2算法作为贝叶斯结构领域的经典算法,具有优异的学习性能,该算法的使用需要满足严格的先验条件^[11]:①提供所有节点 G 的拓扑排序 ρ ;②指定贝叶斯网络中的最大父节点个数 μ 。这2项先验知识通常基于对系统的全面认识和分析,在一般情况下很难获取,因此K2算法的应用具有局限性。多信号流图模型含有变量之间信息流的传递关系,依据对该传递关系的分析可以有效实现对变量的排序并提取出变量可能的父节点个数上限,满足K2算法的使用条件。

传统K2算法的初始模型是一个包含所有节点、但是不含边的无边图,在搜索过程中对节点进行逐个考查,确定其父节点,再添加父节点到当前节点的有向边。K2算法需要通过评分函数对每次添加有向边的操作进行打分,最后寻找出整个网络最优评分值所对应的网络结构。由Cooper和Herskovits提出的CH评分函数^[12]具有可分解、运算复杂度低等优点,因此以CH评分为例,节点 X_i 的家族CH评分为

$$\left\{ \begin{array}{l} \text{CH}(\langle X_i, \pi(X_i) \rangle | D_i) = \\ \sum_{t=1}^{h_i} \left[\lg \frac{\Gamma(a_{it^*})}{\Gamma(a_{it^*} + m_{it^*})} + \sum_{k=1}^{q_i} \lg \frac{\Gamma(a_{ik} + m_{ik})}{\Gamma(a_{ik})} \right] \\ m_{it^*} = \sum_{k=1}^{q_i} m_{ik} \\ a_{it^*} = \sum_{k=1}^{q_i} a_{ik} \end{array} \right. \quad (1)$$

式中: X_i 和 $\pi(X_i)$ 分别为第 i 个故障节点及该节点的父节点组合; D_1 为结构学习所使用的数据集; m_{ik} 为数据集 D_1 中满足 $X_i = k, \pi(X_i) = t$ 的样本个数; a_{ik} 为先验分布的超参数, 在实际使用中需要规定一个等价样本量 α 和先验贝叶斯网络 N_0 , 计算过程如式(2)所示。

$$a_{ik} = \alpha P_{N_0}(X_i = k | \pi(X_i) = t) \quad (2)$$

式中: P 为概率函数。

根据 CH 评分的可分解性, 可以通过对各家评分求和得到完整系统结构 G 的评分, 评分函数如式(3)所示。

$$CH(G | D_1) = \sum_{i=1}^n CH(\langle X_i, \pi(X_i) \rangle | D_1) \quad (3)$$

通过对数据集 D_1 的分析, 找到使得系统评分 $CH(G | D_1)$ 最高的结构 G , 该结构与数据集吻合程度最高, 被认为是最优贝叶斯结构^[13-14]。

2.2 融合算法的原理

先验信息在贝叶斯网络学习过程中有着重要的地位。传统数据驱动的学习方法由于缺乏先验信息, 存在工作效率低下的问题, 而利用多信号流程图模型可以自动获取系统先验信息作为专家知识, 进而建立先验贝叶斯网络, 提前去除大量不合理或无法存在的网络结构, 减小结构学习的搜索空间, 降低了学习的复杂度。同时, 由于数据样本来源的不确定性, 在样本数据不足或者可信度不够高的情况下, 引入多信号流程图知识来对贝叶斯网络的学习结果进行进一步评判, 有利于获得更加符合实际的网络结构。

虽然先验知识的引入通常可以提高学习效率, 但不准确的先验知识也会对结构学习的准确度产生较大的干扰^[15]。多信号流程图模型所包含的先验知识真实、准确, 以其作为专家知识引入建模过程可以提高算法的准确度和效率。本文实现了多信号流图中先验信息的自动提取, 同时结合 K2 算法构建了专家知识与数据分析相融合的结构学习算法, 实现了高性能的诊断贝叶斯网络自动建模过程。算法融合主要表现在 3 个方面:

1) 通过对多信号流程图模型进行解析和编号, 自动获取模型节点的拓扑排序和最大父节点个数, 为使用 K2 算法进行结构学习提供必要的先验知识。

2) 多信号流程图模型的 D -矩阵知识满足朴素贝叶斯模型建立条件^[7], 而朴素贝叶斯模型可以作为系统贝叶斯模型建立的基础。

D -矩阵为由故障变量和测试变量所构成的

0/1 矩阵^[16]:

$$D_{m \times n} = \begin{bmatrix} d_{11} & \cdots & d_{1n} \\ \vdots & & \vdots \\ d_{m1} & \cdots & d_{mn} \end{bmatrix} \quad (4)$$

式中: d_{ij} 为第 i 个故障变量在第 j 个测试变量上的反映, $d_{ij} = 1$ 表示相关, $d_{ij} = 0$ 表示不相关; $F_i = [d_{i1} d_{i2} \cdots d_{in}]$ 为第 i 个故障对象 F_i 与各个测试对象上的相关性。

由于多信号流程图和诊断贝叶斯网络的节点具有相似的节点含义, 前者的故障对象与测试对象可以直接映射到后者, 因此通过 D -矩阵生成的系统朴素贝叶斯模型包含了所有故障节点到测试节点之间的连接关系。现有的研究仅建立了朴素贝叶斯模型, 忽略了故障节点之间的依赖关系。因此, 本文在朴素贝叶斯模型基础上进行基于数据驱动的结构学习, 使得模型更加符合真实情况。同时相比于采用纯 K2 算法进行的结构学习, 该方法仅学习故障节点之间的连接关系, 可以减小学习规模、提升准确性及效率。

3) 对复杂系统进行分解可以提升贝叶斯结构的学习效率, 而多信号流程图知识为系统拆分提供了可能。针对系统规模较大、实现功能复杂的大型贝叶斯网络, 可以引入分解的思想, 即通过合理的划分, 将复杂系统网络结构学习转化为对多个子系统网络的学习, 最后对子网络进行合并得到完整贝叶斯网络结构。

考虑到多信号流程图连接线对诊断贝叶斯网络连接线所具有的包含关系, 在多信号流程图模型中, 若一组由多个故障节点组成的节点群与另一组节点群之间不存在信号的交联和影响关系, 就可以认为它们在贝叶斯网络中无连接, 进而对两组节点群分解, 分别进行结构学习。该方法减小了整个贝叶斯网络的结构学习规模, 同时也可以降低错误连接的风险。

此外, 由于多数结构学习方法都存在迭代计算及样本数据调度等问题, 导致并行计算难以实现。但是基于多信号流程图知识进行的子系统拆分具有独立性, 因此对每个子系统的学习过程可以采用并行计算方法。这一思路为提高学习效率提供了新的途径, 使得学习过程可以同时由多个计算资源共同处理, 显著提升学习效率^[17]。

值得注意的是, 在对多信号流程图模型进行解析时, 需要考虑环路对拓扑排序造成的干扰。为此, 本文算法需要单独对该情况进行拆环处理, 具体操作过程可以参见 2.3 节中的算法步骤。

2.3 算法步骤

基于 2.2 节的原理分析,本文所提出的诊断贝叶斯网络自动建模算法的步骤如下:

步骤 1 获取多信号流图模型的故障变量、测试变量以及 D -矩阵。

步骤 2 获取多信号流图所有变量的拓扑排序 ρ , 排序过程依据信号流的传递顺序进行深度优先搜索,保证信号流的输出端变量编号小于输入端变量对象。

步骤 3 如果多信号流存在环路,对其进行拆环后继续排序,假设拆除 X_i 指向 X_j 的连线 (X_i 为不同于 X_j 的任意故障节点),将其存入连接集合 l 。

步骤 4 遍历多信号流图模型中故障变量的父节点个数,取最大值作为贝叶斯模型的父节点上限 μ 。

步骤 5 以 D -矩阵为基础,建立故障变量到测试变量的朴素贝叶斯网络模型,作为诊断贝叶斯网络的初始结构。

步骤 6 当系统模型复杂,并且由多个局部子系统构成时,依据多信号流图连接关系,对系统进行分割。

步骤 7 已知故障变量排序 ρ 及父节点上限 μ ,读取数据集 D_i ,假设先验分布为均匀分布,根据式(3),即超参数 $a_{ijk}=1$,对故障变量进行 K2 学习,得到含有故障变量间依赖关系的贝叶斯结构网络。

步骤 8 检查 l 中的对象,对于 X_i 指向 X_j 的连线,如果 X_i 与 X_j 无连接,添加该连接进行评分计算,如果 X_i 与 X_j 存在直接或间接连接关系,遍历该环路所有连接关系并打分,最后保留评分较高的连接关系。

2.4 算法效率分析

本文的融合算法不仅有效利用了多信号流图提供的先验信息,解决了纯数据驱动算法无法获得有效先验信息的问题,同时还大大降低了网络结构学习算法的运算复杂度。从算法复杂度来看,K2 算法的时间复杂度为 $O(g h^2 p^2 q)$, g 为实例样本个数, h 为父节点上限个数, p 为节点个数, q 为每一个节点的取值个数^[18]。而本文提出的融合算法由于通过多信号流图生成 D -矩阵,能够迅速建立系统的初始朴素贝叶斯模型,因此测试对象的学习过程明显加快,参数 p 也由此减小。

另外,对于复杂系统,若基于多信号流图对

系统进行拆分,同样可以有效减小参数 p 。假设所有节点拆分成 2 个集合,每个集合的节点数分别为 p_1 和 p_2 ,那么此时的算法复杂度将变为 $O(g h^2 p_1^2 q) + O(g h^2 p_2^2 q)$,并且拆分后的部分子系统的父节点上限个数可能小于系统整体取值 h ,这又将进一步减小计算复杂度。

收敛速度和建模精度是评价算法好坏的两个关键指标。本文中收敛速度通过系统结构学习全过程所用的时间进行衡量,以读取多信号流图模型和数据作为开始,直到输出所有连接关系作为终止。贝叶斯结构学习的精确性能指标通常由以下 4 个因素衡量:缺失边 (Missing Edge, ME)、冗余边 (Redundant Edge, RE)、反向边 (Inverted Edge, IE)、有效边 (Correct Edge, CE),精确度 Accuracy 为

$$Accuracy = \frac{\sum CE}{\sum ME + \sum RE + \sum IE} \quad (5)$$

式中:分母部分又被称为汉明距指标^[19-20]。

3 仿真实验

3.1 对比实验

文献[12]中首次运用 K2 算法对 Alarm 网络进行学习,验证了该算法的可行性。本文同样采用 Alarm 网络进行结构学习,并根据网络信息流向对多信号流图模型进行假设,获得节点编号如图 1 所示(编号顺序不唯一),以及父节点上限个数 4。

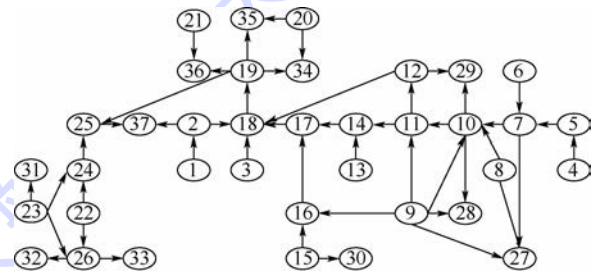


图 1 Alarm 网络结构

Fig.1 Structure of Alarm network

采用常用的网络结构学习算法 Path Condition(PC)算法、Necessary Path Condition(NPC)算法以及 K2 算法与本文提出的融合算法进行对比验证,学习所用数据样本根据 Alarm 网络参数由 Monte Carlo 法通过逻辑抽样生成^[21]。为减小样本数据对学习算法的影响,本文在样本量为 500, 1 000, 3 000, 7 000, 10 000 的情况下分别进行了 5 次仿真实验,表 1 所示为 5 次仿真结果的平均值。表中最后一列为本文提出的融合算法。

表 1 不同算法的学习结果对比

Table 1 Learning outcome comparison among different algorithms

数据规模	错误类型	PC算法	NPC算法	K2算法	融合算法
500	缺失边	11.2	14.0	2.2	1.2
	冗余边	0	0	6.4	4.0
	反向边	7.2	8.2	0	0
	有效边	27.6	23.8	37.4	40.8
1000	缺失边	9.0	10.8	2.0	1.0
	冗余边	1.2	7.8	4.4	3.2
	反向边	8.8	0.8	0	0
	有效边	27.0	26.6	39.6	41.8
3000	缺失边	3.4	4.4	1.0	0
	冗余边	0	0.2	2.0	0.6
	反向边	4.2	5.2	0	0
	有效边	38.4	36.2	43.0	45.4
7000	缺失边	1.8	2.2	1.0	0
	冗余边	0.4	0.4	1.4	0.4
	反向边	5.2	5.6	0	0
	有效边	38.6	37.8	43.6	45.6
10000	缺失边	1.4	1.6	1.0	0
	冗余边	0.2	0	1.2	0
	反向边	4.8	4.6	0	0
	有效边	39.6	39.8	43.8	46.0

假设模型最后一级叶节点是测试组件(即图1中 Alarm 网络的节点 27~37),它与故障节点的连接关系可以通过 D -矩阵获得,因此 Alarm 模型的规模由 37 个节点 46 条边缩小至 26 个节点 27 条边。根据式(5)中所提到的汉明距指标进行衡量,在数据规模为 10 000 时,不正确边数量下降到 0,说明数据规模的增加明显降低了建模中出现错误的可能性。

不同数据规模下的 4 种算法分别计算其精确度,结果如图 2 所示。由表 1 及图 2 可以看出,K2 算法的学习精确度要高于 PC 算法及 NPC 算法,而融合算法的学习精确度比上述 3 种学习算法明显更优。

以往 K2 算法实现中所运用的节点排序等知识需要通过专家指定,而融合算法首先解决了 K2 算法中对先验知识的苛刻限制。其次,融合算法

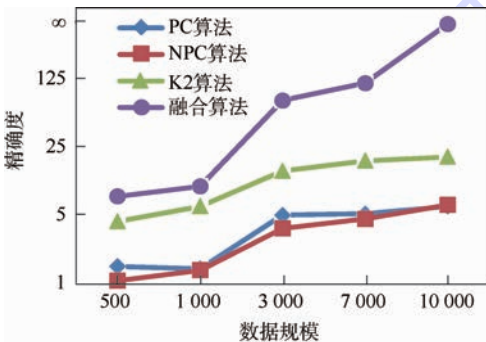


图 2 不同算法的精确度对比

Fig. 2 Precision comparison among different algorithms

利用系统 D -矩阵可以减小模型学习范围,显著提高了算法学习的精确度。例如,在数据规模为 7 000 时,可以计算得融合算法的精确度分别是 PC、NPC、K2 算法的 21.8、24.7、6.3 倍,当数据规模增大到 10 000 时,融合算法达到了 100% 正确的学习结果,此时根据式(5)求得的精确度指标将趋近于正无穷。最后,融合算法对于训练数据规模较小的情况依然有很好的表现。由图 2 可以看出,K2 算法在数据规模下降到 3 000 以下时,有效边的比例将低于总边数的 85%,而融合算法在数据规模仅为 500 的条件下,建模得到有效边的比例仍高于 K2 算法下数据规模为 3 000 的学习结果,这将极大地降低结构学习过程对数据规模的要求。

此外,纯数据驱动的算法虽有较好的建模精确度,但是收敛速度过慢,尤其是当系统规模和学习数据增大时,算法时间复杂度迅速上升。在对比实验中,如图 3 所示,融合算法相比于 K2 算法虽然在节点规模上仅减少了不到 1/3,但是其收敛时间减小了大约 60%。

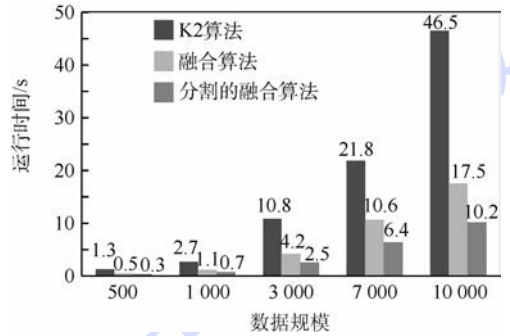


图 3 算法收敛时间对比

Fig. 3 Convergence time comparison among different algorithms

为验证多信号流图对模型的拆分和并行计算具有的潜在贡献,观察图 1 所示 Alarm 系统,根据其信号流的传递关系,可以在节点 18 和节点 19 之间对网络系统进行拆分,此时学习对象分别为节点 1~18 和节点 18~26 组成的网络。由图 3 看出,分割后的融合算法的收敛速度减小为最初的 22%~29%。若该计算过程可以实现并行计算,收敛速度将进一步降低。该算法特性弥补了 K2 算法效率低下的缺点。

3.2 仿真实例

本文所使用系统是某型观瞄系统,根据设计人员所提供的产品信号交联关系及故障模式分析结果可搭建多信号流图模型如图 4 所示,多信号流图模型完全遵循系统物理连接关系,相比于利用专家知识直接建立贝叶斯模型更加容易实现,

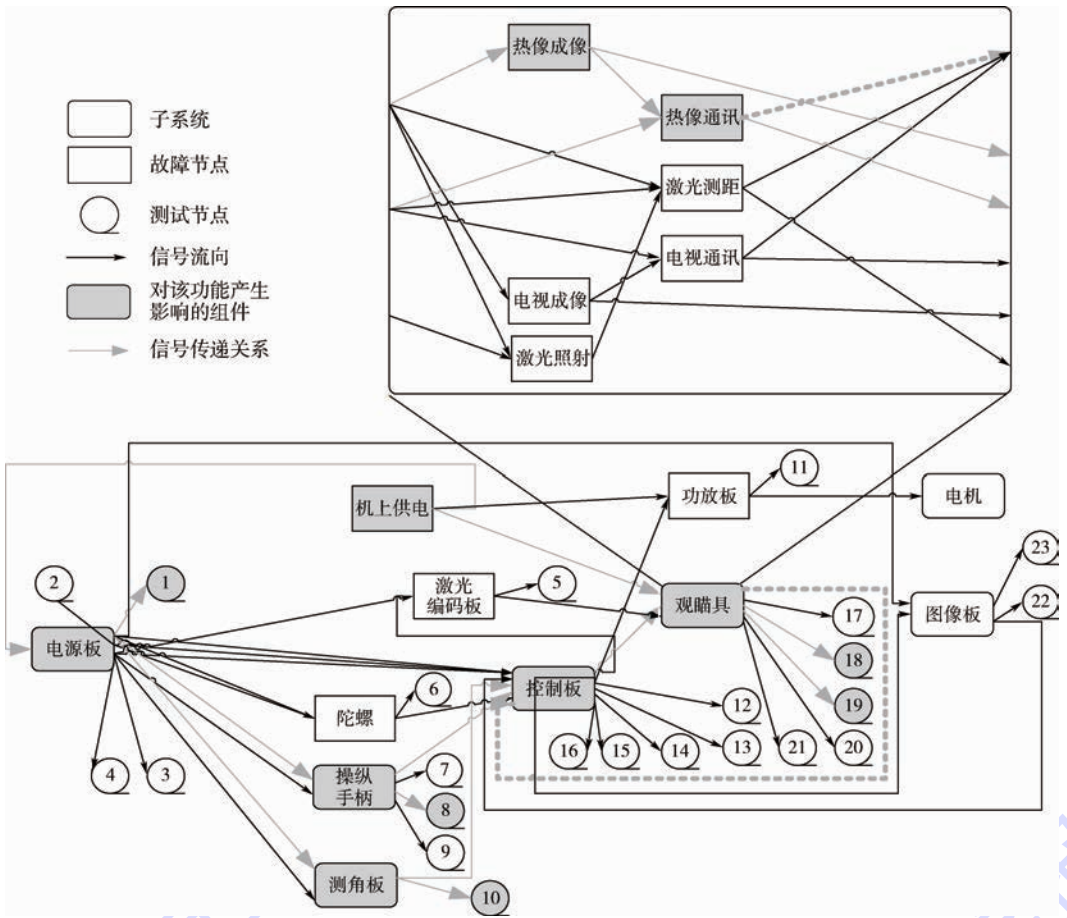


图 4 某型观瞄系统的多信号流图模型

Fig. 4 Multi-signal flow graphs model of an aiming system

并且模型的准确性可以得到保证。图 4 中子系统内部由故障组件构成,如观瞄具的内部结构为图中右上角方框内所示。基于 2.1 节对多信号流图模型可分解性的讨论,本文选取系统主要功能之一的热像观瞄功能进行建模演示,图中控制板与观瞄具间的信号传递产生环路,因此暂不考虑灰色点划线所代表的信号传递关系,以达到拆环的目的。

首先对热像观瞄子系统进行编码,沿灰色实连接线方向由小到大对组件编号,测试组件编码在故障组件全部编码完成后进行,编码结果为: C1 ~ C7 依次为故障组件 {“机上电源”,“电源板”,“测角板”,“操纵手柄”,“控制板”,“热像成像”,“热像通讯”}, C8 ~ C12 依次为测试组件 {“10”,“19”,“18”,“8”,“1”}。其次,寻找输入端最大的输入量作为父节点上限,图 4 中该值出现在控制板输入端及热像通讯输入端,取值为 2。 D -矩阵可以在多信号流图测试性分析过程中获得,矩阵关系如表 2 所示。利用以上信息及系统完整数据,可以进行贝叶斯网络学习。学习步骤如下:

步骤 1 多信号流图测试组件与诊断贝叶斯网络测点一一对应,因此根据表 2 D -矩阵获得系统朴素贝叶斯模型如图 5(a) 所示。

步骤 2 对 C1 ~ C7 节点通过 K2 算法进行学习,得到贝叶斯结构如图 5(b) 所示。

步骤 3 考虑图 1 中的环形模型,灰色点划线信号流由 C7 指向 C5,此时在贝叶斯网络中添加 C7→C5 的有向边不会造成环路连接,因此添加该边并对系统进行打分,评分结果显示新的连接关系优于原连接,因此添加该有向边,获得模型如图 5(d) 所示。

表 2 子系统 D -矩阵

Table 2 D -matrix of a subsystem

测试组件	测试组件 10	测试组件 19	测试组件 18	测试组件 8	测试组件 1
机上电源	0	1	1	0	1
电源板	0	0	0	0	1
测角板	1	0	0	0	0
操纵手柄	0	0	0	1	0
控制板	0	1	0	0	0
热像成像	0	1	1	0	0
热像通讯	0	1	0	0	0

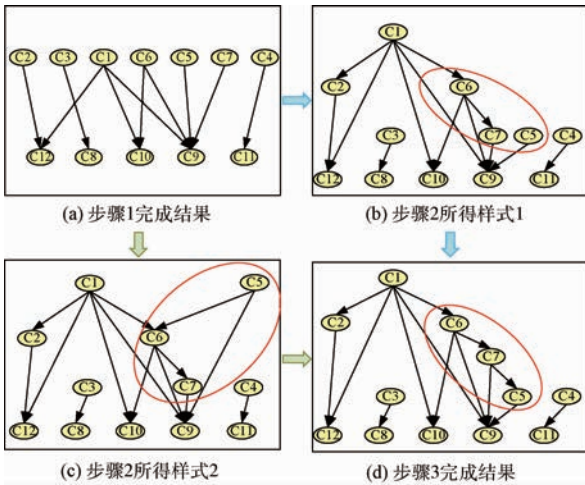


图5 子系统贝叶斯模型的建立过程

Fig. 5 Building process of Bayesian model of a subsystem

受到数据和学习精确度的影响,步骤2中贝叶斯网络学习结构可能如图5(c)所示,此时若添加 $C7 \rightarrow C5$ 有向边会造成网络环路,因此需对 $C5$ 、 $C6$ 和 $C7$ 这3个节点的每种连接关系进行遍历打分,获得分数最高的连接关系依然如图5(d)所示,该网络模型唯一。

在实际工程中,多信号流图中故障节点间的信号传递关系仅有少部分具有因果联系,即在贝叶斯网络中存在连接关系,因此,多信号流图中的环路通常不会造成图5(c)中贝叶斯网络结构矛盾现象,环路越大在贝叶斯网络学习过程中出现矛盾结构的概率就越小。而对于小规模环路而言,遍历该环路的每种连接关系是可以被接受的,同时该遍历过程有机会纠正因编号或数据等问题造成的建模偏差。

通过对系统各个功能区域进行分解,可以得到多个小规模下的贝叶斯网络,由于各个节点编号唯一,因此最终总会得到完整的系统诊断贝叶斯结构。分解建模的过程减小了结构学习的规模和错误率,并且为进一步实现贝叶斯结构的并行学习提供了可能。

4 结 论

本文提出了多信号流图模型与K2算法融合的诊断贝叶斯网络自动建模方法,经实验验证具有以下特点:

- 1) 对已有多信号流图模型进行知识提取,可以获取K2算法所需要的先验信息。
- 2) 充分利用多信号流图中节点的连接关系对模型进行简化和分解,显著改善了纯数据驱动学习方法存在的搜索空间大、运行效率低、建模精

度不够理想等问题。

3) 得益于多信号流图在专家知识表述上的准确性、全面性以及K2算法的优异性能,融合算法可以显著提升贝叶斯网络结构学习的精度和效率,同时也为复杂系统的拆分和并行学习提供了有效途径。

参考文献 (References)

- [1] JENSEN F V. Bayesian networks and decision graphs [M]. New York: Springer, 2001: 32-34.
- [2] EFRON B. Bayes' theorem in the 21st century [J]. Science, 2013, 340(6137): 1177-1178.
- [3] AOUAY S, JAMOUISSI S, BEN AYED Y. Particle swarm optimization based method for Bayesian network structure learning [C] // 2013 5th International Conference on Modeling, Simulation and Applied Optimization (ICMSAO). Piscataway, NJ: IEEE Press, 2013: 1-6.
- [4] 张振海, 王晓明, 党建武, 等. 基于专家知识融合的贝叶斯网络结构学习方法 [J]. 计算机工程与应用, 2014, 50(2): 1-4. ZHANG Z H, WANG X M, DANG J W, et al. Bayesian network structure learning method based on expert knowledge fusion [J]. Computer Engineering and Applications, 2014, 50(2): 1-4 (in Chinese).
- [5] BALABAN E, SAXENA A, NARASIMHAN S, et al. Prognostic health-management system development for electromechanical actuators [J]. Journal of Aerospace Information Systems, 2015, 12(3): 329-344.
- [6] 代京, 于劲松, 张平, 等. 基于多信号流图的诊断贝叶斯网络建模 [J]. 北京航空航天大学学报, 2009, 35(4): 472-475. DAI J, YU J S, ZHANG P, et al. Diagnostic Bayesian networks modeling based on multi-signal flow graphs [J]. Journal of Beijing University of Aeronautics and Astronautics, 2009, 35(4): 472-475 (in Chinese).
- [7] 陈希祥, 邱静, 刘冠军. 测试不确定条件下基于贝叶斯网络的装备测试优化选择技术 [J]. 中国机械工程, 2011, 22(4): 379-384. CHEN X X, QIU J, LIU G J. Test selection of materiel based on Bayesian network under test uncertainty [J]. China Mechanical Engineering, 2011, 22(4): 379-384 (in Chinese).
- [8] Qualtech Systems Inc. Testability, engineering and maintenance system (TEAMS) [EB/OL]. [2010-07-26]. <http://www.teamqi.com/products/teams-designer/>.
- [9] 吴红, 王维平, 杨峰. 融合先验信息的贝叶斯网络结构学习方法 [J]. 系统工程与电子技术, 2012, 34(12): 2585-2591. WU H, WANG W P, YANG F. Structure learning method of Bayesian network with prior information [J]. Systems Engineering and Electronics, 2012, 34(12): 2585-2591 (in Chinese).
- [10] MASEGOSA A R, MORAL S. New skeleton-based approaches for Bayesian structure learning of Bayesian networks [J]. Applied Soft Computing, 2013, 13(2): 1110-1120.
- [11] BOUCHAALA L, MASMOUDI A, GARGOURI F, et al. Improving algorithms for structure learning in Bayesian networks using a new implicit score [J]. Expert Systems with Applica-

tions, 2010, 37(7):5470-5475.

- [12] COOPER G F, HERSKOVITS E. A Bayesian method for the induction of probabilistic networks from data [J]. Machine Learning, 1992, 9(4):309-347.
- [13] CHENG J, GREINER R, KELLY J, et al. Learning Bayesian networks from data: An information-theory based approach[J]. Artificial Intelligence, 2002, 137(12):43-90.
- [14] SILANDER T, ROOS T, MYLLYMÄKI P. Learning locally min-max optimal Bayesian networks [J]. International Journal of Approximate Reasoning, 2010, 51(5):544-557.
- [15] WELLMAN M P, BREESE J S, GOLDMAN R P. From knowledge bases to decision models[J]. The Knowledge Engineering Review, 1992, 7(1):35-53.
- [16] DEB S, PATTIPATI K R, SHRESTHA R. QSI's integrated diagnostics toolset [C] // 1997 IEEE Autotestcon Proceedings AUTOTESTCON, 97. Piscataway, NJ: IEEE Press, 1997:408-421.
- [17] YU K, WANG H, WU X. A parallel algorithm for learning Bayesian networks[M] // Advances in Knowledge Discovery and Data Mining. Berlin Heidelberg: Springer, 2007:1055-1063.
- [18] 曾安, 李晓兵, 杨海东, 等. 基于最小描述长度和 K2 的贝叶斯网络结构学习算法 [J]. 东北师大学报(自然科学版), 2014, 46(3):53-58.
- ZENG A, LI X B, YANG H D, et al. Bayesian network structure learning based on minimum description length and K2 algorithm [J]. Journal of Northeast Normal University (Natural Science

Edition), 2014, 46(3):53-58 (in Chinese).

- [19] 吴永广, 庞世春. K2 & HC 结构学习算法 [J]. 计算机与数字工程, 2014, 42(7):1137-1140.
- WU Y G, PANG S C. K2 & HC structure learning algorithm [J]. Computer & Digital Engineering, 2014, 42(7):1137-1140 (in Chinese).
- [20] TIAN J, PEARL J. Causal discovery from changes [C] // Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence. San Francisco: Morgan Kaufmann Publishers Inc., 2001:512-521.
- [21] TSAMARDINOS I, BROWN L E, ALIFERIS C F. The max-min hill-climbing Bayesian network structure learning algorithm [J]. Machine Learning, 2006, 65(1):31-78.

作者简介:

沈琳 女, 硕士研究生。主要研究方向: 故障预测与健康管理技术。

Tel.: 010-82338693

E-mail: shenlin9177@163.com

于劲松 男, 副教授, 硕士生导师。主要研究方向: 预测与健康管理技术、自动测试系统。

Tel.: 010-82338693

E-mail: yujs@buaa.edu.cn

Automatic learning of Bayesian network structure using graph model and learning algorithm

SHEN Lin¹, YU Jinsong^{1,2,*}, TANG Diyin¹, LIU Hao^{1,3}

(1. School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;

2. Collaborative Innovation Center of Advanced Aero-Engine, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;

3. Unit 93, Army 95809 of PLA, Cangzhou 061736, China)

Abstract: In order to improve the accuracy and efficiency of the data-driven approaches in learning Bayesian network structure, expert knowledge is usually implemented in the learning algorithm. To deal with the lack of effective ways to combine the expert knowledge and the data-driven learning approaches in the existing methods, this paper proposes an automatic learning method for Bayesian network structure learning, which combines multi-signal flow graphs and learning algorithm K2. The method inserts expert knowledge into data-driven learning methods, using the information of relationships between signals from multi-signal flow graphs and the structure learning algorithm K2, to achieve automatic learning of Bayesian network structure. Numerical analysis, compared with other typical network structure learning algorithms, proves that the proposed method significantly lowers the structure learning requirements for learning scale and training data size and provides a higher learning accuracy and computation efficiency. The application of the proposed method is illustrated using a real engineering system and verified the practicability of the algorithm at the same time.

Key words: Bayesian networks; structure learning; multi-signal flow graphs; K2 algorithm; fault diagnosis

Received: 2015-07-02; Accepted: 2015-09-30; Published online: 2015-10-30 17:00

URL: www.cnki.net/kcms/detail/11.2625.V.20151030.1700.010.html

* Corresponding author. Tel.: 010-82338693 E-mail: yujs@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI:10.13700/j.bh.1001-5965.2015.0435



多约束下备件多层多级库存优化模型

蔡芝明^{1,2}, 金家善^{1,2,*}, 陈砚桥^{1,2}

(1. 海军工程大学 舰船动力工程军队重点实验室, 武汉 430033; 2. 海军工程大学 动力工程学院, 武汉 430033)

摘 要: 备件库存优化是提高装备保障效率和战备完好性的有效途径之一, 针对目前部队现有多约束多层多级备件供应模式及工程实践要求的问题, 以舰艇编队索马里护航任务准备阶段备件携带方案优化为研究背景, 以舰艇保障经费、载荷和仓库体积为约束条件, 以舰艇编队期望短缺数最小和可用度最大为目标函数, 运用拉格朗日乘子法和边际效应法求出舰艇编队最优多层多级备件携带方案。通过实例将优化前后基地和后方仓库备件总的质量、体积及费用与对应约束指标上限进行对比分析, 结果表明: 本文模型和方法可以有效解决编队备件携带问题, 同时也可为其他领域解决此类问题提供借鉴。

关键词: 备件; 多层多级; 多约束; 库存优化; 舰艇编队

中图分类号: TJ83; U662.1

文献标识码: A

文章编号: 1001-5965(2016)07-1494-08

备件库存管理是后勤保障中最关键的问题之一, 因为装备系统可用性关系到武器装备的战略实施及库存成本的使用。即使有大量多种多样的备件增加库存成本, 但当出现任何一个备件短缺时对系统的可用性都有非常大的影响, 故采用合理、科学的备件携带方案就显得至关重要。

自 Sherbrooke^[1-3] 提出经典的备件多层级库存理论以来, 国内外学者在此基础上对其进行了大量扩展和模型改进。Levner^[4] 和 Samouei^[5] 等主要采用网络维修流程图的方法, 以最小总库存和运输成本为目标, 同时考虑各保障站点之间的依赖性和相互作用, 对备件多层级库存问题进行了分析; Caglar 等^[6] 建立了以响应时间为约束, 以库存费用最小为目标的价格昂贵备件多级库存模型; Costantino 等^[7] 以备件短缺数最小和系统可用度最大为限制条件, 依据不同维修站点具备不同维修能力, 运用边际效应法对备件多层级库存进行了优化; Topan 和 Bayindir^[8] 采用 (NQ, R) 库存

策略, 对服从泊松分布的 2 级备件携带方案进行了优化; Dada^[9] 以库存费用最小为目标, 建立了低需求率且昂贵备件的 2 级库存模型; Sleptchenko 等^[10] 改进了 VARI-METRIC 模型中无限维修能力的假设条件; Al Hanbali 和 van der Heijden^[11] 采用对系统可用度分区间的形式, 对 2 级备件库存进行了优化, 并验证了算法的精确性; Lau 等^[12] 提出了可用度随时间变化的多级库存模型; Topan 等^[13] 以平均供应时间为约束, 以最小化库存为目标, 采用 (Q, R) 策略对多个备件 2 级库存配制进行了建模分析; Basten^[14] 和樊九九^[15] 等将故障件修理级别分析引入到多级库存模型中, 在降低总库存成本方面取得了显著效果; 王睿等^[16] 以战时舰艇编队任务成功性为约束条件, 以保障费用最小为目标, 运用边际效应法求出 3 级库存保障方案; 孙蕾和左洪福^[17] 以费用为约束条件, 以系统可用度最大为目标, 对航空类备件多级库存进行了优化; 刘任洋等^[18] 以可用度为约束条

收稿日期: 2015-06-30; 录用日期: 2015-08-26; 网络出版时间: 2015-10-19 15:20

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151019.1520.009.html

基金项目: 国防预研基金(51319060103); 中国博士后科学基金(2013T60921)

* 通讯作者: Tel.: 027-65460962 E-mail: jinjiashan401@163.com

引用格式: 蔡芝明, 金家善, 陈砚桥. 多约束下备件多层多级库存优化模型[J]. 北京航空航天大学学报, 2016, 42(7): 1494-1501.
CAI Z M, JIN J S, CHEN Y Q. Optimal inventory modeling of spare parts multi-indenture multi-echelon under multi-constraints [J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1494-1501 (in Chinese).

件,以费用最小为目标,建立了基于横向转运的可修件3级库存优化模型。

可以看出,尽管目前对备件多级库存优化问题的研究非常深入,但在具体工程实践中都存在以下问题:①大部分研究的约束条件都为1个^[4-18],当约束条件增加为2个或以上时,就无法再使用已有的模型和方法求出备件多层级最优库存携带方案。②目前研究多是多层或多级备件库存问题,而对工程实践中出现的多个约束条件下备件多层和多级联合的最优库存问题,截止目前很少有研究提及。③目前多级库存研究大部分从商业供应链的角度,以费用最小为评价指标^[4-6, 9, 13, 16],而部队工程实践中,使用方多以系统可用度为装备效能评估指标,评价指标体系范围有待扩充。④本文产生的工程背景是舰艇索马里护航备件携带利用率非常低的问题(利用率大概为20%左右^[19]),舰艇编队备件携带要受到舰艇载荷、保障经费及仓库体积等多个约束因素影响,已有模型主要针对陆地装备,以降低保障经费^[4-6, 9, 13-18]为目标,并未考虑海军部队中常见的载荷和库存体积方面约束,无法满足海军部队工程实践的需要。⑤备件多层多级混合库存问题是研究的难点和热点。

因此,本文以舰艇保障经费、载荷和仓库体积为约束条件,以舰艇编队期望短缺数最小和可用度最大为目标函数,以舰艇编队索马里护航备件携带为研究背景,建立了相应库存模型,并给出了模型求解步骤和方法,通过案例分析验证了模型可为解决此类问题提供有效借鉴。

1 模型描述

本文产生的工程背景是舰艇编队索马里护航备件携带问题,文献[19]指出编队携带备件大概有80%被原封不动的带回,这种低备件利用率必然引起使用方的重视,故舰艇编队在出海之前,对其携带备件方案进行优化是非常有必要的,且具有重大的经济意义。

编队组成一般为1艘伴随综合保障舰和2艘最新型护卫舰或驱逐舰^[20]。伴随综合保障舰具有中继级的维修能力,相当于1个海上“移动式”后方仓库,因新型驱逐舰或护卫舰配备了专门的修理中队,相当于1个保障基地,这样整个编队便组成了1个后方仓库和2个基地的2级器材保障模式。

新型舰船由很多大型且复杂装备组成,其备件携带方案必然要分2个层级来综合权衡制定,这样才能更好满足工程实践的需要。本文将备件

分为外场可更换单元(Line-Replaceable Unit, LRU)和内场可更换单元(Shop-Replaceable Unit, SRU)。因而编队组成了2层2级的保障模式,由于舰艇备件携带受舰艇载荷及仓库体积的限制,在考虑备件费用约束的同时,还必须考虑备件质量和备件体积约束。为了简化模型求解过程和步骤,对模型的影响条件进行了一些合理的假设:

- 1) 后方仓库或基地内部不进行串件拚修。
- 2) 后方仓库或基地间不进行横向供应。
- 3) 故障件修理工作相对独立,不相互影响。
- 4) 现场可对故障件进行拆卸和定位。
- 5) 维修渠道无限,不存在备件排队等待问题。
- 6) 补给舰和舰艇之间距离很近,一般只有几链,故可以忽略伴随补给舰到舰艇的备件供应时间。
- 7) 假设所有备件重要性相同。
- 8) 所有的SRU都可以在后方仓库进行修理。

2 需求率的确定

2.1 备件需求率模型

不考虑故障间隔时间、误报警率等维修参数的影响,依据备件 k 单装备配置数量 Z_k 、装备在现场配置数量 N_j 、故障间隔时间 $MTBF_k$ 及执行任务周期 T_k 等参数可以得出故障现场需求率为

$$m_k = \frac{T_k N_j Z_k}{MTBF_k} \quad (1)$$

式中: j 为基地编号, $j=1,2,\dots,J$ (0 =后方仓库); k 为LRU的项目编号, $k=1,2,\dots,K$ 。

依据 m_k 可以求出后方仓库或基地的LRU和SRU的需求率。

2.2 各级站点需求率的求解方法

多层多级备件需求率求解过程始于LRU失效并送到基地供应站。如果基地供应站有1件LRU备件,就进行更换,否则基地就发生1次LRU短缺。LRU故障件在基地修理有一定概率:如果LRU设备太复杂,往往就会被送到后方仓库进行修理,同时基地会安排向后方仓库申请该LRU 1件。

若LRU在基地进行修理,则设想将发现1件且仅为1件SRU发生故障。若基地现有库存有1件SRU备件就进行更换,将其安装到LRU,就完成了LRU的修理。基地能否修理SRU有一定概率:在基地修理能力范围之内,就在基地进行修理,反之,则发往后方仓库进行修理,同时向后方

仓库提出该 SRU 的申请。

当完成了 1 件 LRU 的送修和补给时,即使还存在任何未解决的问题,但 1 次备件满足得到了供应;否则,LRU 的现有库存就增加 1 件。若 LRU 未在基地进行修理,在后方仓库就发生 SRU 送修的类似过程。

根据基地 LRU 需求率,再结合给定备件各种参数(如修理概率等),可以计算出基地 j 相应 SRU 任务期内的需求量及后方仓库 LRU 和 SRU 的需求量,具体备件需求和供应渠道流程如图 1 所示。

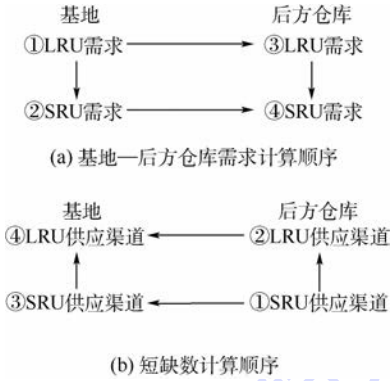


图 1 备件需求及供应渠道流程

Fig. 1 Flowsheet of spares demand and supply channels

基地 j 的 SRU_i 任务周期内需求量为

$$m_{ijk} = m_{jk} r_{jk} q_{ijk} \quad (2)$$

式中: m_{ijk} 为基地 j 的 SRU_i 任务周期内需求量; m_{jk} 为基地 j 的 LRU_k 任务周期内的需求量; r_{jk} 为基地 j 的 LRU_k 在本地修理的概率; q_{ijk} 为 LRU_k 修理产生 SRU_i 需求的概率。

后方仓库 LRU_k 任务周期内的需求量为各基地发生的对后方仓库补给申请的 LRU_k 需求之和:

$$m_{0k} = \sum_{j=1}^J m_{jk} (1 - r_{jk}) \quad (3)$$

式中: m_{0k} 为后方仓库 LRU_k 任务周期内的需求量。

后方仓库 SRU_i 任务周期内的需求量为所有基地补给需求量之和加上后方级修理对应 LRU 所需的 SRU 件数(在 MOD-METRIC 标准模型中忽略了 SRU 的来源),即

$$m_{0ik} = \sum_{j=1}^J m_{ijk} (1 - r_{ijk}) + m_{0k} q_{0ik} \quad (4)$$

式中: r_{ijk} 为基地 j 的 SRU_i 在本基地维修的概率; q_{0ik} 为 LRU_k 在中心仓库修理产生 SRU_i 需求的概率; m_{0ik} 为后方仓库 SRU_i 任务周期需求量。

图 1(b) 为供应渠道及短缺数顺序需要计算基地的期望短缺数。这点正是计算需求量的反向顺序。依据单个使用现场装备可用度与备件短缺

数的关系,可以得出单个装备的备件可用度。本文重点是依据本节给的 1 组备件(1 件 LRU 及其所属第 2 层级的若干 SRU)在基地和后方仓库的需求量计算,相应地可以计算出 K 组备件在基地和后方仓库的需求率。

3 备件库存优化模型

3.1 装备效能评估指标的确定

装备效能评估指标是衡量备件保障方案配置好坏程度的重要参数,常用的指标有可用度和备件期望短缺数等。

备件短缺数是指某一时间内装备不能满足备件需求的个数。若发生一次备件需求不能被满足,则记为备件短缺一次。备件短缺数定义如下:

$$B(X|S) = \begin{cases} X - S & X > S \\ 0 & X \leq S \end{cases} \quad (5)$$

式中: S 为备件库存量; X 为备件需求量; B 为备件短缺函数。

备件短缺数的确定需要通过备件存储量 S 和补给件或需要修理件 D_1 的稳态概率分布来确定。若 D_1 小于库存 S 时,表示库存 S 能够满足备件需求,不会产生备件短缺;只要 D_1 需求超过 S 时,库存无法满足备件需求,就会出现备件短缺。假设某一随机时刻待收库存数为 $(S + L)$ 件时,就会出现 L 件短缺,此时备件期望短缺数 $E_{B_0(S)} = \sum_{x=S+1}^{\infty} (x - S) \Pr\{x\}$ 。

后方仓库或基地库存 S 、后方仓库或基地现有库存量 O_H 、备件短缺数 B_0 及待收件数 D_1 满足如下关系:

$$S = O_H + D_1 - B_0 \quad (6)$$

在备件模型中,备件期望短缺数为 $E_{B_0(S)} = \sum_{x=1}^{\infty} x \Pr\{X=x\}$ 。备件期望短缺数是一个非负的量,当 $S=0$ 时,其值等于某一分布的均值,因此 $E_{B_0(0)} = E(x)$ 。

此时,还需要一个衡量所有随机 x 偏离 $E(x)$ 程度的变量,这一衡量与均值偏离程度的变量称为方差,即 $\text{Var}(x) = E(x^2) - (E(x))^2$, $E(x^2)$ 为 x^2 的期望值,也可以称为 x 的 2 阶样本距。

$\Pr(x)$ 属于何种分布,可以用方差除以均值得到差均比 V_{TMR} 来确定。当 $V_{\text{TMR}} = 1$ 时,分布函数 $\Pr(x)$ 服从泊松分布;当 $0 < V_{\text{TMR}} < 1$ 时,分布函数 $\Pr(x)$ 服从二项分布;当 $V_{\text{TMR}} > 1$ 时,分布函数 $\Pr(x)$ 服从负二项分布^[21]。 V_{TMR} 的定义为 $V_{\text{TMR}} = \text{Var}(x)/E(x)$ (7)

可用度的数学定义是任一时刻装备能够正常工作的程度,装备可用度求解公式为

$$A_j = 100 \prod_{k=1}^K [1 - E_{BO_k(s_k)} / (N_j Z_k)]^{Z_k} \quad (8)$$

式(8)求单个基地可用度最大值,对式(8)两边取对数:

$$\begin{aligned} \ln(A_j/100) &= \sum_{k=1}^K Z_k [1 - E_{BO_k(s_k)} / (N_j Z_k)] \approx \\ &- \sum_{k=1}^K E_{BO_k(s_k)} / N_j \end{aligned} \quad (9)$$

后方仓库和基地组成的整个系统的平均可用度为

$$\bar{A} = \frac{\sum_{j=0}^J (N_j A_j)}{\sum_{j=0}^J N_j} \quad (10)$$

式中: \bar{A} 为系统平均可用度。

3.2 目标函数和约束条件

本文将编队期望短缺数最小和可用度最大作为目标函数,舰艇有效载荷、保障经费及仓库体积作为约束条件。

后方仓库所有备件体积之和必须小于或等于后方仓库体积上限,即

$$\sum_{i=1}^I s_{0i} v_{0i} + \sum_{k=1}^K s_{0k} v_{0k} \leq V_{0T} \quad (11)$$

式中: s_{0i} 为后方仓库中备件SRU_i的存储数量; v_{0i} 为后方仓库中备件SRU_i的体积; s_{0k} 为后方仓库中备件LRU_k的存储数量; v_{0k} 为为后方仓库中备件LRU_k的体积; V_{0T} 为后方仓库体积指标的上限值。

基地j所有备件体积之和必须小于或等于基地仓库体积上限,即

$$\sum_{i=1}^I s_{ij} v_{ij} + \sum_{k=1}^K s_{jk} v_{jk} \leq V_{jT} \quad (12)$$

式中: s_{ij} 为基地j仓库中备件SRU_i的存储数量; v_{ij} 为基地j仓库中备件SRU_i的体积; s_{jk} 为基地j仓库中备件LRU_k的存储数量; v_{jk} 为基地j仓库中备件LRU_k的体积; V_{jT} 为基地j仓库体积指标的上限值。

后方仓库所有备件质量之和必须小于或等于后方仓库质量指标上限,即

$$\sum_{i=1}^I s_{0i} \bar{m}_{0i} + \sum_{k=1}^K s_{0k} \bar{m}_{0k} \leq M_{0T} \quad (13)$$

式中: \bar{m}_{0i} 为后方仓库中备件SRU_i的质量; \bar{m}_{0k} 为后方仓库中备件LRU_k的质量; M_{0T} 为后方仓库质量指标的上限值。

基地j所有备件质量之和必须小于或等于基地仓库质量指标上限,即

$$\sum_{i=1}^I s_{ij} \bar{m}_{ij} + \sum_{k=1}^K s_{jk} \bar{m}_{jk} \leq M_{jT} \quad (14)$$

式中: \bar{m}_{ij} 为基地j仓库中备件SRU_i的质量; \bar{m}_{jk} 为基地j仓库中备件LRU_k的质量; M_{jT} 为基地j仓库质量指标的上限值。

后方仓库所有备件费用之和必须小于或等于后方仓库费用指标上限,即

$$\sum_{i=1}^I s_{0i} c_{0i} + \sum_{k=1}^K s_{0k} c_{0k} \leq C_{0T} \quad (15)$$

式中: c_{0i} 为后方仓库中备件SRU_i的费用; c_{0k} 为后方仓库中备件LRU_k的费用; C_{0T} 为后方仓库保障经费总指标的上限值。

基地j所有备件质量之和必须小于或等于基地仓库质量指标上限,即

$$\sum_{i=1}^I s_{ij} c_{ij} + \sum_{k=1}^K s_{jk} c_{jk} \leq C_{jT} \quad (16)$$

式中: c_{ij} 为基地j仓库中备件SRU_i的费用; c_{jk} 为基地j仓库中备件LRU_k的费用; C_{jT} 为基地j保障经费总指标的上限值。

依据式(11)~式(16),可以得出本文建立的模型为

$$\begin{cases} \max p_s, \\ \min E_{BO(s)} \\ \sum_{i=1}^I s_{0i} v_{0i} + \sum_{k=1}^K s_{0k} v_{0k} \leq V_{0T} \\ \sum_{i=1}^I s_{ij} v_{ij} + \sum_{k=1}^K s_{jk} v_{jk} \leq V_{jT} \\ \sum_{i=1}^I s_{0i} \bar{m}_{0i} + \sum_{k=1}^K s_{0k} \bar{m}_{0k} \leq M_{0T} \\ \sum_{i=1}^I s_{ij} \bar{m}_{ij} + \sum_{k=1}^K s_{jk} \bar{m}_{jk} \leq M_{jT} \\ \sum_{i=1}^I s_{0i} c_{0i} + \sum_{k=1}^K s_{0k} c_{0k} \leq C_{0T} \\ \sum_{i=1}^I s_{ij} c_{ij} + \sum_{k=1}^K s_{jk} c_{jk} \leq C_{jT} \end{cases} \quad (17)$$

主要采用拉格朗日乘子法和边际效应法相结合对模型进行求解和优化。

3.3 库存优化方法

编队在出海执行任务准备阶段,备件携带受到后勤保障经费、载荷及仓库体积的限制,不可能不限品种和数量的携带舰船所需的所有备件,必须要对备件携带方案进行优化,达到在满足所有约束指标的基础上,使编队可用度最大及期望短缺数最小。

本文库存优化方法主要思想是用拉格朗日乘子^[22]将影响库存配置的3个因素转换为拉格朗日乘子,然后采用边际效应分析^[3]求出库存最优

分配方案。一方面,拉格朗日乘子法能将影响库存配置方案的3个因素转换为1个拉格朗日乘子,这为使用边际效应分析奠定了基础;另一方面,边际效应分析具备求解速度快、操作方便及结论精度高的特点,是库存优化中常见的方法。方法开始时,假设所有层级站点备件初始库存量都为0,在每次迭代计算过程中,通过比较边际效应值的大小,确定备件携带种类和数量。模型具体求解步骤如下:

步骤1 由于编队备件携带涉及的影响因素非常多,此种情况下,边际效应法无法使用,引入拉格朗日乘子法,将所有影响因素转换为拉格朗日乘子,即

$$L_s = \lambda_m \bar{m}_s + \lambda_v v_s + \lambda_c c_s \quad (18)$$

式中: λ_m 为备件质量因子; λ_c 为备件购置费用因子; λ_v 为备件体积因子; \bar{m}_s 为备件质量; c_s 为备件购置费用; v_s 为备件体积。

步骤2 只考虑一种约束因素,运用边际效应法求出此时备件携带方案。

$$S_{00} = (s_1, s_2, \dots, s_T) \quad (19)$$

步骤3 据此求出后方仓库或基地总备件质量、体积及费用,从而可求出质量因子、体积因子和费用因子。以费用为例进行说明。

$$\begin{cases} \lambda_c = C_{(s_0)} / C_{(s_0)} = 1 \\ \lambda_v = C_{(s_0)} / V_{(s_0)} \\ \lambda_m = C_{(s_0)} / M_{(s_0)} \end{cases} \quad (20)$$

依据式(20)便可求出相应的拉格朗日乘子。

步骤4 依据式(14)~式(18)构建后方仓库或基地的LRU和SRU拉格朗日乘子组合,即

$$L_T = (L_{s1}, L_{s2}, \dots, L_{sT}) \quad (21)$$

步骤5 重新构建 L_s 为自变量, $\text{Pr}(L_s)$ 为因变量的函数,并做1阶差分,此时边际增量为

$$\Delta \text{Pr}(L_s) = \text{Pr}(L_s + 1) - \text{Pr}(L_s) \quad (22)$$

步骤6 依据式(21)和式(22),计算得到边际效应值,即

$$\Phi = \frac{\Delta \text{Pr}(L_s)}{\Delta L_s(s)} = \frac{\text{Pr}(L_s + 1) - \text{Pr}(L_s)}{L_s(s + 1) - L_s(s)} \quad (23)$$

步骤7 重复步骤1~步骤6,直到再增加1件最优备件时,就会有某个约束值超过给定指标上限,此时停止(只要有1个超过指标上限时就停止)。

4 实例分析

以某舰艇编队准备去索马里执行为期3个月护航任务备件清单中的7项电子类备件为例,构成了1个中心仓库和3个保障基地的2层2级后

勤保障组织结构形式,对应备件参数及各项属性值如表1所示。

表1 备件参数及属性

Table 1 Parameters and attributes of spare parts

备件编号	Z	MTBF/h	c/万元	m/kg	v/m ³	r ₁	r ₂	r ₃
LRU ₁	1	1026	3.7	29.8	2.4	0.8	0.7	0.8
LRU ₂	1	978	7.5	18.7	3.6	0.9	0.9	0.8
LRU ₃	1	1134	3.2	11.6	0.9	0.8	0.8	0.7
LRU ₄	1	654	7.3	32.6	1.1	0.9	0.7	0.7
LRU ₅	2	885	1.8	10.5	0.6	0.6	0.5	0.5
SRU ₁	1	2217	0.7	10.8	0.4	0.6	0.5	0.4
SRU ₂	1	1893	2.7	2.8	1.1	0.5	0.5	0.6

注: Z—单机安装数量; r_i (i = 1, 2, 3)—设备LRU或SRU故障后能在第r_i个基地或仓库被修好的概率。

索马里护航是中国最新型的舰艇,动力装置、武备及特装等都是目前各自领域的最先进装备,绝大部分装备是一样的或相似度非常高,故其备件差异性并不大。舰艇备件数以万计,出于对本文方法介绍的目的,选取了1组相同装备,以1个中心仓库和1条舰艇为例进行研究,并对备件属性值进行了适当的处理,但计算方法和原理是一样的。

依据表1,按照第2节方法,可以得到基地备件需求率为(1.7544, 1.8405, 1.5873, 2.7523, 4.0678, 0.8119, 0.9509),后方仓库备件需求率为(1.2281, 0.7362, 1.1110, 1.9266, 5.6949, 1.2179, 1.3313)。此次索马里护航每条舰艇对于这7项备件分配的保障经费、载荷和仓库体积的指标上限为155万元、710kg及58.5m³,伴随综合保障船对应指标为150万元、775kg及61.5m³。依据文献[3]的边际效应法,循环100次得到在费用约束下的最优备件携带方案,基地和后方仓库分别为8,4,5,5,8,9,8和10,2,4,4,11,12,10。依据备件携带方案,计算此时总的备件质量、费用和体积(见表2)。

表2 费用约束条件下备件方案结果

Table 2 Results of spare parts project under cost constraint

备件方案	c/万元	m/kg	v/m ³	方案结果是否确定
基地	154.4	737.8	60.8	否
后方仓库	149.2	758.3	61.6	否

分析表2中的数据可知,基地的备件费用为154.4万元,小于指标上限155万元,符合要求;但质量为737.8kg,超过指标上限27.8kg,同时基地备件体积为60.8m³,超过指标上限2.3m³,备件质量和体积都不满足要求;后方仓库的备件费用为149.2万元,小于指标150万元上限,符合要求;质量为758.3kg,小于指标775kg上限,符合

要求;同时基地备件体积为 61.6 m^3 , 超过指标上限 0.1 m^3 , 备件体积指标不满足要求。

携带方案成立的前提条件是所有约束条件同时满足要求,若有一项约束值超过了指标上限,备件携带方案就不符合要求,需按第 3.3 节的模型和库存优化方法重新计算。

按照第 3.3 节模型和库存优化原理重新计算得到基地和后方仓库的费用拉格朗日乘子分别为 $(16.0310, 20.5555, 7.9131, 16.9156, 5.5210, 3.9759, 6.0794)$ 和 $(15.1747, 19.7723, 7.5838, 16.1580, 5.2482, 3.7207, 5.8963)$, 按步骤 1 ~ 步骤 6 循环 98 次得到基地和仓库备件携带方案分别为 7, 4, 5, 5, 8, 9, 8 和 10, 2, 4, 4, 11, 11, 10。此时备件费用、质量和体积如表 3 所示。

表 3 多约束条件下备件方案结果

Table 3 Results of spare parts project under multi-constraints

备件方案	$c/\text{万元}$	m/kg	v/m^3	方案结果是否确定
基地	150.7	708.0	58.4	是
后方仓库	148.5	774.5	61.2	是

此时基地备件费用、质量和体积分别为 150.7 万元、708.0 kg 和 58.4 m^3 , 分别小于对应指标 155 万元、710 kg 及 58.5 m^3 的上限, 满足要求。后方仓库的备件费用、质量和体积分别为 148.5 万元, 774.5 kg 和 61.2 m^3 , 分别小于对应指标 150 万元、775 kg 及 61.5 m^3 上限, 满足要求。

下面分别给出上述 2 种情况下的备件期望短缺数 E_{BO} 和费用 c 之间的关系, 如图 2 和图 3 所示。

从图 2 和图 3 可以得出:

1) 舰艇编队备件携带方案受到的影响因素非常多, 只考虑费用约束, 在此基础上得到的备件携带方案并不是最优方案(有 2 个约束值超标)。

2) 备件短缺数曲线上的各个离散点是在各约束因素下的最优值, 且相当于该短缺数下最低

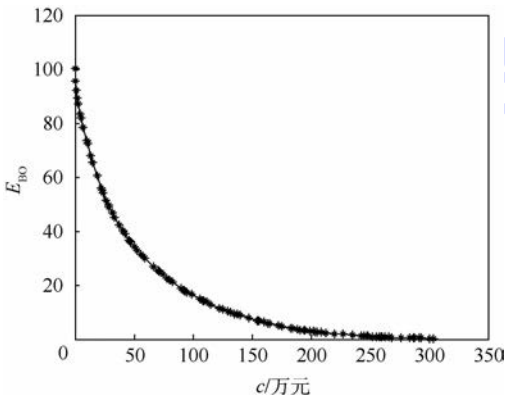
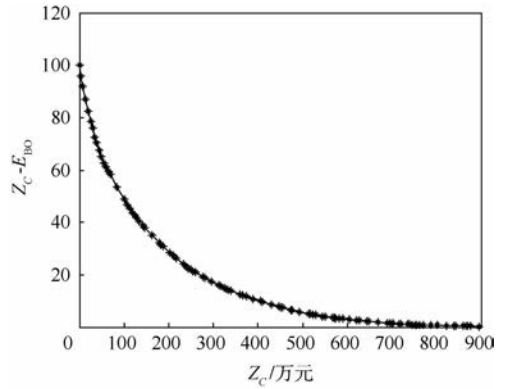


图 2 费用约束条件下的最优短缺数曲线

Fig. 2 Optimal curve of expect back order under

cost constraints



Z_c —资源约束因子。

图 3 多约束条件下的最优短缺数曲线

Fig. 3 Optimal curve of expect back order under multi-constraints

指标值。

3) 短缺数最优曲线能够为方案制定者和决策者制定编队备件携带方案提供参考和依据。以图 3 为例, 当备件短缺数为 0.3954 时, 后方仓库备件费用、质量和体积的值分别为 141 万元、775.8 kg 和 57.6 m^3 , 这是设定约束指标的最低值。

编队可用度在备件期望短缺数基础之上, 依据式(8) ~ 式(10)即可得到, 限于篇幅原因, 不再赘述。

5 结 论

1) 本文以舰艇保障经费、载荷和仓库体积为约束条件, 以备件期望短缺数和可用度为优化函数, 采用拉格朗日乘子和边际效应相结合的方法, 建立了多影响因素下的备件多层多级库存优化模型, 得到了最优方案。

2) 只考虑一种约束因素, 运用边际效应法得到的备件携带方案往往只是对应影响因素下的最优方案, 未考虑的影响因素有可能会超过指标上限, 这种方案显然是欠妥的, 运用本文提供的模型和方法, 求解出的备件库存方案满足所有影响因素指标要求。

3) 解决了多约束因素下多级多层混合的备件库存优化问题, 可为索马里护航备件携带方案制定者提供理论参考依据, 使制定的备件携带方案更加符合工程实践情况。

4) 主要针对定量约束因素进行研究, 若在工程实践中遇到定性因素时, 需要对本文模型进行适当调整, 但求解方法和思路相同(如运用专家群打分法将定性条件转换为定量约束条件, 转入第 3.3 节进行求解, 因篇幅有限, 这里不再详细叙述)。

5) 本文模型是建立在不进行横向转运和串

件拚修基础之上的,备件横向转运和串件拚修是提高保障效能指标及降低保障费用的有效手段,这将是下一步的研究重点和进一步的研究方向。

参考文献 (References)

- [1] SHERBROOKE C C. METRIC: A multi-echelon technique for recoverable item control[J]. *Operations Research*, 1968, 16(1): 122-141.
- [2] SHERBROOKE C C. An evaluator for the number of operationally ready aircraft in a multilevel supply system[J]. *Operations Research*, 1971, 19(3): 618-635.
- [3] SHERBROOKE C C. Optimal inventory modeling of systems: Multi-echelon techniques[M]. Netherland: Springer Science & Business Media, 2006: 6-10.
- [4] LEVNER E, PERLMAN Y, CHENG T C E, et al. A network approach to modeling the multi-echelon spare-part inventory system with backorders and interval-valued demand[J]. *International Journal of Production Economics*, 2011, 132(1): 43-51.
- [5] SAMOUEI P, KHEIRKHAH A S, FATTAHI P. A network approach modeling of multi-echelon spare-part inventory system with backorders and quantity discount[J]. *Annals of Operations Research*, 2015, 226(1): 551-563.
- [6] CAGLAR D, LI C L, SIMCHI-LEVI D. Two-echelon spare parts inventory system subject to a service constraint[J]. *IIE Transactions*, 2004, 36(7): 655-666.
- [7] COSTANTINO F, DI GRAVIO G, TRONCI M. Multi-echelon, multi-indenture spare parts inventory control subject to system availability and budget constraints[J]. *Reliability Engineering & System Safety*, 2013, 119: 95-101.
- [8] TOPAN E, BAYINDIR Z P. Multi-item two-echelon spare parts inventory control problem with batch ordering in the central warehouse under compound Poisson demand[J]. *Journal of the Operational Research Society*, 2012, 63(8): 1143-1152.
- [9] DADA M. A two-echelon inventory system with priority shipments[J]. *Management Science*, 1992, 38(8): 1140-1153.
- [10] SLEPTCHENKO A, VAN DER HEIJDEN M C, VAN HARTEN A. Effects of finite repair capacity in multi-echelon, multi-indenture service part supply systems[J]. *International Journal of Production Economics*, 2002, 79(3): 209-230.
- [11] AL HANBALI A, VAN DER HEIJDEN M. Interval availability analysis of a two-echelon, multi-item system[J]. *European Journal of Operational Research*, 2013, 228(3): 494-503.
- [12] LAU H C, SONG H, SEE C T, et al. Evaluation of time-varying availability in multi-echelon spare parts systems with passivation[J]. *European Journal of Operational Research*, 2006, 170(1): 91-105.
- [13] TOPAN E, BAYINDIR Z P, TAN T. An exact solution procedure for multi-item two-echelon spare parts inventory control problem with batch ordering in the central warehouse[J]. *Operations Research Letters*, 2010, 38(5): 454-461.
- [14] BASTEN R J I, VAN DER HEIJDEN M C, SCHUTTEN J M J. Joint optimization of level of repair analysis and spare parts stocks[J]. *European Journal of Operational Research*, 2012, 222(3): 474-483.
- [15] 樊九九, 郭霖瀚, 杨懿, 等. 考虑维修效能的修理级别优化[J]. *北京航空航天大学学报*, 2015, 41(2): 351-357.
- FAN J J, GUO L H, YANG Y, et al. Optimization of repair level considering maintenance effectiveness[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2015, 41(2): 351-357 (in Chinese).
- [16] 王睿, 雷红伟, 彭英武, 等. 战时任务条件下舰艇装备多级备件优化[J]. *上海交通大学学报*, 2013, 47(3): 398-403.
- WANG R, LEI H W, PENG Y W, et al. Optimization of warship spare parts in the scenario of wartime mission[J]. *Journal of Shanghai Jiao Tong University*, 2013, 47(3): 398-403 (in Chinese).
- [17] SUN L, ZUO H F. Optimal inventory modeling of multi-echelon system for aircraft spares parts[J]. *Information Technology Journal*, 2013, 12(4): 688-695.
- [18] 刘任洋, 李庆民, 李华. 基于横向转运策略的可修件三级库存优化模型[J]. *航空学报*, 2014, 35(12): 3341-3349.
- LIU R Y, LI Q M, LI H. Optimal model of three-echelon inventory for repairable spare parts with lateral transshipments strategy[J]. *Acta Aeronautica et Astronautica Sinica*, 2014, 35(12): 3341-3349 (in Chinese).
- [19] 李予阳. 大数据: 开启智能船舶时代[N]. *经济日报*, 2014-10-20(14).
- LI Y Y. Big data: Open the era of intelligent ship[N]. *Economic Daily*, 2014-10-20(14) (in Chinese).
- [20] 悬崖. 中国海军用效比考量反海盗护航任务[J]. *舰载武器*, 2014(6): 23-27.
- XUAN Y. Chinese navy's cost-effectiveness evaluation in counter piracy and escort operation[J]. *Shipborne Weapons*, 2014(6): 23-27 (in Chinese).
- [21] 孙蕾, 左洪福. 基于 METRIC 的民机初始备件数量确定及配置模型[J]. *中国机械工程*, 2013, 24(23): 3200-3205.
- SUN L, ZUO H F. Optimal allocation modeling for initial spare parts of civil aircraft based on METRIC[J]. *China Mechanical Engineering*, 2013, 24(23): 3200-3205 (in Chinese).
- [22] 王乃超, 康锐. 多约束条件下备件库存优化模型及分解算法[J]. *兵工学报*, 2009, 30(2): 247-251.
- WANG N C, KANG R. An optimization model for inventory spares under multi-constraints and its decomposition algorithm[J]. *Acta Armamentarii*, 2009, 30(2): 247-251 (in Chinese).

作者简介:

蔡芝明 男, 博士研究生. 主要研究方向: 舰艇动力及热力系统的科学管理、综合保障工程。

Tel.: 027-65460962

E-mail: caizhiming1984@163.com

金家善 男, 博士, 教授, 博士生导师. 主要研究方向: 舰艇动力及热力系统的科学管理、综合保障工程。

Tel.: 027-65460962

E-mail: jinjiashan401@163.com

陈砚桥 男, 博士, 讲师. 主要研究方向: 综合保障工程、舰船动力工程。

Tel.: 027-65460962

E-mail: Chen_yanqiao401@163.com

Optimal inventory modeling of spare parts multi-indenture multi-echelon under multi-constraints

CAI Zhiming^{1,2}, JIN Jiashan^{1,2,*}, CHEN Yanqiao^{1,2}

(1. Military Key Laboratory for Naval Ship Power Engineering, Naval University of Engineering, Wuhan 430033, China;

2. College of Power Engineering, Naval University of Engineering, Wuhan 430033, China)

Abstract: One possible way to improve the efficiency of equipment support and readiness is to optimize spare parts inventory, for the existing studies can't adjust to the demands of multi-indenture multi-echelon spare parts supply modes under multi-constraints and meet the requirements of the engineering practice in the army. Optimizing the spare parts carrying project of the warship formation dispatches to fight pirates in Somali waters during the mission preparation phase is taken as the study background. The support costs, tonnage and warehouse space of warship are used as constraint conditions, and the minimum expect back order and maximal availability are used as the target function. Optimal spare parts carrying project of the warship formation, multi-indenture multi-echelon is obtained by Lagrange multiplier method and marginal method. The total spare parts support costs, mass and volume of bases and rear warehouse are compared with the constraints index upper limit before and after optimization by example. The analysis of the example shows that the models and methods of this paper can solve the problems of the warship formation carrying spare parts, and at the same time provide the references for dealing with such problems in other areas.

Key words: spare parts ; multi-indenture multi-echelon ; multi-constraints ; optimal inventory ; warship formation

Received: 2015-06-30; **Accepted:** 2015-08-26; **Published online:** 2015-10-19 15:20

URL: www.cnki.net/kcms/detail/11.2625.V.20151019.1520.009.html

Foundation items: National Defense Pre-research Foundation of China (51319060103); China Postdoctoral Science Foundation (2013T60921)

* **Corresponding author.** Tel. : 027-65460962 E-mail: jiajiashan401@163.com

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0465



加速度场中环路热管温度波动现象分析

薛好¹, 谢永奇^{1,*}, 戴华², 余建祖¹, 马伟³

(1. 北京航空航天大学 航空科学与工程学院, 北京 100083; 2. 郑州幼儿师范高等专科学校, 郑州 453001;
3. 空间物理重点实验室, 北京 100076)

摘 要: 基于离心机系统, 搭建了双储液器环路热管 (DCCLHP) 工作特性实验台, 实验研究了环路热管在地面重力场和过载加速度场中的运行性能, 热载荷范围为 25 ~ 300 W, 过载加速度大小为 3g ~ 9g, 分析了不同热载荷、加速度方向、大小对环路热管温度波动的影响。结果表明, 过载加速度会改变环路热管的启动模式和运行性能, 可抑制或激发冷凝器出口温度波动甚至整个环路的波动, 但相同热载荷下蒸发器运行温度基本一致。回路中气液两相流动自身固有的不稳定性在过载加速度效应下更易受到激发, 加剧回路中流动不稳定, 进而加剧蒸发器向储液器漏热的不稳定。冷凝器的冷却作用可以有效削弱冷凝器段的温度波动。

关键词: 环路热管; 双储液器; 加速度场; 运行性能; 温度波动

中图分类号: V444.3; TB131

文献标识码: A **文章编号:** 1001-5965(2016)07-1502-07

环路热管 (Loop Heat Pipe, LHP) 作为一种高效的两相传热装置, 具有能自启动、反重力运行和远距离传输热量等诸多优点, 已广泛应用于航天器热管理系统和地面电子设备冷却等领域^[1-2]。对于 LHP 的运行原理和运行特性, 国内外众多学者^[3-7] 开展了深入而有成效的研究, 全面了解了其运行规律, 同时还发现 LHP 运行时可能存在温度波动、工质倒流甚至在某些布置方位下无法正常运行的现象。这些特点使得 LHP 的应用受到限制, 尤其在过载加速度环境中, LHP 运行温度会出现波动^[8-10], 甚至蒸发器毛细芯出现烧干的现象^[11]。

双储液器环路热管 (Dual Compensation Chamber Loop Heat Pipe, DCCLHP) 是在常规 LHP 蒸发器两端布置了 2 个储液器, 使其在任何姿态下都能对毛细芯有效供液, 解决常规 LHP 运行姿态受限的问题。它不同姿态和加速度环境下

均能运行的能力已为实验所证实^[12-13]。目前, 对于 DCCLHP 的研究, 主要针对地面重力场中工作特性和运行机理展开, 对其运行时温度波动现象的分析还较少^[14-15]。本文针对 DCCLHP 在地面重力场和过载加速度场中运行时观察到的温度波动现象进行分析, 以期获得有价值的结论。

1 实验装置与方案

1.1 实验装置

实验装置示意图如图 1 所示。冷却水循环子系统由恒温水槽、质量流量计、磁力泵、调节阀、过滤器、换热器和冷板等组成, 为 LHP 的冷凝器提供冷却水; 加热与测量控制子系统由恒压恒流直流电源、电加热膜、Agilent 数据采集仪、Pt100 温度传感器和计算机等组成, 用于对实验件加热、控制及数据采集、记录; 采用 Y53100-3/ZF 型恒加速度离心机模拟加速度环境。

收稿日期: 2015-07-13; 录用日期: 2015-10-14; 网络出版时间: 2015-11-19 10:06

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151119.1006.003.html

基金项目: 中央高校基本科研业务费专项资金 (YWF-14-HKXY-019)

* 通讯作者: Tel.: 010-82318081 E-mail: xyq@buaa.edu.cn

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

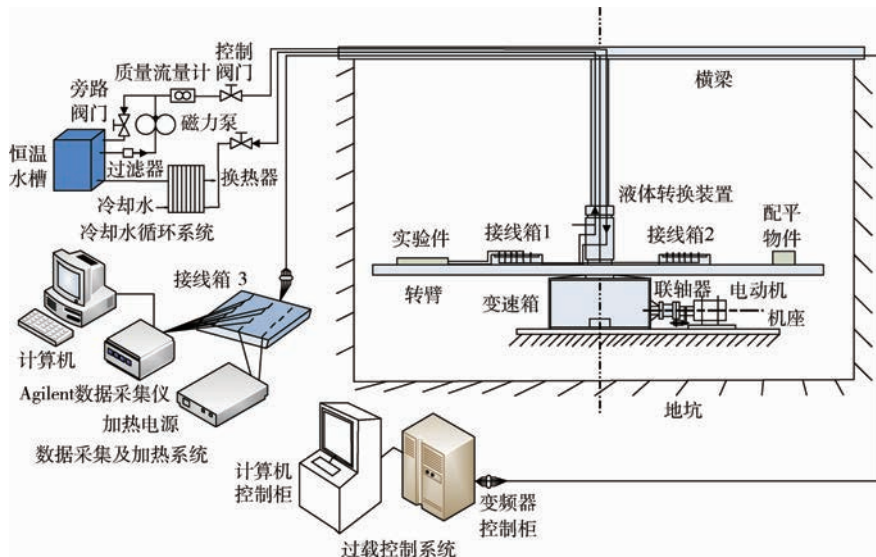


图 1 实验装置示意图

Fig. 1 Schematic of experimental apparatus

所用实验件为一套氨-不锈钢 DCCLHP,如图 2 所示,具体结构参数与安装说明见文献 [13]。采用 Pt100 铂电阻测量 DCCLHP 各部分温度, Pt100 布置位置如图 3 所示,共有 15 个测点,其中 RTD1、RTD2 和 RTD6、RTD7 分别位于储液器 1、2 上下两侧中间位置,RTD3、RTD5 分别位于蒸发器与储液器 1、2 的连接处正上方,RTD4 位于蒸发器上侧中间,RTD8 布置于冷凝器入口,RTD9 和 RTD10 位于冷凝器管线上,RTD11 位于冷凝器出口,RTD12、RTD13 分别位于液体管线中间和出口,RTD15、RTD14 用于测量冷却水进出口温度,

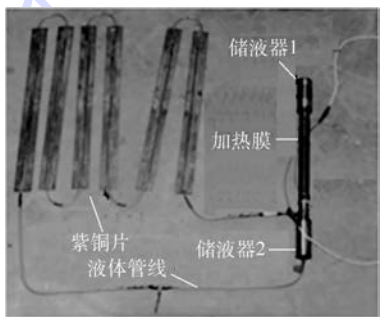


图 2 双储液器环路热管实物

Fig. 2 Picture of testing DCCLHP

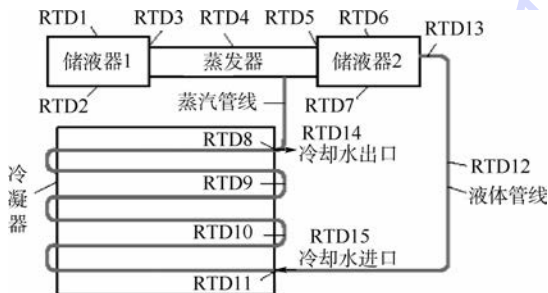


图 3 DCCLHP 测温点布置示意图

Fig. 3 Schematic of DCCLHP temperature measuring points arrangement

RTD16 用来记录环境温度。

1.2 实验方案

实验针对如下 2 种方式进行:①布置方式 A——工装箱体水平安装在转臂上,蒸发器轴线沿转臂径向,储液器 1 靠近转臂旋转轴;②布置方式 B——工装箱体水平安装在转臂上,蒸发器轴线沿转臂径向,储液器 2 靠近转臂旋转轴。

针对上述 2 种加速度方向,同时施加加热载荷和过载加速度。施加到蒸发器的热载荷分别为 25、80、150、200、250 和 300 W。过载加速度分别为 3g、5g、7g 和 9g,离心机运转时间最大不超过 1 h。实验过程中,恒温水槽冷却水温度控制在 $19 \pm 0.5^\circ\text{C}$ 。环境温度为 $23.2 \sim 24.6^\circ\text{C}$ 。实验前对所有 Pt100 在 $10 \sim 60^\circ\text{C}$ 范围内进行了标定,标定后温度测量误差小于 0.5°C ,离心加速度值的控制精度为 $\pm 5\%$ 。

2 实验结果与分析

2.1 地面重力场实验结果与分析

在给定实验条件和工况下,地面重力场中 DCCLHP 仅在 250 W 和 300 W 时运行至稳定,并观察到了环路温度波动现象。

图 4(a) 给出了 250 W 时 DCCLHP 各部分温度随时间变化曲线。由图 4(a) 可以看出,随着热载荷的施加,各测点温度迅速升高,环路内工质循环流动,表明 LHP 已经启动。观察液体管线温度变化 (RTD11 ~ RTD13),在约 200 s 时温度升高至接近 RTD5 和 RTD8,表明气液界面冲出了冷凝器并达到了液体管线出口,这种现象可能是由于工质充装量欠少所致。

在 200 s 时刻以后,RTD11 ~ RTD13 逐渐降

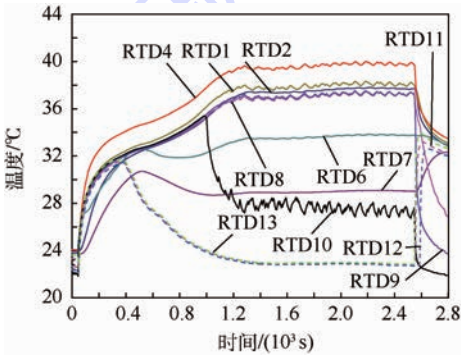
低,气液界面收缩至冷凝器内,工质回流过冷量增大,使得储液器2的温度(RTD6和RTD7)也降低。约400 s时刻,RTD6开始升高,蒸发器向储液器的漏热增大,储液器2内上部气相工质的温度升高,而下部液体工质在回流液体冷却作用下温度略有下降。在500 s时刻,RTD10急剧下降并出现了波动,表明气液界面已后移至RTD10和RTD9之间,冷凝时气液两相流的不稳定性使得冷凝器内两相区流体的温度出现波动。不过,在冷板冷却下,冷凝器出口RTD11未出现波动,这也证明了文献[15]提出的通过强化冷凝器冷却作用来抑制其温度波动的观点。

尽管难以监测环路内压力的变化,仍可以推测,冷凝器内气相工质冷凝为液体时会伴随压力的波动。根据 Clausis-Clapeyron 关系式(1)可知,

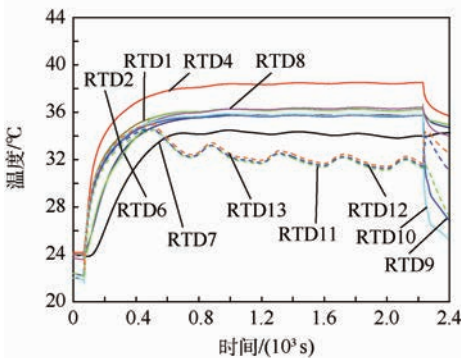
$$\Delta T \sim \frac{\Delta P_{\text{tot}} - \Delta P_w}{(dp/dT)|_{\text{sat}}} \quad (1)$$

式中: ΔT 为毛细芯内外温差; ΔP_{tot} 为回路总压差; ΔP_w 为毛细芯内外压差; dp 和 dT 分别为压力和温度的微小量。外回路压力的波动将导致蒸发器芯部蒸发温度的变化,相应的RTD4、RTD3和RTD5也出现波动,DCCLHP运行在亚稳态。

图4(b)为300 W热载荷时DCCLHP各部分温度随时间变化曲线。可以看出,冷凝器出口至



(a) 地面重力场250 W



(b) 地面重力场300 W

图4 地面重力场250 W和300 W时DCCLHP温度变化曲线

Fig. 4 Temperature fluctuation curves at 250 W and 300 W heat load of DCCLHP in gravity

液体管线出口各点温度波动明显,其余测点波动振幅微小。储液器1上下2个测点温差较小,而储液器2上下2个测点温差较大,约2.0℃。RTD10和RTD11温度变化表明气液界面在冷凝器出口附近来回移动,可能存在部分蒸气进入液体管线。产生这种现象的原因可能是由于工质充装量不足所致。在300 W热载荷工况,DCCLHP以固定热导模式运行,蒸发器RTD4的温度为38.4℃。

2.2 过载加速度场实验

在过载加速度场中,热管在布置方式A时仅在25、80和300 W、过载加速度为5g工况运行至稳定状态,在布置方式B时所有工况均运行至稳定状态。在这些运行稳定的工况中,也观察到了部分测点的温度波动现象。

2.2.1 不同布置方式

图5分别给出了布置方式A和B在25 W、5g时的温度变化曲线。可以看出,在施加加速度的初始阶段,液体管线、储液器2和冷凝器部分测点的温度存在较大变化,而且2种布置方式下环路温度变化截然不同。在图5(a)所示的布置方式A,施加热载荷后,蒸发器和储液器1的温度立即升高,储液器2的温度降低,液体管线和冷凝器测点温度则先降低再升高。这是由于在过载离心力作用下,环路内的液相工质被推向液体管线一侧,储液器1液体工质减小,储液器2液体工质增多,气液分布的变化改变了蒸发器向储液器的漏热,因此,储液器1温度升高,储液器2温度降低。冷凝器内温度较低的液体被推进液体管线,其温度降低;随后毛细芯产生的蒸气使得环路沿程各部分的温度升高,相应的RTD9~RTD13升高。蒸发器温度稳定在25.2℃。在约1500 s时,离心机停止运转,继续加热,蒸发器和储液器各测点温度又升高,而冷凝器各点温度则下降。

在图5(b)中,在施加热载荷和加速度后,蒸发器RTD4迅速升高,而蒸气管线出口RTD8先上升0.3℃后又降低1.8℃,液体管线各测点温度均迅速升高,这种温度变化表明,加速度效应使得蒸发器和储液器内的气液分布发生变化,蒸发器向储液器2的漏热增大,并可能导致芯内产生沸腾,温度较高的气泡经液体引管进入液体管线,使得RTD11温度升高,外回路工质逆向流动,将冷凝器内温度较低的液体推向冷凝器入口,因此RTD8~RTD10的温度降低。另一方面,加速度效应导致外回路压力增大,为平衡压力,蒸发器温度持续升高,直至毛细芯压差达到足以推动工质沿

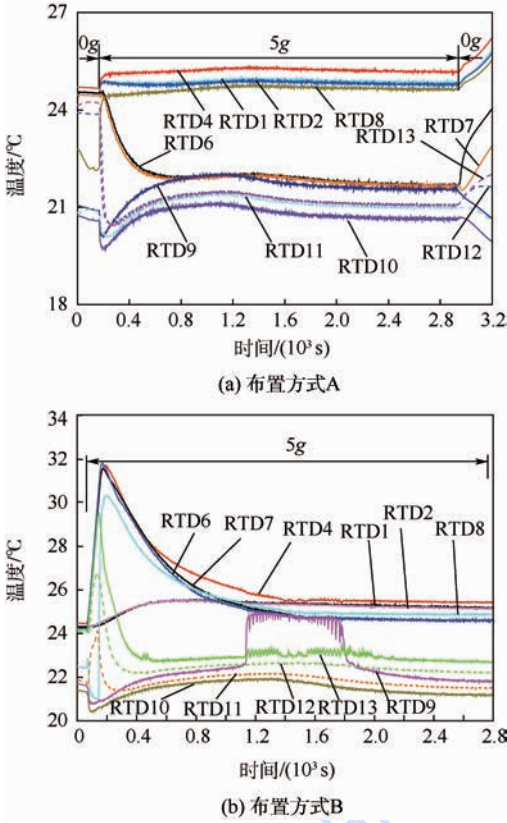


图5 布置方式A和B、25 W、5 g时DCCLHP温度变化曲线

Fig. 5 Temperature fluctuation curves at 25 W and 5 g of DCCLHP under configurations A and B

外回路正向流动,RTD8 ~ RTD10 又逐渐升高,冷凝器出口为过冷液体,使得液体管线的温度迅速下降。随着工质在外回路正向循环,环路的压降降低,蒸发器的温度下降,直至达到平衡状态,蒸发器温度稳定在 25.5℃。

然而,在建立平衡过程中,在 566 s 时刻 RTD9 的温度开始急剧升高至接近储液器 2 的温度,并开始高频小振幅的波动,同时液体管线出口、蒸发器、储液器温度也出现小幅波动。在 890 s 以后,RTD9、RTD13 的温度波动逐渐消失,蒸发器和储液器温度也趋于稳定。导致这种特殊现象的原因还有待于进一步研究,不过冷凝器的温度波动可以诱发蒸发器的温度波动。

图 6 分别为布置方式 A 和 B 在 300 W、5g 时的温度变化曲线。由图 6(a)可以看出,施加热载荷和加速度后,环路各部分温度迅速升高,至 510 s 时,蒸发器温度为 37.8℃,储液器 1 的温度为 34.5℃,储液器 2 的温度为 33.5℃。冷凝器冷凝面积被全部利用,DCCLHP 以固定热导模式运行。随后液体管线温度开始小幅度波动,蒸发器和储液器 1 温度也出现轻微波动,储液器 2 温度降低。至 960 s,蒸发器温度为 37.3℃。对比

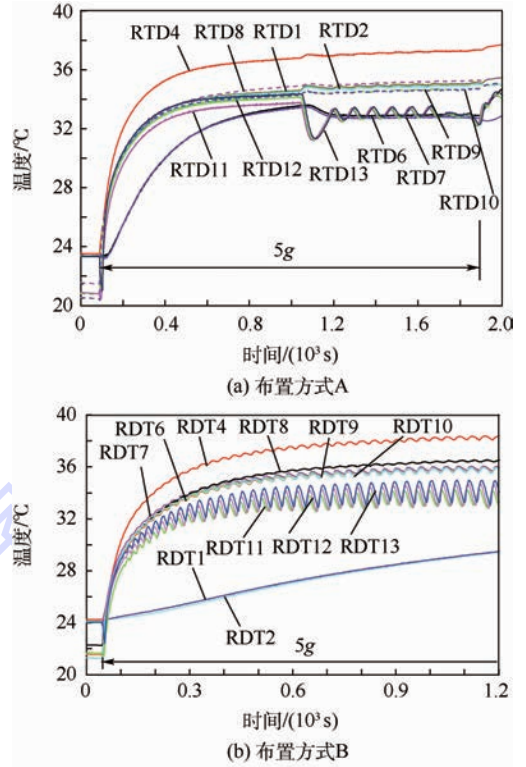


图6 布置方式A和B、300 W、5 g时DCCLHP温度变化曲线

Fig. 6 Temperature fluctuation curves at 300 W and 5 g of DCCLHP under configurations A and B

图 6(a)和图 4(b)可以看出,该方向加速度对冷凝器出口温度波动具有一定的抑制作用。

由图 6(b)可知,在载荷施加 50 s 以后,冷凝器出口 RTD11 温度开始出现小幅波动上升,且振幅逐渐增大,直至变为较为规律的振幅较大的周期性波动,且 RTD11、RTD12 和 RTD13 之间存在约 6 s 的相位差。70 s 以后,储液器 2、冷凝器各点、蒸发器也依次开始小幅低频的波动,蒸发器最终在 38.1 ~ 38.3℃ 之间波动。相对于图 4(b)所示的地面重力场和图 6(a)所示工况,该方向加速度效应激发了环路内工质不稳定流动,使得除储液器 1 外其他各部分测点温度出现了周期性的波动。但蒸发器最终运行温度差别不大。

该布置方式下,储液器 1 的温度低于储液器 2 的温度,这是由于在离心力作用下,液体工质被推入储液器 1,蒸发器向储液器 1 的漏热相对较小,使得 RTD1 和 RTD2 温度较低。

综上所述,过载加速度可以改变 LHP 启动模式或运行性能,可以抑制或激发冷凝器出口温度波动甚至整个环路的波动,但相同热载荷下蒸发器运行温度基本一致。

2.2.2 不同加速度大小

在布置方式 B、250 W 和 300 W、不同过载加速度大小工况下,DCCLHP 各测点均出现了温度波动。图 7 给出了 250 W、3 g 和 5 g 时环路各部分温度变化曲线。

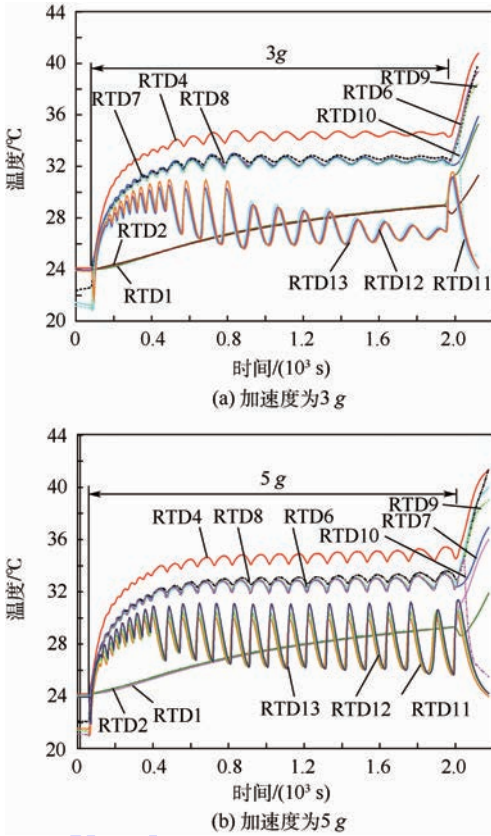


图 7 3 g 和 5 g、250 W 时 DCCLHP 温度变化曲线
Fig.7 Temperature fluctuation curves at 250 W of DCCLHP under 3 g and 5 g conditions

图 7(a)中前 520 s 环路温度变化与图 6(b)所示工况相似,除 RTD1 和 RTD2 之外各测点温度均出现了温度波动。但在 520 s 以后,温度波动振幅增大,频率减小。在 900 s 之后,液体管线 RTD11 ~ RTD13 振幅逐渐减小,蒸发器、冷凝器、储液器 2 的温度波动也随之有所减小。RTD13 温度波动范围为 25.6 ~ 30.9℃,蒸发器温度 RTD4 逐步趋近于 34.5℃。

图 7(b)中前 900 s 各测点温度变化与图 7(a)相似,在 900 s 之后 RTD11 ~ RTD13 的波动频率和振幅基本保持不变,这与图 7(a)中 3 g 工况明显不同,而且其他各测点也保持周期性波动状态。RTD13 温度波动范围为 26.1 ~ 30.2℃,蒸发器温度 RTD4 波动范围为 34.4 ~ 35.3℃。

在上述温度波动中,液体管线的温度波动与蒸发器、冷凝器、储液器 2 的温度波动相位差为 90°,这是由于当蒸发器向储液器 2 漏热增加时,

蒸发器 RTD4 升高,直到处于波峰,为了平衡漏热,冷凝器内气液界面向后移动,回流液体过冷量增加,冷凝器出口 RTD11 降低,直到处于波谷。相反的,当蒸发器向储液器 2 漏热减少时,蒸发器 RTD4 降低,直到处于波谷,为了平衡漏热,冷凝器内气液界面向前移动,回流液体过冷量降低,直到 RTD11 处于波峰。正是蒸发器向储液器 2 漏热的不稳定,导致了这种温度波动现象。

在约 1 500 s 时,离心机停止运转,继续加热,蒸发器和储液器各测点温度又升高,而冷凝器各点温度则下降。

综合分析热载荷为 250 W 和 300 W 各工况,环路温度均存在波动,尤其在过载加速度场中,环路温度波动更为显著。其原因可能是由于工质充装量与储液器容积不匹配,在较大热载荷下气液界面易于冲出冷凝器进入液体管线,导致液体管线温度出现波动;另一方面,回路中气液两相流动自身固有的不稳定性在过载加速度效应下更易受到激发,加剧回路中流动不稳定,进而加剧蒸发器向储液器漏热的不稳定,从而导致了上述温度波动现象。

对于布置方式 B,在 300 W 和 3 g、5 g、7 g、9 g 时均观察到了类似 250 W、5 g 时的温度波动现象,如图 8 所示。在不同过载大小条件下,DCCLHP 蒸发器测点温度 RTD4 波动振幅和周期稍有变化,其中 3 g 过载时波动周期最长,振幅最小;9 g 过载时波动周期最短,振幅最大。对于冷凝器出口测点温度 RTD11,在不同过载下其波动振幅均比 RTD4 温度波动振幅大,但 3 g 过载时其波动振幅最大。

由以上分析可知,在布置方式 B 时,DCCLHP 在 25、200、250 和 300 W 加热功率下均出现了温度波动现象,产生温度波动的功率范围较宽,且在较高加热功率 250 W 和 300 W 时波动剧烈,这对于 DCCLHP 在过载下的应用十分不利,在实际应

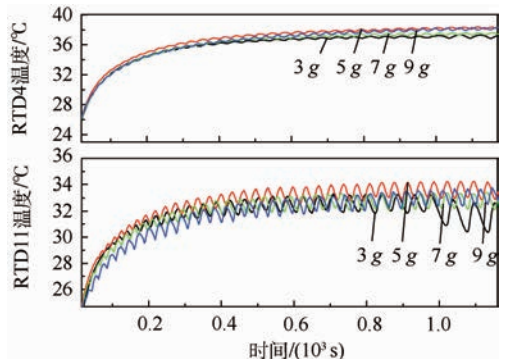


图 8 300 W、不同加速度时 DCCLHP 温度波动
Fig.8 Temperature fluctuations at 300 W of DCCLHP under different acceleration magnitudes conditions

用中应当尽量消除这种温度波动现象。另外,较大的过载使 DCCLHP 蒸发器波动的周期变短,振幅加大,加剧了波动现象。

3 结 论

本文针对热载荷和过载加速度同时施加时双储液器环路热管温度波动现象进行了分析,主要得到如下结论:

1) 冷凝器内气液两相流的不稳定性导致了其两相区流体的温度出现波动;通过强化冷凝器冷却作用可以抑制液体管线温度波动。

2) 相对于地面重力场,布置方式 A 时加速度效应对环路中温度波动具有一定抑制作用,而布置方式 B 则激发了其温度波动,且高功率下温度波动剧烈。对于布置方式 B,加速度增大,运行温度波动的周期变小,振幅增大,但振幅增加量较小。

3) 导致 DCCLHP 温度波动的主要原因可能是加速度效应增强了气液两相流的不稳定和蒸发器向储液器漏热的不稳定。

由于 DCCLHP 在运行时蒸发器与储液器内工质的变化情况难以观察,对于 DCCLHP 的温度波动机理分析还有待于结合可视化进一步深入研究。

参考文献 (References)

- [1] MAYDANIK Y F. Loop heat pipes [J]. Applied Thermal Engineering, 2005, 25(5-6): 635-657.
- [2] AMBIRAJAN A, ADONI A A, VAIDYA J S, et al. Loop heat pipes: A review of fundamentals, operation, and design [J]. Heat Transfer Engineering, 2012, 33(4-5): 387-405.
- [3] BAI L Z, LIN G P, WEN D S, et al. Experimental investigation of startup behaviors of a dual compensation chamber loop heat pipe with insufficient fluid inventory [J]. Applied Thermal Engineering, 2009, 29(8-9): 1447-1456.
- [4] LIN G P, ZHANG H X, SHAO X G, et al. Development and test results of a dual compensation chamber loop heat pipe [J]. Journal of Thermophysics and Heat Transfer, 2006, 20(4): 825-834.
- [5] KU J. Operating characteristics of loop heat pipes; SAE 1999-

01-2007 [R]. Detroit: SAE, 1999.

- [6] LAUNAY S, SARTRE V, BONJOUR J. Parametric analysis of loop heat pipe operation: A literature review [J]. International Journal of Thermal Sciences, 2007, 46(7): 621-636.
- [7] CHUANG P A. An improved steady-state model of loop heat pipes based on experimental and theoretical analyses [D]. Pennsylvania: The Pennsylvania State University, 2003.
- [8] KU J, OTTENSTEIN L, KOBEL M, et al. Temperature fluctuations in loop heat pipe operation [C] // AIP Conference Proceedings. [S. l. : s. n.], 2001, 552: 255-262.
- [9] KU J, OTTENSTEIN L, KAYA T. et al. Testing of a loop heat pipe subjected to variable accelerating forces, Part 1: Start-up; SAE 2000-01-2488 [R]. Detroit: SAE, 2000.
- [10] KU J, OTTENSTEIN L, KAYA T, et al. Testing of a loop heat pipe subjected to variable accelerating forces, Part 2: Temperature stability; SAE 2000-01-2489 [R]. Detroit: SAE, 2000.
- [11] FLEMING A J, THOMAS S K, YERKES K L, et al. Titanium-water loop heat pipe operating characteristics under standard and elevated acceleration fields [J]. Journal of Thermophysics and Heat Transfer, 2010, 24(1): 184-198.
- [12] LIN G P, LI N, BAI L Z, et al. Experimental investigation of a dual compensation chamber loop heat pipe [J]. International Journal of Heat and Mass Transfer, 2010, 53(15-16): 3231-3240.
- [13] XIE Y Q, ZHANG J, XIE L Y, et al. Experimental investigation on the operating characteristics of a dual compensation chamber loop heat pipe subjected to acceleration field [J]. Applied Thermal Engineering, 2015, 81(3): 297-312.
- [14] GONCHAROV K A, KOTLYAROV E Y, SMIRNOV F Y, et al. Investigation of temperature fluctuations in loop heat pipes; SAE 941577 [R]. Detroit: SAE, 1994.
- [15] FENG J T, LIN G P, BAI L Z. Experimental investigation on operating instability of a dual compensation chamber loop heat pipe [J]. Science in China Series E-Technological Sciences, 2009, 52(8): 2316-2322.

作者简介:

薛好 男, 硕士研究生。主要研究方向: 环路热管、强化传热等。

E-mail: buaxh@126.com

谢永奇 男, 博士, 讲师, 硕士生导师。主要研究方向: 飞行器环境控制、电子设备热设计等。

Tel.: 010-82318081

E-mail: xyq@buaa.edu.cn

Investigation of temperature fluctuation in loop heat pipe under acceleration fields

XUE Hao¹, XIE Yongqi^{1,*}, DAI Hua², YU Jianzu¹, MA Wei³

(1. School of Aeronautic Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;

2. Zhengzhou Kindergarten Teachers College, Zhengzhou 453001, China;

3. Science and Technology on Space Physics Laboratory, Beijing 100076, China)

Abstract: Based on the centrifugal machine system, a test apparatus was set up to investigate the operating characteristics of dual compensation chambers loop heat pipe (DCCLHP) under both gravity and acceleration fields. The ranges of the heat load and the magnitude of acceleration are 25 – 300W and 3g – 9g, respectively. The influences of different heat loads, acceleration directions and magnitudes on temperature fluctuations of loop heat pipe were analyzed. The results indicate that the effect of the acceleration can change the start mode and operating performance of the DCCLHP. The temperature fluctuation at the outlet of the condenser and even around the whole loop can be restrained or wakened. But the operating temperature of the evaporator is nearly the same under the same heat load conditions. The inherent flow instability of the gas-liquid two phase flow can be enhanced by the acceleration effect which reinforces this unstable flow in reverse. Furthermore, the instability of the heat leak from the evaporator to the compensation chamber is enhanced. In addition, the cooling effect of the condenser can weaken the temperature fluctuation of the condensing pipe line.

Key words: loop heat pipe; dual compensation chambers; acceleration field; operating characteristics; temperature fluctuation

Received: 2015-07-13; **Accepted:** 2015-10-14; **Published online:** 2015-11-19 10:06

URL: www.cnki.net/kcms/detail/11.2625.V.20151119.1006.003.html

Foundation item: the Fundamental Research Funds for the Central Universities(YWF-14-HKXY-019)

* **Corresponding author.** Tel. : 010-82318081 E-mail: xyq@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0398



含区间参数不确定结构的损伤识别方法

杨君坦¹, 邱志平^{1,*}, 李琦²

(1. 北京航空航天大学 航空科学与工程学院, 北京 100083; 2. 中国科学院沈阳自动化研究所 机器人学国家重点实验室, 沈阳 110016)

摘 要: 针对损伤识别中遇到的含统计信息不足的不确定参数的结构, 以区间向量来描述其中的不确定参数。基于有限元理论给出了应变模态的测试方法, 进一步将应变模态表示为区间参数的函数, 结合一阶 Taylor 展开与区间分析, 计算了参数不确定条件下应变模态的界限, 给出了一种考虑参数不确定的结构损伤识别方法, 并在理论上对比分析了与基于概率方法的区别与联系。为验证方法的合理性, 对简支梁在不同位置损伤的情况进行了数值模拟, 分析了不同损伤程度以及参数不同不确定程度下简支梁应变模态的变化规律, 简要讨论了噪声对识别准确性的影响。为含不确定参数的结构的损伤识别提供了一种新的方法和思路。

关键词: 损伤识别; 不确定参数; 区间分析; 有限元法; 模态分析

中图分类号: V214.3; V215; O326

文献标识码: A

文章编号: 1001-5965(2016)07-1509-09

对于航空航天、土木等工程结构, 尤其是其重要受力部件, 当出现一定损伤而未及时发现时, 损伤往往会迅速发展, 甚至引发严重的后果, 因此对结构早期损伤进行识别与预报显得尤为重要^[1-6]。对于小型、规则的结构, 染色渗透、超声探测和电磁涡流等局部损伤识别方法通常能够取得较好的效果。但对于不可见、狭小封闭空间内的结构, 或损伤的大致区域不明时, 局部损伤识别方法难以取得较好的效果^[3-4]。为此, 全局识别方法得到了广泛研究与应用, 其总体思路是将实际结构视为由质量矩阵、阻尼矩阵和刚度矩阵所构成的力学系统, 当系统发生损伤后, 上述矩阵 (主要是刚度矩阵) 参数就会有相应的改变, 进而引起系统响应或相关静、动力学特性的变化, 识别这些特性的变化, 即可获得系统损伤的情况^[5-6]。

为表征损伤对结构参数的改变, 许多静、动力指标被提了出来。固有频率是结构振动特性中的重要参数, 通常能以很高的精度测得, 早期出现了

许多基于频率的损伤识别研究。Chondros 和 Dimarogonas^[7] 提出以扭簧模拟梁结构的损伤裂纹, 由理论和试验证明了通过检测固有频率的改变量能得知损伤的存在。Cawley 和 Adams^[8-9] 基于一阶 Taylor 展开, 在结构小损伤的前提下证明了损伤前后任意两阶固有频率变化量的比值只是结构损伤位置的函数。此后该结论得到了进一步的研究与改进^[10-11]。然而基于频率的损伤识别方法有着较多不足, 如小损伤的情况下, 频率变化主要体现在高阶, 而获得高阶固有频率较困难, 此外, 由于固有频率反映的系统信息较少, 通常只能发现破损的存在, 但难于对其进行定位, 且不同位置处的损伤可能产生相同的频率改变^[3]。不同于固有频率, 结构位移模态反映了系统更多的信息, 因此位移模态也可作为损伤识别的指标, 常用的方法有模态置信度判据 (MAC) 和坐标模态置信度判据 (COMAC)^[12-13] 等, 这些判据的实质是损伤前后任意两个模态向量夹角余弦值的平

收稿日期: 2015-06-17; 录用日期: 2016-01-08; 网络出版时间: 2016-03-11 10:35

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20160311.1035.001.html

基金项目: 国家自然科学基金 (11372025, 11432002)

* 通讯作者: Tel.: 010-82339628 E-mail: zpqiu@buaa.edu.cn

引用格式: 杨君坦, 邱志平, 李琦. 含区间参数不确定结构的损伤识别方法[J]. 北京航空航天大学学报, 2016, 42(7): 1509-1517.
YANG J T, QIU Z P, LI Q. Damage detection of structures with interval uncertain parameters[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1509-1517 (in Chinese).

方,当结构无损伤时,由模态正交性知 MAC 矩阵是单位阵,而当结构存在损伤时,损伤后模态与损伤前模态不满足正交性,便可由 MAC 矩阵判断结构损伤。由于损伤位置附近自由度上的振型变化量相对较大,利用损伤前后振型相对变化量便可对损伤进行定位。相比于位移模态,模态曲率对结构损伤更加敏感,因此模态曲率也被当作损伤识别的参数。实际中直接利用位移模态参数往往存在着对局部小损伤不够敏感的问题,而模态曲率法大都是基于位移模态的差分,引入了误差。考虑到应变类参数对结构损伤一般比位移类参数更加敏感,而且通常能由应变计直接测出^[14],应变模态指标逐渐在结构损伤识别中得到广泛应用。目前,随着实际工程结构越来越复杂以及学科间的相互交叉,许多新方法不断被提出与应用,如基于小波分析、控制理论和人工智能等的损伤识别方法^[3,5]。

实际工程中由于环境温差、测量误差和材料分散性等因素,结构的参数会存在着不同程度的不确定性^[15-18],这会对结构损伤识别有着不同的影响,甚至影响结果导向^[4,16],因此研究含不确定参数结构的损伤识别有着重要的意义。通常分析不确定性主要是基于概率思想,而该方法中需要解决的一个基础性问题就是结构参数的概率分布,为此人们提出了多种解决方法,Bayesian 模型修正法的基本思想是用参数的概率分布描述结构的不确定性,根据观测数据修正模型参数的不确定性,进而确定模型的后验概率分布^[16],然而由于后验概率密度函数计算的复杂性,Bayesian 方法较难应用于未知参数较多的大规模计算问题。区别于 Bayesian 方法,随机有限元修正法认为模型不确定性、测量误差对修正参数的影响可以通过观测数据的统计平均有效降低,因此利用实测数据和模型参数摄动的随机模拟来获得相应参数的统计特征^[17-18]。但测试数据往往是非常有限的,因此很多情况下仍然必须采用 Monte Carlo 随机模拟的方法,这会降低计算分析效率^[16]。

鉴于此,本文直接将结构不确定参数建模为区间数,实际应用中不需要确定参数的统计信息,而只要知道相应的上下界。以应变模态为损伤识别判据,结合一阶 Taylor 展开与区间分析,提出了一种考虑含区间参数的结构损伤识别方法。通过数值算例对不同损伤程度以及不同不确定量下的损伤识别进行了分析,证明了本文方法的可行性与合理性。本文方法对于含不确定参数结构的损伤识别提供了一种新的思路与方向。

1 基于应变模态的结构损伤识别

对于给定的位移模态,必然有特定的应变分布状况与之对应,这种固有应变分布状态被称为结构的应变模态。国内外学者^[19-20]对应变模态的表达式、正交性等进行了相关研究,下面从有限元角度对应变模态进行推导。

将实际结构离散为有限元模型,设 \mathbf{u}_e 为单元节点位移列向量,则单元内部任意一点的位移为

$$\boldsymbol{\delta} = \mathbf{N}\mathbf{u}_e \quad (1)$$

式中: \mathbf{N} 为形函数矩阵。

利用应变与位移的关系,可得单元内部任意一点的应变:

$$\boldsymbol{\varepsilon} = \mathbf{A}^T \mathbf{N}\mathbf{u}_e = \mathbf{B}\mathbf{u}_e \quad (2)$$

式中: \mathbf{A} 为微分算子矩阵; \mathbf{B} 为应变矩阵。根据有限元理论,节点位移满足如下微分方程:

$$\mathbf{M}\ddot{\mathbf{u}}_e + \mathbf{C}\dot{\mathbf{u}}_e + \mathbf{K}\mathbf{u}_e = \mathbf{f} \quad (3)$$

式中: \mathbf{M} 、 \mathbf{C} 、 \mathbf{K} 和 \mathbf{f} 分别为结构总体的质量矩阵、阻尼矩阵、刚度矩阵和载荷列向量,对于简谐激励,可令 $\mathbf{f} = \mathbf{F}e^{j\omega t}$,其中 \mathbf{F} 为激励力幅值向量, ω 为激励力频率, j 为虚数单位。则频域内稳态位移响应可表示为 $\mathbf{u}_e = \mathbf{U}e^{j\omega t}$,代入式(3)可得

$$(-\omega^2 \mathbf{M} + j\omega \mathbf{C} + \mathbf{K})\mathbf{U} = \mathbf{F} \quad (4)$$

基于模态叠加原理,将式(4)的解表示为

$$\mathbf{U} = \boldsymbol{\Phi}\mathbf{Y}\boldsymbol{\Phi}^T \mathbf{F} \quad (5)$$

式中: $\boldsymbol{\Phi} = [\boldsymbol{\varphi}_1 \ \boldsymbol{\varphi}_2 \ \cdots \ \boldsymbol{\varphi}_k \ \cdots \ \boldsymbol{\varphi}_m]$, $\boldsymbol{\varphi}_k$ 为结构第 k 阶位移模态向量; $\mathbf{Y} = \text{diag}[Y_1 \ Y_2 \ \cdots \ Y_k \ \cdots \ Y_m]$, $Y_k = (k_k - \omega^2 m_k + j\omega c_k)^{-1}$, k_k 、 m_k 和 c_k 分别为模态刚度矩阵、质量矩阵和阻尼矩阵的第 k 个对角元(本文所考虑的是无阻尼结构或比例黏性阻尼结构)。则系统位移响应可表示为

$$\mathbf{u}_e = \mathbf{U}e^{j\omega t} = \boldsymbol{\Phi}\mathbf{Y}\boldsymbol{\Phi}^T \mathbf{F}e^{j\omega t} \quad (6)$$

将式(6)代入式(2),可得

$$\boldsymbol{\varepsilon} = \mathbf{E}\mathbf{e}^{j\omega t} = \mathbf{B}\boldsymbol{\Phi}\mathbf{Y}\boldsymbol{\Phi}^T \mathbf{F}e^{j\omega t} = \boldsymbol{\psi}^\varepsilon \mathbf{Y}\boldsymbol{\Phi}^T \mathbf{F}e^{j\omega t} = \boldsymbol{\psi}^\varepsilon \mathbf{q}^\varepsilon \quad (7)$$

式中: \mathbf{E} 为应变幅值向量; $\boldsymbol{\psi}^\varepsilon = \mathbf{B}\boldsymbol{\Phi}$ 即为系统正应变模态振型矩阵; \mathbf{q}^ε 为应变模态的模态坐标。将式(7)写为分量形式:

$$\boldsymbol{\varepsilon} = \sum_{r=1}^N q_r \boldsymbol{\psi}_r^\varepsilon = \sum_{r=1}^N \frac{\boldsymbol{\psi}_r^\varepsilon \boldsymbol{\varphi}_r^T \mathbf{F}}{k_r - \omega^2 m_r + j\omega c_r} = \mathbf{H}\mathbf{F} \quad (8)$$

式中: N 为系统自由度数; $\boldsymbol{\psi}_r^\varepsilon$ 为第 r 阶应变模态向量; q_r 为位移模态坐标的第 r 个分量; \mathbf{H} 称为应变频响函数矩阵,且有

$$\mathbf{H} = \begin{pmatrix} H_{11} & H_{12} & \cdots & H_{1N} \\ H_{21} & H_{22} & \cdots & H_{2N} \\ \vdots & \vdots & & \vdots \\ H_{N1} & H_{N2} & \cdots & H_{NN} \end{pmatrix} = \sum_{r=1}^N \frac{1}{k_r - \omega^2 m_r + j\omega c_r} \cdot$$

$$\begin{pmatrix} \psi_{1r}^e \varphi_{1r} & \psi_{1r}^e \varphi_{2r} & \cdots & \psi_{1r}^e \varphi_{Nr} \\ \psi_{2r}^e \varphi_{1r} & \psi_{2r}^e \varphi_{2r} & \cdots & \psi_{2r}^e \varphi_{Nr} \\ \vdots & \vdots & & \vdots \\ \psi_{Nr}^e \varphi_{1r} & \psi_{Nr}^e \varphi_{2r} & \cdots & \psi_{Nr}^e \varphi_{Nr} \end{pmatrix} \quad (9)$$

其中: H_{st} 代表在第 t 个自由处施加载荷时对 s 自由度处的频响函数值; ψ_{ir}^e 为第 r 阶应变模态的第 i 个分量; φ_{ir} 为第 r 阶位移模态的第 i 个分量。

由式(9)可以看出,与位移频响矩阵不同,应变频响函数矩阵并不对称,但矩阵的任一列仍包含了应变模态的全部信息。因此在实际测量中,只需要在某一点处施加激励,测得各点处的应变频率响应,对于第 r 阶固有频率处的测量值,可以认为该阶振动模态占主导地位而忽略其他阶次的影响。由于结构模态特性只反映了各自由度上相应量的比值关系,因此只需采取一定的归一化方法即可获得相应的第 r 阶应变模态。对于损伤后的结构,测试的方法与此相同,但由于在损伤区域附近结构应变会有明显的重分布,因此对比完好结构的相应模态就能对损伤进行识别。

2 不确定结构的应变模态区间分析

2.1 区间分析理论

受到腐蚀性环境、制造与测量误差、材料的分散性等因素的影响,结构的几何、物理参数通常存在一定的不确定性。实际工程中,这些不确定参数往往缺乏足够的数据来准确定义结构参数的概率分布,因此采用只需要确定上下界的区间数来描述系统不确定参数是较为可行的一种思路。设结构参数向量可以表示为

$$\mathbf{b} = (b_1, b_2, \dots, b_p) \quad (10)$$

基于区间数学思想,认为结构参数向量 \mathbf{b} 属于某一区间参数向量 \mathbf{b}' , 即 $\mathbf{b} \in \mathbf{b}'$, 写成分量形式 $b_k \in b'_k = [\underline{b}_k, \overline{b}_k] \quad k=1, 2, \dots, p$ (11)

式中: 区间数 b'_k 为区间向量 \mathbf{b}' 的第 k 阶分量; \underline{b}_k 、 \overline{b}_k 分别为其上、下界。区间参数向量 \mathbf{b}' 的名义值 \mathbf{b}^c 和不确定半径 $\Delta \mathbf{b}$ 分别定义如下:

$$\begin{cases} \mathbf{b}^c = [b_1^c, b_2^c, \dots, b_p^c] \\ \Delta \mathbf{b} = (\Delta b_1, \Delta b_2, \dots, \Delta b_p) \end{cases} \quad (12)$$

式中: $b_k^c = (\underline{b}_k + \overline{b}_k)/2$; $\Delta b_k = (\overline{b}_k - \underline{b}_k)/2, k=1, 2, \dots, p$ 。上述参数向量 \mathbf{b} 可以分解为 \mathbf{b}' 的名义值 $\mathbf{b}^c = [b_1^c, b_2^c, \dots, b_p^c]$ 与偏差 $\boldsymbol{\delta} = [\delta_1, \delta_2, \dots, \delta_p]$ 的和的形式:

$$\mathbf{b} = \mathbf{b}^c + \boldsymbol{\delta} \quad (13)$$

其分量满足

$$b_k = b_k^c + \delta_k \quad k=1, 2, \dots, p \quad (14)$$

式中: $\delta_k \in \delta_k^l = [-\Delta b_k, \Delta b_k]$ 为 b_k 的不确定部分。基于有限元分析理论,式(3)中的系统矩阵可以表示为不确定参数的函数,即

$$\begin{cases} \mathbf{K} = \mathbf{K}(\mathbf{b}) \\ \mathbf{M} = \mathbf{M}(\mathbf{b}) \\ \mathbf{C} = \mathbf{C}(\mathbf{b}) \end{cases} \quad (15)$$

则由上述矩阵所求得的结构位移模态矩阵 $\boldsymbol{\varphi}$ 与应变模态矩阵 $\boldsymbol{\psi}^e$ 也是上述参数的函数,即

$$\begin{cases} \boldsymbol{\psi}^e = \boldsymbol{\psi}^e(\mathbf{b}) \\ \boldsymbol{\varphi} = \boldsymbol{\varphi}(\mathbf{b}) \end{cases} \quad (16)$$

由于 \mathbf{b} 在某一有界区间内变动,因此应变模态 $\boldsymbol{\psi}^e$ 也在某一有界区间内,但由于其关于不确定量的函数通常是隐式的,因此很难直接求得 $\boldsymbol{\psi}^e$ 的区间,一个可行的办法是将 $\boldsymbol{\psi}^e$ 在不确定变量中值处进行 Taylor 展开,进一步基于区间扩张理论确定 $\boldsymbol{\psi}^e$ 的上下界。

2.2 应变模态的区间分析方法

基于 Taylor 展开,结合式(16),将结构第 r 阶应变模态向量在结构参数中值 $\mathbf{b}^c = [b_1^c, b_2^c, \dots, b_p^c]$ 处进行一阶 Taylor 展开:

$$\boldsymbol{\psi}_r^e(\mathbf{b}) = \boldsymbol{\psi}_r^e(\mathbf{b}^c) + \sum_{k=1}^p \frac{\partial \boldsymbol{\psi}_r^e(\mathbf{b}^c)}{\partial b_k} \delta_k \quad (17)$$

基于区间运算与区间扩张原理,利用方程式(17),可以得到基于 Taylor 展开的含不确定参数结构的应变模态的上下界:

$$\begin{cases} \overline{\boldsymbol{\psi}}_r^e = \boldsymbol{\psi}_r^e(\mathbf{b}^c) + \sum_{k=1}^p \left| \frac{\partial \boldsymbol{\psi}_r^e(\mathbf{b}^c)}{\partial b_k} \right| \Delta b_k \\ \underline{\boldsymbol{\psi}}_r^e = \boldsymbol{\psi}_r^e(\mathbf{b}^c) - \sum_{k=1}^p \left| \frac{\partial \boldsymbol{\psi}_r^e(\mathbf{b}^c)}{\partial b_k} \right| \Delta b_k \end{cases} \quad (18)$$

式中: $\boldsymbol{\psi}_r^e(\mathbf{b}^c)$ 可通过令 $\mathbf{b} = \mathbf{b}^c$ 并结合式(7)得到。一阶偏导 $\partial \boldsymbol{\psi}_r^e(\mathbf{b}^c)/\partial b_k$ 计算的推导如下,为表达方便,下文中将 $\mathbf{K}(\mathbf{b})$ 、 $\mathbf{M}(\mathbf{b})$ 和 $\boldsymbol{\psi}_r^e(\mathbf{b})$ 等量简记为 \mathbf{K} 、 \mathbf{M} 和 $\boldsymbol{\psi}_r^e$ 。由式(7)易知:

$$\boldsymbol{\psi}_r^e = \mathbf{B} \boldsymbol{\varphi}_r \quad (19)$$

式(19)两边同时关于参数 b_k 求导,可得

$$\frac{\partial \boldsymbol{\psi}_r^e}{\partial b_k} = \mathbf{B} \frac{\partial \boldsymbol{\varphi}_r}{\partial b_k} + \frac{\partial \mathbf{B}}{\partial b_k} \boldsymbol{\varphi}_r \quad (20)$$

由式(2)知,应变矩阵 \mathbf{B} 只与形函数矩阵和微分算子矩阵有关,当 b_k 为结构的几何或物理参数时, $\partial \mathbf{B}/\partial b_k = 0$ 。对于有限自由度系统孤立特征值的情况,结构的特征向量可以张成一个完备的正交空间, $\partial \boldsymbol{\varphi}_r/\partial b_k$ 可由特征向量的线性组合表示:

$$\frac{\partial \boldsymbol{\varphi}_r}{\partial b_k} = \sum_{s=1}^n H_{lks} \boldsymbol{\varphi}_s \quad (21)$$

式中:相关系数 H_{lks} 可由方程式(22)求出:

$$H_{lks} = \begin{cases} -\frac{1}{2} \boldsymbol{\varphi}_l^T \frac{\partial \mathbf{M}}{\partial b_k} \boldsymbol{\varphi}_l & l = s \\ \frac{1}{\lambda_l - \lambda_s} \boldsymbol{\varphi}_s^T \left(\frac{\partial \mathbf{K}}{\partial b_k} - \lambda_l \frac{\partial \mathbf{M}}{\partial b_k} \right) \boldsymbol{\varphi}_l & s \neq l \end{cases} \quad (22)$$

式中: $s \neq l$ 表示 $s=1, 2, \dots, n, s \neq l$; λ_l 为与第 l 模态向量对应的特征值。事实上,通常只需取前几阶特征向量就能取得较好的精度。将式(21)、式(22)代入式(20)即可得 $\partial \boldsymbol{\psi}_r^e / \partial b_k$, 利用式(18)即可得到结构系统各阶应变模态的上下界。至此便得到了含区间参数结构各阶应变模态向量的区间,即

$$(\boldsymbol{\psi}_r^e)^l = [\underline{\boldsymbol{\psi}}_r^e, \overline{\boldsymbol{\psi}}_r^e] = \left[\boldsymbol{\psi}_r^e(\mathbf{b}^c) - \sum_{k=1}^p \left| \frac{\partial \boldsymbol{\psi}_r^e(\mathbf{b}^c)}{\partial b_k} \right| \Delta b_k, \boldsymbol{\psi}_r^e(\mathbf{b}^c) + \sum_{k=1}^p \left| \frac{\partial \boldsymbol{\psi}_r^e(\mathbf{b}^c)}{\partial b_k} \right| \Delta b_k \right] \quad (23)$$

利用本文方法分析不确定条件下结构应变模态的方法流程如图1所示。

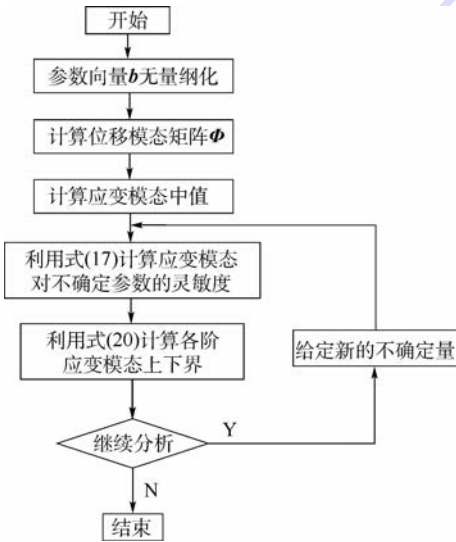


图1 考虑不确定性的应变模态分析流程

Fig. 1 Flowchart of uncertain analysis of strain modal

2.3 与概率方法的比较

下面基于概率方法对不确定结构应变模态进行分析,假定结构变量的总体统计参数已知,即

$$\begin{cases} E(\mathbf{b}) = E(b_k) = \mathbf{b}^c \\ \text{var}(\mathbf{b}) = \text{var}(b_k) = D(\mathbf{b}) = D(b_k) \\ \boldsymbol{\sigma}(\mathbf{b}) = \boldsymbol{\sigma}(b_k) = \sqrt{\text{var}(\mathbf{b})} = \sqrt{D(\mathbf{b})} = \sqrt{\text{var}(b_k)} = \sqrt{D(b_k)} \end{cases} \quad (24)$$

式中: $E(\mathbf{b})$ 、 $\text{var}(\mathbf{b})$ 和 $\boldsymbol{\sigma}(\mathbf{b})$ 分别为不确定参量的期望、方差与标准差。对式(17)两侧同时求期望,并注意到 $E(\delta b_j) = E(b_j - b_j^c) = 0$, 可得

$$E(\boldsymbol{\psi}_{ir}^e(\mathbf{b})) = E(\boldsymbol{\psi}_{ir}^e(\mathbf{b}^c)) + E\left(\sum_{k=1}^p \frac{\partial \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c)}{\partial b_k} \delta_k\right) = \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c) + \sum_{k=1}^p \frac{\partial \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c)}{\partial b_k} E(b_j - b_j^c) = \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c) \quad (25)$$

同样,对式(17)两边同时取方差,由于本文中考虑的各随机变量相互独立,协方差为零,则

$$\begin{aligned} \text{var}(\boldsymbol{\psi}_{ir}^e(\mathbf{b})) &= \sum_{k=1}^p \left(\frac{\partial \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c)}{\partial b_k} \right)^2 D(b_k) + \\ &\sum_{k=1}^m \sum_{l=1}^m \frac{\partial \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c)}{\partial b_k} \cdot \frac{\partial \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c)}{\partial b_l} \text{cov}(b_k, b_l) = \\ &\sum_{k=1}^p \left(\frac{\partial \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c)}{\partial b_k} \right)^2 D(b_k) = \sum_{k=1}^p \left(\frac{\partial \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c)}{\partial b_k} \sigma(b_k) \right)^2 \end{aligned} \quad (26)$$

式中: $\text{cov}(b_k, b_l)$ 为 b_k 与 b_l 的协方差, b_k 和 b_l 为不确定变量的第 k, l 个分量。

由此得到了应变模态响应的期望与方差,取距离应变模态响应均值 m 倍标准差处所确定的区域为应变模态概率解的区间,即有

$$\begin{cases} \overline{\boldsymbol{\varphi}}_{ir}^e(\mathbf{b}) = \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c) + m \sqrt{\sum_{k=1}^p \left(\frac{\partial \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c)}{\partial b_k} \sigma_k \right)^2} \\ \underline{\boldsymbol{\varphi}}_{ir}^e(\mathbf{b}) = \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c) - m \sqrt{\sum_{k=1}^p \left(\frac{\partial \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c)}{\partial b_k} \sigma_k \right)^2} \end{cases} \quad (27)$$

由 Chebyshev 不等式可知,具有有限方差的随机应变模态落在式(27)所确定范围内的概率不小于 $1 - 1/m^2$ 。此时,对应的不确定量范围下由区间方法求得的应变模态的相应区间上下界为

$$\begin{cases} \overline{\boldsymbol{\psi}}_{ir}^e = \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c) + \sum_{k=1}^p \left| \frac{\partial \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c)}{\partial b_k} \right| m \sigma_k \\ \underline{\boldsymbol{\psi}}_{ir}^e = \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c) - \sum_{k=1}^p \left| \frac{\partial \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c)}{\partial b_k} \right| m \sigma_k \end{cases} \quad (28)$$

考虑到对于非负 x_i , 有 $\sum_{i=1}^m x_i \geq \sqrt{\sum_{i=1}^m x_i^2}$, 对比式(27)与式(28)可知:

$$\begin{aligned} \sum_{k=1}^p \left| \frac{\partial \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c)}{\partial b_k} \right| m \sigma_k &= \sum_{k=1}^p \left| \frac{\partial \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c)}{\partial b_k} m \sigma_k \right| \geq \\ &\sqrt{\sum_{k=1}^p \left| \frac{\partial \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c)}{\partial b_k} m \sigma_k \right|^2} = m \sqrt{\sum_{k=1}^p \left(\frac{\partial \boldsymbol{\psi}_{ir}^e(\mathbf{b}^c)}{\partial b_k} \sigma_k \right)^2} \end{aligned} \quad (29)$$

综合式(27)~式(29)可知,在通过概率信息确定的不确定参数区间向量的条件下,由区间方法所得的应变模态区间比概率方法得到的要宽,这通常会使得结果趋于保守,因此当不确定参数的统计信息能够准确获得时,还是采用概率方法,而当不确定量统计信息难于获得时,采用区间估计方法是较为合理的。

3 数值算例

考虑损伤梁结构的应变模态,梁两端简支,结构几何尺寸及梁截面形状如图2所示,以梁轴线

方向为 x 轴,梁横截面所在平面为 yOz 平面。梁的材料为 45Cr 合金钢,材料的弹性模量 E 、密度 ρ 、泊松比 μ 等为不确定参数,且认为其统计信息未知,基于区间思想,上述参数可表示为 $E \in E^I = [1 - \beta_1, 1 + \beta_1] E_c$, $\rho \in \rho^I = [1 - \beta_2, 1 + \beta_2] \rho_c$, $\mu \in \mu^I = [1 - \beta_3, 1 + \beta_3] \mu_c$, 其中 $E_c = 2.06 \times 10^{11}$ Pa、 $\rho_c = 7.90 \times 10^3$ kg/m³ 和 $\mu_c = 0.29$ 分别为上述各参数的名义值, β_1 、 β_2 和 β_3 分别为各参数的不确定量,相互独立。为验证方法,算例中将不确定参数均视为区间数,且变异系数 β 分别取了 2%、4% 和 6% 等不同值。在实际应用中,材料参数取值区间可通过手册、规范等给出的物理性能范围来确定,几何、尺寸的区间可通过加工制造的公差来具体选取。分别在梁的 1/4 跨以及跨中处给定宽度为 1 mm 的不同深度的缺陷来模拟结构损伤,定义裂纹深度与截面原始高度的比值为损伤程度。实际中能够较方便准确测出的通常是正应变,因此以 x 方向的应变分量为识别参量。

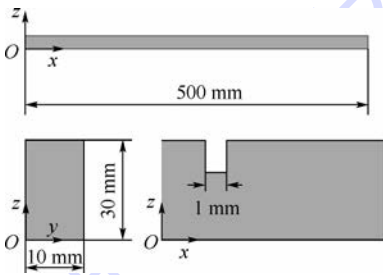


图2 梁结构示意图

Fig. 2 Diagrammatic sketch of beam structure

图3为结构跨中存在不同程度损伤时,结构各阶应变模态的情况。图中横坐标表示沿梁轴线方向的位置,纵坐标为对应的应变,本文中应变模态向量的具体值按质量矩阵进行归一化。

不难看出,由于结构的损伤,应变模态在损伤附近区域会发生突变,且损伤量增加时,应变模态突变量也随之增大。计算发现,结构在完好时和 5% 损伤时一阶固有频率分别为 288.523 Hz 与 288.064 Hz,二者相差约 0.15%,这表明结构小损伤主要影响结构在损伤处局部的参量,对总体特性影响不大,这也是基于固有频率等指标的方法存在缺陷的重要原因。在离损伤位置稍远处,损伤结构的应变与完好结构的相差很小,这说明在实际测试时若应变传感器距损伤位置较远,识别会相对困难,这可以通过合理设置传感器来克服,实际应用中,传感器间距为跨长的 6% 左右时即能较好识别梁的损伤区域^[20],且随着光纤光栅传感器等新型传感器的发展与应用,这一问题将能得到更好的解决。此外,当损伤位于结构某阶应

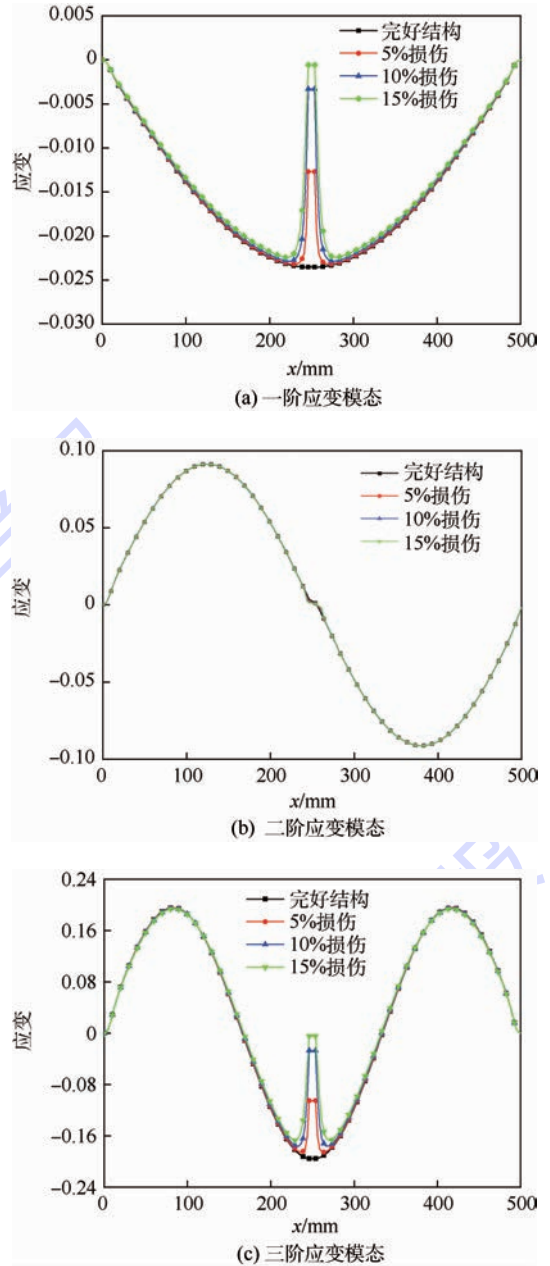


图3 结构跨中不同损伤程度下各阶应变模态

Fig. 3 Various order strain modal under different damage levels at midspan

变模态中应变值较大处(高应变区)时,该阶应变模态对损伤很敏感,能够识别结构较小的损伤;但当损伤位于应变值较小处尤其是模态节点附近时(见图 3(b)),应变模态对损伤很不敏感,难于识别损伤,好在实际应用中可以采用多阶模态向量的组合,因此基本都能较好地识别结构损伤。图4反映的是梁跨中存在 5% 损伤的情况下,结构参数具有不同程度的不确定量时,结构的各阶应变模态的上下界(为方便作图显示,取 $\beta = \beta_1 = \beta_2 = \beta_3$)。

图4为结构跨中 5% 损伤时不同不确定量下结构各阶应变模态的上下界。由图4可以发现,

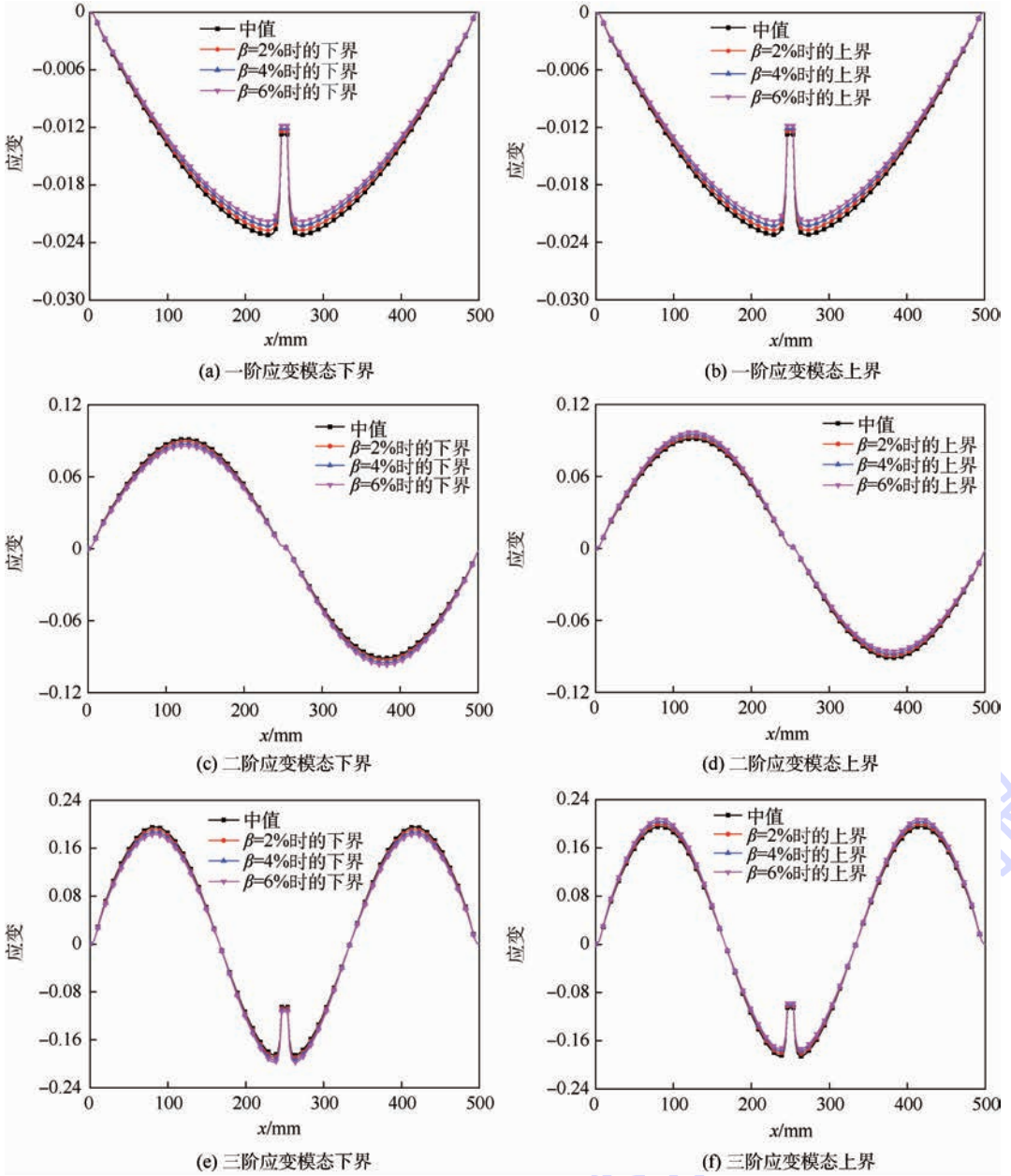


图4 结构跨中5%损伤时不同不确定量下结构各阶应变模态的上下界

Fig. 4 Upper and lower bounds of various order strain modal under different uncertainties with 5% damage at midspan

结构参数存在不同程度的不确定性时,应变模态也会有不同程度的变化。与损伤程度不同所引起的应变模态的变化不同,结构参数不确定对系统应变模态的影响主要表现为总体的影响,而对损伤处局部的应变模态影响相对较小。这是由于结构在损伤附近的应变分布主要受损伤本身控制,而对结构密度、弹性模量和泊松比等总体性参数不够敏感,因此这些参数的不确定性主要会引起结构应变模态整体性的变化。另外,由图4还可以发现模态振型中应变较大的位置,由参数不确定性所引起的应变模态的变化量也越大,对于应变较小的低应变区尤其是振型节点附近,应变值受不确定参数的影响较小。图5反映的是梁结构在1/4跨处存在

5%损伤的情况下,结构参数具有不同程度的不确定量时,结构的各阶应变模态的上下界。

在1/4跨损伤工况下,结构应变模态同样在损伤处发生了突变,同时结构参数不确定量对应变模态的影响与跨中损伤时类似。

在工程应用中,损伤指标不可避免地会受到噪声的影响,本文采用如下的方法计算含噪声应变模态振型^[17],即 $\tilde{\psi}_{ir}^e = \psi_{ir}^e (1 + \text{randn} \times \rho_N)$,其中: ψ_{ir}^e 和 $\tilde{\psi}_{ir}^e$ 分别为考虑噪声前后的应变模态; randn 为均值为0、标准差为1的高斯白噪声; ρ_N 代表噪声水平。将考虑噪声后的应变模态与完好结构的应变模态对比,计算出各点处的应变模态变化率,若变化率最大处出现在实际损伤处,则认为识别

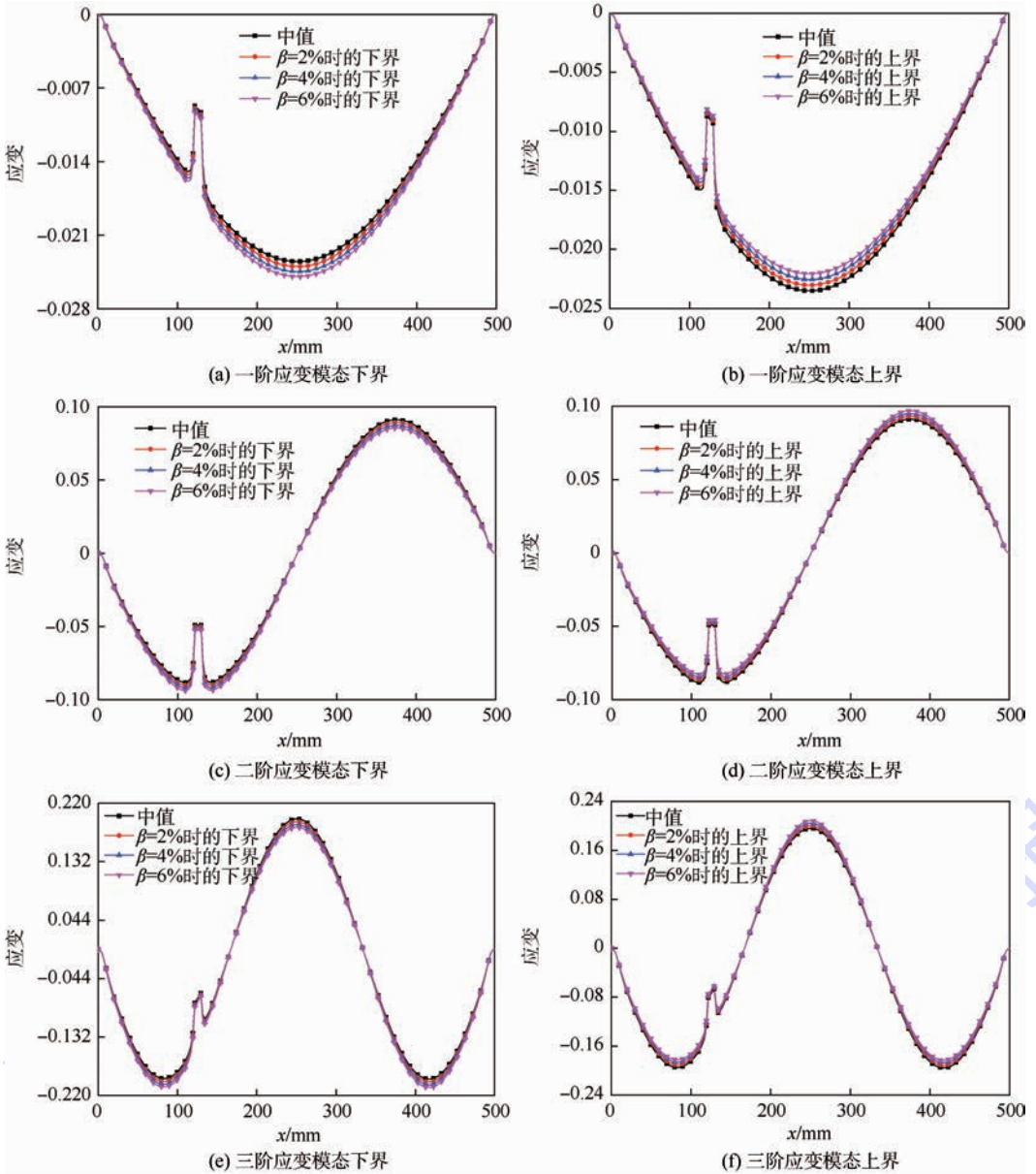


图5 1/4跨5%损伤时不同不确定量下结构各阶应变模态的上下界

Fig. 5 Upper and lower bounds of various order strain modal under different uncertainties with 5% damage at quarter span

成功,重复这一过程 10^4 次,将识别成功的次数除以总次数并定义为识别率。图6给出了结构跨中不同程度损伤时,参数不确定量为2%的情况下,不同噪声水平下的损伤识别率。可以看出,结构损伤程度越大,相应识别的抗噪能力就越强,15%损伤时在2%~3%的噪声水平下仍有较好的识别率。整体来说不确定条件下识别率受噪声影响较大,如何提高识别的抗噪性能仍是后续需要解决的问题。综上对于不同损伤工况的分析表明,基于一阶 Taylor 展开的方法计算应变模态的范围,在结构参数的不确定量较小的情况下,能够较好地识别上述不确定梁结构的损伤情况。但当结构参数的不确定量较大而损伤程度又相对较小时,尤其对于较复杂的结构而言,不确定性、噪声

等对损伤识别有较大影响,甚至可能改变结果导向,这也是后续研究中需要考虑的问题。

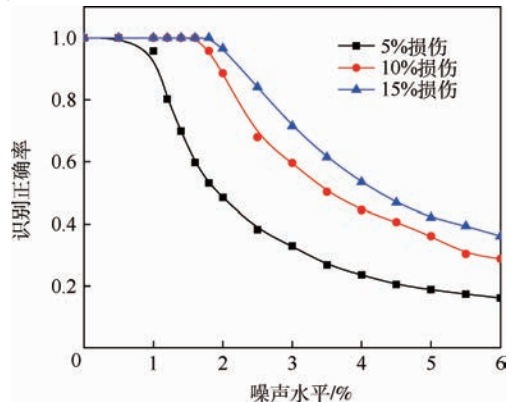


图6 不同噪声水平下的损伤识别正确率

Fig. 6 Correct damage identification rate under different noise levels

4 结 论

1) 应变模态对结构局部损伤较敏感,对于梁结构来说,只要传感器布置合理,该指标能很好地识别结构较小的损伤。

2) 若损伤位于某阶模态的低应变区,该阶模态对损伤不太敏感,实际应用时可采用综合多阶应变模态的方法对损伤进行识别以克服这一问题。

3) 不确定参数主要影响结构特性的总体数值,就本文算例来看,当参数的不确定量小于损伤程度或与损伤程度相当时,不确定性对结构损伤识别结果影响不大;当噪声水平低于 2% ~ 3% 时,本文方法有着较好的识别率,但噪声对识别指标有着较明显的影响,这也是后续研究中需重点考虑的问题。

参考文献 (References)

- [1] DOEBLING S W, FARRAR C R, PRIME M B. A summary review of vibration-based damage identification methods [J]. Shock and Vibration Digest, 1998, 30(2) : 91-105.
- [2] LIU Y, NAYAK S. Structural health monitoring: State of the art and perspectives [J]. Journal of the Minerals, Metals and Materials Society, 2012, 64(7) : 789-792.
- [3] 郑栋梁, 李中付, 华宏星. 结构早期损伤识别技术的现状和发展趋势 [J]. 振动与冲击, 2002, 21(2) : 1-6.
ZHENG D L, LI Z F, HUA H X. A summary review of structural initial damage identification methods [J]. Journal of Vibration and Shock, 2002, 21(2) : 1-6 (in Chinese).
- [4] SALAWU O S. Detection of structural damage through changes in frequency: A review [J]. Engineering Structures, 1997, 19(9) : 718-723.
- [5] OU J, LI H. Structural health monitoring in mainland China: Review and future trends [J]. Structural Health Monitoring, 2010, 9(3) : 219-231.
- [6] 李德葆, 陆秋海. 实验模态分析及其应用 [M]. 北京: 科学出版社, 2001 : 29-36.
LI D B, LU Q H. Experimental modal analysis and its application [M]. Beijing: Science Press, 2001 : 29-36 (in Chinese).
- [7] CHONDROS T G, DIMAROGONAS A D. Identification of cracks in welded joints of complex structures [J]. Journal of Sound and Vibration, 1980, 69(4) : 531-538.
- [8] CAWLEY P, ADAMS R D. The location of defects in structures from measurements of natural frequencies [J]. The Journal of Strain Analysis for Engineering Design, 1979, 14(2) : 49-57.
- [9] ADAMS R D, CAWLEY P, PYE C J, et al. A vibration technique for non-destructively assessing the integrity of structures [J]. Journal of Mechanical Engineering Science, 1978, 20(2) : 93-100.
- [10] ZHU H P, HE B, CHEN X Q. Detection of structural damage through changes in frequency [J]. Wuhan University Journal of Natural Sciences, 2005, 10(6) : 1069-1073.
- [11] 杜思义, 殷学钢, 陈淮. 基于频率二阶摄动的结构损伤识别方法 [J]. 应用力学学报, 2007, 23(4) : 613-617.
DU S Y, YIN X G, CHEN H. Structural damage identification based on second-order perturbation of frequency [J]. Chinese Journal of Applied Mechanics, 2007, 23(4) : 613-617 (in Chinese).
- [12] YI T H, LI H N, GU M. Optimal sensor placement for structural health monitoring based on multiple optimization strategies [J]. Structural Design of Tall and Special Buildings, 2011, 20(7) : 881-900.
- [13] ALLEMANG R J. The modal assurance criterion—twenty years of use and abuse [J]. Sound and Vibration, 2003, 37(8) : 14-23.
- [14] 董聪, 丁辉. 结构损伤识别和定位的基本原理与方法 [J]. 中国铁道科学, 1999, 20(3) : 89-94.
DONG C, DING H. The basic principle and method for recognition and location of structural damage [J]. China Railway Science, 1999, 20(3) : 89-94 (in Chinese).
- [15] QIU Z P, CHEN S H, ELISHAKOFF I. Natural frequencies of structures with uncertain but nonrandom parameters [J]. Journal of Optimization Theory and Applications, 1995, 86(3) : 669-683.
- [16] 宗周红, 牛杰, 王浩. 基于模型确认的结构概率损伤识别方法研究进展 [J]. 土木工程学报, 2012, 45(8) : 121-130.
ZONG Z H, NIU J, WANG H. A review of structural damage identification methods based on finite element model validation [J]. China Civil Engineering Journal, 2012, 45(8) : 121-130 (in Chinese).
- [17] 林秀萍. 基于概率统计方法的结构损伤识别研究 [D]. 重庆: 重庆大学, 2009 : 15-20.
LIN X P. The studies on structural damage detection based on probability statistics method [D]. Chongqing: Chongqing University, 2009 : 15-20 (in Chinese).
- [18] PALMA R, RUS G, GALLEGRO R. Probabilistic inverse problem and system uncertainties for damage detection in piezoelectrics [J]. Mechanics of Materials, 2009, 41(9) : 1000-1016.
- [19] 村井秀晃, 伊藤博幸, 内海荣一. ダイナミックデザインアナリシスの研究(パート2) [J]. コマツ技報, 1979, 25(2) : 73-82.
MURAI S, ITO H, UTSUMI E. Study of dynamic design analysis (Part 2) [J]. Komatsu Technical Report, 1979, 25(2) : 73-82 (in Japanese).
- [20] 顾培英. 基于应变模态技术的结构损伤诊断直接指标法研究 [D]. 南京: 河海大学, 2006 : 28-33.
GU P Y. Study on structural damage diagnosis method based on direct indicators of strain mode technology [D]. Nanjing: Hohai University, 2006 : 28-33 (in Chinese).

作者简介:

杨君坦 男, 硕士研究生。主要研究方向: 计算固体力学、智能材料、结构动力学、结构健康监测。

E-mail: yangjuntan@buaa.edu.cn

邱志平 男, 博士, 教授, 博士生导师。主要研究方向: 计算力学、不确定性、结构动力学、凸模型理论及其结构可靠性。

Tel. : 010-82339628

E-mail: zpqiu@buaa.edu.cn

Damage detection of structures with interval uncertain parameters

YANG Juntan¹, QIU Zhiping^{1,*}, LI Qi²

(1. School of Aeronautic Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;

2. State Key Laboratory of Robotics, Shenyang Institute of Automation Chinese Academy of Sciences, Shenyang 110016, China)

Abstract: To deal with structures with uncertain parameter which often lacks sufficient statistical information in damage detection, the uncertain parameters were modeled as interval numbers. The structural strain modal testing method was introduced using finite element method, and the strain modal was further expressed as the function of the interval parameter. Combined with the first order Taylor expansion and interval analysis, the range of each strain modal under parameter uncertainty is estimated, and thus we obtained an approach for structural damage detection considering uncertainty. Comparison between the proposed method and the traditional probabilistic method was conducted. Numerical simulations of a simply supported beam under different damage were performed by investigating the change of strain modes due to different damage levels and uncertainty levels. The influence rules of damage and uncertainty on strain modal were obtained, and a brief discussion of the impact of noise on damage detection was performed. The work of this paper provides a new technique and thought for damage detection of uncertain structures.

Key words: damage detection; uncertain parameter; interval analysis; finite element method; mode analysis

Received: 2015-06-17; **Accepted:** 2016-01-08; **Published online:** 2016-03-11 10:35

URL: www.cnki.net/kcms/detail/11.2625.V.20160311.1035.001.html

Foundation items: National Natural Science Foundation of China (11372025, 11432002); Defense Industrial Technology Development Program (A0820132001); "111" Project(B07009)

* **Corresponding author.** Tel.: 010-82339628 E-mail: zpqiu@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0458

基于诱导航线的多无人机编队飞行控制方法

吴俊成¹, 周锐^{1,*}, 董卓宁¹, 车军²

(1. 北京航空航天大学 控制一体化技术国家级科技重点实验室, 北京 100083; 2. 中航工业西安飞行自动控制研究所, 西安 710065)

摘 要: 针对无人机编队飞行控制问题, 提出了一种基于诱导航线的协同控制方法, 僚机根据当前相对于期望位置的误差和长机的飞行状态生成诱导航线并进行跟踪; 针对飞行过程中的突发障碍采用改变编队队形的方法进行规避; 针对无人机之间可能出现的碰撞情况, 根据无人机到达碰撞点的时间以及位置分别从高度以及航向两方面进行躲避, 仿真结果表明: 采用此方法无人机能够保持稳定的编队, 并同时能够躲避飞行过程中的障碍以及避免飞机之间的碰撞。

关 键 词: 无人机(UAV); 编队飞行; 诱导航线; 避障; 避撞

中图分类号: V279

文献标识码: A

文章编号: 1001-5965(2016)07-1518-08



近十多年来, 随着计算、通信和传感器等技术的不断发展, 无人机在民用和军事领域发挥着日益重要的作用。通过通信和协调, 多无人机可以组成编队, 以协同的方式执行任务, 在复杂战场环境中, 无人机编队在容错性、可靠性和协作性上比单架无人机具有明显优势。因此无人机编队协同控制越来越受到学术界的重视。

针对无人机编队飞行控制问题, Leader-Follower 法是最经典的一种编队控制方法, 最早见于移动机器人编队控制的文献^[1]中, 随后在无人机编队控制中得到了广泛的应用^[2-4]。该方法实现起来比较简单, 但编队稳定性不够而且误差会逐级放大。虚拟结构法引入刚体运动的概念^[5-7], 将刚体中的各个顶点作为每架无人机的位置。文献[8-10]采用模型预测控制方法对编队进行控制, 具有很好的适应性, 可用于动态变化的编队控制问题, 但是模型复杂、计算量较大。除此之外, 人工势场法^[11]和基于行为法^[12]也经常用于编队控制中, 但控制精度较差。

本文在 Leader-Follower 法的基础上提出了一种基于诱导航线的编队飞行控制方法, 并考虑飞行过程中出现的突发障碍以及编队内飞机的碰撞问题^[13-14]。

1 问题描述

1.1 无人机运动模型

将无人机视为质点, 其在三维平面内的运动模型为

$$\begin{cases} \dot{x} = v \cos \varphi \cos \gamma \\ \dot{y} = v \sin \varphi \cos \gamma \\ \dot{z} = v \sin \gamma \\ \dot{v} = u_v \\ \dot{\varphi} = u_\varphi \\ \dot{\gamma} = u_\gamma \end{cases} \quad (1)$$

式中: x 、 y 和 z 分别代表无人机的经度、纬度和高度; v 为飞机空速; φ 和 γ 分别为无人机的航迹偏航角和俯仰角; $\mathbf{u} = [u_v \quad u_\varphi \quad u_\gamma]^T$ 可以作为无人机的控制输入, 即速度指令、航迹偏航角指令和航

收稿日期: 2015-07-08; 录用日期: 2015-09-25; 网络出版时间: 2015-10-30 17:04

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151030.1704.011.html

基金项目: 国家自然科学基金(61273349, 61175109, 61203223); 航空科学基金(2013ZA18001, 2014ZA18004)

* 通讯作者: Tel.: 010-82339232 E-mail: zhr@buaa.edu.cn

引用格式: 吴俊成, 周锐, 董卓宁, 等. 基于诱导航线的多无人机编队飞行控制方法[J]. 北京航空航天大学学报, 2016, 42(7): 1518-1525. WU J C, ZHOU R, DONG Z N, et al. Formation flight control method of multiple UAVs based on guidance route[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1518-1525 (in Chinese).

迹俯仰角指令。

使用质点模型的同时一般应该加入适当的约束条件,来对模型进行限制:

$$\begin{cases} 0 < v_{\min} \leq v \leq v_{\max} \\ a_{\min} \leq \dot{v} = a \leq a_{\max} \\ |\dot{\varphi}| = |\omega| \leq \omega_{\max} \end{cases} \quad (2)$$

式中: v_{\min} 和 v_{\max} 分别为飞机最小和最大速度限制; a_{\min} 和 a_{\max} 分别为飞机最小和最大加速度限制; a 为飞机加速度; ω 为飞机偏航角速度; ω_{\max} 为飞机最大偏航角速度。

1.2 编队队形

编队队形给出了长机以及各僚机在编队飞行时需要保持的相对位置关系,包括前向和侧向距离,具体表示在长机固联方位角等于长机偏航角的导航坐标系中,如图1所示。

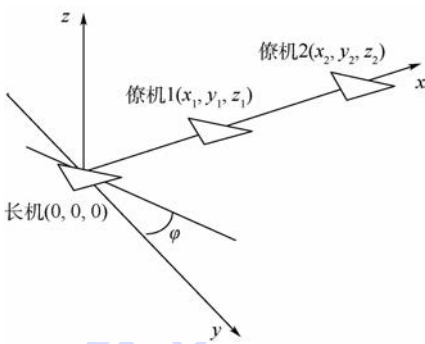


图1 编队坐标系

Fig.1 Coordinate system of formation

1.3 系统框图

每架无人机的飞行控制系统框图如图2所示。

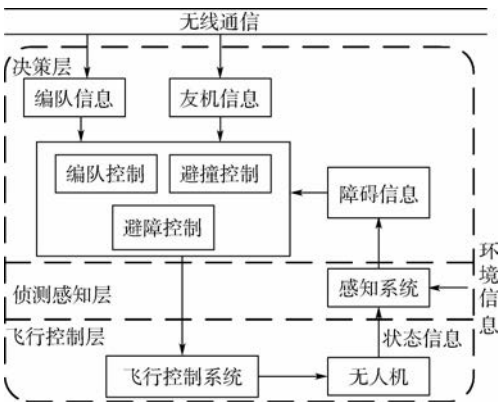


图2 系统框图

Fig.2 System block diagram

整个系统框图分为3层,分别为决策层、侦测感知层和飞行控制层。无人机通过无线通信获得编队信息以及编队中其他飞机信息,通过感知系统获得环境信息进一步获得是否有障碍信息。根

据获得的信息进行编队控制、避障控制以及避撞控制,得出相应控制量,将得出的控制量传到飞行控制系统也就是自动驾驶仪,然后通过飞行控制系统完成对无人机的控制。

2 基于诱导航线的编队飞行控制

形成编队后,长机以期望的速度按照给定的航线飞行,僚机根据相对长机的位置进行编队保持控制,具体是根据当前相对于期望位置的误差和长机的飞行状态生成诱导航线,并跟踪该航线,同时根据长机的速度以及自身相对期望位置的纵向误差调节飞机速度。

2.1 诱导航线生成原理

当僚机跟踪长机的时候采用诱导航线控制方式,根据当前僚机相对期望位置的误差生成一条诱导航线,然后利用航线跟踪方法控制僚机。该控制方式根据位置误差的大小切换不同的航线生成方法,切换航线控制方法的原理就是:当远离期望位置的时候朝期望位置飞行,当处于期望位置附近时,沿经过僚机期望位置并与长机速度方向平行的诱导航线飞行。

航线切换由僚机相对期望位置的误差大小确定,切换范围采用继电器开关线形式,切换范围的大小根据飞机转弯能力设定,如图3所示。

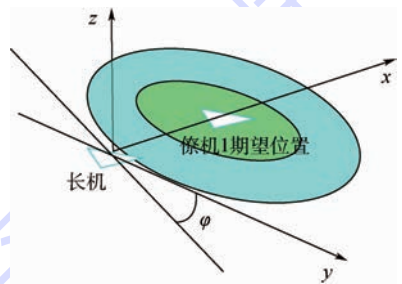


图3 航线切换范围

Fig.3 Range of route switching

较大椭圆的外边界表示从平行诱导航线切换到朝向期望位置诱导航线的切换线,在平行诱导的状态下一旦僚机从较大椭圆内穿出就进行切换。

较小椭圆的外边界表示从朝向期望位置诱导航线切换到平行诱导航线的切换线,在朝向期望位置诱导的状态下一旦僚机从外部穿入较小椭圆就切换。

朝向期望位置诱导航线的生成方法是:在每一个控制周期,以僚机当前位置为起点、以僚机期望位置为终点生成诱导航线,如图4所示。

平行诱导航线的生成方法是:在每一个控制

周期,以僚机期望位置为起点,将过该点平行于长机航向的射线作为诱导航线,如图5所示。

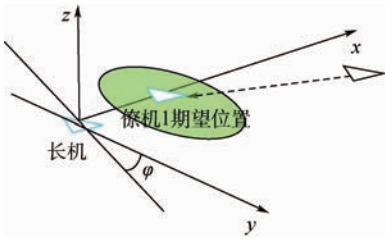


图4 朝向诱导航线

Fig.4 Directing guidance route

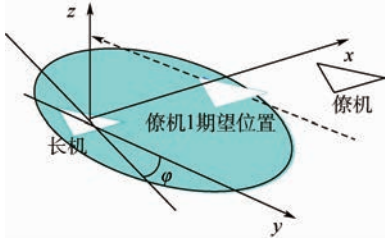


图5 平行诱导航线

Fig.5 Parallel guidance route

僚机跟踪诱导航线的控制方法与跟踪普通航线的方法完全相同。

2.2 诱导航线生成

下面为诱导航线上2个点的计算公式,这里以僚机1为例,僚机1的实际位置为 (x_1, y_1, h_1) ,速度和偏航角为 (v_1, φ_1) ,长机的实际位置为 (x_L, y_L, h_L) ,速度和偏航角为 (v_L, φ_L) 。

$$\begin{cases} x_{11} = x_1 + d_k(x_e - x_1) - \delta_k \sin \varphi_L \\ y_{11} = y_1 + d_k(y_e - y_1) + \delta_k \cos \varphi_L \\ h_{11} = h_1 \\ v_{11} = v_1 \end{cases} \quad (3)$$

$$\begin{cases} x_{12} = x_{11} - \delta_k \sin \varphi_L \\ y_{12} = y_{11} + \delta_k \cos \varphi_L \\ h_{12} = h_e \\ v_{12} = v_e \end{cases} \quad (4)$$

式中: δ_k 为固定的距离;参数 d_k 决定诱导航线位置; h_e 和 v_e 分别为僚机1的期望高度和期望速度; x_e 和 y_e 为僚机1的期望位置:

$$\begin{cases} x_e = x_L - D_F \cos \varphi_L + D_W \sin \varphi_L \\ y_e = y_L + D_F \sin \varphi_L + D_W \cos \varphi_L \end{cases} \quad (5)$$

其中: D_F 和 D_W 分别为僚机在编队中相对于长机的前向距离和侧向距离。

诱导航线产生过程如图6所示。由式(3)可知,诱导航线的位置主要由参数 d_k 决定, d_k 为一个从0到1的变量,根据飞机的滚转能力设定。当 $d_k=0$ 时,以僚机实际位置为基准、以长机飞行方向为航向产生诱导航线;当 $d_k=1$ 时,以僚机期

望位置为基准、以长机飞行方向为航向产生诱导航线。在飞行过程中不断改变 d_k ,使诱导航线逐渐接近僚机的期望位置,僚机通过跟随诱导航线逐渐从实际位置到期望位置移动,最终使得无人机到达期望的位置。

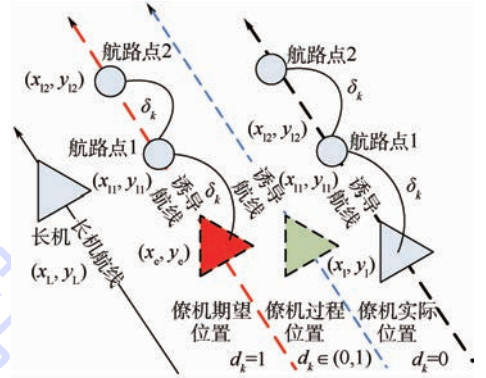


图6 诱导航线产生过程

Fig.6 Production process of guidance route

2.3 僚机诱导航线跟踪

僚机跟踪诱导航线时,主要涉及3个关键变量:期望速度、航向以及侧偏。

期望速度与长僚机在诱导航线上的投影距离(d)、长机速度分量(v_{Ld})和长僚机速度分量(v_{Lg})差有关。

$$v_e = k_1 d + v_{Ld} + k_2 v_{Lg} \quad (6)$$

式中: k_1 和 k_2 为比例系数,与飞机的调速能力有关。

长僚机在诱导航线上的投影距离:

$$d = (y_e - y_1) \cos \varphi_L + (x_e - x_1) \sin(-\varphi_L) \quad (7)$$

长机速度分量:

$$v_{Ld} = v_L \cos(\varphi_L - \varphi_d) \quad (8)$$

长僚机速度分量差:

$$v_{Lg} = v_L \cos(\varphi_L - \varphi_d) - v_1 \cos(\varphi_1 - \varphi_d) \quad (9)$$

式中:

$$\varphi_d = \begin{cases} \arctan(x_1 - x_e, y_e - y_1) & d_e > d_0 \\ \varphi_L & d_e \leq d_0 \end{cases} \quad (10)$$

其中: d_e 为僚机实际位置与期望位置的距离; d_0 为切换范围。可见,长僚机在诱导航线上的投影距离、长机速度分量和长僚机速度分量差越大,则期望速度就越大。

当前航向与期望航向的夹角:

$$\varphi_e = \varphi_1 - \arctan \frac{y_{12} - y_{11}}{x_{12} - x_{11}} \quad (11)$$

侧偏:僚机当前点到诱导航线的有向距离。

2.4 编队避障

无人机编队在飞行过程中面对突发威胁时,要及时反应,在保证顺利躲避障碍的前提下,尽量

减小整个编队的机动动作。

避障流程图如图7所示。

无人机检测到障碍物后,根据当前的相对位置以及障碍物的位置计算无人机恰好绕过障碍物的距离,同时考虑改变位置后与其他无人机是否有碰撞冲突;如果有,计算当前位置再绕过其他无人机的距离,此时还要考虑尽量减小整个编队机动动作。将2次计算的距离作用到无人机原来编队位置上,从而改变编队队形顺利通过障碍。

为了便于说明算法流程,现举例说明。当飞机在躲避障碍时,与其他飞机存在碰撞可能时,如图8所示。

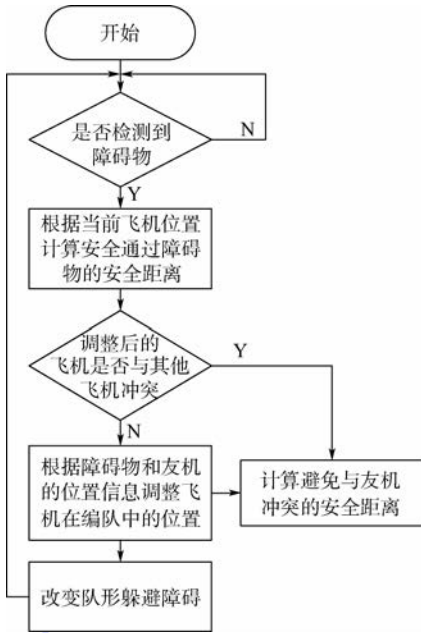


图7 避障算法流程图

Fig.7 Flowchart of obstacle avoidance algorithm

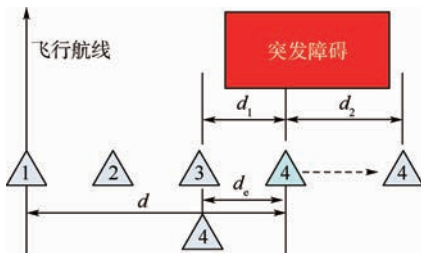


图8 编队避障

Fig.8 Formation of obstacle avoidance

当 $d_1 < d_2$ 时,飞机从左边避障,但会与3号飞机碰撞,会造成编队的扰动,因此需要从右边避障。如果 $d_1 \ll d_2$,4号无人机从右边要飞一个很远的距离,此时需要从左边避障,在改变侧向距离的同时改变前向距离,防止与3号碰撞。

各飞机改变后的相对位置为

$$\begin{cases} D'_F = D_F + d_f \\ D'_W = D_W + d_e \end{cases} \quad (12)$$

式中: d_f 和 d_e 分别为僚机编队中的前向和横向距离。

2.5 编队算法的自动驾驶仪控制

飞机的航向控制,根据飞机距航线的侧偏距离和当前航向差:

$$u_\psi = k_d d + k_\varphi \varphi_e \quad (13)$$

式中: k_d 和 k_φ 为控制参数。

飞机的速度采用PI控制:

$$u_v = K_p (e_v + \frac{1}{T} \int_0^t e_v dt) \quad (14)$$

式中: e_v 为期望速度与当前速度的偏差; K_p 为比例系数; T 为时间。

飞机的高度控制为

$$u_\gamma = k_h (H_e - H) \quad (15)$$

式中: k_h 为比例系数; H_e 为期望高度; H 为高度。

限制条件为

$$\begin{cases} u_{\varphi_{\min}} \leq u_\varphi \leq u_{\varphi_{\max}} \\ u_{v_{\min}} \leq u_v \leq u_{v_{\max}} \\ u_{\gamma_{\min}} \leq u_\gamma \leq u_{\gamma_{\max}} \end{cases} \quad (16)$$

3 编队防撞算法

无人机在编队飞行过程中最为重要的就是防止飞机之间的碰撞,针对这个问题,本文根据无人机的空间位置设计了基于碰撞优先级的防撞算法,算法流程图如图9所示。

防撞算法作用于整个编队飞行过程中,每架无人机都实时获得其他无人机信息,并预测它们可能的碰撞点,根据当前位置与碰撞点位置设计防撞算法,在避障的同时还要考虑优先级问题。

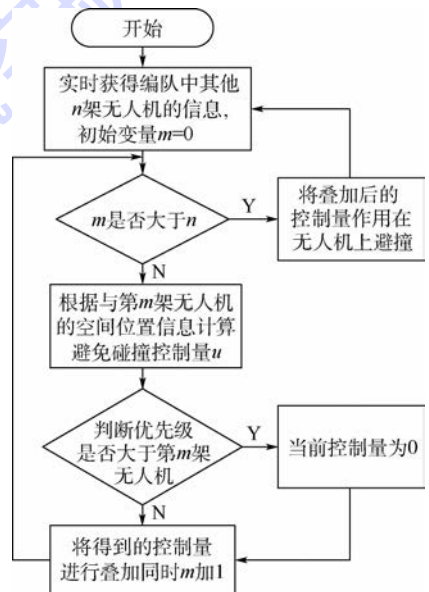


图9 防撞算法流程图

Fig.9 Flowchart of collision avoidance algorithm

3.1 防撞算法设计

避撞时根据飞机在空间中的速度向量以及相对位置(见图 10),分别从高度和航向两方面进行躲避^[15]。其中:

$$e_{ij} = \frac{v_i \times v_j}{\|v_i \times v_j\|} \quad (17)$$

式中: v_i 和 v_j 为 2 架无人机的速度向量; e_{ij} 为方向向量。

根据飞机在碰撞点的垂直距离以及到达碰撞点的时间设计避免碰撞的控制量:

$$u_i = \frac{1}{k_i t_i^2 + k_{ij}(t_i - t_j)^2 + k_{dc} d_{ij}^2} \quad (18)$$

式中: d_{ij} 为 2 架无人机的高度差; t_i 和 t_j 为到达碰撞点的时间; k_i 、 k_{ij} 和 k_{dc} 为控制参数,与飞机的转向以及爬升能力有关。

可以看出随着两机的接近,时间逐渐变短,控制量随之增大,将控制量分别叠加到方向以及高度控制量上:

$$\begin{cases} u'_\psi = u_\psi + k_\psi u_i \\ u'_\gamma = u_\gamma + k_\gamma u_i \end{cases} \quad (19)$$

式中: k_ψ 为偏航控制参数; k_γ 为俯仰控制参数。

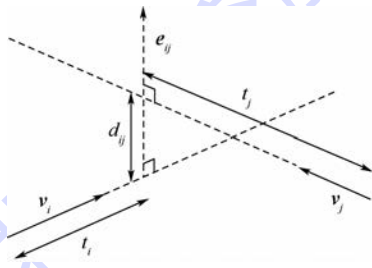


图 10 飞机之间的几何位置

Fig. 10 Geometric position of aircrafts

3.2 基于模糊逻辑的避撞优先级设计

当 2 架无人机有碰撞冲突时,双方都进行避让会造成较大的机动代价,给整个编队带来较大的扰动。因此在避让时要考虑避让优先级问题,高优先级的无人机可以不用避让,低优先级的无人机进行避让。无人机的避撞优先级由无人机的机动潜力、受损程度以及无人机的重要程度决定。

3.2.1 模糊集和隶属度函数的设计

根据无人机受损程度将受损度划分为 10 个等级,即论域为 $[1, 10]$,在此论域的基础上定义 5 个模糊集,分别为受损严重、较重、一般、较轻、轻微。5 个模糊集的隶属函数设计如图 11 所示。图中实线和虚线分别代表了不同模糊子集的隶属度函数,模糊子集在对应区间的隶属度为 1,与相邻区间的隶属度为一条斜线,底端与相邻区间中点连接,顶端与两区间交点连接。这样定义好处是论域上任意一点只隶属于一个模糊子集。

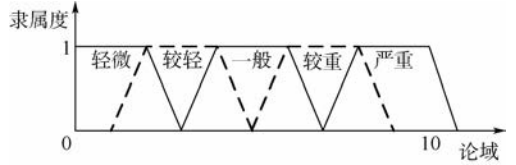


图 11 受损隶属度函数

Fig. 11 Damaged membership function

根据无人机的飞行状况以及避撞方向对无人机的机动潜力进行判断,所谓机动潜力指的是在无人机当前运动的基础上做避撞机动的能力。

例如当无人机当前正向右转弯时,这时检测到威胁在无人机的右方需要向左躲避,在右转弯的基础上向左转执行起来较困难,机动潜力较差。如果在右转弯的基础上向右躲避,执行起来比较容易,机动潜力较强,具体如图 12 所示。在图中侧滑角为 β_1 ,无人机向右转,威胁在左方。

根据侧滑角的大小以及威胁位置建立无人机机动潜力的模糊集。论域为无人机侧滑角的范围 $[\beta_{min}, \beta_{max}]$,在论域的基础上定义 5 个模糊集,分别为很差、差、一般、强、很强。隶属度函数设计与之前类似,如图 13 所示。

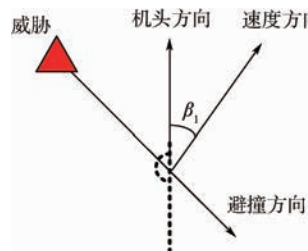


图 12 无人机机动方向

Fig. 12 UAV maneuvering direction

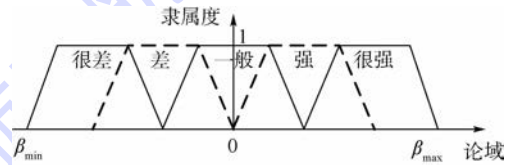


图 13 威胁隶属度函数

Fig. 13 Threat membership function

同理将优先级等级设为 10 个等级,即论域为 $[1, 10]$,在此论域的基础上定义 5 个模糊集,分别为很高、高、一般、低、很低,隶属度与之前设计类似,不再赘述。

3.2.2 模糊推理

模糊逻辑的控制规则形式为“IF-THEN”型,由于知识的前提条件是 2 条,属于复合条件,控制规则集如表 1 所示。

根据无人机受损程度以及当前机动潜力,利用表 1 中的模糊规则得到无人机的避撞优先级。

表 1 控制规则集
Table 1 Set of control rule

受损度	机动潜力				
	很强	强	一般	差	很差
轻微	很高	很高	高	一般	低
较轻	很高	高	高	低	低
一般	高	高	一般	很低	很低
较重	一般	低	很低	很低	很低
严重	低	低	很低	很低	很低

当具有碰撞冲突的 2 架无人机的避让优先级相等时,要根据无人机的战术重要程度进行再判断,例如长机重要程度要高于僚机,同等优先级的僚机要躲避长机。战术重要程度一般在任务开始前就要根据情况设定好,便于后续的应用。

4 仿真结果与分析

为了体现算法的有效性,以下所有仿真都采用了 F15 飞机模型。状态不同的 4 架无人机(UAV1 ~ UAV4)由分散状态形成编队状态,并在形成编队的同时做转弯机动,以横一队形飞行,无人机之间的间距为 300 m,飞行速度为 170 m/s,编队飞行轨迹如图 14 所示,编队飞行状态如图 15 所示。

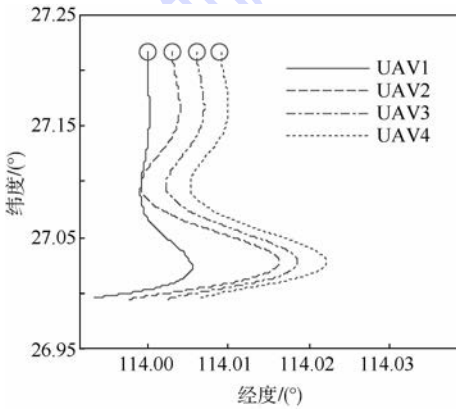


图 14 编队转弯飞行轨迹

Fig. 14 Flight path of formation turning

可以看出,4 架无人机的速度、高度以及偏航角在编队算法的作用下迅速收敛并保持稳定,以横一队形编队飞行。

在飞行过程中遇到的不是简单的障碍,而是一个狭长的山谷,不能容纳 2 架无人机并行飞过,此时在防撞算法的作用下,编队飞行轨迹如图 16 所示,编队飞行状态如图 17 所示。

从图 16 和图 17 中可以看出,横一编队为了躲避障碍,通过速度以及航向的调节依次绕过障碍,形成一个最利于通过障碍的纵一队形,并且各飞机之间保持一个相对安全距离,安全绕过障碍

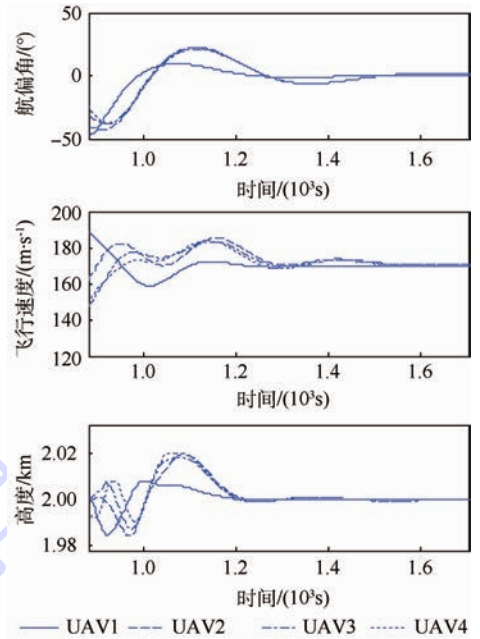


图 15 编队转弯飞行状态

Fig. 15 Flight condition of formation turning

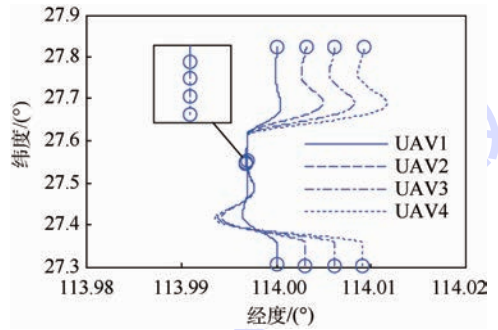


图 16 编队避障飞行轨迹

Fig. 16 Flight path of formation when avoiding obstacle

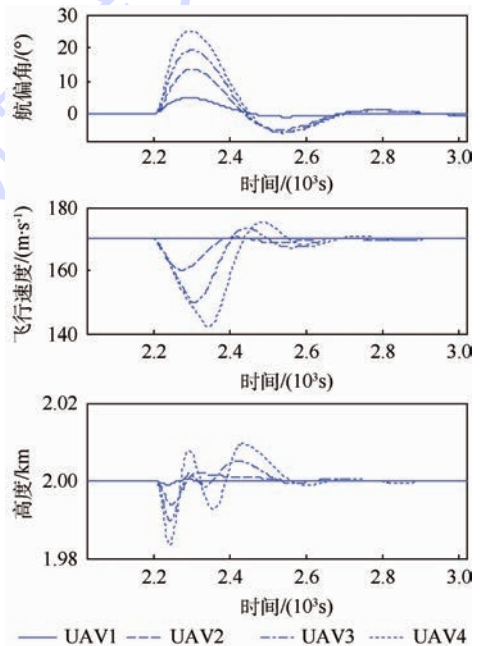


图 17 编队避障飞行状态

Fig. 17 Flight condition of formation when avoiding obstacle

后又重新形成横一队形。

在上述算例的飞行过程中,由于编队要进行机动和躲避障碍使得无人机之间具有碰撞的危险,以 UAV3 为例,在一次机动过程中与 UAV2 有碰撞冲突,根据模糊逻辑规则推理出 UAV2 的优先级要高,因此 UAV3 要进行主动避让。

为了体现避撞算法的有效性,给出未加碰撞消除和增加碰撞消除时的最近邻机距离和控制量变化,飞行状况如图 18 和图 19 所示。

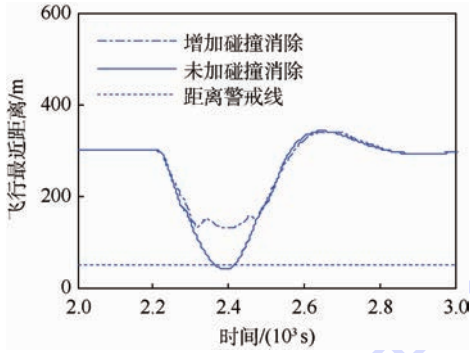


图 18 UAV3 最近邻机距离

Fig. 18 Distance between UAV3 and the nearest neighbor

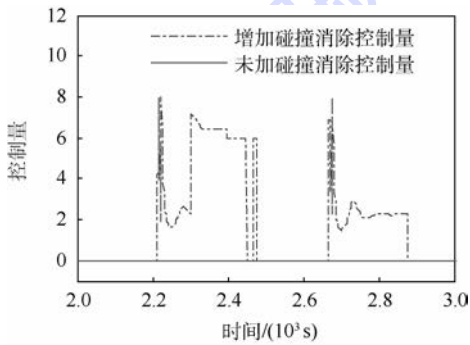


图 19 碰撞消除控制量

Fig. 19 Controlled quantity of collision elimination

从图 18 和图 19 可以看出,未加碰撞消除控制量时,最近邻机距离超过安全距离警戒线,具有碰撞危险,增加碰撞消除控制量之后,可以看出随着距离的接近控制量也随之起作用,控制无人机进行躲避,最终使最近距离保持在警戒距离之外。

5 结论

1) 建立了基于诱导航线的编队飞行控制方法,可以保证编队以指定队形稳定飞行。

2) 根据突发障碍建立规避模型,保证整个编队顺利规避,同时考虑规避过程中无人机之间的碰撞。

3) 建立了飞机之间的避碰模型,并基于模糊逻辑进行了避撞优先级设计,在避撞的同时尽量减小编队的扰动。

参考文献 (References)

- [1] WANG P K C. Navigation strategies for multiple autonomous mobile robots moving in formation[J]. *Journal of Robotic System*, 1991, 9(2): 177-195.
- [2] 丁衍,魏晨,鲍树语. 基于一致性算法的时延多无人机编队分散化控制[J]. *计算机应用*, 2014, 34(S1): 151-155.
DING Y, WEI C, BAO S Y. Formation control for multiple UAVs with time delay based on consensus algorithm[J]. *Journal of Computer Applications*, 2014, 34(S1): 151-155 (in Chinese).
- [3] WEITZ L A, HURTADO J E, SINCLAIR A J. Decentralized cooperative-control design for multivehicle formations[J]. *Journal of Guidance, Control, and Dynamics*, 2008, 31(4): 970-979.
- [4] KRISTIANSEN R, LORIA A, CHAILLET A, et al. Spacecraft relative rotation tracking without angular velocity measurements[J]. *Automatica*, 2009, 45(3): 750-756.
- [5] 邵壮,祝小平,周洲,等. 无人机编队机动飞行时的队形保持反馈控制[J]. *西北工业大学学报*, 2015, 33(1): 26-32.
SHAO Z, ZHU X P, ZHOU Z, et al. A formation keeping feedback control for formation flight of UAVs[J]. *Journal of Northwestern Polytechnical University*, 2015, 33(1): 26-32 (in Chinese).
- [6] REN W, SORENSEN N. Distributed coordination architecture for multi-robot formation control[J]. *Robotics and Autonomous Systems*, 2008, 56(4): 324-333.
- [7] REN W. Decentralization of virtual structure in formation control of multiple vehicle systems via consensus strategies[J]. *European Journal of Control*, 2008, 14(2): 1-11.
- [8] 华思亮,尤优,张红,等. 无人机编队的滚动时域控制[J]. *电光与控制*, 2012, 19(3): 1-5.
HUA S L, YOU Y, ZHANG H, et al. Receding horizon control of UAV formations[J]. *Electronics Optics & Control*, 2012, 19(3): 1-5 (in Chinese).
- [9] FAHIMI F. Non-linear model predictive formation control for groups of autonomous surface vessels[J]. *International Journal of Control*, 2007, 80(8): 1248-1259.
- [10] KEVICAKY T, BORRELLI F, FREGENE K, et al. Decentralized receding horizon control and coordination of autonomous vehicle formations[J]. *IEEE Transactions on Control System Technologies*, 2008, 16(1): 19-33.
- [11] KHATIB O. Real-time obstacle avoidance for manipulators and mobile robots[J]. *International Journal of Robotics Research*, 1986, 5(1): 290-298.
- [12] BALCH T, ARKIN R C. Behavior-based formation control for multirobot teams[J]. *IEEE Transaction on Robotics and Automation*, 1998, 14(6): 926-939.
- [13] 曹建福,凌志浩,高冲,等. 基于群集思想的多智能体编队避障算法研究[J]. *系统仿真学报*, 2014, 26(3): 562-566.
CAO J F, LING Z H, GAO C, et al. Obstacle avoidance and formation control for multi-agent based on swarming[J]. *Journal of System Simulation*, 2014, 26(3): 562-566 (in Chinese).
- [14] 许敬刚,王宏伦,刘畅,等. 无人机动态避撞区建模方法研

究[J]. 电光与控制, 2014, 21(12): 30-35.

XU J G, WANG H L, LIU C, et al. Modeling of dynamic collision avoidance zone for UAVs[J]. Electronics Optics & Control, 2014, 21(12): 30-35 (in Chinese).

- [15] LALISH E, MORGANSEN K A, TSUKAMAKI T. Formation tracking control using virtual structures and deconfliction [C] // Proceedings of the 45th IEEE Conference on Decision and Control 2006. Piscataway, NJ: IEEE Press, 2006: 5699-5705.

作者简介:

吴俊成 男, 硕士研究生。主要研究方向: 无人机编队飞行控制。

E-mail: wujc2013@163.com

周锐 男, 博士, 教授, 博士生导师。主要研究方向: 多无人机协同控制与智能决策。

Tel.: 010-82339232

E-mail: zhr@buaa.edu.cn

Formation flight control method of multiple UAVs based on guidance route

WU Juncheng¹, ZHOU Rui^{1*}, DONG Zhuoning¹, CHE Jun²

(1. Science and Technology on Aircraft Control Laboratory, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;

2. AVIC Automatic Flight Control Research Institute, Xi'an 710065, China)

Abstract: To solve the problem of formation flight control of multiple unmanned aerial vehicles (UAVs), a coordinated control method based on guidance route was proposed. According to the position error between wingman's current location and desired location and the flight status of leader, guidance route was generated for wingman to track. In view of the sudden obstacles during flight, the method of changing formation configuration was used to avoid obstacles. For possible collisions between UAVs, according to UAVs' position and time of arrival at the collision point, UAV avoided collision from two aspects of height and heading. The simulation results demonstrate that using this method, UAV can maintain stable formations, and meanwhile be able to avoid the obstacles during the flight and avoid the collision between aircrafts.

Key words: unmanned aerial vehicle (UAV); formation flight; guidance route; obstacle avoidance; collision avoidance

Received: 2015-07-08; **Accepted:** 2015-09-25; **Published online:** 2015-10-30 17:04

URL: www.cnki.net/kcms/detail/11.2625.V.20151030.1704.011.html

Foundation items: National Natural Science Foundation of China (61273349, 61175109, 61203223); Aeronautical Science Foundation of China (2013ZA18001, 2014ZA18004)

* **Corresponding author.** Tel.: 010-82339232 E-mail: zhr@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0463

基于在线轨迹迭代的自适应再入制导



赵颀, 沈作军*

(北京航空航天大学 航空科学与工程学院, 北京 100083)

摘 要: 针对传统轨迹跟踪制导方法在再入飞行中无法较好适应导航模式切换等突变状况的问题,提出了一种能够有效应对制导系统输入信息不连续性的自适应在线轨迹生成方法。该方法通过实时的多项式拟合以及迭代过程确定满足终端约束条件的高度-速度剖面,并解算出当前飞行状态下所需的攻角与倾侧角指令,从而平稳、精确地将飞行器引导至末端能量管理段。通过对速度与能量、高度、轨迹倾角以及待飞航程等状态量建立解析关系,该方法拥有迭代速度快以及收敛性强的优势。仿真结果显示,该方法对输入信息的误差及跳变等不确定因素的适应性很强,在各类干扰情况下较传统方法拥有更高的制导精度。相较于传统轨迹跟踪制导方法,该方法在实际应用背景下显著地提升了制导的自主性与适应性。

关键词: 再入制导; 自适应性; 多项式拟合; 轨迹迭代; 解析剖面规划

中图分类号: V448.235

文献标识码: A

文章编号: 1001-5965(2016)07-1526-10

再入制导是指再入飞行器在返回大气层时,制导系统通过在线生成指令引导飞行器从初始状态安全、精确地飞向特定终端状态的过程^[1]。由于再入飞行器往往是在复杂的外界环境下进行无动力飞行,因此,制导方法必须采用灵活、可靠的策略有效地耗散能量,从而满足各类严苛的过程约束和终端约束。

再入制导方法按照其生成指令的不同方式可大致分为两大类:即预测-校正制导和轨迹跟踪制导^[2]。就应用角度而言,虽然后者的制导精度和鲁棒性略逊于前者,但由于其计算量较小,因此在目前的再入飞行中往往更受青睐。

历史上,航天飞机采用的阻力加速度制导方法就属于一种典型的轨迹跟踪制导方法(以下简称轨迹跟踪方法)^[3-4]。近年来,为了进一步提高轨迹跟踪方法的自主性和鲁棒性,国内外学者对这类制导方法进行了深入研究。其中,文献[5-6]从不同角度对阻力加速度制导方法进行了改进,

进而避免了该方法在离线规划参考剖面时的繁琐过程;文献[7]则利用拟平衡滑翔条件(Quasi-Equilibrium Glide Condition, QEGC)对再入问题进行了简化,从而使再入轨迹的在线快速规划成为可能;文献[8]提出了一种应用线性二次型调节器(Linear Quadratic Regulator, LQR)的轨迹跟踪方法,拥有较高的鲁棒性;文献[9]提出了一种封闭形式的近似滚动时域控制(Receding Horizon Control, RHC)方法,避免了在线积分和显式的增益调度;文献[10]则提出了一种基于轨迹线性化控制(Trajectory Linearization Control, TLC)理论的再入轨迹跟踪制导律,减少了制导指令的解算时间。

尽管上述各类方法在仿真测试中均表现出了不错的精度和鲁棒性^[7-8],但要在实际飞行中成功实施这些方法,仍依赖于高精度的导航信息。然而,“黑障”区等复杂因素的存在使得再入飞行器的导航精度在相当长的一段时间内难以得到保证^[11]。此时,由于轨迹跟踪方法对导航信息的误

收稿日期: 2015-07-13; 录用日期: 2015-09-06; 网络出版时间: 2015-11-09 09:12

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151109.0912.002.html

* 通讯作者: Tel.: 010-82316572 E-mail: shenzuojun@buaa.edu.cn

引用格式: 赵颀, 沈作军. 基于在线轨迹迭代的自适应再入制导[J]. 北京航空航天大学学报, 2016, 42(7): 1526-1535.
ZHAO D, SHEN Z J. Adaptive reentry guidance based on on-board trajectory iterations[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1526-1535 (in Chinese).

差及不连续性的适应性较差,再入飞行往往面临失败的风险。

针对上述情况,本文提出了一种新型的自适应制导方法。该方法依据实时飞行状态对参考高度-速度剖面进行在线拟合、迭代,并解算出满足终端约束的制导指令,从而有效地消除“黑障”区内累计的导航误差以及飞离“黑障”区时导航信息跳变对制导产生的影响。各类仿真结果显示,该方法较传统轨迹跟踪制导方法拥有更高的自主性和适应性,在干扰情况下也能保持更高的制导精度。此外,该方法拥有较高的运算效率,具备在线实时规划参考剖面的应用潜力。

1 再入制导问题描述

1.1 再入动力学方程

考虑地球自转因素,建立基于北天东坐标系的三自由度无量纲运动方程如下^[12]:

$$\begin{cases} \dot{r} = V \sin \gamma \\ \dot{\theta} = V \cos \gamma \sin \psi / (r \cos \phi) \\ \dot{\phi} = V \cos \gamma \cos \psi / r \\ \dot{V} = -D - \sin \gamma / r^2 + \omega_e^2 \cos^2 \phi \cdot \sin \gamma - \cos \gamma \sin \psi \tan \phi \\ \dot{\gamma} = \frac{1}{V} \left[L \cos \sigma + \left(\frac{V^2}{r} - \frac{1}{r^2} \right) \cos \gamma + 2\omega_e V \cos \phi \sin \psi + \omega_e^2 r \cos \phi (\cos \gamma \cos \phi + \sin \gamma \cos \psi \sin \phi) \right] \\ \dot{\psi} = \frac{1}{V} \left[\frac{L \sin \sigma}{\cos \gamma} + \left(\frac{V^2}{r} \cdot \frac{\cos^2 \gamma}{\cos^2 \phi} + \omega_e^2 r \right) \cdot \frac{\cos \phi \sin \phi \sin \psi}{\cos \gamma} + 2\omega_e V (\sin \phi - \tan \gamma \cos \psi \cos \phi) \right] \end{cases} \quad (1)$$

式中: r 为被地球半径 R_0 无量纲化的地心距; V 为被第一宇宙速度 $V_e = \sqrt{R_0 g_0}$ 无量纲化的相对于地球的速度; g_0 为重力加速度; θ 和 ϕ 分别为弧度单位的经度和纬度坐标; γ 为速度矢量与水平面的夹角,即轨迹倾角; ψ 为从指北方向顺时针计量的速度方向角,即航向角; σ 为倾侧角; ω_e 为无量纲化了的地球自转角速度。所有状态变量均对无量纲化的时间 τ 求导, $\tau = t / \sqrt{R_0 / g_0}$ 。 L 和 D 分别为用 g_0 进行了无量纲化的升力加速度和阻力加速度,其表达式为

$$\begin{cases} L = \rho (V_e V)^2 C_L S_{ref} / (2mg_0) \\ D = \rho (V_e V)^2 C_D S_{ref} / (2mg_0) \end{cases} \quad (2)$$

式中: ρ 为大气密度; S_{ref} 为参考面积; m 为质量; C_L 和 C_D 为升力系数和阻力系数,它们均为攻角 α 、马赫数和高度的函数。

在再入轨迹规划问题中,攻角是根据航程、热防护及配平等要求预先给定的,一般为马赫数的函数:

$$\alpha = f(Ma) \quad (3)$$

1.2 再入轨迹约束条件

再入轨迹约束条件包括过程约束和终端约束。常见的过程约束包括热流率约束、过载约束和动压约束:

$$\begin{cases} \dot{Q} = c \sqrt{\rho} (V_e V)^{3.15} \leq \dot{Q}_{max} \\ N = L \cos \alpha + D \sin \alpha \leq n_{max} \\ q = 0.5 \rho (V_e V)^2 \leq q_{max} \end{cases} \quad (4)$$

式中: c 为与飞行器头部驻点半径有关的常数; n_{max} 为最大方向过载。

再入轨迹的终端约束一般包括飞行器进入末端能量管理段(Terminal Area Energy Management, TAEM)时的高度约束、速度约束和待飞航程约束:

$$\begin{cases} r_f = r_{TAEM} \\ V_f = V_{TAEM} \\ s_f = s_{TAEM} \end{cases} \quad (5)$$

式中: r_{TAEM} 、 V_{TAEM} 和 s_{TAEM} 分别为末端能量管理段高度、速度和待飞航程。

1.3 导航系统模型

在再入飞行段,飞行器一般采用惯性导航系统(Inertial Navigation System, INS)和全球定位系统(Global Positioning System, GPS)的组合导航方式^[13]。其中,INS拥有自主性高和抗干扰能力强的优势,但其导航误差会随时间积累;GPS则通过接受卫星信号解算出高精度的速度与位置信息,从而对INS的导航误差进行校正。

INS的导航误差是由惯性元件的随机噪声、系统的安装误差和刻度因数等因素引起的。本文按文献[14]中近空间飞行器INS误差的仿真结果,对式(1)中的各状态变量 $\mathbf{x} = [r \ \theta \ \phi \ V \ \gamma \ \psi]^T$ 的导航误差按时间进行拟合,有

$$\Delta \mathbf{x} = [\Delta r(t) \ \Delta \theta(t) \ \Delta \phi(t) \ \Delta V(t) \ \Delta \gamma(t) \ \Delta \psi(t)]^T \quad (6)$$

对于轨道再入飞行,“黑障”区一般存在于45~80 km的范围内^[13]。此时,GPS接收的无线电信号被屏蔽,飞行器只能依赖INS进行导航,故在这一阶段制导系统的输入信号为 $\mathbf{x}_{navi} = \mathbf{x} + \Delta \mathbf{x}$ 。而在“黑障”区以外,由于GPS能够精确地解算飞行器的速度和位置信息,因此导航误差可忽略不计,即 $\mathbf{x}_{navi} = \mathbf{x}$ 。制导系统在接收到导航信息 \mathbf{x}_{navi} 后,再计算出跟踪参考状态 \mathbf{x}_c 所需的制导指令。

整个系统的结构图如图 1 所示。

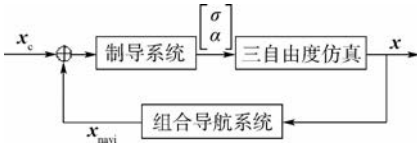


图 1 考虑导航误差的再入飞行结构

Fig. 1 Configuration of reentry flight with navigation error

1.4 再入飞行剖面

结合 1.3 节中组合导航系统的特点,本文按飞行器是否飞离“黑障”区并重新捕获 GPS 信号作为分界点,将再入轨迹分为轨迹跟踪段和在线更新段,如图 2 所示。

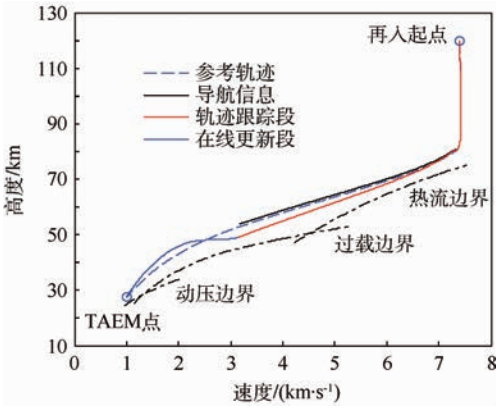


图 2 再入飞行高度-速度剖面

Fig. 2 Altitude-versus-velocity profile of reentry flight

轨迹跟踪段覆盖再入初期到飞行器飞离“黑障”区的飞行阶段。在整个轨迹跟踪段,飞行器都能较好地满足“拟平衡滑翔”条件,轨迹倾角的变化较为缓慢。因此,该段中仍可采用传统的轨迹跟踪制导方法。尽管“黑障”区内不断累积的导航误差会使飞行器的真实飞行状态不断偏离参考轨迹,但此时飞行器尚能较好地维持平稳飞行。

相比于轨迹跟踪段,在线更新段中制导方法面临的考验更为严峻:一方面,飞行器已不能很好地满足“拟平衡滑翔”条件,再入轨迹的力学特性发生了较大变化;另一方面,尽管存在导航信息的融合技术^[13],但是当 GPS 消除“黑障”区内积聚的导航误差后,导航信息仍可能出现跳变。此时,导航信息的误差及不连续性会致使传统轨迹跟踪方法出现控制饱和及能量耗散不均等状况。而在该段实时采用基于轨迹迭代的在线更新方法,可以有效地避免该类风险的发生。

2 制导方法设计

2.1 轨迹跟踪段制导方法

2.1.1 参考轨迹的离线规划

本文中所使用的参考轨迹按文献[7]提出的

“拟平衡滑翔”方法进行设计。该方法的核心是认为再入飞行轨迹的一大部分满足轨迹倾角 γ 较小且变化较缓慢的假设,即 $\cos \gamma = 1$ 和 $\dot{\gamma} = 0$ 。此时,式(1)中的第 5 式可简化为“拟平衡滑翔”条件:

$$L \cos \sigma + \left(V^2 - \frac{1}{r} \right) \frac{1}{r} = 0 \quad (7)$$

在初始下降段,由于大气密度稀薄,不需考虑各类约束的影响,因此飞行器在此段采用常值倾侧角 $|\sigma_0|$ 飞行。飞行从初始下降段过渡到拟平衡滑翔段的条件为

$$\left| \frac{dr}{dV} - \left(\frac{dr}{dV} \right)_{\text{QECC}} \right| \leq \delta \quad (8)$$

式中: $\delta > 0$ 为预先选定的常数; $\frac{dr}{dV}$ 和 $\left(\frac{dr}{dV} \right)_{\text{QECC}}$ 可分别由式(1)和式(7)确定。

飞行进入拟平衡滑翔段后,则可按文献[7]中的方法确定满足式(4)和式(5)中约束条件的倾侧角剖面 $\sigma(V)$,并进一步确定高度、速度、轨迹倾角和待飞航程等纵向状态变量的参考值,即 $\mathbf{x}_{\text{ion}} = [r \ V \ \gamma \ s]^T$ 。

为了增强轨迹跟踪控制的鲁棒性,并降低状态变量的维数,这里可引入能量 e 作为自变量,代替对制导问题不太关键的自变量时间 t 。能量 e 可表示为

$$e = \frac{1}{r} - \frac{V^2}{2} \quad (9)$$

至此,由“拟平衡滑翔”方法确定的纵向参考轨迹状态量和控制量可按均匀选取的能量节点 e_N 为自变量进行存储,即

$$\begin{cases} \mathbf{x}_c = [r_c(e_N) \ \gamma_c(e_N) \ s_c(e_N)]^T \\ \mathbf{u}_c = [|\sigma_c(e_N)| \ \alpha_c(e_N)]^T \end{cases} \quad (10)$$

对于式(10)所存储的参考轨迹,每个节点的参考速度可直接由 $V = \sqrt{2\left(\frac{1}{r} - e\right)}$ 计算得到。

2.1.2 轨迹在线跟踪

按照文献[8],在使用 LQR 方法前,需先对由式(10)存储的参考轨迹按式(11)进行小扰动线性化。

$$\begin{bmatrix} \bar{r}' \\ \bar{\gamma}' \\ \bar{s}' \end{bmatrix} = \mathbf{A}(e) \begin{bmatrix} \bar{r} \\ \bar{\gamma} \\ \bar{s} \end{bmatrix} + \mathbf{B}(e) \begin{bmatrix} |\bar{\sigma}| \\ \bar{\alpha} \end{bmatrix} \quad (11)$$

式中: $\bar{\mathbf{x}} = [\bar{r} \ \bar{\gamma} \ \bar{s}]^T$ 为状态误差量,是经小扰动后的纵向状态变量 $\mathbf{x}_{\text{ion}} = [r \ \gamma \ s]^T$ 与参考量 $\mathbf{x}_c = [r_c \ \gamma_c \ s_c]^T$ 之差。按式(11)确定的 \mathbf{A} 阵和 \mathbf{B} 阵,以及预先选定的权阵 \mathbf{Q} 阵和 \mathbf{R} 阵,可通过 LQR 方法算得对应于每个节点 e_N 的反馈增益 $\mathbf{K}(e_N)$ 。

在线跟踪轨迹时,制导系统通过导航系统输

出的导航信息 \mathbf{x}_{navi} 计算出当前的能量 e_i , 并对由式(10)记录的数表进行插值, 获取纵向参考状态量 $\mathbf{x}_c = [r_c(e_i) \quad \gamma_c(e_i) \quad s_c(e_i)]^T$ 和参考控制量 $\mathbf{u}_c = [|\sigma_c(e_i)| \quad \alpha_c(e_i)]^T$ 。相应地, 也可通过插值确定反馈增益 $\mathbf{K}(e_i)$ 。依据上述参考量和反馈增益, 可计算出跟踪轨迹所需的反馈控制量:

$$\delta \mathbf{u} = [\delta \sigma \quad \delta \alpha]^T = -\mathbf{K}(e_i) \delta \mathbf{x}_{\text{lon}} \quad (12)$$

式中: $\delta \mathbf{x}_{\text{lon}} = \mathbf{x}_{\text{navi/lon}} - \mathbf{x}_c$, 其中 $\mathbf{x}_{\text{navi/lon}} = [r_{\text{navi}} \quad \gamma_{\text{navi}} \quad s_{\text{navi}}]^T$ 为导航系统输出的当前纵向状态变量。轨迹跟踪段最终的制导指令为 $\mathbf{u} = [|\sigma_c + \delta \sigma| \quad \alpha_c + \delta \alpha]^T$, 倾侧角的符号由航向角误差漏斗逻辑确定。

2.2 在线轨迹更新段制导方法

2.2.1 高度-速度剖面的多项式拟合

考虑到再入过程中速度 V 具有良好的单调性, 并为了方便施加终端约束, 本文选择以速度为自变量拟合参考高度-速度剖面, 并进一步解算相应的制导指令。为了兼顾制导精度、状态变量的连续性以及算法的效率, 在线轨迹更新段的高度-速度剖面由四次多项式进行拟合, 即

$$r = a_4 V^4 + a_3 V^3 + a_2 V^2 + a_1 V + a_0 \quad (13)$$

剖面的起点可记为点 T , 该点的高度、速度和轨迹倾角由当前的导航信息确定, 记为 r_T, V_T 和 γ_T 。相应地, 此时的制导指令为 σ_T 和 α_T 。换言之, 对于每个制导周期, 高度-速度剖面的起点 T 由飞行器的实时飞行高度和速度确定。而高度-速度剖面的终点始终由 TAEM 点的高度约束 r_{TAEM} 与速度约束 V_{TAEM} 确定。

通过研究式(1)中再入飞行三自由度动力学方程组的第1式和第4式, 可知在忽略地球自转影响的前提下, 有

$$\frac{dr}{dV} = \frac{V \sin \gamma}{-D - \sin \gamma / r^2} \quad (14)$$

因此, 式(14)可根据导航信息 r_T, V_T 和 γ_T 确定高度-速度剖面在点 T 处的斜率 $(dr/dV)_T$ 。

求解式(13)中5个系数 a_i 所需的拟合条件至此确定了3项: $[V_T \quad r_T]$ 、 $[V_T \quad (dr/dV)_T]$ 以及 $[V_{\text{TAEM}} \quad r_{\text{TAEM}}]$ 。因此, 还需要再选取2个拟合条件, 才能完全确定上述5个拟合系数。

在确定第4个拟合条件时, 应当选择对航程调节能力强的参数作为节点, 以使式(13)中的参考剖面在不同飞行状态下均能满足终端航程约束 s_{TAEM} 。

注意到对于待飞航程有

$$\dot{s} = -\frac{V \cos \gamma \cos \Delta \psi}{r} \quad (15)$$

式中: $\Delta \psi$ 为航向角误差。通常情况下, 可认为飞行器近似沿地球大圆飞行, 即 $\cos \Delta \psi \approx 1$ 。那么, 结合式(1)中的第4式, 有

$$\frac{ds}{dV} = \frac{V \cos \gamma}{r(D + \sin \gamma / r^2)} \quad (16)$$

对于再入轨迹, 无量纲化高度 $r \approx 1$, 因此式(16)表明阻力加速度 D 对待飞航程的变化起主导作用。通过式(2)则容易发现特定飞行器所受阻力加速度 D 主要受到大气密度的影响。换言之, 式(13)中参考剖面的高度分布规律直接决定了飞行器航程的远近。

基于上述特点, 这里可引入点 T 和 TAEM 点之间的速度中点 C , 即

$$V_C = 0.5(V_T + V_{\text{TAEM}}) \quad (17)$$

选取点 C 对应的飞行中点高度 r_C 作为第4个节点, 则第4项拟合条件为 $[V_C \quad r_C]$ 。

由于前4个拟合条件已经基本确定了参考剖面的几何特征, 因此最后一个拟合条件的选择具有一定的自由度。本文在此主要考虑以下两类情况。

第一类情况中, 制导的首要目标是保证倾侧角在起点 T 处的连续性。对式(14)求导可得

$$\frac{d^2 r}{dV^2} = \frac{1}{V} \left(\sin \gamma + \frac{1}{V} \dot{\gamma} V \cos \gamma \right) + \frac{\dot{\gamma} \cos \gamma}{r^2 \dot{V}^3} + \frac{\dot{r}}{\dot{V}^2} D_V - \frac{2r^2 \sin \gamma}{r^3 \dot{V}^3} \quad (18)$$

式中: $D_V = \frac{\partial D}{\partial V}$ 。可见, 对于确定的 r_T, V_T, γ_T 和 α_T , $d^2 r/dV^2$ 在起点 T 处的连续性直接保证了倾侧角指令的连续。因此, 在这类情况中, 第5个拟合条件可选择为 $[V_T \quad (d^2 r/dV^2)_T]$ 。由此确定的剖面有如下形式:

$$r = a_4 v^4 + a_3 v^3 + a_2 v^2 + a_1 v + a_0 \quad (19)$$

式中: $v = V - V_{\text{TAEM}}$ 。此处定义 $\Delta V = V_T - V_{\text{TAEM}}$, $f_T = (dr/dV)_T$, $g_T = (d^2 r/dV^2)_T$, 则式(19)中相应的拟合系数为

$$\begin{cases} a_4 = \frac{g_T}{\Delta V^2} - 6 \frac{f_T}{\Delta V^3} + 2 \frac{r_{\text{TAEM}}}{\Delta V^4} - 16 \frac{r_C}{\Delta V^4} + 14 \frac{r_T}{\Delta V^4} \\ a_3 = -2.5 \frac{g_T}{\Delta V} + 17 \frac{f_T}{\Delta V^2} - 7 \frac{r_{\text{TAEM}}}{\Delta V^3} + 48 \frac{r_C}{\Delta V^3} - 41 \frac{r_T}{\Delta V^3} \\ a_2 = 2g_T - 15 \frac{f_T}{\Delta V} + 9 \frac{r_{\text{TAEM}}}{\Delta V^2} - 48 \frac{r_C}{\Delta V^2} + 39 \frac{r_T}{\Delta V^2} \\ a_1 = -0.5g_T \Delta V + 4f_T - 5 \frac{r_{\text{TAEM}}}{\Delta V} + 16 \frac{r_C}{\Delta V} - 11 \frac{r_T}{\Delta V} \\ a_0 = r_{\text{TAEM}} \end{cases} \quad (20)$$

另一类情况中,为优先保证制导精度,需将终端轨迹倾角约束在某一特定值 γ_{TAEM} 。此时,参考高度-速度剖面的终端斜率可由式(14)确定。则第5个拟合条件相应地取为 $[V_{TAEM} \quad (dr/dV)_{TAEM}]$ 。参考高度-速度剖面此时则可记为

$$r = b_4v^4 + b_3v^3 + b_2v^2 + b_1v + b_0 \tag{21}$$

此处定义 $f_{TAEM} = (dr/dV)_{TAEM}$, 则相应的拟合系数为

$$\begin{cases} b_4 = -2 \frac{f_{TAEM}}{\Delta V^3} + 2 \frac{f_T}{\Delta V^3} - 8 \frac{r_{TAEM}}{\Delta V^4} + 16 \frac{r_C}{\Delta V^4} - 8 \frac{r_T}{\Delta V^4} \\ b_3 = 5 \frac{f_{TAEM}}{\Delta V^2} - 3 \frac{f_T}{\Delta V^2} + 18 \frac{r_{TAEM}}{\Delta V^3} - 32 \frac{r_C}{\Delta V^3} + 14 \frac{r_T}{\Delta V^3} \\ b_2 = -4 \frac{f_{TAEM}}{\Delta V} + \frac{f_T}{\Delta V} - 11 \frac{r_{TAEM}}{\Delta V^2} + 16 \frac{r_C}{\Delta V^2} - 5 \frac{r_T}{\Delta V^2} \\ b_1 = f_{TAEM} \\ b_0 = r_{TAEM} \end{cases} \tag{22}$$

图3中分别绘制了由式(19)和式(21)确定的参考高度-速度剖面以及相应的倾侧角指令。由图3可见,式(19)确定的剖面在起点 T 处变化更为平缓,对应倾侧角指令的连续性也较好。但其中后段的变化加剧,倾侧角指令迅速逼近下边界,不利于轨迹跟踪;相反,式(21)确定的剖面在起点 T 附近较为陡峭,相应的倾侧角指令发生了突变。但其中段以后的变化逐渐放缓,倾侧角指令也较平滑。

基于以上特点,制导方法可在飞行器刚进入在线更新段时选用式(19)进行参考剖面的拟合,

从而保证倾侧角指令的连续性;而当飞行进入在线更新段的中段时,制导方法可选取合适的时机切换到式(21)的剖面拟合,从而更好地保证终端制导精度。

具体的切换时机可如下选取:对于第 i 个制导周期,若假设与参考由式(19)确定的剖面 and 式(21)确定的剖面对应的倾侧角指令分别为 $|\sigma_1^{(i)}|$ 和 $|\sigma_2^{(i)}|$, 则当两组指令之差 $|\Delta\sigma^{(i)}| = ||\sigma_1^{(i)}| - |\sigma_2^{(i)}||$ 不再减小时: $|\Delta\sigma^{(i)}| - |\Delta\sigma^{(i-1)}| > 0$

制导便可由式(19)确定的剖面切换到由式(21)确定的剖面。从图3可见,在线更新段的前半段中,与式(19)确定的剖面对应的倾侧角指令的变化速率在不断增快,而与式(21)确定的剖面对应的倾侧角指令则正好与之相反,即 $|\sigma_1^{(i)}|$ 一直处于追逐 $|\sigma_2^{(i)}|$ 的状态。因此条件式(23)一般在在线更新段的中段得到满足。

需要指出的是,在一些导航误差较小的情况下,式(21)确定的剖面所对应的倾侧角指令只发生 5° 以内的跳变。在这种情况下,在线更新段可以始终采用式(21)确定的剖面进行制导,而无需引入式(19)确定的剖面的应用。

2.2.2 高度-速度剖面的迭代

为使式(19)和式(21)拟合的剖面满足终端航程约束,需要对中点高度 r_C 进行迭代求解。

通过研究发现,在迭代时将能量 e 作为自变量,并将终端待飞航程 s_f 作为迭代指标可以提高算法的效率和收敛性。这是由于待飞航程 s 相对于能量 e 有很好的单调性,且终端航程误差 $s_f - s_{TAEM}$ 的符号有明确的物理意义:当其为正时,表明飞行器航程过近;反之,则说明航程过远。

由式(9)可知,对于能量 e 有

$$\frac{de}{dr} = DV > 0 \tag{24}$$

即整个再入过程中能量 e 单调递增(能量 e 是无量纲化的机械能的负值)。结合式(15),可知:

$$\frac{ds}{de} = -\frac{\cos \gamma}{rD} \tag{25}$$

对于第 k 步迭代以及高度 $r_C^{(k)}$, 通过式(9)和式(19)(或式(21))很容易求出当前能量 e 对应的速度 V 和高度 r 。对式(19)(或式(21))求导一次并结合式(14)则可求得相应的轨迹倾角 γ 。至此,式(25)的右半部分完全确定。将待飞航程 s 从起点 e_T 积分至终点 e_{TAEM} 便可得到终端待飞

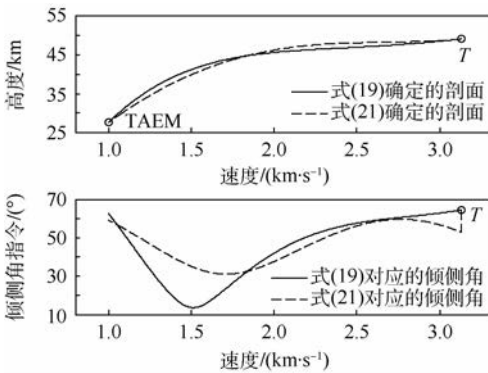


图3 参考高度-速度剖面及倾侧角指令

Fig.3 Reference altitude-versus-velocity profile and bank angle command

航程 $s_f^{(k)}$ 。

第 $k+1$ 步迭代所用的中点高度 $r_C^{(k+1)}$ 可由割线法进行更新:

$$r_C^{(k+1)} = r_C^{(k)} - \frac{r_C^{(k)} - r_C^{(k-1)}}{s_f^{(k)} - s_f^{(k-1)}} (s_f^{(k)} - s_{TAEM}) \quad (26)$$

若第 k 步迭代的终端待飞航程 $s_f^{(k)}$ 满足条件:

$$|s_f^{(k)} - s_{TAEM}| \leq \varepsilon \quad \varepsilon > 0 \quad (27)$$

则对应的中点高度 $r_C^{(k)}$ 所确定的高度-速度剖面即为最终的参考剖面,迭代过程结束。

值得注意的是,虽然以终端待飞航程 s_f 为迭代指标时,迭代过程对初值 $r_C^{(0)}$ 和 $r_C^{(1)}$ 并不敏感,但在中点速度 V_C 所对应的再入走廊中点 $\pm 2 \sim 3$ km 的范围内选取初值仍可在一定程度内提高迭代速度。此外,选取 $\varepsilon = 10^{-5}$ 便可保证良好的制导精度。

2.2.3 制导指令的生成

一旦中点高度 $r_C^{(k)}$ 得以确定,则对于当前速度 V ,通过式(14)和式(19)(或式(21))很容易求得相应的参考高度 r 和参考轨迹倾角 γ 。通过式(3)给定的攻角剖面,则易求得相应的攻角指令 α 。而结合式(1)中的第1、第4和第5式以及式(18)便可求得倾侧角指令 $|\sigma|$ 。倾侧角的符号仍由航向角误差漏斗逻辑确定。

值得注意的是,由式(20)和式(22)可见,当 ΔV 随着飞行不断减小时,多项式拟合的效果将逐渐变差。就物理意义而言, ΔV 对应着飞行器的剩余能量。当 ΔV 较大时,飞行器具有更多的剩余能量,高度-速度剖面也就具备了更大的调节范围;而当剩余能量较少时,高度-速度剖面的调节范围将会减小,而剖面的高度调节范围又直接对应飞行器的航程范围,因此此时终端航程精度也会随之降低。针对本文中的飞行器,通过仿真验证表明,当 $2000 \text{ m/s} \leq V \leq 3200 \text{ m/s}$ 时,在线更新方法拥有良好的制导精度。因此,算法在满足飞行速度 $V \leq 2000 \text{ m/s}$ 时将进行最后一次参考剖面的规划,并求解整个剖面的纵向参考状态和制导指令。上述信息将以式(10)的格式进行存储。随后的飞行中,制导系统通过式(12)对上述剖面进行在线跟踪。

3 在线应用及仿真验证

3.1 在线调整策略

在线应用过程中,在导航误差过大或升阻力所受干扰较强的情况下,上述在线更新方法可能会解算出超过飞行器能力范围的制导指令。此

时,需要对拟合参数进行适当调节。

对于特定飞行器,参考高度-速度剖面的变化范围主要受倾侧角 σ 的限制($0^\circ < \sigma < 90^\circ$),故由式(18)可见,对于确定的 r, V, γ 和 α ,适当地对剖面各点的 d^2r/dV^2 值进行调节,便能解算出合理的倾侧角指令 σ_{cmd} 。

结合文献[15]中的结论,对于本文中规划的高度-速度剖面式(19)和式(21),由于点 T 和 $TAEM$ 点处的倾侧角指令 σ_{cmd} 均较大,故只需对中点 C 附近的 d^2r/dV^2 进行调节以便保证倾侧角指令 σ_{cmd} 在飞行器的能力范围之内。结合式(20)和式(22),并对式(19)和式(21)求导2次,有

$$\left\{ \begin{aligned} \left. \frac{d^2r}{dV^2} \right|_C &= -2.5g_T + 18 \frac{f_T}{\Delta V} - 6 \frac{r_{TAEM}}{\Delta V^2} + \\ &48 \frac{r_C}{\Delta V^2} - 42 \frac{r_T}{\Delta V^2} \\ \left. \frac{d^2r}{dV^2} \right|_C &= 5 \frac{f_{TAEM}}{\Delta V} - 2 \frac{f_T}{\Delta V} + 19 \frac{r_{TAEM}}{\Delta V^2} - \\ &32 \frac{r_C}{\Delta V^2} + 13 \frac{r_T}{\Delta V^2} \end{aligned} \right. \quad (28)$$

由式(28)可知,当式(19)确定的剖面超出飞行器能力范围时,可牺牲一定的制导精度,通过放宽终端航程约束 s_{TAEM} 来调节中点高度 r_C ,进而减小 d^2r/dV^2 ,从而保证相应倾侧角指令 σ_{cmd} 在限制范围内。针对本文中的飞行器,仿真结果表明, s_{TAEM} 每放宽 5 km, σ_{cmd} 的最小值将增加 4° 左右,且 s_{TAEM} 的放宽量一般不必超过 10 km;对于式(21)确定的剖面,则可采用减小 γ_{TAEM} 的策略调节 d^2r/dV^2 的值,进而保证相应的倾侧角指令 σ_{cmd} 得以实现。针对本文中的飞行器,仿真结果表明, γ_{TAEM} 每减小 0.5° , σ_{cmd} 的最小值将增加 2° 左右, γ_{TAEM} 的调节范围一般在 $-12.0^\circ \sim -14.0^\circ$ 之内。需要指出的是,由于参考剖面处于实时更新的状态,因此,除最后一次更新参考剖面外,单个制导周期只需保证起点 T 处对应的倾侧角指令不至于过小($|\sigma| \geq 5^\circ$),便可认为参考剖面满足飞行器的能力范围。

3.2 仿真验证

本文的仿真采用航天飞机模型,再入飞行任务的初始状态参数和过程约束设置分别如表1和表2所示。

所有飞行任务的目标点航向对准锥(HAC)的经、纬度坐标分别为 255.55° 和 12.50° 。终端航程约束 $s_{TAEM} = 110.0 \text{ km}$,高度约束 $r_{TAEM} = 27.5 \text{ km}$,速

度约束 $V_{TAE M} = 1000.0 \text{ m/s}$ 。所有仿真的截止条件均取为 $V \leq 1000.0 \text{ m/s}$ 。

表 1 仿真初始状态参数

Table 1 Initial state parameters of simulations

参数	高度/ km	速度/ ($\text{m} \cdot \text{s}^{-1}$)	经度/ ($^{\circ}$)	纬度/ ($^{\circ}$)	轨迹倾角/ ($^{\circ}$)	航向角/ ($^{\circ}$)
数值	120.0	7400.0	175.5	0	-1.1	74.2

表 2 仿真过程约束

Table 2 Process constraints of simulations

参数	最大热流率/($\text{MW} \cdot \text{m}^{-2}$)	最大过载/ g	最大动压/ kPa
数值	760	2.5	20

3.2.1 常规导航误差下的仿真结果

本文按照文献[14]中的数据对捷联式惯性导航系统的误差进行拟合,并将其折合为三自由度动力学方程组(1)中各状态变量的导航误差。本文中INS的误差参数如下:陀螺仪安装误差为 $10''$,标度因数误差 1.0×10^{-4} ,随机零偏 $0.01(^{\circ})/\text{h}$,一阶马尔可夫过程相关时间 3600 s ,驱动白噪声方差 $0.01(^{\circ})/\text{h}$;加速度计安装误差为 $600''$,标度因数误差 1.5×10^{-3} ,高斯白噪声 $0.01(^{\circ})/\text{h}$,一阶马尔可夫过程相关时间 1800 s ,驱动白噪声方差 $10^{-4}g$ 。其 1000 s 时各状态变量的导航误差如表3所示。

按上述参数设置导航误差,并对传统轨迹跟踪方法和在线更新方法分别进行仿真,可得到表4以及图4~图6中的结果。

表 3 惯性导航系统 1000 s 时的导航误差

Table 3 Navigation error of INS at 1000th second

参数	高度/ km	速度/ ($\text{m} \cdot \text{s}^{-1}$)	经度/ ($''$)	纬度/ ($''$)	轨迹倾角/ ($''$)	航向角/ ($''$)
数值	3.5	30	32.3	32.3	35.0	80.0

表 4 常规导航误差下的终端制导误差

Table 4 Terminal guidance error under normal navigation error

方法	高度/km	航程/km
轨迹跟踪方法	-0.010	2.63
在线更新方法	0.182	-0.77

由表4可见,轨迹跟踪方法与在线更新方法的终端高度精度均较好,而在线更新方法的终端航程精度要高于轨迹跟踪方法。

由图4和图5可见,在满足终端高度约束的情况下,在线跟踪方法可以有效地避免再入轨迹的剧烈变化,保证飞离“黑障”区后飞行过程的平稳过渡。而传统的轨迹跟踪方法则难以达到此目标。

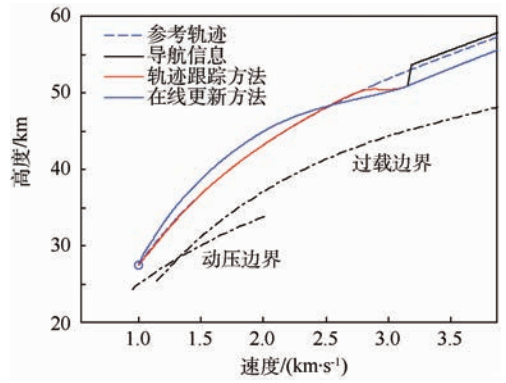


图 4 高度-速度剖面仿真结果

Fig. 4 Simulation results of altitude-versus-velocity profile

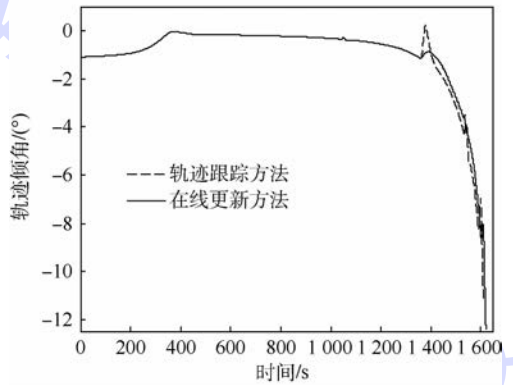


图 5 轨迹倾角历程

Fig. 5 Flight-path angle history

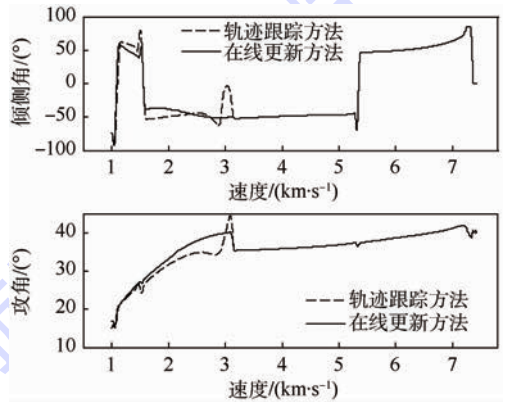


图 6 制导指令历程

Fig. 6 Guidance command history

由图6可见,通过实时规划参考剖面,在线更新方法有效避免了传统轨迹跟踪方法中出现的倾侧角剧烈振荡现象。此外,两种方法的攻角指令均在离开“黑障”区时发生了跳变。这是由于在轨迹跟踪段内,LQR方法需要同时对参考高度和参考航程进行追踪。而由于导航误差的存在,飞行器的实际飞行高度低于参考高度,在此情况下,制导逻辑只能通过逐渐减小攻角来降低阻力,进而跟踪参考航程。

3.2.2 极端导航误差下的仿真结果

由于再入过程中的高超声速气流扰动以及高频动态飞行环境等因素可能会使INS的导航误差进一步显著增加,因而有必要检验在线更新方法在此类恶劣飞行环境下的适应性。为此,本文针对3.2.1节中的INS设置了其极端情况下的导航误差量,如表5所示。

按以上参数设置对在线更新方法进行仿真,可得到表6、图7和图8中的结果。

表5 极端情况下惯性导航系统1000s时的导航误差
Table 5 Navigation error of INS at 1000th second under extreme conditions

参数	高度/ km	速度/ (m·s ⁻¹)	经度/ (")	纬度/ (")	轨迹倾角/ (")	航向角/ (")
数值	5.0	50	48.5	48.5	50.0	120.0

表6 极端导航误差下的终端制导误差

Table 6 Terminal guidance error under extreme navigation error

方法与条件	高度/km	航程/km
轨迹跟踪方法	-0.014	11.74
未考虑倾侧角连续条件	-0.139	1.56
考虑倾侧角连续条件	-0.131	-0.26

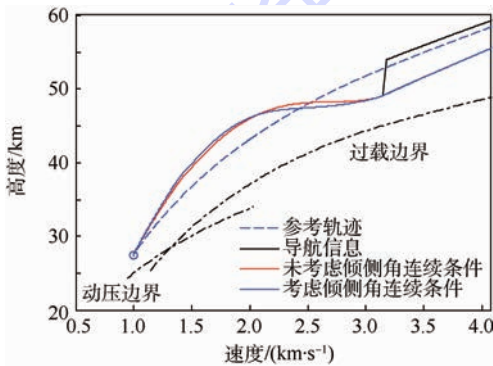


图7 极端导航误差下的高度-速度剖面仿真结果

Fig.7 Simulation results of altitude-versus-velocity profile under extreme navigation error

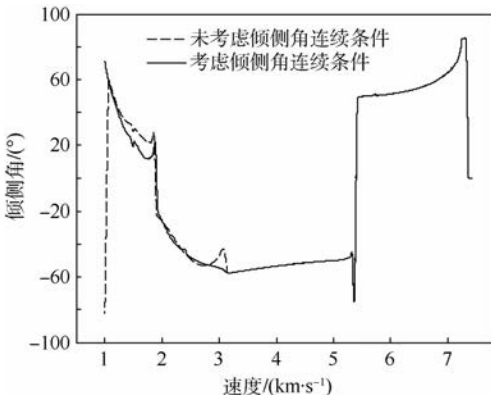


图8 极端导航误差下的倾侧角指令历程

Fig.8 Bank angle command history under extreme navigation error

由表6可见,3种情况下的终端高度精度均较好,但在线更新方法的终端航程精度要远高于轨迹跟踪方法。此外,考虑倾侧角连续条件时的在线更新方法的终端航程精度要略优于不考虑倾侧角连续条件的情况。

由图7可见,相对于未考虑倾侧角连续条件的情况,当在线更新方法引入式(19)确定的剖面时,再入轨迹在在线更新段的起点处拥有更加平滑的过渡段。而在再入轨迹的末端,由于考虑倾侧角连续条件的在线更新方法已经切换到了式(21)确定的剖面,因此两种情况下的末端轨迹具有相似的形态。

由图8可见,相较于提高制导精度,考虑倾侧角连续条件更大的作用在于通过引入式(19)确定的剖面抑制倾侧角突变。从仿真结果来看,未考虑倾侧角连续条件时,倾侧角出现了10°的跳变。同时,倾侧角速率已经接近15(°)/s的限制值;相比之下,考虑倾侧角连续条件时,倾侧角速率小于1(°)/s。在实际飞行中,倾侧角的连续性对应着飞行器姿态角的连续性,因此通过制导方法主动消除不必要的倾侧角突变比起被动地限制倾侧角速率,对降低飞行器姿态控制的难度有着更积极的作用。

3.2.3 气动拉偏情况下的仿真结果

本文挑选了表7中的4种典型气动拉偏情况,对传统轨迹跟踪方法和在线更新方法分别进行了仿真。仿真中的导航误差按表3设置。

按上述参数进行仿真可得如表8、图9和图10所示的结果。

由表8可见,在拉偏的情况下,两种方法的高度误差近似;而在线更新方法的终端航程误差则

表7 气动拉偏参数

Table 7 Aerodynamic dispersion parameters

案例	阻力拉偏/%	升力拉偏/%
案例1	-10	+10
案例2	-5	+5
案例3	+3	-3
案例4	+5	-5

表8 气动拉偏情况下的终端制导误差

Table 8 Terminal guidance error under aerodynamic dispersion conditions

案例	轨迹跟踪方法		在线更新方法	
	高度/km	航程/km	高度/km	航程/km
案例1	-0.450	-28.8	-0.223	-12.23
案例2	-0.087	-14.7	0.184	-4.8
案例3	0.262	8.0	-0.018	-1.4
案例4	-0.199	16.4	0.313	7.3

远远低于传统的轨迹跟踪方法,体现出了其具备更强的航程适应性。值得注意的是,案例4已接近飞行器的性能极限(升阻比减少10%),在线更新方法不适用于更加恶劣的拉偏情况;而对于升阻比增加的情况(案例1和案例2),在线更新方法表现出了更强的适应性。

图9和图10分别绘制了极端拉偏情况下两种方法的高度-速度剖面及相应倾侧角指令。可见,在线更新方法可针对不同的升阻力情况灵活改变参考剖面 and 倾侧角指令,从而满足航程要求。而传统轨迹跟踪方法则在所用情况中均沿同一参考剖面飞行,进而难以对航程实现调节。这也是其终端航程精度远低于在线更新方法的原因。此外,由图10可见案例4中的倾侧角指令在在线更新段的中段已接近飞行器倾侧角调节范围的下边界。即在线更新方法将难以应对比此更为恶劣的拉偏情况。

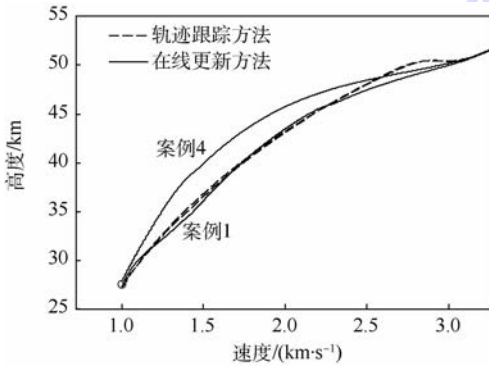


图9 气动拉偏情况下的高度-速度剖面仿真结果

Fig.9 Simulation results of altitude-versus-velocity profile under aerodynamic dispersion conditions

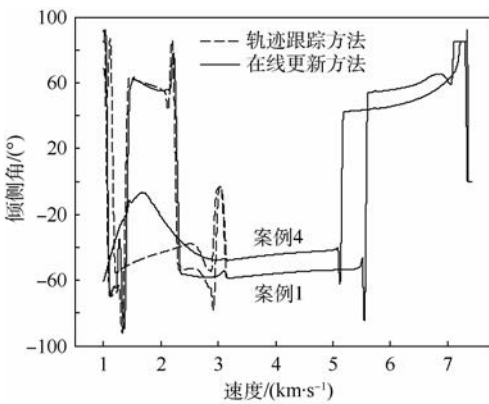


图10 气动拉偏情况下的倾侧角指令历程

Fig.10 Bank angle command history under aerodynamic dispersion conditions

4 结论

本文结合再入飞行中“黑障”区造成的导航信息误差与跳变等实际问题,提出了一种基于轨

迹迭代的自适应在线更新方法。理论分析研究以及轨迹跟踪方法的对比仿真结果表明:

1) 在线更新方法能够结合实时飞行状态快速拟合、迭代出满足终端高度、速度与航程约束的参考高度-速度剖面,并生成相应制导指令。在常规和极端的导航信息误差及跳变情况下均具有良好的自适应性和终端精度。

2) 在线更新方法在拟合参考剖面时充分考虑了各状态量与控制量的连续性,有效避免了传统轨迹跟踪制导方法在飞行器离开“黑障”区时出现的倾侧角剧烈振荡及飞行轨迹骤变等高风险飞行状态,提高了飞行过程的平稳性。

3) 与轨迹跟踪方法相比,在线更新方法在存在气动拉偏的情况下能够更灵活地调节飞行器的飞行状态,从而使其具备更高的航程适应能力和制导精度。

4) 在线更新方法具有较高的运算效率。在普通计算机(2.40 GHz 处理器)上运行时,完成单次高度-速度剖面的规划所需的运算时间在25~50 ms 范围内。而完成整个在线更新段100次左右的高度-速度剖面规划仅需3.2 s。故该方法具备在线实时规划再入轨迹的应用潜力。

因此,本文提出的自适应在线更新方法能够有效应对再入飞行中制导系统输入信息存在的误差及跳变,对提高制导的自主性、适应性和精度具有一定的工程应用价值。

参考文献 (References)

- [1] LU P. Entry guidance: A unified method [J]. Journal of Guidance, Control and Dynamics, 2014, 37(3): 713-728.
- [2] WINGROVE R C. Survey of atmosphere re-entry guidance and control methods [J]. AIAA Journal, 1963, 1(9): 2019-2029.
- [3] HARPOLD J C, GRAVES C A. Shuttle entry guidance [J]. Journal of Astronautical Sciences, 1979, 37(3): 239-268.
- [4] HARPOLD J C, GAVERT D E. Space shuttle entry guidance performance results [J]. Journal of Guidance, Control, and Dynamics, 1983, 6(6): 442-447.
- [5] ROENNEKE A J. Adaptive on-board guidance for entry vehicle: AIAA-2001-37343 [R]. Reston: AIAA, 2001.
- [6] MEASE K D, CHEN D T, TEUFE P, et al. Reduced-order entry trajectory planning for acceleration guidance [J]. Journal of Guidance, Control, and Dynamics, 2002, 25(2): 257-266.
- [7] SHEN Z J, LU P. On-board generation of three-dimensional constrained entry trajectories [J]. Journal of Guidance, Control, and Dynamics, 2003, 26(1): 111-121.
- [8] DUKEMAN G A. Profile-following entry guidance using linear quadratic regulator theory [C] // AIAA Guidance, Navigation, and Control Conference and Exhibit 2002. Reston: AIAA, 2002.

- [9] LU P. Regulation about time-varying trajectories: Precision entry guidance illustrated [J]. *Journal of Guidance, Control, and Dynamics*, 1999, 22 (6): 784-790.
- [10] 朱国栋,沈作军. 基于轨迹线性化控制的再入轨迹跟踪制导 [J]. *北京航空航天大学学报*, 2015, 41 (11): 1975-1982. ZHU G D, SHEN Z J. Trajectory linearization control based tracking guidance design for entry flight [J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2015, 41 (11): 1975-1982 (in Chinese).
- [11] 李瑾. 跨大气层可重复使用运载器再入段自主导航研究 [D]. 北京:北京航空航天大学, 2007: 4. LI J. Study of the autonomous navigation of spanning aerospace reusable launch vehicle [D]. Beijing: Beihang University, 2007: 4 (in Chinese).
- [12] VINH N X, BUSEMANN A, CULP R D. Hypersonic and planetary entry flight mechanics [M]. Ann Arbor, MI: University of Michigan Press, 1980.
- [13] 杨峰,程承,张共愿. 一种亚轨道飞行器再入段组合导航方法 [J]. *宇航学报*, 2010, 31 (3): 729-733. YANG F, CHENG C, ZHANG G Y. Simulation research of sub-orbital vehicle re-entry flight segment navigation technology [J]. *Journal of Astronautics*, 2010, 31 (3): 729-733 (in Chinese).
- [14] 彭慧. 近空间飞行器惯性导航系统误差建模及修正关键技术 [D]. 南京:南京航空航天大学, 2014: 25-28. PENG H. The key technology research on error modeling and compensation of the INS for the HCV [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2014: 25-28 (in Chinese).
- [15] SHEN Z J, LU P. On-board entry trajectory planning for sub-orbital flight [J]. *Acta Astronautica*, 2005, 56 (6): 573-591.

作者简介:

赵嶝 男, 硕士研究生。主要研究方向: 再入制导。

E-mail: andy_zhao1991@sina.com

沈作军 男, 博士, 教授。主要研究方向: 制导技术与飞行控制、轨迹优化。

Tel. : 010-82316572

E-mail: shenzuojun@buaa.edu.cn

Adaptive reentry guidance based on on-board trajectory iterations

ZHAO Di, SHEN Zuojun *

(School of Aeronautic Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: An adaptive method for on-board generation of entry trajectory is presented to effectively tackle the discontinuity of the input data of the guidance system, as the traditional trajectory-tracking guidance method cannot well adapt to saltatorial conditions such as the switch of navigation modes. By real-time polynomial fitting and iteration, the desired altitude-versus-velocity profiles satisfying the final constraints are determined, and then the corresponding angle of attack and bank angle commands are obtained, which guide the vehicle to safely and accurately reach the terminal area of energy management. This method enables the iterations to be fast-converging, as it takes the advantage of building analytical relations between velocity and the state variables including energy, altitude, flight-path angle and range-to-go. Numerical simulations indicate that this method has a strong adaptability against uncertainties such as the error and the saltation of the input data, as well as a better accuracy than the traditional method under various dispersion conditions. It is also proved that this method significantly enhances the autonomy and adaptability of entry guidance under practical application circumstances, compared to the traditional trajectory-tracking guidance method.

Key words: reentry guidance; adaptability; polynomial fitting; trajectory iterations; analytical profile planning

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0486

基于高阶奇异值分解的 LPV 鲁棒控制器设计

孙斌, 杨凌宇*, 张晶

(北京航空航天大学 自动化科学与电气工程学院, 北京 100083)



摘 要: 针对高超声速飞行器线性变参数 (LPV) 模型建模中准确度与复杂度之间的矛盾性, 提出了一种基于网格化张量及高阶奇异值分解 (HOSVD) 的凸多胞 LPV 模型建模及控制方法。首先基于雅可比线性化给出了大包线网格化模型的张量描述形式, 然后提出了一种基于 HOSVD 的多胞 LPV 模型生成算法, 将网格化模型表述为有限个线性时不变 (LTI) 顶点及权重函数的组合, 并基于舍弃的奇异值给出了建模误差的指标, 最后结合某航天飞机再入段六自由度非线性模型进行了大包线鲁棒 LPV 控制器设计与仿真验证, 结果表明该方法可获得计算复杂度低且保证建模精度的 LPV 模型, 设计的鲁棒变增益控制器能够使系统快速跟踪姿态角指令信号, 并能够保证系统的稳定性和鲁棒性。

关键词: 高超声速飞行器; 飞行包线; 多胞线性变参数 (LPV) 系统; 鲁棒变增益控制; 高阶奇异值分解 (HOSVD)

中图分类号: V249.12

文献标识码: A **文章编号:** 1001-5965(2016)07-1536-07

与传统飞行器相比, 高超声速飞行器再入过程需要以十几倍声速跨越约 100 km 的飞行高度, 具有飞行包线跨度大、参数快时变以及强不确定性等特性, 给控制器的设计提出了更高的要求。相对于动态逆^[1]、滑模^[2]等非线性控制器设计过程的复杂性, 基于线性变参数 (LPV) 系统的变增益控制方法能够处理参数大范围快时变系统, 同时具有全局的稳定性保证, 成为高超声速飞行控制的研究热点。

现有 LPV 控制方法在应用中的关键问题在于如何建立准确、计算复杂度低的 LPV 模型。目前 LPV 建模较常用的方法为 Bounding box 方法^[3-4], 该方法先将系统转化为仿射参数依赖型, 然后分析选取系统矩阵中对系统特性影响较大的变量作为变参数, 并直接选取变参数的上下界形成顶点系统, 此方法变参数的选取依赖于设计人

员的经验, 所建立的 LPV 模型在非顶点处的误差没有定量描述, 且存在一定的保守性。文献[5]运用 Qhull 算法在包线范围内基于实际的工作点求取 LPV 多胞顶点, 相较于 Bounding box 方法, 此方法所求得的多胞形区域更小, 保守性降低。对于不具备仿射参数依赖型的 LPV 系统, 文献[6]提出了一种张量积分分解方法求解顶点系统, 此方法要求系统矩阵可表示为状态变量的显示函数, 由于高超声速飞行器的强非线性以及气动参数的不确定性, 系统矩阵的函数表达式是很难得到的。文献[7-8]采用文献[6]提出的张量积分分解理论, 先将飞行器在包线范围内线性化得到一系列线性时不变 (LTI) 系统, 选取状态变量中的变量作为变参数, 然后进行数值拟合得到 LPV 模型, 进而应用高阶奇异值分解 (HOSVD) 理论选取顶点, 此方法数值拟合难度较大, 且精度也很难保

收稿日期: 2015-07-20; 录用日期: 2015-10-16; 网络出版时间: 2015-12-17 10:41

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151217.1041.007.html

基金项目: 国家自然科学基金(61273099); 航空科学基金(20135851041)

* 通讯作者: Tel.: 010-82316873 E-mail: yanglingyu@buaa.edu.cn

引用格式: 孙斌, 杨凌宇, 张晶. 基于高阶奇异值分解的 LPV 鲁棒控制器设计[J]. 北京航空航天大学学报, 2016, 42(7): 1536-1542. SUN B, YANG L Y, ZHANG J. Robust LPV control design based on HOSVD[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1536-1542 (in Chinese).

证。在飞行器 LPV 多胞系统建模过程中,如何同时兼顾到建模准确性、复杂性、保守性以及控制器阶数是值得研究的问题。

本文提出了一种基于网格化张量及 HOSVD 的凸多胞 LPV 模型建模方法,避免了 LPV 系统矩阵函数表达式的拟合过程,给出了奇异值与建模精度和多胞顶点数量的关系,可减少工作量和对设计人员经验的依赖,在保证建模精度的前提下控制顶点数量,提高建模过程的自动化程度。

1 大包线对象的 LPV 建模问题

设对象在整个飞行包线内的模型可表述为如式(1)所示的非线性形式:

$$\begin{cases} \dot{\mathbf{x}} = f_1(\mathbf{x}, \mathbf{u}) \\ \mathbf{y} = f_2(\mathbf{x}, \mathbf{u}) \end{cases} \quad (1)$$

式中: $\mathbf{x} \in \mathbf{R}^m$ 为系统状态; $\mathbf{u} \in \mathbf{R}^n$ 为系统输入信号; $\mathbf{y} \in \mathbf{R}^l$ 为系统输出信号,则非线性系统可转化为如下的 LPV 形式:

$$\begin{bmatrix} \dot{\mathbf{x}} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{A}(\boldsymbol{\rho}(t)) & \mathbf{B}(\boldsymbol{\rho}(t)) \\ \mathbf{C}(\boldsymbol{\rho}(t)) & \mathbf{D}(\boldsymbol{\rho}(t)) \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{u} \end{bmatrix} \quad (2)$$

式中: $\boldsymbol{\rho}(t) = [\rho_1(t) \ \rho_2(t) \ \dots \ \rho_N(t)]^T$ 为随飞行状态或飞行包线变化的未知但可测量或估计的参数。

非线性系统的 LPV 描述形式不是唯一的,其直接影响控制系统的设计复杂度。现有 LPV 系统的控制器综合一般是基于线性变参数轨迹上的一组线性矩阵不等式(LMI),其控制器可以由 LMI 的解来构造,LMI 的数量直接影响到控制器的计算复杂度,若 LPV 系统具有多胞形结构,由于多胞形是凸集,LPV 系统可以由顶点完全描述^[9],则只需对顶点进行控制器的设计即可。因此凸多胞 LPV 系统是解决大包线控制器设计问题的合适的 LPV 描述形式。

定义 1 仿射参数依赖型 LPV 多胞系统。

由变参数 $\boldsymbol{\rho}(t)$ 描述的多胞系统为

$$\Theta := \text{Co}\{\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_r\} := \left\{ \sum_{i=1}^r \alpha_i \mathbf{S}_i : \alpha_i \geq 0, \sum_{i=1}^r \alpha_i = 1 \right\} \quad (3)$$

式中: $r = 2^N$ 为顶点数; \mathbf{S}_i 为多胞系统的顶点; α_i 为多胞系统的凸分解系数。当 LPV 系统的系统矩阵 $\mathbf{A}(\boldsymbol{\rho}(t))$ 、 $\mathbf{B}(\boldsymbol{\rho}(t))$ 、 $\mathbf{C}(\boldsymbol{\rho}(t))$ 、 $\mathbf{D}(\boldsymbol{\rho}(t))$ 仿射依赖于时变参数向量 $\boldsymbol{\rho}(t)$,且 $\boldsymbol{\rho}(t)$ 变化于多胞 Θ 内,称 LPV 系统为多胞系统。

LPV 多胞系统的系统矩阵分解形式为

$$\begin{bmatrix} \mathbf{A}(\boldsymbol{\rho}(t)) & \mathbf{B}(\boldsymbol{\rho}(t)) \\ \mathbf{C}(\boldsymbol{\rho}(t)) & \mathbf{D}(\boldsymbol{\rho}(t)) \end{bmatrix} := \sum_{i=1}^r a_i(t) \begin{bmatrix} \mathbf{A}_i & \mathbf{B}_i \\ \mathbf{C}_i & \mathbf{D}_i \end{bmatrix} \quad (4)$$

式中: \mathbf{A}_i 、 \mathbf{B}_i 、 \mathbf{C}_i 和 \mathbf{D}_i 为 LPV 系统多胞顶点模型系统矩阵;凸分解系数为变参数的函数,可表示为 $\alpha_i(t) = f(\boldsymbol{\rho}(t))$ 。当变参数轨迹在多胞 Θ 内时,则 LPV 控制器系统矩阵 $\mathbf{A}_k(\boldsymbol{\rho}(t))$ 、 $\mathbf{B}_k(\boldsymbol{\rho}(t))$ 、 $\mathbf{C}_k(\boldsymbol{\rho}(t))$ 和 $\mathbf{D}_k(\boldsymbol{\rho}(t))$ 分解形式为

$$\begin{bmatrix} \mathbf{A}_k(\boldsymbol{\rho}(t)) & \mathbf{B}_k(\boldsymbol{\rho}(t)) \\ \mathbf{C}_k(\boldsymbol{\rho}(t)) & \mathbf{D}_k(\boldsymbol{\rho}(t)) \end{bmatrix} := \sum_{i=1}^r \alpha_i(t) \begin{bmatrix} \mathbf{A}_{ki} & \mathbf{B}_{ki} \\ \mathbf{C}_{ki} & \mathbf{D}_{ki} \end{bmatrix} \quad (5)$$

大包线高超声速飞行器 LPV 建模的核心问题可以归结为:①如何获得合适的多胞 LPV 顶点模型及凸分解系数,在保证建模精度的前提下求解最少的多胞顶点和低维度准确的凸分解系数。顶点数量的减少不仅可以降低建模的保守性同时可以降低控制器的阶数,低维度准确的凸分解系数可降低系统的数据存储量、提高在线增益调度时的计算速度,有利于工程化实现。②如何定量描述所求多胞模型的建模精度,使得对于变参数的任意一固定点,满足如下精度要求:

$$\left\| \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} - \begin{bmatrix} \mathbf{A}(\boldsymbol{\rho}) & \mathbf{B}(\boldsymbol{\rho}) \\ \mathbf{C}(\boldsymbol{\rho}) & \mathbf{D}(\boldsymbol{\rho}) \end{bmatrix} \right\| < \varepsilon \quad (6)$$

式中: ε 为系统允许的建模误差大小,式(6)描述了原系统与所求多胞系统的差异化程度。

2 基于 HOSVD 的 LPV 多胞顶点建模

HOSVD 理论^[10]应用于 LPV 系统可以有效地将系统近似为一系列的 LTI 模型,这些 LTI 模型对应于 LPV 多胞系统的顶点。根据多胞系统的顶点性质和实有界定理,设计控制器时只要对其顶点进行设计即可。因此基于 HOSVD 的 LPV 多胞顶点建模方法是 LPV 控制理论中极其重要的部分,其建模精度与复杂度直接影响到 LPV 控制器的设计。HOSVD 理论是将奇异值分解的思想应用于高维张量数据中的一种降维方法,其理论基础是张量展开。

2.1 HOSVD 理论

定义 2 (张量展开)若 \mathbf{A}^E 为一个 $(I_1 \times I_2 \times \dots \times I_N)$ 维实张量,张量中元素可以表示为 $\mathbf{A}_{i_1 i_2 \dots i_N}^E = \{a_{i_1 i_2 \dots i_N}\}$,则其沿第 n 模展开为矩阵 $\mathbf{A}_{(n)}^E$,且 $a_{i_1 i_2 \dots i_N}$ 出现在矩阵 $\mathbf{A}_{(n)}^E$ 的第 i_n 行第 j 列。

$$j = (i_{n+1} - 1)I_{n+2}I_{n+3} \dots I_N I_1 I_2 \dots I_{n-1} +$$

$$(i_{n+2} - 1)I_{n+3}I_{n+4} \dots I_N I_1 I_2 \dots I_{n-1} + \dots +$$

$$(i_N - 1)I_1 I_2 \dots I_{n-1} + (i_1 - 1)I_2 I_3 \dots I_{n-1} +$$

$$(i_2 - 1)I_3 I_4 \dots I_{n-1} + \dots + i_{n-1}$$

定义 3 (HOSVD)对于一个 $(I_1 \times I_2 \times \dots \times I_N)$ 维实张量 \mathbf{A}^E 可表示为乘积形式:

$$\mathbf{A}^E = \mathbf{S}^E \times_1 \mathbf{U}_{(1)} \times_2 \mathbf{U}_{(2)} \times \cdots \times_N \mathbf{U}_{(N)} = \mathbf{S}^E \underset{n=1}{\otimes} \mathbf{U}_{(n)} \quad (7)$$

式中： $\mathbf{U}_{(n)} = [\mathbf{u}_{1_n}^{(n)} \quad \mathbf{u}_{2_n}^{(n)} \quad \cdots \quad \mathbf{u}_{I_n}^{(n)}]$ 为 $(I_n \times I_n)$ 维矩阵， $n = 1, 2, \dots, N$ ， $\mathbf{u}_{i_n}^{(n)}$ 为第 i 个 n 模奇异向量； \mathbf{S}^E 为一个 $(I_1 \times I_2 \times \cdots \times I_N)$ 维实张量，且其子张量满足条件：

$$\begin{cases} \langle \mathbf{S}_{i_n=\alpha}^E, \mathbf{S}_{i_n=\beta}^E \rangle = 0 & \alpha \neq \beta \\ \|\mathbf{S}_{i_n=1}^E\|_F \geq \|\mathbf{S}_{i_n=2}^E\|_F \geq \cdots \geq \|\mathbf{S}_{i_n=I_N}^E\|_F \geq 0 \end{cases}$$

弗罗贝尼乌斯范数定义为 $\|\mathbf{S}_{i_n=i}^E\|_F = (\langle \mathbf{S}_{i_n=i}^E, \mathbf{S}_{i_n=i}^E \rangle)^{\frac{1}{2}}$ ，其可用 $\sigma_i^{(n)}$ 表示，是张量 \mathbf{A}^E 的 n 模奇异值。

定理 1 对一个 n 秩为 $R_n (1 \leq n \leq N)$ 的张量 \mathbf{A}^E 进行 HOSVD，将分解后得到的 n 模奇异值 $\sigma_{I_n+1}^{(n)}, \sigma_{I_n+2}^{(n)}, \dots, \sigma_{R_n}^{(n)}$ 的相应元素变为 0，得到一个新的张量 $\hat{\mathbf{A}}^E$ ，则对于 $I'_n < R_n$ ，有

$$\gamma = \|\mathbf{A}^E - \hat{\mathbf{A}}^E\|^2 \leq \sum_{i_1=I'_1+1}^{R_1} (\sigma_{i_1}^{(1)})^2 + \sum_{i_2=I'_2+1}^{R_2} (\sigma_{i_2}^{(2)})^2 + \cdots + \sum_{i_N=I'_N+1}^{R_N} (\sigma_{i_N}^{(N)})^2 \quad (8)$$

HOSVD 是基于张量展开应用奇异值分解理论来解决高维数据主成分提取问题的方法。相较于仿射参数依赖型 LPV 系统的经验选取法，HOSVD 方法的误差来源于舍弃的奇异值与特征向量，误差有确切的计算方法，可由式 (8) 直接计算出来，有利于控制误差的大小，同时也能够控制顶点数量降低控制器的阶数，更利于工程化实现。

2.2 基于网格化张量的 HOSVD 方法改进

高超声速飞行器气动参数非线性剧烈变化，很难用状态转换法以及函数替换法写出 LPV 模型的具体表达式。相比而言，Jacobian 线性化法更适合于求解高超声速飞行器的 LPV 模型^[11]，其理论基础是一阶 Taylor 展开。

在整个飞行包线内对 $\boldsymbol{\rho}(t)$ 中每一个变参数进行网格化划分，设划分个数分别为 I_1, I_2, \dots, I_N ，则在网格化的参数空间内，可得到变参数的网格化张量形式 $\boldsymbol{\rho}^E \in \mathbf{R}^{I_1 \times I_2 \times \cdots \times I_N}$ 。设非线性系统式 (1) 在全包线内存在一系列平衡点 $\mathbf{x}_e(\boldsymbol{\rho}_e^E)$ ， $\mathbf{u}_e(\boldsymbol{\rho}_e^E)$ ，满足 $f(\mathbf{x}_e(\boldsymbol{\rho}_e^E), \mathbf{u}_e(\boldsymbol{\rho}_e^E)) = 0$ ，则非线性系统可写成网格化线性模型组的形式：

$$\begin{bmatrix} \dot{\mathbf{x}}_\delta \\ \mathbf{y}_\delta \end{bmatrix} = \begin{bmatrix} \mathbf{A}(\boldsymbol{\rho}^E) & \mathbf{B}(\boldsymbol{\rho}^E) \\ \mathbf{C}(\boldsymbol{\rho}^E) & \mathbf{D}(\boldsymbol{\rho}^E) \end{bmatrix} \begin{bmatrix} \mathbf{x}_\delta \\ \mathbf{u}_\delta \end{bmatrix} \quad (9)$$

式中：

$$\begin{cases} \mathbf{x}_\delta = \mathbf{x}(t) - \mathbf{x}_e(\boldsymbol{\rho}^E) \\ \mathbf{u}_\delta = \mathbf{u}(t) - \mathbf{u}_e(\boldsymbol{\rho}^E) \end{cases} \quad (10)$$

其中系统状态矩阵为方程偏导数在平衡点处的值：

$$\begin{cases} \mathbf{A}(\boldsymbol{\rho}^E) = \frac{\partial f_1}{\partial \mathbf{x}} [\mathbf{x}_e(\boldsymbol{\rho}^E), \mathbf{u}_e(\boldsymbol{\rho}^E)] \\ \mathbf{B}(\boldsymbol{\rho}^E) = \frac{\partial f_1}{\partial \mathbf{u}} [\mathbf{x}_e(\boldsymbol{\rho}^E), \mathbf{u}_e(\boldsymbol{\rho}^E)] \\ \mathbf{C}(\boldsymbol{\rho}^E) = \frac{\partial f_2}{\partial \mathbf{x}} [\mathbf{x}_e(\boldsymbol{\rho}^E), \mathbf{u}_e(\boldsymbol{\rho}^E)] \\ \mathbf{D}(\boldsymbol{\rho}^E) = \frac{\partial f_2}{\partial \mathbf{u}} [\mathbf{x}_e(\boldsymbol{\rho}^E), \mathbf{u}_e(\boldsymbol{\rho}^E)] \end{cases} \quad (11)$$

则在每一个固定点 $\boldsymbol{\rho}$ ，系统式 (9) 描述了非线性系统式 (1) 在该点的局部特性。

在网格化的参数空间内，一系列平衡点线性化得到张量形式的系统描述，其中：

$$\begin{aligned} \mathbf{A}^E &\in \mathbf{R}^{I_1 \times I_2 \times \cdots \times I_N \times m \times m} & \mathbf{B}^E &\in \mathbf{R}^{I_1 \times I_2 \times \cdots \times I_N \times m \times m} \\ \mathbf{C}^E &\in \mathbf{R}^{l_1 \times l_2 \times \cdots \times l_N \times l \times m} & \mathbf{D}^E &\in \mathbf{R}^{l_1 \times l_2 \times \cdots \times l_N \times l \times m} \end{aligned}$$

将得到的张量形式的系统描述重新组合成一个张量 $\mathbf{G}^E \in \mathbf{R}^{l_1 \times l_2 \times \cdots \times l_N \times (m+l) \times (m+n)}$ ， \mathbf{G}^E 中的每一个元素均为与变参数相关的变量：

$$\mathbf{G}^E = \begin{bmatrix} \mathbf{A}^E(\boldsymbol{\rho}) & \mathbf{B}^E(\boldsymbol{\rho}) \\ \mathbf{C}^E(\boldsymbol{\rho}) & \mathbf{D}^E(\boldsymbol{\rho}) \end{bmatrix} \quad (12)$$

对张量 \mathbf{G}^E 进行高阶奇异值分解，分别舍弃前 N (变参数的个数) 模中较小的奇异值和对应的特征向量，设 $1 - N$ 模的奇异值中保留的个数分别为 a_1, a_2, \dots, a_N ，得到的顶点个数为 $M = a_1 \times a_2 \times \cdots \times a_N$ ，可以得到 LPV 系统多胞形描述为

$$\begin{bmatrix} \dot{\mathbf{x}}(t) \\ \mathbf{y}(t) \end{bmatrix} = \mathbf{S}_n^E \underset{n=1}{\otimes} \boldsymbol{\omega}_n(\boldsymbol{\rho}_n(t)) \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{u}(t) \end{bmatrix} \quad (13)$$

式 (13) 也可表示为

$$\dot{\mathbf{x}}(t) = \sum_{i_1=1}^{a_1} \sum_{i_2=1}^{a_2} \cdots \sum_{i_N=1}^{a_N} \boldsymbol{\omega}_{1,i_1}(\boldsymbol{\rho}_1(t)) \cdots \boldsymbol{\omega}_{N,i_N}(\boldsymbol{\rho}_N(t)) \cdot (\mathbf{A}_{i_1,i_2,\dots,i_N} \mathbf{x}(t) + \mathbf{B}_{i_1,i_2,\dots,i_N} \mathbf{u}(t)) \quad (14)$$

式中：权重函数 $\boldsymbol{\omega}_{n,i}$ 应满足条件

$$\begin{cases} \forall n, i, \boldsymbol{\rho}_n(t) : \boldsymbol{\omega}_{n,i}(\boldsymbol{\rho}_n(t)) \in [0, 1] \\ \forall n, \boldsymbol{\rho}_n(t) : \sum_{i=1}^{I_N} \boldsymbol{\omega}_{n,i}(\boldsymbol{\rho}_n(t)) = 1 \end{cases} \quad (15)$$

由式 (13) 和式 (14) 可得到顶点张量 \mathbf{S}_n^E 与所求顶点矩阵的对应关系，式 (3) 中凸分解系数 α 与式 (13) 中的权重函数 ω 的关系为

$$\alpha_{i_1, i_2, \dots, i_N} = \boldsymbol{\omega}_{1, i_1}(\boldsymbol{\rho}_1(t)) \boldsymbol{\omega}_{2, i_2}(\boldsymbol{\rho}_2(t)) \cdots \boldsymbol{\omega}_{N, i_N}(\boldsymbol{\rho}_N(t))$$

在误差允许的范围内离散的系统矩阵可由求解出的顶点系统和权重矩阵 \mathbf{U}_n 还原

$$\mathbf{G}^E \approx \underset{\gamma}{\mathbf{S}}_n^E \underset{n}{\otimes} \mathbf{U}_n \quad (16)$$

式中： \mathbf{U}_n 包含每一个变参数对应的权重函数 $\boldsymbol{\omega}_n$ ，误差 γ 可表示为

$$\gamma = (\|S^E - S \otimes_n U_n\|_{L_2})^2 \leq \sum_k \sigma_k^2 \quad (17)$$

与式(3)仿射参数依赖型 LPV 多胞模型的结构类似,基于 HOSVD 建立多胞模型的方法也是由顶点的凸分解得到整个多胞模型,两者的区别在于多胞顶点和凸分解系数的求取方法。相较于仿射参数依赖型 LPV 模型选取变参数上下界组合的方法,本文提出的方法无需依赖于设计者的经验,同时也没有复杂的数据拟合过程,参数的设定只涉及到每一模保留奇异值的个数,在误差允许的范围内选定顶点个数,整个设计过程可以很容易地实现自动化设计。

综上,基于 HOSVD 方法的多胞 LPV 顶点系统的设计流程如图 1 所示。

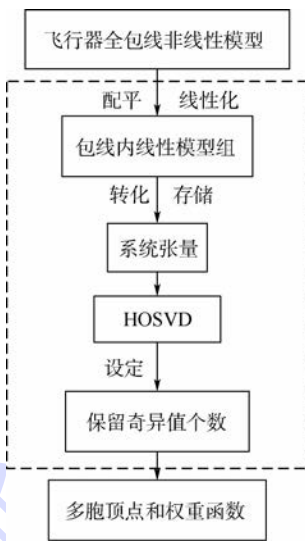


图 1 基于 HOSVD 的 LPV 多胞系统设计流程
Fig. 1 Flowchart of polytopic LPV system design based on HOSVD

根据上述分析,如图 1 所示,可以得到基于 HOSVD 的高超声速飞行器 LPV 系统多胞分解的一般步骤:

- 1) 针对飞行器非线性模型,给定变参数、变量范围、变量划分网格数,在网格化的飞行包线内进行线性化得到线性模型组。
- 2) 将离散化的系统数据转化为可用于 HOSVD 方法的张量 G^E 形式。
- 3) 对 G^E 应用 HOSVD 方法,在每一模中舍弃较小的奇异值以及与之相对应的奇异值向量,得到多胞顶点系统。
- 4) 由 HOSVD 得到的权重矩阵 U_n 的第 i_n ($i_n = 1, 2, \dots, I_n$) 个列向量 u_n 便确定了变量 $\rho_n(t)$ 的权重函数 $\omega_{n,i_n}(\rho_n(t))$,进而得到凸分解系数。

建模的误差大小与 LTI 顶点的个数成反比,选择保留的奇异值数量越多得到的顶点个数越

多,模型越精确。一般情况下,为了便于控制器的分析与设计,同时降低控制器阶数,根据系统可容忍的误差大小,选择有限个 LTI 顶点系统。

3 鲁棒 LPV 自增益控制器设计

3.1 多胞 LPV 系统鲁棒控制器设计

对于某型航天飞机^[12]六自由度非线性模型,选取再入包线:高度 H 为 30 ~ 40 km,马赫数 Ma 为 3 ~ 5。由于航天飞机采用倾斜转弯(Bank-To-Turn, BTT)控制方式,侧滑较小,气动特性除高度、马赫数外,受迎角 α 影响最大,因此选取 H 、 Ma 和 α 作为 LPV 系统的变参数。对变参数进行网格化的划分,网格的划分依据对象的特性与包线的范围,将变参数分别划分 6、3、7 等份。本文选取 $C = I, D = 0$,得到网格化的张量形式 $G^E = [A^E \ B^E]$,对得到的网格化的张量 $G^E \in \mathbf{R}^{6 \times 3 \times 7 \times 6 \times 9}$ 进行高阶奇异值分解,得到的 HOSVD 结果如表 1 所示。

表 1 HOSVD 结果

Table 1 Results of HOSVD

奇异值阶数	一维展开	二维展开	三维展开
一阶	81.5175	81.5185	81.4868
二阶	0.4546	0.2115	2.2289
三阶	0.0006	0.0795	0.4736
四阶	0.0003		0.1174
五阶	0.0001		0.0852
六阶	0.00008		0.0151

本文设定系统顶点建模误差小于 0.1,根据表 1 中的分解结果,每一维上分别保留 2、2、3 个奇异值,通过误差计算式(8)得到实际的误差为 0.028,满足设计要求。在整个顶点选取过程中,只需设定每一维上保留较大奇异值的个数,实现了本文对系统自动选取顶点的要求,并得到一个 12 个顶点的多胞 LPV 系统模型。得到各变参数对应的权重函数曲线如图 2 所示。

图 2 中的曲线描述了权重函数随变参数的实时变化情况,高度、马赫数、迎角对应的权重函数维数与每一维上保留的奇异值个数正好对应。

本文基于上述 HOSVD 过程得到的顶点系统与凸分解系数,对各顶点设计控制器。在考虑系统参数不确定性的情况下,可将每个顶点系统模型描述为增广系统状态空间形式^[13],对增广系统应用 H_∞ 鲁棒控制中的混合灵敏度设计方法^[14],求使系统内稳定且满足闭环性能指标的控制器的。基于此方法设计高超声速飞行器的自增益调度控制系统,控制系统结构如图 3 所示,变参数反馈给

在线调参模块可由变参数对应的权重函数求解得到凸分解系数,实时选择各顶点控制器所占权重,从而得到全包线控制器。

H_∞ 混合灵敏度设计的难点在于如何将系统的闭环性能指标要求表示在尽量简单的加权函数中。本文依据回路成形理论^[15]选择期望的回路传递函数为 ω_s/s ,结合闭环系统性能的要求对回路进行频域整形,选取合适的混合灵敏度加权函数阵,得到多胞顶点的鲁棒控制器。高超声速飞行器三通道的混合灵敏度加权函数阵选取为

$$W_s(s) = \text{diag}\left(\frac{0.8(s+3.5)}{s+0.0035}, \frac{0.8(s+3.5)}{s+0.0035}, \frac{0.8(s+3.5)}{s+0.0035}\right)$$

$$W_u(s) = \text{diag}(0.005, 0.005, 0.005)$$

当LPV系统的顶点控制器确定以后,LPV控制器可由顶点的凸组合求得,全包线控制器形式如式(5)所示,其凸组合是由变参数 $\rho(t)$ 在多胞 Θ 中的位置决定的。

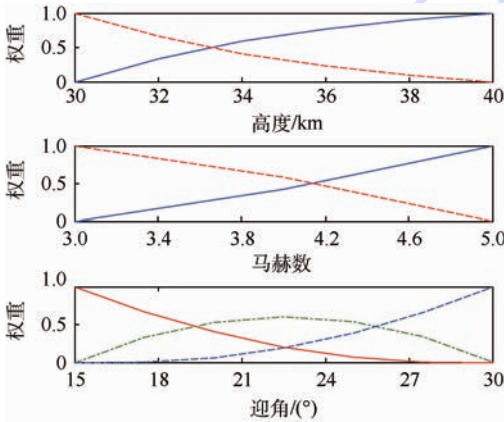
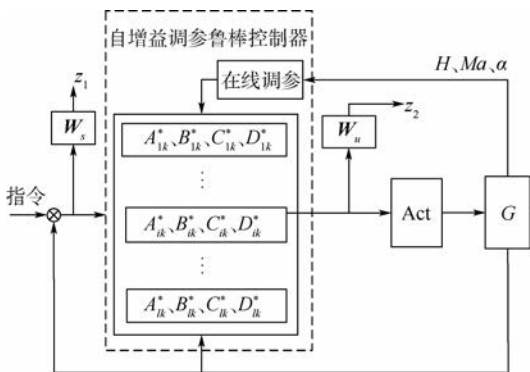


图2 变参数的权重函数

Fig. 2 Weighting functions of varying-parameter



z_1, z_2 —性能指标; Act—作动器; G—被控对象。

图3 基于鲁棒混合灵敏度增益调度控制系统

Fig. 3 Gain scheduling control system based on robust mixed sensitivity

3.2 仿真与分析

在飞行包线区域内选择一状态点 $H=35\text{ km}$, $Ma=4$, $\alpha=20^\circ$,在三通道加入姿态阶跃指令信号进行仿真验证,仿真结果如图4所示。

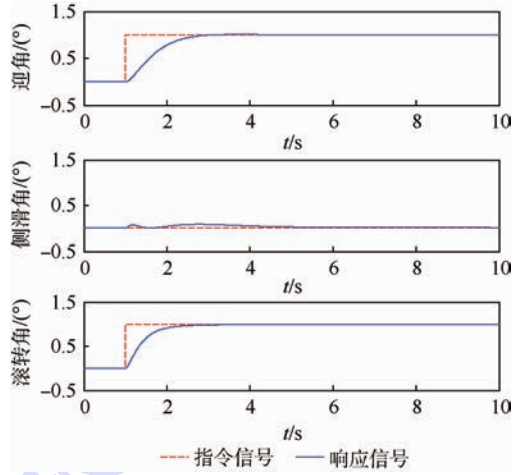


图4 三通道阶跃指令响应

Fig. 4 Three channels' step order response curves

仿真结果表明,在不考虑不确定性的情况下控制器能精确跟踪迎角、侧滑角、滚转角阶跃指令信号,三通道姿态角能够在2s之内跟踪指令信号,响应速度较快,且没有超调量。同时可看出所设计的控制器受通道之间耦合的影响较小。当考虑如表2所示的气动力系数、气动力矩系数、动压、转动惯量、质量的不确定性,对闭环系统进行100次蒙特卡罗仿真,仿真结果如图5所示。

表2 不确定性
Table 2 Uncertainty

不确定性项	不确定性范围
气动力系数	-20% ~ +20%
气动力矩系数	-20% ~ +20%
动压	-30% ~ +30%
转动惯量	-5% ~ +5%
质量	-3% ~ +3%

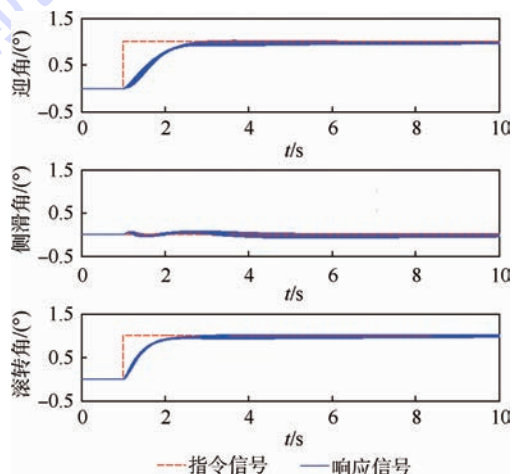


图5 100次蒙特卡罗仿真结果

Fig. 5 Results of 100 Monte-Carlo simulation

从仿真结果可以看出,在强不确定性因素的影响下,100次蒙特卡罗仿真的控制器跟踪性能与标称情况下相近,不确定性对控制器的动态性能影响较小,说明本文设计的LPV变增益控制器对不确定性系统具有良好的跟踪性能和鲁棒性。

根据航天飞机的飞行指令,在包线区域内(见图6)对系统进行仿真,迎角响应如图7所示。

图7中的仿真结果表明在整个飞行包线内能够精确跟踪姿态指令,验证了控制器的增益调度能力。图8中的曲线对应于12个多胞顶点的凸分解系数,其描述了整个在线增益调度过程中各

顶点所占实时状态点的权重,以及顶点控制器所占全包线控制器的权重。

4 结论

本文针对航天飞机再入段大包线对象提出了一种基于网格化张量及HOSVD的LPV多胞模型建模方法,并设计了LPV鲁棒控制器,仿真验证表明:

1) 基于网格化张量的HOSVD方法能够完成LPV多胞模型的建模,在保证建模精度的前提下控制顶点数量,并给出误差的计算方法。

2) 相较于传统的建模方法,基于网格化张量及HOSVD方法能够降低对设计人员经验的依赖,大幅提高建模过程的自动化程度,有利于工程化应用。

3) 基于回路成形理论设计的鲁棒混合灵敏度控制器能够实现在线增益调度,并具有很好的鲁棒性。

参考文献 (References)

- [1] 白辰,任章,樊焱,等. 基于模糊神经网络扰动观测器的RLV动态逆再入姿态控制[J]. 中南大学学报(自然科学版),2013,7(1):58-62.
BAI C, REN Z, FAN Y, et al. Dynamic inversion control for RLV reentry attitude based on fuzzy-neural disturbance observer [J]. Journal of Central South University (Natural Science), 2013,7(1):58-62 (in Chinese).
- [2] HALL C E, SHTESSEL Y B. Sliding mode disturbance observer-based control for a reusable launch vehicle [J]. Journal of Guidance and Dynamics, 2006,29(6):1315-1328.
- [3] LIND R. Linear parameter-varying modeling and control of structural dynamics with aerothermoelastic effects [J]. Journal of Guidance Control and Dynamics, 2002,25(4):733-739.
- [4] YUSUKE Y, TAKASHI S. Attitude control of spacecraft with VSCMGs using LPV modeling technique [C] // SICE Annual Conference. Tokyo: Waseda University, 2011:2678-2683.
- [5] KUMAR A. Convex modeling techniques for aircraft control [D]. Virginia: Virginia Polytechnic Institute and State University, 2000:1-5.
- [6] PETRES Z. Polytopic decomposition of linear parameter-varying models by tensor-product model transformation [D]. Budapest: Budapest University of Technology and Economics, 2006:25-89.
- [7] 秦伟伟,郑志强,刘刚,等. 高超声速飞行器的LPV鲁棒变增益控制 [J]. 系统工程与电子技术, 2011,33(6):1327-1331.
QIN W W, ZHENG Z Q, LIU G, et al. Robust variable gain control for hypersonic vehicles based on LPV [J]. Systems Engineering and Electronics, 2011,33(6):1327-1331 (in Chinese).

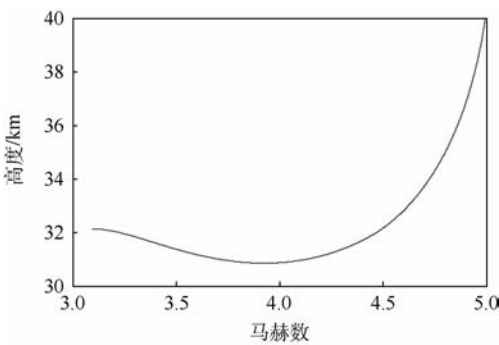


图6 飞行包线

Fig.6 Flight envelope

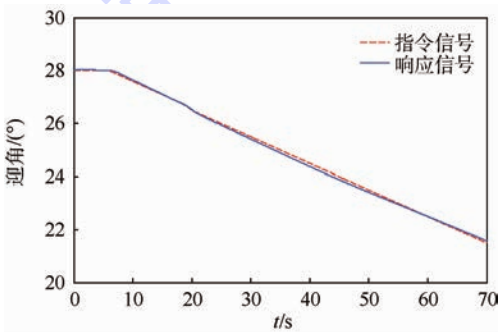


图7 迎角指令跟踪仿真

Fig.7 Angle of attack command tracking simulation

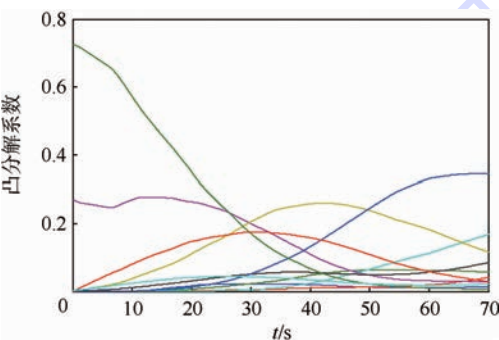


图8 凸分解系数仿真结果

Fig.8 Results of convex parameter decomposition simulation

- [8] 王明昊,刘刚,杨述华.高超声速飞行器的多胞 LPV 系统控制器设计[J].空间控制技术与应用,2013,39(1):15-22.
WANG M H, LIU G, YANG S H. Polytopic-LPV-system-based control design for hypersonic vehicle [J]. Aerospace Control and Application, 2013, 39(1): 15-22 (in Chinese).
- [9] HE C F, YANG L Y, WANG Z C, et al. Linear parameter-varying attitude controller design for a reusable launch vehicle during reentry [C] // Proceedings of 2014 IEEE Chinese Guidance, Navigation and Control Conference. Piscataway, NJ: IEEE Press, 2014: 2723-2728.
- [10] RAJWADE A, RANGARAJAN A, BANERJEE A. Image denoising using the higher order singular value decomposition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(4): 849-862.
- [11] OHARA A, YAMAGUCHI Y, MORITO T. LPV modeling and gain scheduled control of reentry vehicle in approach and landing phase [C] // AIAA Guidance, Navigation, Control and Conference and Exhibit. Reston: AIAA, 2001: 1-11.
- [12] Aerodynamic design data book, Volume 1M: Orbiter vehicle STS-1: SD72-SH-0060 [R]. [S.l.]: Rockwell International, 1980.
- [13] 王振超,杨凌宇,张晶,等.面向大包线导弹的分回路 LPV 姿态控制器设计[J].航空兵器,2015,1(2):14-20.
WANG Z C, YANG L Y, ZHANG J, et al. Loop separate LPV attitude controller design for large envelope missile [J]. Aero Weaponry, 2015, 1(2): 14-20 (in Chinese).
- [14] 夏文艳,高嵩,王洁,等.基于 H_∞ 混合灵敏度的导弹解耦控制器设计方法[J].电子设计工程,2009,17(11):43-45.
XIA W Y, GAO S, WANG J, et al. Decoupling based on the H_∞ mixed sensitivity controller for missile design method [J]. Electronic and Engineering, 2009, 17(11): 43-45 (in Chinese).
- [15] 傅彩芬,谭文,刘吉臻.基于回路成形的鲁棒增益调度控制器设计[J].信息与控制,2005,34(2):152-156.
FU C F, TAN W, LIU J Z. Robust gain scheduling controller based on loop shaping [J]. Information and Control, 2005, 34(2): 152-156 (in Chinese).

作者简介:

孙斌 男,硕士研究生。主要研究方向:先进飞行器控制。

E-mail: sun1990bin@qq.com

Tel.: 010-82316873

杨凌宇 男,博士,副教授,硕士生导师。主要研究方向:先进飞行器控制,容错控制,先进布局飞机多操纵控制方法。

E-mail: yanglingyu@buaa.edu.cn

Tel.: 010-82316873

Robust LPV control design based on HOSVD

SUN Bin, YANG Lingyu*, ZHANG Jing

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

Abstract: Considering the trade-off between the accuracy and complexity of the linear parameter varying (LPV) modeling of hypersonic vehicles, a novel polytopic LPV modelling technique based on grid tensors and higher-order singular value decomposition (HOSVD) was developed. Firstly, through Jacobian linearization, the grid tensor description of a hypersonic vehicle with large flight envelopes was introduced, and a polytopic LPV model was constructed using HOSVD. The resulting model was described by a linear combination of vertices of linear time invariant (LTI) models and corresponding weighting functions. The discarded singular values were treated as a modeling error index. A robust variable gain controller which combines a robust controller and gain scheduling technique was then designed for the 6 degree-of-freedom nonlinear model of the hypersonic vehicle. Simulation results demonstrate that the proposed method not only reduces the computational complexity, but also guarantees the modeling accuracy. Furthermore, under the proposed controller, the closed loop system can track the attitude commands rapidly, and good stability and robustness are achieved.

Key words: hypersonic vehicle; flight envelope; polytopic linear parameter varying (LPV) system; robust variable gain control; higher-order singular value decomposition (HOSVD)

Received: 2015-07-20; Accepted: 2015-10-16; Published online: 2015-12-17 10:41

URL: www.cnki.net/kcms/detail/11.2625.V.20151217.1041.007.html

Foundation items: National Natural Science Foundation of China (61273099); Aeronautical Science Foundation of China (20135851041)

* Corresponding author. Tel.: 010-82316873 E-mail: yanglingyu@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0430

基于拟态电势能的飞行器航路规划方法



何仁珂, 魏瑞轩*, 张启瑞, 许卓凡

(空军工程大学 航空航天工程学院, 西安 710038)

摘 要: 航路规划是现代各类飞行器,特别是无人机(UAV)安全飞行和完成任务的关键要素。对于复杂威胁环境和高维空间的航路规划问题,传统规划方法在规划速度、航路安全性、算法适用性等方面存在一定的应用局限。分析了电荷在电场中移动引起电势能变化的特点规律及电势场分布特性,模拟电势场理论对飞行环境进行威胁建模,建立基于电势威胁场引导的航路节点概率选择机制及安全性评价准则。在此基础上,构建基于拟态电势能的随机采点扩展式航路规划方法。通过与传统航路规划方法的对比仿真实验表明,运用拟态电势能进行航路规划,能够显著缩短路径长度和计算时间,提高规划航路的安全性,对于航路规划的应用很有价值。

关键词: 无人机(UAV); 航路规划; 拟态电势能; 概率选择; 航路安全性

中图分类号: V279+.2

文献标识码: A **文章编号:** 1001-5965(2016)07-1543-07

在无人化技术和装备高速发展的背景下,航路规划成为各类飞行器保证安全飞行和顺利完成任务的关键环节。随着空域中飞行器数量的增加和任务环境不确定性因素趋于复杂,影响飞行安全的威胁因素不断增多,加之飞行器协同任务规划的需求增大^[1],迫切需要能够进行高效、安全航路规划的方法。

传统的航路规划方法可大致分为4类。第1类是基于拓扑图^[2]的方法(如Voronoi图、Laguerre图等)。该方法属于对环境的几何建模,解析过程直观简单,计算复杂度受障碍数量影响较大。第2类是人工势场^[3],属于一种拟物算法,算法的固有特点造成了诸如:落入势场平衡“陷阱”区域、由于相邻障碍间强势场影响造成搜索盲区等一系列问题。第3类是栅格法^[4](如A*、D*算法等),可以通过代价地图计算可行路径,计算代价较大,且仅限于低维离散空间。第4类是

智能优化算法^[5-6](如遗传算法、蚁群算法等),这些寻优方法在解决路径规划问题方面有着优良的性能,其解与迭代次数有很大关系,且易陷入局部最优。近年来提出了一些基于物理学原理的优化算法,如万有引力算法^[7]、拟态物理学方法^[8-9]、类电磁算法^[10]、中心力算法^[11]等。这些算法利用物理规律表征规划要素间的关系,在解决包括航路规划问题在内的优化问题中表现出一定的优势。但算法复杂度大,计算代价高,对于复杂环境和高维空间的快速安全航路规划问题并不适用。

以快速扩展随机树(Rapidly exploring Random Tree, RRT)算法^[12-13]为代表的基于采样的方法为解决高维航路规划的效率和可行性问题提供了一种有效方案。但由于缺少启发要素,基础RRT算法的路径生成带有很大的随机性。文献[14]利用A*算法的代价函数改进RRT算法,但无法避免局部最小问题。文献[15-16]用地形代

收稿日期: 2015-06-29; 录用日期: 2015-09-18; 网络出版时间: 2015-10-30 16:54

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20151030.1654.008.html

基金项目: 航空科学基金(20135896027)

* 通讯作者: Tel.: 029-84787729 E-mail: rxwei369@soho.com

引用格式: 何仁珂, 魏瑞轩, 张启瑞, 等. 基于拟态电势能的飞行器航路规划方法[J]. 北京航空航天大学学报, 2016, 42(7): 1543-1549. HE R K, WEI R X, ZHANG Q R, et al. Mimetic electric potential energy motion planning algorithm for aircraft [J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1543-1549 (in Chinese).

价函数和几何方法优化 RRT 算法生成的路径。文献[17]从减小计算代价角度对算法进行了改进。Karaman 和 Frazzoli^[18]证明了 RRT 算法得到最优解的概率为 0, 并提出改进的 RRT* 算法可解出最优结果。文献[19]提出用滤波方法优化路径生成。由于这些改进方法没有考虑环境威胁的强度分布, 且未综合考虑航路规划的效率和安全性。为此, 本文将电势场理论引入飞行环境的威胁建模, 建立基于电势威胁场引导的航路点生成机制, 进而构建基于拟态电势能导向的随机采点扩展式航路规划方法, 探索基于航路点导向扩展来提高航路规划效率和安全性新机理。

1 基于拟态电势场的飞行环境建模

飞行航路规划的主要目的是为飞行器确定出从起始点到目标点的安全可飞航路, 需要考虑的主要问题是航路上威胁的规避、航路规划的快速性等问题, 以确保规划的航路能够满足任务的实时性和飞行安全性等方面的要求。建立有效描述飞行区域中威胁情况的环境模型, 是实现高效航路规划的基础。

飞行环境中的威胁障碍主要有: 高炮、防空导弹、高山、建筑物、禁飞区等。非攻击性障碍可根据其特性划定界限, 以防止飞行器进入。对于攻击性威胁, 传统环境建模方法常把威胁模型简单建为以火力范围确定半径的圆形或球形, 或建立高斯随机威胁场等。这些方法均无法有效描述威胁强度的变化。防空武器的探测杀伤概率与目标距离有关, 同时考虑到飞行器物理限制和不确定客观因素, 越靠近威胁障碍, 威胁性越大, 即飞行威胁度与相对距离呈正相关。基于这种特性, 可引入模拟电势场来表征威胁强度。

根据电势叠加原理, 在二维坐标直角坐标系中, 设 P 点坐标为 (x, y) , 电荷 q_i 的坐标为 (x_i, y_i) , 则电荷 q_i 在点 P 处的电势为

$$V_P = \sum_{i=1}^n \frac{1}{4\pi\epsilon_0} \cdot \frac{q_i}{\sqrt{(x-x_i)^2 + (y-y_i)^2}} \quad (1)$$

式中: ϵ_0 为真空电容率。

电场强度为

$$\mathbf{E} = -\nabla V = -\mathbf{a}_x \frac{\partial V}{\partial x} - \mathbf{a}_y \frac{\partial V}{\partial y} \quad (2)$$

式中: ∇V 为电势梯度; \mathbf{a}_x 和 \mathbf{a}_y 分别表示沿 x 和 y 轴的正方向。

电场强度是描述电场强弱和电场力的性质的物理量。电场强度的大小取决于激发电场的电荷, 与电场中的受力电荷无关^[20]。模拟电场分

布, 设各威胁源为同种正电荷, 坐标为 (x_i, y_i) , 威胁源 o 在 P 点形成的电势为

$$V_o = \sum_{i=1}^n V_{oi} = \sum_{i=1}^n \frac{q_{oi}}{4\pi\epsilon_0 r_i} \quad r_i \in (0, r_1) \quad (3)$$

式中: r_i 为第 i 个防空阵地与某一点的距离。

由于防区外火力杀伤概率为零, 设 $0 < r_i < r_1$ 为火力覆盖区。设飞行器为受力试探电荷, 不影响电势场分布。引入目标点后, 电势场变化为

$$V_{\text{all}} = \sum_{i=1}^n \frac{q_{oi}}{4\pi\epsilon_0 r_i} - \frac{q_t}{4\pi\epsilon_0 r_t} \quad r_t \in (0, \infty) \quad (4)$$

2 基于拟态电势能引导的航路扩展方法

基于快速扩展随机树思想的航路规划方法, 虽然具有快速向未知区域的搜索倾向, 但随机扩展必然带有很大盲目性, 生成的路径随机性较强。根据拟态电势场的飞行环境建模, 通过引入基于拟态电势能的节点引导机制, 可有效启发路径节点选择, 有效解决随机性带来的盲目扩展问题。

2.1 拟态电势能引导的航路点概率选择原理

飞行器在威胁势场中飞行, 在一定区域范围内受威胁和目标电势的叠加影响。即试探电荷在电场中移动, 受电场力作用而做功, 引起电势能变化。根据电势定义, 对于叠加电场, 在威胁势场某一路径上任意两点 i, j 间有

$$\Delta W_{i,j} = q_0 (V_{\text{all}}(i) - V_{\text{all}}(j)) \quad (5)$$

库仑定律表明, 静电力做功与路径无关, 是保守力, 所以静电场是保守场。如图 1 所示, 做功大小由起点 x_{init} 和终点 x_{goal} 相对电势值决定, 电势差为 ΔV , 且 $\sum_{i=1}^n \Delta V_{i+} + \sum_{j=1}^n \Delta V_{j-} = 0$, 增加的电势差 ΔV_+ 与减小的电势差 ΔV_- 总和相等, 同时

$$W_{ij} = qV_{ij} \quad (6)$$

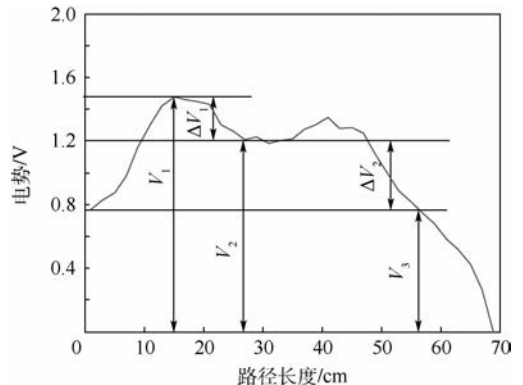


图 1 电势与路径长度示意图

Fig. 1 Schematic diagram of electric potential and path length

因此,电场力做功问题只限于对起点和终点的反映而缺乏对路径的准确描述。基于此,为有效表示路径与功之间的联系,只考虑电势能改变量为正的节点,并引入 δ 为路径比例小量,保证做功相等情况下取最短的路径。设路径长度为 l , 做功为 $w(p)$, 则一条路径的机械功表示为

$$W(p) = q_0 \sum_{i,j=1}^n (V_{\text{all}}(i) - V_{\text{all}}(j)) + \delta l \quad (7)$$

高斯核函数^[21] 可以被认为是在正态分布假设条件下,同类信息间度量距离的工具,在数据挖掘等算法中应用广泛。在威胁势场中,本文将高斯核函数与电势能特性相结合,建立航路节点选择机制,有效启发路径趋向目标点生成。为防止出现局部最小现象,以一定概率保留背离目标方向的点。具体如下:

对于高斯核函数

$$f(x) = \exp\left(-\frac{x(i) - x(j)}{2k^2}\right) \quad (8)$$

式中: k 为指定参数; 设 $x(i) - x(j)$ 为 $V_{\text{all}}(i) - V_{\text{all}}(j)$ 。当 $x(i) - x(j) > 0$, 概率按式(8)计算; 当 $x(i) - x(j) \leq 0$ 时,表明下一状态点朝向目标点生成,概率为 1。

设随机生成 n 条路径,每条路径有 m 个节点。对于第 i 条路径的生成概率 p_i 有

$$p_i = \prod_{j=1}^m (p_{j,\text{rand}}^i) (p_{j,\text{select}}^i) \quad (9)$$

式中: $p_{j,\text{rand}}^i$ 为随机生成点的概率; $p_{j,\text{select}}^i$ 为选择概率。对于 n 条路径来说,每个父结点能随机扩展出下一节点的概率是相同的,都为 $1/n$, 则

$$p_i = \frac{1}{n^m} \prod_{j=1}^m (p_{j,\text{select}}^i) \quad (10)$$

代入式(8)得 $p_i = \frac{1}{n^m} \prod_{j=1}^m f(x_j)$, 即

$$p_i = \frac{1}{n^m} \exp\left(-\frac{\sum \Delta V_j}{(2k^2)^m}\right) \quad (11)$$

将式(7)代入式(11),路径比例小量忽略不计,则

$$p_i = \frac{1}{n^m} \exp\left[-\frac{W(p_i)}{q_0(2k^2)^m}\right] \quad (12)$$

那么可以得出以下结论: $W(p_i)$ 越小,则概率 p_i 越大。即具有最少电势能增量的路径搜索到目标点的概率最大。

2.2 简化运动模型描述

为使规划的路径基本符合飞行器运动约束的要求,本文将位姿点信息融入规划算法。设飞行器的位姿为

$$q(k) = (x_k, y_k, \theta_k) \quad (13)$$

在速度方向上姿态角 θ_k 表示为

$$\tan \theta_k = V_{y_k} / V_{x_k} \quad (14)$$

下一相邻位姿点

$$q(k+1) = (x_{k+1}, y_{k+1}, \theta_{k+1}) \quad (15)$$

设速度 V 不变,单位步长为 ρ , 则时间周期为

$\Delta k = \rho / V$, 下一步位置坐标为

$$\begin{cases} x_{k+1} = x_k + V \cos \theta \times \Delta k \\ y_{k+1} = y_k + V \sin \theta \times \Delta k \end{cases} \quad (16)$$

2.3 算法描述

根据式(12)的结论,本文提出基于电势能引导进行节点概率选择,即在航路扩展时,按 RRT 算法模式以一定步长在父节点周围生长,当势能递减时,将其作为子节点并加入搜索树中,当搜索趋近威胁时,认为树向前扩展失败,以选择概率接受新节点,并累乘概率。重复上述过程,以确保路径趋向目标点扩展,根据路径比例小量选取最优分支。为保证改进算法向未知区域搜索,同时保留距随机生成点单位步长范围内最近的子节点。

首先,对于任务区域 C 进行定义。设在飞行环境中存在若干威胁源,飞行器为单位正电荷,威胁源为负电荷,同一威胁类型电量相同,根据火力覆盖范围确定电场影响范围。设目标点为电量较大的正电荷,电场覆盖任务区域。设 C_{free} 为非威胁区,设 T_k 为一个具有 K 个节点的扩展树, x 为 T_k 的节点,且 $T_k \in C_{\text{free}}$ 。 x_{init} 为起点, x_{goal} 为终点, x_{rand} 为非威胁区空间中一随机电势点,即 $x_{\text{rand}} \in C_{\text{free}}$ 。

然后,进行采点选取新节点。用 x_{rand} 与 K 个节点进行距离计算对比,设 x_{near} 为树中与 x_{rand} 距离最小的节点。设 $p, q \in C_{\text{free}}$, 令 $\text{Dis}(p, q)$ 为状态点的几何距离,这里采用欧氏距离,则 $\text{Dis}(x_{\text{near}}, x_{\text{rand}}) \leq \text{Dis}(x, x_{\text{rand}})$ 。在 x_{rand} 与 x_{near} 两点直线间取 $x_{\text{new}}, x_{\text{new}} \in C_{\text{free}}$, 并满足 $\text{Dis}(x_{\text{near}}, x_{\text{rand}}) = \rho$, 其中 $\rho > 0$, 为树生长的最小单位长度,称为步长。如满足条件,考虑运动约束,保证路径平滑,令 $\alpha = \pi - \theta$, 以 $\alpha > 3\pi/4$ 条件筛选节点。

最后,进行电势能概率选择。计算 x_{near} 与新增节点 x_{new} 的电势差 $\Delta V(x_{\text{near}}, x_{\text{new}})$, 如果 $\Delta V(x_{\text{near}}, x_{\text{new}}) \leq 0$, 则树新增一个节点 x_{new} , 否则以选择概率接受新节点。令 T_{k+1} 为扩展后的树, $T_{k+1} = T_k + x_{\text{new}}$ 。否则重新选取 x_{rand} , 循环上述步骤。

3 仿真实验分析

本节通过仿真实验对比,从规划算法的效率、航路代价和安全性等方面,对基于拟态电势能的

航路规划方法的性能进行分析。仿真实验环境为:软件 MATLAB 7.0;计算机配置:Windows XP 操作系统、CPU 为 Inter Core i3、主频 3.3 GHz。飞行威胁场大小为 50 km × 50 km,设 θ 为常数 1, $\delta = 0.01, k = \frac{1}{\sqrt{2}}, \epsilon_0 = \frac{1}{4\pi}$ 。

首先,在同等条件下与 RRT 算法、A* 算法对比规划时间、航路长度等指标。仿真条件:设定 13 处威胁点,威胁强度由电量大小确定,电量为 +1 C 和 +2 C,目标点电量为 -10 C;飞行器为元电荷,电量为 +1 C。

如图 2 和图 3 所示, (x, y) 表示位置坐标。表 1 列出了各算法性能对比。平均曲率的计算方法为:将路径做近似平滑逼近,按曲率公式对每个节点进行计算求均值得出。根据式(7),目标导向性表示目标点引力对电荷沿路径从起点移动到终点所做的功。由于在 RRT 算法中节点的选取

带有很大的随机性,产生的节点都是建立盲目搜索的基础上的。这种方式带来的最大好处就在于有极大的未知空间搜索优势,但也会因为大面积无效搜索导致计算成本急剧增加,而且产生的路径并非最优。如图 3 所示。目标导向性表示路径上节点的电势之和,描述路径向目标点的趋向性,其值的绝对值越小,则越能以正确方向趋近目标点,且节点数越少。其数据通过位置坐标对应电势大小按路径节点累加取得。可见,A* 算法(以 Manhattan 距离为启发函数)能以最短路径找到目标点,但在复杂环境下计算时间最长,且路径未考虑运动约束;本文提出的算法节点选择收敛性明显增强,扩展树指向较为单一,路径较为平滑,且计算时间较小。

其次,进行路径安全性评价。根据以上的仿真实验结果,对比 3 种方法的路径安全性。基于电势场拟态的威胁环境建模有效表征了威胁强度与距离之间的关系。考虑到飞行器机动性及环境未知因素,路径紧贴威胁边界飞行也存在极大的危险性。因此,综合考察全域内所有威胁对每一飞行节点的影响十分必要。基于此,这里将环境威胁模型拓展为评价指标,即

$$V_{\text{threat}} = \sum_{i=1}^n V_{oi} = \sum_{i=1}^n \frac{q_{oi}}{4\pi\epsilon_0 r_i} \quad r_i \in (0, \infty) \quad (17)$$

设以原点为中心,不同威胁源拟态电荷沿梯度方向的电势线对比如图 4 所示。距威胁源距离越近,威胁度越大。在威胁源中心威胁度急剧增大,距离中心 5 个单位以上时威胁度渐近为 0。

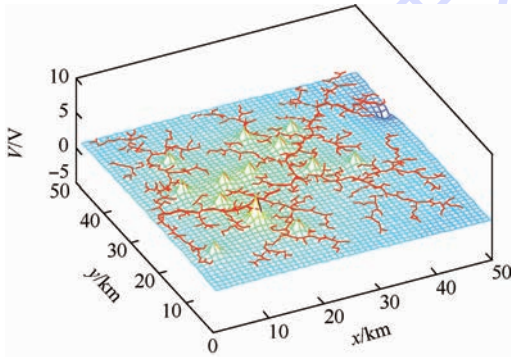


图 2 RRT 算法生成路径

Fig. 2 Path generated by RRT algorithm

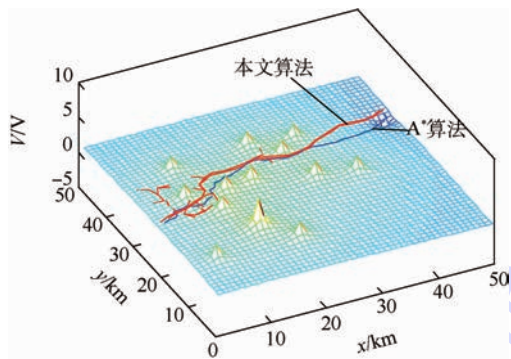


图 3 本文算法与 A* 算法生成路径对比

Fig. 3 Comparison of paths generated by proposed algorithm and A* algorithm

表 1 算法性能对比

Table 1 Algorithm performance comparison

算法	计算时间/s	路径长度/km	平均曲率	目标导向性/J
本文	0.9	78	0.74	-27.0476
A*	9.4	76	0.71	-26.3611
RRT	4.3	92	1.63	-31.8905

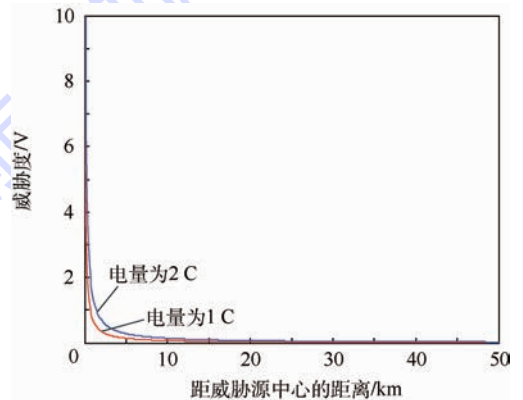


图 4 不同威胁的威胁度对比

Fig. 4 Comparison of threat levels of different threats

因为扩展树节点数和路径长度有很大关系,而不能直接表示威胁度的强弱。为在同一条件下对比 3 种算法生成的路径的威胁度,本文以目标点 x_{goal} 为圆心,分别以等差递增数列为半径作圆,

确定与路径交点的位置坐标。根据威胁度的定义,对每一采样点进行计算,并累加求和。

如图5所示,采样点个数与威胁度大致呈线性关系,采点数对威胁度影响不大。本文算法与RRT算法路径威胁度相似,且均小于A*算法,表明路径距离威胁源较远,路径安全性更佳。

最后,对算法适用性进行对比。通过局部最小测试,与RRT-A*算法进行对比,分析算法的适用性。仿真条件:在任务环境中,设6处威胁源呈反C形排列。起始点坐标(1,22),目标点坐标(30,20)。威胁源和目标点电量分别为1C和-10C。

图6为未考虑运动约束条件 α ,生成的路径与图7相比平滑度更差。对算法进行局部最小测试表明:RRT-A*算法虽然运用A*算法的启发函数改进了RRT算法,航路扩展趋向性更强,但并没有很好地解决A*算法启发后引入的局部最小问题(见图8)。基于拟态电势能启发的算法通过节点概率选择,有效解决了局部最小问题。

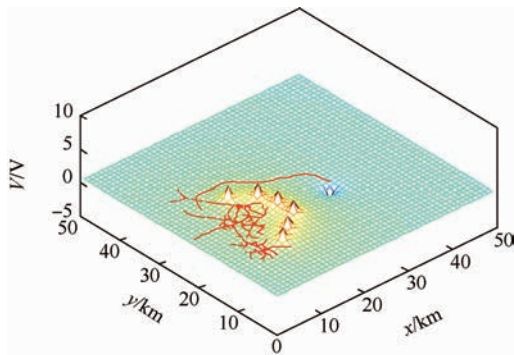


图7 本文算法在考虑约束条件下的局部最小测试
Fig.7 Proposed algorithm's local minimum test with constraint condition

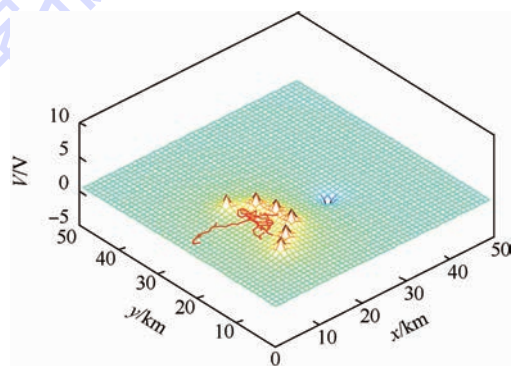


图8 RRT-A*算法的局部最小测试

Fig.8 RRT-A* algorithm's local minimum test

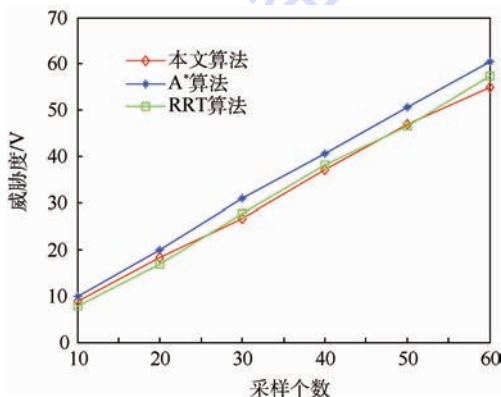


图5 3种算法的威胁度对比

Fig.5 Comparison of three algorithms' threat levels

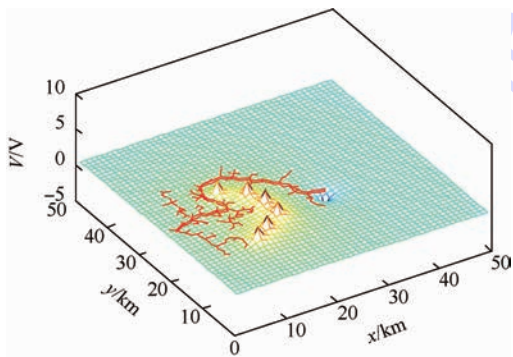


图6 本文算法在未考虑约束条件下的局部最小测试

Fig.6 Proposed algorithm's local minimum test without constraint condition

4 结论

本文依据电势场分布特性进行威胁环境建模,从概率选择的角度建立了基于拟态电势能的航路扩展方法。

1) 该方法改善了复杂环境和高维空间下航路规划算法的计算效率和适用性,突出体现了计算成本低、速度快以及航路安全性高等特点。

2) 本文的研究表明,基于电场力做功的物理原理可为飞行器航路规划问题提供新的机理,但对于动态威胁的实时规划等问题,还需要进一步研究。

参考文献 (References)

[1] 茹常剑,魏瑞轩,沈东. 多无人机协同的稳定控制机理研究[J]. 物理学报, 2014, 63(22): 220202-1-220202-7.
RU C J, WEI R X, SHEN D. Study on stability control mechanism of multiple unmanned aerial vehicle cooperative system [J]. Acta Physica Sinica, 2014, 63(22): 220202-1-220202-7 (in Chinese).

[2] LUGO-CARDENAS I, FLORES G, SALAZAR S, et al. Dubins path generation for a fixed wing UAV[C]// International Conference on Unmanned Aircraft Systems (ICUAS). Piscataway, NJ: IEEE Press, 2014: 339-346.

- [3] CHEN T B, ZHANG Q S. Robot motion planning based on improved artificial potential field [C] // 3rd 2013 International Conference on Computer Science and Network Technology (ICCSNT). Piscataway, NJ: IEEE Press, 2013: 1208-1211.
- [4] SZCZERBA R J, GALKOWSKI P, GLICKTEIN I S, et al. Robust algorithm for real-time route planning [J]. IEEE Transactions on Aerospace and Electronic System, 2000, 36(3): 869-878.
- [5] TUNCER A, YILDIRIM M. Dynamic path planning of mobile robots with improved genetic algorithm [J]. Computers & Electrical Engineering, 2012, 38(6): 1564-1572.
- [6] CHAARI I, KOUBAA A, BENNACEUR H, et al. SmartPATH: A hybrid ACO-GA algorithm for robot path planning [C] // Proceedings 2012 IEEE Congress on Evolutionary Computation (CEC 2012). Piscataway, NJ: IEEE Press, 2012: 1-8.
- [7] SHAMSUDIN H C, ABIDIN A F Z, IRAWAN A, et al. A fast discrete gravitational search algorithm [C] // 4th International Conference on Computational Intelligence, Modelling and Simulation. Piscataway, NJ: IEEE Press, 2012: 24-28.
- [8] SPEARS W M, SPEARS D F, KERR W, et al. An overview of physicomimetics [J]. Lecture Notes in Computer Science, 2005, 3342: 84-97.
- [9] 柴争义, 王秉, 李亚伦. 拟态物理学优化的认知无线网络频谱分配 [J]. 物理学报, 2014, 63(22): 228802-1-228802-6. CHAI Z Y, WANG B, LI Y L. Spectrum allocation of cognitive radio network based on artificial physics optimization [J]. Acta Physica Sinica, 2014, 63(22): 228802-1-228802-6 (in Chinese).
- [10] BIRBIL S I, FANG S C. An electromagnetism-like mechanism for global optimization [J]. Journal of Global Optimization, 2003, 25(3): 263-282.
- [11] RICHARD A, FORMATO J D. Central force optimization: A new nature inspired computational framework for multidimensional search and optimization [J]. Nature Inspired Cooperative Strategies for Optimization, 2008, 129: 221-238.
- [12] DAVE F, ANTHONY S. Anytime RRTs [C] // Proceedings of IEEE/RSJ International Conference on Intelligent, Robots and System. Piscataway, NJ: IEEE Press, 2006: 5798-5803.
- [13] OREN S, DAN H. Asymptotically near-optimal RRT for fast, high-quality, motion planning [C] // IEEE International Conference on Robotics & Automation (ICRA). Piscataway, NJ: IEEE Press, 2014: 4680-4685.
- [14] LI J, LIU S, ZHANG B. RRT-A* motion planning algorithm for non-holonomic mobile robot [C] // SICE Annual Conference 2014. Piscataway, NJ: IEEE Press, 2014: 1833-1838.
- [15] LEE D, SHIM D H. Spline-RRT* based optimal path planning of terrain following flights for fixed-wing UAVs [C] // The 11th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI 2014). Piscataway, NJ: IEEE Press, 2014: 257-261.
- [16] LEE D, SONG H, SHIM D H. Optimal path planning based on spline-RRT* for fixed-wing UAVs operating in three-dimensional environments [C] // International Conference on Control, Automation, and Systems. Piscataway, NJ: IEEE Press, 2014: 22-25.
- [17] VIEIRA H L, GRASSI V. Improving RRT's efficiency through motion primitives generation optimization [C] // 2014 Joint Conference on Robotics: SBR-LARS Robotics Symposium and Robocontrol. Piscataway, NJ: IEEE Press, 2014: 37-42.
- [18] KARAMAN S, FRAZZOLI E. Sampling-based algorithms for optimal motion planning [J]. International Journal of Robotics Research, 2011, 30(7): 846-894.
- [19] RATLIFF N, ZUCKER M, BAGNELL J, et al. CHOMP: Gradient optimization techniques for efficient motion planning [C] // Proceedings of IEEE International Journal Conference on Robotics and Automation (ICRA). Piscataway, NJ: IEEE Press, 2009: 489-494.
- [20] 马文蔚. 物理学 [M]. 5版. 北京: 高等教育出版社, 2006: 149-186. MA W W. Physics [M]. 5th ed. Beijing: Higher Education Press, 2006: 149-186 (in Chinese).
- [21] JUNGHUN S, SONGHWAI O. A cost-aware path planning algorithm for mobile robots [C] // 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, NJ: IEEE Press, 2012: 4724-4729.

作者简介:

何仁珂 男, 硕士研究生。主要研究方向: 导航制导与控制。
Tel: 18700991460

E-mail: lnzrds@163.com

魏瑞轩 男, 博士, 教授。主要研究方向: 导航制导与控制。
Tel.: 029-84787729

E-mail: rxwei369@soho.com

Mimetism electric potential energy motion planning algorithm for aircraft

HE Renke, WEI Ruixuan*, ZHANG Qirui, XU Zhuofan

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

Abstract: Path planning can ensure that the unmanned aerial vehicle (UAV) flies safely and completes mission successfully. In the complex threat environment and high-dimensional space, traditional path planning methods have some limitations in the aspects of the calculation speed, path security and applicability. In order to solve these problems, the electric potential field distribution characteristics and the law of mechanical work driven by electric field force were studied. The mimetism electric potential energy path planning method was proposed, and the environment model based on the electric potential field distribution and path node probability choice mechanism were established. The relationship between threat intensity and distance was described by using electric potential. Combined with electric potential, the path safety evaluation standard was proposed. On this basis, the potential field based sampling-based random path planning method was proposed. The results show that, compared with traditional methods, the method mentioned above can generate optimal path in consideration of non-holonomic differential constraints, significantly shorten the path length and computational time, and improve the path security, which is of great value for application of path planning.

Key words: unmanned aerial vehicle (UAV); path planning; mimetism electric potential energy; probability choice; path security

Received: 2015-06-29; **Accepted:** 2015-09-18; **Published online:** 2015-10-30 16:54

URL: www.cnki.net/kcms/detail/11.2625.V.20151030.1654.008.html

Foundation item: Aeronautical Science Foundation of China(20135896027)

* **Corresponding author.** Tel. : 029-84787729 E-mail: rxwei369@soho.com

http://bhxb.buaa.edu.cn jbuua@buaa.edu.cn

DOI: 10.13700/j.bh.1001-5965.2015.0442

重型直升机-吊挂耦合系统闭环飞行品质分析

朱笑宇, 曹义华*, 曹龙

(北京航空航天大学 航空科学与工程学院, 北京 100083)



摘 要: 以 CH-53A/D 直升机为对象, 研究重型直升机在增稳控制系统工作条件下外挂载荷运输飞行任务中的飞行品质问题。基于单质点吊挂假设和控制系统模型, 建立了直升机-吊挂耦合闭环系统的非线性动力学模型。在该假设下, 吊挂物引入了额外的自由度和约束, 使方程增加为 13 阶的微分方程组。针对闭环状态下的耦合系统, 进行了小扰动线性化处理, 在此基础上进行了吊挂飞行操纵响应的时域特性和频域特性分析, 参照军用旋翼飞行器驾驶品质要求 (ADS-33E) 的品质指标分析了吊挂物质量、吊挂绳长和直升机前飞速度对耦合系统飞行品质的影响。结果表明, 吊挂物质量、吊挂绳长和前飞速度对于系统的时域特性以及各项操纵品质有不同程度的影响。

关键词: 直升机; 吊挂; 闭环; 飞行品质; 控制系统

中图分类号: V275; V212.4

文献标识码: A **文章编号:** 1001-5965(2016)07-1550-07

直升机带外吊挂运输飞行是直升机相比于其他种类飞行器的一项特有功能。吊挂物不受货物外形限制, 直升机还可以在一般运输工具难以到达的地方迅速、高效地完成起降、装卸货物等任务, 大大拓宽了运输范围。与直升机非吊挂飞行状态相比, 直升机吊挂飞行增加了新的载荷和惯性力, 以及因此产生的对直升机质心的力矩。建立合理的非线性飞行动力学模型是进行直升机吊挂飞行研究分析的基础。开展直升机吊挂飞行操纵性的研究对提高吊挂运输效率和飞行安全都有着十分重要的意义。

国外自 1965 年起, 就有关于直升机吊挂飞行的研究。Sampath^[1] 对此有较为系统的综述。Lucassen 和 Sterk^[2] 共同建立了一个简单的三自由度直升机纵向悬停方程, 忽略了吊挂受到的气动力和气动力矩。1986 年, Ronen^[3] 建立了一个新的单点悬挂携带载荷的 CH-53A/D 直升机模型, 用以改善当前的动力学模型, 并开始研究系统的

开环特性。Stucky^[4] 和 Cicolani 等^[5] 分别用数值仿真和飞行试验的方法对带吊挂的直升机本体运动模式进行了研究。稍后, Cliff 和 Bailey^[6] 使用了一个只考虑阻力的单吊挂直升机简化模型来研究吊挂飞行。在他们提供的公式中, 非线性方程在稳定的水平飞行状态下被线性化, 扰动方程就被分成纵向和横向。

国内吊挂飞行的研究起步较晚, 起始于 20 世纪 80 年代, 而且基本停留在理论分析上。崔瑛^[7] 在文献^[8] 的基础上, 研究了直升机气动导数在吊挂引入后的变化。齐万涛和陈仁良^[9] 又继续开展了弹性绳索、吊挂构型和操稳特性等方面的研究和分析。对于带外吊挂直升机品质问题, 由于吊挂物与直升机之间的运动耦合, 使整个系统的运动更为复杂, 崔利等^[10] 建立了带吊挂载荷的直升机飞行动力学模型, 分析了当施加纵向周期变距阶跃输入、横向周期变距阶跃输入以及尾桨距阶跃输入时带吊挂直升机

收稿日期: 2015-07-01; 录用日期: 2015-10-30; 网络出版时间: 2016-01-18 16:42

网络出版地址: www.cnki.net/kcms/detail/11.2625.V.20160118.1642.003.html

* 通讯作者: Tel.: 010-82339537 E-mail: yihuacaobu@126.com

引用格式: 朱笑宇, 曹义华, 曹龙. 重型直升机-吊挂耦合系统闭环飞行品质分析[J]. 北京航空航天大学学报, 2016, 42(7): 1550-1556. ZHU X Y, CAO Y H, CAO L. Heavy helicopter-slung-load coupling system flying qualities in closed-loop state[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(7): 1550-1556 (in Chinese).

的动力学响应历程。

本文在直升机本体模型的基础上,建立了六自由度刚体吊挂物气动模型,得到了完整的直升机-吊挂耦合系统动力学方程,并基于文献[11-12]给出了相应的控制系统模型,建立了完整的闭环直升机-吊挂耦合系统动力学模型。在此基础上,根据军用旋翼飞行器驾驶品质要求(ADS-33E)^[13]对带吊挂物直升机闭环状态下的飞行品质进行计算,分析了吊挂飞行操纵响应的时域特性和频域特性,以及吊挂物质量、吊挂绳长和直升机前飞速度对耦合系统飞行品质的影响。

1 直升机-吊挂耦合系统闭环模型

1.1 直升机-吊挂耦合系统模型及小扰动线性化

1.1.1 线性全量方程

本文所采用的直升机吊挂动力学模型来自于文献[14]。假设吊挂物为质点模型,没有转动自由度,忽略绳索气动力和吊挂物上作用的力矩,其所受的气动力仅为当地来流方向的准静态阻力。忽略绳索质量,绳索为刚性模型,且时刻处于绷紧状态。故在此假设下绳索为连结着直升机挂点和吊挂质心的一条直线,提供一个双侧的完整约束。

直升机本体的状态变量包括前飞速度 $\mathbf{v}_{cg} = [u \ v \ w]^T$ (u, v 和 w 分别为 x, y 和 z 方向速度)、姿态角 $\boldsymbol{\theta} = [\phi \ \theta \ \psi]^T$ (ϕ, θ 和 ψ 分别为滚转角、俯仰角和偏航角)和姿态角速度 $\boldsymbol{\omega} = [p \ q \ r]^T$ (p, q 和 r 分别为滚转角速度、俯仰角速度和偏航角速度),控制变量为 $\mathbf{u} = [\theta_{om} \ B'_1 \ A'_1 \ \theta_{ot}]^T$ (θ_{om} 为飞行员的总距, B'_1 为纵向周期变距, A'_1 为横向周期变距, θ_{ot} 为脚蹬输入)。在直升机的体轴系内衡量吊挂位移的广义坐标即吊挂姿态角,如图1所示。图中, θ_1 和 ϕ_1 分别为吊挂物的后摆角和侧摆角。记吊挂姿态角 $\boldsymbol{\theta}_1 = [\theta_1 \ \phi_1]^T$, 吊挂姿态角速度 $\boldsymbol{\omega}_1 = [\dot{\theta}_1 \ \dot{\phi}_1]^T$, $\dot{\theta}_1$ 和 $\dot{\phi}_1$ 分别为后摆角速度和侧摆角速度。图1和坐标定义参考文献[15]。

直升机-吊挂耦合系统动力学方程组^[15]的矩阵形式为

$$\begin{bmatrix} (m_H + m_L)I_3 & -m_L \tilde{\mathbf{r}}_L & m_L \mathbf{R}_2 & \mathbf{0} \\ m_L \tilde{\mathbf{r}}_h & \mathbf{I} - m_L \tilde{\mathbf{r}}_h \tilde{\mathbf{r}}_L & m_L \tilde{\mathbf{r}}_h \mathbf{R}_2 & \mathbf{0} \\ m_L I_3 & -m_L \tilde{\mathbf{r}}_L & m_L \mathbf{R}_2 & \mathbf{r}_L \end{bmatrix} \begin{bmatrix} \dot{\mathbf{v}} \\ \dot{\boldsymbol{\omega}} \\ \dot{\boldsymbol{\omega}}_1 \\ \lambda_L \end{bmatrix} = \begin{bmatrix} (m_H + m_L)\mathbf{g} + \mathbf{F} - m_L \mathbf{a}_{12} + \mathbf{D}_L - m_H \tilde{\boldsymbol{\omega}} \mathbf{v} \\ \mathbf{M} - m_L \tilde{\mathbf{r}}_h \mathbf{a}_{12} + \tilde{\mathbf{r}}_h \mathbf{D}_L + m_L \tilde{\mathbf{r}}_h \mathbf{g} - \tilde{\boldsymbol{\omega}} (\mathbf{I} \boldsymbol{\omega}) \\ -m_L \mathbf{a}_{12} + \mathbf{D}_L + m_L \mathbf{g} \end{bmatrix} \quad (1)$$

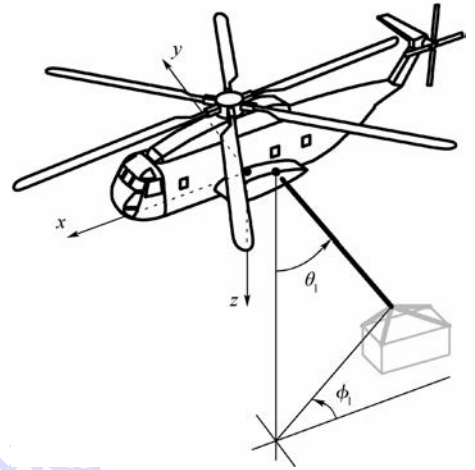


图1 直升机体轴系与吊挂姿态角定义
Fig.1 Definition of helicopter body reference frame and attitude angles of slung-load

方程式(1)中,只有等号左边的向量 $[\dot{\mathbf{v}}_{cg} \ \dot{\boldsymbol{\omega}} \ \dot{\boldsymbol{\omega}}_1 \ \lambda_L]^T$ 中含有加速度项, $\dot{\mathbf{v}}_{cg}$ 为直升机前飞角速度, $\dot{\boldsymbol{\omega}}$ 为直升机姿态角加速度, $\dot{\boldsymbol{\omega}}_1$ 为直升机吊挂姿态角加速度, λ_L 为绳索张力与绳长的比值。上述动力学方程与直升机本体的 Euler 运动学方程联立,即得到直升机-吊挂耦合系统全量动力学方程组。

1.1.2 小扰动线性化

直升机-吊挂耦合系统全量动力学方程的微分代数形式为

$$\mathbf{M}(\mathbf{x}_L) \cdot \begin{bmatrix} \dot{\mathbf{x}}_L \\ \lambda_L \end{bmatrix} = \mathbf{f}(\mathbf{x}_L, \mathbf{u}) \quad (2)$$

式中: \mathbf{x}_L 为耦合系统的状态变量, $\mathbf{x}_L = [\mathbf{v}_{cg} \ \boldsymbol{\omega} \ \boldsymbol{\theta} \ \boldsymbol{\theta}_1 \ \boldsymbol{\omega}_1]^T$ 为13阶向量; $\mathbf{M}(\mathbf{x}_L)$ 为惯性矩阵,详细参见文献[16]。

此时,耦合系统的非线性状态空间方程为

$$\dot{\mathbf{x}}_L = \hat{\mathbf{f}}(\mathbf{x}_L, \mathbf{u}) \quad (3)$$

式中: $\hat{\mathbf{f}}(\mathbf{x}_L, \mathbf{u})$ 为 $[\mathbf{M}(\mathbf{x}_L)]^{-1} \cdot \mathbf{f}(\mathbf{x}_L, \mathbf{u})$ 的前13个分量。

所以,线性化系统的状态矩阵和控制矩阵是 $\frac{\partial \hat{\mathbf{f}}}{\partial \mathbf{x}_L}$ 和 $\frac{\partial \hat{\mathbf{f}}}{\partial \mathbf{u}}$ 。

1.2 带控制系统的直升机-吊挂耦合系统模型

本文所采用的控制系统来自于 NASA CR-3144、CR-3145 报告^[11-12]。 $\boldsymbol{\delta} = [\delta_C \ \delta_B \ \delta_A \ \delta_P]^T$ 分别为飞行员的总距、纵/横向周期变距、脚蹬输入的操纵行程百分比,即操纵输入。 θ_{MR} 、 B_{1s} 、 A_{1s} 和 θ_{TR} 分别为执行机构上的主旋翼总距角、纵/横向变距角(自动倾斜器倾角)和尾旋翼总距角,它们就是旋翼气动模型中的控制变量 θ_{om} 、 B'_1 、 A'_1 和

θ_{ot} 。控制系统按不同操纵对象分为4个通道:总距通道、纵向通道、横向通道、脚蹬通道。总距操纵 δ_c 会交叉进给至横向和脚蹬通道,以在一定程度上抵消由于总距操纵而引起的滚转和偏航力矩。纵向通道与其他3个通道没有关联。直接接收飞行员操纵输入的有7个环节,分别是各操纵通道的机构联动、纵向通道前馈以及总距输入的交叉进给^[16]。

由文献[11-12]给出的控制系统模型,可以写出控制系统执行机构的主旋翼总距角、纵/横向变距角(自动倾斜器倾角)、尾旋翼总距角与系统各变量之间动态关系的频域表达式为

$$\begin{cases} \theta_{om} = 0.129\delta_c \\ B'_1 = 0.272\delta_b + \frac{0.13}{1.4s+1} \cdot \delta_b + 0.32\dot{\theta} + 0.599\theta \\ A'_1 = -0.012\delta_c + 0.124\delta_a - 0.15p - 0.24\phi \\ \theta_{ot} = 0.13\delta_c + 0.266\delta_p + \frac{0.573s}{1.8s+1} \cdot r - \\ \quad 0.0815p + 0.282\dot{v} + 0.215\psi \end{cases} \quad (4)$$

令

$$\begin{cases} c = \frac{0.13}{1.4s+1} \cdot \delta_b \\ l = \frac{0.573}{1.8s+1} \cdot r \end{cases} \quad (5)$$

进一步转化成时域形式可得

$$\begin{cases} \dot{c} = -\frac{1}{1.4}c + \frac{0.13}{1.4}\delta_b \\ \dot{l} = -\frac{1}{1.8}l + \frac{0.573}{1.8}r \end{cases} \quad (6)$$

从方程式(6)可以看到,控制系统的引入使整个系统增加了两阶动态,对应的状态变量可以取为 c 和 l ,记 $\xi = [c \ l]^T$ 。将其代入直升机本体小扰动线性化模型的表达式,变形后可得到加了控制系统后的小扰动线性化模型方程为

$$\begin{bmatrix} \dot{x}_L \\ \dot{\xi} \end{bmatrix} = \begin{bmatrix} A_L + B_L M_1 & B_L M_2 \\ C_\xi & A_\xi \end{bmatrix} \begin{bmatrix} x_L \\ \xi \end{bmatrix} + \begin{bmatrix} B_L M_3 \\ B_\xi \end{bmatrix} \cdot \delta \quad (7)$$

式中: A_L 和 B_L 为直升机-吊挂耦合系统线性化方程的状态矩阵和控制矩阵,即 $\frac{\partial \hat{f}}{\partial x_L}$ 和 $\frac{\partial \hat{f}}{\partial u}$ 。 A_ξ 、 B_ξ 和 C_ξ 分别为 $\dot{\xi}$ 与 ξ 、 δ 、 x_L 相对应的系数矩阵:

$$A_\xi = \begin{bmatrix} -\frac{1}{1.4} & 0 \\ 0 & -\frac{1}{1.8} \end{bmatrix} \quad (8)$$

$$B_\xi = \begin{bmatrix} 0 & \frac{0.13}{1.4} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (9)$$

$$C_\xi = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{0.573}{1.8} & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (10)$$

$$\begin{cases} M_1 = (I - M_{uu})^{-1} \cdot M_{\delta u} \\ M_2 = (I - M_{uu})^{-1} \cdot M_{\xi u} \\ M_3 = (I - M_{uu})^{-1} \cdot M_{x_L u} \end{cases} \quad (11)$$

而 $M_{\delta u}$ 、 $M_{\xi u}$ 、 $M_{x_L u}$ 和 M_{uu} 分别为控制变量 $[\theta_{om} \ B'_1 \ A'_1 \ \theta_{ot}]^T$ 与 δ 、 ξ 、 x_L 和 u 相对应的系数矩阵。具体的矩阵形式为

$$M_{\delta u} = \begin{bmatrix} 0.129 & 0 & 0 & 0 \\ 0 & 0.272 & 0 & 0 \\ -0.012 & 0 & 0.124 & 0 \\ 0.138 & 0 & 0 & 0.266 \end{bmatrix} \quad (12)$$

$$M_{\xi u} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & -\frac{1}{1.8} \end{bmatrix} \quad (13)$$

$$M_{x_L u} = \begin{bmatrix} 0 & 0.32a_{81} & 0 & 0 & 0.282a_{21} \\ 0 & 0.32a_{82} & 0 & 0 & 0.282a_{22} \\ 0 & 0.32a_{83} & 0 & 0 & 0.282a_{23} \\ 0 & 0.32a_{84} & -0.15 & 0.282a_{24} & -0.0815 \\ 0 & 0.32a_{85} & 0 & 0 & 0.282a_{25} \\ 0 & 0.32a_{86} & 0 & 0 & 0.282a_{26} + \frac{0.573}{1.8} \\ 0 & 0.32a_{87} - 0.24 & -0.24 & 0 & 0.282a_{27} \\ 0 & 0.32a_{88} + 0.599 & 0 & 0 & 0.282a_{28} \\ 0 & 0.32a_{89} & 0 & 0 & 0.282a_{29} \\ 0 & 0.32a_{810} & 0 & 0 & 0.282a_{210} \\ 0 & 0.32a_{811} & 0 & 0 & 0.282a_{211} \\ 0 & 0.32a_{812} & 0 & 0 & 0.282a_{212} \\ 0 & 0.32a_{813} & 0 & 0 & 0.282a_{213} \end{bmatrix} \quad (14)$$

$$M_{uu} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.32b_{81} & 0.32b_{82} & 0.32b_{83} & 0.32b_{84} \\ 0 & 0 & 0 & 0 \\ 0.282b_{21} & 0.282b_{22} & 0.282b_{23} & 0.282b_{24} \end{bmatrix} \quad (15)$$

式中: $a_{ij}(i=1,2,\dots,13;j=1,2,\dots,13)$ 为直升机-吊挂耦合系统小扰动线性化模型中状态矩阵的各项; $b_{ij}(i=1,2,\dots,13;j=1,2,3,4)$ 为直升机-吊挂耦合系统小扰动线性化模型中控制矩阵的各项。

则引入控制系统后的新直升机-吊挂耦合线

化模型的状态矩阵 A_{L-aug} 和控制矩阵 B_{L-aug} 分别为

$$A_{L-aug} = \begin{bmatrix} A_L + B_L M_1 & B_L M_2 \\ C_\xi & A_\xi \end{bmatrix} \quad (16)$$

$$B_{L-aug} = \begin{bmatrix} B_L M_3 \\ B_\xi \end{bmatrix} \quad (17)$$

2 基于 ADS-33E 的飞行品质

2.1 带宽分析

本部分选取了直升机在吊挂物质量为 4 t、吊挂绳长分别为 40 m 和 60 m 时的算例。根据带宽定义分别画出了俯仰通道和滚转通道的幅相曲线,算例伯德图如图 2 和图 3 所示。根据规范要求进一步计算了带宽与相位滞后,数据如表 1 所示。

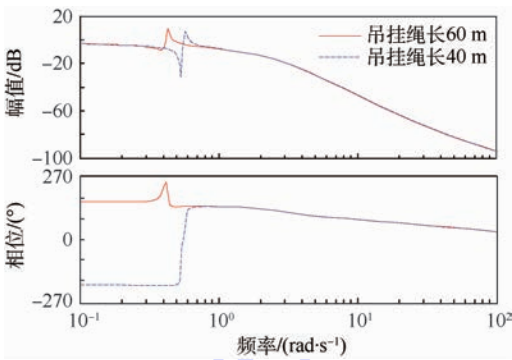


图 2 俯仰通道伯德图

Fig. 2 Bode diagram of pitch channel

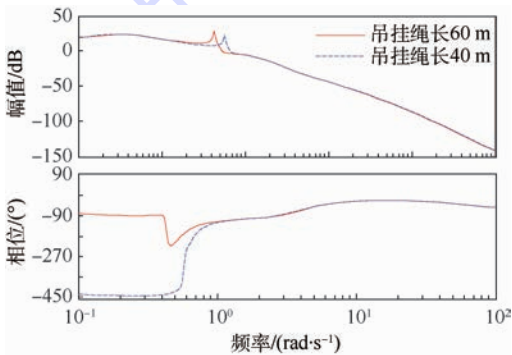


图 3 滚转通道伯德图

Fig. 3 Bode diagram of roll channel

表 1 俯仰和滚转通道带宽和相位延迟对比

Table 1 Comparison of bandwidth and phase delay in pitch and roll channel

吊挂绳长/m	吊挂物质量/t	俯仰通道		滚转通道	
		$\omega_{BW,\theta}/(\text{rad} \cdot \text{s}^{-1})$	$\tau_{p,\theta}/\text{s}$	$\omega_{BW,\phi}/(\text{rad} \cdot \text{s}^{-1})$	$\tau_{p,\phi}/\text{s}$
60	4	0.402	3.8296	0.420	1.0749
40	4	0.574	5.4538	0.611	0.8943
	无吊挂	0.442	0.1844	0.663	0.5389

注: $\omega_{BW,\theta}$ —俯仰通道带宽; $\tau_{p,\theta}$ —俯仰通道相位延迟;

$\omega_{BW,\phi}$ —滚转通道带宽; $\tau_{p,\phi}$ —滚转通道相位延迟。

闭环系统的带宽在频域分析中是很重要的参数,它度量了系统具有一定信号复现能力的频率变化范围。当输入信号频率低于带宽时,系统可以产生足够强度的响应。一般情况下,系统的阶跃响应速度与带宽成正比,调节时间与带宽成反比。因此,在保证系统的合理时域指标的前提下,总是希望系统有较大的带宽。在本算例中,带吊挂闭环状态下,由伯德图可以看出:

1) 幅频特性的峰值 M_r 随绳长增加而增大,这意味着系统的平稳性随绳长增加而变差,阶跃响应的超调量也变大。而幅频响应峰值频率 ω_r 随绳长增加而变小,这是由于绳长增加后,吊挂摆动的固有频率减小而造成的。2 个算例中 CH-53A/D 的带宽均较小,意味着驾驶员可用的、距不稳定边界尚有适当余量的最大操纵频率较小。并且随着吊挂绳长的增加,俯仰通道和滚转通道的带宽均是减小的。这说明,吊挂的绳长越长,系统对阶跃输入的响应越慢,所需的调节时间越长。

2) 对于相位滞后指标,表示驾驶员操纵频率接近相角穿越频率 ω_g 时,直升机走向不稳定的快慢程度。2 个算例的俯仰通道和滚转通道的指标数值均不小,说明相频曲线下落较快,直升机走向不稳定的趋势也较快,但与绳长的关系表现的并不明显。但滚转通道的相位滞后指标明显优于俯仰通道,说明直升机在滚转通道的飞行品质优于俯仰通道,说明吊挂对俯仰通道的影响更严重。

3) 等级方面,对比 ADS-33 规范要求,2 个算例中,直升机在实现目标捕捉及跟踪任务时,俯仰通道和滚转通道的指标均在等级 3 范围。其他任务,当可用感示环境 $UCE > 1$ 时,在等级 3 范围;在 $UCE = 1$ 时,均在等级 2 范围。评价结果表明,CH-53A/D 直升机俯仰(滚转)通道带宽偏小,这使得驾驶员应特别注意直升机的操纵频率,否则操纵频率过大极有可能引发驾驶员诱发振荡。

2.2 轴间耦合特性分析

2.2.1 总距-偏航特性

ADS-33 品质规范规定^[13],为实现某一轴上的响应,飞行员施加的操纵输入不得在另一个或另几个轴上引起不适宜的响应。而对于总距-偏航耦合来说,即在俯仰和滚转姿态基本保持不变的情况下,偏航角速度的响应不宜过大。该品质要求已显示在图 4 中,其中, $|r_1/w(3)|$ 和 r_3/w

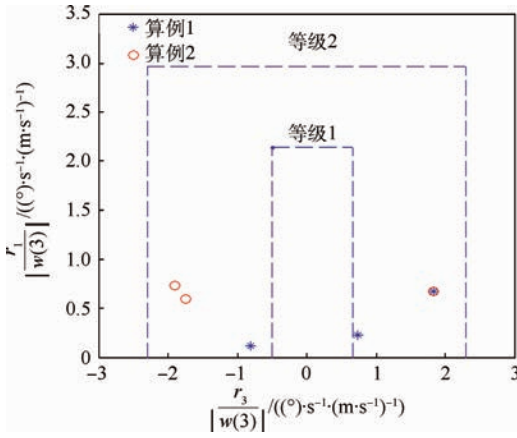


图4 总距-偏航耦合品质指标算例

Fig. 4 Example of yaw due to collective coupling

$|w(3)|$ 为总距-偏航耦合品质的评价指标。

本节选取的2组算例为吊挂质量8t,固定绳长下前飞速度变化和固定前飞速度下吊挂绳长变化的情况。算例1为绳长40m,前飞速度分别为20、30、50m/s;算例2为前飞速度为30m/s,绳长分别为20、40、80m。

由图4可见,所有算例的总距-偏航特性指标均在等级2范围内。说明总距的操纵输入对偏航的角速度有一定影响。

2.2.2 滚转-俯仰耦合特性

迅猛机动中的滚转-俯仰耦合是在时域响应中进行研究。在一个突然的横向操纵机构的阶跃输入后,4s内轴外响应的峰值与4s末目标响应之比,即 θ_ϕ/ϕ_4 ,可以作为品质指标来确定该项耦合品质。ADS-33品质规定等级1为 θ_ϕ/ϕ_4 不超过 ± 0.25 ,等级2为 θ_ϕ/ϕ_4 不超过 ± 0.6 。航向基本保持不变。

本部分选取的算例为前飞速度50m/s、吊挂绳长分别为40m和60m、吊挂质量由1t变化至8t的情况。结果如图5所示,数据点全部在等级1和等级2的范围内,并且由数据可以看出相同的前飞速度和吊挂质量下,滚转-俯仰耦合品质

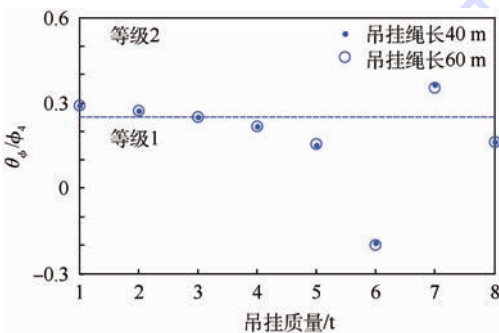


图5 滚转-俯仰耦合品质指标算例

Fig. 5 Example of pitch due to roll coupling

受吊挂绳长的影响不大。综合而言,带吊挂闭环状态下的CH-53A/D的滚转通道对俯仰通道影响较小,能够达到比较满意的滚转-俯仰耦合飞行品质要求。

2.3 垂直轴操纵功效

垂直轴操纵功效是直升机十分重要的飞行品质评价内容之一。品质规范^[13]规定,自总距快速偏离直升机平衡位置起1.5s后,直升机产生的垂向速度 $w_{1.5}$ 将作为垂直轴操纵功效的评价指标。具体的品质要求已显示在图中。在计算本品质指标时,对总距施加1%的阶跃输入,求解由总距操纵引起的垂向速度响应曲线。

2.3.1 垂直轴操纵功效随吊挂质量变化

本节取了2组算例,第1组前飞速度为50m/s,吊挂绳长为60m,吊挂质量由2t变化至6t;第2组前飞速度为50m/s,吊挂绳长为40m,吊挂质量也由2t变化至6t。所得到的 $w_{1.5}$ 均标注在图6中。由图可见,随着吊挂质量增加,2组算例的垂直轴操纵功效品质等级均变差,吊挂质量为2t和3t的算例在等级1范围;吊挂质量为4、5和6t的算例在等级2范围。且同样的前飞速度和吊挂质量下, $w_{1.5}$ 随绳长的变化关系并不明显。

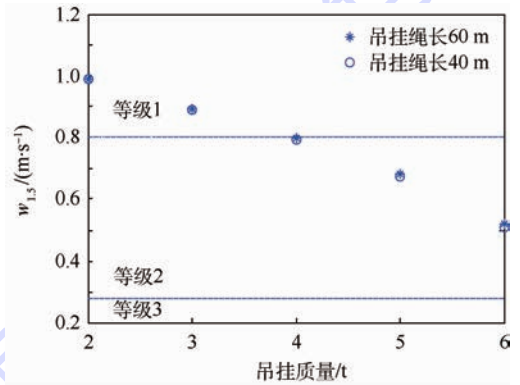


图6 垂直轴操纵功效随吊挂质量变化算例

Fig. 6 Example of vertical control power changing with slung-load mass

2.3.2 垂直轴操纵功效随前飞速度变化

本节取的两组算例为吊挂质量8t,吊挂绳长分别为40m和60m,前飞速度单调递增的情况。所得到的 $w_{1.5}$ 均标注在图7中。由图可见,随着前飞速度的增加,2组算例的垂直轴操纵功效品质呈现先变差后变好的趋势,在速度提高到60m/s后,垂直轴操纵功效品质均在等级1范围内。且同样的前飞速度和吊挂质量下, $w_{1.5}$ 受绳长的影响并不明显。

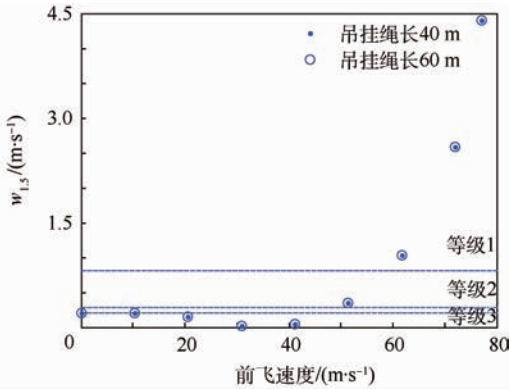


图7 垂直轴操纵功效随前飞速度变化算例

Fig. 7 Example of vertical control power changing with velocity

3 结 论

本文以 CH-53A/D 型直升机为算例,在所建立的飞行动力学模型的基础上,基于直升机单质点吊挂假设和控制系统模型,建立了直升机-吊挂耦合闭环系统的非线性动力学模型,并对系统方程进行了小扰动线性化处理。

在此基础上对直升机-吊挂耦合闭环系统的线性化模型进行了时域特性和频域特性分析,进一步参照 ADS-33E 品质规范分析了吊挂物质量、吊挂绳长和直升机前飞速度对耦合系统飞行品质的影响。结果表明:

1) 耦合系统的带宽随绳长增加而减小,平稳性随绳长增加而变差;阶跃响应的超调量也变大;幅频响应峰值频率 ω_p 随绳长增加而变小;相位滞后指标随绳长变化不明显。综合来看绳长越短,耦合系统的时域特性越好,同时吊挂的引入对俯仰通道的相位滞后指标影响更严重。

2) 对于轴间耦合特性来说,总距的操纵输入对偏航的角速度有一定影响;而相同的前飞速度和吊挂质量下,绳长对滚转-俯仰耦合品质的影响不大。

3) 对于垂直轴操纵功效,随着吊挂质量增加,系统的垂直轴操纵功效品质等级变差;随着前飞速度增加,该项指标呈现先变差后变好的趋势;但同样的前飞速度和吊挂质量下,该项指标受绳长的影响并不明显。

参考文献 (References)

[1] SAMPATH P. Dynamics of a helicopter-slung load system [D]. Maryland: University of Maryland, 1980: 16-21.
[2] LUCASSEN L R, STERK F J. Dynamic stability analysis of a hovering helicopter with a sling load [J]. Journal of the American Helicopter Society, 1965, 10(2): 6-12.
[3] RONEN T. Dynamics of a helicopter with a slung load [D].

Stanford: Stanford University, 1986: 11-26.
[4] STUCKY R A. Mathematical modeling of helicopter slung-load system; DSTO-TR-1257 [R]. [S. l. : s. n.], 2002: 6-37.
[5] CICOLANI L S, CONE A, THERON J N. Flight test and simulation of a cargo container slung load in forward flight [J]. Journal of the American Helicopter Society, 2009, 54 (3) : 032006-1-032006-18.
[6] CLIFF E M, BAILEY D B. Dynamic stability of a translating vehicle with a simple sling load [J]. Journal of Aircraft, 1975, 12 (10) : 773-777.
[7] 崔瑛. 直升机吊挂飞行稳定性分析 [D]. 南京: 南京航空航天大学, 2005: 18-50.
CUI Y. Stability analysis of a helicopter with an external slung load [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2005: 18-50 (in Chinese).
[8] FUSATO D, GUGLIERI G, CELI R. Flight dynamics of an articulated rotor helicopter with an external slung load [J]. Journal of the American Helicopter Society, 2001, 46(1): 3-13.
[9] 齐万涛, 陈仁良. 直升机吊挂飞行稳定性和操纵性分析 [J]. 南京航空航天大学学报, 2011, 43(3): 406-412.
QI W T, CHEN R L. Stability and control characteristic analysis for flight of helicopter with slung-load [J]. Journal of Nanjing University of Aeronautics and Astronautics, 2011, 43(3): 406-412 (in Chinese).
[10] 崔利, 曹义华, 李国知. 直升机吊挂飞行平衡、稳定性与操纵性研究 [J]. 航空动力学报, 2010, 25(10): 2307-2311.
CUI L, CAO Y H, LI G Z. Studies on trims, stability, and control ability of helicopter with slung-load [J]. Journal of Aerospace Power, 2010, 25(10): 2307-2311 (in Chinese).
[11] HEFFLEY R K, JEWELL W F, LEHMAN J M, et al. A compilation and analysis of helicopter handling qualities data, volume one: Data compilation; NASA-CR-3144 [R]. Washington, D. C. : NASA, 1980: 303-357.
[12] HEFFLEY R K. A compilation and analysis of helicopter handling qualities data, Volume two: Data analysis; NASA-CR-3145 [R]. Washington, D. C. : NASA, 1979: 120-142.
[13] Aeronautical design standard performance specification, handling qualities requirements for military rotorcraft: ADS-33E-PRF [S]. [S. l.] : United States Army Aviation and Missile Command, 2000.
[14] STURGEON W R. A mathematical model of the CH-53 helicopter; NASA TM-81238 [R]. Washington, D. C. : NASA, 1980: 1-50.
[15] 曹龙, 曹义华, 李春华. 直升机-吊挂耦合系统平衡特性和稳定性分析 [J]. 北京航空航天大学学报, 2014, 40(9): 1219-1224.
CAO L, CAO Y H, LI C H. Equilibrium characteristics and stability analysis of helicopter-slung-load coupling system [J]. Journal of Beijing University of Aeronautics and Astronautics, 2014, 40(9): 1219-1224 (in Chinese).
[16] 曹龙. 直升机外吊挂动力学建模与飞行性能分析 [D]. 北京: 北京航空航天大学, 2014: 81-98.
CAO L. Dynamics modeling and flying performance analysis of helicopter with slung-load [D]. Beijing: Beihang University, 2014: 81-98 (in Chinese).

作者简介:

朱笑宇 女, 硕士研究生。主要研究方向: 直升机飞行力学与控制。

Tel: 15210967669

E-mail: zhuxiaoyu@buaa.edu.cn

曹义华 男, 博士, 教授, 博士生导师。主要研究方向: 直升机总体设计。

Tel: 010-82339537

E-mail: yihuacaobu@126.com

Heavy helicopter-slung-load coupling system flying qualities in closed-loop state

ZHU Xiaoyu, CAO Yihua*, CAO Long

(School of Aeronautic Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: Flying qualities of a heavy helicopter augmented with an automatic flight control system (AFCS) and flying with externally slung loads are investigated using the mathematical model of the CH-53A/D helicopter. A nonlinear dynamical model of helicopter-slung-load in closed-loop state was built based on single mass-point hypothesis and a control system model. Under this hypothesis, extra degrees of freedom and constraints were brought in by the consideration of the slung-load, which made the set of equations of motion increased to a 13-order one. The coupling system in closed-loop state was linearized under small-perturbed conditions. After this, the slung-load flight control response characteristics in both time and frequency domain were analyzed. The effect of slung-load mass, slung-load rope length and helicopter velocity on coupling system's flight quality was analyzed according to the handling qualities requirements for military rotorcraft indicated by ADS-33E. The final results show that the slung-load mass, slung-load rope length and helicopter velocity have varying degrees of impact on the time domain features and handling qualities of the coupling system.

Key words: helicopter; slung-load; closed-loop; flying qualities; control system

《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人,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: jbuua@buaa.edu.cn

http://bhxb.buaa.edu.cn

http://www.buaa.edu.cn

《北京航空航天大学学报》
第五届编辑委员会

北航学报
赠阅

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7期 2016年7月

JOURNAL OF BEIJING UNIVERSITY OF
AERONAUTICS AND ASTRONAUTICS
(JBUA)

(Monthly, Started in 1956)

Vol.42 No.7 July 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 jbuua@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 Qiping

Edited and Published by Editorial Board of JBUA

Printed by Beijing Kexin Printing Co., Ltd.

Distributed by BUAA Culture Media Group Limited

Telephone (010) 82315594 82338922
82314839

E-mail jbuua@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

