



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



2019-02

Vol.45 No.2

目 次

- 具有初始热变形的转子系统振动响应分析…………… 马艳红, 刘海舟, 邓旺群, 杨海, 洪杰 (227)
- 小视场星敏感器量测延时滤波算法…………… 钱华明, 王迪, 吴永慧, 黄智开 (234)
- 航天刚-弹-液耦合系统的弹-液耦合研究…………… 梁立孚, 郭庆勇 (243)
- 高速运载器燃油热管理系统优化…………… 庞丽萍, 邹凌宇, 阿嵘, 杨晓东, 范俊 (252)
- 基于混沌粒子群优化的北斗/GPS 组合导航选星算法……………
…………… 王尔申, 贾超颖, 曲萍萍, 黄煜峰, 庞涛, 别玉霞, 姜毅 (259)
- 航空发动机转子结构布局优化设计方法…………… 李超, 金福艺, 张卫浩 (266)
- 一种考虑过滤的短纤维增强复合材料 RVE 建模方法…………… 刘丰睿, 骈榕, 赵丽滨, 张建宇 (277)
- 一种新的矩独立重要性测度分析方法及高效算法…………… 巩祥瑞, 吕震宙, 孙天宇, 张雷雷, 封雷 (283)
- 弹性高速飞行器的状态/参数滚动时域估计…………… 陈尔康, 荆武兴, 高长生 (291)
- 数字锁相解调器的优化设计…………… 李勇, 刘泽, 赵鹏飞, 霍继伟, 林阳 (299)
- 基于改进型 Retinex 算法的雾天图像增强技术…………… 张驰, 谭南林, 李响, 李国正, 苏树强 (309)
- 触地关机模式下的着陆器软着陆稳定性研究…………… 董洋, 王春洁, 吴宏宇, 丁宗茂, 满剑锋 (317)
- 蜂群无人机自组网多优先级自适应退避算法…………… 刘炜伦, 张衡阳, 郑博, 高维廷 (325)
- 一种图像缩放算法的 SoC 协同加速设计方法…………… 王鹏, 曹云峰, 许蕾, 丁萌, 张洲宇, 曲金秋 (333)
- 基于地形匹配的直升机低空飞行前视告警方法…………… 张硕俨, 陆洋 (340)
- 未知环境下无人机集群协同区域搜索算法…………… 侯岳奇, 梁晓龙, 何吕龙, 刘流 (347)
- 基于空间两点的视觉自主着陆导引算法设计…………… 魏祥灰, 唐超颖, 王彪 (357)
- 一种新的非高斯随机振动信号的模拟方法…………… 夏静, 袁宏杰, 徐如远 (366)
- 基于改进的动态 Kriging 模型的结构可靠度算法…………… 魏娟, 张建国, 邱涛 (373)
- 芯片互联结构断裂失效的试验研究与统计分析…………… 陈垚君, 景博, 胡家兴, 盛增津, 张钰林 (381)
- 电磁航天器编队悬停鲁棒协同控制方法…………… 张亚博, 师鹏, 张皓, 赵育善 (388)
- 基于半无码的 P(Y) 码自相关 GNSS-R 海面测高方法…………… 樊梦文, 张波, 王峰 (398)
- 基于缓解 HoL 堵塞的单组播混合调度算法…………… 袁龙, 熊庆旭, 萧翰 (405)
- 基于 AVSIMM 算法的高超声速再入滑翔目标跟踪…………… 肖楚晗, 李炯, 雷虎民, 王华吉 (413)
- 考虑临近车道行人对交通流影响的改进跟驰模型……………
…………… 李宏刚, 高哈尔·达吾力, 王帅, 余贵珍, 王朋成 (422)

期刊基本参数: CN 11-2625/V * 1956 * m * A4 * 202 * zh * P * ¥50.00 * 900 * 25 * 2019-02

(编辑 张嵘 李晶 张欣蔚 孙芳 王艳梅 贺伟)

CONTENTS

Vibration response analysis of rotor system with initial thermal deformation
..... MA Yanhong, LIU Haizhou, DENG Wangqun, YANG Hai, HONG Jie (227)

Filtering algorithm of NFOV star sensor measurement delay
..... QIAN Huaming, WANG Di, WU Yonghui, HUANG Zhikai (234)

Elastic-liquid coupling in aerospace rigid-elastic-liquid coupling system
..... LIANG Lifu, GUO Qingyong (243)

Optimization of fuel heat management system for high-speed aircraft
..... PANG Liping, ZOU Lingyu, A Rong, YANG Xiaodong, FAN Jun (252)

BDS/GPS integrated navigation satellite selection algorithm based on chaos particle swarm optimization
..... WANG Ershen, JIA Chaoying, QU Pingping, HUANG Yufeng, PANG Tao, BIE Yuxia, Jiang Yi (259)

Optimized design method of aero-engine rotor structure layout
..... LI Chao, JIN Fuyi, ZHANG Weihao (266)

A filter based modeling method of RVE for short fiber reinforced composite
..... LIU Fengrui, PIAN Rong, ZHAO Libin, ZHANG Jianyu (277)

A new moment-independent importance measure analysis method and its efficient algorithm
..... GONG Xiangrui, LYU Zhenzhou, SUN Tianyu, ZHANG Leilei, FENG Lei (283)

State/parameter moving horizon estimation for elastic hypersonic vehicles
..... CHEN Erkang, JING Wuxing, GAO Changsheng (291)

Optimal design for digital phase-locked demodulator
..... LI Yong, LIU Ze, ZHAO Pengfei, HUO Jiwei, LIN Yang (299)

Foggy image enhancement technology based on improved Retinex algorithm
..... ZHANG Chi, TAN Nanlin, LI Xiang, LI Guozheng, SU Shuqiang (309)

Soft landing stability of lander in mode of shutdown at touchdown
..... DONG Yang, WANG Chunjie, WU Hongyu, DING Zongmao, MAN Jianfeng (317)

An adaptive backoff algorithm for FANETs based on multiple priority differentiation
..... LIU Weilun, ZHANG Hengyang, ZHENG Bo, GAO Weiting (325)

SoC collaborative acceleration design method for image scaling algorithm
..... WANG Peng, CAO Yunfeng, XU Lei, DING Meng, ZHANG Zhouyu, QU Jinqiu (333)

Helicopter forward looking alert method for low-altitude flight based on terrain matching
..... ZHANG Shuoyan, LU Yang (340)

Cooperative area search algorithm for UAV swarm in unknown environment
..... HOU Yueqi, LIANG Xiaolong, HE Lyulong, LIU Liu (347)

Visual guidance algorithm design for autonomous landing based on two points in space
..... WEI Xianghui, TANG Chaoying, WANG Biao (357)

A new simulation method of non-Gaussian random vibration signal
..... XIA Jing, YUAN Hongjie, XU Ruyuan (366)

Structural reliability algorithm based on improved dynamic Kriging model
..... WEI Juan, ZHANG Jianguo, QIU Tao (373)

Experimental research and statistical analysis of fracture failure for interconnected structures in electronic chips
..... CHEN Yaojun, JING Bo, HU Jiaying, SHENG Zengjin, ZHANG Yulin (381)

A robust coordinated control method for hovering of electromagnetic spacecraft formation
..... ZHANG Yabo, SHI Peng, ZHANG Hao, ZHAO Yushan (388)

Semi-codeless based P (Y) code autocorrelation GNSS-R sea surface altimetry method
..... FAN Mengwen, ZHANG Bo, WANG Feng (398)

Packet scheduling algorithm for integrated unicast and multicast traffic based on reducing HoL blocking
..... YUAN Long, XIONG Qingxu, XIAO Han (405)

Hypersonic reentry gliding target tracking based on AVSIMM algorithm
..... XIAO Chuhan, LI Jiong, LEI Humin, WANG Huaji (413)

An improved car-following model considering effect of pedestrians of adjacent lane on traffic flow
..... LI Honggang, Gaoher · DAWULI, WANG Shuai, YU Guizhen, WANG Pengcheng (422)

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

DOI: 10.13700/j.bh.1001-5965.2018.0309

具有初始热变形的转子系统振动响应分析

马艳红^{1,2}, 刘海舟¹, 邓旺群^{3,4}, 杨海^{3,4}, 洪杰^{1,2,*}

(1. 北京航空航天大学 能源与动力工程学院, 北京 100083; 2. 先进航空发动机协同创新中心, 北京 100083;

3. 中国航发湖南航空动力机械研究所, 株洲 412002; 4. 航空发动机振动技术航空科技重点实验室, 株洲 412002)

摘 要: 航空发动机热起动时的温度分布不均会使转子产生初始热变形, 进而引起发动机振动过大, 甚至导致起动失败。针对此问题, 以航空发动机中的典型转子为对象, 根据初始热变形对转子振动的影响建立相应的动力学方程, 并通过模态坐标变换分析初始热变形对转子系统振动响应的影响。结果表明, 初始热变形相当于对转子作用了附加激励, 包括转轴初始弯曲激励、附加不平衡激励和附加陀螺力矩激励, 上述激励均与转速同步。其中, 附加不平衡激励和附加陀螺力矩激励大小与转速有关, 对转子通过各阶临界转速的振动响应均有较大影响; 转轴初始弯曲激励大小与转速无关, 主要影响低阶临界转速的振动响应。

关键词: 转子系统; 热变形; 初始弯曲; 陀螺力矩; 不平衡量

中图分类号: V231.96

文献标识码: A **文章编号:** 1001-5965(2019)02-0227-07

随着航空发动机向高负荷、轻质化的方向发展, 转子系统的振动问题越来越突出。当这类发动机停车后再次起动(即热起动)时, 发动机内的温度分布不均会使转子系统产生初始热变形, 进而使转子振动响应进一步恶化, 甚至导致发动机起动失败。事实上, 国内外对发动机在热起动过程中由初始热变形引起的振动过大的故障屡有报道。朱梓根和晏砺堂^[1]对某涡桨发动机在热起动过程中出现较大的振动响应, 甚至引起转静子碰摩故障的现象进行了研究。张连祥等^[2-3]对某涡扇发动机在热起动过程中振动过大的原因进行了分析。Marinescu^[4]与Baldassarre^[5]等发现, 当发动机热起动时, 轴承上会出现较大的动载荷。因此, 准确描述和分析初始热变形对转子动力学特性的影响具有重要的工程意义。

在早期研究中, 关于初始热变形对转子振动响应影响的研究多以转轴初始热弯曲为重点, 并

以简单转子模型为对象开展理论分析。Nicholas等^[6-7]通过Jeffcott转子模型, 从幅值和相位的角度详细分析了转轴初始弯曲对其不平衡响应的影响规律, 并提出了相应的平衡方法。Shiau和Lee^[8]在Nicholas等^[6-7]的基础上进一步完善, 考虑了转轴初始弯曲对转子在初始不平衡和轮盘初始倾斜下响应的影响。近年来, 研究对象逐渐转变为实际的、较为复杂的转子系统, 但对初始热变形影响的研究仍以转轴初始热弯曲为主。Pennacchi等^[9-10]通过对转子施加与转速同步旋转的弯曲力矩, 研究了带有初始热弯曲的转子的故障诊断方法, 并以实际发动机故障为例, 验证了诊断方法的可行性。Sanches和Pederiva^[11]也进行了相似的研究, 他们对同时带有不平衡量和初始弯曲的转子的诊断方法进行了研究, 认为初始弯曲对转子振动特性的影响可通过与转速同步旋转的激振力来等效。任平珍等^[12-13]利用传递矩阵的

收稿日期: 2018-05-28; 录用日期: 2018-07-13; 网络出版时间: 2018-07-26 18:46

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180725.1730.005.html

基金项目: 国家自然科学基金(51575022); 中央高校基本科研业务费专项资金

* 通信作者. E-mail: hongjie@buaa.edu.cn

引用格式: 马艳红, 刘海舟, 邓旺群, 等. 具有初始热变形的转子系统振动响应分析[J]. 北京航空航天大学学报, 2019, 45(2): 227-233. MA Y H, LIU H Z, DENG W Q, et al. Vibration response analysis of rotor system with initial thermal deformation[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 227-233 (in Chinese).

方法推导了具有热弯曲的转子的振动响应计算方法,并与试验结果进行了对比分析。袁慧群等^[14-15]通过有限元法,计算了某高压转子停车后不同时刻的温度分布、热应力及热变形,并据此计算转子的等效热弯矩来模拟热弯曲对转子振动响应的的影响。

上述文献针对初始热变形引起的转轴初始弯曲对转子振动响应的的影响进行了深入的研究,但 这些研究没有考虑转子初始热变形引起的构件质心变化。此外,针对初始热变形对转子系统振动响应的的影响规律也缺乏详细研究。本文考虑了初始热变形引起的转子构件的质心变化,并根据初始热变形对转子振动响应的的作用效果,建立了相应的动力学方程。通过对方程进行模态坐标变换,分析了初始热变形对转子振动响应的的影响规律。

1 初始热变形转子及其动力学方程

1.1 热状态下的转子变形

航空发动机停车后,虽然在流道内可以采取强制换热措施,但是在转子盘-轴结构上仍会有不均匀的温度分布,其典型分布如图1所示。图中以转子左端下表面的温度为基准,对温度分布进行了无量纲化。转子结构的温度分布为涡轮段高于压气机段,截面上下两端的温度梯度由于与实际的材料、构形和连接界面等因素有关,并无显著规律。

假设转子完全冷却后的温度为 T_0 ,当发动机进行热启动时,由于温度在径向上分布不均,转子截面上端的温度 T_1 与下端的温度 T_2 并不相同,如图2所示。假设截面温度在垂直方向上满足线性分布(线性分布假设只是为了便于说明温度分

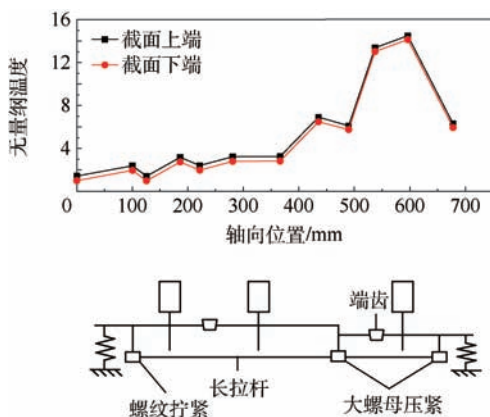


图1 发动机停车后转子温度分布示意图

Fig. 1 Schematic of rotor temperature distribution after engine shutdown

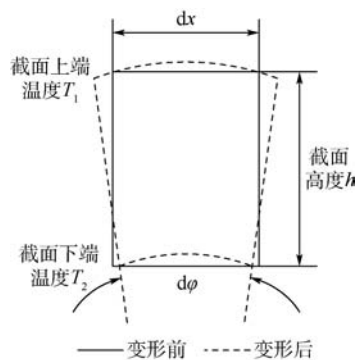


图2 转子微元体热弯曲

Fig. 2 Thermal bow of a rotor element

布对转子弯曲变形的影响,并不一定与实际情况相同),并设转子微元体的转角为 φ ,则由几何关系可得

$$h d\varphi = \alpha(T_2 - T_0) dx - \alpha(T_1 - T_0) dx \quad (1)$$

移项得

$$\frac{d\varphi}{dx} = \frac{\alpha(T_2 - T_1)}{h} \quad (2)$$

式中: α 为材料线膨胀系数; h 为转子截面高度。式(2)说明,转子截面径向的温度梯度通过影响截面的轴向变形使转子发生弯曲变形,且温度梯度越大,转子的弯曲变形越严重。此外,变形还与结构的材料、尺寸等因素有关。

转子截面上的温度梯度还会使截面的径向变形产生差异,进而使其形心偏移。如图3所示,由于转轴和轮盘的结构特征、材料热膨胀系数以及连接界面的影响,当转子产生初始热变形后,轮盘的质心不一定与转轴形心线重合,因此,初始热变形后的轮盘会产生附加质心偏移量 δ 。

设轮盘截面上、下端的位移分别为 $y_{d,u}$ 、 $y_{d,l}$,转轴截面上、下端的位移分别为 $y_{s,u}$ 、 $y_{s,l}$,则轮盘和转轴的中心近似为 $(y_{d,u} + y_{d,l})/2$ 和 $(y_{s,u} + y_{s,l})/2$,因此,轮盘的质心偏移量近似为

$$\delta = \frac{y_{d,u} + y_{d,l}}{2} - \frac{y_{s,u} + y_{s,l}}{2} \quad (3)$$

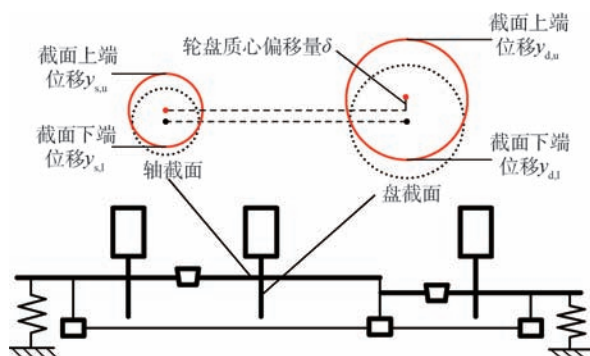


图3 轮盘质心偏移

Fig. 3 Eccentricity of center of mass of a disk

同样,由于转轴和轮盘的结构特征、材料热膨胀系数以及连接界面的影响,当转子产生初始热变形后,轮盘中心线不一定与转轴形心线在轮盘处的切线重合,因此,初始热变形后的轮盘还将产生初始倾斜量 τ ,如图 4 所示。

由上述分析可知,转子温度分布不均引起的转子初始热变形主要有 3 类:转轴弯曲变形、轮盘质心相对旋转中心线偏移、轮盘中心线相对旋转中心线偏移。

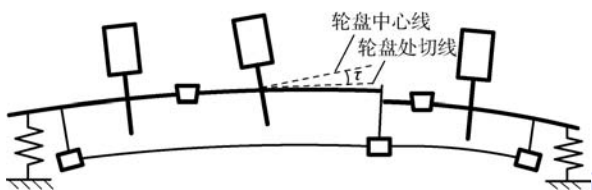


图 4 轮盘初始倾斜
Fig. 4 Initial skewness of a disk

1.2 转子动力学方程

通过有限元法对转子进行振动分析,并采用复数坐标建立动力学方程^[16]:

$$M\ddot{q} - i\omega G\dot{q} + Kq = F \quad (4)$$

式中: M 为质量矩阵; G 为陀螺矩阵; K 为刚度矩阵; ω 为转子转速; q 为转子各节点位移向量, $q = [x_1 + iy_1, \theta_{y1} + i\theta_{x1}, \dots, x_k + iy_k, \theta_{yk} + i\theta_{xk}]^T$; F 为作用于转子上的激励。

设转子由初始热变形引起的转轴初始弯曲、轮盘质心偏移和轮盘初始倾斜分别为 b 、 δ 和 τ 。式(4)的物理意义是各节点处满足力的平衡方程。转轴初始弯曲 b 等效于对转子作用了与转子同步旋转的外力 $Kbe^{i\omega t}$ 。轮盘质心偏移 δ 等效于对转子作用了附加不平衡量 $\varepsilon = m \circ \delta$ (“ \circ ”为 Hadamard 乘积,表示向量各分量相乘),进而引起附加不平衡激励 $\varepsilon\omega^2 e^{i\omega t}$,其中, $m = [\dots, m_k, 0, \dots]^T$ 为轮盘质量列阵, m_k 表示位于节点 k 的轮盘的质量, $\delta = [\dots, \delta_k, 0, \dots]^T$ 为轮盘质心偏移列阵, δ_k 表示位于节点 k 的轮盘的质心偏移量。轮盘初始倾斜会产生陀螺力矩 $(I_p - I_d) \circ \tau\omega^2 e^{i\omega t}$, $I_p = [\dots, 0, I_{p,k}, \dots]^T$ 为轮盘极转动惯量列阵, $I_{p,k}$ 表示位于节点 k 的轮盘的极转动惯量, $I_d = [\dots, 0, I_{d,k}, \dots]^T$ 为轮盘直径转动惯量列阵, $I_{d,k}$ 表示位于节点 k 的轮盘的直径转动惯量, $\tau = [0, \dots, 0, \tau_k, \dots]^T$ 为轮盘初始倾斜角列阵, τ_k 表示位于节点 k 的轮盘的初始倾斜角。将上述表达式代入式(4)可得

$$M\ddot{q} - i\omega G\dot{q} + Kq = [Kb + \varepsilon\omega^2 + (I_p - I_d) \circ \tau\omega^2]e^{i\omega t} \quad (5)$$

由式(5)可知,转子的 3 种初始热变形形式

相当于在转子上作用了 3 种附加激励,且这 3 种激励均与转子同步旋转。其中,转轴初始弯曲激励 $Kbe^{i\omega t}$ 的大小与转速无关,附加不平衡激励 $\varepsilon\omega^2 e^{i\omega t}$ 和附加陀螺力矩激励 $(I_p - I_d) \circ \tau\omega^2 e^{i\omega t}$ 均与轮盘的惯性有关,其大小正比于转速的平方。

需要注意的是,对于实际发动机,轮盘的质心偏移和初始倾斜与转轴和轮盘的结构特征、材料热膨胀系数以及连接界面的影响有关,其沿轴向上的分布特征难以准确确定。因此,最好通过区间数来描述,即 $\varepsilon\omega^2 e^{i\omega t}$ 和 $(I_p - I_d) \circ \tau\omega^2 e^{i\omega t}$ 更准确的表达形式应为 $\varepsilon^1\omega^2 e^{i\omega t}$ 和 $(I_p - I_d) \circ \tau^1\omega^2 e^{i\omega t}$, ε^1 和 τ^1 表示向量中各元素以区间数表示。

2 初始热变形对转子系统振动响应的影响

2.1 转子动力学方程模态坐标变换

转子在转速为 ω 时的自由振动方程为

$$M\ddot{q} - i\omega G\dot{q} + Kq = 0 \quad (6)$$

假设解的形式为 $q = \phi e^{i\Omega t}$, Ω 为转子的涡动角速度, ϕ 为待求的常数列向量,则转子在转速为 ω 时的特征方程为

$$[-\Omega^2 M + \omega\Omega G + K]\phi = 0 \quad (7)$$

令 $\Omega = \omega$,即可求得转子第 i 阶临界转速 $\omega_{cr,i}$ 及其对应振型 ϕ_i 。对于 $\omega_{cr,i}$ 和 ϕ_i ,显然成立

$$[-\omega_{cr,i}^2(M - G) + K]\phi_i = 0 \quad (8)$$

由于采用复数坐标时,陀螺矩阵 G 为对称矩阵,则 $M - G$ 和 K 均为对称矩阵。设 $\Phi = [\phi_1, \phi_2, \dots, \phi_n]^T$,则由振动力学可知, Φ 使得式(9)成立:

$$\Phi^T(M - G)\Phi = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} \quad (9)$$

$$\Phi^T K \Phi = \begin{bmatrix} \omega_{cr,1}^2 & & \\ & \ddots & \\ & & \omega_{cr,n}^2 \end{bmatrix} \quad (10)$$

当转子存在转轴初始弯曲激励、附加不平衡激励和附加陀螺力矩激励时,其动力学方程为

$$M\ddot{q} - i\omega G\dot{q} + Kq = Kbe^{i\omega t} + \varepsilon^1\omega^2 e^{i\omega t} + (I_p - I_d) \circ \tau^1\omega^2 e^{i\omega t} \quad (11)$$

设解的形式为 $q = \Phi\eta e^{i\omega t} = \sum_{i=1}^n \phi_i \eta_i e^{i\omega t}$, $\eta = [\eta_1, \eta_2, \dots, \eta_n]^T$ 表示模态坐标,代入方程式(11)并左乘 Φ^T 得

$$[-\omega^2 \Phi^T(M - G)\Phi + \Phi^T K \Phi]\eta = \Phi^T [Kb + \varepsilon^1\omega^2 + (I_p - I_d) \circ \tau^1\omega^2] \quad (12)$$

即

$$\begin{bmatrix} -\omega^2 + \omega_{cr,1}^2 & & & \\ & \ddots & & \\ & & -\omega^2 + \omega_{cr,n}^2 & \\ & & & \end{bmatrix} \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} = \begin{bmatrix} \phi_1^T [Kb + \varepsilon^1 \omega^2 + (I_p - I_d) \circ \tau^1 \omega^2] \\ \vdots \\ \phi_n^T [Kb + \varepsilon^1 \omega^2 + (I_p - I_d) \circ \tau^1 \omega^2] \end{bmatrix} = \begin{bmatrix} \phi_1^T f \\ \vdots \\ \phi_n^T f \end{bmatrix} \quad (13)$$

因此

$$\eta_i = \frac{\phi_i^T f}{\omega_{cr,i}^2} \cdot \frac{1}{1 - \omega^2/\omega_{cr,i}^2} = \frac{\phi_i^T f}{\omega_{cr,i}^2} \cdot \frac{1}{1 - \lambda_i^2} \quad (14)$$

式中: $f = [f_1, f_2, \dots, f_n]^T$ 为节点载荷列阵; $\lambda_i = \omega/\omega_{cr,i}$ 为第 i 阶频率比。

因此,转子的振动响应可写为

$$q = \Phi \eta e^{i\omega t} = \sum_{i=1}^n \phi_i \frac{\phi_i^T f}{\omega_{cr,i}^2} \cdot \frac{1}{1 - \lambda_i^2} e^{i\omega t} \quad (15)$$

将具体的热变形激励的表达式代入式(15)即可分析其对转子振动响应的影响。

2.2 转子初始热变形对振动响应的影响

图5为航空发动机中的典型转子。其临界转速对应的振型分别为平动振型、俯仰振型(统称刚体振型)和一弯振型,且工作转速一般位于刚体振型临界转速以上,靠近一弯振型临界转速。由式(15)可知,当转子分别通过各阶临界转速时,其振动响应近似为

$$q = \Phi \eta e^{i\omega t} = \sum_{j=1}^n \phi_j \frac{\phi_j^T f}{\omega_{cr,j}^2} \cdot \frac{1}{1 - \lambda_j^2} e^{i\omega t} \approx \phi_k \frac{\phi_k^T f}{\omega_{cr,i}^2} \cdot \frac{1}{1 - \lambda_i^2} e^{i\omega t} \quad (16)$$

相应的转子位移和支承反力分别为

$$q \approx \phi_i \frac{\phi_i^T [Kb + \varepsilon^1 \omega^2 + (I_p - I_d) \circ \tau^1 \omega^2]}{\omega_{cr,i}^2} \cdot \frac{1}{1 - \lambda_i^2} e^{i\omega t} \quad (17)$$

$$F_{brg} \approx k_{brg} x_{i,brg} = \frac{\phi_i^T [Kb + \varepsilon^1 \omega^2 + (I_p - I_d) \circ \tau^1 \omega^2]}{\omega_{cr,i}^2} \cdot \frac{1}{1 - \lambda_i^2} e^{i\omega t} \quad (18)$$

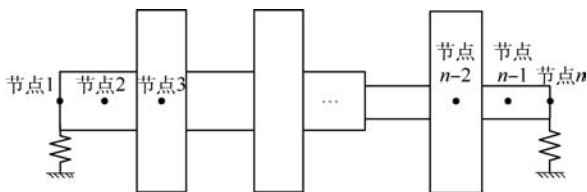


图5 典型航空发动机转子模型

Fig. 5 A model of typical aeroengine rotors

式中: $x_{i,brg}$ 为支点在第 i 阶振型中的位移; k_{brg} 为支点支承刚度。

由式(17)、式(18)可分析转子通过各阶临界转速时的振动响应,设对各阶临界转速均有 $1/(1 - \lambda_i^2) = \text{const}$,则可以看出:

1) 轮盘质心偏移和初始倾斜会在支承反力表达式(18)中引入附加不平衡项 $\varepsilon^1 \omega^2$ 和附加陀螺力矩项 $(I_p - I_d) \circ \tau^1 \omega^2$,当这2个矢量的幅值和方向满足一定的条件时, $x_{i,brg} \phi_i^T [Kb + \varepsilon^1 \omega^2 + (I_p - I_d) \circ \tau^1 \omega^2]$ 的幅值会大于 $x_{i,brg} \phi_i^T Kb$ 的幅值。即与仅考虑转轴初始弯曲的转子相比,当考虑轮盘质心变化时,转子的支承反力可能进一步加大。

2) 附加不平衡激励 $\varepsilon^1 \omega^2$ 和附加陀螺力矩激励 $(I_p - I_d) \circ \tau^1 \omega^2$ 的大小正比于转速的平方,当 ω 分别靠近 $\omega_{cr,i}$ ($i = 1, 2, 3$) 时, $\omega^2/\omega_{cr,i}^2 \approx \text{const}$,即这2种激励对转子通过频率较低的平动、俯仰振型临界转速和频率较高的一弯振型临界转速时的支承反力均可能有较大的影响,由 $x_{i,brg} \phi_i^T [\varepsilon^1 + (I_p - I_d) \circ \tau^1]$ 可知,其影响程度与转子的振型及激励的具体分布特征有关。相反,转轴初始弯曲激励 $Kb e^{i\omega t}$ 大小与转速无关,其主要影响转子通过频率较低的平动、俯仰振型临界转速时的支承反力。

3) 附加不平衡激励 $\varepsilon^1 \omega^2$ 和附加陀螺力矩激励 $(I_p - I_d) \circ \tau^1 \omega^2$ 沿轴向分布具有非确定性,使得振动响应也具有区间分布的特点,振动响应的这种非确定性可能使转子的支承反力进一步加大。因此,如何更准确地确定附加不平衡激励和附加陀螺力矩激励的分布需要进一步的研究,如在实际工程应用中通过统计的方法进行分析等。

3 典型航空发动机转子振动响应分析

如图6所示,燃气发生器转子为两支点支承的大长径比转子,且工作转速靠近弯曲临界转速,是航空发动机中典型的转子系统。以该实际转子为例,分析其存在初始热变形时的振动响应。

该转子支点跨度为 600 mm,最大直径为 200 mm,轴承直径为 50 mm,压气机材料为 TC11,

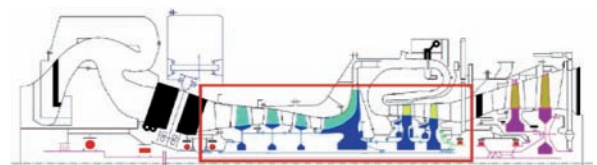


图6 典型航空发动机转子

Fig. 6 A typical aeroengine rotor

涡轮材料为 GH4169。在 ANSYS 中,采用梁单元 BEAM188 模拟转子,质量单元 MASS21 模拟轮盘,弹簧单元 COMBIN14 模拟支点,节点数为 14,单元数为 16。

按照 1.2 节中的热变形转子动力学方程(式(5)),根据转子在给定温度场下初始热变形的有限元计算结果,转轴初始弯曲激励可通过作用于转子轴段中部的径向同步激振力(约 5 750 N)来模拟。对于附加不平衡激励和附加陀螺力矩激励,这里仅考虑第一级涡轮盘。根据初始热变形的计算结果,涡轮盘的附加不平衡量约为 $45 \text{ g} \cdot \text{mm}$,考虑连接结构界面接触等不确定因素的影响,以区间数表示不平衡量变化范围 $[36, 54] \text{ g} \cdot \text{mm}$ (偏差范围取为 20%)。同样地,对于涡轮盘的附加陀螺力矩激励,由于难以准确确定数值,根据计算和统计分析取为 $[0.0008, 0.0012] \text{ rad}$ (偏差范围取为 20%)。上述激励的相位均相同。按照上述激励,计算转子系统在不同转速下的振动响应,如图 7 ~ 图 9 所示。需要说明的是,由于仅是 3 种激励的幅值存在区间分布,且 3 种激励作用的节点均只有 1 个,因此,转子分别在这 3 种激励下的振动响应为激励幅值的单调函数,其区间的上/下界可直接由激励幅值的上/下界确定。

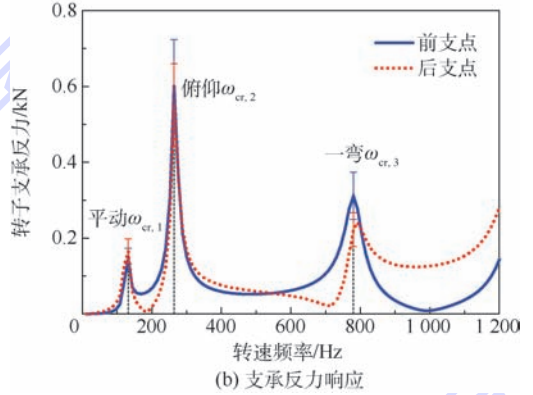
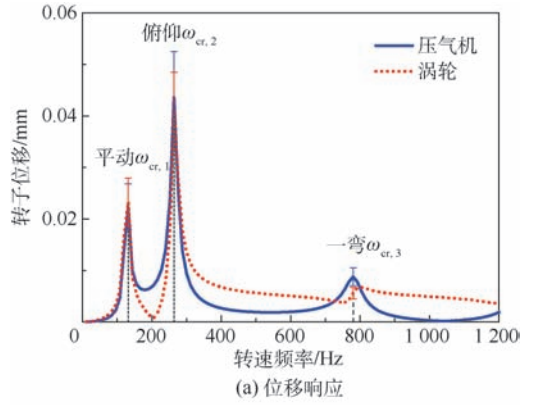


图 8 转子在附加不平衡激励下的振动响应
Fig. 8 Vibration response of rotor under additional unbalance excitation

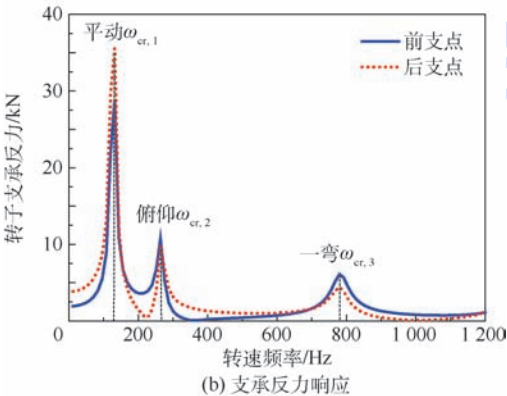
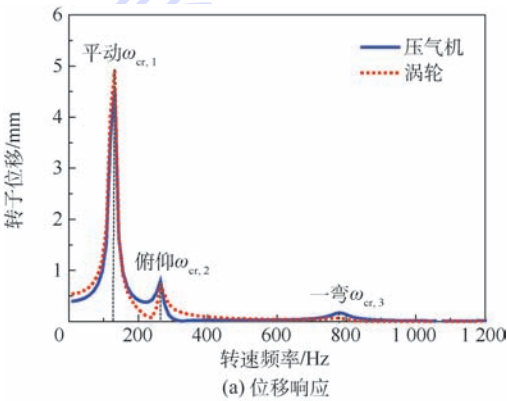


图 7 转子在初始弯曲激励下的振动响应
Fig. 7 Vibration response of rotor under initial bow excitation

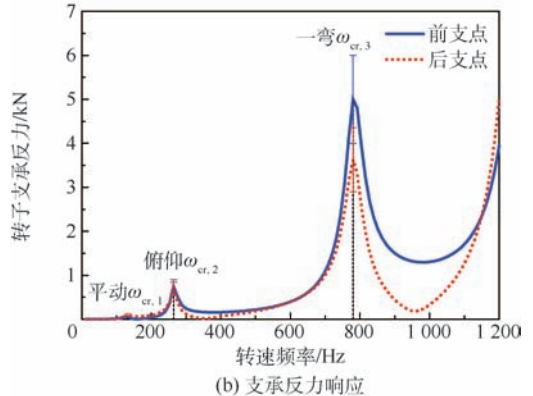
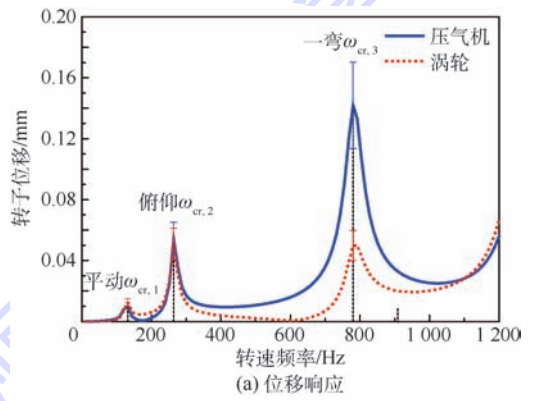


图 9 转子在附加陀螺力矩激励下的振动响应
Fig. 9 Vibration response of rotor under additional gyroscopic moment excitation

由图7~图9可知:①对于本算例,由于3种激励的相位相同,附加不平衡激励和附加陀螺力矩激励会加大支点支承反力的振动响应。由2.2节中的式(18)可知,这2种激励对转子通过各阶临界转速时振动响应的影响程度由 $x_{i,brg} \phi_i^T \cdot [\varepsilon^1 + (I_p - I_d) \cdot \tau^1]$ 决定,即与转子振型及激励位置均有关。在本算例中,附加不平衡激励主要影响转子通过俯仰振型的支承反力,附加陀螺力矩激励主要影响转子通过弯曲振型的支承反力。②附加不平衡激励和附加陀螺力矩激励的大小及相位具有非确定性特征,使得相应的振动响应也呈现区间分布特点,这种非确定性可能使转子不同支点处的支承反力具有分散性。激励的这种非确定性与转子的结构特征和温度分布特征有关,因此,对于由不同材料构件组成的非连续转子系统,工程上可以对盘-轴结构的热变形特性进行统计分析。

4 结 论

基于理论和仿真,本文主要结论如下:

1) 转子结构系统温度分布不均产生的初始热变形会对转子的振动响应产生影响,根据力学特征可分为:转轴初始弯曲、轮盘质心偏移和轮盘初始倾斜,分别使转子产生转轴初始弯曲激励、附加不平衡激励和附加陀螺力矩激励。

2) 当附加不平衡激励和附加陀螺力矩激励的幅值和方向满足一定的条件时,其与转轴初始弯曲激励在支点处引起的振动响应之和大于仅由初始弯曲激励在支点处引起的振动响应,即与仅考虑转轴初始弯曲的转子相比,附加不平衡激励和附加陀螺力矩激励可能加大支点支承反力的振动响应。

3) 附加不平衡激励和附加陀螺力矩激励的大小正比于转速的平方,其对转子通过频率较低和较高的临界转速时的支承反力均可能有较大的影响,具体影响程度与转子的振型及激励的分布特征有关。而转轴初始弯曲激励大小与转速无关,其主要影响转子通过频率较低的临界转速时的支承反力。

4) 由于航空发动机转子结构材料不同且存在连接界面和几何结构突变,其盘-轴结构系统的初始热变形难以确定,使得附加不平衡激励和附加陀螺力矩激励的大小与相位具有非确定性,进而,转子系统振动响应表现出分散性和区间分布特点。因此,在工程应用中可以对具体结构统计

分析其热变形特征。

参考文献 (References)

- [1] 朱梓根,晏砺堂.某型涡轮螺桨发动机转子偏磨故障分析[J].航空学报,1992,13(10):512-516.
ZHU Z G, YAN L T. Troubleshooting of the rotor rubbing case for some turboprop engine[J]. Acta Aeronautica et Astronautica Sinica, 1992, 13(10): 512-516 (in Chinese).
- [2] 张连祥.航空发动机转子热弯曲引发的典型故障分析[J].振动与冲击,2008,27(S):7-9.
ZHANG L X. Analysis of typical aeroengine vibration fault related to thermal bowed rotor[J]. Journal of Vibration and Shock, 2008, 27(S): 7-9 (in Chinese).
- [3] 张连祥,王娟.航空发动机热起动过程中的振动问题分析[J].振动与冲击,2010,29(S):132-134.
ZHANG L X, WANG J. Analysis of aeroengine vibration fault during the hot-starting[J]. Journal of Vibration and Shock, 2010, 29(S): 132-134 (in Chinese).
- [4] MARINESCU G, EHRSAM A. Experimental investigation into thermal behavior of steam turbine components-Part 2: Natural cooling of steam turbines and the impact on LCF life [C] // ASME Turbo-Power for Land, Sea, and Air. New York: ASME, 2012, 4: 1111-1120.
- [5] BALDASSARRE L, FONTANA M. Modelling of rotor bow during hot restart in centrifugal compressors [C] // 39th Turbomachinery Symposium, 2010: 1-8.
- [6] NICHOLAS J C, GUNTER E J, ALLAIRE P E. Effect of residual shaft bow on unbalance response and balancing of a single mass flexible rotor-Part 1: Unbalance response [J]. Journal of Engineering for Power, 1976, 98(2): 171-181.
- [7] NICHOLAS J C, GUNTER E J, ALLAIRE P E. Effect of residual shaft bow on unbalance response and balancing of a single mass flexible rotor-Part 2: Balancing [J]. Journal of Engineering for Power, 1976, 98(2): 182-187.
- [8] SHIAU T N, LEE E K. The residual shaft bow effect on dynamic response of a simply supported rotor with disk skew and mass unbalance [J]. Journal of Vibration, Acoustics, Stress, and Reliability in Design, 1989, 111(2): 170-178.
- [9] BACHSCHMID N, PENNACCHI P, VANIA A. Identification of multiple faults in rotor systems [J]. Journal of Sound and Vibration, 2002, 254(2): 327-366.
- [10] PENNACCHI P, VANIA A. Accuracy in the identification of a generator thermal bow [J]. Journal of Sound and Vibration, 2004, 274(1-2): 273-295.
- [11] SANCHES F D, PEDERIVA R. Theoretical and experimental identification of the simultaneous occurrence of unbalance and shaft bow in a Laval rotor [J]. Mechanism and Machine Theory, 2016, 101: 209-221.
- [12] 任平珍,柴卫东,胡壁刚,等.航空发动机转子热弯曲稳态响应计算方法研究[J].燃气涡轮试验与研究,1996(3): 27-32.
REN P Z, CHAI W D, HU B G, et al. Steady state response calculation of aeroengine rotor with thermal bow [J]. Gas Turbine Experiment and Research, 1996(3): 27-32 (in Chinese).

- [13] 任平珍,陆山,赵明. 转子热弯曲变形及其影响的数值分析方法[J]. 机械科学与技术, 1997, 16(2):95-98.
REN P Z, LU S, ZHAO M. Numerical analysis method of thermal deflection and its affection on vibration response of rotor [J]. Mechanical Science and Technology, 1997, 16(2):95-98 (in Chinese).
- [14] 袁慧群,朱向哲,李东,等. 转子系统瞬态热启动过程动力学特性研究[J]. 振动与冲击, 2009, 28(7):33-37.
YUAN H Q, ZHU X Z, LI D, et al. Dynamic characteristics of transient thermal starting up of a rotor system [J]. Journal of Vibration and Shock, 2009, 28(7):33-37 (in Chinese).
- [15] 朱向哲,贺威,袁慧群. 稳态温度场对转子系统振动特性的影响[J]. 东北大学学报(自然科学版), 2008, 29(1):113-116.
ZHU X Z, HE W, YUAN H Q. Effects of steady temperature field on vibrational characteristics of a rotor system [J]. Journal of Northeastern University (Natural Science), 2008, 29(1):113-116 (in Chinese).
- [16] GENTA G. Dynamics of rotating systems [M]. Berlin: Springer, 2005:170.

作者简介:

马艳红 女,博士,教授,博士生导师。主要研究方向:航空发动机整机动力学、旋转机械振动控制、智能结构与新型阻尼材料等。

洪杰 男,博士,教授,博士生导师。主要研究方向:航空发动机转子动力学、航空发动机整机动力学、智能结构与新型阻尼材料等。

Vibration response analysis of rotor system with initial thermal deformation

MA Yanhong^{1,2}, LIU Haizhou¹, DENG Wangqun^{3,4}, YANG Hai^{3,4}, HONG Jie^{1,2,*}

(1. School of Energy and Power Engineering, Beihang University, Beijing 100083, China; 2. Collaborative Innovation Center of Advanced Aero-engine, Beijing 100083, China; 3. AECC Hunan Aviation Powerplant Research Institute, Zhuzhou 412002, China; 4. Aviation Key Laboratory of Aero-engine Vibration Technology, Zhuzhou 412002, China)

Abstract: Uneven temperature distribution of an aero-engine during hot restart can lead to initial thermal deformation of its rotors, which causes excessive engine vibration, or even a start-up failure. In view of the problem, this paper takes a typical rotor in aero-engines as an object, establishes its dynamic equation based on the effects of thermal deformation on vibration, and analyzes the effects of thermal deformation on vibration response of rotor system through modal coordinate transformation. The results show that thermal deformation is equivalent to additional excitations acting on the rotor, that is, initial shaft bow excitation, additional unbalance excitation and additional gyroscopic moment excitation, all of which are synchronous with the rotating speed. Moreover, the amplitude of unbalance and gyroscopic moment excitation is related to the rotating speed, therefore influencing rotor vibration response at different critical speeds to an equivalent extent, while that of initial shaft bow excitation is not related to the rotating speed and the response of low-order critical rotating speeds is more affected.

Keywords: rotor system; thermal deformation; initial bow; gyroscopic moment; unbalance

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

DOI: 10.13700/j.bh.1001-5965.2018.0279

小视场星敏感器量测延时滤波算法

钱华明*, 王迪, 吴永慧, 黄智开

(哈尔滨工程大学 自动化学院, 哈尔滨 150001)

摘 要: 针对小视场(NFOV)星敏感器用于姿态估计时存在的量测延时情况,提出了一种用于解决量测延时的鲁棒扩展卡尔曼滤波(REKF)算法。根据最小方差准则的思想求解各方差的最小上界,通过最小上界确定滤波增益,设计的REKF算法可以有效解决量测延时问题,提高了姿态估计的精度。对REKF算法进行了仿真验证,结果表明:该算法优于常规加性扩展卡尔曼滤波(AEKF)算法、鲁棒有界时域滤波(RFHF)算法及鲁棒卡尔曼滤波(RKF)算法,能较好解决非线性系统存在的量测延时问题,验证了该算法的有效性。

关键词: 小视场(NFOV)星敏感器;姿态估计;扩展卡尔曼滤波(EKF);鲁棒滤波;量测延时

中图分类号: U666.12

文献标识码: A

文章编号: 1001-5965(2019)02-0234-09

小视场(Narrow Filed of View, NFOV)星敏感器成本较低,并且在极端天气下更加适用,因此小视场星敏感器的姿态估计算法具有重要的研究价值。通过对导航星表进行预处理,导致星图识别过程中识别时间增加,影响星敏感器输出信息的时间,造成量测延时问题。同时,光学系统以及确定姿态等步骤的繁琐性,也导致了姿态敏感器输出信息的同时带有延时现象。

延时问题的存在,导致系统模型中存在着不确定性。鲁棒滤波算法被证实是一种处理模型不确定性问题有效算法^[1-3]。文献[4]讨论了针对有界模型不确定的线性时域系统二阶鲁棒卡尔曼滤波(Robust Kalman Filter, RKF)算法,但该模型不适用于非线性系统。文献[5]提出鲁棒扩展卡尔曼滤波(Robust Extended Kalman Filter, REKF)算法,通过预测-校正形式估计刚体姿态,所提出的姿态估计系统考虑四元数刚体模型,其建模时主要考虑的是加性噪声,但其不适于乘性耦合噪声的环境。文献[6]设计了一种RKF算法

处理乘性耦合噪声,但其不能很好地解决量测延时带来的影响。文献[7]设计了一种RKF算法用于处理状态延迟与量测丢失的时变离散系统,通过求取状态估计协方差的最小上界来计算需要的参数。文献[8]提出了鲁棒有界时域滤波(Robust Finite Horizon Filter, RFHF)算法用于解决确定概率分布情况的延迟系统。但是文献[7-8]这2种算法只适用于线性系统。文献[9]提出了基于5阶球半径规则的容积卡尔曼滤波算法,使其更适用于非线性系统,解决传感器随机延时的问,然而该算法在建模时没有考虑乘性耦合噪声,且与本文所研究的实际情况并不相符。

因此,针对小视场星敏感器与陀螺仪结合的模式,本文提出了REKF算法用于处理量测延时因素引起的模型不确定情况。通过状态扩维理论获得带有量测延时的非线性系统模型,并且建立的系统模型包含乘性耦合噪声。根据最小方差准则近似确定估计误差方差的最小上界范围,通过该最小上界选取滤波增益参数,从而得到量

收稿日期: 2018-05-17; 录用日期: 2018-08-24; 网络出版时间: 2018-09-05 09:34

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180903.1545.005.html

基金项目: 国家自然科学基金(61573113)

*通信作者. E-mail: qianhuam@sina.com

引用格式: 钱华明, 王迪, 吴永慧, 等. 小视场星敏感器量测延时滤波算法[J]. 北京航空航天大学学报, 2019, 45(2): 234-242.
QIAN H M, WANG D, WU Y H, et al. Filtering algorithm of NFOV star sensor measurement delay [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 234-242 (in Chinese).

测延时情况下的 REKF 算法。仿真结果表明,本文算法所能达到的估计精度较原有算法有明显提高。

1 组合姿态算法设计

1.1 系统模型

1.1.1 陀螺的量测模型

在姿态估计系统中,陀螺仪用来测量卫星转过角速率的姿态敏感器。陀螺仪产生的误差因素有很多,如失准角误差、标度因素误差、漂移误差、时间延迟等。大部分的误差可以通过校正来补偿,而随机漂移、噪声因素等是主要的误差源,故陀螺量测模型可表示为^[10]

$$\begin{cases} \tilde{\omega} = \omega + b + \eta_u \\ \dot{b} = \eta_v \end{cases} \quad (1)$$

式中: $\tilde{\omega}$ 为测量到的陀螺角速率; ω 为真实的陀螺角速率; b 为陀螺漂移; η_u 和 η_v 分别为陀螺漂移噪声和陀螺测量噪声,它们是互相独立的、均值为零、服从高斯分布的高斯白噪声, σ_u^2 、 σ_v^2 表示对应的方差。

1.1.2 星敏感器的量测模型

用四元数来表示量测模型为

$$p_b = A(q)p_o + v_s \quad (2)$$

式中: $A(q) = C_o^b$, C_o^b 为轨道坐标系到本体坐标系

之间的姿态矩阵, C_i^o 为惯性坐标系到轨道坐标系之间的姿态矩阵; v_s 为星敏感器测量白噪声,服从 $N(0, \sigma_s^2)$; p_o 为轨道坐标系下的矢量坐标; p_b 为本体坐标系下的输出矢量。

1.2 状态方程与量测方程

由四元数轨道动力学方程可得

$$\dot{q} = \frac{1}{2} \begin{bmatrix} \omega \\ 0 \end{bmatrix} \otimes q = \frac{1}{2} \Omega(\omega) q = \frac{1}{2} \Xi(q) \omega \quad (3)$$

式中: $\omega = [\omega_x \ \omega_y \ \omega_z]^T$ 为输出的陀螺角速率; $q = [q_1 \ q_2 \ q_3 \ q_4]^T = [\rho \ q_4]^T$ 为真实四元数;

$\Omega(\omega) = \begin{bmatrix} -[\omega \times] & \omega \\ -\omega^T & 0 \end{bmatrix}$, $[\omega \times]$ 表示 ω 的反对称矩阵,

$$[\omega \times] = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}; \Xi(q) = \begin{bmatrix} q_4 I_{3 \times 3} + [\rho \times] \\ -\rho^T \end{bmatrix}。$$

为了给出具体的状态方程与量测方程,需选取状态量。这里选取 k 时刻姿态四元数 q_k 和陀螺漂移 b_k 作为状态变量,即 $x_k = [q_k^T \ b_k^T]^T$,那么离散化后的状态方程为

$$x_{k+1} = f(x_k, \tilde{\omega}_k) + \sum_{i=1}^s A_{ik} \eta_{ik} x_k + w_k \quad (4)$$

式中:

$$f(x_k, \tilde{\omega}_k) = \begin{bmatrix} \cos\left(\frac{\|\tilde{\omega} - b_k\| \Delta t}{2}\right) I_{4 \times 4} + \sin\left(\frac{\|\tilde{\omega} - b_k\| \Delta t}{2}\right) \frac{\Omega(\tilde{\omega}_k - b_k)}{\|\tilde{\omega}_k - b_k\|} & \mathbf{0}_{4 \times 3} \\ \mathbf{0}_{3 \times 4} & I_{3 \times 3} \end{bmatrix} \begin{bmatrix} q_k \\ b_k \end{bmatrix}$$

ω_k 为满足均值为零、方差为 $Q_k =$

$\begin{bmatrix} \mathbf{0}_{4 \times 4} & \mathbf{0}_{4 \times 3} \\ \mathbf{0}_{3 \times 4} & \Delta t \sigma_u^2 I_{3 \times 3} \end{bmatrix}$ 的高斯白噪声; $\sum_{i=1}^s A_{ik} \eta_{ik} x_k$ 为

乘性噪声项, s 取3, η_{ik} 表示均值为零、方差为1的噪声, A_{ik} 表示有适当阶数的确定矩阵:

$$A_{ik} = -\frac{\Delta t \sigma_v}{2} \begin{bmatrix} A_{ik}^i & \mathbf{0}_{4 \times 3} \\ \mathbf{0}_{3 \times 4} & \mathbf{0}_{3 \times 3} \end{bmatrix} \quad (5)$$

$$A_{ik}^1 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}$$

$$A_{ik}^2 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

$$A_{ik}^3 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

根据式(2),选取星敏感器输出的姿态四元数作为量测变量,选取3个参考矢量描述姿态估计系统,故系统的量测模型为

$$z_k = \begin{bmatrix} z_k^1 \\ z_k^2 \\ z_k^3 \end{bmatrix} = \begin{bmatrix} A(q_k) r_k^1 \\ A(q_k) r_k^2 \\ A(q_k) r_k^3 \end{bmatrix} + \begin{bmatrix} v_k^1 \\ v_k^2 \\ v_k^3 \end{bmatrix} \quad (6)$$

式中: r_k^i 为星敏感器在参考坐标系下对应的参考向量; v_k^i 为满足均值为零、方差为 $R_k = \sigma_s^2 I_{3 \times 3}$ 的高斯白噪声。

由于星敏感器进行姿态估计时存在量测延时现象,上述建立的星敏感器量测模型没有考虑延

时的存在,使得输出结果并不准确,笔者考虑一步延时的星敏感器量测模型:

$$\mathbf{y}_k = (\mathbf{I} - \mathbf{\Gamma}_k)\mathbf{z}_k + \mathbf{\Gamma}_k\mathbf{z}_{k-1} \quad (7)$$

式中: \mathbf{y}_k 为 k 时刻真实的量测量; $\mathbf{\Gamma}_k$ 为不同的延迟速率, $\mathbf{\Gamma}_k = \text{diag}\{\mu_{k,1}, \mu_{k,2}, \dots, \mu_{k,m}\}$, $\mu_{k,i}(i=1,2,\dots,m)$ 服从伯努利分布(0~1分布),满足:

$$\begin{cases} p(\mu_{k,i} = 1) = E(\mu_{k,i}) = p_{k,i} \\ p(\mu_{k,i} = 0) = 1 - E(\mu_{k,i}) = 1 - p_{k,i} \end{cases} \quad (8)$$

故根据式(4)、式(6)、式(7)可以得出关于量测延时及乘性噪声的非线性离散系统:

$$\begin{cases} \mathbf{x}_{k+1} = f(\mathbf{x}_k, \tilde{\omega}_k) + \sum_{i=1}^s \mathbf{A}_{ik}\boldsymbol{\eta}_{ik}\mathbf{x}_k + \mathbf{w}_k \\ \mathbf{z}_k = h(\mathbf{x}_k) + \sum_{i=1}^r \mathbf{C}_{ik}\boldsymbol{\xi}_{ik}\mathbf{x}_k + \mathbf{v}_k \\ \mathbf{y}_k = (\mathbf{I} - \mathbf{\Gamma}_k)\mathbf{z}_k + \mathbf{\Gamma}_k\mathbf{z}_{k-1} \end{cases} \quad (9)$$

式中: \mathbf{C}_{ik} 为有恰当维数的已知矩阵; $\boldsymbol{\xi}_{ik}$ 为均值为零、方差为1的噪声; \mathbf{v}_k 为零均值高斯白噪声。由于实际的量测输出 \mathbf{y}_k 与 $\mathbf{z}_k, \mathbf{z}_{k-1}$ 两个时刻的量测量均相关,需要利用状态扩维理论,使其得到相同时刻量测量的表达式。

$$\begin{cases} \mathbf{X}_{k+1} = f'(\mathbf{X}_k) + g(\mathbf{X}_k, \boldsymbol{\eta}_k) + \mathbf{W}_k \\ \mathbf{y}_k = \boldsymbol{\gamma}_k(h'(\mathbf{X}_k) + s(\mathbf{X}_k, \boldsymbol{\xi}_k) + \mathbf{V}_k) \end{cases} \quad (10)$$

式中:

$$\mathbf{X}_{k+1} = \begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{x}_k \end{bmatrix}$$

$$f'(\mathbf{X}_k) = \begin{bmatrix} f(\mathbf{x}_k) \\ f(\mathbf{x}_{k-1}) \end{bmatrix}$$

$$g(\mathbf{X}_k, \boldsymbol{\eta}_k) = \begin{bmatrix} \sum_{i=1}^s \mathbf{A}_{ik}\boldsymbol{\eta}_{ik}\mathbf{x}_k \\ \sum_{i=1}^s \mathbf{A}_{ik-1}\boldsymbol{\eta}_{ik-1}\mathbf{x}_{k-1} \end{bmatrix}$$

$$s(\mathbf{X}_k, \boldsymbol{\xi}_k) = \begin{bmatrix} \sum_{i=1}^r \mathbf{C}_{ik}\boldsymbol{\xi}_{ik}\mathbf{x}_k \\ \sum_{i=1}^r \mathbf{C}_{ik-1}\boldsymbol{\xi}_{ik-1}\mathbf{x}_{k-1} \end{bmatrix}$$

$$\mathbf{W}_k = \begin{bmatrix} \mathbf{w}_k \\ \mathbf{w}_{k-1} \end{bmatrix}$$

$$h'(\mathbf{X}_k) = \begin{bmatrix} h(\mathbf{x}_k) \\ h(\mathbf{x}_{k-1}) \end{bmatrix}$$

$$\mathbf{V}_k = \begin{bmatrix} \mathbf{v}_k \\ \mathbf{v}_{k-1} \end{bmatrix}$$

$$\boldsymbol{\gamma}_k = [\mathbf{I} - \mathbf{\Gamma}_k \quad \mathbf{\Gamma}_k]$$

$$\bar{\boldsymbol{\gamma}}_k = E(\boldsymbol{\gamma}_k) = [\mathbf{I} - \bar{\mathbf{\Gamma}}_k \quad \bar{\mathbf{\Gamma}}_k]$$

$$\tilde{\boldsymbol{\gamma}}_k = \boldsymbol{\gamma}_k - \bar{\boldsymbol{\gamma}}_k = [\bar{\mathbf{\Gamma}}_k - \mathbf{\Gamma}_k \quad \mathbf{\Gamma}_k - \bar{\mathbf{\Gamma}}_k]$$

$$E(\mathbf{W}_k\mathbf{W}_k^T) = \bar{\mathbf{Q}}_k = \begin{bmatrix} \mathbf{Q}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_{k-1} \end{bmatrix}$$

$$E(\mathbf{V}_k\mathbf{V}_k^T) = \bar{\mathbf{R}}_k = \begin{bmatrix} \mathbf{R}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{k-1} \end{bmatrix}$$

$$E(\mathbf{W}_k\mathbf{V}_k^T) = \begin{bmatrix} \mathbf{W}_k\mathbf{V}_k^T & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_{k-1}\mathbf{V}_{k-1}^T \end{bmatrix}$$

$\bar{\mathbf{\Gamma}}_k = \text{diag}\{p_{k,1}, p_{k,2}, \dots, p_{k,m}\}$,由 $\mathbf{\Gamma}_k$ 的分布特性可知, $\tilde{\boldsymbol{\gamma}}_k$ 是一个零均值的随机矩阵序列。由文献[11]可知, $h(\mathbf{x}_k)$ 满足:

$$\|h'(\mathbf{X}_k)\| \leq \sigma_1 \|\mathbf{X}_k\| + \sigma_2 \mathbf{I} \quad (11)$$

式中: σ_1, σ_2 为非负标量。

2 鲁棒扩展卡尔曼滤波算法

针对带有量测延时及乘性噪声的离散系统式(9),经过状态扩维理论,获得一个相对简单的模型式(10),提出改进的REKF算法,并且计算预测方差,给出估计误差方差的最小上界。

一步状态预测

$$\hat{\mathbf{X}}_{k+1|k} = f'(\hat{\mathbf{X}}_k) \quad (12)$$

状态更新

$$\hat{\mathbf{X}}_{k+1} = \hat{\mathbf{X}}_{k+1|k} + \mathbf{K}_{k+1}(\mathbf{y}_{k+1} - \bar{\boldsymbol{\gamma}}_k h'(\hat{\mathbf{X}}_{k+1|k})) \quad (13)$$

式中: $\hat{\mathbf{X}}_k$ 为在时间 k 处扩维后的状态估计; $\hat{\mathbf{X}}_{k+1|k}$ 为在时间 $k+1$ 处的一步预测; \mathbf{K}_{k+1} 为滤波器增益。

根据以上分析,对于带有乘性噪声及量测延时的不确定性系统,就是设计式(10)的REKF算法的问题。由于实际的误差方差难以求出,只能通过确定滤波增益参数来确保误差方差的最小上界,即

$$\mathbf{P}_{k+1} = E[(\mathbf{X}_{k+1} - \hat{\mathbf{X}}_{k+1})(\mathbf{X}_{k+1} - \hat{\mathbf{X}}_{k+1})^T] \leq \mathbf{S}_{k+1} \quad (14)$$

2.1 估计误差方差

在REKF算法中,通过选取一些模型不确定参数来表示其对系统的影响。与EKF算法类似,需要对状态函数与量测函数进行泰勒展开,主要区别在于根据上界范围进行滤波器的具体设计。通过这样的计算,使得求得的精度、效果等相较于EKF算法更优良。

一步状态预测误差可写为

$$\tilde{\mathbf{X}}_{k+1|k} = \mathbf{X}_{k+1} - \hat{\mathbf{X}}_{k+1|k} = f'(\mathbf{X}_k) + g(\mathbf{X}_k, \boldsymbol{\eta}_k) + \mathbf{W}_k - f'(\hat{\mathbf{X}}_k) \quad (15)$$

扩维后的状态函数 $f'(\mathbf{X}_k)$ 通过在 $\hat{\mathbf{X}}_k$ 点处泰

勒展开,得到

$$f'(X_k) = f'(\hat{X}_k) + F_k \tilde{X}_k + o(\tilde{X}_k) \quad (16)$$

式中: $F_k = \left. \frac{\partial f'(X_k)}{\partial X_k} \right|_{X_k = \hat{X}_k}$; $o(\tilde{X}_k)$ 为经过展开后的高阶项。引入未知时变矩阵 $\beta_k \in \mathbf{R}^{n \times n}$, 满足 $\beta_k \cdot \beta_k^T = I_{n \times n}$, 具体的缩放矩阵 $B_k \in \mathbf{R}^{n \times n}$, L_k 表示已知的调节矩阵, 故高阶项可表示为^[12]

$$o(\tilde{X}_k) = B_k \beta_k L_k \tilde{X}_k \quad (17)$$

将式(16)、式(17)代入式(15)可得

$$\tilde{X}_{k+1|k} = (F_k + B_k \beta_k L_k) \tilde{X}_k + g(X_k, \eta_k) + W_k \quad (18)$$

由于 η_{ik} 服从高斯分布, 不同时刻之间是互相独立的, 其相关性为 0。而 η_{ik} 与 w_k 是互相独立的高斯白噪声, 均值为零。故一步预测误差方差矩阵可表示为

$$P_{k+1|k} = E[(F_k + B_k \beta_k L_k) \tilde{X}_k \tilde{X}_k^T (F_k + B_k \beta_k L_k)^T] + E[g(X_k, \eta_k) g^T(X_k, \eta_k)] + E(W_k W_k^T) \quad (19)$$

式中:

$$E[(F_k + B_k \beta_k L_k) \tilde{X}_k \tilde{X}_k^T (F_k + B_k \beta_k L_k)^T] = (F_k + B_k \beta_k L_k) E(\tilde{X}_k \tilde{X}_k^T) (F_k + B_k \beta_k L_k)^T \quad (20)$$

$$E[g(X_k, \eta_k) g^T(X_k, \eta_k)] = \sum_{i=1}^s \tilde{A}_{ik} E(X_k X_k^T) \tilde{A}_{ik}^T \quad (21)$$

$$\text{其中: } \tilde{A}_{ik} = \begin{bmatrix} A_{ik} & \mathbf{0} \\ \mathbf{0} & A_{ik-1} \end{bmatrix}.$$

式(19)可简化表示为

$$P_{k+1|k} = (F_k + B_k \beta_k L_k) E(\tilde{X}_k \tilde{X}_k^T) (F_k + B_k \beta_k L_k)^T + \sum_{i=1}^s \tilde{A}_{ik} E(X_k X_k^T) \tilde{A}_{ik}^T + E(W_k W_k^T) \quad (22)$$

已知 $k+1$ 时刻的状态估计误差与真实值和估计值有关, 可以表示为

$$\tilde{X}_{k+1} = \tilde{X}_{k+1|k} - K_{k+1} (y_{k+1} - \bar{\gamma}_k h'(\hat{X}_{k+1|k})) \quad (23)$$

将式(10)代入式(23)中, 可以得到

$$\tilde{X}_{k+1} = \tilde{X}_{k+1|k} - K_{k+1} [\gamma_{k+1} (h'(X_{k+1}) + s(X_{k+1}, \xi_{k+1}) + V_{k+1}) - \bar{\gamma}_k h'(\hat{X}_{k+1|k})] \quad (24)$$

同式(16)相同, 对量测函数泰勒展开, 保留高阶项, 并代入式(24)得

$$h'(X_{k+1}) = h'(\hat{X}_{k+1|k}) + H_{k+1} \tilde{X}_{k+1|k} + o(\tilde{X}_{k+1|k}) \quad (25)$$

式中: $H_{k+1} = \left. \frac{\partial h'(X_{k+1})}{\partial X_{k+1}} \right|_{X_{k+1} = \hat{X}_{k+1|k}}$; $o(\tilde{X}_{k+1|k}) = D_{k+1} \alpha_{k+1} L_{k+1} \tilde{X}_{k+1|k}$, 根据文献[13], α_{k+1} 为未知

的时变矩阵, D_{k+1} 为考虑不确定系统模型误差的比例矩阵。

将式(25)代入式(24)中, 所以有

$$\tilde{X}_{k+1} = (I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1} - K_{k+1} \bar{\gamma}_{k+1} D_{k+1} \alpha_{k+1} L_{k+1}) \tilde{X}_{k+1|k} - K_{k+1} \gamma_{k+1} s(X_{k+1}, \xi_{k+1}) - K_{k+1} \gamma_{k+1} V_{k+1} - K_{k+1} (\gamma_{k+1} - \bar{\gamma}_{k+1}) h'(X_{k+1}) \quad (26)$$

将式(26)代到式(14)中, 得到滤波误差方差矩阵, 表示为

$$P_{k+1} = (I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1} - K_{k+1} \bar{\gamma}_{k+1} D_{k+1} \alpha_{k+1} L_{k+1}) E(\tilde{X}_{k+1|k} \tilde{X}_{k+1|k}^T) (I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1} - K_{k+1} \bar{\gamma}_{k+1} D_{k+1} \alpha_{k+1} L_{k+1})^T + E[K_{k+1} \gamma_{k+1} s(X_{k+1}, \xi_{k+1}) s^T(X_{k+1}, \xi_{k+1}) \gamma_{k+1}^T K_{k+1}^T] + E(K_{k+1} \gamma_{k+1} V_{k+1} V_{k+1}^T \gamma_{k+1}^T K_{k+1}^T) + E[K_{k+1} (\gamma_{k+1} - \bar{\gamma}_{k+1}) h'(X_{k+1}) h'^T(X_{k+1}) (\gamma_{k+1} - \bar{\gamma}_{k+1})^T K_{k+1}^T] - E[(I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1} - K_{k+1} \bar{\gamma}_{k+1} D_{k+1} \alpha_{k+1} L_{k+1}) \tilde{X}_{k+1|k} h'^T(X_{k+1}) \cdot (\gamma_{k+1} - \bar{\gamma}_{k+1})^T K_{k+1}^T] - E[K_{k+1} (\gamma_{k+1} - \bar{\gamma}_{k+1}) h'(X_{k+1}) \tilde{X}_{k+1|k}^T (I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1} - K_{k+1} \bar{\gamma}_{k+1} D_{k+1} \alpha_{k+1} L_{k+1})^T] \quad (27)$$

由式(22)、式(27)可以得到方差矩阵 $P_{k+1|k}$ 、 P_{k+1} , 因为所建立的系统模型考虑了乘性噪声及量测延时时等不确定性因素, 其中包括未知矩阵如 β_k 、 α_{k+1} , 所以并不能实际求出协方差的具体值。根据最小方差准则的要求, 求取预测误差方差 $P_{k+1|k}$ 和滤波更新误差方差 P_{k+1} 的最小上界来近似替代。

2.2 算法设计

根据文献[14]中的 2 条引理(引理 6.1、引理 6.2)构造上界范围, 并且可以求得滤波更新误差方差矩阵 P_{k+1} 的最小上界。

这里引入二阶矩的概念。若对于任意的 t 属于 T , 若 $E[X^2(t)]$ 存在, 则称 $X(t)$ 为二阶矩过程。实际就是对变量的平方求期望, 所以二阶矩的值是大于等于零的。假设 ε_1 为正数, 那么 $\varepsilon_1^{\frac{1}{2}} \tilde{X}_k - \varepsilon_1^{-\frac{1}{2}} \hat{X}_k$ 的二阶矩一定大于等于零。所以有

$$\tilde{X}_k \hat{X}_k^T + \hat{X}_k \tilde{X}_k^T \leq \varepsilon_1 \tilde{X}_k \tilde{X}_k^T + \varepsilon_1^{-1} \hat{X}_k \hat{X}_k^T \quad (28)$$

根据式(28)可以得到

$$E(X_k X_k^T) \leq E[(1 + \varepsilon_1) \tilde{X}_k \tilde{X}_k^T + (1 + \varepsilon_1^{-1}) \hat{X}_k \hat{X}_k^T] \quad (29)$$

$E[g(X_k, \eta_k) g^T(X_k, \eta_k)] \leq$

$$\sum_{i=1}^s \tilde{A}_{ik} [(1 + \varepsilon_1) P_k + (1 + \varepsilon_1^{-1}) \hat{X}_k \hat{X}_k^T] \tilde{A}_{ik}^T \quad (30)$$

假设存在正数 λ 和矩阵 L_k , 满足条件 $0.5\lambda^{-1} \cdot$

$I - L_k P_k L_k^T > 0$, 由文献[15]中的引理可推导出

$$(F_k + B_k \beta_k L_k) P_k (F_k + B_k \beta_k L_k)^T \leq F_k (P_k^{-1} - 2\lambda L_k^T L_k)^{-1} F_k^T + \lambda^{-1} B_k B_k^T \quad (31)$$

将式(29)~式(31)代入到式(22)中, 可以

求出一步预测误差方差矩阵的上界:

$$P_{k+1|k} \leq F_k (P_k^{-1} - 2\lambda L_k^T L_k)^{-1} F_k^T + \lambda^{-1} B_k B_k^T + \sum_{i=1}^s \tilde{A}_{ik} [(1 + \varepsilon_1) P_k + (1 + \varepsilon_1^{-1}) \hat{X}_k \hat{X}_k^T] \tilde{A}_{ik}^T + \bar{Q}_k \quad (32)$$

$$E(s(X_{k+1}, \xi_{k+1}) s^T(X_{k+1}, \xi_{k+1})) =$$

$$\sum_{i=1}^r \tilde{C}_{ik+1} E(X_{k+1} X_{k+1}^T) \tilde{C}_{ik+1}^T \quad (33)$$

将式(33)代入式(27), 得到滤波误差方差矩

阵为

$$P_{k+1} = (I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1} - K_{k+1} \bar{\gamma}_{k+1} D_{k+1} \alpha_{k+1} L_{k+1}) \cdot P_{k+1|k} (I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1} - K_{k+1} \bar{\gamma}_{k+1} D_{k+1} \alpha_{k+1} L_{k+1})^T + K_{k+1} \gamma_{k+1} \sum_{i=1}^r \tilde{C}_{ik+1} E(X_{k+1} X_{k+1}^T) \tilde{C}_{ik+1}^T \gamma_{k+1}^T K_{k+1}^T + K_{k+1} \gamma_{k+1} \bar{R}_{k+1} \gamma_{k+1}^T K_{k+1}^T + K_{k+1} (\gamma_{k+1} - \bar{\gamma}_{k+1}) \cdot E[h'(X_{k+1}) h'^T(X_{k+1})] (\gamma_{k+1} - \bar{\gamma}_{k+1})^T K_{k+1}^T - E[(I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1} - K_{k+1} \bar{\gamma}_{k+1} D_{k+1} \alpha_{k+1} L_{k+1}) \tilde{X}_{k+1} h'^T(X_{k+1}) (\gamma_{k+1} - \bar{\gamma}_{k+1})^T K_{k+1}^T]^T - E[(K_{k+1} (\gamma_{k+1} - \bar{\gamma}_{k+1}) h'(X_{k+1}) \tilde{X}_{k+1}^T \cdot (I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1} - K_{k+1} \bar{\gamma}_{k+1} D_{k+1} \alpha_{k+1} L_{k+1})^T] \quad (34)$$

$$\text{式中: } \tilde{C}_{ik+1} = \begin{bmatrix} C_{ik+1} & \mathbf{0} \\ \mathbf{0} & C_{ik} \end{bmatrix}.$$

同理, 根据二阶矩原理, 假设存在正数 ε_2 , 同样能求解得到

$$(I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1} - K_{k+1} \bar{\gamma}_{k+1} D_{k+1} \alpha_{k+1} L_{k+1}) \cdot \tilde{X}_{k+1} h'^T(X_{k+1}) (\gamma_{k+1} - \bar{\gamma}_{k+1})^T K_{k+1}^T + K_{k+1} (\gamma_{k+1} - \bar{\gamma}_{k+1}) h'(X_{k+1}) \tilde{X}_{k+1}^T \cdot (I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1} - K_{k+1} \bar{\gamma}_{k+1} D_{k+1} \alpha_{k+1} L_{k+1})^T \leq \varepsilon_2 (I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1} - K_{k+1} \bar{\gamma}_{k+1} D_{k+1} \alpha_{k+1} L_{k+1}) \cdot \tilde{X}_{k+1} \tilde{X}_{k+1}^T (I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1} - K_{k+1} \bar{\gamma}_{k+1} D_{k+1} \alpha_{k+1} L_{k+1})^T + \varepsilon_2^{-1} K_{k+1} (\gamma_{k+1} - \bar{\gamma}_{k+1}) h'(X_{k+1}) h'^T(X_{k+1}) K_{k+1}^T \quad (35)$$

由式(15)可得到

$$X_{k+1} = \tilde{X}_{k+1|k} + \hat{X}_{k+1|k} \quad (36)$$

假设存在正数 ε_3 , 同样能求解得到

$$\tilde{X}_{k+1|k} \hat{X}_{k+1|k}^T + \hat{X}_{k+1|k} \tilde{X}_{k+1|k}^T \leq \varepsilon_3 \tilde{X}_{k+1|k} \tilde{X}_{k+1|k}^T + \varepsilon_3^{-1} \hat{X}_{k+1|k} \hat{X}_{k+1|k}^T \quad (37)$$

所以有

$$E(X_{k+1} X_{k+1}^T) \leq (1 + \varepsilon_3) P_{k+1|k} + (1 + \varepsilon_3^{-1}) \hat{X}_{k+1|k} \hat{X}_{k+1|k}^T \quad (38)$$

$$E[h'(X_{k+1}) h'^T(X_{k+1})] \leq 2[\sigma_1^2 \text{trace}((1 + \varepsilon_3) P_{k+1|k} + (1 + \varepsilon_3^{-1}) \hat{X}_{k+1|k} \hat{X}_{k+1|k}^T) + \sigma_2^2] I \quad (39)$$

假设存在正数 μ 和矩阵 L_{k+1} , 满足条件 $\mu^{-1} I -$

$L_{k+1} P_{k+1|k} L_{k+1}^T > 0$, 能够推出

$$(I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1} - K_{k+1} \bar{\gamma}_{k+1} D_{k+1} \alpha_{k+1} L_{k+1}) P_{k+1|k} \cdot (I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1} - K_{k+1} \bar{\gamma}_{k+1} D_{k+1} \alpha_{k+1} L_{k+1})^T \leq (I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1}) (P_{k+1|k} - \mu L_{k+1}^T L_{k+1})^{-1} \cdot (I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1})^T + \mu^{-1} K_{k+1} \bar{\gamma}_{k+1} D_{k+1} D_{k+1}^T \bar{\gamma}_{k+1}^T K_{k+1}^T \quad (40)$$

所以, $k+1$ 时刻的更新后状态误差的方差矩阵为

$$P_{k+1} \leq (1 + \varepsilon_2) [(I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1}) (P_{k+1|k} - \mu L_{k+1}^T L_{k+1})^{-1} (I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1})^T + \mu^{-1} K_{k+1} \bar{\gamma}_{k+1} D_{k+1} D_{k+1}^T \bar{\gamma}_{k+1}^T K_{k+1}^T] + K_{k+1} \gamma_{k+1} \sum_{i=1}^r \tilde{C}_{ik+1} [(1 + \varepsilon_3) P_{k+1|k} + (1 + \varepsilon_3^{-1}) \hat{X}_{k+1|k} \hat{X}_{k+1|k}^T] \cdot \tilde{C}_{ik+1}^T \gamma_{k+1}^T K_{k+1}^T + K_{k+1} \gamma_{k+1} \bar{R}_{k+1} \bar{\gamma}_{k+1}^T K_{k+1}^T + (1 + \varepsilon_2^{-1}) \{K_{k+1} (\gamma_{k+1} - \bar{\gamma}_{k+1}) \cdot 2[\sigma_1^2 \text{trace}((1 + \varepsilon_3) P_{k+1|k} + (1 + \varepsilon_3^{-1}) \hat{X}_{k+1|k} \hat{X}_{k+1|k}^T) + \sigma_2^2] I \cdot (\gamma_{k+1} - \bar{\gamma}_{k+1})^T K_{k+1}^T\} \quad (41)$$

根据式(32)、式(41)得到了一步预测方差

和滤波更新方差的上界, 当 $\varepsilon_1, \varepsilon_2, \varepsilon_3, \lambda, \mu$ 都是正数时, 提出以下 2 个离散 Riccati 方程:

$$\Sigma_{k+1|k} = F_k (\Sigma_k^{-1} - 2\lambda L_k^T L_k)^{-1} F_k^T + \lambda^{-1} B_k B_k^T + \sum_{i=1}^s \tilde{A}_{ik} [(1 + \varepsilon_1) \Sigma_k + (1 + \varepsilon_1^{-1}) \hat{X}_k \hat{X}_k^T] \tilde{A}_{ik}^T + \bar{Q}_k \quad (42)$$

$$\Sigma_{k+1} = (1 + \varepsilon_2) [(I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1}) (\Sigma_{k+1|k}^{-1} - \mu L_{k+1}^T L_{k+1})^{-1} (I - K_{k+1} \bar{\gamma}_{k+1} H_{k+1})^T + \mu^{-1} K_{k+1} \bar{\gamma}_{k+1} D_{k+1} D_{k+1}^T \bar{\gamma}_{k+1}^T K_{k+1}^T] + K_{k+1} \gamma_{k+1} \sum_{i=1}^r \tilde{C}_{ik+1} [(1 + \varepsilon_3) \Sigma_{k+1|k} + (1 + \varepsilon_3^{-1}) \hat{X}_{k+1|k} \hat{X}_{k+1|k}^T] \cdot \tilde{C}_{ik+1}^T \gamma_{k+1}^T K_{k+1}^T + K_{k+1} \gamma_{k+1} \bar{R}_{k+1} \bar{\gamma}_{k+1}^T K_{k+1}^T + (1 + \varepsilon_2^{-1}) \{K_{k+1} (\gamma_{k+1} - \bar{\gamma}_{k+1}) \cdot 2[\sigma_1^2 \text{trace}((1 + \varepsilon_3) \Sigma_{k+1|k} + (1 + \varepsilon_3^{-1}) \hat{X}_{k+1|k} \hat{X}_{k+1|k}^T) + \sigma_2^2] I \cdot (\gamma_{k+1} - \bar{\gamma}_{k+1})^T K_{k+1}^T\} \quad (43)$$

用数学归纳法证明 $\Sigma_{k+1|k}, \Sigma_{k+1}$ 是 $P_{k+1|k}, P_{k+1}$ 的上界范围, 即

$$\begin{cases} P_{k+1|k} \leq \Sigma_{k+1|k} & 0 \leq k \leq n \\ P_{k+1} \leq \Sigma_{k+1} \end{cases} \quad (44)$$

在初始 $k=0$ 时刻, 能够轻易求出 $P_0 = \Sigma_0 > 0$, 根据不等式(29)、式(30)以及等式(43)可得

$$P_{1|0} \leq \Sigma_{1|0} \quad (45)$$

由式(45)及式(44)中的第一项可以得出

$$(P_{1|0}^{-1} - \mu L_1^T L_1)^{-1} \leq (\Sigma_{1|0}^{-1} - \mu L_1^T L_1)^{-1} \quad (46)$$

由式(45)和式(46),可以推断出 $k=0$ 时刻的状态误差方差满足:

$$P_1 \leq \Sigma_1 \quad (47)$$

选取 $k=n-1$ 时刻,假设满足不等式 $P_n \leq \Sigma_n$,那么当 $k=n$ 时,同式(46)类似,有

$$(P_n^{-1} - 2\lambda L_n^T L_n)^{-1} \leq (\Sigma_n^{-1} - 2\lambda L_n^T L_n)^{-1} \quad (48)$$

因此,根据式(44)可以得出

$$\begin{cases} P_{n+1|n} \leq \Sigma_{n+1|n} \\ P_{n+1} \leq \Sigma_{n+1} \end{cases} \quad (49)$$

根据构造的 $P_{k+1|k}$ 和 P_{k+1} 的最小上界确定最优的滤波增益。对式(43) K_{k+1} 项求偏导 $\frac{\partial \Sigma_{k+1}}{\partial K_{k+1}} = 0$ 确定滤波增益的最优值:

$$\begin{aligned} \frac{\partial \Sigma_{k+1}}{\partial K_{k+1}} &= 2(1 + \varepsilon_2) [(I - K_{k+1} - \bar{\gamma}_{k+1} H_{k+1}) \cdot \\ &(\Sigma_{k+1|k}^{-1} - \mu L_{k+1}^T L_{k+1})^{-1} (-\bar{\gamma}_{k+1} H_{k+1})^T + \\ &K_{k+1} (\mu^{-1} \bar{\gamma}_{k+1} D_{k+1} D_{k+1}^T \bar{\gamma}_{k+1}^T)] + \\ &2K_{k+1} \left\{ [\gamma_{k+1} \sum_{i=1}^r \tilde{C}_{ik+1} [(1 + \varepsilon_3) \Sigma_{k+1|k} + \right. \\ &(1 + \varepsilon_3^{-1}) \hat{X}_{k+1|k} \hat{X}_{k+1|k}^T] \tilde{C}_{ik+1}^T \gamma_{k+1}^T] + \\ &\gamma_{k+1} \bar{R}_{k+1} \gamma_{k+1}^T + 2(1 + \varepsilon_2^{-1}) (\gamma_{k+1} - \bar{\gamma}_{k+1}) \cdot \\ &[\sigma_1^2 \text{trace}((1 + \varepsilon_3) \Sigma_{k+1|k} + (1 + \\ &\varepsilon_3^{-1}) \hat{X}_{k+1|k} \hat{X}_{k+1|k}^T) + \sigma_2^2] I (\gamma_{k+1} - \bar{\gamma}_{k+1})^T \left. \right\} \quad (50) \end{aligned}$$

$$\begin{aligned} K_{k+1} &= (1 + \varepsilon_2) (\Sigma_{k+1|k}^{-1} - \\ &\mu L_{k+1}^T L_{k+1})^{-1} (\bar{\gamma}_{k+1} H_{k+1})^T \{ (1 + \\ &\varepsilon_2) \bar{\gamma}_{k+1} H_{k+1} (\Sigma_{k+1|k}^{-1} - \\ &\mu L_{k+1}^T L_{k+1})^{-1} (\bar{\gamma}_{k+1} H_{k+1})^T + \\ &(1 + \varepsilon_2) \mu^{-1} \bar{\gamma}_{k+1} D_{k+1} D_{k+1}^T \bar{\gamma}_{k+1}^T + \\ &\bar{\gamma}_{k+1} \sum_{i=1}^r \tilde{C}_{ik+1} [(1 + \varepsilon_3) \Sigma_{k+1|k} + \\ &(1 + \varepsilon_3^{-1}) \hat{X}_{k+1|k} \hat{X}_{k+1|k}^T] \tilde{C}_{ik+1}^T \bar{\gamma}_{k+1}^T + \\ &\gamma_{k+1} \bar{R}_{k+1} \gamma_{k+1}^T + 2(1 + \varepsilon_2^{-1}) (\gamma_{k+1} - \\ &\bar{\gamma}_{k+1}) [\sigma_1^2 \text{trace}((1 + \varepsilon_3) \Sigma_{k+1|k} + \\ &(1 + \varepsilon_3^{-1}) \hat{X}_{k+1|k} \hat{X}_{k+1|k}^T) + \sigma_2^2] \cdot \\ &I (\gamma_{k+1} - \bar{\gamma}_{k+1})^T \left. \right\}^{-1} \quad (51) \end{aligned}$$

设计滤波器的目的是根据状态模型和不同观测量中的信息计算状态变量的估计值,达到状态量的估计值与真实值误差最小,精度最高;在滤波设计过程中,需要权衡多种不同信息,充分地利用这些信息。REKF 的设计是基于不确定的系统模型进行的,REKF 算法的性能在于不确定性模型是否能完整描述实际系统。为了获得理想的滤波效果,式(1)给出的不确定性模型需要准确给出不确定性参数的大小。

3 仿真实验与分析

3.1 仿真环境

为了验证 REKF 算法的有效性,本节仿真利用传统的加性扩展卡尔曼滤波(AEKF)算法、RKF 算法^[6]、RFHF 算法^[8]与本文所提的 REKF 算法进行比较。实际的星敏传感器精度较高,可以把泰勒展开得到的高阶项忽略, B_k 、 D_{k+1} 设置为 0。调节矩阵 $L_k = \sqrt{\Sigma_k^{-1}}$ 。为了保证系统能达到的精度,参数 λ 和 μ 均设置为 0.0001,参数 ε 设置为 0.1。通过状态模型及量测模型可知,其中包括加性噪声,考虑噪声及延时对系统的影响,可令矩阵 $C_{ik} = 0$ 。仿真时间设为 1 h,量测模型参数 $\sigma_{ij} = 10 \text{ s} (i=1,2,3; j=x,y,z)$ 。

3.2 仿真分析

1) 情况 1

不考虑系统存在延时情况并且假设姿态角及陀螺漂移部分均没有误差存在。假设卫星以角速率 $\omega(t) = [4\pi/3 \ 0 \ 0]^T \text{ rad/s}$ 旋转,初始三轴姿态角设为 $[0 \ 0 \ 0]^T (^\circ)$,卫星轨道角速度 $\Omega_0 = 0.001 \text{ rad/s}$ 。常值漂移设为 $0.1 (^\circ)/\text{h}$,漂移噪声设为 $\sigma_u = \sqrt{10} \times 10^{-10} \text{ rad/s}^{3/2}$,测量噪声设为 $\sigma_v = \sqrt{10} \times 10^{-7} \text{ rad/s}^{1/2}$ 。星敏感器的测量噪声设为 $\sigma_s = 0.005^\circ$ 。初始状态值为 $x_0 = [0 \ 0 \ 0 \ 1 \ 0.1 \ 0.1 \ 0.1]^T$,对应的初始四元数部分为 $q_0 = [0 \ 0 \ 0 \ 1]^T$,初始陀螺漂移部分为 $\beta_0 = [0.1 \ 0.1 \ 0.1]^T (^\circ)/\text{h}$,星敏感器的 3 个参考矢量设为 $r_1 = [1 \ 0 \ 0]^T$, $r_2 = [0 \ 1 \ 0]^T$, $r_3 = [0 \ 0 \ 1]^T$ 。星敏传感器与陀螺的采样频率为 1 Hz。初始的 3 个姿态角误差方差矩阵表示为 $0.01 I_{3 \times 3} (^\circ)^2$,陀螺漂移的误差方差矩阵表示为 $0.04 I_{3 \times 3} ((^\circ)/\text{h})^2$ 。

从图 1 和图 2 可以看出,当假设系统不存在延时情况时,各滤波算法的性能相差不大,基本都能达到 0.0001° 左右的精度。整体来看,REKF 算法和 RKF 算法要略优于 AEKF 算法和 RFHF 算法,这是由于 AEKF 算法和 RFHF 算法在设计时都没有考虑到乘性噪声的影响,导致它们在非线性的估计时存在一定误差,因此滤波精度稍差。

2) 情况 2

假设星敏传感器工作时出现延时现象,设 3 个方向的矢量延时速率各不相同的,并且满足伯努利分布: $\bar{T}_k = \text{diag} \{ [0.2 \ 0.2 \ 0.2 \ 0.1 \ 0.1 \ 0.1 \ 0.05 \ 0.05 \ 0.05] \}$

从图 3 和图 4 可以看出,REKF 算法处理带有

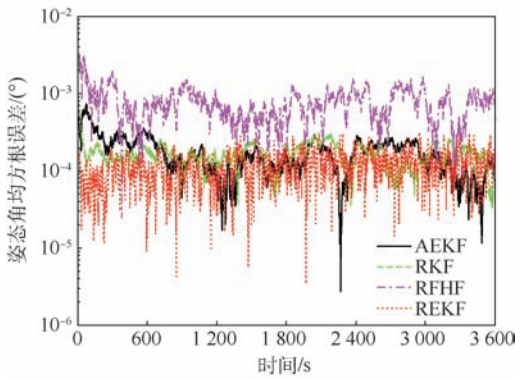


图1 情况1时姿态角均方根误差对比

Fig.1 Comparison of root mean square error of attitude angle in Case 1

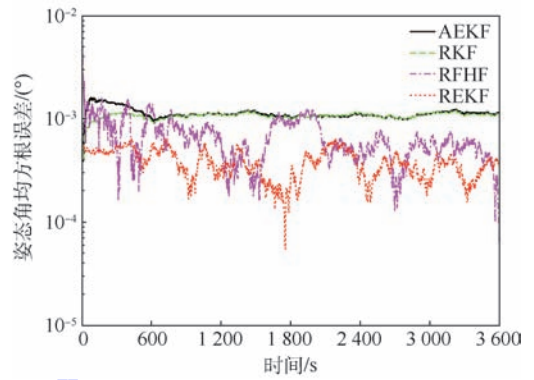


图3 情况2时姿态角均方根误差对比

Fig.3 Comparison of RMSE of attitude angle in Case 2

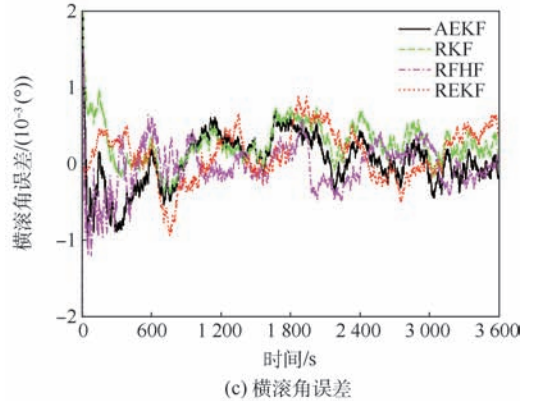
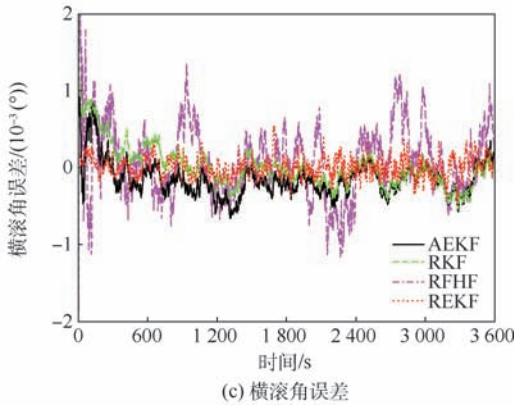
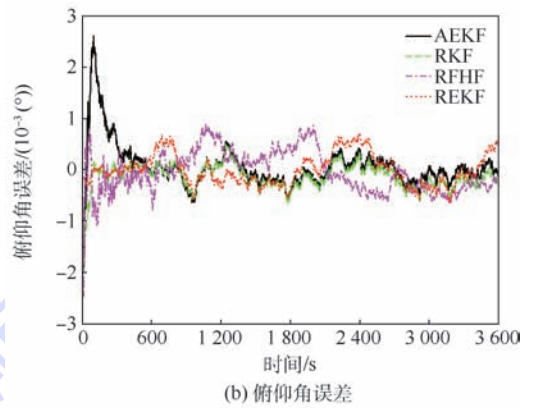
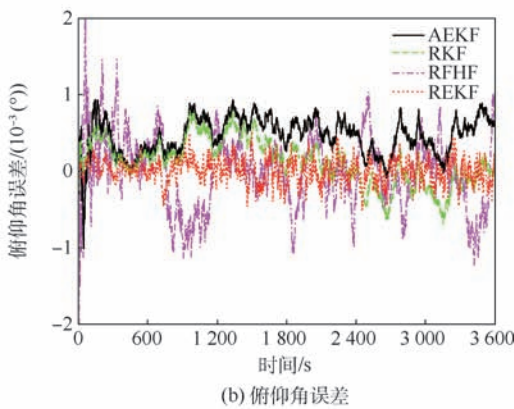
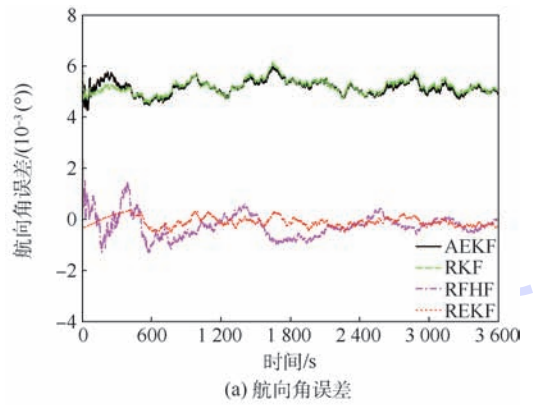
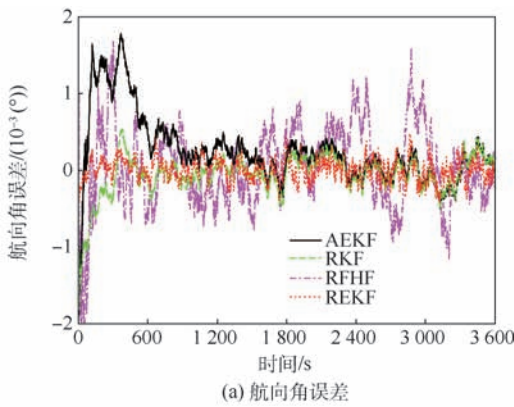


图2 情况1时姿态角误差对比

Fig.2 Comparison of attitude angle error in Case 1

图4 情况2时姿态角误差对比

Fig.4 Comparison of attitude angle error in case 2

延时及乘性噪声的非线性系统的效果要明显优于 RFHF 算法、RKF 算法及 AEKF 算法,这是因为建立了带有乘性噪声和延时的误差模型来表示这种情况,而 AEKF 算法并不适用带有延时的系统,RKF 算法只能满足乘性噪声项所带来的干扰,但没有考虑延时问题。所以,当出现延时, RKF 算法并不能保证系统精度,甚至会带来滤波发散。而 RFHF 算法只考虑了量测延时和量测丢失,没有考虑乘性噪声的影响,因此,在具有乘性噪声的实验环境下不能达到最理想的状态。

4 结 论

针对小视场星敏感器姿态估计时存在的信息延时问题,本文做了如下研究:

- 1) 建立带有延时不确定项的误差模型,该模型考虑到非线性系统同时存在乘性噪声及量测模型延时的情况,对 REKF 滤波算法进行改进。
- 2) 在算法设计时,根据最小均方误差准则要求,通过求取预测误差方差和滤波更新误差方差的最小上界进而确定滤波增益的最优值。REKF 算法的性能在于不确定性模型来表示实际系统,从而达到状态估计误差最小,精度最高。
- 3) 仿真结果表明,REKF 滤波算法可以有效解决量测延时问题,提高姿态估计的精度。

参考文献 (References)

- [1] XIONG K, WEI C L, LIU L D. Robust extended Kalman filtering for nonlinear systems with stochastic uncertainties [J]. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 2010, 40(2):399-405.
- [2] XIONG K, WEI C L, LIU L D. Robust Kalman filtering for discrete-time nonlinear systems with parameter uncertainties [J]. Aerospace Science & Technology, 2012, 18(1):15-24.
- [3] XIONG K, LIU L D, LIU Y. Robust extended Kalman filtering for nonlinear systems with multiplicative noises [J]. Optimal Control Applications & Methods, 2011, 32(1):47-63.
- [4] REZAEI H, ESFANJANI R M, FARSI M. Robust filtering for uncertain networked systems with randomly delayed and lost measurements [J]. IET Signal Processing, 2015, 9(4):320-327.
- [5] INOUE R S, TERRA M H, CERRI J P. Extended robust Kalman filter for attitude estimation [J]. IET Control Theory & Applications, 2016, 10(2):162-172.
- [6] DONG Z, YOU Z. Finite-horizon robust Kalman filtering for un-

certain discrete time-varying systems with uncertain-covariance white noises [J]. IEEE Signal Processing Letters, 2006, 13(8):493-496.

- [7] ZHENG J H, LIU J F. A robust finite-horizon Kalman filter for uncertain discrete time-varying systems with state-delay and missing measurements [J]. International Journal of Grid and Distributed Computing, 2016, 9(3):229-242.
- [8] WANG F, WANG Z D, LIANG J L, et al. Robust finite-horizon filtering for 2-D systems with randomly varying sensor delays [J]. IEEE Transactions on Systems, Man, and Cybernetics: Ssystems, 2018:1-13.
- [9] QIN W T, WANG X G, BAI Y L, et al. Arbitrary-step randomly delayed robust filter with application to boost phase tracking [J]. Acta Astronautica, 2018, 145:304-318.
- [10] FAN Z, YANG J. A research of gyro/star-sensor integrated attitude determination based on particle filter [C] // Third International Conference on Instrumentation, Measurement, Computer, Communication and Control. Piscataway, NJ: IEEE Press, 2013:256-261.
- [11] REIF K, GÜNTHER S, YAZ E, et al. Stochastic stability of the discrete-time extended Kalman filter [J]. IEEE Transactions on Automatic Control, 1999, 44(4):714-728.
- [12] XIONG K, LIU L D, ZHANG H Y. Modified unscented Kalman filtering and its application in autonomous satellite navigation [J]. Aerospace Science & Technology, 2009, 13(4):238-246.
- [13] WANG S, FANG H, TIAN X. Recursive estimation for nonlinear stochastic systems with multi-step transmission delays, multiple packet dropouts and correlated noises [J]. Signal Processing, 2015, 115:164-175.
- [14] 黄蔚. CKF 及鲁棒滤波在飞行器姿态估计中的应用研究 [D]. 哈尔滨:哈尔滨工程大学, 2015.
- [15] HUANG W. Application of CKF and robust filter in aircraft attitude estimation [D]. Harbin: Harbin Engineering University, 2015 (in Chinese).
- [15] XIE L, SOH Y C, DE SOUZA C E. Robust Kalman filtering for uncertain discrete-time systems [J]. IEEE Transactions on Automatic Control, 1994, 39(6):1310-1314.

作者简介:

钱华明 男,博士,教授,博士生导师。主要研究方向:组合导航、星敏感器、信息处理、传感器技术与智能系统技术。

王迪 女,硕士研究生。主要研究方向:飞行器姿态算法。

吴永慧 女,博士研究生。主要研究方向:姿态估计、信息融合。

黄智开 男,博士研究生。主要研究方向:数字信号处理。

Filtering algorithm of NFOV star sensor measurement delay

QIAN Huaming*, WANG Di, WU Yonghui, HUANG Zhikai

(College of Automation, Harbin Engineering University, Harbin 150001, China)

Abstract: Aimed at measurement delay in the narrow field of view (NFOV) star sensor used for attitude estimation, a robust extended Kalman filter (REKF) algorithm is proposed to solve the measurement delay. According to the minimum mean square error criterion, the minimum upper bound of the variance is solved and the filter gain is determined by the minimum upper bound. The designed REKF algorithm can effectively solve the problem of measurement delay and improve the accuracy of attitude estimation. Finally, the simulation results show that the algorithm is superior to the conventional additive robust extended Kalman filter (AEKF), robust finite-horizon filter (RFHF) and robust Kalman filter (RKF) algorithm, which can better solve the problem of measurement delay in nonlinear systems, and the effectiveness of the algorithm is verified.

Keywords: narrow field of view (NFOV) star sensor; attitude estimation; extended Kalman filter (EKF); robust filtering; measurement delay

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

DOI: 10.13700/j.bh.1001-5965.2018.0286

航天刚-弹-液耦合系统的弹-液耦合研究



梁立孚*, 郭庆勇

(哈尔滨工程大学 航天与建筑工程学院, 哈尔滨 150001)

摘 要: 航天刚-弹-液耦合动力学中的弹-液耦合问题是关系到航天刚-弹-液耦合系统液固耦合机理研究和大规模液固耦合建模计算研究的重要课题。通过分析刚-弹-液耦合动力学拟变分原理的泛函,说明刚-弹-液耦合中的刚-弹耦合、刚-液耦合和弹-液耦合的特点。应用 Lagrange 乘子法,将无际边界条件纳入泛函中,通过识别 Lagrange 乘子,逐步说明弹-液耦合的机理。通过对泛函求驻值发现,无际边界条件由原来的先决条件转化为驻值条件,实现了从协调元到杂交元的过渡。通过分析泛函识别 Lagrange 乘子前后的驻值条件,明确了识别 Lagrange 乘子可以有效地减少计算自由度。

关键词: 航天充液系统; 刚-弹-液耦合动力学; 弹-液耦合机理; 液固耦合建模; 无际边界条件

中图分类号: V414

文献标识码: A **文章编号:** 1001-5965(2019)02-0243-09

随着航天事业的发展,特别是大型空间站的建立、空间实验室的出现以及探讨人类长期在宇宙空间居住或者旅行的研究工作的开展,较全面地研究刚-弹-液耦合动力学的任务已经提上了日程^[1-2]。

作为一个航天大国,中国学者在航天充液系统的刚-弹-液耦合机理研究和大规模刚-弹-液耦合模型建模计算的应用研究方面,已经进行了一些有益的探索,并且取得了丰硕的研究成果。杜建镔和王勛成^[3]针对具有旋转周期性的含液容器,将旋转周期方法引入液固耦合动力特性分析,开发出一套有效的液固耦合有限元分析方法。刘习军等^[4]研究了壳液耦合系统的自激振动问题和重力波现象。马兴瑞等^[5]将流固耦合问题作为一个专题进行了研究。邢景棠、周盛和崔尔杰^[6]全面综述了流固耦合力学及其特点、研究分

支、一些进展及进一步发展,并重点阐述了国外学者在这一学科领域的研究成果。王勛成^[7]研究了流固耦合的有限单元法,采用一种简化的方式处理弹-液耦合问题。李磊等^[8]研究了运动圆筒内液体流动不稳定的非线性特征。席如青和曲广吉^[9]研究了充液航天器任意旋转对称偏置贮箱在低微重力情况下的小幅液体晃动问题,并编制了具有工程实用性和通用性的计算机分析软件。李英波和冯正进^[10]建立了三轴定向充液卫星在二维平动和二维摆动扰动作用下的空间摆等效力学模型。王佳栋等^[11]研究了带有环形刚性隔板的部分充液刚性圆柱罐中液体的微幅晃动特性。黄华等^[12]建立了三维质心面等效模型,将贮箱内液体等效为位于液体质心处的质点。贺元军等^[13]建立了在微重力环境、横向激励下圆柱贮箱液固耦合系统的动力学方程,并得到了耦合系统

收稿日期: 2018-05-21; 录用日期: 2018-10-19; 网络出版时间: 2018-11-08 09:44

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20181106.1029.001.html

基金项目: 国家自然科学基金(10272034)

*通信作者: E-mail: lianglifu@hrbeu.edu.cn

引用格式: 梁立孚, 郭庆勇. 航天刚-弹-液耦合系统的弹-液耦合研究[J]. 北京航空航天大学学报, 2019, 45(2): 243-251.

LIANG L F, GUO Q Y. Elastic-liquid coupling in aerospace rigid-elastic-liquid coupling system [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 243-251 (in Chinese).

的幅频特性曲线。包光伟和王政伟^[14]通过建立液体晃动特征问题的泛函极值原理,计算了平放圆柱腔内三维液体晃动的特征频率,并与解析解、实验结果和二维有限元数值解进行了比较。陈建平等^[15]提出了刚-弹-液耦合系统的液体-多体耦合力学模型。高索文等^[16]分析了俯仰激励下矩形贮箱类液固耦合系统动力学特性。苟兴宇等^[17]研究了弹簧-质量系统与圆柱贮箱类液体有限幅晃动系统间的非线性耦合动力学问题,在建立了六自由度非线性耦合动力学模型的基础上,导出了液体有限幅晃动力和力矩解析表达式,说明在终了构形上积分及压力表达式中的非线性项是有限幅晃动作用力、作用力矩非线性的根源。李青等^[18]分别综述了国内外学者在充液航天器液体晃动和液固耦合动力学方面的研究进展,并重点阐述了国外学者在这一学科领域的研究成果。

在刚-弹-液耦合动力学中,存在刚-弹耦合、刚-液耦合、弹-液耦合,它们是互相联系和相互影响的。笔者分别研究了非线性刚-弹耦合动力学^[19]和刚-液耦合动力学^[20],为研究刚-弹-液耦合动力学中存在的刚-弹耦合、刚-液耦合提供了重要的参考。在此基础上,还研究了刚-弹-液耦合动力学^[21],但是没有专门研究弹-液耦合问题。所以,本文重点研究刚-弹-液耦合动力学中的弹-液耦合问题,这是一个关系到航天充液系统液固耦合机理研究和大规模液固耦合模型建模计算研究方面的重要问题。

1 刚-弹-液耦合中的刚-弹耦合、刚-液耦合和弹-液耦合特点

设有如图1所示的刚-弹-液耦合系统。文献[21]给出了刚-弹-液耦合动力学一类变量的Hamilton型的拟变分原理,这里给出一类变量的Hamilton型的拟变分原理改进形式为(用实体张量符号书写)

$$\delta\pi_{\text{req}} - \delta Q_{\text{req}} = 0 \quad (1)$$

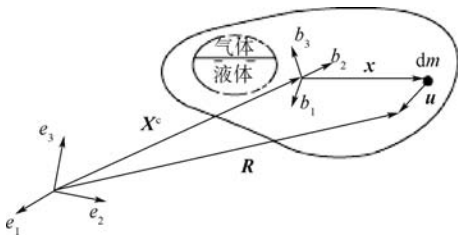


图1 参考标架

Fig. 1 Frame of reference

式中:

$$\begin{aligned} \pi_{\text{req}} = & \int_{t_0}^{t_1} \left[\iiint_{V^e} \rho^e \frac{1}{2} \cdot \frac{d\mathbf{X}^e}{dt} \cdot \frac{d\mathbf{X}^e}{dt} dV + \iiint_{V^q} \rho^q \frac{1}{2} \cdot \frac{d\mathbf{X}^q}{dt} \cdot \frac{d\mathbf{X}^q}{dt} dV + \frac{1}{2} \cdot \frac{d\boldsymbol{\theta}}{dt} \cdot \mathbf{J} \cdot \frac{d\boldsymbol{\theta}}{dt} + \pi_e + \right. \\ & \iiint_{V^e} \rho^e \left(\frac{d\mathbf{X}^e}{dt} + \frac{d\boldsymbol{\theta}}{dt} \times \mathbf{x}^e \right) \cdot \frac{d\mathbf{u}^e}{dt} dV + \\ & \left. \iiint_{V^q} \rho^q \left(\frac{d\mathbf{X}^q}{dt} + \frac{d\boldsymbol{\theta}}{dt} \times \mathbf{x}^q \right) \cdot \frac{d\mathbf{u}^q}{dt} dV + \right. \\ & \left. \pi_q + \mathbf{F} \cdot \mathbf{X}^e + \mathbf{M} \cdot \boldsymbol{\theta} \right] dt \quad (2) \end{aligned}$$

$$\begin{aligned} \pi_e = & \iiint_{V^e} \left[\frac{1}{2} \rho^e \frac{d\mathbf{u}^e}{dt} \cdot \frac{d\mathbf{u}^e}{dt} - A \left(\frac{1}{2} \nabla \mathbf{u}^e + \frac{1}{2} \mathbf{u}^e \nabla + \right. \right. \\ & \left. \left. \frac{1}{2} \nabla \mathbf{u}^e \cdot \mathbf{u}^e \nabla \right) + \mathbf{f}^e \cdot \mathbf{u}^e \right] dV + \iint_{S_\sigma} \mathbf{T}^e \cdot \mathbf{u}^e dS \quad (3) \end{aligned}$$

$$\begin{aligned} \pi_q = & \iiint_{V^q} \left[\frac{1}{2} \rho^q \frac{d\mathbf{u}^q}{dt} \cdot \frac{d\mathbf{u}^q}{dt} + p\mathbf{I} : \nabla \mathbf{u}^q - \mu \left(\nabla \frac{d\mathbf{u}^q}{dt} + \right. \right. \\ & \left. \left. \frac{d\mathbf{u}^q}{dt} \nabla \right) : \nabla \mathbf{u}^q + \mathbf{f}^q \cdot \mathbf{u}^q \right] dV + \iint_{S_f} \mathbf{T}^q \cdot \mathbf{u}^q dS \quad (4) \end{aligned}$$

$$\begin{aligned} \delta Q_{\text{req}} = & \int_{t_0}^{t_1} \left[(\mathbf{X}^e \cdot \delta \mathbf{F} + \boldsymbol{\theta} \cdot \delta \mathbf{M}) + \right. \\ & \left. \iiint_{V^e} \mathbf{u}^e \cdot \delta \mathbf{f}^e dV + \iint_{S_\sigma} \mathbf{u}^e \cdot \delta \mathbf{T}^e dS \right] dt - \\ & \int_{t_0}^{t_1} \left\{ \iiint_{V^q} \left[\nabla \mathbf{u}^q : \delta \mu \left(\nabla \frac{d\mathbf{u}^q}{dt} + \frac{d\mathbf{u}^q}{dt} \nabla \right) - \right. \right. \\ & \left. \left. \mathbf{u}^q \cdot \delta \mathbf{f}^q \right] dV - \iint_{S_f} \mathbf{u}^q \cdot \delta \mathbf{T}^q dS \right\} dt \quad (5) \end{aligned}$$

先决条件为

$$\mathbf{u}^e - \bar{\mathbf{u}}^e = \mathbf{0} \quad (\text{在 } S_u \text{ 上}) \quad (6)$$

式中: ρ^e 为弹性体的质量密度; ρ^q 为液体的质量密度; \mathbf{u}^q 为流体位移; \mathbf{f}^q 为流体所受的体积力; μ 为黏性系数; \mathbf{I} 为单位张量; p 为流体的压强; \mathbf{T}^q 为流体的面积力; ∇ 为梯度算子; S_f 为力学边界面; \mathbf{f}^e 为弹性体所受的体积力; \mathbf{n}^e 为外法线; \mathbf{T}^e 为弹性体的面积力; S_σ 为应力边界面; S_u 为位移边界面; \mathbf{X}^e 为质心矢径; $\boldsymbol{\theta}$ 为角位移; \mathbf{F} 为外力主矢; \mathbf{M} 为外力主矩; V^e 为弹性体的体积; V^q 为液体的体积,认为气体的体积很小,可以忽略, V 为系统的总体积; A 为弹性体的应变能函数; \mathbf{J} 为对质心的转动惯量; \mathbf{x}^e 为质心到弹性体中任意一点的矢径; \mathbf{x}^q 为质心到液体中任意一点的矢径。另外, \mathbf{n}^q 为单位外法线, \mathbf{u}^e 为弹性体位移, $\bar{\mathbf{u}}^e$ 为弹性体的边界位移。

$\delta\pi_{\text{req}}$ 为刚-弹-液耦合动力学一类变量的拟变分原理的定积分形式泛函的变分; π_e 为 π_{req} 的一个组成部分 ($\int_{t_0}^{t_1} \pi_e dt$ 形式上与弹性动力学的 Hamilton 原理的泛函相同); π_q 为 π_{req} 的一个组成部分 ($\int_{t_0}^{t_1} \pi_q dt$ 形式上与流体力学的 Hamilton 原理的泛函相同)。由此可见,刚-弹-液耦合动力学一类变量的 Hamilton 型拟变分原理的定积分形式泛函 π_{req} 不是弹性动力学的 Hamilton 原理的泛函 $\int_{t_0}^{t_1} \pi_e dt$ 、流体动力学的 Hamilton 原理的泛函 $\int_{t_0}^{t_1} \pi_q dt$ 和刚体动力学的 Hamilton 原理的泛函 $\int_{t_0}^{t_1} \left(\iiint_V \rho \frac{1}{2} \cdot \frac{d\mathbf{X}^c}{dt} \cdot \frac{d\mathbf{X}^c}{dt} dV + \frac{1}{2} \cdot \frac{d\boldsymbol{\theta}}{dt} \cdot \mathbf{J} \cdot \frac{d\boldsymbol{\theta}}{dt} + \mathbf{F} \cdot \mathbf{X}^c + \mathbf{M} \cdot \boldsymbol{\theta} \right) dt$ [22] 的组合,而是多出刚-弹耦合项 $\int_{t_0}^{t_1} \left[\iiint_{V^e} \rho^e \left(\frac{d\mathbf{X}^c}{dt} + \frac{d\boldsymbol{\theta}}{dt} \times \mathbf{x}^e \right) \cdot \frac{d\mathbf{u}^e}{dt} dV \right] dt$ 和刚-液耦合项 $\int_{t_0}^{t_1} \left[\iiint_{V^q} \rho^q \left(\frac{d\mathbf{X}^c}{dt} + \frac{d\boldsymbol{\theta}}{dt} \times \mathbf{x}^q \right) \cdot \frac{d\mathbf{u}^q}{dt} dV \right] dt$ 。这便是刚-弹-液耦合中的刚-弹耦合、刚-液耦合的特点,即刚-弹耦合和刚-液耦合是惯性耦合。

该变分原理的泛函隐含弹-液耦合,这是因为弹-液耦合是在弹-液交界面处出现的 [6],在弹-液交界面处,按照一般的力学原理,位移满足协调关系,应力满足平衡关系,这便是刚-弹-液耦合中的弹-液耦合的特点。因为这一问题的重要性,将在第 2 节结合研究弹-液耦合的机理,做进一步的说明。

2 刚-弹-液耦合动力学中弹-液耦合的机理

对于刚-弹-液耦合系统而言,弹-液交界面是整个系统内部出现的。假想将弹-液耦合体划分为 2 个元素,其弹-液交界面 S_w 是无际边界 (inter-element boundary) [23-24],无际边界条件为

$$\mathbf{u}^e - \mathbf{u}^q = \mathbf{0} \quad (\text{在 } S_w \text{ 上}) \quad (7)$$

应用钱伟长先生倡导的 Lagrange 乘子法 [23,25] 来处理无际边界条件的问题。为此,将式 (1) 写成展开形式,引入 Lagrange 乘子 $\boldsymbol{\lambda}$,将无际边界条件式 (7) 纳入泛函中,可得

$$\begin{aligned} \delta\pi_{\text{req}} - \delta Q_{\text{req}} = & \delta \int_{t_0}^{t_1} \left\{ \iiint_{V^e} \rho^e \frac{1}{2} \cdot \frac{d\mathbf{X}^c}{dt} \cdot \frac{d\mathbf{X}^c}{dt} dV + \right. \\ & \iiint_{V^q} \rho^q \frac{1}{2} \cdot \frac{d\mathbf{X}^c}{dt} \cdot \frac{d\mathbf{X}^c}{dt} dV + \frac{1}{2} \cdot \frac{d\boldsymbol{\theta}}{dt} \cdot \mathbf{J} \cdot \frac{d\boldsymbol{\theta}}{dt} + \\ & \mathbf{F} \cdot \mathbf{X}^c + \mathbf{M} \cdot \boldsymbol{\theta} + \iiint_{V^e} \rho^e \left(\frac{d\mathbf{X}^c}{dt} + \frac{d\boldsymbol{\theta}}{dt} \times \mathbf{x}^e \right) \cdot \\ & \frac{d\mathbf{u}^e}{dt} dV + \iiint_{V^q} \left[\frac{1}{2} \rho^e \frac{d\mathbf{u}^e}{dt} \cdot \frac{d\mathbf{u}^e}{dt} - \right. \\ & A \left(\frac{1}{2} \nabla \mathbf{u}^e + \frac{1}{2} \mathbf{u}^e \nabla + \frac{1}{2} \nabla \mathbf{u}^e \cdot \mathbf{u}^e \nabla \right) + \\ & \left. \mathbf{f}^e \cdot \mathbf{u}^e \right] dV + \iint_{S_\sigma} \mathbf{T}^e \cdot \mathbf{u}^e dS + \iint_{S_w} \boldsymbol{\lambda} \cdot (\mathbf{u}^e - \\ & \mathbf{u}^q) dS + \iiint_{V^q} \rho^q \left(\frac{d\mathbf{X}^c}{dt} + \frac{d\boldsymbol{\theta}}{dt} \times \mathbf{x}^q \right) \cdot \frac{d\mathbf{u}^q}{dt} dV + \\ & \iiint_{V^q} \left[\frac{1}{2} \rho^q \frac{d\mathbf{u}^q}{dt} \cdot \frac{d\mathbf{u}^q}{dt} + p \mathbf{I} : \nabla \mathbf{u}^q - \right. \\ & \left. \mu \left(\nabla \frac{d\mathbf{u}^q}{dt} + \frac{d\mathbf{u}^q}{dt} \nabla \right) : \nabla \mathbf{u}^q + \mathbf{f}^q \cdot \mathbf{u}^q \right] dV + \\ & \left. \iint_{S_\tau} \mathbf{T}^q \cdot \mathbf{u}^q dS - (\mathbf{X}^c \cdot \delta \mathbf{F} + \boldsymbol{\theta} \cdot \delta \mathbf{M}) \right\} dt - \\ & \int_{t_0}^{t_1} \left(\iiint_{V^e} \mathbf{u}^e \cdot \delta \mathbf{f}^e dV + \iint_{S_\sigma} \mathbf{u}^e \cdot \delta \mathbf{T}^e dS \right) dt + \\ & \int_{t_0}^{t_1} \left\{ \iiint_{V^q} \left[\nabla \mathbf{u}^q : \delta \mu \left(\nabla \frac{d\mathbf{u}^q}{dt} + \frac{d\mathbf{u}^q}{dt} \nabla \right) - \right. \right. \\ & \left. \left. \mathbf{u}^q \cdot \delta \mathbf{f}^q \right] dV - \iint_{S_\tau} \mathbf{u}^q \cdot \delta \mathbf{T}^q dS \right\} dt = 0 \quad (8) \end{aligned}$$

其先决条件为式 (6)。

进行分部积分,可得

$$\int_{t_0}^{t_1} (\rho^e + \rho^q) \frac{d\mathbf{X}^c}{dt} \cdot \delta \frac{d\mathbf{X}^c}{dt} dt = (\rho^e + \rho^q) \frac{d\mathbf{X}^c}{dt} \cdot \delta \mathbf{X}^c \Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} (\rho^e + \rho^q) \frac{d^2 \mathbf{X}^c}{dt^2} \cdot \delta \mathbf{X}^c dt \quad (9)$$

$$\int_{t_0}^{t_1} \rho^e \frac{d\mathbf{u}^e}{dt} \cdot \delta \frac{d\mathbf{X}^c}{dt} dt = \rho^e \frac{d\mathbf{u}^e}{dt} \cdot \delta \mathbf{X}^c \Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} \rho^e \frac{d^2 \mathbf{u}^e}{dt^2} \cdot \delta \mathbf{X}^c dt \quad (10)$$

$$\int_{t_0}^{t_1} \rho^e \frac{d\mathbf{X}^c}{dt} \cdot \delta \frac{d\mathbf{u}^e}{dt} dt = \rho^e \frac{d\mathbf{X}^c}{dt} \cdot \delta \mathbf{u}^e \Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} \rho^e \frac{d^2 \mathbf{X}^c}{dt^2} \cdot \delta \mathbf{u}^e dt \quad (11)$$

$$\int_{t_0}^{t_1} \rho^e \left(\frac{d\boldsymbol{\theta}}{dt} \times \mathbf{x}^e \right) \cdot \delta \frac{d\mathbf{u}^e}{dt} dt = \rho^e \left(\frac{d\boldsymbol{\theta}}{dt} \times \mathbf{x}^e \right) \cdot \delta \mathbf{u}^e \Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} \rho^e \frac{d}{dt} \left(\frac{d\boldsymbol{\theta}}{dt} \times \mathbf{x}^e \right) \cdot \delta \mathbf{u}^e dt \quad (12)$$

$$\int_{t_0}^{t_1} \rho^e \frac{d\mathbf{u}^e}{dt} \cdot \delta \frac{d\mathbf{u}^e}{dt} dt = \rho^e \frac{d\mathbf{u}^e}{dt} \cdot \delta \mathbf{u}^e \Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} \rho^e \frac{d^2 \mathbf{u}^e}{dt^2} \cdot \delta \mathbf{u}^e dt \quad (13)$$

$$-\int_{t_0}^{t_1} \rho^e \left(\frac{d\mathbf{u}^e}{dt} \times \mathbf{x}^e \right) \cdot \delta \frac{d\boldsymbol{\theta}}{dt} dt = -\rho^e \left(\frac{d\mathbf{u}^e}{dt} \times \mathbf{x}^e \right) \cdot \iiint_{V^q} \nabla \cdot p\mathbf{I} \cdot \delta \mathbf{u}^q dV \quad (22)$$

$$\delta \boldsymbol{\theta} \Big|_{t_0}^{t_1} + \int_{t_0}^{t_1} \rho^e \frac{d}{dt} \left(\frac{d\mathbf{u}^e}{dt} \times \mathbf{x}^e \right) \cdot \delta \boldsymbol{\theta} dt \quad (14) \quad - \iiint_{V^q} \mu \left(\nabla \frac{d\mathbf{u}^q}{dt} + \frac{d\mathbf{u}^q}{dt} \nabla \right) : \delta \nabla \mathbf{u}^q dV =$$

$$\int_{t_0}^{t_1} \rho^q \frac{d\mathbf{u}^q}{dt} \cdot \delta \frac{d\mathbf{X}^c}{dt} dt = \rho^q \frac{d\mathbf{u}^q}{dt} \cdot \delta \mathbf{X}^c \Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} \rho^q \frac{d^2 \mathbf{u}^q}{dt^2} \cdot \delta \mathbf{X}^c dt \quad (15)$$

$$- \iint_{S_w + S_f} \mu \left(\nabla \frac{d\mathbf{u}^q}{dt} + \frac{d\mathbf{u}^q}{dt} \nabla \right) \cdot \mathbf{n}^q \cdot \delta \mathbf{u}^q dS + \iiint_{V^q} \nabla \cdot \mu \left(\nabla \frac{d\mathbf{u}^q}{dt} + \frac{d\mathbf{u}^q}{dt} \nabla \right) \cdot \delta \mathbf{u}^q dV \quad (23)$$

将式(9)~式(23)代入式(8)的变换式中,考虑到位移边界条件的变分式为 $\delta \mathbf{u}^e = \mathbf{0}$,并且按惯例在时域边界 $t = t_0$ 和 $t = t_1$ 处取 $\delta \mathbf{X}^c$ 、 $\delta \boldsymbol{\theta}$ 、 $\delta \mathbf{u}^e$ 、 $\delta \mathbf{u}^q$ 等于零,可得

$$\int_{t_0}^{t_1} \rho^q \left(\frac{d\boldsymbol{\theta}}{dt} \times \mathbf{x}^q \right) \cdot \delta \frac{d\mathbf{u}^q}{dt} dt = \rho^q \left(\frac{d\boldsymbol{\theta}}{dt} \times \mathbf{x}^q \right) \cdot \delta \mathbf{u}^q \Big|_{t_0}^{t_1} - \int_{t_0}^{t_1} \rho^q \frac{d}{dt} \left(\frac{d\boldsymbol{\theta}}{dt} \times \mathbf{x}^q \right) \cdot \delta \mathbf{u}^q dt \quad (17)$$

$$\delta \pi_{\text{req}} - \delta Q_{\text{req}} = \int_{t_0}^{t_1} \left\{ \left(- \iiint_{V^c} \rho^c \frac{d^2 \mathbf{X}^c}{dt^2} dV - \iiint_{V^q} \rho^q \frac{d^2 \mathbf{X}^c}{dt^2} dV - \iiint_{V^c} \rho^c \frac{d^2 \mathbf{u}^c}{dt^2} dV - \iiint_{V^q} \rho^q \frac{d^2 \mathbf{u}^q}{dt^2} dV + \mathbf{F} \right) \cdot \delta \mathbf{X}^c + \right.$$

$$\left. \left[\iiint_{V^c} \rho^c \frac{d}{dt} \left(\frac{d\mathbf{u}^c}{dt} \times \mathbf{x}^c \right) dV + \iiint_{V^q} \rho^q \frac{d}{dt} \left(\frac{d\mathbf{u}^q}{dt} \times \mathbf{x}^q \right) dV - \frac{d}{dt} \left(\mathbf{J} \cdot \frac{d\boldsymbol{\theta}}{dt} \right) + \mathbf{M} \right] \cdot \delta \boldsymbol{\theta} + \iiint_{V^c} \left[-\rho^c \frac{d^2 \mathbf{X}^c}{dt^2} - \rho^c \frac{d}{dt} \left(\frac{d\boldsymbol{\theta}}{dt} \times \mathbf{x}^c \right) - \rho^c \frac{d^2 \mathbf{u}^c}{dt^2} + (\mathbf{I} + \mathbf{u}^c \nabla) \cdot \right.$$

$$\left. \frac{\partial A \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)}{\partial \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)} \cdot \nabla + \mathbf{f}^c \right] \cdot \delta \mathbf{u}^c dV - \iint_{S_\sigma} \left[(\mathbf{I} + \mathbf{u}^c \nabla) \cdot \frac{\partial A \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)}{\partial \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)} \cdot \right.$$

$$\left. \frac{\partial A \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)}{\partial \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)} \cdot \mathbf{n}^c - \mathbf{T}^c \right] \cdot \delta \mathbf{u}^c dS + \iint_{S_w} (\mathbf{u}^c - \mathbf{u}^q) \cdot \delta \boldsymbol{\lambda} dS + \iint_{S_w} \left[-(\mathbf{I} + \mathbf{u}^c \nabla) \cdot \frac{\partial A \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)}{\partial \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)} \cdot \right.$$

$$\left. \frac{\partial A \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)}{\partial \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)} \cdot \nabla \right] \cdot \delta \mathbf{u}^c dV =$$

$$\iint_{S_\sigma + S_u} \left[(\mathbf{I} + \mathbf{u}^c \nabla) \cdot \frac{\partial A \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)}{\partial \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)} \cdot \right.$$

$$\left. \frac{\partial A \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)}{\partial \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)} \cdot \nabla \right] \cdot \delta \mathbf{u}^c dV \quad (21)$$

$$\left. \frac{\partial A \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)}{\partial \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)} \cdot \nabla \right] \cdot \delta \mathbf{u}^c dV - \iint_{V^q} p\mathbf{I} : \delta \nabla \mathbf{u}^q dV = \iint_{S_w + S_f} p\mathbf{I} \cdot \mathbf{n}^q \cdot \delta \mathbf{u}^q dS -$$

$$\iint_{V^q} p\mathbf{I} : \delta \nabla \mathbf{u}^q dV = \iint_{S_w + S_f} p\mathbf{I} \cdot \mathbf{n}^q \cdot \delta \mathbf{u}^q dS - \iint_{V^q} \left[-\rho^q \frac{d^2 \mathbf{X}^c}{dt^2} - \rho^q \frac{d}{dt} \left(\frac{d\boldsymbol{\theta}}{dt} \times \mathbf{x}^q \right) - \rho^q \frac{d^2 \mathbf{u}^q}{dt^2} - \nabla \cdot p\mathbf{I} + \nabla \cdot \mu \left(\nabla \frac{d\mathbf{u}^q}{dt} + \frac{d\mathbf{u}^q}{dt} \nabla \right) + \right.$$

$$f^q \Big] \cdot \delta u^q dV - \iint_{S_r} \left[-pI \cdot n^q + \mu \left(\nabla \frac{du^q}{dt} + \frac{du^q}{dt} \nabla \right) \cdot n^q - T^q \right] \cdot \delta u^q dS \Big\} dt = 0 \quad (24)$$

由于 δX^e 、 $\delta \theta$ 、 δu^e 、 δu^q 和 $\delta \lambda$ 的任意性,故由式(24)可得拟驻值条件为

$$- \iiint_{V^e} \rho^e \frac{d^2 X^e}{dt^2} dV - \iiint_{V^q} \rho^q \frac{d^2 X^e}{dt^2} dV - \iiint_{V^e} \rho^e \frac{d^2 u^e}{dt^2} dV - \iiint_{V^q} \rho^q \frac{d^2 u^q}{dt^2} dV + F = 0 \quad (25-1)$$

$$\iiint_{V^e} \rho^e \frac{d}{dt} \left(\frac{du^e}{dt} \times x^e \right) dV + \iiint_{V^q} \rho^q \frac{d}{dt} \left(\frac{du^q}{dt} \times x^q \right) dV - \frac{d}{dt} (J \cdot \frac{d\theta}{dt}) + M = 0 \quad (25-2)$$

$$- \rho^e \frac{d^2 X^e}{dt^2} - \rho^e \frac{d}{dt} \left(\frac{d\theta}{dt} \times x^e \right) - \rho^e \frac{d^2 u^e}{dt^2} + (I + u^e \nabla) \cdot \frac{\partial A \left(\frac{1}{2} \nabla u^e + \frac{1}{2} u^e \nabla + \frac{1}{2} \nabla u^e \cdot u^e \nabla \right)}{\partial \left(\frac{1}{2} \nabla u^e + \frac{1}{2} u^e \nabla + \frac{1}{2} \nabla u^e \cdot u^e \nabla \right)} \nabla + f^e = 0 \quad (25-3)$$

$$(I + u^e \nabla) \cdot \frac{\partial A \left(\frac{1}{2} \nabla u^e + \frac{1}{2} u^e \nabla + \frac{1}{2} \nabla u^e \cdot u^e \nabla \right)}{\partial \left(\frac{1}{2} \nabla u^e + \frac{1}{2} u^e \nabla + \frac{1}{2} \nabla u^e \cdot u^e \nabla \right)} \cdot n^e = 0 \quad (25-4)$$

$$u^e - u^q = 0 \quad (\text{在 } S_w \text{ 上}) \quad (25-5)$$

$$- (I + u^e \nabla) \cdot \frac{\partial A \left(\frac{1}{2} \nabla u^e + \frac{1}{2} u^e \nabla + \frac{1}{2} \nabla u^e \cdot u^e \nabla \right)}{\partial \left(\frac{1}{2} \nabla u^e + \frac{1}{2} u^e \nabla + \frac{1}{2} \nabla u^e \cdot u^e \nabla \right)} \cdot n^e = 0 \quad (\text{在 } S_w \text{ 上})$$

$$n^e - T^e = 0 \quad (\text{在 } S_r \text{ 上}) \quad (25-4)$$

$$n^e + \lambda = 0 \quad (\text{在 } S_w \text{ 上}) \quad (25-6)$$

$$\left[-pI + \mu \left(\nabla \frac{du^q}{dt} + \frac{du^q}{dt} \nabla \right) \right] \cdot n^q + \lambda = 0 \quad (\text{在 } S_w \text{ 上}) \quad (25-7)$$

$$- \rho^q \frac{d^2 X^e}{dt^2} - \rho^q \frac{d}{dt} \left(\frac{d\theta}{dt} \times x^q \right) - \rho^q \frac{d^2 u^q}{dt^2} + \nabla \cdot \left[-pI + \mu \left(\nabla \frac{du^q}{dt} + \frac{du^q}{dt} \nabla \right) \right] + f^q = 0 \quad (25-8)$$

$$\left[-pI + \mu \left(\nabla \frac{du^q}{dt} + \frac{du^q}{dt} \nabla \right) \right] \cdot n^q - T^q = 0 \quad (\text{在 } S_r \text{ 上}) \quad (25-9)$$

由式(25-6)和式(25-7)解得

$$\lambda = (I + u^e \nabla) \cdot \frac{\partial A \left(\frac{1}{2} \nabla u^e + \frac{1}{2} u^e \nabla + \frac{1}{2} \nabla u^e \cdot u^e \nabla \right)}{\partial \left(\frac{1}{2} \nabla u^e + \frac{1}{2} u^e \nabla + \frac{1}{2} \nabla u^e \cdot u^e \nabla \right)} \cdot n^e \quad (26)$$

$$\lambda = - \left[-pI + \mu \left(\nabla \frac{du^q}{dt} + \frac{du^q}{dt} \nabla \right) \right] \cdot n^q \quad (27)$$

分析式(26)和式(27)的物理意义。不难看出,式(26)表示在 S_w 处弹性体所承受的应力,式(27)表示在 S_w 处液体所承受的应力。并且式(26)和式(27)都与经典分析力学中 Lagrange 乘子是约束力的论述相吻合。将式(26)代入式(27),或者将式(27)代入式(26),均可得到

$$- (I + u^e \nabla) \cdot \frac{\partial A \left(\frac{1}{2} \nabla u^e + \frac{1}{2} u^e \nabla + \frac{1}{2} \nabla u^e \cdot u^e \nabla \right)}{\partial \left(\frac{1}{2} \nabla u^e + \frac{1}{2} u^e \nabla + \frac{1}{2} \nabla u^e \cdot u^e \nabla \right)} \cdot n^e = \left[-pI + \mu \left(\nabla \frac{du^q}{dt} + \frac{du^q}{dt} \nabla \right) \right] \cdot n^q \quad (\text{在 } S_w \text{ 上}) \quad (28)$$

认为弹性体和液体以无际边界面为接触面,二者在此处相互作用,式(28)的物理意义可以解释为:在无际边界 S_w 处,弹性体对液体的作用力和液体对弹性体的作用力大小相等,方向相反。

式(28)又可写为另一种形式:

$$(I + u^e \nabla) \cdot \frac{\partial A \left(\frac{1}{2} \nabla u^e + \frac{1}{2} u^e \nabla + \frac{1}{2} \nabla u^e \cdot u^e \nabla \right)}{\partial \left(\frac{1}{2} \nabla u^e + \frac{1}{2} u^e \nabla + \frac{1}{2} \nabla u^e \cdot u^e \nabla \right)} \cdot n^e + \left[-pI + \mu \left(\nabla \frac{du^q}{dt} + \frac{du^q}{dt} \nabla \right) \right] \cdot n^q = 0 \quad (\text{在 } S_w \text{ 上}) \quad (29)$$

认为弹性力和液体力共同作用在无际边界 S_w 处的无限薄的一层物质面上,式(29)表明:由于刚-弹耦合动力学导致的应力和由于刚-液耦合动力学导致的应力在无际边界面 S_w 处构成平衡应力系。

在以上分析中,建议注意:在无际边界面 S_w 处, $n^e + n^q = 0$ 。

刚-弹-液耦合动力学的拟变分原理式(8)的丰富的内涵反映在其驻值条件中:驻值条件式(25-6)和驻值条件式(25-7)经变换得到式(29),反映弹-液耦合的力的平衡关系;驻值条件式(25-5)反映弹-液耦合的位移协调关系;这是本文的主要研究内容。驻值条件式(25-3)和驻值

条件式(25-4)反映刚-弹耦合特性;驻值条件式(25-8)和驻值条件式(25-9)反映刚-液耦合特性;驻值条件式(25-1)和驻值条件式(25-2)反映连续介质动力学对刚体动力学的影响。由如上论述可以看出,刚-弹-液耦合中弹-液耦合的平衡方程式(29)中的变量,不仅受到弹-液耦合的影响,还受到刚-液耦合(式(25-8)、式(25-9))和刚-弹耦合(式(25-3)、式(25-4))的影响。反之,刚-液耦合(式(25-8)、式(25-9))和刚-弹耦合(式(25-3)、式(25-4))的变量,也要受到弹-液耦合式(29)的影响。

综合以上分析,可以明确刚-弹-液耦合动力学中弹-液耦合的机理。刚-弹-液耦合动力学中弹-液耦合发生在弹、液两相的无际边界面上,在该边界面上,位移满足协调关系式(25-5),应力满足平衡关系式(29)。

应当注意到,虽然刚-弹-液耦合动力学中弹-液耦合是在弹-液交界面实现的,但是,弹-液耦合是受到控制方程式(25-1)~式(25-9)的制约的。

3 弹-液耦合机理与建模计算相结合

关于弹-液耦合的建模计算研究,首先会想到有限元建模计算问题。

结合第2节的研究,由拟驻值条件式(25-5)可以看出,通过应用 Lagrange 乘子法,已经将无际边界条件转化为泛函的拟驻值条件;从有限元建模计算方面考虑问题,这一研究进展实现了从协调元(compatible model)到杂交元(hybrid model)的过渡。国外学者应用 Lagrange 乘子法研究有限元素法基本上到此为止。

按照中国著名应用数学和力学专家钱伟长院士的理论,还要在上述研究的基础上,进一步识别 Lagrange 乘子,即将 Lagrange 乘子用泛函中原有的变量来表示,然后将已经识别的 Lagrange 乘子代入泛函中,用来减少有限元素法的计算自由度。

将 Lagrange 乘子的表达式(26)代入泛函式(8)中,可得

$$\begin{aligned} \delta\pi_{\text{req}} - \delta Q_{\text{req}} = & \delta \int_{t_0}^{t_1} \left\{ \iiint_{V^c} \rho^c \frac{1}{2} \cdot \frac{dX^c}{dt} \cdot \frac{dX^c}{dt} dV + \right. \\ & \iiint_{V^q} \rho^q \frac{1}{2} \cdot \frac{dX^c}{dt} \cdot \frac{dX^c}{dt} dV + \frac{1}{2} \cdot \frac{d\theta}{dt} \cdot \mathbf{J} \cdot \\ & \frac{d\theta}{dt} + \mathbf{F} \cdot \mathbf{X}^c + \mathbf{M} \cdot \theta + \iiint_{V^c} \rho^c \left(\frac{dX^c}{dt} + \right. \\ & \left. \left. \frac{d\theta}{dt} \times \mathbf{x}^c \right) \cdot \frac{d\mathbf{u}^c}{dt} dV + \right. \end{aligned}$$

$$\begin{aligned} & \iiint_{V^c} \left[\frac{1}{2} \rho \frac{d\mathbf{u}^c}{dt} \cdot \frac{d\mathbf{u}^c}{dt} - A \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \right. \right. \\ & \left. \left. \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right) + \mathbf{f}^c \cdot \mathbf{u}^c \right] dV + \iint_{S_r} \mathbf{T}^c \cdot \mathbf{u}^c dS + \\ & \iint_{S_e} (\mathbf{I} + \mathbf{u}^c \nabla) \cdot \\ & \frac{\partial A \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)}{\partial \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)} \cdot \\ & \mathbf{n}^e \cdot (\mathbf{u}^c - \mathbf{u}^q) dS + \iiint_{V^q} \rho^q \left(\frac{dX^c}{dt} + \frac{d\theta}{dt} \times \mathbf{x}^q \right) \cdot \\ & \frac{d\mathbf{u}^q}{dt} dV + \iiint_{V^q} \left[\frac{1}{2} \rho^q \frac{d\mathbf{u}^q}{dt} \cdot \frac{d\mathbf{u}^q}{dt} + \rho \mathbf{I} : \nabla \mathbf{u}^q - \right. \\ & \left. \mu \left(\nabla \frac{d\mathbf{u}^q}{dt} + \frac{d\mathbf{u}^q}{dt} \nabla \right) : \nabla \mathbf{u}^q + \mathbf{f}^q \cdot \mathbf{u}^q \right] dV + \\ & \left. \iint_{S_r} \mathbf{T}^q \cdot \mathbf{u}^q dS - (\mathbf{X}^c \cdot \delta \mathbf{F} + \theta \cdot \delta \mathbf{M}) \right\} dt - \\ & \int_{t_0}^{t_1} \left(\iiint_{V^c} \mathbf{u}^c \cdot \delta \mathbf{f}^c dV + \iint_{S_r} \mathbf{u}^c \cdot \delta \mathbf{T}^c dS \right) dt + \\ & \int_{t_0}^{t_1} \left\{ \iiint_{V^q} \left[\nabla \mathbf{u}^q : \delta \mu \left(\nabla \frac{d\mathbf{u}^q}{dt} + \frac{d\mathbf{u}^q}{dt} \nabla \right) - \right. \right. \\ & \left. \left. \mathbf{u}^q \cdot \delta \mathbf{f}^q \right] dV - \iint_{S_r} \mathbf{u}^q \cdot \delta \mathbf{T}^q dS \right\} dt = 0 \quad (30) \end{aligned}$$

应用与第2节相同的方法,不难求得,其驻值条件为

$$\begin{aligned} & - \iiint_{V^c} \rho^c \frac{d^2 \mathbf{X}^c}{dt^2} dV - \iiint_{V^q} \rho^q \frac{d^2 \mathbf{X}^c}{dt^2} dV - \iiint_{V^c} \rho^c \frac{d^2 \mathbf{u}^c}{dt^2} dV - \\ & \iiint_{V^q} \rho^q \frac{d^2 \mathbf{u}^q}{dt^2} dV + \mathbf{F} = \mathbf{0} \quad (31-1) \end{aligned}$$

$$\begin{aligned} & \iiint_{V^c} \rho^c \frac{d}{dt} \left(\frac{d\mathbf{u}^c}{dt} \times \mathbf{x}^c \right) dV + \iiint_{V^q} \rho^q \frac{d}{dt} \left(\frac{d\mathbf{u}^q}{dt} \times \mathbf{x}^q \right) dV - \\ & \frac{d}{dt} (\mathbf{J} \cdot \frac{d\theta}{dt}) + \mathbf{M} = \mathbf{0} \quad (31-2) \end{aligned}$$

$$\begin{aligned} & - \rho^c \frac{d^2 \mathbf{X}^c}{dt^2} - \rho^c \frac{d}{dt} \left(\frac{d\theta}{dt} \times \mathbf{x}^c \right) - \rho^c \frac{d^2 \mathbf{u}^c}{dt^2} + \\ & (\mathbf{I} + \mathbf{u}^c \nabla) \cdot \\ & \frac{\partial A \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)}{\partial \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)} \cdot \\ & \nabla + \mathbf{f}^c = \mathbf{0} \quad (\text{在 } V^c \text{ 中}) \quad (31-3) \end{aligned}$$

$$\begin{aligned} & (\mathbf{I} + \mathbf{u}^c \nabla) \cdot \\ & \frac{\partial A \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)}{\partial \left(\frac{1}{2} \nabla \mathbf{u}^c + \frac{1}{2} \mathbf{u}^c \nabla + \frac{1}{2} \nabla \mathbf{u}^c \cdot \mathbf{u}^c \nabla \right)} \cdot \\ & \mathbf{n}^e - \mathbf{T}^c = \mathbf{0} \quad (\text{在 } S_e \text{ 上}) \quad (31-4) \end{aligned}$$

$$\mathbf{u}^e - \mathbf{u}^q = \mathbf{0} \quad (\text{在 } S_w \text{ 上}) \quad (31-5)$$

$$-\rho^q \frac{d^2 \mathbf{X}^c}{dt^2} - \rho^q \frac{d}{dt} \left(\frac{d\boldsymbol{\theta}}{dt} \times \mathbf{x}^q \right) - \rho^q \frac{d^2 \mathbf{u}^q}{dt^2} + \nabla \cdot \left[-p\mathbf{I} + \mu \left(\nabla \frac{d\mathbf{u}^q}{dt} + \frac{d\mathbf{u}^q}{dt} \nabla \right) \right] + \mathbf{f}^q = \mathbf{0} \quad (31-6)$$

$$\left[-p\mathbf{I} + \mu \left(\nabla \frac{d\mathbf{u}^q}{dt} + \frac{d\mathbf{u}^q}{dt} \nabla \right) \right] \cdot \mathbf{n}^q - \mathbf{T}^q = \mathbf{0} \quad (\text{在 } S_f \text{ 上}) \quad (31-7)$$

比较式(31)和式(25)可以发现,识别 Lagrange 乘子之后,可以减少计算自由度。具体地说,式(25)和式(8)的 Lagrange 乘子尚未识别,它们仍然作为计算自由度存在于控制方程中;由于这里的 Lagrange 乘子 $\boldsymbol{\lambda}$ 是矢量,需要计算其 3 个分量方可确定,即其占用 3 个计算自由度。式(30)和式(31)的 Lagrange 乘子已经识别,即 Lagrange 乘子已经应用泛函中原有的变量表达出来,从而减少 3 个计算自由度。注意到,这是将刚-弹-液耦合系统划分为弹性体和液体 2 个元素,来处理其无际边界条件问题得到的结果。在有限元建模计算中,弹-液交界面可以离散为很多个元素的无际边界面,因此明确弹-液耦合机理之后,可以有效地减少计算自由度。具体地说,假设弹-液交界面可以离散为 100 个元素的无际边界面,根据前面的分析,明确弹-液耦合机理之后,可以减少 300 个计算自由度。

再引申一步,在有限元建模中,除了以弹-液交界面作为元素的无际边界面之外,还有更多的元素的无际边界面不是弹-液交界面,它们也存在用识别 Lagrange 乘子法减少计算自由度的问题。这样一来,减少的计算自由度就更多了。结合以上 2 种因素,可以大大减少计算自由度。这是中国学者对应用数学和力学的一大贡献。

最后说明,在刚-弹-液耦合动力学的建模计算研究中,存在弹-液交界面处的无际边界条件。把弹-液交界面处的无际边界条件处理好了,其余部分的有限元建模计算与一般的有限元建模计算便没有多大差别了,也就比较容易处理了。二者相结合,便是一个比较完整的刚-弹-液耦合动力学的建模计算研究。对于如何实现有限元计算的问题,可以参阅著名流固耦合专家英籍华人学者邢景棠先生的近期著述^[26]。

4 结 论

航天刚-弹-液耦合动力学是航天器动力学领域最具有挑战性的研究方向之一。

1) 通过分析刚-弹-液耦合动力学拟变分原理的泛函,说明了刚-弹耦合、刚-液耦合和弹-液耦合的特点。

2) 通过识别 Lagrange 乘子,逐步说明了弹-液耦合的机理。

3) 通过分析识别 Lagrange 乘子前后泛函的驻值条件,明确了识别 Lagrange 乘子可以有效地减少计算自由度。

研究工作处于该研究方向的理论基础层面,具有较好的理论意义,对这类系统有限元建模、计算及结果分析、讨论具有重要的参考价值。

参考文献 (References)

- [1] 王照林,刘延柱. 充液系统动力学[M]. 北京:科学出版社,2002.
WANG Z L, LIU Y Z. Fluid filling system dynamics[M]. Beijing: Science Press, 2002 (in Chinese).
- [2] 梁立孚,宋海燕,李海波. 航天分析动力学[M]. 北京:科学出版社,2016.
LIANG L F, SONG H Y, LI H B. Aerospace analytical dynamics [M]. Beijing: Science Press, 2016 (in Chinese).
- [3] 杜建斌,王勳成. 旋转周期性含液容器的流固耦合动力学特性分析[J]. 清华大学学报(自然科学版), 1999, 39(8): 108-116.
DU J B, WANG X C. Fluid solid coupling dynamic characteristics of rotating periodic liquid containing containers[J]. Journal of Tsinghua University (Natural Science Edition), 1999, 39(8): 108-116 (in Chinese).
- [4] 刘习军,张素侠,刘国英,等. 矩形弹性壳液耦合系统中的重力波分析[J]. 力学学报, 2006, 38(1): 106-112.
LIU X J, ZHANG S X, LIU G Y, et al. Gravity wave analysis in rectangular elastic shell and fluid coupling system[J]. Chinese Journal of Theoretical and Applied Mechanics, 2006, 38(1): 106-112 (in Chinese).
- [5] 马兴瑞,王本利,苟兴宇. 航天器动力学——若干问题进展及应用[M]. 北京:科学出版社,2001.
MA X R, WANG B L, GOU X Y. Some problems of spacecraft dynamics: Development and applications[M]. Beijing: Science Press, 2001 (in Chinese).
- [6] 邢景棠,周盛,崔尔杰. 流固耦合力学概述[J]. 力学进展, 1997, 27(1): 19-38.
XING J T, ZHOU S, CUI E J. Fluid solid coupling mechanics [J]. Advances in Mechanics, 1997, 27(1): 19-38 (in Chinese).
- [7] 王勳成. 有限单元法[M]. 北京:清华大学出版社,2003.
WANG X C. The finite element method[M]. Beijing: Tsinghua University Press, 2003 (in Chinese).
- [8] 李磊,任业军,蔡斌,等. 进动圆筒内液体流动不稳定的非线性特征[J]. 力学学报, 2005, 37(1): 64-72.
LI L, REN Y J, CAI B, et al. Nonlinear characteristics of liquid flow instability in precession cylinders[J]. Chinese Journal of Theoretical and Applied Mechanics, 2005, 37(1): 64-72 (in Chinese).

- [9] 席如青,曲广吉. 在低微重力下卫星液体晃动动力学应用研究[J]. 航天器工程,1997,6(1-2):18-26.
XI R Q, QU G J. Application of liquid sloshing dynamics in satellite under low gravity[J]. Spacecraft Engineering, 1997, 6(1-2):18-26(in Chinese).
- [10] 李英波,冯正进. 充液航天器液体晃动等效力学模型的建立[J]. 上海航天,2003(4):1-5.
LI Y B, FENG Z J. Establishment of equivalent mechanical model for liquid sloshing in liquid filled spacecraft[J]. Aerospace Shanghai,2003(4):1-5(in Chinese).
- [11] 王佳栋,周叮,刘伟庆. 带环形隔板圆柱形储罐中液体晃动的解析研究[J]. 振动与冲击,2010,29(2):54-59.
WANG J D, ZHOU D, LIU W Q. Analytical study of liquid sloshing in cylindrical tank with annular baffles[J]. Journal of Vibration and Shock,2010,29(2):54-59(in Chinese).
- [12] 黄华,杨雷,张焱,等. 航天器贮箱大幅液体晃动三维质心面等效模型研究[J]. 宇航学报,2010,31(1):55-59.
HUANG H, YANG L, ZHANG H, et al. The equivalent model of three dimensional centroid surface for large liquid sloshing in spacecraft tanks[J]. Journal of Astronautics,2010,31(1):55-59(in Chinese).
- [13] 贺元军,马兴瑞,王本利. 微重环境下平移圆柱贮箱液固耦合系统的动力响应研究[J]. 西安交通大学学报,2006,40(9):1083-1087.
HE Y J, MA X R, WANG B L. Research on dynamic response of liquid-solid coupling system of translational cylindrical tanks under microgravity environment[J]. Journal of Xi'an Jiao Tong University,2006,40(9):1083-1087(in Chinese).
- [14] 包光伟,王政伟. 液体三维晃动特征问题的有限元数值计算方法[J]. 力学季刊,2003,24(2):185-190.
BAO G W, WANG Z W. Finite element numerical method for three-dimensional sloshing characteristics of fluids[J]. Chinese Quarterly of Mechanics,2003,24(2):185-190(in Chinese).
- [15] 陈建平,周儒荣,虞伟建. 充液系统液体-多体耦合动力响应分析[J]. 力学学报,2004,36(6):724-731.
CHEN J P, ZHOU R R, YU W J. Fluid multi-body coupling dynamic response analysis of fluid filled system[J]. Chinese Journal of Theoretical and Applied Mechanics,2004,36(6):724-731(in Chinese).
- [16] 高索文,尹立中,王本利,等. 矩形贮箱类液固耦合系统转动特性分析[J]. 振动与冲击,2000,19(3):17-19.
GAO S W, YIN L Z, WANG B L, et al. Analysis of rotation characteristics of liquid solid coupling system for rectangular tanks[J]. Journal of Vibration and Shock,2000,19(3):17-19(in Chinese).
- [17] 苟兴宇,李铁寿,马兴瑞,等. 液固耦合系统中液体的有限幅晃动动力及晃动力矩[J]. 应用数学和力学,2001,22(5):465-476.
GOU X Y, LI T S, MA X R, et al. Finite amplitude sloshing power and sloshing torque in liquid solid coupling system[J]. Applied Mathematics and Mechanics,2001,22(5):465-476(in Chinese).
- [18] 李青,王天舒,马兴瑞. 充液航天器液体晃动和液固耦合动力学的研究与应用[J]. 力学进展,2012,42(4):471-480.
LI Q, WANG T S, MA X R. Research and application of liquid sloshing and liquid-solid coupling dynamics in liquid filled spacecraft[J]. Advances in Mechanics,2012,42(4):471-480(in Chinese).
- [19] LIANG L F, SONG H Y. Non-linear and non-conservative quasi-variational principle of flexible body dynamics and application in spacecraft dynamics[J]. Science China Physics, Mechanics and Astronomy,2013,56(11):2192-2199.
- [20] 梁立孚,刘宗民,郭庆勇. 充液系统刚-液耦合动力学功能型拟变分原理[J]. 哈尔滨工程大学学报,2013,34(12):1514-1519.
LIANG L F, LIU Z M, GUO Q Y. The rigid liquid coupled dynamic functional quasi variational principle of the fluid system[J]. Journal of Harbin Engineering University,2013,34(12):1514-1519(in Chinese).
- [21] 冯晓九,梁立孚,宋海燕. 刚-弹-液耦合动力学功能型拟变分原理[J]. 中国科学 E 辑,2017,46(2):195-203.
FENG X J, LIANG L F, SONG H Y. Rigid, elastic fluid coupling dynamics, functional quasi variational principles[J]. Chinese Science E Series,2017,46(2):195-203(in Chinese).
- [22] 梁立孚,郭庆勇. 刚体动力学的拟变分原理及其在应用[J]. 力学学报,2010,42(2):300-305.
LIANG L F, GUO Q Y. Quasi variational principles for rigid body dynamics and their applications[J]. Chinese Journal of Theoretical and Applied Mechanics,2010,42(2):300-305(in Chinese).
- [23] 钱伟长. 广义变分原理[M]. 北京:知识出版社,1985.
QIAN W C. Generalized variational principle[M]. Beijing: Knowledge Press,1985(in Chinese).
- [24] WASHIZU K. Variational method in elasticity and plasticity[M]. New York:Pergamon Press,1982.
- [25] 钱伟长. 弹性理论中广义变分原理的研究及其在有限元计算中的应用[J]. 力学与实践,1979,1(1):16-24.
QIAN W C. The study of the principle of generalized variational principles in elastic theory and its application in finite element calculation[J]. Mechanics in Engineering,1979,1(1):16-24(in Chinese).
- [26] XING J T. Developments of numerical method for linear and nonlinear fluid-solid interaction dynamics with applications[J]. Advances in Mechanics,2016,46(2):95-139.

作者简介:

梁立孚 男,教授,博士生导师。主要研究方向:变分原理及其应用、航天分析动力学。

Elastic-liquid coupling in aerospace rigid-elastic-liquid coupling system

LIANG Lifu*, GUO Qingyong

(College of Aerospace and Civil Engineering, Harbin Engineering University, Harbin 150001, China)

Abstract: This paper focuses on the elastic-liquid coupling problem in the rigid-elastic-liquid coupling dynamics. It is an important topic related to researches on mechanism and large-scale modeling and calculation of liquid-solid coupling in the aerospace rigid-elastic-liquid coupling system. Based on the analysis of the function of the quasi-variational principles for the rigid-elastic-liquid coupling dynamics, characteristics for rigid-elastic coupling, rigid-liquid coupling and elastic-liquid coupling in the rigid-elastic-liquid coupling system were illustrated. Inter-element boundary conditions were plugged into the function by Lagrange multiplier method and the mechanization of the elastic-liquid coupling was illustrated step by step through the identification of Lagrange multipliers. Solving stationary values of function shows that inter-element boundary conditions convert from preconditions to stationary value conditions. It realizes the transition from compatible element model to the hybrid element model. Through analysis on stationary value conditions before and after identification of Lagrange multipliers, it is shown that the identified Lagrange multipliers can effectively reduce the degree of freedom for calculation.

Keywords: aerospace liquid-filled system; rigid-elastic-liquid coupling dynamics; elastic-liquid coupling mechanism; liquid-solid coupling modeling; inter-element boundary condition

Received: 2018-05-21; **Accepted:** 2018-10-19; **Published online:** 2018-11-08 09:44

URL: kns.cnki.net/kcms/detail/11.2625.V.20181106.1029.001.html

Foundation item: National Natural Science Foundation of China (10272034)

* **Corresponding author.** E-mail: lianglifu@hrbeu.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2018.0302



高速运载器燃油热管理系统优化

庞丽萍^{1,*}, 邹凌宇¹, 阿嵘², 杨晓东³, 范俊⁴

(1. 北京航空航天大学 航空科学与工程学院, 北京 100083; 2. 中国空间技术研究院 载人航天总体部, 北京 100094;
3. 北京机电工程研究所, 北京 100074; 4. 陆军航空兵研究所, 北京 101121)

摘 要: 燃油热管理系统设计随着运载器多电化与机载高能电子设备的发展已经得到高度重视, 其中燃油的热承载能力是最关键因素。针对喷气推进式高速运载器, 提出了一种大范围、多任务的燃油热管理系统多目标优化配置方法, 其以热沉利用率最高和燃油质量代价损失最小为目标函数, 以循环回路的燃油最大质量流量、冷却水携带量和机载热负荷发热量为优化变量, 采用改进的遗传算法 NSGA-II, 在不同飞行任务规划下进行双目标优化设计, 所获得的目标函数 Pareto 最优解集, 满足预期的燃油热管理系统模式选择原则, 且通过分析优化变量与优化目标间的相关性, 可以量化燃油热管理系统优化配置准则与可达到的最小燃油质量代价损失, 可应用于支持多热沉重构的机载高效燃油热管理系统。

关键词: 高速运载器; 热管理系统; 燃油热沉; 飞行时长; 消耗性冷却剂

中图分类号: V245.3

文献标识码: A **文章编号:** 1001-5965(2019)02-0252-07

随着高速运载器多电化发展及电子设备集成技术进步, 机载热负荷与能量需求呈指数上升趋势^[1-4], 特别是激光武器、长距雷达、电子对抗平台等高能设备的搭载, 使得机载系统对冷源的需求日益剧增。同时, 外部气动热不断累积, 机身温度随飞行时间增长而持续上升。然而, 有限的机载热沉难以应对上述冷却需求, 使得热问题成为限制高速运载器续航时间和电子设备使用时的技术障碍^[5-6]。

高速运载器热沉来源主要有 3 方面^[7-9]: 燃油热沉、外部空气热沉和消耗性冷却剂热沉。自美国 SR-71 应用燃油为主要热沉起^[10], 先进的喷气式运载器均采用燃油热管理系统, 燃油作为运载器所必需携带的大比热液体工质, 且在飞行中必需加热排出, 是高速运载器机载热沉中的优质热沉之选。但在高马赫数飞行时, 气动热在机内累

积, 最终会导致燃油和外部空气热沉作用逐渐丧失。这种情况下, 为了满足散热量和飞行时长的需求, 引入消耗性冷却剂热沉。因此, 高速运载器飞行时长与热负荷、马赫数、发动机耗油量、携带式消耗性冷却剂等密切相关。

国内外的学者和研究机构均为提升燃油热沉的可用热容量进行了大量的理论和实验研究。一方面, 通过改进燃油成分来提高燃油的热稳定性^[11], 在保持燃油热值的同时提升其发生碳化的温度, 从而增大燃油供给发动机前的可用热容量; 另一方面, 借助先进的脱氧系统^[12-14], 减少燃油中氧分子的含量, 从而提升燃油焦化温度, 在提升燃油系统安全性的同时, 扩大燃油热管理系统热容量。此外, 燃油热管理系统结构改进和优化也是提升热承载能力的重要手段^[15], 例如在 F-22 与 F-35 中采用循环回路附加蒙皮换热器的燃油

收稿日期: 2018-05-28; 录用日期: 2018-08-24; 网络出版时间: 2018-09-06 16:42

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180905.0946.002.html

基金项目: 国家重点研发计划 (2017YFB1201100)

* 通信作者. E-mail: pangliping@buaa.edu.cn

引用格式: 庞丽萍, 邹凌宇, 阿嵘, 等. 高速运载器燃油热管理系统优化[J]. 北京航空航天大学学报, 2019, 45(2): 252-258.

PANG L P, ZOU L Y, A R, et al. Optimization of fuel heat management system for high-speed aircraft [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 252-258 (in Chinese).

热管理系统^[16],其可控性和热承载能力均有提升。

针对马赫数为 3~4 的喷气推进式高速运载器,为进一步提升其热承载能力,适应不同飞行任务需求,本文提出了一种基于消耗性冷却剂的高速运载器燃油热管理系统。采用改进的遗传算法 NSGA-II,针对具体飞行任务需求,给出了热沉利用率最高和燃油质量代偿损失最小的燃油热管理系统的最优配置方案。该方法适用于大范围飞行任务和热负荷需求,优化结果可为高速运载器高效燃油热管理系统设计选型提供参考。

1 高速运载器燃油热管理系统

为了分析影响高速运载器燃油热管理系统飞行热航时的设计因素,本文建立了如图 1 所示的高速运载器燃油热管理系统。该系统将全程利用燃油热沉吸收高温聚 a 烯烃 (PAO)、液压油、发动机滑油、传动滑油的废热。

图中: \dot{m}_r 为循环回路燃油质量流量,kg/s; \dot{m}_i 为燃油油箱出口燃油质量流量,kg/s; \dot{m}_d 为供给发动机推进系统的燃油质量流量,kg/s; \dot{Q}_h 为单位时间内热沉从机载设备等内部热源所吸取的热量,即机载热负荷发热量,kW; \dot{Q}_c 为单位时间内燃油向消耗性冷却剂热沉排散的热量,kW; \dot{Q}_r 为单位时间内由循环回路流入燃油油箱的燃油总焓,kW; \dot{Q}_a 为单位时间内燃油排散到附面层空气热沉的热量,kW; \dot{Q}_1 为单位时间内流出燃油油箱的燃油总焓,kW; \dot{Q}_2 为单位时间内进入循环回路的燃油总焓,kW。

为增加与机载热源换热的燃油流量,增加回流回路,并在加流回路中引入额外热沉,为燃油系

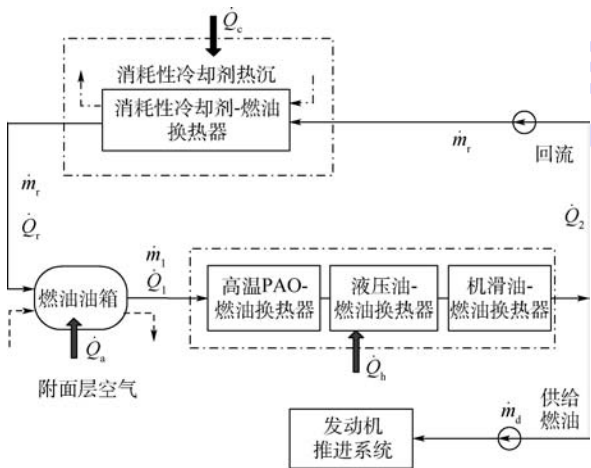


图 1 高速运载器燃油热管理系统架构

Fig. 1 Architecture of fuel heat management system for high-speed aircraft

统提供冷量。升温后的燃油,一部分输送给发动机产生推力,另一部分回流降温,当燃油升温有限时不启动消耗性冷却剂。只有长时间的高马赫数飞行并且燃油温升达到一定值后,燃油热沉无法满足机载散热需求,此时必须启动携带式消耗性冷却剂作为高温燃油的热沉,用来维持整个热管理系统的正常运行。消耗性制冷剂-燃油换热器是一个饱和温度被控制为 60℃ 的水蒸发器^[17-18]。

2 目标函数

2.1 热沉利用率

在燃油热管理系统设计中,为提高系统的经济性,希望所携带的热沉尽可能多的从机身内部热源吸取热量,且在飞行结束后热沉刚好用尽。由此定义经济性指标——热沉利用率,以评估全飞行包线内热沉利用率。

$$COP_{FTMS} = Q_h / \left\{ e_f c_{p,f} (T_{lim} - T_0) \int \dot{m}_d dt + m_w [c_{p,w} (T_{sat} - T_0) + r_w] \right\} \quad (1)$$

式中: Q_h 为热沉从机载设备等内部热源所吸取的总热量,kJ; e_f 为油箱剩余系数,%; $c_{p,f}$ 和 $c_{p,w}$ 分别为油和水的比热容,kJ/(kg·K); T_{lim} 、 T_0 和 T_{sat} 分别为发动机燃油温限、初始温度和水蒸发器饱和温度,K; m_w 为冷却水携带量,kg; r_w 为冷却水汽化潜热,kJ/kg。

对于超声速运载器,不考虑爬升、降落阶段的机外空气冷源作用。由于超声速巡航气动加热影响,其热沉利用率 $COP_{FTMS} < 1$,可通过燃油流量合理配置,减少冷却液的携带量提高热沉利用率。

2.2 燃油质量代偿损失

本文采用起飞总重量法来评价其对飞机系统性能的影响。引入循环回路及消耗性冷却剂所带来的燃油质量代偿损失,其可表示为

$$m_{all} = m_F + m_v + m_p \quad (2)$$

式中: m_{all} 为引入循环回路带来的总的燃油质量代偿损失,kg; m_F 为系统固定管路及设备所引起的燃油质量代偿损失,kg; m_v 为可变质量的消耗性冷却剂所引起的燃油质量代偿损失,kg; m_p 为燃油泵功耗造成的燃油质量代偿损失,kg。

系统固定管路及设备引起的燃油质量代偿损失为

$$m_F = m_r \left(\exp \left(\frac{C_e \tau_0 g}{K} \right) - 1 \right) \quad (3)$$

系统可变质量的消耗性冷却剂引起的燃油质量代偿损失为

$$m_v = \frac{m_{v'}K}{C_e\tau_0g} \left(\exp\left(\frac{C_e\tau_0g}{K}\right) - 1 \right) - m_{v'} \quad (4)$$

系统燃油泵耗功造成的燃油质量代偿损失为

$$m_p = \frac{q_{mP}K}{C_e g} \left(\exp\left(\frac{C_e\tau_0g}{K}\right) - 1 \right) \quad (5)$$

式中: $m_{v'}$ 为系统固定管路及设备质量,kg; C_e 为燃油比耗,kg/(N·s); τ_0 为续航时间,s; g 为重力加速度,m/s²; K 为运载器的气动质量,即升阻比; $m_{v'}$ 为航程中消耗的冷却剂的质量,kg; q_{mP} 为需要附加的燃油消耗量,kg。

3 优化设计

3.1 多目标优化函数及优化变量

在燃油热管理系统的优化配置分析中,系统结构轻量化、经济性和低消耗是关键考虑因素,设计优化目标确定如下:

1) 良好的经济性,即机载热沉利用率最高, $\max(\text{COP}_{\text{FTMS}})$ 。

2) 系统结构轻量化,即燃油质量代偿损失最小化, $\min(m_{\text{all}})$ 。

由此,多目标函数包括了 COP_{FTMS} 和 m_{all} 两个目标,可将目标函数表示为

$$f(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x})] \quad (6)$$

式中: $f_1(\mathbf{x})$ 和 $f_2(\mathbf{x})$ 分别表示 COP_{FTMS} 和 m_{all} ; \mathbf{x} 表示优化设计变量矩阵,即

$$\mathbf{x} = [x_1, x_2, x_3] = [\dot{Q}_h, \dot{m}_{r,\max}, m_w] \quad (7)$$

其中: $\dot{m}_{r,\max}$ 为系统循环回路中燃油的最大质量流量,kg/s。

3.2 约束条件

约束条件即对优化目标和优化变量的约束。对于优化变量 $\dot{m}_{r,\max}$,应满足循环回路系统的动量平衡方程,即

$$\dot{m}_{r,\max} \leq \sqrt{\frac{\Delta P \rho \pi^2 D^4}{8\zeta}} \quad (8)$$

式中: ΔP 为燃油泵增压,Pa; ρ 为燃油密度,kg/m³; D 为燃油管路直径,m; ζ 为燃油管路的阻力损失系数。

冷却水的最大携带量不超过所需最大换热量:

$$m_w \leq \int_0^\tau \frac{\dot{m}_r c_{p,f} (T_{\text{lim}} - T_{\text{sat}})}{r_w} \left(1 - \exp\left(-\frac{UA}{\dot{m}_r c_{p,f}}\right) \right) d\tau \quad (9)$$

式中: U 为蒸发器总换热系数,W/(K·m²); A 为蒸发器面积,m²。

机载热负荷发热量应满足最低热负荷 100 kW 要求,且不超过机载最大热容量:

$$100 \leq \dot{Q}_h < \frac{1}{\tau} [m_0 c_{p,r} (T_{\text{lim}} - T_0) + m_{w,\max} (c_{p,w} (T_{\text{sat}} - T_0) + r_w)] \quad (10)$$

式中: m_0 为燃油初始质量,kg。

综上所述,根据实际系统需求,可将 3 个优化变量的范围设计如下^[19]:

$$\begin{cases} 0 \leq \dot{m}_r \leq 8 \\ 0 \leq m_w \leq 400 \\ 100 \leq \dot{Q}_h \leq 1000 \end{cases} \quad (11)$$

此外增加可行配置方案的热航时与设计航时相等。

3.3 多目标优化方法

基于 3.1 节与 3.2 节内目标函数对燃油热管理系统进行优化配置设计。目标函数及燃油热管理系统模型基于 AMEsim 软件搭建,相关仿真参数详细列于表 1 中。

根据表 1 中相关参数,可在 AMEsim 中计算给定输入变量条件下的目标函数值。优化过程采用 AMEsim 与 Modefrontier 联合仿真,使用带精英策略的非支配排序遗传算法 NSGA-II^[20]进行计算,获得高速运载器优化解集。

通过 NSGA-II 对燃油热管理系统进行多目标优化,设定初始种群,经过选择、交叉、变异运算,最终获得符合约束条件且相对最优解集。NSGA-II 运算过程中的相关参数列于表 2 中。

表 1 燃油热管理系统多目标优化配置仿真参数
Table 1 Simulation parameters for multi-objective optimal configuration of fuel heat management system

参数	数值
仿真时间 τ_{design}/s	4 200
仿真步长 $\Delta\tau/s$	2
升阻比 K	4.62
油箱侧壁面积 A_1/m^2	94.5
油箱底面面积 A_2/m^2	46.25
燃油泵增压 $\Delta P/Pa$	1 000
油箱壁面发射率 ε	0.9
油箱外壁面厚度 δ_3/mm	1.2
燃油初始质量 m_0/kg	23 560
燃油初始温度 $T_0/^\circ C$	20
冷却水初始温度 $T_{w,0}/^\circ C$	20
冷却水饱和温度 $T_{\text{sat}}/^\circ C$	60
燃油泵效率 η	0.8
油箱内壁面厚度 δ_1/mm	1.2

表2 NSGA-II 算法参数设定

Table 2 NSGA-II algorithm parameter setting

参数	设定值
种群个数	10
种群代数	100
交叉概率	0.9
实数向量变异概率	1.0
二进制字符串变异概率	1.0
实数交叉分配指数	20
实数变异分配指数	20

4 多目标优化配置结果与分析

本文以飞行马赫数 $Ma = 4$ 为例进行分析。对于多目标优化问题,当目标函数互不矛盾时,通常可以得到一个绝对最优解。然而,多个目标函数往往不能同时达到最优值,常存在互相矛盾的问题。这种情况下,多目标函数的计算结果中存在一个解集,又叫做 Pareto 最优解集,处于 Pareto 最优解集中的解,其目标函数值都比解集外的可行解对应的目标函数值好。对双目标情况,该解集对应于目标空间的前沿曲线。本文中的 Pareto 最优解集如图 2 所示。在优化中,要求热沉利用率最大化、燃油质量代偿损失最小化,图中右下角深色曲线表示 Pareto 最优前沿。

由图 2 可见:①最优解主要分布在纵轴底部,即同等条件下燃油质量代偿损失最小。②当横轴坐标值小于 0.47 时,对应 Pareto 最优前沿的斜率较小,即随热沉利用率的增大,燃油质量代偿损失的增量较小。③当横轴坐标值大于 0.47 后,Pareto 最优前沿曲线出现拐点,随热沉利用率的进一步增大,燃油质量代偿损失显著增加。

循环回路燃油最大质量流量 $\dot{m}_{r,max}$ 决定单位时间参与机载热负荷散热的最大燃油质量, $\dot{m}_{r,max}$ 越大,燃油热管理系统所能承载的热负荷越高,而 $\dot{m}_{r,max}$ 又会受到管路流通面积的限制,且 $\dot{m}_{r,max}$ 越

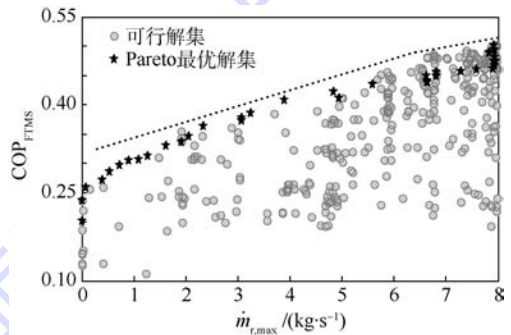
大,所需燃油泵的功率越大,产生的燃油质量代偿损失相应增大。循环回路燃油最大质量流量 $\dot{m}_{r,max}$ 与各优化目标的关系如图 3 所示。

分析图 3 可知:①在图 3(a)中,随 $\dot{m}_{r,max}$ 的增大,所能获得的热沉利用率均处于图中虚线以下,其热沉利用率最大值随 $\dot{m}_{r,max}$ 的提高显著上升。②在图 3(b)中,随 $\dot{m}_{r,max}$ 的增大,燃油质量代偿损失的最小值逐渐上升,由引起燃油质量代偿损失各因素可知,其最小值总在冷却水携带量最小时取得,且燃油质量代偿损失的最小值与 $\dot{m}_{r,max}$ 具有图中所示线性关系。

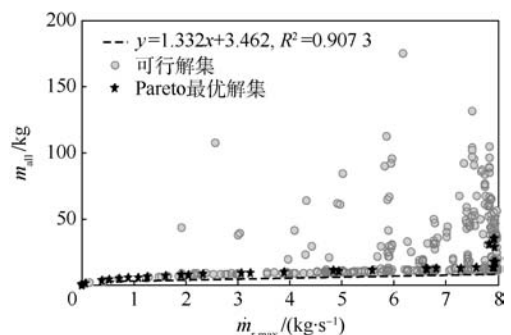
冷却水的携带量 m_w 与各优化目标的关系如图 4 所示。 m_w 的增加有利于增大机载总热沉量,但也会直接引起燃油质量代偿损失中固定质量和可变质量的代偿损失。

由图 4 可知:①在图 4(a)中,当不携带额外冷源时,热沉利用率最大值可达 0.47,与图 2 中前沿曲线拐点横坐标值相对应,随冷却水的质量增大,热沉利用率可继续上升至 0.5,之后不再升高。②在图 4(b)中,随 m_w 的增大,燃油质量代偿损失急速上升,Pareto 最优解则靠近坐标原点附近取得。

机载热负荷发热量 \dot{Q}_h 决定着系统承载能力,其他条件一定时, \dot{Q}_h 越大,热沉利用率越高,但 \dot{Q}_h



(a) 燃油最大质量流量与热沉利用率



(b) 燃油最大质量流量与燃油质量代偿损失

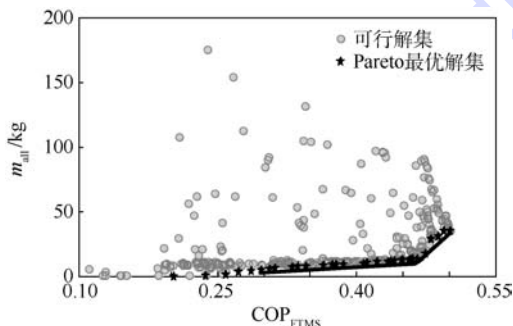
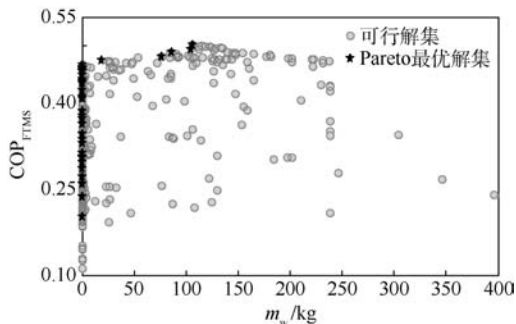


图2 多目标优化配置可行解的目标值空间

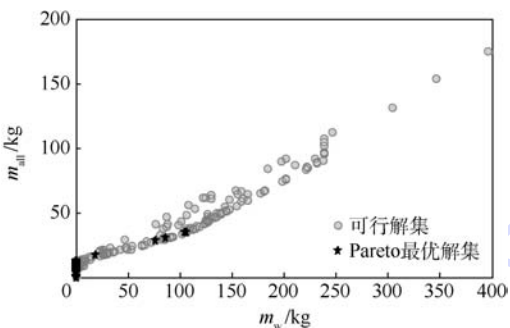
Fig.2 Target value space of optimal configuration of feasible solutions of multiple target

图3 燃油最大质量流量与优化目标之间的关系

Fig.3 Relationship between maximum mass flow rate of fuel and optimization objective



(a) 冷却水的携带量与热沉利用率



(b) 冷却水的携带量与燃油质量代偿损失

图4 冷却水的携带量与优化目标之间的关系

Fig.4 Relationship between cooling water carrying capacity and optimization objective

的增大会导致散热所需的燃油质量流量增大或燃油入口温度较低, 需要提高泵送功率及冷却水的携带量以满足系统需求。机载热负荷发热量 \dot{Q}_h 与各优化目标的关系如图5所示。

分析图5可知: ①在图5(a)中, 热沉利用率与 \dot{Q}_h 大体呈线性关系, 热沉利用率随 \dot{Q}_h 的增大逐渐提高, 但当超过 748 kW 时, 随 \dot{Q}_h 的增大热沉利用率上升斜率略有下降。②在图5(b)中, 随 \dot{Q}_h 的增大, 燃油质量代偿损失先缓慢上升, 超过 748 kW 后上升斜率显著增加。

因此, 在燃油热管理系统参数配置过程中, 需根据机载热负荷发热量匹配循环回路燃油最大质量流量和冷却水的携带量, 从而在保持较高的热沉利用率的同时产生尽量小的燃油质量代偿损失, 实现系统优化配置。

根据 Pareto 最优解集, 高速运载器在不同机载热负荷条件下的优化配置方案如图6所示。

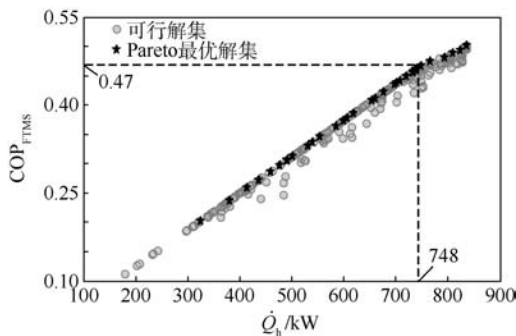
对所得数据点分段拟合, 得到优化配置方案。

$\dot{Q}_h \in [100, 380]$ kW 时,

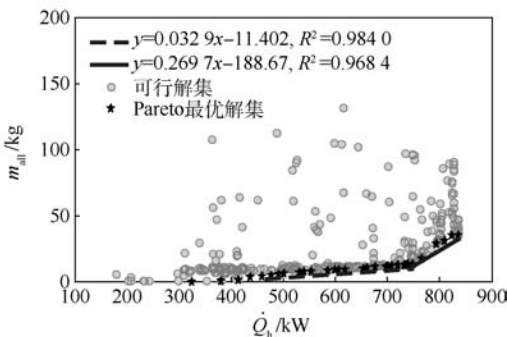
$$\begin{cases} \dot{m}_{r, \max} = 0 \\ m_w = 0 \end{cases}$$

$\dot{Q}_h \in [380, 748]$ kW 时,

$$\begin{cases} \dot{m}_{r, \max} = 5.5 \times 10^{-5} \dot{Q}_h^2 - 0.041 \dot{Q}_h + 7.81 \\ m_w = 0 \end{cases}$$



(a) 机载热负荷发热量与热沉利用率



(b) 机载热负荷发热量与燃油质量代偿损失

图5 机载热负荷发热量与优化目标之间的关系

Fig.5 Relationship between airborne heat load and optimization objective

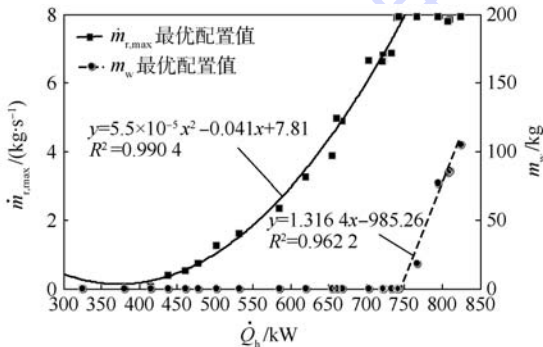


图6 燃油热管理系统参数优化配置方案

Fig.6 Parameter optimal configuration scheme of fuel heat management system

$\dot{Q}_h \in [748, 836]$ kW 时,

$$\begin{cases} \dot{m}_{r, \max} = 8 \\ m_w = 1.3164 \dot{Q}_h - 985.26 \end{cases}$$

5 结论

针对不同飞行任务和机载负荷需求, 本文提出了一种大范围、多任务的燃油热管理系统多目标优化配置方法, 以获得燃油热管理系统循环回路流量和消耗性冷却剂携带量的最优值。

1) 以马赫数为 4 飞行工况为例, 给出了不同机载热负荷需求下, 燃油热管理系统系统的最优

设计结果。通过选取最优配置值,可将系统代偿损失控制在最小值,以提高发动机性能。该方法以热沉利用率最高、燃油代偿损失最小为双目标,以循环回路燃油最大质量流量、冷却水携带量和机载热负荷发热量为优化变量。

3) 采用改进的遗传算法 NSGA-II 实现燃油热管理系统的优化配置设计,所获得的目标函数 Pareto 最优解集,满足预期的燃油热管理系统模式选择原则,且通过深入分析优化变量与优化目标间的相关性,可以量化燃油热管理系统优化配置准则与可达到的最小燃油质量代偿损失。

参考文献 (References)

- [1] MAHEFKEY T, YERKES K, DONOVAN B, et al. Thermal management challenges for future military aircraft power systems [C] // Power Systems Conference. Warrendale, PA: SAE International, 2004: 3204.
- [2] VAN GRIETHUYSEN V J, GLICKSTEIN M R, PETLEY D H, et al. High-speed flight thermal management [M] // MURTHY S N B, CURRAN E T. Developments in high-speed vehicle propulsion. Reston: AIAA, 1996, 165: 517-579.
- [3] DOOLEY M, LUI N, NEWMAN R, et al. Aircraft thermal management heat sink challenge [C] // SAE 2014 Aerospace Systems and Technology Conference. Warrendale, PA: SAE International, 2014: 2193.
- [4] GANEV E, KOERNER M. Power and thermal management for future aircraft [C] // SAE 2013 AeroTech Congress & Exhibition. Warrendale, PA: SAE International, 2013: 2273.
- [5] FISCHER A. Design of a fuel thermal management system for long range air vehicles; AIAA-2005-5647 [R]. Reston: AIAA, 2005.
- [6] YU S, GANEV E. Next generation power and thermal management system [J]. SAE International Journal of Aerospace, 2009, 1(1): 1107-1121.
- [7] DOOLEY M, LUI C, NEWMAN R. Efficient propulsion, power, and thermal management integration [C] // 49th AIAA/ASME/SAE/ASEE Joint Propulsion Conference. Reston: AIAA, 2013: 1-8.
- [8] SPROUSE J. F-22 environmental control/thermal management fluid transport optimization [C] // 30th International Conference on Environmental Systems. Warrendale, PA: SAE International, 2000: 2266.
- [9] 高峰, 袁修干. 高性能战斗机燃油热管理系统 [J]. 北京航空航天大学学报, 2009, 35(11): 1353-1356.
- GAO F, YUAN X G. Fuel thermal management system of high performance fighter aircraft [J]. Journal of Beijing University of Aeronautics and Astronautics, 2009, 35(11): 1353-1356 (in Chinese).
- [10] BLACKBIRD H, FLIGHT A S. SR-71 flight manual [EB/OL]. SR-71 Online, 2001: 1-57 [2001-12-11]. <https://www.sr-71.org/black-bird/manual>.
- [11] EDWARDS T. Liquid fuels and propellants for aerospace propulsion: 1903—2003 [J]. Journal of Propulsion and Power, 2003, 19(6): 1089-1107.
- [12] LAMM F P, VANDERSPURT T H. Catalytic fuel deoxy-generation system: US20060196174 [P]. 2006-09-07.
- [13] MORRIS R, MILLER J, LIMAYE S. Fuel deoxygenation and aircraft thermal management [C] // International Energy Conversion Engineering Conference and Exhibit. Reston: AIAA, 2006: 26-29.
- [14] RHEAUME J, CORDATOS H. Fuel deoxygenation systems: US 20180016025 [P]. 2018-01-18.
- [15] LETLOW J T, JENKINS L C. Development of an integrated environmental control system [C] // International Conference on Environmental Systems. Warrendale, PA: SAE International, 1998: 981544.
- [16] DOMAN D B. Fuel flow topology and control for extending aircraft thermal endurance [J]. Journal of Thermophysics and Heat Transfer, 2018, 32(1): 35-50.
- [17] ROUKIS J G, BOCCHICCHIO R L. The performance of the water bubbler heat sink for the EF-111 aircraft [C] // ASME 7th Intersociety Conference on Environmental Systems. New York: ASME, 1977: 1-5.
- [18] HARRISON W E, BINNS K E, ANDERSON S D, et al. High heat sink fuels for improved aircraft thermal management [C] // International Conference on Environmental Systems. Warrendale, PA: SAE International, 1993: 932084.
- [19] HO Y H, LIN T, HILL B, et al. Thermal benefits of advanced integrated fuel system using JP-8 + 100 fuel [C] // World Aviation Congress. Warrendale, PA: SAE International, 1997: 975507.
- [20] DEB K, PRATAP A, AGARWAL S, et al. A fast and elitist multiobjective genetic algorithm: NSGA-II [J]. IEEE Transactions on Evolutionary Computation, 2002, 6(2): 182-197.

作者简介:

庞丽萍 女,博士,教授。主要研究方向:飞行器环境控制。

邹凌云 男,硕士研究生。主要研究方向:飞行器环境控制。

Optimization of fuel heat management system for high-speed aircraft

PANG Liping^{1,*}, ZOU Lingyu¹, A Rong², YANG Xiaodong³, FAN Jun⁴

(1. School of Aviation Science and Engineering, Beihang University, Beijing 100083, China;

2. Institute of Manned Space System Engineering, China Academy of Space Technology, Beijing 100094, China;

3. Beijing Electro-Mechanical Engineering Institute, Beijing 100074, China; 4. Army Aviation Research Institute, Beijing 101121, China)

Abstract: With the rapid development of multi-electrification of aircraft and airborne high-energy electronic equipment, the design of fuel heat management system has been paid great attention to. The most critical factor is the thermal load capacity of fuel. For jet propulsion high-speed aircraft, this paper presents a multi-objective optimal allocation method for a large-scale and multi-task fuel heat management system. The thermal carrying capacity of fuel decreases with the increase of flight time, due to the dual effect of airborne thermal load and aerodynamic heating. In this paper, the improved genetic algorithm NSGA-II is used to optimize the design of two targets under different flight mission planning. The objective function is heat sink efficiency and fuel compensation loss. The optimization variables are the maximum flow rate of the fuel cycle, the consumption of coolant and the heat load on board. The objective function Pareto optimal solution set is obtained to meet the expected model selection principle of the fuel heat management system. By analyzing the correlation between the optimized variable and the optimization target, the optimization configuration criterion and the minimum fuel compensation loss can be quantified, and the airborne efficient heat management system supporting the multiple heat sink reconstruction is designed.

Keywords: high-speed aircraft; heat management system; fuel heat sink; flight time length; expendable coolant

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

DOI: 10.13700/j.bh.1001-5965.2018.0281

基于混沌粒子群优化的北斗/GPS 组合导航选星算法



王尔申^{1,2,3,*}, 贾超颖¹, 曲萍萍¹, 黄煜峰¹, 庞涛¹, 别玉霞¹, 姜毅⁴

(1. 沈阳航空航天大学 电子信息工程学院, 沈阳 110136; 2. 北京航空航天大学 电子信息工程学院, 北京 100083; 3. 沈阳航空航天大学 辽宁省通用航空重点实验室, 沈阳 110136; 4. 大连海事大学 水上智能交通行业重点实验室, 大连 116026)

摘 要: 全球卫星导航系统(GNSS)接收机在接收信号的过程中会受到诸如建筑物遮挡、信号干扰等因素的影响,无法得到全部可见星。为减轻多星座组合接收机的处理负担,研究利用部分可见卫星进行定位的快速选星算法,提出了一种基于混沌粒子群优化(CPSO)的北斗/GPS组合导航选星算法。首先,对当前历元时刻可见卫星进行连续编码,按照选星数目分组,每个分组视为一个粒子。然后,通过混沌映射初始化粒子种群,选取几何精度因子(GDOP)作为评价粒子优劣的适应度函数;粒子通过粒子群优化算法的速度-位移模型更新自身位置,逐渐趋近空间卫星几何分布较好的卫星组合全局最优解。最后,采集北斗/GPS实际数据对选星算法进行仿真验证和性能比较,结果表明,所提算法在选星颗数多于5颗时,单次选星耗时为遍历法选星的37.5%,选星结果的几何精度因子计算误差在0~0.6之间。该算法可适用于北斗/GPS组合导航定位不同选星颗数的情况。

关键词: 北斗/GPS组合导航;选星;混沌粒子群优化(CPSO);几何精度因子(GDOP);适应度函数

中图分类号: V241.6; TN967.1

文献标识码: A 文章编号: 1001-5965(2019)02-0259-07

随着全球卫星导航系统(Global Navigation Satellite System, GNSS)的不断建设和完善,由多个系统支持的组合导航成为了可能。组合导航使得可见卫星数大大增加,获得远比单一星座更优的卫星几何结构和更多的导航冗余信号,同时也加大了接收机的信号处理负担^[1-2]。此外,导航信号在传播过程中还会受到电离层延迟、对流层延迟、建筑物遮挡、电磁波干扰等因素的影响,接收机很难接收到全部可见卫星。为此,根据导航

需求,在定位之前粗略判别接收机所在位置的可见卫星分布,从中选择一个卫星子集进行定位解算。通过最小化几何精度因子(Geometric Dilution of Precision, GDOP)来获得最佳卫星子集的算法称为最优选星算法;以使接收机与可见卫星构成空间体积尽可能大,选择卫星子集最佳的算法称为最大体积算法。对于单一系统,由于定位和时钟偏差,定位所需的卫星最小数目为4颗^[3],而对于组合导航,由于不同系统的时钟偏差有所

收稿日期: 2018-05-17; 录用日期: 2018-06-29; 网络出版时间: 2018-07-26 19:31

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180725.1854.007.html

基金项目: 国家自然科学基金(61571309,61101161); 中央高校基本科研业务费专项资金(3132016317); 辽宁“百千万人才工程”; 辽宁省教育厅项目(L2014059,L201716,UPRP2018198); 辽宁省高等学校优秀人才支持计划(LR2016069)

* 通信作者. E-mail: wanges_2016@126.com

引用格式: 王尔申, 贾超颖, 曲萍萍, 等. 基于混沌粒子群优化的北斗/GPS组合导航选星算法[J]. 北京航空航天大学学报, 2019, 45(2): 259-265. WANG E S, JIA C Y, QU P P, et al. BDS/GPS integrated navigation satellite selection algorithm based on chaos particle swarm optimization [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 259-265 (in Chinese).

差异,用于定位的卫星数目最少为5颗^[4]。因而,对于组合导航选星算法通常选择超过5颗卫星来尽可能提高定位精度。

现有的快速选星算法主要集中在2个方面:①从所有可见卫星中选择最佳卫星几何分布的某些可见卫星。研究者根据最优选星方案的卫星分布特点,利用卫星高度角和方位角信息实现卫星的区域划分,并对如何分配高仰角和低仰角的选星数目展开讨论^[5-6]。②GDOP的计算。GDOP通常可看作是接收机测量误差与定位误差之间的放大倍数,是用来衡量从 m 颗可见卫星($m > n$)中选择出的 n 颗卫星子集是否具有较好的空间几何分布的重要指标。GDOP的计算存在矩阵相乘和求逆过程,明显增加了选星的计算量。若实现快速选星,可通过求逆引理等方法简化求解GDOP的计算公式^[7],也可以采用优化算法尽可能减少GDOP的计算次数^[8-10],二者都能在一定程度上缩减选星耗时。2010年,Mosavi和Divband^[8]提出用遗传算法(GA)实现选星,通过GA的快速寻优能力减少GDOP的计算次数。然而GA选星存在易陷入局部最优的问题,研究者又对算法进行大量改进,提高了算法选星结果的准确性^[9-10]。但由于GA的计算过程需要调节的参数过多,所以在保证选星准确性的条件下,GA选星速度也会随着可见卫星数目的增加而减慢。

与GA类似,粒子群优化(Particle Swarm Optimization, PSO)算法是一种随机寻优算法。目前该算法已经在目标跟踪、图像处理等诸多领域得到应用^[11-12],但在多星座选星方面少有文献提及。本文针对PSO算法在组合导航选星上的应用进行研究,建立了PSO选星算法模型;引入混沌序列改进算法,提高选星有效性;通过实际的导航数据进行仿真实验,验证所提算法的性能。

1 混沌粒子群优化模型

在多星座组合卫星导航系统中,为了降低接收机的运算量,需要从可见的所有卫星中选出空间结构较优的一组用于定位。精度因子(DOP)常用来衡量卫星空间几何分布情况,其计算公式中存在矩阵相乘和求逆,是选星中比较耗时的运算。而PSO算法能够通过有限次迭代,从全部解空间中快速搜索到符合条件的目标解。因此,运用PSO算法的快速寻优能力,从全部可见卫星组合中快速选取空间几何分布较好的卫星组合,能够减少DOP的计算次数,从而减少选星的耗时。

PSO算法是Eberhart和Kennedy通过模仿鸟

类的觅食行为而产生的^[13],搜索空间的个体通过比较自身所经过的最优位置和种群中其他粒子最优位置,不断调整自身速度,使其向最优解靠拢。个体被称为“粒子”,每个粒子为 d 维空间的一个点,第 i 个粒子可以表示为 $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{id}]$,粒子在运动过程中,会根据自身经验产生个体极值 $\mathbf{pbest} = [p_{i1}, p_{i2}, \dots, p_{id}]$,同时在种群中会产生全局极值 \mathbf{gbest} ,粒子通过2个“极值”不断调整自身位置,位置更新如式(1)所示,使其不断趋近全局最优值。第 i 个粒子的位置变化速度被表示为 $\mathbf{v}_i = [v_{i1}, v_{i2}, \dots, v_{id}]$ 。

$$\begin{aligned} v_{id}(t+1) &= \omega v_{id}(t) + c_1 r_1 (\mathbf{pbest} - \mathbf{x}_{id}(t)) + \\ &\quad c_2 r_2 (\mathbf{gbest} - \mathbf{x}_{id}(t)) \quad (1) \\ \mathbf{x}_{id}(t+1) &= \mathbf{x}_{id}(t) + v_{id}(t+1) \quad i = 1, 2, \dots, N \quad (2) \end{aligned}$$

式中: t 为当前迭代次数; N 为种群中粒子的总数; ω 为惯性权重; c_1 和 c_2 为加速常数,分别调节向 \mathbf{pbest} 和 \mathbf{gbest} 方向的运动速度; r_1 和 r_2 为0~1之间均匀分布的随机数。另外,通过设置微粒的速度范围 $[v_{\min}, v_{\max}]$ 和位置范围 $[x_{\min}, x_{\max}]$,可以对粒子运动的步长进行适当的限定。

惯性权重 ω 对全局和局部搜索的平衡起到了重要作用,其值通常是从0.9~0.4线性递减的^[14]。在迭代过程中,表达式为 $\omega = \omega_{\max} - (\omega_{\max} - \omega_{\min})t/t_{\max}$, ω_{\max} 和 ω_{\min} 分别为最大和最小惯性权重, t_{\max} 为总的迭代次数。

PSO算法的结果容易陷入局部最优,为此引入混沌理论,提高结果的有效性。混沌搜索,即对于给定的优化函数,将变量从混沌空间映射到解空间,然后利用混沌变量进行搜索^[15],可避免搜索结果陷入“局部最优”。

Logistic映射是较为常见的混沌序列产生方法,其表达式为

$$z_{i+1} = f(z_i) = \mu z_i (1 - z_i) \quad i = 0, 1, \dots \quad (3)$$

式中: $z_i \in (0, 1)$ 。当 $z_0 \notin \{0, 0.25, 0.5, 0.75, 1\}$ 、 $\mu = 4$ 时,Logistic映射产生混沌序列。利用Logistic映射初始化均匀分布的初始粒子,能够增强PSO算法的稳定性^[16]。

首先生成(0,1)的随机数 z_i ,根据式(3)更新 z_i ,然后根据公式 $x_i = x_{i\min} + z_i (x_{i\max} - x_{i\min})$,将混沌空间映射到待优化的解空间。

2 北斗/GPS组合导航选星算法

2.1 混沌粒子群优化适应度函数

适应度函数又称为目标函数,是用来评价粒

子优劣的重要衡量标准,适应度函数的选取直接影响算法结果的有效性。本文采用的适应度函数是各可见卫星组合的 GDOP 值,用于评价所选卫星组合的空间几何结构性能。

选星算法主要以用户与可见卫星组合的空间几何分布为衡量标准。针对用户与可见卫星组合的空间几何分布特性的表征,DOP 是应用最广泛的参数,按照 DOP 进行选星可以保证定位精度^[1]。卫星导航定位系统的 DOP 可用 GDOP 与用户等效测距误差(User Equivalent Range Error, UERE)的乘积来表示,即

$$\sigma_p = \text{GDOP} \cdot \sigma_{\text{UERE}} \quad (4)$$

式中: σ_p 为导航定位位置/时间解的精度; σ_{UERE} 为伪距测量值的标准差;GDOP 与用户到卫星的几何结构有关。从式(4)可知,当 σ_{UERE} 确定时,GDOP 越小,定位精度越高。

GDOP 可定义为协方差矩阵的迹的平方根,即

$$\text{GDOP} = \sqrt{\text{trace}\{(\mathbf{H}^T \mathbf{H})^{-1}\}} \quad (5)$$

式中: $\text{trace}(\cdot)$ 表示矩阵的迹; \mathbf{H} 为观测矩阵。假设北斗/GPS 组合导航接收机能接收到 m 颗 GPS 卫星、 k 颗北斗卫星,即可产生 $m+k$ 个观测量,构成的观测矩阵为

$$\mathbf{H} = \begin{bmatrix} e_{G,1} & n_{G,1} & u_{G,1} & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ e_{G,m} & n_{G,m} & u_{G,m} & 1 & 0 \\ e_{B,1} & n_{B,1} & u_{B,1} & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ e_{B,k} & n_{B,k} & u_{B,k} & 0 & 1 \end{bmatrix} \quad (6)$$

式中: (e_i, n_i, u_i) 为接收机近似位置指向第 i 颗卫星单位矢量的方向余弦,下标 G 代表 GPS 星座,可见星数目为 $1, 2, \dots, m$,下标 B 代表北斗星座,可见星数目为 $1, 2, \dots, k$ 。

2.2 混沌粒子群优化选星算法设计

假设某时刻接收机接收到可见卫星总数为 n 颗,从中选取 m 颗,使其 GDOP 值尽可能小。在混沌粒子群优化(CPSO)算法选星中,每个粒子代表一种可见卫星组合,粒子的位置由 m 个元素决定,每个元素代表一颗可见卫星,CPSO 选星算法具体步骤如下:

步骤 1 可见卫星提取。根据导航电文,提取当前时刻仰角大于遮蔽角的卫星(本文中的遮蔽角取值为 5°),得到该时刻的可见卫星总数。

步骤 2 编码。对当前时刻接收机观测到的所有可见卫星进行随机排列,然后将可见卫星从 $1, 2, \dots, n$ 依次连续编码,编码与可见卫星一一对应。

步骤 3 生成初始种群。将 n 颗可见卫星按照每 m 颗为一组进行组合,形成 C_n^m 种可见卫星组合,每种组合方式被视为一个粒子。设定种群大小为 M ,根据式(3),从 C_n^m 种可见卫星组合中混沌搜索 M 个组合,形成初始种群 $G_0 = \{\mathbf{x}_{0i}\}$ ($i=1, 2, \dots, M$);种群中第 i 个粒子表示为 $\mathbf{x}_{0i} = [x_{0i,1}, x_{0i,2}, \dots, x_{0i,m}]$, x 为卫星编号, m 为选星颗数;初始速度 $\mathbf{v}_{0i} = [v_{0i,1}, v_{0i,2}, \dots, v_{0i,m}]$, v 为卫星号的改变量,初始速度设置为 $[0, 0, \dots, 0]$ 。下标“0”表示粒子经过 0 次迭代,即为初始位置和速度。

步骤 4 适应度的计算。本文采用的适应度函数是编码所对应的可见卫星组合的 GDOP,记为粒子的目标值 $f_{it} = \text{GDOP}_i$,下标“ t ”为粒子经过 t 次迭代。将初始种群中的粒子依次代入适应度函数中,得出各粒子的适应度值(即 GDOP)。将种群中 GDOP 最小的粒子设置为初始种群最优位置 \mathbf{gbest} ,每个粒子当前的位置 $\mathbf{g}_{t,i}$ 为初始个体最优位置 \mathbf{pbest} 。

步骤 5 更新。对于每个粒子,根据式(1)和式(2)不断修正种群中粒子的位置 \mathbf{x}_{it} 和速度 \mathbf{v}_{it} ,分别计算新位置对应的目标值 f_{it} ,并更新粒子所经过的最优位置 \mathbf{pbest} 和种群最优位置 \mathbf{gbest} 。直到达到最大迭代次数,终止迭代。

将 CPSO 算法应用到选星过程中,需要明确 3 个量:初始化粒子种群、选取适应度函数和速度位置的更新,其基本步骤流程如图 1 所示。

在算法流程中,粒子更新时需注意 2 个关键点:①可见卫星编号为整数,因此更新过程中必须

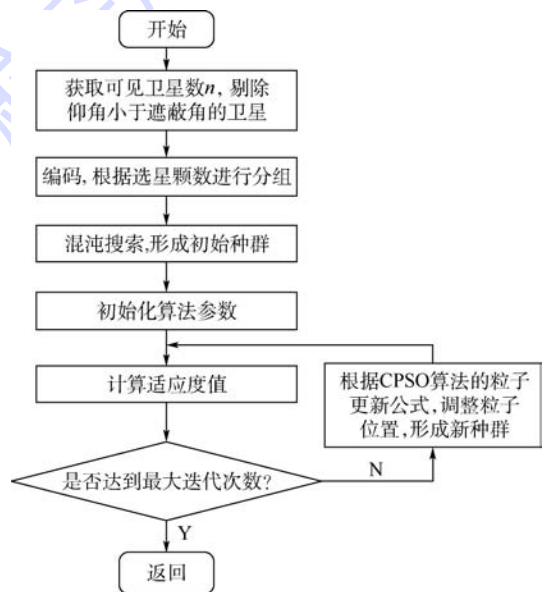


图 1 CPSO 选星算法流程

Fig. 1 Flowchart of CPSO satellite selection algorithm

保证粒子位置中的每个元素都为整数,否则将找不到与编码对应的可见星;②选星颗数为 m 颗,如果粒子位置中的元素有相同的,选星颗数就会少于 m 颗。为了保证选星数目,每次更新都必须判断更新后的粒子位置中是否有相同元素。本文给出的解决办法是在粒子进入迭代循环后,首先判断每个粒子中是否存在相同的元素。如果存在相同元素,则对粒子中元素从小到大排序,找出相同元素的个数及位置,在与之重复的第 k 个元素值上加 k ,然后返回重新判断是否有重复元素,直到该粒子中无相同元素为止。

3 仿真验证与结果分析

CPSO 选星是为了解决北斗/GPS 组合导航下选星颗数大于5的选星问题。本文选取北斗/GPS 接收机坐标为 $[-2\ 279\ 827.315\ 6, 5\ 004\ 704.309\ 4, 3\ 219\ 776.209\ 3]$ m,选星颗数为6。卫星位置由导航星历计算,卫星的截止高度角设为 5° ,仿真开始时间为2016-07-31 00:00:00,仿真时长为3h,仿真步长为1min。

本文将采用遍历法选星得到的GDOP值作为参考。假设全部可见卫星数目为 n ,从中选取 m 颗可见星,有 C_n^m 种组合。所谓遍历法,就是逐个计算 C_n^m 种组合的GDOP值,得出GDOP最小值。北斗/GPS可见卫星数目及其对应的GDOP最小值如图2所示。

由图2中可以看出,同一时刻,接收机接收到北斗/GPS可见卫星数目约为18颗,以从18颗可见卫星中选取6颗为例,需要进行 $C_{18}^6 = 18\ 564$ 次GDOP值的计算,单次选星耗时约为4s左右。

按照PSO选星步骤,设定算法参数:迭代次

数 $MaxIt = 50$,种群大小 $M = 100$,惯性权重 $\omega = 0.7298$,加速常数 $c_1 = c_2 = 1.4962$, $v_{max} = 2$ 。选取某一时刻进行GDOP值的计算,随着迭代次数的增加GDOP值的变化如图3所示。

通过遍历法得出该时刻的GDOP最小值为2.25。图3中的结果显示,PSO算法的收敛速度很快,迭代次数在小于15次时就稳定在2.34附近,且在后续迭代过程中保持不变。很明显,PSO算法出现“早熟”现象,即陷入局部最优解。

采用CPSO算法对同一时刻进行仿真实验,结果如图4所示。

从图4中可以看出,CPSO算法同样在迭代次数低于15次时收敛,并有效改善了PSO选星算法易陷入局部最优的缺点。在迭代次数为50次时,耗时约为1.5s左右,如果按需求适当减少迭代次数,选星耗时也会有所减少。

遍历法、PSO和CPSO选星算法在相同历元的选星耗时及其对应的选星结果(最后一列数字代表卫星号)如表1所示。

从表1中数据得知,基于PSO和CPSO的选星算法在单次选星所用时间约为遍历法选星的

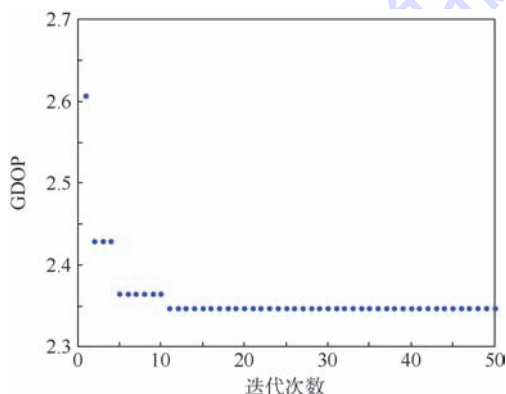


图3 PSO选星时GDOP值随迭代次数的变化

Fig.3 Change of GDOP with iteration numbers of PSO satellite selection

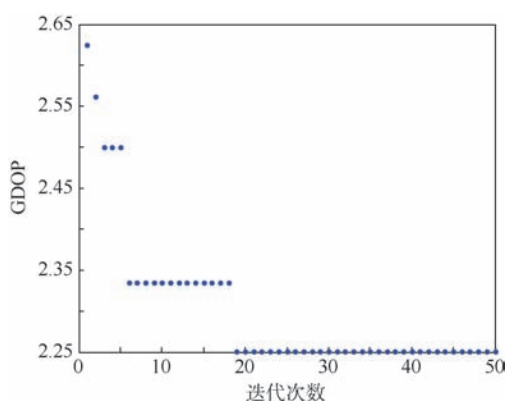


图4 CPSO选星时GDOP值随迭代次数的变化

Fig.4 Change of GDOP with iteration numbers of CPSO satellite selection

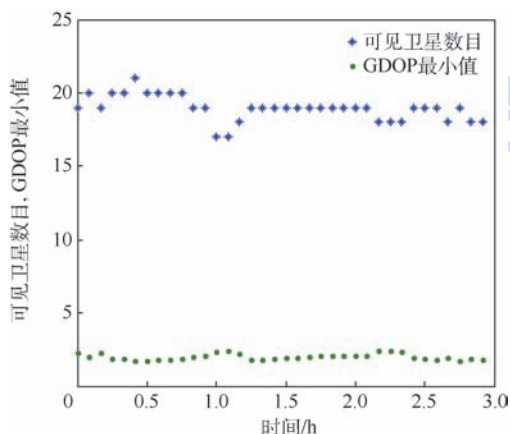


图2 北斗/GPS可见卫星数及对应的GDOP最小值

Fig.2 Number of BDS/GPS visible satellite and their minimum GDOP

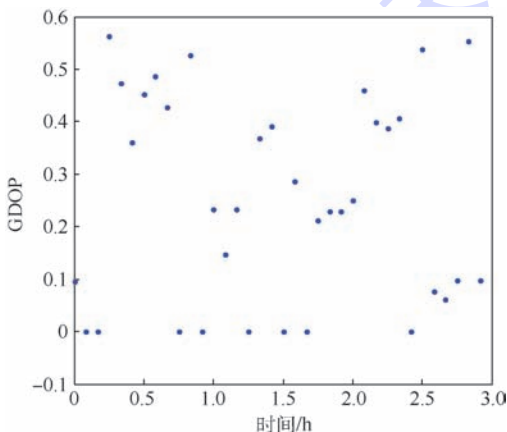
37.5%,但选星结果仍有偏差。为了充分验证算法性能,下面将给出仿真时长为3h的PSO和CPSO选星GDOP计算误差,其误差定义为新算法所得到的GDOP值与遍历法选星所得GDOP值的差值,结果如图5所示。

从图5中可知,PSO和CPSO选星算法的GDOP计算误差均小于等于0.6。对所得数据进行统计,CPSO选星的GDOP计算误差平均值为0.2609,方差为0.0424;PSO选星的GDOP计算误差平均值为0.2632,方差为0.0430。

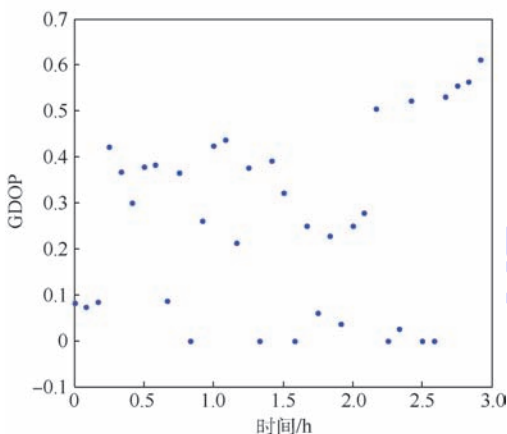
表1 三种选星算法性能对比

Table 1 Performance comparison of three satellite selection algorithms

选星算法	单次耗时/s	GDOP	选星结果
遍历法	4.074 824	2.251 038	92127373839
PSO	1.665 020	2.347 418	92127373842
CPSO	1.435 994	2.333 044	92127333837



(a) 基于PSO选星算法的GDOP计算误差



(b) 基于CPSO选星算法的GDOP计算误差

图5 PSO和CPSO选星的结果误差

Fig. 5 Result error of satellite selection by PSO and CPSO

4 结论

本文提出了一种基于CPSO的北斗/GPS组

合导航选星算法,利用CPSO算法的快速寻优能力,减少GDOP的计算次数,从而实现了北斗/GPS组合导航快速选星,通过对算法进行仿真验证,得到以下结果:

1) 在北斗/GPS组合导航下,选星颗数为6时,算法能够实现快速选星。该算法的单次选星时间约为1.5s,约为遍历法选星的37.5%。

2) 利用混沌方程初始化种群粒子能够提高选星结果的准确性。

3) PSO和CPSO选星算法的GDOP计算误差均小于等于0.6,二者的平均计算误差约为0.26。

本文将CPSO算法应用于组合卫星导航选星过程中,为多星座组合导航快速选星问题提供了新的解决方法。

参考文献 (References)

- [1] 张军. 空地协同的空域监视新技术[M]. 北京:航空工业出版社,2011:36-38.
ZHANG J. Air-ground collaborative airspace surveillance[M]. Beijing: Aviation Industry Press,2011:36-38 (in Chinese).
- [2] 王尔申,杨福霞,庞涛,等. BDS/GPS组合导航接收机自主完好性监测算法[J]. 北京航空航天大学学报,2018,44(4):684-690.
WANG E S, YANG F X, PANG T, et al. BDS/GPS combined navigation receiver autonomous integrity monitoring algorithm [J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(4):684-690 (in Chinese).
- [3] ZHANG M, ZHANG J. A fast satellite selection algorithm; Beyond four satellites[J]. IEEE Journal of Selected Topics in Signal Processing, 2009, 3(5):740-747.
- [4] SWASZEK P F, HARTNETT R J, SEALS K C, et al. Multi-constellation GNSS: New bounds on DOP and a related satellite selection process [C] // Proceedings of the 29th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS + 2016). Washington, D. C.: INST Navigation, 2016:228-235.
- [5] 丛丽, AHMED I A, 谈展中. 卫星导航几何因子的分析和仿真[J]. 电子学报, 2006, 34(12):2204-2208.
CONG L, AHMED I A, TAN Z Z. Analysis and simulation of the GDOP of satellite navigation [J]. Acta Electronica Sinica, 2006, 34(12):2204-2208 (in Chinese).
- [6] 陈灿辉, 张晓林. 一种新的卫星导航系统快速选星方法[J]. 电子学报, 2010, 38(12):2887-2891.
CHEN C H, ZHANG X L. A fast satellite selection approach for satellite navigation system [J]. Acta Electronica Sinica, 2010, 38(12):2887-2891 (in Chinese).
- [7] PHATAK M S. Recursive method for optimum GPS satellite selection [J]. IEEE Transactions on Aerospace & Electronic Systems, 2001, 37(2):751-754.
- [8] MOSAVI M R, DIVBAND M. Calculation of geometric dilution

- of precision using adaptive filtering technique based on evolutionary algorithms[C]//International Conference on Electrical and Control Engineering. Piscataway, NJ: IEEE Press, 2010: 4842-4845.
- [9] 宋丹,许承东,胡春生,等.基于遗传算法的多星座选星方法[J].宇航学报,2015,36(3):300-308.
SONG D,XU C D,HU C S,et al. Satellite selection with genetic algorithm under multi-constellation[J]. Journal of Astronautics,2015,36(3):300-308(in Chinese).
- [10] 霍航宇,张晓林.组合卫星导航系统的快速选星方法[J].北京航空航天大学学报,2015,41(2):273-282.
HUO H Y,ZHANG X L. Fast satellite selection method for integrated navigation systems[J]. Journal of Beijing University of Aeronautics and Astronautics,2015,41(2):273-282(in Chinese).
- [11] 徐小钧,马利华,艾国祥.基于多目标遗传算法的多星座选星方法[J].上海交通大学学报,2017,51(12):1520-1528.
XU X J,MA L H,AI G X. Satellite selection with multi-objective genetic algorithm for multi-GNSS constellations[J]. Journal of Shanghai Jiao Tong University,2017,51(12):1520-1528(in Chinese).
- [12] EBERHART R C,SHI Y H. Particle swarm optimization: Developments,applications and resources[C]//Proceedings of the 2001 Congress on Evolutionary Computation. Piscataway, NJ: IEEE Press,2002:81-86.
- [13] EBERHART R C,KENNEDY J. A new optimizer using Particle swarm theory[C]//Proceeding of the 6th International Symposium on Micro Machine and Human Science. Piscataway, NJ: IEEE Press,1995:39-43.
- [14] SHI Y H,EBERHART R. A modified particle swarm optimizer[C]//IEEE International Conference on Evolutionary Computation. Piscataway, NJ:IEEE Press,1998:69-73.
- [15] 胥小波,郑康锋,李丹,等.新的混沌粒子群优化算法[J].通信学报,2012,33(1):24-30.
XU X B,ZHENG K F,LI D,et al. New chaos-particle swarm optimization algorithm[J]. Journal on Communications,2012,33(1):24-30(in Chinese).
- [16] TIAN D P,SHI Z Z. MPSO:Modified particle swarm optimization and its applications[J]. Swarm & Evolutionary Computation,2018,41:49-68.

作者简介:

王尔申 男,博士,教授,硕士生导师。主要研究方向:卫星导航定位技术、接收机信号处理算法、航空电子技术。

贾超颖 女,硕士研究生。主要研究方向:卫星导航定位技术。

BDS/GPS integrated navigation satellite selection algorithm based on chaos particle swarm optimization

WANG Ershen^{1,2,3,*}, JIA Chaoying¹, QU Pingping¹, HUANG Yufeng¹,
PANG Tao¹, BIE Yuxia¹, JIANG Yi⁴

(1. School of Electronic and Information Engineering, Shenyang Aerospace University, Shenyang 110136, China;

2. School of Electronic and Information Engineering, Beihang University, Beijing 100083, China;

3. Liaoning General Aviation Key Laboratory, Shenyang Aerospace University, Shenyang 110136, China;

4. Key Laboratory of Intelligent Waterway Transport of Ministry of Transport, Dalian Maritime University, Dalian 116026, China)

Abstract: In the process of signal receiving, global navigation satellite system (GNSS) receiver will be affected by factors such as building blockages and signal interference and will not be able to obtain all the visible satellites; moreover, in order to reduce the processing burden of multi-constellation receivers, the fast satellite selection algorithm using partial visible satellites to achieve positioning solution is investigated, and the BDS/GPS integrated navigation satellite selection algorithm based on chaos particle swarm optimization (CP-SO) is proposed. First, the visible satellites are continuously numbered and randomly divided into groups. Each group is regarded as a particle. Then, chaotic maps are used to select several groups from all grouping spaces to form initial population. The geometric dilution of precision (GDOP) is chosen as fitness function to evaluate the particle's quality. In addition, the particle's position is updated by the velocity-displacement model of the PSO algorithm, and it gradually approaches the global optimal solution of the satellite combination with better geometric distribution of the space satellite. Finally, using real navigation data, the algorithm is verified by simulation experiments. The results demonstrate that when the number of selected satellite is more than 5, the time that the proposed algorithm takes to select satellite once is 37.5% of the time that the traversing algorithm takes, and the GDOP error of the selected satellites is between 0 and 0.6. Moreover, the proposed algorithm can be applied to the case of different numbers of selected satellite in BDS/GPS integrated navigation.

Keywords: BDS/GPS integrated navigation; satellite selection; chaos particle swarm optimization (CP-SO); geometric dilution of precision (GDOP); fitness function

Received: 2018-05-17; **Accepted:** 2018-06-29; **Published online:** 2018-07-26 19:31

URL: kns.cnki.net/kcms/detail/11.2625.V.20180725.1854.007.html

Foundation items: National Natural Science Foundation of China (61571309,61101161); the Fundamental Research Funds for the Central Universities (3132016317); Liaoning BaiQianWan Talents Program; Scientific Study Project for Liaoning Province Ministry of Education (L2014059,L201716,UPRP2018198); Program for Liaoning Excellent Talents in University (LR2016069)

* **Corresponding author.** E-mail: wanges_2016@126.com

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

DOI: 10.13700/j.bh.1001-5965.2018.0277

航空发动机转子结构布局优化设计方法

李超*, 金福艺, 张卫浩

(北京航空航天大学 能源与动力工程学院, 北京 100083)



摘 要: 结构是航空发动机功能、性能及可靠性设计水平的综合体现,一切技术要求、性能指标、强度指标或者结构的安全性和可靠性都应建立在合理的结构布局设计上。提出了航空发动机转子结构布局并对其优化设计的观点,同时以典型高推重比涡扇发动机高压转子的结构构型为例,基于试验设计(DOE)的响应面法,应用有限元计算并通过多目标遗传优化算法,分别从抗变形能力、力学环境适应能力以及转子结构效率综合进行相关优化计算,论证了合理的结构布局形式可以大幅度提升转子的力学特性。研究方法对于航空发动机转子系统的初始结构布局设计具有指导意义,可以显著减少结构设计的迭代次数,缩短设计周期。

关键词: 航空发动机; 转子系统; 结构布局; 构型设计; 结构效率

中图分类号: V232.2

文献标识码: A

文章编号: 1001-5965(2019)02-0266-11

目前,中国航空燃气涡轮发动机的研制正处于从测绘仿制到自主研发的过渡发展阶段,由于对结构设计的基础理论研究不足,在设计上大多依赖于传统的工程设计经验。而对于一台高性能、高可靠性的航空发动机来讲,单纯依靠长期积累的设计经验或测绘仿制不能满足其先进性的需求。所谓布局优化设计,即是结构构型的合理设计,只有合理的构型设计才可能最大程度的满足结构效率^[1]的要求。布局设计在飞行器外形^[2]、建筑^[3]、桥梁^[4]和经济^[5]等领域已经有着较为深入的研究,但是对航空发动机结构的布局设计,国内外均鲜有报道。目前,对于航空发动机转子系统的研究多集中在转子动力学特性的计算方法^[6-7],探索转子力学机理,研究影响其力学特征的关键因素^[8],但是实际工作转子结构型式通常较为复杂,影响因素繁多,很难在工程中得到较好的应用,特别是在航空发动机领域,很难大幅度地

对转子结构进行改进和优化设计。

鉴于此,本文以典型高推重比涡扇发动机高压转子的结构构型为例,探索了航空发动机转子结构布局优化设计的方法流程。首先,对各设计变量进行参数关联性研究,对各个待优化的构型参数对结构力学特性的敏感度进行统计分析;然后,在此基础上对结构的构型进行基于响应面的布局优化设计;最后,通过随机验证点充分验证优化结果的准确性。本文方法可以为结构改型提供指导方向,较大程度地摆脱转子结构的经验设计。

1 结构布局优化设计的定义及算例模型

结构布局设计是指基于某些数学优化算法使能够自动识别系统中构型参数的最优或次优布置的技术,由此充分展示结构构型所带来的力学特

收稿日期: 2018-05-16; 录用日期: 2018-08-10; 网络出版时间: 2018-08-24 16:28

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180821.0902.001.html

* 通信作者. E-mail: lichao7715@163.com

引用格式: 李超, 金福艺, 张卫浩. 航空发动机转子结构布局优化设计方法[J]. 北京航空航天大学学报, 2019, 45(2): 266-276.

LI C, JIN F Y, ZHANG W H. Optimized design method of aero-engine rotor structure layout [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 266-276 (in Chinese).

征的变化。简言之,即是通过建立转子构型和其力学特征之间的关联性,进而揭示结构构型和尺寸效应所引起的力学特性的变化,同时附加先进的优化算法,实现转子结构与力学特征一体化设计。结构布局设计可以为航空燃气轮机总体结构初始设计阶段进行结构构型和尺寸控制,提供设计指导的理论依据。本文以典型的高推重比涡扇发动机高压转子结构为例,论述布局优化设计的流程、布局设计方法的可靠性及布局设计的重要意义。

结构布局优化设计可以为结构改型提供指导。图1为2种典型的高推重比涡扇发动机高压转子结构,二者均为1-0-1支承方案。结构I为7级压气机和1级涡轮,前2级压气机为悬臂结构,靠近涡轮处存在一个封沿盘。结构II为5级压气机和1级涡轮,无悬臂结构,封沿盘靠近压气机。通过这2种典型结构构型,可以衍生出其他类型高压转子结构。例如,同结构I相比,F110高压转子总体上是将压气机增加2级,并在压气机后增加封沿盘;结构II则是不存在悬臂的一类结构。以上这些转子结构的改型设计均可以在布局优化设计中得到完整的体现。当然,每家航空发动机集团在结构方案的选择上都有着较强的继承性,有着各自不同的设计风格。在布局优化设计时,应充分考虑传统工程经验和技术的积累,在此基础上进行布局的优化设计,这一点在布局设计中是需要充分考虑的。

图1中,实线勾勒出了转子大概的结构布局线路,盘和叶片等转子构件按照一定的规则分布其上,实心圆点代表构件的集中质量。值得注意

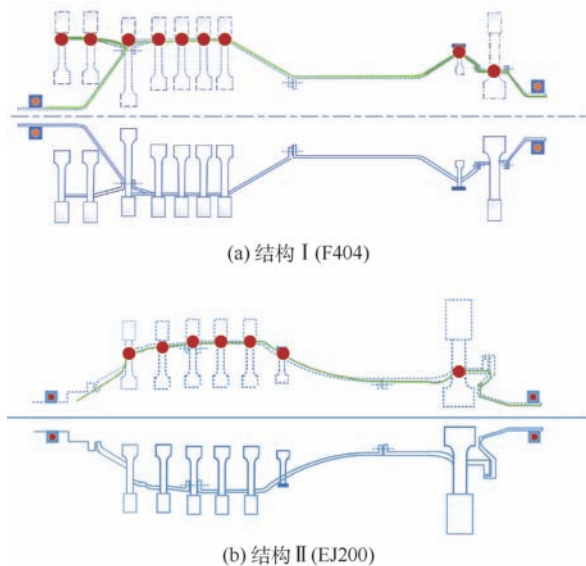


图1 典型转子结构

Fig. 1 Typical rotor structure

的是,转子的结构布局优化设计是对整体构型进行优化设计,并不涉及具体构件的尺寸设计。例如,涡轮盘的构型属于强度设计问题,不在转子结构布局设计范畴。

目前,转子的力学特性计算多为传递矩阵法及有限元法,考虑到传递矩阵法程序的通用性和可扩展性均较差,较难在工程中广泛使用,本文优化示例过程均采用有限元法进行计算,所建立的有限元模型如图2所示。盘和叶片的等效分布质量和转动惯量已经加到相应的位置处(mass21单元),如图2中的A、B、C、D、E、F、G所指。前、后支承位置在抗变形能力优化设计时为简支约束,在力学环境适应能力优化设计时采用弹性支承(combin214轴承单元),其支承刚度如表1所示。其余构型全部采用三维实体solid189高阶单元,不计阻尼。

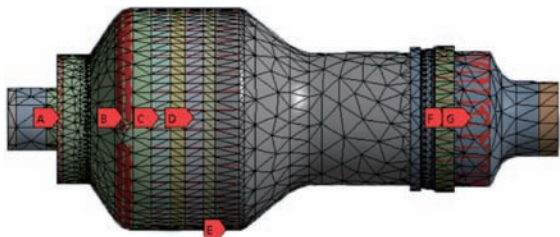


图2 有限元模型

Fig. 2 Finite element model

表1 支承刚度

Table 1 Supporting stiffness 10^7 N/m

支点	刚度
前支点	2.5
后支点	25

2 关键特征参数及其确定方法

关键特征参数的探索是布局设计的首要任务。所谓关键特征参数,是指对转子的力学性能(刚度特征和模态特征等)具有明显影响的构型参数。由于很难把握构型参数相互之间的约束规律,对全体构型参数进行优化设计是不现实的,同时,在后续试验设计(DOE)研究中,设计点的数量会随着输入参数的数量增加而迅速增加,严重降低DOE计算效率。因此,首先需要以一个较优的转子构型参数为基础(可直接由工程实际确定),对各个参数之间进行关联性研究,确定对转子力学性能有明显影响的关键特征参数。图3为转子结构布局的构型尺寸工程图,由于结构的构型尺寸及约束条件参数繁多,图中仅标出了待优化的部分参数。其中, A_1 为前轴颈与水平线的倾

角, R_1 为其倒角半径; A_2 为中间轴段与水平线的倾角, R_2 为其倒角半径, 为方便叙述, 各构型参数在本文里均用其代号表示, 对应位置已标于图 3 中。原始构型参数归纳如表 2 所示。

统计学中的参数关联性研究有多种方法^[9-10], 已广泛应用于优化设计。本文通过拉丁超立方 (Latin hypercube) 法布置样本点, 使用斯珀曼 (Spearman) 法进行数据之间的关联性分析。斯珀曼法是在基于蒙特卡罗关系下, 检查 2 个变量之间的单调关系, 比线性关系限制较少, 具有较高的准确度。通过该方法从众多构型变量中选出较少的参数作为关键特征参数, 这些参数也是后续待优化的设计变量。

得到关键特征参数的目的之一是进行 DOE^[11]。DOE 是研究正确的设计试验计划, 分析试验数据的理论和方法, 通过改变过程的输入因素, 观察其相应的输出响应变化, 从而获取关于此过程的知识, 确定各个输入因素的重要性以及各输入因素如何影响输出响应, 并如何达到最优化的目的。DOE 类型分为很多种, 常见的如稀疏网格设计^[12]、Box-Behnken 试验设计^[13]、田口静态试验设计^[14]、拉丁超立方试验设计^[15]等。DOE 是获取响应面的前提条件, 合适的 DOE 方法不仅

可以提高计算效率, 而且可以有效提高响应面的拟合精度。为了最大程度提高响应面拟合精度, 本文采用稀疏网格响应面法。稀疏网格响应面是一种自适应响应面, 可以对数据点进行自动细化, 精度很高。当然, 每一次细化都需要对构型参数改变的有限元模型进行重新启动分析, 这是以牺牲计算效率为代价的。

转子系统结构效率^[16]内容主要包括承载能力、抗变形能力和力学环境适应能力三方面。承载能力属于强度优化设计, 不属于布局优化设计范畴。布局优化设计主要是针对转子系统的抗变形能力和力学环境适应能力进行构型的寻优设计。由于本文目的仅为阐述转子结构布局优化设计的方法、方法可靠性和重要意义, 不针对某型号的发动机转子, 不针对具体的工程结构, 另外考虑到优化过程中, 为了便于寻找各构型参数对某个力学特征是否存在影响及影响的程度, 因此本文将抗变形能力的寻优目标参数只定义了惯性载荷下的最小等效刚度, 力学环境适应能力的寻优目标参数只定义了一阶弯曲共振时的转子转速, 同时兼顾总质量的变化情况。本文给出了工程综合仿真示例。值得注意的是, 寻优目标参数应该在结构效率的定义范围内, 需要依据实际情况而变化。

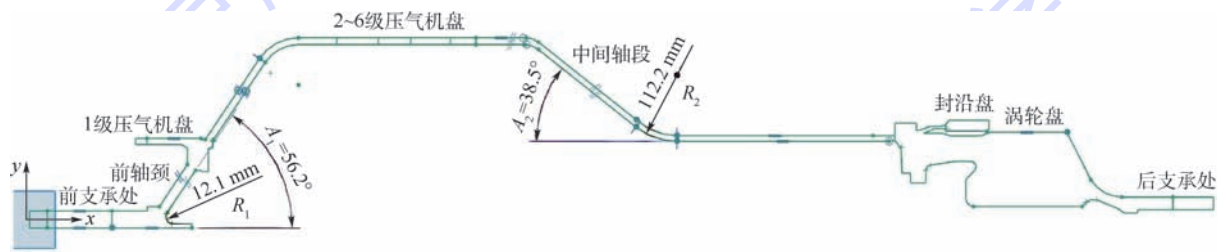


图 3 布局尺寸

Fig. 3 Layout size

表 2 待优化的原始构型参数

Table 2 Original configuration parameters for subsequent optimization

参数	数值
$A_1 / (^\circ)$	56.2
$A_2 / (^\circ)$	38.5
R_1 / mm	12.1
R_2 / mm	112.2

3 基于响应面法的结构布局优化设计方法流程

航空发动机转子系统结构布局优化设计方法流程如图 4 所示。

步骤 1 建立原始 CAD 参数化模型, 并初步

判断对结构力学特征有明显影响的构型参数作为待优化的设计变量, 这是依据技术积累和工程经验得出的。建立模型时应注意图形元素之间的约束条件, 避免构型参数改变时, 转子结构模型的严重失真, 否则将导致结构布局优化设计的失败。

步骤 2 建立有限元模型, 进行有限元求解, 指定优化目标 (如结构效率相关评估参数), 将有限元求解器与 CAD 关联。这里主要是为了使构型参数改变时, 利用 CAD 快速建模和利用有限元求解器快速求解, 以期实现迭代计算。

步骤 3 参数关联性分析, 确定关键特征参数。从步骤 1 选取的待优化设计变量中, 进一步确定哪些是对力学特征有显著影响的构型参数, 作为关键特征参数, 这是后续优化迭代中的设计变量。

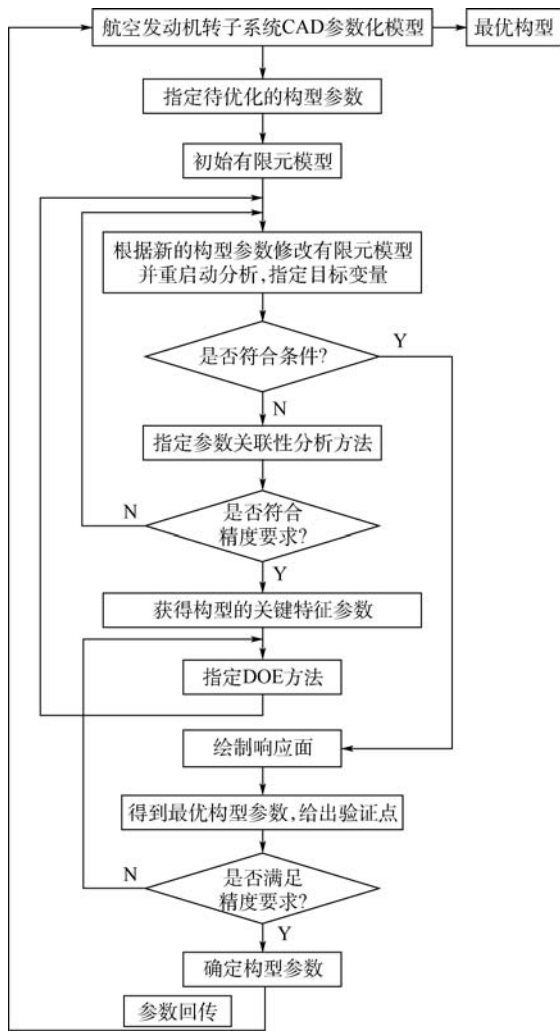


图 4 结构布局优化设计方法流程

Fig. 4 Flow of structure layout optimization design method

步骤 4 进行 DOE, 绘制响应面。确定响应面的拟合优度, 确保数据的准确性。

步骤 5 基于某些数学优化算法, 找到最优或次优的构型参数。

步骤 6 与原构型对比, 验证优化结果。

4 基于结构效率的结构布局优化设计

为了充分证明航空发动机转子系统结构布局优化设计方法的可行性, 本节以典型高推重比涡扇发动机高压转子构型(见图3)为例, 分别从抗

变形能力和力学环境适应能力两方面对转子结构进行结构布局优化设计。构型参数波动的上下限均为在原始构型的基础上 $\pm 10\%$ 。

4.1 抗变形能力

根据航空发动机转子系统的结构特征, 在设计中对转子系统抗变形能力的要求有: ①具有良好的横向弯曲刚度; ②盘轴连接处具有良好的角向刚度。在转子系统刚度设计中, 常采用的大比刚度材料有利于提高结构的抗变形能力, 但是材料的选取通常根据强度设计准则来确定, 对于给定的材料, 通过转子轴截面结构形状和面积的优化, 可以进一步提高转子系统的刚度, 从而达到提高其抗变形能力的目的。虽然连接结构处存在的刚度损伤, 对转子结构的力学特征存在较大影响, 但是由于其力学特征计算的复杂性, 在结构构型设计初期的布局优化中一般不予考虑。

为此, 对转子结构横向施加 1 个重力加速度, 此时转子将发生弯曲变形, 计算整体转子结构横向变形的最大值 (TD_Max) 和转子结构的质量 (TM), 通过寻找合适构型的设计参数, 使得最大变形量和总质量最小。

基于斯珀曼^[17]法的参数关联性分析(进行了 100 组构型参数计算)可以得到如表 3 所示的各参数间的相关性矩阵, 绘制出相应的参数关联性云图, 如图 5 所示, 其中颜色的深浅表示关联度的强弱。

从图 5 中可以较为清晰地看出各参数之间的关联性强弱, 严重影响最大变形量(即转子结构的抗变形能力)的参数为高压转子构型参数 A_2 , 对最大变形量的线性关联系数为 0.998 7, 对总质量线性关联系数为 -0.954, 均接近 1, 明显高于其他值。其次, 构型参数 R_1 和 R_2 的变化对总体质量稍有影响。综上, 可以初步断定, 对转子系统抗变形能力具有显著影响的关键特征参数为 A_2 。

为了减少设计变量个数, 略去一些对力学特征影响不敏感的参数, 同时兼顾对实际转子力学特征的影响。因为 2 个倒角半径 (R_1 和 R_2) 只对

表 3 抗变形能力优化相关性矩阵

Table 3 Correlation matrix of anti-deforming ability optimization

参数	R_2	R_1	A_2	A_1	TD_Max	TM
R_2	1	0.006 8	-0.01	0.041 418	-0.017 436	-0.155
R_1	-0.006 80	1	0.007 8	0.095 346	0.010 111	-0.144
A_2	-0.014 19	0.007 8	1	-0.010 88	0.998 7	-0.954
A_1	0.041 418	0.095 3	-0.01	1	-0.001 296 3	0.094 6
TD_Max	-0.017 43	0.010 1	0.998	-0.001 29	1	-0.953
TM	-0.155 04	-0.145	-0.95	0.094 633	-0.953 4	1

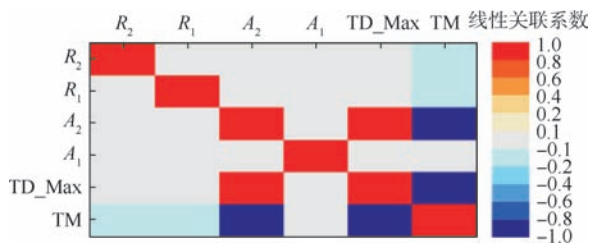


图5 抗变形能力优化参数关联性云图

Fig.5 Contour of parameter correlation about anti-deforming ability optimization

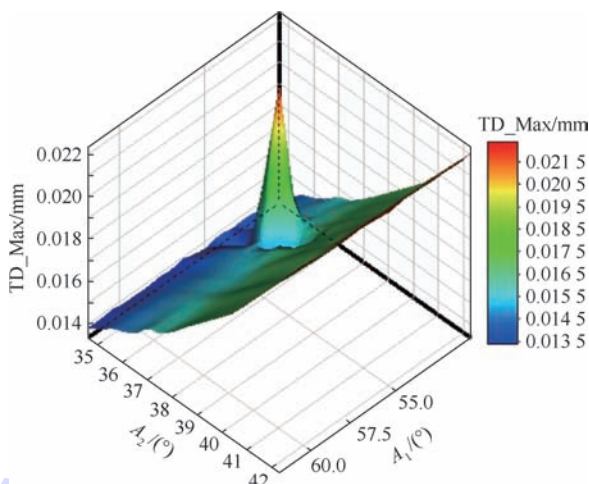
结构质量有少许影响,且影响程度大体相当,因此将这2个构型参数一起去掉。接下来只研究2个锥角(A_1 和 A_2)的变化对转子系统的抗变形能力和总质量的影响程度。

通过有限元迭代计算,可以得到总质量和最大变形量相对于2个锥角(A_1 和 A_2)的稀疏网格响应面,如图6所示。

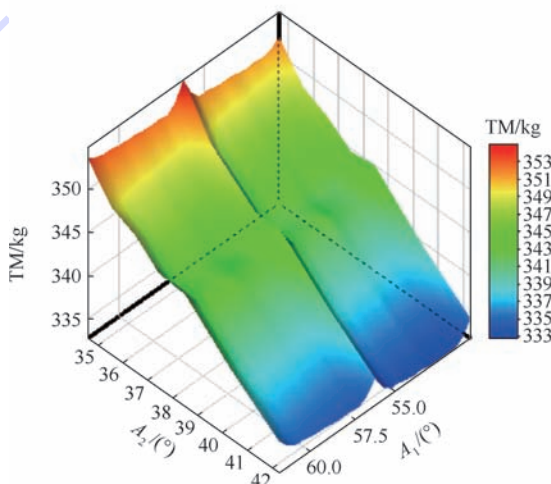
从响应面可以直观地找出使最大变形量和总质量最小时所对应的构型参数。图6(a)所示的最大变形量相对于2个锥角的响应面有着很高的“尖点”,存在迅速上升或下降的最大值点,一定程度上可以反映出,假如构型参数选取的不合理,可能会对转子系统的力学特性产生严重影响。为了验证响应面的准确性,给出了5个验证点,如表4所示。

从随机选取的5个验证点所统计出的相对误差来看,该响应面拟合度很高,几乎零误差,可以为下一步优化提供数据支持。响应面拟合程度的优劣直接影响后续优化结构,影响结构的可靠性,若误差统计数据存在较大的值,则应以更多的数据点对响应面进行细化,直到误差在允许的范围为止。

通过响应面的优化是基于某些先进算法实现,如多目标遗传算法^[18]、自适应多目标^[19]等,相比于有限元计算耗时很少。遗传算法是一种通过模拟自然进化过程搜索最优解的方法,在优化方法中具有独特的优越性,通过前面得到的响应面,基于多目标遗传算法最终得到3组候选设计



(a) 最大变形量响应面



(b) 总质量响应面

图6 抗变形能力优化响应面

Fig.6 Response surface of anti-deforming ability optimization

点,如表5所示,表中“☆”表示优化后总质量最小的一组参数。

从表5可以看到,若以最大变形量最小为优化目标,应选择第3组构型参数,即 $A_1 = 50.646^\circ$, $A_2 = 37.806^\circ$,此时具有较强的抵抗惯性载荷的能力。若以总质量最小为优化目标,应选择第1组构型参数,即 $A_1 = 51.278^\circ$, $A_2 = 38.531^\circ$ 。若折中选取,则应选择第2组构型参数。从表5中的数

表4 抗变形能力响应面误差验证

Table 4 Error verification of response surface of anti-deforming ability

编号	优化变量		验证点(有限元计算)		响应点(响应面计算)		相对误差/%	
	$A_1/(^\circ)$	$A_2/(^\circ)$	TD_Max/mm	TM/kg	TD_Max/mm	TM/kg	TD_Max	TM
1	61.162	40.402	0.019226	340.94	0.019409	347.24	0.95	1.8
2	58.68	40.591	0.019436	336.93	0.019601	341.92	0.85	1.5
3	52.027	39.633	0.018195	343.12	0.018297	351.98	0.56	2.6
4	52.043	41.095	0.020178	335.41	0.020183	340.56	0.02	1.5
5	57.758	39.611	0.01813	350.49	0.018293	355.54	0.90	1.4

表 5 抗变形能力优化设计结果

Table 5 Design results of anti-deforming ability optimization

编号	优化变量		验证点(有限元计算)		响应点(响应面计算)		相对误差/%	
	$A_1/(^\circ)$	$A_2/(^\circ)$	TD_Max/mm	TM/kg	TD_Max/mm	TM/kg	TD_Max	TM
1	51.278	38.531	0.016908	345.57	0.016859	346.91 [*]	0.29	0.4
2	57.409	37.955	0.016546	361.77	0.016267	350.65	1.69	3.1
3	50.646	37.806	0.016264	360.55	0.016174 [*]	351.24	0.55	2.6

据也可以发现,虽然前后锥角仅有 1° 左右的变化,却可以引起最大变形量和总质量的显著改变,特别是质量的变化,大约 4 kg。

考虑第 1 组误差较小,若选取第 1 组设计点,与原始构型对比,质量减少 3.59 kg,在惯性载荷作用下,最大变形量减少 0.004913 mm,如图 7 所示。值得注意的是,本算例计算的是在 1 个重力加速度下的变形量,若在更大的惯性载荷作用下,优化效果会更加明显。

由此可得,经过结构布局优化设计后的转子结构,在惯性力下的最小等效刚度为

$$k_{op} = \frac{346.91 \times 9.8}{0.016921 \times 10^{-3}} = 2 \times 10^8 \text{ N/m}$$

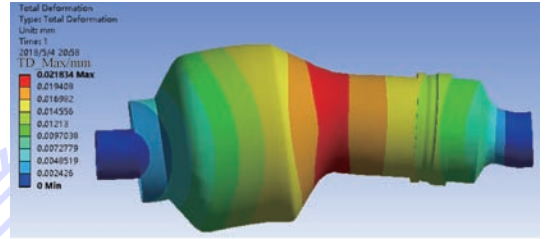
优化前的等效刚度为

$$k = \frac{350.5 \times 9.8}{0.021834 \times 10^{-3}} = 1.56 \times 10^8 \text{ N/m}$$

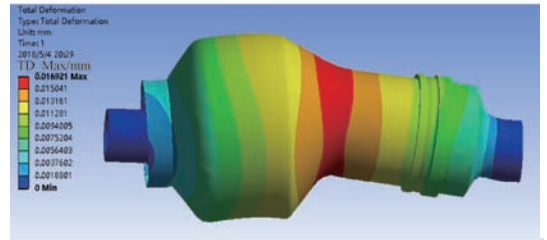
优化前后相比提高了 $4.4 \times 10^7 \text{ N/m}$,整体提升了近 28% 的惯性刚度,可见转子结构布局优化设计的显著效果。

4.2 力学环境适应能力

转子系统的工作转速相对于共振频率之间的稳定裕度一直是结构设计需要考虑的重要问题,在航空发动机转子结构设计初期,常因为结构布局设计的不合理而达不到裕度要求(一般为 20%)。例如,高压转子在工作转速范围内往往存在 2 个共振转速,较低的共振转速是由静不平衡引起的,较高的共振转速是由力偶不平衡引起的。工作转速距离较近的一阶弯曲振型和俯仰振型均应有一定的裕度,这就需要将一阶弯曲共



(a) 优化前(总质量350.5 kg,最大变形量0.021 834 mm)



(b) 优化后(总质量346.91 kg,最大变形量0.016 921 mm)

图 7 抗变形能力优化前后对比

Fig. 7 Comparison of results before and after anti-deforming ability optimization

振频率最大化设计。工程上一般通过调节支承处刚度来调节共振转速,但若是构型设计严重不合理,单凭调节支承刚度是达不到要求的。

仍以图 3 所示的高压转子结构构型为例,以一阶弯曲共振频率最大、振幅最小且结构总质量最小为优化目标,通过对 4 个构型参数(见表 2)的选取使结构趋于“合理化设计”。结构布局优化流程同第 3 节,基于斯珀曼法可以得到参数关联性矩阵(见表 6)及对应的云图(见图 8),观察可知,关键特征参数依旧是 A_1 和 A_2 。

通过有限元迭代计算,可以得到结构总质量、

表 6 力学环境适应能力优化相关性矩阵

Table 6 Correlation matrix of mechanical environment adaptability optimization

参数	R_2	R_1	A_2	A_1	TD_Max	TM	C_S
R_2	1	0.024	-0.050	0.0006	-0.37273	-0.17	0.050649
R_1	0.024	1	0.0038	-0.012	-0.12857	0.175	-0.12987
A_2	-0.05	0.003	1	-0.007	-0.58961	-0.94	-0.96104
A_1	0.006	-0.01	-0.007	1	-0.46364	0.042	-0.17143
TD_Max	-0.37	-0.12	-0.589	-0.463	1	0.679	0.7039
TM	-0.17	0.175	-0.949	0.0428	0.67922	1	0.92727
C_S	0.051	-0.13	-0.961	-0.171	0.7039	0.927	1

注:C_S—临界转速。

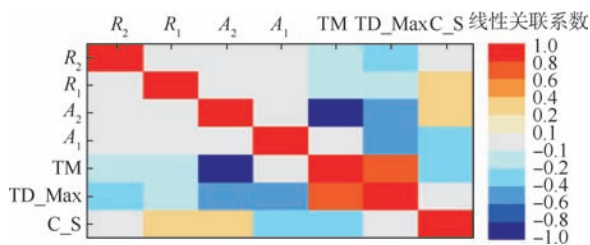


图8 力学环境适应能力优化参数关联性云图

Fig. 8 Contour of parameter correlation about mechanical environment adaptability optimization

一阶弯曲临界转速和一阶弯曲共振最大变形量相对于2个锥角(A_1 和 A_2)的稀疏网格响应面,如图9所示。

响应面误差验证如表7所示。可以看出,响应面误差较小,拟合精度较高,可以为后续优化提供数据支持。基于多目标遗传算法,最终得到了3组候选设计点,如表8所示。

从表8可以看到,优化结果误差较小,第2组设计点的总质量和转子在一阶弯曲共振时的最大变形量较小,第3组的一阶弯曲临界转速最大,可以考虑选择这2组设计点。若选择第3组为设计点,优化前后临界转速如图10所示。

可以看出,相比于优化前,一阶弯曲临界转速增加1372 r/min,但是这是以质量和一阶弯曲共振幅值的增加为代价的,其中质量增加1.51 kg,一阶弯曲共振最大变形量增加0.024 mm。在实际工程中,可以根据要求进行取舍。

需要说明的是,工程中转子结构动力学优化设计包含着更多的内容。例如,在航空发动机转子结构设计中,转子结构的平衡配重是在有限个转速下进行的,因此在飞行包线内仍存在一定的不平衡响应,这时为了保证转子具有足够的抗变形能力,需要结构对不平衡量的敏感性降到最低,即结构对不平衡量的低敏感性设计。例如,低压转子的动挠度曲线形状和弯曲程度随转速而变化,当转速接近某阶临界转速时,该阶阵型分量最大。如果轴承支点位置布置不合理,距离振型节点较远,将会大大增加轴承的支反力,转子系统可靠性急剧降低,这就需要通过调整局部构型以期改变振动振型,从而达到降低轴承支反力的目的。本文所涉及的构型优化仅是为了阐述布局优化设计的实现方法、可行性和重要意义,并不针对具体型号的发动机转子及其出现的具体问题做分析,所以计算结果可能与实际发动机转子存在偏差。

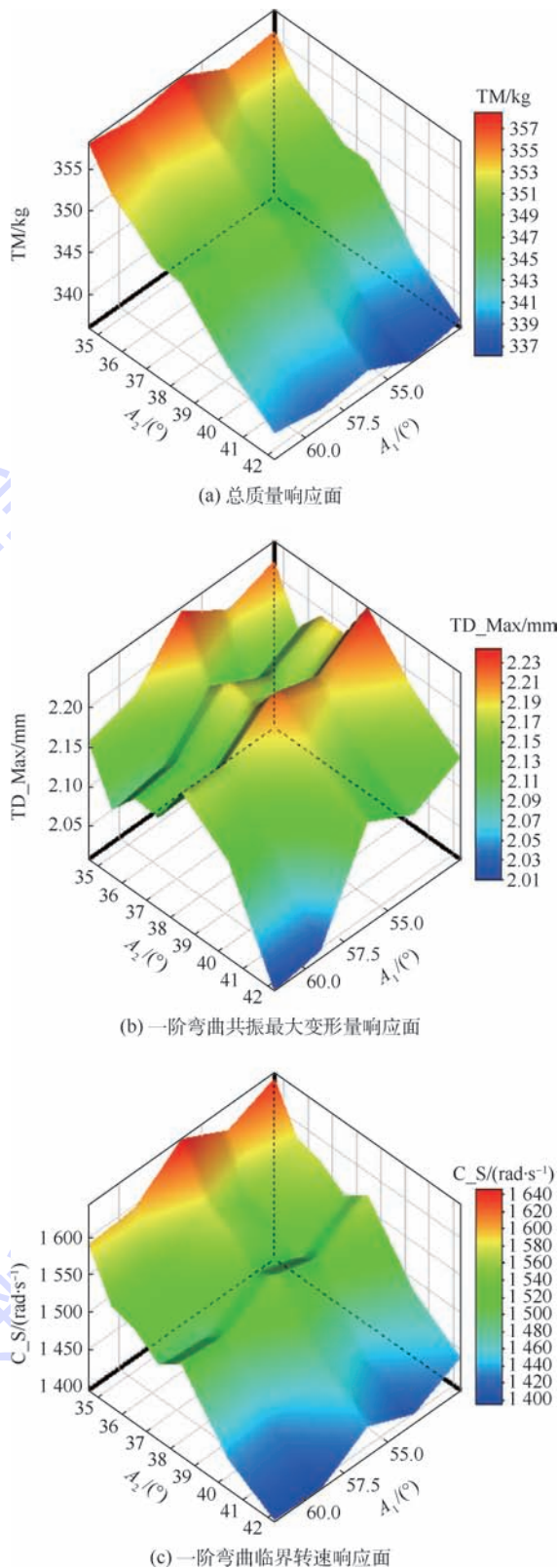


图9 力学环境适应能力优化响应面

Fig. 9 Response surface of mechanical environment adaptability optimization

同时,为了节省篇幅和简化计算,略去了工程上其他诸多方面的布局优化仿真。

表7 力学环境适应能力响应面误差验证

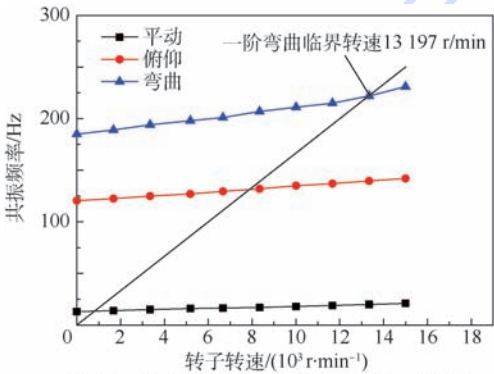
Table 7 Error verification of response surface of mechanical environment adaptability

编号	优化变量		验证点(有限元计算)			响应点(响应面计算)			相对误差/%		
	A ₁ /(°)	A ₂ /(°)	TD_Max/mm	TM/kg	C_S/(r·min ⁻¹)	TD_Max/mm	TM/kg	C_S/(r·min ⁻¹)	TD_Max	TM	C_S
1	52.929	36.22	2.1446	353.95	14 543	2.1548	350.05	14 769	0.48	1.10	1.55
2	52.953	40.724	2.1754	342.91	13 536	2.1422	339.45	13 812	1.53	1.01	2.04
3	59.483	36.251	2.1038	354.49	14 146	2.1234	351.71	14 555	0.93	0.78	2.89
4	59.015	40.384	2.1855	344.37	13 391	2.1105	341.85	13 672	3.43	0.73	2.10
5	50.621	40.172	2.113	343.94	14 662	2.186	341.87	14 255	3.45	0.60	2.78

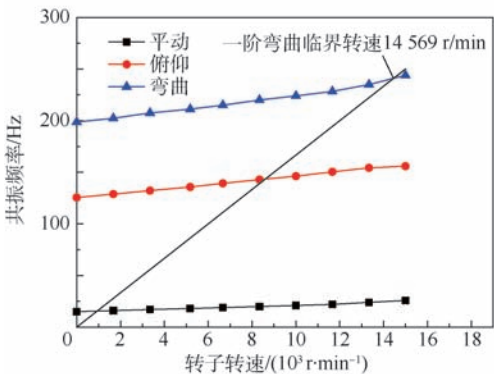
表8 力学环境适应能力优化设计结果

Table 8 Design results of mechanical environment adaptability optimization

编号	优化变量		验证点(有限元计算)			响应点(响应面计算)			相对误差/%		
	A ₁ /(°)	A ₂ /(°)	TD_Max/mm	TM/kg	C_S/(r·min ⁻¹)	TD_Max/mm	TM/kg	C_S/(r·min ⁻¹)	TD_Max	TM	C_S
1	60.774	42.296	2.1357	340.04	13 620	2.0132	338.79	13 342	5.7	0.37	2.04
2	59.369	42.266	2.1241	339.97	13 662	2.0198*	338.29*	13 335	4.91	0.49	2.39
3	56.195	41.615	2.1597	341.37	14 307	2.0806	343	14 569*	3.66	0.48	1.83



(a) 优化前(最大变形量2.1357 mm,总质量339.86 kg,临界转速13 197 r/min)



(b) 优化后(最大变形量2.1597 mm,总质量341.37 kg,临界转速14 569 r/min)

图10 力学环境适应能力优化前后对比

Fig. 10 Comparison of results before and after mechanical environment adaptability optimization

5 工程结构综合优化示例

以两支点高速转子结构系统为背景,为了建立具有通用性的优化流程,掌握结构几何构形和尺寸变化对转子力学特性的影响规律,本节以转

子结构的结构效率为目标,对其进行多目标综合优化设计,使本文提出的结构布局设计方法更具工程应用价值。

对高压转子初始施加国军标所要求的横向过载载荷和机动飞行的角加速度,高压转子工作转速为12 000 r/min,通过有限元法综合考虑结构的抗变形能力和力学环境适应能力,优化的目标是控制转子系统结构总质量的前提下,使转子结构效率损失降到最低。定义结构布局优化的数学表达式如下:

$$\begin{cases} \min \begin{cases} \frac{1}{\sigma_{ave}(\mathbf{X}, \mathbf{f})} \\ w(\mathbf{X}, \mathbf{f}) \\ S_p \\ \frac{1}{S_B} \end{cases} & \mathbf{X} = [x_1, x_2, \dots, x_n] \\ g_j(\mathbf{Y}) \leq 0 & j = 1, 2, \dots, n \end{cases} \quad (1)$$

式(1)即为目标函数。式中: \mathbf{X} 为关键设计参数组成的参数向量; \mathbf{f} 为环境载荷; $w(\mathbf{X}, \mathbf{f})$ 为径向挠曲变形; S_p 为转子俯仰振型对应的临界转速; S_B 为一阶弯曲对应的临界转速; \mathbf{Y} 为结构力学特征参数; $\sigma_{ave}(\mathbf{X}, \mathbf{f})$ 为外加环境载荷下结构的平均应力系数,表达式如下:

$$\sigma_{ave} = \frac{\sum_{i=1}^n \sigma_i m_i}{\sum_{i=1}^n m_i} \quad (2)$$

其中: σ_i 为第*i*个结构单元的应力; m_i 为第*i*个单元的质量。平均应力系数反映了材料的使用效率。

式(1)是一个多参数、多目标的优化问题,也

可以在结构效率的意义下,将其合并成一个参数,定义为结构效率损失系数:

$$\begin{cases} \min \left[a \frac{1}{\sigma_{ave}(X, f)} + bw(X, f) + cS_p + d \frac{1}{S_B} \right] \\ X = [x_1, x_2, \dots, x_n] \\ g_j(Y) \leq 0 \quad j = 1, 2, \dots, n \end{cases} \quad (3)$$

表9 参数关联性矩阵

Table 9 Parameter correlation matrix

参数	A_1	A_2	R_1	R_2	TD_Max
A_1	1	0.275	-0.0286	-0.125	0.214286
A_2	0.275	1	-0.0321	-0.0035	0.78571
R_1	-0.028571	-0.032143	1	-0.1714	0.20357
R_2	-0.125	-0.003571	-0.1714	1	-0.04642
TD_Max	-0.214286	0.78571	0.20357	-0.0464	1

基于有限元法,采用稀疏网格响应面对试验点数据拟合,得到如图11所示的响应面,图中S_E为结构效率损失系数。基于多目标遗传算法,通过响应面可以较为容易地找到极小值点 $(A_1, A_2) = (49.6^\circ, 37.5^\circ)$,该点即是优化后的最优构形点。优化前后的总变形如图12所示。

按照式(3)定义的结构效率损失系数为目标函数,优化后的结构效率损失系数(0.646)同优化前(0.981)相比,下降0.335,使结构总质量基本不变的条件下,转子刚度特性大幅提高。从参数关联性矩阵也可以看出,中间轴段倾角(A_2)对转子刚度特性产生较大影响,合理的倾角选取可以有效提升结构的抗变形能力。从总体变形云图(见图12)上也可以看出,不仅结构效率损失系数有所下降,而且整体变形的分布形式并没有较大变化,这对结构整体性能提升是有意义的。

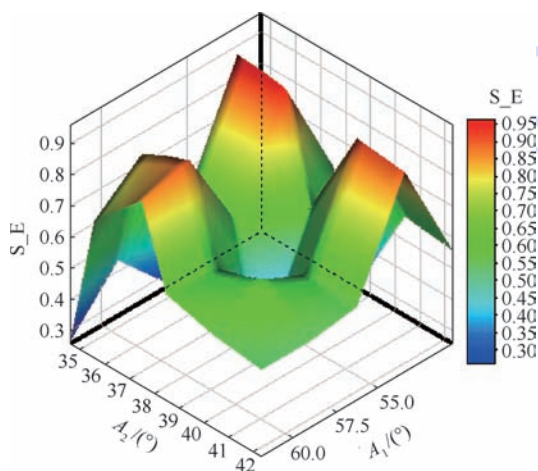
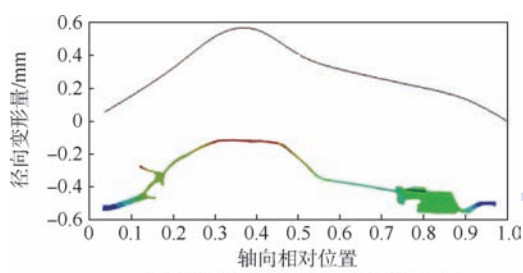


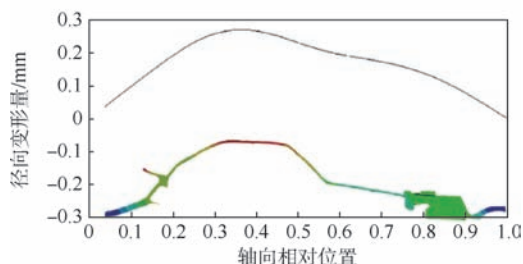
图11 拟合响应面

Fig. 11 Fitting response surface

式中: a 、 b 、 c 、 d 为加权系数,可根据工程实际进行适当选取。可把结构效率损失系数作为结构合理性的度量。以此为目标函数,首先基于斯珀曼法的参数关联性分析得到如表9所示的参数关联性矩阵。可以看出,对结构效率损失系数影响较大的2个构形参数为 A_1 和 A_2 ,故选定 A_1 和 A_2 作为关键特征参数。



(a) 优化前(结构效率损失系数0.981)



(b) 优化后(结构效率损失系数0.646)

图12 优化前后结构总体变形

Fig. 12 Overall structure deformation before and after optimization

6 结论

本文提出的航空发动机转子结构布局设计方法是为适应新一代先进航空发动机转子结构研制顺势而出,是摆脱过多依靠传统设计经验,实现其先进性的有效途径,该方法不仅适合涡扇发动机,对涡轴、涡桨等同样适用。同时,转子结构布局设计方法补充了原转子结构效率的内容,在对转子结构力学特征定量评估的基础上,提供了最优或次优的结构布局易行方法。

1) 航空发动机转子构型的合理设计,特别是中间轴段的合理设计,可以有效提升转子的抗变

形能力和力学环境适应能力。算例结果表明,通过对结构构型的布局优化设计,与原始结构相比,抗变形能力优化设计使高压转子整体提升了近 28% 的惯性刚度,力学环境适应能力优化设计使转子系统的一阶弯曲临界转速提高 1 372 r/min,同时给出了多目标综合优化设计方法,最大程度降低了结构效率损失系数。通过对转子结构布局优化设计,完全可以达到将不合理的结构构型最大程度合理化的效果,这对航空发动机转子系统结构改进设计,乃至整机结构布局优化设计具有重要的借鉴意义。

2) 在航空发动机转子结构布局优化设计时,除了要考虑优化目标达到技术要求外,同时需要兼顾其他力学特征是否发生严重改变,是否依然符合技术要求。例如,本文在力学环境适应能力优化时,虽然一阶弯曲临界转速实现了 1 372 r/min 的增加,但是结构总质量也和一阶弯曲时的最大变形量均有少许增加。

本文仅以航空发动机高压转子构型为例,探索其结构关键特征参数,并对转子结构抗变形能力和力学环境适应能力进行了布局优化设计。实际工程中,可能需要考虑更多的实际环境因素,如高温、高压、流固耦合等复杂环境下对航空发动机双转子系统和整机的结构布局优化设计,这仍有待进一步的研究。

参考文献 (References)

- [1] 郑华强,彭刚,马艳红,等. 航空发动机结构力学性能定量分析方法[J]. 推进技术,2018,39(3):645-652.
ZHENG H Q, PENG G, MA Y H, et al. Quantitative analysis method for mechanical characteristics of structure system in aero-engine [J]. Journal of Propulsion Technology, 2018, 39 (3): 645-652 (in Chinese).
- [2] ARNE S, JULIAN B, SASCHA K, et al. Conceptual investigation of a propulsive fuselage aircraft layout [J]. Aircraft Engineering and Aerospace Technology, 2014, 86(6):464-472.
- [3] LUO Q, WEN Z H. Application of garden design style in Tang dynasty to the design of modern city public gardens: A case study of Tang paradise [J]. Journal of Landscape Research, 2018, 10(2):11-14.
- [4] LIU S, QIAO H. Topology optimization of continuum structures with different tensile and compressive properties in bridge layout design [J]. Structure and Multidisciplinary Optimization, 2011, 43(3):369-380.
- [5] GUTTA P R, CHITHALA V S, MANCHOJU R V, et al. A review on facility layout design of an automated guided vehicle in flexible manufacturing system [J]. Materialstoday: Proceedings, 2018, 05(2):3981-3986.
- [6] WANG Y K, LIAO M F, ZHANG J H. Dynamic load reduction design for inter-shaft bearing of aircraft engine [J]. Journal of Aerospace Power, 2017, 32(2):492-499.
- [7] FU C, REN X M, YANG Y F, et al. An interval precise integration method for transient unbalance response analysis of rotor system with uncertainty [J]. Mechanical System and Signal Processing, 2018, 107:137-148.
- [8] WHANG C, CHOY K L, BACK J. Development and experimental verification of counter-rotating dual rotor/dual generator wind turbine: Generating, yawing and furling [J]. Renewable Energy, 2017, 114(B):644-654.
- [9] 刘子君. 统计学 [M]. 北京:清华大学出版社,2017:148-156.
LIU Z J. Statics [M]. Beijing:Tsinghua University Press,2017:148-156 (in Chinese).
- [10] 鹏翔. 复杂产品设计中参数关联和等效简化方法激起应用 [D]. 杭州:浙江大学,2014:92-145.
PENG X. Method and its application of parameter correlation and equivalent simplification in complex product design [D]. Hangzhou: Zhejiang University, 2014:92-145 (in Chinese).
- [11] 闵亚能. 试验设计 (DOE) 应用指南 [M]. 北京:机械工业出版社,2011:5-17.
MIN Y N. Application guide of DOE [M]. Beijing: China Machine Press, 2011:5-17 (in Chinese).
- [12] XIONG F F, XIONG Y, GREENE S, et al. A new sparse grid based method for uncertainty propagation [C] // International Design Engineering Conference & Computers and Information in Engineering Conference. New York: ASME, 2009:1-11.
- [13] FERREIRA S L C, BRUNS R E, MATOS G D, et al. Box-Behnken design: An alternative for the optimization of analytical methods [J]. Analytica Chimica Acta, 2007, 597(2):179-186.
- [14] 文放怀. 田口方法 [M]. 广州:广东经济出版社,2006:43-101.
WEN F H. Taguchi method [M]. Guangzhou: Guangdong Economic Press, 2006:43-101 (in Chinese).
- [15] HELTON J C, DAVIS F J. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems [J]. Reliability Engineering & System Safety, 2003, 81(1):23-69.
- [16] 马艳红,陈璐璐,张大义,等. 航空发动机转子系统结构效率评估参数及计算方法 [J]. 航空动力学报, 2013, 28(7):1598-1606.
MA Y H, CHEN L L, ZHANG D Y, et al. Assessment parameters and calculation methods of structural efficiency on rotor system in aero engine [J]. Journal of Aerospace Power, 2013, 28 (7): 1598-1606 (in Chinese).
- [17] OSCAR L, OLVERA A, BRUNO D. Population models and simulation methods: The case of the spearman rank correlation [J]. Mzthematical and Statistical Psychology, 2017, 70(6):347-367.
- [18] 玄光男. 遗传算法与工程优化 [M]. 北京:清华大学出版社, 2004:76-88.
XUAN G N. Genetic algorithm and engineering optimization [M]. Beijing: Tsinghua University Press, 2004:76-88 (in Chinese).
- [19] 左益. 基于全局优化和局部学习的进化多目标优化算法 [D]. 西安:西安电子科技大学,2016:29-47.

ZUO Y. Evolutionary multi-objective algorithms based on global optimization and local learning[D]. Xi'an: Xidian University, 2016:29-47 (in Chinese).

金福艺 男, 硕士研究生。主要研究方向: 航空发动机转子动力学。

作者简介:

李超 男, 博士, 讲师。主要研究方向: 航空发动机结构设计、智能结构动力学与控制。

张卫浩 男, 硕士研究生。主要研究方向: 航空发动机静子结构承力系统设计。

Optimized design method of aero-engine rotor structure layout

LI Chao^{*}, JIN Fuyi, ZHANG Weihao

(School of Energy and Power Engineering, Beihang University, Beijing 100083, China)

Abstract: Structure is the comprehensive embodiment of aero-engine function, performance, and reliability design level, and all technical requirements, performance indicators, strength indicators, or structural safety and reliability should be based on reasonable structure layout design. This paper puts forward the viewpoint of the layout and optimization about the aero-engine rotor structure. Taking the structure configuration of the high-pressure rotor of a typical turbofan engine as an example and based on the response surface method of the design of experiment (DOE), the finite element method is used to calculate through multi-objective genetic optimization algorithm. Relevant optimization calculations are carried out from the aspects of deformation resistance, mechanical environment adaptability and rotor structural efficiency respectively. It is demonstrated that reasonable structure layout can greatly enhance the mechanical properties of rotors. The research method is of great guiding significance for the initial structure layout design of the aero-engine rotor system, which can significantly reduce the number of iterations and shorten the design cycle.

Keywords: aero-engine; rotor system; structure layout; configuration design; structural efficiency

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

DOI: 10.13700/j.bh.1001-5965.2018.0295

一种考虑过滤的短纤维增强复合材料 RVE 建模方法

刘丰睿^{1,*}, 骈璐¹, 赵丽滨¹, 张建宇²

(1. 北京航空航天大学 宇航学院, 北京 100083; 2. 重庆大学 航空航天学院, 重庆 400044)

摘 要: 代表性体积单元(RVE)法是短纤维增强复合材料性能预测的常用方法,其RVE生成效率是预测效率的重要影响因素。针对现有的随机顺序吸附(RSA)方法生成RVE时,布尔运算次数多、效率低的问题,通过在布尔运算前加入对纤维形心距的判断,过滤掉一部分随机生成的且与已有纤维相交的纤维,以减少布尔运算次数,提出了一种考虑过滤的随机顺序吸附(FRSA)方法。通过将改进后的FRSA方法在不同RVE参数和方法参数下生成RVE所需布尔运算次数和所需时间与基于布尔运算的随机顺序吸附(BORSA)方法进行比较,证明了FRSA方法的先进性。

关键词: 短纤维增强复合材料; 数值建模; 随机顺序吸附(RSA)方法; 过滤法; 纤维形心距

中图分类号: TB332

文献标识码: A **文章编号:** 1001-5965(2019)02-0277-06

短纤维增强复合材料具有比强度高、比刚度高和可设计性强等特点,在航空航天^[1-2]、建筑^[3]、汽车^[4]、医疗^[5]、船舶^[6]等领域应用广泛。其弹性模量的预测是重要的研究内容,主要的预测方法有均匀化法^[7]、代表性体积单元(Representative Volume Element, RVE)法^[8-10]等。其中,RVE法考虑因素多,能全面预测材料特性,是常用的预测方法。高效地建立RVE是进行材料性能预测的第一步,也是研究难点。

目前建立短纤维增强复合材料RVE的方法有:随机顺序吸附(Random Sequential Adsorption, RSA)方法^[10-12]、基于RSA的纤维生长法^[13]、分子动力学方法^[14]、粒子群优化算法^[15]、蒙特卡罗方法^[16]和图像重建技术^[17]等。其中,RSA方法凭借其操作简单、计算量小的优点被广泛应用。该方法的核心是每次在RVE区域内生成位置随

机、取向按给定函数随机分布的一根纤维后,判断新纤维与已存在纤维是否相交,仅当不相交时认为新纤维可以存在于RVE中,循环此过程直到纤维体积分数达到规定值为止^[10]。纤维相交与否关系的判断通常通过几何方法实现^[18]。考虑到几何方法计算复杂且计算量大,Liu等^[19]采用布尔运算方法判断纤维是否相交,进而提出了基于布尔运算的随机顺序吸附(Boolean Operation based Random Sequential Adsorption, BORSA)方法。但是,布尔运算判断法耗时也较长,并且对一些非常容易判断位置关系的纤维进行布尔运算判断影响了RVE的建立效率。

本文对Liu等^[19]的方法进行了改进,在对纤维进行布尔运算判断前,加入了新纤维与已存在纤维的形心距计算,当形心距小于规定距离时不进行布尔运算,直接重新随机生成纤维位置,这相

收稿日期: 2018-05-22; 录用日期: 2018-07-28; 网络出版时间: 2018-08-01 11:16

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180731.1717.002.html

基金项目: 国家自然科学基金(11772028, 11702012, 11572058, 11372020)

* 通信作者. E-mail: frliu@buaa.edu.cn

引用格式: 刘丰睿, 骈璐, 赵丽滨, 等. 一种考虑过滤的短纤维增强复合材料RVE建模方法[J]. 北京航空航天大学学报, 2019, 45(2): 277-282. LIU F R, PIAN L, ZHAO L B, et al. A filter based modeling method of RVE for short fiber reinforced composite [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 277-282 (in Chinese).

当于过滤掉了部分随机生成的新纤维,进而提出了考虑过滤的随机顺序吸附(Filter based Random Sequential Adsorption, FRSA)方法。然后,采用RVE生成算例验证了方法的先进性。

1 BORSA 方法

Liu 等^[19]提出的 BORSA 方法如下。给定建模所需的纤维体积分数 V_c^{\max} , 纤维长径比 A_R , 纤维直径 D , 单根纤维体积 V_f 和 RVE 模型长、宽、高 (L, W, H) 等参数。定义描述纤维的 5 个参数: 纤维形心的三维坐标 (x, y, z) , 纤维轴线在 XOY 平面内的投影与 X 轴的夹角 θ , 纤维轴线与 Z 轴的夹角 φ 。给定纤维参数取值范围: $x \in (0, L)$, $y \in (0, W)$, $z \in (0, H)$, $\theta \in (0^\circ, 180^\circ)$, $\varphi \in (0^\circ, 180^\circ)$ 。规定最大布尔运算次数为 T_{\max} 。生成 RVE 中所有纤维的步骤如下:

1) 生成第 1 根纤维 $f_i (i = 1)$ 。随机生成符合均布函数的纤维形心位置参数 (x_1, y_1, z_1) , 随机生成符合相应函数的取向参数 (θ, φ) , 并在模型中生成纤维。

2) 计算 RVE 当前纤维体积分数 V_c , 当 $V_c < V_c^{\max}$ 时, $i = i + 1$, 执行步骤 3); 当 $V_c \geq V_c^{\max}$ 时, 输出纤维模型, RVE 建模结束。

3) 随机生成纤维 f_i 的参数 $(x_i, y_i, z_i, \theta_i, \varphi_i)$ 。设定布尔运算次数 $T = 0$ 。

4) $T = T + 1$, 将纤维 f_i 向 $(\pm L, 0, 0), (0, \pm W, 0), (0, 0, \pm H), (\pm L, \pm W, 0), (0, \pm W, \pm H), (\pm L, 0, \pm H), (\pm L, \pm W, \pm H)$ 共 26 个方向复制。然后执行布尔运算: 用 27 个新纤维 f_i 减现有的 $i - 1$ 根纤维, 当 $i - 1$ 根纤维的剩余体积小于 $(i - 1)V_f$ 时, 表明纤维间有相交关系, 进行步骤 5); 当 $i - 1$ 根纤维的剩余体积等于 $(i - 1)V_f$ 时, 认为纤维 f_i 与纤维 $f_j (j = 1, 2, \dots, i - 1)$ 均不相交, 进行步骤 6)。

5) 当 $T < T_{\max}$ 时, 将 f_i 平移到 RVE 内随机位置, 返回步骤 4); 当 $T = T_{\max}$ 时, 则认为已不能在当前模型中放入新纤维, RVE 生成失败, 建模结束。

6) 保存纤维 f_i 的信息, 返回步骤 2)。

由上述过程可见, 每生成一根纤维或对纤维进行一次平移, 均需进行一次布尔运算。由于纤维位置是随机生成的, 其中必然存在新生成的纤维形心位置与现有纤维形心非常接近的情况, 比如小于纤维直径, 此时的新纤维必然与现有纤维相交。如果通过设置过滤函数使得对这部分新纤维不需要进行布尔运算, 则可以提高 RVE 生成速度。

2 FRSA 方法

2.1 FRSA 方法原理

为了实现纤维过滤, 定义新生成纤维与现有纤维形心距最小值 d_i 为

$$d_i = \min(d_{ij}) \quad (1)$$

式中: d_{ij} 为纤维 f_i 和 f_j 的形心距离, 定义为

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \quad (2)$$

并在进行布尔运算前判断 d_i 是否小于一个常数 d_c (d_c 为临界距离)。当 $d_i < d_c$ 时, 认为不必进行布尔运算。当 $d_i \geq d_c$ 时, 认为可以进行布尔运算。从而实现过滤掉一部分纤维。

2.2 临界距离 d_c 对 FRSA 方法的影响

空间中 2 个纤维的距离 d_{ij} 存在 2 个特殊情况:

1) 当 $d_{ij} = D$ 时, 以 f_i 形心为球心的纤维内切球体与以 f_j 形心为球心的纤维内切球体相外切, 如图 1 所示。

2) 当 $d_{ij} = D\sqrt{1^2 + A_R^2}$ 时, f_i 的外切球体与 f_j 的外切球体相外切, 如图 2 所示。

根据上述 2 种情况可以将纤维之间的相交关系分为必然相交、必然不相交和可能相交 3 种情况:

1) 当 $0 \leq d_{ij} < D$ 时, f_i 的内切球体与 f_j 的内切球体相交, 即 f_i 与 f_j 必然相交。

2) 当 $d_{ij} > D\sqrt{1^2 + A_R^2}$ 时, f_i 的外切球体与 f_j 的外切球体相离, 即 f_i 与 f_j 必然不相交。

3) 当 $D < d_{ij} < D\sqrt{1^2 + A_R^2}$ 时, f_i 与 f_j 的内切球体相离, 外切球体相交, f_i 与 f_j 可能相交。 d_{ij} 越小, 相交的可能性就越大。

则 d_c 的取值会导致 4 种过滤状态:

1) 当 $0 < d_c < D$ 时, 过滤掉的 f_i 是一定与 f_j 相交的, 但会有一定与 f_j 相交的 f_i 未被过滤掉。

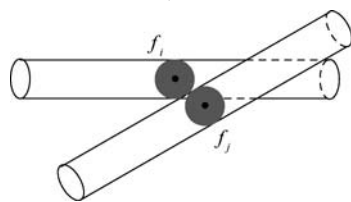


图 1 f_i 内切球体与 f_j 内切球体相外切

Fig. 1 Circumscribed inner spheres of f_i and f_j

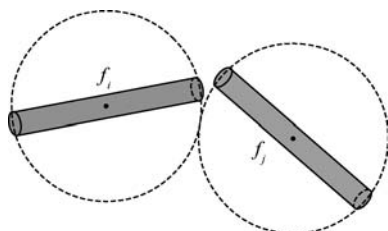


图 2 f_i 外切球体与 f_j 外切球体相外切

Fig. 2 Circumscribed outer spheres of f_i and f_j

2) 当 $d_c = D$ 时,过滤掉的 f_i 是一定与 f_j 相交的,并且一定与 f_j 相交的 f_i 均被过滤掉。

3) 当 $D < d_c < D\sqrt{1^2 + A_R^2}$ 时,过滤掉的纤维 f_i 有的不与 f_j 相交。

4) 当 $d_c > D\sqrt{1^2 + A_R^2}$ 时,所有可能与 f_j 相交的 f_i 都被过滤掉了。

由上述分析可知,如果 $d_c < D$,则会有与 f_j 必然相交的 f_i 未被滤掉,所以 d_c 的最小值应是 D 。状态 2) 和 3) 除了会过滤掉所有必然与 f_j 相交的 f_i 外,还会过滤掉不与 f_j 相交的 f_i ,此时过滤能起到提高建模效率的作用,但是会降低纤维位置的可行域,降低 RVE 的最大纤维体积分数,即提前出现 RVE 的纤维拥塞^[20-21]。 d_c 的取值对本方法影响很大,因此 FRSA 方法应用时需要选择合适的 d_c 取值。

2.3 FRSA 方法步骤

FRSA 方法将生成 RVE 的步骤 4) 改为:计算 f_i 与 $f_j (j=1, 2, \dots, i-1)$ 形心距的最小值 d_i 。当 $d_i < d_c$ 时进行步骤 5)。当 $d_i \geq d_c$ 时, $T = T + 1$, 将纤维 f_i 向 $(\pm L, 0, 0), (0, \pm W, 0), (0, 0, \pm H), (\pm L, \pm W, 0), (0, \pm W, \pm H), (\pm L, 0, \pm H), (\pm L, \pm W, \pm H)$ 共 26 个方向平移。然后执行布尔运算:用 27 个新纤维 f_i 减现有的 $i-1$ 根纤维,当 $i-1$ 根纤维的剩余体积小于 $(i-1)V_f$ 时,表明纤维之间有相交关系,进行步骤 5); 当 $i-1$ 根纤维的剩余体积等于 $(i-1)V_f$ 时,认为纤维 f_i 与纤维 $f_j (j=1, 2, \dots, i-1)$ 均不相交,进行步骤 6)。

需要指出的是,为了避免降低计算效率,上述方法中计算 d_i 时没有对跨边界纤维进行特别处理。

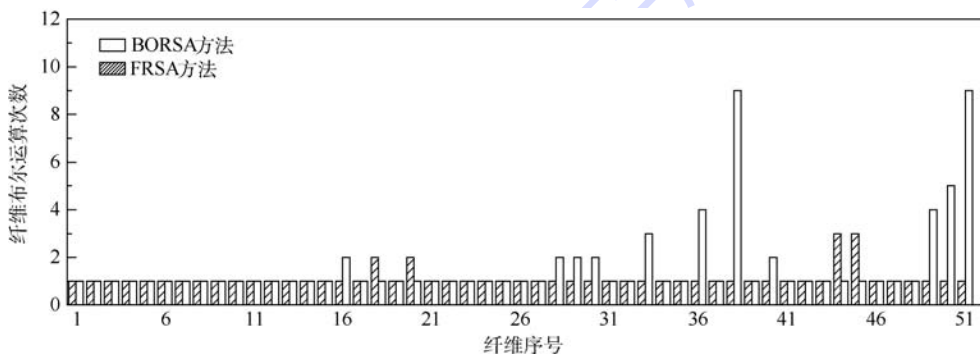


图 3 改进前后 2 种方法在一个 RVE 算例中每根纤维的布尔运算次数

Fig. 3 Number of Boolean operation times per fiber in one RVE example with two methods before and after improvement

3.2 RVE 参数和方法参数对 FRSA 方法先进性的影响

工程中通常需要生成不同纤维体积分数 V_c^{\max} 和不同长径比 A_R 的 RVE; 另外由 2.2 节可知, 临界距离 d_c 对 FRSA 方法效率有影响, 因此本节在

3 FRSA 方法先进性验证

3.1 典型参数下的 FRSA 方法先进性验证

设置 RVE 参数, 取纤维直径 $D = 20 \mu\text{m}$, RVE 长、宽及高 (L, W, H) 均为 $10D$, 纤维长径比 $A_R = 2$, 纤维体积分数 $V_c^{\max} = 8\%$ 。设置方法参数, 令最大布尔运算次数 $T_{\max} = 10^3$, $d_c = \sqrt{5}D$ 。

考虑到过滤法是通过减少布尔运算次数提高建模效率的, 图 3 给出采用 BORSA 方法和 FRSA 方法, 各生成 1 个 RVE 模型时, 生成每根纤维的布尔运算次数。

由图 3 可见, 当 $V_c^{\max} = 8\%$ 时, 一个 RVE 中需要生成 51 根纤维。BORSA 方法中生成 1 根纤维的布尔运算次数最大为 9 次, 而 FRSA 方法中最多为 3 次。总体来看, FRSA 方法明显减少了纤维布尔运算的次数。

采用 2 种方法各生成 10 个 RVE。计算生成一个 RVE 的总布尔运算次数, 并对 10 个 RVE 取平均值, 可知 BORSA 方法和 FRSA 方法生成一个 RVE 的平均布尔运算次数分别为 87.5 和 61.8 次, 改进后 FRSA 方法布尔运算次数降为 BORSA 方法的 70.6%。

建模效率的最好指标是建模时间。比较 2 种方法各生成 10 个算例的总时间, 发现 BORSA 方法耗时 1079 s, 改进后 FRSA 方法耗时 921 s, 可见建模时间降为原方法的 85.4%, 明显提高了 RVE 的建模效率。

采用 FRSA 方法生成的一个 RVE 如图 4 所示, 可以用来进行后续的材料性能分析。

不同的 V_c^{\max} 、 d_c 和 A_R 取值下对 FRSA 方法建模效率进行研究。

3.2.1 V_c^{\max} 对建模效率的影响

选取 V_c^{\max} 的变化范围为 5% ~ 11%, 每种条件均生成 10 个 RVE。计算采用 FRSA 生成一个

RVE 进行的布尔运算次数 T' 和采用 BORSA 生成一个 RVE 进行的布尔运算次数 T'' 的比值 (T'/T''), 采用 FRSA 生成一个 RVE 所用时间 t' 和采用 BORSA 生成一个 RVE 所用时间 t'' 的比值 (t'/t''), 并对 10 次模拟数据取平均值, 结果列于表 1。由表 1 可知, 在不同 V_C^{\max} 情况下 FRSA 方法可将生成 RVE 所需布尔运算次数降低 18% ~ 29%, 所需的时间降低 14% ~ 15%, 改进效果明显。

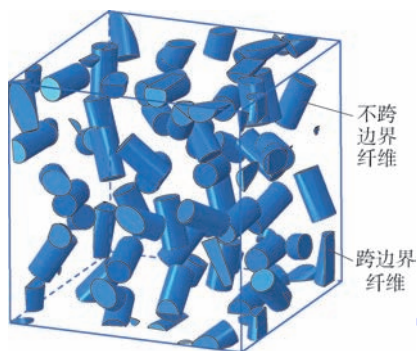


图 4 FRSA 方法生成的 RVE

Fig. 4 RVE generated with FRSA method

表 1 不同 V_C^{\max} 情况下的 RVE 建模效率对比

Table 1 Comparison of RVE modeling efficiency for various V_C^{\max}

$V_C^{\max}/\%$	T'/T''	t'/t''
5	0.76	0.86
8	0.71	0.85
11	0.82	0.86

3.2.2 d_c 对建模效率的影响

在 3.2.1 节的算例中, 设置了 FRSA 方法的临界距离 $d_c = \sqrt{5}D$, 但同时这种做法降低了 RVE 的堵塞极限 ($d_c = \sqrt{5}D$ 时堵塞极限为 11%, 无法生成 $V_C^{\max} > 11\%$ 的 RVE)。如减小 d_c 的取值, 可以提高堵塞极限, 但同时也会降低方法改进效果。为探究 d_c 取值对改进效果的影响, 这里取 $d_c = \sqrt{3.5}D$ 和 $\sqrt{2}D$ 2 种情况分别进行 RVE 生成模拟。

分别计算 2 种方法生成一个 RVE 所需时间的比值 t'/t'' , 并对 10 次模拟数据取平均值, 结果如表 2 所示。

由表 2 可知, 2 种情况下 FRSA 方法生成 RVE 所需时间比原方法均有所减少, 且同一个 V_C^{\max} 时, d_c 取值越大所需时间减少越明显, 建模效率提高的程度越大。结合表 1 结果可知, d_c 取值为 $\sqrt{2}D$ 、 $\sqrt{3.5}D$ 和 $\sqrt{5}D$ 时, 堵塞极限分别为 23%、17% 和 11%, 可见 d_c 越大, 虽然建模效率提高的程度越大, 但纤维堵塞现象出现越早。

3.2.3 A_R 对建模效率的影响

取 $A_R = 4, 6$ 和 8 3 种情况分别进行 RVE 生

成模拟, 以验证 FRSA 方法在不同 A_R 取值下的改进效果。RVE 建模所需时间的比值的 10 次模拟平均值如表 3 所示。

结合表 1、表 2 和表 3, $A_R = 4, 6$ 和 8 3 种情况下结果可见, FRSA 方法生成 RVE 的用时比 BORSA 方法均有所减少, 验证了 FRSA 方法在不同 A_R 取值下均具有先进性。并且, 相同的 d_c 取值条件下, A_R 取值越小, 建模时间减少越明显, 建模效率提高的程度越大。 A_R 取值越大, 建模效率改进效果降低, 是因为在相同 d_c 取值条件下, 当 A_R 取值增大时, 临界距离 d_c 与纤维外切球直径的比值减小, 导致过滤程度降低, 从而导致了方法改进效果的下降。

表 2 不同 d_c 情况下的 RVE 建模时间比

Table 2 Ratios of RVE modeling time for various d_c

$V_C^{\max}/\%$	t'/t''	
	$d_c = \sqrt{3.5}D$	$d_c = \sqrt{2}D$
5	0.84	0.86
8	0.81	0.82
11	0.66	0.72
14	0.56	0.74
17	0.48	0.52
20	—	0.63
23	—	0.37

表 3 不同 A_R 情况下的 RVE 建模时间比

Table 3 Ratios of RVE modeling time for various A_R

A_R	t'/t''		
	$d_c = \sqrt{5}D$	$d_c = \sqrt{3.5}D$	$d_c = \sqrt{2}D$
4	0.74	0.87	0.85
6	0.82	0.92	0.94
8	0.88	0.93	0.94

4 结 论

本文通过在 BORSA 方法进行布尔运算前引入纤维形心距判断, 过滤掉大量可能与已存在纤维相交的新纤维, 提出了 FRSA 方法。在不同 RVE 参数和方法参数下对比了 2 种方法的纤维布尔运算次数和建模时间, 证明了 FRSA 方法的先进性。主要结论如下:

1) 依据形心距可将纤维位置关系分为必然相交、必然不相交和可能相交 3 种情况。

2) 可通过纤维形心距判断, 过滤掉所有必然相交和部分可能相交的纤维。

3) FRSA 方法能够提高 RVE 的建模效率, 且效率提高程度随纤维体积分数增大而增加, 随纤维长径比增大而减少。

4) 增大临界形心距可以增加改进效果,但是更易出现纤维堵塞现象。

参考文献 (References)

- [1] 郭云竹. 热塑性复合材料研究及其在航空领域中的应用[J]. 纤维复合材料, 2016, 33(3): 20-23.
GUO Y Z. Research on thermoplastic composites and its application in aviation[J]. Fiber Composites, 2016, 33(3): 20-23 (in Chinese).
- [2] 郭茹南. 低密度纤维骨架增强泡沫碳的制备与结构性能演化[D]. 哈尔滨: 哈尔滨工业大学, 2016: 1-4.
GUO R N. Preparation and structural properties evolution of low density fiber reinforced foamed carbon[D]. Harbin: Harbin Institute of Technology, 2016: 1-4 (in Chinese).
- [3] 张兴福, 曹永敏, 崔洪涛, 等. 菱镁复合保温建筑墙板研究[J]. 21世纪建筑材料, 2010, 2(5): 20-24.
ZHANG X F, CAO Y M, CUI H T, et al. Study on magnesium composite insulation wallboard[J]. 21st Century Building Materials, 2010, 2(5): 20-24 (in Chinese).
- [4] 孔徐洁. 不同长度玻璃纤维增强复合材料力学性能与界面性能的研究[D]. 上海: 东华大学, 2016: 1-3.
KONG X J. Study on mechanical properties and interfacial properties of glass fiber reinforced composites with different lengths[D]. Shanghai: Donghua University, 2016: 1-3 (in Chinese).
- [5] 孔莉莉, 成玲, 万培培, 等. 苧麻纤维复合材料医用夹板的开发及智能化[J]. 材料导报, 2018, 32(7): 1202-1208.
KONG L L, CHENG L, WAN P P, et al. Development and intelligentization of ramie fiber composite material splints[J]. Materials Review, 2018, 32(7): 1202-1208 (in Chinese).
- [6] 姚树镇. 短切玻璃纤维毡在玻璃钢船艇上的应用[J]. 中外船舶科技, 2011(2): 37-38.
YAO S Z. Application of chopped glass fiber mat on fiberglass boat[J]. Chinese and Foreign Shipbuilding Technology, 2011(2): 37-38 (in Chinese).
- [7] 郑晓霞, 郑锡涛, 隼林虎. 多尺度方法在复合材料力学分析中的研究进展[J]. 力学进展, 2010, 40(1): 41-56.
ZHENG X X, ZHENG X T, GOU L H. Research progress of multi-scale methods in mechanical analysis of composite materials[J]. Advances in Mechanics, 2010, 40(1): 41-56 (in Chinese).
- [8] 刘振宇, 叶燎原, 潘文. 等效体积单元(RVE)在砌体有限元分析中的应用[J]. 工程力学, 2003, 20(2): 31-35.
LIU Z Y, YE L Y, PAN W. Application of equivalent volume element (RVE) in masonry finite element analysis[J]. Engineering Mechanics, 2003, 20(2): 31-35 (in Chinese).
- [9] 谢悦, 宿晓如, 冯春冬, 等. 颗粒增强橡胶复合材料有效力学性能预测分析[J]. 兵器装备工程学报, 2017, 38(7): 142-147.
XIE Y, SU X R, FENG C D, et al. Predictive analysis of effective mechanical properties of particle reinforced rubber composites[J]. Journal of Ordnance Equipment Engineering, 2017, 38(7): 142-147 (in Chinese).
- [10] 田文龙, 齐乐华, 周计明, 等. 基于随机顺序吸附法的 C_{st}/Mg 周期性胞模型的建立及其应用[J]. 稀有金属材料与工程, 2013, 42(8): 1601-1605.
TIAN W L, QI L H, ZHOU J M, et al. The establishment and application of C_{st}/Mg periodic cell model based on random sequential adsorption method[J]. Rare Metal Materials and Engineering, 2013, 42(8): 1601-1605 (in Chinese).
- [11] 霍金星, 齐乐华, 周计明. ABAQUS 二次开发在 C_{st}/Mg 复合材料微观几何建模中的应用[J]. 特种铸造及有色合金, 2011, 31(4): 322-324.
HUO J X, QI L H, ZHOU J M. Application of secondary development of ABAQUS in micro geometry modeling of C_{st}/Mg composites[J]. Special Casting & Nonferrous Alloys, 2011, 31(4): 322-324 (in Chinese).
- [12] KARI S, BERGER H, GABBERT U. Numerical evaluation of effective material properties of randomly distributed short cylindrical fibre composites[J]. Computational Materials Science, 2007, 39(1): 198-204.
- [13] TIAN W, QI L, ZHOU J, et al. Representative volume element for composites reinforced by spatially randomly distributed discontinuous fibers and its applications[J]. Composite Structures, 2015, 131(1): 366-373.
- [14] GHOSSEIN E, LÉVESQUE M. Random generation of periodic hard ellipsoids based on molecular dynamics: A computationally-efficient algorithm[J]. Journal of Computational Physics, 2013, 253: 471-490.
- [15] SCHNEIDER M. The sequential addition and migration method to generate representative volume elements for the homogenization of short fiber reinforced plastics[J]. Computational Mechanics, 2017, 59(2): 247-263.
- [16] GUSEV A A. Representative volume element size for elastic composites: A numerical study[J]. Journal of the Mechanics & Physics of Solids, 1997, 45(9): 1449-1459.
- [17] FAESSEL M, DELISÉE C, BOS F, et al. 3D modelling of random cellulosic fibrous networks based on X-ray tomography and image analysis[J]. Composites Science & Technology, 2005, 65(13): 1931-1940.
- [18] EBERLY D, TOOLS G. Intersection of cylinders[J]. Israel Journal of Mathematics, 2000, 113(1): 231-241.
- [19] LIU H, ZENG D, LI Y, et al. Development of RVE-embedded solid elements model for predicting effective elastic constants of discontinuous fiber reinforced composites[J]. Mechanics of Materials, 2016, 93: 109-123.
- [20] WILLIAMS S R, PHILIPSE A P. Random packings of spheres and spherocylinders simulated by mechanical contraction[J]. Physical Review E, 2003, 67(5): 051301.
- [21] 刘钊, 朱平, 朱超. 长纤维复合材料随机结构生成及其弹性性能预测方法: CN 106815408 A[P]. 2016-12-23.
LIU Z, ZHU P, ZHU C. Random structure generation and prediction of elastic properties of long fiber composites: CN 106815408 A[P]. 2016-12-23 (in Chinese).

作者简介:

刘丰睿 男, 博士, 讲师, 硕士生导师。主要研究方向: 复合材料结构失效分析、多学科优化设计。

骈璐 女, 硕士研究生。主要研究方向: 复合材料结构力学。

多场条件下多功能结构设计。

赵丽滨 女, 博士, 教授, 博士生导师。主要研究方向: 飞行器结构分析与优化设计、复合材料结构的破坏理论和力学设计、

张建宇 男, 博士, 研究员, 博士生导师。主要研究方向: 金属多轴疲劳、复合材料疲劳断裂、疲劳可靠性工程。

A filter based modeling method of RVE for short fiber reinforced composite

LIU Fengrui^{1,*}, PIAN Rong¹, ZHAO Libin¹, ZHANG Jianyu²

(1. School of Astronautics, Beihang University, Beijing 100083, China;

2. College of Aerospace Engineering, Chongqing University, Chongqing 400044, China)

Abstract: The representative volume element (RVE) method is a common numerical modeling method for predicting the material properties of short fiber reinforced composites. The RVE generation efficiency is an important factor affecting the efficiency of the property prediction. For the problem of too many Boolean operation times which cause low efficiency of Boolean operation when the existing random sequential adsorption (RSA) method is used to generate RVE, the centroid distance calculations between the lasted randomly generated fibers and the existing fibers are introduced. Using the centroid distance, some randomly generated fibers, which are intersecting with the existing fibers, are filtered out to reduce the number of Boolean operations. And a filter based random sequential adsorption (FRSA) method for modeling RVE is presented. By comparing the number of Boolean operations and the time cost for generating RVE, the FRSA method has been verified, and the verification is performed with different RVE parameters and method parameters using improved FRSA method and Boolean operation based random sequential adsorption (BORSA).

Keywords: short fiber reinforced composite; numerical modeling; random sequential adsorption (RSA) method; filtering method; centroid distance of the fiber

Received: 2018-05-22; **Accepted:** 2018-07-28; **Published online:** 2018-08-01 11:16

URL: kns.cnki.net/kcms/detail/11.2625.V.20180731.1717.002.html

Foundation items: National Natural Science Foundation of China (11772028, 11702012, 11572058, 11372020)

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

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

DOI: 10.13700/j.bh.1001-5965.2018.0130

一种新的矩独立重要性测度分析方法及高效算法

巩祥瑞^{1,2,*}, 吕震宙², 孙天宇¹, 张雷雷¹, 封雷¹

(1. 西安现代控制技术研究所, 西安 710065; 2. 西北工业大学 航空学院, 西安 710072)



摘 要: 为了更加合理地分析输入随机变量对结构系统失效概率的影响,提出了一种新的矩独立重要性测度分析方法。传统的重要性测度指标只能估计输入随机变量在固定点时对结构系统输出响应的影响,而所提新指标能够充分反映输入随机变量在其分布区域的所有缩减区间上变化时对结构系统输出响应的平均影响程度,更加符合工程实际。为了求解新指标,给出了2种算法:传统的双层重复抽样蒙特卡罗(DLRS MC)方法和自适应超球重要抽样(ARBIS)方法。双层重复抽样蒙特卡罗方法计算结果可以作为对比参照解,但求解效率较低,计算量很大;自适应超球重要抽样方法在满足新指标求解精度的前提下,计算效率得到很大提高。应用数值算例和工程算例证明了所提新指标的意义和所提新算法的高效性。

关键词: 全局灵敏度分析; 重要性测度; 失效概率; 自适应策略; 重要抽样方法

中图分类号: V215.7; TB114.3

文献标识码: A **文章编号:** 1001-5965(2019)02-0283-08

灵敏度分析方法主要研究结构系统的输入变量对输出响应的影响程度^[1],已广泛应用于可靠性工程^[2]、危险分析^[3-4]、环境科学^[5]、计算物理化学^[6-7]等领域。灵敏度分析一般分为局部灵敏度分析方法和全局灵敏度分析方法。局部灵敏度分析方法研究模型输入变量的分布参数对结构系统输出响应的影响程度^[8],只能反映输入变量在某一名义值处对输出响应的影响,具有很大的局限性^[9]。

全局灵敏度分析,也称为重要性测度分析,研究一个或多个输入变量对模型输出响应的影响程度,能够反映输入变量对结构系统输出响应的平均影响程度,因此,该分析方法在实际工程中得到了更广泛的应用。重要性测度分析的核心思想是对影响结构系统输出响应的输入变量进行重要性排序,从而得到对结构系统输出响应影响较大的输入变量,为结构系统的设计改进提供理论

依据。近年来,研究人员已经提出了很多重要性测度分析方法。Saltelli 和 Helton 提出了无参的方法,但缺乏模型的独立性^[9-10]。Sobol^[11]和 Iman^[12]等提出了基于方差的重要性测度分析方法和相应的求解算法,该方法隐含地假设了方差可以充分描述结构系统输出响应的不确定性。然而,相关学者指出随机变量的任何一阶矩只能反映其部分信息,必然会丢失一些信息^[13]。为了分析输入变量在其整个分布区域上对结构系统输出响应的影响程度,Borgonovo^[14]、Liu^[15]和 Li^[16]等又分别提出了矩独立重要性测度分析(Moment-independent Importance Measures analysis, M-IM)方法。

目前,Borgonovo^[14]所提出的矩独立重要性测度分析方法应用最为广泛。但在可靠性分析中,结构系统的失效概率通常是工程技术人员所关注的焦点。Li 等^[16]又提出了基于失效概率的矩独立重要性测度分析方法,该方法能够计算输入变

收稿日期: 2018-03-15; 录用日期: 2018-09-14; 网络出版时间: 2018-10-18 11:09

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20181012.1103.002.html

* 通信作者. E-mail: gxrui1991@126.com

引用格式: 巩祥瑞, 吕震宙, 孙天宇, 等. 一种新的矩独立重要性测度分析方法及高效算法[J]. 北京航空航天大学学报, 2019, 45(2): 283-290. GONG X R, LYU Z Z, SUN T Y, et al. A new moment-independent importance measure analysis method and its efficient algorithm [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 283-290 (in Chinese).

量在其分布区域的任一固定点时,结构系统的无条件失效概率和条件失效概率的平均偏差。然而,在实际工程中将输入变量固定于某一点,这是很难实现的,只能使输入变量在其分布区域的某一区间上缩减变化。

为了解决这一问题,Wei等^[17]提出了一种新的输入变量缩减区间的获取方法,该方法的基本思想是:首先产生2个随机数 q_1 和 q_2 , q_1 和 q_2 服从均匀分布, $q_1 \sim U(0,1)$, $q_2 \sim U(q_1,1)$,然后根据输入变量累积分布函数的逆函数 $F^{-1}(\cdot)$,得到输入变量的缩减区间 $[F^{-1}(q_1),F^{-1}(q_2)]$ 。该方法的优点是很容易获取输入变量的缩减区间,然而它并没有考虑输入变量缩减区间获取的任意性和等可能性。

本文提出了一种新的基于失效概率的矩独立重要性测度分析方法,并给出了一种更加合理的输入变量缩减区间的获取方法。同时,引入自适应超球重要抽样(Adaptive Radial-Based Importance Sampling, ARBIS)方法^[18-19],构建所提新指标的高效求解算法。应用数值算例和工程算例,验证本文所提新指标的合理性和求解算法的高效性。

1 新的矩独立重要性测度指标

1.1 新指标的定义

假设结构系统的极限状态函数为 $Y = g(\mathbf{X})$, \mathbf{X} 为 n 维输入变量,即 $\mathbf{X} = (X_1, X_2, \dots, X_n)$ 。记输入变量 X_i 的基于失效概率的矩独立重要性测度指标为 δ_i^p ,可以表示为

$$\delta_i^p = E_Q(P_f - P_{f|X_i \in U_i(Q)})^2 = E_Q(E(I_f) - E(I_f|X_i \in U_i(Q)))^2 \quad (1)$$

式中: P_f 为非条件失效概率; \mathbf{Q} 为分位数矩阵,用来生成输入变量 X_i 的缩减区间;当 X_i 的分布区间缩减到其所有可能的子区间 $U_i(\mathbf{Q})$ 时,就能够得到条件失效概率 $P_{f|X_i \in U_i(Q)}$; $E_Q(\cdot)$ 为对 \mathbf{Q} 求期望; I_f 为指示函数,当 \mathbf{X} 位于失效域(即 $g(\mathbf{X}) \leq 0$)时, $I_f(x) = 1$,否则 $I_f(x) = 0$ 。

在式(1)中,基于失效概率的矩独立重要性测度指标 δ_i^p 表示了当输入变量 X_i 在其分布区域的所有可能缩减区间上变化时,对结构系统失效概率的平均影响程度。

1.2 输入变量缩减区间获取方法

为了等概率地获取输入变量所有可能的缩减区间,首先定义一个分位数矩阵 $\mathbf{H}_k(3 \times 2)$:

$$\mathbf{H}_k = \begin{bmatrix} 0 & h_{k(1)} & h_{k(2)} \\ h_{k(1)} & h_{k(2)} & 1 \end{bmatrix}^T \quad (2)$$

式中: $h_{k(1)}$ 和 $h_{k(2)}$ 为 $[0,1]$ 区间上的2个随机数,满足 $0 \leq h_{k(1)} < h_{k(2)} \leq 1 (k=1,2,\dots,N)$ 。

$\mathbf{H}_k(k=1,2,\dots,N)$ 组成一个 $(3N \times 2)$ 的分位数矩阵 \mathbf{Q} :

$$\mathbf{Q} = [\mathbf{H}_1 \quad \mathbf{H}_2 \quad \dots \quad \mathbf{H}_k \quad \dots \quad \mathbf{H}_N]^T \quad (3)$$

式中: $\mathbf{H}_k(k=1,2,\dots,N)$ 是被等概率获取的,例如 $P\{\mathbf{H}_k\} = 1/N$ 。

为了更好地理解式(3),可表示为

$$\mathbf{Q} = \begin{bmatrix} 0 & h_{1(1)} & h_{1(2)} & \dots & 0 & h_{N(1)} & h_{N(2)} \\ h_{1(1)} & h_{1(2)} & 1 & \dots & h_{N(1)} & h_{N(2)} & 1 \end{bmatrix}^T = \begin{bmatrix} q_{1(1)} & q_{2(1)} & \dots & q_{3N(1)} \\ q_{1(2)} & q_{2(2)} & \dots & q_{3N(2)} \end{bmatrix}^T \quad (4)$$

式中: $q_{3k-2(1)} = 0, q_{3k-2(2)} = q_{3k-1(1)} = h_{k(1)}, q_{3k-1(2)} = q_{3k(1)} = h_{k(2)}, q_{3k(2)} = 1 (k=1,2,\dots,N)$ 。

因此,根据分位数矩阵 \mathbf{Q} 就可以定义缩减区间矩阵 \mathbf{U} 为

$$\mathbf{U} = \begin{bmatrix} u_{1(1)} & u_{2(1)} & \dots & u_{3N(1)} \\ u_{1(2)} & u_{2(2)} & \dots & u_{3N(2)} \end{bmatrix}^T \quad (5)$$

式中: $u_{j(1)} = F_X^{-1}(q_{j(1)}), u_{j(2)} = F_X^{-1}(q_{j(2)}), j=1,2,\dots,3N; F_X^{-1}(\cdot)$ 为输入变量 \mathbf{X} 的累积分布函数的逆函数。通过获取 $(u_{j(1)}, u_{j(2)}) (j=1,2,\dots,3N)$,输入变量 \mathbf{X} 所有可能的缩减区间就能够被等可能地取得。

对于输入变量 X_i ,其缩减区间可表示为

$$\mathbf{U}_i = \begin{bmatrix} u_{1(1)}^{(i)} & u_{2(1)}^{(i)} & \dots & u_{3N(1)}^{(i)} \\ u_{1(2)}^{(i)} & u_{2(2)}^{(i)} & \dots & u_{3N(2)}^{(i)} \end{bmatrix}^T \quad (6)$$

式中: $u_{j(1)}^{(i)} = F_{X_i}^{-1}(q_{j(1)}), u_{j(2)}^{(i)} = F_{X_i}^{-1}(q_{j(2)}), j=1,2,\dots,3N, F_{X_i}^{-1}(\cdot)$ 为输入变量 X_i 的累积分布函数的逆函数, $X_i \in [u_{j(1)}^{(i)}, u_{j(2)}^{(i)}] (j=1,2,\dots,3N)$ 为输入变量 X_i 的第 j 个缩减区间。

上述是一种新的区间划分技术来获取输入变量在其分布区域上所有可能的缩减区间。本文所提出的这种新的区间获取方法具有2个优点:①输入变量所有可能的缩减区间能够被等可能地获取,从而就能够更加合理地计算当输入变量在其分布区域的缩减区间上变化时对结构系统失效概率的平均影响程度;②全期望公式在这种区间划分技术中成立,基于全期望公式,第2节中就能够将本文提出的新的基于失效概率的矩独立重要性测度指标转化为更加便于计算的基于方差的重要性测度指标。

2 新指标与基于方差的指标的关系

由于 1.2 节中提出的输入变量缩减区间获取方法中全期望公式成立,就能够推导出新的矩独立重要性测度指标与基于方差的重要性测度指标间的关系,同时得到一种新的基于方差的矩独立重要性测度指标表示方式。

输入变量在缩减区间上的全期望公式表示为

$$E(I_F) = E_Q(E(I_F | X_i \in U(Q))) \quad (7)$$

证明过程详见附录 A。

因此,本文所提出的新的基于失效概率的矩独立重要性测度指标 δ_i^p 可表示

$$\delta_i^p = E_Q(E(I_F) - E(I_F | X_i \in U_i(Q)))^2 = V_Q(E(I_F | X_i \in U_i(Q))) \quad (8)$$

从式(8)中可以看出,新的矩独立重要性测度指标能够很容易地转换为基于方差的重要性测度指标,从而新的矩独立重要性测度指标就能够采用基于方差的重要性测度分析方法来进行求解。

3 新指标的高效求解算法

新的矩独立重要性测度指标传统的计算方法是双层重复抽样蒙特卡罗(Double-Loop-Repeat-Set Monte Carlo, DLRS MC)^[17]方法,该方法先用 N 个样本点计算非条件失效概率,然后在每一个子区间上,再产生 N 个新的样本点来计算条件失效概率,最后计算非条件失效概率和条件失效概率差异的平均值。

DLRS MC 方法可以获得高精度的结果值,但其计算量很大,计算效率较低。本节提出了一种新的求解算法——ARBIS 方法,旨在对新的矩独立重要性测度指标进行高效求解。

3.1 自适应超球重要抽样方法

ARBIS 方法最初是用来计算失效概率的,而计算新的基于失效概率的矩独立重要性测度指标的核心就是计算非条件失效概率和条件失效概率。因此,本节基于 ARBIS 方法构建新指标的求解算法。

ARBIS 方法^[18-19]的基本思想是自适应地寻找到一个落在功能函数安全域内的最大的超球。当抽样点落入超球内部时,就将该抽样点归类为安全点,不必再代入结构系统功能函数进行求解。因此,该方法在计算结构系统输出响应的失效概率时,计算量将大幅度降低。同时,ARBIS 方法采用自适应的方法来获取满足条件的最大超球半

径,相比于传统的梯度搜索算法,效率更高,稳健性也更好。

任何一个随机变量通过合适的变换方法,都可以转换到标准正态空间中^[20-21]。因此,以下步骤都假定在标准正态空间中进行。为了更好地理解 ARBIS 方法,给出该方法的一些关键步骤。

步骤 1 设定初始化超球半径 β_0 。为了确保超球与失效域相交,初始半径 β_0 可通过式(9)来确定:

$$\beta_0 = \sqrt{\chi_n^{-2}(1 - p_0)} \quad (9)$$

式中: χ_n^{-2} 表示卡方分布 χ_n^2 的累积分布函数的逆函数; p_0 为抽样点落在超球外部的概率,一般取很小值来确保初始超球与失效域相交。

步骤 2 根据随机变量的分布类型,通过转化得到标准正态空间的样本点 \mathbf{u}_k ($k=1,2,\dots,N$)。

步骤 3 计算落在半径为 β_0 的超球外部样本点的极限状态函数值,确定失效样本点。如果样本点满足 $\|\mathbf{u}_k\| \geq \beta_0$, 计算其极限状态函数值并保存结果。同时,对于满足极限状态函数值 $g(\mathbf{u}_k) \leq 0$ 的样本点,如果总数目为 N_f , 按照式(10)进行失效概率值的累加计算:

$$\hat{P}_{f_y} = \frac{1}{N} \sum_{j=1}^N I_F(\mathbf{u}_j) = \frac{N_f}{N} \quad (10)$$

步骤 4 确定第 i 次迭代的超球半径 β_i 。如果抽样点满足 $g(\mathbf{u}_k) \leq 0$ 且 $\|\mathbf{u}_k\| > \beta_{i-1}$, 则其为落在第 $i-1$ 次迭代超球外的失效样本点,如图 1 中所示(MPP 为最佳设计点)。计算这些点的概率密度函数值,并选取概率密度函数值最大的样本点,在其与坐标原点的连线得到一个以原点为起点的向量。在所得向量方向上采用线性搜索^[18], 由于是内插搜索,一般经过 2~3 次搜索便可确定

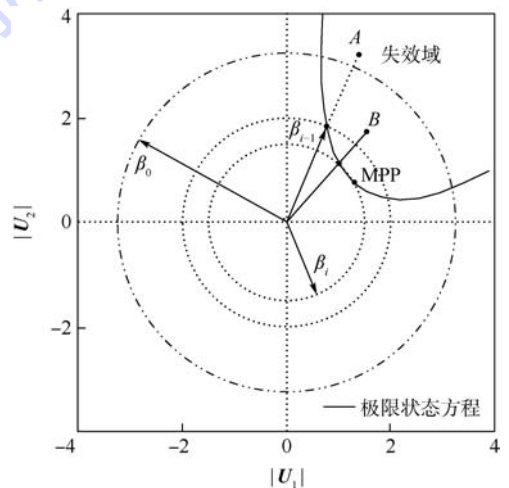


图 1 自适应策略获取最优化半径

Fig. 1 Adaptive strategy for obtaining optimal radius

该向量方向上与结构系统功能函数失效边界的近似交点,该交点到坐标原点的距离即为新的超球半径 β_i 。

步骤 5 筛选位于半径为 β_{i-1} 的超球和半径为 β_i 的超球之间的样本点,确定失效样本点。若样本点满足 $\beta_i < \|\mathbf{u}_k\| \leq \beta_{i-1}$, 计算其对应的结构系统的极限状态函数值,并保留结果。对于满足条件 $g(\mathbf{u}_k) \leq 0$ 的样本点,按照式(10)进行失效概率值的累加计算。

步骤 6 重复步骤 4 和步骤 5,直到收敛条件被满足^[22],即可得到最优的超球半径 β_{opt} 。

3.2 采用自适应超球重要抽样方法求解新指标

为了准确地计算新指标 δ_i^p ,从式(1)可以看出,关键在于计算结构系统的非条件失效概率 P_f 和条件失效概率 $P_{f|X_i \in U_i(Q)}$ 。本节采用 ARBIS 方法来构建 P_f 和 $P_{f|X_i \in U_i(Q)}$ 的求解算法。在构建的新算法中,含有截断输入变量 $X_i \in U_i(Q)$ 的条件失效概率 $P_{f|X_i \in U_i(Q)}$ 将被转化为不含截断输入变量的形式。经过这样的转化后,用于计算非条件失效概率 P_f 的 ARBIS 方法就能够很顺利地用来估计 $P_{f|X_i \in U_i(Q)}$,而不会增加额外的计算量。具体的转化过程如下所述。

当输入变量 X_i 缩减到区间 $U_i(Q)$ 时, X_i 就是一个截断变量,则所有输入变量的联合概率密度函数可表示为

$$f_X^*(\mathbf{x}) = \begin{cases} \frac{1}{K} f_X(\mathbf{x}) & X_i \in U_i(Q) \\ 0 & \text{其他} \end{cases} \quad (11)$$

式中: K 为截断系数,可表示为

$$K = \int_{X_i \in U_i(Q)} f_X(\mathbf{x}) d\mathbf{x} \quad (12)$$

对于式(12)的估计,并不需要计算结构系统的功能函数值,其计算量可以忽略。因此,条件失效概率 $P_{f|X_i \in U_i(Q)}$ 可重新表示为

$$P_{f|X_i \in U_i(Q)} = \int_{g(\mathbf{x}) \leq 0} f_X^*(\mathbf{x}) d\mathbf{x} = \frac{1}{K} \int_{X_i \in U_i(Q)} f_X(\mathbf{x}) d\mathbf{x} \quad (13)$$

从式(13)中可以看出,条件失效概率 $P_{f|X_i \in U_i(Q)}$ 的计算已转化为一个不含截断变量的多模态并联系系统问题。经过这种转化,构建 ARBIS 方法计算新的矩独立重要性测度指标的算法如下:

步骤 1 自适应搜索获取超球半径 β_{opt} ,并计算无条件失效概率 P_f ;同时得到半径为 β_{opt} 的超球外失效点组成的样本点矩阵 \mathbf{B} 。

步骤 2 当输入变量 X_i 缩减到区间 $U_i(Q)$ 时,在样本点矩阵 \mathbf{B} 中搜索输入变量 X_i 落在缩减区间 $U_i(Q)$ 的样本点,记样本点个数为 m 。

步骤 3 计算条件失效概率:

$$P_{f|X_i \in U_i(Q)} = \frac{1}{K} \cdot \frac{m}{N} \quad (14)$$

式中: N 为采用 ARBIS 方法的初始样本点个数。

步骤 4 根据式(1)计算新的矩独立重要性测度指标 δ_i^p 。

从上述步骤可以看出,ARBIS 方法能够很容易地被用来计算新指标 δ_i^p ,只需要一组输入变量的样本点就可以估计结构系统的非条件失效概率 P_f 和条件失效概率 $P_{f|X_i \in U_i(Q)}$,而不需要增加额外的计算量。因此,本节所提出的 ARBIS 方法相比 DLRS MC 方法计算效率得到很大提高。

4 算例分析

4.1 数值算例

某一结构系统的功能函数可表示为

$$G(\mathbf{X}) = 5 - X_1 - X_2^2 - 0.1 \sin(\pi X_2) - X_3^3 - 0.1 \sin(\pi X_3) \quad (15)$$

式中: X_1 、 X_2 和 X_3 为相互独立的输入随机变量,均服从标准正态分布,即 $X_i \sim N(0, 1)$, $i = 1, 2, 3$ 。

采用 DLRS MC 方法和 ARBIS 方法计算的新的矩独立重要性测度指标 δ_i^p 值列于表 1 中,每种方法的计算误差用标准差 SD 表示。

DLRS MC 方法的计算结果可以看作参考精确解。从表 1 可以看出,ARBIS 方法计算得到的结果与 DLRS MC 方法计算结果基本相同。为了说明 2 种算法计算过程的收敛性,图 2 给出了采用 2 种算法计算 δ_i^p 值的标准差。可以看出,当 2 种算法计算结果均收敛时,ARBIS 方法计算量远小于 DLRS MC 方法,计算效率提高很多。

从表 1 可以看出,采用 DLRS MC 方法和 ARBIS 方法计算的新的矩独立重要性测度指标 δ_i^p

表 1 DLRS MC 方法和 ARBIS 方法计算的新的矩独立重要性测度指标值

Table 1 New moment-independent importance measure indices of numerical example computed by DLRS MC and ARBIS methods

随机变量	DLRS MC 方法		ARBIS 方法	
	$\delta_i^p/10^{-2}$	SD/10 ⁻²	$\delta_i^p/10^{-2}$	SD/10 ⁻²
X_1	0.0610	0.010	0.0582	0.048
X_2	0.2264	0.007	0.2167	0.013
X_3	1.2572	0.035	1.2174	0.039
计算量	446000		2619	

方程:

$$g = \sigma_s - \sqrt{\sigma^2 + 3\tau^2} \quad (18)$$

式中: σ_s 为静强度屈服极限, 根据前轴材料特性有 $\sigma_s = 460$ MPa; g 为裕度值。

前轴结构尺寸参数和承受外载看作独立正态随机变量, 分布参数如表 2 所示。

采用 DLRS MC 方法和 ARBIS 方法计算的新的矩独立重要性测度指标结果如表 3 所示, 同时也给出了计算功能函数的总次数。可以看出, 2 种算法计算结果相近。图 4 给出了 DLRS MC 方法和 ARBIS 方法计算结果标准差的收敛曲线。可见, 估计值的标准差都比较小, 因此计算结果可靠稳健。DLRS MC 方法的总次数为 446 000, 而 ARBIS 方法计算总次数仅为 4 475, 因此 ARBIS 方法的计算效率很高。

2 种算法计算的新的矩独立重要性测度指标

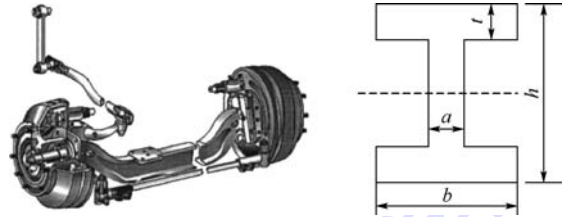


图 3 汽车前轴结构示意图

Fig. 3 Schematic of automobile front axle structure

表 2 汽车前轴结构各输入变量分布参数

Table 2 Distribution parameters of input variables of automobile front axle structure

随机变量	均值	标准差
a/mm	12	0.60
b/mm	65	3.25
t/mm	14	0.70
h/mm	85	4.25
$M/(\text{N} \cdot \text{mm})$	3.5×10^6	7.5×10^5
$T/(\text{N} \cdot \text{mm})$	3.1×10^6	1.5×10^5

表 3 DLRS MC 方法和 ARBIS 方法计算汽车前轴指标结果

Table 3 Results of indices of automobile front axle computed by DLRS MC and ARBIS methods

随机变量	DLRS MC 方法		ARBIS 方法	
	δ_i^p	SD/ 10^{-3}	δ_i^p	SD/ 10^{-3}
a	0.0049	0.276	0.0047	0.249
b	0.0064	0.394	0.0062	0.333
t	0.0163	0.546	0.0157	0.608
h	0.0014	0.173	0.0013	0.149
M	6.49×10^{-5}	0.092	1.95×10^{-5}	0.012
T	0.0095	0.426	0.0092	0.321
计算量	446 000		4 475	

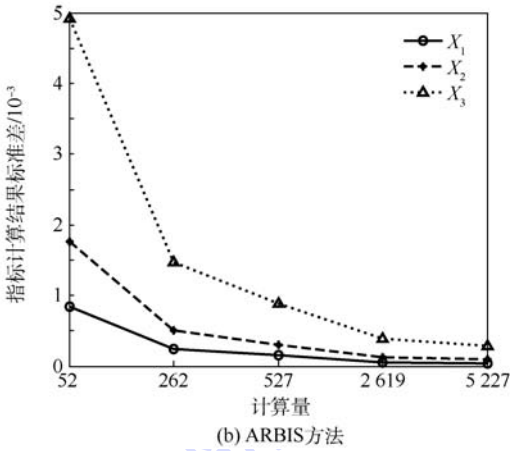
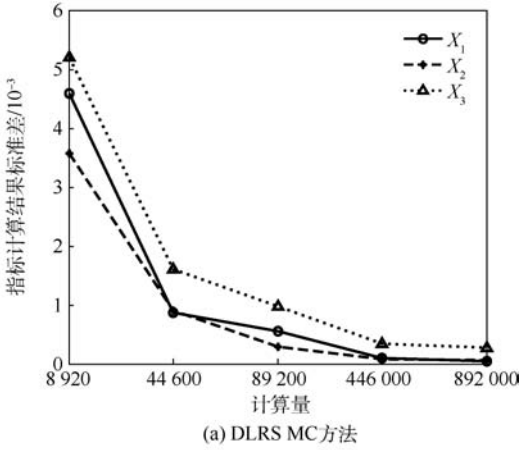


图 2 DLRS MC 方法和 ARBIS 方法计算的新的矩独立重要性测度指标的收敛曲线

Fig. 2 Convergence curves of new moment-independent importance measure indices of numerical example computed by DLRS MC and ARBIS methods

的重要性排序相同: $X_3 > X_2 > X_1$ 。这表明当 3 个输入变量在各自的分布区域内缩减变化时, X_3 对于结构系统失效概率的影响程度最大。

4.2 工程算例

在汽车结构中, 车桥承载着大部分汽车重量, 其通过悬臂与车架相接, 将来自车轮的牵引力和制动力, 还有侧向力经过悬架传递给车架, 起主要承载作用的就是汽车前轴^[23]。由于工字型截断梁能够提高抗弯强度, 因此前轴通常采用工字结构梁。

图 3 为前轴的结构示意图。危险截面处的最大正应力为 $\sigma = M/W_x$, 最大切应力为 $\tau = T/W_p$, M 为弯矩, T 为扭矩, W_x 为截面系数, W_p 为极截面系数, 且有

$$W_x = \frac{a(h-2t)^3}{6h} + \frac{b}{6h}[h^3 - (h-2t)^3] \quad (16)$$

$$W_p = 0.8bt^2 + 0.4[a^3(h-2t)/t] \quad (17)$$

考虑前轴结构静强度失效, 有以下极限状态

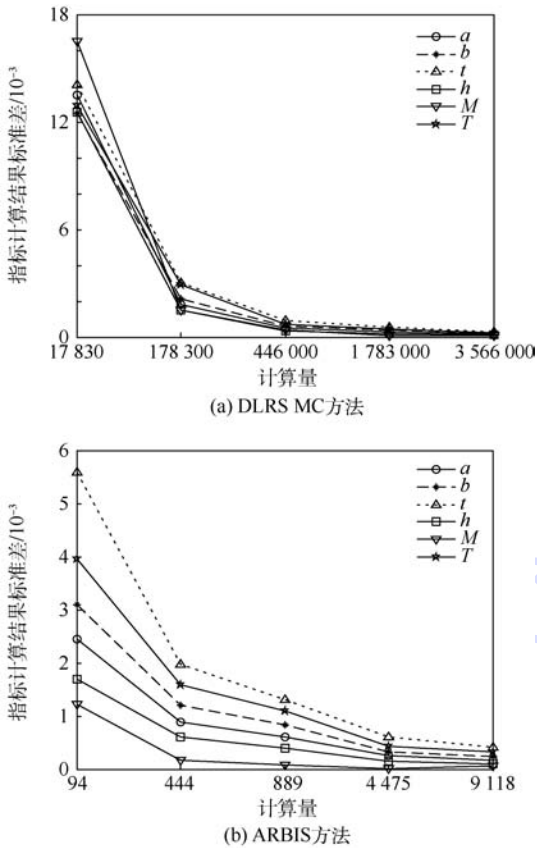


图4 DLRS MC方法和ARBIS方法计算汽车前轴的收敛曲线

Fig.4 Convergence curves of automobile front axle computed by DLRS MC and ARBIS methods

的排序均相同： $t > T > b > a > h > M$ 。这表明当分别固定6个输入随机变量在其各自的分布区域内缩减变化时， t 对汽车前轴结构的失效概率影响程度最大，而 M 对其失效概率的影响几乎可以忽略不计。

5 结论

本文提出了一种新的基于失效概率的矩独立重要性测度分析方法来计算输入随机变量对结构系统输出响应失效概率的平均影响程度。

1) 建立了一种新的区间划分技术来等可能地获得输入随机变量所有可能的缩减区间，并给出了相应证明。

2) 引入了自适应超球重要抽样方法来进行新指标的求解，提高了计算效率。

3) 给出了一个数值算例和一个工程算例，说明了本文所提新指标的意义，同时也验证了新算法的高效性。

参考文献 (References)

- [1] CASTILLO E, MINGUEZ R, CASTILLO C. Sensitivity analysis in optimization and reliability problems [J]. Reliability Engineering & System Safety, 2008, 93(12): 1788-1800.
- [2] WEI P, LV Z, HAO W, et al. Efficient sampling methods for global reliability sensitivity analysis [J]. Computer Physics Communications, 2012, 183(8): 1728-1743.
- [3] SALTELLI A. Sensitivity analysis for importance assessment [J]. Risk Analysis, 2002, 22(3): 579-590.
- [4] BJERAGER P, KRENK S. Parametric sensitivity in first order reliability theory [J]. Journal of Engineering Mechanics, 1989, 115(7): 1577-1582.
- [5] CAMPOLONGO F, SALTELLI A. Sensitivity analysis of an environmental model: An application of different analysis methods [J]. Reliability Engineering & System Safety, 1997, 57(1): 49-69.
- [6] SALTELLI A, RATTO M, TARANTOLA S, et al. Sensitivity analysis for chemical models [J]. Chemical Reviews, 2005, 105(7): 2811-2828.
- [7] BORGONOVO E, TARANTOLA S. Moment independent and variance-based sensitivity analysis with correlations: An application to the stability of a chemical reactor [J]. International Journal of Chemical Kinetics, 2008, 40(11): 687-698.
- [8] VAN GRIENSVEN A, MEIXNER T, GRUNWALD S, et al. A global sensitivity analysis tool for the parameters of multi-variable catchment models [J]. Journal of Hydrology, 2006, 324(1): 10-23.
- [9] SALTELLI A, MARIVOET J. Non-parametric statistics in sensitivity analysis for model output: A comparison of selected techniques [J]. Reliability Engineering & System Safety, 1990, 28(2): 229-253.
- [10] HELTON J C. Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal [J]. Reliability Engineering & System Safety, 1993, 42(2): 327-367.
- [11] SOBOLO I M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates [J]. Mathematics and Computers in Simulation, 2001, 55(1): 271-280.
- [12] IMAN R L, HORA S C. A robust measure of uncertainty importance for use in fault tree system analysis [J]. Risk Analysis, 1990, 10(3): 401-406.
- [13] HELTON J C, DAVIS F J. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems [J]. Reliability Engineering & System Safety, 2003, 81(1): 23-69.
- [14] BORGONOVO E. A new uncertainty importance measure [J]. Reliability Engineering & System Safety, 2007, 92(6): 771-784.
- [15] LIU H, CHEN W, SUDJIANTO A. Relative entropy based method for probabilistic sensitivity analysis in engineering design [J]. Journal of Mechanical Design, 2006, 128(2): 326-336.
- [16] LI L Y, LU Z Z, FENG J, et al. Moment-independent importance measure of basic variable and its state dependent parameter solution [J]. Structural Safety, 2012, 38: 40-47.
- [17] WEI P, LU Z, SONG J. A new variance-based global sensitivity analysis technique [J]. Computer Physics Communications, 2013, 184(11): 2540-2551.
- [18] GROOTEMAN F. Adaptive radial-based importance sampling

method for structural reliability[J]. Structural Safety, 2008, 30(6):533-542.

- [19] AU S K, BECK J L. A new adaptive importance sampling scheme for reliability calculations[J]. Structural Safety, 1999, 21(2):135-158.
- [20] ROSENBLATT M. Remarks on a multivariate transformation [J]. The Annals of Mathematical Statistics, 1952, 23(3):470-472.
- [21] DER KIUREGHIAN A, LIU P L. Structural reliability under incomplete probability information [J]. Journal of Engineering Mechanics, 1986, 112(1):85-104.
- [22] 吕震宙. 结构机构可靠性及可靠性灵敏度分析[M]. 北京: 科学出版社, 2009:160-178.
- LV Z Z. Reliability and reliability sensitivity analysis of structural mechanism[M]. Beijing: Science Press, 2009:160-178 (in Chinese).
- [23] CAI Z Y. Precision design of roll-forging die and its application in the forming of automobile front axles[J]. Journal of Materials Processing Technology, 2005, 168(1):95-101.

作者简介:

巩祥瑞 男, 硕士研究生。主要研究方向: 可靠性工程。

吕震宙 女, 博士, 教授, 博士生导师。主要研究方向: 飞行器可靠性工程、安全工程。

附录 A

由于 $\mathbf{Q} = [\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_k, \dots, \mathbf{H}_N]^T$, 同时

$P\{\mathbf{H}_k\} = 1/N$ 成立, 则有以下推导成立:

$$E_{\mathbf{Q}}(E(I_F | X_i \in \mathbf{U}(\mathbf{Q}))) = \sum_{k=1}^N E_{\mathbf{H}_k}(E(I_F | X_i \in \mathbf{U}(\mathbf{H}_k))) \cdot P\{\mathbf{H}_k\} = \frac{1}{N} \sum_{k=1}^N E_{\mathbf{H}_k}(E(I_F | X_i \in \mathbf{U}(\mathbf{H}_k))) \quad (\text{A1})$$

式中:

$$\mathbf{U}(\mathbf{H}_k) = \begin{bmatrix} F_{X_i}^{-1}(0) & F_{X_i}^{-1}(h_{k(1)}) \\ F_{X_i}^{-1}(h_{k(1)}) & F_{X_i}^{-1}(h_{k(2)}) \\ F_{X_i}^{-1}(h_{k(2)}) & F_{X_i}^{-1}(1) \end{bmatrix} \quad (\text{A2})$$

由指示函数期望 $E(I_F)$ 的定义, 可以进行如下推导:

$$\begin{aligned} E(I_F) &= \int_{\mathbf{X}} I_F(\mathbf{x}) \cdot f_{\mathbf{X}}(\mathbf{x}) \, d\mathbf{x} = \\ &= \int_{X_{-i}} \int_{F_{X_i}^{-1}(0)}^{F_{X_i}^{-1}(1)} I_F(\mathbf{x}) \cdot f_{X_i}(x_i) \, dx_i \cdot f_{X_{-i}}(x_{-i}) \, dx_{-i} = \\ &= \int_{X_{-i}} \int_{F_{X_i}^{-1}(0)}^{F_{X_i}^{-1}(h_{k(1)})} I_F(\mathbf{x}) \cdot f_{X_i}(x_i) \, dx_i \cdot f_{X_{-i}}(x_{-i}) \, dx_{-i} + \\ &= \int_{X_{-i}} \int_{F_{X_i}^{-1}(h_{k(1)})}^{F_{X_i}^{-1}(h_{k(2)})} I_F(\mathbf{x}) \cdot f_{X_i}(x_i) \, dx_i \cdot \\ &= f_{X_{-i}}(x_{-i}) \, dx_{-i} + \int_{X_{-i}} \int_{F_{X_i}^{-1}(h_{k(2)})}^{F_{X_i}^{-1}(1)} I_F(\mathbf{x}) \cdot \\ &= f_{X_i}(x_i) \, dx_i \cdot f_{X_{-i}}(x_{-i}) \, dx_{-i} = \\ &= \int_{X_{-i}} \int_{F_{X_i}^{-1}(0)}^{F_{X_i}^{-1}(h_{k(1)})} I_F(\mathbf{x}) \cdot \frac{f_{X_i}(x_i)}{h_{k(1)}} \, dx_i \cdot \\ &= h_{k(1)} f_{X_{-i}}(x_{-i}) \, dx_{-i} + \int_{X_{-i}} \int_{F_{X_i}^{-1}(h_{k(1)})}^{F_{X_i}^{-1}(h_{k(2)})} I_F(\mathbf{x}) \cdot \\ &= \frac{f_{X_i}(x_i)}{h_{k(2)} - h_{k(1)}} f_{X_i}(x_i) \, dx_i \cdot (h_{k(2)} - h_{k(1)}) \cdot \\ &= f_{X_{-i}}(x_{-i}) \, dx_{-i} + \int_{X_{-i}} \int_{F_{X_i}^{-1}(h_{k(2)})}^{F_{X_i}^{-1}(1)} I_F(\mathbf{x}) \cdot \\ &= \frac{f_{X_i}(x_i)}{1 - h_{k(2)}} f_{X_i}(x_i) \, dx_i \cdot (1 - h_{k(2)}) \cdot \\ &= f_{X_{-i}}(x_{-i}) \, dx_{-i} = E_{\mathbf{H}_k}(E(I_F | X_i \in \mathbf{U}(\mathbf{H}_k))) \end{aligned} \quad (\text{A3})$$

将式(A3)代入式(A1)中, 就可以得到如式(A4)所示的区间上的全期望公式:

$$E_{\mathbf{Q}}(E(I_F | X_i \in \mathbf{U}(\mathbf{Q}))) = \frac{1}{N} \sum_{k=1}^N E(I_F) = E(I_F) \quad (\text{A4})$$

A new moment-independent importance measure analysis method and its efficient algorithm

GONG Xiangrui^{1,2,*}, LYU Zhenzhou², SUN Tianyu¹, ZHANG Leilei¹, FENG Lei¹

(1. Xi'an Institute of Modern Control Technology, Xi'an 710065, China;

2. College of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: In order to analyze the effect of input random variables on the failure probability of structural systems more reasonably, a new moment-independent importance measure analysis method is proposed in this paper. The traditional importance measure index can only estimate the influence of input random variables on the output response of structural systems at fixed points, while the new index proposed in this paper can fully reflect the average influence of input random variables on the output response of structural systems when they change in all reduced intervals of their distribution areas, which is more in line with engineering practice. Seeking to find the new index, this paper presents two algorithms: the conventional double-loop-repeated-set Monte Carlo (DLRS MC) method and adaptive radial-based importance sampling (ARBIS) method. The results of DLRS MC method can be used as a reference solution, yet its calculation process is slow and strenuous. Under the condition of the precision of solving the new index is met, the calculation efficiency of ARBIS method is greatly improved. Finally, a numerical example and an engineering example are given to illustrate the significance of the new index and the efficiency of the proposed algorithm.

Keywords: global sensitivity analysis; importance measure; failure probability; adaptive strategy; importance sampling method

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

DOI: 10.13700/j.bh.1001-5965.2018.0273

弹性高速飞行器的状态/参数滚动时域估计



陈尔康, 荆武兴*, 高长生

(哈尔滨工业大学 航天学院, 哈尔滨 150001)

摘 要: 针对弹性高速飞行器非线性、不确定性和刚体/弹性耦合的特点, 提出了一种基于 QR 分解和滚动时域估计的状态/参数联合估计方法。首先, 通过引入滚动时域策略, 将状态/参数估计问题转化为固定变量数目的优化问题, 能够较好地处理时变参数的估计问题。然后, 利用前向动态规划原理, 将到达代价的计算转化为最小二乘问题, 并利用 QR 分解进行求解, 从而给出了基于 QR 分解的到达代价更新方法。这样使得整个滚动时域估计方法都建立在优化的基础上, 且引入了反馈机制, 提高了估计精度和速度。仿真结果表明: 滚动时域估计的精度明显优于扩展卡尔曼滤波, 且基于 QR 分解的到达代价更新方法在速度上优于传统的基于估计误差协方差的到达代价更新方法。

关键词: 高速飞行器; 弹性; 滚动时域估计; 状态估计; 参数估计; QR 分解

中图分类号: V249.3; TN713

文献标识码: A **文章编号:** 1001-5965(2019)02-0291-08

高速飞行器指高速飞行的有翼或无翼飞行器^[1], 具有强非线性、不确定性、强耦合等特点^[2]。此外由于大量使用轻质材料, 其结构的固有频率较低, 机体弹性形变对其控制有较大影响^[3], 控制系统需要对弹性进行抑制以保证稳定性^[1]。现有研究多将弹性振动看作干扰并在控制回路中加以抑制^[4], 这样虽然能够保证飞行器的稳定, 但难以提高响应速度和控制精度等性能。要实现弹性高速飞行器的精细姿态控制^[3], 需要使用动态面控制、模型预测控制等先进控制方法, 这势必需要引入难以直接测量的弹性状态的反馈, 因此对弹性振动状态进行估计十分重要。除弹性振动状态外, 弹性振动的相关参数如弹性模态固有频率等也十分重要, 能够用于飞行器的故障诊断等。固有频率与飞行器状态相关, 在飞行过程中并非常数, 因此对其进行估计是十分必要的。综上, 需要研究弹性高速飞行器的状态/参数联合估计方法^[5]。

作为一种基于模型的递归滤波器, 扩展卡尔曼滤波(EKF)利用一阶线性化的方法对卡尔曼滤波(KF)进行扩展, 是当前较为成熟和使用广泛的状态/参数估计方法^[5]。虽然 EKF 可以通过将参数扩展为状态量的方式对参数进行估计, 但受限于噪声模型难以估计时变参数; 此外, EKF 也无法处理状态量的约束^[6]。而滚动时域估计(MHE)方法则从最优控制问题的角度出发^[7], 并引入滚动时域策略^[8], 通过求解优化问题实现对状态/参数的估计; 因此滚动时域估计能够较好地处理带约束的估计问题和状态/参数联合估计问题^[5], 近年来得到越来越广泛的应用。文献[9]建立了滚动时域估计渐近稳定和有限稳定的充分条件, 并给出了一种算法实现。文献[10]在化学工程问题中利用滚动时域估计进行状态/参数联合估计。文献[11]将滚动时域估计用于航天器的姿态估计和传感器参数校正。文献[12]则研究了机器人的多速率采样滚动时域估计

收稿日期: 2018-05-15; 录用日期: 2018-07-28; 网络出版时间: 2018-08-22 18:07

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180820.1755.011.html

基金项目: 国家自然科学基金(11572097)

* 通信作者. E-mail: jingwuxing@hit.edu.cn

引用格式: 陈尔康, 荆武兴, 高长生. 弹性高速飞行器的状态/参数滚动时域估计[J]. 北京航空航天大学学报, 2019, 45(2): 291-298. CHEN E K, JING W X, GAO C S. State/parameter moving horizon estimation for elastic hypersonic vehicles [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 291-298 (in Chinese).

问题。

对滚动时域估计的研究可分为两大类:到达代价的计算和优化问题的求解。到达代价的计算直接关系到滚动时域估计的准确性和稳定性。文献[7]推导了滚动时域估计稳定的充分条件,提出利用估计的协方差矩阵计算到达代价,这也是目前计算到达代价的主流思路。基于文献[7]的研究成果,文献[13]采用卡尔曼滤波的误差协方差矩阵计算到达代价,文献[14-15]利用 EKF 计算到达代价,文献[16]则利用无迹卡尔曼滤波(UKF)计算到达代价。

这些方法都是从概率与统计的角度出发,且并未利用滚动时域估计的结果,因此是一类开环的到达代价计算方法,不利于进一步提高估计精度。针对这一问题,本文从动态规划原理出发,将到达代价的计算转化为最小二乘问题,并利用 QR 分解给出计算方法。该方法在到达代价的计算过程中使用了最新估计值,因而形成了反馈机制,有利于提高估计精度与速度。

1 弹性高速飞行器数学模型

本文主要研究弹性高速飞行器的纵向状态/参数估计问题,因此使用如下纵向动力学模型^[17]:

$$\begin{cases} mV\dot{\theta} = L - mg\cos\theta \\ \dot{\vartheta} = \omega_z \\ \dot{\omega}_z = M/I_z \\ \vartheta - \theta - \alpha = 0 \\ \ddot{\eta}_i + 2\zeta_i\omega_i\dot{\eta}_i + \omega_i^2\eta_i = N_i \quad i = 1, 2, \dots, n \end{cases} \quad (1)$$

式中: m 为飞行器质量; V 为飞行器速度大小; θ 为弹道倾角; L 为升力; g 为重力加速度; ϑ 为俯仰角; ω_z 为俯仰角速度; M 为俯仰力矩; I_z 为转动惯量; α 为迎角; η_i 、 ζ_i 、 ω_i 和 N_i 分别为第 i 阶弹性模态的广义坐标、阻尼比、固有频率和广义力。

本文考虑2种传感器布置方案。方案1采用常规传感器输出:角速度信号 ω_m 和过载信号 n_m 。

$$\begin{cases} \omega_m = \omega_z + \sum_{i=1}^n \frac{d\varphi_i(x_s)}{dx} \dot{\eta}_i(t) \\ n_m = n_z + \frac{\sum_{i=1}^n \varphi_i(x_s) \ddot{\eta}_i(t)}{g} \end{cases} \quad (2)$$

式中: $\varphi_i(x)$ 为第 i 阶弹性模态的振型; x_s 为传感器的轴向坐标; n_z 为刚体模态产生的过载。

一般地,弹性高速飞行器仅一阶弹性模态与

刚体运动及控制系统频带耦合^[17-18],因此本文仅考虑一阶模态。忽略长周期模态,且考虑弹性模态固有频率的变化,令

$$\begin{cases} \mathbf{x} = [\vartheta \quad \omega_z \quad \eta_1 \quad \dot{\eta}_1 \quad \omega_1]^T \\ \mathbf{y} = [\omega_m \quad n_m]^T \\ \mathbf{u} = \delta \end{cases} \quad (3)$$

式中: δ 为飞行器的舵偏角。

方案2采用分别安装在弹体前半部(距头部 x_{s1})和后半部(距头部 x_{s2})的2套陀螺^[4]:

$$\begin{cases} \omega_{m1} = \omega_z + \sum_{i=1}^n \frac{d\varphi_i(x_{s1})}{dx} \dot{\eta}_i(t) \\ \omega_{m2} = \omega_z + \sum_{i=1}^n \frac{d\varphi_i(x_{s2})}{dx} \dot{\eta}_i(t) \end{cases} \quad (4)$$

分别将上述信号作差,可得到与刚体运动无关的角速度信号分量 ω_e 和受到弹性振动信号扰动的角速度信号分量 ω_r :

$$\begin{cases} \omega_e = \omega_{m1} - \omega_{m2} \\ \omega_r = \frac{\omega_{m1} + \omega_{m2}}{2} \end{cases} \quad (5)$$

同样可令

$$\begin{cases} \mathbf{x} = [\vartheta \quad \omega_z \quad \eta_1 \quad \dot{\eta}_1 \quad \omega_1]^T \\ \mathbf{y} = [\omega_e \quad \omega_r]^T \\ \mathbf{u} = \delta \end{cases} \quad (6)$$

综上,弹性高速飞行器的数学模型如下:

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \\ \mathbf{y} = \mathbf{h}(\mathbf{x}) \end{cases} \quad (7)$$

2 基于 QR 分解的滚动时域估计

2.1 问题描述

对于式(7)所示的连续系统,可利用差分、多重打靶等方法转化为式(8)所示的离散系统:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{w}_k \\ \mathbf{y}_k = \mathbf{H}(\mathbf{x}_k) + \mathbf{v}_k \end{cases} \quad (8)$$

式中: \mathbf{w}_k 和 \mathbf{v}_k 分别为系统噪声序列和量测噪声序列,通常情况下假设二者服从零均值高斯分布:

$$\begin{cases} \mathbf{w}_k \sim N(0, \mathbf{Q}) \\ \mathbf{v}_k \sim N(0, \mathbf{R}) \end{cases} \quad (9)$$

其中: \mathbf{Q} 和 \mathbf{R} 分别为系统噪声和量测噪声的协方差矩阵。

系统式(8)的状态估计问题可转化为如下优化问题:

$$\min_{\mathbf{x}_0, \{\mathbf{w}_k\}_{k=0}^{T-1}} \Phi_T(\mathbf{x}_0, \{\mathbf{w}_k\}) \quad (10)$$

式中:

$$\Phi_T(\mathbf{x}_0, \{\mathbf{w}_k\}) = \|\mathbf{x}_0 - \hat{\mathbf{x}}_0\|_{\mathbf{P}_0}^2 + \sum_{k=0}^T \|\mathbf{v}_k\|_{\mathbf{R}^{-1}}^2 + \sum_{k=0}^{T-1} \|\mathbf{w}_k\|_{\mathbf{Q}^{-1}}^2 \quad (11)$$

其中: T 为当前时刻; $\hat{\mathbf{x}}_0$ 为系统状态初值的估计值; \mathbf{P}_0 为 $\hat{\mathbf{x}}_0$ 的协方差矩阵; $\|\mathbf{z}\|_{\mathbf{U}}$ 表示 $\mathbf{z}^T \mathbf{U} \mathbf{z}$; $\{\mathbf{w}_k\}_{k=0}^{T-1}$ 表示从 0 时刻到 $T-1$ 时刻的噪声序列。

优化问题式(10)利用了全部的测量数据,因此称全信息估计。全信息估计的计算量会随 T 的增长而迅速增大到无法接受的地步,因此引入滚动时域策略以限制全信息估计的维数,形成滚动时域估计方法。滚动时域估计方法的目标函数 Φ_T 变为

$$\Phi_T(\mathbf{x}_{T-N}, \{\mathbf{w}_k\}) = C_{T-N}(\mathbf{x}_{T-N}) + \sum_{k=T-N}^T \|\mathbf{v}_k\|_{\mathbf{R}^{-1}}^2 + \sum_{k=T-N}^{T-1} \|\mathbf{w}_k\|_{\mathbf{Q}^{-1}}^2 \quad (12)$$

式中: $C_{T-N}(\mathbf{x}_{T-N})$ 为到达代价,根据前向动态规划原理,其定义如下:

$$C_T(\mathbf{z}) = \min_{\mathbf{x}_0, \{\mathbf{w}_k\}_{k=0}^{T-1}} \|\mathbf{x}_0 - \hat{\mathbf{x}}_0\|_{\mathbf{P}_0}^2 + \sum_{k=0}^{T-N-1} \|\mathbf{v}_k\|_{\mathbf{R}^{-1}}^2 + \sum_{k=0}^{T-N-1} \|\mathbf{w}_k\|_{\mathbf{Q}^{-1}}^2 \quad (13)$$

s. t. $\mathbf{x}(T, \mathbf{x}_0, \{\mathbf{w}_k\}) = \mathbf{z}$

其中: $\mathbf{x}(\tau, \mathbf{x}_0, \{\mathbf{w}_k\})$ 表示在初值为 \mathbf{x}_0 且受到噪声序列 $\{\mathbf{w}_k\}$ 的情况下 τ 时刻状态量的取值。

由式(13)可知,到达代价是一个表达式非常复杂的函数^[19]。因此,一般用如下二次函数近似表示到达代价:

$$\Theta_{T-N}(\mathbf{x}_{T-N}) = \Phi_{T-N}^* + \|\mathbf{x}_{T-N} - \bar{\mathbf{x}}_{T-N}\|_{\mathbf{P}_{T-N}}^2 \quad (14)$$

式中: Φ_{T-N}^* 为 $T-N$ 时刻优化指标的最优值; $\bar{\mathbf{x}}_{T-N}$ 为 $T-1$ 时刻对 $T-N$ 时刻状态的估计; \mathbf{P}_{T-N} 为估计误差的权重矩阵。

由于 Φ_{T-N}^* 为常数,因此计算到达代价实际上就是计算 $\bar{\mathbf{x}}_{T-N}$ 和 \mathbf{P}_{T-N} 。目前多数文献多采用卡尔曼滤波及其误差协方差矩阵更新公式计算到达代价,但该方法未利用滚动时域估计的结果,不利于估计精度和速度的提高。本文从优化问题求解的角度出发,给出了一种基于 QR 分解的到达代价计算方法。

2.2 基于 QR 分解的到达代价计算方法

在 $T+1$ 时刻,有

$$\Phi_{T+1}(\mathbf{x}_{T-N+1}, \{\mathbf{w}_k\}) = C_{T-N+1}(\mathbf{x}_{T-N+1}) + \sum_{k=T-N+1}^{T+1} \|\mathbf{v}_k\|_{\mathbf{R}^{-1}}^2 + \sum_{k=T-N+1}^T \|\mathbf{w}_k\|_{\mathbf{Q}^{-1}}^2 = C_{T-N}(\mathbf{x}_{T-N}) + \|\mathbf{v}_{T-N}\|_{\mathbf{R}^{-1}}^2 + \|\mathbf{w}_{T-N}\|_{\mathbf{Q}^{-1}}^2 + \sum_{k=T-N+1}^{T+1} \|\mathbf{v}_k\|_{\mathbf{R}^{-1}}^2 + \sum_{k=T-N+1}^T \|\mathbf{w}_k\|_{\mathbf{Q}^{-1}}^2 \quad (15)$$

因此,利用式(14)所示函数估计到达代价,

根据前向动态规划原理可得

$$C_{T-N+1}(\mathbf{x}_{T-N+1}) = \min_{\mathbf{x}_{T-N}} (\|\mathbf{x}_{T-N} - \bar{\mathbf{x}}_{T-N}\|_{\mathbf{P}_{T-N}}^2 + \|\mathbf{v}_{T-N}\|_{\mathbf{R}^{-1}}^2 + \|\mathbf{w}_{T-N}\|_{\mathbf{Q}^{-1}}^2) \quad (16)$$

s. t. $\begin{cases} \mathbf{w}_{T-N} = \mathbf{x}_{T-N+1} - \mathbf{F}(\mathbf{x}_{T-N}, \mathbf{u}_{T-N}) \\ \mathbf{v}_{T-N} = \mathbf{y}_{T-N} - \mathbf{H}(\mathbf{x}_{T-N}) \end{cases}$

对权重矩阵 \mathbf{P}_{T-N} 、 \mathbf{R}^{-1} 和 \mathbf{Q}^{-1} 作 Cholesky 分解,有

$$\begin{cases} \mathbf{P}_{T-N} = \bar{\mathbf{P}}_{T-N}^T \bar{\mathbf{P}}_{T-N} \\ \mathbf{R}^{-1} = (\bar{\mathbf{R}}^{-1})^T \bar{\mathbf{R}}^{-1} \\ \mathbf{Q}^{-1} = (\bar{\mathbf{Q}}^{-1})^T \bar{\mathbf{Q}}^{-1} \end{cases} \quad (17)$$

利用式(17),式(16)可写为

$$C_{T-N+1}(\mathbf{x}_{T-N+1}) = \min_{\mathbf{x}_{T-N}} \left\| \begin{bmatrix} \bar{\mathbf{P}}_{T-N}(\bar{\mathbf{x}}_{T-N} - \mathbf{x}_{T-N}) \\ \bar{\mathbf{R}}^{-1}(\mathbf{y}_{T-N} - \mathbf{H}(\mathbf{x}_{T-N})) \\ \bar{\mathbf{Q}}^{-1}(\mathbf{x}_{T-N+1} - \mathbf{F}(\mathbf{x}_{T-N}, \mathbf{u}_{T-N})) \end{bmatrix} \right\|_2 \quad (18)$$

式中: $\|\mathbf{z}\|_2$ 表示 $\mathbf{z}^T \mathbf{z}$ 。

为得到式(18)所示最优化问题的解析解,在滚动时域估计得到的最优估计 \mathbf{x}_{T-N}^* 处对非线性函数 \mathbf{F} 和 \mathbf{H} 线性化:

$$\mathbf{x}_{T-N+1} = \mathbf{F}(\mathbf{x}_{T-N}^*) + \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \Big|_{\mathbf{x}_{T-N}^*} (\mathbf{x}_{T-N} - \mathbf{x}_{T-N}^*) = \bar{\mathbf{x}} + \mathbf{F}_x \mathbf{x}_{T-N} \quad (19)$$

式中:

$$\begin{cases} \bar{\mathbf{x}} = \mathbf{F}(\mathbf{x}_{T-N}^*) - \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \Big|_{\mathbf{x}_{T-N}^*} \mathbf{x}_{T-N}^* \\ \mathbf{F}_x = \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \Big|_{\mathbf{x}_{T-N}^*} \end{cases} \quad (20)$$

同理可得

$$\mathbf{y}_{T-N} = \mathbf{H}(\mathbf{x}_{T-N}^*) + \frac{\partial \mathbf{H}}{\partial \mathbf{x}} \Big|_{\mathbf{x}_{T-N}^*} (\mathbf{x}_{T-N} - \mathbf{x}_{T-N}^*) = \bar{\mathbf{H}} + \mathbf{H}_x \mathbf{x}_{T-N} \quad (21)$$

式中:

$$\begin{cases} \bar{\mathbf{H}} = \mathbf{H}(\mathbf{x}_{T-N}^*) - \frac{\partial \mathbf{H}}{\partial \mathbf{x}} \Big|_{\mathbf{x}_{T-N}^*} \mathbf{x}_{T-N}^* \\ \mathbf{H}_x = \frac{\partial \mathbf{H}}{\partial \mathbf{x}} \Big|_{\mathbf{x}_{T-N}^*} \end{cases} \quad (22)$$

将式(19)和式(21)代入式(18)可得

$$C_{T-N+1}(\mathbf{x}_{T-N+1}) = \min_{\mathbf{x}_{T-N}} \left\| \begin{bmatrix} \bar{\mathbf{P}}_{T-N} \bar{\mathbf{x}}_{T-N} \\ \bar{\mathbf{R}}^{-1}(\mathbf{y}_{T-N} - \bar{\mathbf{H}}) \\ -\bar{\mathbf{Q}}^{-1} \bar{\mathbf{x}} \end{bmatrix} \right\|_2 = \left\| \begin{bmatrix} \bar{\mathbf{P}}_{T-N} & \mathbf{0} \\ \bar{\mathbf{R}}^{-1} \mathbf{H}_x & \mathbf{0} \\ \bar{\mathbf{Q}}^{-1} \mathbf{F}_x & -\bar{\mathbf{Q}}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{T-N} \\ \mathbf{x}_{T-N+1} \end{bmatrix} \right\|_2 \quad (23)$$

令

$$\begin{cases} X = \begin{bmatrix} \mathbf{x}_{T-N} \\ \mathbf{x}_{T-N+1} \end{bmatrix} \\ A = \begin{bmatrix} \bar{P}_{T-N} & \mathbf{0} \\ \bar{R}^{-1}H_x & \mathbf{0} \\ \bar{Q}^{-1}F_x & -\bar{Q}^{-1} \end{bmatrix} \\ b = \begin{bmatrix} \bar{P}_{T-N}\bar{\mathbf{x}}_{T-N} \\ \bar{R}^{-1}(y_{T-N} - \bar{H}) \\ -\bar{Q}^{-1}\bar{\mathbf{x}} \end{bmatrix} \end{cases} \quad (24)$$

式(23)所示的到达代价计算问题转换为如下所示的最小二乘问题:

$$C_{T-N+1}(\mathbf{x}_{T-N+1}) = \min_{\mathbf{x}_{T-N}} \|\mathbf{b} - A\mathbf{X}\|_2^2 \quad (25)$$

为求解式(25)所示的最小二乘问题,对 A 作 QR 分解。

$$A = L \begin{bmatrix} R_1 & R_{12} \\ \mathbf{0} & R_2 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad (26)$$

式中: L 为正交矩阵; R_1 和 R_2 为上三角矩阵; R_{12} 为矩阵。

此外为表示方便,设

$$L^T \mathbf{b} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} \quad (27)$$

将式(26)和式(27)代入式(25)可得

$$\begin{aligned} C_{T-N+1}(\mathbf{x}_{T-N+1}) &= \min_{\mathbf{x}_{T-N}} \|\mathbf{b} - A\mathbf{X}\|_2^2 = \\ &= \min_{\mathbf{x}_{T-N}} \left\| \begin{bmatrix} \mathbf{b} \\ L^T \mathbf{b} \end{bmatrix} - \begin{bmatrix} R_1 & R_{12} \\ \mathbf{0} & R_2 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{T-N} \\ \mathbf{x}_{T-N+1} \end{bmatrix} \right\|_2^2 = \\ &= \min_{\mathbf{x}_{T-N}} \left(\left\| \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} - \begin{bmatrix} R_1 & R_{12} \\ \mathbf{0} & R_2 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{T-N} \\ \mathbf{x}_{T-N+1} \end{bmatrix} \right\|_2^2 + \right. \\ &\quad \left. \left\| c_1 - (R_1 \mathbf{x}_{T-N} + R_{12} \mathbf{x}_{T-N+1}) \right\|_2^2 \right) \quad (28) \end{aligned}$$

式(28)中的变量仅有 \mathbf{x}_{T-N} , 因此该最小二乘问题的解为

$$\mathbf{x}_{T-N} = R_1^{-1}(c_1 - R_{12} \mathbf{x}_{T-N+1}) \quad (29)$$

将式(29)代入式(28)可得

$$C_{T-N+1}(\mathbf{x}_{T-N+1}) = \|c_3\|_2^2 + \|\mathbf{x}_{T-N+1} - R_2^{-1}c_2\|_{R_2^T}^2 \quad (30)$$

与式(14)比较即可得到到达代价的更新方程:

$$\begin{cases} \bar{\mathbf{x}}_{T-N+1} = R_2^{-1}c_2 \\ \mathbf{P}_{T-N+1} = R_2^T R_2 \end{cases} \quad (31)$$

综上,到达代价的更新计算方法如下:

步骤1 分别对权重矩阵 P_{T-N} 、 R^{-1} 和 Q^{-1} 作 Cholesky 分解,得到 \bar{P}_{T-N} 、 \bar{R}^{-1} 和 \bar{Q}^{-1} 。

步骤2 线性化函数 F 和 H 。

$$\begin{cases} \mathbf{x}_{T-N+1} = \bar{\mathbf{x}} + F_x \mathbf{x}_{T-N} \\ \mathbf{y}_{T-N} = \bar{H} + H_x \mathbf{x}_{T-N} \end{cases}$$

步骤3 构造矩阵 A 和 b 。

$$\begin{cases} A = \begin{bmatrix} \bar{P}_{T-N} & \mathbf{0} \\ \bar{R}^{-1}H_x & \mathbf{0} \\ \bar{Q}^{-1}F_x & -\bar{Q}^{-1} \end{bmatrix} \\ b = \begin{bmatrix} \bar{P}_{T-N}\bar{\mathbf{x}}_{T-N} \\ \bar{R}^{-1}(y_{T-N} - \bar{H}) \\ -\bar{Q}^{-1}\bar{\mathbf{x}} \end{bmatrix} \end{cases}$$

步骤4 对 A 进行 QR 分解。

$$A = L \begin{bmatrix} R_1 & R_{12} \\ \mathbf{0} & R_2 \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

并计算

$$L^T \mathbf{b} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

步骤5 计算 $\bar{\mathbf{x}}_{T-N+1}$ 和 \mathbf{P}_{T-N+1} 。

$$\begin{cases} \bar{\mathbf{x}}_{T-N+1} = R_2^{-1}c_2 \\ \mathbf{P}_{T-N+1} = R_2^T R_2 \end{cases}$$

即完成对到达代价的更新。

2.3 滚动时域估计问题的求解

由 2.1 节和 2.2 节可知,弹性高速飞行器的状态估计问题转化为固定维数的优化问题。但这需要对系统状态和输入进行采样,因此对于弹性高速飞行器这一连续系统,需要选用合适方法完成系统方程的离散。

考虑到系统非线性较强且采样速率恒定,因此选用离散节点间距恒定且精度较高的多重打靶法。多重打靶法在等间距的离散节点上对状态量和控制量进行离散,认为相邻节点间控制量恒定,并加入节点处状态量相等的约束,从而实现连续系统的离散化,最终将状态估计问题转化为式(32)所示的非线性规划问题。

$$\begin{aligned} &\min_{\mathbf{x}_{T-N}, \{\mathbf{w}_k\}_{k=T-N}^{T-1}} \|\mathbf{x}_{T-N} - \bar{\mathbf{x}}_{T-N}\|_{P_{T-N}}^2 + \\ &\quad \sum_{k=T-N}^T \|\mathbf{v}_k\|_{R^{-1}}^2 + \sum_{k=T-N}^{T-1} \|\mathbf{w}_k\|_{Q^{-1}}^2 \quad (32) \\ &\text{s. t. } \begin{cases} \mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{w}_k \\ \mathbf{y}_k = \mathbf{H}(\mathbf{x}_k) + \mathbf{v}_k \end{cases} \end{aligned}$$

本文采用较为成熟的序列二次规划方法求解非线性规划问题式(32)。

综上,滚动时域估计方法的步骤如下:

步骤 1 初始化。给定 P_0 、 R^{-1} 、 Q^{-1} , 初始状态估计 \bar{x}_0 和滚动时域窗口长度 N 。

步骤 2 当 $k < N$ 时,利用 EKF 更新公式计算状态估计值 \bar{x}_k 和协方差矩阵 P_k 。

步骤 3 当 $k \geq N$ 时,利用序列二次规划方法求解非线性规划问题式(32)。

步骤 4 利用 2.2 节中给出的到达代价更新策略计算 \bar{x}_{T-N} 和 P_{T-N} 。

步骤 5 在 $k+1$ 时刻,获得量测 y_{k+1} ,构造新的量测数据集 $[y_{k-N+2} \dots y_k y_{k+1}]$,返回步骤 3。

3 仿真分析

设窗口长度 $N = 15$, 采样周期 $\Delta t = 15$ s, 量测噪声协方差矩阵为 $R = \text{diag}([2.5 \times 10^{-5} \quad 2.5 \times 10^{-5}])$, 系统噪声协方差矩阵为 $Q = \text{diag}([1 \times 10^{-6} \quad 1 \times 10^{-2} \quad 1 \times 10^{-6} \quad 2.5 \times 10^{-3} \quad 0.36])$ 。系统输入信号如图 1 所示,为了验证不同输入下方法的性能,在前 25 s 为正弦信号,后 25 s 无输入。此外,为了验证滚动时域估计方法对参数的估计性能,对一阶弹性模态固有频率加入了正弦扰动信号。

在量测数据和估计初值相同的情况下,分别采用 EKF、EKF 更新到达代价的滚动时域估计(MHE-EKF)和 QR 分解更新到达代价的滚动时域估计(MHE-QR)对弹性高速飞行器的状态和一阶弹性模态的固有频率进行估计。

在传感器按照方案 1 布置的情况下,分别对上述 3 种方法进行 100 次 Monte Carlo 仿真。估计结果的均方根误差(Root Mean Square Error, RMSE)如图 2 所示。

由图 2(a)~(d)可知,2 种滚动时域估计方法的均方根误差均明显小于 EKF;且受益于优化

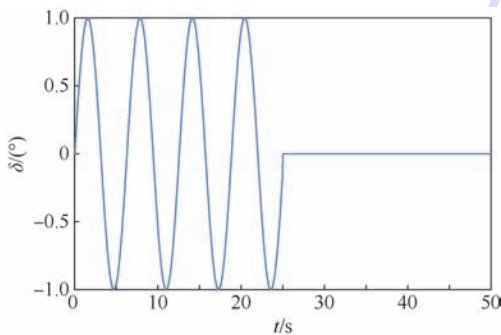


图 1 输入信号
Fig. 1 Input signal

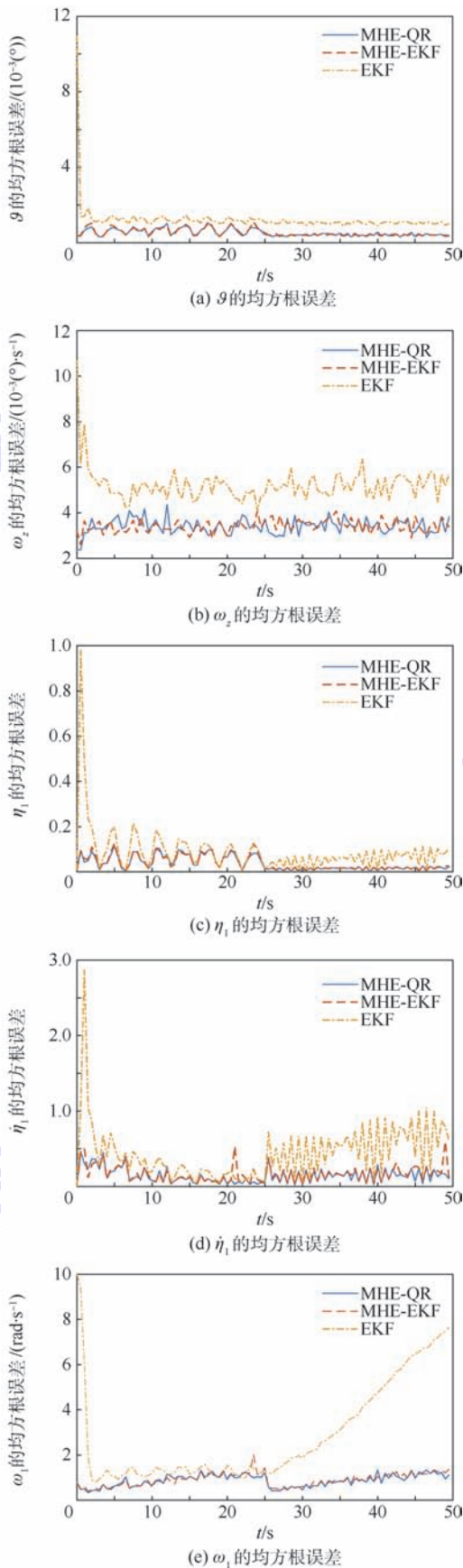


图 2 EKF、MHE-EKF 和 MHE-QR 方法估计结果的均方根误差
Fig. 2 RMSE of estimation results of EKF, MHE-EKF and MHE-QR methods

的思路,滚动时域估计方法的收敛速度快于EKF。由图2(e)可知,在系统输入变为0以后,EKF对弹性模态频率的估计结果出现了发散现象,并导致弹性模态的均方根误差大幅波动;而滚动时域估计并未受到系统输入的影响。以上结果均说明滚动时域估计方法的性能优于EKF。此外,由图2(d)可知,EKF更新到达代价的情况下, $\hat{\eta}_1$ 的均方根误差在某几个时刻存在突然增大的情况,这是由于其更新到达代价存在偏差导致优化结果误差增大,而QR分解更新到达代价的滚动时域估计方法未出现这种现象,表明QR分解更新到达代价的效果更好。

利用MHE-QR方法分别在传感器布置方案1(MHE-QR1)和传感器布置方案2(MHE-QR2)下进行100次Monte Carlo仿真,估计结果的均方根误差如图3所示。

由图3(a)、(b)可知,对于 ϑ 和 ω_z ,方案2与方案1的估计结果类似。由图3(c)~(e)可知,方案2对 η_1 、 $\hat{\eta}_1$ 和 ω_1 的估计结果好于方案1。这是由于方案2的传感器布置利用了飞行器设计模型信息,从而引入了有效信息预估,因而进一步提升了估计效果。

表1为上述方法的均方根误差均值。为验证估计弹性模态频率的必要性,表1还显示了只估计状态量的滚动时域估计方法(MHE-S)和EKF(EKF-S)的均方根误差均值。

由表1可知,MHE-QR和MHE-EKF的精度明显高于EKF。且MHE-QR对 ω_1 的估计精度最高,MHE-EKF的估计误差大于MHE-QR,而EKF对 ω_1 的估计出现了较大的误差。这是由于在滚动时域估计中, ω_1 只是一个优化变量,而噪声项只是提供了改变 ω_1 取值的手段,对 ω_1 的估计并不依赖具体的模型;而EKF将 ω_1 看作状态量,对其的估计精度依赖于模型的准确度,但这类参数的变化并不存在具体的模型,因此EKF很难准确估计时变的参数。此外,只估计状态量的MHE-S和EKF-S的误差均明显大于其他3种同时估计参数和状态量的方法,说明对变化的参数进行估计是十分必要的。

表2显示了3种方法的计算耗时,仿真在Windows 10系统(CPU为i5-7400,主频为3.00 GHz)中MATLAB R2017a环境下完成。EKF的耗时明显短于MHE-QR和MHE-EKF。MHE-QR的平均计算耗时短于MHE-EKF,且MHE-QR的最大计算耗时低于采样周期,具备应用潜力;而MHE-EKF虽然平均计算耗时低于采样周期,但最

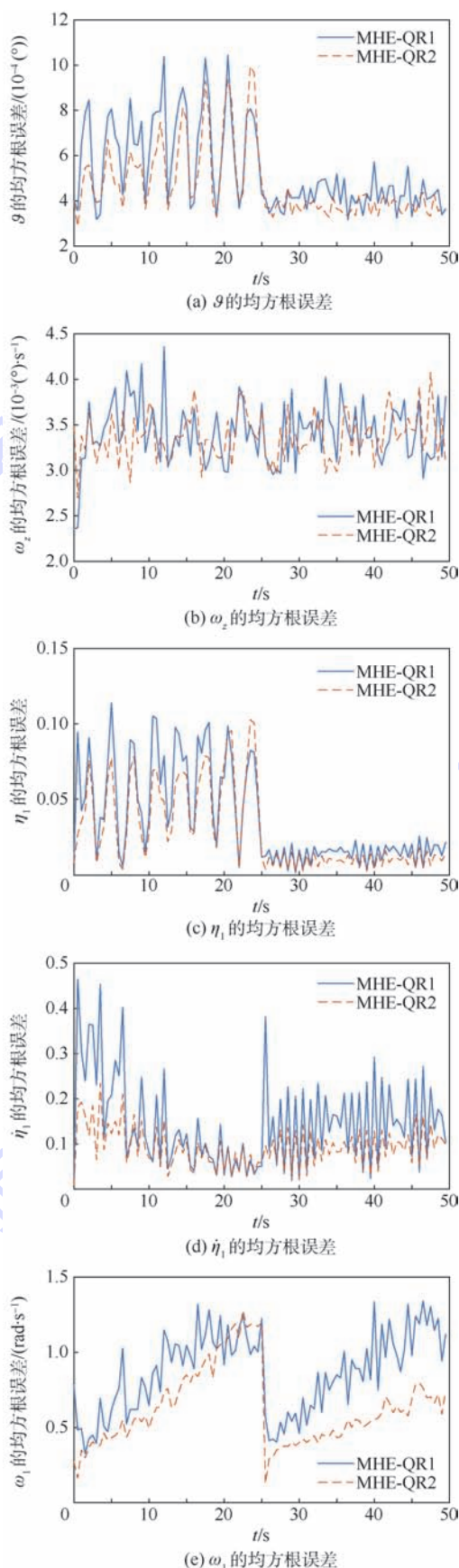


图3 不同方案时MHE-QR方法估计结果的均方根误差

Fig. 3 RMSE of estimation results of different schemes using MHE-QR method

表 1 不同方法估计结果的均方根误差均值

Table 1 Average RMSE mean values of estimation results of different methods

方法	RMSE				
	$\vartheta / (10^{-3} (^\circ))$	$\omega_z / (10^{-3} (^\circ) \cdot s^{-1})$	η_1	$\dot{\eta}_1$	$\omega_1 / (\text{rad} \cdot s^{-1})$
MHE-QR1	0.536	3.4	0.037 7	0.15	0.87
MHE-QR2	0.472	3.3	0.029 6	0.097	0.62
MHE-EKF	0.548	3.4	0.039 0	0.17	0.90
EKF	1.2	5.2	0.088 3	0.47	2.86
MHE-S	3.2	5.7	0.39	1.41	
EKF-S	3.5	7.0	0.37	2.47	

表 2 不同方法的计算耗时

Table 2 Run time of different methods

方法	平均时间/ (10^{-2} s)	最大时间/ (10^{-2} s)
MHE-QR1	2.44	4.74
MHE-QR2	2.35	4.78
MHE-EKF	2.48	7.56
EKF	0.66	1.27

大计算耗时高于采样周期,表明 EKF 更新到达代价在计算效率上不如 QR 分解。

4 结 论

本文提出了一种基于 QR 分解的到达代价更新方法,并将其用于弹性高速飞行器的滚动时域估计中,实现了状态/参数联合估计。

1) 状态/参数联合估计方法的精度远高于只估计状态的方法。由于弹性高速飞行器弹性模式的固有频率并非常数,会随飞行器状态变化而变化,因此对其进行在线估计是非常必要的,能够有效提高状态估计的精度。

2) 滚动时域估计的精度明显高于 EKF。相较于传统的 EKF 更新方法,QR 分解更新到达代价在精度类似的同时,具有更快的计算速度(最大计算耗时优于 EKF)。这得益于 QR 分解更新到达代价的策略利用了滚动时域估计的结果,形成了反馈机制,并通过直接求解优化问题更新到达代价。

3) 传感器采用布置方案 2 时的滚动时域估计结果好于布置方案 1。这是由于方案 2 通过引入有效信息预估而进一步提升了估计效果。

4) QR 分解更新到达代价的滚动时域估计方法的最长计算耗时低于采样速率,具有实际应用的潜力,后续应继续研究更快的优化算法,提高计算速度。

参考文献 (References)

[1] 张超凡,宗群,董琦,等.高超声速飞行器模型及控制若干问题综述[J].信息与控制,2017,46(1):90-102.
ZHANG C F,ZONG Q,DONG Q,et al. A survey of models and control problems of hypersonic vehicles [J]. Information and Control,2017,46(1):90-102 (in Chinese).

[2] 张伸,王青,董朝阳,等.基于跟踪微分器的高超声速飞行器减步控制[J].北京航空航天大学学报,2017,43(10):2054-2062.
ZHANG S,WANG Q,DONG C Y,et al. Reduced step control of hypersonic vehicle based on tracking differentiator [J]. Journal of Beijing University of Aeronautics and Astronautics,2017,43(10):2054-2062 (in Chinese).

[3] 赵贺伟,杨秀霞,沈如松,等.弹性高超声速飞行器预设性能精细姿态控制[J].兵工学报,2017,38(3):501-511.
ZHAO H W,YANG X X,SHEN R S,et al. Prescribed performance fine attitude control for aeroelastic hypersonic vehicle [J]. Acta Armamentarii,2017,38(3):501-511 (in Chinese).

[4] 吴云洁,宋嘉赞,刘晓东,等.推力矢量防空导弹伺服弹性的抑制[J].北京航空航天大学学报,2013,39(11):1480-1485.
WU Y J,SONG J Y,LIU X D,et al. Suppression of aeroservoelasticity in anti-aircraft missile using thrust vector [J]. Journal of Beijing University of Aeronautics and Astronautics,2013,39(11):1480-1485 (in Chinese).

[5] ABDOLLAHPOURI M,TAKACS G,ROHAL-LIKIV B. Real-time moving horizon estimation for a vibrating active cantilever [J]. Mechanical Systems and Signal Processing,2017,86(A):1-15.

[6] 焦志强,李卫华,王鹏.基于多模型与滚动时域估计的机动目标跟踪算法[J].空军工程大学学报(自然科学版),2016,17(2):15-20.
JIAO Z Q,LI W H,WANG P. A multi-model method of tracking maneuvering target based on multiple model and moving horizon estimation [J]. Journal of Air Force Engineering University (Natural Science Edition), 2016, 17 (2): 15-20 (in Chinese).

[7] RAO C V,RAWLINGS J B,LEE J H. Constrained linear state estimation-A moving horizon approach [J]. Automatica,2001,37(10):1619-1628.

[8] 陈伟,孙传杰,冯高鹏,等.基于滚动时域优化的旋转弹解耦控制器设计[J].北京航空航天大学学报,2018,44(4):717-724.
CHEN W,SUN C J,FENG G P,et al. Design of decoupling controller for spinning missile based on receding horizon optimal [J]. Journal of Beijing University of Aeronautics and Astronautics,2018,44(4):717-724 (in Chinese).

[9] RAO C V,RAWLINGS J B,MAYNE D Q. Constrained state estimation for nonlinear discrete-time systems: Stability and moving horizon approximations [J]. IEEE Transactions on Automatic Control,2003,48(2):246-258.

[10] KUHLM P,DIEHL M,KRAUS T,et al. A real-time algorithm for moving horizon state and parameter estimation [J]. Computers

- and Chemical Engineering, 2011, 35(1): 71-83.
- [11] VAN DER STEEN J, DIEHL M, AERTS C, et al. Spacecraft attitude estimation and sensor calibration using moving horizon estimation [J]. Journal of Guidance, Control, and Dynamics, 2013, 36(3): 734-742.
- [12] LIU A D, ZHANG W A, CHEN Z Q, et al. Moving horizon estimation for mobile robots with multirate sampling [J]. IEEE Transactions on Industrial Electronics, 2017, 64(2): 1457-1467.
- [13] 焦志强, 李卫华, 王鹏. 基于量测补偿的多传感器分布式滚动时域估计 [J]. 系统工程与电子技术, 2017, 39(5): 984-990.
- JIAO Z Q, LI W H, WANG P. Distributed moving horizon estimation for multi-sensors system based on measurements compensation [J]. System Engineering and Electronics, 2017, 39(5): 984-990 (in Chinese).
- [14] GAO W, YANG J, LIU J, et al. Moving horizon estimation for cooperative localization with communication delay [J]. The Journal of Navigation, 2015, 68(3): 493-510.
- [15] 赵国荣, 黄婧丽, 苏艳琴, 等. 基于滚动时域估计的飞行器姿态估计及三轴磁强计在线校正 [J]. 物理学报, 2015, 64(21): 210502.
- ZHAO G R, HUANG J L, SU Y Q, et al. Attitude estimation and three-axis magnetometer on-line calibration based on moving horizon estimation [J]. Acta Physica Sinica, 2015, 64(21): 210502 (in Chinese).
- [16] QU C C, HAHN J. Computation of arrival cost for moving horizon estimation via unscented Kalman filtering [J]. Journal of Process Control, 2009, 19(2): 358-363.
- [17] 孟中杰, 闫杰. 高超声速弹性飞行器振动模态自适应抑制技术 [J]. 宇航学报, 2011, 32(10): 2164-2168.
- MENG Z J, YAN J. Adaptive modal suppression for hypersonic aeroelastic vehicles [J]. Journal of Astronautics, 2011, 32(10): 2164-2168 (in Chinese).
- [18] FIORENTINI L, SERRANI A, BOLENDER M A, et al. Nonlinear robust adaptive control of flexible air-breathing hypersonic vehicles [J]. Journal of Guidance, Control, and Dynamics, 2009, 32(2): 401-416.
- [19] SANCHEZ Z, MURILLO M, GIOVANINI L. Adaptive arrival cost update for improving moving horizon estimation performance [J]. ISA Transactions, 2017, 68: 54-62.

作者简介:

陈尔康 男, 博士研究生。主要研究方向: 飞行器动力学与控制。

荆武兴 男, 博士, 教授, 博士生导师。主要研究方向: 飞行器动力学与控制。

State/parameter moving horizon estimation for elastic hypersonic vehicles

CHEN Erkang, JING Wuxing*, GAO Changsheng

(School of Astronautics, Harbin Institute of Technology, Harbin 150001, China)

Abstract: Considering the nonlinearity, uncertainty and rigid/elastic coupling of elastic hypersonic vehicles, a state/parameter joint estimation method based on QR decomposition and moving horizon estimation is proposed. First, this method transforms the state/parameter estimation problem into an optimization problem with fixed-number variables by introducing moving horizon strategy, and it can deal with the time-varying parameter estimation better than Kalman filter. Second, by utilizing the forward dynamic programming principle, the computation of arrival-cost is converted into a least-square problem that is solved by QR decomposition, and the arrival-cost update algorithm based on QR decomposition is given. In this way, the moving horizon estimation is based on optimization, and the feedback mechanism is introduced to improve the estimation accuracy and speed. The simulation results demonstrate that the accuracy of moving horizon estimation is obviously higher than that of extended Kalman filter, and the arrival-cost update strategy based on QR decomposition is better than the traditional arrival-cost update method based on the estimated error covariance in speed.

Keywords: hypersonic vehicle; elasticity; moving horizon estimation; state estimation; parameter estimation; QR decomposition

Received: 2018-05-15; Accepted: 2018-07-28; Published online: 2018-08-22 18:07

URL: kns.cnki.net/kcms/detail/11.2625.V.20180820.1755.011.html

Foundation item: National Natural Science Foundation of China (11572097)

* Corresponding author. E-mail: jingwuxing@hit.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2018.0296

数字锁相解调器的优化设计

李勇, 刘泽*, 赵鹏飞, 霍继伟, 林阳

(北京交通大学 电子信息工程学院, 北京 100044)



摘 要: 数字锁相解调器是电磁无损检测中弱信号处理的重要方法, 针对其性能优化, 设计一种引入卡尔曼滤波器的数字锁相解调器, 该解调器兼具了强抗噪能力及良好的动态跟踪响应特性。利用卡尔曼低通滤波器截止频率随迭代次数下降的特性, 结合零点频率, 设计了数字锁相解调器中的低通滤波环节, 实现了非常窄的锁相带通, 进而提高了抗击强噪声干扰能力。此外, 利用卡尔曼滤波器的迭代预测-更新特点, 实现了数字锁相解调器较小的响应时间, 进而改善了数字锁相解调器的动态跟踪响应特性。仿真和检测实验证明所设计数字锁相解调器具有较高的抗噪能力和良好的动态跟踪响应特性, 在电磁无损检测中能够精确、快速捕捉检测信号中包含的缺陷信息。

关键词: 数字锁相解调器; 卡尔曼滤波; 抗噪; 动态跟踪; 电磁无损检测

中图分类号: TG115.28; TP23

文献标识码: A **文章编号:** 1001-5965(2019)02-0299-10

锁相解调器是一种抗噪能力强、高稳定性的微弱信号检测技术^[1-9], 常应用于多种检测系统^[2-4]。一般情况下, 微弱信号往往隐没在强噪声背景当中, 而且其频率还会与谐波交叠在一起。如果不能有效快速地从强噪声中提取微弱信号, 检测系统的性能将会受到影响。另外, 在一些特殊的应用场景下, 往往还需要考虑其跟踪信号变化的特性。比如, 在电磁钢轨探伤的应用中, 缺陷会引起感应信号的快速突变, 假如不能及时响应跟踪信号的突变, 则检测系统无法有效地识别出钢轨的缺陷^[5-6]。而在实际现场当中, 电磁钢轨探伤往往会耦合外界复杂电磁环境的干扰以及器械运动引起的干扰, 则此时选择响应速度快的滤波器无法对信号进行有效地解调。由此, 兼具实时性和强抗噪的数字锁相解调器对钢轨探伤工程显得尤为重要。

根据数字锁相解调器的原理可知, 解调器的抗噪能力强, 响应特性主要受限于低通滤波器

(LPF) 环节, 而锁相环节仅对目标信号进行锁定移频, 对其没有影响。低通滤波器的系统响应是受到其阶数及截止频率的制约^[1,10]。因此, 需要选择合适的阶数及截止频率优化数字锁相解调器的解调性能。

针对动态响应的优化, 数字锁相解调器将均值滤波器(AF)作为其低通滤波器^[2,11]。根据文献[12]中提出的均值滤波器, 它是一种最简单的有限脉冲响应(FIR)低通滤波器, 时域上每一点的系数相同且等于采样点的倒数。信号在与该均值滤波器进行卷积运算时, 由于其零点频率恰好为锁相移频后的两倍频, 所以此时卷积运算结果为零。对于均值滤波器的卷积来说, 实际上是一个累加的过程, 因此相应所需的计算时间较少, 具有很好的动态响应特性。但是, 当加入强白噪声时, 解调器却无法准确地在强噪声中提取感应信号, 其抗噪能力却变得比较差。

而针对解调抗噪能力的优化, 数字锁相解调

收稿日期: 2018-05-22; 录用日期: 2018-07-27; 网络出版时间: 2018-09-06 16:47

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180905.1309.003.html

基金项目: 国家自然科学基金(61771041)

*通信作者. E-mail: zliu@bjtu.edu.cn

引用格式: 李勇, 刘泽, 赵鹏飞, 等. 数字锁相解调器的优化设计[J]. 北京航空航天大学学报, 2019, 45(2): 299-308.

LI Y, LIU Z, ZHAO P F, et al. Optimal design for digital phase-locked demodulator[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 299-308 (in Chinese).

利用级联阻容 (RC) 滤波器作为其低通滤波器^[5]。RC 滤波器是一种无限脉冲响应 (IIR) 低通滤波器,其优点是可以使用较小的阶数获得较好的抗噪能力且实现简单。此外,多个级联的 RC 滤波器要比单个 RC 滤波器具有更高的抗噪能力。但是在系统的响应阶段,多个级联的 RC 滤波器需要消耗过多的响应时间,导致动态跟踪特性很差,从而难以跟踪突变的信号变化。

已知状态模型和高斯噪声的前提下,卡尔曼滤波器是统计意义上最优的低通滤波器^[13-17]。为了实时处理信号和提高精度,文献[18-22]将卡尔曼滤波器引入数字锁相放大器。文献[18-20]利用卡尔曼滤波器模型建立了等效的数字锁相放大器,以实现实时处理信号准周期性的信号,并且在文献[18]中指出卡尔曼滤波器要比标准低通滤波器更具有优势;文献[21]利用传统锁相放大器对信号解调后引入卡尔曼滤波器,实现对信号的二次滤波,解决了光学膜厚监控系统中经锁相放大输出的监控信号精度较低、极值点附近变化不灵敏等问题;文献[22]也是将卡尔曼滤波器引入在数字锁相放大器之后,对信号进行二次滤波,并且利用 Allan 方差进行了分析,引入卡尔曼滤波器之后提升了原子磁力仪的灵敏度。

结合上述均值滤波器响应速度快的特点和 RC 滤波器抗噪能力强的特点,在保证解调器良好的动态跟踪情况下进一步提高抗噪能力,以实现快速在线电磁无损检测钢轨缺陷,本文对数字锁相解调器的滤波器进行了优化设计。此外,本文还对结合零点频率特点的卡尔曼滤波器被引入到数字锁相放大器中进行分析研究,并且给出调整动态跟踪特性和抗噪能力的方法。

1 数字锁相解调器原理

如图 1 所示,本文引入卡尔曼滤波器的数字锁相解调器与传统锁相解调器一致,主要分为相敏检测和卡尔曼滤波器两部分, $S(k)$ 、 $S_{rs}(k)$ 和 $S_{rc}(k)$ 分别为源信号、正弦参考信号和余弦参考信号, k 为当前时刻, $H_1(z)$ 和 $H_2(z)$ 分别为 2 个通道的低通滤波器, z 为离散域, R' 和 I' 分别为实部和虚部的直流项。

1.1 相敏检测优化及分析

根据文献[6]中设计的电磁传感器,对输入信号采样可得离散序列如下:

$$S(k) = A_0 \sin\left(\frac{2\pi k}{N} + \theta_0\right) \quad (1)$$

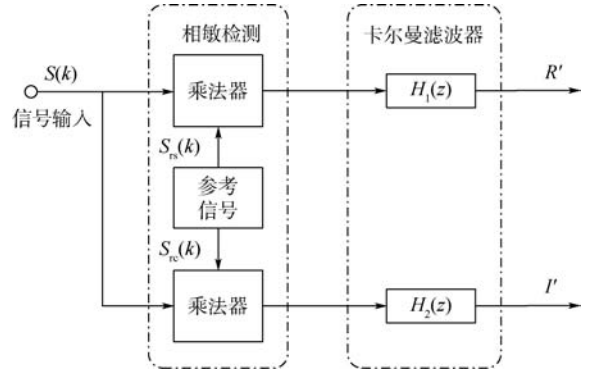


图 1 引入卡尔曼滤波的数字锁相解调器结构

Fig. 1 Structure of digital phase-locked demodulator with Kalman filter

式中: N 为采样频率对源信号采样一个整周期需要的点数; A_0 和 θ_0 分别为源信号的幅值和相位。由于参考信号是同频采样输入信号的相互正交序列,所以参考序列 $S_{rs}(k)$ 和 $S_{rc}(k)$ 计算如下:

$$\begin{cases} S_{rs}(k) = \sin\left(\frac{2\pi k}{N}\right) \\ S_{rc}(k) = \cos\left(\frac{2\pi k}{N}\right) \end{cases} \quad (2)$$

根据正余弦的周期性可知,参考序列满足:

$$\begin{cases} S_{rs}(k) = S_{rs}(k + nN) \\ S_{rc}(k) = S_{rc}(k + nN) \end{cases} \quad (3)$$

$$\begin{cases} S_{rs}(k) = S_{rc}\left(k + \frac{3N}{4}\right) \\ S_{rc}(k) = S_{rs}\left(k + \frac{N}{4}\right) \end{cases} \quad (4)$$

式(3)表明只需要预先存储一个周期的参考序列就可推算任意周期的参考序列。一般情况下,为了获得精确的结果, N 往往取值很大,即周期参考序列越长,则相对于实时计算正余弦参考序列,储存长参考序列会消耗更多的内存。但由于 CPU 直接读取存储在内存的参考值所需的时间要小于实时计算参考值所需的时间,所以在内存允许的条件下预先储存参考序列会节省 CPU 计算的时间。此外,由式(4)可知, $S_{rs}(k)$ 和 $S_{rc}(k)$ 可以相互推算出对方。为了适应内存相对不够的情况下,存储的参考序列可以只是一个整周期的 $S_{rs}(k)$ 或 $S_{rc}(k)$,以此优化相敏检测环节。

将输入的源信号采样值与相应参考序列相乘得

$$R(k) = S(k)S_{rs}(k) \quad (5)$$

$$R(k) = \frac{A_0}{2} \cos \theta_0 - \frac{A_0}{2} \cos\left(\frac{4\pi k}{N} + \theta_0\right) \quad (6)$$

$$I(k) = S(k)S_{rc}(k) \quad (7)$$

$$I(k) = \frac{A_0}{2} \sin \theta_0 + \frac{A_0}{2} \sin\left(\frac{4\pi k}{N} + \theta_0\right) \quad (8)$$

式中: $R(k)$ 和 $I(k)$ 分别为含有交流项和直流项的实部和虚部。

从式(5)~式(7)中可以看出,相敏检测的实质是乘法器,通过将源信号与参考信号相乘使得源信号的频谱迁移。相频谱迁移的结果向 $\pm 2\pi/N$ 方向移动。其中,式(6)和式(8)当中均含有相互正交的直流项,且该直流项包含了目标信号的幅值和相位等信息。在低通滤波器的作用下,由式(6)和式(8)可得

$$\begin{cases} R'(k) \approx \frac{A_0}{2} \cos \theta_0 \\ I'(k) \approx \frac{A_0}{2} \sin \theta_0 \end{cases} \quad (9)$$

则幅值和相位可得

$$\begin{cases} A_0 = 2\sqrt{(R'(k))^2 + (I'(k))^2} \\ \theta_0 = \arctan \frac{I'(k)}{R'(k)} \end{cases} \quad (10)$$

1.2 低通滤波器的优化设计

时域上,卡尔曼滤波是一种模型化的参数迭代估计方法。结合锁相环节的分析,低通滤波器的目的是滤除交流部分而获得直流项。因此,将2个通道的直流项作为卡尔曼滤波器的状态变量,则设当前时刻的状态变量为

$$\mathbf{X}[k] = (R'[k] \quad I'[k])^T \quad (11)$$

式中: $R'[k]$ 和 $I'[k]$ 分别为当前时刻双通道直流项的最优估计值。根据一阶自回归(AR)模型,则状态预测方程和其协方差预测方程可以分布表示为

$$\mathbf{X}[k|k-1] = \mathbf{F}\mathbf{X}[k-1] + \mathbf{W}[k-1] \quad (12)$$

$$\mathbf{P}[k|k-1] = \mathbf{F}\mathbf{P}[k-1]\mathbf{F}^T + \mathbf{Q}[k-1] \quad (13)$$

式中: \mathbf{F} 为状态转移矩阵,由于在一定时间内状态估计值(即直流部分)保持不变,所以令 \mathbf{F} 等于单位矩阵 \mathbf{E} ; $\mathbf{W}[k-1]$ 为均值为0,方差为 σ^2 的预测白噪声,也令 $\sigma^2 = 0$; $\mathbf{Q}[k-1]$ 为系统预测误差的方差,即 $\mathbf{Q}[k-1] = q\mathbf{E}$, q 为估计误差系数。

根据上述可知,获得实际测量值的过程为:先通过ADC对感应线圈上模拟信号采样得到的离散点,再经过相敏运算后得到双通道的 $R[k]$ 和 $I[k]$ 。目标获取的是两通道的直流部分,而其他均为交流干扰部分,且两者相互独立不影响。因此,可得观测方程为

$$\mathbf{S}_{\text{psd}}[k] = (R[k] \quad I[k])^T \quad (14)$$

$$\mathbf{Y}[k] = \frac{1}{N} \sum_{m=0}^{N-1} \mathbf{S}_{\text{psd}}[k-m] + \mathbf{V}[k] \quad (15)$$

式中: m 为序号; $\mathbf{S}_{\text{psd}}[k]$ 为双通道卡尔曼滤波器的输入信号; $\mathbf{V}[k]$ 为观测模型的测量噪声。

由于测量误差来自于ADC采样的量化误差以及电路的高斯白噪声,所以测量误差是稳定地服从高斯分布。通过观测到的 $\mathbf{Y}[k]$ 对预测值 $\mathbf{X}[k|k-1]$ 和 $\mathbf{P}[k|k-1]$ 进行校正,可以得到最优的估计结果

$$\mathbf{X}[k] = \mathbf{X}[k|k-1] + \mathbf{K}(\mathbf{Y}[k] - \mathbf{X}[k|k-1]) \quad (16)$$

$$\mathbf{P}[k] = (\mathbf{E} - \mathbf{K})\mathbf{P}[k|k-1] \quad (17)$$

其中:

$$\mathbf{K} = \mathbf{P}[k|k-1](\mathbf{P}[k|k-1] + \mathbf{R})^{-1} \quad (18)$$

$$\mathbf{R} = \mathbf{E}(\mathbf{V}[k]\mathbf{V}^T[k]) = r\mathbf{E} \quad (19)$$

式中: r 为测量误差系数; $\mathbf{E}()$ 为求期望的运算符; \mathbf{R} 为观测方程误差的协方差,是确定的高斯噪声且不随时间变化; \mathbf{K} 为卡尔曼增益,决定了预测值和测量值对实际最优估计的影响权值,其值在每次迭代过程中通过预测协方差被更新。

2 数字锁相解调器分析与评估

第1节中已经介绍了引入卡尔曼滤波的数字锁相解调器的优化设计。为了验证引入卡尔曼滤波器的数字锁相解调器既能改善解调速度又能使抗噪能力有所提高,本节将会对卡尔曼滤波器的系统进行分析,从而分析讨论数字锁相解调器引入卡尔曼滤波器的适用性及稳定性,并且进行仿真实验以评估解调器的速度以及抗噪能力。

2.1 卡尔曼滤波器的系统分析

假设卡尔曼增益不随迭代次数进行更新,根据式(16),离散系统的结构框图如图2所示,图中 z^{-1} 表示单位时间延迟。

从图2中可知,当仅考虑单通道时,卡尔曼滤波器离散系统的闭环传递函数为

$$H_i(z) = \frac{K_i}{1 - (1 - K_i)z^{-1}} \quad (20)$$

式中: $i=1$ 和 $i=2$ 分别表示数字锁相解调器的通道1或通道2。从式(20)可以看出,此时卡尔曼滤波器已退化成IIR数字滤波器,而IIR数字滤波器的优点恰好在于可以用较少的计算量获得较好的滤波效果。通道 i 的卡尔曼增益 K_i 在每次迭代过程都会被更新,由此可见卡尔曼滤波器实

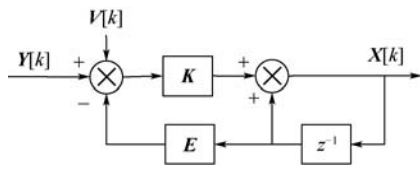


图2 卡尔曼滤波器的系统简化结构

Fig. 2 Simplified system structure of Kalman filter

实际上也是一个自适应的系统。

由于 IIR 滤波器其结构存在反馈,所以卡尔曼增益必然需要满足一定条件才能使得滤波器收敛。卡尔曼滤波器的收敛条件根据卡尔曼滤波器的稳定性求得,由式(20)可得卡尔曼滤波器的闭环特征方程为

$$1 - (1 - K_i)z^{-1} = 0 \tag{21}$$

根据离散系统稳定条件:所有闭环特征根在单位圆内,即

$$|z| = |1 - K_i| < 1 \tag{22}$$

则可得滤波器系统稳定的卡尔曼增益条件为: $0 < K_i < 2$ 。

2.2 卡尔曼滤波器的频域分析

令 $z = e^{j\omega T}$ 代入式(20),则将卡尔曼滤波器从 z 域转化成了 ω 域,即频域;其中, T 为采样间隔。频域形式 $H_i(\omega)$ 表示为

$$H_i(\omega) = \frac{K_i}{1 - (1 - K_i)e^{-j\omega T}} \tag{23}$$

图3是 $H_i(\omega)$ (仅考虑通道1)在不同增益情况下($0 < K_i < 2$)的幅值随频率的响应。图中可以发现,当 $K_i > 1$ 时,卡尔曼滤波器则不再对高频进行抑制反而增强,且随着卡尔曼增益增加而增大;当 $K_i = 1$ 时,卡尔曼滤波器是全通滤波器;当 $K_i < 1$ 时,卡尔曼滤波器就是低通滤波器,并且低通滤波器的截止频率随着卡尔曼增益减少而减小。此外,经过推导可知,当卡尔曼滤波是低通滤波器时,其截止频率 f_L 为

$$f_L = \frac{K_i}{2\pi} f_s \tag{24}$$

式中: f_s 为采样频率。

而根据式(15)可知,测量模型 $S_{psd}[k]$ 实际上是对锁定后的信号进行平均滤波,因此所设计的卡尔曼滤波器具有均值滤波器的性质。从图4中可以发现,均值滤波器随着 N 值的增大,其截

止频率会趋向于低频,对噪声的抑制能力增强。此外,第1个零点频率 $f_m = f_s/N$,也随 N 值增加而减小。由于在零点频率时均值滤波器可以将零点频率抑制为零,所以本文卡尔曼滤波器利用零点频率的这种性质所建立的测量模型能够更加有效地滤除交流部分。另外,前面已经提到, K_1 小于1时,卡尔曼滤波器表现为低通滤波器。从图4中也可以发现,当频率是非零点频率时,均值滤波器对噪声的抑制能力却不如低通滤波器滤波器(卡尔曼增益 K_1 为0.01时),尤其是在高频区域其抑制的能力明显弱于卡尔曼滤波器。并且,此时的低通滤波器其截止频率要比均值滤波器的截止频率小很多。因此,以 $S_{psd}[k]$ 为测量模型的卡尔曼滤波器对噪声的抗干扰能力更具有优势。

图5是卡尔曼增益 K_1 在每次迭代过程中变化的曲线。从图中可以看出,卡尔曼增益 K_1 随着迭代次数 N_i 下降且最终值与 r/q 的比值有关。由于卡尔曼增益随迭代次数减小,且卡尔曼增益始终小于1,所以卡尔曼滤波器可以看做是截止频率随迭代不断减小的低通滤波器,即可以通

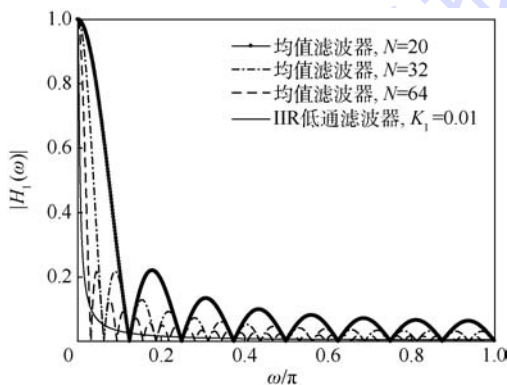


图4 IIR 低通滤波器和均值滤波器的频率响应
Fig.4 Frequency response between IIR low-pass filter and averaging filter

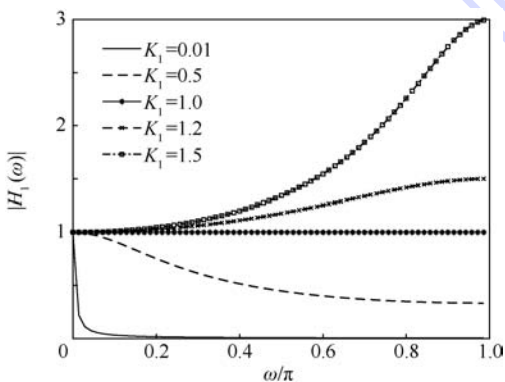


图3 不同卡尔曼增益的频率响应

Fig.3 Frequency response to different Kalman gains

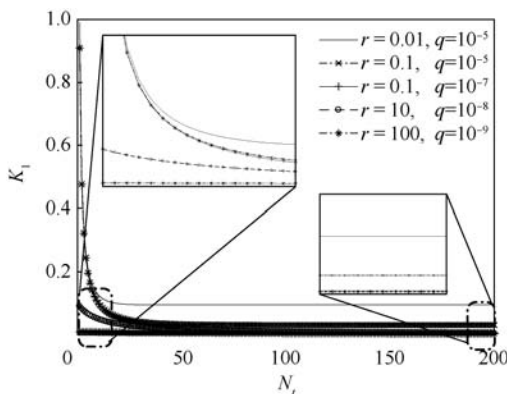


图5 卡尔曼增益随迭代次数的变化

Fig.5 Variation of Kalman gain with number of iterations

过不断实时迭代获得截止频率更低的低通滤波器。

众所周知,数字锁相解调器实质上是一带通滤波器,其中低通滤波器的截止频率越低则数字锁相解调器的通带越窄,解调器的抗噪能力越强。由于迭代次数持续增加使得低通滤波器的截止频率也持续减小,所以数字锁相解调器的通带也是随着迭代次数不断变窄的带通滤波器,其扫频特性所示如图 6 所示。因此,利用卡尔曼滤波器的数字锁相解调器是随实时迭代抗噪能力不断增强。

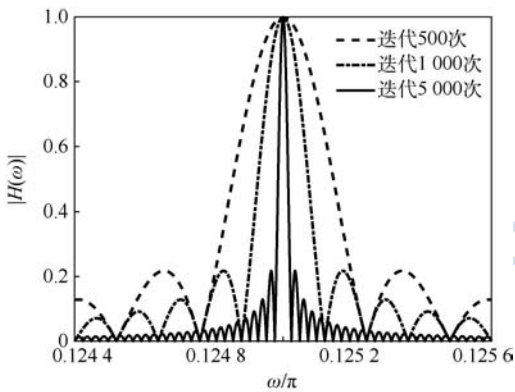


图 6 基于卡尔曼滤波的数字锁相解调器的扫频特性

Fig. 6 Frequency sweep characteristics of digital phase-locked demodulator based on Kalman filter

2.3 性能仿真测试与评估

引言中已经提到了,低通滤波器的响应速度和截止频率是一对矛盾。均值滤波器的优势在于快速性,但抗噪能力弱;而RC低通滤波器,其优

势在于抗噪能力强,但响应时间长、解调速度慢。在 2.2 节已经分析指出,卡尔曼滤波器是截止频率随着迭代次数增加而减小的 IIR 低通滤波器,其具有精度高的优势,而由于利用均值滤波器零点频率的性质建立了测量模型,此时卡尔曼滤波器又具有均值滤波器的特点。因此,本节将从三者的动态响应速度和抗噪能力进行仿真测试。

设仿真时激励信号的频率为 10 kHz,采样频率为 320 kHz,即 N 为 32,且其幅值为 1 V。另外,表 1 列出了 6 种不同情况下的仿真参数,其中白噪声是通过 MATLAB 的 awgn 函数产生并添加在噪声幅值最大为 0.1 V 的实测正弦信号。利用均值滤波器、卡尔曼滤波器和级联 RC 滤波器的数字锁相解调器分别对不同噪声的仿真实验进行解调并对比结果,如图 7 所示。其中,由于卡尔曼滤波器的测量模型实际上是对信号进行平均滤波,

表 1 不同噪声扰动大小以及 r/q 的值

Table 1 Different noise disturbance and value of r/q

编号	噪声幅值/V			SNR/dB	RC 截止频率/kHz	r/q
	DC	干扰频率为 50 Hz	干扰频率为 100 kHz			
1	0	0	0	30	4	10^4
2	1	1	1	30	0.7	10^4
3	1	1	1	10	4	10^4
4	1	1	1	1	0.7	10^4
5	1	1	1	1	0.7	10^7
6	1	1	1	1	0.7	10^9

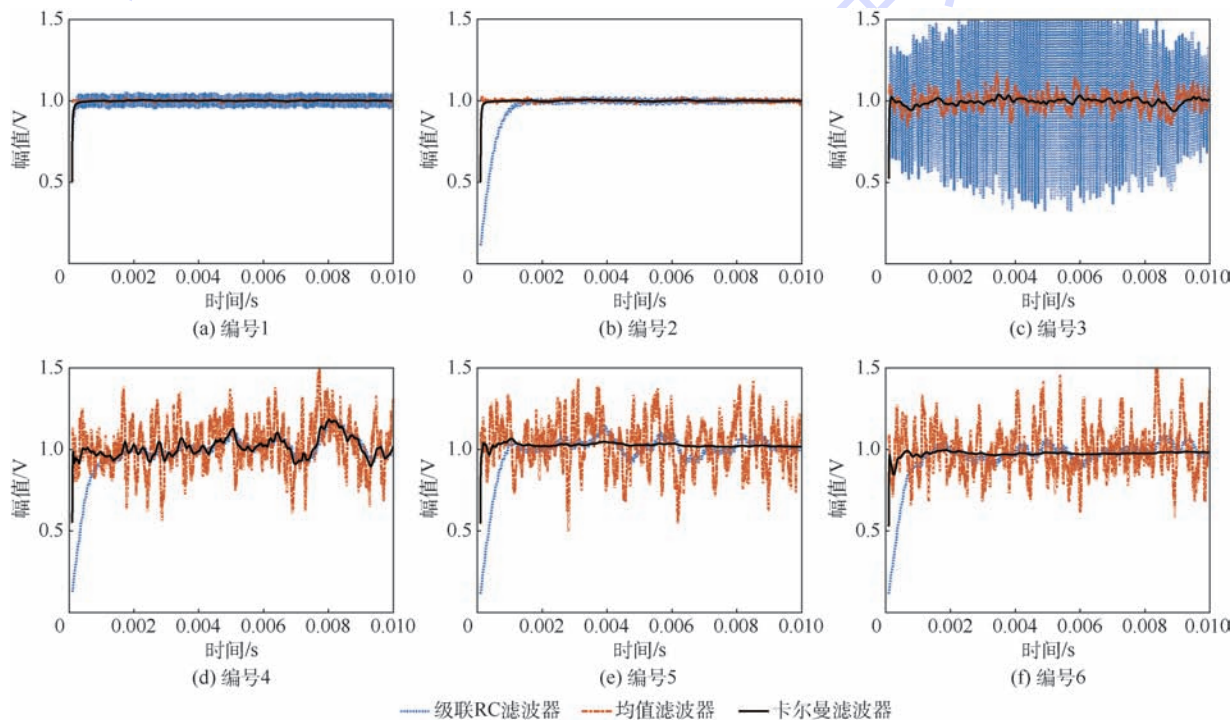


图 7 不同噪声下的解调

Fig. 7 Demodulation under different noises

即均值滤波器平均点数和卡尔曼滤波器中式(15)的累加点数 N 均设置为 32。

根据图 4 可知,均值滤波器的频率响应在 $N=32$ 时其截止频率为 4 kHz。从图 7 可以发现,当 RC 滤波器的截止频率设置为 4 kHz 时,三者的响应时间基本一致,但此时卡尔曼滤波器的抗噪能力要比级联 RC 滤波器和均值滤波器的抗噪能力强。也就是说,在响应时间一致(即动态跟踪性能一致)时,卡尔曼滤波器的抗噪能力是三者中最强的。由于截止频率高会降低 RC 滤波器对低频信号抗扰能力,因此响应时间一致时 RC 滤波器时三者中抗噪能力最差的。但是,当 RC 滤波器的截止频率减为 700 Hz 时,三者的抗噪能力达到一致,甚至当噪声增加时比其他 2 种解调器强。不过此时其响应时间却变长至 0.002 s,而卡尔曼滤波器的响应时间略微长于均值滤波器但几乎一致均远小于 0.002 s。也就是说,相同抗噪能力的同时,卡尔曼滤波器的动态跟踪能力与均值滤波器的一致,且优于级联 RC 滤波器。

上述表明利用卡尔曼滤波器的数字锁相解调器其动态跟踪性能优于级联 RC 滤波器的解调器,并且其抗噪能力强于均值滤波器的解调器。根据图 7 可知,卡尔曼滤波器的动态跟踪性能和抗噪能力跟 r/q 的比值有关。通过调节卡尔曼滤波器 r/q 的比值,抗噪能力会随着比值增大而得到改善,不过却会使其响应时间增大。因此,可以通过调节 r/q 的比值以达到兼具良好动态跟踪和抗噪能力的目的。

3 实验及验证

3.1 在线电磁钢轨缺陷检测实验

为了进一步验证引入卡尔曼滤波的数字锁相解调器的性能,本文利用电动平移台设计了一套在线钢轨缺陷检测系统,并且对比不同滤波器的数字锁相解调器跟踪缺陷信号的结果。钢轨裂纹缺陷检测系统如图 8 所示。图中,钢轨试板经过加工,在其中间位置处有一宽为 0.01 mm、深为 1 mm 的横向裂纹。而传感器则是采用了文献[6]中设计的电磁线圈传感器,利用中间线圈激励产生交变磁场,采集前后检测线圈的电压差分信号。采集模块则是利用了 4 路 16 位、4 MS/s 的 NI PXIe-6124 多功能 DAQ 设备,在信号连接器 BNC-2120 扩展下对线圈的电压差分信号进行采样收集。

实验中,令钢轨试件固定在平移台上,而线圈传感器则悬浮固定在试件上方不动,则试件将随

平移台以 0.3 m/s 速度移动与线圈传感器进行相对运动。缺陷的位置也设置在运动轨迹的中点,并且与试件运动的方向垂直,即垂直于线圈队列。另外,设置正弦激励信号的频率为 10 kHz,经过模拟放大电路放大后通入中间的激励线圈。对前后两线圈的感应信号进行差分放大,用频率为 320 kHz 采样差分信号,如图 9 所示。

从图 9 中可以看出,在未遇到缺陷时线圈上感应的电压幅值平稳,当传感器几何中心靠近缺陷时差分感应电压幅值变化为:增大—变小,而当传感器几何中心离开缺陷时变化为:减小—增大—减小。由于前后线圈的参数不一致,那么阻抗不平衡且对缺陷的敏感性也不一致。根据差分的原理,阻抗不平衡会导致在没有缺陷和缺陷在传感器中心位置时差分感应电压不为零。另外,线圈敏感性的差异导致对缺陷的感应幅值也不同。因此,在遇到缺陷时感应电压的变化不是按照缺陷位置空间对称。虽然线圈阻抗的不平衡导致没有缺陷时也存在电压差,但是不影响线圈遇到缺陷时的变化特征。

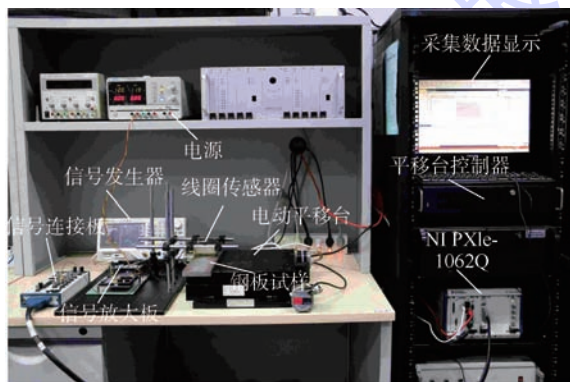


图 8 钢轨缺陷检测系统实验图

Fig. 8 Experimental picture of rail defect detection system

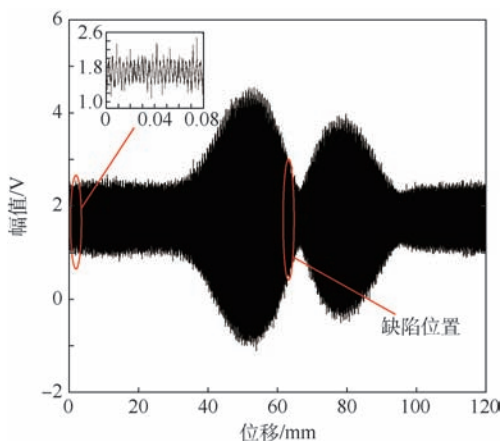


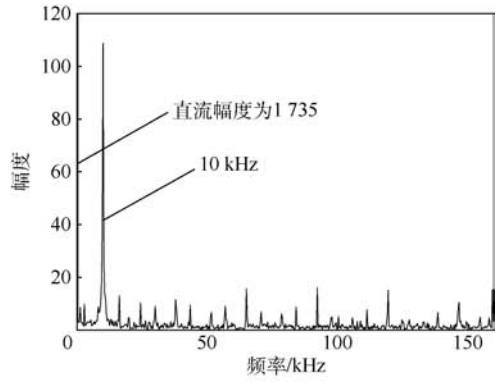
图 9 钢轨裂纹的实测数据

Fig. 9 Measured data of rail crack

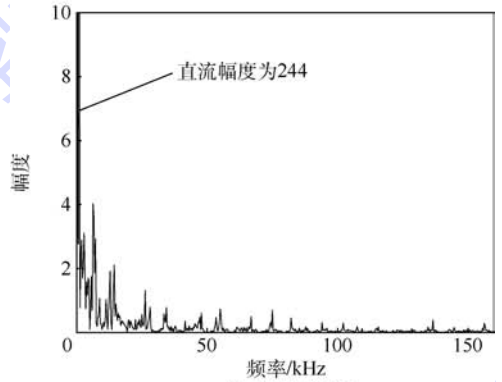
3.2 实测数据解调比较

图 10 显示的是分别引入 3 种数字锁相的解调器实时解调信号幅值的结果,其中 3 种滤波器的参数均与表 1 中编号 2 的设计参数相同。从图 10 中可以看出,3 种解调器均吻合地跟踪了信号的幅值变化。对比 3 种解调器解调的结果发现,在初始解调阶段引入卡尔曼滤波器和均值滤波器响应时间小于引入级联 RC 滤波器的响应时间,这就表明了实际解调中卡尔曼滤波器和均值滤波器的动态跟踪优于级联 RC 滤波器。

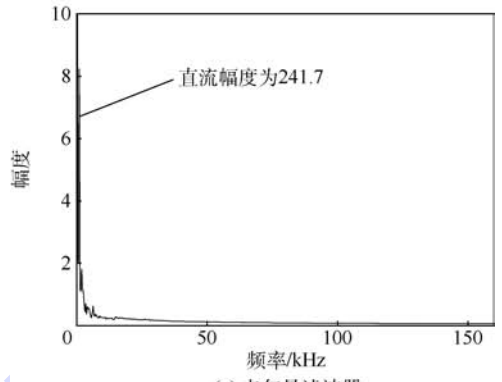
另外,从图 10 中也可以看出,利用卡尔曼滤波器解调的幅值是三者中波动最小。分别绘制实测数据和 3 种解调结果的幅频图,如图 11 所示。从图 11 中可知,实测数据的幅频图中包含频率为 10 kHz 的信号、直流偏置以及大量高频干扰噪声。实际上,实测数据幅频图中频率为 10 kHz 的信号为电磁线圈检测到的的差分感应信号,而其他频段的均为噪声干扰。而从经过 3 种解调器解调后的结果发现,3 种解调器解调后的主要成分均为直流部分,且依然含有的低频成分。这是因为解调器在对激励频率锁定滤波后为直流,但是由于缺陷引起了幅值波动,因此 3 种解调后的幅频图中主要分量为直流且低频段也有部分分量。此外,对比 3 种幅频图发现,利用均值滤波器的解调结果中低频段存在大量的毛刺且高频依然存在一定强度的分量。利用级联 RC 滤波器的解调结果,虽然在其高频部分光滑无分量,但是在其低频部分也依然存在一定强度,尤其在激励频点附近存在较强的分量。而利用卡尔曼滤波的解调结



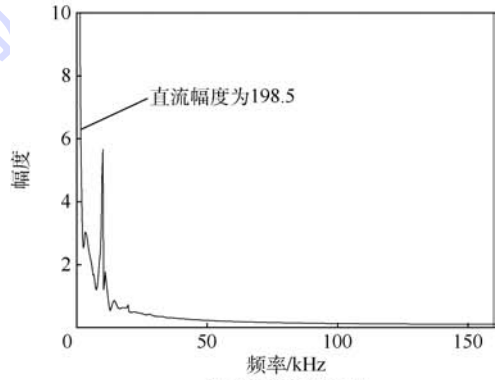
(a) 实测数据



(b) 均值滤波器



(c) 卡尔曼滤波器



(d) 级联RC滤波器

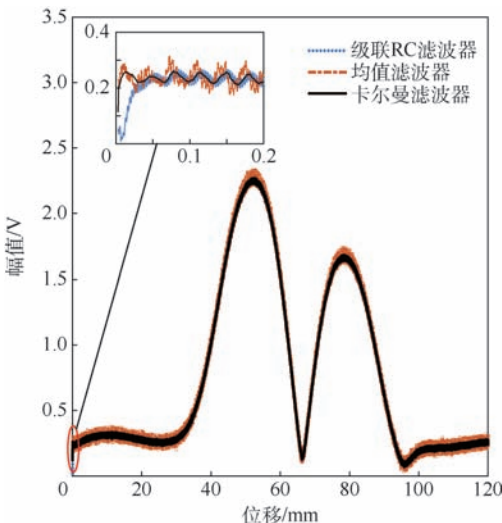


图 10 3 种滤波器的检测钢轨裂纹结果

Fig. 10 Detection results of rail crack with three filters

图 11 实测数据和 3 种解调器解调后的幅频图

Fig. 11 Measured data and amplitude-frequency diagram after demodulation by three demodulators

果几乎不存在高频分量,同时其低频部分也非常光滑基本无毛刺。由此,3种解调器解调结果的幅频分析进一步说明了卡尔曼滤波器时3种滤波器抗噪能力最优的。

3.3 r/q 比值的影响

根据性能仿真的结论,改变 r/q 的比值可以调节引入卡尔曼滤波的数字锁相器的抗噪和动态跟踪性能,以适应不同的噪声环境。增大比值有利于改善解调抗噪性能,但是却不利于解调器的动态跟踪,即比值越大其响应所需的时间越长,则动态跟踪性能越差。为了进一步说明了数字锁相解调器引入卡尔曼滤波后受 r/q 比值影响,进行了不同比值下的解调试验。如图 12 所示,当增大 r/q 比值,利用卡尔曼滤波的锁相解调抗噪能力也随着增高,但是其缺陷的峰值特征递减。此时,峰值特征减小与利用级联 RC 滤波解调器一样,增大比值使得系统的响应时间加长,未能及时响应跟踪缺陷突变引起的幅值变化,即动态响应变差。

因此,为了能够使引入卡尔曼滤波的数字锁相解调器提高识别缺陷的能力,需要实际噪声大小调节 r/q 合适的比值,在抗噪能力和动态跟踪性能之间做一个折中。另外从图 12 可知,如果解调后的幅值特征存在很大噪声以致于影响特征的识别,则此时需要将 r 的值增大,也就是 r/q 的比值变大。而当解调的幅值特征噪声很小,但是特征的峰值很小表明响应的时间过长,动态跟踪的能力受到影响,则此时可适当的增大 q 值,也就是使得 r/q 的值变小。

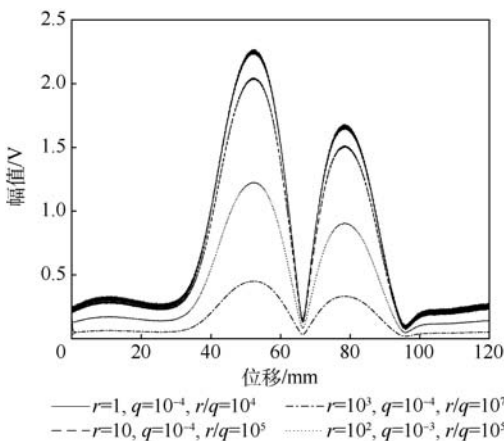


图 12 调节 r/q 比例的解调结果

Fig. 12 Demodulated results while adjusting r/q

4 结论

1) 本文介绍了数字锁相解调器引入卡尔曼滤波的优化设计,从时域和频域上分析,并发现引

入卡尔曼滤波的数字锁相解调器是通带随着迭代次数增加而变窄的带通滤波器。

2) 结合零点频率的性质,通过调节 r/q 的比值,数字锁相解调器引入卡尔曼滤波后能够高效、快速地从噪声中解调出弱的信号。与使用单一均值滤波器和级联 RC 滤波器的解调器相比,引入卡尔曼滤波后的数字锁相解调器在保证解调的动态跟踪性能的同时又能提高抗噪能力。

3) 钢轨裂纹在线电磁检测实验的结果表明,通过调节 r/q 的比值,引入卡尔曼滤波的数字锁相器能够很好地跟踪缺陷突变引起的幅值变化。

参考文献 (References)

- [1] 王化祥,曹章. 基于锁相放大器的低通滤波环节优化设计[J]. 仪器仪表学报,2005,26(7):684-689.
WANG H X, CAO Z. Optimization of low pass filter design in lock-in amplifier[J]. Chinese Journal of Scientific Instrument, 2005, 26(7): 684-689 (in Chinese).
- [2] 尹武良,王奔,王化祥. 电磁层析成像中基于半周期采样的数字解调方法[J]. 天津大学学报(自然科学与工程技术版),2011,44(12):1118-1123.
YIN W L, WANG B, WANG H X. Digital demodulation based on half-period sampling in electromagnetic tomography[J]. Journal of Tianjin University (Science and Technology), 2011, 44(12): 1118-1123 (in Chinese).
- [3] 王野,谭超,董峰. 多激励频率模式的磁感应层析成像系统[J]. 北京航空航天大学学报,2017,43(11):2331-2337.
WANG Y, TAN C, DONG F. Magnetic induction tomography system with multi-excitation frequency mode[J]. Journal of Beijing University of Aeronautics and Astronautics, 2017, 43(11): 2331-2337 (in Chinese).
- [4] 周正干,贺鹏飞,赵翰学,等. 钛合金蜂窝结构蒙皮脱焊缺陷锁相红外热成像检测[J]. 北京航空航天大学学报,2016,42(9):1795-1802.
ZHOU Z G, HE P F, ZHAO H X, et al. Detection of skin desoldering defect in Ti-alloy honeycomb structure using lock-in infrared thermography test[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(9): 1795-1802 (in Chinese).
- [5] LIU Z, ZHU L, KOFFMAN A, et al. Digital lock-in amplifier for precision audio frequency bridge[C] // 2012 Conference on Precision Electromagnetic Measurements. Piscataway, NJ: IEEE Press, 2012: 586-587.
- [6] LIU Z, LI W, XUE F, et al. Electromagnetic tomography rail defect inspection[J]. IEEE Transactions on Magnetics, 2015, 51(10): 1-7.
- [7] MASCIOTTI J M, LASKER J M, HIELSCHER A H. Digital lock-in detection for discriminating multiple modulation frequencies with high accuracy and computational efficiency[J]. IEEE Transactions on Instrumentation & Measurement, 2008, 57(1): 182-189.
- [8] 黄佳亮. 微弱信号检测的噪声和处理方法[J]. 仪器仪表与

- 分析监测,1995(1):38-54.
- HUANG J L. The noise and processing method of weak signal detecting[J]. Instrumentation Analysis Monitoring, 1995(1): 38-54(in Chinese).
- [9] 周浩敏,赵巧转,汤章阳. 谐振式微小型压力传感器数字闭环系统[J]. 北京航空航天大学学报, 2006, 32(11): 1312-1315.
- ZHOU H M, ZHAO Q Z, TANG Z Y. Closed loop systems of resonant small sensors and micro-sensors based on DSP technology[J]. Journal of Beijing University of Aeronautics and Astronautics, 2006, 32(11): 1312-1315(in Chinese).
- [10] SUN S, XU L, CAO Z, et al. Digital recursive demodulator based on Kalman filter[J]. IEEE Transactions on Instrumentation & Measurement, 2017, 66(12): 3138-3147.
- [11] 李刚,张丽君,林凌. 一种新型数字锁相放大器的设计及其优化算法[J]. 天津大学学报(自然科学与工程技术版), 2008, 41(4): 429-432.
- LI G, ZHANG L J, LIN L. Design of a new digital lock-in amplifier and its optimization algorithm[J]. Journal of Tianjin University(Science and Technology), 2008, 41(4): 429-432(in Chinese).
- [12] VAINIO O. Minimum-phase FIR filters for delay-constrained noise reduction[J]. IEEE Transactions on Instrumentation & Measurement, 1999, 48(6): 1100-1102.
- [13] CHENG Y. A linear algebraic approach to Kalman filtering [C]// International Conference on Measuring Technology and Mechatronics Automation. Piscataway, NJ: IEEE Press, 2010: 122-125.
- [14] COOPER W S. Use of optimal estimation theory, in particular the Kalman filter, in data analysis and signal processing[J]. Review of Scientific Instruments, 1986, 57(11): 2862-2869.
- [15] 薛明喜,杨扬,张晨睿,等. 基于自适应 Kalman 滤波的 SAW 测温数据纠错方法[J]. 仪器仪表学报, 2016, 37(12): 2766-2773.
- XUE M X, YANG Y, ZHANG C R, et al. Error correction method for SAW temperature measurement data based on adaptive Kalman filter[J]. Chinese Journal of Scientific Instrument, 2016, 37(12): 2766-2773(in Chinese).
- [16] 刘国海,李沁雪,施维,等. 动态卡尔曼滤波在导航试验状态估计中的应用[J]. 仪器仪表学报, 2009, 30(2): 396-400.
- LIU G H, LI Q X, SHI W, et al. Application of dynamic Kalman filtering in state estimation of navigation test[J]. Chinese Journal of Scientific Instrument, 2009, 30(2): 396-400(in Chinese).
- [17] 张猛,肖曦,李永东. 基于扩展卡尔曼滤波器的永磁同步电机转速和磁链观测器[J]. 中国电机工程学报, 2007, 27(36): 36-40.
- ZHANG M, XIAO X, LI Y D. Speed and flux linkage observer for permanent magnet synchronous motor based on EKF[J]. Proceedings of the CSEE, 2007, 27(36): 36-40(in Chinese).
- [18] COELHO R, ALVES D. Real-time lock-in amplifier implementation using a Kalman filter for quasi-periodic signal processing in fusion plasma diagnostics[J]. IEEE Transactions on Plasma Science, 2009, 37(1): 164-170.
- [19] COELHO R, ALVES D, HAWKES N, et al. Real-time data processing and magnetic field pitch angle estimation of the JET motional Stark effect diagnostic based on Kalman filtering[J]. Review of Scientific Instruments, 2009, 80(6): 164-168.
- [20] ALVES D, COELHO R, KLEIN A, et al. A real-time synchronous detector for the TAE antenna diagnostic at JET[J]. IEEE Transactions on Nuclear Science, 2010, 57(2): 577-582.
- [21] 赵汝进,马孜,姚远程,等. 基于广义卡尔曼滤波的光学膜厚监控信号处理[J]. 激光技术, 2007, 31(4): 412-415.
- ZHAO R J, MA Z, YAO Y C, et al. Monitor signal processing of optic thin film based on generalized Kalman filters[J]. Laser Technology, 2007, 31(4): 412-415(in Chinese).
- [22] 董玉辉. 卡尔曼滤波在原子磁力仪中的应用研究[D]. 哈尔滨:哈尔滨工程大学, 2016.
- DONG Y H. Application research of Kalman filter in optical atomic magnetometer[D]. Harbin: Harbin Engineering University, 2016(in Chinese).

作者简介:

李勇 男,博士研究生。主要研究方向:电磁无损检测及其信号处理。

刘泽 男,博士,教授。主要研究方向:电磁层析成像。

Optimal design for digital phase-locked demodulator

LI Yong, LIU Ze*, ZHAO Pengfei, HUO Jiwei, LIN Yang

(School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China)

Abstract: Digital phase-locked demodulator is an important method for weak signal processing in electromagnetic non-destructive testing. For its performance optimization, a digital phase-locked demodulator with Kalman filter is designed, which has strong anti-noise ability and good dynamic tracking features. By using the characteristic that the cut-off frequency of Kalman low-pass filter decreases with the number of iterations and combining with the zero frequency, we design a low-pass filter in a digital phase-locked demodulator and then realize very narrow phase-locked bandpass, which improves the ability of strong noise resistance. In addition, using the predictive-update feature of the Kalman filter, the smaller response time of the digital phase-locked demodulator is realized, thereby improving the dynamic tracking characteristics of the digital phase-locked demodulator. Simulation and detection experiments show that the designed digital phase-locked demodulator has high noise immunity and good dynamic tracking response. It can accurately and quickly capture the defect information contained in the detection signal in electromagnetic non-destructive testing.

Keywords: digital phase-locked demodulator; Kalman filter; anti-noise; dynamic tracking; electromagnetic non-destructive testing

Received: 2018-05-22; **Accepted:** 2018-07-27; **Published online:** 2018-09-06 16:47

URL: kns.cnki.net/kcms/detail/11.2625.V.20180905.1309.003.html

Foundation item: National Natural Science Foundation of China (61771041)

* **Corresponding author.** E-mail: zliu@bjtu.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2018.0316

基于改进型 Retinex 算法的雾天图像增强技术

张驰¹, 谭南林¹, 李响², 李国正^{1,*}, 苏树强¹

北京交通大学 机械与电子控制工程学院, 北京 100044; 2. 华东交通大学 交通运输与物流学院, 南昌 330013

摘 要: 为增强雾天图像的对比度及颜色和亮度恒常性, 提出了一种改进型 Retinex 算法雾天图像增强算法。使用改进的双边滤波器作为滤波函数, 在保持边缘信息的同时去除噪声的干扰; 并使用 S 型函数曲线对 Retinex 算法中对数域相减去除入射光分量的图像进行颜色恢复处理, 增强整幅图像的对比度和感知特性, 还原图像的色彩信息。实验结果表明, 所提的改进算法能有效提高雾天图像的清晰度和对比度, 相较原雾天图像清晰度提升约 200%, 标准差提升约 110%, 信息熵提升约 10%。同时, 可保持更加真实鲜艳的图像颜色, 计算复杂度较低, 满足实时性要求。

关键词: 图像增强; Retinex; 双边滤波; 边缘信息; 颜色恢复

中图分类号: TP751.1

文献标识码: A

文章编号: 1001-5965(2019)02-0309-08

雾霾天气下所拍摄的图像退化严重, 对比度降低, 细节信息损失, 色调偏移, 不利于分析与识别。对雾天图像进行增强处理, 可提高其可见度, 丰富图像的信息量, 改善图像的品质, 对图像的深入理解以及机器对图像进一步识别和深层次学习有很大的帮助。因此, 对雾天图像的增强处理方法展开研究, 在航空航天^[1]、交通监控^[2]和室外监测等领域有重要的实际工程意义。

现有常用的雾天图像增强算法主要有: 基于暗通道的去雾算法、直方图均衡 (Histogram Equalization, HE) 算法和基于色彩恒常性理论的 Retinex 算法^[3]。从实际应用角度来看, 这 3 类算法均有着各自的局限性: ①暗通道理论在去雾研究中得到了很多应用, 从大气散射模型出发, 对透视率进行细化, 有很好的去雾效果^[4]。结合暗通道理论以及 Retinex 算法对雾天图像进行增强, 可以有效复原图像的清晰度和对比度^[5-6]。基于偏振成像和暗通道技术的结合, 引入了动态偏差因

子, 不仅能有效地提高模糊图像的可见性, 而且能显著地保护图像的细节^[7]。基于 Retinex 算法和引导滤波对夜间图像进行增强, 根据图像的颜色特征, 利用引导滤波器对照明分量进行估计, 并将其应用于边缘保持平滑算子^[8]。该算法处理视觉效果好, 但时空复杂度较高, 且应用环境受限。②HE 算法通过分散灰度值、修改直方图分布来增强图像对比度。有使用显著直方图设计灰色层次映射方法对图像进行增强跟踪, 提升视觉效果^[9]。针对航天领域中无人机的应用条件, 统计直方图数据对图像分别进行白平衡处理和对比度增强处理, 使用自动色阶得到增强图像^[10]。为了改善图像质量并增强视觉效果, 提出基于非线性直方图变换与参数优化的降质图像对比度畸变校正方法^[11]。HE 算法虽然处理速度快, 客观评价参数较优, 但是增强后的图片会有失真现象, 亮度不均, 细节丢失, 主观评价差。③Retinex 是从图像中估计出光照, 得到增强图像, 主要包括单尺度

收稿日期: 2018-06-01; 录用日期: 2018-08-24; 网络出版时间: 2018-09-04 09:09

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180830.1613.006.html

基金项目: 国家自然科学基金 (61527812)

* 通信作者: E-mail: liguozheng@bjtu.edu.cn

引用格式: 张驰, 谭南林, 李响, 等. 基于改进型 Retinex 算法的雾天图像增强技术[J]. 北京航空航天大学学报, 2019, 45(2): 309-316. ZHANG C, TAN N L, LI X, et al. Foggy image enhancement technology based on improved Retinex algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 309-316 (in Chinese).

Retinex (Single Scale Retinex, SSR)、多尺度 Retinex (Multi-Scale Retinex, MSR)、带颜色恢复的多尺度 Retinex (Multi-Scale Retinex with Color Restoration, MSRCR) 等算法^[12]。使用空间自适应基于 α -范数的变分 Retinex 模型,利用亮通道先验和局部方差图生成正则化参数进行图像增强^[13]。采用 MSR 算法和双边滤波结合的方案对图像低频系数进行处理,采用软阈值滤波算法对图像高频系数进行处理进行图像增强^[14]。基于双边滤波和 MSR 算法结合离散小波变换和 Otsu 阈值分割算法进行图像增强^[15]。基于 MSR 算法,提出了一种平衡图像对比度和色彩一致性的显式多尺度表示,并引入直方图截断技术^[16]。为了恢复图像色彩,提出了基于高斯滤波和引导滤波的改进型 MSRCR 图像增强算法,该算法能有效地抑制噪声干扰,有效恢复图像色彩^[17]。基于改进的 MSR 算法,将 RGB 3 个通道分别进行非线性灰度拉伸,得到每个像素点光照阴影关系比例,对雾天图像进行增强^[18]。该算法能够有效提升雾天图像,对比度高,细节丰富,色感一致性好,颜色自然。但是由于其多尺度计算复杂度高,边缘位置模糊,缺乏亮度恒常性,颜色恢复算法简单,会出现反转颜色失真现象,对噪声会有加强效果,因而处理还并不十分理想。

可见,现有 3 类算法对退化图像增强在图像失真、损失细节、去除噪声干扰、运算速度、色彩还原等方面无法做到同时达到很好的效果。为此,本文提出了一种改进型 Retinex 算法,首先以改进的双边滤波法作为中心环绕滤波函数,能够保留边缘信息的同时去除噪声。其次将 S 型 sigmoid 函数应用于雾天图像的颜色还原与增强处理,能够减少计算复杂度并得到更佳的颜色信息。最后通过实验对比了暗通道、HE 和 MSRCR 算法的处理结果,验证了该算法的可行性和有效性。

1 Retinex 算法及双边滤波法

1.1 Retinex 算法

Retinex 算法基于颜色恒常知觉理论,认为得到的雾天图像 $I(x, y)$ 是由入射光分量 $L(x, y)$ 即光照图像和反射光分量 $R(x, y)$ 组成^[14],

$$I(x, y) = L(x, y)R(x, y) \quad (1)$$

将式(1)转换到对数空间进行计算,有利于人眼对图像信息的感知,转换数学公式如下:

$$\ln(R_i(x, y)) = \ln(I_i(x, y)) - \ln(G_i(x, y)I_i(x, y)) \quad (2)$$

式中: $G_i(x, y)$ 为中心环绕函数,现有 Retinex 算法常选用高斯函数,下标 i 为通道数。

Retinex 算法的核心是计算出来精确的照度分量和反射分量,在此基础上能够还原出更加真实的原始图像。对于雾天图片有显著的增强效果,所得到的去雾图像具有较高的局部对比度和较小的颜色失真。SSR 算法难以在细节增强与色感一致性两方面都有好的结果。因此提出 MSR 加入 3 个尺度,在不同雾天条件下都有表现较佳的尺度,但 2 种算法都会出现偏色现象。在 MSR 基础上加入颜色恢复得到 MSRCR 算法,消除图像颜色失真现象。

1.2 双边滤波法

在 Retinex 算法的中心环绕滤波算法中,目标点上的像素值通常是由其所位置上的周围的一个小局部邻居像素的值所决定,所以处理的效果是对整幅图像进行模糊处理。这会导致边缘锐化不足,细节信息丢失。而双边滤波的基本思路是同时考虑将要被滤波的像素点的空域信息和值域信息,与领域中心点距离近、亮度差异较小像素点对滤波影响大,距离远、差异大的像素点影响小,滤波函数为

$$h(x) = k^{-1}(x) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\xi) c(\xi, x) s(f(\xi), f(x)) d\xi \quad (3)$$

式中: $k(x)$ 为权重和,将结果归一化; ξ 为空间像素点; $f(x)$ 为像素值, x 为要求的点; c 为距离函数; s 为相似度函数。为此,考虑将空间高斯函数和值域高斯函数进行结合,使用双边滤波代替高斯滤波作为 MSRCR 中心环绕函数,从而保留边缘处距离相距近但是像素值差异大的像素点特性,进而保持雾天原图像中的边缘信息,改善光晕、过增强问题。

2 改进的 Retinex 算法

雾天的图像有着能见度较低、视觉模糊的特点,图像色彩偏向于灰白色。由于雾气浓度不同、景深和不同波长光线衰减不同,所以整幅图像雾天所造成的模糊程度是不同的。本文目标是在雾天影响衰减的图像恢复,还原色彩,增强整幅图像的对比度,提高图像的可视性。

针对现有 Retinex 算法存在的噪声处理差、增强时间成本高、颜色恢复性差等问题,本文提出如图 1 所示的算法流程, $B(X, Y)$ 为双三次插值结果, a_{ij} 为权重系数, $W(i)$ 和 $W(j)$ 分别为横坐标和纵坐标权重, $S(x)$ 为 sigmoid 函数。

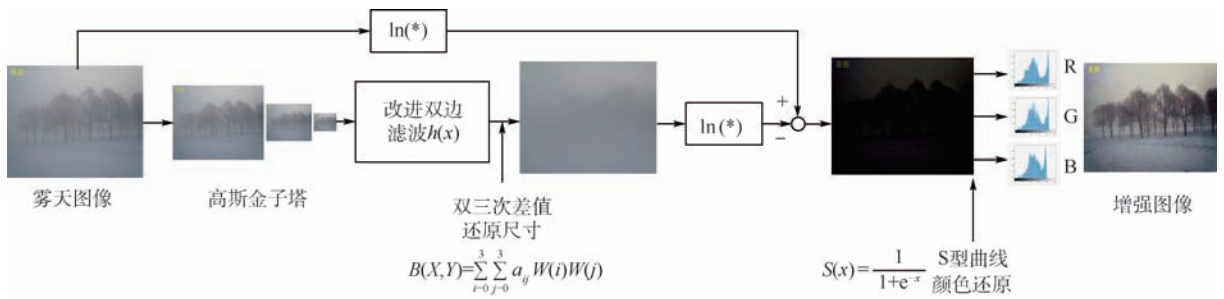


图 1 改进型 Retinex 算法流程图

Fig. 1 Flowchart of improved Retinex algorithm

首先,对 Retinex 算法中心环绕滤波算法进行改进,使用双边滤波法能够提升边缘的效果,同时提升对噪声的处理能力。具体算法实现首先将待处理的原始图像进行高斯金字塔向下采样,由卷积采样与偶数行采样组成,其公式为

$$G_{k+1}(i,j) = \sum_{m=-2}^2 \sum_{n=-2}^2 W(m,n) G_k(2i-m,2j-n) \quad (4)$$

式中: m 和 n 分别为行数和列数, k 为金字塔代数; $W(m,n)$ 为高斯卷积核。

求得滤波使用的核函数,基于空间距离和像素差值得到高斯权重,在中心点为极大值或极小值且周边像素值差异较大时,认为该点为噪声点,对于噪声点的处理时将中心点的权重降低。对 x 轴和 y 轴方向分别进行滤波,由于双边滤波法 2 个权重值相乘,计算复杂度高,还存在归一化问题,所以滤波函数修改为

$$h(x) = k^{-1}(x)(c(\xi,x) + s(f(\xi),f(x)))d\xi \quad (5)$$

因为是对雾天图像进行增强处理,原始光照条件受到雾天大气散射影响缓慢变化,在改进的双边滤波之后,对图像进行对比度压缩,以降低去雾操作对图像的退化影响。图像尺寸还原时,采用双三次差值,用到了四邻域的像素信息,能够产生效果更好,更加精确的插补图形,图像的放大效果很好,在整个算法中也有一定的滤波效果,双三次差值公式为

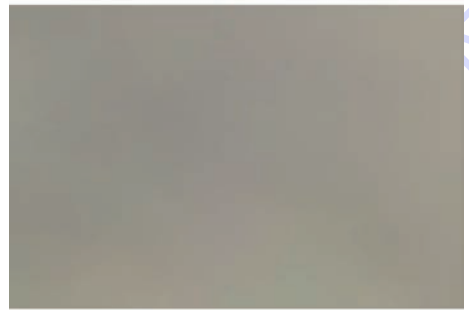
$$B(X,Y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} W(i)W(j) \quad (6)$$

雾天原始图像和通过上述改进型 Retinex 算法得到的入射光图像如图 2 所示。

其后,在对数空间将两图相减,在尺度不变的情况下对于图片的细节增强与颜色恒常性两方面都有好的结果。在计算的时间上,传统多尺度 Retinex 算法使用 3 个不同尺度进行 3 次计算,本



(a) 雾天原始图像



(b) 入射光图像

图 2 雾天原始图像和入射光图像

Fig. 2 Original foggy image and incident light image

文算法计算复杂度低,并且在图像进行高斯金字塔向下采样降低了处理的量级。

3 颜色恢复方法改进

Retinex 算法中传统的颜色恢复方法是利用输入图像中三通道颜色间的比例关系来计算颜色恢复因子矫正图像,该方法会造成像素值的溢出现象。

为了解决还原失真的问题,采用 S 型函数函数来调节计算,根据 S 型函数特点,图像中亮度两端区域范围被压缩,中间亮色区域被拉伸。且人眼对亮度的感知通常是平均亮度越大的地方,感知的亮度差别就越大。将中心亮度区域层次化,就能够提升整幅图像的视觉效果。

将对数域相减结果图像进行S型曲线函数拉伸,因为S型函数的动态压缩能力要高于其他如线性、对数函数^[9]。S型曲线选用sigmoid函数,该函数符合高斯积累特性,并且在低像素值的区段内和对数函数反函数指数函数曲线相似,还可以省去将图像从对数域转换回实数域的过程,sigmoid函数表达式为

$$S(x) = \frac{1}{1 + e^{-x}} \quad (7)$$

在亮色区域和暗色区域进行不同系数的处理,应用于雾天图像颜色还原对比度还原的公式设定为

$$P(x) = \frac{2 \times 255}{1 + e^{-k_1 k_2 x}} \quad (8)$$

式中: $P(x)$ 为颜色还原后的像素值; k_1 为对比度系数,和原雾天图像的标准差成正相关; k_2 为像素值系数,亮度越高系数越小。对图像的颜色还原曲线如图3所示。

在对图像颜色还原后有很好的效果,对原色彩饱和度很高的图像也有着不错的效果,避免了MSRCR算法使得颜色失真的问题。本文方法经过还原后像素值不会小于0,也不会超过255,避免了颜色反转的现象。在处理之后,RGB 3通道的像素值大小关系也不会有变化,在CIE色度值中各个通道的占比变化也相对恒定,减弱了图像的偏色现象,颜色还原效果主要在客观指标中体现。

图4(a)MSRCR算法中颜色还原步骤会使原图颜色饱和和区域出现失真问题,图4(b)本文算法避免了该现象。

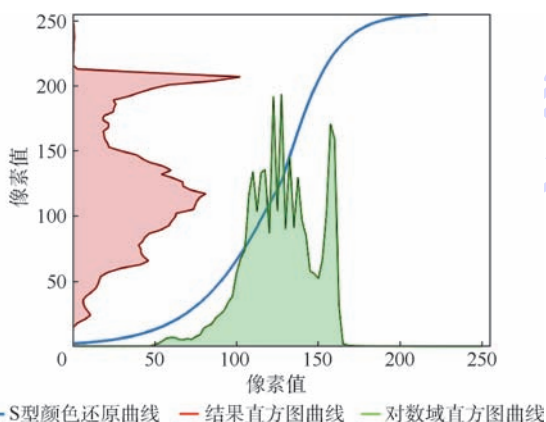
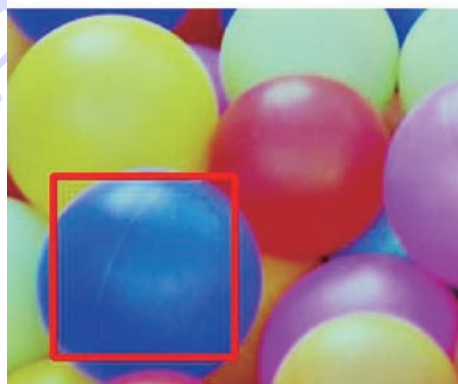


图3 颜色还原直方图曲线示例



(a) MSRCR算法



(b) 本文算法

图4 MSRCR算法和本文算法处理结果

Fig. 4 Processing result of MSRCR algorithm and proposed algorithm

4 实验结果与讨论

为验证本文提出算法的增强效果,使用OpenCV平台进行验证,CPU主频为3.40 GHz,图片尺寸为480像素×360像素。将本文算法的去雾处理结果,与暗通道算法、HE和MSRCR等现有算法进行对比,以验证其可行性和有效性。

由于图像主观评价个体差异大,易受外界环境影响,难以用于科学研究及实际工程之中,针对一幅图像的客观评价至今并没有标准化的图像质量评价标准。为此,本文结合图像物理特性及意义选择以下参数作为评价指标。客观评价指标采用图像的亮度、标准差、信息熵及清晰度。处理目标为雾天图像,雾天的特点为对比度色彩衰减,画面白化,亮度值偏高。亮度值体现彩色图像的整体亮度,去雾效果较好的亮度值相对原图偏低。标准差反应了图像像素值的离散程度,值越高说明对比度越高,算法去雾效果越好,图像标准差std计算公式如下:

$$\text{std} = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (f(i,j) - \text{mean})^2} \quad (9)$$

式中: M 和 N 分别为图像尺寸的宽度和高度; $f(i,j)$ 为图像在坐标 (i,j) 的像素值; mean 为图像的亮度均值。

图像的信息熵代表图像的不确定性,也就是信息量,值越大,含有的信息量越多,信息熵 ent 表达式如下^[5]:

$$\text{ent} = - \sum_{r=1}^{255} \sum_{g=1}^{255} \sum_{b=1}^{255} P_{r,g,b} \text{lb} P_{r,g,b} \quad (10)$$

式中: $P_{r,g,b}$ 为图像中颜色为 (r,g,b) 像素出现的概率。

清晰度选用 Laplace 梯度法,使用 Laplace 算子分别提取水平和垂直方向的梯度,清晰度值越高,说明处理效果越好。图 5 ~ 图 8 是 4 张不同环境下雾天图像的增强结果。

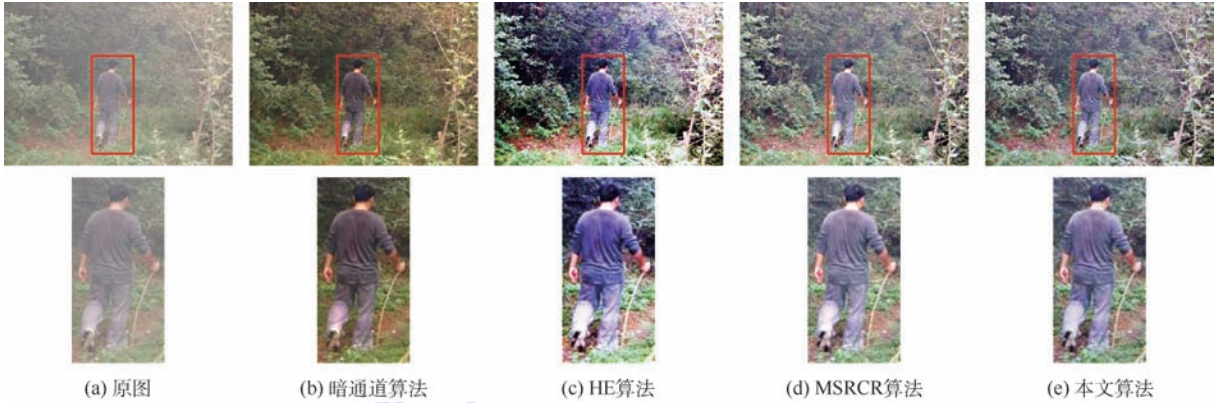


图 5 雾天图像 1 增强效果比较

Fig. 5 Comparison of enhancement effects for foggy image 1

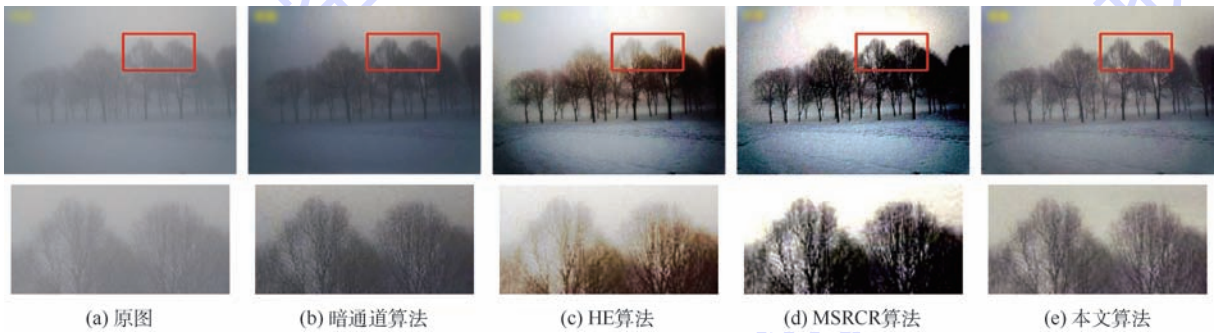


图 6 雾天图像 2 增强效果比较

Fig. 6 Comparison of enhancement effects for foggy image 2



图 7 雾天图像 3 增强效果比较

Fig. 7 Comparison of enhancement effects for foggy image 3



图 8 雾天图像 4 增强效果比较

Fig. 8 Comparison of enhancement effects for foggy image 4

图5~图8从左到右分别是雾天原图,暗通道、HE、MSRCR和本文算法结果。选取不同雾天环境的多幅雾天图像验证本文算法,图8中图像来源于NASA的雾天图片库,本文算法在航空领域中也具有很好的增强效果。从上面多组直观的图像和客观评价结果(见表1~表4)可以看出:

暗通道算法会使得结果图像颜色值偏小,图像中显示过于饱和,该算法最大的问题是计算复杂度过高,无法应用在实时处理环境。

HE算法计算速度快,因为算法就是提升像素之间的差异,所以增强后的图片对比度增加,标准差值大,客观评价效果不错,但是视觉效果差,细节消失,出现图像失真的问题。

MSRCR算法的结果整体比较优异,对不同雾天图像都有较好的增强效果,对比度较低,颜色

表1 雾天图像1客观评价结果

Table 1 Objective evaluation result of foggy image 1

原图及算法	亮度	标准差	信息熵/bit	清晰度	处理时间/ms
原图	151.5	22.9	6.5	6.0	
暗通道算法	98.3	36.7	7.1	10.2	2 185.0
HE算法	129.1	68.9	6.4	10.6	4.1
MSRCR算法	136.4	45.4	7.4	13.9	68.7
本文算法	128.2	52.8	7.5	17.6	39.2

表2 雾天图像2客观评价结果

Table 2 Objective evaluation result of foggy image 2

原图及算法	亮度	标准差	信息熵/bit	清晰度	处理时间/ms
原图	145.5	24.4	6.4	1.4	
暗通道算法	94.8	38.3	6.7	2.1	1 845.5
HE算法	129.4	72.9	6.3	2.6	3.9
MSRCR算法	133.8	46.6	6.5	6.2	56.5
本文算法	129.3	61.0	6.9	6.4	33.4

表3 雾天图像3客观评价结果

Table 3 Objective evaluation result of foggy image 3

原图及算法	亮度	标准差	信息熵/bit	清晰度	处理时间/ms
原图	139.6	57.7	7.2	2.5	
暗通道算法	120.9	69.4	7.2	3.2	2 010.1
HE算法	128.6	74.1	7.1	2.0	3.0
MSRCR算法	90.8	67.4	6.3	4.6	58.1
本文算法	128.0	75.2	7.4	5.1	33.8

表4 雾天图像4客观评价结果

Table 4 Objective evaluation result of foggy image 4

原图及算法	亮度	标准差	信息熵/bit	清晰度	处理时间/ms
原图	116.8	7.4	4.8	1.3	
暗通道算法	99.8	11.6	5.4	2.1	1 564.2
HE	132.8	70.0	4.8	6.3	2.0
MSRCR	118.4	17.7	6.0	4.2	58.2
本文算法	102.7	52.6	5.6	12.8	35.3

较为集中,对颜色饱和区域的还原结果差,运算量较大,且需要配置参数。

本文算法在不同的雾天场景中都有着较好的增强效果,本文算法对雾天增强的结果清晰度较原图提升约200%,标准差提升约110%,信息熵提升约10%。相比于现有算法,除了标准差会有部分图像不如HE算法,其余评价标准都优于现有算法。

现有算法的问题主要为实时性和增强效果无法同时达到最优,部分算法需要调整匹配参数。本文提出的算法在主观视觉和客观评价上都有着较好的处理效果,通过客观图像评价标准可看出该方法对雾天图像增强处理后能够含有更多的信息,颜色更加真实,细节更加丰富,改善了光晕现象。本文提出的算法对不同图像不需修改参数,对不同的雾天条件都有较好的去雾效果,并在实时性上也能达到20帧/s的要求。

5 结论

本文在Retinex算法基础上进行改进,提出了新的图像去雾算法,经实验验证,得出如下结论:

1) 用改进的双边滤波算法作为Retinex算法的滤波函数,在高空金字塔图像进行处理,减小计算量,算法适用性强。

2) 使用改进的S型sigmoid函数颜色还原方法,省去了从对数域转换回实数域的计算过程。通过上述方法的得到得结果在图像亮度、标准差、熵、清晰度等客观图像标准下有更好得结果,计算耗时短,与参数无关,细节突出,颜色还原度高。

未来的研究可以针对雾天和非雾天气进行区分,减小增强的计算复杂度,应用到户外的视频监控场景中。

参考文献 (References)

- [1] 李海波,曹云峰,丁萌,等. 火星沙尘环境光学图像增强方法[J]. 北京航空航天大学学报,2018,44(3):444-453.
LI H B, CAO Y F, DING M, et al. Optical image enhancement method in dust environment on Mars[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(3):444-453 (in Chinese).
- [2] 刘春辉,齐越,丁文锐. 基于大气光鲁棒估计的无人机图像去雾方法[J]. 北京航空航天大学学报,2017,43(6):1105-1111.
LIU C H, QI Y, DING W R. A haze removal method for unmanned aerial vehicle images based on robust estimation of atmospheric light[J]. Journal of Beijing University of Aeronautics and Astronautics, 2017, 43(6):1105-1111 (in Chinese).
- [3] 王浩,张叶,沈宏海,等. 图像增强算法综述[J]. 中国光学,

- 2017,10(4):438-448.
- WANG H,ZHANG Y,SHEN H H, et al. Review of image enhancement algorithms[J]. Chinese Optics,2017,10(4):438-448(in Chinese).
- [4] HE K,SUN J,TANG X. Single image haze removal using dark channel prior[C]//IEEE Conference on Computer Vision and Pattern Recognition,2009. Piscataway, NJ: IEEE Press,2009:1956-1963.
- [5] 赵春丽,董静薇. 基于暗通道及多尺度 Retinex 的雾霾天气图像增强算法[J]. 激光杂志,2018,39(1):104-109.
- ZHAO C L,DONG J W. Image enhancement algorithm of haze weather based on dark channel and multi-scale Retinex[J]. Laser Journal,2018,39(1):104-109(in Chinese).
- [6] 刘海波,杨杰,吴正平,等. 基于暗通道先验和 Retinex 理论的快速单幅图像去雾方法[J]. 自动化学报,2015,41(7):1264-1273.
- LIU H B,YANG J,WU Z P, et al. A fast single image dehazing method based on dark channel prior and Retinex theory[J]. Acta Automatica Sinica,2015,41(7):1264-1273(in Chinese).
- [7] ZHANG W,LIANG J,JU H, et al. Study of visibility enhancement of hazy images based on dark channel prior in polarimetric imaging[J]. Optik-International Journal for Light and Electron Optics,2017,130:123-130.
- [8] JI W,QIAN Z,XU B, et al. A nighttime image enhancement method based on Retinex and guided filter for object recognition of apple harvesting robot[J]. International Journal of Advanced Robotic Systems,2018,15(1):1-12.
- [9] WAN M,GU G,QIAN W, et al. Infrared small target enhancement: Grey level mapping based on improved sigmoid transformation and saliency histogram[J]. Journal of Modern Optics,2018,65(10):1161-1179.
- [10] 黄宇晴,丁文锐,李红光. 基于图像增强的无人机侦察图像去雾方法[J]. 北京航空航天大学学报,2017,43(3):592-601.
- HUANG Y Q,DING W R,LI H G. Haze removal method for UAV reconnaissance images based on image enhancement[J]. Journal of Beijing University of Aeronautics and Astronautics,2017,43(3):592-601(in Chinese).
- [11] 高明,秦世引. 基于非线性直方图变换的对比度畸变图像校正[J]. 北京航空航天大学学报,2016,42(3):514-521.
- GAO M,QIN S Y. Correction of contrast distortion image based on nonlinear transform of histogram[J]. Journal of Beijing University of Aeronautics and Astronautics,2016,42(3):514-521(in Chinese).
- [12] 谢娜. 基于图像增强的图像去雾算法研究[J]. 机械设计与制造工程,2017,46(12):31-33.
- XIE N. Research on the image defogging algorithm based on image enhancement[J]. Machine Design and Manufacturing Engineering,2017,46(12):31-33(in Chinese).
- [13] PARK S,YU S,MOON B, et al. Low-light image enhancement using variational optimization-based Retinex model[J]. IEEE Transactions on Consumer Electronics,2017,63(2):178-184.
- [14] XU K,JUNG C. Retinex-based perceptual contrast enhancement in images using luminance adaptation[C]//IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway, NJ: IEEE Press,2017:1363-1367.
- [15] 刘晓阳,乔通,乔智. 基于双边滤波和 Retinex 算法的矿井图像增强方法[J]. 工矿自动化,2017,43(2):49-54.
- LIU X Y,QIAO T,QIAO Z. Image enhancement method of mine based on bilateral filtering and Retinex algorithm[J]. Industry and Mine Automation,2017,43(2):49-54(in Chinese).
- [16] LIU C,CHENG I,ZHANG Y, et al. Enhancement of low visibility aerial images using histogram truncation and an explicit Retinex representation for balancing contrast and color consistency[J]. ISPRS Journal of Photogrammetry & Remote Sensing,2017,128:16-26.
- [17] MA J,FAN X,NI J, et al. Multi-scale Retinex with color restoration image enhancement based on Gaussian filtering and guided filtering[J]. International Journal of Modern Physics B,2017,31(16-19):1744077.
- [18] YU T H,MENG X,ZHU M, et al. An improved multi-scale Retinex FOG and haze image enhancement method[C]//International Conference on Information System and Artificial Intelligence. Piscataway, NJ: IEEE Press,2017:557-560.

作者简介:

张驰 男,博士研究生。主要研究方向:图像处理、列车控制。

谭南林 男,博士,教授,博士生导师。主要研究方向:控制理论、模式识别。

李响 男,博士,讲师。主要研究方向:信号处理、在线检测。

李国正 男,博士,讲师。主要研究方向:传感器设计、模式识别。

苏树强 男,博士,讲师。主要研究方向:磁性液体、传感器设计。

Foggy image enhancement technology based on improved Retinex algorithm

ZHANG Chi¹, TAN Nanlin¹, LI Xiang², LI Guozheng^{1,*}, SU Shuqiang¹

(1. School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China;

2. School of Transportation and Logistics, East China Jiaotong University, Nanchang 330013, China)

Abstract: In order to enhance the contrast, color and brightness constancy of foggy images, an improved Retinex algorithm of foggy image enhancement is proposed. The algorithm uses an improved bilateral filter as a filter function to remove the noise interference while maintaining the edge information. It uses the S-shape function curve to restore image color by subtracting the incident light component from the logarithmic domain in Retinex algorithm, enhances the contrast and the perceptual characteristics of the whole image, and restores the color information of the image. The experimental results show that the improved algorithm proposed in this paper can effectively improve the clarity and contrast of the foggy image. Compared with the original foggy image, the image clarity is raised by about 200%, the standard deviation is raised by about 110%, and the information entropy is raised by about 10%. At the same time, it can maintain more realistic color information, and the computational complexity is low, which meets the real-time requirements.

Keywords: image enhancement; Retinex; bilateral filtering; edge information; color restore

Received: 2018-06-01; **Accepted:** 2018-08-24; **Published online:** 2018-09-04 09:09

URL: kns.cnki.net/kcms/detail/11.2625.V.20180830.1613.006.html

Foundation item: National Natural Science Foundation of China (61527812)

* **Corresponding author.** E-mail: liguozheng@bjtu.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2018.0318

触地关机模式下的着陆器软着陆稳定性研究

董洋¹, 王春洁^{1,2,*}, 吴宏宇¹, 丁宗茂¹, 满剑锋¹

(1. 北京航空航天大学 机械工程及自动化学院, 北京 100083;

2. 北京航空航天大学 虚拟现实技术与系统国家重点实验室, 北京 100083)



摘 要: 以触地关机软着陆模式下的某型着陆器为研究对象, 建立其软着陆过程的动力学仿真模型。基于仿真模型, 结合优化方法与多岛遗传算法(MIGA)确定了着陆器的极恶劣地形工况参数, 并利用径向基函数(RBF)神经网络建立了反映极恶劣工况下着陆器速度参数与稳定性指标值之间映射关系的代理模型。将着陆器速度参数做离散化处理得到样本点, 利用神经网络模型计算了各样本点对应的软着陆稳定性指标值, 基于计算结果给出了各项软着陆稳定性指标的云图和三维速度稳定性边界, 并得到了综合各项稳定性指标的着陆器速度稳定性边界。分析结果可直观地确定保证着陆器安全着陆的速度取值范围, 为着陆器速度的合理控制提供参考。

关键词: 腿式着陆器; 触地关机; 软着陆; 径向基函数(RBF)神经网络; 稳定性边界

中图分类号: V423.6

文献标识码: A **文章编号:** 1001-5965(2019)02-0317-08

腿式着陆器因其结构简单、可靠性较高等特点, 在各国地外星球探测任务中普遍被选择。目前, 腿式着陆器性能的研究和验证方法主要有物理样机实验和虚拟样机仿真。前者由于成本较高, 且其他星球的重力环境在实验过程中很难模拟, 故难以研究着陆器在多种工况下的性能。随着计算机运算能力的快速增长, 后者的高效率和低成本等优点不断凸显, 逐渐成为着陆器研制过程中不可或缺的一环。

文献[1-3]提出了地球重力下的着陆冲击实验方法, 研究了腿式着陆器在典型工况下的软着陆性能。文献[4-7]建立了着陆器的动力学仿真模型, 研究了典型工况下的着陆器软着陆性能, 其中文献[6]通过对比仿真与实验结果, 证明了建模的准确性。文献[8-9]基于理论分析给出着陆

参数稳定性边界, 其中文献[9]通过实验对边界进行验证。文献[10-11]基于腿式着陆器非0高度关机软着陆模式下的动力学仿真模型, 结合离散化搜寻的仿真策略, 给出了主要工况参数的二维稳定性边界, 其中文献[11]在给出稳定性边界的过程中引入了响应面方法, 提高了计算效率。文献[12-14]在腿式着陆器的动力学仿真模型中添加了主发动机与姿态控制发动机的推力等效模型, 研究了着陆器的触地关机软着陆稳定性, 但未给出直观的稳定性边界。

有鉴于此, 本文以某型着陆器为研究对象, 利用ADAMS软件建立其触地关机软着陆模式下的动力学仿真模型, 结合仿真模型与优化方法确定了多组极恶劣地形工况, 并结合径向基函数(RBF)神经网络代理模型, 给出了速度的三维稳定性边界。

收稿日期: 2018-06-03; 录用日期: 2018-09-03; 网络出版时间: 2018-09-12 15:56

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180911.1031.001.html

基金项目: 国家自然科学基金(51635002)

*通信作者. E-mail: wangcj@buaa.edu.cn

引用格式: 董洋, 王春洁, 吴宏宇, 等. 触地关机模式下的着陆器软着陆稳定性研究[J]. 北京航空航天大学学报, 2019, 45(2): 317-324. DONG Y, WANG C J, WU H Y, et al. Soft landing stability of lander in mode of shutdown at touchdown [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 317-324 (in Chinese).

1 着陆器动力学建模

1.1 着陆器整机模型

本文研究的某型着陆器装有4套中心对称分布于主体的缓冲机构,如图1所示,着陆器是由多个构件通过运动副连接且受到复杂的外力作用的多体系统,分析着陆器软着陆性能,需基于多体动力学理论,建立着陆器的多体动力学模型。其基本动力学方程如下^[15]:

$$\begin{pmatrix} \mathbf{M} & \Phi_q^T \\ \Phi_q & \mathbf{0} \end{pmatrix} \begin{pmatrix} \ddot{\mathbf{q}} \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} \mathbf{Q} \\ \boldsymbol{\gamma} \end{pmatrix} \quad (1)$$

$$\Phi_q = \frac{\partial \Phi}{\partial \mathbf{q}} \quad (2)$$

$$\boldsymbol{\gamma} = -\frac{\partial(\Phi_q \dot{\mathbf{q}})}{\partial \mathbf{q}} \dot{\mathbf{q}} - 2 \frac{\partial \Phi_q}{\partial t} \dot{\mathbf{q}} - \frac{\partial^2 \Phi}{\partial t^2} \quad (3)$$

式中: \mathbf{M} 为广义质量矩阵,由着陆器各构件质量和转动惯量构成; Φ 为运动副约束方程矢量; \mathbf{q} 为广义坐标矢量; Φ_q 为约束方程的雅克比矩阵; \mathbf{Q} 为广义力,主要考虑发动机推力和接触力; $\boldsymbol{\lambda}$ 为拉格朗日乘子矢量; t 为时间。结合所建立的着陆器的动力学模型,利用ADAMS软件进行仿真分析。

缓冲机构主要由主支柱、缓冲杆、左右辅助支柱和足垫组成^[12-14]。缓冲杆在受到冲击时会发生塑性变形,其建模方法参考文献[12]的刚体等效法。如图2所示, α_e 、 M_e 和 S_e 分别为缓冲杆的等效转角、等效弯矩和等效轴向位移,先在有限元软件中建立缓冲杆的模型,进行仿真分析,可以得到图3所示的 α_e - M_e 和 α_e - S_e 曲线,然后将曲线导入刚体模型之中即可完成缓冲杆的等效建模。

辅助支柱内外筒之间的缓冲元件力学特性通过有限元分析获得,进而确定辅助支柱的压缩位移 D_s 和缓冲力 F_s 之间关系如图4所示。将 D_s - F_s 曲线导入刚体模型之中完成其等效建模。

足垫和着陆面的接触采用库伦摩擦模型模拟

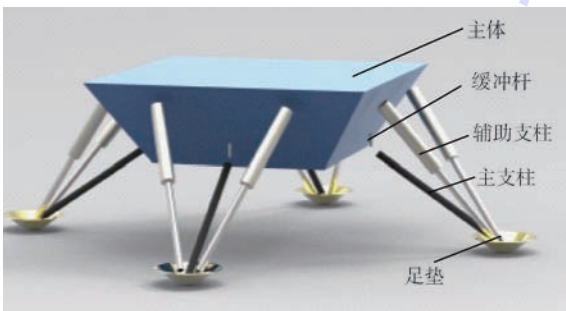


图1 着陆缓冲机构

Fig. 1 Landing buffer mechanism

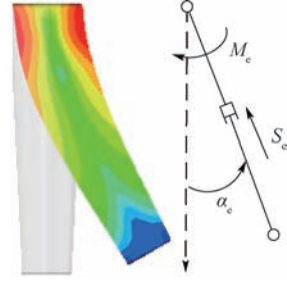


图2 缓冲杆等效模型

Fig. 2 Equivalent model of buffer rod

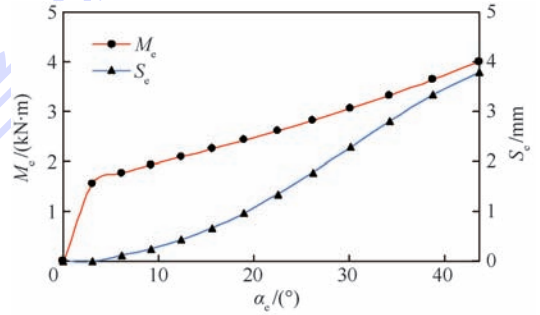


图3 缓冲杆力学特性

Fig. 3 Mechanical behavior of buffer rod

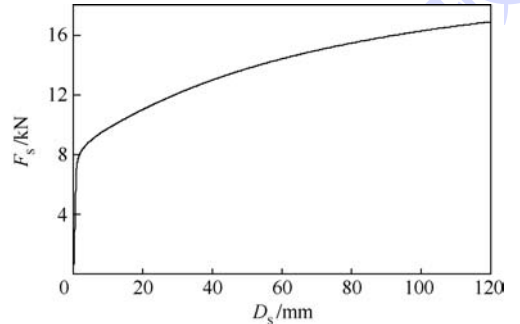


图4 缓冲元件力学特性

Fig. 4 Mechanical behavior of cushion element

切向力,采用非线性弹簧阻尼模型模拟法向碰撞^[12]。

1.2 发动机动力系统等效建模

着陆器主体上装有一个用于制动的主发动机和多台用于姿态调节的姿控发动机。触地关机软着陆模式下,发动机作用对着陆器的软着陆稳定性不能忽视,因此需要建立考虑发动机作用效果的着陆器动力学仿真模型。

1.2.1 姿控发动机控制策略

除主发动机外,触地关机软着陆模式下还有10台姿控发动机处于开启状态,各发动机的位置如图5所示。其中,2台位于主体上方,编号A1、A2,8台位于着陆器主体下方,编号B1~B8。

此外,图5还包括了与姿态控制相关的坐标系。其中, O -XYZ为地面固定坐标系(右手系),

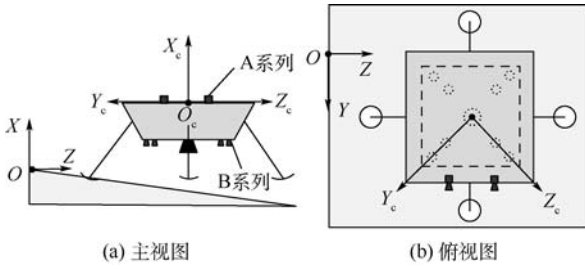


图5 坐标系定义及发动机位置

Fig. 5 Coordinate system definition and engine position

X 轴沿重力方向指向天空, Z 轴水平指向下坡方向。 $O_c-X_c Y_c Z_c$ 为随体控制坐标系(右手系),原点在着陆平台中心, X_c 轴垂直着陆平台表面且指向天空, Z_c 轴在着陆平台内且在俯视图中指向相邻两足垫的中点。定义 $O-XYZ$ 按照 $Z-Y-X$ 的顺序旋转到与 $O_c-X_c Y_c Z_c$ 平行位置时绕三轴转动的角度分别为 $\theta_x, \theta_y, \theta_z$;同时定义 $O_c-X_c Y_c Z_c$ 相对 $O-XYZ$ 的角速度矢量在 $O_c-X_c Y_c Z_c$ 三轴上的投影分别为 $\omega_x, \omega_y, \omega_z$ 。将上述6个参数作为控制输入。

为提升着陆器姿态的稳定性,A1、A2发动机的推力如式(4)和式(5)所示;B系列发动机的推力公式形式相同,均参考文献[12]建立,以B1、B2为例,其推力如式(6)所示:

$$F_{A1} = \begin{cases} \min\{|a_1(\theta_x + k_1\omega_x)|, F_{M1}\} & \theta_x + k_1\omega_x < 0 \\ 0 & \theta_x + k_1\omega_x \geq 0 \end{cases} \quad (4)$$

$$F_{A2} = \begin{cases} \min\{|a_2(\theta_x + k_1\omega_x)|, F_{M1}\} & \theta_x + k_1\omega_x > 0 \\ 0 & \theta_x + k_1\omega_x \leq 0 \end{cases} \quad (5)$$

$$F_{Bi} = \begin{cases} \min\{|b_i(\theta_y + k_2\omega_y)|, F_{M2}\} & \theta_y + k_2\omega_y > 0 \\ 0 & \theta_y + k_2\omega_y \leq 0 \end{cases} \quad (6)$$

$i = 1, 2$

式中: a_i, b_i, k_i 为控制策略中的相关参数; F_{M1} 与 F_{M2} 分别为2个系列发动机推力最大值。

1.2.2 主发动机控制策略

主发动机位于主体下表面中心且喷口方向垂直主体下表面。正常工作时,其推力方程为^[12]

$$F = (-0.7v_x + 2.6)m - \sum_{i=1}^8 F_{Bi} \quad i = 1, 2, \dots, 8 \quad (7)$$

式中: m 为着陆器整机质量; v_x 为竖直速度值。

触地关机软着陆模式下,若足垫与星球表面

接触,则其内部的传感器将发出主发动机关机信号,但主发动机由于关机后效其推力不能马上减小到零,该情况会对着陆稳定性产生影响^[16],因此,为了提高着陆稳定性,姿控发动机会在主发动机收到关机指令之后延迟关机。关机信号发出后,由于信号传输的延迟,主发动机推力下降开始时间会滞后。其推力变化过程如图6所示。图中: t_0 为触地时刻, t_1 时刻发出关机信号, t_2 时刻推力开始下降, t_3 时刻姿控发动机关闭。

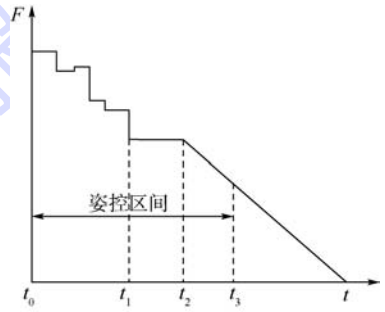


图6 主发动机推力变化过程

Fig. 6 Thrust changing process of main engine

1.2.3 发动机推力等效建模

本文通过在发动机安装位置添加与发动机喷口方向相反的力来模拟发动机的推力作用效果,并利用测量函数实时反馈 $\theta_x, \theta_y, \theta_z, \omega_x, \omega_y, \omega_z$ 与 v_x 的值。在真实情况下,由于传感器按给定周期采集和处理信号,故推力值不随位姿和速度参数连续变化^[17],因此,本文在根据式(7)计算推力值的同时,也按照给定周期给模型中的推力赋值,以模拟发动机的真实工作情况。

2 着陆工况参数与软着陆稳定性判据

2.1 着陆工况参数

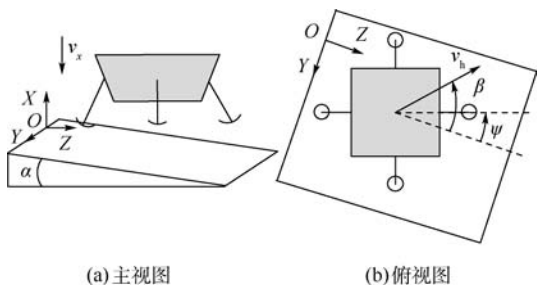
本文考虑的着陆工况参数有2类:

1) 地形参数。包括着陆面坡度(α)、足垫滑移状态(K)和着陆偏航角(ψ)。由于着陆器安装有地貌识别系统,正常情况下会降落在平原地区,坡度一般不超过 12° ,本文考虑着陆面坡度为 $0^\circ \sim 12^\circ$ 。正常情况下足垫可以在着陆面上滑移,但如果足垫遇到石块或者小坑,其将处于卡死状态,本文用 $K=0$ 表示足垫可滑移, $K=1$ 表示足垫卡死。考虑着陆器的几何对称性, ψ 取 $0^\circ \sim 45^\circ$ 即可以表征所有用于评判着陆器软着陆稳定性的姿态。

2) 着陆速度。指着陆器相对地面固定坐标系的速度,为空间矢量。为方便后文的研究,着陆速度采用柱坐标系 (v_h, β, v_x) 描述, v_h ($0 \sim 2.5$ m/s)为速度在 ZOY 平面投影的模长,即水平速度大

小; $\beta(0^\circ \sim 360^\circ)$ 为从 $+X$ 方向看, Z 轴绕到水平速度方向转过的角度; v_x 为竖直速度大小($0 \sim 5 \text{ m/s}$),其方向沿 $-X$ 方向。

部分着陆工况参数如图7所示。



(a)主视图

(b)俯视图

图7 部分着陆工况参数示意图

Fig. 7 Schematic diagram of partial landing conditions parameters

2.2 软着陆稳定性判据

本文重点研究着陆器的软着陆稳定性,根据着陆器构型,确定其稳定性评判参数如下:

1) 缓冲杆最大等效转角 α_{em} (单位为 $^\circ$)。缓冲杆通过塑性变形吸收冲击能量,若等效转角过大其极易发生损坏,故 α_{em} 应小于许用值 α^* 。

2) 辅助支柱最大压缩行程 D_{SM} (单位为mm)。一方面应避免 D_{SM} 超过设计压缩范围,另一方面, D_{SM} 过大易导致着陆平台和坡面夹角过大,不利于探测车的转移,故 D_{SM} 应小于许用值 D^* 。

3) 着陆器质心到4个翻倒平面距离的最小值 L_D (单位为mm)。定义通过相邻两个足垫中心的竖直平面为翻倒平面^[11],为保证着陆器在软着陆过程中不发生翻倒, L_D 应大于许用值 L^* 。

4) 主发动机尾喷管底面中心到着陆面的最小距离 H_M (单位为mm)。触地关机软着陆过程中,主发动机大部分时间处于开启状态,考虑 H_M 过小容易导致粉尘堵塞尾喷管或凸起撞击尾喷管等情况,故 H_M 应大于许用值 H^* 。

3 极恶劣地形的确定

着陆工况参数中, α, K, ψ 依赖于星球的地形,难以主动预测或者控制,而 v_x, v_h, β 作为着陆器自身的参数,具有一定的可控性,为了研究可控参数对着陆器软着陆稳定性的具体影响方式,有必要事先确定 α, K, ψ 的最坏取值。基于动力学模型,本文采用优化方法寻找考虑着陆器水平速度、使稳定性指标 $\alpha_{em}, D_{SM}, L_D, H_M$ 最差的 α, K, ψ 取值。为了降低优化维度、提高计算效率,设定 $v_x = 2.5 \text{ m/s}, v_h = 1.0 \text{ m/s}$,考虑水平速度大小固定的情况下, β 取值的不同可能带来 α, K, ψ 最差取

值的不同,故将 β 作为寻找最差工况的变量之一。综上,以寻找 L_D 最差的工况为例,建立优化模型如下:

$$\min L_D(X) \quad X = [\alpha, K, \psi, \beta]^T \in \Omega \quad (8)$$

式中: Ω 为参数的取值范围。

本文使用多岛遗传算法(MIGA)、结合动力学模型实现优化计算,算法参数配置如表1所示。

经优化计算,得到4组极恶劣工况如表2所示。

表1 优化参数设置

Table 1 Setting of optimal parameters

参数	数值
岛屿种群规模	10
岛数	10
进化代数	25
交叉概率	0.9
迁移概率	0.01
迁移间隔	4

表2 极恶劣工况

Table 2 Worst landing conditions

指标	最优解 $X^* = [\alpha, K, \psi, \beta]^T$
$\max \alpha_{em}$	$[3.713, 1, 0.624, 174.35]^T$
$\max D_{SM}$	$[1.995, 0, 17.32, 138.89]^T$
$\min L_D$	$[12.00, 1, 35.51, 290.63]^T$
$\min H_M$	$[0.405, 0, 13.98, 226.51]^T$

注: α, ψ, β 单位为 $^\circ$;K无量纲。

4 可控工况参数分析

4.1 单稳定性指标的速度稳定性边界

研究 v_x, v_h, β 对着陆器软着陆稳定性的具体影响方式的过程中,需进行大规模仿真计算,为了提高效率,本文引入RBF神经网络代理模型建立极恶劣工况下的 v_x, v_h, β 与4项软着陆稳定性指标值之间的映射关系^[18]。通过随机抽点仿真,将仿真结果导入神经网络模型进行训练,用400个独立的样本测试模型,训练终止准则为均方根相对误差值 R_{MSE} (见式(9))小于0.05。

$$R_{MSE} = \frac{1}{n \bar{y}} \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (9)$$

式中: n 为测试样本点数; y_i 为实际仿真结果; \hat{y}_i 为代理模型计算结果; \bar{y} 为 y_i 的平均值。

得到神经网络模型后,本文采用控制变量法,令 v_x, v_h, β 在取值范围内等间距抽取100个点,共得到 10^6 组样本点,将其代入神经网络模型中,计算得到 10^6 组软着陆稳定性指标值。综合考虑着陆器的软着陆可靠性与模型误差等因素,设定本

文的各项软着陆稳定性指标许用值分别为 $\alpha^* = 30^\circ$ 、 $D^* = 70 \text{ mm}$ 、 $L^* = 50 \text{ mm}$ 、 $H^* = 250 \text{ mm}$, 进而给出了不同 v_x 取值下反映 v_h 、 β 与 α_{em} 、 D_{SM} 、 L_D 、 H_M 关系的云图如图 8 所示, v_h 单位为 m/s 。

此外, 本文还在柱坐标系下给出了 v_x 、 v_h 、 β 三维稳定性边界(许用值等值面), 如图 9 所示。

1) 缓冲杆最大等效转角分析

由图 8(a)可知, 在相同 v_x 与 v_h 下, β 对 α_{em} 影响非常显著, β 越接近 180° 、 α_{em} 越大。因此, 图 8(a)所示的 α_{em} 许用值等值线接近直线, 并且集中在 $120^\circ < \beta < 240^\circ$ 范围内, 即偏坡上方向; 另外随着 v_x 增大, 不稳定区域面积增大。以 α_{em} 为稳

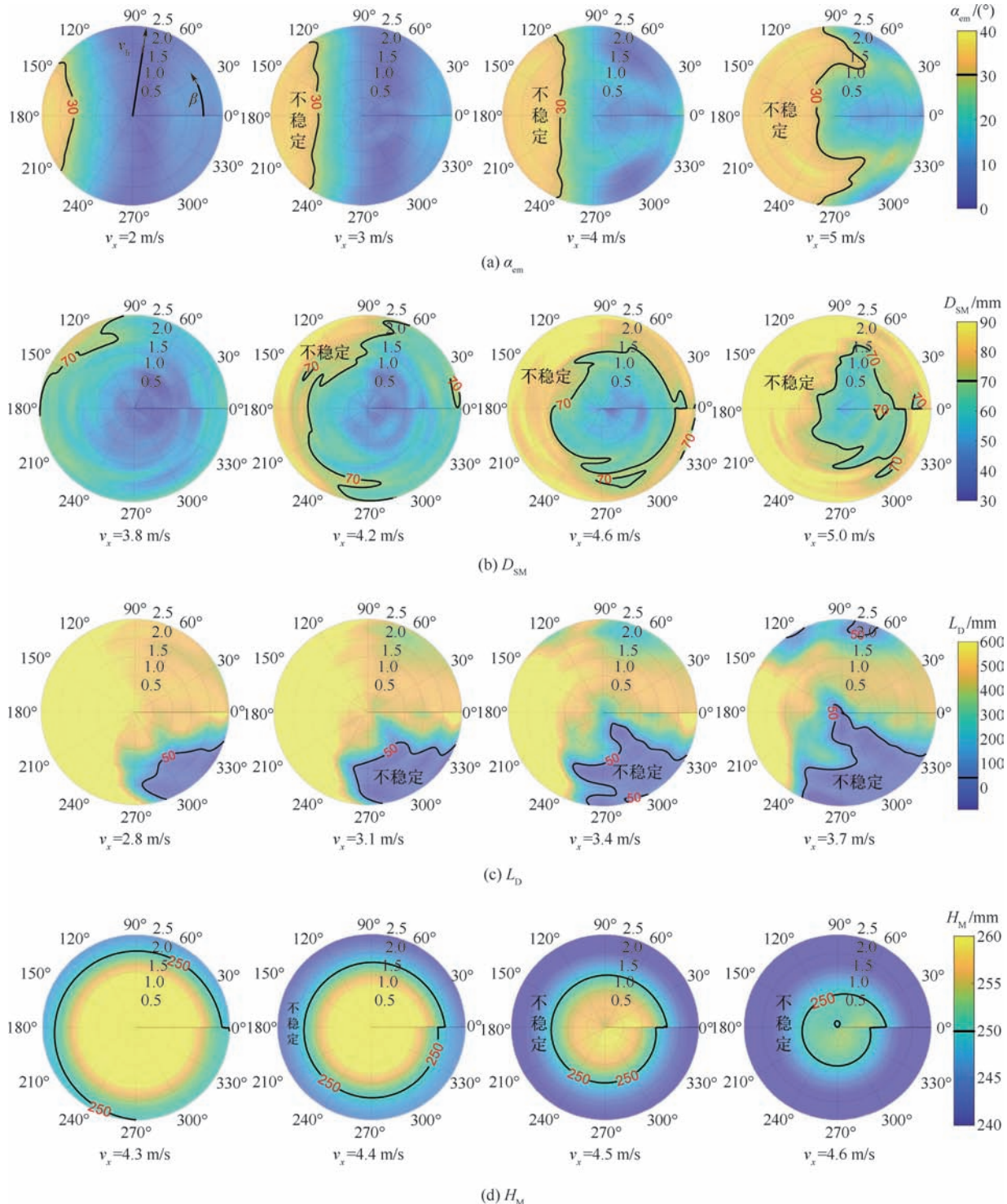


图 8 不同 v_x 下的 α_{em} 、 D_{SM} 、 L_D 及 H_M 云图

Fig. 8 Contour of α_{em} , D_{SM} , L_D and H_M under different v_x

定性指标的速度三维稳定性边界如图9(a)所示,图中曲面为 $\alpha_{em} = 30^\circ$ 的等值面。

2) 辅助支柱最大压缩行程分析

由图8(b)可知,在相同 v_x 下, D_{SM} 的许用值等值线接近环形,可知 β 对 D_{SM} 影响比较小, D_{SM} 的主要影响因素为 v_h 。此外,随着 v_x 增大,不稳定区域面积增大。以 D_{SM} 为稳定性指标的速度三维稳定性边界如图9(b)所示,图中曲面为 $D_{SM} = 70\text{ mm}$ 的等值面。

3) 质心距翻倒平面最小距离分析

由图8(c)可知, L_D 的许用值等值线集中在 $270^\circ < \beta < 360^\circ$ 方向,即偏坡下方向,表明水平速度方向接近坡下方向时着陆器更容易翻倒。以 L_D 为稳定性指标的速度三维稳定性边界如图9(c)所示,图中曲面为 $L_D = 50\text{ mm}$ 的等值面。

4) 尾喷管底面中心到着陆面最小距离分析

由图8(d)可知,在相同 v_x 下, H_M 的许用值等值线非常光滑且接近圆,说明 β 对 H_M 基本没有影响;以 H_M 为稳定性指标的速度三维稳定性边界如图9(d)所示,图中曲面为 $H_M = 250\text{ mm}$ 的等值面,图9(d)直观显示边界集中在 $v_x = 4 \sim 5\text{ m/s}$ 区间,一方面表明该指标主要受 v_x 影响,同时 v_x 小于 4 m/s 时, H_M 不会超过安全阈值。

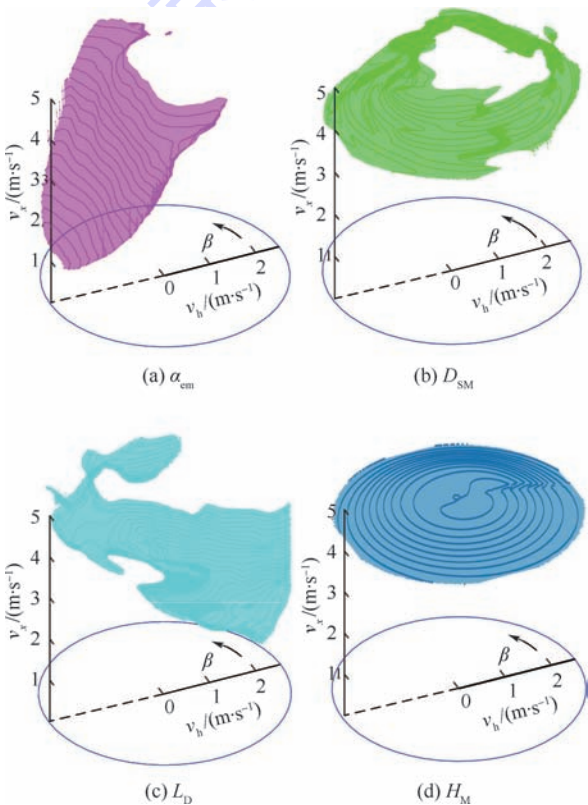


图9 各指标下着陆速度三维稳定性边界

Fig. 9 3D stability boundary determined by landing velocity concerning different indicators

4.2 综合多项稳定性指标的速度稳定性边界

将图9所示的全部三维稳定性边界向同一柱坐标系内整合,得到图10所示的综合三维稳定性边界,同时得到不同竖直速度下的切面图如图11所示,图中稳定区域为各稳定区域的交集, v_h 单位为 m/s 。

将图10所示的三维稳定性边界向同一竖直面内做旋转投影,得到消除 β 影响的4项软着陆稳定性指标的保守速度稳定性边界,综合4条保守边界,可最终确定综合多项稳定性指标的速度稳定性边界,如图12所示。由图12可以直观地确定保证着陆器绝对稳定着陆的速度取值范围。

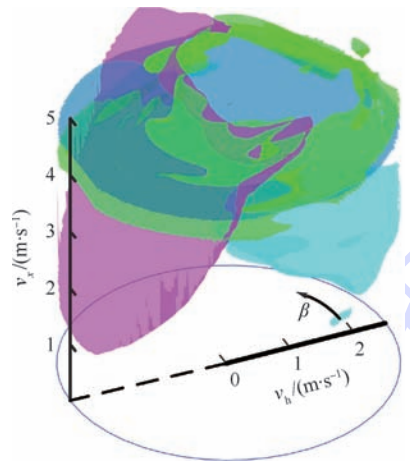


图10 综合指标下着陆速度三维稳定性边界

Fig. 10 3D stability boundary determined by landing velocity considering comprehensive indicators

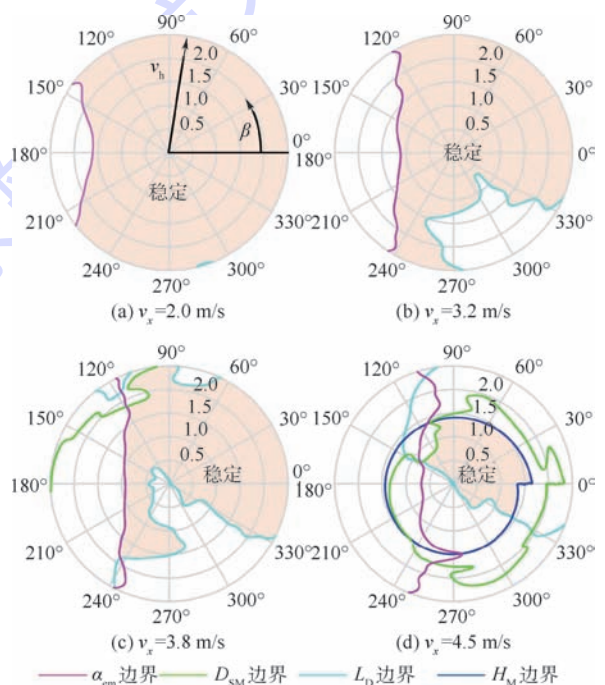
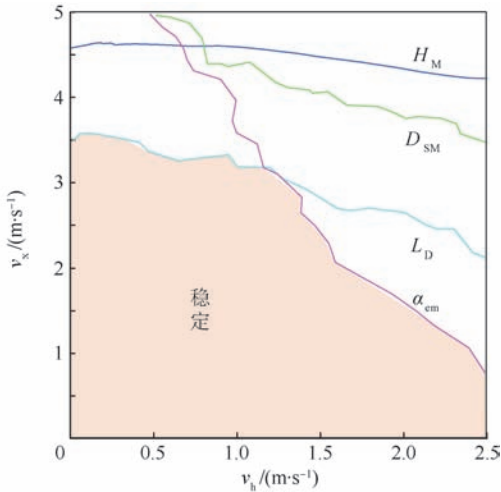


图11 不同 v_x 下综合稳定性边界

Fig. 11 Comprehensive stability boundary under different v_x

图 12 v_x-v_h 综合稳定性边界Fig. 12 v_x-v_h comprehensive stability boundary

5 结 论

本文建立某型着陆器触地关机模式下的动力学仿真模型,着重分析着陆速度对稳定性的影响,得到:

1) 结合动力学仿真模型及多岛遗传优化算法,确定了考虑着陆器水平速度下的极恶劣地形参数。

2) 基于仿真模型计算结果,利用 RBF 神经网络建立了反映极恶劣地形下着陆器速度参数与稳定性指标值之间映射关系的代理模型,进而计算得到了软着陆稳定性指标的速度稳定性边界,确定了保证着陆器安全软着陆的速度取值范围。

3) 本文研究方法具有低成本、高效率等特点,且适用于其他型号着陆器的研制过程。

参考文献 (References)

- [1] 杨建中,曾福明,满剑锋,等.嫦娥三号着陆器着陆缓冲系统设计及验证[J].中国科学:技术科学,2014,44(5):440-449.
YANG J Z,ZENG F M,MAN J F,et al. Design and verification of the landing impact attenuation system for Chang'E-3 lander[J]. Scientia Sinica Technologica,2014,44(5):440-449 (in Chinese).
- [2] CHEN J B,NIE H,WAN J L,et al. Investigation on landing impact dynamic and low-gravity experiments for deep space lander[J]. Science China (Physics, Mechanics & Astronomy),2014,57(10):1987-1997.
- [3] 孙泽洲,张焯,贾阳,等.嫦娥三号探测器地面验证技术[J].中国科学:技术科学,2014,44(4):369-376.
SUN Z Z,ZHANG H,JIA Y,et al. Ground validation technologies for Chang'E-3 lunar spacecraft[J]. Scientia Sinica Technologica,2014,44(4):369-376 (in Chinese).
- [4] 逯运通,宋顺广,王春洁,等.基于刚柔耦合模型的月球着陆器动力学分析[J].北京航空航天大学学报,2010,36(11):1348-1352.
- [5] LU Y T,SONG S G,WANG C J,et al. Dynamic analysis for lunar lander based on rigid-flexible coupled model[J]. Journal of Beijing University of Aeronautics and Astronautics,2010,36(11):1348-1352 (in Chinese).
- [6] MAEDA T,OTSUKI M,HASHIMOYO T,et al. Attitude stabilization for lunar and planetary lander with variable damper[J]. Journal of Guidance, Control, and Dynamics,2016,39(8):1790-1804.
- [7] 曾福明,杨建中,朱汪,等.月球着陆器着陆缓冲性能研究[J].航天器工程,2010,19(5):43-49.
ZENG F M,YANG J Z,ZHU W,et al. Research on landing impact attenuation performance of lunar lander[J]. Spacecraft Engineering,2010,19(5):43-49 (in Chinese).
- [8] 岳帅,聂宏,张明,等.临近空间载人舱着陆动力学及影响因素分析[J].宇航学报,2018,39(3):264-274.
YUE S,NIE H,ZHANG M,et al. Analysis on landing dynamics and influence factors of near space manned capsule[J]. Journal of Astronautics,2018,39(3):264-274 (in Chinese).
- [9] LAVENDER R E. Monte Carlo approach to touchdown dynamics for soft lunar landing:NASA TN D-3117[R]. Washington, D. C.:NASA,1965.
- [10] ROGERS W F. Apollo experience report:Lunar module landing gear subsystem:NASA TN D-6850[R]. Washington, D. C.:NASA,1972.
- [11] 朱汪,杨建中.月球着陆器软着陆机构着陆稳定性仿真分析[J].宇航学报,2009,30(5):1792-1796.
ZHU W,YANG J Z. Touch down stability simulation of landing gear system for lunar lander[J]. Journal of Astronautics,2009,30(5):1792-1796 (in Chinese).
- [12] LIU Y,SONG S,LI M,et al. Landing stability analysis for lunar landers using computer simulation experiments[J]. International Journal of Advanced Robotic Systems,2017,14(6):1-15.
- [13] 吴宏宇,王春洁,丁宗茂,等.两种着陆模式下的着陆器缓冲机构构型优化[J].宇航学报,2017,38(10):1032-1040.
WU H Y,WANG C J,DING Z M,et al. Configuration optimization of landing gear under two kinds of landing modes[J]. Journal of Astronautics,2017,38(10):1032-1040 (in Chinese).
- [14] WU H,WANG C,DING J,et al. Dynamics simulation analysis for novel lander based on two kinds of landing mode[C]// International Conference on Measuring Technology and Mechatronics Automation. Piscataway,NJ:IEEE Press,2017:8-12.
- [15] 丁宗茂,王春洁,吴宏宇,等.探测器触地关机软着陆稳定性分析[J].北京航空航天大学学报,2018,44(3):614-620.
DING Z M,WANG C J,WU H Y,et al. Stability analysis of explorer in soft landing mode of engine shutdown at touchdown[J]. Journal of Beijing University of Aeronautics and Astronautics,2018,44(3):614-620 (in Chinese).
- [16] 陈立平.机械系统动力学分析及 ADAMS 应用教程[M].北京:清华大学出版社,2005:25-30.
CHEN L P. Mechanical system dynamics analysis and ADAMS application tutorial[M]. Beijing:Tsinghua University Press,2005:25-30 (in Chinese).
- [17] 张焯,蔡国飙,许映乔,等.嫦娥三号着陆器软着陆过程中羽

流仿真分析及试验研究[J]. 中国科学:技术科学, 2014, 44(4):344-352.

ZHANG H, CAI G B, XU Y Q, et al. Simulation and experimental study of the plume during the Chang' E-3 lunar landing[J]. Scientia Sinica Technologica, 2014, 44(4): 344-352 (in Chinese).

[17] 张洪华, 关轶峰, 黄翔宇, 等. 嫦娥三号着陆器动力下降的制导导航与控制[J]. 中国科学:技术科学, 2014, 44(4): 377-384.

ZHANG H H, GUAN Y F, HUANG X Y, et al. Guidance navigation and control for Chang' E-3 powered descent[J]. Scientia

Sinica Technologica, 2014, 44(4): 377-384 (in Chinese).

[18] CAROZZA M, RAMPONE S. Function approximation from noisy data by an incremental RBF network[J]. Pattern Recognition, 1999, 32(12): 2081-2083.

作者简介:

董洋 男, 博士研究生。主要研究方向: 航天机构的设计与动力学分析。

王春洁 女, 博士, 教授, 博士生导师。主要研究方向: 数字化设计与仿真分析。

Soft landing stability of lander in mode of shutdown at touchdown

DONG Yang¹, WANG Chunjie^{1,2,*}, WU Hongyu¹, DING Zongmao¹, MAN Jianfeng¹

(1. School of Mechanical Engineering & Automation, Beihang University, Beijing 100083, China;

2. State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100083, China)

Abstract: A lander in the soft landing mode of shutdown at touchdown is taken as the research object, and the dynamic simulation model of its soft landing process is established. Based on the simulation model, the parameters of the worst landing conditions are obtained by combining optimization method and multi-island genetic algorithm (MIGA). The radial basis function (RBF) neural network is used to establish a surrogate model which reflects the mapping relationship between the lander's velocity parameter and the value of stability indicator. The sample points are obtained by discretizing the velocity parameters of the lander, and the soft landing stability performance of each sample point are calculated by using the neural network model. Based on the calculation results, the contour and the three-dimensional velocity stability boundary of each soft landing stability indicator are given, and then a comprehensive stability boundary of lander velocity is obtained. The analysis results can intuitively determine the range of velocity for safe landing, which provides reference for the rational control of lander velocity.

Keywords: leg type lander; shutdown at touchdown; soft landing; radial basis function (RBF) neural network; stability boundary

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

DOI: 10.13700/j.bh.1001-5965.2018.0305

蜂群无人机自组网多优先级自适应退避算法

刘炜伦, 张衡阳*, 郑博, 高维廷

(空军工程大学 信息与导航学院, 西安 710077)

摘 要: 针对现有媒质接入控制 (MAC) 协议退避算法无法为蜂群无人机自组网 (FANETs) 提供区分服务, 且在重负载时性能严重恶化等问题, 提出一种多优先级自适应退避算法。采用忙闲因子自适应机制和最优竞争窗自适应机制, 根据信道忙闲程度和网络状态参数自适应实时更新各优先级竞争窗口 (CW) 长度, 从而使每次退避的竞争窗口可快速收敛到最佳状态, 并实现了多业务区分服务, 得到了最优的系统性能。通过建立不同优先级退避过程的三维 Markov 链模型求解得到了饱和吞吐量下的最优竞争窗自适应因子, 并且理论推导了系统吞吐量和平均 MAC 时延的数学表达式。仿真结果表明, 所提算法在重负载时能够实现多优先级区分服务并有效提高系统的吞吐量性能, 相比区分业务优先级的自适应退避 (PAB) 算法和支持 QoS 的自适应竞争窗口退避算法 (Q-ABACW), 性能均有较大提升。

关键词: 蜂群无人机自组网 (FANETs); 自适应退避; 区分优先级; 忙闲因子; 最优竞争窗口; 忙闲程度

中图分类号: V19; TP393

文献标识码: A

文章编号: 1001-5965(2019)02-0325-08

无人机蜂群作战的概念是美军于 20 世纪 90 年代末率先提出的。无人机蜂群由若干配备多种任务载荷的低成本小型无人机组成, 相比有人战机, 其个体分散性较强, 作战时无人机蜂群可进行专业化分工, 每架无人机的功能不尽相同, 因而可以执行多种任务。由于无人机蜂群作战技术对协同和自主的要求极高, 并且需要建立全新的大规模蜂群指挥控制模式, 因此需要解决协同作战算法、集群个体间通信、远程指挥控制等关键技术^[1-2]。蜂群无人机自组网, 也称飞行自组网 (Flying Ad hoc Networks, FANETs)^[3-5], 是无人机蜂群协同作战的基础和前提, 性能将直接决定协同作战目标能否实现。它的基本思想是: 多无人机间的通信不完全依赖于地面控制站或卫星等基

础通信设施, 而是将无人机作为网络节点, 各节点能够相互转发控制指令, 交换态势感知、情报搜集等数据, 并自动建立一个 Ad hoc 网络。FANETs 采用动态组网、无线中继等技术实现无人机间的互连互通, 具备自组织、自修复能力和高效、快速组网优势, 可使无人机蜂群形成一个整体去执行作战任务。

FANETs 存在大尺度稀疏分布、多业务并存、信道质量不稳定等问题。随着网络负载的增大, 易使信道中产生大量拥塞和冲突, 导致网络性能的下降, 将无法保障控制指令等高优先级业务低时延、高可靠的服务质量 (Quality of Service, QoS) 需求^[6]。针对上述问题, 退避算法作为 FANETs 媒质接入控制 (MAC) 协议的重要组成部分

收稿日期: 2018-05-28; 录用日期: 2018-08-24; 网络出版时间: 2018-09-05 08:45

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180903.1630.007.html

基金项目: 国家自然科学基金 (61701521); 中国博士后科学基金 (2016M603044); 航空科学基金 (20161996010); 陕西省自然科学基金 (2018JQ6074)

* 通信作者. E-mail: hareed@163.com

引用格式: 刘炜伦, 张衡阳, 郑博, 等. 蜂群无人机自组网多优先级自适应退避算法[J]. 北京航空航天大学学报, 2019, 45(2): 325-332. LIU W L, ZHANG H Y, ZHENG B, et al. An adaptive backoff algorithm for FANETs based on multiple priority[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 45(2): 325-332 (in Chinese).

分,其设计的关键在于有效降低信道中的大量冲突,同时兼顾各优先级业务的不同 QoS 需求,保证信息传输的时效性、可靠性。因此,针对 FANETs 设计一种高效、合理的退避算法是有效提升网络性能的关键。

目前移动 Ad hoc 网络广泛采用的是二进制指数退避(Binary Exponential Backoff, BEB)算法^[7-9]。研究表明,该类算法存在无法根据网络负载情况快速选取最佳竞争窗口、在饱和状态时网络性能急剧下降且不能区分服务类别等不足。近年来,许多研究者针对 BEB 算法的不足,根据网络实际需求提出并设计了多种退避算法。例如,文献[10]提出一种增强型退避(Enhanced Backoff, EBO)算法,能够有效降低重负载下网络中的冲突,并提高短期公平性,但无法区分优先级,且不能根据负载情况将竞争窗口收敛到最佳状态;文献[11]提出一种支持 QoS 的自适应竞争窗口退避算法(Adaptive Contention Window Backoff Algorithm for QoS, Q-ABACW),根据分组碰撞概率估计网络中的不同业务的活跃节点数量并动态调整各优先级竞争窗口,实现了多业务区分,但对于大尺度稀疏分布的 FANETs,活跃节点数量难以准确估计;文献[12]提出一种区分业务优先级的自适应退避(Priority Adaptive Backoff, PAB)算法,通过实时调整节点相邻退避阶段的前后转移概率,为多种优先级业务提供区分服务,但其通过载波侦听判定信道忙闲从而决定节点下一阶段退避状态的方法并不准确,无法实际应用于 FANETs。

针对文献[10-12]存在的不足并结合 FANETs 特点,本文提出并设计了一种基于负载反馈和竞争窗口实时更新的多优先级自适应退避算法(Multiple Priority Adaptive Backoff Algorithm,

MPABA)。本文算法采用忙闲因子自适应机制和最优竞争窗自适应机制,通过信道忙闲程度自适应调整各优先级的忙闲自适应因子,并依据网络状态信息自适应调整最优竞争窗自适应因子,从而实时动态更新各优先级竞争窗口长度并使其快速收敛到最佳状态,实现了多优先级业务区分服务、改善了网络在重负载下的性能且能有效保障高优先级业务低时延、高可靠的 QoS 需求。

1 算法描述

由于 FANETs 中各优先级业务具有不同的 QoS 需求,MPABA 可通过忙闲因子自适应机制和最优竞争窗自适应机制,实现对不同优先级业务的区分服务和最优的系统吞吐量性能。该算法的基本原理如图 1 所示,首先对上层产生的业务分组进行纠错编码^[13],然后分别插入各自的优先级队列等待发送。到达队首的分组接入信道前需要进行信道忙闲判定,若信道忙闲程度小于其对应的接入门限,分组立即接入信道,反之执行退避机制,并根据忙闲因子自适应机制和最优竞争窗自适应机制实时更新竞争窗口长度并使其快速收敛到最佳状态。其中,最高优先级(优先级 1)业务具有严格的时效性与可靠性需求,不对其进行接入控制。

该算法主要包含以下 2 大核心机制:

1) 忙闲因子自适应机制。用节点接收机在一段时间内收集到的负载数目对下一时刻的负载数目接收值实时预测,并根据预测值量化信道忙闲程度^[14],然后将结果反馈给接入判定模块,根据信道忙闲程度确定各优先级业务不同的忙闲自适应因子从而实现多业务区分服务。

2) 最优竞争窗自适应机制。为了获得饱和状态下的最优接入参数,假设各优先级业务存在

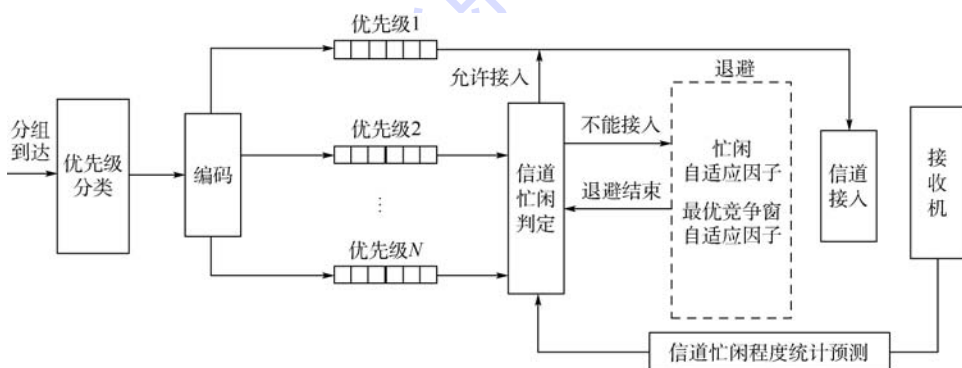


图 1 MPABA 原理

Fig. 1 Principle of MPABA

同一最优竞争窗自适应因子 β_{CWI} , 算法能根据网络状态信息(信道负载、节点数量及信道数目等)自适应调整 β_{CWI} , 以适应动态网络变化, 从而获得更好的网络性能(系统吞吐量最优)。

定义 α_r 表示优先级 $r \in \{2, 3, \dots, R\}$ 业务的忙闲自适应因子, 其大小受负载数目的预测值 G 与各优先级对应的接入门限 A_{th}^r 实时控制, 可将 α_r 量化表示为

$$\alpha_r = \frac{\ln G}{\ln A_{th}^r} \quad (1)$$

且设定 α_r 的最大值为 $\alpha_r^{\max} = l_r$ 。

令 i 表示优先级 r 分组在第 j 次退避阶段时根据信道忙闲程度所确定的忙闲自适应因子 $i \in [1, l_r]$, 其值由式(2)确定, 即

$$i = \min[\lfloor \alpha_r \rfloor, \alpha_r^{\max}] \quad (2)$$

式中: $\lfloor \cdot \rfloor$ 表示向下取整。

将 i 作为竞争窗口调整参数, 构造优先级 r 分组在第 j 个退避阶段的竞争窗口表达式为

$$W_{i,j}^r = \text{Rand int}[\lfloor i\beta_{CWI} \rfloor] \quad (3)$$

2 MPABA 性能分析

2.1 退避过程三维 Markov 链模型

令 p_r 表示优先级 r 的分组到达缓冲区队首时经信道忙闲判定后不能接入的概率, m 表示最大退避次数。令 $g_r(t)$ 表示优先级 r 的分组在 t 时刻的忙闲自适应因子, $s_r(t)$ 表示优先级 r 的分组在 t 时刻所处的退避阶段, $b_r(t)$ 表示优先级 r 的分组在 t 时刻退避计数器的值, 则三维随机过程 $(g_r(t), s_r(t), b_r(t))$ 构成如图 2 所示的离散三维 Markov 链, 其各状态稳态概率为

$$b_{i,j,k}^r = \lim_{t \rightarrow \infty} P^r \{g_r(t) = i, s_r(t) = j, b_r(t) = k\} \quad (4)$$

引入状态 $(0, -1, 0)$ 表示发送缓冲区队首恰好为空的概率(即前一分组服务完成时, 后一分组还未进入队首的瞬时状态), 由图 2 可得, 节点退避状态的一步转移概率可表示为

$$P^r \{i, 0, k | 0, -1, 0\} = \frac{1}{l_r} \cdot \frac{q_r p_r}{W_{i,j}^r} \quad (5)$$

$i \in [1, l_r]; k \in [0, W_{i,j}^r - 1]$

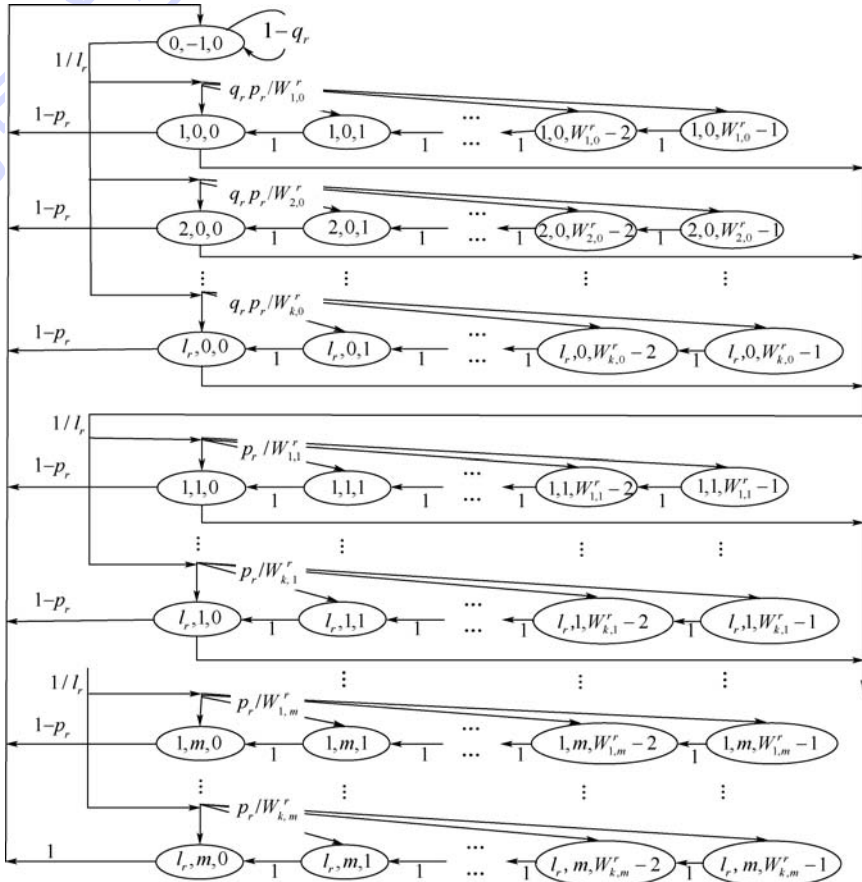


图 2 优先级 r 分组退避状态的三维 Markov 链模型

Fig. 2 Three-dimensional Markov chain model of backoff stage for priority r traffic

$$P^r\{i,j,k|i,j,k+1\} = 1$$

$$i \in [1, l_r]; j \in [0, m]; k \in [0, W'_{i,j} - 2] \quad (6)$$

$$P^r\{p,j,k|i,j-1,0\} = \frac{1}{l_r} \cdot \frac{p_r}{W'_{p,j}}$$

$$i, p \in [1, l_r]; j \in [0, m]; k \in [0, W'_{i,j} - 1] \quad (7)$$

$$P^r\{0, -1, 0|i,j,0\} = 1 - p_r$$

$$i \in [1, l_r]; j \in [0, m] \quad (8)$$

$$P^r\{0, -1, 0|i,m,0\} = 1 \quad i \in [1, l_r] \quad (9)$$

式中: q_r 为单个节点有优先级 r 分组进入相应队列的概率。

式(5)表示发送缓冲区队首分组经信道忙闲判定后首次执行退避的概率;式(6)表示状态转移图中同一行向左转移的概率;式(7)表示状态转移图中分组经信道忙闲判定后无法接入信道并根据式(2)重新确定忙闲自适应因子并进入下一退避阶段的概率;式(8)表示分组经信道忙闲判定后允许接入信道并回到初始状态的概率;式(9)表示分组经 m 次退避后仍不能接入信道,节点将分组丢弃并回到初始状态的概率。

又由图2可得

$$\begin{cases} b'_{i,j-1,0} \cdot p_r = b'_{i,j,0} \rightarrow b'_{i,j,0} = p_r^j \cdot b'_{i,0,0} \\ b'_{0,-1,0} \cdot \frac{1}{l_r} = b'_{i,0,0} \quad i \in [1, l_r] \\ b'_{i,j-1,0} \cdot \frac{1}{l_r} = b'_{p,j,0} \quad i, p \in [1, l_r] \end{cases} \quad (10)$$

结合式(10)及图2推导可得

$$\begin{cases} b'_{0,-1,0} = \sum_{i=1}^{l_r} \left[\sum_{j=0}^{m-1} b'_{i,j,0} (1 - p_r) + b'_{i,m,0} \right] \\ b'_{i,0,k} = b'_{0,-1,0} \cdot \frac{1}{l_r} \cdot \frac{W'_{i,0} - k}{W'_{i,0}} q_r \\ i \in [1, l_r]; k \in [1, W'_{i,0} - 1] \\ b'_{p,j,k} = b'_{i,j-1,0} \cdot \frac{1}{l_r} \cdot \frac{W'_{i,j} - k}{W'_{i,j}} p_r \\ p, i \in [1, l_r]; j \in [1, m]; k \in [1, W'_{i,j} - 1] \end{cases} \quad (11)$$

令 λ_r 表示单个节点中优先级 r 的分组到达率, σ 表示单位时隙长度,则在 σ 内,单个节点有优先级 r 分组进入相应队列的概率 q_r 为

$$q_r = 1 - e^{-\lambda_r \sigma} \quad (12)$$

根据三维 Markov 链的归一化条件,一个节点所有状态概率应满足

$$1 = \sum_{i=1}^{l_r} \sum_{j=0}^m \sum_{k=0}^{W'_{i,j}-1} b'_{i,j,k} + b'_{0,-1,0} \quad (13)$$

联立式(5)~式(13)可求解 $b'_{i,0,0}$ 的表达式。根据式(10)~式(13)可求解图2中每个状

态的取值。

定义 τ^r 表示优先级 r 分组经退避和信道忙闲判定后允许接入的概率,即

$$\tau^r = \sum_{i=1}^{l_r} \sum_{j=0}^m b'_{i,j,0} = \sum_{j=0}^m b'_{1,j,0} = b'_{1,0,0} \frac{1 - p_r^{m+1}}{1 - p_r} \quad (14)$$

由于采用优先抢占式的信道接入策略,节点发送端服务器每次仅能服务一个分组,因此可得分组在任一时隙 σ 能接入信道的概率为

$$P_{\text{access}} = 1 - \prod_{r=2}^R (1 - \tau^r) e^{-\lambda_1 \sigma} \quad (15)$$

则单个节点的分组接入速率 λ_{in} 为

$$\lambda_{\text{in}} = \sum_{r=1}^R \lambda_r P_{\text{access}} \quad (16)$$

2.2 各优先级接入门限求解

由于接入网络的分组会在时域、频域发生碰撞。因此,分组成功接收需保证在同一条信道上,当前分组与其前后一个分组的发送间隔时间同时大于其信道传输时延 T_{send} 。对于单个信道,假设接入网络中的分组在间隔时间上服从参数为 $\lambda_{\text{per}} = N\lambda_{\text{in}}/M$ (N 为节点数量, M 为信道数量)的负指数分布^[15]。定义 P_b 表示分组成功接收概率,则

$$P_b = \frac{e^{-2\lambda_{\text{per}} T_{\text{send}}} (2\lambda_{\text{per}})^0}{0!} = e^{-2\lambda_{\text{per}} T_{\text{send}}} \quad (17)$$

令 N_b 为分组拆分后的突发包个数, M_b 为恢复原分组所需最少突发个数。根据纠错编码机制原理,可得分组成功传输概率 p_{pac} 为

$$p_{\text{pac}} = \sum_{k=M_b}^{N_b} \binom{N_b}{k} P_b^k (1 - P_b)^{N_b - k} \quad (18)$$

假设最高优先级业务所要求的最低传输成功概率为 99%,则令 $p_{\text{pac}} = 99\%$,联立式(17)~式(18)可得当前网络所对应的信道负载为 G_{max} 。

为了保障最高优先级业务的时效性与可靠性需求,需对优先级 $r \in \{2, 3, \dots, R\}$ 业务设置接入门限。设接入网络的优先级 $1, 2, \dots, R$ 业务比例为 $k_1 : k_2 : \dots : k_R$ 。设优先级 r 业务的接入门限为 A_{th}^r ,其含义为其业务恰好无法接入信道时,网络中所对应的负载。 A_{th}^r 需满足

$$A_{\text{th}}^r \left(\frac{\sum_{i=1}^r k_i}{R} \right) = G_{\text{max}} \quad (19)$$

通过式(19)可以求解所有 A_{th}^r 的值,即各优先级业务的接入门限值。

根据接入门限 A_{th}^r 可求解 p_r , 即

$$p_r = P\{G \leq A'_{th}\} = 1 - \sum_{k=0}^{A'_{th}} \frac{e^{-G} G^k}{k!} \quad (20)$$

2.3 系统吞吐量

定义 S 表示系统吞吐量,表示单位时间内信道实际传输成功的所有优先级分组比特数之和,即

$$S = \lambda_{in} N L_{pac} P_b \quad (21)$$

式中: L_{pac} 为数据分组的比特长度。

2.4 平均 MAC 时延

令 $E[D_r]$ 表示优先级 r 分组的平均 MAC 时延,即分组到其相应优先级队首至该分组接入信道前所需时间的平均值,表示为

$$E[D_r] = E[T_r] \sigma \quad (22)$$

式中: $E[T_r]$ 为优先级 r 的分组成功传输所需的平均时隙数,可表示为

$$E[T_r] = \sum_{i=0}^{l_r} \sum_{j=0}^m \left(\frac{W_{i,j} + 1}{2} p_r \right) \quad (23)$$

2.5 最优竞争窗自适应因子求解

为提高接入方案的吞吐量性能,通过式(21)求解可使 S 达到最优的 λ_{in} ,并通过 λ_{in} 求解最优竞争窗自适应因子 β_{cwi} 的值,从而通过自适应调整 β_{cwi} ,使 S 达到最优。

式(21)可表示为

$$S = \lambda_{in} N L_{pac} e^{\frac{-2N\lambda_{in}N_b T_{send}}{M}} \quad (24)$$

通过求吞吐量 S 关于 λ_{in} 的偏导数,令 $\partial S / \partial \lambda_{in} = 0$,易知式(24)存在唯一极大值点可使 S 达到最大。解得该极大值为

$$\lambda_{in} = M \frac{1}{2NN_b T_{send}} \quad (25)$$

联立式(16)和式(25),可得

$$\sum_{r=1}^R \lambda_r p_{access} = M \frac{1}{2NN_b T_{send}} \quad (26)$$

整理式(26)可得

$$A = \prod_{r=2}^R \left(1 - \frac{1}{B_r \beta_{cwi} + C_r} \right) \quad (27)$$

式中:

$$\begin{cases} A = \frac{2NN_b T_{send} - M}{2NN_b T_{send} e^{-\lambda_1 \sigma}} \\ B_r = \left(q_r p_r - \frac{1 - p_r^m}{1 - p_r} \right) \frac{1 - p_r}{1 - p_r^{m+1}} \cdot \frac{l_r - 1}{l_r} \\ C_r = \left[\frac{1 - p_r^m}{1 - p_r} \left(\frac{3}{2} + p_r \right) - \frac{1}{2} q_r p_r + 2 \right] \frac{1 - p_r}{1 - p_r^{m+1}} \end{cases} \quad (28)$$

将式(28)进行整理,并对其取以 e 为底的对

数可得

$$\ln A = \sum_{r=2}^R \ln \left(1 - \frac{1}{B_r \beta_{cwi} + C_r} \right) \quad (29)$$

由于 $q_r, p_r \in (0, 1)$,忽略 p_r^m 和 p_r^{m+1} 的影响,并近似认为 $(l_r - 1)l_r^{-1} \approx 1$,可推导出

$$\frac{1}{B_r \beta_{cwi} + C_r} \approx \frac{1}{\beta_{cwi} - q_r p_r (1 - p_r) \beta_{cwi} + \frac{7}{2} - \frac{1}{2}(q_r + p_r)} \approx \frac{1}{\frac{3}{4} \beta_{cwi} + \frac{11}{4}} \ll 1 \quad (30)$$

因此有 $\ln \left(1 - \frac{1}{B_r \beta_{cwi} + C_r} \right) \sim - \frac{1}{B_r \beta_{cwi} + C_r}$ 成立,并将近似结果代入式(30),最终可得

$$\beta_{cwi} = \frac{4(R - 1)}{-3 \ln A} - \frac{11}{3} \quad (31)$$

由式(27)和式(31)可知,最优竞争窗自适应因子 β_{cwi} 与信道负载(网络中所有节点分组接入速率之和) G_{tra} 、信道数目 M 直接相关。在表1和表2的参数设置下,得到 β_{cwi} 与信道负载、信道数目的关系图3。由图3可知,相同信道数目下的 β_{cwi} 与信道负载呈正比关系;信道负载相同的情况下 β_{cwi} 与信道数目呈反比关系。

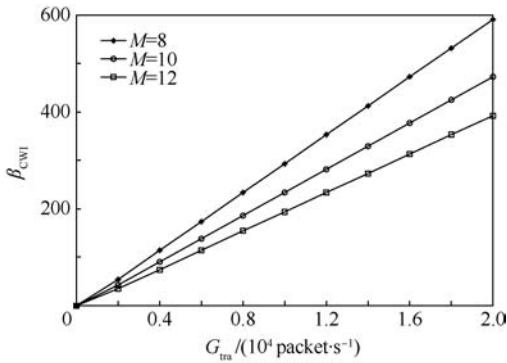
定义信道负载 G_{tra} 表示网络中所有节点的分组接入网络的速率之和。

表 1 仿真参数设置

参数	数值
节点数量	50
信道数量	10
信道传输速率/(Mbit · s ⁻¹)	3
分组长度/bit	1000
编码效率	1/3
单位时隙/μs	100
各优先级最大退避次数	10
分组拆分突发数	28
分组成功接收所需最低突发数	14

表 2 各业务类型相关参数

优先级	信息种类	l_r	$A'_{th}/(\text{packet} \cdot \text{s}^{-1})$
1	武器协同信息		
2	态势感知信息	20	18750
3	网络管理信息	30	5745
4	天气、环境信息	40	1875

图3 β_{CWI} 与 G_{tra} 及 M 的关系Fig.3 Relation of β_{CWI} with G_{tra} and M

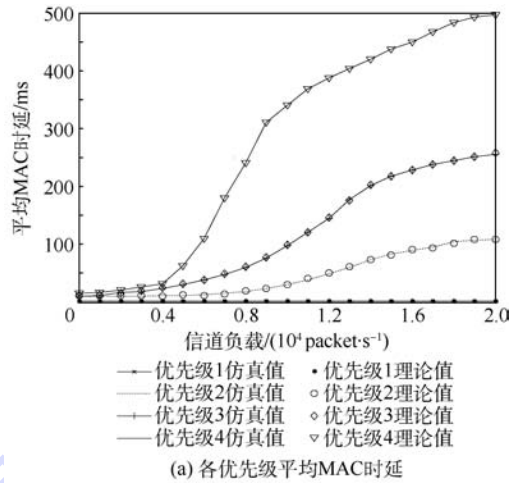
3 仿真分析

下面采用 OMNeT++ 仿真平台对 MPABA 的性能进行仿真分析。仿真场景大小设置为 $200\text{ km} \times 200\text{ km} \times 10\text{ km}$, 所有节点在该场景中的随机分布, 每个节点随机选择目的节点通信, MAC 层采用多信道 ALOHA 协议。仿真结果取 2000 次蒙特卡罗实验的平均值, 具体参数设置见表 1。

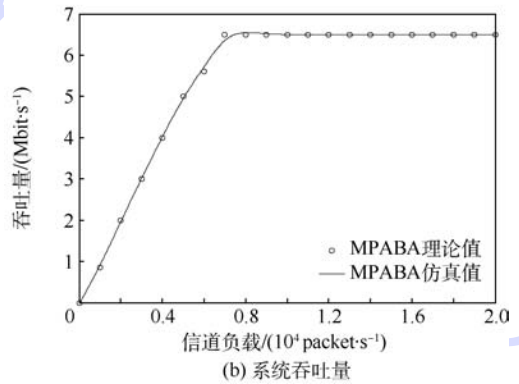
根据 FANETs 的实际应用需求, 对 MPABA 设定 4 种优先级业务, 其中优先级 1 为最高优先级, 其分组到达率固定为 60 packet/s , 优先级 2、3、4 的分组到达率之比为 $1:3:6$ 。在设定优先级 1 分组最低成功传输概率为 99% 的条件下, 在表 1 的参数设置下解得 $G_{max} = 1875\text{ packet/s}$, 其他相关参数如表 2 所示。

首先对 MPABA 性能的数学模型进行仿真验证。不同信道负载下的平均 MAC 时延和系统吞吐量的理论计算和仿真结果如图 4 所示。由图 4(a) 可以看出, 当信道处于轻负载时, 各优先级的平均 MAC 时延均相对较低, 表明此时信道负载小于各优先级接入门限, 各优先级业务均能保证良好的时延性能; 在重负载时, 优先级 1 业务 MAC 时延保持稳定且在 2 ms 以内, 其余各优先级业务的时延均显著增大, 表明此时开始执行退避算法, 由于各优先级业务每次退避时的忙闲自适应因子不同, 导致不同业务每次选取的竞争窗口长度不同, 从而实现了区分服务。由图 4(b) 可以看出, 在重负载时, MPABA 能使吞吐量达到最大且稳定在最大值 6.7 Mbit/s , 表明 MPABA 可自适应调整 β_{CWI} , 使吞吐量达到最优值。由图 4 可知, MPABA 的理论计算结果和仿真实验数据基本吻合, 表明理论计算结果准确有效。

下面将本文提出的 MPABA 与 Q-ABACW^[11]、PAB 算法^[12]的性能进行仿真实验对比, 其中 PAB



(a) 各优先级平均 MAC 时延



(b) 系统吞吐量

图4 信道负载对 MPABA 性能的影响

Fig.4 Influence of channel loads on performance of MPABA 算法、Q-ABACW 均包含 4 种与表 2 相同的优先级业务类别, 在相同仿真参数设置下得到仿真对比结果如图 5 所示。

由图 5(a)、(b) 可知, MPABA 虽然对低优先级业务的时延保障能力较差, 但能为高优先级业务提供严格的时效性保障。由图 5(c) 可知, 相比 MPAB、Q-ABACW, MPABA 在重负载时能够使系统吞吐量保持最优且稳定, 具有明显的性能优势, 可为 FANET 提供较高的系统容量需求。

综合以上仿真结果, 可以得到以下结论:

1) 当信道轻负载时, MPABA 中各优先级业务均可获得较好的 QoS 性能。

2) 当信道重负载时, 低优先级的分组需执行退避算法, 根据信道忙闲程度确定各自的忙闲自适应因子并自适应调整竞争窗口大小, 从而将冲突维持在可控范围内, 为高优先级的分组提供严格的时效性需求。

3) MPABA 可根据网络状况自适应调整 β_{CWI} , 从而可使系统吞吐量达到最优。

4) 对比 PAB 算法、Q-ABACW, MPABA 能为高优先级业务提供更好的服务, 并使得系统吞吐量达到最优。

上具有明显的优势。

参考文献 (References)

[1] 袁政英. 美空军未来 20 年小型无人机发展路线图 [J]. 防务视点, 2016 (10): 58-59.
YUAN Z Y. United States air force in the next 20 years the development of a small UAV roadmap [J]. Defense Point, 2016 (10): 58-59 (in Chinese).

[2] 陈方舟, 黄靖皓, 赵阳辉. 美军无人“蜂群”作战技术发展分析 [J]. 装备学院学报, 2016, 27 (2): 34-37.
CHEN F Z, HUANG J H, ZHAO Y H. Analysis on unmanned swarm fighting system of US armed forces [J]. Journal of Equipment Academy, 2016, 27 (2): 34-37 (in Chinese).

[3] SHARMA V, KUMAR R, KUMAR N. DPTR: Distributed priority tree-based routing protocol for FANETs [J]. Computer Communications, 2018, 122: 129-151.

[4] KHAN M A, SAFI A, QURESHI I M, et al. Flying ad-hoc networks (FANETs): A review of communication architectures, and routing protocols [C] // 2017 First International Conference on Latest trends in Electrical Engineering and Computing Technologies (INTELLECT). Piscataway, NJ: IEEE Press, 2017: 1-9.

[5] BEKMEZCI I, SAHINGOZ O K, TEMEL S. Flying ad hoc networks (FANETs): A survey [J]. Ad hoc Networks, 2013, 11 (3): 1254-1270.

[6] OZGUR K S. Networking models in flying ad-hoc network evaluation: Challenges concepts and challenges [J]. Journal of Intelligent & Robotic Systems, 2014, 74 (1-2): 513-527.

[7] ULLAH A, AHN J S. Performance evaluation of X-MAC/BEB protocol for wireless sensor networks [J]. Journal of Communications and Networks, 2016, 18 (5): 857-869.

[8] DIACONU F. A modified binary exponential backoff algorithm for improving the quality of service in highly populated IEEE 802. 11e networks [C] // International Symposium on Signals, Circuits and Systems ISSCS2013. Piscataway, NJ: IEEE Press, 2013: 1-4.

[9] FIRYAGUNA F, CARVALHO M M. Performance of polling disciplines for the receiver-initiated binary exponential backoff MAC protocol [J]. Ad Hoc Networks, 2015, 59 (1): 1-19.

[10] QI H, HU Z Q, WEN X M. An enhanced MAC backoff algorithm for heavy user loaded WLANs [C] // 2017 IEEE Wireless Communications and Networking Conference (WCNC). Piscataway, NJ: IEEE Press, 2017: 1-6.

[11] IKRAM S, SEUNG-HUN S, BYEONG-HEE R, et al. Performance improvement of QoS-enabled WLANs using adaptive contention window backoff algorithm [J]. IEEE Systems Journal, 2018, 12 (4): 3260-3270.

[12] 卓琨, 张衡阳, 郑博, 等. 一种优先级区分的机载无线网络 MAC 层自适应退避算法 [J]. 航空学报, 2016, 37 (4): 1281-1291.
ZHUO K, ZHANG H Y, ZHENG B, et al. An adaptive backoff algorithm in MAC layer for airborne network based on priority differentiation [J]. Acta Aeronautica et Astronautica Sinica, 2016, 37 (4): 1281-1291 (in Chinese).

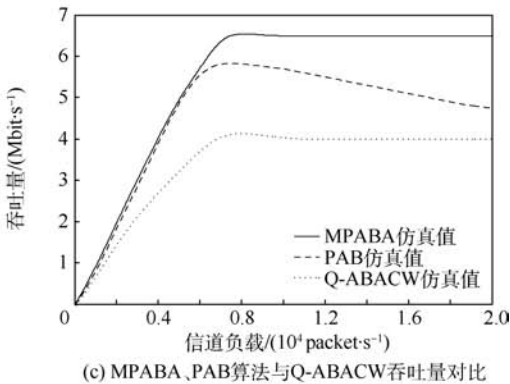
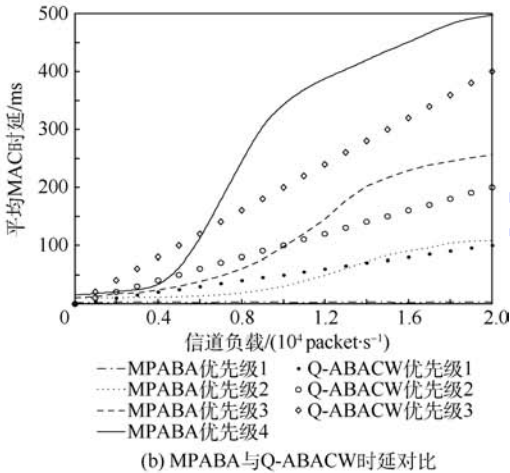
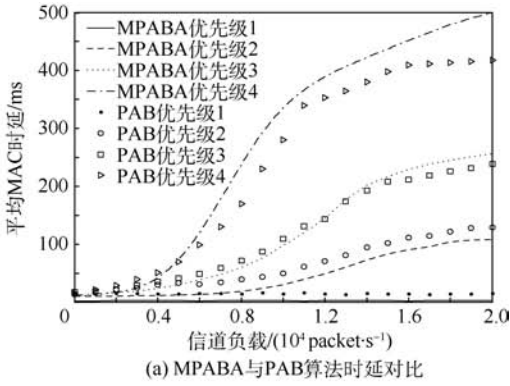


图 5 MPABA、PAB 算法与 Q-ABACW 性能对比
Fig 5 Comparison of performance among MPABA, PAB algorithm and Q-ABACW

4 结 论

本文为 FANETs 提出并设计了一种多优先级自适应退避算法,旨在实现多业务区分服务并有效改善网络在重负载下的性能。

1) 能根据信道忙闲程度和网络状态参数自适应调整各优先级竞争窗口长度,从而可将竞争窗口收敛到最佳状态,为不同优先级业务提供了不同的 QoS 保障能力,得到了最优的系统性能。

2) 仿真实验验证了所建模型准确有效,与 Q-ABACW 和 PAB 算法相比,算法在吞吐量性能

- [13] ZHANG B, HU Z, XING K. Performance of RS-Turbo concatenated code in AOS [C] // 11th International Conference on Electronic Measurement & Instruments. Piscataway, NJ: IEEE Press, 2013: 983-987.
- [14] 刘炜伦, 张衡阳, 郑博. 机载自组网信道占用统计预测机制 [J]. 计算机工程与应用, 2018, 54(15): 78-83.
LIU W L, ZHANG H Y, ZHENG B. Statistical prediction mechanism for channel occupancy in airborne ad hoc network [J]. Computer Engineering and Applications, 2018, 54(15): 78-83 (in Chinese).
- [15] XU D H, ZHANG H Y, ZHENG B, et al. A priority differentiated and multi-channel MAC protocol for airborne networks

[C] // The 8th IEEE International Conference on Communication Software and Networks (ICCSN). Piscataway, NJ: IEEE Press, 2016: 64-70.

作者简介:

刘炜伦 男, 硕士研究生。主要研究方向: 机载自组网。

张衡阳 男, 博士, 副教授, 硕士生导师。主要研究方向: 移动 Ad hoc 网络。

郑博 男, 博士, 讲师。主要研究方向: 移动 Ad hoc 网络。

An adaptive backoff algorithm for FANETs based on multiple priority

LIU Weilun, ZHANG Hengyang*, ZHENG Bo, GAO Weiting

(Information and Navigation College, Air Force Engineering University, Xi'an 710077, China)

Abstract: The existing backoff algorithms of the medium access control (MAC) protocols cannot provide the multiple priority differentiation, and the performance declines sharply under heavy loads in flying Ad hoc networks (FANETs), so a novel adaptive backoff algorithm based on multiple priority differentiation is proposed in this paper. The algorithm adopts a busy/idle factor adaptive mechanism and an optimal contention window (CW) adaptive mechanism, so the length of CW for each priority can be adjusted in real time with the busy degree of channels and network state parameters. Meanwhile, the CW can quickly converge to the best state in every backoff stage, and the multiple priority differentiation can be obtained. Furthermore, the best system throughput performance can be achieved by modeling. The three-dimensional Markov chain model of the backoff process for different priorities is established and the adaptive factor under the saturated throughput is solved by theory. In addition, the mathematical expressions of system throughput and mean MAC delay are also deduced. Simulation results show that the algorithm can achieve the multiple priority differentiation and availably enhance the system throughput, and its performance is superior to the priority adaptive backoff (PAB) algorithm and adaptive CW backoff algorithm for QoS (Q-ABACW).

Keywords: flying Ad hoc networks (FANETs); adaptive backoff; priority differentiation; busy/idle factor; optimal contention window; busy degree

Received: 2018-05-28; **Accepted:** 2018-08-24; **Published online:** 2018-09-05 08:45

URL: kns.cnki.net/kcms/detail/11.2625.V.20180903.1630.007.html

Foundation items: National Natural Science Foundation of China (61701521); China Postdoctoral Science Foundation (2016M603044); Aeronautical Science Foundation of China (20161996010); Natural Science Foundation of Shaanxi Province, China (2018JQ6074)

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

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

DOI: 10.13700/j.bh.1001-5965.2018.0313

一种图像缩放算法的 SoC 协同加速设计方法



王鹏¹, 曹云峰^{1,*}, 许蕾¹, 丁萌², 张洲宇¹, 曲金秋¹

(1. 南京航空航天大学 航天学院, 南京 210016; 2. 南京航空航天大学 民航学院, 南京 211106)

摘 要: 针对无人机自主着陆的跑道检测、识别、跟踪等视觉算法中需要对大量图像进行缩放处理以便后续计算,但又对实时性要求比较高的情况,根据输入输出像素点的映射关系提出了一种适用于硬件加速的图像缩放算法,简化算法结构的同时利用现场可编程门阵列进行模块硬件功能的设计对算法加速,并采用软硬件协同的体系结构搭建实时图像处理系统。实验结果表明,该缩放算法处理精度高、耗时少,且用硬件逻辑实现后,可以进一步提速 171 倍,硬化后的系统可以通过摄像头获取图像数据,实时处理后在显示器中显示,达到 30 帧/s 的处理速度,可以应用于实时性要求较高的图像处理算法中。

关键词: 图像缩放; 实时图像处理; 片上系统 (SoC); 软硬件协同; 无人机视觉

中图分类号: TP391

文献标识码: A

文章编号: 1001-5965(2019)02-0333-07

近年来由于机器视觉的飞速发展,图像处理技术在不同领域下的应用也越来越广泛。通常情况下,工程应用中的图像处理算法需要处理的图像数据很大,处理时间相对较长,因此提高处理速度在实时性要求比较高的应用领域是非常必要的。在算法优化已经无法提高计算速度的前提下,将部分单一、耗时的图像处理操作作用硬件来实现可以有效地提高算法的速度^[1-4]。现场可编程逻辑门阵列 (Field-Programmable Gate Array, FPGA) 具有灵活的开发方式以及很强的并行性和逻辑性,适用于数据量大的简单图像处理,但是其不擅长于比较复杂的算法及逻辑控制,而高端精简指令集机器 (Advanced RISC Machine, ARM) 刚好具有较强的控制能力,能够协调软硬件的资源分配,控制整个系统的运行^[5-8]。

文献[9]中提出了一种基于 ARM + FPGA 的交通标志识别系统,通过设计图像预处理 IP 核,

结合 Zynq-7000 全可编程片上系统 (System on Chip, SoC) 提出了一个新的硬件平台,与现有的基于 FPGA 的解决方案相比,该硬件平台可以实现高达 8 倍的速度。文献[10]中认为方向梯度直方图 (Histogram of Oriented Gradient, HOG) 算法是一个计算密集的任务,并在基于 Zynq SoC 的系统上提出了 3 种不同的实现方法来加速算法,第 1 种采用 OpenCV 在 ARM 上运行 HOG 的方法,第 2 种采用 ARM 和 FPGA 协同实现的方法,第 3 种仅使用 FPGA 实现 HOG 算法,实验结果表明采用 ARM 和 FPGA 的协同方法在处理速度和资源利用综合指标下达到最优。因此笔者采用了 ARM + FPGA 软硬件协同的方式来实现无人机跑道检测算法中图像预处理的硬件加速。

本文针对目前主流计算机处理算法成本高、运行效率低的情况,提出了一种低成本、运行效率高的适用于机载的 SoC 实现方法。针对算法处理

收稿日期: 2018-05-30; 录用日期: 2018-10-15; 网络出版时间: 2018-10-29 09:17

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20181025.1141.002.html

基金项目: 国家自然科学基金 (61673211); 南京航空航天大学博士学位论文创新与创优基金 (BCXJ18-11); 中央高校基本科研业务费专项资金 (kfjj20171502)

* 通信作者. E-mail: cyfac@nuaa.edu.cn

引用格式: 王鹏, 曹云峰, 许蕾, 等. 一种图像缩放算法的 SoC 协同加速设计方法[J]. 北京航空航天大学学报, 2019, 45(2): 333-339. WANG P, CAO Y F, XU L, et al. SoC collaborative acceleration design method for image scaling algorithm [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 333-339 (in Chinese).

中相对耗时的图像缩放处理进行实验,提出了一种适用于硬件模块实现的图像缩放算法,并设计缩放算法知识产权(Intellectual Property, IP)核。采用 Xilinx 公司的 Zynq-7000 系列,该系统集成了 2 个 ARM Cortex A9 核,以及最多可达相当于 500 多万逻辑门可编程逻辑单元,搭建软硬件协同的图像处理操作系统^[11-15],搭载设计的缩放算法 IP 核,完成图像的采集、传输、缩放到显示的过程,达到图像实时处理的要求。

1 系统结构搭建

整个图像处理系统如图 1 所示,系统的核心部分是 Xilinx 公司的 Zynq-7000 芯片,该芯片由处理系统(Processing System, PS)和可编程逻辑(Programmable Logic, PL)两大功能模块组成。PS 和 PL 之间的接口都是基于总线协议(Advanced eXtensible Interface, AXI)设计的,包括高性能总线(High Performance AXI, AXI_HP)和通用总线(General Purpose AXI, AXI_GP)2 种,在图像数据流传输中,部分使用 AXI_HP 设计的接口,提供快速传输的功能,而在控制部分使用 AXI_GP 总线设计的接口。图像采集部分使用的是 OV7670 图像传感器,它具有体积小,工作电压低的特点,可以通过串行摄像机控制总线协议(Serial Camera Control Bus, SCCB)控制输入整帧、取窗口等方式的各种分辨率 8 位影像数据^[9-12]。图像采集模块控制摄像头采集图像数据,并将数据传送到图像处理 IP 核中进行处理,处理后的数据送入数据搬运器(Datamover)中,通过 Datamover 可以实现数据从 PL 到双倍速率同步动态随机存储器(Double Data Rate, DDR)的搬运,也可以实现数据从 DDR

到 PL 的传送,对一些硬件无法实现的复杂图像处理运算,可以将数据传送到 PS 端进行处理,通过 Datamover 传送数据可以减少中央处理器(Central Processing Unit, CPU)的利用率,加速平台处理的速度,通过显示器显示处理后的图像。系统的硬件结构框图如图 1 所示。

2 图像缩放算法设计

传统的图像缩放大多是采用插值算法来实现的,主要包括最邻近插值、双线性插值和双三次插值算法。由于需要用硬件实现图像缩放,考虑到 FPGA 内的逻辑资源有限,而算法越复杂设计时占用的逻辑资源就越多,完成一幅图像缩放所用的时间就越长,影响算法的实时性,因此,笔者在基于无人机自主着陆的跑道检测算法中,考虑在满足缩放精度并且简化算法的情况下,设计了一种适用于 FPGA 硬件实现的图像缩放算法。很多航空电子嵌入式图像处理系统是由固定的图像采集源和显示设备组成,系统中的缩放倍率是固定的,因此根据通用跑道检测算法中显著性区域提取时将输入图像缩小 0.25 倍来进行设计,大大减小了算法的复杂性和硬件开发的难度。

2.1 算法原理

由于考虑将图像以 0.25 进行缩小,该算法的核心思路是选择一个 4 × 4 的像素块求取区域内每个像素点对中心点的影响系数,将该 16 个点的影响集中在一个点上,实现图像的缩小,输入像素和输出像素的映射关系见图 2(a)。

在计算过程中,先选取行像素点求取中心点,这里以图 2(a)中的第 2 行为例,根据 A、B、C、D 距离中心点的距离来确定影响系数并求取中心点 E,如式(1)所示:

$$E = \frac{\frac{1}{4}A + \frac{3}{4}B + \frac{3}{4}C + \frac{1}{4}D}{2} \tag{1}$$

同理可得 E'、e、e'。再分别根据每行获取的中心点求取列中心点 O,即 4 × 4 区域的中心点,见图 2(c)和式(2)。

$$O = \frac{\frac{1}{4}E + \frac{3}{4}E' + \frac{3}{4}e + \frac{1}{4}e'}{2} \tag{2}$$

分别将求取每行中心点获得的等式代入式(2),即可求得每个点对输出点的影响系数,见式(3):

$$\begin{cases} \partial B = \partial C = \partial b' = \partial c' = \partial A' = \partial D' = \partial a = \\ \partial d = 0.046875 \\ \partial A = \partial D = \partial a' = \partial d' = 0.015625 \\ \partial B' = \partial C' = \partial b = \partial c = 0.0140625 \end{cases} \tag{3}$$

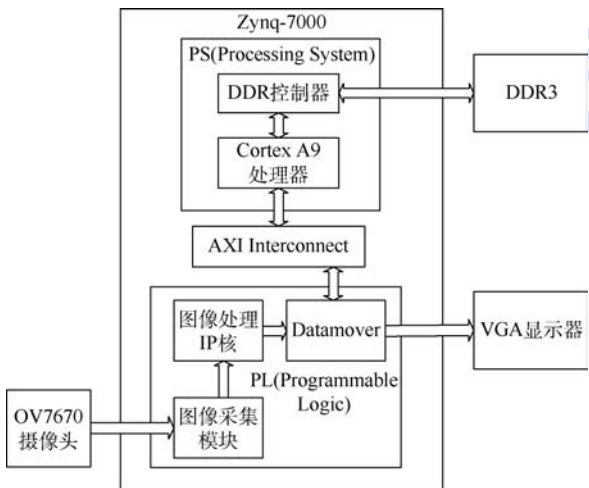


图 1 系统硬件结构示意图

Fig. 1 Schematic diagram of system hardware structure

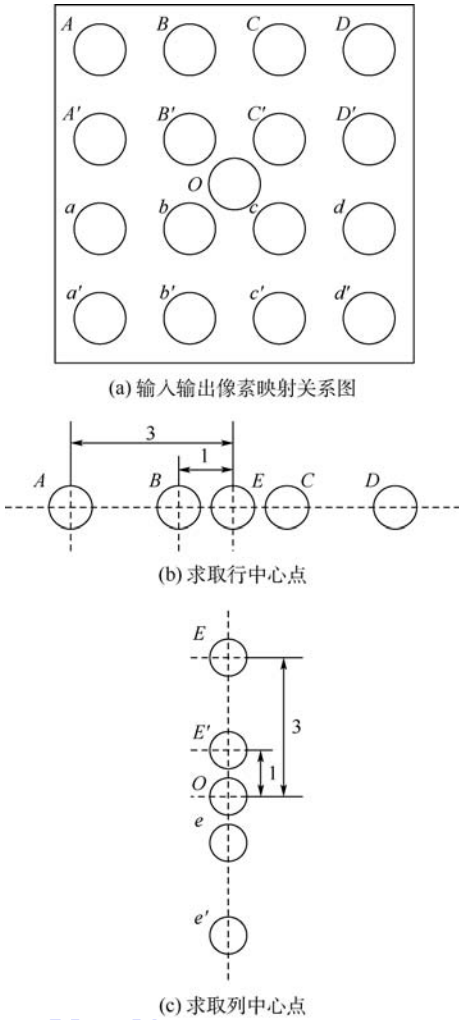


图 2 求取输出点示意图

Fig. 2 Schematic diagram of output point solving

式中: ∂a 为像素点 a 对输出像素点的影响系数, 以此类推。用 4×4 的像素区域遍历输入图像, 通过固定的系数处理该像素区域即可获得缩小后的图像, 因此本文设计的缩放算法在结构上得到了简化, 可以更方便地在 FPGA 中用硬件逻辑语言实现。

2.2 性能测试

算法结构在简化后, 其性能的好坏需要和传统缩放算法作对比, 因此, 笔者根据本文设计的图像缩放算法编写 MATLAB 测试程序, 并且和传

统的最邻近插值、双线性插值和双三次插值算法分别作对比, 从缩放后的效果图、运算时间等 2 个方面作对比。输入图像是用飞行模拟器 (Flight-Gear) 模拟获得的无人机着陆时的跑道图片, 图 3 是 3 种传统算法和本文提出的算法处理效果对比图。

从图 3 中可以看到, 将输入图像按同等倍数缩小时, 4 幅图像都保持了原有图像的基本特征, 但是最邻近插值处理后的图像比较粗糙, 有较明显的锯齿, 而后 3 种图像处理后的效果比较平滑。在图像清晰度上, 本文算法的处理效果相对于双线性插值和双三次插值更清晰。

除此之外, 引入定量分析的指标对实验结果进行评估, 采用文献 [16] 中的将图像缩小后再放大到同一分辨率, 再利用绝对误差均值 (Mean Absolute Error, MAE)、均方误差 (Mean Squared Error, MSE)、峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 以及信噪比 (Signal Noise Ratio, SNR) 对缩放效果进行评估。MAE、MSE、PSNR、SNR 的计算公式分别为

$$MAE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (g(i,j) - f(i,j)) \quad (4)$$

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (g(i,j) - f(i,j))^2 \quad (5)$$

$$PSNR = 10 \lg \frac{255^2}{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (g(i,j) - f(i,j))^2} \quad (6)$$

$$SNR = 10 \lg \frac{\sum_{i=1}^M \sum_{j=1}^N g(i,j)^2}{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (g(i,j) - f(i,j))^2} \quad (7)$$

式中: M 和 N 分别为输入二维图像的行数和列数; $g(i,j)$ 和 $f(i,j)$ 分别为横坐标为 i 、纵坐标为 j 下的输入图像和输出图像像素值。当 MAE、MSE 的值越小, PSNR、SNR 的值越大, 图像的处理效果越好, 质量越高。采用这 4 个指标的客观比较结果如表 1。

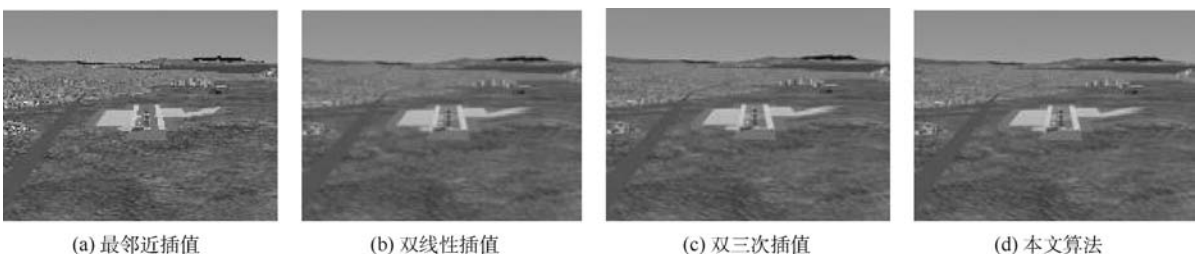


图 3 不同算法处理效果图

Fig. 3 Processing effect diagram of different algorithms

从表 1 中可以看到,评价指标中 MAE 是最小的,MSE 和双三次插值法同时最小,PSNR 和 SNR 都是最大的,所以定量分析的结果是本文提出的算法相较于其他 3 种算法处理效果更好。4 种算法处理的时间见表 2。

从表 2 中可以看到,因为最邻近插值法的算法结构最简单,所以它的耗时最少,而双三次插值的耗时最多。在实际应用中,当基于无人机视觉算法的跑道跟踪过程中需要处理大量图片时,缩放环节节约的总时间是非常可观的。

因此,综合考虑算法的处理效果和运算时间,在该缩放应用处理中,本文提出的算法具有更大的优势,并且由于其结构简单,更利于使用 FPGA 的硬件逻辑编写实现进一步加速。

表 1 定量分析对比

Table 1 Comparison of quantitative analysis

客观评价指标	最邻近插值	双线性插值	双三次插值	本文算法
MAE	0.0274	0.0268	0.0257	0.0252
MSE	0.0026	0.0021	0.0020	0.0020
PSNR	25.9162	26.7568	27.0400	27.0521
SNR	18.0160	18.8565	19.1397	19.1518

表 2 4 种算法的 MATLAB 运行时间对比

Table 2 Comparison of MATLAB running time among four algorithms

算法	运行时间/ms
最邻近插值	3.02
双线性插值	13.73
双三次插值	25.66
本文算法	3.44

3 图像缩放 IP 核设计及仿真

缩放 IP 核的开发软件使用的是 Xilinx 公司的 VIVADO 设计套件,并用 Verilog HDL (Hardware Description Language) 进行硬件逻辑的编写。缩放图像中输入点对输出点的影响系数可以合并为以下 3 个:0.015 625、0.046 875、0.140 625。在硬件实现的时候为了提高处理速度,尽量避免浮点数运算,又要同时保持数据的精度,通常采用去尾法来处理浮点数,这 3 个系数可以写成式(8)的形式:

$$\begin{cases} 0.015625 = (1/2)^6 \\ 0.046875 = (3/2)^6 \\ 0.140625 = (9/2)^6 \end{cases} \quad (8)$$

因此在硬件逻辑转化时将参数扩大 2⁶ 倍转

换为整数运算,在二进制运算中即向左移 6 位,所以参数选用 1、3、9,分别用二进制表示为 4'b0001、4'b0011、4'b1001,最后获取处理后的数据时舍弃后 6 位数据即可。每个图像区域处理块中包含 16 个像素点,用 4 位计数器进行计数,根据计数器的数值对应像素块中点的相对位置,利用 case 语句选取对应参数,将输入数据依次乘上对应参数,并分别存入存储器中,进行累加。

图像缩放模块的仿真时序图如图 4 所示,为了方便观察和计算,测试仿真时输入一个 16 列、8 行的像素矩阵块,前 4 行全部输入测试像素 2,后 4 行全部输入测试像素 6,缩放后的结果应为一个 4 列、2 行的像素块,像素值依次为 2、2、2、2、6、6、6、6,见图 4 仿真波形中的 small_data。

将使用 Verilog HDL 编写的图像缩放 IP 核例化封装并加入到工程的 IP 核库中,在模块设计的时候添加该 IP 核即可完成调用。

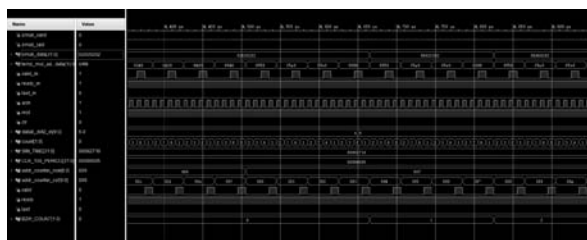


图 4 图像缩放模块仿真波形

Fig.4 Simulation waveform of image scaling module

4 系统平台测试及性能分析

图 5 中只展示了系统部分关键模块图,包括设计的图像缩放 IP 核、ARM 处理器模块、摄像头读取模块以及视频图形阵列 (Video Graphics Array, VGA) 显示模块。

系统测试的主要器材包括一个 Zynq-7000 开发板、OV7670 摄像头和 VGA 显示器。利用 FlightGear 飞行模拟器获得无人机着陆过程的视频,用一台笔记本播放该视频,通过摄像头获取

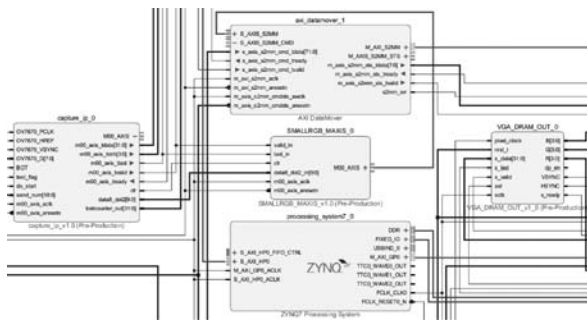


图 5 系统部分模块图

Fig.5 Part of module diagram of system

无人机着陆的图像信息,并在 VGA 中实时显示处理后的结果。实验设备和实验结果见图 6。

图 6(b)中 VGA 显示器中显示了两部分图片,一部分是输入的原图像,一部分是左上角显示的缩小后的处理后的图片。通过串口助手获取摄像头采集到的输入图像以及经过 IP 核处理后的图像数据。在 MATLAB 中将摄像头捕捉到的输入图像用本文提出的缩放算法处理,并和硬件 IP 核处理后的图片作对比,处理时间见表 3,SoC 的处理速度比 MATLAB 快了 171 倍,基本实现了 0 延时。



(a) 系统测试



(b) VGA 显示图

图 6 系统测试图

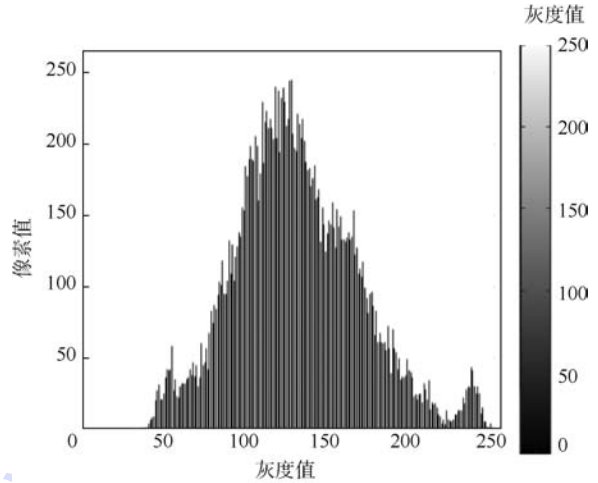
Fig. 6 System test chart

表 3 本文算法的 MATLAB 和 SoC 处理时间对比

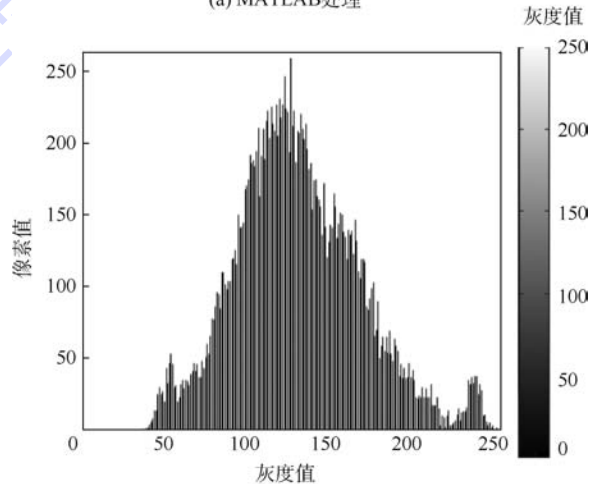
Table 3 Comparison of MATLAB and SoC processing time of proposed algorithm

处理方式	处理时间/ms
MATLAB	3.44
SoC	0.02

MATLAB 和 SoC 处理后的灰度直方图分别如图 7(a)、(b)所示,2 幅图的像素点分布基本一致,处理后图像对应位置像素点的差值绝对值最大为 0.0037,最小为 0。由此可见,利用 SoC 进行图像缩放模块的设计,不仅运行速度大大提升,处理精度也非常高。



(a) MATLAB 处理



(b) SoC 处理

图 7 MATLAB 和 SoC 处理后直方图对比

Fig. 7 Histogram contrast after MATLAB and SoC processing

5 结 论

1) 本文采用了 FPGA + ARM 的 SoC 软硬件处理平台,综合了 FPGA 和 ARM 的优点,具有并行处理的能力和分系统控制能力,应用于图像处理中能够加速算法的处理速度。

2) 针对无人机自主着陆视觉算法中的图像缩放处理,提出了一种适用于硬件加速的图像缩放算法,在处理速度上仅次于结构简单的最邻近插值法,且处理效果最优。

3) 利用硬件描述语言实现本文的缩放算法,并封装成硬件 IP 核,在图像处理系统中的处理速度相较于软件实现提升了 171 倍,且处理后的像素误差最大值为 0.0037。

4) 整个系统采用了 ARM 作为中央控制器协调各 IP 核工作,具有很强的通用性,可以通过增

减IP核来改进系统,适用于进一步的研究,以及图像处理IP核的设计。

参考文献 (References)

- [1] RETTKOWSKI J, BOUTROS A, GOHRINGER D. HW/SW co-design of the HOG algorithm on a Xilinx Zynq SoC[J]. Journal of Parallel and Distributed Computing, 2017, 109(1): 50-62.
- [2] KRYJAK T, KOMORKIEWICZ M, GORGON M. Real-time hardware-software embedded vision system for its smart camera implemented in Zynq SoC[J]. Journal of Real-Time Image Processing, 2016, 12(4): 1-37.
- [3] SENOUCI B, CHARFI I, HEYRMAN B, et al. Fast prototyping of a SoC-based smart-camera: A real-time fall detection case study[J]. Journal of Real-Time Image Processing, 2016, 12(4): 649-662.
- [4] SVETEK A, BLAKE M, HERMIDA M C, et al. The calorimeter trigger processor card: The next generation of high speed algorithmic data processing at CMS[J]. Journal of Instrumentation, 2016, 11(2): 201-210.
- [5] ZARANDY A, NEMETH M, NAGY Z, et al. A real-time multi-camera vision system for UAV collision warning and navigation[J]. Journal of Real-Time Image Processing, 2016, 12(4): 709-724.
- [6] 刘镇弢, 李涛, 黄虎才, 等. 一种用于实时图像处理的众核结构设计[J]. 西安电子科技大学学报, 2015, 42(2): 95-101.
LIU Z T, LI T, HUANG H C, et al. A design of the core structure for real-time image processing[J]. Journal of Xidian University, 2015, 42(2): 95-101 (in Chinese).
- [7] 杨帆, 张皓, 马新文, 等. 基于FPGA的图像处理系统[J]. 华中科技大学学报(自然科学版), 2015, 43(2): 119-123.
YANG F, ZHANG H, MA X W, et al. Image processing system based on FPGA[J]. Journal of Huazhong University of Science and Technology(Natural Science Edition), 2015, 43(2): 119-123 (in Chinese).
- [8] ZHAI X, ALI A S, AMIRA A, et al. MLP neural network based gas classification system on Zynq SoC[J]. IEEE Access, 2016, 4(2): 8138-8146.
- [9] HAN Y, VIRUPAKSHAPPA K, VITORSILVAPINTO E, et al. Hardware/software co-design of a traffic sign recognition system using Zynq FPGAs[J]. Electronics, 2015, 4(4): 1062-1089.
- [10] KELLY C, SIDDIQUI F M, BARDAK B, et al. FPGA soft-core processors, compiler and hardware optimizations validated using HOG[C] // International Symposium on Applied Reconfigurable Computing. Berlin: Springer, 2016, 1: 78-90.
- [11] ALTUNCU M A, GUVEN T, BECERIKLI Y, et al. Real-time system implementation for image processing with hardware/software co-design on the Xilinx Zynq platform[J]. International Journal of Information and Electronics Engineering, 2015, 5(6): 473-477.
- [12] KRAJNIK T, SVAB J, PEDRE S, et al. FPGA-based module for SURF extraction[J]. Machine Vision and Applications, 2014, 25(3): 787-800.
- [13] GAO F, HUANG Z, WANG S, et al. Optimized parallel implementation of face detection based on embedded heterogeneous many-core architecture[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2017, 31(7): 175-180.
- [14] ISHIKAWA S, TAKAHASHI T, WATANABE S, et al. High-speed X-ray imaging spectroscopy system with Zynq SoC for solar observations[J]. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 2017, 22(1): 40-52.
- [15] CAI W, XU Z, LI Z. A high performance surf image feature detecting system based on Zynq[J]. DEStech Transactions on Computer Science and Engineering, 2017, 11(2): 101-110.
- [16] 王博. 数字图像缩放及其质量评价方法研究[D]. 哈尔滨: 哈尔滨工程大学, 2015: 97-105.
WANG B. Digital image zoom and its quality evaluation method [D]. Harbin: Harbin Engineering University, 2015: 97-105 (in Chinese).

作者简介:

王鹏 男, 硕士研究生。主要研究方向: 嵌入式、计算机视觉、无人机先进控制。

曹云峰 男, 硕士, 教授, 博士生导师。主要研究方向: 无人机飞行控制与导航、机器视觉与人工智能、基于模型的系统设计技术。

许蕾 男, 硕士, 讲师。主要研究方向: 嵌入式、无人机智能控制。

丁萌 男, 博士, 副教授, 硕士生导师。主要研究方向: 计算机视觉、无人机导航、制导与控制。

SoC collaborative acceleration design method for image scaling algorithm

WANG Peng¹, CAO Yunfeng^{1,*}, XU Lei¹, DING Meng², ZHANG Zhouyu¹, QU Jinqiu¹

(1. College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;

2. College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: Aimed at the problem that a large number of images need to be scaled in the visual algorithm for the runway detection, recognition and tracking of the unmanned aerial vehicle with high real-time requirement, a new image scaling algorithm suitable for hardware acceleration is proposed based on the mapping relation of the input-output pixel. By simplifying the algorithm structure and using the field programmable gate array to design the hardware function of the module, the algorithm accelerates, and the real-time image processing system is built by the software and hardware cooperative architecture. The experimental results show that the improved scaling algorithm has high precision and less time consumption, and it can speed up by 171 times with the hardware logic. The hardened system can get the image data through the camera, and the real-time processing is displayed in the monitor, which has 30 frame/s processing speed. It can be applied to the image processing algorithm with high real-time requirement.

Keywords: image scaling; real-time image processing; system on chip (SoC); software and hardware collaboration; unmanned aerial vehicle vision

Received: 2018-05-30; **Accepted:** 2018-10-15; **Published online:** 2018-10-29 09:17

URL: kns.cnki.net/kcms/detail/11.2625.V.20181025.1141.002.html

Foundation items: National Natural Science Foundation of China (61673211); Foundation of Innovation and Excellence Fund for Doctoral Dissertations in NUAA (BCXJ18-11); the Fundamental Research Funds for the Central Universities (kfjj20171502)

* **Corresponding author.** E-mail: cyfac@nuaa.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2018.0224

基于地形匹配的直升机低空飞行前视告警方法



张硕俨, 陆洋*

(南京航空航天大学 直升机旋翼动力学国家级重点实验室, 南京 210016)

摘 要: 为解决直升机低空飞行时的防撞告警问题, 提出一种利用地形匹配来修正直升机与地形之间相对位置误差的前视告警方法。首先, 利用直升机飞行动力学模型预测逃逸轨迹, 在此基础上生成告警边界。然后, 将机载地形数据库与雷达探测的高程数据从空域变换到频域, 通过基于互功率谱算法的地形匹配方法计算得到直升机与地形的相对位置误差, 以获取二者之间的准确相对位置。由于地形匹配效果受地形起伏程度影响, 采用地形熵选择匹配区域方法在进行地形匹配前剔除地形起伏较小的匹配区域, 以此提高匹配准确度。最后, 基于UH-60直升机飞行动力学模型设置了1 000组对照算例进行告警方法仿真测试。仿真结果表明: 采用所提告警方法的测试组相比于常规告警方法, 虚警率降低约16%, 告警成功率提高近30%。说明所提的前视告警方法能够有效实现直升机低空飞行防撞告警。

关键词: 直升机; 低空飞行; 前视告警; 地形匹配; 告警成功率; 虚警率

中图分类号: V249.3

文献标识码: A **文章编号:** 1001-5965(2019)02-0340-07

直升机在飞行过程中, 由于缺乏对周围地形感知而发生坠毁的情况, 称为可控飞行撞地。利用直升机前视告警系统可有效避免此类事故的发生^[1-2]。然而, 现有的直升机前视告警系统通常只适用于飞行高度较高的巡航飞行。对于经常需要执行低空飞行任务的直升机, 如医疗救护直升机、武装直升机等, 当其离地高度很低(通常小于100 m)时, 由于告警系统的告警边界过大, 会产生持续虚警。因此, 为减少直升机低空飞行时告警系统的虚警, 需在确保飞行安全的前提下缩小前视告警边界。然而, 由于导航误差与地形数据库误差的存在, 使得直升机与地形的相对位置也存在一定误差。若仅缩小直升机前视告警边界而不做任何修正处理, 将会导致告警失败。可见, 如何准确获取直升机与地形间的相对位置是直升机低空飞行前视告警的关键。

国内外对于准确获取运动物体与地形间相对位置这一问题, 主要集中在导弹、固定翼飞机、潜艇等领域。文献[3]采用地形轮廓匹配算法准确获取导弹与地形的相对位置, 但该算法不允许飞行器进行机动飞行; 文献[4]采用桑迪亚惯性地形辅助导航算法修正战斗机与地形的相对位置误差, 但该算法在地形梯度变化剧烈区域易造成滤波发散。对于直升机而言, 经常需要在复杂的地形上空进行机动低空飞行, 因此上述2种方法都不适用于直升机低空飞行告警。而文献[5]采用的互功率谱算法, 可准确获取潜艇与地形的相对位置, 该算法作为一种性能优良的图像匹配算法, 不仅对地形适应性强, 而且在匹配时允许潜艇作机动航行^[6]。

由于直升机对地形匹配的技术需求与潜艇具有较大的相似性, 因此, 本文基于互功率谱算法,

收稿日期: 2018-04-23; 录用日期: 2018-09-07; 网络出版时间: 2018-10-09 15:04

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180929.1743.002.html

基金项目: 国家自然科学基金(51375229)

* 通信作者. E-mail: luyang@nuaa.edu.cn

引用格式: 张硕俨, 陆洋. 基于地形匹配的直升机低空飞行前视告警方法[J]. 北京航空航天大学学报, 2019, 45(2): 340-346.
ZHANG S Y, LU Y. Helicopter forward looking alert method for low-altitude flight based on terrain matching[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 340-346 (in Chinese).

开展基于地形匹配的直升机低空飞行前视告警方法研究。首先,建立直升机逃逸轨迹模型并确定直升机前视告警边界;然后,基于互功率谱算法进行地形匹配,以准确获取直升机与地形的相对位置,并利用地形熵选择匹配区域提高匹配精度;最后,构建测试算例对本文提出的方法进行仿真验证。

1 直升机低空飞行前视告警方法

直升机前视告警方法的核心是告警边界,目前告警边界的设计方法主要有美国 Honeywell 公司的告警边界设计方法^[7]和基于逃逸轨迹的告警边界设计方法^[8]。其中,Honeywell 的告警边界为考虑通用性,其边界设计趋于保守,低空飞行时虚警率较高,不适用于直升机低空飞行。基于逃逸轨迹的告警边界设计方法则基于直升机的飞行性能,能够在保证安全的基础上最大程度的缩小告警边界以减小虚警率,因此本文采用该方法。

基于逃逸轨迹的告警边界设计方法,首先需要确定直升机的逃逸轨迹,然后在此基础上设计告警边界。

1.1 直升机逃逸轨迹

直升机的逃逸过程通常可分为 3 个阶段:飞行员反应、拉起改出和稳定爬升。其中,飞行员反应阶段保持直升机当前状态飞行,稳定爬升阶段以拉起改出阶段最终时刻的状态为基准进行匀速飞行,这两部分均为直线运动轨迹;而拉起改出为机动飞行状态,其轨迹为复杂的空间曲线,是逃逸轨迹中最为关键的部分,直接决定了逃逸轨迹的整体形态。图 1 示出了直升机逃逸轨迹的 3 个主要组成部分。

由文献[9]可知,直升机逃逸轨迹的拉起改出部分与抛物线方程拟合度较高,该部分逃逸轨迹高度 H 可表示为

$$H = c_0 + c_1x + c_2x^2 \quad (1)$$

式中: x 为轨迹的横向坐标; c_0 、 c_1 和 c_2 为待定拟合系数。拟合系数的大小与直升机的飞行性能相关,通常需要基于飞行试验数据拟合求解,亦可利

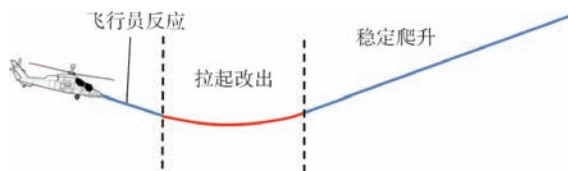


图 1 直升机逃逸轨迹各部分组成示意图

Fig. 1 Schematic diagram of each part of helicopter escape trajectory

用高精度直升机飞行动力学模型获得,本文采用第 2 种方法。

在实际应用时,对于给定机型,首先离线计算各种不同组合飞行状态下逃逸轨迹拉起改出部分的拟合系数,在实际飞行过程中,通过多维插值即可获得该飞行状态下的系数值,从而得到相应逃逸轨迹的拉起改出部分。

1.2 直升机前视告警边界

参考文献[10],将逃逸轨迹中飞行员反应阶段的时间设定为 2 s。此外,由于传感器误差以及地形数据库误差等因素的存在,直升机在飞行过程中应在其下方留有一段安全距离,即最小安全高度。在现有直升机近地告警产品中,如美国 Honeywell 公司的 MXKKII,将最小安全高度的最大值设为 500 feet(对应高度 152.4 m),对于直升机低空飞行而言,显然不能直接采用。因此,本文根据传感器误差、地形数据库误差以及地形平均斜率等因素,将最小安全高度设为 50 m。以 UH-60 直升机为例,设飞行初始状态为:前飞速度为 40 m/s,下降率为 4 m/s,则得到告警边界如图 2 所示。图中: x 、 z 为轨迹坐标。

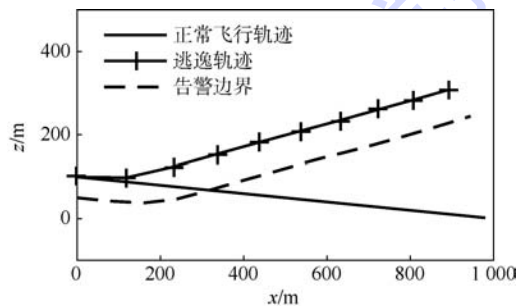


图 2 UH-60 直升机逃逸轨迹及告警边界示意图

Fig. 2 Schematic diagram of escape trajectory and alert envelope for UH-60 helicopter

2 直升机/地形相对位置的准确获取

在确定了直升机前视告警边界后,须进一步准确获取直升机与地形的相对位置,才能进行有效告警。此外,为避免匹配区域选择不当导致的匹配失败,在进行地形匹配之前,利用地形熵选择匹配区域,以提高匹配的准确性。

2.1 直升机低空飞行地形匹配方法

互功率谱算法的基本原理是利用快速傅里叶变换将 2 个相关矩阵由空域变换到频域,再根据傅里叶变换的位移性质,得到一个反映两者相对误差的脉冲函数^[11-12]。将该方法应用于直升机与地形匹配时,须将直升机雷达探测得到的地形

高程矩阵和机载地形数据库的地形高程矩阵进行快速傅里叶变换,得到一个互功率谱,由归一化的互功率谱计算出一个脉冲函数,由该函数即可得到直升机与地形的相对位置误差。

直升机雷达探测的地形高程矩阵和机载地形数据库地形高程矩阵的互功率谱 $G(u, v)$ 可表示为

$$G(u, v) = \frac{F_1^*(u, v)F_2(u, v)}{|F_1^*(u, v)F_2(u, v)|} = e^{-j2\pi(u\Delta x + v\Delta y)} \quad (2)$$

式中: $F_1(u, v)$ 和 $F_2(u, v)$ 为 2 个地形高程矩阵对应的频谱, u, v 为 x, y 对应的频域坐标; $F_1^*(u, v)$ 为 $F_1(u, v)$ 的复共轭; $\Delta x, \Delta y$ 为 2 个地形高程矩阵间的水平偏差值。对式(2)进行傅里叶逆变换后,可以得到一个水平面上的脉冲函数,表示为

$$F^{-1}(e^{-j2\pi(u\Delta x + v\Delta y)}) = \delta(x - \Delta x, y - \Delta y) \quad (3)$$

当直升机与地形的相对位置误差为 0 时,该脉冲函数会产生一个明显的冲激峰,其值为 1,而其他区域则为 0,如图 3(a)所示。当相对位置误差较小时,2 个地形高程矩阵存在的水平偏移量引起的脉冲函数将产生一个明显的冲激峰。相对位置误差越小,2 个地形高程矩阵重叠区域越大,则冲激峰值越接近 1,其他区域的值越接近于 0,如图 3(b)所示。此时冲激峰的位置坐标即为直升机与地形的相对位置误差值。而当相对位置误差很大时,该函数的能量将从单一峰值分散到许多小峰值,如图 3(c)所示。图中: X, Y 为脉冲函数坐标。

2.2 不同地形对匹配方法的影响分析

此处通过仿真测试,考察不同地形起伏程度对互功率谱算法的有效性影响。首先利用马尔可夫算法随机生成一个起伏程度较大的地形^[13],大小为 220×220 矩阵网格,地形高程均值为 487 m,极差为 55 m,方差为 81 m^2 ,如图 4 所示。

在该地形中取出 2 个 200×200 的矩阵作为待匹配地形,两矩阵 X, Y 方向的平移偏差值均为 20。对 2 个待匹配地形矩阵进行快速傅里叶变换,再计算得到脉冲函数,如图 5 所示。该脉冲函数中存在一个明显的冲激峰,其坐标为 (20, 20),峰值为 0.44,由此可得到 2 个待匹配地形矩阵 X, Y 方向的平移偏差值均为 20。

进一步生成一个起伏程度较小的随机地形,地形高程均值为 497 m,极差为 5 m,方差为 1.6 m^2 ,如图 6 所示。

同样在该地形中取出 2 个 200×200 的矩阵作为待匹配地形,两矩阵 X, Y 方向的平移偏差值

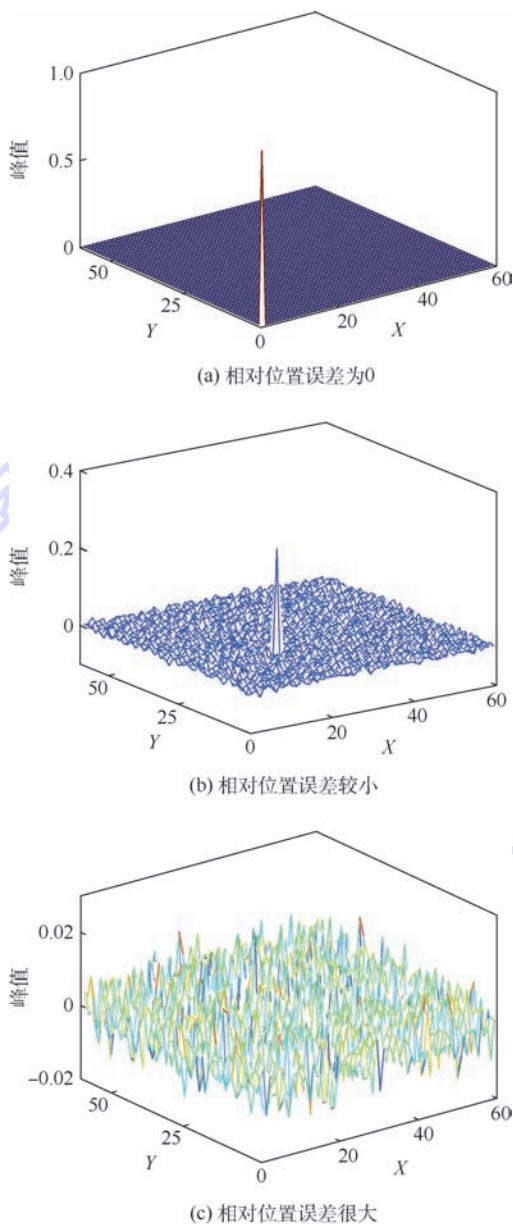


图 3 不同相对位置误差下的脉冲函数
Fig. 3 Pulse functions with different relative position errors

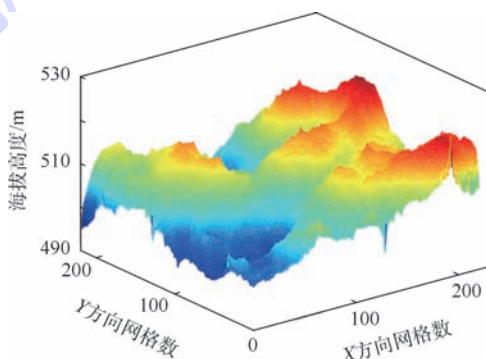


图 4 起伏较大的随机地形
Fig. 4 Undulating random terrain

均为 20。计算得到脉冲函数如图 7 所示。可见,该脉冲函数中不存在明显的冲激峰,此时匹配失败。

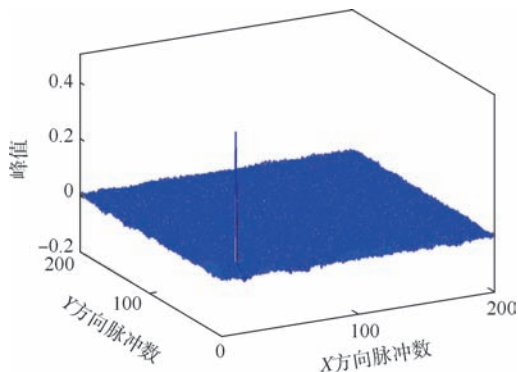


图5 匹配成功时的脉冲函数

Fig. 5 Pulse function on successful matching

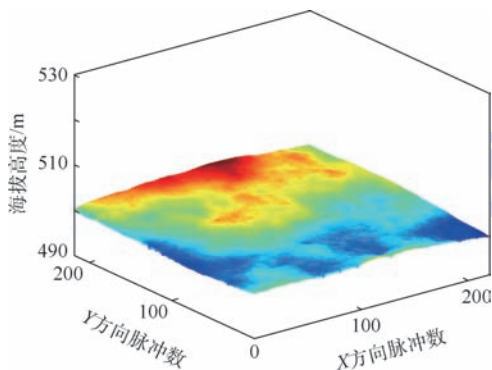


图6 起伏较小的随机地形

Fig. 6 Small undulating random terrain

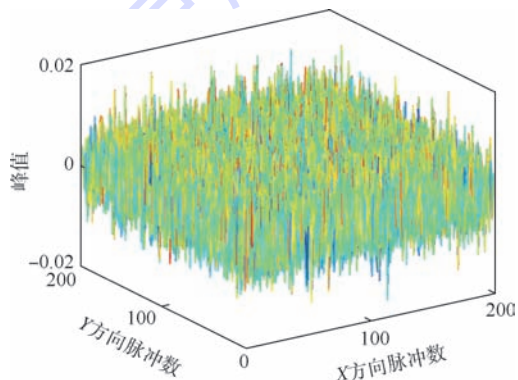


图7 匹配失败时的脉冲函数

Fig. 7 Pulse function on failed matching

上述仿真测试表明,互功率谱算法的匹配效果受地形起伏程度影响很大,地形起伏程度越大,地形越独特,匹配准确度越高,因此山地等地形较适合地形匹配,而平原、沙漠等地形易导致匹配失败。

2.3 地形熵选择匹配区域

针对平原、沙漠等包含信息量较少的区域,本文参考文献[14],采用地形熵选择匹配区域方法来解决此问题。

地形熵能够用来描述地形起伏的剧烈程度,当某区域的局部地形高度值变化较为剧烈时,地

形所包含的信息量比较丰富,此时计算得到的地形熵较小,而越小的地形熵越有利于进行匹配运算。地形熵的定义式为

$$\begin{cases} p(i) = \frac{|h(i)|}{\sum_{i=1}^M |h(i)|} \\ H_{\text{terrain}} = -\sum_{i=1}^M p(i) \lg p(i) \end{cases} \quad (4)$$

式中: p 为无量纲的地形高度值; h 为地形高度值; i 为地形采样点的序号; M 为地形采样点个数的总量; H_{terrain} 为地形熵。

在进行地形匹配前,首先计算直升机当前位置附近区域的地形熵,若地形熵小于预先设定阈值,则进行地形匹配。这样,避免了在平原等区域进行匹配导致的匹配失败,有效提高了地形匹配的准确性。

以图4地形为例,可取出共441个 200×200 的高程矩阵,参考文献[15]将地形熵阈值设为9,计算得到有25个高程矩阵的地形熵大于阈值,而其中有19个匹配失败。将该25个高程矩阵剔除后,剩余416个矩阵均能够匹配成功。

3 算法仿真测试

本文利用MATLAB开展对基于地形匹配的直升机低空飞行前视告警方法的仿真测试。首先给出算法性能的评价指标,然后给出仿真测试的参数设置,最后给出仿真结果及分析。

3.1 算法性能评价指标

根据文献[16],此处采用虚警率和告警成功率作为评价告警方法的性能指标。

虚警是指告警系统在直升机安全飞行时给出了不必要的告警,虚警率定义为

$$P(\text{UA}) = \frac{N_u}{N} \quad (5)$$

式中: N_u 为虚警算例个数; N 为算例总数。

告警成功率则是相对告警失败而言的。告警失败是指告警提示过晚,飞行员得到告警信号时已经无法及时改出而导致撞地事故。告警成功率定义为

$$P(\text{SA}) = \frac{N - N_f}{N} \quad (6)$$

式中: N_f 告警失败算例个数。

表1给出了前视告警方法的性能指标判断标准。

需要说明的是,计算虚警率和告警成功率的算例总数仅包括产生告警的算例。

表1 前视告警方法性能指标判断标准

Table 1 Judgement criteria for performance index of forward looking alert method

告警边界与地形是否相交	正常轨迹与地形是否相交	逃逸轨迹与地形是否相交	告警情况
是	是	否	告警成功
是	是	是	告警失败
是	否	否	虚警

3.2 仿真参数设置

为对比验证本文提出的基于地形匹配的直升机低空飞行前视告警方法的优势,仿真中设置了2个测试组,一个为采用本文方法的匹配组,另一个为采用常规告警方法的对照组,2个测试组的告警边界相同,均采用基于逃逸轨迹的告警边界,最小安全高度设为50 m,但对照组未结合地形匹配算法,无法修正直升机与地形的相对位置误差。

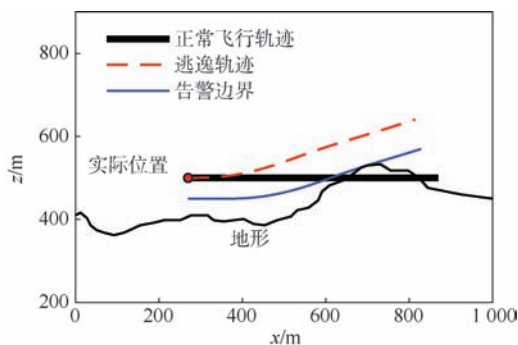
1) 地形参数设置:类似图4,仿真地形采用马尔可夫算法随机生成的500组地形,地形高程均值约为400 m,地形矩阵大小均为 220×220 ,设每个矩阵网格大小为90 m,总大小即为 $19\,800 \text{ m} \times 19\,800 \text{ m}$,2个测试组均在500组地形进行仿真飞行测试。

2) 直升机飞行参数设置:仿真基于UH-60直升机飞行动力学模型。直升机由地形矩阵坐标(20,100)处出发,在500 m海拔高度以40 m/s的速度平飞,无线电高度均值约为100 m,终点为矩阵坐标(160,100)处,每0.5 s主动探测直升机周围 10×10 矩阵网格的地形数据,并计算地形熵,当地形熵小于9时,则进行地形匹配。飞行期间发生告警或到达终点则停止该次仿真测试。

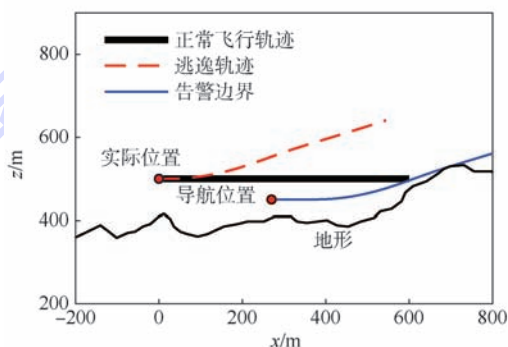
3) 误差参数设置:测试开始时,2个测试组直升机与地形的相对位置误差均为3个矩阵网格,即270 m,匹配组与对照组中各有一半测试组的误差方向沿直升机飞行方向,即直升机导航位置位于实际位置前方,另一半误差方向相反,即直升机导航位置位于实际位置后方。

3.3 仿真结果与分析

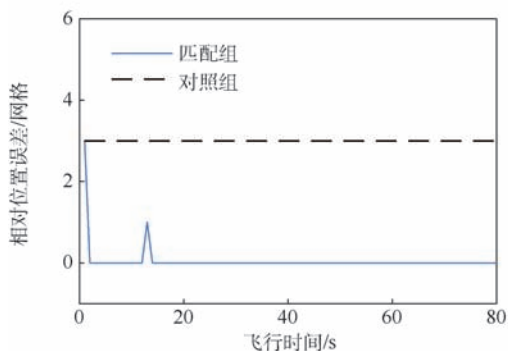
由3.2节可知,在仿真测试中,对照组存在导航位置误差,告警边界以导航位置为准,而正常飞行轨迹和逃逸轨迹则以实际位置为准。取仿真测试中的典型算例为例,如图8和图9所示。图8(a)、图8(b)为测试中同一地形下的2个测试组,其中匹配组告警成功,而对照组产生虚警。图8(a)中,告警边界、正常飞行轨迹与地形相交,而逃逸轨迹与地形不相交,此时告警成功。图8(b)中导航位置在实际位置前方,告警边界与



(a) 匹配组告警成功



(b) 对照组虚警



(c) 相对位置误差对比

图8 对照测试 A

Fig. 8 Controlled trial A

地形相交,而正常飞行轨迹、逃逸轨迹与地形不相交,此时为虚警。图8(c)为该地形下,2个测试组直升机与地形相对位置误差随时间变化的曲线。从图中可以看出,仿真开始时2组的相对位置误差均为3个矩阵网格,但匹配组通过地形匹配修正了这一误差。

图9(a)、图9(b)为仿真测试中同一地形下的另外2个测试组,其中匹配组告警成功而对照组告警失败。图9(a)所述状态与图8(a)相似,为告警成功。而图9(b)中导航位置在实际位置后方,告警边界、正常飞行轨迹、逃逸轨迹均与地形相交,此时告警失败。图9(c)为该地形下,2个测试组直升机与地形相对位置误差随时间变化的曲线。同样可以从图中看出,匹配组通过地形匹配修正了相对位置误差。

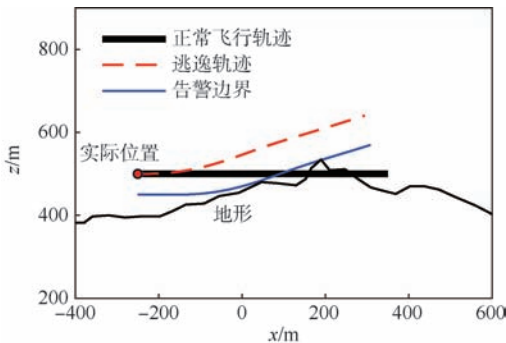
完成对照组、匹配组各 500 组仿真测试算例后,得到的测试结果如表 2 所示。

从表 2 的仿真结果可以看出,对于采用地形匹配的本文告警方法,相比于常规告警方法,虚警率明显降低(约 16%),告警成功率显著提高(近 30%)。

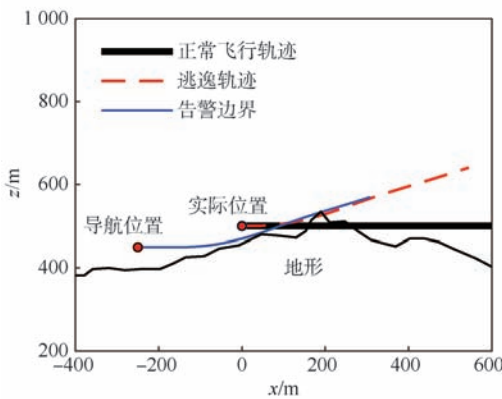
形的相对位置误差,并据此提出了直升机低空飞行前视告警方法,有效实现了直升机低空飞行防撞告警。主要结论如下:

1) 互功率谱算法的匹配效果受地形起伏程度影响,地形起伏程度越大,地形越独特,匹配准确度越高。

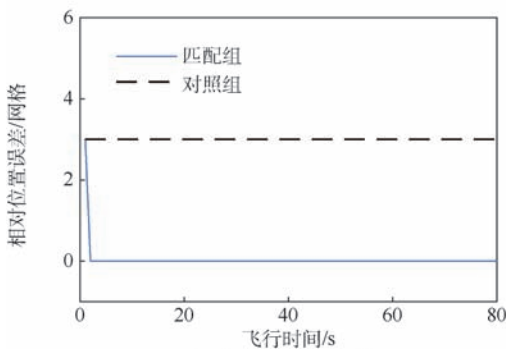
2) 对于直升机低空飞行状态,未结合地形匹配的常规告警方法由于直升机与地形的相对位置不准确,导致告警成功率低、虚警率高。结合地形匹配后的前视告警方法告警成功率显著提高,虚警率明显降低,证明了本文提出的前视告警方法能够很好地适应直升机低空飞行的防撞告警要求。



(a) 匹配组告警成功



(b) 对照组告警失败



(c) 相对位置误差对比

图 9 对照测试 B

Fig.9 Controlled trial B

表 2 告警方法仿真测试结果

Table 2 Simulation test results of alert method

性能指标	匹配组	对照组
告警成功率/%	95.47	65.93
虚警率/%	3.77	19.34

4 结 论

本文将基于互功率谱算法的图像匹配方法应用于直升机地形匹配,有效地修正了直升机与地

参考文献 (References)

- [1] 刘仪. 防范可控飞行撞地事故的最新进展[J]. 国际航空, 2001(12):41-42.
LIU Y. The new development of preventing CFIT[J]. International Aviation, 2001(12):41-42 (in Chinese).
- [2] 杨富民. 直升机地形防撞告警理论与技术研究[D]. 郑州: 解放军信息工程大学, 2012:1-3.
YANG F M. Research on theory and technology of helicopter terrain collision warning[D]. Zhengzhou: Information Technology University of PLA, 2012:1-3 (in Chinese).
- [3] GOLDEN J P. Terrain contour matching (TERCOM): A cruise missile guidance aid[C] // Proceedings of 24th SPIE Annual Technical Symposium. Bellingham; SPIE, 1980, 238:10-18.
- [4] BAIRD C A. Terrain-aided altitude computations on the AFTL/F-16[C] // Proceedings of IEEE Position Location & Navigation Symposium. Piscataway, NJ: IEEE Press, 1990:474-481.
- [5] 徐遵义, 黄东武, 赵洪鑫. 水下地形辅助导航新方法仿真[J]. 舰船科学技术, 2012, 34(6):53-56.
XU Z Y, HUANG D W, ZHAO H L. A new method for underwater terrain-aided navigation[J]. Ship Science and Technology, 2012, 34(6):53-56 (in Chinese).
- [6] 曹闻, 李炳程, 邓子建. 一种基于小波变换的图像配准方法[J]. 测绘通报, 2004(2):16-19.
CAO W, LI B C, DENG Z J. An image registration method based on wavelet transformation[J]. Bulletin of Surveying and Mapping, 2004(2):16-19 (in Chinese).
- [7] ISHIHARA Y, CONNER K. Method and computer program product for helicopter ground proximity warning system: US6583733B2[P]. 2003-12-03.
- [8] 周成中. 基于逃逸轨迹告警原理的直升机地形感知与告警算法研究[D]. 南京: 南京航空航天大学, 2015:7-9.
ZHOU C Z. Study on helicopter terrain awareness and alert algorithm based on escape trajectory alert[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2015:7-9 (in Chinese).
- [9] ANDERSON T, JONES W, BEAMON K. Design and implementation of TAWS for rotary wing aircraft[C] // Proceedings of Aerospace Conference. Aerospace Conference. Piscataway, NJ:

- IEEE Press, 2011: 1-7.
- [10] WILLIAMS S. Towards terrain-aided navigation for underwater robotics[J]. *Advanced Robotics*, 2001, 15(5): 533-549.
- [11] 郑志彬, 叶中付. 基于相位相关的图像配准算法[J]. *数据采集与处理*, 2006, 21(4): 444-449.
ZHENG Z B, YE Z F. Image registration algorithm based on phase-correlation[J]. *Journal of Data Acquisition & Processing*, 2006, 21(4): 444-449 (in Chinese).
- [12] 朱佳媛. 基于FFT相位相关的图像匹配算法与实现[J]. *科技风*, 2013(14): 11.
ZHU J Y. Image matching algorithm and implementation based on phase correlation of FFT[J]. *Technology Wind*, 2013(14): 11 (in Chinese).
- [13] 张继贤, 柳健, 李德仁. 地形生成技术与方法的研究[J]. *中国图象图形学报*, 1997, 2(9): 638-645.
ZHANG J X, LIU J, LI D R. Technical approach to realistic terrain generation[J]. *Journal of Image and Graphics*, 1997, 2(9): 638-645 (in Chinese).
- [14] 王英钧. 地形辅助导航综述[J]. *航空电子技术*, 1997, 5(1): 24-30.
WANG Y J. A review of terrain-assisted navigation[J]. *Avionics Technology*, 1997, 5(1): 24-30 (in Chinese).
- [15] 席雷平, 陈自力, 李小民. 基于地形熵的地形适配区选择准则研究[J]. *弹箭与制导学报*, 2010, 30(4): 247-249.
XI L P, CHEN Z L, LI X M. Discussion on the navigable terrain selection algorithm based on terrain entropy[J]. *Journal of Projectiles Rockets Missiles and Guidance*, 2010, 30(4): 247-249 (in Chinese).
- [16] FITZGERALD T R, BRUNNER M T. Use of high-fidelity simulation in the development of an F/A-18 active ground collision avoidance system[J]. *AIAA Journal*, 1995, 33(6): 68-78.

作者简介:

张硕伊 男, 硕士研究生。主要研究方向: 直升机飞行安全。

陆洋 男, 博士, 副教授, 硕士生导师。主要研究方向: 直升机飞行动力学与控制。

Helicopter forward looking alert method for low-altitude flight based on terrain matching

ZHANG Shuoyan, LU Yang*

(National Key Laboratory of Science and Technology on Rotorcraft Aeromechanics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: In order to solve the problem of terrain awareness when helicopters fly at low altitudes, a forward looking terrain avoidance method is presented using terrain matching, which can correct the relative position error between helicopter and terrain. First, the recovery trajectory of a helicopter is predicted by the helicopter flight dynamics model, based on which the alert envelope is generated. The elevation data of terrain database and radar detection are transformed from airspace to frequency domain. Then, by using the terrain matching method, which is based on cross power spectrum algorithm, the relative position error between the helicopter and the terrain is calculated to obtain the accurate relative position. As the effect of terrain matching can be influenced by topographic relief amplitude, the terrain entropy selection method is adopted to remove the matching area with less topographic relief before terrain matching. Then the matching accuracy is improved. Finally, based on the flight dynamics model of the UH-60 helicopter, 1 000 sets of control case simulation are conducted to verify the alert method. According to the results, the proposed alarm method can significantly reduce the nuisance alert rate by about 16% and increase the successful alert rate greatly by nearly 30%, compared with conventional alert methods. It is shown that the forward looking alert method proposed in this paper can effectively achieve the helicopter collision avoidance alert at low altitude.

Keywords: helicopter; low-altitude flight; forward looking alert; terrain matching; successful alert rate; nuisance alert rate

Received: 2018-04-23; Accepted: 2018-09-07; Published online: 2018-10-09 15:04

URL: kns.cnki.net/kcms/detail/11.2625.V.20180929.1743.002.html

Foundation item: National Natural Science Foundation of China (51375229)

* Corresponding author. E-mail: luyang@nuaa.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2018.0230

未知环境下无人机集群协同区域搜索算法

侯岳奇^{1,2}, 梁晓龙^{1,2,*}, 何吕龙^{1,2}, 刘流^{1,2}

(1. 空军工程大学 国家空管防相撞技术重点实验室, 西安 710051;

2. 空军工程大学 陕西省电子信息系统系统集成重点实验室, 西安 710051)



摘 要: 针对无人机集群在无先验信息的未知环境中协同搜索的问题,提出了一种以覆盖率为实时搜索奖励的无人机集群协同区域搜索算法。首先建立覆盖分布地图(CDM)来描述任务环境,并采用 Hadamard 积实现 CDM 的快速更新,继而基于 CDM 计算覆盖率来定量描述实时搜索效果。将无人机集群视为一个控制系统,基于分布式模型预测控制理论建立系统的预测模型,并将预测周期内最大覆盖率增量设为奖励函数,采用差分进化算法进行求解,得到最优解作为系统的最优输入。仿真结果表明,所提算法能够对区域进行覆盖搜索,在出现突发情况时,覆盖率远高于平行搜索方法。

关键词: 未知环境; 无人机集群; 协同搜索; Hadamard 积; 覆盖率; 分布式模型预测控制

中图分类号: V279⁺.3; V249.122

文献标识码: A **文章编号:** 1001-5965(2019)02-0347-10

随着高精度影像设备与技术的快速发展,携带照相设备的无人机(Unmanned Aerial Vehicle, UAV)在民用和军事领域得到了广泛应用,例如环境监测、战场监视以及目标搜索等^[1-2]。无人机集群通过无人机之间的协同合作,从而实现整体能力上的涌现,即系统涌现出的能力远超系统内单架无人机能力的总和^[3-4]。相比于单架无人机,利用多架无人机协同执行区域搜索任务得到了越来越广泛的关注^[5-6]。

针对协同搜索问题,诸多学者进行了深入的探索并取得丰硕的成果。离线规划方法^[7-9]将任务区域进行分割,设计各个子区域的覆盖搜索航线,飞行航线固定。该方法的优势在于能够实现任务区域的全覆盖,而在无人机故障、火力威胁等突发情况下该方法受限。文献[5]使用笛卡儿栅格描述环境,赋予每个栅格一个值代表目标分布

的不确定性,设计了搜索回报函数和禁飞区回避策略。在有先验信息的情况下,该方法可以实现重点侦察和覆盖搜索。文献[10]在文献[5]的基础上,考虑通信约束,分析了不同通信约束对协同搜索效率的影响。在通信距离和角度有约束的情况下,该方法能够较好地实现协同区域搜索。此类方法依赖先验信息来建立概率地图,然而在实际应用中,先验信息的获取和栅格量化是十分困难的。文献[11]设计了基于信息素的网格回访机制,引导无人机对目标存在概率较大的区域进行回访搜索,使得无人机能够尽早搜索到更多目标。上述方法^[5,10-11]均假设无人机在相邻栅格之间运动,这种“粗粒度”的运动模型虽然简化了协同搜索决策的解空间,但却在一定程度上降低了决策结果的精细程度。文献[12-14]基于分布式模型预测控制框架,采用纳什最优和粒子群优化

收稿日期: 2018-04-25; 录用日期: 2018-10-15; 网络出版时间: 2018-10-26 16:33

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20181025.1141.003.html

基金项目: 国家自然科学基金(61472443,61703427); 陕西省自然科学基金基础研究计划(2017JQ6035)

* 通信作者. E-mail: afeu_lxl@sina.com

引用格式: 侯岳奇, 梁晓龙, 何吕龙, 等. 未知环境下无人机集群协同区域搜索算法[J]. 北京航空航天大学学报, 2019, 45(2): 347-356. HOU Y Q, LIANG X L, HE L L, et al. Cooperative area search algorithm for UAV swarm in unknown environment [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 347-356 (in Chinese).

相结合的算法,有效地降低了协同搜索决策问题的求解规模和通信负担。虽然上述研究在一定程度上使多无人机具备了协同搜索的能力,但仍存在区域覆盖率较低的问题。存在这一问题的原因在于:缺少专门的引导机制,来引导无人机向未覆盖区域进行搜索。

在执行区域搜索任务时,无人机故障和环境中心火力威胁等突发情况的影响是不容忽视的。考虑到传统的离线规划方法难以应对突发情况,本文提出了一种以覆盖率为实时搜索奖励的协同区域搜索算法,以改善一般在线规划搜索方法覆盖率不高的问题。

1 协同搜索问题建模

1.1 问题描述

多无人机协同搜索问题主要分为2类^[15]:第1类是在环境信息已知的条件下,根据先验信息对区域进行搜索,以尽快发现目标,或者对重点区域进行监控;第2类是在环境信息未知的条件下,对区域进行覆盖。本文针对第2类问题展开研究。UAV 集群携带通信设备和照相设备对特定任务区域展开搜索,区域内目标的位置分布未知,如图1所示。

UAV 按照实时规划的航路对任务区域展开搜索,并通过通信组网模块进行实时通信,为实时决策提供信息支撑,通信内容包括 UAV 状态、环境信息和决策信息等。当某架 UAV 出现故障时,传统的离线规划方法难以适应性地完成搜索任务。相比而言,在线规划方法鲁棒性更强,在出现突发情况时,无人机集群能够继续保持协同,自组织地完成搜索任务。因此,本文主要研究如何建立一种高效的区域覆盖在线规划方法,确保 UAV 集群在尽可能短的时间内对区域进行覆盖。

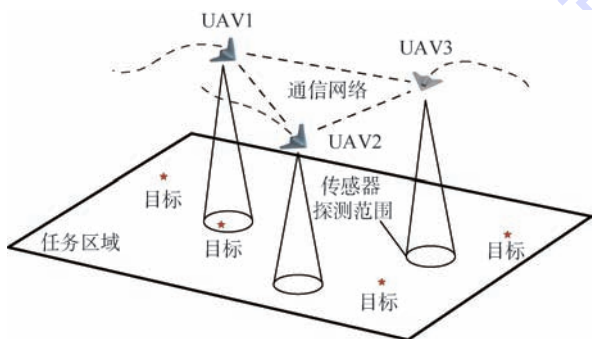


图1 UAV 集群协同搜索示意图

Fig.1 Schematic diagram of UAV swarm cooperative search

1.2 环境模型

设任务区域 Ω 为 $L_x \times L_y$ 的矩形区域,将区域

按照固定间隔 Δd 栅格化为 $M \times N$ 个栅格,如图2所示。

赋予每个栅格一个值 $\mu_{ij}(k)$,用于描述截止到 k 时刻为止栅格 (i,j) 是否已被覆盖。为简化分析,做出如下假设:一旦栅格 (i,j) 处于 UAV 传感器探测范围内,则认为该栅格已被覆盖,且栅格内的目标存在情况完全已知。栅格的状态 $\mu_{ij}(k)$ 表示为

$$\mu_{ij}(k) = \begin{cases} 0 & (i,j) \in \Omega_c(k) \\ 1 & (i,j) \in \Omega_{nc}(k) \end{cases} \quad (1)$$

式中: $\Omega_c(k)$ 为已覆盖的栅格集合; $\Omega_{nc}(k)$ 为未覆盖的栅格集合。任务区域 Ω 为 $\Omega_c(k)$ 和 $\Omega_{nc}(k)$ 的总和。构建覆盖分布地图(Coverage Distribution Map, CDM)来描述任务区域的覆盖分布情况,用矩阵形式来表示覆盖分布地图,定义环境矩阵为 $C(k) = [\mu_{ij}(k)]_{M \times N}$ (2)

在搜索开始前,由于整个区域的环境信息完全未知,且需要对整个区域展开覆盖搜索,因此每个栅格的值 $\mu_{ij}(0) = 1$;随着搜索的进行, $\mu_{ij}(k)$ 实时变化,覆盖分布地图实时更新,并在 UAV 集群内共享,为 UAV 的实时决策提供环境信息。

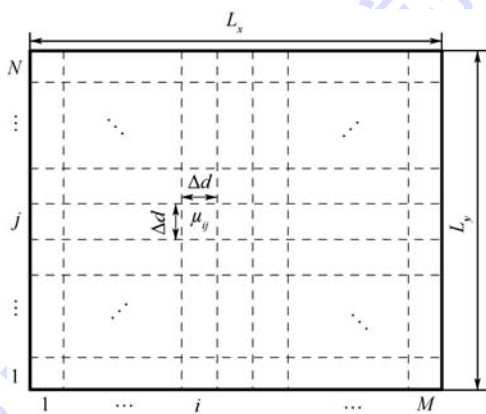


图2 任务区域栅格化

Fig.2 Mission area rasterization

1.3 UAV 模型

为简化分析,假设 UAV 在任务区域上空等高度飞行,并将 UAV 视为二维空间中运动的质点,其运动方程为

$$\begin{bmatrix} x_i(k+1) \\ y_i(k+1) \end{bmatrix} = \begin{bmatrix} x_i(k) \\ y_i(k) \end{bmatrix} + v_0 \Delta t \begin{bmatrix} \cos \varphi_i(k) \\ \sin \varphi_i(k) \end{bmatrix} \quad (3)$$

式中: $[x_i(k), y_i(k)]$ 为 UAV 的位置; $\varphi_i(k) = \varphi_i(k-1) + \Delta \varphi_i(k)$, φ_i 为航向角, $\Delta \varphi_i \in [-\varphi_{\max}, \varphi_{\max}]$ 为航向偏转角,为决策输入, φ_{\max} 为受机动性能限制下的最大转弯角; v_0 为平飞速度; Δt 为时间步长。

将 UAV 集群视为一个控制系统,并将每架

UAV 视为一个子系统。 k 时刻,记子系统 UAV i 的状态为 $\mathbf{p}_i(k) = (x_i(k), y_i(k), \varphi_i(k))$ 。状态方程为

$$\mathbf{p}_i(k+1) = f(\mathbf{p}_i(k), u_i(k)) \quad (4)$$

式中: $u_i(k) = \Delta\varphi_i(k)$ 为子系统 UAV i 的输入; $f(\cdot)$ 为状态转移函数,由式(3)确定。

根据式(4),建立子系统 UAV i 的预测模型为

$$\mathbf{p}_i(k+j|k) = f(\mathbf{p}_i(k+j-1|k), u_i(k+j-1|k)) \quad (5)$$

$$i = 1, 2, \dots, n; j = 1, 2, \dots, H$$

式中: n 为 UAV 架数; H 为预测周期; $\mathbf{p}_i(k+j|k)$ 为基于 $k+j-1$ 时刻 UAV i 状态预测的 $k+j$ 时刻的 UAV i 状态,其值取决于状态 $\mathbf{p}_i(k+j-1|k)$ 和控制输入 $u_i(k+j-1|k)$ 。

从 k 时刻起,给定 H 步预测控制输入后,可以预测出未来 H 步以内 UAV 的航路,如图 3 所示。通过优化预测控制输入,来引导 UAV 尽可能向尚未被覆盖的区域进行搜索,以获得更高的区域覆盖率。

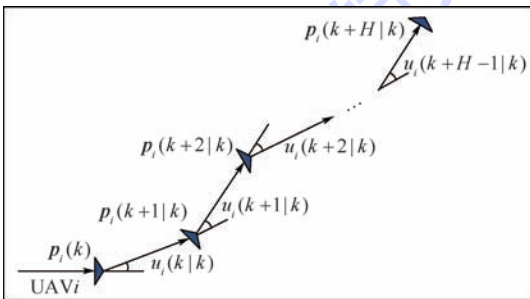


图 3 H 步预测航路图

Fig. 3 H step route prediction

2 覆盖分布地图更新方法

覆盖分布地图描述了任务区域的覆盖搜索情况,是 UAV 进行自主决策的重要依据。因此,如何快速更新覆盖分布地图,对在线实时决策具有重要意义。

覆盖分布地图的更新就是将传感器探测范围内覆盖部分的相应栅格置为 0。通过逐一判断邻近栅格到 UAV 之间的距离是否小于传感器探测距离,并对覆盖范围内的栅格进行逐一赋值,可以实现覆盖分布地图的更新,但这种遍历的方法运算量大、算法复杂度高,不利于实时更新。本文利用 Hadamard 积进行覆盖分布地图更新,操作简单且易于实现,避免了遍历判断和逐一赋值,其流程如图 4 所示。

图 4 中探测矩阵和环境子矩阵的定义,以及对其进行 Hadamard 积的运算过程将在 2.1 节和 2.2 节中详细介绍。

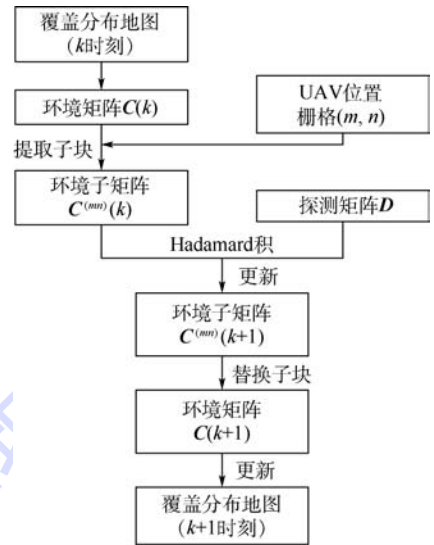


图 4 覆盖分布地图更新流程

Fig. 4 Update process of coverage distribution map

2.1 探测矩阵

本文将传感器探测范围简化为: UAV 所处位置为中心,半径为 R_s 的圆形区域。如图 5 所示,正方形区域记为 Ψ ,将区域 Ψ 栅格化为 $Q \times Q$ 个栅格。

$$Q = 2 \left\lceil \frac{R_s}{\Delta d} - \frac{1}{2} \right\rceil + 1 \quad (6)$$

式中: R_s 为传感器探测半径; $\lceil \cdot \rceil$ 为向上取整函数。

赋予每个栅格一个值 η_{pq} ,用于描述 UAV 传感器能否探测到该栅格。若栅格 (p, q) 处于 UAV 传感器探测范围内,则令 $\eta_{pq} = 0$,反之, $\eta_{pq} = 1$ 。 η_{pq} 表示为

$$\eta_{pq} = \begin{cases} 0 & (p, q) \in \Psi_c \\ 1 & (p, q) \in \Psi_{nc} \end{cases} \quad (7)$$

式中: Ψ_c 为传感器可探测区域; Ψ_{nc} 为不可探测区

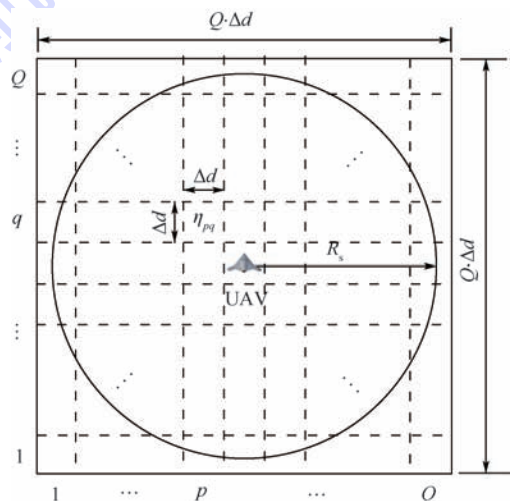


图 5 区域 Ψ 栅格化

Fig. 5 Region Ψ rasterization

域。区域 Ψ 为 Ψ_c 和 Ψ_n 的总和。

以 η_{pq} 为元素建立探测矩阵 $D = [\eta_{pq}]_{Q \times Q}$, 表示 UAV 对邻近栅格的覆盖能力。

2.2 覆盖分布地图的更新

定义环境子矩阵 $C^{(mn)}(k)$: 在环境矩阵 $C(k)$ 中, 以 $\mu_{mn}(k)$ 为中心元素, 维度为 $Q \times Q$ 的子块矩阵, 称为环境子矩阵。

$$C^{(mn)}(k) = [\mu_{pq}^{(mn)}(k)]_{Q \times Q} \quad (8)$$

式中: $(m, n) \in \Omega; p, q = 1, 2, \dots, Q$ 。 $\mu_{pq}^{(mn)}(k)$ 与环境矩阵 $C(k)$ 中的元素 $\mu_{ij}(k)$ 对应关系为

$$\begin{cases} \mu_{pq}^{(mn)}(k) = \mu_{ij}(k) \\ i = p + m - (Q + 1)/2 \\ j = q + n - (Q + 1)/2 \end{cases} \quad (9)$$

当 UAV 处于栅格 (m, n) 内时, 可近似认为 UAV 处于该栅格的中心。此时, 环境子矩阵 $C^{(mn)}(k)$ 与探测矩阵 D 重合, 且维度相等。对上述 2 个矩阵做 Hadamard 积

$$C^{(mn)}(k+1) = C^{(mn)}(k) \circ D = [\mu_{pq}^{(mn)}(k) \times \eta_{pq}]_{Q \times Q} \quad (10)$$

式中: “ \circ ” 为 Hadamard 积运算; $p, q = 1, 2, \dots, Q$ 。

对 $C^{(mn)}(k)$ 和 D 做 Hadamard 积, 对应元素 $\mu_{pq}^{(mn)}(k)$ 与 η_{pq} 相乘, 即对 $C^{(mn)}(k)$ 中的元素 $\mu_{pq}^{(mn)}(k)$ 进行了更新。将更新后的 $C^{(mn)}(k+1)$ 替换到 $C(k)$ 的对应子块中, 即实现了覆盖分布地图的更新。

假设传感器的探测周期为 T_s , 以 T_s 为时间间隔将 UAV 从 k 到 $k+1$ 时刻的直线运动离散为运动点迹。对每个离散点做上述运算, 即可实现一个步长内覆盖分布地图的更新。

2.3 覆盖分布地图的信息融合

文献[16]提出了一种地图信息融合更新方法, 该方法适用于通信理想的情况, 在通信中断或数据丢包的情况下, 部分地图信息无法融合更新, 对搜索过程造成影响。因此, 本文提出基于 Hadamard 积的地图信息融合方法(见图 6), 能够在一定程度上减小通信中断或数据丢包对搜索过程的影响。

在协同搜索过程中, 各 UAV 在本机进行覆盖分布地图更新, 更新完毕后通过通信网络广播发送实现融合共享。 k 时刻, UAV i 本地的覆盖分布地图对应的环境矩阵记为 $C_i(k)$, 并接收到其他 UAV 的广播消息记为 $C_{j \neq i}(k)$ 。基于 Hadamard 积可实现各 UAV 覆盖分布地图的信息融合:

$$C_i(k) = C_i(k) \circ \prod_{j \neq i}^n C_j(k) \quad (11)$$

与文献[16]方法相比, 基于 Hadamard 积的地图信息融合方法共享的探测信息为环境矩阵 $C_i(k)$, 是地图的全局信息。假如 k 时刻数据丢包或通信中断, 当前时刻的探测信息会丢失, 地图信息无法更新。一旦 $k+x$ 时刻通信恢复正常, 环境矩阵 $C_i(k+x)$ 中已包含了 k 到 $k+x$ 时刻的历史探测信息, 经过基于 Hadamard 积的地图信息融合方法更新后, 丢失的历史探测信息就得以恢复。虽然 k 到 $k+x$ 时刻通信中断会影响当前搜索决策, 但是通信恢复后所有历史信息得到恢复, $k+x$ 时刻之后的搜索过程不会受到影响。

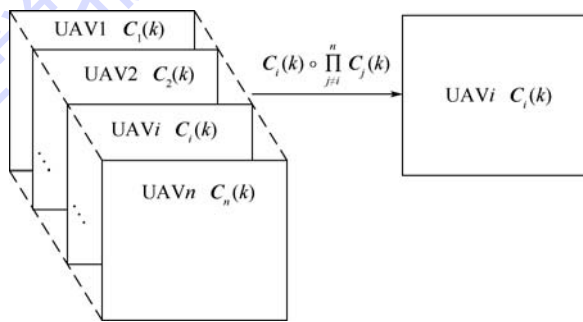


图6 基于 Hadamard 积的地图信息融合示意图

Fig. 6 Schematic diagram of map information fusion based on Hadamard product

3 奖励函数与搜索算法

3.1 搜索奖励函数

协同搜索的关键在于设计一个搜索奖励函数, 对每条预测航路进行评估^[5]。奖励的设定主要是一个步长内覆盖率增量的大小, 并依据边界条件和转弯角度设计惩罚函数。在搜索过程中, UAV 基于当前状态和覆盖分布地图, 利用搜索奖励函数对预测航路进行评估, 并自主地选择奖励值最大的航路作为决策输入。每架 UAV 使用搜索奖励函数 J 选择它的搜索航路:

$$J(p_i(k), u_i(k)) = \omega_1 \gamma J_C(k) + \omega_2 J_T(k) + \omega_3 J_B(k) \quad (12)$$

式中: J_C 为覆盖率增量; J_T 和 J_B 分别为转弯角度和边界距离的惩罚函数; ω_1 、 ω_2 和 ω_3 为相应权重; γ 为重要性因子, 覆盖率的重要性通过调整 γ 来体现, $\gamma \geq 1$ 。

k 时刻的区域覆盖率 $O(k)$ 是已搜索区域 $\Omega_c(k)$ 占任务区域 Ω 的面积比, 即已搜索栅格数量与总栅格数量的比值:

$$O(k) = \frac{MN - \sum_{i=1}^M \sum_{j=1}^N \mu_{ij}(k)}{MN} \quad (13)$$

k 到 $k+1$ 时刻的覆盖率增量是指 $O(k+1)$ 与 $O(k)$ 的差值。实际意义为: k 时刻, 未搜索区域 $\Omega_{nc}(k)$ 中, 在 k 到 $k+1$ 时刻内被搜索到的区域占任务区域 Ω 的面积比。覆盖率奖励函数为

$$J_c(k) = O(k+1) - O(k) = \frac{\sum_{i=1}^M \sum_{j=1}^N (\mu_{ij}(k) - \mu_{ij}(k+1))}{MN} \quad (14)$$

在执行任务过程中, 若转弯角度过大, 导致耗油增大, 影响续航时间。因此, 设计一个惩罚函数, 尽可能减少 UAV 转弯角度过大引起的耗油代价。转弯角度的惩罚函数 J_T 可以表示为

$$J_T(k) = -\frac{|u_i(k)|}{\varphi_{\max}} \quad (15)$$

在搜索过程中, 距离边界越近, 传感器覆盖的有效区域越少, 效率越低。借鉴虚拟势函数的思想, 设计一个惩罚函数, 靠近边界的 UAV 会受到边界的虚拟“斥力”, 距离边界越近则“斥力”越大。因此, 边界距离的惩罚函数 J_B 可以表示为

$$J_B(k) = -\left[\left(\frac{1}{x_i(k)} + \frac{1}{L_x - x_i(k)} \right) + \left(\frac{1}{y_i(k)} + \frac{1}{L_y - y_i(k)} \right) \right] \quad (16)$$

给定 UAV i 的控制输入 $u_i(k)$, 则 UAV i 状态从 $p_i(k)$ 转移到 $p_i(k+1)$, 覆盖分布地图由 $C_i(k)$ 更新为 $C_i(k+1)$, 从而可以计算出奖励函数 J 的值, 优化求解使奖励函数 J 最大, 即解得 k 时刻 UAV i 的最优控制输入 $u_i^*(k)$ 。上述方法也可以推广到 H 步预测的情况, 给定 UAV i 的 H 步控制输入 $U_i(k) = (u_i(k), u_i(k+1), \dots, u_i(k+H-1))^T$, 则 UAV i 状态由转移 $p_i(k)$ 到 $p_i(k+H)$, 覆盖分布地图由 $C_i(k)$ 更新为 $C_i(k+H)$, 将每一步的奖励函数 J 累加, 得到 H 步累积奖励函数 $J^{(H)}$ 为

$$J^{(H)} = \sum_{j=0}^{H-1} \alpha^j J(p_i(k+j), u_i(k+j)) \quad (17)$$

式中: α 为折扣因子, $0 < \alpha < 1$, 表示当前搜索奖励与长期搜索奖励之间的重要性程度。利用优化算法求解使累积奖励函数 $J^{(H)}$ 最大, 即可解得 UAV i 的 H 步最优控制输入 $U_i^*(k)$ 。

3.2 基于 DMPC 和 DE 的协同搜索算法

模型预测控制 (Model Predictive Control, MPC) 是一种利用控制系统模型和优化技术设计预测周期内系统最优控制输入的方法, 核心思想是滚动优化求解^[17]。集中式 MPC 方法依赖中央节点进行决策, 限制了系统规模的扩展和决策速

度, 在实际应用中有一定局限性^[14,18]。考虑到 UAV 子系统之间不存在耦合性, 即不同 UAV 的控制是相对独立的, 它们在系统的动态特性上并没有关联。为提高整个系统的抗毁性和决策速度, 其控制结构可以采用分布式模型预测控制 (Distributed Model Predictive Control, DMPC) 方式^[19], 如图 7 所示。

图 7 中 $\tilde{U}_i(k) = (\tilde{u}_i(k), \tilde{u}_i(k+1), \dots, \tilde{u}_i(k+H-1))^T$ 由 MPC i 求解得到, 为 UAV i 的 H 步预测控制输入, $\tilde{U}_{-i}(k)$ 为其他 UAV 的 H 步预测控制输入。在 DMPC 框架下, UAV 集群不依赖任何中央节点进行决策, 而是每架 UAV 采用局部的模型预测控制器分别进行决策。局部 MPC i 通过通信网络可以获得其他 UAV 的状态信息 $p_{-i}(k)$ 和局部 MPC 的解信息 $\tilde{U}_{-i}(k)$ 。因此, 第 i 个子系统的局部优化问题只与本地状态 $p_i(k)$ 和本机控制输入 $U_i(k)$ 相关, 优化问题的规模大大降低, 求解该优化问题就可以得到 UAV i 的最优预测控制输入 $U_i^*(k)$ ^[14]。对应于图 7 中 UAV i 的本地模型预测控制器 MPC i , 其决策流程如图 8 所示。

MPC i 决策流程分为 3 步。

步骤 1 预测

在预测阶段, 每架 UAV 基于本地覆盖地图和

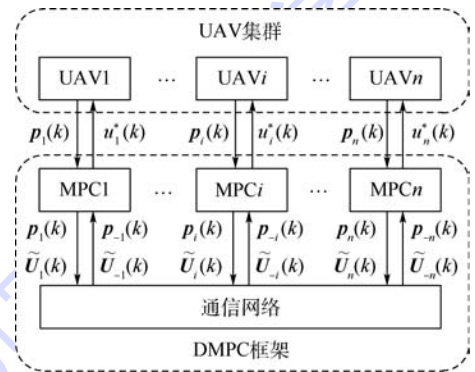


图 7 DMPC 框架示意图

Fig. 7 Schematic diagram of DMPC framework

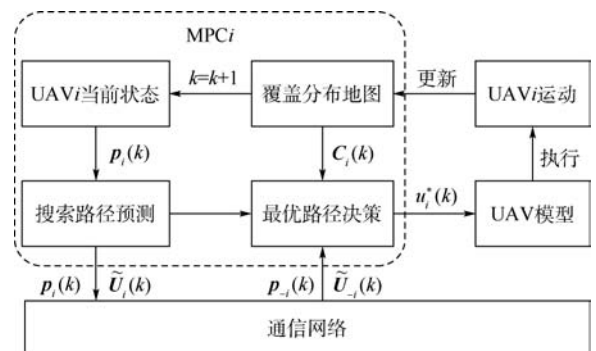


图 8 MPC i 决策流程

Fig. 8 Decision-making process of MPC i

自身状态进行优化求解,而不考虑其他 UAV 的运动,即 UAV 之间不进行协同。根据式(17),求解 UAV i 的 H 步预测控制输入的优化模型可以描述为

$$\tilde{U}_i(k) = \arg \max_{U_i(k)} J^{(H)} \quad (18)$$

利用差分进化(Differential Evolution, DE)算法对上述优化模型进行求解,即解得 UAV i 的 H 步预测控制输入 $\tilde{U}_i(k)$ 。

步骤2 通信

在通信网络理想的情况下,UAV i 将自身状态 $p_i(k)$ 和 H 步预测控制输入 $\tilde{U}_i(k)$ 广播发送,并通过通信网络接收其他 UAV 的状态 $p_{-i}(k)$ 和 H 步预测控制输入 $\tilde{U}_{-i}(k)$ 。

步骤3 决策

在预测阶段,每架 UAV 都不考虑其他 UAV 的运动,以“贪婪”的获得最大搜索奖励。在通信阶段,每架 UAV 通过通信网络获取到其他 UAV 的状态 $p_{-i}(k)$ 和 H 步预测控制输入 $\tilde{U}_{-i}(k)$,进行 UAV 之间的搜索“意图”交互,进而“协同”获得最大搜索奖励,使整个系统获得更高的搜索奖励。

通信完毕后,利用 $p_{-i}(k)$ 和 $\tilde{U}_{-i}(k)$ 将覆盖分布地图 $C_i(k)$ 更新为 $\tilde{C}_i(k)$, $\tilde{C}_i(k)$ 中包含了其他 UAV 的搜索“意图”,是 UAV i 与集群内其他 UAV 进行协同的依据。在决策阶段,只需基于 $\tilde{C}_i(k)$ 对 $U_i(k)$ 进行优化即可,根据式(17),求解 UAV i 的 H 步最优控制输入的优化模型可以描述为

$$U_i^*(k) = \arg \max_{U_i(k)} J^{(H)} \quad (19)$$

利用 DE 算法对上述优化模型进行优化求解,解得 UAV i 的 H 步最优控制输入 $U_i^*(k)$ 。将 $U_i^*(k)$ 的第 1 项 $u_i^*(k|k)$ 作为 k 时刻 UAV i 的控制输入,即 $u_i^*(k) = u_i^*(k|k)$ 。根据 UAV i 的控制输入 $u_i^*(k)$,将其本地覆盖分布地图 $C_i(k)$ 更新为 $C_i(k+1)$,并计算覆盖率。若覆盖率达到设定阈值,算法停止,否则 UAV 集群基于当前的覆盖分布地图和系统状态,重复上述优化过程,返回步骤 1。

搜索决策过程的算法伪代码如下:

1 初始化任务参数

2 for $k = 1$ to k_{\max}

3 预测: $\tilde{U}_i(k) = \arg \max_{U_i(k)} J^{(H)}$

4 发送 $p_i(k)$, $\tilde{U}_i(k)$

5 接收 $p_{-i}(k)$, $\tilde{U}_{-i}(k)$

6 将 $C_i(k)$ 更新为 $\tilde{C}_i(k)$

7 $U_i^*(k) = \arg \max_{U_i(k)} J^{(H)}$ (基于 $\tilde{C}_i(k)$)

8 最优控制输入: $u_i^*(k) = u_i^*(k|k)$

9 将 $C_i(k)$ 更新为 $C_i(k+1)$

10 if $O(k+1) \geq$ 设定阈值

11 break

12 end if

13 end for

控制输入 $U_i(k)$ 中包含 H 个未知变量,此优化问题是一个非线性优化问题。考虑到 DE 算法在求解优化问题,尤其是非线性优化问题中的优势^[20],采用 DE 算法进行子系统本地优化求解,算法细节不再赘述。

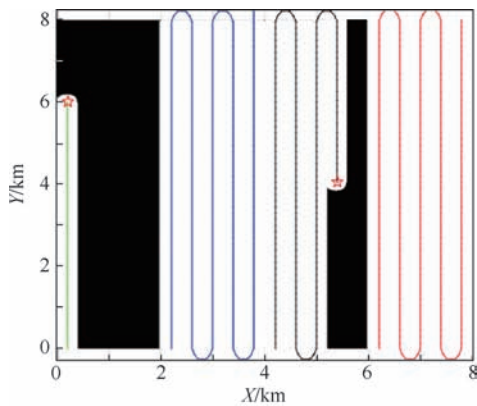
4 仿真分析

为验证本文方法的有效性,本节对其进行仿真验证。仿真环境为 I7-4960,主频 2.60 GHz,16 GB 内存,基于 MATLAB 2014a 为平台进行仿真实验。

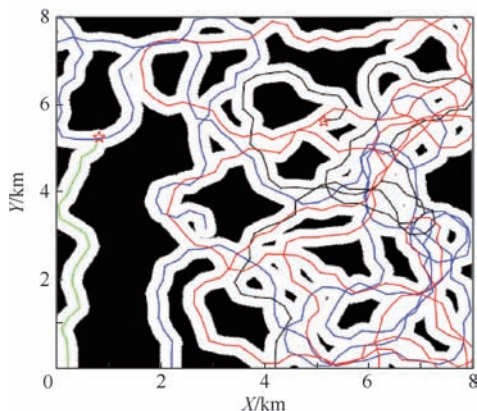
4.1 实验 1

设任务区域为 $8 \text{ km} \times 8 \text{ km}$ 的矩形区域,每个栅格大小为 $20 \text{ m} \times 20 \text{ m}$ 。执行搜索任务的 4 架 UAV 从不同位置进入搜索区域,其进入点坐标分别为 $(200,0) \text{ m}$, $(2200,0) \text{ m}$, $(4200,0) \text{ m}$, $(6200,0) \text{ m}$ 。UAV 之间的通信均为理想条件。UAV 平飞速度 $v_0 = 30 \text{ m/s}$,传感器探测半径 $R_s = 200 \text{ m}$,最大转弯角 $\varphi_{\max} = 60^\circ$,仿真步长 $\Delta t = 10 \text{ s}$,预测步长 $H = 3$, $\omega_1 = 0.9$, $\omega_2 = 0.05$, $\omega_3 = 0.05$, $\gamma = 200$ 。设定仿真环境条件:任务环境为无遮挡的平地区域,任务区域中存在分布未知的火力威胁,有一定几率造成 UAV 设备故障。在仿真中加入突发情况来模拟这一环境条件:运行至 20 步长时,假定 UAV1 设备故障,停止执行任务;运行至 100 步长时,假定 UAV3 设备故障,停止执行任务。为体现本文算法的优势,分别运用平行搜索、随机搜索和本文算法进行对比仿真。仿真结果如图 9 所示(图中黑色区域为 UAV 传感器未覆盖的区域)。

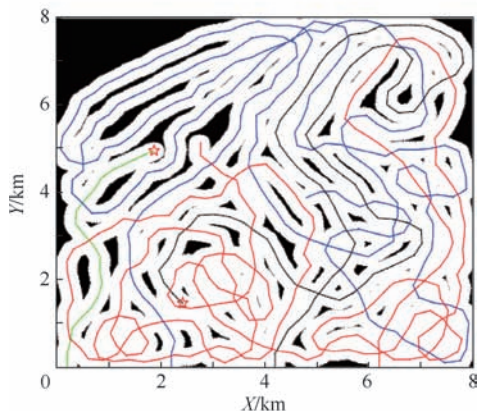
如图 9(a) 所示,在 2 架友机相继发生故障的情况下,UAV2 和 UAV4 只完成了各自预先分配的搜索任务,故障 UAV 未完成的搜索任务得不到继续执行。如图 9(b) 所示,随机搜索方法作为一种无引导机制的在线规划方法,在任务过程中多处存在 UAV 航迹交叉重叠的情况,搜索效率较低。如图 9(c) 所示,在搜索初始阶段,各 UAV 之间保持传感器探测范围尽可能不重叠,实现较高的覆盖率增长。在 2 架友机相继故障的情况下,



(a) 平行搜索



(b) 随机搜索



(c) 本文算法

— 已搜索区域 — UAV2 * UAV4
— UAV1 — UAV3 * 故障点

图 9 3 种搜索方法的仿真结果

Fig. 9 Simulation results of three search methods

UAV2 和 UAV4 通过实时协同,保持原有的搜索策略,充分发挥各自的搜索能力,继续完成搜索任务。总的来看,本文算法各 UAV 探测区域之间重叠部分较少,在出现突发情况时,能够继续完成搜索任务,体现了无人机集群在线协同的优势。

为消除随机因素的影响,在相同仿真条件下,针对本文算法和随机搜索方法使用蒙特卡罗方法进行 500 次仿真,得到平均覆盖率随时间变化曲线如图 10 所示。

如图 10 所示,平行搜索方法和本文算法覆盖

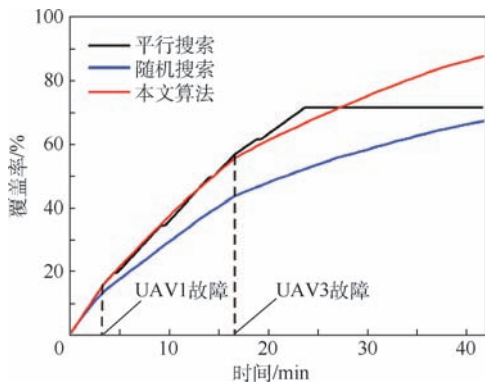


图 10 3 种搜索方法的覆盖率变化曲线

Fig. 10 Coverage rate changing curve of three search methods

率高于随机搜索方法。平行搜索方法的覆盖率变化曲线呈折线状:当某架 UAV 发生故障时,覆盖率曲线的斜率随之降低;当 UAV 到达边界时,执行转弯程序,覆盖率斜率为零。在搜索初期,本文算法与平行搜索方法覆盖率曲线斜率基本一致,体现了较高的搜索效率。当 UAV2 和 UAV4 完成各自的搜索任务后,平行搜索方法的覆盖率保持不变。相比之下,本文算法在 UAV1 和 UAV3 发生故障后,仍能保持覆盖率的稳定增长,最终任务结束时的覆盖率远高于平行搜索方法。

4.2 实验 2

针对本文算法在集群规模较大时的有效性进行验证,采用 10 架无人机组成的无人机集群进行仿真。设任务区域为 15 km × 15 km 的矩形区域,执行搜索任务的 10 架 UAV 的初始位置和航向由程序随机产生,UAV 故障随机指定:设定 UAV2、UAV3、UAV8、UAV9 和 UAV10 分别于仿真步数为 185、120、175、170 和 165 时发生故障,其他仿真条件同实验 1。用本文算法进行仿真,仿真步数 step 为 100 和 230 时的仿真结果如图 11 所示。

由图 11(a)可以看出,在搜索初期,10 架 UAV 保持传感器探测范围尽可能不重叠,以获得较高的覆盖率增长。由图 11(b)可以看出,在搜索后期,由于未搜索区域被分割成多个不规则的形状,UAV 航路出现了部分交叉,但本文算法还是能够引导 UAV 尽可能向未搜索区域移动。搜索过程中覆盖率变化曲线如图 12 所示。

如图 12 所示,在搜索前期,覆盖率增长速度较快,覆盖率随时间变化曲线的斜率保持稳定。到搜索后期,由于出现重复搜索的情况,覆盖率增长放缓,斜率逐渐降低。当搜索时间到达 38.33 min 时,覆盖率达到 90.13%,有效地对区域进行了覆盖。仿真结果表明,在集群规模达

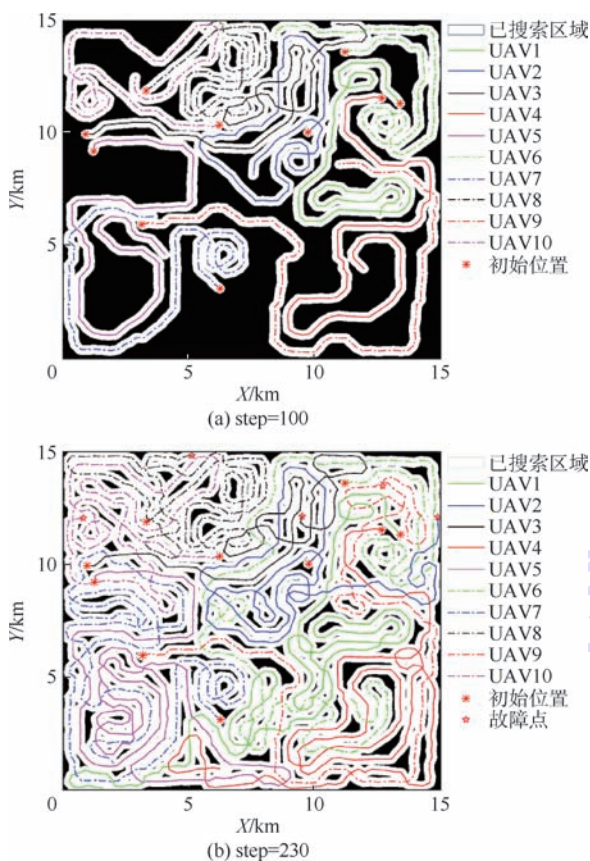


图 11 本文算法仿真结果

Fig. 11 Simulation result of proposed algorithm

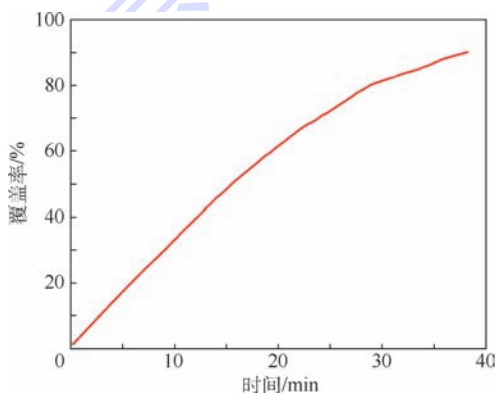


图 12 本文算法覆盖率变化曲线

Fig. 12 Coverage rate changing curve of proposed algorithm

到 10 架的情况下,本文算法能够较好地完成任务。

5 结 论

1) 本文算法在实验仿真条件下能实现较高的区域覆盖率,尤其是在出现突发情况时,覆盖率远高于平行搜索方法,体现了无人机集群在线协同的优势。

2) 以覆盖率作为实时搜索奖励的引导机制,有利于引导 UAV 向未搜索区域运动,并协同各

UAV 之间探测区域重叠部分尽可能少,以实现更高的覆盖率。

3) 利用 Hadamard 积可实现覆盖分布地图的快速更新,避免了遍历判断和逐一赋值,且操作简单、运算速度快,为地图信息实时更新提供了便捷。

4) 采用 DMPC 框架进行滚动优化求解,将长期的搜索奖励考虑在内,并且可以提高系统的抗毁性和决策速度。

本文算法在通信理想的条件下实现了对任务区域的有效覆盖搜索,但针对通信距离和角度约束的情况,尚需进行更加深入的后续研究。

参 考 文 献 (References)

[1] HU J W, XU J, XIE L H. Cooperative search and exploration in robotic networks [J]. Unmanned Systems, 2013, 1 (1): 121-142.

[2] HU J W, XU J, XIE L H. Vision-based multi-agent cooperative target search [C] // International Conference on Control, Automation, Robotics & Vision. Piscataway, NJ: IEEE Press, 2012: 895-900.

[3] 梁晓龙,孙强,尹忠海,等. 大规模无人系统集群智能控制方法综述 [J]. 计算机应用研究, 2015, 32(1) : 11-16.

LIANG X L, SUN Q, YIN Z H, et al. Review on large-scale unmanned system swarm intelligence control method [J]. Application Research of Computers, 2015, 32(1) : 11-16 (in Chinese).

[4] 梁晓龙,张佳强,祝捷,等. 基于 CPS 的空中交通系统架构及能力涌现方法 [J]. 空军工程大学学报 (自然科学版), 2016, 17(1) : 1-7.

LIANG X L, ZHANG J Q, ZHU J, et al. Air traffic control system architecture and ability emergence method based on cyber-physical system [J]. Journal of Air Force Engineering University (Natural Science Edition), 2016, 17(1) : 1-7 (in Chinese).

[5] 吴文超,黄长强,宋磊,等. 不确定环境下的多无人机协同搜索航路规划 [J]. 兵工学报, 2011, 32(11) : 1337-1342.

WU W C, HUANG C Q, SONG L, et al. Cooperative search and path planning of multi-unmanned air vehicles in uncertain environment [J]. Acta Armamentarii, 2011, 32(11) : 1337-1342 (in Chinese).

[6] ZHAO Z Y, LU G S. Receding horizon control for cooperative search of multi-UAVs based on differential evolution [J]. International Journal of Intelligent Computing and Cybernetics, 2012, 5(1) : 145-158.

[7] 于骊男,周锐,夏洁,等. 多无人机协同搜索区域分割与覆盖 [J]. 北京航空航天大学学报, 2015, 41(1) : 167-173.

YU S N, ZHOU R, XIA J, et al. Decomposition and coverage of multi-UAV cooperative search area [J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(1) : 167-173 (in Chinese).

[8] JONES P, VACHTSEVANOS G. Multi-unmanned aerial vehicle coverage planner for area surveillance missions [C] // AIAA Guidance, Navigation, and Control Conference and Exhibit. Re-

- ston: AIAA, 2007: 1-12.
- [9] 陈海,王新民,焦裕松,等.一种凸多边形区域的无人机覆盖航迹规划算法[J].航空学报,2010,31(9):1802-1808.
- CHEN H, WANG X M, JIAO Y S, et al. An algorithm of coverage flight path planning for UAVs in convex polygon areas[J]. Acta Aeronautica et Astronautica Sinica, 2010, 31(9): 1802-1808 (in Chinese).
- [10] 符小卫,魏广伟,高晓光.不确定环境下多无人机协同区域搜索算法[J].系统工程与电子技术,2016,38(4):821-827.
- FU X W, WEI G W, GAO X G. Cooperative area search algorithm for multi-UAVs in uncertainty environment[J]. System Engineering and Electronics, 2016, 38(4): 821-827 (in Chinese).
- [11] KHAN A, YANMAZ E, RINNER B. Information exchange and decision making in micro aerial vehicle networks for cooperative search[J]. IEEE Transactions on Control of Network Systems, 2015, 2(4): 335-347.
- [12] 杜继永,张凤鸣,毛红保,等.多UAV协同搜索的博弈论模型及快速求解方法[J].上海交通大学学报,2013,47(4):667-678.
- DU J Y, ZHANG F M, MAO H B, et al. Game theory based multi-UAV cooperative searching model and fast solution approach[J]. Journal of Shanghai Jiao Tong University, 2013, 47(4): 667-678 (in Chinese).
- [13] TRODDEN P, RICHARDS A. Multi-vehicle cooperative search using distributed model predictive control [C] // Guidance, Navigation, and Control Conference and Exhibit. Reston: AIAA, 2008: 1-11.
- [14] 彭辉,沈林成,朱华勇.基于分布式模型预测控制的多UAV协同区域搜索[J].航空学报,2010,31(3):593-601.
- PENG H, SHEN L C, ZHU H Y. Multiple UAV cooperative area search based on distributed model predictive control[J]. Acta Aeronautica et Astronautica Sinica, 2010, 31(3): 593-601 (in Chinese).
- [15] 郑宏捷.无人机区域侦察航路规划研究[D].长沙:国防科学技术大学,2011:5-8.
- ZHENG H J. Investigation on the UAV path planning problem of area reconnaissance [D]. Changsha: National University of Defense Technology, 2011: 5-8 (in Chinese).
- [16] 刘重,高晓光,符小卫.带信息素回访机制的多无人机分布式协同目标搜索[J].系统工程与电子技术,2017,39(9):1998-2011.
- LIU C, GAO X G, FU X W. Multi-UAVs distributed cooperative target search algorithm with controllable revisit mechanism based on digital pheromone[J]. System Engineering and Electronics, 2017, 39(9): 1998-2011 (in Chinese).
- [17] MAYNE D Q. Model predictive control: Recent developments and future promise [J]. Automatica, 2014, 50(12): 2967-2986.
- [18] KAYACAN E, PESCHEL J M, KAYACAN E. Centralized, decentralized and distributed nonlinear model predictive control of a tractor-trailer system: A comparative study [C] // 2016 American Control Conference. Piscataway, NJ: IEEE Press, 2016: 4403-4408.
- [19] MANSOURI S S, NIKOLAKOPOULOS G, GUSTAFSSON T. Distributed model predictive control for unmanned aerial vehicles [C] // 2015 Workshop on Research, Education and Development of Unmanned Aerial Systems. Piscataway, NJ: IEEE Press, 2015: 152-161.
- [20] DAS S, MULLICK S S, SUGANTHAN P N. Recent advances in differential evolution—An updated survey [J]. Swarm and Evolutionary Computation, 2016, 27: 1-29.

作者简介:

侯岳奇 男,硕士研究生。主要研究方向:航空集群智能决策。

梁晓龙 男,博士,教授,硕士生导师。主要研究方向:航空集群指挥与控制、智能系统、空管智能化。

何吕龙 男,博士研究生。主要研究方向:航空集群编队控制。

刘流 男,硕士研究生。主要研究方向:航空集群编队控制。

Cooperative area search algorithm for UAV swarm in unknown environment

HOU Yueqi^{1,2}, LIANG Xiaolong^{1,2,*}, HE Lyulong^{1,2}, LIU Liu^{1,2}

(1. National Key Laboratory of Air Traffic Collision Prevention, Air Force Engineering University, Xi'an 710051, China;

2. Shaanxi Province Lab. of Meta-synthesis for Electronic & Information System, Air Force Engineering University, Xi'an 710051, China)

Abstract: Aimed at the problem of cooperative search for UAV swarm in an unknown environment without prior information, a cooperative area search algorithm for UAV swarm with coverage rate as real-time search rewards is proposed. First, coverage distribution map (CDM) is established to describe the mission area, and the rapid update of CDM is realized by using Hadamard product. Then, the coverage rate is calculated based on CDM to describe the search results quantitatively. Considering UAV swarm as a control system, a predictive model of the system is established based on the distributed model predictive control theory, and the maximum increment of coverage rate in the predictive period is determined as a reward function. The optimal solution, as the optimal input of system, is obtained by differential evolution algorithm. Simulation results demonstrate that the proposed algorithm can complete the coverage and search of region effectively. In the event of emergencies, its area coverage rate is much higher than that of the parallel search method.

Keywords: unknown environment; UAV swarm; cooperative search; Hadamard product; coverage rate; distributed model predictive control

Received: 2018-04-25; **Accepted:** 2018-10-15; **Published online:** 2018-10-26 16:33

URL: kns.cnki.net/kcms/detail/11.2625.V.20181025.1141.003.html

Foundation items: National Natural Science Foundation of China (61472443,61703427); Natrual Science Basic Research Plan in Shaanxi Province of China(2017JQ6035)

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

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

DOI: 10.13700/j.bh.1001-5965.2018.0285

基于空间两点的视觉自主着陆导引算法设计



魏祥灰, 唐超颖, 王彪*

(南京航空航天大学 自动化学院, 南京 211106)

摘 要: 为提高无人机着陆效率, 从着陆速度向量场和导引律设计两方面研究改进。首先, 基于椭圆设计速度向量场, 实现飞行路程更短、机动性能要求更低的着陆轨迹。然后, 基于像素坐标系与机体坐标系的关系, 设计无人机的航迹方位角指令; 以椭圆切线方向为参考, 结合合作矢量特征, 设计航迹倾斜角指令; 利用图像信息, 设计速度大小指令。最后, 理论比较了传统轨迹与提出轨迹对方向机动性能的要求, 给出了轨迹参数与无人机方向机动性能的关系。利用 Simulink 搭建系统仿真平台, 计算满足要求的合作矢量特征。结果表明, 无人机以曲线轨迹准确软着陆到目标, 满足实际运用的需要。

关键词: 垂直起降无人机 (VTOL UAV); 自主着陆; 视觉导引律; 速度向量场; 方向机动性能

中图分类号: V249.3

文献标识码: A 文章编号: 1001-5965(2019)02-0357-09

垂直起降无人机 (UAV of the type Vertical Take Off and Landing, VTOL UAV) 被越来越广泛地运用到动目标跟踪、紧急着降以及危险救援等任务^[1-2]中, 这对其导引效率与精度要求越来越高^[3]。由于 GPS 信号的局限性, 基于视觉传感器的导引方式受到了越来越多的关注^[4-5]。本文以无人机着陆为背景, 采用视觉传感器, 设计和研究垂直起降无人机的高效率导引算法。

目前, 从实现导引的方式上, 国内外的无人机视觉着陆研究可分为基于位置的导引方式和基于速度的导引方式, 这 2 种方式都需要使用视觉解算当前位置, 容易引入计算误差^[6]。其中基于位置的导引方式^[7-8]是指利用视觉信息解算无人机当前位置, 计算其与预设基准轨迹的位置偏差, 采用纠偏处理, 使无人机按照既定的轨迹飞行。这种导引方式便于理解分析, 被当前大多数研究人员采用^[9-10]。基于速度的导引方式^[11-12]是指建立无人机当前位置与速度指令之间的关系, 不比较

导引过程中的位置误差, 以一定着陆精度完成定点着陆。这种方式具备计算复杂度低、实时性好等特点, 能够快速高效地完成导引任务。Min 等^[13]采用基于速度的导引方式, 设计以时间为基础的纵向加速度指令生成算法, 实现竖直方向上的着陆。

另外, 为实现无人机精准导引, 目前使用的导引法主要有比例导引法、速度向量场法和线性反馈法等^[14]。比例导引律是指飞行器在攻击或者追踪目标的导引过程中, 速度矢量的转动角速度与目标瞄准线 (视线) 的转动角速度成比例的一种导引律^[15], 在导引的末端, 运动轨迹近似为直线, 不能满足垂直起降无人机着陆瞬间速度方向竖直向下的要求, 在无人机着陆领域, 比例导引法使用很少。速度向量场法是一种常用的无人机导引方法。针对垂直起降无人机, 国内外学者几乎都采用“直角型”的转弯方式^[16-17]。由此设计的速度向量场, 需要无人机在“直角点”悬停一段时间再转弯飞行, 这严重影响了无人机着陆效率。

收稿日期: 2018-05-21; 录用日期: 2018-09-19; 网络出版时间: 2018-10-10 16:27

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180930.0938.001.html

基金项目: 南京航空航天大学研究生创新基地 (实验室) 开放基金-中央高校基本科研业务费专项资金 (kfj20170302)

* 通信作者. E-mail: wangbiao@nuaa.edu.cn

引用格式: 魏祥灰, 唐超颖, 王彪. 基于空间两点的视觉自主着陆导引算法设计[J]. 北京航空航天大学学报, 2019, 45(2): 357-365. WEI X H, TANG C Y, WANG B. Visual guidance algorithm design for autonomous landing based on two points in space[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 357-365 (in Chinese).

假若不悬停、直接直角转弯,对无人机的方向机动性能要求过高,难以工程实现。张炳亮^[18]设计了基于圆的速度向量场,但该方案只适用于飞行高度与水平着陆距离相当的情况,不具有普适性。

针对以上无人机着陆效率较低的问题,一方面,从减少着陆路程和降低飞行机动性能要求的角度,设计速度向量场,提高无人机的转弯效率。另一方面,采用基于速度的着陆导引方式,选取合适的视觉合作特征,建立图像信息与速度指令之间的直接关系。避免位置解算及其引入的计算耗时和误差,提高着陆导引效率。

本文首先结合无人机着陆的速度方向以及转弯能力等要求,设计速度向量场,同时讨论相关参数与无人机性能的约束关系;然后,研究图像特征与航迹倾斜角、航迹方位角和速度大小的关系,推导出合理的着陆导引规律;最后,利用 Simulink 仿真平台验证本文设计导引律的正确性和准确性。

1 速度向量场的设计与分析

本节以 $1/4$ 椭圆弧为单位,设计着陆的速度向量场,并分析此时无人机的机动性能要求。以弧上各点的切线方向确定航迹倾斜角的理想值(见图1),从而满足着陆瞬间机体速度方向竖直向下的需求。定义航迹倾斜角 γ 为机体速度 V_b 与水平面的夹角。建立纵平面的坐标系,其中 C_1 点为目标位置,椭圆表达式为

$$\frac{y^2}{a^2} + \frac{x^2}{b^2} = 1 \quad (1)$$

式中: $x \geq 0, y \geq 0, a > 0, b > 0$, 椭圆弧与 y 轴的交点为 $(0, a)$, 与 x 轴交点为 $C_1: (b, 0)$ 。设椭圆曲线上任意一点 M 的坐标为 (x_M, y_M) , 则该点航迹倾斜角 γ_M 为

$$\gamma_M = \arctan[a^2 b^{-2} x_M (a^2 - a^2 b^{-2} x_M^2)^{-1/2}] \quad (2)$$

以 $x=0$ 以及 $x=2b$ 轴上任意点,与 C_1 点作椭圆弧,建立目标着陆点周围的速度向量场(见图2)。

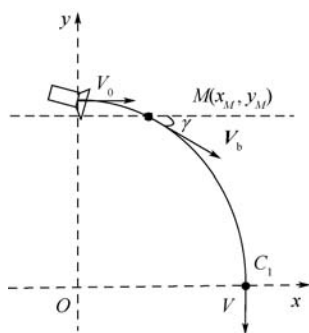


图1 椭圆弧示意图

Fig. 1 Schematic diagram of elliptic arc

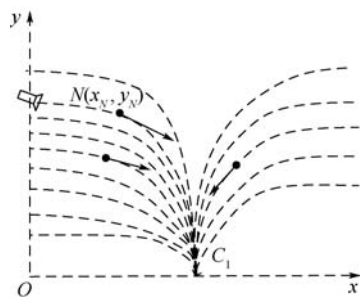


图2 速度向量场示意图

Fig. 2 Schematic diagram of velocity vector field

空间中任意一点 $N(x_N, y_N)$ 的航迹倾斜角 γ_N 表达式为

$$\begin{cases} \gamma_N = \arctan[a_N b^{-1} x_N (b^2 - x_N^2)^{-1/2}] \\ a_N = \frac{y_N}{\sqrt{1 - \frac{x_N^2}{b^2}}} \end{cases} \quad (3)$$

式中: a_N 为点 N 所在椭圆弧与 y 轴交点的纵坐标。

以轨迹中各点的曲率半径为衡量指标,分析椭圆弧对机动性能的要求,半径越小表示对无人机转弯机动能力要求越高。假设着陆的光滑曲线 L 在着陆平面上的表达式为 $y_L = f(x_L)$, 则点 (x_L, y_L) 的曲率半径为

$$R_{(x_L, y_L)} = \left| \frac{(1 + f'(x_L)^2)^{1.5}}{f''(x_L)} \right| \quad (4)$$

若点 (x_L, y_L) 不可求导, 则

$$\lim(R_{(x_L, y_L)}) = 0 \quad (5)$$

将式(1)代入式(4)计算轨迹的曲率半径,得到最大曲率半径 R_{\max} 与最小曲率半径 R_{\min} 分别为

$$R_{\max} = (\max\{a, b\})^2 / \min\{a, b\} \quad (6)$$

$$R_{\min} = (\min\{a, b\})^2 / \max\{a, b\} \quad (7)$$

结合式(7),若无人机固有的转弯机动性能对应的转弯半径为 R_{plane} , 则满足约束条件 $R_{\text{plane}} < R_{\min}$ 时,无人机具有完成既定的轨迹的能力。

2 导引律的设计

本节以一个装载两轴云台的具有速度闭环控制的垂直起降无人机为对象。选取竖直方向上的高度矢量 $\vec{C}_1 \vec{C}_2$ 作为合作特征^[18], 建立其图像信息 $\vec{c}_1 \vec{c}_2$ 与无人机速度指令之间的关系,实现无人机准确着陆,系统结构如图3所示。

2.1 图像特征与速度方向的关系

无人机在三维空间中的速度方向由航迹倾斜角 γ 和航迹方位角 χ 两个角度确定,其中航迹方位角表示速度在水平面的投影与地面固定轴的夹角。

机体坐标系 $X_b Y_b Z_b$ 、相机坐标系 $X_c Y_c Z_c$ 以及像素坐标系 uv 之间的关系如图 4 所示,其中特征点 C_1 的像点 c_1 在相机光轴上,这是由云台的跟瞄功能保证的,在无人机的飞行过程中,相机能完成快速对准,使得光轴 $O_c Z_c$ 一直保持指向点 C_1 。机体坐标系一点 $P(x_b, y_b, z_b)$ 在图像中的像点坐标为 (u, v) ,由相机成像原理可知两坐标关系为

$$\begin{bmatrix} x_b \\ y_b \\ z_b \end{bmatrix} = \mathbf{R}_{bc} \begin{bmatrix} D_u & 0 & -u_0 \\ 0 & D_v & -v_0 \\ 0 & 0 & z_c \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} - \mathbf{T}_{bc} \quad (8)$$

式中: $D_u = -z_c d_u / f$, $D_v = -z_c d_v / f$, f 为相机焦距, d_u 与 d_v 分别为像素坐标系 u 轴、 v 轴方向上单位像素对应的物理长度,一般地, $d_u = d_v$; z_c 为点 P 在相机坐标系下的 Z 轴坐标; u_0 与 v_0 为相机主点位置; \mathbf{R}_{bc} 与 \mathbf{T}_{bc} 分别为机体坐标系与相机坐标系之间的旋转矩阵和平移矩阵。

设旋转角 ψ_{bc} 、 θ_{bc} 、 ϕ_{bc} 分别为相机坐标系绕转动轴 Z'_b 、 Y'_b 、 X_b 的旋转角度,由机体坐标系与相机坐标系定义正方向不同可得 $\psi_{bc} = 90^\circ$,由云台旋转的云台倾角 θ_y 与云台滚转角 ϕ_y ,可以给出 $\theta_{bc} = \theta_y$, $\phi_{bc} = \phi_y$ 。则 3 个角度对应的旋转矩阵表达式为

$$\mathbf{R}_{z'_b} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \phi_y & -\sin \phi_y \\ 0 & \sin \phi_y & \cos \phi_y \end{bmatrix} \quad (9)$$

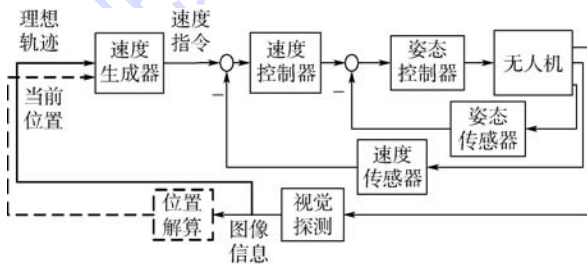


图 3 本文导引律的结构框图

Fig. 3 Structure diagram of proposed guidance law

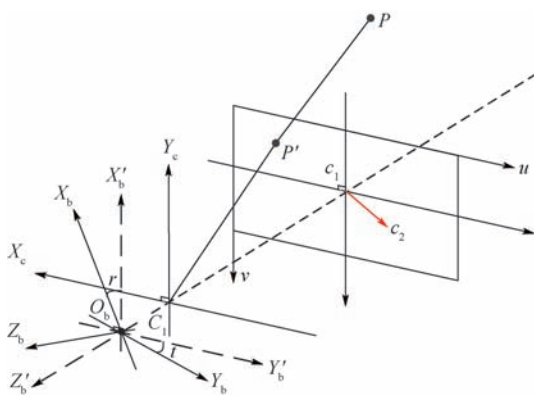


图 4 机体坐标系与像素坐标系的关系示意图

Fig. 4 Schematic diagram of relationship between body coordinate system and image coordinate system

$$\mathbf{R}_{y'_b} = \begin{bmatrix} \cos \theta_y & 0 & \sin \theta_y \\ 0 & 1 & 0 \\ -\sin \theta_y & 0 & \cos \theta_y \end{bmatrix} \quad (10)$$

$$\mathbf{R}_{x_b} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (11)$$

则 \mathbf{R}_{bc} 表达式为

$$\mathbf{R}_{bc} = \mathbf{R}_{z'_b} \mathbf{R}_{y'_b} \mathbf{R}_{x_b} \quad (12)$$

由式(8)可得图像中某一向量 $(\Delta u, \Delta v)$ 对应机体坐标系下的 X_b - Y_b 二维单位矢量 $(\Delta x_b, \Delta y_b)$ 的关系如下:

$$k \begin{bmatrix} \Delta x_b \\ \Delta y_b \end{bmatrix} = \mathbf{R} \begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} \quad (13)$$

式中: k 为尺度因子; \mathbf{R} 的表达式为

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \mathbf{R}_{bc} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad (14)$$

设合作特征 C_1 与 C_2 的像点 c_1 、 c_2 的坐标为 (u_1, v_1) 、 (u_2, v_2) , 则矢量 $\overrightarrow{c_1 c_2}$ 表达式为 $\overrightarrow{c_1 c_2} = (u_2 - u_1, v_2 - v_1)$ (15)

代入式(13), 可得航迹方位角指令 χ_{cmd} 表达式为

$$\chi_{cmd} = \arcsin\left(\frac{\Delta x_b}{\sqrt{(\Delta x_b)^2 + (\Delta y_b)^2}}\right) \quad (16)$$

由式(16), 得到图像信息与航迹方位角指令的关系, 接下来探讨图像信息与航迹倾斜角的关系, 用图像信息表示公式(3)。图 5 展示了无人机与合作特征的空间关系, 其中 α 为相机光轴与世界坐标系 $-Z_w$ 轴的夹角, β 为视线 MC_1 与 MC_2 构成的夹角, α 与 β 反映了机体与目标之间的相对位置关系, 并且满足 $\alpha \geq \beta$ 。

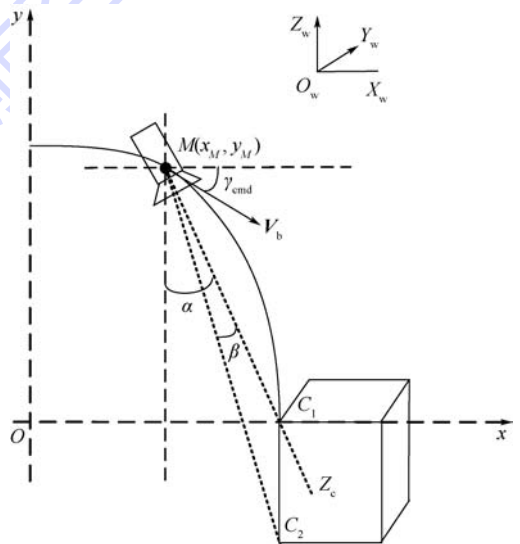


图 5 无人机与合作特征的空间关系

Fig. 5 Spatial relationship between UAV and cooperative characteristics

其中由 ΔMC_1C_2 的正弦定理可得

$$\frac{|\overrightarrow{C_1C_2}|}{\sin \beta} = \frac{|\overrightarrow{C_1M}|}{\sin(\alpha - \beta)} \quad (17)$$

点 $M(x_M, y_M)$ 坐标表达式为

$$x_M = b - |\overrightarrow{C_1M}| \times \sin \alpha \quad (18)$$

$$y_M = |\overrightarrow{C_1M}| \times \cos \alpha \quad (19)$$

将式(17)代入式(18)与式(19),得到 α, β 与点 M 坐标的关系为

$$x_M = b - \frac{|\overrightarrow{C_1C_2}| \sin(\alpha - \beta)}{\sin \beta} \sin \alpha \quad (20)$$

$$y_M = \frac{|\overrightarrow{C_1C_2}| \sin(\alpha - \beta)}{\sin \beta} \cos \alpha \quad (21)$$

将式(20)、式(21)代入式(3),得到点 M 处航迹倾斜角指令 γ_{cmd} 的表达式为

$$\begin{cases} x_M = b - \frac{|\overrightarrow{C_1C_2}| \sin(\alpha - \beta)}{\sin \beta} \sin \alpha \\ y_M = \frac{|\overrightarrow{C_1C_2}| \sin(\alpha - \beta)}{\sin \beta} \cos \alpha \\ \gamma_{cmd} = \arctan(a_M b^{-1} (b^2 - x_M^2)^{-1/2} x_M) \\ a_M = \frac{y_M}{\sqrt{1 - \frac{x_M^2}{b^2}}} \end{cases} \quad (22)$$

式中: b 为无人机在初始位置时与目标点的水平距离:

$$b = \frac{|\overrightarrow{C_1C_2}| \sin(\alpha_0 - \beta_0)}{\sin \beta_0} \sin \alpha_0 \quad (23)$$

式中:下标 0 表示相应变量的初始值; α 的大小由云台系统给出的 2 个转角(倾角 θ_Y 和滚转角 ϕ_Y) 求出:

$$\alpha = \arctan(\sqrt{\tan^2 \theta_Y + \tan^2 \phi_Y}) \quad (24)$$

目标视线夹角 β 与图像信息之间的关系如图 6 所示,其中 $O_c c_1 \perp c_1 c_2$, 得出 β 的表达式为

$$\beta = \arctan\left(\frac{\sqrt{(u_2 - u_1)^2 d_u^2 + (v_2 - v_1)^2 d_v^2}}{f}\right) \quad (25)$$

联列式(22)~式(25)可得点 M 处,航迹倾斜角指令 γ_{cmd} 与图像信息之间的关系表达式为

$$\begin{cases} \beta = \arctan\left(\frac{\sqrt{(u_2 - u_1)^2 d_u^2 + (v_2 - v_1)^2 d_v^2}}{f}\right) \\ x_M = b - \frac{|\overrightarrow{C_1C_2}| \sin(\alpha - \beta)}{\sin \beta} \sin \alpha \\ y_M = \frac{|\overrightarrow{C_1C_2}| \sin(\alpha - \beta)}{\sin \beta} \cos \alpha \\ \gamma_{cmd} = \arctan(a_M b^{-1} (b^2 - x_M^2)^{-1/2} x_M) \\ a_M = \frac{y_M}{\sqrt{1 - \frac{x_M^2}{b^2}}} \\ b = \frac{|\overrightarrow{C_1C_2}| \sin(\alpha_0 - \beta_0)}{\sin \beta_0} \sin \alpha_0 \\ \alpha = \arctan(\sqrt{\tan^2 \theta_Y + \tan^2 \phi_Y}) \end{cases}$$

在实际应用中,角度测量具有精度限制,同时 α, β 的测量精度取决于特征矢量的高度 C 。接下来讨论角度精度与特征高度之间的关系。假设无人机从距离着陆点竖直高度为 h , 水平距离为 l 的位置着陆,则得到本文设计的椭圆着陆轨迹的参数表达式为

$$\begin{cases} x = l \cos \theta \\ y = h \sin \theta \end{cases} \quad (26)$$

式中: $0 < \theta < \pi/2$ 。假设无人机着陆末端允许的竖直高度偏差为 h_T , 即当高度小于等于 h_T 时,无人机可以将速度降为 0。则由 $y = h_T$ 得到着陆末端允许的最大位置偏差 $(\Delta x, \Delta y)$ 为

$$\begin{cases} \Delta x = l - l \cos(\arcsin(h_T/h)) \\ \Delta y = h_T \end{cases} \quad (27)$$

如图 5 所示,设无人机某一点的位置为 (x, y) , 特征点 C_1, C_2 的空间位置分别为 $(b, 0), (b, -C)$ 。由 ΔPC_1C_2 中的余弦定理,可得 α, β 与 (x, y) 的关系为

$$\begin{cases} \beta = \arccos\left(\frac{[(x-b)^2 + y^2 + (x-b)^2 + (y+C)^2 - C^2]}{2\{[(x-b)^2 + y^2][(x-b)^2 + (y+C)^2]\}^{1/2}}\right) \\ \alpha = \arctan\left(\frac{b-x}{y}\right) \end{cases} \quad (28)$$

由式(28)可知,在确定无人机和目标点的位置时,合作矢量的高度 C 与 β 成正相关。将

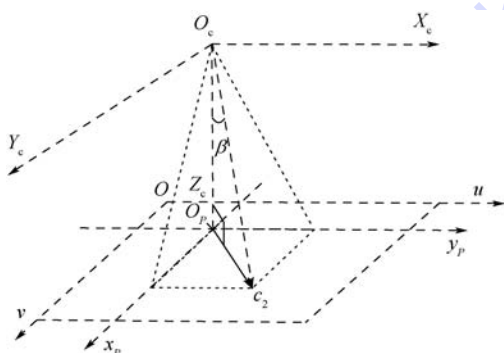


图 6 β 与图像信息之间的关系示意图

Fig. 6 Schematic diagram of relationship between β and image information

式(1)代入式(28)得到 β 与 y 之间的关系:

$$\beta = \arccos(g(y)) \quad (29)$$

经计算 $g(y)$ 在 $x \geq 0, y \geq 0$ 的范围内为关于 y 的单调递减函数,故 β 的大小与 y 为正相关。假设合作特征高度矢量的高度值为 C ,利用式(27)计算得到的偏差 $(\Delta x, \Delta y)$ 代入式(29),可得

$$\Delta\beta = \arccos \frac{[(\Delta x)^2 + (\Delta y)^2] + C\Delta y}{\sqrt{((\Delta x)^2 + (\Delta y)^2)[(\Delta x)^2 + (\Delta y + C)^2]}} \quad (30)$$

若 $(\Delta x, \Delta y)$ 表示无人机能够允许的位置偏差,则相应的 $\Delta\beta$ 表示着陆系统允许的 β 值的测量偏差(即目标视线夹角的精度,用 d_β 表示),由此可得精度 d_β 与 C 的关系:

$$d_\beta = \arccos \frac{[(\Delta x)^2 + (\Delta y)^2] + C\Delta y}{\sqrt{[(\Delta x)^2 + (\Delta y)^2][(\Delta x)^2 + (\Delta y + C)^2]}} \quad (31)$$

同理可得 α 的精度 d_α :

$$d_\alpha = \arctan\left(\frac{\Delta x}{\Delta y}\right) \quad (32)$$

由此可知,在针对不同对象和场景应用时,可以根据应用背景的参数 h, l , 以及无人机和相机传感器的性能参数 h_T, d_β 求出对应的合作矢量的高度 C 。

2.2 图像特征与速度大小的关系

由于 $\alpha \geq \beta$,故本文建立 α 与速度大小指令 $|\mathbf{V}_{cmd}|$ 的关系。将式(1)代入式(28), x_p 为无人机与目标的水平距离,则 α 与 x_p 之间的关系表示为

$$\alpha = \arctan(f(x_p)) \quad (33)$$

经计算 $f(x_p)$ 在 $x_p \geq 0, y_p \geq 0$ 的范围内为关于 x_p 的单调递减函数,故 α 与 x_p 呈负相关。由于 $|\mathbf{V}_{cmd}|$ 与 x_p 也呈负相关。故本文以参数 α_1, α_2 构成速度大小指令 $|\mathbf{V}_{cmd}|$ 的分段函数如式(34)所示,其中 V_0 为无人机初始速度的大小,当 $\alpha_2 > \alpha > \alpha_1$ 时, $|\mathbf{V}_{cmd}|$ 的变化与 α 的变化成比例关系。

$$|\mathbf{V}_{cmd}| = \begin{cases} V_0 & \alpha > \alpha_2 \\ V_0 \frac{\alpha - \alpha_1}{\alpha_2 - \alpha_1} & \alpha_2 > \alpha > \alpha_1 \\ 0 & \alpha < \alpha_1 \end{cases} \quad (34)$$

综上,无人机的机体坐标系下速度指令:

$$\begin{cases} V_{cmd_x} = |\mathbf{V}_{cmd}| \cos \gamma_{cmd} \cos \chi_{cmd} \\ V_{cmd_y} = |\mathbf{V}_{cmd}| \cos \gamma_{cmd} \sin \chi_{cmd} \\ V_{cmd_z} = |\mathbf{V}_{cmd}| \sin \gamma_{cmd} \end{cases} \quad (35)$$

3 仿真验证与分析

3.1 与传统方法比较分析

传统着陆轨迹与本文着陆轨迹如图 7 所示。其中传统着陆方式下的轨迹表达式为

$$\begin{cases} y = a & 0 < x < b \\ x = b & 0 < y \leq a \end{cases} \quad (36)$$

对比本文提出的轨迹表达式(式(1)),得到着陆距离 L_1, L_2 分别为

$$\begin{cases} L_1 = a + b \\ L_2 = a + \frac{(\pi - 2)b}{2} \end{cases} \quad (37)$$

由此可见,在本文提出的着陆距离更短,并且水平方向距离目标越远(b 值越大),本文算法在着陆距离上的优势越明显。

另外,分析比较 2 种轨迹的最小曲率半径,以此评价 2 种着陆方式对无人机的转弯机动性能的需求。本文设计轨迹的最小曲率半径由式(7)给出,当参数 a 或 b 固定,满足 $a = b$ 时,理想轨迹的最小曲率半径 R_{min} 最大。由式(4)、式(5)的曲率半径计算原理,得到传统轨迹的曲率半径 R_T 为

$$R_T = \begin{cases} +\infty & 0 < x < b, y = a \\ \arg(\lim(R_T) = +0) & x = b, y = a \\ +\infty & x = b, 0 < y < a \end{cases} \quad (38)$$

由此可见,与传统着陆方式相比,本文的着陆轨迹对机动性能要求更低。传统方式要求转弯机动性能无限高,而这在实际应用中无法达到,所以工程中通过在转弯处悬停的方式,通过消耗时间,降低对机动性能的高要求。

综上,与传统着陆方式相比,本文的导引方法实现了着陆距离更短,机动性要求更低,效率更高。

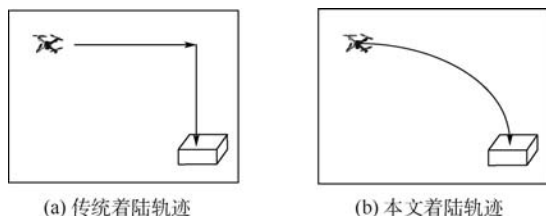


图 7 着陆轨迹的比较

Fig. 7 Comparison of landing trajectory

3.2 系统仿真验证

基于 Simulink 搭建了视觉着陆导引系统仿真平台,将整个系统分为 X450 四旋翼模型、飞行控

制器、着陆导引算法、云台系统、视觉探测算法、小孔成像、目标七大模块,各个模块之间通过各自的输入输出信号建立关联,系统仿真结构框图如图8所示。

图8中,X450四旋翼模型为笔者课题组已有的X450型四旋翼非线性模型,飞行控制器包括机体的速度、姿态、角速率的闭环控制,云台系统由云台模型与云台跟瞄控制器组成,小孔成像模块将世界坐标系中的目标投影到像素坐标系中,视觉探测算法获取目标信息后,将视觉解算信息提供给云台系统。为便于描述,本文假设相机、云台质心与机体质心重合。

参考无人机实际飞行的情况,取 $h_T = 0.1\text{ m}$ 、 $d_p = 0.4^\circ$,代入式(31)计算出合作目标的高度 $C = 1\text{ m}$, $d_\alpha = 0.5^\circ$ 。设世界坐标系下的无人机的初始位置为 $P(0,0,5.0)\text{ m}$,合作特征矢量 \vec{C}_1, \vec{C}_2 的两点: $C_1(1.5,2.0,1.0)\text{ m}$, $C_2(1.5,2.0,0)\text{ m}$,其中, C_1 即为着陆目标点,初始速度 $V_0 = 0.3\text{ m/s}$,角度阈值 $\alpha_1 = d_\alpha = 0.5^\circ$, $\alpha_2 = \alpha_{\max}/3 = 11^\circ$,其中 α_{\max} 的表达式为

$$\alpha_{\max} = \arctan\left(\frac{|\vec{PC}_1|_l}{|\vec{PC}_1|_h}\right) \quad (39)$$

式中: $|\vec{PC}_1|_l$ 与 $|\vec{PC}_1|_h$ 分别为点 P 与 C_1 在水平和垂直方向上的距离。在三维空间中对本文提出的视觉着陆导引算法进行仿真验证,仿真的前5s时间为设备的初始化,包括云台转动对准着陆目标点等。

图9(a)为着陆过程中某时刻的状态,其中曲线表示已飞过的轨迹,左端为无人机的初始位置,右端为当前位置,其中的蓝色图形表示相机的空间姿态,下面的蓝色图形为合作目标,该图可动态地演示视觉导引无人机着陆的全过程。图9(b)为对应的着陆垂直平面内的二维着陆轨迹,其中平行于时间轴的红线表示着陆目标点的高度1m。

无人机的导引指令在着陆过程中的变化如图10所示,航迹倾斜角指令 γ_{cmd} 由初始值逐渐增

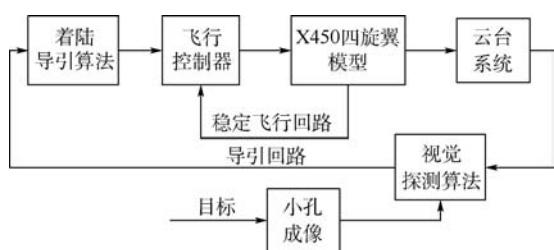
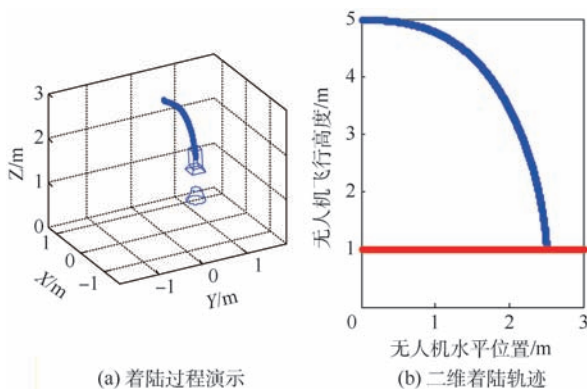


图8 系统仿真平台结构框图

Fig. 8 Structure diagram of system simulation platform

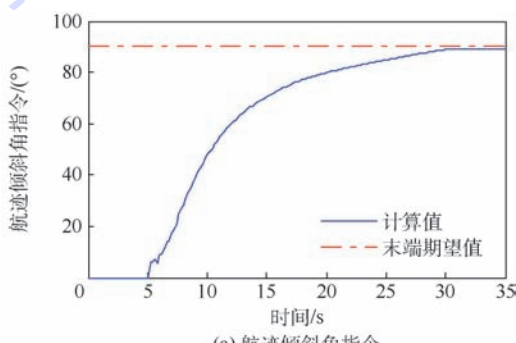


(a) 着陆过程演示

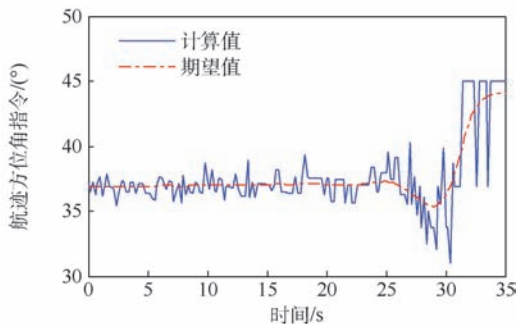
(b) 二维着陆轨迹

图9 无人机着陆过程仿真示意图

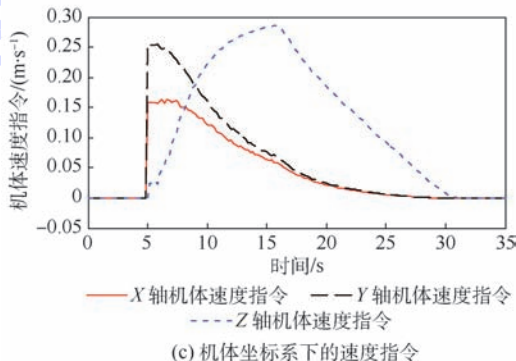
Fig. 9 Schematic diagram of UAV landing process simulation



(a) 航迹倾斜角指令



(b) 航迹方位角指令



(c) 机体坐标系下的速度指令

图10 导引指令随时间的变化

Fig. 10 Variation of guidance command with time

加到末端期望值(图10(a));航迹方位角指令 χ_{cmd} 与期望值曲线走势一致,并在期望值上下波动,这主要受成像误差的影响(图10(b));随着无人机越来越接近目标,速度越来越小(图10(c)),

到达目标位置时,速度方向指令接近竖直向下,大小为零。

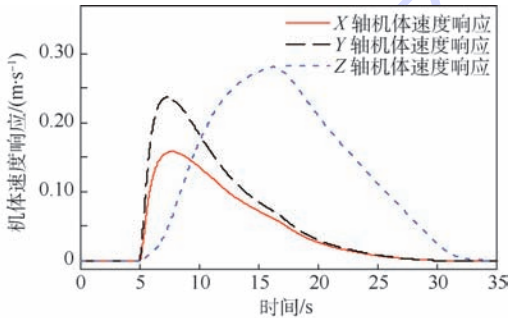
无人机速度响应、位置响应以及航迹倾斜角响应值在着陆过程中的变化如图 11 所示,可见,在本文算法约束下,无人机在接近目标点时,速度缓慢降低,到达目标位置时,速度大小为零,方向竖直向下,实现了垂直软着陆。

航迹倾斜角指令是由图像信息解算的目标视线夹角 β 以及云台偏角 α 计算所得,接下来验证角度的测量准确性。图 12 和图 13 分别表示着陆过程中,两夹角的真实值与测量值的关系。其中真实值与测量值的绝对误差值 $\Delta\beta$ 与 $\Delta\alpha$ 的表达式为

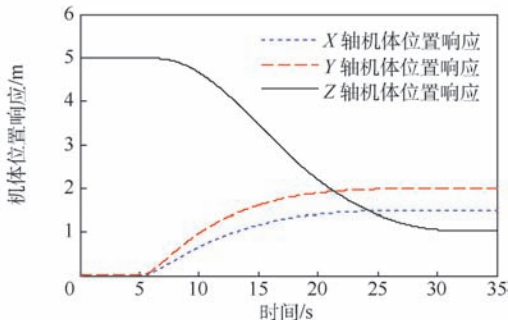
$$\Delta\beta = |\beta - \beta_r| \quad (40)$$

$$\Delta\alpha = |\alpha - \alpha_r| \quad (41)$$

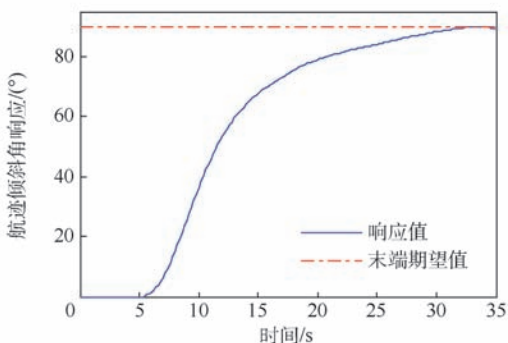
由式(40)得到误差 $\Delta\beta$ 一直保持在 0.14° 内, α 的测量值在着陆始端误差不超过 0.9° , 并且随



(a) 机体速度随时间的响应



(b) 机体位置随时间的响应



(c) 航迹倾斜角响应

图 11 响应随时间的变化

Fig. 11 Variation of response with time

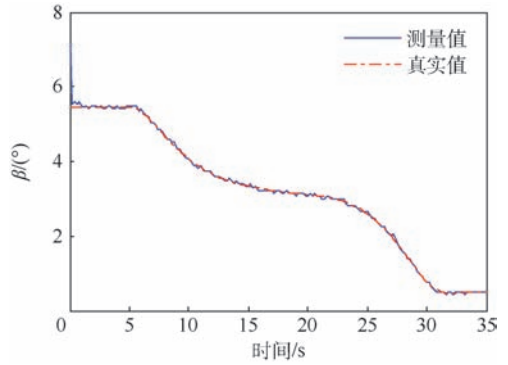


图 12 β 角分析

Fig. 12 β angle analysis

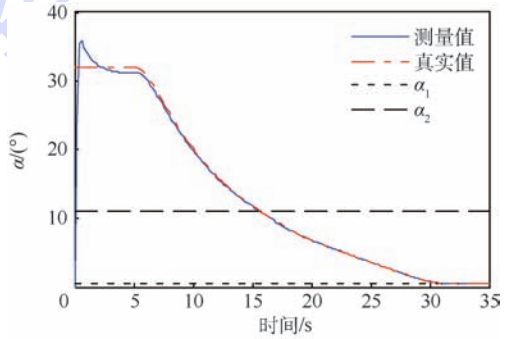


图 13 α 角分析

Fig. 13 α angle analysis

着无人机距离目标越近,很快误差保持在 0.1° 以内。由此可知本文算法能准确地解算出角度值,并且距离目标越近则准确度越高,满足实际应用的要求。

4 结 论

在无人机着陆中的导引研究中,基本上使用基于位置的方式,为提高导引效率,本文设计了基于空间两点的视觉自主着陆导引算法,为今后实现自主、智能的无人机提供理论研究基础。本文的创新性主要概括为以下 2 个方面:

1) 突破传统直角型的转弯方式,设计曲线型速度向量场,导引垂直起降无人机曲线着陆,保证末端约束的前提下,提高无人机的着陆效率。

2) 避免传统导引算法中的位置解算,直接建立图像信息与速度指令之间的关系,避免了位置解算引入的计算量和误差,提高计算效率。

由实验部分可以验证本文算法的有效性和准确性,利用本文给出的合作特征高度与角度测量精度之间的关系,确定出合理的合作目标高度,以目标 $c_1 c_2$ 可探测时的相对位置为初始距离,进行系统仿真。结果表明,无人机以曲线轨迹准确软着陆到目标位置,满足实际运用的需要。

参考文献 (References)

- [1] LIN S, GARRATT M A, LAMBERT A J. Monocular vision-based real-time target recognition and tracking for autonomously landing an UAV in a cluttered shipboard environment[J]. *Autonomous Robots*, 2017, 41(4): 881-901.
- [2] ARAAR O, AOUF N, VITANOV I. Vision based autonomous landing of multirotor UAV on moving platform[J]. *Journal of Intelligent & Robotic Systems*, 2017, 85(2): 369-384.
- [3] SHIRZADEH M, ASL H J, AMIRKHANI A, et al. Vision-based control of a quadrotor utilizing artificial neural networks for tracking of moving targets[J]. *Engineering Applications of Artificial Intelligence*, 2017, 58: 34-48.
- [4] PRAKASH R O, SARAVANAN C. Autonomous robust helipad detection algorithm using computer vision[C] // International Conference on Electrical, Electronics, and Optimization Techniques. Piscataway, NJ: IEEE Press, 2016: 2599-2604.
- [5] TANG D, CHEN Y, KOU K. Navigation method based on the solution to PnP problem for autonomous landing of UAV[C] // Proceedings of 2014 IEEE Chinese Guidance, Navigation and Control Conference. Piscataway, NJ: IEEE Press, 2014, 2315-2320.
- [6] 陈茂武. 基于视觉的无人机自动着陆导航技术与系统研究[D]. 南京: 南京航空航天大学, 2017.
CHEN M W. Research on vision based automatic landing guidance technology and system for UAVs[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2017 (in Chinese).
- [7] MINGU K, YODAN K. Multiple UAVs nonlinear guidance laws for stationary target observation with waypoint incidence angle constraint[J]. *International Journal of Aeronautical and Space Sciences*, 2013, 14(1): 67-74.
- [8] MINGU K, YODAN K. Error dynamics-based Lyapunov guidance law for stationary target observation[J]. *IFAC Proceedings Volumes*, 2011, 44(1): 2042-2047.
- [9] MEDEIROS F L L, GOMES V C F, AQUINO M R C D, et al. A computer vision system for guidance of vtol uavs autonomous landing[C] // 2015 Brazilian Conference on Intelligent Systems (BRACIS). Piscataway, NJ: IEEE Press, 2015: 333-338.
- [10] YANG S, SCHERER S A, ZELL A. An onboard monocular vision system for autonomous takeoff, hovering and landing of a micro aerial vehicle[J]. *Journal of Intelligent & Robotic Systems*, 2013, 69(1-4): 499-515.
- [11] GAUTAM A, SUJIT P B, SARIPALLI S. A survey of autonomous landing techniques for UAVs[C] // International Conference on Unmanned Aircraft Systems. Piscataway, NJ: IEEE Press, 2014: 1210-1218.
- [12] JUNG W, LIM S, LEE D, et al. Unmanned aircraft vector field path following with arrival angle control[J]. *Journal of Intelligent & Robotic Systems*, 2016, 84(1-4): 311-325.
- [13] MIN B M, TAHK M J, SHIM D H, et al. Guidance law for vision-based automatic landing of UAV[J]. *International Journal of Aeronautical and Space Sciences*, 2007, 8(1): 46-53.
- [14] 牛铁峰, 凌黎华. 无人机规避或跟踪空中目标的自适应运动导引方法[J]. *国防科技大学学报*, 2017, 39(4): 116-124.
NIU Y F, LING L H. Adaptive motion guidance method of avoiding or tracking aerial targets for unmanned aerial vehicles[J]. *Journal of National University of Defense Technology*, 2017, 39(4): 116-124 (in Chinese).
- [15] RATNOO A, GHOSE D. Impact angle constrained guidance against nonstationary nonmaneuvering targets[J]. *Journal of Guidance, Control, and Dynamics*, 2010, 33(1): 269-275.
- [16] LEE H, JUNG S, SHIM D H. Vision-based UAV landing on the moving vehicle[C] // 2016 International Conference on Unmanned Aircraft Systems (Icuas). Piscataway, NJ: IEEE Press, 2016: 1-7.
- [17] CHU H P, MA L, SHAO Z J. Trajectory optimization and online re-planning for vertical takeoff vertical landing on lunar surface[C] // Proceedings of 2016 IEEE Chinese Guidance, Navigation and Control Conference. Piscataway, NJ: IEEE Press, 2016: 2999-3004.
- [18] 张炳亮. 基于三维合作目标的视觉探测与导引算法研究[D]. 南京: 南京航空航天大学, 2017.
ZHANG B L. Research on visual detection and guidance algorithm based on 3D cooperative target[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2017 (in Chinese).

作者简介:

魏祥灰 女, 硕士研究生。主要研究方向: 无人机视觉导引。

唐超颖 女, 博士, 副教授, 硕士生导师。主要研究方向: 导引控制与图像处理。

王彪 男, 博士, 副教授, 硕士生导师。主要研究方向: 飞行器控制与视觉导引。

Visual guidance algorithm design for autonomous landing based on two points in space

WEI Xianghui, TANG Chaoying, WANG Biao*

(College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: In order to increase the landing efficiency, improvements are made in two aspects: landing velocity vector field and guidance law design. Landing velocity vector field is designed based on ellipse curve for the requirements of shorter flight path and less maneuverability. Meanwhile, the flight path azimuth angle command is generated based on the relationship between the image coordinate system and the body-fixed frame. With reference to the tangential direction of the ellipse, the flight path elevation angle is tuned and combined with the cooperative vector features. Speed command is calculated using image information. Finally, the requirements of the traditional trajectory and the proposed trajectory on the directional maneuverability are compared in theory. The relationship between the trajectory parameters and the UAV turning performance is then shown. The system simulation platform is built based on Simulink, and the required cooperative vector is calculated. The results show that the UAV accurately lands on the target with curved trajectory, which meets the needs in practical applications.

Keywords: VTOL UAV; autonomous landing; vision-based guidance law; velocity vector field; turning maneuver performance

Received: 2018-05-21; **Accepted:** 2018-09-19; **Published online:** 2018-10-10 16:27

URL: kns.cnki.net/kcms/detail/11.2625.V.20180930.0938.001.html

Foundation item: Fundation of Graduate Innovation Center in NUAA-the Fundamental Research Funds for the Central Universities (kfjj20170302)

* **Corresponding author.** E-mail: wangbiao@nuaa.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2018.0299

一种新的非高斯随机振动信号的模拟方法

夏静, 袁宏杰*, 徐如远

(北京航空航天大学 可靠性与系统工程学院, 北京 100083)



摘 要: 传统的环境振动试验常将随机振动信号假设服从高斯分布,且利用功率谱密度(PSD)来作为试验条件。然而实际环境中结构所受到的振动激励很多呈现非高斯性,且PSD只能描述信号的低阶统计量,无法描述非高斯信号的峭度和偏度等高阶统计量。针对此情况,研究了在窗函数幅值调制法(AMT)基础上利用自PSD和峭度生成非高斯信号的方法。针对调制信号的生成,提出了近似模拟方法。通过Weibull和Beta 2种分布构造调制信号,研究分布参数与目标峭度值之间的关系,并分析2种分布合成目标峭度值的范围。案例验证了仿真生成的非高斯信号与实测外场数据具有相同的PSD、概率密度函数(PDF)和峭度值,进而证明了方法的正确性。

关键词: 幅值调制法(AMT); 近似模拟方法; 峭度; Weibull分布; Beta分布

中图分类号: TP391.9

文献标识码: A **文章编号:** 1001-5965(2019)02-0366-07

工程中常将随机振动信号假设服从高斯分布^[1-2],并且通过功率谱密度(PSD)作为试验条件。然而实际中,结构所受到的振动激励很多不满足高斯分布的假设,比如路面载荷^[3]、加载在结构上的波浪载荷^[4]以及输电线系统受到的风载^[5],而PSD仅仅可以表征信号的低阶统计量,不能够完全表征其高阶统计量,并且非高斯信号对结构往往会产生更严重的破坏,仅仅通过高斯信号来模拟其造成的损伤影响将与实际有较大的偏差。传统的实现非高斯信号的方法是时域波形再现技术(TWR),但是该方法只能复现外场环境的一次试验数据,不能反映外场数据的可变性,因而使用局限性较大。鉴于传统方法的缺陷,许多学者提出了其他非高斯信号的模拟方法。

Steinwolf^[6-7]提出采用相位调制法来生成平稳的非高斯信号进而模拟路面振动激励,并且通过不断地调整相位值,使仿真信号的峭度达到目

标峭度值。Zheng^[8]通过响应非高斯信号的PSD来生成平稳的高斯信号,并利用一种改进的“零记忆”非线性转换方法将生成的高斯信号转化为具有指定峭度的非高斯信号。Yu^[9]提出了利用散粒噪声来生成具有指定的峭度和PSD的非高斯信号,并实现了通过试验振动台来生成该信号进而应用于可靠性加速实验中。蒋瑜等^[10]提出了基于二次相位调制的方法分别生成对称和非对称的超高斯信号,并分析了相位角的变化对偏度和峭度的影响。吴家驹等^[11]研究了利用PSD和峭度产生非高斯随机波形的窗函数调制方法,引入了表征平稳的非高斯随机振动的新的参数——跌宕周期,提出了重构非高斯随机振动的新的方法——Beta分布随机数重新排列技术,该方法既可以生成平稳的非高斯信号也可以生成非平稳的非高斯信号;但文中并没有具体探究Beta分布的参数与目标峭度值之间的关系以及该分布的适用

收稿日期: 2018-05-24; 录用日期: 2018-08-24; 网络出版时间: 2018-09-13 11:40
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180911.1125.003.html

*通信作者. E-mail: yuanhongjie@buaa.edu.cn

引用格式: 夏静, 袁宏杰, 徐如远. 一种新的非高斯随机振动信号的模拟方法[J]. 北京航空航天大学学报, 2019, 45(2): 366-372.
XIA J, YUAN H J, XU R Y. A new simulation method of non-Gaussian random vibration signal[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 366-372 (in Chinese).

范围。徐飞^[12]采用幅值调制法 (AMT) 来仿真生成基于风诱载的非高斯信号,首先构造了服从 Weibull 的含参分布,然后通过不断迭代寻优的方法使得仿真的调制信号与实际信号随时间变化的均方根值 (Root Mean Square, RMS) 之间的误差最小,从而确定分布的参数,合成最终的非高斯信号;但文中并没有分析调制信号参数与合成非高斯信号的峭度之间的关系,以及 Weibull 分布的适用范围,且求解调制参数的过程相对繁琐,误差精度与样本量的大小有很大的关系。

因此,针对现有方法的不足,本文基于窗函数 AMT,提出近似模拟的方法分别生成服从 Weibull 和 Beta 分布的调制信号,进而合成最终的非高斯信号,并分别分析目标非高斯峭度值与 2 种分布的参数之间的关系以及 2 种分布适用范围。

1 非高斯信号的特征描述

一个随机过程可能呈现平稳或非平稳性,高斯或非高斯性。一个各态历经的高斯过程的概率密度函数 (PDF) 分布为

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1)$$

式中: μ 为均值; σ 为标准差。

广义上讲,幅值不服从高斯分布的随机过程均称为非高斯随机过程,其中既包括具有确定的 PDF 的随机过程,如指数分布、Weibull 分布;又包括没有确定 PDF 表达式的分布。针对高斯过程,其特性可由 PSD 完全表征,但是对于非高斯分布,需要采用更高阶统计量即高阶矩和高阶累积量来表征。偏度 s 和峭度 k 来定量表征随机过程的非高斯特性。

偏度的计算式为

$$\begin{cases} s = \frac{m_3}{m_2^{3/2}} \\ m_j = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^j \quad j = 2, 3, \dots \end{cases} \quad (2)$$

式中: m_j 为高阶中心矩。

峭度的计算式为

$$k = \frac{m_4}{m_2^2} \quad (3)$$

偏度是衡量随机信号偏离对称分布的程度,偏度不为零的信号必然服从非对称分布。而峭度表征统计频率曲线接近分布中心时的大致状态,其不仅可以区分高斯和非高斯信号,还可以进一步将非高斯信号分为亚高斯和超高斯信号。相对于高斯分布,超高斯分布的随机过程的幅值分布

中心区域较窄呈现一个大的峰,并在概率密度曲线上呈现较长的拖尾。而亚高斯分布的随机过程幅值中心区域较窄且比较均匀,但整体的分布区较窄。

信号的另一特性指标为自相关函数 (Auto Correlation Function, ACF),自相关函数是信号与延迟后信号之间相似性的一种度量,表征了信号功率随着频率的变化情况,即信号功率在频域的分布状况。其表达式为

$$R_x(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} x(t)x(t-\tau) dt \quad (4)$$

式中: $x(t)$ 为时域的信号; τ 为延迟量。

相关函数在时间域上描述随机过程的统计特征,而功率谱是在频率域上描述随机过程的统计特征。两者互为对傅里叶变换对。

2 非高斯随机信号的模拟

针对风诱载以及路面载荷激励等非高斯随机振动信号的模拟,利用窗函数 AMT 先生成平稳的高斯信号,然后乘以一个调制信号进而合成最终的非高斯信号,具体的过程如下。

2.1 实际外场数据的特征分析

分析采集的实际外场数据的特征,获取其统计特征如 PSD 以及峭度、偏度等信息,以此判断其是否为非高斯信号。

2.2 平稳高斯信号的模拟

基于快速傅里叶逆变换 (IFFT) 的原理,Newland^[13]利用信号的自功率谱和随机相位来构造幅值,进而生成平稳的零均值的高斯时域信号 $g(t)$ 为

$$g(t) = \text{Re} \left\{ \sum_{u=0}^{N-1} A_u e^{i\varphi_u} e^{i2\pi f_u \Delta f} \right\} \quad (5)$$

式中: $\text{Re} \{ \cdot \}$ 表示取实部; A_u 为幅值, $A_u = \sqrt{2S_x(f_u) \Delta f}$, $u = 0, 1, \dots, N-1$, $S_x(f_u)$ 为功率谱密度, φ_u 为 $[-\pi, \pi]$ 上分布的独立相位角; Δf 为频率间隔, f_u 为各频率值, $f_u = u \Delta f$, $u = 0, 1, \dots, N-1$ 。

由式(5)可以看出:当 $N \rightarrow \infty$ 时根据中心极限定理,得到的随机过程近似服从高斯分布,而且它的 PSD 也接近于所要模拟的目标 PSD。因此根据式(1)中外场数据的 PSD,生成平稳的高斯时域信号 $g(t)$ 。

2.3 调制信号的模拟

对于调制信号的生成,Smallwood^[14]建议使用 Beta 分布,因为该分布只产生正值且灵活性强;而徐飞^[12]提出采用 Weibull 信号作为调制信号,

因此分别利用2种分布生成调制信号,研究其参数与目标非高斯信号峭度值之间的关系并探究2种分布的适用范围。由AMT^[15]可知:

$$G(t) = w(t) \cdot g(t) \quad (6)$$

式中: $w(t)$ 为调制信号; $g(t)$ 为高斯信号; $G(t)$ 为非高斯信号。

2.3.1 调制信号参数的确定

利用非高斯信号的峭度公式,归一化的非高斯信号、高斯信号的均值都为0,且高斯信号与调制信号相互独立,所以非高斯信号的峭度为

$$k = \frac{E(Y^4)}{E(Y^2)^2} = \frac{E\{w^4 \cdot g^4\}}{E\{w^2 \cdot g^2\} \cdot E\{w^2 \cdot g^2\}} = \frac{3E_w\{w^4\}}{E_w\{w^2\} \cdot E_w\{w^2\}} \quad (7)$$

由式(7)可知,目标峭度值与调制信号的参数有关,所以分别探究当调制信号服从Weibull分布和Beta分布时,其参数与目标峭度之间的关系。

1) Beta分布

Beta分布是由区间在 $[0,1]$ 的2个正值 α, β 决定,其高阶原点矩为

$$E_w(w^n) = \prod_{i=0}^{n-1} \frac{\alpha + i}{\alpha + \beta + i} \quad (8)$$

假设目标非高斯信号偏度为0,可得调制信号的偏度值也为0,偏度的计算式为

$$s = \frac{2(\beta - \alpha) \sqrt{\alpha + \beta + 1}}{(\alpha + \beta + 2) \sqrt{\alpha\beta}} \quad (9)$$

由此可得 $\alpha = \beta$,代入式(8)可得

$$E_w(w^n) = \prod_{i=0}^{n-1} \frac{\beta + i}{2\beta + i} \quad (10)$$

故根据式(7)和式(10)可得非高斯信号的峭度值为

$$k = \frac{3E_w(w^4)}{E_w(w^2) \cdot E_w(w^2)} = 6 \frac{2\beta^3 + 11\beta^2 + 17\beta + 6}{4\beta^3 + 14\beta^2 + 16\beta + 6} \quad (11)$$

由式(11)可知,一旦给定了目标峭度值,则Beta分布的参数即可确定。

图1为目标峭度与参数 β 之间的关系,由图中可以看出,当峭度值大于6.04时, β 为负值,不满足分布要求,因此调制信号参数的合理范围为 $[0.2, 10.2]$ 。当 β 为 $[0.2, 10.2]$ 时对应的调制信号的PDF如图2所示,由图中可知,当 $\beta > 1$ 时,调制信号的PDF符合Beta分布,对应的非高斯信号的峭度值小于5.4;当 $0.2 < \beta < 1$ 时,调制信号的PDF分布在两侧,对应的非高斯信号的峭度大于5.4,这是不合理的。因此可以得出结论:

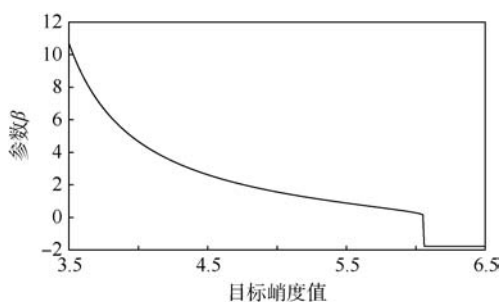


图1 目标峭度与参数 β 的关系

Fig. 1 Relationship between aimed kurtosis and parameter β

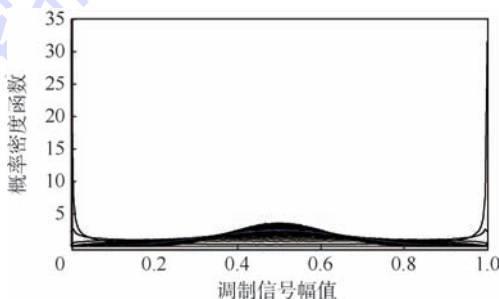


图2 参数 β 为 $[0.2, 10.2]$ 时对应调制信号的PDF
Fig. 2 PDF of modulation signal when β is $[0.2, 10.2]$

当调制信号为Beta分布时,只能合成目标峭度小于5.4的非高斯信号,对应的合理的参数范围为 $[1, 10.2]$ 。

2) Weibull分布

Weibull分布主要是由形状参数 r 和比例参数 λ 的2个正值构成,其高阶原点矩为

$$E_w(w^n) = \lambda^n \Gamma\left(1 + \frac{n}{r}\right) \quad (12)$$

故根据式(7)和式(12)可得非高斯信号的峭度值为

$$k = \frac{3E_w(w^4)}{E_w(w^2) \cdot E_w(w^2)} = \frac{3\Gamma\left(1 + \frac{4}{r}\right)}{\Gamma\left(1 + \frac{2}{r}\right) \cdot \Gamma\left(1 + \frac{2}{r}\right)} \quad (13)$$

由式(13)可知,一旦给定了目标峭度值,则Weibull分布的参数 r 即可确定。

因为仿真非高斯信号与高斯信号具有相同的PSD,即两者具有相同的均方值,则调制信号的均值为1。因此Weibull分布的比例参数 λ 为

$$\lambda = \frac{1}{\sqrt{\Gamma\left(1 + \frac{2}{r}\right)}} \quad (14)$$

目标峭度值为 $[3.5, 12]$ 时,形状参数 r 随着峭度的变化如图3所示。从图中可以看出,随着

峭度增加参数 r 逐渐减小。当峭度小于 6.4 时, r 为负值, 这是不合理的, 故合理的形状参数 r 的范围为 $[1.28, 1.88]$ 。形状参数 r 在 $[1.28, 1.88]$ 时对应的 PDF 如图 4 所示, 可以看出其 PDF 符合 Weibull 分布, 故可得出结论: 当调制信号为 Weibull 分布时, 只能生成目标峭度大于 6.4 的非高斯信号。

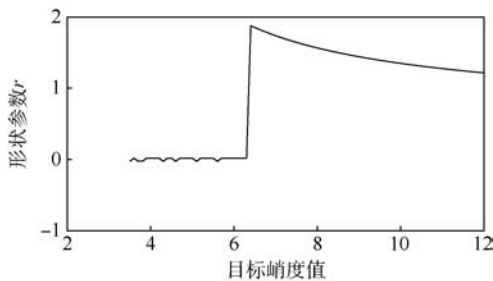


图 3 目标峭度与形状参数 r 的关系

Fig.3 Relationship between aimed kurtosis and shape parameter r

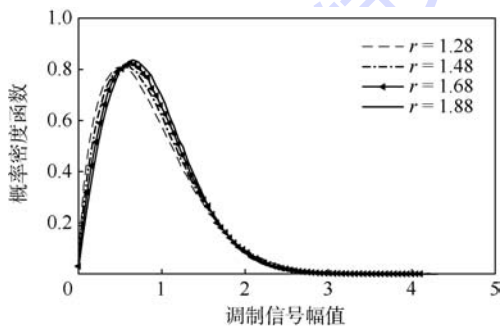


图 4 形状参数 r 为 $[1.28, 1.88]$ 时对应调制信号的 PDF

Fig.4 PDF of modulation signal when shape parameter r is $[1.28, 1.88]$

2.3.2 调制信号的近似模拟方法

本文采用一种新的近似模拟的方法用于生成调制信号进而合成最终的非高斯信号。首先根据实测的非高斯信号的幅值趋势生成满足一定分布的随机数个数为 N 的调制信号, 然后利用 resample 命令进行插值, 每个随机点之间的插值数为 M 个, 保证 $M \times N$ 具有与高斯信号相同的长度。

为了确定插值数 M , 引入了自相关函数来表征。通过实测数据的随时间变化的 RMS 的自相关函数, 即可得到其零穿越时间 t , 信号的采用频率为 f_s , 故每个随机点之间的插值数 $M = tf_s$ 。

图 5 为生成的服从 Weibull 分布的一个调制信号。假设目标峭度值为 8, 产生的样本量为 $N = 20$, 扩增量 $M = 500$ 。

2.4 合成非高斯信号

将 2.2 节中生成的平稳的高斯信号 $g(t)$ 作

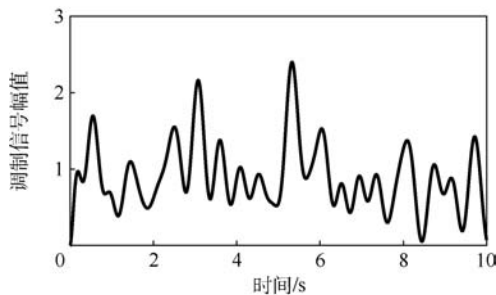


图 5 服从 Weibull 分布的调制信号

Fig.5 Modulation signal following Weibull distribution

为载波信号乘以 2.3 节中生成的调制信号 $w(t)$, 生成最终的非高斯信号 $G(t)$ 。

3 实例分析

3.1 实测数据的特征

通过实际测试得到外场数据, 该数据是由安装在一个运动的车辆上的加速度传感器采集得到, 通过分析得到实测数据的采样频率为 1653 Hz, 采样点为 160 076, 峭度值为 10.95。实测数据及其随时间变化的均方根值 RMS 如图 6 所示。实测数据的 PSD 如图 7 所示, 随时间变化的 RMS 的自相关函数如图 8 所示。由图 7 和图 8 可以看出, 实测信号的 RMS 相关函数零穿越时间为 4 s。

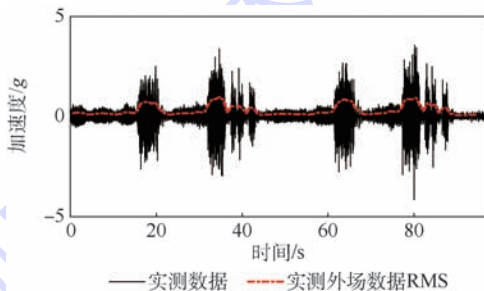


图 6 实测外场数据和随时间变化的 RMS

Fig.6 Measured ourfield data and variation of RMS with time

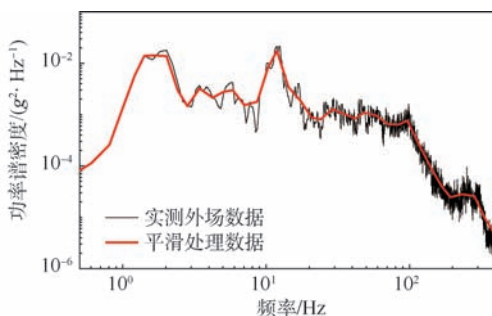


图 7 实测外场数据和平滑处理数据的 PSD

Fig.7 PSD of measured outfield data and smooth data

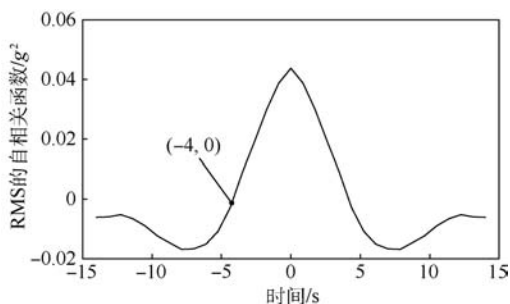


图8 实测外场数据 RMS 的自相关函数

Fig. 8 Auto correlation function of RMS of measured outfield data

3.2 非高斯信号的仿真模拟

3.2.1 高斯信号模拟

因为实测数据的峭度大于3,所以是典型的非高斯信号。根据实测外场数据的 PSD 采用傅里叶逆变换原理生成稳态的高斯信号如图9所示。其峭度值为3.01,信号的 PSD 如图10所示,由图可以看出产生的高斯信号的 PSD 与实测外场数据的 PSD 相同。

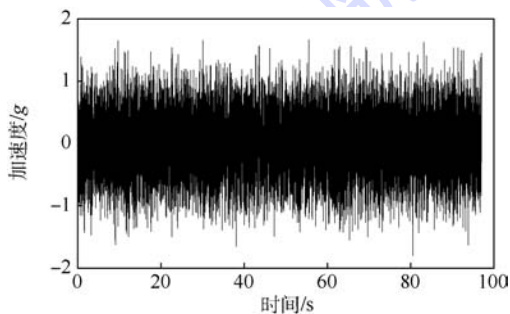


图9 仿真的高斯信号

Fig. 9 Simulated Gaussian signal

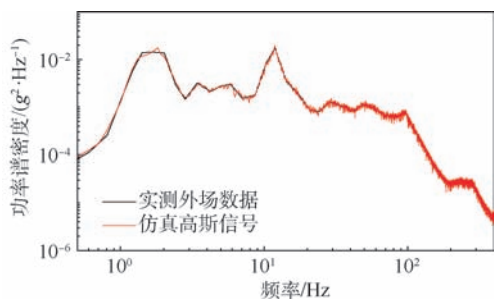


图10 实测外场数据和仿真的高斯信号的 PSD

Fig. 10 PSD of measured outfield data and simulated Gaussian signal

3.2.2 调制信号模拟

根据实测外场数据的峭度值为10.95,其峭度值大于6.4,所以可以采用 Weibull 分布生成调制信号,由式(13)和式(14)可以得到 Weibull 分布的参数为: $r = 1.2779$, $\lambda = 0.8473$ 。为了构造调制信号,根据实测外场数据 RMS 的自相关函数

的零穿越时间为4s,采样频率 $f_s = 1653$ Hz,故采样间隔点数为6612,根据采用点数一般为1024的整数倍,所以插值数 $M = 8192$,根据总的采样点数,可以确定生成的调制信号如图11所示,仿真的自相关函数和实测外场数据 RMS 的自相关函数如图12所示,从图中可以看出两者的自相关函数的零穿越时间很接近,说明仿真的调制信号满足要求。

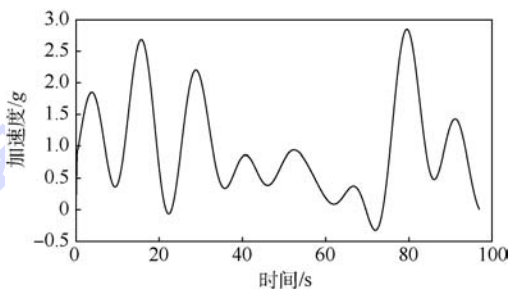


图11 仿真的调制信号

Fig. 11 Simulated modulating signal

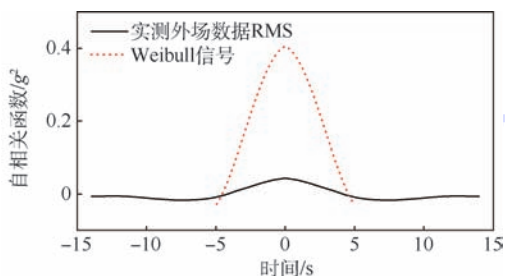


图12 实测外场数据 RMS 和 Weibull 信号的自相关函数

Fig. 12 Auto correlation function of measured outfield data RMS and Weibull signal

3.2.3 仿真的非高斯信号

根据生成的高斯信号和调制信号,利用 AMT 将两者相乘得到最终的非高斯信号。由此产生的非高斯信号如图13所示,其峭度值为11.12,与外场数据的峭度值相近。图14为仿真的非高斯信号的 PSD 和实测外场数据的 PSD,图15为仿真信号的 PDF 和实测外场数据的 PDF,由图可以看出两者的 PSD 和 PDF 很接近,说明方法的准确性。

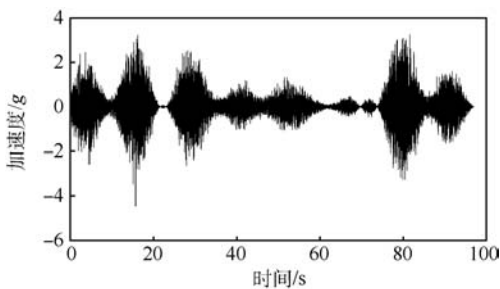


图13 仿真的非高斯信号

Fig. 13 Simulated non-Gaussian signal

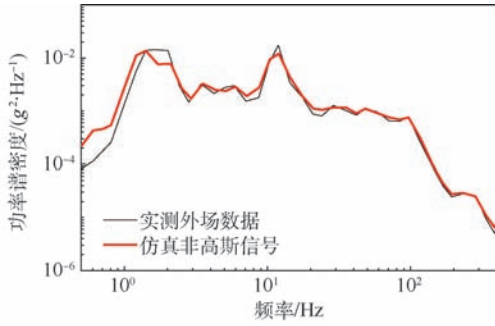


图 14 实测外场数据和仿真非高斯信号的 PSD

Fig. 14 PSD of measured outfield data and simulated non-Gaussian signal

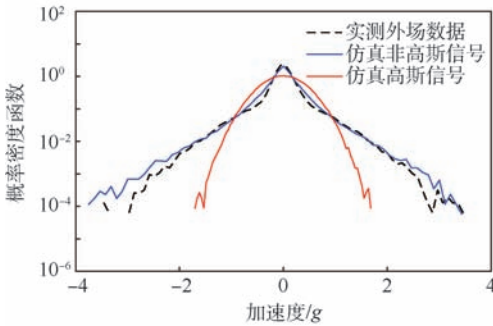


图 15 不同信号的概率密度函数

Fig. 15 PDF of different signals

4 结 论

1) 提出利用 AMT 生成非高斯信号的方法,通过案例得到仿真的信号具有与实测信号相同的 PSD、PDF 和峭度,说明了本文方法的正确性。

2) 对于调制信号,采用近似模拟生成的方法,分别利用 Weibull 分布和 Beta 分布生成调制信号,研究分析了分布参数与目标峭度之间的关系。

3) 分别探究 Weibull 分布和 Beta 分布适合生成非高斯信号峭度值的范围,分析可知 Weibull 分布只适合于生成峭度大于 6.4 的非高斯信号,而 Beta 分布只适合于生成峭度小于 5.4 的非高斯信号。

参考文献 (References)

- [1] BENDAT S,PIERSOL G. Random data: Analysis and measurement procedures [M]. 4th ed. Los Angeles: John Wiley and Sons, Inc. ,2010:1799-1806.
- [2] VINCENT R. On the non-Gaussian nature of random vehicle vibration[C]// Proceedings of the World Congress on Engineering,2017,2:1219-1224.
- [3] KNANI K B. Fatigue damage assessment of a car body-in-white

using a frequency-domain approach[J]. International Journal of Materials & Products Technology,2007,30(1):172-198.

- [4] ZAKHAROV V E, DYACHENKO A I, PROKOFIEV A O. Freak waves as non-linear stage of Stokes wave modulation instability[J]. European Journal of Mechanics,2006,25(5):677-692.
- [5] YI H,BO L. Non-stationary and non-Gaussian characteristics of wind speeds[J]. Wind and Structures,2017,24(1):59-78.
- [6] STEINWOLF A. Closed-loop shaker simulation of non-Gaussian random vibration. Part 1: Discussion and methods[J]. Test Engineering and Management,2006,68(3):10-13.
- [7] STEINWOLF A. Closed-loop shaker simulation of non-Gaussian random vibration. Part 2: Numerical and experimental results [J]. Test Engineering and Management,2006,68(5):14-19.
- [8] ZHENG R H. Control method for multi-input multi-output non-Gaussian random vibration test with cross spectra consideration [J]. Chinese Journal of Aeronautics,2017,30(6):1895-1906.
- [9] YU Y Y. Generation of non-Gaussian random vibration excitation signal for reliability enhancement test[J]. Chinese Journal of Aeronautics,2007,20(3):236-239.
- [10] 蒋瑜,陶俊勇,王得志,等. 一种新的非高斯随机振动数值模拟方法[J]. 振动与冲击,2012,31(19):169-173.
- JIANG Y,TAO J Y,WANG D Z, et al. A novel approach for numerical simulation of a non-Gaussian random vibration[J]. Journal of Vibration and Shock,2012,31(9):169-173 (in Chinese).
- [11] 吴家驹,付玮,张鹏飞,等. 基于 β 分布随机数排序的非高斯振动模拟方法[J]. 强度与环境,2017,44(2):10-17.
- WU J J,FU W,ZHANG P F, et al. Simulation method of non-Gaussian vibration base on re-arrangement of β distribution random number[J]. Structure & Environment Engineering,2017,44(2):10-17 (in Chinese).
- [12] XU F. On the shaker simulation of wind-induced non-Gaussian random vibration[J]. Shock and Vibration,2016,2016(6):1-10.
- [13] NEWLAND D E. An introduction to random vibrations and spectral analysis [M]. 2nd ed. New York: Longman Inc. , 1984: 1102-1203.
- [14] SMALLWOOD D O. Vibration with non-Gaussian noise [J]. Journal of the Institute of Environmental Science and Technology,2009,52(3):13-20.
- [15] RIZZI S A, PRZEKOP A, TURNER T L. On the response of a nonlinear structure to high kurtosis non-Gaussian random loadings[C]// The 8th International Conference on Structural Dynamics,2011:2697-2704.

作者简介:

夏静 女,硕士研究生。主要研究方向:可靠性与环境试验技术。

袁宏杰 男,博士,副教授,硕士生导师。主要研究方向:可靠性评估与验证、环境试验设计等。

A new simulation method of non-Gaussian random vibration signal

XIA Jing, YUAN Hongjie*, XU Ruyuan

(School of Reliability and Systems Engineering, Beihang University, Beijing 100083, China)

Abstract: The traditional environmental vibration tests usually assume that the random vibration signal follows Gaussian distribution and power spectral density (PSD) is used as the test conditions, while in practice, lots of vibration excitations to the structure are non-Gaussian distribution and PSD can only describe the low-order statistics of signals but not for the high-order statistics of the non-Gaussian signals such as the kurtosis and skewness. So a method based on amplitude modulation technique (AMT) of window function to generate the non-Gaussian signal with PSD and kurtosis is developed. An approximated simulation method is provided to generate the modulation signal. Two kinds of statistical distribution, Beta and Weibull, are used to construct modulation signal, the relation between target kurtosis and modulation signal distribution parameters is explored, and the suitable kurtosis range of two distributions is discussed. A case study is presented to show that the synthesized non-Gaussian signal has the same PSD, probability density function (PDF), and kurtosis as outfield measured data, which verifies the correctness of the method.

Keywords: amplitude modulation technique (AMT); approximate simulation method; kurtosis; Weibull distribution; Beta distribution

Received: 2018-05-24; **Accepted:** 2018-08-24; **Published online:** 2018-09-13 11:40

URL: kns.cnki.net/kcms/detail/11.2625.V.20180911.1125.003.html

Foundation item: National Defense Basic Scientific Research Program of China (61325102)

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

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

DOI: 10.13700/j.bh.1001-5965.2018.0301

基于改进的动态 Kriging 模型的结构可靠度算法

魏娟^{1,2}, 张建国^{1,2,*}, 邱涛^{1,2}

(1. 北京航空航天大学 可靠性与系统工程学院, 北京 100083;

2. 北京航空航天大学 可靠性与环境工程技术重点实验室, 北京 100083)

摘 要: 对于复杂航空航天机械产品, 极限状态方程往往表现出隐式、高度非线性的特点, 而且通常需要调用有限元分析, 从而耗费大量时间。将混合粒子群-模拟退火 (PSOSA) 算法应用到 Kriging 模型中相关参数的寻优过程, 提高了预测精度。同时结合动态更新机制, 逐渐加入样本点, 尽可能减少函数的调用次数, 从而提高了计算效率, 并将该算法应用到结构可靠性分析中。通过案例分析, 和传统蒙特卡罗模拟方法、响应面等经典方法进行对比, 所提算法与蒙特卡罗模拟方法计算结果更加接近, 计算时间大大缩短, 效率和精度都明显改进。

关键词: 极限状态函数; 动态更新; Kriging 模型; 粒子群-模拟退火 (PSOSA) 算法; 可靠性

中图分类号: TB114.3

文献标识码: A **文章编号:** 1001-5965(2019)02-0373-08

当现有的可靠性分析方法应用于复杂的工程结构时, 往往面临巨大的挑战^[1], 功能函数通常是高度非线性甚至是隐式的, 而且需要借助有限元分析 (Finite Element Analysis, FEA) 进行评估, 计算量大, 计算时间长^[2-3]。一阶可靠度算法 (First Order Reliability Methodology, FORM)、二阶可靠度算法 (Second Order Reliability Methodology, SORM) 已被广泛用于估算结构系统失效的概率^[4], 但是只能适用于功能函数表达式已知的情况。蒙特卡罗模拟 (Monte Carlo Simulation, MCS) 方法是结构可靠性分析中最简单也是最广泛使用的方法, 计算精度高, 适用于结构可靠性的隐式问题。但由于工程结构的失效概率通常较小, MCS 须生成大量样本以确保准确性, 所带来的计算成本可能会变得非常高, 从而极大地限制了 MCS 在工程中的应用。

为了降低可靠性分析的计算成本, 代理模型越来越多地被用来实现结构可靠性。代理模型作

为一个近似模型可以找到输入和输出之间的潜在关系。如果与功能函数足够接近, 则代理模型可以完全表示功能函数, 以准确评估性能函数的值。常用的方法有响应面方法 (Response Surface Methodology, RSM)^[5]、支持向量机^[6]、神经网络^[7]、Kriging 模型^[8]。其中, Kriging 模型是一种方差估计最小的无偏估计的随机过程算法, 由于其对非线性函数的良好近似能力和独特的误差估计功能^[9], 越来越多地用于可靠度的求解过程。Zhang 等^[10]运用 Kriging 模型进行结构可靠度分析, Liu 等^[11]运用 Kriging 模型和 MCS 方法进行齿轮接触可靠性的分析。刘瞻^[12]和陈志英^[13]等将人工智能优化算法融入到 Kriging 模型的构造中, 进一步提高了预测精度。为了降低复杂结构的可靠性分析的计算成本并实现快速可靠性分析, 尽可能减少样本数量, Ma 等^[14]在采样过程中应用学习函数, 从而构造 Kriging 模型, Li 等^[15]提出一种基于最可能失效点 (Most Probable Point,

收稿日期: 2018-05-28; 录用日期: 2018-08-24; 网络出版时间: 2018-09-07 13:54

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180906.0944.002.html

基金项目: 国家自然科学基金 (51675026)

* 通信作者. E-mail: zjg@buaa.edu.cn

引用格式: 魏娟, 张建国, 邱涛. 基于改进的动态 Kriging 模型的结构可靠度算法[J]. 北京航空航天大学学报, 2019, 45(2): 373-380. WEI J, ZHANG J G, QIU T. Structural reliability algorithm based on improved dynamic Kriging model[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 373-380 (in Chinese).

MPP)的局部抽样方法,提高了 Kriging 的精度。Echard 等^[16]提出动态 Kriging 联合 MCS 的可靠度计算方法。Xia 等^[17]提出了一种基于自适应动态泰勒 Kriging 的高效可靠性优化算法。通过二元粒子群算法优化选择基函数,并确定具有期望拟合误差的最小数量的采样数据。

由于初始训练的高质量样本总是难以获得,本文在训练小样本的基础上,采用修改代理模型拟合误差的动态更新机制。针对复杂结构隐式功能函数的可靠性分析问题,本文将混合粒子群-模拟退火(Particle Swarm Optimization-Simulated Annealing, PSOSA)算法应用到动态 Kriging 模型中相关参数的寻优过程,并将改进的 Kriging 模型应用于可靠度的分析计算问题。该方法运用 PSOSA 算法高效求解最优相关参数,构造 Kriging 模型;同时基于动态更新机制,逐步加入最佳样本点,从而不断修正 Kriging 模型的精度,并结合一次二阶矩法进行可靠度计算。

1 动态 Kriging 模型

1.1 Kriging 模型

近年来,Kriging 模型作为一种新型的响应面模型技术在航天等工程优化领域中得到了广泛应用。对于 k 个 n 维样本点及其对应的响应值 $\mathbf{Y} = [G(\mathbf{x}_1), G(\mathbf{x}_2), \dots, G(\mathbf{x}_k)]^T$, Kriging 模型采用高斯随机过程模型,其插值结果定义为已知样本函数响应值的线性加权^[9],即

$$G(\mathbf{x}) = \sum_{i=1}^p \beta_i f_i(\mathbf{x}) + z(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \boldsymbol{\beta} + z(\mathbf{x}) \quad (1)$$

式中: $\mathbf{x} = [x_1, x_2, \dots, x_k]^T$; $\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_p(\mathbf{x})]^T$ 为回归多项式基函数向量, p 为回归多项式的数量; $\boldsymbol{\beta} = [\beta_1, \beta_2, \dots, \beta_p]^T$ 为多项式参数向量; $z(\mathbf{x})$ 为服从正态分布 $N(0, \sigma^2)$ 的随机过程, σ 为标准差,协方差方程为

$$\text{cov}(z(\mathbf{x}_i), z(\mathbf{x}_j)) = \sigma^2 R(\mathbf{x}_i, \mathbf{x}_j) \quad (2)$$

式中: $R(\mathbf{x}_i, \mathbf{x}_j)$ 为样本点中任意 2 个输入向量 \mathbf{x}_i 和 \mathbf{x}_j 的空间相关方程,通常采用高斯相关方程,其形式为

$$R(\mathbf{x}_i, \mathbf{x}_j) = \exp \left[- \sum_{l=1}^n \theta_l (x_i^l - x_j^l)^2 \right] \quad (3)$$

式中: x_i^l 为输入向量 \mathbf{x}_i 的第 l 个分量; θ_l 为相关性参数,可通过最大似然估计得到

$$\max_{\theta > 0} \left\{ - \left[\frac{1}{2} \ln(|\mathbf{R}|) + \frac{k}{2} \ln(\sigma^2) \right] \right\} \quad (4)$$

式中: \mathbf{R} 为相关矩阵,其中元素 $R_{ij} = R(\mathbf{x}_i, \mathbf{x}_j)$ ($i, j = 1, 2, \dots, k$), 相关参数直接控制着 Kriging 模型的输入输出特性。

基于给定的样本点,多项式参数向量 $\boldsymbol{\beta}$ 与随机过程方差 σ^2 的估计值计算式分别为

$$\hat{\boldsymbol{\beta}} = (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{R}^{-1} \mathbf{Y} \quad (5)$$

$$\hat{\sigma}^2 = \frac{1}{k} (\mathbf{Y} - \mathbf{F} \hat{\boldsymbol{\beta}})^T \mathbf{R}^{-1} (\mathbf{Y} - \mathbf{F} \hat{\boldsymbol{\beta}}) \quad (6)$$

式中: \mathbf{F} 为由回归多项式函数值组成的回归矩阵^[18]。

对于任意未知点 \mathbf{x}_0 , 其函数预测值 $\hat{G}(\mathbf{x}_0)$ 服从正态分布 $N(\mu_{\hat{G}}(\mathbf{x}_0), \hat{\sigma}_{\hat{G}}^2(\mathbf{x}_0))$, 其中最佳线性无偏估计 $\mu_{\hat{G}}(\mathbf{x}_0)$ 及均方差可表示为

$$\mu_{\hat{G}}(\mathbf{x}_0) = \mathbf{f}^T \hat{\boldsymbol{\beta}} + \mathbf{r}_0^T \mathbf{R}^{-1} (\mathbf{Y} - \mathbf{F} \hat{\boldsymbol{\beta}}) \quad (7)$$

$$\mathbf{u} = \mathbf{F}^T \mathbf{R}^{-1} \mathbf{r}_0 - \mathbf{f} \quad (8)$$

$$\hat{\sigma}_{\hat{G}}^2(\mathbf{x}_0) = \hat{\sigma}^2 (1 + \mathbf{u}^T (\mathbf{F}^T \mathbf{R}^{-1} \mathbf{F})^{-1} \mathbf{u} - \mathbf{r}_0^T \mathbf{R}^{-1} \mathbf{r}_0) \quad (9)$$

式中: $\mathbf{r}_0 = [R(\mathbf{x}_0, \mathbf{x}_1), R(\mathbf{x}_0, \mathbf{x}_2), \dots, R(\mathbf{x}_0, \mathbf{x}_k)]^T$ 。

1.2 动态更新机制

为了提高可靠性分析的效率,动态代理模型应符合以下条件:①初始样本数量小;②用于小样本学习的替代模型拟合;③每迭代一次,对 FEA 进行一次调用,以修正代理模型。

本文选择动态 Kriging 模型来减少调用 FEA 的次数。代理模型的计算精度主要取决于对极限状态面的逼近程度。一般来说,使用一些预设的训练样本很难构建具有高拟合精度的响应曲面,而响应曲面的动态更新可用于校正拟合误差。动态 Kriging 可靠度具体算法流程如下:

步骤 1 建立初始的样本集。通过拉丁超立方抽样生成初始样本点,并调用 FEA 计算其响应值。

步骤 2 构造动态 Kriging。首先,通过优化技术寻找最优相关参数;然后,设置最优基函数,并构造代理模型来近似目标函数。

步骤 3 利用 DACE 工具箱可求出模型对各个变量的梯度,结合 FORM 法计算 MPP^[12]。判断 MPP 是否收敛。若是,求解可靠度等指标。若否,进入步骤 4。

步骤 4 上述 MPP 作为新的样本点加入初始样本集,继续迭代。

上述过程实现了可靠度计算过程中, Kriging 模型在实际极限状态响应面附近拟合误差的自适应修正。减少了 FEA 的调用次数,这使计算成本大大降低,有利于减少模型对样本的依赖程度,加快了收敛速度。

2 改进的动态 Kriging 模型

2.1 优化算法

2.1.1 PSO 算法

假设搜索空间 $S \in \mathbb{R}^n$, $\mathbf{X}_k^i = \{X_1^i, X_2^i, \dots, X_n^i\}$ 为粒子 k 在第 i 次迭代的位置, $\mathbf{P}_k^b = \{P_1^b, P_2^b, \dots, P_n^b\}^T$ 为粒子 k 在搜索空间内的最佳位置, $\mathbf{P}^g = \{P_1^g, P_2^g, \dots, P_n^g\}^T$ 为整体最优位置, 在第 i 步迭代时, 粒子的速度和位置就会通过式(10)和式(11)更新:

$$\mathbf{v}_k^{i+1} = w\mathbf{v}_k^i + c_1R_1'(\mathbf{P}_k^b - \mathbf{X}_k^i) + c_2R_2'(\mathbf{P}^g - \mathbf{X}_k^i) \quad (10)$$

$$\mathbf{X}_k^{i+1} = \mathbf{X}_k^i + \mathbf{v}_k^{i+1} \quad (11)$$

式中: \mathbf{v}_k 为第 k 个粒子的速度; R_1' 和 R_2' 为 $0 \sim 1$ 之间的随机数; 非负常数 c_1 和 c_2 为学习因子, 迭代终止条件一般选为最大迭代次数或粒子群迄今为止搜索到的最优位置满足预定最小适应阈值; 系数 w 为权重因子, 实际上控制了搜索空间的搜索, 是跳出目前最小点的随机机制, 其作用是允许粒子在整个搜索区域内移动, 让搜索在局部区域内开展。

本文中系数 w 由式(12)确定:

$$w = \frac{2}{|2 - C - \sqrt{C^2 - 4C}|} \quad (12)$$

式中: $C = c_1 + c_2$ 。

PSO 的工作原理如下: 首先, 粒子群由搜索区域内随机分布的粒子组成, 然后, PSO 通过更新粒子的速度和位置来迭代寻找最优解。当满足收敛标准时, 循环终止。但是, 运用公式 \mathbf{P}^b 和 \mathbf{P}^g 来搜索最优解, 收敛速度快。然而所有粒子有移向最优经验点的趋势, 并在局部区域内搜索, 所以可能会导致收敛过早的问题。因此, 尽管 PSO 比其他方法收敛速度快, 但结果可能精度不高。

2.1.2 SA 算法

模拟退火算法起源于固体退火原理, 模拟了高温金属降温的热力学过程。在搜索过程中具有概率突跳的能力, 能够有效地避免搜索过程中陷入局部极小解。试验点从 \mathbf{x}_i 到 \mathbf{x}_{i+1} 的接受概率为

$$P = \begin{cases} 1 & y(\mathbf{x}_{i+1}) \leq y(\mathbf{x}_i) \\ e^{-\Delta y/T} & \text{其他} \end{cases} \quad (13)$$

式中: y 为目标函数, $\Delta y = y(\mathbf{x}_{i+1}) - y(\mathbf{x}_i)$ 和 T 为控制参数, 可类比为温度, 温度的降低实际上控制了接受概率。

设 φ_j 为每个变量接受的移动次数与试验次数的比值, 则扰动可写为

$$\mathbf{X}_{\text{trial}} = \mathbf{X}_{\text{current}} + R'(-1, 1)\mathbf{V} \quad (14)$$

式中: $R'(-1, 1)$ 为 $-1 \sim 1$ 之间的随机数; \mathbf{V} 为步

长向量。 V_{bi} 是变量上下界之差的一半, 可表示为

$$V_i = \begin{cases} V_{bi}[1 + 2(\varphi_i - 0.6)/0.6] & \varphi_i > 0.6 \\ V_{bi}[1 + 2(0.4 - \varphi_i)/0.4] & \varphi_i < 0.4 \end{cases} \quad (15)$$

式中: V_i 为 \mathbf{V} 中的元素。

理论上证明, 当采用非常缓慢的温度下降率时, 找到全局最小值的概率可以接近 1。

最常见的温度下降规律为

$$T^{i+1} = R_T T^i \quad (16)$$

式中: T_0 为初始温度; R_T 为 $0.8 \sim 0.9999$ 之间的常数。

2.1.3 混合 PSOSA 算法

在混合 PSOSA 算法中, 搜索最优解时, 速度矢量控制粒子在搜索空间内的移动方式。速度矢量有 3 部分组成:

1) 惯性项($w\mathbf{v}_k^i$), 使得粒子保持原来的速度向量, 以防粒子的速度急剧变化。

2) 认知行为($c_1R_1'(\mathbf{P}_k^b - \mathbf{X}_k^i)$), 包含粒子经历过的最优位置。

3) 社会行为($c_2R_2'(\mathbf{P}^g - \mathbf{X}_k^i)$), 给予粒子朝群体最优位置运动的趋势。

本文 PSOSA 算法的主要思想是通过更新群体的群体最佳位置来改善群体的社会行为。当 PSO 循环中最优解没有改进时, 旧的群体最佳位置将被替换为使用 SA 算法计算的新位置。新的最佳位置实际上向群体的社会领导者(即前一个全局最优解所属的粒子)发出信号, 用于更新其方向。PSOSA 算法充分结合了 PSO 算法全局搜索能力和 SA 算法的局部搜索能力, 能达到很好的收敛结果。算法示意图如图 1 所示, 具体算法如下:

步骤 1 初始化。 m 为粒子群粒子的数; i_{max} 为允许最大迭代次数; w 为权重因子; c_1 和 c_2 为学习因子; v_{max} 为速度阈值, T_0 为初始温度; R_T 为温度降低参数。

步骤 2 粒子的位置。可行域均匀分布 m 个粒子; 确定每个粒子的适应度值。

步骤 3 初始最佳位置的确定。确定每个粒子 \mathbf{P}_k^b 和群体的最佳位置 \mathbf{P}^g 。

步骤 4 更新每个粒子的速度和位置:

$$\mathbf{v}_k^{i+1} = w\mathbf{v}_k^i + c_1R_1'(\mathbf{P}_k^b - \mathbf{X}_k^i) + c_2R_2'(\mathbf{P}^g - \mathbf{X}_k^i)$$

$$\mathbf{X}_k^{i+1} = \mathbf{X}_k^i + \mathbf{v}_k^{i+1}$$

步骤 5 更新每个粒子的 \mathbf{P}_k^b 和群体的 \mathbf{P}^g : 若 $f(\mathbf{X}_k^{i+1}) \leq f(\mathbf{P}_k^b)$, 则 $\mathbf{P}_k^b = f(\mathbf{X}_k^{i+1})$; 若 $f(\mathbf{X}_k^{i+1}) \leq f(\mathbf{P}^g)$, 则 $\mathbf{P}^g = f(\mathbf{X}_k^{i+1})$ 。

步骤 6 判断最优解是否有所改进。如果第

$i+1$ 次迭代的 \mathbf{P}^k 没有优于第 i 次迭代的 \mathbf{P}^k , 则进入步骤 7; 否则, 更新温度 $T^{i+1} = T^i R_T$, 返回步骤 4。

步骤 7 SA 算法。 计算一个测试位置, $\mathbf{P}_{\text{trial}}^k = \mathbf{P}^k + R'(-1, 1)\mathbf{V}$, R' 取随机数。如果 $f(\mathbf{P}_{\text{trial}}^k) \leq f(\mathbf{P}^k)$, 则 $\mathbf{P}^k = \mathbf{P}_{\text{trial}}^k$, 更新温度, $T^{i+1} = T^i R_T$ 并继续迭代; 如果 $f(\mathbf{P}_{\text{trial}}^k) > f(\mathbf{P}^k)$, 检测 Metropolis Monte Carlo 接受概率, 如果 $e^{-\Delta f/T} \geq R'(0, 1)$, 则 $\mathbf{P}^k = \mathbf{P}_{\text{trial}}^k$ 并更新温度; 如果 $e^{-\Delta f/T} < R'(0, 1)$, 取目前的 \mathbf{P}^k 为最优解。返回步骤 4。

步骤 8 重复整个过程直到收敛。

该算法主要有以下几个特点:

- 1) 通过改进速度矢量的组成促使群体向整体最优位置所在的搜索空间移动, 这加快了收敛过程。
- 2) PSO 算法和 SA 算法的结合提高了整个寻优机制的质量和稳定性。
- 3) 在 SA 迭代中 PSO 速度公式中的认知成分

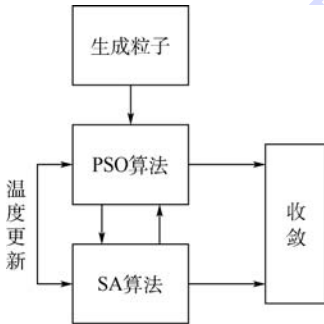


图 1 PSOSA 算法示意图

Fig. 1 Sketch map of PSOSA algorithm

得到了保留。

2.2 Kriging 建模流程

目前 Kriging 建模过程普遍采用的模式搜索法对初始点十分敏感, 容易造成求解 θ^* 无法收敛, Kriging 预测精度很低。主要原因是模式搜索所采用的单点序列搜索方式决定了它只能从一个指定的起点出发, 并收敛于一个距起点较近的局部最优解。为了寻找最优 θ^* , 只能将起点接近于 θ^* , 然而这很难确定。不仅如此, 模式搜索法搜索路径单一, 不能根据似然函数的变化而变化, 因此即使根据经验信息, 使起点接近 θ^* , 搜索路径也有可能偏离导致无法收敛^[13]。

本文一方面运用 PSOSA 算法进行搜索, 种群中各粒子在搜索最优解的过程中实现信息共享, 每个粒子不断根据个体极值与群体极值的变化情况动态调整着自己的搜索方向, 另外, 当 PSO 循环没有改善, 原来全局最优位置被通过 SA 计算得到的新位置代替, 从而在缺乏先验信息的前提下, 也能搜索到最优解 θ^* 。另一方面, 通过自适应动态更新机制减少调用 FEA 的次数, 从而提高计算效率。

其中, 运用 PSOSA 算法搜索极大似然意义下的最优相关参数时:

- 1) 似然函数作为适应度函数。
- 2) 为转化为无约束优化, 粒子采用对数形式。

综上所述, 本文的总体计算流程如图 2 所示。

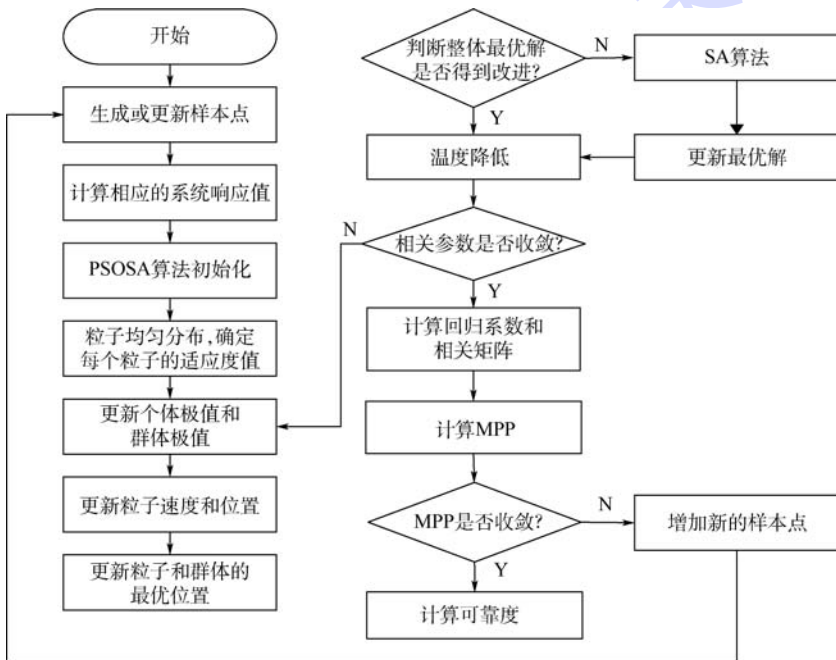


图 2 改进的动态 Kriging 模型流程图

Fig. 2 Flowchart of improved dynamic Kriging model

3 案例分析

3.1 数值算例 1

假设某结构的极限状态函数^[12]为

$$G(\mathbf{x}) = \mathbf{x}_1 \mathbf{x}_2 - 1500 \quad (19)$$

式中: \mathbf{x}_1 和 \mathbf{x}_2 的分布类型为正态分布, $\mathbf{x}_1 \sim N(38, 3.8)$, $\mathbf{x}_2 \sim N(54, 2.7)$, 并且 \mathbf{x}_1 和 \mathbf{x}_2 相互独立, 具体分布参数见表 1。本案例中, $c_1 = 1.5$, $c_2 = 1.5$, $R_T = 0.8$, $w = 2$ 。

表 1 中, β_r 为可靠度系数, β_r 越大, 可靠度越大。

对比 MCS、RSM、经典 Kriging 和 PSO-Kriging^[13] 方法, 由表 1 结果可知, MCS 方法选取样本点 10^8 , 计算失效概率结果为 6.3×10^{-3} ; RSM 方法处理非线性问题有一定的局限性, 计算结果误差较大; 经典 Kriging 方法、PSO-Kriging 方法与本文算法选取同样数量的样本数量, 失效概率与 MCS 方法接近, 但是经典 Kriging 方法与 MCS 方法相比, 误差为 4.76%, PSO-Kriging 方法误差为 4.13%, 本文算法误差为 1.58%。由此看出本文算法精度相比于其他 2 种方法有所提高。

将式(19)转化为标准正态分布变量空间下的功能函数, 即

$$G(u_1, u_2) = G_x(\sigma_{x_1} u_1 + \mu_{x_1}, \sigma_{x_2} u_2 + \mu_{x_2}) = (\sigma_{x_1} u_1 + \mu_{x_1})(\sigma_{x_2} u_2 + \mu_{x_2}) - 1500 \quad (20)$$

式中: u_1 和 u_2 相互独立, 均服从标准正态分布。转换后的极限状态函数 $G(u_1, u_2) = 0$ 为一曲线, $G(u_1, u_2) > 0$ 一侧为安全域, $G(u_1, u_2) < 0$ 一侧为失效域, 如图 3 所示。根据本文所提算法, 首先在给定随机变量空间内, 生成初始设计点, 建立 Kriging 模型, 并通过后续的动态更新机制逐渐增加样本点。初始设计点和拟合过程的样本点如图 4 所示。由图可知, 通过更新机制可以有效选择功能函数附近的样本点, 可以充分利用少量样本点的信息拟合待求功能函数。

表 1 不同方法结果对比 (算例 1)

Table 1 Comparison of results of different methods (Example 1)

方法	样本点	β_r	失效概率/ 10^{-3}
MCS	10^8		6.3
RSM	65	2.3927	8.3
经典 Kriging	40	2.4751	6.7
PSO-Kriging	40	2.4805	6.56
本文	40	2.4893	6.4

3.2 数值算例 2

极限状态函数为 $G(\mathbf{x}) = \mathbf{x}_1^3 + \mathbf{x}_2^3 - 4$, \mathbf{x}_1 和 \mathbf{x}_2 均服从正态分布且相互独立, 此极限状态函数高

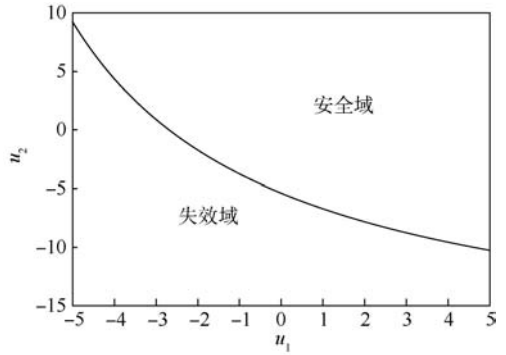


图 3 极限状态函数

Fig. 3 Limit state function

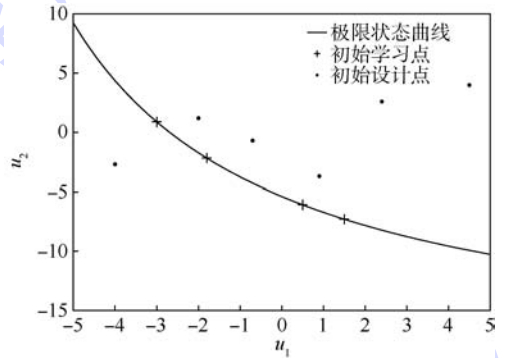


图 4 算法说明图

Fig. 4 Illustration of algorithm

度非线性。 $\mathbf{x}_1 \sim N(3, 1)$, $\mathbf{x}_2 \sim N(2.9, 1)$ 。 $c_1 = 1.2$, $c_2 = 1.2$, $R_T = 0.8$, $w = 2$ 。

将本文计算结果与其他方法进行对比, 结果如表 2 所示。表 2 说明 RSM 在处理高度非线性问题时无法准确收敛和获得正确结果。与其他方法相比, 本文算法的结果与 MCS 结果最接近。其中 MCS 计算时间长达 60.68 s, 本文算法需时 12.18 s。

表 2 不同方法结果对比 (算例 2)

Table 2 Comparison of results of different methods (Example 2)

方法	函数调用次数	β_r	失效概率/ 10^{-3}
MCS	10^6		4.16
RSM	5	1.4729	70.39
经典 Kriging	30	2.6245	4.30
PSO-Kriging	30	2.6307	4.26
本文	30	2.6396	4.15

3.3 工程案例

涡轮盘是现代航空发动机的核心部件, 它在发动机燃烧室内受到高温燃气的推动, 将燃气的热能转化为机械能, 驱动发动机的运转。其可靠性水平的高低直接影响发动机的技术水平。

本文选择某航空发动机涡轮盘为例, 盘上销钉沿周向均匀分布。参数不确定性的主要来源主要包括载荷、材料的随机性。本文取转速均值为

9550 r/min,销轴的材料为 3Cr13,轮盘的材料为 TC11,叶片的载荷施加在销轴上,合力均值为 24925 N.具体参数信息见表 3。

涡轮盘的可靠性水平的高低主要取决于结构危险点处的应力应变分布。由于结构形状、载荷为完全对称的,本文选取轮盘的 1/37 为研究模型。首先以各变量的均值为设置参数,借助 ANSYS 18.1 进行仿真计算,得到结果如图 5 所示,轮盘与销轴交界处应力应变水平最高,是结构的危险点,其中最大应力为 0.834 GPa,最大应变为 0.01017。

当危险点的最大应力超过允许值 $R_{max} = 0.97$ GPa,则结构发生失效,相应的功能函数可表

示为

$$G = R_{max} - S(\omega, E_1, \varepsilon_1, \rho_1, E_2, \varepsilon_2, \rho_2, P) \quad (21)$$

上述功能函数为隐式函数。

采用本文所提算法进行可靠度计算,首先建立 Kriging 模型,其中相关参数 θ^* 的迭代过程见图 6。

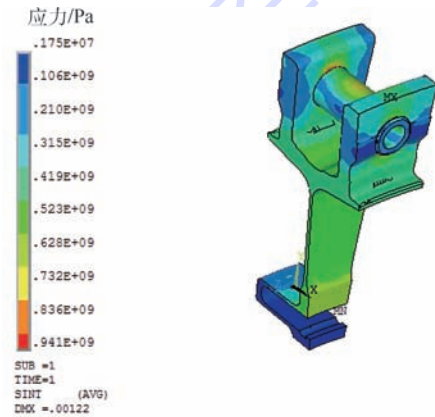
由图 6 可以看出,随着迭代次数的增加,本文所提算法在第 15 次迭代时基本寻得最优相关参数 θ^* 。

利用本文算法与 ANSYS 18.1、Isight 进行联合仿真计算,最终求得可靠度。为了验证本文算法的有效性和工程实用性,将其与 MCS 方法进行对比,结果如表 4 所示。

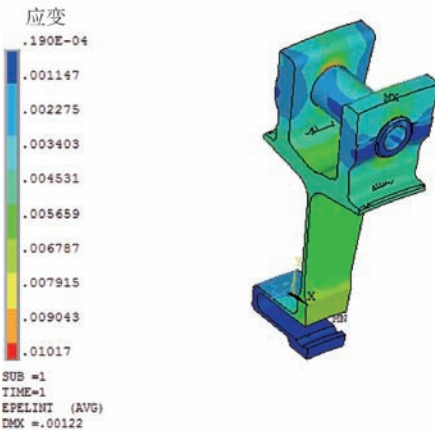
表 3 涡轮盘参数

Table 3 Parameters of turbine disk

参数	涡轮盘的转速 ω	弹性模量 (轮盘) E_1	泊松比 (轮盘) ε_1	密度 (轮盘) ρ_1	弹性模量 (销轴) E_2	泊松比 (销轴) ε_2	密度 (销轴) ρ_2	均布载荷 P
均值	9550 r/min	123 GPa	0.33	4.48 g/cm ³	219 GPa	0.3	7.76 g/cm ³	24925 N
变异系数	0.1	0.015	0.01	0.02	0.015	0.01	0.002	0.1



(a) 应力云图



(b) 应变云图

图 5 应力和应变计算结果

Fig.5 Calculation results of stress and strain

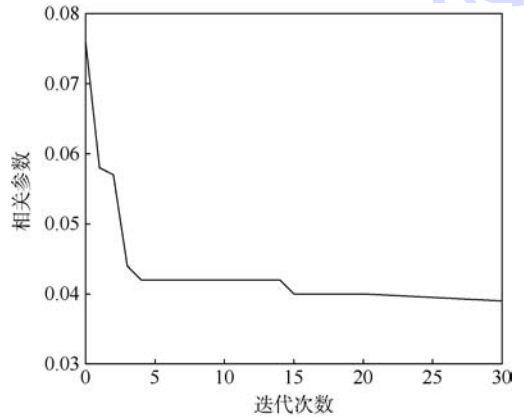


图 6 θ^* 的迭代过程

Fig.6 Iteration process of θ^*

表 4 计算结果对比

Table 4 Comparison of calculation results

方法	函数调用次数	样本点	失效概率/ 10^{-3}
MCS 方法	10^5		3.300
本文算法	15	40	3.352

由表 4 可知,本文算法的结果与 MCS 结果十分接近,误差仅为 1.576%,验证了本文算法的有效性以及准确性,并适用于工程实践。

4 结论

1) 将 PSOSA 算法引入 Kriging 的构造过程, PSO 算法的全局搜索能力及 SA 算法跳出局部极

小值的能力相结合,克服了经典 Kriging 过程的局限性和对初始样本点的依赖,保证了极大似然意义下的最优相关参数的求解精度,使寻优过程更加高效和精确,从而有效保证了 Kriging 模型的预测精度。

2) 本文同时通过动态更新机制,尽可能地减少样本点和调用有限元等数值计算的次数,能够较好地解决功能函数隐式和高度非线性的问题,通过数值案例分析,相比于其他可靠度算法,能够有效提高可靠度计算的效率和精度。通过工程案例分析,本文所提算法可应用于工程实际问题,尤其是功能函数隐式且复杂的问题。有一定的工程实用价值。

参考文献 (References)

- [1] DITLEVSEN O, MADSEN H O. Structural reliability methods [M]. Chichester: John Wiley & Sons Ltd, 2007.
- [2] PELLISSETTI M F, SCHULLER G I. On general purpose software in structural reliability-An overview [J]. Structural Safety, 2006, 28(1-2): 3-16.
- [3] REH S, BELEY J D, MUKHERJEE S, et al. Probabilistic finite element analysis using ANSYS [J]. Steel Construction, 2008, 28(1): 17-13.
- [4] SCHULLER G I, PRADLWARTER H J, KOUTSOURELAKIS P S. A critical appraisal of reliability estimation procedures for high dimensions [J]. Probabilistic Engineering Mechanics, 2004, 19(4): 463-474.
- [5] ZHAO W, QIU Z. An efficient response surface method and its application to structural reliability and reliability-based optimization [J]. Finite Elements in Analysis & Design, 2013, 67(5): 34-42.
- [6] CHERKASSKY V, MA Y. Practical selection of SVM parameters and noise estimation for SVM regression [J]. Neural Networks the Official Journal of the International Neural Network Society, 2004, 17(1): 113-126.
- [7] GOMES H M, AWRUCH A M. Comparison of response surface and neural network with other methods for structural reliability analysis [J]. Structural Safety, 2004, 26(1): 49-67.
- [8] MATHERON G. Principles of geostatistics [J]. Economic Geology, 1963, 58(8): 1246-1266.
- [9] 韩忠华. Kriging 模型及代理优化算法研究进展 [J]. 航空学报, 2016, 37(11): 3197-3225.
HAN Z H. Kriging surrogate model and its application to design optimization; A review of recent progress [J]. Acta Aeronautica et Astronautica Sinica, 2016, 37(11): 3197-3225 (in Chinese).
- [10] ZHANG J, ZHANG L M, TANG W H. Kriging numerical models for geotechnical reliability analysis [J]. Soils and Foundations, 2011, 51(6): 1169-1177.
- [11] LIU H, CUI W M, HE B, et al. The contact reliability analysis of gear rack based on the Kriging and RSM method [C] // Proceedings of the 2015 International Conference on Applied Science and Engineering Innovation. Paris: Atlantis Press, 2015: 279-284.
- [12] 刘瞻, 张建国, 王灿灿, 等. 基于优化 Kriging 模型和重要抽样法的结构可靠度混合算法 [J]. 航空学报, 2013, 34(6): 1347-1355.
LIU Z, ZHANG J G, WANG C C, et al. Hybrid structure reliability method combining optimized Kriging model and importance sampling [J]. Acta Aeronautica et Astronautica Sinica, 2013, 34(6): 1347-1355 (in Chinese).
- [13] 陈志英, 任远, 白广忱, 等. 粒子群优化的 Kriging 近似模型及其在可靠性分析中的应用 [J]. 航空动力学报, 2011, 26(7): 1522-1530.
CHEN Z Y, REN Y, BAI G C, et al. Particle swarm optimized Kriging approximate model and its application to reliability analysis [J]. Journal of Aerospace Power, 2011, 26(7): 1522-1530 (in Chinese).
- [14] MA J G, REN Z Y, ZHAO G X, et al. A new reliability analysis method combining adaptive Kriging with weight index Monte Carlo simulation [J]. IEEE Transactions on Magnetics, 2018, 54(3): 7001904.
- [15] LI X K, QIU H B, CHEN Z Z, et al. A local Kriging approximation method using MPP for reliability-based design optimization [J]. Computers & Structures, 2016, 162: 102-115.
- [16] ECHARD B, GAYTON N, LEMAIRE M. A Kriging improvement of the Monte Carlo simulation; AK-MCS method [M]. Boca Raton: CRC Press-Taylor & Francis Group, 2011: 679-686.
- [17] XIA B, REN Z Y, KOH C S. A novel reliability-based optimal design of electromagnetic devices based on adaptive dynamic Taylor Kriging [J]. IEEE Transactions on Magnetics, 2017, 53(6): 7201504.
- [18] 佟操, 孙志礼, 杨丽, 等. 一种基于 Kriging 和 Monte Carlo 的主动学习可靠度算法 [J]. 航空学报, 2015, 36(9): 2992-3001.
TONG C, SUN Z L, YANG L, et al. An active learning reliability method based on Kriging and Monte Carlo [J]. Acta Aeronautica et Astronautica Sinica, 2015, 36(9): 2992-3001 (in Chinese).

作者简介:

魏娟 女, 硕士研究生。主要研究方向: 机械产品可靠性分析与设计。

张建国 男, 博士, 教授, 博士生导师。主要研究方向: 机械可靠性与安全性。

Structural reliability algorithm based on improved dynamic Kriging model

WEI Juan^{1,2}, ZHANG Jianguo^{1,2,*}, QIU Tao^{1,2}

(1. School of Reliability and Systems Engineering, Beihang University, Beijing 100083, China;

2. Science and Technology on Reliability and Environment Engineering Laboratory, Beihang University, Beijing 100083, China)

Abstract: For complex aerospace machinery products, the limit state functions are often implicit and highly nonlinear, and the reliability calculation usually requires time-consuming finite element analysis. In this paper, the particle swarm optimization-simulated annealing (PSOSA) algorithm is applied to the optimization of the correlation parameters of the dynamic Kriging model, which improves the prediction accuracy. At the same time, with the dynamic update mechanism, sample points are gradually added to reduce the number of function calls as much as possible, thereby improving the calculation efficiency. The algorithm is applied to the structural reliability analysis. The Monte Carlo method, response surface and other classic algorithms are compared, and the results of the proposed algorithm are closer to those of Monte Carlo method, and the calculation time is greatly shortened, which shows that the efficiency and accuracy of the algorithm are improved significantly.

Keywords: limit state function; dynamic update; Kriging model; particle swarm optimization-simulated annealing (PSOSA) algorithm; reliability

Received: 2018-05-28; **Accepted:** 2018-08-24; **Published online:** 2018-09-07 13:54

URL: kns.cnki.net/kcms/detail/11.2625.V.20180906.0944.002.html

Foundation item: National Natural Science Foundation of China (51675026)

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

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

DOI: 10.13700/j.bh.1001-5965.2018.0315

芯片互联结构断裂失效的试验研究与统计分析



陈焱君, 景博*, 胡家兴, 盛增津, 张钰林

(空军工程大学 航空工程学院, 西安 710038)

摘 要: 针对电子芯片互联结构失效机理复杂、寿命数据难以获取且缺乏具有理论依据支撑的寿命分布模型等问题,开展了基于退化试验数据的统计分析研究,并提出了2种统计建模方法。首先,搭建了电子芯片可靠性评估试验台,利用菊花链测试芯片为试验对象获取了互联结构的退化数据和寿命数据;其次,基于断裂力学理论,分析了互联结构断裂失效的失效机理,给出了BGA封装形式的电子芯片互联结构寿命分布服从两参数威布尔分布的理论依据,由此建立了基于寿命数据的电子芯片互联结构寿命分布模型;再次,利用退化量的分布建立了基于退化量数据的电子芯片互联结构寿命分布模型;最后,对2种统计建模方法进行比较分析。结果表明,基于退化量数据的寿命分布模型与基于寿命数据的寿命分布模型吻合度良好,2种分析方法在结果上具有一致性。

关键词: 互联结构; 威布尔分布; 退化量分布; 随机振动; 故障预测与健康管理

中图分类号: TH311

文献标识码: A **文章编号:** 1001-5965(2019)02-0381-07

电子设备在军事及航空航天领域应用越来越广泛,其可靠性要求也不断提高。随着集成电路的设计技术和制造水平的提升,大型电子设备集成度越来越高,功能越来越复杂^[1-2]。同时,电子芯片技术也得到了迅猛的发展,芯片越来越小型化、集成化,其内部的互联结构,即焊点的数目众多且尺寸微小。焊点作为起机械支撑和电气连接的关键部位,一旦某一个焊点产生断裂,则可能导致整个组件的故障或失效,进而产生重大事故。相关资料显示^[3],目前电子设备失效70%的原因是电子封装失效,而电子封装失效最主要的表现形式就是互联结构失效,即焊点失效。因此,互联结构可靠性研究得到了越来越广泛的关注。

大部分电子设备,尤其是机载或车载电子设备经常处于各种振动的服役环境中,振动成了影响焊点可靠性的重要环境因素。统计数据表

明^[4-5],振动载荷下焊点的失效率高达20%,因此针对电子封装的微互连焊点在振动条件下的可靠性研究具有重要的意义。

在往复振动载荷下,电子组件基板和印刷电路板(Printed Circuit Board, PCB)在振动载荷下产生翘曲效应,互联结构受到交变应力,其内部应力集中的部位会首先出现微小的损伤并不断累积,当前,确定互联结构损伤的方法主要基于电参数和应变参数。Kwon等^[6]研究发现焊点发生损伤的时刻电阻信号会产生剧烈变化;Tang等^[7]依据试验监测的电阻信号,提出振动载荷下焊点的失效过程可以分为健康状态、轻度失效、中度失效及彻底失效;陈焱君等^[8]依据BGA(Ball Grid Array)焊点在正弦振动下的电阻数据,建立了单个BGA焊点的退化模型;王文等^[9]根据无铅BGA焊点在随机振动下的全寿命电阻监测数据,利用比例风

收稿日期: 2018-06-01; 录用日期: 2018-09-03; 网络出版时间: 2018-09-10 16:37

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180907.1025.003.html

基金项目: “十三五”装备预研共用技术(41402010102)

* 通信作者: E-mail: jingbo_sensor@163.com

引用格式: 陈焱君, 景博, 胡家兴, 等. 芯片互联结构断裂失效的试验研究与统计分析[J]. 北京航空航天大学学报, 2019, 45(2): 381-387. CHEN Y J, JING B, HU J X, et al. Experimental research and statistical analysis of fracture failure for interconnected structures in electronic chips[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 381-387 (in Chinese).

险模型对焊点剩余寿命进行了预测。

上述研究虽然进行了特征提取、退化建模、寿命预测的系统研究,但主要针对单个焊点开展。然而,对整个芯片而言,是多焊点组成的复杂系统,芯片内不同位置的焊点受到的应力大不相同,因此不能用简单地针对单焊点的研究方法来研究整个芯片焊点的可靠性。焊点全寿命周期内的电参数数据庞大繁杂、失效时间长、难以获得其寿命数据,为了获取较多寿命数据以建立更精确的寿命分布模型需要进行大量试验获得更多样本,时间耗费长、经济性差,而针对退化量分布建立寿命分布模型却能较好的克服这些缺点。退化量分布建模的关键是如何确定退化量分布的类型,余荣斌等^[10]用 β 分布作为退化量分布,进行了光伏组件的退化建模;王立等^[11]结合退化量分布参数估计的平稳时序类型给出了基于退化量分布的寿命预测方法。

鉴于此,本文选取 BGA 封装的菊花链测试芯片为试验对象,搭建了电子芯片可靠性试验台,获取多组随机振动条件下电子芯片全寿命周期内的电阻退化数据寿命的时间数据;依据断裂力学理论,推导了 BGA 封装芯片服从两参数威布尔分布的理论依据,建立了基于寿命数据的电子芯片寿命分布函数;接着利用退化量分布建立了基于退化量分布的寿命分布函数。通过比较发现,2 种统计建模方法结果具有一致性,都可用于具体的工程实践,从而为实现电子芯片剩余使用寿命评估提供重要依据。

1 随机振动载荷下电子芯片环境应力试验

1.1 试验件设计

试验选用具有菊花链连接的测试芯片作为被测对象,其内部所有的 BGA 焊球都被串联起来形成了一个完整的回路,如图 1 所示。

这样,只要有一个焊球产生断裂,整个芯片都会对外表现出失效。测试芯片的尺寸为 $8\text{ mm} \times 8\text{ mm}$,厚度为 0.97 mm ,焊球直径 0.4 mm ,焊球间距 0.8 mm ,通过钎料锡铅焊料焊接在一块尺寸为 $180\text{ mm} \times 90\text{ mm}$,厚度为 0.7 mm 的 PCB 板上,且将测试芯片置于 PCB 板中央,如图 2 所示,能确保对称位置的焊点承受一致的应力载荷。

同时,为了找到试验件的最初共振点和振动特性,采用正弦振动扫频试验法。首先将应变片置于试验件中央靠近芯片处,设置振动台采用正弦扫频振动方式,频率在 $20 \sim 1\,000\text{ Hz}$,功率谱密

度设置为 $0.5\text{ g}^2/\text{Hz}$ (g 为重力加速度),得到的应变响应经过傅里叶变换后结果如图 3 所示(只列举了前三阶固有频率)可以发现,试验件的一阶固有频率在 300 Hz 附近。

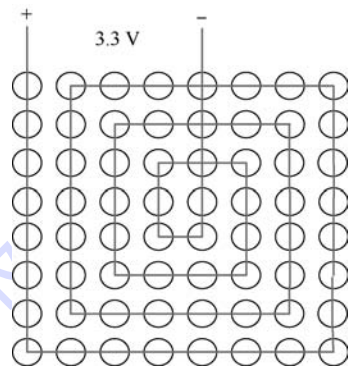


图 1 菊花链测试芯片互联结构连接图
Fig. 1 Connection diagram of interconnected structure of daisy-chain test chip

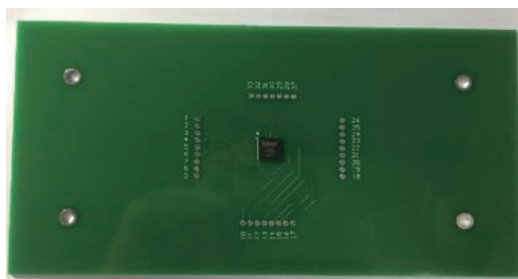


图 2 随机振动试验件
Fig. 2 Test piece for random vibration

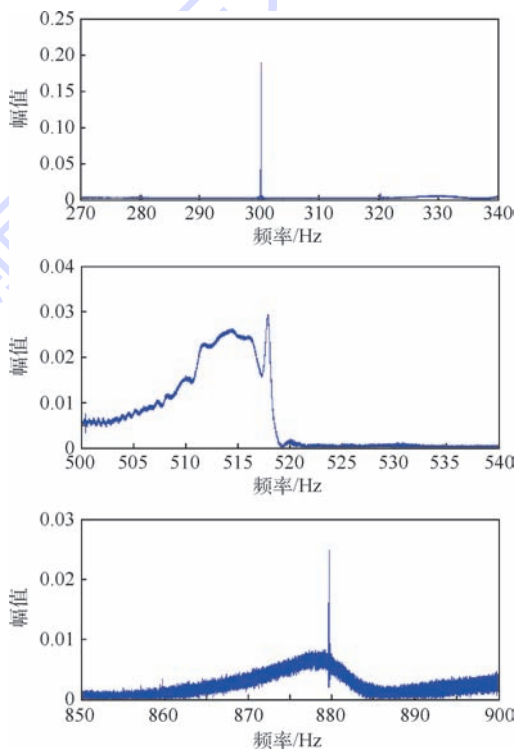


图 3 试验件的振动特性

Fig. 3 Vibration characteristics of test piece

1.2 随机振动试验方法

1.2.1 随机振动载荷谱的确定

机载平台和装备用途的不同,会导致电子装备处在不同的振动环境中,因此,对于不同用途的电子设备,在试验中选择的振动量级也不尽相同。以喷气式飞机机载航电设备中仪表设备为例,发动机噪声对飞机结构的激励、飞机外部结构上的气动湍流、对外部气流敞开的空舱中气动湍流共振以及机动飞行、滑行等都会使电子设备处于振动环境中。其中,仪表板及机身、发动机舱等区域的振动环境均为随机振动^[12],因此本文选择GJB 150.16A—2009^[13]提供的仪表设备参考振动谱型来设置振动试验的载荷谱。其中,加速度功率谱密度 w_0 一般取值为0.01、0.02、0.04、0.06、0.08 g^2/Hz ,但这些功率谱密度值较小,得到完整的全寿命退化数据会耗费大量的时间,因此本文以 $w_0 = 0.8 g^2/\text{Hz}$ 设计加速度功率谱进行试验。同时该标准还指出,仪表设备处于30~1000 Hz的振动频率之间,为了达到最大振动激励并同时更真实的模拟仪表设备服役环境,将载荷谱设置为一阶固有频率处,即290~310 Hz。功率谱密度值 $\pm 3 \text{ dB}$ 设置警戒线, $\pm 6 \text{ dB}$ 设置中断线,当试验振动参数出现异常,会发出警告直至自行中断试验,确保整个试验过程中振动参数控制在规定值 $\pm 3 \text{ dB}$ 以内。

1.2.2 随机振动试验步骤

在菊花链芯片的2个端点引出测试线,接入串联分压电路,菊花链两端的分压便能够反映菊花链内各个焊球的退化状态,这是因为,焊球内的裂纹会随着振动逐渐增大,增大的裂纹导致焊球电阻增大,进而菊花链两端的电压升高。并且只要有一个焊点出现失效,那么整个回路都会被置于高电平,从而反映出芯片的失效。将测试线从芯片引出至PCB板并利用ECONMI-7016型16通道数据采集仪采集测试芯片两端的电压退化数据。

2 电子芯片焊点失效机理及统计建模

2.1 电子芯片焊点失效机理

通过随机振动试验,可以得到连接成菊花链的芯片内互联结构由健康到完全失效的等效电阻分压的动态趋势,以其中的一个试验件的退化过程为例,如图4所示。给定的供电电压为3.3 V,当菊花链测试芯片内的互联结构都完好时,等效电阻很小,分压接近于0 V,当菊花链内任意一个焊球产生裂纹,整个回路的等效电阻增大,分压随

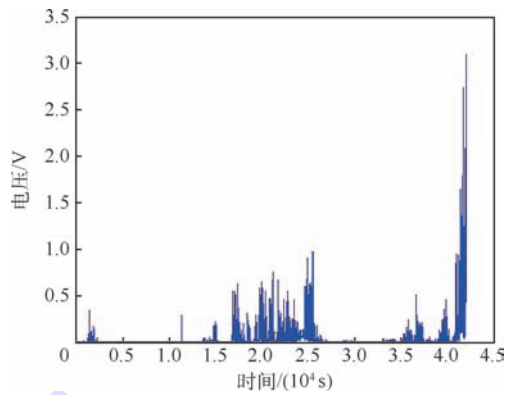


图4 菊花链测试芯片互联结构的退化过程

Fig. 4 Degradation process of interconnected structure of daisy-chain test chip

之增大,当有任意一个焊点完全失效时,整个菊花链回路变为开路状态,电阻变为无穷大,分压达到3.3 V,电子芯片完全失效。

由图4还可以看出,电压测值虽然已经达到3.3 V,但却出现了高低反复振荡,而不是稳定在较高水平,这反映了芯片内失效焊点随着振动出现了反复开合的现象,即焊点在拉力作用下出现开裂、导电面积减小、电阻增大、分压增大的现象;当振动台向反方向运动时,焊点受压,出现了闭合、导电面积增大、电阻减小、分压减小的现象。

2.2 基于寿命数据的电子芯片失效统计建模

随着振动时间的增加,焊球内裂纹的长度也在逐渐增大,令 δ_i 为第 i 次振动循环后的裂纹长度,则一定有 $\delta_1 < \delta_2 < \dots < \delta_n$ 。当裂纹长度达到 δ_m 时,焊点失效,假设 δ_0 为初始裂纹,其存在且极其微小可视为0。则在第 j 个循环下,焊球内部裂纹的增量为 $\delta_j - \delta_{j-1}$ 。依据断裂力学中裂纹长度计算公式Paris公式:

$$\frac{d\alpha}{dN} = c(\Delta k) \quad (1)$$

式中: k 为应力强度因子, Δk 为 k 的增量; α 为裂纹长度; N 为振动循环次数; c 为常数。根据美国国家航空航天局(NASA)^[14]给出的圆截面在拉压弯曲组合作用下圆弧裂纹扩展模型的应力强度因子为

$$K_0 = F_0 \sigma_0 \sqrt{\pi a} \quad (2)$$

式中: a 为裂纹的深度; $F_0 = G(0.752 + 1.286\beta + 0.37Y^3)$, $G = 0.92(2/\pi) \sec \beta((\tan \beta)/\beta)^{1/2}$, $Y = 1 - \sin \beta$, $\beta = \pi a/(4R)$, R 为焊球半径; σ_0 为受到的正应力。

将式(2)代入式(1),可知应力强度因子正比于裂纹长度,即裂纹的增量正比于裂纹长度,因此有理由假定裂纹扩展按照比例效应增长,即

$\frac{\delta_j - \delta_{j-1}}{\delta_{j-1}} = \chi_j$ 。则有

$$\sum_{j=1}^m \chi_j = \sum_{j=1}^m \frac{\delta_j - \delta_{j-1}}{\delta_{j-1}} \quad (3)$$

同时,当每次振动循环下裂纹的增量是微量的情况下有

$$\Delta\delta_j = \delta_j - \delta_{j-1} \rightarrow 0 \quad (4)$$

那么

$$\sum_{j=1}^m \chi_j = \sum_{j=1}^m \frac{\Delta\delta_j}{\delta_{j-1}} \quad (5)$$

当振动循环充分大时,依据微积分定义可知

$$\sum_{j=1}^m \chi_j = \int_{\delta_0}^{\delta_m} \frac{1}{\delta} d\delta = \ln \delta_m - \ln \delta_0 \quad (6)$$

移项得

$$\ln \delta_m = \sum_{j=1}^m \chi_j + \ln \delta_0 \quad (7)$$

由中心极限定理可得 $\sum_{j=1}^m \chi_j$ 的渐进分布为正态分布,由于 $\ln \delta_0$ 为一常量,那么 $\ln \delta_m$ 也服从正态分布,即 δ_m 渐进服从对数正态分布。

假设一个芯片有 n 个焊点,第 i 个焊点的寿命为 $T_i (i=1, 2, \dots, n)$,则每个焊点的寿命分布即为 $F_i(t)$ 。对于一个具有完整功能的电子芯片而言,每一个引脚都需要实现特定的功能,因此如果假设只要有一个焊点出现失效,那么整个芯片就会失效。设芯片的寿命分布为 $F(t)$,则芯片的寿命分布就是所有焊点寿命的极小值分布。即令 $N = \min(T_1, T_2, \dots, T_n)$,令焊点的失效阈值为 z ,则有

$$F(z) = P\{N \leq z\} = 1 - P(N > z) \quad (8)$$

每个焊点的失效状态互不影响,故可以认为 T_i 相互独立,则有

$$F(z) = 1 - (1 - F_1(z))(1 - F_2(z)) \dots (1 - F_n(z)) \quad (9)$$

另外,在振动载荷下,芯片不同位置的焊球所受到的应力不是均匀的,文献[15]表明,位于芯片边角处的焊点受应力更为集中,会优先失效,故 $F_i(t)$ 虽然均服从对数正态分布,但参数却会随着焊球的位置变化而有所不同。定义位置参数 P ,对不同位置焊球的尺度参数进行归一化,则不同位置焊球的寿命分布 $F_i(t)$ 可以归一化为 $F_p(t)$ 。因此有

$$F(z) = 1 - (1 - F_1(z))(1 - F_2(z)) \dots (1 - F_n(z)) \quad (10)$$

可化为

$$F(z) = 1 - (1 - F_p(z))^n \quad (11)$$

文献[16]表明,当 $F_p(t)$ 服从对数正态分布

且 n 充分大时, $F(z) = 1 - (1 - F_p(z))^n$ 的渐进分布为两参数威布尔分布,其概率密度函数为

$$f(t) = \frac{m}{\eta} \left(\frac{t}{\eta}\right)^{m-1} \exp\left[-\left(\frac{t}{\eta}\right)^m\right] \quad (12)$$

式中: m 为形状参数; η 为尺度参数。由此给出了 BGA 封装形式的电子芯片互联结构失效寿命服从威布尔分布的理论依据。

3 基于退化量分布的电子芯片失效统计建模

3.1 基于退化量分布的寿命分布模型

如果把电子芯片的互联结构的退化看作一随时间变化的随机过程 $x(t)$,同一批产品虽然总体来说具有同一种退化规律,但是对于不同的个体而言,在 t 时刻的退化量 $x(t)$ 又会具有较大的差异,但如果在某一时刻 t 确定了 $x(t)$ 在不同个体中的形式 $G(x, t)$,并且能通过试验数据统计出 $G(x, t)$ 随时间变化的规律,就确定了 $G(x, t)$,从而能推导出寿命的累积分布函数 $F(t)$ 。

由于产生了裂纹的 BGA 焊点在振动过程中出现的开开合合的间歇性失效现象,为了增加数据的单调性,消除间歇性失效的影响,对数据进行累加生成,即由原始的退化量数据 x_1, x_2, \dots, x_n 变化为 $x_1, x_1 + x_2, \dots, x_1 + x_2 + \dots + x_n$ 。收集试验产品在 t_1, t_2, \dots, t_m 时刻的性能退化数据,依据中心极限定理,可以导出,经过逐步累加的退化数据在 $t_i (i=1, 2, \dots, M)$ 时刻不同样品的退化量的渐进分布为均值为 $\mu(t)$,标准差为 $\sigma(t)$ 的正态分布。

对它们进行标准化可以得到 $Y = \frac{x(t) - \mu(t)}{\sigma(t)}$

服从标准正态分布。当失效阈值为 z 时

$$F(t) = P(x(t) \geq z) = 1 - P(x(t) \leq z) \quad (13)$$

有

$$F(t) = 1 - P\left(Y \leq \frac{z - \mu(t)}{\sigma(t)}\right) \quad (14)$$

可得

$$F(t) = 1 - \Phi\left(\frac{z - \mu(t)}{\sigma(t)}\right) \quad (15)$$

3.2 基于退化量分布的电子芯片焊点失效数据统计建模

将得到的原始电压数据进行累加生成,得到的退化轨迹如图 5 所示(以 5 个样本中的某一样本为例),累计退化量变成了单调增函数。

取时间 t 的初始值为 5 000 s,每间隔 5 000 s 取 5 个试验样本对应时间下的累计退化量,为一

组,共取20组,取它们的累加退化量的均值 μ 和标准差 σ ,其随时间的分布具有较好的线性度,如图6所示。

对均值和标准差进行最小二乘拟合,得到了均值随和标准差随时间的表达式分别为

$$\mu(t) = 0.02093t + 60.65 \quad (16)$$

$$\sigma(t) = 0.05859t + 169.8 \quad (17)$$

选择完全断裂时芯片两端的分压3.3V为失效阈值,对5组试验数据失效的首达时间下的累计退化量做平均值作为累计退化量的经验阈值,经计算得到阈值为3749V,将该经验值和式(16)、式(17)代入式(15)得到了芯片的寿命的累积分布函数。

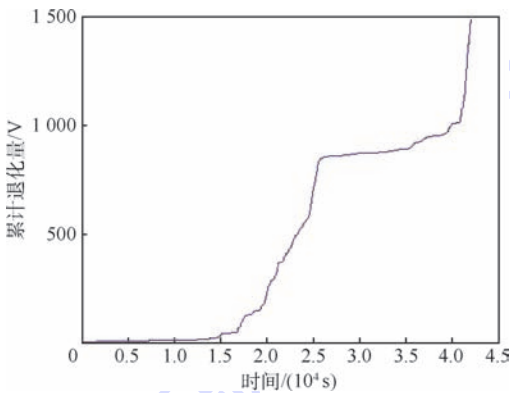
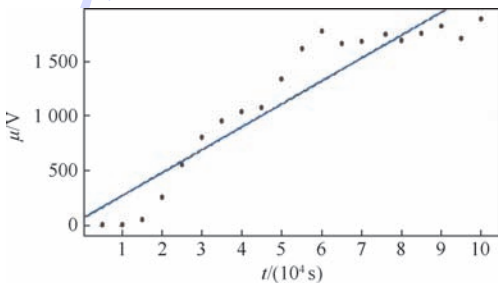
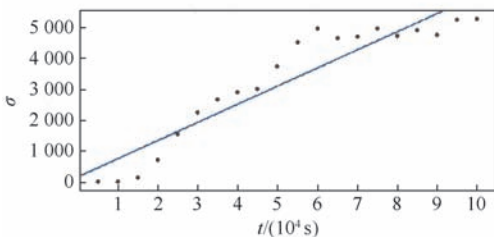


图5 菊花链测试芯片互联结构累计退化量

Fig.5 Accumulated degradation of interconnected structure of daisy-chain test chip



(a) 均值随时间变化



(b) 标准差随时间变化

图6 互联结构累计退化量的均值和标准差随时间变化

Fig.6 Variation of mean value and standard deviation of accumulated degradation of interconnected structure with time

$$F(t) = 1 - \Phi\left(\frac{3749 - (0.02093t + 60.65)}{0.05859t + 169.8}\right) \quad (18)$$

同时,试验件的5组累计失效时间如表1所示。

表1 各组试验件的累计失效时间

Table 1 Accumulative failure time for each group of test piece

试验组	累积失效时间/s
1	60627
2	118618
3	42017
4	165376
5	26115

3.3 对比分析

利用极大似然估计得到了二参数威布尔分布的形状参数和寿命参数如表2所示。

基于退化量数据的电子芯片寿命的累积分布函数和基于寿命数据的电子芯片寿命的累积分布函数如图7所示。

对比分析发现,2种寿命分布建模方法得到的累积分布函数具有较高一致性,基于退化量数据进行统计建模可以大幅缩短试验时间,但因需要引入经验值导致精度上会有所缺失。基于寿命数据的统计建模方法具有较高的精度,但因失效时间长而导致寿命数据难以获取。

表2 两参数威布尔分布的参数估计

Table 2 Parameter estimation of two-parameter Weibull distribution

参数	数值
形状参数 m	1.667
尺度参数 η	93002

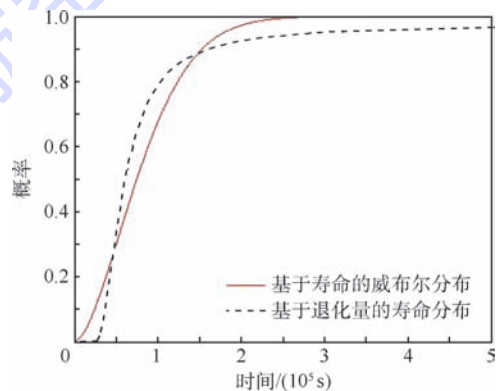


图7 菊花链测试芯片互联结构的寿命分布

Fig.7 Lifetime distribution of interconnected structure of daisy-chain test chip

4 结论

本文针对电子芯片互联结构失效机理复杂、

寿命数据难以获取且缺乏具有理论依据支撑的寿命分布模型等问题,开展了电子芯片互联结构的可靠性试验。结合断裂力学理论,推导了电子芯片互联结构寿命分布服从两参数威布尔分布的理论依据,并建立了基于寿命数据的电子芯片失效的统计模型。利用退化数据,建立了基于退化数据的电子芯片失效统计模型。2种模型结果均以电子芯片寿命的累计分布函数给出。由此得到了以下结论:

1) 电子芯片互联结构的寿命分布的渐进分布服从两参数威布尔分布,利用其分布能够建立基于寿命数据的电子芯片失效统计模型。

2) 基于退化数据的电子芯片统计模型与基于寿命数据的电子芯片统计模型结果互相吻合,具有较高的一致性。

3) 基于退化量数据进行统计建模可以大幅缩短试验时间,但因需要引入经验值导致精度上会有所缺失。基于寿命数据的统计建模方法具有较高的精度,但因寿命数据难以获取会给建模带来困难。

参考文献 (References)

- [1] 胡家兴,景博,汤巍,等. 电子设备无铅焊点的热疲劳评估进展与展望[J]. 空军工程大学学报(自然科学版),2016,17(6):7-12.
- HU J X, JING B, TANG W, et al. Evaluation and prospect of thermal fatigue on lead-free solder joints in electronic equipment[J]. Journal of Air Force Engineering University(Natural Science Edition),2016,17(6):7-12(in Chinese).
- [2] 周德俭. 电子制造中的电器互连技术[M]. 北京:电子工业出版社,2010:1-5.
- ZHOU D J. Manufacturing technology for electronics[M]. Beijing:Electronic Industry Press,2010:1-5(in Chinese).
- [3] GUO Q, ZHAO M, WANG H F. SMT solder joint's semi-experimental fatigue model[J]. Mechanics Research Communications,2005,32(3):351-358.
- [4] 胡家兴,景博,汤巍,等. 无铅微焊点的热效应仿真及可靠性分析[J]. 电子元件与材料,2016,35(3):81-84.
- HU J X, JING B, TANG W, et al. Thermo effect simulation and reliability analysis of lead-free micro solder joints[J]. Electronic Components and Materials,2016,32(3):81-84(in Chinese).
- [5] PANG J H L, WONG F L, HENG K T, et al. Combined vibration and thermal cycling fatigue analysis for SAC305 lead free solder assemblies[C]//2013 IEEE 63rd Electronic Components and Technology Conference. Piscataway, NJ: IEEE Press, 2013:1300-1307.
- [6] KWON D, AZARIAN M H, PECHT M. Remaining-life prediction of solder joints using RF impedance analysis and Gaussian process regression[J]. IEEE Transactions on Components Packaging & Manufacturing Technology,2015,5(11):1602-1609.
- [7] TANG W, JING B, HUANG Y F, et al. Multistate degradation model for prognostics of solder joints under vibration conditions[J]. Chinese Journal of Electronics,2016,25(4):779-785.
- [8] 陈垚君,景博,常雅男,等. 振动载荷下板级焊点退化状态表征与分析[J]. 微电子学,2017,47(6):866-871.
- CHEN Y J, JING B, CHANG Y N, et al. Failure characterization and analysis of board level solder interconnects under vibration load[J]. Microelectronics,2017,47(6):866-871(in Chinese).
- [9] 王文,孟光,刘芳,等. 基于比例风险模型的板级无铅焊点跌落寿命分析[J]. 振动与冲击,2011,30(3):124-128.
- WANG W, MENG G, LIU F, et al. Lifetime analysis of lead-free solder joints under drop impact using proportional hazards model[J]. Journal of Vibration and Shock,2011,30(3):124-128(in Chinese).
- [10] 余荣斌,刘桂雄,徐欢. 基于 β 分布统计法的光伏组件性能退化可靠度估算[J]. 仪器仪表学报,2015,36(11):2586-2593.
- YU R B, LIU G X, XU H. β distribution uniform expression based photovoltaic modules reliability evaluation with degradation data distribution[J]. Chinese Journal of Scientific Instrument,2015,36(11):2586-2593(in Chinese).
- [11] 王立,李晓阳,姜同敏. 基于退化量分布时序分析的产品寿命预测[J]. 北京航空航天大学学报,2011,37(4):492-498.
- WANG L, LI X Y, JIANG T M. Life prediction of product based on degradation amount distribution time series analysis[J]. Journal of Beijing University of Aeronautics and Astronautics,2011,37(4):492-498(in Chinese).
- [12] 祝耀昌,李韵,徐俊,等. 民用飞机机载设备振动实验要求和应用分析(一)[J]. 装备环境工程,2016,13(4):54-61.
- ZHU Y C, LI Y, XU J, et al. Vibration test requirements and application analysis of civil aircraft equipment(1)[J]. Equipment Environmental Engineering,2016,13(4):54-61(in Chinese).
- [13] 军用装备实验室环境试验方法:第16部分:振动实验:GJB 150.16A—2009[S]. 北京:中国人民解放军总装备部. Military equipment laboratory environmental methods: Part 16 vibration experiment: GJB 150.16A—2009[S]. Beijing: PLA General Armament Department(in Chinese).
- [14] FORMAN R G, SHIVAKUMAR V. Growth behavior of surface crack in the circumferential plane of solid and hollow cylinders[M]//UNDERWOOD J, CHAIT R, WILHEM D, et al. Fracture Mechanics: Seventeenth Volume. West Conshohoken: ASTM International,1986:59-74.
- [15] KIM Y K, HWANG D S. PBGA packaging reliability assessments under random vibrations for space applications[J]. Microelectronics Reliability,2015,55(1):172-179.
- [16] 钱萍. 航天电连接器综合应力加速寿命试验与统计分析的研究[D]. 杭州:浙江大学,2009.
- QIAN P. Research on multiple stresses accelerated life test and statistical analysis of aerospace electronic connector[D]. Hangzhou: Zhejiang University,2009(in Chinese).

作者简介:

陈焱君 男,硕士研究生。主要研究方向:电子封装可靠性。

景博 女,博士,教授,博士生导师。主要研究方向:故障诊断与健康管理、测试性设计、传感器网络与信息融合。

Experimental research and statistical analysis of fracture failure for interconnected structures in electronic chips

CHEN Yaojun, JING Bo*, HU Jiaying, SHENG Zengjin, ZHANG Yulin

(School of Aeronautic Engineering, Airforce Engineering University, Xi'an 710038, China)

Abstract: Aimed at the problems of complex failure mechanism of interconnected structures in electronic chips, difficulty in obtaining lifetime data, and lack of a theoretically supported life distribution model, a statistical analysis based on degraded experimental data was conducted. First, an electronic chip reliability evaluation testbed is built, and the daisy-chain test chip is used as the test object. Degradation data and lifetime data of interconnected structures are acquired. Then, based on the theory of fracture mechanics, the failure mechanism of interconnection structure is analyzed. The theoretical basis of the two-parameter Weibull distribution of the life distribution of the interconnection structure of electronic chips in BGA package is given. The life distribution model of the interconnection structure of electronic chips based on the life data is established. Next, using the distribution of degradation, a lifetime distribution model of the interconnected structures in electronic chips is established. Finally, the two statistical modeling methods are compared. The results show that the lifetime distribution model based on degradation data is in good agreement with that based on lifetime data.

Keywords: interconnected structures; Weibull distribution; distribution of degradation data; random vibration; failure prediction and health management

Received: 2018-06-01; **Accepted:** 2018-09-03; **Published online:** 2018-09-10 16:37

URL: kns.cnki.net/kcms/detail/11.2625.V.20180907.1025.003.html

Foundation item: National Defence Pre-research Foundation of China (41402010102)

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

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

DOI: 10.13700/j.bh.1001-5965.2018.0282

电磁航天器编队悬停鲁棒协同控制方法

张亚博¹, 师鹏^{1,*}, 张皓², 赵育善¹

(1. 北京航空航天大学 宇航学院, 北京 100083; 2. 中国科学院空间应用工程与技术中心
中国科学院太空应用重点实验室, 北京 100094)

摘 要: 针对电磁航天器编队近地轨道悬停问题, 提出一种在缺少参考轨道准确信息时的协同控制方法。用 TH 方程描述航天器间的相对运动, 选择与参考轨道同周期的圆轨道为标称轨道。将参考轨道相对于标称圆轨道的偏差、地球非球形引力、大气阻力及其他天体引力等参数单独归类, 视其为不确定量, 构成不确定系统。通过引入一致性理论, 在电磁作用模型和动力学方程均存在不确定性的条件下, 针对航天器编队悬停的目标设计了鲁棒协同控制律。考虑能量消耗最优和均衡以及轨道姿态解耦, 给出了通过优化进行磁矩配置的方案。仿真结果表明, 所设计的鲁棒协同控制律能够实现编队电磁航天器高精度悬停, 所给出的磁矩配置方案能够实现磁矩的合理分配。

关键词: 电磁航天器; 编队飞行; 一致性协议; 鲁棒控制; 悬停

中图分类号: V412.4

文献标识码: A **文章编号:** 1001-5965(2019)02-0388-10

编队悬停是指空间中的多个伴随航天器在控制力或自然力的作用下, 相对参考航天器或其他自然天体保持相对静止的状态。编队悬停技术在航天器在轨服务、深空探测、航天器跟踪监视等领域有着广泛的应用^[1-3]。

航天器悬停编队的实现依赖于准确的轨道和姿态控制。在近距离悬停时, 传统基于冲量原理的推力器会造成羽流污染, 某些任务甚至会严格限制目标与悬停航天器连线方向的点火^[4]; 再者编队中成员航天器悬停轨道通常具有强非开普勒性, 要求推力器长时间工作^[5], 必然导致传统航天器编队寿命有限。因此利用库仑力^[6]、洛伦兹力^[7]和电磁力^[8]等非接触力实现航天器无工质编队的控制受到了众多学者的关注。相比于星间库仑力、洛伦兹力, 通电电磁线圈间能够产生任意

方向的电磁力, 可以实现更多的控制自由度, 本文选择星间电磁相互作用来实现航天器编队的控制。电磁航天器通过改变其上安装的 3 个正交电磁线圈的电流以改变航天器受到的电磁作用, 进而实现航天器编队的相对轨道控制^[9]。

通过引入电磁相互作用之后较好地解决了传统推力器存在的问题, 但电磁作用固有的非线性、相互性以及空间环境的不确定干扰决定了电磁航天器编队控制具有强非线性、强耦合性以及不确定性的特点^[10]。此外, 电磁力是系统内力, 无法实现编队质心的机动。这些因素对控制的设计提出了更高的要求。

针对电磁航天器编队的动力学与控制问题, 国内外学者进行了丰富的研究。就双电磁航天器编队控制, Elias 等^[11]引入双电磁航天器编队动

收稿日期: 2018-05-18; 录用日期: 2018-07-05; 网络出版时间: 2018-07-19 17:10

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180718.1806.001.html

基金项目: 国家自然科学基金(11572019); 中国科学院太空应用重点实验室开放基金(LSU-2016-07-02); 中国科学院百人计划

*通信作者. E-mail: shipeng@buaa.edu.cn

引用格式: 张亚博, 师鹏, 张皓, 等. 电磁航天器编队悬停鲁棒协同控制方法[J]. 北京航空航天大学学报, 2019, 45(2): 388-397.
ZHANG Y B, SHI P, ZHANG H, et al. A robust coordinated control method for hovering of electromagnetic spacecraft formation [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 388-397 (in Chinese).

力学非线性模型并对其进行了线化,分析了线性模型的闭环稳定性和可控性,设计了最优线性反馈控制律。黄涣等^[12]基于 Kane 方法建立了电磁航天器六自由度耦合非线性动力学模型,分析了 3 种平衡态构型的控制需求,并设计了线性二次型调节器(LQR)控制器。徐增文等^[13]针对双星电磁航天器构型保持设计了自适应控制律,并给出了双星磁矩配置解析表达式。邵龙飞等^[14]将双星编队运动规划问题转化为标准优化问题进行了求解,并指出多电磁航天器的编队控制可以转化为分阶段的双电磁航天器控制。Kwon 等^[15]用双星地面实验验证了利用超导线圈产生电磁力进而进行位置保持和轨道跟踪的可行性。

对于多电磁航天器编队控制,Ahsun 和 Miller^[16]给出了编队相对运动二维动力学模型,并分别基于人工势函数法和最优控制方法研究了电磁航天器编队的构型保持问题。Miller 等^[17]基于近地轨道动力学模型,采用非线性自适应控制方法研究了电磁航天器编队的构型保持和重构问题,并提出了通过优化配置磁矩的方案。胡敏等^[18]分析了地磁场对近地轨道电磁编队的影响,建立了高精度相对运动动力学模型,在此基础上提出了一种非线性反馈控制方法。Zeng 和 Hu^[19]在动力学模型存在不确定性和外部摄动的情况下,设计了电磁航天器编队轨道跟踪的线性终端滑模控制。Zhang 等^[20]建立了六自由度电磁航天器动力学模型,给出了航天器轨道姿态控制的自适应终端滑模控制。

尽管以上研究较好地解决了电磁航天器编队的相对构型保持和轨道姿态跟踪问题,但文献中采用的动力学模型大多数只考虑了外部摄动的不确定性,而未考虑动力学模型的内部不确定性和参数摄动,在实际的任务中该参数摄动是不可避免的。再者,当参考轨道是椭圆时,系统呈现动态时变特性,文献中设计的控制律均需要参考轨道实时的准确信息,获取这些信息一些情况下是十分困难甚至是不可能的。此外,文献中设计的控制律并未考虑编队成员航天器之间的信息交互,导致编队的协同能力较差并产生冗余控制。

考虑到上述问题,本文针对电磁航天器编队近地轨道悬停控制,提出了一种在缺少参考轨道准确信息时的协同控制方法。首先,用 TH 方程描述航天器间的相对运动,选择与参考轨道同周期的圆轨道作为标称轨道,将参考轨道相对于标称轨道的偏差、地球非球形引力、大气阻力以及其他天体引力等参数单独归类,视其为不确定量,构

成不确定系统。然后,针对动力学方程存在内部和外部不确定性,通过引入滑模控制理论和一致性理论,就电磁航天器编队悬停的目标设计了鲁棒协同控制律。最后,考虑到能量消耗最优和均衡以及轨道姿态解耦的问题,给出了通过优化进行磁矩配置的方案,并进行了数值仿真验证。

1 电磁航天器编队动力学建模

1.1 电磁远场作用模型

电磁航天器通过其上安装的 3 个正交电磁线圈产生电磁相互作用,如图 1 所示。图中 i_1 、 i_2 为线圈电流; a_1 、 a_2 为线圈半径; dl_1 、 dl_2 为线圈长度微元; d 为线圈微元间的距离; s 为线圈 2 的微元于线圈 1 的中心的距离。可以采用毕奥-萨伐尔定律精确计算 2 个通电线圈之间的电磁力和电磁力矩,设线圈 2 位于线圈 1 的电磁场中,对线圈 2 上的电流微元在线圈 1 的电磁场中进行积分,即可求得线圈 2 受到线圈 1 的电磁力 F_2 和力矩 T_2 分别为

$$F_2 = \frac{\mu_0 i_1 i_2}{4\pi} \oint \left(\oint \frac{d \times dl_1}{d^2} \right) \times dl_2 \quad (1)$$

$$T_2 = \frac{\mu_0 i_1 i_2}{4\pi} \oint a_2 \times \left(\oint \left(\frac{d \times dl_1}{d^2} \right) \times dl_2 \right) \quad (2)$$

式中: $\mu_0 = 4\pi \times 10^{-7} \text{N/A}^2$ 为真空磁导率。

式(1)、式(2)形式复杂,不易在工程中直接应用。研究发现^[21],2 个线圈之间的距离大于 5 倍的线圈半径时,将线圈视为偶极子得到的远场电磁力模型误差在 10% 内,能够基本满足精度需求。远场电磁力模型如图 2 所示。该模型将线圈电磁场视为偶极子,其大小和方向用磁矩描述,磁矩的求解公式为

$$\mu = nISl \quad (3)$$

式中: n 为线圈的匝数; I 为电流值; S 为线圈的面积; l 为载流平面法线方向单位向量,服从右手定则。

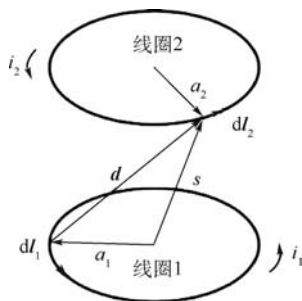


图 1 电磁线圈相互作用示意图

Fig.1 Schematic diagram of electromagnetic coil interaction

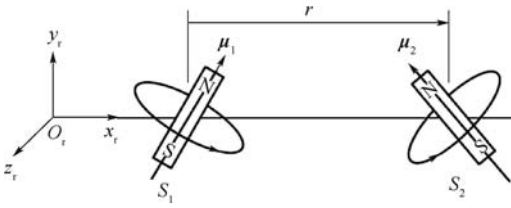


图2 磁偶极子相互作用示意图

Fig.2 Schematic diagram of magnetic dipole interaction

远场模型下,两线圈之间的电磁力和电磁力矩可以分别为

$$F_2 = -\frac{3\mu_0}{4\pi} \left[-\frac{\mu_1 \cdot \mu_2}{r^5} r - \frac{\mu_1 \cdot r}{r^5} \mu_2 - \frac{\mu_2 \cdot r}{r^5} \mu_1 + 5 \frac{(\mu_1 \cdot r)(\mu_2 \cdot r)}{r^7} r \right] \quad (4)$$

$$T_2 = \mu_2 \frac{\mu_0}{4\pi} \left(3 \frac{\mu_1 \cdot r}{r^5} r - \frac{\mu_1}{r^3} \right) \quad (5)$$

由式(4)、式(5)可以看出,远场模型下电磁力是线圈磁矩和线圈之间距离的强非线性函数,这使得在电磁力已知的情况下,求解磁矩变得较为困难。从式(4)和式(5)也可以看出,电磁力和电磁力矩是耦合产生的,在多航天器编队(航天器数量大于2的编队)中,这种耦合性将导致设计控制律时产生控制欠驱动问题。本文设计控制律时,采用轨道优先策略对其进行解耦,将电磁力矩考虑为干扰因素,其可以用反动量轮或力矩陀螺进行消除。本文研究重点是电磁航天器编队的轨道悬停控制,因此不考虑其姿态控制,并假设航天器的姿态已经调整到了期望的状态。

1.2 相对运动动力学模型

为描述航天器之间的相对运动,首先引入2个参考坐标系,即地心惯性坐标系 S_i 和参考轨道坐标系 S_o 。如图3所示,地心惯性坐标系坐标原点位于地球中心, X 轴位于赤道平面,指向春分点方向, Z 轴垂直于赤道平面,指向北极为正, Y 轴与 X 轴、 Z 轴构成右手系。参考轨道坐标系坐标原点位于参考航天器质心, x 轴沿径向,从地球中心指向参考航天器质心, y 轴在轨道平面

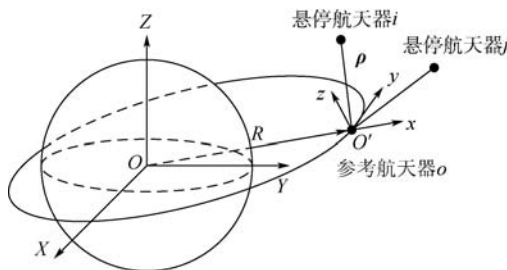


图3 航天器编队悬停参考系

Fig.3 Spacecraft formation hovering coordinate system

内与 x 轴垂直,指向参考航天器前进方向, z 轴与 x 轴、 y 轴构成右手坐标系。

假设对参考航天器不施加控制且编队中成员航天器之间的距离远小于航天器质心到地心的距离,则在参考轨道坐标系下,悬停编队中的成员航天器相对参考航天器的线性相对运动动力学方程为

$$\ddot{\rho}_i + D_e \dot{\rho}_i + K_e \rho_i = u_{mhi} + u_{dhi} - u_{do} \quad (6)$$

式中:下标 $i=1,2,\dots,N$ 为悬停编队中航天器的编号; $\rho_i = [x_i \ y_i \ z_i]^T$ 为悬停航天器 i 在参考轨道坐标系 S_o 中的位置; u_{mhi} 为悬停航天器 i 受到的远场电磁力; u_{dhi} 为悬停航天器 i 受到的外部扰动,包括电磁力建模误差、气动力、太阳光压以及地磁场和其他天体引力引起的扰动; u_{do} 参考航天器受到的外部扰动,包括气动力、太阳光压扰动以及其他天体引力。式(6)中的各系数矩阵为

$$D_e = \begin{bmatrix} 0 & -2\omega & 0 \\ 2\omega & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (7)$$

$$K_e = \begin{bmatrix} -\omega^2 - \frac{2\mu^*}{R^3} & -\dot{\omega} & 0 \\ \dot{\omega} & -\omega^2 + \frac{\mu^*}{R^3} & 0 \\ 0 & 0 & \frac{\mu^*}{R^3} \end{bmatrix} \quad (8)$$

其中: R 为参考航天器质心到地心的距离; ω 、 $\dot{\omega}$ 分别为参考航天器的轨道角速度和角加速度; μ^* 为地球引力常数。

2 电磁航天器编队悬停控制方法

2.1 鲁棒协同控制律设计

滑模控制能够有效克服系统的不确定性,对干扰和未建模动态具有很强的鲁棒性,且具有响应迅速、物理实现简单等优点。本文将以滑模控制方法为基础实现航天器的悬停。

对于一般的椭圆参考轨道,式(6)表示的相对运动动力学模型是一个系数时变的微分方程,式中变化的参数都与参考轨道相关,若按式(6)设计控制律将需要获得参考轨道实时的准确参数,带来很多的不便。本文拟将一个与参考轨道同周期的圆轨道作为基准,将参考轨道和标称圆轨道的偏差、地球非球形摄动以及其他各种摄动单独归类,作为不确定量进行处理,由此得到的动力学方程为

$$\ddot{\rho}_i + (D_e + \Delta D) \dot{\rho}_i + (K_e + \Delta K) \rho_i = u_{mhi} + u_{di} \quad (9)$$

式中: $\mathbf{u}_{di} = \mathbf{u}_{dhi} - \mathbf{u}_{di}$ 为悬停航天器相对参考航天器的外部不确定量; $\Delta \mathbf{D}$ 为椭圆轨道相对于圆轨道的偏差以及动力学方程的参数摄动; $\Delta \mathbf{K}$ 为参考轨道相对于圆轨道的偏差、地球非球形引力摄动以及动力学方程的参数摄动。与椭圆轨道同周期的圆轨道的半径为椭圆的半长轴 a , 因此矩阵 \mathbf{D}_c 、 \mathbf{K}_c 的值为

$$\mathbf{D}_c = \begin{bmatrix} 0 & -2\sqrt{\mu^*/a^3} & 0 \\ 2\sqrt{\mu^*/a^3} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (10)$$

$$\mathbf{K}_c = \begin{bmatrix} -3\mu^*/a^3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \mu^*/a^3 \end{bmatrix} \quad (11)$$

不确定矩阵可表示为

$$\Delta \mathbf{D} = \sum_{i=1}^k \alpha_i \mathbf{E}_i \quad |\alpha_i| \leq a_i \quad (12)$$

$$\Delta \mathbf{K} = \sum_{j=1}^h \beta_j \mathbf{F}_j \quad |\beta_j| \leq b_j \quad (13)$$

$$|\mathbf{u}_{di}| \leq \mathbf{f}_i(x, t) \quad \mathbf{f}_i(x, t) \geq 0 \quad (14)$$

式中: \mathbf{E}_i 、 \mathbf{F}_j 为已知矩阵, 存在参数摄动的位置元素为 1, 其余元素为 0; α_i 、 β_j 为不确定矩阵 $\Delta \mathbf{D}$ 、 $\Delta \mathbf{K}$ 对应位置的摄动参数; a_i 、 b_j 分别为 α_i 、 β_j 模的上界; $\mathbf{f}_i(x, t)$ 为外部不确定量 \mathbf{u}_{di} 模的上界; k 、 h 分别为矩阵 $\Delta \mathbf{D}$ 、 $\Delta \mathbf{K}$ 不为零的元素的个数。

小偏心率参考轨道本质是时变的, 动力学模型中存在真近点角, 涉及到实时进行开普勒方程的求解, 同此还涉及到线性时变系统的控制, 在设计控制律时存在着一定的难度。这里通过用一个同周期的圆参考轨道代替这个小偏心率椭圆, 使得系统由非定常变成定常, 分析设计比较方便。这么一来, 就会出现一个小周期项的动力学偏差, 这里把它也作为不确定性进行处理。另外, 由于文中把时变的部分都当作小扰动, 那么即使对参考卫星的建模存在一定误差, 上面设计的控制仍然可以保证不错的效果。因此, 基于式(9)的模型设计的控制律将具有更强的鲁棒性。

令悬停航天器 i 标称状态为 ρ_i^* 、 $\dot{\rho}_i^*$, 可以求得当前状态和标称状态的偏差为

$$\tilde{\rho}_i = \rho_i - \rho_i^*$$

$$\tilde{\dot{\rho}}_i = \dot{\rho}_i - \dot{\rho}_i^*$$

选取滑动面为

$$\mathbf{s}_i = \tilde{\rho}_i + \Lambda \tilde{\dot{\rho}}_i \quad (15)$$

式中: $\Lambda = \text{diag}\{\lambda_r\}$, $\lambda_r > 0$, $r = 1, 2, 3$, 显然滑动面上的状态运动是稳定的。接下来设计到达运动,

对式(15)的滑动面函数求导, 得

$$\dot{\mathbf{s}}_i = \ddot{\tilde{\rho}}_i + \Lambda \dot{\tilde{\rho}}_i = -(\mathbf{D}_c + \Delta \mathbf{D}) \dot{\rho}_i + \Lambda \dot{\rho}_i - (\mathbf{K}_c + \Delta \mathbf{K}) \rho_i + \mathbf{u}_{mhi} + \mathbf{u}_{di} - \ddot{\rho}_i^* - \Lambda \dot{\rho}_i^* \quad (16)$$

本文研究电磁航天器的编队悬停问题, 有 $\ddot{\rho}_i^* = \mathbf{0}$, $\dot{\rho}_i^* = \mathbf{0}$, 式(16)可简化为

$$\dot{\mathbf{s}}_i = \ddot{\tilde{\rho}}_i + \Lambda \dot{\tilde{\rho}}_i = -(\mathbf{D}_c + \Delta \mathbf{D}) \dot{\rho}_i + \Lambda \dot{\rho}_i - (\mathbf{K}_c + \Delta \mathbf{K}) \rho_i + \mathbf{u}_{mhi} + \mathbf{u}_{di} \quad (17)$$

选择到达控制律为

$$\dot{\mathbf{s}}_i = -\mathbf{H}_i \mathbf{s}_i - \boldsymbol{\varepsilon}_i \text{sat}(\mathbf{s}_i) \quad (18)$$

式中: \mathbf{H}_i 和 $\boldsymbol{\varepsilon}_i$ 为对角元素均为正的对角矩阵, 且 $\mathbf{H}_i, \boldsymbol{\varepsilon}_i \in \mathbf{R}^{3 \times 3}$ 。为了消除滑动运动在滑膜面上的抖颤, 在此处选择饱和函数 $\text{sat}(x)$ 代替符号函数:

$$\text{sat}(x) = \begin{cases} 1 & x > \varphi \\ x/|\varphi| & -\varphi \leq x \leq \varphi \\ -1 & x < -\varphi \end{cases} \quad (19)$$

式中: φ 为饱和函数边界层厚度。对比式(17)、式(18)得到等效控制为

$$\mathbf{u}_{mhi} = -\mathbf{H}_i \mathbf{s}_i - \boldsymbol{\varepsilon}_i \text{sat}(\mathbf{s}_i) + (\mathbf{D}_c - \Lambda) \dot{\rho}_i + \mathbf{K}_c \rho_i + \Delta \mathbf{D} \dot{\rho}_i + \Delta \mathbf{K} \rho_i - \mathbf{u}_{di} \quad (20)$$

式(20)控制中存在不确定参数, 无法直接使用, 为消去不确定参数, 首先任选 2 个向量 \mathbf{z}_{1i} 、 \mathbf{z}_{2i} , 构造式(21)作为实际控制:

$$\mathbf{u}_{mhi} = -\mathbf{H}_i \mathbf{s}_i - \boldsymbol{\varepsilon}_i \cdot \text{sat}(\mathbf{s}_i) + (\mathbf{D}_c - \Lambda) \dot{\rho}_i + \mathbf{K}_c \rho_i - \mathbf{z}_{1i} - \mathbf{z}_{2i} \quad (21)$$

为确定 \mathbf{z}_{1i} 、 \mathbf{z}_{2i} , 将式(21)代入式(17)中得

$$\dot{\mathbf{s}}_i = -\mathbf{H}_i \mathbf{s}_i - \boldsymbol{\varepsilon}_i \cdot \text{sat}(\mathbf{s}_i) - \mathbf{z}_{1i} - \Delta \mathbf{D} \dot{\rho}_i - \Delta \mathbf{K} \rho_i - (\mathbf{z}_{2i} - \mathbf{u}_{di}) \quad (22)$$

又滑模控制到达条件为

$$\text{diag}(\mathbf{s}_i) \text{diag}(\dot{\mathbf{s}}_i) < 0 \quad (23)$$

将式(22)代入式(23)之后写成分量形式得

$$\begin{aligned} s_i^r \dot{s}_i^r &= -\mathbf{H}_i^{rr} (s_i^r)^2 - \boldsymbol{\varepsilon}_i^{rr} \cdot \\ &|s_i^r| - s_i^r z_{1i}^r - s_i^r \sum_{j=1}^k \alpha_j v^r(\mathbf{E}_j) \dot{\rho}_i^r - \\ &s_i^r \sum_{j=1}^h \beta_j v^r(\mathbf{F}_j) \rho_i^r - s_i^r (z_{2i}^r - \mathbf{u}_{di}^r) \end{aligned} \quad (24)$$

式中: 上标 $r = 1, 2, 3$ 为向量的第 r 个分量; $v^r(\ast)$ 为矩阵 (\ast) 的第 r 行, 若取 \mathbf{z}_{1i} 、 \mathbf{z}_{2i} 为式(25)、式(26)。将式(25)、式(26)代入式(24), 比较各项大小, 消去不确定项, 可使得到达条件式(23)得到满足。

$$\mathbf{z}_{1i}^r = \left(\sum_{j=1}^m a_j |v^r(\mathbf{E}_j) \dot{\rho}_i^r| + \sum_{k=1}^h b_k |v^r(\mathbf{F}_k) \rho_i^r| \right) \times \text{sat}(s_i^r) \quad (25)$$

$$\mathbf{z}_{2i}^r = |\mathbf{f}_i^r(x, t)| \text{sat}(\mathbf{s}_i^r) \quad (26)$$

式(21)~式(26)构成航天器编队悬停的控制律,该控制律已经能够实现航天器编队的悬停任务,且由于动力学模型中考虑了航天器可能受到的所有的误差,同时给出了模型误差的上界。因此上面给出的控制对不确定干扰具有较强的鲁棒性,但该控制律只考虑了编队中成员航天器自身的位置和速度偏差信息,没有用到其他航天器的状态信息,从编队系统层面上来看,所设计的控制律只能达到局部的良好性能,从全局视角审视,存在控制的浪费,因此对该控制律进行改进,以利用航天器之间的信息交换实现全局协同。

为了清晰地描述航天器编队中成员航天器之间的关系,在此引入图论的相关概念。一个加权无向图 $G(V, E, \mathbf{A})$ 由节点集 $V = \{1, 2, \dots, N\}$, 边集 $E \subseteq V \times V$ 和加权邻接矩阵 $\mathbf{A} = a_{ij} \in \mathbf{R}$ 组成。节点表示编队中的成员航天器,边表示编队中成员航天器之间的通信关系,邻接矩阵表示编队中成员航天器的通信性能。无向图中若两航天器 i 、 j 之间可以通信,则边 $(i, j) \in E$ 、 $(j, i) \in E$ 且 $a_{ij} = a_{ji} > 0$, 否则 $a_{ij} = 0$ 。一般认为航天器和自身不存

在通信,即 $a_{ii} = 0$ 。如果图中的任何 2 个节点之间存在路径,则称该无向图是连通的。

在上述图论的基础上,考虑到编队中成员航天器之间的信息交互以及通信延时,对上述鲁棒控制器增加一协同项,得到协同控制器如式(27)所示:

$$\begin{aligned} \mathbf{u}_{mhi} = & -\mathbf{H}_i \mathbf{s}_i - \boldsymbol{\varepsilon}_i \cdot \text{sat}(\mathbf{s}_i) + (\mathbf{D}_c - \boldsymbol{\Lambda}) \dot{\boldsymbol{\rho}}_i + \\ & \mathbf{K}_c \boldsymbol{\rho}_i - \mathbf{z}_{1i} - \mathbf{z}_{2i} - \sum_{j=1}^N a_{ij} (\mathbf{s}_i - \mathbf{s}_j(t - \tau_{ij})) \end{aligned} \quad (27)$$

式中: τ_{ij} 为电磁航天器 i 、 j 之间的通信延时。

2.2 鲁棒协同控制器稳定性证明

本节分析和证明上述控制器(27)的稳定性,选取 Lyapunov 函数为式(28),将其对时间求导得到式(29)。将式(22)、式(25)、式(26)代入式(28),可以得到式(29),双画线部分小于等于零,因此对于 Lyapunov 函数的证明只需要证后面部分,即式(30)。

$$V = \frac{1}{2} \sum_{i=1}^N \mathbf{s}_i^T \mathbf{s}_i + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N a_{ij} \int_{t-\tau_{ij}}^t \mathbf{s}_j^T(\tau) \mathbf{s}_j(\tau) d\tau \quad (28)$$

$$\begin{aligned} \dot{V} = & \sum_{i=1}^N \mathbf{s}_i^T \dot{\mathbf{s}}_i + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N a_{ij} [\mathbf{s}_j^T \mathbf{s}_j - \mathbf{s}_j^T(t - \tau_{ij}) \mathbf{s}_j(t - \tau_{ij})] = \\ & \sum_{i=1}^N \underline{\underline{-\mathbf{s}_i^T \mathbf{H}_i \mathbf{s}_i - \boldsymbol{\varepsilon}_i |\mathbf{s}_i| - \mathbf{s}_i^T \mathbf{z}_{1i} - \mathbf{s}_i^T \Delta \mathbf{D} \dot{\boldsymbol{\rho}}_i - \mathbf{s}_i^T \Delta \mathbf{K} \boldsymbol{\rho}_i - \mathbf{s}_i^T (\mathbf{z}_2 - \mathbf{u}_{di}) -}} \\ & \sum_{i=1}^N \mathbf{s}_i^T \sum_{j=1}^N a_{ij} [\mathbf{s}_i - \mathbf{s}_j(t - \tau_{ij})] + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N a_{ij} [\mathbf{s}_j^T \mathbf{s}_j - \mathbf{s}_j^T(t - \tau_{ij}) \mathbf{s}_j(t - \tau_{ij})] \end{aligned} \quad (29)$$

$$\begin{aligned} \dot{V}_1 = & -\sum_{i=1}^N \mathbf{s}_i^T \sum_{j=1}^N a_{ij} [\mathbf{s}_i - \mathbf{s}_j(t - \tau_{ij})] + \\ & \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N a_{ij} [\mathbf{s}_j^T \mathbf{s}_j - \mathbf{s}_j^T(t - \tau_{ij}) \mathbf{s}_j(t - \tau_{ij})] = \\ & -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N a_{ij} \mathbf{s}_i^T \mathbf{s}_i + \sum_{i=1}^N \sum_{j=1}^N a_{ij} \mathbf{s}_i^T \mathbf{s}_j(t - \tau_{ij}) - \\ & \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N a_{ij} [\mathbf{s}_j^T(t - \tau_{ij}) \mathbf{s}_j(t - \tau_{ij})] \leq 0 \end{aligned} \quad (30)$$

式(30)进一步推导得到 $\dot{V}_1 \leq 0$, 因此得到 Lyapunov 函数对时间的导数 $\dot{V} \leq 0$, 由 Barbalat 引理推论可知,当 $t \rightarrow \infty$ 时 $\mathbf{s}_i \rightarrow \mathbf{0}$, 由于所设计的滑动面上的运动是稳定的,进而可以得到 $\tilde{\boldsymbol{\rho}}_i \rightarrow \tilde{\boldsymbol{\rho}}_j \rightarrow \mathbf{0}$ 且 $\tilde{\boldsymbol{\rho}}_i \rightarrow \tilde{\boldsymbol{\rho}}_j \rightarrow \mathbf{0}$ 。从证明过程也可以看出,本文设计的控制律的稳定性和通信拓扑结构的连通性没有关系。当其联通时,系统将有更强的协同能力。

3 电磁航天器编队磁矩优化配置

2.1 节建立的航天器编队悬停控制律可以得到航天器在任意位置悬停需要的控制加速度。本节讨论控制加速度的具体实现问题。电磁航天器编队系统有 N 个成员航天器,编号为 $i = 1, 2, 3, \dots, N$, 成员航天器均能够产生任意指向的磁矩。对于编号为 i 的航天器,其受到的电磁力表达式为

$$\mathbf{F}_i = \sum_{j=1}^N \mathbf{F}_{ij} = \sum_{j=1}^N \mathbf{F}_{ij}(\boldsymbol{\mu}_i, \boldsymbol{\mu}_j, \mathbf{r}_{ij}) \quad i \neq j \quad (31)$$

式中: \mathbf{F}_i 为编号为 i 的航天器受到电磁力的合力; \mathbf{F}_{ij} 为 j 号航天器对 i 号航天器的电磁作用; $\boldsymbol{\mu}_i$ 、 $\boldsymbol{\mu}_j$ 分别为 i 号航天器和 j 号航天器的磁矩; \mathbf{r}_{ij} 为 i 号航天器相对于 j 号航天器的位置,从 j 号航天器指向 i 号航天器。由于电磁力为系统内力,故存在式(32)、式(33)成立:

$$\mathbf{F}_{ij} + \mathbf{F}_{ji} = \mathbf{0} \quad (32)$$

$$\sum_{j=1}^N \mathbf{F}_i = \mathbf{0} \quad (33)$$

可以看出整个编队系统受到的电磁力的合力为零,显然在只有相互作用电磁力的情况下,系统质心无法改变,无法实现任意位置的悬停,为此在系统中任选一颗航天器,在该航天器上施加额外推力,通过该推力来改变质心位置。设推力施加在 N 号航天器上,则满足:

$$\mathbf{F}_N = \sum_{j=1}^{N-1} \mathbf{F}_{Nj} + \mathbf{f}_0 \quad (34)$$

式中: \mathbf{f}_0 为施加在 N 号航天器上的悬停外力。

从上述分析中可以看出,电磁航天器编队系统一共受到独立的约束 $3N$ 个,但是相互独立的变量有 $3N+3$ (包含 $3N$ 个磁矩变量和 3 个外力变量) 个,系统的变量有冗余的自由度,因此可以将该问题转化为带约束的优化问题进行求解。取目标函数表达式为

$$\min \left(\sum_{i=1}^N \boldsymbol{\mu}_i^T \mathbf{W}_{1i} \boldsymbol{\mu}_i + \sum_{i=1}^N (\boldsymbol{\mu}_i - \bar{\boldsymbol{\mu}})^T \mathbf{W}_{2i} (\boldsymbol{\mu}_i - \bar{\boldsymbol{\mu}}) + \sum_{i=1}^N \gamma_i |\boldsymbol{\tau}_i| + \mathbf{f}_0^T \mathbf{W}_0 \mathbf{f}_0 \right) \quad (35)$$

式中: $\bar{\boldsymbol{\mu}}$ 为编队中各航天器磁矩的平均值, $\bar{\boldsymbol{\mu}} = \frac{1}{N} \sum_{i=1}^N \boldsymbol{\mu}_i$; \mathbf{W}_{1i} 和 \mathbf{W}_{2i} 为权重系数,通过选择该矩阵可以实现磁矩分配的能量最小和能量均衡; $\boldsymbol{\tau}_i$ 为 i 号航天器受到的电磁干扰力矩,该力矩数量级较小; γ_i 为选择的配平系数; $\mathbf{f}_0^T \mathbf{W}_0 \mathbf{f}_0$ 表示所用的外力最小, \mathbf{W}_0 为其配平系数。

4 数值仿真

4.1 仿真参数设置

为验证所设计控制律的控制性能以及磁矩分配方案的可行性,本文对 4 颗电磁航天器组成的编队进行了仿真计算,航天器编队采用的编队构型如图 4 所示。

初始时刻 4 颗航天器位于棱长 15 m 的正四面体的 4 个顶点上,参考航天器位于正四面体的中心。假设编队成员航天器的位置和速度在初始时刻受到随机扰动,在运行期间一直存在低于扰动上限的噪声。仿真选择的参考轨道参数,各个状态期望值和初始摄动值分别如表 1 和表 2 所示。

成员航天器的质量均为 100 kg,线圈半径均为 1 m,线圈匝数均为 100。控制器的参数选择为 $\mathbf{H}_i = 10^{-2} \mathbf{I}^{3 \times 3}$, $\boldsymbol{\varepsilon}_i = 10^{-7} \mathbf{I}^{3 \times 3}$, $\mathbf{A}_i = 10^{-2} \mathbf{I}^{3 \times 3}$ 。不确

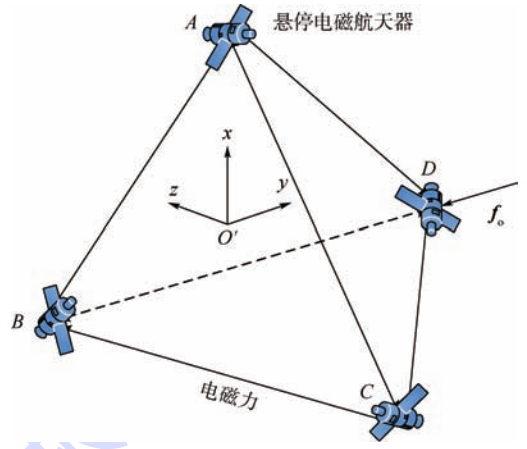


图 4 电磁航天器悬停编队构型

Fig. 4 Hovering formation configuration of electromagnetic spacecraft

表 1 参考轨道参数

Table 1 Reference orbital parameters

轨道参数	数值
a/km	7371
e	0.01
$i/(\circ)$	45
$\Omega/(\circ)$	45
$\bar{\omega}/(\circ)$	30
$\theta/(\circ)$	0

表 2 状态初值信息

Table 2 Initial value information of state

参数	期望值 \mathbf{X}	初始摄动值 δx
位置/m	$15 \times [-\sqrt{6}/4 \quad 0 \quad 0]$	$5 \times \text{rand}(-1, 1)$, $\text{rand}(-1, 1)$ 表示 $(-1, 1)$ 区间的 随机数
	$15 \times [-\sqrt{6}/12 \quad -\sqrt{3}/6 \quad 1/2]$	
	$15 \times [-\sqrt{6}/12 \quad -\sqrt{3}/6 \quad -1/2]$	
	$15 \times [-\sqrt{6}/12 \quad \sqrt{3}/3 \quad 0]$	
速度/ ($\text{m} \cdot \text{s}^{-1}$)	$[0 \quad 0 \quad 0]$	$0.01 \times \text{rand}(-1, 1)$

定参数上限的选择参考了文献[21],分别取 $a_i = 10^{-5}$, $b_j = 10^{-7}$, $f_i = 10^{-7}$ 。考虑 3 颗航天器链接成环,另外一颗航天器连入该三角形环的通信拓扑结构,权值均取为 $a_{ij} = 10^{-3}$ 。成员航天器通信延时为 5 s。在磁矩求解时,考虑到能量最优和均衡、电磁力和力矩的解耦以及系统所需外力最小,为使得式(35)每一项的数量及相当,配置磁矩参数取值为 $\mathbf{W}_{1i} = 10^{-11} \mathbf{I}^{3 \times 3}$, $\mathbf{W}_{2i} = 10^{-9} \mathbf{I}^{3 \times 3}$, $\gamma_i = 10^2$, $\mathbf{W}_0 = 10^5 \mathbf{I}^{3 \times 3}$,然后用 MATLAB 进行仿真。

4.2 仿真结果

仿真结果如图 5 ~ 图 11 所示。图 5 和图 6 为相对位置跟踪误差曲线,图 7 和图 8 为相对速度跟踪误差曲线。图 9 ~ 图 11 分别描述了控制磁矩、电磁力、编队悬停外力的变化规律。

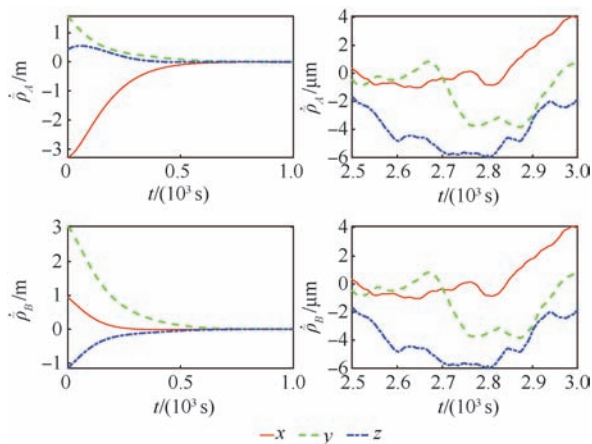


图5 AB相对位置跟踪误差曲线

Fig. 5 Tracking error curves of AB relative position

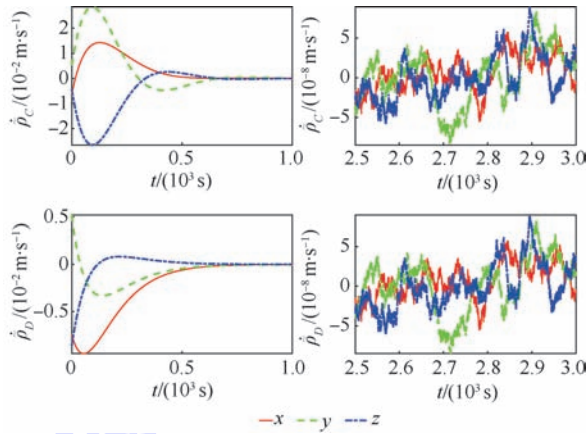


图8 CD相对速度跟踪误差曲线

Fig. 8 Tracking error curves of CD relative velocity

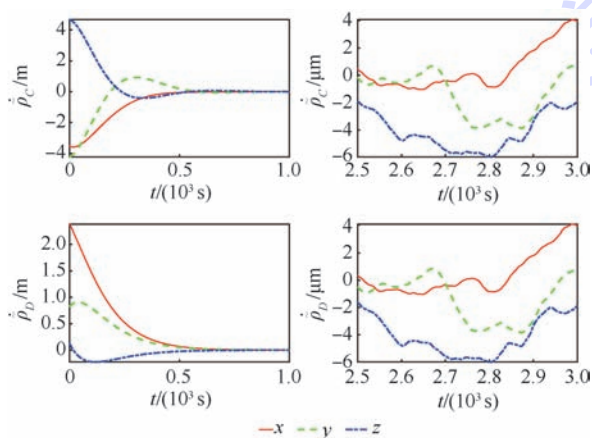


图6 CD相对位置跟踪误差曲线

Fig. 6 Tracking error curves of CD relative position

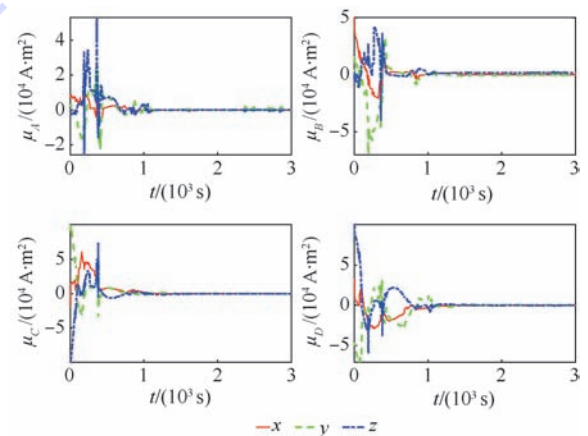


图9 控制磁矩曲线

Fig. 9 Curves of control magnetic moment

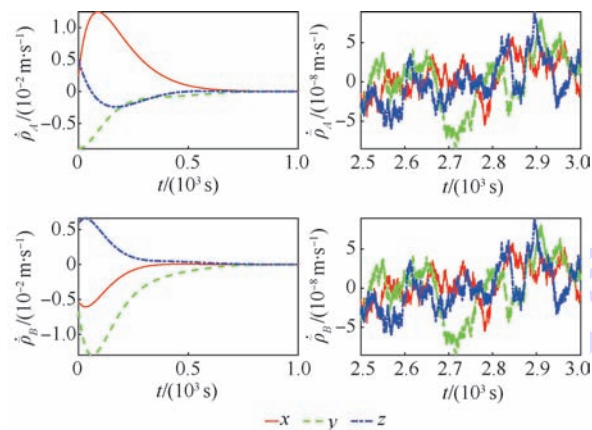


图7 AB相对速度跟踪误差曲线

Fig. 7 Tracking error curves of AB relative velocity

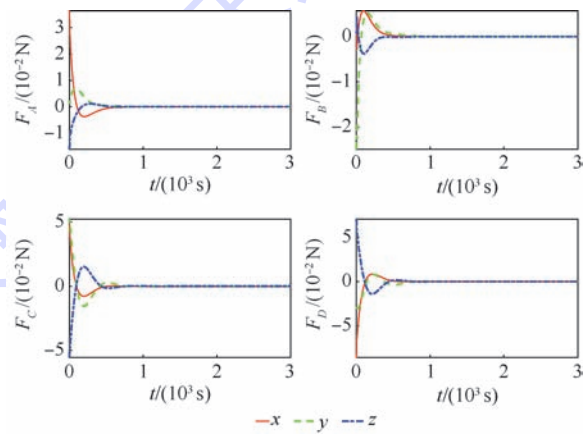


图10 电磁力曲线

Fig. 10 Curves of electromagnetic force

从图5、图6可以看出,在位置初值、速度初值均存在随机扰动并且状态信息存在噪声的情况下,本文设计的鲁棒协同控制律利用航天器之间的电磁力能够在700s内回到期望轨道,且位置跟踪误差在 10^{-6} 数量级。从放大的结果也可以看出,控制器中的协同项最终使得编队中所有航天

器的位置跟踪误差趋于一致,这有利于误差的管理。

图7和图8表明,协同控制律可以使成员航天器对速度的跟踪误差达到 10^{-8} 数量级,具有较高的精度,同时也说明设计的控制器具有较强的鲁棒性和协同性。

从图9可以看出,优化之后的控制磁矩数

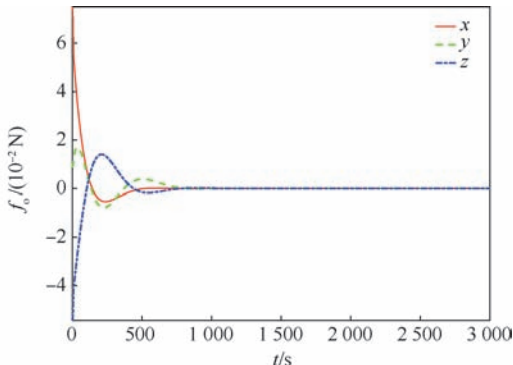


图 11 编队悬停外力曲线

Fig. 11 Curve of formation hovering external force

量均小于 10^5 , 该磁矩可以使用超导线圈实现。图 10 表明编队悬停控制所需的电磁力在 10 mN 量级,一方面这样的小推力可以使得控制精度有所提升,另一方面也表明电磁力只能实现近距离的编队。图 11 为编队悬停时为了改变航天器质心所需要的外力,从图中可以看出其大小和所需的电磁力相当。

4.3 有无协同控制效果比较

为了分析在控制律中引入协同项的效果,在同样的初值下对未引入协同的情况也进行了仿真。选取文献[22]中的暂态构型维持误差 σ 作为评价指标,其具体形式为

$$\sigma = \|\rho_1\|_2 - \|\rho_3\|_2 + \|\rho_1\|_2 - \|\rho_1 - \rho_2\|_2 + \|\rho_1 - \rho_2\|_2 - \|\rho_2 - \rho_3\|_2 + \|\rho_2\|_2 - \|\rho_1 - \rho_3\|_2 \quad (36)$$

式中: σ 越小表明暂态构型维持能力越强。从图 12 中可以看出,在编队成员均存在 5 s 通信延时的情况下,尽管在 200 s 附近有无协同指标略低于有协同指标的情况,但是从整体趋势来看,有协同的控制器的表现出了更好的性能。

为分析设计的控制律的能量消耗情况,本文采用系统各成员航天器所受力的模之和作为能

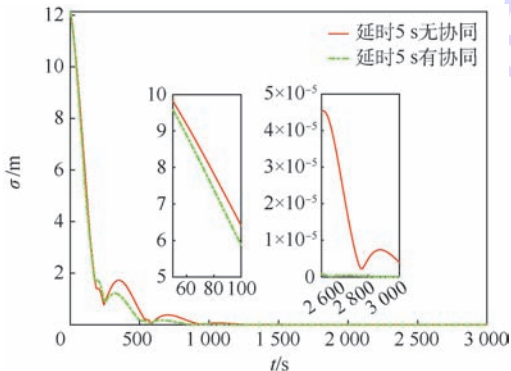


图 12 暂态构型维持位置误差曲线

Fig. 12 Curve of transient configuration position keeping error

量评价指标,记为 E 。在相同初值下对编队开环控制的能量指标进行了计算,结果如图 13 所示。

图中箭头所标数字为一个周期内不同控制方式能量指标 \bar{E} 的总和。从图中可以看出,在初始时刻闭环控制需要的能量较多,在 400 s 之后,开环控制需要的能量将超过闭环控制。在一个周期内开环控制需要的总能量约闭环控制的 2.05 倍。

以上仿真结果表明本文设计的电磁航天器编队鲁棒协同控制律能够有效实现航天器编队任意位置的悬停,对系统存在的内部和外部不确定性具有较强的鲁棒性,文中给出的基于能量最优和均衡的磁矩配置方案能够实现磁矩的有效分配。

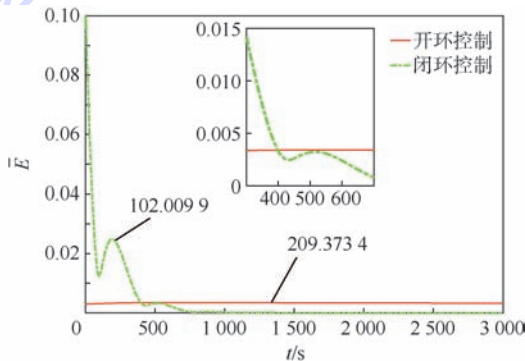


图 13 能量消耗对比

Fig. 13 Comparison of energy consumption

5 结 论

本文针对近地近圆轨道附近电磁航天器编队悬停问题,设计了一种有效处理绝对轨道信息缺失的控制方法,具体如下所述。在电磁力模型和动力学模型均存在不确定性的条件下,通过引入滑模控制理论和一致性理论,设计了鲁棒协同控制律。考虑到能量最优和控制均衡以及轨道姿态的解耦,给出了磁矩的优化配置方案。分析证明了所设计的协同控制律的稳定性及鲁棒性。仿真结果表明,位置跟踪误差和速度跟踪误差分别能达到 10^{-6} 、 10^{-8} 数量级,所给出的磁矩分配方案能够实现磁矩的合理分配,控制中引入协同项后可以提高编队的协同能力,且能平衡各成员航天器之间能量消耗。

参考文献 (References)

[1] 林来兴,黎康. 卫星对空间目标悬停的轨道动力学与控制方法研究[J]. 中国空间科学技术, 2008, 28(1): 9-12.
LIN L X, LI K. Orbit dynamics and control of satellites hovering over space target[J]. Chinese Space Science and Technology, 2008, 28(1): 9-12 (in Chinese).
[2] 朱亚文,闫野. 椭圆轨道卫星空间任意位置悬停的方法

- [J]. 中国空间科学技术, 2010, 30(6): 17-23.
- ZHU Y W, YAN Y. Hovering method at any selected position over space target on elliptical orbit[J]. Chinese Space Science and Technology, 2010, 30(6): 17-23 (in Chinese).
- [3] 林来兴, 张小琳. 纳型卫星编队飞行技术现状及发展趋势[J]. 航天器工程, 2017, 26(5): 65-73.
- LIN L X, ZHANG X L. Current status and developing trends of nanosatellites formation flying [J]. Spacecraft Engineering, 2017, 26(5): 65-73 (in Chinese).
- [4] 张皓, 师鹏, 李保军, 等. 利用库仑力实现悬停轨道的新方法研究[J]. 宇航学报, 2012, 33(1): 68-75.
- ZHANG H, SHI P, LI B J, et al. Hover orbit using inter-spacecraft coulomb forces[J]. Journal of Astronautics, 2012, 33(1): 68-75 (in Chinese).
- [5] 李俊峰, 龚胜平. 非开普勒轨道动力学与控制[J]. 宇航学报, 2009, 30(1): 47-53.
- LI J F, GONG S P. Dynamics and control of the non-Keplerian orbits [J]. Journal of Astronautics, 2009, 30(1): 47-53 (in Chinese).
- [6] HUANG J, LI C, MA G, et al. Coulomb control of a triangular three-body satellite formation using nonlinear model predictive method [C] // Proceedings of the 33rd Chinese Control Conference. Piscataway, NJ: IEEE Press, 2014: 7685-7690.
- [7] POLLOCK G E, GANGESTAD J W, LONGUSKI J M. Inclination change in low-earth orbit via the geomagnetic Lorentz force [J]. Journal of Guidance, Control, and Dynamics, 2010, 33(5): 1387-1395.
- [8] KWON D W. Propellantless formation flight applications using electromagnetic satellite formations [J]. Acta Astronautica, 2010, 67(9-10): 1189-1201.
- [9] KONG E M C, KWON D W, SCHWEIGHART S A, et al. Electromagnetic formation flight for multi-satellite arrays [J]. Journal of Spacecraft & Rockets, 2004, 41(4): 659-666.
- [10] 魏伟, 左敏, 苏婷立, 等. 航天器对接的全局渐近稳定控制[J]. 清华大学学报(自然科学版), 2016, 56(1): 106-110.
- WEI W, ZUO M, SU T L, et al. Global asymptotically stable control for spacecraft docking [J] Journal of Tsinghua University (Science and Technology), 2016, 56(1): 106-110 (in Chinese).
- [11] ELIAS L M, KWON D W, SEDWICK R J, et al. Electromagnetic formation flight dynamics including reaction wheel gyroscopic stiffening effects [J]. Journal of Guidance, Control, and Dynamics, 2007, 30(2): 499-511.
- [12] 黄涣, 杨乐平, 朱彦伟, 等. 双星电磁编队的动力学平衡态稳定性与控制[J]. 国防科技大学学报, 2013, 35(3): 12-17.
- HUANG H, YANG L P, ZHU Y W, et al. Stability and control of dynamics equilibrium for two-spacecraft electromagnetic formation [J]. Journal of National University of Defense Technology, 2013, 35(3): 12-17 (in Chinese).
- [13] 徐增文, 师鹏, 赵育善. 双电磁航天器编队构型保持自适应控制 [J]. 北京航空航天大学学报, 2015, 41(12): 2302-2308.
- XU Z W, SHI P, ZHAO Y S. Adaptive control for two-spacecraft electromagnetic formation keeping [J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(12): 2302-2308 (in Chinese).
- [14] 邵龙飞, 师鹏, 赵育善. 电磁航天器编队动力学建模与运动规划方法 [J]. 北京航空航天大学学报, 2015, 41(4): 737-743.
- SHAO L F, SHI P, ZHAO Y S. Dynamics modeling and motion programming for electromagnetic formation flight [J]. Journal of Beijing University of Aeronautics and Astronautics, 2015, 41(4): 737-743 (in Chinese).
- [15] KWON D W, SEDWICK R J, LEE S I, et al. Electromagnetic formation flight testbed using superconducting coils [J]. Journal of Spacecraft & Rockets, 2011, 48(1): 124-134.
- [16] AHSUN U, MILLER D W. Dynamics and control of electromagnetic satellite formations [C] // American Control Conference. Piscataway, NJ: IEEE Press, 2007: 1730-1735.
- [17] MILLER D W, AHSUN U, RAMIREZRIBEROS J L. Control of electromagnetic satellite formations in near-earth orbits [J]. Journal of Guidance, Control and Dynamics, 2010, 33(6): 1883-1891.
- [18] 胡敏, 曾国强, 党朝辉. 近地轨道集群航天器电磁编队飞行非线性反馈控制方法 [J]. 空间科学学报, 2012, 32(3): 417-423.
- HU M, ZENG G Q, DANG Z H. Nonlinear feedback control of fractionated spacecraft electromagnetic formation flying in low-earth orbit [J]. Chinese Journal of Space Science, 2012, 32(3): 417-423 (in Chinese).
- [19] ZENG G, HU M. Finite-time control for electromagnetic satellite formations [J]. Acta Astronautica, 2012, 74(3): 120-130.
- [20] ZHANG J, YUAN C, JIANG D, et al. Adaptive terminal sliding mode control of electromagnetic spacecraft formation flying in near-earth orbits [J]. Advances in Mechanical Engineering, 2014, 2014(1): 512583.
- [21] 齐彧. 电磁航天器编队飞行关键技术研究 [D]. 北京: 北京航空航天大学, 2017.
- QI Y. Research on key technology of electromagnetic satellites formation flight [D]. Beijing: Beihang University, 2017 (in Chinese).
- [22] ZHANG B Q, SONG S M. Decentralized coordinated control for multiple spacecraft formation maneuvers [J]. Acta Astronautica, 2012, 74: 79-97.

作者简介:

张亚博 男, 硕士研究生。主要研究方向: 航天器动力学与控制、卫星编队飞行技术。

师鹏 男, 博士, 讲师, 硕士生导师。主要研究方向: 航天器总体设计、航天器动力学与控制。

张皓 男, 博士, 副研究员, 硕士生导师。主要研究方向: 航天器动力学与控制。

赵育善 男, 博士, 教授, 博士生导师。主要研究方向: 航天器动力学与控制。

A robust coordinated control method for hovering of electromagnetic spacecraft formation

ZHANG Yabo¹, SHI Peng^{1,*}, ZHANG Hao², ZHAO Yushan¹

(1. School of Astronautics, Beihang University, Beijing 100083, China;

2. Key Laboratory of Space Utilization, Technology and Engineering Center for Space Utilization, Chinese Academy of Sciences, Beijing 100094, China)

Abstract: Aimed at the problem of near-Earth orbit hovering of electromagnetic spacecraft formation, a coordinated control method is proposed in the absence of accurate information of reference orbit. The relative motion between spacecraft is described by the TH equation, and the circular orbit of the same cycle as the reference orbit is selected as the nominal orbit. The deviation of the reference orbit from the nominal circular orbit, the Earth's non-spherical gravity, the atmospheric resistance and the other celestial gravitation are classified separately, and they are considered as uncertain and constitute an uncertain system. By introducing the consistency theory, the robust coordinated control law is designed for the target of spacecraft formation hovering under the condition that the electromagnetic action model and the dynamic equation are all uncertain. Considering the optimal and balanced energy consumption and the decoupling of orbital attitude, a scheme of magnetic moment distribution through optimization is given. The simulation results show that the designed robust coordinated control law can achieve high-precision hovering of electromagnetic spacecraft formation. The proposed magnetic moment configuration scheme can realize the rational distribution of magnetic moments.

Keywords: electromagnetic spacecraft; formation flying; consistency protocol; robust control; hovering

Received: 2018-05-18; **Accepted:** 2018-07-05; **Published online:** 2018-07-19 17:10

URL: kns.cnki.net/kcms/detail/11.2625.V.20180718.1806.001.html

Foundation items: National Natural Science Foundation of China (11572019); the Open Research Fund of Key Laboratory of Space Utilization of Chinese Academy of Sciences (LSU-2016-07-02); 100 Talents Program of the Chinese Academy of Sciences

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

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

DOI: 10.13700/j.bh.1001-5965.2018.0294

基于半无码的 P(Y) 码自相关 GNSS-R 海面测高方法

樊梦文, 张波*, 王峰

(北京航空航天大学 电子信息工程学院, 北京 100083)



摘 要: 针对传统 GPS C/A 码延迟测高方法测高精度低和互相关测高方法时延波形信噪比低的问题, 研究了基于半无码的 P(Y) 码自相关海面测高方法。首先, 通过理论测高精度模型, 对比分析了 C/A 码和 P(Y) 码的测高精度; 然后, 描述了 L1、L2 双频反射信号基于半无码的 P(Y) 码自相关处理架构, 并论述了测高模型和 L1、L2 双频电离层时延误差修正方法; 最后, 利用模拟的 L1、L2 反射信号对比分析了基于半无码的 P(Y) 码自相关、C/A 码自相关以及信号互相关方法的海面测高精度。结果表明: 相比于 C/A 码自相关和信号互相关测高精度, 基于半无码的 P(Y) 码自相关的测高精度分别提高了 3.97 倍和 1.47 倍。

关键词: 半无码; P(Y) 码; 海面测高; 双频; 电离层时延校正

中图分类号: P237

文献标识码: A **文章编号:** 1001-5965(2019)02-0398-07

GNSS-R (Global Navigation Satellite System-Reflection) 技术通过接收、处理经地球物理表面反射的导航卫星 L 波段信号, 进行地表物理参数的探测和反演, 是一种新型的无源被动式遥感手段^[1]。通过大量理论研究和试验论证, 该技术已被证实可以用于海面高度测量^[2]、海面风场反演^[3]、海冰遥感^[4]和土壤湿度探测^[5]等。海面高度作为海洋表面的重要参数, 在海洋循环、气候研究以及灾害预警中具有重要作用。

自 1993 年 Martin-Neira^[6]提出了 PARIS (Passive Reflectometry and Interferometry System) 概念, 首次论证了 GPS 反射信号进行海面测高的可行性以来, 基于 GNSS-R 的海面测高一直是国内外学者的研究热点之一。根据天线配置的不同, GNSS-R 海面测高主要分单天线^[7]和双天线^[8]模式。相比于单天线测高模式, 双天线观测可配置在更广泛的应用场景中, 尤其可搭载在低轨卫星

上, 实现全球海面高度测量, 具有广阔的应用前景。GNSS-R 海面测高的基础观测量是 GNSS 反射和直射信号时延差, 可通过跟踪直射、反射信号载波或伪随机码得到。基于载波的测量具有厘米级测高精度, 但是要求反射信号具有强相干性, 即反射表面光滑, 主要应用在湖面、平静海面等场景, 且需要求解载波相位整周模糊度, 在处理算法上难度较大^[9]。基于伪随机码的海面测高技术对海面散射条件要求较低, 处理流程简单, 但是其测高精度和伪随机信号带宽成正比^[10]。直射和反射信号互相关技术因其能够充分利用信号带宽, 提高海面测高精度而被广泛研究和应用^[11]。但是, 该技术存在两方面的不足: 一方面, 由于经过长距离空间衰减的直射和反射信号均为弱信号, 为了提高互相关时延波形的信噪比, 直射和反射信号天线均需采用高增益天线; 另一方面, 为了降低不同卫星信号的互干扰, 直射和反射信号天

收稿日期: 2018-05-21; 录用日期: 2018-07-28; 网络出版时间: 2018-08-27 09:21

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180824.1525.004.html

基金项目: 国家自然科学基金 (41774028)

* 通信作者. E-mail: bozhang@buaa.edu.cn

引用格式: 樊梦文, 张波, 王峰. 基于半无码的 P(Y) 码自相关 GNSS-R 海面测高方法 [J]. 北京航空航天大学学报, 2019, 45 (2): 398-404. FAN M W, ZHANG B, WANG F. Semi-codeless based P(Y) code autocorrelation GNSS-R sea surface altimetry method [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45 (2): 398-404 (in Chinese).

线均需采用窄波束、多波束天线实现不同导航卫星信号的空间隔离和同步接收。上述两方面无疑增加了天线设计和研制难度,增大了测量设备的体积和质量。军用加密码,如 GPS P(Y)码具有更宽的信号带宽,测量精度高,但是由于被加密而无法直接实施本地自相关处理。半无码技术是未知加密码的确切结构和码速率,但是已知加密码是由伪随机码和一种未知的调制码模二和而成,且在 C/A 码已经进入稳态跟踪的情况下,通过对未知调制码进行估计,实现本地随机码和接收信号相关的过程。Carreno-Luengo^[12]和 Lowe^[13]等将半无码技术应用到 GNSS-R 海面测高中,得到了 P(Y)码的相关时延波形,有效提高了测高精度。相比于互相关技术,基于半无码技术的海面测高对直射和反射信号接收天线降低了要求,且通过对不同频点反射信号上的同一加密码进行跟踪,可完成电离层等时延误差的校正。

本文主要对 P(Y)码测高性能进行了分析,论述了基于半无码的 P(Y)码自相关测高方法和双频电离层修正模型,并在机载仿真环境下,对比分析了基于半无码的 P(Y)码自相关、信号互相关以及 C/A 码自相关测高方法的测高精度。

1 P(Y)码测高精度

测高精度模型描述了测高精度与时延观测量之间的一种解析关系,表征了反射信号时延波形测量的不确定性对海面测高的影响。对于复杂的粗糙海面,时延波形的不确定性可用散射区内的统计特征来进行描述^[14]。Martin-Neira 等^[8]通过对海面统计特性、信号特性的研究,提出了基于算术平方差的测高精度模型:

$$\sigma_h = \frac{c}{2 \sin \theta_{sp} \sqrt{N_{inc}}} \cdot \frac{P_z(0)}{P'_z(0)} \sqrt{\left(1 + \frac{1}{\text{SNR}}\right)^2 + \left(\frac{1}{\text{SNR}}\right)^2} \quad (1)$$

式中: σ_h 为测高精度; c 为真空中的光速; θ_{sp} 为镜面反射点处的卫星高度角; N_{inc} 为非相干累加次数; $P_z(0)$ 和 $P'_z(0)$ 分别为镜面反射点处平均功率的幅值和一阶微分值;SNR为镜面反射点处时延波形的信噪比。为了描述式(1)中信号自相关功率波形平均幅值与一阶微分值的比值,即信号的自相关特性对测高精度的影响,定义在镜面反射点处的测高灵敏度为^[15-16]

$$S_h = \frac{P'_z(0)}{c P_z(0)} = \frac{U(0)}{c \int_0^{\tau_c} U(-\tau) d\tau} \quad (2)$$

式中: τ_c 为测距码的周期; $U(\cdot)$ 为信号的自相关功率; τ 为时间延迟。信号入射到海面时与随机起伏的海面产生散射效应,使得其自相关特性发生改变。通常情况下,假设海面高度变化服从高斯分布,则反射信号的相关功率可表示为^[17]

$$U(\tau) = \int_{-\infty}^{\infty} U_0(\tau + \tau_z) f_{sp}(z) dz = \int_{-\infty}^{\infty} U_0(\tau + \tau_z) \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{z^2}{2\sigma^2}\right) dz \quad (3)$$

式中: $f_{sp}(z)$ 为海面高度变化的概率密度函数, z 为海洋表面高度变化量; σ 为表面高度的标准差; $U_0(\cdot)$ 为入射信号自相关功率,对于 P(Y)码而言,可近似为三角函数的平方; τ_z 为海面高度变化引起的信号相关时延,是关于海洋表面高度变化量 z 与卫星高度角 θ 的函数:

$$\tau_z = \frac{2z}{c} \sin \theta \quad (4)$$

通常,海面粗糙度采用有效波高 H_s 表示^[17],即存在 $H_s = 4\sigma$ 。从式(2)、式(3)可看出,测高灵敏度一方面反映了信号特性对测高的影响,另一方面体现了信号对海面高度变化的响应,即海面粗糙度对测高精度的影响。

根据式(3),C/A 码与 P(Y)码在不同海面粗糙度情况下的反射信号归一化处理后的相关功率仿真结果如图 1 所示。从图中可以看出,随着粗糙度的增大,信号自相关功率 $U(\tau)$ 的峰值减小,宽度增大;相比于 C/A 码,P(Y)码反射信号归一化相关功率对海面粗糙度变化更为敏感,即海面粗糙度对 P(Y)码测高精度的影响更大。

图 2 为不同信噪比差 ΔSNR ($\Delta\text{SNR} = \text{SNR}_{CA} - \text{SNR}_P$)条件下(假设 C/A 码时延波形的信噪比为 10 dB),C/A 码信号与 P(Y)码信号测高精度的比值 σ_{CA}/σ_P 随粗糙度的变化趋势。从图中可以看出,随着海面粗糙度的增大, σ_{CA}/σ_P 减小,即 P(Y)码测高精度随海面粗糙度的增大而下降的速度较 C/A 码更快,受海面粗糙度的影响更大。因此,一方面在利用 P(Y)码进行海面测高时,需对海面粗糙度引起的测高误差进行修正;另一方面可以利用 P(Y)码反射信号时延功率对海面粗糙度的敏感性进行海面有效波高的测量。P(Y)码信号的发射功率较 C/A 码低 3 dB,且其第一码片区域也较 C/A 码小,使得在相同的接收条件下,P(Y)码的接收功率低。在利用直射信号测距时,P(Y)码的测距精度较 C/A 码提高了近 10 倍,但是,利用 P(Y)码进行海面测高时,相比于 C/A 码,精度提升幅度小于 10。

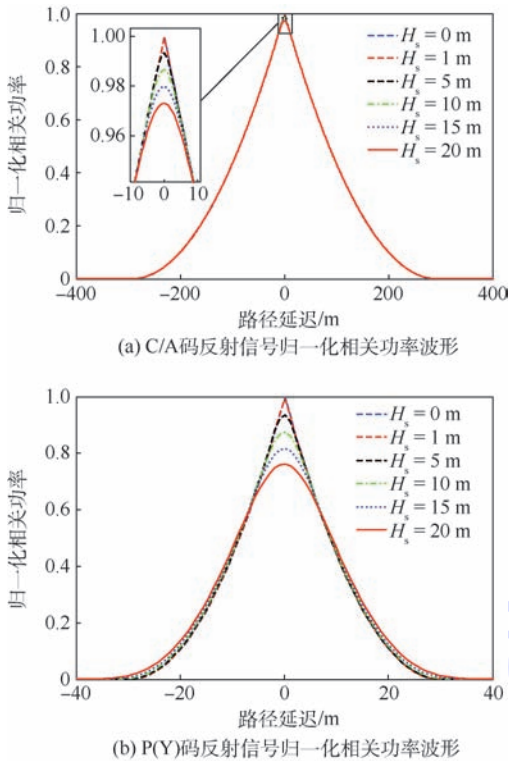


图1 不同海面粗糙度条件下的 GPS C/A 码与 P(Y) 码反射信号归一化相关功率曲线

Fig.1 Curves of normalized correlation power of GPS C/A code and P(Y) code reflected signals under different sea surface roughness conditions

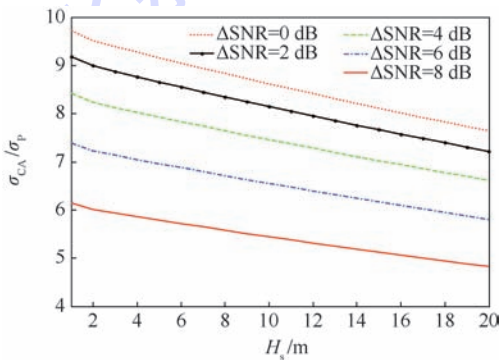


图2 不同 Δ SNR 条件下 GPS C/A 码与 P(Y) 码信号测高精度的比值随海面粗糙度的变化

Fig.2 Variation of ratio of GPS C/A code to P(Y) code signal altimetry precision with sea surface roughness under different Δ SNR conditions

2 测高方法设计与误差分析

2.1 信号处理

半无码技术是在 Y 码的确切结构和码速率未知,但是已知 Y 码是由 P 码和一种加密、未知的 W 码模二和而成,即 $Y = P \oplus W$ (P 码根据 ICD-GPS-200^[18] 可以确切得到,而 W 码根据相关

试验测得其码速率近似为 500 kHz) 的前提下,在 C/A 码已经进入稳态跟踪的情况下,首先对来波信号的 W 码进行估计,然后和本地 P 码模二和之后估计 P(Y),实现本地随机码和接收信号相关的过程。如图 3 所示,基于半无码的 P(Y) 码相关包括 L1 直射与 L1、L2 反射 3 个中频信号 S_{IF} 处理通道。在直射通道中,首先,采用闭环的方式对 GPS L1 直射信号进行跟踪,得到 L1 信号的载波频率和当前时间;然后,根据当前时间产生与 L1 信号中 P 码相位对齐的本地 P 码,将该 P 码与 L1 信号进行相关,并对相关结果进行累加估计获得携带导航电文的 W 码;最后,将估计得到的本地 W 码与本地 P 码进行模二和得到本地 P(Y) 码。反射通道复用直射通道中估计的 P(Y) 码分别与 L1、L2 两路反射信号进行相关处理,并对相关结果进行非相干累加得到反射信号时延相关波形。基于半无码的 P(Y) 码相关方法,利用本地估计的 P(Y) 码与反射信号进行相关处理,与互相关方法中将直射、反射两路弱信号直接相关相比,在接收天线增益相同的条件下可以获得更高的波形信噪比。此外,基于半无码的 P(Y) 码相关处理技术可同时对 L1 和 L2 反射信号进行开环跟踪,可实现电离层延时误差的双频校正,有效提高测高精度。

2.2 测高模型

在低海拔接收平台,如岸基或机载条件下,根据几何光学原理,反射信号相对于直射信号的几何路径时延可表示为

$$\Delta\rho = 2h\sin\theta \quad (5)$$

式中: θ 为卫星高度角; h 为接收机距离海面的高度。由于存在电离层时延误差 ρ_{atm} 、海面粗糙度引起的散射时延误差 ρ_{sca} 以及随机测量误差 ρ_n 等,实际测量的反射信号相对于直射信号的时延差可表示为

$$\rho = \Delta\rho + \rho_{\text{atm}} + \rho_{\text{sca}} + \rho_n \quad (6)$$

随着海面粗糙度的增加,反射信号时延波形的峰值后移,产生散射时延误差。为了消除或减小散射时延误差,一种有效的方法是跟踪时延波形前沿中对粗糙度不敏感点,并将不敏感点对应的时延映射到镜面反射点对应时延。在 GNSS-R 测高研究中,最接近前沿斜率峰值的点通常在时延波形峰值的 75% 处,因此通常从该点附近选取粗糙度不敏感点^[19]。本文根据模拟信号时延波形前沿斜率的特征选取波形前沿峰值的 74% 所对应的点作为对海面粗糙度不敏感点。

在高海拔接收条件下,由于反射信号和直射

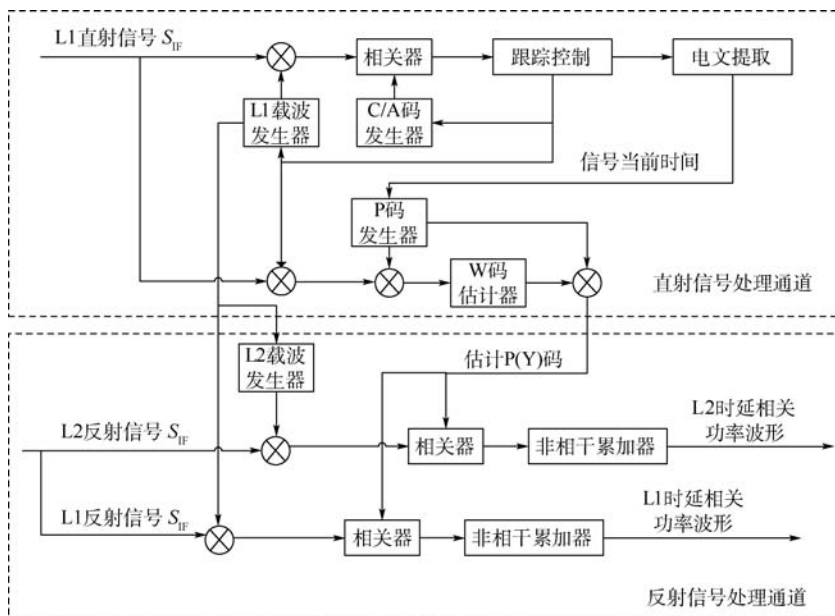


图 3 基于半无码的 P(Y)码自相关测高方法框图

Fig. 3 Block diagram of semi-codeless based P(Y) code autocorrelation altimetry method

信号经历的路径不一样,导致直射和反射信号的电离层误差无法校正,即 ρ_{aim} 不为零。根据电离层时延误差是信号频率的函数这一特点,可利用双频观测进行电离层延迟误差校正^[20]。基于半无码的 P(Y)码自相关测高方法可利用 L1、L2 信号进行双频电离层时延误差修正。若仅考虑电离层延迟误差, L1 和 L2 频点的反射信号相对于直射信号的实际时延可表示为^[11]

$$\begin{bmatrix} \rho_{L1} \\ \rho_{L2} \end{bmatrix} = \begin{bmatrix} 1 & 40.28 \frac{1}{f_{L1}^2} \\ 1 & 40.28 \frac{1}{f_{L2}^2} \end{bmatrix} \begin{bmatrix} \rho_E \\ I' \end{bmatrix} \quad (7)$$

式中: ρ_E 为仅消除了电离层误差的值; I' 为沿电磁波传播路径的电子总量; f 为载波频率。通过对式(7)进行求解,可得双频电离层时延误差修正后的反射信号相对于直射信号的几何路径时延为

$$\rho_E = -2.5457\rho_{L1} + 1.5457\rho_{L2} \quad (8)$$

假设利用 2 个频率的时延波形测量的时延精度相同,可得到利用 L1、L2 双频电离层修正后的时延测高精度为 $\sigma_{\rho_E} = 2.9783\sigma_p$, 即利用双频观测可以有效校正电离层引起的固有时延误差,但在试验测量中引入了电离层估计的随机误差,增大了时延估计的随机误差,降低了时延测量的精度。

3 机载 GNSS 反射信号测高仿真分析

3.1 机载 GNSS 反射信号的仿真

本文利用现有的 GPS L1/L2 直射信号模拟

产生如表 1 所示场景的 GNSS 海洋反射信号,对基于半无码的 P(Y)码自相关测高方法进行仿真验证。

如图 4 所示,海面反射信号可近似为不同散射单元对应的反射信号的叠加,且不同散射单元对应的反射信号具有不同的散射强度、时延和多普勒频率^[21]。在机载条件下,不同散射单元对应反射信号的多普勒频率变化较小,在仿真过程中

表 1 反射信号仿真参数

Table 1 Reflected signal simulation parameters

参数	数值
机载飞行高度/m	600
卫星高度角/(°)	30
采样频率/MHz	58
相干积分时间/ms	1
非相干累加次数	1000
风速/(m·s ⁻¹)	3

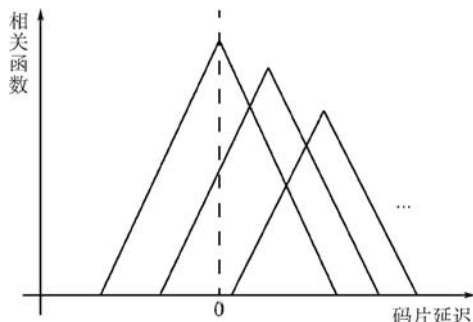


图 4 等时延信号相关结果示意图

Fig. 4 Schematic diagram of iso-delay signal correlation results

可以假设所有散射单元对应反射信号的多普勒频率等于镜面反射点反射信号的多普勒频率。具体仿真流程如下：

1) 根据几何关系计算得到镜面反射点处的反射信号相对于直射信号的时延 τ 。

2) 以镜面反射点处的时延 τ 为基准,以 $\Delta\tau$ 为不同散射单元的时延间隔(本文设置为 0.5 个 P 码长度),将实采的直射 L1、L2 信号分别做 n (本文取 $n = 16$) 路等间隔时延,得到 n 路等间隔时延信号。

3) 对如表 1 所示场景条件下仿真的 Z-V 模型进行拟合得到 n 路不同时延信号的幅度权重^[22],即

$$\{a_1, a_2, \dots, a_n\} = \arg \min \left\{ \left[\sum_{i=1}^n a_i A_i(\tau - i\Delta\tau) - Y(\tau) \right]^2 \right\} \quad (9)$$

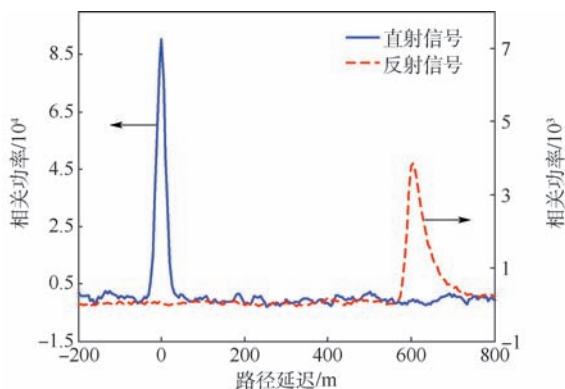
式中: $a_i A_i(\tau - i\Delta\tau)$ 为第 i 路反射信号的相关函数; $Y(\tau)$ 为 Z-V 模型仿真的时延波形。

4) 将 n 路等间隔时延信号分别加权后进行累加得到模拟的反射信号。

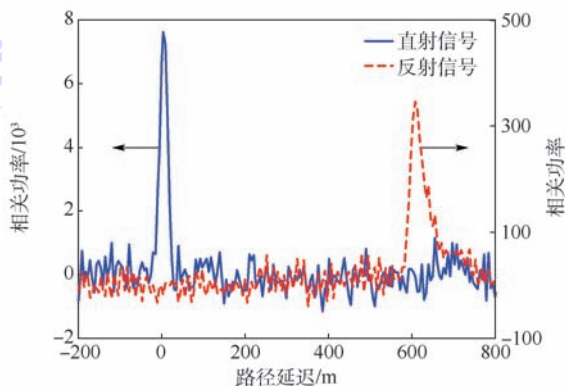
5) 利用基于半无码的 P(Y) 码相关方法对直射 L1 信号和模拟的 L1、L2 反射信号进行相关处理得到时延波形。

3.2 测高精度分析

本文选择 GPS 15 号卫星信号对基于半无码的 P(Y) 码自相关、C/A 码自相关和信号互相关测高技术进行对比分析。图 5 为基于半无码的 P(Y) 码自相关和信号互相关处理得到的时延相关功率,其中相干积分时间为 1 ms,非相干累加次数为 1000。通过计算可得,相比于信号互相关时延波形,基于半无码的 P(Y) 码相关时延波形具有更高的信噪比,提高了 10.4 dB(见图 5)。图 6 为处理 20 s 数据得到的 20 组基于半无码的 P(Y) 码自相关、C/A 码自相关以及信号互相关海面高度反演的测量结果。从图中可以看出,相比于 C/A 自相关和信号互相关海面高度反演结果,基于半无码的 P(Y) 码自相关反演的海面高度具有更小的波动幅度,即测高精度更高。如表 2 为海面高度反演的均方根误差。在相同的接收环境下,基于半无码的 P(Y) 码自相关比 C/A 码自相关和信号互相关的海面高度测量精度分别提高了 3.97 倍和 1.47 倍。



(a) 基于半无码的 P(Y) 码自相关时延



(b) 信号互相关时延

图 5 时延相关功率波形

Fig. 5 Delay correlation power waveform

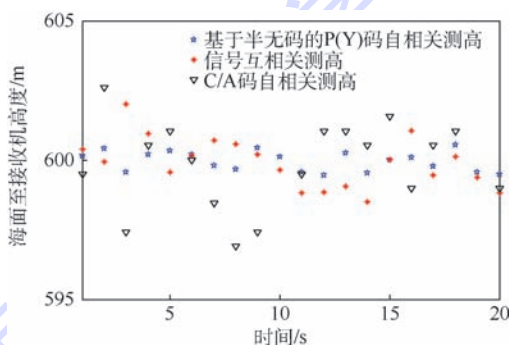


图 6 3 种测高方法处理模拟信号得到的高度测量结果

Fig. 6 Analog signal processing results of three altimetry methods

表 2 3 种测高方法处理模拟信号得到的高度均方根误差

Table 2 Height mean square root errors obtained by analog signal processing using three altimetry methods

测高方法	均方根误差/cm
基于半无码的 P(Y) 码自相关测高	36.2202
信号互相关测高	89.3739
C/A 码自相关测高	180.0400

4 结 论

本文对基于半无码的 P(Y)码自相关 GNSS-R 海面测高方法进行了研究,首先介绍了测高精度模型,并对比分析了 C/A 码和 P(Y)的测高精度;然后描述了基于半无码的 P(Y)码自相关处理架构、测高模型以及双频电离层时延误差修正方法;最后,建立了 L1、L2 反射信号模拟平台,对基于半无码的 P(Y)码自相关、C/A 码自相关以及信号互相关海面测高方法进行了对比分析。结果表明:

1) 在相同的接收环境下,基于半无码的 P(Y)码自相关时延波形比信号互相关时延波形具有更高的信噪比。

2) 相比于 C/A 码自相关和信号互相关测高结果,基于半无码的 P(Y)码自相关测高精度分别提高了 3.97 倍和 1.47 倍。

基于 P(Y)码反射信号时延相关功率对有效波高的敏感性,本文后续将进一步研究该方法在海面有效波高测量中的应用。

参考文献 (References)

- [1] 杨东凯,张其善. GNSS 反射信号处理基础与实践 [M]. 北京:电子工业出版社,2012:7-14.
YANG D K, ZHANG Q S. GNSS reflected signal processing: Fundamentals and applications [M]. Beijing: Publishing House of Electronics Industry, 2012:7-14 (in Chinese).
- [2] LOWE S T, ZUFFADA C, CHAO Y, et al. 5-cm-precision aircraft ocean altimetry using GPS reflections [J]. Geophysical Research Letters, 2002, 29 (10): 1375.
- [3] ZAVOROTNY V U, VORONOVICH A G. Scattering of GPS signals from the ocean with wind remote sensing application [J]. IEEE Transactions on Geoscience & Remote Sensing, 2002, 38 (2): 951-964.
- [4] KOMJATHY A, MASLANIK J, ZAVOROTNY V U, et al. Sea ice remote sensing using surface reflected GPS signals [C] // IEEE International Geoscience and Remote Sensing Symposium. Piscataway, NJ: IEEE Press, 2000: 2855-2857.
- [5] MASTERS D, AXELRAD P, KATZBERG S. Initial results of land-reflected GPS bistatic radar measurements in SMEX02 [J]. Remote Sensing of Environment, 2004, 92 (4): 507-520.
- [6] MARTIN-NEIRA M. A passive reflectometry and interferometry system (PARIS) application to ocean altimetry [J]. ESA Journal, 1993, 17 (4): 331-355.
- [7] ROUSSEL N, RAMILLIEN G, FRAPPART F, et al. Sea level monitoring and sea state estimate using a single geodetic receiver [J]. Remote Sensing of Environment, 2015, 171: 261-277.
- [8] RIUS A, CARDELLACH E, MARTIN-NEIRA M. Altimetric analysis of the sea-surface GPS-reflected signals [J]. IEEE Transactions on Geoscience & Remote Sensing, 2010, 48 (4): 2119-2127.
- [9] MARTIN-NEIRA M, COLMENAREJO P, RUFFINI G, et al. Altimetry precision of 1 cm over a pond using the wide-lane carrier phase of GPS reflected signals [J]. Canadian Journal of Remote Sensing, 2002, 28 (3): 394-403.
- [10] WILMHOF B, LALEZARI F, ZAVOROTNY V, et al. GPS ocean altimetry from aircraft using the P(Y) code signal [C] // IEEE International Geoscience and Remote Sensing Symposium. Piscataway, NJ: IEEE Press, 2007: 5093-5096.
- [11] MARTIN-NEIRA M, D'ADDIO S, BUCK C, et al. The PARIS ocean altimeter in-orbit demonstrator [J]. IEEE Transactions on Geoscience & Remote Sensing, 2011, 49 (6): 2209-2237.
- [12] CARRENO-LUENGO H, CAMPS A, RAMOS-PÉREZ I, et al. Experimental evaluation of GNSS-reflectometry altimetric precision using the P(Y) and C/A signals [J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2014, 7 (5): 1493-1500.
- [13] LOWE S T, MEEHAN T, YOUNG L. Direct signal enhanced semicodeless processing of GNSS surface-reflected signals [J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2017, 7 (5): 1469-1472.
- [14] WAGNER C, KLOKOČNÍK J. The value of ocean reflections of GPS signals to enhance satellite altimetry: Data distribution and error analysis [J]. Journal of Geodesy, 2003, 77 (3-4): 128-138.
- [15] LI W, RIUS A, FABRA F, et al. Revisiting the GNSS-R waveform statistics and its impact on altimetric retrievals [J]. IEEE Transactions on Geoscience & Remote Sensing, 2018, 56 (5): 2854-2871.
- [16] 于成辉. 基于 GNSS 反射信号的测高精度增强方法研究与实现 [D]. 北京:北京航空航天大学, 2013.
YU C H. Research and implementation of achieving better altimetry precision based on GNSS-R [D]. Beijing: Beihang University, 2013 (in Chinese).
- [17] HAJJ G A, ZUFFADA C. Theoretical description of a bistatic system for ocean altimetry using the GPS signal [J]. Radio Science, 2016, 38 (5): 1089.
- [18] GPS Joint Program Office. Navstar GPS space segment/navigation user interfaces: IS-GPS-200 [S]. Washington: Navstar GPS Space Segment, 2013: 21-31.
- [19] CARDELLACH E, RIUS A, MARTÍN-NEIRA M, et al. Consolidating the precision of interferometric GNSS-R ocean altimetry using airborne experimental data [J]. IEEE Transactions on Geoscience & Remote Sensing, 2014, 52 (8): 4992-5004.
- [20] 严伏朝,王永平,谢世杰. GPS 定位的电离层误差 [J]. 测绘通报, 2000 (9): 10-12.
YANG F C, WANG Y P, XIE S J. Ionospheric error in positioning with GPS [J]. Bulletin of Surveying and Mapping, 2000 (9): 10-12 (in Chinese).
- [21] 王晔,孟斌,李潇,等. 一种 GPS 海洋反射信号测试系统 [C] // 第三届中国卫星导航学术年会电子文集, 2012: 44-47.
WANG Y, MENG B, LI X, et al. A GNSS-R testing system based on GPS signal generator [C] // Third session China Satellite Navigation Conference Corpus, 2012: 44-47 (in Chinese).

- [22] ZAVOROTNY V U, VORONOVICH A G. Scattering of GPS signals from the ocean with wind remote sensing application[J]. IEEE Transactions on Geoscience & Remote Sensing, 2000, 38(2): 951-964.

洋遥感中的应用。

作者简介:

樊梦文 女, 硕士研究生。主要研究方向: 卫星导航及其在海

张波 男, 博士, 讲师。主要研究方向: 扩频通信、卫星导航及其应用等。

王峰 男, 博士研究生。主要研究方向: 导航卫星信号处理及其在海洋遥感中的应用。

Semi-codeless based P(Y) code autocorrelation GNSS-R sea surface altimetry method

FAN Mengwen, ZHANG Bo*, WANG Feng

(School of Electronic and Information Engineering, Beihang University, Beijing 100083, China)

Abstract: In this paper, the semi-codeless based P(Y) code autocorrelation sea surface altimetry method was studied due to the fact that the conventional GPS C/A code delay altimetry method is less precise, and that the signal-to-noise ratio of delay waveform in interference altimetry method is low. First, the height measurement accuracy of the C/A code and P(Y) code was compared and analyzed through theoretical model for describing the height measurement accuracy. Then, the architecture of semi-codeless based P(Y) code autocorrelation processing for L1 and L2 dual-frequency reflected signals was described. Meanwhile, the altimetry model and dual-frequency ionospheric time-delay error correction methods are discussed. Finally, using simulated L1 and L2 reflection signals, the measurement accuracy is compared and analyzed regarding the proposed method, C/A code autocorrelation method and signal cross-correlation method. The results show that the accuracy of the proposed method is 3.97 times and 1.47 times higher than that of the C/A code autocorrelation method and signal cross-correlation method, respectively.

Keywords: semi-codeless; P(Y) code; sea surface altimetry; dual-frequency; ionospheric delay correction

Received: 2018-05-21; Accepted: 2018-07-28; Published online: 2018-08-27 09:21

URL: kns.cnki.net/kcms/detail/11.2625.V.20180824.1525.004.html

Foundation item: National Natural Science Foundation of China (41774028)

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

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

DOI: 10.13700/j.bh.1001-5965.2018.0297

基于缓解 HoL 堵塞的单组播混合调度算法

袁龙, 熊庆旭*, 萧翰

(北京航空航天大学 电子信息工程学院, 北京 100083)



摘 要: 针对联合输入交叉队列(CICQ)结构的单组播混合调度研究不多,且没有针对性研究头分组(HoL)堵塞问题,提出了以缓解HoL堵塞为目标的一种新的单组播混合调度算法,即单组播低HoL堵塞(MULHB)算法,使交换机尽量逼近work-conserving状态。该算法还充分考虑了单组播之间的差异性,利用权重裁决单组播之间的竞争,避免“饿死”现象发生。同时,还给出了一种新的组播分组入队算法,即动态组播分组入队(DMQ)策略,该策略在不乱序的前提下,允许新到达分组选择合适的队列入队。仿真结果表明,在不同业务下,DMQ-MULHB算法的通过率及平均时延均优于现有主流的单组播混合调度算法,尤其在非均匀业务下,该算法性能接近输出排队(OQ)调度。

关键词: 分组交换; 联合输入交叉队列(CICQ); work-conserving; 调度算法; 组播; HoL堵塞

中图分类号: TP393

文献标识码: A **文章编号:** 1001-5965(2019)02-0405-08

随着网络应用的发展,网络所承载的业务类型从单一类型逐渐向多元化发展,交换机不仅需要支持单播业务或纯组播业务分组转发,还需要有效调度单组播混合业务。

输出排队(Output Queued, OQ)结构的通过率为100%,可提供QoS确保功能,但存在 N 倍加速比,不易于扩展^[1]。输入排队(Input Queued, IQ)结构加速比为1,可扩展性强,但需采用虚拟输出排队(Virtual Output Queuing, VOQ)^[2]结构克服头分组(Head of Line, HoL)堵塞^[3]问题。对于单播业务只需在每个输入端口配置 N 个虚拟队列,但组播业务分组扇出类型多达 $2^N - 1$ 种,则需要每个输入端口放置 $2^N - 1$ 个队列^[4],难以在大规模交换机中实用化。目前的研究通常配置 k ($1 < k \ll 2^N - 1$)个组播队列,以平衡交换机性能和复杂度,但无法完全避免HoL堵塞。

随着集成电路技术的发展,联合输入交叉排

队(Combined Input and Crossbar Queued, CICQ)被普遍认为最值得期待的高性能分组交换结构,但CICQ结构单组播混合业务调度算法研究不多。早期主要以单播调度或组播调度为主,在单播调度中^[5-7],文献[7]提出的交叉缓存队列均衡(Crossbuffer Queue Balance, CQB)算法从定量的角度出发,以交换机实现work-conserving状态为追求目标,使交换机的性能得到显著提升。在组播调度中,Mhamdi和Hamdi提出了组播交叉结点轮询(Multicast Cross-Points Round Robin, MXRR)算法^[8],由于输入端口中仅采用单FIFO(First In First Out)队列,存在严重的HoL堵塞问题。文献[9]提出了轮询(Round Robin, RR)、最小扇出优先(Minimum Residue First, MRF)、最大服务优先(Maximum Service First, MSF)以及最大服务率优先(Maximum Ratio of Service First, MRSF)等调度算法,并在输入端口采用多个组播队列,从结构

收稿日期: 2018-05-23; 录用日期: 2018-08-10; 网络出版时间: 2018-09-04 09:09

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180830.1513.003.html

基金项目: 国家自然科学基金(61271196)

*通信作者. E-mail: qxxiong@buaa.edu.cn

引用格式: 袁龙, 熊庆旭, 萧翰. 基于缓解HoL堵塞的单组播混合调度算法[J]. 北京航空航天大学学报, 2019, 45(2): 405-412.
YUAN L, XIONG Q X, XIAO H. Packet scheduling algorithm for integrated unicast and multicast traffic based on reducing HoL blocking[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 405-412 (in Chinese).

上缓解了组播 HoL 堵塞,提高了交换机的通过率降低了平均分组时延。为了满足单播组播混合业务调度的需求,Mhamdi 和 Vassiliadis 提出了单组播轮询调度 (Multicast and Unicast Round robin Scheduling, MURS) 算法^[10],采用不同的优先级轮询调度单组播业务。虽然算法复杂度低,但整体性能欠佳。Yi 等提出的低代价组播调度 (Low-Cost Multicast Scheduling, LCMS) 算法^[11],考虑了单组播业务之间的差异性,基于权重裁决单组播之间的竞争,算法的性能有一定提升,但输入端口仅采用一个组播队列,导致了较严重的 HoL 堵塞。梁佳诚等提出的均衡交叉节点缓存单组播调度 (Multicast and Unicast Crossbuffer Balance, MUCB) 算法^[12],与以往基于轮询或者基于权重的调度算法不同,而是采用了均衡交叉节点缓存占用的思路,使交换机最大程度地逼近 work-conserving 状态。虽然 MUCB 算法使交换机的性能有了较好的改善,但没有采取有效机制以缓解 HoL 堵塞。

为了缓解组播 HoL 堵塞问题,通常在输入端口配置 k 个组播队列。因此,新到达分组需要按照一定的入队策略选择合适的队列入队。在已有研究中,多数基于三大准则^[9]提出相应的入队策略,主要分为 2 类:

1) 静态入队策略,如文献[13]提出 Module 入队策略,利用分组扇出数取余的方法选择队列,即: $i = F \bmod k$; 实现简单,但扇出数目相同却不相关的分组会进入同一缓存队列,且对于小扇出分组均衡性差。文献[14]提出加权取模 (Weighted Module, WM) 入队策略改善了 Module 中小扇出组播均衡效果差的问题。文献[15]提出特征向量法 (vector) 选择分组扇出与队列特征向量距离最小的队列入队,增加了头分组扇出差异性,但负载均衡的效果一般。

2) 动态入队策略,如文献[9]提出的面向分组轮询分配 (Cell oriented Round-Robin Assignment, CRRA)、面向突发轮询分配 (Burst oriented Round-Robin Assignment, BRRA)、面向最短队列分组优先 (Cell oriented Shortest Queue First, CSQF) 和面向最短突发队列优先 (Burst oriented Shortest Queue First, BSQF) 等入队策略,存在分组失序问题。

本文通过研究 HoL 堵塞与 work-conserving 之间的关系,基于 CICQ 结构提出一种新的单组播混合调度算法,即单组播低 HoL 堵塞 (Multicast and Unicast with Low HoL Blocking, MULHB) 算

法,其基本思想就是以缓解 HoL 堵塞为目标,使交换机逼近 work-conserving 状态。同时,为了进一步缓解 HoL 堵塞,本文提出一种新的组播分组入队策略,即动态组播分组入队 (Dynamic Multicast Queuing, DMQ) 策略,使新到达分组根据缓存队列的状态可以动态选择合适的队列入队。

本文首先对 CICQ 结构单组播混合调度中的主要问题进行了分析;然后,具体说明了 DMQ 策略以及 MULHB 算法;最后,给出了本文提出算法的仿真结果,并和当前主流的单组播混合调度算法进行比较和分析。

1 问题分析

本文讨论的 CICQ 交换结构如图 1 所示。每个输入端口分别配置 N 个单播队列和 k ($1 < k \leq 2^N - 1$) 个组播队列。这种缓存配置方式相比较于单 FIFO 配置方式的性能有了较好的提升,但由于 k 个组播队列不能完全覆盖组播分组所有的扇出类型,无法彻底解决组播 HoL 堵塞问题,仅能从结构上缓解该问题。

CICQ 结构 work-conserving 状态是指,任一时隙中若交换机中存在去往某个输出端口的分组,则该时隙必有分组离开该输出端口。由于输入端口中组播队列的数目受到限制,因此,在调度过程中,属于同一组播队列的分组扇出去向不完全一致。此时,若组播队列的后续分组中存在使交换机满足 work-conserving 状态的扇出去向,而头分组中不存在,则发生了由 HoL 堵塞引起的非 work-conserving 问题。由于考虑整个组播队列中全部后续分组的扇出过于复杂,本文算法中仅考虑头分组对次分组的堵塞造成的影响。

为了尽量解决由 HoL 堵塞引起的非 work-conserving 状态,在每一时隙中,分组入队后、输入调度前,应使所有输入端口中存在若干不属于同一端口的头分组的扇出去向可以完全覆盖所有传输需求但对应节点缓存列为空的输出端口。为了达到上述条件,首先,可以通过入队策略动态调整新到达分组在队列排队状态。已有实用的静态入队策略由于选择队列的方式固定,分组入队后,相同扇出类型的分组不能根据调度的需求重新选择合适的队列入队。同时,仿真统计发现在调度过程中,输入端口内的缓存队列出现大量为空的状态。因此,入队策略应当使新到达分组根据当前缓存队列的状态动态选择合适的队列入队,尤其存在可选的空队列时,应充分选择空队列入队,使输入端口中头分组的扇出差异性增大。其次,为

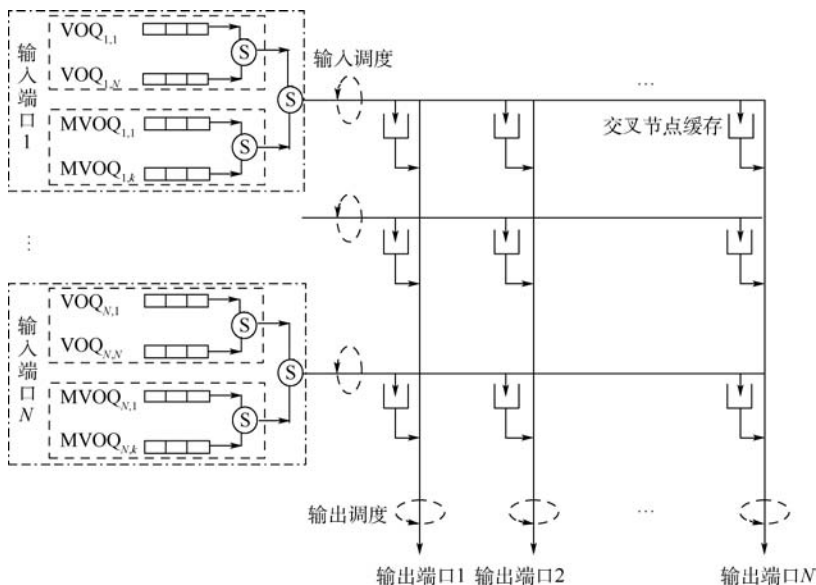


图 1 CICQ 交换结构

Fig. 1 CICQ switch architecture

为了使交换机达到高通过率及低时延的目标,应尽可能使交换机处于 work-conserving 状态。从文献[7]中给出交换机实现 work-conserving 的条件可知,在输入调度之后,所有的有传输需求的输出端口对应节点缓存列均不为空,从而使得输出端口总有分组离开。在此基础上,输入调度后,应使有助于交换机逼近 work-conserving 状态的次分组尽快成为头分组。最后,输出调度应尽快调度节点缓存中可能导致发生 HoL 堵塞的分组离开交换机。

此外,在输入端口中采用的缓存配置方式对单播和组播分组进行隔离处理。由于 2 种业务之间存在一定的差异性,当输入调度在选择分组进入节点缓存内时,需要采用一定的裁决机制解决单、组播分组之间的竞争问题,以避免“饿死”现象发生。

2 DMQ-MULHB 调度算法

由第 1 节分析可知,为使交换机避免发生由 HoL 堵塞引起的非 work-conserving 状态,本文采用新的入队策略控制分组入队,同时结合相应的输入输出调度算法使交换机尽可能在分组入队后、输入调度前,所有输入端口中存在若干不属于同一端口的头分组的扇出去向可以完全覆盖有传输需求但对应节点缓存列为空的输出端口。本节将从入队策略、输入调度和输出调度等环节详细分析 DMQ-MULHB 调度算法的设计。

2.1 入队策略

本文提出的 DMQ 策略在处理组播分组入队

时,对于任意输入端口 i ,首先,为 k 个组播队列分别构造“特征向量” \mathbf{v}_k ,任意 $i \neq j, \exists \|\mathbf{v}_i \cdot \mathbf{v}_j\| = 0$ 。其次,采用长度为 $2^N - N - 1$ 的 2 个数组(哈希缓存数组 H 和哈希序号数组 F)分别记录输入端口中不同扇出类型分组的缓存数量及各自的入队序号,初始值均为 0。然后,将分组扇出对应一个二进制串 $L = [x_0, x_1, \dots, x_j, \dots, x_N], x_j = 1$ 表示分组有去往输出端口 j 的扇出,否则 $x_j = 0$ 。因此,利用哈希思想,任意分组的扇出对应的二进制串可转成唯一对应的十进制数作为该组播分组的查询索引值 p 。

当输入端口 i 有组播分组到达时,选择入队列表的具体步骤如下:

- 1) 分组到达时,若 $H[p] \neq 0$,选择队列号等于 $F[p]$ 值的队列入队,入队结束;否则进入步骤 2)。
- 2) 若存在空队列,选择分组扇出与队列特征向量差异最小的队列,进入步骤 5);否则进入步骤 3)。
- 3) 若存在队长为 1 的队列,选择分组与头分组扇出差异最小的队列,进入步骤 5);否则进入步骤 4)。
- 4) 选择分组与队尾分组扇出差异最小的队列;若存在多个,则选择最短队列。
- 5) $F[p]$ 更新为所选队列序号, $H[p]$ 加 1,选择 $F[p]$ 对应队列入队,入队结束。当索引值为 p 的头分组在输入端口完全扇出时, $H[p]$ 减 1。

由上述步骤可知,首先,步骤 1) 限定了新到达分组重新选择队列入队的条件,只有当该输入端口中相同扇出类型的分组缓存数目为零时,才

能重新选择队列入队,该机制避免发生乱序;其次,步骤2)允许重新选择队列的新到达组播分组进入空队列增加端口中头分组扇出差异,使队列头分组获得更多调度机会,以缓解了 HoL 堵塞,使交换机尽量逼近 work-conserving 状态。最后,步骤3)~步骤5)对于缓存分组具有均衡作用,同时尽量减少相邻分组间的扇出差异,当前一分组成为头分组时,可以缓解头分组对次分组造成的 HoL 堵塞。

2.2 MULHB 调度算法

本文提出 MULHB 算法,在输入调度过程中,首先,找出有传输需求且对应节点缓存列为空的输出端口;其次,找出扇出去向包含上述输出端口且头分组扇出去向类型数最少的输入端口;再次,从上述输入端口中找出满足工作保持状态的头分组;最后,从中优先选择可以完全扇出且对应次分组有助于交换机逼近 work-conserving 状态的头分组进行传输。在输入调度中尽量使交换机既在当前时隙实现 work-conserving 状态,又使输入端口中头分组扇出在后续时隙尽可能覆盖所有有传输需求的输出端口,从而避免交换机发生由 HoL 堵塞引起的非 work-conserving 状态。另外,在裁决单、组播分组之间的竞争时,组播分组的权重以头、次分组扇出、节点缓存状态以及头分组等待时间作为权重因子,单播分组的权重以等待时间作为权重因子。基于权重比较完成,避免“饿死”现象发生。在输出调度过程中,输出调度优先选择其次分组有助于交换机在后续时隙逼近 work-conserving 状态的头分组对应的节点缓存中的分组进行输出。

方便起见,下面给出算法中相关符号的含义: O 为输出端口集合,且按端口号升序排列; I 为输入端口集合,且按端口号升序排列; V_{ij} 为输入端口 i 中去往输出端口 j 的单播队列; M_{ik} 为输入端口 i 中第 k 个组播队列; X_{ij} 为输入端口 i 和输出端口 j 对应交叉节点缓存; b_j 为输出端口 j 对应节点缓存列中分组数目; M_i 为输入端口 i 中所有组播队列; J 为头分组扇出包含输出端口 j 的所有输入端口集合; X_{ij} 为空,同时 V_{ij} 不为空或 M_i 中所有非空头分组中有去往输出端口 j ; D_i 为输入端口 i 所有头分组扇出种类数; D_{ij}^V 为 V_{ij} 头分组的等待时间; W_{ij}^V 为 V_{ij} 头分组权重, $W_{ik}^V = D_{ik}^V + 1$; w 为组播权重因子; A_{ik} 为 M_{ik} 头分组到达时扇出数; C_{ik} 为 M_{ik} 头分组当前扇出数; D_{ik}^M 为 M_{ik} 头分组等待时间; W_{ij}^M 为 M_{ik} 头分组权重, $W_{ik}^M = 2wA_{ik}A_{ik} \lceil C_{ik}(D_{ik}^M + 1) \rceil$, w 设为 0.5; W_{ij}^X 为 X_{ij} 分组权重, $W_{ij}^X = W_{ij}^V + W_{ik}^M$ 。其

中 M_{ik} 扇出去向包含输出端口 j ; $Y = \{y_j\}$, 若 $b_j = 0$ 且 $\sum_{i=1}^N q_{ij} > 0$, 则 $y_j = 1$, 否则 $y_j = 0$; R_{ik} 为 M_{ik} 的次分组和头分组可调度扇出去向并集数目与头分组可调度扇出去向数目比值; T_{ik} 为 M_{ik} 头分组可调度扇出去向数目与剩余扇出去向数目的比值; S_{ik} 为 M_{ik} 次分组扇出去向包含有传输需求且对应节点缓存列中分组数小于等于 1 的输出端口数。

输入调度算法:

1) 初始化集合 O 包含所有输出端口,集合 I 包含所有输入端口。

2) 若集合 O 为空,则结束该时隙输入调度。

3) 从集合 O 的第 1 个元素开始,找出 $\max\{y_j\}$ 对应输出端口 j ,若 $\max\{y_j\} = 0$,进入步骤 11)。

4) 若集合 J 为空,从集合 O 中剔除输出端口 j ,回到步骤 2)。

5) 从 J 的第 1 个元素开始,选择其中 $\min\{D_i\}$ 对应输入端口 i ;找出输入端口 i 中存在扇出去向包含输出端口 j 且可完全扇出的组播头分组,进入下一步;否则,进入步骤 8)。

6) 若步骤 5) 找出的头分组中存在次分组满足 $S_{ik} > 0$,则选择 $\max\{S_{ik}\}$ 对应的组播队列,进入步骤 9);否则,进入步骤 7)。

7) 若步骤 5) 找出的头分组对应的次分组满足 $\max\{S_{ik}\} = 0$,则从中选择 $\max\{R_{ik}\}$ 对应的组播队列;若存在多个,则选择 $\max\{D_{ik}^M\}$ 对应的组播队列,进入步骤 9)。

8) 若输入端口 i 中扇出去向包含输出端口 j 的头分组均不可完全扇出,则选择 $\max\{T_{ik}\}$ 对应的组播队列;若存在多个 $\max\{T_{ik}\}$,则选择 $\max\{R_{ik}\}$ 对应的组播队列;若存在多个 $\max\{R_{ik}\}$,则选择 $\max\{D_{ik}^M\}$ 对应组播队列,进入步骤 9)。

9) 若 V_{ij} 头分组非空,则调度 $\max\{W_{ij}^V, W_{ij}^M\}$ 对应的头分组,否则调度 $\max\{W_{ij}^M\}$ 对应的头分组。

10) 从集合 I 中剔除输入端口 i ,更新 U, J, Y ,回到步骤 3)。

11) 若集合 I 为空,该时隙输入调度结束。

12) 从集合 I 的第 1 个元素开始,若 M_i 中存在可完全扇出的头分组,进入步骤 13),否则,进入步骤 15)。

13) 若 M_i 中存在次分组满足 $S_{ik} > 0$,则选择 $\max\{S_{ik}\}$ 对应的组播队列,进入步骤 16);否则,进入步骤 14)。

14) 若 $\max\{S_{ik}\} = 0$,则选择 $\max\{R_{ik}\}$ 对应的组播队列;若存在多个,则选择 $\max\{D_{ik}^M\}$ 对应

的组播队列,进入步骤 16)。

15) 若 M_i 中头分组均不可完全扇出,则选择 $\max\{T_{ik}\}$ 对应的组播队列;若存在多个 $\max\{T_{ik}\}$,则选择 $\max\{R_{ik}\}$ 对应的组播队列;若存在多个 $\max\{R_{ik}\}$,则选择 $\max\{D_{ik}^M\}$ 对应的组播队列,进入步骤 16)。

16) 若 V_{ij} 头分组非空,调度 $\max\{W_{ij}^V, W_{ij}^M\}$ 对应的头分组,否则调度 $\max\{W_{ij}^M\}$ 对应的头分组;回到步骤 10)。

输出调度算法:

1) 初始化集合 O 包含全部输出端口,集合 I 包含全部输入端口。

2) 若集合 O 为空,则结束该时隙输出调度。

3) 从 O 的第 1 个元素开始,从输入端口中找出所有包含输出端口 j 扇出去向且状态不为空的 X_{ij} 对应的队列;若存在,进入下一步;否则进入步骤 7)。

4) 若步骤 3) 中找出的队列中存在次分组满足 $S_{ik} > 0$,则选择 $\max\{S_{ik}\}$ 对应的组播队列;否则,进入步骤 6)。

5) 若 $\max\{S_{ik}\}$ 对应的队列存在多个,则选择 $\max\{R_{ik}\}$ 对应的输入端口 i ;若存在多个,则选择 $\max\{T_{ik}\}$ 对应的输入端口 i ;若存在多个,则选择 $\max\{W_{ij}^M\}$ 对应的输入端口 i 。

6) 输出 X_{ij} 中分组,从集合 O 中剔除输出端口 j ,回到步骤 2)。

7) 从非空 X_{ij} 中选择 $\max\{W_{ij}^X\}$ 对应的分组输出,从集合 O 中剔除输出端口 j ,回到步骤 2)。

3 仿真及对比分析

本节将基于 CICQ 交换结构对 DMQ-MULHB 算法进行仿真并对其性能进行评估。仿真采用 Bernoulli 和 ON-OFF 两种业务作为业务源模型。交换结构的端口规模为 16×16 ,仿真时间为 100 万个时隙,每个输入端口设有 8 个组播队列。其中 ON-OFF 业务的平均突发长度为 16。为了方便对性能进行分析,同时给出了最大扇出优先和最大服务比率优先(Maxfanout First and Maximum Ratio of Service First, MF-MRSF)算法、MUCB 算法以及 OQ 调度的仿真结果进行对比,其中 MF-MRSF 算法、MUCB 算法均采用 Module 入队策略。

3.1 流量模型

在本次仿真中,假设单播业务占比 f_u ,组播业务占比 f_m ,组播分组平均扇出为 $|\phi|$,输入负载为 λ ,输出负载为 μ ,仿真结果中的负载均为输出负

载,在单组播混合调度中,输入与输出负载之间的关系满足 $\mu = \lambda(f_u + |\phi|f_m)$ 。

仿真中的非均匀业务为弱对角业务,对于输入端口 i ,其到达分组去向分布为

$$u_{ij} = \begin{cases} u_i(1 - \omega)/N & j = i \\ u_i[\omega + (1 - \omega)/N] & j \neq i \end{cases}$$

其中: u_{ij} 为输入端口 i 的分组到达时,去往输出端口 j 的概率; u_i 为输入负载; ω 为非均匀因子,仿真中将 ω 设为 0.5。

3.2 均匀业务

低负载情况下,不同算法的性能差异性相对较小,本文给出高负载情况下的通过率及平均时延对比。

表 1、表 2 给出了均匀业务下, $f_m = 0.8$ 时,各算法在归一化负载为 0.90、0.95、0.99 的通过率。可以看出,不同负载下, MULHB 算法的通过率高于 MF-MRSF 及 MUCB 算法。

图 2 给出了均匀业务下平均时延的对比结果。从图中可以看出, MULHB 算法的时延性能优于 MF-MRSF 及 MUCB 算法。

表 1 均匀 Bernoulli 业务通过率

算法	归一化负载		
	0.90	0.95	0.99
MF-MRSF	0.9999876	0.9999887	0.9910755
MUCB	0.9999844	0.9999905	0.9999422
MULHB	0.9999863	0.9999913	0.9999651
OQ	0.9999890	0.9999939	0.9999671

表 2 均匀 ON-OFF 业务通过率

算法	归一化负载		
	0.90	0.95	0.99
MF-MRSF	0.9854138	0.9577737	0.9373589
MUCB	0.9988132	0.9965875	0.9849949
MULHB	0.9994254	0.9985551	0.9881961
OQ	0.9995213	0.9992420	0.9943861

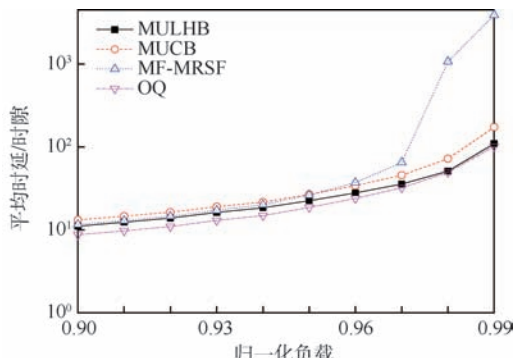
3.3 非均匀业务

表 3、表 4 给出了非均匀业务下, $f_m = 0.8$ 时,不同算法的通过率。MULHB 算法的通过率要高于 MF-MRSF 及 MUCB 算法。

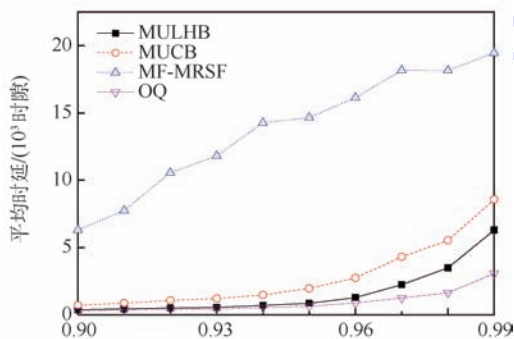
图 3 给出了非均匀业务下平均时延的对比结果。从图中可以看出, MULHB 算法的时延性能显著优于 MF-MRSF 及 MUCB 算法,并且接近 OQ 调度。

通过仿真结果对比可以看出,无论是在均匀

业务下还是非均匀业务下,DMQ-MULHB 算法的性能相对于 MUCB 算法、MF-MRSF 算法,都有了明显的提升。尤其在非均匀业务下,DMQ-MULHB 算法的性能接近 OQ 调度。



(a) 均匀Bernoulli业务平均时延



(b) 均匀ON-OFF业务平均时延

图2 均匀 Bernoulli 和 ON-OFF 业务平均时延

Fig.2 Average delay under uniform Bernoulli and ON-OFF traffic

表3 非均匀 Bernoulli 业务通过率

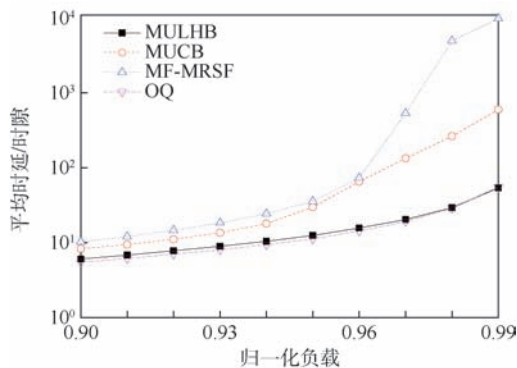
Table 3 Throughput under non-uniform Bernoulli traffic

算法	归一化负载		
	0.90	0.95	0.99
MF-MRSF	0.999 993 8	0.999 961 1	0.980 287 5
MUCB	0.999 994 0	0.999 974 9	0.999 345 6
MULHB	0.999 995 5	0.999 989 6	0.999 944 4
OQ	0.999 996 0	0.999 990 6	0.999 949 7

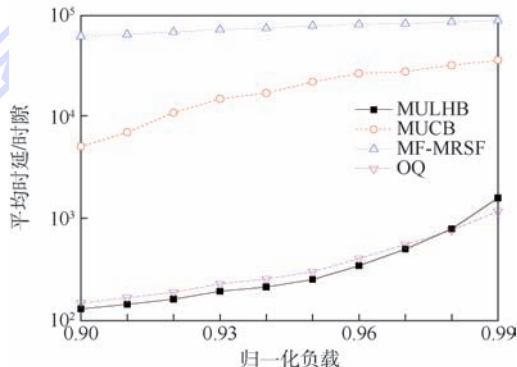
表4 非均匀 ON-OFF 业务通过率

Table 4 Throughput under non-uniform ON-OFF traffic

算法	归一化负载		
	0.90	0.95	0.99
MF-MRSF	0.869 107 1	0.828 709 2	0.800 881 3
MUCB	0.992 866 1	0.959 425 3	0.930 353 8
MULHB	0.999 941 4	0.999 837 4	0.997 646 2
OQ	0.999 946 0	0.999 813 6	0.998 512 5



(a) 非均匀Bernoulli业务平均时延



(b) 非均匀ON-OFF业务平均时延

图3 非均匀 Bernoulli 和 ON-OFF 业务平均时延

Fig.3 Average delay under non-uniform Bernoulli and ON-OFF traffic

4 结论

1) 本文基于 CICQ 交换结构分析了单组播混合调度中 HoL 堵塞、节点缓存状态以及使交换机逼近 work-conserving 状态之间的关系。

2) 与已有研究中基于轮询机制或者基于最大权重匹配的调度算法不同,本文以缓解 HoL 堵塞问题为目标,尽量使交换机逼近 work-conserving 状态。

在上述分析的基础之上,本文提出了 DMQ 入队策略以及 MULHB 算法,通过仿真与现有性能较好的 MUCB 算法以及 MF-MRSF 算法进行对比,在均匀业务下,DMQ-MULHB 算法的性能更好。尤其在非均匀业务下,DMQ-MULHB 算法性能的表现有了显著的提升,并且接近 OQ 调度。

参考文献 (References)

[1] 熊庆旭. 输入排队结构交换机分组调度研究[J]. 通信学报,2005,26(6):118-129.
XIONG Q X. Research on packet scheduling in input-queued switches[J]. Journal on communications,2005,26(6):118-129 (in Chinese).

- [2] KAROL M, HLUCHYJ M, MORGAN S. Input versus output queueing on a space-division packet switch [J]. IEEE Transactions on communications, 1987, 35 (12): 1347-1356.
- [3] MARSAN M A, BIANCO A, GIACCONE P, et al. Optimal multicast scheduling in input-queued switches [C] // IEEE International Conference on Communications. Piscataway, NJ: IEEE Press, 2001: 2021-2027.
- [4] NABESHIMA M. Performance evaluation of a combined input-and-crosspoint-queued switch [J]. IEICE Transactions on Communications, 2000, 83 (3): 737-741.
- [5] WANG X, WANG Y, LI S, et al. Novel high performance scheduling algorithms for crosspoint buffered crossbar switches [J]. IEICE Transactions on Information and Systems, 2015, 98 (12): 2105-2115.
- [6] GAO Z, ZENG H, XIA Y, et al. SBF-GWF scheduling for combined input-crosspoint-queued (CICQ) switches [C] // International Conference on Computer Sciences and Convergence Information Technology. Piscataway, NJ: IEEE Press, 2012: 404-408.
- [7] 张元昊, 熊庆旭. CICQ 结构中逼近 work-conserving 的分组调度算法 [J]. 北京航空航天大学学报, 2016, 42 (11): 2481-2487.
ZHANG Y H, XIONG Q X. A New work-conserving-based packet scheduling algorithm for CICQ switches [J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42 (11): 2481-2487 (in Chinese).
- [8] MHAMDI L, HAMDI M. Scheduling multicast traffic in internally buffered crossbar switches [C] // IEEE International Conference on Communications. Piscataway, NJ: IEEE Press, 2004, 2: 1103-1107.
- [9] SUN S, HE S, ZHENG Y, et al. Multicast scheduling in buffered crossbar switches with multiple input queues [C] // 2005 Workshop on High Performance Switching and Routing. Piscataway, NJ: IEEE Press, 2005: 73-77.
- [10] MHAMDI L, VASSILIADIS S. Integrating uniand multicast scheduling in buffered crossbar switches [C] // 2006 Workshop on High Performance Switching and Routing. Piscataway, NJ: IEEE Press, 2006: 9076939.
- [11] YI P, LI H, XU J, et al. Scheduling multicast and unicast traffic in buffered crossbar switches [C] // 2006 IET International Conference on Wireless, Mobile and Multimedia Networks. London: IET, 2006: 1-4.
- [12] 梁佳诚, 熊庆旭, 闫付龙, 等. 基于 Work-Conserving 的 CICQ 结构中单组播分组调度算法 [J]. 北京航空航天大学学报, 2017, 42 (11): 144-150.
LIANG J C, XIONG Q X, YAN F L, et al. Packet scheduling algorithm for mixed unicast and multicast traffic in CICQ switches based on work-conserving [J]. Journal of Beijing University of Aeronautics and Astronautics, 2017, 42 (11): 144-150 (in Chinese).
- [13] MHAMDI L, VASSILIADIS S. Integrating uni-and multicast scheduling in buffered crossbar switches [C] // 2006 Workshop on High Performance Switching and Routing, HPSR 2006. Piscataway, NJ: IEEE Press, 2006: 99-104.
- [14] 高雅, 邱智亮, 张健. 一种支持小扇出多播业务均衡的信元排队策略 [J]. 北京邮电大学学报, 2014, 37 (5): 91-95.
GAO Y, QIU Z L, ZHANG J. A cell assignment algorithm for balancing multicast traffic with small fanout [J]. Journal of Beijing University of Posts and Telecommunications, 2014, 37 (5): 91-95 (in Chinese).
- [15] 陈庶樵, 扈红超, 郭云飞, 等. 一种支持单组播的 MCICQ 交换结构及其性能仿真 [J]. 系统仿真学报, 2009, 21 (13): 4003-4008.
CHEN S Q, HU H C, GUO Y F, et al. Uni-multicast support MCICQ switch and its performance simulation [J]. Journal of System Simulation, 2009, 21 (13): 4003-4008 (in Chinese).

作者简介:

袁龙 男, 硕士研究生。主要研究方向: 卫星网络、高性能交换技术。

熊庆旭 男, 博士, 教授, 博士生导师。主要研究方向: 通信网络、高性能交换技术、语义通信、无线传感网络。

萧翰 男, 博士研究生。主要研究方向: 卫星网络、高性能交换技术。

Packet scheduling algorithm for integrated unicast and multicast traffic based on reducing HoL blocking

YUAN Long, XIONG Qingxu^{*}, XIAO Han

(School of Electronic and Information Engineering, Beihang University, Beijing 100083, China)

Abstract: In view of the problem that few researches have been done on the mixed scheduling of unicast and multicast traffic for the combined input and crossbar queued (CICQ) architecture, and there is no research on the head of line (HoL) blocking problem, a new scheduling algorithm for mixed multicast and unicast, i. e. unicast with low HoL blocking (MULHB) algorithm, is proposed, which aims to reduce HoL blocking so that the switch can operate in work-conserving state as much as possible. In addition, to avoid the phenomenon of "starvation", the proposed algorithm considers the difference between unicast traffic and multicast traffic and uses weights to complete the arbitration between unicast and multicast. At the same time, this paper also proposes a dynamic multicast cell assignment algorithm named dynamic multicast queuing (DMQ), which allows the arrival multicast cell to select the appropriate queue without disorder. Simulation results show that the performance in terms of through rate and average packet delay obtained by DMQ-MULHB is much better than that of the existing popular algorithms under the different traffic patterns, and especially under the non-uniform traffic pattern, the performance is close to the output queuing (OQ) scheduling.

Keywords: packet switching; combined input and crossbar queued (CICQ); work-conserving; scheduling algorithm; multicast; HoL blocking

Received: 2018-05-23; **Accepted:** 2018-08-10; **Published online:** 2018-09-04 09:09

URL: kns.cnki.net/kcms/detail/11.2625.V.20180830.1513.003.html

Foundation item: National Natural Science Foundation of China (61271196)

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

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

DOI: 10.13700/j.bh.1001-5965.2018.0314

基于 AVSIMM 算法的高超声速再入滑翔目标跟踪



肖楚晗¹, 李炯^{2,*}, 雷虎民², 王华吉¹

(1. 空军工程大学 研究生学院, 西安 710051; 2. 空军工程大学 防空反导学院, 西安 710051)

摘 要: 针对跟踪高超声速目标的交互式多模型 (IMM) 算法中存在模型数量过多, 模型之间竞争导致滤波精度降低的问题, 在自适应网格交互式多模型 (AGIMM) 算法的基础上, 提出了一种自适应变结构交互式多模型 (AVSIMM) 算法跟踪高超声速再入滑翔目标。根据高超声速无动力再入滑翔目标当前机动状态的角速度参数, 在自适应调整当前时刻模型集中参数的同时, 针对 AGIMM 算法运动学模型的单一性, 设计了具有多种跟踪滤波运动学模型的 AVSIMM 算法, 通过模型集参数与算法结构的双重自适应调整实现了对目标高精度的跟踪。仿真结果表明, 与 AGIMM 算法相比, 所设计的 AVSIMM 算法不仅对结构和参数都具有更强的自适应性, 同时提高了高超目标的跟踪精度和跟踪效率。

关键词: 临近空间; 目标跟踪; 滤波; 自适应变结构交互式多模型 (AVSIMM); 自适应网格交互式多模型 (AGIMM) 算法

中图分类号: TJ765.1

文献标识码: A

文章编号: 1001-5965(2019)02-0413-09

近年来, 临近空间高超声速飞行器的快速发展对中国的空天防御安全构成了极大的威胁, 开展针对此类高速目标的拦截策略研究迫在眉睫^[1]。与传统的飞机类目标以及弹道类目标不同, 高超声速再入滑翔式飞行器 (Hypersonic Reentry-Glide Vehicle, HRGV) 具有高机动、高速、打击范围广等特点, 现阶段几乎无防御系统能够完成对其的拦截任务, 故研究 HRGV 对中国的国防事业有着重大意义。

HRGV 的机动范围在距地面 20 ~ 60 km、空气稀薄的临近空间内, 飞行速度极快, 与大气相互作用形成等离子鞘套并产生湍流尾迹等特殊的电磁/光学现象, 对拦截方的探测预警系统的要求极高^[1-2], 高超声速目标的跟踪难度大。但同时热流、动压、过载等硬性条件将高超目标的机动限制

在一定范围内^[3-5], 且高超声速滑翔式目标的机动能力有限。这些限制都为高超声速目标的跟踪提供了基础。HRGV 的典型代表是美国军队研制的高超声速技术飞行器 (Hypersonic Technology Vehicle, HTV-2) 和先进高超声速武器 (Advanced Hypersonic Weapon, AHW), 其中 AHW 因其外形结构更适合有效负载, 实用性更高。本文以 AHW 为研究对象展开讨论。

高超声速目标的运动模式多变, 运动模式与跟踪模型难以匹配, 且目标运动速度快、范围大, 因此对目标的跟踪精度难以满足拦截方的要求。为了提高跟踪精度, 国内外相关文献主要围绕跟踪方法与目标运动模型展开研究。

在目标运动模型方面, 文献 [6] 针对作周期性的滑跃机动的临近空间高超声速目标, 设计了

收稿日期: 2018-05-30; 录用日期: 2018-08-10; 网络出版时间: 2018-09-03 11:00

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20180830.1541.005.html

基金项目: 国家自然科学基金 (61573374, 61503408)

* 通信作者. E-mail: gracefull001@126.com

引用格式: 肖楚晗, 李炯, 雷虎民, 等. 基于 AVSIMM 算法的高超声速再入滑翔目标跟踪[J]. 北京航空航天大学学报, 2019, 45(2): 413-421. XIAO C H, LI J, LEI H M, et al. Hypersonic reentry gliding target tracking based on AVSIMM algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 413-421 (in Chinese).

一种运动学模型,该模型将加速度建模为正弦波自相关的零均值随机过程。文献[7]基于“当前”统计模型的思想,提出了一种可实时修正转弯角速度的交互式多模型(IMM)算法,用于高超声速目标跟踪。但以上研究采用了较为单一的模型进行目标跟踪,无法完全覆盖高超声速目标的机动模式,故多采用多模型对高超目标进行跟踪。文献[8-9]将白噪声模型(如常速(CV)模型、常加速度(CA)模型)与Morkov过程模型(如Singer模型)相结合,提出了基于CV-CA-Singer的IMM算法,分别对X-51飞行器的平飞加速巡航和俯冲攻击段以及滑跃式机动飞行的高超声速目标进行跟踪。文献[10]则在文献[6]的基础上针对滑跃式机动的高超声速目标,将正弦波模型与CV^[11]、CA^[12-13]模型组成IMM算法。

然而若采用固定结构的交互式多模型(FSIMM)算法,模型集中很多模型在特定的时间内与系统的有效模式差别较大,且这些“多余”模型的不必要的竞争降低了跟踪精度^[14],对跟踪模型数量和类型的综合选取限制了FSIMM算法的应用。为此,国外学者提出了基于变结构的多模型估计(VSMM)理论的跟踪方法,以突破FSIMM算法的局限性并提高了跟踪算法的精度。其中,文献[15]提出了3种模型集自适应方法:激活有向图(AD)方法,自适应网格(AG)方法,有向图切换(DS)方法。文献[16]基于DS方法利用加速度与角速度为模型的参数,同时依据模型的后验概率对子模型实时地进行调整,从而实现模型集自适应的变结构算法。但模型可跳转的模式是有限的,即有向图结构是固定,算法的自适应程度受限于有向图结构的复杂程度。与文献[16]相比,文献[17-19]均增强了算法的自适应性,并选取角速度为模型参数。其中,文献[17]提出了利用自适应网格交互式多模型(AGIMM)算法跟踪高机动模型。文献[18]则借鉴了文献[17]的思想,将AGIMM算法运用到对临近空间高超声速目标的跟踪,通过转弯角速度自适应,结合了多模型粒子滤波算法,调整模型集中3个模型的参数,从而完成滤波跟踪算法的自适应。文献[19]基于有向图的切换,根据卡尔曼滤波公式,提出了一种角速度滤波方法,并根据角速度的估计值修正有向图。但文献[17-19]的算法中均存在运动学滤波模型较为单一,无法完全覆盖所有机动模式的不足。AGIMM算法中每个模型的运动学跟踪模型始终采用的是匀速转弯(CT^[20])模型,然而当转弯角速率很小,接近于0时,可将目标运动近

似为直线运动,此时,若仍然只采用CT模型为跟踪系统的运动学模型进行滤波,则会对滤波精度造成影响。针对上述AGIMM算法运动学模型单一,无法适用于所有运动模式的问题,本文提出了一种自适应变结构交互多式模型(AVSIMM)算法跟踪高超声速目标,同时选择更优的跟踪运动模型并进行比较,通过对滤波算法与跟踪模型的双重自适应调整,最终改善了目标跟踪算法的跟踪性能。

1 高超声速目标运动学模型

高超声速飞行器在纵向平面内发生机动时,描述其运动的状态方程的形式为

$$\mathbf{X}(k+1) = \mathbf{F}(k)\mathbf{X}(k) + \mathbf{\Gamma}(k)\mathbf{w}(k) \quad (1)$$

当目标的状态变量为 $\mathbf{X} = [x, \dot{x}, y, \dot{y}]^T$ 时,状态量分别为二维平面内 x, y 方向的位置、速度; \mathbf{F} 为系统的状态转移矩阵,而 $\mathbf{\Gamma}$ 则为噪声矩阵;取 $\mathbf{w}_{CV} \sim N(0, \mathbf{Q}_{CV})$, $\mathbf{w}_{CT_2} \sim N(0, \mathbf{Q}_{CT_2})$ 表示过程演化噪声, \mathbf{Q} 为协方差噪声,下标CV和CT₂分别表示CV模型和二维CT模型。

此时当系统采用常速(CV)模型和匀速转弯(CT)模型进行目标跟踪时,假设系统的采样时间间隔为 T ,目标的转弯角速度为 ω ,则系统的离散化状态转移矩阵为

$$\mathbf{F}_{CV} = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

$$\mathbf{F}_{CT_2} = \begin{bmatrix} 1 & \frac{\sin(\omega T)}{\omega} & 0 & -\frac{(1 - \cos(\omega T))}{\omega} \\ 0 & \cos(\omega T) & 0 & -\sin(\omega T) \\ 0 & \frac{1 - \cos(\omega T)}{\omega} & 1 & \frac{\sin(\omega T)}{\omega} \\ 0 & \sin(\omega T) & 0 & \cos(\omega T) \end{bmatrix} \quad (3)$$

当目标的状态变量为 $\mathbf{X} = [x, \dot{x}, \ddot{x}, y, \dot{y}, \ddot{y}]^T$ 时,状态量分别为二维平面内 x, y 方向的位置、速度、加速度,取 $\mathbf{w}_{Singer} \sim N(0, \mathbf{Q}_{Singer})$, $\mathbf{w}_{CT_3} \sim N(0, \mathbf{Q}_{CT_3})$,下标Singer和CT₃分别表示Singer模型和三维CT模型。此时当系统采用CT模型和Singer^[21]模型时,假设系统的采样时间间隔为 T ,目标的转弯角速度为 ω ,此时系统的离散化状态转移矩阵为

$$F_{\text{Singer}} = \begin{bmatrix} 1 & T & \frac{\alpha T - 1 + e^{-\alpha T}}{\alpha T} & 0 & 0 & 0 \\ 0 & 1 & \frac{1 - e^{-\alpha T}}{\alpha} & 0 & 0 & 0 \\ 0 & 0 & e^{-\alpha T} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & T & \frac{\alpha T - 1 + e^{-\alpha T}}{\alpha T} \\ 0 & 0 & 0 & 0 & 1 & \frac{1 - e^{-\alpha T}}{\alpha} \\ 0 & 0 & 0 & 0 & 0 & e^{-\alpha T} \end{bmatrix} \quad (4)$$

$$F_{\text{CT}_3} = \begin{bmatrix} 1 & \frac{\sin(\omega T)}{\omega} & \frac{1 - \cos(\omega T)}{\omega^2} & 0 & 0 & 0 \\ 0 & \cos(\omega T) & \frac{\sin(\omega T)}{\omega} & 0 & 0 & 0 \\ 0 & -\omega \sin(\omega T) & \cos(\omega T) & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \frac{\sin(\omega T)}{\omega} & \frac{1 - \cos(\omega T)}{\omega^2} \\ 0 & 0 & 0 & 0 & \cos(\omega T) & \frac{\sin(\omega T)}{\omega} \\ 0 & 0 & 0 & 0 & -\omega \sin(\omega T) & \cos(\omega T) \end{bmatrix} \quad (5)$$

式中: α 为机动时间常数的倒数,而机动时间常数又与目标的机动时间持续的长短有关。

2 AVSIMM 算法

为解决 AGIMM 算法中运动学滤波模型单一的问题,考虑添加 CV、CA、Singer 等运动学模型,用于描述直线运动的模型作为系统的。但在直接增添运动学模型的同时,会增加模型集中的模型数量,进而加强模型之间的竞争关系,降低跟踪精度,故设计了一种 AVSIMM 算法:设置一个合适的门限值,当滤波更新得到的角速度值小于门限值时,自适应选择跟踪系统的运动学模型为 CV、CA 或 Singer 模型。至此,在模型集参数自适应的基础上,实现了算法结构的自适应。

初始时刻,取能够覆盖所有机动空间的粗网格模型集的 3 个模型参数为 $D_0 = \{\omega_L, \omega_C, \omega_R\} = \{\omega_{\min}, \omega_{\min} + \frac{\omega_{\max} - \omega_{\min}}{2}, \omega_{\max}\}$ 。假设 $k (k = 0, 1, \dots)$ 时刻,目标以角速度 ω (未知) 在 $[\omega_{\min}, \omega_{\max}]$ 范围内机动。模型集中状态向量的集合为

$X_k = \{x_k^L, x_k^C, x_k^R\}$, 误差协方差矩阵集合为 $P_k = \{P_k^L, P_k^C, P_k^R\}$, 模型转换概率为

$$P_{\text{LCR}} = \begin{bmatrix} P_{\text{LL}} & P_{\text{LC}} & P_{\text{LR}} \\ P_{\text{CL}} & P_{\text{CC}} & P_{\text{CR}} \\ P_{\text{RL}} & P_{\text{RC}} & P_{\text{RR}} \end{bmatrix} \quad (6)$$

在每个循环时间步长 ($k \rightarrow k+1$) 内,模型集

按照以下步骤进行自适应调整。

步骤 1 更新模型集网格中心的 ω_k^C 参数:

$$\omega_{k+1}^C = \mu_k^L \omega_k^L + \mu_k^C \omega_k^C + \mu_k^R \omega_k^R \quad (7)$$

式中: $\{\omega_k^L, \omega_k^C, \omega_k^R\}$ 为模型集, μ_k^L, μ_k^C 和 μ_k^R 为模型的交互后验概率。

步骤 2 根据更新的网格中心 ω_{k+1}^C , 调整模型间距。

1) 若 $\mu_k^C = \max\{\mu_k^L, \mu_k^C, \mu_k^R\}$, 即模型集无跳跃

$$\omega_{k+1}^L = \begin{cases} \omega_{k+1}^C - \frac{\lambda_k^L}{2} & \mu_k^L < t_1 \\ \omega_{k+1}^C - \lambda_k^L & \text{其他} \end{cases} \quad (8)$$

$$\omega_{k+1}^R = \begin{cases} \omega_{k+1}^C + \frac{\lambda_k^R}{2} & \mu_k^R < t_1 \\ \omega_{k+1}^C + \lambda_k^R & \text{其他} \end{cases} \quad (9)$$

式中: $\lambda_k^L = \max\{\omega_k^C - \omega_k^L, \delta_\omega\}$, $\lambda_k^R = \max\{\omega_k^R - \omega_k^C, \delta_\omega\}$, δ_ω 为模型最小间距; t_1 为检测阈值。

2) 若 $\mu_k^L = \max\{\mu_k^L, \mu_k^C, \mu_k^R\}$, 即模型集向左跳跃

$$\omega_{k+1}^R = \omega_{k+1}^C + \lambda_k^R \quad (10)$$

$$\omega_{k+1}^L = \begin{cases} \omega_{k+1}^C - 2\lambda_k^L & \mu_k^L > t_2 \\ \omega_{k+1}^C - \lambda_k^L & \text{其他} \end{cases} \quad (11)$$

式中: t_2 为检测阈值。

3) 若 $\mu_k^R = \max\{\mu_k^L, \mu_k^C, \mu_k^R\}$, 即模型集向右跳跃

$$\omega_{k+1}^L = \omega_{k+1}^C - \lambda_k^L \quad (12)$$

$$\omega_{k+1}^R = \begin{cases} \omega_{k+1}^C + 2\lambda_k^R & \mu_k^R > t_2 \\ \omega_{k+1}^C + \lambda_k^R & \text{其他} \end{cases} \quad (13)$$

至此,获得更新的模型集 $D_{k+1} = \{\omega_{k+1}^L, \omega_{k+1}^C, \omega_{k+1}^R\}$ 。

步骤3 k 时刻误差协方差矩阵与状态估计值融合。

状态与误差协方差矩阵采用的是交互式多模型算法,具体过程如下:

1) 获得 k 时刻的混合概率 $\mu_{k|k}^{ij}$

$$\mu_{k|k}^{ij} = \frac{1}{c_j} p_{ij} \mu_k^i \quad i, j = 1, 2, \dots, r \quad (14)$$

式中: $\bar{c}_j = \sum_{i=1}^r p_{ij} \mu_k^i$ 为归一化常数; p_{ij} 为模型 i 转移到模型 j 的状态转移概率。此处因模型集中有3个模型,故取 $r=3$,下文同理。

2) k 时刻每个模型的状态估计和误差协方差矩阵初始化

$$\hat{X}_{k|k}^{0j} = \sum_{i=1}^r \hat{X}_k^i \mu_{k|k}^{ij} \quad j = 1, 2, \dots, r \quad (15)$$

$$\hat{P}_{k|k}^{0j} = \sum_{i=1}^r \left[P_{k|k}^i + (\hat{X}_k^i - \hat{X}_{k|k}^{0j}) (\hat{X}_k^i - \hat{X}_{k|k}^{0j})^T \right] \mu_{k|k}^{ij} \quad j = 1, 2, \dots, r \quad (16)$$

步骤4 更新状态转移矩阵 F_{k+1} 。

根据更新的 $D_{k+1} = \{\omega_{k+1}^L, \omega_{k+1}^C, \omega_{k+1}^R\}$,将转弯角速度与门限值作比较,再选择合适的状态转移矩阵。取门限 $G = 0.002 \text{ rad/s}$ 。

1) 若 $\omega_{k+1}^L < G$,则更新状态转移矩阵 F_{k+1} 为可描述直线运动的运动学模型,如 CV、CA、Singer 模型等,否则取状态转移矩阵 F_{k+1} 为 CT 模型。

2) 若 $\omega_{k+1}^C < G$,则更新 F_{k+1} 为可描述直线运动的运动学模型,否则取状态转移矩阵 F_{k+1} 为 CT 模型。

3) 若 $\omega_{k+1}^R < G$,则更新 F_{k+1} 为可描述直线运动的运动学模型,否则取状态转移矩阵 F_{k+1} 为 CT 模型。

步骤5 扩展卡尔曼滤波。

分别对模型集中的每个模型进行扩展卡尔曼滤波(EKF),得到每个模型的状态估计值与误差协方差矩阵估计值,即 $\{\hat{X}_{k+1|k}^L, \hat{X}_{k+1|k}^C, \hat{X}_{k+1|k}^R\}$ 和 $\{P_{k+1}^L, P_{k+1}^C, P_{k+1}^R\}$ 。

系统的状态与误差协方差估计为

$$\hat{X}_{k+1|k}^i = F_{k+1}^i \hat{X}_k^i \quad (17)$$

$$P_{k+1|k}^i = F_{k+1}^i P_k^i (F_{k+1}^i)^T + Q_k \quad (18)$$

系统的量测预测为

$$\hat{Z}_{k+1|k}^i = h(\hat{X}_{k+1|k}^i) \quad (19)$$

量测矩阵的雅可比矩阵为

$$H_{k+1} = \left. \frac{\partial h(x, v)}{\partial x} \right|_{x=\hat{x}_{k+1|k}, v_{k+1}=0} \quad (20)$$

新息协方差矩阵为

$$S_{k+1}^i = H_{k+1} P_{k+1|k}^i H_{k+1}^T + R_{k+1} \quad (21)$$

量测一步预测为

$$\tilde{z}_{k+1}^i = Z_{k+1}^i - \hat{Z}_{k+1|k}^i \quad (22)$$

状态与误差协方差估计为

$$\hat{X}_{k+1|k+1}^i = \hat{X}_{k+1|k}^i + P_{k+1|k}^i H_{k+1}^T (S_{k+1}^i)^{-1} \tilde{z}_{k+1}^i \quad (23)$$

$$P_{k+1}^i = P_{k+1|k}^i (I - H_{k+1}^T (S_{k+1}^i)^{-1} H_{k+1} P_{k+1|k}^i) \quad (24)$$

式中: $i=1, 2, \dots, r$,代表模型集中的第 i 个模型。

步骤6 $k+1$ 时刻误差协方差矩阵与状态估计值融合。

1) 模型似然函数更新:

$$L_{k+1}^i = |2\pi S_{k+1}^i|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\tilde{z}_{k+1}^i)^T (S_{k+1}^i)^{-1} \tilde{z}_{k+1}^i\right] \quad i = 1, 2, \dots, r \quad (25)$$

2) 模型概率更新:

$$\mu_{k+1}^j = \frac{1}{\sum_{j=1}^r L_{k+1}^j \bar{c}_j} L_{k+1}^j \bar{c}_j \quad j = 1, 2, \dots, r \quad (26)$$

3) 误差协方差矩阵与状态估计值融合:

$$\hat{X}_{k+1|k+1} = \hat{X}_{k+1|k+1}^L \mu_{k+1}^L + \hat{X}_{k+1|k+1}^C \mu_{k+1}^C + \hat{X}_{k+1|k+1}^R \mu_{k+1}^R \quad (27)$$

$$P_{k+1} = P_{k+1}^L \mu_{k+1}^L + P_{k+1}^C \mu_{k+1}^C + P_{k+1}^R \mu_{k+1}^R \quad (28)$$

AVSIMM 算法采用了直线运动模型与协调转弯模型自适应,与 AGIMM 算法相比,转弯协调模型仍采用 CT 模型,而在描述直线运动时,可选择 CV、CA、Singer 模型。

Singer 模型是一个 Markov 过程模型,通过选择 Singer 模型中的参数 $0 < \alpha T < \infty$,能够使 Singer 模型介于 CV 模型和 CA 模型之间,较之二者, Singer 模型具有更宽的覆盖面^[14]。同时,在关于目标机动的在线信息不足或不准确的情况下,相比较 CV、CA 模型, Singer 模型是最理想的模型。

下文将把基于 CT、Singer 模型的 AVSIMM I 算法与基于 CT、CV 模型的 AVSIMM II 算法以及文献[18]中只基于 CT 模型的 AGIMM 算法的仿真结果进行比较。同时,考虑到临近空间高超声速目标与普通高机动目标存在一定的差异性,而

本文设计的方法的理论基础不是针对的临近空间目标的,故为了体现设计方法的性能,还将与文献[10]中针对临近空间高超声速目标提出的 IMM 算法作比较。

3 算法对比与仿真

本文以跟踪临近空间高超声速再入滑翔弹 AHW 为研究背景。因现阶段无 AHW 的真实实验数据,故利用通用飞行器(Common Aerial Vehicle, CAV)模型^[22]作为参考,结合基本的动力学方程,利用 GPOPS^[23]软件近似模拟为 AHW 的机动轨迹。考虑到 AHW 主要是在纵向平面内大范围的跳跃机动,本文针对该机动方式研究平面内的目标跟踪问题^[16,24-25]。同时考虑到假设在纵向二维平面内对高超声速飞行器进行跟踪,设目标在(1500,60) km 处开始机动,仿真得到目标的真实轨迹如图1所示。

在二维平面内,设机动目标与基站的相对位置关系如图1所示。图中 d 为目标与基站之间的相对距离, φ 为目标俯仰角,规定逆时针为正。故可得高超声速模型的量测方程为

$$\begin{cases} d = \sqrt{x^2 + y^2} \\ \varphi = \arctan\left(\frac{y}{x}\right) \end{cases} \quad (29)$$

离散化得

$$z_k = \mathbf{h}(x_{k+1}) + \mathbf{V}_{k+1} \quad (30)$$

式中:

$$\mathbf{h}(x_{k+1}) = \begin{bmatrix} \sqrt{x_{k+1}^2 + y_{k+1}^2} \\ \arctan\left(\frac{y_{k+1}}{x_{k+1}}\right) \end{bmatrix} \quad (31)$$

$$\mathbf{V}_{k+1} = \begin{bmatrix} R_d(k+1) & 0 \\ 0 & R_\varphi(k+1) \end{bmatrix} \quad (32)$$

目标在(1500,60) km 处开始机动,其真实轨迹如图2所示。

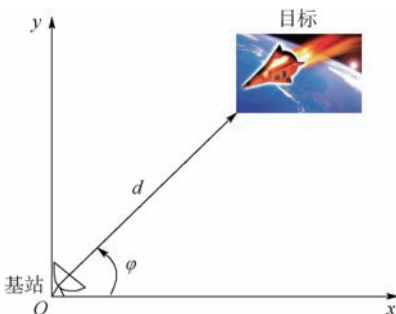


图1 AHW 目标与基站相对位置示意图

Fig.1 Schematic diagram of relative position of

AHW target to base station

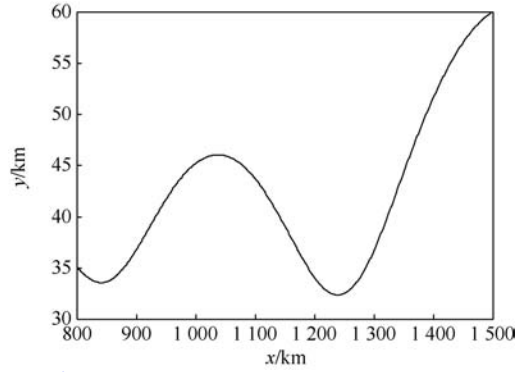


图2 目标运动真实轨迹

Fig.2 True trajectory of target moving

设定观测基站坐标为(0,0),跟踪时间为277 s,采样周期为 $T=0.5$ s,量测距离标准差为50 m,俯仰角的标准差为 0.05° ,即量测协方差矩阵为 $\text{diag}(0.05 \text{ km}, 0.05\pi/180 \text{ rad})$ 。同时,考虑到 EKF 算法的局限性,所有的位置、速度、加速度单位均为 km、km/s、km/s²。

AVSIMM I 算法中,目标的状态变量为 $\mathbf{X} = [x, \dot{x}, \ddot{x}, y, \dot{y}, \ddot{y}]^T$, $\mathbf{w}_{\text{CT}} \sim N(0, \mathbf{Q}_{\text{CT}})$, $\mathbf{w}_{\text{Singer}} \sim N(0, \mathbf{Q}_{\text{Singer}})$ 。模型的初始状态估计值与状态更新值均为 $[1500 + 0.1, -3 + 0.01, 2 \times 10^{-4} + 10^{-3}, 60 + 0.1, -0.1 + 0.01, -9 \times 10^{-3} + 10^{-3}]^T$,初始模型更新协方差矩阵为 $\text{diag}(0.1, 10^{-2}, 10^{-3}, 0.1, 10^{-2}, 10^{-3})$ 。

AVSIMM II 算法中,目标的状态变量为 $\mathbf{X} = [x, \dot{x}, y, \dot{y}]^T$, $\mathbf{w}_{\text{CT}} \sim N(0, \mathbf{Q}_{\text{CT}})$, $\mathbf{w}_{\text{CV}} \sim N(0, \mathbf{Q}_{\text{CV}})$ 。模型的初始状态估计值与状态更新值均为 $[1500 + 0.1, -3 + 0.01, 60 + 0.1, -0.1 + 0.01]^T$,初始模型更新协方差矩阵为 $\text{diag}(0.1, 10^{-2}, 0.1, 10^{-2})$ 。

2种 AVSIMM 算法对模型集中3个模型的初始化均为 $D_0 = \{\omega_0^L, \omega_0^C, \omega_0^R\} = \{\omega_{\min}, \omega_{\min} + \frac{\omega_{\max} - \omega_{\min}}{2}, \omega_{\max}\} = \{-0.3, -0.05, 0.2\}$ 。根据式(2)~式(7),设置变量初始误差协方差矩阵为 $\text{diag}(0.1, 10^{-2}, 0.1, 10^{-2})$ 。最小模型间距为 $\delta_\omega = \frac{0.1\pi}{180} \text{ rad/s}$,检测门限 $t_1 = 0.2 \text{ s}, t_2 = 0.9 \text{ s}$ 。

参照文献[18],AGIMM 算法中,目标的状态量为 $\mathbf{X} = [x, \dot{x}, y, \dot{y}]^T$, $\mathbf{w}_{\text{CT}} \sim N(0, \mathbf{Q}_{\text{CT}})$ 。状态变量初值和状态更新均为 $[1500 + 0.1, -3 + 0.01, 60 + 0.1, -0.1 + 0.01]^T$ 。

参照文献[10],设计的 CV-CA-Singer 模型的 IMM 算法中,目标的状态量为 $\mathbf{X} = [x, \dot{x}, x, x, y,$

$\dot{y}, \ddot{y}, \ddot{y} \dots]$, $w_{\text{sin}} \sim N(0, Q_{\text{sin}})$, $w_{\text{cv}} \sim N(0, Q_{\text{cv}})$, $w_{\text{ca}} \sim N(0, Q_{\text{ca}})$, 其中 Q_{sin} 表达式如文献[10]所述, 取目标的机动频率为 $1/200$ Hz, 设目标的加速度方差为 0.09^2 。状态变量初值为 $[1500 + 0.1, -3 \times 10^{-4} + 10^{-5}, 3 \times 10^{-6} + 10^{-6}, 60 + 0.1, -0.1 + 0.01, -10^{-2} + 10^{-3}, 1.6 \times 10^{-5} + 10^{-5}]$ 。

4种算法中, 每个模型集包含3个模型, 设定初始时刻3个模型的初始概率为 $[0.01, 0.98, 0.01]^T$ 。初始模型转移概率矩阵为

$$P_{\text{LCR}} = \begin{bmatrix} 0.85 & 0.1 & 0.5 \\ 0.1 & 0.8 & 0.1 \\ 0.05 & 0.1 & 0.85 \end{bmatrix} \quad (33)$$

跟踪目标的位置和速度的均方根误差(RMSE)的计算公式为

$$P_{\text{RMSE}} = \left\{ \frac{1}{M} \sum_{i=1}^M [(x_k - \hat{x}_i)^2 + (y_k - \hat{y}_{k,i})^2] \right\}^{\frac{1}{2}} \quad (34)$$

$$V_{\text{RMSE}} = \left\{ \frac{1}{M} \sum_{i=1}^M [(V_{x_k} - V_{\hat{x}_{k,i}})^2 + (V_{y_k} - V_{\hat{y}_{k,i}})^2] \right\}^{\frac{1}{2}} \quad (35)$$

3.1 跟踪轨迹对比

根据以上设定的仿真条件, 考虑到观测矩阵的非线性性, 利用EKF算法得到基于AVSIMM I算法、AVSIMM II算法、AGIMM算法、IMM算法的目标跟踪轨迹。进行100次蒙特卡罗仿真, 结合跟踪目标的真实运动轨迹, 得到对比如图3所示。

由图3易知, 4种算法都能够实现对临近空间高超声速目标的有效跟踪。为了完成对4种算法的进一步比较, 下面将对4种算法的速度误差、位置误差、速度均方根误差以及位置均方根误差的比较进行具体讨论与分析。

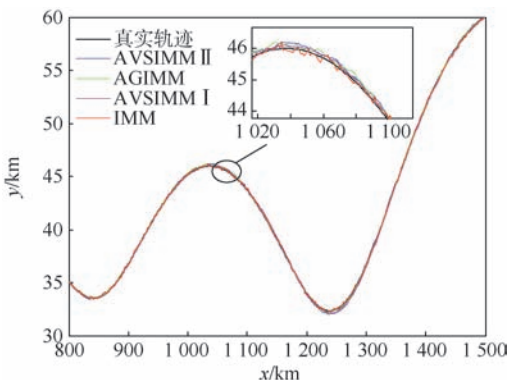


图3 目标真实轨迹与4种算法估计轨迹

Fig.3 True trajectory of target and estimation trajectories in four algorithms

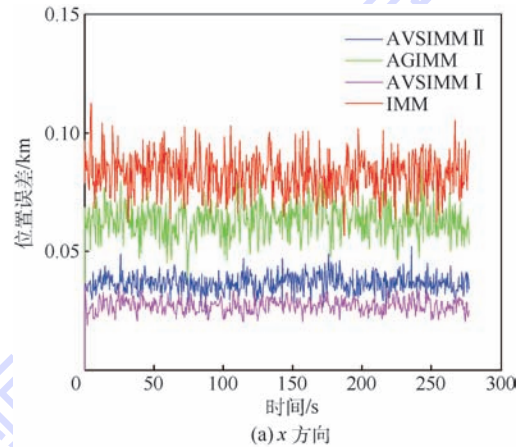
3.2 位置与速度的估计误差结果对比

基于AVSIMM I算法、AVSIMM II算法、AGIMM算法以及IMM算法对目标位置的估计误差如图4所示, 速度误差估计如图5所示。

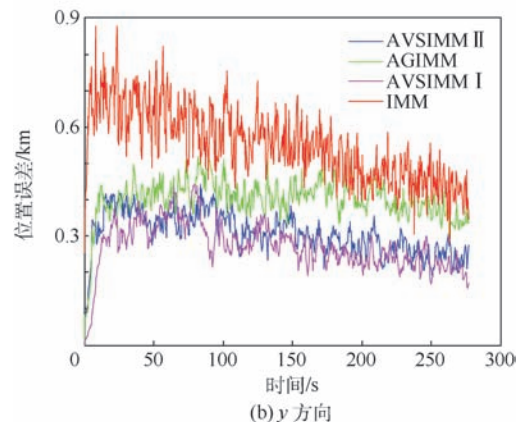
结合表1和图4可知, AVSIMM算法中x方向与y方向的位置误差要小于AGIMM算法与IMM算法中x方向与y方向的位置误差。由表1与图5可知, AVSIMM算法x方向与y方向的速度误差要小于AGIMM算法的x方向与y方向的速度误差, 而AGIMM算法在x方向上的速度误差大于IMM算法, 而其在y方向上的速度误差则小于IMM算法。而结合表1与图4、图5比较AVSIMM I算法与AVSIMM II算法的位置与速度误差可知, 前者的速度与位置误差均小于后者。综上, AVSIMM I算法的速度与位置误差最小, 跟踪性能最优。

根据式(34), 可以得到基于AVSIMM I、AVSIMM II、AGIMM算法以及IMM算法的目标位置跟踪的均方根误差(RMSE), 如图6所示。

根据式(35), 可以得到基于AVSIMM I、AVSIMM II、AGIMM算法、IMM算法的目标速度



(a) x方向



(b) y方向

图4 目标跟踪位置误差

Fig.4 Position error of target tracking

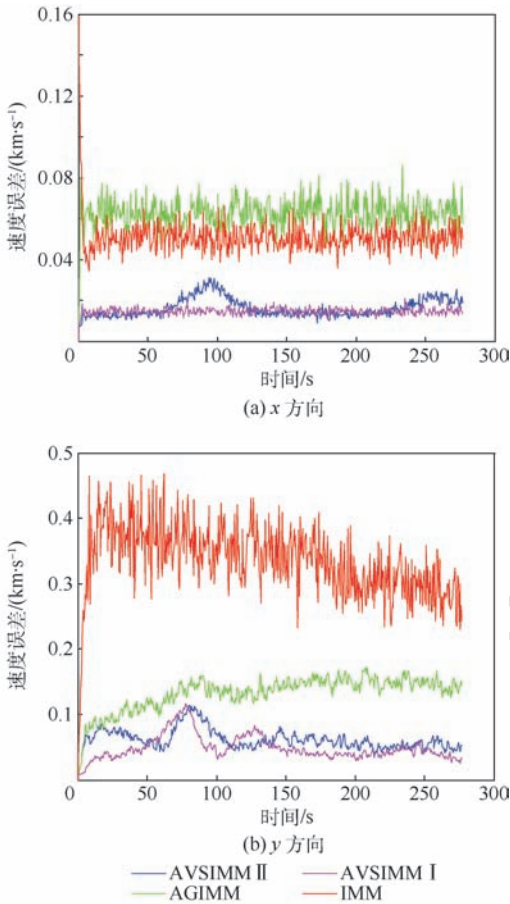


图 5 目标速度误差

Fig. 5 Velocity error of target tracking

表 1 位置、速度估计误差绝对值之和

Table 1 Sum of absolute value of position and velocity estimation error

算法	x 方向位置	y 方向位置	x 方向速度	y 方向速度
	误差/ km	误差/ km	误差/($\text{km} \cdot \text{s}^{-1}$)	误差/($\text{km} \cdot \text{s}^{-1}$)
AVSIMM I	14.97	150.44	8.11	27.02
AVSIMM II	20.27	169.79	9.02	34.19
AGIMM	35.05	222.14	34.59	73.63
IMM	45.83	306.13	28.30	185.63

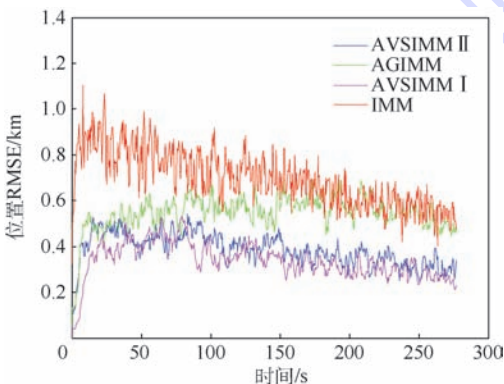


图 6 位置均方根误差

Fig. 6 Root mean square error of position

跟踪的 RMSE,如图 7 所示。

结合表 2 与图 6 可知,AVSIMM 算法的位置均方根误差要小于 AGIMM 算法,而 AGIMM 算法的位置均方根误差则小于 IMM 算法,结合表 2 与图 7 可知,AVSIMM 算法与 AGIMM 算法、IMM 算法相比,AVSIMM 算法的速度均方根误差更小。同时由表 2 与图 6、图 7 可知,AVSIMM I 算法的位置均方根误差与速度均方根误差均小于与 AVSIMM II 算法。同时由表 2 易知,AVSIMM 算法仿真的时长虽略长于 AGIMM 算法,但 AVSIMM 算法的速度与位置均方根误差更小,且 IMM 算法的仿真时间最长,故 AVSIMM 算法在单位时间内的跟踪效率更高,实用性更强。同样的,AVSIMM I 算法的跟踪效率高于 AVSIMM II 算法。

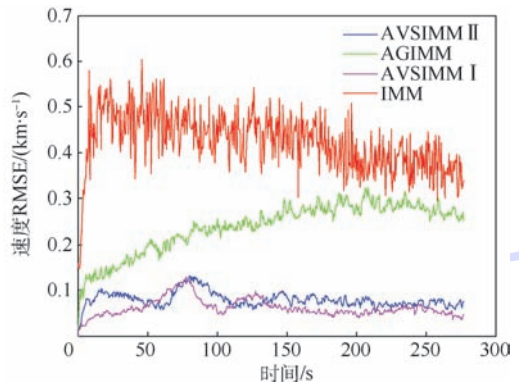


图 7 速度均方根误差

Fig. 7 Root mean square error of velocity

表 2 位置、速度均方根误差之和及仿真时间

Table 2 Sum of root mean square error of position and velocity and simulation time

算法	位置	速度 RMSE/	单次仿真 时间/s
	RMSE/km	($\text{km} \cdot \text{s}^{-1}$)	
AVSIMM I	188.0	34.50	1.47
AVSIMM II	213.70	44.15	1.45
AGIMM	302.79	132.30	1.44
IMM	3 896.73	234.55	4.32

4 结 论

本文设计了一种同时具有结构与参数自适应的交互式多模型的高精度滤波算法,研究了对临近空间高超声速再入滑翔目标的跟踪问题。

通过仿真结果可以得到如下结论:

1) AVSIMM 算法在算法结构上明显优于 AGIMM 算法与文献[10]中设计的针对临近空间高超声速目标的 IMM 算法。

2) AVSIMM I 算法在运动学模型上明显优于 AVSIMM II 算法。故设计的 2 种 AVSIMM 算法

中, AVSIMM I 算法的跟踪效果最佳。

本文算法重点讨论了对高超声速目标的跟踪情况,但考虑到目标的加速度气动参数信息也与目标机动情况紧密相关,故下一步工作考虑将加速度与气动参数扩维到状态向量中进行滤波算法设计。

参考文献 (References)

- [1] 李海宁,雷虎民,翟岱亮,等. 面向跟踪的吸气式高超声速飞行器动力学建模[J]. 航空学报,2014,35(6):1651-1664.
LI H N, LEI H M, ZHAI D L, et al. Tracking oriented dynamics modeling of air breath hypersonic vehicles[J]. Acta Aeronautica et Astronautica Sinica, 2014, 35(6): 1651-1664 (in Chinese).
- [2] 雍恩米,钱炜棋,何开锋. 基于雷达跟踪仿真的滑翔式再入弹道突防性能分析[J]. 宇航学报,2012,33(10):1370-1376.
YONG E M, QIAN W Q, HE K F. Penetration ability analysis for glide reentry trajectory based on radar tracking[J]. Journal of Astronautics, 2012, 33(10): 1370-1376 (in Chinese).
- [3] TAUQEER R S, HE L S, XU D J. Optimal trajectory analysis of hypersonic boost-glide waverider with heat load constraint[J]. Aircraft Engineering and Aerospace Technology, 2015, 87(1): 67-79.
- [4] ZHAO J, ZHOU R. Reentry trajectory optimization for hypersonic vehicle satisfying complex constraints[J]. Chinese Journal of Aeronautics, 2013, 26(6): 1544-1553.
- [5] 黄长强,国海峰,丁达理. 高超声速滑翔飞行器轨迹优化与制导综述[J]. 宇航学报,2014,35(4):369-379.
HUANG C Q, GUO H F, DING D L. A survey of trajectory optimization and guidance for hypersonic gliding vehicle[J]. Journal of Astronautics, 2014, 35(4): 369-379 (in Chinese).
- [6] 王国宏,李俊杰,李志准,等. 基于 SCT-IMM 的临近空间高超声速机动目标跟踪模型[J]. 现代雷达,2013,35(4):15-19.
WANG G H, LI J J, LI Z H, et al. Tracking model based on the SCT-IMM near space hypersonic target[J]. Modern Radar, 2013, 35(4): 15-19 (in Chinese).
- [7] 苗伟,李昌玺,吴聪. 基于修正转弯模型的交互多模型跟踪算法[J]. 现代防御技术,2015,43(3):113-118.
MIAO W, LI C X, WU C. Interactive multiple model tracking algorithm based on the modified model of turning[J]. Modern Defence Technology, 2015, 43(3): 113-118 (in Chinese).
- [8] 关欣,赵静,张政超,等. 一种可行的高超声速飞行器跟踪算法[J]. 电讯技术,2011,51(8):80-84.
GUAN X, ZHAO J, ZHANG Z C, et al. A feasible tracking algorithm for hypersonic aircrafts[J]. Telecommunication Engineering, 2011, 51(8): 80-84 (in Chinese).
- [9] 张远,吴昊. 临近空间目标运动建模与跟踪方法研究[J]. 火控雷达技术,2013,42(4):22-27.
ZHANG Y, WU H. Study on modeling of near-space moving targets and its tracking algorithm[J]. Fire Control Radar Technology, 2013, 42(4): 22-27 (in Chinese).
- [10] 李俊杰,王国宏,张翔宇,等. 临近空间高超声速滑跃式机动目标跟踪的 IMM 算法[J]. 电光与控制,2015,22(9):15-19.
LI J J, WANG G H, ZHANG X Y, et al. An IMM algorithm for tracking near-space hypersonic target with slippage leap maneuvering[J]. Electronics Optics & Control, 2015, 22(9): 15-19 (in Chinese).
- [11] ZHU W L, XU Z P, LI B, et al. Research on the observability of bearing-only target tracking based on multiple sonar sensors [C]//Proceedings of 2012 2nd International Conference on Instrumentation, Measurement, Computer, Communication and Control (IMCCC). Piscataway, NJ: IEEE Press, 2012: 631-634.
- [12] ZHU W L, XU Z P, LI B, et al. Research on the observability of bearing-only tracking for moving target in constant acceleration based on multiple sonar sensors [C]//Proceedings of 2012 International Symposium on Information Science and Engineering. Piscataway, NJ: IEEE Press, 2013: 3-6.
- [13] 何友,修建娟,关欣,等. 雷达数据处理及应用[M]. 北京:电子工业出版社,2013:36-40.
HE Y, XIU J J, GUAN X, et al. Radar data processing and application[M]. Beijing: Electronics Industry Press, 2013: 36-40 (in Chinese).
- [14] 韩崇昭,朱洪艳,段战胜,等. 多源信息融合[M]. 2版. 北京:清华大学出版社,2010:1-14.
HAN C Z, ZHU H Y, DUAN Z S, et al. Multi-source information fusion[M]. 2nd ed. Beijing: Tsinghua University Press, 2010: 1-14 (in Chinese).
- [15] LI X R, BAR-SHALON Y. Multiple-model estimation with variable structure [J]. IEEE Transactions on Automatic Control, 1996, 41(4): 478-493.
- [16] 肖松,李智淮,谭贤四,等. 临近空间高超声速飞行器 DG-VSMM 跟踪算法[J]. 弹道学报,2013,25(2):22-27.
XIAO S, LI Z H, TAN X S, et al. DG-VSMM tracking algorithm for near-space hypersonic vehicle [J]. Journal of Ballistics, 2013, 25(2): 22-27 (in Chinese).
- [17] JILKOV V P, ANGELOVA D S, SEMERDJIEV T Z A. Design and comparison of mode-set adaptive IMM algorithms for maneuvering target tracking [J]. IEEE Transactions on Aerospace and Electronic Systems, 1999, 35(1): 343-350.
- [18] 秦雷,李君龙,周荻. 基于 AGIMM 的临近空间机动目标跟踪滤波算法[J]. 系统工程与电子技术,2015,37(5):1009-1014.
QIN L, LI J L, ZHOU D. Tracking filter algorithm for near space target based on AGIMM [J]. Systems Engineering and Electronics, 2015, 37(5): 1009-1014 (in Chinese).
- [19] 田野,蒋宏,梁国威,等. 基于角速度修正的变结构多模型目标跟踪算法[J]. 电光与控制,2012,19(1):46-53.
TIAN Y, JIANG H, LIANG G W, et al. Turn rate amended VSMM algorithm for target tracking [J]. Electronic Optics & Control, 2012, 19(1): 46-53 (in Chinese).
- [20] LI X R, JILKOV V P. Survey of maneuvering target tracking. Part I: Dynamic models [J]. IEEE Transactions on Aerospace and Electronic Systems, 2003, 39(4): 1345-1353.
- [21] SINGER R A. Estimating optimal tracking filter performance for manned maneuvering targets [J]. IEEE Transactions on Aero-

- space and Electronic Systems,1970,6(4):473-483.
- [22] WANG Y L,TANG W,ZHANG Y, et al. Geometry parameterization and aerodynamic analysis of a common aero vehicle[J]. Acta Aerodynamica Sinica,2007,25(30):368-371.
- [23] PATTERSON M A,RAO A V. GPOPS-II: A MATLAB software for solving multiple-phase optimal control problems using HP-adaptive Gaussian quadrature collocation methods and sparse nonlinear programming[J]. ACM Transactions on Mathematical Software,2014,41(1):1-37.
- [24] 李君龙,周荻. 基于交互式多模型算法跟踪临近空间目标[J]. 系统工程与电子技术,2014,36(7):1243-1249.
- LI J L,ZHOU D. Tracking for near space target based on IMM algorithm[J]. Systems Engineering and Electronics,2014,36(7):1234-1249(in Chinese).
- [25] 聂晓华,张夫鸣,徐一鸣. NSHV 机动目标跟踪的自适应模型算法[J]. 系统工程与电子技术,2016,38(3):506-511.
- NIE X H,ZHANG F M,XU Y M. Adaptive model algorithm for maneuvering target tracking of NSHV[J]. Systems Engineering and Electronics,2016,38(3):506-511(in Chinese).

作者简介:

肖楚晗 女,硕士研究生。主要研究方向:临近空间高超声速目标跟踪。

李炯 男,博士,副教授。主要研究方向:导航制导与控制。

Hypersonic reentry gliding target tracking based on AVSIMM algorithm

XIAO Chuhan¹, LI Jiong^{2,*}, LEI Humin², WANG Huaji¹

(1. Master Student College, Air Force Engineering University, Xi'an 710051, China;

2. Air and Missile Defense College, Air Force Engineering University, Xi'an 710051, China)

Abstract: Aimed at the decrease of filtering accuracy caused by too many models and competitions between them in interactive multi-model (IMM) algorithm for tracking hypersonic targets, an adaptive variable structure interactive multi-model (AVSIMM) algorithm is proposed based on adaptive grid IMM (AGIMM) algorithm for tracking hypersonic non-powered reentry gliding targets. Based on the angular velocity parameters of the real-time maneuvering state of gliding hypersonic target without power load, AVSIMM algorithm with various dynamic models is designed to focus on the unicity of models in AGIMM algorithm. AVSIMM algorithm realizes high-precision tracking through adaptive adjustments in both model set parameters and algorithm structure. Simulation results show that, compared with AGIMM algorithm, the designed AVSIMM algorithm not only is more adaptive in structure and parameters but also improves accuracy and efficiency in tracking hypersonic targets.

Keywords: near space; target tracking; filter; adaptive variable structure interactive multi-model (AVSIMM); adaptive grid interactive multi-model (AGIMM)

Received: 2018-05-30; Accepted: 2018-08-10; Published online: 2018-09-03 11:00

URL: kns.cnki.net/kcms/detail/11.2625.V.20180830.1541.005.html

Foundation items: National Natural Science Foundation of China (61573374,61503408)

* Corresponding author. E-mail: graceful1001@126.com

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

DOI: 10.13700/j.bh.1001-5965.2018.0662

考虑临近车道行人对交通流影响的改进跟驰模型

李宏刚¹, 高哈尔·达吾力¹, 王帅^{2,*}, 余贵珍¹, 王朋成¹

1. 北京航空航天大学 交通科学与工程学院, 北京 100083; 2. 北京航空航天大学 计算机学院, 北京 100083)

摘 要: 基于实际人车混合行驶的情形, 提出了一个改进的车辆动力学模型。该模型考虑了邻车道或横向的行人、自行车等对主干道车辆的驾驶行为影响。基于经典的最优速度模型, 将主干道车辆与行人之间的横向距离和纵向距离作为参数引入最优速度模型中。为了验证所提模型的稳定性和有效性, 利用线性稳定性理论, 推导出模型的稳定性条件和非稳定性条件, 绘制中性稳定性曲线, 直观描述了交通流稳定区域大小。理论结果表明: 考虑横向行人干扰因素的改进跟驰模型比传统的只考虑单一车道车辆因素的跟驰模型更加的稳定, 并且不同参数的变化, 所引起的稳态区域也会发生变化。采用了更加真实的优化速度方程, 通过仿真实验来描述车辆的驾驶行为; 仿真实验列举了2种实际场景: 行人稀少和行人较多时。分别绘制了车辆的速度-时间变化曲线以及车辆的时空图, 实验结果表明: 横向行人的确会干扰车辆的正常驾驶; 不同场景下, 行人数量的多寡也会对车辆的驾驶行为造成不同程度的影响。

关键词: 跟驰模型; 最优速度模型; 线性稳定性理论; 人车混合交通; 横向行人干扰因素

中图分类号: U46

文献标识码: A

文章编号: 1001-5965(2019)02-0422-07

几十年来, 交通流得到了广泛的研究, 国内外学者提出了很多模型, 如流体模型、跟驰模型和细胞自动化模型等^[1-3]。由于具有直观的物理意义, 车辆跟驰模型引起了很多研究者的关注。大多数的跟驰模型都假定车辆只受到前车驾驶行为的影响, 但在实际复杂的交通环境中, 不仅有车辆还包括行人和自行车的混合交通流。截止目前, 人车混合流交通行为的研究相对较少, 本文旨在补充这一领域的内容。

自 Reuschel^[4] 和 Pipes^[5] 首次提出单车道跟随引导模型以来, 在过去几十年中, 已经开发了多种的跟驰模型。其中大多数都是仅考虑到周围车辆带来的影响。例如, Chandler 等^[6] 提出了“加利

福尼亚”模型, 该模型考虑了车辆在跟驰行为中存在的延迟时间; Newell^[7] 利用速度-车头间距关系的一般数学方程来模拟车辆跟驰行为; Bando 等^[8] 开发了最优速度(OV)模型, 该模型由于其简单、直观和创新性而被广泛采用; Jiang 等^[9] 通过考虑正速度差异因子, 建立了全速差(FVD)模型, FVD模型因可用于表征较为真实的汽车跟踪行为, 而被认为优于OV模型和广义力(GF)模型^[10]。因此, 许多研究人员都是通过改进FVD模型, 来更好地描述真实的交通环境。例如, Zhou^[11] 通过引入驾驶员的视角来扩展FVD模型, 通过有效地解释一些复杂的交通现象, 使新模型显得更加符合实际。Yu 和 Shi^[12] 将相对速度

收稿日期: 2018-11-16; 录用日期: 2018-11-23; 网络出版时间: 2018-11-30 15:33

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20181128.1001.001.html

基金项目: 国家自然科学基金(U1564212)

*通信作者. E-mail: wangshuai@buaa.edu.cn

引用格式: 李宏刚, 高哈尔·达吾力, 王帅, 等. 考虑临近车道行人对交通流影响的改进跟驰模型[J]. 北京航空航天大学学报, 2019, 45(2): 422-428. LI H G, Gaoher · DAWULI, WANG S, et al. An improved car-following model considering effect of pedestrians of adjacent lane on traffic flow [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(2): 422-428 (in Chinese).

波动纳入其中对 FVD 模型进行改进,使得能满足先进自适应巡航控制系统的设计。

最近十年来,人车混合交通的情况引起学者的兴趣,但其中的大多数都专注于不受控制的十字路口和人行横道。例如,Helbing 等^[13]通过观察过马路的行人,从宏观动态的角度研究了行人和车辆流动的振荡和延迟;Chen 等^[14]开发了一种改进的双向行人模型来描述不受控制的中间人行横道上车辆和行人之间的相互作用;Zhang 和 Chang^[15]设计了一个合作规划模型,将行人和车辆纳入疏散网络;Ito 和 Nishinari^[16]基于元胞自动机(CA)模型描述了拥挤交叉路口的车辆和行人之间的相互作用。

毋庸置疑,横跨公路的行人肯定会对车辆的行为产生影响,而行走相邻车道上的行人也会对驾驶行为产生一些影响。行驶过程中驾驶员必须注意周围车辆以及相邻车道上的行人或自行车,以避免任何潜在的冲突。在机动车和非机动车混合行驶的复杂环境中,驾驶行为可能有很大差异。驾驶员会依据车辆与行人之间距离的变化来调整相应的速度,以此保证车辆的安全驾驶。然而目前在这个方向存在的研究工作很少。虽然 Jiang 和 Wu^[17-18]使用格子模型模拟了进入狭窄通道的汽车中的类似现象,但因为他们的关注点在疏散时行人的行为,所以并不能获取混合交通流中车辆的动态特征。

考虑到在相邻车道上存在行人或自行车的情况在城市中很常见,因此,有必要研究这种人车行为。本文提出一种改进跟驰模型,该模型考虑了行人和自行车在邻车道上对跟驰车辆产生的影响。改进后的模型对城市复杂混合车辆-行人交通的安全性和控制性有很大的帮助。例如,由于在这种复杂的驾驶环境中存在异常驾驶行为,该模型可用于确定更适合于由行人和自行车包围的城市街道的限制车速。此外,该模型还可用于改善对行人和自行车在相邻车道上行驶的时间估计。这对于城市交通的控制是非常必要的。

1 模型构建

在本节中,假设车辆与行人混合行驶在同一条道路上,并且二者之间无任何显著的物理隔离,如图1所示,其中 x_n 和 s_m 分别为汽车 n 和行人 m 的位置; $\Delta x_n = x_{n+1} - x_n$ 为汽车 n (即跟驰车辆)与汽车 $n+1$ (即前向车辆)之间的距离; $\Delta x_{n,m}^W$ 为

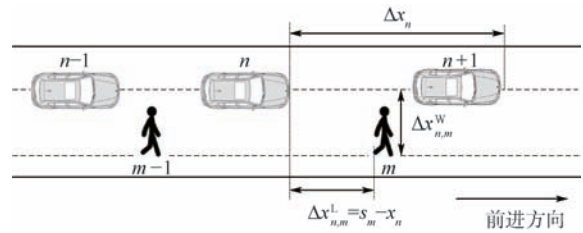


图1 行人车辆混合流的示意图

Fig.1 Sketch map of pedestrian-vehicle mixed flow

车辆 n 与行人 m 之间的横向行驶距离; $\Delta x_{n,m}^L = s_m - x_n$ 为车辆 n 与行人 m 之间的纵向行驶距离。

本文基于 Newell^[7]和 Whitham^[19]提出的优化速度模型,引入车辆与横向行人之间的横向与纵向车头间距,得到如下改进跟驰模型:

$$\frac{dx_n(t + \tau)}{dt} = V(\Delta x_n(t), \Delta v_n(t), \Delta x_{n,m}^L(t), \Delta x_{n,m}^W(t)) \quad (1)$$

式中: V 为考虑连续车辆的车间距与相对速度,以及车与邻道行人的横向与纵向距离的最优速度函数; $\Delta v_n(t) = v_{n+1}(t) - v_n(t)$ 为前车与后车的速度差(即相对速度); τ 为延迟时间。

对式(1)进行显式化,本文采用文献[20]中提到的最优速度函数:

$$V(\Delta x_n) = \frac{v_{\max}}{2} (\tanh(\Delta x_n - h_c) + \tanh(h_c)) \quad (2)$$

式中: h_c 为安全距离; v_{\max} 为车辆最大速度。可知

$V(\Delta x_n)$ 为单调递增函数,因为 $V'(\Delta x_n) = \frac{v_{\max}}{2} (1 - \tanh^2(\Delta x_n - h_c))$ 为 $V(\Delta x_n)$ 的一阶导数,并且无论 Δx_n 的值如何,它都是非负的。另可知式(2)

具有上限,即 $\lim_{\Delta x_n \rightarrow \infty} V(\Delta x_n) = \frac{v_{\max}}{2} (1 + \tanh(h_c))$ 。

正如 Sawada^[21]和 Li 等^[22]所提出的理论,多参数最优速度函数可以通过在单个参数的最优速度函数中加入可调权值来进行线性化。因此,将式(1)的右侧线性化可以得到

$$V(\Delta x_n, \Delta v_n, \Delta x_{n,m}^L, \Delta x_{n,m}^W) = pV(\Delta x_n) + qV(\Delta x_{n,m}^L) + \gamma V(\Delta x_{n,m}^W) + \lambda \Delta v_n \quad (3)$$

式中: p 、 q 和 γ 分别为连续车辆车头间距、车辆和行人之间的横向间距以及纵向间距的最优函数的权重; λ 为连续车辆之间相对速度的响应系数。注意: λ 为一个与时间、速度和位置都无关的常数。

将式(1)进行泰勒展开,可以得到以下微分方程:

$$\frac{d^2 x_n(t)}{dt^2} = \alpha \left[pV(\Delta x_n) + qV(\Delta x_{n,m}^L) + \gamma V(\Delta x_{n,m}^W) - \frac{dx_n(t)}{dt} \right] + \mu \Delta v_n \quad (4)$$

式中: α 为驾驶员反应度, 且 $\alpha = 1/\tau, \mu = \lambda/\tau$ 。

注意, $p + q + \gamma = 1$ 。同时本文假设:

$$\begin{cases} q = 0 & \Delta x_{n,m}^L > l_c \\ \gamma = 0 & \Delta x_{n,m}^W > d_c \end{cases}$$

式中: l_c 和 d_c 分别为观察到的车辆与行人之间预定的临界纵向和横向间距。如果跟驰车辆和前方行人之间的横向行驶距离太大(即大于临界距离 d_c), 则当车辆超过行人时, 行人对车辆没有影响。类似地, 如果观察到的汽车与前方行人之间的纵向行驶距离大于临界距离 l_c , 则相邻道路上的行人对驾驶员没有影响。 $p + q + \gamma = 1$ 表明除了这3个参数之外, 驾驶员不会考虑周围其他影响因素^[23]。

为了计算方便, 式(4)需要使用非对称前向离散化方法, 得到如下差分方程:

$$x_n(t + 2\tau) - x_n(t + \tau) = \tau [pV(\Delta x_n(t)) + qV(\Delta x_{n,m}^L(t)) + \gamma V(\Delta x_{n,m}^W(t))] + \lambda(\Delta x_n(t + \tau) - \Delta x_n(t)) \quad (5)$$

2 线性稳定性分析

在稳定性分析中, 首先假定车辆和行人的流量是均匀的, 不存在换道和超车行为。根据文献[10]所述, 在一个理想的均匀状态下, 每个行人和车辆将保持一定的空间留给前面和后面的行人。因此, 本文给出式(5)的均匀稳态解如下:

$$x_n^{(0)} = hn + V(h)t \quad h = L/N_1 \quad (6)$$

$$s_m^{(0)} = h_p m + U(h_p)t \quad h_p = L/N_2 \quad (7)$$

式中: h 和 h_p 分别为连续车辆和连续行人之间的间距; N_1 和 N_2 分别为混合流动中车辆和行人的数量; V 和 U 分别为车辆和行人的最优速度; L 为道路长度。

通过在式(6)和式(7)的解中添加小偏差 $y_n(t)$ 和 $\delta_m(t)$, 给出车辆和行人位置的方程分别为

$$x_n(t) = x_n^{(0)}(t) + y_n(t) \quad (8)$$

$$s_m(t) = s_m^{(0)}(t) + \delta_m(t) \quad (9)$$

式中: $|y_n(t)| \ll 1, |\delta_m(t)| \ll 1$ ^[8]。

为了更好地分析线性稳定性, 式(5)可以改写为

$$\Delta x_n(t + 2\tau) - \Delta x_n(t + \tau) = \tau [p(V(\Delta x_{n+1}(t)) - V(\Delta x_n(t))) + q(V(\Delta x_{n+1,m+1}^L(t)) - V(\Delta x_{n,m}^L(t))) + r(V(\Delta x_{n+1,m+1}^W(t)) -$$

$$V(\Delta x_{n,m}^W(t))) + \lambda(\Delta x_{n+1}(t + \tau) - \Delta x_{n+1}(t) - \Delta x_n(t + \tau) + \Delta x_n(t)) \quad (10)$$

基于式(5)、式(6)和式(8), 可将式(10)改写为

$$\Delta y_n(t + 2\tau) - \Delta y_n(t + \tau) = \tau [pV'(\Delta y_{n+1}(t) - \Delta y_n(t)) + q\beta V'(\Delta y_{n+1}(t) - \Delta y_n(t)) + \gamma\eta V'(\Delta y_{n+1}(t) - \Delta y_n(t))] + \lambda(\Delta y_{n+1}(t + \tau) - \Delta y_{n+1}(t) - \Delta y_n(t + \tau) + \Delta y_n(t)) \quad (11)$$

式中: V' 为 $V(\Delta x_n)$ 在 $\Delta x_n = h$ 时的一阶导; β 为平均强度系数, 描述了连续车辆的车头间距与观察到的汽车和前方行人的纵向间距之间的干扰关系; η 为描述观察到的汽车与前方行人之间的干扰关系的平均强度系数。注意: 本文也应用了平均场理论, 该理论把任何一个因子的所有相互作用替换为平均或有效的单一相互作用, 可以将2种类型的粒子(即行人和车辆)从宏观视角合并成单个物体^[24]。

假设一个显式函数 $\Delta y_n(t) = A \exp(ikn + zt)$, 方程式(11)可写为

$$e^{2zt} - e^{zt} = \tau [pV'(e^{ik} - 1) + q\beta V'(e^{ik} - 1) + \gamma\eta V'(e^{ik} - 1)] + \lambda(e^{ik+zt} - e^{ik} - e^{zt} + 1) \quad (12)$$

式中: A 为波动幅度; k 为波数, 其取值范围是 $0 \leq k \leq \pi$; $z = -i\omega$, ω 为波动的角频率。

通过将 z 展开为 $z = z_1(ik) + z_2(ik)^2 + \dots$ ^[25], 式(12)可转换为

$$\begin{cases} z_1 = (p + q\beta + \gamma\eta)V' \\ z_2 = -\frac{3\tau z_1^2}{2} + \frac{z_1}{2} + \lambda z_1 \end{cases} \quad (13)$$

对于长波模式, 如果 z_2 为负, 则均匀稳定流会变得不稳定; 如果 z_2 为正, 则均匀流保持稳定状态。在 $z_2 = 0$ 时设置中性稳定条件

$$\tau = \frac{2\lambda + 1}{3[(p + q\beta + \gamma\eta)V']} \quad (14)$$

式中: 延迟时间 τ 称为临界值, 通常写为 τ_c 。它的倒数被称为临界敏感度 ($\alpha_c = 1/\tau_c$)。对于长波小扰动, 可以导出以下不稳定和稳定的条件:

$$\tau > \frac{2\lambda + 1}{3[(p + q\beta + \gamma\eta)V']} \quad (15)$$

$$\tau < \frac{2\lambda + 1}{3[(p + q\beta + \gamma\eta)V']} \quad (16)$$

在式(1)中, 最优速度 V 的导数 V' 有一个最大值, 即 $v_{\max}/2$ 。将车辆最大速度输入式(16), 发现当满足如下不等式 $\tau < \frac{2(2\lambda + 1)}{3[(p + q\beta + \gamma\eta)v_{\max}]}$

时,均匀车流在任何车辆密度下都是稳定的。需要注意的是,本文中式(14)的稳定性条件同 Bando 提出 OV 模型稳定性条件^[8]以及 Yu 提出的相对速度(RV)模型的稳定性结果^[26]相似。

图2显示了具有 (λ, p, q, γ) 不同组合的中性稳定性曲线, $v_{\max} = 3 \text{ m/s}$, $h_c = 4 \text{ m}$, $\beta = 0.5$, $\eta = 0.1$ 。实线对应于 $(0, 1, 0, 0)$, 与 OV 模型一致;虚线对应于 $(0.2, 1, 0, 0)$, 与 RV 模型一致;圆线对应于 $(0.2, 0.8, 0.1, 0.1)$, 考虑了相邻行人的相互作用。从这3条曲线中可以看出,改进模型的特征曲线(即带有圆圈标记的线)具有最大的稳定区域(即最小的不稳定区域)。该发现意味着考虑相邻车道行人影响的跟驰模型将比传统跟驰模型具有更好的稳定性。

在图3中为车头间距-敏感度的中性稳定性曲线,每条曲线的顶点表示临界点。在图中呈现为当 $p = 0.6$ 是定值时,随着 q 的增大以及 γ 的减小,稳定区域增大。曲线以外的区域表示交通流畅通的稳定状态;曲线以内区域表示交通流量拥挤的不稳定状态。图3表明相对于横向

的行人影响因素,驾驶员更关注前车的运动变化行为。

3 仿真实验

为了更好地模拟真实的交通行为,本文应用了 Konishi 等^[27]提出的另一个最优速度函数,具体方程如下:

$$V^{\text{op}}(\Delta x_n(t)) = \frac{v_{\max}}{2} \left[1 + H \left(2 \frac{\Delta x_n(t) - \psi}{\xi} \right) \right] \quad (17)$$

式中: ξ 为距离参数; ψ 为车辆安全距离;饱和度函数 $H(x)$ 被描述为

$$H(x) = \begin{cases} 1 & x > 1 \\ x & -1 \leq x \leq 1 \\ -1 & x < -1 \end{cases} \quad (18)$$

实际上,方程式(17)是由 $V(\Delta x) = 16.8 \times (\tanh(0.086(\Delta x - 25)) + 0.913)$ 推导出来的,其来源于文献[28-30]中提到的日本高速公路上的观测数据。这些值与本文中车辆的仿真参数一致,如表1所示。表2是行人的相关仿真参数。

假设一个由11辆车组成的车队在混合行人车道上行驶,用以验证受到相邻道路行人影响的交通模式的变化。为了更好地验证所提出的模型,提供了2种模拟情景:

情景1 街道上只有少数行人的情景。车队中车辆逐个经过前向相邻的行人。在这种情况下,本文研究了几个受相邻行人影响的车辆特征的变化。这种情况适用于在夜间或雨天在路边行走的少数行人的实际交通情况。

情景2 大量行人行走的情景。在这种情况下

表1 车辆仿真参数

Table 1 Simulation parameters of vehicles

参数	数值	物理意义
ψ/m	25.0	车辆安全距离
ξ/m	23.3	距离参数
$v_{\max}/(\text{m} \cdot \text{s}^{-1})$	33.6	车辆最大速度
$v/(\text{m} \cdot \text{s}^{-1})$	20	车辆初始化速度
α_c/s^{-1}	2.0	敏感度
y_{\min}/m	7.02	最小车头时距
T/s	0.1	抽样时间

表2 行人仿真参数

Table 2 Simulation parameters of pedestrians

参数	数值	物理意义
$v^p/(\text{m} \cdot \text{s}^{-1})$	1.5	行人初始速度
$v_{\max}^p/(\text{m} \cdot \text{s}^{-1})$	3	行人最大速度
φ/m	7.6	行人安全距离

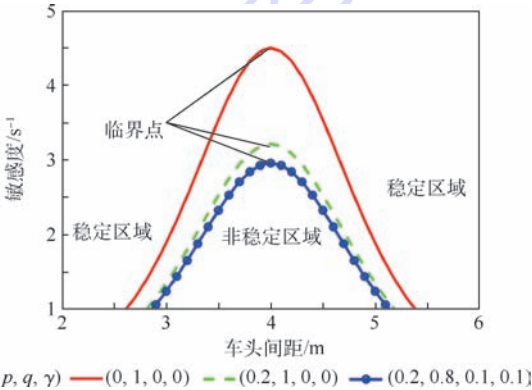


图2 不同参数组合 (λ, p, q, γ) 下的中性稳定性曲线

Fig.2 Neutral stability curves with different parameter sets of (λ, p, q, γ)

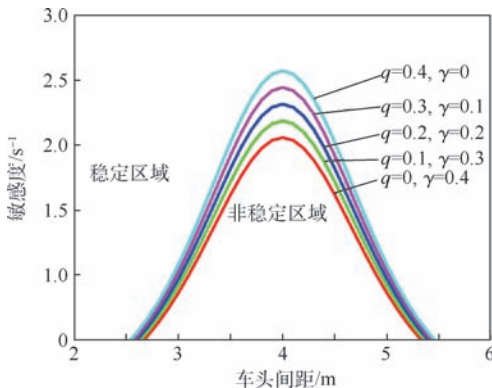


图3 不同 q 和 γ 值的中性稳定性曲线

($v_{\max} = 3 \text{ m/s}$, $h_c = 4 \text{ m}$)

Fig.3 Neutral stability curves with different values of

q and γ ($v_{\max} = 3 \text{ m/s}$, $h_c = 4 \text{ m}$)

下,本文提出研究以下的3个时间间隔内受相邻行人影响的车辆行为:车辆开始进入,车辆驶过,车辆离开。适用于行人较多的道路。

对于这2种情况,初始条件设置为: $v_n(0) = 20 \text{ m/s}$, $x_{11}(0) = 0$ (即车队的最后一辆车位于原点),当 $n \neq 1, \Delta x_1(t) = v\xi/v_{\max} - \xi/2 + \psi, \Delta x_n(0) = v\xi/v_{\max} - \xi/2 + \psi$ 。假设所有车辆在接近相邻行人之前具有相同的初始速度和车头间距。也就是说,所有车辆在早期均匀运行且没有额外的干扰。当它们开始超过考虑了行人道的相邻行人时,交通流的特征将发生巨大变化。

3.1 情景 1

图4和图5分别通过假设 $p = 0.6, q = 0.3, \gamma = 0.1$ 和 $p = 0.6, q = 0.2, \gamma = 0.2$ 时第2、第5、第8和第11辆(即最后一辆汽车)的速度随时间的变化情况。

图4与图5表示当车辆超过相邻行人时,均匀行驶车辆的交通冲击波产生波动,且所选择的车辆的速度变化是不同的。该图基本上可证明

振荡行为和交通拥堵的演变。可以看到在 $p = 0.6, q = 0.2, \gamma = 0.2$ 的条件下(图5),速度波动的幅度相对较小。这表明与图4中的情况相比,汽车能更平稳地运行并且恢复稳定状态。

3.2 情景 2

假设行人沿着路边均匀且稳定地行走。当领先的车辆遇到行人队尾时,车辆的振荡行为发生。为了更好地观察和分析,本文只选择其中的4辆车,图6为情景2下的第2、第5、第8和第11辆车的速度随时间的变化。

在 $100 \sim 400 \text{ s}$ 期间,车辆遇到行人,然后车速减慢。在 $400 \sim 700 \text{ s}$ 期间,所有车辆都在人群队列旁行驶,其速度仍然相对较低。在这个阶段,行人之间的间距比车辆车头间距小得多,车速比所有行人都快。从宏观角度来看,行人流可以看作是一个长长的障碍。事实上,在这个时期车流受到每个行人的影响,由于间隔极短(行人间距离短),因此每辆车的速度波动曲线看起来像一条直线。在 $700 \sim 1000 \text{ s}$ 期间,汽车连续经过人群,并通过一系列复杂的速度波动后将其速度恢复到初始速度。

为了进一步分析交通流量行为,图7显示了随着时间的变化,第1辆车和其余每辆跟驰车之间的距离变化。如图7所示,左边第1条线为第2辆车与第1辆车的距离差,向右以此类推,由于距离波动,每条曲线有2个凸起。当第1辆车遇到行人并且必须降低速度时,发生第1次扰动导致前车和后车之间的距离变小。当前车离开人群并恢复到其初始速度时,发生第2次扰动,导致前车和后车之间的距离变大。

经仿真实验发现,在行人与车辆混合车道上,交通流会受到相邻行人的影响。速度和车头间距随着时间的推移而变化。结果表明,当行人在邻

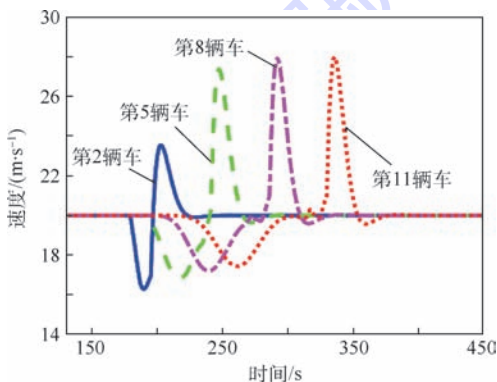


图4 当 $p = 0.6, q = 0.3, \gamma = 0.1$ 时,车辆速度随时间变化曲线

Fig. 4 Variation curves of vehicle velocity with time for $p = 0.6, q = 0.3, \gamma = 0.1$

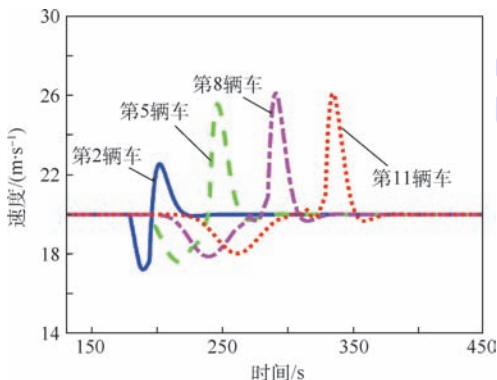


图5 当 $p = 0.6, q = 0.2, \gamma = 0.2$ 时,车辆速度随时间变化曲线

Fig. 5 Variation curves of vehicle velocity with time for $p = 0.6, q = 0.2, \gamma = 0.2$

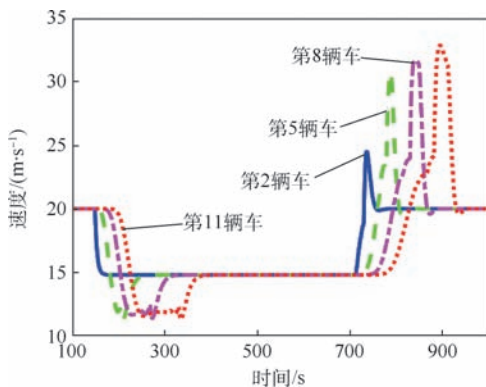


图6 场景2下的车辆速度随时间变化曲线

Fig. 6 Variation curves of vehicle velocity with time in scene 2

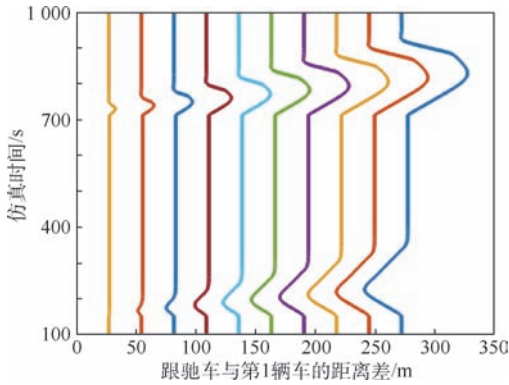


图 7 通过行人人群时的车队时空图

Fig. 7 Space-time plots for a motorcade passing by pedestrian crowd

近的车道上行走时,司机通常会放慢车速,以确保行人和车辆自身的安全。这表明,通过考虑相邻行人的横向和纵向影响,可以改善交通安全,有助于解决交通拥堵问题。

4 结 论

1) 本文提出了一种改进跟驰模型,该模型考虑了基于 OV 模型的行人和自行车在相邻车道上行走时对行驶车辆产生的影响。特别地,将车辆与行人之间的横向和纵向距离引入到传统的跟驰模型中。

2) 利用所提出的新模型,进一步进行线性稳定性分析,获得一系列中性稳定性曲线。

3) 通过仿真评估 2 种实际场景的模型能够成功描述混合交通流量行为,提高交通流量,避免交通事故和堵塞。

参考文献 (References)

- [1] NAGATANI T. The physics of traffic jams[J]. Reports on Progress in Physics, 2002, 65(9): 1331.
- [2] TREIBER M, KESTING A, HELBING D. Delays, inaccuracies and anticipation in microscopic traffic models[J]. Physica A: Statistical Mechanics and Its Applications, 2006, 360(1): 71-88.
- [3] NGODUY D. Analytical studies on the instabilities of heterogeneous intelligent traffic flow[J]. Communications in Nonlinear Science and Numerical Simulation, 2013, 18(10): 2699-2706.
- [4] REUSCHEL A. Vehicle movements in a platoon with uniform acceleration or deceleration of the lead vehicle[J]. Zeitschrift des Oesterreichischen Ingenieur-und Architekten-Vereines, 1950, 95: 50-62.
- [5] PIPES L A. An operational analysis of traffic dynamics[J]. Journal of Applied Physics, 1953, 24(3): 274-281.
- [6] CHANDLER R E, HERMAN R, MONTROLL E W. Traffic dynamics: Studies in car following[J]. Operations Research, 1958, 6(2): 165-184.
- [7] NEWELL G F. Nonlinear effects in the dynamics of car following[J]. Operations Research, 1961, 9(2): 209-229.
- [8] BANDO M, HASEBE K, NAKAYAMA A, et al. Dynamical model of traffic congestion and numerical simulation[J]. Physical Review E, 1995, 51(2): 1035-1045.
- [9] JIANG R, WU Q, ZHU Z. Full velocity difference model for a car-following theory[J]. Physical Review E, 2001, 64(1): 017101.
- [10] HELBING D, TILCH B. Generalized force model of traffic dynamics[J]. Physical Review E, 1998, 58(1): 133-138.
- [11] ZHOU J. An extended visual angle model for car-following theory[J]. Nonlinear Dynamics, 2015, 81(1): 549-560.
- [12] YU S, SHI Z. An improved car-following model considering relative velocity fluctuation[J]. Communications in Nonlinear Science and Numerical Simulation, 2016, 36: 319-326.
- [13] HELBING D, JIANG R, TREIBER M. Analytical investigation of oscillations in intersecting flows of pedestrian and vehicle traffic[J]. Physical Review E, 2005, 72(4): 046130.
- [14] CHEN P, WU C, ZHU S. Interaction between vehicles and pedestrians at uncontrolled mid-block crosswalks[J]. Safety Science, 2016, 82: 68-76.
- [15] ZHANG X, CHANG G. A dynamic evacuation model for pedestrian-vehicle mixed-flow networks[J]. Transportation Research Part C: Emerging Technologies, 2014, 40: 75-92.
- [16] ITO H, NISHINARI K. Totally asymmetric simple exclusion process with a time-dependent boundary: Interaction between vehicles and pedestrians at intersections[J]. Physical Review E, 2014, 89(4): 042813.
- [17] JIANG R, WU Q. The moving behavior of a large object in the crowds in a narrow channel[J]. Physica A: Statistical Mechanics and Its Applications, 2006, 364: 457-463.
- [18] JIANG R, WU Q. Interaction between vehicle and pedestrians in a narrow channel[J]. Physica A: Statistical Mechanics and Its Applications, 2006, 368(1): 239-246.
- [19] WHITHAM G B. Exact solutions for a discrete system arising in traffic flow[C] // Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences. London: The Royal Society of London, Mathematical, Physical and Engineering Sciences, 1990: 49-69.
- [20] GE H X, DAI S Q, DONG L Y, et al. Stabilization effect of traffic flow in an extended car-following model based on an intelligent transportation system application[J]. Physical Review E, 2004, 70(6): 066134.
- [21] SAWADA S. Generalized optimal velocity model for traffic flow[J]. International Journal of Modern Physics C, 2002, 13(1): 1-13.
- [22] LI Y, ZHANG L, ZHENG H, et al. Evaluating the energy consumption of electric vehicles based on car-following model under non-lane discipline[J]. Nonlinear Dynamics, 2015, 82(1): 629-641.
- [23] PENG G H, SUN D H. A dynamical model of car-following with the consideration of the multiple information of preceding cars[J]. Physics Letters A, 2010, 374: 1694-1698.
- [24] TANG T, HUANG H, GAO Z. Stability of the car-following model on two lanes[J]. Physical Review E, 2005, 72(6):

- 066124.
- [25] LI Z P, LIU Y C. Analysis of stability and density waves of traffic flow model in an its environment[J]. The European Physical Journal B: Condensed Matter and Complex Systems, 2006, 53 (3):367-374.
- [26] YU X. Analysis of the stability and density waves for traffic flow [J]. Chinese Physics, 2002, 11 (11):1128-1134.
- [27] KONISHI K, KOKAME H, HIRATA K. Coupled map car-following model and its delayed-feedback control[J]. Physical Review E, 1999, 60(4):4000-4007.
- [28] BANDO M, HASEBE K, NAKANISHI K, et al. Analysis of optimal velocity model with explicit delay[J]. Physical Review E, 1998, 58(5):5429-5450.
- [29] BANDO M, HASEBE K, NAKANISHI K, et al. Phenomenological study of dynamical model of traffic flow[J]. Journal de Physique I, 1995, 5(11):1389-1399.
- [30] TADAKI S, KIKUCHI M, SUGIYAMA Y, et al. Coupled map traffic flow simulator based on optimal velocity functions[J]. Journal of the Physical Society of Japan, 1998, 67(7):2270-2276.

作者简介:

李宏刚 男,博士研究生。主要研究方向:智能交通系统、交通跟驰模型。

王帅 女,助理研究员。主要研究方向:人工智能、智慧城市。

An improved car-following model considering effect of pedestrians of adjacent lane on traffic flow

LI Honggang¹, Gaohaer · DAWULI¹, WANG Shuai^{2,*}, YU Guizhen¹, WANG Pengcheng¹

(1. School of Transportation Science and Engineering, Beihang University, Beijing 100083, China;

2. School of Computer Science and Engineering, Beihang University, Beijing 100083, China)

Abstract: This paper proposes an improved vehicle dynamics model based on the current situation of mixed-pedestrian-vehicle driving in real life. The model considered the influence of pedestrians, bicycles, etc. on the adjacent lanes or in lateral directions on the driving behavior of the main road vehicles. Based on the classical optimal velocity model, this paper modelled the actual problem and introduced the lateral distance and longitudinal distance between the main road vehicle and the pedestrian as parameters into the optimal velocity model, and expanded to obtain the improved model. In order to verify the stability and effectiveness of the model, this paper used the linear stability theory to derive the stability and non-stable conditions of the model, draw a neutral stability curve, and visually describe the size of the traffic flow stability area. It is shown that the improved model considering the lateral pedestrian interference factor is more stable than the traditional car-following model considering only the single-lane vehicle factor, and the steady-state region caused by the change of different parameters also changes. This paper adopted a more realistic optimization of the velocity equation and described the driving behavior of the vehicle through simulation experiments. The simulation experiment lists two kinds of practical scenarios: when pedestrians are scarce and when pedestrians are more. The velocity-time curve of the vehicle and the space (distance)-time curve of the vehicle were plotted separately. The experimental results show that the lateral pedestrians do interfere with the normal driving of the vehicle; in different scenarios, the number of pedestrians will also affect the driving behavior of the vehicle to varying degrees.

Keywords: car-following model; optimal velocity model; linear stability theory; mixed-pedestrian-vehicle traffic; lateral pedestrian interference factor

Received: 2018-11-16; Accepted: 2018-11-23; Published online: 2018-11-30 15:33

URL: kns.cnki.net/kcms/detail/11.2625.V.20181128.1001.001.html

Foundation item: National Natural Science Foundation of China (U1564212)

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

《北京航空航天大学学报》征稿简则

《北京航空航天大学学报》是北京航空航天大学主办的以航空航天科学技术为特色的综合性自然科学学术期刊(月刊)。本刊以反映航空航天领域研究成果与动态、促进学术交流、培养科技人才和推动科技成果向社会生产力转化为办刊宗旨。本刊为中国自然科学技术核心期刊,并被 Ei Compendex 等国内外权威文献检索系统收录。本刊向国内外公开发行人,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: jbuaa@buaa.edu.cn

http://bhxb.buaa.edu.cn

http://www.buaa.edu.cn

《北京航空航天大学学报》 第五届编辑委员会

主任(主编): 赵沁平

(以下按姓氏笔划为序)

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45卷第2期 2019年2月

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS (JBUA)

(Monthly, Started in 1956)

Vol.45 No.2 February 2019

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 50.00元/期

ISSN 1001-5965



9 771001 596199