ISSN 1001-5965
CN 11-2625/V

北京航空航天大學学报

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

2023-09
Vol.49 No.9

北京航空航天大学学报

第 49 卷　第 9 期　（总第 367 期）　2023 年 9 月

目　　次

基于 TRL 的非标准接口机载射频线缆测试方法 …… 石旭东，李瑞蒲，赵宏旭，张浩天（2207）
新冠疫情背景下航空物流网络的鲁棒优化 …… 张锦，张哲睿，洪治潮，杨文广，闫妍（2218）
锂离子电池热失控气体快速检测及危险性分析方法 ……
…… 张青松，刘添添，郝朝龙，曲奕润，张伟，陈达（2227）
航班保障因果分析框架 …… 邢志伟，张林，罗谦，夏欢，文涛，张涛（2234）
基于显著感知与一致性约束的目标跟踪算法 …… 国强，吴天昊，徐伟，CHORNOGOR Leonid（2244）
10 cm 口径发散磁场离子推力器放电模型 …… 孙明明，耿海，胡竟，杨福全，龙建飞（2258）
超声速民机发动机短舱布局对声爆的影响 …… 肖天航，徐雅楠，朱震浩，邓双厚（2267）
利用改进遗传算法的软件故障定位辅助测试用例生成方法 …… 杨波，何宇泽，许福，陈志泊（2279）
基于 TESLA 协议的 BDSBAS 电文认证技术 …… 陈潇，田翔，罗瑞丹，刘婷，宋佳慧，吴海涛（2289）
基于变权重-正态云模型的飞机轮胎滑水风险 …… 李岳，胡宇祺，蔡靖，戴轩（2299）
基于多 Agent 的航空机群保障仿真评估分析 …… 丁刚，崔利杰，韩诚，王新河，张峰（2306）
基于 GEMD 与改进 PCNN 的红外与可见光图像融合 …… 杨艳春，李小苗，党建武，王阳萍（2317）
推进剂储罐裂纹缺陷非概率可靠性分析方法 …… 辛腾达，崔村燕，刘阳，同江，段永胜（2330）
基于虚拟飞行的混合翼身融合布局操稳特性 …… 王延灵，冯帅，卜忱，沈彦杰，陈昊，芦士光（2337）
SERF 原子自旋惯性测量检测误差分析及抑制 …… 邢力，全伟，宋天骁，蔡庆中，叶文（2345）
基于脊背特征的发动机低转速特性扩展方法 …… 王佳美，郭迎清，于华锋（2351）
含间隙全动舵面的非线性颤振模式及被动抑制方法 …… 白刘月，吴志刚，杨超（2361）
跨音涡轮转子叶尖间隙内流动分析与建模 …… 轩笠铭，邹正平，曾飞（2374）
车路耦合荷载下沥青混凝土路面振动响应 …… 董倩，程少锋，张献民，包伊婷（2385）
四旋翼无人机的自适应故障诊断与估计 ……
…… 王莉娜，刘贞报，院金彪，党庆庆，江飞鸿，王宝栋（2395）
机器人辅助视网膜下注射系统的设计与实现 …… 王朝董，广晨汉，王丽强，宗俊杰，郑昱，杨洋（2406）
基于综合评价优化的民机顶层需求指标权衡 …… 范周伟，余雄庆，戴亚林（2415）
面向制造成本的变刚度复合材料结构优化设计 …… 陈世泽，李道春，向锦武（2423）
基于直角网格伴随自适应的声爆预测 …… 朱震浩，肖天航，徐雅楠，邓双厚，张珍铭，余雄庆（2432）
基于多重检测的无人机抗遮挡目标跟踪算法 …… 张博恒，柴栋栋，孟令博，孙明健（2442）
中国民用航空器 CO_2 减排潜力的区域划分 …… 曾雯，胡荣，宋文，刘志昊，张军峰（2455）
IFB 不确定性对双频码载分歧监测的影响分析 …… 康凯，王志鹏，方堃，朱衍波（2463）

模块化抛物面折展机构设计与优化 …… 何天宇，董洋，王辉，赵军鹏，王春洁（2473）
DES 与 DDES 在湍流分离中的原理与性能研究 …… 宋汉奇，张恺玲，马鸣，阎超（2482）
一种历史信息特征敏感的行人迭代检测方法 …… 戴佩哲，刘翔，张星，尚岩峰，赵静文，王诗雨（2493）
基于 L_1-ITD 的尾座式无人机姿态控制 …… 王姝旸，张晶，杨凌宇（2501）
民用飞机客舱 CO_2 分布特性分析 …… 高经诚，陈维建，胡万俊，冯诗愚，刘豪正（2510）
基于弱监督的遥感图像镶嵌质量盲评价 …… 潘林朋，谢凤英，赵薇薇，周颖，刘畅，王艳（2518）
新型变摩擦滑橇式飞行器滑跑纠偏 …… 王宇晟，印寅，梁涛涛，魏小辉（2527）
基于 CFD/CSD 耦合的高速射弹尾拍载荷特性研究 …… 许云涛，檀大林，杨超（2539）

期刊基本参数： CN 11-2625/V * 1956 * m * A4 * 340 * zh * P * ￥50.00 * 400 * 35 * 2023-09

（**编　辑**　张　嵘　孙　芳　苏　磊　卞欢欢　王　茜　李亚泰　李艳霞）

JOURNAL OF BEIJING UNIVERSITY OF
AERONAUTICS AND ASTRONAUTICS

Vol. 49 No. 9 （Sum 367） September 2023

CONTENTS

Non-standard interface aviation RF cable test method based on TRL
…… SHI Xudong，LI Ruipu，ZHAO Hongxu，ZHANG Haotian （2207）

Robust optimization of aviation logistics network in context of COVID-19 pandamic
…… ZHANG Jin，ZHANG Zherui，HONG Zhichao，YANG Wenguang，YAN Yan （2218）

Rapid detection analysis method of thermal runaway gas composition and risk of lithium ion battery
…… ZHANG Qingsong，LIU Tiantian，HAO Chaolong，QU Yirun，ZHANG Wei，CHEN Da （2227）

Causal analysis framework of flight service
…… XING Zhiwei，ZHANG Lin，LUO Qian，XIA Huan，WEN Tao，ZHANG Tao （2234）

Target tracking algorithm based on saliency awareness and consistency constraint
…… GUO Qiang，WU Tianhao，XU Wei，CHORNOGOR Leonid （2244）

Discharge model of divergent magnetic field ion thruster of 10 cm diameter
…… SUN Mingming，GENG Hai，HU Jing，YANG Fuquan，LONG Jianfei （2258）

Effect of engine nacelle layout on sonic boom of supersonic transport
…… XIAO Tianhang，XU Yanan，ZHU Zhenhao，DENG Shuanghou （2267）

Using improved genetic algorithm for software fault localization aided test case generation
…… YANG Bo，HE Yuze，XU Fu，CHEN Zhibo （2279）

Design of message authentication based on TESLA protocol for BDSBAS
…… CHEN Xiao，TIAN Xiang，LUO Ruidan，LIU Ting，SONG Jiahui，WU Haitao （2289）

Hydroplaning risk of aircraft tire based on variable weight theory-normal cloud model
…… LI Yue，HU Yuqi，CAI Jing，DAI Xuan （2299）

Simulation evaluation and analysis of aircraft group support based on multi-agent
…… DING Gang，CUI Lijie，HAN Cheng，WANG Xinhe，ZHANG Feng （2306）

Infrared and visible image fusion based on GEMD and improved PCNN
…… YANG Yanchun，LI Xiaomiao，DANG Jianwu，WANG Yangping （2317）

Non-probabilistic reliability analysis method for propellent tank with crack defect
…… XIN Tengda，CUI Cunyan，LIU Yang，TONG Jiang，DUAN Yongsheng （2330）

Stability characteristics of hybrid wing-body aircraft based on virtual flight
…… WANG Yanling，FENG Shuai，BU Chen，SHEN Yanjie，CHEN Hao，LU Shiguang （2337）

Error analysis and suppression of probe system for SERF atomic spin co-magnetometer
…… XING Li，QUAN Wei，SONG Tianxiao，CAI Qingzhong，YE Wen （2345）

Extension method of engine low speed characteristics based on backbone features
…… WANG Jiamei，GUO Yingqing，YU Huafeng （2351）

Nonlinear flutter modes and flutter suppression of an all-movable fin with freeplay
…… BAI Liuyue，WU Zhigang，YANG Chao （2361）

Analyzing and modeling flow in tip clearance of transonic turbine rotor
…… XUAN Liming，ZOU Zhengping，ZENG Fei （2374）

Vibration response of asphalt concrete pavement under vehicle-road coupled load
…… DONG Qian，CHENG Shaofeng，ZHANG Xianmin，BAO Yiting （2385）

Adaptive fault diagnosis and estimation for quadrotor UAV
…… WANG Lina，LIU Zhenbao，YUAN Jinbiao，DANG Qingqing，JIANG Feihong，WANG Baodong （2395）

Design and implementation of robot-assisted subretinal injection system
…… WANG Zhaodong，GUANG Chenhan，WANG Liqiang，ZONG Junjie，ZHENG Yu，YANG Yang （2406）

Trade-off for top-level requirements of commercial aircraft using comprehensive evaluation and optimization
…… FAN Zhouwei，YU Xiongqing，DAI Yalin （2415）

Design optimization of tow-steered composite structure targeting on manufacturing cost
…… CHEN Shize，LI Daochun，XIANG Jinwu （2423）

Adjoint-based adaptive Cartesian mesh refinement for sonic boom prediction
…… ZHU Zhenhao，XIAO Tianhang，XU Yanan，DENG Shuanghou，ZHANG Zhenming，YU Xiongqing （2432）

Anti-occlusion target tracking algorithm of UAV based on multiple detection
…… ZHANG Boheng，CHAI Dongdong，MENG Lingbo，SUN Mingjian （2442）

Regional classification of CO_2 emission reduction potential of China's civil aircraft
…… ZENG Wen，HU Rong，SONG Wen，LIU Zhihao，ZHANG Junfeng （2455）

Impact of IFB uncertainty on dual-frequency code carrier divergence monitoring
…… KANG Kai，WANG Zhipeng，FANG Kun，ZHU Yanbo （2463）

Design and optimization of modular parabolic deployable mechanism
…… HE Tianyu，DONG Yang，WANG Hui，ZHAO Junpeng，WANG Chunjie （2473）

Theory and performance research of DES and DDES in turbulent separation
…… SONG Hanqi，ZHANG Kailing，MA Ming，YAN Chao （2482）

An iterative pedestrian detection method sensitive to historical information features
…… DAI Peizhe，LIU Xiang，ZHANG Xing，SHANG Yanfeng，ZHAO Jingwen，WANG Shiyu （2493）

Attitude control law based on L_1-ITD for a tail-sitter UAV
…… WANG Shuyang，ZHANG Jing，YANG Lingyu （2501）

Analysis of CO_2 distribution characteristics in cabin of civil aircraft
…… GAO Jingcheng，CHEN Weijian，HU Wanjun，FENG Shiyu，LIU Haozheng （2510）

Weak supervision based blind remote sensing image mosaic quality assessment
…… PAN Linpeng，XIE Fengying，ZHAO Weiwei，ZHOU Ying，LIU Chang，WANG Yan （2518）

Taxiing deviation-correction control of a new variable-friction equipped-skid aircraft
…… WANG Yucheng，YIN Yin，LIANG Taotao，WEI Xiaohui （2527）

Study on tail-slap load characteristics of high-speed projectile based on CFD/CSD coupling
…… XU Yuntao，TAN Dalin，YANG Chao （2539）

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0645

基于 TRL 的非标准接口机载射频线缆测试方法

石旭东，李瑞蒲，赵宏旭*，张浩天
(中国民航大学 电子信息与自动化学院，天津 300300)

摘　　要：机载射频（RF）线缆是传输信号的主要介质，其回波损耗等性能参数影响飞机通导系统的工作，目前存在多种非标准接口线缆，需采用精密的转接引线，但其对阻抗、传输损耗等参数要求较高，且易发生性能衰减，导致测试结果产生较大误差。因此，针对应用精密转接引线进行测试时存在的问题，提出了基于 TRL 的非标准接口机载射频线缆测试方法。利用级联传输参数矩阵的方法对转接引线与待测线缆进行建模；利用改进的 TRL 校准方法对转接引线的参数进行去嵌入获取待测线缆的散射参数，进而评估机载射频线缆的性能。应用所提方法进行机载射频线缆测试实验，结果表明：除个别谐振点外，回波损耗最大误差约为 1 dB，传输损耗最大误差约为 0.2 dB，验证了所提方法的可行性和有效性。

关　键　词：去嵌入；TRL 校准；传输线理论；S 参数；级联
中图分类号：V260; TM931
文献标志码：A　　**文章编号**：1001-5965(2023)09-2207-11

机载射频(radio frequency, RF)线缆由于具有良好的电气性能和机械性能，在通信、电子对抗及导航中被广泛应用[1]，特别是在民用飞机通导系统中被用来传输射频信号。随着民用飞机上射频产品越来越多，出现了许多非标准接口射频线缆，回波损耗、传输损耗及特征阻抗是其重要的电气性能指标。使用矢量网络分析仪测量射频线缆的散射参数(S 参数)是评估射频线缆性能的一种重要方法，一根线缆的 S 参数为二阶矩阵，常用 $\boldsymbol{S}_{21}$ 或 $\boldsymbol{S}_{12}$ 表示传输损耗，用 $\boldsymbol{S}_{11}$ 或 $\boldsymbol{S}_{22}$ 表示回波损耗，但矢量网络分析仪通常采用标准 SMA 同轴接口，无法与非标准接口射频线缆连接[2]。

GB/T 17738.1—2013[3]、IEC 60966-1[4] 等主流标准采用精密转接引线实现非标准接口射频线缆与测量仪器的连接，由于精密转接引线对测试结果影响较小，对检测结果不进行后续校准。但精密转接引线对阻抗、传输损耗等参数的要求较高，开发成本高昂，且随使用时长增加易发生性能衰减的情况，会给测试结果引入较大误差。采用非精密转接引线实现非标准接口与标准接口之间的转接，通过去嵌入的方法移除转接引线带来的寄生参量的影响，具有成本低、准确度高等优点。目前，飞机总装部门并不针对单一线缆做性能测试，而是装配完成后进行系统联调，一旦出现故障再逐一排故，浪费装配时间。采用去嵌入测试方法在机载射频线缆装配前对其进行性能评估，可使用任意长度且损耗较大的非精密转接引线，降低测试成本，提高了装配效率。

Bianco 等[5] 对去嵌入方法进行了理论研究，但由于理论不够完善而未得到广泛应用。Franzen 和 Speciale[6] 提出了直通-短路-延迟 (thru-short-delay, TSD) 校准，通过设置直通、短路、延迟 3 种校准件

收稿日期：2021-10-29；**录用日期**：2022-01-14；**网络出版时间**：2022-02-24 15:53
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220222.1850.002.html
基金项目：国家自然科学基金 (51377161)
* **通信作者**. E-mail：zhx2581@163.com

引用格式：石旭东，李瑞蒲，赵宏旭，等. 基于 TRL 的非标准接口机载射频线缆测试方法 [J]. 北京航空航天大学学报，2023，49（9）：2207-2217. SHI X D，LI R P，ZHAO H X，et al. Non-standard interface aviation RF cable test method based on TRL[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（9）：2207-2217（in Chinese）.

来实现对两端转接引线的去嵌入。Engen 和 Hoer[7]在 TSD 校准的基础上进行改进，提出了直通-反射-延迟线 (thru-reflect-line, TRL) 校准，该方法与 TSD 校准的区别主要在于反射校准件可以是理想的短路或开路。TRL 校准由于校准件制作简单被广泛应用于微波元件去嵌入测试[8-11]和矢量网络分析仪校准工作中[7]，但 TRL 校准在非标准射频线缆测试中应用存在局限性，如测量频率的限制导致延迟校准件的长度过短或过长，反射校准件中反射系数的相位难以保证在±90°内[12]。

许多学者针对 TRL 校准存在的问题进行了研究。廖进昆和刘仁厚[13]对 TRL 校准进行了优化，提出了延迟线-反射-延迟线 (line-reflect-line, LRL) 校准，旨在解决难以制作零长度直通校准件的问题。文献 [14-15] 利用多延迟校准件提高 TRL 校准的精度。文献 [16] 针对非理想 TRL 校准件带来的误差进行分析，得到了 S 参数偏差的解析表达式。文献 [17] 提出 TRL 校准的简化模型——直通-延迟线 (thru-line, TL) 校准，在两端转接引线对称的情况下不再需要反射校准件。文献 [18]对 TRL 校准进行了相应改进，并将其应用于二极管的测试中，为应用 TRL 校准测试非标准接口射频线缆提供了思路。文献 [19] 应用 TRL 校准对传输线进行测试，但未解决其局限性问题，应用非均匀传输线算法实现对转接引线的去嵌入，但测试对象为无损传输线，当转接引线的损耗较大时，该方法将不再适用。

综上所述，目前 TRL 校准技术在微波元件的测试领域已得到应用，但其在非标准接口射频线缆测试工作中应用较少且存在一定的局限性。因此，本文首先利用级联传输参数矩阵的方法对转接引线与待测线缆进行建模；然后在传统 TRL 校准的基础上对 TRL 校准在非标准接口射频线缆测试中存在的问题进行改进，采用虚拟传输线插入解决宽带测试和低频测试时延迟校准件相位不达标的问题，采用归一化阻抗解决转接引线特征阻抗与测量仪器端口阻抗不一致的问题，采用 TL 校准解决非理想反射校准件对测试结果带来较大误差的问题；最后应用改进的 TRL 校准测试非标准接口射频线缆，通过仿真和实验验证了本文工作的有效性。

1 级联模型与传统 TRL 校准方法

1.1 级联模型

非标准接口射频线缆的测试模型如图 1 所示。由于非标准接口射频线缆无法与矢量网络分析仪连接，需使用转接引线，转接引线一端可与测量仪器连接，另一端可与非标准接口射频线缆连接。但矢量网络分析仪获取的 S 参数不仅为待测线缆 S 参数，也包括转接引线 S 参数，将转接引线的参数移除得到待测线缆 S 参数的过程称为去嵌入，即为将测试参考面从测量面转移到校准面的过程。

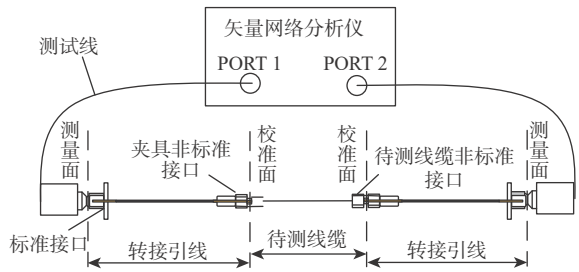


图 1 非标准接口射频线缆测试示意图

Fig. 1 Diagram of non-standard interface RF cable test

非标准接口射频线缆测试的级联模型如图 2 所示。左转接引线、待测线缆、右转接引线的 S 参数分别用 $\boldsymbol{S}_{\mathrm{L}}$、$\boldsymbol{S}_{\mathrm{D}}$、$\boldsymbol{S}_{\mathrm{R}}$ 来表示，矢量网络分析仪测得的 S 参数用 $\boldsymbol{S}_{\mathrm{M}}$ 表示，其中，左右转接引线 $\boldsymbol{S}_{\mathrm{L}}$ 和 $\boldsymbol{S}_{\mathrm{R}}$ 的矩阵表达形式为

$$\begin{cases} \boldsymbol{S}_{\mathrm{L}} = \begin{bmatrix} \boldsymbol{S}_{11\mathrm{L}} & \boldsymbol{S}_{12\mathrm{L}} \\ \boldsymbol{S}_{21\mathrm{L}} & \boldsymbol{S}_{22\mathrm{L}} \end{bmatrix} \\ \boldsymbol{S}_{\mathrm{R}} = \begin{bmatrix} \boldsymbol{S}_{11\mathrm{R}} & \boldsymbol{S}_{12\mathrm{R}} \\ \boldsymbol{S}_{21\mathrm{R}} & \boldsymbol{S}_{22\mathrm{R}} \end{bmatrix} \end{cases} \tag{1}$$

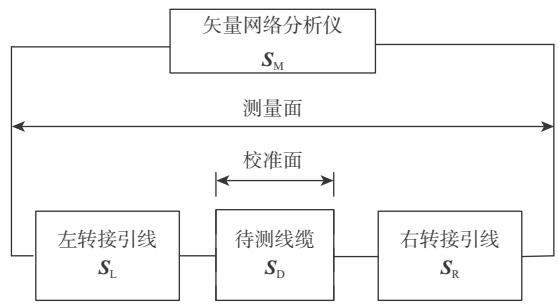


图 2 非标准接口射频线缆测试级联模型

Fig. 2 Cascade model of non-standard interface RF cable test

S 参数虽可以表现微波网络特征，但不适用于级联运算。因此，使用传输参数(T 参数)来进行运算，再进行去嵌入处理，将 T 参数转换为 S 参数。从测量面获取的 T 参数 $\boldsymbol{T}_{\mathrm{M}}$ 如下：

$$\boldsymbol{T}_{\mathrm{M}} = \boldsymbol{T}_{\mathrm{L}}\boldsymbol{T}_{\mathrm{D}}\boldsymbol{T}_{\mathrm{R}} \tag{2}$$

式中：$\boldsymbol{T}_{\mathrm{L}}$、$\boldsymbol{T}_{\mathrm{R}}$ 分别为左、右转接转接引线的 T 参数；$\boldsymbol{T}_{\mathrm{D}}$ 为待测线缆的 T 参数。

转接引线和待测线缆均为二端口元件，其传输矩阵均为二阶矩阵，传输矩阵 $\boldsymbol{T}_{\mathrm{L}}$ 和 $\boldsymbol{T}_{\mathrm{R}}$ 如下：

$$\begin{cases} \boldsymbol{T}_{\mathrm{L}} = r_{22}\begin{bmatrix} a & b \\ c & 1 \end{bmatrix} \\ \boldsymbol{T}_{\mathrm{R}} = \rho_{22}\begin{bmatrix} \alpha & \beta \\ \gamma_1 & 1 \end{bmatrix} \end{cases} \tag{3}$$

通过矩阵运算, 得到待测线缆传输参数矩阵 $\boldsymbol{T}_{\mathrm{D}}$ 如下:

$$\begin{cases}\boldsymbol{T}_{\mathrm{D}}=\boldsymbol{T}_{\mathrm{L}}^{-1}\boldsymbol{T}_{\mathrm{M}}\boldsymbol{T}_{\mathrm{R}}^{-1}=Q\begin{bmatrix}1 & -b\\ -c & a\end{bmatrix}\boldsymbol{T}_{\mathrm{M}}\begin{bmatrix}1 & -\beta\\ -\gamma_1 & \alpha\end{bmatrix}\\ Q=\dfrac{1}{r_{22}\rho_{22}}\cdot\dfrac{1}{a\alpha}\cdot\dfrac{1}{1-b\dfrac{c}{a}}\cdot\dfrac{1}{1-\gamma_1\dfrac{\beta}{\alpha}}\end{cases}\tag{4}$$

1.2　传统 TRL 校准方法

TRL 校准通过设置直通、反射和延迟 3 种校准件, 结合矩阵运算来获取式 (4) 中的未知参量实现去嵌入。3 种校准件的连接方式如图 3 所示。图中: a_1、b_1 为左转接引线信号流图中的左端节点, a_2、b_2 为右转接引线信号流图中的右端节点。

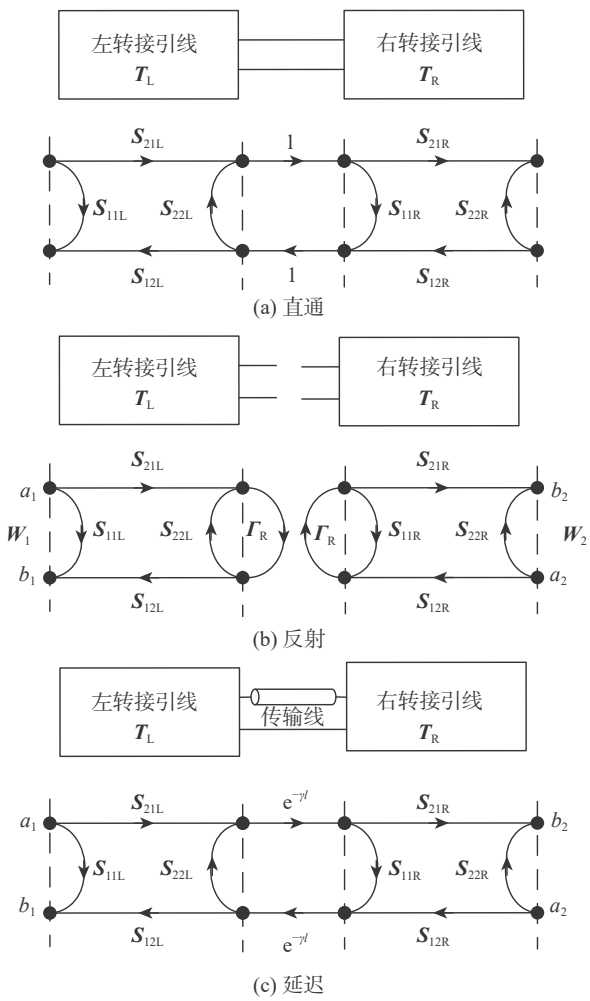


图 3　3 种校准件的连接框图及对应的信号流图

Fig. 3　Connection diagram of three calibration kits and corresponding signal flow diagram

直通校准件的标准为两端转接引线直接相连, 传输系数为 1; 反射校准件的标准为两端转接引线末端状态为短路或断路, 反射系数最好为 1, 使电磁波在转接引线末端发生全反射; 延迟校准件的标准为在两转接引线之间加入一段长度已知、特性参数与转接引线一致的传输线,该传输线为无反射的传输线。

在测量面获取的直通校准件、延迟校准件 T 参数分别为 $\boldsymbol{T}_{\mathrm{MT}}$、$\boldsymbol{T}_{\mathrm{ML}}$, 其表达式如下

$$\boldsymbol{T}_{\mathrm{MT}}=\boldsymbol{T}_{\mathrm{L}}\boldsymbol{T}_{\mathrm{R}}\tag{5}$$

$$\boldsymbol{T}_{\mathrm{ML}}=\boldsymbol{T}_{\mathrm{L}}\boldsymbol{T}_{\mathrm{line}}\boldsymbol{T}_{\mathrm{R}}\tag{6}$$

式中: $\boldsymbol{T}_{\mathrm{line}}$ 为在左右转接引线中间增加的传输线 T 参数, 具体表示为

$$\boldsymbol{T}_{\mathrm{line}}=\begin{bmatrix}\exp(-\gamma l) & 0\\ 0 & \exp(\gamma l)\end{bmatrix}\tag{7}$$

其中: l 为增加传输线的长度; γ 为传播常数。

反射校准件 2 个端口的反射系数 $\boldsymbol{W}_1$ 和 $\boldsymbol{W}_2$ 如下:

$$\begin{cases}\boldsymbol{W}_1=\dfrac{a\boldsymbol{\Gamma}_{\mathrm{R}}+b}{c\boldsymbol{\Gamma}_{\mathrm{R}}+1}\\ \boldsymbol{W}_2=-\dfrac{\alpha\boldsymbol{\Gamma}_{\mathrm{R}}-\gamma}{\beta\boldsymbol{\Gamma}_{\mathrm{R}}-1}\end{cases}\tag{8}$$

式中: $\boldsymbol{\Gamma}_{\mathrm{R}}$为末端反射信号分量。

利用直通校准件和延迟校准件构造矩阵 $\boldsymbol{T}_{\mathrm{TL}}$, 如下:

$$\boldsymbol{T}_{\mathrm{TL}}=\boldsymbol{T}_{\mathrm{ML}}\boldsymbol{T}_{\mathrm{MT}}^{-1}=\begin{bmatrix}t_{11} & t_{12}\\ t_{21} & t_{22}\end{bmatrix}\tag{9}$$

结合式 (5) 构造如下矩阵等式:

$$\boldsymbol{T}_{\mathrm{TL}}\boldsymbol{T}_{\mathrm{L}}=\boldsymbol{T}_{\mathrm{L}}\boldsymbol{T}_{\mathrm{line}}\tag{10}$$

该矩阵等式的方程组形式为

$$\begin{cases}t_{11}a+t_{12}c=a\mathrm{e}^{-\gamma l}\\ t_{21}a+t_{22}c=c\mathrm{e}^{-\gamma l}\\ t_{11}b+t_{12}=b\mathrm{e}^{\gamma l}\\ t_{21}b+t_{22}=\mathrm{e}^{\gamma l}\end{cases}\tag{11}$$

利用式 (11) 的解对式 (5)、式 (8) 进一步求解, 即可求解出全部未知参量, 完成去嵌入。

2　改进的 TRL 校准方法

文献 [12] 提出设计 TRL 校准件时, 需要满足以下要求:

1) 延迟校准件与直通校准件的传输损耗相位差必须在 20°～160°之内, 以避免相位模糊发生。

2) 当测试频率最高值与最低值之比大于 8 时, 需要多个延迟校准件以满足要求 1)。

3) 转接引线特征阻抗应与测量仪器端口阻抗一致。

为更好地满足 TRL 校准件设计的需求, 本文对 TRL 校准进行了改进, 改进的 TRL 校准步骤如图 4 所示, 主要为以下 3 个步骤:

步骤 1　判断延迟校准件和直通校准件的传输损耗相位差是否满足要求, 若不满足需使用虚拟传

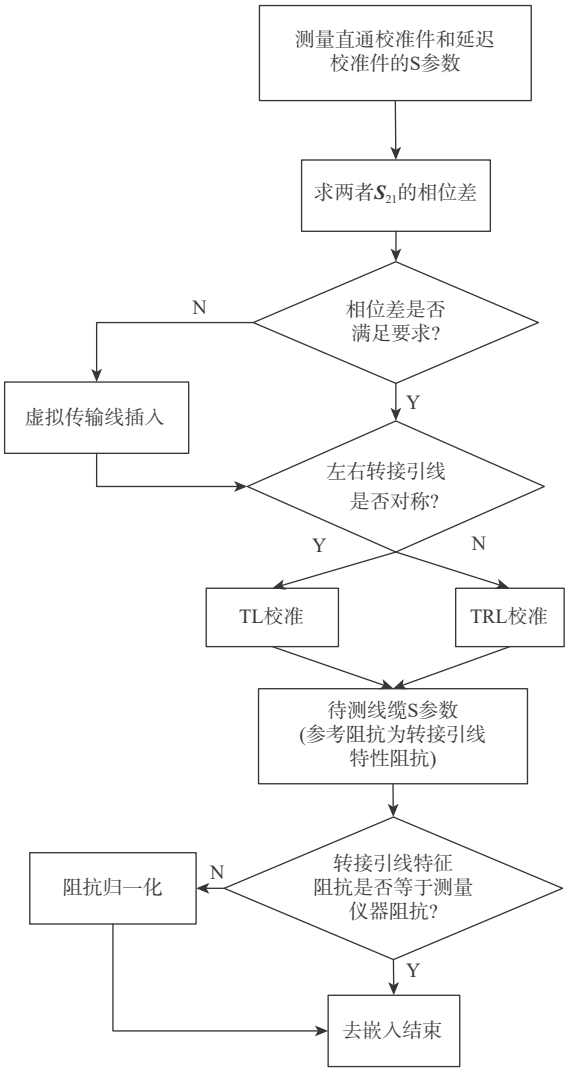


图4 改进的TRL校准流程

Fig. 4 Flow diagram of improved TRL calibration

输线插入进行修正。

步骤2 根据左右转接引线是否对称选择TL校准或TRL校准。

步骤3 去嵌入后的待测线缆S参数是以转接引线特征阻抗为参考阻抗的计算结果。为满足TRL校准件设计要求3),应判断转接引线特征阻抗与测量仪器端口阻抗是否相等,若不满足需进行阻抗归一化处理,得到以测量仪器端口阻抗为参考阻抗的待测线缆S参数。

2.1 虚拟传输线插入对延迟校准件的优化

本文采用虚拟传输线插入的方法,可虚拟构造任意长度的传输线,解决测试频率范围过宽或测试频率过高、过低时延迟校准件相位不达标的问题。虚拟传输线插入的计算方法如下:

$$\boldsymbol{T}_{\mathrm{Lnew}}=\begin{bmatrix}\exp\left[-\gamma(l_0/l_1)\right] & 0\\ 0 & \exp\left[\gamma(l_0/l_1)\right]\end{bmatrix} \tag{12}$$

式中:l_0为满足相位要求的延迟线长度;l_1为实际可提供的延迟线长度。通过在T参数矩阵中进行比例运算,可获得长度为l_0的延迟线T参数矩阵$\boldsymbol{T}_{\mathrm{Lnew}}$。在仿真软件中分别对直通校准件和延迟校准件进行建模,计算两者相位差,通过虚拟传输线插入来获取满足相位要求的延迟校准件。

设置2段相同长度的同轴线作为左右转接引线,测量频率为100～500 MHz。将左右转接引线直接相连作为直通校准件,左右转接引线经一段与转接引线特性相同、长度为5 cm的同轴线连接作为延迟校准件。由测量频率及同轴线缆介质的相对介电常数计算可得,需长度为17 cm的传输线才可满足延迟校准件与直通校准件传输损耗相位差在20°～160°范围内的要求。应用虚拟传输线插入修正延迟校准件,观察其相位差是否达到要求以验证虚拟传输线插入方法的可行性。

延迟校准件经虚拟传输线插入修正前后与直通校准件传输损耗的相位差如图5所示。

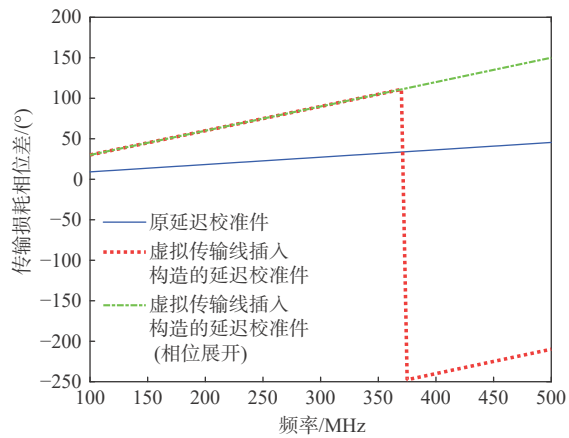


图5 虚拟传输线插入前后延迟校准件与直通校准件传输损耗相位差

Fig. 5 Transmission loss phase-difference of line standard and through standard before and after virtual transmission line insertion

图5中,蓝色实线为延迟校准件未经虚拟传输线插入与直通校准件传输损耗的相位差,观察可得其值在5°～50°范围内,未满足要求。由于在相位计算过程中将相位折叠到了$(-\pi,\pi)$的区间内,造成了相位跳跃的情况,如红色虚线所示的经虚拟传输线插入修正后的延迟校准件与直通校准件的传输损耗相位差在375 MHz处发生相位跳跃。经相位展开后,相位差在30°～150°范围内,满足对延迟校准件的要求,避免了相位模糊的发生。

当测试终止频率与起始频率之比大于8时,如测试频率范围为0.1～1 GHz,需制作2个延迟校准件,此时可应用虚拟传输线插入方法对另一根传输线进行构造,减少延迟校准件的制作数量。

测试频率范围宽时, 虚拟传输线插入前后延迟校准件与直通校准件传输损耗相位差如图 6 所示。图中: 蓝线是延迟线长度为 237 mm 的延迟校准件与直通校准件的传输损耗相位差, 当频率高于 400 MHz 时, 不满足要求。利用虚拟传输线插入技术构造延迟线长度为 75 mm 的延迟校准件, 其与直通校准件的传输损耗相位差如图 6 中的绿线所示。为满足宽频带测量中关于延迟校准件和直通校准件的传输损耗相位差要求, 设置频率分界为 300 MHz, 可观察到在频率分界 300 MHz 前延迟校准件满足相位要求。在频率分界 300 MHz 后, 构造得到的延迟校准件与直通校准件传输损耗相位差在 35°～140°之间, 满足相位要求。

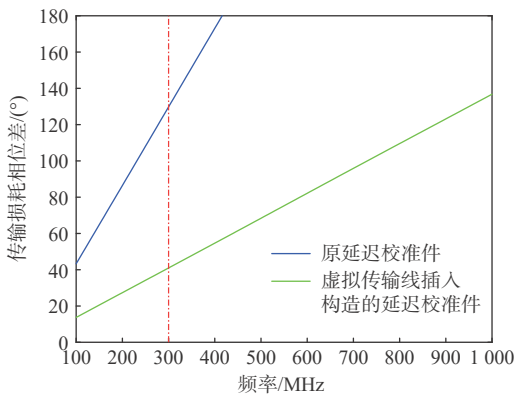


图 6　测试频率范围宽时虚拟传输线插入前后延迟校准件与直通校准件传输损耗相位差

Fig. 6　Transmission loss phase-difference of line standard and through standard before and after virtual transmission line insertion under wide range of test frequencies

2.2　阻抗归一化在 TRL 校准中的应用

为满足 TRL 校准件中转接引线特征阻抗应与测量仪器端口阻抗一致的要求, 若转接引线特征阻抗和测量仪器端口阻抗相差较大时, 应采用阻抗归一化处理。

由图 4 可知, 应先判断转接引线阻抗与测量仪器端口阻抗是否一致, 可通过直通校准件回波损耗的驻波比进行判断。当驻波比的值小于 1.2 时, 可认为转接引线阻抗和测量仪器端口阻抗一致性较好; 当驻波比等于 1 时, 可认为两者阻抗相等。

进行阻抗归一化的前提是转接引线特征阻抗和测量仪器端口阻抗已知, 测量仪器端口阻抗可查阅所使用的矢量网络分析仪资料得知, 转接引线特征阻抗可利用回波损耗计算, 如下:

$$\boldsymbol{Z}=\frac{1+|\boldsymbol{S}_{11\mathrm{T}}|}{1-|\boldsymbol{S}_{11\mathrm{T}}|}\boldsymbol{Z}_0 \tag{13}$$

式中: $\boldsymbol{Z}$ 为转接引线特征阻抗; $\boldsymbol{Z}_0$ 为测量仪器端口阻抗; $\boldsymbol{S}_{11\mathrm{T}}$ 为直通校准件回波损耗。

当转接引线特征阻抗和测量仪器端口阻抗已知后, 可按照式 (14) 进行阻抗归一化:

$$\boldsymbol{T}_{\mathrm{new}}=\frac{1}{4\boldsymbol{Z}\boldsymbol{Z}_0}\begin{bmatrix}\boldsymbol{Z}_0+\boldsymbol{Z} & \boldsymbol{Z}-\boldsymbol{Z}_0\\ \boldsymbol{Z}-\boldsymbol{Z}_0 & \boldsymbol{Z}_0+\boldsymbol{Z}\end{bmatrix}\boldsymbol{T}_{\mathrm{old}}\begin{bmatrix}\boldsymbol{Z}_0+\boldsymbol{Z} & \boldsymbol{Z}_0-\boldsymbol{Z}\\ \boldsymbol{Z}_0-\boldsymbol{Z} & \boldsymbol{Z}_0+\boldsymbol{Z}\end{bmatrix} \tag{14}$$

式中: $\boldsymbol{T}_{\mathrm{old}}$ 为未阻抗归一化的待测线缆 T 参数矩阵; $\boldsymbol{T}_{\mathrm{new}}$ 为以测量仪器端口阻抗为参考且经阻抗归一化后的待测线缆 T 参数矩阵。

2.3　TL 校准去嵌入

实际测试中, 反射系数为 1 的反射校准件难以制作, 且反射校准件中反射系数的相位难以满足要求, 非标准反射校准件会给 TRL 校准造成较大误差。

左右转接引线对称时, 非标准接口射频线缆测试级联模型如图 7 所示。

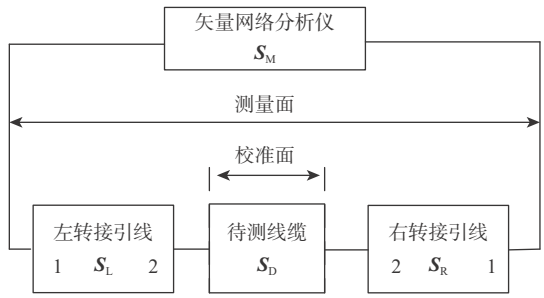


图 7　转接引线对称时非标准接口线缆测试级联模型

Fig. 7　Cascade model of non-standard interface cable test under symmetrical test leads

根据左右转接引线的对称性, 左右转接引线 T 参数关系为

$$\boldsymbol{T}_{\mathrm{R}}=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\boldsymbol{T}_{\mathrm{L}}\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} \tag{15}$$

由式 (15) 可知, 右端转接引线参数可通过左端转接引线参数计算得到, 减少了式 (4) 中未知量的个数, 此时反射校准件提供的数据是冗余的, 因此可以不使用反射校准件。

TL 校准简化了 TRL 校准步骤, 消除了不理想反射校准件引入的误差, 但仅适用于转接引线对称的情况, 实际中转接引线的非对称性可能会对去嵌入精度造成一定影响。

3　数值仿真验证

3.1　仿真条件及仿真对象

基于 ADS 仿真软件构建左右转接引线及待测线缆的电磁仿真模型, 仿真模型的参数参照机载射频线缆进行设置, 仿真模型如图 8 所示。

采用同轴线缆作为转接引线, 左右转接引线为对称结构, 采用长度 10 cm 的同轴线作为待测线缆。在 ADS 中建立直通、反射与延迟校准件的仿

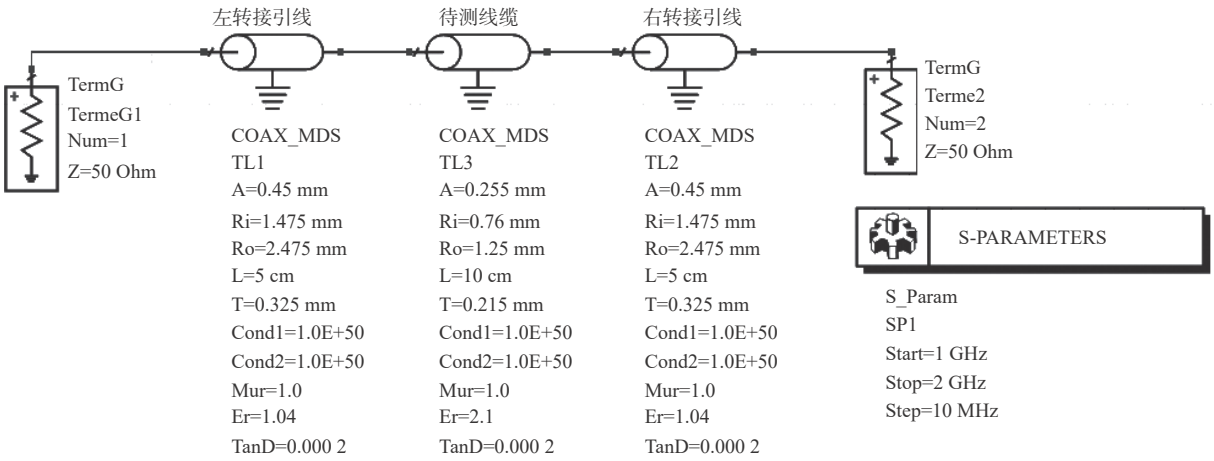


图 8 仿真模型

Fig. 8 Simulation model

真模型。直通校准件将左右转接引线直接连接，反射校准件将左右转接引线经大电阻接地，形成理想开路状态。由测量频率 1～2 GHz 计算可得满足测试需求的延迟校准件中延迟线长度为 3.29 cm，为验证虚拟传输线插入的有效性，将延迟线长度设置为 6 cm。

3.2 方法验证

应用改进 TRL 校准对待测线缆进行测试，将测试结果与参考值对比，验证改进 TRL 校准的有效性。

虚拟传输线插入前后延迟校准件与直通校准件传输损耗相位差如图 9 所示。蓝线为延迟线长度 6 cm 的延迟校准件与直通校准件传输损耗的相位差，观察可得，当频率超过 1.6 GHz 时，相位差未在 20°～160°的范围。应用虚拟传输线插入构造新的延迟校准件，其与直通校准件传输损耗的相位差在 20°～160°内，满足对延迟校准件的要求。

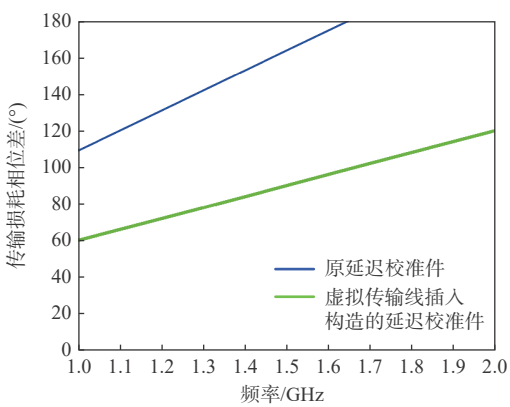


图 9 虚拟传输线插入前后延迟校准件与直通校准件传输损耗相位差仿真结果

Fig. 9 Simulation results of transmission loss phase-difference of line standard and through standard before and after virtual transmission line insertion

直通校准件回波损耗的驻波比如图 10 所示。直通校准件回波损耗的驻波比均在 1.07 之下且不等于 1，说明阻抗匹配程度较好，但转接引线特征阻抗不等于测量仪器端口阻抗。

采用改进 TRL 校准对待测线缆进行去嵌入测试，由于待测线缆为对称的二端口元件，$\boldsymbol{S}_{11}=\boldsymbol{S}_{22}$，$\boldsymbol{S}_{21}=\boldsymbol{S}_{12}$，在图 11 中仅展示 $\boldsymbol{S}_{11}$ 和 $\boldsymbol{S}_{21}$ 的幅值和相位。

去嵌入前后结果如图 11 所示。红色虚线是以转接引线特征阻抗作为参考阻抗得到的去嵌入结果，绿色虚线是以测量仪器端口阻抗为参考阻抗归一化后的结果，黑色实线是未去除转接引线参数的结果，蓝色实线是作为参考值的待测线缆参数。

观察图 11 可得，去嵌入后的结果和参考值很接近，由于转接引线特征阻抗不等于测量仪器端口阻抗，经归一化处理后去嵌入精度进一步提高。

TRL 校准与 TL 校准去嵌入结果如图 12 所示。由于左右转接引线是对称的，TRL 校准与 TL 校准的去嵌入结果相同。

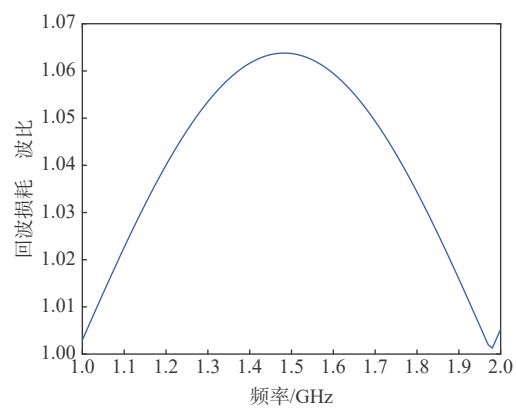


图 10 直通校准件回波损耗驻波比的仿真结果

Fig. 10 Simulation results of voltage standing wave ratio of through standard

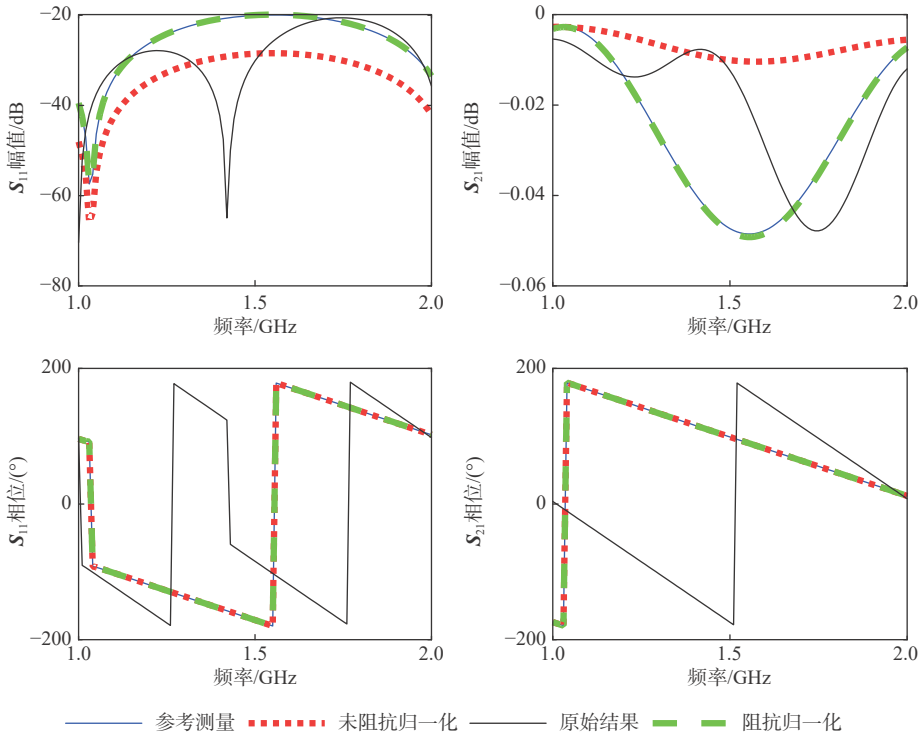


图 11　去嵌入前后结果对比

Fig. 11　Comparison of results before and after de-embedding

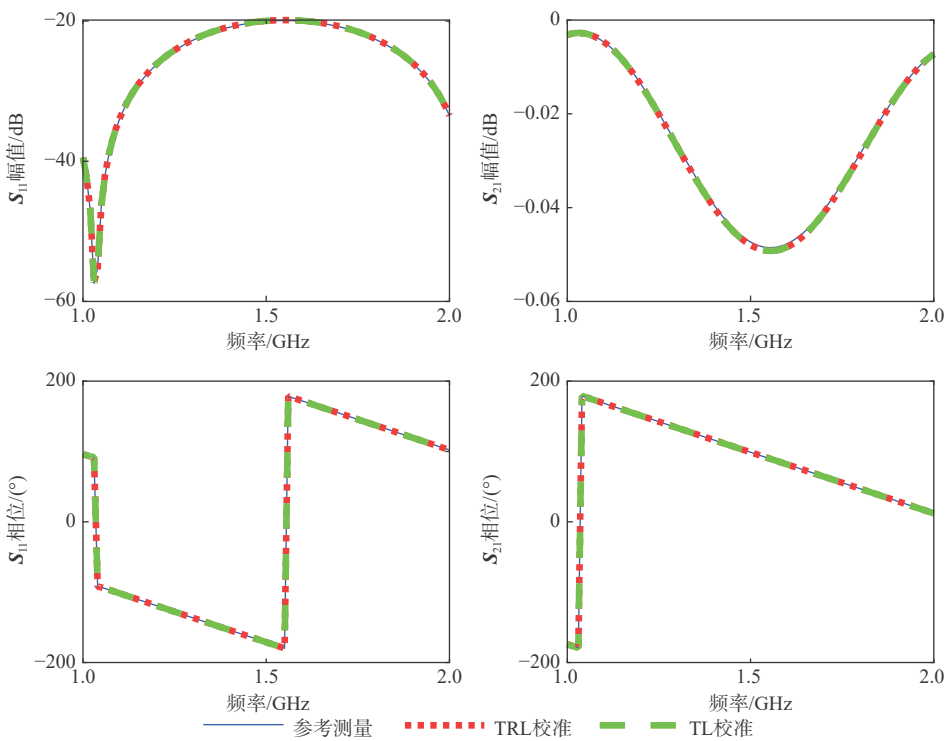


图 12　TRL 校准与 TL 校准去嵌入结果对比

Fig. 12　Comparison of TRL and TL calibrations de-embedding

4　实验验证

为验证改进 TRL 校准的有效性, 构建了由矢量网络分析仪、测试线缆及 4 种测量件组成的实验平台, 如图 13 和图 14 所示。

测试线缆为高精度同轴线缆, 起到端口延伸的作用, 其连接器端口类型为 SMA 内螺内孔型。为方便获取待测线缆的 S 参数作为参考值, 采用两端连接器为 SMA 外螺内针型的射频线缆 RG316 作为待测线缆, 待测线缆的接口可直接与测试线缆相连。左右转接引线采用长度为 5 cm 的同轴线缆 RG58, 其连接器一端为可与测试线缆连接的 SMA 外螺内

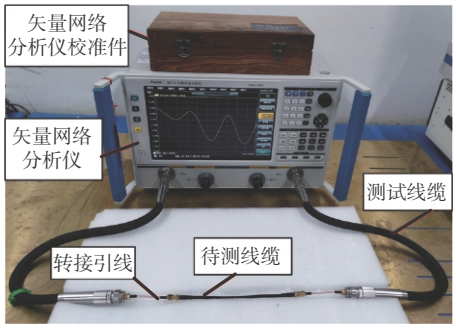


图 13 实验平台

Fig. 13 Experimental platform

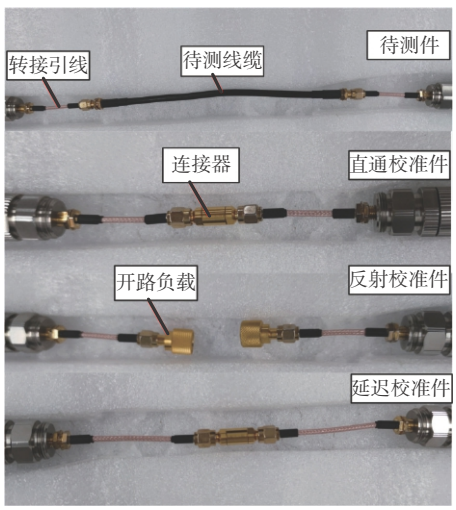


图 14 4 种测量件

Fig. 14 Four measurement components

针型，另一端为可与待测线缆连接的 SMA 内螺内孔型。

4 种测量件为直通校准件、延迟校准件、反射校准件和待测件。直通校准件由左右转接引线经高性能双 SMA 外螺内针型连接器连接构成；延迟校准件由 2 根长度为 7 cm 的 RG58 同轴线缆经高性能双 SMA 外螺内针型连接器连接构成；反射校准件由端接开路负载的左右转接引线构成；待测件由待测线缆和左右转接引线连接构成。

计算延迟校准件与直通校准件传输损耗相位差，其结果如图 15 所示。相位差满足要求，不需要进行虚拟传输线插入处理。

直通校准件回波损耗的驻波比如图 16 所示。驻波比最高为 1.24，说明转接引线特征阻抗和测量仪器端口阻抗差距较大，需进行阻抗归一化处理。

利用式 (13) 对转接引线特征阻抗进行计算，其结果如图 17 所示。传输线分布参数是随频率变化而改变的，因此其特征阻抗也随之改变。实际测试中的传输线为有损传输线，且存在一定非均匀性，因此所得结果存在一定误差。

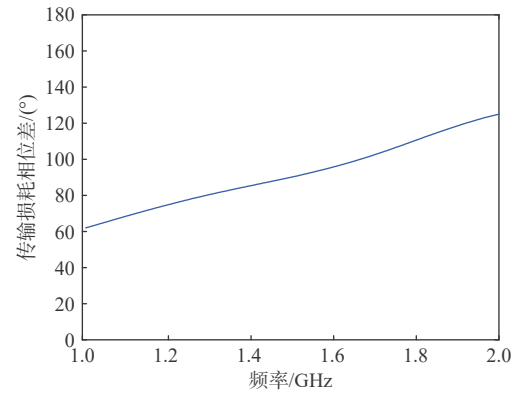


图 15 延迟校准件与直通校准件传输损耗相位差

Fig. 15 Transmission loss phase-difference of line standard and through standard

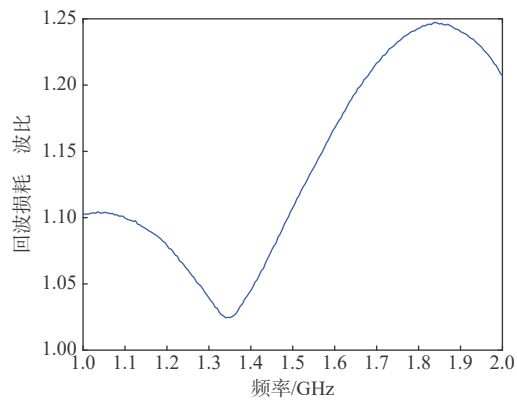


图 16 直通校准件回波损耗驻波比的实验结果

Fig. 16 Experimental result of voltage standing wave ratio of through standard

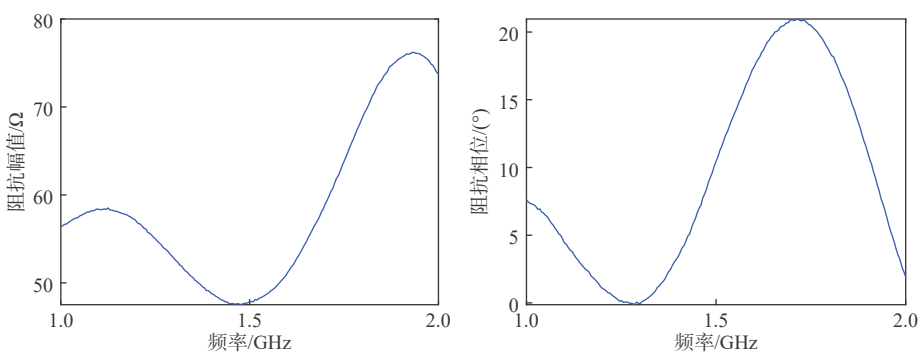


图 17 转接引线特征阻抗

Fig. 17 Characteristic impedance of test lead

应用 TRL 校准去嵌入的结果如图 18 所示。可观察到去嵌入结果和参考测量的差距很大, 非理想反射校准件会给测试结果带来误差, 当反射校准件的性能不理想时, TRL 校准将会失效。

应用 TL 校准去嵌入的结果如图 19 所示。红色虚线为未阻抗归一化的去嵌入结果, 绿色虚线为以测量仪器端口阻抗为参考阻抗归一化后的结果。

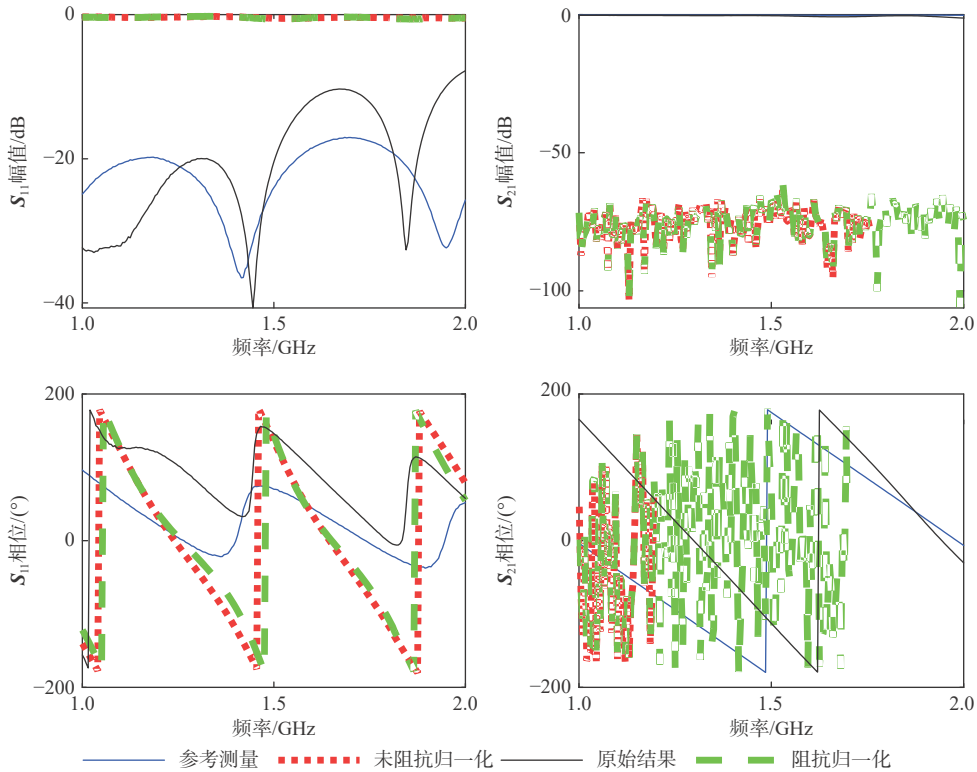


图 18　TRL 校准去嵌入结果

Fig. 18　Results of TRL calibration de-embedding

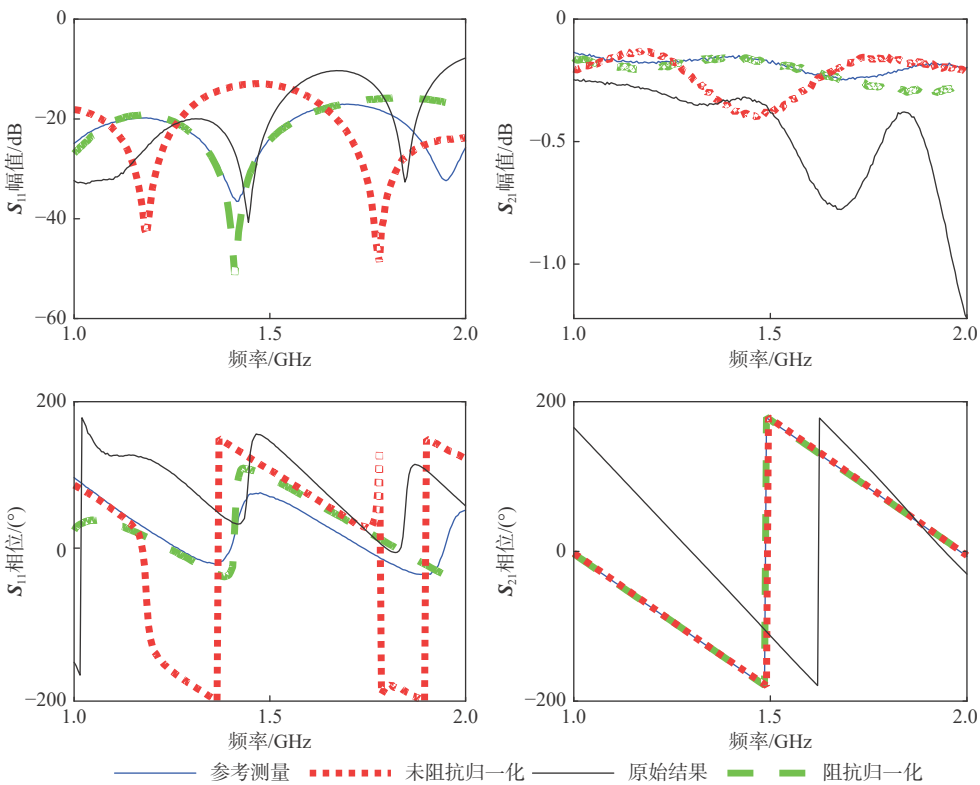


图 19　TL 校准去嵌入结果

Fig. 19　Results of TL calibration de-embedding

对比 S_{21} 中绿色虚线和蓝色实线可知，幅值的最大误差为 0.2 dB，相位几乎完全重合。S_{11} 的幅值除个别谐振点外，最大误差为 1 dB，相位结果中绿色虚线和蓝色实线趋势基本一致，误差较小。造成误差的原因可能有：①左右转接引线具有一定非对称性；②转接引线不是理想均匀传输线，在计算其特征阻抗时存在误差，阻抗归一化计算中进一步引入一定误差；③延迟校准件中延迟线不是无反射的理想传输线。

5 结 论

1）将改进的 TRL 校准应用于非标准接口机载射频线缆测试，通过对比测试结果和参考值发现，除个别谐振点外，传输损耗幅值误差不超过 0.2 dB，回波损耗幅值误差不超过 1 dB，验证了本文方法的有效性。

2）对传统 TRL 校准的不足进行了相应改进。应用虚拟传输线插入技术可在测试频率过高或过低时降低延迟校准件的制作精度，当测试频率范围较宽时，减少延迟校准件的制作数量；当转接引线特征阻抗和测量仪器端口阻抗不一致时，应用阻抗归一化可提高测试精度；转接引线对称时，可应用 TL 校准算法进行测试，不再需要制作反射校准件，消除了非理想反射校准件引入的误差。

参考文献（References）

[1] 邹维明, 刘蕾, 姜睿智, 等. 非标准接口射频电缆性能测试细节控制浅谈[J]. 计测技术, 2017, 37(S1): 269-271.
ZOU W M, LIU L, JIANG R Z, et al. Talking about the detailed control of the performance test of the non-standard interface RF cable[J]. Measurement Technology, 2017, 37(S1): 269-271(in Chinese).

[2] 高维胜. 浅谈宇航级电连接器射频性能测试技术[J]. 机电元件, 2018, 38(2): 38-43.
GAO W S. Talking about the testing technology of RF performance of aerospace electrical connector[J]. Electromechanical Components, 2018, 38(2): 38-43(in Chinese).

[3] 中华人民共和国国家质量监督检验检疫总局, 中国国家标准化管理委员会. 射频射频电缆组件第1部分: 总规范 一般要求和试验方法: GB/T 17738.1—2013[S]. 北京: 中国标准出版社, 2014: 5-9.
General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Radio frequency and coaxial cable assemblies. Part 1: Generic specification-General requirements and test methods: GB/T 17738.1—2013[S]. Beijing: Standards Press of China, 2014: 5-9(in Chinese).

[4] IEC. Radio frequency and coaxial cable assemblies. Part 1: Generic specification-General requirements and test methods: IEC 60966-1[S]. Geneva: IEC, 2019.

[5] BIANCO B, PARODI M, RIDELLA S, et al. Launcher and microstrip characterization[J]. IEEE Transactions on Instrumentation and Measurement, 1976, IM-25(4): 320-323.

[6] FRANZEN N R, SPECIALE R A . A new procedure for system calibration and error removal in automated S-parameter measurements[C]//Proceedings of the 5th European Microwave Conference. Piscataway: IEEE Press, 1975.

[7] ENGEN G F, HOER C A. Thru-reflect-line: An improved technique for calibrating the dual six-port automatic network analyzer[J]. IEEE Transactions on Microwave Theory and Techniques, 1979, 27(12): 987-993.

[8] 李新伟, 易磊. 非标器件S参数去嵌入测试方法研究[J]. 计量与测试技术, 2014, 41(6): 30-32.
LI X W, YI L. Research on the testing method of S device parameters of non-standard[J]. Metrology & Measurement Technique, 2014, 41(6): 30-32(in Chinese).

[9] 彭安虎, 韩周安, 李志强. T/R阵列自动测试系统去嵌入技术研究[J]. 中国测试, 2021, 47(7): 92-98.
PENG A H, HAN Z A, LI Z Q. Research on de-embedding technology of T/R array automatic test system[J]. China Measurement & Testing Technology, 2021, 47(7): 92-98(in Chinese).

[10] 吴宇, 高源慈, 钟催林, 等. 微波射频差分探针去嵌入理论研究[J]. 电子测量技术, 2020, 43(4): 16-22.
WU Y, GAO Y C, ZHONG C L, et al. Microwave RF differential probe de-embedding theory research[J]. Electronic Measurement Technology, 2020, 43(4): 16-22(in Chinese).

[11] 孟文营, 穆卓, 王身云. 一种PIN二极管参数的提取方法与应用[J]. 电子元件与材料, 2019, 38(4): 32-36.
MENG W Y, MU Z, WANG S Y. Parameter extraction of PIN diode and its applications[J]. Electronic Components and Materials, 2019, 38(4): 32-36(in Chinese).

[12] 刘迪. 怎样设计和验证TRL校准件及具体过程[J]. 电子产品世界, 2008(3): 123-126.
LIU D. How to design and verify TRL Cal Kit also implementation of TRL calibration[J]. Electronic Engineering & Product World, 2008(3): 123-126(in Chinese).

[13] 廖进昆, 刘仁厚. LRL校准法及其在微波测量中的应用[J]. 电子科技大学学报, 2000, 29(2): 149-152.
LIAO J K, LIU R H. LRL calibration method and its application to microwave measurement[J]. Journal of University of Electronic Science and Technology of China, 2000, 29(2): 149-152(in Chinese).

[14] ZUIGA-JUAREZ J E, REYNOSO-HERNANDEZ J A, MAYA-SANCHEZ M C. An improved multiline TRL method[C]//Proceedings of the 67th ARFTG Microwave Measurements Conference. Piscataway: IEEE Press, 2008: 139-142.

[15] MARKS R B. A multiline method of network analyzer calibration[J]. IEEE Transactions on Microwave Theory and Techniques, 1991, 39(7): 1205-1215.

[16] WANG M, ZHAO Y J, JIN Y M, et al. Sensitivity analysis of multiport S-parameter due to non-ideal TRL calibration standards[J]. Radio Science, 2017, 52(9): 1096-1105.

[17] STEER M B, GOLDBERG S B, RINNE G, et al. Introducing the through-line deembedding procedure[C]//Proceedings of the International Microwave Symposium Digest. Piscataway: IEEE Press, 1992: 1455-1458.

[18] 易波, 王为, 刘培国, 等. 基于改进TRL校准算法的二极管参数测量[J]. 中国舰船研究, 2015, 10(2): 121-124.

YI B, WANG W, LIU P G, et al. Measurement of PIN diode parameters based on TRL calibration algorithm[J]. Chinese Journal of Ship Research, 2015, 10(2): 121-124(in Chinese).

[19] 梁子健. 基于非均匀传输线重构的滤波器设计与电缆测试技术研究[D]. 广州: 华南理工大学, 2017.
LIANG Z J. Design of filters and cable test technology based on reconstruction of non-uniform transmission lines[D]. Guangzhou: South China University of Technology, 2017(in Chinese) .

Non-standard interface aviation RF cable test method based on TRL

SHI Xudong，LI Ruipu，ZHAO Hongxu*，ZHANG Haotian

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

Abstract: Aircraft radio frequency (RF) cables are the main medium for transmitting signals. There are currently various non-standard interface cables, and in order to use standard interface instruments for performance testing, precise adapter leads are required. However, the precise adapter leads have high requirements on parameters such as impedance and transmission loss, and precise adapter cables are prone to performance degradation, resulting in large errors in test results. Therefore, this paper proposes a non-standard interface aircraft RF cable test method based on TRL for the problems that exist when applying precision adapter leads. Firstly, the cascade transmission parameter method is used to model the test leads and the cable under test, and then the improved TRL calibration method is used to de-embed the test leads to obtain the scattering parameters of the tested cable, so as to evaluate the performance of the aircraft RF cable. The technique was used to test aircraft RF cables, and the findings show that, with the exception of a few resonance spots, the maximum error in return loss is approximately 1 dB and the maximum error in transmission loss is approximately 0.2 dB. This validates the method's viability and efficacy.

Keywords: de-embedding; TRL calibration; transmission line theory; S-parameters; cascade

Received：2021-10-29；**Accepted**：2022-01-14；**Published Online**：2022-02-24 15:53

URL：kns.cnki.net/kcms/detail/11.2625.V.20220222.1850.002.html

Foundation item：National Natural Science Foundation of China (51377161)

* **Corresponding author**. E-mail：zhx2581@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0664

新冠疫情背景下航空物流网络的鲁棒优化

张锦[1, 2, 3, *]，张哲睿[1]，洪治潮[1]，杨文广[1]，闫妍[4]

(1. 西南交通大学 交通运输与物流学院，成都 610031；2. 西南交通大学 综合交通大数据应用技术国家工程实验室，成都 610031；
3. 西南交通大学 综合交通运输智能化国家地方联合工程实验室，成都 610031；4. 西安财经大学 管理学院，西安 710100)

摘 要：为应对新冠疫情影响背景下客机航线中断导致客机腹舱运力减少的风险，以最小的成本提升航空物流网络的鲁棒性，提出不确定客机腹舱运力情境下全货机航线配机及路径优化问题研究。以网络总成本最小化为目标，考虑“客机腹舱+全货机”的运输方式、城市间的航空货运需求、运输距离等因素，建立了客机腹舱运力不确定环境下航空物流网络鲁棒优化模型，并采用列与约束生成（C&CG）算法对模型进行有效求解。通过S货运航空公司案例数据，研究航空物流网络航线优化布局问题。通过案例分析证明，运用两阶段鲁棒优化方法，可有效降低航空物流网络受客机腹舱运力减少导致的成本变动；在客机航线完全中断的状况下，所提方案可以降低航空物流网络20.7%的成本。运用两阶段鲁棒优化方法有利于决策者灵活进行全货机配机选线，在客机腹舱运力不确定影响背景下兼顾航空物流网络的经济性和鲁棒性。

关 键 词：航空货运；航线路径优化；不确定性；鲁棒优化；列与约束生成算法
中图分类号：V221^{+}.3; TB553
文献标志码：A **文章编号**：1001-5965(2023)09-2218-09

现阶段，中国航空货物运输的运载方式主要为腹舱载货，腹舱货运运输量约占总运输量的70%。中国航空网络存在货运枢纽和客运枢纽高度重叠的特征，航空物流网络依赖于客运网络特性显著。然而，由于航空货物运输和航空旅客运输有着不同的特性，且客机航班规划往往满足于客运需求而非货运需求，现阶段航空货物运输依然受限于航空客运网络。特别在新冠疫情背景下，航空物流网络供需不匹配性进一步体现。根据中国民用航空局统计，2020年2月，新冠疫情的强传染性导致全国范围内客运航班大量取消，客运航班执行比例仅为30%左右，同时导致航空物流网络中各航段上客机腹舱运力被大幅削减。2020年6月后，全国各航段上的客机航班基本恢复正常水平，然而成都、北京、西安、杭州等地由于新冠疫情的反复，导致局部地区的航空客运网络发生中断，也使国内航空客运网络始终处于不确定状态下。由于新冠疫情对于客机腹舱运力不确定影响导致航空物流网络呈现非鲁棒性，航空货运企业急需识别现有航空物流网络中存在的关键节点和薄弱环节，在未来航空物流网络规划过程中考虑采用合理的风险态度配置全货机航线以降低运营风险，构建相对鲁棒的航空物流网络。

目前，国内外对于航空网络规划的研究主要集中于航空网络结构[1]、航空网络枢纽点选址[2-3]、航空航线优化[4]等方面。对于航空航线优化的研究，可根据其参数是否固定分为确定型航线网络优化与不确定型航线网络优化。在确定型航线网络优化方面，多采用定性及定量分析方法，建立评价指标体系或数学模型，通过决策树法、Frank-Wolfe算法[5]、遗传算法、禁忌搜索算法[6]等求解，确定最优航线分布。在不确定型航线网络优化方面，相关学者采用了模糊规划、随机规划、鲁棒优化等一系列方

收稿日期：2021-11-05；录用日期：2022-02-19；网络出版时间：2022-03-18 13:22
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220316.1342.002.html
*通信作者. E-mail：zhjswjtu@home.swjtu.edu.cn

引用格式：张锦，张哲睿，洪治潮，等. 新冠疫情背景下航空物流网络的鲁棒优化[J]. 北京航空航天大学学报，2023，49(9)：2218-2226.
ZHANG J，ZHANG Z R，HONG Z C，et al. Robust optimization of aviation logistics network in context of COVID-19 pandamic[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2218-2226 (in Chinese).

法。鲁棒优化常用于考虑不确定条件下, 包括设施失效及需求不确定等情况, 尤其适用于缺乏历史数据及难以进行数据预测的情况, 近年来两阶段鲁棒优化方法[7-8]也能有效弥补单阶段鲁棒优化方法过于保守的问题。对于鲁棒优化方法, 现有研究多采用 Benders 分解法[9]、列生成算法[10]、列与约束生成(column-and-constraint generation, C&CG)算法[11]等精确算法及禁忌搜索算法[12]、遗传算法等启发式算法求解。现有研究中, 姜涛[13]从航空企业的角度, 研究航线网络的鲁棒优化问题, 考虑航空运输需求和成本的不确定性, 建立各种“轴-辐”式航线网络的鲁棒优化模型; 张冲[14]考虑航空运输环境的不确定性, 构建了需求和成本不确定的联盟航线网络鲁棒优化模型。现有研究在航线路径优化问题的不确定参数方面考虑了成本、需求、服务时间等, 但在研究航空物流网络时考虑货机和客机腹舱共用现状较少, 且较少考虑航段上客机减少对航空物流网络产生的影响。

基于此, 本文充分考虑了航空物流网络“客机腹舱+全货机”的实际构建模式, 以及新冠疫情下客机腹舱运力变动存在的不确定性, 构建了客机腹舱运力不确定环境下航空物流网络鲁棒优化模型。采用两阶段鲁棒优化方法解决新冠疫情不确定性问题, 可在第二阶段的配流决策中获取新冠疫情导致航线中断对第一阶段的设计决策产生最大影响的场景, 从而不断对第一阶段设计决策进行迭代验证调整, 避免单阶段鲁棒优化方法获取方案过度保守的问题。同时, 决策者可通过不确定集考虑新冠疫情的不同阶段差异, 采取不同的风险态度选取适配的规划方案。

1　问题描述

在有向图$G=(N,\Omega)$上建模, 其中, N为节点集, Ω为网络中的弧集。基于多个航空货运基地O的位置, 进行航线路径规划。实际中, 航空公司选取客机进行货运作业时, 通常选取直达航班, 可理解为由客机腹舱构成“点-点”式客机腹舱货运网络。任一货机具备从航空货运基地, 经过多个节点经停, 最终返回出发的航空货运基地, 形成闭合环状航线的特点。因此, 建模流程可描述为: 将部分机场确定为航空货运基地, 限制飞机只可由航空货运基地出发并最终飞回同一航空货运基地。同时, 考虑新冠疫情背景下进行网络鲁棒优化, 即新冠疫情发生情景存在不确定性, 客机腹舱运力对航线受影响的程度不同, 航线数量存在差异, 且不同航线发生中断后对航空物流网络的影响程度也存在差异。由于该不确定性, 管理者规划网络时需采取不同风险态度制定合理的货机航线布局方案及货物配流方案。具体不确定客机腹舱运力情境下货机航线网络优化问题如图 1 所示。

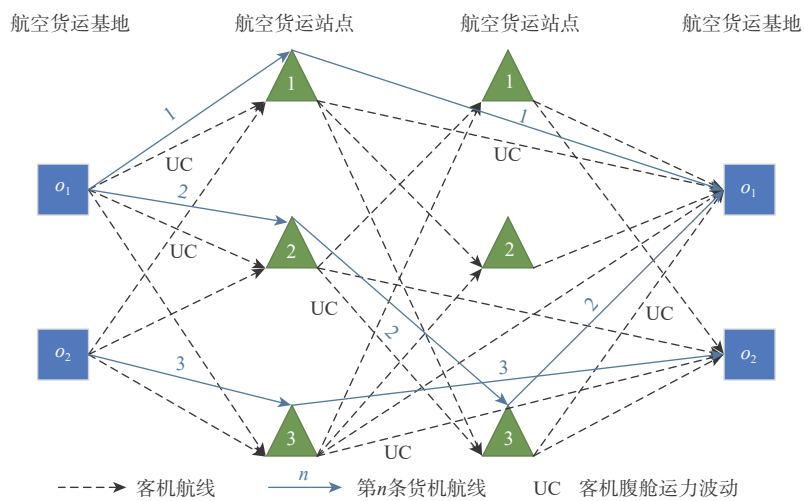


图 1　不确定客机腹舱运力下货机航线路径优化

Fig. 1　Route optimization of cargo aircraft under uncertainty for capacity of passenger aircraft bellies

2　模型假设及建立

2.1　模型假设

1) 运输过程中, 各航段的全货机单位运输成本相同, 全货机的单位运输成本与飞机的类型相关, 假设所用的全货机类型一致。

2) 运输过程中, 客机腹舱的单位租赁成本相同。

3) 货物可以在混合网络中的任何航段进行中转运输。

4) 考虑全货机在每一机场的装卸、等待时间及全货机航权约束等问题, 一般每天全货机可飞行 3 个航段。

5) 航空物流网络正常运行时除少数不支持腹舱运输的航段外,其他航段上都有腹舱运输。

2.2 不确定集构建

采用多面体不确定集刻画客机腹舱运力的不确定性,与盒子不确定集相比,多面体不确定集不会过于保守[15]。由于需要在考虑客机腹舱货运能力减少情形下构建鲁棒的航空物流网络,在此情形中,正向偏差值对系统成本几乎不产生影响,小于客机腹舱运力期望值时才会产生额外的成本。因此,本文研究的不确定客机腹舱运力范围为 $[\overline{q}_{ij}-\hat{q}_{ij},\overline{q}_{ij}]$,即构建的多面体不确定集为

$$q_{ij}\in\left[\overline{q}_{ij}-\hat{q}_{ij},\overline{q}_{ij}\right]\qquad \overline{q}_{ij}>0,\hat{q}_{ij}>0 \tag{1}$$

$$U=\left\{\tilde{q}_{ij}\middle|\tilde{q}_{ij}=\overline{q}_{ij}-\hat{q}_{ij}z_{ij},\forall(ij)\in\Omega\right\} \tag{2}$$

$$Z=\left\{z_{ij}\middle|\sum_{(ij)\in\Omega}z_{ij}\leqslant\Gamma,0\leqslant z_{ij}\leqslant 1,\forall(ij)\in\Omega\right\} \tag{3}$$

式中:q_{ij}为航段 $[i,j]$ 上的最大客机腹舱货运容量;U和Z为不确定参数集合;z_{ij}表示偏差程度;Ω为航段集合,$\Omega\{q_{ij}(ij)|i,j\in N\}$;$\tilde{q}_{ij}$为航段$[i,j]$上不确定客机腹舱运力参数;$\hat{q}_{ij}$为航段$[i,j]$上新冠疫情发生后客机腹舱运力最大偏离量;$\overline{q}_{ij}$为航段$[i,j]$上正常状态下客机腹舱运力均值;$\Gamma$为不确定预算水平,表示决策者对客机腹舱运力不确定性的风险态度。Γ在实证问题中的现实意义可以近似理解为:决策者在规划网络过程中预判可能发生客机腹舱运力大幅下降的航段数量,故Γ值的定义区间由网络中航段集合大小决定($0\leqslant\Gamma\leqslant|\Omega|$);若$\Gamma=0$,即所有航段上腹舱运力均取对应航段未受新冠疫情影响时的均值量,模型等价于确定性模型,此时不确定预算水平最低,决策者对风险态度无所谓;若$\Gamma=|\Omega|$,对应任意航段间客机腹舱运力变化均取最大偏差,此时不确定预算水平最高,即决策者风险态度高度重视。

2.3 模型建立

采用鲁棒优化模型,考虑“客机腹舱+全货机”的运输方式、城市间的航空货运需求、运输距离等因素,解决新冠疫情背景下腹舱运力不确定的航空物流网络优化问题。

1) 目标函数。

$$\min\left\{\sum_{l=1}^{L}\sum_{i=1,j=1}^{N}C_{\mathrm{P}}X_{[i,j]}^{l}+\sum_{i=1,j=1}^{N}\sum_{l=1}^{L}C_{\mathrm{F}}d_{ij}X_{[i,j]}^{l}+C_{o}B\right\}+\max_{z}\min_{Q_{ij}^{\mathrm{F}},Q_{ij}^{\mathrm{K}},S_{ij}}\left\{\sum_{i=1,j=1}^{N}C_{\mathrm{K}}d_{ij}Q_{ij}^{\mathrm{K}}+\sum_{i=1,j=1}^{N}C_{\mathrm{S}}S_{ij}\right\} \tag{4}$$

式中:O为航空货运基地集合,$o\in O$;N为节点集合,$i,j\in N$;L为货机集合,$l\in L$;C_{P}为货机在任一航段的固定飞行成本;C_o为航空货运基地o的建设成本;C_{F}为单架货机的单位运输成本;C_{K}为客机腹舱单位租赁成本;B为航空基地数量;d_{ij}为点i到点j的距离;Q_{ij}^{F}为点i到点j上的货机货运量;Q_{ij}^{K}为点i到点j上的客机腹舱货运量;C_{S}为未运输货物的单位损失成本;S_{ij}为点i到点j上未运输的需求;$X_{[i,j]}^{l}$为第l架货机是否经过$[i,j]$航段,货机l经过$[i,j]$航段则取1,反之则取0。式(4)指总体航空物流网络的系统总成本最小,其中,第1部分为飞机在航段上运行的固定成本,第2部分为每一航段的货机运输成本(此处考虑无论货机满载与否,承担货机全航段运输费用),第3部分为航空货运基地的固定成本,第4部分为客机腹舱租赁成本,第5部分为无人承运部分的损失惩罚成本。

2) 约束条件。

$$\sum_{j=1}^{N}X_{[i,j]}^{l}=x_{i}^{l} \tag{5}$$

$$\sum_{i=1}^{N}\sum_{j=1}^{N}X_{[i,j]}^{l}\leqslant M\sum_{o=1}^{O}x_{o}^{l} \tag{6}$$

$$\sum_{i=1}^{N}X_{[i,s]}^{l}-\sum_{j=1}^{N}X_{[s,j]}^{l}=0\qquad\forall s\in N \tag{7}$$

$$\sum_{i=1,j=1}^{N}X_{[i,j]}^{l}\leqslant 3 \tag{8}$$

$$\left(\sum_{l=1}^{L}X_{[i,j]}^{l}-1\right)Q_{\mathrm{F}}\leqslant Q_{ij}^{\mathrm{F}}\leqslant\sum_{l=1}^{L}X_{[i,j]}^{l}Q_{\mathrm{F}} \tag{9}$$

$$Q_{ij}^{\mathrm{K}}\leqslant\overline{q}_{ij}-\hat{q}_{ij}z_{ij} \tag{10}$$

$$Q_{ij}^{\mathrm{F}}+Q_{ij}^{\mathrm{K}}+S_{ij}=W_{ij} \tag{11}$$

$$\sum_{(ij)\in\Omega}z_{ij}\leqslant\Gamma \tag{12}$$

$$0\leqslant z_{ij}\leqslant 1,\forall(ij)\in\Omega \tag{13}$$

$$X_{[i,j]}^{l}\in\{0,1\},x_{i}^{l}\in\{0,1\} \tag{14}$$

$$Q_{ij}^{\mathrm{K}},Q_{ij}^{\mathrm{F}},S_{ij}\geqslant 0 \tag{15}$$

式中:W_{ij}为点i到点j的需求量;Q_{F}为货机最大运载容量;x_{i}^{l}为第l架货机是否由i点起飞,货机l由i点起飞则取1,反之则取0;M为足够大的常数。

式(5)~式(7)为货机起终点限制,表示货机必须从同一航空货运基地起飞,且最终飞回起始基

地; 式(7)为流平衡约束, 即所有需求点服务的货机必从此需求点离开; 式(8)为货机每天最大飞行航段约束; 式(9)为ij段所有货机的运量约束; 式(10)为客机腹舱运量约束; 式(11)为航段运输需求量为航段内货机运量与客机腹舱运量及未运输量的总和; 式(12)、式(13)表示不确定集; 式(14)、式(15)为变量相关约束。

3　模型求解

3.1　模型等价转化

由于模型目标函数中包含“max-min”结构形式, 无法直接求解。基于模型双层结构特点, 将原问题拆分为主问题与子问题。其中, 拆分主问题为

$$\min\left\{\sum_{l=1}^{L}\sum_{i=1,j=1}^{N}C_{\mathrm{P}}X_{[i,j]}^{l}+\sum_{i=1,j=1}^{N}\sum_{l=1}^{L}C_{\mathrm{F}}d_{ij}X_{[i,j]}^{l}+C_{o}B+\theta\right\}$$

$$\text{s.t.}\begin{cases}\displaystyle\sum_{j=1}^{N}X_{[i,j]}^{l}=x_{i}^{l}\\ \displaystyle\sum_{i=1}^{N}\sum_{j=1}^{N}X_{[i,j]}^{l}\leqslant M\sum_{o=1}^{O}x_{o}^{l}\\ \displaystyle\sum_{i=1}^{N}X_{[i,s]}^{l}-\sum_{j=1}^{N}X_{[s,j]}^{l}=0\qquad\forall s\in N\\ \displaystyle\sum_{i=1,j=1}^{N}X_{[i,j]}^{l}\leqslant 3\\ X_{[i,j]}^{l}\in\{0,1\},x_{i}^{l}\in\{0,1\}\end{cases}\tag{16}$$

式中: θ为子问题最优目标函数值。

子问题为

$$\max_{z}\min_{Q_{ij}^{\mathrm{F}},Q_{ij}^{\mathrm{K}},S_{ij}}\left\{\sum_{i=1,j=1}^{N}C_{\mathrm{K}}d_{ij}Q_{ij}^{\mathrm{K}}+\sum_{i=1,j=1}^{N}C_{\mathrm{S}}S_{ij}\right\}$$

$$\text{s.t.}\begin{cases}\left(\displaystyle\sum_{l=1}^{L}X_{[i,j]}^{l}-1\right)Q_{\mathrm{F}}\leqslant Q_{ij}^{\mathrm{F}}\leqslant\sum_{l=1}^{L}X_{[i,j]}^{l}Q_{\mathrm{F}}\\ Q_{ij}^{\mathrm{K}}\leqslant\overline{q}_{ij}-\hat{q}_{ij}z_{ij}\\ Q_{ij}^{\mathrm{F}}+Q_{ij}^{\mathrm{K}}+S_{ij}=W_{ij}\\ \displaystyle\sum_{(ij)\in\Omega}z_{ij}\leqslant\Gamma\\ 0\leqslant z_{ij}\leqslant 1,\forall(ij)\in\Omega\\ Q_{ij}^{\mathrm{K}},Q_{ij}^{\mathrm{F}},S_{ij}\geqslant 0\end{cases}\tag{17}$$

子问题根据对偶理论, 以最优单纯形乘子α、β、ρ、λ分别对应式(17)中第 1 ～ 3 个约束, 转化为其对偶问题, 即 DRO:

$$\max_{\alpha,\beta,\lambda,\rho}\left\{-\sum_{i=1,j=1}^{N}\sum_{l=1}^{L}X_{[i,j]}^{l}Q_{\mathrm{F}}\alpha_{ij}+\left(\sum_{l=1}^{L}X_{[i,j]}^{l}-1\right)Q_{\mathrm{F}}\beta_{ij}-\sum_{i=1,j=1}^{N}(\overline{q}_{ij}-\hat{q}_{ij}z_{ij})\rho_{ij}+\sum_{i=1,j=1}^{N}W_{ij}\lambda_{ij}\right\}$$

$$\text{s.t.}\begin{cases}-\alpha_{ij}+\beta_{ij}+\lambda_{ij}\leqslant 0\\ -\rho_{ij}+\lambda_{ij}\leqslant C_{\mathrm{K}}d_{ij}\\ \lambda_{ij}\leqslant C_{\mathrm{S}}\\ \alpha_{ij},\beta_{ij},\rho_{ij}\geqslant 0\end{cases}\tag{18}$$

由于对偶问题模型 DRO 内存在线性决策变量的乘积, 是 NP-hard 问题, 难以直接求解。Gabrel 等[16]提出并证明了将此类问题转化为混合整数规划表达方式的方法。即若$\Gamma\in\mathbf{Z}^{+}$, 则 DRO 存在最优解$(\alpha^{*},\beta^{*},\lambda^{*},\rho^{*},z^{*})$且$z^{*}\in\{0,1\}$。引入连续辅助变量$\omega_{ij}=z_{ij}\rho_{ij}\geqslant 0$及足够大的常数$M$线性化目标函数中线性决策变量的乘积, 则可得以下推论模型 DRO-MIP:

$$\mathrm{DR}(\overline{\overline{x}})=\max_{\alpha,\beta,\lambda,\rho}\left\{-\sum_{i=1,j=1}^{N}\sum_{l=1}^{L}X_{[i,j]}^{l}Q_{\mathrm{F}}\alpha_{ij}+\left(\sum_{l=1}^{L}X_{[i,j]}^{l}-1\right)\cdot Q_{\mathrm{F}}\beta_{ij}-\sum_{i=1,j=1}^{N}(\overline{q}_{ij}\rho_{ij}-\hat{q}_{ij}\omega_{ij})+\sum_{i=1,j=1}^{N}W_{ij}\lambda_{ij}\right\}$$

$$\text{s.t.}\begin{cases}-\alpha_{ij}+\beta_{ij}+\lambda_{ij}\leqslant 0\\ -\rho_{ij}+\lambda_{ij}\leqslant C_{\mathrm{K}}d_{ij}\\ \lambda_{ij}\leqslant C_{\mathrm{S}}\\ \alpha_{ij},\beta_{ij},\rho_{ij}\geqslant 0\\ z_{ij}\in\{0,1\},\forall(ij)\in\Omega\\ \omega_{ij}\leqslant\rho_{ij},\forall(ij)\in\Omega\\ \omega_{ij}+(1-z_{ij})M\geqslant\rho_{ij},\forall(ij)\in\Omega\\ \omega_{ij}\leqslant Mz_{ij},\forall(ij)\in\Omega\\ \omega_{ij}\geqslant 0,\forall(ij)\in\Omega\end{cases}\tag{19}$$

3.2　求解算法流程

Zeng 和 Zhao[17]提出采用 C&CG 算法能够有效求解不确定集为多面体集时的两阶段鲁棒优化问题。C&CG 算法的计算流程主要为在第m次迭代过程中, 基于主问题求得最优解, 代入主问题最优解求解子问题得到最坏情景$\overline{Z}^{m+1}$, 将最坏情景对应约束作为紧约束加到主问题当中, 直至上下界收敛, 即可获取鲁棒解及鲁棒解在各中断情形下的最优配流解。

因此, 结合模型形成 C&CG 算法框架如下:

1) 令 UB = $+\infty$, LB = $-\infty$, $m=0$, σ为足够小的正数。

2) 求解如下主问题:

$$\min\left\{\sum_{l=1}^{L}\sum_{i=1,j=1}^{N}C_{\mathrm{P}}X_{[i,j]}^{l}+\sum_{i=1,j=1}^{N}\sum_{l=1}^{L}C_{\mathrm{F}}d_{ij}X_{[i,j]}^{l}+C_{o}B+\theta\right\}$$

$$\text{s.t.}\begin{cases}\displaystyle\sum_{j=1}^{N}X_{[i,j]}^{l}=x_{i}^{l}\\ \displaystyle\sum_{i=1}^{N}\sum_{j=1}^{N}X_{[i,j]}^{l}\leqslant M\sum_{o=1}^{O}x_{o}^{l}\\ \displaystyle\sum_{i=1}^{N}X_{[i,s]}^{l}-\sum_{j=1}^{N}X_{[s,j]}^{l}=0\qquad\forall s\in N\\ \displaystyle\sum_{i=1,j=1}^{N}X_{[i,j]}^{l}\leqslant 3\\ \theta\geqslant\displaystyle\sum_{i=1,j=1}^{N}C_{\mathrm{K}}d_{ij}Q_{ij}^{\mathrm{K}r}+\sum_{i=1,j=1}^{N}C_{\mathrm{S}}S_{ij}^{r}\\ \qquad\forall r=1,\cdots,m\\ \left(\displaystyle\sum_{l=1}^{L}X_{[i,j]}^{l}-1\right)Q_{\mathrm{F}}\leqslant Q_{ij}^{\mathrm{F}r}\leqslant\sum_{l=1}^{L}X_{[i,j]}^{l}Q_{\mathrm{F}}\\ \qquad\forall r=1,\cdots,m\\ Q_{ij}^{\mathrm{K}r}\leqslant\bar{q}_{ij}-\hat{q}_{ij}z_{ij}^{r}\qquad\forall r=1,\cdots,m\\ Q_{ij}^{\mathrm{F}r}+Q_{ij}^{\mathrm{K}r}+S_{ij}^{r}=W_{ij}^{r}\qquad\forall r=1,\cdots,m\\ X_{[i,j]}^{l}\in\{0,1\},x_{i}^{l}\in\{0,1\}\\ Q_{ij}^{\mathrm{K}r},Q_{ij}^{\mathrm{F}r},S_{ij}^{r}\geqslant 0\qquad\forall r=1,\cdots,m\end{cases}\tag{20}$$

求得主问题最优解为：$\left(x_{i}^{l(m+1)},X_{[i,j]}^{(m+1)l},\theta^{m+1},Q_{ij}^{\mathrm{K}1},\cdots,Q_{ij}^{\mathrm{K}m};Q_{ij}^{\mathrm{F}1},\cdots,Q_{ij}^{\mathrm{F}m};S_{ij}^{1},\cdots,S_{ij}^{m}\right)^{*}$，同时更新主问题的最优目标函数值至下界：$\mathrm{LB}=\left\{\sum_{l=1}^{L}\sum_{i=1,j=1}^{N}C_{\mathrm{P}}X_{[i,j]}^{(m+1)l}+\sum_{i=1,j=1}^{N}\sum_{l=1}^{L}C_{\mathrm{F}}d_{ij}X_{[i,j]}^{(m+1)l}+\theta^{m+1}\right\}$。

3）将$\left(x_{i}^{l(m+1)},X_{[i,j]}^{(m+1)l}\right)^{*}$代入对应的相关子问题DRO-MIP，以得到最坏情景下的$\bar{Z}^{m+1}$及其目标函数值$\mathrm{opt}\left(Q_{ij}^{\mathrm{F}(m+1)},Q_{ij}^{\mathrm{K}(m+1)},S_{ij}^{m+1}\right)$，同时更新上界UB：$\min\left\{\mathrm{UB},\sum_{l=1}^{L}\sum_{i=1,j=1}^{N}C_{\mathrm{P}}X_{[i,j]}^{(m+1)l}+C_{o}B+\sum_{i=1,j=1}^{N}\sum_{l=1}^{L}C_{\mathrm{F}}d_{ij}X_{[i,j]}^{(m+1)l}+\mathrm{opt}(Q_{ij}^{\mathrm{F}(m+1)},Q_{ij}^{\mathrm{K}(m+1)},S_{ij}^{m+1})\right\}$。

4）如果$\mathrm{gap}=(\mathrm{UB}-\mathrm{LB})/\mathrm{UB}\leqslant\sigma$时，输出$\left(x_{i}^{l(m+1)},X_{[i,j]}^{(m+1)l},\theta^{m+1},Q_{ij}^{\mathrm{K}1},\cdots,Q_{ij}^{\mathrm{K}m};Q_{ij}^{\mathrm{F}1},\cdots,Q_{ij}^{\mathrm{F}m};S_{ij}^{1},\cdots,S_{ij}^{m}\right)^{*}$，并终止算法；否则，更新$m=m+1$，并将创建新变量$\left(Q_{ij}^{\mathrm{K}(m+1)},Q_{ij}^{\mathrm{F}(m+1)},S_{ij}^{m+1}\right)$并将子问题求得的最坏情景$\bar{Z}^{m+1}$对应约束作为一个紧约束加到主问题中，并重新开始步骤2。

其中，所得$\left(x_{i}^{l(m+1)},X_{[i,j]}^{(m+1)l}\right)^{*}$为鲁棒物流网络下的最优全货机航线布局方案，而$\left(Q_{ij}^{\mathrm{K}1},\cdots,Q_{ij}^{\mathrm{K}m};Q_{ij}^{\mathrm{F}1},\cdots,Q_{ij}^{\mathrm{F}m};S_{ij}^{1},\cdots,S_{ij}^{m}\right)$为面对不同情景$\bar{Z}^{1},\cdots,\bar{Z}^{m}$时的最优配流情况。

4 实证分析

4.1 参数设定

选取S货运航空公司服务的10个重要节点城市(北京、广州、上海、深圳、杭州、沈阳、郑州、武汉、成都、西安)作为案例，选取深圳、杭州作为其航空货运基地进行分析。根据货机机队机型，设定最大载重量为Q_{T}=30 t，每段航程所需的固定费用(包含飞机折旧等)为C_{P}=45 000元/架次，航空货运基地建设成本C_{o}=300 000元[18]。根据S货运航空公司实际运营情况，设定货机单位运输成本为C_{F}=1 500元/10^{2}km，客机腹舱租赁单位成本为C_{K}=150元/(t·10^{2}km)，未承运惩罚成本为C_{S}=20 000元/t。运输需求采用S货运航空公司2019年某日的OD数据，其中包含90个有效航段。平均客机腹舱运力取值参照2019年某日客机航班飞行情况，客机腹舱运力偏差值参考2020年3月某日的客机航班情况。采用Python编程实现并通过Gurobi9.1.2进行求解，同时设定$\sigma\leqslant10^{-4}$时终止计算。所有实验在R7-4800H CPU@2 900 MHz，16 GB内存电脑上运行计算。

4.2 算法性能比较分析

为评估C&CG算法在求解两阶段鲁棒优化问题时的求解效果，引入求解鲁棒优化模型比较常见的Benders分解法[9]求解鲁棒优化模型。Benders分解法与C&CG算法的不同之处主要在于：Benders分解法在每次迭代中都会向主问题中添加相关的Benders割，即

$$\theta\geqslant\left\{-\sum_{i=1,j=1}^{N}\sum_{l=1}^{L}X_{[i,j]}^{l}Q_{\mathrm{F}}\alpha_{ij}^{r}+\left(\sum_{l=1}^{L}X_{[i,j]}^{l}-1\right)Q_{\mathrm{F}}\beta_{ij}^{r}-\sum_{i=1,j=1}^{N}(\bar{q}_{ij}-\hat{q}_{ij}z_{ij})\rho_{ij}^{r}+\sum_{i=1,j=1}^{N}W_{ij}\lambda_{ij}^{r}\right\}\tag{21}$$

主要设置Γ取不同值情景下，对比Benders分解法与C&CG算法在3 600 s内求得解的gap值；以及2种算法所获的UB值(即在限制时间内所获得的最优可行解的目标函数值)。2种算法的性能对比结果如表1所示。

由于Benders分解法在求解本案例问题时效率非常低，考虑选取5个重要节点城市(北京、广州、上海、深圳、杭州)作为算例，比较2个算法在$\Gamma=5$情景下的收敛性能。在小算例情景下，C&CG算法在1 s内通过2次迭代即可求解，而Benders分解法需经过160 s通过12次迭代求解。2个算法具体的求解收敛过程如图2和图3所示。由此可知，C&CG算法相比于Benders算法迭代次数较少，求解时间短，C&CG算法的总体性能显著更优。

表 1　算法性能对比结果

Table 1　Algorithm performance comparison result

算法	Γ	迭代次数	gap/%	UB/10^4	时间/s
Benders算法	0	18	40	603	3 600
	5	14	62	690	3 600
	10	14	51	831	3 600
	15	15	49	811	3 600
C&CG算法	0	1	0	541	0.8
	5	24	0	640	840
	10	39	0	664	3 367
	15	17	0	666	1 987

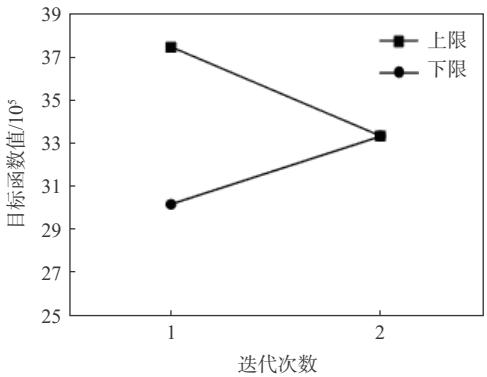


图 2　C&CG 算法迭代次数相关的收敛过程

Fig. 2　Convergence process related to number of iterations of C&CG algorithm

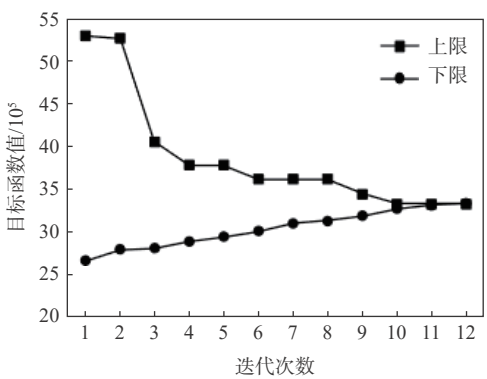


图 3　Benders 分解法迭代次数相关的收敛过程

Fig. 3　Convergence process related to number of iterations of Benders algorithm

4.3　结果分析

分析Γ取[0,90]区间内某一值获得航线网络规划最优决策后, 不同程度腹舱运力减少情景下Γ取不同值时航线网络最优解的系统总成本, 分析结果并评估各决策方案的经济性及鲁棒性效果。为刻画不同程度的腹舱运力减少情景, 引入客机腹舱运力减少系数ζ, 假设$\tilde{q}_{ij}=\bar{q}_{ij}-\zeta\hat{q}_{ij}(0\leqslant\zeta\leqslant1)$, 为便于观察, ζ以 0.1 间隔取值。

4.3.1　不同不确定预算水平下求解方案

案例中$\Gamma=0$时对应于确定性模型, 即对应于客机腹舱运力确定情景下航线网络最优方案, 此时模型仅从经济性角度出发, 所选航段采用货机承运经济性优于客机腹舱租赁; $\Gamma\neq0$时为鲁棒性模型, 对应于客机腹舱运力减少下决策者采取不同决策态度的航线网络鲁棒方案, Γ的取值对应于决策者认为某日客机腹舱运力存在不确定的航段数量, 如$\Gamma=90$时, 即决策者认为所有航段均存在运力偏差, 决策者决策态度完全保守。

不同不确定预算水平Γ下货机航线方案配置结果如表 2 所示, 列举了Γ分别为 0、5、10、15、20 时的货机航线路径方案。结果显示, 决策者采取不同的决策态度进行决策时, 网络配置的航线数量随不确定预算水平的增长而增加。案例中当$\Gamma\geqslant15$时, 货机航线保持不变, 说明在案例条件下仅有部分航段客机腹舱运力变化对决策产生影响, 且模型可自动筛选出这类航段, 降低了对于航段运力减少信息的依赖。

表 2　不同不确定预算水平下求解的航线网络方案

Table 2　Route network solutions under different uncertain budget levels

方案	Γ	货机航线路径
1	0	①3-0-8-3;②3-2-5-3;③4-2-8-4; ④3-7-0-3;⑤3-0-9-3;⑥3-0-7-3;
2	5	①3-9-0-3;②3-0-7-3;③3-0-8-3; ④3-2-5-3;⑤3-6-3;⑥3-0-9-3; ⑦4-2-8-4;⑧4-1-0-4;
3	10	①3-4-3;②3-7-0-3;③4-1-0-4; ④3-9-2-3;⑤3-6-3;⑥3-0-8-3; ⑦4-2-8-4;⑧3-0-9-3;⑨3-0-7-3; ⑩3-2-5-3;
4	15	①4-7-0-4;②3-5-2-3;③3-6-3; ④3-9-0-3;⑤3-7-1-3;⑥3-0-8-3; ⑦3-4-3;⑧4-1-7-4;⑨3-2-5-3; ⑩3-0-9-3;⑪3-0-7-3;⑫4-2-8-4;
5	20	①4-7-0-4;②3-5-2-3;③3-6-3; ④3-9-0-3;⑤3-7-1-3;⑥3-0-8-3; ⑦3-4-3;⑧4-1-7-4;⑨3-2-5-3; ⑩3-0-9-3;⑪3-0-7-3;⑫4-2-8-4;

4.3.2　不确定预算方案在不同客机腹舱运力减少情景中的成本分析

分析客机腹舱运力产生不同程度减少的影响下, 决策者采取不同决策策略后航空物流网络系统总成本变化情况, 图 4～图 6 仅选取有差异、有代表性的结果。图 4 和图 5 分别为ζ取不同值时航线网络总成本的变化。可以发现, 在不同程度客机腹舱运力减少的影响下, 不确定预算水平Γ增加时, 总

成本的变化情况呈非线性变化，主要由于模型会优先选择对子问题目标函数影响最大的场景。

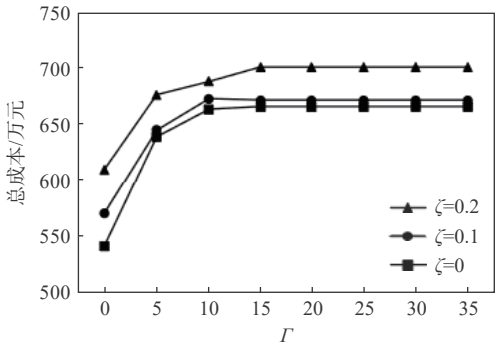


图4 ζ为0、0.1和0.2时各航空物流网络成本比较

Fig. 4 Comparison of cost of each aviation logistics network plan when ζ is 0, 0.1 and 0.2 respectively

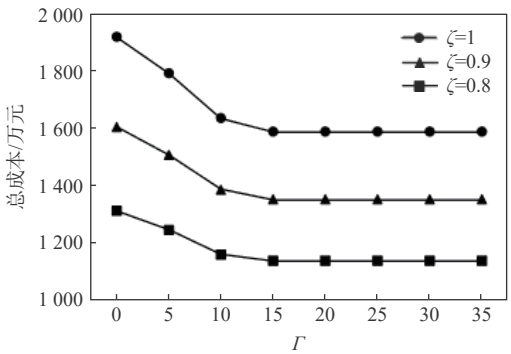


图5 ζ为0.8、0.9和1时各航空物流网络成本比较

Fig. 5 Comparison of cost of each aviation logistics network plan when ζ is 0.8, 0.9 and 1 respectively

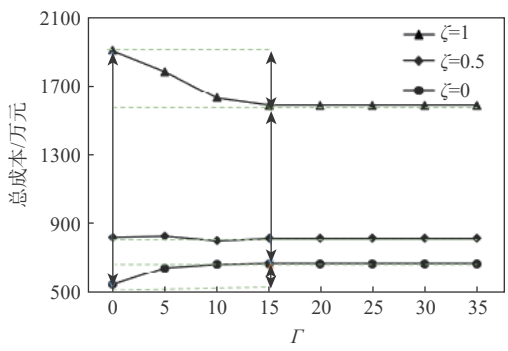


图6 ζ为0、0.5和1时各航空物流网络成本比较

Fig. 6 Comparison of cost of each aviation logistics network plan when ζ is 0, 0.5 and 1 respectively

如图4和图5所示，ζ分别取值0～0.2和0.8～1时，ζ取值越小，航空物流网络的系统总成本越低，即客机腹舱运力资源越充足，网络总成本越小。同时随着不确定预算水平的增加，系统的总成本变化趋势并不相同，这主要与客机腹舱运力成本、货机成本、惩罚成本及各类运输方式的流量分配情况相关。

如图4所示，各航段内客机腹舱运力较充足时（ζ趋近于0时），Γ趋近于0时，航空物流网络规划方案货机布局较少，大多航段采用客机腹舱，在航段货运量较少时采用客机腹舱货运成本低，因此航空物流网络总成本较低。

如图5所示，各航段内客机运力较匮乏时（ζ趋近于1时），Γ趋近于0时，航空物流网络货机配置较少，各航段未满足运输量上升，总成本较高。

如图6所示，案例中ζ取0.5时，各方案的航空物流网络系统总成本差异较小。同时，在不同客机腹舱运力减少情况下（$0 \leqslant \zeta \leqslant 1$），确定性模型（$\Gamma = 0$）取得的最优方案对应总成本标准差为$4.5\times10^6$；当$\Gamma \geqslant 15$时，获得的航空物流网络方案面对各客机腹舱运力减少情况对应总成本标准差为3.0×10^6。因此，$\Gamma \geqslant 15$时所得航空物流网络方案面对各类客机腹舱运力减少总成本变动较$\Gamma = 0$时所获方案小，获得的航空物流网络鲁棒性更强。当客机航线完全中断时（ζ为1时），$\Gamma \geqslant 15$时所得航空物流网络方案较$\Gamma = 0$时所获方案每日可降低成本约330万元，即可降低20.7%的成本。

以上结果表明，在现实运行过程中，决策者需根据腹舱运力网络中断情景发生概率，综合权衡航空物流网络的鲁棒性与经济性选择对应策略。在新冠疫情广泛传播状况下，航空物流网络受新冠疫情影响发生航段中断数量大情景概率大，此时决策者风险态度应高度重视，即使Γ取值较大，所获方案对腹舱运力依赖较低，能有效保障航空物流网络鲁棒性。新冠疫情影响环节后，各航段上同时发生大面积腹舱运力变动的概率低，此时新冠疫情的影响通常表现为少量航段中断，此时决策者风险态度无所谓，即使Γ取值较小，所获方案能有效利用客机腹舱运力降低网络成本，在保障航空物流网络运力需求的同时兼顾了航空物流网络的经济性。

5 结 论

本文主要解决新冠疫情背景下客机腹舱运力减少时航空物流网络配机及路径优化问题。采用多面体不确定集描述客机腹舱运力可能出现的减少情况。针对新冠疫情背景下客机腹舱运力不确定性问题，提出了一种两阶段航线优化方法，并通过C&CG算法有效求解了鲁棒模型。

1）不同客机腹舱运力减少情况对航空物流网络系统成本影响结果不同，决策者可通过两阶段鲁棒优化方法充分考虑客机腹舱运力减少对航空货运的影响程度，从而灵活进行货机选线，以在客机正常运行场景下增加少量网络运营成本为代价，大幅降低客机中断场景造成的损失，从而保障航空物流网络的鲁棒性。

2) 部分航段货机的增减会显著影响航空物流网络的鲁棒性, 采用两阶段鲁棒优化方法可以自动寻找对航空物流网络影响较大的航段, 有效辅助决策者进行航线网络规划决策。

3) 在新冠疫情背景下, 决策者可以根据自身风险态度转变, 兼顾航空物流网络的经济性和鲁棒性。相比于确定性模型, 充分考虑航段间客机数量减少甚至中断状况; 相比于单阶段鲁棒优化方法, 决策者可以灵活调整应对风险态度, 灵活控制成本, 搭建合理的航空物流网络。

4) 通过实证研究验证了客机腹舱运力不确定环境下航空物流网络鲁棒优化模型在国内航空物流网络优化问题中的有效性。同时, 由于国际航空物流网络运输结构特点及腹舱运力不确定变动情况对比国内航空物流网络具有相似性, 模型对于国际航空物流网络鲁棒优化问题也有一定的参考价值, 未来可对国际航空物流网络鲁棒优化问题进行进一步验证研究。

参考文献（References）

[1] DU W, ZHOU X, LORDAN O, et al. Analysis of the Chinese airline network as multi-layer networks[J]. Transportation Research Part E:Logistics and Transportation Review, 2016, 89: 108-116.

[2] O'KELLY M E. A quadratic integer program for the location of interacting hub facilities[J]. European Journal of Operational Research, 1987, 32(3): 393-404.

[3] ZHANG J, YAN Y, TANG Q Y. Hub location of air cargo company in air alliance[J]. Journal of Southeast University (English Edition), 2020, 36(3): 341-348.

[4] YAN Y, ZHANG J, TANG Q Y. Aircraft fleet route optimization based on cost and low carbon emission in aviation line alliance network[J]. Journal of Intelligent & Fuzzy Systems, 2020, 39(1): 1163-1182.

[5] YU S N, YANG Z Z, YU B. Air express network design based on express path choices: Chinese case study[J]. Journal of Air Transport Management, 2017, 61: 73-80.

[6] 高娇蛟. 我国快递企业航空运输网络的优化设计研究[D]. 北京: 北京交通大学, 2011: 40-46.
GAO J J. Research on optimization design of air transport network of express enterprises in China[D]. Beijing: Beijing Jiaotong University, 2011: 40-46(in Chinese).

[7] CHENG C, QI M, ZHANG Y, et al. A two-stage robust approach for the reliable logistics network design problem[J]. Transportation Research Part B:Methodological, 2018, 111: 185-202.

[8] LING A, SUN J, XIU N, et al. Robust two-stage stochastic linear optimization with risk aversion[J]. European Journal of Operational Research, 2017, 256(1): 215-229.

[9] MARTINS D S, MORABITO R, CAMARGO R S. Benders decomposition applied to a robust multiple allocation incomplete hub location problem[J]. Computers & Operations Research, 2018, 89: 31-50.

[10] 梁金鹏, 不确定条件下的城市轨道交通客流控制策略及接驳公交网络优化研究[D]. 北京: 北京交通大学, 2020: 85-96.
LIANG J P. Research on passenger flow control policy and bus bridging service optimization under uncertainty in urban rail transit[D]. Beijing: Beijing Jiaotong University, 2020: 85-96(in Chinese).

[11] WANG Z, QI M. Robust service network design under demand uncertainty[J]. Transportation Science, 2020, 54(3): 676-689.

[12] 陈焕平. 内支线集装箱班轮航线鲁棒优化模型[D]. 大连: 大连海事大学, 2020: 26-30.
CHEN H P. Robust optimization model of container liner routes in feeder line network[D]. Dalian: Maritime Affairs University of Dalian, 2020: 26-30(in Chinese).

[13] 姜涛. 航空公司中枢辐射航线网络鲁棒优化设计问题研究[D]. 南京: 南京航空航天大学, 2007.
JIANG T. Research on the robust optimization design of airline's hub and radial route network[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2007(in Chinese).

[14] 张冲. 航空联盟网络设计与优化研究[D]. 南京: 南京航空航天大学, 2019: 40-49.
ZHANG C. Research on design and optimization of aviation alliance network[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2019: 40-49(in Chinese).

[15] BERTSIMAS D, SIM M. Robust discrete optimization and network flows[J]. Mathematical Programming, 2003, 98(1-3): 49-71.

[16] GABREL V, LACROIX M, MURAT C, et al. Robust location transportation problems under uncertain demands[J]. Discrete Applied Mathematics, 2014, 164: 100-111.

[17] ZENG B, ZHAO L. Solving two-stage robust optimization problems using a column-and-constraint generation method[J]. Operations Research Letters, 2013, 41(5): 457-461.

[18] 闫妍, 马啸来. 航线联营下基于转运的飞机航线路径优化[J]. 北京航空航天大学学报, 2023, 49(1): 115-127.
YAN Y, MA X L. Air freight route planning based on transshipment under air alliance[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(1): 115-127(in Chinese).

Robust optimization of aviation logistics network in context of COVID-19 pandamic

ZHANG Jin[1, 2, 3, *], ZHANG Zherui[1], HONG Zhichao[1], YANG Wenguang[1], YAN Yan[4]

(1. School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 610031, China;
2. National Engineering Laboratory of Integrated Transportation Big Data Application Technology, Southwest Jiaotong University, Chengdu 610031, China;
3. National United Engineering Laboratory of Integrated and Intelligent Transportation, Southwest Jiaotong University, Chengdu 610031, China;
4. School of Management, Xi'an University of Finance and Economics, Xi'an 710100, China)

Abstract: To deal with the risk of reducing the capacity of passenger aircraft bellies due to the interruption of passenger aircraft routes in the context of COVID-19 pandemic, and to improve the robustness of the aviation logistics network at a minimum cost, a study on the allocation and route optimization of all-cargo aircraft routes under the situation of uncertain passenger aircraft belly storage capacity was proposed. With the objective of minimizing the total cost of the route network while considering transportation mode, air cargo demand, and transportation distance between cities, this study established a robust optimization model for the aviation logistics network under the uncertain passenger aircraft bellies. The column-and-constraint generation (C&CG) algorithm was effectively used to solve the model. In addition, this paper investigated the optimized layout of aviation logistics network routes using the case study and data of company S. The results show that the two-stage robust optimization method for aviation logistics network can effectively reduce cost changes caused by reduced passenger aircraft belly capacity, and the robust solution can reduce the cost of the aviation logistics network by 20.7% in the extreme case of complete disruption of passenger routes. Therefore, this study concluded that the use of the two-stage robust optimization method enables decision-makers to flexibly select cargo aircraft routing while considering the economy and robustness of the aviation logistics network under the influence of uncertainty surrounding passenger aircraft belly capacity.

Keywords: air cargo; aircraft route optimization; uncertainty; robust optimization; column-and-constraint generation algorithm

Received: 2021-11-05; **Accepted**: 2022-02-19; **Published Online**: 2022-03-18 13:22
URL: kns.cnki.net/kcms/detail/11.2625.V.20220316.1342.002.html
* **Corresponding author**. E-mail: zhjswjtu@home.swjtu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0668

锂离子电池热失控气体快速检测及危险性分析方法

张青松*，刘添添，郝朝龙，曲奕润，张伟，陈达
(中国民航大学 民航热灾害防控与应急重点实验室，天津 300300)

摘 要：为实现对锂离子电池热失控气体成分及危险性进行快速分析，提出了一种基于激光拉曼光谱气体检测技术结合经验公式的分析方法。利用自主设计搭建的锂离子电池气体检测平台，使用不同荷电状态（75%、100%）和不同正极材料（LCO、NCM）的18650型锂离子电池对热失控气体成分及爆炸危险性分析方法进行验证。结果表明：相同荷电状态（SOC）下，NCM电池比LCO电池在热失控后释放的可燃气体量更多，100%荷电状态的NCM电池在热失控后容器内部CO占比高达22.02%，随着荷电状态的增加，锂离子电池热失控生成的可燃气体量增加。不同实验工况下，锂离子电池热失控后的气体仍具有较高的爆炸风险。研究结果证明了激光拉曼光谱气体检测技术检测锂离子电池热失控气体的可行性，为热失控气体成分和爆炸危险性的快速检测及分析提供了理论依据和技术支撑。

关 键 词：锂离子电池；热失控；拉曼光谱；气体检测；爆炸极限
中图分类号：X934
文献标志码：A **文章编号**：1001-5965(2023)09-2227-07

随着人们环保意识的提高，新能源汽车正在逐步取代传统燃油汽车，而锂离子电池因其具有自放电率低、使用寿命长及比容量高等优势，成为目前主流的汽车动力电源[1]，锂离子航空运输量逐年上升，未来将成为电动飞机动力电源主要选择之一。然而，制约锂离子电池应用范围进一步扩大的主要瓶颈是安全性[2]。近年来，锂离子电池热失控引发的火灾事故逐年上升，国内外研究人员针对锂离子电池热失控开展了大量研究[3]。

锂离子电池热失控诱因主要包括3种滥用情况：热滥用、机械滥用及电滥用[4-6]。当电池长期暴露在这些滥用情景中时，可能会触发电池内部材料自放热反应，从而在短时间内导致温度急剧上升并释放出大量可燃及有毒气体，这些气体具有严重的燃烧爆炸风险，对周围人员及设备产生严重威胁。因此，需要对锂离子电池热失控气体成分及产生机理进行研究。

郭志慧等[7]通过外部加热的方式针对NCM电池开展热失控实验，并通过气相色谱-质谱联用仪(GC-MS)对热失控的气体进行分析，发现CO和C_2H_4的含量与电池荷电状态(state of charge, SOC)呈正相关关系，CO_2与H_2的含量与电池SOC呈负相关关系。张青松等[8]通过实验研究发现，相同SOC下NCM电池热失控后CO生成量随环境压力的降低而升高。Fernandes等[9]在过充测试中使用傅里叶红外变换(Fourier transform infrared, FTIR)光谱对$LiFePO_4$电池释放的气体组分进行分析。Yuan等[10]使用加速量热仪(accelerating rate calorimeter, ARC)对NCM、LPF及LTO等3种正极材料的电池进行热失控实验，使用GC对热失控后的气体成分进行分析后发现，电池热失控产生的主要气体有H_2、CO、CO_2、CH_4、C_2H_2、C_2H_4及C_2H_6。Golubkov等[11]对3种不同正极材料的锂离子电池热失控特性进行了研究，使用GC对热失控后的气体成分进

收稿日期：2021-11-05；**录用日期**：2021-12-13；**网络出版时间**：2022-03-03 07:20
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220301.1707.002.html
基金项目：国家自然科学基金民航联合基金重点支持项目(U2033204)；中央高校基本科研业务费项目中国民航大学专项(3122023022)
* **通信作者**. E-mail：nkzqsong@126.com

引用格式：张青松，刘添添，郝朝龙，等. 锂离子电池热失控气体快速检测及危险性分析方法[J]. 北京航空航天大学学报，2023，49(9)：2227-2233. ZHANG Q S，LIU T T，HAO C L，et al. Rapid detection analysis method of thermal runaway gas composition and risk of lithium ion battery[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2227-2233 (in Chinese).

行分析。Sun 等[12] 对不同正极材料的商用锂电池燃烧产物的毒性进行研究，结果表明，100% SOC 的电池在热失控后生成的有毒气体种类最多，有毒气体的排放与电池材料、电池容量及电池 SOC 高度相关。Said 等[13] 研究了正极材料及气体环境对锂电池热失控传播的影响，使用气体传感器对实验过程中产生的 CO、CO_2、O_2 及烃类气体浓度进行测量。

目前，研究人员主要使用 GC-MS、FTIR 等设备对热失控气体成分进行分析，但存在以下不足：GC-MS 检测灵敏度高，可以对气体成分进行全面的分析，但分析时间较长，在单次实验中只能进行一次分析；FTIR 无法对 O_2、H_2 等双原子分子进行分析。综上，本文提出一种基于激光拉曼光谱气体检测技术的锂离子电池热失控气体检测方法，并通过实验验证其可行性。针对正极材料为 LCO 和 NCM 的锂离子电池，在 SOC 为 75% 和 100 % 的状态下开展实验，分析锂离子电池热失控后释放的气体成分，基于分析结果对锂离子电池热失控气体爆炸下限进行计算。

1 本文方法

1.1 气体成分分析

激光拉曼光谱气体检测技术是一种使用单频率激光对混合气体进行非接触、无损伤的检测手段。由于不同气体的分子结构不同，其产生的拉曼特征峰不同，可以对烃类和非烃类气体同时进行检测。此外，在对多种气体混合样品进行检测前，不需要进行不同气体的预先分离，从而可以减少不同气体成分分离造成的误差[14]。

激光拉曼光谱气体检测系统示意图如图 1 所示。激光器最大输出功率为 500 mW，使用 24 V 直流电源进行供电，激光器可提供波长为 532 nm 的激光作为激发光源，激光通过芯径 105 μm 的激发光纤传输至拉曼探头，并对锂离子电池热失控气体进行照射，照射产生的拉曼散射信号通过芯径 600 μm 的接收光纤传输至拉曼光谱仪，最终通过软件在电脑上形成拉曼光谱，单次测量时间为 5 s。

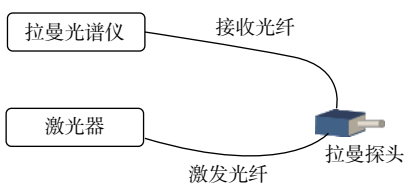


图 1 激光拉曼光谱气体检测技术原理示意图

Fig. 1 Schematic diagram of laser Raman spectroscopy gas detection technology

根据研究可知，锂离子电池热失控主要气体有 CO_2、CO、H_2 及烃类气体[15]，使用标准气体确认不同气体的拉曼特征峰，如图 2 所示。为确认气体特征峰高与气体体积分数之间的关系，使用 N_2 作为平衡气体，并配置 31 种气体体积分数构建气体定量模型。以 CO 为例，特征峰与气体体积分数关系如图 3 所示。为了更清楚地展现拉曼特征峰峰高与气体体积分数的关系，图 3 中只展示 11 组体积分数气体的拉曼特征峰。根据气体标定结果，拉曼

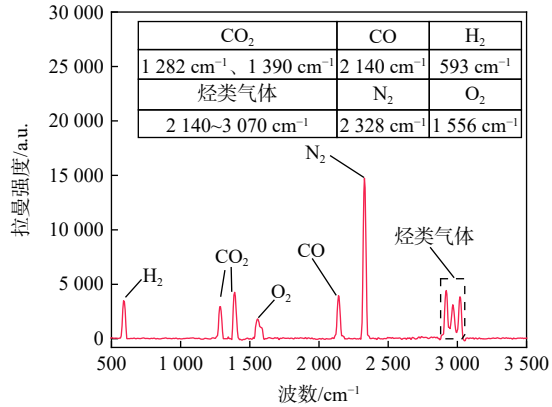
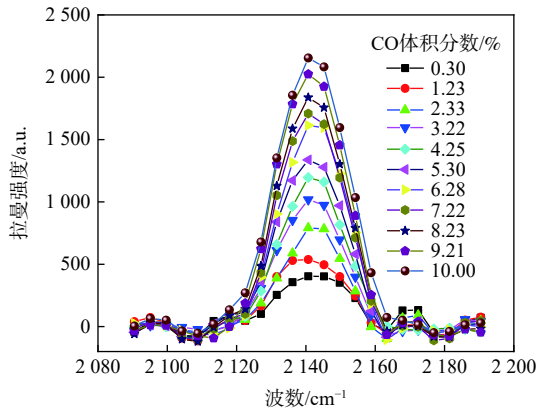
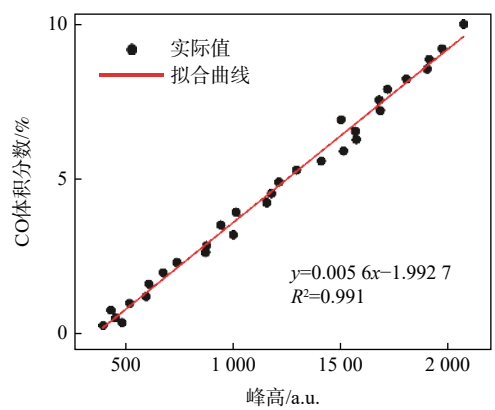


图 2 不同气体特征峰峰位

Fig. 2 Characteristic peak positions of different gases



(a) 11种CO体积分数与拉曼特征峰高曲线



(b) CO拉曼特征峰高与体积分数的定量回归曲线

图 3 拉曼信号与气体体积分数关系

Fig. 3 Relationship between Raman signal and gas concentration

光谱检测下限可达 0.3%, 为了减少随机噪声对测量结果的影响, 在分析气体结果时对每个样品进行 3 次光谱采集, 最终形成一个平均光谱数据并使用气体定量模型进行分析。

1.2　气体爆炸危险性分析

锂离子电池排放气体为多种气体混合, 混合气体爆炸极限可以根据 Le Chatelier 经验公式进行计算。当电池气体种类及体积分数已知的情况下, 爆炸下限通过式(1)和式(2)进行计算:

$$L_{\mathrm{ELmix}}=\frac{1}{\sum_{i=1}^{n}\frac{X_i}{L_{\mathrm{EL}i}}} \tag{1}$$

$$L'_{\mathrm{ELmix}}=L_{\mathrm{ELmix}}\frac{1+B/(1-B)}{100+L_{\mathrm{ELmix}}B/(1-B)}\times 100\% \tag{2}$$

式中: L_{ELmix} 为可燃气体混合物的爆炸下限, %; $L_{\mathrm{EL}i}$ 为气体混合物中一种组分的爆炸下限, %; X_i 为气体混合物中一种组分的摩尔分数或体积分数, %; n 为气体种类数; L'_{ELmix} 为含惰性气体的可燃气体混合物的爆炸下限, %; B 为惰性气体的体积分数。

2　验证实验设计

2.1　实验样品

实验选取正极材料分别为 LCO($LiCoO_2$) 和 NCM($LiNi_{0.8}Co_{0.1}Mn_{0.1}O_2$) 的 18650 型锂离子电池作为实验电池, 2 种电池的电解液主要由六氟磷酸锂($LiPF_6$)、碳酸二甲酯 (DMC) 和碳酸乙烯酯 (EC) 构成, 电池具体性能参数如表 1 所示。电池通过蓝电

表 1　实验电池参数

Table 1　Experimental battery parameters

正极材料	容量/(mA·h)	额定电压/V	截止电压/V	充电电压/V
LCO	2 200	3.6	2.65	4.2
NCM	2 600	3.635	2.5	4.2

CT3001 D 型电池充放电测试装置进行充放电。25% SOC 的电池首先以 0.2 C(440 mA, 520 mA)的放电速率放电至截止电压 (2.65 V, 2.5 V), 再以 0.5 C (1 100 mA, 1 300 mA)的充电速率充至相应 75% SOC; 100% SOC 的电池在放电至截止电压后, 以 0.5 C(1 100 mA, 1 300 mA)的充电速率将电压充至 4.2 V, 再通过恒压充电直至电流达到 0.02 C(44 mA, 52 mA)。为了保持电池稳定性, 在充放电结束后放入 25 ℃ 恒温箱静置 24 h。

2.2　实验流程

实验装置如图 4 所示。反应舱容积 5 L, 通过螺栓和垫片进行密封, 电池使用圆形加热圈进行加热, 加热圈内径 18 mm, 高度 65 mm, 加热功率 40 W, 加热圈为不闭合结构, 一个 K 型铠装热电偶放置于电池表面未被加热圈覆盖的部分用来测量电池表面温度, 另一个在电池安全阀上方 30 mm 处放置 K 型热电偶用来测量电池排出气体的温度。压力传感器监测实验舱内部压力变化, 使用抽气泵将反应舱内电池生成气体在管路中循环, 使气体循环管路与反应舱内气体体积分数达到一致。激光通过拉曼探头对气体池中的待测气体样品进行照射并形成拉曼光谱。

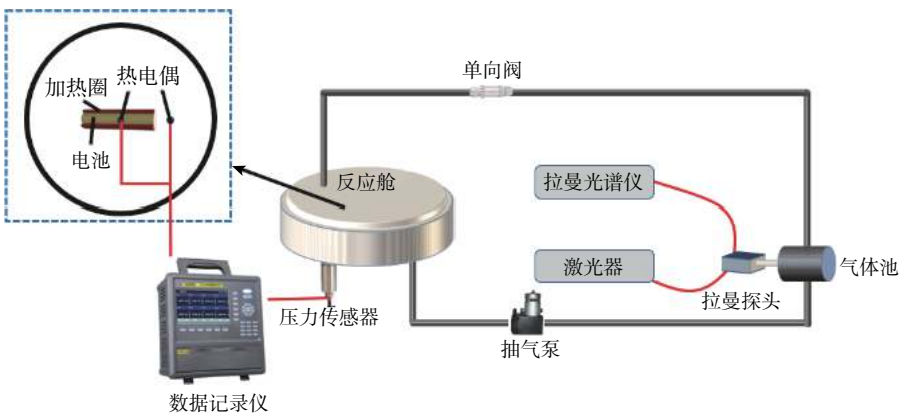


图 4　实验装置示意图

Fig. 4　Schematic diagram of experimental device

3　验证结果与讨论

3.1　锂离子电池热失控行为

不同实验工况下锂离子电池表面温度、释放气体温度、反应舱内部压力及电池电压变化如图 5 所示。锂离子电池热失控可分为 3 个阶段: ①阶段 1, 电池安全阀开启前, 电池表面温度在加热装置的作用下稳定升高; ②阶段 2, 电池安全阀开启, 电池内部反应生成的气体及部分电解液在压力作用下高速喷出, 反应舱内压力出现小幅度上升, 快速喷出的气体带走了部分热量, 并对电池有一定冷却作用, 此时电池表面温度出现小幅度下降, 随后电池

温升速度明显高于上一阶段；③阶段 3，电池发生热失控，内部发生多种放热副反应，生成大量气体，内部压力剧烈上升，电池内部材料随着快速喷出的气体电池表面温度、电池释放气体温度及反应舱内部压力急剧上升，电池电压完全消失。

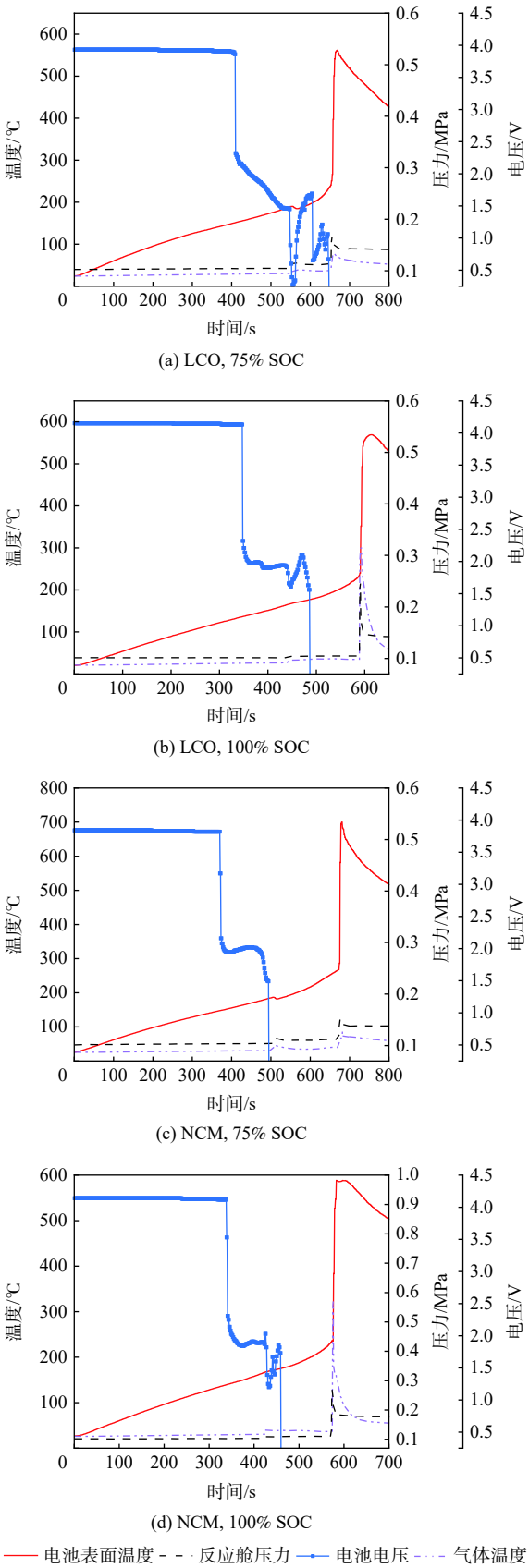


图 5 不同实验工况下温度、电压、压力曲线

Fig. 5 Temperature, voltage and pressure curves under different experimental conditions

不同实验工况下，电池安全阀打开温度 T_1、热失控起始温度 T_2、热失控时容器内部的压力变化 ΔP、释放气体最高温度 T_{gmax} 及热失控最高温度 T_{max} 如表 2 所示。可知，在电池正极材料相同时，安全阀打开温度、热失控起始温度随着 SOC 的升高而降低；热失控时容器压力变化、气体最高温度随 SOC 的升高而升高。锂离子电池在充电时，正极材料失去锂，导致正极材料中的 Ni 和 Co 等金属元素化合价上升，锂离子通过隔膜嵌入到石墨负极中。随着电池 SOC 的增加，电池正极材料失去的锂离子数量增加，高化合价的金属阳离子数量越多，而高价态的金属阳离子较为活泼，导致正极热稳定性变差；此外，随着 SOC 的增加，锂离子从嵌入到负极的数量增加，导致插层碳的不稳定性提高，反应活化能降低[16]。因此，电池 SOC 越高，就越容易发生热失控[17]。SOC 相同时，LCO 电池的安全阀打开温度与热失控起始温度高于 NCM 电池，说明相同条件下正极材料为 NCM 的电池更容易发生热失控。NCM 电池在热失控时的容器压力变化、气体最高温度及热失控最高温度高于 LCO 电池，说明热失控后产生气体量更大，释放热量更多。

3.2 锂离子电池热失控气体成分分析

由于锂离子电池热失控产生的主要气体成分为 H_2、CO、CO_2 及烃类气体，而烃类气体种类较多，本文对 H_2、CO、CO_2 及空气中的固有气体 N_2 和 O_2 进行标定，通过相减的方式计算出烃类气体的体积分数。根据标定后的拉曼特征峰峰高-气体体积分数模型计算不同实验工况下锂离子电池热失控气体体积分数如表 3 所示。锂离子电池热失控后生成大量气体，导致 N_2 与 O_2 含量明显下降。由于 N_2 作为惰性气体在热失控过程中不参与反应，而 N_2 体积分数随着 SOC 的增加逐渐降低，说明随着 SOC 的增加热失控气体量增加。100% SOC 的 NCM 电池在热失控后容器内的 O_2 含量大幅度下降，在热失控时发生了剧烈的燃烧现象，导致 CO_2 体积分数高于其余实验工况下的电池，但由于电池处于受限空间内，氧气在消耗后无法得到补充，燃烧不完全，导致 CO 占比明显高于其余电池。随着 SOC 的升高，电池储能的电能增加，内部物质热稳定性下降，热失控时的反应更加剧烈并最终导致各类生成的气体含量呈上升趋势。SOC 为 75% 时，2 种电池热失控后的 CO、CO_2、H_2 浓度相差不大，但 NCM 电池热失控后烃类气体含量明显

表 2　锂离子电池在不同实验条件下的实验数据

Table 2　Experimental data of lithium ion battery under various experimental conditions

正极材料	T_1/℃		T_2/℃		ΔP/MPa		T_{gmax}/℃		T_{max}/℃	
	75%SOC	100%SOC	75%SOC	100%SOC	75%SOC	100%SOC	75%SOC	100%SOC	75%SOC	100%SOC
LCO	189.03	164.77	266.43	233.82	0.038 6	0.141 8	77.9	300.03	561.65	569.24
NCM	186.24	162.59	264.95	232.98	0.051 5	0.165 6	85.69	321.26	699.95	588.3

表 3　锂离子电池热失控气体分析结果

Table 3　Analysis results of thermal runaway gas of lithium ion battery

正极材料	N_2体积分数/%		O_2体积分数/%		H_2体积分数/%		CO_2体积分数/%		CO体积分数/%		C_xH_y体积分数/%	
	75%SOC	100%SOC	75%SOC	100%SOC	75%SOC	100%SOC	75%SOC	100%SOC	75%SOC	100%SOC	75%SOC	100%SOC
LCO	67.89	54.54	14.31	11.26	5.92	8.81	4.86	5.30	3.58	11.95	3.44	8.14
NCM	59.75	39.89	14.57	4.71	5.66	11.56	5.21	7.25	4.21	22.02	9.31	14.57

高于 LCO 电池, 说明在热失控时发生的副反应类型较多。其中, H_2 主要来源于黏结剂聚偏氟乙烯(PVDF)在高温下与嵌入负极中的锂发生反应后的产物[18]:

$$—CH_2—CF_2+Li\longrightarrow LiF+—C=CF—+0.5H_2 \quad (3)$$

CO 则是由锂与电解液及 CO_2 的还原反应生成[19]:

$$DMC+2Li^++2e^-\longrightarrow CH_3OLi+CO \quad (4)$$

$$EC+2Li^++2e^-\longrightarrow CH_2OLi+CO \quad (5)$$

$$DEC+2Li^++2e^-\longrightarrow CH_3CH_2OLi+CO \quad (6)$$

$$2CO_2+2Li^++2e^-\longrightarrow Li_2CO_3+CO \quad (7)$$

SEI 中的亚稳定成分, 如 $ROCO_2Li$ 在 90～120 ℃范围热解, 释放 CO_2。此外, 电解液燃烧后同样会生成 CO_2[20]:

$$ROCO_2Li\longrightarrow ROLi+CO_2 \quad (8)$$

$$Electrolyte+O_2\longrightarrow H_2O+CO_2 \quad (9)$$

烃类气体则主要由 SEI 的分解及电解液与负极中镶嵌的 Li 反应生成[21]:

$$EC+2Li^++2e^-\longrightarrow Li_2CO_3+C_2H_4 \quad (10)$$

$$DMC+2Li^++2e^-+H_2\longrightarrow Li_2CO_3+CH_4 \quad (11)$$

$$(CH_2OCO_2Li)_2\longrightarrow Li_2CO_3+C_2H_4+CO_2+0.5O_2 \quad (12)$$

3.3　锂离子电池热失控气体爆炸危险性分析

为方便计算, 烃类气体爆炸下限可由 CH_4 替代[13]。H_2、CO、烃类气体(假设为 CH_4)的爆炸下限分别为 5%、12.5% 和 4%。结合式(1)和式(2)计算不同实验工况下锂离子电池热失控气体爆炸下限, 如图 6 所示。可知, 在不考虑惰性气体 CO_2 的情况下, 100% SOC 的 NCM 电池热失控气体爆炸下限最高, 主要原因是: 爆炸下限较高的 CO 占比较高, 导致混合气体的爆炸下限高于其余电池。电池种类相同时, 在不考虑惰性气体的情况下, 电池 SOC 越高, 爆炸下限越高, 混合气体爆炸危险性越低, 结合表 2 可得, 虽然高 SOC 的电池热失控产生威胁较大, 但低 SOC 的电池在热失控后残余的气体仍具有较高的爆炸危险性。在考虑惰性气体对混合气体爆炸下限的影响后, 所有实验工况下的混合气体爆炸下限均出现上升, 其中 75% SOC 的 LCO 电池爆炸下限上升最明显, 虽然 75% SOC 的 LCO 电池热失控后实验舱内的 CO_2 占比最低, 但除去实验开始前已经存在的 N_2 与 O_2 后, 其生成的 CO_2 占比在 4 种实验工况中最高, 导致其爆炸下限上升幅度最大。无论是否考虑惰性气体的影响, 各实验工况下的气体爆炸下限均低于 CO, 因此, 在处理锂电池热失控事故时应考虑残余气体二次爆炸的风险。

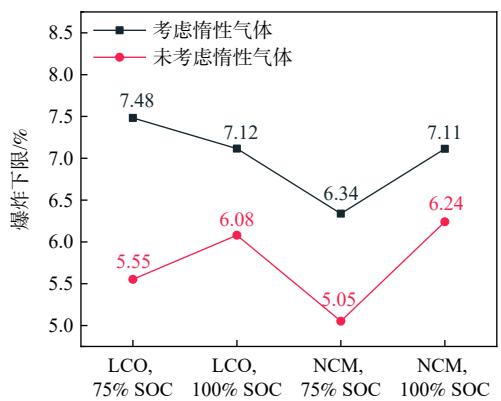


图 6　热失控气体爆炸下限

Fig. 6　Lower explosion limit of thermal runaway gas

4　结　论

1) 本文分析了激光拉曼光谱气体检测技术应用于锂离子电池热失控气体成分分析的可行性, 结合爆炸极限经验公式构建热失控气体成分及危险性快速分析方法, 通过实验对方法可行性进行验证。

2) 拉曼光谱分析结果表明, H_2、CO、CO_2 及烃类气体生成量随 SOC 的升高而升高, 相同 SOC 下

NCM电池热失控后生成的可燃气体量更多，100% SOC的NCM电池在热失控时出现燃烧现象并消耗了大量O_2。

3）热失控时生成的CO_2会显著提高混合气体的爆炸下限，在不考虑惰性气体对热失控气体爆炸下限的影响时，气体爆炸下限随SOC上升而上升，较低SOC的电池热失控生成的气体更容易达到爆炸条件。

参考文献（References）

[1] 华政, 梁风, 姚耀春, 等. 电动汽车电池的发展现状与趋势[J]. 化工进展, 2017, 36(8): 2874-2881.
HUA Z, LIANG F, YAO Y C, et al. Status and development trend for battery of electric vehicles[J]. Chemical Industry and Engineering Progress, 2017, 36(8): 2874-2881(in Chinese).

[2] 芮新宇, 冯旭宁, 韩雪冰, 等. 锂离子电池热失控蔓延问题研究综述[J]. 电池工业, 2020, 24(4): 193-201.
RUI X Y, FENG X N, HAN X B, et al. Review on the thermal runaway propagation of lithium-ion batteries[J]. Chinese Battery Industry, 2020, 24(4): 193-201(in Chinese).

[3] 张青松, 罗星娜, 程相静, 等. 基于锂离子电池温降指数的细水雾添加剂筛选方法[J]. 北京航空航天大学学报, 2020, 46(6): 1073-1079.
ZHANG Q S, LUO X N, CHENG X J, et al. Method for screening fine water mist additive based on temperature drop index of lithium-ion battery[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(6): 1073-1079(in Chinese).

[4] 许辉勇, 范亚飞, 张志萍, 等. 针刺和挤压作用下动力电池热失控特性与机理综述[J]. 储能科学与技术, 2020, 9(4): 1113-1126.
XU H Y, FAN Y F, ZHANG Z P, et al. Thermal runaway characteristics and mechanisms of Li-ion batteries for electric vehicles under nail penetration and crush[J]. Energy Storage Science and Technology, 2020, 9(4): 1113-1126(in Chinese).

[5] 张青松, 赵启臣. 过充循环对锂离子电池老化及安全性影响[J]. 高电压技术, 2020, 46(10): 3390-3397.
ZHANG Q S, ZHAO Q C. Effects of overcharge cycling on the aging and safety of lithium ion batteries[J]. High Voltage Engineering, 2020, 46(10): 3390-3397(in Chinese).

[6] 贺元骅, 郭君, 王海斌, 等. 振动环境对锂离子电池热失控危险性的影响[J]. 科学技术与工程, 2020, 20(29): 12242-12246.
HE Y H, GUO J, WANG H B, et al. Impact of vibration on thermal runaway risk of lithium-ion battery[J]. Science Technology and Engineering, 2020, 20(29): 12242-12246(in Chinese).

[7] 郭志慧, 崔潇丹, 赵林双, 等. 高镍三元锂离子电池火灾及气体爆炸危险性实验[J]. 储能科学与技术, 2022, 11(1): 193-200.
GUO Z H, CUI X D, ZHAO L S, et al. Experimental research on fire and gas explosion hazards of high nickel lithium-ion battery [J]. Energy Storage Science and Technology, 2022, 11(1): 193-200 (in Chinese).

[8] 张青松, 刘添添, 赵洋. 受限空间环境压力对三元锂离子电池热失控影响[J]. 中国安全生产科学技术, 2021, 17(6): 36-40.
ZHANG Q S, LIU T T, ZHAO Y. Effect of environmental pressure in confined space on thermal runaway of ternary lithium ion battery[J]. Journal of Safety Science and Technology, 2021, 17(6): 36-40(in Chinese).

[9] FERNANDES Y, BRY A, DE-PERSISI S. Identification and quantification of gases emitted during abuse tests by overcharge of a commercial Li-ion battery[J]. Journal of Power Sources, 2018, 389: 106-119.

[10] YUAN L M, DUBANIEWICZ T, ZLOCHOWER I, et al. Experimental study on thermal runaway and vented gases of lithium-ion cells[J]. Process Safety and Environmental Protection, 2020, 144: 186-192.

[11] GOLUBKOV A W, FUCHS D, WAGNER J, et al. Thermal-runaway experiments on consumer Li-ion batteries with metal-oxide and olivine-type cathodes[J]. RSC Advances, 2014, 4: 3633-3642.

[12] SUN J, LI J, ZHOU T, et al. Toxicity, a serious concern of thermal runaway from commercial Li-ion battery[J]. Nano Energy, 2016, 27: 313-319.

[13] SAID A O, LEE C, STOLIAROV S I. Experimental investigation of cascading failure in 18650 lithium ion cell arrays: Impact of cathode chemistry[J]. Journal of Power Sources, 2020, 446: 227347.

[14] BULDAKOV M A, KOROLKOV V A, MATROSOV I I, et al. Analyzing natural gas by spontaneous Raman scattering spectroscopy[J]. Journal of Optical Technology, 2013, 80(7): 426-430.

[15] SOMANDEPALLI V, MARR K, HORN Q. Quantification of combustion hazards of thermal runaway failures in lithium-ion batteries[J]. SAE International Journal of Alternative Powertrains, 2014, 3(1): 98-104.

[16] PENG Y, YANG L Z, JU X Y, et al. A comprehensive investigation on the thermal and toxic hazards of large format lithium-ion batteries with $LiFePO_4$ cathode[J]. Journal of Hazardous Materials, 2020, 381: 120916.

[17] BAK S M, HU E Y, ZHOU Y N, et al. Structural changes and thermal stability of charged $LiNi_xMn_yCo_zO_2$ cathode materials studied by combined in situ time-resolved XRD and mass spectroscopy[J]. ACS Applied Materials & Interfaces, 2014, 6(24): 22594-22601.

[18] PASQUIER A D, DISMA F, BOWMER T, et al. Differential scanning calorimetry study of the reactivity of carbon anodes in plastic Li-ion batteries[J]. Journal of the Electrochemical Society, 1998, 145(4): 472-477.

[19] YOSHIDA H, FUKUNAGA T, HAZAMA T, et al. Degradation mechanism of alkyl carbonate solvents used in lithium-ion cells during initial charging[J]. Journal of Power Sources, 1997, 68(2): 311-315.

[20] KIM G H, PESARAN A, SPOTNITZ R. A three-dimensional thermal abuse model for lithium-ion cells[J]. Journal of Power Sources, 2007, 170(2): 476-489.

[21] DIAZ F, WANG Y F N, WEYHE R, et al. Gas generation measurement and evaluation during mechanical processing and thermal treatment of spent Li-ion batteries[J]. Waste Management, 2019, 84: 102-111.

Rapid detection analysis method of thermal runaway gas composition and risk of lithium ion battery

ZHANG Qingsong*, LIU Tiantian, HAO Chaolong, QU Yirun, ZHANG Wei, CHEN Da

(Key Laboratory of Civil Aviation Thermal Hazards Prevention and Emergency Response, Civil Aviation University of China, Tianjin 300300, China)

Abstract: In order to quickly analyze the composition and risk of thermal runaway gas in lithium ion batteries, an analysis method based on laser Raman spectroscopy and the empirical formula is proposed. Using the independently designed gas detection platform of the lithium ion battery, the composition and explosion risk of runaway gas released by a 18650 lithium ion battery with different charge states (75%,100%) and different cathode materials (LCO,NCM) were studied. According to the experimental findings, the NCM battery releases more flammable gas following thermal runaway than the LCO battery under the same state of charge (SOC). After thermal runaway, the proportion of CO in the container of NCM battery with 100% SOC is as high as 22.02%. With the increase of SOC, the amount of combustible gas generated by the thermal runaway of lithium ion batteries increases. Under different experimental conditions, the gas after the thermal runaway of the lithium-ion battery still had a high explosion risk. The findings demonstrate the viability of using laser Raman spectroscopy to identify lithium-ion battery thermal runaway gas, and they offer a theoretical foundation and technical backing for the quick identification and evaluation of the composition of thermal runaway gas and explosion danger.

Keywords: lithium ion battery; thermal runaway; Raman spectroscopy; gas detection; explosion limit

Received: 2021-11-05; **Accepted**: 2021-12-13; **Published Online**: 2022-03-03 07:20

URL: kns.cnki.net/kcms/detail/11.2625.V.20220301.1707.002.html

Foundation items: Key Program of the Joint Fund for Civil Aviation Research with National Natural Science Foundation of China (U2033204); Central University Basic Research Business Fee Project Special Project of Civil Aviation University of China (3122023022)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0679

航班保障因果分析框架

邢志伟[1]，张林[1]，罗谦[2,*]，夏欢[2]，文涛[2]，张涛[2]

(1. 中国民航大学 电子信息与自动化学院，天津 300300；2. 中国民用航空局第二研究所 工程技术研究中心，成都 610041)

摘 要：航班离港延误是民航界一致关注并希望解决的重点及难点问题。针对航班离港延误产生原因不明晰及因果影响强度不确定的问题，提出了BLCNS-LV-IDA因果分析框架，从因果推断的角度出发，在航班地面保障业务范畴领域内，从定性和定量2个层面对航班离港推出延误进行因果分析。以航班离港推出延误时长为目标变量，基于特征选择和极大祖先图贪婪搜索(GSMAG)算法结合的BLCNS局部因果网络结构发现算法，构建因果网络模型；根据得到的因果网络，基于LV-IDA算法对各等价网络中的边缘进行因果效应评估。实验结果表明：BLCNS局部因果网络结构发现算法在处理变量较多的大样本数据集时有一定的优势，在50节点集的1 000、5 000、10 000样本量下，F_1值较基线算法分别提高了−0.303、0.008、0.132，呈明显上升趋势，且在运行时间方面至少缩短16.29%。节点间的因果效应明确了各节点对航班延误的具体影响强度，为航班保障的精细化管理、减少航班延误提供了指导。

关 键 词：航班离港延误；航班地面保障；特征选择；局部因果结构学习；因果效应评估
中图分类号：V351; TP181
文献标志码：A **文章编号**：1001-5965(2023)09-2234-10

航班地面保障是航班运行的核心环节，指航空器上轮挡到撤轮挡2个时刻间开展的旅客、货邮、航油、机务等不同场景的复杂服务过程，是决定航班准点的关键因素之一[1]。航班地面保障各工作节点间高度相关，前一节点的作业完成情况会对后续节点的作业链产生影响，并波及到航班的推出。鉴于上述航班地面保障的复杂性和高度相关性，为实现机场航班地面保障的精细化运行管理，提高机场航班准点率，将因果推理这种用于解释性分析的强大建模工具引入到航班保障的运行研究中，以辨识航班地面保障推出延误的原因节点，并估计其对应的因果效应强度，从定性和定量2个层次对航班延误进行因果分析。

贝叶斯网络是一种以概率统计为基础，先验知识同样本数据相结合的图模型，是进行因果定性分析的经典方法。基于贝叶斯网络模型，目前国内外学者[2-5]主要侧重于航班到达系统运行状态的评估、航班延误原因变量的定性分析、航班延误波及预测等问题的研究，并未从定量的角度对航班延误进行深层次分析。此外，贝叶斯网络模型以因果充分性假设为基础，未考虑潜在变量的存在，对数据的完备性要求过高，与实际存在偏差。除贝叶斯网络模型外，决策树、支持向量机、随机森林、神经网络等机器学习算法[6-9]同样作为经典方法被用于航班延误预测领域。上述经典方法虽然可以提高航班延误的预测精度，但由于机器学习算法的“黑箱特性”，对航班延误产生机理无法解释。为解决实际业务场景中数据完备性不足的问题，Triantafillou和Tsamardinos[10]提出一种基于分数的、可处理混淆的全局因果网络结构发现算法，即极大祖先图贪

收稿日期：2021-11-10；**录用日期**：2021-12-13；**网络出版时间**：2022-01-11 10:36
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220110.1932.006.html
基金项目：国家重点研发计划(2018YFB1601200)；四川省青年科技创新研究团队专项计划(2019JDTD0001)
*通信作者. E-mail：luoqian@caacetc.com

引用格式：邢志伟，张林，罗谦，等. 航班保障因果分析框架[J]. 北京航空航天大学学报，2023，49(9)：2234-2243.
XING Z W，ZHANG L，LUO Q，et al. Causal analysis framework of flight service[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2234-2243 (in Chinese).

婪搜索 (greedy search for maximal ancestral graph, GSMAG) 算法。与快速因果推断 (fast causal inference, FCI) 算法、真正快速因果推断 (really fast causal inference, RFCI) 算法等采用 FCI 定向规则识别混淆变量的方法不同, GSMAG 算法基于混合图的评分函数结合贪婪搜索方法对所有可能的 MAG 空间进行搜索, 构建完整的因果图网。在处理小样本数据集时, GSMAG 算法及其延伸算法 M3HC 效果显著, 但在处理高维数据时, GSMAG 等全局结构学习算法仍存在“维度灾难”。

针对上述航班延误预测经典方法所局限的数据完备性和“黑箱特性”, 以及全局因果结构学习所面临的“维度灾难”, 本文提出了基于 Boruta 的局部因果网络结构发现算法 (Boruta local causal network structure discovery algorithm, BLCNS), 仅关注于感兴趣变量的因果结构, 避免“维度灾难”, 并将该算法与 LV-IDA 算法相结合构成 BLCNS-LV-IDA 因果分析框架用于航班延误的因果分析。

1　BLCNS-LV-IDA 因果分析框架

BLCNS-LV-IDA 因果分析框架如图 1 所示。本文提出适用于航班保障的因果分析框架对航班保障过程进行深度解析。首先, 基于 Boruta 的特征选择算法得到目标变量的候选变量集, 在候选变量集间构建目标变量的局部因果祖先图的等价类; 然后, 对每个等价类模型中原因变量与航班延误间的因果强度进行估计, 以完成对航班推出延误的定性与定量分析。

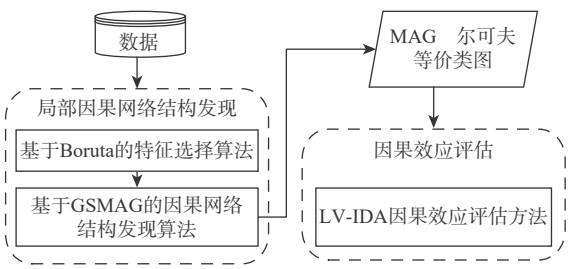


图 1　BLCNS-LV-IDA 因果分析框架

Fig. 1　BLCNS-LV-IDA causality analysis framework

2　BLCNS 局部因果网络结构发现算法

机场实际航班保障过程是一个大型网络结构。因果网络结构发现算法在构建航班保障的因果网络时, 算法的空间和时间复杂度会随节点的数量急剧增加, 且所构建的因果图存在大量与航班推出延误无关的冗余信息, 故可通过将特征选择与因果发现相结合的方法, 仅关注与航班推出延误相关的节点, 消除数据特征空间存在的冗余特征和噪声特征, 构建局部的航班保障因果图谱。

设 $\boldsymbol{D}$ 为一个包含连续变量 $\boldsymbol{V}$ 和 N 个独立同分布样本的数据集, 其中, $\boldsymbol{V}=\{v_i : i=1,\cdots,n\}$ 表示变量数为 n 的随机向量, 其服从多元正态分布 $N(0,\boldsymbol{\Sigma})$, 且协方差矩阵 $\boldsymbol{\Sigma}$ 是正定的。令 $m\in\{1,2,\cdots,m_{\text{tree}}\}$, 根据 Ahmed 等[11] 的理论, Boruta 算法是一种基于随机森林分类的包装型特征选择方法, 其通过对属性重要性进行多次评估, 并计算对应的 Z-score 值,识别出输入特征中的不相关噪声变量并将其删除。属性重要性计算如下:

$$M=\frac{1}{m_{\text{tree}}}\sum_{m=1}^{m_{\text{tree}}}\frac{\sum_{i\in O}I(v_j=f(v_i))}{|O|}-\frac{\sum_{i\in O}I(v_j=f(v_i^j))}{|O|}\tag{1}$$

式中: m_{tree} 为随机森林中树的数量; M 为损失精度的平均值, 由对象之间属性值的随机置换造成的分类准确率的损失衡量; O 为基于 bootstrap 聚合的每个训练样本的预测误差; $I(v_j=f(v_i))$ 为置换前的预测值; $I(v_j=f(v_i^j))$ 为置换后的预测值。

当连续变量 $\boldsymbol{V}$ 经 Boruta 算法特征选择后, 将与目标变量相关的属性放入 $Q=\{P_i,i=1,\cdots,P\}$, 相应的数据集 $\boldsymbol{\Delta}$, 经 Boruta 算法特征选择后, 形成包含连续变量 $\boldsymbol{P}$ 的新数据集。令 G 表示 MAG, 根据 Richardson 和 Spirtes[12] 的理论, 适用于 MAG 的 BIC 渐进评分准则如下:

$$\text{BIC}(\hat{\boldsymbol{\Sigma}},G)=-2\ln(l_G(\hat{\boldsymbol{\Sigma}}|G))+\ln N(2|\boldsymbol{P}|+|\boldsymbol{E}|)\tag{2}$$

式中: $l_G(\hat{\boldsymbol{\Sigma}}|G)$ 为图 G 的似然函数; $|\boldsymbol{P}|$ 和 $|\boldsymbol{E}|$ 分别为图 G 的顶点数和边数; 矩阵 $\hat{\boldsymbol{\Sigma}}$ 为协方差矩阵 $\boldsymbol{\Sigma}$ 根据与 G 关联的正态线性模型的最大似然估计。简而言之, MAG G 定义了一组线性方程 $\boldsymbol{P}=\boldsymbol{B}\cdot\boldsymbol{P}+\boldsymbol{\varepsilon}$。其中, $\boldsymbol{B}=[\beta_{ij}]$ 为矩阵, 且 $(\boldsymbol{I}-\boldsymbol{B})$ 是可逆的,如果 $P_j\to P_i$ 不在 G 中, 则 $\beta_{ij}=0$; $\boldsymbol{\varepsilon}$ 为使 $\boldsymbol{\Omega}=\text{Cov}(\boldsymbol{\varepsilon})=[w_{ij}]$ 正定的随机误差向量, 如果 $P_j\leftrightarrow P_i$ 不在 G 中, 则 $w_{ij}=0$。给定经验协方差矩阵 $\boldsymbol{S}$, 最大似然估计 $\hat{\boldsymbol{B}}$ 和 $\hat{\boldsymbol{\Omega}}$ 可以通过使用 Drton 等[13] 提出的残差迭代条件拟合 (residual iterative conditional fitting, RICF) 算法最大对数似然函数 $l_G(\boldsymbol{B},\boldsymbol{\Omega}|\boldsymbol{\Sigma})$ 来确定,相应的协方差矩阵计算为 $\hat{\boldsymbol{\Sigma}}=(\boldsymbol{I}-\hat{\boldsymbol{B}})\hat{\boldsymbol{\Omega}}((\boldsymbol{I}-\hat{\boldsymbol{B}})^{-1})^{\text{T}}$。

根据 Tian[14] 和 Richardson[15] 提出的图 G 的 c 分量分数分解特性, $l_G(\hat{\boldsymbol{\Sigma}}|G)$ 是可分解, c 分量对应于图 G 双向边的连通分量。根据图 G 的 c 分量可将似然函数分解为式(3), 而式(3)中每个 s_k 对应一个 c 分量[16]:

$$l_G(\hat{\boldsymbol{\Sigma}}|G)=-\frac{N}{2}\sum_k s_k\tag{3}$$

$$s_k=|c_k|\ln(2\pi)+\ln\left(|\hat{\boldsymbol{\Sigma}}G_k|\Big/\prod_{j\in\text{pa}_{G_k}}\sigma_{kj}^2\right)+\frac{N-1}{N}\text{tr}\left[\boldsymbol{\Sigma}_{G_k}^{-1}\boldsymbol{S}_{G_k}-|\hat{\text{pa}}_G(C_k)\backslash\{C_k\}|\right]\tag{4}$$

式中：C_k为c分量k中的节点；$\mathrm{pa}_G(C_k)$为C_k中节点的父节点；G_k为仅由$C_k \cup \mathrm{pa}_G(C_k)$中节点组成的图谱，$\widehat{\mathrm{pa}}_G(C_k)\backslash\{C_k\}$中的变量间没有任何边；$G_k$作为下标时是对参与$G_k$的矩阵行列的限制；$\sigma_{kj}^2$为$\boldsymbol{\Sigma}_{G_k}$对应父节点$k$的对角线项。

图2为BLCNS局部因果网络结构发现算法的流程。该算法分为2个阶段。第1阶段为特征选择阶段，通过Boruta算法筛选出与目标变量相关的最优特征集并存储在Q集合当中，具体操作如图3所示。为每个输入特征随机生成影子(替换)属性，多

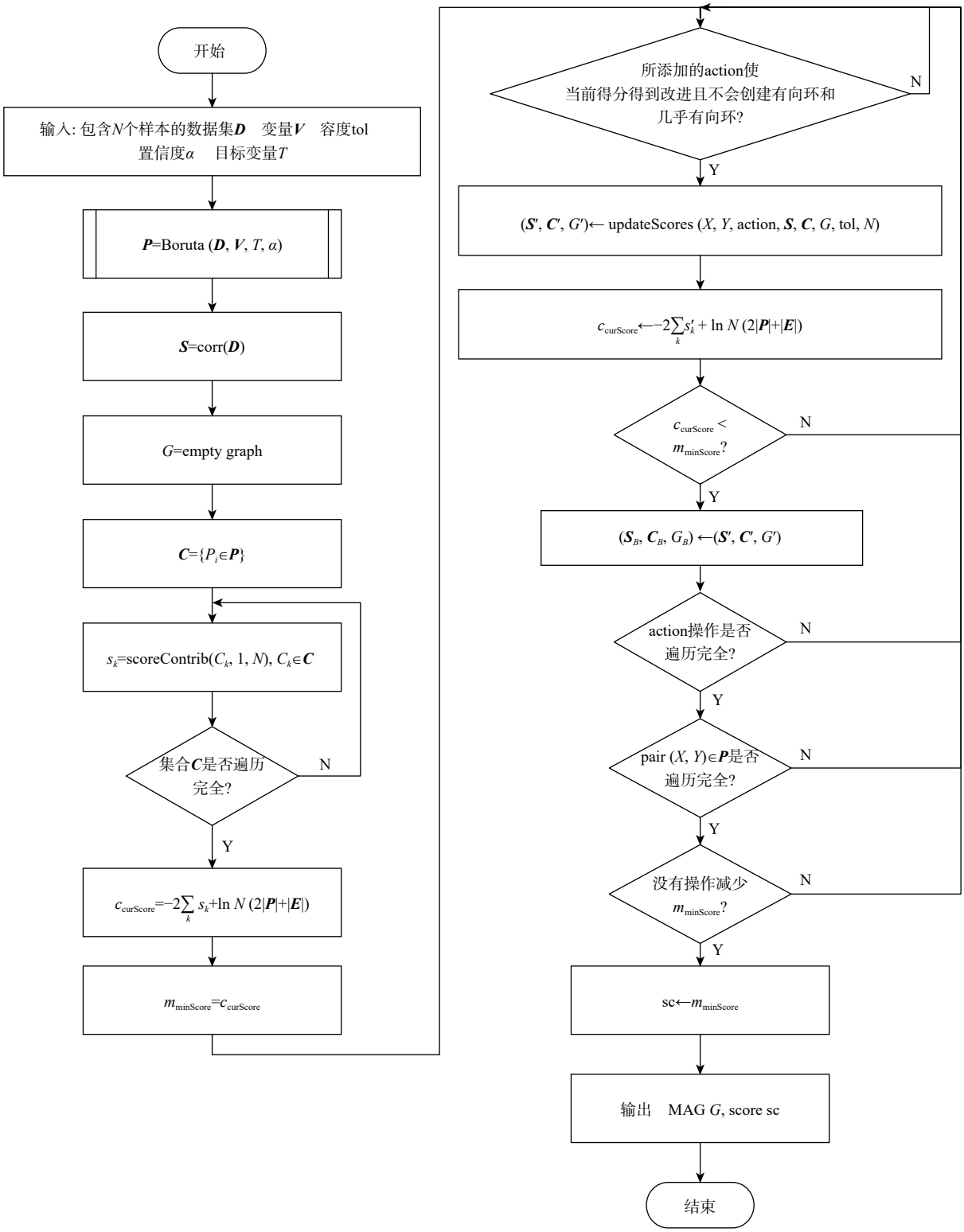


图2 BLCNS局部因果网络结构发现算法流程

Fig. 2 BLCNS local causal network structure discovery algorithm process

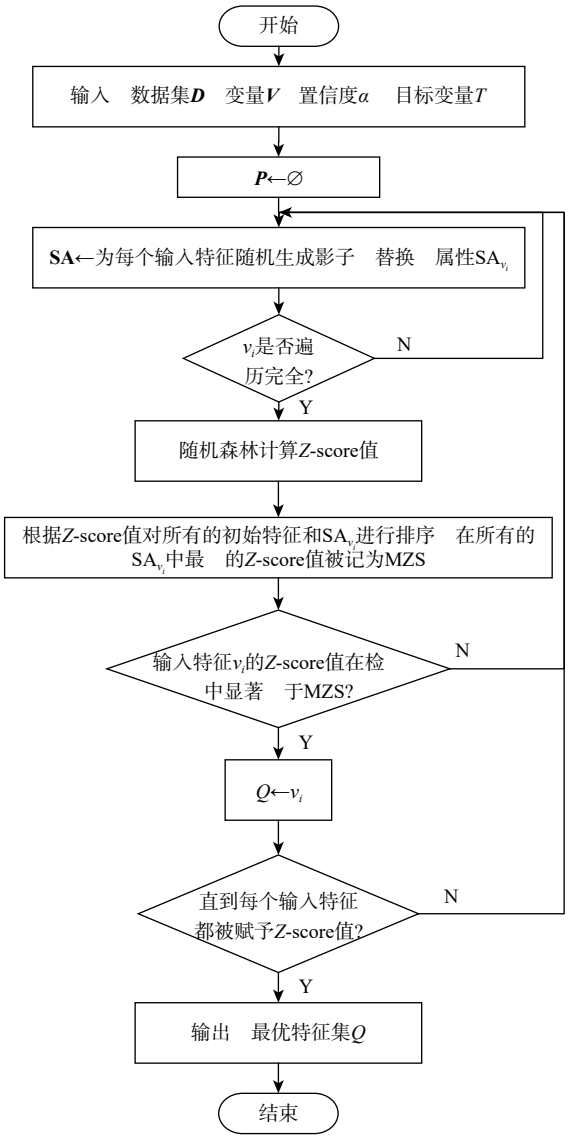


图 3　Boruta 算法流程

Fig. 3　Boruta algorithm flow chart

次利用随机森林算法计算 *Z*-sore 值, 并根据 *Z*-score 值对初始特征和影子属性进行排序。在所有的影子属性中, 最高的 *Z*-score 值被称为 MZS。*Z*-score 值在统计学检验中显著高于 MZS 的输入特征时, 作为与目标变量相关的特征, 反之删除。此外, 该算法还标识出一部分 *Z*-score 值较高但在统计学上并未显著高于 MZS 的弱相关变量, 这些变量可根据实际需要进行选择。第 2 阶段为祖先图构建阶段。设置空图作为初始模型, 在可能的 MAG 空间中搜索得分最高的图形, 在每次迭代过程中, 算法在每对变量$(X,Y)\in Q\cup T$间进行单边更改操作得到新的网络图谱, 操作包括 addLeft、addRight、addBidirected、orientLeft、orientRight、orientBidirected、remove、reverse。对每个新图, updateScores 函数都会更新 *c* 分量, 并计算其分值, 通过图 2 算法中的函数 scoreContrib 实现式(2), 用于计算 *c* 分量的分数, 以评估整个图的似然性和 BIC 分值。一旦某个动作提高了当前分数, 则更新当前的网络模型和分数, 直到进行任何操作也不会对当前分数进行改进, 该过程终止, 并返回当前的网络模型。

3　基于 LV-IDA 的因果效应估计

设 X 为二元处理变量, Y 为结果变量, X 对 Y 的因果效应 $\mathrm{CE}(X,Y)$ 是 X 的变化而导致的 Y 的变化。将$X=1$和$X=0$分别表示为x和x', 则$\mathrm{CE}(X,Y)$可定义为

$$\mathrm{CE}(X,Y)=E(Y|\mathrm{do}(x))-E(Y|\mathrm{do}(x')) \tag{5}$$

式中: $\mathrm{do}(x/x')$表示干预对变量 X 的影响。以航班保障为例, X 为航班保障的加航油环节的持续作业时间; Y 为航班推出延误时间; $\mathrm{CE}(X,Y)$的含义就为当航油的作业时间由干预发生单位时长的变化时, 航班推出延误期望时长的变化。

在此基础上, Maathuis 等[17] 提出了真正的因果结构已知时计算干预影响结果的方法, 即 IDA 算法。当没有隐藏变量或满足因果充分性的前提下, 通过 PC 算法等因果网络结构发现算法, 从数据中可以学到一个因果 DAG 的马尔可夫等价类, 而对于等价类中的每个 DAG, IDA 算法识别出相应变量对间的因果效应。当存在隐藏变量时, Malinsky 和 Spirtes[18] 提出基于 MAG 的 LV-IDA 算法, 该算法是对 IDA 算法的扩展, 如图 4 所示。

在估计因果效应时, 需要消除或调整其他变量的影响, 以获得对因果效应的无偏估计θ_{ik}, Θ_i^{L}作为存放θ_{ik}的 $p\times m$因果效应矩阵, 其每个元素都是 PAG 代表的某个 MAG 中 X_i对 Y 的因果效应θ_{ik},列对应协变量, 行对应等价类中的 MAG。Pearl 的后门准则是识别调整集合的重要准则。

对于代表线性结构方程模型的 DAG, X_i对 $Y, Y\notin \mathrm{pa}(X_i,G)$的因果效应为系数$\beta_{i|\mathrm{pa}(X_i,G)}$。$\beta_{i|\mathrm{pa}(X_i,G)}$表示 Y 对 X_i和父集合 $\mathrm{pa}(X_i,G)$的线性回归中 X_i前的回归系数, 一般地, 对于任何集合$S\subseteq\{X_1,\cdots,X_p,Y\}\backslash\{X_i\}$, X_i对 Y 的因果效应可为$\beta_{i|S}$。Malinsky 和 Spirtes[18] 将X_i的父母节点考虑进X_i和Y间的因果效应是因为集合 $\mathrm{pa}(X_i,G)$满足 Pearl 的后门准则。Maathuis 和 Colombo[19] 将此规则扩展成适用于 MAG、PAG 等的广义准则。图 3 算法的第 5～8 行为后门准则识别阶段, 该部分将无法识别的因果效应边以“NA”代替, 而将可识别部分的X_i的回归系数$\beta_{i|S}$作为该边的因果效应存放到Θ_i^{L}中。

possibleDe(X,G)为 G 中 X 的可能后代节点, 如果G为 MAG, 则 possibleDe(X,G)与子孙节点De(X,G)是等价的。设 X 和 Y 是混合图G中 2 个不同的顶点, 如果$V\neq X$且在 X、V之间存在碰撞路径, 使得这

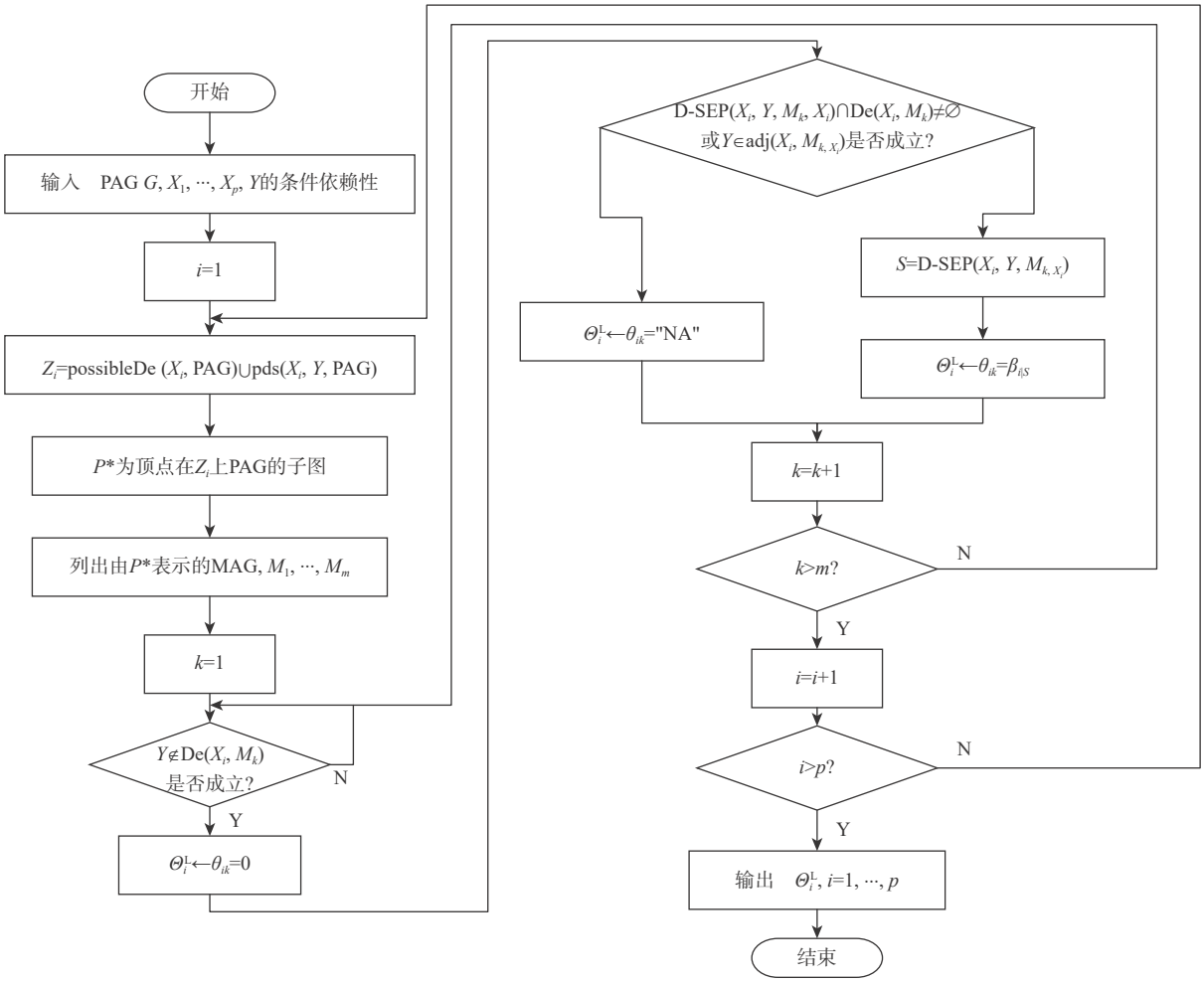


图4 LV-IDA算法流程

Fig. 4 LV-IDA algorithm flow chart

条路径上每个路径的每个顶点都是G中X或Y的祖先，则$V \in$ D-SEP(X,Y,G)。 pds(X_i,Y,G)为Possible-D-SEP(X_i,Y,G)的简称，且仅当G中X_i和V之间存在路径π，对于π上的每个子路径$< X_m, X_l, X_h >$，X_l是G中子路径上的碰撞器，或者三元组$< X_m, X_l, X_h >$的每对顶点都是相邻的，则可令$V \in$ pds(X_i,X_j,G)。

4 实验分析

BLCNS-LV-IDA因果分析框架可分为2部分：因果网络结构发现部分和因果效应估计部分。因果网络结构发现部分，分别在10、20和50个变量上随机模拟50个DAG网络，并以真实的PAG进行记录。为了控制DAG的稀疏性，每个变量最多有3个父节点，而对于每个DAG网络，将10%的变量设为隐藏变量（即分别设有1、2和5个隐藏变量），任意1个变量作为目标变量。因果效应估计部分采用真实的航班保障数据，构建航班保障关于延误的局部因果网络图，并计算因果效应。实验在R语言4.1.1版本中完成。

4.1 BLCNS局部因果网络结构发现

为了验证BLCNS因果网络结构发现算法的性能，本节分别在10、20和50这3个不同变量数的数据集上各生成1 000条、5 000条和10 000条数据，图5为上述仿真数据经Boruta算法筛选后的重要变量、不重要变量和弱相关变量的数量对比。

由图5可以看出，1 000条数据量时，不重要变量和弱相关变量共占据初始变量的0.50、0.53、0.75；5 000条数据量时，占据初始变量的0.38、0.29、0.52；10 000条数据量时，占据初始变量的0.25、0.24、0.48。综上，不重要和弱相关变量占据初始变量的0.24～0.75。采用Boruta算法进行初步筛选后，与目标变量不相关或关联较弱的变量被全部清除，仅保留相关或弱相关变量作为因果网络构建模型的输入，大幅降低了算法的空间复杂度。通过将Boruta算法与GSMAG算法相结合构建目标变量的BLCNS局部因果网络模型，并将其与M3HC算法[20]进行比较，以对BLCNS进行评估。M3HC算法为因果充分性

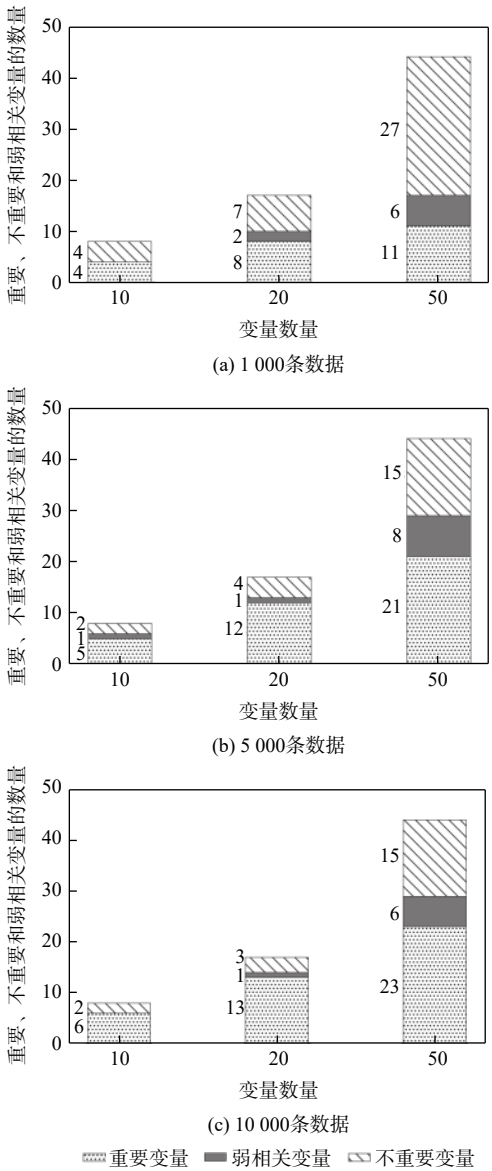


图 5　不同数据集时重要变量、不重要变量和弱相关变量的数量对比

Fig. 5　Comparison of number of important variables, unimportant variables and weakly correlated variables in different datasets

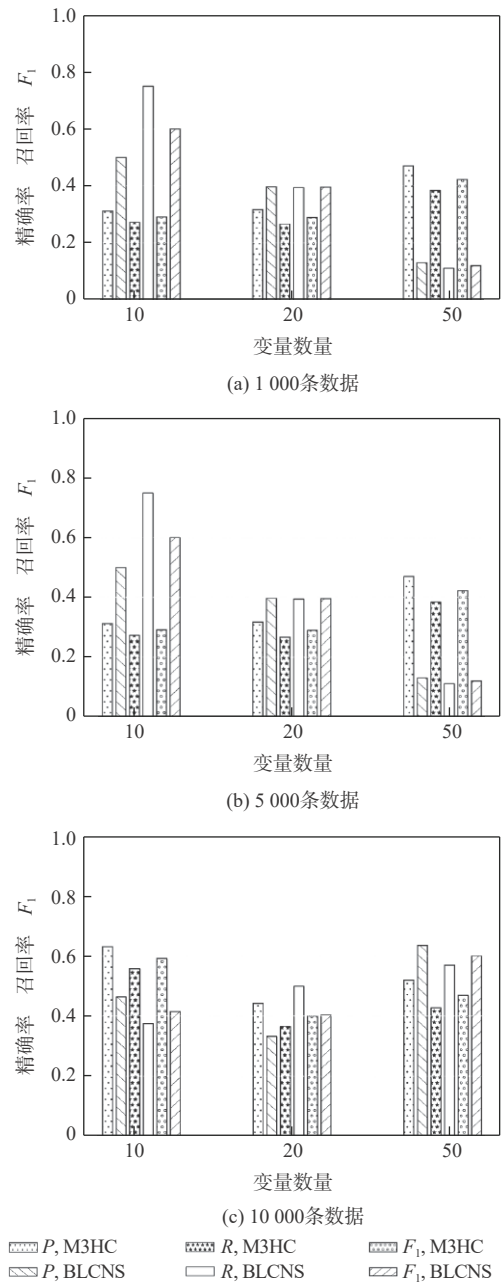


图 6　不同数据集时 BLCNS 与 M3HC 算法准确率、召回率和 F_1 值的比较

Fig. 6　Comparison of accuracy, recall and F_1 value of BLCNS and M3HC algorithms in different datasets

假设不成立时, 基于评分的混合因果网络发现算法, 其将 MMPC 算法与 GSMAG 算法相结合, 构建完整的因果网络结构, 与原始的 GSMAG 算法相比,算法在准确率等方面均有一定的提升。将 M3HC 算法作为基准算法, 在准确率 P、召回率 R、F_1 值和运行时间等方面与其进行比较, 如图 6 和图 7 所示。准确率、召回率、F_1 值计算如式(6)～式(8)所示, 具体数值如表 1～表 3 所示。

$$P=\frac{C}{I} \tag{6}$$

$$R=\frac{C}{C+F} \tag{7}$$

$$F_1=\frac{2P}{P+R} \tag{8}$$

式中: P 为准确率, 指正确预测的个数 C 占预测总个数 I 的百分比, 用于评估算法输出中误报的数量; R 为召回率, 指正确预测的数量 C 占给定真实节点总数的比例, F 为未被正确识别的样本节点数量; F_1 指标综合了准确率和召回率的均衡指标。

将图 6 与表 1～表 3 相结合可以看出, 在样本变量数量过多而样本量过少的情况下, 如样本变量数达到 50, 而样本集的数量仅有 1 000, M3HC 算法

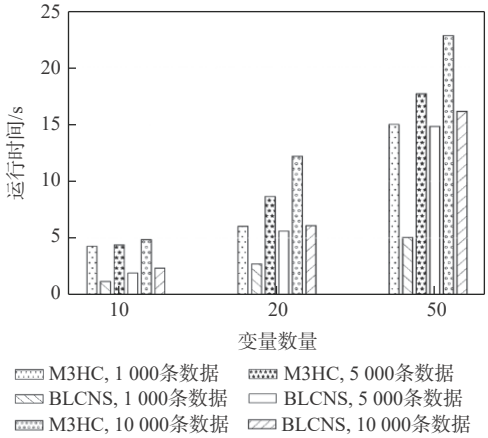


图7 不同数据集时BLCNS与M3HC算法运行时间比较

Fig. 7 Comparison of running time of BLCNS and M3HC algorithms in different datasets

明显优于BLCNS算法，其F_1值比BLCNS算法高0.303，而随着样本量的升高，BLCNS算法的F_1值逐步与M3HC算法持平，甚至略有超出，超出值可达0.132，这是因为对于变量多的大样本数据集，BLCNS算法通过特征选择的方式将与目标变量无关的变量进行过滤，缩小了网络构建的空间，此外相对于大样本数据集，GSMAG在处理中小型数据集时，准确率与召回率更高，相比于M3HC算法,在处理大样本数据集时优势更为明显。同理，对于变量较少的数据集，BLCNS所依赖的特征选择算法无法过滤较多的冗余变量，故其在处理小样本数据集时发挥并不稳定。而在运行时间方面，BLCNS算法较M3HC算法分别可缩短72.85%、55.16%、66.40%，优势较为明显。

表1 1 000条数据时BLCNS与M3HC算法的准确率、召回率、F_1值和运行时间

Table 1 Accuracy, recall, F_1 value and run time of BLCNS and M3HC algorithms in 1 000 data

变量数量	准确率		召回率		F_1		运行时间/s	
	M3HC	BLCNS	M3HC	BLCNS	M3HC	BLCNS	M3HC	BLCNS
10	0.312	0.500	0.272	0.750	0.291	0.600	4.263	1.157
20	0.317	0.397	0.266	0.394	0.289	0.396	6.032	2.705
50	0.470	0.130	0.384	0.111	0.423	0.12	14.992	5.038

表2 5 000条数据时BLCNS与M3HC算法的准确率、召回率、F_1值和运行时间

Table 2 Accuracy, recall, F_1 value and run time of BLCNS and M3HC algorithms in 5 000 data

变量数量	准确率		召回率		F_1		运行时间/s	
	M3HC	BLCNS	M3HC	BLCNS	M3HC	BLCNS	M3HC	BLCNS
10	0.738	0.750	0.682	1.000	0.709	0.857	4.386	1.899
20	0.412	0.389	0.354	0.304	0.381	0.341	8.668	5.601
50	0.632	0.600	0.520	0.560	0.571	0.579	17.699	14.815

表3 10 000条数据时BLCNS与M3HC算法的准确率、召回率、F_1值和运行时间

Table 3 Accuracy, recall, F_1 value and run time of BLCNS and M3HC algorithms in 10 000 data

变量数量	准确率		召回率		F_1		运行时间/s	
	M3HC	BLCNS	M3HC	BLCNS	M3HC	BLCNS	M3HC	BLCNS
10	0.632	0.464	0.558	0.375	0.593	0.415	4.839	2.337
20	0.442	0.333	0.366	0.500	0.400	0.404	12.178	6.087
50	0.520	0.636	0.428	0.570	0.469	0.601	22.839	16.150

4.2 基于真实数据集的BLCNS-LV-IDA因果分析

本节根据中国某机场6月至8月的航班保障日志，提炼出航班保障各节点的工作时长与各航班的延误时间，形成初始数据集，而对于初始数据集中空白部分，以多重插补的方法进行填充，将填充后的数据集导入BLCNS算法进行因果网络结构的发现，以LV-IDA算法进行因果效应估计，该机场的航班地面保障的因果关系及因果效应如图8所示。图中：各节点采用标号来表示。以标号v_{18}代表的航班延误为例，变量v_2为变量v_{18}的原因变量，v_2代表开始值机与航班到站的间隔时长，其余节点的标号如表4所示。依据LV-IDA变量v_2对变量v_{18}的因果效应为<NA,0.063|8,8>，其中，“|”后面的8分别表示“NA”和0.063出现的次数，“NA”表明目标变量的协变量集不足排除(使用后门标准)某些MAG中的混杂。通过其他估计值可以看出变量v_2对延误的因果效应为0.063，根据式(3)可知，开始值机与航班到站的间隔时长对航班推出延误的因果影响强

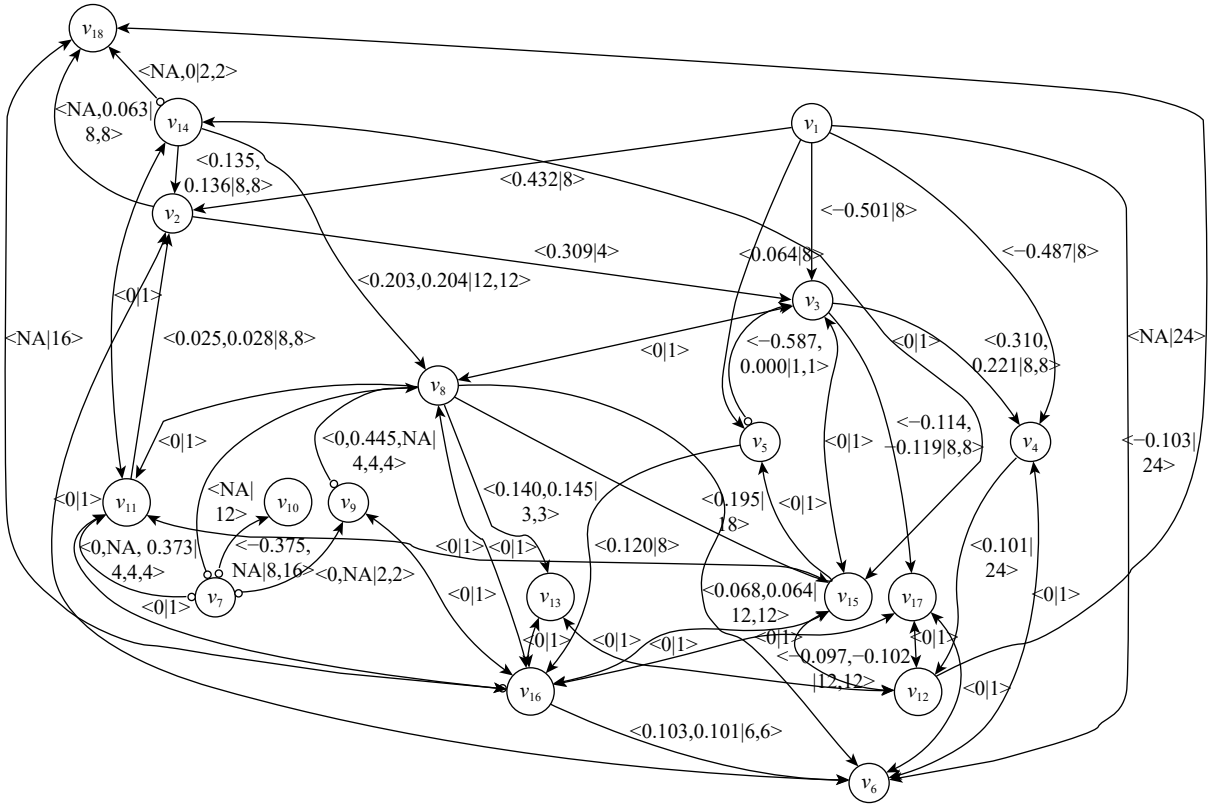


图 8　航班保障因果图

Fig. 8　Flight service causal graph

度为 0.063, 即当开始值机与航班到站的间隔时长在减少或者增加 1 h 时, 航班延误缩短或延长 0.06 h。

表 4　航班保障各节点标识

Table 4　Identification of flight service nodes

节点变量	节点含义
v_1	机位变更—开始值机
v_2	开始值机—到达本站
v_3	开始值机—截止值机
v_4	到达本站—进港桥电人员到位
v_5	进港桥电人员到位—进港桥电车辆到位
v_6	卸机人员到位—卸机车辆到位
v_7	放行签字放行—开始登机
v_8	装机关货舱门—上客登机结束
v_9	进港桥电撤离设备—上客登机结束
v_{10}	客舱清洁开始清洁—客舱清洁结束
v_{11}	客舱清洁结束—上客开始登机
v_{12}	上客开始登机—上客登机结束
v_{13}	靠桥梯人员到位—靠桥梯廊桥检查
v_{14}	上客人员到位—下客开舱门
v_{15}	下客结束—上客车辆到位
v_{16}	上客车辆到位—上客开始登机
v_{17}	牵引车开始作业—牵引车作业结束
v_{18}	航班延误

航班值机之所以能影响航班延误, 其深层次原因在于: 航班开始值机的时间本身会影响旅客到达、办理行李托运等时间, 而这些业务制约着航班地面保障的正常运转, 当提前航班开始值机时间时, 旅客到达机场、办理行李业务的时间会更积极, 对减轻航班延误具有一定的积极意义。

从以上研究可以看出, BLCNS-LV-IDA 因果分析框架对航班延误的分析和改善具有相当的价值和意义, 但其仍存在一定缺陷, 该分析框架的第 2 阶段采用 LV-IDA 算法进行因果效应估计, 其优势在于可在因果不充分的条件下进行效应估算, 但并不能对因果网中所有的边缘进行识别, 变量v_{18}的原因变量共有 4 个: $\{v_2, v_{12}, v_{14}, v_{16}\}$,但其仅能识别变量$v_2$和变量$v_{18}$间的因果效应。

5　结　论

1) 本文提出了 BLCNS-LV-IDA 因果分析框架, 将因果网络结构发现方法和因果效应估计方法相结合, 应用于实际问题的研究。

2) 结合航班地面保障的复杂性, 在考虑因果充分性假设不成立时, 提出基于 Boruta 的特征选择的局部因果网络发现算法, 提升算法效率。但本文框架仍存在一定缺陷, 其所基于的 LV-IDA算法并不能识别因果网络中所有的边缘, 故这也是在接下来所要研究的重点问题。

参考文献（References）

[1] 李彪, 王立文, 邢志伟. 过站航班地面保障流程效能评估[J]. 系统工程与电子技术, 2020, 42(7): 1543-1549.
LI B, WANG L W, XING Z W. Effectiveness evaluation of ground support process for transit flight[J]. Systems Engineering and Electronics, 2020, 42(7): 1543-1549(in Chinese).

[2] RODRÍGUEZ-SANZ Á, COMENDADOR F G, VALDÉS R A, et al. Assessment of airport arrival congestion and delay: Prediction and reliability[J]. Transportation Research Part C:Emerging Technologies, 2019, 98: 255-283.

[3] 曹卫东. 基于改进贝叶斯网络结构学习的航班延误波及分析[D]. 天津: 天津大学, 2009.
CAO W D. Flight delay propagation analysis based on improving Bayesian networks structure[D]. Tianjin: Tianjin University, 2009 (in Chinese).

[4] 许保光, 刘倩倩, 高敏刚. 基于机场繁忙程度的航班延误波及分析[J]. 中国管理科学, 2019, 27(8): 87-95.
XU B G, LIU Q Q, GAO M G. The flight delay propagation analysis based on airport busy state[J]. Chinese Journal of Management Science, 2019, 27(8): 87-95(in Chinese).

[5] 刘玉洁. 基于贝叶斯网络的航班延误与波及预测[D]. 天津: 天津大学, 2009.
LIU Y J. Flight delay the estimation of flight delay and propagation based on Bayesian networks[D]. Tianjin: Tianjin University, 2009 (in Chinese).

[6] 周覃, 高强, 马农. 基于聚类分析和CHAID决策树算法的航班延误预测研究[J]. 武汉理工大学学报, 2017, 39(11): 32-40.
ZHOU T, GAO Q, MA N. Flight delay prediction based on clustering analysis and CHAID decision tree algorithm[J]. Journal of Wuhan University of Technology, 2017, 39(11): 32-40(in Chinese).

[7] REBOLLO J J, BALAKRISHNAN H. Characterization and prediction of air traffic delays[J]. Transportation Research Part C:Emerging Technologies, 2014, 44: 231-241.

[8] 陈超. 基于GP-LVM和LS-SVM航班延误等级预测研究[J]. 无线互联科技, 2020, 17(6): 10-11.
CHEN C. Study on flight delay grade prediction based on GP-LVM and LS-SVM[J]. Wireless Internet Technology, 2020, 17(6): 10-11 (in Chinese).

[9] KHANMOHAMMADI S, TUTUN S, KUCUK Y. A new multi-level input layer artificial neural network for predicting flight delays at JFK airport[C]//Conference on Engineering Cyber Physical Systems: Applying Theory to Practice. Berlin: Springer, 2016, 95: 237-244.

[10] TRIANTAFILLOU S, TSAMARDINOS I. Score-based vs constraint-based causal learning in the presence of confounders[C]// Proceeding of "Causation: Foundation to Application" Workshop, Uncertainty in Artificial Intelligence, 2016: 59-67.

[11] AHMED A A M, DEO R C, GHAHRAMANI A, et al. LSTM integrated with Boruta-random forest optimiser for soil moisture estimation under RCP4.5 and RCP8.5 global warming scenarios[J]. Stochastic Environmental Research and Risk Assessment, 2021, 35: 1851-1881.

[12] RICHARDSON T, SPIRTES P. Ancestral graph Markov models[J]. The Annals of Statistics, 2002, 30(4): 962-1030.

[13] DRTON M, EICHLER M, RICHARDSON T S. Computing maximum likelihood estimates in recursive linear models with correlated errors[J]. Journal of Machine Learning Research, 2009, 10: 2329-2348.

[14] TIAN J. Identifying direct causal effects in linear models[C]//Proceedings of the 20th National Conference on Artificial Intelligence. Reston: AIAA, 2005: 346-353.

[15] RICHARDSON T S. A factorization criterion for acyclic directed mixed graphs[EB/OL]. (2014-06-26)[2021-11-01]. https://arxiv.org/abs/1406.6764.

[16] NOWZOHOUR C, MAATHUIS M, BÜHLMANN P. Structure learning with bow-free acyclic path diagrams[EB/OL]. (2017-12-02)[2021-11-01]. https://arxiv.org/abs/1508.01717v2.

[17] MAATHUIS M H, KALISCH M, BÜHLMANN P. Estimating high-dimensional intervention effects from observational data[J]. The Annals of Statistics, 2009, 37(6A): 3133-3164.

[18] MALINSKY D, SPIRTES P. Estimating causal effects with ancestral graph Markov models[C]//Conference on Probabilistic Graphical Models. Brookline: Microsoft Publishing, 2016: 299-309.

[19] MAATHUIS M H, COLOMBO D. A generalized back-door criterion[J]. The Annals of Statistics, 2015, 43(3): 1060-1088.

[20] TSIRLIS K, LAGANI V, TRIANTAFILLOU S, et al. On scoring maximal ancestral graphs with the max-min hill climbing algorithm[J]. International Journal of Approximate Reasoning, 2018, 102: 74-85.

Causal analysis framework of flight service

XING Zhiwei[1]，ZHANG Lin[1]，LUO Qian[2,*]，XIA Huan[2]，WEN Tao[2]，ZHANG Tao[2]

(1. School of Electronic Information and Automation，Civil Aviation University of China，Tianjin 300300，China；
2. Engineering Technology Research Center，The Second Research Institute of CAAC，Chengdu 610041，China)

Abstract：Flight departure delay is a key and difficult problem that the civil aviation industry is unanimously concerned about. Aiming at the problem of unclear reasons for flight delays, the BLCNS-LV-IDA causality analysis framework is proposed. In the discipline of flight ground support business, causal analysis of flight departure delays is done on both a qualitative and quantitative level from the standpoint of causal inference. Firstly, a causal network model is built with the departure delay time as the goal variable using the BLCNS local causal structure learning algorithm, which combines feature selection with greedy search for maximal ancestral graph (GSMAG). Secondly, the causal effect of each edge in each equivalent network is evaluated based on the LV-IDA algorithm according to the obtained causal network. The experimental results show that the BLCNS causality discovery method has certain advantages in dealing with large sample data sets of multiple variables. In a data set with 50 nodes, compared with the baseline algorithm, the F_1 value of BLCNS increased by −0.303, 0.008, and 0.132, respectively, with sample sizes of 1 000, 5 000, and 10 000. It presents a clear upward trend. In addition, the running time of BLCNS has been shortened by 16.29%. The causal effect between nodes clarifies the specific strength of each node's impact on flight delays. It provides guidance for the fine management of flight guarantees and reducing flight delays.

Keywords：flight departure delay; flight ground support; feature selection; local causal structure learning; causal effect estimation

Received：2021-11-10；**Accepted**：2021-12-13；**Published Online**：2022-01-11 10：36

URL：kns.cnki.net/kcms/detail/11.2625.V.20220110.1932.006.html

Foundation items：National Key R & D Program of China (2018YFB1601200); Sichuan Youth Science and Technology Innovation Research Team Program (2019JDTD0001)

* **Corresponding author**. E-mail：luoqian@caacetc.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0688

基于显著感知与一致性约束的目标跟踪算法

国强[1,2,*]，吴天昊[1,2]，徐伟[1,2]，CHORNOGOR Leonid[1,3]

(1. 哈尔滨工程大学 信息与通信工程学院，哈尔滨 150001；2. 先进船舶通信与信息技术工业和信息化部重点实验室，哈尔滨 150001；
3. 哈尔科夫国立大学 无线电物理学院，哈尔科夫 61166)

摘　　要：针对空间正则相关滤波 (SRDCF) 算法正则权重固定和模型退化等问题，提出了一种基于显著感知与一致性约束的目标跟踪算法。提取方向梯度直方图特征、浅层特征及中层特征进行融合，提升物体外观模型的表达能力；通过显著性检测算法获得初始帧的显著感知参考权重，建立正则权重在相邻2帧之间的关联；最小化实际一致性响应与理想一致性响应之间的差异，约束一致性水平防止滤波器模板退化；提出一种动态约束策略，进一步提高跟踪器在复杂场景下的适应性。在OTB2015、TempleColor128和UAV20L公开数据集上对所提算法进行测试，实验结果表明：相比于SRDCF算法，所提算法在OTB2015数据集上距离精度提高了0.108，AUC提高了0.077，速度为22.41帧/s，实时性较好。

关 键 词：机器视觉；目标跟踪；相关滤波；显著感知；一致性约束
中图分类号：TP391.4
文献标志码：A　　**文章编号**：1001-5965(2023)09-2244-14

目标跟踪是一个机遇与挑战并存的机器视觉研究方向[1]，其结合了机器学习、图像处理、凸优化等理论，在智能视觉导航、智能视频监控、人机交互及三维重建等军事及民用领域应用广泛。机遇主要体现在经过十几年的研究，基于相关滤波的目标跟踪算法性能得到不断的提升，但仍然存在进一步深入研究和探索的空间。挑战主要体现在跟踪过程中会经历遮挡、目标快速运动、光照改变等不利的因素。因此，如何研究出性能优异的跟踪器仍然是一项富有挑战性的工作。

近期，基于相关滤波的目标跟踪算法凭借其跟踪精度高和实时跟踪的优势，受到国内外学者的广泛关注。Bolme等[2]提出了最小输出误差平方和(minimum output sum of squared error, MOSSE)目标跟踪算法，将通信领域的相关滤波引入到目标跟踪中，通过傅里叶变换将时域的复杂运算转换到频域进行加速求解，算法速度高达669帧/s。Henriques等[3]提出了核相关滤波(kernelized correlation filters, KCF)目标跟踪算法，将核函数和多通道的HOG特征引入到相关滤波器中，不仅增强了模型的判别能力，而且保持了较快的跟踪速度。之后，Bertinetto等[4]融合HOG特征和全局颜色直方图特征的互补优势，进一步提升算法的精度。

基于手工特征的相关滤波算法实时性较好，但在相似背景干扰或遮挡等条件下，易发生跟踪错误或漂移。随着深度学习的发展，卷积神经网络提取的深度特征应用于相关滤波领域来提高跟踪的精度和鲁棒性。Danelljan等[5]以空间正则相关滤波(spatially regularized discriminative correlation filters, SRDCF)算法为基线提出DeepSRDCF算法，采用单层卷积特征代替HOG特征，跟踪性能得到明显提升。刘芳等[6]提出了一种基于残差学习的自适应

收稿日期：2021-11-16；**录用日期**：2022-03-25；**网络出版时间**：2022-05-17 17:06
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220517.1418.003.html
基金项目：国家重点研发计划(2018YFE0206500)；国家自然科学基金(62071140)；国家国际科技合作专项(2015DFR10220)
* **通信作者**. E-mail：guoqiang@hrbeu.edu.cn

引用格式：国强，吴天昊，徐伟，等. 基于显著感知与一致性约束的目标跟踪算法[J]. 北京航空航天大学学报，2023，49(9)：2244-2257.
GUO Q，WU T H，XU W，et al. Target tracking algorithm based on saliency awareness and consistency constraint[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2244-2257 (in Chinese).

无人机目标跟踪算法, 通过构建深度网络提取目标特征, 将提取的深度特征融入 KCF 算法, 进一步提高跟踪定位的精度。Ma 等[7] 提出分层卷积跟踪 (hierarchical convolutional features for visual tracking, HCF) 算法, 将不同层的深度特征分别训练相关滤波器并进行由粗到细的融合。Danelljan 等[8] 提出连续卷积操作跟踪 (continuous convolution operators for visual tracking, CCOT) 算法, 采用多层深度特征, 通过连续空域插值运算, 有效结合不同层次的深度特征训练模型, 从而提高模型的适应性。此外, Danelljan 等[9] 在 CCOT 算法的基础上提出高效卷积操作跟踪 (efficient convolution operators for tracking, ECO) 算法, 从模型、样本集和更新策略 3 个方面进行优化, 有效提升了算法的速度和精度。Qi 等[10] 提出的对冲深度跟踪 (hedged deep tracking, HDT) 算法采用 6 层卷积特征, 通过自适应学习每层预测位置的权重, 将 6 层预测位置进行加权融合, 提高跟踪定位的精度。虽然上述深度相关目标跟踪算法取得了优异的跟踪性能, 但提取特征耗费大量时间, 本文通过将手工特征和轻量化卷积网络提取的深度特征进行融合, 满足实时性的前提下提升算法的鲁棒性。

除了对特征进行改进, 边界效应和滤波器模板退化依然影响着相关滤波算法的性能。边界效应是由样本的周期性假设引起的。为了缓解边界效应, Galoogahi 等[11] 提出了背景感知相关滤波算法, 通过引入裁剪矩阵, 在扩大的搜索区域上进行样本裁剪, 保证样本的质量。Danelljan 等[12] 提出了 SRDCF 算法, 采用固定的空间正则权重惩罚滤波器系数, 一定程度上缓解了边界效应。Feng 等[13] 提出了基于动态显著正则化的相关滤波算法, 利用目标的显著性信息指导空间正则权重在线更新, 由于该算法未建立相邻 2 帧滤波器之间的联系, 在目标发生大幅形变和遮挡的条件下, 滤波器易退化导致跟踪失败。Fu 等[14] 提出了一致性约束相关滤波算法, 在 SRDCF 算法的基础上, 引入响应图中隐藏的时间信息避免模型退化, 但是空间正则权重没有自适应的学习能力, 在遇到目标出视野和相似背景干扰时易跟踪失败。

本文在 SRDCF 算法的基础上, 提出了基于显著感知与一致性约束的相关滤波算法。为了提高目标跟踪算法的实时性, 同时避免速度提升造成精度下降的问题, 将手工特征与轻量化网络提取的深度特征进行融合, 提升物体外观模型的表征能力。利用显著性检测算法构造自适应空间正则权重, 更有效地缓解边界效应。充分利用响应图中隐藏的时间信息构造时间一致性约束项, 同时根据响应图的质量调整一致性约束的强度, 防止滤波器模板退化, 提升算法的鲁棒性与稳健性。

1　空间正则相关滤波算法

相关滤波算法主要分为训练和检测 2 个阶段。训练阶段通过循环矩阵对基础样本进行循环移位获得训练样本集, 不可避免地产生边界效应问题。SRDCF 算法在相关滤波框架下, 通过引入空间正则项来缓解边界效应。目标函数如下:

$$E(\boldsymbol{H})=\frac{1}{2}\left\|\boldsymbol{y}-\sum_{k=1}^{K}\boldsymbol{x}_t^k*\boldsymbol{h}_t^k\right\|_2^2+\frac{\lambda}{2}\sum_{k=1}^{K}\left\|\boldsymbol{w}\cdot\boldsymbol{h}_t^k\right\|_2^2 \tag{1}$$

式中: $\boldsymbol{y}\in\mathbf{R}^{MN\times1}$为理想高斯响应, 作为训练样本的标签函数, M和N分别为图像特征的长和宽; $\boldsymbol{x}_t^k\in\mathbf{R}^{T\times1}(k=1,2,\cdots,K)$为第$t$帧第$k$个通道提取的特征, $T=MN$, K为每帧图像中总的通道数量; $\boldsymbol{h}_t^k\in\mathbf{R}^{T\times1}$为第$t$帧训练的第$k$个通道的滤波器; $\boldsymbol{H}=[\boldsymbol{h}_1,\boldsymbol{h}_2,\cdots,\boldsymbol{h}_k]$; "$*$" 表示卷积运算; λ为正则项参数; 等式右端第 2 项为空间正则项, $\boldsymbol{w}\in\mathbf{R}^{T\times1}$为碗状的空间正则权重, 表达式为

$$w(i,j)=a+b[(i-M/2)/\varpi]^2+b[(j-N/2)/h]^2 \tag{2}$$

其中: $i=1,\cdots,M$, $j=1,\cdots,N$; a和b为固定的预设参数; ϖ和h分别为跟踪目标的宽和高。

SRDCF 算法定义的正则化权重图是不合理的, 忽视了样本信息的重要性, 在视频第 1 帧确定正则化权重图后, 未在后续的视频序列中进行更新。由于目标形状在实际跟踪过程中随着时间的推移频繁发生变化, 通过各像素点与权重图中心距离来定义权重图中的取值大小, 很有可能许多非目标区域距离目标中心较近却未受到惩罚, 降低滤波器的判别能力。同时, 该算法记录了从视频开始到结束的多幅样本信息, 降低算法的跟踪速度, 在目标被遮挡时导致样本信息错误造成误差的累积。

为此, 本文通过多特征融合的方式构造鲁棒的外观模型; 通过构造显著感知参考权重, 引入与目标建立联系的自适应空间正则项, 增强滤波器惩罚系数的可靠性; 引入时间一致性约束项使滤波器学习到相邻帧之间的信息, 降低算法在处理不准确样本时发生过拟合的风险。

2　本文算法

2.1　多特征融合

近期, Wang 等[15] 提出了理解和诊断视觉跟踪系统, 通过实验发现特征提取在目标跟踪过程中发挥着重要作用。常见的目标跟踪算法大多采用单

一特征，不能对目标的外观进行准确的建模。不同于SRDCF[12]和DeepSRDCF[5]算法，前者仅利用手工特征对目标进行表达，后者采用单层卷积特征代替手工特征，跟踪性能得到明显提升，但是跟踪速度缓慢。因此，本文将手工特征和深度特征进行融合，选取HOG特征作为手工特征，该特征可以有效描述目标的轮廓和形状信息，具有较好的几何和光照不变性。为了进一步提高定位的精度和跟踪的鲁棒性，利用预训练深度学习网络提取视觉目标的分层深度特征。文献[5]经过实验发现底层特征保留了更多细粒度的空间特征，具有丰富的纹理信息，适合目标跟踪过程中的精确定位。随着网络深度的增加，特征图的分辨率不断降低，但是中高层特征具有丰富的语义信息，能够更好地应对遮挡、相似背景干扰等挑战，提升跟踪的鲁棒性。由于预训练网络提取的深度特征通道数较多，易造成跟踪速度缓慢，利用文献[16]提出的轻量化网络CF_VGG进行底层和中层特征提取，在不损失跟踪精度的前提下，不仅能够降低特征维数，而且可以大大节省特征提取时间。同时，本文采用串行特征融合策略，具体方式为：先提取31维的HOG特征，再选择CF_VGG的Conv2和Conv3进行底层和中层特征提取，对应的特征维度分别为32维和64维，最终将3种特征进行融合，对应的维度为127维。

2.2 显著感知自适应空间正则化

由于SRDCF算法中的空间正则权重未与目标建立联系，同时不能自适应更新，导致目标在经历快速运动和大幅形变等挑战时易跟踪失败。文献[13]提出了动态显著性目标跟踪算法，将目标的内容信息引入到空间正则化权重图中，提升了基线算法SRDCF的跟踪精度和鲁棒性。由于目标的形状和大小往往随时间不断变化，本文利用显著性检测算法获得目标的显著性检测结果，将目标的变化信息引入到正则化权重图中，构建显著感知自适应空间正则项，实现正则权重时序的动态更新。

考虑到本文算法对实时性的要求，采用文献[17]提出的谱残差算法对初始帧进行显著性检测。

如图1所示，首先裁剪比跟踪框大2.5倍的同心区域作为显著性检测区域，获得初始帧的显著性检测结果$\boldsymbol{S}(o)$。$\boldsymbol{S}(o)$求解如下：将第1帧图像进行傅里叶变换，分别得到幅频响应$\boldsymbol{A}(f)$和相频响应$\boldsymbol{P}(f)$。将$\boldsymbol{A}(f)$取对数与平均幅频响应相减，得到谱残差$\boldsymbol{U}(f)$。将$\boldsymbol{U}(f)$与$\boldsymbol{P}(f)$相加进行傅里叶反变换后，与高斯函数$\boldsymbol{g}(o)$进行卷积运算得到显著性响应$\boldsymbol{S}(o)$。公式如下：

$$\boldsymbol{U}(f)=\lg(\boldsymbol{A}(f))-\boldsymbol{h}_n(f)*\lg(\boldsymbol{A}(f)) \tag{3}$$

$$\boldsymbol{S}(o)=\boldsymbol{g}(o)*F^{-1}[\exp(\boldsymbol{U}(f)+\boldsymbol{P}(f))]^2 \tag{4}$$

式中：f为频率；“$*$”表示卷积运算；F^{-1}表示傅里叶反变换；o表示整幅图像；$\boldsymbol{h}_n(f)$为局部平均滤波器。

$\boldsymbol{S}(o)$与余弦窗相乘，抑制边界框外的错误检测结果，同时将$\boldsymbol{S}(o)$裁剪成与空间正则权重大小相同的显著性图$\boldsymbol{S}'(o)\in\mathbf{R}^{T\times1}$。将$\boldsymbol{S}'(o)$中的响应值通过阈值0.1重映射到正则权重$\boldsymbol{S}_{\mathrm{f}}$。经过实验发现，$\boldsymbol{S}'(o)$中响应强度小于0.1的区域被认为是背景，因此给予较大的惩罚；大于0.1的区域被认为是目标，因此给予较小的惩罚。与另一个预定义的正则化权重图$\boldsymbol{S}_{\mathrm{r}}$进行相交，最终获得初始帧的显著感知参考权重$\boldsymbol{w}_{\mathrm{ref}}$。其中，$\boldsymbol{S}_{\mathrm{r}}$中大于目标跟踪框外的所有区域给予较大的惩罚，跟踪框内的区域给予较小的惩罚。通过相交可以将跟踪框内的背景区域施加精确的惩罚，同时考虑了空间先验和目标的上下文信息。

$$\boldsymbol{w}_{\mathrm{ref}}=\boldsymbol{S}_{\mathrm{f}}\cap\boldsymbol{S}_{\mathrm{r}} \tag{5}$$

式中：

$$\boldsymbol{S}_{\mathrm{f}}=\begin{cases}10^5 & \boldsymbol{S}'(o)<0.1\\ 10^{-3} & \boldsymbol{S}'(o)>0.1\end{cases}$$

$$\boldsymbol{S}_{\mathrm{r}}=\begin{cases}10^5 & o\in\boldsymbol{\Omega}_{\mathrm{non}}\\ 10^{-3} & o\in\boldsymbol{\Omega}_{\mathrm{obj}}\end{cases}$$

式中：$\boldsymbol{\Omega}_{\mathrm{non}}$属于目标跟踪框外的区域；$\boldsymbol{\Omega}_{\mathrm{obj}}$属于目标跟踪框内的区域。

将参考权重加入到目标函数中得到自适应的空间正则权重，公式如下：

$$E_{\mathrm{ASR}}=\frac{\lambda_1}{2}\sum_{k=1}^{K}\left\|\boldsymbol{w}\cdot\boldsymbol{h}_t^k\right\|_2^2+\frac{\lambda_2}{2}\sum_{k=1}^{K}\left\|\boldsymbol{w}-\boldsymbol{w}_{\mathrm{ref}}\right\|_2^2 \tag{6}$$

式中：λ_1和λ_2为空间正则化系数；$\boldsymbol{w}$和$\boldsymbol{w}_{\mathrm{ref}}$分别为自

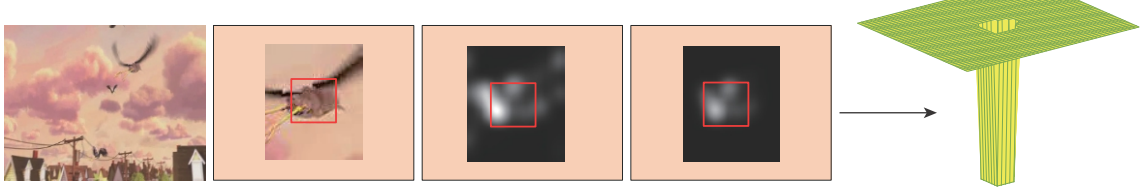
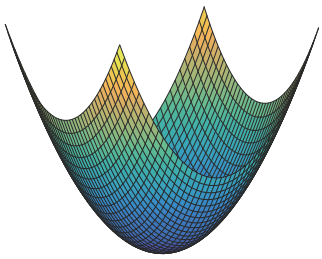


图1 初始帧显著感知参考权重建模

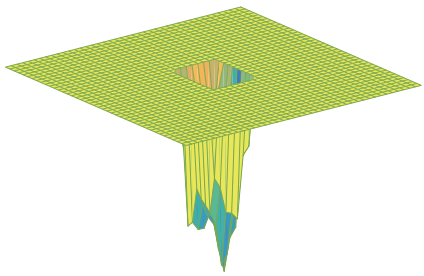
Fig. 1 Initial frame saliency aware reference weight modelling

适应空间正则权重和显著感知参考权重。

如图 2(a)所示, 空间正则化权重呈现负高斯形状, 在跟踪过程中是固定不变的。图 2(b)对应本文求解的自适应空间正则权重, 该权重在跟踪过程中随着目标的变化进行自适应更新, 更加准确地突出目标区域抑制背景区域, 有效缓解了边界效应。



(a) SRDCF算法对应的空间正则权重



(b) 本文算法对应的自适应空间正则权重

图 2　不同目标跟踪算法空间正则权重的可视化结果

Fig. 2　Visualization results of spatial regularization weights of different target tracking algorithms

2.3　动态一致性约束正则化

目标跟踪过程中, 相邻 2 帧图像之间的连续性隐含着很强的时间序列相关性, 已有相关滤波算法利用时间信息进一步提升跟踪器的性能。时空正则相关滤波目标跟踪 (learning spatial-temporal regularized correlation filters for visual tracking, STRCF) 算法[18]在 SRDCF 算法的基础之上引入时间正则项, 以元素方式惩罚滤波器系数的变化, 实现模型的被动更新。文献 [19] 采用 STRCF 算法中的时间正则项使相关滤波器学习到相邻帧之间的信息, 降低模型在处理不准确样本时发生过拟合的风险。然而, STRCF 算法使用前一帧的信息约束相关滤波器的更新, 这种被动更新方案并不准确。当目标被部分或者完全遮挡时, 该算法仍然使用错误的信息更新相关滤波器, 导致错误信息逐渐累积, 最终跟丢目标。

不同于 STRCF 算法中采用的时间正则项, 本文充分利用响应图中隐藏的时间信息, 有效缓解滤波器模板退化。时间序列相关性本质上是由物体位置连续性引起的, 最终体现为时域内响应图之间的连续性。在不发生剧烈外观变化的前提下, 将检测响应$\boldsymbol{R}_t$和目标响应$\boldsymbol{R}'_t$进行卷积运算获得实际一致性响应, 该结果具有很强的时间序列相关性。通过最小化实际一致性响应与理想一致性响应之间的平方误差, 使得相邻 2 帧滤波器之间实现平滑约束, 从而注入丰富的时间信息。由于在经历遮挡、剧烈形变等挑战时目标外观变化的速率是不同的, 通过判断响应图的质量动态调整一致性约束强度, 能够使当前帧训练的滤波器模板充分学习目标的变化信息, 更好地适应外观变化的速率。

首先, 为了研究响应之间的相似性, 检测响应$\boldsymbol{R}_t$和目标响应$\boldsymbol{R}'_t$的公式为

$$\begin{cases} \boldsymbol{R}_t = \sum\limits_{k=1}^{K} \boldsymbol{h}_{t-1}^k * \boldsymbol{z}_t^k \\ \boldsymbol{R}'_t = \sum\limits_{k=1}^{K} \boldsymbol{h}_t^k * \boldsymbol{x}_t^k \end{cases} \tag{7}$$

式中: $\boldsymbol{h}_{t-1}^k$和$\boldsymbol{h}_t^k$分别为第$t-1$帧和第t帧训练的第k个通道的滤波器; $\boldsymbol{z}_t^k$和$\boldsymbol{x}_t^k$分别为第t帧中第k个通道提取的检测样本和训练样本的特征; "$*$" 表示卷积运算。

为充分利用跟踪过程中的时间信息, 将检测响应与目标响应之间进行卷积运算, 评估响应之间的一致性水平, 约束相邻 2 帧滤波器保持一致, 公式如下:

$$\boldsymbol{C}_t = \boldsymbol{R}'_t * \boldsymbol{R}_t\left[\Delta\tau_{m,n}\right] \tag{8}$$

式中: $\boldsymbol{C}_t \in \mathbf{R}^{T\times 1}$为相邻 2 帧的一致性响应; $\left[\Delta\tau_{m,n}\right]$为移位运算符, m和n分别为相邻 2 帧响应图的峰值位置。为了消除运动的影响,通过移位运算符将响应图的峰值进行重合。

理想情况下, 检测响应和目标响应通过构建岭回归方程与理想高斯响应$\boldsymbol{y}$进行约束。考虑到$\boldsymbol{C}_t$的生成过程, 理想一致性响应$\boldsymbol{l}_{\mathrm{f}}$由 2 个高斯响应进行卷积运算生成, 作为实际一致性响应的约束标签, 具体公式为$\boldsymbol{l}_{\mathrm{f}} = \boldsymbol{y} * \boldsymbol{y}$。最终构建的动态一致性约束正则项如下:

$$E_{\mathrm{DCC}} = \frac{\rho}{2}\left\|\boldsymbol{l}_t - \boldsymbol{R}'_t * \boldsymbol{R}_t\left[\Delta\tau_{m,n}\right]\right\|_2^2 \tag{9}$$

式中: ρ为一致性约束正则系数; $\boldsymbol{l}_t = d_t\boldsymbol{l}_{\mathrm{f}}$, d_t为动态正则系数用于调整一致性约束的强度, 公式为

$$d_t = d_{\min} + \frac{\zeta}{\phi}\left(d_{\max} - d_{\min}\right) \tag{10}$$

其中: $d_{\min}$和$d_{\max}$分别为理想高斯响应标签对应的最小和最大幅值; ϕ为归一化系数; ζ[20]指标用于评估响应图的质量, 公式如下:

$$\zeta = \frac{\left|F_{\max} - F_{\min}\right|}{\mathrm{mean}\left(\sum\limits_{i,j}\left(F_{i,j} - F_{\min}\right)^2\right)} \tag{11}$$

式中: $F_{\max}$为响应图的最大值; $F_{\min}$和$F_{i,j}$分别为响应

图的最小值和响应图中第i行、第j列取值。当ζ突然减小时，即目标发生了遮挡或者丢失。因此，引入ζ指标能够正确评估响应图的质量，更好地调节一致性约束强度。

由于相关滤波类的算法属于模板匹配类的算法，当目标发生快速运动或者剧烈形变时，跟踪器易跟丢目标，此时放松一致性约束强度，使滤波器可以从外观变化中进行学习。相反，如果目标外观变化是平滑的，需要加强一致性约束强度。不同于STRCF算法中约束相邻2帧滤波器保持高度一致，本文充分利用响应图的质量构建动态调节因子调整一致性约束强度，更好地应对不同的挑战场景。动态调节过程如图3所示。Dragonbaby序列第20帧和第102帧中目标外观未发生剧烈变化，因此，3个算法均可以准确跟踪目标，动态标签$\boldsymbol{l}_t$的值较大；第62帧和第84帧由于目标发生剧烈运动导致外观改变，此时$\boldsymbol{l}_t$的值较小，降低一致性约束强度，使跟踪器充分学习目标的外观变化，从而准确的跟踪目标。如图4所示，本文算法在该序列的平均中心位置误差是最小的，平均重叠率是最高的。

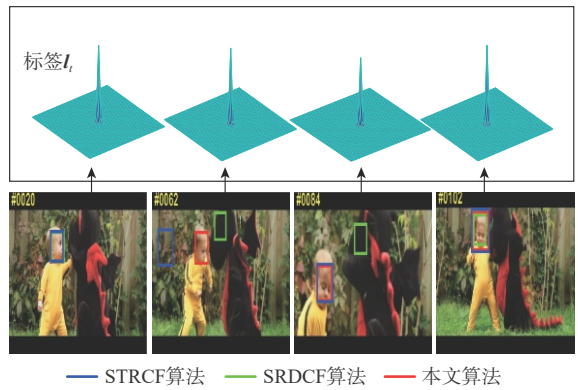
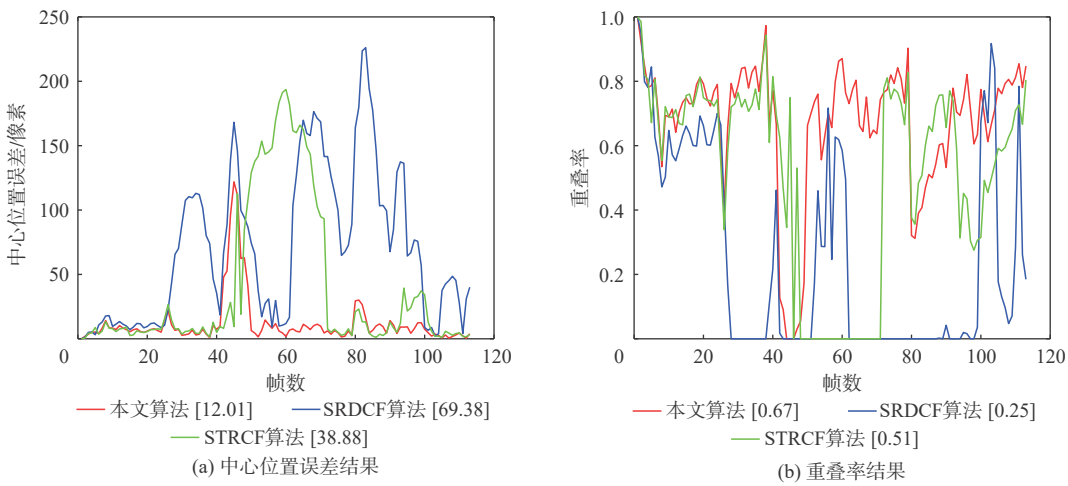


图3 动态标签$\boldsymbol{l}_t$的调整过程

Fig. 3 Adjustment process of dynamic label $\boldsymbol{l}_t$



(a) 中心位置误差结果 (b) 重叠率结果

图4 不同算法逐帧比较的中心位置误差和重叠率

Fig. 4 Frame-by-frame comparison of center location errors and overlap rate of different algorithms

2.4 基于显著感知和动态一致性约束的时空正则模型求解

2.4.1 目标函数模型

本文最终构建的目标函数如下：

$$E(\boldsymbol{H},\boldsymbol{w})=\frac{1}{2}\left\|\boldsymbol{y}-\sum_{k=1}^{K}\boldsymbol{x}_t^k*\boldsymbol{h}_t^k\right\|_2^2+\frac{\lambda_1}{2}\sum_{k=1}^{K}\left\|\boldsymbol{w}\cdot\boldsymbol{h}_t^k\right\|_2^2+\frac{\lambda_2}{2}\left\|\boldsymbol{w}-\boldsymbol{w}_{\mathrm{ref}}\right\|_2^2+\frac{\rho}{2}\left\|\boldsymbol{l}_t-\left(\sum_{k=1}^{K}\boldsymbol{x}_t^k*\boldsymbol{h}_t^k\right)*\boldsymbol{R}_t\left[\Delta\tau_{m,n}\right]\right\|_2^2 \tag{12}$$

式中：第1项为岭回归分类误差项，第2项和第3项为自适应空间正则项，第4项为动态一致性约束正则项。

优化求解是在一维情况下导出的，可以使优化过程更为简便，很容易扩展到二维情况，公式如下：

$$E(\boldsymbol{H},\boldsymbol{w})=\frac{1}{2}\left\|\boldsymbol{y}-\boldsymbol{X}_t\boldsymbol{h}_t\right\|_2^2+\frac{\lambda_1}{2}\left\|\boldsymbol{w}\cdot\boldsymbol{h}_t\right\|_2^2+\frac{\lambda_2}{2}\left\|\boldsymbol{w}-\boldsymbol{w}_{\mathrm{ref}}\right\|_2^2+\frac{\rho}{2}\left\|\boldsymbol{l}_t-\boldsymbol{R}^{\mathrm{T}}\boldsymbol{X}_t\boldsymbol{h}_t\right\|_2^2 \tag{13}$$

式中：$\boldsymbol{y}=[\boldsymbol{y}_1,\boldsymbol{y}_2,\cdots,\boldsymbol{y}_T]$；$\boldsymbol{X}_t=[\boldsymbol{x}_t^1,\boldsymbol{x}_t^2,\cdots,\boldsymbol{x}_t^K]$；$\boldsymbol{h}_t=[\boldsymbol{h}_t^1,\boldsymbol{h}_t^2,\cdots,\boldsymbol{h}_t^K]$；$\boldsymbol{R}^{\mathrm{T}}$表示由检测响应$\boldsymbol{R}_t\left[\Delta\tau_{m,n}\right]$循环移位生成的循环矩阵。

式(13)对应简化的增广拉格朗日公式如下：

$$L(\boldsymbol{H},\boldsymbol{w},\hat{\boldsymbol{G}},\hat{\boldsymbol{S}})=\frac{1}{2}\left\|\hat{\boldsymbol{y}}-\hat{\boldsymbol{X}}_t\hat{\boldsymbol{g}}_t\right\|_2^2+\frac{\lambda_1}{2}\left\|\boldsymbol{w}\cdot\boldsymbol{h}_t\right\|_2^2+\frac{\lambda_2}{2}\left\|\boldsymbol{w}-\boldsymbol{w}_{\mathrm{ref}}\right\|_2^2+\frac{\rho}{2}\left\|\hat{\boldsymbol{l}}_t-\hat{\boldsymbol{X}}_t^T\hat{\boldsymbol{g}}_t\right\|_2^2+\frac{\mu}{2}\left\|\hat{\boldsymbol{g}}_t-\sqrt{T}\boldsymbol{F}\boldsymbol{h}_t+\hat{\boldsymbol{s}}_t\right\|_2^2 \tag{14}$$

式中：$\boldsymbol{F}$为离散傅里叶矩阵；“^”表示物理量的离散傅里叶变换；μ为正则项的步长参数。

子问题$\boldsymbol{H}$的求解：假设给定$\boldsymbol{w}$、$\hat{\boldsymbol{G}}$、$\hat{\boldsymbol{S}}$，求解$\boldsymbol{H}$的

目标函数如下:

$$\arg\min_{\boldsymbol{h}_t}\frac{\lambda_1}{2}\|\boldsymbol{w}\cdot\boldsymbol{h}_t\|_2^2+\frac{\mu}{2}\left\|\hat{\boldsymbol{g}}_t-\sqrt{T}\boldsymbol{F}\boldsymbol{h}_t+\hat{\boldsymbol{s}}_t\right\|_2^2 \tag{15}$$

令 $\boldsymbol{W}=\mathrm{diag}(\boldsymbol{w})\in\mathbf{R}^{T\times T}$为 $\boldsymbol{w}$的对角矩阵, 则有 $\boldsymbol{w}\cdot\boldsymbol{h}_t=\boldsymbol{W}\boldsymbol{h}_t$。式(15)对$\boldsymbol{h}_t$求偏导, 解得

$$\boldsymbol{h}_t=\frac{\mu T(\boldsymbol{s}_t+\boldsymbol{g}_t)}{\lambda_1(\boldsymbol{W}\cdot\boldsymbol{W})+\mu T} \tag{16}$$

子问题$\hat{\boldsymbol{G}}$的求解: 假设给定$\boldsymbol{H}$、$\boldsymbol{w}$、$\hat{\boldsymbol{S}}$, 求解$\hat{\boldsymbol{G}}$的目标函数如下:

$$\arg\min_{\hat{\boldsymbol{g}}_t}\frac{1}{2}\left\|\hat{\boldsymbol{y}}-\hat{\boldsymbol{X}}_t\hat{\boldsymbol{g}}_t\right\|_2^2+\frac{\rho}{2}\left\|\hat{\boldsymbol{l}}_t-\hat{\boldsymbol{X}}_t^r\hat{\boldsymbol{g}}_t\right\|_2^2+\frac{\mu}{2}\left\|\hat{\boldsymbol{g}}_t-\sqrt{T}\boldsymbol{F}\boldsymbol{h}_t+\hat{\boldsymbol{s}}_t\right\|_2^2 \tag{17}$$

最终解得$\boldsymbol{V}_j(\hat{\boldsymbol{g}}_t)$的封闭解为

$$\boldsymbol{V}_j(\hat{\boldsymbol{g}}_t)=\frac{1}{\mu}\left(\boldsymbol{V}_j(\hat{\boldsymbol{x}}_t)\hat{\boldsymbol{y}}_j+\rho\boldsymbol{V}_j(\hat{\boldsymbol{x}}_t^r)\boldsymbol{V}_j(\hat{\boldsymbol{l}}_t)+\mu\boldsymbol{V}_j(\hat{\boldsymbol{h}}_t)-\mu\boldsymbol{V}_j(\hat{\boldsymbol{s}}_t)\right)-\frac{\boldsymbol{V}_j(\hat{\boldsymbol{x}}_t)}{\mu\boldsymbol{B}}\left(\hat{\boldsymbol{s}}_{\boldsymbol{x}_t}\hat{\boldsymbol{y}}_j+\rho\hat{\boldsymbol{s}}_{\boldsymbol{x}_t}\boldsymbol{V}_j(\hat{\boldsymbol{r}}_t)\boldsymbol{V}_j(\hat{\boldsymbol{l}}_t)+\mu\hat{\boldsymbol{s}}_{\boldsymbol{h}_t}-\mu\hat{\boldsymbol{s}}_{\boldsymbol{s}_t}\right) \tag{18}$$

式中: $\hat{\boldsymbol{r}}_t$为移位检测响应$\boldsymbol{R}_t[\Delta\tau_{m,n}]$的离散傅里叶变换。

式(14)和式(18)的求解过程如附录 A 所示。

为了便于表达, 令

$$\begin{cases}\hat{\boldsymbol{s}}_{\boldsymbol{x}_t}=\boldsymbol{V}_j(\hat{\boldsymbol{x}}_t)\boldsymbol{V}_j(\hat{\boldsymbol{x}}_t)^{\mathrm{T}}\\ \hat{\boldsymbol{s}}_{\boldsymbol{h}_t}=\boldsymbol{V}_j(\hat{\boldsymbol{x}}_t)\boldsymbol{V}_j(\hat{\boldsymbol{h}}_t)^{\mathrm{T}}\\ \hat{\boldsymbol{s}}_{\boldsymbol{s}_t}=\boldsymbol{V}_j(\hat{\boldsymbol{x}}_t)\boldsymbol{V}_j(\hat{\boldsymbol{s}}_t)^{\mathrm{T}}\\ \boldsymbol{B}=\dfrac{\mu}{1+\rho\boldsymbol{V}_j(\hat{\boldsymbol{r}}_t)\boldsymbol{V}_j(\hat{\boldsymbol{r}}_t)^{\mathrm{T}}}+\boldsymbol{V}_j(\hat{\boldsymbol{x}}_t)\boldsymbol{V}_j(\hat{\boldsymbol{x}}_t)^{\mathrm{T}}\end{cases} \tag{19}$$

拉格朗日乘子更新为

$$\hat{\boldsymbol{s}}_t^{i+1}=\hat{\boldsymbol{s}}_t^i+\mu\left(\hat{\boldsymbol{g}}_t^{i+1}-\hat{\boldsymbol{h}}_t^{i+1}\right) \tag{20}$$

式中: $\hat{\boldsymbol{g}}_t^{i+1}$和$\hat{\boldsymbol{h}}_t^{i+1}$分别为上述 2 个子问题在第$i$+1次迭代时所获得的解。步进参数$\mu$的更新方式为: $\mu^{(i+1)}=\min\left(\mu_{\max},\delta\mu^{(i)}\right)$, $\delta=10$, $\mu_{\max}=10\,000$。

子问题$\boldsymbol{w}$的求解: 假设给定$\boldsymbol{h}_t$, 采用交替方向乘子法(alternating direction method of multipliers, ADMM)[21]对$\boldsymbol{w}$进行求解。目标函数为

$$\arg\min_{\boldsymbol{w}}\frac{\lambda_1}{2}\|\boldsymbol{w}\cdot\boldsymbol{h}_t\|_2^2+\frac{\lambda_2}{2}\|\boldsymbol{w}-\boldsymbol{w}_{\mathrm{ref}}\|_2^2 \tag{21}$$

引入辅助变量$\boldsymbol{u}=\boldsymbol{w}$, 式(21)对应的增广拉格朗日形式为

$$L(\boldsymbol{u},\boldsymbol{w},\boldsymbol{n})=\frac{\lambda_1}{2}\|\boldsymbol{w}\cdot\boldsymbol{h}_t\|_2^2+\frac{\lambda_2}{2}\|\boldsymbol{u}-\boldsymbol{w}_{\mathrm{ref}}\|_2^2+\frac{\psi}{2}\|\boldsymbol{w}-\boldsymbol{u}+\boldsymbol{n}\|_2^2 \tag{22}$$

式中: ψ为步进参数, 更新方式为: $\psi^{(i+1)}=\min(\psi_{\max},\vartheta\psi^{(i)})$, $\vartheta=10$, $\psi_{\max}=10\,000$; $\boldsymbol{n}$为拉格朗日乘子。

由于该模型仍为可微凸函数, 同样采用 ADMM 分解为 3 个子问题求解:

$$\boldsymbol{w}=\frac{\boldsymbol{u}-\boldsymbol{n}}{\boldsymbol{I}+\dfrac{\lambda_1}{\psi}\boldsymbol{h}_t\cdot\boldsymbol{h}_t} \tag{23}$$

$$\boldsymbol{u}=\frac{\lambda_2\boldsymbol{w}_{\mathrm{ref}}+\psi(\boldsymbol{w}+\boldsymbol{n})}{\lambda_2+\psi} \tag{24}$$

$$\boldsymbol{n}^{(i+1)}=\boldsymbol{n}^{(i)}+\boldsymbol{w}^{(i+1)}-\boldsymbol{u}^{(i+1)} \tag{25}$$

2.4.2　目标定位

目标定位阶段, 将学习到的相关滤波器$\hat{\boldsymbol{h}}_{t-1}$与当前帧输入$\hat{\boldsymbol{z}}_t$进行运算, 得到响应结果为

$$\boldsymbol{R}_t=F^{-1}\sum_{k=1}^{K}\left(\hat{\boldsymbol{z}}_t^{\,k}\cdot\hat{\boldsymbol{h}}_{t-1}^k\right) \tag{26}$$

取响应结果最大值所在的位置对当前帧进行定位。本文采用文献 [22] 中的尺度估计策略, 将目标位置和尺度估计分开求解。通过求解 5 个尺度下的响应最大值快速估计出当前目标的最佳尺度。

2.4.3　模型更新

视觉跟踪过程中, 目标和背景外观模型随着时间的推移会发生相应的变化, 为了更好地适应外观的改变, 目标的外观模型必须及时更新, 防止跟踪漂移。每帧都进行模型更新易导致过拟合, 不仅降低跟踪速度, 同时使训练的外观模型更易受到干扰。因此, 采用每隔 2 帧的稀疏模型更新策略, 在减少算法计算量的同时, 进一步增强算法在不同跟踪场景下的适应性, 表达式如下:

$$\hat{\boldsymbol{x}}_t^{\mathrm{model}}=(1-\eta)\hat{\boldsymbol{x}}_{t-2}^{\mathrm{model}}+\eta\hat{\boldsymbol{x}}_t \tag{27}$$

式中: $\hat{\boldsymbol{x}}_t^{\mathrm{model}}$为当前帧更新之后的外观模型; $\hat{\boldsymbol{x}}_t$为当前帧对应的观测模型; $\hat{\boldsymbol{x}}_{t-2}^{\mathrm{model}}$为前 2 帧对应的外观模型; η为外观模型的在线学习速率。

2.5　算法步骤

基于显著感知与一致性约束的目标跟踪算法流程如图 5 所示, 步骤如下:

步骤 1　如图 1 所示, 在初始帧裁剪目标区域, 采用式 (5) 获得初始帧的显著感知参考权重。根据第 1 帧框选的目标训练滤波器模板。

步骤 2　将跟踪过程分为训练和检测 2 个阶段。在训练阶段, 将第$t-1$帧训练获得的滤波器模板$\boldsymbol{h}_{t-1}$与第t帧的搜索区域进行运算获得检测响应$\boldsymbol{R}_t$。$\boldsymbol{R}_t$中最大值所处的位置作为第t帧的目标中心。以第t帧的目标中心提取特征与待求解的滤波器模板$\boldsymbol{h}_t$进行运算获得目标响应$\boldsymbol{R}_t'$。将$\boldsymbol{R}_t$和$\boldsymbol{R}_t'$进行卷积运算获得实际一致性响应$\boldsymbol{C}_t$。利用检测响应$\boldsymbol{R}_t$计算ζ, 构建动态一致性约束标签$\boldsymbol{l}_t$。通过求解式(12)获得第t帧的最优滤波器模板$\boldsymbol{h}_t$。

步骤 3　在检测阶段, 利用式(26)将第$t+1$帧提取的特征与求得的$\boldsymbol{h}_t$进行运算获得检测响应$\boldsymbol{R}_{t+1}$,

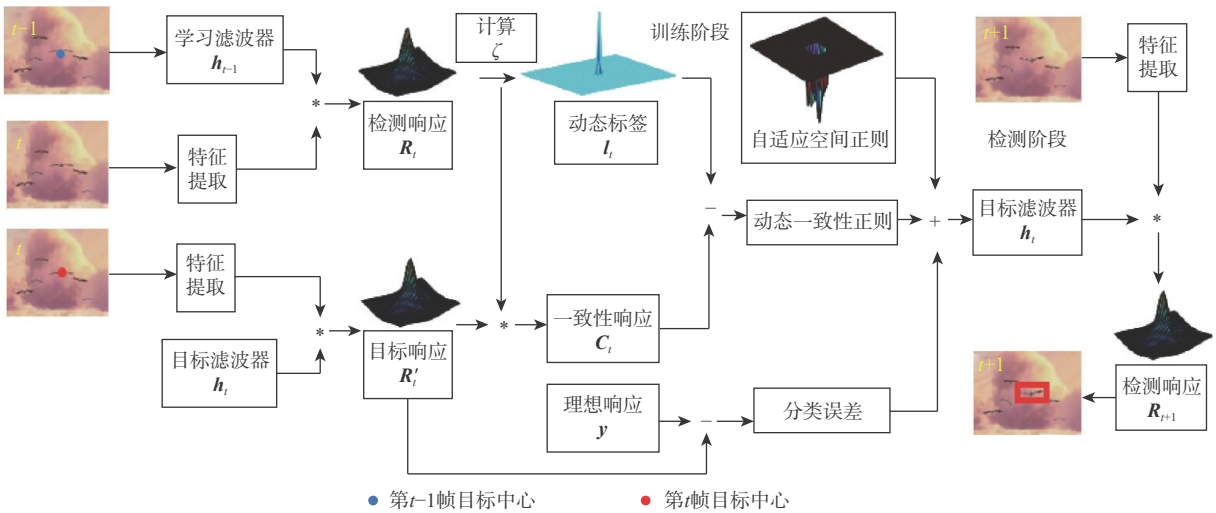


图5 本文算法的整体流程

Fig. 5 Overall flow chart of the proposed algorithm

$\boldsymbol{R}_{t+1}$中最大值所处的位置作为第$t+1$帧的目标中心。

步骤 4 重复步骤 2 和步骤 3，得到每一帧的目标位置，直至完成整个视频序列的跟踪。

3 实验与结果分析

3.1 实验环境与参数设置

本文实验所用的操作系统为 Windows10，所有实验都在 i7-10750H CPU(2.6 GHz)，16 GB RAM，NVIDIA GTX 1650 电脑上进行，编程软件为 MATLAB 2016b。目标定位阶段，利用 CF_VGG 的 Conv2 和 Conv3 分别提取浅层和中层特征，同时结合 HOG 特征对目标的外观进行建模，且每帧图像中总的通道数量 K 为 127。尺度估计阶段，通过构建基于 HOG 特征的 5 个不同尺度的金字塔对目标进行尺度估计，尺度步长为 1.01。空间正则化系数λ_1和λ_2设置为 1.2 和 0.001，动态一致性约束正则系数ρ设置为 0.008，对于动态约束策略，$d_{\min}$和$d_{\max}$分别设置为 0.6 和 1.2，归一化系数ϕ和权重系数φ分别设置为 50 和 100，模型学习速率η设置为 0.018 5。利用 ADMM 算法优化求解相关滤波器和自适应空间正则权重。经过实验发现，ADMM 迭代次数设置为 2 时可以收敛获得最优解。其余参数与 SRDCF 算法[12]一致。

3.2 数据集和评估指标

为了进一步证明本文算法的有效性，在 OTB 2015[23]、TempleColor128[24]、UAV20L[25] 3 个不同的数据集上对目标跟踪算法进行测试和评估。本文使用距离精度 DP、成功率曲线下面积 AUC 及速度 3 种指标进行定量分析。其中，距离精度为实验所得目标位置与实际目标位置间的欧氏距离小于 20 像素的帧数占总帧数的比例；重叠成功率 OSR 为实验所得跟踪框与实际跟踪框的重叠部分面积与二者总面积的比值大于指定阈值的帧数占总帧数的百分比，该阈值取 50%。此阈值取值不同易产生不同的实验结果，因此，本文采用成功率曲线下面积进行算法的评估。跟踪速度的计算方法为视频总帧数除以总时间。

3.3 定量实验

在 3 个不同的数据集上分别对算法进行评估，得到各自的距离精度曲线和 AUC 曲线，并与当前几种先进的目标跟踪算法进行比较，这些算法包括：

1）基于预训练深度特征与相关滤波算法结合的方法 HCF[7]、HDT[10]。

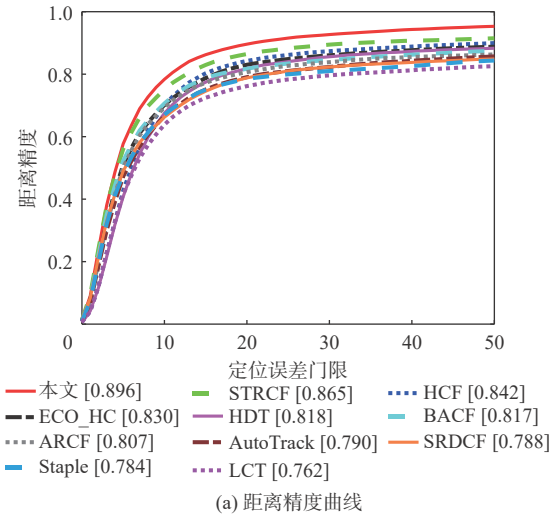
2）基于传统手工特征的相关滤波算法 STRCF[18]、SRDCF[12]、BACF[11]、ECO_HC[9]、LCT[26]、Staple[4]。

3）基于传统手工特征适合于无人机领域的相关滤波算法 ARCF[27]、AutoTrack[28]。

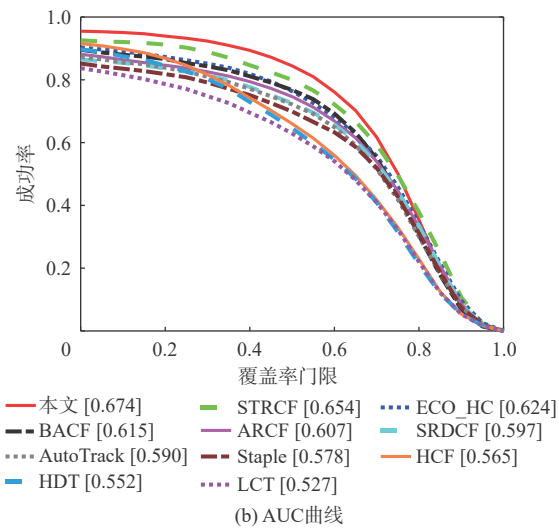
3.3.1 OTB2015 数据集

图 6 为本文算法和对比算法在 OTB2015 数据集上的跟踪精度和 AUC 对比曲线。本文算法对应的 DP 和 AUC 分别为 0.896 和 0.674。相比于基线算法 SRDCF，DP 和 AUC 分别提高了 0.108 和 0.077。与同类型的 STRCF 算法相比，DP 和 AUC 分别提高了 0.031 和 0.02。

与采用深度特征的 HCF 和 HDT 算法相比，本文算法 DP 分别提高了 0.054、0.078；AUC 分别提高了 0.109、0.122。尽管 HCF 算法采用 3 层卷积特征融入到相关滤波器中取得了很好的跟踪性能，但是该算法未使用底层的高分辨率特征，因此在跟踪过程中鲁棒性较好，定位精度略差。同时，HCF 算法没有进行尺度估计，因此对应的 AUC 相对较低。HDT 算法采用 Hedge 算法自适应加权不同尺度估计，因此对应的 AUC 较低。表 1



(a) 距离精度曲线



(b) AUC曲线

图 6　不同算法在 OTB2015 数据集上的距离精度和 AUC 曲线

Fig. 6　Distance precision and AUC curves of different algorithms on OTB2015 dataset

表 1　不同算法在 3 个数据集上的平均跟踪速度对比

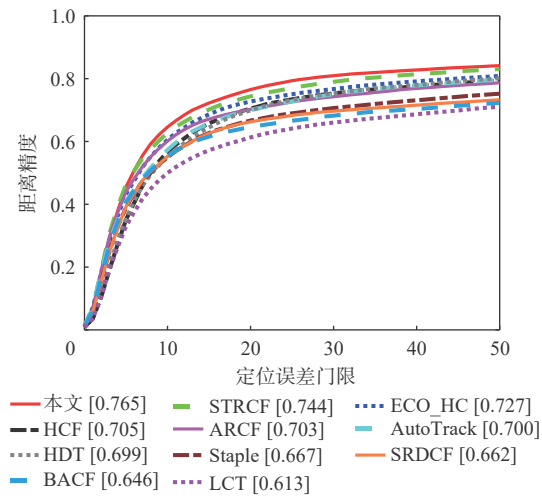
Table 1　Comparison of average tracking speed of different algorithms on 3 datasets　　帧/s

算法	OTB2015	TempleColor128	UAV20L
本文	22.41	20.84	19.49
BACF	35.43	36.75	27.13
STRCF	24.28	21.31	19.17
SRDCF	7.60	8.42	6.04
Staple	82.80	79.16	73.71
ARCF	17.80	16.49	16.08
AutoTrack	28.44	27.52	23.46
ECO_HC	59.39	58.64	63.21
LCT	22.78	25.26	29.54
HCF	1.97	1.92	6.70
HDT	3.69	3.57	2.88

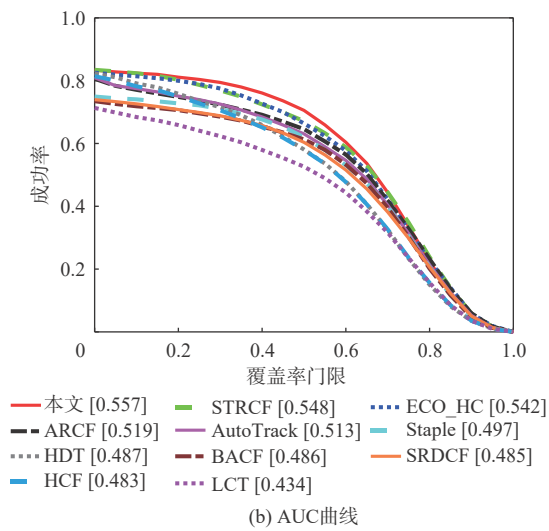
为不同算法在 OTB2015 数据集上的平均跟踪速度对比。本文算法的速度为 22.41 帧/s, 满足实时性的要求。与 HCF、HDT 算法的速度相比, 本文算法在速度上依然有着很大的优势, 主要得益于采用了轻量化的卷积网络进行深度特征提取, 在不损失跟踪精度的前提下, 有效节省了特征提取的时间。

3.3.2　TempleColor128 数据集

图 7 为不同算法在 TempleColor128 数据集上的跟踪精度和 AUC 对比曲线。可以看出, 本文算法在该数据集上获得了最好的 DP 和 AUC, 分别为 0.765 和 0.557。与基线算法 SRDCF 相比, DP 和 AUC 分别提升了 0.103 和 0.072; 同时, 本文算法对应的 DP 和 AUC 得分明显高于基于深度特征的 HCF 和 HDT 算法, 主要得益于将手工特征和轻量化网络提取的深度特征进行融合, 有效解决了彩色视频序列中的跟踪难题。通过构建尺度金字塔从而实现尺度自适应追踪。



(a) 距离精度曲线



(b) AUC曲线

图 7　不同算法在 TempleColor128 数据集上的距离精度和 AUC 曲线

Fig. 7　Distance precision and AUC curves of different algorithms on TempleColor128 dataset

3.3.3 UAV20L 数据集

为了进一步验证本文算法在无人机航拍数据集上的适用性,选取 UAV20L 长时视频序列进行测试。

如图 8 所示,本文算法在该数据集上仍然获得了最好的 DP 和 AUC,分别为 0.621、0.445。相比于基线算法 SRDCF,DP 和 AUC 分别提高了 0.114 和 0.102。与此同时,相比于适用于无人机领域的 ARCF 和 AutoTrack 跟踪器,本文算法在精确率和成功率方面均有很大程度的提升。由于该数据集中的目标尺寸和拍摄角度变化大,视频序列长,需要跟踪器具有较快的学习能力。由图 8 可知,本文算法的跟踪性能明显优于基于深度特征的 HCF 和 HDT 算法,主要得益于本文算法通过构建动态一致性正则项有效缓解了滤波器模板退化,通过将位置和尺度分开求解,从而对目标的尺度估计更为准确。最终证明了本文算法的有效性及优秀的跟踪性能。

为了进一步对比各算法的跟踪性能,对 OTB 2015 数据集的 11 个视频属性进行了定量分析。这些属性包括光照改变 (IV)、尺度变化 (SV)、遮挡 (OC)、形变 (DEF)、运动模糊 (MB)、快速运动 (FM)、平面内旋转 (IPR)、平面外旋转 (OPR)、出视野 (OV)、背景杂波 (BC)、低分辨率 (LR)。

如表 2 和表 3 所示,可以看出,除了低分辨率和平面内旋转场景以外,本文算法在剩余场景中的 DP 排名第一。同时,除了低分辨率场景以外,本文算法在剩余场景中的 AUC 排名第一。与基线算法 SRDCF 相比,本文算法在 11 种不同场景下的 DP 和 AUC 均有一定程度的提升。通过采用多特征融合的方式构建目标的外观模型,因此,在光照改变及背景杂波等场景中跟踪效果较好。通过构建显著感知自适应空间正则项,使得正则权重与目标建立联系,同时进行自适应更新,因此,在应对遮挡及目标发生快速运动等场景时获得了较高的

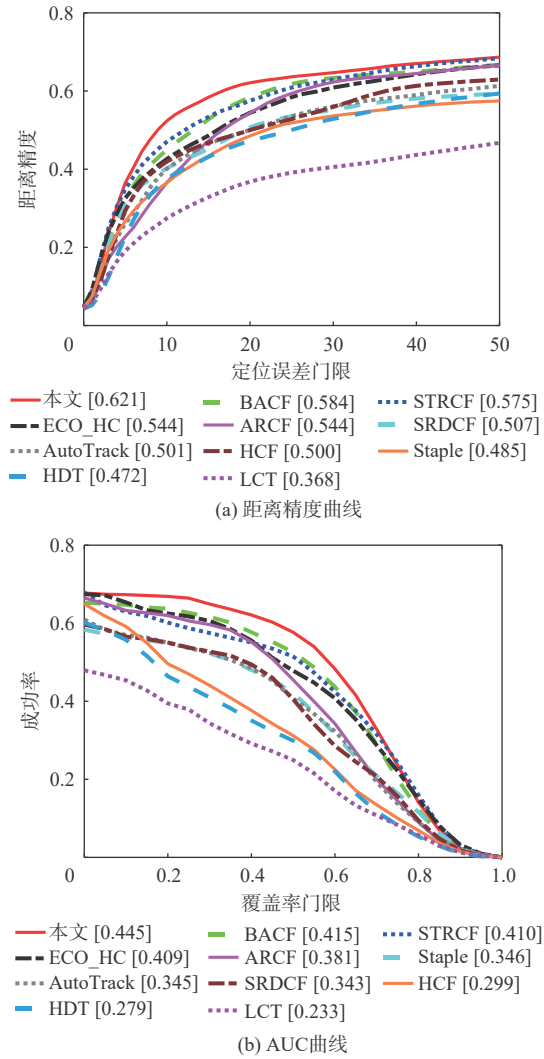


图 8 不同算法在 UAV20L 数据集上的距离精度和 AUC 曲线

Fig. 8 Distance precision and AUC curves of different algorithms on UAV20L dataset

表 2 不同算法在 OTB2015 数据集中 11 个视频属性上的距离精度

Table 2 Precision rate of different algorithms on OTB2015 dataset with 11 video attributes

算法	IV	SV	OC	DEF	MB	FM	IPR	OPR	OV	BC	LR
本文	**0.881**	**0.861**	**0.841**	**0.895**	**0.867**	**0.853**	0.856	**0.882**	**0.870**	**0.897**	0.782
BACF	0.803	0.769	0.730	0.764	0.745	0.790	0.792	0.781	0.756	0.805	0.741
STRCF	0.837	0.840	0.810	0.841	0.826	0.802	0.811	0.850	0.766	0.872	0.737
SRDCF	0.781	0.743	0.727	0.730	0.767	0.769	0.742	0.740	0.603	0.775	0.663
Staple	0.778	0.724	0.724	0.747	0.700	0.710	0.768	0.738	0.668	0.749	0.610
ARCF	0.763	0.770	0.737	0.767	0.757	0.768	0.785	0.769	0.671	0.760	0.749
AutoTrack	0.783	0.742	0.735	0.735	0.735	0.746	0.777	0.766	0.696	0.755	0.773
ECO_HC	0.775	0.792	0.777	0.793	0.770	0.799	0.762	0.801	0.764	0.807	0.847
LCT	0.743	0.678	0.678	0.685	0.670	0.681	0.781	0.746	0.592	0.734	0.537
HCF	0.830	0.798	0.776	0.790	0.804	0.815	**0.864**	0.816	0.677	0.843	0.831
HDT	0.809	0.774	0.744	0.802	0.783	0.779	0.799	0.787	0.616	0.789	**0.849**

注:黑色加粗字体代表当前属性下的最佳跟踪值。

表 3　不同算法在 OTB2015 数据集中 11 个视频属性上的成功率

Table 3　Success rate of different algorithms on OTB2015 dataset with 11 video attributes

算法	IV	SV	OC	DEF	MB	FM	IPR	OPR	OV	BC	LR
本文	**0.678**	**0.639**	**0.647**	**0.664**	**0.673**	**0.651**	**0.624**	**0.652**	**0.638**	**0.682**	0.550
BACF	0.622	0.572	0.565	0.571	0.575	0.599	0.582	0.578	0.548	0.605	0.532
STRCF	0.652	0.631	0.614	0.605	0.652	0.628	0.602	0.626	0.583	0.647	0.538
SRDCF	0.607	0.559	0.554	0.541	0.594	0.597	0.541	0.547	0.461	0.582	0.495
Staple	0.593	0.518	0.541	0.548	0.540	0.540	0.549	0.534	0.476	0.561	0.399
ARCF	0.600	0.561	0.560	0.583	0.605	0.591	0.562	0.559	0.500	0.588	0.512
AutoTrack	0.604	0.542	0.555	0.559	0.585	0.583	0.554	0.555	0.534	0.562	0.540
ECO_HC	0.603	0.592	0.587	0.587	0.604	0.618	0.553	0.587	0.560	0.603	**0.589**
LCT	0.517	0.428	0.476	0.481	0.516	0.507	0.529	0.505	0.446	0.528	0.299
HCF	0.550	0.485	0.533	0.530	0.585	0.570	0.566	0.540	0.474	0.585	0.439
HDT	0.529	0.477	0.522	0.540	0.577	0.555	0.536	0.528	0.453	0.544	0.463

注: 黑色加粗字体代表当前属性下的最佳跟踪值。

DP 和 AUC 得分。通过利用响应图中隐藏的时间信息及根据响应图的质量动态调整一致性约束的强度, 可以更好地应对遮挡及目标的剧烈形变等挑战, 尤其面临目标的大幅形变, 相比于 STRCF 和基线算法, 本文算法在 DP 和 AUC 方面均有很大程度的提升。就 DP 而言, HCF 算法在平面内旋转场景中的 DP 排名第一, 主要归因于该算法采用 3 层深度特征来表达目标, 通过 3 种响应的线性加权得到的最终响应来估计目标位置。尽管 ECO_HC 采用手工特征对目标进行表达, 但是该算法从模型、样本集及更新策略方面进行优化, 因此, 该算法在低分辨率场景中 AUC 排名第一。

在实际跟踪过程中, 目标跟踪算法的实时性依然是一个重要的评价指标。如表 1 所示, 本文算法在 3 个数据集上的跟踪速度分别为 22.41, 20.84, 19.49 帧/s, 满足实时性要求。

3.4　定性实验

将本文算法和 10 种对比算法在 3 个数据集中分别选取 2 个具有挑战性的视频序列进行对比, 包括 Bird1、Girl2、Airport_ce、Surf_ce1、car6、person20。图 9 显示了这些序列的定性比较结果。

1) Bird1 序列。该序列包含了形变、快速运动、出视野 3 种挑战。第 105 帧时目标将要穿越云层, SRDCF 和 ECO_HC 算法跟丢目标。第 203 帧时目标经历长时间的云层遮挡, 仅有本文算法和 Staple 算法可以跟踪目标, 其他算法均跟踪失败。一直到视频结束, 本文算法可以准确跟踪目标, 主要得益于采用了多特征融合的方式, 有效提高了算法对复杂场景的适应性。通过学习获得的自适应参考权重能够对目标和背景区域施加准确的惩罚, 更好地应对目标的快速运动。

2) Girl2 序列。该序列包含 5 种不同的挑战, 最主要的挑战是目标经历了短暂的完全遮挡。第 101 帧时所有算法均能正常跟踪目标。之后由于行人从目标身旁经过造成目标的遮挡, 导致很多算法的跟踪框漂移至行人。一直到第 346 帧时有些算法能够重新跟踪到目标, 但是本文算法的跟踪框更为精准, 主要得益于将位置和尺度估计分开求解, 利用动态一致性约束项有效抑制模型的退化, 提高算法的跟踪准确性。

3) Airport_ce 序列。该序列包含尺度变化、遮挡 2 种挑战。第 81 帧时由于其他行人从目标经过造成局部的遮挡, 此时仅有本文算法可以准确跟踪目标, 其他算法的跟踪框由于算法的局限性导致跟踪失败。一直到第 148 帧, 本文算法的跟踪框始终未发生漂移, 主要原因在于提取的深度特征具有丰富的语义信息, 能够有效应对遮挡等挑战。

4) Surf_ce1 序列。该序列包含 5 种不同的挑战, 最主要的挑战在于目标会经历海浪、飞鸟的遮挡及目标的剧烈形变。第 62 帧时第 1 只飞鸟从冲浪者身旁飞过, SRDCF 算法错把飞鸟当成目标进行跟踪。第 95 帧时第 2 只飞鸟再次从冲浪者身边飞过造成目标的遮挡, 此时 LCT、SRDCF、BACF 算法均发生跟踪漂移。第 209 帧时目标受到海浪的遮挡同时发生剧烈形变, 本文算法依旧保持稳定的跟踪状态。

5) car6 序列。该序列为无人机在高空拍摄的行驶中的汽车, 主要面临尺度变化、出视野等 8 种不同的挑战。第 70 帧时汽车平稳行驶, 所有算法均可以准确跟踪目标。第 1 438 帧时, 汽车经历了短暂的出视野, 此时 SRDCF 算法已经跟丢目标。第 3 272 帧时, 由于视野的改变汽车的尺度逐渐变小, 本文算法将位置和尺度分开求解, 因此可以实现目标尺度的准确估计。第 4 722 帧时, 仅有 BACF 和

图 9 部分视频仿真结果

Fig. 9 Partial results of video simulation

本文算法能够实现准确的跟踪。

6）person20 序列。该序列包含 7 种不同的挑战，在跟踪过程中无人机不断改变拍摄角度，造成了小目标及相似目标影响的复杂情况。第 31 帧之后，由于目标尺度的改变及受到部分遮挡的影响，很多算法尽管能够跟踪目标，但不能实现尺度的自适应估计。本文算法和 BACF 算法能够准确跟踪目标，因此对小目标的跟踪性能更稳定。

3.5 消融实验

为了验证构建动态正则系数d_t的有效性，使用不同的评价指标测试对跟踪结果的影响。$\boldsymbol{I}_t$表示动态正则系数$d_t=1$，即采用静态一致性约束标签，lt_APCE、lt_PSR、lt_PSRM、lt_Fmax 分别表示采用ζ、ϱ、ι、$F_{\max}$指标构建动态一致性约束标签，ϱ[2]用来衡量目标响应图的峰值强度，$\iota=\varrho+\varphi F_{\max}$同时考虑$\varrho$和$F_{\max}$。由图 10可知，采用$\zeta$指标在 OTB2015 数据集上可以获得优异的 DP 和 AUC，更有效地评估响应图的质量，从而动态调节一致性约束强度。

为了验证本文算法中各个组分的有效性，图 11给出了对比验证实验。其中，Baseline 代表基线算法 SRDCF；Baseline+MF 代表在基线算法的基础上采用多特征融合的方式构建外观模型；Baseline+MF+ASR 代表在 Baseline+MF 的基础上添加显著感知自

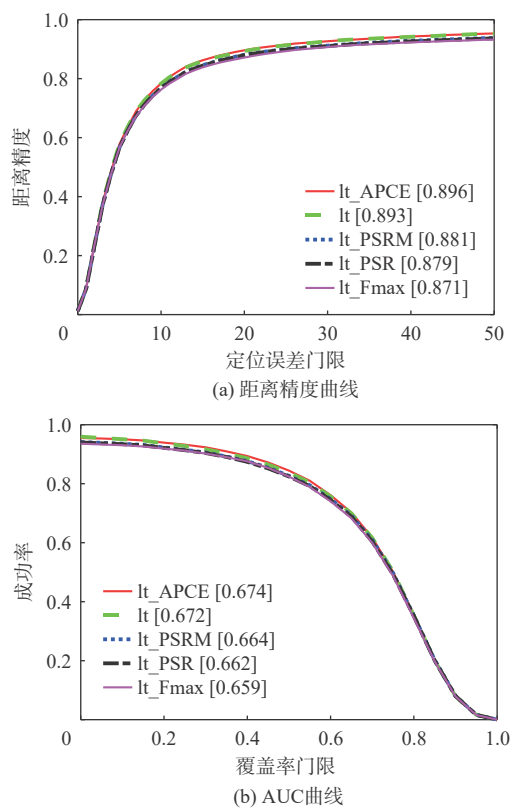


图 10 OTB2015 数据集上不同评价指标的消融实验结果

Fig. 10 Results of ablation experiments with different evalution indicators on OTB2015 dataset

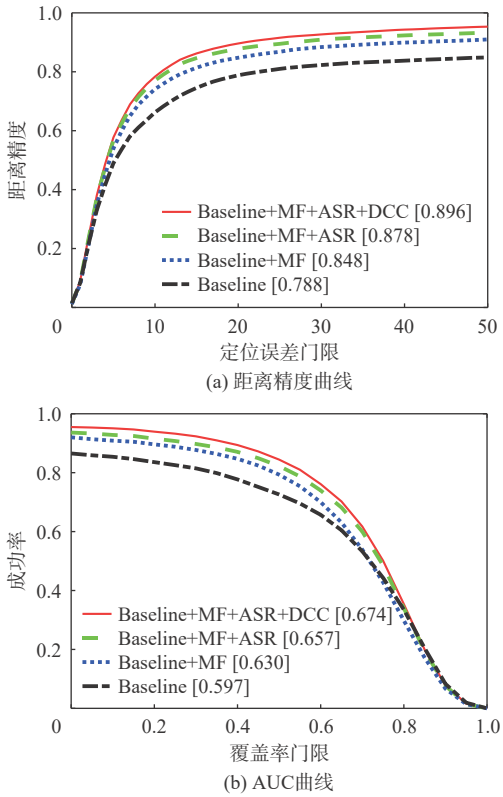


图 11　OTB2015 数据集上不同组分的消融实验结果

Fig. 11　Results of ablation experiments of different components on OTB2015 dataset

适应空间正则项; Baseline+MF+ASR+DCC 代表本文最终所提算法。可以看出, 在 OTB2015 数据集上, Baseline+MF 算法、Baseline+MF+ASR 算法及 Baseline+MF+ASR+DCC 算法相比于基线算法在 DP 方面分别提高 0.006、0.009、0.108; 在 AUC 方面分别提高 0.033、0.06、0.077。充分说明本文算法的各个组分对性能提升有着较大贡献, 同时说明本文算法的有效性和优越性。

4　结　论

本文针对空间正则算法存在的问题及为了更好地应对不同场景下的跟踪挑战提出了相应的改进算法。针对单一特征存在的缺陷, 将手工特征和深度特征进行融合, 在 3 个不同数据集进行测试时能够有效解决彩色视频序列中的跟踪难题, 同时满足实时性要求; 利用显著性检测算法构造自适应空间正则权重有效应对目标的快速运动和剧烈形变等挑战, 在 OTB2015 数据集上获得 0.878 的 DP 和 0.657 的 AUC; 构造动态一致性正则项实现模型的被动更新, 在 OTB2015 数据集上获得 0.896 的 DP 和 0.674 的 AUC。由于本文算法为不同特征分配相同权重, 该方式不利于充分发挥不同特征的优势; 同时经过实验发现一致性约束正则系数的选取会对最终结果产生一定的影响。因此, 未来工作将进一步探索自适应多特征融合及自适应调节正则化参数。

参考文献（References）

[1] 孟晓燕, 段建民. 基于相关滤波的目标跟踪算法研究综述[J]. 北京工业大学学报, 2020, 46(12): 83-106.
MENG X Y, DUAN J M. Advances in correlation filter-based object tracking algorithms: A review[J]. Journal of Beijing University of Technology, 2020, 46(12): 83-106(in Chinese).

[2] BOLME D S, BEVERIDGE J R, DRAPER B A, et al. Visual object tracking using adaptive correlation filters[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2010: 2544-2550.

[3] HENRIQUES J F, RUI C, MARTINS P, et al. High-speed tracking with kernelized correlation filters[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3): 583-596.

[4] BERTINETTO L, VALMADRE J, GOLODETZ S, et al. Staple: Complementary learners for real-time tracking[C]//Proceedings of the International Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 1401-1409.

[5] DANELLJAN M, HAGER G, KHAN F S, et al. Convolutional features for correlation filter based visual tracking[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2015: 58-66.

[6] 刘芳, 孙亚楠, 王洪娟, 等. 基于残差学习的自适应无人机目标跟踪算法[J]. 北京航空航天大学学报, 2020, 46(10): 1874-1882.
LIU F, SUN Y N, WANG H J, et al. Adaptive UAV target tracking algorithm based on residual learning[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(10): 1874-1882 (in Chinese).

[7] MA C, HUANG J B, YANG X, et al. Hierarchical convolutional features for visual tracking[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2015: 3074-3082.

[8] DANELLJAN M, ROBINSO A, KHAN F S, et al. Beyond correlational filters: Learning continuous convolution operators for visual tracking[C]//Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2016: 472-488.

[9] DANELLJAN M, BHAT G, KHAN F S, et al. ECO: Efficient convolution operators for tracking[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 6931-6939.

[10] QI Y, ZHANG S P, QIN L, et al. Hedged deep tracking[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 4303-4311.

[11] GALOOGAHI H K, FAGG A, LUCEY S. Learning background-aware correlation filters for visualtracking[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 1144-1152.

[12] DANELLJAN M, HAGER G, KHAN F S, et al. Learning spatially regularized correlation filters for visual tracking[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2015: 4310-4318.

[13] FENG W, HAN R Z, GUO Q, et al. Dynamic saliency-aware regularization for correlation filter-based object tracking[J]. IEEE Transactions on Image Processing, 2019, 28(7): 3232-3245.

[14] FU C H, YANG X X, LI F, et al. Learning consistency pursued correlation filters for real-time UAV tracking[C]//Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2020: 1-8.

[15] WANG N Y, SHI J P, YEUNG D Y, et al. Understanding and diagnosing visual tracking systems[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2015: 3101-3109.

[16] WANG N, ZHOU W G, SONG Y B, et al. Real-time correlation tracking via joint model compression and transfer[J]. IEEE Transactions on Image Processing, 2020, 29: 6123-6135.

[17] HOU X D, ZHANG L Q. Saliency detection: A spectral residual approach[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2007: 1-8.

[18] LI F, TIAN C, ZUO W M, et al. Learning spatial-temporal regularized correlation filters for visual tracking[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 4904-4913.

[19] 胡昭华, 韩庆, 李奇. 基于时间感知和自适应空间正则化的相关滤波跟踪算法[J]. 光学学报, 2020, 40(3): 0315003.
HU Z H, HAN Q, LI Q. Correlation filter tracking algorithm based on temporal awareness and adaptive spatial regularization[J]. Acta Optica Sinica, 2020, 40(3): 0315003(in Chinese).

[20] WANG M M, LIU Y, HUANG Z Y. Large margin object tracking with circulant feature maps[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 4800-4808.

[21] BOYD S, PARIKH N, CHU E, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers[J]. Foundations and Trends in Machine Leaning, 2010, 3(1): 1-122.

[22] DAI K N, WANG D, LU H C, et al. Visual tracking via adaptive spatially-regularized correlation filters[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2019: 4670-4679.

[23] WU Y, YANG J L, MING H. Object tracking benchmark[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1834-1848.

[24] LIANG P P, BLASCH E, LING H B. Encoding color information for visual tracking: Algorithms and benchmark[J]. IEEE Transactions on Image Processing, 2015, 24(12): 5630-5644.

[25] MUELLER M, SMITH N, GHANEM B. A benchmark and simulator for UAV tracking[C]//Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2016: 445-461.

[26] MA C, YANG X K, ZAHNG C Y, et al. Long-term correlation tracking[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2015: 5388-5396.

[27] HUANG Z Y, FU C H, LI Y M, et al. Learning aberrance repressed correlation filters for real-time UAV tracking[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2019: 2891-2900.

[28] LI Y M, FU C H, DING F Q, et al. AutoTrack: Towards high-performance visual tracking for UAV with automatic spatio-temporal regularization[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 11920-11929.

附录 A

根据帕塞瓦尔定理，引入辅助变量$\hat{\boldsymbol{g}}_t = \sqrt{T}\boldsymbol{F}\boldsymbol{h}_t$，$\hat{\boldsymbol{g}}_t \in \mathbf{C}^{KT\times 1}$。将式(13)从时域转换到频域：

$$E(\boldsymbol{H},\boldsymbol{w},\hat{\boldsymbol{G}}) = \frac{1}{2}\left\|\hat{\boldsymbol{y}} - \hat{\boldsymbol{X}}_t\hat{\boldsymbol{g}}_t\right\|_2^2 + \frac{\lambda_1}{2}\|\boldsymbol{w}\cdot\boldsymbol{h}_t\|_2^2 + \frac{\lambda_2}{2}\|\boldsymbol{w}-\boldsymbol{w}_{\text{ref}}\|_2^2 + \frac{\rho}{2}\left\|\hat{\boldsymbol{l}}_t - \hat{\boldsymbol{X}}_t^r\hat{\boldsymbol{g}}_t\right\|_2^2 \tag{A1}$$

式中：$\hat{\boldsymbol{X}}_t^r = \left[\text{diag}\left(\hat{\boldsymbol{r}}_t^1\cdot\hat{\boldsymbol{x}}_t^1\right),\cdots,\text{diag}\left(\hat{\boldsymbol{r}}_t^K\cdot\hat{\boldsymbol{x}}_t^K\right)\right]$；$\hat{\boldsymbol{y}} = \sqrt{T}\boldsymbol{F}\boldsymbol{y}$；$\hat{\boldsymbol{X}}_t = \sqrt{T}\boldsymbol{F}\boldsymbol{X}_t$；$\hat{\boldsymbol{l}}_t = \sqrt{T}\boldsymbol{F}\boldsymbol{l}_t$。

由于式(A1)是可微凸函数，采用ADMM来求解全局最优解，降低计算复杂度，加快算法的跟踪速度。式(A1)对应的增广拉格朗日形式为

$$L(\boldsymbol{H},\boldsymbol{w},\hat{\boldsymbol{G}},\hat{\boldsymbol{\xi}}) = E(\boldsymbol{H},\boldsymbol{w},\hat{\boldsymbol{G}}) + \hat{\boldsymbol{\xi}}_t^{\mathrm{T}}\left(\hat{\boldsymbol{g}}_t - \sqrt{T}\boldsymbol{F}\boldsymbol{h}_t\right) + \frac{\mu}{2}\left\|\hat{\boldsymbol{g}}_t - \sqrt{T}\boldsymbol{F}\boldsymbol{h}_t\right\|_2^2 \tag{A2}$$

式中：$\hat{\boldsymbol{\xi}}_t \in \mathbf{C}^{KT\times 1}$为拉格朗日乘子的傅里叶变换。引入$\boldsymbol{s}_t = \boldsymbol{\xi}_t/\mu$，式(A2)化简为式(14)。

直接对式(17)进行优化求解很困难，由于预测标签$\hat{\boldsymbol{y}}$的第j个元素仅取决于所有通道中$\hat{\boldsymbol{X}}_t$和$\hat{\boldsymbol{g}}_t$的第j个元素，式(17)可以分解成$M\times N$个子问题，每个子问题对应的目标函数为

$$\arg\min_{\boldsymbol{V}_j(\hat{\boldsymbol{g}}_t)} \frac{1}{2}\left\|\hat{\boldsymbol{y}}_j - \boldsymbol{V}_j^{\mathrm{T}}(\hat{\boldsymbol{x}}_t)\boldsymbol{V}_j(\hat{\boldsymbol{g}}_t)\right\|_2^2 + \frac{\rho}{2}\left\|\boldsymbol{V}_j(\hat{\boldsymbol{l}}_t) - \boldsymbol{V}_j^{\mathrm{T}}(\hat{\boldsymbol{x}}_t^r)\boldsymbol{V}_j(\hat{\boldsymbol{g}}_t)\right\|_2^2 + \frac{\mu}{2}\left\|\boldsymbol{V}_j(\hat{\boldsymbol{g}}_t) - \boldsymbol{V}_j(\hat{\boldsymbol{h}}_t) + \boldsymbol{V}_j(\hat{\boldsymbol{s}}_t)\right\|_2^2 \tag{A3}$$

式中：$\boldsymbol{V}_j(\hat{\boldsymbol{g}}_t) \in \mathbf{R}^{K\times 1}$表示对$K$个通道的$j$点处取值进行向量化。式(A3)对$\boldsymbol{V}_j(\hat{\boldsymbol{g}}_t)$求偏导，解得$\boldsymbol{V}_j(\hat{\boldsymbol{g}}_t)$的封闭解为

$$\boldsymbol{V}_j(\hat{\boldsymbol{g}}_t) = \left((1+\rho\boldsymbol{V}_j(\hat{\boldsymbol{r}}_t)\boldsymbol{V}_j(\hat{\boldsymbol{r}}_t)^{\mathrm{T}})\boldsymbol{V}_j(\hat{\boldsymbol{x}}_t)\boldsymbol{V}_j(\hat{\boldsymbol{x}}_t)^{\mathrm{T}} + \mu\right)^{-1}\boldsymbol{Q} \tag{A4}$$

且$\boldsymbol{Q} = \boldsymbol{V}_j(\hat{\boldsymbol{x}}_t)\hat{\boldsymbol{y}}_j + \rho\boldsymbol{V}_j(\hat{\boldsymbol{x}}_t^r)\boldsymbol{V}_j(\hat{\boldsymbol{l}}_t) + \mu\boldsymbol{V}_j(\hat{\boldsymbol{h}}_t) - \mu\boldsymbol{V}_j(\hat{\boldsymbol{s}}_t)$。

由于式(A4)中求逆运算量大，导致跟踪效率变差。因此，通过谢尔曼-莫里森公式进行求解，提高跟踪的效率。最终得到简化的式(18)。

Target tracking algorithm based on saliency awareness and consistency constraint

GUO Qiang[1, 2, *], WU Tianhao[1, 2], XU Wei[1, 2], CHORNOGOR Leonid[1, 3]

(1. College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China;
2. Key Labotatory of Advanced Marine Communication and Information Technology, Ministry of Industry and Information Technology, Harbin 150001, China;
3. Schoolof Radio Physics, Kharkiv V.N. Karazin National University, Kharkiv 61166, Ukraine)

Abstract: Aimed at the problems that the spatially regularized discriminative correlation filtering (SRDCF) algorithm with fixed regularization weight and model degradation, a correlation filtering tracking algorithm based on saliency awareness and consistency constraint was proposed. Firstly, the histogram of the oriented gradient feature, shallow feature, and the middle feature was extracted to improve the expression ability of the appearance model. Secondly, the regularization weight between the two adjacent frames was associated after the saliency detection algorithm determined the saliency-awareness reference weight of the initial frame.Furthermore, to prevent the degradation of the filter model,the difference between the practical and the scheduled ideal consistency map was minimized and the consistency level was constrained. In addition, a dynamic constraint strategy was proposed to further improve the adaptability of the tracker in complex scenarios. The algorithm is tested on the public OTB2015, TempleColor128, and UAV20L benchmarks. Experimental results show that compared with SRDCF, the proposed algorithm improves the accuracy by 0.108 and the success rate by 0.077 on OTB2015, with a speed of 22.41 frames per second, and has a good real-time effect.

Keywords: machine vision; target tracking; correlation filtering; saliency awareness; consistency constraint

Received: 2021-11-16; **Accepted**: 2022-03-25; **Published Online**: 2022-05-17 17:06

URL: kns.cnki.net/kcms/detail/11.2625.V.20220517.1418.003.html

Foundation items: National Key R & D Program of China (2018YFE0206500); National Natural Science Foundation of China (62071140); International Science & Technology Cooperation Program of China (2015DFR10220)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0648

10 cm 口径发散磁场离子推力器放电模型

孙明明[1,*]，耿海[1]，胡竟[1]，杨福全[1]，龙建飞[2,3]

(1. 兰州空间技术物理研究所 真空技术与物理重点实验室，兰州 730000；2. 国科大杭州高等研究院基础物理与数学科学学院，杭州 310024；3. 南华大学 核科学技术学院，衡阳 421001)

摘 要：为研究 10 cm 口径发散磁场离子推力器内部的放电过程并对后续工程改进提供参考，采用 COMSOL 多物理场耦合软件建立推力器放电模型，获得关键放电参数，并根据试验结果进行验证。模拟结果表明：放电室内部上游磁极和下游磁极之间形成具有强烈发散特性的磁场，并在正交电场的影响下，使电子发生以磁力线为导向中心的霍尔漂移运动；放电室内部气体压强分布均匀且基本在 0.1～0.11 Pa 范围内，大部分区域的中性原子密度约为 1.5×10^{19} m^{-3}，流体速度在 0.2～0.9 m/s 的范围内且呈现明显的黏滞流特性；电子密度峰值出现在阴极出口区域，约为 8.57×10^{18} m^{-3}，阳极壁面附近及栅极上游区域的等离子密度约为 6.8×10^{17} m^{-3}。试验结果显示：采用 E×B 探针测量得到双核离子占总束流离子比为 14.1%，根据 COMSOL 计算值得到的 0.353 mA 束流理论值与 0.323 mA 的束流实测值比对误差为 9%，误差主要来自于仿真条件设置及试验测量。研究结果可为离子推力器工程化改进提供快速的放电参数分析及优化设计参考。

关 键 词：发散磁场；离子推力器；放电模型；等离子体密度；电子温度

中图分类号：V439.4

文献标志码：A **文章编号**：1001-5965(2023)09-2258-09

发散磁场离子推力器即 Kaufman 型离子推力器，是目前国际上主流的离子推力器类型之一，其特点是阴极磁极与放电室上游磁极之间形成具有强发散特征的轴向磁场，以保护位于放电室壁面附近的圆柱筒形阳极[1-3]。发散磁场离子推力器放电室内部的放电区域可分为耦合放电区(阴极出口平面、阴极极靴、阴极挡板所形成的区域)和主放电区(放电室、阴极挡板、阳极和栅极所构成的区域)，对于其放电过程来说，电子向阳极壁面的运动过程是由受轴向磁场梯度影响的迁移-扩散过程所决定的。基于上述放电特性，发散磁场离子推力器可通过调节推进剂的供给量和阳极电流，并通过闭环调节电磁铁电流来调整轴向磁场的发散特性即梯度，实现通过影响放电室内部等离子体密度，并进一步连续控制推力产生大小的目的。

不同于环形会切(ring-cusp)磁场离子推力器(通过多组环形永磁体构成沿阳极壁面的磁场等势面，约束电子使其最大化地参与放电，典型如美国 XIPS 系列[4]和兰州空间技术物理研究所的 LIPS 系列离子推力器[5])的固定工作点设计(常应用于轨道保持或轨道转移)、推力调节范围窄等特点，发散磁场离子推力器具有其独特优势，即推力连续可调、高精度及推力闭环控制，主要用于对产生推力具有连续需求的航天器。Milder[6]、Goebel[7]、Brophy[8]等在对 SERT-Ⅱ发散磁场离子推力器的研究过程中建立了放电室 0-D 计算模型，通过计算认为对于发散磁场构型，由于磁力线与阳极不相交，且当原初电子能量过高时将不参与 Bohm 扩散过程。电子

收稿日期：2021-10-29；**录用日期**：2022-01-07；**网络出版时间**：2022-02-14 11:38

网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220214.1108.001.html

基金项目：国家自然科学基金 (61901202)；中国博士后科学基金 (2019M653906)；湖南省教育厅重点项目 (19A440)；湖南省自然科学基金 (2021JJ30564)；南华大学博士启动基金 (190XQD013)

* **通信作者**. E-mail：smmhappy@163.com

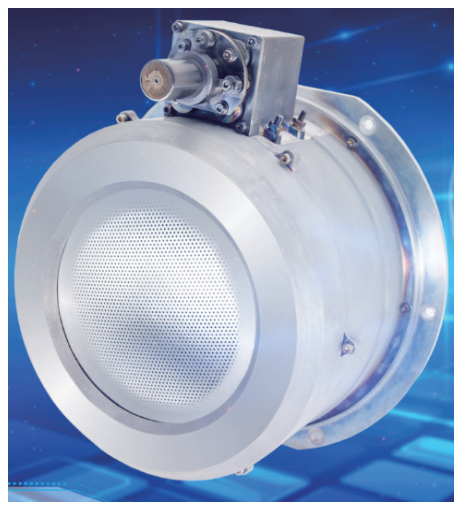
引用格式：孙明明，耿海，胡竟，等. 10 cm 口径发散磁场离子推力器放电模型 [J]. 北京航空航天大学学报，2023，49（9）：2258-2266.
SUN M M，GENG H，HU J，et al. Discharge model of divergent magnetic field ion thruster of 10 cm diameter[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（9）：2258-2266（in Chinese）.

首先需通过粒子间碰撞以穿过轴向磁场并沉积至阳极, 电子温度在 1～10 eV 的范围内, 且越靠近栅极区域越高。其次, 穿过磁场后的等离子体仍受双极扩散效应影响, 若轴向磁场强度高于 50 Gs(1 Gs=10^{-4} T), 则轴向电场影响几乎可忽略, 此时的离子损失率仅为 Bohm 电流的 1/10, 损失至阳极的离子电流可忽略。最后, 由于发散磁场离子推力器放电室内设计有阴极挡板, 从而迫使耦合放电区内的原初电子偏离放电室轴向, 导致放电室轴向的等离子体密度具有显著梯度, 且耦合放电区内的等离子体密度最高, 达到了 10^{17} m^{-3} 量级。James[9]、Milligan[10]等在 T5 发散磁场离子推力器(口径为 10 cm, 可产生 1～21 mN 的连续可调推力)的研制过程中发现, 阴极磁极和阴极挡板构成的挡板通道对放电室性能有显著影响, 挡板通道尺寸可制约进入主放电区的原初电子的能量、密度分布, 进而影响离子推力器的放电性能, 通过试验认为 10 cm 离子推力器的挡板直径应在 20～40 mm 范围内, 基于 0-D 模型计算得到耦合放电区的平均电子密度[10] 约为 3.0×10^{17} m^{-3}, 通过磁场分析获得的挡板通道内的平均磁感应强度约为 150 Gs。Ozaki 等[11-12] 对 ETS-Ⅷ卫星搭载的 12 cm 口径发散磁场离子推力器进行的放电优化研究中发现, 当耦合放电区的磁场强度高于 50 Gs 时, 由于耦合放电区的阴极电势降低, 使得放电损耗显著降低, 放电室内的工质利用率可提升至 90%以上, 相比环形会切磁场离子推力器, 放电电压更低, 可降低至约 30 V, 而较低的放电电压使得等离子体电势与阳极电势呈现明显的反比特性。2019 年, 中国启动了超低轨道监测卫星的研制任务, 为了补偿超低轨道(高度为 200～350 km)存在的大气阻尼, 选择了兰州空间技术物理研究所研制的 10 cm 口径发散磁场离子推力器[13-14], 该推力器的推力调节范围为 1～25 mN, 推力分辨率为±10 μN, 在轨主工况点为 20 mN工况点。根据国外研究结论来看[6-8], 发散磁场离子推力器的工作性能(推力、比冲及效率)受气体电离过程、电离效率和磁场构型的影响相比环形会切磁场离子推力器更大, 并且气体放电稳定性和等离子体分布均匀性是发散磁场离子推力器实现推力宽范围、高精度、高分辨率调节的基础。目前, 国内大部分研究针对环形会切磁场离子推力器, 且对于发散磁场离子推力器的放电性能研究较少, 而国外的相关研究年代较早且主要采用 0-D 放电模型, 结果仅能反映放电室内等离子体参数的大致范围和变化趋势, 无法获取较为精确的放电参数及分布特性。因此, 建立起国内发散磁场离子推力器的放电模型对推力器设计改进具有重要的理论基础和工程指引价值。

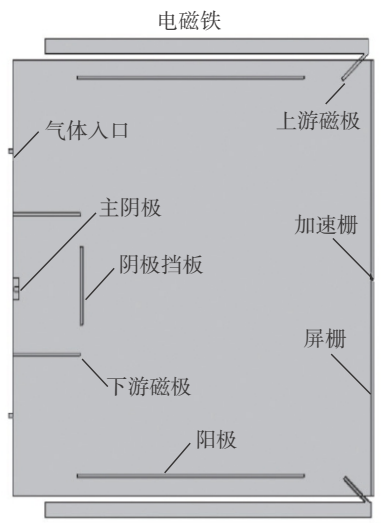
本文采用 COMSOL 多物理场耦合软件对国产 10 cm 口径发散磁场离子推力器建立推力器放电模型, 针对 20 mN 的主工况点进行放电过程仿真以获得关键放电参数, 对现有放电室设计和放电性能进行评价, 并基于试验参数进行模型的验证及误差分析。

1　放电室结构及工作参数

10 cm 口径发散磁场推力器外形如图 1(a) 所示, 其中, 离子推力器主体结构为圆柱形, 由安装环隔成前后 2 段, 前段为放电室结构, 后段用于布置供气管路及供电电缆。图 1(b) 为离子推力器放电室结构。放电室由呈 60°均布的 6 根螺线管电磁铁组件固定, 起到中心承力的作用, 以提高推力器的结构刚度和稳定性。阴极位于放电室几何中心并提供放电所需初级电子, 在发散磁场构型下, 为使



(a) 离子推力器外形



(b) 放电室结构

图 1　10 cm 口径发散磁场离子推力器及放电室结构示意图

Fig. 1　Schematic diagram of structure of 10 cm diameter divergent magnetic field ion thruster and discharge chamber

阴极发射的初级电子受到磁场强约束并向最佳放电区域扩散，以提高电离率和改善束流均匀性，参考英国T5离子推力器设计[10]，在下游磁极端面增加阴极挡板，以控制阴极发射电子的分布（阴极挡板和阴极触持极为同一电位，几乎不会造成电子损失[15]）。阳极为筒状结构，且距离放电室壁面约7 mm，如图1(b)所示，上游磁极为电磁铁的延伸段，沿磁极方向并深入放电室约10 mm，且与放电室壁面呈40°夹角。离子光学系统采用双栅极离子引出系统，且栅极为凹面结构。对于电离过程来说，凹面栅极带来的影响主要是电场等势线分布，由于凹面栅极的拱高仅为5 mm，栅极直径为105 mm，从几何建模角度考虑可忽略凹面影响。经仿真验证，凹面栅极和平面栅极的电势分布差别较小，为便于建模和实现较快的计算收敛，本文将栅极模型建为平面结构。

离子推力器在20 mN主工况点的工作参数如表1所示，其中，屏栅电势为放电室电势最低处，上游磁极、下游磁极与阴极挡板的电势均与屏栅电势相同。阴极触持极和阳极电势相对屏栅电势分别高12 V和35 V。表1所列参数将作为仿真过程的边界条件加载至模型，以计算并获得离子推力器的关键放电参数。

表1 10 cm口径发散磁场离子推力器20 mN主工况点的工作参数

Table 1 Parameters of 10 cm diameter divergent magnetic field ion thruster with 20 mN main work point

参数	数值
阴极流量/(mg·s^{-1})	0.08
阳极流量/(mg·s^{-1})	0.5
中和器流量/(mg·s^{-1})	0.08
屏栅电势/V	1 150
加速栅电势/V	−150
触持电势/V	1 162
励磁电流/A	0.56
励磁电势/V	1 150
挡板电势/V	1 150
上磁极电势/V	1 150
下磁极电势/V	1 150
阳极电势/V	1 185

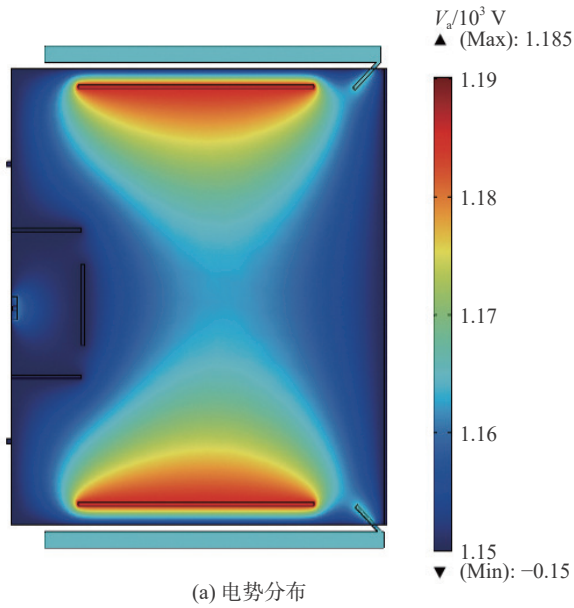
2 放电模型边界条件及仿真结果

放电过程模拟采用COMSOL多物理场耦合软件，分别调用了软件的静电模块（ES）、磁场模块（MF）、层流模块（SPF）和漂移扩散模块（DD）。其中，静电、磁场和层流模块设置为稳态分析，作为前端求解，漂移扩散模块设置为瞬态模块，在获得电场、磁场和中性原子密度分布的基础上完成后端求解，整个求解过程为耦合求解。根据求解步骤，先计算离子推力器放电室电势及磁场分布，根据表1所示离子推力器供电关系，分别在阳极、阴极触持极、屏栅、加速栅及磁铁等部件施加电压，根据电场强度与电势的关系，离子运动方向为沿电场线方向。此外，对于离子推力器来说，等离子体的自洽电势大小仅为0～2.0 V，等离子体运动产生的自洽电场几乎不会影响粒子运动[16-17]，因此，本文在电场模拟过程中假设等离子体处于准中性电离，且忽略了等离子体自洽电场影响。

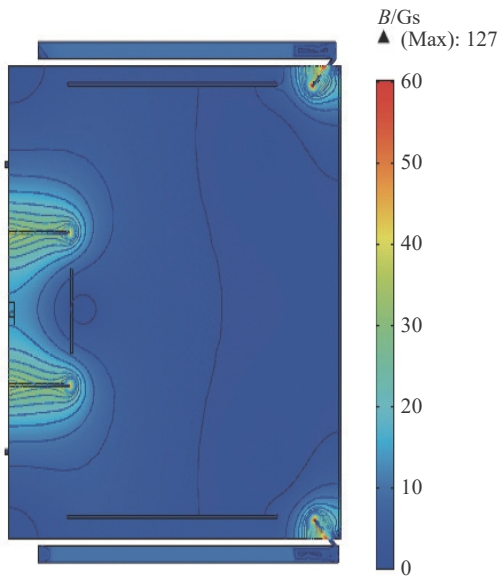
图2(a)为根据表1参数设置得到的放电室内部电势V_a的分布。由于离子推力器采用电磁铁设计，下游磁极和上游磁极均为同组线圈绕制而成，区别主要在于线圈匝数和磁芯材料（10 cm离子推力器采用4J29导磁材料）。因此，根据上下游磁极匝数、面积和电流值，以电流密度边界的形式加载至模型得到的放电室内部磁场强度B的分布如图2(b)所示。

从图2(a)结果来看，阳极电势最高，并在阳极中间区域形成了相对较为均匀的电势分布，在这种电势分布下，电子能够更有效地向阳极运动，并在阳极附近与进入放电室的中性气体发生碰撞，产生的离子则在电场作用下向栅极区域运动。从图2(b)结果来看，放电室内部磁场呈现出强烈的发散特性，并且在下游磁极和上游磁极之间形成了轴向磁场。图2(b)显示，靠近阴极磁极和放电室磁极的磁场强度分别为69.7 Gs和40.1 Gs，试验测量得到，在距离阴极磁极和放电室磁极1 mm处（磁极端面垂直方向）的磁场强度分别为62.5 Gs和32.3 Gs，二者比对误差较小，证明磁场分析结果是准确的。此外，从磁场构型来看，阴极发射的初级电子，将穿过挡板通道以一定的倾斜角度向阳极运动。综合图2(a)和图2(b)所示结果，电场和磁场基本呈现出正交特性，因此，放电室内部电子将以磁力线为导向中心做霍尔漂移，并且电子在向阳极的运动过程中，旋转半径会由于磁场强度的降低而增大，从而进一步提高与中性原子的碰撞概率。上述电磁场分布特性是发散磁场离子推力器有别于其他类型离子推力器的最大特点，即通过电磁场来控制初级电子向放电区域的扩散及电离过程，以获得较高的电离率和束流均匀性[18]。

进一步考虑放电室内部的中性气体压强及密度分布。根据表1参数，阳极和主阴极入口边界均



(a) 电势分布



(b) 磁场分布

图 2　放电室电势分布和磁场分布

Fig. 2　Electric potential distribution and magnetic field distribution in discharge chamber

设置为质量流率边界, 由于在径向方向上, 双栅极的间隙会引起一定的中性气体泄漏, 计算模型需考虑该泄漏情况以提高模拟的准确性。根据 Goebel 和 Katz[18] 的研究结果, 双栅间中性气体泄漏流量 Q_f可表示为: $Q_f = n_u \bar{c} A_S \eta_c / 4$, n_u、$\bar{c}$、A_S和η_c分别为栅极上游中性原子密度(大部分离子推力器在 10^{18} m^{-3} 量级[8])、原子热速度(约 300 m/s)、栅极开孔面积和 Clausing 因子(即气体通过管道的传输几率), 对于离子推力器来说, 该值范围在 0.5～0.6 之间[19]。代入参数计算后, 泄漏气体流量在 0.027～0.032 mg/s 范围内, 根据 10 cm 离子推力器阳极气体流量为 0.5 mg/s, 阴极和中和器气体流量均为 0.08 mg/s, 得到泄漏气体流量占比在 4.1%～4.8% 范围内, 因此, 对表 1 模型的入口流量在原来基础除去 4.1%～4.8% 的泄漏气体流量, 该部分不参与放电过程。由于放电室内部气体稀薄且流速一般均远小于 0.3 *Ma*, 流体属性设置为不可压缩流并忽略湍流影响[20]。流体温度设置为 500 K 以模拟离子推力器工作时内部气体的受热温度[21], 放电室内部所有流体边界均设置为无滑移, 即壁面流体的运动速度为 0。栅极模型设置为气体出口, 且在上述计算过程中已经考虑了栅极开孔率造成的气体泄漏, 因而气体出口边界根据实测结果设置为压强边界, 数值为 0.08 Pa, 流体运动方向设置为与出口面垂直。根据上述边界设置, 并假设放电室内部流体均为稳定性流动, 得到放电室内部的中性流体特性如图 3(a)～图 3(d) 所示。

从图 3(a) 结果来看, 放电室内部压强 P 大部分在 0.1～0.11 Pa 范围内, 而阳极和主阴极进气端的压强明显高于放电室其他区域, 这主要是由气体的扩散过程所决定。如图 3(b) 所示, 放电室内部中性原子密度 n_0 大部分在 1.5×10^{19} m^{-3} 量级, 并且其分布特性与压强分布基本一致。图 3(c) 为流体速度 v_0, 从结果来看, 放电室的 3 个主要进气端流体速度在 0.2～0.9 m/s 范围内, 靠近阳极和放电室壁面的流体速度则快速衰减, 这是由边界条件的设置及气体的黏滞特性所决定(氙气黏滞系数为 2.017×10^{-4} g/(cm·s))[22]。图 3(d) 为放电室内部气体的雷诺数, 雷诺数与数值计算结果离散化的稳定性具有紧密联系, 计算结果显示, 放电室内部雷诺数在 20～88.8 范围内, 远小于 1 200, 因此可认为放电室内部流体处于黏滞流状态。

漂移扩散模块采用瞬态分析过程, 时间范围为 10^{-8}～1 s, 以保证有足够的时长使得等离子体放电过程达到稳态。漂移扩散模块主要用来计算放电室内部(如电子温度、电子密度及碰撞概率等)关键气体放电参数。对于放电室内部等离子体鞘层电势, 根据 Goebel 和 Katz[18] 的研究结论, 对于准中性电离过程, 假设离子密度 n_i 与电子密度 n_e 相等, 即 n_i=n_e, 则鞘层电势可表示为

$$\phi = \frac{kT_e}{e} \ln\left(\frac{A_a}{A}\sqrt{\frac{2M}{\pi m_e}}\right) \tag{1}$$

式中: e 为电子电量; k 为 Boltzmann 常数; T_e 为电子温度(1 eV 对应 11 600 K); M 和 m_e 分别为离子质量和电子质量; A_a 和 A 分别为电子损失面积和离子损失总面积, 本文将 A_a 和 A 分别默认为阳极和屏栅面积。

在计算过程中, 假设电子温度为 5 eV(参考

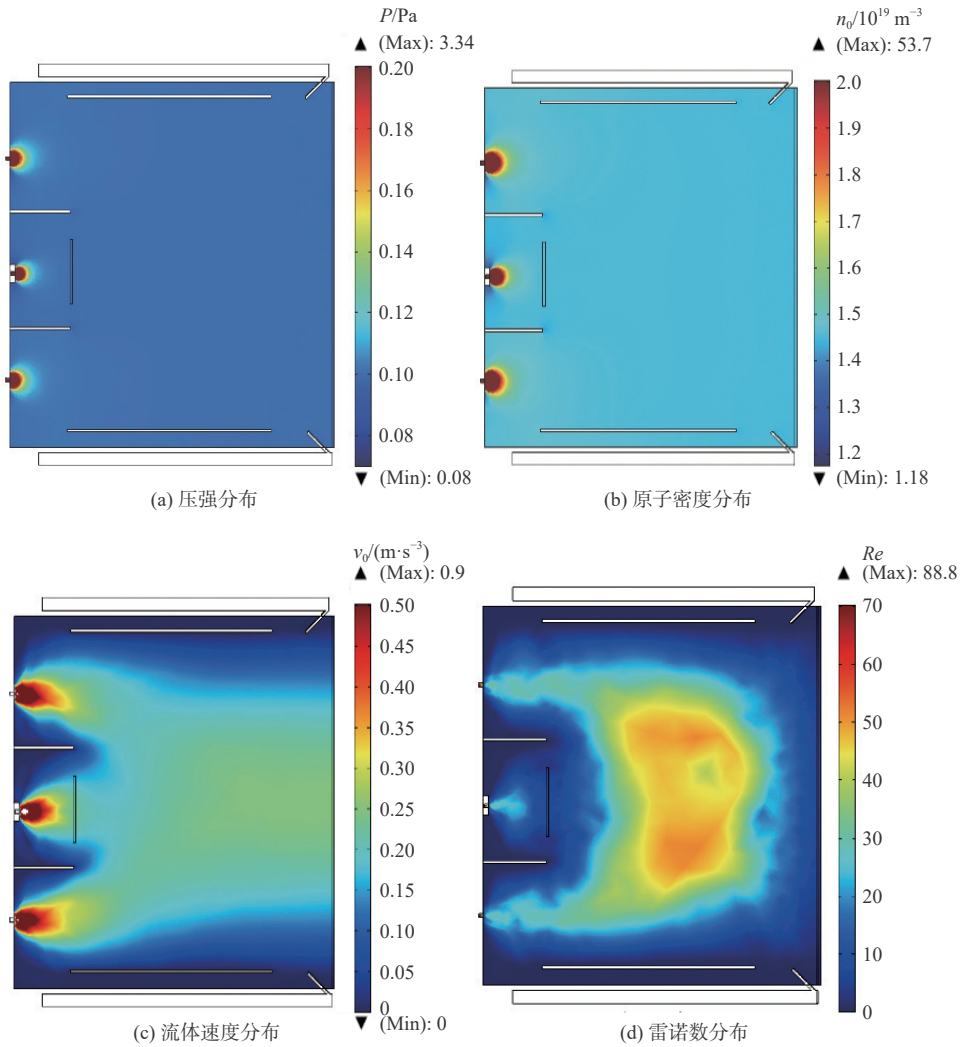


图3 放电室内部中性流体特性仿真结果

Fig. 3 Simulation of neutral fluid characteristics inside discharge chamber

NSTAR推力器在放电达到稳态时的电子温度[22]），由此得到放电模型的鞘层电势约为36.5 V，与表1所示阳极相对屏栅的电势，即35 V的电势差值基本相等，因此在仿真过程中，模型将阳极壁面作为鞘层边界以简化仿真过程。漂移扩散模块主要理论基础包括电子连续性方程、电子通量方程和能量平衡方程。其中，电子连续性方程可表示为

$$\frac{\partial n_e}{\partial t}+\nabla\Gamma_e=R_e-(u\cdot\nabla)n_e \tag{2}$$

式中：Γ_e为放电室内电子通量；u为电子迁移速度；R_e为电子产率，个/(m^3·s)。

放电室内部电子来源主要是弹性碰撞、激发碰撞和电离碰撞。Miller等[23]给出了不同碰撞类型的反应率系数表述形式，本文将弹性、激发和电离碰撞的反应率系数分别定义为k_1、k_2和k_3，反应率表述形式均为$\langle\sigma v\rangle$（其中，σ为碰撞截面，v为由电子温度所决定的粒子间碰撞频率[23]）。由于弹性碰撞主要进行能量传递，几乎不会产生二次电子，在仿真过程中忽略了弹性碰撞的电子产率。电子激发产率$r_2=k_2n_0n_e$，电离产率$r_3=k_3n_0n_e$，因此总电子产率R_e为

$$R_e=r_2+r_3=k_2n_0n_e+k_3n_0n_e \tag{3}$$

电子通量Γ_e根据迁移扩散方程可表示为

$$\Gamma_e=-(\mu_eE)n_e-D_e\cdot\nabla n_e \tag{4}$$

式中：E为放电室电场强度，可表示为$E=-\nabla V_a$（V_a如图2(a)所示）；μ_e为电子迁移系数，$\mu_e=e/[m_e\nu(1+\Omega_e^2)]$。同时，模型中考虑了电子和中性原子间的碰撞频率$\nu_{en}$及电子和离子间的碰撞频率$\nu_{ei}$[23]，而由于电子与电子、离子与离子及离子和中性原子间的碰撞过程所产生的电子相比电子和中性原子的电离过程几乎可忽略不计，模型中忽略了上述碰撞过程。式(4)中霍尔漂移修正项即Ω_e可表示为$\Omega_e=eB/(m_e\nu)$（B如图2(b)所示），从表达形式可看出[18]，对于发散磁场离子推力器，在上游磁极和下游磁极之间形成的强烈发散磁场区域，电子的霍尔漂移运动非常明显。D_e为电子扩散系数，与电子迁移系数

μ_e之间存在 Einstein 关系，即$D_e=\mu_e T_e$。模型中的能量平衡方程可表示为

$$\frac{\partial n_\varepsilon}{\partial t}+\nabla\Gamma_\varepsilon+E\Gamma_e=S_{en}-(u\cdot\nabla)n_\varepsilon \tag{5}$$

式中：n_ε和Γ_ε分别为电子能量密度和电子能量密度通量，2 个变量均与μ_e、n_e 和T_e存在紧密联系；S_{en}为电子碰撞过程的能量损失，由弹性碰撞能量损失d_{e1}、一级和二级电离能损失d_{e2}和d_{e3}组成。

迁移扩散模块的参数设置如表 2 所示，电子密度和电子温度的计算初值分别设置为 1×10^{18} m^{-3} 和 2.5 eV。表中：σ_{en} 为电子和原子的碰撞截面，ln Λ 为库伦碰撞对数。

表 2　迁移扩散模块的参数设置

Table 2　Parameter setting of drift diffusion module

参数	表达式
k_1	$1.99^{-14}T_e^{0.93}e^{-0.41/Te}$
k_2	$1.93^{-19}e^{-11.6/Te}/\left[T_e^{0.5}(8eT_e/\pi/m_e)^{0.5}\right]$
k_3	$10^{-20}\left(3.97+0.643T_e-0.036\,8T_e^2\right)e^{-12.127/Te}(8eT_e/\pi/m_e)^{0.5}$
r_1	$k_1n_0n_e$
r_2	$k_2n_0n_e$
r_3	$k_3n_0n_e$
μ_e	$e/[m_e(\nu_{en}+\nu_{ei})]$
R_e	r_2+r_3
S_{en}	$-e(r_1d_{e1}+r_2d_{e2}+r_3d_{e3})$
σ_{en}	$6.6^{-19}(T_e/4-0.1)/\left[1+(T_e/4)^{1.6}\right]$
lnΛ	$23^{0.5}\ln\left(10^{-6}n_e/T_e^3\right)$
ν_{en}	$\sigma_{en}n_0(8kT_e/\pi/m_e)^{0.5}$
ν_{ei}	$2.9^{-12}n_e\ln\Lambda/T_e^{1.5}$
d_{e1}	0
d_{e2}	8.31
d_{e3}	12.13

根据表 2 的参数设置，以及前述静电、磁场和层流模块，得到电势分布、中性粒子密度、流体速度等参数，计算得达到放电稳定后的电子密度n_e及电子温度T_e，分别如图 4(a) 和图 4(b) 所示。在仿真过程中，根据等离子体密度的动态变化结果，认为放电过程大约在 10^{-3}～10^{-2} s 左右即达到平衡态，并且从电子密度的变化趋势来看，挡板和阴极极靴组成的挡板通道，明显将大部分电子制约在靠近阳极壁面附近，此处电子密度最高且也是阳极的进气端，从而可提高电离效率，限于篇幅，本文未给出电子密度的动态变化情况。

如图 4(a) 所示，电子密度在阳极进气端和耦合放电区域附近最高，约为 8.57×10^{18} m^{-3}，而随着电子在电场和发散磁场作用下向放电室上游磁极运动过程中，电离效应逐渐减弱，阳极壁面及栅极上游区域的等离子密度约为 6.8×10^{17} m^{-3} 量级，造成图 4(a)

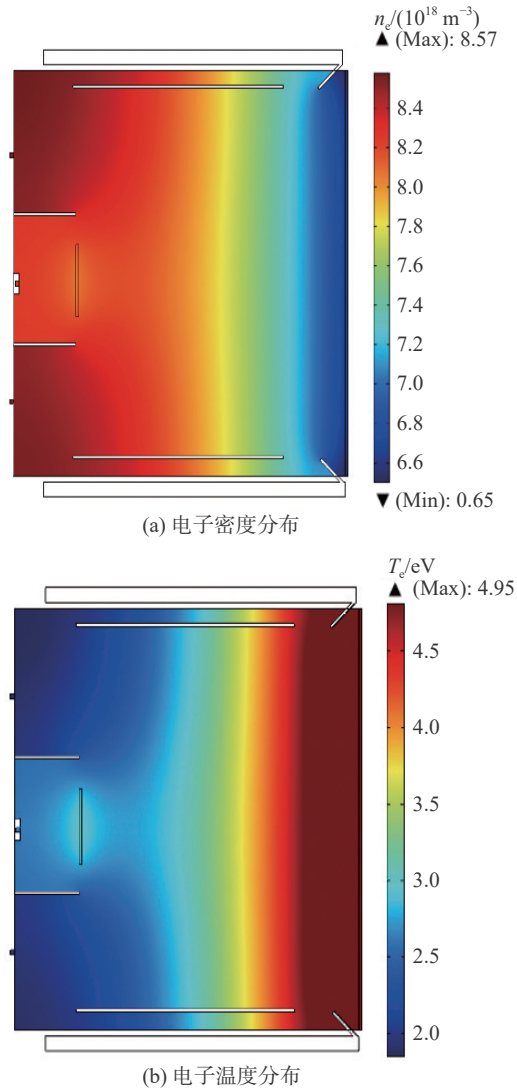


图 4　电子密度和电子温度分布

Fig. 4　Distribution of electron density and temperature

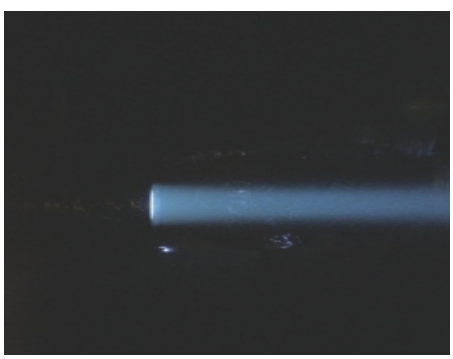
所示电子密度分布特性的原因与中性原子密度分布(见图 3(b))具有紧密关系，即在准中性电离状态下，电子密度与原子密度呈正比。在靠近栅极区域，由于设置为气体和电子的出口边界，且磁场对等离子体的影响相比耦合放电区域要减弱许多，导致电子密度远低于耦合放电区域。图 4(b) 给出了放电室内部的电子温度分布，从结果来看，在图 2(a) 所示的电势分布影响下，电子温度沿离子推力器轴向逐渐升高，从阴极出口处的 3 eV 升高到 5 eV 左右。电场分布是造成图 4(b) 所示的电子温度趋势的主要原因，从图 2(a) 的电势分布来看，上游磁极至下游磁极区间内的电势明显高于栅极上游区域，电场线方向指向阳极，由此决定了电子在向出口运动的过程中，电子能量处于持续提升状态。

3　试验验证

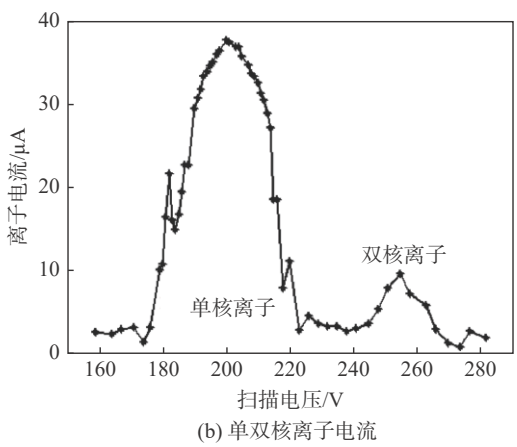
由于目前尚未对放电室内部的等离子体参数

进行测量，针对本文建立的放电模型，在假设栅极上游等离子体密度全部被引出的情况下，通过屏栅电流即束流大小来验证在20 mN工况（束流实测值为0.376 mA）下模型计算结果的准确性。由于束流中实际含有大量二价Xe离子，而放电模型的计算结果是基于等离子体中离子类型均为一价Xe离子的基础上，束流实测值需进行单双核离子比例修正后，计算结果才具备可比性。

图5(a)为离子推力器处于稳态工作时的引束流状态。可以看出，在凹面栅极的影响下，羽流区束流呈现明显先聚焦后发散的特点，因此发散磁场离子推力器在羽流近场区域具有更高的推力密度。图5(b)为离子推力器在20 mN工况下，采用E×B探针测量得到的1～300 V扫描电压下的单核和双核Xe离子分布特性。图中，横坐标为E×B探针加载的扫描电压，用来在探针内部形成扫描电场，带电离子进入后，在电场作用下发生偏转，由于带电量不同致使偏转距离不同，从而分辨不同带电粒子的电流值。根据图5(b)所示结果，采用Gauss拟合后可分别得到单核和双核Xe离子电流的峰值和中值，图5(b)所示曲线的2个波峰面积比即为单双核离子比例。拟合后得到单核Xe离子波峰为37.26 μA，中值为23.15 V，双核Xe离子波峰为20.15 μA，中值为7.01 V，由此计算得到双核和单核离子比例为16.3%，双核离子占总束流离子比例为14.1%。



(a) 束流引出



(b) 单双核离子电流

图5 离子推力器试验验证

Fig. 5 Test verification of ion thruster

束流实测结果为0.376 mA，在考虑到双核离子占总束流离子比例为14.1%的基础上，得到单核离子束流为0.323 mA。栅极上游等离子体密度计算值约为6.8×10^{17} m^{-3}，根据Goebel和Katz[18]得到的束流与栅极上游等离子体密度关系，即

$$I_b = \frac{1}{2}n_e e\sqrt{\frac{kT_e}{M}}A_s T_s \tag{6}$$

式中：I_b为推力器束流；A_s和T_s分别为栅极面积和栅极有效透过率（10 cm离子推力器栅极透过率为67%）。

将电子温度（根据图4(b)取为4.95 eV）和上游等离子体密度模拟值代入式(6)后计算得到单核离子束流值为0.353 mA，与0.323 mA的实测值比对误差为9%。误差认为来自于多个方面，如模拟过程中的边界条件设置、模型网格划分或E×B探针的测量误差，但总体而言，比对结果证明模型具有较好的精度。

将束流计算结果延伸至推力特性，根据推力表述形式，即$F = 1.65K_\alpha K_\beta I_b\sqrt{V_b}$，$K_\alpha$为发散角修正系数，$K_\beta$为双核离子修正系数，$V_b$为束电压。通过试验测量得到20 mN工况下的束流发散角和双核离子修正系数分别为0.991和0.962，由此基于仿真和实测束流值，计算得到的推力值分别为17.2 mN和20.05 mN，误差与束流误差相同，模型的计算结果可用于推力器优化设计过程中的推力指标预估，并提供性能指标参考。

4 结 论

采用COMSOL多物理场耦合分析软件建立了国内研制的10 cm口径发散磁场离子推力器放电模型，并针对20 mN的在轨主工况点进行了关键放电参数计算，并与20 mN工况的实际引出束流进行了验证比对，研究表明：

1）10 cm离子推力器放电室内部具有强烈发散特性的轴向磁场与电场分布呈现正交特点，电子以磁力线为导向中心并做霍尔漂移，主要电离区域集中在靠近阴极的放电区域及下游磁极和上游磁极间形成的发散磁场区域。

2）放电室内部压强大部分在0.1～0.11 Pa范围内，而放电室大部分区域的中性原子密度约为1.5×10^{19} m^{-3}且呈现均匀分布特点，阴极和阳极进气端的流体速度在0.2～0.9 m/s的范围内，且流体呈现出明显黏滞流特性。

3）电子密度峰值出现在阴极放电区域，约为8.57×10^{18} m^{-3}，阳极壁面及栅极上游区域的等离子密度约为6.8×10^{17} m^{-3}。采用E×B探针测量得到双

核离子占总束流离子比为 14.1%, 根据上级等离子体密度模拟值得到的 0.353 mA 束流理论值与 0.323 mA 的束流实测值, 比对误差为 9%, 误差认为主要来自于仿真条件设置及试验测量。

后续将针对 10 cm 离子推力器的更多工况点进行测试, 并根据测试结果进行模型的验证工作。

参考文献（References）

[1] GRAY H, SMITH P, FERN D. Design and development of the UK-10 ion propulsion system: AIAA-1996-3084[R]. Reston: AIAA, 1996.

[2] BASSNER H, KILLINGER R, MARX M, et al. Ion propulsion for drag compensation of GOCE: AIAA-2000-3417[R]. Reston: AIAA, 2000.

[3] CORBETT M, EDWARDS C. Thrust control algorithms for the GOCE ion propulsion assembly: IEPC-2007-228[R]. Florence: IEPC, 2007.

[4] TIGHE G, CHIEN R, SOLIS E, et a1. Performance evaluation of the XIPS-25 cm thruster for application to NASA discovery missions: AIAA-2006-4666[R]. Reston: AIAA, 2006.

[5] 赵以德, 吴宗海, 张天平, 等. 离子推力器多模式化研究[J]. 推进技术, 2020, 41(1): 187-193.
ZHAO Y D, WU Z H, ZHANG T P, et al. Research on multi-mode realization of ion thruster[J]. Journal of Propulsion Technology, 2020, 41(1): 187-193(in Chinese).

[6] MILDER N. A survey and evaluation of research on the discharge chamber plasma of a Kaufman thruster: AIAA-1969-494[R]. Reston: AIAA, 1969.

[7] GOEBEL D. Ion source discharge performance and stability[J]. Physics of Fluids, 1982, 25(6): 1093-1102.

[8] BROPHY R, WILBUR P. Calculation of plasma properties in ion sources: AIAA-1985-2006[R]. Reston: AIAA, 1985.

[9] JAMES A. Particle based plasma simulation for an ion engine discharge chamber[D]. Dayton: Wright State Niversity, 2007: 87-88.

[10] MILLIGAN D, GABRIEL S. Generation of experimental plasma parameter maps around the baffle aperture of a Kaufman (UK-25) ion thruster[J]. Acta Astronautica, 2009, 64(2): 952-968.

[11] OZAKI T, KASAI Y, NAKAGAWA T, et al. In-orbit operation of 20 mN class xenon ion engine for ETS-Ⅷ: IEPC-2007-084[R]. Florence: IEPC, 2007.

[12] OZAKI T, NISHIDA E, KASAI Y, et al. Development status of xenon ion engine subsystem for ETS-Ⅷ: AIAA-2003-2215[R]. Reston: AIAA, 2003.

[13] 杨福全, 王成飞, 胡竟, 等. 超低轨道卫星应用离子电推进技术方案[J]. 中国空间科学技术, 2021, 1(3): 1-8.
YANG F Q, WANG C F, HU J, et al. Ion-electric propulsion technology scheme for ultra-low orbit satellites[J]. Chinese Space Science and Technology, 2021, 1(3): 1-8(in Chinese).

[14] 杨福全, 胡竟, 郭德洲, 等. LIPS-100离子推力器挡板通道面积优化研究[J]. 真空与低温, 2021, 2(5): 395-399.
YANG F Q, HU J, GUO D Z, et al. Study on optimization of baffle channel area of LIPS-100 ion thruster[J]. Vacuum and Cryogenics, 2021, 2(5): 395-399(in Chinese).

[15] 吴辰宸, 顾左. Kaufman型离子推力器束流调节理论与试验研究[J]. 真空科学与技术学报, 2015, 35(11): 1368-1373.
WU C C, GU Z. Theoretical and experimental studies of beam current regulation of Kaufman-type ion thruster[J]. Chinese Journal of Vacuum Science and Technology, 2015, 35(11): 1368-1373(in Chinese).

[16] 陈娟娟, 张天平, 刘明正, 等. LIPS-200离子推力器放电室原初电子动力学行为的数值模拟研究[J]. 推进技术, 2015, 36(1): 155-160.
CHEN J J, ZHANG T P, LIU M Z, et al. Investigation on dynamical behavior of primary electrons in LIPS-200 ion thruster discharge chamber[J]. Journal of Propulsion Technology, 2015, 36(1): 155-160(in Chinese).

[17] 龙建飞, 张天平, 杨威, 等. 离子推力器推力密度特性[J]. 物理学报, 2018, 67(2): 022901.
LONG J F, ZHANG T P, YANG W, et al. Thrust density characteristics of ion thruster[J]. Acta Physica Sinica, 2018, 67(2): 022901(in Chinese).

[18] GOEBEL D, KATZ I. Fundamentals of electric propulsion: Ion and hall thrusters[M]. New York: Wiley, 2005: 142-145.

[19] CLAUSING P. The flow of highly rarefied gases through tubes of arbitrary length[J]. Journal of Vacuum Science and Technology, 1971, 8(2): 636-646.

[20] SUM M M, ZHANG T P, WEN X D, et al. Plasma characteristics in the discharge region of a 20 A emission current hollow cathode[J]. Plasma Science and Technology, 2018, 20(1): 025503.

[21] 孙明明, 张天平, 吴先明. 20 cm离子推力器放电室流场计算模拟[J]. 强激光与粒子束, 2015, 27(5): 054003.
SUM M M, ZHANG T P, WU X M. Flow field simulation of 20 cm diameter ion thruster discharge chamber[J]. High Power Laser and Particle Beams, 2015, 27(5): 054003(in Chinese).

[22] WIRZ R, KATZ I. Plasma processes of DC ion thruster discharge chambers: AIAA-2005-3690[R]. Reston: AIAA, 2005.

[23] MILLER J, PULLINS S, LEVANDIER D, et al. Xenon charge exchange cross sections for electrostatic thruster models[J]. Journal of Applied Physics, 2002, 91(3): 984-991.

Discharge model of divergent magnetic field ion thruster of 10 cm diameter

SUN Mingming[1, *], GENG Hai[1], HU Jing[1], YANG Fuquan[1], LONG Jianfei[2, 3]

(1. Science and Technology on Vacuum Technology and Physics Laboratory, Lanzhou Institute of Physics, Lanzhou 730000, China;
2. School of Fundamental Physics and Mathematical Sciences, Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou 310024, China;
3. School of Nuclear Science and Technology, University of South China, Hengyang 421001, China)

Abstract: A discharge model of the divergent magnetic field ion thruster (Kaufman thruster) of 10 cm diameter is established with COMSOL multi-physical field coupling software to study the internal discharge process of the thruster and to provide reference for subsequent engineering improvement. The key discharge parameters are obtained and verified through experiments. The calculation results show a magnetic field formed between the upstream and downstream magnetic poles in the discharge chamber with strong divergent characteristics. Due to the orthogonal electric field, the Hall drift of the electrons occur with the magnetic field line as the guiding center. The gas pressure in the discharge chamber is uniformly distributed in the range of 0.1−0.11 Pa, with the neutral atom density in most areas about 1.5×10^{19} m^{-3}, and the fluid velocity in the range of 0.2−0.9 m/s. The fluid shows obvious characteristics of viscous flow. The peak electron density appears in the cathode outlet region, which is about 8.57×10^{18} m^{-3}, while the plasma density near the anode wall and the upstream of the screen grid is about 6.8×10^{17} m^{-3}. The results measured by E×B probe show that the proportion of divalent ions in the total beam ions is 14.1%. The comparison error between the theoretical beam current value of 0.353 mA and the measured beam current value of 0.323 mA is 9%, which is mainly due to the simulation setting and the measurement error. The discharge model established in this paper can provide rapid discharge parameter analysis and reference for optimal design of the thruster.

Keywords: divergent magnetic field; ion thruster; discharge model; plasma density; electron temperature

Received: 2021-10-29; **Accepted**: 2022-01-07; **Published Online**: 2022-02-14 11:38

URL: kns.cnki.net/kcms/detail/11.2625.V.20220214.1108.001.html

Foundation items: National Natural Science Foundation of China (61901202); China Postdoctoral Science Foundation (2019M653906); Key Projects of the Education Department of Hunan Province (19A440); Hunan Provincial Natural Science Foundation of China (2021JJ30564); Doctoral Startup Fund of University of South China (190XQD013)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0687

超声速民机发动机短舱布局对声爆的影响

肖天航*，徐雅楠，朱震浩，邓双厚
(南京航空航天大学 航空学院，南京 210016)

摘　　要：超声速民机较高的飞行速度导致发动机短舱与机翼机身存在强烈的气动干扰，且高速喷流会改变主流的激波系结构，进而影响声爆强度。以超声速民机为研究对象，对不同短舱布局开展数值模拟研究，揭示发动机短舱位置、数量对近场过压信号特性的影响规律。数值模拟运用有限体积法在直角网格上求解流体控制方程进行，通过伴随自适应网格加密生成高分辨率网格以精确捕捉近场过压信号。结果表明：发动机进气道唇口激波、尾喷口后缘激波及尾喷流与机体机翼引起的激波系存在强烈的相互干扰，一定程度上增加了近场过压幅值，从而增强了声爆强度。短舱沿弦向前移及置于机翼上方可以有效降低声爆强度，沿展向外移通过抑制尾部各激波的合并也可有效降低声爆强度；在相同总推力前提下，相比三发构型，双发构型能有效降低后体激波强度，但较大尺寸的短舱引起较强的进气道唇口激波。综合考虑喷流噪声和气动阻力因素，翼下双发布局对于新一代超声速民机并非最佳选项。

关　键　词：超声速民机；短舱布局；声爆；网格自适应；笛卡儿网格
中图分类号：V221^{+}.3; TB553
文献标志码：A　　**文章编号**：1001-5965(2023)09-2267-12

随着技术的进步和全球商务旅行的蓬勃发展，超声速民机或将重返历史舞台，近年来，全球各先进航空大国一直保持着相关技术研究的热情。其中，低声爆设计是公认的未来超声速民机商业化的关键技术问题之一[1]。

地面声爆强度的准确预测是降低声爆水平和低声爆设计的前提。声爆传播距离长，且受大气变化和地球湍流边界层等因素影响，其预测存在较大困难[2]。国内诸多学者，如王迪[3]、钱战森[4]、冯晓强[5]、王刚[6]、乔建领[7-8]、张绎典[9]与陈鹏[10]等对声爆预测技术进行了研究和总结，至今已发展了多种声爆预测方法，包括简化的声爆预测方法、波形参数法、基于CFD近场计算与基于增广Burgers方程或KZK方程的远场声爆传播相结合的声爆预测方法。针对超声速民机的低声爆气动布局优化设计问题，国内也开展了大量的相关研究。冯晓强[11]和李占科[12]等先后对静音锥的低声爆效果展开了研究；冯晓强等[13]基于SGD (Seebass-George-Darden)方法对超声速低声爆布局进行了反设计，并提出了一种“梭式”布局静音超声速民机方案；韩忠华[14-15]、乔建领[16]等应用代理模型优化方法实现了全局的低声爆优化；黄江涛等[17]通过求解流场声爆耦合伴随方程对超声速公务机进行了低声爆布局优化设计。

由于较高的飞行速度，超声速民机的短舱与机翼和机身之间存在不可忽视的气动干扰现象，且发动机高速尾喷流也会改变主流场中的复杂激波系结构，从而影响超声速民机的声爆特性。因此，有必要对超声速民机的短舱布局展开研究。就该问题而言，目前已有相关学者开展了前期工作。Putnam和Capone[18]在1.2 m×1.2 m(相当于4 ft×4 ft)的超声

收稿日期：2021-11-16；**录用日期**：2022-04-10；**网络出版时间**：2022-04-21 12:26
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220420.1604.005.html
基金项目：航天进入减速与着陆技术实验室开放基金 (EDL19092111)；国家自然科学基金 (11672133)；陕西省自然科学基础研究计划 (2021JQ-078)
*通信作者. Email：xthang@nuaa.edu.cn

引用格式：肖天航，徐雅楠，朱震浩，等. 超声速民机发动机短舱布局对声爆的影响[J]. 北京航空航天大学学报，2023，49(9)：2267-2278. XIAO T H，XU Y N，ZHU Z H，et al. Effect of engine nacelle layout on sonic boom of supersonic transport[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2267-2278 (in Chinese).

速压力风洞中，探索了单独尾喷管及带有4个翼下短舱的超声速飞机构型的激波系结构，以及其对声爆传播的影响规律，结果表明，尾喷流与机翼后缘激波存在强烈的相互干扰，改变了近场压力分布特征。Bui[19]采用数值模拟方法研究了单独喷管的落压比(nozzle pressure ratio, NPR)对声爆特性的影响，发现高度欠膨胀喷流能够显著降低声爆压力特性中的尾部激波强度。基于发动机喷流与流场激波结构存在相互干扰的事实，Castner[20-21]和Howe[22]通过求解Euler方程，对不同短舱布局和喷管落压比对翼身组合体近场流场的干扰情况进行了探究。Castner和Lake[23]还针对该问题对比了黏性湍流计算和无黏流计算的近场过压信号分布，两者在幅值上稍有差别，但可认为基于Euler方程的流场求解满足要求且效率更高。

可以看出，前期的研究都基于简单的翼身组合体，且仅模拟了发动机的喷流；近年来，国内外学者在考虑复杂外形和发动机进气道的情况下，探究了短舱布局对流场和声爆的影响。冯晓强等[24]在一种“梭式”布局超声速民机构型[13]的基础上，通过改变翼下短舱展向和弦向位置，探究其对声爆强度和气动性能的影响，结果表明，通过将短舱位置合理布置，可以有效降低声爆的同时改善气动性能，但短舱位置对声爆性能的影响未展开详细论述。Kim和Lee[25]对带动力短舱布局的影响进行了数值模拟研究，但只考虑了短舱弦向位置的影响。因此，目前对超声速民机短舱布局的研究仍具有一定的局限性，并未全面探讨短舱与机翼、机身不同相对位置及发动机数量对声爆强度的影响。

针对复杂构型的超声速民机模型，本文通过数值模拟方法较全面地研究了不同短舱布局超声速民机的声爆特性，主要包括发动机短舱与机翼的相对位置和发动机数量的影响。采用在直角网格上运用有限体积法解算Euler方程求解流场，通过伴随自适应方法提高局部网格分辨率，以较为准确地捕捉不同短舱布局的近场过压信号。由于近场过压信号是远场声爆传播的输入，可以直接决定地面声爆强度，通过分析近场过压幅值来对比和评估不同短舱布局的声爆特性。

1 计算模型与工况

本文采用的超声速民机基础构型为第一届AIAA声爆预测会议提供的LM1021构型的风洞试验缩比模型[26]，缩放比率为0.008，模型机身长568.96 mm，参考面积为21 404.86 mm^2；发动机尾喷管的几何外形根据文献[20]中的nozzle6等比例缩放得到，其喷管直径d为14 mm。整体外形如图1所示。图1和表1给出了不同发动机短舱布局形式，包括翼上式与翼下式、三发与双发，以及不同展向与弦向位置的具体参数，其中，展向位置改变的同时保证翼下短舱外露于机翼的长度不变。

针对三发与双发布局，为了保证总推力相同，对风扇入口和尾喷口面积进行缩放。假设在截面积改变前后忽略短舱外阻力的变化，只需保证内流对短舱内表面的作用力，即内推力相同。内流控制体相关参数如图1所示，V_0、p_0、A_0分别为发动机进气道捕捉截面上的气流流速、压强、截面积；V_9、p_9、A_9分别为喷口截面上的流速、压强、截面积。

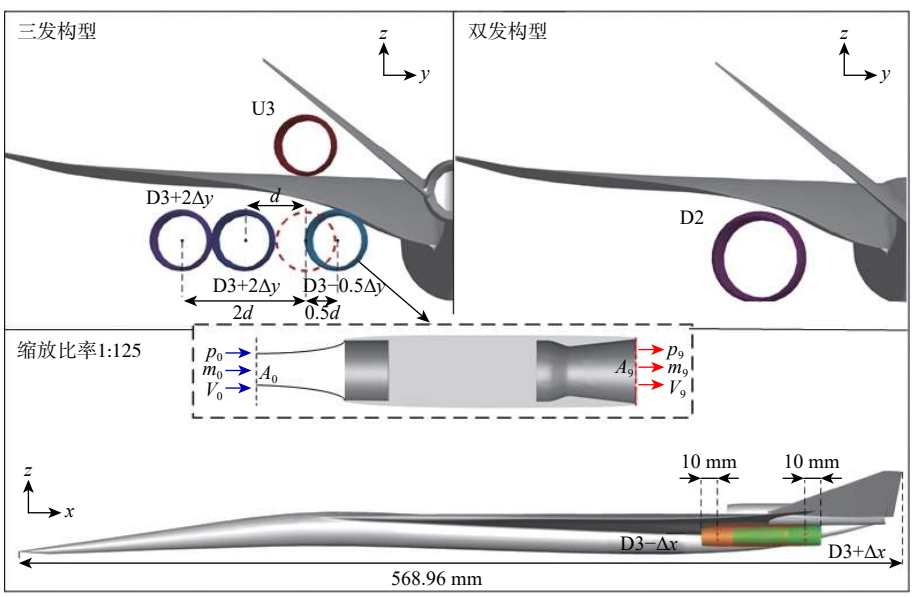


图1 LM1021基础构型及不同发动机短舱布局示意图

Fig. 1 Baseline of LM1021 with powered nacelles and configurations with various nacelle layouts

表 1　不同构型短舱相对基础构型的偏移量

Table 1　Offsets of nacelles from baseline configuration

构型	展向偏移量/d	轴向偏移量/d	纵向偏移量/d
基础构型	0	0	0
D3−Δx	0	−0.7	0
D3+Δx	0	0.7	0
D3−0.5Δy	−0.5	0.12	0
D3+Δy	1	0.5	0
D3+2Δy	2	1	0
U3	0	0	1.54
D2	0	0	0

注: 偏移量均为以喷管直径d为基准的相对量。编号含义如下: 轴向从机头到机尾方向为正, 展向从翼根到翼梢方向为正, 采用右手坐标系。D (down)表示短舱置于机翼下方, U (up) 表示短舱置于机翼上方; 3或2表示发动机数量; Δx/Δy表示短舱沿弦向/展向移动, 其中, Δy等于一个喷管直径; +/−表示发动机沿对应坐标轴正/负向移动。

发动机推力表达式为[27]

$$F_{\text{in}} = \dot{m}_9 V_9 - \dot{m}_0 V_0 + (p_9 - p_0)A_9 \tag{1}$$

式中: $\dot{m}_0$和$\dot{m}_9$分别为进气道捕捉的气流和尾喷管喷流的质量流量。

若质量流量守恒, 式(1)可简化为

$$F_{\text{in}} = \rho_0 V_0 (V_9 - V_0)A_0 + (p_9 - p_0)A_9 \tag{2}$$

式中: ρ_0为发动机进气道捕捉截面上的气流密度。

由于给定的发动机风扇入口和尾喷口的落压比、总温比和马赫数不随发动机数量变化, 可近似认为 $\rho_0 V_0(V_9-V_0)$ 和 (p_9-p_0) 也不随发动机数量变化, 分别记为常数 K_1 和 K_2。三发与双发状态下, 每个发动机产生的推力分别为 F_1 和 F_2; 比例因子 k 为双发时每个发动机横截面积和三发时每个发动机横截面积之比, 发动机各截面面积同比例变化, 则有$3F_1 = 2F_2$, 即

$$3(K_1A_0 + K_2A_9) = 2(K_1kA_0 + K_2kA_9) \tag{3}$$

进而可知, 当三发改为双发时, 为保证推力近似不变, 每个发动机进口流量捕获面积和尾喷口面积需放大 1.5 倍。

2　数值计算方法

2.1　流体控制方程及求解

通过求解 Euler 方程进行流场计算, 该方程在空间上采用基于直角网格的有限体积法进行离散, 运用 Roe 迎风格式[28] 计算控制体交界面对流通量, 为保证二阶空间精度, 交界面两侧流场值由最小二乘法或高斯积分法重构的流场梯度线性插值得到, 并使用 Venkatakrishnan 限制器[29] 防止新的极值出现; 时间方向上运用后向差分隐式格式离散, 在时间方向上运用全隐式推进求解, 离散后的方程采用 GMRES 方法迭代求解。

2.2　网格生成与边界条件

对求解区域的网格划分是实现流体控制方程离散求解的重要部分, 计算网格的合理设计和高质量生成是流场数值计算的前提条件, 对于流场激波的捕捉尤为如此。笛卡儿直角网格具有单元规则、生成过程简单、易于实现网格自适应等优势, 利用前期研究基础[30-31], 对计算域生成直角网格, 并运用网格自适应技术进行局部加密, 以提高捕捉目标特征的精度。

为提高计算精度, 对于初始网格的生成, 一方面, 将计算域基于马赫角进行旋转, 使直角网格单元与激波面大致平行, 以利于捕捉激波; 另一方面, 在机体下方、发动机喷管后方设置局部加密区域, 以更好地捕捉机身下方激波与膨胀波流场结构和发动机喷流对激波传播的影响。计算域和对称面网格如图 2 所示, 初始网格的总网格量约为 400 万。边界条件的设置如图 2 所示, 远场设置为压力远场, 物面设置为滑移壁面, 发动机风扇入口截面设置为压力出口, 尾喷口设置为压力入口。

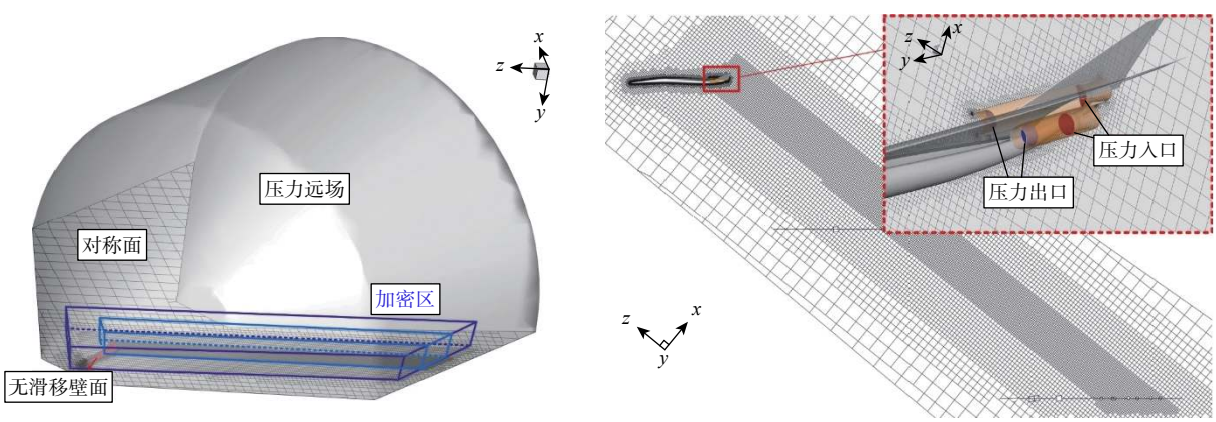


图 2　初始网格与边界条件

Fig. 2　Initial mesh and boundary conditions

2.3 直角网格伴随自适应技术

为提高计算精度，采用基于伴随误差分析的直角网格自适应方法对初始网格进行自适应加密。相比于常规的基于流场特征的网格自适应方法[32]，伴随自适应建立了与目标函数的直接关系，能够精准地估计目标函数全局误差，并对结果修正，计算精度更高[33]。详细的伴随自适应方法可参考文献[34]，本节只简单介绍其基本思路和步骤，流程如图3所示。

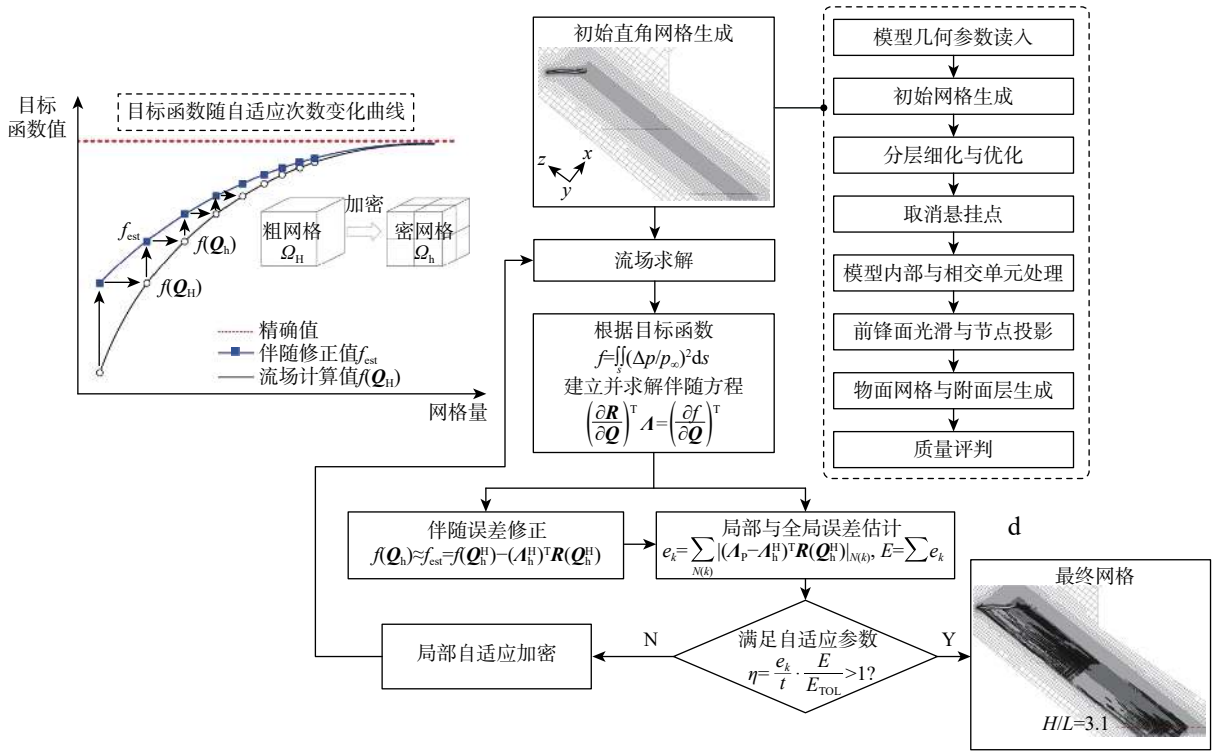


图3 直角网格伴随自适应技术流程

Fig. 3 Flow chart for adjoint-based adaptive Cartesian mesh refinement

首先，基于初始网格求解流场。针对用于声爆预测的近场流场计算，目标函数选取近场过压信号分布，即过压值的平方在指定区域上的积分：

$$f=\iint_s(\Delta p/p_\infty)^2\mathrm{d}s \tag{4}$$

式中：过压值 $\Delta p/p_\infty=(p-p_\infty)/p_\infty$，$p$ 为流场当地压强，p_∞为远场来流压强。指定截面选为机身轴线正下方 H/L=1.4 与 H/L=3.1 两处，其中，H 为截面距机身的垂直高度，L 为机身长度。

根据求得的流场解和目标函数建立伴随方程：

$$\left(\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{Q}}\right)^{\mathrm{T}}\boldsymbol{\Lambda}=\left(\frac{\partial f}{\partial \boldsymbol{Q}}\right)^{\mathrm{T}} \tag{5}$$

式中：$\boldsymbol{Q}$ 为流场解；$\boldsymbol{R}(\boldsymbol{Q})$ 为流场残差；$\boldsymbol{\Lambda}$为伴随变量；f为目标函数。由此，通过伴随方程建立了流场变量、局部残差和目标函数之间的关系。

用下标或上标 H、h 分别标识粗网格和密网格，假定在粗网格 Ω_H 上计算得到的目标函数值为 $f(\boldsymbol{Q}_\mathrm{H})$。为提高预测精度，在粗网格的基础上，通过单元细分的方式进行网格加密，得到密网格 Ω_h，假设密网格上得到的目标函数值为 $f(\boldsymbol{Q}_\mathrm{h})$。通过求解粗网格上的伴随方程得到相应的伴随变量，再从粗网格通过重构算子 $\boldsymbol{L}_\mathrm{h}^\mathrm{H}$ 得到密网格上的近似值 $\boldsymbol{Q}_\mathrm{h}^\mathrm{H}=\boldsymbol{L}_\mathrm{h}^\mathrm{H}\boldsymbol{Q}_\mathrm{H}$、$\boldsymbol{\Lambda}_\mathrm{h}^\mathrm{H}=\boldsymbol{L}_\mathrm{h}^\mathrm{H}\boldsymbol{\Lambda}_\mathrm{H}$，从而在不需要直接求解密网格流场的情况下，得到密网格下的近似预测值 f_{est}，以避免大量复杂的密网格计算，即

$$f(\boldsymbol{Q}_\mathrm{h})\approx f_{\mathrm{est}}=f(\boldsymbol{Q}_\mathrm{h}^\mathrm{H})-\left(\boldsymbol{\Lambda}_\mathrm{h}^\mathrm{H}\right)^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{Q}_\mathrm{h}^\mathrm{H}) \tag{6}$$

式中：f_{est} 为可计算的目标函数（即图3左上角曲线所示）。

同时，还可得到剩余误差项 $|f(\boldsymbol{Q}_\mathrm{h})-f_{\mathrm{est}}|$：

$$f_{\mathrm{err}}=\left(\boldsymbol{\Lambda}_\mathrm{h}-\boldsymbol{\Lambda}_\mathrm{h}^\mathrm{H}\right)^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{Q}_\mathrm{h}^\mathrm{H}) \tag{7}$$

为了避免求解式(7)中密网格上的伴随变量 $\boldsymbol{\Lambda}_\mathrm{h}$，通过高阶多项式插值算子 $\boldsymbol{P}_\mathrm{h}^\mathrm{H}$，将粗网格伴随变量 $\boldsymbol{\Lambda}_\mathrm{H}$ 插值得到密网格伴随变量近似值，即 $\boldsymbol{\Lambda}_\mathrm{h}\approx\boldsymbol{P}_\mathrm{h}^\mathrm{H}\boldsymbol{\Lambda}_\mathrm{H}\triangleq\boldsymbol{\Lambda}_\mathrm{P}$。

于是，通过单元细分的方式，每个粗网格单元 k 的局部误差 e_k 和目标估计的全局误差 E 为

$$\begin{cases}e_k=\displaystyle\sum_{N(k)}\left|\left(\boldsymbol{\Lambda}_\mathrm{P}-\boldsymbol{\Lambda}_\mathrm{h}^\mathrm{H}\right)^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{Q}_\mathrm{h}^\mathrm{H})\right|_{N(k)}\\ E=\displaystyle\sum e_k\end{cases} \tag{8}$$

式中：N 为粗网格单元 k 对应区域的加密网格单元数。

最后，结合局部和全局自适应参数，给定一个综合自适应参数：

$$\eta=\frac{e_k}{t}\cdot\frac{E}{E_{\mathrm{TOL}}}>1 \tag{9}$$

式中: E_{TOL} 为用户设定的目标函数全局误差容忍上限; t 为粗网格单元的平均局部误差上限, $t=E_{\mathrm{TOL}}/N$。当 $\eta>1$ 时, 认为粗网格单元 k 需要被加密, 直至自适应参数满足条件后, 结束加密, 得到最终网格。

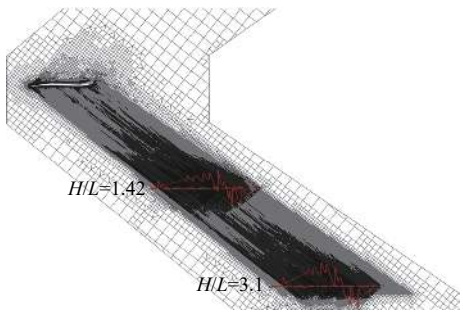
3　算例验证

为验证数值计算方法及网格生成技术在发动机喷流与激波相互干扰数值模拟中的可行性与准确性, 本节选取了 2 个有风洞试验数据的模型进行算例验证。

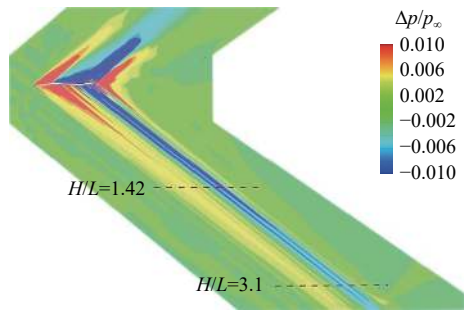
3.1　LM1021 模型

本节对图 1 所示的 LM1021 模型进行数值模拟, 计算来流马赫数为 1.6, 测量机身下方 H/L=1.42 和 H/L=3.1 处的过压信号, 并与风洞试验结果[35] 进行对比验证。

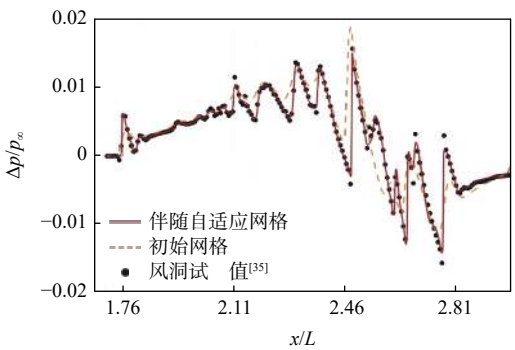
图 4(a)、(b) 分别为伴随自适应得到的最终网格及流场过压值云图, 最终得到的网格量约为 1 500 万。可以看出, 在指定平面附近激波处都进行了网格的自适应加密, 从而较好地捕捉到了流场中的激波与膨胀波系。



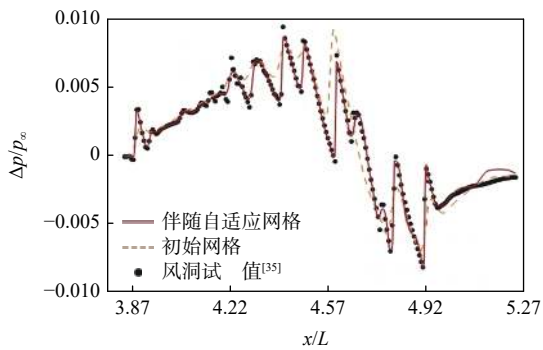
(a) 自适应加密后的网格



(b) 过压值云图



(c) H/L=1.42处过压信号



(d) H/L=3.1处过压信号

图 4　LM1021 模型计算结果

Fig. 4　Computational results for LM1021 model

图 4(c)、(d) 分别将模型下方 H/L=1.4 及 H/L=3.1 处的过压值与试验结果进行对比, 结果表明, 相比于初始网格, 基于伴随的自适应网格与试验值基本重合, 对激波处压力跃升的捕捉有明显改善, 具有较好的计算精度。因此, 上述数值方法和网格生成策略对于超声速流场的数值模拟是可行且可信的, 能够满足近场过压信号的计算精度要求。

3.2　Biconvex 模型

本节选取第三届 AIAA 声爆预测会议提供的带有前锥体的喷管与双凸面平尾组成的 Biconvex 模型[36], 用于验证上述数值方法和网格生成方法对超声速流场中发动机喷流模拟的可行性与准确性。该模型总长 568.93 mm, 喷管直径为 38 mm, 置于喷管上方的平尾翼根、翼梢弦长及展长分别为 38 mm、13.33 mm 及 22.86 mm。计算来流马赫数为 1.6, 喷管落压比为 8, 测量机体正下方 H/L=0.67 处过压信号。Biconvex 几何外形及伴随自适应得到的最终网格如图 5 所示, 网格量约为 300 万。

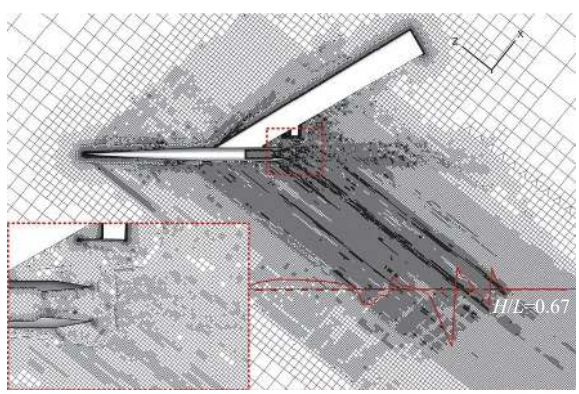
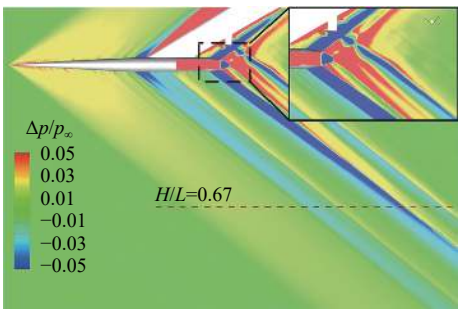


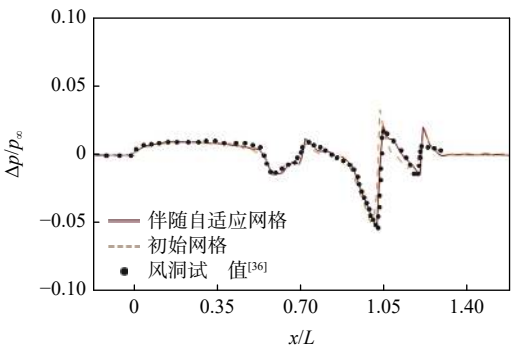
图 5　Biconvex 几何外形及伴随自适应得到的最终网格

Fig. 5　Biconvex geometry and its adaptively refined mesh by adjoint method

图6为计算得到的对称面过压值云图和H/L=0.67处过压信号与试验值[36]的对比。从图6(a)可以看出，尾翼后缘膨胀波在向下传播时与喷管及喷流带来的复杂波系结构产生了强烈的相互干扰，对声爆特性造成影响。图6(b)表明，在1.05<x/L<1.32之间的区域，初始网格与试验值稍有差别，但是伴随自适应后的网格得到的结果与试验值吻合良好，较精确地捕捉到了喷流与尾翼激波相互干扰带来的影响，验证了算法的计算精度。



(a) 过压值云图



(b) H/L=0.67处过压信号

图6 Biconvex模型计算结果

Fig. 6 Calculation results of biconvex model

4 计算结果及分析

对表1所示的各个构型展开数值模拟分析，通过对比其近场过压信号来分析发动机短舱布局和喷流对声爆强度的影响。发动机风扇入口和尾喷口边界条件参考第二届AIAA声爆预测会议提供的C25D发动机边界条件[37]，详细的来流条件及发动机动力参数如表2所示，迎角为0°。

4.1 发动机进气道及尾喷流的影响

对比分析有无发动机喷流的结果，以探究尾喷流对流场声爆强度的影响。图7(a)、7(b)为有无喷流2种情况下的过压值云图，可以发现，发动机尾喷流对前半机身下方的激波与膨胀波系基本没有影响，而后半部分则受到了显著影响，形成了复杂的流场结构。图7(c)、7(d)给出的尾部区域的马赫数云图展示了其复杂性，从左向右对对称面上的激波与膨胀波系依次进行分析：首先，机翼下方的发动机进气道唇口激波与机翼下表面和后半机身引起的膨胀波相互干扰；其次，尾翼前缘激波与翼下发动机尾喷管后缘激波，在小段距离后两者合并成一道强激波，削弱并截断了上翘的机身尾部引起的膨胀波，尾翼的后半部分都在该激波区域的影响范围内，因此对称面上尾翼引起的膨胀波也被其削弱；然后，机尾发动机喷管后缘激波与机身尾部激波、尾翼后缘激波合并成一道激波；最后，欠膨胀状态的喷流导致膨胀波与激波系，但是由于之后网格尺寸较大，捕捉不再精确。

表2 自由来流条件及发动机动力参数

Table 2 Conditions of freestream and propulsion system parameters

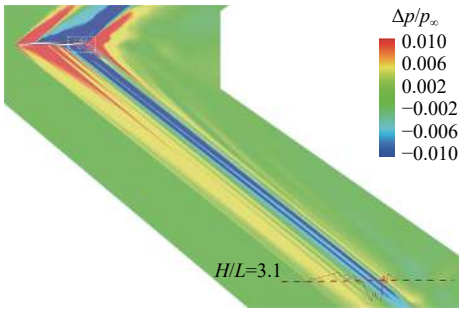
参数	数值
来流马赫数	1.6
来流静温/K	211.96
发动机风扇入口静压比	3.260 6
发动机风扇入口密度比	14.54
发动机风扇入口马赫数	0.609 3
雷诺数	8.1×10^6
来流静压/Pa	25 894.936
发动机尾喷口落压比	14.54
发动机尾喷口总温比	7.872 2
发动机尾喷口马赫数	0.662 0

图8更直观地对比了2个构型在H/L=3.1处的过压信号。可以发现，在x/L<4.82区域并未受到发动机喷流的影响，2个工况下的过压信号分布几乎重合，发动机喷流与翼身组合体带来的激波与膨胀波系相互干扰的区域主要集中在x/L>4.82区域。喷流的引入明显增大了过压值，尤其是在x/L=4.89处，由于上述较强的喷管后缘激波的加入，过压值达到峰值。另外，尽管在距离机身较近的尾部区域有着复杂的流场结构，但是随着距离的增加，激波与膨胀波逐渐削弱和相互合并，在H/L=3.1截面只形成了x/L=4.89处一道强激波。而对于无喷流构型，在x/L=4.99和x/L=5.06处还能明显捕捉到机身尾部激波和后体激波。对于x/L>5.06区域，无喷流构型的过压值基本不再变化，而带有喷流的构型，则由于欠膨胀尾喷流的存在而产生额外的激波系，声爆强度必定大于无喷流的情况。

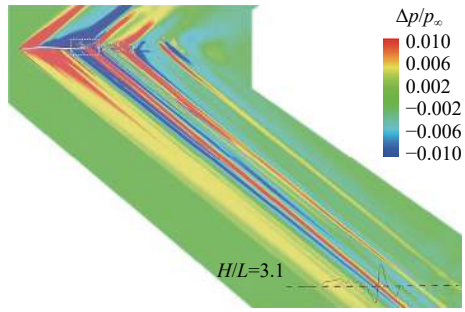
4.2 发动机短舱位置的影响

4.2.1 弦向位置

对于弦向位置的影响，将4.1节中带喷流构型作为基础构型，分别将机翼下方的2个发动机沿轴向向前、向后移动0.7倍的喷管直径，得到构型D3−Δx与D3+Δx，图9为2种构型的近场过压值与

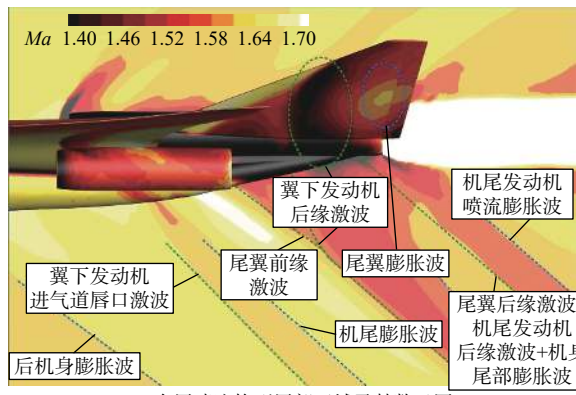


(a) 无尾喷流构型过压值云图



(b) 有尾喷流构型过压值云图

(c) 无尾喷流构型尾部区域马赫数云图



(d) 有尾喷流构型尾部区域马赫数云图

图 7　有无尾喷流构型近场过压值与马赫数云图

Fig. 7　Near-field over pressure and Mach number contours of configurations without and with powered nacelles

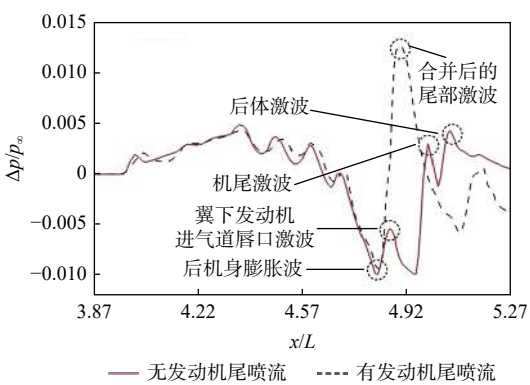


图 8　有无尾喷流构型在 H/L=3.1 处过压信号对比

Fig. 8　Comparison of pressure signatures at H/L=3.1 between configurations without and with powered nacelles

尾部区域马赫数云图, 图 10 给出了 2 种构型在 H/L=3.1 处的过压信号, 并与基础构型进行对比。结果显示, 前半部分流场结构基本没有变化, 主要的影响区域同样集中在后半流场。

对于短舱前移的 D3−Δx 构型, 如图 9(c) 所示, 此时翼下短舱进气道唇口激波并没有与后方的激波合并, 在监视截面 x/L=4.82 处的后机身膨胀波之后出现了一道较小的激波。随后, 喷管后缘激波与尾翼前缘激波在机身下方合并成一道较强的激波, 即在 x/L=4.89 处形成了过压峰值(见图 10)。

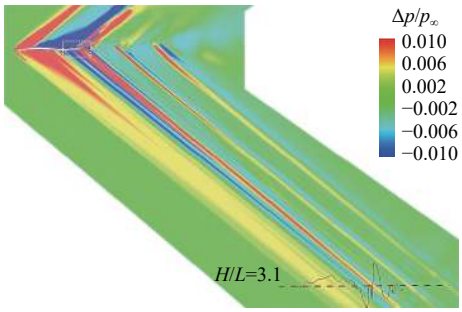
随着短舱的后移, 一方面, 进气道唇口激波后移, 使其与机翼下表面和后半机身的膨胀波之间的相互干扰减小, 图 10 中 x/L=4.82 处的后机身膨胀波强度随之增加。另一方面, 喷管后缘激波的影响区域也后移, 对尾翼前缘激波和尾翼膨胀波的影响都减小。因此, 在紧靠机身下方的尾翼前缘激波与翼下发动机后缘激波之间出现了一道由尾翼表面引起的较弱的膨胀波(如图 9(d) 中虚线方框所示), 但随后便被翼下短舱后缘激波截断。

随着短舱的后移, 其喷管后缘激波与机尾处发动机喷流带来的激波结构之间的相互干扰也会增强, 在向下传播的过程中, 其逐渐合并, 并促进与尾翼前后缘激波的合并, 从而导致对称面上尾部合并的激波位置前移和强度增加, 表现为图 10 所示的约 x/L=4.89 处的过压峰值的变化。

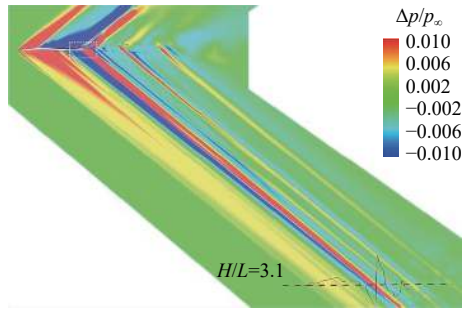
综上, 短舱前移有利于削弱后半机身和机翼带来的膨胀波, 避免尾翼前后缘激波与机身上方短舱后缘激波的合并, 进而降低后体激波强度, 因此, 能够降低声爆强度。

4.2.2　展向位置

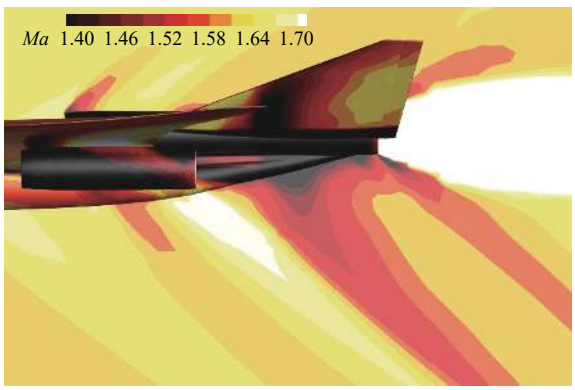
基于基础构型, 保证翼下短舱外露于机翼的长度不变, 沿展向分别内移 0.5 个喷管直径、外移 1 个与 2 个喷管直径, 得到 D3−0.5Δy、D3+Δy 与 D3+2Δy 构型, 计算结果如图 11 和图 12 所示。通过对比分析发现, 当翼下短舱距离机身较近时, 发动机进气道唇口激波对机翼与机身引起的膨胀波相互干扰明显(见图 11(a)), 于是在 x/L=4.82 膨胀波之后存在



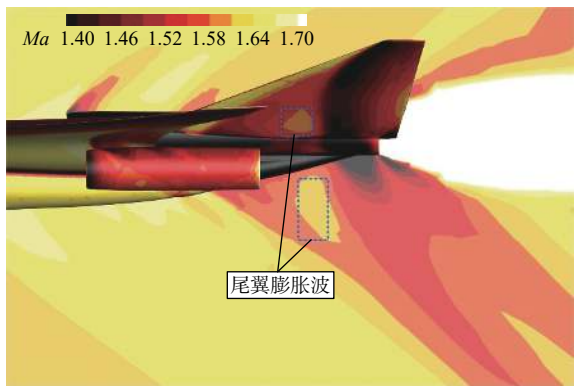
(a) 短舱前移D3−Δ*x*构型过压值云图



(b) 短舱后移D3+Δ*x*构型过压值云图



(c) 短舱前移D3−Δ*x*构型尾部区域　赫数云图



(d) 短舱后移D3+Δ*x*构型尾部区域　赫数云图

图 9　改变短舱弦向位置的 D3−Δ*x* 与 D3+Δ*x* 构型近场过压值与马赫数云图

Fig. 9　Near-field over pressure and Mach number contours of D3−Δ*x* and D3+Δ*x* configurations with different chordal nacelle layouts

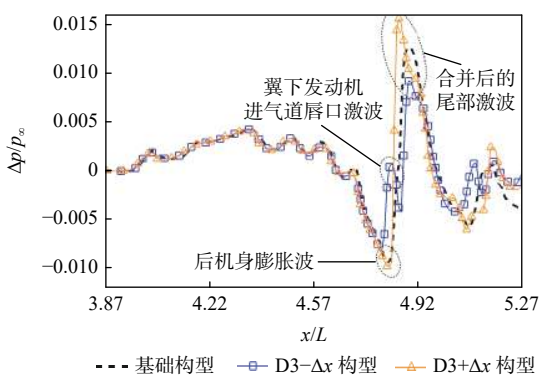


图 10　D3−Δ*x* 与 D3+Δ*x* 构型在 *H*/*L*=3.1 处的过压信号对比

Fig. 10　Comparison of pressure signatures at *H*/*L*=3.1 between D3−Δ*x* and D3+Δ*x* configurations

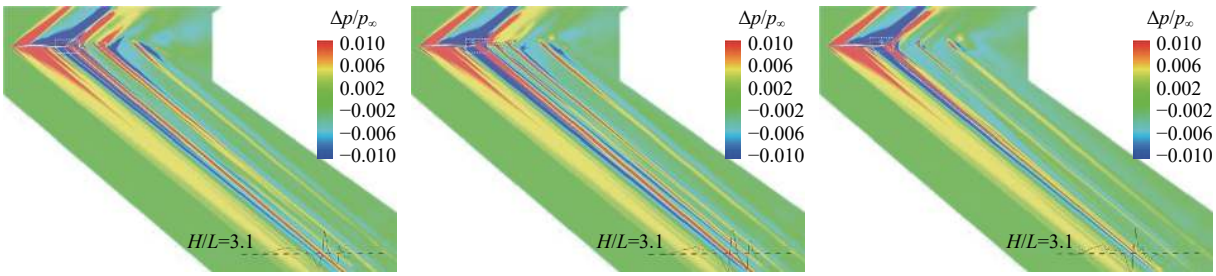
一个较小的唇口激波，随着翼下短舱的外移，该激波在对称面上的强度及对后机身膨胀波的影响逐渐减弱。

对于 D3+Δ*y* 构型，唇口激波影响区域移动至尾翼中后部分，与机尾激波合并后又与尾翼前缘激波再次合并为一道强激波。对于 D3+2Δ*y* 构型，翼下短舱后缘激波对尾翼的影响区域继续后移至尾翼后缘且强度降低，该激波与机尾激波合并后的激波强度相对减弱，以致其无法与尾翼前缘激波合并，尾翼引起的膨胀波区域增大（如图 11(b)中虚线方框所示），故图 12 中该构型下 *x*/*L*=4.82 膨胀波后也观察到一个较小的激波，但其是尾翼前缘激波。对于约 *x*/*L*=4.89 处合并后的尾部激波，D3−0.5Δ*y* 构型的喷管距离机身最近而激波强度最强，而 D3+Δ*y* 构型的机尾处由于存在多个激波合并导致激波强度要大于基础构型，但仍小于喷管内移构型，D3+2Δ*y* 构型喷管后缘激波在对称面中影响最小，强度最弱。

综上，将翼下短舱沿展向向外移动一定距离，使机尾各激波不再合并，有利于降低后体激波强度，从而降低声爆强度。

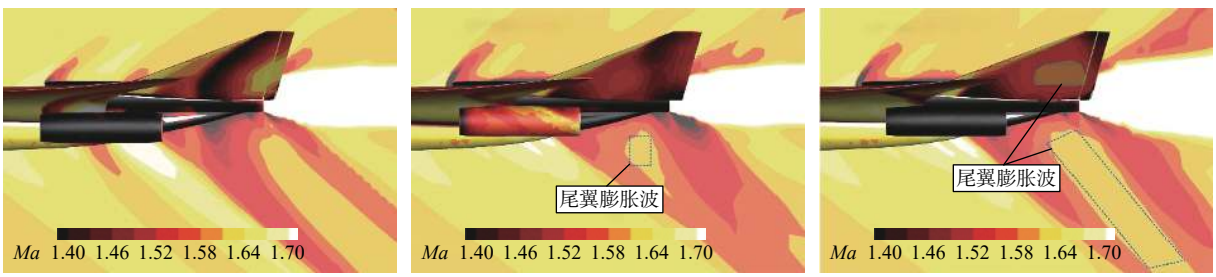
4.2.3　翼上式布局

翼上式短舱布局 U3 将机翼处的短舱放置于机翼上方，与基础构型相比，不改变机翼与短舱的展向、弦向位置及纵向距离，图 13(a)、(b) 为其近场过压值与尾部区域马赫数云图，图 14 对比了翼上式与基础翼下式 2 个构型在 *H*/*L*=3.1 处的近场过压信号。由于发动机放于机翼上方，发动机进气道唇口激波被机翼阻隔，机翼及后半机身引起的膨胀波没有受到影响，在图 14 中体现为在 *x*/*L*=4.82 附近完整的膨胀波。对比图 7(d) 和图 13 (b) 所示两者尾部区域的马赫数云图，发动机上置后短舱与尾翼的间距减小，此时尾翼同时受到翼上短舱后缘激波及其尾喷流的影响，从而完全截断了尾翼表面引起的膨胀波在对称面上的影响。另外，由于短舱上置，其后缘激波对机身正下方的影响明显减弱，于是图 14 所示的近场过压信号在 *x*/*L*=4.89 的后体激波强度降低。因此，翼上式布局有利于降低声爆强



D3−0.5Δ*y*构型　　D3+Δ*y*构型　　D3+2Δ*y*构型

(a) 近场过压值云图



D3−0.5Δ*y*构型　　D3+Δ*y*构型　　D3+2Δ*y*构型

(b) 尾部区域　赫数云图

图 11　改变短舱展向位置的 D3−0.5Δ*y*、D3+Δ*y* 与 D3+2Δ*y* 构型近场过压值与马赫数云图

Fig. 11　Near-field over pressure and Mach number contours of D3−0.5Δ*y*, D3+Δ*y* and D3+2Δ*y* configurations with different spanwise nacelle layouts

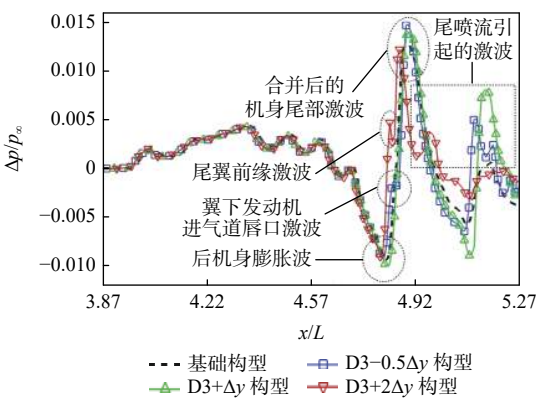
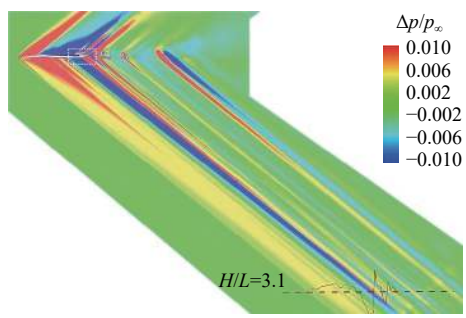
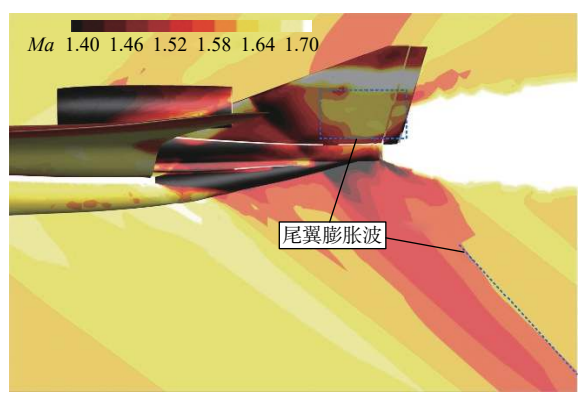


图 12　D3−0.5Δ*y*、D3+Δ*y* 与 D3+2Δ*y* 构型在 *H*/*L*=3.1 处的过压信号对比

Fig. 12　Comparison of pressure signatures at *H*/*L*=3.1 between D3−0.5Δ*y*, D3+Δ*y* and D3+2Δ*y* configurations



(a) 翼上式布局U3构型过压值云图



(b) 翼上式布局U3构型尾部区域　赫数云图

图 13　翼上式布局 U3 构型近场过压值与马赫数云图

Fig. 13　Near-field over pressure and Mach number contours of U3 configuration with nacelles mounted on upper wing surface

度, 但该布局由于不能充分利用进气道唇口激波与机翼下表面之间的有利干扰而气动性能有所削弱, 且尾喷流会直接影响到尾翼, 需要对结构布置及热防护额外考虑, 此外舵效也会受到影响。

4.3　发动机数量的影响

在保证总推力相等的前提下, 将基础构型机尾处的发动机去除, 同时按第 1 节所述等比例放大进气道与喷口面积, 得到双发的 D2 构型以研究发动机数量对声爆强度的影响, 图 15 和图 16 给出了双发构型的流场计算结果和近场过压信号。由于去除了机尾的发动机, 在过压值云图中, 双发构型的机身尾部区域激波与膨胀波系发生了很大变化。进一步对比图 7(d) 与图 15(b) 2 个构型的尾部区域马赫数云图发现, 双发构型更大的短舱尺寸使得进气道唇口激波强度增加, 并与机身腹部产生的激波合并, 于是在图 16 所示 *H*/*L*=3.1 处的近场过压信号

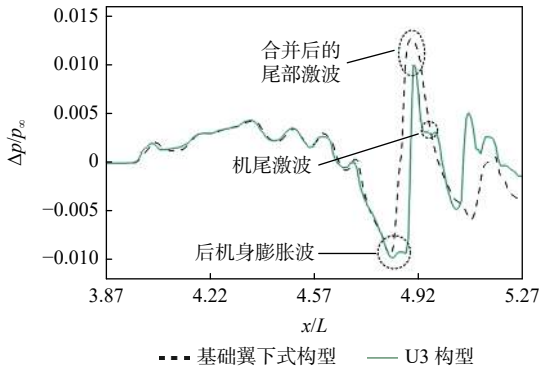
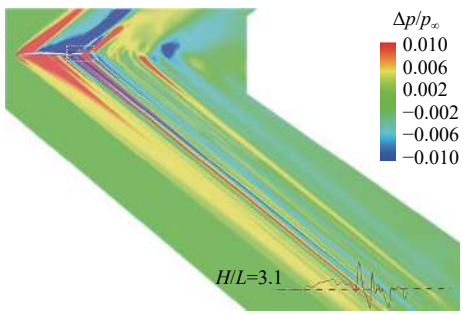
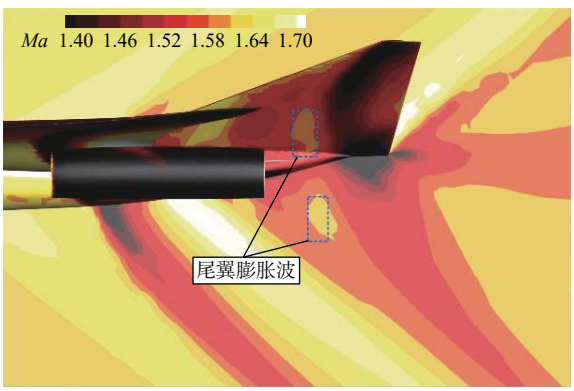


图 14 翼上式布局 U3 构型在 H/L=3.1 处的过压信号

Fig. 14 Pressure signatures at H/L=3.1 of U3 configuration with nacelles mounted on upper wing surface



(a) 双发布局D2构型过压值云图



(b) 双发布局D2构型尾部区域 赫数云图

图 15 双发布局 D2 构型近场过压值与马赫数云图

Fig. 15 Near-field over pressure and Mach number contours of D2 configuration with two nacelles

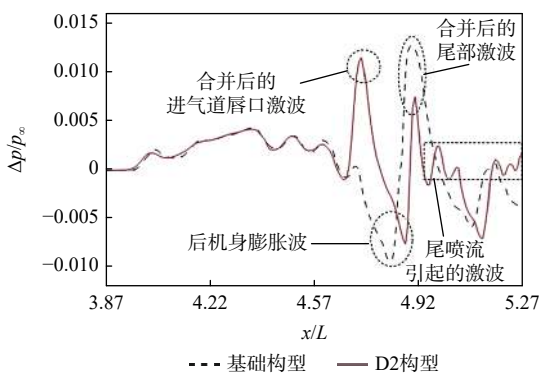


图 16 双发布局 D2 构型在 H/L=3.1 处的过压信号

Fig. 16 Pressure signatures at H/L=3.1 of D2 configuration with two nacelles

x/L=4.73 处出现了一道较强的激波，并直接截断了机翼及后半机身引起的膨胀波。此外，由于双发构型下，喷流的欠膨胀状态更严重，使得尾翼上受其影响的区域后移且范围更小，喷流带来的激波结构主要影响尾翼后缘激波。

另外，由于机尾处喷流的去除，双发构型下机尾处复杂的激波结构也随之消失。因此，如图 16 所示，双发构型在 x/L=4.89 处的后体激波强度降低。综上，尽管双发构型减小了后体激波，但是又带来了较强的进气道唇口激波；考虑到双发构型的发动机尺寸与噪声更大，气动阻力也更大。因此，翼下双发布局对于新一代超声速民机并非最佳选项。

5 结 论

1）发动机进气道及尾喷流的引入，带来了进气道唇口激波、尾喷口后缘激波及喷流引起的激波和膨胀波，与主流激波结构存在强烈的相互干扰，增加近场过压幅值，声爆强度将随之增加，但可以通过合理的短舱布局有效降低声爆强度。

2）短舱沿弦向前移与沿展向外移一定距离，都能有效削弱后半机身和机翼带来的膨胀波，抑制后体激波的合并，延长激波之间的间隔，从而降低声爆强度。对于翼上式布局，由于机翼的遮挡和短舱后缘激波影响区域的减小，声爆强度能够得到显著的抑制。

3）保证推力近似相同的前提下，双发构型能有效降低后体激波强度，但是会额外形成一道较强的进气道唇口激波；考虑到双发构型尺寸、噪声更大，气动阻力也更大，翼下双发布局对于新一代超声速民机并非最佳选项。

参考文献（References）

[1] 朱自强, 兰世隆. 超声速民机和降低音爆研究[J]. 航空学报, 2015, 36(8): 2507-2528.
ZHU Z Q, LAN S L. Study of supersonic commercial transport and reduction of sonic boom[J]. Acta Aeronautica et Astronautica Sinica, 2015, 36(8): 2507-2528(in Chinese).

[2] 钱战森, 韩忠华. 声爆研究的现状与挑战[J]. 空气动力学学报, 2019, 37(4): 601-619.
QIAN Z S, HAN Z H. Progress and challenges of sonic boom research [J]. Acta Aerodynamica Sinica, 2019, 37(4): 601-619(in Chinese).

[3] 王迪, 冷岩, 杨龙, 等. 基于广义Burgers方程的声爆传播特性大气湍流影响研究[J]. 航空学报, 2023, 44(2): 626318.
WANG D, LENG Y, YANG L, et al. Atmospheric turbulence effects on sonic boom propagation based on augmented Burgers equation[J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(2): 626318 (in Chinese).

[4] 王迪, 钱战森, 冷岩. 广义Burgers方程声爆传播模型高阶格式离散研究[J]. 航空学报, 2022, 43(1): 124916.

WANG D, QIAN Z S, LENG Y. High-order scheme discretization of the sonic boom propagation model based on augmented Burgers equation[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(1): 124916(in Chinese).

[5] 冯晓强, 李占科, 宋笔锋. 超声速客机音爆问题初步研究[J]. 飞行力学, 2010, 28(6): 21-23.
FENG X Q, LI Z K, SONG B F. Preliminary analysis on the sonic boom of supersonic aircraft[J]. Flight Dynamics, 2010, 28(6): 21-23 (in Chinese).

[6] 王刚, 马博平, 雷知锦, 等. 典型标模音爆的数值预测与分析[J]. 航空学报, 2018, 39(1): 169-181.
WANG G, MA B P, LEI Z J, et al. Simulation and analysis for sonic boom on several benchmark cases[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(1): 169-181(in Chinese).

[7] 乔建领, 韩忠华, 丁玉临, 等. 基于非线性Burgers方程的超声速客机远场声爆预测方法[C]//首届中国空气动力学大会, 2018.
QIAO J L, HAN Z H, DING Y L, et al. Sonic boom prediction method for supersonic transports based on nonlinear Burgers equation[C]//The First Chinese Conference of Aerodynamics, 2018(in Chinese).

[8] 乔建领, 韩忠华, 丁玉临, 等. 基于广义Burgers方程的超声速客机远场声爆高精度预测方法[J]. 空气动力学学报, 2019, 37(4): 663-674.
QIAO J L, HAN Z H, DING Y L, et al. Sonic boom prediction method for supersonic transports based on augmented Burgers equation[J]. Acta Aerodynamica Sinica, 2019, 37(4): 663-674(in Chinese).

[9] 张绎典, 黄江涛, 高正红. 基于增广Burgers方程的音爆远场计算及应用[J]. 航空学报, 2018, 39(7): 122039.
ZHANG Y D, HUANG J T, GAO Z H. Far field simulation and applications of sonic boom based on augmented Burgers equation[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(7): 122039(in Chinese).

[10] 陈鹏, 李晓东. 基于Khokhlov-Zabolotskaya-Kuznetsov方程的声爆频域预测法[J]. 航空动力学报, 2010, 25(2): 359-365.
CHEN P, LI X D. Frequency domain method for predicting sonic boom propagation based on Khokhlov-Zabolotskaya-Kuznetsov equation[J]. Journal of Aerospace Power, 2010, 25(2): 359-365(in Chinese).

[11] 冯晓强, 宋笔锋, 李占科. 低声爆静音锥设计方法研究[J]. 航空学报, 2013, 34(5): 1009-1017.
FENG X Q, SONG B F, LI Z K. Research of low sonic boom quiet spike design method[J]. Acta Aeronautica et Astronautica Sintca, 2013, 34(5): 1009-1017(in Chinese).

[12] 李占科, 刘秧, 丁玉临, 等. 静音锥对超声速民机低声爆效果的影响[J]. 西北工业大学学报, 2019, 37(1): 203-210.
LI Z K, LIU Y, DING Y L, et al. Influence of quiet spike on supersonic transport for low boom effect[J]. Journal of Northwestern Polytechnical University, 2019, 37(1): 203-210(in Chinese).

[13] 冯晓强, 李占科, 宋笔锋. 超声速客机低音爆布局反设计技术研究[J]. 航空学报, 2011, 32(11): 1980-1986.
FENG X Q, LI Z K, SONG B F. A research on inverse design method of a lower sonic boom supersonic aircraft configuration[J]. Acta Aeronautica et Astronautica Sinica, 2011, 32(11): 1980-1986 (in Chinese).

[14] 韩忠华, 许晨舟, 乔建领, 等. 基于代理模型的高效全局气动优化设计方法研究进展[J]. 航空学报, 2020, 41(5): 25-65.
HAN Z H, XU C Z, QIAO J L, et al. Recent progress of efficient global aerodynamic shape optimization using surrogate-based approach[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(5): 25-65(in Chinese).

[15] 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).

[16] 乔建领, 韩忠华, 宋文萍. 基于代理模型的高效全局低音爆优化设计方法[J]. 航空学报, 2018, 39(5): 121736.
QIAO J L, HAN Z H, SONG W P. An efficient surrogate-based global optimization for low sonic boom design[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(5): 121736(in Chinese).

[17] 黄江涛, 张绎典, 高正红, 等. 基于流场/声爆耦合伴随方程的超声速公务机声爆优化[J]. 航空学报, 2019, 40(5): 122505.
HUANG J T, ZHANG Y D, GAO Z H, et al. Sonic boom optimization of supersonic jet based on flow/sonic boom coupled adjoint equations[J]. Acta Aeronautica et Astronautica Sinica, 2019, 40(5): 122505(in Chinese).

[18] CAPONE F J, PUTNAM L E. Experimental determination of equivalent solid bodies to represent jets exhausting into a Mach 2.20 externalstream:NASA-TN-D-5553[R]. Washington,D.C.:NASA,1969.

[19] BUI T. CFD analysis of the nozzle jet plume effects on sonic boom signature: AIAA 2009-1054[R]. Reston: AIAA, 2009.

[20] CASTNER R. Analysis of plume effects on sonic boom signature for isolated nozzle configurations: AIAA 2008-3729[R]. Reston: AIAA, 2008.

[21] CASTNER R. Analysis of exhaust plume effects on sonic boom for a 59-degree wing body model: AIAA 2011-917[R]. Reston: AIAA, 2011.

[22] HOWE D. Engine placement for sonic boom mitigation: AIAA-2002-0148[R]. Reston: AIAA, 2002.

[23] CASTNER R, LAKE T. Exhaust plume effects on sonic boom for a delta wing and swept wing-body model: AIAA 2012-1033[R]. Reston: AIAA, 2012.

[24] 冯晓强, 李占科, 宋笔锋, 等. 基于混合网格的声爆/气动一体化设计方法研究[J]. 空气动力学学报, 2014, 32(1): 30-37.
FENG X Q, LI Z K, SONG B F, et al. Optimization of sonicboom and aerodynamic based on structured/unstructured hybrid grid[J]. Acta Aerodynamica Sinica, 2014, 32(1): 30-37(in Chinese).

[25] KIM H, LEE M. Flow simulation of a supersonic airplane with installed engine nacelle[J]. Aerospace Science and Technology, 2021, 117: 106900.

[26] CLIFF S E, DURSTON D, CHAN W M, et al. Computational and experimental assessment of models for the first AIAA sonic boom prediction workshop: AIAA 2014-0560[R]. Reston: AIAA, 2014.

[27] 廉筱纯, 吴虎. 航空发动机原理[M]. 西安: 西北工业大学出版社, 2005.
LIAN X C, WU H. Principle of aeroengine[M]. Xi'an: Northwestern Polytechnical University Press, 2005(in Chinese).

[28] ROE P L. Approximate riemann solvers, parameter vectors, and difference schemes[J]. Journal of Computational Physics, 1981, 43(2): 357-372.

[29] VENKATAKRISHNAN V. Convergence to steady state solutions of the Euler equations on unstructured grids with limiters[J]. Journal of Computational Physics, 1995, 118(1): 120-130.

[30] 吕凡熹, 肖天航, 余雄庆. 基于自适应直角网格的二维全速势方程有限体积解法[J]. 计算力学学报, 2016, 33(3): 424-430.
LYU F X, XIAO T H, YU X Q. A finite volume method for 2D full-potential equation on adaptive Cartesian grids[J]. Chinese Journal of Computational Mechanics, 2016, 33(3): 424-430(in Chinese).

[31] 吕凡熹, 李正洲, 邓经枢, 等. 面向飞行器概念设计的全速域气动分析工具[J]. 空气动力学学报, 2017, 35(5): 625-632.
LYU F X, LI Z Z, DENG J S, et al. An aerodynamic analysis tool for aircraft conceptual design at full speed range[J]. Acta Aerodynamica Sinica, 2017, 35(5): 625-632(in Chinese).

[32] WARREN G, ANDERSON W, THOMAS J, et al. Grid convergence for adaptive methods: AIAA-91-1592[R]. Reston: AIAA, 1991.

[33] 李立, 白文, 梁益华. 基于伴随方程方法的非结构网格自适应技术及应用[J]. 空气动力学学报, 2011, 29(3): 309-316.
LI L, BAI W, LIANG Y H. An adjoint-based method for unstructured mesh adaptation and its applications[J]. Acta Aerodynamica Sinica, 2011, 29(3): 309-316(in Chinese).

[34] 朱震浩, 肖天航, 徐雅楠, 等. 基于直角网格伴随自适应的声爆预测[J/OL]. 北京航空航天大学学报, 2022(2022-01-25) [2022-04-01]. https://link.cnki.net/urlid/11.2625.V.20220124.1709.004.
ZHU Z H, XIAO T H, XU Y N, et al. Adjoint-based adaptive cartesian mesh refinement for sonic boom prediction[J/OL]. Journal of Beijing University of Aeronautics and Astronautics, 2022(2022-01-25)[2022-04-01]. https://link.cnki.net/urlid/11.2625.V.20220124.1709.004(in Chinese).

[35] PARK M A, MORGENSTERN J M. Summary and statistical analysis of the first AIAA sonic boom prediction workshop[J]. Journal of Aircraft, 2016, 53(2): 578-598.

[36] WINSKI C S, CARTER M B, ELMILIGUI A A, et al. Computational and experimental study of plume and shock interaction effects on sonic boom in the NASA Ames 9×7 supersonic wind tunnel: AIAA 2018-0331[R]. Reston: AIAA, 2018.

[37] PARK M A, NEMEC M. Nearfield summary and statistical analysis of the second AIAA sonic boom prediction workshop[J]. Journal of Aircraft, 2019, 56(3): 851-875.

Effect of engine nacelle layout on sonic boom of supersonic transport

XIAO Tianhang*, XU Yanan, ZHU Zhenhao, DENG Shuanghou

(College of Aerospace Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: The strong aerodynamic interaction between the nacelles and the wing-body at supersonic speed, and the effect of the exhaust plume on the mainstream shock wave structure could have a great influence on the sonic boom loudness. This paper carries out a numerical investigation on the effect of nacelle layout on thesonic boom for a supersonic transport configuration where the supersonic flows and near field pressure signatures are computed by using a finite volume solver and an adjoint-based Cartesian adaptive mesh refinement method. Various relative positions of nacelles with respect to the wing andthe different number of nacelles are considered. Results indicate that the shock and expansion waves caused by the wing-body combination are significantly interfered with by the lip shock at the inlet, the nozzle trailing edge shock, and the expansion and aft shock of the nozzle plume. This increases the amplitude of the pressure signatures and subsequently increases the sonic boom loudness.Compared with the baseline, the ground loudness can be effectively suppressed by moving the nacelles forward along the chord, outward along the span-wise, and mounting above the wing. Under the same overall force and with the same number of engines and nacelles, the configuration with twin nacelles can reduce the aft shock, but the larger nacelle results in a stronger lip shock at the intake, increasing the loudness of the sonic boom. Thus, considering the engine noise and aerodynamic drag, it is not recommended to deploy twin nacelles for next-generation supersonic transport.

Keywords: supersonic transport; nacelle layout; sonic boom; adaptive mesh; Cartesian mesh

Received: 2021-11-16; **Accepted**: 2022-04-10; **Published Online**: 2022-04-21 12:26

URL: kns.cnki.net/kcms/detail/11.2625.V.20220420.1604.005.html

Foundation items: Opening Foundation of EDL Laboratory (EDL19092111); National Natural Science Foundation of China (11672133); Natural Science Basic Research Programof Shaanxi (2021JQ-078)

* **Corresponding author**. Email: xthang@nuaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0524

利用改进遗传算法的软件故障定位辅助测试用例生成方法

杨波[1, 2, 3]，何宇泽[3]，许福[1, 2, *]，陈志泊[1, 2]

(1. 北京林业大学 信息学院，北京 100083；2. 国家林业和草原局 林业智能信息处理工程技术研究中心，北京 100083；
3. 北方工业大学 信息学院，北京 100144)

摘 要：在软件故障自动化定位过程中，如果在给定的测试用例集合情况下故障在故障疑似度列表种排名不高，比较有效的方法为补充新的测试用例。如何在较小的代价下提升故障的可疑度排名是一项具有挑战的工作。提出一种基于改进遗传算法(IGA)的测试用例生成方法，利用软件故障定位的疑似故障排名，来辅助生成软件故障定位过程中的测试用例。依次对所提方法进行阐述和分析，在6个C程序和2个Python程序上开展实验，实验结果表明：所提方法自动生成的测试用例能够有效地帮助提高故障定位的效率。

关 键 词：故障定位；测试用例生成；遗传算法；适应度函数；随机测试
中图分类号：V221+.3; TB553
文献标志码：A **文章编号**：1001-5965(2023)09-2279-10

传统的手工软件测试用例生成在较大程度上依靠测试人员的经验和水平，这项工作费时费力。为此许多研究者提出了大量的自动化测试用例生成方法。目前已经有许多关于测试用例生成的研究，主要包括基于搜索的测试用例生成方法[1-3]、基于约束求解的测试用例生成方法[4-6]、基于需求的测试用例生成方法[7]、基于符号执行的测试用例生成方法[8]和随机测试(random testing, RT)用例生成方法[9]等。

基于搜索的测试用例生成方法来源于人工智能的研究。对于求解一些很难精确求解的问题，往往会把其转换成为一个在问题空间中进行搜索的过程。基于搜索的测试用例生成方法是根据以往搜索的结果所得到的信息，产生下一步搜索动作。搜索算法是处理NP-hard问题的有效途径。常用的算法有爬山法[10]、模拟退火算法[11]和遗传算法[12](genetic algorithm, GA)。GA是应用到测试用例生成中一种比较典型的算法，该算法模拟自然过程中的基因组合过程，通过将测试用例的属性与特征进行提取，转化为数学上的函数求解问题。在进行求解的过程中利用适应度函数，来对测试用例的好坏进行评价。然后进行多次迭代使得新生成的测试用例满足测试需求。许多研究者利用GA来生成测试用例，例如，文献[13]提出一种利用GA针对软件产品线的测试用例生成方法。该方法能够针对软件产品线的某种特定产品来生成测试用例，这样生成的测试用例还可以重用到其他同族的软件。文献[14]提出一种融合GA和禁忌搜索算法的测试用例生成方法，该方法既利用了GA的全局搜索能力，也考虑了具有局部搜索能力的禁忌搜索算法。文献[15]提出一种基于路径约束的测试用例方法，该方法利用遗传算法来对程序输入参数进行搜索。文献[16]提出一种自适应的蚁群算法来自动地生成测试用例，该方法利用了路径覆盖准则

收稿日期：2022-06-22；**录用日期**：2022-08-15；**网络出版时间**：2022-11-18 17:11
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20221117.1931.005.html
基金项目：中央高校基本科研业务费专项资金(BLX202003)；国家自然科学基金(61502011)；北京林业大学热点追踪项目(BLRD202124)
* **通信作者**. E-mail：xufu@bjfu.edu.cn

引用格式：杨波，何宇泽，许福，等. 利用改进遗传算法的软件故障定位辅助测试用例生成方法[J]. 北京航空航天大学学报，2023，49(9)：2279-2288.
YANG B，HE Y Z，XU F，et al. Using improved genetic algorithm for software fault localization aided test case generation[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2279-2288 (in Chinese).

来构建适应度函数。文献[17]针对多种测试覆盖准则，提出一种基于超空间的GA来解决测试用例优化的问题。文献[18]使用改进散点搜索策略，结合灰度预测进化算法自动生成路径覆盖的测试用例。

在测试用例生成过程中，大多数测试覆盖要求可以归结为一个或多个逻辑表达式，这些表达式就是一些约束条件。因此，产生达到覆盖要求的测试用例的问题就转化成为求解满足约束条件值的过程。根据约束问题中变量值域的不同，约束问题又可以分为布尔约束问题、有限约束问题和数值约束问题。文献[19]用约束求解的方法来生成测试数据，利用静态单一取值(static single-assignment, SSA)和控制依赖，将被测程序转换为一个约束系统，接着去寻找执行给定语句的路径，从而生成想要的测试数据。文献[20]利用迭代逼近法，求解满足选定路径上谓词的输入值。在每次迭代中执行与选定路径相关的语句，这样可以得到一个线性的约束集。接下来利用高斯消去法求解线性约束集来获得一个输入增量，再进行迭代，最终产生选定路径的测试数据。文献[21]利用自然语言处理的方法来生成验收测试的测试用例。文献[22]提出一种用于测试用例分配、选择和生成的自适应软件测试策略，其中利用了基于操作配置文件和覆盖范围的组合，主要用来实现被测程序的可靠性。基于符号执行的测试用例生成方法不是采用完全指定的输入值，而是抽象地将其表示为符号，从而能够分析出特定路径的约束条件和计算式[8]。基于符号执行的测试用例生成方法需要建立约束系统，而建立约束系统的方式主要有2种：一种是从前向后替换，另一种是从后向前替换。值得注意的是，基于符号执行的测试用例生成方法通常需要进行比较复杂的代数运算，并且很难处理一些情况，比如依赖输入变量的循环条件、数组元素下标和函数调用。这让基于符号执行的测试用例生成方法难以得到更加广泛的应用。

RT用例生成方法也是应用较多的自动测试用例生成技术[9]，RT用例生成方法主要考虑到手工进行的测试用例生成与测试人员的经验密切相关。RT用例生成方法则通过随机的方法，来自动地生成测试用例，这样开销会较小。

文献[23]提出一种自适应的RT用例生成方法。利用自适应方法，使得随机生成的测试用例能够比较均匀地分布在输入空间内。另外，也有一些研究对目前的方法进行了分析和评价，例如，文献[24]评估了目前用来进行测试用例生成的13种典型的进化算法，实验结果显示有些算法在某些问题上表现比其他算法好，但是没有一种算法能够在所有的问题上都表现最佳。文献[25]针对测试用例生成的搜索算法中，面临参数选择困难的问题，提出一种基于转变熵的测试用例生成方法。为了进一步分析各种搜索算法的参数选择，文献[26]对基于搜索的测试生成工具(EvoSuite)的1 200个超参数配置进行了调优。

目前已有的测试用例生成方法都取得了一定的效果，但是缺乏利用软件故障定位的信息来进行测试用例生成的工作。而在软件故障定位的过程中，存在大量与程序密切相关的信息，这些信息包括程序的语句执行情况、变量、控制流和数据流等，这些信息是否处于成功或失败执行的测试用例中，对于设计测试用例非常关键。尤其在已有测试用例难以对故障进行较高准确的定位，且需要补充新的测试用例时，这些信息将会提供测试用例生成的重要依据。另外，目前许多软件故障定位的实验程序中都有初始的测试用例，在利用这些测试用例进行故障定位时，发现有时候故障定位的效果不是很好，真实包含故障的语句有时候在疑似故障排名中不高。为此本文希望能够借用已有的故障定位信息，利用GA来进行搜索，自动生成能够促进故障定位进度的测试用例。

基于此，本文提出一种基于改进遗传算法(improved genetic algorithm, IGA)的测试用例生成方法，通过随机选取或生成少量初始测试用例，来作为IGA的初始种群。同时，利用这些初始种群的测试用例，来执行被测程序，获得执行后的程序频谱信息。接下来将程序频谱信息转化为覆盖矩阵和结果向量，然后通过XGBoost模型进行分析，可以获得被测程序所包含语句的疑似故障的重要性排名[27]。这些排名会应用在IGA的适应度函数构造当中。最后，利用IGA迭代生成新的测试用例。需要注意的是，本文工作是对文献[28]所做工作的扩展，主要是对其方法和实验的相关内容进行了细化。

1 研究方法

1.1 方法框架

本文方法框架如图1所示。从图可以看出，首先会获取初始的测试用例，然后利用这些测试用例去执行待测程序，这些待测程序已经提前插装好了。利用测试用例执行完待测程序之后，会得到程序中每条语句覆盖的信息，这样就可以得到覆盖矩阵与结果向量。

这些覆盖的矩阵与结果向量会经过XGBoost

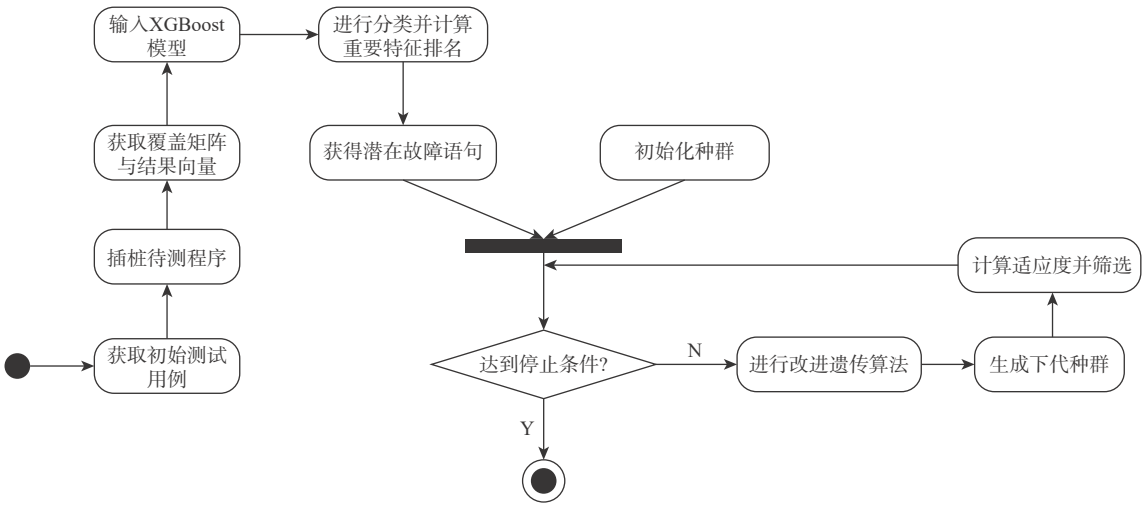


图 1　本文方法框架

Fig. 1　Framework of proposed method

模型进行训练, 训练的过程中会将每条语句作为其中的特征, 这样会得到一个语句的重要性排名。这个排名会显示出可能包含故障的语句列表, 即潜在故障语句。接下来将潜在故障语句与初始化总群放在一起。利用 IGA 生成新的测试数据, 这些测试数据会根据适应度函数来进行筛选, 最终获得生成的测试用例。

本文所提 IGA 框架中, 考虑了软件故障定位的信息, 这有可能可以帮助生成揭示故障的测试用例, 对于生成新的测试用例非常重要。但是, 利用已有基于 GA 的技术, 很难利用软件故障定位信息。为此, 对 GA 进行了改进, 在其适应度函数中加入了疑似故障语句的位置参数, 这样在生成新的测试用例时, 尽量提高对这些疑似故障语句的覆盖程度。

1.2　疑似故障语句排名计算

在利用测试用例执行程序 Q 之后, 会获得覆盖矩阵和结果向量, 这些信息会作为输入进入机器学习模型。模型训练结束之后, 利用虚拟测试用例矩阵作为预测数据输入模型, 来生成最终的预测结果, 这个结果会成为生成疑似故障语句的依据。本文选择 XGBoost 来作为训练的机器学习模型。

覆盖矩阵为一个数据集合, 该数据的每一行元素都可看成是一个样本。结果向量看成是该行元素的标签, 代表在这一行元素的样本下, 程序 Q 是执行成功还是执行失败的情况。

覆盖矩阵的每一行中的每一列元素都对应程序 Q 当中的一条语句。这条语句对于 XGBoost 模型而言, 是一个特征。XGBoost 模型会将这些特征进行分割, 同时将代表语句的特征进行排序。然后, XGBoost 模型会使用贪心算法, 对这些特征进行扫描, 得到所有的备选分割点。实际上, XGBoost 模型训练的过程可以看成是选择最优分割点的过程。

采用 XGBoost 模型中的 gain 准则, 来计算数据集合中特征的重要性, 特征越重要, 表示该特征对于程序中故障的贡献度越高, 这样包含这个特征的语句包含故障的可疑度也越高。

1.3　基于 IGA 的测试用例生成

GA 中的适应度函数可以用来评价种群中个体的分数。此外, 适应度函数可以与选择、交叉、变异操作进行结合, 从而引导种群的进化方向。

为了利用程序执行测试用例过程中产生的故障定位信息, 本文提出 IGA, 首先对 GA 的适应度函数进行更改, 主要体现在 IGA 的适应度函数, 包括 3 个参数, 分别为疑似故障语句与测试用例执行时所覆盖的语句的最小距离、条件覆盖数目和语句覆盖率。然后以最小距离、条件覆盖数目和语句覆盖率 3 个参数加权后相加来构造最终的适应度函数。

其次, IGA 还体现在遗传算子上。GA 中有一个比较重要的概念为种群的多样性, 种群多样性是种群中个体基因广泛分布的体现, 这是种群在整体进化过程中必须确保的一个关键方面。GA 在迭代的过程中, 会产生个体之间趋同的情况, 这样个体之间的相似度会增加, 从而影响 GA 的优化搜索能力。为此本文为了防止多样性下降, 减少早熟现象, 提出了动态变化的交叉和变异概率。即在 GA 进行进化的过程中, 可以根据个体的多样性, 动态地调整交叉和变异概率, 来保证 GA 的搜索性能。

1.3.1　适应度函数

IGA 的适应度函数包括 3 个参数, 分别为疑似

故障语句与测试用例执行时所覆盖的语句最小距离、条件覆盖数目和语句覆盖率，适应度函数由这3个参数进行加权来表示。

1）疑似故障语句与测试用例执行时所覆盖的语句最小距离

利用初始测试用例，执行被测程序之后，XGBoost模型会对特征的重要性进行分类和排名，并返回疑似故障语句的列表。列表中的每个元素都是一条潜在的包含故障的语句，其中包含故障的可疑程度依次递减。假设疑似故障语句中包含的语句 i 的列表长度为 l_i，则权重 w_i 可表示为

$$w_i = r_i \frac{1}{2l_i} \tag{1}$$

式中：r_i 为第 i 条语句的可疑度，r_0 为疑似故障语句列表的第1条语句，初始值取为1，疑似故障语句列表的下一条语句的疑似度为 $r_i - w_{i-1}$。

对于每一条疑似故障的语句 i，与潜在故障语句的距离分数为

$$d_{\text{score}} = \sum_{i=1}^{n} \frac{1}{1+\min(|d_i|)} w_i \tag{2}$$

式中：d_i 为覆盖语句与第 i 个疑似故障语句之间的距离；n 为疑似故障语句列表的长度。

2）条件覆盖数目

这里的条件指的是被测程序中分支语句中的各个部分。条件的值可以取真或假。条件覆盖数目指的是在给定的测试用例集执行之后，被测程序被覆盖到的条件数目。

3）语句覆盖率

本文的语句覆盖率和软件测试中的语句覆盖率是同一个概念，指的是给定的测试用例执行后，被测程序被覆盖的语句数，与被测程序总语句数的比值。

4）适应度函数

适应度函数由与潜在故障语句的距离、条件覆盖数目和语句覆盖率3个参数的加权来表示，对于GA中的种群个体，其计算式为

$$\text{fit}(s_i) = \alpha d_{\text{score}}(s_i) + \beta C_{\text{cov}}(s_i) + \gamma C_{\text{con}}(s_i) \tag{3}$$

式中：α、β、γ 为3个参数对应的权重，这个预先设定好，3个权重满足：$\alpha+\beta+\gamma=1$。根据不同的权重可以表示出3个参数不同的重要性。$C_{\text{cov}}(s_i)$ 为语句覆盖的数目；$C_{\text{con}}(s_i)$ 为条件覆盖的数目。

1.3.2 改进遗传算子的设计

1）种群多样性

种群多样性是种群中个体基因广泛分布的体现，这是种群在整体进化过程中需要保证的一个方面。种群多样性的计算式为

$$V = \frac{\sum_{i=1}^{N} (f_{s_i} - f_{\text{avg}})^2}{N} \tag{4}$$

式中：f_{s_i} 为种群个体 s_i 的适应函数值。另外，本文还考虑了种群中所有个体适应函数值的平均值，记为 f_{avg}，其计算式为

$$f_{\text{avg}} = \frac{\sum_{i=1}^{N} f_{s_i}}{N} \tag{5}$$

2）自适应的交叉概率

交叉是GA非常重要的部分，其把2个父体部分结构加以替换，生成新个体的操作，交叉概率 P_c 为其中的关键参数，能够在一定程度上影响种群的进化，最终影响GA的搜索能力。本文采用自适应的交叉概率，主要是希望能够根据父代种群的多样性来增加或保存迭代过程中的交叉概率。

自适应的交叉概率取值过程为：当目前种群中个体的适应度函数值小于上一代种群时，将考虑增加交叉概率值，具体是用基本交叉概率值加上概率变化量的值。当目前种群中个体的适应度函数值大于或等于上一代种群时，不改变交叉概率值。

3）自适应的变异算子

变异是GA种群进化的一种辅助方式，改变变异算子也能够在一定程度上影响种群的进化，最终影响GA的搜索能力。本文采用自适应的变异算子，主要是希望能够根据父代种群的多样性来增加或保存迭代过程中的变异概率。

首先按照变异概率，来从种群中选择一个个体。然后选择其染色体上的一个或多个基因点位进行变异操作。常见的变异包括Uniform变异、Boundary变异和基本位变异等，本文选择的是基本位变异。

1.3.3 IGA流程

基于改进的自适应函数、自适应的交叉概率与变异算子，可以得到IGA，其算法流程如下所示。

算法：基于IGA

输入：初始的种群 P、预先设置好的基本交叉概率值、交叉概率的变化量、基本的变异概率、变异概率的变化量、停止条件与最大迭代的次数。

输出：下一代种群 P'

1．初始化迭代的次数 round=0

2．将 P 赋给 P'

3．当round小于最大的迭代次数且不满足停止条件时，循环执行如下步骤：

3.1．对种群中的个体，计算适应度函数值

3.2．计算当前种群的多样性

3.3. 动态调整交叉概率和变异概率

3.4. 对种群中的个体进行编码

3.5. 生成下一代种群赋给P', 即测试数据。

4. 输出P'

从算法流程可以看出, 其输入为初始的种群、预先设置好的基本交叉概率值、交叉概率的变化量、基本的变异概率、变异概率的变化量。另外, 还有停止条件与最大迭代的次数。

算法的输出是下一代种群。利用 IGA 来进行迭代时, 每一轮的迭代过程都会对种群中的个体计算适应度函数值, 通过适应度来计算当前种群的多样性。

根据种群中个体的适应度函数值, 来动态调整交叉概率和变异概率。种群中的交叉概率能够在一定程度上影响种群的进化, 如果目前种群中个体的适应度函数小于上一代种群, 交叉概率值将会适当增加。反之则不改变, 这样做可以增加或保存迭代过程中的交叉概率。变异概率也能起到影响种群进化的效果。如果目前种群中个体的适应度函数小于上一代种群, 变异概率值将会适当增加。反之则不改变。接下来, 再进行迭代, 直到停止条件得到满足。

值得注意的是, 交叉概率和变异概率能够对种群的进化起到很关键的作用。如果交叉概率为 100%, 则后代将都由交叉产生。如果交叉概率为 0%, 则后代都是从旧种群的染色体中复制产生。另外, 如果变异概率为 100%, 则整个染色体都会发生变化。如果交叉概率为 0%, 则没有变化。

最后, 对种群中的个体进行编码, 就可以得到利用 IGA 生成的测试用例。

2 实　验

2.1 研究问题

问题 1　分别用初始的测试用例和新生成的测试用例, 去执行被测程序, 谁的覆盖率会更高一些。

问题 2　利用已有的故障定位方法, 分别用初始的测试用例和新生成的测试用例, 去进行故障定位, 谁的效果会更好一些。

2.2 评价准则

为了对 2 个研究问题进行分析, 提出如下评价准则:

1) 最高语句覆盖率

指的是给定的测试用例执行后, 被测程序被覆盖的语句数, 与被测程序总语句数的比值的最大值。

$$\text{sta_con}_{\max} = \max(\text{sta_con}) \tag{6}$$

2) 平均语句覆盖率

指的是给定的测试用例执行后, 被测程序被覆盖的语句数, 与被测程序总语句数的比值的平均值。

$$\text{sta_con}_{\text{avg}} = \text{average}(\text{sta_con}) \tag{7}$$

3) 平均条件覆盖数

指的是在给定的测试用例集执行之后, 被测程序被覆盖到的条件数目的平均值。

$$\text{con_num}_{\text{avg}} = \text{average}(\text{con_num}) \tag{8}$$

4) 平均故障语句覆盖率

指的是执行给定的的测试用例与所覆盖的故障数目的平均数与程序中包含故障数的比值。其计算式为

$$\text{fault_sta}_{\text{avg}} = \frac{\text{average}(\text{fault_sta})}{\text{all_fault_sta}} \tag{9}$$

5) 故障定位时需要检查的语句数

一般指的是提供的疑似故障列表中, 故障语句所在的行数。

2.3 实验设计

1) 实验对象

实验对象选择的是 SIR(software infrastructure repository) 中的 6 个程序 (见 https://sir.csc.ncsu.edu/php/) 和 2 个 Python 程序。SIR 中的 6 个程序都有初始的测试用例, 且里面包含的故障所在的语句也是确定的。另外, 选择的 2 个 Python 程序分别为字符串比较的程序 Difflib 与压缩程序 Zipfile, 其中 Difflib 代码行数为 1 937 行, Zipfile 代码行数为 2 440 行。表 1 为所选取的实验程序信息。

表 1　实验程序信息

Table 1　Experimental objects informations

程序	语句数	测试用例数目
Print_token1(Pt1)	563	4 130
Print_token2(Pt2)	508	4 115
Tcas	173	1 607
Schedule1(Sch1)	410	2 650
Schedule2(Sch2)	307	2 710
Tot_info(Tot)	406	1 052
Difflib	1 937	100
Zipfile	2 440	100

2) 实验设置

IGA 的实验设置经过多轮实验, 最终选取的参数为: 采用的是两点交叉, 基本交叉概率为 0.6, 交叉概率变化量为 0.2。变异算子选择的是单点位变异, 基本变异概率为 0.03, 变异概率变化量为 0.01。适应度函数的参数中, α=0.45, β=0.35, γ=0.2。迭代次数选择 15 轮。

对于问题 1, 将基于 IGA 与 GA[26]、RT[9, 27] 及基

于路径引导的模糊测试用例生成方法(简称 Fuzzy)[28]进行了对比。之所以选择与这些方法进行比较,主要是因为 RT 和 Fuzzy 是比较经典的测试用例生成方法。另外,基于 IGA 生成的测试用例也与 Siemens 本身提供的测试用例执行的效果进行了比较。

GA 选取的参数为:交叉概率为 0.7,变异概率为 0.04。适应度函数的参数中,α=0,β=0.6,γ=0.4。

对于问题 2,从 SIR 中的 6 个程序中,每个程序选择了其中具有代表性的故障版本数目 4 个。在这个版本的基础上分别利用初始的测试用例集和新生成的测试用例集,来比较最终故障定位的结果。

采用的故障定位方法有 4 个,分别为 XGBoost 故障定位方法[18]、D^{3}Star 故障定位方法[19]、GP19 故障定位方法[20]和 Ochiai1[21]故障定位方法。

2.4 实验结果及分析

1)对问题 1 的回答

对问题 1 的回答具体如图 2~图 5 所示,可以看出,基于 IGA 和 GA 生成的测试用例,无论是最高语句覆盖率,还是平均语句覆盖率、平均条件覆盖数、平均故障语句覆盖率,均表现得都要比 RT 和 Fuzzy 好。

具体而言,对于最高语句覆盖率和平均语句覆盖率,在大部分情况下,基于 IGA 要比基于 GA 生成的效果更好,所生成的测试用例作用在被测程序上产生的覆盖率更高。

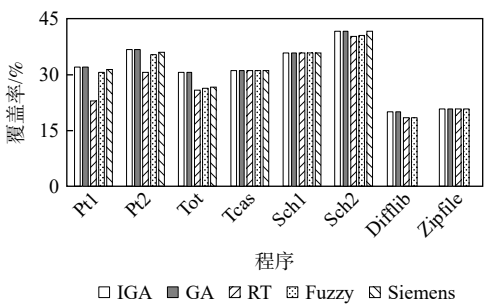


图 2 最高语句覆盖率

Fig. 2 Highest statement coverage

对于平均语句覆盖率,基于 IGA 生成的测试用例,同 SIR 中的 6 个程序所提供的测试用例相比,分别高 12.88%、11.31%、11.35%、4.04%、6.08% 和 5.99%。

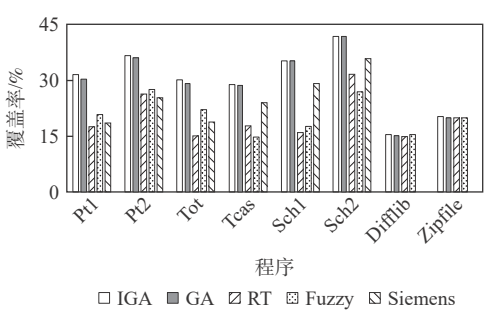


图 3 平均语句覆盖率

Fig. 3 Average statement coverage

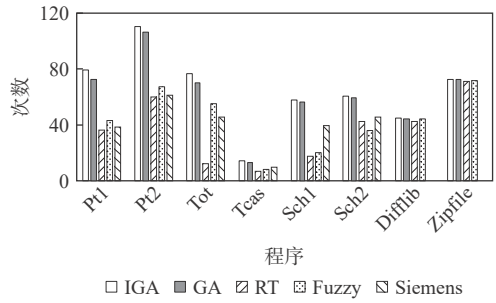


图 4 平均条件覆盖数

Fig. 4 Average condition coverage

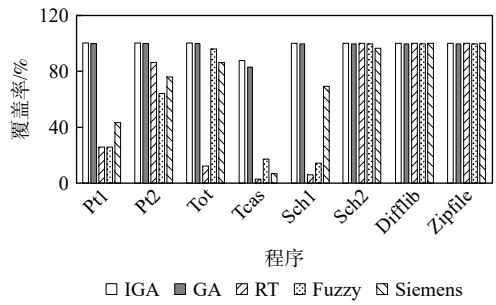


图 5 故障语句覆盖率

Fig. 5 Faulty statement coverage

为此,可以认为,基于 IGA 生成的测试用例,同其他 4 种方法产生的测试用例相比较,在最高语句覆盖率和平均语句覆盖率方面,效果要更好一些。

对于平均条件覆盖数目和故障语句覆盖率,在大部分情况下,基于 IGA 也要比基于 GA 生成的测试用例产生的效果更好,所生成的测试用例作用在被测程序上产生的覆盖率更高。

具体而言,对于平均条件覆盖数目,基于 IGA 生成的测试用例,同 SIR 中的 6 个程序所提供的测试用例相比,分别高 40.77%、49.37%、30.8%、4.34%、18.44% 和 14.65%。

对于故障语句覆盖率,基于 IGA 生成的测试用例,同 SIR 中的 6 个程序所提供的测试用例相比,分别高 56.18%、23.7%、13.47%、80.15%、30.72% 和 3.92%。

为此,可以认为,基于 IGA 生成的测试用例,同其他 4 种方法产生的测试用例相比,在平均条件覆盖数目和故障语句覆盖率方面,效果要更好一些。

对产生这样的结果进行进一步的分析,结果发现基于 IGA 生成的测试用例之所以能有较好的辅助故障定位的效果,主要原因是多次利用故障定位的疑似故障语句集合这一信息,这一信息在多轮故障定位过程中会出现变化,每次变化都会给下一次

的故障定位提供相关的依据。所以在一定程度上增加了平均条件覆盖和故障语句覆盖率。

2) 对问题 2 的回答

表 2～表 9 为基于 IGA 的测试用例生成方法所生成的测试用例, 与使用基准测试用例使用 XGBoost、D^{3}Star、GP19 和 Ochiai1 4 种故障定位方法的定位结果比较。

表 2　测试用例应用到故障定位的效率比较（Print_token1 程序）

Table 2　Comparison of the efficiency of test case applied to fault location(Print_token1)

故障定位方法	原始检查语句数	目前检查语句数	检查语句的减少率 /%
XGBoost	62	6	90.32
D^{3}Star	360	97	73.05
GP19	297	173	41.75
Ochiai1	297	97	67.34

表 3　测试用例应用到故障定位的效率比较（Print_token2 程序）

Table 3　Comparison of the efficiency of test case applied to fault location(Print_token2)

故障定位方法	原始检查语句数	目前检查语句数	检查语句的减少率/%
XGBoost	160	16	90.00
D^{3}Star	61	52	14.80
GP19	182	52	71.43
Ochiai1	64	61	4.69

表 4　测试用例应用到故障定位的效率比较（Tot_info 程序）

Table 4　Comparison of the efficiency of test case applied to fault location(Tot_info)

故障定位方法	原始检查语句数	目前检查语句数	检查语句的减少率/%
XGBoost	49	24	51.02
D^{3}Star	357	80	77.59
GP19	369	280	21.57
Ochiai1	357	80	77.59

表 5　测试用例应用到故障定位的效率比较（Schedule1 程序）

Table 5　Comparison of the efficiency of test case applied to fault location(Schedule1)

故障定位方法	原始检查语句数	目前检查语句数	检查语句的减少率/%
XGBoost	8	6	50.00
D^{3}Star	20	26	−30.00
GP19	128	139	−8.59
Ochiai1	20	26	−30.00

表 6　测试用例应用到故障定位的效率比较（Schedule2 程序）

Table 6　Comparison of the efficiency of test case applied to fault location(Schedule2)

故障定位方法	原始检查语句数	目前检查语句数	检查语句的减少率/%
XGBoost	132	41	68.94
D^{3}Star	92	16	82.61
GP19	92	16	82.61
Ochiai1	92	16	82.61

表 7　测试用例应用到故障定位的效率比较（Tcas 程序）

Table 7　Comparison of the efficiency of test case applied to fault location(Tcas)

故障定位方法	原始检查语句数	目前检查语句数	检查语句的减少率/%
XGBoost	59	5	91.52
D^{3}Star	72	32	55.00
GP19	84	60	28.57
Ochiai1	72	32	55.00

表 8　测试用例应用到故障定位的效率比较（Difflib 程序）

Table 8　Comparison of the efficiency of test case applied to fault location(Difflib)

故障定位方法	原始检查语句数	目前检查语句数	检查语句的减少率/%
XGBoost	51	51	0
D^{3}Star	136	47	65.44
GP19	136	47	65.44
Ochiai1	136	47	65.44

表 9　测试用例应用到故障定位的效率比较（Zipfile 程序）

Table 9　Comparison of the efficiency of test case applied to fault location(Zipfile)

故障定位方法	原始检查语句数	目前检查语句数	检查语句的减少率/%
XGBoost	121	93	23.14
D^{3}Star	336	146	56.55
GP19	336	146	56.55
Ochiai1	336	146	56.55

可以看出, 对于最好情况, 除了在 Schedule1 程序上定位结果略有下降(降低了 4 名)以外, 在其他各程序上均能保持原有的定位效果或带来一定提升。

对于平均情况和最差情况, 使用新测试用例进行故障定位的结果除了在 Schedule 的 2 个版本上有所下降(降低了 4 名和 9 名)外, 在其他版本上均

获得了较大幅度的提升。

造成下降的原因是基于IGA生成的测试用例能够覆盖的路径数量多于SIR中的6个程序所提供的测试用例，也就是说基于IGA生成的独特测试用例的种类更多，而这些测试用例存在巧合正确性，虽然覆盖了潜在故障语句，但是并没有触发故障，导致XGBoost模型判断时，将这些语句也列入潜在故障语句从而导致了故障定位需要检查的语句数的小幅下降。在Print_tokens1的平均情况和最差情况上，使用新测试用例进行定位的故障排名分别提高了56名(90.32%)和102名(63.75%)。对于所有程序版本，新测试用例在基本保证原有定位效果或带来较大提升的同时。

对结果进行进一步的分析，可以看出基于IGA生成的新测试用例能够有更大的概率覆盖错误的语句，这样使得在进行疑似故障的计算中，将这些包含故障的语句能够排到前列。

为此，可以认为，基于IGA生成的测试用例，在已采用的4种故障定位方法效果方面，产生了不错的效果。

2.5 影响实验有效性分析

基于IGA在2.4节中的实验中，与GA、RT、Fuzzy进行了对比，另外，基于IGA生成的测试用例也与Siemens本身提供的测试用例执行的效果进行了比较，取得更好的效果。另外，在利用新生成的测试用例进行故障定位的效果上，也比使用原始的测试用例取得了更好的效果。

但是，考虑到实验程序还比较少，并且故障的个数也有限，故后面还需要进行更多的实验，来进行进一步的验证。另外，尽管本文复现了GA、RT、Fuzzy等方法，但是与原始实现相比，可能也存在一定的差距。

3 结 论

1）IGA可通过改变传统GA的适应度函数、交叉概率和变异算子等参数，将其应用在软件的测试用例自动生成中。

2）在实验中，将本文的基于IGA所生成的测试用例与一般GA、RT、基于路径引导的模糊测试用例生成方法所生成的测试用例及SIR中的程序集所提供的基准测试用例进行对比，其最高覆盖率、平均覆盖率、故障语句覆盖次数和条件覆盖数目均具备一定优势，IGA能够以较小的迭代次数达到较好的效果，执行效率更高。

3）本文将新生成的测试用例用于4种已有的故障定位方法，在大部分情况下其故障定位效果要优于使用基准测试用例的定位效果，因此，可以说明本文的基于IGA测试用例生成方法能够有效辅助频谱故障定位。

本文将机器学习模型XGBoost和GA进行了组合，这种组合主要解决的是2个问题，分别是利用XGBoost模型来获取疑似语句的重要性排名，用GA去进行测试用例的生成，这种组合可能会有多种方式，也进行了多种尝试，可能还有更好的组合方式，这也是后续研究的方向。

另外，未来会将本文方法，应用在较大规模的程序上，在应用的过程中对方法进行进一步的完善。

参考文献（References）

[1] TRACEY N J. A search-based automated test-data generation framework for safety-critical software[D]. York: The University of York, 2001.

[2] HARMAN M, JONES B F. Search-based software engineering[J]. Information and Software Technology, 2001, 43(14): 833-839.

[3] RUAN H, ZHANG J, YAN J. Test data generation for C programs with string-handling functions[C]// 2008 2nd IFIP/IEEE International Symposium on Theoretical Aspects of Software Engineering. Piscataway: IEEE Press, 2008: 219-226.

[4] DEMILLI R A, OFFUTT A J. Constraint-based automatic test data generation[J]. IEEE Transactions on Software Engineering, 1991, 17(9): 900-910.

[5] PENG C, RAJAN A. Automated test generation for OpenCL kernels using fuzzing and constraint solving[C]//Proceedings of the 13th Annual Workshop on General Purpose Processing Using Graphics Processing Unit. New York: ACM, 2020.

[6] YANG F, FAN Y, XIAO P, et al. Test data generation method based on multiple convergence direction adaptive PSO[J]. Software Quality Journal, 2022, 31: 279-303.

[7] 杨波, 吴际, 徐珞, 等. 一种软件测试需求建模及测试用例生成方法[J]. 计算机学报, 2014, 37(3): 522-538.
YANG B, WU J, XU L, et al. An approach of modeling software testing requirements and generating test case[J]. Chinese Journal of Computers, 2014, 37(3): 522-538(in Chinese).

[8] BOYER R S, ELSPAS B, LEVITT K N. SELECT—a formal system for testing and debugging programs by symbolic execution[J]. ACM SIGPLAN Notices, 1975, 10(6): 234-245.

[9] MARCINIAK J J. Encyclopedia of software engineering[M]. Hoboken: JohnWiley & Sons, Inc., 2002.

[10] HARMAN M, MCMINN P. A theoretical & empirical analysis of evolutionary testing and hill climbing for structural test data generation[C]//Proceedings of the 2007 International Symposium on Software Testing and Analysis. New York: ACM, 2007: 73-83.

[11] METROPOLIS N, ROSENBLUTH A W, ROSENBLUTH M N, et al. Equation of state calculations by fast computing machines[J]. The Journal of Chemical Physics, 1953, 21(6): 1087-1092.

[12] PARGAS R P, HARROLD M J, PECK R R. Test-data generation using genetic algorithms[J]. Software Testing, Verification and Reliability, 1999, 9(4): 263-282.

[13] LI X L, WONG W E, GAO R Z, et al. Genetic algorithm-based test generation for software product line with the integration of fault localization techniques[J]. Empirical Software Engineering, 2018, 23(1): 1-51.

[14] SHEN X J, WANG Q, WANG P P, et al. Automatic generation of test case based on GATS algorithm[C]//2009 IEEE International Conference on Granular Computing. Piscataway: IEEE Press, 2009: 496-500.

[15] NOSRATI M, HAGHIGHI H, VAHIDI ASL M. Test data generation using genetic programming[J]. Information and Software Technology, 2021, 130: 106446.

[16] SAHOO R R, RAY M. PSO based test case generation for critical path using improved combined fitness function[J]. Journal of King Saud University - Computer and Information Sciences, 2020, 32(4): 479-490.

[17] DI NUCCI D, PANICHELLA A, ZAIDMAN A, et al. A test case prioritization genetic algorithm guided by the hypervolume indicator[J]. IEEE Transactions on Software Engineering, 2020, 46(6): 674-696.

[18] CAI G C, SU Q H, HU Z B. Automated test case generation for path coverage by using grey prediction evolution algorithm with improved scatter search strategy[J]. Engineering Applications of Artificial Intelligence, 2021, 106: 104454.

[19] GOTLIEB A, BOTELLA B, RUEHER M. Automatic test data generation using constraint solving techniques[J]. ACM SIGSOFT Software Engineering Notes, 1998, 23(2): 53-62.

[20] GUPTA N, MATHUR A P, SOFFA M L. Automated test data generation using an iterative relaxation method[J]. ACM SIGSOFT Software Engineering Notes, 1998, 23(6): 231-244.

[21] WANG C H, PASTORE F, GOKNIL A, et al. Automatic generation of acceptance test cases from use case specifications: An NLP-based approach[J]. IEEE Transactions on Software Engineering, 2022, 48(2): 585-616.

[22] BERTOLINO A, MIRANDA B, PIETRANTUONO R, et al. Adaptive test case allocation, selection and generation using coverage spectrum and operational profile[J]. IEEE Transactions on Software Engineering, 2021, 47(5): 881-898.

[23] CHEN T Y, LEUNG H, MAK I K. Adaptive random testing[C]// Annual Asian Computing Science Conference. Berlin: Springer, 2004: 320-329.

[24] CAMPOS J, GE Y, ALBUNIAN N, et al. An empirical evaluation of evolutionary algorithms for unit test suite generation[J]. Information and Software Technology, 2018, 104: 207-235.

[25] ZAMANI S, HEMMATI H. A cost-effective approach for hyper-parameter tuning in search-based test case generation[C]//2020 IEEE International Conference on Software Maintenance and Evolution. Piscataway: IEEE Press, 2020: 418-429.

[26] ZAMANI S, HEMMATI H. A pragmatic approach for hyper-parameter tuning in search-based test case generation[J]. Empirical Software Engineering, 2021, 26(6): 1-35.

[27] YANG B, HE Y Z, LIU H, et al. A lightweight fault localization approach based on XGBoost[C]//2020 IEEE 20th International Conference on Software Quality, Reliability and Security. Piscataway: IEEE Press, 2020: 168-179.

[28] YANG B, HE Y Z, FENG L X. Automated test case generation for effective spectrum-based fault localization[J]. Journal of Physics: Conference Series, 2020, 1576(1): 012042.

Using improved genetic algorithm for software fault localization aided test case generation

YANG Bo[1, 2, 3], HE Yuze[3], XU Fu[1, 2, *], CHEN Zhibo[1, 2]

(1. School of Information, Beijing Forestry University, Beijing 100083, China;
2. Engineering Research Center for Forestry Goriented Intelligent Information Processing, National Forestry and Grassland Administration, Beijing 100083, China;
3. School of Information, North China of Science and Technology, Beijing 100144, China)

Abstract: The ranking of suspected faults in the process of automatic software fault localization will be continuously created and is determined after the execution of existing test cases. Sometimes the program units corresponding to the fault are ranked lower in the ranking of suspected failures based on the existing test cases. If it is necessary to improve the suspected fault ranking of the program unit corresponding to the fault, supplementary test cases are a feasible method. This article suggests a technique for creating test cases based on a genetic algorithm that can make use of knowledge about the location of software faults. The paper analyzes and analyzes the methods used. Based on the joint experiment of the paper on 6 C programs and 2 Python programs, experimental results show that the test cases automatically generated by this method can effectively help improve the efficiency of fault location.

Keywords: fault localization; test case generation; genetic algorithm; fitness function; random test

Received: 2022-06-22; **Accepted**: 2022-08-15; **Published Online**: 2022-11-18 17:11

URL: kns.cnki.net/kcms/detail/11.2625.V.20221117.1931.005.html

Foundation items: Fundamental Research Funds for the Central Universities (BLX202003); National Natural Science Foundation of China (61502011); Focus Tracking Project of Beijing Forestry University (BLRD202124)

* **Corresponding author**. E-mail: xufu@bjfu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0669

基于 TESLA 协议的 BDSBAS 电文认证技术

陈潇[1,2]，田翔[2]，罗瑞丹[1]，刘婷[1,*]，宋佳慧[1,2]，吴海涛[1]

(1. 中国科学院空天信息创新研究院，北京 100094；2. 中国科学院大学 电子电气与通信工程学院，北京 100049)

摘 要：星基增强系统（SBAS）能为航空、航海等生命安全领域提供完好性增强服务，针对SBAS服务的欺骗风险，面向北斗星基增强系统（BDSBAS）发展，提出基于中国商用密码标准算法的时间效应流丢失容错认证机制（TESLA）电文认证方案。阐述SBAS电文认证架构与TESLA认证原理，开展基于中国商用密码标准算法的BDSABS认证电文编排及空中密钥更新（OTAR）的电文播发设计，针对L5I与L5Q开展认证时间间隔和最大认证延迟等指标的理论分析，进一步从OTAR信息权重和解调错误率2方面仿真首次认证时间指标。仿真结果能为基于TESLA协议的BDSBAS电文认证提供一定的理论支持。

关 键 词：北斗星基增强系统；电文认证；消息认证码；时间效应流丢失容错认证机制；空中密钥更新

中图分类号：V19; X949

文献标志码：A **文章编号**：1001-5965(2023)09-2289-10

以2011年伊朗捕获美国无人机、2012年德州大学团队对“白玫瑰”游艇开展的全球导航卫星系统(global navigation satellite system,GNSS)欺骗测试等事件为代表，GNSS信号格式公开导致其易受欺骗攻击的脆弱性逐步显现，导航服务安全引起广泛关注。星基增强系统(satellite-based augmentation system, SBAS)核心是增强GNSS完好性能力，满足航空、高铁等高完好性应用需求[1]。面向民航等生命安全领域，SBAS遭受欺骗攻击将导致用户完好性告警错误，进而引发飞行风险[2-3]。

SBAS提供信息增强服务，针对SBAS的欺骗攻击主要为生成式欺骗，即通过篡改SBAS电文影响SBAS服务性能。因此，欧洲和美国提出了SBAS导航电文认证(navigation message authentication, NMA)方法，保障用户能够认证信号是否来自真实的地球静止轨道(geosynchronous Earth orbit,GEO)卫星，从而在系统端赋予信号抗欺骗能力[4-5]。SBAS电文认证概念最早于2003年由美国学者斯科特提出，即基于广域增强系统(wide area augmentation system, WAAS)提供广播电文认证服务[6]。随后，欧洲和美国研究团队提出基于时间效应流丢失容错认证机制(timed efficient streaming loss-tolerant authentication protocol, TESLA)和椭圆曲线数字签名算法(elliptic curve digital signature algorithm, ECDSA)的SBAS认证方案，并开展2种认证方案评估及其对SBAS服务性能的影响[4,7-9]。2019年起，美国和欧洲联合工作组弹性技术小组共同推进SBAS电文认证标准化，开展基于欧盟和美国密码标准的ECDSA-I、ECDSA-Q、TESLA-I及TESLA-Q这4种认证方案的仿真[10]。2021年5月，美国、欧洲和日本正式启动SBAS电文认证标准化，计划在2022年底完成草案设计。2021年11月，美国联邦航空管理局计划开展WAAS电文认证标准的制定。随着中国北斗星基增强系统(Beidou satellite-based augmentation system, BDSBAS)的建设与发展，国内相关单位前期已经开展了基于中国商用密码标准SM2算法的ECDSA电文认证方法设计与仿真[11]，但是针对TESLA认证方法研究较少。

收稿日期：2021-11-05；**录用日期**：2022-03-27；**网络出版时间**：2022-04-08 09:36
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220407.1606.001.html
*通信作者. E-mail：liuting101015@aircas.ac.cn

引用格式：陈潇，田翔，罗瑞丹，等. 基于TESLA协议的BDSBAS电文认证技术[J]. 北京航空航天大学学报，2023，49(9)：2289-2298.
CHEN X，TIAN X，LUO R D，et al. Design of message authentication based on TESLA protocol for BDSBAS[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2289-2298 (in Chinese).

本文面向BDSBAS设计了基于中国商用密码标准SM3杂凑算法的TESLA认证方案，并开展仿真试验。首先，阐述SBAS电文认证架构和TESLA电文认证原理，然后，提出基于TESLA协议的电文认证方案，最后，针对TESLA协议对核心指标进行仿真与分析。

1 SBAS电文认证架构

1.1 SBAS系统架构

SBAS系统由空间段、地面段和用户段3部分组成。空间段为GEO卫星，主要产生SBAS增强信号。地面段包括监测站、数据处理中心和上注站。用户端为GNSS/SBAS接收机。相较于现有SBAS系统架构，在地面段新增认证服务中心以支持SBAS电文认证。加入SBAS电文认证后的系统架构如图1所示。

认证服务中心产生消息认证码（message authentication code, MAC）或数字签名等认证电文，由数据处理中心将认证电文与差分完好性数据产品打包生成SBAS电文，上注至GEO卫星进行广播。用户端接收GEO卫星播发的SBAS信号并解调获得认证电文，根据认证电文验证SBAS信号是否从真实GEO播发及SBAS电文是否被伪造。由于SBAS欺骗者难以伪造数字签名或MAC，导航电文认证技术将提升导航信号欺骗防护能力。

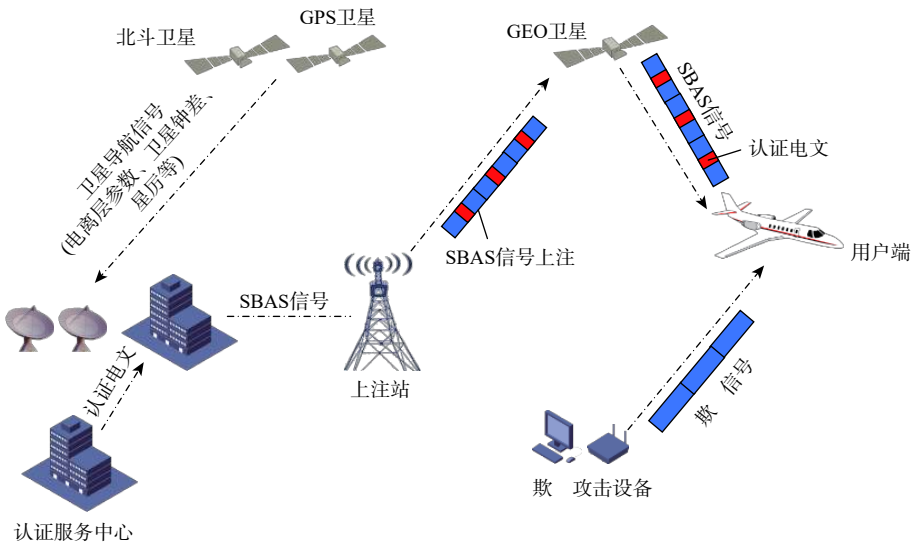


图1 加入电文认证后的SBAS系统构架

Fig. 1 SBAS architecture after adding authentication

1.2 TESLA电文认证原理

TESLA协议是一种广播认证协议，能够适用于带宽有限的卫星导航广播信号[8,12-14]。TESLA协议由Perrig设计[15]，是基于MAC的广播认证协议，该协议采用对称加密方法，其关键在于利用密钥延迟发布来保障播发密钥的安全性。基于TESLA的电文认证原理图如图2所示。

SBAS地面段利用SBAS电文和对称密钥产生

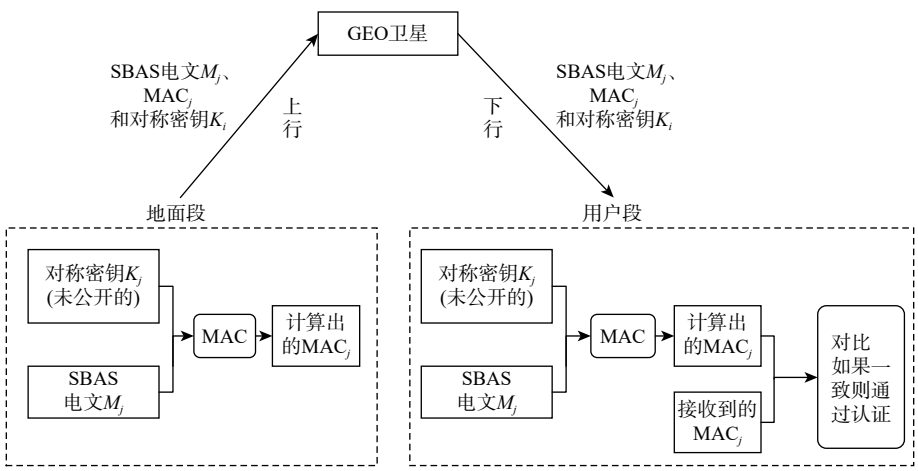


图2 基于TESLA的电文认证原理图

Fig. 2 Message authentication principle of TESLA

MAC, 并通过地面上行至 GEO 卫星。其中, 在广播 MAC 延迟 δ 时间后发送用于生成 MAC 的对称密钥。

SBAS 用户段接收 GEO 卫星广播的 SBAS 电文及 MAC 并进行存储, 待接收到对应延迟对称密钥后产生本地 MAC, 并与 SBAS 电文广播的 MAC 对比, 若两者一致表示认证通过。

2　基于 TESLA 的 BDSBAS 认证电文设计

2.1　密码算法

目前中国采用的密码标准为商用密码标准, TESLA 涉及 SM3 杂凑算法。具体包括 3 个步骤[16]:

步骤 1　分组处理。SM3 杂凑算法将长度不超过 2^{64} bit 的消息按照 512 bit 分组处理, 最终压缩成 256 bit 的摘要值。

步骤 2　消息扩展。每一个长度为 512 bit 的消息块都被划分成 16 个字 $M^{(i)}$, 其中 $0 \leqslant i \leqslant 5$。根据算法定义, $M^{(i)}$ 将扩展成 132 个字: $W_t(0 \leqslant t \leqslant 67)$, $W_t^{'}$ $(0 \leqslant t \leqslant 63)$。

步骤 3　压缩模块。扩展模块中生成的字被送入压缩模块, 也是整个算法的核心模块。模块利用 8 个 32 bit 寄存器变量 A、B、C、D、E、F、G 和 H, 其初始值为 $H_0, H_1, \cdots, H_7$。SS_1、SS_2、TT_1和TT_2为中间变量。每个压缩模块经过 64 轮的迭代压缩, 过程如下:

$$\mathrm{SS}_1 \leftarrow ((A <<< 12) + E + (T_j <<< (j \bmod 32))) <<< 7 \tag{1}$$

$$\mathrm{SS}_2 \leftarrow \mathrm{SS}_1 \oplus (A <<< 12) \tag{2}$$

$$\mathrm{TT}_1 \leftarrow \mathrm{FF}_j(A,B,C) + D + \mathrm{SS}_2 + W_j^{'} \tag{3}$$

$$\mathrm{TT}_2 \leftarrow \mathrm{GG}_j(E,F,G) + H + \mathrm{SS}_1 + W_j \tag{4}$$

$$D \leftarrow C \tag{5}$$

$$C \leftarrow B <<< 9 \tag{6}$$

$$B \leftarrow A \tag{7}$$

$$A \leftarrow \mathrm{TT}_1 \tag{8}$$

$$H \leftarrow G \tag{9}$$

$$G \leftarrow F <<< 19 \tag{10}$$

$$F \leftarrow E \tag{11}$$

$$E \leftarrow P_0(\mathrm{TT}_2) \tag{12}$$

式中: T_j 为常量, 随 j 的变化取不同值。

消息经过 64 轮迭代后, 计算哈希值的中间量 $H_j^{(i)}$ 的计算式为

$$H_0^{(i)} = A \oplus H_0^{(i-1)} \tag{13}$$

$$H_1^{(i)} = B \oplus H_1^{(i-1)} \tag{14}$$

$$\vdots$$

$$H_7^{(i)} = H \oplus H_7^{(i-1)} \tag{15}$$

式中: j 为迭代压缩轮数; $\leftarrow$为左向赋值运算符; $<<< k$为 32 位循环左移 k 位运算; $\oplus$为 32 位异或运算; +为 mod 2^{32} 位算术加运算; mod 为模运算; FF_j 和 GG_j 为布尔函数; P_0 为置换函数。

SM3 算法不断重复上述 512 bit 块的迭代压缩过程, 直到最后一个消息块 $M^{(N)}$ 处理完成输出 $H^{(N)}$, 则最终整个明文哈希值为

$$H^{(N)} = H_0^{(N)} \| H_1^{(N)} \| H_2^{(N)} \| \cdots \| H_7^{(N)} \| \tag{16}$$

式中: N 为消息分组块数量; $\|$为逻辑或运算符号。

2.2　认证电文格式设计

随着 SBAS 双频多星座(dual frequency multi constellation, DFMC)标准化推进, 未来 L5 频点将增加 SBAS 认证电文[12,17-18]。L5 信号包括 I 路和 Q 路, 其中 I 路为数据通道, 电文速率为 250 bit/s。Q 路为导频通道, 未调制电文。L5 认证电文可选播发信道包括 L5I 和 L5Q。

BDSBAS 采用 TESLA 需要产生密钥链和 MAC, 其中密钥链采用 SM3 杂凑算法, MAC 运算采用哈希运算消息认证码(hash-based message authentication code, HMAC)算法[19]。HMAC 算法是一种使用单向散列函数来构造消息认证码的方法, 其中单向散列函数采用 SM3 算法。

密钥链产生基于单向的 SM3 杂凑算法, 产生的哈希值作为密钥, 哈希值长度为 256 bit。HMAC 算法的输入为密钥和 SBAS 电文, 输出为 MAC, 其长度为 256 bit。

由于电文帧长度仅为 250 bit, 难以同时将密钥和 MAC 放入同一帧电文, 因此, 需要对 TESLA 密钥和 MAC 进行截断。从安全性角度, TESLA 密钥长度不小于 115 bit, MAC 长度不小于 30 bit[12]。其中, MAC 代表暴力破解难度, 30 bit 被欺骗攻击破解的概率为 $1/2^{30}$。TESLA 电文结构如图 3 所示, 电文结构还包括 24 bit 循环冗余校验(cyclic redundancy check, CRC)。

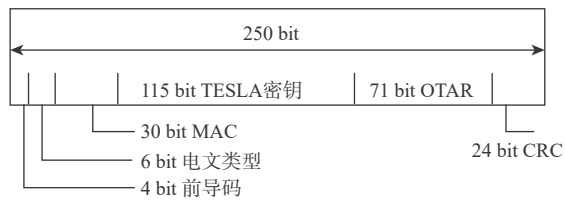


图 3　TESLA 电文结构

Fig. 3　TESLA message structure

1) L5I

L5I完好性电文的播发周期为6 s,从完好性电文的认证需求出发,需要每6 s认证完好性电文。因此,在现有L5I电文基础之上,需要每6 s增加1帧认证电文。针对L5I电文每6 s的电文帧,5帧用于SBAS差分及完好性电文传输,1帧用于SBAS认证电文传输。

图4为L5I TESLA电文结构,M_i为现有SBAS差分及完好性电文,Key为延迟密钥,OTAR为空中密钥更新(over the air rekeying, OTA)电文。由于TESLA采用密钥延迟播发,与MAC对应Key延迟6 s播发。

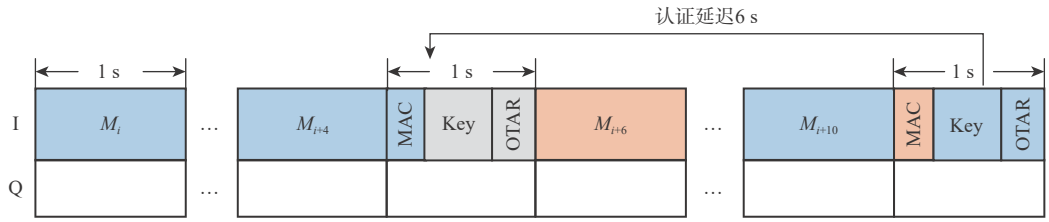


图4 L5I TESLA电文结构

Fig. 4 L5I TESLA message structure

2) L5Q

相比较L5I,L5Q播发认证电文能够提供更大带宽。图5为L5Q TESLA电文结构,在i时刻L5I电文为M_i,L5Q电文为M_i对应的MAC_i、上一秒密钥key_{i-1}及OTAR参数。在$i+1$时刻播发i时刻的key_i、当前时刻MAC_{i+1}及OTAR参数。针对M_i认证需要收齐MAC_i和key_i,故认证延迟为1 s。

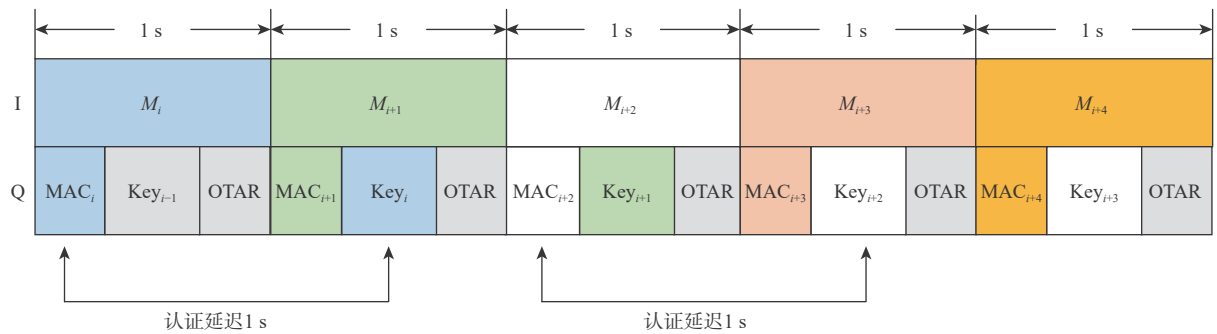


图5 L5Q TESLA电文结构

Fig. 5 L5Q TESLA message structure

2.3 OTAR设计与播发策略

OTAR(Over the air rekeying)利用无线信号发布和更新密钥。TESLA认证方法采用3层的密钥架构,其OTAR主要包括当前密钥链的盐值、密钥链的根密钥或中间密钥、盐值和根密钥的数字签名、当前系统公钥、当前系统公钥的证书管理机构(certifying authority, CA)签名、密钥链或系统公钥的到期声明及下一周期的密钥相关信息。

盐值是用于提升散列函数安全性而产生的随机数;根密钥和中间密钥用于验证当前密钥真实性;根密钥数字签名采用系统私钥签名,用户接收系统公钥验证根密钥的正确性;当前系统公钥的CA数字签名主要验证系统公钥的安全性;公钥到期声明主要用于提醒用户在当前公钥即将到期时停止使用,并即将采用新的公钥。所有OTAR播发电文类型(OTAR message type, OMT)如表1所示。

表1 TESLA OTAR播发电文类型

Table 1 TESLA OTAR message type

OTAR播发电文类型	定义	长度/bit
OMT1	当前盐值、密钥链的根密钥或中间密钥	30+115
OMT2	当前盐值和根密钥的数字签名	512
OMT3	当前系统公钥	512
OMT4	当前系统公钥的CA数字签名	512
OMT5	当前盐值/密钥链/系统公钥/认证机构公钥到期声明	48
OMT6	到期声明的CA数字签名	512
OMT7	下一个盐值、密钥链的根密钥或中间密钥	30+115
OMT8	下一个盐值和根密钥或者中间密钥的数字签名	512
OMT9	下一个系统公钥	512
OMT10	下一个系统公钥的CA数字签名	512
OMT11	下一个认证机构公钥的3项签名证书	512*3

OTAR电文结构如图6所示。每帧认证电文中OTAR电文长度为71 bit,其中包括4 bit的OMT

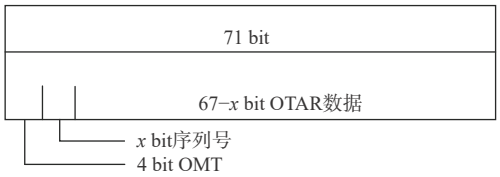


图 6　OTAR 电文结构

Fig. 6　OTAR message structure

类型、x bit 的序列号及 67−x bit 的 OTAR 数据。OMT 主要指 OMT1～OMT11, 详见表 1。序列号主要指当前 OMT 播发的分包顺序, 比特数由 OMT 信息大小决定。OTAR 数据主要传输具体 OMT 信息。

OTAR 播发策略采用标准分组公平队列算法[11], 其思想是根据每类 OMT 电文信息优先级设置不同权重, 并按照权重参数为信息分配相称的带宽, 实现重要电文的快速接收。

2.4　核心指标

SBAS 电文认证核心指标包括认证时间间隔(time between authentication, TBA)、最大认证延迟(max authentication latency, MAL)和首次认证时间(time to first authentication, TTFA)[4,18]。

1) TBA。TBA 是指完成 2 次 SBAS 电文认证的时间间隔。在 TESLA-I 情况下, 每 5 帧电文加入 1 帧认证电文, TBA 代表从接收第 1 帧电文到接收完成第 6 帧认证电文的时间。在 TESLA-Q 情况下, Q 路认证电文与 I 路电文相对应, TBA 代表接收完整认证电文的时间。

2) MAL。MAL 是指接收机从接收到一帧待认证数据到输出认证结果所需最大时间延迟, 由于 TESLA 采用密钥延迟播发机制, MAL 至少需要接收 2 帧认证电文。

3) TTFA。TTFA 是指从接收机开机开始, 第 1 次完成电文认证所需时间。TTFA 用于接收 Key、MAC 及密钥相关参数。

TESLA-I 和 TESLA-Q 的 TBA 与 MAL 如图 7 和图 8 所示, 图中 k_i(i=0,1,2,3,4)代表第 i 个时刻, 每个时刻间隔为 1 s。

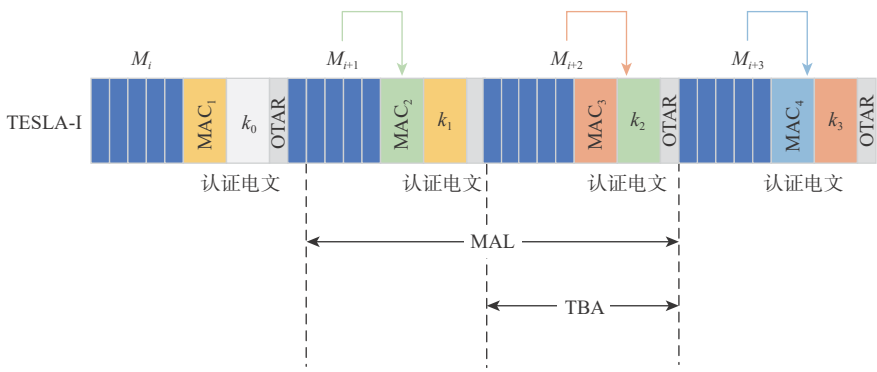


图 7　TESLA-I 最大认证延迟和认证时间间隔

Fig. 7　MAL and TBA of TESLA-I

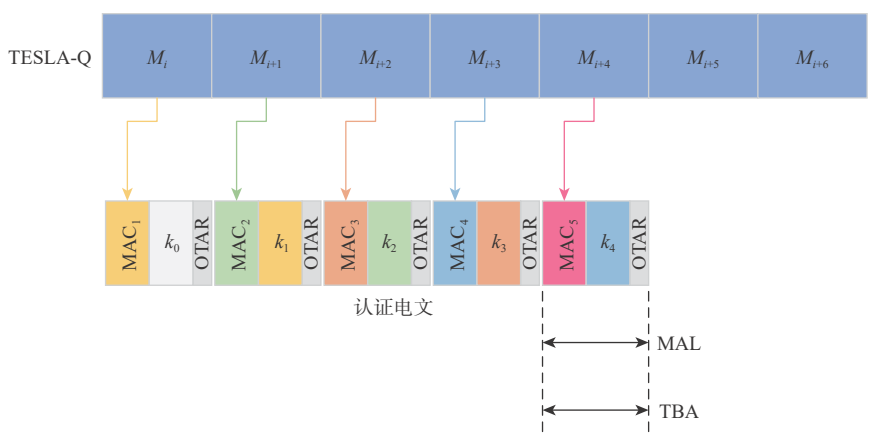


图 8　TESLA-Q 最大认证延迟和认证时间间隔

Fig. 8　MAL and TBA of TESLA-Q

3　仿真分析

3.1　核心指标分析

TBA 和 MAL 主要与认证电文设计相关, 属于设计指标。TTFA 与 OTAR 播发策略相关, 需要深入讨论。

1) TBA。针对 TESLA-I, 由于完好性电文的播发周期为 6 s, 在每 6 s 内加入一帧认证电文, 故

TESLA-I的TBA为6 s。针对TESLA-Q，由于L5Q电文余量大，I路增强电文和对应的Q路认证电文可以并行播发，故TESLA-Q的TBA为1 s。

2）MAL。由于TESLA采用密钥延迟播发的策略，实现认证需要匹配SBAS待认证电文、MAC和Key。针对TESLA-I，MAL需要接收第1个6 s周期的SBAS待认证电文和MAC_i，第2个6 s周期的Key_i，故TESLA-I的MAL为11 s。TESLA-Q的MAL为1 s。

3）TTFA。用于接收认证所需信息，包括MAC、Key和OTAR参数。其中密钥链相关信息更新周期至少以月为单位，系统公钥相关信息更新周期至少以半年为单位，认证机构公钥相关信息更新周期至少以年为单位。分成如下3种情况：

①当前密钥链相关参数已知，该情况无需接收OTAR电文。该情况下，TTFA不大于MAL，TESLA-I为6～12 s，TESLA-Q为1～2 s。

②仅需接收密钥链及系统公钥等重要电文，用于建立对当前密钥链的初始信任。

③所有OTAR参数都需要接收，该情况较少。

TTFA需要考虑OTAR参数接收，因此，对重要OTAR电文和全部OTAR电文参数接收时间进行仿真。

3.2 仿真试验

针对TTFA仿真试验采用基于MATLAB软件的蒙特卡罗工具包，该工具包通过配置OTAR参数，实现OTAR电文仿真与用户接收的全流程[20]。

仿真试验包括重要OTAR电文和全部OTAR电文的接收时间评估。重要OTAR电文指OMT1～OMT4电文（见表1）。通常情况下接收重要OTAR电文即可完成认证。

3.3 仿真分析

仿真试验包括权重分配和解调错误率2个内容。权重分配涉及分配OTAR参数不同的权重，权重较高的OTAR参数，将缩短对应参数的接收时间。解调错误率主要是接收机由于无意干扰、信噪比下降等因素导致电文解调出现帧错误。

3.3.1 权重分配

权重分配用于验证在不同权重时OTAR电文的接收时间。OMT1和OMT2涉及根密钥信息，具有最高优先级，设置最大权重W_1。OMT3和OMT4为系统公钥及签名，其更新周期小于OMT1和OMT2，设置权重为OMT1和OMT2的一半，其他OTAR电文设置较小权重W_r。

仿真试验基于不同的权重比W_1/W_r条件，分析权重比W_1/W_r变化对重要OTAR电文和全部OTAR电文的接收时间影响。在该项仿真试验中，不考虑解调错误率情况（即解调错误率为0），通过设置5个典型的权重比$W_1/W_r=\{10^0,10^1,10^2,10^3,10^4\}$，统计$10^4$次仿真数据，并绘制了TESLA-I和TESLA-Q权重比影响分析图，图9和图10分别为权重比W_1/W_r变化对TESLA-I和TESLA-Q的影响。接收全部OTAR电文的平均时间、最大时间、最小时间为T_{ave}、T_{max}、T_{min}；接收重要OTAR电文的平均时间、最大时间、最小时间为t_{ave}、t_{max}、t_{min}。

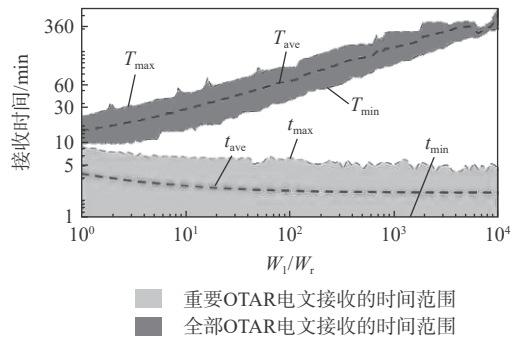


图9 TESLA-I权重比影响分析

Fig. 9 Impact analysis of weight ratio of TESLA-I

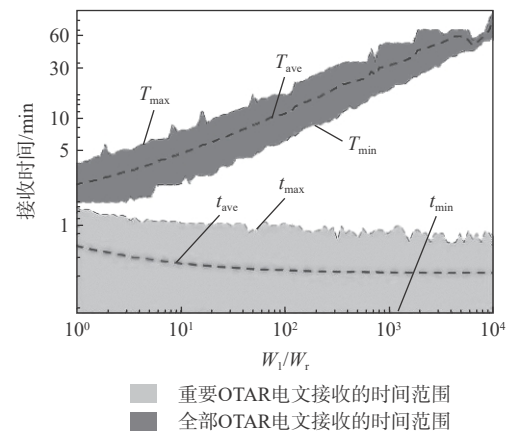


图10 TESLA-Q权重比影响分析

Fig. 10 Impact analysis of weight ratio of TESLA-Q

表2为图9中5个权重比W_1/W_r下的接收时间结果。对于TESLA-I，当权重比W_1/W_r从10^0变化到10^4过程中，接收重要OTAR电文的最大时间t_{max}从552 s降到252 s，平均时间t_{ave}从235.5 s降到131.5 s；同时可以看到接收全部OTAR电文平均时间T_{ave}急剧增长，从873.4 s增长到32 457.2 s。

表3为图10中5个权重比W_1/W_r下的接收时间结果。对于TESLA-Q，趋势与TESLA-I一致。相比较TESLA-I每6 s播发认证电文帧，TESLA-Q的播发周期为1 s，因此，仿真结果在时间方面呈现6倍关系。

根据上述仿真结果，权重比W_1/W_r代表相同时间重要OTAR电文传输的比重，随着权重比的增

表 2　TESLA-I 权重比对比结果

Table 2　Comparison results of weight ratio for TESLA-I

W_l/W_r	t_{ave}/s	t_{max}/s	t_{min}/s	T_{ave}/s
10^0	235.5	552	54	873.4
10^1	161.2	384	54	1 688.2
10^2	138.3	378	54	3 973.5
10^3	132.7	330	54	11 505.7
10^4	131.5	252	54	32 457.2

表 3　TESLA-Q 权重比对比结果

Table 3　Comparison results of weight ratio for TESLA-Q

W_l/W_r	t_{ave}/s	t_{max}/s	t_{min}/s	T_{ave}/s
10^0	39.2	92	9	145.6
10^1	26.9	64	9	281.3
10^2	23.1	63	9	662.2
10^3	22.1	55	9	1 917.6
10^4	21.9	42	9	5 409.5

加, 重要 OTAR 电文的接收平均时间越短。相对而言, 由于其他 OTAR 电文占比减小, 全部 OTAR 电文的接收时间急剧增加。

从实际仿真结果看, 权重比 W_l/W_r 大于 10^2 时, 重要 OTAR 电文接收所需平均时间下降不明显(TESLA-I 从 138.3 s 下降到 131.5 s, TESLA-Q 从 23.1 s 下降到 21.9 s), 但是全部 OTAR 电文接收所需平均时间增长很快(TESLA-I 从 3 973.5 s 上升到 32 457.2 s, TESLA-Q 从 662.2 s 增长至 5 409.5 s)。

3.3.2　解调错误率分析

解调错误率试验仿真验证在不同帧错误率(packet error ratio, PER)下的 OTAR 电文接收时间。权重分析仿真试验未考虑解调错误率, 针对实际情况设置解调错误率, 设 PER 从 10^{-5} 至 10^{-1}, 设权重比 W_l/W_r 为 10^2, 统计 10^4 次仿真数据的 OTAR 电文接收时间。

图 11 为 TESLA-I PER 影响分析, 图 12 为 TESLA-Q PER 影响分析。横轴为 PER, 纵轴为 OTAR 电文所需接收时间。

表 4 为图 11 中 TESLA-I 关键点位仿真情况。PER 从 10^{-5} 到 10^{-3} 的增长过程中, 接收重要 OTAR 电文所需的平均时间几乎不变, 139.2 s 上升至 139.5 s, 变化不到 1 s; 接收全部 OTAR 电文变化也不大, 从 3 986.3 s 上升至 3 979.3 s。当 PER>10^{-3} 后, 接收平均时间开始随 PER 的增大而剧烈上涨。表 5 为图 12 中 TESLA-Q 关键点位仿真情况。与 TESLA-I 结果类似, TESLA-I 时间大体是 TESLA-Q 的 6 倍。

PER 从 10^{-5} 到 10^{-3} 的增长过程中, PER 相对较低, 误码大概率落在 OMT1～OMT4 等重要 OTAR 电文, 由于初始设置权重比 W_l/W_r=10^2, 重要 OTAR 电文播发频度较快, 因此, 对整个接收时间影响不大。但当 PER>10^{-3} 后, 误码落在其他 OTAR 电文的概率变大, 会导致全部 OTAR 电文接收时间急剧

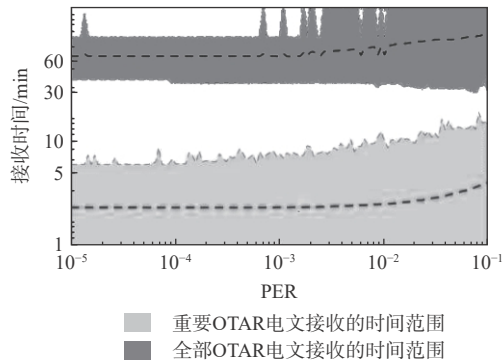


图 11　TESLA-I 帧错误率影响分析

Fig. 11　PER analysis of TESLA-I

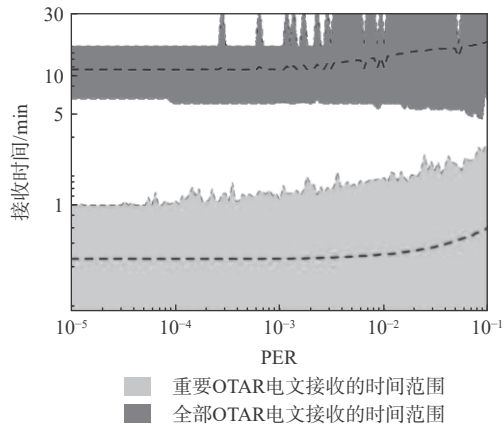


图 12　TESLA-Q 帧错误率影响分析

Fig. 12　PER analysis of TESLA-Q

表 4　TESLA-I 帧错误率对比结果

Table 4　Comparison results of TESLA-I PER

PER	t_{ave}/s	t_{max}/s	t_{min}/s	T_{ave}/s
10^{-1}	239.7	1 056	54	6 553.5
10^{-2}	150.7	594	54	5 004.1
10^{-3}	139.5	498	54	3 986.3
10^{-4}	139.0	384	54	3 985.6
10^{-5}	139.2	360	54	3 979.3

表 5　TESLA-Q 帧错误率对比结果

Table 5　Comparison results of TESLA-Q PER

PER	t_{ave}/s	t_{max}/s	t_{min}/s	T_{ave}/s
10^{-1}	40.6	164	9	1 087.3
10^{-2}	25.1	98	9	828.1
10^{-3}	23.3	74	9	663.3
10^{-4}	23.2	65	9	663.4
10^{-5}	23.2	60	9	663.8

增加。

3.4 仿真结果

表6为基于ECDSA和TESLA方案的TBA和MAL指标，其中TBA和MAL主要为设计指标。可以看出本文与斯坦福TESLA-I[14]方案设计指标一致。而TESLA-Q方案设计指标优于ECDSA-Q[11]方案，主要是因为MAC和延迟密钥的长度小于数字签名长度。

表6 基于ECDSA和TESLA方案的TBA和MAL指标

Table 6 TBA and MAL indicators based ECDSA and TESLA soulution

方案	认证方式	数字签名长度/bit	密钥	长度/bit	TBA/s	MAL/s
TESLA-I	MAC	30	延迟密钥	115	6	11
TESLA-Q	MAC	30	延迟密钥	115	1	1
ECDSA-Q[11]	数字签名	512	签名私钥 公钥	256 512	3	4
斯坦福TESLA-I[14]	MAC	30	延迟密钥	115	6	11
斯坦福ECDSA-Q[14]	数字签名	448	签名私钥 公钥	112 224	2	2

表7为不接收OTAR电文、仅接收重要OTAR电文和接收全部OTAR电文这3种情况的TTFA指标。其中后2种情况仿真条件为权重比 $W_1/W_r=10^2$ 和PER=10^{-5}。美国采用的CA签名为448 bits，小于本文采用的512 bits，故在接收重要OTAR电文方面，斯坦福TESLA-I方案接收平均时间和最小时间略好于本文方案。同时，本文针对OTAR参数进行了优化设计(详见表1)，在接收全部OTAR电文方面，接收平均时间本文方案TESLA-I略好于斯坦福方案。

表7 首次认证时间性能指标

Table 7 TTFA performance indicators s

方案	不接收OTAR电文时间	t_{max}	t_{min}	t_{ave}	T_{ave}	T_{min}	T_{ave}
TESLA-I	6～12	360	54	139.2	6 000	2 406	3 979.3
TESLA-Q	1～2	60	9	23.2	1 000	401	663.9
TESLA-I	6～12	264	66	134.2	6 000	3 858	4 989.4

4 结 论

1）本文提出一种基于中国商用密码标准SM3杂凑算法的BDSBAS TESLA认证方案。中国和美国采用不同密码标准产生密钥链，在密钥长度截短为115 bit的情况下，针对TBA、MAL和TTFA等核心指标进行分析与仿真，结果表明基于SM3杂凑算法标准的TESLA-I方案与斯坦福TESLA-I方案的仿真结果相当。

2）受益于L5Q更多的可用比特资源，TESLA-Q相对比TESLA-I，电文认证的性能更加优异。L5Q目前未调制电文，利用L5Q调制认证电文带来众多技术挑战，如国际电联频率协调(609决议、与GNSS-L5兼容性)、I路/Q路功率分配及工程实施(空间段、地面段、用户段)等，因此，还需要进一步分析与研究。

3）从OTAR的仿真试验结果可以得出，权重比最优值是10^2:1，此时重要OTAR电文的接收平均时间曲线开始变的平缓，继续增加权重比接收重要OTAR电文的平均时间下降不明显，但是接收全部OTAR电文的平均时间将激增。同时，也可以得出PER可选范围建议小于10^{-3}，当大于10^{-3}时接收重要OTAR电文的平均时间和接收全部OTAR电文的平均时间都将大幅增加。

4）在SBAS密码安全方面，美国和欧洲均优先采用本国的密码标准完成了SBAS电文认证方案设计与仿真。本文方案将为中国商用密码标准未来纳入SBAS电文认证标准提供技术支撑。

参考文献（References）

[1] 谭述森. 北斗卫星导航系统的发展与思考[J]. 宇航学报, 2008, 29(2): 391-396.
TAN S S. Development and thought of compass navigation satellite system[J]. Journal of Astronautics, 2008, 29(2): 391-396(in Chinese).

[2] PSIAKI M L, HUMPHREYS T E. GNSS spoofing and detection[J]. Proceedings of the IEEE, 2016, 104(6): 1258-1270.

[3] CHEN Y, GAO W G, CHEN X, et al. Advances of SBAS authentication technologies[J]. Satellite Navigation, 2021, 2(1): 1-7.

[4] DALLA CHIARA A, DA BROI G, POZZOBON O, et al. Authentication concepts for satellite-based augmentation systems[C]//Proceedings of the 29th International Technical Meeting of the Satellite Division of the Institute of Navigation. Manassas: Institute of Navigation, 2016: 3208-3221.

[5] ENGE P, WALTER T. Digital message authentication for SBAS and APNT[C]//Proceedings of the 27th International Technical Meeting of the Satellite Division of The Institute of Navigation. Manassas: Institute of Navigation , 2014: 1328-1336.

[6] SCOTT L. Anti-spoofing & authenticated signal architectures for civil navigation systems[C]//Proceedings of the 16th International Technical Meeting of the Satellite Division of The Institute of Navigation. Manassas: Institute of Navigation, 2003: 1543-1552.

[7] FERNáNDEZ-HERNáNDEZ I. GNSS authentication: Design parameters and service concepts[C]//Proceedings of the European Navigation Conference. Alphen aan den Rijn: EUGIN, 2014.

[8] NEISH A, WALTER T, POWELL J D. SBAS data authentication: a concept of operations[C]//Proceedings of the 32nd International Technical Meeting of the Satellite Division of The Institute of Navigation. Manassas: Institute of Navigation, 2019: 1812-1823.

[9] DALLA CHIARA A, DA BROI G, POZZOBON O, et al. SBAS authentication proposals and performance assessment[C]//Proceedings of the 30th International Technical Meeting of the Satellite Division of the Institute of Navigation. Manassas: Institute of Navigation, 2017: 2106-2116.

[10] FERNÁNDEZ-HERNÁNDEZ I, WALTER T, NEISH A M, et al. SBAS message authentication: A review of protocols, figures of merit and standardization plans[C]//Proceedings of the 2021 International Technical Meeting of the Institute of Navigation. Manassas: Institute of Navigation, 2021: 111-124.

[11] 穆盛林, 陈颖, 刘婷, 等. 面向BDSBAS电文认证的OTAR播发策略设计[J]. 北京航空航天大学学报, 2021, 47(7): 1453-1461.
MU S L, CHEN Y, LIU T, et al. Design of message authentication and OTAR broadcast strategy for BDSBAS[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(7): 1453-1461(in Chinese).

[12] NEISH A, WALTER T, ENGE P. Parameter selection for the TESLA keychain[C]//Proceedings of the 31st International Technical Meeting of the Satellite Division of the Institute of Navigation. Manassas: Institute of Navigation, 2018: 2155-2171.

[13] NEISH A, WALTER T, ENGE P. Quantum-resistant authentication algorithms for satellite-based augmentation systems[J]. Navigation, 2019, 66(1): 199-209.

[14] NEISH A, WALTER T, POWELL J D. Design and analysis of a public key infrastructure for SBAS data authentication[C]//Proceedings of the ION 2019 Pacific PNT Meeting. Manassas: Institute of Navigation, 2019: 964-988.

[15] PERRIG A, CANETTI R, TYGAR J D, et al. Efficient authentication and signing of multicast streams over lossy channels[C]//Proceeding 2000 IEEE Symposium on Security and Privacy. Piscataway: IEEE Press, 2002: 56-73.

[16] 中华人民共和国国家质量监督检查检验检疫总局, 中国国家标准化管理委员会. SM3密码杂凑算法: GB/T 32905—2016. [S]. 北京: 中国标准出版社, 2017.
General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Information security techniques—SM3 cryptographic hash algorithm: GB/T 32905—2016 [S]. Beijing: Standards Press of China, 2017(in Chinese)

[17] Satellite Based Augmentation System Interoperability Working Group. SBAS L5 DFMC interface control document [S]. Montreal: SBAS IWG, 2015

[18] FERNÁNDEZ-HERNÁNDEZ I, CHÂTRE E, DALLA CHIARA A, et al. Impact analysis of SBAS authentication[J]. Navigation, 2018, 65(4): 517-532.

[19] 结城浩著. 图解密码技术[M]. 周自恒译. 第3版. 北京: 人民邮电出版社, 2016.
HIROSHI Y K. Graphic cryptography technology[M]. ZHOU Z H translated. 3rd ed. Beijing: Posts & Telecom Press, 2016 (in Chinese).

[20] NEISH A. Establishing trust through authentication in satellite based augmentation systems[D]. Stanford: Stanford University, 2020.

Design of message authentication based on TESLA protocol for BDSBAS

CHEN Xiao[1, 2], TIAN Xiang[2], LUO Ruidan[1], LIU Ting[1, *], SONG Jiahui[1, 2], WU Haitao[1]

(1. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China;
2. School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: Aviation, navigation, and other life safety areas can benefit from the integrity augmentation services offered by SBAS. In response to the deception risk of SBAS services and the development of the BeiDou satellite-based augmentation system (BDSBAS), this paper proposes a message authentication method based on the timed efficient streaming loss-tolerant authentication (TESLA) protocol. Firstly, this article introduces SBAS service security risks and the concept of message authentication and then explains the principle of TESLA, key chain generation, and operation process. The Over the Air Rekeying (OTAR) message broadcast design and authentication message format are created in accordance with the domestic commercial cryptographic standard hash algorithm. Finally, a simulation test is carried out for the L5I channel and the L5Q channel. The simulation results can provide theoretical support for the BDSBAS authentication based on the TESLA scheme.

Keywords: Beidou satellite-based augmentation system; navigation message authentication; message authentication code; the timed efficient streaming loss-tolerant authentication; over the air rekeying

Received: 2021-11-05; **Accepted**: 2022-03-27; **Published Online**: 2022-04-08 09:36
URL: kns.cnki.net/kcms/detail/11.2625.V.20220407.1606.001.html
* **Corresponding author**. E-mail: liuting101015@aircas.ac.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0680

基于变权重-正态云模型的飞机轮胎滑水风险

李岳[1]，胡宇祺[1,2]，蔡靖[1,3,*]，戴轩[1]

(1. 中国民航大学 交通科学与工程学院，天津 300300；2. 中国国际航空股份有限公司地面服务部，北京 101309；
3. 民航机场智能建造与工业化工程技术研究中心，天津 300456)

摘　　要：针对飞机轮胎滑水行为的随机性和模糊性特征，提出基于正态云模型的滑水安全评价分析方法。引入变权理论动态调整权值，采用惩罚性变权函数降低常权权值对评价结果的主观影响。构建飞机轮胎滑水流固耦合仿真分析模型，选取飞机轮载、滑行速度、积水厚度、道面摩擦系数及刻槽深度作为风险影响因素，基于单因素云模型数字特征及变权向量求解综合隶属度，建立多元决策下的飞机滑水风险等级及划分标准。以某山区多雨机场轮胎滑水事件为实例进行验证，结果表明：以传统临界滑水速度指标进行条件判定仅得出允许起降的一般结论；对比常权和变权评价结果，工况1安全系数由1.09提高至1.17，工况2由2.09提高至2.94，可定量描述道面起降环境差异，滑水风险仍在可接受范围内，变权评价结果偏于保守；工况3安全系数由3.13提高至3.74，滑水风险上升至Ⅳ级，即使道面积水厚度符合上限要求，轮胎滑水发生几率仍有可能显著提高，与实际风险情况基本一致，对道面运行安全分级管理具备参照性。

关 键 词：正态云；变权重；飞机滑水；综合隶属度；流固耦合；机场工程
中图分类号：V351.11
文献标志码：A　　**文章编号**：1001-5965(2023)09-2299-07

随着机场建设规模提升及航班运输总量的增加，恶劣起降条件下发生飞机轮胎滑水事件几率随之增大，可引发冲偏出跑道等严重事故[1]，滑水已成为威胁航空运行安全的热点研究课题。飞机轮胎滑水属于一类多因素复杂系统风险评价工作，涉及飞机起降性能、降雨气候条件、道面结构设计与表面特性等多方面，具有较强的随机性和模糊性。

当前，飞机滑水安全问题研究大多采用工程仿真方法进行定性-半定量分析[2]。以美国国家航空航天局(National Aeronautics and Space Administration, NASA)临界滑水速度作为风险评价指标，通过有限元计算求解不同道面积水厚度、轮胎磨损程度、道面摩擦特性和平整度条件下临界滑水速度特征值，作为飞机起降安全评价的依据[3-9]。不足之处在于未将飞机滑水事件视为完全模糊随机过程，仅通过比较临界滑水速度与飞机接地速度峰值得出二元评价结论(滑水或不滑水)，与事故分析“发生频次低、危险性高”的统计特征不符[10]，缺乏系统的多因素飞机滑水风险分析方法。

云评价模型在处理此类随机模糊问题方面具备一定优势。云评价模型由中国李德毅院士提出，该模型综合概率论和模糊数学方法，旨在通过算法解决定性描述向定量数值表达之间的不确定性转换问题，在航空安全及稳定性评估中得到应用[11-13]。文献[14]将云模型理论引入空管安全运营保障能力评价，通过逆向云发生器得到安全运行保障能力的评价云。文献[15]针对多因素影响下的飞机重着陆复杂系统风险问题，提出因素与系统间交互作

收稿日期：2021-11-10；录用日期：2021-12-27；网络出版时间：2022-01-27 12:58
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220126.0946.001.html
基金项目：天津市自然科学基金(21JCQNJC00850)；中央高校基本科研业务费专项资金(3122019103)
*通信作者. E-mail：jcai@cauc.edu.cn

引用格式：李岳，胡宇祺，蔡靖，等. 基于变权重-正态云模型的飞机轮胎滑水风险[J]. 北京航空航天大学学报，2023，49(9)：2299-2305.
LI Y，HU Y Q，CAI J，et al. Hydroplaning risk of aircraft tire based on variable weight theory-normal cloud model[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2299-2305 (in Chinese).

用模型，结合各个风险因素以多维云理论生成 5 朵四维风险等级云模型。文献 [16] 提出基于云模型的客机风险评估方法，将主客观权重进行加权平均得到各因素的权重。文献 [17] 引入变权理论克服常权方法局限性，突出关键因素对综合评价结果的影响。上述研究工作为飞机起降滑水风险评价提供了分析思路和研究手段。

针对飞机滑水风险问题特征，本文提出基于正态云模型的滑水安全评价分析方法，引入变权理论动态调整权值，考察多种滑水影响因素相对重要程度及状态均衡度，建立轮胎滑水风险等级及划分标准，对某山区多雨机场滑水事件进行实例例证，验证本文方法的可行性和适用性。

1 变权重-正态云模型原理

1.1 正态云评价模型

云评价模型以模糊理论为基础，旨在考察随机事件判断过程中的不确定性。常用云模型有对称云、半云、组合云、一维云和正态云等类型。由中心极限定理可知，多个相互独立自变量之和的随机变量服从正态分布。据此，本文采用普适性正态云模型进行滑水安全评价。

根据云模型理论，若存在一个可用精确数值描述的论域 T 且 $T=\{x\}$；假定 A 是 T 上的一个定性概念，对任意元素 x，都存在一个具有稳定倾向的随机数值 $\mu(x)\in[0,1]$，即为 x 对 A 的隶属度。单个 $(x,\mu(x))$ 称为云滴，云滴是论域空间中的一个随机变量，整体服从泛正态分布，其在 T 上的分布称为云。对于定性概念 A，正态云模型隶属度满足：

$$\mu(x)=\exp\left[-\frac{(x-E_x)^2}{2(E_{n'})^2}\right] \tag{1}$$

式中：E_x 和 $E_{n'}$ 为云模型的期望和熵，且 $E_{n'}$ 服从 $E_{n'}\sim(E_n,H_e^2)$，E_n 为熵的期望值，H_e 为超熵（即 $E_{n'}$ 的熵），三者共同构成云评价模型基本数字特征，如图 1 所示。

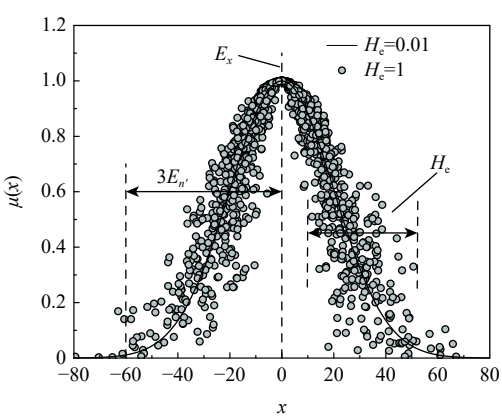


图 1 云评价模型数字特征

Fig. 1 Digital features of cloud evaluation model

采用正向云发生器生成云模型，基于云模型的数字特征得出隶属度结果，完成定性与定量信息转化[18]。单影响因素云图总体云滴数量不少于 3 000 个时，隶属度结果误差低于 5%[19]。

1.2 变权重计算与综合隶属度

飞机滑水影响因素众多，风险因素的赋权权值对评价结果准确性至关重要。采用层次分析法确定因素权值，评价结果受主观影响较大。对此引入变权理论，考察各因素间相对重要次序和状态均衡度，动态调整常权权值。

假定因素状态向量 $\boldsymbol{X}=[x_1,x_2,\cdots,x_n]$，常权 $\boldsymbol{W}=[w_1,w_2,\cdots,w_n]$，根据变权理论定义存在一组映射 $[0,1]^n\to[0,1]$，$\boldsymbol{X}\to\boldsymbol{W}(\boldsymbol{X})$，该映射满足归一性、连续性及单调性条件。$\boldsymbol{W}(\boldsymbol{X})=[w_1(\boldsymbol{X}),w_2(\boldsymbol{X}),\cdots,w_n(\boldsymbol{X})]$ 为特定状态向量 $\boldsymbol{X}$ 下的变权值。

本文选用惩罚性变权函数，定义 n 维递减向量 $\boldsymbol{S}:[0,1]^n\to[0,1]$，$\boldsymbol{X}\to\boldsymbol{S}(\boldsymbol{X})=[S_1(\boldsymbol{X}),S_2(\boldsymbol{X}),\cdots,S_n(\boldsymbol{X})]$。映射 $\boldsymbol{S}$ 为一个 n 维状态变权向量，惩罚性变权向量应满足以下条件：

1）$x_i\geqslant x_j\to S_i(\boldsymbol{X})\leqslant S_j(\boldsymbol{X})$；

2）$S_i(\boldsymbol{X})$ 对每个变元连续；

3）常权向量 $\boldsymbol{W}$ 在满足上述定义前提下有：

$$\boldsymbol{W}(\boldsymbol{X})=\frac{\boldsymbol{W}\cdot\boldsymbol{S}(\boldsymbol{X})}{\sum_{j=1}^{n}w_jS_j(\boldsymbol{X})} \tag{2}$$

式中：$\boldsymbol{W}\cdot\boldsymbol{S}(\boldsymbol{X})=[w_1S_1(\boldsymbol{X}),w_2S_2(\boldsymbol{X}),\cdots,w_nS_n(\boldsymbol{X})]$ 为 Hardarmard 乘积；状态变权向量 $\boldsymbol{S}(\boldsymbol{X})$ 可由均衡函数 $f(x)$ 计算得到[13]。通过飞机滑水正态云评价模型得出单个影响因素隶属度 $\mu_j(x)$ 后，结合变权向量 $\boldsymbol{W}(\boldsymbol{X})$，综合隶属度 U_t 的计算式为

$$U_t=\sum_{j=1}^{n}\mu_j(x)w_j(\boldsymbol{X}) \tag{3}$$

2 飞机轮胎滑水风险分析

2.1 滑水风险等级量化过程

基于变权重-正态云模型的飞机轮胎滑水风险评价流程如图 2 所示。风险等级量化的基本思路是：分析飞机滑水行为影响因素，根据风险等级划分标准，提取单因素正态云模型数字特征，生成云图并得出单因素隶属度。在此基础上，依托变权重理论计算各因素权值，获得综合隶属度结果及当前飞机滑水风险等级，实现风险定量评价目标。

2.2 滑水风险影响因素及分级标准

飞机接地滑行时滚动轮胎与湿滑道面发生流固耦合相互作用，影响因素间交互关联，随机模糊

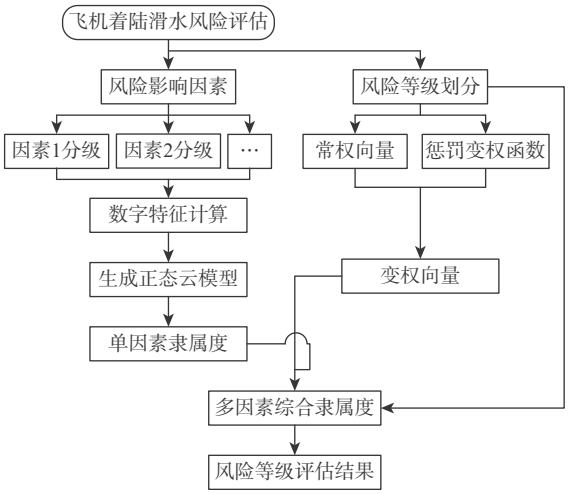


图 2　滑水风险评价流程

Fig. 2　Assessment procedure of hydroplaning risk

特征突出, 涵盖飞机滑行性能、积水水文特征及道面运行状况等多方面。

采用 ABAQUS 软件建立飞机轮胎滑水仿真分析模型, 如图 3 所示。综合以往研究经验, 选取飞机轮载、滑行速度、积水厚度, 道面摩擦系数和刻槽深度共 5 项风险影响因素[2-3, 20-21], 共同描述飞机接地滑水状态。

滚动飞机轮胎模型与下部积水(含空气层)相互独立离散, 采用耦合欧拉–拉格朗日(coupled Lagrangian Eulerian, CEL)算法模拟流固耦合作用过程; 调节轮胎前缘的水泵区流速模拟轮胎与道面相对运动状态。该分析方法在以往研究中已广泛应用, 计算精度得到充分验证, 详细建模参数及模型验证过程参见文献 [21]。

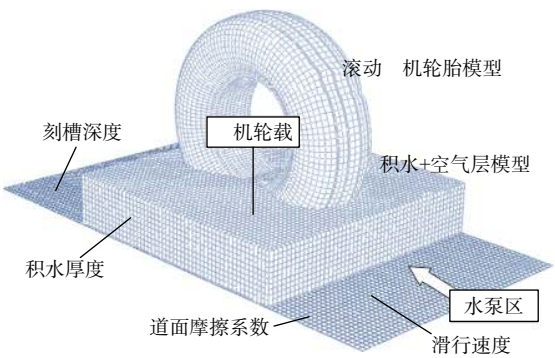


图 3　飞机轮胎滑水流固耦合仿真模型

Fig. 3　Simulation model of aircraft tire hydroplaning on fluid-solid coupring analysis

以积水厚度因素为例, 由图 3 所示仿真模型得出 0～15 mm 范围内临界滑水速度结果, 对照机载快速存取记录器(quick access record, QAR)数据实现不同工况下滑水风险量化, 定义滑水风险指标(hydroplaning risk index, HRI)为

$$H_{\mathrm{HRI}}=\frac{K(v_{\mathrm{c}})-K(v_0)}{K(v_1)-K(v_0)}\times 100\% \tag{4}$$

式中: $K(v)$ 为飞机滑行距离函数, 可由 QAR 获得; v_{c} 为临界滑水速度; v_0 为飞机接地速度; v_1 为跑道脱离速度, 以概率形式量化单因素影响下飞机滑水风险。通过控制变量法分别计算 5 项风险影响因素下 HRI 值, 参照双重不确定性行为安全评价分级方法, 将单因素对飞机滑水行为影响程度等分为: 可忽略(Ⅰ级)、可接受(Ⅱ级)、可接受但需整改(Ⅲ级)和不可接受(Ⅳ级)4 个等级[22-23]。

通过大量滑水仿真计算确定各风险等级划分标准及因素取值范围如表 1 所示。

表 1　飞机滑水风险因素及分级

Table 1　Aircraft hydroplaning evaluation index and classification

风险等级	滑行速度/(km·h^{-1})	积水厚度/mm	飞机轮载/kN	刻槽深度/mm	道面摩擦系数
Ⅰ级	190～200	0～5	123.1～138.0	5～6	0.7～0.8
Ⅱ级	200～230	5～8	107.3～123.1	4～5	0.6～0.7
Ⅲ级	230～260	8～11	91.7～107.3	3～4	0.5～0.6
Ⅳ级	260～300	11～15	75.8～91.7	0～3	0～0.5

2.3　正态云模型数字特征提取

根据不同因素取值范围, 确定各风险等级正态云模型数字特征, 数字特征的计算式为

$$\begin{cases} E_x=\dfrac{1}{2}(c_{\max}+c_{\min}) \\ E_{n'}=\dfrac{1}{3}(c_{\max}-c_{\min}) \\ H_{\mathrm{e}}=k \end{cases} \tag{5}$$

式中: $c_{\min}$ 和 $c_{\max}$ 分别对应 4 个等级标准区间的左右边界值; E_x 表征了目标概念在论域 T 的中心值; $E_{n'}$ 量化了云滴的离散程度和取值范围; 超熵 H_{e} 反映了云滴的不确定性; k 为当前超熵取值, 根据经验取 0.01[24]。计算各风险等级数字特征结果如表 2 所示。依据表 2, 在 Python 环境下, 指定种子数为 0, 采用 numpy 随机生成数字建立数据库, 按式(1)计算隶属度, 生成多组 4 朵一维单因素云图, 如图 4 所示。

如图 4(a)和图 4(b)所示, 积水厚度和滑行速度单因素云图自左向右安全评价等级逐步提高, 因素影响由“可忽略”逐步偏移为“不可接受”, 表明当积水越厚时跑道起降条件越恶劣, 高速接地滑行飞机安全性越差。典型 3 mm 至 13 mm 积水厚度范围对应安全等级Ⅱ级和Ⅲ级云图曲线, 与现行《航空承运人湿跑道和污染跑道运行管理规定》中关于允许飞机起降条件相符合。图 4(c)～图 4(e)中单因素云图自左向右安全评价等级逐步降低, 表明单轮

表2 正态云模型数字特征

Table 2 Digital features of normal cloud model

影响因素	滑行速度/(km·h^{-1})	积水厚度/mm	飞机轮载/kN	刻槽深度/mm	道面摩擦系数
I级	(195,3.33,0.01)	(2.5,1.67,0.01)	(129.8,5.7,0.01)	(5.5,0.33,0.01)	(0.75,0.03,0.01)
II级	(215,10,0.01)	(6.5,1,10.01)	(113.0,5.5,0.01)	(4.5,0.33,0.01)	(0.65,0.03,0.01)
III级	(245,10,0.01)	(9.5,1,0.01)	(97.1,5.1,0.01)	(3.5,0.33,0.01)	(0.55,0.03, 0.01)
IV级	(280,13.3,0.01)	(13,1.33,0.01)	(82.8,4.5,0.01)	(1.5,1,0.01)	(0.25,0.166, 0.01)

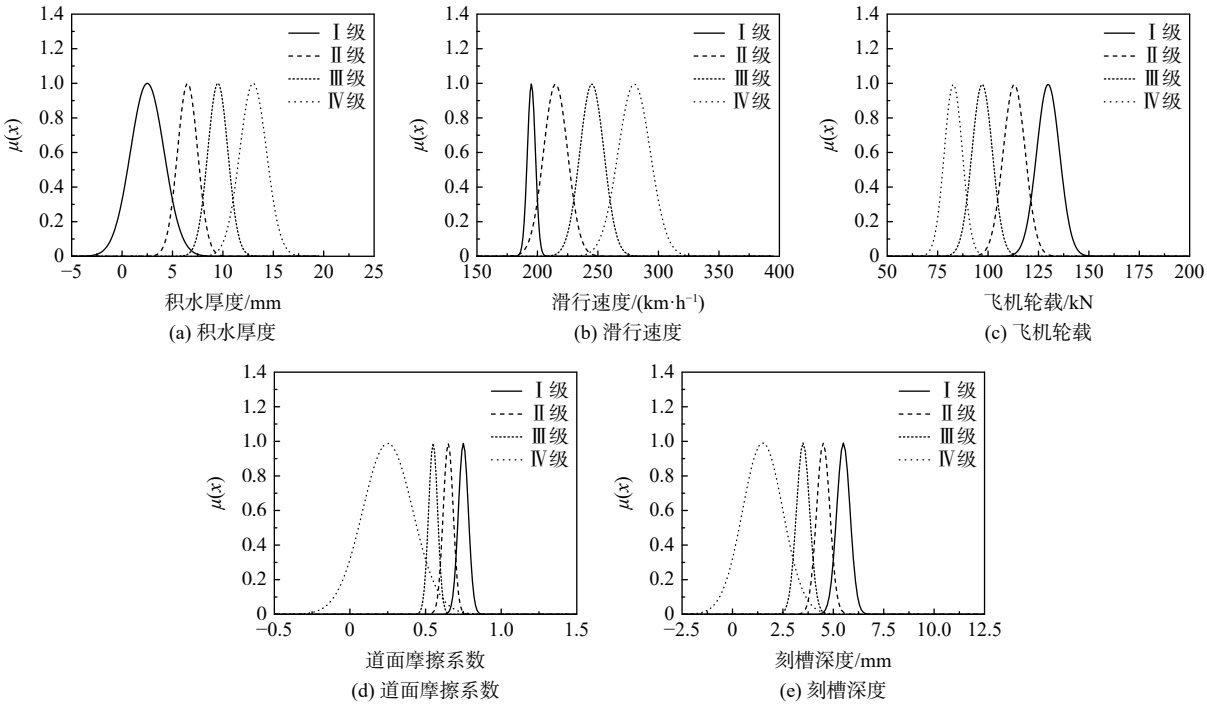


图4 单影响因素典型云图

Fig. 4 Typical cloud chart of single influencing factor

轴载越低轮胎越容易在动水压强作用下发生水滑行为,道面摩擦系数及刻槽深度降低则跑道表面抗滑性能越差,飞机滑水可能性越高。

图4表明单因素影响下飞机滑水风险等级关系,仍需通过变权理论进行多因素叠加,以获得符合实际跑道复杂运行环境的综合隶属度。

3 案例应用

某南方山区4E等级机场经过多年运行,道面运行状况存在不同程度劣化。气象资料显示该机场周边区域降雨强度可达2.0 mm/min,道面频繁处于湿滑污染起降运行状态。飞行员曾数次报告轮胎滑水事件及事故征候,飞机滑水风险问题突出,对此,采用本文建立的变权重-正态云模型进行飞机滑水安全评价。通过现场考察确定3组轮胎水滑事件对应的跑道运行参数如表3所示。

考虑不同因素量纲差异,为简便计算将表中数值归一化处理,获得的状态向量矩阵 $\boldsymbol{X}$ 为

$$\boldsymbol{X}=\begin{bmatrix} 0.293\,3 & 0.187\,5 & 0.439\,5 & 0.428\,6 & 0.376\,2 \\ 0.333\,3 & 0.312\,5 & 0.318\,5 & 0.357\,1 & 0.326\,7 \\ 0.373\,4 & 0.500\,0 & 0.242\,0 & 0.214\,3 & 0.297\,1 \end{bmatrix} \tag{6}$$

表3 轮胎水滑事件影响因素基本参数

Table 3 Basic parameters of influencing factors of aircraft tire hydroplaning symptoms

工况	滑行速度/(km·h^{-1})	积水厚度/mm	飞机轮载/kN	刻槽深度/mm	道面摩擦系数
1	220	3	138.8	6	0.76
2	250	5	100.4	5	0.66
3	280	8	76.5	3	0.60

3.1 常权及变权计算

变权向量由因素常权向量及状态变权向量联合计算求得。因此,求解因素变权向量应首先确定对应因素常权。基于流固耦合仿真结果,采用层次分析法数值评价得到多因素判断矩阵 $\boldsymbol{J}$ 为

$$\boldsymbol{J}=\begin{bmatrix} 1 & \frac{1}{5} & \frac{1}{9} & \frac{1}{7} & \frac{1}{3} \\ 5 & 1 & \frac{1}{5} & \frac{1}{3} & 3 \\ 9 & 5 & 1 & 3 & 7 \\ 7 & 3 & \frac{1}{3} & 1 & 5 \\ 3 & \frac{1}{3} & \frac{1}{7} & \frac{1}{5} & 1 \end{bmatrix} \tag{7}$$

得出常权向量 $\boldsymbol{W}$ 为

$$\boldsymbol{W}=[0.052\,0, 0.201\,2, 0.239\,9, 0.408\,0, 0.098\,9] \tag{8}$$

状态变权向量 $\boldsymbol{S}(\boldsymbol{X})$ 形式多样, 常见的有算数均值型、指数型、对数型等。指数型状态变权向量可以通过给定不同的均衡力来确定合适的参数, 更具灵活性[25]。惩罚性变权向量的实质为当影响因素状态值处于不均衡组态时, 将通过降低优势因素权值来达到安全判断的目的。

据此, 本文选用惩罚性状态变权向量, 强调劣势因素在飞机湿滑道面滑行时不利影响, 判定结果较为保守。n 维指数型状态变权向量第 i 项可表示为[26]

$$\boldsymbol{S}_i[x_1, x_2, \cdots, x_n] = f(\beta x_i - \bar{x}) = \mathrm{e}^{-\alpha(\beta x_i - \bar{x})} \tag{9}$$

式中: $\bar{x}$为状态向量平均值, $\alpha \geqslant 0$且α取值越大则惩罚效果越明显, 分析相关研究内容可知在分析因素较少的情况下α取值范围为 0～10, 本文选用$\alpha = 10$, 强调劣势因素对整体评价的影响, $1/n \leqslant \beta \leqslant 1$, $i = 1, 2, \cdots, n$, 此处取$\beta = 0.5$[25]。

结合常权向量 $\boldsymbol{W}$ 及状态变权向量 $\boldsymbol{S}(\boldsymbol{X})$ 计算得到变权向量 $\boldsymbol{W}(\boldsymbol{X})$ 为

$$\begin{cases} \boldsymbol{W}_1(\boldsymbol{X}) = [0.066\,5, 0.436\,9, 0, 147\,7, 0.265\,3, 0.083\,6] \\ \boldsymbol{W}_2(\boldsymbol{X}) = [0.198\,7, 0.206\,7, 0, 204\,4, 0.189\,0, 0.201\,2] \\ \boldsymbol{W}_3(\boldsymbol{X}) = [0.185\,3, 0.131\,7, 0, 231\,6, 0.235\,6, 0.215\,8] \end{cases} \tag{10}$$

3.2　安全等级计算

通过计算确定不同工况组合的变权向量, 得出各因素对整体滑水行为权重。根据式(1)计算不同影响因素的隶属度结合不同工况的变权向量利用式(3)得出各工况下综合隶属度。

在此基础上, 引入分级矩阵$\boldsymbol{N} = [1, 2, 3, 4]$得到不同工况下的道面安全系数。最终道面安全等级计算式为

$$L = \boldsymbol{N}\boldsymbol{U}^{\mathrm{T}} \tag{11}$$

式中: $\boldsymbol{U} = [u_1, u_2, \cdots, u_m]$

为验证采用惩罚性变权函数合理性, 表 4 列举了基于临界滑水速度判定和常权向量得出的道面滑水风险评价结果, 作为本文方法的对照。结果表明:

1) 3 组工况下道面积水厚度均在现行允许起降规定范围内, 采用传统临界滑水速度指标进行滑水条件判定, 仅能得出允许起降的一般结论, 无法体现道面运行条件差异及滑水风险变化。

2) 基于常权函数判定工况 1～工况 3 安全评价结果分别为“可忽略”、“可接受”和“可接受但需整改”, 可作为道面安全分级管控参照, 各因素对评价分级主观作用较为显著。

3) 选用惩罚性变权函数评价分级整体高于同工况下常权函数判定结果, 评价结论偏于保守; 工况 3 道面滑水安全等级由Ⅲ级上升为Ⅳ级(不可接受), 道面满足积水厚度上限要求时, 发生轮胎滑水几率仍可显著提高, 符合实际运行情况。

表 4　不同工况下飞机滑水风险安全性评价

Table 4　Safety evaluation of aircraft hydroplaning under different conditions

工况	评价方法	安全系数	安全等级	规范判定
1	常权	1.09	Ⅰ级	允许起降
	本文方法	1.17	Ⅰ级	
2	常权	2.09	Ⅱ级	允许起降
	本文方法	2.94	Ⅲ级	
3	常权	3.13	Ⅲ级	允许起降
	本文方法	3.73	Ⅳ级	

4　结　论

1) 考虑层次分析法确定常权对评价结果的主观影响, 采用惩罚性变权函数动态调整权值, 避免危险指标被优势指标中和, 由单因素云模型及变权向量求解综合隶属度, 实现多元决策下的飞机滑水风险分级。

2) 本文方法可考察不同参数组合下轮胎滑水风险等级差异并做出定量描述, 与实际风险情况基本一致, 较传统二元滑水条件判定方法, 对道面运行安全分级管理更具参照性。

参考文献（References）

[1] 余治国, 李曙林, 朱青云. 机轮动力滑水机理分析[J]. 空军工程大学学报(自然科学版), 2004, 5(5): 9-11.
YU Z G, LI S L, ZHU Q Y. Mechanism analysis of an aircraft tire dynamic hydroplaning[J]. Journal of Air Force Engineering University (Natural Science Edition), 2004, 5(5): 9-11(in Chinese).

[2] 蔡靖, 许诤. 沟槽磨损对飞机轮胎滑水影响仿真分析[J]. 中国民航大学学报, 2020, 38(2): 38-43.
CAI J, XU Z. Simulation analysis on influence of groove abrasion on aircraft hydroplaning[J]. Journal of Civil Aviation University of China, 2020, 38(2): 38-43(in Chinese).

[3] 朱兴一, 庞亚凤, 杨健, 等. 湿滑条件下基于真实纹理道面的机轮

着陆滑水行为解析[J]. 中国公路学报, 2020, 33(10): 159-170.
ZHU X Y, PANG Y F, YANG J, et al. Analysis on the hydroplaning of aircraft tire under real texture pavement conditions[J]. China Journal of Highway and Transport, 2020, 33(10): 159-170(in Chinese).

[4] PURBA J H. A fuzzy-based reliability approach to evaluate basic events of fault tree analysis for nuclear power plant probabilistic safety assessment[J]. Annals of Nuclear Energy, 2014, 70: 21-29.

[5] 朱晟泽. 基于路面宏观纹理的轮胎抗滑行为数值模拟研究[D]. 南京: 东南大学, 2017: 89-114.
ZHU S Z. Numerical simulation of tire skid resistance based on pavement macro-texture[D]. Nanjing: Southeast University, 2017: 89-114 (in Chinese).

[6] 冯停. 湿滑路面轮胎滑水机理研究[D]. 青岛: 青岛理工大学, 2018: 51-62.
FENG T. The study of tire hydroplaning mechanism on wet road[D]. Qingdao: Qingdao Tehcnology University, 2018: 51-62 (in Chinese).

[7] 黄晓明, 刘修宇, 曹青青, 等. 积水路面轮胎部分滑水数值模拟[J]. 湖南大学学报(自然科学版), 2018, 45(9): 113-121.
HUANG X M, LIU X Y, CAO Q Q, et al. Numerical simulation of tire partial hydroplaning on flooded pavement[J]. Journal of Hunan University (Natural Sciences), 2018, 45(9): 113-121(in Chinese).

[8] FWA T F, PASINDU H R, ONG G P. Critical rut depth for pavement maintenance based on vehicle skidding and hydroplaning consideration[J]. Journal of Transportation Engineering, 2012, 138(4): 423-429.

[9] BENEDETTO A. A decision support system for the safety of airport runways: The case of heavy rainstorms[J]. Transportation Research Part A:Policy and Practice, 2002, 36(8): 665-682.

[10] Aviation Safety Network. Accident description [DB/OL]. (2020-08-01) [2020-10-01].https://aviation-safety.net/database/events/dblist.php?Event=LTLW.

[11] 徐征捷, 张友鹏, 苏宏升. 基于云模型的模糊综合评判法在风险评估中的应用[J]. 安全与环境学报, 2014, 14(2): 69-72.
XU Z J, ZHANG Y P, SU H S. Application of risk assessment on fuzzy comprehensive evaluation method based on the cloud model[J]. Journal of Safety and Environment, 2014, 14(2): 69-72(in Chinese).

[12] 王贺, 刘高峰, 王慧敏. 基于云模型的城市极端雨洪灾害预警研究[J]. 水利经济, 2014, 32(4): 59-62,74.
WANG H, LIU G F, WAGN H M. Warning of urban extreme rainstorm flood disasters based on cloud model[J]. Journal of Economics of Water Resources, 2014, 32(4): 59-62,74(in Chinese).

[13] LI D Y, LIU C Y, GAN W Y. A new cognitive model: Cloud model[J]. International Journal of Intelligent Systems, 2009, 24(3): 357-375.

[14] 唐家文, 董兵, 王超峰. 基于云模型的空管安全运行保障能力评价[J]. 航空工程进展, 2021, 12(4): 59-67.
TANG J W, DONG B, WANG C F. Evaluation on safe operation support ability of air traffic mangement based on cloud model[J]. Advances in Aeronautical Science and Engineering, 2021, 12(4): 59-67(in Chinese).

[15] 史佳辉, 徐吉辉, 陈玉金, 等. 基于交互作用矩阵—多维云模型的飞机重着陆风险评估方法研究[J]. 系统工程与电子技术, 2021, 43(10): 3026-3032.
SHI J H, XU J H, CHEN Y J, et al. Research on risk assessment method of aircraft heavy landing based on interaction matrix—multidimensional cloud model[J]. Systems Engineering and Electronics, 2021, 43(10): 3026-3032(in Chinese).

[16] 刘伟, 韩莉. 基于云模型的综合保障评估方法研究[J]. 科技创新与应用, 2017(5): 67-69.
LIU W, HAN L. Research on comprehensive support evaluation method based on cloud model[J]. Technology Innovation and Application, 2017(5): 67-69(in Chinese).

[17] 华攸金, 李希建, 陈刘瑜. 基于变权与正态云理论的煤矿安全评价及应用[J]. 煤矿安全, 2020, 51(3): 239-242.
HUA Y J, LI X J, CHEN L Y. Coal mine safety evaluation and application based on variable weight and normal cloud theory[J]. Safety in Coal Mines, 2020, 51(3): 239-242.

[18] 王洪利, 冯玉强. 基于云模型具有语言评价信息的多属性群决策研究[J]. 控制与决策, 2005, 20(6): 679-681.
WANG H L, FENG Y Q. On multiple attribute group decision making with linguistic assessment information based on cloud model[J]. Control and Decision, 2005, 20(6): 679-681(in Chinese).

[19] LIN C J, ZHANG M, LI L P, et al. Risk assessment of tunnel construction based on improved cloud model[J]. Journal of Performance of Constructed Facilities, 2020, 34(3): 04020028.

[20] 宗一鸣. 湿滑道面条件下轮胎力学行为与飞机着陆安全问题研究[D]. 天津: 中国民航大学, 2017.
ZONG Y M. Study on the mechanical properties of aircraft tire and safety problem in landing on wet-pavement[D]. Tianjin: Civil Aviation University of China, 2017 (in Chinese).

[21] 牛亚东, 张思祥, 田广军, 等. 机场跑道摩擦系数影响因素研究[J]. 应用力学学报, 2021, 38(2): 715-720.
NIU Y D, ZHANG S X, TIAN G J, et al. Research on influencing factors of friction coefficient in airport runway[J]. Chinese Journal of Applied Mechanics, 2021, 38(2): 715-720(in Chinese).

[22] 孙晓, 杨淑芳, 武磊, 等. 基于博弈论-云模型的隧道围岩稳定性分析模型研究[J]. 有色金属工程, 2020, 10(10): 112-119.
SUN X, YANG S F, WU L, et al. Study on the stability analysis model of tunnel surrounding rock based on the combined weighting cloud model of game theory[J]. Nonferrous Metals Engineering, 2020, 10(10): 112-119(in Chinese).

[23] OH C W, KIM T W, JEONG H Y, et al. Hydroplaning simulation for a straight-grooved tire by using FDM, FEM and an asymptotic method[J]. Journal of Mechanical Science and Technology, 2008, 22(1): 34-40.

[24] 梁辉如, 王永东, 彭浩, 等. 基于正态云理论的软弱隧道围岩分级[J]. 重庆交通大学学报(自然科学版), 2021, 40(11): 82-87.
LIANG H R, WANG Y D, PENG H, et al. Classification of soft surrounding rock of tunnel based on normal cloud theory[J]. Journal of Chongqing Jiaotong University (Natural Science), 2021, 40(11): 82-87(in Chinese).

[25] 胡皓, 李德清, 赵娜, 等. 因素状态值为模糊数的变权综合决策方法[J]. 军械工程学院学报, 2014, 26(5): 75-78.
HU H, LI D Q, ZHAO N, et al. Variable weights synthesis decision making based on fuzzy numbers[J]. Journal of Ordnance Engineering College, 2014, 26(5): 75-78(in Chinese).

[26] 李德清, 李洪兴. 状态变权向量的性质与构造[J]. 北京师范大学

学报(自然科学版), 2002, 38(4): 455-461.
LI D Q, LI H X. The properties and construction of state variable weight vectors[J]. Journal of Beijing Normal University (Natural Science), 2002, 38(4): 455-461(in Chinese).

Hydroplaning risk of aircraft tire based on variable weight theory-normal cloud model

LI Yue[1], HU Yuqi[1, 2], CAI Jing[1, 3, *], DAI Xuan[1]

(1. School of Transportation Science and Engineering, Civil Aviation University of China, Tianjin 300300, China;
2. Ground Service Department, Air China Ltd., Beijing 101309, China;
3. Engineering Research Center of Intelligent Construction and Industrialization, CAAC, Tianjin 300456, China)

Abstract: Aiming at the randomness and fuzziness features of hydroplaning behavior of aircraft tires, a new safety evaluation methodology of tire hydroplaning is established on the basis of normal cloud model. The variable weight theory is adopted in order to adjust the weight value dynamically. The penalty function of variable weight is used to reduce the subjective impact of constant weight value on evaluation results. A fluid-solid coupling model of aircraft tire landing on wet pavement is developed for numerical analysis. The selection of five independent variables as the influencing determinants of hydroplaning risk includes aircraft wheel load, taxiing speed, water-film thickness, pavement friction coefficient, and groove depth. The comprehensive membership degree is then calculated based on the digital features of the single-factor cloud model and variable weight vector. The risk level and classification standard of aircraft tire hydroplaning is then established by multivariate decision. The case study of several accident symptoms of tire hydroplaning that took place in an airport located in a rainy mountainous area indicates that only a simplified conclusion of allowed/not-allowed landing or taking off of aircraft can be obtained by using critical hydroplaning speed as an evaluation criterion. By comparing the evaluation results of constant weight and variable weight theory, the safety factor of Case 1 is increased from 1.09 to 1.17 and that of Case 2 is increased from 2.09 to 2.94. The difference in airport runways under different running conditions can be quantitatively described. Although the risk levels of tire hydroplaning are both within the acceptable range, evaluation results derived from variable weight theory can be more conservative. As for Case 3, the safety factor is increased from 3.13 to 3.74 and the risk level is increased to level IV(unacceptable). Although the water-film thickness on the pavement surface complies with the top limit criteria, it is still possible to greatly increase the likelihood of tire hydroplaning, which is compatible with the actual risk scenario. The conclusions can be used as a reference for the classified safety management of pavement operation.

Keywords: normal cloud; variable weight; aircraft hydroplaning; comprehensive membership degree; fluid-solid coupling; airport engineering

Received: 2021-11-10; **Accepted**: 2021-12-27; **Published Online**: 2022-01-27 12:58
URL: kns.cnki.net/kcms/detail/11.2625.V.20220126.0946.001.html
Foundation items: Tianjin Natural Science Foundation (21JCQNJC00850); Fundamental Research Funds for the Central Universities (3122019103)
* **Corresponding author**. E-mail: jcai@cauc.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0685

基于多 Agent 的航空机群保障仿真评估分析

丁刚[1]，崔利杰[1,*]，韩诚[2]，王新河[2]，张峰[2]

(1. 空军工程大学 装备管理与无人机工程学院，西安 710051；
2. 西北工业大学 力学与土木建筑学院，西安 710129)

摘 要：航空机群保障涉及要素广、策略多、交互性强，仿真分析方法是开展飞机保障决策与评估研究的热点和难点。提出一种机群保障仿真评估方法，建立多智能体(Agent)功能类型与交互关系模型；对多 Agent 结构进行定义，并建立模型；以五型飞机构成的航空机群为仿真对象，以任务成功率为保障指标进行案例验证与仿真计算分析。结果表明：各型飞机平均故障间隔时间（MTBF）、设备故障平均修复时间（MTTR）对任务成功率的影响总体趋势相似，随着 MTBF 的增加任务成功率提升，随着 MTTR 的增加任务成功率降低。所提方法能够为航空机群维修保障智能决策提供一种可行、有效的方法手段，支撑基于模型的智能决策优化实现。

关 键 词：航空机群；保障；多智能体；仿真评估；敏感性分析
中图分类号：V221^{+}.3; TB553
文献标志码：A **文章编号**：1001-5965(2023)09-2306-11

在现代战争中，多平台、高精密设备已成为主导，多设备协同作战相较于单设备作战效果愈发显著[1]。航空机群是面向执行一定任务的航空机群，由相互关联、功能互补的航空保障设备或航空机群，按照能力需求牵引、综合集成优化原则形成的支撑航空机群保障功能的有机整体[2-3]。航空机群的保障能力直接影响和制约航空机群的战斗力，有效评估航空机群保障性一直以来都是航空设备领域的关注重点之一[4]。针对机群保障体系规模宏大、复杂度高的特点，采用仿真的方法探究体系能力的作用机理，研究各类敏感性因素对机群能力的贡献度，能够有效检验与评估机群的作战能力和效能[5-6]。

智能体(Agent)[7]仿真理论和方法有效地实现了复杂系统运行过程仿真，近年来，相关学者运用 Agent 技术解决了诸多方面的问题，如行为控制[8-9]、军事仿真[10-13]、人工智能[14-15]、智能进化[16]和控制架构[17]等。多 Agent 系统[18]是一组自主的、相互作用的实体，这些实体共享一个环境，利用传感器感知，利用执行器行动，并赋予预先设计的行为，且不断学习，使得多 Agent 的性能逐步提高，这种系统提供了一种分析问题的分布式视角，将控制权限分布在各个 Agent 上。多 Agent 解决方案被广泛应用于生产物流车间调度[19-20]、人机交互[21-22]、强化学习[23]、疫情防控[24]和智能监控[25]等领域。

近年来对设备保障体系建模与仿真方面的研究有很多。文献[26]构建了基于面向对象的 PetriNet (object-oriented PetriNet, OPN)设备维修保障系统 Petri 网模型，并利用 CPN-Tools(colored

收稿日期：2021-11-15；录用日期：2022-03-04；网络出版时间：2022-03-22 15:24
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220321.1807.002.html
基金项目：技术基础科研计划(212KJ43005)
*通信作者. E-mail：baffulo@sina.com

引用格式：丁刚，崔利杰，韩诚，等. 基于多 Agent 的航空机群保障仿真评估分析[J]. 北京航空航天大学学报，2023，49(9)：2306-2316. DING G，CUI L J，HAN C，et al. Simulation evaluation and analysis of aircraft group support based on multi-agent[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2306-2316 (in Chinese).

PetriNet-tools)验证了其动态性能。文献 [27] 从系统工程建模和仿真的角度, 建立了军事设备保障仿真概念模型。文献 [28] 提出一种美国国防部架构框架 (Department of Defense Architecture Framework, DoDAF) 建模框架, 该框架基于 DoDAF 元模型 (DoDAF meta model, DM2), 并定义了所有建模概念及其关系。文献 [29] 以数据为中心的体系结构建模思路, 提出体系结构的语义模型, 并构建作战视图、系统视图、能力视图本体模型。

以上研究大多基于离散系统, 采用离散系统仿真方法, 体系的状态是由离散事件驱动的, 难以发现与反映机群保障的非线性、“涌现”性、自适应性及自组织性等特点[30]。多 Agent 技术是一种自底向上的建模方法, 将多 Agent 技术应用于机群保障体系仿真中能够具有很强的适应性, 不少学者尝试将多 Agent 技术应用于设备保障体系中。文献 [31] 针对设备保障运行的实际情况, 采用多 Agent 仿真的方法, 构建了设备保障体系供应保障系统。文献 [32] 设计了基于多 Agent 的导弹设备军民一体化维修保障决策支持系统的功能模块和体系结构, 并分析了决策支持系统的决策过程。文献 [33] 将基于 Agent 的离散事件仿真建模技术, 引入设备综合保障系统可靠性、维修性、保障性 (reliability, maintainability and supportability, RMS) 的建模与仿真研究中。文献 [30] 提出了基于多 Agent 分布式建模与仿真的体系架构, 并基于 JavaAgent 开发框架 (JavaAgent DEvelopment Framework, JADE) 开源平台设计并实现了基于多 Agent 的设备保障体系分布式仿真软件原型系统。

通过以上研究发现, 研究的对象主要集中于单飞机层面的设备保障, 针对航空机群保障领域, 还缺乏一种有效的建模与仿真方法和系统。多 Agent 技术可以模拟航空机群保障系统的每个单元, 并实现相互之间复杂的交互关系。目前将多 Agent 技术应用于航空机群保障建模方法的相关研究还比较少, 因此, 本文针对航空机群保障展开研究, 提出一种基于多 Agent 的航空机群保障仿真评估分析方法, 以五型 40 架飞机组成空战作战机群, 以任务成功率为机群保障评估指标, 建立多 Agent 机群保障仿真模型及其求解算法, 最后进行仿真试验, 验证本文方法的合理性和有效性。

1　航空机群多 Agent 功能类型与交互关系建模

1.1　多 Agent 功能类型设计

多 Agent 模型主要用来表示 2 类实体。一类是保障系统中具有心智概念和理性特征的实体, 通过规划、推理和决策实现 Agent 功能, 表达指挥人员、设备管理人员、设备乘员、保障人员和维修保障管理人员的智能行为, 学习优化任务下达、保障指挥、保障监控、维修实施、器材供应等过程, 需要综合各种约束进行资源的调度和分配, 其协作关系也非常复杂; 另一类是保障系统中通过相对简单的代理之间交互实现体系的智能行为, 内部不需要复杂的推理模型, 如保障设备、备品备件、保障设施等。航空机群保障性研究的目的是解决构成航空机群的系统及其要素相互作用、相互影响产生的随时间推移而涌现出的体系属性、状态及行为变化。通过对这些属性、状态和行为的分析预测, 采取一定的控制方法, 引导体系产生出有益的或期望的涌现。进行航空机群保障多 Agent 建模的起点是体系任务, 主线是随时间线推进不断发生的各种事件, 基础是体系运行涉及的众多实体及其行为, 重点是形式化的准确描述各种场景下的实体约束, 通过对这些约束的调整来控制引导体系涌现。其中多 Agent 建模的难点则在于面向研究目标刻画各个 Agent 之间的交互关系, 设计 Agent 内部的规则、推理和决策模型。

1.2　面向航空机群的多 Agent 交互关系设计

面向航空机群保障研究目标, 本文构建的多 Agent 模型包括体系保障性 Agent、保障要素 Agent 和保障行为 Agent 3 大类。体系保障性 Agent 模型包括保障任务 Agent、保障监控 Agent、保障指挥 Agent 和飞机选取 Agent; 保障要素 Agent 模型包括组织机构 Agent、机群 Agent、使用人员 Agent、维修机构 Agent、仓库 Agent、保障资源 Agent、保障设备 Agent 和保障人员 Agent; 保证行为 Agent 模型包括预防性维修 Agent、修复性维修 Agent、器材供应 Agent 和弹药供应 Agent。

以保障任务 Agent 为例, 其目标为保障任务的生成。功能包括从作战指挥机关接收任务想定, 解析作战任务、分解任务阶段、产生保障任务、生成保障设备体系对象。交互包括向指挥机关反馈任务完成情况; 向保障监控 Agent 传输任务数据; 向保障指挥 Agent 发送任务状态消息, 接收决策消息等。航空机群保障 Agent 基本功能类型和交互关系如图 1 所示。

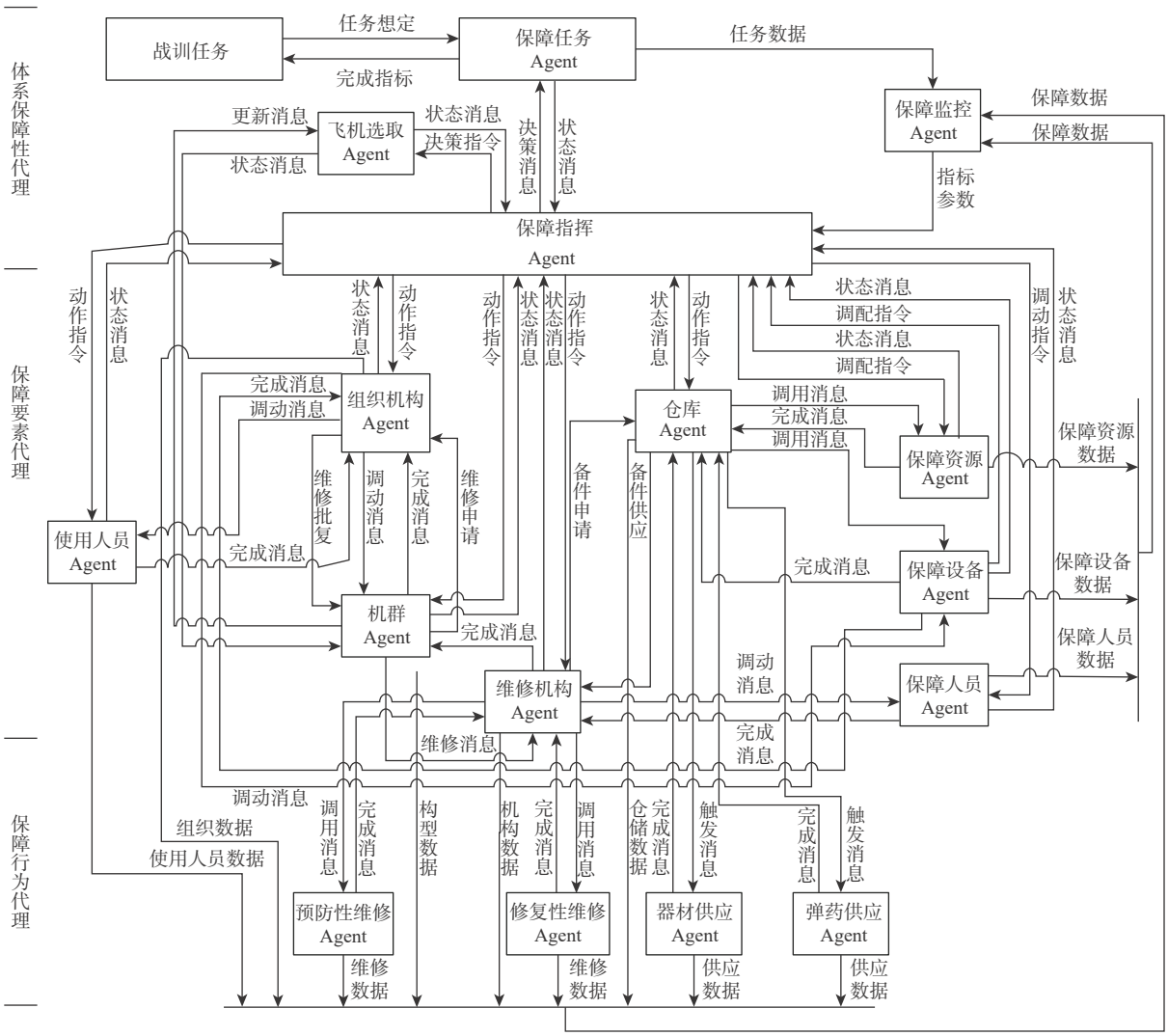


图 1 设备保障 Agent 基本功能类型和交互关系

Fig. 1 Basic function types and interaction relationships of equipment support Agent

2 航空机群的多 Agent 结构定义与控制模型设计

航空机群保障建模涉及众多 Agent 内部结构、交互属性与学习算法的设计，为了实现传统由保障指挥人员和保障人员人工控制和决策的功能，自然考虑到将复杂的要素控制、流程模拟和逻辑决策固化到各类 Agent 内部实现的算法模块中。以 Agent 基本结构框架为基础，从易于构造和实现的角度出发，采取一种基于目标驱动具有内部状态的 Agent 结构[34]，内部状态作为历史信息与当前感知共同决定下一步动作，使 Agent 反应更具理性；同时通过目标完成度的评价，实现 Agent 的自适应。以飞机选取 Agent 设计、机群 Agent 和保障监控 Agent 设计为例，阐明 Agent 设计架构及与其他 Agent 的交互设计。

2.1 飞机选取 Agent 设计

航空机群中飞机选取重点考虑阶段飞机任务要求，结合飞机动用的影响因素 (年度总飞行时间、飞机平均规定寿命值和飞机余寿等)，选取执行任务的飞机。飞机选取 Agent 首先从保障任务 Agent 获取决策指令，解析各机型在各任务阶段所需机型数量，根据规则库获取年度总飞行时间 T_{all}、平均规定寿命 T_{aver}、飞机总数 N 和阶段剩余寿命 T_{r} 等。通过飞机选取 Agent 内部推理确定年度机队总余寿与总寿命比值(以下简称“余寿比”)，进而确认飞机出动策略，最后给出满足各任务阶段数量要求的飞机编号，并将该数据记录下来，用于与其他 Agent 交互。飞机选取 Agent 功能设计如图 2 所示。

第 n 架飞机的年度标准余寿比 Q 为

$$Q=\left(\frac{T_{\text{all}}}{N}+\frac{0.8T_{\text{aver}}}{N}\cdot\frac{0.8}{N}\right)n \tag{1}$$

飞机出动的计算式为

$$\max_{M}(T_{\text{r}}-Q) \tag{2}$$

式中：M 为本科目计划出动的飞机架数。

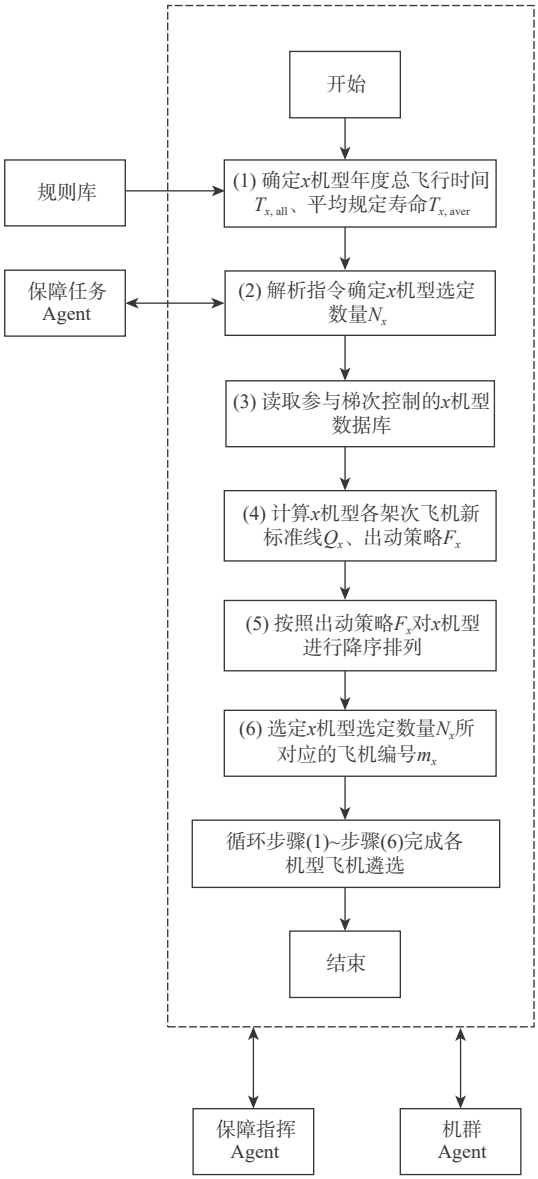


图 2　飞机选取 Agent 功能设计

Fig. 2　Aircraft selection Agent function design

2.2　机群 Agent 设计

航空机群 Agent 作用在于构建完成体系任务的符合各任务阶段的航空机群构型, 并给出机群的实时状态。功能包括向保障任务 Agent 获取各机型数量及各任务阶段时间; 向保障指挥 Agent 提供构型状态, 并接收来自保障指挥 Agent 的命令; 根据动作指令向组织机构 Agent 发出维修申请, 并能接收来自组织机构 Agent 的调用信息; 向维修机构 Agent 发出维修消息, 并能接收维修机构 Agent 发送的完成信息; 向保障监控 Agent 提供保障数据。

航空机群状态矩阵即考虑飞机发生故障或战损, 保障指挥机构调动保障资源和库存物流资源组织维修保障活动, 修复完毕后立即返航到指定空域整个任务阶段内的状态序列。

例如有 4 架 A 型飞机, $[0,T]$时间段内状态序列时序分别为

$$\left[0\cdots t_{a1}\cdots t_{b1}\cdots t_{a1}^{1}\cdots t_{b1}^{1}\cdots t_{a1}^{2}\cdots t_{b1}^{2}\cdots t_{a1}^{n}\cdots t_{b1}^{n}\cdots T\right]$$
$$\left[0\cdots t_{a2}\cdots t_{b2}\cdots t_{a2}^{1}\cdots t_{b2}^{1}\cdots t_{a2}^{2}\cdots t_{b2}^{2}\cdots t_{a2}^{n}\cdots t_{b2}^{n}\cdots T\right]$$
$$\left[0\cdots t_{a3}\cdots t_{b3}\cdots t_{a3}^{1}\cdots t_{b3}^{1}\cdots t_{a3}^{2}\cdots t_{b3}^{2}\cdots t_{a3}^{n}\cdots t_{b3}^{n}\cdots T\right]$$
$$\left[0\cdots t_{a4}\cdots t_{b4}\cdots t_{a4}^{1}\cdots t_{b4}^{1}\cdots t_{a4}^{2}\cdots t_{b4}^{2}\cdots t_{a4}^{n}\cdots t_{b4}^{n}\cdots T\right]$$

式中: a 和 b 分别为 A 型飞机故障开始时刻和维修后返回任务状态的时刻; t_a^n和t_b^n分别为 A 型飞机第 n 次故障时间和维修后返回的时间, 且满足: $t_b^n = t_a^n + t_{\text{ammt}}^n + t_{\text{amldt}}^n$, 其中, t_{ammt}^n和t_{amldt}^n分别为 A 型飞机第 n 次故障的修复时间和后勤延误时间。

生成 4 架 A 型飞机状态矩阵, 矩阵的行代表 4 架飞机; 列代表飞机任务状态, 状态时间间隔为 1 min, 如果时刻在无故障区间, 状态标记为 1, 非无故障状态标记为 0, 即

若$t \in [t_{a1},t_{b1}] \cup \left[t_{a1}^{1},t_{b1}^{1}\right] \cup \cdots \left[t_{a1}^{n},t_{b1}^{n}\right]$, 则$a_{ij}=0$;

若$t \notin [t_{a1},t_{b1}] \cup \left[t_{a1}^{1},t_{b1}^{1}\right] \cup \cdots \left[t_{a1}^{n},t_{b1}^{n}\right]$, 则$a_{ij}=1$;

则可以得到飞机状态矩阵为

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & \cdots & a_{1T} \\ a_{21} & a_{22} & \cdots & a_{2n} & \cdots & a_{2T} \\ a_{31} & a_{32} & \cdots & a_{3n} & \cdots & a_{3T} \\ a_{41} & a_{42} & \cdots & a_{4n} & \cdots & a_{4T} \end{bmatrix}$$

假设某任务阶段时间为 [0, t], 对应机群构型为这 4 架飞机, 此阶段任务需要 4 架飞机中至少 3 架无故障工作。则遍历 4 架飞机状态矩阵, t 时刻对应矩阵的第 j 列, 判断 4 架飞机完好状态架数State_A的值:

如果$\text{State}_\text{A}=a_{1j}+a_{2j}+a_{3j}+a_{4j}\geqslant 3$,则阶段任务可以完成;

如果$\text{State}_\text{A}=a_{1j}+a_{2j}+a_{3j}+a_{4j}<3$, 则阶段任务失败。

对于不同任务阶段构型的飞机计算步骤相同。如增加 S 型飞机 6 架, 要求至少 5 架无故障工作, 则同样构建 6 架 S 型飞机状态矩阵, 遍历状态矩阵 $\boldsymbol{A}$ 和状态矩阵 $\boldsymbol{S}$, 判断State_A和 6 架飞机完好状态架数State_S的值。

如果 $\text{State}_\text{A}=a_{1j}+a_{2j}+a_{3j}+a_{4j}\geqslant 3$, 且 $\text{State}_\text{S}=s_{1j}+s_{2j}+s_{3j}+s_{4j}+s_{5j}+s_{6j}\geqslant 5$同时满足, 则阶段任务可以完成, 否则阶段任务失败。

航空机群 Agent 外部交互及内部工作原理如图 3 所示。

2.3　航空机群保障监控 Agent 设计

航空机群保障监控 Agent 从机群 Agent 获取机群状态信息, 通过解析状态信息获取机型体系状态时间序列和任务构型参数, 输入总仿真次数 N_s, 通过该 Agent 内部推理判断, 记录任务成功次数 n_s, 最终计算出任务成功率并记录, 用于与其他 Agent 交互。

以计算任务成功率为例, 通过生成的机群时间

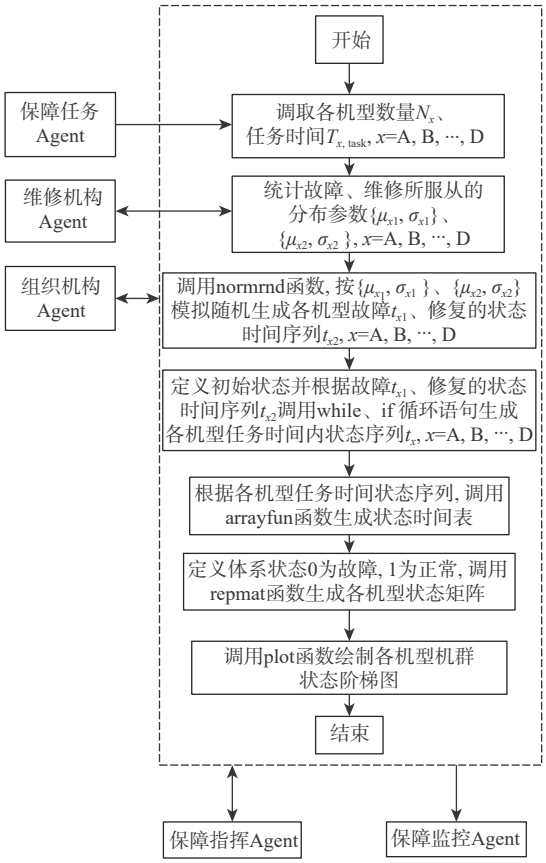


图3 航空机群Agent外部交互和内部工作原理

Fig. 3 Aircraft group Agent external interaction and internal working principle

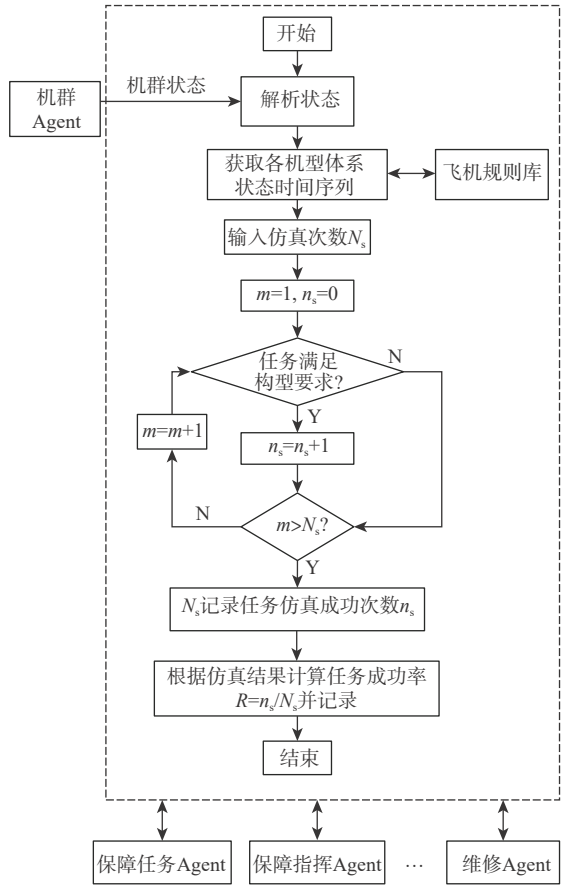


图4 保障监控Agent设计

Fig. 4 Guarantee monitoring Agent design

序列, 以各机型任务阶段构型为约束, 通过判断在仿真过程中各机型是否满足构型要求,来确认任务成功与失败, 任务成功率 R 的计算式为

$$R=\frac{n_s}{N_s} \tag{3}$$

保障监控Agent设计如图4所示。

Agent设计在流程和方法上类似, 区别在于功能与智能化要求, 输入与输出、交互Agent设计, 内部实现则主要体现在学习算法和行为模型的不同。

3 航空机群保障仿真评估分析

3.1 航空机群保障案例想定

某次作战行动要求空中作战集团出动A、B、C、D、E五型飞机共40余架组成空中编队执行任务, 任务持续时间约26 h, 包括预警探测阶段12 h、前出侦查阶段8 h和远程攻击阶段6 h。仿真计算初始输入数据如表1所示, 表中 n 为要求出动的飞机架数, k 为任务完成要求的最少完好飞机架数。

表1 仿真计算初始输入

Table 1 Initial input of simulation calculation

机型	数量/架	滞空时间/min	任务时间/min			构型要求(k/n)			MTBF/h	MTTR/h
			第1阶段	第2阶段	第3阶段	第1阶段	第2阶段	第3阶段		
A	3	210	12	8	6	1/2	1/2	2/3	[300,20]	[10,2]
B	10	210			6			7/10	[330,25]	[6,2]
C	16	90	12	8	6	5/8	7/10	10/16	[200,30]	[5,1]
D	8	180		8	6		4/6	6/8	[300,50]	[7,2]
E	6	210			6			4/6	[350,30]	[8,2]

本文考虑一定任务条件下的平均故障时间间隔(mean time between failure, MTBF)和平均修复时间(mean time to repair, MTTR)值, 取值来源于设备实际使用过程中质量控制部门统计得到的样本, 进

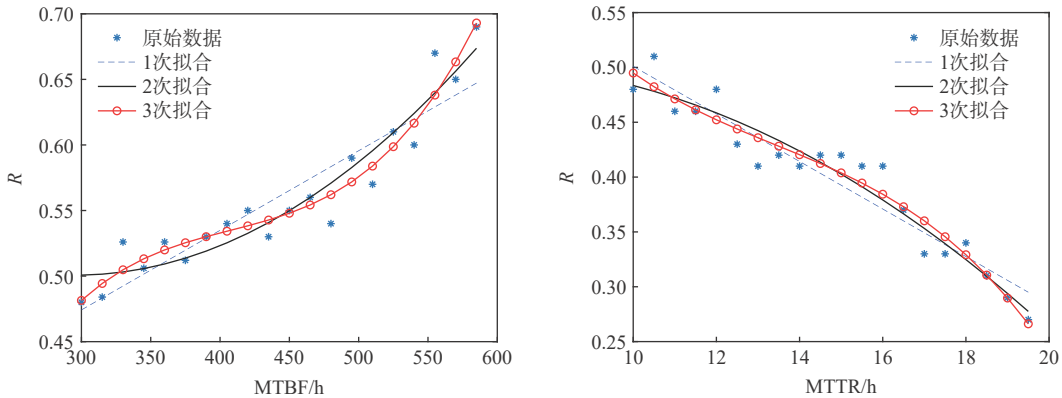
行概率拟合得出。

3.2　任务成功率敏感性分析

3.2.1　各型飞机可靠性维修性任务成功率敏感性分析

对各型飞机基本作战单元典型任务场景进行仿真数据敏感性分析, 将 MTBF 和 MTTR 服从的正态分布均值进行步长为 5% 的调整, 以任务成功率为输出指标, 分别进行 30 次仿真, 对数据进行拟合, 得到各型飞机任务成功率敏感性分析图, 如图 5～图 15 所示。

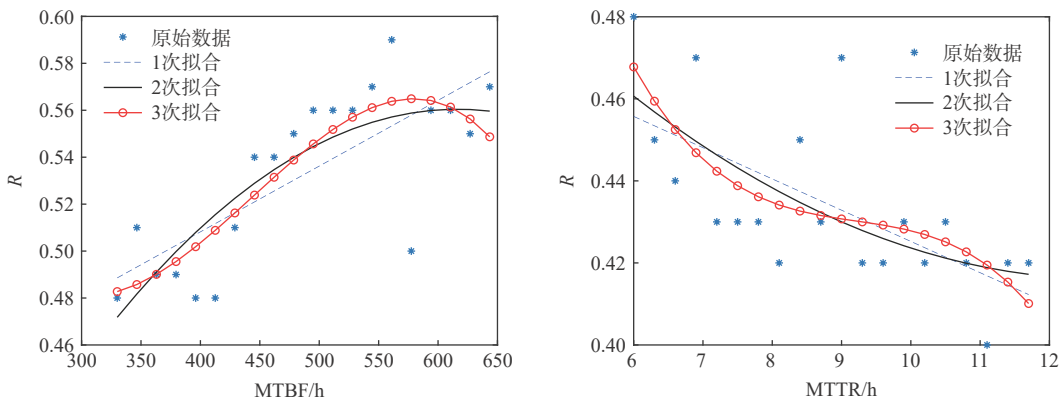
通过仿真结果可知, 各型飞机 MTBF 和 MTTR 对任务成功率的影响总体趋势相似, 随着 MTBF 的增加任务成功率提升, 随着 MTTR 的增加任务成功率降低。由于各型飞机初始可靠性、维修性、任务强度参数要求不同, 拟合曲线在局部呈现不同的变



(a) A型-MTBF　　(b) A型-MTTR

图 5　A 型-MTBF 和 A 型-MTTR 步长 5% 任务成功率敏感性分析

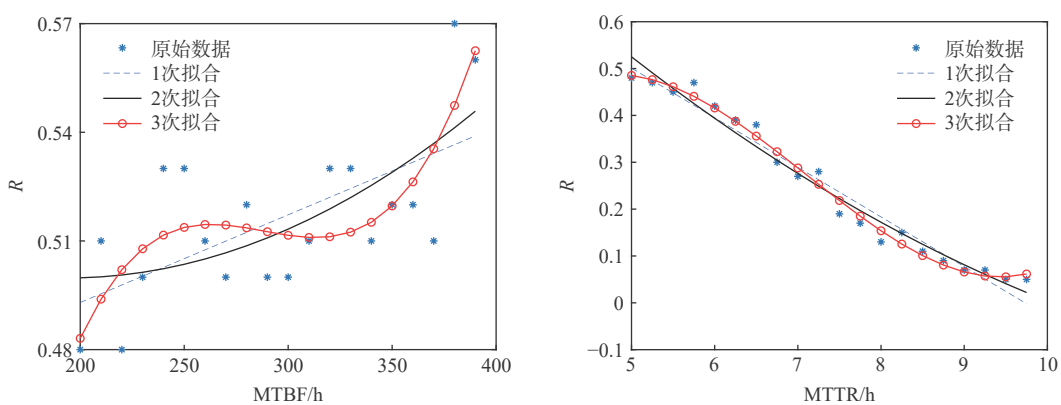
Fig. 5　Sensitivity analysis of A type-MTBF and A type-MTTR step size 5% mission success rate



(a) B型-MTBF　　(b) B型-MTTR

图 6　B 型-MTBF 和 B 型-MTTR 步长 5% 任务成功率敏感性分析

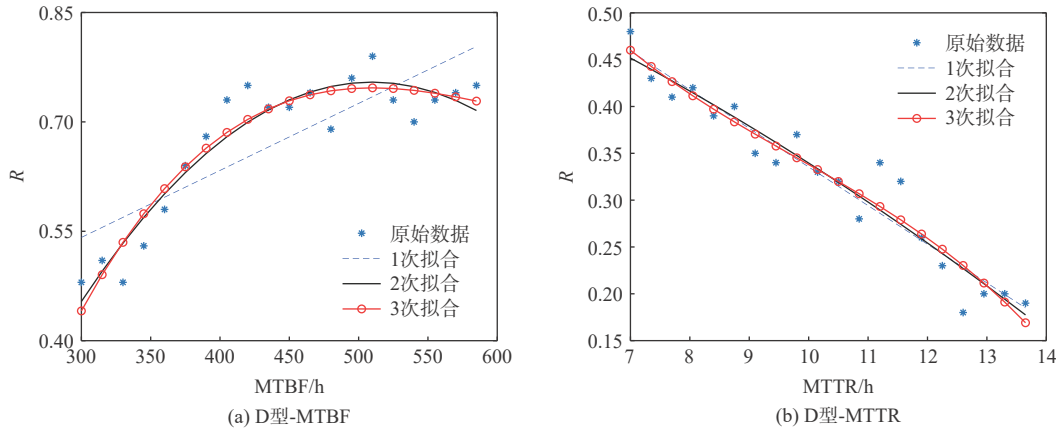
Fig. 6　Sensitivity analysis of B type-MTBF and B type-MTTR step size 5% mission success rate



(a) C型-MTBF　　(b) C型-MTTR

图 7　C 型-MTBF 和 C 型-MTTR 步长 5% 任务成功率敏感性分析

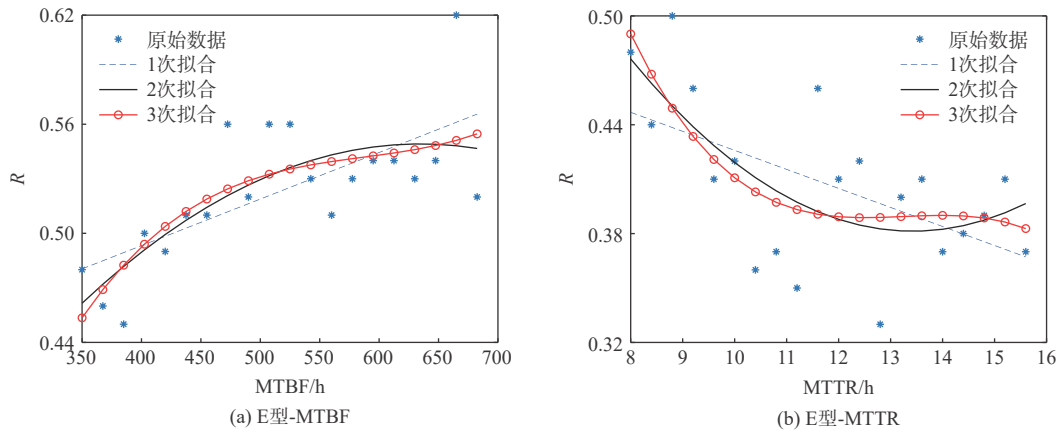
Fig. 7　Sensitivity analysis of C type-MTBF and C type-MTTR step size 5% mission success rate



(a) D型-MTBF (b) D型-MTTR

图 8 D 型-MTBF 和 D 型-MTTR 步长 5% 任务成功率敏感性分析

Fig. 8 Sensitivity analysis of D type-MTBF and D type-MTTR step size 5% mission success rate



(a) E型-MTBF (b) E型-MTTR

图 9 E 型-MTBF 和 E 型-MTTR 步长 5% 任务成功率敏感性分析

Fig. 9 Sensitivity analysis of E type-MTBF , and E type-MTTR step size 5% mission success rate

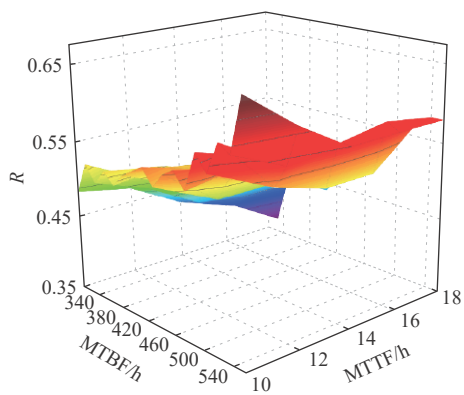


图 10 A 型-MTBF-5% 和 A 型-MTTR-20% 步长任务成功率敏感性分析

Fig. 10 Sensitivity analysis of A type-MTBF-5% and A type-MTTR-20% step size mission success rate

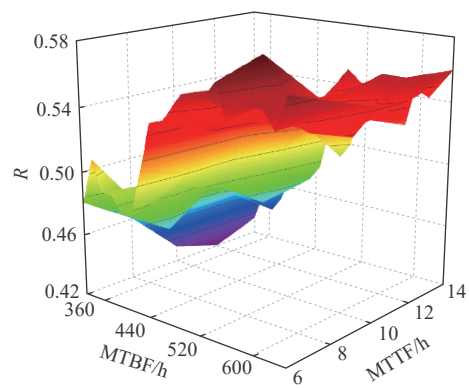


图 11 B 型-MTBF-5%、B 型-MTTR-20% 步长任务成功率敏感性分析

Fig. 11 Sensitivity analysis of B type-MTBF-5% and B type-MTTR-20% step size mission success rate

化规律。以 A 型飞机为例，当 A 型飞机 MTBF 服从的均值增加时，故障间隔时间增加，体系任务成功率呈现上升的趋势，从初始 0.48 上升至 0.69，当 MTBF=585 h 时，任务成功率最高达到 0.69；当 MTTR 服从的均值增加，修复时间逐步增加，体系任务成功率呈现下降趋势，从初始 0.48 下降至 0.27，当 MTTR=19.5 h 时，任务成功率最低到 0.27；当对 MTBF 服从的均值按照 5%，MTTR 服从的均值同步按照 20% 步长进行变化，即 MTBF 增加、MTTR 减少，通过得到的三维颜色映射曲面图进行综合敏感性分

析得知, 体系任务成功率呈现上升趋势, 当 [MTBF, MTTR] 取值为 [300,18] h 时任务成功率取最小值, 为 0.37, 当 [MTBF, MTTR] 取值为 [585,12] h 时任务成功率取最大值, 为 0.69。

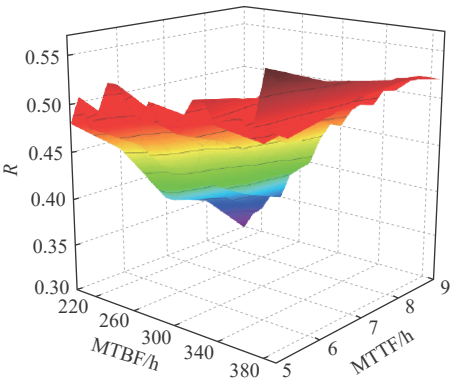


图 12　C 型-MTBF-5%、C 型-MTTR-20% 步长任务成功率敏感性分析

Fig. 12　Sensitivity analysis of C type-MTBF-5% and C type-MTTR-20% step size mission success rate

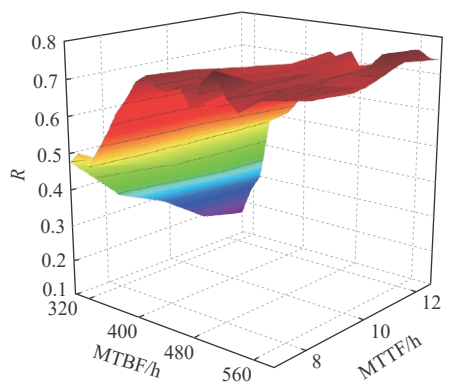


图 13　D 型-MTBF-5%、D 型-MTTR-20% 步长任务成功率敏感性分析

Fig. 13　Sensitivity analysis of D type-MTBF-5% and D type-MTTR-20% step size mission success rate

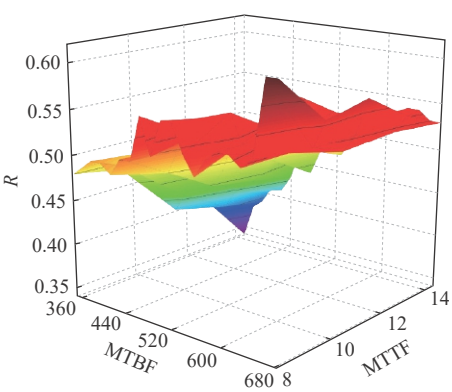


图 14　E 型-MTBF-5%、E 型-MTTR-20% 步长任务成功率敏感性分析

Fig. 14　Sensitivity analysis of E type-MTBF-5% and E type-MTTR-20% step size mission success rate

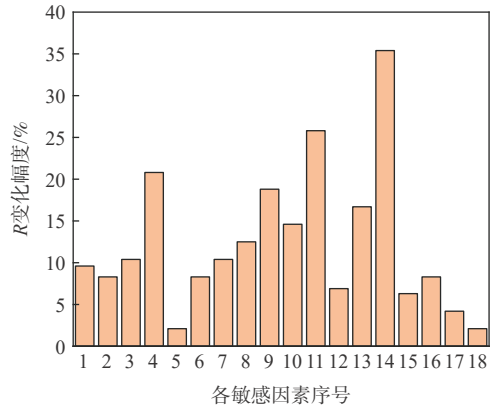


图 15　综合敏感性分析图

Fig. 15　Comprehensive sensitivity analysis chart

3.2.2　机型与任务强度敏感性对比分析

对各机型任务强度及各阶段任务时间进行步长为 20% 的调整, 进行仿真计算。各敏感性因素步长 20% 综合敏感性分析图如图 15 所示, 计算统计数据如表 2 所示。

表 2　任务强度敏感性分析

Table 2　Mission intensity sensitivity analysis

类别	序号	敏感性因素	*R*		*R*变化幅度/%
			初始值	变化值	
各机型MTBF	1	A型-MTBF增加20%(300～360)	0.48	0.526	9.6
	2	B型-MTBF增加20%(330～396)	0.48	0.52	8.3
	3	C型-MTBF增加20%(200～240)	0.48	0.53	10.4
	4	D型-MTBF增加20%(300～360)	0.48	0.58	20.8
	5	E型-MTBF增加20%(350～420)	0.48	0.49	2.1
各机型MTTR	6	A型-MTTR增加20%(10～12)	0.48	0.44	−8.3
	7	B型-MTTR增加20%(6～7.2)	0.48	0.43	−10.4
	8	C型-MTTR增加20%(5～6)	0.48	0.42	−12.5
	9	D型-MTTR增加20%(7～8.4)	0.48	0.39	−18.8
	10	E型-MTTR增加20%(8～9.6)	0.48	0.41	−14.6
任务强度	11	A型飞机任务强度降低20%(2/3→1/3)	0.48	0.60	25.8
	12	B型飞机任务强度降低20%(5/10→4/10)	0.48	0.51	6.9
	13	C型飞机任务强度降低20%(4/10→3/10)	0.48	0.56	16.7
	14	D型飞机任务强度降低20%(6/8→5/8)	0.48	0.65	35.4
	15	E型飞机任务强度降低20%(4/6→3/6)	0.48	0.51	6.3
阶段任务时间	16	第1阶段时间减少20%(9.6 h)	0.48	0.52	8.3
	17	第2阶段时间减少20%(6.4 h)	0.48	0.50	4.2
	18	第3阶段时间减少20%(4.8 h)	0.48	0.49	2.1

4 结 论

1）本文方法不仅适用于航空机群，同样可推广至其他类似体系和复杂系统中。所提Agent模型构建与实现，针对不同实体或行为进行分析提炼后可推广至其他仿真活动中。本文围绕典型想定所开展的敏感度分析只列举了个别变量，且仅以随机变量的形式进行抽样计算，如对其他变量进行分析，还需对其取值空间和不确定性形式进行细化，以便得到更加精准的结果，为实际保障决策与评估提供科学有效的支持。

2）本文方法能够为体系化作战开展预先保障推演或实时保障决策提供一种研究思路。另外，提出的保障功能与交互行为机制、Agent模型实现、敏感性评估等可作为体系级保障仿真的一种有益探索，支撑保障决策向更准、更快、更科学转变。

参考文献（References）

[1] 汤润泽, 张承龙, 李林林, 等. 多武器跨域智能协同对空作战应用及关键技术[J]. 现代防御技术, 2021, 49(2): 26-34.
TANG R Z, ZHANG C L, LI L L, et al. Research on applications and key technologies of multi-weapons in cross-domain intelligent coordination air combat[J]. Modern Defence Technology, 2021, 49(2): 26-34(in Chinese).

[2] 丁刚, 张琳, 崔利杰, 等. 航空装备体系保障性分析仿真运行控制[J]. 舰船电子工程, 2021, 41(1): 58-62.
DING G, ZHANG L, CUI L J, et al. Simulation operation control of aviation equipment system of system supportability analysis[J]. Ship Electronic Engineering, 2021, 41(1): 58-62(in Chinese).

[3] 丁刚, 张琳, 崔利杰, 等. 基于任务的航空装备保障体系概念建模研究[J]. 军事运筹与系统工程, 2020, 34(1): 39-46.
DING G, ZHANG L, CUI L J, et al. Research on concept modeling of mission-based aviation equipment support system of systems[J]. Military Operations Research and Systems Engineering, 2020, 34(1): 39-46(in Chinese).

[4] 崔利杰, 丛继平, 丁刚, 等. 基于不确定性的航空装备体系保障性评估[J]. 北京航空航天大学学报, 2021, 47(12): 2452-2461.
CUI L J, CONG J P, DING G, et al. Supportability evaluation of aviation equipment system based on uncertainty[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(12): 2452-2461(in Chinese).

[5] 杨东昌, 马永忠, 宋科. 武器装备体系需求论证研究[J]. 中国设备工程, 2021(12): 250-251.
YANG D C, MA Y Z, SONG K. Research on requirement demonstration of weapon equipment system[J]. China Plant Engineering, 2021(12): 250-251(in Chinese).

[6] 姜忠钦, 张明智, 张斌, 等. 基于仿真的武器装备体系论证关键问题研究[J]. 装备指挥技术学院学报, 2009, 20(5): 19-23.
JIANG Z Q, ZHANG M Z, ZHANG B, et al. Study on key questions of weapon system-of-systems demonstration based on simulation[J]. Journal of the Academy of Equipment Command & Technology, 2009, 20(5): 19-23(in Chinese).

[7] KAVIARI F, MESGARI M S, SEIDI E, et al. Simulation of urban growth using agent-based modeling and game theory with different temporal resolutions[J]. Cities, 2019, 95: 102387.

[8] NGUYEN N D, NGUYEN T, NAHAVANDI S. Multi-agent behavioral control system using deep reinforcement learning[J]. Neurocomputing, 2019, 359: 58-68.

[9] HUANG J, CAO M, ZHOU N, et al. Distributed behavioral control for second-order nonlinear multi-agent systems[J]. IFAC-PapersOnLine, 2017, 50(1): 2445-2450.

[10] 黄柯棣, 刘宝宏, 黄健, 等. 作战仿真技术综述[J]. 系统仿真学报, 2004, 16(9): 1887-1895.
HUANG K D, LIU B H, HUANG J, et al. A survey of military simulation technologies[J]. Acta Simulata Systematica Sinica, 2004, 16(9): 1887-1895(in Chinese).

[11] 李宏亮, 程华, 金士尧. 基于Agent的复杂系统分布仿真建模方法的研究[J]. 计算机工程与应用, 2007, 43(8): 209-213.
LI H L, CHENG H, JIN S Y. Research of modeling method for Agent-based complex systems distributed simulation[J]. Computer Engineering and Applications, 2007, 43(8): 209-213(in Chinese).

[12] 饶明波. 基于Agent战场仿真实体模型研究与实现[D]. 南京: 南京航空航天大学, 2011.
RAO M B. Agent-based battlefield simulation entity model research and implementation[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2011 (in Chinese).

[13] 陈坚, 廖守亿, 邓方林. 基于多Agent系统的计算机生成兵力建模研究[J]. 系统工程与电子技术, 2008, 30(10): 1924-1928.
CHEN J, LIAO S Y, DENG F L. On modeling computer generated forces based on multi-agent systems[J]. Systems Engineering and Electronics, 2008, 30(10): 1924-1928(in Chinese).

[14] 阿纳达·佩雷兹·卡斯塔诺. 人工智能实战[M]. 敖富江, 周云彦, 李博, 等, 译, 北京: 清华大学出版社, 2019: 27-29, 51-73.
AMALDO P C. Practical artificial intelligence[M]. AO F J, ZHOU Y Y, Li B, et al. translated. Beijing: Tsinghua University Press, 2019: 27-29, 51-73 (in Chinese).

[15] 武强. 多智能体强化学习在城市交通信号控制中的研究与应用[D]. 兰州: 兰州大学, 2020.
WU Q. Research and application of multi-agent reinforcement learning in traffic signal control[D]. Lanzhou: Lanzhou University, 2020 (in Chinese).

[16] 梁鹏, 王海燕, 王耀辉. 由反应式Agent构造智能Agent的方法[J]. 东北电力大学学报(自然科学版), 2008, 28(2): 94-97.
LIANG P, WANG H Y, WANG Y H. The method by reflection type agent structure intelligence agent[J]. Journal of Northeast Dianli University (Natural Science Edition), 2008, 28(2): 94-97(in Chinese).

[17] MOSTAFA S A, MUSTAPHA A, GUNASEKARAN S S, et al. An agent architecture for autonomous UAV flight control in object classification and recognition missions[J]. Soft Computing, 2023, 27(1): 391-404.

[18] TAJALLI S Z, KAVOUSI-FARD A, MARDANEH M. Multi-agent-based optimal power scheduling of shipboard power systems[J]. Sustainable Cities and Society, 2021, 74: 103137.

[19] SHPILEVOY V, SHISHOV A, SKOBELEV P, et al. Multi-agent system “Smart Factory” for real-time workshop management in

aircraft jet engines production[J]. IFAC Proceedings Volumes, 2013, 46(7): 204-209.

[20] LU Z Y, ZHUANG Z L, HUANG Z Z, et al. A framework of multi-agent based intelligent production logistics system[J]. Procedia CIRP, 2019, 83: 557-562.

[21] ZHANG X, SUN Y C, ZHANG Y J, et al. Multi-agent modelling and situational awareness analysis of human-computer interaction in the aircraft cockpit: A case study[J]. Simulation Modelling Practice and Theory, 2021, 111: 102355.

[22] LIU X, GE S S. Optimized control for human-multi-robot collaboration via multi-agent adaptive dynamic programming[J]. IFAC-PapersOnLine, 2020, 53(2): 9207-9212.

[23] ZHANG L X, LI J C, ZHU Y A, et al. Multi-agent reinforcement learning by the actor-critic model with an attention interface[J]. Neurocomputing, 2022, 471: 275-284.

[24] CASTRO B M, MELO Y D AD, SANTOS N F D, et al. Multi-agent simulation model for the evaluation of COVID-19 transmission[J]. Computers in Biology and Medicine, 2021, 136: 104645.

[25] VALLEJO D, CASTRO-SCHEZ J J, GLEZ-MORCILLO C, et al. Multi-agent architecture for information retrieval and intelligent monitoring by UAVs in known environments affected by catastrophes[J]. Engineering Applications of Artificial Intelligence, 2020, 87: 103243.

[26] NING H, SHI Y B, GAO X J. Dynamic simulation of the equipment repair support system based on the OPN model[J]. Procedia Engineering, 2011, 23: 284-289.

[27] ZHANG W J, KANG R, GUO L H, et al. Study on military equipment support modeling and simulation[J]. Chinese Journal of Aeronautics, 2005, 18(2): 142-146.

[28] GRADY J O. System modeling using the DoD architecture framework[C]//System Requirements Analysis. Amsterdam: Elsevier, 2006: 275-287.

[29] 李亢, 李新明, 刘东. 基于语义元模型的装备体系结构建模方法[J]. 系统工程与电子技术, 2015, 37(11): 2503-2512.
LI K, LI X M, LIU D. Modeling method for weapon system of systems structure based on semantic meta-model[J]. Systems Engineering and Electronics, 2015, 37(11): 2503-2512(in Chinese).

[30] 尹丽丽, 寇力, 范文慧. 基于多Agent的装备保障体系分布式建模与仿真方法[J]. 系统仿真学报, 2017, 29(12): 3185-3194.
YIN L L, KOU L, FAN W H. Distributed modeling and simulation method of equipment support system based on multi agent[J]. Journal of System Simulation, 2017, 29(12): 3185-3194(in Chinese).

[31] 邢彪, 曹军海, 宋太亮. 基于多Agent仿真的装备保障体系供应保障系统设计与实现[J]. 指挥控制与仿真, 2016, 38(2): 102-105.
XING B, CAO J H, SONG T L. Design and implementation for the supply support system of equipment support system of systems based on multi-agent simulation[J]. Command Control & Simulation, 2016, 38(2): 102-105(in Chinese).

[32] 尹延涛, 徐衡博, 高杰, 等. 基于Multi-Agent的导弹装备军民一体化维修保障决策支持系统研究[J]. 质量与可靠性, 2012(4): 12-14.
YIN Y T, XU H B, GAO J, et al. Research on multi-agent-based decision support system for military-civilian integrated maintenance support of missile equipment[J]. Quality and Reliability, 2012(4): 12-14(in Chinese).

[33] 黄海, 曹军海, 单志伟. 基于MAS的装备综合保障仿真系统设计[J]. 装甲兵工程学院学报, 2008, 22(3): 17-21.
HUANG H, CAO J H, SHAN Z W. Design of the MAS-based material integrated support simulation system[J]. Journal of Academy of Armored Force Engineering, 2008, 22(3): 17-21(in Chinese).

[34] 王步云, 张国. 一种适用于人工生命作战仿真的混合Agent结构[J]. 系统仿真学报, 2010, 22(11): 2515-2518.
WANG B Y, ZHANG G. Mixed agent architecture for warfare-simulation based on artificial life[J]. Journal of System Simulation, 2010, 22(11): 2515-2518(in Chinese).

Simulation evaluation and analysis of aircraft group support based on multi-agent

DING Gang[1]，CUI Lijie[1,*]，HAN Cheng[2]，WANG Xinhe[2]，ZHANG Feng[2]

(1. School of Equipment Management and UAV Engineering，Air Force Engineering University，Xi'an 710051，China；
2. School of Mechanics，Civil Engineering and Architecture，Northwestern Polytechnical University，Xi'an 710129，China)

Abstract：Aircraft group support involves a wide range of elements, multiple strategies, and strong interaction. Its simulation analysis method is a hot and difficult point in conducting research on aircraft group support decision-making and evaluation. This paper first establishes the multi-agent function type and interaction relationship model; Secondly, defines the multi-agent structure and establish the model; Finally, Using the aircraft group consisting of five types of aircraft as the simulation object, the case validation and simulation analysis are conducted based on the mission success rate as the guarantee indicators. The results show that the mean time between failures (MTBF) and the mean time to repair (MTTR) of various types of fighters have similar effects on the mission success rate. As the MTBF increases, the mission success rate increases. As the MTTR increases, the mission success rate decreases. The proposed support simulation evaluation method can provide a feasible and effective method for the intelligent decision-making of aircraft group maintenance support, and support the realization of model-based intelligent decision-making optimization.

Keywords：aircraft group；support；multi-agent；simulation evaluation；sensitivity analysis

Received：2021-11-15；**Accepted**：2022-03-04；**Published Online**：2022-03-22 15：24

URL：kns.cnki.net/kcms/detail/11.2625.V.20220321.1807.002.html

Foundation item：Basic Scientific Research Program of Technology (212KJ43005)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0756

基于GEMD与改进PCNN的红外与可见光图像融合

杨艳春*，李小苗，党建武，王阳萍
(兰州交通大学 电子与信息工程学院，兰州 730070)

摘 要：针对传统图像融合方法因分解工具的局限性使融合图像边缘出现伪影与亮度、对比度下降的问题，提出一种基于梯度保边多层级分解(GEMD)与改进脉冲耦合神经网络(PCNN)的红外与可见光图像融合方法。利用梯度双边滤波器(GBF)与梯度滤波器(GF)构造一种多层级分解模型，将源图像分解为3层特征图与1个基础层，且每层特征图有细、粗2个结构；根据各特征图所包含信息特点，分别采用在输入刺激中引入改进拉普拉斯算子来增强对图像中弱细节信息捕捉的PCNN、区域能量和对比度显著的融合规则进行子图融合，得到各子特征融合图像与基础层融合图像；将各子融合图像进行叠加，获得最终融合图像。实验结果表明：所提方法在视觉效果方面与定量评价方面均有所提高，提高了红外与可见光融合图像的亮度与对比度信息。

关 键 词：图像融合；红外与可见光图像；脉冲耦合神经网络；双边滤波器；梯度滤波器
中图分类号：TP391
文献标志码：A **文章编号**：1001-5965(2023)09-2317-13

图像融合是把多幅图像中相关联的互补信息进行有效地提取，获得一幅信息更加丰富全面的综合图像[1]。红外图像是通过热辐射成像，能够较好地突出目标物体，在夜间及恶劣天气等弱光条件下，红外图像可以进行全天时的探测，但是分辨率较低，达不到亚像素(面阵摄像机的成像面以像素为最小单位)级别的分辨率。可见光图像是利用反射特性成像，分辨率和对比度比较高，然而容易受到天气、光照及遮挡等因素的制约。红外与可见光成像各有优势和局限，将两者获得的图像融合能够克服单一传感器的局限性实现优势互补，可以获得对场景更为精确的理解和解释。在电力设备监测、军事侦察及工业检测等实际工程中有广泛的应用[2-3]。

红外与可见光图像融合方法按照融合理论和策略差异可分为2类：基于多尺度分析方法和基于显著区域分析方法。基于多尺度分析方法先将源图像分解成不同尺度的子图像，再对各个子图像分别融合，这类方法能很好地表现图像的固有特征；基于显著区域分析方法属于特征级融合，这类方法是根据人眼视觉注意机制原理设计的，所以能在最终融合结果中保留相对完整的显著目标区域，提高融合图像的视觉效果。基于多尺度分析方法以其良好的特性在红外与可见光图像融合中得到学者们的广泛关注[4-6]，基于保边滤波的方法是多尺度变换的一种常用方法，具有边缘保持特性、运算效率高等特点。文献[7]为解决传统融合方法中存在低对比度、边缘模糊和噪声的挑战，采用一种基于模糊梯度阈值函数和全局优化的梯度滤波器(gradient filter, GF)处理图像，在去除小的渐变纹理

收稿日期：2022-09-02；**录用日期**：2023-01-02；**网络出版时间**：2023-02-06 14:56
网络出版地址：kns.cnki.net/kcms/detail//11.2625.V.20230206.1154.002.html
基金项目：长江学者和创新团队发展计划资助(IRT_16R36)；国家自然科学基金(62067006)；甘肃省科技计划项目(18JR3RA104)；甘肃省高等学校产业支撑计划项目(2020C-19)；兰州市科技计划项目(2019-4-49)；甘肃省教育厅青年博士基金(2022QB-067)；甘肃省自然科学基金(23JRRA847, 21JR7RA300)；兰州交通大学天佑创新团队(TY202003)；兰州交通大学—天津大学联合创新基金(2021052)
* **通信作者**. E-mail：yangyanchun102@ sina.com

引用格式：杨艳春，李小苗，党建武，等. 基于GEMD与改进PCNN的红外与可见光图像融合[J]. 北京航空航天大学学报，2023，49(9)：2317-2329. YANG Y C，LI X M，DANG J W，et al. Infrared and visible image fusion based on GEMD and improved PCNN[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2317-2329 (in Chinese).

和噪声的同时保持图像整体亮度和边缘梯度；文献[8]提出一种基于动态范围压缩增强和非下采样剪切波变换的融合方法，用动态范围压缩增强方法增强弱可见光图像，利用非下采样剪切波变换提取红外与可见光图像的低频和高频系数，接着对高频系数实施硬阈值收缩，抑制高频中的噪声，可以有效保留源图像的边缘特征和纹理细节，显著提高融合图像的清晰度和对比度；文献[9]利用高斯和双边滤波器混合分解源图像，通过在可见光图像中直接引入分解后的多尺度红外图像特征，增强可见光图像的背景场景和细节，融合过程中根据不同的信息尺度水平自适应地选用不同的融合策略，来实现更好的融合结果；文献[10]针对图像融合方法在性能、时间和鲁棒性之间的权衡问题提出了一种基于联合双边滤波器和局部梯度能量算子的方法，利用联合双边滤波器的保边性将图像分解为捕获丰富梯度的结构层和包含能量信息的强度层，并利用结构张量和局部梯度能量算子来融合，将源图像更多的信息保存到融合图像中。这些图像融合方法取得了较好的图像融合效果，但这类方法在多尺度分解过程中面临着如何平衡边缘保留、纹理细节信息平滑、亮度对比度变化捕捉的问题，如高斯滤波、均值滤波对图像的细节纹理结构能较好地平滑，却不能较好地保留边缘信息，引导滤波器、双边滤波器等保边滤波器虽可以高效地保持图像中的边缘信息，却不能准确地捕获图像的亮度、对比度信息变化，而GF对图像中的亮度、对比度信息则比较敏感。

针对传统融合方法中边缘出现伪影及亮度对比度信息被削弱的问题，本文设计了一种能在平滑细节信息的同时有效保留图像边缘与整体亮度、对比度的梯度双边滤波器(gradient bilateral filter, GBF)。针对分解方法使融合图像亮度对比度下降问题，结合GBF与GF设计一种新的分解模型—梯度保边多层级分解(gradient edge-preserving multi-level decomposition, GEMD)模型。GEMD将源图像分解为3层特征图与1个基础层图，依据各特征图特性分别采用基于改进脉冲耦合神经网络(pulse coupled neural network, PCNN)的融合规则与基于区域能量的融合规则对各特征图进行融合，基础层采用基于对比度显著性的融合规则进行融合，最后将各特征融合图与基础层融合图叠加重构得到融合图像。通过实验验证，本文方法融合图像的效果更优，能在改善边缘伪影的同时有效地提高融合图像的亮度与对比度。

1 相关知识

1.1 双边滤波器

双边滤波器是一种非线性保边滤波器，对于明显的边缘结构信息能够较好地保留，对于局部小的纹理细节有一定的平滑作用[11]。双边滤波器不仅考虑了像素之间在空间几何位置上的差异，还考虑了像素点之间灰度值的差异[12]，使双边滤波器能够在有效去除图像噪声的同时保留边缘结构。源图像$\boldsymbol{I}$在像素点(x, y)处的双边滤波表示为

$$\boldsymbol{I}_{\mathrm{B}}(x,y)=\frac{\sum\limits_{(i,j)\in\boldsymbol{U}}\mathrm{e}^{-\frac{(x-i)^2+(y-j)^2}{2\sigma_{\mathrm{s}}^2}}\mathrm{e}^{-\frac{\|\boldsymbol{I}(x,y)-\boldsymbol{I}(i,j)\|^2}{2\sigma_{\mathrm{r}}^2}}\boldsymbol{I}(x,y)}{\boldsymbol{\Omega}(x,y)} \tag{1}$$

$$\boldsymbol{\Omega}(x,y)=\sum_{(i,j)\in\boldsymbol{U}}\mathrm{e}^{-\frac{(x-i)^2+(y-j)^2}{2\sigma_{\mathrm{s}}^2}}\mathrm{e}^{-\frac{\|\boldsymbol{I}(x,y)-\boldsymbol{I}(i,j)\|^2}{2\sigma_{\mathrm{r}}^2}} \tag{2}$$

式中：$\boldsymbol{I}(x, y)$和$\boldsymbol{I}(i, j)$分别为(x, y)处和(i, j)处的灰度值；$\boldsymbol{I}_{\mathrm{B}}(x, y)$为双边滤波后的结果；$\boldsymbol{U}$为输入图像矩阵的滤波窗口；$\boldsymbol{\Omega}(x, y)$为$(x, y)$处的归一化因子；$\mathrm{e}^{-\frac{(x-i)^2+(y-j)^2}{2\sigma_{\mathrm{s}}^2}}$为空间距离因子；$\mathrm{e}^{-\frac{\|\boldsymbol{I}(x,y)-\boldsymbol{I}(i,j)\|^2}{2\sigma_{\mathrm{r}}^2}}$为灰度差值因子；$\sigma_{\mathrm{s}}$为控制空间滤波的强度，其值越大表示平滑的空间范围越广；σ_{r}为控制灰度域滤波强度，其值越大表示可平滑的边缘灰度跨度越大。当灰度值相差较小时，灰度差值因子趋于1，$\boldsymbol{I}(i, j)$的权重由空间距离因子决定；当灰度差值较大时，点(x, y)与点(i, j)处相关性较低，灰度差值因子趋于0，对应权重也趋于0，避免了边缘结构信息被削弱。故双边滤波器在滤除部分细节图像的同时可以很好地保留边缘信息。

将双边滤波器中灰度差因子的源图像$\boldsymbol{I}$换成了另外一个矩阵大小同$\boldsymbol{I}$一样但不同于输入的引导图像$\boldsymbol{g}$，称为联合双边滤波器。联合双边滤波器将引导图$\boldsymbol{g}$中的边缘结构信息有条件地与输入图$\boldsymbol{I}$相结合再转移到滤波结果图中[13]，但其对图像的亮度与对比度信息不敏感。

1.2 经典脉冲耦合神经网络

PCNN作为一个单层的神经网络，在图像融合中具有广泛的应用。PCNN空间信息与每个神经元高度相连，经过多次迭代，图像的空间特征信息得到了很好地保存[14]。PCNN通常包含3个部分：接收域、调制域和脉冲发生器，如图1所示，简化的PCNN表达式为

$$\begin{cases} \boldsymbol{F}_{ij}[n]=\boldsymbol{I}_{ij} \\ \boldsymbol{L}_{ij}[n]=\mathrm{e}^{(-\alpha_l)}\boldsymbol{L}_{ij}[n-1]+V_l\sum_{kl}\boldsymbol{W}_{ijkl}\boldsymbol{Y}_{kl}[n-1] \\ \boldsymbol{U}_{ij}[n]=\boldsymbol{F}_{ij}[n](1+\beta\boldsymbol{L}_{ij}[n]) \\ \boldsymbol{\theta}_{ij}[n]=\mathrm{e}^{(-\alpha_\theta)}\boldsymbol{\theta}_{ij}[n-1]+V_\theta\boldsymbol{Y}_{ij}[n-1] \\ \boldsymbol{Y}_{ij}[n]=\begin{cases}1 & \boldsymbol{U}_{ij}[n]>\boldsymbol{\theta}_{ij}[n-1]\\ 0 & \text{其他}\end{cases} \end{cases} \tag{3}$$

式中: $\boldsymbol{I}_{ij}$ 为外部输入刺激, 为图像在 (i,j) 位置像素点的灰度值; $\boldsymbol{F}_{ij}[n]$ 为 (i,j) 位置处的神经元馈送输入; $\boldsymbol{L}_{ij}[n]$ 为链接输入; V_l 为链接输入的振幅增益; $\boldsymbol{W}_{ijkl}$ 为突触连接权; β 为链接强度, $\boldsymbol{U}_{ij}[n]$ 为内部状态信号; $\boldsymbol{\theta}_{ij}[n]$ 为动态阈值; V_θ 为动态阈值的振幅增益; α_l 和 α_θ 为时间衰减常数; $\boldsymbol{Y}_{ij}[n]$ 为 PCNN 的输出; n 代表第 n 次迭代。

为方便表示, 本文中 $\boldsymbol{I}_{ij}$=$\boldsymbol{I}(i,j)$, 其他符号类似。

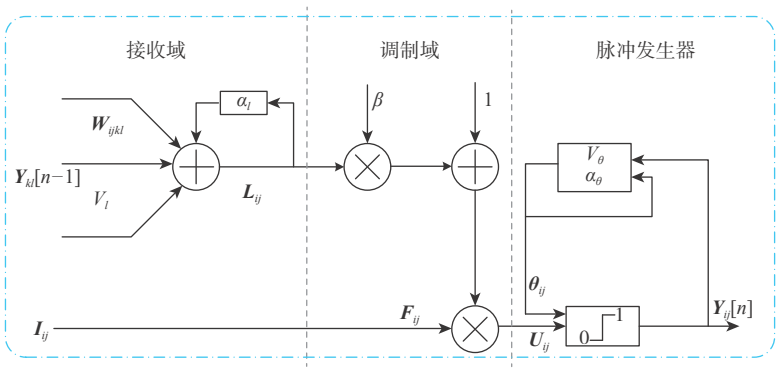


图 1　脉冲耦合神经网络模型

Fig. 1　PCNN model

2　改进模型

2.1　分解模型

2.1.1　梯度双边滤波器

针对图像平滑、边缘结构保留、亮度对比度信息削弱之间的平衡问题, 设计了 GBF。文献 [7] 从亮度和梯度分离的角度, 提出一种能在平滑图像的同时保持图像整体亮度、对比度的 GF。GBF 将双边滤波器中用来计算图像灰度差的图像改进为梯度滤波后的图像, GBF 利用 GF 在滤波过程中对亮度与对比度信息敏感性, 将捕捉的图像中亮度与对比度变化信息有条件地转移到输出结果中, 使得 GBF 滤波图像在保留边缘的同时可提高对亮度与对比度信息变化的敏感性。GBF 有效地结合了双边滤波器的保边特性与 GF 对亮度、对比度变化敏感的优势, 能在平滑掉图像纹理结构信息的同时保留图像的边缘信息, 而且能增强对图像亮度与对比度信息变化的捕捉能力。对源图像 $\boldsymbol{I}$ 梯度双边滤波的表达式为

$$\boldsymbol{I}_{\mathrm{GBF}}(x,y)=\frac{\sum_{(i,j)\in\boldsymbol{U}}\mathrm{e}^{-\frac{(x-i)^2+(y-j)^2}{2\sigma_{\mathrm{s}}^2}}\mathrm{e}^{-\frac{\|\boldsymbol{G}(x,y)-\boldsymbol{G}(i,j)\|^2}{2\sigma_{\mathrm{r}}^2}}\boldsymbol{I}(x,y)}{\boldsymbol{\Omega}(x,y)} \tag{4}$$

式中: $\boldsymbol{G}(x,y)$ 为图像 $\boldsymbol{I}$ 经过梯度滤波的图像在位置 (x,y) 处的灰度值; $\boldsymbol{G}(i,j)$ 为在位置 (i,j) 处的梯度滤波图像灰度值。此时

$$\boldsymbol{\Omega}(x,y)=\sum_{(i,j)\in\boldsymbol{U}}\mathrm{e}^{-\frac{(x-i)^2+(y-j)^2}{2\sigma_{\mathrm{s}}^2}}\mathrm{e}^{-\frac{\|\boldsymbol{G}(x,y)-\boldsymbol{G}(i,j)\|^2}{2\sigma_{\mathrm{r}}^2}} \tag{5}$$

为了方便表示, 本文将图像 $\boldsymbol{I}$ 梯度双边滤波的过程表达为

$$\boldsymbol{F}_{\mathrm{GBF}}(\boldsymbol{I})=\mathrm{GBF}(\boldsymbol{I},\boldsymbol{G}(\boldsymbol{I}),s,r) \tag{6}$$

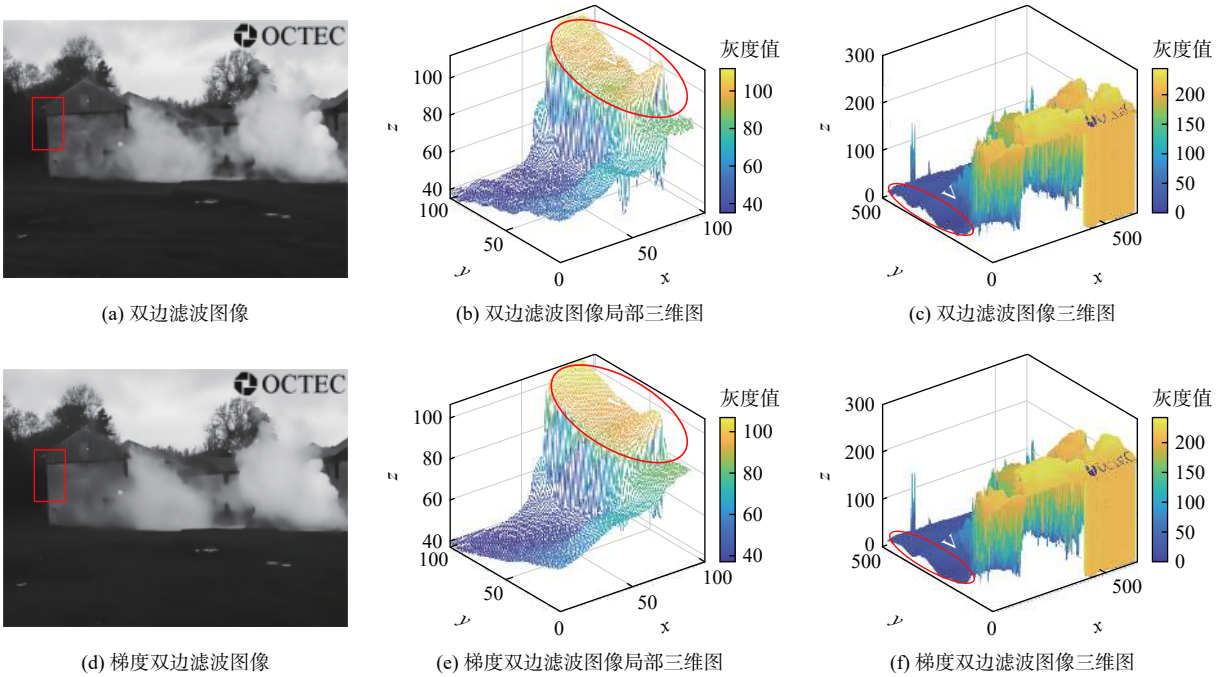
式中: s 为标准差; r 为滤波窗口半径。

为验证 GBF 的性能, 选取一幅图像进行滤波性能分析, 实验结果如图 2 所示, 图 2 (a) 和图 2 (d) 分别为双边滤波器、GBF 滤波后的图像, 对比可知, 图 2 (d) 在平滑掉图像细节纹理信息的同时可以很好地保持图像的边缘结构。图 2 (b) 和图 2 (e) 分别为图 2 (a) 和图 2 (d) 红色框图中的局部区域的三维图, 图 2 (b) 中红色椭圆区域整体呈曲面, 但上面有高低不平的突出部分, 而图 2 (e) 中在保持曲面弧度的同时平滑掉了曲面上的突出信息, 使得曲面更加平滑。图 2 (c) 和图 2 (f) 分别为图 2 (a) 和图 2 (d) 的三维网格图像, 从红色椭圆部分可以看出图 2 (c) 中的边缘毛刺多, 高低起伏, 而图 2 (f) 中的边缘在保持同样曲线弧度的同时平滑过渡。综上表明 GBF 有着更优异的平滑保边性。

2.1.2　梯度保边多层级分解

GBF 在对图像细节信息平滑的同时对边缘可有效地保持, GF 在去除小的渐变纹理和噪声的同时保持图像的整体亮度, 结合两者的特性, 提出了一种新的分解模型, 即 GEMD 模型, 如图 3 所示。

图 3 中: $\boldsymbol{I}_{\mathrm{in}}$ 为输入图像, $\boldsymbol{I}_{\mathrm{GBF}}^{i}$为第 i 次的梯度双



(a) 双边滤波图像 (b) 双边滤波图像局部三维图 (c) 双边滤波图像三维图

(d) 梯度双边滤波图像 (e) 梯度双边滤波图像局部三维图 (f) 梯度双边滤波图像三维图

图 2 梯度双边滤波器性能分析图

Fig. 2 Performance analysis diagram of GBF

边滤波图像，$\boldsymbol{I}_{\mathrm{GF}}^{i}$为第 i 次的梯度滤波图像，$\boldsymbol{l}_i$ 为第 i 层分解的细结构特征图，$\boldsymbol{c}_i$ 为第 i 层分解的粗结构特征图，$\boldsymbol{b}$ 为基础层。输入图像 $\boldsymbol{I}_{\mathrm{in}}$ 经过多层级分解后各层特征图与基础层的数学表达式为

$$\boldsymbol{l}_1 = \boldsymbol{I}_{\mathrm{in}} - \boldsymbol{I}_{\mathrm{GF}}^{1} \tag{7}$$

$$\boldsymbol{l}_i = \boldsymbol{I}_{\mathrm{GBF}}^{i-1} - \boldsymbol{I}_{\mathrm{GF}}^{i} \qquad i = 2,3 \tag{8}$$

$$\boldsymbol{c}_i = \boldsymbol{I}_{\mathrm{GF}}^{i} - \boldsymbol{I}_{\mathrm{GBF}}^{i} \qquad i = 1,2,3 \tag{9}$$

$$\boldsymbol{b} = \boldsymbol{I}_{\mathrm{GBF}}^{3} \tag{10}$$

图像 $\boldsymbol{I}_{\mathrm{in}}$ 经过 GEMD 模型分解过程的通式为

$$(\boldsymbol{l}_i, \boldsymbol{c}_i, \boldsymbol{b}) = \mathrm{GEMD}(\boldsymbol{I}_{\mathrm{in}}, \boldsymbol{I}_{\mathrm{GBF}}^{i}, \boldsymbol{I}_{\mathrm{GF}}^{i}) \qquad i = 1,2,3 \tag{11}$$

利用 GEMD 模型将源图像分解为 3 层特征图与 1 个基础层图，每层又分为细结构特征图 $\boldsymbol{l}_i$ 与粗结构特征图 $\boldsymbol{c}_i$。第 1 层的细结构特征图由源图像与梯度滤波图像作差得到，其中包含了大量的细节纹理信息，而其他剩余特征图均由 GBF 与 GF 在不同滤波次数后作差得到，其包含的信息为同一种特征的边缘结构信息，而不同层次是对上一层中未捕捉到的信息经过滤波再次优化捕捉，以达到最优的融合结果。经过实验分析，本文分解层数取 3，而基础层是 GBF 经 3 次滤波后的图像，其平滑掉了大量的纹理细节信息，保留了各源图像的边缘与亮度、对比度信息。因此，在融合分解各特征图与基础层图时，需要根据各子图中包含的特征信息选取不同的融合规则以达到更优的融合效果。图 4 对 GBF 的性能再次进行验证，图 4 (a) 和图 4 (d) 分别为双边滤波图、梯度双边滤波图与源图像的作差图，可见图 4 (d) 含有更多的细节信息，即在 GEMD 模型分解过程中可丰富各特征图中细节信息，使融合图像所含信息更多。

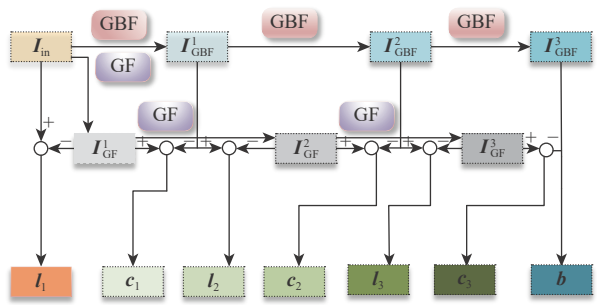


图 3 梯度保边多层级分解模型

Fig. 3 Gradient edge-preserving multi-level decomposition model

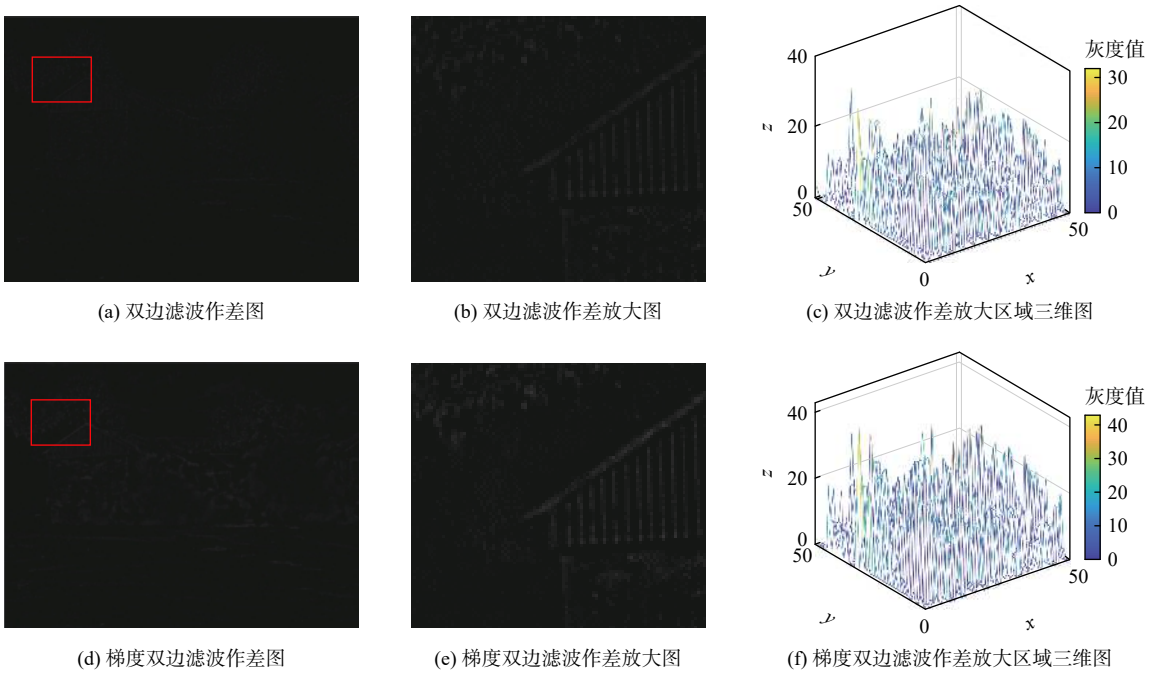
2.2 改进脉冲耦合神经网络

经典 PCNN 通常将图像中 (i,j) 位置的灰度值作为该模型的外部输入刺激，而本文在输入刺激中引入了改进的拉普拉斯算子 (sum of modified laplacian, SML)，为了适应纹理元素大小的可能变化，SML 在像素之间使用了一个可变的间距来计算，表达式为

$$\begin{aligned}\nabla^2_{\mathrm{ML}}\boldsymbol{I}(x,y) = {} & |2\boldsymbol{I}(x,y) - \boldsymbol{I}(x-\mathrm{step},y) - \\ & \boldsymbol{I}(x+\mathrm{step},y)| + |2\boldsymbol{I}(x,y) - \boldsymbol{I}(x,y-\mathrm{step}) - \\ & \boldsymbol{I}(x,y+\mathrm{step})|\end{aligned} \tag{12}$$

$$\mathrm{SML}(\boldsymbol{I}(x,y),N) = \sum_{i=x-N}^{i=x+N}\sum_{j=y-N}^{j=y+N}\nabla^2_{\mathrm{ML}}\boldsymbol{I}(x,y)$$

$$\text{for } \nabla^2_{\mathrm{ML}}\boldsymbol{I}(x,y) \geqslant T \tag{13}$$



(a) 双边滤波作差图　(b) 双边滤波作差放大图　(c) 双边滤波作差放大区域三维图

(d) 梯度双边滤波作差图　(e) 梯度双边滤波作差放大图　(f) 梯度双边滤波作差放大区域三维图

图 4　差图对比图

Fig. 4　Difference map comparison map

式中: step 为变换步长, 值为 1; N 为窗口的半径; T 为阈值, 本文中为 0。$\boldsymbol{F}_{ij}[n]$ 可改进为

$$\boldsymbol{F}_{ij}[n]=\boldsymbol{W}_{ij}^{1}\boldsymbol{I}_{ij}+\boldsymbol{W}_{ij}^{2}\mathrm{SML}(\boldsymbol{I}_{ij}) \tag{14}$$

$$\boldsymbol{W}_{ij}^{1}=0.5+0.5\left(\boldsymbol{I}_{ij}-\mathrm{SML}(\boldsymbol{I}_{ij})\right) \tag{15}$$

$$\boldsymbol{W}_{ij}^{2}=0.5+0.5\left(\mathrm{SML}(\boldsymbol{I}_{ij})-\boldsymbol{I}_{ij}\right) \tag{16}$$

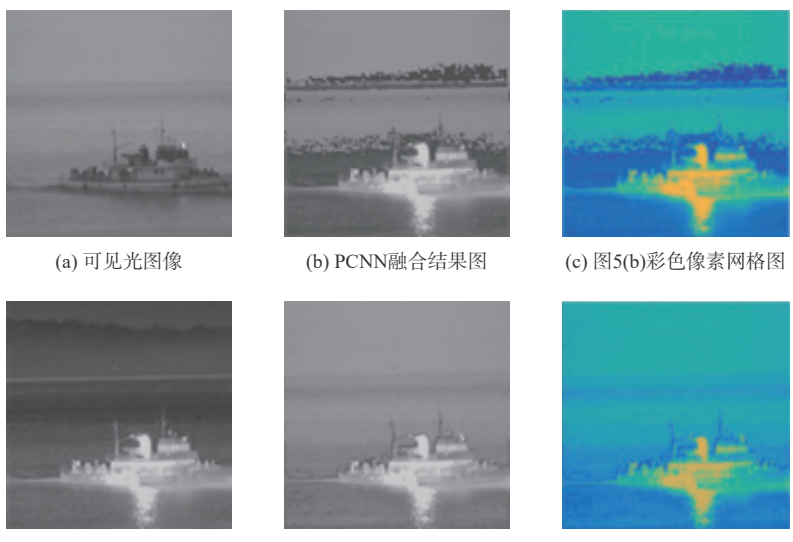
PCNN 中迭代次数 n 的设置会影响 PCNN 的脉冲同步性, 一般根据经验或多次实验来确定迭代次数 n, 本文经过多次实验取 n=200; 在改进 PCNN 中为了反映出点火的幅度差异, 引入了新的脉冲激活函数: Swish 函数, Swish 函数值变化敏感度高, 在同步脉冲激发时, 能够更好地体现出点火幅度之间特征效果的不同。采用 Swish 函数来计算每次迭代过程中子带系数的点火输出度, Swish 激活函数的表达式为

$$\boldsymbol{\gamma}_{ij}[n]=\frac{\boldsymbol{U}_{ij}[n]-\boldsymbol{\theta}_{ij}[n-1]}{1+\mathrm{e}^{-\left(\boldsymbol{U}_{ij}[n]-\boldsymbol{\theta}_{ij}[n-1]\right)}} \tag{17}$$

所以 $\boldsymbol{Y}_{ij}[n]$ 可变换为

$$\boldsymbol{Y}_{ij}[n]=\begin{cases}1 & \boldsymbol{\gamma}_{ij}[n]>0\\ 0 & \text{其他}\end{cases} \tag{18}$$

改进 PCNN 的性能分析如图 5 所示, 选取一组可见光与红外源图像如图 5 (a) 和图 5 (d)。图 5 (b)



(a) 可见光图像　(b) PCNN融合结果图　(c) 图5(b)彩色像素网格图

(d) 红外图像　(e) 改进PCNN融合结果图　(f) 图5(e)彩色像素网格图

图 5　改进 PCNN 性能分析

Fig. 5　Improved PCNN performance analysis

和图 5 (e) 分别为 PCNN 与改进 PCNN 对 2 张源图像融合的结果图。图 5 (b) 所示融合图中出现了斑驳，融合效果不够理想，而图 5 (e) 改进 PCNN 融合图像效果好，理想的融合了红外与可见光图像，通过彩色像素网格图像能更直观地表现出改进 PCNN 的性能。

3 本文方法

本文方法结构图如图 6 所示，主要分为 3 部分：

1）源图像多层级分解。将已配准的红外与可见光图像利用 GEMD 分解为 3 层特征图与 1 个基础层图，其中每层特征图包含 1 个细特征图与 1 个粗特征图。

2）子图融合。根据各层特征图与基础层成图原理的不同分别采用基于改进 PCNN、基于区域能量和基于对比度显著性的融合规则对各特征图、基础层图像融合，得到子融合图。

3）图像重构。将各子融合图像进行叠加实现重构，得到最终融合结果图。

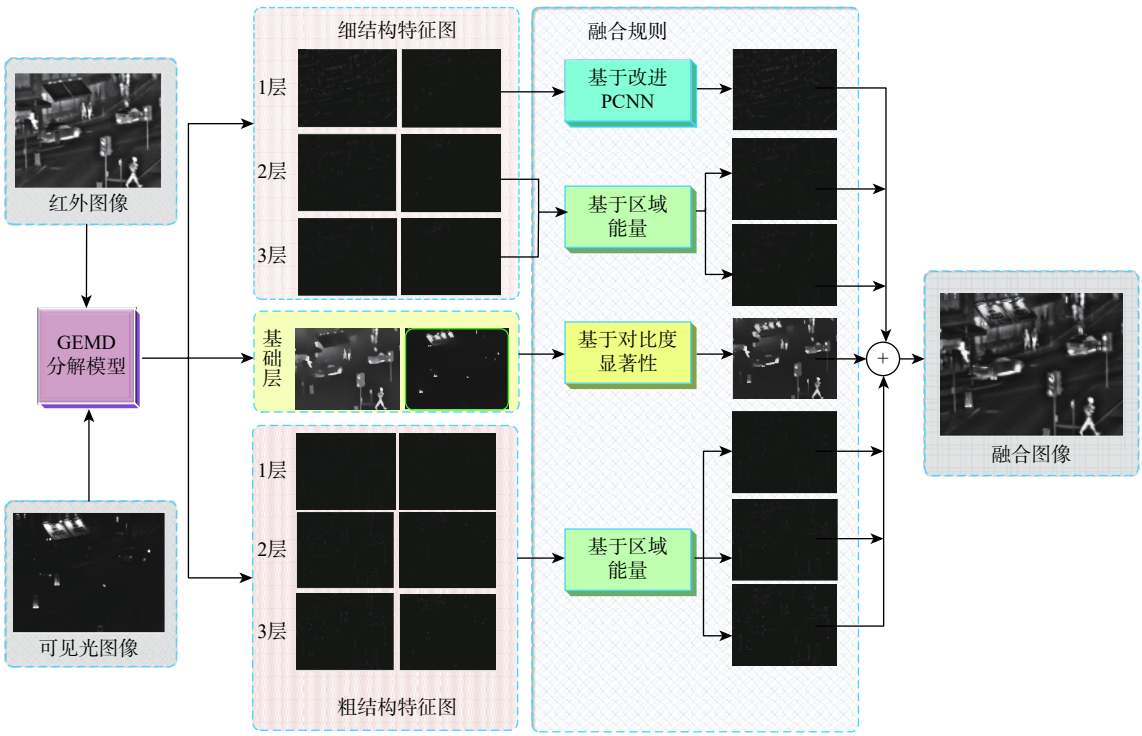


图 6 本文方法结构图

Fig. 6 Structure diagram of proposed method

3.1 基于改进 PCNN 的融合规则

红外与可见光图像经 GEMD 分解的第 1 层细结构特征图 $\boldsymbol{l}_1$ 是由源图像与梯度滤波器滤波后的图像作差得到的，该特征图中包含大量的细节纹理结构特征信息，改进 PCNN 通过 SML 增强特征图中的弱细节信息来提高融合权重图的准确性，使融合图像含有更丰富的信息，故 $\boldsymbol{l}_1$ 特征图采用基于改进 PCNN 的融合规则进行融合。将红外与可见光图像分解后的第 1 层细特征图 ($\boldsymbol{l}_{1\mathrm{I}}$、$\boldsymbol{l}_{1\mathrm{V}}$) 作为改进 PCNN 的输入刺激，根据式 (16) 求得馈送输入 $\boldsymbol{F}_{ij}[n]$，并对 $\boldsymbol{F}_{ij}[n]$ 进行归一化处理使值在 [0,1]，计算式 (3) 中各项并累加触发次数为

$$\boldsymbol{T}_{ij}[n]=\boldsymbol{T}_{ij}[n-1]+\boldsymbol{Y}_{ij}[n] \tag{19}$$

每个神经元总触发次数为 $\boldsymbol{T}_{ij}[M]$，M 为总迭代次数，改进 PCNN 红外图像的触发次数为 $\boldsymbol{T}_{\mathrm{I}}$，可见光图像的触发次数为 $\boldsymbol{T}_{\mathrm{V}}$，最后通过取最大值规则得到第 1 层细特征融合图 $\boldsymbol{f}_1$。

$$\boldsymbol{f}_1=\begin{cases}\boldsymbol{l}_{1\mathrm{I}} & \boldsymbol{T}_{\mathrm{I}}[M]\geqslant\boldsymbol{T}_{\mathrm{V}}[M]\\ \boldsymbol{l}_{1\mathrm{V}} & \text{其他}\end{cases} \tag{20}$$

3.2 基于区域能量的融合规则

GEMD 分解后剩余 5 组特征图包含源图像大量的边缘结构信息，能量是图像的显著特征，反映了图像信号的变化，利用区域能量进行融合能够体现像素之间的相关性，故这 5 组特征图使用基于区域能量的融合规则进行融合。特征图 $\boldsymbol{l}$ 的局部能量值 $\boldsymbol{E}$ 的计算式为

$$\boldsymbol{E}(i,j)=\sum_{i=1}^{w}\sum_{j=1}^{w}(\boldsymbol{l}(i+x,j+y))^2\boldsymbol{W}_{\mathrm{e}}(x,y) \tag{21}$$

式中：(i,j) 为特征图 $\boldsymbol{l}$ 的像素位置；w 为局部窗口的大小，值为 3，$\boldsymbol{W}_{\mathrm{e}}$=[1,1,1;1,1,1;1,1,1] 为 3×3 的滤波器模板，通过用 3×3 的窗口对区域能量进行滤波，来选择邻域内的最大值 $\boldsymbol{E}_{\max}(i,j)$，其表达式为

$$\boldsymbol{E}_{\max}(i,j)=\max\{\boldsymbol{E}(i+x,j+y)|1\leqslant x,y\leqslant 3\} \tag{22}$$

$$\boldsymbol{L}(i,j)=\begin{cases}1 & \boldsymbol{E}_{\max \mathrm{I}}(i,j)\boldsymbol{l}_{\mathrm{I}}>\boldsymbol{E}_{\max \mathrm{V}}(i,j)\boldsymbol{l}_{\mathrm{V}}\\ 0 & \text{其他}\end{cases} \tag{23}$$

得到剩余 5 组特征图的融合系数 $\boldsymbol{f}_k(i,j)$ 为

$$\boldsymbol{f}_k(i,j)=\boldsymbol{L}_k(i,j)\boldsymbol{l}_{k\mathrm{I}}(i,j)+(1-\boldsymbol{L}_k(i,j))\boldsymbol{l}_{k\mathrm{V}}(i,j)$$
$$k=2,3,4,5,6 \tag{24}$$

式中: $\boldsymbol{l}_{k\mathrm{I}}(i,j)$ 和 $\boldsymbol{l}_{k\mathrm{V}}(i,j)$ 分别为红外图像、可见光图像在第 k 组结构特征图在像素点 (i,j) 处的灰度值。

3.3　基于对比度显著性的融合规则

基础层反映了源图像的亮度、对比度分布, 一般不受噪声的影响。对比度反映一幅图像中最高亮度水平和最低亮度水平之间的差异, 亮度差异越明显, 对比度就越高, 因此, 像素值相对于图像平均值越亮或越暗, 其对图像对比度的贡献越大, 对比度显著性越强, 因此, 根据基础层的亮度分布构造对比度显著性图。采用 2 范数来度量像素值与平均值的偏差, 并表示每个像素的显著性。基础层图 $\boldsymbol{b}$ 的对比度显著性映射计算表达式为

$$\boldsymbol{S}_{\mathrm{SA}}=\mathrm{norm}(\|\boldsymbol{b}-\mathrm{mean}(\boldsymbol{b})\|_2) \tag{25}$$

式中: mean(*) 表示图像矩阵的平均值, 采用 2 范数消除符号的影响; norm(*) 函数拉开显著值的距离, 使显著性映射的效率越来越高。

$$\mathrm{norm}(x)=\frac{x-\min(x)}{\max(x)-\min(x)} \tag{26}$$

对基础层的显著性映射进行以下操作, 直接得到基础层的权重矩阵 $\boldsymbol{w}_1$ 和 $\boldsymbol{w}_2$ 为

$$\begin{cases}\boldsymbol{w}_1=0.5\boldsymbol{I}_{\mathrm{d}}+0.5(\boldsymbol{S}_{\mathrm{SA_I}}-\boldsymbol{S}_{\mathrm{SA_V}})\\ \boldsymbol{w}_2=0.5\boldsymbol{I}_{\mathrm{d}}+0.5(\boldsymbol{S}_{\mathrm{SA_V}}-\boldsymbol{S}_{\mathrm{SA_I}})\end{cases} \tag{27}$$

式中: $\boldsymbol{I}_{\mathrm{d}}$为单位矩阵。

只有当 2 个源图像的对比度有显著差异时, 一个接近 0, 另一个接近 1 时, 权重矩阵才会接近选择最大方法。然而, 在大多数情况下, 权重矩阵接近于平均法, 基础层的最终权值矩阵在空间平滑度上与基础层相似, 重构得基础层融合子图 $\boldsymbol{B}$ 为

$$\boldsymbol{B}=\boldsymbol{w}_1\boldsymbol{b}_{\mathrm{I}}+\boldsymbol{w}_2\boldsymbol{b}_{\mathrm{V}} \tag{28}$$

3.4　图像重构

将各子融合图像进行叠加得到最终的融合图像为

$$\boldsymbol{F}=\boldsymbol{f}_1+\boldsymbol{f}_k+\boldsymbol{B}\quad k=2,3,4,5,6 \tag{29}$$

4　实验结果与分析

4.1　参数设置

为了从视觉效果上分析本文方法的有效性, 本文在公开的 TNO 数据集 (https://figshare.com/articles/TNO Image Fusion Dataset/1008029) 上进行了测试, 均取得较好的融合效果。本文选用 6 组已配准大小分别为 640×480、256×256、256×201、256×256、253×256、512×512 的红外 (波段范围为 3～14 μm) 与可见光图像对本文方法的可行性与有效性进行验证, 实验环境为操作系统 Win10, 编程环境为 MATLAB R2016b。经多次实验分析, 当 GBF 参数 s=0.001, r=15 时效果好, 因此, 将其作为本文方法的默认设置, 改进 PCNN 的初始化条件为: $\boldsymbol{Y}_{ij}[0]=0$, $\boldsymbol{U}_{ij}[0]=0$ 和 $\boldsymbol{\theta}_{ij}[0]=0$。

4.2　参数分析

4.2.1　分解级数

GEMD 中分解级数 i 的作用旨在使融合图像中有更丰富的特征信息, 该参数对融合结果有着较为重要的作用, 分别就 i=1,2,3,4,5,6等取值评估实验。实验表明, 当 i=3 时, 本文方法融合效果最佳, i 值过小融合图像亮度对比度偏低, i 值过大时亮度信息过于丰富, 且会增加算法耗时。因此, 本文选择值 i=3 作为分解级数。实验分析过程如图 7 所示。

图 7　分解级数分析图

Fig. 7　Decomposition series analysis diagram

4.2.2 对比度显著性的权重值

式(27)中为融合基础层图像分配权重时，第1个0.5是为使权重 $\boldsymbol{w}_1$ 与 $\boldsymbol{w}_2$ 求和后值为1取的值，第2个0.5是经多次实验分析取的最优值，分析过程如图8所示。通过对比权值 k=0.1～0.9时实验融合效果，观察融合结果图像，随着权重值的增大，目标对象人物与天空亮度逐渐增强，但当权值过小时图像中红外信息不显著，人物整体轮廓模糊，当权重值过大时图像整体过亮，对比度信息过于突出，保真度低。当权重值为0.5时目标人物与建筑平滑自然，亮度对比度更符合人眼观察综合效果最优，故本文 k 取值为0.5。

k=0.1 k=0.2 k=0.3 k=0.4 k=0.5 k=0.6 k=0.7 k=0.8 k=0.9

图8 权重取值分析图

Fig. 8 Weight value analysis diagram

4.3 对比实验分析

将本文方法与GF方法[7]、滚动引导滤波器(rolling guide filter, RGF)方法[15]、卷积神经网络(convolutional neural network, CNN)方法[16]、基于卷积稀疏性的形态分量分析(convolutional sparsity based morphological component analysis ,CSMCA)方法[17]、背景重构概述(background reconstruction general, BRG)方法[18]和渐进照明感知融合(progressive illumination aware fusion, PIAF)方法[19]进行实验对比。GF基于模糊梯度阈值函数优化分解图像，提高了图像整体亮度和边缘梯度；RGF基于视觉显著性映射和加权最小二乘优化提出一种多尺度红外与可见光图像融合方法；CNN利用卷积网络获得源图像像素活动信息的权值图，并采用局部相似度融合策略对分解系数进行自适应融合；CSMCA将卷积稀疏表示和形态成分分析集成到1个统一卷积稀疏的形态成分分析模型，实现源图像的多分量和全局稀疏表示；BRG将提取的红外明亮特征保存到可见光图像中，保存了大量原始可见光图像信息；PIAF利用光照概率构建光照感知损失来指导融合网络的训练，自适应地保持显著目标的强度分布并保留背景中的纹理信息。选用的实验源图像如图9所示。

通过评价指标对融合方法的性能来进行评判，现阶段评价指标主要分为2类：主观评价和客观评价[20]。主观评价是指人眼对融合结果图的直观感

(a1) img1 (a2) img2 (a3) img3 (a4) img4 (a5) img5 (a6) img6

(a) 6组可见光源图像

(b1) img1 (b2) img2 (b3) img3 (b4) img4 (b5) img5 (b6) img6

(b) 6组红外源图像

图9 实验源图像

Fig. 9 Experimental source images

受, 有好、一般、差; 客观评价指标是指进行量化分析, 对融合图像进行计算或计算融合图像与源图像之间的关系来得到客观数值, 最终达到对融合方法性能优劣的评价, 本文选用平均梯度 (average gradient, AG)、信息熵 (information entropy, IE)、图像清晰度 (image definition, ID)、标准差 (standard deviation, SD)、视觉信息保真度 (visual information fidelity, VIFF)、边缘信息保持度 (edge information retention, EIR)6 种被广泛使用的客观评价指标对本文方法的有效性进行评价。AG 衡量图像中纹理变化和细节反差; IE 度量融合后的图像所含信息量; ID 衡量融合图像的清晰程度; SD 度量融合图像的对比度情况; VIFF 通过利用信息论的知识, 计算失真图像与源图像之间的交互信息, 从而对失真图像的质量进行评价, 该评价指标取得的结果与人眼的视觉效果较一致, 并且精确性高; EIR 计算从源图像转移到融合图像中的边缘结构信息量; 6 个评价指标值均越大表明融合结果图像效果越优。

4.3.1　主观评价分析

本文方法与各对比方法的融合结果如图 10 所示。第 1 组中(img1), BRG 融合图像边缘出现明显的伪影, GF、RGF、PIAF 中背部树枝的轮廓模糊, 而本文方法中整个图像对比度比其他 6 种方法的融合结果图高, 同时对图像的亮度信息有更高效地保留, 如放大图中人型轮廓; 第 2 组中(img2), GF、RGF、CNN、CSMCA、BRG、PIAF 融合结果图中背部的树枝条部分边缘模糊、图像的清晰度不高, 而本文方法融合中树枝部分的边界明显清晰度有明显地提高, 同时对细节纹理信息部分有更多地保留, 如红框中的区域, 在可见光源图像中可以清楚地看见树叶的边缘轮廓, 而各对比方法中的枝干均变得模糊, 本文方法中边缘更加清晰且亮度保真度高; 第 3 组中(img3), GF 融合结果马路上的纹理信息被过度平滑, RGF 中右上角部分的道路边缘部分信息缺失导致呈现出边缘弯曲, CNN、BRG、PIAF 融合结果图对比度都比较低, CSMCA 中红绿灯中第 2 颗灯的亮度偏低不突出, 而本文方法图像整体清晰较高, 将红外图像中的亮度特征很好地转移到了融合结果中, 同时对图像的边缘信息有着高效地保持, 如红框左边的路人信息, 其他 6 种对比方法中 2 人之

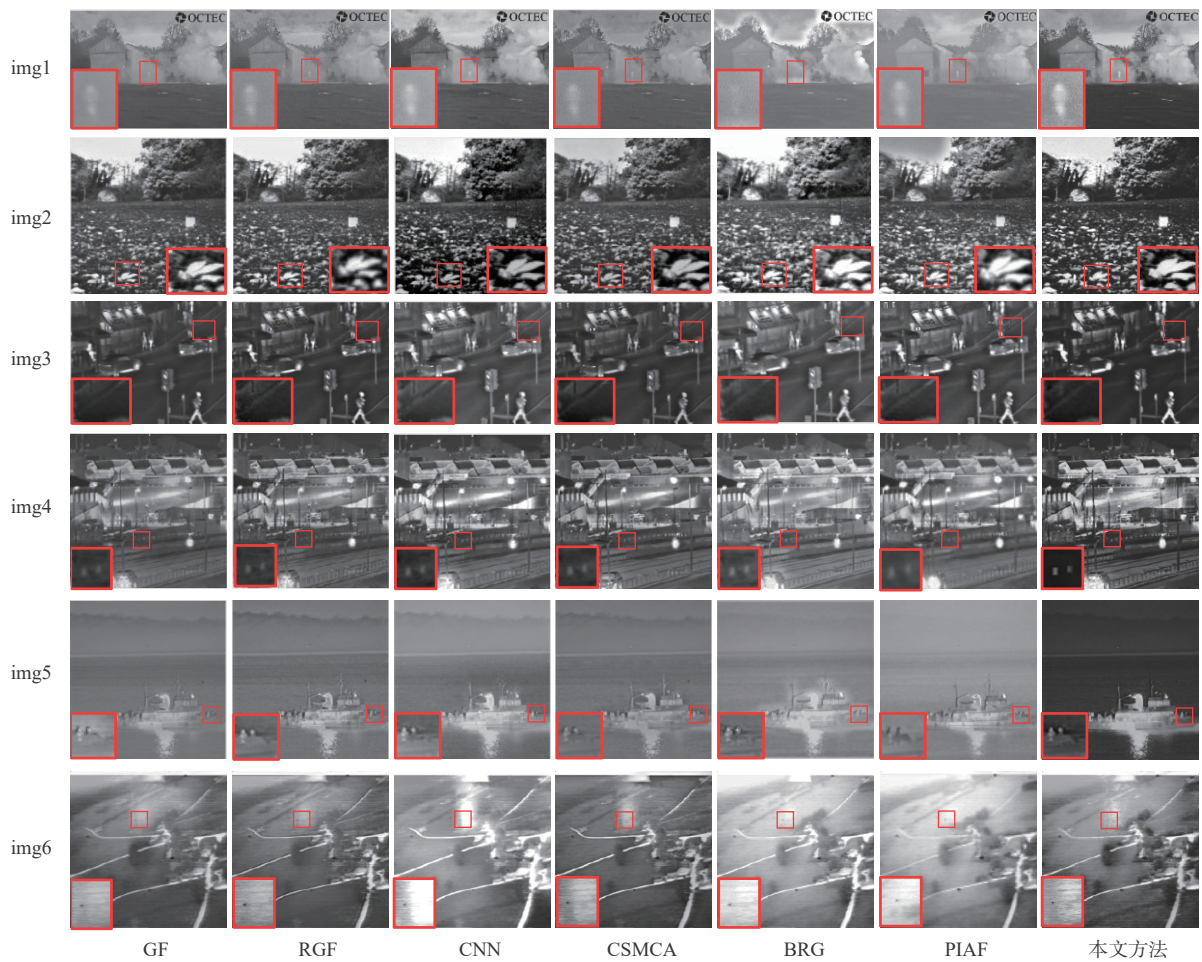


图 10　实验结果

Fig. 10　Experimental results

间的边界轮廓模糊导致难以辨认出形状，本文方法中2人之间轮廓边界明显；第4组中(img4)，GF、RGF、CNN、CSMCA、BRG融合结果图中道路边的栏杆模糊不清，出现了细节信息缺失，GF、RGF中图像上部房屋边界模糊，对可见光图像中的可视信息缺失，PIAF融合结果图清晰度低，栏杆与房屋的边缘存在不同程度的伪影，而本文方法中可以清晰的辨认出栏杆；第5组中(img5)，GF、RGF、CNN、BRG中轮船的细节信息几乎观察不到，如红色框中轮船上的人物轮廓，在各个对比方法融合结果模糊不清，难以辨认，PIAF中人物亮度低，而本文方法中可以明显地观察到人物信息；第6组中(img6)，GF、RGF、CNN、CSMCA中的右上角树木部分均出现不同程度的缺失，GF、RGF中最下方道路旁的树木被过度平滑掉，PIAF融合结果图中各道路亮度均偏低，融合效果不够理想，而本文方法中的树木轮廓清晰分明。综上可知，本文方法可以更高效地保留源图像的边缘轮廓信息与红外图像中的亮度特征，得到效果更优的融合结果图。

4.3.2 客观评价分析

选取AG、IE、ID、SD、MIPP、EIR这6个常见的评价指标进行定量分析，各指标值越大表示融合图像更优[21]。表1为实验图像各评价指标值，最优结果加粗显示，可知6组图像中每一组中的AG、

表1 实验图像各评价指标值

Table 1 The evaluation index values of experimental images

图像编号	方法	AG	IE	ID	SD	VIFF	EIR/10^9
	GF	2.969 3	6.726 1	4.456 8	40.306 9	0.540 5	6.706 5
	RGF	4.243 6	6.733 3	6.215 2	35.730 4	0.554 7	7.705 0
	CNN	4.007 7	7.014 1	6.219 9	52.012 7	0.681 0	10.010 0
img1	CSMCA	3.487 3	6.455 9	5.134 8	31.781 5	0.524 6	7.781 5
	BRG	3.818 3	5.844 4	6.029 2	43.128 9	0.589 7	1.396 6
	PIAF	2.860 6	6.532 2	2.436 5	31.996 8	0.354 6	7.206 5
	本文方法	**5.311 7**	**7.266 2**	**8.998 6**	**59.429 0**	**0.771 1**	**11.347 0**
	GF	11.744 9	7.342 0	13.692 8	59.628 1	0.703 8	7.809 4
	RGF	14.308 4	7.478 2	17.091 3	62.938 8	0.768 7	10.007 0
	CNN	14.020 7	7.370 8	16.357 7	68.026 7	0.809 3	10.163 0
img2	CSMCA	14.143 2	**7.536 2**	16.396 4	57.280 0	0.777 7	10.006 0
	BRG	15.623 1	7.435 5	18.167 2	65.564 7	0.808 2	12.375 0
	PIAF	12.648 5	7.254 6	16.621 8	60.298 4	0.754 8	11.657 0
	本文方法	**17.720 0**	7.467 5	**21.022 4**	**75.696 1**	**0.905 7**	**16.571 0**
	GF	5.006 5	6.210 6	5.630 6	31.645 2	0.473 1	2.130 2
	RGF	6.537 1	6.498 5	7.559 0	34.597 7	0.551 6	2.966 8
	CNN	6.324 0	6.692 5	7.230 9	37.386 8	0.614 3	2.820 5
img3	CSMCA	5.906 1	6.527 1	6.621 8	32.846 7	0.539 9	2.583 6
	BRG	6.311 2	6.715 0	6.963 9	38.289 2	0.624 1	2.748 4
	PIAF	7.100 7	**6.778 1**	8.032 0	38.376 9	0.647 1	3.483 7
	本文方法	**7.560 5**	6.616 1	**8.452 9**	**46.211 9**	**0.711 2**	**4.344 6**
	GF	6.567 4	7.138 8	7.713 2	41.808 9	0.464 7	3.089 3
	RGF	8.562 0	7.193 2	10.403 9	43.818 2	0.519 8	4.648 8
	CNN	8.744 5	7.420 3	10.297 7	51.704 3	0.627 5	4.871 2
img4	CSMCA	8.994 9	7.411 8	10.416 5	46.175 2	0.451 2	5.317 1
	BRG	8.785 1	7.432 1	10.367 6	48.122 2	0.513 5	5.034 9
	PIAF	8.433 9	7.305 6	9.700 2	49.038 8	0.586 0	5.045 7
	本文方法	**11.467 6**	**7.632 2**	**13.993 3**	**63.170 5**	**0.645 0**	**7.944 6**
	GF	1.816 8	5.331 7	2.280 2	15.154 6	0.235 8	1.031 2
	RGF	2.848 8	5.474 0	3.547 6	17.434 0	0.326 1	1.051 3
	CNN	2.559 2	**6.355 5**	3.236 3	26.934 6	0.265 2	1.119 3
img5	CSMCA	2.224 6	5.298 3	2.736 0	14.506 9	0.150 7	1.247 4
	BRG	2.800 4	5.871 2	3.301 3	19.904 9	0.115 7	**1.356 2**
	PIAF	1.925 3	5.723 7	2.344 0	16.537 9	0.149 7	1.037 1
	本文方法	**3.303 7**	5.716 3	**4.177 9**	**35.782 9**	**0.343 8**	1.165 5
	GF	3.882 8	7.187 9	5.463 6	45.742 8	0.464 0	6.209 7
	RGF	5.507 1	7.226 4	8.045 8	45.546 7	0.500 0	10.246 0
	CNN	4.831 8	7.115 2	6.862 5	51.000 2	0.537 8	9.130 1
img6	CSMCA	5.257 8	7.216 7	7.055 0	46.525 0	**0.623 3**	**11.943 0**
	BRG	3.993 6	7.422 1	5.674 7	50.337 4	0.392 5	7.132 2
	PIAF	3.777 0	7.591 7	5.319 8	52.362 0	0.246 1	5.767 8
	本文方法	**5.509 7**	**7.633 9**	**8.202 7**	**54.625 6**	0.484 6	9.510 2

注：加粗数据表示最优结果

ID、SD 评价指标均为本文方法的最高, 表明本文方法在梯度信息保留与提高清晰度方面有着一定的优势, 各指标中 VIFF 在 6 组图像中有 5 组处于最优, EIR 在 6 组中 4 组处于最优, IE3 组处于最优, 综合所述, 本文方法优于绝大部分的对比方法。

图 11 为本文方法与各对比方法的各指标折线图, 可以更直观地看出各指标的比较情况, 本文方法的 AG、ID、SD、VIFF、EIR 均明显高于其他对比方法, 信息熵中大部分也优于其他对比方法, 说明本文方法融合结果图细节纹理保留效果更好、图像更清晰、对比度信息更高、更符合人眼视觉效果、保边性更优。综合主观与客观评价指标, 本文方法在性能上优于其他对比方法, 可得到更优的融合结果图像。

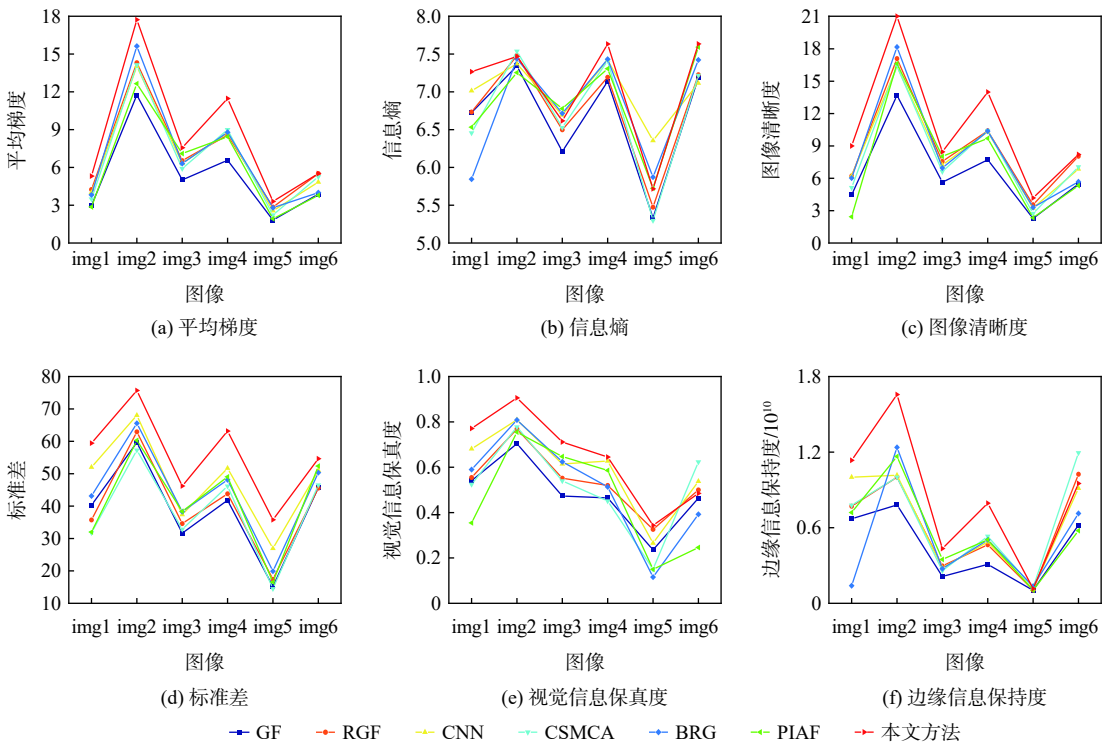


图 11　各指标折线图

Fig. 11　Line chart of each indicator

5　结　论

1) 提出的 GBF 可在平滑细节信息的同时有效保留边缘亮度对比度信息, 提高分解模型中特征图中的信息。

2) 提出的 GEMD 模型可较好地分离出各源图像中的细节纹理信息、边缘结构信息、亮度对比度信息。不同融合规则的选择使得融合图像保留了大量源图像有效信息, 提高了融合图像的亮度与对比度, 客观评价指标上也有所提高。

3) 改进 PCNN 可增强对图像中弱细节信息的捕捉敏感性, 提高权重图的准确度, 使融合图像保留更多源图像的有效信息。

本文对参数处理较多后续会再对其进行优化提高方法效率。

参考文献 (References)

[1] 沈瑜, 陈小朋. 基于DLatLRR与VGG Net的红外与可见光图像融合[J]. 北京航空航天大学学报, 2021, 47(6): 1105-1114.
SHEN Y, CHEN X P. Infrared and visible image fusion based on latent low-rank representation decomposition and VGG Net[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(6): 1105-1114(in Chinese).

[2] LI G F, LIN Y J, QU X D. An infrared and visible image fusion method based on multi-scale transformation and norm optimization[J]. Information Fusion, 2021, 71: 109-129.

[3] 朱雯青, 汤心溢, 张瑞, 等. 基于边缘保持和注意力生成对抗网络的红外与可见光图像融合[J]. 红外与毫米波学报, 2021, 40(5): 696-708.
ZHU W Q, TANG X Y, ZHANG R, et al. Infrared and visible image fusion based on edge-preserving and attention generative adversarial network[J]. Journal of Infrared and Millimeter Waves, 2021, 40(5): 696-708(in Chinese).

[4] 刘佳妮, 金伟其, 李力, 等. 自适应参考图像的可见光与热红外彩色图像融合算法[J]. 光谱学与光谱分析, 2016, 36(12): 3907-3914.
LIU J N, JIN W Q, LI L, et al. Visible and infrared thermal image fusion algorithm based on self-adaptive reference image[J]. Spectroscopy and Spectral Analysis, 2016, 36(12): 3907-3914(in Chinese).

[5] LUO Y Y, HE K J, XU D, et al. Infrared and visible image fusion

based on visibility enhancement and hybrid multiscale decomposition[J]. Optik, 2022, 258: 168914.

[6] GUO Z Y, YU X T, DU Q L. Infrared and visible image fusion based on saliency and fast guided filtering[J]. Infrared Physics & Technology, 2022, 123: 104178.

[7] MA J Y, ZHOU Y. Infrared and visible image fusion via gradientlet filter[J]. Computer Vision and Image Understanding, 2020, 197-198: 103016.

[8] 王满利, 王晓龙, 张长森. 基于动态范围压缩增强和NSST的红外与可见光图像融合算法[J]. 光子学报, 2022, 51(9): 277-291.
WANG M L, WANG X L, ZHANG C S. Infrared and visible image fusion algorithm based on dynamic range compression enhancement and NSST[J]. Acta Photonica Sinica, 2022, 51(9): 277-291(in Chinese).

[9] ZHOU Z Q, WANG B, LI S, et al. Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters[J]. Information Fusion, 2016, 30: 15-26.

[10] LI X S, ZHOU F Q, TAN H S, et al. Multimodal medical image fusion based on joint bilateral filter and local gradient energy[J]. Information Sciences, 2021, 569: 302-325.

[11] HU J W, LI S T. The multiscale directional bilateral filter and its application to multisensor image fusion[J]. Information Fusion, 2012, 13(3): 196-206.

[12] 周志强, 汪渤, 李立广, 等. 基于双边与高斯滤波混合分解的图像融合方法[J]. 系统工程与电子技术, 2016, 38(1): 8-13.
ZHOU Z Q, WANG B, LI L G, et al. Image fusion based on a hybrid decomposition via bilateral and Gaussian filters[J]. Systems Engineering and Electronics, 2016, 38(1): 8-13(in Chinese).

[13] 刘峰, 沈同圣, 马新星. 交叉双边滤波和视觉权重信息的图像融合[J]. 仪器仪表学报, 2017, 38(4): 1005-1013.
LIU F, SHEN T S, MA X X. Image fusion via cross bilateral filter and visual weight information[J]. Chinese Journal of Scientific Instrument, 2017, 38(4): 1005-1013(in Chinese).

[14] 戴进墩, 刘亚东, 毛先胤, 等. 基于FDST和双通道PCNN的红外与可见光图像融合[J]. 红外与激光工程, 2019, 48(2): 67-74.
DAI J D, LIU Y D, MAO X Y, et al. Infrared and visible image fusion based on FDST and dual-channel PCNN[J]. Infrared and Laser Engineering, 2019, 48(2): 67-74(in Chinese).

[15] MA J L, ZHOU Z Q, WANG B, et al. Infrared and visible image fusion based on visual saliency map and weighted least square optimization[J]. Infrared Physics & Technology, 2017, 82: 8-17.

[16] LIU Y, CHEN X, CHENG J A, et al. Infrared and visible image fusion with convolutional neural networks[J]. International Journal of Wavelets, Multiresolution and Information Processing, 2018, 16(3): 1850018.

[17] LIU Y, CHEN X, WARD R K, et al. Medical image fusion via convolutional sparsity based morphological component analysis[J]. IEEE Signal Processing Letters, 2019, 26(3): 485-489.

[18] ZHANG Y, ZHANG L J, BAI X Z, et al. Infrared and visual image fusion through infrared feature extraction and visual information preservation[J]. Infrared Physics & Technology, 2017, 83: 227-237.

[19] TANG L F, YUAN J T, ZHANG H, et al. PIAFusion: A progressive infrared and visible image fusion network based on illumination aware[J]. Information Fusion, 2022, 83-84: 79-92.

[20] 张小利, 李雄飞, 李军. 融合图像质量评价指标的相关性分析及性能评估[J]. 自动化学报, 2014, 40(2): 306-315.
ZHANG X L, LI X F, LI J. Validation and correlation analysis of metrics for evaluating performance of image fusion[J]. Acta Automatica Sinica, 2014, 40(2): 306-315(in Chinese).

[21] 王晓文, 赵宗贵, 汤磊. 一种新的红外与可见光图像融合评价方法[J]. 系统工程与电子技术, 2012, 34(5): 871-875.
WANG X W, ZHAO Z G, TANG L. Novel quality metric for infrared and visible image fusion[J]. Systems Engineering and Electronics, 2012, 34(5): 871-875(in Chinese).

Infrared and visible image fusion based on GEMD and improved PCNN

YANG Yanchun*, LI Xiaomiao, DANG Jianwu, WANG Yangping

(School of Electronic and Information Engineering, Lan zhou Jiaotong University, Lan zhou 730070, China)

Abstract: Because of the limitations of decomposition tools in traditional image fusion methods, artifacts, decreased brightness, and contrast appear on fused image edges. A gradient edge-preserving multi-level decomposition(GEMD)-based approach for infrared and visible image fusion, as well as an enhanced pulse-coupled neural network(PCNN), are described. The gradient bilateral filter(GBF) is proposed based on the bilateral filter and the gradient filter(GF), which can preserve the edge structure, brightness, and contrast information while smoothing the detail information. Firstly, the source images are divided into three layers of feature maps and a base layer, and then a multi-level decomposition model is built using a gradient bilateral filter and gradient filter. Each layer of feature maps has two different structures, thin and thick. Then, according to the characteristics of the information contained in each feature map, the PCNN, which introduces an improved Laplacian operator in the input stimulus to enhance weak details captured in the image, regional energy, and contrast saliency are respectively adopted to obtain the fusion images of each sub-feature and the fusion image of the base layer as the fusion rules. Finally, the sub-fusion images are superimposed to obtain the final fusion image. Through experimental verification, the proposed method has improved both visual effect and quantitative evaluation and improved the brightness and contrast information of infrared and visible fusion images.

Keywords: image fusion; infrared and visible image; pulse coupled neural network; bilateral filter; gradient filter

Received: 2022-09-02; **Accepted**: 2023-01-02; **Published Online**: 2023-02-06 14:56

URL: kns.cnki.net/kcms/detail//11.2625.V.20230206.1154.002.html

Foundation items: Program for Changjiang Scholars and Innovative ResearchTeam (IRT_16R36); National Natural Science Foundation of China (62067006); Gansu Provincial Science and Technology Plan Project (18JR3RA104); Gansu Province Higher Education Industry Support Program Project (2020C-19); Lanzhou Science and Technology Plan Project (2019-4-49); Gansu Provincial Department of Education: Youth Doctoral Fund Project (2022QB-067); Natural Science Foundation of Gansu Province (23JRRA847, 21JR7RA300); Tianyou Innovation Team of Lanzhou Jiaotong University (TY202003); Lanzhou Jiaotong University-Tianjin University Joint Innovation Fund Project (2021052)

* **Corresponding author**. E-mail: yangyanchun102@ sina.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0651

推进剂储罐裂纹缺陷非概率可靠性分析方法

辛腾达[1,*]，崔村燕[1]，刘阳[1]，同江[1]，段永胜[2]

(1. 北京跟踪与通信技术研究所，北京 100094；2. 航天工程大学 宇航科学与技术系，北京 101416)

摘 要：在推进剂储罐服役期内，准确地对裂纹缺陷进行分析，掌握储罐存在裂纹缺陷情况下的可靠状态，既可保证发射场的安全，亦可有效避免不必要的恐慌，为应急预案的制定提供参考。基于区间理论与失效评定图理论，提出一种非概率失效评定图（NFAD）可靠性分析方法。有效解决了工程实际中难以准确获得失效评定点与失效评定曲线情况下，推进剂储罐裂纹缺陷的可靠性分析问题。结合实例参数对所提方法进行验证，结果表明：无需精确失效评定图与失效评定曲线，所提方法可对储罐裂纹缺陷的任意状态进行分析，并可以充分考虑分析中的不确定性，将传统失效评定图方法失效或可靠的二元逻辑状态细化为3种情况，可靠性指标为0表示失效状态，可靠性指标大于0小于1表示可靠度，可靠性指标大于等于1表示安全裕度。

关 键 词：推进剂储罐；裂纹缺陷；非概率可靠性；失效评定图；可靠性指标

中图分类号：V555.+1

文献标志码：A **文章编号**：1001-5965(2023)09-2330-07

在日常定期对推进剂储罐检查时，掌握储罐存在裂纹缺陷情况下的可靠状态，既可保证发射场的试验安全，亦可有效避免不必要的恐慌，为应急预案的制定提供参考。1999年，由9个国家的17个组织共同编著完成的“欧洲工业结构完整性评定方法”，标志着失效评定图理论的正式确立[1]。目前，在英国R6[2]、英国BS 7910[3]、美国API 579-1[4]及中国GB/T 19624-2019[5]等结构可靠性评估规范和标准中，失效评定图已成为各国工业部门对含缺陷结构进行评估分析的重要方法。

文献[6]以离散元法为失效评估图量化基础，结合J积分，对弱砂岩水力裂缝扩展的非弹性程度进行评价。文献[7]利用失效评定图理论对钢壳圆柱结构常见的缺陷类型进行评估，并确定每种缺陷的可接受程度及临界尺寸。文献[8]采用失效评定图理论对大型直管及各种管径和裂纹尺寸弯头韧性断裂的起裂载荷进行评定。文献[9]结合扩展有限元和失效评定图法，对含裂纹缺陷的液化石油气瓶进行失效分析。文献[10]应用失效评定图法，分别对含轴向裂纹和环向裂纹的高压管道系统的安全评估问题进行研究。文献[11]结合克里格回归、支持向量机分类及失效评定图理论，提出一种结构损伤容限失效分析的方法。为失效评定图理论的实际应用发展提供了重要参考。

随着失效评定图理论研究的深入，结合材料特性与载荷条件对评定区域进行划分的区失效评定图[12]，以参考应变作为评估图横坐标的应变失效评定图[13]，基于循环内聚力模型的循环失效评定图[14]，结合工程应用实际引入概率参数的概率失效评定图[15]，基于缺口应力强度因子的缺口失效评定图[16]及考虑塑性约束效应的约束失效评定图[17]等方法，进一步丰富了失效评定图理论的研究内容。

然而，在传统失效评定图理论应用中，失效评定点与失效评定曲线的确定需要大量试验数据为

收稿日期：2021-10-31；**录用日期**：2022-01-27；**网络出版时间**：2022-03-08 10:40
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220307.1439.006.html
基金项目：发射综合安全评估与试验验证技术 (ZRHT2020T06)
*通信作者. E-mail：xtd0701@126.com

引用格式：辛腾达，崔村燕，刘阳，等. 推进剂储罐裂纹缺陷非概率可靠性分析方法[J]. 北京航空航天大学学报，2023，49(9)：2330-2336.
XIN T D，CUI C Y，LIU Y，et al. Non-probabilistic reliability analysis method for propellent tank with crack defect[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2330-2336 (in Chinese).

基础，而在工程实际中难以获得其精确值[18]。因此，为有效解决参数不确定情况下结构裂纹缺陷可靠性分析问题，本文基于区间理论与失效评定图理论，提出一种裂纹缺陷非概率失效评定图（non-probability failure assessment diagram, NFAD）可靠性分析方法。

1　非概率失效评定图模型建立

在工程实际中，传统失效评定图理论将结构及缺陷参数均作为确定值，分析结果只存在 2 种情况，即可靠与失效，在失效评定点与失效评定曲线难以准确确定的情况下，可能会得到完全相反的分析结果。因此，本文基于区间理论与失效评定图理论，充分考虑分析存在的不确定性，建立 NFAD 模型，无需精确的失效评定点与失效评定曲线，将传统失效评定图理论的二值状态转换为可靠、非完全可靠和失效三值状态，如图 1 所示。

$f_1(L_r)$、$f_2(L_r)$和$f_3(L_r)$分别表示非精确失效评定曲线，根据失效评定点落于非概率失效评定图不同区域时的状态，本文将 NFAD 划分为 3 个区域，即将横坐标L_r、纵坐标K_r及失效评定曲线围成的最小区域定义为可靠域，将全部失效评定曲线以外的区域定义为失效域，可靠域与失效域之间的区域定义为不确定域。$\underline{L}_r$与$\overline{L}_r$分别为失效评定点载荷比L_r的最小值与最大值，$\underline{K}_r$与$\overline{K}_r$分别为失效评定点断裂比K_r的最小值与最大值，本文将$(\underline{L}_r,\overline{L}_r)$与$(\underline{K}_r,\overline{K}_r)$围成的区域定义为失效评定矩。当$L_r$与$K_r$的范围确定后，无需精确的失效评定点，即可确定失效评定矩。$(\underline{L}_r,\underline{K}_r)$（$C$ 点）与$(\overline{L}_r,\overline{K}_r)$（$D$ 点）分别为失效评定矩的最小坐标点与最大坐标点，$(\underline{L}_{rp},\underline{K}_{rp})$（$E$ 点）与$(\overline{L}_{rp},\overline{K}_{rp})$（$F$ 点）分别为 CD 直线与失效评定曲线相交的最小坐标点与最大坐标点。

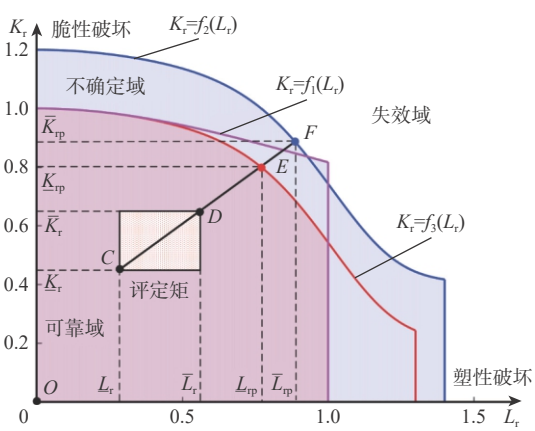


图 1　非概率失效评定图模型

Fig. 1　Non-probabilistic failure assessment diagram model

2　非概率可靠性分析方法设计

当失效评定点载荷比L_r和断裂比K_r的范围已知时，$\underline{L}_{rp}$与$\underline{K}_{rp}$的关系式为

$$\begin{cases}\underline{K}_{rp}=\dfrac{\overline{K}_r-\underline{K}_r}{\overline{L}_r-\underline{L}_r}\underline{L}_{rp}\\ \underline{K}_{rp}=f_1\left(\underline{L}_{rp}\right)\end{cases}\tag{1}$$

同理，可得$\overline{L}_{rp}$与$\overline{K}_{rp}$的关系式为

$$\begin{cases}\overline{K}_{rp}=\dfrac{\overline{K}_r-\underline{K}_r}{\overline{L}_r-\underline{L}_r}\overline{L}_{rp}\\ \overline{K}_{rp}=f_2\left(\overline{L}_{rp}\right)\end{cases}\tag{2}$$

据失效评定图理论可知，裂纹缺陷的可靠状态可由失效评定点与失效临界点坐标值间的相互关系确定。在 NFAD 模型中，失效评定点及失效临界点均非传统意义上的单一固定值，相互关系无法通过其坐标值进行直接比较。

由于存在$L_r\geqslant 0$，$K_r\geqslant 0$，$L_{rp}\geqslant 0$，$K_{rp}\geqslant 0$，且直线 CD 的斜率始终大于零，可知直线 CD 上的点将单调变化，则(L_r,K_r)和(L_{rp},K_{rp})之间仅存在 2 种关系，即

$$L_r\leqslant L_{rp}\text{且}K_r\leqslant K_{rp}\tag{3}$$

$$L_r> L_{rp}\text{且}K_r> K_{rp}\tag{4}$$

在失效评定矩确定的情况下，对于失效评定矩内任意一点(L_r,K_r)均存在：

$$\sqrt{\underline{L}_r^2+\underline{K}_r^2}\leqslant\sqrt{L_r^2+K_r^2}\leqslant\sqrt{\overline{L}_r^2+\overline{K}_r^2}\tag{5}$$

因此，本文以失效评定点到原点的距离$\sqrt{L_r^2+K_r^2}$代替失效评定点的坐标值，则仅需要确定失效评定矩内失效评定点的边界坐标值$(\underline{L}_r,\underline{K}_r)$与$(\overline{L}_r,\overline{K}_r)$，即可将失效评定矩转换为区间变量$S\left(L_r,K_r\right)$：

$$S\left(L_r,K_r\right)=\left[\sqrt{\underline{L}_r^2+\underline{K}_r^2},\sqrt{\overline{L}_r^2+\overline{K}_r^2}\right]\tag{6}$$

据式（6）可知，当$(\underline{L}_r,\underline{K}_r)$与$(\overline{L}_r,\overline{K}_r)$已知时，无需失效评定点的分布形式，即可确定$S\left(L_r,K_r\right)$区间的中值$S^{\mathrm{c}}\left(L_r,K_r\right)$与半径$S^{\mathrm{r}}\left(L_r,K_r\right)$，表达式分别为

$$S^{\mathrm{c}}\left(L_r,K_r\right)=\frac{\overline{S}\left(\overline{L}_r,\overline{K}_r\right)+\underline{S}\left(\underline{L}_r,\underline{K}_r\right)}{2}\tag{7}$$

$$S^{\mathrm{r}}\left(L_r,K_r\right)=\frac{\overline{S}\left(\overline{L}_r,\overline{K}_r\right)-\underline{S}\left(\underline{L}_r,\underline{K}_r\right)}{2}\tag{8}$$

同理，可将失效临界点转换为区间变量$R\left(L_{rp},K_{rp}\right)$：

$$R\left(L_{rp},K_{rp}\right)=\left[\sqrt{\underline{L}_{rp}^2+\underline{K}_{rp}^2},\sqrt{\overline{L}_{rp}^2+\overline{K}_{rp}^2}\right]\tag{9}$$

据式（9）可知，在$(\underline{L}_r,\underline{K}_r)$与$(\overline{L}_r,\overline{K}_r)$已知时，即可确定$R\left(L_{rp},K_{rp}\right)$区间的中值$R^{\mathrm{c}}\left(L_{rp},K_{rp}\right)$与半径$R^{\mathrm{r}}\left(L_{rp},K_{rp}\right)$，表达式分别为

$$R^{\mathrm{c}}\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)=\frac{\overline{R}\left(\overline{L}_{\mathrm{rp}},\overline{K}_{\mathrm{rp}}\right)+\underline{R}\left(\underline{L}_{\mathrm{rp}},\underline{K}_{\mathrm{rp}}\right)}{2} \tag{10}$$

$$R^{\mathrm{r}}\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)=\frac{\overline{R}\left(\overline{L}_{\mathrm{rp}},\overline{K}_{\mathrm{rp}}\right)-\underline{R}\left(\underline{L}_{\mathrm{rp}},\underline{K}_{\mathrm{rp}}\right)}{2} \tag{11}$$

由式(6)～式(11)可知，NFAD方法中的失效评定矩与失效临界点均可转换为区间变量的形式，如图2所示。

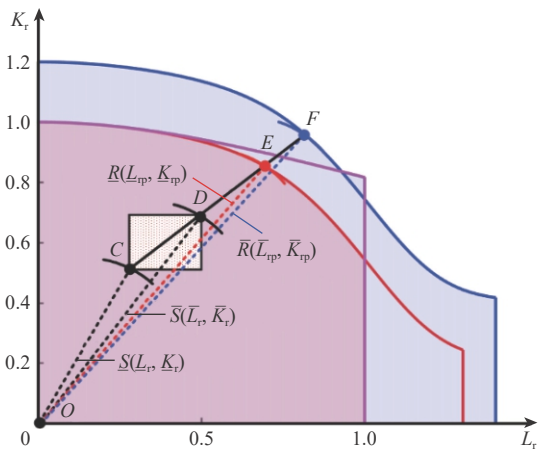


图2 非概率失效评定图区间变量转换

Fig. 2 Interval variable conversion of non-probabilistic failure assessment diagram

因此，状态函数可表示为

$$M=R\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)-S\left(L_{\mathrm{r}},K_{\mathrm{r}}\right) \tag{12}$$

式中：$M>0$表示可靠状态；$M<0$表示失效状态；$M=0$表示临界状态函数。

根据区间理论，$R\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)$与$S\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)$可表示为

$$R\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)=R^{\mathrm{c}}\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)+R^{\mathrm{r}}\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)\delta_{\mathrm{R}} \tag{13}$$

$$S\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)=S^{\mathrm{c}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)+S^{\mathrm{r}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)\delta_{\mathrm{S}} \tag{14}$$

式中：$0\leqslant\delta_{\mathrm{R}}\leqslant1$与$0\leqslant\delta_{\mathrm{S}}\leqslant1$均为标准化区间变量，则据式(13)与式(14)，即可将区间变量$R\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)$与$S\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)$转换为标准化区间变量的形式，如图3所示。

则状态函数可表示为

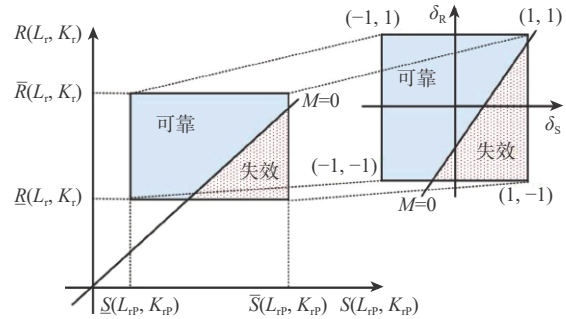


图3 标准化区间转换

Fig. 3 Conversion of normalized interval

$$M=R^{\mathrm{r}}\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)\delta_{\mathrm{R}}-S^{\mathrm{r}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)\delta_{\mathrm{S}}+R^{\mathrm{c}}\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)-S^{\mathrm{c}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right) \tag{15}$$

临界状态函数$M=0$可表示为

$$\delta_{\mathrm{R}}=\frac{S^{\mathrm{r}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)}{R^{\mathrm{r}}\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)}\delta_{\mathrm{S}}-\frac{R^{\mathrm{c}}\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)-S^{\mathrm{c}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)}{R^{\mathrm{r}}\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)} \tag{16}$$

$M=0$在标准化区间内的斜率可表示为

$$k=\frac{S^{\mathrm{r}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)}{R^{\mathrm{r}}\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)} \tag{17}$$

不同状态下，$M=0$与标准化区间的位置关系不同。

1）当$\underline{R}\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)\geqslant\overline{S}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)$时，对于标准化区间内$\forall\ (\delta_{\mathrm{S}},\delta_{\mathrm{R}})$，$\exists\ M\geqslant0$，储罐处于完全可靠状态。此时，将NFAD可靠性指标η_{d}定义为

$$\eta_{\mathrm{d}}=\frac{\sqrt{\underline{L}_{\mathrm{rp}}^2+\underline{K}_{\mathrm{rp}}^2}}{\sqrt{\overline{L}_{\mathrm{r}}^2+\overline{K}_{\mathrm{r}}^2}} \tag{18}$$

据式(18)可知，此时存在$\eta_{\mathrm{d}}\geqslant1$，表示储罐的安全裕度。

2）当$\underline{R}\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)<\overline{S}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)$且$S^{\mathrm{c}}(L_{\mathrm{r}},K_{\mathrm{r}})<R^{\mathrm{c}}\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)$时，$M=0$将标准化区间分为2部分，$M>0$为可靠域，$M<0$为失效域，但$R\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)$区间中值仍大于$S\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)$区间中值，储罐处于非完全可靠状态。

当$k>1$时，可靠性指标η_{d}可表示为

$$\eta_{\mathrm{d}}=\begin{cases}1-\dfrac{\left[\overline{S}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)-\underline{R}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right)\right]^2}{4R^{\mathrm{r}}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right)S^{\mathrm{r}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)} & \underline{R}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right)<\overline{S}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)\leqslant\overline{R}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right)\\ \dfrac{R^{\mathrm{c}}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right)-S^{\mathrm{c}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)}{S^{\mathrm{r}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)} & \overline{R}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right)<\overline{S}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right),S^{\mathrm{c}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)\leqslant R^{\mathrm{c}}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right)\end{cases} \tag{19}$$

当$k=1$时，可靠性指标η_{d}可表示为

$$\eta_{\mathrm{d}}=1-\frac{\left[\overline{S}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)-\underline{R}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right)\right]^2}{4R^{\mathrm{r}}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right)S^{\mathrm{r}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)}\qquad \underline{R}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right)<\overline{S}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right),\ S^{\mathrm{c}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)\leqslant R^{\mathrm{c}}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right) \tag{20}$$

当$0<k<1$时，可靠性指标η_{d}可表示为

$$\eta_{\mathrm{d}}=\begin{cases}1-\dfrac{\left[\overline{S}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)-\underline{R}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right)\right]^{2}}{4R^{\mathrm{r}}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right)S^{\mathrm{r}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)} & \underline{S}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)\leqslant\underline{R}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right)<\overline{S}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)\\ \dfrac{R^{\mathrm{c}}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right)-S^{\mathrm{c}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)}{S^{\mathrm{r}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)} & \underline{R}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right)<\underline{S}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right),S^{\mathrm{c}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)\leqslant R^{\mathrm{c}}\left(L_{\mathrm{rp}},R_{\mathrm{rp}}\right)\end{cases}\tag{21}$$

3）当$S^{\mathrm{c}}\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)\geqslant R^{\mathrm{c}}\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)$时，但$R\left(L_{\mathrm{rp}},K_{\mathrm{rp}}\right)$区间中值不大于$S\left(L_{\mathrm{r}},K_{\mathrm{r}}\right)$区间中值，此时将 NFAD 可靠性指标$\eta_{\mathrm{d}}$定义为

$$\eta_{\mathrm{d}}=0\tag{22}$$

3　可靠性指标验证分析

本节以含屈服平台和不含屈服平台的失效评定曲线为例，分别如式(23)与式(24)所示，结合示例参数，通过与传统失效评定图方法进行对比分析，验证 NFAD 方法可靠性指标的有效性和可行性。

$$\begin{cases}f_1\left(L_{\mathrm{r}}\right)=\left(1+0.5L_{\mathrm{r}}^{2}\right)^{-0.5}\left[0.3+0.7\mathrm{e}^{-0.65L_{\mathrm{r}}^{6}}\right]\\ L_{\mathrm{r}}^{\max}=1.25\end{cases}\tag{23}$$

$$\begin{cases}f_2\left(L_{\mathrm{r}}\right)=\left(1+0.5L_{\mathrm{r}}^{2}\right)^{-0.5}\\ L_{\mathrm{r}}^{\max}=1\end{cases}\tag{24}$$

为充分验证 NFAD 方法可靠性指标对可靠状态、非完全可靠状态及失效状态进行分析的有效性、可行性及其与传统失效评定图方法之间的区别，将$\underline{K}_{\mathrm{r}}$、$\overline{K}_{\mathrm{r}}$、$\underline{L}_{\mathrm{r}}$及$\overline{L}_{\mathrm{r}}$由小到大分别取 10 组示例参数进行分析验证，如表 1 所示。

表 1　示例参数

Table 1　Sample parameters　　MPa

序号	$\underline{K}_{\mathrm{r}}$	$\overline{K}_{\mathrm{r}}$	$\underline{L}_{\mathrm{r}}$	$\overline{L}_{\mathrm{r}}$
1	0.5	0.6	0.6	0.7
2	0.5	0.7	0.6	0.7
3	0.6	0.7	0.7	0.8
4	0.6	0.8	0.7	0.8
5	0.7	0.8	0.8	0.9
6	0.7	0.9	0.8	0.9
7	0.8	0.9	0.9	1.0
8	0.8	1.0	0.9	1.0
9	0.9	1.0	1.0	1.1
10	0.9	1.1	1.0	1.1

据表 1 所示参数，可分别求得 NFAD 方法可靠性指标及取不同失效评定图或不同失效评定曲线情况下的传统失效评定图方法可靠性指标，可靠性分析结果如表 2 所示。

表 2　可靠性分析结果

Table 2　Reliability analysis results

示例	η_{d}	状态	f_{s1}	状态	f_{s2}	状态	f_{s3}	状态	f_{s4}	状态
1	1.21	可靠	1.21	可靠	1.38	可靠	1.44	可靠	1.64	可靠
2	1.12	可靠	1.12	可靠	1.22	可靠	1.44	可靠	1.64	可靠
3	1.05	可靠	1.05	可靠	1.19	可靠	1.21	可靠	1.38	可靠
4	0.98	非完全可靠		失效	1.07	可靠	1.21	可靠	1.38	可靠
5	0.65	非完全可靠		失效	1.04	可靠	1.05	可靠	1.18	可靠
6	0.05	非完全可靠		失效		失效	1.05	可靠	1.18	可靠
7	0	失效		失效		失效		失效	1.04	可靠
8	0	失效		失效		失效		失效	1.04	可靠
9	0	失效		失效		失效		失效		失效
10	0	失效		失效		失效		失效		失效

注：f_{s1},f_{s2},f_{s3},f_{s4}分别为$\left(\overline{L}_{\mathrm{r}},\overline{K}_{\mathrm{r}}\right)$, $f_1\left(L_{\mathrm{r}}\right)$;$\left(\overline{L}_{\mathrm{r}},\overline{K}_{\mathrm{r}}\right)$,$f_2\left(L_{\mathrm{r}}\right)$;$\left(\underline{L}_{\mathrm{r}},\underline{K}_{\mathrm{r}}\right)$,$f_1\left(L_{\mathrm{r}}\right)$;$\left(\underline{L}_{\mathrm{r}},\underline{K}_{\mathrm{r}}\right)$,$f_2\left(L_{\mathrm{r}}\right)$传统失效评定图方法的可靠性指标。

如表 2 示例 1～示例 3 所示，NFAD 方法的分析结果与最为保守的失效评定图方法分析结果一致，结构处于可靠状态。安全系数[19]$f_{\mathrm{s}}=OB/OA\geqslant1$与可靠性指标$\eta_{\mathrm{d}}=OE/OA\geqslant1$之间的关系，如图 4 所示。

据图 4 可知，当失效评定矩完全落于可靠域时，存在$\eta_{\mathrm{d}}>1$及$f_{\mathrm{s}}>1$，当$\left(\overline{L}_{\mathrm{r}},\overline{K}_{\mathrm{r}}\right)$落于最保守失效评定曲线上时，存在$\eta_{\mathrm{d}}=1$及$f_{\mathrm{s}}=1$，表明当结构处于可靠状态，NFAD 方法的分析结果与传统失效评定图方法最保守结果一致。

如表 2 示例 4～示例 8 所示，对于传统失效评定图方法，取不同失效评定点和失效评定曲线时，可靠性分析结果不同。这意味着将失效评定点和失效评定曲线作为固定值时，如果失效评定点及失效评定曲线不能精确获得，可能会导致完全相反的分析结果。而此时 NFAD 方法可对裂纹缺陷状态

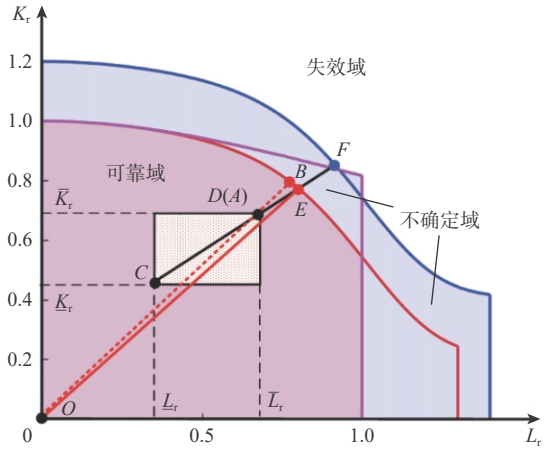


图 4 可靠性指标与安全系数间关系

Fig. 4 Relation between reliability index and safety factor

进行详细的分析，将其划分为可靠状态、非完全可靠状态及失效状态。

如表 2 示例 9～示例 10 所示，NFAD 方法与取不同失效评定点和失效评定曲线情况下传统失效评定图方法分析结果一致，结构处于失效状态。

据以上分析可知，NFAD 方法可以对裂纹缺陷的任意状态进行分析。与将结构分为失效或可靠的二元状态的传统失效评定图方法相比，NFAD 方法可以充分考虑分析中的不确定性，并将结构状态细分为 3 种情况。当$\eta_d \geqslant 1$时，结构处于可靠状态，当$0 < \eta_d < 1$时，结构处于非完全可靠状态，当$\eta_d = 0$时，结构处于失效状态。

4 实例分析

据 GB/T 19 624-2004 标准即可求得结构裂纹缺陷的应力强度因子K_I为

$$K_I = \sqrt{\pi d}(\sigma_m f_m + \sigma_d f_d) \tag{25}$$

式中：σ_m为薄膜应力；σ_d为弯曲应力；对于表面裂纹，f_m与f_d为与裂纹尺寸及壁厚相关的系数；对于埋藏裂纹，f_m与f_d为与裂纹尺寸、壁厚及裂纹中心与壳厚中心偏移量e相关的系数。

对于沿壁厚方向线性分布的应力，σ_m与σ_d可分别表示为

$$\sigma_m = \frac{\sigma_1 + \sigma_2}{2} \tag{26}$$

$$\sigma_d = \frac{\sigma_1 - \sigma_2}{2} \tag{27}$$

式中：σ_1与σ_2分别为内壁与外壁处应力。

由于推进剂储罐半径远大于壁厚，可将储罐内壁与外壁应力视为等值，当应力为区间变量时，薄膜应力区间中值σ_m^c取为储罐应力区间中值σ^c。同时存在弯曲应力最大值等于应力区间半径，即$\sigma_d^{max} = \sigma^r$，在保守情况下，将弯曲应力区间中值$\sigma_d^c$取为应力区间半径$\sigma^r$。

当已知储罐材料断裂韧性K_C的情况下，可得断裂比K_r为

$$K_r = \frac{K_I}{K_C} \tag{28}$$

对于半椭圆形表面裂纹，载荷比L_r可表示为

$$L_r = \frac{\sigma_d + \sqrt{\sigma_d^3 + 9(1-\zeta_1)^2\sigma_m^2}}{3(1-\zeta_1)^2\sigma_s} \tag{29}$$

式中：σ_s为材料屈服强度；$\zeta_1 = d_1 c_1 / B(c_1 + B)$，$d_1$为表面裂纹深度，$2c_1$为表面裂纹长度，$B$为壳厚，式(29)的适用范围为$d_1/B \leqslant 0.8$。

对于椭圆形埋藏裂纹，载荷比L_r可表示为

$$L_r = \frac{(3\zeta_2\sigma_m + \sigma_d)}{3\left[(1-\zeta_2)^2 + 4\zeta_2\vartheta\right]\sigma_s} + \frac{\sqrt{(3\zeta_2\sigma_m + \sigma_d)^2 + 9\left[(1-\zeta_2)^2 + 4\zeta_2\vartheta\right]\sigma_m^2}}{3\left[(1-\zeta_2)^2 + 4\zeta_2\vartheta\right]\sigma_s} \tag{30}$$

式中：$\zeta_2 = 2d_2c_2/B(c_2 + B)$，$\vartheta = v/B$，$2d_2$为埋藏裂纹深度，$2c_2$为埋藏裂纹长度，式(30)的适用范围为$2d_2/B \leqslant 0.8$，$v$为裂纹据表面最近距离，其与$d_2$及偏移量$e$之间的关系可表示为

$$\frac{v}{B} + \frac{e}{B} + \frac{d_2}{B} = \frac{1}{2} \tag{31}$$

某型推进剂储罐应力参数如表 3 所示。为充分验证 NFAD 方法的有效性，表示结构随裂纹尺寸变化由可靠状态到非完全可靠状态再到失效状态的过程。c_1^c取 0.005～0.02 m，$c_1^r = c_1^c \times 10\%$，$d_1^c$取 0.000 2～0.001 4 m，$d_1^r = d_1^c \times 10\%$，$e = 0.1$时，可求得储罐表面裂纹可靠性指标，如图 5 所示。

据图 5 可知，当裂纹尺寸已知的前提下，储罐裂纹缺陷的可靠性均可应用 NFAD 方法进行分析，储罐表面裂纹深度处与长度处可靠性指标均随d_1^c与c_1^c的增大而减小。

c_2^c取 0.005～0.02 m，$c_2^r = c_2^c \times 10\%$，埋藏裂纹$d_2^c$取 0.000 1～0.000 7 m，$d_2^r = d_2^c \times 10\%$，$e = 0.1$时，可求得储罐表面裂纹可靠性指标，如图 6 所示。

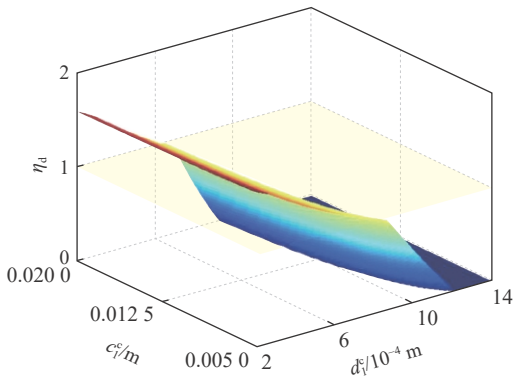
据图 6 可知，当裂纹尺寸已知的前提下，储罐裂纹缺陷的可靠性均可应用 NFAD 方法进行分析，储罐埋藏裂纹深度处与长度处可靠性指标随d_1^c与c_1^c的增大而减小。

综上所述，NFAD 方法可对储罐裂纹缺陷可靠

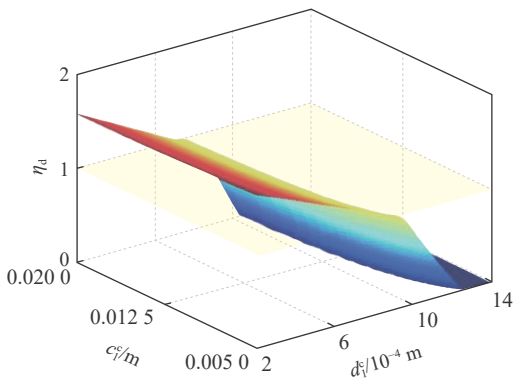
表 3 某型推进剂储罐应力参数

Table 3 Stress parameters of a certain propellant tank

σ_m^c	σ_m^r	σ_s^c	σ_s^r	σ_d^c	σ_d^r
101.469	30.147	300	15	30.147	3.015



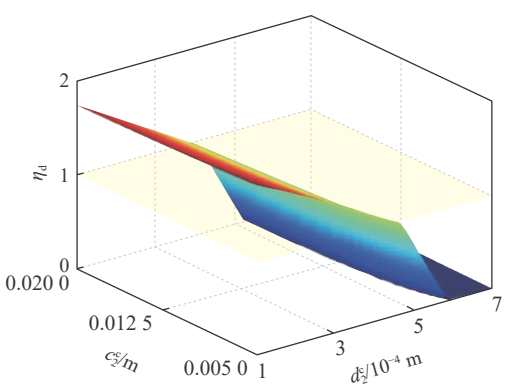
(a) 裂纹深度处



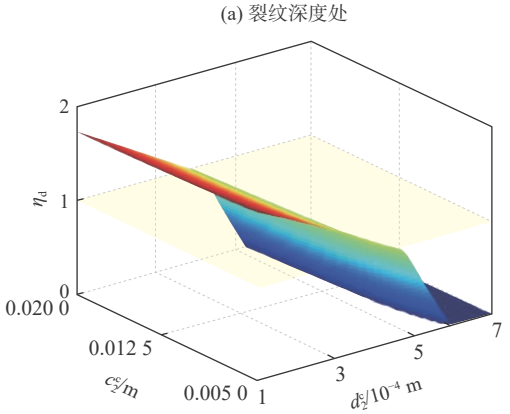
(b) 裂纹长度处

图 5　储罐表面裂纹可靠性指标

Fig. 5　Reliability index of tank surface crack



(a) 裂纹深度处



(b) 裂纹长度处

图 6　储罐埋藏裂纹可靠性指标

Fig. 6　Reliability index of embedded crack of tank

性进行分析。当储罐裂纹缺陷处于可靠状态时, 可靠性指标$\eta_d \geqslant 1$表示其安全裕度。当储罐处于非完全可靠状态时, 可靠性指标$0 < \eta_d < 1$表示其可靠度; 当储罐处于失效状态时, 可靠性指标$\eta_d = 0$。NFAD 方法无需参数的精确值即可对储罐裂纹缺陷的可靠性进行分析。

5　结　论

1) 结合区间理论与失效评定图理论, 将传统失效评定图方法的二值逻辑状态细化为 3 种, 即可靠状态、非完全可靠状态及失效状态。

2) 定义了一种全新的非概率失效评定图可靠性指标$\eta_d \in [0, +\infty)$, $\eta_d \geqslant 1$表示安全裕度; $0 < \eta_d < 1$表示可靠度; $\eta_d = 0$表示失效状态。

3) 结合实例参数, 对 NFAD 方法可靠性指标与传统失效评定图方法可靠性指标进行了对比分析, 并验证可靠性指标的有效性与可行性。

4) 无需失效评定点及失效评定曲线的精确值, 应用 NFAD 方法, 即可对结构的非概率可靠性进行分析。

参考文献 (References)

[1] BANNISTER A C, RUIZ OCEJO J, GUTIERREZ-SOLANA F. Implications of the yield stress/tensile stress ratio to the SINTAP failure assessment diagrams for homogeneous materials[J]. Engineering Fracture Mechanics, 2000, 67(6): 547-562.

[2] British Energy Generation Limited. Assessment of the integrity of structures containing defects : R6 Revision 4[S]. Gloucester: British Energy Generation Limited, 2001.

[3] British Standards Institute. Guide on methods for assessing the acceptability of flaws in metallic structures: BS 7910[S]. London: BSI Standards Limited, 2013.

[4] American Petroleum Institute and American Society of Mechanical Engineers. Fitness-for-service: API 579-1/ASME FFS-1[S]. Washington, D.C. : API Publishing Services, 2016.

[5] 国家市场监督管理总局, 国家标准化管理委员会. 在用含缺陷压力容器安全评定: GB/T 19624—2019[S]. 北京: 中国标准出版社, 2019.
State Administration for Market Regulation, Standardization Administration of the People's Republic of China. Safety assessment of in-service pressure vessels containing defects: GB/T 19624 — 2019[S]. Beijing: Standards Press of China, 2019 (in Chinese).

[6] WEI S J, TOMAC I. Numerical evaluation of failure assessment diagram (FAD) for hydraulic fracture propagation in sandstone[J]. Engineering Fracture Mechanics, 2022, 263: 108311.

[7] RADU D, SEDMAK A, BĂNCILĂ R. Determining the crack acceptability in the welded joints of a wind loaded cylindrical steel shell structure[J]. Engineering Failure Analysis, 2018, 91: 341-353.

[8] AINSWORTH R A, GINTALAS M, SAHU M K, et al. Application of failure assessment diagram methods to cracked straight pipes and elbows[J]. International Journal of Pressure Vessels and Piping,

2016, 148: 26-35.

[9] KINGKLANG S, DAODON W, UTHAISANGSUK V. Failure investigation of liquefied petroleum gas cylinder using FAD and XFEM[J]. International Journal of Pressure Vessels and Piping, 2019, 171: 69-78.

[10] ANDREWS R, COSHAM A, MACDONALD K. Application of BS 7910 to high pressure pipelines[J]. International Journal of Pressure Vessels and Piping, 2018, 168: 148-155.

[11] CHOCAT R, BEAUCAIRE P, DEBEUGNY L, et al. Damage tolerance reliability analysis combining Kriging regression and support vector machine classification[J]. Engineering Fracture Mechanics, 2019, 216: 106514.

[12] PLUVINAGE G, BOULEDROUA O, HADJ MELIANI M, et al. Corrosion defect analysis using domain failure assessment diagram[J]. International Journal of Pressure Vessels and Piping, 2018, 165: 126-134.

[13] XIAO Q D, LIU Y H, DAI Y W. Developments of strain-based failure assessment diagram applications: Measuring reference strain by displacement and a modified assessment method[J]. International Journal of Mechanical Sciences, 2018, 140: 27-36.

[14] ROTH S, KUNA M. Prediction of size-dependent fatigue failure modes by means of a cyclic cohesive zone model[J]. International Journal of Fatigue, 2017, 100: 58-67.

[15] RADU V, ROTH M. Probabilistic fracture mechanics applied for DHC assessment in the cool-down transients for CANDU pressure tubes[J]. Nuclear Engineering and Design, 2012, 253: 211-218.

[16] BOULEDROUA O, ZELMATI D, HASSANI M. Inspections, statistical and reliability assessment study of corroded pipeline[J]. Engineering Failure Analysis, 2019, 100: 1-10.

[17] JIN H J, WU S J. Effect of plasticity constraint on structural integrity assessment of pressure vessel welds[J]. International Journal of Pressure Vessels and Piping, 2015, 134: 72-81.

[18] BERGANT M A, YAWNY A A, PEREZ IPIÑA J E. A comparison of failure assessment diagram options for Inconel 690 and Incoloy 800 nuclear steam generators tubes[J]. Annals of Nuclear Energy, 2020, 140: 107310.

[19] FUENTES J D, CICERO S, IBÁÑEZ-GUTIÉRREZ F T, et al. On the use of British standard 7910 option1 failure assessment diagram to non-metallic materials[J]. Fatigue & Fracture of Engineering Materials & Structures, 2018, 41(1): 146-158.

Non-probabilistic reliability analysis method for propellent tank with crack defect

XIN Tengda[1,*], CUI Cunyan[1], LIU Yang[1], TONG Jiang[1], DUAN Yongsheng[2]

(1. Beijing Institute of Tracking and Telecommunications Technology, Beijing 100094, China;
2. Department of Aerospace Science and Technology, Space Engineering University, Beijing 101416, China)

Abstract: During the service period of the propellant tank, the accurate analysis of its crack defects and reliable state with crack defects can not only guarantee the safety of the launch site, but also effectively avoid unnecessary panic, providing a reference for emergency plan formulation. Based on interval theory and failure assessment diagram theory, a non-probabilistic failure assessment diagram (NFAD) is proposed. The reliability analysis of the crack defect of propellant tanks can be effectively conducted when it is difficult to accurately obtain the failure evaluation point and curve in engineering practice. The proposed method is verified with the example parameters. The results show that this method can analyze any state of tank crack defects without accurate failure assessment point and curve, fully considering the uncertainty in the analysis. The binary logic states of failure or reliability of the traditional failure assessment chart method are divided into three cases with he reliability index is equal to 0, the reliability index is greater than 0 and less than 1, and the reliability index is greater than or equal to 1, represent the failure state, reliability degree and safety margin respectively.

Keywords: propellant tank; crack defect; non-probabilistic reliability; failure assessment diagram; reliability index

Received: 2021-10-31; **Accepted**: 2022-01-27; **Published Online**: 2022-03-08 10:40

URL: kns.cnki.net/kcms/detail/11.2625.V.20220307.1439.006.html

Foundation item: Launch Comprehensive Safety Assessment and Test Verification Technology(ZRHT2020T06)

* **Corresponding author**. E-mail: xtd0701@126.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0684

基于虚拟飞行的混合翼身融合布局操稳特性

王延灵[1,2,*]，冯帅[1,2]，卜忱[1,2]，沈彦杰[1,2]，陈昊[1,2]，芦士光[1,2]

(1. 航空工业空气动力研究所，哈尔滨 150001；2. 低速高雷诺数气动力航空科技重点实验室，哈尔滨 150001)

摘　　要：针对混合翼身融合布局面临的三轴运动耦合和大迎角失稳等潜在操稳问题，研制三自由度虚拟飞行试验系统，基于动力学相似模型开展纵向和横航向开环虚拟飞行试验，对飞机的本体操稳特性进行研究。结果表明：该混合翼身融合布局飞机纵向和航向开环操纵均存在三轴运动耦合现象。纵向操纵会引起大迎角极限环失稳现象，迎角振荡平衡位置约为28°、振荡幅值约为2.56°、振荡主频率为0.55 Hz，振荡过程中气动力呈现非定常特性；V型尾翼偏航操纵响应呈现横向运动幅值最大、偏航运动次之、俯仰运动最小的特点。

关 键 词：虚拟飞行；混合翼身融合布局；操稳特性；俯仰极限环振荡；运动耦合

中图分类号：V221^{+}.3; TB553

文献标志码：A　　**文章编号**：1001-5965(2023)09-2337-08

混合翼身融合布局融合了常规布局与翼身融合体的特点，机身前半部中央体与机翼融合一体，后部分机身保持圆形与尾翼连接。该布局形式作为未来商用飞机的潜在构型之一，具有气动效率高、噪声低、排放少等优点。不仅如此，此种布局延长了机身后体增加控制力臂，提升尾部操纵面的控制效率，有利于解决纵向和航向的操稳问题[1]。

混合翼身融合布局延续了翼身融合布局高升阻比等气动特性的同时，也存在运动失稳、操纵面带来的运动耦合等潜在的操稳问题，有必要对其操稳特性进行更为准确的评估和预测[2]。

风洞虚拟飞行试验技术是一种约束线运动放开角运动的风洞飞行试验，通过人为操纵或控制程序进行各种科目飞行试验，并实时记录飞行数据，是开展飞行器操稳特性与飞行控制研究的重要试验手段之一[3-4]。世界主要航空大国基于虚拟飞行试验平台开展了操稳特性相关研究。英国克兰菲尔德大学和布里斯托大学基于虚拟飞行试验平台开展飞机动稳定性和控制特性的研究[5-6]。英国布里斯托大学Lowenberg项目团队在1.1 m量级开口风洞开展0.6 m展长BAE Hawk模型单自由度和双自由度俯仰运动研究，发现非线性极限环振荡现象，并采用动态增益策略对试验模型进行俯仰反馈控制，改善了其整个试验迎角区间的动态响应特性[7]，近年来该项目团队基于五自由度运动机构对非线性迎角区间的横航向偏离特性进行研究[8]。文献[9]在俄罗斯中央流体动力学研究院低速风洞中开展大迎角机翼摇滚及尾旋特性研究，验证纵向及横航向增稳系统的控制效果。文献[10]在单自由度机构上对飞翼布局飞机航向控制增稳特性进行研究。

国内虚拟飞行试验技术从2010年起陆续开展相关研究工作。中国空气动力研究与发展中心基于开发的虚拟飞行试验平台先后开展尾旋特性研究[11]、飞机结冰动力学特性与自适应控制[12]、倾转旋翼飞机悬停/前飞过渡过程飞行控制验证[13]及飞翼布局飞机先进飞机控制技术验证[14]等研究工作。航天空气动力技术研究院开展了类F-16飞行器的虚拟飞行试验，测量到大迎角俯仰失稳和负迎角横航向失稳等非定常试验现象[15]。南京航空航

收稿日期：2021-11-12；录用日期：2022-05-07；网络出版时间：2022-06-09 16:15
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220609.1028.001.html
*通信作者. E-mail：wangyanlingqdy@163.com

引用格式：王延灵，冯帅，卜忱，等. 基于虚拟飞行的混合翼身融合布局操稳特性[J]. 北京航空航天大学学报，2023，49(9)：2337-2344. WANG Y L，FENG S，BU C，et al. Stability characteristics of hybrid wing-body aircraft based on virtual flight[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2337-2344 (in Chinese).

天大学在 NH-2 低速风洞中开发了虚拟飞行试验系统，并针对翼身融合布局的偏离特性开展虚拟飞行验证试验[16-17]。综上所述，国内外研究机构应用虚拟飞行试验平台针对常规布局、飞翼布局和翼身融合布局等飞行器开展了操稳特性研究，取得了一定的研究成果。然而，从现有文献来看，采用风洞虚拟飞行试验研究混合翼身融合布局飞机面临的操稳特性问题鲜有报道。

本文针对混合翼身融合布局的潜在操稳问题，基于 4 m 量级的动力学相似模型开展三自由度开环虚拟飞行试验，并对飞机本体的俯仰极限环振荡和三轴耦合运动现象进行分析，为混合翼身融合布局气动优化和控制律设计提供一定参考。

1 混合翼身融合布局民机

混合翼身融合民机采用下单翼前缘延伸大边条、中度翼身融合、V 型尾翼（以下简称“V 尾”）和翼吊双发的布局特征，如图 1 所示，几何参数如表 1 所示。

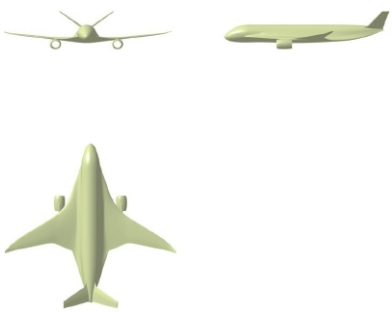


图 1 混合翼身融合布局民机

Fig. 1 HWB civil aircraft

表 1 混合翼身融合布局民机几何参数

Table 1 Geometric parameters of hybrid wing body civil aircraft

机翼面积S/m^2	展长L/m	平均气动弦长c/m	参考中心与机头距离/m
26.802	36	10.413	14.445

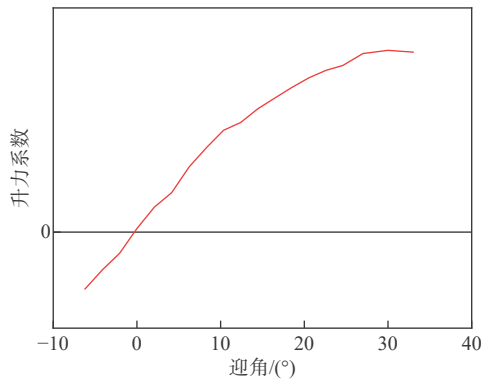
2 静态测力试验

为了解该混合翼身融合布局民机的气动特性，在 FL-51 风洞中进行静态测力试验，如图 2 所示，模型缩比为 1/15。

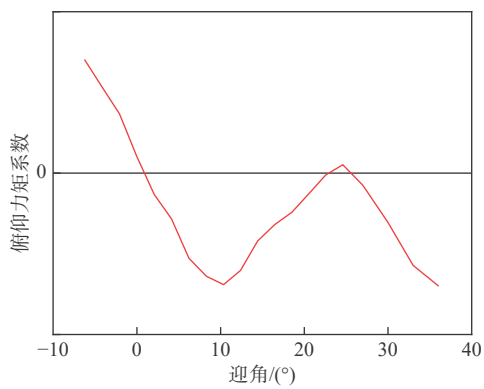
图 3 为零侧滑（侧滑角β=0°）下飞机升力系数和俯仰力矩系数的纵向静态测力试验结果，从图 3(b) 可知飞机迎角在−5°～10°范围内是俯仰静稳定的，在 10°～25°范围内，转为静不稳定。原因主要为飞机布局形式为大边条翼和翼身融合的组合导致气动焦点随迎角增加前移、纵向稳定性减小，

图 2 FL-51 风洞试验

Fig. 2 FL-51 wind tunnel test



(a) 升力系数



(b) 俯仰力矩系数

图 3 纵向静态测力试验结果（β=0°）

Fig. 3 Longitudinal static force wind tunnel test results（β=0°）

而迎角大于 25°，飞机又重新进入静稳定状态。

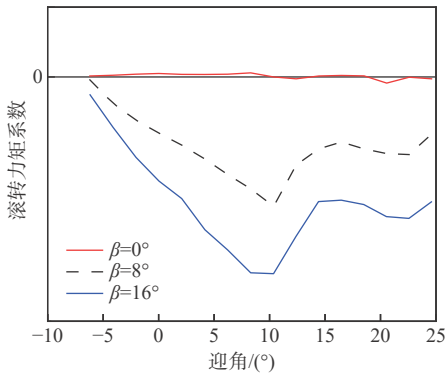
图 4 为横航向静态测力试验结果，图 4(a)和图 4(b)结果表明：该飞机迎角在−5°～25°，侧滑角在 0°～16°范围内具有横向静稳定性；迎角在−5°～20°，侧滑角β在 0°～16°范围内具有航向静稳定性。

图 4(c)为零侧滑下 V 尾舵面偏转 10°带来的横航向气动增量，为了表明舵面偏转带来的操纵特性，根据动力学方程推演得到计算式为

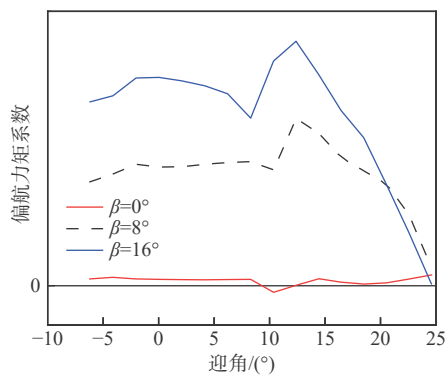
$$k = \left| \frac{\Delta C_l}{I_x} \cdot \frac{I_z}{\Delta C_n} \right| \tag{1}$$

式中：ΔC_l为滚转力矩系数增量；ΔC_n为偏航力矩系数增量；I_x、I_z分别为x、z轴惯性矩；k为滚转与偏航

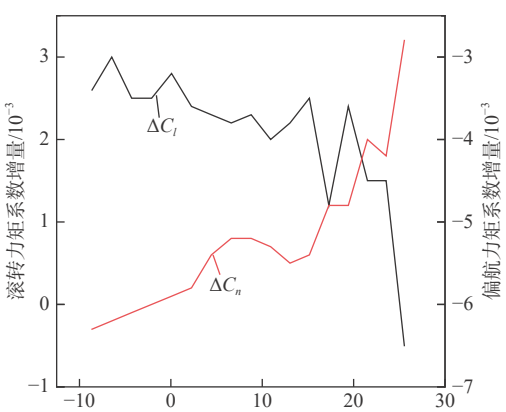
响应速度等效比例因子。经计算可知迎角在 0°～10°时, k 值均大于 1, 即 V 尾偏转不仅会带来飞机的偏航运动, 还伴有滚转运动, 且滚转运动响应速度高于偏航运动, 表明 V 尾带来的偏航滚转耦合较强。



(a) 滚转力矩系数



(b) 偏航力矩系数



(c) V尾舵效增量系数 (δ_f=10°)

图 4　横航向静态测力试验结果

Fig. 4　Lateral static force wind tunnel test results

3　虚拟飞行试验方法

3.1　相似准则与模拟方法

相似准则是虚拟飞行试验的理论基础, 为了使试验结果能反映全尺寸飞机的稳定与控制特性, 虚拟飞行试验必须在满足相似准则的状态下开展。根据动力学方程和流动方程, 动力学相似参数和流动相似参数可经推导得到。动力学相似参数主要包括几何相似、质量相似、惯量相似; 流动相似参数主要包括斯特劳哈尔数、弗劳德数、雷诺数和马赫数。虚拟飞行试验中除了几何相似, 弗劳德数 (Fr) 是必须满足的相似准则, 对应的参数相似关系如表 2 所示[18], 表中: V 为飞行速度, l 为飞机特征线性尺寸, m 为飞机质量, J 为转动惯量, g 为重力加速度, ρ 为空气密度, N 为缩比模型与全尺寸飞机的特征长度之比。

表 2　虚拟飞行试验参数相似准则（模型/原型机）[18]

Table 2　Scaling parameters for wind tunnel virtual flight test(sub-scale model/full-scale model prototype)[18]

参数	比例	参数	比例
线性尺寸	N	弗劳德数V^2/gl	1
线速度	$N^{0.5}$	质量$m/l^3\rho$	1
线加速度	1	惯量$J/l^5\rho$	1
角速度	$N^{-0.5}$	时间t	$N^{0.5}$
角加速度	N^{-1}		

3.2　虚拟飞行试验系统

混合翼身融合布局民机虚拟飞行试验在 8 m 量级 FL-10 闭口风洞中进行。该风洞闭口试验段截面尺寸为 8 m×6 m, 长度为 20 m, 空风洞最大风速为 110 m/s[19], 虚拟飞行试验模型的缩比为 1/9。

图 5 为试验模型在风洞中的安装状态。缩比模型通过低阻尼三自由度转动机构(见图 6)与腹撑支杆相连接, 并安装于风洞试验段中。三自由度转动机构将模型的线位移约束, 具有 3 个角运动自由度, 其中滚转角运动范围为±45°、俯仰角运动范围为±45°、偏航角运动范围为±180°。

图 5　试验模型在风洞中的安装状态

Fig. 5　Installation state of test model in wind tunnel

虚拟飞行试验系统示意图如图 7 所示, 缩比模型内部装有机载姿态测量传感器、飞行控制律计算机和驱动操纵面偏转的舵机。机载姿态测量传感器用于测量飞机运动参数, 飞行控制计算机采集机载姿态测量传感器的数据, 运行控制律算法, 通过舵面控制缩比模型的姿态。试验控制地面站可以实时显示试验过程中的相关数据, 包括迎角、侧滑

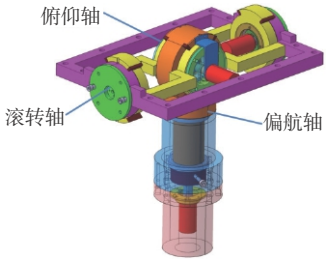


图 6 三自由度转动机构

Fig. 6 3-DOF rotate device

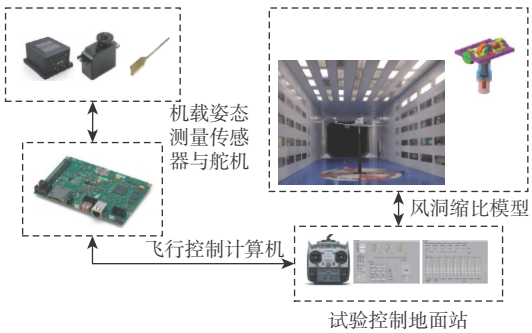


图 7 虚拟飞行试验系统示意图

Fig. 7 Diagram of virtual flight test system

角、姿态角和角速度及舵面实时信息，同时地面站可以发出典型激励指令。

4 典型风洞试验结果分析

针对该混合翼身融合民机存在的潜在操稳问题，分别开展纵向和横航向开环虚拟飞行试验，试验风速范围为 20～42 m/s。试验过程中，横向指令对应副翼偏转；左、右 V 尾作为升降舵和方向舵的共同操纵面，其中左、右 V 尾同时向下偏转视为升降舵正向偏转，左 V 尾向下偏转、右 V 尾向上偏转视为方向舵正向偏转。

4.1 三轴耦合运动特性

为了评估该民机的三轴耦合运动特性，以下试验均采用三轴开环控制方式。

4.1.1 俯仰操纵特性

操纵模型升降舵从 8°到−6°间隔 2°连续变化，相应的平衡迎角变化范围为−3°～33°。试验过程中响应曲线如图 8 所示。图 8(a)为迎角和升降舵的响应曲线，当升降舵从 8°间隔变化到−4°时，迎角均振荡衰减至平衡位置，而当升降舵为−6°时，迎角迅速增大，出现连续振荡现象，振荡平衡位置约为 28°。改变升降舵至 0°，迎角振荡消除。

图 8(b)为纵向操纵过程中横航向运动响应曲线。可以看出，迎角迅速增大进入振荡的过程中，伴随着滚转偏离运动和大幅度侧滑角振荡。此现象表明在大迎角非线性运动区域，该民机纵向操纵会带来横航向的运动，三轴耦合特性较强。

图 8(b)中显示升降舵恢复至 0°后(图 8 中蓝色实线后区域)，滚转角ϕ仍然持续偏离，最终稳定至−40°附近。为了解释这一持续偏离现象，图 9 给出滚转角持续运动过程中的三轴姿态角速度时间历程曲线。结合图 8 和图 9 结果，分析滚转角偏离原因为：在侧滑角不断增加的过程中(见图 8(b))，该民机由于横向静稳定性会产生负向滚转力矩与负向滚转角速度(图 9 中蓝色实线)，在已有的负滚转角的基础上滚转运动持续同方向(负向)偏离。

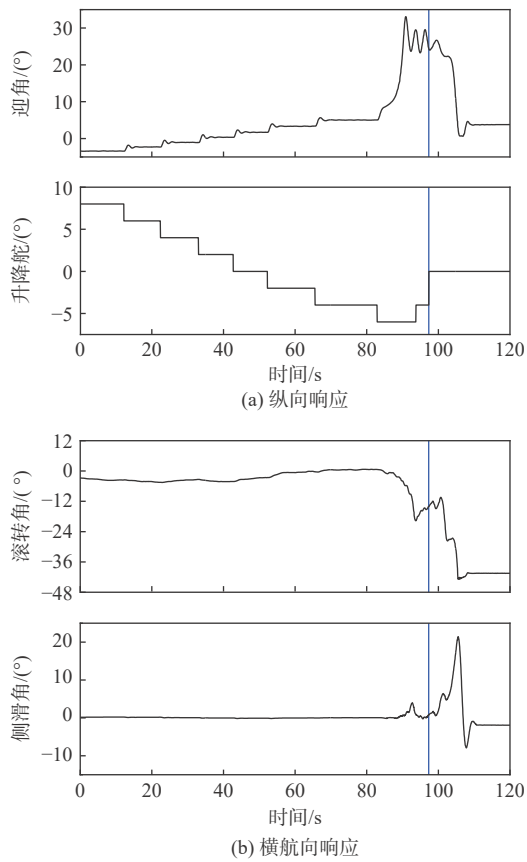


图 8 纵向操纵三轴开环响应结果

Fig. 8 Three-axis open-loop responses results of longitudinal control

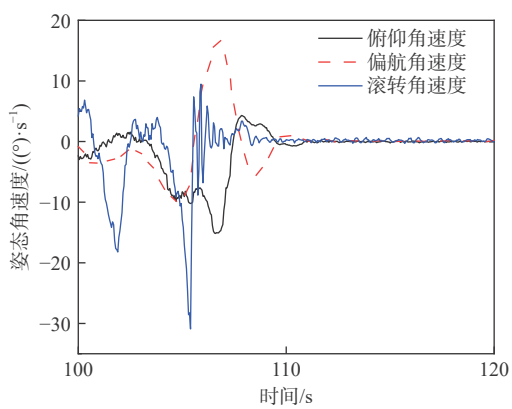


图 9 姿态角速度变化曲线

Fig. 9 Velocity variation curve of angular rates

为了消除上述大迎角三轴耦合运动, 开展控制增稳措施研究。首先, 保持俯仰开环控制方式, 偏航、滚转采用经典比例-积分-微分(proportional integral derivative, PID)控制进行零度保持。升降舵同样为2°间隔连续变化, 迎角响应曲线如图 10 中区域 1 所示。可以看出, 升降舵为−6°时, 迎角出现连续振荡现象, 并仍伴有滚转角和侧滑角的大幅度改变, 分析原因为横航向控制增益效果不足, 无法完全抑制横航向耦合运动。

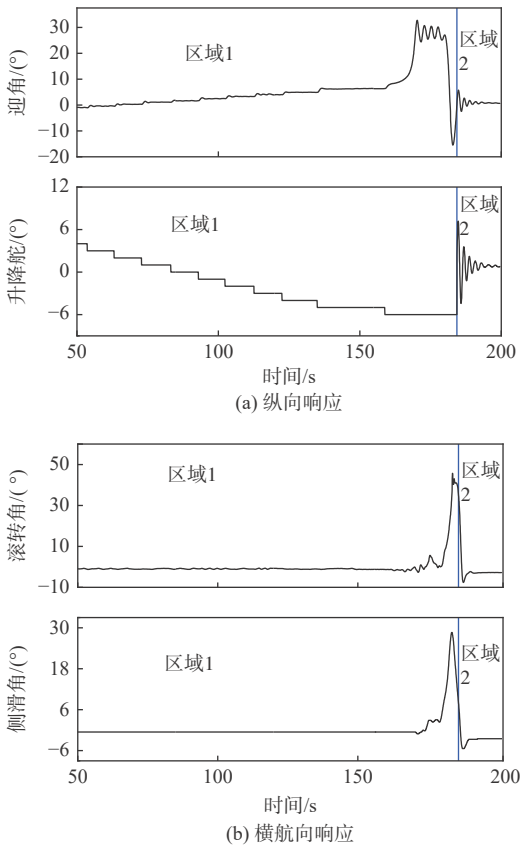


图 10　纵向操纵响应结果

Fig. 10　Responses results of longitudinal control

同时, 图 10 区域 1 中反映了另一种运动现象即: 不进行外部操纵输入, 迎角在振荡一段时间后会自行迅速降低。为分析此现象成因, 图 11 给出左、右 V 尾舵机编码器实时测量结果。可以看出, 为了控制滚转角和侧滑角为 0°, 横航向闭环控制解算的左、右 V 尾舵面产生较大偏转, 由于 V 尾共同控制俯仰和偏航运动, 升降舵量值也随之改变, 导致迎角大幅下降。此种现象再次说明 V 尾操纵带来的俯仰和偏航耦合运动较为明显, 在设计控制增稳措施时要着重考虑。

进一步, 不改变横航向控制增益, 对俯仰运动实施闭环控制(图 10 中区域 2), 迎角、侧滑角和滚转角均振荡衰减至 0°, 试验停止。

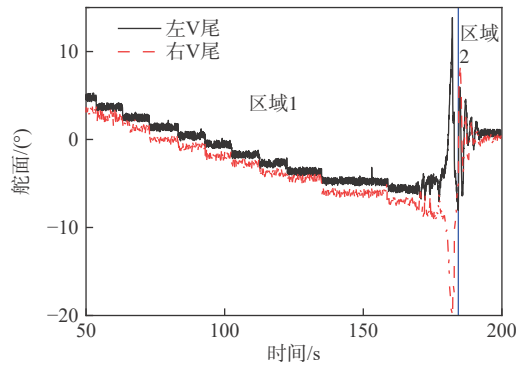


图 11　V 尾编码器测量结果

Fig. 11　V tail encoder measurement results

4.1.2　偏航操纵特性

针对 V 尾偏航操纵带来的运动耦合问题, 对方向舵采用连续方波激励, 幅值为±2°。三轴姿态角响应曲线如图 12 所示, 运动历程结果表明方向舵偏转后飞机呈现明显的三轴耦合状态, 响应特性呈现滚转角运动幅度最大、偏航角次之、俯仰角最小的特点, 其中滚转角运动范围为±20°, 而偏航角运动范围为±6°, 此现象与第 2 节中的静态舵效试验分析结果一致。

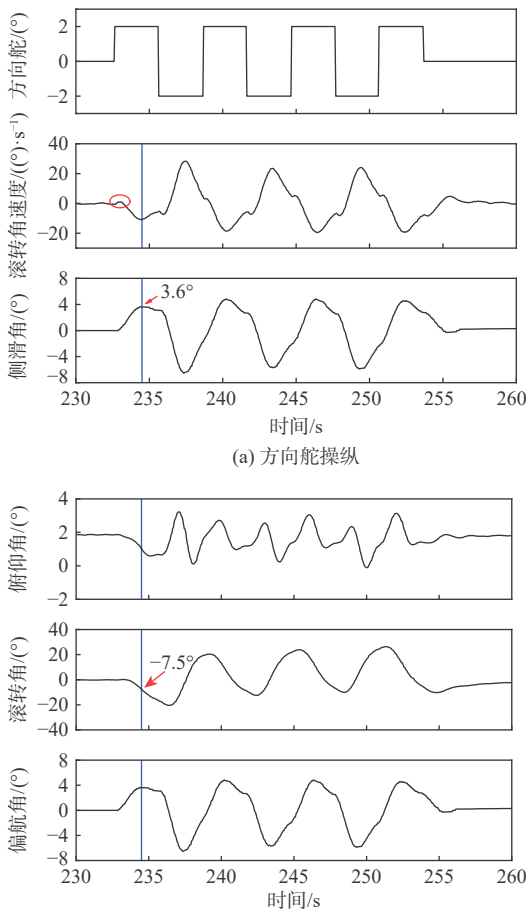


图 12　偏航开环试验结果

Fig. 12　Yaw open-loop test results

图 11 中蓝色实线显示滚转角偏转幅值大于侧滑角，分析原因为：方向舵操纵产生的偏航操纵力矩产生较大的侧滑角（如图 12(a)中第 3 条曲线），横向稳定性使得飞机产生较大的滚转力矩和滚转角速度，从而产生较大的滚转角。

由飞机运动特性可知，方向舵正向偏转产生正向侧滑运动，同时带来负向滚转运动。但图 12(a)中红色虚线显示，方向舵正向偏转带来了小幅度的正向滚转角速度，即引起了非单调操纵响应。分析原因为：V 尾力矩作用点位于机身上方较高位置，方向舵正向偏转会产生正向滚转力矩，从而引起短时间内的滚转角速度正向变化。

4.2 俯仰大迎角失稳特性

针对纵向开环操纵试验中出现的大迎角振荡现象，开展单自由度纵向标准激励试验，获取迎角响应历程，从而研究该飞机的俯仰大迎角失稳特性。

在配平迎角 8°位置对升降舵施加正弦激励，幅值为±4°，迎角产生振荡现象如图 13 所示，振荡平衡位置约为 28°。再次，在配平迎角 8°位置对升降舵施加“3−2−1−1”激励，幅值为±2°，图 13 中结果显示不同幅值的激励信号产生相同的迎角振荡现象，且 2 次振荡平衡位置一致，迎角振荡平衡位置 28°与静态测力中俯仰力矩为零的迎角位置 26°（见图 3(b)）较为接近。存在差异原因可能为：虚拟飞行试验存在配平升降舵偏角，使得俯仰力矩为零位置不同。

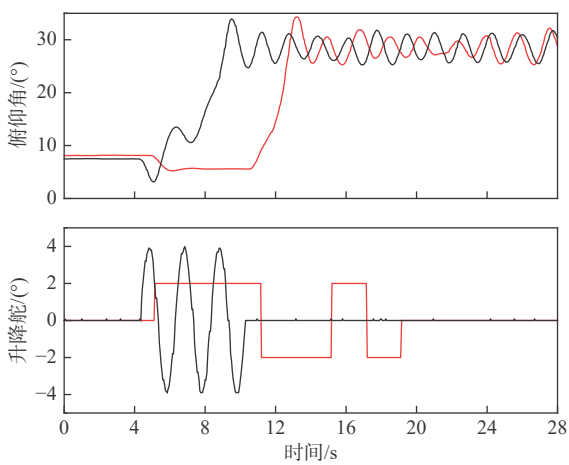


图 13 迎角扰动响应曲线

Fig. 13 Perturbation responses curves of angle of attack

图 14 和图 15 分别为俯仰极限环振荡的相平面曲线和振荡频率特性，可以看出振荡主频为 0.55 Hz，振荡幅值为 2.56°。

为进一步探究俯仰极限环持续振荡的受力特性，根据振荡运动历程，得到飞机运动过程中受到的俯仰力矩系数 C_m 为

$$C_m = \frac{I_y \dot{q}}{QSc} \tag{2}$$

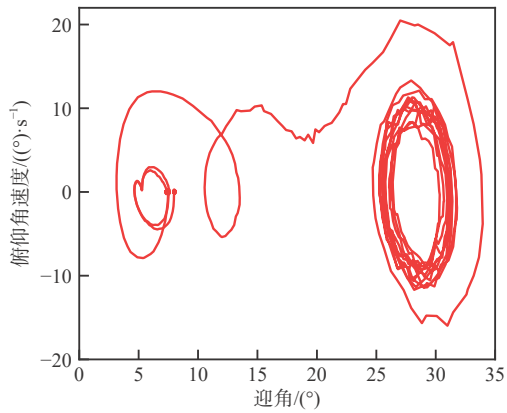


图 14 俯仰极限环振荡相平面曲线

Fig. 14 Pitch limit cycle oscillation phase plan curve

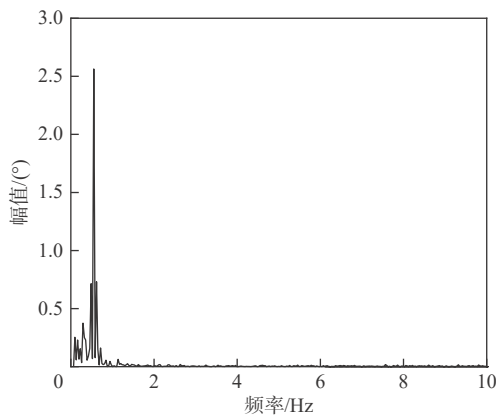


图 15 俯仰极限环振荡频率特性

Fig. 15 Oscillation frequency characteristics of pitch limit cycle

式中：I_y为 y 轴的惯性矩；$\dot{q}$为俯仰角加速度；Q 为来流动压；S 为机翼面积；c 为平均气动弦长。

图 16 为不同时间段的俯仰力矩系数迟滞曲线，可以发现飞机运动过程中的不同时间范围内，迟滞曲线呈现“8”字、“单向”或双“8”字环，方向如图中箭头所示，迟滞环的形状和交叉点的位置随时间变化而不同，体现了运动过程中的非定常特性。时间段 t_3 内，迟滞曲线在迎角 26°和 29°附近存在 2 个交叉点，构成了双“8”字环，方向为逆时针−顺时针−逆时针。

从流动机理角度分析，极限环振荡主要是迎角增大使得翼身融合布局背风区表面发生规流动分离和再附着过程，导致俯仰动稳定性降低，从而产生极限环振荡现象，振荡过程中由于流动的非定常特性使得俯仰迟滞曲线呈现不同形状。

本节针对虚拟飞行试验中的俯仰极限环振荡现象进行时域与频域特性分析，获得了俯仰振荡的平衡角、振幅和主频率，并且从俯仰极限环振荡的力矩特性曲线分析可看出俯仰方向具有明显的非定常特性。上述分析不仅能够验证虚拟飞行试验的可靠性，同时说明虚拟飞行试验能够针对混合翼

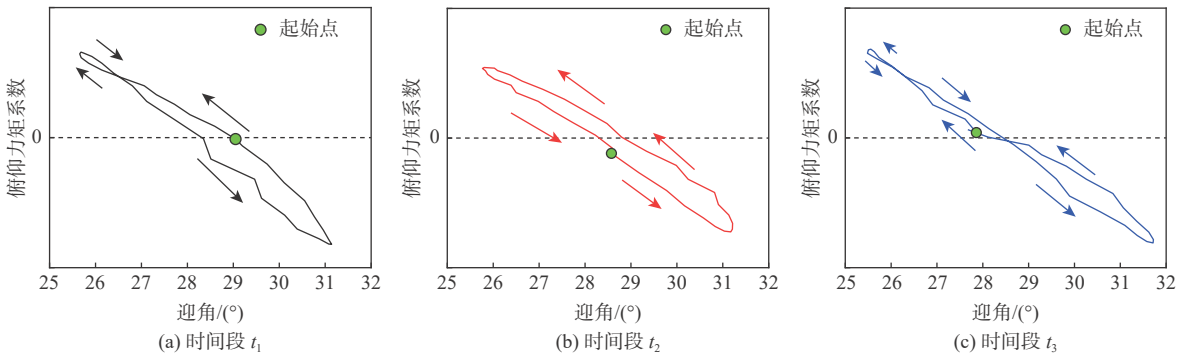


图 16　俯仰极限环振荡俯仰力矩特性

Fig. 16　Characteristics of oscillating pitching moment of pitching limit cycle

身融合等先进布局的操稳特性进行评估分析。

5　结　论

1）为评估混合翼身融合民机操稳特性，开展三自由度虚拟飞行试验。试验结果表明，该混合翼身融合民机纵向和航向开环操纵均会引起三轴耦合运动，尤其是大迎角区域。

2）升降舵俯仰开环操纵存在大迎角极限环振荡现象，迎角振荡平衡位置约为 28°、振荡幅值约为 2.56°、振荡主频率为 0.55 Hz；偏航操纵呈现横向运动幅值最大，偏航运动次之，俯仰运动最小的特点，其中横向幅值达到±20°、偏航幅值为±6°。

3）俯仰极限环振荡过程中俯仰力矩迟滞曲线在不同时间段内，呈现“8”字、“单向环”或双“8”字环，且迟滞环交叉点的位置随时间变化而不同，体现了俯仰运动的非定常特性。

参考文献（References）

[1] KAWAI R. Acoustic prediction methodology and test validation for an efficient low-noise hybrid wing body subsonic transport: NF1676L-14465[R]. Washington, D. C. : NASA Langley Research Center, 2011.

[2] REIST T A, ZINGG D W, RAKOWITZ M, et al. Multifidelity optimization of hybrid wing–body aircraft with stability and control requirements[J]. Journal of Aircraft, 2019, 56(2): 442-456.

[3] LU Y, ZHANG S G, ZHANG Z J, et al. Multiple hierarchy risk assessment with hybrid model for safety enhancing of unmanned subscale BWB demonstrator flight test[J]. Chinese Journal of Aeronautics, 2019, 32(12): 2612-2626.

[4] RATLIFF C, MARQUART E. An assessment of a potential test technique: AIAA-1995-3415[R]. Reston: AIAA, 1995.

[5] MAGILL J C, CATALDI P, MORENCY J R, et al. Demonstration of a wire suspension for wind-tunnel virtual flight testing[J]. Journal of Spacecraft and Rockets, 2009, 46(3): 624-633.

[6] CARNDUFF S, ERBSLOEH S, COOKE A, et al. Development of a low cost dynamic wind tunnel facility utilizing MEMS inertial sensors[C]//Proceedings of the 46th AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2008.

[7] GATTO A, LOWENBERG M H. Evaluation of a three degree of freedom test rig for stability derivative estimation[J]. Journal of Aircraft, 2006, 43(6): 1747-1761.

[8] GATTO A. Application of a pendulum support test rig for aircraft stability derivative estimation[J]. Journal of Aircraft, 2009, 46(3): 927-934.

[9] GRISHIN I, KHRABROV A, KOLINKO A, et al. Wind tunnel investigation of critical flight regimes using dynamically scaled actively controlled model in 3 dof gimbal[C]//29th Congress of the International Council of the Aeronautical Sciences.Washington, D. C. : I CAS, 2014.

[10] STENFELT G, RINGERTZ U. Yaw control of a tailless aircraft configuration[J]. Journal of Aircraft, 2010, 47(5): 1807-1811.

[11] 郭林亮, 祝明红, 傅澔, 等. 水平风洞中开展飞机尾旋特性研究的理论分析[J]. 航空学报, 2018, 39(6): 122030.
GUO L L, ZHU M H, FU H, et al. Theoretical analysis of research on aircraft spin characteristic in horizontal wind tunnel[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(6): 122030(in Chinese).

[12] 郭林亮, 祝明红, 吴海瀛, 等. 结冰条件下大型民机操稳特性研究与风洞虚拟飞行验证[C]//中国力学大会论文集, 2019: 3261-3271.
GUO L L, ZHU M H, WU H Y, et al. Study on stability and controllability of large civil aircraft under icing conditions and verification via virtual flight in wind tunnel[C]//Proceedings of Chinese Conference of Theoretical and Applied Mechanics, 2019: 3261-3271.

[13] 聂博文, 杨洪森, 刘志涛, 等. 倾转四旋翼无人机飞行控制系统集成与风洞虚拟飞行试验应用[C]//中国空气动力学会论文集, 2019: 162-167.
NIE B W, YANG H S, LIU Z T, et al. Flight control system integration and wind tunnel virtual flight test of a quad-rotor aircraft[C]//Proceeding of Chinese Aerodynamics Research Society, 2019: 162-167(in Chinese).

[14] 郭天豪. 飞翼布局飞机虚拟飞行试验报告[R]. 绵阳: 中国空气动力研究与发展中心, 2019.
GUO T H. Report of wind tunnel virtual flight for a flying wing[R]. Mianyang: China Aerodynamics Research and Develoment Center, 2019(in Chinese).

[15] 张石玉, 赵俊波, 付增良, 等. 类F-16飞行器风洞虚拟飞行试验研究[J]. 实验流体力学, 2020, 34(1): 49-54.
ZHANG S Y, ZHAO J B, FU Z L, et al. Wind tunnel based virtual flight testing research of F-16 fighter[J]. Journal of Experiments in Fluid Mechanics, 2020, 34(1): 49-54(in Chinese).

[16] 付军泉, 史志伟, 耿玺, 等. 基于试验分岔分析的翼身融合飞行器纵向稳定性[J]. 航空学报, 2022, 43(1): 124931.
FU J Q, SHI Z W, GENG X, et al. Longitudinal stability of blended-wing-body aircraft based on experimental bifurcation analysis[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(1): 124931(in Chinese).

[17] 付军泉, 史志伟, 陈杰, 等. 一种翼身融合布局飞行器的偏离特性分析[J]. 实验流体力学, 2020, 34(6): 32-37.
FU J Q, SHI Z W, CHEN J, et al. Departure characteristics of blended-wing-body aircraft[J]. Journal of Experiments in Fluid Mechanics, 2020, 34(6): 32-37(in Chinese).

[18] 岑飞, 聂博文, 刘志涛, 等. 低速风洞带动力模型自由飞试验[J]. 航空学报, 2017, 38(10): 121214.
CEN F, NIE B W, LIU Z T, et al. Low speed wind tunnel free-flight test of powered sub-scale aircraft[J]. Acta Aeronautica et Astronautica Sinica, 2017, 38(10): 121214(in Chinese).

[19] 李周复. 风洞试验手册[M]. 北京: 航空工业出版社, 2015: 4-5.
LI Z F. Handbook of wind tunnel test[M]. Beijing: Aviation Industry Press, 2015: 4-5 (in Chinese).

Stability characteristics of hybrid wing-body aircraft based on virtual flight

WANG Yanling[1, 2, *], FENG Shuai[1, 2], BU Chen[1, 2], SHEN Yanjie[1, 2], CHEN Hao[1, 2], LU Shiguang[1, 2]

(1. AVIC Aerodynamics Research Institute, Harbin 150001, China;
2. Key Laboratory of Aviation Technology for Aerodynamics for Low Speed and High Reynolds Number, Harbin 150001, China)

Abstract: A three-degree-of-freedom wind tunnel virtual flight test system was created in an effort to address the possible stability issues of three-axis kinematic coupling and high angle-of-attack instability presented by the hybrid wing-body aircraft. Based on the dynamic similarity model, longitudinal and transverse open-loop virtual flight tests were carried out to study the stability characteristics of the aircraft. The research results show that there is a three-axis motion coupling phenomenon in both the longitudinal and directional open-loop manipulation of the hybrid wing-body aircraft. Longitudinal manipulation will cause the instability of the limit cycle at high angles of attack. The equilibrium position of oscillation is about 28°, the oscillation amplitude is about 2.56°, and the main oscillation frequency is 0.55 Hz. The aerodynamic force presents unsteady characteristics during the oscillation process; The lateral motion amplitude of the yaw manipulation is the largest, followed by yaw motion and pitch motion.

Keywords: virtual flight; hybrid wing body; stability characteristics; pitch limit cycle oscillation; kinematic coupling

Received: 2021-11-12; **Accepted**: 2022-05-07; **Published Online**: 2022-06-09 16:15
URL: kns.cnki.net/kcms/detail/11.2625.V.20220609.1028.001.html
* **Corresponding author**. E-mail: wangyanlingqdy@163.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0691

SERF 原子自旋惯性测量检测误差分析及抑制

邢力[1]，全伟[2,3]，宋天骁[4]，蔡庆中[3,*]，叶文[5]

(1. 中国计量科学研究院 热工计量科学研究所，北京 100029；2. 北京航空航天大学 前沿科学技术创新研究院，北京 100191；3. 北京航空航天大学 仪器科学与光电工程学院，北京 100191；4. 北京航天发射技术研究所，北京 100076；5. 中国计量科学研究院 力学与声学计量科学研究所，北京 100029)

摘 要：在无自旋交换弛豫 (SERF) 原子自旋惯性测量装置中，检测系统的性能是决定输出信号灵敏度和稳定性极限的关键因素。为有效抑制 SERF 原子自旋惯性测量的低频随机噪声，在横向电子自旋极化率稳态解和旋光角表达式的基础上，建立检测相关的输出信号误差机理模型，明确了检测系统影响最终测量信号噪声的主要因素。研究结果表明：检测光入射气室时的初始光强作为信号背景直接引起刻度系数的波动而非作用于电子自旋，而检测光的非理想线偏振导致电子自旋产生横向抽运效果和横向光频移作用，从而引入测量误差。针对检测系统影响测量信号噪声的主要参数，设计了参数优化路径：先通过优化检测光频率，提高惯性测量刻度系数，后通过检测光功率的优化，减小横向抽运效应、横向光频移作用及检测光背景波动。实验表明：通过对比输出信号的 Allan 方差，优化后 SERF 原子自旋惯性测量的零偏不稳定性减小了 1.8 倍，速率斜坡噪声系数从 0.124 (°)/h^2 减小至 0.041 (°)/h^2，达到了抑制测量信号低频随机噪声的效果。

关 键 词：无自旋交换弛豫；惯性测量；检测系统；误差机理；横向抽运效应；横向光频移；低频随机噪声

中图分类号：V249.3; TN294

文献标志码：A **文章编号**：1001-5965(2023)09-2345-06

近年来为满足高分辨率对地观测、水下勘察、航空航天等领域对惯性测量器件稳定性和灵敏度需求的提升，亟需探索兼具超高灵敏、高稳定性能的惯性测量器件[1-2]。惯性测量器件自诞生以来经历了机电式[3]、光电式[4-5]两代的变革，然而近十年来其性能的发展进入了瓶颈期，灵敏度与稳定性的进一步提升亟待技术手段上的突破。随着量子操控与光学探测技术的发展，利用原子效应敏感角速率信息的方法为突破惯性测量器件性能提供了新的可能，寄希望其能实现超高灵敏与超高稳定的惯性测量[6-7]，原子式的惯性测量方法得到了国内外学者的广泛关注。目前原子式惯性测量器件主要采用原子干涉[8]或原子自旋两大原理，基于原子自旋原理的惯性测量方法主要有核磁共振[9]和无自旋交换弛豫 (spin-exchange relaxation-free，SERF)[10]这2类。其中，SERF 原子自旋惯性测量使碱金属原子的价电子在高密度、近零磁场的环境下实现了自旋交换弛豫的大幅抑制，惰性气体核自旋与电子自旋通过自旋交换碰撞超极化后，可以实现角速率的敏感且其自旋磁矩发生变化，而 SERF 态下电子自旋对核自旋产生的磁矩变化具有超高灵敏的特性[11]，从而可实现超高灵敏的惯性测量。

SERF 原子自旋惯性测量校早于 2005 年由普林斯顿大学 Romalis 等[10]实现，并达到灵敏度为 3.5×10^{-5} (°)·s^{-1}·Hz$^{-\frac{1}{2}}$@0.5 Hz，低频角速率漂移为 0.04 (°)/h 的性能指标。随后通过对装置结构及真

收稿日期：2021-11-16；录用日期：2021-12-27；网络出版时间：2022-02-15 11:01
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220214.1617.008.html
基金项目：国家自然科学基金 (61803015)；中国博士后科学基金 (2021M703049)
*通信作者. E-mail：qingzhong_cai@buaa.edu.cn

引用格式：邢力，全伟，宋天骁，等. SERF 原子自旋惯性测量检测误差分析及抑制 [J]. 北京航空航天大学学报，2023，49（9）：2345-2350.
XING L，QUAN W，SONG T X，et al. Error analysis and suppression of probe system for SERF atomic spin co-magnetometer[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（9）：2345-2350（in Chinese）.

空系统的改进，普林斯顿大学于2011年将惯性测量灵敏度提升至8.7×10^{-6} (°)·s^{-1}·$Hz^{-\frac{1}{2}}$@0.2 Hz的水平[12]，并搭建了南极实验平台，克服了地球自转对惯性测量的影响，使低频惯性测量灵敏度达到了2.8×10^{-6} (°)·s^{-1}·$Hz^{-\frac{1}{2}}$[13]。2009年美国Twinleaf公司基于惯性测量装置的理论研究，开展了小型化低漂移SERF原子自旋惯性测量器件的研究，并于2018年发布了研究进展报告，其采用直径为10 mm的球形气室实现了稳定性为0.01 (°)/h的性能指标，此外基于阳极键合技术制成直径为2 mm的小型化气室和垂直腔面发射激光器(vertical-cavity surface-emitting laser, VCSEL)，实现了1.6 cm^3的物理表头体积[14-15]，证明了SERF原子自旋惯性测量装置具备高精度、小体积的潜力，是新一代惯性测量器件的发展方向之一。

目前，SERF原子自旋惯性测量装置主要由抽运系统、磁屏蔽及磁补偿系统、无磁电加热系统、检测系统、原子源组成。原子自旋进动的测量主要由检测系统实现，检测方法通常基于极化了的碱金属原子对线偏振检测激光的双折射原理，因此，检测系统的性能决定了最终输出信号灵敏度及稳定性的极限，同时非理想检测光中的圆偏振成分会与原子自旋相互作用，从而引入测量误差。因此，检测系统的误差机理分析与噪声抑制，对提升惯性测量的信噪比和稳定性十分关键，本文首先利用横向电子自旋极化率的稳态解与旋光角的输出方程，分析并建立检测系统的误差模型，明确检测系统探测到的电压信号的表达式，给出惯性测量输出信号的简化误差模型，并理清检测光对输出信号造成影响的主要因素及形式，最终根据理论分析提出检测系统误差的参数优化路径，实现惯性测量输出信号信噪比的提升及低频随机噪声的抑制。

1 检测系统误差机理模型

SERF原子自旋惯性测量中通常采用存在一定频率失谐的线偏振检测光，对电子自旋极化率在检测光方向的投影进行探测。被抽运光极化了的原子自旋具有旋光效应，因此，线偏振检测光入射碱金属气室时，会产生双折射现象[16]，即由于线偏振检测光可被视为左旋和右旋圆偏振光的合成，而极化了的原子自旋对其折射率不同。因此，检测光入射碱金属气室后，与入射前相比会产生旋光角的变化，最终通过偏振检测方法可实现对此旋光角的探测，从而得到横向电子自旋极化率P_x^e的信息，检测光旋光角θ可表示为[17]

$$\theta=\frac{\pi}{2}lnr_e cP_x^e f\left[\frac{-(\nu_{probe}-\nu_1)}{(\nu_{probe}-\nu_1)^2+(\Gamma_1/2)^2}+\frac{(\nu_{probe}-\nu_2)}{(\nu_{probe}-\nu_2)^2+(\Gamma_2/2)^2}\right] \tag{1}$$

式中：l为检测光通过的气室有效长度；n为饱和蒸气压下碱金属原子数密度；r_e为电子的经典半径；c为光速；f为振子强度，本实验中取常数值1/3；ν_1和ν_2分别为碱金属原子D1线和D2线的跃迁频率；ν_{probe}为检测光的频率；Γ_1和Γ_2分别为D1线和D2线压力展宽的半高全宽。

P_x^e可通过描述SERF原子自旋惯性测量动力学过程的Bloch方程[18]稳态解获得，其与角速率Ω在一定动态范围内成正比，即有$P_x^e\approx\gamma_e P_z^e R_{tot}^e\left(\tilde{\Omega}_y+b\right)/\left(R_{tot}^{e2}+\gamma_e^2(L_z+\delta B_z)^2\right)$。其中$\tilde{\Omega}_y=(1/\gamma_n-Q/\gamma_e)\Omega$；$\gamma_e$和$\gamma_n$分别为电子自旋和核自旋的旋磁比，$Q$为电子的减慢因子，$P_z^e$为沿抽运光的纵向电子自旋极化率，$L_z$为纵向光频移，$\delta B_z$为纵向磁补偿后的剩余磁场，$R_{tot}^e$为电子自旋总弛豫率，其值为自旋交换弛豫率$R_{se}^{en}$、自旋破坏率$R_{sd}^e$、抽运光产生的纵向抽运率$R_p$及检测光产生的横向抽运率$R_m$之和，$b$为包括横向剩余磁场$B_m$、横向光频移$L_m$、横向抽运率$R_m$等相关的误差项[17]。通常采用法拉第调制、光弹调制和偏振平衡分束等检测方法，本文以最具普适性的偏振平衡分束方法为例建立检测系统的误差机理模型，基本光路示意图如图1所示，其由半波片与偏振分光棱镜(polarization beam splitter, PBS)组成差分模块，将透射光与反射光束进行差分作为包含旋光角的输出信号，该检测方法探测到的电压信号可描述为

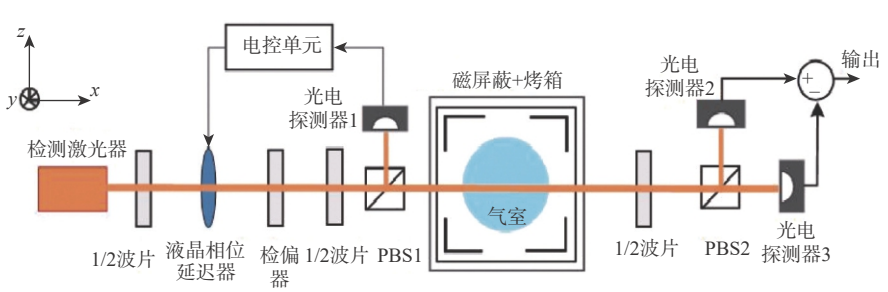


图1 基于偏振平衡分束的检测光光路示意图

Fig. 1 Schematic diagram of probe beam based on polarization balance difference optical path

$$S_{\text{output}} = K_{\text{PD}} I_0 \theta = K_{\text{PD}} \pi I_0 l n r_{\text{e}} c P_x^{\text{e}} f \left[\frac{(\nu_{\text{probe}} - \nu_1)}{(\nu_{\text{probe}} - \nu_1)^2 + (\Gamma_1/2)^2} + \frac{(\nu_{\text{probe}} - \nu_2)}{(\nu_{\text{probe}} - \nu_2)^2 + (\Gamma_2/2)^2} \right] = K_{\text{PD}} \pi I_0 l n r_{\text{e}} c f \cdot \frac{\gamma_{\text{e}} P_{\text{e}}^z R_{\text{tot}}^{\text{e}}}{R_{\text{tot}}^{\text{e}2} + \gamma_{\text{e}}^2 (L_z + \delta B_z)^2} (\tilde{\Omega}_y + b) \cdot \left[\frac{(\nu_{\text{probe}} - \nu_1)}{(\nu_{\text{probe}} - \nu_1)^2 + (\Gamma_1/2)^2} + \frac{(\nu_{\text{probe}} - \nu_2)}{(\nu_{\text{probe}} - \nu_2)^2 + (\Gamma_2/2)^2} \right] \tag{2}$$

式中: K_{PD}为光电二极管与前置放大器的放大倍数, 光电二极管将光强信号转换为电流信号, 然后通过前置放大器转换为电压信号, 仿真中可取 1×10^5 V/W; I_0为入射气室时的检测光初始光强, 光路中通过液晶相位延迟器组成光强稳定闭环控制模块以确保I_0的稳定, 有效抑制检测光强波动对背景信号的影响。

通过推导出的式 (2) 可以分析出, 检测光引入惯性测量输出信号的误差主要可分为 2 类: 影响刻度系数稳定性的非直接作用于原子的误差, 以及直接作用于电子自旋而产生的电子极化率误差。由于光路光学器件及玻璃气室偏振性能的非理想性[19], 使检测光并非理想线偏振, 而是具有一定椭圆偏振度s_{m}的偏振光, 本文的实验装置中, 入射气室前检测光的椭偏度主要取决于偏振分光棱镜的偏振性能, 使用偏振分析仪测得其椭偏度一般约为 0.02°; 入射气室后高温下的玻璃气室对检测光有一定的退偏作用, 使其椭偏度一般约为 0.1°。非理想线偏振检测光作用于原子自旋会引起横向抽运效果和横向光频移虚拟磁场, 从而使电子自旋极化率在横向的投影产生误差$\delta P_{\text{probe}}^{\text{e}}$。因此, 可将检测系统相关的 SERF 原子自旋惯性测量输出信号误差模型简化为

$$S_{\text{output}} = K_{\text{g}} (I_0 + \delta I_0)(P_x^{\text{e}} + \delta P_{\text{probe}}^{\text{e}})[K(\nu_{\text{probe}}) + \delta K_{\text{wave}}] \tag{3}$$

式中: $K_{\text{g}} = K_{\text{PD}} \pi l n r_{\text{e}} c f$为与检测系统无关的惯性测量刻度系数项; δK_{wave}为检测光频率波动引起的刻度系数误差。仅考虑检测系统引入的误差影响时, 导致电子自旋横向极化率误差$\delta P_{\text{probe}}^{\text{e}}$的横向抽运率可表示为[20]

$$R_{\text{m}} = \frac{I_0 \sigma(\nu_{\text{probe}})}{h \nu_{\text{probe}}} \tag{4}$$

式中: h为普朗克常数; $\sigma(\nu_{\text{probe}})$为频率相关的吸收截面积, 可以看出$R_{\text{m}}$受检测光光强和频率的共同影响。横向电子自旋极化率的稳态解中, 与R_{m}相关项的表达式为$P_x^{\text{e}}(R_{\text{m}}) = R_{\text{tot}}^{\text{e}} R_{\text{m}} s_{\text{m}} / (R_{\text{tot}}^{\text{e}2} + \gamma_{\text{e}}^2 (L_z + \delta B_z)^2)$, 当$s_{\text{m}} \neq 0$时, R_{m}产生的影响则不为 0, 本文实验系统中s_{m}可近似视为不为 0 的常值, 因此, 通过减小R_{m}可以抑制该误差的影响。同时考虑检测光对碱金属原子的 D1 线和 D2 线总的能级分裂作用, 则检测光引起的横向光频移可表示为

$$L_{\text{m}} = \frac{\Phi(\nu_{\text{probe}}) r_{\text{e}} c f}{A \gamma_{\text{e}}} \cdot \left(\frac{-(\nu_{\text{probe}} - \nu_1)}{(\nu_{\text{probe}} - \nu_1)^2 + (\Gamma_1/2)^2} + \frac{2(\nu_{\text{probe}} - \nu_2)}{(\nu_{\text{probe}} - \nu_2)^2 + (\Gamma_2/2)^2} \right) s_{\text{m}} \tag{5}$$

式中: $\Phi(\nu_{\text{probe}})$为单模激光器的光子通量; A为光斑面积, 光强的大小决定了$\Phi(\nu_{\text{probe}})/A$的大小。当检测光椭圆偏振度$s_{\text{m}}$不为 0 时, 横向光频移则不为 0, 并且在光斑大小不变的情况下, 光功率和频率都对光频移产生影响。由式 (4) 和式 (5) 可知, 检测光频率对R_{m}和L_{m}的影响为非线性, 而光强对其的影响呈线性并且正相关, 因此, 合理的减小检测光功率能直接抑制R_{m}和L_{m}引起的误差。此外, 式 (3) 中δI_0为入射气室的检测光光强本身的波动, 其虽不作用于原子自旋, 但会直接导致惯性测量刻度系数的波动; $K(\nu_{\text{probe}}) = \frac{(\nu_{\text{probe}} - \nu_1)}{(\nu_{\text{probe}} - \nu_1)^2 + (\Gamma_1/2)^2} + \frac{(\nu_{\text{probe}} - \nu_2)}{(\nu_{\text{probe}} - \nu_2)^2 + (\Gamma_2/2)^2}$为与检测光频率相关的刻度系数项, 检测光频率的波动则会引入δK_{wave}, 其影响呈洛伦兹线型。基于 3 个大气压下 K-Rb-^{21}Ne 混合抽运系统对输出信号随检测光波长的变化进行仿真, 如图 2 所示, 可以看出通过合理优化波长可以有效增大惯性测量的刻度系数, 提高输出信号的信噪比。

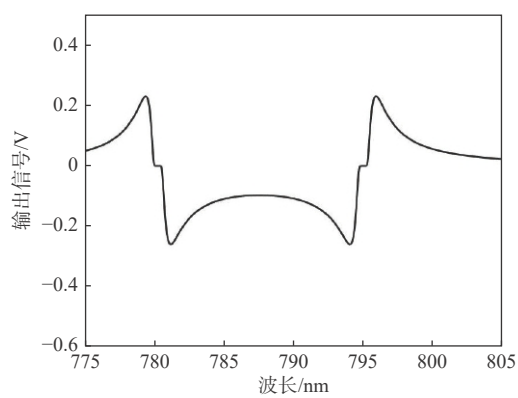


图 2　输出信号随检测光波长变化关系仿真

Fig. 2　Simulation of relationship between output signal and probe light wavelength

因此, 基于所建立的检测系统误差模型, 在保持检测光光斑大小和偏振度不变的情况下, 首先通过优化检测光频率(波长)增大$K(\nu_{\text{probe}})$的值, 而后通过优化检测光功率, 抑制横向抽运效应及横向光频移的作用, 并进一步减小初始光强波动δI_0的影响,

最终提高惯性测量信号的信噪比。

2 实验与结果分析

2.1 检测光频率的优化

通过式 (1) 和图 2 仿真结果可知，由于 Rb 原子的 D1 线为 $5^2S_{1/2}$～$5^2P_{1/2}$ 的跃迁线，其波长在 794.98 nm 附近，当检测光波长与之对应时，将被气室中的碱金属原子吸收，此时检测光不能穿透气室，导致检测灵敏度大幅降低，惯性测量输出信号近似于 0。当检测光波长相较于碱金属原子跃迁线存在一定失谐量时，输出信号随波长的变化反应出惯性测量刻度系数先增大后减小的趋势，因此，本文选取 Rb 原子 D1 线的红移部分进行检测光频率（波长）的优化。基于 2 个大气压的 K-Rb-^{21}Ne 气室得到检测光频率优化实验结果如图 3 所示，检测激光器采用分布反馈半导体激光管 (distributed feedback laser diode, DFB)，其波长可调范围为 794.6 ～796 nm。实验中检测光功率与光斑大小保持不变，由结果可知检测光波长在 795.6 ～796 nm 时，SERF 惯性测量刻度系数基本不再增大，说明其已经达到刻度系数优化的峰值，由图 2 仿真结果可知，若波长继续增大，则刻度系数将呈减小趋势，因此，优化后选择 795.68 nm 波长点，可得到检测激光可调频率范围内，惯性测量最大刻度系数为 26.13 V/((°)·s^{-1})。

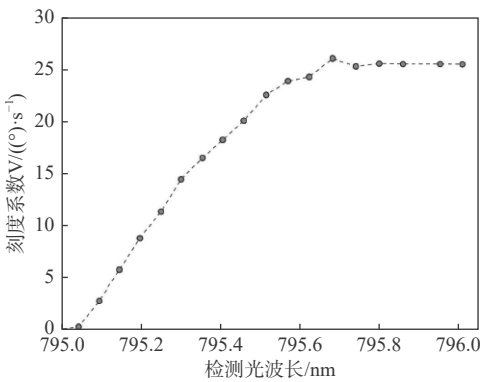


图 3 检测光波长与惯性测量刻度系数的实测关系

Fig. 3 Measurement results of relationship between probe light wavelength and scale coefficients for co-magnetometer

2.2 检测光功率的优化

通过 2.1 节对检测光频率的优化使刻度系数得到有效增大后，保持频率不变，对检测光功率进行优化，抑制横向抽运效应和横向光频移的作用，以及检测背景（即初始入射检测光强 I_0）波动，达到提高惯性测量输出信号信噪比的效果。入射气室的检测光为直径 1 mm 的圆形光斑，在检测激光光斑大小不变的情况下，改变施加于 DFB 激光管的控制电流可使激光光源功率变化，如图 4 所示，对电流与入射气室的检测激光功率进行标定，可以看出两者之间为很好的线性变化关系，通过实测点的拟合可得到关系表达式，则能通过改变控制电流大小直接调节入射气室的检测光功率，并且经过液晶稳光强等光路的衰减，在控制电流允许的范围内，入射气室的检测光功率最大为 2 mW。

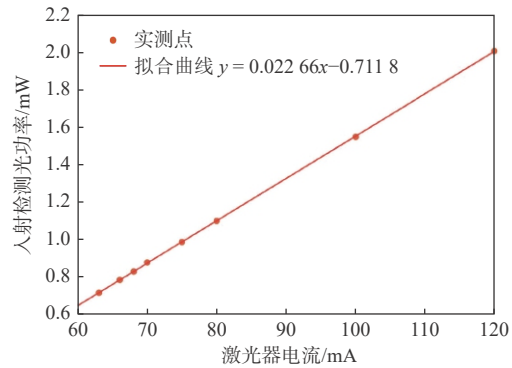


图 4 激光器控制电流与检测光功率关系标定

Fig. 4 Measurement results of relationship between laser control current and probe light power

由式 (3)～式 (5) 可知，入射气室的检测光光强会直接影响横向抽运率 R_m、横向光频移 L_m 和输出信号的检测背景。在光斑大小保持不变的条件下，减小检测光功率可有效抑制 R_m、L_m 与检测背景 δI_0 对原子自旋的影响，然而由于惯性测量输出信号 S_{output} 的强度与检测光初始入射光强 I_0 成正比，减小检测光功率会导致惯性测量刻度系数线性降低，因此，最终输出信号信噪比的提升，需要通过检测光功率的合理优化来实现。图 5 为检测光功率优化惯性测量输出信号 Allan 方差分析，在不同的检测光功率下，对长时的静态 SERF 惯性测量输出信号进行 Allan 方差分析，对比输出信号中低频随机噪声的抑制情况，可分析得到检测光功率优化的效果。采用 Allan 方差方法[21] 可将输出信号中的噪声按时间相关特性进行分类分析，其可描述为

$$\sigma^2(\tau)=\frac{R^2\tau^2}{2}+\frac{K^2\tau}{2}+\frac{2}{\pi}B^2\ln 2+\frac{N^2}{\tau}+\frac{3D^2}{\tau^2} \tag{6}$$

式中：τ 为相关时间；R 为速率斜坡系数；K 为速率随机游走系数；B 为零偏不稳定性系数；N 为角度随机游走系数；D 为量化噪声系数。以提高 SERF 原子自旋惯性测量的稳定性为目标，则应关注信号中的低频随机噪声，其主要表现为零偏不稳定性和速率斜坡噪声，分别对应图 5 中的 A 区域和 B 区域数据段。零偏不稳定性（A 区域）斜率为 0，该系统中其相关时间在 100 s 左右，其值越小、相关时间越长，代表系统信噪比水平越好；速率斜坡（B 区域）为信号中包含的单调趋势的缓慢漂移噪声，抑制其斜率可提升系统的长期稳定性、重复性。由图 5 的实验结果分析可知，当入射气室的检测光功率为 2 mW

时, 信号 Allan 方差的零偏不稳定性最差, 直到光功率降低至 0.78 mW, 输出信号 Allan 方差中的零偏不稳定性最优, 并且速率斜坡噪声也得到了一定程度的抑制。当继续降低检测光功率时, Allan 方差的零偏不稳定性和速率斜坡噪声逐渐增大, 表明由于惯性测量刻度系数随I_0的减小, 信号的信噪比明显降低, 不利于提升系统的长期稳定性。为了更直观的对比检测光功率对信号中低频随机噪声的影响, 图 6 为不同检测光功率下 Allan 方差的零偏不稳定性和速率斜坡系数。可以看出通过检测光功率的优化, 得到最优检测光功率为 0.78 mW, 零偏不稳定性得到了 1.8 倍的抑制, 速率斜坡噪声系数从 0.124 (°)/h^2 减小至 0.041 (°)/h^2, 得到了 3 倍的抑制效果, 因此, 证明检测光功率的优化可有效抑制横向抽运率、横向光频移和光强引起的趋势项漂移, 从而提高输出信号的信噪比及稳定性。

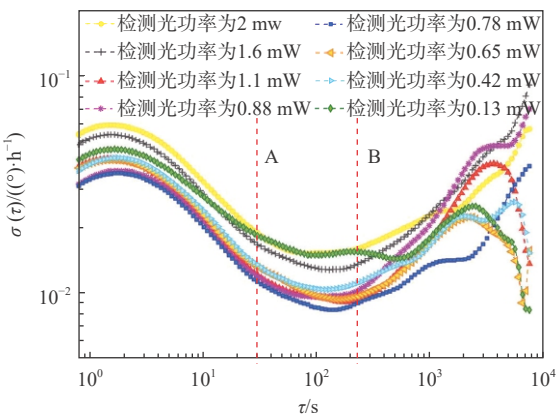


图 5　检测光功率优化惯性测量输出信号 Allan 方差分析

Fig. 5　Allan variance analysis for output signals of the co-magnetometer optimized by probe light power

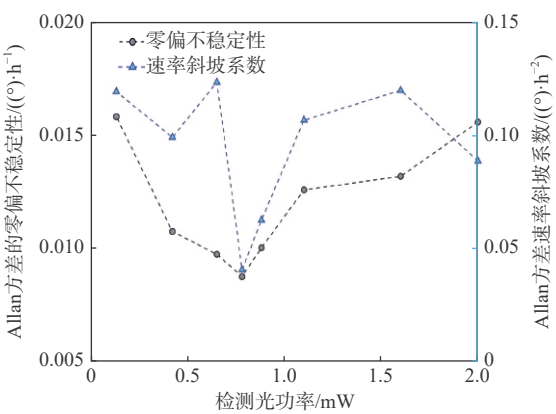


图 6　不同检测光功率下 Allan 方差的零偏不稳定性和速率斜坡系数

Fig. 6　Bias instability and rate ramp coefficient of Allan variance under different probe light power

3　结　论

1) 明确检测光引入输出信号的误差主要可分为 2 类: 影响刻度系数稳定性的非直接作用于原子的误差, 以及直接作用于电子自旋而产生的电子极化率误差。

2) 在检测光椭圆偏振度不为 0 的情况下, 检测光功率和频率都对横向抽运效应和横向光频移产生影响, 其中检测光频率的影响为非线性, 而光功率对其的影响呈线性并且正相关。

3) 检测光频率对惯性测量的刻度系数的影响呈现洛伦兹线形, 因此, 优化检测光频率可有效增大惯性测量刻度系数, 并且存在最优失谐频率点。

4) 信号本底噪声水平主要取决于检测光功率的噪声水平, 同时由于检测光为非理想线偏振, 在检测光光斑大小不变的情况下, 光功率的增大虽然会线性增大刻度系数, 但还会导致横向抽运率和横向光频移的增大, 从而引起测量误差, 因此, 优化检测光功率可有效抑制惯性测量信号的低频随机噪声。

5) 在误差模型建立的基础上, 对关键参数进行优化, 最终输出信号中的零偏不稳定性抑制了 1.8 倍, 速率斜坡噪声系数则得到了 3 倍的抑制, 有效提高了惯性测量的信噪比及稳定性。

参考文献 (References)

[1] 郭雷, 房建成. 导航制导与传感技术研究领域若干问题的思考与展望[J]. 中国科学:信息科学, 2017, 47(9): 1198-1208.
GUO L, FANG J C. Recent prospects on some problems of navigation guidance and sensing technology[J]. Scientia Sinica (Informationis), 2017, 47(9): 1198-1208(in Chinese).

[2] 薛连莉, 陈少春, 陈效真. 2017年国外惯性技术发展与回顾[J]. 导航与控制, 2018, 17(2): 1-9.
XUE L L, CHEN S C, CHEN X Z. Development and review of foreign inertial technology in 2017[J]. Navigation and Control, 2018, 17(2): 1-9(in Chinese).

[3] HU Z X, GALLACHER B J. Effects of nonlinearity on the angular drift error of an electrostatic MEMS rate integrating gyroscope[J]. IEEE Sensors Journal, 2019, 19(22): 10271-10280.

[4] LEFÈVRE H C. The fiber-optic gyroscope[M]. Boston: Artech House, 1993.

[5] LEE B. Review of the present status of optical fiber sensors[J]. Optical Fiber Technology, 2003, 9(2): 57-79.

[6] KETTERLE W. Nobel lecture: When atoms behave as waves: Bose-Einstein condensation and the atom laser[J]. Reviews of Modern Physics, 2002, 74(4): 1131-1151.

[7] CHU S. Cold atoms and quantum control[J]. Nature, 2002, 416(6877): 206-210.

[8] GUSTAVSON T L, BOUYER P, KASEVICH M A. Precision rotation measurements with an atom interferometer gyroscope[J]. Physical Review Letters, 1997, 78(11): 2046-2049.

[9] WOODMAN K F, FRANKS P W, RICHARDS M D. The nuclear magnetic resonance gyroscope: A review[J]. Journal of Navigation, 1987, 40(3): 366-384.

[10] KORNACK T W, GHOSH R K, ROMALIS M V. Nuclear spin

gyroscope based on an atomic comagnetometer[J]. Physical Review Letters, 2005, 95(23): 230801.

[11] ALLRED J C, LYMAN R N, KORNACK T W, et al. High-sensitivity atomic magnetometer unaffected by spin-exchange relaxation[J]. Physical Review Letters, 2002, 89(13): 130801.

[12] SMICIKLAS M A, BROWN J M, CHEUK L W, et al. New test of local Lorentz invariance using a ^{21}Ne-Rb-K Comagnetometer[J]. Physical Review Letters, 2011, 107(17): 171604.

[13] SMICIKLAS M A, ROMALIS M V. Test of Lorentz invariance with Rb-^{21}Ne co-magnetometer at the south pole[C]//Proceedings of the Sixth Meeting on CPT and Lorentz Symmetry. Bloomington: World Scientific Publishing House, 2014: 95-98.

[14] ROMALIS M V, KORNACK T. Chip-scale combinatorial atomic navigator (C-SCAN) low drift nuclear spin gyroscope: 125560431[R]. Princeton: Princeton University, 2018.

[15] LIMES M E, SHENG D, ROMALIS M V. 3He-129Xe comagnetometery using 87Rb detection and decoupling[J]. Physical Review Letters, 2018, 120(3): 033401.

[16] BUDKER D, GAWLIK W, KIMBALL D F, et al. Resonant nonlinear magneto-optical effects in atoms[J]. Reviews of Modern Physics, 2002, 74(4): 1153-1201.

[17] XING L, ZHAI Y Y, FAN W F, et al. Miniaturized optical rotation detection system based on liquid crystal variable retarder in a K-Rb-^{21}Ne gyroscope[J]. Optics Express, 2019, 27(26): 38061.

[18] VASILAKIS G. Precision measurements of spin interactions with high density atomic vapors[D]. Princeton: Princeton University, 2011.

[19] CHOUDHURI A R. Classical electrodynamics[J]. Current Science, 2015, 109(3): 632-633.

[20] KORNACK T W. A test of CPT and Lorentz symmetry using a K-3He co-magnetometer[D]. Princeton: Princeton University, 2005.

[21] 张桂才. 光纤陀螺原理与技术[M]. 北京: 国防工业出版社, 2008.
ZHANG G C. The principles and technologies of fiber-optic gyroscope[M]. Beijing: National Defense Industry Press, 2008 (in Chinese).

Error analysis and suppression of probe system for SERF atomic spin co-magnetometer

XING Li[1], QUAN Wei[2, 3], SONG Tianxiao[4], CAI Qingzhong[3, *], YE Wen[5]

(1. Division of Thermophysics and Process Measurements, National Institute of Metrology, Beijing 100029, China;
2. Research Institute for Frontier Science, Beihang University, Beijing 100191, China;
3. School of Instrumentation Science and Opto-electronics Engineering, Beihang University, Beijing 100191, China;
4. Beijing Institute of Space Launch Technology, Beijing 100076, China;
5. Division of Mechanics and Acoustic Metrology, National Institute of Metrology, Beijing 100029, China)

Abstract: The performance of the probe system is the key factor to determine the sensitivity and stability limit of the spin-exchange relaxation-free (SERF) atomic spin co-magnetometer for inertial measurement. In order to suppress the low-frequency random noise in SERF auto spin co-magnetometer, the error mechanism model for the probe system is established based on the steady-state solution of transverse electron spin polarization and optical rotation angle. The main factors affecting the output signal of the probe system are clarified. According to model analysis, the initial probe light intensity incident on the vapor cell as the signal background directly causes the fluctuation of the scale coefficient rather than the electron spins. In addition, the non-ideal linear polarization of probe light affects the electron spins in a transverse pumping manner and causes a light shift, both of which can cause measurement error. Aiming at the main parameters affecting the noise, the optimization path has been proposed. The probe light frequency has been first optimized to increase the scale coefficient of inertial measurement. Then the transverse pumping rate, light shift, and background fluctuation have been reduced by optimizing the probe light intensity. According to the analysis of Allan variance, the bias instability of SERF auto spin co-magnetometer is suppressed by 1.8 times, and the coefficients of rate ramp are reduced from 0.124 (°)/h^2 to 0.041 (°)/h^2. Therefore, the effect of reducing the low-frequency random noise in the output signal is achieved.

Keywords: spin-exchange relaxation-free; inertial measurement; probe system; error mechanism; transverse pumping effect; transverse light-shift; low-frequency random noise

Received: 2021-11-16; **Accepted**: 2021-12-27; **Published Online**: 2022-02-15 11:01

URL: kns.cnki.net/kcms/detail/11.2625.V.20220214.1617.008.html

Foundation items: National Natural Science Foundation of China (61803015); China Postdoctoral Science Foundation (2021M703049)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0634

基于脊背特征的发动机低转速特性扩展方法

王佳美，郭迎清*，于华锋
(西北工业大学 动力与能源学院，西安 710072)

摘 要：旋转部件在高空低转速时，其工作状态受来流的吹动作用可能会发生变化，此时压气机处在特殊的“搅拌机”或“涡轮”工作状态，使得发动机的动态计算中效率插值出现不连续的问题。为解决此问题，采用美国国家航空航天局（NASA）和通用电气公司（GE）联合开发的针对旋转部件特性转化的脊背特征方法，通过分析低转速下旋转部件脊背特征及非脊背特征的变化趋势，提出基于脊背特征的旋转部件低转速范围特性的扩展方法，并有效规避了效率特性在低转速下插值的失效。以某型军用涡扇发动机为例，计算其处于不同飞行条件下的发动机风车工作状况，结果表明：所提方法能够反映出低转速下压气机压比小于1的特殊工作状态，且不同飞行条件下的风车特性计算合理。

关 键 词：低转速；效率；不连续；脊背点；非脊背点；风车
中图分类号：V235.13; TB553
文献标志码：A **文章编号**：1001-5965(2023)09-2351-10

风车运行[1]是涡轮基组合循环(turbine based combined cycle, TBCC)发动机的一种特有工作状态，TBCC发动机在发生正向转级及逆向转级时，其中的涡轮发动机会先后经历高空熄火冷却和再次点火启动的过程，在这2种情况下，涡轮发动机均会在一段时间内处于风车运行状态。由于风车下发动机燃烧室熄火，不产生动力，压气机部件直接在来流的吹动作用下低转速运行，使得压气机的抽吸和做功能力大大减弱，流量、压比减小，受压比减小的影响，压气机从前至后流量系数的减小幅度逐渐放缓。当工作转速下降过多时，压气机的前几级工作在正攻角下，而后几级工作在负攻角下，此时，压气机的前几级仍然对来流起到减速增压的作用，后几级则起到膨胀加速的作用[2]，转速越低，压气机中表现为涡轮状态的工作级越多，最终导致压气机整体呈现出涡轮态，涡轮态下压气机的出口总温和总压均小于进口。在压气机从压气机的工作状态转为涡轮的过程中，受摩擦损失及叶型损失的影响，会在一段时间内出现气体的动能损耗为热能而使得压气机出口温度升高的状态，即搅拌机状态[3]，对应压气机的出口总温大于进口总温，出口总压则小于进口总压。压气机由搅拌机状态转至涡轮状态的过程中，在实际耗功为0的这一工作点下效率会出现数值奇点问题，导致发动机低转速瞬态运行下，效率的变化是不连续的[4-5]。在利用部件法建模的发动机整机计算中，只有建立起能够有效反映低转速下压气机部件的工作状态变化并且消除不连续参数的部件特性，才可以正确模拟发动机处于风车下的工作过程。

目前国内外常见的对旋转部件低转速的特性扩展方法主要包括以下几种：①基于相似理论的指数外推方法[6]，该方法将低转速下的流体视为不可压缩流体，根据其速度三角形相似，得到相似工作点的流量、压比及效率的指数关系，并通过适当改

收稿日期：2021-10-26；**录用日期**：2022-01-14；**网络出版时间**：2022-02-21 17:18
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220221.1524.001.html
基金项目：国家科技重大专项基金 (J2019-Ⅴ-0003)
* **通信作者**. E-mail：yqguo@nwpu.edu.cn

引用格式：王佳美，郭迎清，于华锋. 基于脊背特征的发动机低转速特性扩展方法 [J]. 北京航空航天大学学报，2023，49（9）：2351-2360. WANG J M，GUO Y Q，YU H F. Extension method of engine low speed characteristics based on backbone features [J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（9）：2351-2360（in Chinese）.

变指数来反映气流的可压缩性，从而由已有的换算转速线外推得到低换算转速下的特性。需要指出的是当对效率进行外推时，由于效率在压比小于1之后会出现不连续的问题，此时效率的定义已经不再适用[7]，为了克服效率的不连续，通常利用扭矩[8-9]这一参数代替效率进行外推。由于指数外推方法仅利用了最低2条转速线的指数信息，并只借助单独一条转速线往下延伸，因此，在低转速特性的扩展上难免会出现较大的误差，这种数学意义上的外推也无法对低转速下旋转部件的特殊状态作出合理解释。②根据旋转部件处于“锁定”状态下的零转速特性[10]和已知转速线之间进行内插[11]。其中，零转速特性需要在已知部件几何参数的前提下，由计算流体力学(computational fluid dynamics, CFD)仿真计算得到，或者根据部件试验特性得到。相比指数外推的方法而言，基于CFD仿真[12]或试验[13]得到的零转速特性为特性扩展提供了更为准确的插值基点，但对叶栅的几何参数信息及试验条件要求较高，较难获得。③借助支持向量机等人工智能方法实现从高转速区域特性关系至低转速区域的回归预测[14]，人工智能的方法可以有效地利用所有的已知特性线数据，避免了指数外推中选择单一参考转速线产生的依赖性误差，在逼近能力上大大优于指数外推法，但该方法也只能在一定的低转速范围内符合物理规律，无法外推至压气机的特殊工作状态。

为了从物理意义上了解旋转部件的工作特性，从而给出更能代表旋转部件实际工作状况的特性趋势，本文基于1984年美国国家航空航天局(National Aeronautics and Space Administration, NASA)和通用电气公司(General Electric Company, GE)联合开发的针对旋转部件特性转化的脊背特征方法[15]，从旋转部件基元级的速度三角形入手，对低转速下基元级的脊背特征及非脊背特征进行分析，从而实现了旋转部件低转速区域的特性扩展。本文方法可以有效地解释及模拟旋转部件中低转速下的状态变化。

1 脊背特征方法的建模机理

低转速下旋转部件发生状态变化时，效率变化的不连续使得传统的效率插值不再适用，此时需要以能代替效率的连续变化参数来反映部件特性[16]。脊背特征方法中借助损失这一特性来表征效率变化，损失的定义如下：

$$L=|\Delta H_{\mathrm{id}}-\Delta H| \tag{1}$$

式中：ΔH_{id}为压气机的理想耗功或涡轮的理想做功；ΔH为压气机的实际耗功或涡轮的实际做功。在全工况工作条件下，旋转部件的损失都是有限的正值且变化连续，通过将效率的除式定义转化为损失的减式定义，可以有效地消除效率定义式中分母近0出现的奇点问题，进而实现全转速插值。

本文方法基于旋转部件的基元级速度三角形，从来流处于不同攻角下的损失变化出发，在流量系数、功系数及损失的概念下对部件的特性进行推导，以最小损失点的变化趋势搭建起特性扩展的基本框架，同时由于最小损失点近似刻画了损失变化的二次极点位置，因此，可借助最小损失点对其他工作点的损失特性进行直线拟合外推，从而实现低转速下的旋转部件特性扩展。

2 压气机脊背特征方法的推导

2.1 压气机脊背特性

轴流式压气机部件中的旋转工作叶轮及其后的一排静子叶片构成了压气机的基元级叶栅，基元级是压气机与外界进行能量交换的基本单元。气流流经压气机基元级时，轴功通过动叶排被传递给气流，使得气流的压力和绝对动能均增加，之后在静叶叶栅中减速增压，气流的动能逐渐转化为静压的提高，静子本身对气流并不做功，压气机基元级转子的速度三角形如图1所示。本文方法以基元级推导得到的特性关系来近似表征部件整体的特性关系。

图1中，C为气流的绝对速度，W为气流的相对速度，U为动叶叶尖的圆周速度，下标1和2分别为动叶的进口和出口截面，下标a和u分别为绝对速度或相对速度的轴向和周向分量，α和β分别为气流的绝对速度和相对速度与圆周方向的夹角。脊背特征方法中压气机的流量系数φ、功系数ψ及损失系数C_L的定义如下：

$$\varphi=\frac{C_{\mathrm{a}}}{U} \tag{2}$$

$$\psi=\frac{\Delta H}{U^2/2} \tag{3}$$

$$C_L=\psi-\psi_{\mathrm{id}}=\frac{\Delta H-\Delta H_{\mathrm{id}}}{U^2/2} \tag{4}$$

式中：ψ_{id}为理想功系数，也称作压力系数。由于$\Delta H=U\Delta C_{\mathrm{u}}$，功系数可被进一步表示为

$$\psi=\frac{U\Delta C_{\mathrm{u}}}{U^2/2}=2-2\varphi\left(\cot\alpha_1+\frac{C_{2\mathrm{a}}}{C_{1\mathrm{a}}}\cot\beta_2\right) \tag{5}$$

假设$\cot\alpha_1+C_{2\mathrm{a}}/C_{1\mathrm{a}}\cdot\cot\beta_2$为一常数，则功系数$\psi$与流量系数$\varphi$成直线关系[17]，此时基元级压力系数曲线上斜率恰好等于功系数和流量系数所成直线斜率的点即为最小损失点，如图2所示。定义最

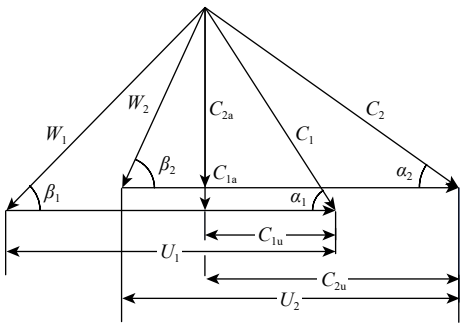


图 1　压气机基元级转子速度三角形

Fig. 1　Velocity triangle of compressor stage

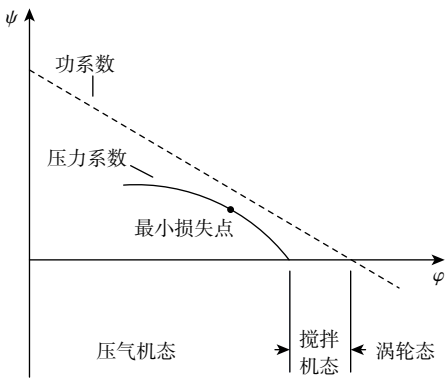


图 2　压气机脊背点的位置

Fig. 2　The position of backbone point of compressor

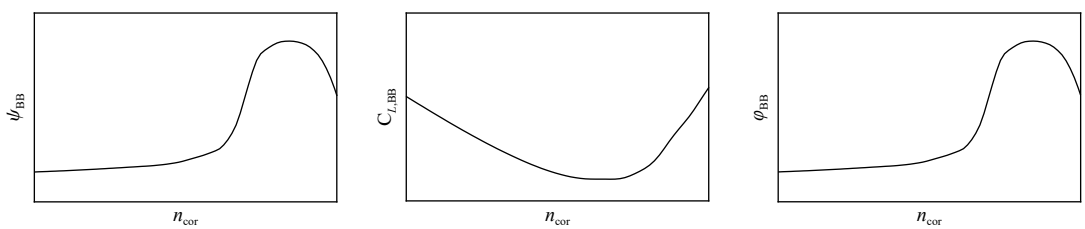
小损失点为脊背点, 压气机的基元级在不同换算转速下都存在一个脊背点。

脊背点为其他工作点的特性变化搭建了基本的位置框架, 在找到已有特性线上的脊背点之后, 需要对脊背点的特性变化规律进行分析, 以实现低转速工作线上脊背点的特性位置外延。图 3 为压气机的脊背点特性, 具体分析如下:

1) 对于脊背点的功系数而言, 其大小取决于扭速和转速的相对变化, 在低转速区域, 由于此时气流可近似为不可压气流, 基元级的速度三角形基本保持相似, 此时功系数变化很小; 随着转速增加, 气流的可压缩性增强, 扭速随之增加, 由此功系数逐渐增加, 并在接近设计转速时达到最大, 之后随着转速的进一步增加, 气流分离导致的损失也进一步增加, 使得功系数反而逐渐下降, 呈现出如图 3(a)所示的变化趋势, 其中, n 为转子转速(百分比物理转速, (实际转速/设计超速)×100%), 下标 BB 和 cor 分别为脊背点和换算参数。

2) 对于脊背点的损失系数而言, 设计点转速在 1 左右, 气流的叶盆分离和叶背分离均会增加, 从而使得设计转速附近的损失最小, 远离设计转速之后损失会逐渐增加, 呈现出如图 3(b)所示的趋势。

3) 对于脊背点的流量系数而言, 其大小取决于轴向速度和转速的相对变化, 低转速时由于速度三角形近似相似, 因此, 流量系数变化很小, 随着转速的增加, 压气机抽吸能力增加, 使得流量系数逐渐增加, 并在接近设计转速时达到最大, 转速进一步增加时, 压气机后边级出现堵塞情况, 流量系数反而减小, 呈现出如图 3(c)所示的趋势。



(a) 功系数　(b) 损失系数　(c) 流量系数

图 3　压气机的脊背点特性

Fig. 3　The backbone characteristics of compressor

2.2　压气机非脊背特性

得到压气机的脊背特性之后, 需要对各个转速线上的非脊背特性进行外扩, 非脊背特性根据非脊背点与脊背点之间的关系得到, 压气机的非脊背点特性如图 4 所示。

1) 对于非脊背点的损失特性而言, 在同一转速下, 进气攻角的变化引起的损失系数的变化与流量系数的变化近似成二次关系, 由此损失系数的变化与功系数的变化成二次关系, 即

$$C_L - C_{L,BB} = c_1(\psi - \psi_{BB})^2 \tag{6}$$

式中: c_1 为损失系数变化与功系数变化二次关系的系数。

又由于脊背点处于极点位置, 则可将各转速下的损失系数和功系数的平方在脊背点的左右两侧分别近似外扩为如图 4(a)所示的直线关系[18]。

2) 对于非脊背点的流量特性而言, 多级轴流式压气机由于存在后面级堵塞的情况, 因此, 各转速下流量系数存在最大值, 这种情况下马赫数 Ma 恒定为 1, 为了更加明确地表示堵塞之后的流量特性, 借助流量函数 $q(Ma)$ 来表征流量特性, 在等转速的

情况下，$q(Ma)$等价于流量系数。由于各转速下的流量系数与功系数成直线关系，也即$q(Ma)$与功系数成如图4(b)所示的直线关系，又由于脊背点的功系数ψ_{BB}及流量系数φ_{BB}在各个转速下都为一定值，此时非脊背点与脊背点之间的流量函数满足以下直线关系：

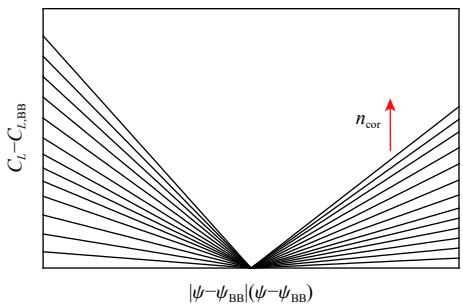
$$q(Ma)=c_2(\psi-\psi_{BB})+q(Ma_{BB}) \tag{7}$$

式中：c_2为流量函数与功系数关系的系数。

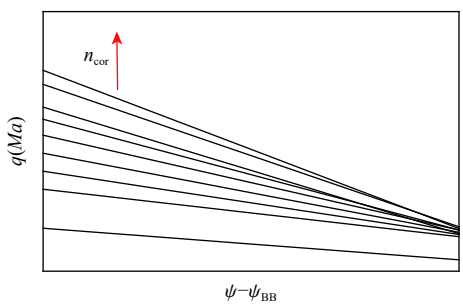
选择脊背点为计算参考点，在已知脊背点流量函数的前提下，非脊背点的流量系数可直接由下式得到：

$$\frac{\varphi}{\varphi_{BB}}=\frac{q(Ma)}{q(Ma_{BB})} \tag{8}$$

根据压气机设计状态的进口参数，将原始特性中已知的压比、效率及换算流量特征参数，在压气机的进口截面转换为相应的流量系数、功系数及损失系数，等换算转速下按照最小损失的含义找寻脊背点(近似于二次曲线找极小值点)，并将脊背点的特性趋势按照如图3所示进行外延，将各换算转速线上的非脊背特性按照如图4所示进行低转速的外延。



(a) 损失特性



(b) 流量特性

图4 压气机的非脊背点特性

Fig. 4 The off-backbone characteristics of compressor

当换算转速减小时，压气机的抽吸能力减弱，此时换算流量会相应减小，同时低转速下进口流量系数增加的方向(也即功系数减小的方向)会缓解压气机部件“前喘后涡”的状态，而使得低转速范围相对高转速不易发生堵塞情况，因此，随着换算转速的减小，非脊背点的流量特性线逐渐变缓，由于流量特性线的截距代表了脊背点的流量函数，在换算转速减小时，截距也会相应地变小；非脊背点的损失变化与流量变化基本一致，因此，在低转速下流量特性变缓的同时，可认为损失变化也是相应变缓的。为了简化非脊背特性的外延操作，可令压气机中非脊背流量特性及非脊背损失特性中的直线斜率随着换算转速的改变线性变化，需要注意的是，由于零转速下静摩擦扭矩的存在，损失特性线及流量特性线斜率的绝对值不可能降低至0。

在得到如图3和图4所示中的变化关系图后，即可在已知压气机部件的换算转速和$\psi-\psi_{BB}$的前提下，得到脊背点的功系数、流量系数及损失系数，进而插值得到当前工作点的流量特性及损失特性。

3 涡轮脊背特征方法的推导

3.1 涡轮脊背特性

在利用脊背特征方法对涡轮部件的特性进行扩展时，同样利用基元级推导得到的特性关系来近似表征部件整体的特性关系。涡轮基元级中转子的速度三角形如图5所示。

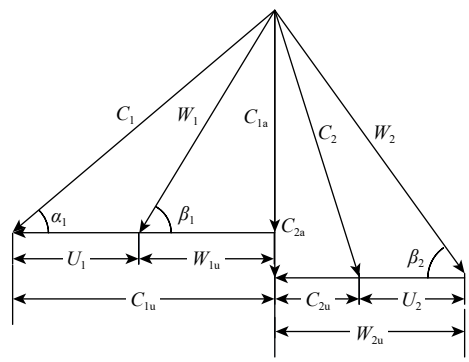


图5 涡轮的基元级转子速度三角形

Fig. 5 Velocity triangle of turbine stage

涡轮中流量系数的定义与压气机相同，功系数定义为

$$\psi=\frac{\Delta H}{2U^2} \tag{9}$$

损失系数定义为

$$C_L=\frac{\Delta H_{id}-\Delta H}{T^*} \tag{10}$$

式(10)可进一步推导为

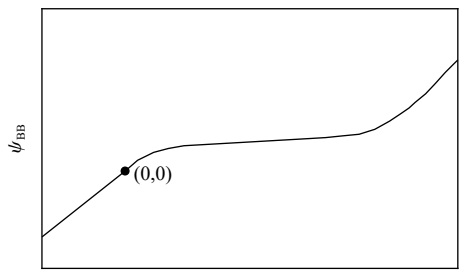
$$C_L=\frac{\Delta H_{id}-\Delta H}{T^*}=\frac{\Delta H/\eta_T-\Delta H}{T^*}=\frac{\Delta H}{T^*}\left(\frac{1}{\eta_T}-1\right) \tag{11}$$

式中：η_T为涡轮部件的效率；T^*为进气总温。根据压气机中定义的脊背点含义，可知涡轮中的脊背点即为等相似功下的最高效率点。涡轮中脊背点的特性如图6所示，具体分析如下：

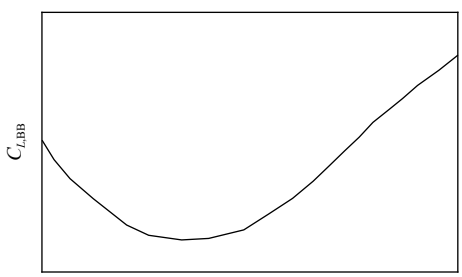
1) 对于脊背点的功系数而言，受涡轮工作状态变化引起的相似功正负性的变化，可知随着相似功的增加，功系数逐渐从负值变为正值，并经过原点。低相似功下，转子叶片效率的迅速变化使得功

系数随着相似功的增加而迅速增加, 在设计点相似功附近, 效率基本保持不变, 此时功系数的增加相对变缓, 相似功进一步增加时, 转速基本保持不变, 因此功系数的变化又逐渐加快, 呈现出如图 6(a)所示的变化趋势。

2) 对于脊背点的损失系数而言, 由于涡轮是否已经发生状态变化, 其实际轴功始终小于理想轴功, 因此, 损失在全相似功范围内均为正值, 在相似功等于 0 的附近, 损失最小, 随着相似功绝对值的增加, 损失也逐渐增加, 呈现出如图 6(b)所示的变化趋势。



(a) 功系数



(b) 损失系数

图 6　涡轮的脊背点特性

Fig. 6　The backbone characteristics of turbine

3.2　涡轮非脊背特性

涡轮非脊背点的特性同样需要根据非脊背点与脊背点之间的关系得到。

1) 对于非脊背点的损失特性而言, 关键在于探讨非脊背点与脊背点之间的效率差值与功系数差值之间的关系, 涡轮基元级的效率定义如下[19]:

$$\eta_T=\frac{\Delta H}{\Delta H_{id}}=\frac{\Delta H}{\Delta H+\Delta H_N+\Delta H_B} \tag{12}$$

式中: ΔH_N 为静叶中产生的损失; ΔH_B 为动叶中的损失, 定义如下:

$$\begin{cases}\Delta H_N=\dfrac{T_{2,id}^*}{T_1}\left(h_1-h_{1,id}\right)=\dfrac{T_{2,id}^*}{T_1}\cdot\dfrac{C_1^2}{2}\left(\dfrac{1}{\eta_N}-1\right)\\ \Delta H_B=\dfrac{T_2^*}{T_2}\left(h_2-h_{2,id}\right)=\\ \qquad\dfrac{T_2^*}{T_2}\left(\dfrac{W_2^2}{2}\left(\dfrac{1}{\eta_B}-1\right)+\dfrac{W_1^2}{2}(1-\cos^2 i)\right)\end{cases} \tag{13}$$

式中: h 为转子静焓; η_N 和 η_B 分别为静叶及动叶在无攻角损失下的效率; T 为转子静温; T_2^*为转子的出口总温; 下温 id 表示理想(等熵)条件下; $W_1^2(1-\cos^2 i)/2$ 为动叶中因为攻角 i 而产生的额外损耗, 来源于文献 [20] 中最佳进气方向垂直方向的动能的全部损失这一定义, 在最佳进气方向为 op 角度时, 攻角损失项可被进一步表示为

$$\frac{W_1^2}{2}(1-\cos^2 i)=\frac{U_1^2}{2}\cos^2\beta_{1,op}\left(\frac{\varphi_1}{\varphi_{1,op}}-1\right)^2 \tag{14}$$

将静叶损失及动叶损失代入基元级的效率表达式, 可以得到以下效率和功系数之间的关系式[21]:

$$\psi\left(\frac{1}{\eta_T}-1\right)=A\left[2\psi+\left(\frac{R_2}{R_1}\right)^2\right]^2+B(\psi-\psi_{op})^2 \tag{15}$$

式中:

$$\begin{cases}A=\dfrac{1}{4}\cdot\left\{\dfrac{T_2^*}{T_2}\left[\left(\dfrac{1}{\eta_B}-1\right)\dfrac{C_{2a}^2}{C_{1a}^2}\cdot\dfrac{1}{\sin^2\beta_2}\right]+\dfrac{T_{2,id}^*}{T_1}\left[\left(\dfrac{1}{\eta_N}-1\right)\dfrac{1}{\sin^2\alpha_1}\right]\right\}\Big/\left(\cot\alpha_1+\dfrac{R_2}{R_1}\cdot\dfrac{C_{2a}}{C_{1a}}\cot\beta_2\right)^2\\ B=\dfrac{T_2^*}{T_2}\cdot\cos^2\beta_{1,op}\Big/\left[2\psi_{op}+\left(\dfrac{R_2^2}{R_1^2}\right)\right]^2\end{cases}$$

其中, R 为转子半径; 将 A、B 视作常数, 对其进行求导, 即可得到一定相似功下脊背点的效率 η_{BB} 及功系数 ψ_{BB}:

$$\begin{cases}\eta_{BB}=\left\{A\left[2\psi_{BB}+\left(R_2/R_1\right)^2\right]^2+B(\psi_{BB}-\psi_{op})^2/\psi_{BB}+1\right\}^{-1}\\ \psi_{BB}=\sqrt{\left(A\left(\dfrac{R_2}{R_1}\right)^4+B\psi_{op}^2\right)\Big/(4A+B)}\end{cases} \tag{16}$$

将式(16)代入式(15), 则可得到非脊背点与脊背点之间的损失关系:

$$\frac{\psi}{\psi_{BB}}\left(\frac{1}{\eta}-\frac{1}{\eta_{BB}}\right)=\psi_{BB}(4A+B)\left(\frac{\psi}{\psi_{BB}}-1\right)^2 \tag{17}$$

图 7 为涡轮的非脊背特性, 在不同的相似功下, A 和 B 为不同的常数, 且脊背点代表了效率的二次极点位置, 因此, 非脊背点与脊背点之间的损失关系可由图 7(a)所示的线性关系给出。

2) 对于非脊背点的流量特性而言, 由于涡轮在导叶喉部处于临界或超临界的情况下会出现堵塞的情况, 此时流量在某一转速下不再随着落压比的增加而增加, 马赫数恒定为 1, 为了表征涡轮发生堵塞之后的流量状况, 利用马赫数代替流量系数对非

脊背点的流量特性进行推导。由涡轮的基元级速度三角形，得到涡轮的单位轴功为

$$\Delta H = U_1 C_{1u} + U_2 C_{2u} = U_1 \left(C_{1u} + \frac{R_2}{R_1} C_{2u} \right) \tag{18}$$

进一步推导得到功系数与流量系数之间的关系式为

$$\psi = \frac{1}{2}\varphi \left(\cot\alpha_1 + \frac{R_2}{R_1} \cdot \frac{C_{2a}}{C_{1a}} \cot\beta_2 \right) - \frac{1}{2} \left(\frac{R_2}{R_1} \right)^2 \tag{19}$$

假设 $\cot\alpha_1 + (R_2/R_1)(C_{2a}/C_{1a})\cot\beta_2 = K_1$ 为一常数，则功系数与流量系数成直线关系。

为了引入马赫数，将流量系数进行改写，在导叶喉部亚临界的情况下，转子进口的绝对气流角 α_1 变化较小，近似于一常数[22]，此时，得到：

$$\varphi = \frac{C_{1a}}{U_1} = \frac{C_1 \sin\alpha_1}{a^*} \cdot \frac{a^*}{U_1} = K_2 \frac{C_1}{a^*} \cdot \frac{1}{n/\sqrt{T^*}} \tag{20}$$

式中：K_2 为一常数。定义滞止马赫数 $Ma^* = C_1/a^*$，a^* 为滞止声速。功系数与流量系数之间的关系可被进一步推导为

$$Ma^* n/\sqrt{T^*} = \frac{1}{K_1 K_2} \left[\left(\frac{60}{2\pi R_1} \right)^2 \frac{\Delta H}{T^*} + \left(\frac{R_2}{R_1} \right)^2 \left(n/\sqrt{T^*} \right)^2 \right] \tag{21}$$

则在等换算转速的情况下，$Ma^* n_{cor}$ 与 $\Delta H/T^*$ 成直线关系，如图 7(b)所示。

与压气机相同，涡轮中非脊背点流量系数的获取也要选择一个参考点，此处选择的是同一换算转速下的堵塞流量点，根据滞止马赫数与马赫数之间的关系：

$$Ma/Ma^* = \sqrt{T^*/T} = \sqrt{1 + (k-1)Ma/2} \tag{22}$$

可知堵塞时 Ma_{bl}^*=0.925 8。将流量函数推导为

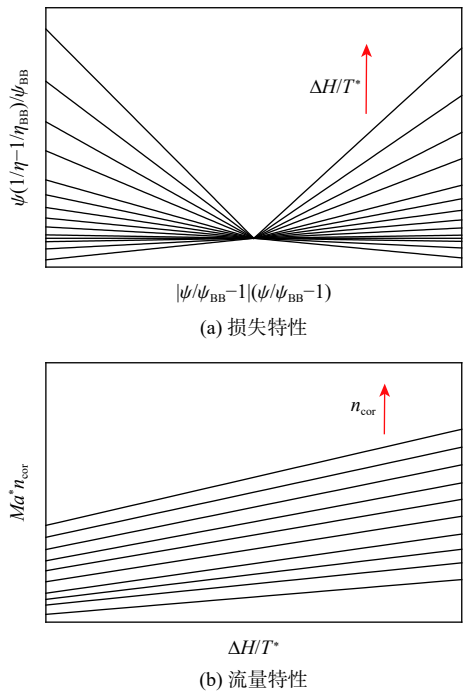
$$q(Ma) = \left(\frac{k+1}{2} \right)^{\frac{k+1}{2(k-1)}} Ma^* \left(1 - \frac{k-1}{2} Ma^{*2} \right)^{\frac{1}{k-1}} \tag{23}$$

式中：k 为比热比常数。在已知堵塞点的换算流量的前提下，非脊背点的换算流量可直接由下式得到：

$$\frac{\dot{m}_{cor}}{\dot{m}_{cor,bl}} = \frac{q(Ma)}{q(Ma_{bl})} = \frac{Ma^* \left(1 - \frac{k-1}{2} Ma^{*2} \right)^{\frac{1}{k-1}}}{Ma_{bl}^* \left(1 - \frac{k-1}{2} Ma_{bl}^{*2} \right)^{\frac{1}{k-1}}} \tag{24}$$

式中：$\dot{m}$为流量；下标 bl 表示堵塞点。涡轮中堵塞点的换算流量满足如图 8 所示的变化趋势，即由于导叶中的总压损失随着换算转速增大而增大进而导致堵塞点的换算流量减小。

根据涡轮设计状态的进口参数，将原始特性中已知的相似功、效率及换算流量特征参数，转换为相应的流量系数、功系数及损失系数，等相似功下按照最小损失的含义找寻脊背点(近似于二次曲线找极小值点)，并将脊背点的特性趋势按照如图 6 所示进行外延，将各换算转速线上的非脊背特性按照如图 7 所示进行低转速的外延。



(a) 损失特性

(b) 流量特性

图 7 涡轮的非脊背点特性

Fig. 7 The off-backbone characteristics of turbine

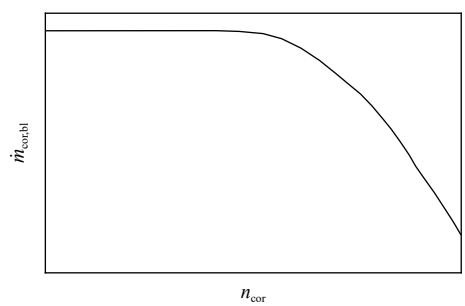


图 8 涡轮堵塞点换算流量的变化趋势

Fig. 8 The trend of maximum value of turbine flow function

非脊背点的损失变化与流量变化基本一致，随着相似功的减小，静子叶片中的压降减小，使得静子出口的速度减小，从而流量减小，因此，非脊背损失特性左右两分支斜率的绝对值也相应的减小；在同一相似功下，换算转速的减小对进气马赫数的影响很小，涡轮的非脊背流量特性主要受换算转速的影响，因此，随着换算转速的减小，非脊背流量特性也逐渐减小，同时，由于低转速下效率变化较快，在固定范围的相似功变化下，导致动叶中落压比的变化较小，因而随着换算转速的减小，非脊背流量特性也是逐渐变缓的。为了简化操作，可将涡轮中的非脊背流量特性及非脊背损失特性的直线斜率做随换算转速或相似功变化的均匀外延，需要注意的

是, 当转速为 0 时, 由于存在静摩擦力矩, 此时流量>0, 而当相似功为 0 时, 脊背点的功系数 ψ_{BB} =0, 由式(17)可知此时非脊背点损失特性的斜率为 0, 当相似功进一步减小至负值时, 由式(25)可知, 由于损失在全相似功范围内恒>0, 此时的$\frac{\psi}{\psi_{BB}}\left(\frac{1}{\eta}-\frac{1}{\eta_{BB}}\right)$<0。

$$\frac{\psi}{\psi_{BB}}\left(\frac{1}{\eta}-\frac{1}{\eta_{BB}}\right)=\frac{\psi}{\psi_{BB}}\left(\frac{\Delta H_{id}}{\Delta H}-\frac{\Delta H_{BB,id}}{\Delta H_{BB}}\right)=\frac{\psi}{\psi_{BB}}\left(\frac{\Delta H+C_L}{\Delta H}-\frac{\Delta H_{BB}+C_{L,BB}}{\Delta H_{BB}}\right)=\frac{\psi}{\psi_{BB}}\frac{1}{\Delta H}\left(C_L-C_{L,BB}\right) \tag{25}$$

利用图 6～图 8 的关系图, 即可在已知涡轮的换算转速和$\psi-\psi_{BB}$的前提下, 得到脊背点的功系数、流量系数及损失系数, 进而插值得到当前工作点的流量特性及损失特性。

4　计算与分析

以某型军用涡扇发动机为例, 利用脊背特征方法对风车减速运行过程及不同高度和马赫数下的稳态风车运行情况进行仿真分析, 由于脊背特征方法中的特性图均是按照换算转速及功系数的偏差$\psi-\psi_{BB}$进行插值构造的, 所以发动机平衡方程中的自变量中也要作相应的更替, 同时由于此时旋转部件直接插值得到的是损失系数 C_L, 可将损失系数转化为效率进行部件出口参数的计算(此时的效率与不连续性无关), 或者直接利用等熵功及实际功直接计算部件的压比及温比。

图 9 中给出了 16 km/2.5*Ma* 飞行条件下发动机直接由 98% 相对高压物理转速熄火的风车减速的性能参数响应, 可以看到, 该飞行条件下受大的来流速度影响, 发动机的推力 F_N 在 98% 高压转速下已经是负值, 熄火后燃烧室不产生动力, 从而高低压转速下降, 风扇和压气机的抽吸能力减弱, 压比 r_p下降, 导致推力进一步减小; 低压部件不仅要受到高压转速下降的影响, 还要受来流冲压阻力的影响, 导致低压转子转速迅速降低, 且下降速度快于高压转子转速, 但是由于内涵的流动阻力相比外涵要大得多, 导致内涵流量的下降比外涵更大, 因此, 涵道比r_{BP}随之增加。由图 9(b)和图 9(f)可知, 随着转速的下降, 风扇的压比逐渐到达小于 1 的区域, 风扇的效率 η_f 也从有限范围 (0,1) 逐步变化至−∞之后再次瞬变至 (1,+∞), 即脊背特征法有效地模拟了压气机部件状态变化时的效率不连续问题, 使得发动机的风车计算能够到达旋转部件的特殊状态。

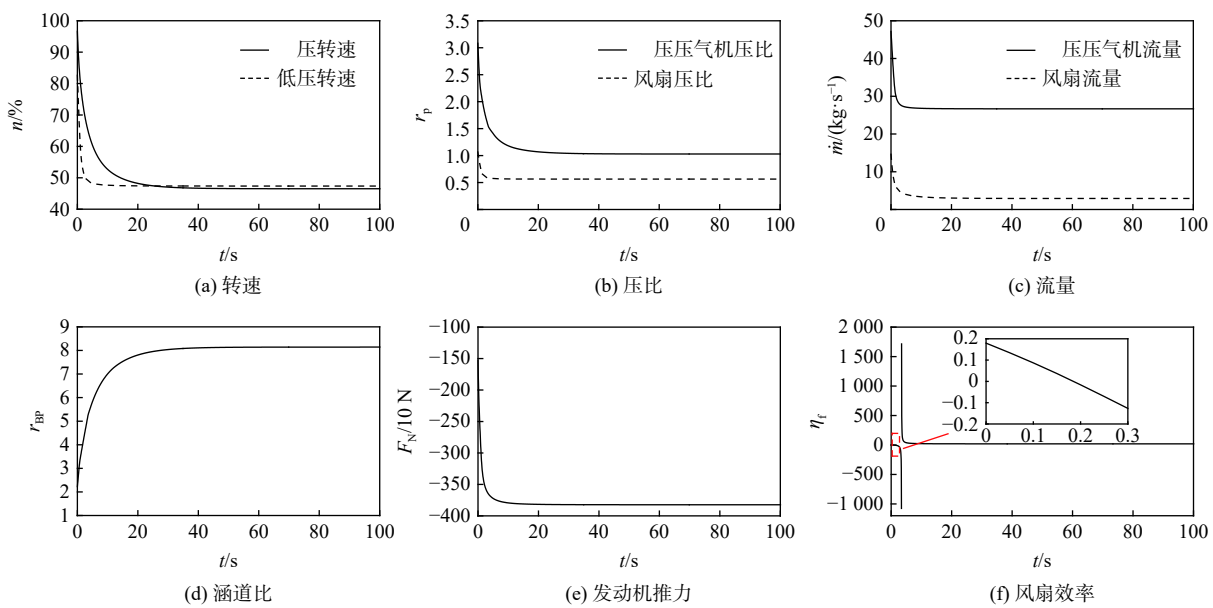


图 9　风车减速时发动机的性能参数响应

Fig. 9　The response of engine performance parameter during deceleration of windmilling

图 10 为不同高度及不同马赫数下的发动机风车特性, 可以看到在同一高度下, 随着马赫数的增加, 来流的冲压作用增强[23], 从而发动机的高压转速 n_c 和低压转速 n_f 增加, 核心机的流通能力受高压转速增加的影响而大幅提高, 使得内涵的流量增加, 涵道比r_{BP}随之下降, 受到来流速度的影响, 推力 F_N 减小。对工作在正常状态的高压压气机而言, 转速越大, 做功能力越强, 因此, 高压压气机的压比$r_{p,c}$随着马赫数的增加而增加, 但由于风扇直接受到来流的吹动作用, 且来流质量流量随着马赫数的增加而增加, 加剧了风扇内气流的负攻角流动[24], 从而使得风扇越来越远离正常工作状态, 压比$r_{p,f}$反而减小。在同一马赫数下, 随着高度的增加, 来流的温度及压力均会下降, 使得气流的冲压

作用减弱，因此，发动机的高低压转速均下降，相对应的高压压气机的压比减小，风扇的压比反而增加，核心机的流通能力降低从而使得涵道比r_{BP}增加。

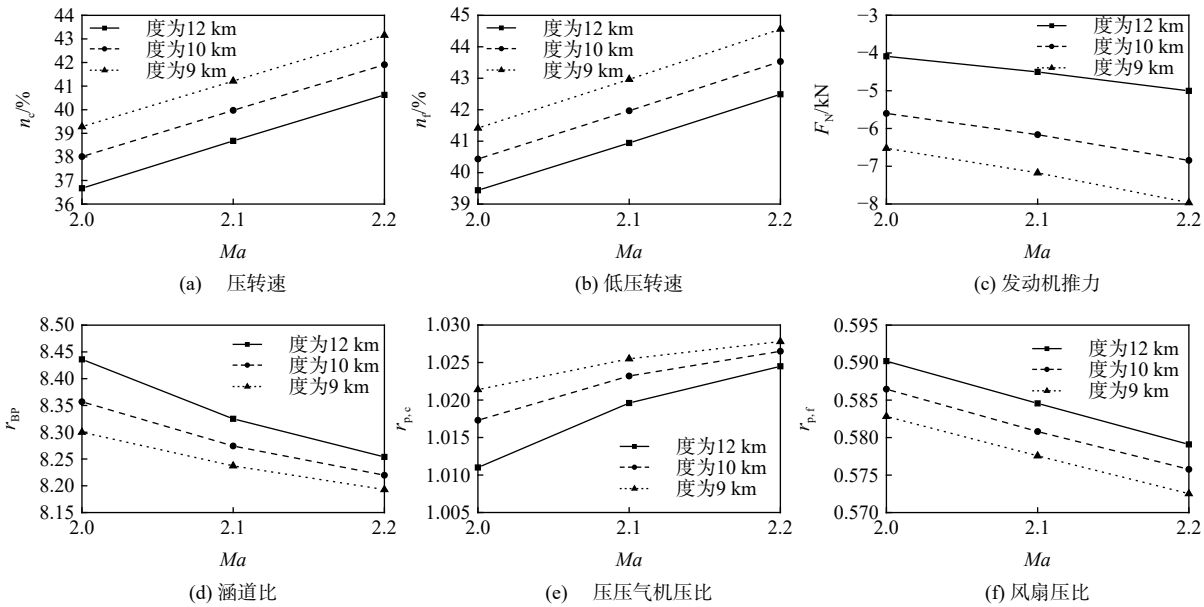


图 10 不同高度及马赫数下的风车特性

Fig. 10 Windmilling characteristics at different altitudes and Mach numbers

5 结 论

1）由于低转速下旋转部件发生状态变化时带来的效率不连续问题，旋转部件特性扩展的关键在于寻找低转速下仍为连续变化的有限值参数来代替效率参数，脊背特征法中部件的损失特性符合这一要求。通过对低转速下基元级的脊背特征及非脊背特征进行分析，能够更为客观地实现旋转部件低转速区域的特性扩展。

2）发动机处于风车减速时，由于核心机失去了动力来源，低压转速同时受到高压转速下降及来流冲压阻力的影响，导致低压转速的下降速度较高压转速更快；内涵的流动阻力相比外涵要大得多，导致内涵流量的下降比外涵更大，涵道比随之增加。

3）对发动机处于不同高度及马赫数下的风车稳态进行计算可知，同一高度下马赫数的增加使得发动机的高低压转速增加，涵道比降低，推力减小。对高压压气机而言，由于其风车稳态时工作在正常状态，压比随着转速的增加而增加，但由于风扇直接受到来流的吹动作用，且来流马赫数增加时来流的质量流量随之增加，加剧了风扇内气流的负攻角流动，风扇的压比反而随马赫数的增加而减小。同一马赫数下，高度增加时来流的冲压作用减弱，此时发动机的性能变化恰好与同一高度下马赫数增加时发动机的性能变化相反。

参考文献（References）

[1] BRAIG W, SCHULTE H, RIEGLER C. Comparative analysis of the windmilling performance of turbojet and tbrbofan engines[J]. Journal of Propulsion and Power, 1999, 15(2): 326-333.

[2] WALKER C L, FENN D B. Investigation of power extraction characteristics and braking requirements of a windmilling turbojet engine[J]. Australian Veterinary Journal, 1952, 48(5): 258-62.

[3] BINDER N, COURTY-AUDREN S K, DUPLAA S, et al. Theoretical analysis of the aerodynamics of low-speed fans in free and load-controlled windmilling operation[J]. Journal of Turbomachinery, 2015, 137(10): 101001.

[4] KURZKE J. How to get component maps for aircraft gas turbine performance calculations[C]//ASME 1996 International Gas Turbine and Aeroengine Congress and Exhibition. Birmingham: American Society of Mechanical Engineers, 1996.

[5] XU S Y, ZHU Z L, LIU Z D, et al. Research on the starting characteristics and control law of two spool turbofan engine[J]. Energy Procedia, 2019, 158: 1765-1771.

[6] 王松岭, 张学镭, 陈海平, 等. 基于相似定律外推压气机通用特性曲线的方法[J]. 动力工程, 2007, 27(2): 169-173.
WANG S L, ZHANG X L, CHEN H P, et al. Method based on similarity laws for extrapolating generalized performance curves of compressors[J]. Journal of Power Engineering, 2007, 27(2): 169-173(in Chinese).

[7] GAUDET S R, DONALD GAUTHIER J E. A simple sub-idle component map extrapolation method[C]//Proceedings of ASME Turbo Expo 2007: Power for Land, Sea, and Air. Montreal: ASME, 2009: 29-37.

[8] JONES G, CURNOCK B. Compressor characteristics in gas turbine performance modelling: 2001-GT-0384 [R]. New Orleans : ASME , 2001.

[9] 王占学, 王永杰, 乔渭阳, 等. 涡扇发动机低转速部件特性扩展和风车状态性能模拟[J]. 推进技术, 2006, 27(2): 146-149.
WANG Z X, WANG Y J, QIAO W Y, et al. Extrapolating com-

ponent maps into the low speed and simulation of windmilling performance of turbofan engine[J]. Journal of Propulsion Technology, 2006, 27(2): 146-149(in Chinese).

[10] HOWARD J. Sub-idle modelling of gas turbines: Altitude relight and windmilling[D]. Cranfield : Cranfield University, 2007.

[11] ZACHOS P K, ASLANIDOU I, PACHIDIS V, et al. A sub-idle compressor characteristic generation method with enhanced physical background[J]. Journal of Engineering for Gas Turbines and Power, 2011, 133(8): 1.

[12] GOTO T, KATO D, OHTA Y, et al. Unsteady flow structure in an axial compressor at windmill condition[C]//Proceedings of ASME Turbo Expo 2014. Düsseldorf: ASME, 2014.

[13] ZACHOS P K, PENGUE F, PACHIDIS V, et al. Flowfield investigation of a compressor cascade at high incidence—part 2: Numerical analysis[C]//Proceedings of ASME Turbo Expo 2009: Power for Land, Sea, and Air. Orlando : ASME, 2010: 355-362.

[14] 饶高, 苏三买, 翟向博. 指数外推法和支持向量机相结合的压气机特性扩展方法[J]. 航空动力学报, 2017, 32(3): 749-755.
RAO G, SU S M, ZHAI X B. Method of compressor characteristic extension combining exponent extrapolation method with support vector machine[J]. Journal of Aerospace Power, 2017, 32(3): 749-755(in Chinese).

[15] CONVERSE G L, GIFFIN R. Extended parametric representation of compressor fans and turbines. Volume 1: CMGEN user's manual: NASA-CR-174645 [R]. Washington D. C. : NASA, 1984.

[16] RIEGLER C, BAUER M, KURZKE J. Some aspects of modelling compressor behavior in gas turbine performance calculations[C]// ASME Turbo Expo 2000: Power for Land, Sea, and Air. Munich: ASME, 2000.

[17] 马文通, 苏明, 余南华. 变几何多级轴流式压气机特性估算[J]. 中国电机工程学报, 2008, 28(11): 72-76.
MA W T, SU M, YU N H. Characteristic estimation method of variable geometry multistage axial-flow compressors[J]. Proceedings of the CSEE, 2008, 28(11): 72-76(in Chinese).

[18] SETHI V, DOULGERIS G, PILIDIS P, et al. The map fitting tool methodology: Gas turbine compressor off-design performance modeling[J]. Journal of Turbomachinery, 2013, 135(6): 061010.

[19] DIXON S L, HALL C A . Fluid mechanics and thermodynamics of turbomachinery[M]. Oxford : Pergamon Press, 1998.

[20] FLAGG E E. Analytical procedure and computer program for determining the off-design performance of axial flow turbines: NASA-CR-710 [R]. Washington D. C. : NASA, 1967.

[21] CONVERSE G L. Extended parametric representation of compressor fans and turbines volume Ⅱ- PART user's manual: NASA-CR-174646 [R]. Washington D. C. : NASA, 1984.

[22] 施洋. 民用大涵道比涡扇发动机全状态性能模型研究[D]. 西安: 西北工业大学, 2017.
SHI Y. A research on full states performance model for civil high bypass turbofan engine[D]. Xi'an: Northwestern Polytechnical University, 2017 (in Chinese).

[23] 高扬, 田晓平, 李秋锋. 基于换算扭矩特性的混排涡扇发动机风车状态性能模拟[J]. 现代机械, 2017(6): 10-13.
GAO Y, TIAN X P, LI Q F. Simulation of windmilling performance for mixed flow turbofan engine based on converted torque characteristics[J]. Modern Machinery, 2017(6): 10-13(in Chinese).

[24] PRASAD D, LORD W K. Internal losses and flow behavior of a turbofan stage at windmill[J]. Journal of Turbomachinery, 2010, 132(3): 1.

Extension method of engine low speed characteristics based on backbone features

WANG Jiamei，GUO Yingqing*，YU Huafeng

(School of Power and Energy，Northwestern Polytechnical University，Xi'an 710072，China)

Abstract：The blowing action of the incoming flow may cause the functioning state of the rotating components to alter when they operate at high altitudes but low speeds. The compressor is in a special "mixer" or "turbine" working state at such a condition, which makes a discontinuous change of efficiency during the dynamic operation of the engine. In order to solve this problem, the method based on the backbone features developed by National Aeronautics and Space Administration (NASA) and General Electric Company (GE) was used to transform the characteristics of rotating parts. It is possible to expand the low-speed range characteristics of rotating parts and successfully prevent the invalidation of the interpolation of efficiency characteristics at low speeds by studying the changing trends of the backbone and off-backbone features at low speeds. Taking a military turbofan engine as an example, the windmilling characteristics under different flight conditions are calculated. The results show that this method can reflect the special working conditions of the compressor with a pressure ratio of less than 1 at low speeds, and the calculation of windmilling characteristics under different flight conditions is reasonable.

Keywords：low speeds; efficiency; discontinuity; backbone; off-backbone; windmilling

Received：2021-10-26；**Accepted**：2022-01-14；**Published Online**：2022-02-21 17：18

URL：kns.cnki.net/kcms/detail/11.2625.V.20220221.1524.001.html

Foundation item：National Science and Technology Major Project of China (J2019-Ⅴ-0003)

* **Corresponding author**. E-mail：yqguo@nwpu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0162

含间隙全动舵面的非线性颤振模式及被动抑制方法

白刘月，吴志刚*，杨超
(北京航空航天大学 航空科学与工程学院，北京 100191)

摘 要：间隙非线性环节广泛存在于飞行器的结构当中，由其引发的极限环振荡（LCO）往往造成结构的疲劳破坏，但目前关于全动舵面非线性颤振机理和被动抑制方法的研究相对较少。针对典型含扭转间隙小展弦比全动舵面开展分析，提出了2种典型的非线性颤振模式，并探索相应的非线性颤振被动抑制方法。基于描述函数法和活塞理论建立舵面的动力学模型，考虑到间隙非线性环节导致刚度降低的效果，计算对比了2种不同扭转刚度下全动舵面的非线性颤振特性，进而提出了2种不同的非线性颤振模式，其中，模式Ⅱ不存在稳定的极限环振荡过程，可以有效抑制低于线性颤振边界的极限环振荡问题；研究了舵面根部弯曲扭转刚度和质量特性对非线性颤振模式和极限环初始动压的影响，提出相应的非线性颤振被动抑制策略。针对所提算例，数值计算结果表明：通过调整舵面根部弯曲扭转刚度或质量特性可以提高极限环初始动压，甚至改变非线性颤振模式，从而达到非线性颤振抑制的目的。

关 键 词：间隙非线性；非线性颤振；全动舵面；极限环振荡；颤振抑制
中图分类号：V215.3
文献标志码：A **文章编号**：1001-5965(2023)09-2361-13

飞行器结构中往往有大量的轴承、铰链、螺栓等支撑连接结构，由于加工制造误差，加上使用磨损、松动等原因，结构间隙不可避免，间隙的大小也具有不确定性。间隙非线性环节往往使得飞行器表现出复杂的非线性颤振特性：极限环振荡（limit cycle oscillations, LCO）、准周期、混沌等现象[1-2]。极限环振荡大致可以分为超临界颤振和亚临界颤振2种[3]。超临界颤振某种意义上提高了颤振速度，在风速大于原线性颤振速度时，系统表现为有限幅值的极限环振荡而非发散；而亚临界颤振则使系统在低于线性颤振速度时便有可能表现出大幅值极限环振荡。

有关舵面间隙非线性颤振问题一直是国内外学者们研究关注的重点。舵面通过作动器机构与舵机连接，舵机内部的间隙[4-5]和连接机构的间隙最终表现为舵面在舵轴弯曲方向或者扭转方向上的间隙，容易诱发极限环振荡。Laurenson和Trn[6]对超声速舵面开展了极限环的计算与分析，采用描述函数法处理非线性环节，得到了极限环振荡幅值随风速的变化规律。Lee和Tron[7]采用同样的处理方式，对CF-18飞机折叠机翼转轴和襟翼转轴处的间隙非线性问题开展研究，发现了极限环的跳跃和滞回等复杂现象。Lee和Kim[8]使用虚拟质量法建模，从时域角度研究了根部含扭转间隙全动舵面的非线性颤振问题，发现系统的响应对初始条件比较敏感，观察到了极限环振荡和混沌现象。Conner[9]、Tang[10]等对典型的二维机翼非线性气动弹性问题开展研究，发现系统分岔点与间隙大小无关，指出频率比低于某个临界值时，系统只存在不稳定极限环分枝，但未进一步研究该临界值的影响因素；Dowell等[11]在研究立方刚度、间隙和摩擦非线性环节对某三维机翼的影响时，发现库伦摩擦可能带

收稿日期：2022-03-18；录用日期：2022-05-20；网络出版时间：2022-05-31 15:21
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220530.2125.005.html
*通信作者. E-mail：wuzhigang@buaa.edu.cn

引用格式：白刘月，吴志刚，杨超. 含间隙全动舵面的非线性颤振模式及被动抑制方法[J]. 北京航空航天大学学报，2023，49(9)：2361-2373. BAI L Y，WU Z G，YANG C. Nonlinear flutter modes and flutter suppression of an all-movable fin with freeplay[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2361-2373 (in Chinese).

来一种对系统有利的极限环特性，使非线性情况下系统发散边界高于线性状况。1999年，Lee等[12]全面总结了有关二维机翼非线性气动弹性系统的研究成果，汇总了间隙、迟滞、立方刚度等非线性环节的研究方法和成果。近期，Panchal和Benaroya[13]针对舵面间隙非线性问题的研究做了充分的总结，Bueno等[14]也对近几十年来舵面结构非线性问题的研究现状进行了综述。

赵永辉和胡海岩[15]针对二维机翼模型建立分段线性运动方程，通过仿真发现了极限环振荡幅值频率的跳跃现象。李道春和向锦武[16]采用时域积分的方法研究了舵面间隙对系统开闭环响应的影响。杨智春等[17]对于带集中非线性环节的机翼气动弹性问题做了全面充分的总结，详细阐述了建模分析方法，讨论总结了相关研究成果。隋鑫等[18]针对二维机翼应用稳定性分析、Hopf分岔理论和数值方法分析了系统的非线性颤振特性，采用多项式拟合的方法处理间隙非光滑特性，从而求解系统的Hopf分岔点，研究了马赫数和间隙值对分岔点的影响。王翔宇等[19]针对机副翼间隙非线性问题求解提出了双准谐波平衡方法，给出极限环振荡的解析解，且计算效率与描述函数法相当。大多数研究以二维机翼为主要对象，缺少对三维对象的研究。

近年来，针对三维全动舵面间隙非线性问题的研究逐渐增多，主要集中在间隙辨识、动力学建模、数值仿真及风洞试验等方面。一类研究对象为含折叠间隙的全动舵面[20-24]，另一类则针对根部含弯曲或扭转间隙全动舵面[25]，研究的重心逐渐由单一非线性问题向复杂非线性问题过渡。韩景龙等[26]利用间隙非线性环节设计了风洞颤振抑制系统，将全动舵面线性颤振转变成极限环振荡，并通过试验进行了验证。Firouz-Abadi等[27]研究了根部具有3个方向间隙的全动舵面非线性气动弹性响应；田玮等[28]研究了间隙非线性和气动非线性对高超声速全动舵面的影响，并对多自由度间隙非线性情况下系统的响应进行数值仿真；李家旭等[29]研究了三维控制舵在舵轴弯曲和扭转2个方向单独或同时存在间隙非线性时系统的动力学响应。

有关非线性颤振抑制方法的研究，大致可以分为主动控制和被动控制2个方面。主动控制从控制律的角度出发，利用系统的状态反馈[30]，通过闭环控制达到抑制颤振的目的[17]。李道春和向锦武[31]基于状态依赖Riccati方程推导了非线性控制律，实现对间隙非线性二维机翼的抑制；贾尚帅和丁千[32]基于微分几何法和二次型最优控制设计了非线性控制器，推迟了分岔速度，但未改变系统的亚临界颤振特性；杨国伟等[33]针对间隙非线性机翼跨声速颤振提出了一种时滞反馈控制额方法，有效实现了颤振主动抑制。从仿真结果来看，主动控制技术可以有效抑制含间隙非线性二维机翼的非线性颤振问题，但是额外的控制律会增加控制系统的复杂性，有待进一步应用于工程。被动控制主要是从调整结构质量、刚度和阻尼角度出发，质量平衡法是一种传统有效的方法，现有的研究[34]主要针对线性颤振的抑制；除此之外，非线性动力吸振器（nonlinear energy sink, NES）也被应用于非线性颤振被动抑制研究[35-37]，取得了较好的效果，但应用于工程仍有一定的挑战。

综上所述，对于舵面间隙非线性颤振问题的研究，早期研究以二维机翼居多，采用数学方法对临界或过临界状态开展分析研究；针对三维全动舵面的研究主要集中在非线性环节辨识、动力学建模及数值仿真方面，但是缺少对这类对象非线性颤振特性规律性的探索。在颤振抑制方面主要是基于理想数学模型，结合现代控制理论进行数学仿真，缺乏实用性，关于被动抑制方法的研究提及较少。

基于此，本文以一个典型的含扭转间隙三维全动舵面为研究对象，通过数值方法研究结构主要参数对其非线性颤振特性的影响，提出了2种典型的非线性颤振模式。2种模式的区别在于是否存在稳定的极限环振荡过程，通过转变模式可以有效抑制极限环振荡。从这个角度出发，研究非线性颤振被动抑制策略，探究全动舵面根部弯曲扭转刚度和质量特性对非线性颤振模式极限环初始动压的影响，从模态频率和振型角度提出了非线性颤振模式的近似评估方法。

1 间隙非线性颤振分析方法

1.1 研究对象

本文所讨论的研究对象为典型含扭转间隙小展弦比全动舵面，图1为全动舵面的简化模型示意图。对于自身刚度较大的舵面，可以将其视为刚体[27-29]，在根部（O点）可以绕x轴和y轴转动，分别模拟舵轴发生弯曲和扭转变形时舵面的运动，O点施加了2个弹簧约束，用于模拟舵面根部的弯曲刚度K_B和扭转刚度K_T。

本文主要考虑全动舵面在扭转方向上存在中心对称间隙的情况，定义半间隙角度大小为δ，对应总间隙大小为2δ。当舵面扭转角为q_T时，对应扭转非线性弹性回复力f_{NL}的表达式为

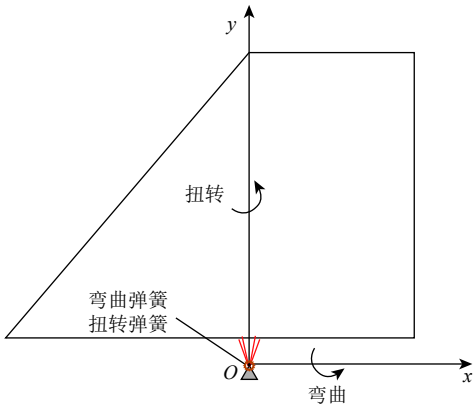


图 1　全动舵面简化模型

Fig. 1　A simplified model of all-movable fin

$$f_{\mathrm{NL}}(q_{\mathrm{T}})=\begin{cases}K_{\mathrm{T}}(q_{\mathrm{T}}-\delta) & q_{\mathrm{T}}>\delta\\ 0 & -\delta\leqslant q_{\mathrm{T}}\leqslant\delta\\ K_{\mathrm{T}}(q_{\mathrm{T}}+\delta) & q_{\mathrm{T}}<-\delta\end{cases}\tag{1}$$

1.2　气动弹性动力学建模

将舵面视为刚体, 采用 2 个弹簧建模的思路基于子结构方法中的分枝模态法[38]。对于 1 个弹性全动舵面, 可以将其划分为 2 个分枝: 分枝 1 为刚性舵面与弹性支撑, 分枝 2 为弹性舵面与刚性支撑, 如图 2 所示。

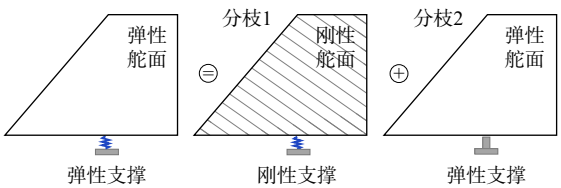


图 2　弹性舵面分枝模态

Fig. 2　Branching modes of an elastic fin

根据模态叠加法, 线性舵面各点位移响应为

$$\boldsymbol{x}=\boldsymbol{\Phi}_{\mathrm{B}}q_{\mathrm{B}}+\boldsymbol{\Phi}_{\mathrm{T}}q_{\mathrm{T}}+\sum_{i=1}^{n}\boldsymbol{\Phi}_{\mathrm{E}}^{i}q_{\mathrm{E}}^{i}\tag{2}$$

式中: q_{B}和q_{T}分别为分枝 1 中刚性舵面绕根部弯曲和扭转两阶假设模态的广义坐标; $\boldsymbol{\Phi}_{\mathrm{B}}$和$\boldsymbol{\Phi}_{\mathrm{T}}$为相应的模态振型函数; q_{E}^{i}为分枝 2 固支弹性舵面自身第 i 阶模态坐标; $\boldsymbol{\Phi}_{\mathrm{E}}^{i}$为对应的模态振型函数; n 为分枝 2 中考虑的模态总个数。

模态坐标下舵面的振动方程为[6, 39]

$$\left[\begin{array}{cc:c}I_{xx} & -I_{xy} & \boldsymbol{R}\\ -I_{xy} & I_{yy} & \\ \hdashline \boldsymbol{R}^{\mathrm{T}} & & \boldsymbol{M}_{\mathrm{E}}\end{array}\right]\left[\begin{array}{c}q_{\mathrm{B}}\\ q_{\mathrm{T}}\\ \hdashline \boldsymbol{q}_{\mathrm{E}}\end{array}\right]+\left[\begin{array}{cc:c}K_{\mathrm{B}} & 0 & 0\\ 0 & K_{\mathrm{T}} & \\ \hdashline 0 & & \boldsymbol{K}_{\mathrm{E}}\end{array}\right]\left[\begin{array}{c}q_{\mathrm{B}}\\ q_{\mathrm{T}}\\ \hdashline \boldsymbol{q}_{\mathrm{E}}\end{array}\right]=\boldsymbol{F}\tag{3}$$

式中: I_{xx}、I_{xy}和I_{yy}为参考坐标系下舵面的转动惯量; $\boldsymbol{R}$为分枝 1 与分枝 2 惯性耦合项; $\boldsymbol{M}_{\mathrm{E}}$和$\boldsymbol{K}_{\mathrm{E}}$分别为分枝 2 的广义质量和广义刚度矩阵; $\boldsymbol{q}_{\mathrm{E}}$为分枝 2 的广义模态坐标向量; $\boldsymbol{F}$为气动力向量。

当舵面自身刚度较大时, 分枝 2 中模态的频率远高于分枝 1, 此时可以忽略分枝 2 高阶弹性模态的影响, 只保留分枝 1, 视舵面为刚体, 最终简化为一个二自由度模型, 即

$$\begin{bmatrix}I_{xx} & -I_{xy}\\ -I_{xy} & I_{yy}\end{bmatrix}\begin{bmatrix}\ddot{q}_{\mathrm{B}}\\ \ddot{q}_{\mathrm{T}}\end{bmatrix}+\begin{bmatrix}K_{\mathrm{B}} & 0\\ 0 & K_{\mathrm{T}}\end{bmatrix}\begin{bmatrix}q_{\mathrm{B}}\\ q_{\mathrm{T}}\end{bmatrix}=\begin{bmatrix}Q_{\mathrm{B}}\\ Q_{\mathrm{T}}\end{bmatrix}\tag{4}$$

式中: Q_{B}和Q_{T}分别为弯曲和扭转广义气动力。

存在扭转间隙非线性环节时, 气动弹性动力学方程为

$$\begin{bmatrix}I_{xx} & -I_{xy}\\ -I_{xy} & I_{yy}\end{bmatrix}\begin{bmatrix}\ddot{q}_{\mathrm{B}}\\ \ddot{q}_{\mathrm{T}}\end{bmatrix}+\begin{bmatrix}K_{\mathrm{B}}q_{\mathrm{B}}\\ f_{\mathrm{NL}}\end{bmatrix}=\begin{bmatrix}Q_{\mathrm{B}}\\ Q_{\mathrm{T}}\end{bmatrix}\tag{5}$$

气动力采用二阶活塞理论[40], 考虑翼型厚度时, 广义气动力时域形式为

$$\boldsymbol{Q}=\boldsymbol{A}_0\boldsymbol{q}+\boldsymbol{A}_1\dot{\boldsymbol{q}}\tag{6}$$

式中:

$$\boldsymbol{A}_0=-\frac{1}{2}\rho V^2\left(\frac{4}{Ma}\right)(\boldsymbol{E}+G\boldsymbol{C})$$

$$\boldsymbol{A}_1=-\frac{1}{2}\rho V\left(\frac{4}{Ma}\right)(\boldsymbol{B}+G\boldsymbol{D})$$

$$\boldsymbol{E}=\begin{bmatrix}\iint_S\frac{\partial\Phi_{\mathrm{B}}}{\partial x}\Phi_{\mathrm{B}}\mathrm{d}x\mathrm{d}y & \iint_S\frac{\partial\Phi_{\mathrm{T}}}{\partial x}\Phi_{\mathrm{B}}\mathrm{d}x\mathrm{d}y\\ \iint_S\frac{\partial\Phi_{\mathrm{B}}}{\partial x}\Phi_{\mathrm{T}}\mathrm{d}x\mathrm{d}y & \iint_S\frac{\partial\Phi_{\mathrm{T}}}{\partial x}\Phi_{\mathrm{T}}\mathrm{d}x\mathrm{d}y\end{bmatrix}$$

$$\boldsymbol{B}=\begin{bmatrix}\iint_S\Phi_{\mathrm{B}}^2\mathrm{d}x\mathrm{d}y & \iint_S\Phi_{\mathrm{T}}\Phi_{\mathrm{B}}\mathrm{d}x\mathrm{d}y\\ \iint_S\Phi_{\mathrm{B}}\Phi_{\mathrm{T}}\mathrm{d}x\mathrm{d}y & \iint_S\Phi_{\mathrm{T}}^2\mathrm{d}x\mathrm{d}y\end{bmatrix}$$

$$\boldsymbol{C}=\begin{bmatrix}\iint_S\frac{\partial H}{\partial x}\cdot\frac{\partial\Phi_{\mathrm{B}}}{\partial x}\Phi_{\mathrm{B}}\mathrm{d}x\mathrm{d}y & \iint_S\frac{\partial H}{\partial x}\cdot\frac{\partial\Phi_{\mathrm{T}}}{\partial x}\Phi_{\mathrm{B}}\mathrm{d}x\mathrm{d}y\\ \iint_S\frac{\partial H}{\partial x}\cdot\frac{\partial\Phi_{\mathrm{B}}}{\partial x}\Phi_{\mathrm{T}}\mathrm{d}x\mathrm{d}y & \iint_S\frac{\partial H}{\partial x}\cdot\frac{\partial\Phi_{\mathrm{T}}}{\partial x}\Phi_{\mathrm{T}}\mathrm{d}x\mathrm{d}y\end{bmatrix}$$

$$\boldsymbol{D}=\begin{bmatrix}\iint_S\frac{\partial H}{\partial x}\Phi_{\mathrm{B}}^2\mathrm{d}x\mathrm{d}y & \iint_S\frac{\partial H}{\partial x}\Phi_{\mathrm{T}}\Phi_{\mathrm{B}}\mathrm{d}x\mathrm{d}y\\ \iint_S\frac{\partial H}{\partial x}\Phi_{\mathrm{B}}\Phi_{\mathrm{T}}\mathrm{d}x\mathrm{d}y & \iint_S\frac{\partial H}{\partial x}\Phi_{\mathrm{T}}^2\mathrm{d}x\mathrm{d}y\end{bmatrix}$$

$$G=Ma\left(\frac{\gamma+1}{2}\right)$$

$$\boldsymbol{q}=\begin{bmatrix}q_{\mathrm{B}}\\ q_{\mathrm{T}}\end{bmatrix}$$

$$\mathbf{Q}=\begin{bmatrix}Q_{\mathrm{B}}\\ Q_{\mathrm{T}}\end{bmatrix}$$

其中: ρ为气体密度; Ma为马赫数; V为来流速度; γ为比热比; H为由舵面中剖面量起的厚度分布函数。

对于频域情况, 引入谐振荡假设$\boldsymbol{q}=\bar{\boldsymbol{q}}\mathrm{e}^{\mathrm{i}\omega t}$, 化简

得到

$$\boldsymbol{Q}=\frac{1}{2}\rho V^2\boldsymbol{A}\bar{\boldsymbol{q}} \tag{7}$$

式中：$\boldsymbol{A}$为广义气动力影响系数矩阵；$\boldsymbol{A}=-(4/Ma)\cdot((\boldsymbol{E}+G\boldsymbol{C})+i(k/b)(\boldsymbol{B}+G\boldsymbol{D}))$；$k=\omega b/V$，$k$为减缩频率，$b$为半参考弦长；$\omega$为谐振荡频率。

1.3 非线性颤振问题的求解

频域常见的线性颤振求解方法主要有v-g法、p-k法等，本文不做赘述。考虑非线性环节时，常规的线性颤振求解方法不再适用，本文采用描述函数法进行处理。描述函数法是一种频域等效线性化方法，认为非线性系统的输出主要由一次谐波分量组成。对于本文这种具有扭转方向中心对称间隙的全动舵面（半间隙大小为δ），其等效扭转刚度$\bar{K}_{\mathrm{T}}$表达式为[41]

$$\frac{\bar{K}_{\mathrm{T}}}{K_{\mathrm{T}}}=\begin{cases}0 & A\leqslant\delta\\ \dfrac{1}{\pi}(\pi-2t_1-\sin 2t_1) & A>\delta\end{cases} \tag{8}$$

式中：$t_1=\arcsin(\delta/A)$，A为舵面扭转假设模态广义坐标极限环振荡幅值。可以看出，等效扭转刚度$\bar{K}_{\mathrm{T}}$是由极限环振荡幅值A与半间隙δ的比值决定的。

图3展示了等效刚度系数$\bar{K}_{\mathrm{T}}/K_{\mathrm{T}}$随$A/\delta$的变化情况。当$A/\delta\leqslant 1$时，舵面在间隙内振动，对应$\bar{K}_{\mathrm{T}}$为0，$A/\delta>1$时，随着$A/\delta$的增大，$\bar{K}_{\mathrm{T}}$逐渐增大，最终趋于线性刚度$K_{\mathrm{T}}$。

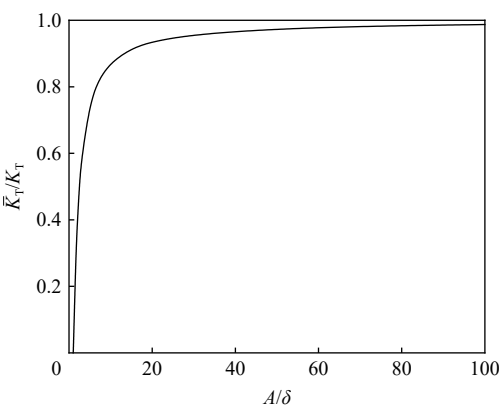


图3 $\bar{K}_{\mathrm{T}}/K_{\mathrm{T}}$随$A/\delta$变化情况

Fig. 3 $\bar{K}_{\mathrm{T}}/K_{\mathrm{T}}$ with variation of A/δ

经过描述函数法的处理，便可以采用常规的线性颤振求解方法对非线性颤振问题进行求解。首先，给定系统发生极限环振动时扭转广义坐标的幅值A；然后，根据式(8)计算出等效扭转刚度$\bar{K}_{\mathrm{T}}$，代替式(4)的K_{T}；其次，求解出扭转刚度为$\bar{K}_{\mathrm{T}}$时的线性颤振速度V_{F}、动压Q_{F}及频率ω_{F}。上述结果表示，来流速度为V_{F}时，系统可能发生扭转广义坐标幅值为A、频率为ω_{F}的极限环振荡。通过给定不同的幅值A可以计算得到一系列极限环状态点，最终描绘出系统的非线性颤振特性。

从时域角度求解系统响应时，本文采用四阶龙格库塔积分方法。线性情况下，系统状态空间方程可以写为

$$\begin{bmatrix}\dot{\boldsymbol{q}}\\ \ddot{\boldsymbol{q}}\end{bmatrix}=\begin{bmatrix}\boldsymbol{0} & \boldsymbol{I}\\ \boldsymbol{M}^{-1}(\boldsymbol{A}_0-\boldsymbol{K}) & \boldsymbol{M}^{-1}\boldsymbol{A}_1\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\\ \dot{\boldsymbol{q}}\end{bmatrix} \tag{9}$$

式中：$\boldsymbol{M}=\begin{bmatrix}I_{xx} & -I_{xy}\\ -I_{xy} & I_{yy}\end{bmatrix}$；$\boldsymbol{K}=\begin{bmatrix}K_{\mathrm{B}} & 0\\ 0 & K_{\mathrm{T}}\end{bmatrix}$。

考虑非线性情况时，需要考虑非线性恢复力f_{NL}的作用，方程变为

$$\begin{bmatrix}\dot{\boldsymbol{q}}\\ \ddot{\boldsymbol{q}}\end{bmatrix}=\begin{bmatrix}\boldsymbol{0} & \boldsymbol{I}\\ \boldsymbol{M}^{-1}(\boldsymbol{A}_0-\hat{\boldsymbol{K}}) & \boldsymbol{M}^{-1}\boldsymbol{A}_1\end{bmatrix}\begin{bmatrix}\boldsymbol{q}\\ \dot{\boldsymbol{q}}\end{bmatrix}+\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{M}^{-1}\boldsymbol{F}_{\mathrm{NL}}\end{bmatrix} \tag{10}$$

式中：$\boldsymbol{F}_{\mathrm{NL}}=\begin{bmatrix}0\\ -f_{\mathrm{NL}}\end{bmatrix}$；$\hat{\boldsymbol{K}}=\begin{bmatrix}K_{\mathrm{B}} & 0\\ 0 & 0\end{bmatrix}$。

获得时域响应数据后，截取稳定极限环振荡区间的数据开展计算分析，从而得到极限环振荡的幅值和频率大小。

2 非线性颤振模式

本节基于某典型小展弦比全动舵面的数值结果来具体说明2种不同的颤振模式。

2.1 模态分析和线性颤振分析

根据研究对象的有限元模型（见图4）给出相关结构参数，如表1所示。表2列出了本文所建立的简化模型与实际有限元模型模态频率对比结果，图5展示了前2阶模态振型。有限元模型考虑了舵面自身弹性，材料为铝合金，简化模型由于视舵面为

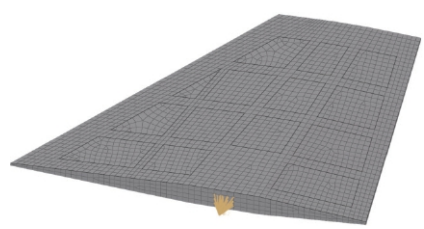


图4 全动舵面有限元模型

Fig. 4 Finite element model of all-movable fin

表1 全动舵面结构参数

Table 1 Structural parameters of all-movable fin

参数	数值
质量m/kg	3.794
质心$[x_{\mathrm{c}},y_{\mathrm{c}}]$/m	[0.012, 0.154]
转动惯量$I_{xx}/(\mathrm{kg\cdot m^2})$	1.275×10^{-1}
转动惯量$I_{xy}/(\mathrm{kg\cdot m^2})$	2.042×10^{-2}
转动惯量$I_{yy}/(\mathrm{kg\cdot m^2})$	4.035×10^{-2}
弯曲弹簧刚度$K_{\mathrm{B}}/(\mathrm{N\cdot m\cdot rad^{-1}})$	9.0×10^{3}
扭转弹簧刚度$K_{\mathrm{T}}/(\mathrm{N\cdot m\cdot rad^{-1}})$	3.0×10^{3}

表 2　模态频率计算结果对比

Table 2　Comparison of modal frequency calculation results　　Hz

模态阶数	简化模型	有限元模型
1	37.77	37.31
2	50.68	50.23
3		407.12
4		517.66
5		573.28

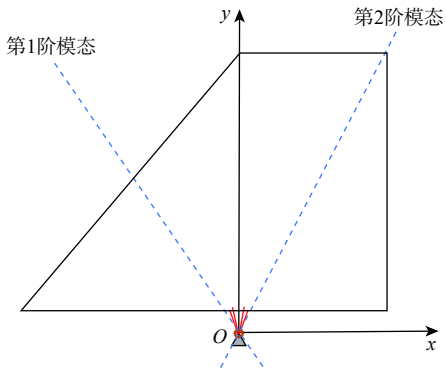


图 5　全动舵面模态振型

Fig. 5　Mode shapes of all-movable fin

刚体，前 2 阶模态频率略有偏高，此外，可以发现有限元模型第 3 阶及之后的模态频率远高于前 2 阶。

采用非匹配的 p-k 法对本文所建立的简化模型进行线性颤振分析。飞行高度设为 18 km，对应大气密度为 0.122 kg/m^3，参考马赫数 3.0。同时基于有限元模型，利用 ZAERO 软件分别计算截取前 2 阶和前 5 阶模态时舵面的线性颤振特性，计算结果对比如表 3 所示。结果表明，对于自身刚度较大的小展弦比全动舵面而言，与舵面自身弹性相关的高阶模态对其颤振特性影响不大，本文所建立的简化模型适用于该类型对象颤振问题的分析研究。

从描述函数法角度来看，舵面根部存在扭转间隙时，其等效扭转刚度降低。因此，在保持弯曲刚度K_B不变的情况下，先对舵面扭转刚度K_T进行变参分析，舵面线性颤振动压Q_F随刚度比$\xi = K_T/K_B$的变化结果如图 6 所示。

本文所给舵面结构参数对应于 A 点，刚度比 $\xi_A = 0.333$，线性颤振动压 $Q_A = 53.9$ kPa；随着扭转刚度K_T的降低，Q_F先减小后增大，记Q_F随K_T变化的极

表 3　线性颤振结果

Table 3　Linear flutter results

计算方法	ω_F/Hz	V_F/(m·s^{-1})	Q_F/kPa
本文方法	44.14	939.78	53.87
ZAERO(前2阶)	42.72	928.83	52.63
ZAERO(前5阶)	42.72	928.85	52.63

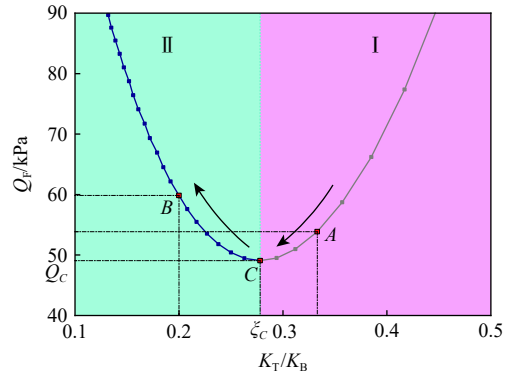


图 6　线性颤振动压随刚度比变化

Fig. 6　Linear flutter dynamic pressure with variation of stiffness ratio

小值点为 C 点，对应颤振动压 $Q_C = 49.1$ kPa，刚度比 $\xi_C = 0.278$。C 点将整个计算区域划分为 2 部分：右侧区域，记作Ⅰ区域，线性颤振动压随扭转刚度的降低先减小后增大；左侧Ⅱ区域，线性颤振动压随着扭转刚度的降低而增大。分别选取Ⅰ区域中 A 点、Ⅱ区域中 B 点 ($Q_B = 59.8$ kPa, $\xi_B = 0.2$) 对全动舵面的非线性颤振特性开展计算分析。

2.2　舵面非线性颤振分析

对于 A 点，计算结果如图 7 所示。先从极限环振荡幅值间隙比 A/δ 频域计算结果来看，来流动压 $Q < Q_{L0}$ 时，舵面始终处于稳定状态，即受到扰动后振动响应收敛；当Q增大到Q_{L0}时，极限环开始出现，记Q_{L0}为极限环初始动压。随着Q的增大，极限环振荡幅值也不断增大，最终发散，记Q_{LF}为极限环发散动压。在来流动压达到Q_{L0}时，系统发生了明显的分岔，出现 2 个极限环分枝，上面的分枝稳定，下面灰色部分的分枝不稳定。从极限环振荡频率变化来看，稳定极限环的频率随着来流动压的增大从ω_{L0}不断增大趋于ω_{LF}，称ω_{L0}和ω_{LF}分别为极限环

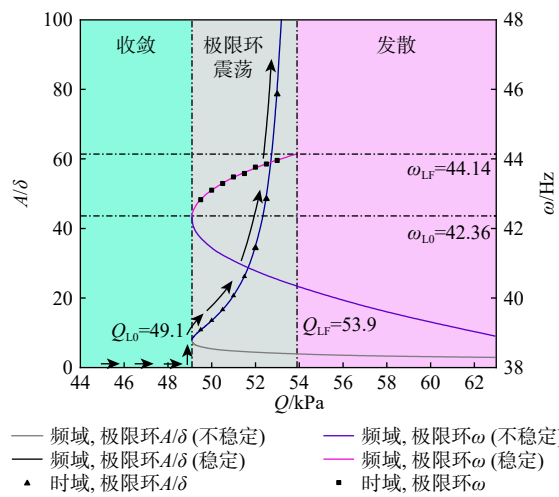
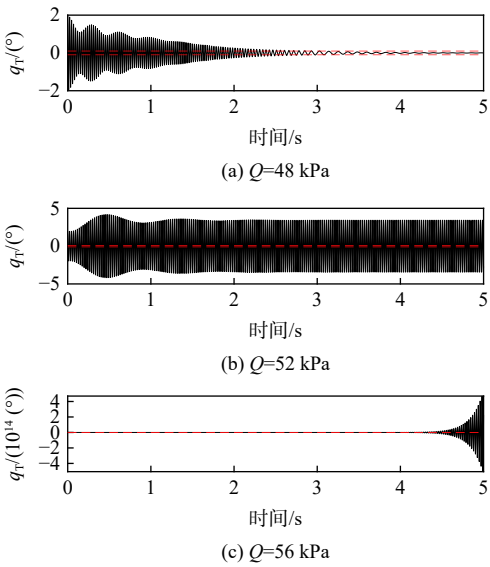


图 7　状态 A 非线性颤振特性

Fig. 7　Nonlinear flutter characteristic of state A

初始频率和发散频率。

以上是频域分析结果，在时域仿真方面，设定半间隙δ为0.1°，考虑非线性系统具有初值敏感性，本文取较大的初始扭转角q_{T0}为2.0°，其他参数均为0。从图7结果来看，时域分析所得极限环振荡幅值间隙比A/δ、极限环振荡频率ω的结果与频域计算结果基本一致。图8给出了3个动压下扭转广义坐标q_T的位移响应结果，红色虚线代表间隙上下边界。从结果来看，动压48 kPa时，系统处于稳定状态，受到扰动后，在气动阻尼作用下振动收敛；动压52 kPa时，出现了稳定的极限环振荡，振荡幅值约为3.44°，振荡幅值间隙比A/δ约为34.4，频率约为43.76 Hz；动压56 kPa时，系统振荡发散。



(a) Q=48 kPa　(b) Q=52 kPa　(c) Q=56 kPa

图8　q_T时域响应结果(状态A)

Fig. 8　Time domain response results of q_T (State A)

对于B点状态，计算结果如图9所示。先从极限环振荡幅值间隙比A/δ频域计算结果来看，来流动压$Q<Q_{LF}$时，舵面始终处于稳定状态，即受到扰动后在气动阻尼主要作用下振动收敛；当Q达到Q_{LF}后，舵面受到一定强度扰动后便会振荡发散，没有稳定极限环出现，只存在不稳定的极限环分枝。从极限环振荡频率变化来看，不稳定极限环振荡频率随着动压的增大而降低。

由于时域分析当中较难观测到不稳定极限环，本文未能给出具体的不稳定极限环振荡幅值和频率的数据。图10给出了2个动压下扭转广义坐标q_T的位移响应结果，分别对应稳定和不稳定2种系统状态。从结果来看，稳定系统受到扰动时响应收敛，不稳定系统受到一定扰动后响应发散，与频域分析结果一致。

实际中，可以通过对应的线性系统颤振动压随刚度比的变化曲线(见图6)及刚度比随极限环振荡

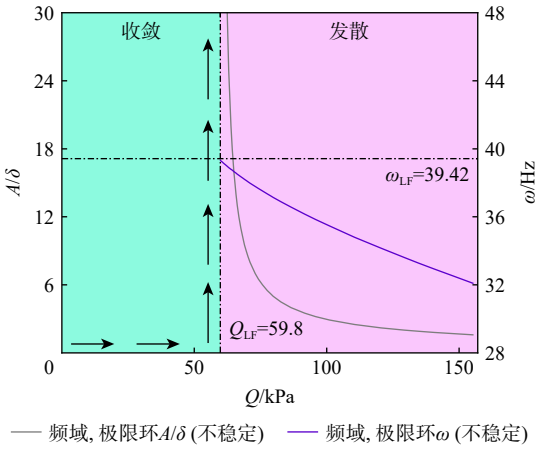
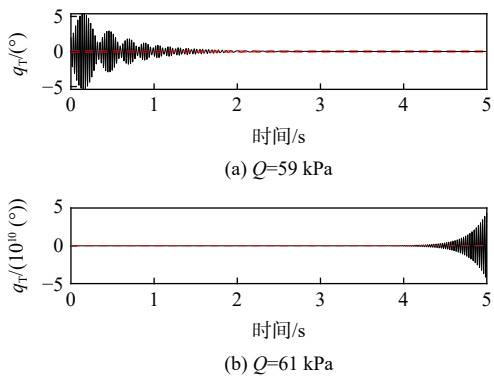


图9　状态B非线性颤振特性

Fig. 9　Nonlinear flutter characteristic of state B



(a) Q=59 kPa　(b) Q=61 kPa

图10　q_T时域响应结果(状态B)

Fig. 10　Time domain response results of q_T (State B)

幅值间隙比的变化曲线(见图3)判断出系统的非线性颤振特性。例如，A点状态，随着扭转广义坐标振荡幅值间隙比A/δ的增大，系统的等效扭转刚度$\bar{K}_T$从零不断增大趋近于线性扭转刚度K_T，对应图3中刚度比$\bar{K}_T/K_B$从零不断增大到ξ_A，对应图6中颤振动压先减小到Q_C后增大到Q_A，最终A点非线性颤振特性表现结果如图7所示。可以发现，A点状态极限环发散动压Q_{LF}对应其线性系统的颤振动压Q_A，极限环初始动压Q_{L0}对应线性系统颤振动压随扭转刚度变化的极小值Q_C，对于Ⅰ区域中的其他点，极限环的初始动压均为Q_{L0}。

2.3　非线性颤振模式总结

从以上结果及讨论可以看出，当舵面弯扭刚度取值分别位于Ⅰ和Ⅱ2个不同区域时，其非线性颤振特性具有很大的区别。本节针对典型含扭转间隙小展弦比全动舵面定义2种不同的非线性颤振模式，如下：

1）Ⅰ型。表现出典型的亚临界颤振特性，低于线性颤振边界存在稳定的极限环分枝。具体表现为：随着来流动压的增大，舵面一开始受到扰动后

振动收敛，达到极限环初始动压（系统分岔点）时才有可能出现极限环振荡，且随着动压的增大，极限环振荡幅值也不断增大，当动压达到原线性系统的颤振动压后，系统振荡发散。

2）Ⅱ型。低于线性颤振边界不存在极限环振荡，仅存在高于线性颤振边界的不稳定极限环分枝。具体表现为：随着来流动压增大，舵面一开始受到扰动后振动收敛，当动压增大到线性系统的颤振动压之后振荡发散，不存在稳定极限环振荡的过程。

对于间隙这类常引发极限环振荡的非线性环节，广泛存在于飞行器的结构当中，不可避免，且随着飞行器的使用老化，间隙大小也具有不确定性，不仅带来安全隐患，也伴随着较高的监测和维护成本[13]。Ⅱ型的非线性颤振模式有效避免了低于线性颤振边界的极限环振荡过程，间隙存在与否及间隙特性的变化对其影响较小。因此，通过合理设计全动舵面的参数，使其具有Ⅱ型非线性颤振模式，可以有效避免线性颤振边界以下极限环振荡现象的发生。

3　非线性颤振被动抑制方法

转变非线性颤振模式，即选取合适的刚度比使其位于图 6 中Ⅱ区域，因此，C 点，即线性系统颤振动压 Q_{F} 随扭转刚度 K_{T} 变化的极小值点可以成为研究的切入点，本文称 ξ_C 为临界刚度比。此外，由于所有Ⅰ型极限环的初始动压 Q_{L0} 都等于动压 Q_C，C 点的动压 Q_C 也决定了极限环出现的最小动压，即出现分岔的位置。除了转变非线性颤振模式，提高极限环初始动压 Q_{L0} 也是一种可行的非线性颤振抑制方法，可以推迟极限环的发生。因此，本节主要研究根部弯扭刚度和舵面质量特性对 C 点状态参数的影响。

3.1　根部弯曲扭转刚度

首先研究了弯曲刚度 K_{B} 不变、扭转刚度 K_{T} 变化时，非线性颤振特性随刚度比 ξ 的变化，结果如图 11 所示。随着扭转刚度的降低，ξ 减小，舵面非线性颤振模式由Ⅰ型逐渐转变为Ⅱ型，对于Ⅰ型非线性颤振，扭转刚度的降低不改变极限环初始动压 Q_{L0} 的大小，但会导致极限环发散动压 Q_{LF} 的降低；对于Ⅱ型非线性颤振，扭转刚度降低将提高 Q_{LF}。因为 C 点是颤振动压随扭转刚度变化的极小值点，所以 ξ_C 与 K_{T} 无关，是不变的。

同理，研究扭转刚度 K_{T} 不变、弯曲刚度 K_{B} 对极限环特性的影响，结果如图 12 所示。为便于研究 K_{B} 对临界刚度比 ξ_C 的影响，将动压 Q 和振荡幅值间隙比 A/δ 互换坐标位置，此外，一并给出了等效刚度比 $\bar{\xi}=\bar{K}_{\mathrm{T}}/K_{\mathrm{B}}$ 随 A/δ 的变化。通过分析可以发现，随

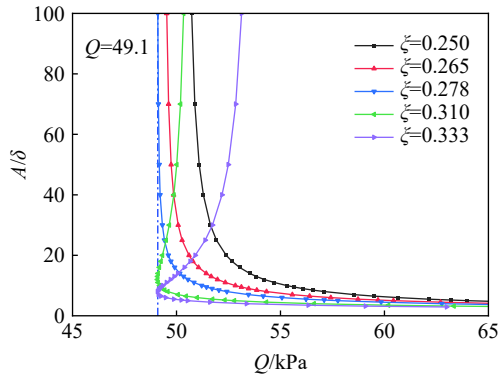


图 11　非线性颤振特性随刚度比变化（改变 K_{T}）

Fig. 11　Nonlinear flutter characteristic with variation of stiffness ratio (Changing K_{T})

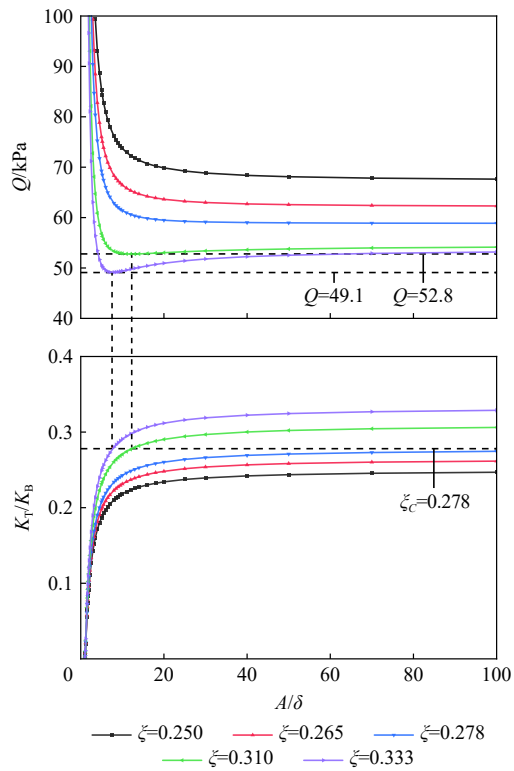


图 12　非线性颤振特性随刚度比变化（改变 K_{B}）

Fig. 12　Nonlinear flutter characteristic with variation of stiffness ratio (Changing K_{B})

着弯曲刚度的增加，刚度比 ξ 减小，舵面非线性颤振模式由Ⅰ型逐渐转变为Ⅱ型；对于Ⅰ型非线性颤振模式，弯曲刚度的增加会提高极限环初始动压 Q_{L0}，推迟极限环的发生，也会提高极限环发散动压 Q_{LF}；对于Ⅱ型非线性颤振模式，提高弯曲刚度将提高极限环发散动压 Q_{LF}；上下图对应来看，Ⅰ型非线性颤振模式分岔点对应的等效刚度比 $\bar{\xi}$ 是不变的，即临界刚度比 ξ_C 与 K_{B} 无关。

ξ_C 是和 K_{T}、K_{B} 取值无关的。因此，当只调节舵面根部弯曲和扭转刚度时，如果舵面刚度比 $\xi \leqslant \xi_C$，舵面表现Ⅱ型非线性颤振模式，$\xi > \xi_C$ 时，舵面表现

Ⅰ型非线性颤振模式。本文记舵面根部扭转弯曲刚度比值为ξ_C时，舵面的模态为临界模态。由于ξ_C与K_T、K_B无关，对于不同弯曲扭转刚度的舵面，其临界模态振型节线是不变的，可以通过实际舵面振型节线位置判断其非线性颤振模式类型。

降低全动舵面根部扭转弯曲刚度比ξ，将有助于促进非线性颤振模式由Ⅰ型向Ⅱ型转变。而降低刚度比的途径，提高弯曲刚度比降低扭转刚度更为有利。对于Ⅰ型非线性颤振模式，在未改变非线性颤振模式之前，提高弯曲刚度可以提高极限环初始动压Q_{L0}，而降低扭转刚度不改变Q_{L0}且导致Q_{LF}减小，降低扭转刚度可能带来静气弹发散的问题，需要注意。除了降低ξ，提高临界刚度比ξ_C也会促进非线性颤振模式的转化，由于ξ_C与舵面根部弯扭刚度无关，需对舵面质量特性对ξ_C的影响进行研究。

3.2 舵面质量特性

从舵面质量特性方程中来看，最终归结为 3 个惯量参数 (I_{xx},I_{yy},I_{xy})，但是单独对这 3 个参数变参分析是缺乏实际意义的。因此，本文针对实际舵面模型，在舵面上不同位置添加不同质量的配重，以改变舵面的质量分布，研究其对非线性颤振特性的影响。

在不同位置分别添加 50 g、150 g 和 250 g 配重，对舵面进行非线性颤振分析，研究舵面非线性颤振的模式，临界刚度比ξ_C、极限环初始动压Q_{L0}及发散动压Q_{LF}随配重位置的变化。

图 13～图 15 给出了不同质量配重情况下临界刚度比ξ_C随配重位置的变化情况。红线上ξ_C值为 0.278，与无配重情况下的ξ_C大小一致，即如果在红线位置添加配重，则不改变ξ_C的大小。舵面根部扭转弯曲刚度比ξ为 0.333，图中灰色部分代表$\xi_C>0.333$的情况下，在该区域中添加相应配重之后，舵面的非线性颤振模式将转变为Ⅱ型。综合来看，配重位置越靠近舵面根部前后缘，ξ_C越大，舵面非线性颤振模式越趋向于Ⅱ型，越靠近翼梢中部，ξ_C越小。

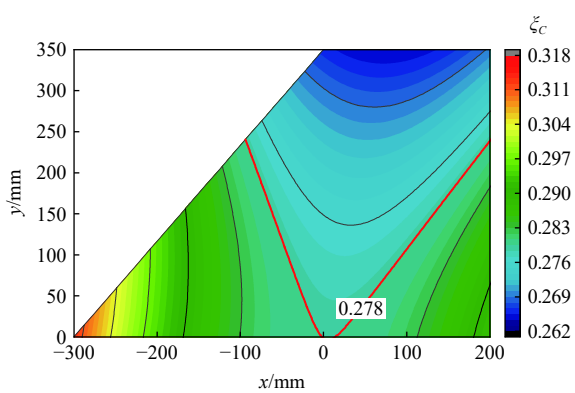


图 13 ξ_C随配重位置变化(配重 50 g)

Fig. 13 ξ_C with variation of weight position (50 g)

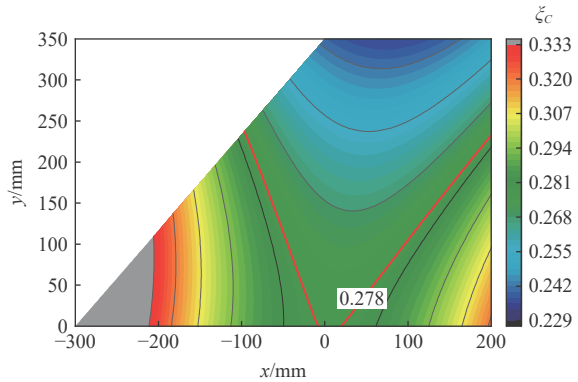


图 14 ξ_C随配重位置变化(配重 150 g)

Fig. 14 ξ_C with variation of weight position (150 g)

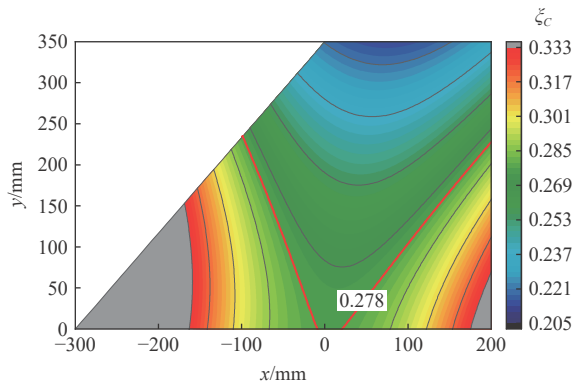


图 15 ξ_C随配重位置变化(配重 250 g)

Fig. 15 ξ_C with variation of weight position (250 g)

图 16～图 18 给出了不同质量配重情况下极限环初始动压Q_{L0}随配重位置变化情况，蓝线和红线上Q_{L0}分为49.1 kPa和53.9 kPa，对应未施加配重时舵面的极限环初始动压Q_{L0}和发散动压Q_{LF}的值，灰色区域表示非线性颤振模式变为Ⅱ型，不存在极限环。综合来看，配重位置越靠近前缘中部，Q_{L0}越小，越靠近后缘梢部，Q_{L0}越大。

图 19～图 21 给出了不同质量配重情况发散动

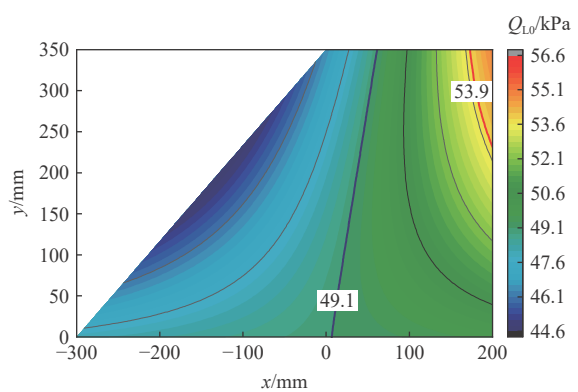


图 16 Q_{L0}随配重位置变化(配重 50 g)

Fig. 16 Q_{L0} with variation of weight position (50 g)

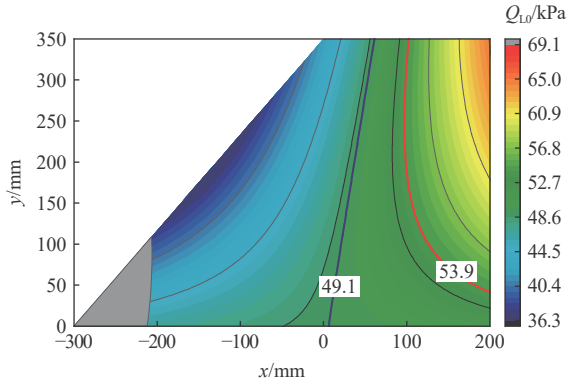


图 17　Q_{L0}随配重位置变化(配重 150 g)

Fig. 17　Q_{L0} with variation of weight position (150 g)

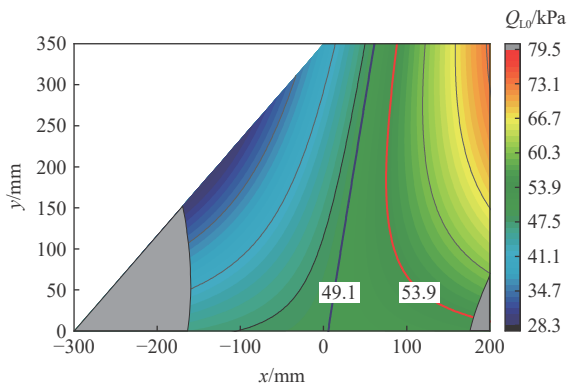


图 18　Q_{L0}随配重位置变化(配重 250 g)

Fig. 18　Q_{L0} with variation of weight position (250 g)

压Q_{LF}随配重位置变化情况，灰色半透明区域表示非线性颤振模式为Ⅱ型。综合来看，配重位置越靠近前缘中部，Q_{LF}越小，越靠近舵面后缘梢部，Q_{LF}越大。

通过上述计算分析，本节研究了配重质量、位置变化对临界刚度比ξ_C、极限环初始动压Q_{L0}及发散动压Q_{LF}的影响。总的来看，配重质量越大，影响越明显。对于本文所研究的对象，从调整舵面质量特性入手，抑制舵面非线性颤振可能有效的策略如下：

1）如果以提高舵面非线性颤振极限环初始动压Q_{L0}和发散动压Q_{LF}为目的，可以在舵面后缘翼梢

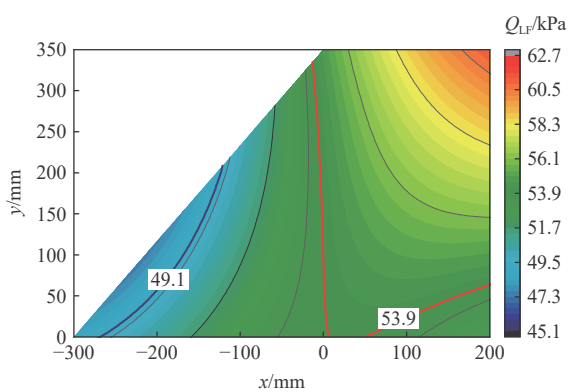


图 19　Q_{LF}随配重位置变化(配重 50 g)

Fig. 19　Q_{LF} with variation of weight position (50 g)

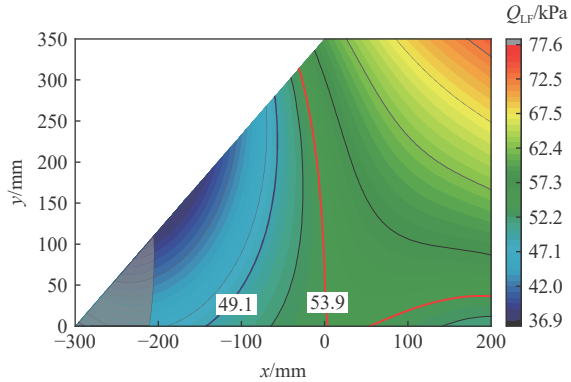


图 20　Q_{LF}随配重位置变化(配重 150 g)

Fig. 20　Q_{LF} with variation of weight position (150 g)

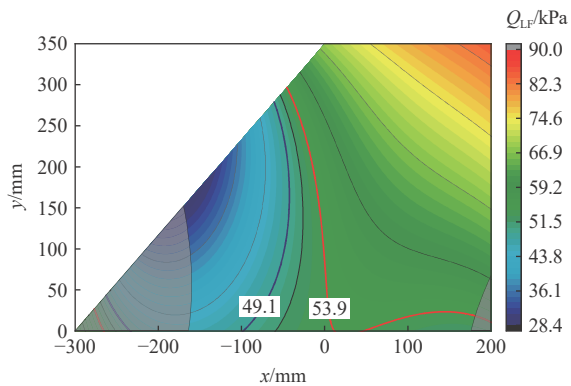


图 21　Q_{LF}随配重位置变化(配重 250 g)

Fig. 21　Q_{LF} with variation of weight position (250 g)

位置施加配重，小质量配重便可达到推迟极限环发生的目的。

2）改变舵面非线性颤振模式，需要在舵面根部后缘位置施加配重，且配重质量较大(见图 21 中后缘灰色区域)。

对于本文算例，从调整舵面质量特性入手，相较于改变非线性颤振模式，提高极限环初始动压Q_{L0}更容易实现。一种更好的方法是同时调整舵面的质量特性和根部弯曲扭转刚度，使舵面刚度比低于临界刚度比，且极限环发散动压Q_{LF}满足相关指标要求。

3.3　非线性颤振模式近似评估方法

本文所研究的全动舵面颤振类型为典型的弯扭耦合型颤振，大多数情况下，当弯曲和扭转频率差越小，颤振动压越低。在防颤振设计当中，也常依据频率分离原则，使耦合的模态频率分开，从而提高颤振动压。3.2 节的讨论分析主要基于所选对象的数值仿真结果，提出的非线性颤振抑制方法适用性也有限。因此，本节从模态频率和振型角度提出非线性颤振模式近似评估方法，为舵面设计和改进方向提供指导。

首先考虑状态 A、B、C 这 3 种不同情况下全动舵面的前两阶模态振型，如图 22 所示。可以发现，

随着扭转刚度的降低，即刚度比ξ的降低，两阶模态振型节线发生了顺时针旋转。通过 3.1 节的分析可知，当舵面质量特性不变时，改变舵面根部刚度不改变临界刚度比ξ_C，即不改变对应的临界模态的节线位置，即状态 C 情况下的模态振型节线位置，且当线性舵面模态节线与临界模态节线重合时，颤振动压最低。因此，对于实际的舵面，当其质量特性及相关计算条件确定后，可以根据颤振动压随扭转刚度的变化求出临界模态的节线位置，对于具有不同弯曲和扭转刚度的舵面，当前两阶模态节线分别位于临界节线左侧时，其非线性颤振模式为Ⅰ型，右侧时则为Ⅱ型。

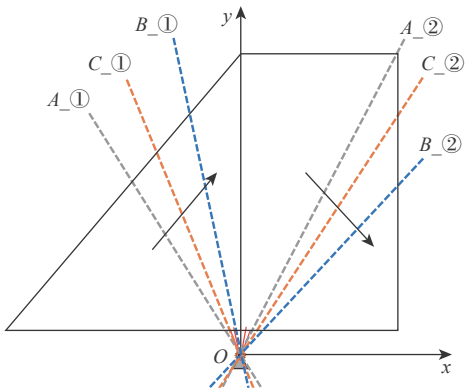


图 22 模态振型随扭转刚度变化

Fig. 22 Mode shapes with variation of torsion stiffness

从图 22 中模态节线发生顺时针的旋转来看，对于Ⅰ型非线性颤振模式，大多数情况下，其第 2 阶模态振型节线相比于第 1 阶更靠近旋转轴，因此，扭转刚度的降低导致第 2 阶模态频率降低更多，即频率差的减小，更多地起到降低颤振动压的效果，慢慢转化为Ⅱ型。对于Ⅱ型非线性颤振模式，大多数情况下，其第 1 阶模态振型节线相比于第 2 阶更靠近旋转轴，因此，扭转刚度的降低导致第 1 阶模态频率降低更多，即频率差的扩大，更多地起到提高颤振动压的效果。

因此，如果全动舵面在扭转方向存在间隙，通过合理设计舵面的前两阶模态，使得第 1 阶模态振型节线更靠近旋转轴，可以有效避免低于线性颤振边界极限环的发生。

除了通过模态振型节线与临界节线的关系来近似判断非线性颤振模式之外，也可以从模态频率差入手，求出其对扭转刚度的极小值点，从而近似估计颤振动压随扭转刚度变化的极小值点，来快速估计临界刚度比ξ_C。

由结构动力学方程：

$$\begin{bmatrix} I_{xx} & -I_{xy} \\ -I_{xy} & I_{yy} \end{bmatrix} \begin{bmatrix} q_{\mathrm{B}} \\ q_{\mathrm{T}} \end{bmatrix} + \begin{bmatrix} K_{\mathrm{B}} & 0 \\ 0 & K_{\mathrm{T}} \end{bmatrix} \begin{bmatrix} q_{\mathrm{B}} \\ q_{\mathrm{T}} \end{bmatrix} = 0 \tag{11}$$

得到模态频率差$\Delta\omega(\Delta\omega>0)$与系统结构参数之间的关系为

$$(\Delta\omega)^2 = \frac{K_{\mathrm{B}}I_{yy} + K_{\mathrm{T}}I_{xx} - 2\sqrt{K_{\mathrm{T}}K_{\mathrm{B}}\left(I_{xx}I_{yy} - I_{xy}^2\right)}}{2\left(I_{xx}I_{yy} - I_{xy}^2\right)} \tag{12}$$

对扭转刚度K_{T}求偏导有

$$\frac{\partial(\Delta\omega)^2}{\partial K_{\mathrm{T}}} = \frac{I_{xx} - \sqrt{\dfrac{K_{\mathrm{B}}}{K_{\mathrm{T}}}\left(I_{xx}I_{yy} - I_{xy}^2\right)}}{2\left(I_{xx}I_{yy} - I_{xy}^2\right)} \tag{13}$$

令$\dfrac{\partial(\Delta\omega)^2}{\partial K_{\mathrm{T}}} \leqslant 0$，得到

$$\frac{K_{\mathrm{T}}}{K_{\mathrm{B}}} \leqslant \frac{I_{xx}I_{yy} - I_{xy}^2}{I_{xx}^2} \tag{14}$$

记$\eta = \left(I_{xx}I_{yy} - I_{xy}^2\right)/I_{xx}^2$，当舵面的刚度比$\xi \leqslant \eta$时，频率差$\Delta\omega$随着扭转刚度$K_{\mathrm{T}}$的降低而增大。图 23 给出了模态频率、频率差和颤振动压随扭转刚度的变化情况。从图中可以看出，本文模型$\eta = 0.291$，而分岔点对应的临界刚度比$\xi_C = 0.278$，两者具有一定的误差。由于上述分析过程没有考虑气动力的影响，因此η只能用于估计临界刚度比ξ_C。

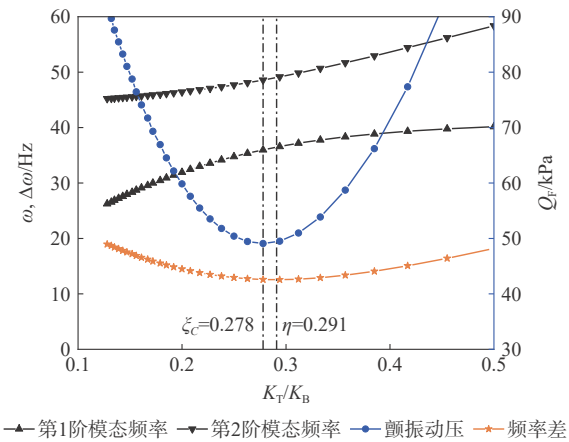


图 23 模态频率、频率差和颤振动压随刚度比变化

Fig. 23 Modal frequency, frequency difference and flutter dynamic pressure with variation of stiffness ratio

考虑舵面不同位置施加 50 g 配重时，计算η的变化如图 24 所示。通过对比图 13 可以发现，η和ξ_C变化趋势几乎一致，因此施加配重调整舵面质量特性时，可以通过参数η快速估计临界刚度比ξ_C的变化情况，从而判断非线性颤振模式的转变趋势。

除了通过转变模式抑制非线性颤振，本文也从模态振型角度研究了Q_{L0}随配重位置的变化规律，如图 25 所示。可以发现，当在第 2 阶临界模态节线附近施加配重时，有助于提高Q_{L0}，在第 1 阶临界模态节线附近施加配重将导致Q_{L0}的降低。从频率差角度来看，当在第 2 阶临界模态节线附近施加配重时，更多地导致第 1 阶临界模态频率下降，频率

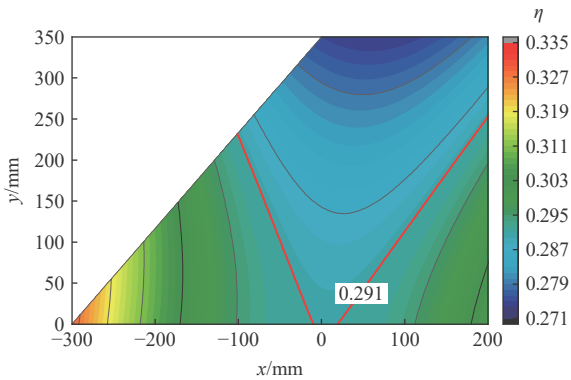


图 24　η随配重位置变化(配重 50 g)

Fig. 24　η with variation of weight position (50 g)

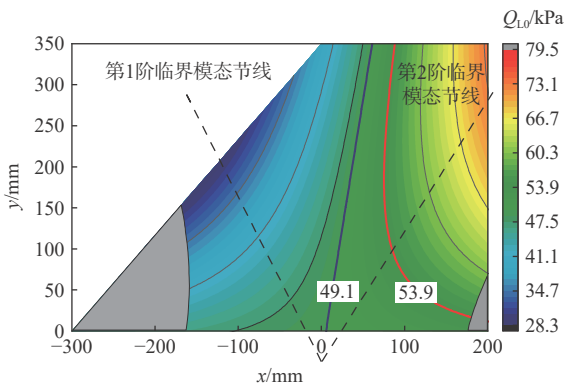


图 25　临界模态节线位置与Q_{L0}变化

Fig. 25　Q_{L0} with variation of weight position (250 g)

差增大，某种程度上提高了临界状态的颤振动压，即极限环初始动压Q_{L0}，反之类似。

总的来看，从调节舵面质量特性角度抑制非线性颤振，有以下 2 点实用性较高的策略：

1）施加配重以改变舵面非线性颤振模式时，可以以配重位置对η的影响为参考，预估配重施加的大致方位，通过调节配重质量和降低根部扭转弯曲刚度比，直到改变非线性颤振模式，且发散动压满足相关指标要求为止。

2）可以通过在舵面第 2 阶临界模态节线附近施加配重以达到提高极限环初始动压Q_{L0}的作用，推迟极限环的产生。

4　结　论

本文针对含扭转间隙全动舵面开展非线性颤振分析，主要结论如下：

1）对于典型含扭转间隙小展弦比全动舵面，存在 2 种不同的非线性颤振模式，其中，Ⅰ型模式存在着稳定的极限环振荡范围，极限环初始动压低于线性系统颤振动压，Ⅱ型模式不存在稳定极限环振荡范围，其发散动压不低于线性系统颤振动压。

2）舵面扭转弯曲刚度比与临界刚度比的关系决定了舵面非线性颤振模式的类型，降低舵面根部扭转弯曲刚度比或者提高临界刚度比均可促进非线性颤振模式从Ⅰ型向Ⅱ型转变。

3）临界刚度比与根部弯扭刚度无关，受舵面质量特性影响。当舵面质量特性改变时，可以通过参数η近似判断临界刚度比的变化趋势。

4）通过调整舵面根部弯曲扭转刚度或者质量特性可以提高极限环初始动压，甚至改变非线性颤振模式，从而达到抑制非线性颤振的效果。

需要说明的是，本文研究的对象主要是自身刚度较大的小展弦比全动舵面，对于自身弹性较大的舵面，其自身弹性可能会对非线性颤振特性产生一定的影响，有待进一步开展研究。

参考文献（References）

[1] 杨超, 吴志刚, 万志强, 等. 飞行器气动弹性原理[M]. 2版. 北京: 北京航空航天大学出版社, 2016: 165-171.
YANG C, WU Z G, WAN Z Q, et al. Flight vehicle aeroelasticity Principle[M]. 2rd ed. Beijing: Beihang University Press, 2016: 165-171(in Chinese).

[2] ZHAO L C, YANG Z C. Chaotic motions of an airfoil with non-linear stiffness in incompressible flow[J]. Journal of Sound and Vibration, 1990, 138(2): 245-254.

[3] 赵永辉. 气动弹性力学与控制[M]. 北京: 科学出版社, 2007: 7-10.
ZHAO Y H. Aeroelastic mechanics and control[M]. Beijing: Science Press, 2007: 7-10(in Chinese).

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

[5] 卢晋, 吴志刚, 杨超. 电动舵机模块化建模及动刚度仿真[J]. 北京航空航天大学学报, 2021, 47(4): 765-778.
LU J, WU Z G, YANG C. Modular modeling and dynamic stiffness simulation of electromechanical actuator[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(4): 765-778(in Chinese).

[6] LAURENSON R M, TRN R M. Flutter analysis of missile control surfaces containing structural nonlinearities[J]. AIAA Journal, 1980, 18(10): 1245-1251.

[7] LEE B H K, TRON A. Effects of structural nonlinearities on flutter characteristics of the CF-18 aircraft[J]. Journal of Aircraft, 1989, 26(8): 781-786.

[8] LEE I, KIM S H. Aeroelastic analysis of a flexible control surface with structural nonlinearity[J]. Journal of Aircraft, 1995, 32(4): 868-874.

[9] CONNER M D, TANG D M, DOWELL E H, et al. Nonlinear behavior of a typical airfoil section with control surface freeplay: A numerical and experimental study[J]. Journal of Fluids and Structures, 1997, 11(1): 89-109.

[10] TANG D, DOWELL E H, VIRGIN L N. Limit cycle behavior of an airfoil with a control surface[J]. Journal of Fluids and Structures,

1998, 12(7): 839-858.

[11] PADMANABHAN M A, DOWELL E H, PASILIAO C L. Computational study of aeroelastic limit cycles due to localized structural nonlinearities[J]. Journal of Aircraft, 2018, 55(4): 1531-1541.

[12] LEE B H K, PRICE S J, WONG Y S. Nonlinear aeroelastic analysis of airfoils: Bifurcation and chaos[J]. Progress in Aerospace Sciences, 1999, 35(3): 205-334.

[13] PANCHAL J, BENAROYA H. Review of control surface freeplay[J]. Progress in Aerospace Sciences, 2021, 127: 100729.

[14] BUENO D D, WAYHS-LOPES L D, DOWELL E H. Control-surface structural nonlinearities in aeroelasticity: A state of the art review[J]. AIAA Journal, 2022, 60(6): 3364-3376.

[15] 赵永辉, 胡海岩. 具有操纵面间隙非线性二维机翼段的气动弹性分析[J]. 航空学报, 2003, 24(6): 521-525.
ZHAO Y H, HU H Y. Aeroelastic analysis of a two-dimensional airfoil with control surface freeplay nonlinearity[J]. Acta Aeronautica et Astronautica Sinica, 2003, 24(6): 521-525(in Chinese).

[16] 李道春, 向锦武. 控制面间隙对非线性二元机翼气动弹性响应的影响[J]. 航空学报, 2009, 30(8): 1385-1391.
LI D C, XIANG J W. Effect of control surface freeplay on nonlinear aeroelastic responses of an airfoil[J]. Acta Aeronautica et Astronautica Sinica, 2009, 30(8): 1385-1391(in Chinese).

[17] 杨智春, 田玮, 谷迎松, 等. 带集中非线性的机翼气动弹性问题研究进展[J]. 航空学报, 2016, 37(7): 2013-2044.
YANG Z C, TIAN W, GU Y S, et al. Advance in the study on wing aeroelasticity with concentrated nonlinearity[J]. Acta Aeronautica et Astronautica Sinica, 2016, 37(7): 2013-2044(in Chinese).

[18] 隋鑫, 韩敬永, 刘博, 等. 含间隙舵面非线性动力学分析[J]. 导弹与航天运载技术, 2020(3): 107-110.
SUI X, HAN J Y, LIU B, et al. Nonlinear dynamical analysis of the rudder with gaps[J]. Missiles and Space Vehicles, 2020(3): 107-110(in Chinese).

[19] WANG X Y, WU Z G, SUN Y K, et al. A novel method for estimating three-domain limit cycles in a 3D wing-aileron model with freeplay in aileron deflection[J]. Journal of Fluids and Structures, 2021, 105: 103286.

[20] WU Z G, YANG N, YANG C. Identification of nonlinear structures by the conditioned reverse path method[J]. Journal of Aircraft, 2015, 52(2): 373-386.

[21] YANG N, WANG N, ZHANG X, et al. Nonlinear flutter wind tunnel test and numerical analysis of folding fins with freeplay nonlinearities[J]. Chinese Journal of Aeronautics, 2016, 29(1): 144-159.

[22] 何昊南, 于开平, 唐宏, 等. 有间隙折叠舵面的振动实验与非线性建模研究[J]. 力学学报, 2019, 51(5): 1476-1488.
HE H N, YU K P, TANG H, et al. Vibration experiment and nonlinear modelling research on the folding fin with freeplay[J]. Chinese Journal of Theoretical and Applied Mechanics, 2019, 51(5): 1476-1488(in Chinese).

[23] HE H N, TANG H, YU K P, et al. Nonlinear aeroelastic analysis of the folding fin with freeplay under thermal environment[J]. Chinese Journal of Aeronautics, 2020, 33(9): 2357-2371.

[24] 王强, 马志赛, 张欣, 等. 基于模态综合法的含间隙折叠舵面动态特性分析[J]. 航空学报, 2020, 41(5): 223507.
WANG Q, MA Z S, ZHANG X, et al. Dynamic characteristic analysis for a folding fin with freeplay nonlinearities based on mode synthesis method[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(5): 223507(in Chinese).

[25] SUN Y K, YANG C, WU Z G, et al. Nonlinear system identification of an all movable fin with rotational freeplay by subspace-based method[J]. Applied Sciences, 2020, 10(4): 1205.

[26] XUAN C W, HAN J L, ZHANG B, et al. Hypersonic flutter and flutter suppression system of a wind tunnel model[J]. Chinese Journal of Aeronautics, 2019, 32(9): 2121-2132.

[27] FIROUZ-ABADI R D, ALAVI S M, SALARIEH H. Analysis of non-linear aeroelastic response of a supersonic thick fin with plunging, pinching and flapping free-plays[J]. Journal of Fluids and Structures, 2013, 40: 163-184.

[28] TIAN W, YANG Z C, ZHAO T. Nonlinear aeroelastic characteristics of an all-movable fin with freeplay and aerodynamic nonlinearities in hypersonic flow[J]. International Journal of Non-Linear Mechanics, 2019, 116: 123-139.

[29] 李家旭, 田玮, 谷迎松. 考虑间隙非线性的控制舵非线性气动弹性分析[J]. 航空工程进展, 2020, 11(6): 827-835.
LI J X, TIAN W, GU Y S. Nonlinear aero-elastic analysis of control fin with free-play nonlinearity[J]. Advances in Aeronautical Science and Engineering, 2020, 11(6): 827-835(in Chinese).

[30] 宋晨, 吴志刚, 杨超. 二元机翼滑模变结构控制颤振主动抑制[J]. 北京航空航天大学学报, 2010, 36(12): 1400-1403.
SONG C, WU Z G, YANG C. Sliding mode variable structure control of flutter suppression for two-dimensional wing[J]. Journal of Beijing University of Aeronautics and Astronautics, 2010, 36(12): 1400-1403(in Chinese).

[31] 李道春, 向锦武. 间隙非线性气动弹性颤振控制[J]. 北京航空航天大学学报, 2007, 33(6): 640-643.
LI D C, XIANG J W. Flutter control of aeroelasticity with freeplay nonlinearity[J]. Journal of Beijing University of Aeronautics and Astronautics, 2007, 33(6): 640-643(in Chinese).

[32] 贾尚帅, 丁千. 含间隙超音速二元弹翼非线性颤振与主动控制[J]. 中国科学:物理学 力学 天文学, 2013, 43(4): 390-400.
JIA S S, DING Q. Nonlinear flutter and active control of the supersonic two-dimensional missile wings with a freeplay[J]. Scientia Sinica: Physica, Mechanica & Astronomica, 2013, 43(4): 390-400(in Chinese).

[33] 聂雪媛, 郑冠男, 杨国伟. 含间隙非线性机翼跨声速颤振时滞反馈控制[J]. 北京航空航天大学学报, 2021, 47(10): 1980-1988.
NIE X Y, ZHENG G N, YANG G W. Time delay feedback control for transonic flutter of airfoil with free-play nonlinearity[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(10): 1980-1988(in Chinese).

[34] 史晓鸣, 梅睿, 张晓宏, 等. 配重位置对全动舵面颤振抑制效果研究[J]. 应用力学学报, 2016, 33(2): 268-272.
SHI X M, MEI R, ZHANG X H, et al. Effect of balance mass location on flutter of an all-moving rudder[J]. Chinese Journal of Applied Mechanics, 2016, 33(2): 268-272(in Chinese).

[35] LEE Y S, VAKAKIS A F, BERGMAN L A, et al. Suppression aeroelastic instability using broadband passive targeted energy transfers. Part 1: Theory[J]. AIAA Journal, 2007, 45(3): 693-711.

[36] LEE Y S, KERSCHEN G, MCFARLAND D M, et al. Suppressing aeroelastic instability using broadband passive targeted energy transfers. Part 2: Experiments[J]. AIAA Journal, 2007, 45(10):

2391-2400.

[37] TIAN W, ZHAO T, GU Y S, et al. Nonlinear flutter suppression and performance evaluation of periodically embedded nonlinear vibration absorbers in a supersonic FGM plate[J]. Aerospace Science and Technology, 2022, 121: 107198.

[38] 王永岩. 动态子结构方法理论及应用[M]. 北京: 科学出版社, 1999: 33-43.
WANG Y Y. Theory and application of dynamic substructure method[M]. Beijing: Science Press, 1999: 33-43(in Chinese).

[39] 杨宁. 间隙非线性结构的气动弹性建模与分析方法研究[D]. 北京: 北京航空航天大学, 2015: 119-120.
YANG N. Studies on aeroelastic modeling and analysis methods of structures with freeplay nonlinearity[D]. Beijing: Beihang University, 2015: 119-120(in Chinese).

[40] ASHLEY H, ZARTARIAN G. Piston theory-A new aerodynamic tool for the aeroelastician[J]. Journal of the Aeronautical Sciences, 1956, 23(12): 1109-1118.

[41] SHEN S F. An approximate analysis of nonlinear flutter problems[J]. Journal of the Aerospace Sciences, 1959, 26(1): 25-32.

Nonlinear flutter modes and flutter suppression of an all-movable fin with freeplay

BAI Liuyue，WU Zhigang*，YANG Chao

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

Abstract：Freeplay nonlinearity widely exists in the structure of aircraft, which can bring limit cycle oscillations (LCO), leading to the fatigue failure of the structure. There is relatively little research on the mechanism and passive suppression of the nonlinear flutter for all-movable fins. Therefore, this paper proposes two different nonlinear flutter modes and provides passive suppression methods. The dynamic model is developed using the describing function method and piston theory. Considering that the torsional stiffness is reduced due to the existence of the freeplay, nonlinear flutter characteristics under different torsional stiffnesses are compared and two different nonlinear flutter modes are defined, in which mode Ⅱ doesn't contain stable LCO, effectively suppressing the LCO below the linear flutter boundary. Then the influence of bending and torsional stiffnesses and mass moments of inertia on nonlinear flutter modes and the dynamic pressure of LCO appearance is studied, and some nonlinear flutter passive suppression methods are proposed. The findings indicate that reasonable adjustment of the bending and torsional stiffnesses or mass moments of inertia can raise the dynamic pressure of LCO appearance and even change the nonlinear flutter mode, so as to suppress the nonlinear flutter.

Keywords：freeplay nonlinearity; nonlinear flutter; all-movable fins; limit cycle oscillations; flutter suppression

Received：2022-03-18；**Accepted**：2022-05-20；**Published Online**：2022-05-31 15:21

URL：kns.cnki.net/kcms/detail/11.2625.V.20220530.2125.005.html

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0635

跨音涡轮转子叶尖间隙内流动分析与建模

轩笠铭[1, 2]，邹正平[2, 3, *]，曾飞[1, 2, 4]

(1. 北京航空航天大学 能源与动力工程学院，北京 100191；2. 航空发动机气动热力国防重点实验室，北京 102206；
3. 北京航空航天大学 航空发动机研究院，北京 102206；4. 中国航发湖南机械动力研究所，株洲 412002)

摘 要：为明确跨音涡轮叶尖泄漏流动机理，进一步提升涡轮效率，对跨音条件下叶顶喷气对平叶尖及凹槽叶尖性能的影响进行研究，并探讨跨音条件下平叶尖及凹槽叶尖间隙内部的流动状态。结果表明：跨音条件下，叶顶喷气可以增加平叶尖叶栅的气动效率，而刮削涡仍是凹槽内的主控流动结构；喷气流量的增加对平叶尖的总泄漏流量影响有限，但会增加凹槽叶尖的总泄漏流量。在更高负荷情况下，平叶尖间隙内呈跨音速流动特征，具体状态与叶片负荷、叶片厚度有关；凹槽叶尖条件下，泄漏流动在吸力侧肋条上方快速膨胀至超声速状态。基于此，建立可用于跨音条件下的泄漏流量预测模型。

关 键 词：跨音涡轮；泄漏流动；叶尖间隙；叶顶喷气；流动模型
中图分类号：V231.3
文献标志码：A **文章编号**：1001-5965(2023)09-2374-11

涡轮作为航空发动机的关键部件之一，提升涡轮的效率可以有效提高航空发动机的性能。Denton[1]指出，叶尖泄漏流动引起的损失约占转子通道内部总损失的1/3。随着涡轮叶片负荷不断增加，泄漏损失进一步增大，因此，有效组织叶尖泄漏流动成为改善涡轮性能和提高其可靠性的关键[2-3]。

Rains[4]较早地对叶轮机泄漏流动进行了研究，指出在间隙内部叶片周向的压力梯度远大于流向方向，泄漏流动方向可以认为近似与中弧线方向垂直。Bindon[5]、Yaras和Sjolander[6]对泄漏损失的分布情况进行了初步研究，随后更多的学者对泄漏流动结构和损失细节进行了关联[7-9]。为更好地控制泄漏损失，研究人员提出了凹槽叶尖方案[10]。邹正平等[11-13]指出凹槽内的刮削涡结构可以起到气动篦齿的封严效果，降低了泄漏流动的流通面积，同时降低了泄漏流动的能量。Krishnababu等[14]和Kegali等[15]的实验结果表明，凹槽叶顶可以有效提升涡轮的气动性能。

随着航空发动机涡轮的负荷不断增加，目前典型的高压涡轮通常在跨音环境下工作，此时，在轴向中前部，间隙内仍保持亚声速状态，压力侧有清晰的分离泡结构；在轴向中部以后，间隙内出现激波结构对整个叶尖端区的气动热力结构产生显著的影响，激波的出现及泄漏流量变化规律的改变使得跨音条件下的间隙流动机制明显不同于亚声速条件下的间隙流动[16-18]。首先，由于间隙内部存在分离泡并与机匣形成收缩扩张通道，超声速气流经过通道最小面积位置后将继续加速；此外，该加速流动促使流动分离迅速再附，抑制了分离泡的发展[19]。其次，在分离再附位置及其下游，间隙内部超声速流动中还会出现数条斜激波，并以正激波结尾[20-21]。激波的出现也使得间隙内部超声速流动特性完全不同于亚声速泄漏流动，这将导致当地泄漏量不再随着进出口压差的变化而改变，而仅是进口总温、

收稿日期：2021-10-26；**录用日期**：2022-01-11；**网络出版时间**：2022-01-29 14:25
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220129.1034.001.html
基金项目：国家自然科学基金(51676005)
*通信作者. E-mail：zouzhengping@buaa.edu.cn

引用格式：轩笠铭，邹正平，曾飞. 跨音涡轮转子叶尖间隙内流动分析与建模[J]. 北京航空航天大学学报，2023，49(9)：2374-2384.
XUAN L M，ZOU Z P，ZENG F. Analyzing and modeling flow in tip clearance of transonic turbine rotor[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2374-2384 (in Chinese).

总压和间隙高度的函数。最后, 跨音环境对叶尖泄漏流动的控制也有较大影响。Wheeler 和 Saleh[22] 对跨音条件下平叶尖和凹槽叶尖的泄漏流动进行了分析, 发现跨音条件下凹槽叶尖的射流系数仍小于平叶尖, 对泄漏流动有一定的抑制作用。Duan 和 He[23] 认为, 跨音条件下, 凹槽一方面会堵塞泄漏流动, 但另一方面会增加泄漏流进入吸力侧时的膨胀损失, 在这 2 方面间需要作出平衡。

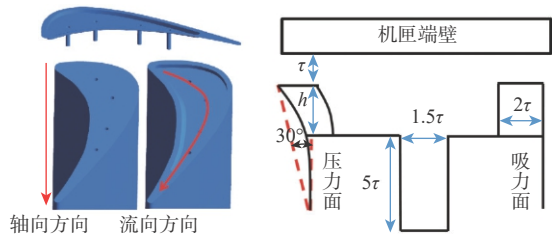
叶顶冷气的引入大幅改善了叶尖的传热性能, 同时, 叶顶喷入间隙的冷气也会对泄漏流动结构产生影响[24-29]。Gao 等[30] 认为叶顶喷气对泄漏损失的影响主要体现在以下方面: 叶顶喷气对泄漏流的阻挡、叶顶喷气导致的泄漏流和主流速度差的减小及喷气会带来的额外损失。Zhou 和 Hodson[31] 发现对于凹槽叶尖, 叶顶喷气对凹槽叶尖的气动性能影响有限, 且此时泄漏损失和泄漏量不成正比。

虽然以上学者对跨音涡轮泄漏流动及叶顶喷气对涡轮性能的影响做出了一定研究, 但对跨音环境及喷气对泄漏流动的综合影响尚缺乏研究。本文通过模拟机匣相对运动条件下平面叶栅三维流场, 研究跨音条件下, 平叶顶及凹槽叶顶在有无喷气条件下间隙内的流动结构及其对泄漏流量的影响, 并基于流场分析结果建立泄漏流量的预测模型, 为更好地快速分析叶尖气动性能提供参考。

1　研究方法

本文以某跨音高压涡轮转子的 Low Stagger (LS)叶型为研究对象, 其叶型参数如下: 弦长为 43.63 mm, 轴向弦长为 26.681 mm, 栅距为 36.794 mm, 进口角为 27.43°, 出气角为 71.5°。设计工况下, 出口雷诺数为 1.1×10^6, 进口马赫数为 0.18, 出口马赫数为 0.78[32]。定义吹风比为喷气孔进口动量与叶栅进口气流动量之比, 喷气孔吹风比为 2。为了考虑叶尖机匣相对运动对泄漏流动的影响, 设定上端壁与叶尖的相对运动速度为 300 m/s。凹槽叶尖叶片采用了压力面肋条前倾 30°的造型方法。图 1 为平叶尖和凹槽叶尖的几何示意图及参数, 4 个气膜孔分别位于轴向的 30%、50%、70% 和 80% 位置, τ 为间隙高度, 取 1% 叶高, 即 0.5 mm, 凹槽深度 h 为 1.5τ, 肋条宽度为 2τ, 压力侧前倾的基准面为 97.5% 叶高位置处, 肋条前倾的基准截面为垂直于流向方向的截面。肋条最大前倾角度为 30°, 并在前缘和尾缘附近逐渐过渡减小至 0°, 同时压力面肋条和叶片压力面主体部分通过光滑过渡的方式连接。

计算所采用的网格由商用软件 ICEM CFD 生成, 整个流体域内为一体化结构网格, 网格总数约为单通道 2.98×10^6, 网格示意图如图 2 所示。通过调整间隙径向网格数改变总体网格数, 如表 1 所示, 其变化对无量纲泄漏流量的影响如图 3 所示。本文中, 泄漏量通过对吸力侧间隙出口截面进行积分得到。可以看出, 随着网格数量的增加, 间隙泄漏量的变化逐渐平缓。当总网格数量达到 3×10^6 左右时, 间隙泄漏量变化小于 1%, 可以认为网格 4 达到计算要求。



(a) 几何示意图　　(b) 几何参数示意图

图 1　平叶尖和凹槽叶尖的几何示意图及参数

Fig. 1　Geometry and parameters of flat tip and squealer tip

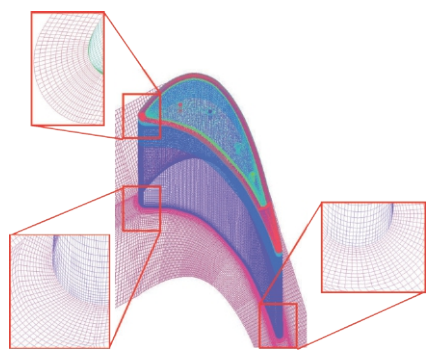


图 2　网格示意图

Fig. 2　Sketch figure of mesh

表 1　网格参数

Table 1　Grid parameters

网格编号	间隙径向网格数量	网格总数/10^6
网格1	9	2.51
网格2	15	2.67
网格3	21	2.84
网格4	26	2.98
网格5	31	3.14

采用商用软件 ANSYS CFX 求解三维定常黏性雷诺平均 Navier-Stokes 方程, 数值求解时选择切应力输运(shear stress transport, SST)k-ω 湍流模型封闭方程组。图 4 为仿真与实验结果比较, 在 50% 叶高处计算得到的叶片表面压力分布与实验值对比如图 4(a)所示, 图中纵轴为等熵马赫数, 其值按式(1)计算得到, 叶顶区域通过与另一叶栅的实验测试值对比得到, 如图 4(b)所示, 该叶栅基本参数与实验方法见文献 [33]。其中, FT 为平叶尖, ST 为凹槽叶尖, 压力系数计算方法如式(2)所示。

$$Ma_{is}=\sqrt{\frac{2}{k-1}\left[\left(\frac{p_{cin}^{*}}{p}\right)^{\frac{k-1}{k}}-1\right]} \tag{1}$$

$$C_p=\frac{p^{*}-p_{out}}{p_{in}^{*}-p_{out}} \tag{2}$$

式中：p_{cin}^{*}为叶栅进口总压；p为当地压力；k为常数，取1.4，其值旨在描述在等熵流动条件下当地的压力及流场的可压缩性情况；p^{*}为当地总压；p_{out}为间隙出口压力，p_{in}^{*}为间隙进口总压。

对比模拟和实验结果可知，本文的模拟结果准确可靠。

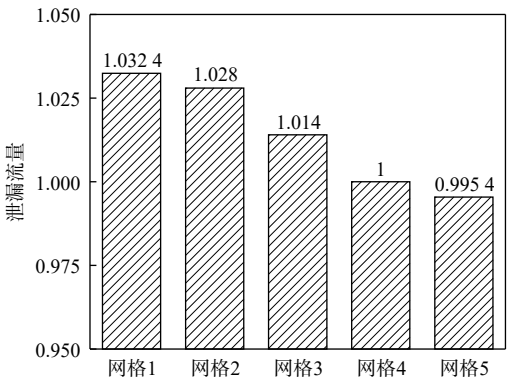
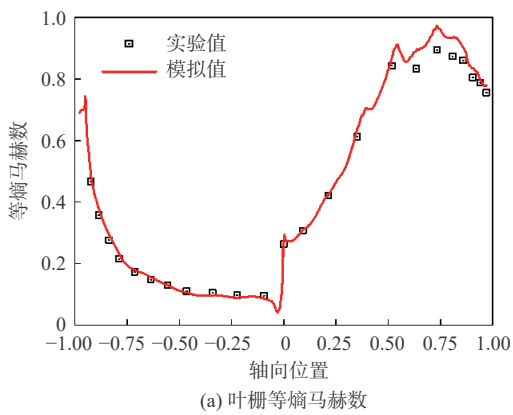
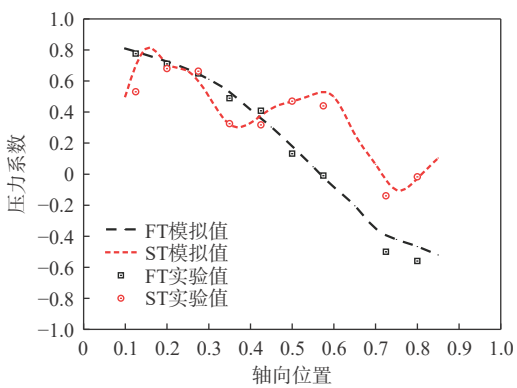


图3 网格无关性验证

Fig. 3 Grid independence verification



(a) 叶栅等熵马赫数



(b) 叶顶中弧线压力系数分布对比

图4 仿真与实验结果

Fig. 4 CFD and experiment results

2 结果与讨论

2.1 叶顶喷气条件下凹槽叶尖与平叶尖性能对比

本文采用Young和Horlock[34]提出的带冷气的涡轮叶栅效率计算方法评估叶栅效率，式(3)与式(4)给出考虑了冷气和主流的压力掺混的加权等效压力计算方法，式(5)体现了冷气的压力对叶栅效率带来的影响。

$$\frac{P_{0w}}{P_{01g}}=\prod_i\left(\frac{P_{01c}}{P_{01g}}\right)^{(\phi_i R_c/R)} \tag{3}$$

$$R=(1-\phi)R_g+\phi R_c \tag{4}$$

$$(\varepsilon_c)_{WP}=\frac{h_{02}-h_2}{h_{0w}-h_{2sw}} \tag{5}$$

式中：P_{01g}为主流进口总压；P_{01c}为冷气的进口总压；P_{0w}为考虑冷气和主流掺混后的总压；下标c表示考虑冷气条件；ϕ为冷气的质量分数；R为气体常数；R_g为主流气体的气体常数；h_{02}为实际状态下出口总焓；h_2为实际状态下出口静焓；h_{0w}为利用式(3)和式(4)计算考虑两股气流掺混后的总焓；h_{2sw}为等熵膨胀后的静焓，则叶栅效率ε_c为叶栅气流的实际焓降与考虑压力掺混后气体的等熵焓降之比，下标WP表示经过压力加权计算后的效率。

图5对比了不同条件下叶栅的气动效率，jet为有叶顶喷气的工况，可以看出，叶顶喷气可以增加平叶尖叶栅的气动效率，而对凹槽叶尖性能的提升较小；无论有无喷气，凹槽叶尖气动性能均优于平叶尖。

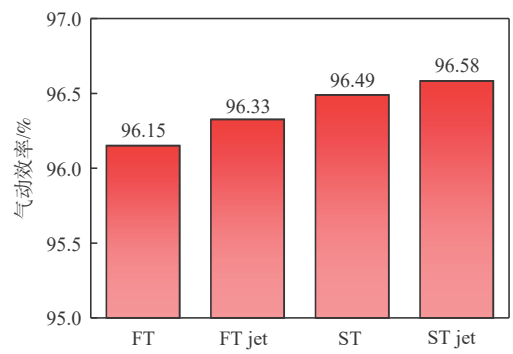
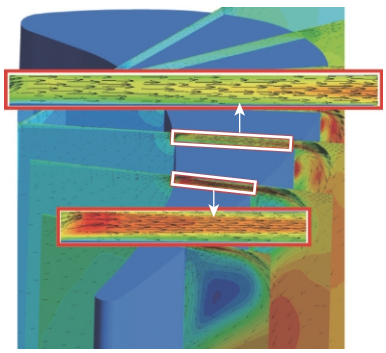


图5 不同条件下叶栅的气动效率

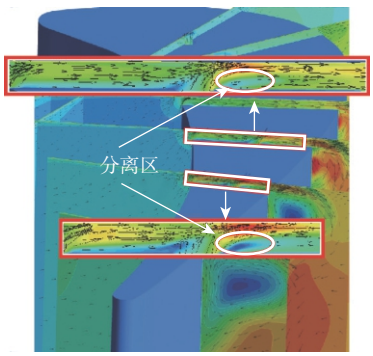
Fig. 5 Aerodynamic performance of cascade under different conditions

对间隙内部的流动状态做具体分析，如图6所示。平叶尖条件下，泄漏流在流入间隙的过程中迅速加速，且加速程度从前到后逐渐增加，当叶栅在设计状态下(出口马赫数为0.78)时，平叶尖中后部的泄漏流在间隙内接近声速，间隙上下边界层随着流动的发展厚度不断增加，泄漏流动最终在达到声速后不再加速。此外，如图6(b)所示，喷气会在孔

下游形成分离区, 并降低泄漏流的流通面积。因此, 喷气会降低喷气附近区域的泄漏流量。图 7 为喷气条件下平叶尖各截面压力系数云图及吸力侧出口马赫数云图, 如图 7(a)所示, 喷出的气体会冲击在机匣上形成高压滞止区, 减小了当地的周向压力梯度, 同时受到喷气的影响, 吸力侧间隙截面上出现了明显的低速区, 其宽度约为喷气孔直径的 2 倍。图 7(b)为图 7(a)中框出的出口截面, 图中的流线还表明, 在平叶尖条件下, 泄漏流动在间隙内近似于垂直中弧线的二维流动。



(a) 无喷气条件



(b) 喷气条件

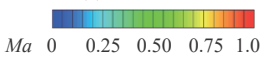
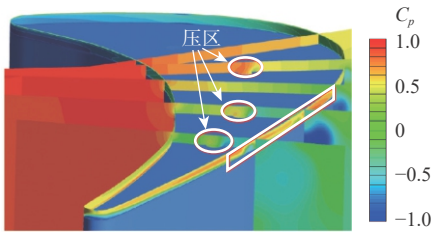
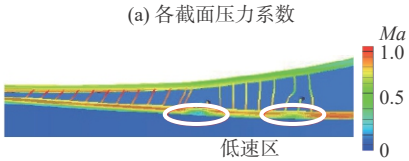


图 6　平叶尖各截面马赫数

Fig. 6　Mach number contours of each section of flat tip



(a) 各截面压力系数

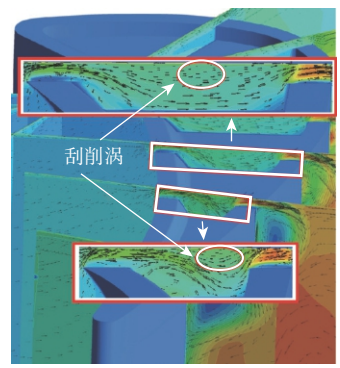


(b) 吸力侧出口马赫数系数

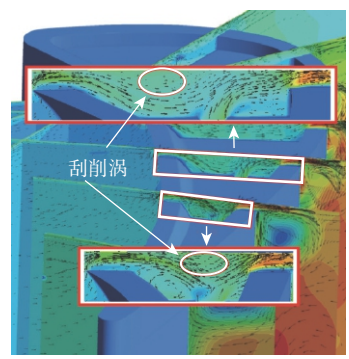
图 7　喷气条件下平叶尖各截面压力系数及吸力侧出口马赫数

Fig. 7　Pressure coefficient contours of each section of flat tip and Mach number contour of tip clearance outlet with tip injection

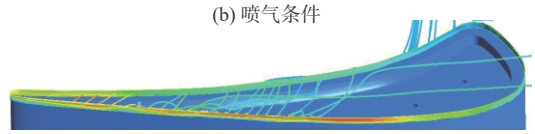
凹槽叶尖条件各截面的马赫数分布如图 8 所示。结果表明, 凹槽叶尖间隙内部马赫数相较平叶尖大幅下降, 整个凹槽内的泄漏流动均保持在亚声速状态, 泄漏流主要的加速位置在吸力侧肋条上方, 随后并进入主流形成泄漏涡。当喷气位置在刮削涡下游时, 刮削涡受到喷气的影响会向上游移动。图 8(c)为凹槽进出口截面马赫数分布及间隙内部流线图, 结果表明凹槽叶尖间隙出口并没有明显的低速区; 无论有无喷气, 刮削涡是凹槽内的主控流动结构。



(a) 无喷气条件



(b) 喷气条件



(c) 喷气条件出口截面

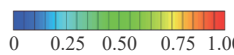


图 8　凹槽叶尖各截面马赫数

Fig. 8　Mach number contours of each section of squealer tip

图 9 为泄漏流量沿轴向的分布, 通过图中对比表明, 平叶尖 FT 的轴向中后部, 泄漏流量在喷气孔附近的位置存在巨大波动, 并总体上显著降低泄漏流量; 喷气孔越接近尾缘, 叶片厚度越小, 这一效果越明显。凹槽叶尖 ST 条件下, 凹槽腔内的刮削涡在轴向 60%～85% 范围内起到了控制泄漏流动的作用, 降低了泄漏流的流量。而凹槽内的喷气对凹槽内部的流动结构和泄漏流的分布影响都较为有限, 并会在轴向 60%～80% 范围内使泄漏流量略有增加。

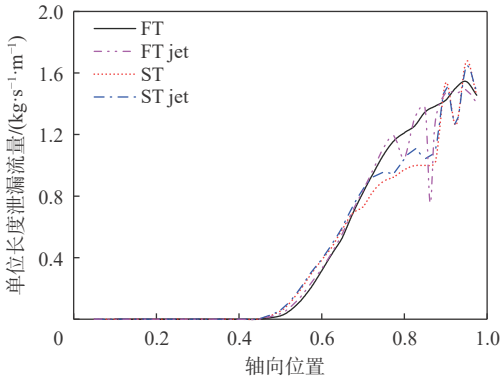


图 9 泄漏流量沿轴向的分布

Fig. 9 Axial wise distribution of leakage flow

以 80% 轴向位置的孔 4 为例，分析喷气吹风比变化对 82.5%～87.5% 位置累积泄漏流量的影响，结果如图 10 和图 11 所示，其纵坐标以无喷气条件下当地的泄漏流量为 1 进行无量纲化。结果表明，平叶尖条件下，随着吹风比的变化泄漏流量基本保持不变，喷气流量的增加值约等于主流流量的减小值；凹槽叶尖条件下，随着吹风比的增加，主流泄漏流量略有减少，总泄漏流量增加。

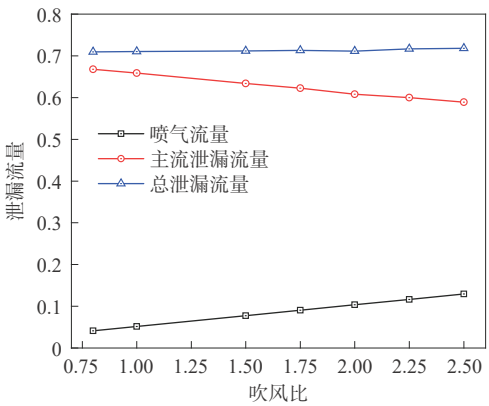


图 10 平叶尖流向 82.5%～87.5% 位置无量纲泄漏流量随孔 4 吹风比变化

Fig. 10 Variation of normalized leakage flow at 82.5%−87.5% posisiton along flat tip with hole 4 blowing ratio

2.2 马赫数对间隙流动状态的影响

随着高压涡轮的负荷不断提高，涡轮通道内的马赫数也随之升高，间隙内的流动有可能进入跨音状态，这会对泄漏损失产生明显的影响。图 12 为叶栅效率随出口马赫数变化趋势，可见，在出口马赫数为 0.7～0.9 范围内时，采用叶顶喷气和凹槽叶尖几何能有效提升叶栅的气动效率，当出口马赫数达到 1.1 时，叶尖的几何和喷气条件对效率的影响基本消失。

图 13 为平叶尖条件下流向中后部各截面马赫数，其中，出口马赫数为 1.1。在 70% 流向位置后，间隙内的流动开始进入跨音状态，且在间隙的进口

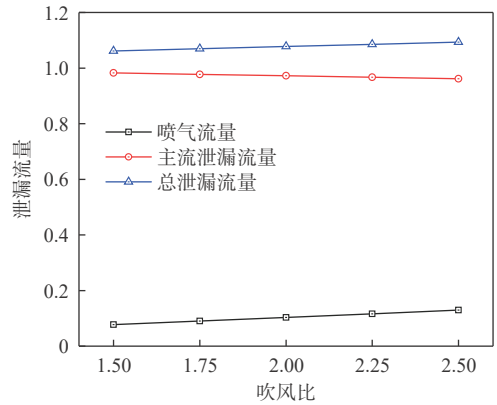


图 11 凹槽叶尖流向 82.5%～87.5% 位置无量纲泄漏流量随孔 4 吹风比变化

Fig. 11 Variation of normalized leakage flow at 82.5%−87.5% posisiton along squealer tip with hole 4 blowing ratio

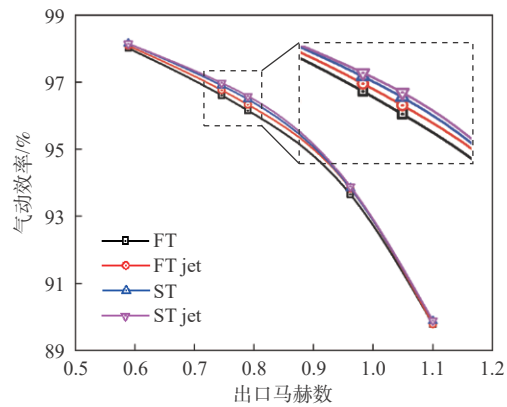


图 12 叶栅气动效率随出口马赫数变化趋势

Fig. 12 Variation of cascade efficiency with outlet Mach number

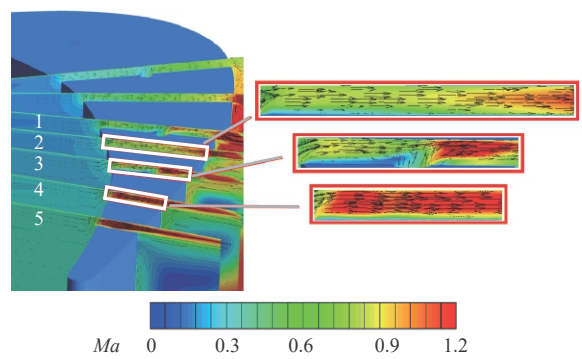


图 13 平叶尖喷气条件下间隙各截面马赫数

Fig. 13 Mach number contours of each section of flat tip with tip injection

和出口出现了复杂的波系结构，气流以超声速状态在间隙出口进入主流。喷气条件下，以截面 3 为例，喷气位置上游保持亚声速状态，喷气下游泄漏流的马赫数则大于 1。此外，喷气仍会导致下游产生一个分离区，并减小了泄漏流的流通面积。

为更好地分析间隙内的流场情况，对间隙内不同截面的轴向密度梯度进行分析，如图 14 所示，其

中, 出口马赫数为 1.1, 截面 1～截面 5 分别为 70%、75%、80%、85%、90% 流向截面, 由图 4(a)的负荷分布特征可知, 从截面 1～截面 5 叶片负荷依次减少。图 14 中, 浅色为密度增大, 深色为密度降低, 则在梯度较大的区域, 认为当地有激波或膨胀波产生。

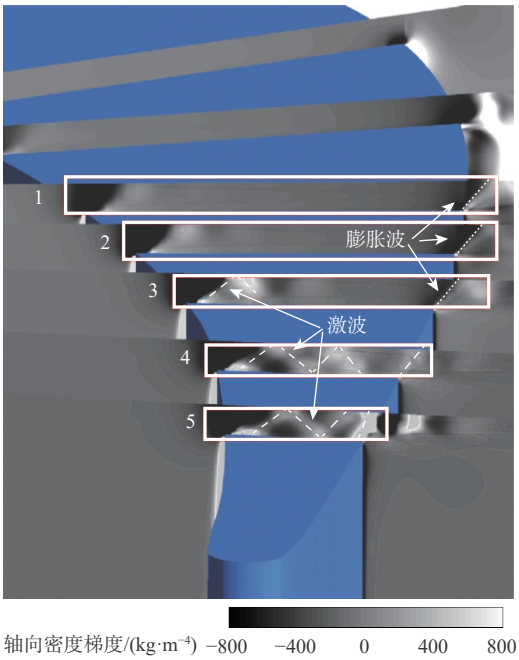


图 14　平叶尖条件间隙内密度梯度

Fig. 14　Density gradient contours of each section of flat tip

具体分析各截面的流动状况如图 15 所示。图 15(a)为 70% 截面流动状况, 该截面也是叶栅负荷最大的位置。泄漏流在进入平叶尖间隙时首先产生一个分离泡, 随后叶尖和端壁壁面的边界层开始发展, 泄漏流的流通面积不断减小, 气流不断加速, 当流道的截面积减小到一定程度时, 泄漏流的速度达到当地声速, 不再继续增加。当泄漏流动接近吸力侧出口时, 由于出口效应, 边界层开始变薄, 泄漏流动的流通面积开始重新增加, 则此时流动变为超声速; 泄漏流在流过分离泡后到间隙出口截面的这一流动过程类似一维拉伐尔喷管中由亚声速加速至超声速的过程。由于此时间隙出口的压力仍高于主流在这一区域的压力, 因此, 在间隙出口处有一道膨胀波。75% 截面处的流动状态与 70% 截面类似, 但由于叶片中后部的厚度不断减小, 没有足够的长度使边界层充分发展, 因此, 泄漏流动刚达到壅塞状态即达到了间隙出口。在 80% 截面, 当叶片厚度近一步减小, 边界层的发展更不充分, 间隙内的喉口位于分离泡斜上方的位置, 由于气流在分离泡上方加速至过膨胀的状态, 在分离泡快消失的位置出现了斜激波。在分离泡下游, 泄漏流动都处于超声速状态。更接近尾缘的在 85% 截面, 随着吸力侧压力的上升, 间隙出口由膨胀波变为激波, 间隙内部也出现了明显的激波反射现象; 其中, 激波反射会导致激波与边界层干涉作用, 使贴近叶尖的边界层迅速增厚, 并降低泄漏流动的马赫数, 最终降低间隙出口的激波强度。在 90% 截面, 叶片已经很薄, 边界层没有足够的空间发展, 间隙内部出现剧烈的激波反射现象, 最终泄漏流动穿过一道角度接近正激波的激波进入主流, 速度由超声速变为亚声速, 随即与主流进行跨音掺混, 由于泄漏流动是由有限空间进入空间较大的主流通道, 因此, 在通过间隙出口的激波后会再次膨胀, 重新变为超声速状态。

对于有喷气的截面, 如图 16 所示, 喷气后形成的分离区的上方会成为该截面流动面积最小的位置, 也就是喉口处, 在喉口的上游处于亚声速状态, 下游则处于超声速状态。

对凹槽叶尖而言, 如图 17 所示, 虽然泄漏流在间隙出口的马赫数在前中部沿流向不断增加, 凹槽内的流动速度仍保持在亚声速状态, 泄漏流直至吸力侧肋条上方才变为超声速进入主流。图 18 为凹槽叶尖 80% 流向截面流动示意图, 泄漏流在流经吸力侧肋条时形成的分离区的上方是泄漏流动过程中横截面积最小的位置。

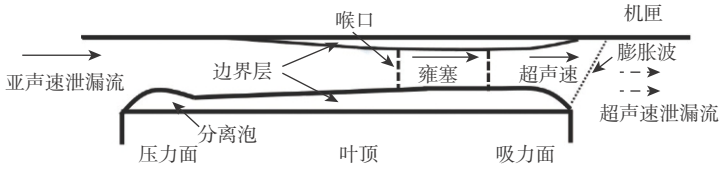
图 19 为在出口马赫数为 1.1 条件下, 不同叶尖条件下单位轴向长度泄漏流量沿流向的分布, 分布趋势与设计工况下类似, 说明喷气对平叶尖有较大影响, 而对凹槽叶尖的影响较小。通过对比图 19 与图 9, 可以认为跨声速条件下间隙内各截面已达到流通能力的极值, 负荷的变化对泄漏流量的影响很小。图 19 中轴向中后部轴向长度泄漏流量沿流向分布的变化主要是由于单位轴向长度对应的弧长长度越来越大导致的。

2.3　跨音环境考虑喷气条件的叶尖泄漏预测模型

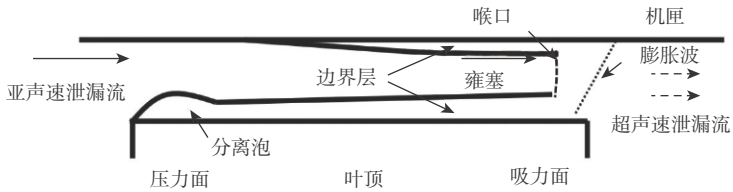
在涡轮转子通道中, 泄漏损失与泄漏量的大小总体上呈正相关关系, 建立精确的泄漏流预测方法是准确预测泄漏损失的前提。本节基于气体动力学方程与相关工程流体力学公式, 建立跨声速条件下泄漏流预测模型。

由 2.2 节分析可知, 平叶尖和凹槽叶尖的泄漏流动在间隙内的特征完全不同, 因此, 对这 2 种情况分别进行讨论。

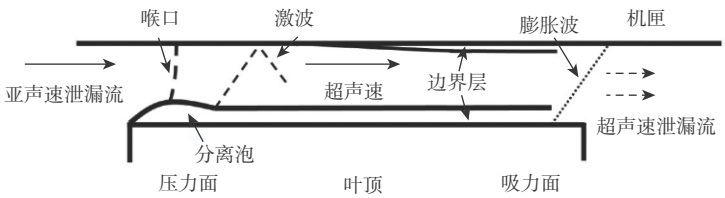
对平叶尖间隙内进行分析, 由图 7 分析可知, 将平叶尖泄漏流动在各个截面近似于二维流动。首先判断间隙内的流动是否达到临界状态(Ma=1)。在间隙前中部, 间隙出口位置为最小截面, 整个流动可以近似为收缩喷管。以该裁面压力侧压力为



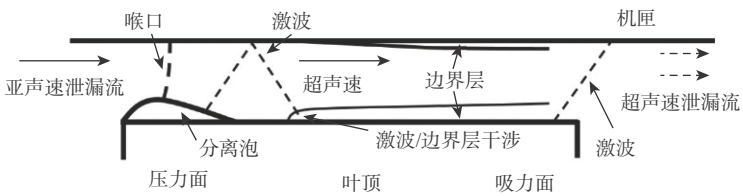
(a) 70%截面1流向截面



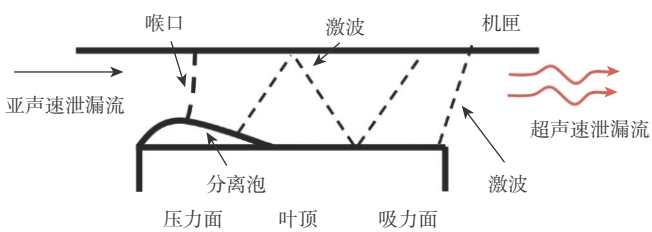
(b) 75%截面2流向截面



(c) 80%截面3流向截面



(d) 85%截面4流向截面



(e) 90%截面5流向截面

图 15 平叶尖各截面泄漏流流动示意图

Fig. 15 Flow state diagram at each section of flat tip

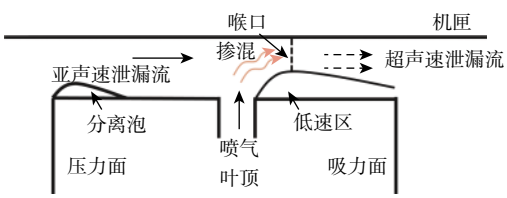


图 16 平叶尖截面 3 喷气条件下泄漏流流动示意图

Fig. 16 Flow state diagram in section 3 under flat tip jet condition

间隙进口总压 p_{in}^{*}，吸力侧压力为出口静压 p_{out}，通过气体动力学方程可知，当 $p_{in}^{*}/p_{out}\geqslant 1.892\,9$ 时，间隙内为跨声速状态，反之，则处于亚声速状态。

平叶尖条件下，分离泡导致的维纳收缩效应使得泄漏流动通道相对收缩，而后的边界层在分离泡尾端的基础上发展而来。

层流边界层厚度方程为[35-36]

$$\delta \approx 5.0\left(\frac{\gamma x}{U_{\infty}}\right)^{0.5} \tag{6}$$

式中：γ为运动黏度；x为下游位置距平板前缘距离；U_{∞}为泄漏流速度。

将叶片厚度 d 代入式(6)，当 $d>6\tau$ 时，有：$\delta>\tau$，则间隙内呈典型的平板间间隙流动，速度剖面近似呈抛物线型，则整个间隙的等效流通面积 $\sigma=2/3\tau$。

而当叶片趋近尾缘区域时，叶片厚度减小，上下边界层发展的厚度不及分离泡的厚度时，喉口位置从出口位置移动到分离泡上方，分离泡下游出现激波，并出现激波边界层干涉效应，间隙内边界层

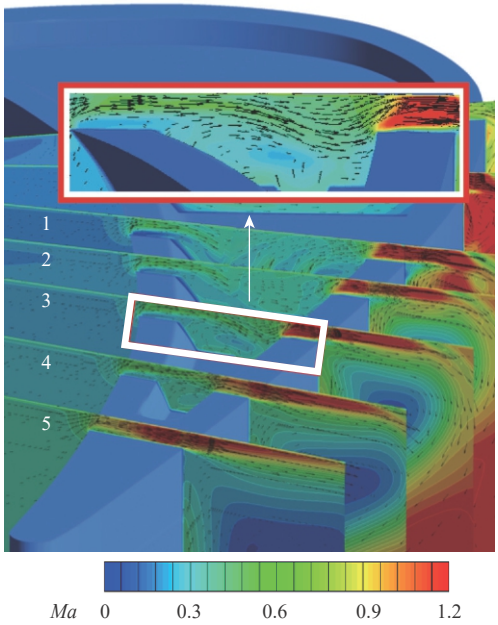


图 17　凹槽叶尖喷气条件下间隙各截面马赫数分布(出口 *Ma*=1.1)

Fig. 17　Mach number contours of each section of squealer tip with tip injection (Outlet *Ma*=1.1)

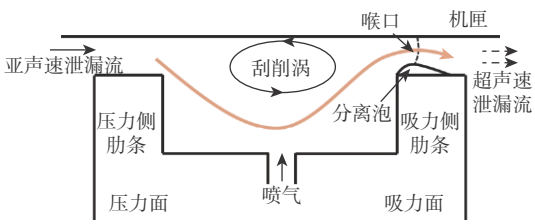


图 18　凹槽叶尖喷气条件下裁面 3 流动示意图

Fig. 18　Flow state diagram in section 3 with jet of squealer tip

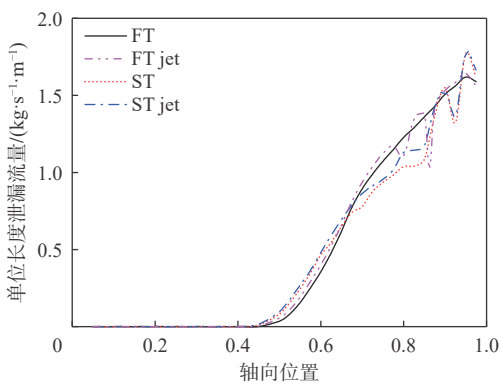


图 19　泄漏流动沿轴向分布

Fig. 19　Axial wise leakage axial distribution

厚度显著减小, 当上壁面反射的激波与叶顶的边界层干涉时, 叶顶边界层流动呈现分离再附情况, 转捩为湍流边界层, 且厚度从 0 开始重新增长。边界层厚度计算式为

$$\delta_{99}=0.037\,0\left(\frac{\gamma}{U_\infty}\right)^{0.2}x^{0.8} \tag{7}$$

式中: 下标 99 表示边界层边界为主流速度的 99%。

此时间隙内部流动变为缩放喷管, 喉口位置为间隙入口分离泡上端, 如图 15(c)所示, 此时可认为喉口面积为$\sigma\approx 0.8\tau$。

对各阶段的损失进行评估, 流体流入间隙的速度为v_{in}, 实际流动中, 泄漏流流入间隙的速度约等于主流在压力面侧的速度, 有$v_{\mathrm{in}}\approx v_{\mathrm{ps}}$。由工程流体力学知识可知, 维纳收缩效应会导致的总压损失为[36]

$$\Delta p^*_{\mathrm{loss1}}=0.5v^2_{\mathrm{ps}} \tag{8}$$

并由式(8), 即求出间隙内的进口总压为

$$p^*_{\mathrm{in}}=p^*-\Delta p^*_{\mathrm{loss1}} \tag{9}$$

当间隙内不存在跨音流动时, 出口 *Ma* 可由式(10)计算得到:

$$\frac{p_{\mathrm{ss}}}{p^*_{\mathrm{in}}}=\left(1+\frac{k-1}{2}Ma\right)^{-\frac{k}{k-1}} \tag{10}$$

式中: p_{ss} 为间隙吸力侧出口静压; k 为常数, 取 1.4; 则该截面泄漏流量为

$$m_{\mathrm{leak}}=K\frac{p^*_{\mathrm{in}}}{\sqrt{T^*}}A_{\mathrm{e}}q(Ma) \tag{11}$$

式中: T^*为气流进口总温; K 为常数, 取 0.040 4 s·K$^{0.5}$/m; $A_{\mathrm{e}}=\sigma$; $q(Ma)$为流量函数。

当间隙内达到临界状态时, $q(Ma)=1$, 截面泄漏流量为

$$m_{\mathrm{leak}}=K\frac{p^*_{\mathrm{in}}}{\sqrt{T^*}}A_{\mathrm{e}} \tag{12}$$

对于凹槽叶尖, 喉口处为吸力侧肋条上方, 则需考虑 2 段肋条的维纳收缩效应和凹槽内的掺混损失, 其中维纳收缩效应的损失为

$$\Delta p^*_{\mathrm{loss2}}=0.5v^2_{\mathrm{ps}}+0.5v^2_{\mathrm{cav}} \tag{13}$$

式中: v_{cav} 为凹槽内泄漏流流动速度, 可由射流理论求出 v_{cav}, 如下:

$$v_{\mathrm{cav}}=\left(\frac{\tau}{\tau+h}\right)^{0.5}v_{\mathrm{ps}} \tag{14}$$

考虑$v_{\mathrm{in}}\approx v_{\mathrm{ps}}$, 则凹槽内的掺混损失为

$$\Delta p^*_{\mathrm{loss3}}=\left(1-\frac{\tau}{\tau+h}\right)^2v^2_{\mathrm{ps}} \tag{15}$$

则吸力侧肋条上方的总压为

$$p^*_{\mathrm{out}}=p^*-\Delta p^*_{\mathrm{loss2}}-\Delta p^*_{\mathrm{loss3}} \tag{16}$$

与平叶尖类似, 当 $p^*_{\mathrm{in}}/p_{\mathrm{out}}\geqslant 1.892\,9$ 时, 肋条上方为跨声速状态, 反之, 则处于亚声速状态。利用式(10)可求出泄漏流的出口 *Ma*, 进而求出流出间隙的速度v_{out}。最后, 考虑凹槽内部流动对泄漏流动的影响, 泄漏流在流向方向具备一定的速度, 其值约等于v_{ps}, 可求得泄漏流动与流向方向的夹角θ为

$$\theta = \arccos\left(\frac{v_{ps}}{v_{out}}\right) \tag{17}$$

则截面泄漏流量为

$$m_{leak} = K\frac{p^*_{out}}{\sqrt{T^*}}A_e q(Ma)\sin\theta \tag{18}$$

由于喉口位置为吸力侧肋条分离泡上端，有 $A_e \approx 0.8\tau$。

在叶片尾缘处叶尖不再有凹槽部分，此时采用平叶尖的计算方法来估算当地截面的泄漏流量。

考虑叶顶喷气的情况，在平叶尖条件下，由 2.1 节分析结果表明，叶栅中前部的喷气对泄漏流分布的影响有限，中后部的喷气孔在流向的影响范围约为 2 倍的喷气孔直径，在这一范围内，泄漏流流量约为无喷气条件下的 2/3。对于凹槽叶尖，喷气泄漏流动的影响较小，根据图 11 的结果，可以估计泄漏流量的增量约为喷气流量值的 2/3。

基于本节分析，对各截面泄漏流量沿流向方向积分，可得到叶尖累积泄漏流量为

$$M_{leak} = \int_0^s m_{leak}\mathrm{d}s \tag{19}$$

其预测结果与模拟计算值对比如图 20 所示。结果表明本文模型整体预测与数值模拟计算结果的误差小于 5%，可以准确预测泄漏流量。

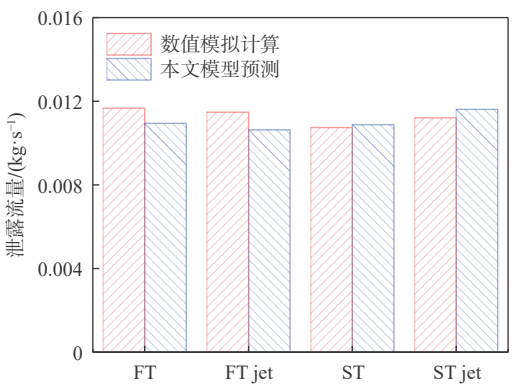


图 20 本文模型预测与数值模拟计算结果

Fig. 20 Proposed model prediction and numerical simulation calculation results

3 结 论

1）跨音条件下，叶顶喷气可以增加平叶尖叶栅的气动效率，而对凹槽叶尖性能的提升较小。平叶尖条件下，轴向中后部位置的叶顶喷气可以在喷气孔上方区域滞止形成高压区，减小泄漏流动的驱动力，并在孔下游形成分离区域，有效降低泄漏流动的流通面积。对于凹槽叶尖，叶顶喷气对刮削涡结构的影响整体有限，刮削涡仍是凹槽内的主控流动结构。

2）平叶尖条件下，喷气流量的改变对总泄漏量影响有限，喷气流量的增加值和主流泄漏量的减小值基本相等。凹槽叶尖条件下，喷气流量越大，总泄漏量越大。无论有无喷气，凹槽叶尖的气动性能均优于平叶尖。

3）随着叶栅负荷增大，平叶尖间隙内出现跨音流动情况。当叶片厚度大于 6τ 时，由于间隙上下壁面边界层不断发展，有效流通面积不断减小，间隙内呈壅塞状态；在接近尾缘区域，叶片厚度开始减小，间隙内出现激波，此时间隙进口的分离泡上方为整个流动的喉口位置。凹槽叶尖条件下，吸力面肋条上游区域全部保持在亚声速状态，气流在吸力面肋条上方迅速膨胀，最终以超声速状态进入主流。

4）在明确跨音条件下平叶尖和凹槽叶尖间隙内流动规律的基础上，综合考虑叶片负荷、叶片厚度和叶顶喷气的影响，并考虑间隙内的流动损失，利用气体动力学计算式给出各截面的泄漏流量，从而建立可用于跨音条件下的泄漏流量预测模型。

参考文献（References）

[1] DENTON J D. The 1993 IGTI scholar lecture: Loss mechanisms in turbomachines[J]. Journal of Turbomachinery, 1993, 115(4): 621-656.

[2] HAN J C. Fundamental gas turbine heat transfer[J]. Journal of Thermal Science and Engineering Applications, 2013, 5(2): 021007.

[3] ZOU Z P, WANG S T, LIU H X, et al. Axial turbine aerodynamics for aero-engines[M]. Berlin: Springer, 2018.

[4] RAINS D A. Tip clearance flows in axial flow compressors and pumps[D]. Pasadena: California Institute of Technology, 1954.

[5] BINDON J P. The measurement and formation of tip clearance loss[J]. Journal of Turbomachinery, 1989, 111(3): 257-263.

[6] YARAS M I, SJOLANDER S A. Prediction of tip-leakage losses in axial turbines[C]//Proceedings of ASME 1990 International Gas Turbine and Aeroengine Congress and Exposition. Brussels: International Gas Turbine Institute , 2015.

[7] GAO J, ZHENG Q, XU T B, et al. Inlet conditions effect on tip leakage vortex breakdown in unshrouded axial turbines[J]. Energy, 2015, 91: 255-263.

[8] HUANG A. Loss mechanisms in turbine tip clearance[D]. Cambridge : Massachusetts Institute of Technology, 2011.

[9] VIRDI A S, ZHANG Q, HE L, et al. Aerothermal performance of shroudless turbine blade tips with relative casing movement effects[J]. Journal of Propulsion and Power, 2015, 31(2): 527-536.

[10] BUNKER R S. Axial turbine blade tips: Function, design, and durability[J]. Journal of Propulsion and Power, 2006, 22(2): 271-285.

[11] 邹正平, 姚李超, 轩笠铭, 等. 涡轮转子凹槽叶尖泄漏流动气动热力特征[J]. 推进技术, 2020, 41(9): 1975-1987.
ZOU Z P, YAO L C, XUAN L M, et al. Aero-thermodynamic characteristics of tip leakage flow in turbine rotor with squealer tip[J]. Journal of Propulsion Technology, 2020, 41(9): 1975-1987(in Chinese).

[12] ZOU Z P, SHAO F, LI Y R, et al. Dominant flow structure in the squealer tip gap and its impact on turbine aerodynamic performance[J]. Energy, 2017, 138: 167-184.

[13] 杜金霖, 邹正平, 黄霖, 等. 计及机匣相对运动的涡轮叶片叶顶凹槽流动研究[J]. 实验流体力学, 2021, 35(2): 22-35.
DU J L, ZOU Z P, HUANG L, et al. Experimental and numerical study of squealer tip flow field considering relative casing motion[J]. Journal of Experiments in Fluid Mechanics, 2021, 35(2): 22-35(in Chinese).

[14] KRISHNABABU S K, NEWTON P J, DAWES W N, et al. Aerothermal investigations of tip leakage flow in axial flow turbines—part I: Effect of tip geometry and tip clearance gap[J]. Journal of Turbomachinery, 2009, 131(1): 011006.

[15] KEGALJ M, SCHMID G, WARTZEK F, et al. Experimental and numerical investigation of tip leakage flow in a 1 1/2 stage turbine rig comparing flat and cavity-squealer tip geometries[C]//Proceedings of ASME Turbo Expo 2012: Turbine Technical Conference and Exposition. Copenhagen: International Gas Turbine Institute, 2012.

[16] MOORE J, MOORE J G, HENRY G S, et al. Flow and heat transfer in turbine tip gaps[J]. Journal of Turbomachinery, 1989, 111(3): 301-309.

[17] JIANG S J, LI Z G, LI J, et al. Numerical investigations on the heat transfer performance of transonic squealer tip with different film cooling layouts[C]//Proceedings of ASME Turbo Expo 2020: Turbomachinery Technical Conference and Exposition. [S. l.] : International Gas Turbine Institute , 2021.

[18] WANG T Y, XUAN Y M, HAN X S. The effects of tip gap variation on transonic turbine blade tip leakage flow based on VLES approach[J]. Aerospace Science and Technology, 2021, 111: 106542.

[19] WHEELER A P S, ATKINS N R, HE L. Turbine blade tip heat transfer in low speed and high speed flows[J]. Journal of Turbomachinery, 2011, 133(4): 041025.

[20] MOORE J, ELWARD K M. Shock formation in overexpanded tip leakage flow[J]. Journal of Turbomachinery, 1993, 115(3): 392-399.

[21] ARISI A, XUE S, NG W F, et al. Numerical investigation of aerothermal characteristics of the blade tip and near-tip regions of a transonic turbine blade[C]//Proceedings of ASME Turbo Expo 2014: Turbine Technical Conference and Exposition. Düsseldorf : International Gas Turbine Institute , 2014.

[22] WHEELER A P S, SALEH Z. Effect of cooling injection on transonic tip flows[J]. Journal of Propulsion and Power, 2013, 29(6): 1374-1381.

[23] DUAN P H, HE L. Optimization of turbine cascade squealer tip cooling design by combining shaping and flow injection[J]. Journal of Turbomachinery, 2021, 143(11): 111007.

[24] MA H, ZHANG Q, HE L, et al. Cooling injection effect on a transonic squealer tip—part II: Analysis of aerothermal interaction physics[J]. Journal of Engineering for Gas Turbines and Power, 2017, 139(5): 052507.

[25] WANG Y B, SONG Y P, YU J Y, et al. Effect of cooling injection on the leakage flow of a turbine cascade with honeycomb tip[J]. Applied Thermal Engineering, 2018, 133: 690-703.

[26] ZHOU Z H, CHEN S W, LI W H, et al. Thermal performance of blade tip and casing coolant injection on a turbine blade with cavity and winglet-cavity tip[J]. International Journal of Heat and Mass Transfer, 2019, 130: 585-602.

[27] 胡建军, 张铎, 张香兰, 等. 一种复合方法抑制叶尖泄漏的试验及模拟研究[J]. 北京航空航天大学学报, 2018, 44(11): 2283-2291.
HU J J, ZHANG D, ZHANG X L, et al. Experiment and simulation study on tip leakage suppression by a compound method[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(11): 2283-2291(in Chinese).

[28] 王维杰, 卢少鹏, 马海腾, 等. 高压涡轮尾切凹槽叶尖冷却换热特性[J]. 航空动力学报, 2019, 34(10): 2131-2139.
WANG W J, LU S P, MA H T, et al. Cooling and heat transfer characteristics of high-pressure turbine blade with cutback squealer tip[J]. Journal of Aerospace Power, 2019, 34(10): 2131-2139(in Chinese).

[29] 杨蓓洁, 谭晓茗, 单勇, 等. 肋条结构对气膜冷却凹槽叶尖流动换热的影响[J]. 航空动力学报, 2021, 36(7): 1462-1471.
YANG B J, TAN X M, SHAN Y, et al. Effects of rib structure on flow and heat transfer characteristics of film cooling rotor blade with squealer tip[J]. Journal of Aerospace Power, 2021, 36(7): 1462-1471(in Chinese).

[30] GAO J, ZHENG Q, ZHANG Z Y, et al. Aero-thermal performance improvements of unshrouded turbines through management of tip leakage and injection flows[J]. Energy, 2014, 69: 648-660.

[31] ZHOU C, HODSON H. The tip leakage flow of an unshrouded high pressure turbine blade with tip cooling[J]. Journal of Turbomachinery, 2011, 133(4): 041028.

[32] NICHOLSON J H, FOREST A E, OLDFIELD M L G, et al. Heat transfer optimised turbine rotor blades: An experimental study using transient techniques[C]//Proceedings of ASME 1982 International Gas Turbine Conference and Exhibit. London: International Gas Turbine Institute, 2015.

[33] ZENG F, DU J L, HUANG L, et al. An experimental method for squealer tip flow field considering relative casing motion[J]. Chinese Journal of Aeronautics, 2020, 33(7): 1942-1952.

[34] YOUNG J B, HORLOCK J H. Defining the efficiency of a cooled turbine[J]. Journal of Turbomachinery, 2006, 128(4): 658-667.

[35] SCHLICHTING H. Boundary-layer theory[M]. 7th ed. New York: McGraw-Hill, 1979.

[36] MUNSON B R, YOUNG D F, OKIISHI T H. Fundamentals of fluid mechanics[M]. New York: Wiley, 1990.

Analyzing and modeling flow in tip clearance of transonic turbine rotor

XUAN Liming[1, 2], ZOU Zhengping[2, 3, *], ZENG Fei[1, 2, 4]

(1. School of Energy & Power Engineering, Beihang University, Beijing 100191, China;
2. National Key Laboratory of Science and Technology on Aero-Engine Aero-Thermodynamics, Beijing 102206, China;
3. Research Institute of Aero-Engine, Beihang University, Beijing 102206, China;
4. Hunan Aviation Powerplant Research Institute, Zhuzhou 412002, China)

Abstract: In order to clarify the mechanism of the tip leakage flow in transonic turbines and further improve turbine efficiency, the effect of tip injection on the performance of a flat tip and a squealer tip is studied under the transonic condition, and the flow state in the clearance between a flat tip and squealer tip is discussed. The results show that under transonic conditions, tip jet can increase the aerodynamic efficiency of the flat tip cascade, and that the scraping vortex is still the dominant flow structure in the cavity of the squealer tip. The increase of jet flow has limited effect on the total leakage flow of the flat tip, but will increase the total leakage flow of the squealer tip. With higher load, the transonic flow in the flat tip clearance is related to the blade load and thickness. For squealer tips, the leakage flow expands rapidly to the supersonic state above the suction side squealer. Finally, a leakage flow prediction model for transonic conditions is established.

Keywords: transonic turbine; leakage flow; tip clearance; tip injection; flow model

Received: 2021-10-26; **Accepted**: 2022-01-11; **Published Online**: 2022-01-29 14:25
URL: kns.cnki.net/kcms/detail/11.2625.V.20220129.1034.001.html
Foundation item: National Natural Science Foundation of China (51676005)
* **Corresponding author**. E-mail: zouzhengping@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0654

车路耦合荷载下沥青混凝土路面振动响应

董倩[1,2]，程少锋[1]，张献民[1,3,*]，包伊婷[1]

(1. 中国民航大学 交通科学与工程学院，天津 300300；2. 天津大学 建筑工程学院，天津 300072；
3. 南京航空航天大学 民航学院，南京 210016)

摘　　要：为研究沥青混凝土路面的振动特性，采用滤波白噪声法拟合出路面平整度时域模型，基于1/4车辆模型，考虑车辆–道路的耦合作用，分析不同车辆速度、不同路面平整度等级时，车辆对路面的实时动荷载；建立道路三维有限元模型，研究车辆随机动荷载作用下道路的振动响应，分析道路各结构层参数对道路表面中心振动基频的影响。结果表明：路面平整度等级、车辆行驶速度对车–路耦合系统影响显著，当路面平整度等级由A级变化至C级时，同一行驶速度下的车辆动荷载增大了20%；道路振动基频随土基动模量呈对数关系增加，土基动模量由60 MPa增大至260 MPa，道路振动基频由5.61 Hz增大至10.80 Hz，振动基频增幅高达48.06%；在常用动模量变化范围内，面层、基层、垫层的动模量对基频的影响较小；随着面层、基层与垫层厚度的增加，道路基频呈线性减小的趋势，面层厚度对振动频率影响的敏感性大于基层厚度，基层厚度对振动频率影响的敏感性大于垫层厚度；在常用结构层厚度变化范围内，振动基频分别减小9.28%、18.05%与12.75%。试验结果证明：振动基频计算较正确，计算结果可为道路承载力快速测试提供理论支持。

关　键　词：沥青混凝土路面；车辆动荷载；有限元；结构层厚度；结构层弹性模量；振动基频
中图分类号：U411
文献标志码：A　　**文章编号**：1001-5965(2023)09-2385-10

目前，国内外常用的路面承载力测试方法在模拟加载、中断交通、模量反算结果上有时存在较大误差[1-5]等弊端。为解决上述弊端，近年来国内外学者提出通过在道路表面布置各种传感器，采集车辆通过时道路的变形等信息，以此反算道路结构承载力的方法[6-10]，并取得了一定的研究成果。Lajnef等[6]利用应变等智能传感器收集道路在日常运营中的变形等信息，实现了对道路结构的长期监测；Nielsen等[8]通过在已建成的道路表面布置加速度传感器，使用固定加载车加载，收集道路振动信息，实现根据振动加速度反算道路基顶模量；董倩[9]利用加速度传感器采集飞机荷载作用下跑道表面振动加速度，通过对加速度数据进行二次积分获得路面的弯沉值，再借助计算机模拟计算等手段，确定结构承载力；Ryynänen等[10]研究了使用传感器来加强公路设施管理的技术，并指出在长期监测工程中，加速度传感器是最稳定的。但随着研究深入，Arraigada等[11]发现加速度二次积分所得的误差无法消除，当车速较低时，对于距离车辆位置较远的传感器，加速度积分得到的动位移精度不满足要求。研究表明，加速度数据快速傅里叶变换获得的振动基频直接快速，可以避免二次积分带来的误差，且结构振动基频与其本身强度有直接关系[12-13]。因此，可探索通过道路振动频率反算其承载力的技术途径与理论模型。

基于此，本文建立了1/4车辆模型，采用滤波白噪声法拟合路面的平整度，考虑车路耦合作用，分析了不同车辆速度、不同路面平整度等级下车辆对

收稿日期：2021-11-01；**录用日期**：2022-01-21；**网络出版时间**：2022-02-15 14:36
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220214.1748.012.html
基金项目：国家自然科学基金(51178456)；中央高校基本科研业务费(3122017039)
* **通信作者**. E-mail：xmzhang@cauc.edu.cn

引用格式：董倩，程少锋，张献民，等. 车路耦合荷载下沥青混凝土路面振动响应[J]. 北京航空航天大学学报，2023，49(9)：2385-2394. DONG Q，CHENG S F，ZHANG X M，et al. Vibration response of asphalt concrete pavement under vehicle-road coupled load[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2385-2394 (in Chinese).

路面的实时动荷载，将此动荷载作用于路面结构三维有限元模型，分析了路面中心位置处的振动基频与各结构层厚度、弹性模量间的相关关系，相应结果旨在为路面结构的检测评估提供理论参考。

1 不同等级路面平整度曲线

平整度是评价路面施工质量和服务水平的一个重要指标，美国试验与材料协会（American Society for Testing and Materials, ASTM）将其定义为道路表面相对于理想平面的竖向偏差。国内外学者研究道路表明，路面平整度是具有零均值、各态历经平稳的高斯随机过程[14]，滤波白噪声法可模拟此过程[15]。因此，将路面偏离水平面的高低波动描述为满足一定条件的高斯白噪声，拟合出路面平整度激励的时域模型[16]。基于滤波白噪声的路面平整度一阶微分方程如下：

$$\dot{x}_g(t) = -2\pi n_1 v_0 x_g(t) + 2\pi n_0 \sqrt{G_q(n_0)v_0}\,\omega(t) \quad (1)$$

式中：n_1为空间频率的下限，取 0.01 m^{-1}；$\omega(t)$为标准高斯白噪声；$x_g(t)$为经滤波白噪声生成的路面平整度时域信号；v_0为车辆行驶速度，m/s；n_0为基准空间频率，取 0.1 m^{-1}；$G_q(n_0)$为基准空间频率n_0对应的功率谱密度 (power spectral density, PSD)，m^2/m^{-1}，长春汽车研究所编制的《车辆振动输入路面平度表示方法》中[17]，规定了不同路面等级下的 PSD，如表 1 所示[18]，其中，n_0=0.1 m^{-1}。

表 1 不同等级路面功率谱密度[18]

Table 1 Grading standard of road roughness by PSD[18] m^2/m^{-1}

等级	下限值	几何平均值	上限值
A	8	16	32
B	32	64	128
C	128	256	512
D	512	1 024	2 048
E	2 048	4 096	8 192
F	8 192	16 384	32 768
G	32 768	65 536	131 072
H	131 072	262 144	524 288

在 MATLAB/Simulink 环境下，建立基于滤波白噪声的路面平整度仿真模型，如图 1 所示，包括一个限带白噪声模块，2 个增益模块，一个积分器和一个道面时域模型观察窗。因需要的限带白噪声单边功率谱密度为 1，但 MATLAB/Simulink 的限带白噪声模块假设的是双边功率谱，因此，限带白噪声的噪声功率谱密度应设置为 0.5。限带白噪声模块的采样时间设置为 1/10v，该采样频率能保证 1 s 内车辆以滑行速度 v 行驶过的路程内包含 10 个采样点，与标准规定的最大空间采样频率 10 m^{-1} 一致。

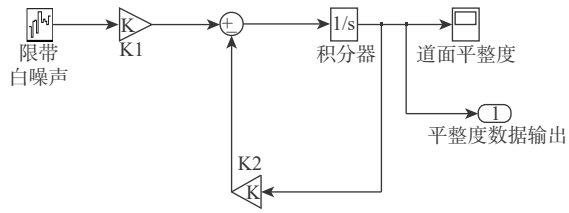


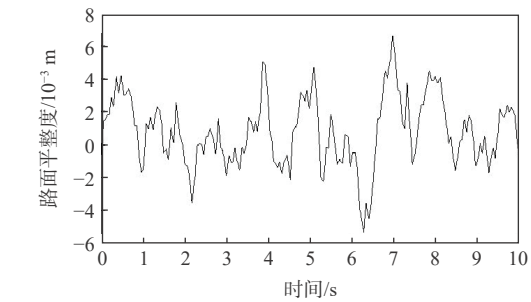
图 1 路面平整度时域仿真模型

Fig. 1 Time domain model of pavement roughness

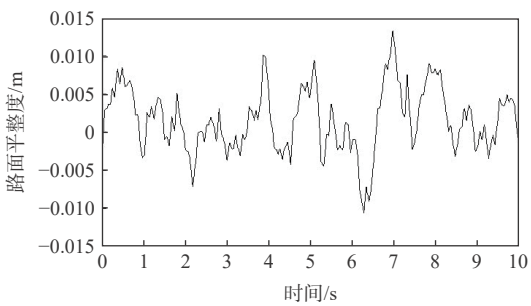
左起第 1 个增益模块 K1 大小为$2\pi n_0\sqrt{G_q(n_0)v_0}$；第 2 个增益模块 K2 大小为$2\pi n_1 v_0$。

应用路面平整度时域仿真模型，输入对应参数，当行车速度为 20 m/s 时，可得到 A、B、C 级路面平整度时域仿真曲线如图 2 所示。

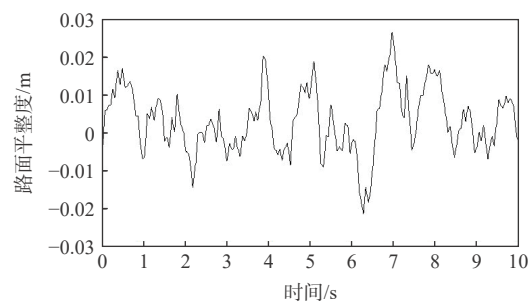
由图可以看出，随着路面等级变差，A、B、C 这 3 级路面的平整度值变大。求解模拟所得 3 种路面的平整度曲线进行功率谱密度分析，可发现与本级别道面标准功率谱密度符合较好。



(a) A级



(b) B级



(c) C级

图 2 不同等级路面的平整度曲线

Fig. 2 Pavement roughness of three grades

2　随机荷载的求解

车辆行驶于路面时, 路面平整度的随机性是车辆产生随机动荷载的根本原因。与此同时, 车身系统的振动将引起道路结构的变形及振动, 而道路的变形与振动将反过来加强车身系统的振动, 这种车路系统的互相影响即为车路系统的耦合作用。本节首先建立二自由度车辆质阻弹系统, 建立振动平衡方程, 然后将路面平整度作为约束条件, 求解车辆随机动荷载。

2.1　车辆平衡方程的建立

研究表明, 1/4 车辆模型在车辆垂直振动方面的研究很成功[19], 因此, 采用公路工程中常用的 1/4 车辆模型, 模型有 2 个自由度, 如图 3 所示。

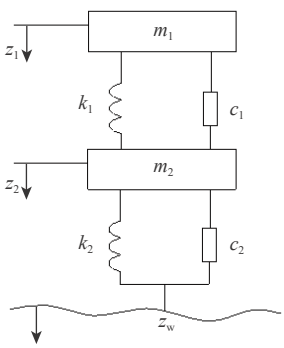


图 3　二自由度车辆模型

Fig. 3　Vehicle model with two-degree-of-freedom

二自由度 1/4 车辆模型振动微分方程如下:

$$\begin{cases} m_1\ddot{z}_1 + c_1(\dot{z}_1-\dot{z}_2) + k_1(z_1-z_2) = 0 \\ m_2\ddot{z}_2 - c_1(\dot{z}_1-\dot{z}_2) - k_1(z_1-z_2) + \\ \quad c_2(\dot{z}_2-\dot{z}_w) + k_2(z_2-z_w) = 0 \\ p_t = k_2(z_2-z_x) + c_2(\dot{z}_2-\dot{z}_w) - m_w\ddot{z}_w \end{cases} \tag{2}$$

式中: m_1和m_2分别为簧载质量和非簧载质量, kg; k_1和k_2分别为簧载刚度和非簧载刚度, N/m; c_1和c_2分别为簧载和非簧载阻尼, (N·s)/m^{-1}; m_w为轮胎质量, kg; z_1为簧载部分绝对垂直位移, m; z_2为非簧载部分绝对垂直位移, m; z_w为轮胎绝对垂直位移, m; p_t为车辆作用于路面上的动荷载, N。

2.2　路面平衡方程的建立

根据弹性多层体系理论建立沥青混凝土路面的有限元模型进行分析研究[20]。假设: ① 各层都由均匀、各向同性的材料组成, 其泊松比和弹性模量分别为μ_i和E_i; ② 土基在水平方向和向下的深度方向均为无限, 其上的各结构层在水平方向无限, 深度方向为有限; ③ 层间完全连续接触。则路面平衡方程如下:

$$\boldsymbol{m}\ddot{\boldsymbol{y}} + \boldsymbol{c}\dot{\boldsymbol{y}} + \boldsymbol{k}\boldsymbol{y} = \boldsymbol{p} \tag{3}$$

式中: $\boldsymbol{m}$、$\boldsymbol{c}$和$\boldsymbol{k}$分别为路面有限元模型的总体质量、阻尼和刚度矩阵; $\boldsymbol{y}$、$\dot{\boldsymbol{y}}$、$\ddot{\boldsymbol{y}}$分别为路面模型的位移、速度和加速度矢量; $\boldsymbol{p}$为作用于路面模型上的荷载矢量, 包括随机动荷载与静荷载。

2.3　车辆–路面系统平衡方程的建立

车辆–路面耦合状态下轮胎竖向位移是路面振动产生的竖向位移与路面平整度之和, 表达式为

$$Z_w = Z_{(x)} + y_w \tag{4}$$

式中: $Z_{(x)}$为路面不平整函数; y_w为路面模型中车辆轮胎作用位置处路面竖向位移; Z_w为车辆轮胎总的竖向位移。

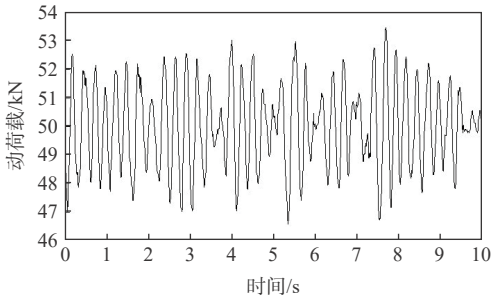
根据式 (2)～式 (4) 可以得到车路耦合系统的振动微分方程。ANSYS 中建立 1/4 车辆模型, 以随机的路面平整度作为约束条件, 车速的变化通过加载时间调节, 采用 Full 法求解车辆随机动荷载。常用 10 t 加载车参数如表 2 所示, 轮组结构为单轴双轮组。当加载车以 20 m/s 的速度行驶于 A、B、C 级路面时, 以 30 m/s 的速度行驶于 B 级路面时, 加载车每个轮胎上动荷载如图 4 所示。

由图可知, 当速度不变时, 车辆正常行驶过程中产生的动荷载随着路面平整度情况的变差而变大, 路面平整度等级直接影响着车路之间的相互作用力, 当路面等级由 A 到 C 时, 车辆动荷载最大值由 54 kN 增大到 64 kN 左右, 增大了 20%。当车辆行驶于同一平整度等级的路面结构时, 随车辆行驶速度的增大, 其动荷载也增大, 汽车由 20 m/s 增大至 30 m/s, 车辆动荷载由 56.2 kN 增大至 60 kN。

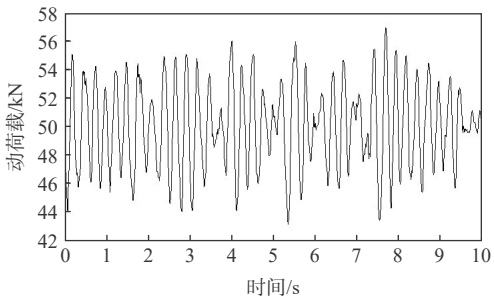
表 2　10 t 加载车辆参数

Table 2　Parameters of truck with 10 tons' weight

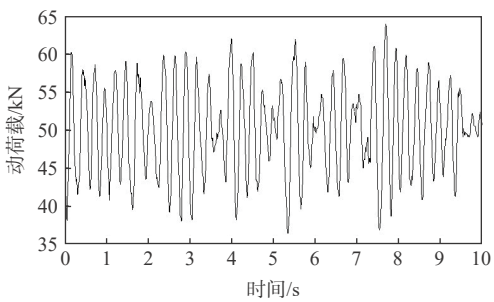
参数	数值
簧载部分质量m_1/kg	4 450
非簧载部分质量m_2/kg	550
簧载部分刚度系数k_1/(N·m^{-1})	1 000 000
非簧载部分刚度系数k_2/(N·m^{-1})	1 750 000
簧载部分阻尼系数c_1/(N·s·m^{-1})	15 000 000
非簧载部分阻尼系数c_2/(N·s·m^{-1})	2 000 000



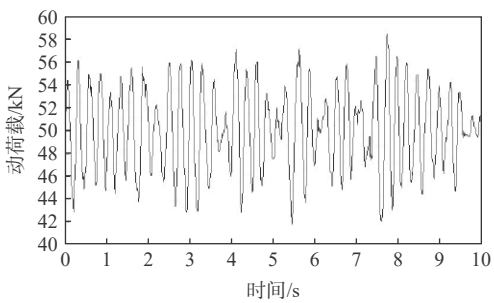
(a) A级, 速度为20 m/s



(b) B级, 速度为20 m/s



(c) C级, 速度为20 m/s



(d) B级, 速度为30 m/s

图 4 A、B、C 级路面随机荷载

Fig. 4 Random load of grade A, B and C

3 有限元模型的建立

3.1 路面结构参数

研究表明, 在动态荷载条件下, 采用动态模量能更好的反映结构受力状态。当温度为 15 ℃、胎压为 0.7 MPa、车速在 60～80 km/h 的车辆荷载作用下, 沥青混凝土动静态模量之间统计关系式[21] 为

$$\begin{cases} E_{d1} = 2.445E_{s1} \\ R^2 = 0.984 \end{cases} \quad (5)$$

式中: E_{s1}为沥青混凝土静态模量, MPa; E_{d1}为沥青混凝土动态模量, MPa; R^2 为回归系数。

水泥稳定碎石动静态模量之间统计关系[22] 为

$$\begin{cases} E_{d2} = 1.486E_{s2} + 463.03 \\ R^2 = 0.832 \end{cases} \quad (6)$$

式中: E_{s2}为水泥稳定碎石基层静态模量, MPa; E_{d2}为水泥稳定碎石基层动态模量, MPa。

二灰稳定碎石动静态模量之间统计关系[22] 为

$$\begin{cases} E_{d3} = 1.814E_{s3} - 353.2 \\ R^2 = 0.864 \end{cases} \quad (7)$$

式中: E_{s3}为二灰稳定碎石基层静态模量, MPa; E_{d3}为二灰稳定碎石基层动态模量, MPa。

土基动静态模量转换关系[23] 为

$$E_{s4} = 0.33E_{d4} \quad (8)$$

式中: E_{s4}为土基静态模量, MPa; E_{d4}为土基动态模量, MPa。

将路面结构简化为面层、基层、垫层和地基构成的弹性层状体系, 根据实际情况及《公路沥青路面设计规范》[24], 选取合理的路面结构参数, 如表 3 所示。

表 3 路面结构和材料参数

Table 3 Structural and material parameters of road

结构层	厚度/m	动态模量/MPa	泊松比	密度/(kg.m^{-3})
沥青混凝土	0.20	4 000	0.30	2 400
水泥稳定碎石	0.40	8 000	0.25	2 200
二灰稳定碎石	0.30	2 000	0.30	2 000
土基	9	140	0.35	1 900

3.2 单元尺寸及约束条件

基于弹性层状体系理论建立道路三维有限元模型, 通过试算确定道路模型尺寸为 46 m×9.9 m×16 m(x×y×z)[25-26]。x 方向为行车方向, y 方向为路面结构的深度方向, z 方向为路面横断面宽度方向。

面层、基层、底基层网格尺寸为 0.23 m×0.1 m×0.16 m, 土基网格尺寸为 0.46 m×0.3 m×0.32 m。网格划分后的模型如图 5 所示。边界条件的确定参

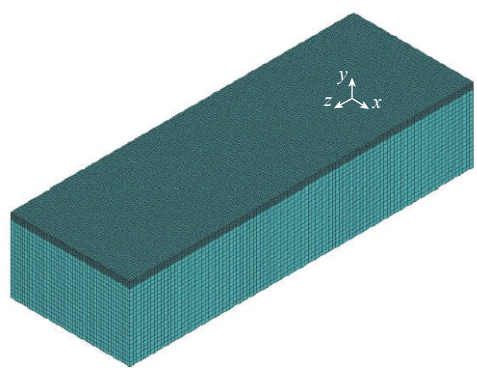


图 5 道路有限元模型

Fig. 5 Finite element model of pavement

考文献 [27]: x=0 m 及 x=46 m 处只约束 x 方向; y=0 m 处固定约束; z=0 m 及 z=16 m 处只约束 z 方向。

假设车辆沿道路中心线行驶, 采取阶跃加载对单元组施加均匀分布荷载, 单元组的面积即为轮胎接地面积[26]。考虑到边界影响效应, 从沿路面 x 轴方向 11.5 m 开始加载, 直至 x=34.5 m 时结束加载, 车辆行车速度可通过加载时间调节, 如图 6 所示。

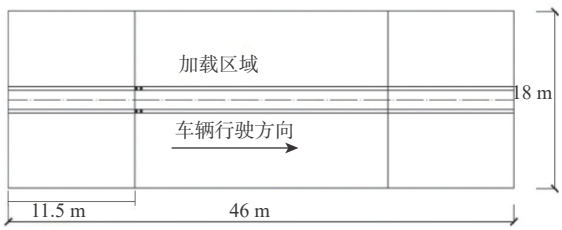


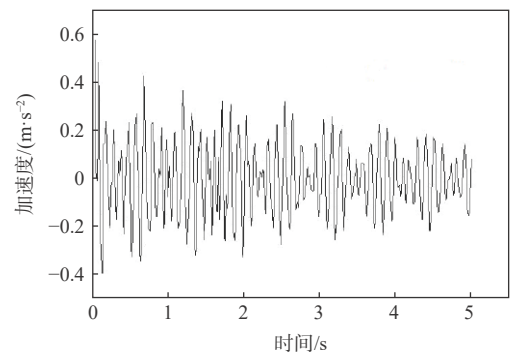
图 6　车辆移动荷载加载轨迹

Fig. 6　loading path of vehicle load

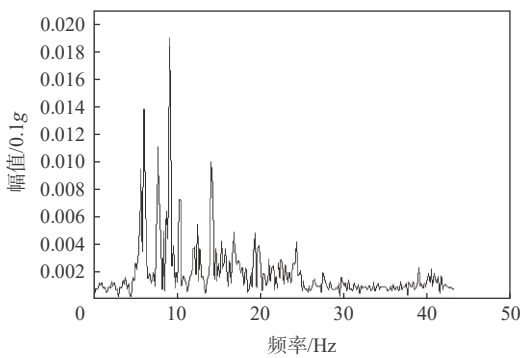
4　计算结果分析

4.1　车辆移动荷载作用下的振动基频

选取路面中心点为观测点, 提取汽车移动荷载作用下观测点处的振动加速度, 对加速度数据进行快速傅里叶变换, 可得到道路振动的一阶频率为 6.08 Hz, 如图 7 所示, 其中, E_{d4}=140 MPa。



(a) 加速度



(b) 振动基频

图 7　路面中心位置处加速度与振动基频

Fig. 7　Acceleration and fundamental frequency at center of road surface

提取面层、基层、垫层及土基不同深度处的振动加速度数据, 经过快速傅里叶变换得到道路由上至下不同深度处的振动基频及振幅, 如表 4 所示, 其中 x=23, z=8。由表 4 可知, 不同深度处的振动基频不变, 仅幅值逐渐减小, 且在路基中减小速率最大。究其原因为, 随着深度增加, 振动能量逐步消散, 消散速度与材料阻尼等呈正相关。

表 4　不同深度处振动基频及幅值

Table 4　Fundamental frequencies and amplitudes at different depths

y/m	振动基频/Hz	幅值/0.1g
9.9	6.08	0.014
9.7	6.08	0.013
9.3	6.08	0.013
9.0	6.08	0.013
7.0	6.08	0.010
5.0	6.08	0.009
3.0	6.08	0.007
1.0	6.08	0.004

4.2　面层厚度与模量对道路振动基频的影响

逐一变化各结构层某一参数, 其他参数均保持不变, 探讨道路表面中心位置处一阶振动频率随参数的变化规律。

先探讨面层参数对振动的影响。土基动态模量取 100 MPa, 水泥稳定碎石基层、二灰稳定碎石垫层参数与表 3 中保持一致, 沥青混凝土面层厚度 H_1由 0.14 m 增加至 0.22 m, 动态模量E_{d1}由 3 500 MPa 增加至 5 500 MPa 时, 对应的道路基频统计如表 5 所示。由表 5 可知: 当沥青混凝土面层模量为 3 500 MPa, 其厚度由 0.14 m 增加至 0.22 m 时, 道路表面中心位置处基频由 6.57 Hz 减少到 5.96 Hz, 减

表 5　不同道面参数时的振动基频

Table 5　Fundamental frequencies under different pavement parameters

面层厚度/m	面层模量/MPa	振动基频/Hz	面层厚度/m	面层模量/MPa	振动基频/Hz
0.14	3 500	6.57	0.20	3 500	6.08
	4 000	6.58		4 000	6.08
	4 500	6.58		4 500	6.08
	5 000	6.58		5 000	6.09
	5 500	6.59		5 500	6.09
0.16	3 500	6.40	0.22	3 500	5.96
	4 000	6.40		4 000	5.96
	4 500	6.40		4 500	5.97
	5 000	6.40		5 000	5.98
	5 500	6.41		5 500	5.98
0.18	3 500	6.26			
	4 000	6.26			
	4 500	6.27			
	5 000	6.27			
	5 500	6.27			

少了 0.61 Hz, 减幅为 9.28%。对表 4 数据进行回归分析可知, 当面层弹性模量不变时, 振动基频 f 与面层厚度 H_1 呈线性负相关。当面层厚度不变, 动弹性模量由 3 500 MPa 增大至 5 500 MPa, 道路表面中心位置处基频变化较小。

在沥青混凝土道路结构中, 面层弹性模量的变化对道路基频的影响微弱, 面层厚度对道路基频的影响相对较大。究其原因为: 道路振动基频反映的是车辆荷载影响范围内道路结构的整体刚度, 而对于道路这类连续且在纵向上无限延伸的结构, 面层材料参数的变化对结构整体强度影响微弱。

4.3 基层模量和基层厚度

当土基动态模量为 100 MPa, 沥青混凝土面层、二灰稳定碎石垫层参数与表 3 中保持一致, 基层厚度 H_2 由 0.3 m 增大至 0.5 m, 间隔步距取 0.05 m; 基层动态模量 E_{d2} 由 2 000 MPa 增大至 14 000 MPa, 间隔步距取 3 000 MPa。道路表面中心位置的振动频率如图 8 所示。

由图可知, 当基层动态模量为 8 000 MPa, 随着面层厚度的增加, 道路表面中心位置的振动基频由 6.65 Hz 减小至 5.45 Hz, 减少了 1.2 Hz, 减幅为 18.05%。对同一模量下不同基层厚度的振动基频数据回归分析, 可知道路表面中心位置处的振动基频 f 与基层厚度 H_2 呈负相关关系。当基层厚度不变, 随着模量的增加, 道路基频变化很小。

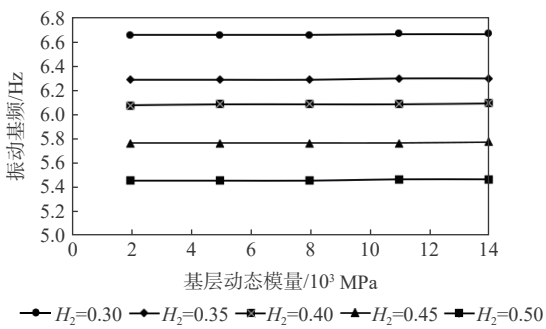


图 8 基层厚度 H_2 和动态模量 E_{d2} 对基频的影响

Fig. 8 Effect of base thickness H_2 and modulus E_{d2} on fundamental frequency

4.4 垫层模量和基层厚度

当土基动态模量为 100 MPa, 沥青混凝土面层、水泥稳定碎石基层参数与表 3 中保持一致, 垫层厚度 H_3 由 0.2 m 增大至 0.4 m, 间隔步距取 0.05 m; 垫层动态模量 E_{d3} 由 1 000 MPa 增大至 5 000 MPa, 间隔步距取 1 000 MPa。道路表面中心位置的振动频率如图 9 所示。

由图可知: 当垫层动态模量为 3 000 MPa, 随着面层厚度的增加, 道路表面中心位置的振动基频由 6.51 Hz 减小至 5.68 Hz, 减少了 0.83 Hz, 减幅为 12.75%。对同一模量下不同垫层厚度的振动基频数据回归分析, 可知道路表面中心位置处的振动基频 f 与垫层厚度 H_3 呈负相关关系。当垫层厚度不变, 随着模量的增加, 道路基频几乎不变。

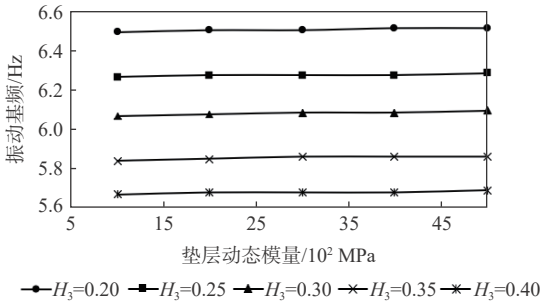


图 9 垫层厚度 H_3 和动态模量 E_{d3} 对基频的影响

Fig. 9 Effect of cushion thickness H_3 and modulus E_{d3} on fundamental frequency

绘制出振动基频与面层、基层、垫层厚度间的相关关系, 如图 10 所示, 回归表达式如下:

$$\begin{cases} f = -7.70H_1 + 7.64 \\ R^2 = 0.996\,8 \end{cases} \tag{9}$$

$$\begin{cases} f = -5.84H_2 + 8.38 \\ R^2 = 0.994\,8 \end{cases} \tag{10}$$

$$\begin{cases} f = -4.18H_3 + 7.32 \\ R^2 = 0.997\,7 \end{cases} \tag{11}$$

式中: f 为振动基频, Hz; H_1 为面层厚度, m ; H_2 为面层厚度, m ; H_3 为垫层厚度, m。

由图 10 及式(9)～式(11)可知, 由上到下 3 个结构层回归曲线的斜率分别为 7.7、5.84、4.18, 说明由上至下各结构层厚度的变化对振动基频影响的敏感性逐渐降低。这与振动能量的消散规律密不可分, 相比于基层、垫层, 面层是振动能量消散最快的区域, 因此, 面层厚度的变化对振动基频的影响更加敏感, 这与文献 [28] 的研究结论一致。

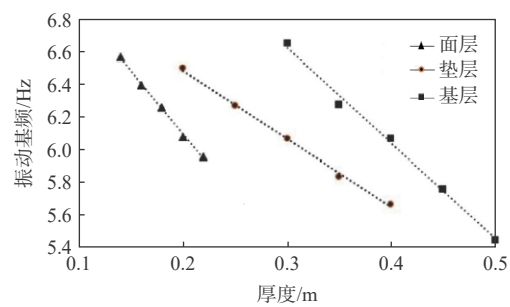


图 10 各结构层厚度-基频回归曲线

Fig. 10 Thickness-fundamental frequency regression curve of each structural layer

4.5 土基模量

土基动模量 E_{d4} 由 60 MPa 增大至 260 MPa, 其他结构层参数与表 3 保持一致。不同土基动模量下道路表面中心位置的振动基频如表 6 所示, 土基

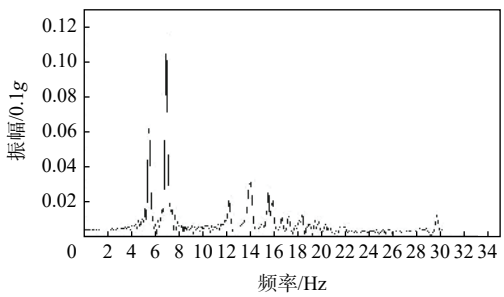
动态模量分别为 60 MPa、260 MPa 时, 频谱曲线如图 11 所示, 根据频谱曲线可读取振动基频。

从表 6 可看出, 土基动态模量从 60 MPa 增大至 260 MPa, 道路表面中心位置处的振动基频由 5.61 Hz 增大至 10.80 Hz, 增大了 5.19 Hz, 增幅高达 48.06%。对不同土基动态模量下的振动基频进行回归分析, 可知振动基频与土基动态模量呈对数关系, 随着土基动模量的增加, 基频增大的幅度逐渐变小, 如图 12 所示。究其原因为随土基动态模量增大, 结构整体刚度增大, 荷载扩散更快, 使得参与振动的土基厚度减少。

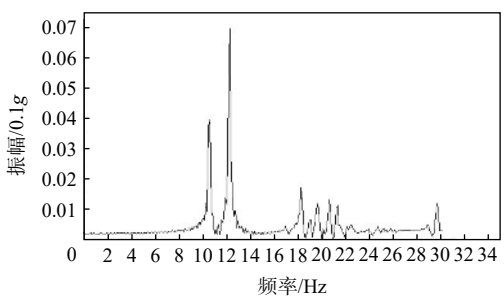
表 6　不同土基动态模量下道路基频

Table 6　Fundamental frequency under different soil foundation dynamic moduli

土基动态模量/MPa	基频/Hz	土基动态模量/MPa	基频/Hz
60	5.61	180	9.10
100	7.01	220	9.88
140	8.19	260	10.80



(a) E_{d4}=60 MPa



(b) E_{d4}=260 MPa

图 11　不同土基动态模量时频谱曲线

Fig. 11　Spectrum curve with different dynamic modulus of soil foundation

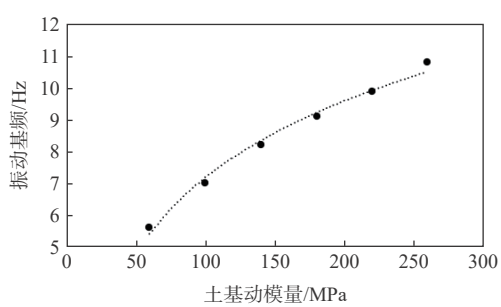


图 12　振动基频-土基动态模量回归曲线

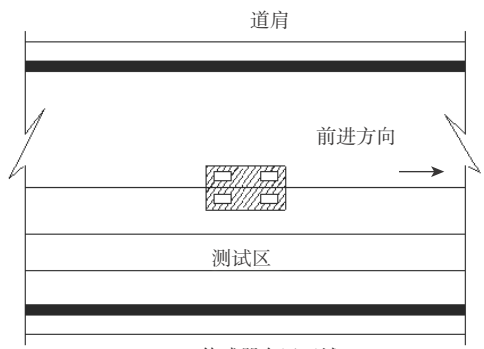
Fig. 12　Regression curve of fundamental frequency-dynamic modulus of soil foundation

振动基频回归表达式为

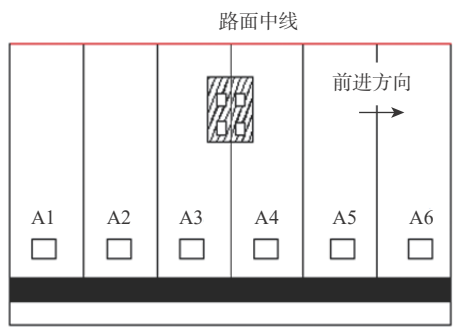
$$\begin{cases} f = 3.488 \ln E_{d4} - 8.887 \\ R^2 = 0.988\,9 \end{cases}$$

5　现场试验

选取河北省某等级为 A 级的新建沥青混凝土道面进行实测分析, 在测试区内布置传感器, 采集加载车通过时路面动态响应信号, 分析道路的振动基频情况, 传感器的布置情况如图 13 所示, 其中, A1～A6 为传感器。



(a) 传感器布置区域



(b) 传感器布置示意图

(c) 现场布置

图 13　传感器布置

Fig. 13　Layout of sensors

由重 2.4 t、后轴重 1.935 t 的测试车加载, 以 40 km/h 的速度匀速通过待测区域。试验路段的结构层构成及厚度如表 7 所示。试验场地地基经过了统一的压实处理。

加载车辆经过时，A1 传感器采集到的振动加速度信号如图 14 所示，对信号进行快速傅里叶变换，得到道面结构的振动基频为 8.74 Hz。

表 7 道路结构及参数

Table 7 Structures and Parameters of road

结构层	静态模量/MPa	动态模量/MPa	厚度/m	密度/($kg \cdot m^{-3}$)
沥青混凝土面层	2 000	4 890	0.2	2 400
水泥稳定碎石基层	5 000	7 893.03	0.4	2 200
二灰稳定碎石底基层	800	1 098	0.3	2 000
土基	55	167	∞	1 900

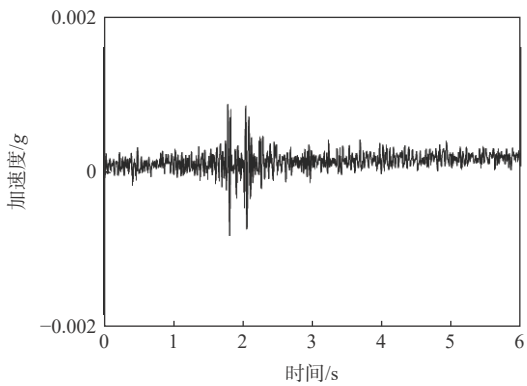


图 14 A1 传感器的加速度响应曲线

Fig. 14 Acceleration response curve at A1 sensor

对采集的加速度数据进行频谱分析，可知其他传感器所在位置处的振动基频如表 8 所示。由表可知，传感器 A1～A6 所在位置处，振动基频基本不变，分别为 8.64 Hz 与 8.74 Hz，这是由数据采集的误差导致的。

表 8 中数据表明了计算结果的正确性，同时验证了本文方法用于结构承载力的评估的可行性。

表 8 道路振动基频

Table 8 Fundamental frequencies of test road

传感器	振动基频/Hz	传感器	振动基频/Hz
A1	8.74	A4	8.64
A2	8.64	A5	8.64
A3	8.74	A6	8.74

6 结 论

1）路面平整度等级、车辆行驶速度对车路耦合系统影响显著。当路面平整度等级由 A 级变化至 C 级时，同一行驶速度下的车辆动荷载增大了 20% 左右；当汽车行驶于同一等级路面、速度由 20 m/s 增大至 30 m/s 时，动荷载增加 6.3% 左右。

2）土基动态模量对道路表面中心位置处的振动基频影响最大，当土基动态模量由 60 MPa 增大至 260 MPa，振动基频增幅高达 48.06%，振动频率随土基动态模量呈对数关系增加。

3）面层厚度、基层厚度、垫层厚度与道路表面中心位置处的振动基频呈线性负相关规律变化，且斜率逐渐降低，说明由上至下 3 个结构层厚度的变化对振动基频影响的敏感度是逐步降低的，符合振动能量的消散。而面层、基层、垫层的动态模量在常用范围内变化时，对此位置振动基频的影响较小。

4）路基为道路结构的薄弱部分，通常是路基的损坏或强度降低导致道路承载力的降低。现场试验过程中，通过测试车辆荷载作用下道路表面中心位置处的振动基频便可直观快速的获得道路路基强度，进而对道路的承载力进行快速评估。但本文方法目前的研究成果不能精确判断面层、基层、垫层的强度，需在后续的研究中进一步深化。

5）在新建道路上开展现场试验，验证了振动基频计算的正确性。

参考文献（References）

[1] WANG H, XIE P Y, JI R, et al. Prediction of airfield pavement responses from surface deflections: Comparison between the traditional backcalculation approach and the ANN model[J]. Road Materials and Pavement Design, 2021, 22(9): 1930-1945.

[2] TAREFDER R A, AHMED M U. Modeling of the FWD deflection basin to evaluate airport pavements[J]. International Journal of Geomechanics, 2014, 14(2): 205-213.

[3] HAMIM A, YUSOFF N I M, CEYLAN H, et al. Comparative study on using static and dynamic finite element models to develop FWD measurement on flexible pavement structures[J]. Construction and Building Materials, 2018, 176: 583-592.

[4] LI M Y, WANG H. Prediction of asphalt pavement responses from FWD surface deflections using soft computing methods[J]. Journal of Transportation Engineering Part B-Pavements, 2018, 144(2): 04018014.

[5] LI M Y, WANG H. Development of ANN-GA program for backcalculation of pavement moduli under FWD testing with viscoelastic and nonlinear parameters[J]. International Journal of Pavement Engineering, 2019, 20(4): 490-498.

[6] LAJNEF N, RHIMI M, CHATTI K, et al. Toward an integrated smart sensing system and data interpretation techniques for pavement fatigue monitoring[J]. Computer-Aided Civil and Infrastructure Engineering, 2011, 26(7): 513-523.

[7] SKAR A, KLAR A, LEVENBERG E. Load-independent characterization of plate foundation support using high-resolution distributed fiber-optic sensing[J]. Sensors, 2019, 19(16): 3518.

[8] NIELSEN J, LEVENBERG E, SKAR A. Inference of pavement properties with roadside accelerometers[C]//Proceedings of the 9th International Conference on Maintenance and Rehabilitation of Pavements. Berlin: Springer, 2020: 719-728.

[9] 董倩. 基于飞机滑行刚性道面位移场的跑道承载力研究[D]. 天津: 中国民航大学, 2013.

DONG Q. Studies on carrying capacity of rigid pavement based on displacement field in conditions of aircraft taxing[D]. Tianjin: Civil Aviation University of China, 2013 (in Chinese).

[10] RYYNÄNEN T, PELLINEN T, BELT J. The use of accelerometers in the pavement performance monitoring and analysis[J]. IOP Conference Series:Materials Science and Engineering, 2010, 10: 012110.

[11] ARRAIGADA M, PARTL M N, ANGELONE S M, et al. Evaluation of accelerometers to determine pavement deflections under traffic loads[J]. Materials and Structures, 2009, 42(6): 779-790.

[12] 张存巍. 基于车辆和道路结构耦合振动的道路振动频率研究[D]. 天津: 中国民航大学, 2016.
ZHANG C W. Studies on road surface vibration frequency based on vehicle-road coupling vibration system[D]. Tianjin: Civil Aviation University of China, 2016 (in Chinese).

[13] 薛海, 李强. 地铁车辆天线梁振动加速度及动应力试验[J]. 北京交通大学学报, 2015, 39(4): 33-36.
XUE H, LI Q. Test study on vibration and dynamic stress of subway vehicle 's antenna beam[J]. Journal of Beijing Jiaotong University, 2015, 39(4): 33-36(in Chinese).

[14] 张献民, 胡鹏. 随机荷载作用下刚性路面动态响应研究[J]. 振动与冲击, 2015, 34(19): 126-130.
ZHANG X M, HU P. Dynamic response of a rigid pavement under random loads[J]. Journal of Vibration and Shock, 2015, 34(19): 126-130(in Chinese).

[15] 殷珺, 陈辛波, 吴利鑫, 等. 滤波白噪声路面时域模拟方法与悬架性能仿真[J]. 同济大学学报 (自然科学版), 2017, 45(3): 398-407.
YIN J, CHEN X B, WU L X, et al. Simulation method of road excitation in time domain using filtered white noise and dynamic analysis of suspension[J]. Journal of Tongji University (Natural Science), 2017, 45(3): 398-407(in Chinese).

[16] 雷继超, 石鑫刚, 蔡良才, 等. 滤波白噪声法的单轮起落架滑跑模型[J]. 空军工程大学学报(自然科学版), 2020, 21(3): 12-18.
LEI J C, SHI X G, CAI L C, et al. A quarter landing gear taxiing model based on filtered white noise method[J]. Journal of Air Force Engineering University (Natural Science Edition), 2020, 21(3): 12-18(in Chinese).

[17] 中华人民共和国机械工业部. 机械振动 道路路面谱测量数据报告: GB/T 7031-2005[S]. 北京: 中国质检出版社, 1995.
Ministry of Machine-Building and Electronics Industry. Mechanical vibration-road surface profiles-reporting of measured data: GB/T 7031-2005[S]. Beijing: China Quality Inspection Press, 1995(in Chinese).

[18] 赵济海, 王哲人, 关朝雳. 路面不平度的测量分析与应用[M]. 北京: 北京理工大学出版社, 2000.
ZHAO J H, WANG Z R, GUAN C L. Measurement analysis and application of road roughness[M]. Beijing: Beijing Insititute of Technology Press, 2000(in Chinese).

[19] 周玉民, 谈至明, 刘伯莹. 1/4车-路耦合动力学模型研究[J]. 同济大学学报(自然科学版), 2012, 40(3): 408-413.
ZHOU Y M, TAN Z M, LIU B Y. Quarter vehicle-road coupling dynamics models[J]. Journal of Tongji University (Natural Science), 2012, 40(3): 408-413(in Chinese).

[20] 李倩, 刘俊卿. 基于车-路相互作用的沥青路面平整度劣化研究[J]. 振动与冲击, 2018, 37(6): 76-81.
LI Q, LIU J Q. Asphalt pavement evenness deterioration analysis based on the vehicle-pavement interaction[J]. Journal of Vibration and Shock, 2018, 37(6): 76-81(in Chinese).

[21] 孔伟斌. 行车荷载作用下路面结构动态响应敏感性分析[J]. 北方交通, 2017(3): 87-90.
KONG W B. The analysis on pavement structure dynamic response sensitivity under the action of vehicle load[J]. Northern Communications, 2017(3): 87-90(in Chinese).

[22] 沙爱民, 胡力群. 半刚性基层材料的结构特征[J]. 中国公路学报, 2008, 21(4): 1-5.
SHA A M, HU L Q. Structural characteristics of semi-rigid base course material[J]. China Journal of Highway and Transport, 2008, 21(4): 1-5(in Chinese).

[23] 曾胜. 路面性能评价与分析方法研究[D]. 长沙: 中南大学, 2003.
ZENG S. Research on the pavement performance evaluation and analyze methods[D]. Changsha: Central South University, 2003(in Chinese).

[24] 中华人民共和国交通运输部. 公路沥青路面设计规范: JTG D50—2017[S]. 北京: 人民交通出版社, 2017.
Ministry of Tran Sport of the People's Republic of China. Specifications for design of highway asphalt pavement: JTG D50—2017[S]. Beijing: China Communications Press, 2017 (in Chinese).

[25] 张献民, 张存巍, 张靖. 水泥混凝土面层结构的振动响应试验研究[J]. 公路交通科技, 2016, 33(9): 1-6.
ZHANG X M, ZHANG C W, ZHANG J. Experimental study on vibration response of cement concrete pavement surface course[J]. Journal of Highway and Transportation Research and Development, 2016, 33(9): 1-6(in Chinese).

[26] 程少锋. 基于道路振动频率快速检测道面承载能力的研究[D]. 天津: 中国民航大学, 2018.
CHENG S F. Research on fast detection of pavement bearing capacity based on vibration frequency[D]. Tianjin: Civil Aviation University of China, 2018(in Chinese).

[27] 卢正, 王长柏, 付建军, 等. 交通荷载作用下公路路基工作区深度研究[J]. 岩土力学, 2013, 34(2): 316-321.
LU Z, WANG C B, FU J J, et al. Research on influence depth of road subgrade induced by vehicle loads[J]. Rock and Soil Mechanics, 2013, 34(2): 316-321(in Chinese).

[28] 付梓君. 车—路耦合作用下沥青路面振动规律及能量谱表征研究[D]. 西安: 长安大学, 2018.
FU Z J. Study on the vibration regularity and energy spectrum characterization of asphalt pavement under the coupling of vehicles and roads[D]. Xi'an: Changan University, 2018 (in Chinese).

Vibration response of asphalt concrete pavement under vehicle-road coupled load

DONG Qian[1, 2], CHENG Shaofeng[1], ZHANG Xianmin[1, 3, *], BAO Yiting[1]

(1. College of Transportation Science and Engineering, Civil Aviation University of China, Tianjin 300300, China;
2. School of Civil Engineering, Tianjin University, Tianjin 300072, China;
3. College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: To study the vibration characteristics of asphalt concrete pavement, the time domain model of road roughness was simulated by the white noise filtering method, the coupling effect of vehicle and road with different vehicle speeds and different road roughness levels was analyzed based on the quarter vehicle mode, and the real-time dynamic load of vehicle on the road surface was determined to study the vibration characteristics of Tarmac pavements. The three-dimensional finite element model of the road is established, the vibration response of the road under the action of the vehicle random dynamic load is analyzed, and the influence of the road structure-layer parameters on the fundamental vibrational frequency of the road surface center is analyzed. The results show that the pavement roughness level and vehicle speed have significant influence on the vehicle-road coupling system. When the pavement roughness level changes from class A to class C, the dynamic load of vehicles under the same speed increases by 20%; the fundamental vibrational frequency of the road increases with the logarithmic relation of dynamic subgrade modulus, the subgrade modulus increases from 60 MPa to 260 MPa, the fundamental vibrational frequency of the road increases from 5.61 Hz to 10.80 Hz, with an increase of 48.06%. The dynamic modulus of the surface layer, base layer and cushion have little influence on the fundamental frequency, which are almost negligible. As the thickness of surface layer, base layer and cushion increases, the fundamental frequency of road shows a linear decrease trend. Within the commonly used variation range of the thickness of structure layer, the fundamental vibrational frequency decreases by 9.28%, 18.05% and 12.75% respectively.

Keywords: asphalt concrete pavement; vehicle dynamic load; finite-element; thickness of structural layer; elastic modulus of structural layer; fundamental frequency

Received: 2021-11-01; **Accepted**: 2022-01-21; **Published Online**: 2022-02-15 14:36

URL: kns.cnki.net/kcms/detail/11.2625.V.20220214.1748.012.html

Foundation items: National Natural Science Foundation of China (51178456); The Fundamental Research Funds for the Central Universities (3122017039)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0656

四旋翼无人机的自适应故障诊断与估计

王莉娜[1,*]，刘贞报[1]，院金彪[1]，党庆庆[1]，江飞鸿[1,2]，王宝栋[1]

(1. 西北工业大学 民航学院，西安 710072；2. 航空工业第一飞机设计研究院 飞控系统设计研究所，西安 710089)

摘　　要：针对四旋翼无人机执行器常见故障，提出一种基于自适应技术和观测器的鲁棒故障检测和估计（FDE）方法。在故障检测阶段，设计非线性诊断观测器，通过解析函数推导出阈值，确保所提检测方法的鲁棒性，并对所设计的观测器和残差评估函数进行证明。在故障估计阶段，提出基于切换ρ-修正的自适应律来准确估计检测到故障的方案。该方案不仅能够同时估计系统状态和残差信号，而且能估计未知故障的特征和大小。通过线性矩阵不等式进行设计参数的计算。利用2种故障场景分别进行仿真验证，同时在4种情况下讨论所提方法的有效性。基于四旋翼无人机硬件在环实验台验证了所提方法的可行性。

关 键 词：故障检测和估计；无人机；自适应技术；观测器；执行器故障

中图分类号：V279; V241

文献标志码：A　　**文章编号**：1001-5965(2023)09-2395-11

近年来，无人机因其在各种军事活动和民用应用中的潜力而备受业界广泛关注[1-2]。四旋翼无人机是一类特殊而被广泛应用的无人机，其配备有4个螺旋桨，分别安装在4个无刷直流电机的轴上。旋翼的旋转产生了维持四旋翼无人机所需的推力。此外，通过系统地改变旋翼速度来控制四旋翼无人机姿态，然后产生作用在机身框架上的力矩。然而，四旋翼系统存在严重的多变量耦合和非线性，对外界干扰也很敏感，也会产生部件退化或马达、螺旋桨等的损坏，以上因素都可能导致不同的四旋翼无人机故障，如执行器增益故障、电压控制故障、结构损坏、物理老化、疲劳和泄漏，这些都不可避免地影响了无人机系统的控制。故障的发生会对闭环控制系统的跟踪性能和稳定性造成不良影响。因此，飞行的稳定性、可靠性和安全性成为控制系统的重要指标[1-4]。为了达到这些要求，研究人员一方面研究了基于四旋翼无人机的故障检测和隔离的容错控制系统[5-7]。另外一方面，在不需要故障诊断机制的情况下，通过使用自适应控制算法开发了四旋翼执行器故障的容错控制方法[8-9]。为了确保传感器和执行器故障下的可靠性，研究了容错控制方法，以补偿故障、干扰、不确定性和未建模动态的影响，并将稳定性和其他性能保持在可接受的水平。然而，无人机的可靠性与安全性一直是故障检测与诊断和容错控制领域的热点问题，对故障发生时间、位置及其严重性信息的检测不足是设计容错控制系统的难点。因此，在设计容错控制系统的第1阶段非常有必要获得关于故障的严重程度、规模和行为的更多信息[10-13]，对这些信息的需求是开发故障检测和估计技术的主要动机。故障检测和估计可以提供更多有关故障功能的信息，如故障的严重程度、频率和时间特征等，有助于识别故障模式，从而设计容错控制系统。因此，设计一种对外部干扰不敏感，对故障高度敏感的故障诊

收稿日期：2021-11-02；**录用日期**：2022-02-13；**网络出版时间**：2022-03-18 17:19

网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220317.1828.004.html

基金项目：国家自然科学基金(61672430)；国家留学基金创新型人才国际合作培养项目(201906290246)；航空科学基金(ASFC-2018ZC53026)；陕西省重点研发计划(2019ZDLGY14-02-010)；深圳市基础研究资助项目(JCYJ20190806152203506)；深圳市高等院校稳定支持计划(GXWD202012311 65807008,20200830220334001)；北京空间飞行器总体设计部项目(JSZL2020203B004)

* **通信作者**. E-mail：lena@mail.nwpu.edu.cn

引用格式：王莉娜，刘贞报，院金彪，等. 四旋翼无人机的自适应故障诊断与估计 [J]. 北京航空航天大学学报，2023，49（9）：2395-2405.
WANG L N，LIU Z B，YUAN J B，et al. Adaptive fault diagnosis and estimation for quadrotor UAV[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（9）：2395-2405 （in Chinese）.

断方案是故障检测与估计领域的首要问题之一。

基于解析模型的方法在航空领域故障诊断中使用较早，因为在设计飞行控制系统的过程中就会建立相对完整的数学模型，基于该数学模型使用观测器估计系统的状态或参数，计算估计值与实际值之间的残差，最后判断和处理残差，从而判断出系统是否故障。基于模型的诊断方法是根据系统数学模型所蕴含的内在解析余度关系进行诊断的方法。近些年来基于模型的故障检测和诊断方法研究由于能获得较为完备的结论，所以得到了特别的重视并取得了快速发展[6]。通常在故障检测中，往往考虑最多的是突变故障，而忽略微小故障，对微小故障检测和隔离可以避免潜在的严重甚至灾难性故障发生[14]。执行器卡死、电压控制失效或系统结构损坏等突变性故障的出现对系统有明显的影响。由于突变故障与扰动明显不同，因此，可以通过选择合适的对应阈值来检测突变故障。比如传统的滑模观测器(sliding mode observer, SMO)[15]和Thau观测器[16]，都可以有效的处理这些故障。然而，这些观测器对微小故障并不适用。四旋翼飞行过程中受到噪声、气流干扰和振动信号的影响，这些影响通常很大。因此，检测和估计微小故障变得非常困难。在系统满足利普希茨条件的情况下，设计自适应滑模观测器，利用线性矩阵不等式确定观测器参数，最终通过仿真验证无人机微小故障检测方法[17]。为减少微小故障的影响，通过构造未知输入观测器的方法设计自适应容错控制算法[18]。为实现执行器的微小故障诊断，针对一类非线性系统提出了基于神经网络滑模观测器来进行微小故障诊断[6]。设计一种新的非线性观测器方法来进行闭环系统的微小故障诊断，最后通过仿真验证该方法的有效性[19]。

通过故障诊断程序获得的详细故障信息对基于状态的维护、冗余管理和故障转移非常有价值。利用滑模观测器设计了一种执行器故障估计方法。假设悬停飞行的扰动很小，以获得欧拉角速率和机身角速率之间的线性关系[20-21]。使用2级卡尔曼滤波器[5]、观测器[22-24]和滑模技术[25-26]获得一些较满意的实验结果。然而，这些故障估计方法大多数是基于线性化模型开发的[27]，很少考虑实际应用过程中不可避免的外部干扰[28]；还有大部分方法是通过假设欧拉角速率和机体角速度之间是线性关系开发的。而且大量文献中使用的故障估计阈值是手动选择的固定常数，在实际控制中的自由度很受限制。此外，自适应技术需要设计成输出反馈形式，其必须对不确定性和外部干扰的影响具有鲁棒性。然而，这一领域的相关研究仍然很少，需要进一步研究先进的理论和实验工作。

针对一类具有执行器故障、建模和测量不确定性的利普希茨非线性系统，本文提出一种基于自适应技术和观测器的鲁棒故障检测和估计(fault detection and estimation, FDE)方法。首先，设计非线性观测器来检测故障的发生。然后，通过残差评估函数进行适当的阈值评估，从而做出故障检测决策。当残差评估函数超过分析阈值时，检测到故障并产生报警信号。在报警信号产生时，自适应估计方法被激活用以估计检测到的故障特征。基于切换ρ-修正的鲁棒自适应律对未知故障函数进行估计，避免了参数漂移。此外，FDE方法的设计步骤以线性矩阵不等式(linear matrix inequality, LMI)问题的形式表示，为设计参数的计算提供了一种有效的方法。最后，利用四旋翼无人机模型进行软件仿真和硬件在环半实物实验(hardware in the loop simulation, HILS)验证了本文方法的有效性。

1 相关工作

1.1 四旋翼无人机模型

图1为四旋翼无人机布局，四旋翼无人机安装了4个电机M1～M4，在俯视视角中，M1和M3是逆时针旋转，而M2和M4是顺时针旋转，产生沿机体z轴反方向的力$F_n(n=1,2,\cdots,4)$。$\boldsymbol{p}$、$\boldsymbol{q}$、$\boldsymbol{r}$分别为3个方向的角速度。给定输入指令需要映射到螺旋桨速度以执行各自的运动，如图2所示。由于电机的旋转，每个电机产生相对于机体的反扭矩T_n。根据电机产生的推力和扭矩与电机角速度$W_n(n=1,2,\cdots,4)$平方的关系，可以得到如下关系[1-2]：

$$F_n = b_F W_n^2 \tag{1}$$

$$T_n = -k\mathrm{sgn}(W_n)W_n^2 \tag{2}$$

式中：b_F为电机推力系数；k为电机扭矩系数。

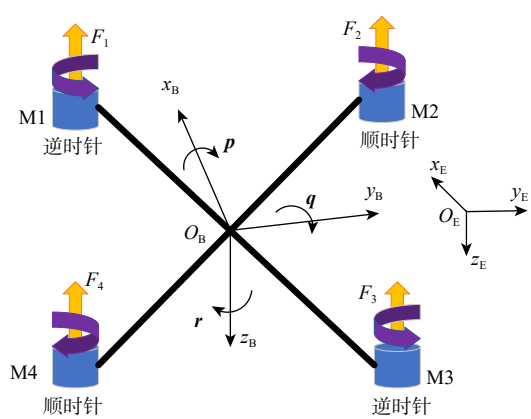


图1 四旋翼无人机布局

Fig. 1 Quadrotor UAV layout diagram

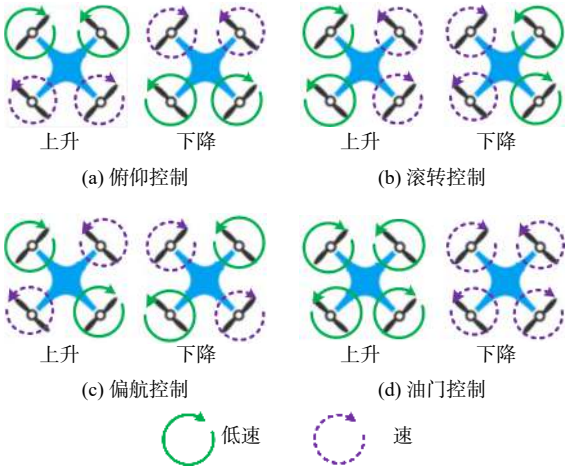


图 2　四旋翼无人机控制模式

Fig. 2　Quadrotor UAV control model

结合图 1, 利用式 (1) 和式 (2) 可以得到无人机的总推力和扭矩的关系式:

$$\begin{bmatrix} U \\ T_\phi \\ T_\theta \\ T_\psi \end{bmatrix} = \boldsymbol{\Pi} \begin{bmatrix} W_1^2 \\ W_2^2 \\ W_3^2 \\ W_4^2 \end{bmatrix} \tag{3}$$

式中: $\left[W_1^2, W_2^2, W_3^2, W_4^2\right]^{\mathrm{T}}$为给定电机转速; U、T_ϕ、T_θ和T_ψ分别为总推力、滚转扭矩、俯仰扭矩和偏航扭矩。$\boldsymbol{\Pi}$为推力和扭矩与电机角速度的映射矩阵, 具体表达式为

$$\boldsymbol{\Pi} = \begin{bmatrix} b_F & b_F & b_F & b_F \\ b_F l & -b_F l & -b_F l & b_F l \\ b_F l & b_F l & -b_F l & -b_F l \\ k & -k & k & -k \end{bmatrix} \tag{4}$$

式中: $l = d/\sqrt{2}$, d为电机几何中心到无人机质心之间的距离。

研究中使用的动力学模型考虑了质量、电机产生的推力和作用在四转子体上的扭矩。考虑已知干扰和模型不确定性的情况下, 根据牛顿第二定律和欧拉动力学方程, 构建无人机模型如下[2]:

$$\dot{\boldsymbol{p}}_{\mathrm{E}} = \boldsymbol{v}_{\mathrm{E}} \tag{5}$$

$$\dot{\boldsymbol{v}}_{\mathrm{E}} = \frac{1}{m}\boldsymbol{R}_{\mathrm{EB}}(\boldsymbol{\eta}) \left(\begin{bmatrix} 0 \\ 0 \\ -U \end{bmatrix} - \begin{bmatrix} c_d v_B \\ c_d v_B \\ c_d v_B \end{bmatrix} \right) + \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} + \boldsymbol{\xi}_v + \boldsymbol{\chi}_v \tag{6}$$

$$\dot{\boldsymbol{\eta}} = \boldsymbol{R}_{\eta}(\boldsymbol{\phi}, \boldsymbol{\theta})\boldsymbol{\omega} \tag{7}$$

$$\dot{\boldsymbol{\omega}} = \begin{bmatrix} \dfrac{J_y - J_z}{J_x}\boldsymbol{qr} \\ \dfrac{J_z - J_x}{J_y}\boldsymbol{pr} \\ \dfrac{J_x - J_y}{J_z}\boldsymbol{pq} \end{bmatrix} + \begin{bmatrix} \dfrac{1}{J_x}T_\phi \\ \dfrac{1}{J_y}T_\theta \\ \dfrac{1}{J_z}T_\psi \end{bmatrix} + \boldsymbol{\xi}_\omega + \boldsymbol{\chi}_\omega \tag{8}$$

式中: $\boldsymbol{p}_{\mathrm{E}}$、$\boldsymbol{v}_{\mathrm{E}}$、$\boldsymbol{\eta}$和$\boldsymbol{\omega} = \left[\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r}\right]^{\mathrm{T}}$分别为惯性位置、惯性速度、欧拉角和角速率; $c_d v_{\mathrm{B}}$为阻力。$\boldsymbol{R}_{\mathrm{EB}}$和$\boldsymbol{R}_{\eta}$均为旋转矩阵; m为无人机质量; J_x、J_y、J_z为 3 个方向上的惯量; $\boldsymbol{\chi}_v$和$\boldsymbol{\chi}_\omega$为建模不确定性; $\boldsymbol{\xi}_v$和$\boldsymbol{\xi}_\omega$为已知干扰。

1.2　问题描述

针对 1.1 节的四旋翼无人机模型, 用微分方程式(9)表示, 该模型包含了模型的非线性干扰、不确定性、测量不确定性及执行器故障。

$$\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u} + \boldsymbol{\xi}(\boldsymbol{x}, \boldsymbol{u}) + \boldsymbol{\chi}(\boldsymbol{x}, \boldsymbol{u}, t) + \\ \qquad \boldsymbol{\beta}(t - T)F_{\mathrm{a}}(\boldsymbol{x}, \boldsymbol{u}) \\ \boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} + \boldsymbol{D}\boldsymbol{d}(t) \end{cases} \tag{9}$$

式中: $\boldsymbol{x} \in \mathbf{R}^n$为状态量; $\boldsymbol{u} \in \mathbf{R}^q$为输入向量; $\boldsymbol{y} \in \mathbf{R}^m$为输出向量; $\boldsymbol{A}$、$\boldsymbol{B}$、$\boldsymbol{C}$、$\boldsymbol{D}$ 为已知矩阵; $\boldsymbol{\xi}(\boldsymbol{x}, \boldsymbol{u})$为模型的非线性干扰; $\boldsymbol{\chi}(\boldsymbol{x}, \boldsymbol{u})$为模型的不确定性和系统状态的干扰; $\boldsymbol{d}(t)$为测量不确定性; $\boldsymbol{\beta}(t - T)F_{\mathrm{a}}(\boldsymbol{x}, \boldsymbol{u})$表示$F_{\mathrm{a}}(\boldsymbol{x}, \boldsymbol{u})$故障发生的情况下, 系统动态性的改变。$F_{\mathrm{a}}(\boldsymbol{x}, \boldsymbol{u})$为系统执行器故障, $\boldsymbol{\beta}(t - T) = \mathrm{diag}(\beta_1(t - T), \beta_2(t - T), \cdots, \beta_n(t - T))$, T为未知故障发生时间, t为任意时刻, $\beta_i(t - T)$为第 i 个执行器故障对第 i 个状态方程的影响。

考虑增益故障的情况, 分别为有增益故障和无增益故障:

$$\beta_i(t - T) = \begin{cases} 0 & t < T \\ 1 & t \geqslant T \end{cases}$$

考虑微小故障的情况, 分别为无微小故障和有微小故障:

$$\beta_i(t - T) = \begin{cases} 0 & t < T \\ 1 - \mathrm{e}^{-k(t-T)} & t \geqslant T \end{cases}$$

假设非线性系统式 (9) 存在执行器微小故障或增益故障。采用乘性建模方法对这类故障进行建模, 非线性系统式 (9) 的微分方程可以描述为

$$\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}\boldsymbol{u} + \boldsymbol{\xi}(\boldsymbol{x}, \boldsymbol{u}) + \boldsymbol{\chi}(\boldsymbol{x}, \boldsymbol{u}, t) + \\ \qquad \boldsymbol{\beta}(t - T)\boldsymbol{\varphi}(\boldsymbol{x}, \boldsymbol{u})\boldsymbol{\theta} \\ \boldsymbol{y} = \boldsymbol{C}\boldsymbol{x} + \boldsymbol{D}\boldsymbol{d}(t) \end{cases} \tag{10}$$

式中: $\boldsymbol{\theta} \in \mathbf{R}^{p \times 1}$为故障幅度和大小的未知向量函数, 并假定$\boldsymbol{\theta}$有界$\|\boldsymbol{\theta}\| \leqslant \bar{\theta}$。$\boldsymbol{\varphi}(\boldsymbol{x}, \boldsymbol{u}) \in \mathbf{R}^{n \times p}$为执行器故障的函数结构。

本文的目的是提出一种基于自适应技术和观测器的鲁棒 FDE 方法, 该方法可以通过处理输入和状态信息来确定故障的发生、特征和严重程度。故障检测包括 2 个主要过程: 残差生成和残差评估。在残差生成过程中, 非线性诊断观测器被用来观测系统的状态并产生残差信号。残差信号反映了实际系统的状态与其标准模型估计状态之间的差异。在残差评估过程中, 残差评估函数与通过分析的阈值进行比较。如果残差评估函数超过阈值, 则检测到故障并产生报警信号。在报警产生时, 自适应故障估计方法被激活, 用以对故障的大小和严重

程度进行精确估计。本文方法提出之前需要做以下假设:

假设 1 $\boldsymbol{\xi}(\boldsymbol{x},\boldsymbol{u})$和$\boldsymbol{\varphi}(\boldsymbol{x},\boldsymbol{u})$均为$\boldsymbol{x}$的利普希茨函数,满足如下条件:

$$\|\boldsymbol{\xi}(\boldsymbol{x},\boldsymbol{u})-\boldsymbol{\xi}(\hat{\boldsymbol{x}},\boldsymbol{u})\|\leqslant\kappa_1\|\boldsymbol{x}-\hat{\boldsymbol{x}}\|$$

$$\|\boldsymbol{\varphi}(\boldsymbol{x},\boldsymbol{u})-\boldsymbol{\varphi}(\hat{\boldsymbol{x}},\boldsymbol{u})\|\leqslant\kappa_2\|\boldsymbol{x}-\hat{\boldsymbol{x}}\|$$

其中,κ_1和κ_2为利普希茨常量;$\hat{\boldsymbol{x}}\in\mathbf{R}^n$为非线性观测器的状态向量。

假设 2 假设式 (9) 模型不确定性$\boldsymbol{\chi}(\boldsymbol{x},\boldsymbol{u},t)$是非结构化的,同时也是未知非线性函数,但其是有界的,并满足$\left\|\boldsymbol{\chi}(\boldsymbol{x},\boldsymbol{u},t)\right\|\leqslant\bar{\chi}(\boldsymbol{x},\boldsymbol{u},t)$。

假设 3 未知测量不确定性满足范数有界:$\|\boldsymbol{d}(t)\|\leqslant\bar{d}_1,\left\|\dot{\boldsymbol{d}}(t)\right\|\leqslant\bar{d}_2$,$\bar{d}_1$和$\bar{d}_2$为已知常量。

2 自适应故障检测和估计方法

2.1 故障检测方法

为了估计系统的状态并产生用于故障检测的残差信号,设计如下非线性诊断观测器:

$$\begin{cases}\dot{\hat{\boldsymbol{x}}}=\boldsymbol{A}\hat{\boldsymbol{x}}+\boldsymbol{B}\boldsymbol{u}+\boldsymbol{\xi}(\hat{\boldsymbol{x}},\boldsymbol{u})+\boldsymbol{\beta}(t-T)\boldsymbol{\varphi}(\hat{\boldsymbol{x}},\boldsymbol{u})\hat{\boldsymbol{\theta}}+\\ \qquad\boldsymbol{P}(\boldsymbol{y}-\boldsymbol{C}\hat{\boldsymbol{x}})\\ \hat{\boldsymbol{y}}=\boldsymbol{X}\hat{\boldsymbol{x}}\end{cases}\tag{11}$$

式中:$\hat{\boldsymbol{\theta}}$为未知故障向量的估计函数,表示被检测故障的严重程度、大小和幅值。$\boldsymbol{P}$为非线性观测器的增益矩阵。

假设矩阵对$(\boldsymbol{A},\boldsymbol{C})$是可观测的,因此,非线性观测器的增益矩阵$\boldsymbol{P}$满足 Riccati 方程式 (12),且存在$\boldsymbol{X}$的对称正定解,表达式为

$$(\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C})^{\mathrm{T}}\boldsymbol{X}+\boldsymbol{X}(\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C})+(\varepsilon+1)\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}+\kappa_1^2\boldsymbol{I}<0\tag{12}$$

式中:ε为一个正常数。定义$\boldsymbol{\sigma}=\boldsymbol{x}-\hat{\boldsymbol{x}}$为状态误差,那么由式 (10) 和式 (11) 可以得到状态误差的变化量:

$$\begin{aligned}\dot{\boldsymbol{\sigma}}=&(\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C})\boldsymbol{\sigma}+\boldsymbol{\xi}(\boldsymbol{x},\boldsymbol{u})-\boldsymbol{\xi}(\hat{\boldsymbol{x}},\boldsymbol{u})+\boldsymbol{\chi}(\boldsymbol{x},\boldsymbol{u},t)+\\ &\boldsymbol{\beta}(t-T)\boldsymbol{\varphi}(\boldsymbol{x},u)\boldsymbol{\theta}-\boldsymbol{\beta}(t-T)\boldsymbol{\varphi}(\hat{\boldsymbol{x}},\boldsymbol{u})\hat{\boldsymbol{\theta}}+\boldsymbol{P}\boldsymbol{D}\boldsymbol{d}\end{aligned}\tag{13}$$

当系统处于无故障模式下,$\boldsymbol{\theta}$和$\hat{\boldsymbol{\theta}}$均为$\mathbf{0}$,从而,状态误差的变化量有如下形式:

$$\begin{aligned}\dot{\boldsymbol{\sigma}}=&(\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C})\hat{\boldsymbol{x}}+\boldsymbol{\xi}(\boldsymbol{x},\boldsymbol{u})-\boldsymbol{\xi}(\hat{\boldsymbol{x}},\boldsymbol{u})+\\ &\boldsymbol{\chi}(\boldsymbol{x},\boldsymbol{u},t)+\boldsymbol{P}\boldsymbol{D}\boldsymbol{d}\end{aligned}\tag{14}$$

对式 (14) 两边同时积分可得

$$\begin{aligned}\boldsymbol{\sigma}=&\mathrm{e}^{\boldsymbol{\Delta}t}\boldsymbol{\sigma}(0)+\int_0^t\mathrm{e}^{\boldsymbol{\Delta}(t-\tau)}(\boldsymbol{\xi}(\boldsymbol{x}(\tau),\boldsymbol{u}(\tau))-\\ &\boldsymbol{\xi}(\hat{\boldsymbol{x}}(\tau),\boldsymbol{u}(\tau))+\boldsymbol{\chi}(\boldsymbol{x}(\tau),\boldsymbol{u}(\tau),\tau)+\boldsymbol{P}\boldsymbol{D}\boldsymbol{d})\mathrm{d}\tau\end{aligned}\tag{15}$$

式中:$\boldsymbol{\Delta}=\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C}$;$\boldsymbol{d}$为不确定项。

残差评估的主要方法是选择适当的大于 0 的阈值J_{th},并使用以下逻辑关系进行故障检测:

$\|\boldsymbol{J}_r(t)\|<J_{\mathrm{th}}\Rightarrow$无故障发生,无报警;

$\|\boldsymbol{J}_r(t)\|\geqslant J_{\mathrm{th}}\Rightarrow$发生故障,报警 。其中,$\boldsymbol{J}_r(t)$为残差评估函数,定义为如下形式:

$$\|\boldsymbol{J}_r(t)\|=\left[\int_t^{t+T_1}\boldsymbol{r}^{\mathrm{T}}(t)\boldsymbol{r}(t)\mathrm{d}t\right]^{1/2}\tag{16}$$

式中:$\boldsymbol{r}=\boldsymbol{y}-\hat{\boldsymbol{y}}$为残差信号。$T_1$为有限的评估时间窗口。

引理 1[29] 由于$\boldsymbol{\Delta}=\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C}$是稳定的,存在大于 0 的常数$\alpha$和 b,满足$\mathrm{e}^{\boldsymbol{\Delta}t}\leqslant\alpha\mathrm{e}^{-bt}$。

定理 1 考虑非线性故障系统式 (10)、故障诊断观测器式 (11) 和评估函数式 (16),当评估函数$\|\boldsymbol{J}_r(t)\|$超过阈值函数J_{th}时,可以判定故障的发生。阈值函数为

$$\begin{aligned}J_{\mathrm{th}}=&\left[\int_t^{t+T_1}\left(\frac{\alpha(\bar{\chi}+\|\boldsymbol{P}\|\|\boldsymbol{D}\|\bar{d}_1)\|\boldsymbol{C}\|}{b-\alpha\kappa_1}+\right.\right.\\ &\left.\left.\left(\alpha\varepsilon-\frac{\alpha(\bar{\chi}+\|\boldsymbol{P}\|\|\boldsymbol{D}\|\bar{d}_1)}{b-\alpha\kappa_1}\right)\|\boldsymbol{C}\|\mathrm{e}^{-(b-\alpha\kappa_1)\tau}+\|\boldsymbol{D}\|\bar{d}_1\right)^2\mathrm{d}\tau\right]^{1/2}\end{aligned}\tag{17}$$

式中:ε为一个常数界,满足$\|\dot{\boldsymbol{x}}(0)\|=\|\boldsymbol{x}(0)\|\leqslant\varepsilon$。

证明 在 1.2 节假设 1~假设 3 存在的情况下,应用三角不等式,推导状态误差的范数上界如下:

$$\begin{aligned}\|\boldsymbol{\sigma}(t)\|\leqslant&\mathrm{e}^{\boldsymbol{\Delta}t}\|\boldsymbol{\sigma}(0)\|+\int_0^t\mathrm{e}^{\boldsymbol{\Delta}(t-\tau)}\|\boldsymbol{\xi}(\boldsymbol{x}(\tau),\boldsymbol{u}(\tau))-\\ &\boldsymbol{\xi}(\hat{\boldsymbol{x}}(\tau),\boldsymbol{u}(\tau))+\boldsymbol{\chi}(\boldsymbol{x}(\tau),\boldsymbol{u}(\tau),\tau)+\boldsymbol{P}\boldsymbol{D}\boldsymbol{d}\|\mathrm{d}\tau\leqslant\\ &\mathrm{e}^{\boldsymbol{\Delta}t}\varepsilon+\kappa_1\int_0^t\mathrm{e}^{\boldsymbol{\Delta}(t-\tau)}\|\boldsymbol{x}(\tau)-\hat{\boldsymbol{x}}(\tau)\|\mathrm{d}\tau+\\ &(\bar{\chi}+\|\boldsymbol{P}\|\|\boldsymbol{D}\|\bar{d}_1)\int_0^t\mathrm{e}^{\boldsymbol{\Delta}(t-\tau)}\mathrm{d}\tau\end{aligned}\tag{18}$$

在 1.2 节假设 2 存在的情况下,ε总是存在。把状态估计的初始条件设置为 0,即$\|\hat{\boldsymbol{x}}(0)\|=0$,那么$\|\boldsymbol{\sigma}(0)\|=\|\boldsymbol{x}(0)\|\leqslant\varepsilon$。把引理 1 中的关系代入式 (18) 中,得到:

$$\begin{aligned}\|\boldsymbol{\sigma}(t)\|\leqslant&\alpha\varepsilon\mathrm{e}^{-bt}+\alpha\kappa_1\int_0^t\|\boldsymbol{\sigma}(\tau)\|\mathrm{e}^{-b(t-\tau)}\mathrm{d}\tau+\\ &\frac{\alpha(\bar{\chi}+\|\boldsymbol{P}\|\|\boldsymbol{D}\|\bar{d}_1)}{b}(1-\mathrm{e}^{-bt})\end{aligned}\tag{19}$$

应用 Bellman-Gronwall 定理,式 (19) 可以转换为

$$\begin{aligned}\|\boldsymbol{\sigma}(t)\|\leqslant&\frac{\alpha(\bar{\chi}+\|\boldsymbol{P}\|\|\boldsymbol{D}\|\bar{d}_1)}{b-\alpha\kappa_1}+\\ &\left(\alpha\varepsilon-\frac{\alpha(\bar{\chi}+\|\boldsymbol{P}\|\|\boldsymbol{D}\|\bar{d}_1)}{b-\alpha\kappa_1}\right)\mathrm{e}^{-(b-\alpha\kappa_1)t}\end{aligned}\tag{20}$$

残差信号$\boldsymbol{r}=\boldsymbol{C}\boldsymbol{\sigma}$,对式 (20) 应用三角不等式,残差信号范数的上界可以表示为

$$\begin{aligned}\|\boldsymbol{r}(t)\|\leqslant&\frac{\alpha(\bar{\chi}+\|\boldsymbol{P}\|\|\boldsymbol{D}\|\bar{d}_1)\|\boldsymbol{C}\|}{b-\alpha\kappa_1}+\\ &\left(\alpha\varepsilon-\frac{\alpha(\bar{\chi}+\|\boldsymbol{P}\|\|\boldsymbol{D}\|\bar{d}_1)}{b-\alpha\kappa_1}\right)\|\boldsymbol{C}\|\mathrm{e}^{-(b-\alpha\kappa_1)t}+\|\boldsymbol{D}\|\bar{d}\end{aligned}\tag{21}$$

因此,可得阈值评价函数为

$$J_{\text{th}}=\left[\int_{t}^{t+T_1}\left(\frac{\alpha(\bar{\chi}+\|\boldsymbol{P}\|\|\boldsymbol{D}\|\bar{d}_1)\|\boldsymbol{C}\|}{b-\alpha\kappa_1}+\left(\alpha\varepsilon-\frac{\alpha(\bar{\chi}+\|\boldsymbol{P}\|\|\boldsymbol{D}\|\bar{d}_1)}{b-\alpha\kappa_1}\right)\|\boldsymbol{C}\|\mathrm{e}^{-(b-\alpha\kappa_1)\tau}+\|\boldsymbol{D}\|\bar{d}_1\right)^2\mathrm{d}\tau\right]^{1/2}\tag{22}$$

即证

估计误差的范数上界是通过$\sup_{\xi=\mathbf{0},\ \chi}\|\boldsymbol{\sigma}(t)\|$计算的[29], 这是在没有故障的情况下, 模型和测量不确定性对估计误差范数可能影响的最坏情况评价。

2.2　故障估计方法

在故障估计方法中关键的步骤是自适应率的设计。估计初值为 0, 故障诊断产生报警信息后, 故障估计方法被调用来识别故障的类型和幅度。

引理 2　对于系统式 (10) 和非线性观测器式 (11), 如果 $\boldsymbol{Y}$ 为任意对称矩阵, 那么有如下不等式成立:

$$2\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{Y}(\boldsymbol{\xi}(\boldsymbol{x},\boldsymbol{u})-\boldsymbol{\xi}(\hat{\boldsymbol{x}},\boldsymbol{u}))\leqslant\kappa_1^2\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{Y}\boldsymbol{\sigma}+\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\sigma}\tag{23}$$

证明　在 1.2 节假设 2 存在的前提下, 可以得到如下不等式:

$$2\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{Y}(\boldsymbol{\xi}(\boldsymbol{x},\boldsymbol{u})-\boldsymbol{\xi}(\hat{\boldsymbol{x}},\boldsymbol{u}))\leqslant 2\left\|\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{Y}(\boldsymbol{\xi}(\boldsymbol{x},\boldsymbol{u})-\boldsymbol{\xi}(\hat{\boldsymbol{x}},\boldsymbol{u}))\right\|\leqslant 2\kappa_1\left\|\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{Y}\right\|\|\boldsymbol{\sigma}\|\tag{24}$$

利用二次不等式, 式 (24) 可以变换为如下形式:

$$2\kappa_1\left\|\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{Y}\right\|\|\boldsymbol{\sigma}\|\leqslant(\kappa_1^2\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{Y}\boldsymbol{\sigma})+\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\sigma}\tag{25}$$

即证

定理 2　在 1.2 节假设 1～假设 3 成立的前提下, 考虑非线性系统式 (10)、状态观测器式 (11) 及其估计误差式 (15), 存在正定矩阵$\boldsymbol{Q}\in\mathbf{R}^{m\times m}$使得如下不等式成立:

$$\boldsymbol{\Pi}=\begin{bmatrix}\boldsymbol{\Theta}_1 & (\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C})^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{H}+\boldsymbol{\Gamma}\boldsymbol{P}\boldsymbol{D} & \boldsymbol{0}\\ (\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C})^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{H}+\boldsymbol{\Gamma}\boldsymbol{P}\boldsymbol{D} & \boldsymbol{\Theta}_2 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & -\rho_s\boldsymbol{I}\end{bmatrix}<\boldsymbol{0}\tag{26}$$

$$\dot{\hat{\boldsymbol{\theta}}}(t)=\gamma\boldsymbol{\varphi}^{\mathrm{T}}(\hat{\boldsymbol{x}},\boldsymbol{u})\boldsymbol{C}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{\nu}(t)-\gamma\rho_s\hat{\boldsymbol{\theta}}(t)\tag{27}$$

式中: $\boldsymbol{\Gamma}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{C}$; $\boldsymbol{Y}=(\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}})^{-1}\boldsymbol{C}\boldsymbol{\Gamma}\boldsymbol{C}^{\mathrm{T}}(\boldsymbol{C}\boldsymbol{C}^{\mathrm{T}})^{-1}$; $\boldsymbol{U}=\boldsymbol{C}^{-1}\boldsymbol{D}$;

$$\boldsymbol{\Theta}_1=(\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C})^{\mathrm{T}}\boldsymbol{\Gamma}+\boldsymbol{\Gamma}(\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C})+2\kappa_1^2\boldsymbol{I}+2\kappa_2^2\bar{\theta}^2\boldsymbol{I}+3\boldsymbol{\Gamma}\boldsymbol{\Gamma}+\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}$$

$$\boldsymbol{\Theta}_2=3\boldsymbol{V}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\Gamma}\boldsymbol{U}+\boldsymbol{D}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{U}+\boldsymbol{V}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{P}\boldsymbol{D}+\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{U}$$

$$\rho_s=\begin{cases}0 & \|\boldsymbol{\theta}(t)\|<N\\ \rho_0\left(\dfrac{\|\boldsymbol{\theta}(t)\|}{N}-1\right) & N\leqslant\|\boldsymbol{\theta}(t)\|\leqslant 2N\\ \rho_0 & \|\boldsymbol{\theta}(t)\|>2N\end{cases}$$

其中, $\gamma>0$代表学习率, ρ_0和N均为设计参数。

证明　选取李亚普诺夫函数:

$$\boldsymbol{V}=\boldsymbol{\nu}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{\nu}+\frac{1}{\gamma}\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\tilde{\boldsymbol{\theta}}\tag{28}$$

对李亚普诺夫函数求导:

$$\dot{\boldsymbol{V}}=\dot{\boldsymbol{\nu}}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{\nu}+\boldsymbol{\nu}^{\mathrm{T}}\boldsymbol{Y}\dot{\boldsymbol{\nu}}-\frac{2}{\gamma}\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\dot{\hat{\boldsymbol{\theta}}}\tag{29}$$

式中: $\tilde{\boldsymbol{\theta}}=\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}$为参数估计误差, 由此得到:

$$\rho_s\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\hat{\boldsymbol{\theta}}=\rho_s\tilde{\boldsymbol{\theta}}^{\mathrm{T}}(\boldsymbol{\theta}-\tilde{\boldsymbol{\theta}})=\rho_s\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\boldsymbol{\theta}-\rho_s\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\tilde{\boldsymbol{\theta}}=-\frac{1}{2}\rho_s\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\tilde{\boldsymbol{\theta}}-\frac{1}{2}\rho_s[(\tilde{\boldsymbol{\theta}}-\boldsymbol{\theta})^{\mathrm{T}}(\tilde{\boldsymbol{\theta}}-\boldsymbol{\theta})]+\frac{1}{2}\rho_s\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\theta}\tag{30}$$

考虑状态误差的动态性, 令 $\boldsymbol{\Gamma}=\boldsymbol{C}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{C}$, $\dot{\boldsymbol{V}}$ 可以写为

$$\begin{aligned}\dot{\boldsymbol{V}}=&\boldsymbol{\sigma}^{\mathrm{T}}\left[(\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C})^{\mathrm{T}}\boldsymbol{\Gamma}+\boldsymbol{\Gamma}(\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C})\right]\boldsymbol{\sigma}+\\ &2(\boldsymbol{\xi}(\boldsymbol{x},\boldsymbol{u})-\boldsymbol{\xi}(\hat{\boldsymbol{x}},\boldsymbol{u}))^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\sigma}+\\ &2(\boldsymbol{\varphi}(\boldsymbol{x},\boldsymbol{u})\boldsymbol{\theta}-\boldsymbol{\varphi}(\hat{\boldsymbol{x}},\boldsymbol{u})\hat{\boldsymbol{\theta}})^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\sigma}+\\ &2\boldsymbol{\chi}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\sigma}+\boldsymbol{d}^{\mathrm{T}}\boldsymbol{D}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\sigma}+\\ &2\dot{\boldsymbol{d}}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\sigma}+\boldsymbol{\sigma}^{\mathrm{T}}(\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C})^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{d}+\\ &2(\boldsymbol{\xi}(\boldsymbol{x},\boldsymbol{u})-\boldsymbol{\xi}(\hat{\boldsymbol{x}},\boldsymbol{u}))^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{d}+\\ &2(\boldsymbol{\varphi}(\boldsymbol{x},\boldsymbol{u})\boldsymbol{\theta}-\boldsymbol{\varphi}(\hat{\boldsymbol{x}},\boldsymbol{u})\hat{\boldsymbol{\theta}})^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{d}+\\ &2\boldsymbol{\chi}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{d}+\boldsymbol{d}^{\mathrm{T}}\boldsymbol{D}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{d}+\\ &2\dot{\boldsymbol{d}}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{d}+\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{P}\boldsymbol{D}\boldsymbol{d}+\boldsymbol{d}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{P}\boldsymbol{D}\boldsymbol{d}+\\ &\boldsymbol{d}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}(\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C})\boldsymbol{\sigma}+2(\boldsymbol{\varphi}(\hat{\boldsymbol{x}},\boldsymbol{u})\tilde{\boldsymbol{\theta}})^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{C}\boldsymbol{\sigma}+\\ &2(\boldsymbol{\varphi}(\hat{\boldsymbol{x}},\boldsymbol{u})\tilde{\boldsymbol{\theta}})^{\mathrm{T}}\boldsymbol{C}^{\mathrm{T}}\boldsymbol{Y}\boldsymbol{D}\boldsymbol{d}-\frac{2}{\gamma}\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\dot{\hat{\boldsymbol{\theta}}}\end{aligned}\tag{31}$$

由引理 2 可得如下不等式成立:

$$2((\boldsymbol{\varphi}(\boldsymbol{x},\boldsymbol{u})-\boldsymbol{\varphi}(\hat{\boldsymbol{x}},\boldsymbol{u}))\boldsymbol{\theta})^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\sigma}\leqslant\kappa_2^2\bar{\theta}^2\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\sigma}+\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\Gamma}\boldsymbol{\sigma}\tag{32}$$

$$2(\boldsymbol{\xi}(\boldsymbol{x},\boldsymbol{u})-\boldsymbol{\xi}(\hat{\boldsymbol{x}},\boldsymbol{u}))^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\sigma}\leqslant\kappa_1^2\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\sigma}+\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\Gamma}\boldsymbol{\sigma}\tag{33}$$

$$2((\boldsymbol{\varphi}(\boldsymbol{x},\boldsymbol{u})-\boldsymbol{\varphi}(\hat{\boldsymbol{x}},\boldsymbol{u}))\boldsymbol{\theta})^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{d}\leqslant\kappa_2^2\bar{\theta}^2\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\sigma}+\boldsymbol{d}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{d}\tag{34}$$

$$2(\boldsymbol{\xi}(\boldsymbol{x},\boldsymbol{u})-\boldsymbol{\xi}(\hat{\boldsymbol{x}},\boldsymbol{u}))^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{d}\leqslant\kappa_1^2\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\sigma}+\boldsymbol{d}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{d}\tag{35}$$

利用式 (32)～式 (35), 结合自适应率式 (27) 和式 (31), $\dot{\boldsymbol{V}}$ 可以写为

$$\begin{aligned}\dot{\boldsymbol{V}}=&\boldsymbol{\sigma}^{\mathrm{T}}[(\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C})^{\mathrm{T}}\boldsymbol{\Gamma}+\boldsymbol{\Gamma}(\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C})]\boldsymbol{\sigma}+\kappa_1^2\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\sigma}+\\ &\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\Gamma}\boldsymbol{\sigma}+\kappa_2^2\bar{\theta}^2\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\sigma}+\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\Gamma}\boldsymbol{\sigma}+\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\Gamma}\boldsymbol{\sigma}+\\ &\boldsymbol{d}^{\mathrm{T}}\boldsymbol{D}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\sigma}+\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\sigma}+\\ &\boldsymbol{\sigma}^{\mathrm{T}}(\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C})^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{d}+\kappa_1^2\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\sigma}+\\ &\boldsymbol{d}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{d}+\kappa_2^2\bar{\theta}^2\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\sigma}+\boldsymbol{d}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{d}+\\ &\boldsymbol{d}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{d}+\boldsymbol{d}^{\mathrm{T}}\boldsymbol{D}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{d}+\\ &\boldsymbol{d}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{U}\boldsymbol{d}+\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{P}\boldsymbol{D}\boldsymbol{d}+\\ &\boldsymbol{d}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}(\boldsymbol{A}-\boldsymbol{P}\boldsymbol{C})\boldsymbol{\sigma}+\boldsymbol{d}^{\mathrm{T}}\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma}\boldsymbol{P}\boldsymbol{D}\boldsymbol{d}+\\ &2\boldsymbol{\chi}^{\mathrm{T}}\boldsymbol{\chi}+2\dot{\boldsymbol{d}}^{\mathrm{T}}\dot{\boldsymbol{d}}-\rho_s\tilde{\boldsymbol{\theta}}^{\mathrm{T}}\tilde{\boldsymbol{\theta}}-\\ &\rho_s[(\tilde{\boldsymbol{\theta}}-\boldsymbol{\theta})^{\mathrm{T}}(\tilde{\boldsymbol{\theta}}-\boldsymbol{\theta})]+\rho_s\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\theta}\leqslant\boldsymbol{l}^{\mathrm{T}}\boldsymbol{\Pi}\boldsymbol{l}+\boldsymbol{\Psi}\end{aligned}\tag{36}$$

式中：$\boldsymbol{l}^{\mathrm{T}}=[\ \boldsymbol{\sigma}^{\mathrm{T}}\quad \boldsymbol{d}^{\mathrm{T}}\quad \tilde{\boldsymbol{\theta}}^{\mathrm{T}}\ ]$，则可得：

$$\boldsymbol{\Pi}=\begin{bmatrix}\boldsymbol{\Theta}_1 & (\boldsymbol{A}-\boldsymbol{PC})^{\mathrm{T}}\boldsymbol{\Gamma H}+\boldsymbol{\Gamma PD} & \boldsymbol{0}\\ (\boldsymbol{A}-\boldsymbol{PC})^{\mathrm{T}}\boldsymbol{\Gamma H}+\boldsymbol{\Gamma PD} & \boldsymbol{\Theta}_2 & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0} & -\rho_s\boldsymbol{I}\end{bmatrix}<\boldsymbol{0} \tag{37}$$

在不等式 (36) 中，$\dot{\boldsymbol{V}}$还可以表示为：$\dot{\boldsymbol{V}}\leqslant -\lambda(-\boldsymbol{\Pi}).\|\boldsymbol{l}(t)\|^2+\boldsymbol{\Psi}$。如果$\lambda(-\boldsymbol{\Pi})\|\boldsymbol{l}(t)\|^2>\boldsymbol{\psi}$，那么$\dot{\boldsymbol{V}}<\boldsymbol{0}$。根据李亚普诺夫理论，$\boldsymbol{\ell}(t)$可以收敛到一个较小集合。因此，状态估计误差和故障估计误差最终一致有界。

证毕

为了利用线性矩阵不等式求解，基于 Schur 引理[30]，将式 (37) 转换为关于$\boldsymbol{\Gamma}$的线性矩阵不等式如下：

$$\begin{bmatrix}\boldsymbol{\Theta} & \boldsymbol{M}_1 & \boldsymbol{0} & \boldsymbol{\Gamma} & \boldsymbol{\Gamma U} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{M}_1 & \boldsymbol{M}_2 & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma} & \boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma U}\\ \boldsymbol{0} & \boldsymbol{0} & -\rho_s\boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{\Gamma} & \boldsymbol{0} & \boldsymbol{0} & -3\boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{\Gamma U} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & -2\boldsymbol{I} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma U} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{I}\end{bmatrix}<\boldsymbol{0} \tag{38}$$

式中：

$$\boldsymbol{\Theta}=(\boldsymbol{A}-\boldsymbol{PC})^{\mathrm{T}}\boldsymbol{\Gamma}+\boldsymbol{\Gamma}(\boldsymbol{A}-\boldsymbol{PC})+2\kappa_1^2\boldsymbol{I}+2\kappa_2^2\bar{\boldsymbol{\theta}}^2\boldsymbol{I}$$

$$\boldsymbol{M}_1=(\boldsymbol{A}-\boldsymbol{PC})^{\mathrm{T}}\boldsymbol{\Gamma U}+\boldsymbol{\Gamma PD}$$

$$\boldsymbol{M}_2=\boldsymbol{D}^{\mathrm{T}}\boldsymbol{P}^{\mathrm{T}}\boldsymbol{\Gamma U}+\boldsymbol{U}^{\mathrm{T}}\boldsymbol{\Gamma PD}$$

3 仿真和实验

3.1 仿真分析

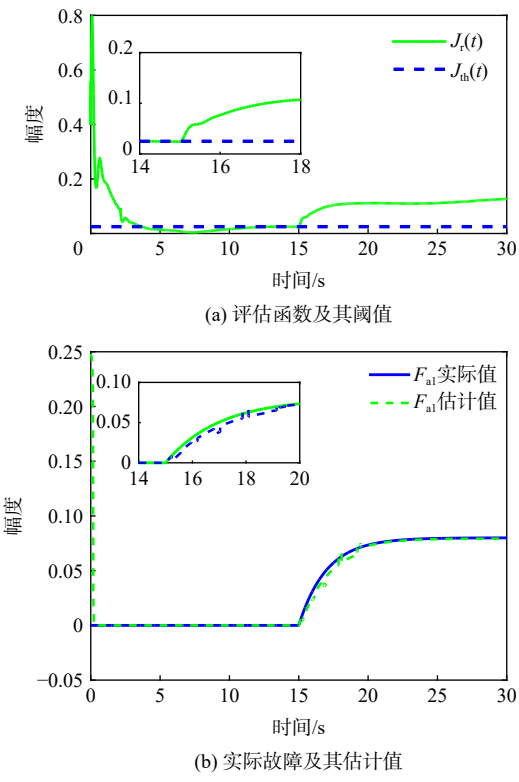
在本节中，将通过数值仿真验证本文方法在四旋翼无人机姿态控制系统中的性能。为了验证本文方法的有效性，将其性能与文献 [11] 中的滑模观测器故障诊断方法进行了比较。为了评估本文方法的性能，考虑了 2 种类型的故障：四旋翼执行机构微小故障和增益故障。

场景 1：执行机构老化。异常摩擦故障(F_{a1})：四旋翼执行器机构机械疲劳老化，异常摩擦会导致微小故障的发生。微小故障函数的特点是斜率很小，变化周期很长。在仿真中为了模拟执行器机构的微小故障，通过驱动电机上控制电压的变化模拟微小故障函数$F_{\mathrm{a1}}=(1-\mathrm{e}^{-0.5\times(t-15)})\times0.8\times0.1$，微小故障对应$\beta_i(t-T)=1-\mathrm{e}^{-0.5\times(t-15)}$，$\|\boldsymbol{\varphi}(\boldsymbol{x},\boldsymbol{u})\|=0.8$ 为微小故障的上界。$\|\boldsymbol{\theta}\|\in[0\ 1]$，当$\|\boldsymbol{\theta}\|=0$时，表明系统没有故障。当$\|\boldsymbol{\theta}\|\neq0$时，表明系统存在故障。$\|\boldsymbol{\theta}\|$取 0.1，代表系统在$t=15$ s时开始出现 10% 的老化摩擦故障。

场景 2：执行器增益故障。四旋翼在飞行过程中会受到外界气流扰动的影响，导致驱动电压无法保持在精确的基准电压上，出现跳变和失效。在仿真中模拟驱动电压的部分失效，模拟故障函数$F_{\mathrm{a2}}=-0.25\varepsilon(t-10)$。增益故障对应$\beta_i(t-T)=1$，$\|\boldsymbol{\varphi}(\boldsymbol{x},\boldsymbol{u})\|=-\varepsilon(t-10)$，$\|\boldsymbol{\theta}\|\in[0\ 1]$，当$\|\boldsymbol{\theta}\|=0$时，表明系统没有故障。当$\|\boldsymbol{\theta}\|\neq0$时，表明系统存在故障。$\|\boldsymbol{\theta}\|$取 0.25，代表系统在$t=10$ s时开始出现 25% 的失效故障。

仿真步骤具体如下：

步骤 1 在没有模型和测量不确定性的情况下，研究本文方法对故障类型为F_{a1}和F_{a2}的诊断和估计性能。图 3 为场景 1 下验证本文方法，图 4 为场景 2 下验证本文方法，故障诊断结果如图 3(a)和图 4(a)所示，诊断时间分别为 15.2 s 和 10.25 s。故障估计结果如图 3(b)和图 4(b)所示。由于不考虑不确定性，在故障发生时，评估函数偏离 0。因此，其检测到故障并产生报警信号。在故障估计时，本文方法被激活以对检测到的故障进行估计。从图 3 中可以看出本文方法对实际故障的估计是比较准确的。从图中可知，一旦评估函数超过阈值，就检测到故障，并且利用自适应估计方法来估计故障的幅度和严重性。仿真结果表明，本文方法不仅能够同时生成残差信号和检测故障，而且能够在可接受的时间内给出故障的准确估计。



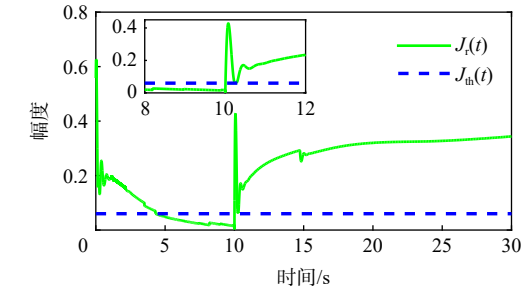
(a) 评估函数及其阈值

(b) 实际故障及其估计值

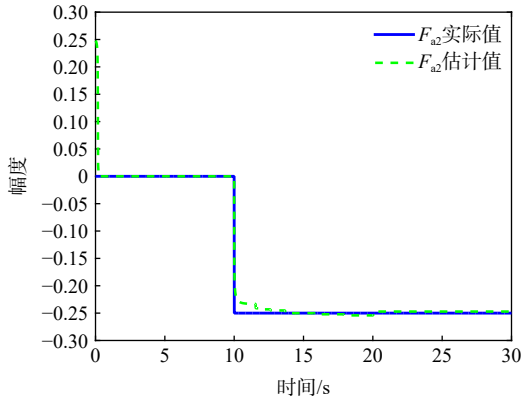
图 3 场景 1 下验证本文方法

Fig. 3 Verify Droposed method in scenario 1

步骤 2 为了验证本文方法对参数不确定性的鲁棒性，假设模型不确定性的上界为$\bar{\chi}=0.1$。图 5



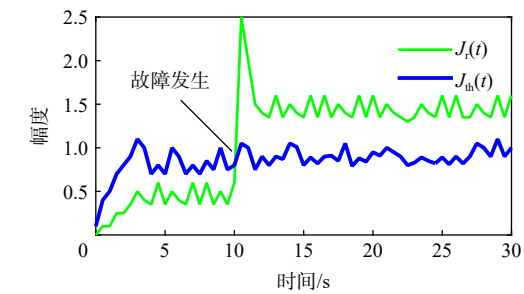
(a) 评估函数及其阈值



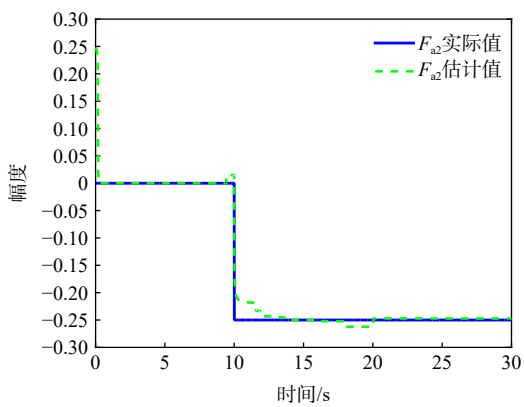
(b) 实际故障及其估计值

图 4　场景 2 下验证本文方法

Fig. 4　Proposed method verification in scenario 2



(a) 评估函数及其阈值



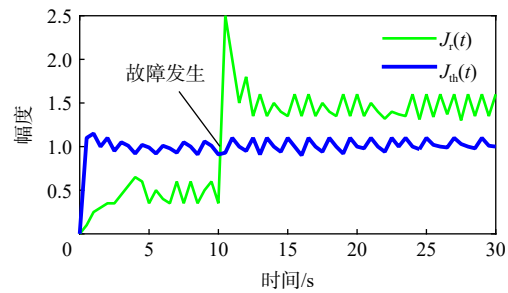
(b) 实际故障及其估计值

图 5　故障和模型不确定性存在情况下的方法验证

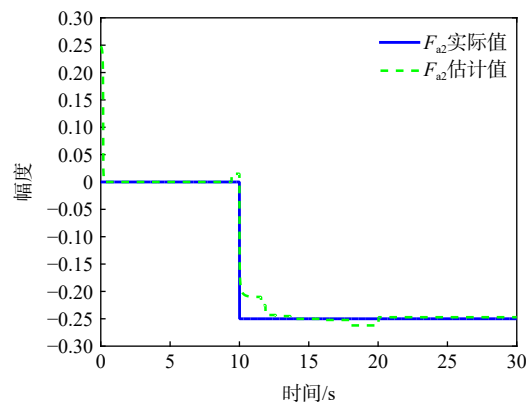
Fig. 5　Scheme verification in the presence of fault and model uncertainties

为故障和模型不确定性存在情况下的方法验证, 由图可知, 一旦故障发生, 评估函数趋势变大并超过故障检测的阈值。在对故障估计时, 自本文方法被激活以对发生的故障进行建模, 从中可以看出本文提出的切换 ρ-修正型自适应律, 可以良好的应用于存在不确定性时的鲁棒故障估计。

步骤 3　为了验证本文方法对未知测量不确定性的鲁棒性, 假设四旋翼无人机受到的干扰为$d_1(t) = 0.01\sin 200\,t$。图 6 为故障和测量不确定性存在情况下的方法验证。



(a) 评估函数及其阈值



(b) 实际故障及其估计值

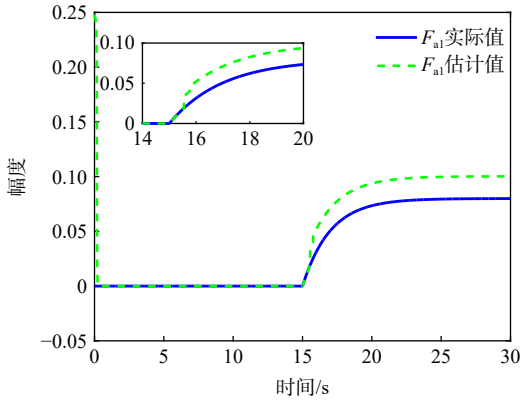
图 6　故障和测量不确定性存在情况下的方法验证

Fig. 6　Scheme verification in the presence of fault and measurement uncertainty

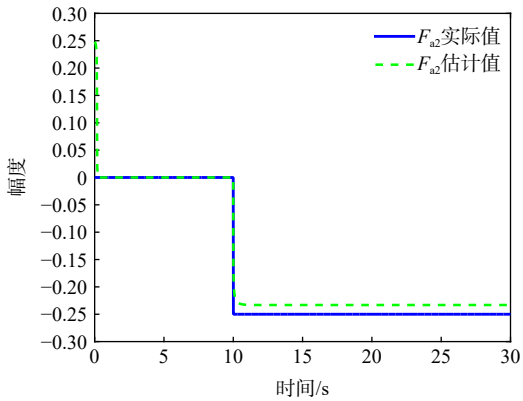
步骤 4　为了突出本文方法的有效性和优越性, 将故障估计结果与文献 [11] 中的方法进行比较, 给出了以上 2 种故障场景下的故障估计比对结果, 如图 7 所示。图 7 显示了文献 [11] 中的方法对 F_{a1} 和F_{a2}的估计结果。通过比较可以得出本文方法不仅能准确估计未知故障, 而且收敛速度快。针对故障检测时间和均方根误差 MSE 这 2 个指标, 将本文方法与文献 [11] 的方法进行比较, 如表 1 所示。本文方法对不确定性具有鲁棒性。由图 6 和图 7 可知, 文献 [11] 的方法估计效果不好, 偏差过大, 相比之下图 6 中本文方法估计效果更好。另外, 文献 [11] 中的方法对于建模和测量不确定性没有稳健的性能。

3.2　HILS 分析

图 8 为四旋翼无人机硬件在环仿真框图。整个测试环境主要由 3 大部分构成: ①四旋翼和飞行



(a) 场景1下实际故障及其估计值



(b) 场景2下实际故障及其估计值

图 7 基于文献 [11] 中方法的故障估计

Fig. 7 Fault estimation based on method in literature[11]

控制系统(部分功能);②xPC 实时环境;③半实物实时虚拟环境。四旋翼的运动可以通过飞行控制系统传输到目标计算机上,然后利用目标计算机对系统状态的捕获,生成控制命令,同时把命令传输给飞行控制板卡。四旋翼的运动状态同时也显示在虚拟环境中。

四旋翼和飞行控制系统具体由以下关键环节组成:四旋翼无人机、飞行控制系统、电池单元及球节附着结构。四旋翼和飞行控制系统通过接收控制指令,控制电机。xPC 实时环境上传飞行航点,接收姿态位置信息,发送状态给虚拟环境。本文所采用的为 DJI F450 四旋翼无人机。四旋翼无人机的框架总重量约为 710g,包括机身框架、Lipo 电池、电调、无刷直流电机和螺旋桨。型号为 Pixhawk 的飞行控制硬件系统是基于 TMS320F4 核心处理器进行研发的,已经进行多次试飞实验验证,具备良好的任务执行性能和可靠性。由于四旋翼无人机是一个开环不稳定系统,可能会对附近的人造成危险,因此,安全性是需要考虑的最重要因素之一。所以,把飞机固定在一个球形接头的顶部。这种配置使四旋翼无人机可以自由偏航运动,并有大约 45°的俯仰和滚转运动自由度。这种连接意味着四旋翼无人机上有 3 个可用的自由度,即滚转、俯仰和偏航运动。

HILS 实验台以 MATLAB xPC Target 作为实时仿真环境,这个实时仿真环境是基于 Simulink 平台的,成本也比较低。其由运行 Simulink 的主计算机和运行实时应用程序的目标计算机组成。在主计

表 1 本文方法与文献 [11] 中方法比较

Table 1 Performances of proposed method and reported approach of literature [11]

方法	场景1故障诊断时间/s	场景2故障诊断时间/s	场景1的MSE	场景2的MSE
本文方法	15.2	10.25	$6.672\,1\times10^{-4}$	$3.821\,1\times10^{-4}$
文献[11]中的方法	15.5	10.5	5.124 5	3.582 8

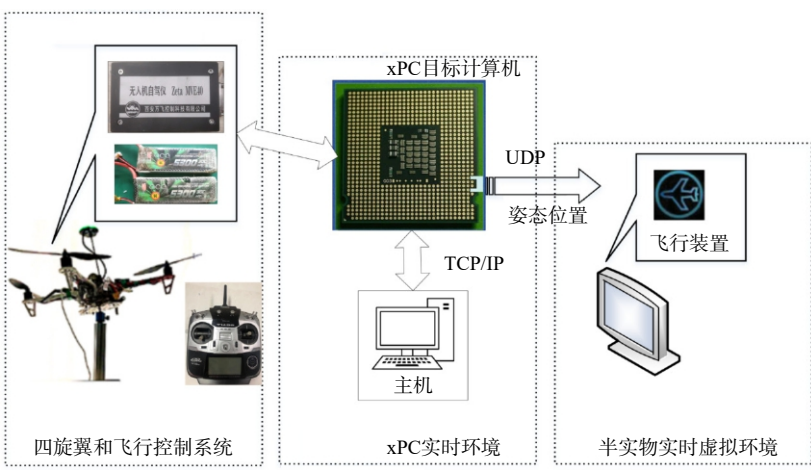


图 8 四旋翼无人机硬件在环仿真框图

Fig. 8 Quadrotor UAV HILS diagram

算机上搭建仿真模型, 而在目标计算机运行无人机的控制算法。2 个计算机通过基于传输控制协议/网际协议(transmission control protocol/internet protocol, TCP/IP)的无线网进行连接。目标计算机包含一个低功耗高性能的 CPU, 实现对复杂飞行控制任务的处理。主计算机和目标计算机的关系如图 9 所示。

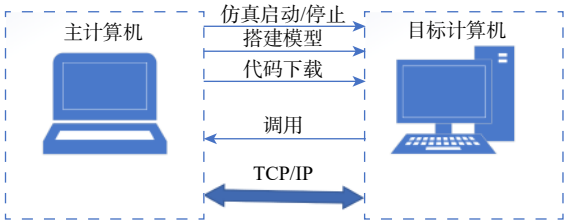
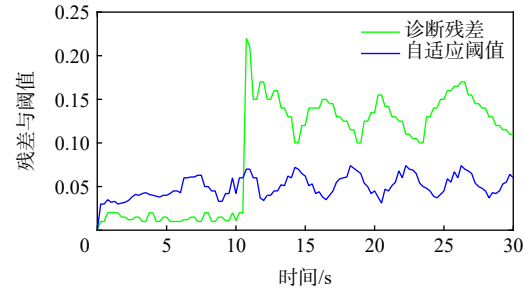


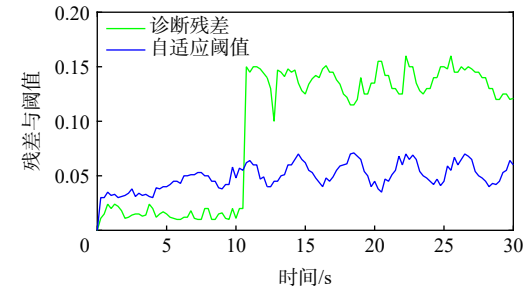
图 9　计算机之间的关系

Fig. 9　Connection between computers

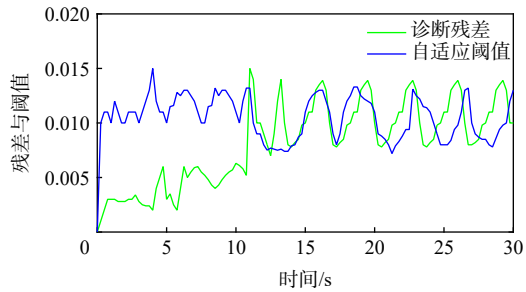
针对仿真分析中场景 2, 本文利用 HILS 实验台来进行进一步验证本文方法的有效性。故障诊断仿真模块首先在主计算机上构建和编译, 然后在目标计算机上下载并执行。仿真模块由 5 部分组成: ①参考模块为四旋翼生成所需的轨迹。②控制器模块将参考信号与反馈飞行状态进行比较, 运行控制算法, 然后生成相应的控制命令。③在执行器模块中, 创建一个串行通信模块, 将电机速度的命令传送到机载飞行控制系统中。④动力学模块包含 2 个部分: 基于非线性动态模型的四旋翼位置子系统和直接从机载飞行控制系统中获取的姿态信号模块。⑤显示模块负责管理 HILS 系统中的所有飞行状态。其不仅通过用户数据报协议(user datagram protocol, UDP)模块将飞行状态发送给虚拟显示环境, 还存储了所有实时模拟数据进行离线分析。本文在执行器模型中电机部分加入了故障信号, 通过故障注入系统对电机的影响来模拟实际飞行中的执行器故障, 以代码的形式控制特定电机表现出故障特性, 并将对应代码下载到目标计算机中, 通过无人机的姿态等数据判断故障注入前后的飞行效果, 并与仿真结果相对比。在实际仿真模拟软件中, 四旋翼无人机在恒定高度上沿椭圆轨迹飞行, 同时运行实时执行器故障诊断程序。根据场景 2, 当t = 10 s的时候给电机 1 注入 25% 的推力失效故障。图 10 为 3 个分量上的故障诊断和自适应阈值, 图 11 为故障估计结果。从图中可以看出, 在执行器故障发生之前, 所有分量的残差都低于其相应的阈值。在执行器故障发生后, 至少有一个分量超过其对应阈值, 故障诊断时间为 10.35 s, 从而可以得出, 本文方法在实际飞行仿真中可以及时的诊断故障发生的时间, 同时对故障进行良好的估计。



(a) ***p***分量机体角速率的故障诊断结果



(b) ***q***分量机体角速率的故障诊断结果



(c) ***r***分量机体角速率的故障诊断结果

图 10　3 个分量上的故障诊断和自适应阈值

Fig. 10　Fault diagnosis result

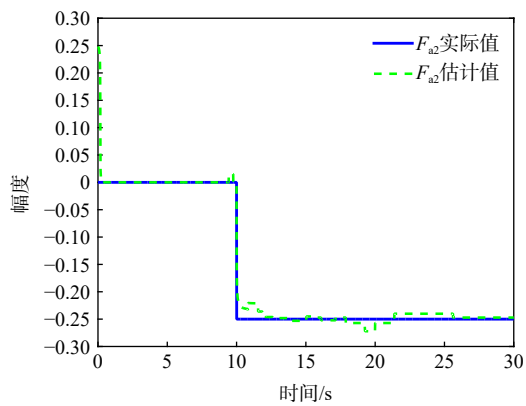


图 11　故障估计结果

Fig. 11　Fault estimation result

4　结　论

1) 通过仿真, 验证了本文方法对实际故障的估计是比较准确的。

2) 通过和文献方法进行比较, 从多个方面验证本文方法不仅能准确估计未知故障, 而且收敛速度快。

3) 通过 HILS 实验台验证了本文方法在实际飞

行仿真中可以及时的诊断故障发生的时间，同时对故障进行良好的估计。

参考文献（References）

[1] 全权. 多旋翼飞行器设计与控制[M]. 北京: 电子工业出版社, 2018: 70-85.
QUAN Q. Design and control of multi-rotor aircraft[M]. Beijing: Publishing House of Electronics Industry Press, 2018: 70-85 (in Chinese).

[2] ZHANG Y M, CHAMSEDDINE A, RABBATH C A, et al. Development of advanced FDD and FTC techniques with application to an unmanned quadrotor helicopter testbed[J]. Journal of the Franklin Institute, 2013, 350(9): 2396-2422.

[3] 申富媛, 李炜. 四旋翼无人机执行器可重构性量化评价方法研究[J]. 北京航空航天大学学报, 2020, 46(11): 2077-2086.
SHEN F Y, LI W. Quantitative reconfigurability evaluation method of actuator for quadrotor UAV[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(11): 2077-2086(in Chinese).

[4] 刘贞报, 马博迪, 高红岗, 等. 基于形态自适应网络的无人机目标跟踪方法[J]. 航空学报, 2021, 42(4): 524904.
LIU Z B, MA B D, GAO H G, et al. Adaptive morphological network based UAV target tracking algorithm[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(4): 524904(in Chinese).

[5] AMOOZGAR M H, CHAMSEDDINE A, ZHANG Y M. Experimental test of a two-stage Kalman filter for actuator fault detection and diagnosis of an unmanned quadrotor helicopter[J]. Journal of Intelligent & Robotic Systems, 2013, 70(1): 107-117.

[6] 权璐, 姜斌, 杨蒲. 基于神经网络滑模观测器的飞控系统故障诊断[J]. 扬州大学学报(自然科学版), 2019, 22(2): 51-55.
QUAN L, JIANG B, YANG P. Research on fault diagnosis algorithm of flight control system based on neural network sliding mode observer[J]. Journal of Yangzhou University (Natural Science Edition), 2019, 22(2): 51-55(in Chinese).

[7] ZHANG K, JIANG B, SHI P. Fast fault estimation and accommodation for dynamical systems[J]. IET Control Theory & Applications, 2009, 3(2): 189-199.

[8] 杨荟憬, 姜斌, 张柯. 四旋翼直升机姿态系统的直接自修复控制[J]. 控制理论与应用, 2014, 31(8): 1053-1060.
YANG H L, JIANG B, ZHANG K. Direct self-repairing control for four-rotor helicopter attitude systems[J]. Control Theory & Applications, 2014, 31(8): 1053-1060(in Chinese).

[9] AVRAM R C, ZHANG X D, MUSE J. Nonlinear adaptive fault-tolerant quadrotor altitude and attitude tracking with multiple actuator faults[J]. IEEE Transactions on Control Systems Technology, 2018, 26(2): 701-707.

[10] FREEMAN P, PANDITA R, SRIVASTAVA N, et al. Model-based and data-driven fault detection performance for a small UAV[J]. IEEE/ASME Transactions on Mechatronics, 2013, 18(4): 1300-1309.

[11] ZHANG J A, SWAIN A K, NGUANG S K. Robust sliding mode observer based fault estimation for certain class of uncertain nonlinear systems[J]. Asian Journal of Control, 2015, 17(4): 1296-1309.

[12] 王发威, 董新民, 陈勇, 等. 多操纵面飞机舵面损伤的快速故障诊断[J]. 航空学报, 2015, 36(7): 2350-2360.
WANG F W, DONG X M, CHEN Y, et al. Fast fault diagnosis of multi-effectors aircraft with control surface damage[J]. Acta Aeronautica et Astronautica Sinica, 2015, 36(7): 2350-2360(in Chinese).

[13] HAJIYEV C, SOKEN H E. Robust adaptive Kalman filter for estimation of UAV dynamics in the presence of sensor/actuator faults[J]. Aerospace Science and Technology, 2013, 28(1): 376-383.

[14] 李娟, 周东华, 司小胜, 等. 微小故障诊断方法综述[J]. 控制理论与应用, 2012, 29(12): 1517-1529.
LI J, ZHOU D H, SI X S, et al. Review of incipient fault diagnosis methods[J]. Control Theory & Applications, 2012, 29(12): 1517-1529(in Chinese).

[15] TAN C P, EDWARDS C. Sliding mode observers for detection and reconstruction of sensor faults[J]. Automatica, 2002, 38(10): 1815-1821.

[16] CEN Z H, NOURA H, SUSILO T B, et al. Robust fault diagnosis for quadrotor UAVs using adaptive thau observer[J]. Journal of Intelligent & Robotic Systems, 2014, 73(1): 573-588.

[17] LIU C, JIANG B, ZHANG K. Incipient fault detection using an associated adaptive and sliding-mode observer for quadrotor helicopter attitude control systems[J]. Circuits Systems and Signal Processing, 2016, 35(10): 3555-3574.

[18] CHEN W T, SAIF M. Unknown input observer design for a class of nonlinear systems: An LMI approach[C]//2006 American Control Conference. Piscataway: IEEE Press, 2006.

[19] WU Y K, JIANG B, ZHOU D H, et al. ToMFIR-based detection and estimation for incipient actuator faults in a class of closed-loop nonlinear systems[J]. IFAC Proceedings Volumes, 2014, 47(3): 1096-1101.

[20] TANG Y, LI Y. Dynamic modeling for high-performance controller design of a UAV quadrotor[C]//IEEE International Conference on Information and Automation. Piscataway: IEEE Press, 2015: 3112-3117.

[21] ZHANG K, JIANG B, YAN X G, et al. Incipient fault detection based on robust threshold generators: A sliding mode interval estimation approach[J]. IFAC-Papers, 2017, 50(1): 5067-5672.

[22] KE Y J, WANG K L, CHEN B M. Design and implementation of a hybrid UAV with model-based flight capabilities[J]. IEEE/ASME Transactions on Mechatronics, 2018, 23(3): 1114-1125.

[23] SUN C Y, YIN Y Z, MA H J. A dynamic fault detection method for nonlinear process[C]//2021 IEEE 10th Data Driven Control and Learning Systems Conference. Piscataway: IEEE Press, 2021: 108-113.

[24] 尹艳辉, 王付永, 刘忠信, 等. 带有完全分布式观测器的多智能体系统自适应容错一致性[J]. 控制理论与应用, 2021, 38(7): 1082-1090.
YIN Y H, WANG F Y, LIU Z X, et al. Fully distributed observer-based adaptive fault-tolerant consensus control for multi-agent systems[J]. Control Theory & Applications, 2021, 38(7): 1082-1090(in Chinese).

[25] ZHANG J Q, REN Z H, DENG C, et al. Adaptive fuzzy global sliding mode control for trajectory tracking of quadrotor UAVs[J]. Nonlinear Dynamics, 2019, 97(1): 609-627.

[26] EDWARDS C, TAN C P. A comparison of sliding mode and unknown input observers for fault reconstruction[J]. European Journal

of Control, 2006, 12(3): 245-260.

[27] NIAN X H, CHEN W Q, CHU X Y, et al. Robust adaptive fault estimation and fault tolerant control for quadrotor attitude systems[J]. International Journal of Control, 2020, 93(3): 725-737.

[28] AVRAM R C, ZHANG X D, MUSE J. Quadrotor actuator fault diagnosis and accommodation using nonlinear adaptive estimators[J]. IEEE Transactions on Control Systems Technology, 2017, 25(6): 2219-2226.

[29] TANG Y R, LI Y M. Dynamic modeling for high-performance controller design of a UAV quadrotor[C]//2015 IEEE International Conference on Information and Automation. Piscataway: IEEE Press, 2015: 3112-3117.

[30] DING S X. Model-based fault diagnosis techniques: Design schemes, algorithms and tools[M]. Berlin: Springer, 2013: 80-152.

Adaptive fault diagnosis and estimation for quadrotor UAV

WANG Lina[1,*], LIU Zhenbao[1], YUAN Jinbiao[1], DANG Qingqing[1], JIANG Feihong[1,2], WANG Baodong[1]

(1. School of Civil Aviation, Northwestern Polytechnical University, Xi'an 710072, China;
2. Flight Control System Design Institute, AVIC the First Aircraft Design Institute, Xi'an 710089, China)

Abstract: A fault detection and estimation(FDA) method based on adaptive technology and observer is designed for common actuator faults of quadrotor unmanned aerial vehicle. In the stage of fault detection, the nonlinear diagnostic observer is designed, and the threshold value is derived by analytic function to ensure the robustness of the proposed detection method. Moreover, the designed observer and residual evaluation function are proved. In the stage of fault estimation, an adaptive law based on switching ρ-correction is proposed to accurately estimate the detected faults. This scheme can not only simultaneously estimate the system state and residual signal, but also estimate the characteristics and magnitude of unknown faults. The design parameters are calculated by linear matrix inequality method. The simulation results are verified in two fault scenarios, and the validity of the proposed method is discussed in four cases. The experimental results show the effectiveness of the proposed method by using the hardware-in-loop simulation of quadrotor unmanned aerial vehicle.

Keywords: fault detection and estimation; unmanned aerial vehicle; adaptive technology; observer; actuator fault

Received: 2021-11-02; **Accepted**: 2022-02-13; **Published Online**: 2022-03-18 17:19

URL: kns.cnki.net/kcms/detail/11.2625.V.20220317.1828.004.html

Foundation items: National Natural Science Foundation of China (61672430); China Scholarship Fund International Cooperation Training Program for Innovative Talents (201906290246); Aeronautical Science Foundation of China (ASFC-2018ZC53026); Key R & D Program of Shaanxi Province (2019ZDLGY14-02-010); Shenzhen Fundamental Research Program (JCYJ20190806152203506); Shenzhen Basic Research Funding Project (GXWD202012311 65807008, 20200830220334001); Beijing Institute of Space Systems Engineering (JSZL2020203B004)

* **Corresponding author**. E-mail: lena@mail.nwpu.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0667

机器人辅助视网膜下注射系统的设计与实现

王朝董[1]，广晨汉[1]，王丽强[2]，宗俊杰[1]，郑昱[1]，杨洋[1,*]

(1. 北京航空航天大学 机械工程及自动化学院，北京 100191；2. 解放军总医院 眼科医学部，北京 100853)

摘 要：视网膜下注射人胚胎干细胞是治疗视网膜变性的一种有效术式，该术式对医生手术操作的精准性、稳定维持能力、安全性均提出了很高的要求。为此，提出一种基于主从式机器人的视网膜下注射系统，辅助医生完成视网膜下注射操作。构建从手机器人的正逆运动学模型，然后进行视网膜下注射手术运动分析，确定从手机器人的运动要求。建立主手和从手机器人之间的运动映射关系，依据运动映射关系推导建立远程运动中心（RCM）点位置调整运动、术中RCM运动的状态切换和分离过程的速度映射模型。通过离体猪眼球视网膜下穿刺注射实验对机器人的精确性和稳定维持能力进行验证，结果表明：机器人辅助操作系统末端注射针具有稳定维持能力和运动的精确性，机器人辅助操作比徒手操作对视网膜造成的创伤更小，注射更稳定。

关 键 词：机器人辅助视网膜手术；主从控制；远程运动中心；眼内手术；显微手术

中图分类号：TP242.3

文献标志码：A **文章编号**：1001-5965(2023)09-2406-09

视网膜病变是眼科常见的一类疾病，其中老年性黄斑变性 (age-related macular degeneration, AMD) 等是目前严重危害人们视力的疾病，是发达国家老年人失明的主要原因[1]。在现代医学中，尚无有效的治疗方法[2]。细胞替代疗法已成为治疗该类疾病的一种可行策略[3-5]。该治疗方法需要将干细胞移植到视网膜下[6]。

以视网膜下干细胞移植为例，眼科手术普遍面临手术空间狭小、操作精度高、医生手生理振颤等方面带来的制约[7-9]。机器人技术的进步为解决上述问题提供了可行的途径。美国霍普金斯大学采用协同式控制方式开发了稳定手眼机器人（steady hand eye robot, SHER）。共有SHER 1.0、SHER 2.0、SHER 2.1这3代，使用时医生直接握持安装于机器人末端的手术器械进行手术操作，手术系统具备微力感知、人手振颤过滤等功能，并开展了眼内力感知、模拟眼底注药等研究[10-11]；荷兰埃因霍温大学的研究者们针对视网膜手术研发了Preceyes机器人，2018年在医生的操作下Preceyes机器人成功完成病人视网膜表面上0.01 mm厚的再生膜摘除，完成了全球较早机器人辅助操作的视网膜临床手术[12]；比利时鲁汶大学针对视网膜静脉血管插管术研制了一套主从式机器人，2016年在医生的控制下，该机器人成功完成了全球较早人体视网膜静脉插管手术[13-14]；北京航空航天大学分别在2016年开发了角膜移植手术机器人、2014年开发了第1代主从视网膜手术机器人、2018年开发了第2代主从式视网膜手术机器人[15-17]。

本文提出一种基于主从式机器人的视网膜下注射系统，用于辅助医生完成视网膜下注射干细胞。机器人采用摇杆作为主手，采用包含远程运动中心（remote centre of motion, RCM）机构的机器人为从手，同时采用眼科显微镜为医生提供视觉反馈。分析了机器人正逆运动学模型，建立了主手和

收稿日期：2021-11-05；**录用日期**：2022-02-13；**网络出版时间**：2022-03-25 10:47
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220324.1616.004.html
基金项目：国家重点研发计划智能机器人专项 (2007YFB1302700)
*通信作者. E-mail：yang_mech@buaa.edu.cn

引用格式：王朝董，广晨汉，王丽强，等. 机器人辅助视网膜下注射系统的设计与实现 [J]. 北京航空航天大学学报，2023，49（9）：2406-2414. WANG Z D，GUANG C H，WANG L Q，et al. Design and implementation of robot-assisted subretinal injection system [J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（9）：2406-2414 （in Chinese）.

从手机器人之间的运动映射关系, 解决了该术式眼内操作时巩膜约束运动、准确定位注射位置、注射时稳定器械、微流量注射等问题。最后通过离体猪眼球视网膜下穿刺注射实验对机器人的精确性和稳定维持能力进行验证。

1　机器人辅助视网膜下注射系统

本文机器人系统组成如图 1 所示。相较于视网膜血管相关手术, 视网膜下穿刺注射手术的操作复杂度较低, 故选择操纵简便的摇杆作为主手, 进行速度控制。在满足动作要求前提下简化控制程序, 以提高控制效率。

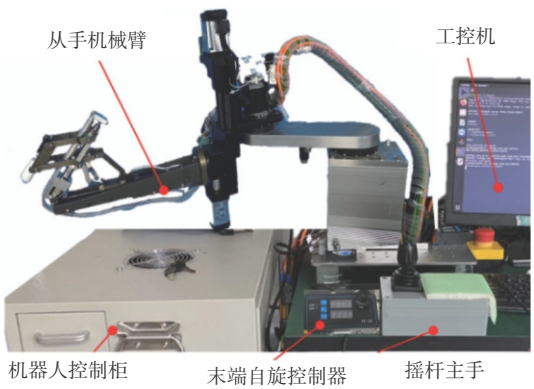


图 1　机器人系统组成

Fig. 1　Composition of robot system

机器人系统由摇杆主手、末端自旋控制器、从手机械臂、机器人控制柜、工控机组成。

摇杆主手产生 6 轴控制信号, 注射针自旋为单独自由度由末端自旋控制器控制。摇杆主手、末端自旋控制器两者共同完成 7 个自由度控制。

从手机械臂作为手术动作执行系统, 机械结构分为 Scara 机械臂(机器人前三自由度)和 RCM 机构(后四自由度), Scara 机械臂完成术中 RCM 点位置调整, RCM 机构完成 RCM 点与巩膜口重合后的注射针位置调整、姿态调整、穿刺注射等动作。

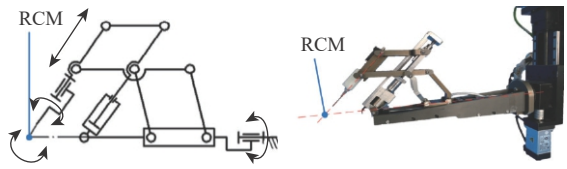
控制系统实现摇杆指令信号获取、RCM 初始位置信息采集、机器人各关节编码器信息的采集、处理、发送, 控制机器人运动。

1.1　RCM 机构

本文在笔者课题组 RCM 机构[18] 基础上优化, 机器人辅助视网膜下注射系统采用的 RCM 机构, 可以实现手术器械围绕巩膜口(RCM 点)约束的俯仰、偏摆、自旋及沿器械轴线的进给退出运动, RCM 机构的运动简图和实物图如图 2 所示。

1.2　末端微流注射系统

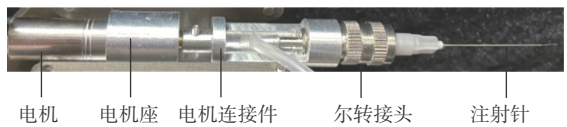
微流注射系统如图 3 所示, 图 3(a) 从右至左依次为: 注射针、鲁尔转接头、电机连接件、电机座、电机; 图 3(b) 为微流泵。注射针螺旋连接至鲁尔内



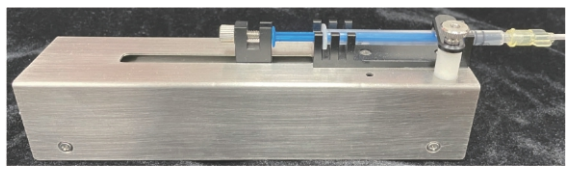
(a) 机构简图　　(b) 实物图

图 2　RCM 机构运动简图和实物图

Fig. 2　Schematic diagram and physical diagram of RCM



(a) 注射系统套件连接



(b) 微流泵

图 3　微流注射系统示意

Fig. 3　Schematic diagram of microfluidic injection system

旋接头, 鲁尔转接头与电机连接件通过螺纹连接, 电机连接件与电机连接, 电机安装在电机座上, 电机座安装至机器人。

视网膜下注射胚胎干细胞的研究中, 视网膜下注射干细胞的个数为 10^6, 注射浓度为 10 000 个/μL, 即注射液总容量为 100 μL, 其中微流泵吸液速度为 100 μL/min[4]。本系统采用的微型注射泵 RSP01-BG2。微流注射泵注射速度、流量可调, 从μL/h 至 mL/s 流量量级进行设定, 更换不同容量注射器满足流量设定要求。

末端注射针选取特种针头, 型号为 41G, 针尖处直径为 80 μm, 注射针如图 4 所示。

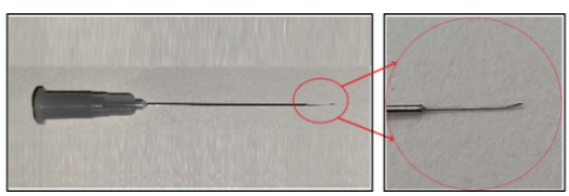


图 4　注射针

Fig. 4　Injection needle

1.3　系统性能指标

机器人辅助视网膜下注射系统是对笔者课题组前期成果(系统的组成和控制方法)[8] 的优化提升, 机器人性能指标如表 1 所示, 本文立足于验证机器人系统在离体猪眼球上进行眼底注射操作的可行性和优越性, 机器人精度建模、计算并不属于本文的研究重点。因此, 在后续的研究中, 会直接给出精度方面的研究结果。

表 1 机器人性能指标

Table 1 Robot performance index

参数	数值
RCM点*xyz*运动范围/mm	±50
RCM点*xyz*运动精度/mm	0.01
俯仰运动范围/(°)	±30
翻滚运动范围/(°)	±90
俯仰/翻滚运动精度/rad	0.000 5
注射针自身轴向的旋转范围/(°)	±180
注射针自身轴向的旋转精度/(°)	< 1

2 机器人运动学模型及工作空间

2.1 正运动学及末端工作空间

2.1.1 机器人正运动学分析

本节介绍机器人的运动学模型，机器人坐标系如图 5 所示。机器人系统中关节 1 高度为 d_1，长度为 a_1，关节 2 距离关节 1 的高度为 d_2，长度为 a_2，关节 3 距离关节 2 的高度为 d_3，RCM 点与关节 3 的距离为 a_3，注射针末端弯折点距 RCM 点的距离为 d_6。根据 D-H 参数定义得到 D-H 参数表如表 2 所示，其中，i 为连杆编号。

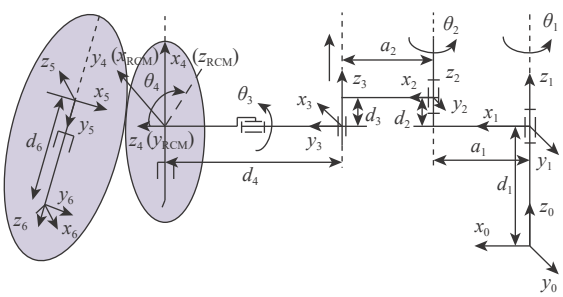


图 5 机器人坐标系

Fig. 5 Robot coordinate system

表 2 D-H 参数表

Table 2 D-H parameter table

i	z_{i-1}到z_i沿x_{i-1}距离a_{i-1}	z_{i-1}到z_i绕x_{i-1}角度α_{i-1}/(°)	x_{i-1}到x_i沿z_i距离d_i	x_{i-1}到x_i绕z_i角度θ_i
1	0	0	d_1	θ_1
2	a_1	0	d_2	θ_2
3	a_2	0	d_3	−90°
4	0	−90	d_4	−90°+θ_3
5	0	−90	0	θ_4
6	0	−90	d_6	0

根据 D-H 参数表得到 Scara 机械臂和 RCM 机构的正运动学方程，如下：

$$
{}_0^3\boldsymbol{T}={}_0^1\boldsymbol{T}_1^2\boldsymbol{T}_2^3\boldsymbol{T}=\begin{bmatrix}\sin\theta_1\sin\theta_2 & \cos\theta_1\cos\theta_2 & 0 & a_1\cos\theta_1+a_2\cos\theta_1\cos\theta_2\\ -\cos\theta_1\cos\theta_2 & \sin\theta_1\sin\theta_2 & 0 & a_1\sin\theta_1+a_2\sin\theta_1\sin\theta_2\\ 0 & 0 & 1 & d_{123}\\ 0 & 0 & 0 & 1\end{bmatrix} \tag{1}
$$

$$
{}_3^6\boldsymbol{T}={}_3^4\boldsymbol{T}_4^5\boldsymbol{T}_5^6\boldsymbol{T}=\begin{bmatrix}-\cos\theta_4\sin\theta_3 & -\cos\theta_3 & -\sin\theta_3\sin\theta_4 & -d_6\sin\theta_3\sin\theta_4\\ -\sin\theta_4 & 0 & -\cos\theta_4 & d_4-d_6\cos\theta_4\\ \cos\theta_3\cos\theta_4 & \sin\theta_3 & -\cos\theta_3\sin\theta_4 & -d_6\cos\theta_3\sin\theta_4\\ 0 & 0 & 0 & 1\end{bmatrix} \tag{2}
$$

$$
{}_6^z\boldsymbol{T}=\begin{bmatrix}\cos\theta_6 & -\sin\theta_6 & 0 & 0\\ \sin\theta_6 & \cos\theta_6 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix} \tag{3}
$$

$$
\boldsymbol{P}_o={}_0^3\boldsymbol{T}\begin{bmatrix}0\\ d_4\\ 0\\ 1\end{bmatrix}=\begin{bmatrix}a_1\cos\theta_1+(a_2+d_4)\cos\theta_1\cos\theta_2\\ a_1\sin\theta_1+(a_2+d_4)\sin\theta_1\sin\theta_2\\ d_{123}\end{bmatrix} \tag{4}
$$

式中：${}_0^3\boldsymbol{T}$为 Scara 机械臂的正运动学方程；$\boldsymbol{P}_o$ 为机器人 RCM 点相对于基座标系$Ox_0y_0z_0$的三维坐标。除此之外，还需要求解机器人末端点的坐标。故还需要求解 RCM 机构的正运动学方程，如式 (2) 所示，${}_3^6\boldsymbol{T}$为 RCM 机构的正运动学方程；${}_i^j\boldsymbol{T}$为该行的齐次变换矩阵；d_{123} 为 d_1+d_2+d_3。

2.1.2 末端 RCM 工作空间分析

机器人正运动学中经变换矩阵${}_0^3\boldsymbol{T}$计算 $\boldsymbol{P}_o$，手术时 $\boldsymbol{P}_o$ 与巩膜口重合保持固定。眼内手术动作由 RCM 机构实现，基于蒙特卡罗法经 RCM 机构变换矩阵${}_3^6\boldsymbol{T}$计算末端 RCM 工作空间如图 6～图 9 所示。图中，由器械刺入点处向外扩展，图中灰色球面模拟眼球，半径为 25 mm。

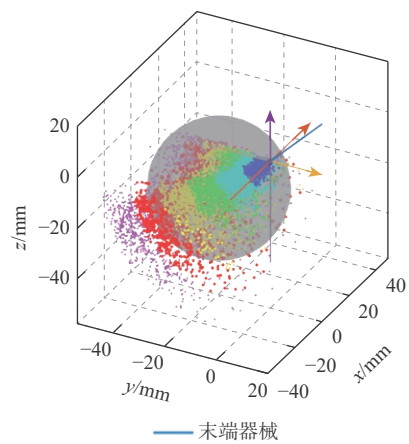


图 6 末端工作空间向视图

Fig. 6 End workspace direction view

仿真结果显示机器人末端工作空间完全覆盖眼底穿刺注射手术涉及区域，器械姿态中轴线由 RCM 点连线确定，另一姿态参数则是注射针的自旋，范围在表 1 中已列出。

2.2 机器人逆运动学分析

机器人的运动中，Scara 机械臂运动和 RCM 运动进行分别控制。Scara 机械臂实现 RCM 点在基

坐标系的坐标位置(P_x, P_y, P_z), 即粗定位运动; RCM 运动即 RCM 点与巩膜开口重合后的精定位及穿刺阶段所需的动作, RCM 机构的运动学求解与 Scara 机械臂分别进行, 由逆运动学计算出各关节变量参数, Scara 机械臂逆运动学求解如下, 由 2.1 节中正运动学模型可得

$$d_3 = P_z - d_1 - d_2 \tag{5}$$

$$a_1\cos\theta_1 + (a_2+d_4)\cos(\theta_1+\theta_2) = P_x \tag{6}$$

$$a_1\sin\theta_1 + (a_2+d_4)\sin(\theta_1+\theta_2) = P_y \tag{7}$$

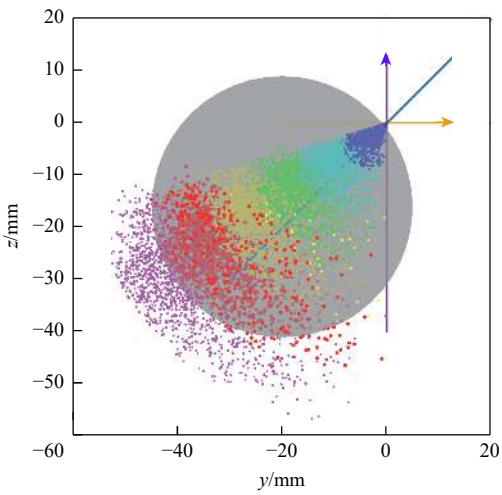


图 7　末端工作空间 *y*-*z* 投影

Fig. 7　*y*-*z* projection of end workspace

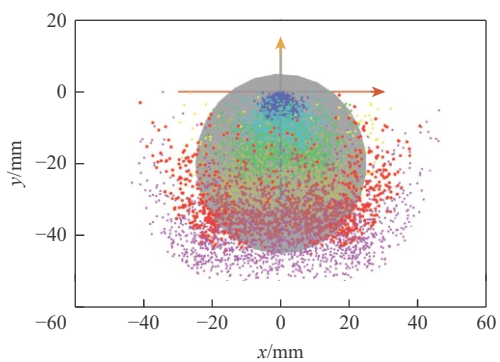


图 8　末端工作空间 *x*-*y* 投影

Fig. 8　*x*-*y* projection of end workspace

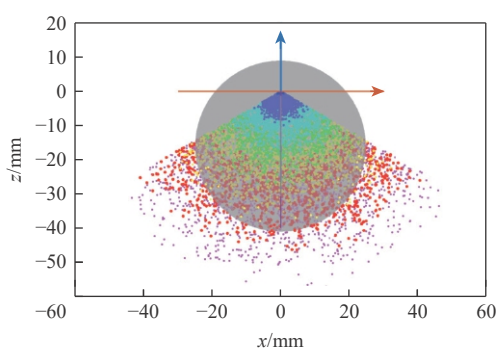


图 9　末端工作空间 *x*-*z* 投影

Fig. 9　*x*-*z* projection of end workspace

$$a_1^2 + (a_2+d_4)^2 + 2a_1(a_2+d_4)\cos\theta_2 = P_x^2 + P_y^2 \tag{8}$$

$$\theta_2 = \pm\arccos\frac{P_x^2+P_y^2-a_1^2-(a_2+d_4)^2}{2_{a2}(a_2+d_4)} \tag{9}$$

由式 (6) 和式 (7) 可得

$$[a_1+(a_2+d_4)\cos\theta_2]\cos\theta_1 - (a_2+d_4)\sin\theta_2\sin\theta_1 = P_x \tag{10}$$

$$[a_1+(a_2+d_4)\cos\theta_2]\sin\theta_1 - (a_2+d_4)\sin\theta_2\cos\theta_1 = P_y \tag{11}$$

由三角函数求得:

$$\sin\theta_1 = \frac{(a_1+(a_2+d_4)\cos\theta_2)P_y - (a_2+d_4)\sin\theta_2 P_x}{(a_1+(a_2+d_4)\cos\theta_2)^2 + (a_2+d_4)^2\sin^2\theta_2} \tag{12}$$

$$\cos\theta_1 = \frac{(a_1+(a_2+d_4)\cos\theta_2)P_x - (a_2+d_4)\sin\theta_2 P_y}{(a_1+(a_2+d_4)\cos\theta_2)^2 + (a_2+d_4)^2\sin^2\theta_2} \tag{13}$$

$$\theta_1 = \arctan\frac{\sin\theta_1}{\cos\theta_1} = \arctan\frac{(a_1+(a_2+d_4)\cos\theta_2)P_y - (a_2+d_4)\sin\theta_2 P_x}{(a_1+(a_2+d_4)\cos\theta_2)P_x - (a_2+d_4)\sin\theta_2 P_y} \tag{14}$$

因此, Scara 机械臂的逆运动学求解方程分别为式 (5)、式 (9) 和式 (12)。

图 10 为 RCM 机构简图, 图中, L 为器械末端到关节铰链的长度, l_1 为定长连杆长度, l_2 为进给连杆长度, x_1 为水平连杆长度。除去 θ_6 自旋自由度外的 3 个自由度进行逆运动学求解, 即通过${}^0_3\boldsymbol{T}$ 中机器人末端器械的末端点位置坐标 P_x、P_y、P_z 得到 x_1、θ_5、θ_3, 由式 (2) 可得关系式 (15), 求解得到 θ_3、θ_4、d_6 分别如式 (16)～式 (18) 所示。

$$\begin{cases} -d_6\sin\theta_3\sin\theta_4 = P_x \\ d_4 - d_6\cos\theta_4 = P_y \\ -d_6\cos\theta_3\sin\theta_4 = P_z \end{cases} \tag{15}$$

求解 θ_3、θ_4、d_6 表达式分别为

$$\theta_3 = \arctan\frac{P_x}{P_z} \tag{16}$$

$$\theta_4 = \arctan\frac{P_x}{(P_y-d_4)\sin\theta_3} \tag{17}$$

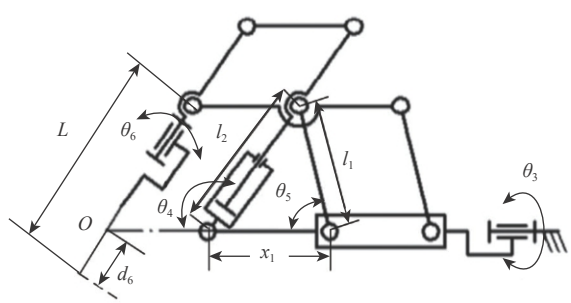


图 10　RCM 机构简图

Fig. 10　Sketch map of RCM mechanism

$$d_6 = -\frac{P_x}{\sin\theta_3 \sin\theta_4} \tag{18}$$

由图 10 所示几何关系可得式 (19)～式 (22), 整理可得式 (23) 和式 (24)。

$$l_2 = L - d_6 \tag{19}$$

$$\cos(\pi - \theta_4) = \frac{x_1^2 + l_2^2 - l_1^2}{2_{x1} l_2} \tag{20}$$

$$\sin\theta_5 = \frac{l_2 \sin\theta_4}{l_1} \tag{21}$$

$$\cos\theta_5 = \frac{x_1^2 + l_1^2 - l_2^2}{2_{x1} l_1} \tag{22}$$

$$x_1 = -l_2\cos\theta_4 \pm \sqrt{l_2^2\cos^2\theta_4 - l_2^2 + l_1^2} \tag{23}$$

$$\theta_5 = \arctan\frac{2x1 l_2 \sin\theta_4}{x_1^2 + l_1^2 - l_2^2} \tag{24}$$

3 机器人主从映射模型

本节介绍机器人系统的主从映射模型, 机器人的映射关系分为 2 个阶段: ①机器人 RCM 点调整阶段; ②RCM 末端运动阶段。

3.1 主从运动映射模型

本节介绍手术所需运动、主从自由度映射模型。图 11 为手术关键阶段示意图。注射针初始位置(针尖与 RCM 点重合)时 RCM 点可视, 即注射针针尖, 如图 11(a) 所示。根据医生分析, 可将从手机器人手术运动分解为 3 个阶段。

第 1 阶段: 粗定位阶段。注射针位于初始位置, Scara 机械臂运动至针尖(RCM 点)与巩膜套管口重合, 如图 11(a) 和图 11(b) 所示, 重合后进入第 2 阶段。

第 2 阶段: 精定位阶段。机器人 RCM 运动(俯仰、偏摆、进给)实现针尖由初始位置(RCM 点)运动至目标穿刺点上方。

第 3 阶段: 穿刺注射阶段。视网膜穿刺、微流泵注射。所需运动与第 2 阶段一致, 都为 RCM 运动。如图 11(b) 和图 11(c) 所示穿刺完成, 穿刺后微流泵进行自动注射。

综上, 实现手术所需运动共计 7 个自由度: 粗定位阶段: Scara 机械臂完成 RCM 点在机器人坐标系下的 x、y、z 方向平移; 精定位、穿刺注射阶段: RCM 机构完成围绕巩膜开口的俯仰 $\boldsymbol{p}$、偏摆 $\boldsymbol{y}$、沿器械轴线的进给 $\boldsymbol{f}$ 及注射针自旋 $\boldsymbol{r}$, 该运动即为 RCM 运动, 如图 12 所示。

图 13 为机器人系统及手术阶段。主手摇杆具备 3 个方向轴及 1 个按钮 B, B=0 和 B=1 分别对应按钮开/闭状态。各轴示意如图 13 中ⓐ所示, 根据按钮按起按下状态, 摇杆等效为 6 轴自由度矢量分别为 $\boldsymbol{X}_0$、$\boldsymbol{Y}_0$、$\boldsymbol{Z}_0$、$\boldsymbol{X}_1$、$\boldsymbol{Y}_1$、$\boldsymbol{Z}_1$。自旋控制器单元左右旋记为 $\boldsymbol{r}$, 综上共计 7 自由度对应上文提及的 7 自由度。

RCM 点在 O_{base} 基坐标系下的 x、y、z 方向平移, 记为 $\boldsymbol{X}_{\text{RCM}}$、$\boldsymbol{Y}_{\text{RCM}}$、$\boldsymbol{Z}_{\text{RCM}}$; RCM 运动中的俯仰 $\boldsymbol{p}$、偏摆 $\boldsymbol{y}$、进给 $\boldsymbol{f}$ 运动, $\boldsymbol{r}$ 为 RCM 运动中的自旋指令, 如图 13 中ⓓ所示。

$\boldsymbol{X}_1$、$\boldsymbol{Y}_1$、$\boldsymbol{Z}_1$ 映射至 $\boldsymbol{X}_{\text{RCM}}$、$\boldsymbol{Y}_{\text{RCM}}$、$\boldsymbol{Z}_{\text{RCM}}$; $\boldsymbol{X}_0$、$\boldsymbol{Y}_0$、$\boldsymbol{Z}_0$ 映射至俯仰 $\boldsymbol{p}$、偏摆 $\boldsymbol{y}$、进给 $\boldsymbol{f}$; 注射针自旋采用自旋控制器控制不在摇杆映射关系内。综上控制映射关系得到

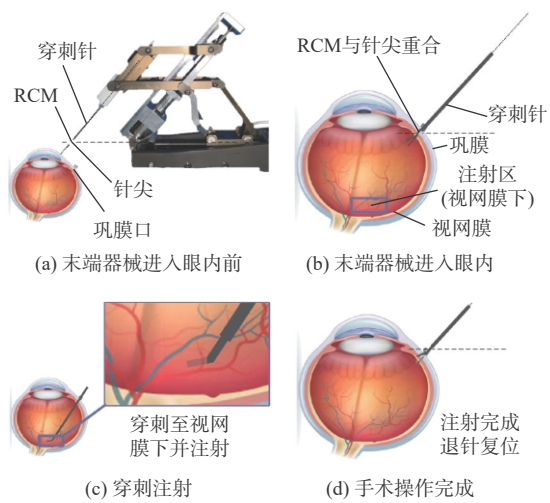
$$\begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{Y}_1 \\ \boldsymbol{Z}_1 \end{pmatrix}_{\text{base}} \Rightarrow \begin{pmatrix} \boldsymbol{X}_{\text{RCM}} \\ \boldsymbol{Y}_{\text{RCM}} \\ \boldsymbol{Z}_{\text{RCM}} \end{pmatrix}_{\text{base}} \qquad B = 1 \tag{25}$$

$$\begin{pmatrix} \boldsymbol{X}_0 \\ \boldsymbol{Y}_0 \\ \boldsymbol{Z}_0 \end{pmatrix}_{\text{base}} \Rightarrow \begin{pmatrix} \boldsymbol{p} \\ \boldsymbol{y} \\ \boldsymbol{f} \end{pmatrix}_{\text{RCM}} \qquad B = 0 \tag{26}$$

3.2 速度映射模型

3.2.1 粗定位速度映射

第 1 阶段粗定位主手摇杆按钮按下, 切换至粗



(a) 末端器械进入眼内前　(b) 末端器械进入眼内

(c) 穿刺注射　(d) 手术操作完成

图 11　手术关键阶段示意图

Fig. 11　Schematic diagram of key stages of surgery

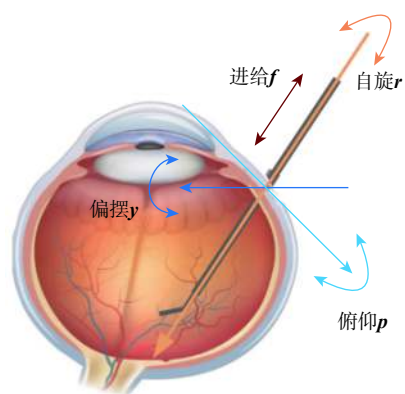


图 12　RCM 运动示意图

Fig. 12　Schematic representation of RCM movement

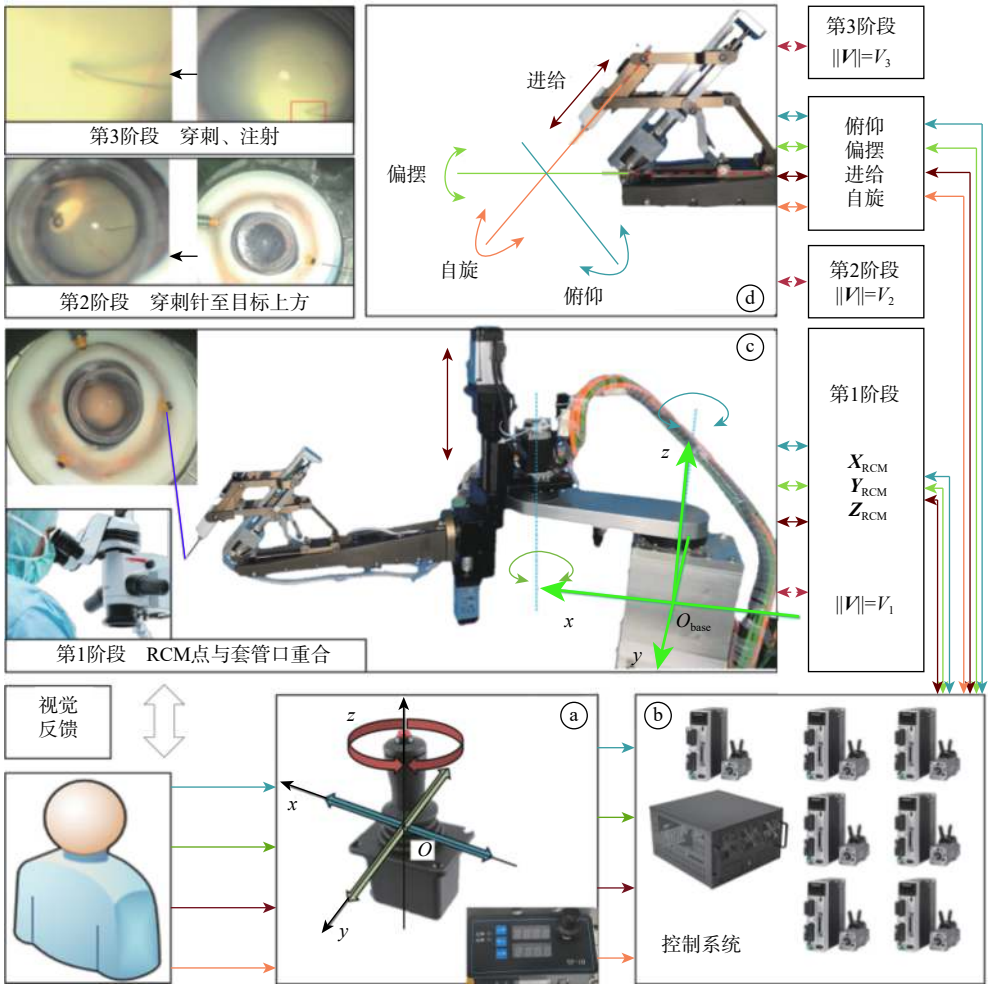


图 13　机器人系统及手术阶段

Fig. 13　Robotic system and surgical stages

定位模式, 末端 RCM 机构姿态保持, 根据式 (25) 的映射关系得: 主手摇杆($\boldsymbol{X}_1$, $\boldsymbol{Y}_1$, $\boldsymbol{Z}_1$)映射至 RCM 点在基坐标系下进行 x、y、z 这 3 个方向的平动, 第 1 阶段速度映射关系为

$$\boldsymbol{V}=\begin{bmatrix}V_x\\V_y\\V_z\end{bmatrix}=\begin{bmatrix}k_xV_{1x\max}\\k_yV_{1y\max}\\k_zV_{1z\max}\end{bmatrix}\tag{27}$$

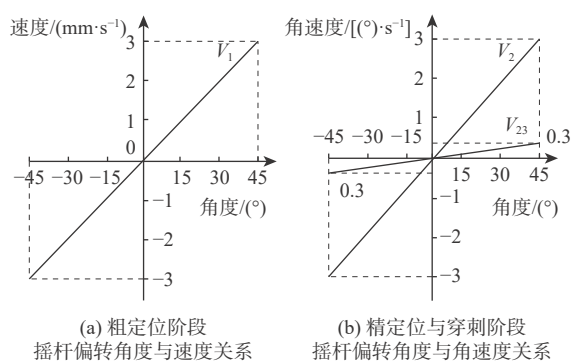
其中:

$$\begin{pmatrix}k_x\\k_y\\k_z\end{pmatrix}=\begin{pmatrix}\alpha/\alpha_{x\max}^{-1}\\\beta/\beta_{y\max}^{-1}\\\gamma/\gamma_{z\max}^{-1}\end{pmatrix}\tag{28}$$

式中: $\boldsymbol{V}$为从手机械臂速度; V_x、V_y、V_z分别为 RCM 点 x、y、z 方向的速度, α、β、γ 为摇杆 x、y、z 三轴偏离初始零位的角度, $\{\alpha,\beta,\gamma\}\in[-45°,45°]$, $\alpha_{x\max}$、$\beta_{y\max}$、$\gamma_{z\max}$ 为 45°; 即 $\{k_x, k_y, k_z\}\in[-1,1]$。图 14 为主手摇杆轴偏移角度与从手末端速度关系, 粗定位阶段摇杆轴偏离角度-速度关系如图 14(a) 所示。

3.2.2　RCM 运动速度映射

第 2 阶段精定位, 按钮按起状态进入精定位模式的 RCM 运动, 此模式主从映射关系根据式 (26) 得:主手摇杆($\boldsymbol{X}_1$, $\boldsymbol{Y}_1$, $\boldsymbol{Z}_1$)映射至穿刺器械在 RCM 坐



(a) 粗定位阶段
摇杆偏转角度与速度关系

(b) 精定位与穿刺阶段
摇杆偏转角度与角速度关系

图 14　主手摇杆轴偏移角度与从手末端速度关系

Fig. 14　Relationship between axis offset angle of master joystick and speed of slave end

标系下的($\boldsymbol{p}$, $\boldsymbol{y}$, $\boldsymbol{f}$), 为表达一致, 第 2 阶段的角速度和速度则记为$\boldsymbol{\omega}_2$, $\boldsymbol{p}$、$\boldsymbol{y}$、$\boldsymbol{f}$记为$\boldsymbol{\omega}_{p2}$、$\boldsymbol{\omega}_{y2}$、$\boldsymbol{\omega}_{f2}$, 映射关系为

$$\boldsymbol{\omega_2}=\begin{bmatrix}\omega_{p2}\\\omega_{y2}\\\omega_{f2}\end{bmatrix}=\begin{bmatrix}k_x\omega_{p2\max}\\k_y\omega_{y2\max}\\k_z\omega_{f2\max}\end{bmatrix}\tag{29}$$

第 3 阶段穿刺执行中运动映射关系相同于第 2 阶段, 速度通过改变$\boldsymbol{\omega}_{3\max}$实现速度减小。第 2 阶

段和第3阶段RCM的末端运动速速如图14中ⓑ所示。

图15为机器人主从映射关系，需要说明的是，在实际控制中，控制系统在生成从机械臂的速度指令后，先将速度指令积分为位移指令，然后再依据逆运动学模型式(5)、式(9)和式(14)解算从机械臂各个关节的控制指令，并控制从机械臂各关节完成预期运动，目的是考虑到手术安全性的要求，需要实时监测机器人的位移以避免损伤极其脆弱的眼底组织。

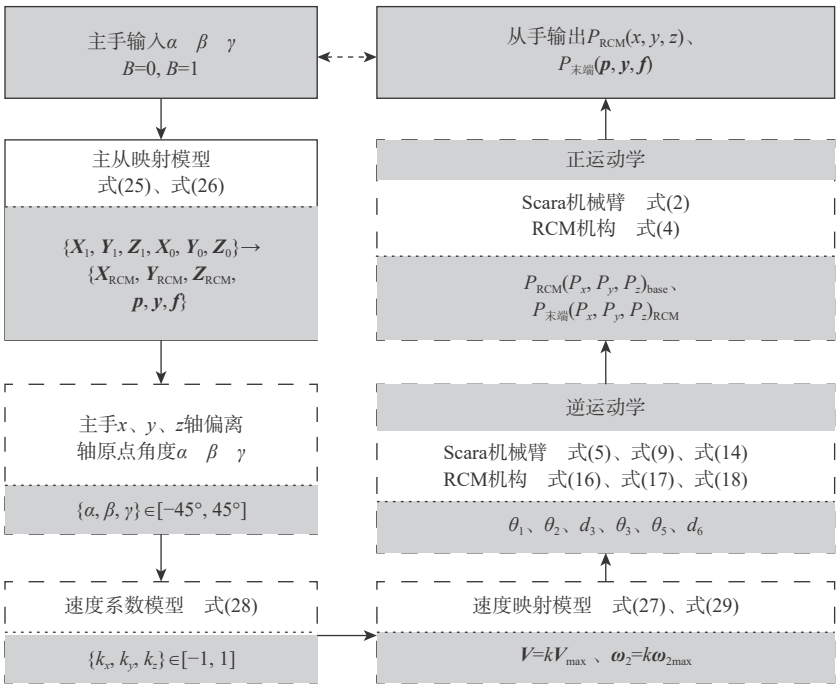


图15 机器人主从映射关系

Fig. 15 Master-slave mapping relationship of robot

4 实验及结果讨论

4.1 实验设置

通过离体猪眼球视网膜下穿刺注射实验对机器人系统及控制方法进行验证。图16为手术场景图，除搭建的手术机器人外还有特种注射针(41G)、显微镜、眼内光源(玻切机提供)、眼压泵及手术必备品。

实验在宰杀4 h内的离体猪眼球中进行，2组实验都对猪眼球视网膜下穿刺、注射操作，时长为10 min。第1组实验操作者徒手进行；第2组实验操作者使用摇杆控制从手机器人、微流泵完成。每组选取5只猪眼球进行实验。实验记录采用Leica M220 F12医用显微镜。

实验结果从3个方面进行评价，分别为：漏液程度、穿刺孔洞创伤、实验后眼球完整程度。选取此评价标准的依据来自视网膜下注射的医学实验评价规则[19]。

渗液程度即注射物由穿刺孔和注射针间隙渗漏到玻璃体腔的程度。间隙形成的一个原因是注射针的颤抖形成拉扯，造成穿刺孔径直径因弹性变

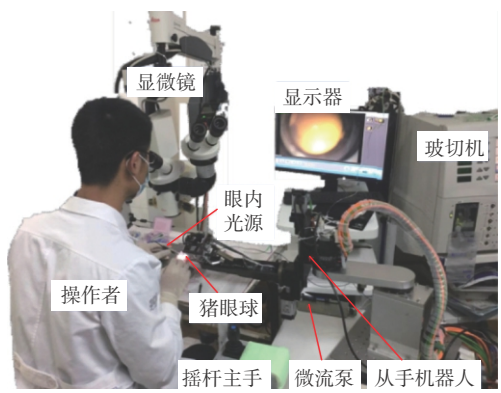


图16 手术场景图

Fig. 16 Operation scene diagram

形而大于注射针。另一部分原因则是颤抖造成的永久撕裂性创伤，发生渗液。

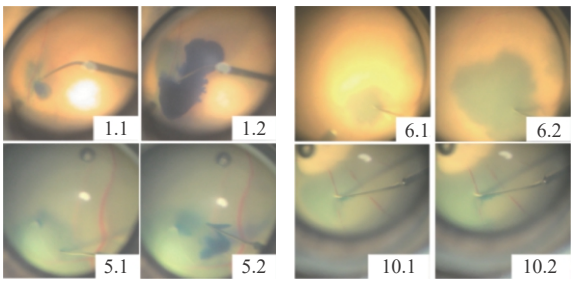
穿刺孔洞创伤程度指由注射针对视网膜造成的永久性创伤，包括注射必需的注射针直径大小和由于注射针颤抖而形成的永久性撕裂创伤。

眼球完整性指实验结束后眼球是否有实验过程中非必需的破坏如眼底刺穿、巩膜口撕裂等。

4.2 实验结果

总计对10只猪眼球进行视网膜穿刺及注射，

其中徒手操作和机器人辅助各 5 只, 部分图像展示如图 17 所示, *X*.1 表示穿刺注射开始图像, *X*.2 表示注射结束图像, *X* 表示为眼球编号, 1～5 号猪眼球为徒手操作、6～10 号猪眼球为机器人辅助操作, “+” 符号个数越多, 表示创伤越大, 漏液越严重, 分析结果如表 3 所示。



(a) 徒手操作图像　　(b) 机器人辅助操作图像

图 17　注射结束前后的图像

Fig. 17　Images before and after injection

表 3　实验结果

Table 3　Experimental result

操作方式及编号	玻璃体腔漏液	视网膜创伤孔大小	眼球完整性
徒手操作1	+++	+++	完整
徒手操作2	+++	++	完整
徒手操作3	+	++	完整
徒手操作4	+++	+++	完整
徒手操作5	+	++++	完整
机器人辅助6	+	+	完整
机器人辅助7	++	++	完整
机器人辅助8	+	+	完整
机器人辅助9	+	+	完整
机器人辅助10	+	+	完整

4.3　讨　论

眼科手术中, 医生手部的生理学颤抖幅值为 100 μm 左右[20], 其大小已超过注射针头直径大小。在徒手穿刺后进行注射时需要维持手部固定, 但是生理性颤抖导致难以维持该动作。手部的生理性颤抖带动注射针在穿刺孔内振动, 进而导致对视网膜的弹性拉扯和永久性撕裂形成渗液。

机器人辅助操作避免手部颤抖的影响, 提高穿刺成功后注射针的稳定维持能力。因此, 6～10 号眼球中未出现穿刺孔洞被拉扯变大和玻璃体腔出现渗液的现象, 进而注射针对视网膜的永久性创伤直径也越小。

5　结　论

1) 为解决眼内操作时巩膜约束运动、注射时稳定器械、微流量(微流速)注射等问题提供了一种可行方案, 初步实现了视网膜下机器人辅助注射。

2) 通过离体猪眼球视网膜下注射实验验证了机器人辅助操作系统末端注射针稳定维持能力和运动的精确性, 机器人辅助操作比徒手操作对视网膜造成的创伤更小, 注射更稳定。

参考文献（References）

[1] MILLER J W, D'ANIERI L L, HUSAIN D, et al. Age-related macular degeneration (AMD): A view to the future[J]. Journal of Clinical Medicine, 2021, 10(5): 1124.

[2] HANDA J T, RICKMAN C B, DICK A D, et al. A systems biology approach towards understanding and treating non-neovascular age-related macular degeneration[J]. Nature Communications, 2019, 10(1): 106-116.

[3] ZHU J, LAMBA D A. Small molecule-based retinal differentiation of human embryonic stem cells and induced pluripotent stem cells[J]. Bio-Protocol, 2018, 8(12): e2882.

[4] SHIRAI H, MANDAI M, MATSUSHITA K, et al. Transplantation of human embryonic stem cell-derived retinal tissue in two primate models of retinal degeneration[J]. Proceedings of the National Academy of Sciences of the United States of America, 2016, 113(1): 81-90.

[5] YE K, TAKEMOTO Y, ITO A, et al. Reproducible production and image-based quality evaluation of retinal pigment epithelium sheets from human induced pluripotent stem cells[J]. Scientific Reports, 2020, 10(1): 1748-1760.

[6] CHAO J R, LAMBA D A, KLESERT T R, et al. Transplantation of human embryonic stem cell-derived retinal cells into the subretinal space of a non-human primate[J]. Translational Vision Science & Technology, 2017, 6(3): 4.

[7] HE C Y, YANG Y. Multipoint force-constrained admittance control for retinal surgical robot[J]. Journal of Mechanical Engineering, 2021, 57(9): 12-18.

[8] XIAO J J, YANG Y, SHEN L J, et al. A robotic system for retinal vascular bypass surgery[J]. Robot, 2014, 36(3): 293-299(in Chinese).

[9] HE C Y, YANG Y, LIANG Q F, et al. Applications and research progress of robot assisted eye surgery[J]. Robot, 2019, 41(2): 265-275(in Chinese).

[10] ÜNERI A, BALICKI M A, HANDA J, et al. New steady-hand eye robot with micro-force sensing for vitreoretinal surgery[C]//2010 3rd IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics. Piscataway: IEEE Press, 2010: 814-819.

[11] HE X C, ROPPENECKER D, GIERLACH D, et al. Toward clinically applicable steady-hand eye robot for vitreoretinal surgery[C]// Proceedings of ASME 2012 International Mechanical Engineering Congress and Exposition. New York: ASME, 2013: 145-153.

[12] EDWARDS T L, XUE K, MEENINK H C M, et al. First-in-human study of the safety and viability of intraocular robotic surgery[J]. Nature Biomedical Engineering, 2018, 2(9): 649-656.

[13] GIJBELS A, SMITS J, SCHOEVAERDTS L, et al. In-human robot-assisted retinal vein cannulation, D world first[J]. Annals of Biomedical Engineering, 2018, 46(10): 1676-1685.

[14] GIJBELS A, WILLEKENS K, ESTEVENY L, et al. Towards a

clinically applicable robotic assistance system for retinal vein cannulation[C]//2016 6th IEEE International Conference on Biomedical Robotics and Biomechatronics. Piscataway: IEEE Press, 2016: 284-291.

[15] SU P, DENG S J, HUANG L, et al. Analysis and evaluation of a robotic trephination in penetrating keratoplasty[J]. Journal of Medical Devices, 2016, 10(2): 024503.

[16] XIAO J J, HUANG L, SHEN L J, et al. Design and research of a robotic aided system for retinal vascular bypass surgery[J]. Journal of Medical Devices, 2014, 8(4): 044501.

[17] HE C Y, HUANG L, YANG Y, et al. Research and realization of a master-slave robotic system for retinal vascular bypass surgery[J]. Chinese Journal of Mechanical Engineering, 2018, 31(1): 1-10.

[18] HUANG L, YANG Y, SU P, et al. Type synthesis of 1R1T remote center of motion mechanisms[J]. Journal of Mechanical Engineering, 2015, 51(13): 131-136(in Chinese).

[19] SCRUGGS B A, JIAO C H, CRANSTON C M, et al. Optimizing donor cellular dissociation and subretinal injection parameters for stem cell-based treatments[J]. Stem Cells Translational Medicine, 2019, 8(8): 797-809.

[20] SINGH S P N, RIVIERE C N. Physiological tremor amplitude during retinal microsurgery[C]//Proceedings of the IEEE 28th Annual Northeast Bioengineering Conference. Piscataway: IEEE Press, 2002: 171-172.

Design and implementation of robot-assisted subretinal injection system

WANG Zhaodong[1], GUANG Chenhan[1], WANG Liqiang[2], ZONG Junjie[1], ZHENG Yu[1], YANG Yang[1,*]

(1. School of Mechanical Engineering and Automation, Beihang University, Beijing 100191, China;
2. Department of Ophthalmology, PLA General Hospital, Beijing 100853, China)

Abstract: Subretinal injection of human embryonic stem cells is an effective procedure for treating retinal degeneration, which imposes high demands on the precision, stability, and safety of the surgical procedure. To address these challenges, a robot-assisted retinal surgery system based on a master-slave control approach is proposed. The system utilizes a master robot to control a slave robot during the subretinal injection operation. By constructing the direct and inverse kinematic models of the slave robot, the precise motion requirements for the subretinal injection are determined. Additionally, the system incorporates the concept of the remote center of motion (RCM) to adjust the position of the surgical instruments during remote motion. A velocity mapping model is derived to facilitate the adjustment of the RCM point, state transitions, and separation process during the procedure.The validation of the system 's accuracy and stability is conducted through ex vivo porcine eyeball subretinal puncture injection experiments, simulating the conditions of intraocular surgery. The results demonstrate that the robot-assisted system exhibits stable needle maintenance and precise motion, indicating its superiority over manual procedures in terms of reducing trauma to the retina and achieving more stable injections. This technology represents a advancement in microsurgery techniques, enhancing the precision and safety of subretinal injection procedures for the treatment of retinal degeneration.

Keywords: robot-assisted retinal surgery; master-slave control; remote center of motion; intraocular surgery; Microsurgery

Received: 2021-11-05; **Accepted**: 2022-02-13; **Published Online**: 2022-03-25 10:47

URL: kns.cnki.net/kcms/detail/11.2625.V.20220324.1616.004.html

Foundation item: National Key R&D Program for Intelligent Robots (2007YFB1302700)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0670

基于综合评价优化的民机顶层需求指标权衡

范周伟[1,2]，余雄庆[1,*]，戴亚林[1]

(1. 南京航空航天大学 航空学院 飞行器先进设计技术国防重点学科实验室，南京 210016；
2. 中国商用飞机有限责任公司 上海飞机设计研究院，上海 201210)

摘 要：确定顶层需求指标是民机设计流程之初的重要工作。针对民机顶层设计工作，提出一套民机顶层需求定量权衡方法。根据民机顶层需求与概念设计之间的关系，将概念设计方案的特性参数视为一组可行的顶层需求。将顶层需求的权衡分析问题转换为对飞机总体设计参数的优化问题。选择“经济性、舒适性、环保性、适应性、安全性/可靠性”这5条判断准则，建立民机顶层需求综合评价模型，评估顶层需求的综合竞争力。以竞争力最大为目标，优化概念方案，确定最优的顶层需求组合。在宽体客机案例中，得到了比初始需求更合理的顶层需求，验证了所提方法的可行性。

关 键 词：民机；顶层需求；综合评价；优化；权衡分析
中图分类号：V221+.7
文献标志码：A **文章编号**：1001-5965(2023)09-2415-08

民航客机的研发工作是一项复杂的大系统工程，如何确定合理的顶层需求指标对项目的成败影响巨大。然而民机的设计过程中涉及多个学科且各学科间相互交叉耦合，为顶层需求的定义带来了困难。在以往的飞机设计中，顶层需求指标的确定主要依赖于用户的期望和设计人员的经验，指标合理性缺乏严格的理论依据和科学的方法。随着系统工程方法在航空领域的应用，飞机顶层需求受到了越来越多的关注，工业界迫切需要一种对飞机顶层需求定量权衡分析的方法。

根据系统工程的一般技术流程[1]，顶层需求可以通过对客户等利益相关方的需求分析而得到。Eres等[2]提出了一套价值驱动的方法，在用户需求与工程指标之间建立联系。黄俊等[3]建立了作战飞机总体设计的综合评价准则，评价设计方案能否充分反映用户的需求。

然而，由利益相关方需求转换为飞机顶层需求的过程中，如何确定顶层需求的合理取值仍是待解决的问题。Parker[4]基于工程项目的经验，总结了飞机设计过程中任务需求和飞机概念设计之间的关系，Parker[4]认为利用飞机概念设计方法可以辅助定义合理的飞机顶层任务需求。Peteilh等[5]以顶层需求中的起飞场长需求为例，使用基于统计数据的模型进行优化，确定起飞场长需求的合理取值，该方法仅对起飞场长需求进行量化分析，对于多条顶层需求的同时权衡仍需要继续研究。

本文提出一套对民机顶层需求进行定量分析权衡的方法。本文方法基于顶层需求与概念设计的关系，通过对飞机总体参数进行优化，得到最为合适的飞机顶层需求指标组合。

1 顶层需求和概念设计

飞机顶层需求驱动飞机概念设计[6]。民机的顶层需求包括设计航程、商载能力、巡航马赫数、起

收稿日期：2021-11-05；录用日期：2022-01-05；网络出版时间：2022-01-25 18:15
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220121.2017.002.html
基金项目：国家自然科学基金(12032011)；民用飞机专项科研项目
*通信作者. E-mail：yxq@nuaa.edu.cn

引用格式：范周伟，余雄庆，戴亚林. 基于综合评价优化的民机顶层需求指标权衡[J]. 北京航空航天大学学报，2023，49(9)：2415-2422.
FAN Z W，YU X Q，DAI Y L. Trade-off for top-level requirements of commercial aircraft using comprehensive evaluation and optimization[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2415-2422 (in Chinese).

降距离等。这些顶层需求是概念设计的输入，根据顶层需求可以确定飞机概念设计所需的各种参数。

同时，顶层需求也是概念设计必须达到的技术指标[7]。概念设计方案的相关特性参数，必须满足顶层需求指标的要求。如果概念设计方案的特性参数无法满足顶层需求，必须对方案进行调整。在某些情况下，基于现有的技术水平无法设计出满足所有顶层需求的概念方案，则顶层需求不具备可行性，必须对顶层需求做出调整[8]。

正是因为顶层需求和飞机概念设计方案特性参数之间的这层关系，可以认为已知的飞机概念设计方案特性参数必定对应了一组可行的顶层需求。而以特性参数的取值作为顶层需求指标，至少存在一个能够满足顶层需求指标的可行方案。

对于飞机这样的复杂系统，其顶层需求定义往往不是顺序的。通过顶层需求定义与概念设计之间的迭代，可以对顶层需求进行权衡。

在顶层需求尚未明确时，可以将概念设计方案的特性参数视为待选的顶层需求指标。如果遍历概念设计的所有设计空间，则可以找到一组最为合理的顶层需求。通过优化算法，可以加快搜索速度。然而，顶层需求条目众多，不同顶层需求之间甚至存在冲突，如何确定优化的目标成为定义优化问题的关键。因此，需要建立顶层需求综合评价模型，以顶层需求的综合评价结果作为优化目标。

2 顶层需求的竞争力评估

顶层需求综合评价模型对各项顶层需求进行综合考虑，形成以“竞争力”描述的评估结果。

不同顶层需求对竞争力影响的权重大小不同，必须选择合适的判断准则，对顶层需求的权重进行判断。

2.1 判断准则

关于民机的顶层需求的判断准则，已经有不少研究作为参照。Dirks和Meller[9]从经济性、通用性、性能、舒适性、市场适应性这几个方面对民机进行评价。索欣诗等[10]采用层次分析法，以经济性、舒适性、环保性、适应性这4个方面作为判断准则，建立了民机总体方案的综合评价模型。宋笔锋等[11]建立了大型客机安全性、经济性、舒适性、环保性设计与评价准则体系。这些研究从不同角度出发，最后建立的判断准则存在一定的差异。

结合民机的特点，在中国大型民用客机研制所追求的“新四性”（安全性、经济性、舒适性、环保性）[12]的基础上，选择5条判断准则。

1）经济性。对于民机而言，更高的经济性能够节省客户的使用成本，带来更多的经济收益。经济性判断准则与飞机的使用成本、油耗等有着直接的关联。除此之外，航程、有效载荷等代表飞机运输能力的参数也与经济性准则有关。

2）舒适性。根据统计，在航班票价和离港时间相接近时，乘客往往会选择舒适性更高的航班[13]。提高飞机的舒适性，有助于航空公司提高上座率、增加经济收益。

舒适性跟乘客的主观感受有关。影响乘客舒适性体验的因素有很多。Richards和Jacobson[14]研究发现，增大座椅宽度或腿部空间可以大幅提高乘客的满意度。

3）环保性。随着民众环保意识的日益增强，民航飞机的环保性也受到越来越高的重视。民机对环境造成的污染主要包括噪声污染和尾气排放污染[15]。

4）适应性。航空公司在考虑新机型选购时，对飞机适应性的评估尤为重要[16]。飞机的适应性主要体现飞机对目标航线和目标机场的适应能力。航线适应性的评价指标包括航程、商载能力等。机场适应性的评价指标包括起降场长、进近速度等。

5）安全性/可靠性。对于民机而言，保障飞行的安全至关重要，而可靠性设计是确保安全的关键。飞机的维修性、测试性、保障性与可靠性紧密相关，同样属于“安全性/可靠性”准则。

确定民机顶层需求判断准则之后，采用群组综合层次分析法[10]，通过综合多位专家的打分结果，计算判断准则的权重。判断准则的层次结构如图1所示。

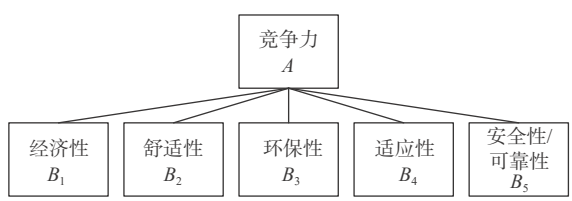


图1 民机顶层需求判断准则层次结构

Fig. 1 Hierarchy of criteria for top-level aircraft requirements

2.2 顶层需求的权重

根据判断准则，可以对顶层需求的权重进行打分。利用质量功能展开（quality function deployment, QFD）[17]中的关系度矩阵，建立飞机顶层需求与各项判断准则之间的关系。图2为关系度矩阵的主要组成，其中，ω_i为第i条判断标准的权重，$R_{i,j}$为第i条判断标准与第j条需求之间的关系度，W_j为第j条需求的相对权重。W_j的计算式为

$$W_j=\frac{\sum_{i=1}^{m}(\omega_i R_{i,j})}{\sum_{j=1}^{n}\sum_{i=1}^{m}(\omega_i R_{i,j})} \tag{1}$$

式中: $R_{i,j}$ 的取值为“0, 1, 3, 5, 7, 9”, 表示的相关程度分别为: 无关、弱相关、较弱相关、中相关、较强相关、强相关。

			1	2	3	4
	权重	顶层需求 / 判断准则	需求1	需求2	需求3	需求4
1	ω_1	准则1	$R_{1,1}$	$R_{1,2}$	$R_{1,3}$	$R_{1,4}$
2	ω_2	准则2	$R_{2,1}$	$R_{2,2}$	$R_{2,3}$	$R_{2,4}$
3	ω_3	准则3	$R_{3,1}$	$R_{3,2}$	$R_{3,3}$	$R_{3,4}$
		相对权重	W_1	W_2	W_3	W_4

图 2　关系度矩阵的主要组成

Fig. 2　Main components of relationship matrix

2.3　顶层需求指标的标准化

不同需求指标取值的数量级往往相差很大, 不具备一致性, 难以对各项需求指标进行综合评价。采用式 (2) 的非线性 S 型可导函数[18], 将民机顶层需求指标的取值标准化为 (0, 1) 区间内的数值。

$$\overline{V}_j=\frac{1}{1+T\mathrm{e}^{-kS\left[\frac{6}{V_{j\max}-V_{j\min}}\left(V_j-\frac{V_{j\max}+V_{j\min}}{2}\right)\right]}} \tag{2}$$

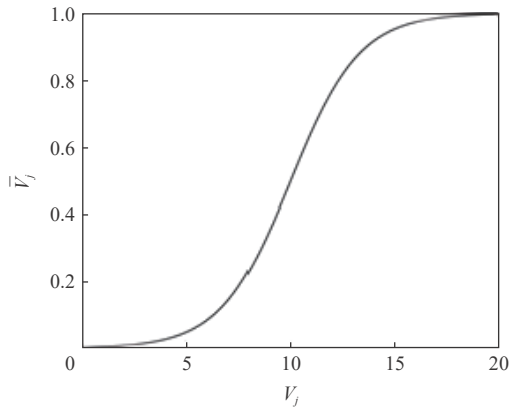
式中: V_j为变换前的顶层需求指标值, $\overline{V}_j$为归一化后的数值; $V_{j\min}$和$V_{j\max}$分别为顶层需求指标的最小和最大值, 可以通过统计类似的飞机产品数据获得; S 和 T 为曲线形状调节参数, S一般取 1～2, T一般取 1; k 为曲线的方向调节参数, 对于取值越大越好的顶层需求, k 值为 1, 对于取值越小越好的顶层需求, k 值为−1。

图 3 为 S 型可导函数, 其中, V_j取值在闭区间 [0, 20], S=2, T=1。从图中可以看出, 标准化处理后的数值分布可以体现需求指标的饱和特性。在当前技术能力限制下, 一味增大或减小某项需求指标的值, 所带来的收益增量趋于平缓。

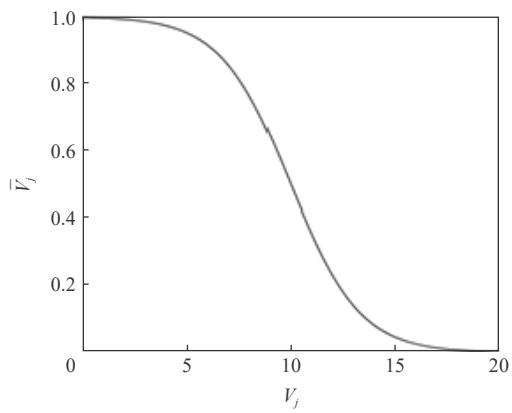
2.4　竞争力计算

标准化后的顶层需求指标数值$\overline{V}_j$反映了该项需求的严格程度。$\overline{V}_j$值越接近 1, 则该项需求越严格, 但也意味着满足该项需求的飞机在同类机型中的表现更加优异。而对于整个飞机而言, 由于其产品的复杂性, 满足某项需求的同时, 可能会牺牲其他的需求。根据 2.2 节的介绍, 顶层需求的权重各不相同。因此, 整个飞机的竞争力为综合考虑各项顶层需求后的结果。通过式 (3) 计算整个飞机的竞争力 A。A 的值越大, 代表飞机的竞争力越强。

$$A=\sum_{j=1}^{n}(\overline{V}_j W_j) \tag{3}$$



(a) k=1, 单调递增



(b) k=−1, 单调递减

图 3　S 型可导函数

Fig. 3　S-shaped derivative function

3　顶层需求的权衡

将飞机不同顶层需求的权衡问题转换为竞争力的优化问题。图 4 为顶层需求权衡的架构, 其可通过设计结构矩阵[19] 描述。

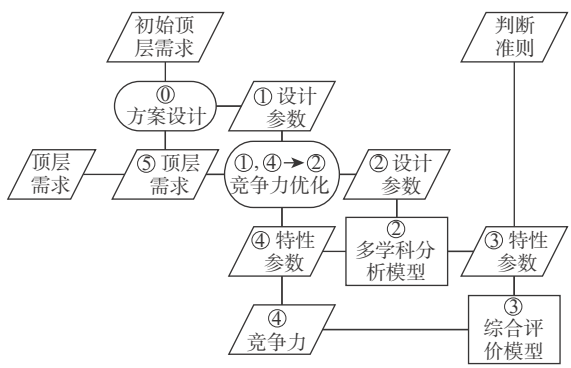


图 4　顶层需求权衡架构

Fig. 4　Architecture of trade-off of top-level aircraft requirements

根据初步拟定的顶层需求, 设计出初始的飞机概念方案。对概念方案进行分析, 得到概念方案的

特性参数。将特性参数视为一组可行的顶层需求，通过综合评价模型计算顶层需求的竞争力。改变概念方案的设计参数，经过多学科分析模型的计算，得到不同的特性参数，对应了多组可行的顶层需求。对顶层需求的竞争力进行优化。最终得到竞争力最优的顶层需求指标。

3.1 多学科分析模型

竞争力计算所依据的各项特性参数由多学科分析模型获得。采用一款民机总体参数多学科分析程序[20]，从推进系统、几何、气动、性能、噪声、排放及成本等多个学科，对民机概念方案进行分析，得到飞机的各项特性参数。

在该程序中，缺少对飞机“安全性/可靠性”参数的分析。签派可靠度是民机的关键可靠性指标之一，能够综合反映飞机各系统的故障率、故障间隔、维修时间等信息。因此，以签派可靠度作为对飞机“安全性/可靠性”评估的补充。采用Fielding[21]总结的方法，对飞机的签派可靠度进行计算。对于长航程飞机，延误率 D_{R} 的计算如下：

$$D_{\mathrm{R}}=C_1+\left(C_2W_{\mathrm{T}}\times10^{-6}\right)+\left(C_3W_{\mathrm{T}}/t\times10^{-6}\right)+\left(C_4\frac{P}{100t}\right)+\left(\frac{C_5P}{100}\right)+\left(10^{\left(2.29(W_{\mathrm{T}}/t)10^{-6}\right)-0.887}\right)+\left(0.307\,5\lg V_{\mathrm{APP}}^2W_{\mathrm{L}}\times10^{-8}\right)+\left(\frac{T_{\mathrm{SL}}C_6}{100\,000}\right)\tag{4}$$

式中：W_{T} 为飞机的最大起飞质量，lb；t 为平均航线时间，h；P 为乘客数；V_{APP} 为进近速度，ft/s；W_{L} 为飞机的最大着陆质量，lb；T_{SL} 为飞机海平面最大总静推力，lb。对于装配大风扇发动机的远程客机，$C_1=-0.012\,9$，$C_2=0.235$，$C_3=2.175$，$C_4=-0.010\,2$，$C_5=0.041\,9$，$C_6=1.142\,6$。

考虑到随着时间的推移，航空技术不断的发展，飞机的延误率不断下降，对延误率计算式(4)进行修正。根据B767的数据，飞机首次交付时间自1969年起，每往后一年，飞机的延误率降低1%。

最终，飞机的签派可靠度为100%减去飞机的延误率。

3.2 竞争力优化

定义优化问题的三要素是：优化目标、设计变量和约束条件。民机竞争力优化问题以竞争力最大为优化目标。以飞机的机翼几何参数、发动机参数和客舱布置参数等总体设计参数作为设计变量。

优化过程中，通过不断调整设计变量的取值，生成不同的设计方案，经过多学科分析模型的计算，得到设计方案的特性参数(例如，直接使用成本、起飞场长等)。将特性参数视为可行的顶层需求指标，传入综合评价模型，计算得到该组顶层需求的竞争力。

在常规的飞机设计方案优化问题中，往往以给定的顶层需求作为约束，但在竞争力优化问题中，顶层需求是可变的。因此，将适航条例中明确规定的顶层需求作为约束条件，如约束飞机起飞2阶段单发失效的爬升梯度不小于2.4%。除此之外，还有一些设计中的物理约束必须满足，如飞机的任务燃油体积必须小于油箱的容积。

4 案例研究

以一款宽体客机作为研究对象，对其顶层需求进行权衡。根据对利益相关方(主要为航空公司)需求的分析，初步拟定了飞机顶层需求。

4.1 初步顶层需求

通过市场分析和用户需求分析，初步拟定飞机的设计航程为12 000 km，巡航马赫数为0.85。从顶层需求中选择14条需求进行权衡分析。

根据所确定的5条判断标准，使用群组综合层次分析法[10]对判断标准的权重进行打分。最终综合判断矩阵和权重如表1所示。

表1中，B_i 为各项判断准则，ω 为判断准则的权重。综合判断矩阵的一致性比例(consistency ratio, CR)为0.074，小于0.1，满足一致性要求。

表1 综合判断矩阵和权重

Table 1 Comprehensive judgement matrix and weiglas

判断准则	矩阵数据					ω
	B_1	B_2	B_3	B_4	B_5	
B_1	1	5.472	5.472	3	1	0.335
B_2	0.187	1	1.472	0.264	0.173	0.060
B_3	0.187	0.843	1	0.2	0.165	0.051
B_4	0.333	4.045	5	1	0.242	0.163
B_5	1	5.955	6.427	4.91	1	0.391

通过建立如图5所示的判断准则与顶层需求之间的关系度矩阵，由判断准则的权重计算得到顶层需求的相对权重。

通过统计同类机型的数据，确定顶层需求的合理取值范围，得到图5中各项顶层需求的最大和最小值。

通过与竞争机型比较，初步拟定的顶层需求指标及顶层需求的相对权重如表2所示。

4.2 初步概念方案

根据初步顶层需求，设计形成飞机的初步概念方案。图6为飞机初步概念方案的几何外形。飞机采用常规布局，机翼位置为下单翼。飞机装配2台涡扇发动机，发动机短舱采取翼吊式布局。概

				1	2	3	4	5	6	7	8	9	10	11	12	13	14
			改进方向	↓	↓	↑	↑	↑	↑	↓	↓	↓	↑	↓	↓	↓	↑
	相对权重/%	权重	顶层需求 / 判断准则	座公里油耗	座公里成本	座椅宽度	座椅排距	客舱高度	设计航程	展长	起飞场长	着陆场长	单发失效二阶段爬升梯度	进近速度	起降总噪声	航线 NO_x 排放量	签派可靠度
1	33.5	0.335	经济性	7	9				5								
2	6	0.06	舒适性			9	9	7							7		
3	5.1	0.051	环保性												9	9	
4	16.3	0.163	适应性						7	3	9	7	5	7			
5	39.1	0.391	安全性/可靠性								1	1	3	1			7
			权重	2.345	3.015	0.54	0.54	0.42	2.816	0.489	1.858	1.532	1.988	1.532	0.879	0.459	2.737
			相对权重/%	11.09	14.26	2.553	2.553	1.986	13.31	2.312	8.785	7.243	9.4	7.243	4.156	2.17	12.94
			权重排名	4	1	10	10	14	2	12	6	7	5	7	9	13	3
			最小值	0.019 6	0.35	475	775	1.85	9 700	60	2 440	1 700	0.024	64	258	1 200	0.95
			最大值	0.022	0.4	525	850	2.3	15 000	65	3 000	2 000	0.04	82	297	1 400	0.995

图 5　判断准则与顶层需求之间的关系度矩阵

Fig. 5　Relationshis between criteria and top-level aircraft requirements

表 2　初始顶层需求

Table 2　Initial top-level aircraft requirements

顶层需求	初步估值	相对权重/%
座公里油耗/(kg·km^{-1}·座$^{-1}$)	0.022	11.09
座公里成本/(元·km^{-1}·座$^{-1}$)	0.38	14.26
座椅宽度/mm	480	2.553
座椅排距/mm	800	2.553
客舱高度/m	2	1.986
设计航程/km	12 000	13.31
展长/m	61	2.312
起飞场长/m	2 700	8.785
着陆场长/m	1 800	7.243
单发失效二阶段爬升梯度/%	3.5	9.40
进近速度/(m·s^{-1})	70	7.243
起降总噪声/dB	278	4.156
航线NO_x排放/kg	1 300	2.17
签派可靠度/%	98	12.94

念方案的主要设计参数如表 3 所示。

4.3　竞争力优化

通过优化民机竞争力的方式对顶层需求进行权衡。优化的目标函数为竞争力 A 最大。竞争力 A 综合考虑了各项顶层需求的取值和权重, 其结果通过式 (3) 由标准化后的顶层需求指标加权后计算得到。

图 6　飞机初步概念方案的几何外形

Fig. 6　Geometric shape of preliminary conceptual plan of aircraft

表 3　概念方案主要设计参数

Table 3　Main design parameters of aircraft concept

部件	设计参数	取值
机翼	参考面积/m^2	355
	展弦比	10.5
	1/4弦线后掠角/(°)	32
机身	机身长/m	60
	机身宽/m	5.9
	机身高/m	6.1
发动机	最大静推力/kN	334
	涵道比	11
	总压比	52

优化问题的设计变量及其取值范围如表4所示。其中，客舱高度及经济舱的座椅宽度、排距既是航空公司和乘客所关注的顶层需求，也是影响飞机机身尺寸的设计参数，因此，将其作为设计变量，研究其对飞机总体性能的影响。

优化的约束条件为：① 适航条款规定的起飞第2阶段单发失效爬升梯度大于或等于2.4%；② 飞机油箱容积大于需用燃油的体积；③ 飞机翼展小于4E级飞行区限制的65 m。

采用遗传算法寻找优化问题的最优解。优化过程中，竞争力随优化代数的变化如图7所示。经过133代优化，竞争力变化趋于收敛。

表4 设计参数取值范围

Table 4 Range of para mefers

范围	单发海平面最大静推力/kN	机翼参考面积/m^2	机翼展弦比	机翼1/4弦线后掠角/(°)	机翼翼根相对厚度	经济舱座椅宽度/mm	经济舱座椅排距/mm	客舱高度/m
下界	230	300	8.5	25	0.13	475	775	1.8
初始值	334	355	10.5	32	0.14	480	800	2
上界	400	430	11	36	0.15	525	850	2.3

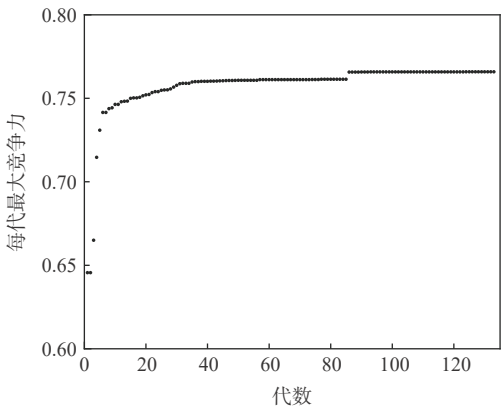


图7 竞争力随优化代数的变化

Fig. 7 Change with generation

4.4 顶层需求指标的确定

从竞争力优化结果中选择竞争力最优的解，对应的顶层需求指标如表5所示。从表5中可以看出，在初始顶层需求的基础上，更为严格的座公里油耗需求和座公里成本需求可以为飞机带来更高的竞争力。

飞机的经济舱座椅宽度可以减小到475 mm，因为更窄的座椅虽然会降低乘客的舒适度体验，但有利于减小客舱宽度，从而减小机身外形尺寸，带来结构质量和气动上的收益。飞机的座椅排距和客舱高度可以适当增加，因为座椅排距和客舱高度

表5 顶层需求优化前后对比

Table 5 Comparison of top-level aircraft requirements

座公里油耗/(kg·km^{-1}·座$^{-1}$)		变化/%	座公里成本/(元·km^{-1}·座$^{-1}$)		变化/%
初步估值	优化结果		初步估值	优化结果	
0.022	0.020	−9.1	0.38	0.369	−2.9
经济舱座椅宽度/mm		变化/%	经济舱座椅排距/mm		变化/%
初步估值	优化结果		初步估值	优化结果	
480	475.0	−1.0	800	822.2	2.8
客舱高度/m		变化/%	设计航程/km		变化/%
初步估值	优化结果		初步估值	优化结果	
2	2.19	9.5	12 000	11 956	−0.4
展长/m		变化/%	起飞场长/m		变化/%
初步估值	优化结果		初步估值	优化结果	
61	61.1	0.2	2 700	2 397	−11.2
着陆场长/m		变化/%	单发失效二阶段爬升梯度/%		变化/%
初步估值	优化结果		初步估值	优化结果	
1 800	1 725	−4.2	3.5	3.6	2.9
进近速度/(m·s^{-1})		变化/%	起降总噪声/dB		变化/%
初步估值	优化结果		初步估值	优化结果	
70	71	1.4	278	276.8	−0.4
航线NO_x排放		变化/%	签派可靠度/%		变化/%
初步估值	优化结果		初步估值	优化结果	
1 300	1 251	−3.8	98	98.3	0.3

对机身尺寸的影响相较于座椅宽度相对较小, 在初始顶层需求基础上, 增加座椅排距和客舱高度在提高乘客舒适度上的收益高于造成的飞机适应性的损失。

飞机的起飞场长和着陆场长需求在优化前后有较大的变化。初步拟定的起降场长需求过于宽松, 限制了飞机能够投入运营的机场。根据优化后的起降需求指标, 飞机在海平面标准大气条件下的起飞场长应该小于或等于 2 397 m, 着陆场长应该小于或等于 1 725 m, 能够适应更复杂的机场运营环境。

在优化后的顶层需求中, 飞机起飞单发失效 2 阶段爬升梯度需求指标提升到 3.9%, 该项需求指标严于适航条例所规定的 2.4%。更严格的爬升梯度需求保证了飞机具有更高的安全性。

优化后航线 NO_x 排放量需求指标相较于初始排放需求减小了 3.8%, 飞机的航程需求指标相较于初始航程需求降低了 0.4%。在单位航程内, 优化后的 NO_x 排放量需求比初始 NO_x 排放量需求更加严格, 飞机的环保性更高。

表 6 为设计变量优化前后对比, 变量取值为竞争力最优的方案。飞机的发动机推力和机翼展弦比增加。更大的发动机推力带来更优秀的起飞性能和爬升性能, 提高飞机的适应性和安全性。更高的展弦比有利于减小巡航时的诱导阻力, 降低油耗, 提高飞机的经济性。

表 6　设计变量优化前后对比

Table 6　Comparison of variables

单发海平面最大静推力/kN		变化/%	机翼参考面积/m^2		变化/%	机翼展弦比		变化/%	机翼1/4弦线后掠角/(°)		变化/%
初始值	优化值		初始值	优化值		初始值	优化值		初始值	优化值	
334	341.5	2.2	355	339.1	−4.5	10.5	11.0	4.8	32	30.4	−5.0
机翼翼根相对厚度		变化/%	经济舱座椅宽度/mm		变化/%	经济舱座椅排距/mm		变化/%	客舱高度/m		变化/%
初始值	优化值		初始值	优化值		初始值	优化值		初始值	优化值	
0.14	0.131	−6.4	480	475.0	−1.0	800	822.2	2.8	2	2.19	9.5

飞机的机翼面积和机翼后掠角减小。由于机翼的相对厚度减小, 控制激波阻力所需要的机翼后掠角有所减小。更小的机翼面积和机翼后掠角有利于减轻飞机的结构质量。

5　结　论

1) 以一款宽体客机为例, 对客机的顶层需求进行权衡, 获得了更合理、更准确的顶层需求指标取值, 验证了基于综合评价优化的权衡方法的有效性。在权衡之后的顶层需求中, 飞机的经济性明显改善。

2) 在舒适性方面, 牺牲了座椅宽度, 但增加了座椅排距和客舱高度。根据新的需求, 飞机具备更小的噪声和更低的排放, 环保性有所提升。在适应性上, 飞机的航程略有减小, 但起降场长明显降低, 起飞单发失效 2 阶段爬升梯度有所提升, 机场适应性获得改善。更高的起飞单发失效 2 阶段爬升梯度也意味着飞机安全性的提高。同时飞机的签派可靠度略有提升, 具备更高的可靠性。

参考文献（References）

[1] WALDEN D D, ROEDLER G J, FORSBERG K, et al. Systems engineering handbook: A guide for system life cycle processes and activities[M]. 4th ed. San Diego: Wiley, 2015: 47-56.

[2] ERES M H, BERTONI M, KOSSMANN M, et al. Mapping customer needs to engineering characteristics: An aerospace perspective for conceptual design[J]. Journal of Engineering Design, 2014, 25(1-3): 64-87.

[3] 黄俊, 武哲, 孙惠中, 等. 作战飞机总体设计评价准则和评估方法研究[J]. 航空学报, 2000, 21(1): 70-73.
HUANG J, WU Z, SUN H Z, et al. Study on evaluation criteria and assessment methodology for conceptual/preliminary design of combat aircraft[J]. Acta Aeronautica et Astronautica Sinica, 2000, 21(1): 70-73(in Chinese).

[4] PARKER J. Mission requirements and aircraft sizing[C]//Proceedings of the Aircraft Systems, Design and Technology Meeting. Reston: AIAA, 1986.

[5] PETEILH N, KLEIN T, DRUOT T Y, et al. Challenging top level aircraft requirements based on operations analysis and data-driven models, application to takeoff performance design requirements[C]//Proceedings of the AIAA Aviation 2020 Forum. Reston: AIAA, 2020.

[6] TORENBEEK E. Advanced aircraft design: Conceptual design, analysis and optimization of subsonic civil airplanes[M]. Chichester: John Wiley and Sons, Ltd. , 2013: 31-33.

[7] RAYMER D P. Aircraft design: A conceptual approach[M]. 6th ed. Washington, D.C.: AIAA, Inc. , 2018: 9-15.

[8] SADRAEY M H. Aircraft design: A systems engineering approach[M]. Chichester: John Wiley and Sons, Ltd., 2013: 12-15.

[9] DIRKS G, MELLER F. Multidisciplinary design optimization - Enhanced methodology for aircraft and technology evaluation[C]//Proceedings of the 8th Symposium on Multidisciplinary Analysis and Optimization. Reston: AIAA, 2000.

[10] 索欣诗, 李晓勇, 余雄庆, 等. 客机总体方案的综合评价方法[J]. 南京航空航天大学学报, 2018, 50(4): 558-564.
SUO X S, LI X Y, YU X Q, et al. Comprehensive evaluation meth-

od for conceptual design of commercial aircraft[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2018, 50(4): 558-564(in Chinese).

[11] 宋笔锋, 张彬乾, 韩忠华. 大型客机总体设计准则与概念创新[J]. 航空学报, 2008, 29(3): 583-595.
SONG B F, ZHANG B Q, HAN Z H. The study of concept design criteria for large-scale passenger aircraft with new technologies[J]. Acta Aeronautica et Astronautica Sinica, 2008, 29(3): 583-595(in Chinese).

[12] 陈迎春, 张美红, 张淼, 等. 大型客机气动设计综述[J]. 航空学报, 2019, 40(1): 522759.
CHEN Y C, ZHANG M H, ZHANG M, et al. Review of large civil aircraft aerodynamic design[J]. Acta Aeronautica et Astronautica Sinica, 2019, 40(1): 522759(in Chinese).

[13] VINK P, BRAUER K. Aircraft interior comfort and design[M]. Boca Raton: CRC Press, 2011.

[14] RICHARDS L G, JACOBSON I D. Ride quality assessment III: Questionnaire results of a second flight programme[J]. Ergonomics, 1977, 20(5): 499-519.

[15] CHAI X A, YU X Q, WANG Y. Tradeoff study between cost and environmental impact of aircraft using simultaneous optimization of airframe and engine cycle[J]. International Journal of Aerospace Engineering, 2017, 2017: 1-10.

[16] CLARK P. Buying the big jets: Fleet planning for airlines[M]. Hampshire: Ashgate Publishing Limited, 2007: 29-41.

[17] 郭婧宜, 孙宇锋, 吴寒雪, 等. 基于STT与模糊QFD的装备质量特性分解方法[J]. 北京航空航天大学学报, 2012, 38(8): 1090-1095.
GUO J Y, SUN Y F, WU H X, et al. Decomposition method of equipment quality characteristic based on STT and fuzzy QFD[J]. Journal of Beijing University of Aeronautics and Astronautics, 2012, 38(8): 1090-1095(in Chinese).

[18] 董彦非, 王礼沅, 王卓健, 等. 基于空战模式和AHP法的空战效能评估模型[J]. 系统工程与电子技术, 2006, 28(6): 885-888.
DONG Y F, WANG L Y, WANG Z J, et al. Air combat effectiveness assessment model based on operational pattern and analytic hierarchy process[J]. Systems Engineering and Electronics, 2006, 28(6): 885-888(in Chinese).

[19] LAMBE A B, MARTINS J R R A. Extensions to the design structure matrix for the description of multidisciplinary design, analysis, and optimization processes[J]. Structural and Multidisciplinary Optimization, 2012, 46(2): 273-284.

[20] CHAI X, YU X Q, WANG Y. Multipoint optimization on fuel efficiency in conceptual design of wide-body aircraft[J]. Chinese Journal of Aeronautics, 2018, 31(1): 99-106.

[21] FIELDING J. Introduction to aircraft design[M]. Cambridge: Cambridge University Press, 1999: 255-258.

Trade-off for top-level requirements of commercial aircraft using comprehensive evaluation and optimization

FAN Zhouwei[1, 2], YU Xiongqing[1, *], DAI Yalin[1]

(1. Key Laboratory of Fundamental Science for National Defense-Advanced Design Technology of Flight Vehicle, College of Aerospace Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;
2. Shanghai Aircraft Design and Research Institute, Commercial Aircraft Corporation of China, Ltd., Shanghai 201210, China)

Abstract: As the key process of aircraft design, top-level requirements should be carefully considered at the beginning of design activities. A quantitative trade-off method for top-level requirements of commercial aircraft has been proposed. According to the relationship between top-level requirements and aircraft conceptual design, characteristics of aircraft concept can be regarded as a set of feasible top-level requirements. The trade-off of top-level requirements is transformed to the optimization of design parameters. With the consideration of five criteria, including economy, comfort, environmental protection, flexibility, and safety/reliability, a comprehensive evaluation model for top-level requirements of commercial aircraft is established. The optimal set of top-level requirements can be found by the optimization for the best result of the comprehensive evaluation model. In the case of wide-body commercial aircraft, more reasonable top-level requirements are obtained, which reveals the feasibility of the comprehensive evaluation and optimization method.

Keywords: commercial aircraft; top-level requirement; comprehensive evaluation; optimization; trade-off analysis

Received: 2021-11-05; **Accepted**: 2022-01-05; **Published Online**: 2022-01-25 18:15

URL: kns.cnki.net/kcms/detail/11.2625.V.20220121.2017.002.html

Foundation items: National Natural Science Foundation of China (12032011); Special Scientific Research Project for Civil Aircraft of the Ministry of Industry and Information

* **Corresponding author**. E-mail: yxq@nuaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0677

面向制造成本的变刚度复合材料结构优化设计

陈世泽*，李道春，向锦武
(北京航空航天大学 航空科学与工程学院，北京 100191)

摘　　要：传统的复合材料结构设计方法未能有效利用复合材料的面内设计空间，变刚度结构的实用化为复合材料结构效能的提升提供了新的机会。使用制造工艺成本模型，以制造成本为目标函数，建立结构性能约束和制造约束，通过路径函数法和平移法对一型大展弦比机翼进行变刚度设计，建立了以制造成本为目标的变刚度复合材料结构设计优化方法，实现制造成本减少23.87%，结构减重35.58%，并分析了不同纤维路径对变刚度结构构型设计、性能、质量和制造过程的影响。

关 键 词：复合材料；变刚度；制造成本；结构优化；遗传算法
中图分类号：V221+.3
文献标志码：A　　**文章编号**：1001-5965(2023)09-2423-09

随着复合材料结构的设计、分析和制造技术的不断发展，在铺贴面内纤维路径进行设计的变刚度复合材料(tow-steered composite, TSC)结构具备了实现的可能。变刚度复合材料通过释放复合材料面内设计空间，能够按照载荷分布设置纤维方向，使铺丝路径与关键受力路径一致，以改善应力分布，实现更优异的屈曲性能、静强度、振动特性及更轻的结构质量，从而获得更高的结构效率。

目前对变刚度复合材料结构的研究主要关注2个方面：结构减重和性能提升效应。减重是飞行器设计工作的核心目标，变刚度结构的概念一经提出，立刻被研究者应用于减重方面的设计和研究。Weisshaar[1]研究了复合材料前掠翼的铺丝方向角对机翼升力特性和发散速度的影响，发现100°～140°的区间对前掠翼较为理想。Stanford和Jutte[2]通过将美国国家航空航天局(National Aeronautics and Space Administration, NASA)通用研究模型(common research model, CRM)蒙皮的主铺丝方向角从传统固定值中释放，发现在上蒙皮旋转10.32°，下蒙皮旋转3.55°时，可以获得减重5%的最优结构。在主方向角被从固定值释放以后，研究者开始关注纤维路径设计对结构效率的作用和影响。Wang等[3]对纤维角线性变化的复合材料翼盒进行变刚度优化设计，发现变刚度结构可以在同等性能下减重13.2%。Alhajahmad和Mittelstedt[4]对复合材料机身壁板在屈曲性能约束下的变刚度设计，发现使用圆弧曲线设计铺丝路径可以实现减重18%，Brooks等[5]使用CRM以最小质量为目标对具有相同结构的金属、直线纤维复材机翼和变刚度复材机翼进行优化和对比分析，进一步验证了变刚度结构在减重方面相较于金属和直线纤维结构的优势。

另一方面，现有研究从不同的角度和性能指标出发，研究了变刚度结构的性能提升潜力。Pereira等[6]以层合板为对象，研究发现变刚度设计可以将自振频率提升15%。Butler等[7]在屈曲优化的基础上，结合实验发现相较于直纤维板，变刚度层合板抗冲击性能可提升9%。Zhang等[8]对复合材料机翼蒙皮使用遗传算法和矩阵摄动方法寻找最优铺丝路径，实现机翼颤振速度提高13.49%。Stodieck等[9]使用CRM，逐一分析了在静强度、突风响应、颤振速度、控制效率约束下对机翼蒙皮变刚度优化设计取得的结果，得出在综合性能约束下变刚度设计可

收稿日期：2021-11-10；录用日期：2022-01-02；网络出版时间：2022-01-29 07:02
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220128.1217.001.html
*通信作者. E-mail：chensz@cgwic.com

引用格式：陈世泽，李道春，向锦武. 面向制造成本的变刚度复合材料结构优化设计[J]. 北京航空航天大学学报，2023，49(9)：2423-2431.
CHEN S Z，LI D C，XIANG J W. Design optimization of tow-steered composite structure targeting on manufacturing cost[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2423-2431 (in Chinese).

以最大减重10.5%。而如果以性能最优为目标[10]，变刚度设计的颤振速度可以提升7%、突风载荷减小24%。除基于自动铺丝(automatic fiber placement, AFP)技术实现变刚度以外，李飞和聂国隽[11]对基于连续丝束剪切(continuous tow shearing, CTS)技术的变刚度层合板进行分析，发现合理设置铺丝路径还可以提高层合板的临界屈曲温度。

总体来看，变刚度结构因其对复合材料设计空间的释放而带来的结构效能提升潜力已得到验证，但与传统的直线纤维结构一样，这些研究仍主要聚焦于减重和性能提升2个方面，而未对变刚度设计对制造过程的影响和制造成本展开深入考察。

值得注意的是，变刚度结构设计会使本就长期存在的性能提升/成本增加的“性价比”问题更加突出。在变刚度结构从理论研究转向量产飞行器结构应用以前，有2个问题尚待研究者解答：

1) 纤维的曲线铺贴直接导致结构复杂度提升，这将对制造过程产生什么影响。这种影响与性能提升或减重的收益是否匹配。

2) 在满足设计要求且性能指标不变的前提下，可否通过变刚度设计，实现结构制造成本的降低，从而突破复合材料发展的一个关键瓶颈：使复合材料结构在具备性能优势的同时也具备成本优势。

一些组织和研究者已开始尝试回答上述问题。NASA Langley研究中心[12]对用于机身筒段或火箭箭体的整体编织复合材料筒状结构进行变刚度设计和试制，并在过程中分析变刚度的影响。在其加工过程中发现，要保证制件的质量，必须降低纤维的铺贴速度，并增加用于每层压实的时间。Xiao等[13]设计了一种手工取代AFP实现曲线铺贴的制造方法，希望能够降低设备投入并缩短加工时间，但并未对手工工艺取代实际带来的成本影响进行计算。孔斌等[14]研究认为变刚度复合材料结构需要对丝束路径进行更加精确的控制，会导致铺丝速度降低而提高制造成本。Hao等[15]在对变刚度结构的设计和研究中强调了变刚度结构因其复杂性而对AFP工艺过程提出更高的要求，须施加更严格的约束。Sohouli等[16]对复合材料风力发电机叶片采取了变刚度设计，并对不同设计方案的成本/效能比进行测算，但其制造成本仅包括原材料价格，并未考虑制造过程。

对于处在应用探索阶段的变刚度结构，制造成本的测算和控制有关的研究非常有限，本文提出的2个对变刚度结构实用化和复合材料的进一步发展具有关键影响的问题也仅做出了有限而初步的回答。作为以制造成本为目标的复合材料结构优化设计研究的一部分，文献[17]已经建立了基于工序流程仿真的复合材料制造工艺成本模型，并进一步建立了直线纤维结构以制造成本为目标的优化方法，完成了直线纤维机翼的制造成本优化。本文希望在此基础上，进一步建立以制造成本为目标的变刚度结构优化设计方法，分析变刚度对复合材料制造成本的作用与影响，从而建立完整的复合材料结构制造成本优化设计方法，为变刚度结构的实用化及复合材料结构的设计和研究提供有益参考。

1 基本思路和方法

为了避免初始方案存在优化目标即制造成本值上的冗余，本文选择以完成制造成本优化的直线纤维复合材料机翼作为初始方案，以保证优化结果完全来自变刚度设计。同时，变刚度结构对制造成本产生直接影响的参数除外形尺寸、材料选取、铺层厚度、铺层方案设计以外，增加了纤维脱离固定值约束的主方向角和沿此方向角的铺丝路径。为了保证结构性能符合要求，优化过程将静强度和气动弹性指标设置为约束条件。在此情况下，为合理控制计算量，避免在完全释放的设计空间内，在一步计算中考虑全部设计变量和约束条件可能导致的困难和问题，本文采取分步优化策略，将设计变量、约束条件、目标函数合理分离至各优化步，具体步骤为：①完成直线纤维结构设计优化，获得优化方案和有限元模型；②选取工艺路线，确定目标函数，即制造成本估值函数；③根据静强度计算、模态分析和气动弹性分析结果，建立性能约束条件；④根据变刚度结构加工过程，明确各项制造约束；⑤选取多种铺丝路径定义方法，明确不同纤维路径的设计变量；⑥执行分步优化，获得不同铺丝路径下满足静强度、气动弹性和工艺要求的设计方案；⑦比较分析不同类型铺丝路径的结果，获得具有最佳制造成本的变刚度结构。

变刚度复合材料优化设计流程如图1所示。其中，分步优化方法分为2步，如图2所示。具体流程为：①静强度约束下的优化设计，目标函数为最小制造成本，设计变量为铺层厚度、纤维主方向角，约束条件主要包括位移、扭转等静强度约束和制造约束；②气动弹性约束下的优化设计，在第1步优化结果上，以纤维路径和铺层厚度为设计变量，气动弹性指标和制造约束为约束条件，继续以制造成本为目标函数，进行二级优化。

根据变刚度复合材料结构的制造要求，自动化铺贴技术是实现变刚度的基本前提，因此，无需考察手工铺贴方式的制造成本。同时变刚度设计改变层内铺丝方式，对热压罐及非热压罐固化、胶

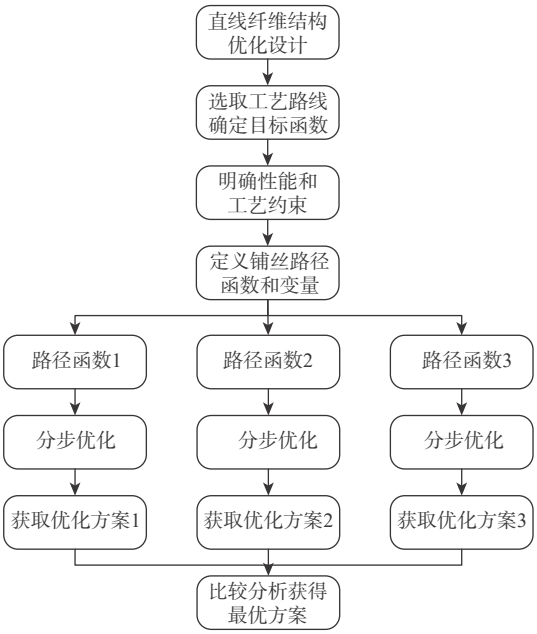


图 1　变刚度复合材料优化设计流程

Fig. 1　Flow chart of variable stiffness composite structure design optimization

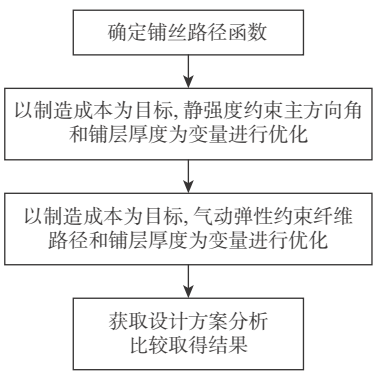


图 2　分步优化方法

Fig. 2　Two step optimization method

接、装配等工艺的成本影响仅通过零件整体外形如厚度等参数的改变而体现, 在原材料单价和自动铺贴工艺固定情况下, 后续工艺不同路线的成本差异与变刚度设计无直接相关性。因此, 仅选取一条自动化铺贴的工艺路线, 建立变刚度结构的优化设计方法, 不再开展工艺路线间的横向比较分析。如需考察变刚度结构在不同工艺路线上的制造成本, 只需置换目标函数即可。

2　机翼有限元模型

本文采用完成制造成本优化的 T800 碳纤维直线纤维 NACA0018 型机翼, 基本参数如表 1 所示。

机翼有限元模型如图 3 所示。使用 MSC.Patran/Nastran 软件进行静强度和气动弹性分析, 静强度结果如图 4 和图 5 所示。颤振速度通过 Nastran 颤振分析模块, p-k 法求解, 颤振计算结果如图 6 所示。

表 1　机翼基本参数

Table 1　Basic parameters of wing

参数	初始值
翼展/m	20
弦长/m	1.75
翼肋数量	19
蒙皮厚度/mm	3.3
蒙皮层数	22
翼肋厚度/mm	1.8
翼肋层数	12
梁边长/mm	101
翼尖最大位移/m	0.323
翼尖最大扭角/(°)	2.44
颤振速度/(m·s^{-1})	206
结构质量/kg	535.08

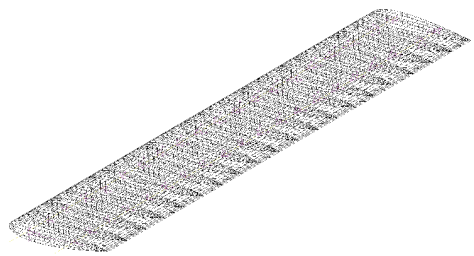


图 3　机翼有限元模型

Fig. 3　Finite element model of wing

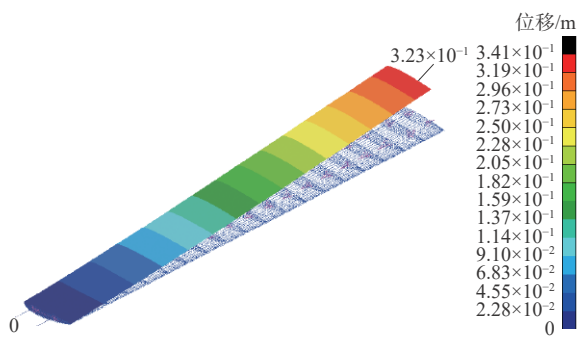


图 4　翼尖最大位移

Fig. 4　Maximum displacement at wing tip

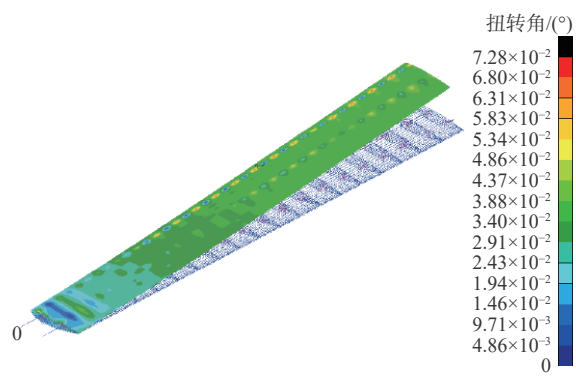


图 5　最大扭转角

Fig. 5　Maximum torsion angle

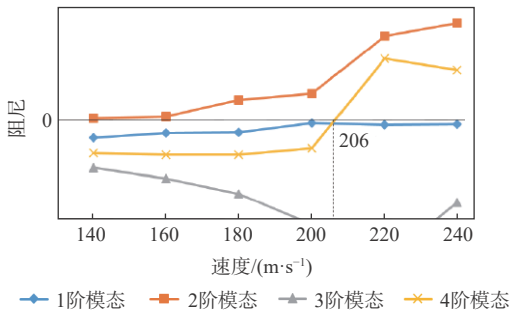


图 6 颤振计算结果

Fig. 6 Flutter calculation results

3 结构优化设计

3.1 目标函数

选定初始方案的过程中实际上已经确定了目标函数的选取，该机翼的工艺路线为：自动铺贴干纤维，树脂转移注塑(resin ttanfer molding, RTM)，胶接。根据文献 [17] 已建立的制造成本评估模型，制造成本以“等效工时(effective working hour, EWH)”的形式表达，EWH 由根据特征尺寸计算得到的加工时间经结构复杂性修正和不同工艺间的工时成本差异修正后得到，其优点在于可以通过时间单位而非货币价值来表征成本，能够简化计算并提高评估准确性。以制造成本为目标的变刚度结构优化设计的目标函数计算式为

$$t_3 = t_{\mathrm{dry,ATL}} + t_{\mathrm{RTM}} + t_{\mathrm{glue}} \tag{1}$$

式中：$t_{\mathrm{dry,ATL}}$ 为干纤维自动铺贴工艺的制造成本；t_{RTM} 为 RTM 工艺的制造成本；t_{glue} 为胶接工艺的制造成本。其计算式如式(2)～式(8)所示。

$$t_{\mathrm{dry,ATL}} = C_{\mathrm{ATL,dry}} + C_{\mathrm{s}} C_{\mathrm{c}} t_{\mathrm{ply}} N_{\mathrm{ply}} \tag{2}$$

$$C_{\mathrm{ATL,dry}} = \left(P_{\mathrm{dry,ATL}} / H_{\mathrm{manual}}\right) \rho_{\mathrm{f}} h_{\mathrm{l}} S V_{\mathrm{f}} (1 + 0.1) \tag{3}$$

$$t_{\mathrm{ply}} = \frac{S}{w v_0 K_{\mathrm{a}} K_{\mathrm{r}}} + \frac{L\sin\theta + W\cos\theta}{w} \cdot \left(\frac{v_0 K_{\mathrm{a}} K_{\mathrm{r}}}{a} + t_{\mathrm{delay}}\right) \tag{4}$$

式中：$C_{\mathrm{ATL,dry}}$ 为干纤维自动铺贴工艺的原材料成本；C_{s} 为结构复杂性系数；C_{c} 为工时成本因子；t_{ply} 为单层纤维铺贴时间；N_{ply} 为铺层总数；w 为纤维带宽度；L 和 W 分别为铺贴面长和宽；θ 为纤维方向角；a 为铺带头加速度；t_{delay} 为工序中除铺贴外其他动作时间，包括转向、切断、进给/脱离等，由所使用的设备性能决定；K_{a} 由纤维弯曲的最尖锐角度决定；K_{r} 由曲率半径决定；$P_{\mathrm{dry,ATL}}$ 为干纤维价格；H_{manual} 为手工工艺的综合小时费率；ρ_{f} 为干纤维密度；h_{l} 为零件厚度；V_{f} 为零件干纤维体积比。

$$t_{\mathrm{RTM}} = C_{\mathrm{RTM}} + 2.5 C_{\mathrm{s}} \frac{S h_{\mathrm{v}}}{0.024\,6} + 1.092\,5 C_{\mathrm{s}} \tag{5}$$

$$C_{\mathrm{RTM}} = (P_{\mathrm{resin}} / H_{\mathrm{manual}}) \rho_{\mathrm{r}} h_{\mathrm{v}} S (1 - V_{\mathrm{f}})(1 + r_{\mathrm{waste}}) \tag{6}$$

式中：C_{RTM} 为 RTM 工艺的原材料成本；P_{resin} 为树脂价格；ρ_{r} 为树脂密度；h_{v} 为注射垂直距离；r_{waste} 为废料率。

$$t_{\mathrm{glue}} = C_{\mathrm{s}} \frac{S_{\mathrm{glu}}}{2.04} + 1.16 + C_{\mathrm{adhesive}} \tag{7}$$

$$C_{\mathrm{adhesive}} = (P_{\mathrm{adh}} / H_{\mathrm{manual}}) S_{\mathrm{glu}} \rho_{\mathrm{a}} h_{\mathrm{a}} (1 + r_{\mathrm{scrap}}) \tag{8}$$

式中：C_{adhesive} 为胶接材料成本；S_{glu} 为胶接面积；P_{adh} 为胶接材料价格；ρ_{a} 为胶接材料密度；h_{a} 为涂胶厚度；r_{scrap} 为废料率。

初始方案以“等效工时”表达的制造成本为 1 730.75 EWH。

3.2 优化变量

确定目标函数以后，选取变刚度结构优化设计的变量。初始机翼已经完成了对蒙皮厚度、肋厚度、梁截面边长、蒙皮(0°，±45°，90°)各角度铺层数、肋(0°，±45°，90°)各角度铺层数、蒙皮铺层顺序、肋铺层顺序的优化。在此基础上的变刚度设计，需要引入的设计变量有 2 个。

3.2.1 主方向角

在传统的直线纤维结构设计和优化过程中，由于纤维的铺放角被限制在0°，±45°，90°上，优化过程一般将各角度的铺层比例和铺层顺序设置为优化变量。实际上，研究表明更丰富的铺层角度组成可以形成更具效率的复合材料结构。自由主方向角将释放更多设计空间，使纤维可以更精确地沿着期望的传力方向铺设，这本身也是复合材料结构优化设计的主要任务。特别是针对不同展弦比、不同载荷工况、不同内部构造的复合材料零部件，固定的铺丝角度很可能无法充分利用各向异性和刚度方向性，从而将优化方案限制在“次优解”上。主方向角的优化结果也将作为铺丝路径优化的基础，优化方案的铺丝路径函数，实际就是基于主方向角对局部刚度分布进行细化调整。

初始方案采用0°，±45°，90°纤维角，铺层顺序已经过优化，因此，将主方向角变量 θ 的变化范围设定为

$$\begin{cases} \theta_0 - 45^\circ \leqslant \theta \leqslant \theta_0 + 45^\circ \\ \theta_0 = 0^\circ, 45^\circ, -45^\circ, 90^\circ \end{cases} \tag{9}$$

式中：θ_0 为初始方案中该层的铺丝角，可使铺丝主方向角的变化范围覆盖整个层内设计空间。

3.2.2 纤维路径

纤维曲线铺放的路径建立方法包括铺层参数法和路径函数法。本文建立的以制造成本为目标函数和优化设计方法均使用尺寸和角度参数作为设计变量，与路径函数法中的设计变量一致。同时，路径函数法有利于保证纤维连续性，可以直观

地施加制造约束条件, 能够获得最优连续纤维路径的唯一解。

本文使用路径函数法建立 3 种铺丝路径。

1) 线性函数。当纤维按照线性函数变化时, 函数表达式为

$$\theta(x_1)=\begin{cases}\dfrac{2}{d}(T_0-T_1)x_1+T_1 & 0\leqslant x_1<\dfrac{2}{d}\\ \dfrac{2}{d}(T_1-T_0)x_1+T_1-2(T_1-T_0) & \dfrac{2}{d}\leqslant x_1\leqslant d\end{cases} \tag{10}$$

2) 抛物线函数。当纤维路径按照抛物线变化时, 函数表达式为

$$\theta(x_1)=\frac{4(T_1-T_0)}{d^2}\left(x_1-\frac{d}{2}\right)^2+T_0 \qquad 0\leqslant x_1\leqslant d \tag{11}$$

3) 正弦函数。纤维按照正弦函数变化的表达式为

$$\theta(x_1)=(T_0-T_1)\sin\left(\frac{\pi}{d}x_1\right)+T_1 \qquad 0\leqslant x_1\leqslant d \tag{12}$$

式中: x_1 为沿纤维方向的位置坐标; T_0 和 T_1 分别为初始和终止位置的方向角; d 为特征长度。

建立由以上函数定义的纤维参考路径后, 通过平移法使纤维铺满整个表面。由此可定义变刚度结构优化设计的设计变量, 包括: 纤维主方向角 θ; 路径函数参数 T_0、T_1; 蒙皮厚度, 按照单层厚度 0.15 mm 圆整; 肋厚度, 按照单层厚度 0.15 mm 圆整; 梁截面边长。

3.3　约束条件

3.3.1　性能约束

以制造成本为目标进行优化设计的基本思路是在许用性能范围内寻找成本最低的结构。因此, 对变刚度结构的优化设计仍将性能指标设置为约束条件, 继续沿用初始方案所经历的优化过程中的性能约束值, 使本文建立的变刚度设计方法可以在保证结构性能不降低的前提下通过变刚度降低制造成本。

性能约束条件为: ①2.5g 过载纵向拉起, 马赫数为 0.5, 海平面高度, 翼尖最大位移不超过半翼展的 0.341 m; ②2.5g 过载纵向拉起, 马赫数为 0.5, 海平面高度, 翼尖扭转角不超过 4.17°; ③颤振速度大于或等于 205 m/s。

3.3.2　制造约束

变刚度结构的铺放路径在设计过程中确定, 因此, 必须在设计过程中根据当前常见的或计划在制造环节使用的材料要求和设备能力, 建立明确的制造约束, 才能保证设计结果可以加工实现。定义制造约束条件为: ①相邻单元同一铺层纤维需连续; ②均匀对称铺层; ③纤维最小转向半径为 250 mm。

目前可用的变刚度复合材料制造技术主要有 AFP 和 CTS 这 2 种, 由于 CTS 技术距离工业化应用和实际结构的加工尚有较大距离, 且铺贴速率仅能达到当前 AFP 技术的 10% 以下[18], 根据文献 [17] 建立的制造成本评估模型可以预见 CTS 的铺放速率将会大幅度增加制造成本, 制造成本最优结果不可能通过 CTS 技术实现, 因此, 本文选取 AFP 技术为铺贴方法, 现阶段不对 CTS 工艺进行成本计算, 待其发展成熟后, 可将其制造约束条件加入并将其成本估算函数模块纳入制造成本模型中。

约束条件①和②是复合材料结构的一般工艺约束, 通过设计变量控制。约束条件③基于当前主流 AFP 设备最高能力, 结合曲线铺贴对避免出现褶皱、间隙和重叠等质量问题的要求建立。以此建立的约束值表征并不直观, 不能直接施加在优化算法中, 因此, 需要将其转换为纤维路径函数曲率的限制。约束条件③可定义为

$$\frac{\left(1+\theta'^2\right)^{3/2}}{|\theta''|}\leqslant 0.25 \tag{13}$$

3.4　有限元模型修正

变刚度结构不同于直线纤维结构, 使用解析方法不能得到显式解, 只能使用 Ritz 法等半解析数值方法。此型机翼在 MSC.Patran/Nastran 软件中建立有限元模型并进行静强度和气动弹性计算, 在变刚度设计时, 采用纤维路径函数建立以后, 将有限元模型近似离散的方法。

使用路径函数法建立的纤维路径铺放角连续变化, 但在每个有限元单元内, 可以将铺设角近似为固定值, 从而将连续的路径函数离散。根据路径函数, 赋予每个单元中心点坐标值, 计算得到该点的纤维角度并使之在单元内固定, 建立近似曲线路径, 有限元模型离散[14] 如图 7 所示。

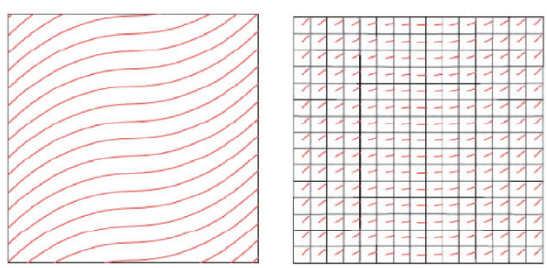


图 7　有限元模型离散[14]

Fig. 7　Discretization of finite model[14]

经离散近似后的模型可以按照直线纤维结构的方法, 调用 MSC.Patran/Nastran 对应模块进行静强度和气动弹性分析。

3.5 优化流程

采用分步策略，分 2 步对该机翼蒙皮进行变刚度优化。首先，选取一个路径函数，对有限元模型进行离散；然后，第 1 步以 3.2.2 节中的纤维主方向角 θ、蒙皮厚度、肋厚度、梁截面边长为设计变量，以 3.3.1 节中的①和②为性能约束，3.3.2 节中的①和②为制造约束，以目标函数最小进行自由主刚度角优化。第 2 步在第 1 步取得的结果的基础上，优化变量不变，性能约束改为③，并增加制造约束③，继续以目标函数最小进行优化，即在颤振速度约束下进行气动弹性剪裁，获得该种路径函数下制造成本最优的设计。优化流程如图 8 所示。

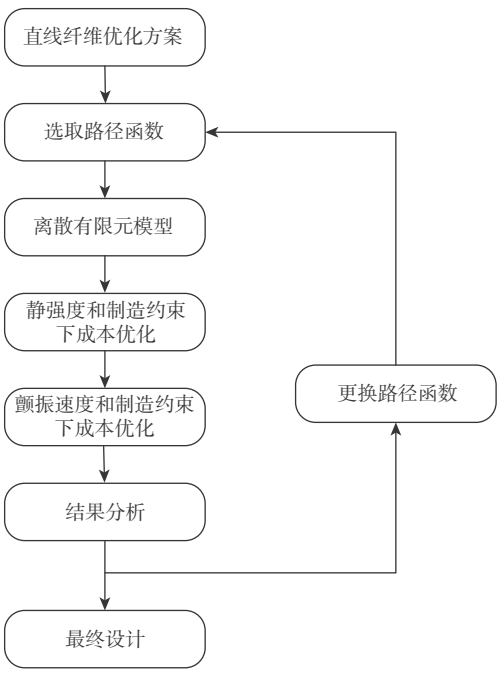


图 8 优化流程

Fig. 8 Flow chart of optimization

3.6 优化程序

在 Isight 平台上搭建多岛遗传算法程序，以 1°为变动分辨率，在机翼上生成主方向角、T_0、T_1 的不同组合，形成初始种群。选择个体，根据主方向角参数调用 MSC.Patran/Nastran 修改有限元模型，进行静强度计算，判断是否满足约束条件，满足条件的个体根据 T_0、T_1 的值形成路径函数，由程序计算路径函数是否满足制造约束，符合约束的个体进入第 2 步优化，自动调用 MSC.Patran/Nastran 将有限元模型离散，执行颤振速度计算，判断是否符合气动弹性约束，满足全部约束条件的个体计算目标函数值，然后进化和迁移，重复以上过程直至结果收敛。优化程序结构如图 9 所示。

本文多岛遗传算法子种群规模取为 15，共 10 个岛，进化 20 代，交叉概率为 1.0，变异概率为 0.01，岛间迁移率为 0.01，迁移间隔代数为 5，精英个体数量为 1。

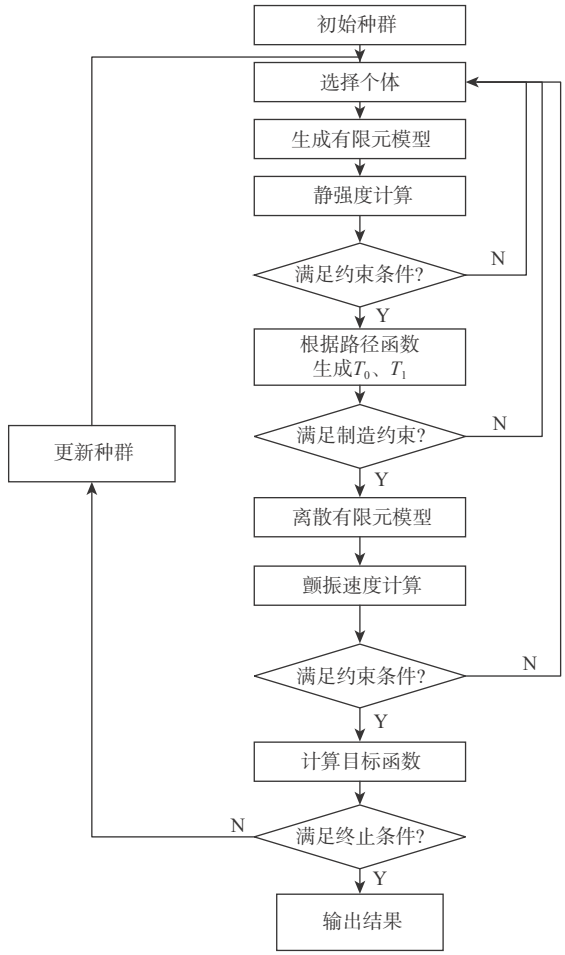


图 9 优化程序结构

Fig. 9 Structure of optimization program

4 优化结果

以制造成本为目标的变刚度机翼优化结果如表 2 所示。

4.1 线性函数

线性函数结果收敛过程如图 10 所示。线性路径铺丝的机翼制造成本由 1 730.75 EWH 降低至 1 326.17 EWH，降幅为 23.38%。蒙皮铺层数由 22 层降低至 16 层，厚度降至 2.4 mm，梁截面边长由 101 mm 降低至 83 mm，翼肋则没有发生厚度变化。

蒙皮铺层方案在自由主方向角的优化步中由初始方案的 $[0/0/0/0/-45/-45/-45/90/90/45/45]_s$ 改变为 $[0/0/0/-35/90/90/21/21]_s$，然后在第 2 步中得到最优路径 T_0=47°，T_1=36°。

性能指标上，翼尖最大扭角由 2.44°增加到 3.19°，最大位移由 0.323 m 增加至 0.327 m，两者均在性能

表 2　优化结果

Table 2　Optimization results

方案	制造成本/EWH	蒙皮厚度/mm	蒙皮层数	翼肋厚度/mm	翼肋层数	梁边长/mm	翼尖最大位移/m	翼尖最大扭角/(°)	颤振速度/(m·s^{-1})	结构质量/kg
线性函数	1 326.17	2.4	16	1.8	12	83	0.327	3.19	205	351.87
抛物线函数	1 331.92	1.5	10	1.8	12	84	0.325	3.09	205	346.52
正弦函数	1 317.66	1.8	12	1.8	12	83	0.324	3.17	205	344.68
初始方案	1 730.75	3.3	22	1.8	12	101	0.323	2.44	206	535.08

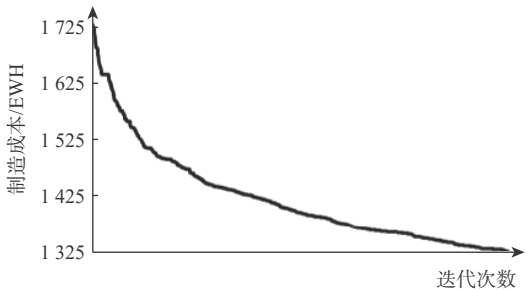


图 10　线性函数结果收敛过程

Fig. 10　Convergence process of linear function

约束的范围内略有增加, 颤振速度为 205 m/s, 满足约束条件。结构质量则由 535.08 kg 降低至 351.87 kg, 降幅为 34.24%。

4.2　抛物线函数

抛物线函数结果收敛过程如图 11 所示。在抛物线函数下, 机翼制造成本降低幅度略小, 降至 1 331.92 EWH, 降幅为 23.04%。蒙皮厚度降低最多, 降低至 10 层, 厚度为 1.5 mm, 梁截面边长则略大于线性函数结果, 由 101 mm 降低至 84 mm, 翼肋厚度同样没有发生变化。

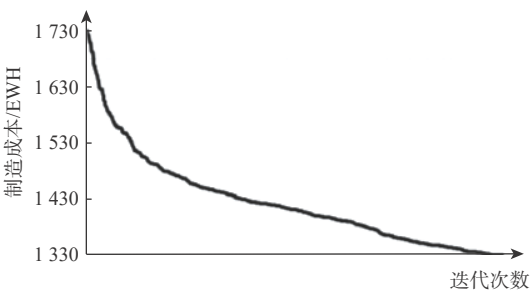


图 11　抛物线函数结果收敛过程

Fig. 11　Convergence process of parabola function

在此种路径函数的设计中, 蒙皮铺层方案改变为 $[0/0/-35/90/21]_s$, 在路径优化中得到最优方案 T_0=52°, T_1=28°。

性能指标上, 翼尖最大扭角和翼尖最大位移仍然均略有增加, 但增幅小于线性路径函数的设计, 最大扭角此时增加到 3.09°, 最大位移则为 0.325 m, 颤振速度仍为约束值 205 m/s。结构质量在此方案中进一步降低, 为 346.52 kg, 降幅达到 35.23%。

4.3　正弦函数

正弦函数结果收敛过程如图 12 所示。正弦函数实现了 3 种方案中的最低制造成本, 在此方案中成本降至 1 317.66 EWH, 降幅为 23.87%。蒙皮厚度则介于其他 2 种方案之间, 铺层共 12 层, 厚度为 1.8 mm, 梁截面边长与线性函数结果相同, 减至 83 mm, 翼肋厚度与其他方案一致, 未发生任何变化。

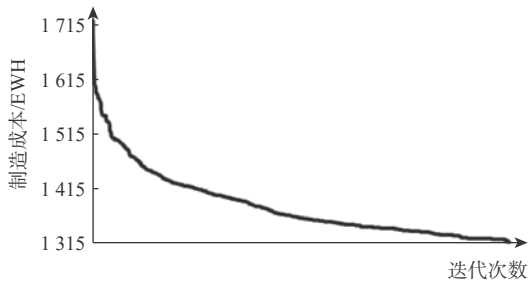


图 12　正弦函数结果收敛过程

Fig. 12　Convergence process of sine function

正弦函数路径设计下的蒙皮铺层方案改变为 $[0/0/-35/90/21/21]_s$, 并得到最优曲线路径控制参数 T_0=53°, T_1=26°。

性能指标方面, 翼尖最大扭角较初始方案仍有所增加, 增幅小于线性路径但大于抛物线路径, 最大扭角此时为 3.17°, 最大位移则为 0.324 m, 为 3 个方案中最低, 颤振速度保持在约束值 205 m/s。结构质量在此方案中达到最小值 344.68 kg, 降幅达到 35.58%。

4.4　结果分析

根据以上优化设计结果, 采用变刚度设计可以实现降低制造成本的目标。在文献 [17] 已经以相同目标函数和性能约束条件完成以制造成本为目标的剪裁优化的直线纤维机翼上, 进一步降低了制造成本, 最大降幅达到 23.87%。

同时, 由于在优化设计过程中, 梁的截面边长和蒙皮的厚度均有所下降, 对此型机翼以制造成本为目标的优化设计同时实现了结构减重, 最大减重达到 35.58%。这种效应是复合材料结构增材制造的属性导致的: 减少的铺层数直接降低了材料成本, 也降低了自动铺贴工序的加工时间, 同时因厚

度下降，对固化等其他工艺和所需辅料的成本也产生了影响。制造成本和结构质量在变刚度复合材料结构中仍具有较强的正相关性，因此，由基于工艺过程仿真的制造成本模型建立的目标函数，在变刚度结构中可以实现降本同时减重的效果。

制造成本的降低幅度小于结构减重幅度，出现这种现象的原因是，在变刚度结构最重要的曲线铺贴环节，纤维的曲线路径使AFP铺放速度低于直线纤维，由此增加了在一个铺贴面上AFP工艺所需的铺贴时间，铺层数减少所带来的成本效益被工艺时间的增加部分抵消。进一步考察此现象，线性函数方案材料费用为578.83 EWH，抛物线函数方案为569.53 EWH，正弦函数方案为568.02 EWH，材料费用占总费用比例分别由初始方案的50.86%降低至43.65%、42.76%和43.10%。而在工艺成本上，线性函数为747.34 EWH，占比由初始方案的49.14%增加至56.35%，抛物线函数为762.39 EWH，占比57.24%，正弦函数749.64 EWH，占比56.89%。可见，在3种优化方案中，成本整体下降的同时材料费用占比下降，而工艺成本占比上升，这一点验证了包括本文在内的现阶段研究者对于变刚度结构的一个重要顾虑：变刚度复合材料虽然能够提升结构效率，但也会同时带来制造复杂度的提升，因此，确有必要评估变刚度对制造过程的影响，设置制造约束以保证结构可实现也是变刚度结构设计工作的必要内容。

对比3种铺丝路径函数所取得的结果，对这种影响进一步分析。线性函数虽然在3种方案中铺层厚度最厚，所需材料最多、铺贴层数最多，但仍然取得了低于抛物线函数的制造成本。这是因为线性函数所产生的最大曲率较低、最小纤维转向半径较大所致，铺贴线性函数路径的纤维对AFP过程的速度影响较小。抛物线函数虽然将蒙皮层数降至最低，获得了最薄的蒙皮，材料费用及占比也低于线性函数，但由于这种路径会产生较尖锐的纤维转角，其工艺成本反而更高，导致总成本也最高，从制造成本的角度来看并没有获得最高的效率。正弦函数虽然具有3种方案里最大的T_0、T_1差，意味着具有最大的铺丝角变化范围，但由于其周期性曲线整体变动较平滑，因此，在材料费用与抛物线函数接近的条件下取得了最低的制造成本，实现了3种方案中最优的综合效益。

从性能指标的角度来看，3种方案的颤振速度都收敛于约束值上，而静强度指标均高于初始方案。以制造成本为目标的结构优化思路的出发点就是在满足许用性能要求的条件下实现最低的制造成本，变刚度结构实现了这一目标。在直线纤维结构中，由于铺丝角只能选择4个值，而结构的铺层顺序和层数也都是离散的变量，因此，优化过程中结构变化呈现跳跃型而非线性变化，在多个约束条件下经常产生整体满足条件但部分指标值比较高，比例高于约束值的结果，这是直线纤维结构不能完整利用复合材料设计空间的具体体现。

在此机翼的变刚度优化中，初始方案翼尖最大位移和最大扭角与性能约束之间的剩余空间因自由主方向角和曲线路径2个面内设计变量的加入而被利用起来。由此可见，变刚度设计能够使复合材料结构的优化过程更加线性，充分的利用直线纤维结构离散设计变量下剩余的设计空间来降低成本。

5 结 论

1）本文建立了以制造成本为目标的变刚度复合材料结构优化设计方法，在已完成制造成本优化的直线纤维机翼上，实现进一步降低制造成本23.87%。

2）制造成本变刚度优化设计能够在降低制造成本的同时，进一步降低结构质量，在此机翼上最大降幅35.58%。依靠目标函数的结构和建立的制造约束，本文方法将变刚度导致的结构复杂性问题纳入考虑，能够在提升结构效率、避免变刚度带来制造问题的同时找到成本最低的设计方案，实现了预期目标，变刚度设计对制造过程带来的影响也可以通过降本与减重数据及不同纤维路径间的对比分析得到。

3）变刚度设计能够充分利用设计空间，包括铺贴面内的设计空间和直线纤维设计在性能约束内因变量离散而剩余的设计空间。虽然有待于充分、全面的实验来验证纤维路径变化对复合材料损伤模式的具体影响，并进一步建立冲击阻抗和损伤容限等约束条件，但本文通过有限元模拟，建立了以制造成本为目标的变刚度复合材料结构优化设计方法，能够在同等力学和气动弹性性能要求下，获得制造成本最低的结构。本文方法的建立有助于复合材料优势的挖掘和竞争力的提升，与直线纤维结构设计方法相结合，可形成架构完整的复合材料制造成本优化设计方法。

参考文献（References）

[1] WEISSHAAR T A. Aeroelastic tailoring of forward swept composite wings[J]. Journal of Aircraft, 1981, 18(8): 669-676.

[2] STANFORD B K, JUTTE C V. Comparison of curvilinear stiffeners and tow steered composites for aeroelastic tailoring of aircraft wings[J]. Computers & Structures, 2017, 183: 48-60.

[3] WANG Z X, WAN Z Q, GROH R M J, et al. Aeroelastic and local

buckling optimisation of a variable-angle-tow composite wing-box structure[J]. Composite Structures, 2021, 258: 113201.

[4] ALHAJAHMAD A, MITTELSTEDT C. Minimum weight design of curvilinearly grid-stiffened variable-stiffness composite fuselage panels considering buckling and manufacturing constraints[J]. Thin-Walled Structures, 2021, 161: 107526.

[5] BROOKS T R, MARTINS J R R A, KENNEDY G J. Aerostructural tradeoffs for tow-steered composite wings[J]. Journal of Aircraft, 2020, 57(5): 787-799.

[6] PEREIRA D A, SALES T P, RADE D A. Multi-objective frequency and damping optimization of tow-steered composite laminates[J]. Composite Structures, 2021, 256: 112932.

[7] BUTLER R, BAKER N, LIU W L, et al. Damage tolerance of buckling optimized variable angle tow panels[C]//Proceedings of the 50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. Reston: AIAA, 2009: 2443.

[8] ZHANG B, CHEN K L, ZU L. Aeroelastic tailoring method of tow-steered composite wing using matrix perturbation theory[J]. Composite Structures, 2020, 234: 111696.

[9] STODIECK O, COOPER J E, WEAVER P M, et al. Aeroelastic tailoring of a representative wing box using tow-steered composites[J]. AIAA Journal, 2017, 55(4): 1425-1439.

[10] STODIECK O, COOPER J E, WEAVER P M, et al. Optimization of tow-steered composite wing laminates for aeroelastic tailoring[J]. AIAA Journal, 2015, 53(8): 2203-2215.

[11] 李飞, 聂国隽. 基于连续丝束剪切技术的变角度复合材料层合板的热屈曲分析[J]. 力学季刊, 2019, 40(2): 265-273.
LI F, NIE G J. Thermal buckling analysis of VAT composite laminates fabricated by continuous tow shearing[J]. Chinese Quarterly of Mechanics, 2019, 40(2): 265-273(in Chinese).

[12] WU K, TATTING B, SMITH B, et al. Design and manufacturing of tow-steered composite shells using fiber placement[C]//Proceedings of the 50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. Reston: AIAA, 2009: 2700.

[13] XIAO Z F, ACKERMANN A, HARRISON P. Manual 2-dimensional fabric steering, for the manufacture of variable stiffness panels[J]. Procedia Manufacturing, 2020, 47: 24-28.

[14] 孔斌, 顾杰斐, 陈普会, 等. 变刚度复合材料结构的设计、制造与分析[J]. 复合材料学报, 2017, 34(10): 2121-2133.
KONG B, GU J F, CHEN P H, et al. Design, manufacture and analysis of variable-stiffness composite structures[J]. Acta Materiae Compositae Sinica, 2017, 34(10): 2121-2133(in Chinese).

[15] HAO P, LIU D C, WANG Y, et al. Design of manufacturable fiber path for variable-stiffness panels based on lamination parameters[J]. Composite Structures, 2019, 219: 158-169.

[16] SOHOULI A, YILDIZ M, SULEMAN A. Cost analysis of variable stiffness composite structures with application to a wind turbine blade[J]. Composite Structures, 2018, 203: 681-695.

[17] CHEN S Z, LI D C, XIANG J W, et al. Composite manufacturing cost model targeting on design optimization[J]. Applied Composite Materials, 2020, 27(5): 673-691.

[18] KIM B C, POTTER K, WEAVER P M. Continuous tow shearing for manufacturing variable angle tow composites[J]. Composites Part A:Applied Science and Manufacturing, 2012, 43(8): 1347-1356.

Design optimization of tow-steered composite structure targeting on manufacturing cost

CHEN Shize*, LI Daochun, XIANG Jinwu

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

Abstract: Due to a number of limitations, traditional composite structure design approaches are unable to completely use the design space of composite. The realization of tow-steered composite brings new opportunities for improvement of the performance of composites. In this research, a tow-steered composite structure design optimization method is established. The optimization method was based on the manufacturing process cost model, manufacturing cost was set as the target function, performance and manufacturing constraints were established, and the fiber path function method and shifted method were selected. On a high respect ratio wing, the tow-steered design optimization realized a 23.87% deduction on manufacturing cost and a 35.58% deduction on structural weight. The research further analyzed the impacts fiber path exerts on configuration, performance, weight, and manufacturing processes.

Keywords: composite; tow-steer; manufacturing cost; structure optimization; genetic algorithm

Received: 2021-11-10; **Accepted**: 2022-01-02; **Published Online**: 2022-01-29 07:02

URL: kns.cnki.net/kcms/detail/11.2625.V.20220128.1217.001.html

* **Corresponding author**. E-mail: chensz@cgwic.com

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0689

基于直角网格伴随自适应的声爆预测

朱震浩[1]，肖天航[1,*]，徐雅楠[1]，邓双厚[1]，张珍铭[2]，余雄庆[1]

(1. 南京航空航天大学 航空学院，南京 210016；2. 上海机电工程研究所，上海 201109)

摘　　要：声爆预测是超声速民机设计关键技术之一，近场过压信号的精确计算是声爆预测的基础和重要环节。针对超声速民机声爆预测问题，发展了基于直角网格伴随自适应的超声速近场求解方法。运用有限体积法求解流体控制方程，以指定近场位置过压值平方的积分作为目标函数，通过求解离散伴随方程获得目标函数对流场残差的敏感度，评估和修正目标函数全局误差，并驱动网格自适应加密，通过局部加密逐渐减小各个网格单元的剩余误差，提高目标函数的计算精度。选取69°后掠三角翼翼身组合体和美国国家航空航天局（NASA）的C25D超声速民机模型进行算例验证并与试验数据对比，结果表明：与流场特征自适应网格方法相比，所提方法能够以更少的网格量，更精确地捕捉流场中的激波和膨胀波，得到更精确的近场过压信号分布，是一种可用于声爆预测的高精度、高效率的超声速近场求解方法。

关 键 词：超声速民机；声爆预测；伴随；误差估计；网格自适应；直角网格

中图分类号：V221^{+}.3; TB553

文献标志码：A　　**文章编号**：1001-5965(2023)09-2432-10

超声速民机能够大大节约飞行时间，一直存在着可观的市场需求。虽然自“协和”号最后一次飞行后，没有任何新的超声速民用飞机问世，但全球各先进航空大国一直没有放弃其市场需求和相关技术的研究。技术的进步和全球商务旅行的蓬勃发展，或将促使超声速民机重新崛起。根据国内外学者，如文献[1-5]对超声速民机相关技术发展的综述，低声爆设计是公认的超声速民机商业化的关键问题之一。

地面声爆强度的准确预测，是评估和降低声爆水平的前提。声爆传播距离长，且受大气变化和地球湍流边界层等因素影响，声爆预测存在较大的困难。文献[3,6-11]对声爆预测技术进行了研究和总结，至今已发展出多种声爆预测技术。其中一种兼顾效率与精度的声爆预测方法是基于计算流体力学(computational fluid dynamics, CFD)近场计算和远场声爆预测相结合的方法，该方法把声爆的计算分为2个区域，即近场声爆的产生与远场声爆的传播[12]。利用近场CFD计算取得距离机身3～5倍机身长度处的较弱的过压信号，作为远场声爆传播程序的输入。很显然，该方法对地面声爆强度的预测精度取决于近场过压信号的精度，因此，对近场CFD计算格式和网格分辨率提出了较高的要求。另一方面，在超声速民机多学科设计优化中，声爆是其中重要的设计约束或优化目标，这也对近场计算网格生成和CFD计算提出了自动化和效率的要求。美国国防高级研究计划局(Defense Advanced Research Projects Agency, DARPA)开展的修形声爆试验(shaped sonic boom demonstrator, SSBD)项目进行的CFD分析[13-14]也表明高精度的声爆预测对高分辨率网格的依赖性，但是较大的网格量导致的计算时间增加，使声爆约束难以包含进项目开发的多学科设计优化框架中。因此，自动化程度高、既快

收稿日期：2021-11-16；**录用日期**：2022-01-05；**网络出版时间**：2022-01-25 11:22

网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220124.1709.004.html

基金项目：国家自然科学基金(11672133)；陕西省自然科学基础研究计划(2021JQ-078)；航空科学基金(F2021110)；航天进入减速与着陆技术实验室开放基金(EDL19092111)

* **通信作者**. E-mail：xthang@nuaa.edu.cn

引用格式：*朱震浩，肖天航，徐雅楠，等. 基于直角网格伴随自适应的声爆预测[J]. 北京航空航天大学学报，2023，49(9)：2432-2441.*
ZHU Z H，XIAO T H，XU Y N，et al. Adjoint-based adaptive Cartesian mesh refinement for sonic boom prediction[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2432-2441 (in Chinese).

速又能保证精度的近场过压信号流场计算成为超声速民机设计亟需解决的关键问题和研究热点之一。

目前, 国内外已开展了不少面向声爆预测的超声速近场计算技术及应用研究。文献 [14-16] 比较了黏性湍流计算和无黏流计算的近场过压信号分布, 认为声爆预测更多关注流场激波的传播, 而非精确的气动力, 对声爆预测而言, 效率更高的基于 Euler 方程的流场求解即可满足需要。文献 [17-19] 开展了基于结构网格、非结构网格的近场计算研究, 通过合理的网格疏密布置, 其计算可取得较满意的精度; 但就网格质量和自动化程度而言, 相比于结构网格和传统四面体非结构网格, 直角网格生成自动化程度高, 网格质量好, 网格量少且容易实现自适应加密, 更满足方案设计和多学科设计优化对效率和自动化程度的需求[20-22]。文献 [23-25] 建立了基于流场特性的网格自适应技术, 利用流场梯度作为判据, 自适应地在局部特征区域, 如激波、边界层、尾流等, 进行局部网格加密以减小流场局部误差、提高近场过压信号捕捉精度; 但流场特征处的局部误差并不一定完全关联升、阻力或近场过压信号等全局函数的误差, 减小某些流场特征的局部误差并不能保证全局函数的精度。针对此, 文献 [26] 提出基于伴随误差分析的网格自适应技术, 即通过伴随方程建立局部误差与目标函数全局误差之间的关系, 利用伴随方程获得目标函数对流场残差的敏感度, 估计目标函数全局误差, 并驱动局部网格自适应, 提高目标函数的计算精度。文献 [27-30] 分别将此伴随自适应技术应用于非结构网格和直角网格, 最大程度降低网格量, 同时提高了近场过压信号捕捉精度, 证实该方法在声爆预测上的可行性和有效性。

基于自编的直角网格生成与流场求解程序, 发展了伴随自适应直角网格技术, 并将其用于超声速近场过压信号的捕捉, 以实现高效、精确的声爆预测。选取 2 个有风洞试验数据的超声速模型进行测试, 比较 4 种网格划分策略, 即常规初始网格(局部加密)、旋转计算域初始网格(局部加密)、在旋转计算域初始网格的基础上采用基于流场特征的网格自适应及基于伴随的网格自适应, 捕捉飞机正下方近场过压信号, 并与试验数据对比, 以验证和说明直角网格伴随自适应技术在声爆预测上的有效性。

1　数值计算方法

1.1　初始直角网格生成

直角网格兼具传统结构与非结构网格的优点, 包括网格单元规则、生成过程简单、易于实现网格自适应等。直角网格的生成过程是利用 quadtree (二维) 或 octree(三维)不断细分计算域的过程, 采用课题组前期研究和开发的直角网格生成方法与程序[20-22], 对本文算例自动生成直角网格, 此处简单介绍其基本思路和步骤:

1) 几何模型和数据结构的管理。将输入的几何模型处理成 STL 格式作为几何输入文件, 通过交替数字树 (alternating digital tree, ADT) 数据结构来管理物面边, 建立八叉树数据结构定义网格的拓扑关系。

2) 网格初始化和几何自适应。从八叉树的根节点开始生成初始均匀网格, 将树中的每个叶子节点加密, 直至所有叶子节点层数加密至设定的初始层数。但均匀网格不能精确描述几何外形, 需进行几何自适应局部加密, 单元是否加密主要基于模型表面曲率并结合用户设定的最小单元尺寸来判断。此外, 也可设定区域(如长方体、球体区域)进行空间单元的局部加密。

3) 网格单元的分类与光顺。为满足后续处理的需要, 将几何自适应生成的网格单元分为 3 类: 流场内非切割单元、固壁内非切割单元和物面边界单元。使用 ADT 算法搜索与物面相交的单元, 确定物面边界单元。随后流场内外非切割单元使用光线投射算法进行类型判断, 并使用染色算法快速确定其单连通域内其他单元的类型。最后删除固壁内非切割单元。分类完后, 对不同网格进行适当加密和粗化, 以方便后续物面边界单元的切割处理和保证计算的稳健性。

4) 物面边界切割处理。采用切割法处理物面边界单元, 建立各种物面与直角网格单元交叉的切割处理算法, 通过切割、连接、合并处理生成贴体的切割单元。

1.2　流体控制方程及求解

采用无黏欧拉方程求解流场, 其三维积分形式为

$$\frac{\partial}{\partial t}\int_{\Omega}\boldsymbol{Q}\mathrm{d}V+\oint_{\partial\Omega}\boldsymbol{F}(\boldsymbol{Q})\mathrm{d}S=\boldsymbol{0} \tag{1}$$

式中: $\boldsymbol{Q}=[\rho, \rho u, \rho v, \rho w, \rho E]^{\mathrm{T}}$ 为流场守恒变量, ρ 为密度, u、w、v 为速度在 x、y、z 轴上的分量, E 为总能量; $\boldsymbol{F}(\boldsymbol{Q})$ 为对流通量; Ω 为任意网格单元。式(1)在空间上采用有限体积法进行离散, 运用 Roe 迎风格式计算控制体交界面对流通量, 为保证二阶空间精度, 交界面两侧流场值由最小二乘法或高斯积分法重构的流场梯度线性插值得到, 并使用 Venkatakrishnan 限制器防止新的极值出现; 时间方向上运用后向差分隐式格式离散, 得到的离散方程可表示为

$$\left(\frac{V}{\Delta t}\boldsymbol{I}+\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{Q}}\right)\Delta\boldsymbol{Q}+\boldsymbol{R}(\boldsymbol{Q})=\boldsymbol{0} \tag{2}$$

式中：$\boldsymbol{R}(\boldsymbol{Q})_i=\sum_{j=1}^{N_e}\boldsymbol{F}_{ij}\Delta S$为第 i 个控制体的流场残差，F_{ij} 为控制体 i 和控制体 j 交接面处的通量，N_e 为邻居控制体。离散后的方程采用广义最小残量（generalized minimal RESidual, GMRES）方法迭代求解，收敛后 $\boldsymbol{R}$ 趋于 0。详细的流场求解方法可参考文献 [31]。

2 伴随误差估计

本节简要介绍基于直角网格的伴随误差估计方法。

2.1 目标函数

基于伴随的误差估计方法建立局部误差与所关心的目标函数的直接关系，利用伴随方程可直接度量局部误差对全局误差的影响，从而对目标函数的全局误差进行估计。目标函数可以是物面积分的升力、阻力和动量系数等气动力，也可以是流场中某一关注的变量，如涡扇发动机风扇入口的总压损失或质量流量或是流场任一指定截面上的压力特征。对用于声爆预测的近场流场计算而言，目标函数取为近场过压信号分布，即过压值的平方在指定区域上的积分：

$$f=\iint_S(\Delta p/p_\infty)^2\mathrm{d}s \tag{3}$$

式中：过压值$\Delta p/p_\infty=(p-p_\infty)/p_\infty$，$p$ 为当地压强，p_∞ 为来流压强。

2.2 伴随方程的建立与求解

目标函数 f 可表示为欧拉方程得到的流场解 $\boldsymbol{Q}=[\rho,\rho u,\rho v,\rho w,\rho E]^{\mathrm{T}}$ 的函数：

$$f=f(\boldsymbol{Q}) \tag{4}$$

对其求导，可得

$$\frac{\partial f}{\partial \boldsymbol{Q}}=\left(\frac{\partial f}{\partial \boldsymbol{R}}\right)^{\mathrm{T}}\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{Q}} \tag{5}$$

定义伴随变量 $\boldsymbol{\Lambda}$ 为目标函数对流场残差的梯度，即敏感度：

$$\boldsymbol{\Lambda}=\frac{\partial f}{\partial \boldsymbol{R}} \tag{6}$$

结合式 (5) 和式 (6)，可建立伴随方程为

$$\left(\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{Q}}\right)^{\mathrm{T}}\boldsymbol{\Lambda}=\left(\frac{\partial f}{\partial \boldsymbol{Q}}\right)^{\mathrm{T}} \tag{7}$$

通过伴随方程，建立了流场变量、局部残差和目标函数之间的关系。为方便式 (7) 的求解，引入伪时间，构造与流体控制方程类似的时间推进求解的离散方程为

$$\left(\frac{V}{\Delta\tau}\boldsymbol{I}+\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{Q}}\right)^{\mathrm{T}}\Delta\boldsymbol{\Lambda}=\left(\frac{\partial f}{\partial \boldsymbol{Q}}\right)^{\mathrm{T}}-\left(\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{Q}}\right)^{\mathrm{T}}\boldsymbol{\Lambda}^n \tag{8}$$

式中：$\Delta\boldsymbol{\Lambda}=\boldsymbol{\Lambda}^{n+1}-\boldsymbol{\Lambda}^n$，方程左侧雅克比矩阵用一阶精度格式构造，右侧则用二阶格式以保证精度，伴随方程的求解与主控欧拉方程求解一致，采用 GMRES 方法迭代求解。

2.3 伴随误差修正

考虑平均网格尺寸为 H 的粗网格 Ω_H，在该网格上的目标函数值为 $f(\boldsymbol{Q}_H)$，其中 $\boldsymbol{Q}_H$ 为该粗网格上得到的离散解向量。为了提高预测精度，在粗网格的基础上，通过单元细分的方式进行加密（见图 1），得到平均网格尺寸为 h、网格量约为原来 8 倍的密网格 Ω_h，密网格下得到的函数值为 $f(\boldsymbol{Q}_h)$。尽管该值更加接近目标值，但是由于网格量的增加，计算所需时间也会大大增加。为了减小计算资源，只求解粗网格的流场，然后通过合适的方法对密网格上的解进行预测，并计算离散误差$|f(\boldsymbol{Q}_h)-f(\boldsymbol{Q}_H)|$。

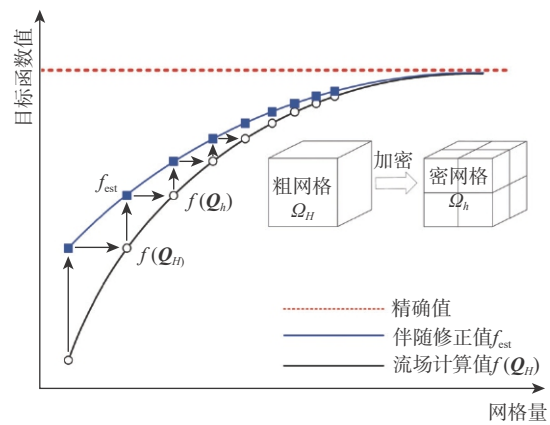


图 1 直角网格伴随自适应过程目标函数变化曲线

Fig. 1 Functional convergence for adjoint-based adaptive Cartesian mesh refinement

将密网格下的函数值 $f(\boldsymbol{Q}_h)$ 在 $\boldsymbol{Q}_h^H$ 处泰勒展开，忽略高阶项后为

$$f(\boldsymbol{Q}_h)\approx f(\boldsymbol{Q}_h^H)+\left.\frac{\partial f(\boldsymbol{Q}_h)}{\partial \boldsymbol{Q}_h}\right|_{\boldsymbol{Q}_h^H}(\boldsymbol{Q}_h-\boldsymbol{Q}_h^H) \tag{9}$$

式中：$\boldsymbol{Q}_h^H$ 为粗网格下的流场解通过重构得到的密网格下的流场解；$\partial f(\boldsymbol{Q}_h)/\partial \boldsymbol{Q}_h\big|_{\boldsymbol{Q}_h^H}$ 为密网格下的目标函数值对通过粗网格流场解重构得到的细网格流场解的敏感度。

类似地，密网格下的残差方程也可近似为

$$\boldsymbol{R}(\boldsymbol{Q}_h)=\boldsymbol{0}\approx\boldsymbol{R}(\boldsymbol{Q}_h^H)+\left.\frac{\partial \boldsymbol{R}(\boldsymbol{Q}_h)}{\partial \boldsymbol{Q}_h}\right|_{\boldsymbol{Q}_h^H}(\boldsymbol{Q}_h-\boldsymbol{Q}_h^H) \tag{10}$$

结合式 (9) 和式 (10)，即可得到通过粗网格下的函数值与流场解，预测的细网格下的函数值为

$$f(\boldsymbol{Q}_h)\approx f(\boldsymbol{Q}_h^H)-\left(\boldsymbol{\Lambda}_h|_{\boldsymbol{Q}_h^H}\right)^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{Q}_h^H) \tag{11}$$

式中：表征目标函数对流场残差的敏感度的离散伴随变量 $\boldsymbol{\Lambda}_{\boldsymbol{Q}_h^H}$ 可通过求解以下伴随方程得到：

$$\left(\frac{\partial \boldsymbol{R}(\boldsymbol{Q}_h)}{\partial \boldsymbol{Q}_h}\bigg|_{\boldsymbol{Q}_h^H}\right)^{\mathrm{T}}\boldsymbol{\Lambda}_h|_{\boldsymbol{Q}_h^H}=\left(\frac{\partial f(\boldsymbol{Q}_h)}{\partial \boldsymbol{Q}_h}\bigg|_{\boldsymbol{Q}_h^H}\right)^{\mathrm{T}} \tag{12}$$

显然，要求解式 (11)，仍然需要求解密网格上的伴随方程式 (12)，而这又是需要避免的。于是，类似地，不直接求解密网格上的伴随变量，而是用粗网格上的伴随变量$\boldsymbol{\Lambda}_H$，通过插值的方法近似，重构的伴随变量用$\boldsymbol{\Lambda}_h^H$表示。

$$\left(\frac{\partial \boldsymbol{R}(\boldsymbol{Q}_H)}{\partial \boldsymbol{Q}_H}\right)^{\mathrm{T}}\boldsymbol{\Lambda}_H=\left(\frac{\partial f(\boldsymbol{Q}_H)}{\partial \boldsymbol{Q}_H}\right)^{\mathrm{T}} \tag{13}$$

将重构的伴随变量用$\boldsymbol{\Lambda}_h^H$代入式 (11)，即在不需要直接求解密网格流场情况下，得到密网格下的预测值为

$$f(\boldsymbol{Q}_h)\approx f_{\mathrm{est}}=f(\boldsymbol{Q}_h^H)-\left(\boldsymbol{\Lambda}_h^H\right)^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{Q}_h^H) \tag{14}$$

式中：f_{est}为可计算的目标函数（见图 1），等于插值得到的密网格函数值$f(\boldsymbol{Q}_h^H)$减去伴随修正项$\left(\boldsymbol{\Lambda}_h^H\right)^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{Q}_h^H)$。

为了兼顾插值的鲁棒性和准确性，采用高阶多项式插值法，获得密网格下的重构解，即$\boldsymbol{Q}_{\mathrm{P}}$和$\boldsymbol{\Lambda}_{\mathrm{P}}$，将其代入式 (14)，则可得到高阶目标函数的估计值为

$$f(\boldsymbol{Q}_h)\approx f_{\mathrm{est}}=f(\boldsymbol{Q}_{\mathrm{P}})-(\boldsymbol{\Lambda}_{\mathrm{P}})^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{Q}_{\mathrm{P}}) \tag{15}$$

需要注意的是，为了保证目标函数的准确性，要求粗网格具有一定的分辨率，使得采用插值方法能够较好的获得细网格上的重构解。

2.4　误差估计

尽管通过式 (14) 得到了目标函数在密网格下修正后的估计值，即使插值足够精确，对于非线性的目标函数仍会存在一定的剩余误差。直角网格伴随自适应技术就是将剩余误差作为自适应参数，通过局部加密逐渐减小各个网格单元的剩余误差，以提高目标函数估计值的计算精度。

结合式 (11) 和式 (14)，得到目标函数修正后的剩余误差f_{err}为

$$f(\boldsymbol{Q}_h)-f_{\mathrm{est}}\approx f_{\mathrm{err}}=\left(\boldsymbol{\Lambda}_h-\boldsymbol{\Lambda}_h^H\right)^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{Q}_h^H) \tag{16}$$

由式（16）可知，伴随方法不仅给出了修正项，同时也得到了剩余误差项。通过单元细分的方式，每个粗网格单元k的局部误差上限可以通过剩余误差确定，为粗网格单元k细分的密网格单元剩余误差之和：

$$e_k=\sum_{N(k)}\left|\left(\boldsymbol{\Lambda}_h-\boldsymbol{\Lambda}_h^H\right)^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{Q}_h^H)\right|_{N(k)} \tag{17}$$

式中：$N(k)$为粗网格单元k细分的密网格单元数量。

同样，采用高阶多项式插值代替计算代价较大的$\boldsymbol{\Lambda}_h$，可得

$$e_k=\sum_{N(k)}\left|(\boldsymbol{\Lambda}_{\mathrm{P}}-\boldsymbol{\Lambda}_{\mathrm{L}})^{\mathrm{T}}\boldsymbol{R}(\boldsymbol{Q}_{\mathrm{P}})\right|_{N(k)} \tag{18}$$

最后，定义目标估计的全局误差为每个粗网格单元的局部误差之和：

$$E=\sum e_k \tag{19}$$

3　直角网格自适应策略

为了对比直角网格伴随自适应方法与常规基于流场特征的网格自适应方法，本节简要介绍两者自适应判据的选择及自适应策略。

3.1　基于流场特征的网格自适应

基于流场特征的网格自适应一般在大梯度区域对网格进行加密，如激波、边界层、尾流等，认为局部误差与流场特征量相关，利用流场特征作为自适应判据，标记需要加密的区域，通过局部网格加密以减小流场局部误差。因此，该方法的关键在于建立合适的自适应参数以精确的捕捉相应的流场特征，以超声速民机关注的激波捕捉为例，一般使用压力梯度、密度梯度或速度梯度作为判据。首先定义自适应参数为

$$\tau_{\mathrm{d}k}=|\nabla g|\,l_k^{\frac{r+1}{r}} \tag{20}$$

式中：g为流场压力、密度或速度量；l_k为该粗网格单元k体积的算术立方根；r为调节参数。以所有粗网格单元的$\tau_{\mathrm{d}k}$的均方差作为基准：

$$\tau_{\mathrm{d}}=\sqrt{\sum_{k=1}^{K}\tau_{\mathrm{d}k}^2\Big/K} \tag{21}$$

式中：K为粗网格网格量。

基于流场特征的自适应判据为

$$\eta=\frac{\tau_{\mathrm{d}k}}{\tau_{\mathrm{d}}} \tag{22}$$

当η>1时，则认为粗网格单元k需要被加密。

3.2　基于直角网格伴随自适应

与基于流场特征的网格自适应不同，如 2.4 节所述，基于直角网格伴随自适应关注的是目标函数，通过局部加密网格来减小剩余误差，从而提高目标函数伴随修正后的精度。基于式 (18) 和式 (19)，在得到粗网格上的局部和全局误差估计之后，基于设定的目标函数全局误差容忍上限T_{TOL}，和每个粗网格单元的平均局部误差上限t=T_{TOL}/K，定义局部误差自适应参数为

$$\eta_k=\frac{e_k}{t} \tag{23}$$

当η_k>1时，表明目标函数的局部误差大于平均局部误差上限，此粗网格单元将被加密。

对于目标函数的全局误差，同样定义一个全局误差自适应参数为

$$\eta_g = \frac{E}{T_{TOL}} \tag{24}$$

当 η_g>1 时，表明目标函数的全局误差大于给定的容忍上限，网格需要进一步加密。

因为单独使用局部或全局自适应参数，会导致网格被过度加密或加密程度不够，所以结合局部和全局自适应参数，给定一个综合自适应参数：

$$\eta = \eta_k \eta_g \tag{25}$$

当 η>1 时，则认为粗网格单元 k 需要被加密。

4 算例验证与应用

选取 2 个有风洞试验数据的模型进行测试，采用直角网格伴随自适应技术获得飞机正下方近场过压信号，并与试验数据对比，以验证直角网格伴随自适应方法的快速性与准确性。

4.1 69°后掠三角翼翼身组合体模型算例

第 1 个算例为文献 [32] 开展的一系列不同平面构型的翼身组合体声爆试验中的 Model 4 三角翼翼身组合体，几何外形及尺寸如图 2 所示。计算来流马赫数为 1.68，攻角为 0，测量机身正下方 0.54 m（H/L=3.1）和 0.63 m（H/L=3.6）处的过压信号。

采用 4 种不同网格划分策略建立近场流场计算网格，4 种不同网格划分策略的对称面网格及过压值云图如图 3 所示。4 种计算网格分别为常规初始网格（见图 3(a)，初始局部加密，不采用网格自适应），旋转计算域初始网格（见图 3(b)，将计算域旋转一定角度，再生成网格，使得网格单元面与激波面平行，不采用网格自适应），在旋转计算域初始网格的基础上采用基于压力梯度的网格自适应（见图 3(c)）及基于直角网格伴随自适应（见图 3(d)）进行网格加密，通过比较分析计算得到的近场压力分布及计算耗时，考察这 4 种网格划分策略的效率与准确性。对于基于直角网格伴随自适应，目标函数式 (3) 中的指定截面即为关注的 H/L=3.1 和 H/L=3.6 截

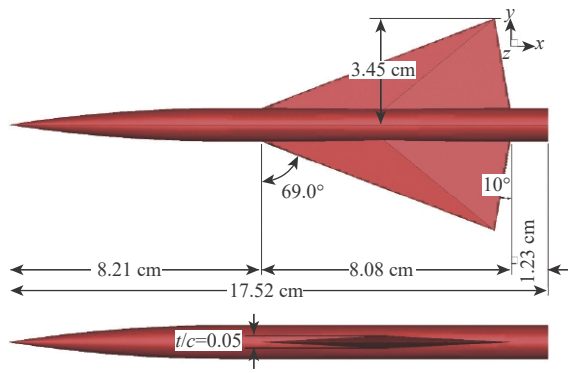
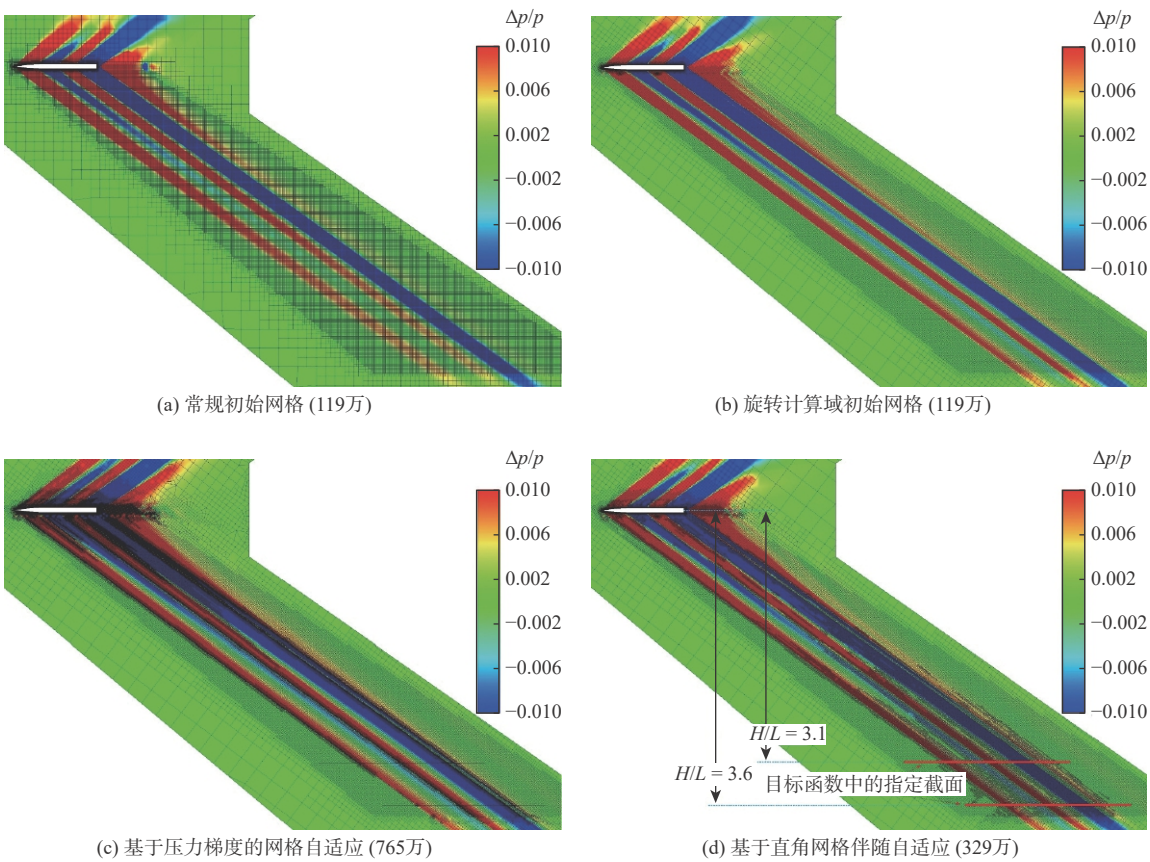


图 2 69°后掠三角翼翼身组合体 (Model 4) 模型几何外形[32]

Fig. 2 Geometry of 69° Delta Wing-Body (Model 4)[32]



(a) 常规初始网格 (119万)　(b) 旋转计算域初始网格 (119万)

(c) 基于压力梯度的网格自适应 (765万)　(d) 基于直角网格伴随自适应 (329万)

图 3 4 种不同网格划分策略的对称面网格及过压值云图

Fig. 3 Final meshes generated by four strategies and corresponding over-pressure contours at symmetry plane

面，图 3 中给出的是最终网格。

4 种不同网格计算结果如图 3 所示，对比图 3(a)和图 3(b) 可以发现，当通过旋转使直角网格单元面与激波面平行时，机身正下方激波及膨胀波的强度与位置的捕捉明显更加精确，尤其是距离机身较远处。图 3(c) 和图 3(d) 所示的网格自适应得到的最终网格，都明显在激波处进行了局部加密，使得激波和膨胀波的捕捉更加细致，尤其是机翼所导致的。但是基于压力梯度的网格自适应得到的网格在空间分布及过渡上更加流畅、激波处的加密程度更加精细，而基于直角网格伴随自适应则相反；基于直角网格伴随自适应加密的范围主要集中在所关注的区域，表明其他区域对目标函数的误差贡献较小，而基于压力梯度的网格自适应的加密范围明显更广，如对非关注截面的紧靠机身的区域也进行了较精细的加密。因此，基于压力梯度的网格自适应得到的最终网格的网格量约为基于直角网格伴随自适应的 2 倍。

以机身正下方 H/L=3.1 截面为例，图 4 为 4 种网格与风洞试验得到的过压信号对比，其中 3 次压力突变分别表示头部激波、机翼激波及尾部激波。通过对比，常规体加密的初始网格得到的结果最差，仅仅大致捕捉到了激波位置，对过压值峰值及激波导致的压力突变都没有捕捉到，激波处的压力变化很平缓。而在相同网格量下，仅将计算域基于马赫角旋转后再生成网格，对压力突变的捕捉明显提高，逐渐逼近试验值，主要因为此时激波处的直角网格单元的通量和梯度的计算更加准确。但此时激波处压力变化仍较平缓，且过压值幅值也不精确。最后，对比 2 种网格自适应方法的结果可以发现，两者对于激波位置及其导致的压力阶跃的捕捉都更加精确。但是两者在头部激波和尾部激波幅值的捕捉上与试验值相比仍有误差，且如图 4 中椭圆虚线框内所示，基于直角网格伴随自适应对于机翼附近的激波及各激波与膨胀波之间压力的变化的捕捉更加精确，明显优于压力梯度网格自适应。

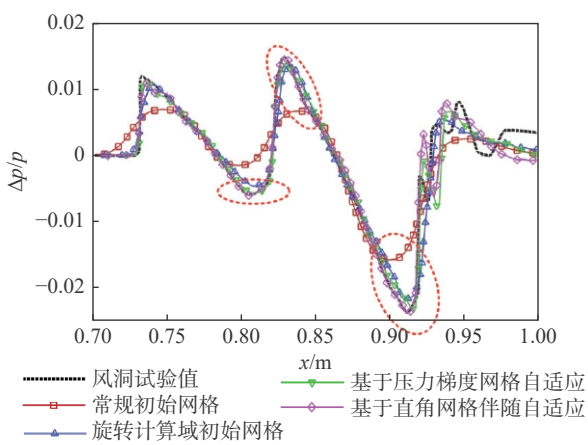


图 4　H/L=3.1 处 4 种网格及风洞试验得到的过压信号对比

Fig. 4　Comparison of pressure signatures obtained from numerical simulations of four different grids and wind tunnel tests at H/L=3.1

为了进一步研究直角网格伴随自适应，图 5 为不同自适应次数（adapoive mesh refinement, AMR）得到的 H/L=3.1 处过压信号，并与风洞试验进行了对比。可以发现，随着自适应次数的增加，与试验值的吻合度越来越高，且在第 5 次自适应后，基本与试验值重合。采用机翼参考面积无量纲化目标函数的收敛曲线如图 6(a) 所示，也可以发现，自适应 5 次后，目标函数趋于恒定值，变化较小。从图 6(a) 中还能够看出，随着自适应次数的增加，粗网格下得到的伴随修正值能够较好的预测下一步加密后的目标函数值。图 6(b) 对比了基于直角网格伴随自适应与压力梯度网格自适应的相对误差收敛情况，其中基于压力梯度网格自适应的相对误差是以采用 Richardson 外插法[33] 得到的目标函数外插值为基准，基于直角网格伴随自适应的相对误差定义为

$$R_{\text{error}} = |f(\boldsymbol{Q}_h) - f_{\text{est}}| = \left| f(\boldsymbol{Q}_h) - \left[f(\boldsymbol{Q}_h^H) - (\boldsymbol{\Lambda}_h^H)^{\mathrm{T}} \boldsymbol{R}(\boldsymbol{Q}_h^H) \right] \right| \tag{26}$$

不难发现，相同网格量下，基于直角网格伴随自适应的计算精度比基于压力梯度网格自适应高一阶多，误差修正效果显著，且更大的斜率表明收敛速度更快。这意味着基于直角网格伴随自适应的声爆预测方法能够以更少的网格量更快的达到目标计算精度。

为了更直观的验证基于直角网格伴随自适应技术的高效性，表 1 为 4 种网格最终网格量、计算精度和计算耗时明细，所有工况都是在 12 线程、主频为 4.1 GHz 的工作站进行计算，其中前 2 种网格均计算了 3 000 步并收敛，压力梯度自适应进行了 15 次，基于直角网格伴随自适应 11 次后终止。不难看出，基于马赫角旋转计算域得到的网格能够显

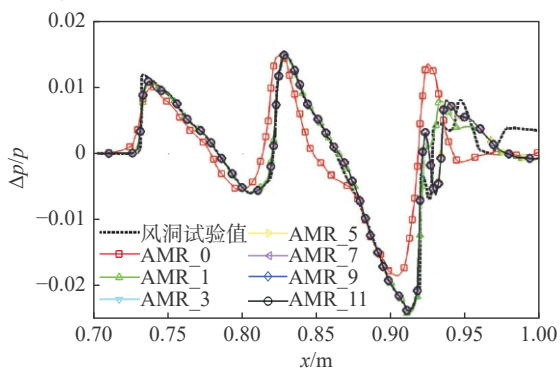
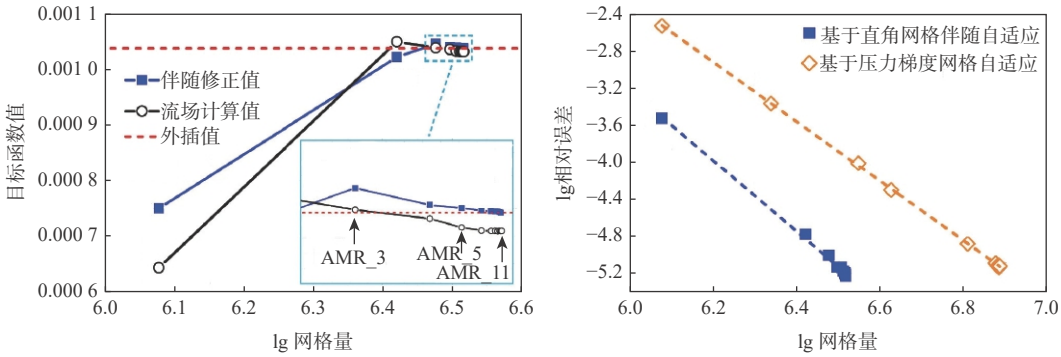


图 5　不同自适应次数得到的 H/L=3.1 处过压信号

Fig. 5　Pressure signatures at H/L=3.1 measured after different refinement times



(a) 目标函数收敛曲线　　(b) 相对误差收敛曲线对比

图 6　*H*/*L*=3.1 处目标函数收敛曲线和直角网格伴随自适应与压力梯度自适应的误差收敛对比

Fig. 6　Functional convergence using adjoint-based Cartesian mesh refinement and comparison of relative error for adjoint-based and pressure gradient-based mesh refinement at *H*/*L*=3.1

表 1　4 种网格最终网格量、计算精度和计算耗时明细

Table 1　Comparison of final grid size, computational accuracy and cost for four grid strategies

网格	网格量/10^4	计算精度	计算耗时/h
常规初始网格	119	很差	2
旋转计算域初始网格	119	一般	2
基于压力梯度网格自适应	765	较精确	28
基于直角网格伴随自适应	329	更精确	12 (6)

著提高激波的捕捉精度。基于直角网格伴随自适应由于加密区域仅集中于对目标函数敏感度较大的区域，网格量更少，于是计算耗时更少，且由图 6(b)可知其计算精度更高。由图 5 和图 6 也可看出，伴随自适应 5 次即可达到所需精度，那么计算耗时仅为 6 h，只有常规压力梯度自适应的 1/4。因此，由于其快速性与准确性，该方法在超声速民机声爆预测中有着很好的发展前景。

4.2　带动力短舱的 C25D 模型算例

第 2 个算例为带动力短舱的 C25D 超声速民机模型[34-35]，以验证基于直角网格伴随自适应技术应用于复杂外形的真实超声速民机的可行性。该模型取自第 2 届 AIAA 声爆预测会议，几何外形如图 7 所示，作为特征长度的机身总长 L_{ref}=32.92 m，参考面积 S_{ref}=37.16 m^2(半模)。风洞试验由文献 [35] 提供，计算工况：来流马赫数为 1.6；攻角为 3.375°；来流压力为 10 674.3 Pa；发动机风扇入口截面总压比为 3.260 6，马赫数为 0.609 3；发动机尾喷口截面总压比为 14.54，总温比为 7.872 2，马赫数为 0.662。

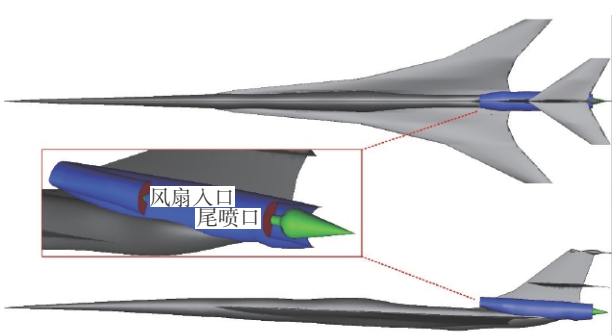
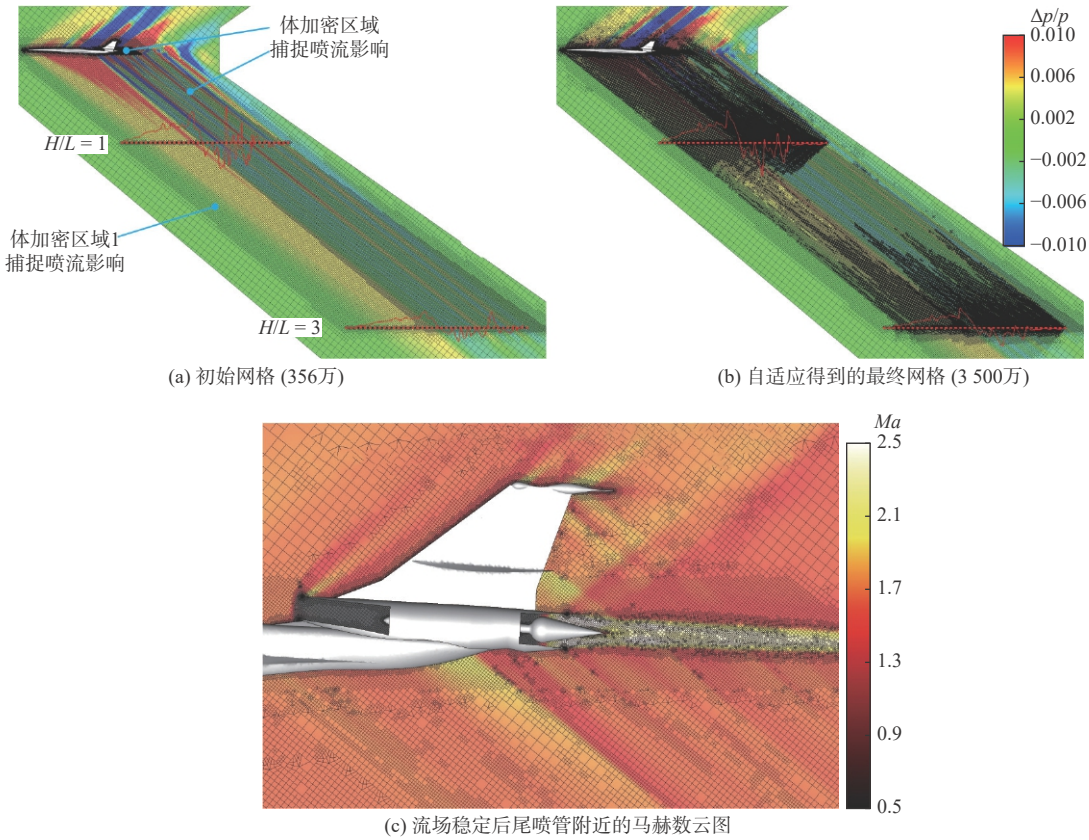


图 7　带动力短舱的 C25D 几何外形[34-35]

Fig. 7　Geometry of C25D configuration with powered nacelle[34-35]

初始网格如图 8(a) 所示，为了保证粗网格具有一定的分辨率，设置了 3 个体加密区域，其中一个用于捕捉常规激波及膨胀波的传播，另外 2 个网格单元更小的区域用于捕捉发动机喷流对激波的传播的影响，总网格量为 356 万。根据计算得到的过压值云图及 2 个高度处的过压信号，发现尾部后方区域过压信号的变化很剧烈，这意味着发动机喷流会导致额外激波的产生，且对飞机尾部激波与膨胀波系的传播造成明显的影响，但由于此时网格分辨率不高，并没有得到较好的捕捉。图 8(b) 为网格自适应 20 次后得到的最终网格，可以发现在关注的 2 个高度附近的激波处进行了精细的局部加密，更精确地捕捉到压力的变化，但是网格量也增加到了 3 500 多万。通过图 8(c) 的马赫数云图可以发现，发动机喷流区域网格也得到一定的加密，尽管对于喷流内部的菱形激波及马赫盘结构都没有捕捉到，但是能够较好地捕捉处于轻微过膨胀状态的发动机喷流的外轮廓边界，且尾部区域复杂的流场结构足以说明发动机喷流会与上方尾翼造成的激波相互干扰，波浪形的喷流边界也会带来额外的激波与膨胀波系。

由于一般声爆预测的近场过压信号取为机身下方 3 倍特征长度处，以 *H*/*L*=3 为例，目标函数收敛曲线及不同自适应次数得到的 *H*/*L*=3 处的过压信号，如图 9 所示。类似地，随着自适应次数的增加，目标函数逐渐收敛，伴随修正值对下一步密网格的预测也越来越准确。13 次自适应后，网格量达到了 1 800 多万，目标函数基本不再有较大的变化，逐渐收敛，且随后的过压信号也基本与试验值重合(见图 9(b))，因此，可以认为此时已满足要求。对

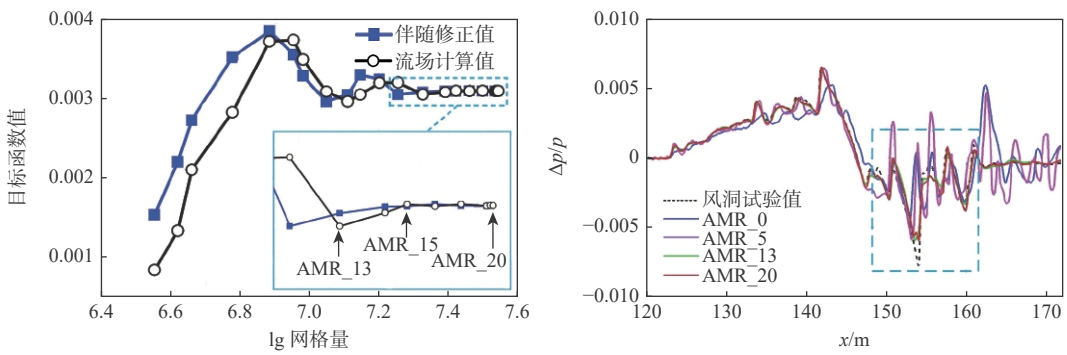


(a) 初始网格 (356万)　　(b) 自适应得到的最终网格 (3 500万)

(c) 流场稳定后尾喷管附近的马赫数云图

图 8　C25D 模型初始网格与最终的自适应网格示意及各自计算结果

Fig. 8　The initial and final refined mesh for C25D model as well as the corresponding results



(a) 目标函数收敛曲线　　(b) H/L=3处过压信号

图 9　目标函数收敛曲线及不同自适应次数得到的 H/L=3 处的过压信号

Fig. 9　Functional convergence and pressure signatures at H/L=3 measured after different refinement times

于 x<148 m 的区域，机头、机翼及发动机之前的机身造成的激波与膨胀波的捕捉，较为容易，经过 5 次自适应后已较精确，但是对于尾部激波，即 x>148 m 的图中虚线框内区域，捕捉较难，即使是目前 20 次自适应得到的网格，对该区域内激波位置及峰值的捕捉与试验值都有偏差。这主要是因为对发动机喷流的模拟不够精确，而喷流与尾翼的相互干扰对激波和膨胀波的影响又较大，所以必将出现偏差。调研发现，对于考虑发动机喷流的声爆预测，仅以过压值为目标函数，发动机入口、风扇入口及尾喷口的质量流量可能难以收敛，从而导致不准确的喷流模拟。文献 [30] 加入了与发动机进出口质量流量相关的目标函数，结果表明进气道及喷流区域的网格得到了更好的局部加密，精确捕捉到尾喷流内部的菱形激波与马赫盘结构，因此，对尾部过压信号的捕捉更加精确。此外，由于采用 Euler 方程求解流场，无法模拟得到发动机高速喷流与主流之间强烈的剪切和湍流掺混现象，可能对机身下方的波系存在影响。

5　结　论

1）基于马赫角旋转计算域再生成网格，使得流

场中直角网格单元与激波面大致平行，相同网格量下能够极大提高激波的捕捉精度。

2）本文方法与关注的目标函数直接相关，仅在对目标函数误差敏感度影响较大的区域进行局部加密，虽然不一定能捕捉到所有的流场细节特征，但是能对目标函数的全局误差进行精确的估计并对计算结果进行修正。

3）相比于常规的流场特征网格自适应，相同网格量下直角网格伴随自适应技术计算精度更高、收敛更快，对关注高度处的近场过压信号的捕捉更加精确，因此，是一种可用于声爆预测的高精度、高效率的超声速近场求解方法。

4）对于带动力短舱的超声速民机近场过压信号的计算，对喷流导致的激波与膨胀波系的捕捉还不够精确。一方面由于Euler方程无法精确模拟高速喷流与主流之间的剪切和湍流掺混现象；另一方面因为仅以过压值为目标函数，发动机入口、风扇入口及尾喷口的质量流量可能难以收敛，从而导致不准确的喷流模拟，若加入与发动机质量流量相关的目标函数，可能可以提高喷流的捕捉精度，有待进一步研究。

参考文献（References）

[1] 朱自强, 兰世隆. 超声速民机和降低音爆研究[J]. 航空学报, 2015, 36(8): 2507-2528.
ZHU Z Q, LAN S L. Study of supersonic commercial transport and reduction of sonic boom[J]. Acta Aeronautica et Astronautica Sinica, 2015, 36(8): 2507-2528(in Chinese).

[2] 韩忠华, 乔建领, 丁玉临, 等. 新一代环保型超声速客机气动相关关键技术与研究进展[J]. 空气动力学学报, 2019, 37(4): 620-635.
HAN Z H, QIAO J L, DING Y L, et al. Key technologies for next-generation environmentally-friendly supersonic transport aircraft: A review of recent progress[J]. Acta Aerodynamica Sinica, 2019, 37(4): 620-635(in Chinese).

[3] 钱战森, 韩忠华. 声爆研究的现状与挑战[J]. 空气动力学学报, 2019, 37(4): 601-619.
QIAN Z S, HAN Z H. Progress and challenges of sonic boom research[J]. Acta Aerodynamica Sinica, 2019, 37(4): 601-619(in Chinese).

[4] SMITH H. A review of supersonic business jet design Issues[J]. The Aeronautical Journal, 2007, 111(1126): 761-776.

[5] SUN Y C, SMITH H. Review and prospect of supersonic business jet design[J]. Progress in Aerospace Sciences, 2017, 90: 12-38.

[6] 钱战森, 刘中臣, 冷岩, 等. OS-X0试验飞行器声爆特性飞行测量与数值模拟分析[J]. 空气动力学学报, 2019, 37(4): 675-682.
QIAN Z S, LIU Z C, LENG Y, et al. Flight measurement and numerical simulation of sonic boom signature of OS-X0 experimental aircraft[J]. Acta Aerodynamica Sinica, 2019, 37(4): 675-682(in Chinese).

[7] 王刚, 马博平, 雷知锦, 等. 典型标模音爆的数值预测与分析[J]. 航空学报, 2018, 39(1): 121458.
WANG G, MA B P, LEI Z J, et al. Simulation and analysis for sonic boom on several benchmark cases[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(1): 121458(in Chinese).

[8] 冯晓强, 李占科, 宋笔锋. 超音速客机音爆问题初步研究[J]. 飞行力学, 2010, 28(6): 21-23.
FENG X Q, LI Z K, SONG B F. Preliminary analysis on the sonic boom of supersonic aircraft[J]. Flight Dynamics, 2010, 28(6): 21-23(in Chinese).

[9] 兰世隆. 超声速民机声爆理论、预测和最小化方法概述[J]. 空气动力学学报, 2019, 37(4): 646-654.
LAN S L. Overview of sonic boom theory, prediction and minimization methods for supersonic civil aircraft[J]. Acta Aerodynamica Sinica, 2019, 37(4): 646-654(in Chinese).

[10] PLOTKIN K J. Review of sonic boom theory[C]//12th Aeroacoustic Conference. Reston: AIAA, 1989: 1105.

[11] MAGLIERI D J, BOBBITT P J, PLOTKIN K J, et al. Sonic boom: Six decades of research[M]. Hampton: Langley Research Center, 2014.

[12] PLOTKIN K J. State of the art of sonic boom modeling[J]. The Journal of the Acoustical Society of America, 2002, 111(1): 530-536.

[13] HAERING E, MURRAY J, PURIFOY D, et al. Airborne shaped sonic boom demonstration pressure measurements with computational fluid dynamics comparisons[C]//Proceedings of the 43rd AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2005: 9.

[14] MEREDITH K, DAHLIN J, GRAHAM D, et al. Computational fluid dynamics comparison and flight test measurement of F-5E off-body pressures[C]//Proceedings of the 43rd AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2005: 6.

[15] CASTNER R. Analysis of exhaust plume effects on sonic boom for a 59-degree wing body model[C]//Proceedings of the 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition. Reston: AIAA, 2011: 917.

[16] KIRZ J. DLR TAU simulations for the third AIAA sonic boom prediction workshop near-field cases[C]//Proceedings of the AIAA Scitech 2021 Forum. Reston: AIAA, 2021: 472.

[17] PARK M A, CAMPBELL R L, ELMILIGUI A A, et al. Specialized CFD grid generation methods for near-field sonic boom prediction[C]//Proceedings of the 52nd Aerospace Sciences Meeting. Reston: AIAA, 2014: 115.

[18] ISHIKAWA H, MAKINO Y, ITO T, et al. Sonic boom prediction using multi-block structured grids CFD code considering jet-on effects[C]//Proceedings of the 27th AIAA Applied Aerodynamics Conference. Reston: AIAA, 2009: 3508.

[19] CLIFF S E, THOMAS S D. Euler/experiment correlations of sonic boom pressure signatures[J]. Journal of Aircraft, 1993, 30(5): 669-675.

[20] LYU F X, XIAO T H, YU X Q. A fast and automatic full-potential finite volume solver on Cartesian grids for unconventional configurations[J]. Chinese Journal of Aeronautics, 2017, 30(3): 951-963.

[21] 吕凡熹, 肖天航, 余雄庆. 基于自适应直角网格的二维全速势方程有限体积解法[J]. 计算力学学报, 2016, 33(3): 424-430.
LYU F X, XIAO T H, YU X Q. A finite volume method for 2D Full-Potential equation on adaptive Cartesian grids[J]. Chinese Journal of Computational Mechanics, 2016, 33(3): 424-430(in Chinese).

[22] 吕凡熹, 李正洲, 邓经枢, 等. 面向飞行器概念设计的全速域气动分析工具[J]. 空气动力学学报, 2017, 35(5): 625-632.
LYU F X, LI Z Z, DENG J S, et al. An aerodynamic analysis tool

for aircraft conceptual design at full speed range[J]. Acta Aerodynamica Sinica, 2017, 35(5): 625-632(in Chinese).

[23] NAYANI S. Evaluation of grid modification methods for on-and off-track sonic boom analysis[C]//Proceedings of the 51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition. Reston: AIAA, 2013: 798.

[24] CLIFF S E, THOMAS S D, MCMULLEN M S, et al. Assessment of unstructured Euler methods for sonic boom pressure signatures using grid refinement and domain rotation methods: TM-2008-214568[R]. Washington, D.C.: NASA, 2008.

[25] OZCER I, KANDIL O. Fun3D/OptiGRID coupling for unstructured grid adaptation for sonic boom problems[C]//Proceedings of the 46th AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2008: 61.

[26] VENDITTI D A, DARMOFAL D L. Grid adaptation for functional outputs: Application to two-dimensional inviscid flows[J]. Journal of Computational Physics, 2002, 176(1): 40-69.

[27] JONES W, NIELSEN E, PARK M. Validation of 3D adjoint based error estimation and mesh adaptation for sonic boom prediction[C]//Proceedings of the 44th AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2006: 1150.

[28] NEMEC M, AFTOSMIS M, WINTZER M. Adjoint-based adaptive mesh refinement for complex geometries[C]//Proceedings of the 46th AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2008: 725.

[29] WINTZER M, NEMEC M, AFTOSMIS M. Adjoint-based adaptive mesh refinement for sonic boom prediction[C]//Proceedings of the 26th AIAA Applied Aerodynamics Conference. Reston: AIAA, 2008: 6593.

[30] NEMEC M, RODRIGUEZ D L, AFTOSMIS M J. Adjoint-based mesh adaptation and shape optimization for simulations with propulsion[C]//Proceedings of the AIAA Aviation 2019 Forum. Reston: AIAA, 2019: 3488.

[31] XIAO T H, QIN N, LUO D M, et al. Deformable overset grid for multibody unsteady flow simulation[J]. AIAA Journal, 2016, 54(8): 2392-2406.

[32] HUNTON L W, HICKS R M, MENDOZA J P. Some effects of wing planform on sonic boom: TN D-7160[R]. Washington, D.C.: NASA, 1973.

[33] CELIK I, KARATEKIN O. Numerical experiments on application of Richardson extrapolation with nonuniform grids[J]. Journal of Fluids Engineering, 1997, 119(3): 584-590.

[34] ORDAZ I, WINTZER M, RALLABHANDI S K. Full-carpet design of a low-boom demonstrator concept[C]//Proceedings of the 33rd AIAA Applied Aerodynamics Conference. Reston: AIAA, 2015: 2261.

[35] WINTZER M, ORDAZ I. Under-track CFD-based shape optimization for a low-boom demonstrator concept[C]//Proceedings of the 33rd AIAA Applied Aerodynamics Conference. Reston: AIAA, 2015: 2260.

Adjoint-based adaptive Cartesian mesh refinement for sonic boom prediction

ZHU Zhenhao[1], XIAO Tianhang[1,*], XU Yanan[1], DENG Shuanghou[1], ZHANG Zhenming[2], YU Xiongqing[1]

(1. College of Aerospace Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;
2. Shanghai Electro-Mechanical Engineering Institute, Shanghai 201109, China)

Abstract: In view of the sonic boom prediction for supersonic transport, this paper develops a numerical method for supersonic near-field based on a finite volume scheme and an adjoint-based Cartesian adaptive mesh refinement for the governing equations. By solving the adjoint equation on an embedded-boundary Cartesian mesh, it can be determined that how sensitive the chosen output functional is to the residual of the discretized Euler equations, such as the surface integral of the related pressure signatures. Then, the computed adjoint variables render a correction term to improve the accuracy of the function on the coarse mesh and direct estimation of the remaining error to form an error estimate and a localized refinement parameter, according to which the Cartesian grid cells are locally refined subsequently. As a result, local mesh refinement steadily reduces each cell's remaining error, improving the output function's computational correctness. The 69° delta wing-body and NASA C25D configuration with powered nacelle are employed as validation cases where the computed results are compared with the documented experimental tests. Results show that compared with feature-based mesh refinement, the present adjoint-based refinement method captures the shock and expansion waves and the near-field pressure signatures with higher accuracy at a less computational cost, indicating its effectiveness and efficiency for sonic boom prediction.

Keywords: supersonic transport; sonic boom prediction; adjoint; error estimate; adaptive mesh refinement; Cartesian mesh

Received: 2021-11-16; **Accepted**: 2022-01-05; **Published Online**: 2022-01-25 11:22

URL: kns.cnki.net/kcms/detail/11.2625.V.20220124.1709.004.html

Foundation items: National Natural Science Foundation of China (11672133); National Science Foundation of Shanxi Province (2021JQ-078); Aeronautical Science Foundation of China (F2021110); Open Foundations of EDL Laboratory (EDL19092111)

* **Corresponding author**. E-mail: xthang@nuaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0693

基于多重检测的无人机抗遮挡目标跟踪算法

张博恒[1]，柴栋栋[2]，孟令博[3]，孙明健[1,3,*]

(1. 哈尔滨工业大学 航天学院，哈尔滨 150001；2. 北京航天试验技术研究所，北京 100074；
3. 哈尔滨工业大学 信息科学与工程学院，威海 264209)

摘 要：针对无人机（UAV）目标跟踪过程中遇到目标被障碍物遮挡时跟踪效果不佳的问题，提出一种多重检测的抗遮挡目标跟踪算法。在基于时空正则化相关滤波算法的框架下通过融合多种置信度函数，设计了一种响应置信度判别方法；为了具体了解目标被遮挡情况，将响应差值变化和响应梯度变化结合在一起作为判断是否更新滤波模板参数的依据；设计了一种融合分块思想与金字塔尺度池的尺度估计方法来解决目标在图像中尺度大小变化问题。所提算法在UAV数据集上相较于其他7种算法有不错的表现，在跟踪过程中面对目标遮挡、尺度变化和快速移动问题的跟踪精度和成功率上都有明显的提升。结果表明：所提算法能够更好地应对UAV在目标跟踪过程中出现的目标遮挡和尺度变化的问题，具有良好的快速性、准确性和鲁棒性。

关 键 词：无人机目标跟踪；相关滤波；抗遮挡；多重检测；尺度自适应
中图分类号：V279+.2; TP391
文献标志码：A **文章编号**：1001-5965(2023)09-2442-13

目标跟踪近些年来在视频监控、无人飞行器、无人驾驶、医学影像等领域的广泛应用使其成为计算机视觉方向研究的热点之一[1]。随着无人机行业的兴起，无人机与目标跟踪技术的结合使其应用更加广泛[2]。通常来说，无人机目标跟踪的本质就是确定目标在航拍视频的每一帧图像中的具体坐标位置及尺度信息，而后确定其运动轨迹，使无人机和目标保持一定距离进行跟踪。近些年来，研究人员在目标跟踪方向取得了显著进展，但是目标跟踪算法仍存在着很多挑战，例如外界光照变化、目标被遮挡及尺度变化等复杂场景下都没有得到很好的解决，难以实现实时跟踪的速度和长期跟踪的要求。

目标跟踪算法总体上可以分为基于生成式方法的跟踪和基于判别式方法的跟踪。随着技术的进步，主流的目标跟踪算法的实现已经由生成式变为判别式方法，其中基于相关滤波类目标跟踪算法发展尤为突出[3]。相关滤波类算法由于通过傅里叶变换在频域中进行计算，所以处理运行速度较快，满足无人机目标跟踪最基本的实时性的需要。首先，Bolme等[4]提出相关性的概念，并在目标跟踪算法中引入相关滤波的思想，提出最小输出误差平方和(minimum output sum of squared error, MOSSE)算法。然后，Henriques等[5]提出检测跟踪的循环结构(circulant structure of tracking-by- detection with kernels, CSK)算法，为了增加训练样本的可靠性，在相关滤波类算法的基础上增加循环矩阵的概念进而增加了训练样本的数量，与此同时，为了强调跟踪目标包含的特征，引入了核函数。接着，核相关滤波(kernel correlation filter, KCF)算法[6]则是利用梯度直方图(histogram of oriented gradients, HOG)特征并且融合多通道特征，融合了更多的特征就使得跟踪结果更加准确。但是这些算法在应对目标跟踪过程中出现的特殊情况时，表现就不尽如人意。随

收稿日期：2021-11-18；录用日期：2022-01-02；网络出版时间：2022-01-29 11:33
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220128.1757.007.html
*通信作者. E-mail：sunmingjian@hit.edu.cn

引用格式：张博恒，柴栋栋，孟令博，等. 基于多重检测的无人机抗遮挡目标跟踪算法[J]. 北京航空航天大学学报，2023，49(9)：2442-2454. ZHANG B H，CHAI D D，MENG L B，et al. Anti-occlusion target tracking algorithm of UAV based on multiple detection[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2442-2454 (in Chinese).

着人工智能技术的发展, 判别式目标跟踪方法有了新的研究方向即利用深度学习来完成目标跟踪。多域卷积神经网络 (multi-domain convolutional neural network, MDNet)[7] 算法是一种轻量级的网络模型, 利用学习得到图像的卷积特征来表示目标。Danelljan 等[8] 使用分解卷积算子, 很好地降低了卷积计算复杂度, 在不丢失精度的前提下提升了速度。Wang 等[9] 提出通过使用以树结构为基础的递归神经网络 (recurrent neural networks, RNN) 来学习层次特征用于视觉目标跟踪, RNN 比其他深度神经网络 (如卷积神经网络) 具有更少的参数; Jiang 等[10] 提出一种基于多智能体深度强化学习的多目标跟踪新算法, 以解决现有跟踪算法中目标数量变化、非因果性和非实时性等问题。但是, 深度学习算法过于依赖 GPU 等硬件设备, 在大多数单 CPU 的无人机平台上使用很难有良好的表现。因此, 算法计算复杂度相对较低, 运行速度和可靠性较好的基于相关滤波类目标跟踪算法, 在无人机平台上有着更好的应用。

在目标跟踪过程中, 由于目标移动的随机性和不确定性, 会存在许多的挑战, 比如: 目标遮挡、尺度变换、快速移动、形变等。其中比较困难的就是目标被遮挡且目标尺度发生变化的情况, 此时特征提取较难, 滤波模板很容易提取到错误的参数信息, 训练出来的跟踪器模型也会失真, 产生漂移问题[11]。

解决目标被遮挡问题的方法之一是确定合适的跟踪器更新策略。大多数跟踪器采用线性插值的方法对模板参数进行更新, 此方法容易产生漂移, 导致目标跟踪的失败。文献 [12] 中使用一种峰值旁瓣比作为判断跟踪器模型更新标准的方法, 这个方法考虑到了干扰的问题, 但没有考虑到目标的细节信息; 文献 [13] 提出将平均峰值相关能量的置信度作为判别标准, 此置信度可以用来衡量目标跟踪的鲁棒性, 但是也有其不足的地方。

为了解决无人机在目标跟踪过程中目标被遮挡和目标尺度发生变化产生的遮挡问题, 本文将时空正则化的相关滤波算法[14] 作为基本框架, 提出一种采用多重检测的无人机目标跟踪抗遮挡算法。首先, 结合响应差值变化和响应梯度变化定义一种响应置信度, 进而提出一种基于该置信度判别的多重检测方法, 增加跟踪器更新的可靠性; 其次, 设计一种融合分块思想与金字塔尺度池的尺度估计方法, 解决算法应对目标尺度变化问题; 最后, 本文算法在 UAV123[15]、OTB100[16] 和 DTB70[17] 无人机开源数据集上与其他主流算法进行对比实验, 在特殊场景下, 特别是目标被遮挡和目标尺度发生变化时的跟踪效果有了较好的提升。

1　算法原理与分析

1.1　基于时空正则化的相关滤波算法原理

相关滤波目标跟踪的本质是通过图像的当前帧所标定的目标, 结合前序帧的信息训练出模板滤波器, 从而确定跟踪目标再下一帧的准确位置。相关滤波目标跟踪算法的表达式为

$$h = x \otimes \omega \tag{1}$$

式中: h为输出的响应图; x为样本即输入的图像数据; ω为滤波模板; $\otimes$为卷积运算符号。

所以根据目标函数获得滤波模板就是相关滤波目标跟踪算法的核心。

当目标快速运动时, 目标可能会移动到图像的边缘处, 造成目标信息的缺失, 跟踪器就会错判或误判跟踪目标。为了解决类似问题, 抑制跟踪过程中模型的漂移, 基于时空正则化的相关滤波跟踪 (spatial-temporal regularized correlation filters, STRCF) 算法在原有目标函数岭回归模型上引入空间正则项, 用系数矩阵来起到抑制作用, 优化了岭回归函数模型; 同时, 引入时间正则项, 对前后帧获得的滤波模板进行正则化处理, 使得单个图像上的训练效果更加完美, 最终的目标函数为

$$\varepsilon(\omega,u) = \arg\min \left\| \sum_{k=1}^{K} \omega^k \otimes x^k - y \right\|_2^2 + \sum_{k=1}^{K} \left\| w \odot \omega^k \right\|_2^2 + \frac{u}{2} \sum_{k=1}^{K} \left\| \omega_t^k - \omega_{t-1}^k \right\|_2^2 \tag{2}$$

式中: k为目标特征通道数; ω^k为训练的第 k 个特征通道的滤波模板, 其特征数量与图像样本特征数量相同, 可以很好地表示目标的外观表征; y为目标期望响应; w为空间正则项系数, 通过对滤波模板进行正则化可以很好地训练出模板的模型; $\odot$为哈达玛积; u为一个超参数, 表示时间正则项系数因子; ω_t^k为第t帧图像训练的第k个特征通道的滤波模板。

本文算法中, 在式 (2) 基础上将空间正则项系数和时间正则项系数由固定值变为自适应系数, 空间正则项系数受响应图的局部响应变化而改变, 时间正则项系数因子受相邻图像全局响应变化而变化, 因此, 算法对于解决因快速移动产生的边界效应问题也有着一定提升作用。

将式(2)通过方向交替乘子(alternating direction method of multipliers, ADMM)法对滤波模板参数进行迭代求解, 首先, 可得拉格朗日方程为

$$\varepsilon(\omega,g,\alpha)=\arg\min\left\|\sum_{k=1}^{K}\omega^k\otimes x^k-y\right\|_2^2+\sum_{k=1}^{K}\left\|w\odot g^k\right\|_2^2+\sum_{k=1}^{K}\alpha^k(g^k-\omega^k)+\frac{\beta}{2}\sum_{k=1}^{K}\left\|g_t^k-\omega_t^k\right\|_2^2+\frac{u}{2}\sum_{k=1}^{K}\left\|\omega_t^k-\omega_t^{k-1}\right\|_2^2 \tag{3}$$

式中：α为拉格朗日乘数因子；β为步长参数；g为辅助变量；t为第t帧图像；ω_t^k为第t帧图像训练的第k个特征通道的滤波模板；g_t^k为第t帧图像训练的第k个特征通道的辅助变量。然后，对目标函数子问题进行迭代求解。

对滤波器模板ω_t最优求解表达式为

$$\omega_{t+1}=\arg\min\left\|\sum_{k=1}^{K}\omega_t^k\otimes x_t^k-y\right\|_2^2+\beta\|\omega_t-g_t+\alpha\|_2^2+\mu\|\omega_t-\omega_{t-1}\|_2^2 \tag{4}$$

式中：x_t^k为输入的第t帧第k个特征通道的图像。

辅助变量g_t最优求解表达式为

$$g_{t+1}=\arg\min\sum_{k=1}^{K}\left\|w\odot g_t^k\right\|_2^2+\beta\|\omega_t-g_t+\alpha\|_2^2 \tag{5}$$

子问题α_t最优求解表达式为

$$\alpha_{t+1}=\alpha_t+\omega_{t+1}-g_{t+1} \tag{6}$$

经过迭代求解，即可得到滤波器模板ω_t，完成目标跟踪的核心任务。滤波模板ω_t与当前帧图像x_t进行相关操作获得响应图的最大响应位置即为目标所在区域。

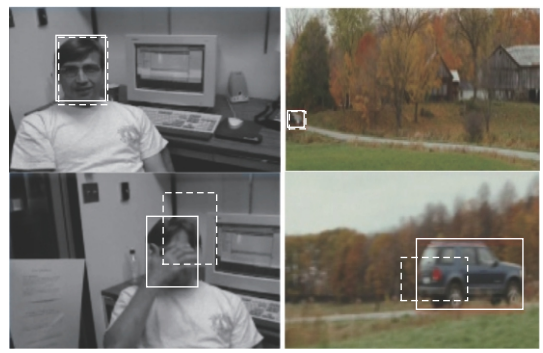
STRCF算法是一种在线训练的算法，通过将之前帧的目标区域作为训练集，训练得到当前帧的滤波器模板，再通过滤波器模板与当前帧进行相关操作得到目标所在的区域。因此，滤波器模板的更新就显得格外重要，滤波器模板的更新算法会对最终的跟踪效果产生很大的影响。常用的一种更新算法是线性插值法，这种算法采用固定频率来对模板进行更新，但是对目标遮挡和背景干扰特别敏感，容易产生漂移的现象；另一种是利用峰值旁瓣比（peak side lobe ratio, PSR）或平均峰值相关能量（avenue peak correlated energy, APCE）作为置信度，通过设定置信度阈值来判断是否对模板进行更新。这种算法考虑到图像的全局响应的变化，这样就可以更加准确地分辨出目标被遮挡的情况。

1.2 STRCF算法处理无人机目标跟踪遮挡问题的性能分析

无人机在进行目标跟踪时难免会遇到或多或少的目标遮挡及伴随的尺度变化的干扰。解决这种问题的方法之一就是选择合适的滤波器模板更新策略。

STRCF算法通常采用PSR或APCE置信度作为滤波模板的更新依据。而且由于算法自身的快速性所以可以满足无人机跟踪时的实时性要求。但是当目标被遮挡程度较大或是目标尺度发生较大变化时，就会导致滤波模板的更新失败，从而导致无人机目标跟踪的失败。

图1为STRCF算法在部分数据集中的跟踪效果，其中虚线框表示STRCF算法处理的跟踪结果，实线框表示目标的实际位置。可见当目标受到较大遮挡或目标尺度发生大幅度变化时，由于滤波器模型的更新判别标准始终保持一致，所以STRCF算法处理结果的误差就会增大，从而导致无人机目标跟踪的失败。



(a) 目标被遮挡 (b) 目标尺度变化

图1 STRCF算法在部分数据集中的跟踪效果

Fig. 1 Tracking results of STRCF algorithm in selected data sets

可见STRCF算法虽然具有非常优异的快速性和准确性，但是在解决无人机在目标被遮挡及目标尺度发生变化时的效果并不是很理想，所以需要设计出一种算法既能保证跟踪的快速性和准确性，也要能够处理目标被遮挡的情况完成无人机目标跟踪的任务。

2 抗遮挡目标跟踪算法

2.1 多重检测抗遮挡目标跟踪算法的总体框架

为了能够保留STRCF算法在目标跟踪时的快速性和准确性，还能够使得在跟踪过程中有效地克服目标被遮挡的挑战，本文在时空正则化目标跟踪算法的基础上进行改进，本文算法总体框架如图2所示。首先，对第1帧图像提取HOG特征和CN特征，选定感兴趣的目标区域；其次，通过使用循环位移矩阵扩充训练集，利用高斯响应对滤波响应进行训练，得到第1帧的滤波模板；接着，下一帧图像与滤波模板进行相关操作得到输出相关响应图，通过

尺度自适应检测获取目标的精确尺度, 进而输出目标位置; 最后, 为了得到精确的滤波器模板抑制漂移现象, 需要通过基于多重检测的模型更新策略来判断是否需要对滤波器模板参数进行更新。无人机处理系统通过循环完成这一任务实时得到目标的精确位置, 从而完成目标跟踪。

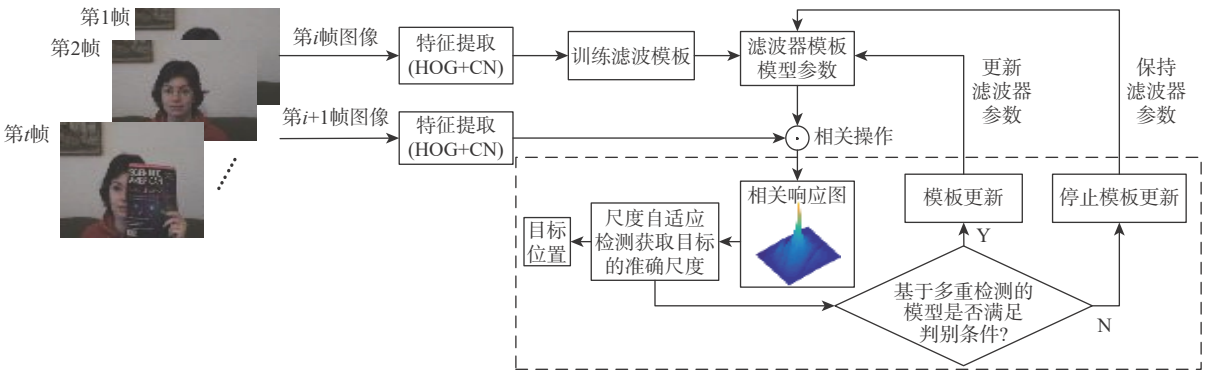


图 2　本文算法总体框架

Fig. 2　Overall framework of proposed algorithm

本文算法改进的内容有 2 部分(见图 2 中虚线部分): ①为了应对无人机跟踪时目标被遮挡的情况, 确定一种滤波器模型的更新策略。设计一种多重检测置信度标准, 用图像的响应变化和响应变化梯度作为判别目标受遮挡的情况, 定义一种响应置信度的判别标准, 通过设定阈值来决定滤波模板参数是否需要更新; ②为了解决目标尺度变化的问题, 在多尺度检测的基础上融合分块的思想, 进而优化尺度因子使得尺度池的范围能够更加准确的获得目标的尺寸, 这样既可以得到抗遮挡的效果同时也克服了目标尺度变化对目标跟踪造成的困难。

2.2　多重检测模型更新策略

常用的模型更新策略对目标受遮挡的情况有一定作用, 例如: STRCF 算法中采用 PSR 和 APCE 作为置信度, 来更新滤波器模板的指标参数, 但均是以单一置信度指标作为阈值, 进而判断跟踪目标受遮挡的情况来决定是否能够继续进行滤波器模型参数的更新。这种阈值设定的单一性往往不能对模型受遮挡的情况进行准确的判断, 使得滤波器模型进行错误的更新, 最终导致跟踪的失败。

根据 PSR 置信度和 APCE 置信度判别的核心思想, 提出一种多重检测置信度的设计, 定义一种响应置信度判别方法, 将响应变化梯度情况加入到判别标准中, 再根据实际的测试信息设定合适的阈值从而判定滤波模板参数是否需要更新。首先, 在 APCE 置信度基础上构造新的置信度判别式为

$$V_{\mathrm{q}}=\frac{\left|\sum_{i=1}^{n}\left(R_{\max}^{i}-R_{\min}^{i}\right)\right|^{2}}{n^{2}\sum_{w,h}\left(R_{w,h}-R_{\min}\right)^{2}/wh} \tag{7}$$

式中: V_{q}为当前所选区域中最大的响应与最小的响应之间的变化程度; $R_{\max}^{i}$为第i帧图像最大的响应值; $R_{\min}^{i}$为第i帧图像最小的响应值; n为选择的前n图像帧, 一般设置为 3。

接着, 定义图像响应变化梯度为

$$V_{\mathrm{DIFR}}^{t}=V_{\mathrm{q}}^{t}-V_{\mathrm{q}}^{t-1} \tag{8}$$

式中: V_{DIFR}^{t}为第t序号的图像响应变化梯度; V_{q}^{t}和V_{q}^{t-1}分别为第t帧和第$t-1$帧图像响应变化。同时考虑图像的目标和背景信息变化和其梯度的变化情况, 可以更加可靠地处理目标被遮挡问题。根据新的置信度标准结合分析经验响应变化来确定阈值, 其中响应变化阈值计算式为

$$V_{\mathrm{q}}^{\mathrm{thres}}=\frac{M_{1}^{a}}{10^{b}} \tag{9}$$

式中: $V_{\mathrm{q}}^{\mathrm{thres}}$为响应变化阈值; a和b为权重参数, 均设置为 30; M_1为连续图像帧之间的响应变化平均值, 计算式为

$$M_{1}=\frac{1}{m}\sum_{t=n-m+1}^{n}V_{\mathrm{q}}^{t} \tag{10}$$

式中: m为连续的图像个数; V_{q}^{t}为第t帧的响应变化。根据式 (7) 计算得到的响应变化值和式 (9) 计算的阈值进行比较判断跟踪是否遇到遮挡。

定义响应梯度阈值$V_{\mathrm{DIFR}}^{\mathrm{thres}}$为

$$V_{\mathrm{DIFR}}^{\mathrm{thres}}=(-1)^{c}M_{2}-d \tag{11}$$

式中: $V_{\mathrm{DIFR}}^{\mathrm{thres}}$为响应梯度阈值; c和d为权重参数, 分别设置为 3 和−20, d的参数设置可以检测到梯度的负向变化情况; M_2为最近连续的响应梯度变化值的平均值, 表达式为

$$M_{2}=\frac{1}{m}\sum_{t=n-m+1}^{n}V_{\mathrm{DIFR}}^{t} \tag{12}$$

当响应变化值V_{q}小于响应变化阈值$V_{\mathrm{q}}^{\mathrm{thres}}$时, 表

明被跟踪目标受到了遮挡干扰，继续判断当响应值变化梯度V_{DIFR}小于其阈值$V_{\mathrm{DIFR}}^{\mathrm{thres}}$时，确定目标收到了较大的遮挡干扰情况，模板滤波器模型参数停止更新，避免产生漂移。为了避免误判，再增加设定一个对比阈值防止模型始终停留在不更新阶段，当响应值梯度变化小于对比阈值时，应使跟踪器更新模型。对比阈值为

$$V_{\mathrm{comp}}^{\mathrm{thres}} = V_{\mathrm{DIFR}}^{\mathrm{thres}} - f \tag{13}$$

式中：$V_{\mathrm{comp}}^{\mathrm{thres}}$为对比阈值；$f$为模型不更新时的梯度可减小的最大范围值，设置为20，可以根据实际进行调整。

由此可以得到模型更新的判断条件为

$$\begin{cases} V_{\mathrm{q}} > V_{\mathrm{q}}^{\mathrm{thres}} \\ V_{\mathrm{DIRF}} > V_{\mathrm{DIRF}}^{\mathrm{thres}} \end{cases} \tag{14}$$

$$\begin{cases} V_{\mathrm{q}} > V_{\mathrm{q}}^{\mathrm{thres}} \\ V_{\mathrm{DIRF}} < V_{\mathrm{comp}}^{\mathrm{thres}} \end{cases} \tag{15}$$

当响应变化值V_{q}小于响应变化阈值$V_{\mathrm{q}}^{\mathrm{thres}}$，且响应变化梯度$V_{\mathrm{DIFR}}$小于响应变化梯度阈值$V_{\mathrm{DIFR}}^{\mathrm{thres}}$时，或当响应变化值$V_{\mathrm{q}}$小于响应变化阈值$V_{\mathrm{q}}^{\mathrm{thres}}$，且响应变化梯度$V_{\mathrm{DIFR}}$小于对比阈值$V_{\mathrm{comp}}^{\mathrm{thres}}$时，对跟踪滤波器模型参数进行更新。否则不更新，维持之前帧的滤波器模型参数。

在完成对无人机目标跟踪抗遮挡算法改进之后，需要解决无人机目标跟踪时的尺度变化问题，所谓尺度变化问题就是当目标距离机载相机的距离发生变化时，目标成像的尺度也会发生变化，进而会对目标的具体位置标定和滤波模板的更新产生影响。因此，尺度变化问题也是无人机实现目标跟踪任务的挑战之一。

2.3 尺度自适应检测方法

尺度变化在目标跟踪过程中是一个较为困难的挑战，当目标尺寸大小发生变化或遭遇遮挡时，跟踪算法的感兴趣区域无法完全包含目标，这会使后续的目标信息提取不完整，降低了训练的跟踪器模型精度。

解决目标跟踪时的尺度变化问题的重要手段是使用尺度估计的方法，常用的方法有金字塔估计法[18]。金字塔模型的尺度检测首先确定目标的中心位置，然后在进行计算，金字塔的尺寸以$M\times N$大小的图像块为基准，M、N分别为目标区域框的宽和高，使用S个尺度数量对基准图像进行缩放，σ为尺度缩放因子，从而获得一个$\sigma M\times\sigma N\times S$的金字塔结构的尺度池，在经过滤波器相关处理后，获取不同的响应图，峰值最大的响应图对应的尺度特征图即为所求的尺度特征，确定下一时刻图像中目标尺寸的数值，使目标跟踪算法更加精确，且能适应目标的尺度变化。其中尺度滤波器的尺寸固定为$M\times N$，且数量大小同样为S，这样的目的是为了保证求出的响应图具有相同的结构。σ具体可表示为σ^n，其中n的定义为

$$n \in \left\{ \left[-\frac{S-1}{2} \right], -\frac{S-2}{2}, \cdots, \left[\frac{S-1}{2} \right] \right\} \tag{16}$$

这种金字塔尺度估计方法也有自身的缺点，这是因为尺度个数和缩放因子的取值往往不会很准确，所以会导致尺度估计的偏差。因此，可以采取将图像分块与多尺度检测相结合设计出新的尺度自适应检测方法，从而可以对尺度因子进行优化更加逼近于真实值。算法步骤是：①采用分块思想将图像分为9等份；②把图像的中心目标块作为起点，相连的4个分块形成4个子块，分别对这4个子块进行相关操作确定峰值位置；③通过4个峰值位置确定目标的尺寸。就完成了尺度估计的任务。

其中4个子块包含了目标的主要信息，分别计算4个子块的响应峰值位置，用P_1、P_2、P_3、P_4表示，进而通过这4个位置确定目标的尺寸大小，计算式为

$$S_{\mathrm{o}} = |P_2 - P_1| \times |P_3 - P_4| \tag{17}$$

式中：S_{o}为分块化的目标尺寸计算；定义分块目标跟踪的尺寸因子S_1为

$$S_1 = \frac{S_{\mathrm{o}}^i}{S_{\mathrm{o}}^{i-1}} \tag{18}$$

式中：S_{o}^i为第i帧的目标尺寸。通过前后2帧之间的尺度变化求出的分块尺度因子，再与尺度池的尺度因子σ^n进行加权融合，获得最终的尺度因子：

$$S_2 = (1-\theta)\sigma^n + \theta S_1 \tag{19}$$

式中：S_2为自适应尺度因子；θ为加权系数，初始值设置为0.08。改进的尺度因子优化了尺度池的范围，能够更加准确地获取目标尺寸。

本文利用分块和尺度估计设计的尺度自适应检测方法就可以准确的获得跟踪目标的尺度大小，有效地克服了无人机在目标跟踪时遇到的目标尺度变化的挑战。

2.4 本文算法流程

图3为本文算法的流程，首先，通过无人机平台上携带的摄像机获取目标及背景的图像，而后通过本文算法可以同时进行目标位置的输出和目标尺度的估计，并对滤波模型更新策略进行优化，可以克服目标被遮挡和目标尺度发生变化的困难，从而获取目标的准确位置区域，完成无人机目标跟踪任务。

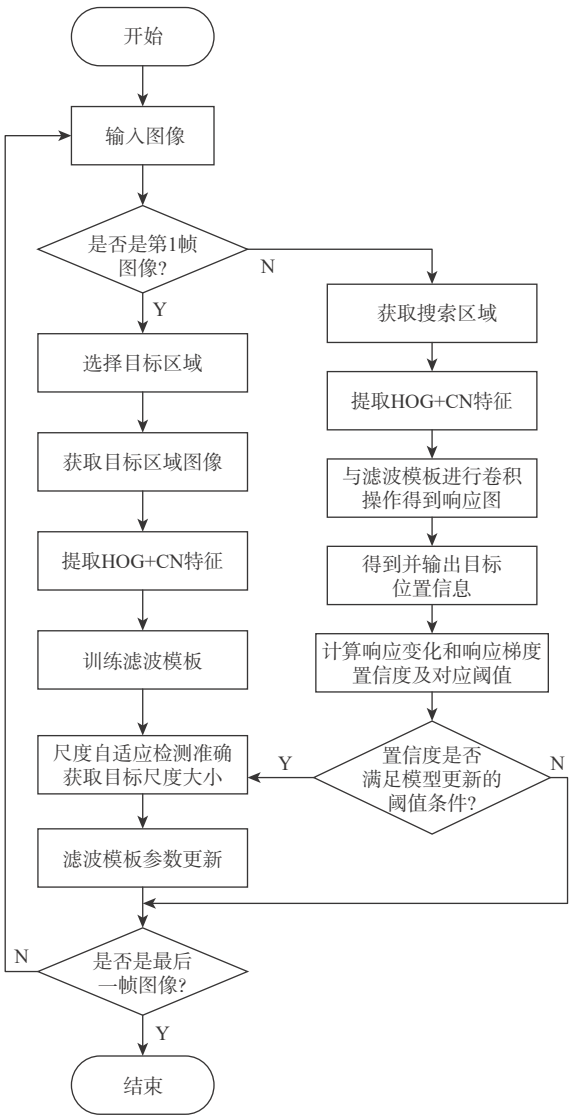


图 3　本文算法流程

Fig. 3　Flow chart of proposed algorithm in this article

3　实验结果及分析

3.1　实验环境、参数设置及评价指标

本文程序采用 C++语言, 实验运行环境为 MATLAB 2018b, Windows10 64 位操作系统, 平台的处理器为 i7-5500U CPU@2.4GHz。

仿真测试选用 UAV123、OTB100 和 DTB70 开源图像数据集进行无人机目标跟踪性能的实验。UAV123、OTB100 和 DTB70 这 3 个数据集均是当前无人机目标跟踪算法主流的开源测试数据集, 具有较高的权威性和通用性。UAV123 是无人机航拍数据集, 跟踪对象涉及类型很广泛。同时 UAV123 数据集共标注了 12 种挑战任务, 依次为纵横比变化、背景干扰、尺度变化、完全遮挡、部分遮挡、相机运动、相似目标、低分辨率、光照变化、视角变化、脱离视野和图像内旋转。多种挑战使得检验无人机目标跟踪算法的优劣更有说服力; OTB100 数据集共有 98 个视频序列包含 100 个测试场景, 定义了 11 个挑战任务类型, 包括光照变化、比例变化、快速运动等; DTB70 由无人机航拍视频构成, 包括 70 个图像序列, 同时也有 11 个挑战任务。

因为多重检测模型更新策略中需要对置信度和阈值设定的参数进行设置, 所以我们设定的是基于前 3 张图像的响应值信息来对设计的置信度进行求取, 并且由a和b来确定响应变化阈值, c和d来确定响应变化梯度阈值。经过实际测试, 取n为 3, a和b均为 30, c和d设定分别为 3 和−20。同时, 对于尺度自适应检测方法, 选择的初始权重θ的值为 0.08。其他设置参数保持不变。

测试结果的评价指标是精确度和成功率。精确度评价指标用来衡量无人机在目标跟踪时的准确性, 成功率评价指标用来衡量无人机在目标跟踪时的可靠性。其中精确度表示算法计算出的跟踪目标位置的中心坐标与目标真实位置的中心坐标的距离小于设定值的视频帧数的占比。通常将平均精确度作为评价指标; 成功率则是表示跟踪成功的帧数所占百分比, 目标跟踪的结果区域和与目标所在真实区域的相交程度, 相交程度即为重叠率, 若重叠率大于设定的重叠率阈值则表示目标跟踪成功。

3.2　实验结果分析

在 UAV123、OTB100 和 DTB70 开源数据集上对本文算法进行测试实验, 其中, 本文算法中的各个参数和各个阈值设定值如 3.1 节所述。测试结果与目前比较流行的目标跟踪算法: STRCF、尺度递归相关滤波器(scale-recurrent correlation filter, SRDCF)[19]、CSK[20]、KCF、深度结构化自相关跟踪(deep structured-self correlation tracking, DSST)[18]、尺度自适应均值漂移(scale adaptive mean-shift, SAMF)[21] 和 AutoTrack[22] 进行对比。

8 种算法在 UAV123 无人机数据集上应对快速移动、遮挡、尺度变化和总体性能的精度曲线和成功率曲线如图 4 所示, 表 1 为 8 种算法在 UAV123 无人机数据集上的具体指标数据对比。由图 4 和表 1 可知, 本文算法在 3 种挑战和总体的性能上都有良好的表现, 本文算法的精度达到 62.2%, 成功率达到 60.8%, 都明显高于 STRCF 算法, 分别提升了 3.1% 和 7.9%。与 2020 年的主流算法 AutoTrack 相比, 在目标遮挡和尺度变化挑战中的精度和成功率也有显著的提高。由此, 可以看出本文算法在无人机数据集 UAV123 上无论精度还是成功率都达到最优或次优的效果。在 UAV123航拍数据集上的出

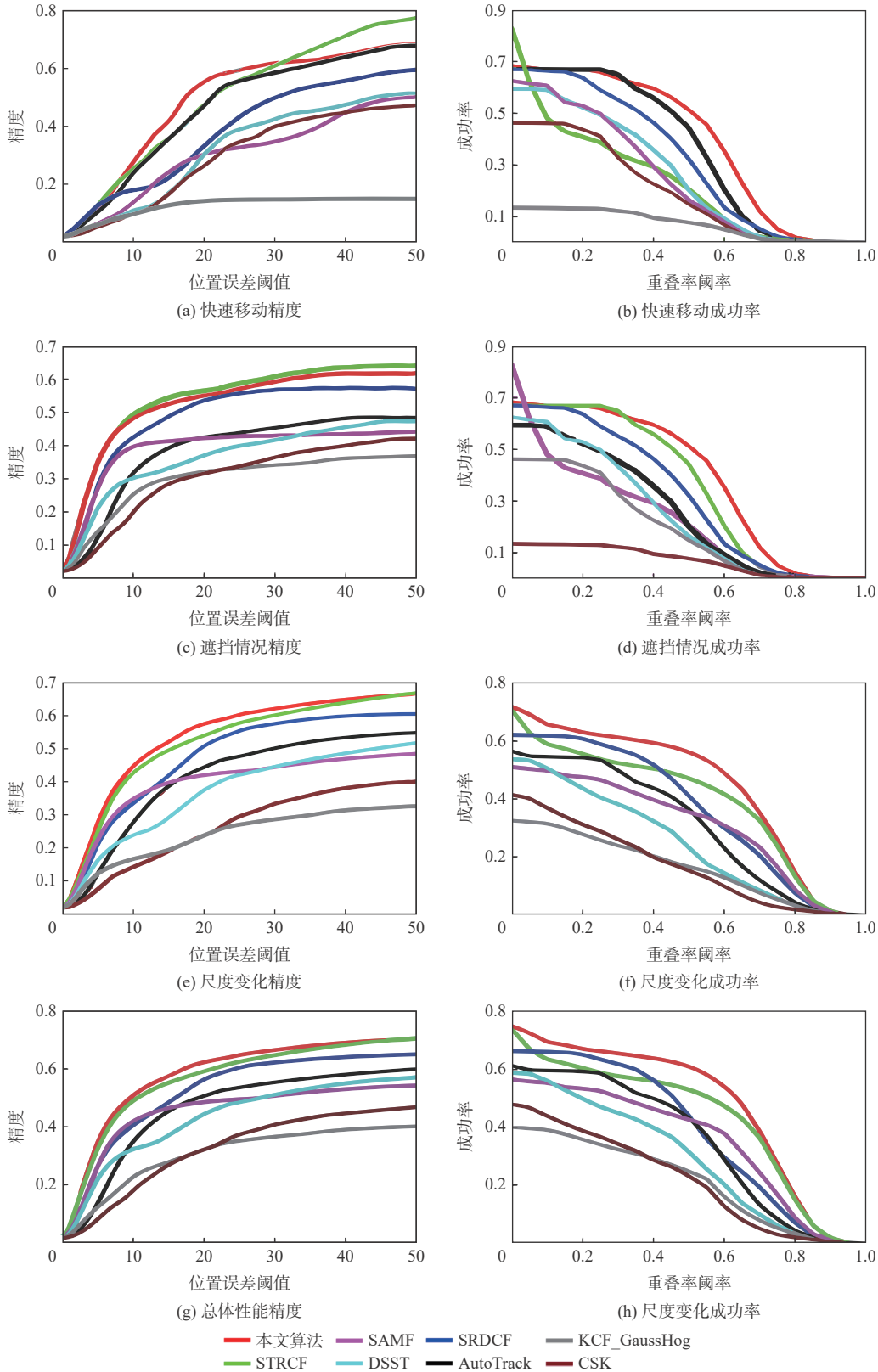


图 4 8 种算法在 UAV123 上的精度和成功率

Fig. 4 Accuracy and success rate of 8 algorithms on UAV123 databases

色表现验证了本文算法在实际无人机系统上完成目标跟踪抗遮挡与应对目标尺度变化挑战时具有优良的性能。

8 种算法在 DTB70 航拍数据集上应对 3 种挑战和总体性能的成功率和精度曲线如图 5 所示，表 2 为 8 种算法具体的评价指标数据。其中应对目标遮挡挑战，本文算法精度提高了 0.7%，成功率提高了 1.2%；应对尺度变化挑战，精度提高了 3.9%，成功率提高了 4.3%；总体性能表现上，本文算法较最优秀的算法精度提高了 6.6%，成功率提高了 3.7%。

表 1　8 种算法在 UAV123 数据集上的性能对比

Table 1　Performance comparison of 8 algorithms on UAV123 dataset　%

算法	精度				成功率			
	快速运动	遮挡	尺度变化	总体性能	快速运动	遮挡	尺度变化	总体性能
CSK	26.2	31.6	23.8	32.1	14.6	22.8	15.2	23.2
KCF	14.1	32.2	23.8	32.2	7.6	26.6	16.5	24.8
DSST	30.1	37.1	37.6	44.5	20.0	28.9	23.1	31.6
SAMF	30.3	42.2	42.0	48.5	16.4	37.5	35.3	42.5
SRDCF	33.0	53.6	50.8	56.3	32.1	39.4	40.9	44.0
STRCF	47.3	56.5	54.0	59.1	20.4	51.2	47.0	52.9
AutoTrack	46.9	42.5	44.4	50.6	44.1	32.3	36.4	43.3
本文算法	55.4	55.0	57.4	62.2	51.1	53.1	55.9	60.8

从上述指标可以看出, DTB70 数据集相较于 UAV123 数据集其图像序列场景较简单一些, 所以本文算法在整体指标上都有所提高。在快速运动和遮挡问题上, 本文算法相较于其他的相关滤波算法仍然具有很高的精度和成功率, 验证了本文算法的有效性。

在开源数据集 OTB100 上进行测试实验, 图 6 为 8 种算法在 OTB100 数据集上精度与成功率对比。表 3 为 8 种算法在该数据集上的评价指标数据。在目标遮挡挑战中, 本文算法较 AutoTrack 跟踪器精度提高了 0.5%, 成功率提高了 1%, 在目标尺度变化挑战中, 精度提高了 3.5%, 成功率提高了 4.1%。总体性能评价指标中, 较其他表现最优秀的算法准度提高了 2.8%, 成功率提高了 7.5%。本文算法在 OTB100 数据集上的突出表现表明本文算法在处理其他场合的目标跟踪问题也有良好的性能, 体现了本文算法的广泛性。

为了更直观的展现本文算法在应对无人机目标跟踪测试中的优异表现, 选取 OTB100 数据集的 Persons 图像序列中的部分关键图像帧的跟踪结果, 目标遮挡挑战中各算法跟踪效果如图 7 所示, 红色框代表本文算法确定的目标区域, 其他颜色框分别代表对比的 7 种算法确定的结果。视频图像包括了部分遮挡和完全遮挡的挑战。在开始阶段, 目标周围没有障碍物没有发生遮挡现象, 所有算法均能够成功跟踪目标; 第 65 帧～第 75 帧, 目标发生了部分遮挡, CSK、SRDCF、DSST、AutoTrack 和 STRCF 算法都有了不同程度的跟踪误差; 第 80 帧～第 85 帧, 目标发生了完全遮挡, 此时 SRDCF 和 DSST 算法完全跟踪失败; 第 130 帧～第 140 帧, 目标脱离遮挡后只有本文算法和 SAMF 算法依旧跟踪相对成功, 但是 SAMF 相较于本文算法也存在一定的尺度误差。因此, 本文算法采用多重检测的模板更新策略, 当出现较严重的目标遮挡的情况即响应变化和响应变化梯度较大时, 及时地停止了对滤波模板的更新, 依旧使用前一帧滤波模板来确定目标位置, 防止其进行错误的更新而使得跟踪的失败。

当面对比较严重的遮挡和尺度变化的挑战时, 算法在 OTB100 数据集的 CarScale 图像序列上部分关键图像帧的跟踪结果, 目标尺度变化和遮挡挑战中各算法跟踪效果如图 8 所示, 红色框代表本文算法确定的目标区域位置, 其他颜色框分别代表对比的 7 种算法。视频图像包括了尺度变化和遮挡的挑战。在开始阶段, 目标周围没有障碍物没有发生尺度变化和遮挡现象, 所有算法均能够成功跟踪目标; 第 70 帧～第 140 帧, 目标发生了尺度变化, CSK、SRDCF、SAMF 和 DSST 算法都没有完全跟踪汽车的尺度变化, 有了不同程度的跟踪误差; 第 150 帧～第 180 帧, 目标发生了遮挡, 此时 CSK 算法完全跟踪失败, 其他 6 种算法还能够跟踪目标, 但存在尺度上的误差; 第 180 帧～第 200 帧, 目标脱离遮挡后本文算法和 KCF 算法跟踪效果相对成功, 但是本文算法处理尺度变化问题效果更好, 精度更高。因此, 本文算法采用尺度自适应检测在面对遮挡和尺度变化综合挑战时能够有较好的表现, 通过图像分块和尺度缩放的思想更为准确的确定跟踪目标的尺度, 有效的提高了跟踪算法在应对尺度变化和遮挡的问题的鲁棒性和准确性。

为了进一步验证本文算法在高空中对于地面目标跟踪效果, 本文继续对无人机在高空俯视条件下对道路上受到遮挡的目标进行跟踪的实际效果进行验证。图 9 为在高空中跟踪地面被遮挡的目标测试效果, 其中红色框表示算法标识出来的跟踪框。对图像帧进行分析: 视频前 700 帧目标被成功跟踪; 在第 700 帧图像目标发生了较大的尺度变化挑战, 本文算法根据尺度自适应检测, 重新确定跟踪目标的尺度大小, 依旧可以保持良好的跟踪效

果；第1 300帧～第1 400帧图像出现了目标被遮挡的现象，根据多重检测的模型更新策略，在第1 368帧时，检测判定目标被完全遮挡，这时停止对滤波模板的更新保持之前帧的模板，在之后检测判定目标未被遮挡时重新更新滤波模板就如第1 380帧所示；1 380帧之后目标又被很好的跟踪。从结果可以看出，本文算法可以较好得应对无人机在高空追踪地面受遮挡和发生尺度变化的目标的情况。

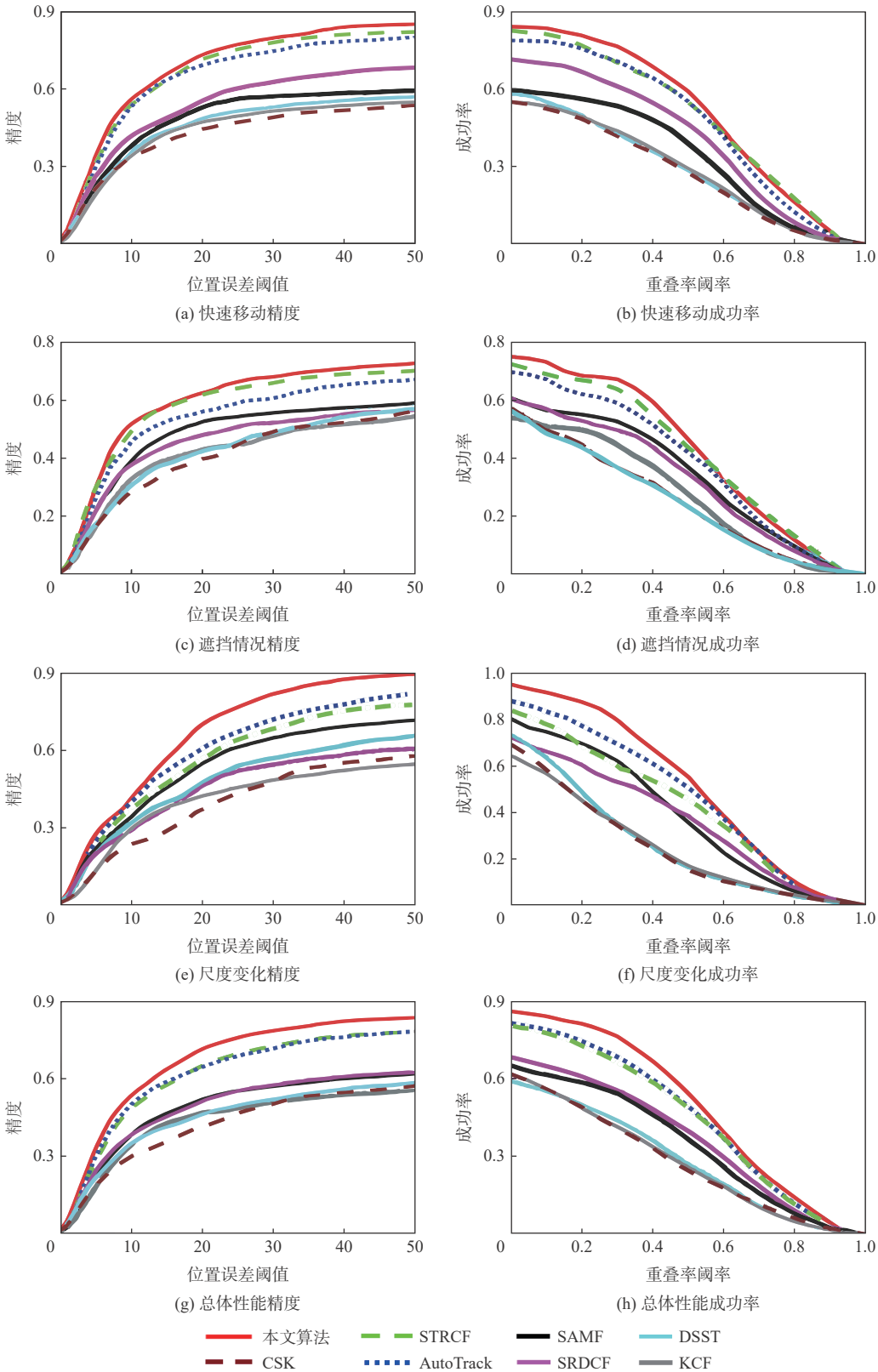


图5 8种算法在DTB70上的精度和成功率

Fig. 5 Accuracy and success rate of 8 algorithms on DTB70 databases

表 2　8 种算法在 DTB70 数据集上的性能对比

Table 2　Performance comparison of 8 algorithms on DTB70 dataset　%

算法	精度				成功率			
	快速运动	遮挡	尺度变化	总体性能	快速运动	遮挡	尺度变化	总体性能
CSK	44.3	39.7	37.0	41.5	27.2	24.4	23.9	27.2
KCF	47.0	42.9	42.2	46.3	28.0	27.0	24.0	28.0
DSST	48.5	42.3	47.3	46.3	28.1	24.4	25.5	27.6
SAMF	52.8	52.5	54.9	52.0	33.4	32.5	37.3	33.9
SRDCF	55.4	47.8	46.2	51.2	39.8	31.0	35.9	36.3
STRCF	71.3	61.7	56.8	64.9	46.7	40.0	41.7	43.7
AutoTrack	69.1	55.9	60.6	64.5	45.7	37.2	45.9	44.2
本文算法	73.1	62.4	70.0	71.5	49.0	41.2	50.2	47.9

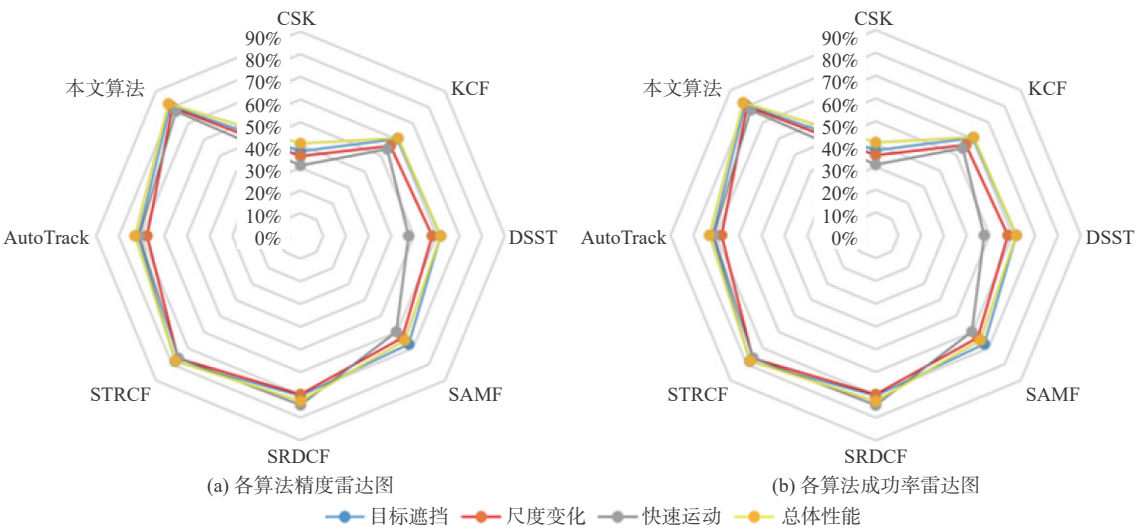


图 6　8 种算法在 OTB100 数据集上精度与成功率对比

Fig. 6　Comparison of accuracy and success rate of eight algorithms on OTB100 dataset

表 3　8 种算法在 OTB100 数据集上的性能对比

Table 3　Performance comparison of 8 algorithms on OTB100 dataset　%

算法	精度				成功率			
	快速运动	遮挡	尺度变化	总体性能	快速运动	遮挡	尺度变化	总体性能
CSK	31	37.2	35.1	40.7	29.9	27.4	27	33.6
KCF	54	60.4	56	60.8	43.6	44.2	38.2	46.3
DSST	47.6	61.7	58	61.7	38.8	42.6	39.4	46.8
SAMF	59.7	67.4	63.4	65	51.1	56.6	52	55
SRDCF	74.5	70.4	69.8	72.8	68.9	61.9	62.7	65.8
STRCF	76.1	78.1	76.3	77.9	70.8	69.9	67.7	70.4
AutoTrack	71	71.5	67.6	72.9	65.1	62.5	57.9	64.9
本文算法	77.6	81	79.8	82.2	71.1	70.7	71.8	73.8

图 7　目标遮挡挑战中各算法跟踪效果

Fig. 7　Tracking results of each algorithm in target occlusion challenge

图 8 目标尺度变化和遮挡挑战中各算法跟踪效果

Fig. 8 Tracking results of each algorithm in occlusion challenge of target scale change

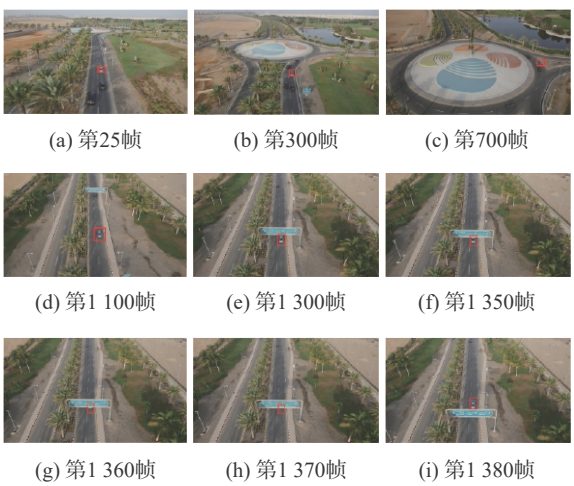


图 9 在高空中跟踪地面被遮挡的目标测试效果

Fig. 9 Test results of tracking the obscured target on ground at high altitude

本文算法在时空正则化相关滤波算法的基础上增加了多重检测的模型更新策略和尺度自适应检测等步骤，增加了计算复杂度，而本文算法采用自适应矩估计(adaptive moment estimation, ADAM)算法迭代求解的方法，有效地提升了计算速度。为了验证本文算法对于跟踪实时性的性能，本文在3种数据集上对8种算法的实时性性能进行测试，得到每种算法的平均实时性能如表4所示。由此可见，较为简单的CSK和KCF算法的实时性能最好，由于DSST和SAMF算法复杂度较大，所以实时性能较差。本文算法在3种数据集上的平均实时性能可以达到93帧/s，在特定数据集上可以达到100帧/s，远远可以满足无人机目标跟踪最低限度30帧/s的要求。所以，尽管本文算法在计算复杂度上有所增加，但是依然可以满足无人机目标跟踪的实时性要求。

表 4 各个算法在3种数据集上的平均实时性能

Table 4 Average real-time performance of each algorithm on three kinds of datasets

算法	实时性能/(帧·s^{-1})
CSK	118
KCF	112
DSST	32
SAMF	26
SRDCF	88
STRCF	101
AutoTrack	110
本文算法	93

4 结 论

1) 本文算法在i7-5500U, 2.4 GHz的处理器上进行测试，测试的处理速度能够达到100 FPS，因此，本文算法能够应用到无人机的机载电脑上，其快速性可以保证无人机对目标的实时跟踪。

2) 在数据集UAV123、DTB70和OTB100上进行仿真测试，本文算法相较于其他7种主流算法在处理跟踪过程中遇到的目标遮挡、目标尺度变化等挑战时的跟踪性能均有一定的提升，本文算法的快速性和准确性都很优异，同时具备较好的鲁棒性。在高空俯视条件下对地面目标进行跟踪测试中也体现了很好的抗遮挡效果，因此，本文算法在无人机目标跟踪上具有良好的应用前景。

参考文献（References）

[1] BONATTI R, HO C, WANG W S, et al. Towards a robust aerial cinematography platform: Localizing and tracking moving targets in unstructured environments[C]// 2019 IEEE/RSJ International Con-

ference on Intelligent Robots and Systems. Piscataway: IEEE Press, 2020: 229-236.

[2] LI R, PANG M J, ZHAO C, et al. Monocular long-term target following on UAVs[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE Press, 2016: 29-37.

[3] LUKEŽIC A, VOJÍR T, ZAJC L C, et al. Discriminative correlation filter with channel and spatial reliability[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 4847-4856.

[4] BOLME D S, BEVERIDGE J R, DRAPER B A, et al. Visual object tracking using adaptive correlation filters[C]// 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2010: 2544-2550.

[5] HENRIQUES J F, CASEIRO R, MARTINS P, et al. Exploiting the circulant structure of tracking-by-detection with kernels[C]// European Conference on Computer Vision. Berlin: Springer, 2012: 702-715.

[6] HENRIQUES J F, CASEIRO R, MARTINS P, et al. High-speed tracking with kernelized correlation filters[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3): 583-596.

[7] NAM H, HAN B. Learning multi-domain convolutional neural networks for visual tracking[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 4293-4302.

[8] DANELLJAN M, BHAT G, KHAN F S, et al. ECO: Efficient convolution operators for tracking[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 6931-6939.

[9] WANG L, LIU T, WANG B, et al. Learning hierarchical features for visual object tracking with recursive neural networks[C]// 2019 IEEE International Conference on Image Processing. Piscataway: IEEE Press, 2019: 3088-3092.

[10] JIANG M X, HAI T, PAN Z G, et al. Multi-agent deep reinforcement learning for multi-object tracker[J]. IEEE Access, 2019, 7: 32400-32407.

[11] 潘振福, 朱永利. 使用PSR重检测改进的核相关目标跟踪方法[J]. 计算机工程与应用, 2017, 53(12): 196-202.
PAN Z F, ZHU Y L. Improvement for tracker with kernelized correlation filters with PSR redetection[J]. Computer Engineering and Applications, 2017, 53(12): 196-202(in Chinese).

[12] 侯志强, 韩崇昭. 视觉跟踪技术综述[J]. 自动化学报, 2006, 32(4): 603-617.
HOU Z Q, HAN C Z. A survey of visual tracking[J]. Acta Automatica Sinica, 2006, 32(4): 603-617(in Chinese).

[13] MA H Y, ACTON S T, LIN Z L. SITUP: Scale invariant tracking using average peak-to-correlation energy[J]. IEEE Transactions on Image Processing, 2020, 29: 3546-3557.

[14] LI F, TIAN C, ZUO W M, et al. Learning spatial-temporal regularized correlation filters for visual tracking[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 4904-4913.

[15] UZKENT B, SEO Y. EnKCF: Ensemble of kernelized correlation filters for high-speed object tracking[C]// 2018 IEEE Winter Conference on Applications of Computer Vision. Piscataway: IEEE Press, 2018: 1133-1141.

[16] LIU S, WANG S, LIU X Y, et al. Fuzzy detection aided real-time and robust visual tracking under complex environments[C]// IEEE Transactions on Fuzzy Systems. Piscataway: IEEE Press, 2020: 90-102.

[17] HUANG Z Y, FU C H, LI Y M, et al. Learning aberrance repressed correlation filters for real-time UAV tracking[C]// 2019 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2020: 2891-2900.

[18] DANELLJAN M, HÄGER G, SHAHBAZ KHAN F, et al. Accurate scale estimation for robust visual tracking[C]// Proceedings of the British Machine Vision Conference 2014. Nottingham: British Machine Vision Association, 2014: 1064-1075.

[19] DANELLJAN M, HÄGER G, KHAN F S, et al. Learning spatially regularized correlation filters for visual tracking[C]// 2015 IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2016: 4310-4318.

[20] HENRIQUES J F, CASEIRO R, MARTINS P, et al. Exploiting the circulant structure of tracking-by-detection with kernels[C]//Proceedings of the 12th European conference on Computer Vision—Volume Part IV. New York: ACM, 2012: 702-715.

[21] LI Y, ZHU J K. A scale adaptive kernel correlation filter tracker with feature integration[C]//European Conference on Computer Vision. Berlin: Springer, 2015: 254-265.

[22] LI Y M, FU C H, DING F Q, et al. AutoTrack: Towards high-performance visual tracking for UAV with automatic spatio-temporal regularization[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 11923-11932.

Anti-occlusion target tracking algorithm of UAV based on multiple detection

ZHANG Boheng[1]，CHAI Dongdong[2]，MENG Lingbo[3]，SUN Mingjian[1, 3, *]

(1. College of Astronautics，Harbin Institute of Technology，Harbin 150001，China；
2. Beijing Institute of Space Test Technology，Beijing 100074，China；
3. School of Information Science and Engineering，Harbin Institute of Technology，Weihai 264209，China)

Abstract：In this research, a multi-detection anti-occlusion target tracking method is suggested to address the issue of poor tracking effect when unmanned aerial vehicle (UAV) targets are occluded by obstacles during the target tracking process. A response confidence discrimination method is designed by fusing various confidence functions under the framework of a spatiotemporal regularization correlation filtering algorithm. In order to understand the occlusion situation, the change of response difference and response gradient are combined together as the basis to judge whether to update the filter template parameters. A scale estimation method combining the block idea and pyramid scale pool is designed to solve the problem of the scale variation of the target in the image. Compared with the other seven algorithms, the proposed algorithm performs well in the UAV data set, and significantly improves the tracking accuracy and success rate in the face of target occlusion, scale change, and fast movement problems in the tracking process. The findings demonstrate that the multi-detection-based anti-occlusion target tracking algorithm has good speed, accuracy, and robustness and can more effectively address the issues of target occlusion and dimension change in the process of UAV target tracking.

Keywords：unmanned aerial vehicle target tracking；correlation filtering；anti-occlusion；multiple detection；scale adaptation

Received：2021-11-18；**Accepted**：2022-01-02；**Published Online**：2022-01-29 11：33
URL：kns.cnki.net/kcms/detail/11.2625.V.20220128.1757.007.html
* **Corresponding author**. E-mail：sunmingjian@hit.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0647

中国民用航空器 CO_2 减排潜力的区域划分

曾雯，胡荣*，宋文，刘志昊，张军峰
(南京航空航天大学 民航学院，南京 211106)

摘　　要：对中国进行科学的区域划分，是研究具有区域差异性与空间相关性航空器 CO_2 排放问题并提出差异化 CO_2 减排措施的重要基础。基于2007—2016年中国部分地区民用航空器运行数据，在构建 CO_2 减排潜力分类模型的基础上，通过Kruskal算法得到省域网络关系最小生成树；采用谱聚类算法，以“最大化区划优度”为目标，划分航空器 CO_2 减排潜力的不同区域。结果表明：最优区域划分为四分区，区划优度为0.35；航空器 CO_2 减排潜力高的区域主要分布于中国西南和中南地区，潜力低的区域主要分布于东北和华北地区；针对各区域 CO_2 减排特征，从民航局、机场和航空公司等不同主体提出了差异化的区域 CO_2 减排措施。

关 键 词：民用航空器；CO_2 减排潜力；区域划分；最小生成树；谱聚类
中图分类号：U8
文献标志码：A　　**文章编号**：1001-5965(2023)09-2455-08

CO_2 的过量排放是造成全球气候变暖问题的主要原因。2018年，交通运输部门排放的 CO_2 占全球总排放量的24.64%[1]，是能源需求和 CO_2 排放的主要贡献者。航空运输一方面具有高 CO_2 排放强度的特点，约为私家车的2倍或高铁的4倍[1-2]；另一方面，又具有远距离运输、空间跨度大的特点，所涉省市的地理位置、社会经济水平等具有不同特点，这就造成了航空器 CO_2 排放分布的区域差异性与空间相关性并存的现象[1,3]。因此，为制定富有针对性的区域 CO_2 减排目标与措施，势必需要开展兼顾 CO_2 排放区域差异性与空间相关性的区域划分研究。

当前涉及 CO_2 排放区域划分的研究，大致可分为2类：①直接使用国家相关规划中的区域划分结果。例如：东-中-西部、六大行政区域和八大经济区域等区域划分。不少学者选择其中一种划分结果对 CO_2 排放的区域差异性、空间演化特征和收敛性等开展研究[4-6]。但以上区域划分未充分考虑区域间的 CO_2 排放差异，制约了研究成果的针对性。②基于 CO_2 排放相关指标，采用定性或定量方法进行区域划分。常见的指标有年均 CO_2 排放量、CO_2 排放密度、CO_2 排放强度、能源强度、能耗综合指标、人均GDP等。常用的方法包括：依据指标水平定性划分、主成分分析、聚类方法等[7-12]。但文献[7-12]主要以截面数据为主，侧重依据指标水平开展区域划分，忽视了地理空间的邻接关系。

航空器 CO_2 排放是民航 CO_2 排放的最主要来源，诸多学者从 CO_2 排放量测算、时空演化特征、驱动因素等角度开展了卓有成效的研究，但很少有关民航 CO_2 减排潜力及其区域划分的研究。尽管其他领域的 CO_2 排放区域划分成果丰富，但考虑到航空器 CO_2 排放的空间相关性特征，现有区域划分方法与结果无法直接嫁接采用，需做进一步改进。

基于此，首先，本文采用省域航空器 CO_2 排放强度与换算旅客吞吐量指标，构建区域航空器 CO_2 减排潜力分类模型；然后，通过Kruskal算法得到省域网络关系最小生成树；最后，采用谱聚类算法，以“最大化区划优度”为目标，开展中国民用航空器 CO_2 减排潜力的区域划分，并开展区域 CO_2 减排特征及措施讨论。

收稿日期：2021-10-29；**录用日期**：2022-01-05；**网络出版时间**：2022-02-11 11:14
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220210.1710.001.html
基金项目：国家自然科学基金 (U1933117,U2033203)
*通信作者. E-mail：hoorong@nuaa.edu.cn

引用格式：曾雯，胡荣，宋文，等. 中国民用航空器 CO_2 减排潜力的区域划分 [J]. 北京航空航天大学学报，2023，49（9）：2455-2462.
ZENG W，HU R，SONG W，et al. Regional classification of CO_2 emission reduction potential of China's civil aircraft[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（9）：2455-2462 （in Chinese）.

1 区域划分的优化目标

区域划分的目的是使每个子区域内部的差异尽可能小，子区域间的差异尽可能大。据此，兼顾区域内与区域间异质度设置优化目标。

1.1 区域内异质度

区域内异质度用以衡量某子区域内部的差异性水平。其取值越小，表明该子区域内差异越小。可由式(1)求得：

$$\bar{H}_i=\sum_{s=1}^{S}\sum_{t=1}^{T}\sum_{j=1}^{J}(X_{(i,j,s,t)}-\bar{X}_{(i,s,t)})^2\Big/J \tag{1}$$

式中：$\bar{H}_i$为子区域i的区域内异质度；$X_{(i,j,s,t)}$为子区域i内省域j的第s维指标第t年的值；$\bar{X}_{(i,s,t)}$为子区域i内所有省域的第s维指标第t年的均值；I为子区域数量；J为子区域i涵盖省域的数量；S为指标维数；T为时间跨度。

1.2 区域间异质度

区域间异质度用以衡量某子区域与其他子区域间的差异性水平。其取值越大，表明该子区域与其他子区域间的差异越大。可由式(2)和式(3)求得：

$$\bar{X}_{(s,t)}=\sum_{i}^{I}\bar{X}_{(i,s,t)}\Big/I \tag{2}$$

$$\tilde{H}_i=\sum_{s=1}^{S}\sum_{t=1}^{T}\sum_{j=1}^{J}(X_{(i,j,s,t)}-\bar{X}_{(s,t)})^2\Big/J \tag{3}$$

式中：$\tilde{H}_i$为子区域i与其他子区域间异质度；$\bar{X}_{(s,t)}$为整个区域第s维指标第t年的均值。

1.3 区划优度

综合考虑区域内与区域间异质度信息，以“子区域内异质度尽可能小、子区域间异质度尽可能大”为区域划分的优化目标，采用式(4)进行计算并将其命名为“区划优度”。易知，区化优度取值越大，则区域划分结果越合理。

$$H=\sum_{i}^{I}(\tilde{H}_i-\bar{H}_i)\Big/I \tag{4}$$

式中：H为区划优度。

2 区域划分的方法步骤

2.1 构建航空器 CO_2 减排潜力分类模型

2.1.1 航空器 CO_2 排放量测算模型

航空器运行过程可划分为2个阶段：起飞着陆(landing and take-off, LTO)阶段与爬升巡航下降(climb cruise and descent, CCD)阶段。

针对LTO阶段，国际民用航空组织(International Civil Aviation Organization, ICAO)基于大量地面试验与机场实践的数据，构建了ICAO航空器CO_2排放测算模型。该模型LTO阶段包含了进近、滑行、起飞和爬升4个运行阶段；针对CCD阶段，考虑到该过程中推力水平不是标准化的且巡航时长占整个阶段比重很大，在核算大量航班CCD阶段的CO_2排放量时，可简化为核算巡航阶段的CO_2排放量[13]；基于此，航空器CO_2排放量核算共涵盖5个运行阶段的CO_2排放量，计算式为

$$E=\sum_{q=1}^{5}(n\hat{t}_q f_q\bar{I}) \tag{5}$$

式中：E为某航班CO_2排放量，kg；q为航空器运行阶段(进近、滑行、起飞、爬升和巡航)；n为航空器发动机数，个；$\hat{t}_q$为航空器在第q个运行阶段的运行时间，s；f_q为航空器在第q个运行阶段的单发燃油流量，kg/s；$\bar{I}$为CO_2排放指数。

将省域j第t年内所有机场全部类型航空器LTO阶段CO_2排放量与按省域分配的CCD阶段CO_2排放量加总[13]，即可得到省域j第t年航空器CO_2排放总量E_{jt}。

2.1.2 航空器 CO_2 排放强度测算模型

将某省域内航空器CO_2排放强度定义为实现单位换算旅客吞吐量运输而产生的CO_2排放量，计算式为

$$C_{\mathrm{EI}jt}=E_{jt}/T_{jt} \tag{6}$$

式中：$C_{\mathrm{EI}jt}$为省域j第t年航空器CO_2排放强度，kg/人次；T_{jt}为省域j第t年换算旅客吞吐量，人次(90 kg货邮吞吐量等同于1人次旅客吞吐量)。

2.1.3 航空器 CO_2 减排潜力分类模型

航空器CO_2减排潜力是指通过一系列减排措施，能够降低的CO_2排放量，可通过相应的指标加以衡量。选取航空器CO_2排放强度和换算旅客吞吐量2类指标，并选择中位数作为指标水平高低分类间隔，进而可得到区域航空器CO_2减排潜力分类模型，如图1所示。象限I的区域具有高CO_2排放强度与高换算旅客吞吐量，说明区域可降低的CO_2排放量大，减排潜力高。象限II、IV的区域具有高(低)换算旅客吞吐量与低(高)CO_2排放强度，相较于象限I，可减排的空间较小，因此，减排潜力适中。象限III的区域具有低换算旅客吞吐量与低CO_2排放强度，显然其CO_2减排潜力最低。

基于航空数据和航班信息提供商(Official Airways Guide, OAG)数据库(https://www.oag.cn/)中全国民航航班运行数据与中国民用航空局网站(http://www.caac.gov.cn)2007—2016年的年度机场

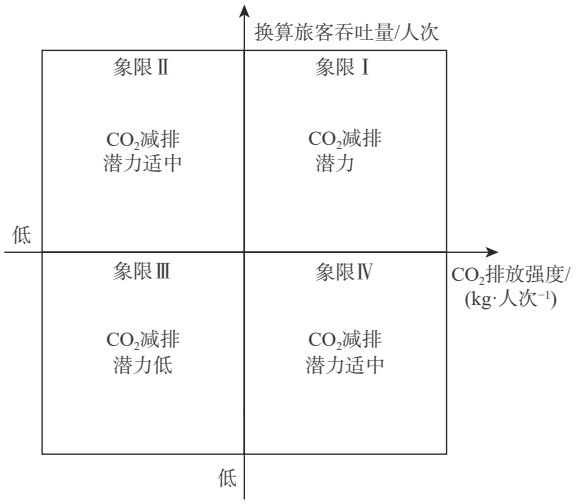


图 1　区域航空器 CO_2 减排潜力分类模型

Fig. 1　Classification model of CO_2 emission reduction potential of regional aircraft

生产公报中旅客和货邮吞吐量数据, 可计算得出 2007—2016 年中国部分地区航空器 CO_2 排放强度与换算旅客吞吐量。航空器 CO_2 减排潜力相关指标统计量如表 1 所示, 可知各省域指标水平差异性明显, 进一步说明航空器 CO_2 减排潜力存在区域差异性。

表 1　航空器 CO_2 减排潜力相关指标统计量

Table 1　Index statistics of Aircraft CO_2 emission reduction potential

指标	CO_2排放量/10^8 kg	CO_2排放强度/(kg·人次$^{-1}$)	货邮吞吐量/10^7 kg	旅客吞吐量/10^6人次	换算旅客吞吐量/10^6人次
最大值	33.20	73.17	386.92	114.40	149.45
最小值	0.23	10.91	0.53	0.82	0.93
均值	5.50	25.43	37.77	21.48	25.68
方差	27.81	101.46	5 185.26	508.76	889.26

2.2　度量省域间多维指标属性距离

省域之间指标属性距离代表不同省域之间航空器 CO_2 减排潜力的差异性, 属性距离越大, 差异性越大。属性距离用于构建省域邻接关系网络图中连接边的权值, 是进行区域划分的前提条件。航空器 CO_2 减排潜力区域划分指标的数据为面板数据, 而综合欧式距离是面板数据聚类中常见的衡量样本间属性距离的方法, 该方法不仅可以提取样本间指标维度的相似关系, 并且考虑了样本间时序变动关系。因此, 考虑到航空器 CO_2 减排潜力区域划分需要同时提取样本之间指标维和时间维的相似性信息, 故采用综合欧式距离法度量不同省域之间的属性距离。

设面板数据 X 中省域 j 的第 s 维指标在第 t 年的值记为 $X_{(j,s,t)}$, 省域 j 与 j' $(j,j'=1,2,\cdots,J, j\neq j')$ 间的综合欧式距离 $D_{jj'}$ 由标准化后的绝对指标距离 $D^1_{jj'}$、增量指标距离 $D^2_{jj'}$、时间波动指标距离 $D^3_{jj'}$ 构成[14]。

1) 绝对指标距离 $D^1_{jj'}$ 为

$$D^1_{jj'}=\sum_{s=1}^{S}\sum_{t=1}^{T}\left(X_{(j,s,t)}-X_{(j',s,t)}\right)^2 \tag{7}$$

2) 增量指标距离 $D^2_{jj'}$ 为

$$D^2_{jj'}=\sum_{s=1}^{S}\sum_{t=2}^{T}\left(\frac{X_{(j,s,t)}-X_{(j,s,t-1)}}{X_{(j,s,t-1)}}-\frac{X_{(j',s,t)}-X_{(j',s,t-1)}}{X_{(j',s,t-1)}}\right)^2 \tag{8}$$

3) 时间波动指标距离 $D^3_{jj'}$ 为

$$D^3_{jj'}=\sum_{s=1}^{S}\left(\frac{\bar{X}_{js}}{S_{js}}-\frac{\bar{X}_{j's}}{S_{j's}}\right)^2 \tag{9}$$

式中: $\bar{X}_{js}(\bar{X}_{j's})$ 为省域 $j(j')$ 的第 s 维指标在 T 年间的均值; $S_{js}(S_{j's})$ 为省域 $j(j')$ 的第 s 维指标在 T 年间的方差。

4) 综合欧式距离 $D_{jj'}$ 为

$$D_{jj'}=z(D^1_{jj'})+z(D^2_{jj'})+z(D^3_{jj'}) \tag{10}$$

式中: $z(D^1_{jj'})$、$z(D^2_{jj'})$、$z(D^3_{jj'})$ 为对应于 Min-max 标准化后的 $D^1_{jj'}$、$D^2_{jj'}$、$D^3_{jj'}$。

2.3　建立省域邻接关系网络图

设研究范围内省域无向网络连通图为 $G^*(V,E)$, 其中 V 为顶点集(每个省域代表 1 个顶点), E 为边集(边长为省域间的属性距离 $D_{jj'}$), 由此构建的是完整的网络图。

研究表明, 一个区域的 CO_2 排放行为不仅取决于自身内部因素的控制, 而且还受周边地区 CO_2 排放的关联作用, 存在空间相关性[4,15]。因此, 以各省域空间邻接关系作为邻接网络图建立的依据, 剔除 $G^*(V,E)$ 中不相邻省市之间的边(海南省在空间上不存在邻接关系, 定义与广东省邻接), 可以建立全国范围(不含港澳台地区)航空器 CO_2 减排潜力省域间的邻接关系网络图, 如图 2 所示。

2.4　生成最小生成树

常见的最小生成树 (minimum spanning tree, MST) 的生成算法有 Prim、Kruskal 等。2.3 节构建的航空器 CO_2 减排潜力省域邻接关系网络图为稀疏图, 更适合采用 Kruskal 算法[16]。具体步骤如下:

步骤 1　构建初始无向连通图 $G(V,E)$。

步骤 2　将 $G(V,E)$ 中所有的边按权值大小进行升序排序。

步骤 3　按权值从小到大依次判断是否选取边, 如果这条边不会与之前选择的所有连接边组成

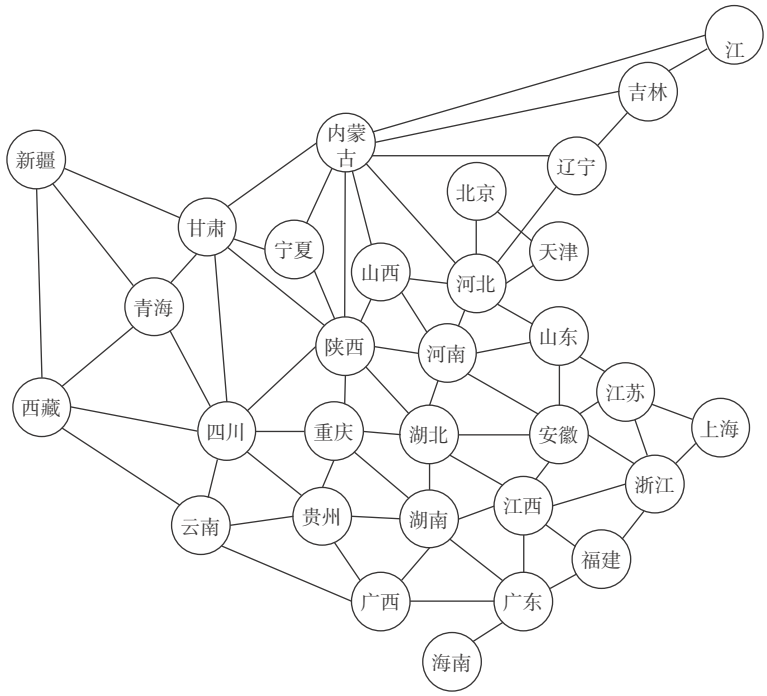


图2 航空器 CO_2 减排潜力省域邻接关系网络

Fig. 2 Provincial adjacency network of aircraft CO_2 emission reduction potential of aircraft

回路，就可以作为MST的一部分；反之，舍去。

步骤4 重复步骤3直到遍历完所有的边，使得具有 v 个顶点的 $G(V,E)$ 筛选出来 $v-1$ 条边，构成 $G(V,E)$ 的最小生成树。

按照上述步骤，可得到中国部分地区航空器 CO_2 减排潜力连续MST，如图3所示。

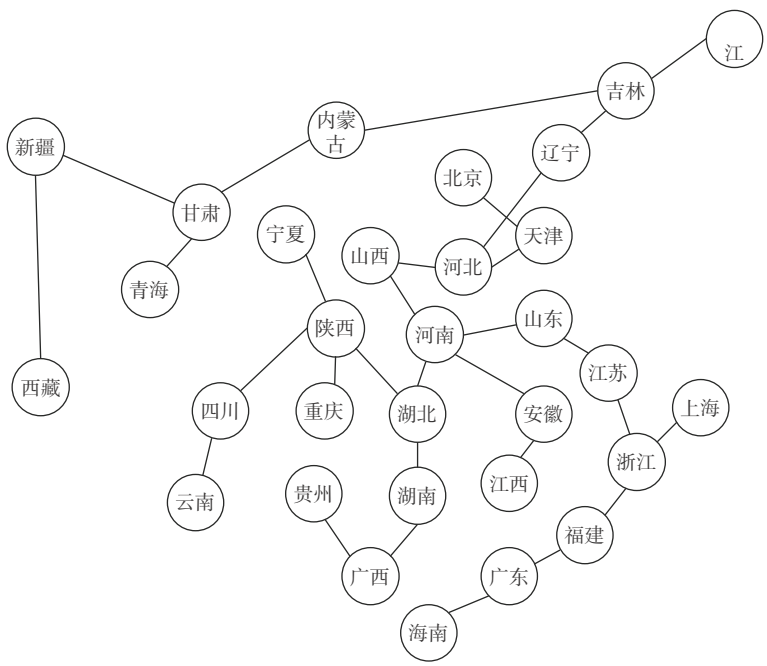


图3 航空器 CO_2 减排潜力连续最小生成树

Fig. 3 Continuous MST of aircraft CO_2 emission reduction potential of aircraft

2.5 截断最小生成树

在中国航空器 CO_2 减排潜力连续MST的基础上，采用谱聚类算法，以最大化区划优度为目标，截断MST，划分中国航空器 CO_2 减排潜力的不同区域。

2.5.1 谱聚类

谱聚类可以解决稀疏相似矩阵的聚类问题，通常比经典聚类算法（如K-means和模糊聚类法（fuzzy clustering method, FCM））产生的结果更优[17]。

将中国部分地区作为一组样本点，利用谱聚类

将其划分为k个簇$G_1, G_2, \cdots, G_k$的步骤如下:

步骤 1　构建空间相似度矩阵$\boldsymbol{A}$, 其元素为

$$A(j,j') = \mathrm{e}^{\frac{-D_{jj'}}{2\sigma^2}} \tag{11}$$

式中: $A(j,j')$为省域j和省域j'之间的相似度; σ为缩放参数。

步骤 2　计算标准化的拉普拉斯矩阵$\boldsymbol{L}=\boldsymbol{Q}^{-1/2}\boldsymbol{A}\boldsymbol{Q}^{-1/2}$, $\boldsymbol{Q}$为一个对角矩阵, 元素$Q(j,j)$为$\boldsymbol{A}$的第j行元素的总和。

步骤 3　对拉普拉斯矩阵$\boldsymbol{L}$进行特征值分解, 得到k个特征向量, 并通过将特征向量纵向堆叠来形成矩阵$\boldsymbol{U}_k=[\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_k]^{\mathrm{T}}$。

步骤 4　通过将$\boldsymbol{U}_k$的行进行归一化操作, 形成矩阵$\boldsymbol{Y}_k$。

步骤 5　将$\boldsymbol{Y}_k$的每一行视为一个点, 并使用K-means将其分为k个簇$C_1, C_2, \cdots, C_k$。

步骤 6　当且仅当矩阵$\boldsymbol{Y}_k$中的行j被分配给簇C_j时, 才将省域j分配给簇G_j。

基于 MST 构建$\boldsymbol{A}$, 按照本节所述步骤, 输入分区数k截断 MST, 可得相应区域划分结果。根据“最大化区划优度”为目标, 确定最优区域划分数量, 得到最优区域划分结果。

2.5.2　确定最优区域划分结果

由式(4)计算不同分区数下的区划优度, 如图 4 所示。当分区数为 4 时, 区划优度最大。因此, 综合考虑研究区域省份数量与分区数的关系, 以及分区数目过多将增加政策制定与实施的难度, 最终确定航空器 CO_2 减排潜力最优分区数目为 4。

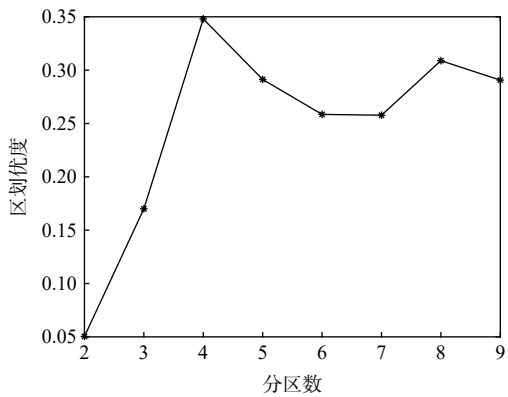


图 4　不同分区数下的区划优度

Fig. 4　Regional classification superiority with different partition numbers

在确定了最优分区数目 4 后, 可得谱聚类算法截断 MST 后形成的子树, 如图 5 所示, 每个子树对应区域划分后的子区域。

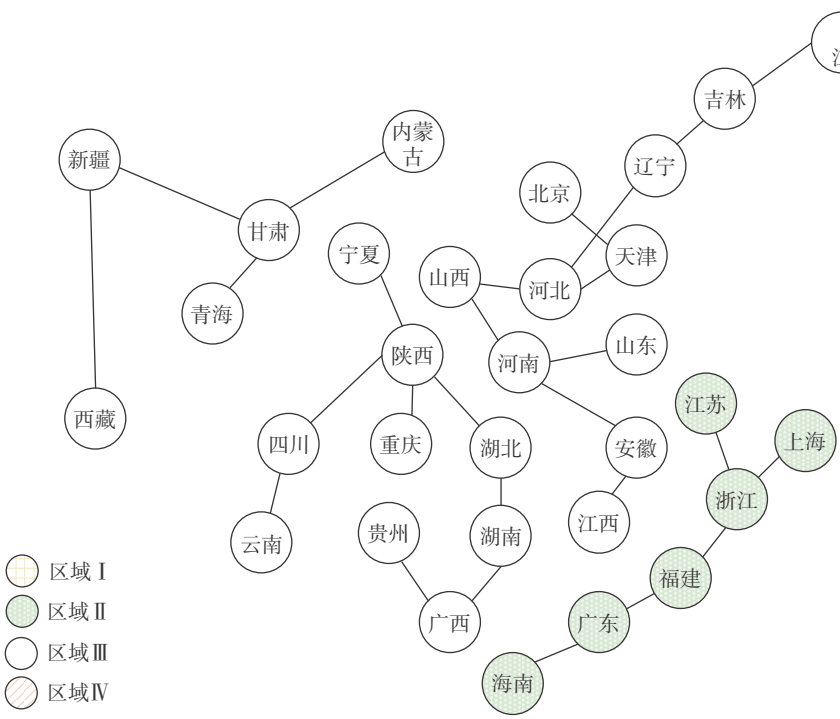


图 5　航空器 CO_2 减排潜力最小生成树截断示意图

Fig. 5　The truncated MST digram of aircraft CO_2 emission reduction potential of aircraft

3　区域划分的结果讨论

3.1　区域划分结果

计算各子区域航空器 CO_2 排放强度与换算旅客吞吐量, 基于航空器 CO_2 减排分类模型, 可得各子区域航空器 CO_2 减排潜力, 如图 6 所示。区域 I、II、III、IV 的减排潜力分别对应于象限 I、II、III、IV 所属的类型。

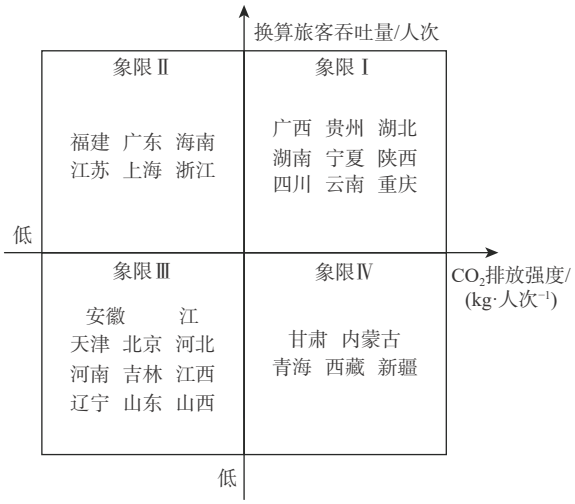


图 6 四分区下区域航空器 CO_2 减排潜力

Fig. 6 CO_2 emission reduction potential of regional aircraft under four partions

四分区与中国已有的一些区域划分的区划优度如表 2 所示。可以发现，相较于已有的区域划分，四分区的区划优度最大，区划优度为 0.35，进一步说明综合指标水平与地理空间邻接关系的区域划分结果更优。

表 2 不同区域划分方法的区划优度

Table 2 Regional classification superiority with different partition methods

分区方法	区划优度
四分区(本文方法)	0.35
东中西分区	0.16
六大行政分区	0.07
八大经济分区	0.25

3.2 区域减排特征

根据航空器 CO_2 减排潜力分区结果，各子区域航空器 CO_2 排放特征分析如下：

区域 I：CO_2 减排潜力最高，涵盖省域多位于中国西南、中南地区，地形复杂，为航空运输发展创造了条件。换算旅客吞吐量和起降架次在 4 个区域中排名第 2，分别为 2.14×10^7 人次与 2.08×10^5 架次，CO_2 排放强度处于较高水平(23.63 kg/人次)。区域内有 5 个省域(全国共 8 个)起降架次年均增长率高于换算旅客吞吐量的年均增长率，表明该区域存在航班客座率(载运率)下滑的现象，一定程度上增加了区域 CO_2 排放强度；此外，远距离航线占比低也会使区域 CO_2 排放强度处于较高水平[18]，该区域国际航线占比低，仅为最高区域(区域 II)的 56.85%。

区域 II：CO_2 减排潜力适中，涵盖省域位于中国东南沿海地区，交通、科技、管理、经济水平均处于全国领先地位，地区人均 GDP 在 4 个区域中最高(3.01×10^4 元/人)，衍生出高民航出行需求。此区域涵盖中国重点打造的长三角与珠三角 2 个世界级机场群，换算旅客吞吐量、起降架次与国际航线占比在 4 个区域中最高，分别达到 5.63×10^7 人次、3.68×10^5 架次与 29.53%，而 CO_2 排放强度在 4 个区域最低(19.43 kg/人次)。但高民航出行需求使该区域面临着航班时刻、空域资源紧张等情况，这会增加航班延误等待时间，从而增加航空器燃油消耗量。

区域 III：CO_2 减排潜力最低，涵盖省域多位于中国华北和东北地区，总体来看，区域 CO_2 排放强度较低，排名第 3。但江西省与河南省国际航线占比偏低，分别为 11.85% 与 12.41%；此外，京津冀地区面临航班时刻与空域资源紧张的问题。

区域 IV：CO_2 减排潜力适中，涵盖省域靠近中国西部和北部边境地区，空域资源相对富裕；青海、西藏地势较高，拥有 12 个高高原机场，占全国总量的 60%。此区域 CO_2 排放强度在 4 个区域中最高(41.66 kg/人次)，但换算旅客吞吐量、起降架次、地区人均 GDP 与国际航线占比均为最低。受限于经济条件，此区域民航出行需求不足。进一步可知，此区域多高高原机场分布，航空器不能满载起飞，存在客座率低的情况，并且由于气候复杂多变，难以保证机场出港准点率，会增加航空器延误等待时间，油耗量也会同步增加。

3.3 区域减排措施

3.2 节分析表明：航空器 CO_2 排放具有明显的区域差异性特征，需确定富有差异化的 CO_2 减排重点及配套措施：

区域 I：①通过优化收益管理、机票超售等市场手段，实施航班共享、航班合并及增加经停点等运营策略，提高航班客座率(载运率)。②推进成渝世界级机场群的构建，以成渝地区为中心带动整个区域的对外开放发展，提高国际航线占比。

区域 II：①通过时刻交换、代码共享、共同经营等方式优化航班时刻。②引进先进的通信、导航、监控和情报技术，提高空中交通服务能力。③通过优化空域结构与管制手段，提升航班飞行效率。

区域 III：①江西和河南需加快融入“一带一路”建设，开辟更多通往“一带一路”沿线国家的航线，扩大远距离航线占比。②京津冀地区需统筹做好航班时刻与空域结构优化，提升域内航班的运行效率，减少航班的延误与绕飞。

区域 IV：①鼓励航空公司采用 A319 等高原机场适应性更佳的机型运营高原航线。②加强高高原机场新技术的研发和应用，提高空地运营效率、减少航班延误，从而提高节能减排效果[19]。

4　结　论

1）基于 2007—2016 年省域航空器 CO_2 排放强度与换算旅客吞吐量, 构建了航空器 CO_2 减排潜力分类模型; 并综合指标水平与地理空间邻接关系, 构建了区域划分模型。

2）航空器 CO_2 减排潜力最优区域划分为四分区, 区划优度达到 0.35, 且优于中国已有的区域划分。

3）航空器 CO_2 减排潜力最高的区域主要分布于中国西南和中南地区, 最低的区域主要分布于东北和华北地区。

4）建议中国民用航空局、机场和航空公司根据航空器 CO_2 排放区域差异性特征落实空中交通管制优化、节能减排新技术应用、航班共享策略实施等相关 CO_2 减排措施。

参考文献（References）

[1] JACOB T, JACOPO T, APOSTOLOS P, et al. CO_2 emissions from fuel combustion: IEA 2020[R]. Paris: OECD Publishing, 2020: 9-12.

[2] WU C T, HE X H, DOU Y. Regional disparity and driving forces of CO_2 emissions: Evidence from China's domestic aviation transport sector[J]. Journal of Transport Geography, 2019, 76: 71-82.

[3] 朱佳琳. 中国民用航空器碳排放时空演化特征及驱动因素研究[D]. 南京: 南京航空航天大学, 2020: 37-43.
ZHU J L. Spatial-temporal evolution characteristics and driving factors of carbon emissions from civil aircraft in China[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2020: 37-43(in Chinese).

[4] HU R, ZHU J L, ZHANG Y, et al. Spatial characteristics of aircraft CO_2 emissions at different airports: Some evidence from China[J]. Transportation Research Part D:Transport and Environment, 2020, 85: 102435.

[5] LIU J G, LI S J, JI Q. Regional differences and driving factors analysis of carbon emission intensity from transport sector in China[J]. Energy, 2021, 224: 120178.

[6] 曾晓莹, 邱荣祖, 林丹婷, 等. 中国交通碳排放及影响因素时空异质性[J]. 中国环境科学, 2020, 40(10): 4304-4313.
ZENG X Y, QIU R Z, LIN D T, et al. Spatio-temporal heterogeneity of transportation carbon emissions and its influencing factors in China[J]. China Environmental Science, 2020, 40(10): 4304-4313(in Chinese).

[7] 宋德勇, 徐安. 中国城镇碳排放的区域差异和影响因素[J]. 中国人口·资源与环境, 2011, 21(11): 8-14.
SONG D Y, XU A. Regional difference and influential factors of China's urban carbon emissions[J]. China Population, Resources and Environment, 2011, 21(11): 8-14(in Chinese).

[8] 王建伟, 李娉, 高洁. 中国交通运输碳减排区域划分[J]. 长安大学学报(自然科学版), 2012, 32(1): 72-77.
WANG J W, LI P, GAO J. Region division in China for transportation carbon emission reduction[J]. Journal of Chang'an University (Natural Science Edition), 2012, 32(1): 72-77(in Chinese).

[9] 武旭, 杜奕, 贾传峻. 基于全距离完全度量聚类的交通运输能耗区域划分[J]. 交通运输系统工程与信息, 2019, 19(2): 7-12.
WU X, DU Y, JIA C J. Regionalization of China's transportation energy consumption based on full-order-constrained complete linkage partitioning[J]. Journal of Transportation Systems Engineering and Information Technology, 2019, 19(2): 7-12(in Chinese).

[10] 张彬, 姚娜, 刘学敏. 基于模糊聚类的中国分省碳排放初步分析[J]. 中国人口·资源与环境, 2011, 21(1): 53-56.
ZHANG B, YAO N, LIU X M. Preliminary fuzzy clustering analysis on carbon emission in different provinces of China[J]. China Population, Resources and Environment, 2011, 21(1): 53-56(in Chinese).

[11] 涂正革, 谌仁俊. 中国碳排放区域划分与减排路径: 基于多指标面板数据的聚类分析[J]. 中国地质大学学报(社会科学版), 2012, 12(6): 7-13.
TU Z G, CHEN R J. Regional division of carbon emission and emission reduction path in China: Cluster analysis based on multi-index panel data[J]. Journal of China University of Geosciences (Social Sciences Edition), 2012, 12(6): 7-13(in Chinese).

[12] CHANG K L, DU Z F, CHEN G J, et al. Panel estimation for the impact factors on carbon dioxide emissions: A new regional classification perspective in China[J]. Journal of Cleaner Production, 2021, 279: 123637.

[13] 朱佳琳, 胡荣, 张军峰, 等. 中国航空器碳排放测算与演化特征研究[J]. 武汉理工大学学报(交通科学与工程版), 2020, 44(3): 558-563.
ZHU J L, HU R, ZHANG J F, et al. Research on the measurement and evolution characteristics of aircraft carbon emissions in China[J]. Journal of Wuhan University of Technology (Transportation Science & Engineering), 2020, 44(3): 558-563(in Chinese).

[14] 杨娟, 谢远涛. 面板数据聚类的复合方法与应用[M]. 北京: 对外经济贸易大学出版社, 2016: 19-20.
YANG J, XIE Y T. Composite method and application of panel data clustering[M]. Beijing: University of International Business and Economics Press, 2016: 19-20(in Chinese).

[15] 刘贤赵, 郭若鑫, 张勇, 等. 中国省域碳排放空间依赖结构的非参数估计及其实证分析[J]. 中国人口·资源与环境, 2019, 29(5): 40-51.
LIU X Z, GUO R X, ZHANG Y, et al. Nonparametric estimation and empirical analysis of spatial dependence structure of provincial carbon emissions in China[J]. China Population, Resources and Environment, 2019, 29(5): 40-51(in Chinese).

[16] FAVATI P, LOTTI G, MENCHI O, et al. Construction of the similarity matrix for the spectral clustering method: Numerical experiments[J]. Journal of Computational and Applied Mathematics, 2020, 375: 112795.

[17] MUR A, DORMIDO R, DURO N, et al. Determination of the optimal number of clusters using a spectral clustering optimization[J]. Expert Systems with Applications, 2016, 65: 304-314.

[18] 郭鹏程. 中国民航客运碳排放时空演变与影响因素研究[D]. 兰州: 西北师范大学, 2020: 31-34.
GUO P C. Spatial-temporal evolution and influencing factors of carbon emissions of civil aviation passenger transport in China[D]. Lanzhou: Northwest Normal University, 2020: 31-34(in Chinese).

[19] 胡荣, 王德芸, 冯慧琳, 等. 碳达峰视角下的机场航空器碳排放预测[J]. 交通运输系统工程与信息, 2021, 21(6): 257-263.
HU R, WANG D Y, FENG H L, et al. Prediction of aircraft CO_2 emission from perspective of CO_2 emission peak[J]. Journal of Transportation Systems Engineering and Information Technology, 2021, 21(6): 257-263(in Chinese).

Regional classification of CO_2 emission reduction potential of China's civil aircraft

ZENG Wen, HU Rong*, SONG Wen, LIU Zhihao, ZHANG Junfeng

(College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: Scientific regional classification is fundamental for investigating aircraft CO_2 emissions with regional differences and spatial relativity, and for proposing differentiated CO_2 emission reduction measures. Using aircraft operation data from China's some provinces over a ten-year time span (2007–2016), this paper first establishes a model of CO_2 emission reduction potential. Then we employ the Kruskal algorithm to obtain the minimum spanning tree of the provincial network relationship. Next the spectral clustering algorithm is used to divide the area of CO_2 emission reduction potential with the goal of maximizing the regional classification superiority. Our analysis shows that the optimal classification involves four regions with the regional classification superiority of 0.35. The region with the highest CO_2 emission reduction potential is mainly in southwest and central south China, and the region with the lowest potential is mainly in northeast and north China. According to the regional characteristics, differentiated CO_2 emission reduction measures are proposed for different subjects such as civil aviation administration, airports, and airlines.

Keywords: civil aircraft; CO_2 emission reduction potential; regional classification; minimum spanning tree; spectral clustering

Received: 2021-10-29; **Accepted**: 2022-01-05; **Published Online**: 2022-02-11 11:14
URL: kns.cnki.net/kcms/detail/11.2625.V.20220210.1710.001.html
Foundation items: National Natural Science Foundation of China (U1933117,U2033203)
* **Corresponding author**. E-mail: hoorong@nuaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0649

IFB不确定性对双频码载分歧监测的影响分析

康凯[1]，王志鹏[1]，方堃[2,*]，朱衍波[1]

(1. 北京航空航天大学 电子信息工程学院，北京 100191；2. 北京航空航天大学 前沿科学技术创新研究院，北京 100191)

摘 要：码载分歧（CCD）监测器是卫星导航地基增强系统（GBAS）引入的完好性监测器之一，用于监测码伪距与载波相位观测值之间的不一致性。双频平滑技术的引入改变了CCD监测器的检验统计量等参数，进而影响了监测性能。针对双频GBAS中存在的频间偏差（IFB），基于北斗双频观测数据，分析IFB不确定性对于双频码载分歧监测性能的影响。研究结果表明:在IFB不确定性的影响下，监测器的阈值增大了26.1%，导致其在发生CCD故障情况下的漏检率显著增大，最小可检测故障值增大了26.9%，监测器的灵敏度降低；最坏情况下系统完好性损失概率从低于10^{-14}增大到接近10^{-8}，同时机载端为满足漏检率性能安全要求所引入的延迟值更大，导致CCD监测器的故障响应变慢，双频GBAS的完好性受到影响。

关 键 词：频间偏差；码载分歧；地基增强系统；双频监测器；完好性

中图分类号：V249.32

文献标志码：A **文章编号**：1001-5965(2023)09-2463-10

地基增强系统(ground-based augmentation systems, GBAS)基于全球导航卫星系统(global navigation satellite system, GNSS)误差的时空相关性，将差分校正量和完好性信息广播给飞机，为飞机提供精确、安全、可靠的进近引导服务[1]。C类GBAS进近服务类型(GBAS approach service type C, GAST-C)使用单频信号进行定位和监测，受电离层异常等因素的影响较大，必须对定位误差进行非常保守的包络，限制了系统的性能，只能支持国际民航组织(International Civil Aviation Organization, ICAO)要求的CAT I类进近服务[2]。为了应对电离层异常威胁，D类GBAS进近服务类型(GBAS approach service type D, GAST-D)在GAST-C的基础上增加了若干完好性监测器，使系统可以满足CAT II/III类精密进近要求[3-4]。码载分歧(code carrier divergence, CCD)监测器是GAST-D引入的一类典型监测器，用来监测用户接收到的码伪距与载波相位测量值的不一致性。Xie[5]最早提出CCD监测算法和监测阈值的确定方法，并对监测器性能进行了初步测试[5]。Simili和Pervan[6]进一步完善了CCD监测算法和流程，并针对CCD故障完成了系统的完好性分析。目前，CCD监测算法已经被写入航空无线电技术委员会(Radio Technical Commission for Aeronautics, RTCA)制定的GBAS机载最低运行性能标准DO-253D[3]和欧洲民用航空设备组织(European Organization for CivilAviation Equipment, EUROCAE)制定的GBAS地面设备最低运行性能规范ED-114A[7]。

随着GNSS各星座的不断发展及多个频段新体制信号的加入，双频多星座(dual-frequency multi-constellation, DFMC)GBAS可以使用双频平滑技术来应对电离层活动对于系统的威胁[8]。目前主要使用2种双频平滑技术，即无分歧平滑(divergence-free smoothing, Dfree)和无电离层平滑(ionosphere-free smoothing, Ifree)[9]。Dfree使用单频伪距与双频

收稿日期：2021-10-31；**录用日期**：2022-02-13；**网络出版时间**：2022-02-21 09:00
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220219.1405.001.html
基金项目：国家重点研发计划 (2020YFB0505602)；国家自然科学基金 (61871012,62022012)；民航安全能力建设基金 (CAAC Contract 2021(77)，CAAC Contract 2020(123))；北京市科技创新基金 (Z191100001119134)

*通信作者. E-mail：fangkun@buaa.edu.cn

引用格式：康凯，王志鹏，方堃，等. IFB不确定性对双频码载分歧监测的影响分析[J]. 北京航空航天大学学报，2023，49(9)：2463-2472. KANG K，WANG Z P，FANG K，et al. Impact of IFB uncertainty on dual-frequency code carrier divergence monitoring[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2463-2472 (in Chinese).

载波相位测量值进行组合，虽然能在保持单频系统噪声水平的同时提供更精确的监测，但却只能消除部分电离层延迟误差。Ifree使用2个频率的伪距与载波相位测量值进行组合，可以完全消除电离层延迟误差，然而付出的代价却是测量噪声成倍增加，同时会在测量值中引入频间偏差（inter-frequency bias, IFB）[10]。目前提出的F类GBAS进近服务类型（GBAS approach service type F, GAST-F）也分为GAST-F SF模式和GAST-F IF模式[11]，分别对应Dfree和Ifree，2种模式的转换取决于是否检测到电离层异常[12-13]。Jiang等[14]在单频CCD监测器的基础上提出针对GAST-F的双频CCD监测器，扩展到双频系统的CCD监测器是对于复杂故障空间下、多种故障源的联合监测。

IFB是不同频率的信号在接收机端产生的硬件延迟差异，属于差分码偏差（differential code bias, DCB）的接收机端分量，而卫星端分量通常称为卫星端群延迟（time group delay, TGD）[15]。IFB与接收机的结构参数有关，同时也会随着观测卫星（或其仰角）的不同而发生变化，双频载波相位观测量中的IFB与伪距相比可以忽略[10]。Ma和Maruyama[16]提出一种结合电离层总电子含量（total electron content, TEC）快速估计接收机IFB的方法，该方法在后续电离层研究中得到广泛应用。Holaschutz等[17]结合双频电离层观测值和三阶电离层球谐波模型，使用卡尔曼滤波对GPS监测站的DCB进行估计，但并没有对IFB和TGD进行分离。Li等[18]提出中国科学院测量与地球物理研究所DCB（institute of geodesy and geophysics DCB, IGGDCB）方法，结合广义三角级数局域电离层模型和最小二乘估计，使用TGD“零均值”条件分离得到IFB值。Liu等[19]借助北斗信号模拟器和接收机，使用“零伪距”法直接对IFB进行硬件层面的标定和校准。根据文献[16-19]中对IFB修正方法的研究可知，目前IFB难以精确修正，修正后的不确定度最小为0.1 m左右，该IFB不确定性会作为Ifree平滑伪距中的误差项，对使用Ifree伪距相关检验统计量的双频CCD监测器产生影响。

首先，介绍目前单频和双频GBAS中的CCD监测器算法，进而，对GBAS站参考接收机IFB不确定性进行估计，推导其与监测器检验统计量之间的关系，接着，基于北斗双频观测数据完成标称情况下的监测器特性建模，最后，结合仿真结果对双频CCD监测器性能受IFB不确定性的影响进行验证和分析。

1 码载分歧监测

CCD故障是在用户端观测到的一种典型故障，表现为码伪距与载波相位测量值的不一致性。针对此类故障，目前已经提出单频和双频CCD监测器，分别应用于GAST-D和GAST-F完好性监测。

1.1 单频CCD监测器

CCD监测器的输入量是码减载波（code minus carrier, CMC）的变化率。C_{CMC}的计算式为[14]

$$C_{\mathrm{CMC}}=\rho-\varphi \tag{1}$$

式中：ρ为原始伪距观测量；φ为原始载波相位观测量。

k时刻CCD监测器的输入量$z(k)$为[14]

$$z(k)=\frac{C_{\mathrm{CMC}}(k)-C_{\mathrm{CMC}}(k-1)}{\Delta} \tag{2}$$

式中：Δ为数据的采样间隔。

CCD监测器的核心结构是级联一阶低通滤波器，其时域表达式为[6]

$$\begin{cases} d_1(k)=\dfrac{\tau_1-T_0}{\tau_1}d_1(k-1)+\dfrac{1}{\tau_1}[z(k)-z(k-1)] \\ d_2(k)=\dfrac{\tau_2-T_0}{\tau_2}d_2(k-1)+\dfrac{T_0}{\tau_1}d_1(k-1) \end{cases} \tag{3}$$

式中：τ_1和τ_2分别为2个一阶低通滤波器的滤波时间常数；T_0为数据的采样率；d_1和d_2为滤波器的输出，即CCD监测器的检验统计量。

通常在拉普拉斯域内分析CCD监测器的稳态响应，其表达式为[6]

$$F_{\mathrm{m}}=\frac{1}{(\tau_1 s+1)(\tau_2 s+1)} \tag{4}$$

式中：s为拉普拉斯算子。

滤波器输出的检验统计量为[6]

$$\hat{D}(s)=F_{\mathrm{m}}Z(s) \tag{5}$$

达到稳态后[14]：

$$\lim_{s\to 0}F_{\mathrm{m}}Z(s)=\frac{s^2}{(\tau_1 s+1)(\tau_2 s+1)}\cdot\frac{d}{s^2}=d \tag{6}$$

稳态值d就是待检测的码载波分歧的速率值，将其与阈值比较来对CCD故障进行监测。

1.2 双频CCD监测器

北斗B1I和B3I这2个频率上的原始伪距观测量为

$$P_1=R+t^s+t_{\mathrm{r}}+I_1+T+M_1+\varepsilon_1+I_{\mathrm{FB}}+T_{\mathrm{GD}} \tag{7}$$

$$P_3=R+t^s+t_{\mathrm{r}}+I_3+T+M_3+\varepsilon_3 \tag{8}$$

式中：R为卫星与用户之间的真实距离；t^s为卫星钟差；t_{r}为接收机钟差；I为电离层延迟，与导航信号的频率有关；T为对流层延迟；M为多径误差；ε为热噪声；I_{FB}为接收机频间偏差；T_{GD}为卫星端群延迟。

载波相位测量值与式(7)和式(8)类似。北斗卫星导航系统（BeiDou navigation satellite system, BDS）卫星钟差以B3I频率的天线相位中心为时空

零基准, 钟差中包含 B3I 频率的硬件延迟。因此, 可以将 B1I 与 B3I 的伪距硬件延迟不一致的部分仅写在 B1I 的伪距表达式中。

目前的 GAST-F 处理模式, 正常情况下系统使用单频观测量进行定位, 使用 Dfree 平滑进行监测。当机载端监测到电离层异常时, 系统将从 Dfree 平滑模式转换为 Ifree 平滑模式, 使用 Ifree 平滑进行定位和监测[11], GBAS 载波平滑处理过程如图 1 所示。

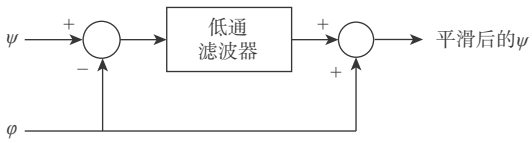


图 1　GBAS 载波平滑处理过程

Fig. 1　GBAS carrier smoothing process

其中 ψ 代表码伪距。对于 Ifree 平滑, ψ 和 φ 分别为[13]

$$\psi_{\text{Ifree}} = P_3 - \frac{1}{\alpha}(P_3 - P_1) \tag{9}$$

$$\varphi_{\text{Ifree}} = \varphi_3 - \frac{1}{\alpha}(\varphi_3 - \varphi_1) \tag{10}$$

式中: α为与所使用信号的频率相关的系数, 表达式为

$$\alpha = 1 - f_3^2 / f_1^2 \tag{11}$$

其中: f_1为 B1I 的中心频率; f_3为 B3I 的中心频率。

实际上, 由于 GBAS 是一个差分系统, 卫星端 TGD 会在机载端差分校正的过程中被消除掉, 因此, 后续分析中省略了 TGD 项。Ifree 平滑后的伪距、载波相位表达式为[10]

$$\rho_{\text{Ifree}} = R + t^s + t_r + T + \eta_{\rho_3} - \frac{1}{\alpha}\left(\eta_{\rho_3} - \eta_{\rho_1}\right) + \frac{1}{\alpha} I_{\text{FB}} \tag{12}$$

$$\varphi_{\text{Ifree}} = R + t^s + t_r + T + \eta_{\varphi_3} - \frac{1}{\alpha}\left(\eta_{\varphi_3} - \eta_{\varphi_1}\right) + \varepsilon_{\text{N}} \tag{13}$$

式中: η_ρ和η_φ分别为原始伪距和载波相位观测量中的多径和热噪声; ε_{N}为载波相位整周模糊度误差, 在前后差分过程中被消除, 因此, 忽略此项。

双频 CCD 监测器与单频 CCD 监测器的区别主要在于检验统计量的变化, 在双频 CCD 监测中, 针对 Ifree 平滑滤波方式一共有 3 类原始输入量, 分别为[14]

$$\chi_{\text{IF}_\rho} = \frac{\gamma}{\gamma - 1}\rho_1 - \frac{1}{\gamma - 1}\rho_3 \tag{14}$$

$$\chi_{\text{IF}_\varphi} = \frac{\gamma}{\gamma - 1}\varphi_1 - \frac{1}{\gamma - 1}\varphi_3 \tag{15}$$

$$\chi_{\text{IF}} = \chi_{\text{IF}_\rho} - \chi_{\text{IF}_\varphi} \tag{16}$$

式中: γ为 Ifree 组合的频率因子[14]:

$$\gamma = f_3^2 / f_1^2 \tag{17}$$

由于式 (14) 和式 (15) 需要针对卫星移动、接收机和卫星钟差及对流层延迟进行补偿和校正[14], 主要对式 (16) 中 χ_{IF}进行研究, 此输入量更符合 CCD 监测器的设计初衷。

2　IFB 不确定性的影响

IFB 是不同频率的信号在接收机端产生的硬件延迟差异, 其与接收机的结构参数有关。Ifree 平滑使用 2 个频率的伪距观测值引入了 IFB, 虽然双频 CCD 监测器是对码载分歧的速率进行监测, 但是受周围环境的影响, IFB 会在一定范围内随机变化, 存在不确定性, 并最终影响监测性能, 带来潜在的完好性风险。

2.1　IFB 不确定性估计

结合式 (7) 和式 (8) 所示 2 个频率的伪距测量值可以得到 2 个频率上伪距的差值:

$$P_1 - P_3 = I_1 - I_3 + I_{\text{FB}} + T_{\text{GD}} \tag{18}$$

理论上, 式 (18) 中还应该包括这 2 个频率之间天线群延迟变化量(group delay variations, GDV)的差, 其是所观测卫星的仰角、方位角和所用频率的函数[20]。但根据 Beer 等[21] 的研究, 即使在最坏情况下, 1 561.098 MHz(B1I 的中心频率)和 1 268.520 MHz (B3I 的中心频率)对应的 GDV 差值与 IFB 相比也要小得多。因此, 式 (18) 中并未考虑 GDV。

需要使用电离层总电子含量(total electron content, TEC)产品将式 (18) 中的 2 个频率电离层延迟的差值从双频伪距差值中扣除。使用 TEC 产品计算电离层延迟的方法为[17]

$$I = \frac{40.28}{f^2} F V_{\text{TEC}} \tag{19}$$

式中: F为倾斜因子; f为所使用信号对应的频率; V_{TEC}为垂直方向 TEC 值。

根据 TEC 与电离层延迟的关系, 将式 (18) 改写为

$$I_{\text{FB}} + T_{\text{GD}} = P_1 - P_3 - 40.28\left(\frac{1}{f_1^2} - \frac{1}{f_3^2}\right) F V_{\text{TEC}} \tag{20}$$

使用 IGS GNSS 联合实验室(The Multi-GNSS Experiment, MGEX)公布的卫星端 DCB 产品将 TGD 扣除掉, 从而实现式 (20) 中 IFB 与 TGD 的分离。

同时, 使用中国东营胜利机场(ZSDY)GBAS 站于 2019 年 11 月 24 日至 2020 年 1 月 2 日收集到的 32 天 B1I 和 B3I 观测数据估计地面接收机 IFB 的标准差, 将其作为 IFB 不确定性, 最终得到的平均结果为 0.14 m。采用相同的方法对用户接收机的 IFB 不确定度进行标定, 得到的结果为 0.22 m。虽

然 IFB 统计特性存在波动性，但通过采集东营 GBAS 站实际运行的观测数据进行较长时间的 IFB 不确定性统计分析，所得的结果与参考文献 [15] 中的近似，因此具有一定的典型性。

2.2 IFB 不确定性与检验统计量

IFB 时间序列存在不确定性，即相邻历元之间的 IFB 也会出现差值，此误差会存在于检验统计量 χ_{IF} 及整个检测流程中。将式 (12) 和式 (13) 代入式 (15) 中，可得

$$\chi_{\mathrm{IF}}=\eta_{\rho_3}-\frac{1}{\alpha}\left(\eta_{\rho_3}-\eta_{\rho_1}\right)-\eta_{\varphi_3}+\frac{1}{\alpha}\left(\eta_{\varphi_3}-\eta_{\varphi_1}\right)+\frac{1}{\alpha}\varepsilon_{\mathrm{IFB}} \tag{21}$$

式中：η_ϕ 比 η_ρ 的量级小很多，通常忽略；$\varepsilon_{\mathrm{IFB}}$ 为 IFB 前后历元做差后剩余的残差，即 IFB 不确定性带来的误差。

双频 CCD 监测器的检验统计量是监测器其他参数的解算基础，此 IFB 不确定性带来的测距误差会进一步影响监测器在标称情况下的阈值设定、故障情况下的漏检率（probability of missed detection, PMD）、完好性损失（loss of integrity, LOI）概率及 PMD 安全性能等。

3 标称情况下的监测器特性

分析 IFB 不确定性对双频 CCD 监测器的影响，需要从标称情况和故障情况 2 方面展开。首先，使用标称情况下的 B1I 和 B3I 双频观测数据，计算标称情况下检验统计量的统计特性，对双频 CCD 监测器的阈值建模。

3.1 检验统计量特性

选用东营机场 GBAS 站于 2021 年 1 月 20 日收集到的 B1I 和 B3I 数据进行 CCD 监测器的建模和性能分析，数据采样间隔 Δ 为 0.5 s，2 个一阶滤波器的时间常数均为 30 s。将 Ifree CMC 输入到双频 CCD 监测器中，得到的检验统计量时间序列如图 2 所示。北斗 GEO 卫星对应的检验统计量较为稳定。受可见性影响，IGSO 卫星和 MEO 卫星对应的检验统计量在卫星进入视野和退出视野时会出现明显峰值，但都在 5×10^{-3} m/s 内。

3.2 阈值设定

双频 CCD 监测器监测阈值的确定方法是对标称情况下检验统计量的标准差进行膨胀包络，使其满足为 CCD 故障分配的完好性风险概率要求[6]，即

$$T_{\mathrm{CCD}}=k_{\mathrm{ffd}}I_{\mathrm{NF}}\sigma_{\mathrm{d}} \tag{22}$$

式中：k_{ffd} 为使监测器满足完好性要求的膨胀系数，取值为 5.83[3]；I_{NF} 为用来包络厚尾分布的膨胀系数；σ_{d} 为标称情况下检验统计量的标准差。

图 2 表明在观测期间内并没有 CCD 故障出现，属于标称情况，可以用于双频 CCD 监测器阈值的计算。

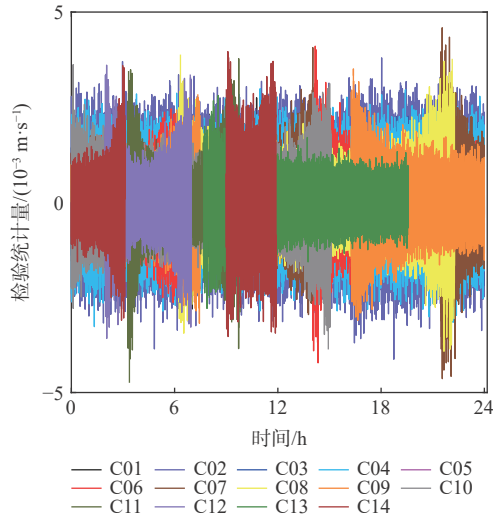
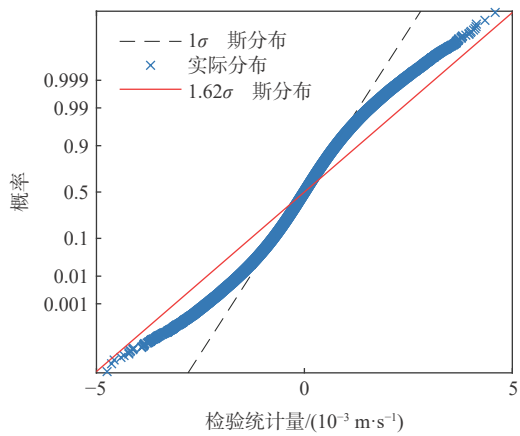


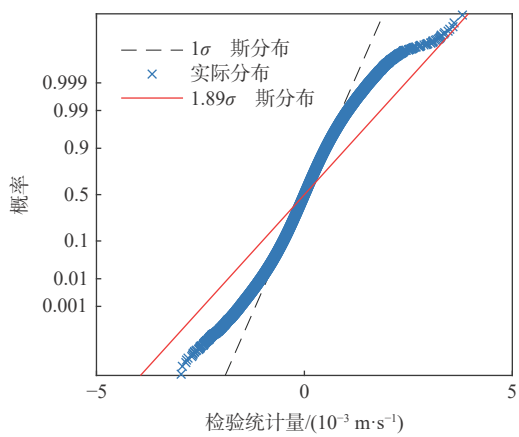
图 2 检验统计量时间序列

Fig. 2 Time series of test statistics

图 3 为校验统计量分布，图 3(a) 为原始检验统计量的分位数图（quantile quantile plot, Q-Q），蓝色数据点表示原始检验统计量，符合零均值的高斯分布，但是存在厚尾的情况，因此，需要使用INF膨胀



(a) 原始的检验统计量分布



(b) 扣除IFB不确定性后的检验统计量分布

图 3 检验统计量分布

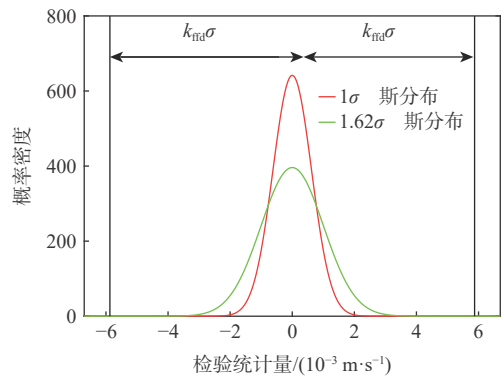
Fig. 3 Distribution of test statistics

系数对其标准差进行包络。灰色虚线对应原始检验统计量的标准差, 为 6.22×10^{-4} m/s, 红色实线对应膨胀后的标准差, INF为 1.62, 包络后的标准差为 0.001 0 m/s。

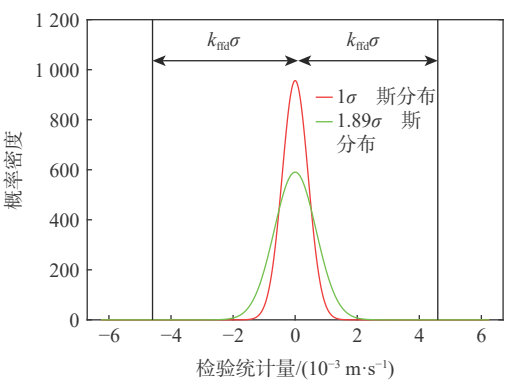
为了分析 CCD 监测阈值受 IFB 不确定性的影响, 使用第 2.1 节中的标定结果将 IFB 不确定性从原始输入量中扣除, 形成不包含 IFB 不确定性的检验统计量。此处理方法的前提是 IFB 与其他误差不相关。需要注意, 由此得到的新检验统计量仅用于分析双频 CCD 监测器受 IFB 不确定性的影响, 在实际 CCD 监测中无法得到。

图 3(b) 为扣除 IFB 不确定性后的检验统计量累积分布函数(cumulative distribution function, CDF)分布图, 图 3(b) 中的蓝色数据点表明此时的检验统计量仍符合零均值的高斯分布, 且存在厚尾的情况, 但是由于标准差减小, 曲线的斜率更大。灰色虚线对应包络前的标准差, 为 4.17×10^{-4} m/s, 红色实线对应膨胀包络后的标准差, INF为 1.89, 膨胀后的标准差为 7.88×10^{-4} m/s。

图 4 为双频 CCD 监测器阈值, 采用式 (22) 中的计算方法, 得到双频 CCD 监测器的阈值为 0.005 8 m/s, 如图 4(a) 中的黑色实线所示。使用扣除 IFB 不确定性后的检验统计量标准差计算得到的阈值为 0.004 6 m/s,



(a) 包含IFB不确定性时的阈值



(b) 扣除IFB不确定性后的阈值

图 4　双频 CCD 监测器阈值

Fig. 4　Threshold of dual-frequency CCD monitor

如图 4(b) 所示。受 IFB 不确定性的影响, 双频 CCD 监测器的阈值增大了 26.1%。

2 种情况的阈值设定对比如表 1 所示, 即 IFB 不确定性导致双频 CCD 监测器的阈值增大了 0.12 m/s。由于使用k_{ffd}膨胀使监测器满足完好性要求, 所以二者在标称情况下的误警率不受影响。但是对于 CCD 故障情况, 阈值的增大则会导致监测性能明显下降。

表 1　监测阈值比较

Table 1　Comparison of monitoring threshold

仿真试验场景	原始标准差/(10^{-4} m·s^{-1})	包络标准差/(10^{-4} m·s^{-1})	阈值/(10^{-4} m·s^{-1})
包含IFB不确定性	6.22	10	58
扣除IFB不确定性	4.17	7.88	46

4　故障情况下的监测性能

CCD 故障以均值偏移故障为主[22], 且出现故障的概率非常小, 所以基于仿真试验, 对 CCD 均值偏移情况下双频 CCD 监测器性能受 IFB 不确定性的影响进行验证分析。

4.1　漏检率

对于给定的 CCD 均值偏移故障, PMD 定义为故障条件下检验统计量小于监测阈值的概率。3.2 节提到的 2 种检验统计量对应的 2 个监测阈值, 其对应的 PMD 如图 5 所示。

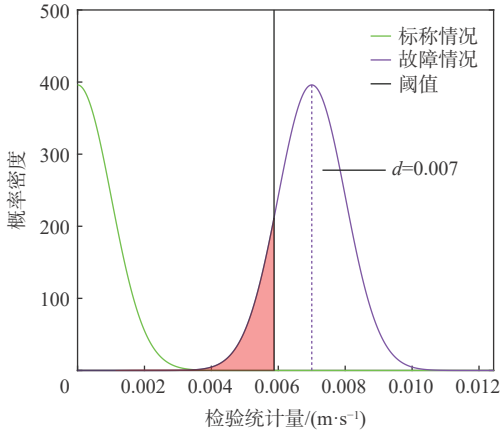
以故障 d=0.007 m/s 为例, 图 5 中 PMD 对应紫色概率分布中位于黑色实线左侧的概率, 即图中红色区域。对比图 5(a) 和图 5(b), 对于给定的 0.007 m/s 均值偏移故障, 使用包含 IFB 不确定性的原始检验统计量计算阈值, 双频 CCD 监测器的 PMD 达到 11.68%; 而使用扣除 IFB 不确定性的检验统计量计算阈值, PMD 仅为 0.12%。不同故障情况下的 PMD 对比如表 2 所示。

表中: PMD_1表示包含 IFB 不确定性的仿真场景对应的漏检率, PMD_2表示扣除 IFB 不确定性的仿真场景对应的漏检率。表 2 数据表明, 对于同一 CCD 故障, IFB 不确定性会导致双频 CCD 监测器的 PMD 显著增大, 尤其对于偏移较小的故障, 增大的更加明显。

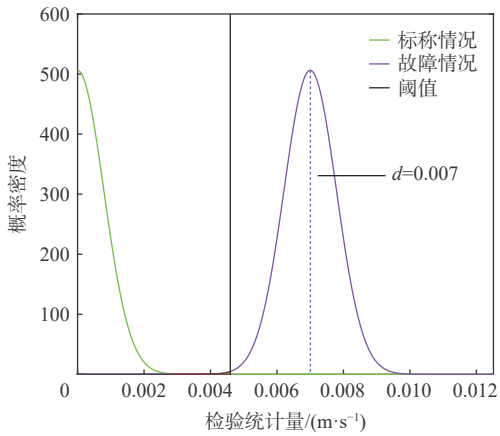
4.2　最小可检测故障

最小可检测故障表示在给定的 PMD 要求下, 监测器能够检测出的最小故障值, 其表征监测器的检测灵敏度。

图 6 为 2 种监测阈值对应的双频 CCD 监测器的最小可检测故障值, 其值为图中曲线与 10^{-9}PMD 对应横轴的交点, 其中蓝色曲线对应扣除 IFB 不确定性的监测器, 其最小可检测故障为 0.009 3 m/s; 红



(a) 包含IFB不确定性时的PMD



(b) 扣除IFB不确定性后的PMD

图 5 PMD 对比

Fig. 5 Comparison of PMD

表 2 不同监测阈值对应的漏检率对比

Table 2 Comparison of PMD related to different monitoring thresholds

监测阈值/(m·s^{-1})	PMD_1	PMD_2
7×10^{-5}	11.68	0.12
8×10^{-5}	1.39	7.99×10^{-4}
9×10^{-7}	6.87	1.18×10^{-4}
1×10^{-7}	1.33	3.62×10^{-7}
1.1×10^{-10}	9.96	2.29×10^{-8}

色曲线对应包含 IFB 不确定性的监测器，其最小可检测故障为 0.011 8 m/s，增大了 26.9%。

对比二者可以得出结论，IFB 不确定性导致双频 CCD 监测器的最小可检测故障值增大，即监测器的灵敏性降低。

4.3 完好性损失概率

在分析系统的完好性损失概率之前，需要对 CCD 故障导致的误差进行分析。由于 GAST-F 涉及 2 个频率下的伪距和载波相位观测量，具有多种故障模式，在此只对单故障模式进行分析，即某一频率的伪距或载波相位观测量出现故障。多故障

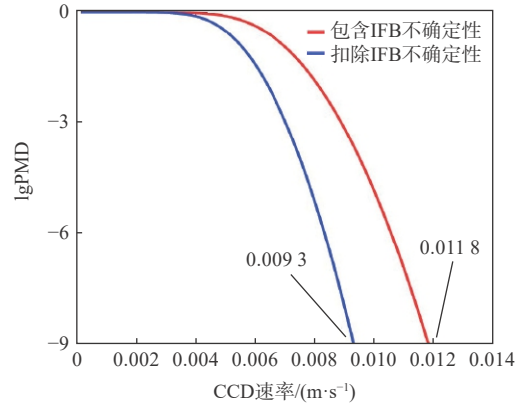


图 6 最小可检测故障对比

Fig. 6 Comparison of minimum detectable fault

模式在大多数情况下是单故障模式的线性叠加。CCD 误差分析的仿真条件如表 3 所示。

表 3 CCD 故障误差分析的仿真条件

Table 3 Simulation conditions of CCD fault error analysis

参数	数值
CCD监测器滤波时间常数τ_m/s	30
Hatch滤波的时间常数τ/s	100
CCD故障发生的时刻t_f/s	450
地面双频CCD监测器初始化时刻$t_{g,0}$/s	0
机载双频CCD监测器初始化时刻$t_{a,0}$/s	600
从机载滤波器初始化到将测量值应用于定位解的时间延迟t_{delay}/s	0, 50
地面端延迟τ_G/s	1.5
数据及完好性信息的更新率f_C/Hz	2
仿真时间t/s	[0∶0.5∶1 800]
故障d/(m·s^{-1})	0.001

监测器初始化是指由于失锁等原因导致的监测器重启初始化。当$t_{a,0}>t_f$，即 CCD 故障发生在机载监测器初始化之前，此时监测效果处于最坏情况，因此，仅分析双频 CCD 监测器在最坏情况下的监测性能。仿真时间 t 为 1 800 s，以 0.5 s 步进。置入故障 d=0.001 m/s，对应于双频 CCD 监测器的最小可监测故障值。

地面端和机载端都会对 CCD 故障进行监测，图 7 为由 CCD 故障导致的平滑误差，MC1 表示 B1I 频率伪距出现 CCD 故障导致的误差，MC3 表示 B3I 频率伪距出现 CCD 故障导致的误差，MP1 表示 B1I 频率载波相位出现 CCD 故障导致的误差，MP3 表示 B3I 频率载波相位出现 CCD 故障导致的误差。对比图中各条曲线可以看出，无论是在机载端还是地面端，MC1 和 MC3 故障造成的平滑误差更大，且误差不会收敛；相比而言，MP1 和 MP3 故障造成的平滑误差要小的多，并且误差最终会收敛为一个恒定值。

图 8 为机载端使用地面端广播的差分校正量后

得到的由于 CCD 故障导致的差分误差。对比图 8(a)和图 8(b)，当t_{delay}=0 时，差分误差需要经过 600 s 左右才会从最大值 2.35 m 收敛为 0，而如果在机载端应用了t_{delay}=50 s，差分误差将大幅度减小，最大值为 1.39 m，但是误差收敛到 0 的时间没有明显变化。基于以上分析得到的差分误差，进一步分析 IFB 不确定性影响下系统完好性损失的概率。

故障情况下，系统完好性损失的概率定义为：

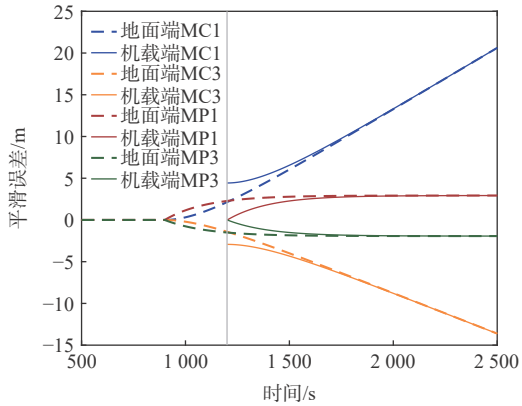
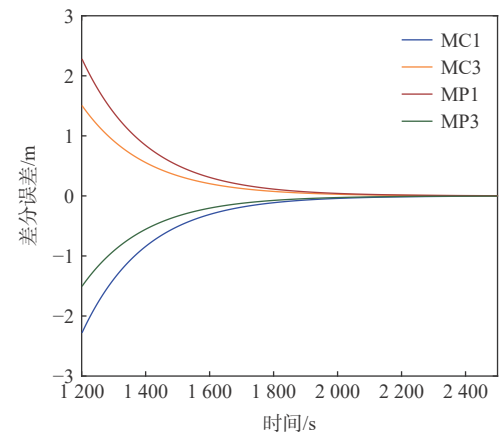
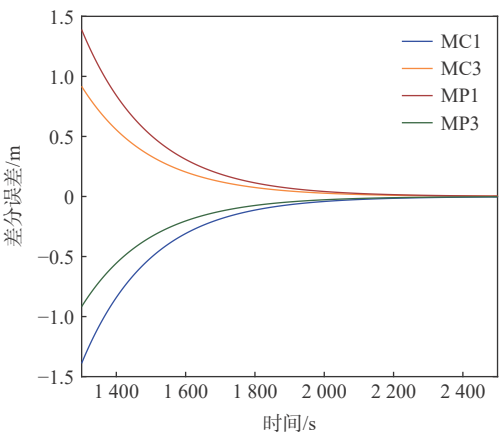


图 7　地面端和机载端平滑误差

Fig. 7　Smoothing error of ground terminal and airborne terminal



(a) 机载端不引入t_{delay}



(b) 机载端引入t_{delay}=50 s

图 8　差分误差

Fig. 8　Difference error

定位误差大于保护级且监测器检验统计量小于阈值的概率，如式 (23) 所示[6]：

$$P(L|f_k)=P(|e_v|>V|f_k)\cdot P\left(|q_k|<k_{\mathrm{ffd,mon}}\sigma_{r,k}|f_k\right) \tag{23}$$

式中：e_v为差分定位误差的垂直分量；V为垂直保护级；q_k为监测器对于卫星 k 的检验统计量；$k_{\mathrm{ffd,mon}}\sigma_{r,k}$为监测阈值。

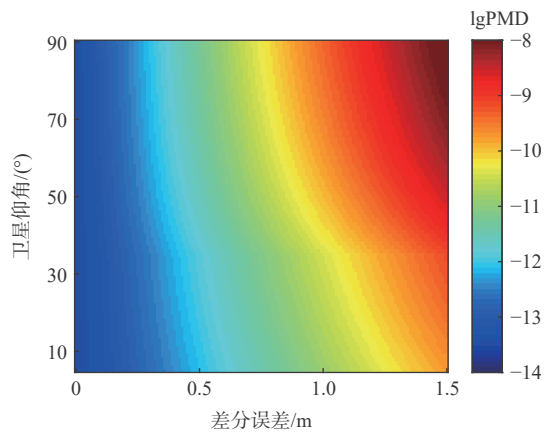
将式 (23) 化简，最终得到[6]

$$P\left(L|f_k\right)=Q\left\{\frac{|b_k|}{\sigma_k}-k_{\mathrm{ffmd}}\right\}\cdot Q\left\{k_{\mathrm{ffd,mon}}-\frac{|r_k|}{\sigma_{r,k}}\right\} \tag{24}$$

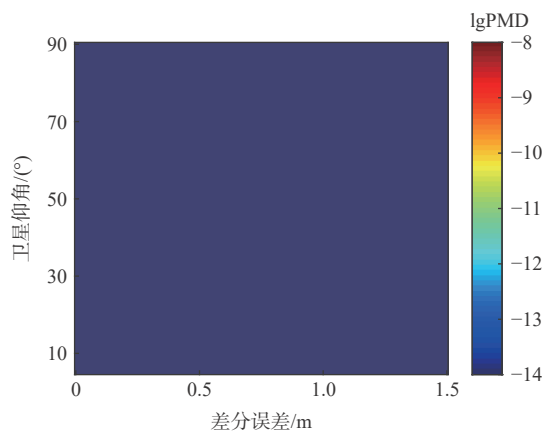
式中：Q为标准正态分布的右尾函数(Q 函数)；b_k为由于故障导致的差分误差；σ_k为卫星 k 用于计算保护级的非相关误差项；r_k为故障值；$\sigma_{r,k}$为标称情况下检验统计量的标准差。

对于双频 CCD 监测器，设定的仿真条件为：b_k=[0:0.01:1.5] m，r_k=0.01 m/s。

图 9 为完好性损失概率。图 9(a) 为使用包含 IFB 不确定性的检验统计量构造的双频 CCD 监测器对应的完好性损失概率，横坐标变量为差分误差E_R，纵坐标变量为卫星仰角，不同颜色代表概率大



(a) 包含IFB不确定性时



(b) 扣除IFB不确定性后

图 9　完好性损失概率

Fig. 9　Probability related to loss of integrity

小。图 9(a) 中大部分区域能够满足 10^{-9} 的完好性要求，但是在高仰角位置却达到了 10^{-8}，不满足完好性要求；当从检验统计量中扣除掉 IFB 不确定性后，如图 9(b) 所示，在最坏情况下完好性损失概率也在 10^{-14} 以下，符合完好性要求。对比结果表明，IFB 不确定性导致双频 CCD 监测器对应的系统完好性损失概率上升，给双频 GBAS 带来了一定的完好性风险。

4.4 PMD 性能安全验证

ICAO 提出伪距域 PMD 要求的目的是对航空导航领域的各类监测器进行性能测试，以检验其是否能满足飞行器安全需求。对于特定的差分误差 E_{R}，PMD 计算式为[14]

$$\mathrm{PMD}(|E_{\mathrm{R}}(t)|)=\mathrm{PMD}\left(f^{-1}(E_{\mathrm{R}}(t+T_{\mathrm{TTDABA}}-\tau_{\mathrm{G}}))\right) \tag{25}$$

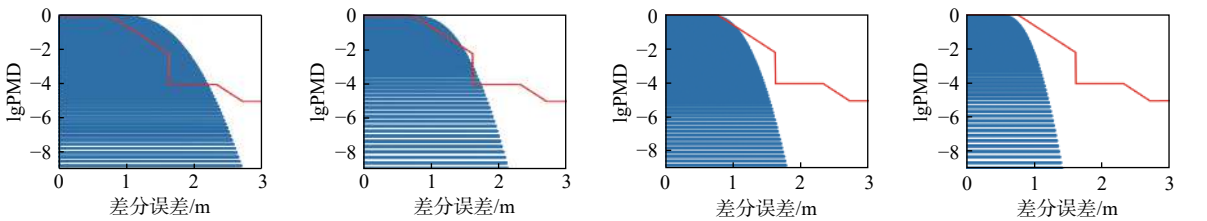
式中：f^{-1}为概率 1 减去对应的 Q 函数；T_{TTDABA}为差分误差达到E_{R}的时间点与地面站发送最后 1 比特完好性信息的时间点之间的差值，GAST-D 中规定的告警时间是 2.5 s，TTADBA 要求值为 1.5 s，留 1 s 裕度给可能出现的延迟和消息丢失，GAST-F 延用此值[14]。为了验证 IFB 不确定性对于双频 CCD 监测器性能安全的影响，开展了仿真试验，仿真条件与表 4 中所示基本相同，唯一改变的是置入故障 d 改为从 0～0.02 m/s 以 0.000 1 m/s 步进，即 d=[0:0.000 1:0.02] m/s。

表 4 为满足 PMD 要求在机载端引入的最小t_{delay}值

Table 4 Minimum t_{delay} introduced in airborn to meet PMD requirements

故障模式	包含IFB不确定性/s	扣除IFB不确定性/s
MC1 (MP1)	50	22
MC3 (MP3)	6	0

图 10 为t_{delay}=0 时对应的双频 CCD 监测器性能安全验证结果，其中红色曲线为 ICAO 规定的 PMD 要求，蓝色曲线为监测器对应的 PMD 值，不超过红色曲线便满足性能要求。从图 10 可以看出，如果机载端t_{delay}=0，在 MC1 或 MP1 故障模式下，无论是否扣除 IFB 不确定性，双频 CCD 监测器的性能安全都不符合 PMD 要求；在 MC3 或 MP3 故障模式下，包含 IFB 不确定性的双频 CCD 监测器仍不满足 PMD 要求，但是扣除 IFB 不确定性后双频 CCD 监测器的却可以满足 PMD 要求。可见，如果t_{delay}=0，则 IFB 不确定性会导致双频 CCD 监测器在 MC3 或 MP3 故障模式下的 PMD 符合性发生变化，机载端需要引入一定值的t_{delay}。

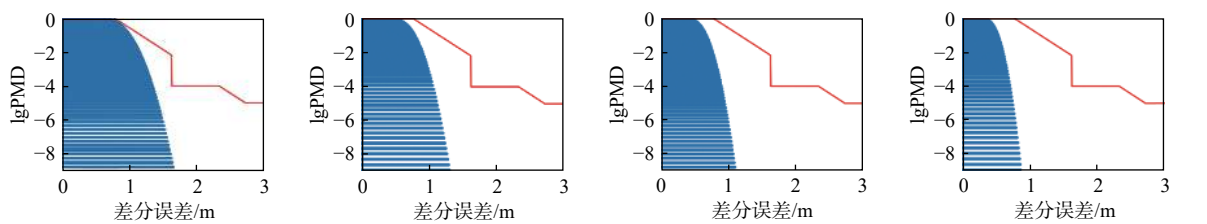


(a) 包含IFB不确定性 (MC1/MP1) (b) 扣除IFB不确定性 (MC1/MP1) (c) 包含IFB不确定性 (MC3/MP3) (d) 扣除IFB不确定性 (MC3/MP3)

图 10 监测器性能安全验证(t_{delay}=0)

Fig. 10 Monitor performance safety verification (t_{delay}=0)

图 11 为t_{delay}=50 s 时对应的双频 CCD 监测器性能安全验证结果，此时所有故障情况下都满足 PMD 要求。对比图 11(a) 和图 11(b) 可以看出，在 MC1 或 MP1 故障模式下，扣除 IFB 不确定性之后的双频 CCD 监测器性能安全曲线距离 PMD 要求曲线有一定的裕度，这说明t_{delay}可以适当减小，进一步仿真表明此时满足 PMD 要求的t_{delay}最小值为 22 s。对比图 11(c) 和图 11(d) 可以得出相似结论。



(a) 包含IFB不确定性 (MC1/MP1) (b) 扣除IFB不确定性 (MC1/MP1) (c) 包含IFB不确定性 (MC3/MP3) (d) 扣除IFB不确定性 (MC3/MP3)

图 11 监测器性能安全验证(t_{delay}=50 s)

Fig. 11 Monitor performance safety verification (t_{delay}=50 s)

t_{delay}的大小表征了双频 CCD 监测器的响应时间，t_{delay}越小，监测器对于 CCD 故障的检测时间越短。仿真结果表明，为满足 ICAO 对 PMD 的要求，IFB 不确定性会导致机载端t_{delay}值增大，即导致双

频 CCD 监测器的检测时间变长, 响应变慢。各种故障情况下机载端引入的最小t_{delay}值如表 4 所示。机载端所需引入的最小t_{delay}值对于 B1I 和 B3I 频率单故障分别增大了 28 s 和 6 s。

5 结　论

1) 受 IFB 不确定性的影响, 双频 CCD 监测器的阈值T_{CCD}由 0.004 6 m/s 增长到 0.005 8 m/s, 增大了 26.1%。

2) 故障情况下, 双频 CCD 监测器的 PMD 随着故障值的减小而显著增大。

3) 双频 CCD 监测器的最小可检测故障值增大了 26.9%, 灵敏度降低。

4) 最坏情况下, 系统完好性损失的概率从低于 10^{-14} 增大到接近 10^{-8}。

5) 为满足 PMD 安全要求, 机载端所需要引入的最小t_{delay}值对于 B1I 和 B3I 频率单故障分别增大了 28 s 和 6 s, 双频 CCD 监测器的响应变慢。

参考文献（References）

[1] 谢钢. GPS原理与接收机设计[M]. 北京: 电子工业出版社, 2017.
XIE G. Principles of GPS and receiver design[M]. Beijing: Publishing House of Electronics Industry, 2017 (in Chinese).

[2] Radio Technical Commission for Aeronautics (RTCA) Special Committee 159 (SC-159). Minimum operational performance standards for GPS local area augmentation system airborne equipment: DO-253C[S]. Washington, D. C. : RTCA, 2008.

[3] Radio Technical Commission for Aeronautics (RTCA) Special Committee 159 (SC-159). Minimum operational performance standards for GPS local area augmentation system airborne Equipment: DO-253D[S]. Washington, D. C. : RTCA, 2017.

[4] International Civil Aviation Organization (ICAO) Navigation Systems Panel (NSP). GBAS CAT II/III development baseline SARPs[S]. Montreal: ICAO, 2010.

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

[6] SIMILI D V, PERVAN B. Code-carrier divergence monitoring for the GPS local area augmentation system[C]// 2006 IEEE/ION Position, Location, and Navigation Symposium. Piscataway: IEEE Press, 2006.

[7] European Organization for Civil Aviation Equipment (EUROCAE). Minimum operational performance specification for global navigation satellite ground based augmentation system ground equipment to support category I operations: ED-114A[S]. SaintDenis: EUROCAE, 2013.

[8] KONNO H. Dual-frequency smoothing for CAT III LAAS performance assessment considering ionosphere anomalies[C]//Proceedings of the 20th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2007). Fort Worth: The Institute of Navigation, 2007: 424-437.

[9] SUNG Y T, LIN Y W, YEH S J, et al. A dual-frequency ground based augmentation system prototype for GPS and BDS[C]// Proceedings of the 32nd International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2019). Fort Worth: The Institute of Navigation, 2019: 628-636.

[10] CIRCIU M S, MEURER M, FELUX M, et al. Evaluation of GPS L5 and Galileo E1 and E5a performance for future multifrequency and multiconstellation GBAS[J]. Navigation, 2017, 64(1): 149-163.

[11] FELUX M, CIRCIU M, BELABBAS B, et al. Concept for a dual frequency dual constellation GBAS[C]// Proceedings of the 28th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2015). Fort Worth: The Institute of Navigation, 2015: 1519-1525.

[12] KONNO H, PULLEN S, RIFE J, et al. Ionosphere monitoring methodology for hybrid dual-frequency LAAS[C]// Proceedings of the 19th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2006). Fort Worth: The Institute of Navigation, 2006: 409-424.

[13] KONNO H, PULLEN S, RIFE J, et al. Evaluation of two types of dual-frequency differential GPS techniques under anomalous ionosphere conditions[C]// Proceedings of the 2006 National Technical Meeting of the Institute of Navigation. Fort Worth: The Institute of Navigation, 2006: 735-747.

[14] JIANG Y P, MILNER C, MACABIAU C. Code carrier divergence monitoring for dual-frequency GBAS[J]. GPS Solutions, 2017, 21(2): 769-781.

[15] BAO Y D, LI J, LIU H N. Extraction and analysis of DCB daily variation based on single station dual-frequency observations[C]// Proceedings of the 9th China Satellite Navigation Academic Annual Conference. Berlin: Springer, 2018.

[16] MA G, MARUYAMA T. Derivation of TEC and estimation of instrumental biases from GEONET in Japan[J]. Annales Geophysicae, 2003, 21(10): 2083-2093.

[17] HOLASCHUTZ D, BISHOP R H, HARRIS R B, et al. Inter-frequency bias estimation for the GPS monitor station network[C]// International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2008). Savannah: The Institute of Navigation, 2008: 2405-2415.

[18] LI Z S, YUAN Y B, LI H, et al. Two-step method for the determination of the differential code biases of COMPASS satellites[J]. Journal of Geodesy, 2012, 86(11): 1059-1076.

[19] LIU Y H, LI X H, ZHANG H J, et al. Calculation and accuracy evaluation of TGD from IFB for BDS[J]. GPS Solutions, 2016, 20(3): 461-471.

[20] MURPHY T, GEREN P, PANKASKIE T. GPS antenna group delay variation induced errors in a GNSS based precision approach and landing systems[C]// 20th International Technical Meeting of the Satellite Division. Fort Worth: The Institute of Navigation, 2007: 2974-2989.

[21] BEER S, WANNINGER L, HEBELBANTH A. Estimation of absolute GNSS satellite antenna group delay variations based on those of absolute receiver antenna group delays[J]. GPS Solutions, 2021, 25(3): 1-10.

[22] THOMAS Z. A new evaluation of maximum allowable errors and missed detection probabilities for LAAS ranging source monitors[C]// Proceedings of the 58th Annual Meeting of The Institute of Navigation and CIGTF 21st Guidance Test Symposium (2002). Fort Worth: The Institute of Navigation, 2002: 187-194.

Impact of IFB uncertainty on dual-frequency code carrier divergence monitoring

KANG Kai[1], WANG Zhipeng[1], FANG Kun[2,*], ZHU Yanbo[1]

(1. School of Electronic Information Engineering, Beihang University, Beijing 100191, China;
2. Research Institute for Frontier Science, Beihang University, Beijing 100191, China)

Abstract: Code carrier divergence (CCD) monitor is one of the integrity monitors introduced by ground-based augmentation systems (GBAS), which is used to monitor the inconsistency between code pseudorange and carrier phase observations. Dual-frequency smoothing technology changes the test statistics and other parameters of CCD monitor, which affect the monitoring performance. Considering the inter-frequency bias (IFB) introduced in dual-frequency GBAS, the impact of IFB uncertainty on the dual-frequency CCD monitoring is analyzed based on the dual-frequency observation data of BDS B1I and B3I signals. The results show that under the influence of IFB uncertainty, the threshold of the monitor increases by 26.1%, resulting in a significant increase of the probability of missed detection (PMD) in the case of CCD fault. And the minimum detectable fault increases by 26.9%, which means a decrease in the sensitivity of the monitor. Meanwhile, the probability of the loss of integrity in the worst case increases from less than 10^{-14} to nearly 10^{-8}, and the delay introduced by the airborne to meet the PMD requirement is larger, resulting in a slower response of CCD monitor and impacts the integrity of dual-frequency GBAS.

Keywords: inter-frequency bias; code carrier divergence; ground-based augmentation systems; dual-frequency monitor; integrity

Received: 2021-10-31; **Accepted**: 2022-02-13; **Published Online**: 2022-02-21 09:00

URL: kns.cnki.net/kcms/detail/11.2625.V.20220219.1405.001.html

Foundation items: National Key R & D Program of China (2020YFB0505602); National Natural Science Foundation of China (61871012,62022012); Civil Aviation Security Capacity Building Fund Project (CAAC Contract 2021(77),CAAC Contract 2020(123)); Beijing Nova Program of Science and Technology (Z191100001119134)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0652

模块化抛物面折展机构设计与优化

何天宇[1]，董洋[1]，王辉[1]，赵军鹏[1,*]，王春洁[1,2]

(1. 北京航空航天大学 机械工程及自动化学院，北京 100191；
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 100191)

摘 要：抛物面天线作为卫星上重要的功能组件，是实现遥感测量和无线通信等功能的基础。为了解决模块化抛物面天线折展机构构型设计与刚度匹配困难的问题，提出一种基于剪刀机构的肋单元折展机构的构型方法，并设计了肋单元折展机构。通过运动学分析，验证了该机构通过参数的修改，可以实现同一折展机构展开状态包络不同抛物面拟合球面。给出一种包络锥法实现多模块折展机构展开过程的求解思路，为可转动肋单元的运动副配置提供依据。构建多模块折展机构的有限元模型，以提高固有频率和降低整体质量为优化目标，对肋单元折展机构的构型参数进行优化。采用非劣排序遗传算法完成优化迭代计算,得到具有高刚度、轻量化的折展机构设计参数，在相同条件下优化后调和平均固有频率提高了67.4%，天线整体质量下降了35.2%。

关 键 词：模块化抛物面天线；构型综合；运动学分析；有限元分析；优化设计
中图分类号：V443+.4; TH122
文献标志码：A **文章编号**：1001-5965(2023)09-2473-09

模块化抛物面可展开天线具有口径大、展收比大、型面精度高、可拓展能力强、质量轻等特点,受到广泛研究与应用[1]。运用模块化的思想根据需求可调整模块的组合形式,使得整体的设计思路多元化。由于模块化抛物面天线机构部件多、耦合关系复杂,因此,需要将设计与优化方法应用于整个天线产品的研制过程中。

文献[2-3]阐述了构建模块化抛物面天线的设计理论。Onoda等[4]提出一种适合空间应用的锁定机构数量少、展收比高的二维可展桁架结构,并设计了桁架平台进行试验验证。Guo等[5]针对抛物面反射面的几何分割问题,提出一种基于对称六边形分割的模块化桁架式展开天线机构设计方法。Xu等[6]提出一种曲面天线机构最小组合单元的划分方法,使最终折叠状态下所有节点的姿态保持一致,得到可靠性和稳定性更好的天线展开机构。刘兆晶[7]针对变形模块之间连接会出现偏差的情况,提出空间等包络圆法,构建多层模块的准确连接。陈向阳和关富玲[8]提出一个利用伸缩杆驱动的六棱柱展开单元,整体机构简单,展开可靠度高,结构刚度好。岳建如等[9]利用拉伸弹簧构造了一个可展伸缩杆六棱柱元,展开后所有展开机构均不需锁定,具有良好的几何适应性。Hu等[10-11]建立3RR-3RRR四面体单元的参数化模型和约束矩阵,分析了单元的可延拓性。崔吉[12]和余丹敏[13]分别归纳现有空间可展开天线结构,构建新型单元模块和抛物面天线支撑桁架模型,并比较各种模块的运动学动力学特性及展开后的结构特性。田大可等[14]基于误差反向传播神经网络和遗传算法对多模块构架式天线的结构参数进行了优化。史创[15]提出基于可展开四棱柱单元的索杆张拉式天线机构并综合优化设计。王辉等[16]提出一种新型的张紧绳索多层设计方法,通过优化张紧力参数,提高天线结构刚度的同时尽量降低张紧力引起的形面误差。

收稿日期：2021-10-31；**录用日期**：2022-01-25；**网络出版时间**：2022-02-22 09:37
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220221.1844.002.html
基金项目：国家自然科学基金(51635002)
*通信作者. E-mail：zhaojunpeng@buaa.edu.cn

引用格式：何天宇，董洋，王辉，等. 模块化抛物面折展机构设计与优化[J]. 北京航空航天大学学报，2023，49(9)：2473-2481.
HE T Y，DONG Y，WANG H，et al. Design and optimization of modular parabolic deployable mechanism[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2473-2481 (in Chinese).

文献 [1-16] 的研究主要集中在天线的结构优化、构型综合、动力学建模仿真及形面精度调整等方面，对于模块化抛物面折展机构的基本模块和基础肋单元机构的设计与优化研究尚不充分。对此，本文提出一种基于剪刀机构的新型肋单元展开机构构型设计方法，并通过包络锥法解算多模块展开机构的展开过程，进一步通过优化肋单元构型参数提高多模块展开机构的整体性能。

1 模块化抛物面折展机构总体构型

本文提出的折展机构由多个基本模块组成，每个基本模块包括 6 个围绕中心支柱分布的肋单元，图 1 为模块化抛物面折展机构组成示意图。肋单元与基本模块公共支柱连接的部分为肋单元公共端，另一端为肋单元自由端。在多模块折展机构中，肋单元自由端可以与其他基本模块的肋单元自由端通过铰链连接或悬空。相互连接的肋单元自由端共线，顶端滑块上侧与底端固定块下侧各有一个铰链，铰链轴线沿自由端轴线方向，提供该方向转动自由度。当有多个自由端铰接于同一位置时，铰链采用同轴分层布置的形式。该构型采用球面来代替抛物面确定肋单元接点，而非保证展开后的肋单元接点均在抛物面上。为保证工作表面为抛物面，工作表面与拟合球面的误差用支撑杆来补偿。

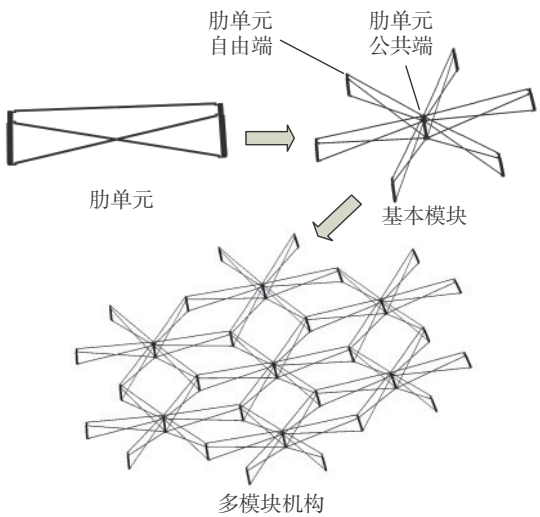


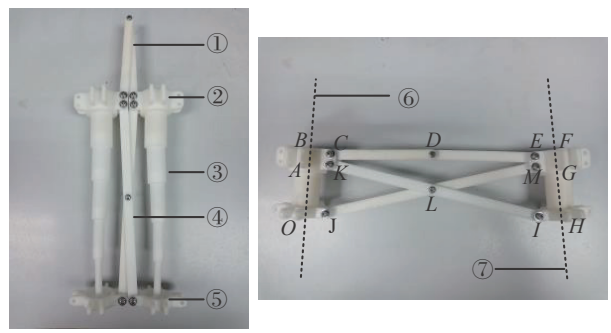
图 1 模块化抛物面折展机构组成示意图

Fig. 1 Composition diagram of modular parabolic deployable mechanism

2 肋单元展开机构设计

2.1 肋单元构型设计

本文提出一种基于剪刀机构的肋单元展开机构构型设计方法，并设计了一种以口径任意变化的肋单元机构。图 2 为肋单元折展机构实物图，其由 2 根剪刀杆、2 根折杆、2 块顶端滑块、2 块底端固定块及 2 组多级套筒组成。剪刀杆两端分别与两侧的顶端滑块和底端固定块转动副连接。折杆一端与另一折杆转动副连接，另一端与顶端滑块转动副连接，折杆在肋单元折展机构完全展开后用于提供抛物面形面拟合中间插值点，并确定展开最终状态，起到对剪刀机构的辅助限位作用。两侧顶端滑块与底端固定块均通过多级套筒提供的移动副连接。通过顶端滑块与底端固定块之间的移动副运动实现肋单元展开机构在一个平面内的折展过程。此外，多级套筒在肋单元展开状态下收拢，约束顶端滑块和底端固定块位置，提高展开状态下的结构刚度。



(a) 收拢状态 (b) 展开状态

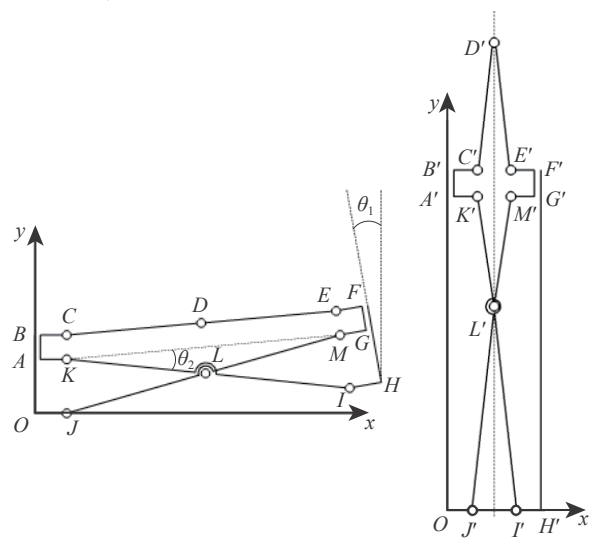
①折杆；②顶端滑块；③多级套筒；④剪刀杆；⑤底端固定块；⑥公共端轴线；⑦自由端轴线

图 2 肋单元折展机构实物图

Fig. 2 Physical picture of rib unit deployable mechanis

为满足天线展开机构展开后包络曲面拟合抛物面的要求，肋单元展开后两侧滑块移动副所在直线需呈一定夹角θ_1。另外肋单元收拢状态两侧滑块移动副所在直线需平行。

根据展开运动需求，采用解析法对肋单元进行尺度综合，如图 3 所示。



(a) 展开状态 (b) 收拢状态

图 3 肋单元折展机构展开收拢状态简图

Fig. 3 Schematic diagram of unfolding and folding state of rib unit deployable mechanism

设肋单元展开状态, 点 B、点 F 在抛物面拟合球面半径为 R 的大圆上, 弦 BF(肋单元跨距) 对应的圆心角为θ_1。由肋单元折展机构的对称性可以得到: $l_{BC}=l_{EF}$、$l_{CD}=l_{DE}$、$l_{AK}=l_{MG}$、$l_{OJ}=l_{IH}$、$l_{KL}=l_{LM}$、$l_{JL}=l_{LI}$。为使得杆件种类减少、杆件尺寸简单, 令 $l_{BC}=l_{AK}$。由肋单元收拢状态和展开状态杆件尺寸关系可得

$$\begin{cases} l_{CD}=R\sin\theta_1/2-l_{BC}\cos\theta_1/2 \\ \theta_2=\arccos\dfrac{(R+l_{AB})\sin\theta_1/2-l_{BC}\cos\theta_1/2}{l_{KL}} \\ \dfrac{l_{JL}}{l_{KL}}=\dfrac{2l_{BC}-2l_{OJ}+l_{C'E'}}{l_{C'E'}} \\ \dfrac{l_{JL}\cos\theta_2+l_{OJ}\cos\theta_1/2}{\tan\theta_1/2}=(l_{JL}+l_{KL})\sin\theta_2+ \\ \qquad (l_{BC}-l_{OJ})\sin\theta_1/2+(R+l_{AB})\cos\theta_1/2 \end{cases} \tag{1}$$

式中: θ_1、R、l_{OJ}、l_{BC}、l_{AB}、$l_{C'E'}$为确定的独立设计参数, 由式 (1) 可解出l_{CD}、l_{KL}、l_{JL}、θ_2, 进一步求出全部杆件尺寸关系。

2.2　肋单元运动学分析

将肋单元折展机构的驱动设置为顶端滑块(图 3 中 AB 点之间) 沿多级套筒轴向的移动副运动, 即可通过解析方式求解滑块移动到任意位置的肋单元关键坐标点位置。由于肋单元的对称性, 展开过程中肋单元始终存在中心对称轴 DL。

由肋单元展开任意时刻, A 点坐标为$(0,l_{OA})$得到以下坐标点$B(0,l_{OA}+l_{AB})$、$C(l_{BC},l_{OA}+l_{AB})$、$K(l_{BC},l_{OA})$、$J(l_{OJ},0)$。

根据杆件尺寸可求解得到 $L(x_L,y_L)$为

$$\begin{cases} (x_K-x_L)^2+(y_K-y_L)^2=l_{KL}^2 \\ (x_J-x_L)^2+(y_J-y_L)^2=l_{JL}^2 \end{cases} \tag{2}$$

由矢量关系可得

$$\begin{cases} x_M=x_L+\dfrac{l_{KL}}{l_{JL}}(x_L-x_J) \\ y_M=y_L+\dfrac{l_{KL}}{l_{JL}}(y_L-y_J) \end{cases} \tag{3}$$

求解得到 $M(x_M,y_M)$, 并求出 KM 中点 $N(x_N,y_N)$。用 N、L 坐标构造肋单元中心对称轴 DL 的解析式为

$$\frac{y_N-y_L}{x_N-x_L}x-y+\frac{y_N-y_L}{x_N-x_L}x_L+y_L=0 \tag{4}$$

由对称关系可以求得 E、F、G、H、I 点坐标。可根据式 (5) 求解得到 D 点坐标为

$$\begin{cases} (x_C-x_D)^2+(y_C-y_D)^2=l_{CD}^2 \\ (x_D-x_E)^2+(y_D-y_E)^2=l_{CD}^2 \end{cases} \tag{5}$$

通过本节关键点求解, 得到顶端滑块在驱动行程内运动的肋单元展开轨迹, 如图 4 所示。结果显示机构展开过程平稳, 能实现收拢和展开状态的平稳过渡。

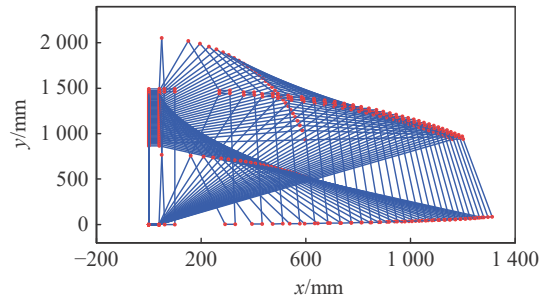


图 4　肋单元折展机构展开轨迹图

Fig. 4　Unfolding track diagram of rib unit deployable mechanism

确定单个肋单元跨距即完全展开状态下 BF 长度的参数为拟合球面半径 R 和两侧滑块移动副所在直线所呈夹角θ_1。在肋单元展开过程中, 每一状态下的 BF 长度均对应不同的过程拟合球面半径 R'和过程两侧滑块所在直线夹角θ_1', 表达式如式 (6) 所示, 二者随 AB 行程变化的曲线如图 5 所示。

$$\begin{cases} \theta_1'=\arctan\dfrac{x_H-x_G}{y_G-y_H} \\ R'=\dfrac{l_{BF}}{2\tan\theta_1'/2} \end{cases} \tag{6}$$

肋单元在展开过程中瞬时中间态可以看作是另一拟合球面 R'下的展开末状态, 中间每一状态都对应不同的拟合球面, 折展全过程对应的拟合球面半径的取值范围为$[R,+\infty)$。只需通过修改肋单元机构中铰链点 D 的位置即可实现跨距 BF 的修改, 达到同一折展机构展开状态包络不同抛物面拟合球面的效果。

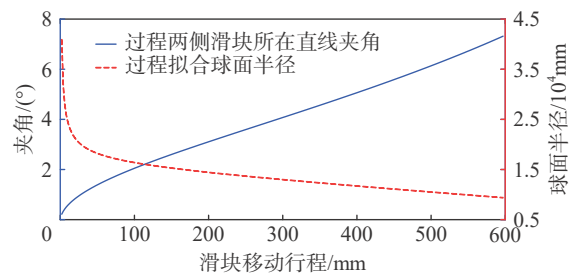


图 5　过程拟合球面半径和过程两侧滑块所在直线夹角随滑块行程变化曲线

Fig. 5　Curves of fitting spherical radius and included angle of line changed with slider stroke

3　多模块折展机构设计

3.1　多模块包络锥构建

多模块折展机构由多个基本模块分层布置拼接组成。相邻基本模块的肋单元自由端(图 3 中 FH 轴线)共线, 通过顶端滑块上侧 F 处和底端固定块下侧 H 处的 2 个铰链连接, 传递各肋单元顶端滑块的运动。不与任何基本模块连接的肋单元自由端则悬空。展开过程中每个基本模块的肋单元同

一时刻展开状态相同，所有基本模块的关键点（B点及F点）均应落在过程拟合球面上。由于正六边形无法密铺在球面表面，因此，除中心基本模块外，每个基本模块中6个肋单元之间的夹角并不均匀分布。文献[7]提出等包络圆法进行展开状态几何建模。本文在此基础上提出包络锥法，既可以用于多模块折展机构构型设计也可以针对展开过程进行运动学分析。

在任意展开状态下，以基本模块的肋单元公共端为旋转轴，将肋单元自由端所在直线作为母线绕轴旋转2π，得到的扫掠曲面即为该基本模块在该展开状态下的包络锥。所有包络锥的顶点重合于一点，即拟合球面的球心，图6为多模块机构包络锥示意图。各基本模块的肋单元自由端所在直线均为该模块包络锥的母线。改变基本模块中相邻肋单元所在平面的夹角只会造成肋单元自由端在包络锥曲面上的移动。

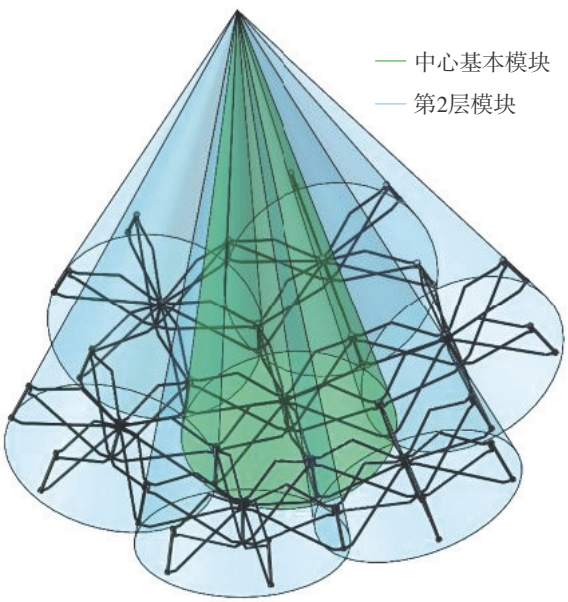


图6 多模块机构包络锥示意图

Fig. 6 Schematic diagram of envelope cone of multi- module mechanism

以双层七模块为例求解多模块折展机构中各基本模块包络锥，约定：

1）中心基本模块各相邻肋单元所在平面的夹角均为π/3。

2）与中心模块连接的任一第2层模块的2个肋单元所在平面间夹角为π/3。

3）任一第2层模块不与其他第2层模块连接的2个肋单元所在平面夹角为π/3。

4）任一第2层模块关于其肋单元公共端所在直线与中心模块肋单元公共端所在直线确定的平面对称。

此时第2层模块中有2个肋单元相对于原位置产生偏角τ。对应的肋单元称为可转动肋单元，其与基本模块内肋单元公共端相接的一侧需要添加一组转动副，以保证天线整体展开过程的实现。

首先构建中心基本模块展开过程任意时刻的包络锥。包络锥的半顶角为当前时刻两侧滑块所在直线夹角θ_1'，包络锥顶点为坐标系原点，可构建包络锥解析式为

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{T}\boldsymbol{X}=\boldsymbol{0} \tag{7}$$

式中：

$$\boldsymbol{T}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&-\tan\theta_1'\end{bmatrix}$$

$$\boldsymbol{X}=[x,y,z]^{\mathrm{T}},\ z\leqslant 0$$

围绕中心基本模块阵列分布的5个第2层基本模块对应的包络锥，可由中心基本模块包络锥经过2次绕坐标轴旋转变换得到。中心基本模块包络锥绕y轴逆时针旋转2η得到第2层1号包络锥，第2层1号包络锥绕z轴顺时针旋转$(n-1)\pi/3$得到第2层n（n=1, 2, ⋯, 6）号包络锥。可构建第2层n号包络锥解析式为

$$(\boldsymbol{R}_1\boldsymbol{R}_2\boldsymbol{X})^{\mathrm{T}}\boldsymbol{T}(\boldsymbol{R}_1\boldsymbol{R}_2\boldsymbol{X})=0 \tag{8}$$

式中：

$$\boldsymbol{R}_1=\boldsymbol{R}_y(2\eta)$$

$$\boldsymbol{R}_2=\boldsymbol{R}_z\left[-(n-1)\frac{\pi}{3}\right]$$

η可由中心基本模块包络锥和第2层1号包络锥几何关系求解得到。用任意长度R_r截取两包络锥，得到O^0O和O^1O两圆锥体并且两高相等，图7为

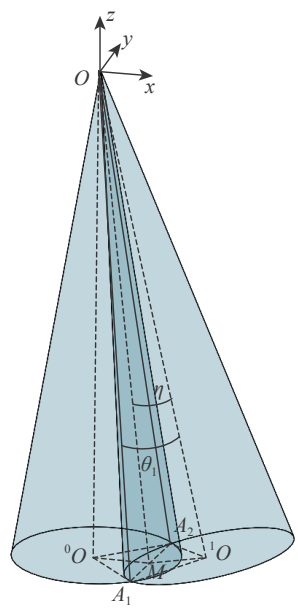


图7 中心基本模块和第2层模块1包络锥示意图

Fig. 7 Schematic diagram of envelope cone of central module and layer 2 module 1

中心基本模块和第 2 层模块 1 包络锥示意图。

由于中心基本模块的相邻肋单元所在平面夹角为 $\pi/3$, 故可得到 $\angle A_1{}^0OA_2=\pi/3$。由几何关系可以得到:

$$\begin{cases} {}^0OM = O^0O\tan\eta \\ {}^0OA_1 = O^0O\tan\theta_1' \\ {}^0OM = \cos\dfrac{\pi}{6}\,{}^0OA_1 \end{cases} \tag{9}$$

整理可得

$$\eta = \arctan\left(\frac{\sqrt{3}}{2}\tan\theta_1'\right) \tag{10}$$

因此, 在约定多模块机构组接方案后, 任一展开时刻各包络锥的参数与几何关系仅由当前时刻两侧滑块所在直线夹角 θ_1' 确定。

3.2　包络锥法多模块运动学求解

用包络锥联立求解各基本模块肋单元自由端所在直线解析式, 并求解第 2 层模块中 2 个肋单元相对于原位置产生偏角 τ 随天线展开的变化曲线。

如图 8 所示, 选取第 2 层模块的包络锥 i 和包络锥 i+1, 联立求解得到的交线即为第 2 层模块 i 和模块 i+1 的公共肋单元自由端所在直线。

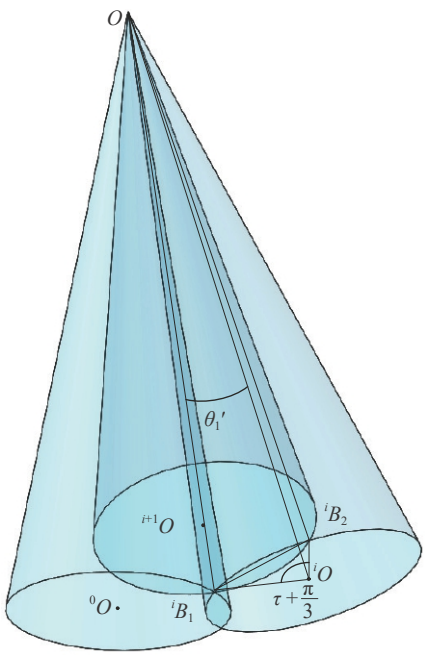


图 8　中心基本模块和第 2 层模块 i 及模块 i+1 包络锥示意图

Fig. 8　Schematic diagram of envelope cone of central module,layer 2 module i and module i + 1

由于第 2 层一共 6 个模块, 规定 i=6 时 i+1 表示第 1 个模块。由于所有包络锥共用同一个顶点 (即坐标系原点), 可用待定系数法表达交线方程:

$$\begin{cases} (\boldsymbol{R}_1\boldsymbol{R}_3\boldsymbol{X})^{\mathrm{T}}\boldsymbol{T}(\boldsymbol{R}_1\boldsymbol{R}_3\boldsymbol{X}) = 0 \\ (\boldsymbol{R}_1\boldsymbol{R}_4\boldsymbol{X})^{\mathrm{T}}\boldsymbol{T}(\boldsymbol{R}_1\boldsymbol{R}_4\boldsymbol{X}) = 0 \\ \boldsymbol{X} = \boldsymbol{A}_i x \quad i = 1,2,\cdots,6 \end{cases} \tag{11}$$

式中:

$$\boldsymbol{R}_1 = \boldsymbol{R}_y(2\eta)\ \boldsymbol{R}_3 = \boldsymbol{R}_z\left(-i\frac{\pi}{3}\right)$$

$$\boldsymbol{R}_4 = \boldsymbol{R}_z\left(-(i+1)\frac{\pi}{3}\right)$$

$$\boldsymbol{A}_i = [1,\ a_i,\ b_i]^{\mathrm{T}}$$

其中: a_i、b_i 为交线方程的待定系数。

通过改变肋单元的滑块行程可以得到天线展开各中间状态下的各基本模块间相连接肋单元自由端所在直线解析式。

用任意长度 R_r 截取第 2 层模块 i 和 i+1 的包络锥及交线, 可得到交点 ${}^iB_1$、${}^iB_2$ 和分别以 ${}^iO$、${}^{i+1}O$ 为圆心的两圆锥底面。

$$\begin{cases} ({}^i\boldsymbol{B}_j)^{\mathrm{T}}({}^i\boldsymbol{B}_j) = R_{\mathrm{r}}^2 \\ {}^i\boldsymbol{B}_j = \boldsymbol{A}_{ij}{}^ix_j \end{cases} \quad i = 1,2,\cdots,6;\ \ j = 1,2 \tag{12}$$

式中: ${}^i\boldsymbol{B}_j$ 为交点坐标; $\boldsymbol{A}_{ij} = [1,\ a_{ij},\ b_{ij}]^{\mathrm{T}}$ 为交线方程的待定系数。

由圆锥 O^iO 的几何关系可得到偏角 τ 的解析表达式为

$$\tau = \frac{\pi}{3} - \arccos\frac{2^il_1{}^2 - {}^il_2{}^2}{2^il_1{}^2} = \frac{\pi}{3} - \arccos\left[1 - \left(1 - \frac{1 + a_{i1}a_{i2} + b_{i1}b_{i2}}{\sqrt{(1+a_{i1}{}^2+b_{i1}{}^2)(1+a_{i2}{}^2+b_{i2}{}^2)}}\right) \Big/ (\sin\theta_1')^2\right] \tag{13}$$

式中: ${}^il_1$ 为点 ${}^iB_1$ 和点 ${}^iO$ 之间的距离; ${}^il_2$ 为点 ${}^iB_1$ 和点 ${}^iB_2$ 之间的距离。

由本节推导出的单个肋单元中 θ_1' 与滑块行程的关系, 可得到偏角 τ 与滑块行程的关系。偏角随滑块行程单调上升。偏角大小确定了转动副的运动范围, 可用于辅助配置可转动肋单元的转动副。

4　构型参数优化

在多模块折展机构设计的基础上, 通过对机构的构型参数进行优化, 达到天线展开状态结构刚度高, 天线整体结构质量轻的设计目标。

4.1　有限元模型

抛物面天线主要由多个杆件、铰链及张紧绳索组成。在太空中, 抛物面天线完全展开后, 负责连接杆件的铰链会实现位置锁定, 天线折展机构间的张紧绳索产生张紧力[16]。因此, 需要综合考虑杆件、绳索、铰链等的特性构建有限元模型。

本文抛物面天线采用环形封闭绳索。各模块相邻肋单元自由端间靠一条封闭环形绳索张紧, 绳索从两侧肋单元自由端内穿过, 并分别连接两侧肋单元自由端的上端和下端。环形封闭绳索可提供肋单元间张紧力, 同时提供肋单元自由端顶部与底

部纵向张紧，起到辅助锁定肋单元自由端套筒的作用。

建立抛物面天线展开状态的有限元模型计算固有频率，如图9所示。所有杆件均采用梁单元环形截面等效建模；绳索均采用桁架等效建模[17]。杆件与绳索的属性参数设置如表1和表2所示。为了仿真铰链锁定的真实状态，有限元模型中采用Bushing元来代替铰链，可以很好地控制铰链的自由度，更接近铰链的真实状态[16]。在铰链位置配置质量点，模拟铰链质量对天线结构刚度的影响。

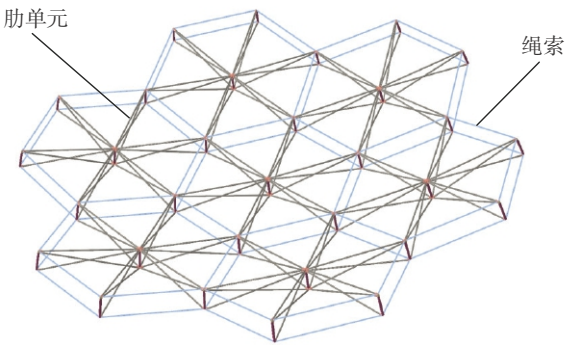


图9 七模块抛物面折展机构有限元模型

Fig. 9 Finite element model of seven module parabolic deployable mechanism

表1 杆件与绳索材料参数

Table 1 Material parameters of rods and rope

类型	弹性模量/MPa	密度/(10^3 kg·m^{-3})	泊松比
杆件	7×10^4	2.84	0.3
绳索	1.5×10^5	7.9	0.3

表2 杆件与绳索截面参数

Table 2 Section parameters of rods and rope

杆件截面编号	杆件外圆半径	杆件厚度/mm	应用杆件(见图3)	绳索截面半径/mm	绳索截面张紧压强/MPa
1	r_1	1	*CD*、*DE*、*KI*、*JM*	0.75	113.32
2	r_2	1	*OJ*、*AK*、*BC*、*EF*、*MG*、*IH*	0.75	113.32
3	r_3	12	*AB*、*FG*	0.75	113.32
4	r_3	4	*OA*、*GH*	0.75	113.32

表2中*OA*杆和*GH*杆对应多级套筒的收拢状态，截面厚度即为多级套筒收拢状态的总壁厚。

4.2 优化数学模型

性能评估即优化目标综合参考天线结构刚度和天线整体质量。本文采用天线调和平均固有频率作为结构刚度评价标准，将若干阶天线固有频率通过表达式整合为调和平均固有频率得到式(14)[18]：

$$F=\left(\frac{1}{m}\sum_{i=1}^{m}\lambda_i^{-1}\right)^{-1} \tag{14}$$

式中：F为天线前m阶调和平均固有频率；λ_i为天线的第i阶固有频率。

本文选取肋单元构型参数l_{OJ}、l_{BC}、l_{AB}及截面尺寸r_1作为设计变量，选取七模块抛物面折展机构展开状态前6阶调和平均固有频率F和整体质量M为目标函数。给定独立设计变量的另外5个参数按表3中数据为确定值。

表3 固定构型参数

Table 3 Fixed configuration parameters

θ_1/(°)	R/mm	$l_{C'E'}$/mm	r_2/mm	r_3/mm
7.318	9 402	20	11	13

优化数学模型为

$$\begin{cases}\text{find}\ \ \boldsymbol{X}=[l_{OJ},l_{BC},l_{AB},r_1]^{\mathrm{T}}\\ \min\ \begin{cases}M(\boldsymbol{X})\\ -F(\boldsymbol{X})\end{cases}\\ \text{s.t.}\ \ \boldsymbol{\Phi}(\boldsymbol{X})>\boldsymbol{0}\end{cases} \tag{15}$$

式中：

$$\boldsymbol{\Phi}(\boldsymbol{X})=\begin{bmatrix}l_{BC}-l_{OJ}\\ l_{JL}-l_{KL}\\ l_{KL}-l_{CD}-l_{AB}\sin\theta_1/2\\ l_{OA'}-l_{OA}\\ 5l_{OA}-l_{OA'}\\ 2l_{CD}-l_{AB}-l_{OA}\end{bmatrix}$$

前4个约束条件为肋单元机构构型尺寸约束，保证机构的合理性与可运动性。后2个约束条件为肋单元实用性约束，确保移动副多级套筒层数不超过5层，且肋单元展开状态横向展宽大于纵向展宽。自变量取值范围如表4所示。

表4 优化数学模型自变量取值范围

Table 4 Range of independent variables of optimization mathematical model

类型	l_{OJ}/mm	l_{BC}/mm	l_{AB}/mm	r_1/mm
下界	24	24	20	5
上界	40	40	30	20

为提高优化效率，解决自变量搜索范围内不可行点过多的问题，对决定肋单元构型的自变量、取值范围内点全因子遍历，找寻可行域范围，保留满足约束条件的自变量取值范围为搜索空间。用过程变量等效替代原自变量，缩小优化参数范围，如图10所示。

在自变量l_{OJ}、l_{BC}构成的搜索平面内，可行点集中在沿45°角平分线的一侧，自变量及其可行域可用通过式(16)转化成过程变量表达：

$$\begin{cases}l_{OJ}=x\cos\dfrac{\pi}{4}-y\sin\dfrac{\pi}{4}+24\\ l_{BC}=x\sin\dfrac{\pi}{4}+y\cos\dfrac{\pi}{4}+24\end{cases} \tag{16}$$

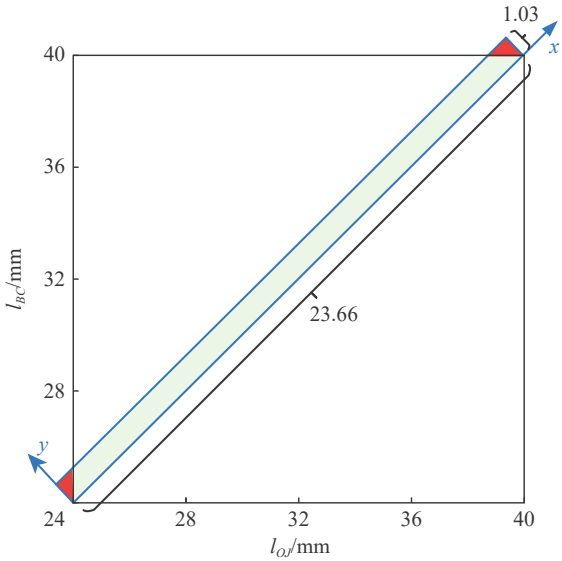


图 10　可行域及优化参数范围

Fig. 10　Feasible region and optimization parameter range

用过程变量替代原变量后, 优化数学模型为

$$\begin{cases} \text{find} & \boldsymbol{X}'=[x,y,l_{AB},r_1]^{\mathrm{T}} \\ \min & \begin{cases} M(\boldsymbol{X}') \\ -F(\boldsymbol{X}') \end{cases} \\ \text{s.t} & \boldsymbol{\Phi}'(\boldsymbol{X}')>\boldsymbol{0} \end{cases} \tag{17}$$

式中:

$$\boldsymbol{\Phi}'(\boldsymbol{X}')=\begin{bmatrix} 5l_{OA}-l_{OA'} \\ 2l_{CD}-l_{AB}-l_{OA} \\ l_{OJ}-24 \\ 40-l_{BC} \end{bmatrix}$$

自变量取值范围如表 5 所示。

表 5　优化数学模型更新后自变量取值范围

Table 5　Range of independent variables after updating the optimization mathematical model

类型	x/mm	y/mm	l_{AB}/mm	r_1/mm
下界	0	0	20	5
上界	23.6	1.03	30	20

本文优化问题为典型多目标优化问题, 目标函数的取值差异较明显, 采用多目标优化算法 NSGA-Ⅱ, 基于帕累托最优理论完成迭代计算, 获得多目标优化问题的最优解集。优化算法的参数设置如表 6 所示。

表 6　优化算法参数

Table 6　Optimization algorithm parameters

种群数	进化代数	交叉概率	交叉指数	变异指数
24	40	0.9	10	20

4.3　优化结果与性能分析

经过迭代计算得到七模块抛物面折展机构展开状态性能的帕累托前沿如图 11 所示。

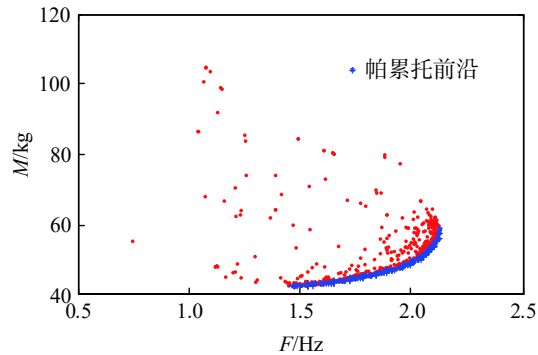


图 11　帕累托前沿

Fig. 11　Pareto front

综合考虑折展机构的质量特性和展开状态频率特性, 选取帕累托解集中的点 $\boldsymbol{X}'=[22.129, 0.191, 21.295, 7.944]^{\mathrm{T}}$作为最优解。最优解变量取值对应折展机构质量为 50.293 kg, 调和平均固有频率为 2.036 Hz, 综合性能优于其他自变量取值条件对应的性能。将过程变量 x、y 替换回自变量l_{OJ}、l_{BC}, 可得到最优解对应的设计变量$\boldsymbol{X}=[39.513, 39.782, 21.295, 7.944]^{\mathrm{T}}$。

将最优解和初始值自变量代入有限元模型对比验证, 得到对比结果如表 7 所示。

表 7　优化前后对比

Table 7　Comparison before and after optimization

类型	l_{OJ}/mm	l_{BC}/mm	l_{BC}/mm	l_{BC}/mm	l_{BC}/mm	l_{BC}/mm
初始值	39	40	25	12.5	1.216	77.565
最优解	39.513	39.782	21.295	7.944	2.036	50.293

经过构型参数优化后, 天线调和平均固有频率提高了 67.4%, 天线整体质量下降了 35.2%, 整体性能更优。

5　结　论

1) 本文方法实现了预定的展开状态肋单元两侧顶端落在拟合球面指定位置, 收拢状态肋单元两侧直线平行紧凑的目标。通过解析法对肋单元展开机构进行运动学分析, 得出过程拟合球面半径与驱动滑块行程的变化关系。

2) 提出包络锥法, 运用该方法实现了多模块折展机构任意状态下的肋单元自由端与公共端所在直线及各关键点坐标求解。解算七模块抛物面折展机构第 2 层模块可转动肋单元产生的偏角, 确定了肋单元公共端的运动副配置。

3) 构建七模块抛物面天线折展机构的展开状态有限元模型, 以调和平均固有频率最大和整体质量最轻为目标, 运用非劣排序算法对肋单元构型参数进行优化设计。

本文方法为模块化抛物面折展机构的设计提供了理论指导。基于剪刀机构的肋单元设计方法，可推广到其他模块化展开机构设计中；包络锥法可推广到任意肋单元机构构建的多层多模块抛物面天线折展机构设计与分析中。

参考文献（References）

[1] 陈传志, 董家宇, 陈金宝, 等. 空间大型星载抛物面天线研究进展[J]. 航空学报, 2021, 42(1): 523833.
CHEN C Z, DONG J Y, CHEN J B, et al. Large spaceborne parabolic antenna researchp progress[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(1): 523833(in Chinese).

[2] NATORI M C, HIRABAYASHI H, OKUIZUMI N, et al. A structure concept of high precision mesh antenna for space VLBI observation[C]// 43rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. Reston: AIAA, 2002: 1359-1362.

[3] YONEZAWA K, HOMMA M. Attitude control on ETS-8 mobile communication satellite with large deployable antennas[C]//21st International Communications Satellite Systems Conference and Exhibit. Reston: AIAA, 2003: 1-5.

[4] ONODA J, FU D Y, MINESUGI K. Two-dimensional deployable hexapod truss[J]. Journal of Spacecraft and Rockets, 1996, 33(3): 416-421.

[5] GUO J W, ZHAO Y S, XU Y D, et al. Design and analysis of truss deployable antenna mechanism based on a novel symmetric hexagonal profile division method[J]. Chinese Journal of Aeronautics, 2021, 34(8): 87-100.

[6] XU Y D, CHEN L L, LIU W L, et al. Type synthesis of the deployable mechanisms for the truss antenna using the method of adding constraint chains[J]. Journal of Mechanisms and Robotics, 2018, 10(4): 041002.

[7] 刘兆晶. 模块化可展开抛物面天线支撑机构设计与研制[D]. 哈尔滨: 哈尔滨工业大学, 2011: 12-13.
LIU Z J. Design and development of modular deployable parabolic antenna support mechanism[D]. Harbin: Harbin Institute of Technology, 2011: 12-13 (in Chinese).

[8] 陈向阳, 关富玲. 六棱柱单元可展抛物面天线结构设计[J]. 宇航学报, 2001, 22(1): 75-78.
CHEN X Y, GUAN F L. A large deployable hexapod paraboloid antenna[J]. Journal of Astronautics, 2001, 22(1): 75-78(in Chinese).

[9] 岳建如, 关富玲, 陈向阳. 大型可展构架式星载抛物面天线结构设计[J]. 浙江大学学报, 2001, 35(3): 238-243.
YUE J R, GUAN F L, CHEN X Y. A large deployable hexapod paraboloid antenna[J]. Journal of Zhejiang University, 2001, 35(3): 238-243(in Chinese).

[10] HU F, SONG Y P, HUANG Z R, et al. Malleability and optimization of tetrahedral metamorphic element for deployable truss antenna reflector[J]. AIP Advances, 2018, 8(5): 055217.

[11] HU F, SONG Y P, XU Y D, et al. Synthesis and optimization of modular deployable truss antenna reflector[J]. Aircraft Engineering and Aerospace Technology, 2018, 90(8): 1288-1294.

[12] 崔吉. 大尺度可展开式抛物面天线机构的设计与展开性能分析[D]. 哈尔滨: 哈尔滨工业大学, 2012: 9-11.
CUI J. Design and unfolding performance analysis of largescale deployable parabolic antenna mechanism[D]. Harbin: Harbin Institute of Technology, 2012: 9-11(in Chinese).

[13] 余丹敏. 大口径可展抛物面天线支撑桁架结构优化设计及性能研究[D]. 哈尔滨: 哈尔滨工业大学, 2020: 16-24.
YU D M. Structural optimization design and performance study of support truss for large-aperture deployable paraboloid antenna[D]. Harbin: Harbin Institute of Technology, 2020: 16-24(in Chinese).

[14] 田大可, 郭宏伟, 邓宗全, 等. 多模块构架式空间可展开天线结构参数优化[J]. 华中科技大学学报(自然科学版), 2012, 40(3): 49-53.
TIAN D K, GUO H W, DENG Z Q, et al. Optimization of structural parameters for space deployable truss antenna with multi-module[J]. Journal of Huazhong University of Science and Technology (Nature Science), 2012, 40(3): 49-53(in Chinese).

[15] 史创. 基于四棱柱可展单元的索杆张拉式天线机构研究[D]. 哈尔滨: 哈尔滨工业大学, 2019: 14-39.
SHI C. Research on cable-strut tension antenna mechanism based on quadrangular prism deployable element[D]. Harbin: Harbin Institute of Technology, 2019: 14-39 (in Chinese).

[16] 王辉, 何天宇, 都显琛, 等. 单元构架式抛物面天线张紧绳索多层设计方法[J]. 北京航空航天大学学报, 2022, 48(4): 657-664.
WANG H, HE T Y, DU X C, et al. Multi-layer design method of tension rope for paraboloid antenna[J]. Journal of Beijing University of Aeronautics and Astronautics, 2022, 48(4): 657-664(in Chinese).

[17] 刘荣强, 田大可, 邓宗全, 等. 多模块构架式空间可展开天线背架的模态分析[J]. 北京理工大学学报, 2011, 31(6): 685-690.
LIU R Q, TIAN D K, DENG Z Q, et al. Modal analysis of truss structure for deployable truss antenna with multi-module[J]. Transactions of Beijing Institute of Technology, 2011, 31(6): 685-690(in Chinese).

[18] MA Z D, KIKUCHI N, CHENG H C. Topological design for vibrating structures[J]. Computer Methods in Applied Mechanics and Engineering, 1995, 121(1-4): 259-280.

Design and optimization of modular parabolic deployable mechanism

HE Tianyu[1], DONG Yang[1], WANG Hui[1], ZHAO Junpeng[1, *], WANG Chunjie[1, 2]

(1. School of Mechanical Engineering and Automation, Beihang University, Beijing 100191, China;
2. State Key Laboratory of Virtual Reality and Systems, Beihang University, Beijing 100191, China)

Abstract: As an important functional component of a satellite, a parabolic antenna is a basis for realizing remote sensing measurement and wireless communication. In order to solve the difficulty of configuration design and stiffness matching of modular parabolic antenna deployable mechanism, a configuration method of rib element deployable mechanism based on scissors mechanism is proposed, and the rib element deployable mechanism is designed. Through kinematic analysis, it is verified that the same deployable mechanism can envelope different paraboloid-fitted spheres through the modification of parameters. The configuration of kinematic pairs of rotatable rib parts is based on a concept for solving the unfolding process of the multi-module deployable mechanism by the envelope cone approach.The configuration parameters of the rib element deployable mechanism are improved to increase natural frequency and decrease overall quality, and the finite element model of the multi-module deployable mechanism is built. The non-inferior sorting genetic algorithm is used to complete the optimization iterative calculation, and the design parameters of deployable mechanisms with high structural stiffness and lightweight are obtained. Under the same conditions, the harmonic average natural frequency is increased by 67.4%, and the overall quality of the antenna is reduced by 35.2% after optimization.

Keywords: modular parabolic antenna; configuration synthesis; kinematic analysis; finite element analysis; optimal design

Received: 2021-10-31; **Accepted**: 2022-01-25; **Published Online**: 2022-02-22 09:37

URL: kns.cnki.net/kcms/detail/11.2625.V.20220221.1844.002.html

Foundation item: National Natural Science Foundation of China (51635002)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0653

DES与DDES在湍流分离中的原理与性能研究

宋汉奇[1]，张恺玲[1]，马鸣[2]，阎超[1,*]

(1. 北京航空航天大学 航空科学与工程学院，北京 100191；2. 空间物理重点实验室，北京 100076)

摘 要：随着工程上流动结构的日益复杂，兼具雷诺平均Navier-Stokes（RANS）方法高效率与大涡模拟（LES）高精度的分离涡模拟（DES）类混合方法成为现阶段工程中最有效的湍流模拟方法之一。围绕DES类混合方法中的DES与延迟分离涡模拟（DDES）方法开展工作，分析二者开关函数构造上的不同，研究延迟因子作用机理，并考察DES与DDES方法的求解能力。研究表明：DES与DDES方法在模拟表现上存在一定差异，DDES方法通过引入延迟因子，保护RANS求解区域，改善模化应力不足，降低了DDES方法对交界面系数C_{DES}敏感程度；DDES方法在计算过程中容易出现过度保护，导致求解瞬时涡结构能力不如DES方法，分析与延迟因子引入比重及开关函数构造形式有关。

关 键 词：分离涡模拟；延迟分离涡模拟；开关函数；延迟因子；交界面系数；过度保护
中图分类号：V211.3
文献标志码：A **文章编号**：1001-5965(2023)09-2482-11

近几十年间，计算流体力学(computational fluid dynamics, CFD)迅速发展，成绩斐然，已成为飞行器研制和空气动力学研究的重要手段[1]。CFD的使用很大程度上降低了飞行器设计的风险和费用，同时可以为科研人员提供更加丰富的流场信息。

在实际飞行器使用过程中，遇到的气体流动大多是湍流。因此，湍流模型一直是CFD发展中十分关键的一环。美国国家航空航天局(National Aeronautics and Space Administration, NASA)在《2030年CFD远景规划》的报告中指出：湍流模型是影响CFD计算精度的重要因素之一，必须对湍流模型开展持续深入的研究[2]。

目前根据湍流数值模拟的精细程度，可以将湍流模拟方法分成3个层次[3]：直接数值模拟(direct numerical simulation, DNS)方法、大涡模拟(large eddy simulation,LES)方法及雷诺平均Navier-Stokes(Reynolds average Navier-Stockes, RANS)方法。其中，DNS与LES方法计算精度较高，给出的流场信息也较为全面。但是，DNS与LES方法计算需要网格量较大，计算效率较低。在当前计算条件下，DNS与LES方法距离实际工程的应用还有一段距离。RANS方法计算需要网格量小，可以较快地给出计算结果，是迄今为止最常用的方法。然而，RANS计算精度不如DNS与LES方法，对于非定常大分离等复杂流动难以给出满意结果。

对于实际工程中的湍流模拟，RANS方法计算效率高，但精度低；LES方法精度高，但计算效率低。这种情况下，将二者结合起来或许是最有效的方式，因此，RANS/LES混合方法诞生。RANS/LES混合方法兼具RANS方法的高效率和LES方法的高精度特点[4]。可以在较小的计算量下(相比于LES方法)，保证较高的计算精度(相比于RANS方法)[5]。这种方法一问世便得到广泛关注，许多研究者依据不同的基本模型发展了各式各样的

收稿日期：2021-10-31；**录用日期**：2022-01-21；**网络出版时间**：2022-02-10 13:16
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220209.1857.001.html
基金项目：国家数值风洞项目(NNW2019ZT1-A03)；国家自然科学基金(11721202)
* **通信作者**. E-mail：yanchao@buaa.edu.cn

引用格式：宋汉奇，张恺玲，马鸣，等. DES与DDES在湍流分离中的原理与性能研究[J]. 北京航空航天大学学报，2023，49(9)：2482-2492. SONG H Q，ZHANG K L，MA M，et al. Theory and performance research of DES and DDES in turbulent separation[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2482-2492 (in Chinese).

RANS/LES 混合方法[6]。其中 DES 类混合方法便是提出时间最早, 目前应用最广的方法。

DES 类混合方法最早由 Spalart 等[7] 在 1997 年提出, 即 Spalart-Allmaras 分离涡模拟(Spalart-Allmaras-detached eddy simulation SA-DES)方法。该方法在接近壁面的区域采用 SA 模型计算, 在远离壁面区域表现出 LES 方法求解的特性。之后参照文献 [7], Strelets[8] 在 Menter 剪切应力输运(shear stress transport, SST)模型基础上发展了 SST-DES 方法。DES 方法对于复杂流动表现出很好的适应性, 在后向台阶、翼型大迎角分离等流动中都给出了较满意的计算结果。然而 DES 方法也存在自己的不足, 其中较为明显的是模化应力不足问题。模化应力不足问题的本质在于 DES 方法对网格尺度依赖程度高, 导致在网格加密过程中容易出现 LES 方法入侵 RANS 方法计算区域的情况。为了降低 DES 方法对网格的依赖, Menter 等[9] 提出边界层保护思想。采用 SST 模型自带的边界层识别函数 F_1 或 F_2 作为延迟因子, 保护 RANS 方法计算区域, 即 SST-DDES 方法,DDES 是延迟分离涡模拟(delayed detached eddy simulation, DDES)。之后在 2006 年, Spalart 等[10] 参考引进延迟因子的思想, 提出 SA-DDES 方法。DDES 方法通过引入延迟因子来考虑当地湍流量对计算分区的影响, 以此提高划分 RANS 和 LES 方法计算区域的准确度。不过在这一过程中, 延迟因子引入的比重和开关函数的数学构造形式十分重要。倘若这 2 个方面设计的不合理, 容易导致 DDES 方法求解瞬时涡结构的能力下降。

本文主要研究 DES 类混合方法中的 DES 与 DDES 方法。首先, 对 2 种方法的原理进行研究, 分析二者的开关函数构造形式。同时以 SA-DDES 方法为例, 在湍流边界层的各个分层对其延迟因子进行数学上的化简, 通过其化简的表达式分析延迟因子在 DDES 方法中的作用。之后选取后向台阶流动算例, 分别采用基于 SA 和 SST 的 DES 与 DDES 方法进行计算。分析 DES 与 DDES 方法在模化应力不足问题和交界面敏感问题上的表现。同时采用 Q 准则涡量图和功率谱分析方法对比二者对于瞬时涡结构的捕捉能力。进一步利用基于 SST 模型的 DES 与 DDES 方法计算超音速凹腔-压缩拐角流动算例。同时引入另一种更改了开关函数的 SST-DDES-Fd 方法, 对比 3 种方法的时均结果与瞬时流场, 探究 2 种开关函数哪一种更适合基于 SST 模型的 DDES 方法。最后, 结合 2 个算例的计算结果, 综合分析 DES 与 DDES 方法在湍流分离模拟中的表现差异。

1　原理简介

1.1　SA-DES/DDES

SA-DES 方法是 Spalart 等[7] 在 1997 年提出的一种 DES 类混合方法。其控制方程为

$$\frac{\partial \tilde{\nu}}{\partial t}+U_j\frac{\partial \tilde{\nu}}{\partial x_j}=C_{b1}(1-f_{t2})\Omega\tilde{\nu}+\left\{C_{b1}[(1-f_{t2})f_{v2}+f_{t2}]\frac{1}{\kappa^2}-C_{w1}f_w\right\}\left(\frac{\tilde{\nu}}{\tilde{d}}\right)-\frac{C_{b2}}{\sigma}\tilde{\nu}\frac{\partial^2\tilde{\nu}}{\partial x_j^2}+\frac{1}{\sigma}\cdot\frac{\partial}{\partial x_j}\left\{\left[\mu+(1+C_{b2})\tilde{\nu}\right]\frac{\partial\tilde{\nu}}{\partial x_i}\right\}\tag{1}$$

式中: $\tilde{\nu}$为与涡黏性有关的变量; κ 为卡门常数; f_{t2}、f_{v2}、f_w为控制方程所需函数; C_{b1}、C_{w1}、C_{b2}、σ 为控制方程所需常数; U_j为速度; x_j、x_i为坐标; Ω为涡量张量不变量; μ为动力黏性系数。

$\tilde{d}$的表达式为

$$\tilde{d}=d_w-f\cdot\max(0,d_w-C_{DES}\Delta)\tag{2}$$

式中: d_w为壁面距离, 也是原来 SA 模型方程中模型系数; Δ为 3 个方向网格间距的最大值, 即 $\Delta=\max(\Delta_x,\Delta_y,\Delta_z)$, Δ_x、Δ_y、Δ_z 分别为 x、y、z 这 3 个方向的网格间距; C_{DES}为交界面系数; f为 SA-DES 方法的开关函数, 其表达式为

$$f=\begin{cases}0 & d_w<C_{DES}\Delta\\ 1 & d_w\geqslant C_{DES}\Delta\end{cases}\tag{3}$$

在靠近壁面的地方, $d_w<C_{DES}\Delta$, SA-DES 方法的开关函数取值为 0, 方法保持 RANS 模式进行计算; 在远离壁面的区域, $d_w\geqslant C_{DES}\Delta$, SA-DES 方法的开关函数取值为 1, 方法切换为 LES 模式。式(1)中其他各项具体意义及模型常数详见文献 [11]。

SA-DDES 方法是在 SA-DES 方法基础上对开关函数进行更改[8], 其更改之后的开关函数 f_d和$\tilde{d}$的表达式为

$$f_d=1-\tanh[(8r_d)^3]\tag{4}$$

$$\tilde{d}=d_w-f_d\cdot\max(0,d_w-C_{DES}\Delta)\tag{5}$$

式中: f_d为 SA-DDES 方法的开关函数; r_d 为延迟因子。

式(4)中, SA-DDES 方法的开关函数 f_d采用了双曲正切的数学形式。同 SA-DES 方法的分段函数形式的 f相比, 可以在 RANS 与 LES 方法的交界区域有更好的过渡。与此同时, SA-DDES 方法还在开关函数 f_d中引入延迟因子r_d, 其表达式为

$$r_d=\frac{\nu_t+\nu}{\sqrt{U_{i,j}U_{i,j}}\kappa^2d_w^2}\tag{6}$$

式中: ν_t为运动湍流涡黏性系数; ν为运动黏性系数; $U_{i,j}$ 为速度梯度; κ为卡门常数, 数值为 0.41。r_d在对

数层以内约等于1，然后至黏性顶层逐渐减小为0。从中可以看出相较于SA-DES方法，SA-DDES方法在开关函数中通过延迟因子引入了当地的湍流量，以此作为判断计算分区的二次判据。

1.2 SST-DES/DDES

SST-DES方法是Strelets[8]于2001年在两方程SST模型基础上提出的DES类混合方法，其模型方程为

$$\frac{\partial k}{\partial t}+U_j\frac{\partial k}{\partial x_j}=\frac{1}{\rho}P_k-\beta^*k\omega F_{\mathrm{DES}}+\frac{1}{\rho}\cdot\frac{\partial}{\partial x_j}\left[\left(\mu+\frac{\mu_{\mathrm{t}}}{\sigma_k}\right)\frac{\partial k}{\partial x_j}\right] \tag{7}$$

$$\frac{\partial \omega}{\partial t}+U_j\frac{\partial \omega}{\partial x_j}=\frac{1}{\rho}P_\omega-\beta\omega^2+\frac{1}{\rho}\cdot\frac{\partial}{\partial x_j}\left[\left(\mu+\frac{\mu_{\mathrm{t}}}{\sigma}\right)\frac{\partial \omega}{\partial x_j}\right]+2(1-F_1)\frac{1}{\sigma_{\omega_2}}\cdot\frac{1}{\omega}\cdot\frac{\partial k}{\partial x_j}\cdot\frac{\partial \omega}{\partial x_j} \tag{8}$$

$$S_{\mathrm{D},k}=-\beta^*k\omega F_{\mathrm{DES}} \tag{9}$$

式中：k和ω为湍动能和比耗散律；μ_{t}为动力湍流涡黏性系数；P_k和P_ω分别为式(7)和式(8)的生成项；β^*、σ_k、σ_ω、σ_{ω_2}为SST模型常数；ρ为密度；F_{DES}为SST-DES方法的开关函数，其表达式为

$$F_{\mathrm{DES}}=\max\left[\frac{L_{\mathrm{RANS}}}{C_{\mathrm{DES}}\Delta},1\right] \tag{10}$$

式中：L_{RANS}为RANS模型的长度尺度，$L_{\mathrm{RANS}}=\sqrt{k}/(\beta^*\omega)$。$F_{\mathrm{DES}}$在近壁区取值为1，在远离壁面区域取值大于1。式(7)～式(9)中其他参数和表达式详见文献[12]。

SST-DDES方法是由Menter等[9]提出，对SST-DES方法的开关函数进行修改以降低方法对于网格的依赖程度。其更改过后的开关函数表达式为

$$F_{\mathrm{DDES}}=\max\left[(1-F_{\mathrm{SST}})\frac{L_{\mathrm{RANS}}}{C_{\mathrm{DES}}\Delta},1\right] \tag{11}$$

式中：F_{SST}等于F_1或F_2，F_1和F_2为SST模型自带的边界层识别函数，其取值范围与r_{d}相似，是SST-DDES方法的延迟因子。二者的表达式为

$$F_1=\tanh(\mathrm{arg1}^4) \tag{12}$$

$$F_2=\tanh(\mathrm{arg2}^2) \tag{13}$$

$$\mathrm{arg1}=\min[\max(\mathrm{arg1a},\mathrm{arg1b}),\mathrm{arg1c}] \tag{14}$$

$$\mathrm{arg2}=\max(2\mathrm{arg1a},\mathrm{arg1b}) \tag{15}$$

$$\mathrm{arg1a}=\sqrt{k}/(0.09\omega y) \tag{16}$$

$$\mathrm{arg1b}=500\mu_1/(\rho\omega y^2) \tag{17}$$

$$\mathrm{arg1c}=4\rho k/y^2\sigma_{\omega_2}\mathrm{CD} \tag{18}$$

$$\mathrm{CD}=\max\left(\frac{2\rho}{\sigma_{\omega_2}\omega}\cdot\frac{\partial k}{\partial x_i}\cdot\frac{\partial \omega}{\partial x_j},10^{-20}\right) \tag{19}$$

式中：y为SST模型计算所用的壁面距离；arg1、arg1a、arg1b、arg1c、arg2、CD为SST模型计算所需函数。

根据DES与DDES的构造原理可以分析出，DES类混合方法通过开关函数将RANS模型与LES模型结合在一起进行计算。在靠近壁面的地方，DES类混合方法采用传统RANS模型，而随着壁面距离的增加，DES类混合方法由RANS切换为LES。分别利用RANS和LES方法对其擅长的区域进行计算，使得模拟方法在保证精度的同时又具有较高的计算效率，这是DES类混合方法最根本的思想与目的。

1.3 延迟因子分析

DES与DDES方法的主要差别是2种方法在开关函数上的不同，DDES方法在其开关函数中引入了延迟因子。本文以SA-DDES方法的延迟因子r_{d}为例，对延迟因子的作用进行分析。

依据边界层内速度与壁面距离的无量纲关系式及湍流量的取值，可以对r_{d}的数学形式进行一定程度的化简。在线性底层，速度与壁面距离二者的无量纲关系为

$$u^+=y^+ \tag{20}$$

式中：u^+为无量纲速度；y^+为无量纲壁面距离。

在式(20)中，可以先简单认为在该层中，流动的速度梯度是在一个有限的区间内变化的系数，并且区间的上、下限相差不大，同时在线性底层，湍流涡黏性系数相比于运动黏性系数通常可以忽略，所以可以把延迟因子表达式化简为

$$r_{\mathrm{d}}\propto\frac{\nu}{\kappa^2d_{\mathrm{w}}^2} \tag{21}$$

在线性底层，流动的黏性力占主导地位。因此，分式的分子是不可忽略的存在，而此时壁面距离较小，导致分子与分母的数量级相差无几甚至壁面距离的数量级可能稍小一些，所以此时r_{d}的取值在1附近或大于1。

在对数层，其速度的无量纲u^+和壁面距离的无量纲y^+二者具备对数关系：

$$u^+=\frac{1}{\kappa}\ln y^++C \tag{22}$$

式中：C为常数。

依据式(22)，可以初步认为在对数层，流动的速度梯度与壁面距离的倒数正相关，此时延迟因子r_{d}可以化简为

$$r_{\mathrm{d}}\propto\frac{\nu_{\mathrm{t}}+\nu}{\kappa d_{\mathrm{w}}} \tag{23}$$

式(23)与式(21)的分析几乎一样，其数值大小也在1左右。

在湍流边界层的外层，流动黏性的影响及速度梯度的取值都将变得较小，此时可以先简单地认为二者比值的极限趋近于 1，即

$$\lim_{y\to\delta}\left(\frac{\nu_t+\nu}{\sqrt{U_{i,j}U_{i,j}}}\right)=1 \tag{24}$$

式中：δ 为边界层的厚度。

延迟因子可以简化为

$$r_d \propto \frac{1}{\kappa^2 d_w^2} \tag{25}$$

此时的延迟因子受壁面距离的影响较大，在远离壁面的区域，延迟因子的数值将变得较小。从式(21)、式(23)及式(25)中可以看出：在整个计算过程中，SA-DDES 方法通过引入当地湍流量控制开关函数取值进一步划分 RANS 和 LES 方法的求解区域。整个计算过程不再仅仅依靠算例的网格尺度。

在 SST-DDES 方法中，引入 SST 模型自带的边界层识别函数 F_1 或 F_2 作为延迟因子，基本原理与化简思想和 SA-DDES 方法相差无几，只不过考虑的湍流量是湍动能和耗散率。因此，这里不再对 SST-DDES 方法延迟因子的化简进行描述。

2　后向台阶流动算例

2.1　算例概述

后向台阶流动虽然几何模型简单，但分离再附的流动结构复杂，是对湍流模型性能进行评估的典型算例。选取后向台阶算例以 Driver 和 Seegmiller [13] 的实验为参照，来流马赫数为 0.128，温度为 298 K，来流雷诺数为 36 000，台阶高度 h=1 cm，所有长度都以台阶高度为参考进行无量纲化。采用 Roe 格式进行空间离散，lower-upper symmetric Gauss-Seidel (LU-SGS) 进行时间推进。采用定常计算获得稳定的初始流场，而后进行 20 000 步非定常计算，非定常计算的物理时间步长为 3.6×10^{-7} s。后向台阶流动局部网格划分如图 1 所示。

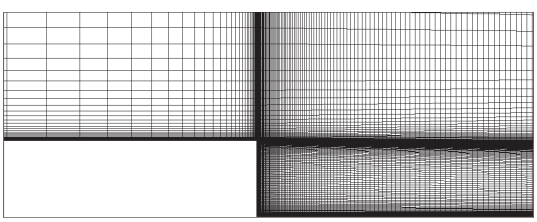


图 1　后向台阶流动局部网格划分

Fig. 1　Local mesh generation of backward step flow

2.2　时均结果分析

图 2 不同方法计算的时均上壁面摩擦系数分布，C_f 为摩擦系数。

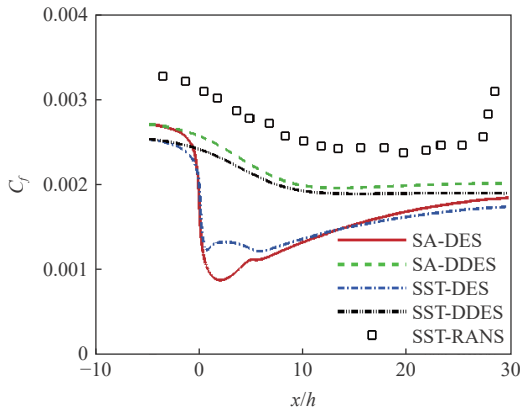


图 2　不同方法计算的时均上壁面摩擦系数分布

Fig. 2　Distribution of time-averaged top wall friction coefficients calculated with different methods

参考兰利研究中心湍流模型资源(Langley Research Center Turbulence Modeling Resource,LaRC-TMR)数据，SST RANS 模型在该雷诺数下可以给出与实验数据较为贴合的可靠计算结果。从图 2 中可以看出 DES 方法计算的上壁面摩擦系数明显小于 SST 和 DDES 方法计算结果，分析是由于 DES 方法在上壁面出现了模化应力不足问题，最终导致计算上壁面摩擦系数偏小。为了对模化应力不足有直观认识，给出 SA-DES 等 4 种方法计算的雷诺应力。图 3 为 DES 与 DDES 方法计算得出的不同站位雷诺应力分布，$u'v'/U_{ref}^2$ 表示雷诺应力的无量纲化。

从图 3 中可以看出，在整个上壁面近壁区 DES 方法计算的雷诺应力均小于 DDES 方法。由此可以证明在后向台阶流动的上壁面近壁区，DES 方法在计算过程中出现了模化应力不足问题。进一步对二者计算的雷诺应力的差值进行分析，以 x/h=4 站位为例。将该站位 y/h=8.99 处的 DES 与 DDES 方法计算的雷诺应力的差值作为假想的剪切应力差值 $\Delta\tau_\omega$ 代入摩擦系数计算公式：

$$\Delta C_f=\frac{2(\Delta\tau_\omega)}{\rho_{ref}U_{ref}^2} \tag{26}$$

式中：ρ_{ref} 为参考密度；U_{ref} 为参考速度。

从式(26)的计算中可以得出 $\Delta C_f=1.5\times10^{-3}$。直接从图 3 中读取的 x/h=4 站位处的摩擦系数差值约为 1.3×10^{-3}。从中可以看出上壁面摩擦系数的差值与 ΔC_f 的数值十分接近。因此，可以进一步证明在上壁面，DES 与 DDES 方法计算摩擦系数的差异就是模化应力不足问题造成的。总体来看，由于靠近台阶的地方流动结构复杂，上壁面的网格在接近台阶的区域进行了加密。而 DES 方法对于计算分区的判断完全依赖于网格尺度，在加密的过程中出现 LES 入侵 RANS 计算区域的情况。但是当地的

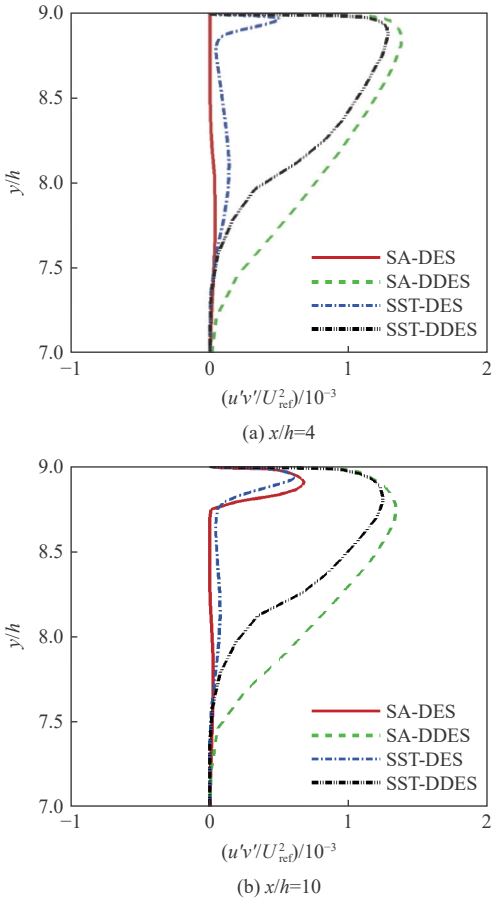


图3 DES与DDES方法计算的不同站位雷诺应力分布

Fig. 3 Distribution of Reynolds stress at different stations calculated by DES and DDES methods

网格分辨率无法满足LES方法计算的条件，造成DES方法在计算过程中出现模化应力不足，最终导致计算摩擦系数偏小。而DDES方法引入延迟因子，在网格加密的过程中很好地保护了RANS方法计算的区域，改善了模化应力不足的问题。

2.3 交界面敏感分析

对于DES类方法中的模型系数C_{DES}，其数值大小在相同网格条件下决定了LES和RANS方法计算的交界面位置所在。当混合方法切换到LES的时候，C_{DES}和Δ的乘积相当于此时方法的湍流混合尺度，类比于大涡模拟中的$C_{DES}\Delta$。在最早的DES方法中，依据各向同性湍流实验将其标定为0.65[4]，而传统的大涡模拟中C_{DES}=0.18，小于DES方法中的0.65。因此，分析适当调小C_{DES}数值可以使得DES类方法的LES模式更加贴近真正的大涡模拟，计算流场涡结构的能力将会提高。同时，根据非定常流场的变化，有时需要调整LES的模型系数以更适用于流场求解。那么在调节LES模型系数的过程中，对于近壁区RANS方法计算的影响便是本文关注的问题。由于C_{DES}的数值大小可以决定混合方法RANS与LES的交界面位置，因此，将该部分内容称为交界面敏感分析，将C_{DES}称为交界面系数。

研究C_{DES}为0.65和0.35时SA-DES和SA-DDES方法的表现，其中SA-DES-C_{DES}和SA-DDES-C_{DES}分别代表C_{DES}为0.35时的2种方法。

图4为不同方法计算的时均上壁面摩擦系数分布，以及其与SST-RANS计算结果的对比。

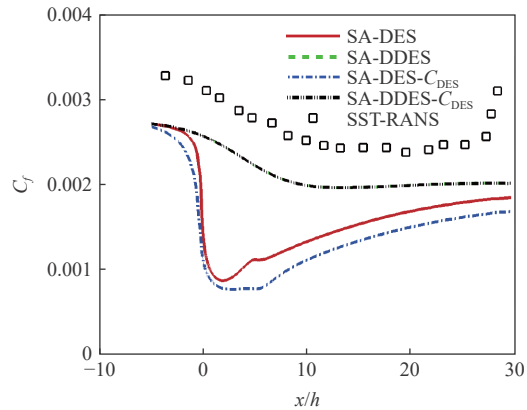
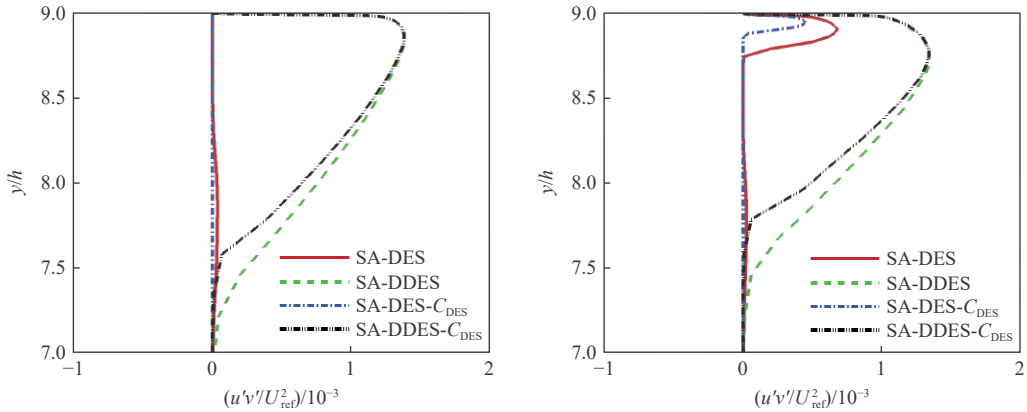


图4 不同方法计算的时均上壁面摩擦系数分布及与SST-RANS方法计算结果的对比

Fig. 4 Distribution of time averaged top wall friction coefficients calculated with different methods and comparison with SST-RANS method calalations

从图4中可以看出，更改C_{DES}之后，SA-DES方法表现出较强的敏感性：C_{DES}=0.35要明显小于C_{DES}=0.65的上壁面摩擦系数。与之对比，SA-DDES方法就表现的非常稳定。进一步参照2.2节对模化应力不足的分析，对比更改C_{DES}前后的雷诺应力。图5为更改C_{DES}前后方法计算的上壁面近壁区不同站位雷诺应力分布。

在x/h=10处，y/h在8.5～9.0区间内，SA-DES方法在更改前后出现较大的波动，而SA-DDES方法则十分稳定。在x/h=4处、x/h=10，y/h在7.0～8.0区间内，SA-DDES方法的前后变化貌似要大于SA-DES方法；但此时SA-DES方法本身计算的雷诺应力就较小，所以本文采用更改C_{DES}前后的相对变化率来评估2种方法。通过计算可以得出，SA-DDES方法的相对变化率是33%左右，而SA-DES方法则约为96%，因此，SA-DDES方法前后变化依旧要比SA-DES方法小。分析是SA-DDES方法引入延迟因子，使得方法的计算更加稳定。综合计算结果来看，DES方法对于交界面系数C_{DES}较为敏感。当根据流场需求调节LES的模型系数时，会对近壁区RANS方法计算产生较大的影响。与之对比，DDES方法通过延迟因子保护了RANS方法计算的区域，降低了对交界面系数C_{DES}的敏感程度。



(a) x/h=4　　(b) x/h=10

图 5　更改 C_{DES} 前后 DES 与 DDES 方法计算的不同站位雷诺应力分布

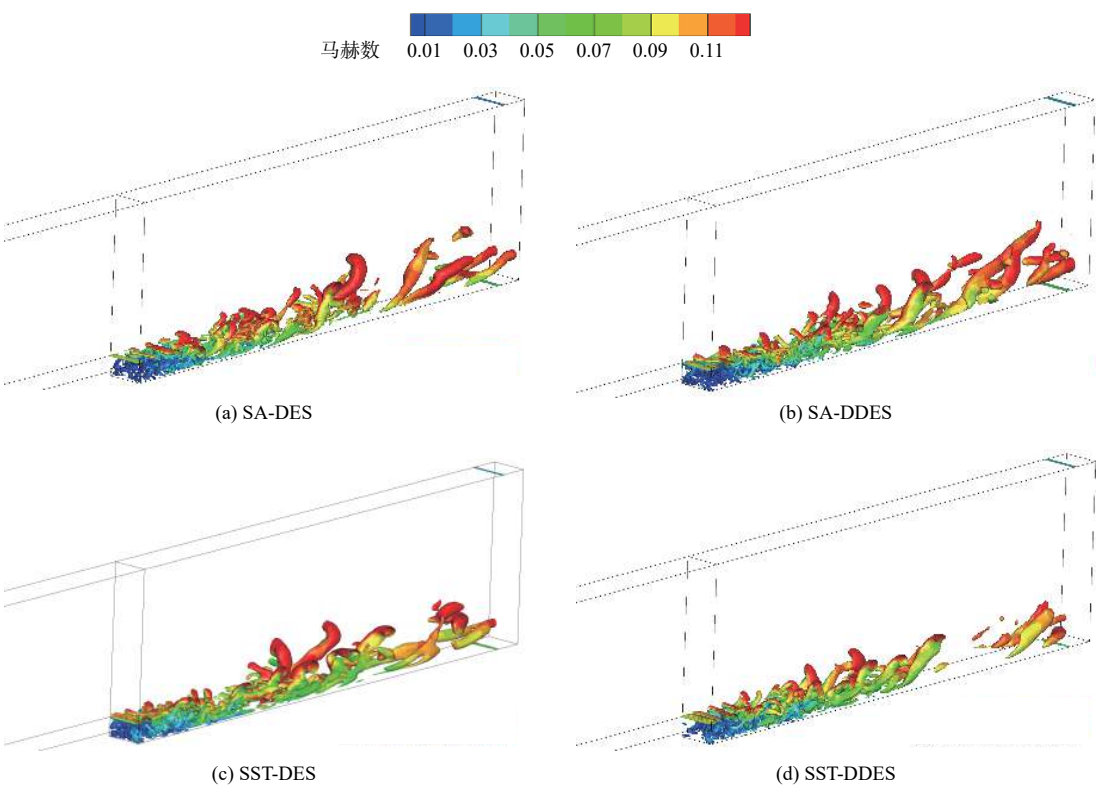
Fig. 5　Distribution of Reynolds stress at different stations calculated by DES and DDES methods before and after changing C_{DES}

2.4　瞬时流场分析

图 6 为 4 种 DES 类方法计算的 Q 准则瞬时涡量图。

后向台阶的分离涡结构是典型 Kelvin-Helmholt 不稳定性表现[14]。从 Q 准则瞬时涡量图中可以看出，基于 SA 模型的 DES 与 DDES 方法计算的涡结构相差不大；而基于 SST 模型的 DES 与 DDES 方法则表现出一定差异。为了进一步对比 SST-DES 与 SST-DDES 方法计算的瞬时流场，采用功率谱分析的方式，如图 7 所示。

图 7 为 SST-DES 与 SST-DDES 方法功率谱分析（power spectral density,PSD）图。从图 7 中可以看出：在 10^4～10^5 频率段内，2 个站位 SST-DES 方法的 PSD 值都要高于 SST-DDES 方法，其中在 x/h=10 站位处则更加明显。越高的频率往往对应于更加细小的涡结构，而更高的 PSD 值则代表更强的涡结构捕捉能力。从 2 个站位的 PSD 对比中可以得出结论：SST-DDES 方法对于更细小的涡结构的捕捉能力要逊于 SST-DES 方法。参考 Gritskevich 等[15-16]的分析：以 F_1 或 F_2 为延迟因子构造的开关函数在某些情况下不能很好地适用于 DDES 方法，造成了计算出现过度保护。



(a) SA-DES　　(b) SA-DDES

(c) SST-DES　　(d) SST-DDES

图 6　4 种 DES 类方法计算的 Q 准则瞬时涡量图

Fig. 6　Q-criterion instantaneous vorticity diagram calculation by four DES-like methods

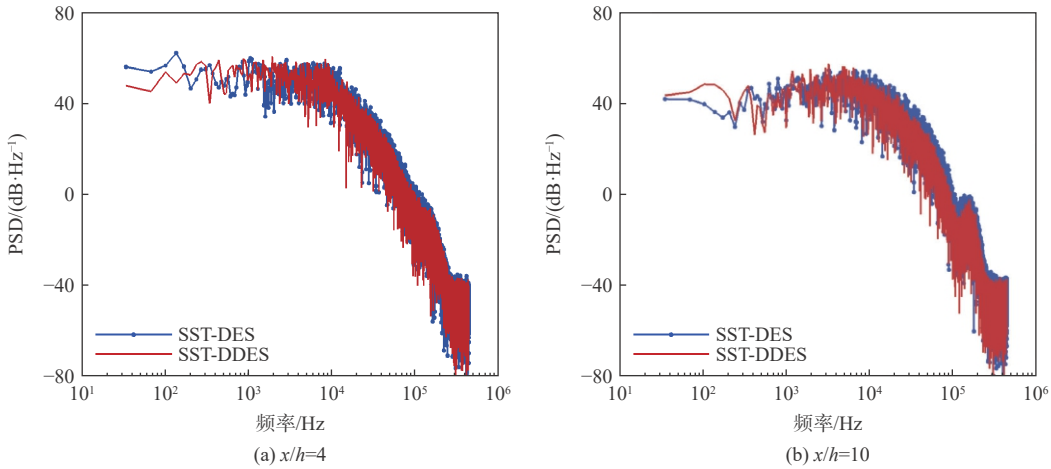


图 7 SST-DES 与 SST-DDES 方法功率谱分析图

Fig. 7 Power spectrum analysis plots of SST-DES and SST-DDES method

3 超音速凹腔-压缩拐角流动算例

3.1 算例概述

超音速凹腔−压缩拐角外形是未来高超声速飞行器发动机部分的典型部件，通过合理的设计来保证发动机内部燃烧的稳定性[17]。算例主要参考 Settles 等[18]试验，来流马赫数为 2.92，来流压强为 21 240 Pa，来流温度为 95.37 K。固定的物理时间步长为 1.5×10^{-7} s，首先定常计算获得稳定流场结构，之后进行非定常计算。采用 Roe 格式进行空间离散，LU-SGS 进行时间推进。图 8 为超音速凹腔-压缩拐角算例局部网格划分。

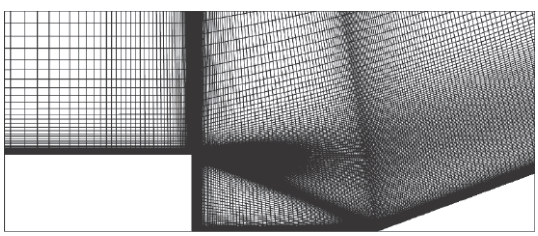


图 8 超音速凹腔-压缩拐角算例局部网格划分

Fig. 8 Local mesh generation of supersonic cavity-compression corner

3.2 SST-DDES-Fd 方法简介

引入 Gritskevich 等[15-16]提出的一种开关函数 f_{d_c}，将其应用到基于 SST 模型的 DDES 方法中，即 SST-DDES-Fd 方法。将计算结果与 SST-DES 和 SST-DDES 方法进行对比，分析 f_{d_c} 和 F_{DDES} 哪一种开关函数更适合基于 SST 模型的 DDES 方法。本节对 SST-DDES-Fd 方法的原理进行简单介绍。

SST-DDES-Fd 方法的构造原理和 SST-DDES 方法相差不多，基本方程一致。而其开关函数的构造及在方法中的应用形式与 SA-DDES 方法比较接近，表达式为

$$L = L_{RANS} - f_{d_c}\cdot\max(0, L_{RANS} - C_{DES}\Delta) \tag{27}$$

$$f_{d_c} = 1 - \tanh[(20r_{d_c})^3] \tag{28}$$

式中：f_{d_c} 和 r_{d_c} 分别为开关函数和延迟因子；L 为 SST-DDES-Fd 方法的长度尺度。延迟因子 r_{d_c} 的表达式为

$$r_{d_c} = \frac{\nu_t + \nu}{\kappa^2 y^2 \sqrt{0.5(S^2+\Omega^2)}} \tag{29}$$

式中：S 和 Ω 分别为应变率和涡量张量不变量。从式 (30) 中可以看出，SST-DDES-Fd 方法的开关函数依旧采用了双曲函数的数学形式，不过对于当地湍流量的引入相较于 SST-DES 方法有了一定的变化。同时相较于 SA-DDES 方法的开关函数，二者数学形式虽然相似，但 SST-DDES-Fd 方法增大了延迟因子的 r_{d_c} 比重。

3.3 时均结果分析

图 9 为时均壁面摩擦系数分布。其中，x'为距离压缩拐角偏折点的长度。从图中可以看出，3 种 DES 类方法计算的摩擦系数均大于实验数据且偏差较大。3 种方法相比，SST-DDES-Fd 方法计算结

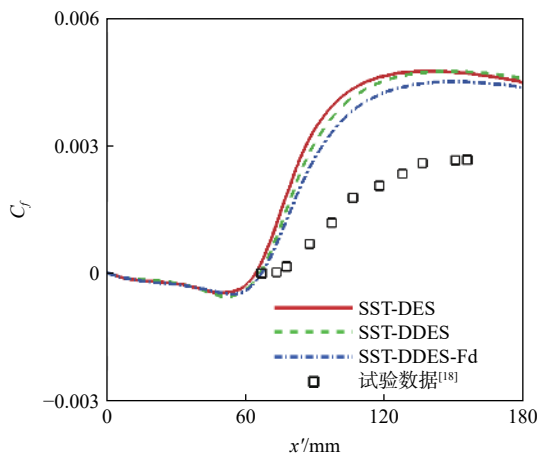


图 9 时均壁面摩擦系数分布

Fig. 9 Distribution of time-averaged wall friction coefficient

果更贴近实验数据, 总体来看, 三者没有特别明显的差别。

图 10 为在 x/h 分别为 1.5 和 3.5 这 2 个站位时均流向速度分布。

在 x/h=1.5 站位, 位置处于算例的凹腔之内, 流动结构复杂, 因此, 三者计算的速度出现较大差异, 其中, SST-DDES 和 SST-DDES-Fd 在该站位计算的速度与实验值较为符合, 而 SST-DES 计算偏小; 在 x/h=3.5 站位处, 位置位于凹腔之后的压缩拐角的壁面上, 流动结构相比于凹腔内部相对简单, 所以 3 种 DES 类方法计算速度差异不大, 其中以 SST-DDES-Fd 方法计算数值最为贴近实验值。综合来看, SST-DDES-Fd 方法计算的时均速度效果最好, 数值更加精确。

从时均结果分析中可以看出, SST-DES 方法对于时均结果的计算不如另外 2 种 DDES 方法。进一步考察 3 种方法计算的时均涡黏性等线图, 如图 11 所示。

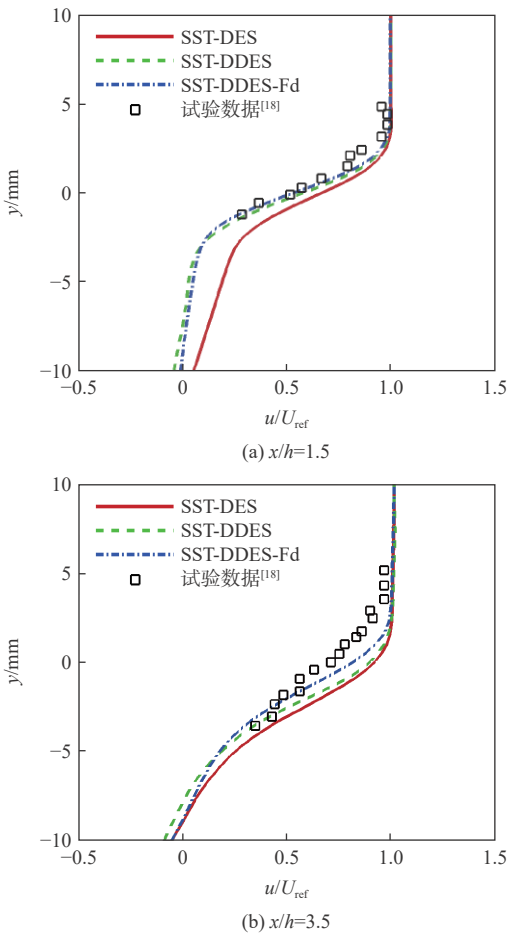


图 10　不同站位时均流向速度分布

Fig. 10　Distributions of time-averaged streamwise velocity at different locations

从图 11 中可以看出, 3 种方法计算的涡黏性大小关系为: SST-DES < SST-DDES < SST-DDES-Fd。

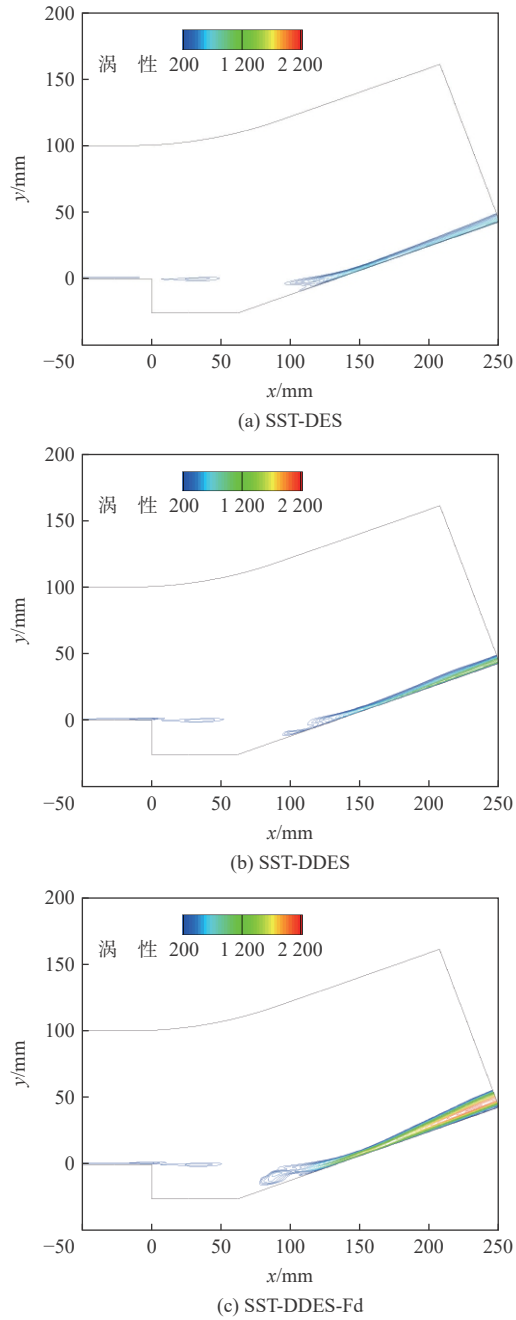


图 11　时均涡黏性等值线图

Fig. 11　Time-averaged eddy viscosity contour map

模型计算的雷诺应力与涡黏性系数具有正相关的关系, 因此, SST-DES 方法计算的雷诺应力也应是三者中的最小值。由此可以分析, SST-DES 方法在超音速凹腔-压缩拐角计算中也出现了一定程度的模化应力不足问题。同时, SST-DDES-Fd 方法计算的涡黏性要大于 SST-DDES 方法, 其对于 RANS 区域的保护程度要大于 SST-DDES 方法。就这一点而言, 理论上 SST-DDES-Fd 方法相较于 SST-DDES 方法更容易出现过度保护的问题。

3.4　瞬时流场分析

图 12 为 3 种 DES 类方法在 2 个瞬时状态下的 Q

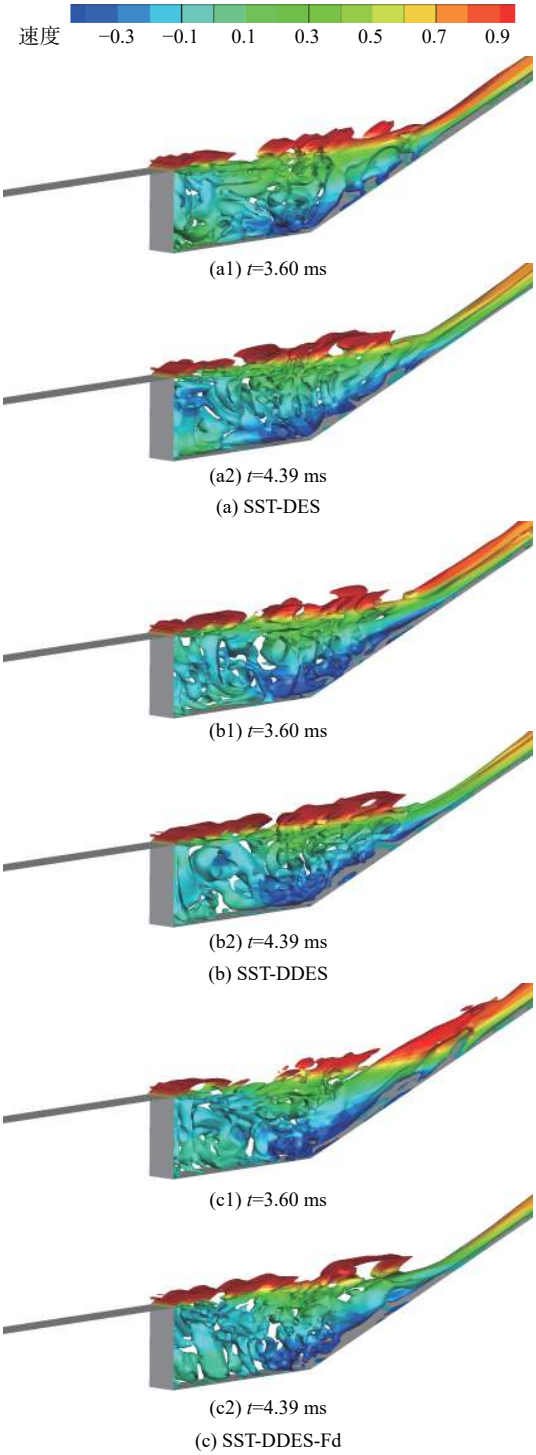


图 12 3 种 DES 方法在 2 个瞬时状态下的 Q 准则涡量图

Fig. 12 Q-criterion instantaneous vorticity diagram for 3 DES methods at 2 transient states

准则涡量图。在同一瞬时，SST-DES 与 SST-DDES 方法计算的涡结构相差不大：就凹腔内部涡结构而言，二者既可以计算出主体大涡结构，也可以计算出较小的三维涡结构，而在拐角之后，二者均可以计算出紧贴壁面的三维马蹄涡；对于 SST-DDES-Fd 方法，其在凹腔内部的回流区计算的涡结构与另外 2 种方法相差不多，但在拐角处计算的涡结构相较于另外 2 种方法则略显粗糙；在再附之后的斜坡面上 SST-DDES-Fd 方法仅仅计算出了卷起的大涡管，并没有计算出另外 2 种方法的平行马蹄涡结构。结合上文对 3 种方法的涡黏性分析可以得出 SST-DDES-Fd 方法在一定程度上过度保护 RANS 计算区域，导致一些精细的涡结构没有计算出来。

从瞬时流场的计算中可以看出，SST-DDES 方法并未出现过度保护的问题，而 SST-DDES-Fd 方法则出现了一定程度的过度保护。分析应当是 SST-DDES-Fd 方法的开关函数数学形式不当及过分加大了延迟因子引入比重导致计算出现了过度保护。不过，对于过度保护的影响因素还需要进一步深入研究。从计算结果来看，随着计算步数的增加，SST-DDES-Fd 方法的过度保护现象有所减弱。这与 Spalart 等[10] 在 2006 年对 SA-DDES 方法中 RANS 入侵 LES 计算区域问题的观点有些类似。从这一角度上来讲，过度保护对 DDES 方法来说并非“灾难性的”，其会增加非定常计算时间但对计算结果的影响不大。因此，过度保护能否真的作为评价 DES 类混合方法的一个标准可能就要视使用者的需求而定，在本文工作中暂时将其定为一个评价标准。因此，综合超音速凹腔-压缩拐角算例的计算结果来看，F_{DDES} 更适合作为基于 SST 模型的 DDES 方法的开关函数。

4 结 论

1）在原理研究方面，通过对 DES 与 DDES 方法开关函数的对比，分析 DES 方法与 DDES 方法的主要差别便是延迟因子的引入。进一步分析延迟因子得出：理论上延迟因子可以保护 RANS 计算区域，帮助 DDES 方法改善模化应力不足问题。

2）DDES 方法通过引入延迟因子，降低了对交界面系数的敏感性，提高了对壁面摩擦的计算精度；同时，在超音速凹腔−压缩拐角流动算例中，DDES 方法计算的时均速度更准确。综合对比，DDES 方法计算时均结果更加精确。

3）在瞬时结果对比中，DES 方法可以捕捉到比 DDES 方法更细小的涡结构。分析是由于过度引入延迟因子及不适当的开关函数构造形式，导致 DDES 方法在计算过程中出现过度保护。

下一步计划中，将围绕 DDES 方法的延迟因子比重开展工作，在一定范围内对比不同比重的计算结果，以期为 DDES 方法提供更精确、更具普适性的开关函数。此外，对于交界面敏感问题分析，本文目前只对比了 C_{DES}=0.65 和 C_{DES}=0.35 时基于 SA 模型的 DES 与 DDES 方法的表现。在接下来工作中，将进一步扩大 C_{DES} 的取值范围进行对比分析，

同时将对比模型扩展到 SST-DES 与 SST-DDES 方法。

参考文献（References）

[1] 阎超, 屈峰, 赵雅甜, 等. 航空航天CFD物理模型和计算方法的评述与挑战[J]. 空气动力学学报, 2020, 38(5): 829-857.
YAN C, QU F, ZHAO Y T, et al. Review of development and challenges for physical modeling and numerical scheme of CFD in aeronautics and astronautics[J]. Acta Aerodynamic Sinica, 2020, 38(5): 829-857(in Chinese).

[2] SLONTNICK J, KHODADOUST A, ALONSO J, et al. CFD vision 2030 study: A path to revolutionary computational aerosciences: NASA/CR-2014-218178 [R]. Washing ton, D. C. : NASA, 2014.

[3] 张兆顺, 崔桂香, 许春晓. 湍流理论与模拟[M]. 第2版. 北京: 清华大学出版社, 2017: 173-174.
ZHANG Z S, CUI G X, XU C X. Theory and modeling of turbulence[M]. 2nd ed. Beijing: Tsinghua University Press, 2017: 173-174 (in Chinese).

[4] 肖志祥, 罗堃宇, 刘健. 宽速域RANS-LES混合方法的发展及应用[J]. 空气动力学学报, 2017, 35(3): 338-353.
XIAO Z X, LUO K Y, LIU J. Developments and applications of hybrid RANS-LES methods for wide-speed-range flows[J]. Acta Aerodynamica Sinica, 2017, 35(3): 338-353(in Chinese).

[5] 杜若凡, 阎超, 韩政, 等. DDES延迟函数在超声速底部流动中的性能分析[J]. 北京航空航天大学学报, 2017, 43(8): 1585-1593.
DU R F, YAN C, HAN Z, et al. Performance of delayed functions in DDES for supersonic base flow[J]. Journal of Beijing University of Aeronautics and Astronautics, 2017, 43(8): 1585-1593(in Chinese).

[6] FRÖHLICH J, VON TERZI D. Hybrid LES/RANS methods for the simulation of turbulent flows[J]. Progress in Aerospace Sciences, 2008, 44(5): 349-377.

[7] SPALART P R, JOU W H, STRELETS M, et al. Comments on the feasibility of LES for wings and on a hybrid RANS/LES approach[C]//Proceedings of 1st AFOSR International Conference on DNS/LES, Advances in DNS/LES. Columbus: Greyden Press, 1997: 137-147.

[8] STRELETS M. Detached eddy simulation of massively separated flows: AIAA-2001-0879 [R]. Reston: AIAA, 2001.

[9] MENTER F R, KUNTZ M, LANGTRY R. Ten years of industrial experience with the SST turbulencemodel[C]//Heatand Mass Transfer 4. Redding: Begell House, 2003: 625-632.

[10] SPALART P R, DECK S, SHUR M L, et al. A new version of detached-eddy simulation, resistant to ambiguous grid densities[J]. Theoretical and Computational Fluid Dynamics, 2006, 20(3): 181-195.

[11] SPALART P, ALLMARAS S. A one-equation turbulence model for aerodynamic flows[C]// 30th Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 1992.

[12] MENTER F R. Two-equation eddy-viscosity turbulence models for engineering applications[J]. AIAA Journal, 1994, 32(8): 1598-1605.

[13] DRIVER D M, SEEGMILLER H L. Features of a reattaching turbulent shear layer in divergent channelflow[J]. AIAA Journal, 1985, 23(2): 163-171.

[14] 张露, 李杰, 牟永飞, 等. 基于DES类混合方法模拟后台阶分离流动[J]. 西北工业大学学报, 2017, 35(6): 983-989.
ZHANG L, LI J, MOU Y F, et al. Simulating unsteady flow over a backward facing step with advanced detached-eddy-simulation methods[J]. Journal of Northwestern Polytechnical University, 2017, 35(6): 983-989(in Chinese).

[15] GRITSKEVICH M S, GARBARUK A V, SCHÜTZE J, et al. Development of DDES and IDDES formulations for the k-ω shear stress transport model[J]. Flow, Turbulence and Combustion, 2012, 88(3): 431-449.

[16] GRITSKEVICH M S, GARBARUK A V, MENTER F R. Fine-tuning of DDES and IDDES formulations to the k–ω shear stress transport model[J]. Progress in Flight Physics, 2013, 5: 23-42.

[17] ZHOU L, ZHAO R, YUAN W. An investigation of interface conditions inherent in detached-eddy simulation methods[J]. Aerospace Science and Technology, 2018, 74(3): 46-55.

[18] SETTLES G S, WILLIAMS D R, BACA B K, et al. Reattachment of a compressible turbulent free shear layer[J]. AIAA Journal, 1982, 20(1): 60-67.

Theory and performance research of DES and DDES in turbulent separation

SONG Hanqi[1]，ZHANG Kailing[1]，MA Ming[2]，YAN Chao[1,*]

(1. School of Aeronautic Science and Engineering，Beihang University，Beijing 100191，China；
2. Key Laboratory of Space Physics，Beijing 100076，China)

Abstract：With the increasing complexity of flow structure in engineering, detached eddy simulation (DES) has become one of the most effective methods for turbulence simulation. DES is a hybrid method, combining Reynolds averaged Navier-Stokes (RANS) and large eddy simulation (LES) and thus possessing the high efficiency of RANS and high prevision of DES. This research focuses on DES and delayed detached eddy simulation (DDES), analyzing the differences in the structure of shielding functions, and the action mechanism of delay factors. Backward step flow and supersonic cavity compression corner flow are selected to compare and analyze the solving ability of DES and DDES. The results show that DDES protects the RANS solution area by introducing the delay factor, improving modeled-stress depletion and reducing the sensitivity to C_{DES}. However, DDES is prone to over protection in the calculation process, resulting in lesser ability to solve the instantaneous vortex structure than DES method. The analysis is related to the introduction of delay factors and the construction form of shielding functions.

Keywords：detached eddy simulation；delayed detached eddy simulation；shielding functions；delay factors；interface coefficient；over protection

Received：2021-10-31；**Accepted**：2022-01-21；**Published Online**：2022-02-10 13：16

URL：kns.cnki.net/kcms/detail/11.2625.V.20220209.1857.001.html

Foundation items：National Numerical Wind Tunnel Project (NNW2019ZT1-A03); National Natural Science Foundation of China (11721202)

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

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0665

一种历史信息特征敏感的行人迭代检测方法

戴佩哲[1]，刘翔[1,*]，张星[2]，尚岩峰[3]，赵静文[1]，王诗雨[1]

(1. 上海工程技术大学 电子电气工程学院，上海 201620；2. 上海工程技术大学 管理学院，上海 201620；
3. 公安部第三研究所 物联网技术研发中心，上海 200031)

摘　　要：基于深度学习的目标检测算法通常需要使用非极大值抑制等后处理方法对预测框进行筛选，无法在行人拥挤的场景下平衡模型的检测精度和召回率。虽然迭代检测的方法可以解决非极大值抑制等方法带来的问题，但是重复检测同样会限制模型的性能。提出了一种历史信息特征敏感的行人迭代检测方法。引入带权重的历史信息特征（WHIC），提高特征的区分度；利用历史信息特征提取模块（HIFEM）得到不同尺度的历史信息特征，并融合进主网络中进行多尺度检测，增强了模型对历史信息特征的敏感度，有效抑制重复检测框的产生。实验结果表明：所提方法在拥挤场景的行人检测数据集 CrowdHuman 和 WiderPerson 上取得了最优的检测精度和召回率。

关 键 词：机器视觉；目标检测；特征融合；卷积神经网络；深度学习
中图分类号：TP391.4
文献标志码：A　　**文章编号**：1001-5965(2023)09-2493-08

行人检测的应用场景中目标通常比较拥挤，行人之间的相互重叠和遮挡都会对目标检测算法产生巨大的挑战。如何在拥挤的场景下对行人进行精确的检测是一个学术界和工业界都十分关注的课题，但是目前通用的目标检测算法在行人拥挤场景下的检测精度和召回率都难以满足要求[1]。目前，基于深度学习的目标检测算法大致可以分为2个类别：两阶段的检测算法和单阶段的检测算法。两阶段的检测算法先产生一组候选目标框，再对其进行分类和定位。例如，Girshick 等提出的基于区域的卷积神经网络（region-based convolutional neural network, R-CNN）[2] 采用选择性搜索算法对输入图像提取出 2 000 个候选区域，再对候选区域进行裁剪缩放后输入 CNN 进行分类和定位。He 等提出了空间金字塔池化网络（spatial pyramid pooling Net, SPP-Net）[3]，使得网络能够输入任意尺度的候选区域而不需要对候选区域进行裁剪和缩放。Girshick 提出了 Fast R-CNN[4]，利用了多任务损失函数，将分类损失和定位边界框回归损失统一训练学习。为了解决选择性搜索算法速度慢的问题，Ren 等提出了 Faster R-CNN[5]，采用区域建议网络（region proposal network, RPN）取代选择性搜索算法，实现了端到端的训练，极大提高了检测速度。虽然两阶段的检测算法精度高，但是训练步骤繁琐、检测速度过慢。为了实现目标检测算法的实时性，基于回归的单阶段检测算法舍弃了 RPN 网络来生成候选区域，直接在图像的多个位置回归出候选框和类别。单阶段的检测算法主要有 YOLOv1～YOLOv4[6-9] 和 SSD[10] 2个系列。单阶段的检测算法使用锚框机制提高了网络的定位能力，但是带来了与锚框相关的复杂计算和超参数设计。Law 和 Deng 提出了 CornerNet[11]，直接预测物体的左上角和右下角来得到检测框，将

收稿日期：2021-11-05；**录用日期**：2022-01-27；**网络出版时间**：2022-02-15 14:47
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220214.1838.014.html
基金项目：国家重点研发计划 (2017YFC0821603)；上海市自然科学基金 (19ZR1421500)
*通信作者. E-mail：xliu@sues.edu.cn

引用格式：戴佩哲，刘翔，张星，等. 一种历史信息特征敏感的行人迭代检测方法 [J]. 北京航空航天大学学报，2023，49（9）：2493-2500.
DAI P Z，LIU X，ZHANG X，et al. An iterative pedestrian detection method sensitive to historical information features [J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（9）：2493-2500 （in Chinese）.

目标检测问题当作关键点检测问题来解决。Zhou 等提出了 ExtremeNet[12]，选取物体上下左右 4 个极值点和 1 个中心点预测分支。Tian 等提出了基于全卷积的单阶段目标检测器全卷积一阶段目标检测(full convolutional one-stage object detection, FCOS)[13]，采用逐像素预测方式解决目标检测问题。最近，Zhang 等[14]提出 anchor-based 和 anchor-free 方法之间的内在区别是对正负训练样本的定义不同所导致的，并且提出了一种自适应的训练样本选择算法根据目标的统计特征自动选择正负训练样本。

以上基于深度学习的目标检测算法都需要采用非极大值抑制等后处理方法去除冗余的检测框，但是非极大值抑制等方法在多个同类别的目标出现重叠和遮挡时，无法判断出大量检测框是否属于需要进行抑制的冗余检测框还是对目标正确的检测结果[15]。Ge 等基于两阶段目标检测器设计了 PS-RCNN[16]，先使用一个 R-CNN 对图像中未被遮挡的目标进行检测，将检测结果转换成目标形状的挡板，用其来过滤掉已经被检测出的目标后，再使用另外一个 R-CNN 来检测那些被遮挡的目标，但是，PS-RCNN 只能应用于两阶段的检测算法，并不能移植到单阶段的检测算法上，不具备通用性。Rukhovich 等提出了迭代检测器(iterative detector, IterDet)[17]，设计了一种迭代的网络结构，通过引入历史信息特征和设计模型的训练方式使得模型学习利用第 1 次检测结果的信息对输入图像进行第 2 次检测，从而提高模型在拥挤场景下的精确度和召回率。虽然 IterDet 的迭代机制可以方便地移植到两阶段或者单阶段的算法上，但是模型依旧不能处理重复检测的问题，使得模型的检测精度和召回率受到了限制。

本文提出了一种历史信息特征敏感的行人迭代检测方法。首先，引入带权重的历史信息特征(weighted historical information characteristics, WHIC)，对不同遮挡程度的像素位置赋予相应的权重，提高特征的区分度；然后，通过历史信息特征提取模块(historical information feature extraction module, HIFEM)得到多尺度的历史信息特征，并且融入主网络中进行检测，增强了模型对历史信息特征的敏感度。本文方法有效抑制了迭代检测中重复检测框的产生，显著提高了模型的检测精度与召回率。

1 IterDet 迭代检测

迭代检测的思想本质上是将 1 次困难的目标检测拆分成 2 次相对简单的检测，通过对前后 2 次检测的结果进行合并得到最优的结果集。如图 1 所示，IterDet 将第 1 次的检测结果转化成历史信息特征后融合进主网络中再进行第 2 次检测。

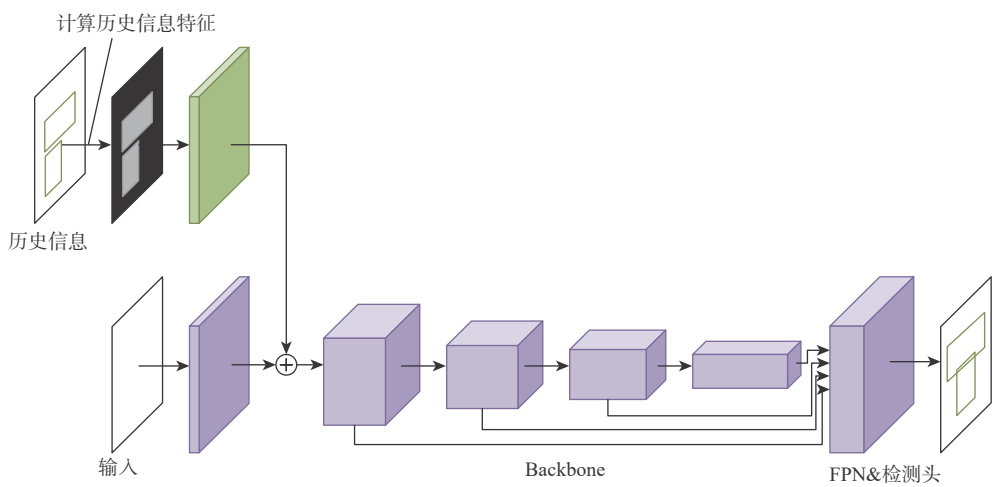
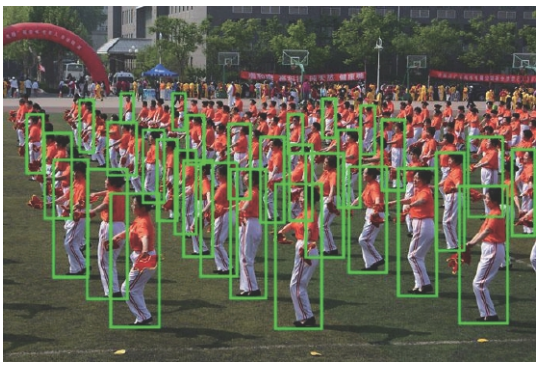


图 1 IterDet 迭代网络结构

Fig. 1 Structure of IterDet iterative network

通过实验发现，IterDet 的设计无法将第 1 次的检测结果和第 2 次的检测结果做到有效隔离，第 2 次检测中会产生许多第 1 次检测已经得到的重复检测框。重复检测框的产生说明了迭代检测机制的失效。

图 2 展示了 IterDet 前后 2 次检测中得到的检测框。其中，图 2(a)的绿色方框为第 1 次检测产生的检测框；图 2(b)为第 2 次检测产生的检测框，紫色框为重复检测框，黄色框为新产生的检测框。从图 2 可以看出，IterDet 在第 2 次检测中产生了许多重复检测框，被重复检测框遮挡的目标会被非极大值抑制等后处理方法直接过滤掉，最终模型的检测精度和召回率都会受到严重的限制。



(a) 第1次检测产生的检测框

(b) 第2次检测产生的检测框

图 2　IterDet 2 次检测得到的检测框展示

Fig. 2　Display of detection box for two tests of IterDet

2　历史信息特征敏感的迭代机制

2.1　带权重的历史信息特征

迭代机制的核心就是历史信息特征, 在训练过程中需要利用历史信息特征使得模型获得对其敏感的权重, 在检测过程中需要利用对历史信息特征敏感的模型进行第 2 次检测。因此, 如何设计历史信息特征是迭代机制的关键问题。

IterDet 的历史信息特征设计非常粗糙, 先统计所有像素点上覆盖的检测框个数, 再将其转换成灰度图上对应的灰度值, 最终作为模型的历史信息特征。这样的历史信息特征区分度不高, 最终会导致模型对历史信息特征不够敏感。

本文提出用 WHIC 来提高特征的区分度。基于检测框的个数将像素位置的遮挡程度划分为轻度遮挡、中度遮挡、重度遮挡, 对不同的遮挡程度设置不同的权重系数, 每一个像素位置上的值为覆盖的检测框个数与权重系数的乘积, 通过引入权重系数使得不同遮挡程度的像素位置具有区分度更高的特征。权重系数的设定为: 当像素点上覆盖的检测框数量在 5 以内时, 该像素位置属于轻度遮挡, 权重系数设置为 4; 当像素点上覆盖的检测框数量超过 5 但在 10 以内时, 该像素位置属于中度遮挡, 被遮挡的目标缺乏足够的像素信息, 权重系数设置为 8; 当像素点上覆盖的检测框数量超过 10 时, 该像素位置属于重度遮挡, 将权重系数设置为 12。

引入权重系数的具体方法为: 首先, 将原始图像输入模型, 得到模型第 1 次检测出的检测框结果 $R=\{(x_k,y_k,w_k,h_k)\}_{k=1}^{n}$, 其中, x_k、y_k为第 k 个检测框的左上角坐标, w_k和h_k分别为坐标框的宽和高, 统计每个像素位置上被R中的检测框所覆盖过的检测框数量N_{xy}, 如式(1)所示; 然后, 根据N_{xy}计算对应像素位置上的权重系数W_{xy}, 如式(2)所示; 最后, 定义一张与原图尺寸大小一致的历史信息特征图H_{xy}, H_{xy}中对应像素位置的值为N_{xy}与W_{xy}的乘积, 如式(3)所示。

$$N_{xy}=\sum_{k=1}^{R}1,\ x_k\leqslant x\leqslant x_k+w_k, y_k\leqslant y\leqslant y_k+h_k \tag{1}$$

$$W_{xy}=\begin{cases}0 & N_{xy}=0\\ 4 & 0<N_{xy}\leqslant 5\\ 8 & 5<N_{xy}\leqslant 10\\ 12 & N_{xy}>10\end{cases} \tag{2}$$

$$H_{xy}=N_{xy}W_{xy} \tag{3}$$

2.2　历史信息特征提取模块

CNN 中浅层特征的像素信息比较丰富, 但是语义信息不足, 深层特征的像素信息不足, 但是语义信息比较丰富[18], 并且浅层特征的像素信息会随着特征提取网络的卷积操作逐渐消失[19]。

历史信息特征是训练历史信息特征敏感模型的关键, 如果网络中浅层的历史信息特征随着网络的加深而消失, 迭代机制就会失效, 模型也无法利用多尺度检测提高检测精度和召回率[20]。虽然采用特征金字塔网络(feature pyramid network, FPN)[21]或路径聚合网络(path aggregation network, PANet)[22]的特征融合手段可以将深层特征的语义信息与浅层特征的像素信息进行融合, 从而弥补特征不平衡的缺陷, 但是这些特征融合的方法是按照一定的顺序对相邻层的特征进行融合, 导致模型只关注顺序相邻的特征层之间的信息融合[23], 浅层的历史信息特征依旧会随着网络的加深而丢失。

本文提出用 HIFEM 来解决浅层的历史信息特征消失的问题。模型通过 HIFEM 对输入的历史信息特征进行单独的特征提取, 从而得到多种尺度的历史信息特征, 并将其融合进主网络中。HIFEM 主要由 4 个残差块conv1、conv2_x、conv3_x、conv4_x组成, 残差块的网络参数如表 1 所示。该模块还具

表 1　HIFEM 网络参数

Table 1　Parameters of HIFEM

层名称	输出尺寸	层参数
conv1	112×112	7×7, 64, stride 2
conv2_x	56×56	3×3 max pool, stride 2 $\begin{pmatrix}3\times3, & 256\\ 3\times3, & 256\end{pmatrix}\times2$
conv3_x	28×28	$\begin{pmatrix}3\times3, & 512\\ 3\times3, & 512\end{pmatrix}\times2$
conv4_x	14×14	$\begin{pmatrix}3\times3, & 1\,024\\ 3\times3, & 1\,024\end{pmatrix}\times2$

有一条只有最大值池化操作的分支对浅层的历史信息特征进行尺度缩放后与对应的残差块的输出进行特征融合。HIFEM的网络结构见图3。HIFEM利用残差块对输入的历史信息特征进行卷积操作后得到多尺度的深层语义信息特征，同时利用最大值池化操作分支保留浅层的像素信息进行特征融合，其中，特征融合采用逐像素相加的方式在保留浅层像素信息的基础上融合了更加深层的语义信息。

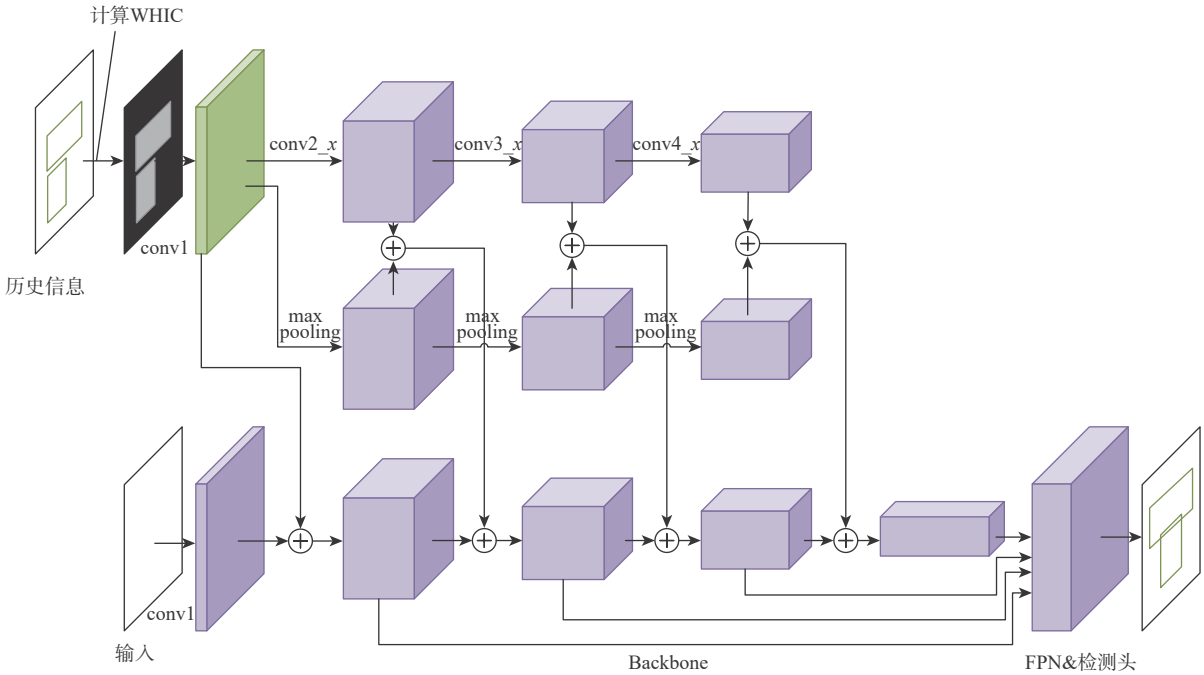


图3 HIFEM网络结构

Fig. 3 Network structure of HIFEM

2.3 模型训练

模型选取ResNet-50作为主网络，并且在每一个特征层中融合HIFEM输出的尺度一致的历史信息特征，在训练过程中将主网络与HIFEM同时进行反向传播。模型的具体训练过程可以分为如下3个步骤：

步骤1 将数据集中的真值框B分为2部分：B_{old}和B_{new}。

步骤2 将B_{old}直接转化成历史信息特征图H_{xy}后，输入HIFEM，得到多个尺度的历史信息特征，再将多个尺度的历史信息特征融合进主网络中，特征融合的方式为对应像素直接相加。通过主网络进一步提取特征，利用检测头得到预测框P。

步骤3 通过预测框P与B_{new}之间的误差反向传播优化模型的损失函数，对主网络ResNet-50与HIFEM同时进行训练。

3 实验结果与分析

3.1 实验环境与数据集

本文的相关实验基于Pytorch深度学习框架，操作系统为Ubuntu，GPU为GTX2080。为了实验结果的公平性，将模型的参数与IterDet模型的参数保持一致。

CrowdHuman数据集[24]是一个复杂的行人图像数据集，训练集有15 000张图像，验证集有4 370张图像，测试集有5 000张图像。其中，每张图像平均拥有23个行人，并且每个行人都有3个标注的真值框：全身框、可见身体框和头部框，本文实验选取全身框标注的真值框。数据集中的真值框不但重叠得非常严重，并且很多目标都处于图像的边缘位置。WiderPerson数据集[25]包含的场景更加广泛与复杂。其中，训练集有8 000张图像，验证集有1 000张图像，测试集有4 382张图像，其中测试集的标注非公开。本文将模型在测试集中产生的重复检测框的平均数量作为衡量重复检测的实验指标。

选取单阶段检测算法RetinaNet与两阶段检测算法Faster R-CNN作为Baseline，选择ResNet-50作为特征提取网络。

3.2 最优权重系数

本文基于IterDet针对不同的遮挡程度设置了5组权重系数，在2个不同的数据集上对单阶段检测算法与两阶段检测算法进行最优权重系数的实验。其中，考虑灰度值的范围大小，每组的轻度遮挡权重系数以2作为步长，灰度值的范围为2～10；考虑不同遮挡程度的区分度，每组的不同遮挡程度权重系数以该组轻度遮挡系数作为步长。

实验结果如表 2～表 5 所示。其中, 表 2、表 3 为在 CrowdHuman 数据集上分别对 RetinaNet 的 IterDet 与 Faster R-CNN 的 IterDet 的实验结果, 表 4、表 5 为在 WiderPerson 数据集上分别对 RetinaNet 的 IterDet 与 Faster R-CNN 的 IterDet 的实验结果。由 4 组实验数据可知, 引入权重系数可以提高特征的区分度, 模型的平均重复检测框个数减少, 检测精度与召回率也得到了提高。4 组实验中, 权重系数设置为 4、8、12 的模型在检测精度、召回率及实验最优的两个指标为召回率(%)和检测精度(%)的指标上都达到了最优; 但是随着权重系数数值的增大, 检测精度与召回率都会下降。分析上述结果可知, 权重系数虽然可以提高特征的区分度, 从而抑制重复检测的产生, 但是权重系数的数值过大也会干扰原始图像的像素信息, 导致模型的检测精度与召回率下降。

3.3　消融实验

本文采用 HIFEM 对增强模型的历史信息特征敏感度设置消融实验进行验证。

表 6 和表 7 分别为 CrowdHuman 和 WiderPerson 数据集上的实验结果。上述 2 组实验的结果表明, HIFEM 的设计有效增强了模型对于历史信息特征的敏感度, 能够充分利用历史信息特征进行多尺度上的检测, 从而抑制重复检测框的产生, 并且大幅度提高了模型的检测精度与召回率。

表 2　CrowdHuman 数据集上基于 RetinaNet+IterDet 的权重系数实验结果

Table 2　Experimental results of weight coefficient based on RetinaNet+IterDet on CrowdHuman dataset

轻度遮挡权重系数	中度遮挡权重系数	重度遮挡权重系数	召回率	检测精度	平均重复检测框个数
			91.49	84.77	11.89
2	4	6	92.39	85.45	8.87
4	8	12	93.26	85.59	8.56
6	12	18	92.13	85.16	8.59
8	16	24	88.19	85.11	7.18
10	20	30	80.76	81.13	6.12

表 3　CrowdHuman 数据集上基于 Faster R-CNN+IterDet 的权重系数实验结果

Table 3　Experimental results of weight coefficient based on Faster R-CNN+IterDet on CrowdHuman dataset

轻度遮挡权重系数	中度遮挡权重系数	重度遮挡权重系数	召回率	检测精度	平均重复检测框个数
			95.80	88.08	8.45
2	4	6	96.14	89.18	7.34
4	8	12	96.54	89.56	6.12
6	12	18	96.34	88.65	5.06
8	16	24	95.16	85.17	5.12
10	20	30	90.46	83.14	4.06

表 4　WiderPerson 数据集上基于 RetinaNet+IterDet 的权重系数实验结果

Table 4　Experimental results of weight coefficient based on RetinaNet+IterDet on WiderPerson dataset

轻度遮挡权重系数	中度遮挡权重系数	重度遮挡权重系数	召回率	检测精度	平均重复检测框个数
			95.35	90.23	8.66
2	4	6	95.44	91.59	7.15
4	8	12	96.13	92.43	6.21
6	12	18	95.89	91.87	6.34
8	16	24	90.14	88.56	5.22
10	20	30	88.19	85.34	5.12

表 5　WiderPerson 数据集上基于 Faster R-CNN+IterDet 的权重系数实验结果

Table 5　Experimental results of weight coefficient based on Faster R-CNN+IterDet on WiderPerson dataset

轻度遮挡权重系数	中度遮挡权重系数	重度遮挡权重系数	召回率	检测精度	平均重复检测框个数
			97.15	91.95	5.65
2	4	6	97.16	92.59	4.49
4	8	12	97.60	93.14	4.45
6	12	18	94.17	90.15	4.19
8	16	24	90.23	88.71	4.01
10	20	30	88.59	84.88	3.67

3.4　对比实验

本文方法采用 WHIC 增加特征的区分度, 并且通过 HIFEM 提高模型对历史信息特征的敏感度, 与IterDet 和PS-RCNN 在CrowdHuman 与WiderPerson 数据集上进行对比。对比实验在检测精度与召回率这 2 个指标上对不同方法的性能进行评估。表 8 和表 9 分别为 CrowdHuman和 WiderPerson 数据集上的实验结果。结果表明, 本文方法在检测精度与召回率上均取得了最优的结果。

采用本文方法对随机选取的拥挤场景图像进行检测, 效果如图 4 所示。可以很明显看到, 本文方法对 2 次检测做出了有效的隔离, 几乎没有重复检测的产生, 特别是对于被严重遮挡的目标, 模型同样可以对其进行有效的检测。

表 6　CrowdHuman 数据集上的消融实验结果

Table 6　Ablation experimental results on CrowdHuman dataset

检测器	召回率		检测精度		平均重复检测框个数	
	IterDet	IterDet+HIFEM	IterDet	IterDet+HIFEM	IterDet	IterDet+HIFEM
RetinaNet	91.49	96.76	84.77	88.98	11.89	3.34
Faster R-CNN	95.80	97.10	88.08	91.10	8.45	2.21

表7 WiderPerson数据集上的消融实验结果

Table 7 Ablation experimental results on WiderPerson dataset

检测器	召回率		检测精度		平均重复检测框个数	
	IterDet	IterDet+HIFEM	IterDet	IterDet+HIFEM	IterDet	IterDet+HIFEM
RetinaNet	95.35	97.60	90.23	94.70	8.66	4.12
Faster R-CNN	97.15	98.23	91.95	95.40	5.65	1.81

表8 CrowdHuman数据集对比实验结果

Table 8 Results of comparison experiment on CrowdHuman dataset

检测器	召回率				检测精度			
	Baseline	PS-RCNN	IterDet	本文方法	Baseline	PS-RCNN	IterDet	本文方法
RetinaNet	93.80	91.49		97.40	80.83	84.77		89.73
Faster R-CNN	90.24	93.77	95.80	97.98	84.95	86.05	88.08	91.15

表9 WiderPerson数据集对比实验结果

Table 9 Results of comparison experiment on WiderPerson dataset

检测器	召回率				检测精度			
	Baseline	PS-RCNN	IterDet	本文方法	Baseline	PS-RCNN	IterDet	本文方法
RetinaNet	90.20	95.35		98.87	89.12	90.23		95.99
Faster R-CNN	93.60	94.71	97.15	98.67	88.89	89.96	91.95	96.67

(a) 第1次检测产生的检测框 (b) 第2次检测产生的检测框

图4 检测效果

Fig. 4 Detection effect

4 结 论

本文提出了一种历史信息特征敏感的行人迭代检测方法，优化了历史信息特征的设计，并提出了更加有效的网络结构对历史信息特征进行有效的提取与融合。

1) 权重系数的引入提高了特征的区分度, 丰富了历史信息特征的信息, 对模型的性能有一定的提升。

2) 历史信息特征提取模块在训练过程中有助于提高模型对历史信息特征的敏感度, 在推断过程中能够有效利用多尺度的历史信息特征进行检测。实验证明, 历史信息特征提取模块大幅度提高了模型在 CrowdHuman 与 WiderPerson 数据集上的检测精度与召回率。

参考文献（References）

[1] 邱博, 刘翔, 石蕴玉, 等. 一种轻量化的多目标实时检测模型[J]. 北京航空航天大学学报, 2020, 46(9): 1778-1785.
QIU B, LIU X, SHI Y Y, et al. A lightweight multi-target real-time detection model[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(9): 1778-1785(in Chinese).

[2] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]// IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2014: 580-587.

[3] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.

[4] GIRSHICK R. Fast R-CNN [C]//IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2016: 1440-1448.

[5] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[6] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 779-788.

[7] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 6517-6525.

[8] REDMON J, FARHADI A. YOLOv3: An incremental improvement[EB/OL].(2018-04-08)[2021-06-06],https://arxiv.org/abs/1804.02767.

[9] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: Optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2021-06-06]. https://arxiv.org/abs/2004.10934.

[10] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//European Conference on Computer Vision. Berlin: Springer, 2016: 21-37.

[11] LAW H, DENG J. CornerNet: Detecting objects as paired keypoints[J]. International Journal of Computer Vision, 2020, 128(3): 642-656.

[12] ZHOU X Y, ZHUO J C, KRÄHENBÜHL P. Bottom-up object detection by grouping extreme and center points[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 850-859.

[13] TIAN Z, SHEN C H, CHEN H, et al. FCOS: Fully convolutional one-stage object detection[C]//IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2020: 9626-9635.

[14] ZHANG S F, CHI C, YAO Y Q, et al. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 9756-9765.

[15] 罗会兰, 陈鸿坤. 基于深度学习的目标检测研究综述[J]. 电子学报, 2020, 48(6): 1230-1239.
LUO H L, CHEN H K. Survey of object detection based on deep learning[J]. Acta Electronica Sinica, 2020, 48(6): 1230-1239(in Chinese).

[16] GE Z, JIE Z, HUANG X, et al. PS-RCNN: Detecting secondary human instances in a crowd via primary object suppression [EB/OL]. (2020-03-16) [2021-06-06]. http://arxiv.org/abs/2003.07080.

[17] RUKHOVICH D, SOFIIUK K, GALEEV D, et al. IterDet: Iterative scheme for object detection in crowded environments[C]// Structural, Syntactic, and Statistical Pattern Recognition. Beilin: Springer, 2021: 344-354.

[18] 王海, 王宽, 蔡英凤, 等. 基于改进级联卷积神经网络的交通标志识别[J]. 汽车工程, 2020, 42(9): 1256-1262.
WANG H, WANG K, CAI Y F, et al. Traffic sign recognition based on improved cascade convolution neural network[J]. Automotive Engineering, 2020, 42(9): 1256-1262(in Chinese).

[19] 郑浦, 白宏阳, 李伟, 等. 复杂背景下的小目标检测算法[J]. 浙江大学学报 (工学版), 2020, 54(9): 1777-1784.
ZHENG P, BAI H Y, LI W, et al. Small target detection algorithm in complex background[J]. Journal of Zhejiang University (Engineering Science), 2020, 54(9): 1777-1784(in Chinese).

[20] 马立, 巩笑天, 欧阳航空. Tiny YOLOV3目标检测改进[J]. 光学精密工程, 2020, 28(4): 988-995.
MA L, GONG X T, OUYANG H K. Improvement of Tiny YOLOV3 target detection[J]. Optics and Precision Engineering, 2020, 28(4): 988-995(in Chinese).

[21] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 936-944.

[22] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 8759-8768.

[23] PANG J M, CHEN K, SHI J P, et al. Libra R-CNN: Towards balanced learning for object detection[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 821-830.

[24] SHAO S, ZHAO Z J, LI B X, et al. CrowdHuman: A benchmark for detecting human in a crowd[EB/OL]. (2018-04-30)[2021-06-06]. https://arxiv.org/abs/1805.00123.

[25] ZHANG S F, XIE Y L, WAN J, et al. WiderPerson: A diverse dataset for dense pedestrian detection in the wild[J]. IEEE Transactions on Multimedia, 2020, 22(2): 380-393.

An iterative pedestrian detection method sensitive to historical information features

DAI Peizhe[1]，LIU Xiang[1,*]，ZHANG Xing[2]，SHANG Yanfeng[3]，ZHAO Jingwen[1]，WANG Shiyu[1]

(1. School of Electronic and Electric Engineering，Shanghai University of Engineering Science，Shanghai 201620，China；
2. School of Management，Shanghai University of Engineering Science，Shanghai 201620，China；
3. Internet of Things R&D Technology Center，The Third Research Institute of the Ministry of Public Security，Shanghai 200031，China)

Abstract：Object detection algorithms based on deep learning usually need to use post-processing methods such as non-maximum suppression to filter the prediction box, and can not balance the detection accuracy and recall rate of the model in the crowded pedestrian scene. Although the iterative detection method can solve the problems caused by non-maximum suppression methods, repeated detection will also limit the performance of the model. In this paper, a pedestrian iterative detection method sensitive to historical information features is proposed. Firstly, the weighted historical information characteristics (WHIC) is introduced to improve the feature discrimination. Second, the historical information feature extraction module (HIFEM) suggested in this paper is utilized to obtain and fuse historical information features of various scales into the main network for multi-scale detection, increasing the sensitivity of the model to the historical information features. This method can effectively suppress the generation of repeated detection frames. Experimental results show that the proposed method achieves the best detection accuracy and recall on CrowdHuman and WiderPerson.

Keywords：machine vision; object detection; feature fusion; convolution neural network; deep learning

Received：2021-11-05；**Accepted**：2022-01-27；**Published Online**：2022-02-15 14：47

URL：kns.cnki.net/kcms/detail/11.2625.V.20220214.1838.014.html

Foundation items：National Key R&D Program of China (2017YFC0821603); Natural Science Foundation of Shanghai (19ZR1421500)

* **Corresponding author**. E-mail：xliu@sues.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0681

基于 L_1-ITD 的尾座式无人机姿态控制

王姝旸，张晶*，杨凌宇
(北京航空航天大学 自动化科学与电气工程学院，北京 100191)

摘 要：针对尾座式无人机在垂直起降（VTOL）阶段对风干扰敏感的问题，提出了一种基于 L_1-ITD 的尾座式无人机姿态控制方法。建立了垂直起降阶段尾座式无人机六自由度非线性模型，设计了基于 L_1 自适应的无人机姿态控制器，可抑制扰动对系统性能的影响，实现垂直起降阶段存在风干扰和模型不确定时对无人机姿态的良好控制。在此基础上，针对 L_1 自适应控制方法对量测噪声敏感及无法直接获取有效微分信号的问题，引入改进跟踪微分器，在快速精确跟踪信号的同时抑制了量测噪声的影响。仿真结果验证了所提方法的有效性。

关 键 词：L_1 自适应控制；跟踪微分器；尾座式无人机；姿态控制；量测噪声
中图分类号：V249.122
文献标志码：A **文章编号**：1001-5965(2023)09-2501-09

尾座式无人机凭借着其起降灵活、定点悬停、机动性好等优势，受到各军事强国的重视，在军事与民用领域应用广泛[1]。垂直起降(vertical take-off and landing, VTOL)模式是尾座式无人机较为特殊的飞行模式，该模式中，无人机保持机头竖直朝上、机尾朝下的姿态来实现垂直起降、定点滞空等飞行状态，因此，该模式最能反映尾座式无人机的特点。小型尾座式无人机在垂直起降飞行阶段面临低空飞行环境复杂、抗风抗扰能力差等问题[2]，仅依靠螺旋桨产生升力和推动力，在受到外部扰动时易陷入振荡或发散，这也是尾座式无人机飞行过程中姿态控制的难点所在。因此，必须对尾座式无人机垂直起降阶段姿态控制律进行研究。

针对存在扰动时的无人机姿态控制问题，文献 [3-6] 采用标称控制器和干扰观测器相结合的设计方法，对扰动进行估计补偿，但非线性干扰观测器要求系统干扰具有慢变特性，限制其应用范围。韩京清[7] 提出了自抗扰控制方法，不依赖于精确模型，具有跟踪速度快、控制精度高和抗干扰能力强的特点。文献 [8-9] 采用自抗扰控制方法解决无人机受扰时的姿态控制问题，但控制参数整定较为繁琐。

自适应控制方法针对被控对象的不确定性，通过自动调节控制参数达到快速准确跟踪参考输入的目的，因此常与其他理论方法结合，解决无人机受扰时姿态控制问题。文献 [10] 提出了一种基于神经网络的自适应控制方法，可以估计四旋翼无人机在狭窄区域或近壁、近表面飞行时所受的不确定力，使无人机在存在干扰的条件下表现出令人满意的性能，但计算成本较高。文献 [11] 提出了一种自适应快速非奇异终端滑模控制方法，抑制了扰动对系统性能的影响，姿态角跟踪误差在±4.3°范围内。L_1 自适应控制[12-14] 是对模型参考自适应控制方法的一种改进，通过引入投影算子对控制器中的可调机构参数进行限幅，并在控制律设计环节引入低通滤波器，实现对参考输入信号的重构，在快速自适应的同时保证系统具有良好的鲁棒性[15]。目前，该方法已应用于 X-48B、无人战斗机、高超声速飞行器、四旋翼飞行器等控

收稿日期：2021-11-11；**录用日期**：2022-01-21；**网络出版时间**：2022-02-16 09:05
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220215.1630.003.html
* **通信作者**. E-mail：zhangjing2013@buaa.edu.cn

引用格式：王姝旸，张晶，杨凌宇. 基于 L_1-ITD 的尾座式无人机姿态控制 [J]. 北京航空航天大学学报，2023，49（9）：2501-2509.
WANG S Y，ZHANG J，YANG L Y. Attitude control law based on L_1-ITD for a tail-sitter UAV[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（9）：2501-2509（in Chinese）.

制对象[16]，也有将该方法用于尾座式无人机姿态控制之中。刘玉焘[1]在简化垂直起降无人机模型的基础上，采用L_1自适应控制方法对垂直起降无人机姿态进行控制，仿真结果验证了L_1自适应控制对于参数摄动和外界干扰有较强的抑制作用，能有效跟踪输入信号，控制效果良好。文献[17]提出了一种新颖的考虑输入约束的L_1自适应控制结构，并将其与前馈补偿器、线性二次型调节器结合，对尾座式无人机姿态进行控制，在无人机控制面偏转存在约束的情况下，最大限度地消除超调，显著提高系统性能。文献[18]将L_1自适应控制方法和反步法结合，对无人机姿态进行稳定控制，控制效果良好，同时降低了控制系统设计的复杂度。但以上L_1自适应控制方法均未考虑传感器量测噪声带来的影响，引入量测噪声后操纵机构的控制指令剧烈抖动，可能会导致执行器件损伤。另外，L_1自适应控制方法对量测噪声敏感，信号混杂随机噪声时难以获得其有效微分信号，直接影响控制效果及系统稳定性，故需与其他方法结合抑制量测噪声的影响。

针对量测噪声的问题，韩京清和王伟于1994年提出了跟踪微分器[19]（tracking differentiator, TD）的概念，通过适当选取跟踪参数，可以快速准确地跟踪输入信号，同时得到其微分信号。文献[20-22]对跟踪微分器进行了改进，综合了线性跟踪微分器和非线性跟踪微分器的优点，兼顾了快速性和准确性的要求，具有良好的动态响应及较强的滤波能力。刘海亭[23]在文献[22]的基础上提出了改进跟踪微分器（improved tracking differentiator, ITD），减少了系统参数，调参简单，利于实际工程应用。

本文将L_1自适应控制与改进跟踪微分器结合，提出了基于L_1-ITD的尾座式无人机姿态控制方法。该方法实现了垂直起降阶段存在风干扰及模型不确定情况下对无人机姿态的良好控制，在快速精确跟踪信号的同时抑制了量测噪声的影响。

1 尾座式无人机建模

尾座式无人机如图1所示。该无人机主要由机身、机翼、副翼、方向舵、升降舵、小翼、螺旋桨和电子设备构成，具有质量轻、尺寸小、对风干扰敏感的特点。垂直起降阶段，该无人机飞行速度低，主要通过螺旋桨的下洗气流流过小翼产生气动力和力矩。通过控制左右螺旋桨的转速差、螺旋桨的倾角和左右小翼的倾角对无人机飞行姿态进行控制。

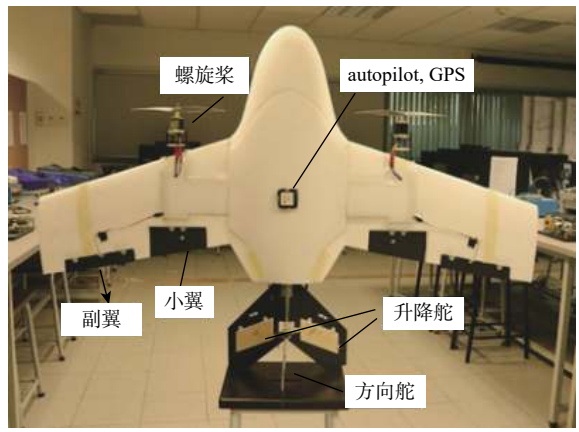


图1 尾座式无人机

Fig. 1 Tail-sitter UAV

1.1 坐标系

1.1.1 地面坐标系$O_{x,y,z}$

地面坐标系的原点为在地面的任意一点O，选定后，过O点在地平面内选择一个方向为x轴的方向，z轴过O点与地面垂直指向地心，y轴在地平面内与x轴、z轴垂直，具体方向根据右手定则确定。

1.1.2 机体坐标系O_{x_b,y_b,z_b}

将无人机的质心作为机体坐标系的原点，y_b和z_b轴分别指向无人机的翼尖和机尾，x_b轴方向根据右手定则确定[24]。

1.1.3 螺旋桨坐标系$O_{x_{r1},y_{r1},z_{r1}}$和$O_{x_{r2},y_{r2},z_{r2}}$

2个螺旋桨坐标系原点选取为左右2个螺旋桨的中心。对于每个螺旋桨坐标系，x_r轴与y_r轴在相应的螺旋桨平面上。z_r轴指向与螺旋桨推力相反的方向，且x_r轴、y_r轴与z_r轴的方向遵守右手定则。当螺旋桨不倾斜时，螺旋桨坐标系的坐标轴平行于机体坐标系的坐标轴。

1.2 运动学方程

$$\begin{bmatrix}\dot{\phi}\\ \dot{\theta}\\ \dot{\psi}\end{bmatrix}=\begin{bmatrix}1 & \sin\phi\tan\theta & \cos\phi\tan\theta\\ 0 & \cos\phi & -\sin\phi\\ 0 & \sin\phi/\cos\theta & \cos\phi/\cos\theta\end{bmatrix}\begin{bmatrix}p\\ q\\ r\end{bmatrix} \tag{1}$$

式中：p、q、r分别为无人机的滚转角速度、俯仰角速度、偏航角速度；ϕ、θ、ψ分别为无人机的滚转角、俯仰角、偏航角。

1.3 力方程组

位置$\boldsymbol{p}=[x_p,y_p,z_p]^{\mathrm{T}}$，速度$\boldsymbol{v}=[u_p,v_p,w_p]^{\mathrm{T}}$，力方程组如下：

$$\begin{bmatrix}\dot{x}_p\\ \dot{y}_p\\ \dot{z}_p\end{bmatrix}=\begin{bmatrix}u_p\\ v_p\\ w_p\end{bmatrix} \tag{2}$$

$$\begin{bmatrix}\dot{u}_p\\ \dot{v}_p\\ \dot{w}_p\end{bmatrix}=\begin{bmatrix}0\\ 0\\ g\end{bmatrix}+\frac{1}{m}\boldsymbol{R}_{\mathrm{BE}}\boldsymbol{R}_{\mathrm{RB}}\begin{bmatrix}0\\ 0\\ -T_{\mathrm{prop}}\end{bmatrix}+\frac{1}{m}\boldsymbol{F}_{\mathrm{surface}} \tag{3}$$

式中: x_p、y_p、z_p为无人机的位置坐标; u_p、v_p、w_p为无人机的线速度; g 为重力加速度; T_{prop}为螺旋桨转动产生的推力; $\boldsymbol{F}_{\text{surface}}$为无人机控制面偏转所产生的力; $\boldsymbol{R}_{\text{RB}}$为机体坐标系与螺旋桨坐标系间的转换矩阵; $\boldsymbol{R}_{\text{BE}}$为机体坐标系与地面坐标系间转换矩阵; m为无人机质量。

假设在垂直起降阶段, 2 个螺旋桨的倾角δ_{prop}大小与方向均相同, 则矩阵$\boldsymbol{R}_{\text{RB}}$可以写为

$$\boldsymbol{R}_{\text{RB}}=\begin{bmatrix}\cos\delta_{\text{prop}} & 0 & \sin\delta_{\text{prop}}\\ 0 & 1 & 0\\ -\sin\delta_{\text{prop}} & 0 & \cos\delta_{\text{prop}}\end{bmatrix} \tag{4}$$

1.4　力矩方程组

在机体坐标系中, 无人机力矩方程如下:

$$\boldsymbol{I}_{\text{f}}\dot{\boldsymbol{\Omega}}_{\text{b}}=-\boldsymbol{\Omega}_{\text{b}}\times(\boldsymbol{I}_{\text{f}}\boldsymbol{\Omega}_{\text{b}})+\boldsymbol{M}_{\text{fin}}+\boldsymbol{M}_{\text{prop}}+\boldsymbol{M}_T \tag{5}$$

式中: $\boldsymbol{I}_{\text{f}}$为无人机惯性矩阵, $\boldsymbol{I}_{\text{f}}=\text{diag}\{If_{xx},If_{yy},If_{zz}\}$; $\boldsymbol{\Omega}_{\text{b}}$为姿态角速度; $\boldsymbol{M}_{\text{fin}}$为小翼偏转产生的力矩; $\boldsymbol{M}_{\text{prop}}$为左右螺旋桨速度差产生的力矩; $\boldsymbol{M}_T$为螺旋桨倾转产生的推力矢量力矩。

$$\boldsymbol{M}_{\text{prop}}=\boldsymbol{R}_{\text{RB}}\begin{bmatrix}0\\0\\m_{\text{prop2}}-m_{\text{prop1}}\end{bmatrix} \tag{6}$$

$$\boldsymbol{M}_T=\boldsymbol{r}_{\text{prop}}\times\begin{bmatrix}\boldsymbol{R}_{\text{RB}}\begin{matrix}0\\0\end{matrix}\\T_{\text{prop}}\end{bmatrix} \tag{7}$$

$$\boldsymbol{M}_{\text{fin}}=\begin{bmatrix}0\\0\\\dfrac{1}{2}\rho_{\text{air}}C_L S l V^2\end{bmatrix} \tag{8}$$

式中: $m_{\text{prop}i}$为螺旋桨坐标下表示的螺旋桨倾转产生的力矩; $\boldsymbol{r}_{\text{prop}}$为从无人机重心到螺旋桨的距离向量; T_{prop}为螺旋桨转动产生的拉力; ρ_{air}为空气密度; C_L为偏航力矩系数; S为小翼参考面积; l为沿y_{b}轴方向小翼中心到无人机重心的力臂; V为螺旋桨产生的滑流速度。

2　无人机 L_1 自适应控制

2.1　无人机模型状态空间描述

L_1 自适应控制需要被控对象的全部状态信息, 故需将姿态回路描述成如下形式:

$$\begin{cases}\dot{\boldsymbol{x}}=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}u+\boldsymbol{\sigma}'\\ y=\boldsymbol{c}\boldsymbol{x}\end{cases} \tag{9}$$

首先, 对式 (1) 进行简化[1]:

$$\begin{cases}\dot{\phi}=p+q\sin\phi\tan\theta+r\cos\phi\tan\theta=p+d_\phi\\ \dot{\theta}=q\cos\phi-r\sin\phi=q+d_\theta\\ \dot{\psi}=(q\sin\phi+r\cos\phi)/\cos\theta=r+d_\psi\end{cases} \tag{10}$$

式中:

$$\begin{cases}d_\phi=q\sin\phi\tan\theta+r\cos\phi\tan\theta\\ d_\theta=q(\cos\phi-1)-r\sin\phi\\ d_\psi=q\sin\phi/\cos\theta-r(1-\cos\phi/\cos\theta)\end{cases} \tag{11}$$

无人机垂直起降阶段, 可将d_ϕ、d_θ、d_ψ视为小量干扰。将无人机力矩方程 (6) 代入式 (11) 可得

$$\begin{cases}\ddot{\phi}=\dfrac{If_{yy}-If_{zz}}{If_{xx}}qr+\dfrac{M_{\text{roll}}}{If_{xx}}+\dot{d}_\phi\\ \ddot{\theta}=\dfrac{If_{zz}-If_{xx}}{If_{yy}}pr+\dfrac{M_{\text{pitch}}}{If_{yy}}+\dot{d}_\theta\\ \ddot{\psi}=\dfrac{If_{xx}-If_{yy}}{If_{zz}}pq+\dfrac{M_{\text{yaw}}}{If_{zz}}+\dot{d}_\psi\end{cases} \tag{12}$$

式中: M_{roll}、M_{pitch}、M_{yaw}分别为滚转、俯仰、偏航通道的合外力矩。

1) 滚转通道。

$$\begin{cases}\boldsymbol{x}_{\text{roll}}=\begin{bmatrix}\phi & \dot{\phi}\end{bmatrix}^{\text{T}}\\ u_{\text{roll}}=M_{\text{roll}}/If_{xx}\\ \boldsymbol{\sigma}'_{\text{roll}}=\begin{bmatrix}0\\ \dfrac{If_{yy}-If_{zz}}{If_{xx}}qr+\dot{d}_\phi\end{bmatrix}\end{cases} \tag{13}$$

2) 俯仰通道。

$$\begin{cases}\boldsymbol{x}_{\text{pitch}}=\begin{bmatrix}\theta & \dot{\theta}\end{bmatrix}^{\text{T}}\\ u_{\text{pitch}}=M_{\text{pitch}}/If_{yy}\\ \boldsymbol{\sigma}'_{\text{pitch}}=\begin{bmatrix}0\\ \dfrac{If_{zz}-If_{xx}}{If_{yy}}pr+\dot{d}_\theta\end{bmatrix}\end{cases} \tag{14}$$

3) 偏航通道。

$$\begin{cases}\boldsymbol{x}_{\text{yaw}}=\begin{bmatrix}\psi & \dot{\psi}\end{bmatrix}^{\text{T}}\\ u_{\text{yaw}}=M_{\text{yaw}}/If_{zz}\\ \boldsymbol{\sigma}'_{\text{yaw}}=\begin{bmatrix}0\\ \dfrac{If_{xx}-If_{yy}}{If_{zz}}pq+\dot{d}_\psi\end{bmatrix}\end{cases} \tag{15}$$

此外, $\boldsymbol{A}=\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$, $\boldsymbol{b}=\begin{bmatrix}0\\1\end{bmatrix}$, $\boldsymbol{c}=\begin{bmatrix}1 & 0\end{bmatrix}$。

2.2　L_1 自适应控制器设计

L_1 自适应控制系统由被控对象、状态预测器、自适应律、控制律 4 部分组成[12-14], 结构如图 2 所示。与传统模型参考自适应控制相比, L_1 自适应控制在控制律部分增加低通滤波器, 不仅能保证输出的渐近稳定性, 还能使系统控制信号、输入信号、输出信号暂态性能良好[25]。

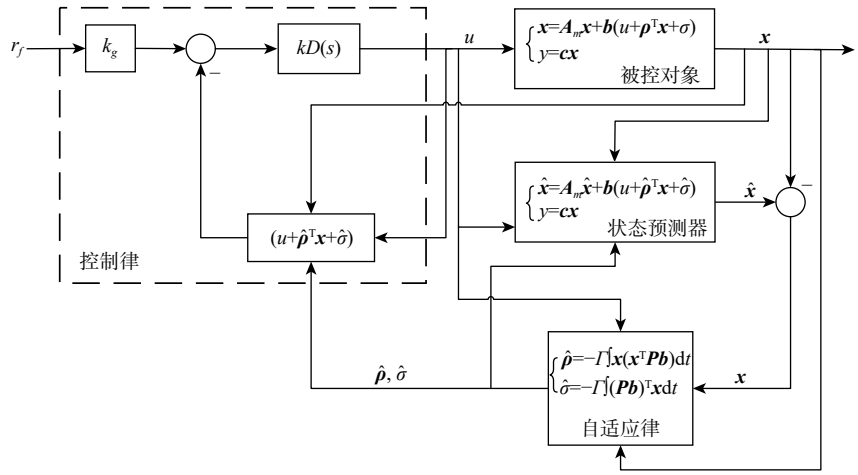


图2 L_1自适应控制系统结构

Fig. 2 L_1 adaptive control system structure

1）被控对象。

由式(9)可得

$$\begin{cases}\dot{\boldsymbol{x}}(t)=\boldsymbol{A}_m\boldsymbol{x}(t)+\boldsymbol{b}(u(t)+\boldsymbol{\rho}^{\mathrm{T}}(t)\boldsymbol{x}(t)+\sigma(t))\\ y(t)=\boldsymbol{c}\boldsymbol{x}(t)\end{cases} \tag{16}$$

式中：$\boldsymbol{\rho}$、σ描述了被控对象的参数不确定性；$\boldsymbol{A}_m=\boldsymbol{A}-\boldsymbol{b}\boldsymbol{K}$且满足霍尔维茨条件；$\boldsymbol{\rho}$与被控对象本身的不确定性相关，$\sigma$与系统存在的扰动相关。

2）状态预测器。

$$\begin{cases}\dot{\hat{\boldsymbol{x}}}(t)=\boldsymbol{A}_m\hat{\boldsymbol{x}}(t)+\boldsymbol{b}(u(t)+\hat{\boldsymbol{\rho}}^{\mathrm{T}}(t)\boldsymbol{x}(t)+\hat{\sigma}(t))\\ y(t)=\boldsymbol{c}\boldsymbol{x}(t)\end{cases} \tag{17}$$

式中：$\hat{\boldsymbol{x}}(t)$、$\hat{\boldsymbol{\rho}}(t)$、$\hat{\sigma}(t)$分别对应$\boldsymbol{x}$、$\boldsymbol{\rho}$、$\sigma$的估计值。

3）自适应律。

定义误差$\tilde{\boldsymbol{x}}(t)=\hat{\boldsymbol{x}}(t)-\boldsymbol{x}(t)$，表示为

$$\dot{\tilde{\boldsymbol{x}}}(t)=\boldsymbol{A}_m\tilde{\boldsymbol{x}}(t)+\boldsymbol{b}(u(t)+\tilde{\boldsymbol{\rho}}^{\mathrm{T}}(t)\boldsymbol{x}(t)+\tilde{\sigma}(t)) \tag{18}$$

式中：$\tilde{\boldsymbol{\rho}}(t)=\hat{\boldsymbol{\rho}}(t)-\boldsymbol{\rho}(t)$；$\tilde{\sigma}(t)=\hat{\sigma}(t)-\sigma(t)$。

定义误差方程的能量函数为

$$V(t)=\frac{1}{2}\tilde{\boldsymbol{x}}^{\mathrm{T}}(t)\boldsymbol{P}\tilde{\boldsymbol{x}}(t)+\frac{1}{2}1/\Gamma(\tilde{\boldsymbol{\rho}}^{\mathrm{T}}(t)\tilde{\boldsymbol{\rho}}(t)+\tilde{\sigma}(t))^2 \tag{19}$$

式中：Γ为自适应增益；$\boldsymbol{P}$为使$\boldsymbol{A}_m^{\mathrm{T}}\boldsymbol{P}+\boldsymbol{P}\boldsymbol{A}_m=-\boldsymbol{I}_n$成立的正定对称阵。

通过使误差方程的能量函数的导数负定来保证误差在李雅普诺夫意义下稳定，推导得到参数$\hat{\boldsymbol{\rho}}(t)$、$\hat{\sigma}(t)$的表达式为

$$\begin{cases}\hat{\boldsymbol{\rho}}(t)=-\Gamma\int\boldsymbol{x}(t)(\tilde{\boldsymbol{x}}^{\mathrm{T}}(t)\boldsymbol{P}\boldsymbol{b})\mathrm{d}t\\ \hat{\sigma}(t)=-\Gamma\int(\boldsymbol{P}\boldsymbol{b})^{\mathrm{T}}\tilde{\boldsymbol{x}}(t)\mathrm{d}t\end{cases} \tag{20}$$

4）控制律。

控制律环节包括2个部分：参考输入的重新构建与低通滤波器$D(s)=1/s$。为保证系统稳定，低通滤波器带宽k需满足式(21)，其中，$C(s)=k/(s+k)$，$\boldsymbol{H}(s)=(s\boldsymbol{I}-\boldsymbol{A}_m)^{-1}\boldsymbol{b}$。为使$\lim\limits_{t\to\infty}y=r_f$，需在参考输入前乘增益$k_g$，如式(22)所示，最终控制律如式(23)所示，其中，$\bar{r}_f(s)$是$\bar{r}_f(t)=\hat{\boldsymbol{\rho}}^{\mathrm{T}}(t)\boldsymbol{x}(t)-\hat{\sigma}(t)$的拉氏变换。

$$\max\sum_{i=1}^{n}|\rho_i|\cdot\|\boldsymbol{H}(s)(1-C(s))\|_{L_1}<1 \tag{21}$$

$$k_g=\lim_{s\to0}\frac{1}{\boldsymbol{c}(s\boldsymbol{I}-\boldsymbol{A}_m)^{-1}\boldsymbol{b}}=-\frac{1}{\boldsymbol{c}\boldsymbol{A}_m^{-1}\boldsymbol{b}} \tag{22}$$

$$u(s)=C(s)\left(k_g r^f(s)-\bar{r}^f(s)\right) \tag{23}$$

2.3 L_1自适应控制方法稳定性分析

设定参考系统为[15]

$$\begin{cases}\dot{\boldsymbol{x}}_{\mathrm{ref}}(t)=\boldsymbol{A}_m\boldsymbol{x}_{\mathrm{ref}}(t)+\boldsymbol{b}(u_{\mathrm{ref}}(t)+\boldsymbol{\rho}^{\mathrm{T}}(t)\boldsymbol{x}_{\mathrm{ref}}(t)+\sigma(t))\\ y_{\mathrm{ref}}(t)=\boldsymbol{c}\boldsymbol{x}_{\mathrm{ref}}(t)\end{cases} \tag{24}$$

$\boldsymbol{x}_{\mathrm{ref}}(s)$为$\boldsymbol{x}_{\mathrm{ref}}(t)$的拉氏变换，可表示为

$$\boldsymbol{x}_{\mathrm{ref}}(s)=\boldsymbol{H}(s)\left(\frac{s}{s+k}\bar{r}_{\mathrm{ref}}(s)-\frac{k}{s+k}\cdot\frac{r_f(s)}{\boldsymbol{c}\boldsymbol{A}_m^{-1}\boldsymbol{b}}\right) \tag{25}$$

由式(25)，$\boldsymbol{x}(s)$可表示为

$$\boldsymbol{x}(s)=\boldsymbol{H}(s)\left(\frac{s}{s+k}\bar{r}_f(s)-\frac{k}{s+k}\cdot\frac{r_f(s)}{\boldsymbol{c}\boldsymbol{A}_m^{-1}\boldsymbol{b}}-\frac{k}{s+k}\tilde{\eta}(s)\right) \tag{26}$$

定义误差$\boldsymbol{e}(t)=\boldsymbol{x}(t)-\boldsymbol{x}_{\mathrm{ref}}(t)$，$\boldsymbol{e}(s)$为$\boldsymbol{e}(t)$的拉氏变换。$\tilde{\eta}(s)$为$\tilde{\eta}(t)=u(t)+\tilde{\boldsymbol{\rho}}^{\mathrm{T}}(t)\boldsymbol{x}(t)+\tilde{\sigma}(t)$的拉氏变换。则有

$$\boldsymbol{e}(s)=(s\boldsymbol{I}-\boldsymbol{A}_m)^{-1}\boldsymbol{b}\left(\frac{s}{s+k}\boldsymbol{\rho}^{\mathrm{T}}\boldsymbol{e}(s)-\frac{k}{s+k}\tilde{\eta}(s)\right) \tag{27}$$

两边取L_1范数可得

$$\|\boldsymbol{e}\|_{L_\infty} \leqslant -\left\|\frac{k}{s+k}\right\|_{L_1} \Bigg/ 1-\left\|\boldsymbol{c}(s\boldsymbol{I}-\boldsymbol{A}_m)^{-1}\boldsymbol{b}\frac{s}{s+k}\right\|_{L_1}\cdot\max\sum_{i=1}^{n}|\rho_i|\|\tilde{\boldsymbol{x}}\|_{L_\infty} \tag{28}$$

根据能量函数可得

$$\|\tilde{\boldsymbol{x}}\|_{L_\infty} \leqslant \sqrt{\max\sum_{i=1}^{n}4\rho_i^2+4\sigma_{\max}^2 \Big/ \min(\boldsymbol{P})\Gamma} \tag{29}$$

可以看出, 当Γ足够大时, $\|\tilde{\boldsymbol{x}}\|_{L_\infty}$趋向于零, 同时$\|\boldsymbol{e}\|_{L_\infty}$趋向于 0, 说明自适应增益足够大时, 系统具有良好的稳态性能及暂态性能。

3　基于 L_1-ITD 的无人机姿态控制

由于 L_1 自适应控制方法设计时需要姿态角及其微分信号, 而对含有量测噪声的信号难以直接微分, 引入 ITD 来实现姿态信号的跟踪与微分, 并与 L_1 自适应控制方法结合, 实现基于 L_1-ITD 的无人机姿态控制。

控制器整体结构如图 3 所示。针对无人机垂直起降阶段易受风扰影响的问题, 在滚转、俯仰、偏航通道分别设计 L_1 自适应控制器, 对系统的不确定性进行估计和补偿; 针对 L_1 自适应控制对量测噪声敏感的问题, 将 L_1 自适应控制方法与 ITD 结合, 减小传感器量测噪声对系统性能的影响; 通过控制指令解算模块由力矩指令生成无人机的螺旋桨转速差、螺旋桨倾角和左右小翼偏角, 实现基于 L_1-ITD 的尾座式无人机姿态控制。

ITD 具体形式如式 (30) 所示[23]。该跟踪微分器与传统跟踪微分器相比, 结构简单, 易于调参, 可以快速准确跟踪输入信号, 具有良好的动态响应及较强的滤波能力。

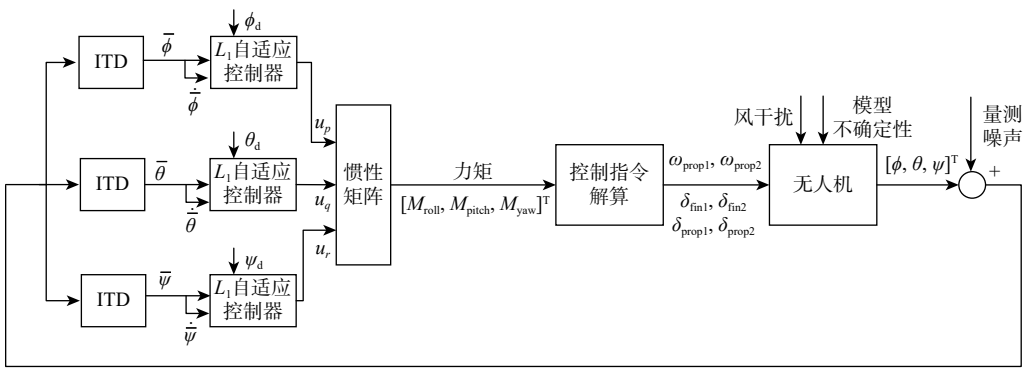


图 3　控制器整体结构

Fig. 3　Structure of controller

$$\begin{cases}\dot{\bar{x}}_1(t)=\bar{x}_2(t)\\ \dot{\bar{x}}_2(t)=R^2\Big[-a_0(\bar{x}_1(t)-v(t))-\\ \quad a_1|\bar{x}_1(t)-v(t)|^j\mathrm{sign}(\bar{x}_1(t)-v(t))-\\ \quad b_0\dfrac{\bar{x}_2(t)}{R}-b_1\left|\dfrac{\bar{x}_2(t)}{R}\right|^j\mathrm{sign}\left(\dfrac{\bar{x}_2(t)}{R}\right)\Big]\end{cases} \tag{30}$$

式中: $\bar{x}_1$为滤去噪声后的姿态角信号; $\bar{x}_2$为$\bar{x}_1$的微分信号; R、a_0、a_1、b_0、b_1、j为调节参数, 且满足$R, a_0, a_1, b_0, b_1>0$, $1>j>0$。

将 ITD 状态变量 $\bar{\boldsymbol{x}}=\begin{bmatrix}\bar{x}_1 & \bar{x}_2\end{bmatrix}^{\mathrm{T}}$作为 L_1 自适应控制方法的输入, 由式 (23) 和式 (30) 可以得到无人机姿态控制律, 如式 (31) 所示, 其中, $\bar{j}'(s)$为$\bar{j}'(t)=\hat{\boldsymbol{\rho}}^{\mathrm{T}}(t)\bar{\boldsymbol{x}}(t)-\hat{\sigma}(t)$的拉氏变换。

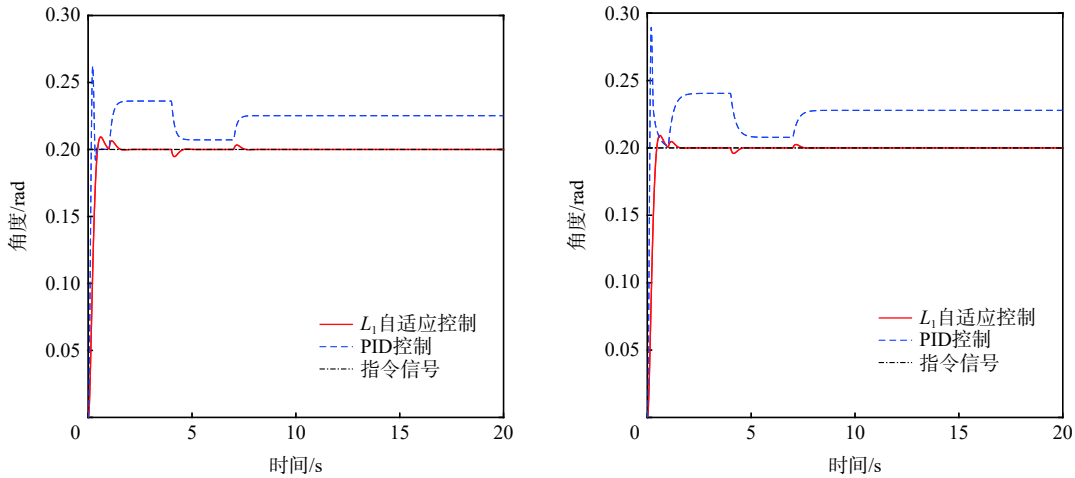
$$u(s)=C(s)\left(k_g j(s)-\bar{j}'(s)\right) \tag{31}$$

4　仿真验证

对第 2 节的无人机 L_1 自适应控制方法进行仿真验证。考虑到实际操纵机构的物理性能, 螺旋桨倾角和小翼偏角的限制范围取为±0.5 rad, 螺旋桨转速限制范围取为±7.4 rad/s。系统扰动选取为分段常值风干扰$\varepsilon(t-1)-0.8\varepsilon(t-4)+0.5\varepsilon(t-7)$、正弦高频风干扰$0.3\sin t-0.8\sin(2t)$, 以及±30% 模型不确定性。$L_1$ 自适应控制器参数选取为$k=70, \Gamma=3\times10^6$。

图 4 为加入分段常值风干扰时 L_1 自适应控制与 PID 控制对比效果曲线, 图 5 为加入正弦高频风干扰时 L_1 自适应控制与 PID 控制对比效果曲线。可知, 在加入风干扰及模型不确定性的情况下, 传统 PID 控制方法无法保证姿态角准确跟踪控制指令, 而 L_1 自适应控制方法可以实现对姿态控制指令的快速准确跟踪。

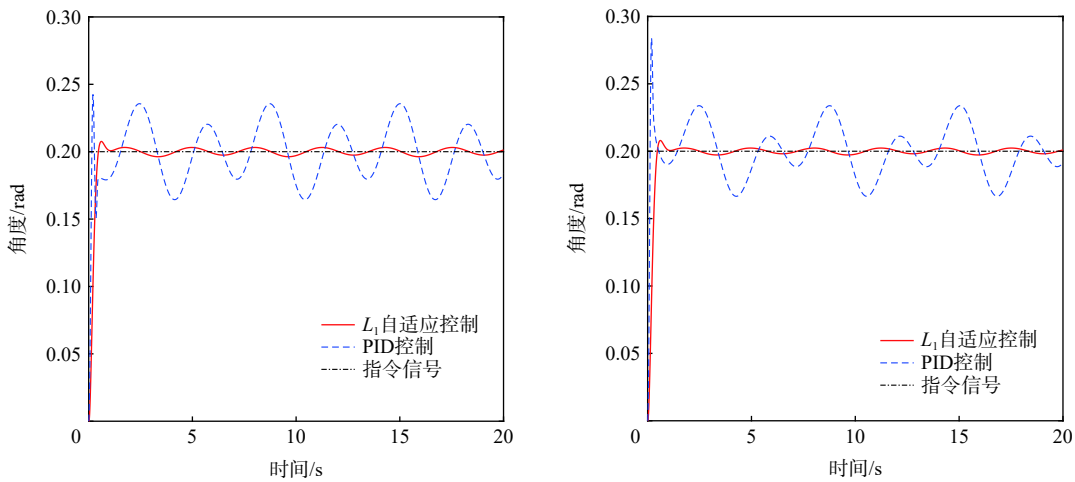
进一步加入±3% 传感器量测噪声进行仿真验证。结合实际垂直起降情况选取紊流风和常值风组合形式的风干扰, 如图 6 所示。L_1-ITD 控制器参数选取为 $k=70$, $\Gamma=3\times10^5$, $R=25$, $b_0=1$, $b_1=1$, $a_0=4$, $a_1=4$。传统 L_1 自适应控制方法的姿态响应如图 7 所示。基于 L_1-ITD 控制方法的无人机姿态



(a) 俯仰角 (b) 偏航角

图 4 分段常值风干扰下 L_1 自适应控制与 PID 控制效果对比曲线

Fig. 4 Under segmented constant wind disturbance comparison curves between L_1 adaptive control and PID control



(a) 俯仰角 (b) 偏航角

图 5 正弦高频风干扰下 L_1 自适应控制与 PID 控制效果对比曲线

Fig. 5 Under sinusoidal high-frequency wind disturbance comparison curves between L_1 adaptive control and PID control

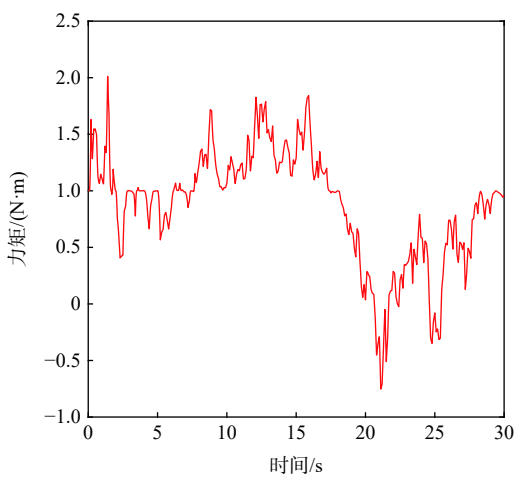


图 6 组合风干扰

Fig. 6 Combined wind disturbance

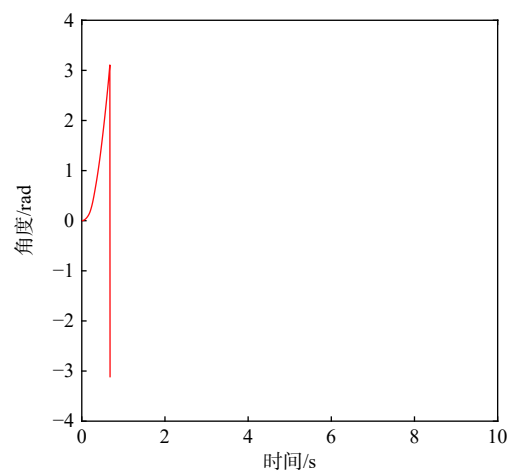


图 7 加入量测噪声后 L_1 自适应控制方法效果曲线

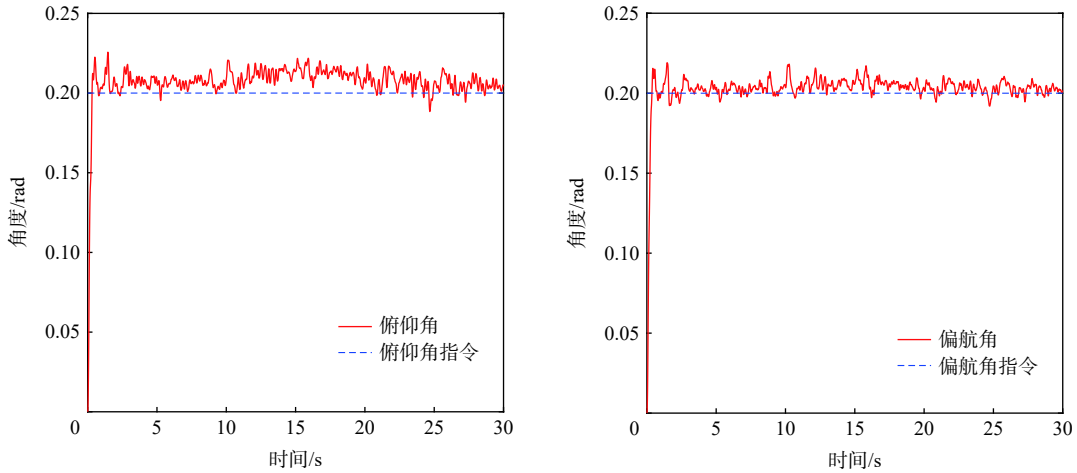
Fig. 7 L_1 adaptive control curve after adding measurement noise

响应如图 8 所示，相应的操纵机构控制指令变化如图 9 所示。

由图 7 可知，传统 L_1 自适应控制对姿态信号直接微分时会导致有效信号淹没于噪声中，将上述信

号作为 L_1 自适应控制器的输入会导致系统发散。

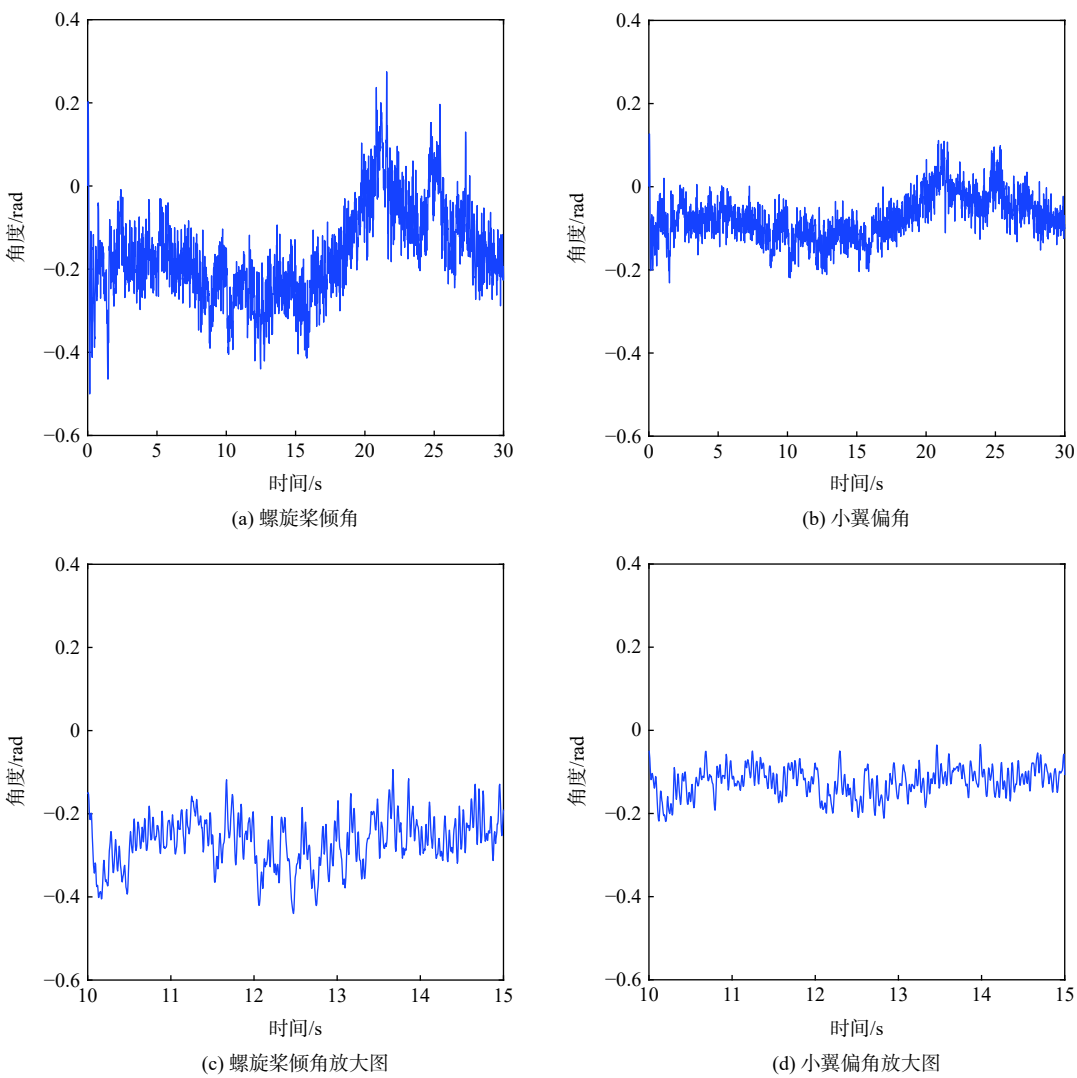
由图 8、图 9 可知, 采用本文提出的 L_1-ITD 控制方法, 在加入±3% 姿态量测噪声和组合风干扰情况下, 可以保证姿态角快速准确跟踪控制指令, 误差在±0.02 rad 范围之内, 且各操纵机构控制指令和变化速率均在物理限制范围内。



(a) 俯仰角　　(b) 偏航角

图 8　基于 L_1-ITD 控制效果

Fig. 8　Control effect based on L_1-ITD



(a) 螺旋桨倾角　　(b) 小翼偏角

(c) 螺旋桨倾角放大图　　(d) 小翼偏角放大图

图 9　各操纵机构控制指令变化

Fig. 9　Changes in control commands of each operating mechanism

5 结 论

1）针对尾座式无人机垂直起降阶段姿态控制问题，本文设计了基于L_1自适应控制方法的尾座式姿态控制律，在考虑风干扰及模型不确定的情况下，实现对姿态控制指令精准追踪。

2）针对L_1自适应控制方法对量测噪声敏感及存在量测噪声时操纵机构控制指令剧烈抖动导致执行器件损伤等问题，本文提出了基于L_1-ITD的尾座式无人机姿态控制方法，在系统存在风扰、模型不确定性及量测噪声的情况下，无人机仍可快速跟踪指令，姿态控制精度满足要求。

参考文献（References）

[1] 刘玉焘. 尾座式无人机的飞行控制器设计[D]. 哈尔滨: 哈尔滨工业大学, 2014: 1-33.
LIU Y T. Design of flight controller for a tail-sitter UAV[D]. Harbin: Harbin Institute of Technology, 2014: 1-33(in Chinese).

[2] 王刚强. 大气扰动下无人直升机悬停/小速度段位置控制律设计[D]. 南京: 南京航空航天大学, 2013: 1-3.
WANG G Q. Research on position control technology of the unmanned helicopter in hover and low speed state with atmospheric disturbance[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2013: 1-3(in Chinese).

[3] 王正熙, 陈洋, 郑秀娟, 等. 风扰下基于气动参数估计的四旋翼无人机控制[J]. 信息与控制, 2018, 47(6): 663-670.
WANG Z X, CHEN Y, ZHENG X J, et al. Quadrotor UAV control with disturbance based on aerodynamic parameter estimation[J]. Information and Control, 2018, 47(6): 663-670(in Chinese).

[4] KIM M, KIM Y. Error dynamics-based guidance law of UAVs for target observation under wind disturbance: AIAA 2012-2593[R]. Reston: AIAA, 2012.

[5] WASLANDER S, WANG C. Wind disturbance estimation and rejection for quadrotor position control: AIAA 2009-1983[R]. Reston: AIAA, 2009.

[6] WANG Y N, YU J L, LI Q D, et al. Control strategy for the transition flight of a tail-sitter UAV[C]//2017 Chinese Control Conference. Piscataway: IEEE Press, 2017: 3504-3509.

[7] 韩京清. 自抗扰控制器及其应用[J]. 控制与决策, 1998, 13(1): 19-23.
HAN J Q. Auto-disturbances-rejection controller and it's applications[J]. Control and Decision, 1998, 13(1): 19-23(in Chinese).

[8] WANG Q Q, XIONG H J, QIU B. The attitude control of transmission line fault inspection UAV based on ADRC[C]//2017 International Conference on Industrial Informatics-Computing Technology, Intelligent Technology, Industrial Information Integration. Piscataway: IEEE Press, 2018: 186-189.

[9] ZHU G J, QI J T, WU C. Landing control of fixed-wing UAV based on ADRC[C]//2019 Chinese Control Conference. Piscataway: IEEE Press, 2019: 8020-8025.

[10] LEE D, AWAN A, KIM S, et al. Adaptive control for a VTOL UAV operating near a wall: AIAA 2012-4835[R]. Reston: AIAA, 2012.

[11] LIAN S K, MENG W, LIN Z M, et al. Adaptive attitude control of a quadrotor using fast nonsingular terminal sliding mode[J]. IEEE Transactions on Industrial Electronics, 2022, 69(2): 1597-1607.

[12] CAO C Y, HOVAKIMYAN N. Guaranteed transient performance with L_1 adaptive controller for systems with unknown time-varying parameters and bounded disturbances: Part I[C]//2007 American Control Conference. Piscataway: IEEE Press, 2007: 3925-3930.

[13] LUO J, CAO C, HOVAKIMYAN N. L_1 adaptive controller for a class of systems with unknown nonlinearities[C]//2010 American Control Conference. Piscataway: IEEE Press, 2010: 1659-1664.

[14] CAO C, HOVAKIMYAN N. L_1 adaptive controller for nonlinear systems in the presence of unmodelled dynamics: Part II[C]//2008 American Control Conference. Piscataway: IEEE Press, 2008: 4099-4104.

[15] 秦奇. 飞行器纵向轨迹的L_1自适应控制[D]. 北京: 北京理工大学, 2015: 1-39.
QIN Q. L_1 adaptive control for aircraft longitudinal trajectory[D]. Beijing: Beijing Institute of Technology, 2015: 1-39(in Chinese).

[16] 陈海, 何开锋, 钱炜祺. 基于非线性L_1自适应动态逆的飞行器姿态角控制[J]. 控制理论与应用, 2016, 33(8): 1111-1118.
CHEN H, HE K F, QIAN W Q. Attitude control of flight vehicle based on a nonlinear L_1 adaptive dynamic inversion approach[J]. Control Theory & Applications, 2016, 33(8): 1111-1118(in Chinese).

[17] ZHONG J Y, SONG B F, LI Y B, et al. L_1 adaptive control of a dual-rotor tail-sitter unmanned aerial vehicle with input constraints during hover flight[J]. IEEE Access, 2019, 7: 51312-51328.

[18] 甄红涛, 齐晓慧, 李杰, 等. 四旋翼无人机L_1自适应块控反步姿态控制器设计[J]. 控制与决策, 2014, 29(6): 1076-1082.
ZHEN H T, QI X H, LI J, et al. Quadrotor UAV L_1 adaptive block backstepping attitude controller[J]. Control and Decision, 2014, 29(6): 1076-1082(in Chinese).

[19] 韩京清, 王伟. 非线性跟踪-微分器[J]. 系统科学与数学, 1994, 14(2): 177-183.
HAN J Q, WANG W. Nonlinear tracking-differentiator[J]. Journal of Systems Science and Mathematical Sciences, 1994, 14(2): 177-183(in Chinese).

[20] 史永丽, 侯朝桢. 改进的非线性跟踪微分器设计[J]. 控制与决策, 2008, 23(6): 647-650.
SHI Y L, HOU C Z. Design of improved nonlinear tracking differentiator[J]. Control and Decision, 2008, 23(6): 647-650(in Chinese).

[21] 刘延泉, 郭佳颖. 双曲正切跟踪微分器设计及相平面分析[J]. 电力科学与工程, 2017, 33(10): 74-78.
LIU Y Q, GUO J Y. Design and phase plane analysis of a hyperbolic tangent tracking differentiator[J]. Electric Power Science and Engineering, 2017, 33(10): 74-78(in Chinese).

[22] 王新华, 陈增强, 袁著祉. 全程快速非线性跟踪-微分器[J]. 控制理论与应用, 2003, 20(6): 875-878.
WANG X H, CHEN Z Q, YUAN Z Z. Nonlinear tracking-differentiator with high speed in whole course[J]. Control Theory & Applications, 2003, 20(6): 875-878(in Chinese).

[23] 刘海亭. 基于滑模变结构控制的机械臂轨迹跟踪[D]. 天津: 天津大学, 2018.
LIU H T. Trajectory tracking of robotic manipulators based on sliding mode variable structure control[D]. Tianjin: Tianjing University, 2018(in Chinese).

[24] KE Y J, WANG K L, CHEN B M. Design and implementation of a hybrid UAV with model-based flight capabilities[J]. IEEE/ASME Transactions on Mechatronics, 2018, 23(3): 1114-1125.

[25] LIU X X, LI Y, YUAN M Q, et al. Gust alleviation controller for elastic aircraft based on L_1 adaptive control[C]//2017 Chinese Automation Congress. Piscataway: IEEE Press, 2018: 5382-5385.

Attitude control law based on L_1-ITD for a tail-sitter UAV

WANG Shuyang，ZHANG Jing*，YANG Lingyu

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

Abstract：Aiming at the problem that the tail-sitter UAV is sensitive to wind disturbance in the vertical take-off and landing (VTOL) stage, an attitude control law based on L_1-ITD is proposed. Firstly, a 6-DOF nonlinear model of a tail-sitter UAV in VTOL stage is established, and an L_1 adaptive attitude controller for the UAV is designed. The controller can suppress the influence of disturbance on system performance and achieve good attitude control performance with wind disturbance and model uncertainty in VTOL stage. Then, aiming at the problem that L_1 adaptive control method is sensitive to measurement noise and cannot directly obtain effective differential signal, an improved tracking differentiator is used to track the signal quickly and accurately while suppressing the influence of measurement noise. Finally, the simulation results demonstrate the effectiveness of the proposed control method.

Keywords：L_1 adaptive control; tracking differentiator; tail-sitter UAV; attitude control; measurement noise

Received：2021-11-11；**Accepted**：2022-01-21；**Published Online**：2022-02-16 09：05

URL：kns.cnki.net/kcms/detail/11.2625.V.20220215.1630.003.html

* **Corresponding author**. E-mail：zhangjing2013@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0683

民用飞机客舱 CO_2 分布特性分析

高经诚[1]，陈维建[1]，胡万俊[2]，冯诗愚[1,*]，刘豪正[1]

(1. 南京航空航天大学 航空学院，南京 210016；2. 四川泛华航空仪表电器有限公司 质量管理部，成都 610599)

摘　　要：为保证乘客舒适环境与电子设备等正常工作，采用计算流体力学（CFD）方法，对民用飞机座舱内流场进行模拟，通过改变回风和新风温度、巡航高度及回风比，得到不同工况下舱内 CO_2 体积分数、温度场及速度场分布，并与已有数据对比，验证了计算模型的正确性。研究结果显示：地面标准天及巡航状态舱内流场基本符合要求，冬天供风 50 ℃ 及回风比 0.5 时，舱内部分区域温度较高，且 CO_2 体积分数升高并接近标准值，可根据实际需要确定环控系统方案。

关　键　词：飞机座舱；CO_2；计算流体力学；数值模拟；民机适航

中图分类号：V243

文献标志码：A　　**文章编号**：1001-5965(2023)09-2510-08

近年来，人类飞机出行的需求日益增大，环控系统成为现代飞机的必须系统。民用飞机客舱内部空间狭窄密闭，人员密度大，若气流组织方案设计不合理将威胁乘客健康[1]，且相关疾病传播的可能性也增加[2]。

CO_2 作为人体代谢产物，其浓度可间接反映室内通风效果好坏，当 CO_2 体积分数较高时，会引起血液中碳酸浓度增大，造成酸中毒，且不同温度下人体对 CO_2 耐受程度也不同[3]。民机适航条款 25 部第 831 条规定飞行期间舱内 CO_2 体积分数不得超过 0.5% 体积含量(海平面当量)。另外，舱内温度场、速度场情况同样值得重视，要求舱内横向、纵向温度差值不能过大[4]。

Mazumdar 和 Chen 建立了一维分析模型，模拟舱内污染物传播情况[5]。波音公司针对飞机座舱建立二维模型，并用 Fluent 模拟计算座舱内部气流组织情况[6]。但这些研究均不是三维模型，其实际应用性有待提高。

代炳荣等采用计算流体力学(computational fluid dynamics, CFD)方法，建立简化座舱三维模型，研究了无乘客工况下，不同送风方式对舱内流场的影响[7]；黄彦祺采用 CFD 方法，利用改造的室内环境模拟程序，研究不同送风方案对气流组织的影响[8]；陆月星和李沃森以座舱内空气为冷源，得到了舱内不同位置空气对人体各部位散热的影响[9]；王乔蓬研究了不同送风角度和行李架覆面结构对座舱内气流组织的影响[10]。当前相关研究涵盖了一维至三维的数值模拟，分别包含座舱供风方式、角度及行李架覆面结构等因素的影响，但未能同时综合考虑供风条件、舱内结构、巡航高度及空调回风比等各因素对舱内流场的影响。

民机适航条款规定通风系统必须能提供足够量的未被污染的空气，目前中国民机研制正在逐渐追求环控系统能源的节约，也需要在不同情况下满足适航要求。因此，为验证不同新风+回风温度、巡航高度下舱内流场情况，以及不同回风比对舱内污染物分布的综合影响，本文研究以 CFD 方法为基础，基于现有国内外研究成果，利用数值方法计算

收稿日期：2021-11-12；**录用日期**：2022-03-04；**网络出版时间**：2022-03-23 09:30

网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220321.2040.006.html

基金项目：国家自然科学基金 (U1833121)；中央高校基本科研业务费专项资金；江苏高校优势学科建设工程；南京航空航天大学研究生科研与实践创新计划 (xcxjh20210108,xcxjh20220102)

* **通信作者**. E-mail：shiyuf@nuaa.edu.cn

引用格式：高经诚，陈维建，胡万俊，等. 民用飞机客舱 CO_2 分布特性分析 [J]. 北京航空航天大学学报，2023，49（9）：2510-2517.

GAO J C，CHEN W J，HU W J，et al. Analysis of CO_2 distribution characteristics in cabin of civil aircraft[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49（9）：2510-2517（in Chinese）.

进行设计和评估，能够节约实验成本，缩短实验周期，为民用飞机环控系统气流组织方案提供参考。

1　座舱 CO_2 及物理模型

1.1　座舱 CO_2 来源

1）空调新风。民用飞机的环控系统大多采用天花板送风、侧壁送风、地板处送风及个人通风的方式[11-12]，其送风口送入座舱的空气和地面大气成分基本类似，故全新风的情况下，可将送风口 CO_2 体积分数近似看为 0.03%，而有回风时则利用 UDF 功能，即用户自定义函数，按回风比计算。

2）人体口鼻。除了环控系统正常通风产生的 CO_2 以外，机上乘客及机组人员呼吸也会产生大量的 CO_2，查阅资料并换算可得机上人员每人由口鼻处产生的呼吸质量流量约为 1.72×10^{-4} kg/s，而舱段模型中共有 8 名乘客，因此，口鼻质量流量设为 1.376×10^{-3} g/s，而其中 CO_2 体积分数约为 4.17%。

3）空调回风。为了节省能源，实现能源循环利用，当代飞机的环控系统通常会在气流组织方案上采用有回风的形式，即有一定百分比的风量送回舱内，因此该部分回风中的 CO_2 又会进入舱内，对舱内空气环境造成影响[13]。本文研究假设客舱模型内乘客开启全部空调风口。

1.2　座舱几何模型

本文选取典型支线飞机客舱为模型，取对称面一侧的 4 排座椅舱段为研究对象，如图 1 所示，所取舱段每排座椅有 2 名乘客，设有天花板送风口和靠近地处的回风口。

对模型适当简化，基于本文仅研究飞机座舱内部流场，因此，可保留座舱内壁面为边界，简化窗户、外蒙皮的部分，并不考虑壁面厚度。

同时，对人椅部分进行适当简化，如图 2 所示。

1.3　网格划分

基于飞机座舱属于复杂几何模型，为保证网格质量，研究利用 ICEM 软件对飞机座舱模型进行结构化的网格划分。所得人椅部分网格和座舱整体网格划分如图 3、图 4 所示。

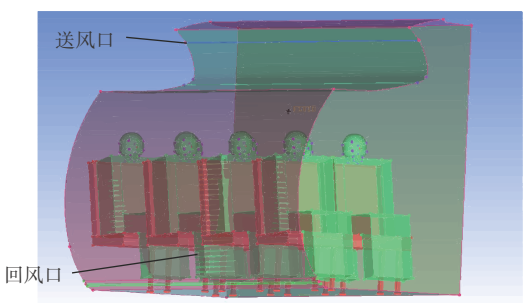


图 1　座舱几何模型

Fig. 1　Geometric model of cabin

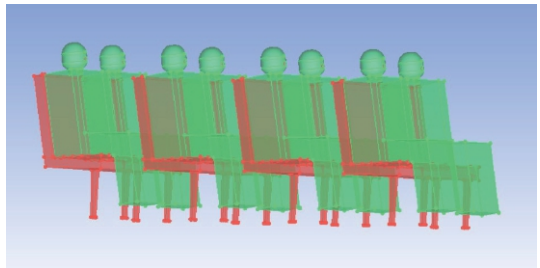


图 2　人椅几何模型

Fig. 2　Geometric model of man and chair

模型网格划分后生成网格数量为 370 万，其中网格比例因子为 1，最大边长为 100 mm，网格最小内角为 24.84°，网格整体质量均在 0.4 及以上。可见网格质量达到预期要求。

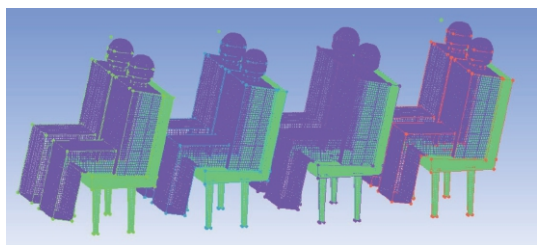


图 3　人椅部分网格划分

Fig. 3　Meshing of human and chair

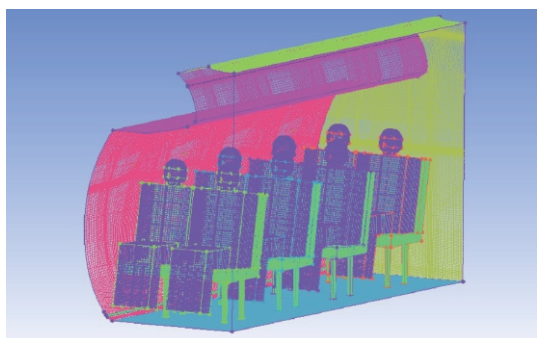


图 4　整体网格生成

Fig. 4　Global mesh generation

1.4　网格无关性验证

在生成网格时，同时对舱段模型划分了 3 套网格：19 万、370 万和 2 300 万，所有网格均质量良好。通过试算巡航高度有回风时舱内流场情况，比较不同数量网格计算所得结果。其中，人体中心 CO_2 分布对比如图 5～图 7 所示。

从图 5～图 7 中可见，后 2 种网格数量的流场计算结果偏差不大，当网格数量大于 370 万时，舱段及人体周围 CO_2 分布基本不变。同时对比了速度场和温度场，也趋于稳定，故选用 370 万数量的网格可兼顾计算收敛性和网格无关性，网格无关性验证成功。

(a) 283 K, 回风比0.3

(b) 283 K, 回风比0.5

(c) 303 K, 回风比0.3

(d) 303 K, 回风比0.5

图 5 网格数量 19 万时 CO_2 分布

Fig. 5 CO_2 distribution when 190 000 grids

(a) 283 K, 回风比0.3

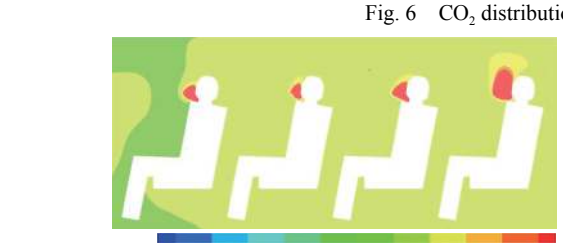
(b) 283 K, 回风比0.5

(c) 303 K, 回风比0.3

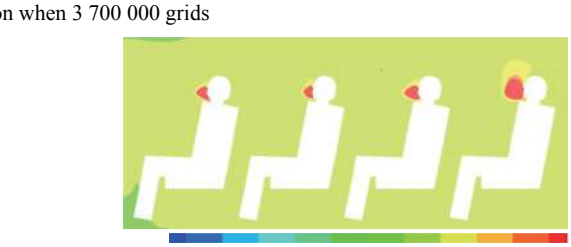
(d) 303 K, 回风比0.5

图 6 网格数量 370 万时 CO_2 分布

Fig. 6 CO_2 distribution when 3 700 000 grids



(a) 283 K, 回风比0.3



(b) 283 K, 回风比0.5

(c) 303 K, 回风比0.3

(d) 303 K, 回风比0.5

图 7 网格数量 2 300 万时 CO_2 分布

Fig. 7 CO_2 distribution when 23 000 000 grids

2　模拟方法和边界条件

本文针对多工况下座舱内 CO_2 的分布特性及温度场和速度场的分布情况, 并着重关注舱内乘客人体周围的 CO_2 体积分数情况。舱内供风参数为: 天花板供风, 夏天送风温度 10 ℃, 冬天送风 30 ℃、40 ℃、50 ℃。模拟飞机在地面标准天状态及 7 620 m 巡航时的舱内空气流场。

2.1　计算方法

本文采用耦合求解计算方法, 压力等各参数的离散均选用二阶迎风格式, 最终根据残差大小及出口流量稳定情况来判定收敛。

除此之外, 本文对典型工况进行了有回风条件的模拟计算, 即通过 UDF 功能编译函数, 载入 Fluent 中分别进行了回风比 0.3 和 0.5 的计算, 与全新风工况对比分析。

2.2　自然对流和多组分设定

开启能量方程, 打开浮升力及重力影响作用, 设定气体为不可压缩的理想气体。由于所取模型为客机的一个舱段, 可将前后壁面及中轴面全部设为对称面。通过以上设定模拟舱内的自然对流。

简化舱内多组分气体为一般气体和 CO_2, 即向混合气体中添加主组分空气及副组分 CO_2。

2.3　湍流模型选取

本文采用 RNG k-ε(2 equ)湍流模型来模拟舱内流动, 该湍流模型由统计实验得出, 相对于标准 k-ε 模型来说精度更高, 可应用于更广泛的范围, 具有更高的可信度, 且该模型提供了针对低雷诺数的湍流模型计算[14]。

2.4　边界条件参数

1) 尺寸。计算时调整尺寸放大 1 000 倍, 即模拟以 m 为单位的真实飞机舱段的流场情况。

2) 气体入口。为方便污染物浓度的设定, 将人体口鼻处和座舱送风口均设定为质量流量入口。其中, 送风口的质量流量根据民机客舱气流组织方案, 设为 0.058 kg/s, 方向为垂直壁面, 温度由多工况条件决定。口鼻处质量流量设定为 0.001 376 kg/s, 方向为 30°斜向下, 温度为 31 ℃。

3) 气体出口。回风口设定为压力出口。

4) 人椅表面。设定人体表面为 30 ℃, 座椅 24 ℃左右。

5) 其他边界条件温度设定。夏天舱外温度高于舱内, 故将侧壁温度设为 27 ℃, 相当于通过设置侧壁壁温的方式引入外部热载荷, 为保证舱内温度维持在 24 ℃ 左右, 设定进口温度并进行试算, 以确定天花板、行李架等其他边界的温度。冬天舱外温度低于舱内, 可将侧壁设为略微较低温度, 并在此 3 种送风温度的工况下进行试算调整, 保证舱内温度在 24 ℃ 左右[15]。

采用以上方法模拟各种工况下的舱内空气流场, 并获得舱内 CO_2 分布特性及温度场、速度场的云图, 对比并分析所得结果。

3　数值模拟结果分析

对计算结果进行后处理, 并取图 8 所示的几种典型的切面示意图, 包括位于人体口鼻前面的 x_1 切面, 2 个位于人体中心的 y 切面(y_1 为靠侧壁处切面, y_2 为靠过道处切面), z_1、z_2、z_3 分别为回风口处、人体腿部上方、口鼻高度处的切面。

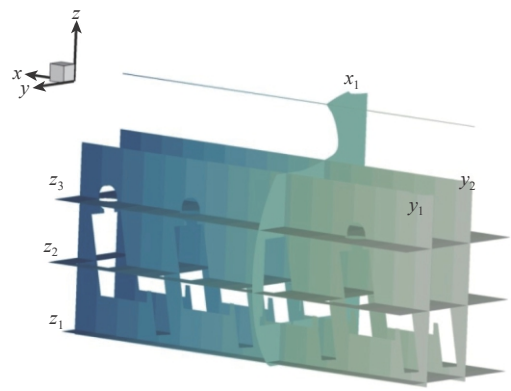


图 8　典型切面位置示意图

Fig. 8　Schematic diagram of typical section position

利用 Tecplot 对以上所有切面在不同工况下的温度、速度、CO_2 分布情况进行后处理, 并对比分析流场, 以下为几个典型切面的空气流场云图。

3.1　送风口温度对舱内流场影响

图 9～图 11 为夏天送风 10 ℃ 及冬天送风 30 ℃、40 ℃ 及 50 ℃ 时地面标准天状态舱内流场大致情况。

从图 9～图 11 中可见, 夏天供风 10 ℃ 时, 空气对流最强, 人体腿部以下气流吹袭感觉较强, 座舱中过道气流速度略高; 冬天供风 40 ℃ 和 50 ℃ 时, 口鼻处 CO_2 的扩散效应略微增强, 且近舱壁一侧 CO_2 体积分数略高于内排。

同时, 大部分情况下舱内温度可维持在 24 ℃, 只有冬天供风 50 ℃ 时, 舱内部分区域尤其是人体头部周围温度略高, 超过舒适范围, 值得引起注意。

3.2　巡航高度对舱内流场影响

图 12 和图 13 为 7 620 m 巡航状态下舱内流场的 CO_2 分布情况。

从图 12～图 13 中可见, 巡航高度 7 620 m 时, 舱内温度场和速度场基本可控制与地面标准天相符, CO_2 体积分数略有升高但不显著, 不会对乘客造成影响, 其他参数均仍满足要求。

3.3　回风比对舱内流场影响

在考虑回风的情况下, 通过 UDF 进行出口回风和进口新风的 CO_2 组分计算, 作为进口的实际

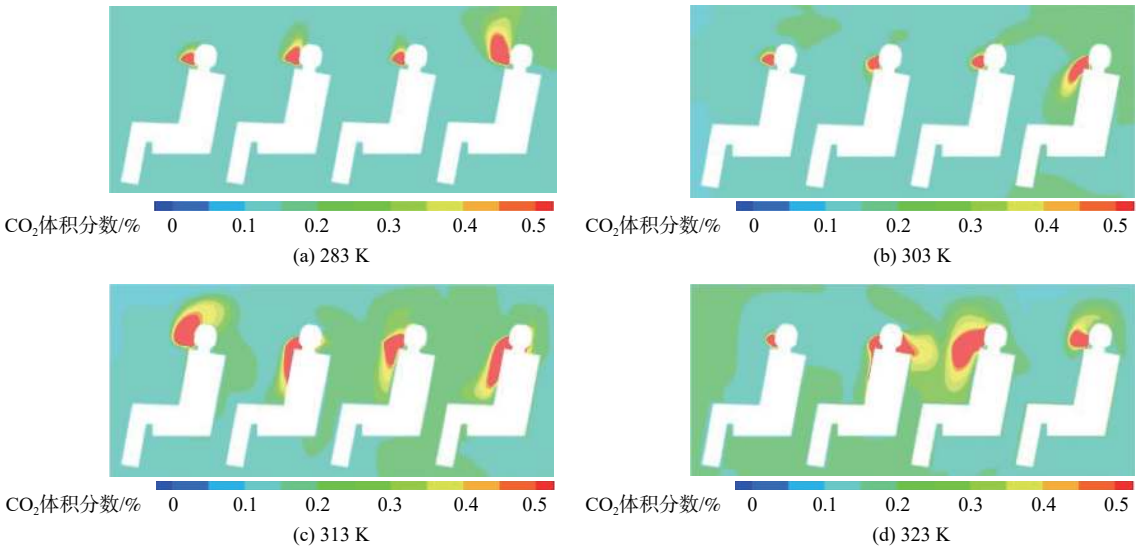


图 9 近壁侧人体中心 CO_2 对比(地面)

Fig. 9 Comparison of carbon dioxide in center of human body near the wall side (ground)

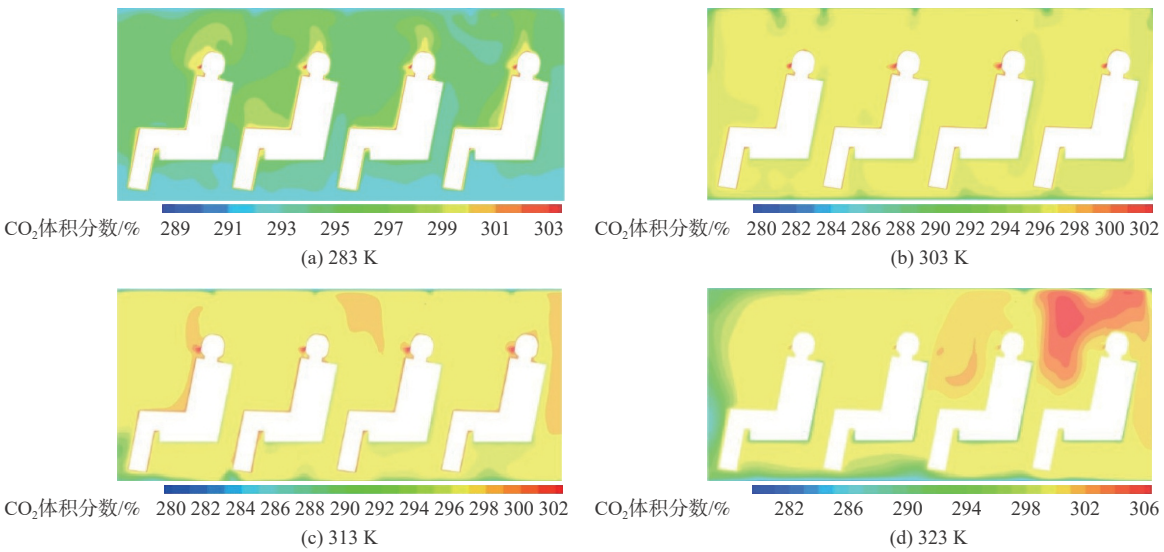


图 10 过道侧人体中心温度场对比(地面)

Fig. 10 Comparison of temperature field of human body center at aisle side (ground)

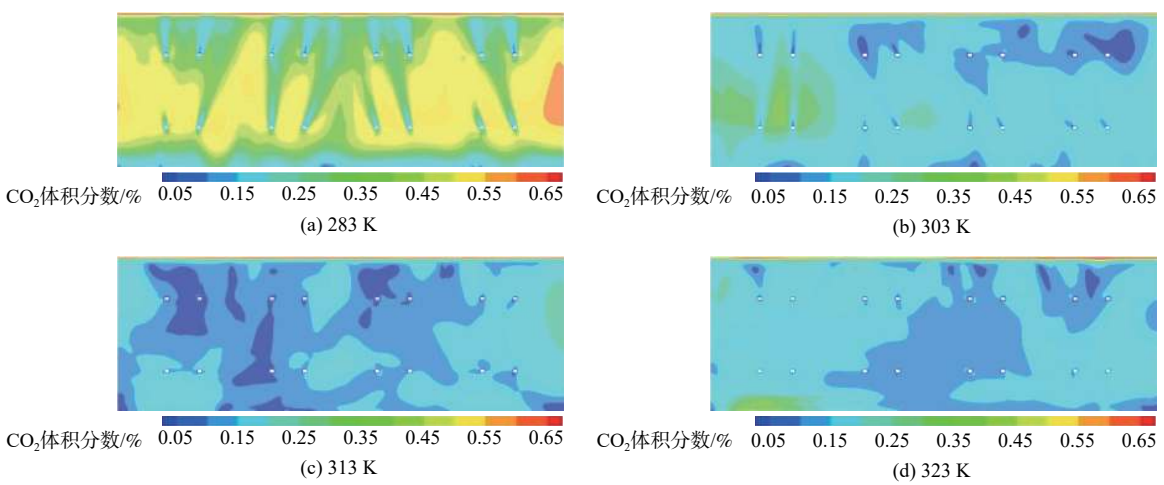


图 11 回风口切面速度场对比(地面)

Fig. 11 Comparison of sectional velocity field of return air outlet (ground)

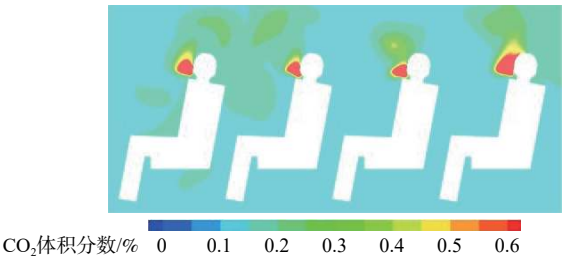


图 12　供风 10 ℃ 时巡航高度人体中心 CO_2 分布(283 K)

Fig. 12　Distribution of CO_2 in center of human body at cruise altitude at 10 ℃ of air supply

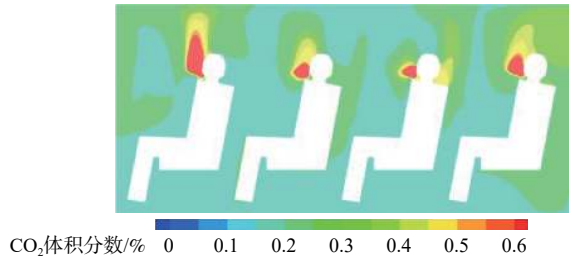


图 13　供风 30 ℃ 时巡航高度人体中心 CO_2 分布(303 K)

Fig. 13　Distribution of CO_2 in center of human body at cruise altitude at 30 ℃ of air supply

CO_2 组分。针对飞机在 7 620 m 巡航高度情况下, 分别计算回风比 0.3 和 0.5 的夏天送风 10 ℃ 和冬天送风 30 ℃ 的工况, 所得结果如图 14～图 16 所示。

从图 14～图 16 中可见, 回风比为 0.3 时舱内流场无明显变化, 而回风比 0.5 时舱内 CO_2 体积分数明显增大, 且扩散效应增强。

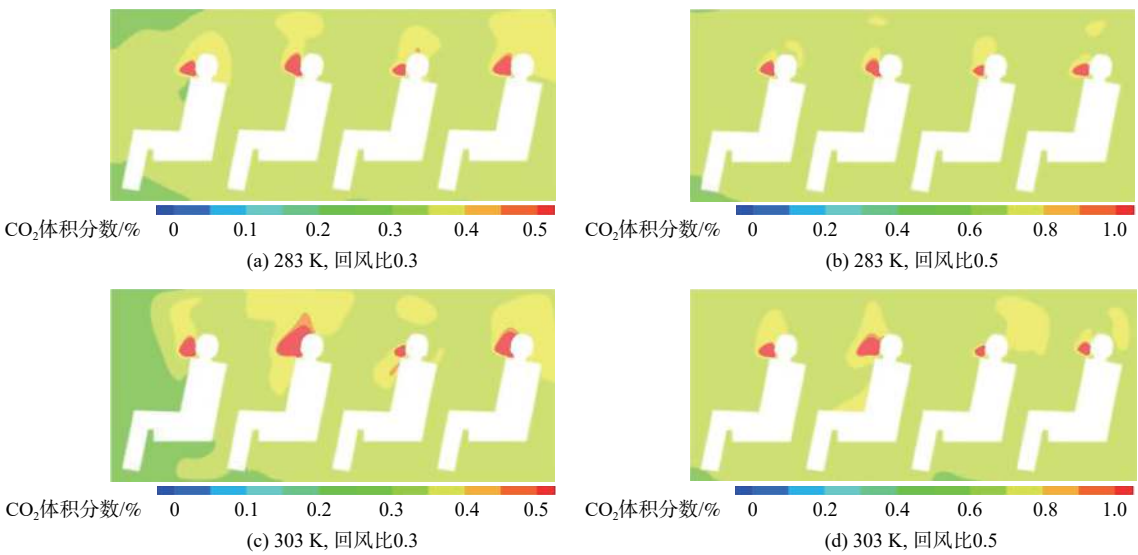


图 14　有回风时人体中心 CO_2 分布对比(巡航)

Fig. 14　Comparison of CO_2 distribution in human body center with return air (Cruise)

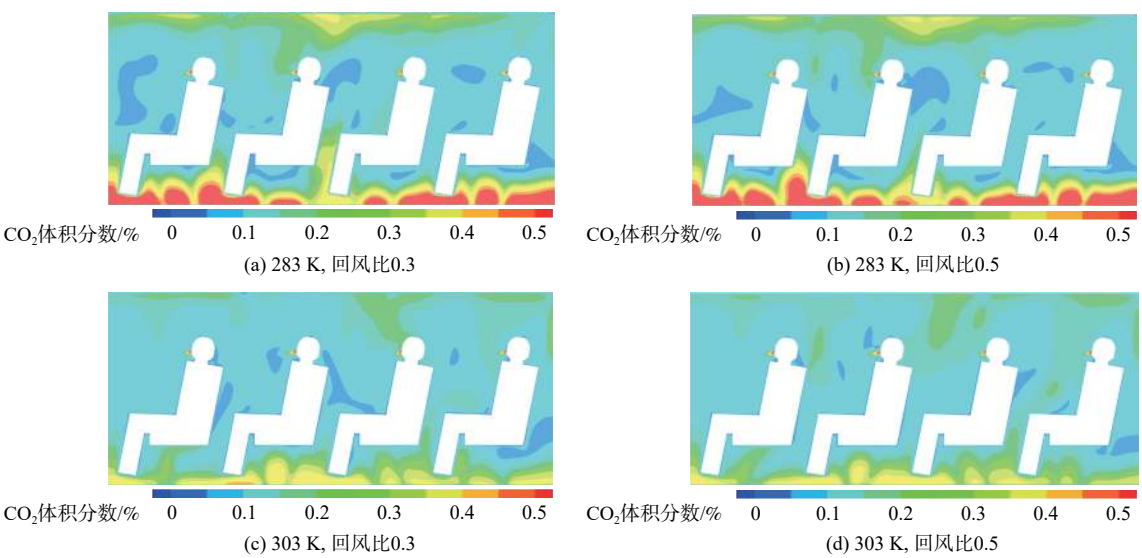


图 15　有回风时人体中心速度场对比(巡航)

Fig. 15　Comparison of velocity field of human body center with return air (Cruise)

(a) 283 K, 回风比0.3

(b) 283 K, 回风比0.5

(c) 303 K, 回风比0.3

(d) 303 K, 回风比0.5

图 16 有回风时人体中心温度场对比(巡航)

Fig. 16 Comparison of temperature field of human body center with return air (Cruise)

4 结 论

本文分析不同工况舱内空气流场,并与已有数据对比,验证数值方法合理性和计算模型正确性。大部分工况下舱内的 CO_2 体积分数、温度场和速度场基本符合客舱设计和运行要求;综合考虑各因素对舱内流场影响,当供风温度达到 50 ℃、回风比为 0.5 时,无论地面还是巡航高度都出现 CO_2 体积分数明显升高,舱内部分位置温度较高,值得引起重视。

1) 回风比为 0.5 时,舱内 CO_2 体积分数明显提高,并接近适航允许的标准值,若回风量进一步增大而供风量持续减小,CO_2 体积分数极易超标,对乘客造成影响。因此,需要关注飞机迫降及环控空调系统出现障碍时的舱内环境。

2) 巡航高度 7 620 m 时,舱内流场基本符合要求。

3) 大部分情况舱内温度场符合要求,夏天供风 10 ℃ 时,舱内对流较强;冬天供风 50 ℃ 时,舱内部分区域温度略高。

民用飞机研制需进行相应的适航,同时也不断向节约能源的方向发展。本文利用数值方法计算,研究了民用飞机供风方式、舱内结构、巡航高度及回风比等因素对座舱流场环境的共同影响。CFD 模拟仿真的方法可用于设计和评估舱内环控系统,节约实验成本,缩短实验周期,并与已有数据相互验证。本文所述研究可为民用飞机舱内流场研究提供依据。

参考文献(References)

[1] LI M X, ZHAO B, TU J Y, et al. Study on the carbon dioxide lockup phenomenon in aircraft cabin by computational fluid dynamics[J]. Building Simulation, 2015, 8(4): 431-441.

[2] 崔燚. 座舱高密度群体环境中病毒传播的数值分析[J]. 科技信息, 2013(6): 46,48.
CUI Y. Numerical analysis of virus transmission in high-density crowd environment in cockpit[J]. Science & Technology Information, 2013(6): 46,48(in Chinese).

[3] 张英杰. 狭小空间内温度和二氧化碳对人体舒适性影响实验研究[D]. 重庆: 重庆大学, 2018.
ZHANG Y J. The experimental research into the influence of temperature and carbon dioxide to human comfort in small space[D]. Chongqing: Chongqing University, 2018(in Chinese).

[4] 周琼瑶. 飞机座舱非均匀环境人体热舒适研究[D]. 南京: 南京航空航天大学, 2019.
ZHOU Q Y. Research on thermal comfort for non-uniform environment in aircraft cabin[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2019(in Chinese).

[5] MAZUMDAR S, CHEN Q. A one-dimensional analytical model for airborne contaminant transport in airliner cabins[J]. Indoor Air, 2009, 19(1): 3-13.

[6] ABOOSAIDI F, WARFIELD M J, CHOUDHURY D. Numerical analysis of airflow in aircraft cabins[C]//International Conference On Environmental Systems. Warrendale: SAE International, 1991: 709-721.

[7] 代炳荣, 刘义国, 余刚, 等. 飞机座舱空气质量数值模拟研究[J]. 交通运输工程学报, 2016, 16(3): 108-115.
DAI B R, LIU Y G, YU G, et al. Numerical simulation research on air quality of aircraft cabin[J]. Journal of Traffic and Transportation Engineering, 2016, 16(3): 108-115(in Chinese).

[8] 黄彦祺. 飞机座舱内流场的数值模拟研究[D]. 重庆: 重庆大学, 2013.
HUANG Y Q. Study on numerical simulation of air-flow based on aircraft cabin[D]. Chongqing: Chongqing University, 2013(in Chinese).

[9] 陆月星, 李洪森. 飞机座舱散热的数值模拟研究[C]//第十一届全国流体力学学术会议. 北京: 中国力学学会, 2020: 957.
LU Y X, LI H M. Research on numerical simulation of aircraft cabin heat dissipation [C]//Summary of the 11th National Symposium on Fluid Mechanics, Professional Committee of Fluid Mechanics,

Chinese Society of Mechanics. Beijing: The Chinese Society of Theoretical and Applied Mechanics, 2020:957(in Chinese).

[10] 王乔蓬. 民航客机座舱内气流组织的数值模拟与结构优化研究[D]. 长沙: 长沙理工大学, 2019.
WANG Q P. Numerical simulation of air distribution and structure optimization in civil aircraft cabin[D]. Changsha: Changsha University of Science & Technology, 2019(in Chinese).

[11] 卢国栋. 大型客机座舱送风方式研究[J]. 科技创新导报, 2018, 15(9): 9-12.
LU G D. Study on air supply mode in cockpit of large passenger aircraft[J]. Science and Technology Innovation Herald, 2018, 15(9): 9-12(in Chinese).

[12] 卢国栋. 送风方式对座舱环境的影响[D]. 南京: 南京航空航天大学, 2016.
LU G D. The influence of air supply mode on the cabin environment[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2016(in Chinese).

[13] 陈希远, 于天鹏, 欧阳建亮, 等. 考虑再循环风的飞机座舱引气污染仿真研究[J]. 科学技术与工程, 2017, 17(19): 295-300.
CHEN X Y, YU T P, OUYANG J L, et al. A study of bleed air contaminant in aircraft based on recirculation air[J]. Science Technology and Engineering, 2017, 17(19): 295-300(in Chinese).

[14] 杨林谦, 魏雅川, 宋国磊等. 湍流模型的选择与评价[C]//中国航天电子技术研究院科学技术委员会2020年学术年会. 北京: 中国航天电子技术研究院, 2020: 742-747.
YANG L Q, WEI Y C, SONG G L, et al. Selection and evaluation of turbulence models [C]//Proceedings of the 2020 Academic Annual Meeting of the Science and Technology Committee of the Chinese Academy of Aerospace Electronics Technology. Beijing: China Aerospace Electronics Technology Research Institute, 2020:742-747(in Chinese).

[15] 唐慧儒, 李闯, 李晗. 民用航空器座舱高低温环境下适航符合性研究[J]. 民用飞机设计与研究, 2021(2): 124-129.
TANG H R, LI C, LI H. Research on airworthiness compatibility of civil aircraft cabin under high or low temperature environment[J]. Civil Aircraft Design & Research, 2021(2): 124-129(in Chinese).

Analysis of CO_2 distribution characteristics in cabin of civil aircraft

GAO Jingcheng[1], CHEN Weijian[1], HU Wanjun[2], FENG Shiyu[1,*], LIU Haozheng[1]

(1. School of Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;
2. Quality Management Department of Sichuan Fanhua Aviation Instrument Electric Appliance Co., Ltd., Chengdu 610599, China)

Abstract: The interior of civil aircraft cockpit is relatively closed, the air circulation is not smooth, the personnel density is large, and there are organic heat sources such as personnel and electronic equipment. It is essential to have a firm grasp of the flight envelope's CO_2 concentration, temperature distribution, and air velocity in order to certify civil aircraft as airworthy. The research introduced in this paper uses the computational fluid dynamics (CFD) method to simulate the flow field in the civil aircraft cockpit. The distribution of CO_2 concentration, temperature field, and velocity field in the cockpit under various operating conditions is determined by varying the return air and fresh air temperatures, cruise altitude, and return air ratio. Compared with the existing literature data, the correctness of the calculation model is verified. The research results show that the flow field in the cabin in the ground standard day and cruise state basically meets the requirements. When the air supply is 50 ℃ and the return air ratio is 0.5 in winter, the temperature in some areas of the cabin is higher, and the CO_2 concentration increases and approaches the standard value. The environmental control system scheme can be determined according to the actual needs.

Keywords: aircraft cabin; CO_2; computational fluid dynamics; numerical simulation; civil aircraft airworthiness

Received: 2021-11-12; **Accepted**: 2022-03-04; **Published Online**: 2022-03-23 09:30

URL: kns.cnki.net/kcms/detail/11.2625.V.20220321.2040.006.html

Foundation items: National Natural Science Foundation of China (U1833121); The Fundamental Research Funds for the Central Universities; Priority Academic Program Development of Jiangsu Higher Education Institutions; Nanjing University of Aeronautics and Astronautics Graduate Research and Practice Innovation Program (xcxjh20210108,xcxjh20220102)

* **Corresponding author**. E-mail: shiyuf@nuaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0694

基于弱监督的遥感图像镶嵌质量盲评价

潘林朋[1]，谢凤英[1,*]，赵薇薇[2]，周颖[2]，刘畅[1]，王艳[2]
(1. 北京航空航天大学 宇航学院，北京 100191；2. 北京市遥感信息研究所，北京 100192)

摘 要：遥感图像镶嵌是遥感图像解译的一项重要研究内容，然而受成像时间、角度及地物纹理的影响，镶嵌图像经常会有颜色不一致、地物结构错位等情况。针对遥感图像镶嵌缝两侧出现的颜色差异、地物结构错位等质量问题，设计了双分支网络进行遥感图像镶嵌质量的盲评价，2个分支网络分别用于镶嵌缝两侧颜色差异评价和结构错位评价，综合2个网络的输出实现遥感图像镶嵌质量的综合评价。由于获得图像的质量真值需要耗费大量的人力物力，为了减少训练卷积网络所需要的数据量，提出了一种基于两阶段训练的弱监督学习策略，第1阶段在仿真的镶嵌数据集上以颜色差异量和结构错位量为客观真值对网络进行预训练，学习与质量评价有关的先验知识，第2阶段在有主观质量真值的数据集上进行微调。在建立的带有质量真值的仿真数据集和真实数据集上的实验结果表明：所提方法能够有效评价遥感图像镶嵌的质量，性能优于对比方法。

关 键 词：遥感图像；镶嵌质量评价；弱监督学习；两阶段训练；颜色差异；结构错位
中图分类号：TP751.1
文献标志码：A **文章编号**：1001-5965(2023)09-2518-09

在遥感图像的实际应用中，为了统一分析、研究和利用遥感信息，通常需要大范围的高分辨率遥感图像。而单幅遥感图像覆盖范围有限，通常不能满足需求，为此需要将多幅遥感图像拼接成1幅大场景图像，从而遥感图像镶嵌技术应运而生并不断发展。

一般来说，待镶嵌的遥感图像需要具有一定比例的重合部分[1]，通常为30%～60%。理论上，镶嵌完美的遥感图像需要满足以下2个条件：①镶嵌痕迹要尽可能地减弱到人眼视觉系统无法察觉的程度；②镶嵌后生成图像的亮度特性和几何特性需要与镶嵌前的图像保持一致。然而，由于成像时间、成像位置、摄影镜头的畸变、卫星成像姿态的变化等因素，即使在镶嵌时进行了图像配准与融合，仍然有可能出现镶嵌缝两侧颜色不一致、地物错位等情况。对遥感图像的镶嵌质量进行评价，既可以为后续任务提供可靠的大场景遥感图像，又可以判断图像镶嵌算法的优劣，为遥感图像镶嵌算法的改进提供反馈。

目前，关于遥感图像的质量评价主要是对云雾视觉质量评价[2]、遥感图像噪声评价[3]、遥感图像模糊度评价[4]等，而关于遥感图像镶嵌质量的无参考评价的研究尚不多见。

遥感图像的镶嵌质量体现在拼接缝两侧的颜色是否一致、地物纹理是否有错位重影等方面，尽管目前对于遥感图像镶嵌质量没有专门的研究，但有一些针对自然场景图像拼接质量评价的研究可以借鉴。Yu等[5]认为拼接图像的质量优劣主要取决于重叠区域，即拼接图像中结构不一致和重影集中出现的地方，在重叠区域提取多方向梯度特征、显著图特征及相位一致性特征来完成对拼接图像的质量评价。Zhu等[6]提出了一种基于Gabor滤波器的拼接图像质量评价方法，计算图像拼接前后的

收稿日期：2021-11-18；**录用日期**：2022-03-11；**网络出版时间**：2022-04-25 19:05
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220425.1220.003.html
基金项目：国家重点研发计划(2019YFC1510905)；国家自然科学基金(61871011)
*通信作者. E-mail：xfy_73@buaa.edu.cn

引用格式：潘林朋，谢凤英，赵薇薇，等. 基于弱监督的遥感图像镶嵌质量盲评价[J]. 北京航空航天大学学报，2023，49(9)：2518-2526.
PAN L P，XIE F Y，ZHAO W W，et al. Weak supervision based blind remote sensing image mosaic quality assessment [J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2518-2526 (in Chinese).

边缘差分图像, 通过统计边缘差分图像的信息得到对拼接图像的评分。Ling 等[7] 利用卷积稀疏编码和特征选择算法获得了表示拼接图像的特征集, 并通过回归的方式对拼接图像进行评价。

2012 年, AlexNet[8] 网络在著名的 ImageNet[9] 图像分类竞赛上大放异彩, 使人们意识到了卷积神经网络在图像处理领域的重要应用前景, 一些学者开始利用卷积神经网络进行自然场景图像的质量评价。Kang 等[10] 设计了具有 1 个卷积层和 2 个全连接层的卷积网络, 采用图像块作为网络的输入来完成图像质量的评价。Bosse 等[11] 设计了具有 10 个卷积层和 2 个全连接层的网络来完成全参考的图像质量评价, 而且对网络的结构稍加调整也能完成无参考的图像质量评价。Zhu 等[12] 使用元学习的方法进行无参考的图像质量评价。Hou 等[13] 采用多任务学习框架对拼接图像进行无参考的质量评价。Lin 等[14] 引入弱监督学习的思路进行质量评价, 第 1 阶段在无质量真值的合成数据集上以结构相似性(structural similarity, SSIM)等全参考质量评价指标作为真值训练特征提取器, 第 2 阶段在带有质量真值的数据集上训练回归网络, 从而完成无参考的图像质量评价。Liu 等[15] 利用相对质量关系已知的图像对训练网络对图像质量进行排序, 再在带有质量真值的数据集上微调获得最终的质量评价模型。

卷积神经网络的性能高度依赖于可获得数据集的规模[16], 训练数据不足很容易导致网络对特定数据过拟合, 而通过主观质量评价获得图像的质量真值需要耗费大量的人力物力, 如 LIVE Challenge[17] 数据库, 仅包含接近 1 200 张图像, 就需要超过 350 000 个主观质量得分, 因此获得与 ImageNet 规模相当的图像质量评价数据集几乎是不可能完成的任务。通常, 可以通过数据增强的方法解决数据量不足的问题, 但是除了对输入图像水平翻转, 其他的数据增强方法(如旋转、裁剪、添加噪声等)都会显著地改变图像的主观质量。为了解决训练数据不足的问题, 一些学者假设图像的质量在图像空间均匀分布, 将训练图像分割成不同的图像块, 并将图像的质量真值作为图像块的质量真值, 从而大幅扩充训练集的规模[10-12]。但是以上假设对于镶嵌图像并不成立, 对于镶嵌图像, 重叠区域对于图像质量的影响程度显然要高于非重叠区域的影响。因此, 针对遥感图像镶嵌质量评价任务, 难以获得足够的有质量真值标注的数据来完成网络的训练。

针对遥感图像镶嵌过程中出现的颜色差异和地物结构错位 2 种质量失真情况, 本文设计了双分支图像质量回归网络来完成镶嵌图像的质量评价。同时, 为了解决训练数据不足的问题, 提出了一种基于两阶段训练的弱监督学习策略, 第 1 阶段在仿真数据集上, 以仿真图像的错位量和颜色差异量为质量真值对网络进行预训练, 使网络能够预先学习到与镶嵌质量有关的先验知识, 第 2 阶段在有主观质量真值的数据集上对网络进行微调。最终获得的评价模型能够有效评价遥感图像镶嵌的质量。

1　本文算法

1.1　数据集构建

为了完成遥感图像镶嵌的质量评价, 创建了 3 个数据集: ①用于网络第 1 阶段预训练的具有客观质量真值的大规模仿真数据集; ②用于第 2 阶段微调的带有主观质量真值的小规模仿真数据集; ③从真实镶嵌图像上裁剪获得的带有主观质量真值的真实失真数据集。

1.1.1　原始数据来源

本文使用的原始无失真的数据集为 NWPU-RESISC45[18], 该数据集是西北工业大学建立的用于遥感图像场景分类的公开基准, 共有 31 500 幅 256 像素×256 像素大小的图像, 包含 45 个不同的场景, 每个场景 700 张图像。45 个场景类别包括机场、桥梁、丛林、教堂、商业区、沙漠、森林、高速公路、工业区、草地、宫殿等。图 1 为一组来自该数据集中的图像。从该数据集中选出 120 幅具有代表性的场景图像用于生成带有主观质量真值的小规模仿真数据集, 剩余的 31 380 幅图像用于生成带有客观质量真值的大规模仿真数据集。

图 1　一组来自 NWPU-RESISC45[18] 数据集中的图像

Fig. 1　A set of images from NWPU-RESISC45[18] dataset

1.1.2　带客观质量真值的大规模仿真数据集生成

按照 1.1.1 节中的数据划分, 生成带客观质量真值的大规模仿真数据集。对于一幅原始图像, 仿真过程分为以下 4 个步骤:

1) 将图像分割成具有一定重叠区域的 2 幅图像。

2) 在x方向和y方向上随机产生错位量。

3) 将其中 1 幅图像转换到 HSL 色彩空间, 随机产生色调 H 和亮度 L 的差异。

4) 采用渐入渐出方法拼接 2 部分图像。

对于一幅图像, 均有 0.5 的概率发生错位、色调

差异和亮度差异，最终生成的仿真数据是原始数据集的10倍，图2为一组仿真生成的图像。

对于这部分仿真数据，将仿真时的颜色差异量和结构错位量作为其质量真值，用于后续的网络预训练。

(a) 颜色差异 (b) 结构错位 (c) 混合失真

图2 仿真生成的一组图像

Fig. 2 A set of images generated by simulation

1.1.3 带主观质量真值的小规模数据集生成

本节生成的数据集包括带主观质量真值的仿真数据集和带主观质量真值的真实数据集2部分。

将1.1.1节中选出的120幅具有代表性的场景图像作为原始图像。失真类型包括结构错位、颜色差异和混合失真，仿真过程与1.1.2节中的方法类似，区别在于在这次仿真中手动设置失真的参数量来使视觉感知质量尽可能地从1分(非常差)到5分(很好)线性变化。本文从一个小规模的数据集中选出合适的失真参数，将其应用到整个数据集上。

本文邀请了9位在遥感图像处理领域中有经验的专家使用标准的退化评级方法对图像质量进行打分，图像得分分为5个等级，即完全感知不到图像失真(5分)、可感知图像失真但不妨碍观看(4分)、能清楚地看出图像失真略微妨碍观看(3分)、妨碍观看图像(2分)、非常妨碍观看图像(1分)，将每个专家对图像的得分取均值作为图像的质量真值。最终建立的带有主观质量真值的仿真数据集包括颜色差异1 450张、结构错位700张、混合失真350张。

本文还从真实的镶嵌图像镶嵌缝附近裁剪了部分图像来建立真实镶嵌图像数据集。同样邀请专家针对图像的结构错位、颜色差异和综合评价3个方面按照上述标准对图像进行打分，从而获得真实镶嵌图像的质量真值。最终建立的带有主观质量真值的真实镶嵌图像数据集包括1 437张图像。

1.2 镶嵌质量评价网络结构设计

如图3所示，遥感图像镶嵌时，由于成像时间及天气情况的变化，即使采用融合方法进行处理，镶嵌缝两侧仍然会出现颜色差异现象，而且由于卫星成像姿态的变化，2幅图像经常存在仿射变换，从而导致拼接缝两侧的地物出现错位现象。

由于对图像的颜色差异评价和结构错位评价需要提取不同的特征，为此以ResNet-18[19]为基础设计了如图4所示的双分支网络。网络包含2个独立的分支，分别对镶嵌图像的颜色差异问题和结构错位问题进行学习。对图像的综合评价需要同时考虑到图像的颜色差异和结构错位2个方面，而且随着网络层数的加深，网络学习到的特征越抽象，语义性越强。为了充分利用2个分支网络学习到的质量感知特征，设计了图4中所示的融合策

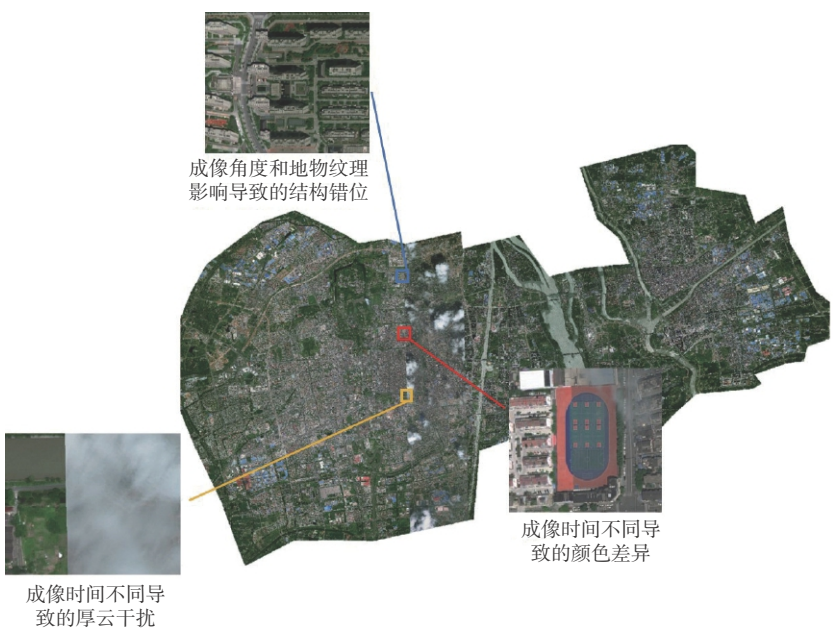


图3 实验中使用的一幅具有结构错位和颜色差异的真实遥感镶嵌图像

Fig. 3 A real remote sensing mosaic image with structural dislocation and color difference used in experiment

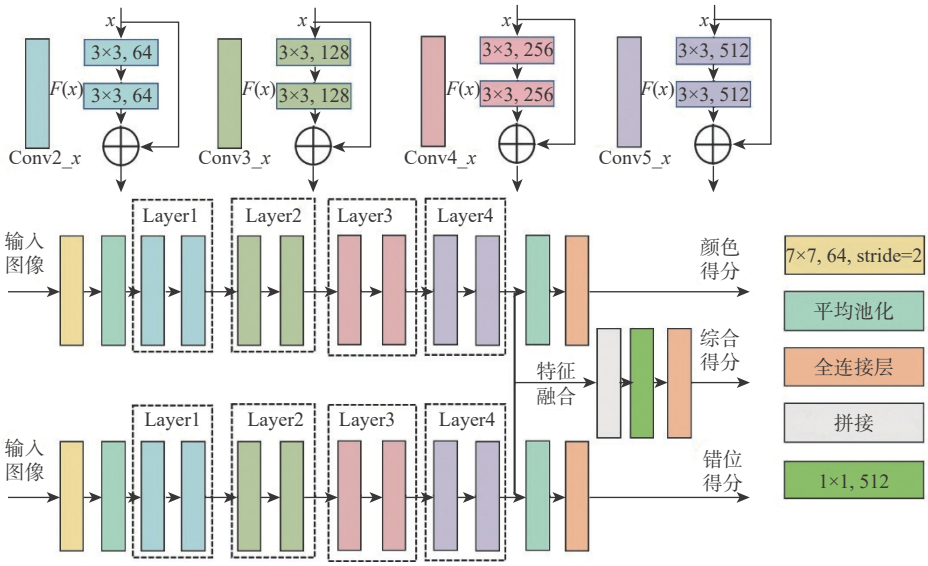


图 4　双分支卷积神经网络

Fig. 4　Two-branch convolutional neural network

略。首先, 提取 2 个分支网络最后 1 个卷积层的输出, 在通道维度上连接; 然后, 通过1×1的卷积核融合通道信息并降低通道维度; 最后, 通过全连接层回归得到图像的综合质量评分。

1.3　基于两阶段训练的弱监督学习

获取图像的主观质量真值的过程通常耗费大量的人力物力, 因此, 获得大规模的图像质量评价数据集几乎是不可能完成的任务, 而训练数据不足又会在很大程度上影响卷积神经网络的性能, 限制了深度学习算法在图像质量评价领域的应用。通常, 可以使用 ImageNet 数据集上预训练的模型作为初始参数来减少训练网络需要的数据, 但是 ImageNet 是用于自然场景图像分类任务的数据集, 在该数据集上得到的预训练模型不是专门为遥感图像镶嵌质量评价任务进行的初始训练。虽然 Lin[14] 和 Liu[15] 等也提出了其他的针对图像质量评价任务的预训练方式, 但是 Lin 等[14] 所提的方法需要计算 SSIM 等全参考评价指标作为预训练时的真值, 而镶嵌图像在发生结构错位时, 由于失真图像和参考图像之间图像内容不对应, 会导致无法计算全参考评价指标。而 Liu 等[15] 所提的方法进行第 1 阶段训练时, 要求每次输入同一张参考图像生成的同一类型失真的图像对, 将这个图像对之间的质量等级对比作为真值进行预训练, 这种方法只能针对单一失真类型进行评价, 无法用于综合失真的情况。因此, 这 2 种预训练方式都不适合镶嵌图像质量评价任务。本文提出了如图 5 所示的基于两阶段的弱监督学习策略, 第 1 阶段在带有客观质量真值的仿真数据集上对网络进行预训练, 第 2 阶段在带有主观质量真值的数据集上对网络进行微调, 来获得最终的评

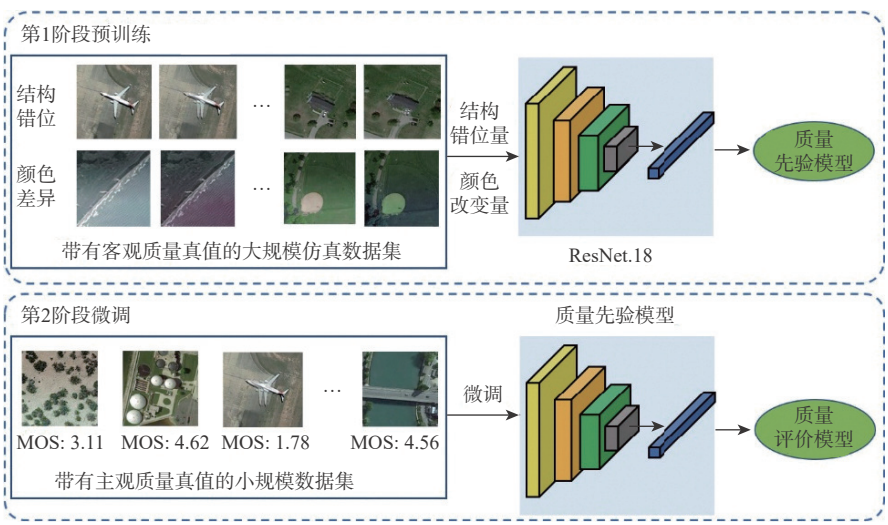


图 5　两阶段训练策略

Fig. 5　Two-stage training strategy

价模型。

图像失真过程影响图像的质量，如果网络训练过程中能够学习到失真过程的相关信息，将会有助于图像质量的预测。由此，设计了第1阶段的预训练，即在大规模仿真数据集上以图像的错位量和颜色差异量为训练真值分别对错位分支和颜色分支进行训练。由于这一阶段的训练不需要图像的主观质量真值，本文仿真生成大量不同场景、不同失真程度的图像，为卷积网络的训练提供足够的训练数据。通过这一阶段的训练，网络不仅能够识别到与遥感图像镶嵌质量评价相关的2个重要因素：结构错位程度和颜色差异程度，还能学习到不同场景的知识从而避免对特定场景过拟合。

第1阶段训练完成后，为了将学习到的先验知识映射到图像的质量得分，需要在带有主观质量真值的小规模数据集上进行微调。受益于第1阶段的预训练，这一阶段不需要大量的数据，可以在规模很小的数据集上完成训练。对仿真的镶嵌图像进行镶嵌质量评价时，只需要用带主观真值的仿真数据集对网络进行微调。同样，对真实的镶嵌图像进行质量评价时，只需要用带主观真值的真实数据集进行微调。本文设计的两阶段训练方案，第1阶段训练使用的仿真数据集并没有主观质量真值，本质上属于无标签样本，该阶段的训练为第2阶段使用少量带有主观质量真值标签的样本进行微调时提供了更合理的初始参数。也就是说，本文方法首先在大量无标签样本上进行第1阶段的训练，然后在少量具有真实标签的样本上进行微调，因此该方法属于弱监督学习策略中的半监督学习。

2个阶段训练时使用的损失函数为平均平方误差（mean square error, MSE）。假设$\boldsymbol{Y}=(y_1,\cdots,y_N)$和$\hat{\boldsymbol{Y}}=(\hat{y}_1,\cdots,\hat{y}_N)$分别为真实标签和预测标签，$N$为样本个数，则平均平方误差定义为

$$L_{\mathrm{MSE}}(\boldsymbol{Y},\hat{\boldsymbol{Y}})=\frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{y}_i)^2 \tag{1}$$

在第1阶段，结构错位分支的训练真值为x和y 2个方向上的错位量，颜色差异分支的训练真值为色调H差异和亮度I差异。第2阶段微调的质量真值为图像的平均主观质量得分（mean opinion score, MOS）。

2 实验验证

2.1 评估指标

评估质量评价方法性能的2个常用指标是皮尔逊线性相关系数（Pearson linear correlation coefficient, PLCC）和斯皮尔曼秩相关系数（Spearman rank order correlation coefficient, SROCC），PLCC衡量算法预测值和图像质量真值之间的线性程度，对于N个失真图像，第i个图像的质量真值记为y_i，算法的预测值记为$\hat{y}_i$，则PLCC的计算公式如下：

$$P_{\mathrm{PLCC}}=\frac{\sum_{i=1}^{N}(y_i-\bar{y})(\hat{y}_i-\bar{\hat{y}})}{\sqrt{\sum_{i=1}^{N}(y_i-\bar{y})^2}\sqrt{\sum_{i=1}^{N}(\hat{y}_i-\bar{\hat{y}})^2}} \tag{2}$$

式中：$\bar{y}$和$\bar{\hat{y}}$分别为图像质量真值和预测值的平均值。

S_{SROCC}的计算公式如下：

$$S_{\mathrm{SROCC}}=1-\frac{6\sum_{i=1}^{N}(v_i-p_i)^2}{N(N^2-1)} \tag{3}$$

式中：v_i和p_i分别为y_i和$\hat{y}_i$在图像质量真值和预测值中的排序位置。S_{SROCC}能够衡量预测值和真实值之间的单调性。

2.2 消融实验

为了证明两阶段训练策略中第1阶段训练的必要性，在仿真数据集上进行了相应的消融实验。首先，在带有客观质量真值的大规模仿真数据集上完成第1阶段的预训练；然后，对第1阶段预训练得到的先验模型和ImageNet上预训练的模型进行比较。2种模型都在带有主观质量真值的小规模仿真数据集上进行微调。微调时，按照原始图像对数据集进行随机划分，以80%的数据作为训练集，10%的数据作为验证集，10%的数据作为测试集，使用Adam[20]优化器，mini-batch的大小为64，学习率为0.001，weight_decay为1×10^{-4}，训练的epoch为40，在验证集上进行模型参数选择，在测试集上进行测试，数据集划分及微调过程重复10次，并取均值作为最终的结果，实验结果如表1所示。可以看出，对于仿真图像，两阶段训练策略能够显著提高模型的评价性能，尤其是结构错位类型的失真和混合失真图像。

为了验证本文方法对真实镶嵌图像质量评价的有效性，在带有主观质量真值的真实数据集上对第1阶段预训练得到的先验模型和ImageNet上预训练的模型进行了比较。2种模型同样在带有主观质量真值的真实失真数据集上进行微调。微调时，将1.1.3节中的真实镶嵌图像按照8∶1∶1的比例随机划分为训练集、验证集和测试集，采用随机水平翻转对训练数据进行增强，其余参数设置与仿真数据集上的实验设置相同，数据集划分及微调过程重复10次，并取均值作为最终的结果，实验结果如表2所示。可以看出，在大规模仿真数据集上进行预训练的模型效果要优于ImageNet上预训练的结果，表

表 1　在带有主观质量真值的仿真数据集上的评估结果

Table 1　Evaluation results on simulation dataset with subjective quality score

	S_{SROCC}			P_{PLCC}		
	颜色差异	结构错位	混合失真	颜色差异	结构错位	混合失真
两阶段训练策略	0.814 3	0.763 5	0.862 4	0.823 6	0.826 1	0.903 8
ImageNet	0.765 6	0.332 6	0.760 4	0.752 4	0.354 2	0.764 3

表 2　在带有主观质量真值的真实数据集上的评估结果

Table 2　Evaluation results on authentic dataset with subjective quality score

	S_{SROCC}			P_{PLCC}		
	颜色差异	结构错位	混合失真	颜色差异	结构错位	混合失真
两阶段训练策略	0.603 6	0.601 8	0.613 2	0.886 2	0.890 1	0.905 6
ImageNet	0.574 6	0.544 1	0.574 5	0.876 3	0.887 2	0.870 1

明在大规模仿真数据集上进行预训练能够提高模型对真实镶嵌图像的评价性能。

2.3　与其他评价方法的对比实验

为验证本文方法的有效性, 与其他无参考质量评价方法进行对比, 包括针对自然场景图像的 3 种无参考质量评价方法(NIQE[21]、BRISQUE[22]、SSEQ[23])和 1 种基于深度学习的针对拼接图像的无参考质量评价方法(SPIQA[13])。其中, NIQE 方法无需训练, 而对于 BRISQUE、SSEQ 和 SPIQA 方法, 为了公平对比, 用本文的数据集对其重新训练。

2.3.1　在仿真数据集上的对比

用带主观真值的小规模仿真数据集对方法进行测试, 即随机选择其中的 80% 用作训练图像进行网络微调, 10% 作为验证集进行参数选择, 剩余的 10% 作为测试集进行测试, 上述过程重复 10 次, 并取均值作为最终的结果, 实验结果如表 3 所示。可以看出, NIQE 和 BRISQUE 方法针对自然场景图像的通用类型失真进行质量评价, 不适于遥感图像镶嵌质量的评价, 因此, 整体评价准确率较低。SSEQ 方法由于提取了更多类型的特征, 在某些方面的评价准确率甚至超过了 SPIQA 方法, 但是对颜色差异失真的评价准确率比较低。SPIQA 方法专门针对图像拼接质量进行评价, 在对比方法中取得了较好的评价结果。与这些对比方法相比, 本文方法专门针对遥感图像镶嵌质量评价而设计, 在 3 种不同失真类型的数据集上均取得了最好的评价结果。

表 3　不同方法在仿真数据集上的性能对比

Table 3　Performance comparison of different algorithms on simulation dataset

	S_{SROCC}			P_{PLCC}		
	颜色差异	结构错位	混合失真	颜色差异	结构错位	混合失真
NIQE	0.006 5	−0.054 2	0.196 5	0.356 4	0.228 4	0.435 6
BRISQUE	0.024 5	0.170 5	0.203 9	0.223 1	0.275 9	0.304 3
SSEQ	0.034 1	0.695 6	0.745 8	0.214 2	0.805 4	0.844 3
SPIQA	0.674 2	0.375 9	0.680 1	0.687 2	0.386 5	0.725 6
本文	0.814 3	0.763 5	0.862 4	0.823 6	0.826 1	0.903 8

在对比方法中, SPIQA 和 SSEQ 方法取得了比其他对比方法更好的结果。为了直观对比不同的方法, 绘制了 SPIQA、SSEQ 及本文方法在测试数据集上的散点图。图 6 给出了 3 种方法在混合失真类型的仿真数据集上的散点图和线性回归结果。可以看出, 与 SPIQA 和 SSEQ 两种方法相比, 本文方法的评价结果与主观质量评价得分相关性更好。

2.3.2　在真实数据集上的对比

对于带主观真值的真实数据集, 同样将数据集按照 8: 1: 1 的比例划分成训练集、测试集和验证集, 并重复 10 次实验取均值作为最终的结果, 实验结果如表 4 所示。可以看出, 在所有的方法中, 本文方法取得了最好的结果。SPIQA、SSEQ 及本文方法在混合失真的真实失真数据集上的散点图和线性回归结果如图 7 所示。可以看出, 本文方法的评价结果与主观质量评价得分的相关性最好, 进一步表明了本文方法对真实失真的镶嵌图像能够做出有效评价。此外, 从表 3 和表 4中也可以看出, 本文方法相比于颜色差异和结构错位在混合失真评价指标上有更好的效果, 这主要是因为对混合失真样本的评价利用的是颜色差异分支和结构错位分支提取的高层语义信息, 并且对 2 部分信息进行了

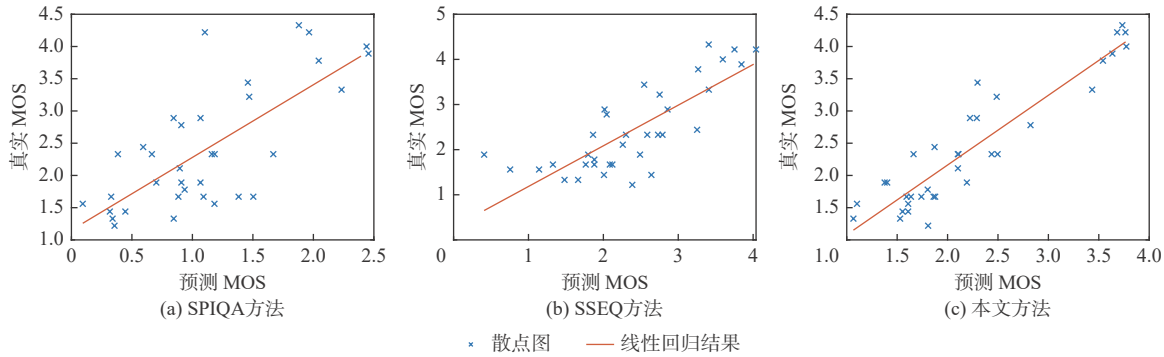


图 6 不同方法在仿真数据集上的散点图和线性回归结果

Fig. 6 Scatter plot and linear regression results of different algorithms on simulation dataset

表 4 不同方法在真实数据集上的性能对比

Table 4 Performance comparison of different algorithms on authentic dataset

	S_{SROCC}			P_{PLCC}		
	颜色差异	结构错位	混合失真	颜色差异	结构错位	混合失真
NIQE	−0.003 2	0.022 5	0.002 3	0.112 5	0.083 4	0.082 1
BRISQUE	0.323 1	0.336 9	0.325 0	0.796 4	0.785 6	0.798 5
SSEQ	0.348 9	0.354 2	0.324 1	0.774 5	0.764 4	0.778 9
SPIQA	0.435 6	0.375 9	0.436 4	0.846 9	0.822 1	0.852 3
本文	0.603 6	0.601 8	0.613 2	0.886 2	0.890 1	0.905 6

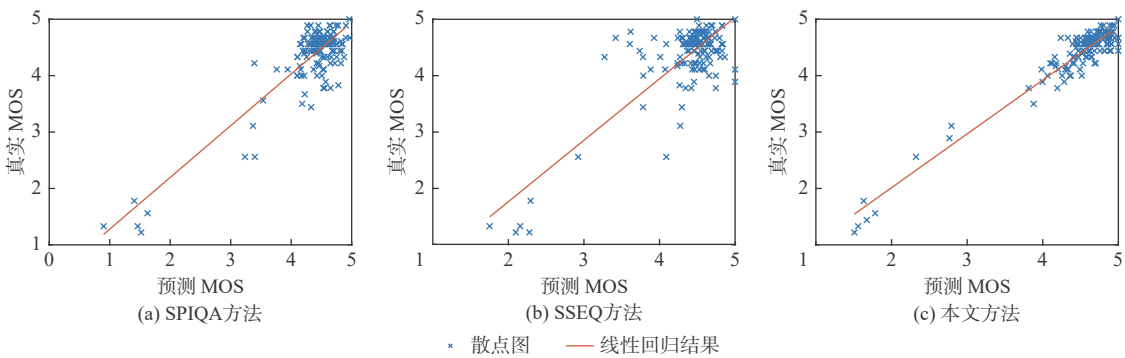


图 7 不同方法在真实数据集上的散点图和线性回归结果

Fig. 7 Scatter plot and linear regression results of different algorithms on authentic dataset

融合，训练时只需要训练特征融合模块和最后的全连接层。而对颜色差异和结构错位的评价利用的则是原始图像，训练时需要对整个网络进行训练。相较来说，混合失真评价学习难度更低，其获得的指标也会更好一些。

3 结 论

1）本文方法通过主观质量评价实验，建立了遥感图像镶嵌质量评价数据集，为后续的实验提供了数据支持。

2）设计的双分支卷积神经网络能够有效评价遥感图像镶嵌的质量，在仿真数据集和真实数据集上性能都优于对比方法。

3）提出的基于两阶段的弱监督学习策略能够有效提升网络的评价性能。

参考文献（References）

[1] 周清华, 潘俊, 李德仁. 遥感图像镶嵌接缝线自动生成方法综述[J]. 国土资源遥感, 2013, 25(2): 1-7.
ZHOU Q H, PAN J, LI D R. Overview of automatic generation of mosaicking seamlines for remote sensing images[J]. Remote Sensing for Land & Resources, 2013, 25(2): 1-7(in Chinese).

[2] 秦曼君. 多光谱遥感图像去雾算法研究[D]. 北京：北京航空航天大学, 2018: 7-13.
QIN M J. Research on the fog removal algorithm of multispectral remote sensing image[D]. Beijing: Beihang University, 2018: 7-13(in Chinese).

[3] 徐冬宇. 高光谱遥感图像质量评价算法研究[D]. 杭州: 浙江大学, 2019: 5-10.
XU D Y. Research on quality evaluation algorithm of hyperspectral remote sensing image[D]. Hangzhou: Zhejiang University, 2019: 5-10(in Chinese).

[4] 徐冬宇, 厉小润, 赵辽英, 等. 基于多模型融合的高光谱图像质量

评价[J]. 激光与光电子学进展, 2019, 56(2): 92-101.
XU D Y, LI X R, ZHAO L Y, et al. Hyperspectral image quality evaluation based on multi-model fusion[J]. Laser & Optoelectronics Progress, 2019, 56(2): 92-101(in Chinese).

[5] YU S J, LI T S, XU X Y, et al. NRQQA: A no-reference quantitative quality assessment method for stitched images[C]//Proceedings of the ACM Multimedia Asia. New York: ACM, 2019: 1-6.

[6] ZHU S D, ZHANG Y Z, TAO L, et al. A novel method for quality assessment of image stitching based on the Gabor filtering[C]//2018 IEEE International Conference on Information and Automation (ICIA). Piscataway: IEEE Press, 2019: 1605-1610.

[7] LING S Y, CHEUNG G, LE CALLET P. No-reference quality assessment for stitched panoramic images using convolutional sparse coding and compound feature selection[C]//2018 IEEE International Conference on Multimedia and Expo (ICME). Piscataway: IEEE Press, 2018: 1-6.

[8] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.

[9] DENG J, DONG W, SOCHER R, et al. ImageNet: A large-scale hierarchical image database[C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2009: 248-255.

[10] KANG L, YE P, LI Y, et al. Convolutional neural networks for no-reference image quality assessment[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2014: 1733-1740.

[11] BOSSE S, MANIRY D, MÜLLER K R, et al. Deep neural networks for no-reference and full-reference image quality assessment[J]. IEEE Transactions on Image Processing, 2018, 27(1): 206-219.

[12] ZHU H C, LI L D, WU J J, et al. MetaIQA: Deep meta-learning for no-reference image quality assessment[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 14131-14140.

[13] HOU J W, LIN W S, ZHAO B Q. Content-dependency reduction with multi-task learning in blind stitched panoramic image quality assessment[C]//2020 IEEE International Conference on Image Processing (ICIP). Piscataway: IEEE Press, 2020: 3463-3467.

[14] LIN H H, HOSU V, SAUPE D. DeepFL-IQA: Weak supervision for deep IQA feature learning[EB/OL]. (2021-01-20)[2021-11-01]. https//arxiv.org/abs/2001.08113.

[15] LIU X L, VAN DE WEIJER J, BAGDANOV A D. RankIQA: Learning from rankings for No-reference image quality assessment[C]//2017 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE Press, 2017: 1040-1049.

[16] KIM J, ZHENG H, GHADIYARAM D, et al. Deep convolutional neural models for picture-quality prediction: Challenges and solutions to data-driven image quality assessment[J]. IEEE Signal Processing Magazine, 2017, 34(6): 130-141.

[17] GHADIYARAM D, BOVIK A C. Massive online crowdsourced study of subjective and objective picture quality[J]. IEEE Transactions on Image Processing, 2016, 25(1): 372-387.

[18] CHENG G, HAN J W, LU X Q. Remote sensing image scene classification: Benchmark and state of the art[J]. Proceedings of the IEEE, 2017, 105(10): 1865-1883.

[19] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 770-778.

[20] KINGMA D P, BA J. Adam: A method for stochastic optimization[EB/OL]. (2017-01-30)[2021-11-01]. https://arxiv.org/abs/1412.6980.

[21] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a "completely blind" image quality analyzer[J]. IEEE Signal Processing Letters, 2013, 20(3): 209-212.

[22] MITTAL A, MOORTHY A K, BOVIK A C. No-reference image quality assessment in the spatial domain[J]. IEEE Transactions on Image Processing, 2012, 21(12): 4695-4708.

[23] LIU L X, LIU B, HUANG H, et al. No-reference image quality assessment based on spatial and spectral entropies[J]. Signal Processing:Image Communication, 2014, 29(8): 856-863.

Weak supervision based blind remote sensing image mosaic quality assessment

PAN Linpeng[1], XIE Fengying[1,*], ZHAO Weiwei[2], ZHOU Ying[2], LIU Chang[1], WANG Yan[2]

(1. School of Astronautics, Beihang University, Beijing 100191, China; 2. Beijing Institute of Remote Sensing, Beijing 100192, China)

Abstract: A remote sensing image mosaic is an important research content of remote sensing image interpretation. However, affected by the imaging time, angle, and object texture, mosaic images often suffer inconsistent colors and structure dislocation. Aiming at the above quality problems, a double-branch network is designed to perform the blind assessment of the remote sensing image mosaic quality. The branch network are used to assess the color difference and structural dislocation respectively. Finally, the output of the branch networks is integrated to realize the comprehensive assessment. A weakly supervised learning technique based on two-stage training is proposed to lower the quantity of images in network training because it requires a lot of labor and material resources to determine the true score of the image quality. Firstly, to gain the previous knowledge associated with quality assessment, the network is initially pre-trained on the simulated mosaic dataset, which uses color change and structural dislocation as the objective quality score. Secondly, fine-tuning is performed on the dataset with subjective scores. Secondly, fine-tuning is performed on the dataset with subjective score. The experiment results on the established simulation dataset and authentic dataset show that the proposed method can effectively assess the quality of remote sensing image mosaic and outperforms the comparison algorithms.

Keywords: remote sensing images; mosaic quality assessment; weakly supervised learning; two stage training; color difference; structural dislocation

Received: 2021-11-18; **Accepted**: 2022-03-11; **Published Online**: 2022-04-25 19:05

URL: kns.cnki.net/kcms/detail/11.2625.V.20220425.1220.003.html

Foundation items: National Key R&D Program of China (2019YFC1510905); National Natural Science Foundation of China (61871011)

* **Corresponding author**. E-mail: xfy_73@buaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2021.0700

新型变摩擦滑橇式飞行器滑跑纠偏

王宇晟[1,2]，印寅[1,2]，梁涛涛[1,2]，魏小辉[1,2,*]

(1. 南京航空航天大学 机械结构力学及控制国家重点实验室，南京 210016；
2. 南京航空航天大学 飞行器先进设计技术国防重点学科实验室，南京 210016)

摘　　要：相较于轮式起落架，结构轻巧紧凑的滑橇式起落架更适用于扁平化的高超声速飞行器，然而前轮后橇布局所固有的地面力学特性使得飞行器在中速滑跑阶段存在严重的航向失稳问题。针对滑橇式飞行器滑跑航向稳定性较差的问题，提出了一种具备航向增稳功能的新型变摩擦滑橇式起落架。建立了滑橇式飞行器的非线性地面滑跑动力学模型，考虑了气动载荷、地面载荷和纠偏机构模型；基于摩擦特性试验结果得到了滑橇及摩擦材料摩擦系数的多参数拟合公式，并将其代入所提滑跑模型以提高精度；引入积分视线（ILOS）法建立了滑橇式飞行器滑跑纠偏控制系统，并采用粒子群算法优化控制参数。试验结果表明：滑橇及摩擦材料的摩擦系数均随速度和压强的增加先增大后减小；典型初始工况下的滑跑仿真结果验证了基于变摩擦滑橇纠偏控制的有效性；当摩擦材料摩擦系数从0.3增至0.4时，飞行器地面滑跑安全边界提升16.6%。

关 键 词：滑橇式起落架；摩擦系数；地面滑跑；纠偏控制；积分视线引导律
中图分类号：V226
文献标志码：A　　**文章编号**：1001-5965(2023)09-2527-12

高超声速技术是航空航天史上的重要里程碑之一，相比于常规飞行器，高超声速飞行器有速度快、全球到达、反侦察和突防能力强等优点[1-3]，但由于其扁平化的气动布局[4]、严苛的质量要求和极限的着陆工况，传统的轮式起落架存在机轮尺寸大、收纳空间不足[5]、系统质量大和滑跑制动时受热严重[6]等问题。相比于轮式起落架，滑橇式起落架因其机构紧凑、质量轻、占据空间小和耐高温等优点，更适用于高超声速飞行器。

然而，传统滑橇式飞行器的地面力学特性导致其在地面滑跑阶段航向稳定性较差。飞行器前起和主起的侧向力变化所产生的偏航力矩决定了飞行器地面滑跑的方向稳定性。对于采用前轮后橇布局的飞行器，由于前轮侧向力的变化速率显著快于主起落架滑橇，滑跑过程中更易发生偏航，严重时可能会出现甩尾甚至冲出跑道。国内外学者对滑橇式起落架的地面滑跑性能鲜有研究，采用滑橇式起落架的固定翼飞行器也较少。美国X-15高超声速飞行器采用了船型钢质滑橇型主起落架[7]，其着陆场地选定为广阔的干涸湖床以避免冲出跑道；美国“追梦者”航天飞机采用舱门起落架一体的滑橇式前起落架，成功完成了无动力滑翔着陆和滑跑制动的试验[8]。

目前，关于轮式飞行器滑跑纠偏控制已存在广泛且深入的研究，对滑橇式飞行器同样具备指导意义。轮式飞行器通常采用前轮转向、主轮差动刹车和方向舵偏转进行滑跑纠偏。Lemay等[9]提出了一种增益调度方法，通过前轮实现对飞行器低速阶段滑跑和机动的控制。王勇和王英勋[10]建立了包含差动刹车与方向舵的联合滑跑纠偏系统。尹乔之

收稿日期：2021-11-19；**录用日期**：2022-02-16；**网络出版时间**：2022-03-22 16:57
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20220321.2030.005.html
基金项目：国家自然科学基金(52172368)；机械结构力学及控制国家重点实验室自主研究课题(MCMS-I-0221Y02)；中央高校基本科研业务费专项资金(NF2018001)
* **通信作者**. E-mail：wei_xiaohui@nuaa.edu.cn

引用格式：*王宇晟，印寅，梁涛涛，等. 新型变摩擦滑橇式飞行器滑跑纠偏[J]. 北京航空航天大学学报，2023，49(9)：2527-2538.*
WANG Y C，YIN Y，LIANG T T，et al. Taxiing deviation-correction control of a new variable-friction equipped-skid aircraft[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2527-2538 (in Chinese).

和聂宏等[11]设计了一种联合纠偏控制律，以充分发挥差动刹车与前轮转向的控制能力。李继广等[12]基于反步控制为菱形翼布局的飞行器设计了一种起飞滑跑控制器，对航向和滚转进行了有效控制。李旭波等[13]采用视线法制导及预测控制的联合应用使飞行器在滑跑阶段有效跟踪跑道中心线。在中速阶段，由于前轮转向与方向舵的控制效率较低，差动刹车至关重要，针对在此阶段缺乏有效控制手段的滑橇飞行器，本文提出了一种具备航向增稳功能的变摩擦滑橇式起落架。

变摩擦滑橇的核心在于主动控制滑橇装置的等效摩擦系数，摩擦系数受到材料副配对性质、法向载荷大小、滑动速度、温度状况和摩擦表面接触几何特性等因素的影响[14]。常用的分析方法有计算法、有限元法和试验法等。魏兵辉等[15]通过有限元技术对不同工况下轨道-滑靴的摩擦系数进行计算，结合实验进行验证和分析。魏龙等[16]通过理论模拟计算和试验，研究了工作参数和端面形貌分形参数对接触式机械密封端面摩擦系数的影响。Wu 等[17]通过对湿式离合器摩擦系数的实验数据挖掘和理论分析，得到了不同影响因素对摩擦系数的影响和敏感性。陶润和侯之超[18]通过试验测得滚动轴承在不同径向力和转速下的摩擦力矩，并基于最小二乘法拟合得到了经验公式。试验法仍然是分析摩擦特性最直接、最准确的方法，采用计算法和有限元法分析材料摩擦特性时，也需要通过试验进行验证，因此本文开展摩擦特性试验。

本文针对传统滑橇式起落架的缺陷，提出了一种具备航向增稳功能的新型变摩擦滑橇式起落架。首先，考虑运动过程中多种非线性因素的影响，建立了滑橇式飞行器的地面滑跑动力学模型；然后，开展滑橇装置地面摩擦特性研究，总结滑橇装置与混凝土道面摩擦系数的影响因素及变化规律；其次，引入积分视线（integral line of sighe, ILOS）法，结合 PID 控制器，实现对跑道中心线的跟踪控制；最后，典型工况下的仿真结果验证了基于变摩擦滑橇机构的纠偏控制有效性。

1 增稳型滑橇式起落架

根据某型飞行器缩比试验样机设计的滑橇式主起落架如图 1 所示。起落架由外筒、活塞杆、缓冲器、扭力臂、滑橇、纠偏机构及纠偏驱动装置组成，其中，纠偏机构由连杆、摇臂、安装板和摩擦材料组成，纠偏驱动装置由力传感器、电机、减速器、联轴器和顶针组成。

其工作原理为：飞行器着陆，滑橇接触道面，飞行器滑跑制动，若飞行器滑跑方向正确，则摩擦材料保持在上位，与道面分离；当飞行器滑跑方向发生偏差时，纠偏驱动装置通过摇臂和连杆将摩擦材料下压接触道面产生压力，由于摩擦材料摩擦系数大于滑橇，能够增大单侧起落架与道面的摩擦力，使飞行器受到偏航力矩，进而对飞行器的滑跑方向进行纠正。

当滑橇式起落架收入机身时，滑橇和摩擦板能够向前折叠，节省收放空间。起落架收上状态的侧视图如图 2 所示。相比于轮式起落架，滑橇式起落架能节省约 57% 的垂向收放空间，更适用于采用扁平化气动布局的高超声速飞行器。

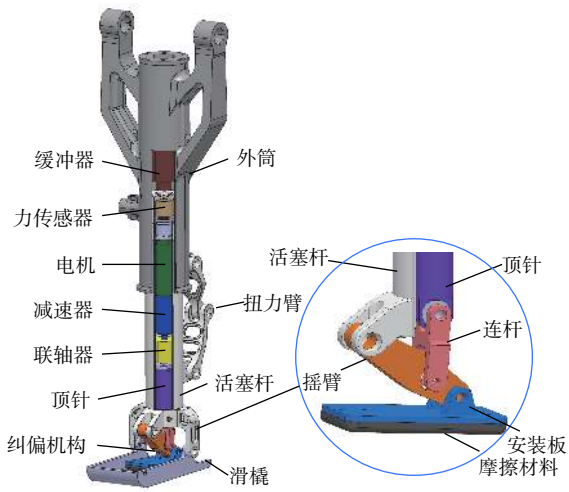


图 1 新型滑橇式起落架示意图

Fig. 1 Schematic diagram of a new type of skid landing gear

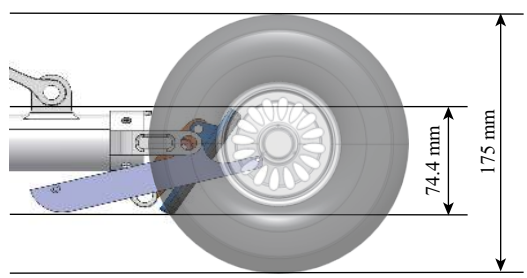


图 2 收放空间对比

Fig. 2 Comparison of space occupation

2 滑跑动力学模型

由于飞行器地面滑跑过程真实情况十分复杂，为了便于仿真计算同时较好地模拟实际情况，做出如下合理假设[19]：

1）将飞行器看作刚体，不考虑各部件之间的相对运动和相互作用力。

2）机体坐标系的$O_b x_b y_b$为飞行器的对称平面，飞行器外形及质量均对称。

3）由于跑道平整且缓冲器刚度较大，飞行器地

面滑跑被视为平面运动, 其垂向、滚转及俯仰的轻微姿态变化被忽略不计。

在跑道上滑跑时,飞行器所受的力除空气动力外,还有地面的作用力[20]。地面作用力包括地面支持力、前轮滚动摩擦力、前轮的侧向力和滑橇的滑动摩擦力。飞行器地面滑跑过程的受力分析如图 3 所示。图中: $\boldsymbol{P}_{\mathrm{n}}$、$\boldsymbol{P}_{\mathrm{ml}}$和$\boldsymbol{P}_{\mathrm{mr}}$分别为前轮、左滑橇和右滑橇受到的地面支持力, 向上为正; $\boldsymbol{Q}_{\mathrm{n}}$、$\boldsymbol{F}_{\mathrm{n}}$分别为前轮受到的滚动摩擦力和侧向力; $\boldsymbol{Q}_{\mathrm{ml}}$、$\boldsymbol{Q}_{\mathrm{mr}}$、$\boldsymbol{F}_{\mathrm{ml}}$、$\boldsymbol{F}_{\mathrm{mr}}$分别为左滑橇和右滑橇受到的滑动摩擦力的分量; $\boldsymbol{Q}_{\mathrm{n}}$、$\boldsymbol{Q}_{\mathrm{ml}}$、$\boldsymbol{Q}_{\mathrm{mr}}$在跑道平面内沿机体轴$x_{\mathrm{b}}$, 向后为正; $\boldsymbol{F}_{\mathrm{n}}$、$\boldsymbol{F}_{\mathrm{ml}}$、$\boldsymbol{F}_{\mathrm{mr}}$在跑道平面内沿机体轴$y_{\mathrm{b}}$, 向右为正; O_{b}为机体坐标系原点; a_{n}和a_{m}分别为前起和主起到飞行器重心的纵向距离; b_{w}为 2 个主滑橇之间的距离; ψ为飞行器相对跑道中心线的偏航角, 右偏为正; $\boldsymbol{V}$为飞行器滑跑速度。

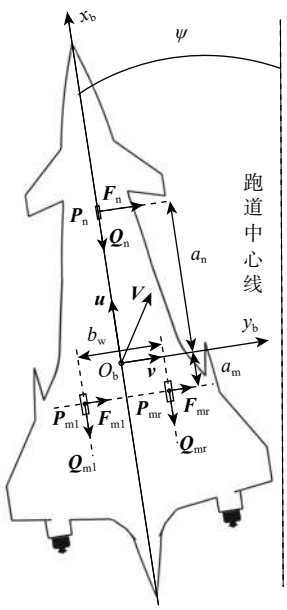


图 3　飞行器受到的地面力

Fig. 3　Ground forces on aircraft

飞行器地面滑跑的动力学方程如下:

$$\begin{cases} \dot{\boldsymbol{u}} = (-\boldsymbol{D}\cos\beta + \boldsymbol{C}\sin\beta)/m + \boldsymbol{r}\boldsymbol{v} + \boldsymbol{F}_x/m \\ \dot{\boldsymbol{v}} = (\boldsymbol{D}\sin\beta + \boldsymbol{C}\cos\beta)/m - \boldsymbol{r}\boldsymbol{u} + \boldsymbol{F}_y/m \\ \dot{\boldsymbol{r}} = (\boldsymbol{N} + \boldsymbol{M}_z)/I_z \end{cases} \tag{1}$$

$$\begin{cases} \boldsymbol{F}_x = -(\boldsymbol{Q}_{\mathrm{n}} + \boldsymbol{Q}_{\mathrm{ml}} + \boldsymbol{Q}_{\mathrm{mr}}) \\ \boldsymbol{F}_y = \boldsymbol{F}_{\mathrm{n}} + \boldsymbol{F}_{\mathrm{ml}} + \boldsymbol{F}_{\mathrm{mr}} \\ \boldsymbol{M}_z = \boldsymbol{F}_{\mathrm{n}}a_{\mathrm{n}} - (\boldsymbol{F}_{\mathrm{ml}} + \boldsymbol{F}_{\mathrm{mr}})a_{\mathrm{m}} + (\boldsymbol{Q}_{\mathrm{mr}} - \boldsymbol{Q}_{\mathrm{ml}})b_{\mathrm{w}}/2 \end{cases} \tag{2}$$

运动学方程如下:

$$\begin{cases} \dot{\psi} = r \\ \dot{x}_{\mathrm{d}} = u\cos\psi - v\sin\psi \\ \dot{y}_{\mathrm{d}} = u\sin\psi + v\sin\psi \end{cases} \tag{3}$$

式中: $\boldsymbol{D}$和$\boldsymbol{C}$分别为气动阻力和侧力; m为飞行器质量; $\boldsymbol{F}_x$、$\boldsymbol{F}_y$分别为地面力在机体轴x_{b}、y_{b}的分量; $\boldsymbol{N}$为气动偏航力矩; $\boldsymbol{M}_z$为各地面作用力在机体轴z_{b}的合力矩; β为气动侧滑角; $\boldsymbol{r}$为偏航角速率; $\boldsymbol{u}$、$\boldsymbol{v}$分别为飞行器速度在机体轴x_{b}、y_{b}下的分量; x_{d}、y_{d}分别为飞行器重心在地面坐标轴x_{g}、y_{g}的位置分量; I_z为飞行器相对于机体轴z_{b}的转动惯量。

3　滑跑纠偏控制系统

3.1　差动纠偏控制系统设计

通过有效控制左右两侧滑橇装置的等效摩擦系数, 飞行器能够实现对滑跑方向的纠正。为提高纠偏控制的性能, 引入 ILOS 法。LOS 法[21] 的控制原理如图 4 所示。

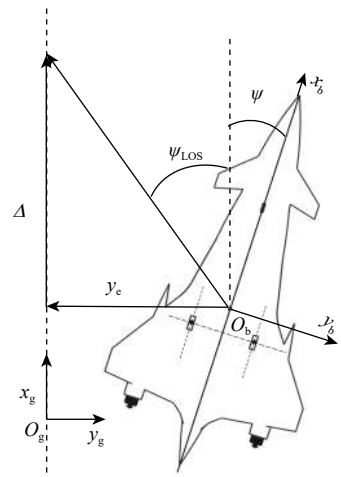


图 4　LOS 控制原理

Fig. 4　Line-of-sight control principle

图 4 中, y_{e}为飞行器距跑道中心线的横向偏差, 即侧偏距离; O_{g}为地面坐标系原点; Δ为可视距离; 视线角(LOS 角)ψ_{LOS}可通过式(4)计算得到:

$$\psi_{\mathrm{LOS}} = \arctan\left(\frac{-y_{\mathrm{e}}}{\Delta}\right) \tag{4}$$

若飞行器的航向能保持对准视线角, 即飞行器偏航角ψ与视线角ψ_{LOS}保持一致, 飞行器将逐渐沿跑道中心线滑跑[22]。当飞行器在滑跑过程中受到侧风等非对称载荷时, 传统 LOS 法无法消除纠偏过程中的横向偏差, 因此引入带有横向偏差积分项的 ILOS 法[23], 改进公式为

$$\psi_{\mathrm{ILOS}} = -\arctan\left(\frac{y_{\mathrm{e}} + y_{\mathrm{eint}}}{\Delta}\right) \tag{5}$$

$$\dot{y}_{\mathrm{eint}} = \frac{k y_{\mathrm{e}}}{\sqrt{\Delta^2 + (y_{\mathrm{e}} + y_{\mathrm{eint}})^2}} \tag{6}$$

式中: ψ_{ILOS}为 ILOS 的视线角; y_{eint}为横向偏差的积分项。

参数k满足如下条件:

$$k = \begin{cases} k_1|\boldsymbol{V}|, & y_{\mathrm{e}}y_{\mathrm{eint}} \leqslant 0 \\ k_2 \mathrm{e}^{-\rho|y_{\mathrm{e}}|}, & y_{\mathrm{e}}y_{\mathrm{eint}} > 0 \end{cases} \tag{7}$$

式中：$k_1>1$；$k_2<|V|$；ρ为收敛速率。

为使偏航角ψ跟踪积分视线角ψ_{ILOS}，使得两者偏差$\psi_{\mathrm{e}}=\psi-\psi_{\mathrm{ILOS}}$逐渐趋向 0，采用 PID 控制方法，以$\psi_{\mathrm{e}}$为输入量，输出纠偏指令$F_{\mathrm{c}}$，控制律如下：

$$F_{\mathrm{c}}=K_1\psi_{\mathrm{e}}+K_2\dot{\psi}_{\mathrm{e}}+K_3\int\psi_{\mathrm{e}}\mathrm{d}t \tag{8}$$

式中：K_1、K_2和K_3分别为 PID 控制的比例系数、微分系数和积分系数。

$$\begin{cases}F_{\mathrm{cl}}=F_{\mathrm{c}},F_{\mathrm{cr}}=0 & F_{\mathrm{c}}>0\\ F_{\mathrm{cr}}=-F_{\mathrm{c}},F_{\mathrm{cl}}=0 & F_{\mathrm{c}}\leqslant 0\end{cases} \tag{9}$$

式中：F_{cl}、F_{cr}分别为左、右主起纠偏机构的纠偏指令。

PID 控制原理如图 5 所示。

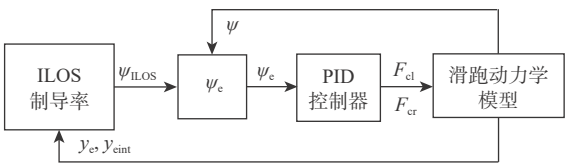


图 5 PID 控制结构

Fig. 5 PID control structure

通过设置合理的 PID 控制参数，便可得到稳定收敛的控制规律，使ψ_{e}趋于 0。当输出纠偏指令F_{c}为正时，左起落架纠偏驱动装置启动，摩擦材料与地面接触产生摩擦力，右侧起落架摩擦材料保持在上位，纠偏指令为负时，右起落架纠偏驱动装置启动，摩擦材料与地面接触产生摩擦力，左侧起落架摩擦材料保持在上位。通过调整电机输出力矩改变摩擦材料受到的地面支持力，使力传感器采集到顶针的垂向载荷达到纠偏指令的值。

在实际情况中，从电机接收到控制信号，到摩擦材料下压达到指定的正压力，存在动态响应延迟。在仿真模型中，采用一阶惯性传递函数进行模拟：

$$F_{\mathrm{s}}=\frac{K_{\mathrm{i}}}{1+T_{\mathrm{i}}s}F_{\mathrm{c}} \tag{10}$$

式中：F_{s}为力传感器采集到的垂向载荷；K_{i}为一阶惯性环节的比例系数；T_{i}为延迟系数。

3.2 纠偏机构受力分析

假设纠偏机构零件均为刚性件，忽略摩擦材料的压缩变形和摩擦过程中的磨损，忽略各运动副连接的间隙与摩擦。当摩擦材料接触道面产生压力时，纠偏机构的受力简图如图 6 所示。图中：P_{p}和Q_{p}分别为摩擦材料受到的道面支持力和沿机体轴x_{b}的摩擦力分量；θ_1为摇臂与道面之间的夹角；θ_2为摇臂与连杆之间的夹角；l_1和l_2分别为摇臂下半部分和上半部分的长度。

根据摇臂的力矩平衡式 (11) 可以得到F_{s}与P_{p}，Q_{p}之间的关系，根据滑橇式起落架地面摩擦特性试验，结合摩擦材料受到的正压力、飞行器滑跑速度

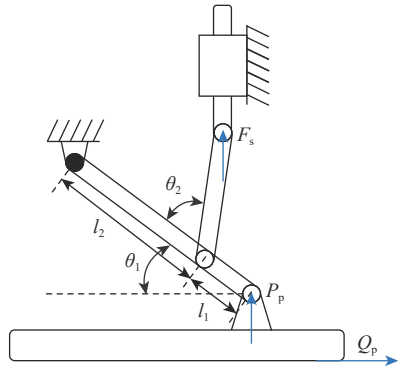


图 6 纠偏机构受力简图

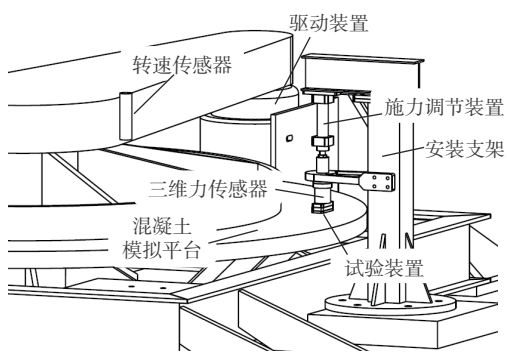
Fig. 6 Force diagram of deviation correction mechanism

和偏航角可以得到Q_{p}与P_{p}之间的关系。

$$P_{\mathrm{p}}(l_1+l_2)\cos\theta_1+Q_{\mathrm{p}}(l_1+l_2)\sin\theta_1=F_{\mathrm{s}}l_2\sin\theta_2/\sin(\theta_1+\theta_2) \tag{11}$$

4 滑橇式起落架地面摩擦特性试验

对于变摩擦滑橇式起落架，准确识别滑跑和纠偏过程中滑橇和摩擦材料的摩擦系数非常关键，因此，开展滑橇式起落架地面摩擦特性试验，试验方案如图 7(a) 所示，实物台架如图 7(b) 所示。试验装置和施力调节装置固定在安装支架上，试验装置的试验件包括滑橇和各种摩擦材料，需要时可进行更换；模拟平台的转盘能在驱动装置的带动下转动，



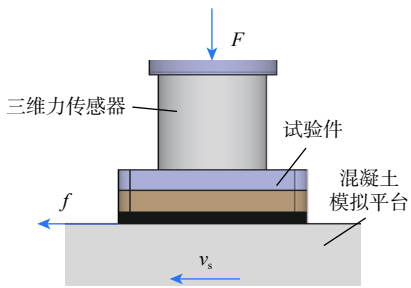
(a) 试验方案示意图

(b) 实物台架

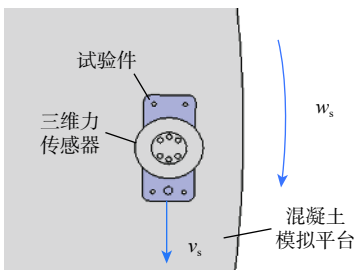
图 7 摩擦试验方案

Fig. 7 Friction test scheme

模拟平台上铺有混凝土涂层, 用于模拟机场的混凝土道面; 试验装置在施力调节装置的作用力下向下接触模拟平台表面产生摩擦力, 施力调节装置可以调节试验装置的安装角度和受到的正压力; 三维力传感器能够同时测量试验装置受到的正压力F和摩擦力f, 如图 8(a) 所示; 转速传感器能够测量转盘的转速w_s, 进而换算成试验件和转盘的相对线速度v_s, 模拟飞行器的航向速度, 如图 8(b) 所示。



(a) 侧视图



(b) 俯视图

图 8　摩擦试验装置

Fig. 8　Friction test equipment

主要的试验步骤如下:

1)试验开始前, 控制施力调节装置, 使得试验装置高于模拟平台上表面。

2)开启驱动装置, 使模拟平台逐渐加速至指定转速, 随后关闭驱动装置。

3)调节施力调节装置, 使试验装置接触模拟平台上表面并使正压力达到指定值, 模拟平台开始减速直至刹停, 由动态测试仪记录此过程中转速传感器和三维力传感器的数据。

4)待试验装置冷却至室温, 调整模拟平台的转速、施力调节装置施加的正压力、试验装置的安装角度或更换试验件, 进行不同工况的试验。

5　数据及仿真分析

5.1　摩擦试验结果分析

摩擦材料和金属滑橇的部分摩擦系数原始数据如图 9 所示。可以看出, L-1S 材料的摩擦系数在 0.5 左右, G-60 材料的摩擦系数在 0.35 左右, 滑橇的摩擦系数在 0.14 左右。

变摩擦滑橇式起落架要求摩擦材料的摩擦系数显著大于滑橇摩擦系数, 因此, 本文选用 L-1S 材料作为滑橇的摩擦材料。进一步对 L-1S 材料和滑橇的摩擦数据进行分析, 并将向下正压力折算成底面的压强P, 得到不同压强和速度下滑橇和摩擦材料的摩擦系数值, 将其作为试验点, 用高次多项式进行曲面拟合, 得到滑橇和摩擦材料摩擦系数的拟合曲面和拟合多项式。采用速度的 3 次、压强的 4 次多项式能较好地拟合两者的摩擦系数, 拟合多项式为

$$\begin{aligned}\mu = {} & p_{00} + p_{10}v_s + p_{01}P + p_{20}v_s^2 + p_{11}Pv_s + \\ & p_{02}P^2 + p_{30}v_s^3 + p_{21}Pv_s^2 + p_{12}P^2v_s + p_{03}P^3 + \\ & p_{31}Pv_s^3 + p_{22}P^2v_s^2 + p_{13}P^3v_s + p_{04}P^4 \end{aligned} \tag{12}$$

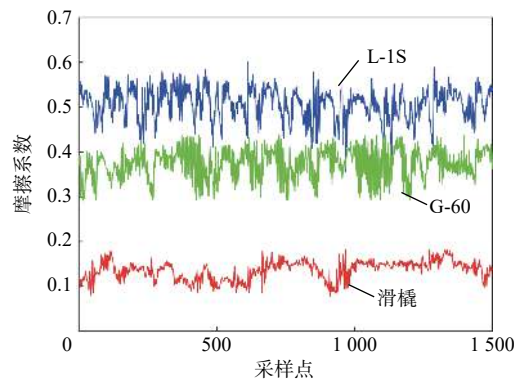


图 9　摩擦系数原始数据

Fig. 9　Raw data of friction coefficient

拟合参数如表 1 所示, 拟合曲面分别如图 10 和图 11 所示。

从图 10 和图 11 中可以看出, 滑橇的摩擦系数随速度和压强的增加先增大再减小, 在速度v_s = 15 m/s、压强P = 0.35 MPa左右摩擦系数达到极大值; 摩擦材料的摩擦系数也随速度和压强的增加先增大再减小, 在速度v_s = 10 m/s、压强P = 0.1 MPa左右达到极大值。

表 1　拟合多项式的系数

Table 1　Coefficients of fitting polynomials

系数	滑橇摩擦系数μ_s	材料摩擦系数μ_p
p_{00}	0.067	0.390 7
p_{10}	0.005 29	0.013 64
p_{01}	1.009×10^{-6}	1.798×10^{-6}
p_{20}	-3.418×10^{-4}	-6.458×10^{-4}
p_{11}	1.404×10^{-8}	-6.084×10^{-8}
p_{02}	-6.93×10^{-12}	-1.092×10^{-11}
p_{30}	5.825×10^{-6}	8.209×10^{-6}
p_{21}	-5.442×10^{-6}	2.545×10^{-9}
p_{12}	2.115×10^{-14}	2.671×10^{-14}
p_{03}	1.938×10^{-17}	3.134×10^{-17}
p_{31}	-4.945×10^{-12}	-3.235×10^{-11}
p_{22}	1.729×10^{-15}	-6.44×10^{-16}
p_{13}	-1.438×10^{-19}	3.3×10^{-20}
p_{04}	-1.814×10^{-23}	-3.337×10^{-23}

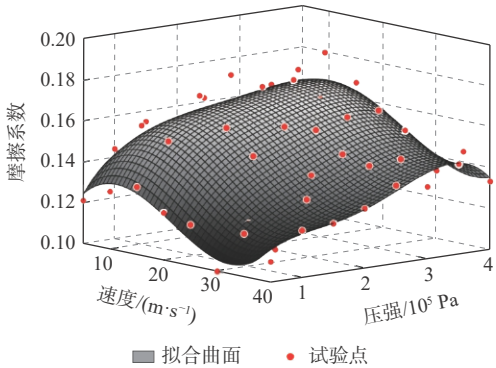


图 10 滑橇摩擦系数的拟合曲面

Fig. 10 Fitting surface of skid friction coefficient

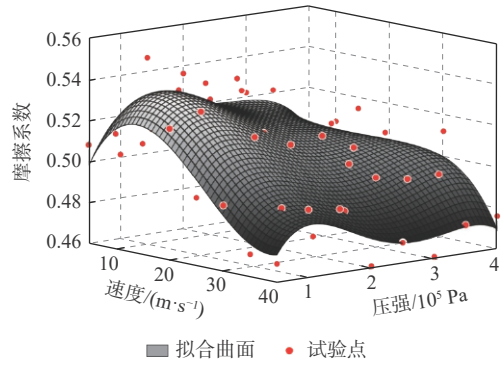


图 11 摩擦材料摩擦系数的拟合曲面

Fig. 11 Fitting surface of material friction coefficient

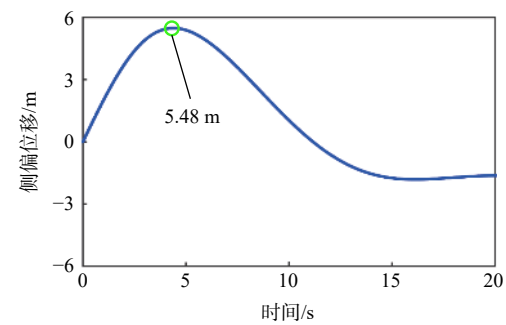
5.2 纠偏仿真分析

将滑橇和摩擦材料的摩擦系数拟合多项式代入仿真纠偏模型，以滑橇受到的支持力和飞行器滑跑速度为输入，得到滑橇实时的摩擦系数；以电机驱动力和飞行器滑跑速度为输入，通过纠偏机构受力关系式计算出摩擦材料的正压力，从而得到摩擦材料实时的摩擦系数。

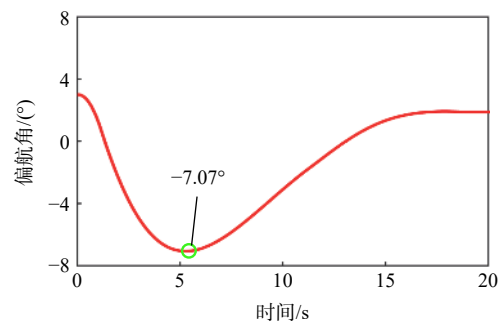
5.2.1 初始偏航角和侧偏距离下的纠偏响应

假设飞行器初始航向速度为 40 m/s，给飞行器一定的初始偏航角和侧偏距离，模拟其滑跑减速并纠偏直至刹停的过程，仿真结果如图 12 和图 13 所示。图 12(a) 和图 12(b) 展示了 3°的初始偏航角下，侧偏距离和偏航角的变化曲线，在纠偏过程过，最大侧偏距离达到 5.48 m，最大偏航角为 7.07°；图 12(c) 展示了纠偏过程中左、右主起落架摩擦材料摩擦系数 μ_{pl}、μ_{pr}的变化曲线，图 12(d) 展示了纠偏过程中左、右主起落架滑橇摩擦系数 μ_{sl}、μ_{sr}的变化曲线。

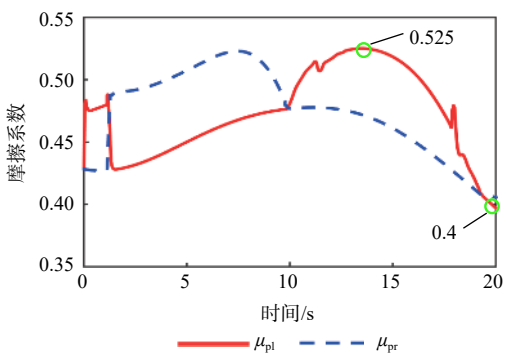
图 13(a) 和 13(b) 分别展示了 3 m 的初始侧偏距离下，侧偏距离和偏航角的变化曲线，纠偏过程中最大偏航角为−2.76°；图 13(c) 展示了纠偏过程中左、右主起落架摩擦材料摩擦系数 μ_{pl}、μ_{pr}的变化曲线，图 13(d) 展示了纠偏过程中左、右主起落架滑橇摩擦系数 μ_{sl}、μ_{sr}的变化曲线。从图 12 和图 13 中可以看出，当飞行器出现较小的侧偏距离和偏航角时，飞行器可以通过纠偏机构差动纠偏调整滑跑方向并回到跑道中心线；纠偏过程中，摩擦材料和滑橇的摩擦系数随着滑跑速度和正压力的变化而不断变化，摩擦材料摩擦系数在 0.4～0.528 之间变



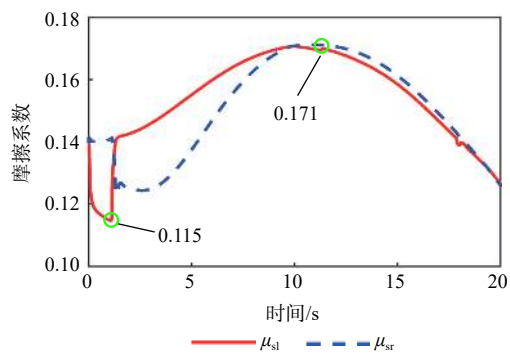
(a) 侧偏距离变化曲线



(b) 偏航角变化曲线



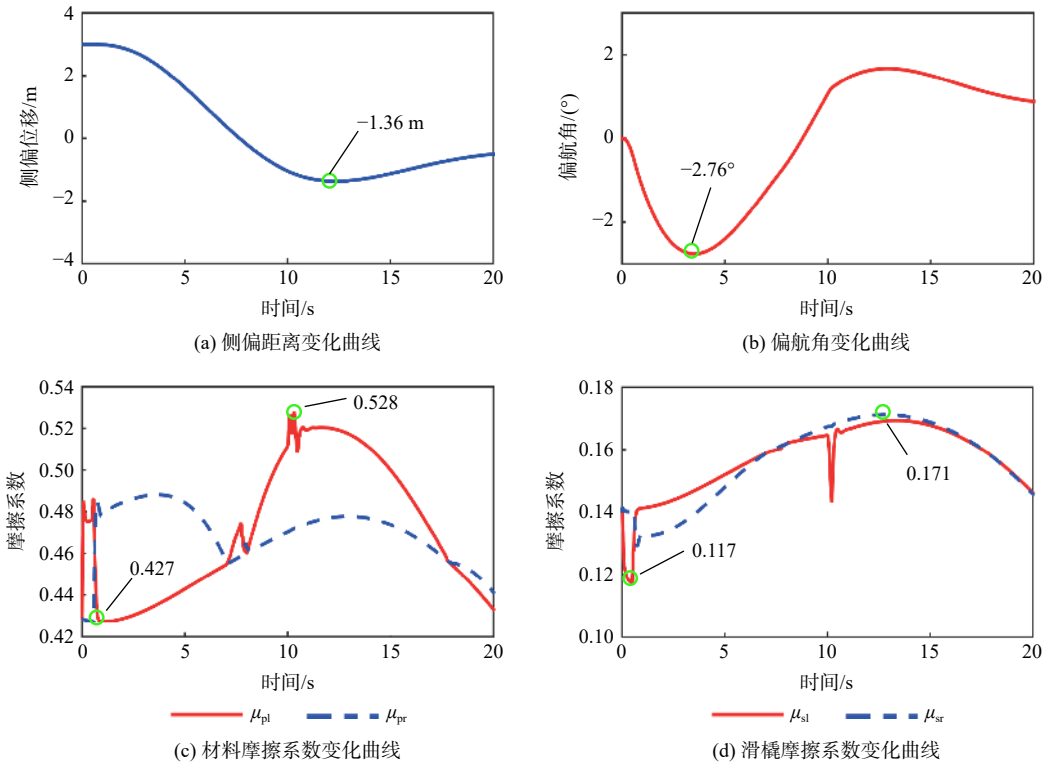
(c) 材料摩擦系数变化曲线



(d) 滑橇摩擦系数变化曲线

图 12 初始偏航角下的纠偏响应

Fig. 12 Deviation-correction responses to initial yaw angle



(a) 侧偏距离变化曲线　(b) 偏航角变化曲线

(c) 材料摩擦系数变化曲线　(d) 滑橇摩擦系数变化曲线

图 13　初始侧偏距离下的纠偏响应

Fig. 13　Deviation-correction responses to initial lateral displacement

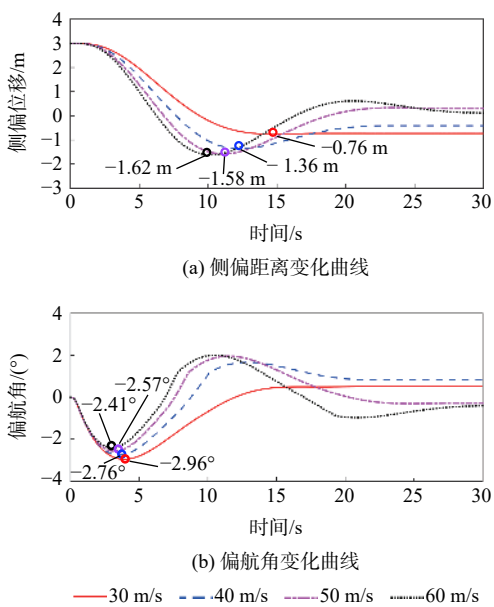
化, 滑橇摩擦系数在 0.115～0.171 之间变化。

5.2.2　不同初始工况的影响性分析

1) 不同初始航向速度。进一步研究不同工况下飞行器的纠偏性能。假设飞行器存在初始侧偏距离 3 m, 图 14 展示了初始航向速度为 30～60 m/s 时侧偏距离和偏航角的变化曲线。

从图 14 中可以看出, 飞行器均能纠正 3 m 的初始侧偏距离并回到跑道中心线; 当初始航向速度从 30 m/s 逐渐增加至 60 m/s 时, 侧偏距离最大偏差从 0.76 m 逐渐增大到 1.62 m, 而最大偏航角从 2.96°减小至 2.41°。

2) 不同初始侧偏距离。假设飞行器初始航向速度为 40 m/s, 图 15 展示了初始侧偏距离 2～8 m 时侧偏距离和偏航角的变化曲线。可以看出, 飞行

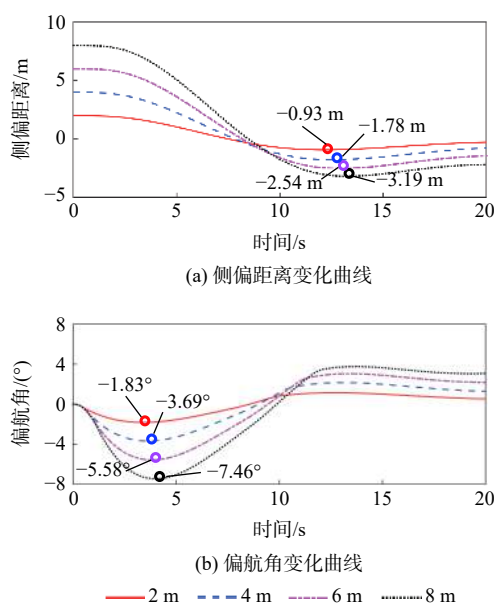


(a) 侧偏距离变化曲线

(b) 偏航角变化曲线

图 14　不同初始航向速度下的纠偏响应

Fig. 14　Deviation-correction responses to different initial velocities



(a) 侧偏距离变化曲线

(b) 偏航角变化曲线

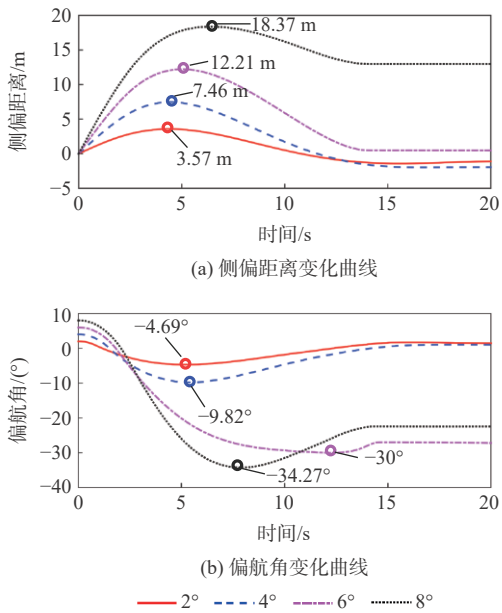
图 15　不同初始侧偏距离下的纠偏响应

Fig. 15　Deviation-correction responses to different initial lateral displacements

器在 2～8 m 的初始侧偏距离下均能成功纠偏并回到跑道中心线附近；当初始侧偏距离逐渐增大时，纠偏过程中侧偏距离的最大偏差从 0.93 m 增大至 3.19 m，最大偏航角从 1.83°增大至 7.46°。

3）不同初始偏航角。假设飞行器初始航向速度为 40 m/s，图 16 展示了初始偏航角为 2°～8°时侧偏距离和偏航角的变化曲线。

从图 16 中可以看出，飞行器在 2°和 4°的初始偏航角下能成功纠偏；当初始偏航角增大到 6°之后，偏航角的最大偏差将达到 30°，最大侧偏距离将达到 12.21 m，增加至 8°之后，最大侧偏距离继续增加至 18.37 m，偏航角的最大偏差进一步增大，且飞行器将无法回到跑道中心线附近。相比于初始侧偏距离，飞行器更难以纠正初始偏航角，因此，在飞行器着陆前，应调整姿态，尽量减小初始偏航角。



(a) 侧偏距离变化曲线

(b) 偏航角变化曲线

图 16 不同初始偏航角下的纠偏响应

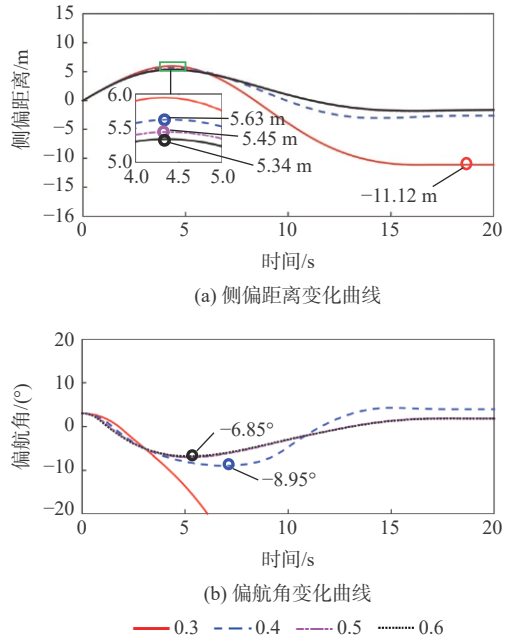
Fig. 16 Deviation-correction responses to different initial yaw angles

5.2.3 不同摩擦系数的影响性分析

摩擦材料是纠偏机构的关键部件，为了研究摩擦材料摩擦系数对滑跑纠偏性能的影响，更换摩擦系数不同的摩擦材料进行纠偏仿真。

飞行器初始航向速度为 40 m/s，初始偏航角为 3°，使用的摩擦材料摩擦系数约为 0.5，更换平均摩擦系数分别为 0.3、0.4 和 0.6 的摩擦材料进行仿真，仿真结果如图 17 所示。

从图 17 中可以看出，当摩擦系数为 0.3 时，飞行器将出现很大的偏航角且无法纠正方向；当摩擦系数从 0.4 增大到 0.6 后，侧偏距离最大值从 5.63 m 减小为 5.34 m，最大偏航角从 8.95°减小为 6.85°，纠偏性能逐渐增强。在此种工况下，选用摩擦系数为



(a) 侧偏距离变化曲线

(b) 偏航角变化曲线

图 17 不同摩擦材料下的纠偏响应

Fig. 17 Deviation-correction responses to different friction materials

0.5 和 0.6 的摩擦材料，纠偏效果基本没有变化。

5.2.4 滑跑安全边界分析

为了进一步定量分析飞行器纠偏系统的纠偏能力，对不同初始航向速度、初始侧偏距离和初始偏航角的工况进行批量仿真，假设跑道中心线左右 15 m 为安全范围，若飞行器在滑跑纠偏过程中一直处于安全范围，则代表纠偏成功[24]。批量仿真结果如图 18 所示。当飞行器的初始工况在包络面内时，飞行器能够在安全范围内滑跑，随着初始航向速度的增大，由侧偏距离与偏航角构成的二维安全边界面积逐渐缩小。

进一步研究不同摩擦系数下的滑跑安全边界，图 19(a)～图 19(d) 分别展示了选用平均摩擦系数为 0.3、0.4、0.5 和 0.6 的摩擦材料时飞行器滑跑的安全边界。从 0.3 至 0.6，摩擦系数每提升 0.1，飞行器滑跑的安全边界范围提升率为 16.6%、7.3% 和 1.5%。

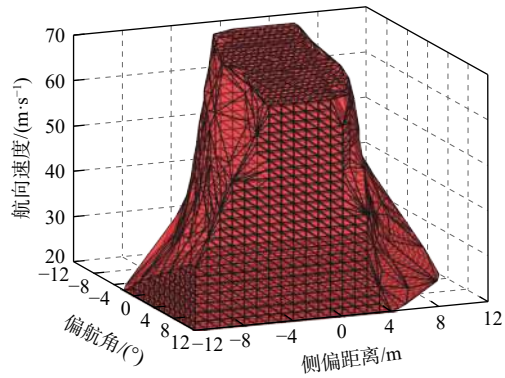
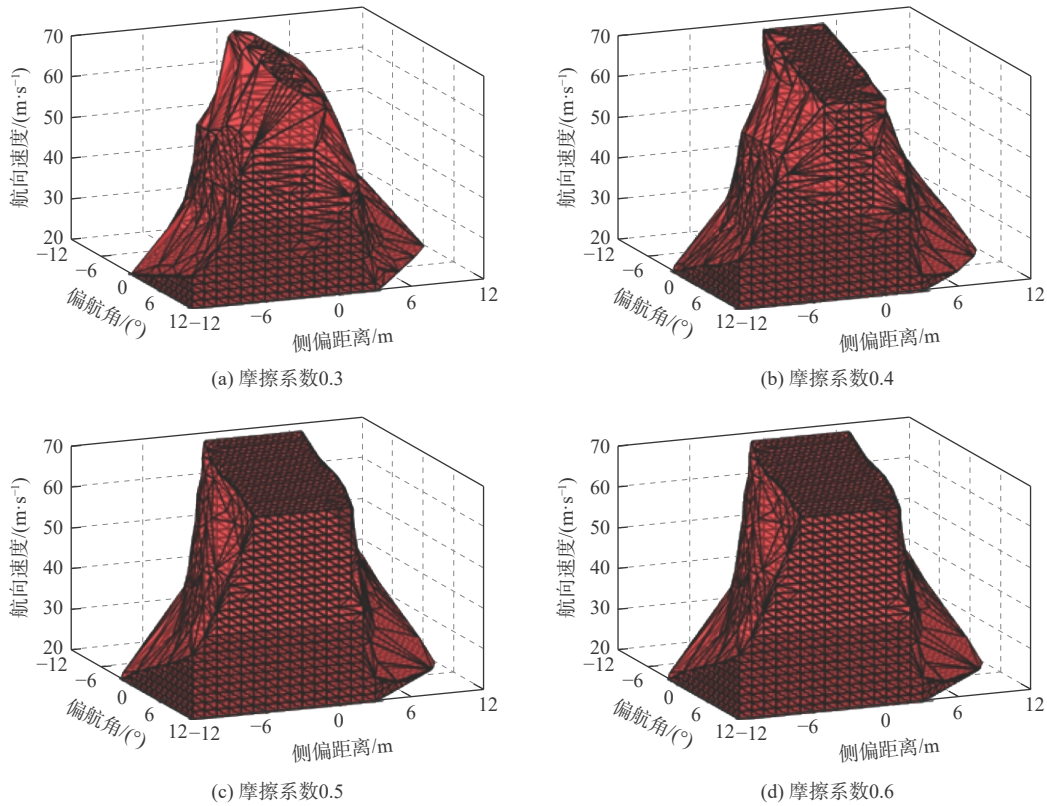


图 18 滑跑安全边界

Fig. 18 Safe set of taxiing



(a) 摩擦系数0.3　(b) 摩擦系数0.4

(c) 摩擦系数0.5　(d) 摩擦系数0.6

图 19　不同摩擦材料下的滑跑安全边界

Fig. 19　Safe sets of taxiing to different friction materials

5.2.5　控制系统优化分析

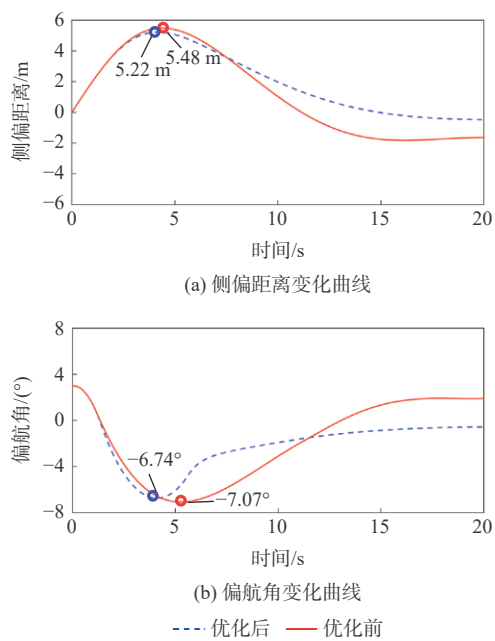
为了进一步提高纠偏性能, 采用最优化方法调整控制参数, 基于粒子群算法的多目标最优 PID 控制是一种有效的控制方法[25]。粒子群算法通过群体中个体之间的协作和信息共享来寻找最优解, 本文选用粒子群算法[26]来寻找最优 PID 控制系数。

假设在一个 D 维的搜索空间中, 由n个粒子组成的种群$\boldsymbol{X}=(\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n)$, 其中第$i$个粒子表示为一个 D 维向量$\boldsymbol{X}_i=(x_{i1},x_{i2},\cdots,x_{iD})^{\mathrm{T}}$, 代表第$i$个粒子在 D 维搜索空间中的位置, 根据目标函数即可计算出每个粒子位置$\boldsymbol{X}_i$对应的适应度值。第i个粒子速度$\boldsymbol{V}_i=(v_{i1},v_{i2},\cdots,v_{iD})^{\mathrm{T}}$, 其个体极值$\boldsymbol{Q}_i=(q_{i1},q_{i2},\cdots,q_{iD})^{\mathrm{T}}$, 种群的群体极值为$\boldsymbol{Q}_{\mathrm{g}}=\left(q_{\mathrm{g}1},q_{\mathrm{g}2},\cdots,q_{\mathrm{g}D}\right)^{\mathrm{T}}$。在每次迭代过程中, 粒子通过个体极值和群体极值更新自身的速度和位置, 即

$$\begin{cases} v_{id}^{t+1}=\omega v_{id}^{t}+c_1r_1(q_{id}^{t}-x_{id}^{t})+c_2r_2(q_{\mathrm{g}d}^{t}-x_{id}^{t}) \\ x_{id}^{t+1}=x_{id}^{t}+v_{id}^{t+1} \end{cases} \tag{13}$$

式中: ω为惯性因子; $d=1,2,\cdots,D$; $i=1,2,\cdots,n$; t为当前迭代次数; c_1、c_2为非负的常数, 称为加速度因子; r_1、r_2为分布于[0,1]区间的随机数。

假设飞行器初始航向速度为 40 m/s, 以初始偏航角为 3°的工况为例, 首先产生一定数量的粒子群, 将粒子群中的粒子依次赋值给 PID 控制器的 3 个参数, 然后进行滑跑仿真, 得到该组参数的性能指标J, 令$J=\alpha_1y_{\mathrm{emax}}+\alpha_2\psi_{\max}$, 其中, α_1、α_2为非负常数, y_{emax}和$\psi_{\max}$分别为滑跑过程中侧偏距离和偏航角绝对值的最大值。用该性能指标J作为粒子的适应值, 进行粒子群更新, 并不断迭代得到最优的 PID 系数, 优化前后的仿真结果如图 20 所示。最大侧



(a) 侧偏距离变化曲线

(b) 偏航角变化曲线

图 20　优化前后的纠偏响应

Fig. 20　Deviation-correction responses before and after optimization

偏距离减小了4.7%，最大偏航角也减小了4.7%，优化后偏航角的收敛速度明显加快，侧偏距离和偏航角的稳态误差也显著减小，经过优化后，飞行器的纠偏性能得以提高。

6 结 论

针对传统滑橇式起落架地面滑跑阶段航向稳定性较差的问题，设计了一种具备航向增稳功能的新型变摩擦滑橇式起落架，建立了飞行器地面滑跑纠偏模型，开展了滑橇式起落架地面摩擦特性试验，并将摩擦特性代入滑跑模型中。典型工况的仿真结果验证了新型滑橇的纠偏能力。通过综合分析得到了以下结论：

1）本文提出了一种具备航向增稳功能的新型变摩擦滑橇式起落架，能够有效节省起落架的收放空间，并可在铺设混凝土道面的常规机场跑道完成着陆滑跑，为机身收纳空间不足的高超声速飞行器提供了一种可行的起落架设计方案。

2）滑橇和摩擦材料的摩擦系数随着压强和速度的增大都是先增大后减小的，当速度为15 m/s、压强为0.35 MPa时，滑橇的摩擦系数达到极大值；当速度为10 m/s、压强为0.1 MPa时，摩擦材料的摩擦系数达到极大值。得到了滑橇和摩擦材料摩擦系数的拟合公式，并将其代入滑跑模型中，得到了摩擦系数的变化曲线。

3）滑跑纠偏仿真结果显示，采用新型变摩擦滑橇式起落架的飞行器能够适应不同初始航向速度、侧偏距离和偏航角的工况，滑跑的安全边界随着初始航向速度增大而收缩，机构纠偏性能随着摩擦材料摩擦系数从0.3增大到0.5逐渐增强，进一步增大到0.6后提升不明显，优化后的最大侧偏距离和偏航角减小4.7%。

参考文献（References）

[1] 张丽静, 刘东升, 于存贵, 等. 高超声速飞行器[J]. 航空兵器, 2010, 17(2): 13-16.
ZHANG L J, LIU D S, YU C G, et al. Hypersonic aircraft[J]. Aero Weaponry, 2010, 17(2): 13-16(in Chinese).

[2] 杜善义, 方岱宁, 孟松鹤, 等. “近空间飞行器的关键基础科学问题”重大研究计划结题综述[J]. 中国科学基金, 2017, 31(2): 109-114.
DU S Y, FANG D N, MENG S H, et al. Review of the achievements of major research plan on “key fundamental scientific problems on hypersonic vehicle”[J]. Bulletin of National Natural Science Foundation of China, 2017, 31(2): 109-114(in Chinese).

[3] 黄伟, 罗世彬, 王振国. 临近空间高超声速飞行器关键技术及展望[J]. 宇航学报, 2010, 31(5): 1259-1265.
HUANG W, LUO S B, WANG Z G. Key techniques and prospect of near-space hypersonic vehicle[J]. Journal of Astronautics, 2010, 31(5): 1259-1265(in Chinese).

[4] 李宪开, 王霄, 柳军, 等. 水平起降高超声速飞机气动布局技术研究[J]. 航空科学技术, 2020, 31(11): 7-13.
LI X K, WANG X, LIU J, et al. Research on the aerodynamic layout design for the horizontal take-off and landing hypersonic aircraft[J]. Aeronautical Science & Technology, 2020, 31(11): 7-13(in Chinese).

[5] 朱林昊, 邹泽铧, 印寅, 等. 面向狭小空间的起落架收放轨迹智能设计方法[J]. 北京航空航天大学学报, 2021, 47(12): 2560-2570.
ZHU L H, ZOU Z H, YIN Y, et al. Intelligent design method of landing gear retraction and extension trajectory for narrow space[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(12): 2560-2570(in Chinese).

[6] 张猛. 着陆过程航空轮胎热-力学特性研究[D]. 哈尔滨: 哈尔滨工业大学, 2019: 53-57.
ZHANG M. Study on thermal-mechanical characteristics of aircraft tires during landing[D]. Harbin: Harbin Institute of Technology, 2019: 53-57(in Chinese).

[7] 段卓毅. X系列飞行器概览[M]. 北京: 航空工业出版社, 2017: 267-351.
DUAN Z Y. Survey of X-planes[M]. Beijing: Aviation Industry Press, 2017: 267-351(in Chinese).

[8] SEEDHOUSE E. SpaceX[M]. SpaceX's Dragon: America's Next Generation Spacecraft. Berlin: Springer, 2016: 1-9.

[9] LEMAY D, CHAMAILLARD Y, BASSET M, et al. Gain-scheduled yaw control for aircraft ground taxiing[J]. IFAC Proceedings Volumes, 2011, 44(1): 12970-12975.

[10] 王勇, 王英勋. 无人机滑跑纠偏控制[J]. 航空学报, 2008, 29(增刊1): 142-149.
WANG Y, WANG Y X. Lateral deviation correction control for UAV taxiing[J]. Acta Aeronautica et Astronautica Sinica, 2008, 29(Sup 1): 142-149(in Chinese).

[11] YIN Q Z, NIE H, WEI X, et al. Aircraft electric anti-skid braking and combined direction control system using co-simulation and experimental methods:[J]. Proceedings of the Institution of Mechanical Engineers, Part G:Journal of Aerospace Engineering, 2020, 234(2): 173-191.

[12] 李继广, 董彦非, 杨芳, 等. 基于反步控制方法的菱形翼无人机起飞滑跑控制[J]. 北京航空航天大学学报, 2020, 46(3): 496-504.
LI J G, DONG Y F, YANG F, et al. Takeoff taxiing control of joined-wing UAV based on backstepping method[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(3): 496-504(in Chinese).

[13] LI X B, WEI X H, WANG Y A, et al. Nonlinear tracking deviation correction control of airplanes in high-speed landing taxiing[J]. Journal of Guidance, Control, and Dynamics, 2020, 43(12): 2413-2424.

[14] 黄平. 摩擦学教程[M]. 北京: 高等教育出版社, 2008: 60-74.
HUANG P. Tribology course[M]. Beijing: Higher Education Press, 2008: 60-74(in Chinese).

[15] 魏兵辉, 孙琨, 王云鹏, 等. 高速干滑动摩擦系数的有限元仿真计算研究[J]. 西安交通大学学报, 2020, 54(6): 82-89.
WEI B H, SUN K, WANG Y P, et al. Finite element prediction for friction coefficient of high speed dry sliding[J]. Journal of Xi'an

Jiaotong University, 2020, 54(6): 82-89(in Chinese).

[16] 魏龙, 张鹏高, 刘其和, 等. 接触式机械密封端面摩擦系数影响因素分析与试验[J]. 摩擦学学报, 2016, 36(3): 354-361.
WEI L, ZHANG P G, LIU Q H, et al. Influencing factors analysis and experiments of friction coefficient between the end faces for contact mechanical seals[J]. Tribology, 2016, 36(3): 354-361(in Chinese).

[17] WU B Z, QIN D T, HU J J, et al . Experimental data mining research on factors influencing friction coefficient of wet clutch[J]. Journal of Tribology' 2021, 143(12): 121802.

[18] 陶润, 侯之超. 受径向力滚动轴承摩擦力矩的测试和函数拟合[J]. 清华大学学报(自然科学版), 2014, 54(6): 744-749.
TAO R, HOU Z C. Measurement and curve fitting of the friction torque of rolling bearings subjected to radial loads[J]. Journal of Tsinghua University (Science and Technology), 2014, 54(6): 744-749 (in Chinese).

[19] SADIEN E, ROOS C, BIROUCHE A, et al. Control design oriented modeling of an on-ground aircraft[C]//2018 17th European Control Conference (ECC). Limassol: IEEE Press, 2018: 2757-2762.

[20] 李洁玉. 高速水平回收无人机综合纠偏方法研究[D]. 南京: 南京航空航天大学, 2017: 9-13.
LI J Y. Research on lateral deviation correction control for high-speed UAV of horizontal recovery[D]. Nanjing: Nanjing University of Aeronautics and Astronautics , 2017: 9-13(in Chinese).

[21] SUJIT P B, SARIPALLI S, SOUSA J B. Unmanned aerial vehicle path following: A survey and analysis of algorithms for fixed-wing unmanned aerial vehicless[J]. IEEE Control Systems Magazine, 2014, 34(1): 42-59.

[22] PANTELEY E, LORIA A. On global uniform asymptotic stability of nonlinear time-varying systems in cascade[J]. Systems & Control Letters, 1998, 33(2): 131-138.

[23] 瞿洋, 徐海祥, 余文曌, 等. 基于ILOS的欠驱船舶循迹控制[J]. 武汉理工大学学报(交通科学与工程版), 2016, 40(5): 834-838.
QU Y, XU H X, YU W Z, et al. Integral line-of-sight guidance for path following of underactuated marine surface vessels[J]. Journal of Wuhan University of Technology (Transportation Science & Engineering), 2016, 40(5): 834-838(in Chinese).

[24] LIANG T T, YIN Q Z, FANG W G, et al. The maximum taxiing safe set of the wheelskid aircraft under op-timal control of rudder[J]. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 2021, 235(15): 2274-2287.

[25] 王海涛, 蒋经凯. 飞机起落架主动控制与滑跑性能优化研究[J]. 西北工业大学学报, 2013, 31(2): 223-228.
WANG H T, JIANG J K. Active control of landing gear and taxiing performance optimization[J]. Journal of Northwestern Polytechnical University, 2013, 31(2): 223-228(in Chinese).

[26] EBERHART R, KENNEDY J. A new optimizer using particle swarm theory[C]//Micro Machine and Human Science, 1995. MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science. Piscataway: IEEE Press, 1995: 39-43

Taxiing deviation-correction control of a new variable-friction equipped-skid aircraft

WANG Yucheng[1, 2], YIN Yin[1, 2], LIANG Taotao[1, 2], WEI Xiaohui[1, 2, *]

(1. State Key Laboratory of Mechanics and Control of Mechanical Structures, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;
2. Key Laboratory of Fundamental Science for National Defense-Advanced Design Technology of Flight Vehicle, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

Abstract: Flat hypersonic aircraft are better suited to skid landing gear than wheeled landing gear because of its light weight and compact design. However, the inherent characteristic of the wheel skid layout causes the problem of heading instability during mid-speed taxiing. Aiming at the problem, a variable-friction skid landing gear with a directional stability augmentation function is designed. Firstly, a nonlinear ground taxiing model is established, considering aerodynamic forces, ground forces, correction mechanism model, and deviation-correction control system.The ground friction characteristics experiment findings are then used to create the multi-parameter fitting polynomials of the friction coefficients of the skid and material, which are then incorporated into the dynamic model discussed earlier. Finally, an integral line of sight (ILOS) guidance is introduced to realize the effective tracking of the runway centerline during the rollout phase. Particle swarm optimization is introduced to optimize control parameters. Results show that the friction coefficients of the skid and the material first increase and then decrease with the increase of speed and pressure. The effectiveness of the deviation-correction control system is verified by taxiing simulation results under typical initial conditions. The safety set of taxiing increases by 16.6% when the friction coefficient of the friction material increases from 0.3 to 0.4.

Keywords: skid landing gear; friction coefficient ground; taxiing; deviation-correction control; integral line of sight guidance law

Received: 2021-11-19; **Accepted**: 2022-02-16; **Published Online**: 2022-03-22 16:57

URL: kns.cnki.net/kcms/detail/11.2625.V.20220321.2030.005.html

Foundation items: National Natural Science Foundation of China (52172368); The Research Fund of State Key Laboratory of Mechanics and Control of Mechanical Structures (MCMS-I-0221Y02); The Fundamental Research Funds for the Central Universities (NF2018001)

* **Corresponding author**. E-mail: wei_xiaohui@nuaa.edu.cn

http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2022.0939

基于CFD/CSD耦合的高速射弹尾拍载荷特性研究

许云涛[1,2]，檀大林[2]，杨超[1,*]

(1. 北京航空航天大学 航空科学与工程学院，北京 100191；2. 北京机电工程研究所，北京 100074)

摘 要：为提高射弹尾拍载荷的预测精度，建立了一套基于计算流体动力学（CFD）/计算结构动力学（CSD）双向耦合分析的计算方法和程序。射弹流体计算主控方程采用耦合SSTk-ω湍流模型和Schnerr-Sauer空化模型的Navier-Stokes方程，射弹结构计算采用基于模态叠加法简化的结构动力学方程，流固耦合界面插值采用径向基函数法，网格变形采用弹簧网格法。分别对泡型计算方法和流固耦合方法进行验证，在此基础上，计算对比1 000 m/s速度下射弹刚体和弹性体的尾拍泡型、结构变形和尾拍流体载荷特性差异。计算表明：弹性体尾拍过程，射弹泡型会产生弹身“二次拍击”、“局部沾湿”和沾湿面积增大等特殊现象，结构变形由弹性一弯模态主导，较大的变形引起流体载荷增大27%～105%，尾拍姿态角增大13%，尾拍频率增加20%，流固耦合效应对尾拍泡型、尾拍载荷和尾拍弹道均产生了较强的影响。

关 键 词：射弹；CFD/CSD耦合；尾拍；载荷；刚体；弹性体
中图分类号：TJ630
文献标志码：A **文章编号**：1001-5965(2023)09-2539-08

空化是一种液体特有的现象，与物体有高速相对运动的液体，常因相变而产生空化流动[1]。充分利用空化流动的特点在航行体表面生成覆体超空泡，以大幅降低摩擦阻力的技术，称为超空泡减阻技术[2]。超空泡射弹就是一种利用了超空泡减阻技术的动能兵器，因其阻力较小、航行速度可高达1 000 m/s，可以实现对水下目标的精准、快速毁伤攻击[3-5]。

由于高速发射时存在无法避免的微小扰动，超空泡射弹在水下运动时，尾部通常以周期性撞击空泡壁来维持稳定，这种现象称为尾拍现象[6]。由于尾拍运动是超空泡射弹维持水下稳定航行的最主要方式，且尾拍载荷对射弹的弹道特性、弹体的强度和运动稳定性都有重要影响，尾拍问题受到了广泛的关注和研究。

超空泡射弹的长细比通常大于15且运动速度极高，运动过程的结构变形效应不可避免，结构变形条件下的流体动力载荷必然发生变化。研究结果表明，射弹的流体动力载荷与弯曲变形之间存在正反馈[7-8]。因此，高速射弹在运动过程中存在较强的结构与流动耦合问题。

为研究高速射弹在运动过程中的结构响应与流场特性，学者们研究发展了不同的数值方法。计算流体动力学（computational fluid dynamics，CFD）方法可以准确模拟尾拍流场特性，但其采用刚体假设，无法耦合计算结构动力学（computational structure dynamics，CSD），导致结构响应精确度偏低[2,5-6]；耦合欧拉拉格朗日（coupled Euler Lagrangian，CEL）方法可以准确地模拟结构响应，但CEL方法难以考虑空化模型，流场求解精度较CFD方法偏低[9-10]。CFD/CSD双向耦合计算方法可避免CEL方法和CFD方法的缺点[8,11]，可以更精确地模拟结构响应与流场特性，但空泡流场与结构弹性振动的耦合作用导致计算效率较低，计算耗时

收稿日期：2022-11-24；录用日期：2023-02-26；网络出版时间：2023-04-03 15:44
网络出版地址：kns.cnki.net/kcms/detail/11.2625.V.20230403.1102.001.html
*通信作者. E-mail：yangchao@buaa.edu.cn

引用格式：许云涛，檀大林，杨超. 基于CFD/CSD耦合的高速射弹尾拍载荷特性研究[J]. 北京航空航天大学学报，2023，49(9)：2539-2546. XU Y T，TAN D L，YANG C. Study on tail-slap load characteristics of high-speed projectile based on CFD/CSD coupling[J]. Journal of Beijing University of Aeronautics and Astronautics，2023，49(9)：2539-2546 (in Chinese).

巨大。

鉴于模态叠加法可以高效求解结构响应，本文发展了一种结合模态叠加法的高效率高精度CFD/CSD双向耦合计算方法。完成了典型超空泡射弹尾拍过程的流固耦合仿真计算，获得了该对象尾拍过程中刚体和弹性体的空泡流型、结构变形、尾拍流体载荷的差异，可为尾拍现象的研究提供参考。

1 流固耦合模型

1.1 流体计算方程

超空泡射弹尾拍过程涉及空气、液态水、水蒸汽三相流动，本文基于流体体积（volume of fluid, VOF）模型模拟多相流场。VOF模型是典型的均质多相流模型，在超空泡流动的数值仿真计算中应用广泛[8,12-14]。流体求解满足以下基本方程。

连续性方程：

$$\frac{\partial \rho_{\mathrm{m}}}{\partial t}+\frac{\partial\left(\rho_{\mathrm{m}} u_j\right)}{\partial x_j}=0 \tag{1}$$

动量方程：

$$\frac{\partial\left(\rho_{\mathrm{m}} u_i\right)}{\partial t}+\frac{\partial\left(\rho_{\mathrm{m}} u_i u_j\right)}{\partial x_j}= -\frac{\partial p}{\partial x_i}+\frac{\partial}{\partial x_j}\left[\left(\mu_{\mathrm{m}}+\mu_t\right)\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)\right] \tag{2}$$

式中：$\rho_{\mathrm{m}}=\sum_{k=1}^{n}\alpha_k\rho_k$和$\mu_{\mathrm{m}}=\sum_{k=1}^{n}\alpha_k\mu_k$分别为按体积分数确定的混合物密度和动力黏性系数，$\alpha_k$为体积分数；$u_i$为第$i$相的速度；$p$为压力；$n$为多相流的相数。

选取Schnerr-Sauer模型模拟尾拍过程中的空化现象，蒸汽相体积分数控制方程为

$$\frac{\partial\left(\alpha_{\mathrm{v}}\rho_{\mathrm{v}}\right)}{\partial t}+\frac{\partial\left(\alpha_{\mathrm{v}}\rho_{\mathrm{v}} v_{\mathrm{m}}\right)}{\partial x_i}=\frac{\rho_{\mathrm{v}}\rho_{\mathrm{l}}}{\rho_{\mathrm{m}}}\alpha_{\mathrm{v}}\left(1-\alpha_{\mathrm{v}}\right)\frac{3}{R}\sqrt{\frac{2}{3}\cdot\frac{p_{\mathrm{v}}-p}{\rho_{\mathrm{l}}}} \tag{3}$$

式中：α_{v}为气相体积分数；ρ_{v}为气相密度；v_{m}为混合相速度；ρ_{l}为液相密度；ρ_{m}为混合相密度；R为气泡半径；p_{v}为饱和蒸气压。右端表示蒸汽相和液相的质量传输速率，下标v、l分别表示蒸汽相与液相。

湍流模型采用剪切应力传输SSTk-ω模型，该模型对壁面切应力计算准确，计算结果相对稳定[15]。

1.2 结构动力学方程

射弹结构模型离散后可通过非保守系统的拉格朗日方程得到如下的结构动力学方程[16]：

$$\boldsymbol{M}\ddot{\boldsymbol{x}}+\boldsymbol{C}\dot{\boldsymbol{x}}+\boldsymbol{K}\boldsymbol{x}=\boldsymbol{F} \tag{4}$$

式中：$\boldsymbol{M}$、$\boldsymbol{C}$和$\boldsymbol{K}$分别为结构的质量、阻尼和刚度矩阵；$\boldsymbol{F}$为水动外激励载荷。

本文采用模态叠加法简化上述方程。通过求解特征方程$\boldsymbol{K}\boldsymbol{\varphi}_i=\lambda_i\boldsymbol{M}\boldsymbol{\varphi}_i$，得到系统的特征值$\lambda_1,\cdots,\lambda_N$和特征向量$\boldsymbol{\varphi}_1,\cdots,\boldsymbol{\varphi}_N$，保留对系统响应贡献较大的$n$个低阶振型，组成振型矩阵$\boldsymbol{\Phi}=\left[\boldsymbol{\varphi}_1,\cdots,\boldsymbol{\varphi}_n\right]$。引入广义坐标$\boldsymbol{q}$，且$\boldsymbol{x}=\boldsymbol{\Phi}\boldsymbol{q}$，式(4)可近似为如下$n$阶方程[16]：

$$\bar{\boldsymbol{M}}\ddot{\boldsymbol{q}}+\bar{\boldsymbol{C}}\dot{\boldsymbol{q}}+\bar{\boldsymbol{K}}\boldsymbol{q}=\bar{\boldsymbol{F}} \tag{5}$$

式中：$\bar{\boldsymbol{M}}=\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{M}\boldsymbol{\Phi}$，$\bar{\boldsymbol{C}}=\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{C}\boldsymbol{\Phi}$和$\bar{\boldsymbol{K}}=\boldsymbol{\Phi}^{\mathrm{T}}\boldsymbol{K}\boldsymbol{\Phi}$分别为广义坐标下的质量、阻尼和刚度矩阵；$\bar{\boldsymbol{F}}$为广义外激励载荷。

1.3 流固耦合方法

流固耦合问题描述的是流体与固体之间的相互作用，涉及到流场运动、结构运动及流固交界面的物理量之间的联系。采用松耦合算法在时域内分别求解流体、结构动力学方程，交错时间推进获得耦合系统的响应，其流程如图1所示。

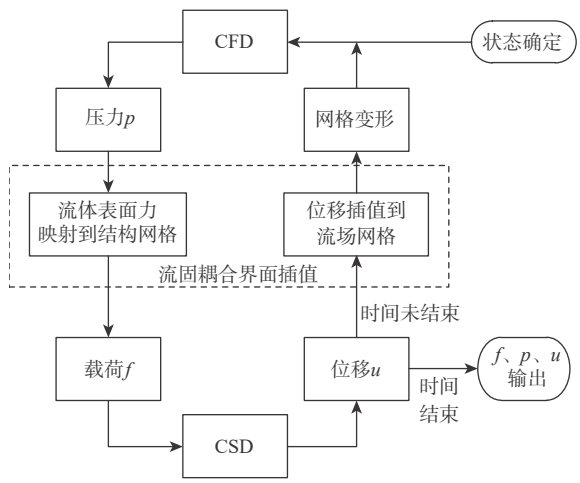


图1 CFD/CSD双向流固耦合迭代流程

Fig. 1 Two-way CFD/CSD fluid-structural coupling iterative process

步骤1 建立结构点与流体表面网格点的插值关系。

步骤2 计算流场获得流体表面力分布。

步骤3 利用虚功原理确定结构力与流体表面力分布之间的关系，并计算结构力。

步骤4 求解结构运动方程，获得刚体运动位移和结构点弹性变形。

步骤5 基于结构点与流体表面网格点的插值关系，更新流场物面网格点坐标及流场域网格。

步骤6 重复步骤2～步骤5完成下个时间步的计算，直至结束。

本文流体计算采用Fluent 16.2，结构模态计算

采用 Abaqus 6.14, 流固耦合界面插值采用径向基函数法[17], 网格变形采用弹簧网格动网格[18]方法。本文分析方法可保证每一个迭代循环内的力平衡、能量守恒和物面光滑连续。

2　流固耦合模型验证

本节分别采用经典的标模算例, 验证泡型计算和流固耦合计算的准确性。

2.1　泡型计算方法验证

本节以 Hrubes 射弹的泡型试验数据, 验证空泡流型算法[19]。Hrubes 射弹的空化器直径为 1.42 mm, 长度为 157.4 mm, 试验时水下航行速度为 970 m/s、航行深度为 4 m, 空化数约为 0.000 29。图 2 为数值模拟泡型与 Hrubes 试验数据[19-20]及经典的 Logvinovich 公式[21]和 Savchenko 公式[22](适用空化数范围分别为 0～0.25 和 0.012～0.57)的泡型对比。可知, 本文对象空化数较小, 数值模拟泡型与 Logvinovich 公式及试验结果更加接近, 趋势合理; 弹尾位置的泡型直径模拟误差小于 3%, 精度较高。综上, 验证了本文泡型计算方法的准确性。

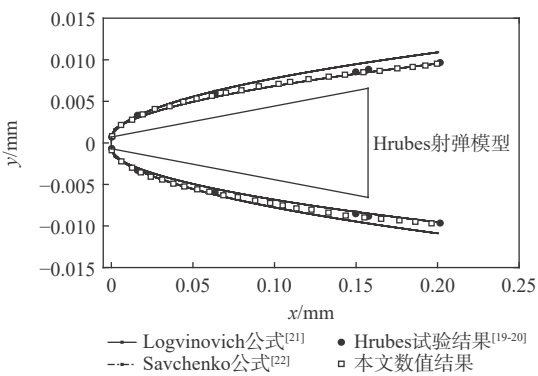


图 2　本文泡型与试验数据及经验公式对比

Fig. 2　Comparison of cavity shape in this paper with experimental data and empirical formula

2.2　流固耦合方法验证

本节以 AGARD 445.6 标模的颤振试验数据, 验证流固耦合算法[23]。采用基于密度的求解器, 梯度计算选择 Green-Gauss 方法, 通量计算基于 Roe-FDS 格式, 湍流模型采用 S-A 模型, 空间离散采用二阶迎风格式。以马赫数为 0.96 为例, 选取前 2 阶模态, 保持马赫数和密度不变, 通过改变来流速度的方法寻找颤振边界。对第 1 阶模态施加大小为 10^{-4} 的扰动, 图 3 给出了仿真速度 V 分别为 300 m/s、309 m/s 和 305 m/s 时, 广义位移的衰减、发散和等幅运动情况。即该马赫数下的颤振速度为 305 m/s, 与试验值 309 m/s 的误差为−1.3%, 两者基本一致。综上, 验证了本文流固耦合方法的准确性。

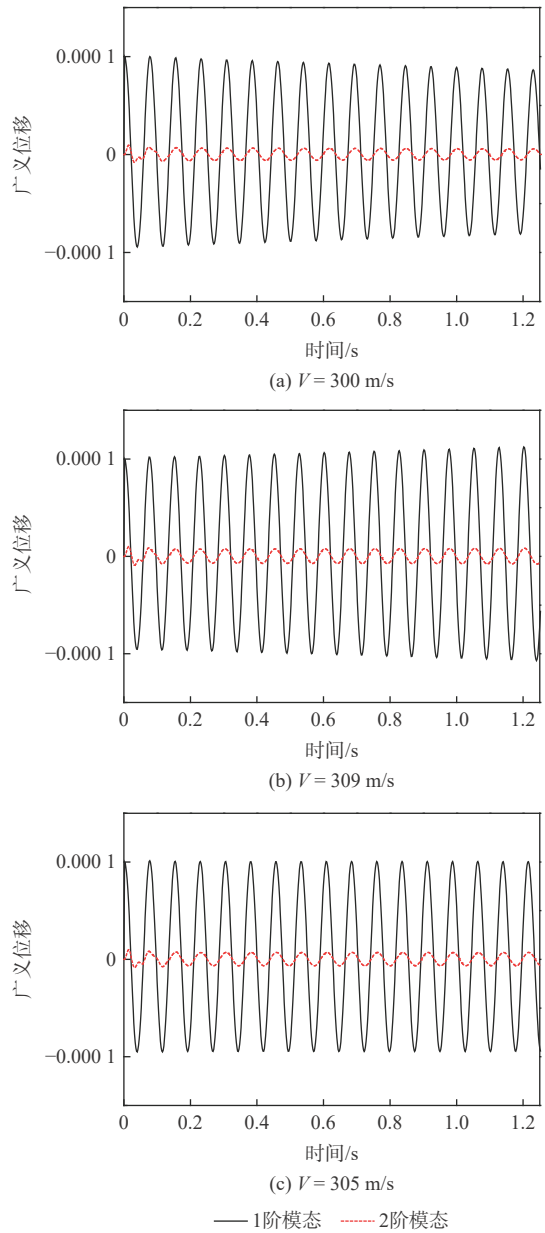


图 3　广义位移时间曲线 (Ma = 0.96)

Fig. 3　Curves of generalized displacement time (Ma = 0.96)

3　超空泡射弹尾拍载荷分析

3.1　研究对象

本文选择大长细比超空泡射弹模型, 开展尾拍双向流固耦合特性研究, 具体尺寸如图 4 所示。射弹外形分为钨合金锥段和钢柱段 2 部分, 具体材料参数如表 1 所示。射弹总质量为 0.086 kg, 质心距前端面 77.4 mm, 转动惯量为 1.82×10^{-4} kg·m^2。射弹

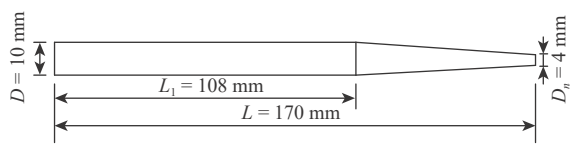


图 4　超空泡射弹模型

Fig. 4　Model of supercavitating projectile

表 1 超空泡射弹材料参数

Table 1 Material parameters of supercavitating projectile

部位	材料	弹性模量/GPa	密度/(kg·m^{-3})	泊松比
锥段	钨合金	312	17 500	0.3
柱段	钢	210	7 850	0.3

总长度 L 为 170 mm, 弹径 D 为 10 mm, 空化器直径 D_n 为 4 mm。

本文坐标原点位于弹体质心, x 轴与射弹轴线重合并指向弹体头部, y 轴位于弹体对称面且垂直弹轴向上, z 轴由右手定则确定。

如图 5 所示, 流场计算采用 20 L×20 L×31 L 的六面体流场域, 流域入口及上下左右边界与弹头中心点距离均为 10 L, 流域出口距弹尾为 20 L, 六面体边界条件均设置为压力出口, 参考压力设置为 101 325 Pa。为提高效率, 减少不必要的网格数量, 整个流体计算域随弹体一起平动和转动。

流体域计算采用六面体结构网格, 首先建立 100 万、180 万、460 万 3 套不同数量的网格, 进行流场网格无关性验证。图 6 给出了典型工况下 3 套网格的泡型对比。可知, 3 套网格的泡型基本一致, 与 460 万网格的空泡直径对比, 100 万网格的最大差异约 3.5%, 180 万网格的最大差异约 1.3%。为兼顾精度和效率, 本文选用 180 万网格进行计算, 网格如图 7 所示。

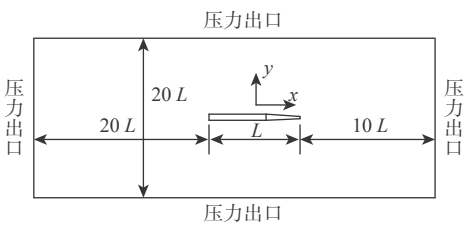


图 5 流体计算域及边界条件示意图

Fig. 5 Schematic diagram of fluid calculation domain and boundary conditions

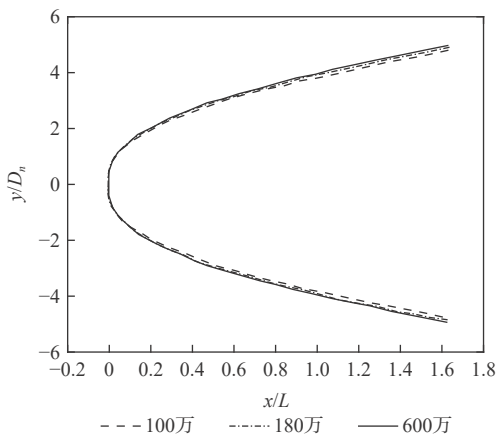


图 6 流场网格无关性分析

Fig. 6 Independent analysis of flow field mesh

固体域计算采用六面体网格, 总网格量 2 万, 网格如图 8 所示。流固耦合计算时, 刚体运动选取平动、升沉和俯仰 3 阶刚体模态, 弹性变形选取弹体 y 向一弯、y 向二弯和 y 向三弯 3 阶弹性模态, 图 9 给出了 3 阶弹性模态的频率及振型。

针对 xOy 平面内水平速度为 1 000 m/s、深度为 4 m 的工况开展基于模态的双向流固耦合计算, 具体为: 在计算初段, 通过用户定义函数 (user-defined function, UDF) 将弹体的升沉和俯仰模态设置为 0, 仅考虑 x 向平动运动, 待泡型稳定后放开弹体的升沉和俯仰模态, 同时施加 4°的初始扰动角度

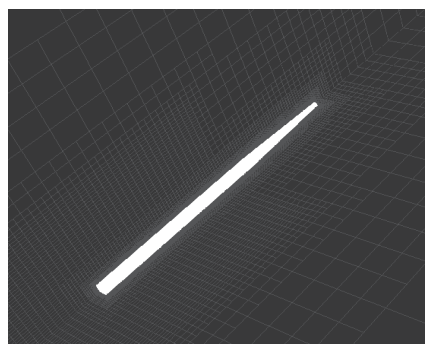


图 7 流场网格模型

Fig. 7 Mesh of flow field

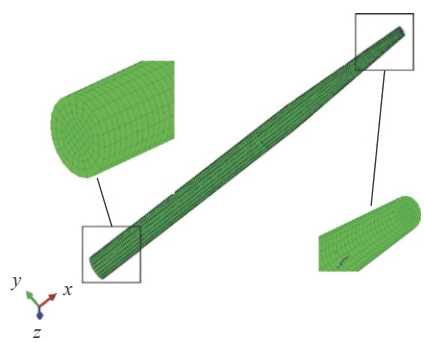


图 8 结构网格模型

Fig. 8 Mesh of structural model

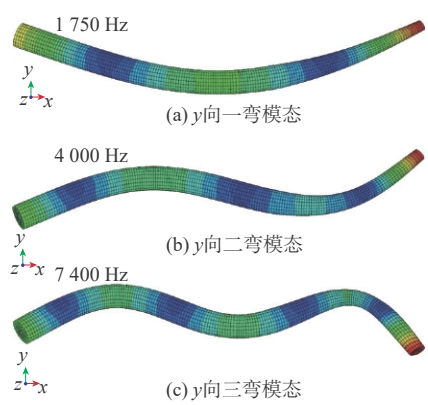


图 9 3 阶弹性模态

Fig. 9 First three order elastic modes

进行尾拍运动仿真。分别针对刚体尾拍(结构仅考虑 3 阶刚体模态)和弹性体尾拍(结构采用 3 阶刚体模态+3 阶弹性模态)2 种情况开展仿真分析。

3.2　泡型及结构变形特性

图 10 为射弹在刚体和弹性体状态, 半个稳态尾拍周期的泡型变化对比(分别对应刚体的 0.012 6～0.013 4 s 和弹性体的 0.010 72～0.011 52 s)。图中深色部分表示液体域, 白色部分表示空化区域。

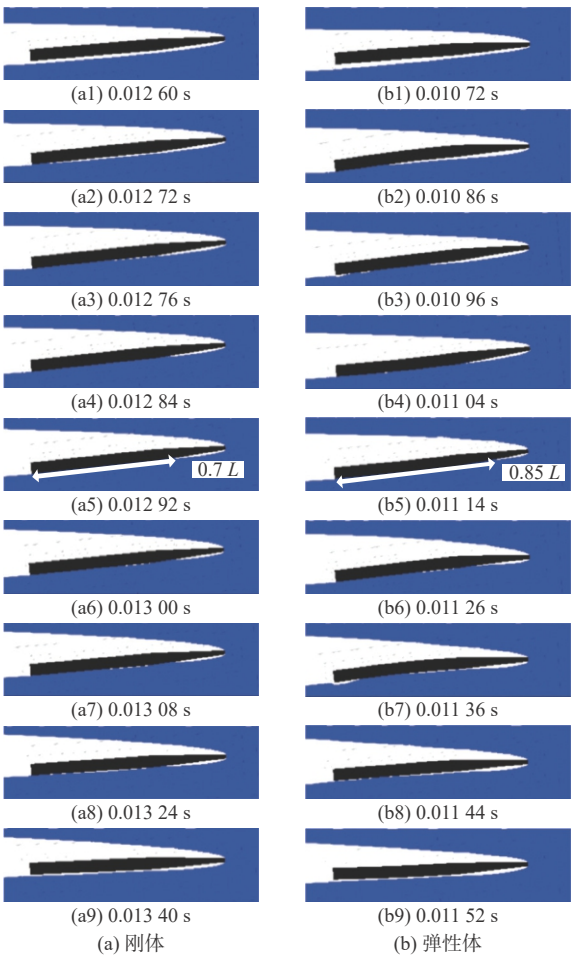


图 10　刚体与弹性体尾拍运动泡型对比

Fig. 10　Comparison of cavity profiles between rigid body and elastic body during tail-slapping

刚体尾拍可以分为如下阶段: 图 10(a1)为头部沾湿阶段, 此阶段弹身均位于空泡内, 弹体存在绕质心的逆时针转动角速度; 图 10(a2)～(a8)为头部和弹身均沾湿阶段, 这一阶段弹身受水动力作用产生绕质心的恢复力矩, 具体过程为: 除头部沾湿外, 弹身尾部先触水沾湿, 且随着逆时针转角不断增加, 沾湿长度由尾部向前不断增加, 在图 10(a5)时刻达到最大沾湿长度 0.7 L, 此后弹体转为顺时针转动, 弹身沾湿长度不断减小, 并最终与空泡完全分离; 图 10(a9)时刻重回头部沾湿阶段, 此时弹体与泡壁完全分离, 射弹转为顺时针向上拍动。

弹性体尾拍与刚体尾拍在图 10(b1)、图 10(b10)头部沾湿阶段泡型类似, 但在图 10(b2)～(b9)的弹身沾湿阶段存在较大差异, 具体为: 弹性体尾拍的图 10(b2)时刻同样弹身尾部先发生沾湿, 在沾湿水动力冲击下, 弹身弹性模态被显著激发, 会产生图 10(b3)时刻所示的弹身尾部与空泡分离、弹身中部“局部沾湿”的现象, 且沾湿面积向两侧不断扩展, 在图 10(b4)时刻发生弹身“二次沾湿”的现象, 并在图 10(b5)时刻达到最大沾湿长度 0.85 L, 随后弹体开始顺时针转动, 在与泡壁发生分离的过程中, 由于弹性效应, 图 10(b6)～(b8)时刻继续出现明显的“局部沾湿”现象。

为研究弹性体尾拍的射弹变形规律, 沿弹长方向设置了 18 个测点, 统计了一个稳定尾拍周期内, 射弹各测点的 y 向变形峰值, 如图 11 所示。可知, 在尾拍过程中, 弹体各测点 y 向变形的正负幅值基本一致, 且分布规律与射弹的弹体 y 向一弯振型相近, 表明该射弹的弹性变形主要由弹性一弯模态主导。射弹头部的变形最大, 幅值达±2 mm, 振幅与头部直径相当, 进而导致弹体在尾拍过程中呈现出较为明显的流固耦合效应。

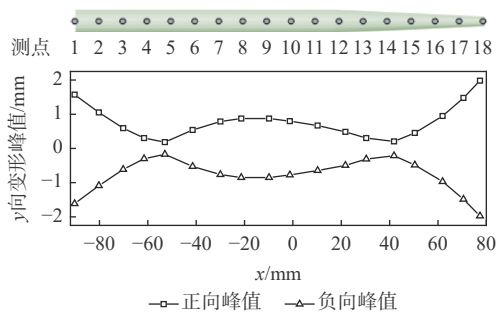


图 11　典型节点 y 向变形的峰值分布

Fig. 11　Distribution of peak deformation of typical nodes

3.3　尾拍流体载荷特性

射弹在刚体和弹性体尾拍过程中的流体载荷及俯仰姿态角对比如图 12 所示。可知, 刚体和弹性体尾拍的水动力及姿态角变化趋势较为一致, 主要分为 2 个阶段: 第 1 个阶段的俯仰角变化范围在±3°以内, 此阶段虽然拍动俯仰角有所变化, 但射弹仅头部沾湿, 弹身始终被空泡包裹, 此时弹体主要受轴向力作用, 法向力和俯仰力矩接近于 0; 第 2 阶段时俯仰角大于 3°, 此阶段射弹与空泡壁面发生碰撞, 随着俯仰角的增大, 弹体的沾湿长度与沾湿面积不断增加, 受到的水动力也不断增加。

刚体尾拍与弹性体尾拍的流体载荷存在 2 个典型差异。首先, 从趋势上, 在单个尾拍周期内, 刚体尾拍的流体载荷仅含单个冲击峰, 弹性体尾拍则

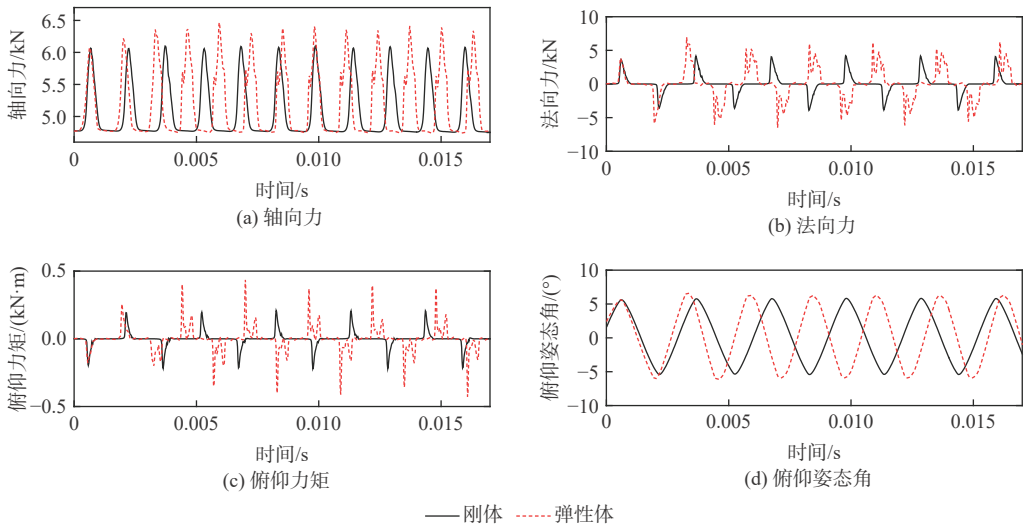


图 12 刚体与弹性体尾拍流体载荷及俯仰角对比

Fig. 12 Comparisons of loads and pitch angle between rigid body and elastic body during tail-slapping

普遍存在多个冲击峰，这是由于弹性体尾拍过程射弹变形显著，诱发了 3.2 节的“二次尾拍”现象所致；其次，从峰值上，如表 2 所示弹性体尾拍的流体载荷较刚体尾拍高 27%～105%，这是由于弹性弹体与空泡强耦合作用下，弹体沾湿长度显著增大所致。此外，在高幅值高频载荷作用下，弹性体的尾拍姿态角较刚体增大 13%，尾拍频率增大 20%，流固耦合效应对尾拍弹道也产生了较强影响。

表 2 刚体与弹性体尾拍流体载荷特性对比

Table 2 Comparison of load characteristics between rigid body and elastic body during tail-slapping

类型	尾拍引起的轴向力增量峰值/kN	尾拍法向力峰值/kN	俯仰力矩峰值/(kN·m)	尾拍姿态角峰值/(°)	尾拍频率/Hz
刚体	1.35	4.21	0.21	5.83	329
弹性体	1.72	6.95	0.43	6.59	394
变化量	0.37	2.74	0.22	0.76	65

注：尾拍引起的轴向力增量峰值、尾拍法向力峰值、俯仰力矩峰值、尾拍姿态角峰值、尾拍频率的变化百分比分别为27%，65%，105%，13%，20%。

4 结 论

本文通过耦合 CFD 求解器和 CSD 求解器，建立了超空泡射弹尾拍过程流固耦合数值模型。计算对比了 1 000 m/s 速度下超空泡射弹刚体尾拍和弹性体尾拍在泡型演化、结构变形及流体载荷特性的差异，结论如下：

1）弹性体尾拍过程，弹体结构与空泡会产生较强的流固耦合效应，射弹尾部出现了特殊的“二次拍击”和弹身“局部沾湿”现象，流体载荷也呈现明显的多峰值现象。

2）弹性体尾拍过程，射弹的弹性变形主要由弹性一弯模态主导，最大变形与头部直径相当，变形显著，变形引起弹体沾湿面积由 0.7 L 增大至 0.85 L，导致流体载荷增大 27%～105%。

3）弹性体尾拍过程，在高幅值高频载荷激励下，弹体的尾拍姿态角增大 13%，尾拍频率增大 20%，流固耦合效应对尾拍弹道也产生了较强影响。

参考文献（References）

[1] 潘森森, 彭晓星. 空化机理[M]. 北京: 国防工业出版社, 2013.
PAN S S, PENG X X. Physical mechanism of cavitation[M]. Beijing: National Defense Industry Press, 2013(in Chinese).

[2] 熊天红, 易文俊. 高速射弹超空泡减阻试验研究与数值模拟分析[J]. 工程力学, 2009, 26(8): 174-178.
XIONG T H, YI W J. Experimental research and numerical simulation of supercavity drag reduction of a high speed projectile[J]. Engineering Mechanics, 2009, 26(8): 174-178(in Chinese).

[3] 曹伟, 魏英杰, 王聪, 等. 超空泡技术现状、问题与应用[J]. 力学进展, 2006, 36(4): 571-579.
CAO W, WEI Y J, WANG C, et al. Current status, problems and applications of supercavitation technology[J]. Advances in Mechanics, 2006, 36(4): 571-579(in Chinese).

[4] 杨莉, 张庆明. 超空泡技术的应用现状和发展趋势[J]. 战术导弹技术, 2006(5): 6-10.
YANG L, ZHANG Q M. Current application and perspectives on supercavitation technology research[J]. Tactical Missile Technology, 2006(5): 6-10(in Chinese).

[5] 黄闯. 跨声速超空泡射弹的弹道特性研究[D]. 西安: 西北工业大学, 2017.
HUANG C. Research of trajectory characteristics of supersonic-supercavitating projectiles[D]. Xi'an: Northwestern Polytechnical University, 2017(in Chinese).

[6] 魏英杰, 何乾坤, 王聪, 等. 超空泡射弹尾拍问题研究进展[J]. 舰船科学技术, 2013, 35(1): 7-15.
WEI Y J, HE Q K, WANG C, et al. Review of study on the tail-slap problems of supercavitating projectile[J]. Ship Science and Technology, 2013, 35(1): 7-15(in Chinese).

[7] 郝常乐, 党建军, 陈长盛, 等. 基于双向流固耦合的超空泡射弹入水研究[J]. 力学学报, 2022, 54(3): 678-687.
HAO C L, DANG J J, CHEN C S, et al. Numerical study on water entry process of supercavitating projectile by considering bidirectional fluid structure interaction effect[J]. Chinese Journal of Theoretical and Applied Mechanics, 2022, 54(3): 678-687(in Chinese).

[8] CHEN C, ZHANG A M, CHEN J Q, et al. SPH simulations of water entry problems using an improved boundary treatment[J]. Ocean Engineering, 2021, 238: 109679.

[9] LI Y, SUN T Z, ZONG Z, et al. Dynamic crushing of a dedicated buffer during the high-speed vertical water entry process[J]. Ocean Engineering, 2021, 236: 109526.

[10] 张劲生. 超空泡运动体尾拍冲击振动特性的研究[D]. 哈尔滨: 哈尔滨工业大学, 2010.
ZHANG J S. Research on impact and vibration characteristic of supercavitating high-speed bodies with tail-slaps[D]. Harbin: Harbin Institute of Technology, 2010(in Chinese).

[11] 梁景奇, 王瑞, 徐保成, 等. 攻角对高速射弹入水过程影响研究[J]. 兵器装备工程学报, 2020, 41(7): 23-28.
LIANG J Q, WANG R, XU B C, et al. Research on influence of angle of attack on process of high-speed water-entry projectile[J]. Journal of Ordnance Equipment Engineering, 2020, 41(7): 23-28(in Chinese).

[12] 王强, 侯宇, 金子涵, 等. 基于双向流固耦合的柔性扑翼变形数值研究[J]. 飞行力学, 2021, 39(3): 27-33.
WANG Q, HOU Y, JIN Z H, et al. Numerical study on the deformation of flexible flapping wing based on bidirectional fluid-solid coupling[J]. Flight Dynamics, 2021, 39(3): 27-33(in Chinese).

[13] 秦杨, 易文俊, 管军. 超空泡射弹高速倾斜入水的空化流动数值模拟[J]. 兵器装备工程学报, 2019, 40(7): 99-104.
QIN Y, YI W J, GUAN J. Numerical simulation for cavitation flow of high-speed rotating body oblique entrying into water[J]. Journal of Ordnance Equipment Engineering, 2019, 40(7): 99-104(in Chinese).

[14] MENTER F R. Two-equation eddy-viscosity turbulence models for engineering applications[J]. AIAA Journal, 1994, 32(8): 1598-1605.

[15] 谢亮, 徐敏, 李杰, 等. 基于CFD/CSD耦合的颤振与动载荷分析方法[J]. 振动与冲击, 2012, 31(3): 106-110.
XIE L, XU M, LI J, et al. Flutter and dynamic analysis based on CFD/CSD coupling method[J]. Journal of Vibration and Shock, 2012, 31(3): 106-110(in Chinese).

[16] 许云涛, 吴志刚, 杨超. 地面颤振模拟试验中的非定常气动力模拟[J]. 航空学报, 2012, 33(11): 1947-1957.
XU Y T, WU Z G, YANG C. Simulation of the unsteady aerodynamic forces for ground flutter simulation test[J]. Acta Aeronautica et Astronautica Sinica, 2012, 33(11): 1947-1957(in Chinese).

[17] 林言中, 陈兵, 徐旭. 基于径向基函数插值的气动弹性计算方法[J]. 北京航空航天大学学报, 2014, 40(7): 953-958.
LIN Y Z, CHEN B, XU X. Numerical method of aeroelasticity based on radial basis function interpolation[J]. Journal of Beijing University of Aeronautics and Astronautics, 2014, 40(7): 953-958(in Chinese).

[18] 周璇, 李水乡, 陈斌. 非结构动网格生成的弹簧-插值联合方法[J]. 航空学报, 2010, 31(7): 1389-1395.
ZHOU X, LI S X, CHEN B. Spring-interpolation approach for generating unstructured dynamic meshes[J]. Acta Aeronautica et Astronautica Sinica, 2010, 31(7): 1389-1395(in Chinese).

[19] HRUBES J D. High-speed imaging of supercavitating underwater projectiles[J]. Experiments in Fluids, 2001, 30(1): 57-64.

[20] 陈伟善. 高速超空泡射弹尾拍运动特性数值研究[D]. 南京: 南京理工大学, 2020.
CHEN W S. Numerical study on the motion characteristics of the tail-slap of the high-speed supercavitating projectile[D]. Nanjing: Nanjing University of Science and Technology, 2020(in Chinese).

[21] 罗格维诺维奇. 自由边界流动的水动力学[M]. 施红辉译. 上海: 上海交通大学出版社, 2012: 97-126.
LOGVINOVICH G V. Hydrodynamics of free-boundary flow[M]. SHI H H, translated. Shanghai: Shanghai Jiao Tong University Press, 2012: 97-126(in Chinese).

[22] SAVCHENKO Y N. Supercavitation-problems and perspectives[C]// 4th International Symposium on Cavitation, 2001: 1-8.

[23] JK YATES E C. AGARD standard aeroelastic configuration for dynamic response I-wing-445.6: NASA-TM-100492[R]. Washington, D. C. :NASE, 1988.

Study on tail-slap load characteristics of high-speed projectile based on CFD/CSD coupling

XU Yuntao[1, 2], TAN Dalin[2], YANG Chao[1, *]

(1. School of Aeronautic Science and Engineering, Beihang University, Beijing 100191, China;
2. Beijing Institute of Mechanical and Electrical Engineering, Beijing 100074, China)

Abstract: A calculating approach based on the two-way analysis of computational fluid dynamics (CFD)/computational structure dynamics (CSD) coupling is devised to increase the predictability of the projectile's tail-slap load. The main governing equation of fluid calculation adopts the Navier-Stokes equation with SST k-ω turbulent flow model and Schnerr-Sauer cavitation model, structural calculations use simplified structural dynamic equations based on the modal superposition method, interpolation of fluid-structural coupling interface using radial basis function method, mesh deformation using spring stretching method. Both the fluid-structural coupling approach and the cavity shape computation method were independently verified. Further, the differences in cavity shape, structural deformation, and characteristics of tail-slap loads between the rigid body and the elastic body of the projectile at a speed of 1 000 m/s were calculated and compared. Simulation results show that during the tail-slapping process of the elastic body, the cavity shape of the projectile will produce special phenomena such as 'secondary slap', 'partial wetting' and a wetting area increased. Structural deformation is dominated by the first-bending mode. The larger deformation causes the dynamic load to increase by 27%～105%, the amplitude of the pitch angle to increase by 13%, and the tail-slapping frequency to increase by 20%. The fluid-structural coupling effect has a strong influence on the cavity shape, loads, and ballistic of tail-slapping.

Keywords: projectile; CFD/CSD coupling; tail-slap; load; rigid body; elastic body

Received: 2022-11-24; **Accepted**: 2023-02-26; **Published Online**: 2023-04-03 15:44
URL: kns.cnki.net/kcms/detail/11.2625.V.20230403.1102.001.html
* **Corresponding author**. E-mail: yangchao@buaa.edu.cn

《北京航空航天大学学报》征稿简则

《北京航空航天大学学报》是北京航空航天大学主办的以航空航天科学技术为特色的综合性自然科学学术期刊(月刊)。本刊以反映航空航天领域研究成果与动态、促进学术交流、培养科技人才和推动科技成果向社会生产力转化为办刊宗旨。本刊为中国自然科学技术核心期刊,并被 Ei Compendex 等国内外权威文献检索系统收录。本刊向国内外公开发行,为进一步提高办刊质量和所刊出文章的学术水平,特制定本简则。

1 论文作者及内容

1.1 本刊面向海内外所有学者。

1.2 主要刊载与航空航天科学技术有关的材料科学及工程、飞行器设计与制造、宇航科学与工程、信息与电子技术、控制技术和自动化工程、流体力学和动力工程、计算机科学及应用技术、可靠性工程与失效分析等领域的研究文章。航空航天科学技术民用方面以及具有航空航天工程背景的应用数学、应用物理、应用力学和工程管理等方面的文章也在本刊优先考虑之列。

2 来稿要求

2.1 论文应具有创新性、科学性、学术性和可读性。

2.2 论文为原创作品,尚未公开发表过,并且不涉及泄密问题。若发生侵权或泄密问题,一切责任由作者承担。

2.3 主题明确,数据可靠,图表清晰,逻辑严谨,文字精练,公式符号标准规范,标点符号正确。

2.4 文稿撰写顺序:中文题名(一般不超过 20 个汉字),作者中文姓名、单位、所在城市、邮政编码(作者顺序始终不得更改),中文摘要(包括目的、方法、结果及结论),中文关键词(5 ~ 8 个),中图分类号,引言,正文,参考文献,英文题名,作者英文姓名、单位、所在城市、邮政编码、国别,英文摘要,英文关键词。首、末页下角注明基金项目名称及编号,作者信息。

2.5 作者请登录本刊网页进行在线投稿。

3 稿件的审核、录用与版权

3.1 来稿须经专家两审和主编、编委讨论后决定刊用与否。

3.2 若来稿经审查后认定不宜在本刊发表,将及时告知作者。如果在投稿满 3 个月后仍未收到本刊任何通知,作者有权改投其他刊。在此之前,请勿一稿多投,否则一切后果自负。

3.3 来稿一经作者签字并在本刊刊出,即表明所有作者都已经认可其版权转至本刊编辑部。本刊在与国内外文献数据库或检索系统进行交流及合作时,不再征询作者意见。

邮寄地址:100191 北京市海淀区学院路 37 号 《北京航空航天大学学报》编辑部

办公地点:北京航空航天大学办公楼 405,407 房间

电　　话:(010)82316698,82316981,82315499,82317448

E-mail: jbuaa@ buaa. edu. cn

http://bhxb. buaa. edu. cn

http://www. buaa. edu. cn

《北京航空航天大学学报》
第五届编辑委员会

主　任（主　编）：赵沁平

（以下按姓氏笔画为序）

副主任（副主编）：丁希仑　王少萍　孙志梅　李秋实　李焕喜　杨嘉陵
苗俊刚　相　艳　徐立军　钱德沛　曹晋滨

编　委：马殿富　王　琪　王　聪　邓小燕　王青云　王荣明　刘　宇
刘　红　江　洁　刘　强　闫　鹏　朱天乐　刘铁钢　齐铂金
陈万春　邹正平　苏东林　杨世春　沈成平　邱志平　宋知人
杨树斌　张晓林　杨晓奕　杨继萍　李惠峰　吴新开　张瑞丰
杨照华　宋凝芳　周　锐　林宇震　林贵平　战　强　姚仰平
胡庆雷　赵秋红　段海滨　赵巍胜　席　平　郭　宏　徐　洁
徐世杰　郭洪波　康　锐　翟　锦　熊华钢

北京航空航天大学学报

Beijing Hangkong Hangtian Daxue Xuebao

（原《北京航空学院学报》）

（月刊　1956年创刊）

第49卷 第9期　2023年9月

JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS (JBUAA)

(Monthly, Started in 1956)

Vol.49 No.9　September 2023

主管单位　中华人民共和国工业和信息化部
主办单位　北京航空航天大学
主　　编　赵沁平
编辑出版　《北京航空航天大学学报》编辑部
邮　　编　100191
地　　址　北京市海淀区学院路37号
印　　刷　北京科信印刷有限公司
发　　行　《北京航空航天大学学报》编辑部
发行范围　国内外发行
联系电话　(010) 82316698 82316981
82315499 82317448
电子信箱　jbuaa@ buaa.edu.cn

Administrated by　Ministry of Industry and Information Technology of the People's Republic of China
Sponsored by　Beijing University of Aeronautics and Astronautics (BUAA)
(Beijing 100191, P. R. China)
Chief Editor　ZHAO Qinping
Edited and Published by　Editorial Board of JBUAA
Printed by　Beijing Kexin Printing Co., Ltd.
Distributed by　Editorial Board of JBUAA
Telephone　(010) 82316698 82316981
82315499 82317448
E-mail　jbuaa@buaa.edu.cn
http://bhxb.buaa.edu.cn

中国标准连续出版物号：ISSN 1001-5965 / CN 11-2625/V　　国内定价：50.00元／期

ISSN 1001-5965