



ISSN 1001-5965

CN 11-2625/V

# 北京航空航天大学 学报

JOURNAL OF BEIJING UNIVERSITY OF  
AERONAUTICS AND ASTRONAUTICS



**2023-03**

Vol.49 No.3

目 次

基于 SABA 优化的 Volterra 级数空战目标机动轨迹预测 ..... 李战武, 彭明毓, 高春庆, 杨爱武, 徐安, 方诚喆 (503)

基于倒置残差注意力的无人机航拍图像小目标检测 ..... 刘树东, 刘业辉, 孙叶美, 李懿霏, 王娇 (514)

单目标概率约束规划的微种群免疫优化算法 ..... 李静, 张仁崇, 潘春燕, 杨凯 (525)

基于 MSST 及 HOG 特征提取的雷达辐射源信号识别 ..... 全大英, 唐泽雨, 陈赞, 楼维中, 汪晓锋, 章东平 (538)

超低轨卫星气动舵机辅助姿态控制方法设计 ..... 王涛, 焦洪臣, 刘杰, 陈乐宇, 张迎春 (548)

航天遥感器用泵驱两相流体回路热真空试验研究 ..... 孟庆亮, 赵振明, 陈祥贵, 朱许 (559)

基于修正似然滤波的无人机编队相对导航方法 ..... 苏炳志, 王磊, 张红伟, 汪海涵, 石璐璐 (569)

面向基准约束与余量约束的配准算法 ..... 朱燊, 肖世宏, 陈志同 (580)

直反信号协同的 GNSS-R BSAR 距离多普勒成像算法 ..... 吴世玉, 杨东凯, 王峰, 苗铎 (588)

涡轴发动机时延鲁棒串级 PI 控制器设计 ..... 陈义峰, 郭迎清, 毛皓天 (597)

多种群合作学习的多模态多目标路径规划算法 ..... 赵萌, 路辉, 王诗琪, 杨思琦, 王赞 (606)

基于线性约束最小方差的稳健波束形成算法 ..... 吕岩, 曹菲 (617)

火箭整流罩锥壳夹层结构不确定性轻量化设计 ..... 董欣心, 岳振江, 王志, 刘莉 (625)

基于 Logistic 回归麻雀算法的图像分割 ..... 陈刚, 林东, 陈飞, 陈祥宇 (636)

基于电感特征的开关磁阻电机电流斩波控制策略 ..... 陈越, 蒋启龙, 王金锁, 姚卫丰 (647)

GNSS 外辐射源雷达低慢小目标探测概率 ..... 苗铎, 杨东凯, 许志超, 王峰, 吴世玉 (657)

大行程柔性微定位平台的伴生转动分析 ..... 孟刚, 黄河, 吴伟官, 居勇健, 曹毅 (665)

基于 CNN 机翼气动系数预测 ..... 吕召阳, 聂雪媛, 赵奥博 (674)

自适应变异麻雀搜索优化算法 ..... 唐延强, 李成海, 宋亚飞, 陈晨, 曹波 (681)

基于 M-AFSA 的 MPRM 逻辑电路面积优化 ..... 邵艺璇, 何振学, 周宇豪, 霍志胜, 肖利民, 王翔 (693)

基于动态特征选择的遥感图像目标检测算法 ..... 陈超, 赵巍 (702)

改进 YOLOv4 的表面缺陷检测算法 ..... 李彬, 汪诚, 丁相玉, 巨海娟, 郭振平, 李卓越 (710)

一种基于组合赋权法的小波去噪质量评价方法 ..... 李晋斐, 赵冬青, 王栋民, 蔡聪聪, 贾晓雪, 张乐添 (718)

柔性涡流阵列传感器孔边裂纹监测技术 ..... 樊祥洪, 缙百勇, 陈涛, 何宇廷, 崔荣洪, 喻健 (726)

基于循环神经网络的 SPMA 协议信道状态智能检测改进算法 ..... 张彦晖, 吕娜, 缪竞成, 高旗, 王翔, 陈卓 (735)

# JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS

Vol. 49 No. 3 ( Sum 361 ) March 2023

## CONTENTS

Air combat maneuver trajectory prediction of target based on Volterra series optimized by SABA algorithm	
..... LI Zhanwu, PENG Mingyu, GAO Chunqing, YANG Aiwu, XU An, FANG Chengzhe	( 503 )
Small object detection in UAV aerial images based on inverted residual attention	
..... LIU Shudong, LIU Yehui, SUN Yemei, LI Yifei, WANG Jiao	( 514 )
Micro immune optimization algorithm for single objective probabilistic constrained programming	
..... LI Jing, ZHANG Renchong, PAN Chunyan, YANG Kai	( 525 )
Radar emitter signal recognition based on MSST and HOG feature extraction	
..... QUAN Daying, TANG Zeyu, CHEN Yun, LOU Weizhong, WANG Xiaofeng, ZHANG Dongping	( 538 )
Design of attitude control method for ultra-low-orbit satellite with pneumatic steering gear	
..... WANG Tao, JIAO Hongchen, LIU Jie, CHEN Leyu, ZHANG Yingchun	( 548 )
Thermal vacuum test study of mechanically pumped two-phase loop for space remote sensor	
..... MENG Qingliang, ZHAO Zhenming, CHEN Xianggui, ZHU Xu	( 559 )
Relative navigation method based on modified likelihood filtering for unmanned aerial vehicle formation	
..... SU Bingzhi, WANG Lei, ZHANG Hongwei, WANG Haihan, SHI Lulu	( 569 )
A registration algorithm with datum constraints and allowance constraints	
..... ZHU Yu, XIAO Shihong, CHEN Zhitong	( 580 )
GNSS-R BSAR range-Doppler imaging algorithm based on synchronization of direct and echo signal	
..... WU Shiyu, YANG Dongkai, WANG Feng, MIAO Duo	( 588 )
Design of time-delay robust cascade PI controller for turboshaft engine	
..... CHEN Yifeng, GUO Yingqing, MAO Haotian	( 597 )
A multimodal multi-objective path planning algorithm based on multi-swarm cooperative learning	
..... ZHAO Meng, LU Hui, WANG Shiqi, YANG Siyi, WANG Zan	( 606 )
Robust beamforming based on linear constraint minimum variance algorithm	
..... LYU Yan, CAO Fei	( 617 )
Uncertainty lightweight design of sandwich structure of rocket fairing cone	
..... DONG Xinxin, YUE Zhenjiang, WANG Zhi, LIU Li	( 625 )
Image segmentation based on Logistic regression sparrow algorithm	
..... CHEN Gang, LIN Dong, CHEN Fei, CHEN Xiangyu	( 636 )
Current chopping control strategy of switched reluctance motor based on inductance characteristics	
..... CHEN Yue, JIANG Qilong, WANG Jinsuo, YAO Weifeng	( 647 )
Low-altitude, slow speed and small target detection probability of passive radar based on GNSS signals	
..... MIAO Duo, YANG Dongkai, XU Zhichao, WANG Feng, WU Shiyu	( 657 )
Parasitic rotation of large stroke compliant micro-positioning platform	
..... MENG Gang, HUANG He, WU Weiguan, JU Yongjian, CAO Yi	( 665 )
Prediction of wing aerodynamic coefficient based on CNN	
..... LYU Zhaoyang, NIE Xueyuan, ZHAO Aobo	( 674 )
Adaptive mutation sparrow search optimization algorithm	
..... TANG Yanqiang, LI Chenghai, SONG Yafei, CHEN Chen, CAO Bo	( 681 )
Area optimization of MPRM circuits based on M-AFSA	
..... SHAO Yixuan, HE Zhenxue, ZHOU Yuhao, HUO Zhisheng, XIAO Limin, WANG Xiang	( 693 )
Remote sensing target detection based on dynamic feature selection	
..... CHEN Chao, ZHAO Wei	( 702 )
Surface defect detection algorithm based on improved YOLOv4	
..... LI Bin, WANG Cheng, DING Xiangyu, JU Haijuan, GUO Zhenping, LI Zhuoyue	( 710 )
A quality evaluation method for wavelet denoising based on combinatorial weighting method	
..... LI Jinfei, ZHAO Dongqing, WANG Dongmin, CAI Congcong, JIA Xiaoxue, ZHANG Letian	( 718 )
Hole edge crack monitoring technology of flexible eddy current array sensor	
..... FAN Xianghong, GOU Baiyong, CHEN Tao, HE Yuting, CUI Ronghong, YU Jian	( 726 )
Improved intelligent detection algorithm for SPMA protocol channel state based on recurrent neural network	
..... ZHANG Yanhui, LYU Na, MIAO Jingcheng, GAO Qi, WANG Xiang, CHEN Zhuo	( 735 )

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0287

# 基于 SABA 优化的 Volterra 级数空战 目标机动轨迹预测

李战武<sup>1,\*</sup>, 彭明毓<sup>2</sup>, 高春庆<sup>3</sup>, 杨爱武<sup>2</sup>, 徐安<sup>1</sup>, 方诚喆<sup>2</sup>

(1. 空军工程大学 航空工程学院, 西安 710038; 2. 空军工程大学 研究生院, 西安 710038;

3. 中国人民解放军 94582 部队, 信阳 464194)

**摘 要:** 目标机动轨迹预测是空战态势感知和目标威胁评估的重要前提。针对传统目标机动轨迹预测模型复杂度大、预测精度低等问题, 通过分析并结合目标机动轨迹时序数据所具备的混沌特性, 引入 Volterra 泛函级数模型进行目标机动轨迹预测。为解决 Volterra 泛函级数模型中存在高阶核函数难以求解的问题, 利用变异机制和自适应步长控制机制改进蝙蝠算法的寻优能力, 进而构建了一种基于自适应蝙蝠算法 (SABA) 优化的 Volterra 泛函级数目标机动轨迹预测模型, 并利用优化后不同阶数的 Volterra 泛函级数模型对目标未来机动轨迹进行预测。仿真实验中, 通过与其他优化算法改进的 Volterra 泛函级数模型的预测精度对比, 验证了所提预测模型的可行性, 同时也说明了二阶 Volterra 泛函级数模型更加适用于目标机动轨迹预测。

**关键词:** 轨迹预测; Volterra 泛函级数模型; 核参数优化; 自适应蝙蝠算法; 截断阶数

**中图分类号:** V19

**文献标志码:** A

**文章编号:** 1001-5965(2023)03-0503-11

近年来, 随着空战进程的加快, 空战目标的状态预测需求随之愈加迫切, 因而信息化战场下, 如何通过海量数据进行分析已成为一个热点问题。在空战对抗过程中, 飞行员可通过海量的目标状态数据进行挖掘分析, 挖掘出目标状态信息中利于作战决策的价值信息, 从而及时、有效地获取目标潜在的作战规律, 客观地分析、总结敌方的作战方式及特点, 进而实现对战场态势的实时感知与获得优势占位<sup>[1]</sup>。因此, 通过空战目标状态数据进行目标机动预测, 对飞行员作战决策及获取先敌优势具有重要意义。

目前, 目标机动轨迹预测主要有 2 类途径。其中, 一类的主要思想是利用卡尔曼滤波算法的传统预测思想进行轨迹预测。该方法具备优异的估计能力, 对于运动状态变化大且随机性高、运动方式多样的目标轨迹有较强的适用性, 因而能够实现目

标实时的状态估计与预测, 并适用于有限维度的线性和非线性目标时空运动轨迹。例如, 姜佰辰等<sup>[2]</sup>针对目标运动模式变化多及机动幅度较大的情况, 提出了一种基于多项式卡尔曼滤波的运动轨迹预测算法; 赵帅兵等<sup>[3]</sup>针对历史位置信息存在缺失的情况, 提出了一种具有系统噪声估计的改进卡尔曼滤波算法, 对目标机动轨迹进行预测; 乔少杰等<sup>[4]</sup>针对传统轨迹预测算法精度和实时性不足, 提出了一种基于卡尔曼滤波的动态轨迹预测算法; 翟岱亮等<sup>[5]</sup>针对高阶目标运动模型存在的高度非线性、数据处理难度大及预测精度低等问题, 提出了一种改进的交互多模型轨迹预测算法。上述预测算法都只适用于目标运动特性相对简单的轨迹预测问题, 但是在空战过程中, 目标的运动往往是高度复杂的时序过程, 传统的轨迹预测算法不能准确地学习目标的机动特性, 同时也存在建模复杂度高和算法适应性

收稿日期: 2021-06-01; 录用日期: 2021-09-03; 网络出版时间: 2021-09-13 10:29

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210913.0933.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210913.0933.001.html)

\*通信作者. E-mail: [afeulzw@189.cn](mailto:afeulzw@189.cn)

**引用格式:** 李战武, 彭明毓, 高春庆, 等. 基于 SABA 优化的 Volterra 级数空战目标机动轨迹预测 [J]. 北京航空航天大学学报, 2023, 49 (3): 503-513.  
LI Z W, PENG M Y, GAO C Q, et al. Air combat maneuver trajectory prediction of target based on Volterra series optimized by SABA algorithm [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 503-513 (in Chinese).

差的问题,并且算法的预测精度难以满足空战对抗需求。

另一类方法主要以智能算法为核心,结合大数据建立目标机动轨迹预测模型。例如,杨彬和贺正洪<sup>[6]</sup>针对广义回归神经网络(generalized regression neural network, GRNN)模型对非线性数据有较好的预测性能,且容错性及鲁棒性较好,将该预测模型运用于高超声速飞行器的轨迹预测当中;针对BP神经网络存在对初值敏感且全局搜索能力较差的问题,谭伟<sup>[7]</sup>和甘旭升<sup>[8]</sup>等分别利用遗传算法(genetic algorithm, GA)、粒子群算法对预测模型参数进行寻优此外,目标机动轨迹预测的本质是对时间序列数据进行预测,因而具有高度非线性和时变性的特点<sup>[9]</sup>。理论和实践表明,Volterra泛函级数模型对于非线性系统可以很好的表征,同时对时间序列可以进行精确预测。然而,高阶Volterra泛函级数模型核函数的求解是其应用瓶颈。目前,Volterra泛函级数参数辨识方法有传统的最小二乘法<sup>[10-13]</sup>及智能启发式优化算法<sup>[14-18]</sup>。但是,最小二乘法在辨识参数时必须满足目标函数是连续可导的条件,同时搜索时利用梯度信息而容易陷入局部极值。因此,可以采用智能算法克服传统辨识算法的缺点,但传统的智能算法在结构和优化的实时性方面仍存在很多的不足。

蝙蝠算法(bat algorithm, BA)是受蝙蝠回声定位行为启发而提出的一种新型搜索优化算法<sup>[19]</sup>,其模型较为简单,且算法参数少、普适性强。因此,本文将蝙蝠算法引入到Volterra泛函级数模型的优化中,同时为解决基本蝙蝠算法中存在的易陷入局部最优和求解精度低的问题<sup>[20-21]</sup>,运用变异机制和自适应步长控制机制改进蝙蝠算法,并将其用于优化Volterra泛函级数模型,构建一种基于改进蝙蝠算法优化的Volterra泛函级数目标机动轨迹预测模型。通过仿真实验证明,该模型的预测性能优于基于基本蝙蝠算法、粒子群算法、蚁群(ant colony optimization, ACO)算法优化的Volterra泛函级数模型和反向传播神经网络(back-propagation neural network, BPNN)预测模型,同时也验证了该预测方法的有效性。

## 1 非线性Volterra泛函级数模型

在非线性系统中,Volterra泛函级数模型常常用来表征输入 $u(t)$ 到输出 $y(t)$ 的映射,其映射关系可以表述为<sup>[22]</sup>

$$y(t) = \sum_{n=1}^{\infty} y_n(t) = y_1(t) + y_2(t) + \dots + y_n(t) + \dots \quad (1)$$

式中:

$$\begin{cases} y_1(t) = \int_{-\infty}^{\infty} h_1(\tau_1)u(t-\tau_1)d\tau_1 \\ y_2(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_2(\tau_1, \tau_2)u(t-\tau_1)u(t-\tau_2)d\tau_1 d\tau_2 \\ \vdots \\ y_n(t) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} h_n(\tau_1, \tau_2, \dots, \tau_n) \prod_{i=1}^n u(t-\tau_i) d\tau_1 d\tau_2 \dots d\tau_n \end{cases} \quad (2)$$

其中: $h_n(\tau_1, \tau_2, \dots, \tau_n)$ 为非线性Volterra泛函级数模型的核函数。

在采用Volterra泛函级数模型对时间序列进行预测时,需先对式(1)离散化处理,得到经过离散化之后的级数模型为

$$y(k) = \sum_{n=1}^N y_n(k) + e(k) \quad k = 1, 2, 3, \dots \quad (3)$$

式中: $k$ 为Volterra泛函级数的阶数; $N$ 为最高阶数; $e(k)$ 为级数模型的截断误差; $y_n(k)$ 可以表述为

$$y_n(k) = \sum_{m_1=0}^{M_1-1} \dots \sum_{m_n=0}^{M_n-1} h_n(m_1, m_2, \dots, m_n) \cdot \prod_{i=1}^n u(k-m_i) \quad (4)$$

其中: $h_n(m_1, m_2, \dots, m_n)$ 为Volterra泛函级数的核函数; $m_i$ 为级数模型的记忆长度。

在采用Volterra泛函级数模型进行预测时,采用无穷形式将会导致其无法精确仿真计算。因此,通常对其进行截断处理,即确定Volterra泛函级数的阶数 $k$ 及模型的记忆长度 $m_i$ 。一般情况下,采用3阶截断形式即可近似描述大多数非线性系统,则可将Volterra泛函级数模型进行如下简化:

$$y(k) = \sum_{n=1}^3 y_n(k) + e(k) = y_1(k) + y_2(k) + y_3(k) + e(k) \quad (5)$$

由于Volterra泛函级数的时域核函数具有对称特性<sup>[23]</sup>,即满足:

$$h_n^{sym}(\tau_1, \tau_2, \dots, \tau_n) = \frac{1}{n!} \sum_{i_1, i_2, \dots, i_n} h_n(\tau_{i_1}, \tau_{i_2}, \dots, \tau_{i_n}) \quad (6)$$

式中: $i_1, i_2, \dots, i_n$ 为 $1, 2, \dots, n$ 中的任意排列。

基于Volterra泛函级数核函数的对称性,可将其各个分量表述为

$$\begin{cases} y_1(k) = a_1(m_1) \sum_{m_1=0}^{N-1} h_1(m_1)u(k-m_1) \\ y_2(k) = a_2(m_1, m_2) \sum_{m_1=0}^{N-1} \sum_{m_2=0}^{N-1} h_2(m_1, m_2)u(k-m_1)u(k-m_2) \\ y_n(k) = a_3(m_1, m_2, m_3) \sum_{m_1=0}^{N-1} \sum_{m_2=0}^{N-1} \sum_{m_3=0}^{N-1} h_3(m_1, m_2, m_3) \cdot \\ u(k-m_1)u(k-m_2)u(k-m_3) \end{cases} \quad (7)$$

式中: $a_i$ 为基于Volterra泛函级数核函数的对称性引入的权系数,并且满足如下关系:

$$a_1(m_1) = 1 \tag{8}$$

$$a_2(m_1, m_2) = \begin{cases} 1 & m_1 = m_2 \\ 2 & m_1 \neq m_2 \end{cases} \tag{9}$$

$$a_3(m_1, m_2, m_3) = \begin{cases} 1 & m_1 = m_2 = m_3 \\ 3 & m_1 \neq m_2 = m_3 \text{ 或} \\ & m_1 = m_2 \neq m_3 \text{ 或} \\ & m_1 \neq m_2 \neq m_3 \\ 6 & m_1 \neq m_2 \text{ 且 } m_2 \neq m_3 \text{ 且} \\ & m_1 \neq m_3 \end{cases} \tag{10}$$

基于上述表述形式, 可将三阶 Volterra 泛函级数模型表示为

$$y(k) = \sum_{m_1=0}^{N-1} a_1(m_1)h_1(m_1)u(k-m_1) + \sum_{m_1=0}^{N-1} \sum_{m_2=0}^{N-1} a_2(m_1, m_2) \cdot h_2(m_1, m_2)u(k-m_1)u(k-m_2) + \sum_{m_1=0}^{N-1} \sum_{m_2=0}^{N-1} \sum_{m_3=0}^{N-1} a_3(m_1, m_2, m_3)h_3(m_1, m_2, m_3) \cdot u(k-m_1)u(k-m_2)u(k-m_3) + e(k) \tag{11}$$

当模型取合适的记忆长度时, 在误差允许的范围,  $e(k)$  可忽略不计。

## 2 改进的蝙蝠算法

### 2.1 基本蝙蝠算法

蝙蝠算法是 Yang<sup>[19]</sup> 在 2010 年提出的模拟蝙蝠觅食特征的启发式群智能随机搜索算法。该算法通过模拟蝙蝠在觅食过程中向外发出的一些超声波, 根据其特征(包括脉冲频率  $f$ 、响度  $A$  及发射率  $R$  等)的变化而建立的寻优模型。在算法寻优迭代的过程中, 可将蝙蝠个体  $i$  的参数变化描述为

$$f_i = f_{\min} + (f_{\max} - f_{\min})\beta \tag{12}$$

$$v_i^d = v_i^{d-1} + (x_i^{d-1} - x_*)f_i \tag{13}$$

$$x_i^d = x_i^{d-1} + v_i^d \tag{14}$$

式中:  $d$  为算法此刻正进行的迭代次数;  $\beta \in [0, 1]$  为随机数, 且服从均匀分布;  $v$  为速度;  $x_*$  为算法此时最优的全局解。

在算法的全局最优解附近, 应对其进行局部搜索, 采取如下策略:

$$x_i^d = x_* + \varepsilon A^d \tag{15}$$

式中:  $\varepsilon$  为  $[-1, 1]$  之间的随机数;  $A^d$  为当前蝙蝠种群的响度平均值。

算法的脉冲响度  $A$  和脉冲发射率  $R$  按照如下策略进行更新:

$$A_i^{d+1} = \alpha A_i^d \tag{16}$$

$$R_i^{d+1} = R_{\max} [1 - \exp(-\gamma d)] \tag{17}$$

式中:  $\alpha$  为蝙蝠响度的衰减系数;  $\gamma$  为脉冲频率的增

加系数。其中,  $\alpha$  和  $\gamma$  都是常量, 同时通过选取合适的  $\alpha$  和  $\gamma$  可以有效平衡算法全局搜索及局部搜索的能力。随着算法深入, 可以得到

$$A_i^d \rightarrow 0, R_i^d \rightarrow R_{\max}, d \rightarrow \infty \tag{18}$$

### 2.2 自适应的改进蝙蝠算法

蝙蝠群体在接近觅食目标时, 为不被发现且能够实时监察目标的运动状态, 采取提高超声波发射率  $R$ , 同时减小脉冲响度  $A$  的策略。因此, 基于上述分析可以看出, 该算法的寻优能力与脉冲响度  $A$  及脉冲发射率  $R$  的动态变化息息相关。寻优初期, 局部极值可能出现多个, 此时的脉冲响度  $A$  较大, 因而在解空间内应采取更多全局搜索策略。此时, 可将脉冲响度与算法的全局搜索能力相结合, 以使算法实现全局搜索能力的自适应调整。寻优后期, 算法初步进入极值区域, 发射率  $R$  较高, 因而应采取更多的局部搜索策略。此时, 可将发射率  $R$  与算法的局部寻优能力相结合, 以使算法实现更优的局部寻优能力。基于上述分析, 基于自适应蝙蝠算法 (self-adaptive bat algorithm, SABA) 利用自适应步长控制机制和变异机制<sup>[20]</sup> 改进传统的蝙蝠算法。

#### 2.2.1 自适应步长控制机制

对于蝙蝠个体  $i$  的速度更新方式, 可借鉴粒子群算法<sup>[24]</sup> 中的速度更新规则方式。在原有的更新机制中, 对速度更新加入了惯性权重因子, 即

$$v_i^d = \omega v_i^{d-1} + f_1 r_1 (h_* - x_i^{d-1}) + f_2 r_2 (x_* - x_i^{d-1}) \tag{19}$$

式中:  $\omega$  为惯性权重因子, 其可以影响蝙蝠算法的全局搜索及局部搜索的能力, 改进算法参照粒子群算法中的调整策略, 采取线性递减的惯性权重方法, 算法前期  $\omega$  的数值较大, 可以使算法具备较强的全局搜索能力, 算法后期  $\omega$  的数值应选取较小值, 使算法在最优解附近具备较强的局部寻优能力, 同时加快收敛;  $h_*$  为蝙蝠个体的历史最优解;  $x_*$  为当前蝙蝠种群的最优解;  $f_1$  和  $f_2$  为脉冲频率;  $r_1$  和  $r_2$  为取值在  $(0, 1)$  之间的均匀分布的随机数。

通过对改进后的蝙蝠个体速度更新公式分析可以看出, 不仅惯性权重对蝙蝠速度更新有影响, 同时脉冲频率  $f_1$  和  $f_2$  也对蝙蝠速度更新有影响。算法初期, 选取较大的脉冲频率  $f_1$  可以增加群体的多样性, 改善全局寻优能力; 在算法后期的局部寻优中, 选取较小的脉冲频率  $f_2$  能够提高算法的局部寻优能力, 同时加快收敛。因此, 本文提出一种自适应改变脉冲频率的机制, 以提高脉冲频率  $f_1$  和  $f_2$  随算法进程的自适应学习能力, 其更新公式如下:

$$f_1 = \alpha \left( 1 - e^{-|F_{\text{avg}} - F_{\text{best}}|} \right) + \gamma \left( 1 - \frac{d}{d_{\max}} \right) + f_{\min} \tag{20}$$

$$f_2 = 3 - \alpha \left( 1 - e^{-|F_{\text{avg}} - F_{\text{best}}|} \right) - \gamma \left( 1 - \frac{d}{d_{\max}} \right) - f_{\min} \tag{21}$$

式中:  $F_{\text{avg}}$  为蝙蝠种群的所有个体适应度的平均值;

$F_{\text{best}}$ 为当前迭代次数下蝙蝠种群中的最优个体;脉冲频率 $f_i$ 在不同的迭代次数 $d$ 中,由不同的适应度平均值 $F_{\text{avg}}$ 决定,其2种影响因素权重分别为 $\gamma$ 和 $\alpha$ ;  $f_{\text{min}}$ 为脉冲频率最小控制量。

针对算法在局部寻优后期阶段过程中的搜索策略,改进的蝙蝠算法将当前迭代次数 $d$ 融入至主要寻优参数脉冲发射率 $R$ 与脉冲响度当中。其结合步骤如下:

1) 选取一个随机参数 $\beta \in [0, 1]$ 。当 $\beta$ 小于蝙蝠脉冲发射率 $R$ 时,算法将进入局部寻优阶段。随着算法进程的不断深入,迭代次数增加,蝙蝠脉冲频率也将增加,从而在算法后期阶段,局部搜索最优解的概率也会增大。

2) 算法的局部搜索策略可以描述为

$$x_i^{d'} = \begin{cases} x_* + A' \text{sg}(d') \delta & d' < 0.4 \\ x_* + A' s 0.1^{g(d')} \delta & d' \geq 0.4 \end{cases} \quad (22)$$

$$s = \frac{\|X_{\text{max}} - X_{\text{min}}\|}{p} \quad (23)$$

$$d' = d/d_{\text{max}} \quad (24)$$

式中: $d'$ 为算法进程评价指标; $d_{\text{max}}$ 为算法迭代的最大次数; $x_*$ 为种群的最优全局解; $A'$ 为当前蝙蝠群体响度的平均值; $X_{\text{max}}$ 和 $X_{\text{min}}$ 分别为可行解域的上边界和下边界; $p$ 为种群的大小; $s$ 为平衡局部搜索步长和待优化问题规模的自适应参数; $\delta \in [-1, 1]$ 为随机量。

3) 为了加速算法收敛,算法引入分段函数 $g(d')$ :

$$g(d') = \begin{cases} e^{0.4-d'} & 0 \leq d' < 0.4 \\ e^{d'-0.4} & 0.4 \leq d' < 0.6 \\ 2e^{d'+0.4} & 0.6 \leq d' \leq 1 \end{cases} \quad (25)$$

引入分段函数 $g(d')$ 是为了在算法初期具有较大的搜索步长,从而扩大算法的搜索区域;进入算法局部寻优阶段时,调整较小的搜索步长,有利于提升寻优精度。

## 2.2.2 变异机制

2.2.1节引入的自适应步长控制机制仅仅能够有效改善算法的全局搜索能力,但不能保证算法一定能够搜索到全局最优解,并且在算法后期,随着脉冲频率 $f_i$ 的减小,导致算法的全局搜索能力下降,因此算法仍可能陷入局部极值解当中。主要原因是:在基本的蝙蝠算法中,算法后期的个体响度减少为0,发射率增大为1,使得蝙蝠个体无法采取变异操作。因此,对响度 $A$ 及发射率 $R$ 进行相关限定约束,使算法后期也可变异操作而不易陷入局部最优。脉冲响度 $A$ 和发射率 $R$ 的更新规则为

$$A_i^{d+1} = \frac{f_1}{f_{\text{max}}} \quad (26)$$

$$R_i^{d+1} = \frac{f_2}{f_{\text{max}}} \quad (27)$$

式中: $f_{\text{max}}$ 为最大脉冲频率限制。其中,脉冲频率 $f_1$ 和 $f_2$ 由当前迭代次数及种群适应度平均共同决定,因此随着算法深入,脉冲响度及发射率产生相应的自适应变化。

结合上述对蝙蝠觅食过程特点的分析,为实现变异操作,本文提出一种结合脉冲响度的变异机制。其变异操作过程如下:对于随机数 $\beta_1 \in [0, 1]$ ,如果满足 $\beta_1 < A$ 条件,同时算法中的蝙蝠个体未进入局部搜索阶段,则继续生成一个随机数 $\beta_2 \in [0, 1]$ ;如果随机数 $\beta_2$ 满足条件 $\beta_2 \geq \rho \in [0, 1]$ ( $\rho$ 为阈值),则对该蝙蝠个体进行随机重置。

上述自适应步长控制机制及变异机制是在全面分析了解蝙蝠觅食行为特征的基础上,充分结合蝙蝠觅食时发出的超声波特征参数,以避免算法过早寻优过程中寻到局部最优值,同时也提升了后期的最优解求解精度。SABA算法的具体流程如下<sup>[20-21]</sup>:

**步骤1** 对蝙蝠种群参数进行初始化处理。其中,包括种群的大小 $m$ 、最大的迭代次数 $d_{\text{max}}$ 、脉冲频率增加系数 $\gamma$ 、脉冲响度衰减系数 $\alpha$ 、惯性权重因子 $\omega$ 、搜索脉冲的频率范围 $[f_{\text{min}}, f_{\text{max}}]$ 、最大发射率 $R_{\text{max}}$ 等。

**步骤2** 对第 $i$ 只蝙蝠个体的位置 $x_i$ 、速度 $v_i$ 、发射率 $R_i$ 及响度 $A_i$ 分别进行初始化,同时根据目标函数求解第 $i$ 只蝙蝠个体的适应度值。

**步骤3** 根据式(20)和式(21)更新蝙蝠的脉冲频率 $f$ ,根据式(26)和式(27)更新蝙蝠的脉冲响度 $A$ 及发射率 $R$ 。

**步骤4** 根据式(19)对蝙蝠个体的速度进行更新,根据式(14)对蝙蝠个体的位置进行更新,依据定义的适应度函数计算蝙蝠个体适应度。

**步骤5** 判断蝙蝠个体是否符合进行局部搜索操作的条件。如果符合,则根据式(22)~式(25)转入局部搜索操作,计算蝙蝠个体的适应度,并转到步骤7。

**步骤6** 判断蝙蝠个体是否符合变异的条件。如果符合,则对蝙蝠个体采取变异操作,并得到每个蝙蝠个体的适应度值。

**步骤7** 判断算法是否达到结束条件。如果未达到,转至步骤3。

## 3 基于SABA优化的Volterra泛函级数模型

Volterra泛函级数模型是非线性动力学系统中最基本的模型,具有线性和非线性的特性,同时综合线性滤波器、非线性滤波器及预测的功能,因而具备满足系统实时性要求的能力<sup>[22-23]</sup>。利用截断之后的Volterra泛函级数模型对目标机动轨迹进行预

测的关键在于确定核参数, 其辨识精度直接影响目标机动轨迹预测的精度。对 Volterra 泛函级数核参数进行辨识实质上是一个多维度参数优化问题, 可通过改进后的 SABA 算法优化 Volterra 泛函级数模型的核参数。

将 Volterra 泛函级数核参数作为待优化变量, 且将模型输入的轨迹时序数据与核参数向量之间得到乘积值。其中, SABA 算法的适应度函数为其乘积值与真实输出值向量的均方差。SABA 算法通过不断迭代进行寻优的方式, 使得预测模型输出不断逼近真实值, 从而得到核参数。采用 SABA 算法优化 Volterra 泛函级数, 待优化的核参数向量可以表示为

$$\mathbf{H} = [h_1(0), h_1(1), \dots, h_1(M_1 - 1), h_2(0, 0), h_2(0, 1), \dots, h_2(M_2 - 1, M_2 - 1), h_3(0, 0, 0), h_3(0, 0, 1), \dots, h_3(M_3 - 1, M_3 - 1, M_3 - 1)] \quad (28)$$

将核参数向量作为 SABA 算法的个体位置, 算法搜索空间维度为

$$D = M_1 + \frac{M_2(M_2 + 1)}{2} + \frac{M_3(M_3 + 1)(2M_3 + 1)}{12} \quad (29)$$

预测模型的输入矩阵为

$$\mathbf{U} = [\mathbf{u}(k), \mathbf{u}(k + 1), \dots, \mathbf{u}(k + L - 1)]^T \quad (30)$$

式中:  $L$  为数据的个数;  $\mathbf{u}(k)$  为第  $k$  时刻的输入向量, 其可以描述为

$$\mathbf{u}(k) = [u(k), u(k - 1), \dots, u(k - M_1 + 1), u^2(k), u(k)u(k - 1), \dots, u^2(k - M_3 + 1)]^T \quad (31)$$

定义 SABA 算法的适应度函数为

$$\text{fitness}(\mathbf{X}_i) = F[\hat{\mathbf{y}} - \mathbf{y}_e] = \frac{1}{L} \sum_{k=1}^L [\hat{y}(k) - y_e(k)]^2 = \frac{1}{L} \sum_{k=1}^L [\mathbf{u}^T(k) \hat{\mathbf{H}}^d - y_e(k)]^2 \quad (32)$$

式中:  $F(*)$  为计算预测输出  $\hat{\mathbf{y}}$  和理想输出  $\mathbf{y}_e$  的均方误差 (MSE);  $\hat{\mathbf{H}}^d$  为算法第  $d$  次迭代所优化的核参数向量值。

基于 SABA 算法优化的 Volterra 泛函级数目标机动轨迹预测流程如图 1 所示。

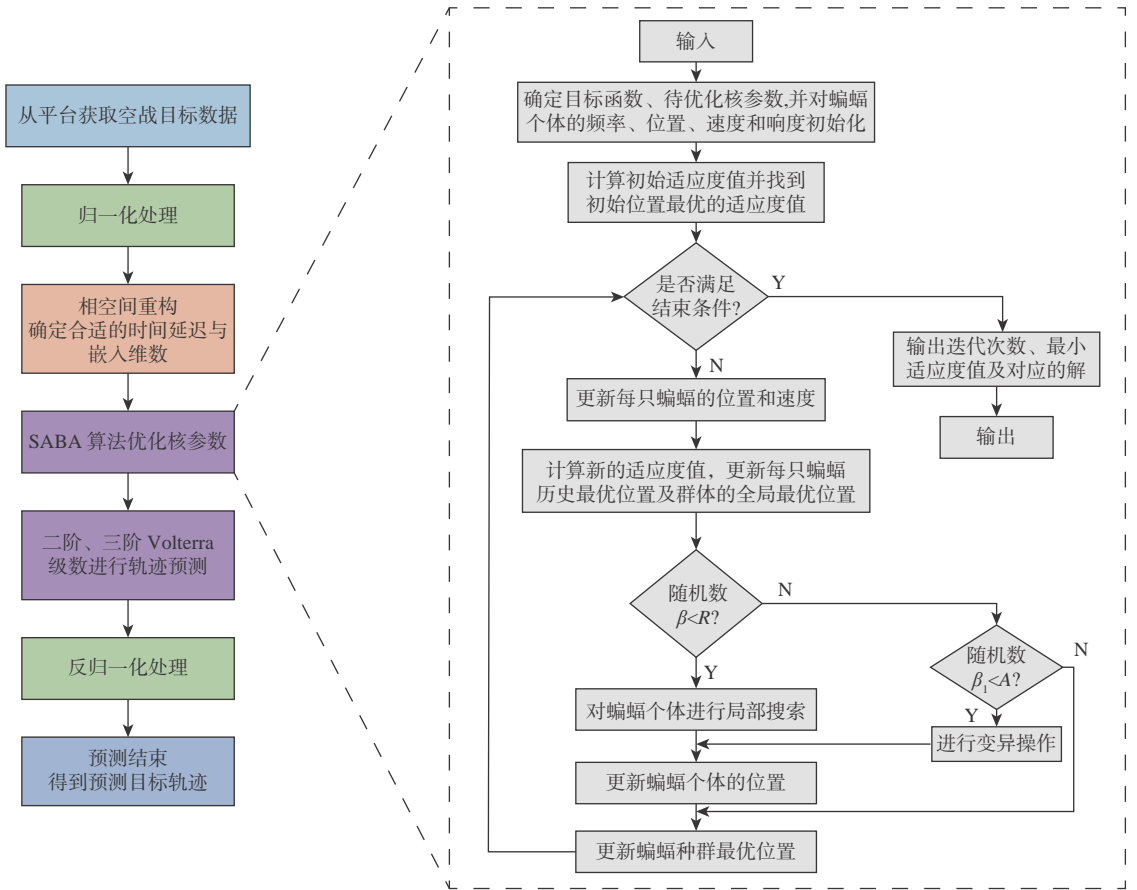


图 1 基于 SABA 算法优化的 Volterra 泛函级数目标机动轨迹预测流程

Fig. 1 Flow chart of target maneuver trajectory prediction based on Volterra functional series optimized by SABA algorithm

## 4 仿真实验

### 4.1 数据预处理

仿真通过基于 4 种智能算法优化的 Volterra 泛

函级数模型进行目标机动轨迹的预测, 并进行对比分析。同时, 为了更好地发挥预测模型的性能, 对目标历史机动轨迹数据进行了归一化和反归一化处理<sup>[25]</sup>。假设目标历史机动轨迹的原始数据为  $\mathbf{X} =$

$(x_1, x_2, \dots, x_N), X \in \mathbf{R}^N$ , 进行归一化的方法如下:

$$x'(t) = \frac{x(t) - x_{\min}}{x_{\max} - x_{\min}} \quad t = 1, 2, \dots, N \quad (33)$$

式中:  $N$ 为目标历史机动轨迹时间序列的长度;  $x(t)$ 为待归一化处理的初始数据;  $x'(t) \in [0, 1]$ 为经过归一化处理之后的轨迹数据;  $x_{\max}$ 为原始数据  $X$  中的最大值;  $x_{\min}$ 为  $X$  中的最小值。

假设基于智能算法优化的Volterra泛函级数预测模型仿真输出为  $Y = (y_1, y_2, \dots, y_N), Y \in [0, 1]$ , 对其进行反归一化的方法为

$$y'(t) = y(t)(y_{\max} - y_{\min}) + y_{\min} \quad t = 1, 2, \dots, N \quad (34)$$

式中:  $y(t)$ 为模型得到的预测值;  $y'(t)$ 为经过反归一化处理之后的预测轨迹值;  $y_{\max}$ 为预测得到序列  $Y$  中的最大值;  $y_{\min}$ 为预测得到序列  $Y$  中的最小值。

为获得原始数据的相空间结构, 可通过相空间重构进行处理<sup>[25]</sup>。该处理可利用一维时序数据逆向重构出原始系统的相空间结构。Takens 定理<sup>[26]</sup>证明了在原始的系统, 其任一分量的演化均由其他及其相关的分量确定。假设获得的分量为  $x$ , 研究的时间序列为  $x(t), t = 1, 2, \dots, N$ , 则其重构的相空间可表示为

$$X(t) = [x(t), x(t-\tau), \dots, x(t-(m-1)\tau)] \\ t = (m-1)\tau + 1, \dots, N \quad (35)$$

式中:  $\tau$ 为时间延迟;  $m$ 为嵌入维数。选择合适的时

间延迟和嵌入维数可以对未来  $t + \eta$  时刻的目标机动轨迹值  $x(t + \eta)$  进行预测, 其中,  $\eta$  为预测步长。当  $\eta = 1$  时, 即对目标机动轨迹进行单步预测; 当  $\eta > 1$  时, 即对目标机动轨迹进行多步预测。

### 4.2 目标机动轨迹时间序列预测

仿真采用 4 种智能算法优化的 Volterra 模型和 BPNN 模型对目标机动轨迹时间序列进行单步预测并分析。首先, 对目标机动轨迹时间序列进行相空间重构<sup>[26]</sup>。根据 C-法<sup>[26]</sup> 确定目标机动轨迹  $x$  方向嵌入维数  $m_x = 6$  和时间延迟  $\tau_x = 8$ ;  $y$  方向的时间延迟  $\tau_y = 7$ , 嵌入维数  $m_y = 7$ ;  $z$  方向的时间延迟  $\tau_z = 5$ , 嵌入维数  $m_z = 7$ 。其中, 采用单步预测方式, 即设置预测步数  $\eta = 1$ 。仿真中, 采用 4 种智能算法优化的二阶和三阶 Volterra 模型对目标机动轨迹的三维坐标进行空间位置预测, 预测结果经平滑后如图 2 所示。

通过图 2 中预测结果及误差对比可以看出, 5 种预测模型的预测值与目标机动轨迹的实际值的变化大体上保持一致, 说明 4 种智能算法优化的二阶、三阶 Volterra 级数模型和 BPNN 模型可以对目标机动轨迹进行有效预测。但是从整体上可以看出, 基于 SABA 算法优化的 Volterra 泛函级数模型预测结果较其他模型更加接近真实值, 二阶 Volterra 模型比三阶 Volterra 模型更接近真实值, 但各个预

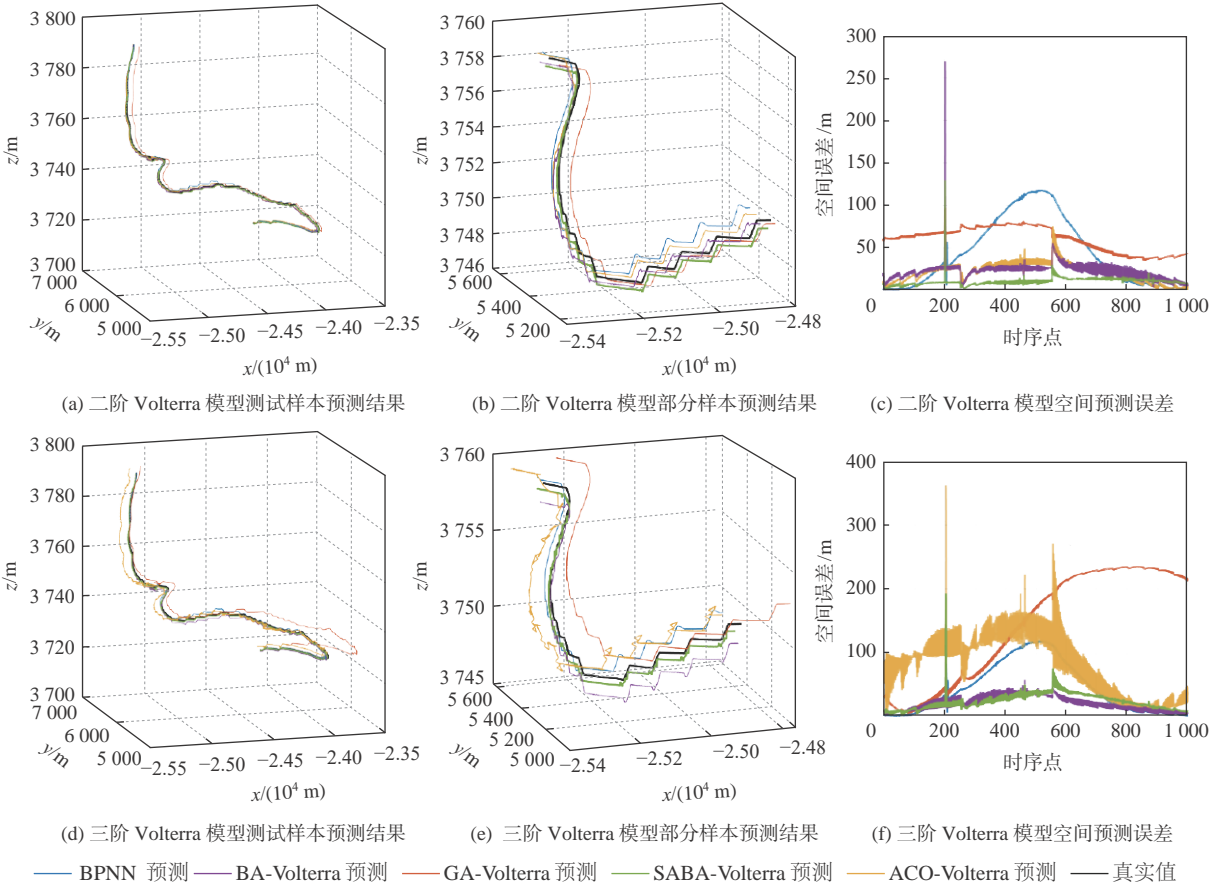


图 2 基于二、三阶 Volterra 模型目标空间预测结果对比

Fig. 2 Comparison of target space prediction results based on second-order and third-order Volterra models

测模型的预测精度具体效果还需要进行进一步的误差分析仿真实验。

### 4.3 模型性能评价指标

仿真中, 主要对比 SABA-Volterra 模型与其他预测模型在预测精度方面的性能差异。仿真实验利用平均绝对误差 ( $M_{ad}$ )、均方误差 ( $M_{sc}$ )、平均绝对百分比误差 ( $M_{ape}$ ) 及相关系数 ( $C_{or}$ ) 等 4 个指标进行衡量对比。其性能指标定义如下:

$$M_{ad} = \frac{1}{n} \sum_{k=1}^n |y(k) - \hat{y}(k)| \quad (36)$$

$$M_{sc} = \frac{1}{n} \sum_{k=1}^n |y(k) - \hat{y}(k)|^2 \quad (37)$$

$$M_{ape} = \frac{1}{n} \sum_{k=1}^n \left| \frac{y(k) - \hat{y}(k)}{\hat{y}(k)} \right| \quad (38)$$

$$C_{or} = \frac{\sum_{k=1}^n |\hat{y}(k) - \bar{\hat{y}}| \cdot |y(k) - \bar{y}|}{n\sigma(\hat{y})\sigma(y)} \quad (39)$$

式中:  $\hat{y}(k)$  为模型输出的目标轨迹预测值;  $y(k)$  为样本中目标轨迹的真实值;  $\bar{y}$  为  $y$  的平均值;  $\sigma(\hat{y})$  为  $\hat{y}$  的标准差;  $\bar{\hat{y}}$  为  $\hat{y}$  的平均值;  $\sigma(y)$  为  $y$  的标准差;  $n$  为样本个数。

限于篇幅, 仿真实验以 3 个坐标构成三维空间位置精度进行计算分析。其中, 空间相关系数取 3 个坐标预测结果与真实值的相关系数的平均值。

为进一步分析 Volterra 泛函级数的预测性能, 对 SABA-Volterra 算法进行多步预测仿真, 并与 BA-Volterra<sup>[27]</sup>、SABA-Volterra、ACO-Volterra<sup>[16]</sup>、GA-Volterra、BPNN<sup>[7]</sup> 等预测模型进行对比分析。如表 1~表 3 所示, 分别进行了 2、4 和 8 步预测。其中, 考虑到 Volterra 泛函级数的实时性和单步预测的优异性能, 采用单步递归方法进行多步预测仿真<sup>[28]</sup>。即将  $t$  时刻下得到的单步预测结果  $y_t$  与前时序序列  $X_t = (x_{t-m}, x_{t-m+1}, \dots, x_t)$  组合成新的时序序列  $X_{t+1} = (x_{t-m+1}, \dots, x_t, y_t)$  作为输入, 直至预测所需步数。该多步预测方法对算法的单步预测性能要求更高, 以减少多步递归导致的预测误差累积。

通过图 2 和表 1~表 3 可以看出: ①基于智能算法优化的 Volterra 泛函级数模型, 可以利用目标历史机动轨迹数据有效地反映出目标未来运动趋势, 并可以较高精度地实现对目标未来轨迹一步预测, 且预测结果较为稳定, 实现简便、快捷; ②基于 SABA 优化后的 Volterra 泛函级数的单步预测较其他算法优化后的 Volterra 及 BPNN 预测模型有更高的

表 1 不同预测模型进行 2 步预测的性能对比

Table 1 Performance comparison of two-step prediction with different prediction models

算法	阶数	$M_{ad}$	$M_{sc}$	$M_{ape}$	$C_{or}$
BA-Volterra	二阶	21.913 7	$6.563 3 \times 10^2$	0.000 9	0.998 3
BA-Volterra	三阶	30.502 1	$1.460 4 \times 10^3$	0.001 2	0.997 6
SABA-Volterra	二阶	15.433 7	$2.424 6 \times 10^2$	0.000 6	0.998 7
SABA-Volterra	三阶	25.187 4	$8.296 9 \times 10^2$	0.001 0	0.997 8
ACO-Volterra	二阶	24.432 0	$8.006 2 \times 10^2$	0.001 0	0.997 9
ACO-Volterra	三阶	112.590 1	$1.684 0 \times 10^4$	0.004 3	0.989 1
GA-Volterra	二阶	35.425 2	$1.355 7 \times 10^3$	0.001 4	0.997 5
GA-Volterra	三阶	202.458 3	$5.289 1 \times 10^4$	0.008 0	0.983 6
BPNN		83.799 4	$9.162 4 \times 10^3$	0.003 3	0.997 1

表 2 不同预测模型进行 4 步预测的性能对比

Table 2 Performance comparison of four-step prediction with different prediction models

算法	阶数	$M_{ad}$	$M_{sc}$	$M_{ape}$	$C_{or}$
BA-Volterra	二阶	85.605 2	$8.851 1 \times 10^3$	0.003 3	0.989 9
BA-Volterra	三阶	134.292 2	$2.223 0 \times 10^4$	0.005 3	0.988 6
SABA-Volterra	二阶	27.907 5	$1.269 8 \times 10^3$	0.001 1	0.998 2
SABA-Volterra	三阶	27.959 7	$1.104 6 \times 10^3$	0.001 1	0.998 3
ACO-Volterra	二阶	30.599 9	$1.143 9 \times 10^3$	0.001 3	0.997 8
ACO-Volterra	三阶	114.262 2	$1.695 3 \times 10^4$	0.004 5	0.989 0
GA-Volterra	二阶	53.713 0	$3.015 4 \times 10^3$	0.002 1	0.997 0
GA-Volterra	三阶	210.482 2	$5.956 8 \times 10^4$	0.008 2	0.983 4
BPNN		221.763 6	$7.253 6 \times 10^4$	0.008 7	0.983 0

表3 不同预测模型进行8步预测的性能对比

Table 3 Performance comparison of eight-step prediction with different prediction models

算法	阶数	$M_{ad}$	$M_{sc}$	$M_{ape}$	$C_{or}$
BA-Volterra	二阶	108.002 4	$1.475\ 9 \times 10^4$	0.004 2	0.988 7
BA-Volterra	三阶	125.673 3	$1.900\ 8 \times 10^4$	0.004 9	0.988 8
SABA-Volterra	二阶	39.527 5	$1.589\ 9 \times 10^3$	0.001 5	0.997 6
SABA-Volterra	三阶	71.762 4	$6.519\ 5 \times 10^3$	0.002 8	0.996 8
ACO-Volterra	二阶	294.702 3	$1.395\ 2 \times 10^5$	0.011 4	0.981 1
ACO-Volterra	三阶	545.395 2	$3.806\ 3 \times 10^5$	0.021 2	0.968 9
GA-Volterra	二阶	100.653 2	$1.049\ 8 \times 10^4$	0.003 9	0.989 0
GA-Volterra	三阶	430.597 7	$2.278\ 4 \times 10^5$	0.016 7	0.975 3
BPNN		369.869 7	$1.923\ 3 \times 10^5$	0.014 4	0.977 3

预测精度,从而实现较优的单步递归方法下的多步预测性能;③基于4种智能算法优化的二阶Volterra泛函级数模型具有运算速度快、预测精度高的优点,且硬件容易实现。

对于目标机动轨迹时间序列,二阶Volterra泛函级数模型能够对其进行有效预测的本质在于:二阶截断Volterra泛函级数综合了目标机动轨迹时间序列中的线性和非线性因素,更符合系统的非线性本质<sup>[29]</sup>;而当Volterra泛函级数的截断阶数取值较大时,其导致存在大量的非线性耦合,而使得Volterra泛函级数模型的系数数量呈几何级数增长,不仅确定模型参数困难,而且非线性项会造成智能算法优化的Volterra泛函级数参数不佳,进而造成Volterra泛函级数的系数不收敛或者不稳定<sup>[30]</sup>,最终影响预测精度。

#### 4.4 算法优化性能对比

对目标机动轨迹的三维坐标分别进行建模与预测时,GA-Volterra算法、ACO-Volterra算法、BA-Volterra算法及SABA-Volterra算法的收敛曲线如

图3所示。其中,纵坐标为对数坐标轴,且设定为各预测算法在空间位置中寻优适应度的对数值。

在图3(a)中,二阶情况下ACO-Volterra、BA-Volterra、GA-Volterra和SABA-Volterra算法的适应度函数值达到较小值且稳定时的迭代次数分别为149、74、230、60;在图3(b)中,三阶情况下4种算法的适应度函数值达到较小值且稳定时,其所迭代的次数分别为47、63、199、169。仿真中,将Volterra泛函级数的预测误差设定为智能优化算法的适应度函数,即适应度函数值越大,Volterra泛函级数的预测误差越大。通过上述分析可知,本文提出的SABA-Volterra算法模型能够实现较高的预测精度及较快的收敛速度,同时也进一步验证了二阶Volterra更加适合于空战目标机动轨迹预测。

#### 4.5 算法实时性能对比

考虑到空战对抗中对预测的实时性要求高,进一步分析SABA-Volterra模型的实时性能,并与其他算法进行对比。各算法的训练及预测所需时间如图4所示。其中,迭代次数设置为250次。

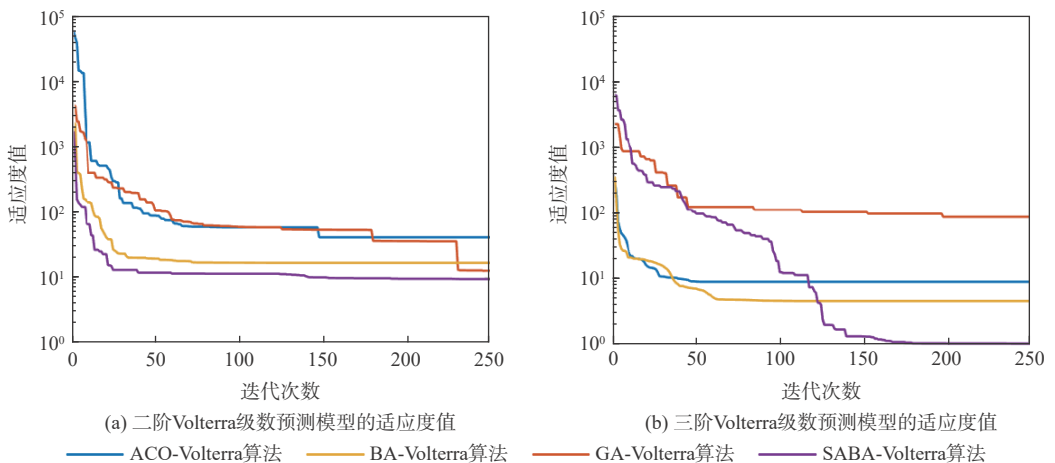


图3 不同智能算法优化预测模型的空间适应度函数值比较

Fig. 3 Comparison of spatial fitness function values of prediction models optimized by different intelligent algorithms

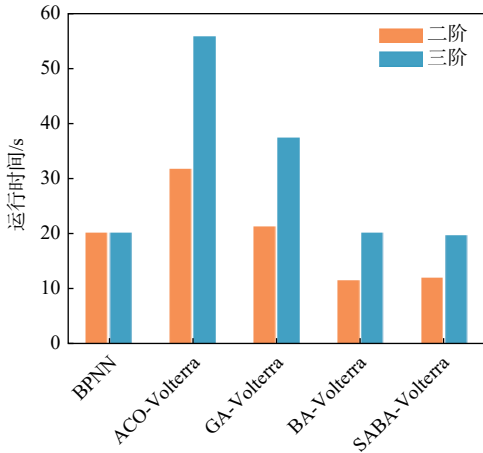


图4 不同智能算法优化预测模型的运行时间对比

Fig. 4 Comparison of running time of prediction models optimized by several intelligent algorithms

如图4所示,改进后的SABA-Volterra级数仍保持较高的实时性,并优于其他传统算法。同时,采取二阶的Volterra泛函级数模型,相比于三阶能实现更高的实时性。这是由于三阶Volterra模型需要更多的核函数,从而提高了算法运算的复杂度。因此,综合预测精度与实时性对比,二阶模型相比三阶模型更适用于空战目标机动轨迹预测。

## 5 结论

针对空战目标机动轨迹预测问题,引入Volterra泛函级数模型,针对其存在的不足,提出一种改进的蝙蝠算法对其优化,构建一种基于SABA算法优化的Volterra泛函级数目标机动轨迹预测模型。通过理论分析和仿真实验可以得出如下结论:

1) 改进后的SABA算法具有自适应步长控制机制和变异机制等能力,可适应不同问题的解空间,从而能够自适应调整合适的步长,进而提高算法求解精度。仿真结果表明,SABA算法的寻优性能较优异。

2) 改进的蝙蝠算法能够确定合适的Volterra泛函级数核系数,所构建的基于SABA优化的Volterra泛函级数模型能对目标机动轨迹实现较高精度的实时预测。

3) 不同的Volterra泛函级数模型具有不同的预测精度。针对不同的预测问题,需要通过仿真实验,得出最佳的截断阶数。

## 参考文献 (References)

[1] 寇英信,李战武,陈哨东,等.火控系统在航空作战中的作用——作战飞机之“魂”[J].电光与控制,2013,20(12):1-5.  
KOU Y X, LI Z W, CHEN S D, et al. The important role of fire

control system in air combat—Soul of fighters[J]. Electronics Optics & Control, 2013, 20(12): 1-5(in Chinese).

[2] 姜佰辰,关键,周伟,等.基于多项式卡尔曼滤波的船舶轨迹预测算法[J].信号处理,2019,35(5):741-746.  
JIANG B C, GUAN J, ZHOU W, et al. Vessel trajectory prediction algorithm based on polynomial fitting Kalman filtering[J]. Journal of Signal Processing, 2019, 35(5): 741-746(in Chinese).

[3] 赵帅兵,唐诚,梁山,等.基于改进卡尔曼滤波的控制河段船舶航迹预测[J].计算机应用,2012,32(11):3247-3250.  
ZHAO S B, TANG C, LIANG S, et al. Track prediction of vessel in controlled waterway based on improved Kalman filter[J]. Journal of Computer Applications, 2012, 32(11): 3247-3250(in Chinese).

[4] 乔少杰,韩楠,朱新文,等.基于卡尔曼滤波的动态轨迹预测算法[J].电子学报,2018,46(2):418-423.  
QIAO S J, HAN N, ZHU X W, et al. A dynamic trajectory prediction algorithm based on Kalman filter[J]. Acta Electronica Sinica, 2018, 46(2): 418-423(in Chinese).

[5] 翟岱亮,雷虎民,李炯,等.基于自适应IMM的高超声速飞行器航迹预测[J].航空学报,2016,37(11):3466-3475.  
ZHAI D L, LEI H M, LI J, et al. Trajectory prediction of hypersonic vehicle based on adaptive IMM[J]. Acta Aeronautica et Astronautica Sinica, 2016, 37(11): 3466-3475(in Chinese).

[6] 杨彬,贺正洪.一种GRNN神经网络的高超声速飞行器航迹预测方法[J].计算机应用与软件,2015,32(7):239-243.  
YANG B, HE Z H. Hypersonic vehicle track prediction based on GRNN[J]. Computer Applications and Software, 2015, 32(7): 239-243(in Chinese).

[7] 谭伟,陆百川,黄美灵.神经网络结合遗传算法用于航迹预测[J].重庆交通大学学报,2010,29(1):147-150.  
TAN W, LU B C, HUANG M L. Track prediction based on neural networks and genetic algorithm[J]. Journal of Chongqing Jiaotong University, 2010, 29(1): 147-150(in Chinese).

[8] 甘旭升,端木京顺,孟月波,等.基于粒子群优化的WNN飞行数据气动力建模[J].航空学报,2012,33(7):1209-1217.  
GAN X S, DUANGMU J S, MENG Y B, et al. Aerodynamic modeling from flight data based on WNN optimized by particle swarm[J]. Acta Aeronautica et Astronautica Sinica, 2012, 33(7): 1209-1217(in Chinese).

[9] SLATTERY R, ZHAO Y. Trajectory synthesis for air traffic automation[J]. Journal of Guidance, Control, and Dynamic, 1997, 20(2): 232-238.

[10] 张家良,曹建福,高峰.大型装备传动系统非线性频谱特征提取与故障诊断[J].控制与决策,2012,27(1):135-138.  
ZHANG J L, CAO J F, GAO F. Feature extraction and fault diagnosis of large-scale equipment transmission system based on nonlinear frequency spectrum[J]. Control and Decision, 2012, 27(1): 135-138(in Chinese).

[11] 张华君.基于递推批量最小二乘的Volterra级数辨识方法[J].小型微型计算机,2004,25(12):2282-2285.  
ZHANG H J. Volterra series identification method based on recursive least square algorithm[J]. Mini-Micro Systems, 2004, 25(12): 2282-2285(in Chinese).

[12] 孔祥玉,韩崇昭,马红光,等.基于Volterra级数的全解耦RLS自适应辨识算法[J].系统仿真学报,2004,16(4):807-809.  
KONG X Y, HAN C Z, MA H G, et al. Fully decoupled RLS

- adaptive identification algorithm based on Volterra series[J]. *Journal of System Simulation*, 2004, 16(4): 807-809(in Chinese).
- [13] 孔祥玉, 韩崇昭, 马红光, 等. 一种总体最小二乘算法及在Volterra滤波器中的应用[J]. *西安交通大学学报*, 2004, 38(4): 339-342.
- KONG X Y, HAN C Z, MA H G, et al. Total least square algorithm and its application to Volterra filter[J]. *Journal of Xi'an Jiao Tong University*, 2004, 38(4): 339-342(in Chinese).
- [14] 唐浩, 屈梁生, 温广瑞. 基于Volterra级数的转子故障诊断研究[J]. *中国机械工程*, 2009, 20(4): 447-454.
- TANG H, QU L S, WEN G R. Fault diagnosis for rotor system based on Volterra series[J]. *China Mechanical Engineering*, 2009, 20(4): 447-454(in Chinese).
- [15] ABBAS H M, BAYOUMI M M. Volterra system identification using adaptive genetic algorithms[J]. *Applied Soft Computing*, 2005, 5(1): 75-86.
- [16] 李志农, 唐高松, 肖尧先, 等. 基于自适应蚁群优化的Volterra核辨识算法研究[J]. *振动与冲击*, 2011, 30(10): 35-38.
- LI Z N, TANG G S, XIAO Y X, et al. Volterra series identification method based on adaptive ant colony optimizations[J]. *Journal of Vibration and Shock*, 2011, 30(10): 35-38(in Chinese).
- [17] 李志农, 蒋静, 陈金刚, 等. 基于量子粒子群优化的Volterra核辨识算法研究[J]. *振动与冲击*, 2013, 32(3): 60-63.
- LI Z N, JIANG J, CHEN J G, et al. Volterra series identification method based on quantum particle swarm optimization[J]. *Journal of Vibration and Shock*, 2013, 32(3): 60-63(in Chinese).
- [18] 李志农, 蒋静, 冯辅周, 等. 基于量子粒子群优化Volterra时域核辨识的隐Markov模型识别方法[J]. *仪器仪表学报*, 2011, 32(12): 2693-2698.
- LI Z N, JIANG J, FENG F Z, et al. Hidden Markov model recognition method based on Volterra kernel identified with particle swarm optimization[J]. *Chinese Journal of Scientific Instrument*, 2011, 32(12): 2693-2698(in Chinese).
- [19] YANG X. A new metaheuristic bat-inspired algorithm[C]//*Nature Inspired Cooperative Strategies for Optimization*. Berlin: Springer, 2010, 284: 65-74.
- [20] 吕石磊, 黄永霖, 陈海强, 等. 基于自适应步长的改进蝙蝠算法[J]. *控制与决策*, 2018, 33(3): 557-567.
- LV S L, HUANG Y L, CHEN H Q, et al. Improved bat algorithm using self-adaptive step[J]. *Control and Decision*, 2018, 33(3): 557-567(in Chinese).
- [21] FISTER I, FISTER D, YANG X S. A hybrid bat algorithm[J]. *Elektrotehniski Vestnik*, 2013, 80(1): 1-7.
- [22] TAN L, JIANG J. Adaptive second-order Volterra filtered-XRLS algorithms with sequential and partial updates for non-linear active noise control[C]//*Proceedings of 4th IEEE Conference on Industrial Electronics and Applications*. Piscataway: IEEE Press, 2009: 1625-1630.
- [23] CHENG C M, PENG Z K, ZHANG W M, et al. Volterra-series-based nonlinear system modeling and its engineering applications: A state-of-the-art review[J]. *Mechanical Systems and Signal Processing*, 2017, 87: 340-364.
- [24] 卫晓娟, 丁旺才, 李宁洲, 等. 基于改进粒子群算法的Volterra模型参数辨识[J]. *振动与冲击*, 2015, 34(21): 105-112.
- WEI X J, DING W C, LI N Z, et al. Parametric identification of nonlinear Volterra model based on improved PSO algorithm[J]. *Journal of Vibration and Shock*, 2015, 34(21): 105-112(in Chinese).
- [25] SONG J, MENG D, WANG Y. Analysis of chaotic behavior based on phase space reconstruction methods[C]//*Proceedings of the IEEE 6th International Symposium on Computational Intelligence and Design*. Piscataway: IEEE Press, 2014: 414-417.
- [26] 陆振波, 蔡志明, 姜可宇. 基于改进的C-C方法的相空间重构参数选择[J]. *系统仿真学报*, 2007, 19(11): 2527-2529.
- LU Z B, CAI Z M, JIANG K Y. Determination of embedding parameters for phase space reconstruction based on improved C-C methods[J]. *Journal of System Simulation*, 2007, 19(11): 2527-2529(in Chinese).
- [27] 赵玮, 王强, 何晓晖, 等. 基于BA优化核参数的非线性Volterra滤波方法研究[J]. *仪表技术与传感器*, 2019(6): 77-81.
- ZHAO W, WANG Q, HE X H, et al. Research on nonlinear Volterra filter method with kernel parameters optimized based on BA[J]. *Instrument Technique and Sensor*, 2019(6): 77-81(in Chinese).
- [28] 姜学鹏, 洪贝. 基于AP的Volterra级数自适应多重回归及其多步预测应用[J]. *系统工程与电子技术*, 2014, 36(12): 2652-2655.
- JIANG X P, HONG B. Multi-step predicting model based on multi-recursive AP algorithm of Volterra series[J]. *Systems Engineering and Electronics*, 2014, 36(12): 2652-2655(in Chinese).
- [29] 韩敏. 混沌时间序列预测理论与方法[M]. 北京: 中国水利水电出版社, 2007: 24-30.
- HAN M. Theory and method of chaotic time series prediction[M]. Beijing: China Water Resources and Hydropower Press, 2007: 24-30(in Chinese).
- [30] 张家树, 肖先赐. 混沌时间序列的自适应高阶非线性滤波预测[J]. *物理学报*, 2000, 49(7): 1222-1226.
- ZHANG J S, XIAO X C. Prediction of chaotic time series by using adaptive higher-order nonlinear Fourier infrared filters[J]. *Acta Physica Sinica*, 2000, 49(7): 1222-1226(in Chinese).

## Air combat maneuver trajectory prediction of target based on Volterra series optimized by SABA algorithm

LI Zhanwu<sup>1,\*</sup>, PENG Mingyu<sup>2</sup>, GAO Chunqing<sup>3</sup>, YANG Aiwu<sup>2</sup>, XU An<sup>1</sup>, FANG Chengzhe<sup>2</sup>

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

2. Graduate School, Air Force Engineering University, Xi'an 710038, China;

3. Unit 94582 of PLA, Xinyang 464194, China)

**Abstract:** Target maneuver trajectory prediction is an important prerequisite for air combat situation awareness and target threat assessment. Aiming at the problems of high complexity and low prediction accuracy of traditional target maneuvering trajectory prediction model, Volterra functional series model was introduced to predict the target maneuvering trajectory by analyzing and combining the chaotic characteristics of target maneuvering trajectory time series data. To solve the problem that it is difficult to solve the high-order kernel function in Volterra functional series model, the mutation mechanism and adaptive step control mechanism were used to improve the optimization ability of bat algorithm. Then, a Volterra functional series target maneuver trajectory prediction model based on self-adaptive bat algorithm (SABA) optimization was constructed, and the future maneuvering trajectory of the target was predicted by using the optimized Volterra series model with different orders. In the simulation experiment, the feasibility of the prediction model is verified by comparing with the prediction accuracy of the Volterra series prediction model improved by other optimization algorithms, and the second-order Volterra series model is proved to be more suitable for target maneuver trajectory prediction.

**Keywords:** trajectory prediction; Volterra functional series model; kernel parameter optimization; self-adaptive bat algorithm; truncation order

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2022.0362

# 基于倒置残差注意力的无人机航拍 图像小目标检测

刘树东, 刘业辉, 孙叶美, 李懿霏, 王娇\*

(天津城建大学 计算机与信息工程学院, 天津 300384)

**摘要:** 针对无人机航拍图像背景复杂、小尺寸目标较多等问题, 提出了一种基于倒置残差注意力的无人机航拍图像小目标检测算法。在主干网络部分嵌入倒置残差模块与倒置残差注意力模块, 利用低维向高维的特征信息映射, 获得丰富的小目标空间信息和深层语义信息, 提升小目标的检测精度; 在特征融合部分设计多尺度特征融合模块, 融合浅层空间信息和深层语义信息, 并生成4个不同感受野的检测头, 提升模型对小尺寸目标的识别能力, 减少小目标的漏检; 设计马赛克混合数据增强方法, 建立数据之间的线性关系, 增加图像背景复杂度, 提升算法的鲁棒性。在VisDrone数据集上的实验结果表明: 所提模型的平均精度均值比DSHNet模型提升了1.2%, 有效改善了无人机航拍图像小目标漏检、误检的问题。

**关键词:** 目标检测; 无人机图像; 倒置残差; 注意力; 多尺度特征融合

**中图分类号:** TP391.4

**文献标志码:** A

**文章编号:** 1001-5965(2023)03-0514-11

近年来, 深度学习技术表现出优异的性能, 基于深度学习的无人机航拍图像检测已被广泛应用于各行各业<sup>[1]</sup>, 如城市规划、环境监测和灾害检测等。由于无人机飞行高度变化较快、拍摄角度和位置具有特殊性, 航拍图像具有背景复杂、小目标较多的特点, 影响算法的检测精度, 采用深度学习的方法提升目标检测精度仍需进一步研究。

基于深度学习的目标检测算法主要分为两阶段检测算法和单阶段检测算法2类。其中, 两阶段检测算法先生成候选区域, 再对候选区域进行分类和定位, 如R-CNN系列<sup>[2-4]</sup>算法、Mask R-CNN<sup>[5]</sup>算法和Cascade R-CNN算法<sup>[6]</sup>等, 但这些算法在检测过程中存在候选区域, 通常会产生较大的计算开销和计算成本。单阶段检测算法通过初始锚点框对目标预测类别并定位, 可以在不进行候选区域的情

况下完成端到端的目标检测, 使得网络结构更加简单且快速运行, 具有良好的实用性能, 如YOLO系列<sup>[7-10]</sup>算法、SSD<sup>[11]</sup>算法、RetinaNet<sup>[12]</sup>算法等。与两阶段检测算法相比, 单阶段检测算法具有较高的实时性, 更能满足无人机航拍图像目标检测的需求。目前, 基于单阶段检测算法的无人机航拍图像目标检测越来越受到人们的关注。Liu等<sup>[13]</sup>基于YOLOv3检测框架提出UAV-YOLO, 通过增加浅层卷积改善整个网络结构, 丰富空间信息, 但对于小目标存在漏检和误检问题。Liang等<sup>[14]</sup>提出一种基于特征融合和尺度缩放的单阶段检测模型FS-SSD, 利用反卷积模块和特征融合模块进行预测, 并通过上下文分析进一步提高检测精度。Zhang等<sup>[15]</sup>提出一种实时检测模型SlimYOLOv3, 通过对YOLOv3进行稀疏训练及通道剪枝的方式来减少模型的计算量

收稿日期: 2022-05-16; 录用日期: 2022-08-19; 网络出版时间: 2022-10-18 10:31

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20221017.1624.010.html](https://kns.cnki.net/kcms/detail/11.2625.V.20221017.1624.010.html)

基金项目: 天津市科技计划(20YDTPJC01310)

\*通信作者. E-mail: [wangjiaoq@163.com](mailto:wangjiaoq@163.com)

**引用格式:** 刘树东, 刘业辉, 孙叶美, 等. 基于倒置残差注意力的无人机航拍图像小目标检测[J]. 北京航空航天大学学报, 2023, 49(3): 514-524.  
LIU S D, LIU Y H, SUN Y M, et al. Small object detection in UAV aerial images based on inverted residual attention[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(3): 514-524 (in Chinese).

和参数,以达到实时检测的效果,但对小目标的检测精度不高。裴伟等<sup>[16]</sup>提出一种改进的SSD无人机航拍图像目标检测算法,通过提出的特征融合机制将网络的浅层视觉特征与深层语义特征进行有效结合,改善了SSD算法存在的重复检测问题,但由于网络深度增大了计算量。刘婷婷等<sup>[17]</sup>在YOLOv3模型的基础上,通过轻量化主干网络和融合场景上下文信息提升检测精度,但对于小目标存在漏检情况。虽然目前算法在评价指标上取得了一定的提高,但仍存在小目标检测精度较低的问题。

目前,轻量化网络结构在几乎不影响性能的同时使得参数量明显下降,其主要采用深度可分离卷积(depthwise separable convolution, DS-Conv)和倒置残差结构实现。倒置残差结构<sup>[18-19]</sup>将残差块的“降维-提取-升维”转换为“升维-提取-降维”模式,能够实现将特征图的通道进行扩张,丰富特征数量,进一步提升检测精度,减少深度卷积提取特征较少的问题。倒置残差注意力(inverted residuals attention, IRA)模块添加了基于通道注意力的压缩激励模块,提高了通道之间的相关性,提升了模型检测精度,在表面缺陷检测和隧道渗漏水检测方面均取得了良好的效果。刘艳菊等<sup>[20]</sup>提出一种具有实时性的钢条表面缺陷检测算法,利用倒置残差注意力优化YOLOv4的特征提取网络,并采用K-Means聚类加快收敛速度,有效提升了检测精度。周中等<sup>[21]</sup>提出一种基于语义分割的隧道渗漏水图像识别算法,利用倒置残差注意力对DeepLabv3+语义分割算法进行优化,在降低参数量的同时提升图像目标识别精度。在此基础上,本文在倒置残差注意力模块中添加有效通道注意力(efficient channel attention, ECA-Net)<sup>[22]</sup>模块,通过不降维的逐通道全局平均池化来减少信息丢失,并建立通道之间的依赖关系,更多地关注小目标特征。倒置残差(inverted residuals, IR)模块采用 $3\times 3$ 卷积替换深度可分离卷积,避免在获取丰富浅层特征的同时减少内存访问开销,并提升模型训练速度。

针对上述问题,本文提出了一种基于倒置残差注意力的无人机航拍图像小目标检测算法模型,改进YOLOv5x特征提取阶段的CSPDarknet53(C3)模块以提升特征提取能力,设计多尺度特征融合(multi-scale feature fusion, MFF)模块以加强特征融合,设计马赛克混合(mosaic-mixup, MM)数据增强方法以实现数据增强,提升无人机航拍图像的检测精度。

1) 设计倒置残差注意力C3(IRAC3)模块,利用

倒置残差中“升维-提取-降维”的模式,进一步在倒置残差注意力模块中采用ECA-Net模块,通过跨通道交互策略实现通道间的信息交流,减少信息丢失问题,并建立通道之间的依赖关系,突出感兴趣区域特征的表达能力,获取丰富的小目标的细节信息,并通过深度可分离卷积<sup>[23]</sup>的通道和特征区域分离降低模型参数量。设计倒置残差C3(IRC3)模块,倒置残差模块在倒置残差注意力模块的基础上采用 $3\times 3$ 卷积替换深度可分离卷积,在激活函数之前增加通道数,将低维特征映射到高维空间,提升浅层特征的提取能力,减少非线性激活函数造成的空间信息损失,获取丰富的小目标位置信息,在获取丰富浅层特征的同时减少内存访问开销,并提升模型训练速度。

2) 设计多尺度特征融合模块,将提取的深层语义信息经过上采样放大后与浅层空间信息融合,生成用于极小目标的检测头,增强特征图中小尺寸目标的学习能力和对感兴趣区域的定位能力。通过结合3个不同感受野的检测头,提升模型在多尺度目标下的识别能力,进一步减少小目标的漏检情况。

3) 设计马赛克混合数据增强方法,利用4张训练图像进行裁剪拼接,形成1张新的训练图像,并将同一批次内不同的新的训练图像逐像素线性混合相加,建立数据样本之间的线性关系,增加数据集中图像的背景复杂度,提升数据集的泛化性,使模型可以在复杂背景的干扰下更专注于对目标物体的提取,提升了模型的鲁棒性。与YOLOv5x算法相比,本文算法更好地适用于无人机航拍图像中小目标的检测。

## 1 本文算法

本文提出的基于倒置残差注意力的无人机航拍图像小目标检测算法包括特征提取模块与多尺度特征融合模块,整体模型结构如图1所示。图中:SPP为空间金字塔池化。首先,通过特征提取模块从输入图像中提取低级空间信息和高级语义特征;然后,利用多尺度特征融合模块将提取的多层次特征进行融合,获得不同尺度的检测头,生成密集的边界框,并预测类别分数;最后,采用非极大值抑制算法消除冗余的预测候选框,得到最终结果。

### 1.1 特征提取模块

为进一步提高模型的特征提取能力,对YOLOv5x特征提取阶段的C3模块进行改进。将倒置残差模块嵌入到第3层和第5层的C3模块中,称为IRC3模块;将倒置残差注意力模块嵌入到第7层和

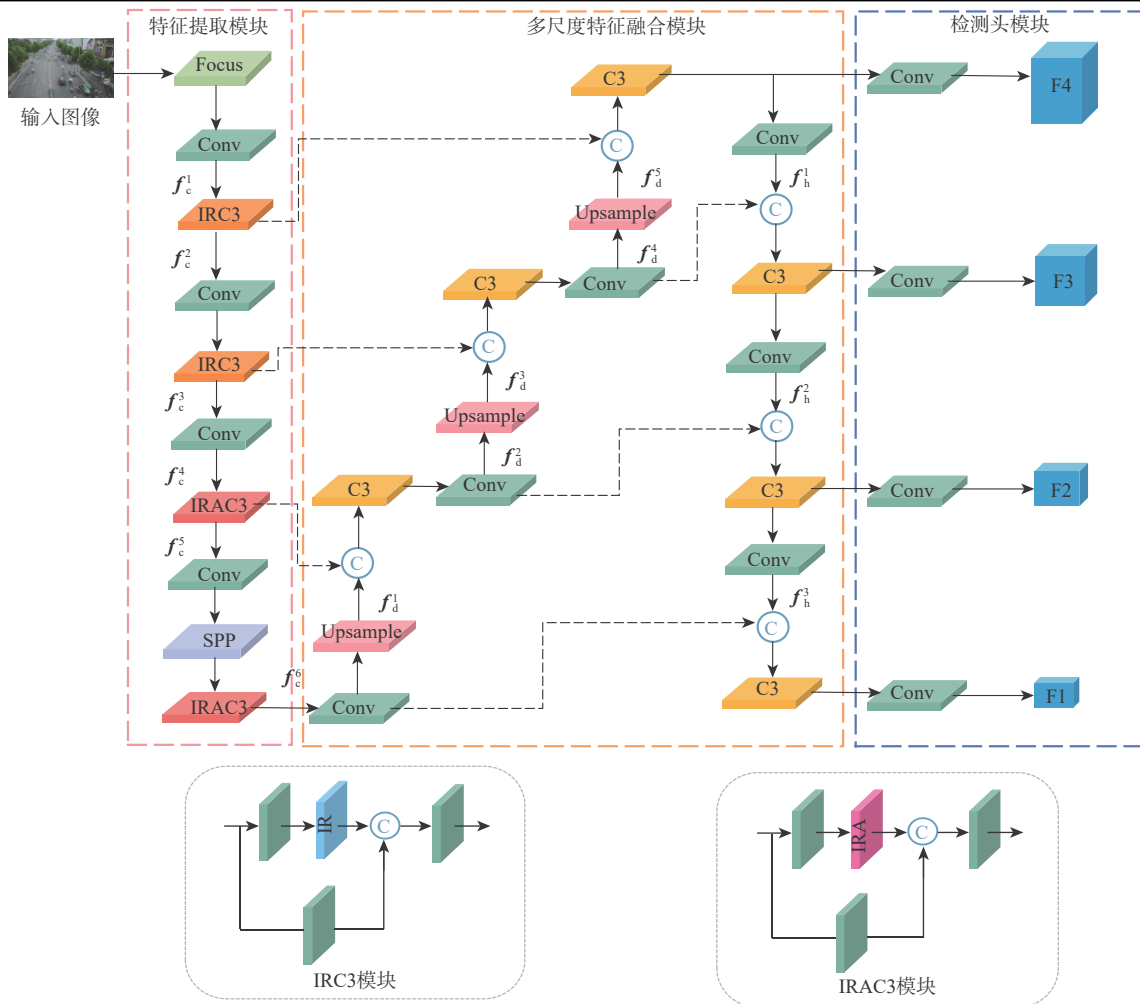


图1 基于倒置残差注意力的无人机航拍图像小目标检测模型结构

Fig. 1 Structure of small object detection in UAV aerial image based on inverted residual attention

第10层的C3模块中,称为IRAC3模块。通过引入本文设计的IRC3模块和IRAC3模块,提取不同尺度的特征,提高模型对特征的可分辨性,使得特征提取模块能够更有指向性地提取小目标的特征。

### 1.1.1 IRC3 模块

IRC3模块利用倒置残差模块将残差块中“降维-提取-升维”转换为“升维-提取-降维”的模式,以获取丰富浅层空间中微小目标的位置信息,模块结构如图2所示。

倒置残差模块的网络结构如图2中虚线框部分所示。首先,利用扩张因子 $a(a=2)$ 对输入特征图实现 $C \times a$ 的通道数扩张,获取丰富的浅层特征;然后,利用线性变换实现高维到低维通道的映射,

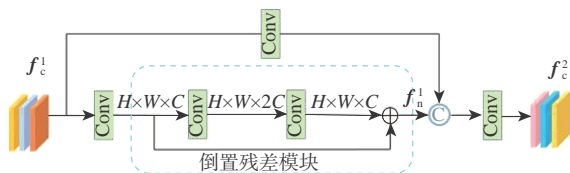


图2 IRC3 模块

Fig. 2 IRC3 module

有效防止通道压缩时因非线性激活函数引起的特征信息丢失问题;最后,通过残差操作将恒等映射与输入特征相结合。

对于输入的无人机航拍图像,首先通过卷积提取特征,获得的特征图 $f_c^1$ 可表示为

$$f_c^1 = \text{Conv}(\text{Conv}_{\text{Focus}}(x)) \quad (1)$$

式中:Conv表示卷积操作;Conv<sub>Focus</sub>表示切片卷积操作。

经IRC3模块后获得特征图 $f_c^2$ ,其获取过程可表示为

$$f_n^1 = \text{Conv}(\text{Conv}(\text{Conv}(f_c^1))) + \text{Conv}(f_c^1) \quad (2)$$

$$f_c^2 = \text{Conv}(F_c(\text{Conv}(f_c^1), f_n^1)) \quad (3)$$

式中: $f_n^1$ 表示在经过IRC3模块后,所得到的特征图; $F_c$ 表示级联操作(见图2中©符号所示)。经过2个IRC3模块后得到特征图 $f_c^3$ ,再经过一层卷积层得到特征图 $f_c^4$ ,送入IRAC3模块。

### 1.1.2 IRAC3 模块

IRAC3模块中,设计倒置残差注意力模块(见

图 3), 引入深度可分离卷积和 ECA-Net 模块。首先, 利用  $k \times k$  的卷积核扩展后的特征图在二维平面内进行逐通道深度卷积 (depthwise convolution, D-Conv) 操作 (见图 4(a)), 减少模型参数量。然后, 通过 ECA-Net 模块建立通道和权重的关系, 突出小目标特征的显著性, 通过跨信道交互的方式, 在显著降低模块复杂度的同时增强特征的表达能力, 提取

更丰富的小目标区域语义特征, 提升模型细节信息的提取能力。进一步, 利用  $1 \times 1$  的卷积对特征图进行逐点卷积 (pointwise convolution, P-Conv) 操作 (见图 4(b)), 将高维空间信息映射到低维空间, 减少特征丢失。最后, 进行随机深度 (stochastic depth, SD)<sup>[24]</sup> 计算, 减少前向传播过程和梯度计算, 缓解梯度消失现象, 提高整体网络的泛化性。

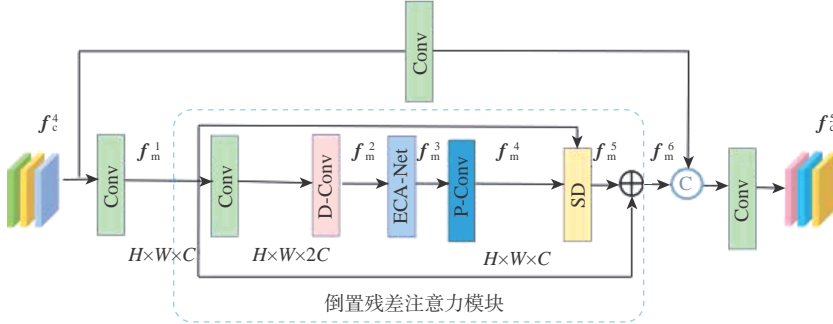


图 3 IRAC3 模块  
Fig. 3 IRAC3 module

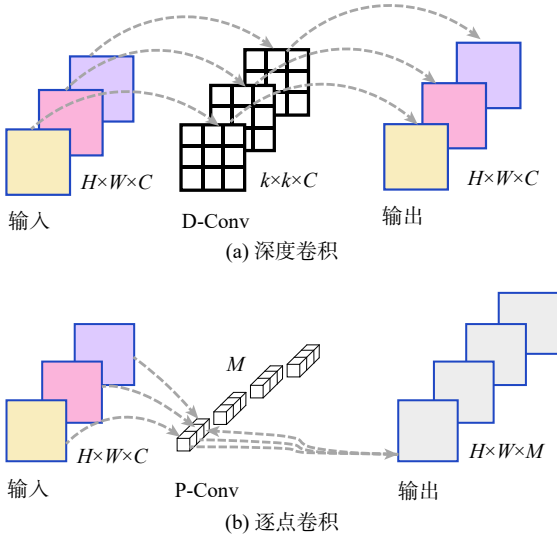


图 4 深度可分离卷积模块  
Fig. 4 Depthwise separable convolution module

倒置残差注意力模块的特征提取过程具体可分为如下 4 步完成:

1) 为进一步减少模型参数和提高检测效率, 该模块设计使用参数量少且计算复杂度低的深度卷积替代  $3 \times 3$  的标准卷积。将特征图  $f_m^4$  输入该模块后, 先通过卷积得到特征图  $f_m^1$ , 在倒置残差注意力模块中经过卷积和深度卷积得到特征图  $f_m^2$ , 如下:

$$\hat{G}_{q,l,c} = \sum_{i,j} \hat{K}_{i,j,c} \text{Conv}(f_m^1)_{q+i-1,l+j-1,c} \quad (4)$$

式中:  $\hat{K}$  为深度卷积核;  $i, j$  为卷积核大小;  $q, l$  为特征图大小;  $\hat{K}$  中的第  $c$  个卷积核应用于与之相乘特征中的第  $c$  个通道,  $\hat{G}$  为  $f_m^2$  的第  $c$  个通道。

2) ECA-Net 模块如图 5 所示, 其提升了通道特征利用率, 并突出了小目标的显著性。

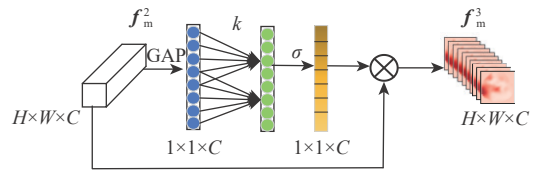


图 5 ECA-Net 模块  
Fig. 5 ECA-Net module

将特征图  $f_m^2$  输入到 ECA-Net 模块中, 首先, 该特征图在不降低维度的情况下进行逐通道全局平均池化 (global average pooling, GAP), 进而利用共享权重的快速一维卷积进行特征学习, 在特征学习过程中通过当前通道及  $k$  个邻域通道来捕获局部跨通道信息交互, 其中,  $k$  为一维卷积核大小, 同时也表示局部跨通道交互覆盖率; 然后, 通过 sigmoid 激活函数获得相应通道的权重; 最后, 将其与输入特征相乘作为下一层的输入:

$$\omega = \sigma(\text{Conv}_{1D}(H_{\text{GAP}}(f_m^2))) \quad (5)$$

式中:  $\text{Conv}_{1D}$  表示一维卷积;  $H_{\text{GAP}}$  表示全局平均池化;  $\sigma$  表示 sigmoid 激活函数。通过激活函数后获得相应的权重  $\omega$  与特征图  $f_m^2$  相乘, 得到特征图  $f_m^3$  为

$$f_m^3 = f_m^2 \cdot \omega \quad (6)$$

这种捕捉跨通道信息交互的方式选择性地增强特征, 提高了模型对航拍图像细节信息的提取能力, 进而提高模型的检测精度。

3) 利用逐点卷积对特征图  $f_m^3$  应用单个卷积核, 来共同创建一个深度输出的线性组合, 得特征图

$f_m^4$ 为

$$f_m^4 = \text{Conv}_{P\text{-Conv}}(f_m^3) \quad (7)$$

深度卷积与逐点卷积构成深度可分离卷积。当输入数据为 $H \times W \times C$ 时,共有 $M$ 个大小为 $k \times k$ 的卷积核,深度可分离卷积的参数量为

$$P_{\text{DS-Conv}} = k \times k \times C + C \times M \quad (8)$$

标准卷积的参数量为

$$P_{\text{Conv}} = k \times k \times C \times M \quad (9)$$

深度可分离卷积与标准卷积参数量比值为

$$N_P = \frac{P_{\text{DS-Conv}}}{P_{\text{Conv}}} = \frac{1}{M} + \frac{1}{k^2} \quad (10)$$

式中: $C$ 为输入通道数; $M$ 为输出通道数; $P_{\text{DS-Conv}}$ 为深度可分离卷积的参数量; $P_{\text{Conv}}$ 为标准卷积的参数量; $N_P$ 为倍数。

由式(10)可知,当输出特征的通道数很大时,深度可分离卷积参数量大约为传统卷积的 $1/k^2$ ,一般 $k > 1$ 。可以证明,相较传统卷积,利用深度可分离卷积可以减少模型参数量,提升无人机航拍图像目标检测的检测速度。

4) 由于学习深度的增加,特征信息在训练的过程中被提取过滤,当特征信息到达深层时会逐渐衰减,使特征信息难以被有效训练。通过引入随机深度模块在不激活一部分模块的情况下,促使高层的模块能接收更多来自浅层的信息,提高模型的特征表达能力。在随机深度模块中,引入随机变量 $b_l, b_l$ 表示伯努利随机变量,将 $f_m^4$ 乘以 $b_l$ ,对残差部分进行随机丢弃,如下:

$$f_m^5 = \text{ReLU}(b_l f_m^4 + f_m^1) \quad (11)$$

式中: $\text{ReLU}$ 代表激活函数。当 $b_l = 1$ 时,将这2部分求和后经过激活函数输出,当 $b_l = 0$ 时,残差部分没有被激活,整个结构为一个恒等映射:

$$f_m^5 = \text{ReLU}(f_m^1) \quad (12)$$

为保留浅层特征,以便浅层特征被更好的利用,利用线性衰减规律得到每一层的生存概率:

$$P_l = 1 - \frac{l}{L}(1 - P_L) \quad (13)$$

式中: $L$ 为残差模块的总数量; $P_l$ 为第 $l$ 层在训练中的生存概率,并满足线性衰减规律,即从 $P_0 = 1$ 线性衰减至 $P_L = 0.5$ 决定了随机变量 $b_l$ 的取值。选取更显著的浅层特征生成深层特征,增强了反向传播的梯度,较少的网络深度减少训练时间,有效缓解了过拟合的问题,最终经过 $1 \times 1$ 卷积得出特征 $f_c^5$ :

$$f_c^5 = \text{Conv}(F_c(\text{Conv}(f_c^4), f_m^6)) \quad (14)$$

## 1.2 多尺度特征融合模块

无人机航拍图像中小尺寸目标居多且目标之

间比较密集,如图6所示。按照相对尺寸大小,绘制了数据集的目标分布图,如图7所示。可以看出,目标分布图中的目标尺寸与图像尺寸相比小于0.12%的小目标数量占总目标数量的72%。浅层特征图分辨率较大包含空间特征细节更丰富,但感受野小所反映出的语义表征能力较弱。随着网络层数的加深,深层卷积在提取小目标特征信息时容易丢失关键位置信息,但感受野较大能获得高级语义信息。因此,为进一步加强对小目标的检测性能,本文提出了多尺度特征融合模块,将浅层空间信息和深层语义相融合。在多尺度特征融合模块中,在原始YOLOv5x模型3个检测头的基础上增加1个检测头,生成了4种不同尺度的检测头,分别用于无人机航拍图像中极小目标、小目标、中目标及大目标的检测。

在多尺度特征融合模块中,如图1虚线框所示,先通过3次上采样(Upsample)得到特征图 $f_d^3$ ,再与主干网络中 $f_c^2$ 进行级联操作,最终经过C3模块和卷积层进行特征提取获得特征图 $F_4$ ,其特征尺寸为原图像的1/4,用于极小目标的检测。同时,将特征提取模块中的特征图 $f_c^3$ 与多尺度特征融合模块的特征图 $f_d^3$ 级联,经过C3模块和卷积层后得到特征图 $f_d^4$ ,进一步再与特征图 $f_m^1$ 进行级联操作后再



图6 无人机航拍图像

Fig. 6 UAV aerial image

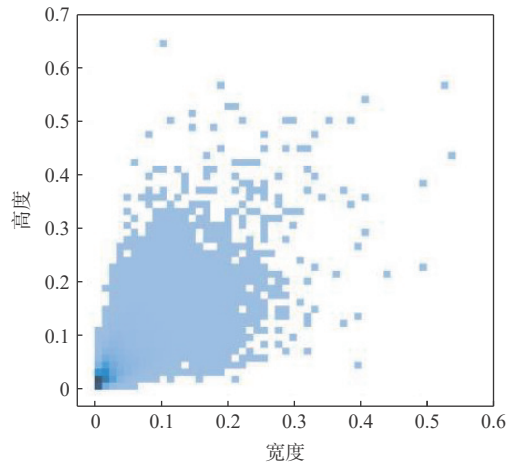


图7 目标分布图像

Fig. 7 Object distribution image

次经过 C3 模块和卷积层获得特征图 F3, 其特征图尺寸为原图像的 1/8, 用于小目标的检测。同理, 通过这种方式还可以获得特征图 F2 和 F1, 其特征尺寸为原图像的 1/16 和 1/32, 用于中目标和大目标的检测, 特征图 F1、F2、F3、F4 的具体过程如下:

$$F1 = \text{Conv}(F_{C3}(F_c(f_h^3, \text{Conv}(f_c^6)))) \quad (15)$$

$$f_d^2 = \text{Conv}(F_{C3}(F_c(f_h^1, f_c^5))) \quad (16)$$

$$F2 = \text{Conv}(F_{C3}(F_c(f_h^2, f_d^2))) \quad (17)$$

$$f_d^4 = \text{Conv}(F_{C3}(F_c(f_h^3, f_c^3))) \quad (18)$$

$$F3 = \text{Conv}(F_{C3}(F_c(f_h^1, f_d^4))) \quad (19)$$

$$F4 = \text{Conv}(F_{C3}(F_c(f_d^5, f_c^2))) \quad (20)$$

式中:  $F_{C3}$  为 C3 模块操作。

通过多尺度特征融合的方式, 将浅层网络中丰富的位置信息和纹理信息更好地与深层网络的语义特征信息相融合, 增强模型在小目标下的多尺度特征学习能力, 从而提升模型在复杂场景下的小目标检测能力。

### 1.3 数据增强模块

为提升复杂背景干扰下的小目标检测精度, 进而提升模型的鲁棒性, 本文在 YOLOv5x 原有的马赛克数据增强方法上增加混合<sup>[25]</sup>数据增强方法, 采用马赛克数据增强和混合数据增强相结合的方式, 称为马赛克混合数据增强方法。

该方法首先确定高为  $h$ 、宽为  $w$  的图像模板作为输出图像尺寸, 同时在宽高方向随机生成 2 条分割线, 将选取的 4 张图像按照图像模板经过随机裁剪后进行拼接操作。同理, 再选取另外 4 张图像进行随机裁剪并进行拼接, 拼接后的图像形成新的训练样本。其主要利用随机裁剪丰富了数据集中目标的特征, 使模型更易于学习, 并通过拼接的方式保留了图像的目标特征, 极大程度地丰富了检测目标的背景, 有效减少了由于训练背景相似导致模型泛化性降低的问题。然后通过逐像素线性相加将 2 类图像混合, 融合示意图如图 8 所示。该样本的混合过程是利用贝塔分布生成的融合系数进行图像融合, 生成后的融合训练样本在尺寸上与原始的训练样本相同。本文通过阈值对生成的融合训练

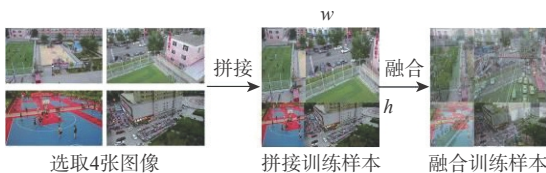


图 8 融合增强方法过程

Fig. 8 Fusion enhancement method process

样本进行控制, 由于在融合过程中每个批次样本都会随机产生相应的权重, 而权重在  $N$  个批次中的期望值近似为 0.5, 将本文阈值设置为 0.5。融合过程为

$$\lambda = \text{Beta}(\alpha, \beta) \quad (21)$$

$$\bar{x} = \lambda x_i + (1 - \lambda) x_j \quad (22)$$

$$\bar{y} = \lambda y_i + (1 - \lambda) y_j \quad (23)$$

式中:  $x_i, x_j$  代表同一批次内不同的训练样本;  $y_i, y_j$  分别对应该样本的标签;  $\lambda$  为由参数  $\alpha, \beta$  的贝塔分布计算出的混合系数, 服从  $\text{Beta}(\alpha, \beta)$  分布;  $\bar{x}$  为混合后的批次样本;  $\bar{y}$  为混合后的批次样本对应的标签。

通过该方法生成的训练样本计算量小, 且扩展了训练数据的空间分布, 在保持检测速度不变的情况下, 降低了不同分辨率下识别能力的衰弱速度, 提高了模型对航空图像中目标检测的泛化性, 提升了模型的鲁棒性。

## 2 实验结果与分析

### 2.1 数据集

本文在无人机航拍图像目标检测公开数据集 Vis-Drone 上进行了训练和测试, 该数据集包含 10 209 张图像, 其中, 包括 6 471 张训练集图像, 548 张验证集图像, 3 190 张测试集图像, 并对 10 个类别的对象进行了丰富的标注, 包括行人 (Pedestrian)、人 (Person)、汽车 (Car)、货车 (Van)、公共汽车 (Bus)、卡车 (Truck)、摩托车 (Motor)、自行车 (Bicycle)、遮阳篷三轮车 (Awning-tricycle) 和三轮车 (Tricycle)。该数据集的图像分为  $1\,360 \times 765$  像素和  $960 \times 540$  像素 2 种不同的图像尺寸, 本文使用验证集来评估提出的算法。

### 2.2 实验环境

实验采用的硬件环境为: GPU 为 NVIDIA TITAN V12 GB, CPU 为 Intel Xeon(R) Gold 5115 CPU @ 2.40 GHz; 软件环境为 Windows 7 操作系统, 选用 Pytorch 为深度学习框架。采用 YOLOv5x 进行实验, 分别在训练集和验证集上进行训练和测试, 输入图像尺寸大小为  $832 \times 832$ , 批次为 2, Epochs 为 120, 初始学习率为 0.01, 终止学习率为 0.2, 采用随机梯度下降策略, 动量和权重衰减分别为 0.937 和 0.000 5。

### 2.3 评价指标

本文使用平均精度 (AP)、平均精度均值 (mAP0.5: 0.95) 和模型参数作为衡量模型性能的相关指标。AP 为单一类别检测精度的评价指标。对各类别的 AP 值相加再除以类别数得到 mAP。mAP0.5: 0.95 的计算方法是: 设置 10 个 IoU 阈值, 从 0.5 到 0.95, 步长为 0.05, 并计算每个 IoU 对应的 mAP, 再对所

有 mAP 求平均值。本文将 mAP<sub>0.5:0.95</sub> 简称为 mAP。文中 mAP<sub>0.5</sub> 为 IoU=0.5 时的 mAP, mAP<sub>0.75</sub> 为 IoU=0.75 时的 mAP, IoU 为交并比。AP 和 mAP 计算过程如下:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (24)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (25)$$

$$\text{AP} = \int_0^1 \text{Precision}(\text{Recall}) d(\text{Recall}) \quad (26)$$

$$\text{mAP} = \frac{\sum_{i=1}^N \text{AP}_i}{N} \quad (27)$$

式中:  $N$  为类别个数; TP 表示正样本被正确标识为正样本, 即真阳性; FP 表示负样本被错误标识为正样本, 即假阳性; FN 表示正样本被错误识别为负样本, 即假的负样本。

## 2.4 总体性能分析

为验证本文算法的有效性, 针对不同模型进行了消融实验。模型 1 只采用多尺度特征融合模块进行训练; 模型 2 在模型 1 的基础上添加数据增强模块; 模型 3 在模型 2 的基础上添加倒置残差注意力模块; 模型 4 在模型 3 的基础上添加倒置残差模块, 即为本文提出的最终模型。如表 1 所示, 所有实验在 VisDrone 数据集上进行测试, 分别以 mAP、mAP<sub>0.5</sub>、mAP<sub>0.75</sub>、参数量及检测速度作为衡量标准。表中: FPS 表示帧/s。以原始 YOLOv5x 为基准, 分别添加相应的模块改进, 通过对客观评价指标的计算比较不同模型性能。

由表 1 可以看出, 模型 1 在加入多尺度特征融合模块后, 增强了模型在小目标下的多尺度特征学习能力, 其 mAP 相比于基准模型 YOLOv5x 提高了 1.2%, 由于小目标检测层的增加, 参数量略有提升, 检测速度有所下降。模型 2 相比模型 1 的 mAP 提高了 0.8%, 其参数量和检测速度基本不变, 验证了

表 1 不同模型的客观指标对比

Table 1 Comparison of objective indicators of different models

模型	mAP/%	mAP <sub>0.5</sub> /%	mAP <sub>0.75</sub> /%	参数量/ $10^6$	检测速度/FPS
YOLOv5x	23.4	35.7	25.1	83.2	41.3
模型 1	24.6	38.6	26.2	86.7	28.7
模型 2	25.4	39.7	27.1	85.5	28.7
模型 3	26.8	41.4	28.8	69.3	25.6
模型 4	27.4	42.4	29.0	72.5	23.4

数据增强模块的有效性。模型 3 在加入倒置残差注意力模块后, 其 mAP 相比模型 2 提升了 1.4%, 并且参数量降低了  $16.2 \times 10^6$ , 由于倒置残差注意力模块中深度可分离卷积具有大量数据读写操作, 在检测速度方面略有降低。模型 4 相比模型 3, 虽然参数量略有提升, 检测速度略有降低, 但 mAP 提升了 0.6%, 进一步验证了在使用倒置残差注意力模块的基础上增加倒置残差模块的有效性。综上所述, 在 4 个模块同时加入时达到的效果最佳, mAP 相比于基准模型 YOLOv5x 提升 4.0%, 参数量下降  $10.7 \times 10^6$ , 进一步验证了模型的有效性。

从图 9 中可以看出, YOLOv5x 基准模型存在漏检情况, 对小目标检测效果不佳, 模型 1 和模型 2 降低了部分小尺寸目标的漏检情况, 仍存在漏检和误检问题。模型 3 中误检问题减少, 并进一步降低了小目标漏检情况, 相比之下, 模型 4 达到了较好的检测效果, 较大程度上降低了小目标漏检。

## 2.5 算法对比分析

本节将本文模型与当前经典的 RetinaNet、Faster R-CNN 和 Cascade R-CNN 模型进行了比较, 并与这些比较算法增加 MMF<sup>[26]</sup>、SimCal<sup>[27]</sup>、BGS<sup>[28]</sup> 和 DSHNet<sup>[29]</sup> 后的模型进行了比较。采用不同的主干网络(backbone), 平均精度均值 mAP 的对比结果如表 2 所示。在检测速度方面, 将本文模型与当前经

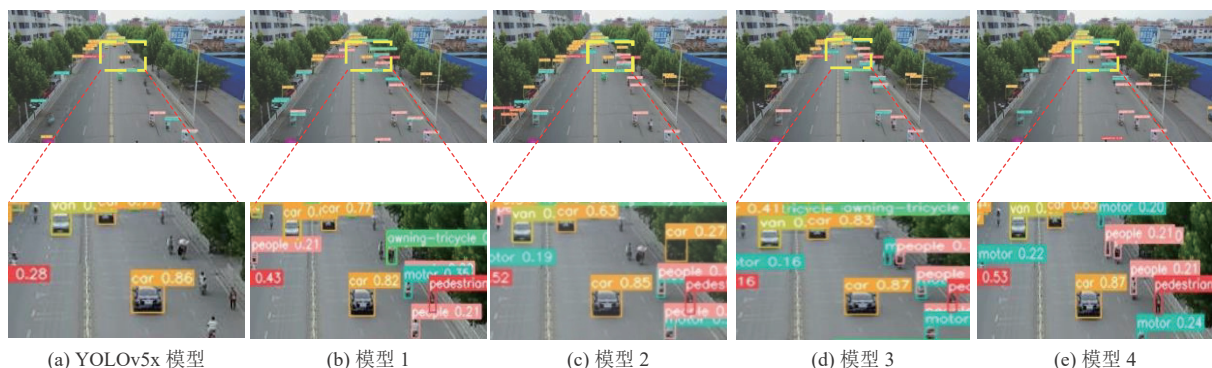


图 9 不同模型检测结果

Fig. 9 Detection results of different models

典的 RetinaNet、Faster R-CNN和 Cascade R-CNN 在此基础上增加 DSHNet 后的模型进行了比较, 实验同样在 VisDrone 数据集上进行训练, 采用相同的验证集图像进行验证, 在相同硬件环境 NVIDIA

TITAN V12 GB GPU 上进行测试, 其模型的平均精度均值 mAP 与检测速度的对比结果如表 3 所示。视觉效果如图 10 所示, 展示了 3 幅图像的检测结果及其部分放大图。

表 2 不同算法的检测结果对比

Table 2 Comparison of detection results of different algorithms

算法	backbone	mAP/%	AP/%									
			Pedestrian	Person	Bicycle	Car	Van	Truck	Tricycle	Awning-tricycle	Bus	Motor
RetinaNet	R50	13.9	13.0	7.9	1.4	45.5	19.9	11.5	6.3	4.2	17.8	11.8
Faster R-CNN	X101	22.4	21.3	15.5	7.9	52.0	29.5	20.5	14.7	8.9	32.1	21.6
Cascade R-CNN	R50	23.2	22.2	14.8	7.6	54.6	31.5	21.6	14.8	8.6	34.9	21.4
Faster R-CNN+MMF	R50	22.6	21.6	15.3	9.6	51.5	28.5	20.4	15.9	7.5	33.7	21.6
Faster R-CNN+SimCal	R50	20.0	18.7	13.8	5.7	51.0	28.4	16.4	13.6	5.9	27.0	19.4
Faster R-CNN +BGS	R50	23.0	21.8	16.0	8.1	51.8	31.1	19.8	15.0	8.4	36.1	21.5
RetinaNet+DSHNet	R50	16.1	14.1	8.9	1.3	48.2	24.8	14.2	8.8	6.0	21.6	13.1
Faster R-CNN+DSHNet	R50	24.6	22.5	16.5	10.1	52.8	32.6	22.1	17.5	8.8	39.5	23.7
Faster R-CNN+DSHNet	X101	25.8	23.3	16.7	11.4	53.7	33.1	23.8	19.5	11.1	40.0	25.5
Cascade R-CNN+DSHNet	R50	26.2	23.2	16.1	11.2	55.5	33.5	25.2	19.1	10.0	43.0	25.1
本文模型	CSPDarknet53	27.4	28.9	6.0	9.5	60.1	36.2	34.5	16.7	17.2	47.2	18.0

表 3 不同算法的平均精度均值与检测速度结果对比

Table 3 Comparison of average accuracy and detection speed of different algorithms

算法	backbone	mAP/%	检测速度/FPS
RetinaNet+DSHNet	R50	16.1	19
Faster R-CNN+DSHNet	R50	24.6	22.5
Cascade R-CNN+DSHNet	R50	26.2	15
本文模型	CSPDarknet53	27.4	23.4

遮阳篷三轮车等小目标类别中取得了优秀的检测指标; 另一方面, 相比最优的 DSHNet 模型, mAP 增长了 1.2%。由表 3 看出, 本文模型在平均精度均值和检测速度方面均得到了有效提升。从图 10 可以看出, 对于小目标密集区域的检测结果, DSHNet 存在漏检情况, 而本文模型能够在小目标密集区域充分挖掘其特征信息, 对于小目标密集区域的检测结果要优于 DSHNet 模型, 降低了漏检情况, 在处理无人机航拍图像目标检测任务时具有较大优势。

2.6 结果分析

为进一步验证模型的有效性, 图 11 给出了本文模型在不同背景下的部分航空图像目标检测结果。其中, 图 11(a)为背景简单、目标单一时的检

从表 2 可以看出, 在经典 RetinaNet、Faster R-CNN 和 Cascade R-CNN 基础上增加 DSHNet 模型后, 客观评价指标较高; 与 DSHNet 模型相对比, 本文模型取得了最优的综合性能。一方面, 在行人、汽车、

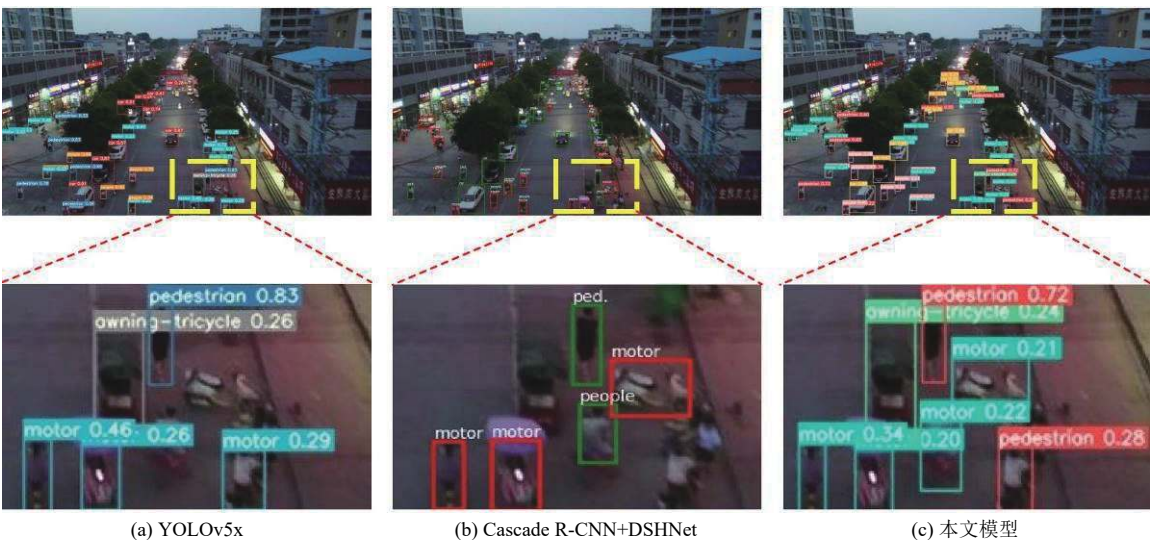


图 10 三种模型检测结果

Fig. 10 Detection results of three models

测结果及区域放大图;图 11(b)为背景复杂、小目标密集时的检测结果及区域放大图。从图 11 可以看出,对于以下这 2 种情况,本文模型都可以较好地完成目标检测,对于场景复杂、目标较小的情况

下仍能完全检测,没有出现汽车和行人的误检情况。综上所述,从主观视觉和评价指标来看,目标检测准确率有所提升,对于背景复杂、密集小目标的漏检情况有所改善。

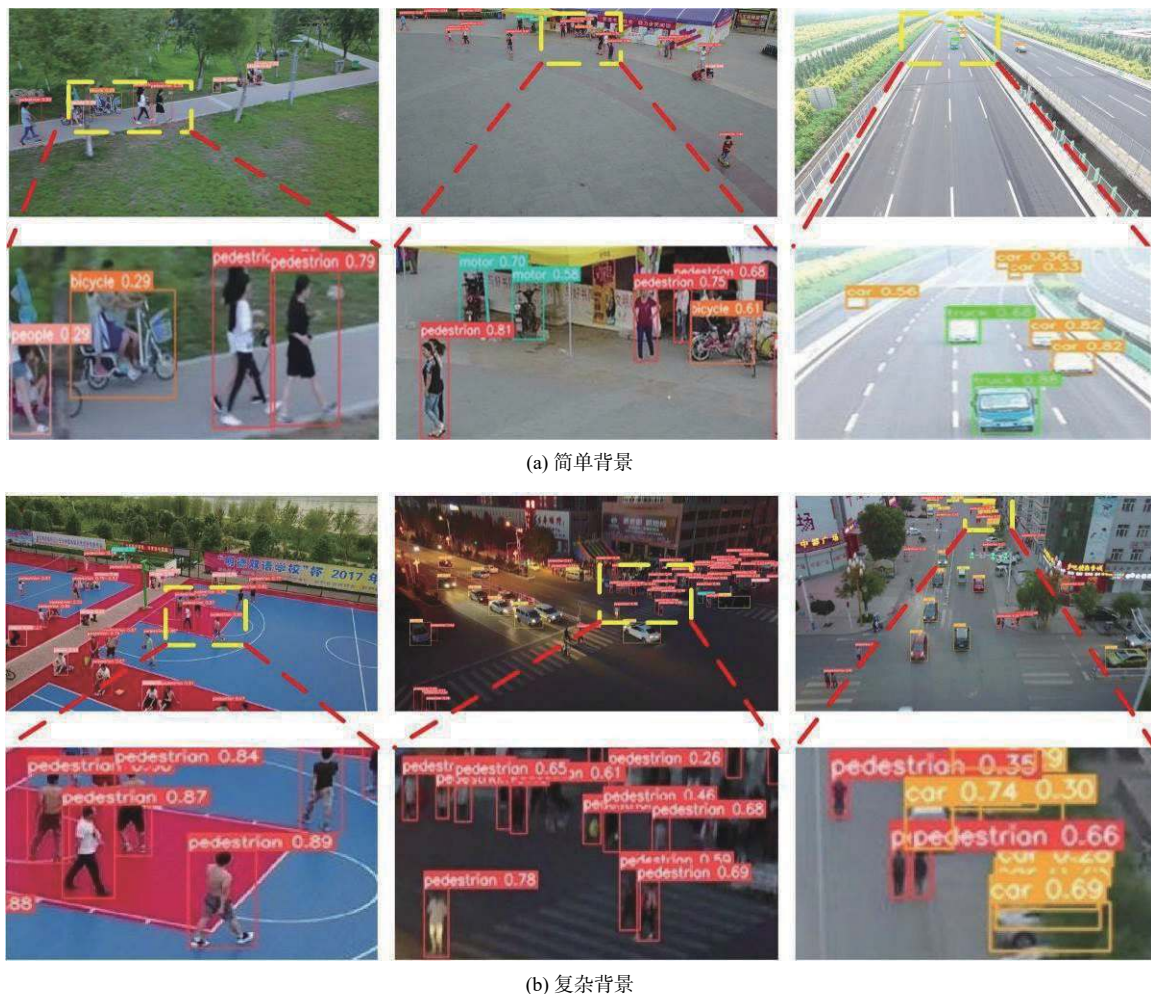


图 11 本文模型在不同背景下的检测结果

Fig. 11 Detection results of the proposed model under different backgrounds

### 3 结论

针对无人机航拍图像中存在的小尺寸目标检测精度低的问题,本文提出一种基于倒置残差注意力的无人机航拍图像小目标检测算法。

1) 将倒置残差和倒置残差注意力模块添加到 YOLOv5x 特征提取阶段的 CSPDarknet53(C3)模块中,获取丰富的空间信息和语义特征,使模型更多地关注图像中的小目标区域,同时利用通道之间的信息交互增强了模型的特征表达能力,提升了小目标的检测精度。

2) 设计了多尺度特征融合模块,将不同感受野的浅层空间信息和深层语义信息相融合,有效改善了小目标的漏检问题。

3) 通过马赛克混合数据增强方法,对混合后的训练样本进行线性混合相加,丰富了训练样本的多

样性,增强了模型在复杂背景干扰下目标特征提取的能力。

4) 实验结果表明,在背景复杂的情况下,本文算法对小尺寸目标具有更强的辨识能力,平均精度均值方面达到最优,相比 DSHNet 算法提升了 1.2%,减少了漏检和误检。

下一步将继续研究高效目标检测算法,保证精准度的同时,进一步提升检测的实时性。

### 参考文献 (References)

[1] WU X, LI W, HONG D, et al. Deep learning for UAV-based object detection and tracking: A survey[EB/OL]. (2021-10-25)[2022-05-01].<https://arxiv.org/abs/2110.12638>.

[2] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]/Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2014:

- 580-587.
- [3] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2015: 1440-1448.
- [4] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. Piscataway: IEEE Press, 2015: 91-99.
- [5] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2961-2969.
- [6] CAI Z, VASCONCELOS N. Cascade R-CNN: Delving into high quality object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 6154-6162.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 779-788.
- [8] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 7263-7271.
- [9] REDMON J, FARHADI A. YOLOv3: An incremental improvement[EB/OL]. (2018-04-08)[2022-05-01].<https://arxiv.org/abs/1804.02767>.
- [10] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2022-05-01].<https://arxiv.org/abs/2004.10934?sid=NDAqzT>.
- [11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2016: 21-37.
- [12] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2980-2988.
- [13] LIU M, WANG X, ZHOU A, et al. UAV-YOLO: Small object detection on unmanned aerial vehicle perspective[J]. *Sensors*, 2020, 20(8): 2238.
- [14] LIANG X, ZHANG J, ZHUO L, et al. Small object detection in unmanned aerial vehicle images using feature fusion and scaling-based single shot detector with spatial context analysis[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2019, 30(6): 1758-1770.
- [15] ZHANG P, ZHONG Y, LI X. SlimYOLOv3: Narrower, faster and better for real-time UAV applications[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2019: 37-45.
- [16] 裴伟, 许晏铭, 朱永英, 等. 改进的 SSD 航拍目标检测方法[J]. *软件学报*, 2019, 30(3): 738-758.
- PEI W, XU Y M, ZHU Y Y, et al. The target detection method of aerial photography images with improved SSD[J]. *Journal of Software*, 2019, 30(3): 738-758(in Chinese).
- [17] 刘婷婷, 苗华, 李琳, 等. 融合场景上下文的轻量级目标检测网络[J]. *激光与光电子学进展*, 2021, 58(20): 127-135.
- LIU T T, MIAO H, LI L, et al. Lightweight target detection network integrating scene context[J]. *Laser & Optoelectronics Progress*, 2021, 58(20): 127-135(in Chinese).
- [18] HOWARD A, SANDLER M, CHEN B, et al. Searching for mobileNetV3[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2019: 1314-1324.
- [19] TAN M, LE Q. EfficientNetV2: Smaller models and faster training[EB/OL].(2021-06-23)[2022-05-01].<https://arxiv.org/abs/2104.00298v2>.
- [20] 刘艳菊, 王秋霖, 赵开峰, 等. 基于卷积神经网络的热轧钢条表面实时缺陷检测[J]. *仪器仪表学报*, 2021, 42(12): 211-219.
- LIU Y J, WANG Q J, ZHAO K F, et al. Real-time defect detection of hot rolling steel bar based on convolution neural network[J]. *Chinese Journal of Scientific Instrument*, 2021, 42(12): 211-219(in Chinese).
- [21] 周中, 张俊杰, 龚琛杰, 等. 基于深度语义分割的隧道渗漏水智能识别[J]. *岩石力学与工程学报*, 2022, 41(10): 2082-2093.
- ZHOU Z, ZHANG J J, GONG C J, et al. Automatic identification of tunnel leakage based on deep semantic segmentation[J]. *Chinese Journal of Rock Mechanics and Engineering*, 2022, 41(10): 2082-2093(in Chinese).
- [22] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 11534-11542.
- [23] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications[EB/OL]. (2017-04-17)[2022-05-01].<https://arxiv.org/abs/1704.04861>.
- [24] HUANG G, SUN Y, LIU Z, et al. Deep networks with stochastic depth[C]//Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2016: 646-661.
- [25] ZHANG H, CISCHE M, DAUPHIN Y N, et al. Mixup: Beyond empirical risk minimization[EB/OL]. (2018-04-27)[2022-05-01].<https://arxiv.org/abs/1710.09412>.
- [26] ZHANG X, IZQUIERDO E, CHANDRAMOULI K. Dense and small object detection in UAV vision based on cascade network[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2019: 118-126.
- [27] WANG T, LI Y, KANG B, et al. The devil is in classification: A simple framework for long-tail instance segmentation[C]//Proceedings of the European Conference on Computer Vision. Berlin: Springer, 2020: 728-744.
- [28] LI Y, WANG T, KANG B, et al. Overcoming classifier imbalance for long-tail object detection with balanced group softmax[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 10991-11000.
- [29] YU W P, YANG T J N, CHEN C. Towards resolving the challenge of long-tail distribution in UAV images for object detection[C]//Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision. Piscataway: IEEE Press, 2021: 3258-3267.

## Small object detection in UAV aerial images based on inverted residual attention

LIU Shudong, LIU Yehui, SUN Yemei, LI Yifei, WANG Jiao\*

(School of Computer and Information Engineering, Tianjin Chengjian University, Tianjin 300384, China)

**Abstract:** Aiming at the problems of complex background and too many small-size targets in UAV aerial images, a small target detection algorithm based on inverted residual attention is proposed. Firstly, an inverted residual module and an inverted residual attention module are embedded into the backbone network, while rich spatial information and deep semantic information of small targets are obtained by feature information mapping from low dimension to high dimension, thus improving the accuracy of small target detection; Secondly, in feature fusion, a multi-scale feature fusion module is established to fuse the shallow spatial information and deep semantic information, and to generate four detection heads with different sensory fields, which improves the recognition of small-size targets and reduces missed detection of small targets; Finally, a mosaic mixed data enhancement method is designed to establish the linear relationship between the data, increase the complexity of the image background and improve the robustness of the algorithm. The experimental results on data set VisDrone show that the mean average precision of this algorithm is 1.2% higher than that of DSHNet, which means that the proposed algorithm could effectively reduce missed detection and false detection of small targets in UAV aerial images.

**Keywords:** object detection; UAV images; inverted residual; attention; multi-scale feature fusion

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0288

# 单目标概率约束规划的微种群免疫优化算法

李静<sup>1</sup>, 张仁崇<sup>1,\*</sup>, 潘春燕<sup>2</sup>, 杨凯<sup>3</sup>

(1. 贵州商学院 计算机与信息工程学院, 贵阳 550014; 2. 黔南民族师范学院 数学与统计学院, 都匀 558000;

3. 贵州大学 贵州省大数据产业研究院, 贵阳 550025)

**摘 要:** 针对无先验随机分布信息的单目标概率约束规划, 探讨了微种群免疫优化算法。算法设计中, 受危险理论启发设计微种群免疫优化算法进化框架; 借助估计值的误差幅度, 提出2个方法分别估计概率值和目标值; 依据个体间的优劣关系, 划分群体为3个类型子群协同进化; 构建生命周期模型, 设计自适应的交叉与变异概率、变异策略, 结合交叉算子促进子群信息有效交流, 并沿不同方向协同进化。数值实验统计结果说明: 所提算法拥有良好的搜索效率、搜索效果及降噪能力, 具有一定的竞争力和应用潜力。

**关键词:** 概率约束规划; 免疫优化; 危险理论; 自适应采样; 微种群

**中图分类号:** TP301.6

**文献标志码:** A

**文章编号:** 1001-5965(2023)03-0525-13

在实际工程优化过程中, 涌现出许多设计问题, 如锌精矿混合成分<sup>[1]</sup>、第四方物流路径优化<sup>[2]</sup>、鲁棒动态调度<sup>[3-4]</sup>、电力系统故障诊断<sup>[5]</sup>、分布式微电网优化<sup>[6]</sup>、水资源管理<sup>[7]</sup>等, 这些问题均被描述为单目标概率约束规划 (single objective probabilistic constrained programming, SOPCP), 其模型主要有以下5个特征: ①随机变量(噪声)一般呈复杂或未知分布; ②目标函数在给定置信水平下由概率不等式确定; ③约束条件至少含一个概率约束或一个机会约束; ④随机函数通常是非线性、多模态或非凸; ⑤决策变量通常在非凸区域取值。SOPCP由于存在以上特征, 导致其难以进行求解。因此, 探讨出具有效率和效果突出、应用潜力高、抑噪能力强等优势的先进智能优化技术求解 SOPCP, 已成为亟待解决的技术难关。

在噪声处理方面, 若噪声先验概率分布已知时, 如噪声服从指数分布、正态分布、某些特殊非正态分布、均匀分布等, 则凸性近似<sup>[8-9]</sup>、对数凹变换<sup>[10]</sup>、

离散阶跃变换<sup>[11]</sup>等数学方法能转化部分 SOPCP 为确定性解析模型, 从而借助现有规划方法和静态智能算法求解, 但由于模型转换复杂且需已知先验概率分布, 致使它们在实际应用中受到很大限制。若噪声分布复杂或未知时, 蒙特卡罗随机模拟因不受限噪声分布而被广泛用于 SOPCP 的求解中。就其样本采样方法而言, 重要采样<sup>[12]</sup>、拉丁超立方采样<sup>[12]</sup>均可处理机会约束, 但此类方法实现较为困难且效率低下; 静态采样、动态采样<sup>[13]</sup>、自适应采样<sup>[14-20]</sup>是处理机会约束和估计目标函数值的常用方法。静态采样要求所有个体都附加相同且足够大的样本量以获得近似解, 尽管该方法简单, 但会导致昂贵的计算资源消耗; 动态采样要求个体依据其质量动态获取样本量, 该方法样本总量小、效率高, 但设计困难; 自适应采样优于静态采样和动态采样, 其要求个体自适应附加样本量, 即质量越高的个体附加的样本量越大, 反之则个体附加的样本量越小, 这使得其在随机优化的背景下变得越来越流行。

收稿日期: 2021-06-02; 录用日期: 2021-09-24; 网络出版时间: 2021-10-12 09:05

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20211011.1811.002.html](https://kns.cnki.net/kcms/detail/11.2625.V.20211011.1811.002.html)

基金项目: 贵州省科技计划(黔科合基础[2020]1Y423, 黔科合基础[2019]1178); 贵州省大数据应用工程研究中心(黔教合KY字[2017]022); 贵州省教育厅青年科技人才成长项目(黔教合KY字[2018]276, 黔教合KY字[2018]429)

\*通信作者. E-mail: [zhangrenchong1990@163.com](mailto:zhangrenchong1990@163.com)

**引用格式:** 李静, 张仁崇, 潘春燕, 等. 单目标概率约束规划的微种群免疫优化算法[J]. 北京航空航天大学学报, 2023, 49(3): 525-537.  
LI J, ZHANG R C, PAN C Y, et al. Micro immune optimization algorithm for single objective probabilistic constrained programming[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(3): 525-537 (in Chinese).

从智能优化和工程应用角度来看, SOPCP 的研究成果依据算法类型主要分为 3 类: 遗传算法<sup>[1-3,5,7,21-22]</sup>、粒子群优化算法<sup>[6,23-27]</sup> 和免疫优化算法<sup>[16-20,28]</sup>。遗传算法是在静态采样下求解 SOPCP 问题的一种简单且实用的优化方法, 其中个体的适应度通过随机模拟、模糊模拟、神经网络及同步扰动随机逼近等方法来估计, 在工程优化设计中得到较好的应用, 但其计算效率、求解质量和模型逼近还有待改进。例如, Liu<sup>[21]</sup> 采用神经网络训练获取 SOPCP 的近似模型, 提出了几种适用于求解 SOPCP 的改进混合遗传算法; 卢福强等<sup>[2]</sup> 构建了第四方物流路径优化的 SOPCP 模型, 引入移民算子增强种群多样性, 且加入精英算子保存最优个体, 从而获得改进型遗传算法求解。

粒子群优化算法具有结构简单、参数少、效率高等优点, 在静态采样下与随机模拟、不确定函数模拟、神经网络结合已成为求解 SOPCP 问题的替代方法, 但该方法效率较低, 解质量极大依赖于训练样本分布及规模, 全局勘探能力不足。为克服此缺点, 肖宁<sup>[23]</sup> 将粒子的惯性权重设置为零, 增强局部搜索能力, 通过加入随机粒子丰富种群多样性。此外, 粒子群优化算法已广泛应用于可描述为 SOPCP 的工程优化问题<sup>[6,24-27]</sup>。

近年来, 免疫优化算法被初步应用于求解 SOPCP 问题, 且已有少量研究成果。在静态采样方面, 段富等<sup>[28]</sup> 利用随机模拟及神经网络逼近模型处理目标函数和机会约束概率值的估计问题, 采用双变异和双克隆策略, 提出改进的克隆选择算法求解。在自适应采样方面, 张著洪等<sup>[16-17]</sup> 针对一般类型的 SOPCP 问题, 基于备受关注的危险理论, 提出 2 个具有自适应采样、繁殖、变异的微种群免疫优化算法求解, 该方法效率较高, 效果令人满意; 同时, 张著洪等<sup>[18-20]</sup> 还针对单目标、多目标的概率优化问题分别设计了免疫优化算法求解, 证明免疫优化算法求解此类问题具有良好潜力。

除上述算法外, 部分其他启发式算法在静态采样下亦可有效求解 SOPCP 问题, 如量子进化算法<sup>[29]</sup>、帝国竞争算法<sup>[30]</sup> 等。

综上所述, 由于目标函数和约束条件存在随机变量, 导致 SOPCP 的求解一直是个难题。尽管一些智能算法在静态采样下能有效求解, 但在效率和求解质量上的缺陷明显限制了其实际应用; 虽然 2 个自适应采样的微种群免疫优化算法<sup>[16-17]</sup> 可高效求解 SOPCP 问题, 但其效率和寻优效果需进一步改善。因此, 本文从生物免疫系统危险理论的免疫应答中获取灵感, 探讨具有参数少、效率高、进化

能力强、噪声抑制效果好、应用潜力高等优点的微种群免疫优化算法 (micro immune optimization algorithm,  $\mu$ IOA-III), 同时提出目标值、机会约束的估计法分别高效估算目标值和机会约束的概率值。

本文  $\mu$ IOA-III 与  $\mu$ IOA<sup>[16]</sup>、 $\mu$ IOA-II<sup>[17]</sup> 借助危险理论启发和不同设计灵感, 设计微种群免疫优化算法求解 SOPCP 问题。其中,  $\mu$ IOA 是初次探讨问题求解的微种群免疫优化算法;  $\mu$ IOA-II 对  $\mu$ IOA 结构的复杂性和适应性进行改进, 设计新变异策略提高解搜索能力, 使用新的排序方法划分种群, 将个体支配概念扩展到一般情况, 引入个体生命周期模型防止劣质个体在进化过程中停留太久; 在  $\mu$ IOA-II 的基础上,  $\mu$ IOA-III 构建新生命周期模型, 基于此模型引入交叉算子加强群体内个体间信息交流, 设计变异策略增加算法搜索能力, 同时改进机会约束概率值的误差幅度和设计目标值的近似误差幅度, 提出 2 个新方法分别估计机会约束概率值和目标值。

## 1 问题描述与概率估计

### 1.1 问题描述

考虑如下 SOPCP 问题:

$$\min_{\mathbf{x} \in D} \bar{f}(\mathbf{x})$$

$$\text{s.t.} \begin{cases} \Pr\{f(\mathbf{x}, \boldsymbol{\zeta}) \leq \bar{f}(\mathbf{x})\} \geq \alpha \\ p_i(\mathbf{x}) \equiv \Pr\{G_i(\mathbf{x}, \boldsymbol{\xi}_i) \leq 0\} \geq \beta_i \quad 1 \leq i \leq I \\ g_j(\mathbf{x}) \leq 0, h_k(\mathbf{x}) = 0 \quad 1 \leq j \leq J, 1 \leq k \leq K \end{cases}$$

式中:  $D = \prod_{i=1}^p [a_i, b_i]$  为  $\mathbf{R}^p$  中有界决策闭区域;  $\mathbf{x}$  为  $D$  内的决策向量;  $\boldsymbol{\zeta}$  和  $\boldsymbol{\xi}_i$  为概率分布未知或复杂的随机向量;  $\alpha \in (0, 1)$  和  $\beta_i \in (0, 1)$  为置信水平;  $\Pr\{\cdot\}$  为概率算子;  $f(\mathbf{x}, \boldsymbol{\zeta})$  和  $G_i(\mathbf{x}, \boldsymbol{\xi}_i)$  分别为随机目标函数和随机约束函数;  $g_j(\mathbf{x})$  和  $h_k(\mathbf{x})$  为确定约束函数。在点  $\mathbf{x}$  处, 若随机向量  $\boldsymbol{\zeta}$  和  $\boldsymbol{\xi}_i$  附加一定样本量, 则  $\bar{f}(\mathbf{x})$  和  $p_i(\mathbf{x})$  的值可采用蒙特卡罗随机模拟进行估计<sup>[21]</sup>。若  $\mathbf{x}$  满足以上约束条件, 则称其为可信解。为有效判定候选解  $\mathbf{x}$  的可信性, 设计如下经验约束违背量:

$$\Gamma(\mathbf{x}) = G(\mathbf{x}) + H(\mathbf{x}) \quad (1)$$

式中:

$$\begin{cases} G(\mathbf{x}) = \frac{1}{I} \sum_{i=1}^I \max\{\beta_i - \hat{p}_i(\mathbf{x}) + \Delta_i(\mathbf{x}), 0\} \\ H(\mathbf{x}) = \frac{1}{J} \sum_{j=1}^J \max\{g_j(\mathbf{x}), 0\} + \frac{1}{K} \sum_{k=1}^K h_k^2(\mathbf{x}) \\ \hat{p}_i(\mathbf{x}) = n_i^{-1}(\mathbf{x}) \sum_{k=1}^{n_i(\mathbf{x})} I\{G_i(\mathbf{x}, \boldsymbol{\xi}_i^k) \leq 0\} \end{cases} \quad (2)$$

其中:  $\hat{p}_i(\mathbf{x})$  为  $p_i(\mathbf{x})$  在样本量  $n_i(\mathbf{x})$  下的估计值;  $I\{\cdot\}$  为

指示函数, 若条件为假则取 0, 反之取 1;  $\Delta_i(\mathbf{x})$  为改进的误差幅度, 其设计如下:

$$\begin{cases} \Delta_i(\mathbf{x}) = \max\{\psi_i(\mathbf{x}), \min\{1/n_i(\mathbf{x}), 1 - \hat{p}_i(\mathbf{x})\}\} \\ \psi_i(\mathbf{x}) = u_{(1+\delta)/2} \sqrt{\hat{p}_i(\mathbf{x})(1 - \hat{p}_i(\mathbf{x}))/n_i(\mathbf{x})} \end{cases} \quad (3)$$

式中:  $u_{(1+\delta)/2}$  表示标准正态分布的  $(1+\delta)/2$  分位数,  $\delta \in (0, 1)$ ;  $\psi_i(\mathbf{x})$  为  $p_i(\mathbf{x})$  与  $\hat{p}_i(\mathbf{x})$  的绝对误差幅度<sup>[14]</sup>, 为更贴近真实误差幅度, 被改进为  $\Delta_i(\mathbf{x})$ 。一旦样本量  $n_i(\mathbf{x})$  增大, 则误差幅度  $\Delta_i(\mathbf{x})$  将变小, 这有助于设计算法 1, 使个体自适应附加样本量高效获取估计值  $\hat{p}_i(\mathbf{x})$ , 并给出经验约束违背量  $\Gamma(\mathbf{x})$ 。若  $\Gamma(\mathbf{x}) = 0$ , 称  $\mathbf{x}$  为经验可信解, 经算法 2 可快速获得  $\mathbf{x}$  的经验目标值  $\hat{f}(\mathbf{x})$ ; 反之, 则称  $\mathbf{x}$  为经验非可信解。为有效比较候选解  $\mathbf{x}$  和  $\mathbf{y}$ , 引入如下经验支配概念<sup>[17]</sup>。

**定义 1** 对于候选解  $\mathbf{x}, \mathbf{y} \in D$ , 如果其满足下列 4 个条件之一, 则称  $\mathbf{x}$  经验支配  $\mathbf{y}$  (简记  $\mathbf{x} < \mathbf{y}$ ):

- 1)  $\Gamma(\mathbf{x}) = 0 = \Gamma(\mathbf{y}) \wedge \hat{f}(\mathbf{x}) < \hat{f}(\mathbf{y})$ 。
- 2)  $\Gamma(\mathbf{x}) = 0 < \Gamma(\mathbf{y})$ 。
- 3)  $0 < \Gamma(\mathbf{x}) < \Gamma(\mathbf{y})$ 。
- 4)  $0 < \Gamma(\mathbf{x}) = \Gamma(\mathbf{y}) \wedge G(\mathbf{x}) < G(\mathbf{y})$ 。

在以上 4 个条件中, 若  $\mathbf{x}$  和  $\mathbf{y}$  满足条件 4), 此时可用  $G(\mathbf{x})$  和  $G(\mathbf{y})$  确定  $\mathbf{x}$  和  $\mathbf{y}$  的经验支配关系。

**定义 2**  $\mathbf{x}^* \in D$  称为 SOPCP 问题的经验最优可信解, 如果对于  $\forall \mathbf{x} \in D$ , 均有  $\mathbf{x}^* < \mathbf{x}$ 。

### 1.2 噪声处理

由于目标函数和约束函数存在分布未知或复杂的随机向量, SOPCP 问题极难或不能转化为解析模型, 而随机模拟却是处理此类随机规划问题的有效工具。根据大数定律可知, 样本量足够大时, 机会约束的概率值和目标函数的估计值均逼近真实值, 但不可避免会导致高计算复杂度; 反之, 每个候选解都附加一个小样本量, 非可信解易被误认为是可信解, 且优质和劣质的候选解难以辨别, 使得优化质量严重受到噪声影响。为快速计算机会约束的概率值和目标函数的估计值, 现提出以下 2 种估计算法。

#### 1.2.1 概率值估计

为快速有效地判定候选解  $\mathbf{x}$  是否经验可信, 提出了机会约束概率值估计法。算法设计中, 先计算确定约束的约束违背量  $H(\mathbf{x})$ , 在固定小样本量下快速初始估计所有机会约束的概率值  $\hat{p}_i(\mathbf{x})$ 。若  $\mathbf{x}$  不满足确定约束条件, 则判定  $\mathbf{x}$  为非可信解; 否则, 借助误差幅度  $\Delta_i(\mathbf{x})$ , 自适应采样更新所有估计值  $\hat{p}_i(\mathbf{x})$ , 在较大样本量下使  $\hat{p}_i(\mathbf{x})$  逼近真实值, 依式 (1) 计算经验约束违背量  $\Gamma(\mathbf{x})$ , 判定  $\mathbf{x}$  是否为经验可信解。

#### 算法 1 机会约束概率值估计法。

**步骤 1** 输入参数: B 细胞  $\mathbf{x}$ , 最大迭代次数  $G_{\max}$ , 样本量  $M$ , 置信水平  $\beta_i$ , 参数  $\delta_{\min}, \delta_{\max}, 1 \leq i \leq I$ 。

**步骤 2** 置  $T_{\max} \leftarrow M^2 \ln G_{\max}, T_{\min} \leftarrow M^2/2$ ; 依据式 (2) 计算确定约束的约束违背量  $H(\mathbf{x})$ 。

**步骤 3** 置  $T_{\beta_i} = (1 + 2|\beta_i - 0.5|)T_{\min}, n_i(\mathbf{x}) \leftarrow T_{\min}, \delta \leftarrow \delta_{\max}$ , 随机产生  $n_i(\mathbf{x})$  个样本向量  $\xi_i^k$ , 获取估计值  $\hat{p}_i(\mathbf{x})$  及误差幅度  $\Delta_i(\mathbf{x}), k = 1, 2, \dots, n_i(\mathbf{x}), i = 1, 2, \dots, I$ 。

**步骤 4** 若  $H(\mathbf{x}) = 0$ , 则自适应采样更新概率值:

**步骤 4.1** 置  $i \leftarrow 1, l \leftarrow 0$ 。

**步骤 4.2** 若  $n_i(\mathbf{x}) \geq T_{\max}$ , 则置  $l \leftarrow l + 1$ , 转步骤 4.7。

**步骤 4.3** 若  $\beta_i > \hat{p}_i(\mathbf{x})$ , 转步骤 5。

**步骤 4.4** 若  $n_i(\mathbf{x}) < T_{\beta_i}$ , 转步骤 4.6。

**步骤 4.5** 若  $\beta_i + \Delta_i(\mathbf{x}) < \hat{p}_i(\mathbf{x})$ , 则置  $l \leftarrow l + 1$ , 转步骤 4.7; 否则, 转步骤 4.6。

**步骤 4.6** 随机产生  $M$  个样本向量  $\xi_i^k, k = 1, 2, \dots, M$ , 更新第  $i$  个机会约束的概率估计值  $\hat{p}_i(\mathbf{x})$ :

$$\hat{p}_i(\mathbf{x}) \leftarrow \frac{n_i(\mathbf{x})\hat{p}_i(\mathbf{x}) + \sum_{k=1}^M I\{G_i(\mathbf{x}, \xi_i^k) \leq 0\}}{n_i(\mathbf{x}) + M} \quad (4)$$

同时更新  $\Delta_i(\mathbf{x})$ , 置  $n_i(\mathbf{x}) \leftarrow n_i(\mathbf{x}) + M$ 。

**步骤 4.7** 若  $i < I$ , 则  $i \leftarrow i + 1$ , 并转步骤 4.2。

**步骤 4.8** 若  $l < I$ , 则转步骤 4.1; 否则, 转步骤 5。

**步骤 5** 置  $\delta \leftarrow \delta_{\min}$ , 依据式 (3) 重新计算误差幅度  $\Delta_i(\mathbf{x}), i = 1, 2, \dots, I$ ; 依据式 (1) 计算经验约束违背量  $\Gamma(\mathbf{x})$ , 并输出  $\Gamma(\mathbf{x})$ 。

注:  $\delta_{\max}$  和  $\delta_{\min}$  选取较大值,  $\delta_{\max}$  避免非可信个体因采样不足被误判为经验可信, 且个体可信的概率至少为  $\delta_{\min}$ 。

在以上描述中, 该算法的计算复杂度主要取决于步骤 3 和步骤 4, 其在最坏情况下为  $O(IT_{\max} + J + K)$ 。步骤 4 被设计自适应控制并赋予各机会约束不同样本量, 使算法缩减样本总量并快速更新估计值  $\hat{p}_i(\mathbf{x})$ ; 若  $\mathbf{x}$  不满足确定约束条件或不满足某个机会约束条件或明显满足所有机会约束条件, 抑或是满足终止条件时, 经步骤 5 计算并输出  $\Gamma(\mathbf{x})$  有效判定候选解  $\mathbf{x}$  经验可信性。此外, 样本量  $T_{\beta_i}$  被设计与  $\beta_i$  有关, 若概率值  $p_i(\mathbf{x})$  与置信水平  $\beta_i$  接近 1 (或 0) 时, 则估计值  $\Delta_i(\mathbf{x})$  将非常小, 极易使非可信解  $\mathbf{x}$  在式 (1) 下因采样不足被误判为经验可信解, 此时  $\mathbf{x}$  被要求在较大样本量  $T_{\beta_i}$  下更新估计值  $\hat{p}_i(\mathbf{x})$ 。

#### 1.2.2 目标值估计

为抑制噪声对目标函数的影响, 提出目标值估计法。该算法在小样本量下, 借助蒙特卡罗随机模拟, 通过线性、加权、继承等方式强化噪声抑制效

果,初始快速估计所有经验可信解  $\mathbf{x}$  的经验目标值  $\hat{f}(\mathbf{x})$ ; 在样本量上限  $M_n$  下,借助目标值近似误差幅度  $\Lambda(\mathbf{x})$  辨析个体间优劣关系,并基于此设计自适应采样更新经验目标值  $\hat{f}(\mathbf{x})$ ,使  $\hat{f}(\mathbf{x})$  逼近真实值。

**算法 2** 目标值估计法。

**步骤 1** 输入参数: 群体  $Q = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{|Q|}\}$ , 置信水平  $\alpha$ , 样本量  $m, M_n$ , 参数  $\delta_{\max}, m \ll M_n$ 。

**步骤 2** 置  $m(\mathbf{x}_i) \leftarrow m, \delta \leftarrow \delta_{\max}, i=1, 2, \dots, |Q|$ 。

**步骤 3** 估计  $Q$  中每个 B 细胞  $\mathbf{x}$  的经验目标值。

**步骤 3.1** 令  $s = m(\mathbf{x})$ ; 随机产生  $s$  个样本向量  $\zeta^k$  及其观测值  $\hat{f}_s(\mathbf{x}, \zeta^k) (1 \leq k \leq s)$ , 采用蒙特卡罗随机模拟估计 B 细胞  $\mathbf{x}$  的目标值  $\hat{f}_s(\mathbf{x})$ 。

$$\hat{f}_s(\mathbf{x}) = \hat{f}_s(\mathbf{x}, \zeta^v) + (\alpha(s-1) - [\alpha(s-1)]) \cdot (\hat{f}_s(\mathbf{x}, \zeta^{v+1}) - \hat{f}_s(\mathbf{x}, \zeta^v)) \quad (5)$$

式中:  $\hat{f}_s(\mathbf{x}, \zeta^v)$  为  $s$  个  $\hat{f}_s(\mathbf{x}, \zeta^k)$  中第  $v$  个最小观测值,  $v = [\alpha(s-1)] + 1$ 。

**步骤 3.2** 更新  $\mathbf{x}$  的经验目标值  $\hat{f}(\mathbf{x})$  及经验目标平方值  $\hat{F}(\mathbf{x})$ 。若  $s = m$ , 则  $\hat{f}(\mathbf{x}) = \hat{f}_s(\mathbf{x}), \hat{F}(\mathbf{x}) = \hat{f}_s^2(\mathbf{x})$ ; 否则  $\hat{f}(\mathbf{x}), \hat{F}(\mathbf{x})$  依据式(6)、式(7)更新:

$$\hat{f}(\mathbf{x}) \leftarrow \frac{(s-m)\hat{f}(\mathbf{x}) + 2\hat{f}_s(\mathbf{x})}{s-m+2} \quad (6)$$

$$\hat{F}(\mathbf{x}) \leftarrow \frac{\sum_{k=m}^{s-1} k\hat{F}(\mathbf{x}) + s\hat{f}_s^2(\mathbf{x})}{\sum_{k=m}^{s-1} k + s} \quad (7)$$

**步骤 3.3** 借助数理统计<sup>[31]</sup>, 将  $\hat{F}(\mathbf{x}) - \hat{f}^2(\mathbf{x})$  作为  $\bar{f}(\mathbf{x})$  经验目标值的近似偏差平方和, 进而设计  $\bar{f}(\mathbf{x})$  和  $\hat{f}(\mathbf{x})$  的误差幅度  $\Lambda(\mathbf{x})$  如下:

$$\Lambda(\mathbf{x}) = u_{(1+\delta)/2} [(s-m)^{-1} (\hat{F}(\mathbf{x}) - \hat{f}^2(\mathbf{x}))]^{1/2} \quad (8)$$

**步骤 3.4** 若  $s < M_n$ , 则更新  $Q$  中所有 B 细胞  $\mathbf{x}$  的样本量, 置  $m(\mathbf{x}) \leftarrow m(\mathbf{x}) + 1$ , 转步骤 3.1。

**步骤 4** 自适应更新经验目标值。

**步骤 4.1** 升幂排列  $Q$  中的经验目标值  $\hat{f}(\mathbf{x})$ 。

**步骤 4.2** 置  $i \leftarrow 1, j \leftarrow 0$ 。

**步骤 4.3** 若  $M_n \leq m(\mathbf{x}_i)$ , 置  $j \leftarrow j+1$ , 转步骤 4.5。

**步骤 4.4** 若  $\hat{f}(\mathbf{x}_{i+1}) - \hat{f}(\mathbf{x}_i) < \Lambda(\mathbf{x}_i) \wedge i < |Q|$  或者  $\hat{f}(\mathbf{x}_i) - \hat{f}(\mathbf{x}_{i-1}) < \Lambda(\mathbf{x}_i) \wedge i > 1$ , 则置  $m(\mathbf{x}_i) \leftarrow m(\mathbf{x}_i) + 1, \mathbf{x}_i$  在样本量  $m(\mathbf{x}_i)$  下, 按照步骤 3.1~步骤 3.3 更新  $\hat{f}(\mathbf{x}_i), \hat{F}(\mathbf{x}_i), \Lambda(\mathbf{x}_i)$ , 转步骤 4.1; 否则,  $j \leftarrow j+1$ , 转步骤 4.5。

**步骤 4.5**  $i \leftarrow i+1$ , 若  $i \leq |Q|$ , 转步骤 4.3。

**步骤 4.6** 若  $j < |Q|$ , 转步骤 4.1。

**步骤 5** 输出  $Q$  中全体 B 细胞  $\mathbf{x}$  的经验目标值  $\hat{f}(\mathbf{x})$ 。

注:  $\delta_{\max}$  选取较大值, 利于设计算法自适应分配

样本, 并避免劣质个体在小样本下被误判为优质个体。

算法 2 的计算复杂度主要由步骤 3 和步骤 4 决定。在最坏情况下, 算法 2 的计算复杂度为  $O(|Q|M_n^2(\log_2 M_n + \log_2 |Q|))$ , 主要依赖于  $Q$  和  $M_n$ 。其中, 步骤 3 的计算复杂度为  $O(|Q|s^2 \log_2 s)$ , 步骤 4 的计算复杂度为  $O(s^2 |Q| \log_2 |Q|)$ ,  $s \leq M_n$ 。

以上算法描述中, 步骤 3 经蒙特卡罗随机模拟和式(5)线性化初步抑制噪声获取估计值  $\hat{f}_s(\mathbf{x})$ ; 再经式(6)继承、加权方式增强抑噪效果更新估计值  $\hat{f}(\mathbf{x})$ ; 设计  $\hat{F}(\mathbf{x})$  经式(7)继承、加权方式获得, 进而设计目标值近似误差幅度  $\Lambda(\mathbf{x})$  (见式(8))。步骤 4 被设计使群体内所有个体循环依据经验目标值的误差幅度辨析其优势, 并自适应附加样本量更新经验目标值; 若个体与群内其他个体的目标值偏差较大使其易于辨析, 其目标值可在小样本量下估计; 反之, 其目标值在较大样本量下估计。

## 2 危险理论与微种群免疫优化算法

### 2.1 危险理论

危险理论<sup>[32]</sup>是一种新流行的免疫学理论, 与传统“自我-非我”模型区分异己的免疫理论应答机理不同。该理论认为, 免疫应答启动仅取决于机体内受损、凋亡细胞(即感染细胞)释放的危险信号, 对机体无害的内源、外源物质不作应答。具体而言, 当机体内正常或生命力减弱的细胞非正常凋亡或因外源物质受损时, 免疫系统发送信号 1 给辅助性 T 细胞, 发送信号 0 激活抗原呈递细胞(APC)释放信号 2, 辅助性 T 细胞接收信号 1、信号 2 后被完全激活产生淋巴因子, 促使机体发生免疫应答清理感染细胞, 帮助机体恢复健康。免疫系统包括 3 种细胞, 即未感染、易感染和已感染细胞。

### 2.2 算法设计

以上危险理论的描述过程, 为研究智能算法求解 SOPCP 问题提供思路和灵感,  $\mu$ IOA-III 的运行机制如图 1 所示。初始评价是经算法 1 判定个体是否经验可信, 在固定小样本量下经算法 2 估计经验可信个体的目标值; 群体分割是依据个体的经验约束违背量和经验目标值, 将进化群分割成 3 个子群: 已感染、易感染和未感染子群。各子群个体依据其质量实施克隆变异, 质量越高的个体繁殖克隆越多, 且交叉概率、变异概率及变异幅度越小; 同时, 采用交叉算子加强各子群个体间的信息交流, 促进协同进化, 经变异产生多样、高质量的新个体。经初始评价后筛选优质的新个体, 在较大样本量下与对应父代个体重新评价并比较, 随机生成新个体替换进化已达生命周期的个体, 产生新进化群体。

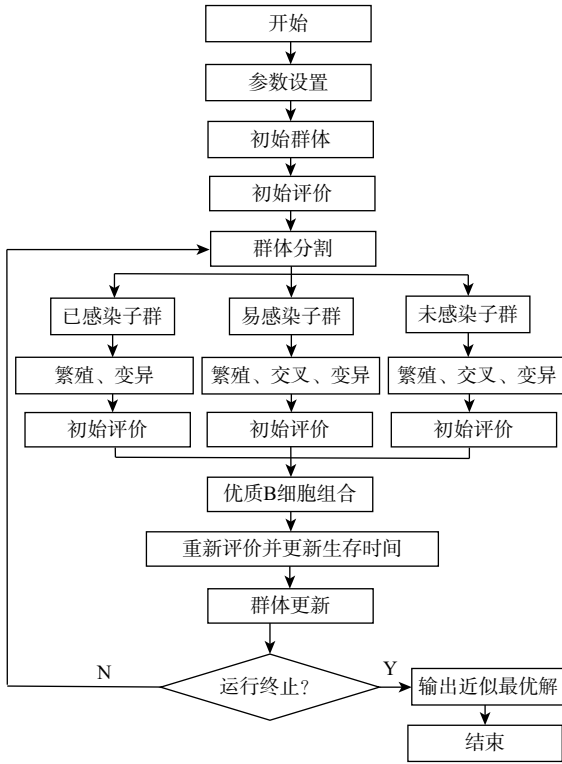


图 1 μIOA-III 的流程

Fig. 1 Flowchart of μIOA-III

为便于算法的描述, 结合以上 SOPCP 问题和图 1, 将此问题视为危险信号, 候选解视为 B 细胞。基于此, μIOA-III 的描述如下。

**算法 3** μIOA-III。

**步骤 1** 输入参数: 样本量  $m$ 、 $M$ , 群体规模  $N$ , 置信水平  $\alpha$ 、 $\beta$ , 参数  $W$ , 最大迭代数  $G_{\max}$ ,  $m \leq M$ ,  $1 \leq i \leq I$ 。

**步骤 2** 初始化: 置  $n \leftarrow 1$ , 产生  $N$  个随机 B 细胞的初始群体  $A = \{x_1, x_2, \dots, x_N\}$ , 生存时间  $w(x_i) \leftarrow 0$ 。

**步骤 3** 经验约束违背量: 经算法 1 估算群体  $A$  中所有 B 细胞  $x$  的经验约束违背量  $\Gamma(x)$ 。

**步骤 4** 经验目标值:  $M_n \leftarrow (M + m)[2 - \cos(\pi n / G_{\max})]$ ; 在小样本量  $M_1$  下经算法 2 获取  $A$  中所有经验可信 B 细胞  $x$  的经验目标值  $\hat{f}(x)$ 。

**步骤 5** 排序: 借助定义 1 的经验支配, 优劣排序  $A$  中所有 B 细胞, 其中排在最前的 B 细胞质量最好。

**步骤 6** 群体分割:  $A$  中第 1 个经验可信 B 细胞视为已感染细胞构成子群  $B_1$ , 其余经验可信 B 细胞视为易感染细胞构成子群  $B_2$ ; 所有经验非可信 B 细胞视为未感染细胞构成子群  $B_3$ 。

**步骤 7** 克隆: 群体  $A$  中第  $i$  个 B 细胞繁殖  $N-i+1$  个克隆。

**步骤 8** 交叉与变异。

**步骤 8.1** 交叉、变异概率: B 细胞  $x$  的变异概率  $p_m(x)$ 、交叉概率  $p_c(x)$  分别设计如下:

$$p_m(x) = \begin{cases} r_n \exp\left(\frac{\hat{f}(x) - \bar{f}_{\max}}{\bar{f}_{\max} - \bar{f}_{\min}}\right) & \Gamma(x) = 0 \\ r_n + (1 - r_n)(1 - nG_{\max}^{-1}) \exp\left(\frac{\Gamma(x) - \bar{g}_{\max}}{\bar{g}_{\max} - \bar{g}_{\min}}\right) & \text{其他} \end{cases} \quad (9)$$

$$p_c(x) = p_m(x) \cdot \max\{1 - w(x)/W, 0\} \quad (10)$$

式中:  $r_n = r_n A^{-1} + (1 - r_n A^{-1})(1 - nG_{\max}^{-1})$ ,  $r_n$  为群体  $A$  中经验可信 B 细胞的数量;  $\bar{f}_{\min}$ 、 $\bar{f}_{\max}$  分别为经验可信 B 细胞中最小、最大经验目标值;  $\bar{g}_{\min}$ 、 $\bar{g}_{\max}$  分别为经验非可信 B 细胞中最小、最大经验约束违背量。

**步骤 8.2** 交叉: 易感染子群  $B_2$ 、未感染子群  $B_3$  中 B 细胞  $x_i$  的克隆分别与  $B_1 \cup B_2$ 、 $B_1 \cup B_2 \cup B_3$  中随机选取较优的 B 细胞  $z (z < x_i)$ , 依据概率  $p_c(x_i)$  执行交叉操作获得  $x'_i$ 。

**步骤 8.3** 变异: 已感染子群  $B_1$  和易感染子群  $B_2$  中 B 细胞  $x_i$  的克隆经交叉生成的  $x'_i$  依据概率  $p_m(x_i)$  分别执行变异分布指数为  $G_{\max}^{0.5} + w(x_i) + 1$  和  $\sqrt{n} + w(x_i) + 1$  的多项式变异。未感染子群  $B_3$  的克隆经交叉生成  $x'_i$  依据  $p_m(x_i)$  执行分布指数为  $\Delta(x_i) = 1 - \zeta_{nu}^{(1-w(x_i))/(2W)}$  的非均匀变异<sup>[20]</sup>,  $\zeta_{nu}$  服从  $(0, 1)$  上的均匀分布。  $x'_i$  经变异后获得  $x''_i$ , 若  $x''_i$  的分量越界, 则执行越界处理<sup>[20]</sup>。  $x_i$  经克隆、交叉、变异获得子群  $C_i, i = 1, 2, \dots, N$ 。

**步骤 9** 初始评价:  $C_i$  中每个 B 细胞经算法 1 估算其经验约束违背量;  $C_i$  中经验可信 B 细胞在小样本量  $M_1$  下经算法 2 估计其经验目标值,  $i = 1, 2, \dots, N$ 。

**步骤 10** 重新评价并更新生存时间。

**步骤 10.1** 重新评价:  $x_i$  与其对应子群  $C_i$  中最优 B 细胞  $z$  构成优质子群  $D_i$ ; 若  $D_i$  中的 2 个 B 细胞均为经验可信 B 细胞, 需经算法 2 在较大样本量  $M_n$  下重新估计经验目标值,  $i = 1, 2, \dots, N$ 。

**步骤 10.2** 精英选择与更新生存时间: 选择每个子群  $D_i$  中最优 B 细胞  $y_i$  构成群体  $E$ ; 其中, 若  $D_i$  的 2 个 B 细胞  $z_i, x_i$  满足  $z_i < x_i$ , 则  $y_i = z_i, w(y_i) \leftarrow 0$ ; 否则,  $y_i = x_i, w(y_i) \leftarrow w(x_i) + 1$ 。经步骤 5, 优劣排序  $E$  中所有 B 细胞。

**步骤 10.3** 重置生存时间: 若已感染 B 细胞  $y_1$  生存时间  $w(y_1)$  达到生命周期  $W + N$ , 重置  $w(y_1) \leftarrow 0$ ; 若易感染 B 细胞  $y_2$  生存时间  $w(y_2)$  达到生命周期  $W + N - 1$ , 重置  $w(y_2) \leftarrow 0$ ; 若易感染或未感染 B 细胞  $y_i$  生存时间  $w(y_i)$  达到生命周期  $W + N - i + 1$ , 则随机生成新 B 细胞替换, 经算法 1 和算法 2 初始评价新 B 细胞, 重置  $w(y_i) \leftarrow 0, i = 3, 4, \dots, N$ 。

**步骤 10.4** 再次重新评价: 若步骤 10.3 随机生成新经验可信 B 细胞, 则群体  $E$  中所有经验可信 B 细胞经算法 2 在较大样本量  $M_n$  下重新估计经验目标值; 经步骤 5, 优劣排序  $E$  中所有 B 细胞。

**步骤 11** 群体  $E$  构成新进化群  $A$ ; 若  $G_{\max} \leq n$ , 则输出  $A$  中最优质的  $B$  细胞, 结束; 否则  $n \leftarrow n+1$ , 转步骤 5。

在以上  $\mu$ IOA-III 算法描述中, 步骤 6 将群体  $A$  分割成 3 类感染子群, 经步骤 7 使不同质量个体有差异繁殖克隆。步骤 8 构建生命周期模型, 基于此模型设计自适应的交叉与变异概率、变异分布指数; 采用交叉算子促进群体间个体的信息交流, 以及多项式变异、非均匀变异指导个体展开多方向进化, 产生多样性丰富、质量高的子代个体。设计步骤 10 筛选子代最优个体与对应父代个体, 在较大样本量下更新经验目标值, 有效辨析并挑选优质个体, 随机产生新个体替换达到生命周期的劣质个体, 获得多样性足够的新进化种群  $A$ 。在设计中, 进化个体仅 1 次经算法 1 估算其经验约束违背量, 使远离约束边界的个体用小样本量估算, 提高算法运行效率; 边界可信邻域内的个体用较大样本量估算, 使个体经验可信性不因采样不足被误判, 利于搜索优质个体。

### 2.3 计算复杂度

经由以上  $\mu$ IOA-III 描述, 其计算复杂度主要由步骤 5、步骤 8~步骤 10 确定。步骤 5 排序的计算复杂度为  $O(M \log_2 N)$ ; 步骤 7 至多克隆  $S=(N+1)N/2$  个, 步骤 8 执行交叉变异后的计算复杂度为  $O(pS)$ , 步骤 9 初始评价的计算复杂度为  $O(S(M_1^2(\log_2 M_1 + \log_2 S) + IT_{\max} + J + K))$ , 步骤 10 重新评价和排序的计算复杂度为  $O(NM_n^2 \log_2 M_n + M \log_2 N)$ 。  $\mu$ IOA-III 在最坏情形下的计算复杂度为

$$O_n = O(pN^2) + O(N^2(M_1^2(\log_2 M_1 + \log_2 N^2) + IT_{\max} + J + K)) + O(N(M_n^2 \log_2 M_n + \log_2 N)) + O(N(M_n^2 \log_2 M_n + NIT_{\max} + NJ + NK + Np)) \quad (11)$$

由上述分析可知,  $\mu$ IOA-III 的计算复杂度由  $M_n$ 、 $T_{\max}$ 、 $N$ 、 $p$ 、 $I$ 、 $J$ 、 $K$  确定, 主要取决于  $T_{\max}$ 、 $M_n$  和  $N$ ;  $T_{\max}$  和  $M_n$  主要依赖于样本量  $M$ ,  $M_n$  随算法运行逐渐增大。因此, 为使  $\mu$ IOA-III 高效率运行,  $M$  和  $N$  被要求取较小值。

## 3 数值实验

本文所有实验均在 Windows XP 系统配置为 CPU/2.20 GHz、RAM/1.99 GB 的 Visual C++ 平台上进行。为分析  $\mu$ IOA-III 的算法性能, 包括解质量、搜索效率和噪声抑制能力, 选取代表性的、竞争性的智能优化算法 SSGA-A<sup>[33]</sup>、SSGA-B<sup>[33]</sup>、FROFI<sup>[34]</sup>、C<sup>2</sup>oDE<sup>[35]</sup>、SPSO<sup>[23]</sup>、GA<sup>[21]</sup>、 $\mu$ IOA<sup>[16]</sup> 及  $\mu$ IOA-II<sup>[17]</sup>, 在 2 个理论测试问题、2 个工程问题上与  $\mu$ IOA-III 充分测试比较。以上算法具有相同的终止准则, 即最大

迭代次数为 300。为降低随机因素对算法分析的影响, 各算法对每个测试问题均独立运行 100 次, 输出的解需在样本量  $10^4$  下重新评价以使其经验约束违背量和经验目标值接近理论值; 之后, 将获取的结果用于统计分析。由于  $\mu$ IOA 和  $\mu$ IOA-II 具有自适应分配样本量能力, 不需要为每个个体分配固定样本量; 其余 6 种比较算法的所有个体均附加相同样本量 300<sup>[33]</sup>, 其种群规模均设置为 30; 以上 8 种比较算法的其他最佳参数设置均与原文献相同。

$\mu$ IOA-III 的主要可调参数为种群规模  $N$ 、样本量  $M$ 、生命周期参数  $W$ 。由于  $\mu$ IOA-III 是微种群优化算法, 要求种群大小  $N$  取较小的值 (2~6 之间)<sup>[18,36]</sup>, 特别是  $N=5$  的小种群算法已有相关研究<sup>[18,36-38]</sup>。样本量  $M$  直接影响算法搜索效果和效率, 若取值过小, 噪声抑制效果不理想, 解质量受影响; 相反, 算法效率低。  $W$  是控制个体生存时间的参数, 若  $W$  很大, 劣质个体将长时间停留; 相反, 获得一些有价值个体易消失, 可取值在 4~6 之间。经过实验测试后, 设置  $N=5$ ,  $m=2$ ,  $M=10$ ,  $W=5$ ,  $\delta_{\min}=0.99$ ,  $\delta_{\max}=0.9^9$ 。

### 3.1 理论实验

**问题 1** 多分布非线性概率优化问题<sup>[21]</sup>。

$$\max \quad \bar{f}(\mathbf{x})$$

$$\text{s.t.} \quad \begin{cases} \Pr \left\{ \sum_{i=1}^3 \zeta_i x_i \geq \bar{f}(\mathbf{x}) \right\} \geq 0.9 \\ \Pr \left\{ \sum_{i=1}^3 \eta_i x_i^2 \leq 8 \right\} \geq 0.8 \\ \Pr \left\{ \sum_{i=1}^3 \tau_i x_i^3 \leq 15 \right\} \geq 0.85 \\ x_1, x_2, x_3 \geq 0 \\ \zeta_1 \sim U(1,2), \zeta_2 \sim N(1,1), \zeta_3 \sim \text{Exp}(1) \\ \eta_1 \sim U(2,3), \eta_2 \sim N(2,1), \eta_3 \sim \text{Exp}(2) \\ \tau_1 \sim U(3,4), \tau_2 \sim N(3,1), \tau_3 \sim \text{Exp}(3) \end{cases}$$

此问题常被用于测试算法的降噪能力、搜索效果<sup>[21,23,28,39]</sup>, 其包含 3 个连续决策变量 ( $x_1, x_2, x_3$ ), 9 个服从均匀分布、正态分布或指数分布的随机变量; 目标函数和 2 个约束函数是非线性的且受 3 种噪声影响, 致使候选解  $\mathbf{x}$  是否可信难以确定, 目标值  $\bar{f}(\mathbf{x})$  难以准确估计, 问题求解难度极大。9 种算法独立求解此问题 100 次的统计结果如表 1 所示, 目标值的箱线图和平均搜索曲线如图 2 和图 3 所示。表 1 中: min 为最小值, max 为最大值, mean 为平均值, St.Dev 为均方差, CI 为置信区间, IAE 为平均约束违背量, FR 为 100 个最优解中经验可信解的比率, AR 为平均运行时间。图 3 中, 纵轴表示第  $n$  次迭代时 100 个最小经验目标值的平均值。

表 1 问题 1 的统计结果比较

Table 1 Comparison of statistical results for problem 1

算法	min	max	mean	St.Dev	CI	IAE	FR/%	AR/s
SSGA-A	1.883 0	2.323 5	2.216 3	0.078 4	[2.203 3,2.229 4]	0.002 1	83	3.30
SSGA-B	1.944 8	2.341 6	2.222 1	0.078 8	[2.209 0,2.235 2]	0.002 12	84	3.13
FROFI	2.165 3	2.344 9	2.274 0	0.037 8	[2.267 7,2.280 2]	0.003 1	78	6.52
C <sup>2</sup> oDE	1.911 4	2.322 2	2.183 4	0.089 0	[2.168 6,2.198 2]	0.001 0	94	20.78
SPSO	2.133 4	2.353 3	2.302 7	0.029 9	[2.297 8,2.307 7]	0.011 4	47	15.44
GA	2.203 1	2.346 6	2.307 5	0.024 9	[2.303 3,2.311 6]	0.005 1	67	20.18
μIOA	2.102 4	2.309 8	2.238 3	0.047 7	[2.230 4,2.246 2]	0	100	0.84
μIOA-II	2.150 5	2.319 1	2.265 8	0.031 9	[2.260 5,2.271 1]	0	100	0.54
μIOA-III	2.132 2	2.316 3	2.261 4	0.037 2	[2.255 2,2.267 5]	0	100	0.49

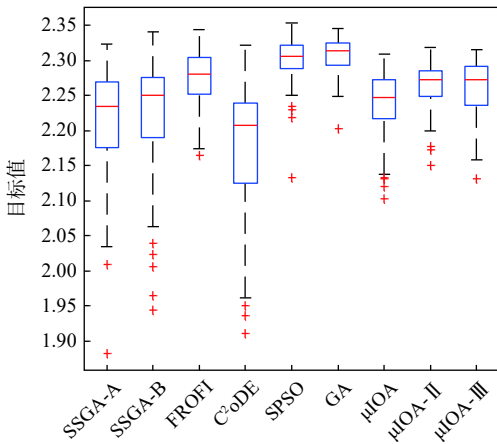


图 2 问题 1 的箱线图

Fig. 2 Box plots for problem 1

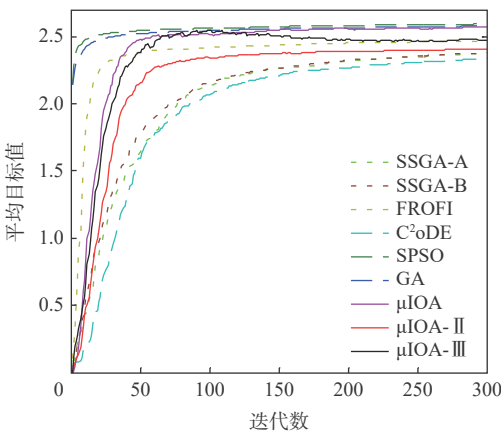


图 3 问题 1 的搜索曲线

Fig. 3 Search curves for problem 1

由表 1 第 7 列和第 8 列可知, 所有算法经 100 次独立运行后都能找到可信解, 其中, μIOA-III、μIOA-II 和 μIOA 在每次运行中均能找到可信解, C<sup>2</sup>oDE 次之, 其他比较算法仅能找到部分可信解, 且在某些次运行中有很小的约束违背量。表 1 第 2 列~第 5 列显示, 所有算法均获得相近的目标值, 表现出稳定的搜索性能。但结合第 7 列约束违背量,

μIOA-III、μIOA-II、μIOA 寻到最优解是可信的, 使其解质量相对较好; 比较算法 FROFI 有时难以寻到可信解, 即使获得的目标值较大, 其解的质量也相对较差。因此, 从解质量的角度来看, μIOA-III 和 μIOA-II 解质量最好, μIOA 次之且优于 FROFI、C<sup>2</sup>oDE、SSGA-A、SSGA-B、SPSO 较差。结合图 2、图 3 及以上分析获知, μIOA-III 可快速、稳定求解上述问题, 且解质量高。以上分析说明, 9 种不同算法均能抑制噪声对解搜索的影响, 主要原因归结为: ①μIOA-III、μIOA-II 和 μIOA 针对约束条件、目标函数存在的噪声, 分别设计相应的自适应采样策略抑制噪声; ②其他比较算法在个体附加相同的固定大样本容量 (300) 下, 可以控制噪声对算法寻优的影响。

此外, 表 1 第 9 列清楚地说明 μIOA-III 在执行效率方面优于其他比较算法, 至少是 SSGA-A 和 SSGA-B 的 5 倍、FROFI 的 12 倍、SPSO 的 30 倍、C<sup>2</sup>oDE 和 GA 的 40 倍; 尽管 μIOA-II 和 μIOA 也是快速的搜索算法, 但效率不如 μIOA-III。

问题 2 多模态概率约束优化问题。

$$\max \bar{f}(\mathbf{x})$$

$$\text{s.t.} \begin{cases} \Pr \left\{ \sum_{k=1}^3 x_k \sin(k\pi x_k) + \zeta \geq \bar{f}(\mathbf{x}) \right\} \geq 0.9 \\ \Pr \left\{ \sum_{i=1}^3 \xi_i x_i \leq 10 \right\} \geq 0.7 \\ \Pr \left\{ \sum_{i=1}^3 \varsigma_i x_i^2 \leq 100 \right\} \geq 0.8 \\ x_1, x_2, x_3 \geq 0, \\ \zeta \sim N(0,1), \xi_1 \sim U(0.8,1.2) \\ \xi_2 \sim U(1,1.3), \xi_3 \sim U(0.8,1), \varsigma_1 \sim N(1,0.5) \\ \varsigma_2 \sim \text{Exp}(1.2), \varsigma_3 \sim \log_2 N(0.8,0.6) \end{cases}$$

此问题是通过修改高度非线性的多模态约束优化问题<sup>[39]</sup>而获得的, 包含 3 个连续决策变量(x<sub>1</sub>,

$x_2, x_3$ ), 存在的7个服从正态分布、均匀分布、指数分布或对数正态分布的随机变量, 极大影响候选解可信性的判定和目标值的估计。此问题可用于测试算法是否有效抑制噪声, 是否有效判定候选解可

信性, 以及全局搜索、跳出局部搜索的能力。与上述实验一样, 每种算法直接求解该问题获得的统计结果如表2所示, 而对应的箱线图和平均搜索曲线分别如图4和图5所示。

表2 问题2的统计结果比较

Table 2 Comparison of statistical results for problem 2

算法	min	max	mean	St.Dev	CI	IAE	FR/%	AR/s
SSGA-A	6.954 1	8.597 5	7.940 1	0.323 0	[7.886 4,7.993 7]	0.002 7	83	2.98
SSGA-B	6.884 7	8.546 8	7.932 1	0.335 4	[7.876 4,7.987 8]	0.003 8	79	2.98
FROFI	6.796 9	8.545 9	7.819 4	0.305 5	[7.768 7,7.870 1]	0.000 5	92	5.87
C <sup>2</sup> oDE	7.906 3	8.597 2	8.318 5	0.141 9	[8.295 0,8.342 1]	0.002 4	84	17.42
SPSO	7.290 1	8.655 9	8.065 2	0.342 4	[8.008 3,8.122 0]	0.012 6	61	9.01
GA	7.602 9	8.629 3	8.172 9	0.271 9	[8.127 7,8.218 0]	0.005 7	66	11.52
μIOA	7.317 2	8.342 7	7.972 6	0.160 7	[7.945 9,7.999 3]	$8.50 \times 10^{-5}$	99	1.08
μIOA-II	7.290 4	8.277 9	7.969 7	0.159 0	[7.943 3,7.996 1]	$3.56 \times 10^{-5}$	99	0.91
μIOA-III	7.631 6	8.309 2	8.027 3	0.150 4	[8.002 3,8.052 3]	0	100	0.70

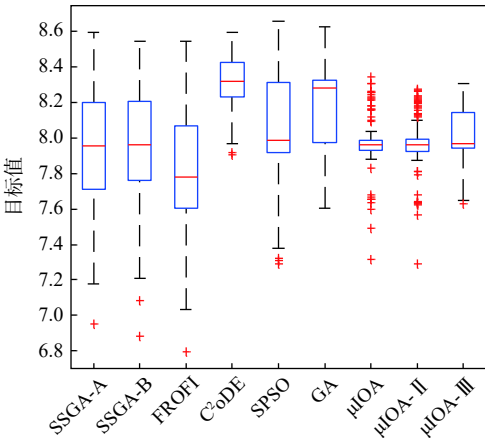


图4 问题2的箱线图

Fig. 4 Box plots for problem 2

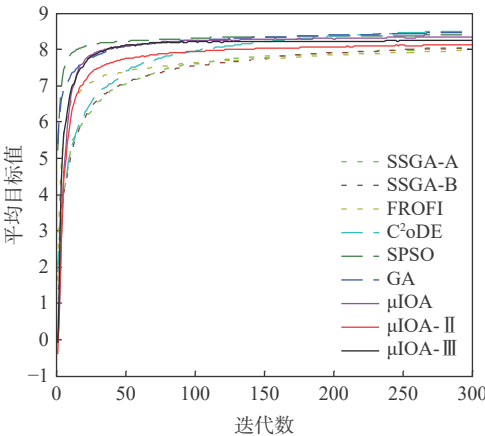


图5 问题2的搜索曲线

Fig. 5 Search curves for problem 2

由表2第7列和第8列获知, 所有算法均能搜寻到可信解, 其中μIOA-III、μIOA-II和μIOA找到可信解的概率最高, 但μIOA-II、μIOA有时难以找到可信解。同时, 结合第4列平均目标值可知, μIOA-III、

μIOA-II和μIOA获得满意的目标值; 尽管C<sup>2</sup>oDE、GA、SPSO的平均目标值占优, 但其部分最优解不满足约束条件, 使得解质量相对较差, 因此噪声抑制、约束处理能力需要改进。结合第5列均方差、第6列置信区间及图4表明, μIOA-III和比较算法有相近的稳定性, 但μIOA-III和C<sup>2</sup>oDE解搜索的稳定性最好。结合以上分析和图5可知, μIOA-III求解上述多模态概率约束优化问题, 算法搜索稳定、收敛速度快且所获解质量高。

就平均运行时间而言, μIOA-III的执行效率最高, μIOA-II、μIOA次之; μIOA-III的效率至少是SSGA-A和SSGA-B的3倍、FROFI的7倍、SPSO和GA的11倍、C<sup>2</sup>oDE的23倍。

3.2 工程应用

问题3 饲料混合成分问题。

min  $\bar{f}(x)$

$$\begin{cases}
 \Pr\{24.55x_1 + 26.75x_2 + 39.00x_3 + 40.50x_4 + \\
 \zeta \leq \bar{f}(x)\} \geq 0.9 \\
 \Pr\left\{\sum_{i=1}^4 x_i \xi_i \geq 21\right\} \geq 0.95 \\
 2.3x_1 + 5.6x_2 + 11.1x_3 + 1.3x_4 \geq 5 \\
 \sum_{i=1}^4 x_i = 1, x_i \geq 0 \\
 \zeta \sim N(0, 1) \\
 \xi_1 \sim N(12, 0.53), \xi_2 \sim N(11.9, 0.44) \\
 \xi_3 \sim N(41.8, 4.5), \xi_4 \sim N(52.1, 0.79)
 \end{cases}$$

此问题是通过修改原4维饲料混合成分优化模型<sup>[33]</sup>获得的, 饲料主要由大麦、燕麦、芝麻片、花生粉混合获得且所占比例为 $x_1$ 、 $x_2$ 、 $x_3$ 、 $x_4$ , 原料中蛋白质含量百分比为随机变量 $\xi_1$ 、 $\xi_2$ 、 $\xi_3$ 、 $\xi_4$ , 脂

肪含量百分比为 2.3、5.6、11.1、1.3, 每吨原料成本为 24.55、26.75、39.00、40.50 荷兰盾; 在每吨饲料需求 5% 脂肪、21% 蛋白质的条件下, 构建以成本最小化为目标函数在噪声环境下求解。该问题存在 5 个不同分布特征的随机变量, 特别是噪声强度较

大的  $\xi_3$  对解搜索的质量有很大影响, 可以测试算法对含噪声的线性规划问题的应用能力。所有算法分别执行 100 次求解此问题, 统计结果如表 3 所示, 相应的箱线图和平均搜索曲线如图 6 和图 7 所示。

表 3 问题 3 的统计结果比较

Table 3 Comparison of statistical results for problem 3

算法	min	max	mean	St.Dev	CI	IAE	FR/%	AR/s
SSGA-A	30.768 5	31.401 0	30.980 3	0.112 0	[30.961 7,30.998 9]	0.002 3	74	2.83
SSGA-B	30.716 9	31.328 1	30.948 4	0.112 4	[30.929 7,30.967 1]	0.002 5	76	2.79
FROFI	30.715 8	31.158 5	30.943 1	0.087 7	[30.928 6,30.957 7]	0.002 2	76	5.88
C <sup>2</sup> oDE	30.704 7	30.906 2	30.790 5	0.040 4	[30.783 8,30.797 2]	0.006 5	39	17.51
SPSO	30.671 6	31.445 7	30.963 3	0.147 0	[30.938 9,30.987 7]	0.010 7	28	7.43
GA	30.753 1	31.348 4	31.041 9	0.119 4	[31.022 1,31.061 7]	0.002 4	83	6.88
$\mu$ IOA	30.790 5	31.499 9	31.004 8	0.113 5	[30.986 0,31.023 7]	$4.54 \times 10^{-5}$	99	0.72
$\mu$ IOA-II	30.821 3	31.177 7	30.978 1	0.078 2	[30.965 1,30.991 1]	0	100	0.93
$\mu$ IOA-III	30.785 7	31.132 6	30.918 5	0.060 8	[30.908 4,30.928 6]	0	100	0.75

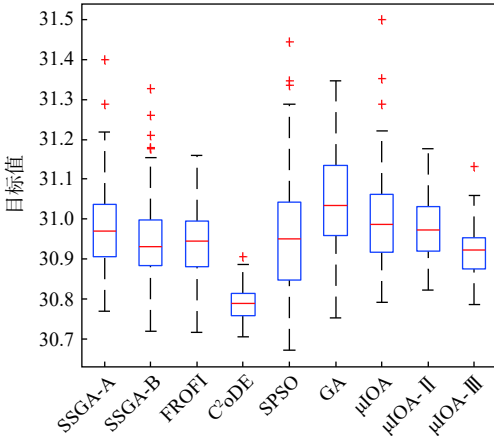


图 6 问题 3 的箱线图

Fig. 6 Box plots for problem 3

大, 说明以上算法在寻找可信解的过程中,  $\mu$ IOA-III、 $\mu$ IOA-II 和  $\mu$ IOA 能有效抑制噪声干扰, C<sup>2</sup>oDE 和 SPSO 的搜索性能受噪声干扰大。经由 IAE、FR、mean 的值及图 6 可知,  $\mu$ IOA-III 和 C<sup>2</sup>oDE 的平均目标值最小, 但 C<sup>2</sup>oDE 因多数最优解违反约束条件获得较小的平均目标值最小, 使得其解质量相对较差。为此,  $\mu$ IOA-III 求解上述问题获得的解质量优于其他算法。进一步, 结合 St.Dev、CI 的值及图 7 表明,  $\mu$ IOA-III 和比较算法有相近的搜索效果和稳定性, 但  $\mu$ IOA-III 相对更好, 且具有良好的收敛性。

在执行效率方面,  $\mu$ IOA-III、 $\mu$ IOA-II、 $\mu$ IOA 的执行效率最高, SSGA-A 和 SSGA-B 高于 FROFI、SPSO、GA、C<sup>2</sup>oDE 效率最低。

问题 4 混凝土泵车设计的费效权衡问题<sup>[40]</sup>。

$$\max \bar{f}(x)$$

$$\text{s.t.} \begin{cases} \Pr \{LCE/LCC \geq \bar{f}(x)\} \geq 0.95 \\ \Pr \{LCE \geq 70\} \geq 0.95 \\ \Pr \{LCC \leq 500\} \geq 0.95 \end{cases}$$

$$LCE = 0.441 8(x_2 + \varepsilon_2) + 0.028x_3 + 0.028x_4 + 0.066 3x_5 + 0.163 1(x_8 + \varepsilon_8) + 0.062 3x_{10} + 0.272 7x_{12}$$

$$LCC = 4.26 \times 10^{-5} \{0.029 3x_1 + 0.103 9(x_2 + \varepsilon_2) + 0.103 8x_3 + 0.092 9x_4 + 0.084 4x_5 + 0.098 7(x_6 + \varepsilon_6) + 0.090 7(x_7 + \varepsilon_7) + 0.098 8(x_8 + \varepsilon_8) + 0.092 6x_9 + 0.080 1x_{10} + 0.062 4x_{11} - 0.020 8x_{12}\}^{1.959} + \xi$$

$$x_1 \in [0.85, 1], x_2 \in [42, 56], x_3 \in [21.3, 36.6]$$

$$x_4 \in [10, 15.2], x_5 \in [4, 5], x_6 \in [11 700, 14 000]$$

$$x_7 \in [27 000, 42 000], x_8 \in [140, 170]$$

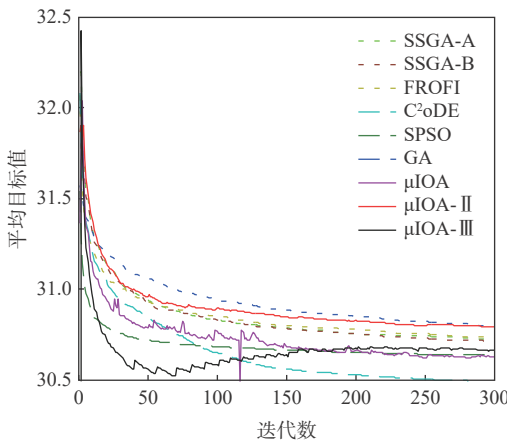


图 7 问题 3 的搜索曲线

Fig. 7 Search curves for problem 3

经由表 3 的 IAE 和 FR 值可知,  $\mu$ IOA-III 和  $\mu$ IOA-II 每次运行均找到可信解,  $\mu$ IOA 次之 (99%), C<sup>2</sup>oDE 和 SPSO 找到可信解的概率低且平均约束违背量略

$$x_9 \in [2\,000, 2\,200], x_{10} \in [265, 300]$$

$$x_{11} \in [380, 400], x_{12} \in [10.52, 14.256]$$

$$\varepsilon_2 \sim N(0, 1.44), \varepsilon_6 \sim N(0, 12), \varepsilon_7 \sim N(0, 16)$$

$$\varepsilon_8 \sim N(0, 4), \xi \sim N(0, 25)$$

为优化混凝土泵车设计,文献[40]选取底盘、臂架和泵送系统等重要部件中12个关键设计参数,以泵车生命周期效能LCE最大、费用LCC最小构造目标函数获得SOPCP问题。此问题包含4个连续

设计参数( $x_1, x_3, x_4, x_{11}$ )、8个整型设计参数( $x_2, x_5, \dots, x_{10}, x_{12}$ ),其中4个整型设计参数和费用LCC受方差较大的噪声影响,加之目标函数、约束函数是高度非线性,致使在大决策空间内寻求问题最优解变得极为困难,该问题可以测试算法对复杂噪声环境下工程问题的应用能力。实验与上述方式相同,各算法所得的统计结果如表4所示,箱形图和平均搜索曲线分别如图8和图9所示。

表4 问题4的统计结果比较

Table 4 Comparison of statistical results for problem 4

算法	min	max	mean	St.Dev	CI	IAE	FR/%	AR/s
SSGA-A	0.164 0	0.166 0	0.165 2	0.000 4	[0.165 1, 0.165 3]	0	100	6.25
SSGA-B	0.163 6	0.166 1	0.165 2	0.000 5	[0.165 1, 0.165 3]	0	100	6.20
FROFI	0.163 7	0.165 8	0.165 0	0.000 4	[0.165 0, 0.165 1]	0	100	13.16
C <sup>2</sup> oDE	0.165 9	0.166 4	0.166 2	0.000 1	[0.166 2, 0.166 3]	0	100	38.17
SPSO	0.163 9	0.166 3	0.165 8	0.000 4	[0.165 8, 0.165 9]	0	100	26.94
GA	0.161 3	0.165 5	0.163 5	0.000 8	[0.163 4, 0.163 7]	0	100	26.54
μIOA	0.165 8	0.166 5	0.166 3	0.000 1	[0.166 3, 0.166 3]	0	100	3.71
μIOA-II	0.165 4	0.166 4	0.166 0	0.000 2	[0.166 0, 0.166 1]	0	100	2.09
μIOA-III	0.165 3	0.166 4	0.166 0	0.000 2	[0.166 0, 0.166 1]	0	100	1.58

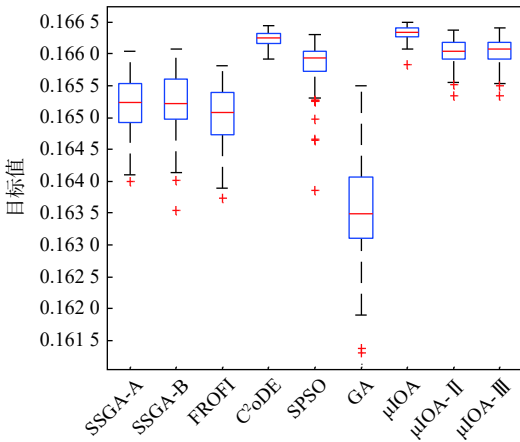


图8 问题4的箱线图

Fig. 8 Box plots for problem 4

从表4中IAE和FR值可以看出,所有算法在每次运行中都能找到可信解。由mean值,所有算法所获最优的平均目标值之间存在微小差异,即解质量没有明显差异;但从图8可看出,μIOA-III、μIOA-II、μIOA、C<sup>2</sup>oDE及SPSO的解质量最好且算法搜索更加稳定。进一步,由图9表明,9种算法求解该问题均是收敛的。

此外,效率是明显不同,由AR值看出,μIOA-III的效率最高,其效率稍优于μIOA-II、μIOA,至少是其他算法的2.9倍以上,而C<sup>2</sup>oDE、SPSO和GA需要更多的时间来寻找最优解。

### 3.3 灵敏度分析

μIOA-III主要有3个可调参数:  $N$ 、 $M$ 、 $W$ , 选择多模态概率约束优化问题2来测试算法的搜索性能是否在很大程度上依赖于它们参数的设置。为此,  $N$ 、 $M$ 、 $W$ 组成如表5所示的7种不同的参数组合,充分测试算法对参数设置的敏感性;算法在每种参数组合下独立运行100次,得到的统计结果如表5所示。

经由表5可知, μIOA-III在7种参数组合下均独立运行100次求解多模态概率约束优化问题,每次都能以很高的概率寻到满意的可信最优解,且解质量相近,表明算法μIOA-III对参数  $N$ 、 $M$ 、 $W$  的敏感度不高。在相同的迭代次数下,当生命周期参数  $W$  增大时,局部搜索能力增强,解质量略微提高;当群体规模  $N$ 、样本量  $M$  增大时,解质量提高,效率

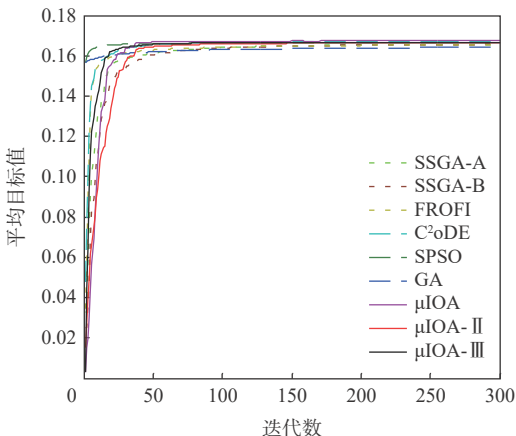


图9 问题4的搜索曲线

Fig. 9 Search curves for problem 4

降低; 当样本量  $M$  减小时, 算法抑制噪声能力降低, 使算法有时难以寻到可信解。为此, 综合考虑

算法的抑噪效果、寻优能力、运行效率, 选取参数组合:  $N=5, M=10, W=5$ 。

表 5  $\mu$ IOA-III 在不同参数设置下求解问题 2 的统计结果比较

Table 5 Comparison of statistical results for solving problem 2 of  $\mu$ IA-III in different parameter settings

参数组合	min	max	mean	St.Dev	CI	IAE	FR/%	AR/s
$N=4, M=10, W=5$	7.381 4	8.322 5	7.936	0.177 9	[7.906 4, 7.965 4]	0	100	0.41
$N=5, M=10, W=5$	7.631 9	8.374 7	8.022 9	0.134 8	[8.000 5, 8.045 3]	0	100	0.60
$N=6, M=10, W=5$	7.846 6	8.344 6	8.057 9	0.139 0	[8.034 8, 8.081 0]	0	100	0.88
$N=5, M=10, W=4$	7.612 1	8.315 4	8.007 4	0.146 0	[7.983 2, 8.031 7]	0	100	0.60
$N=5, M=10, W=6$	7.653 5	8.339 7	8.030 8	0.138 4	[8.007 8, 8.053 8]	0	100	0.63
$N=5, M=8, W=5$	7.607 9	8.304 4	7.984 3	0.143 6	[7.960 4, 8.008 1]	$1.19 \times 10^{-4}$	99	0.43
$N=5, M=12, W=5$	7.585 2	8.342 3	8.042 2	0.164 2	[8.015 0, 8.069 5]	0	100	0.79

## 4 结 论

单目标概率约束规划问题是一类具有综合工程应用背景的随机规划难题, 为此, 受生物免疫学中危险理论运行机制启发, 设计了微种群免疫优化算法( $\mu$ IOA-III)求解。经理论分析和实验分析, 得出以下结论:

1) 危险理论是探索智能优化算法求解 SOPCP 问题的一种有用的生物启发理论。

2)  $\mu$ IOA-III 的计算复杂度主要取决于群体规模、机会约束和目标函数的样本量上限。

3)  $\mu$ IOA-III 可以稳定、高效地搜索到满意质量的解, 具有种群小、可调参数少、自适应性等特点, 是求解 SOPCP 问题的一种有竞争力和应用潜力的优化算法。

4)  $\mu$ IOA-III 中机会约束概率值估计法能自适应为 SOPCP 问题中各机会约束分配不同数量的样本, 以有效抑制噪声干扰并估计概率值; 目标值估计法能自适应为进化种群各个体分配不同数量的样本, 以有效抑制噪声干扰并估算目标值。

5)  $\mu$ IOA-III 引入个体生命周期、交叉算子及自适应的交叉与变异概率、多项式变异、非均匀变异、分布指数等, 可以有效增强算法进化能力。

此外, 尽管该算法具有一些优异的性能, 特别是其运行效率, 但其工程应用仍需进一步研究。

### 参考文献 (References)

[1] CHEN Y, LI Y, SUN B, et al. A chance-constrained programming approach for a zinc hydrometallurgy blending problem under uncertainty[J]. *Computers & Chemical Engineering*, 2020, 140: 106893.

[2] 卢福强, 陈伟东, 毕华玲, 等. 考虑随机需求和多种运输方式的第四方物流路径问题[J]. *计算机集成制造系统*, 2020, 26(10): 2864-2876.

LU F Q, CHEN W D, BI H L, et al. Fourth party logistics routing problem considering stochastic demand and multiple transportation modes[J]. *Computer Integrated Manufacturing Systems*, 2020,

26(10): 2864-2876(in Chinese).

[3] LONG J, SUN Z, PARDALOS P M, et al. A robust dynamic scheduling approach based on release time series forecasting for the steelmaking-continuous casting production[J]. *Applied Soft Computing*, 2020, 92: 106271.

[4] LONG J, SUN Z, HONG Y, et al. Robust dynamic scheduling with uncertain release time for the steelmaking-continuous casting production[C]//2018 International Conference on Sensing, Diagnostics, Prognostics, and Control. Piscataway: IEEE Press, 2018: 531-536.

[5] SONG H, DONG M, HAN R, et al. Stochastic programming-based fault diagnosis in power systems under imperfect and incomplete information[J]. *Energies*, 2018, 11(10): 2565.

[6] HONG Z, ZHANG Q, GONG T, et al. Peak load regulation and cost optimization for microgrids by installing a heat storage tank and a portable energy system[J]. *Applied Sciences*, 2018, 8(4): 567.

[7] ZHANG H, HA M, ZHAO H, et al. Inexact multistage stochastic chance constrained programming model for water resources management under uncertainties[J]. *Scientific Programming*, 2017, 2017: 1-14.

[8] AHMED S. Convex relaxations of chance constrained optimization problems[J]. *Optimization Letters*, 2014, 8(1): 1-12.

[9] SUN X L, BAI X D, ZHENG X J. A survey on probabilistically constrained optimization problems[J]. *Operations Research Transactions*, 2012, 16(3): 65-73.

[10] 刘宝碇, 赵瑞清. 随机规划与模糊规划[M]. 北京: 清华大学出版社, 1998: 79-87.

LIU B D, ZHAO R Q. Stochastic programming and fuzz programming[M]. Beijing: Tsinghua University Press, 1998: 79-87(in Chinese).

[11] YANG L, YANG Z, LI G, et al. Optimal scheduling of an isolated microgrid with battery storage considering load and renewable generation uncertainties[J]. *IEEE Transactions on Industrial Electronics*, 2019, 66(2): 1565-1575.

[12] HOMEN-DE-MELLO T, BAYRAKSAN G. Monte Carlo sampling-based methods for stochastic optimization[J]. *Surveys in Operations Research and Management Science*, 2014, 19(1): 56-85.

[13] ZHANG Z H. Noisy immune optimization for chance-constrained programming problems[J]. *Applied Mechanics and Materials*, 2011, 48-49: 740-744.

[14] ZHANG Z H, WANG L, LIAO M. Adaptive sampling immune al-

- gorithm solving joint chance-constrained programming[J]. *Journal of Control Theory and Applications*, 2013, 11(2): 237-246.
- [15] YANG K, ZHANG Z H. Adaptive sampling detection based immune optimization approach and its application to chance constrained programming[M]//SUN H, YANG C Y, LIN C W, et al. Genetic and evolutionary computing. Berlin: Springer International Publishing, 2015: 19-28.
- [16] ZHANG Z H, LI L, ZHANG R C. Danger theory based micro immune optimization algorithm solving probabilistic constrained optimization[C]//2017 2nd IEEE International Conference on Computational Intelligence and Applications. Piscataway: IEEE Press, 2017: 103-107.
- [17] ZHANG Z H, ZHANG R C. Danger theory inspired micro-population immune optimization for probabilistic constrained programming[J]. *Evolving Systems*, 2020, 11(2): 333-348.
- [18] 张著洪, 张仁崇. 求解概率优化问题的微种群免疫优化算法[J]. *北京航空航天大学学报*, 2016, 42(9): 1785-1794.  
ZHANG Z H, ZHANG R C. Micro immune optimization algorithm for solving probabilistic optimization problems[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2016, 42(9): 1785-1794(in Chinese).
- [19] 张仁崇, 张著洪. 非线性多目标概率约束规划免疫优化算法[J]. *北京航空航天大学学报*, 2020, 46(5): 900-914.  
ZHANG R C, ZHANG Z H. Immune optimization algorithm for nonlinear multi-objective probabilistic constrained programming [J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2020, 46(5): 900-914(in Chinese).
- [20] 张仁崇, 潘春燕, 武星, 等. 非线性多目标概率优化问题的自适应采样免疫优化算法[J]. *电子学报*, 2021, 49(4): 647-660.  
ZHANG R C, PAN C Y, WU X, et al. Adaptive sampling immune optimization algorithm for nonlinear multi-objective probabilistic optimization problems[J]. *Acta Electronica Sinica*, 2021, 49(4): 647-660(in Chinese).
- [21] LIU B D. Theory and practice of uncertain programming[M]. 3rd ed. Berlin: Springer International Publishing, 2009: 50-53.
- [22] 沈依婷, 张菁, 武鹏, 等. 含电动汽车的配电网双重不确定性网架规划方法[J]. *中国电力*, 2020, 53(4): 139-146.  
SHEN Y T, ZHANG J, WU P, et al. Bi-uncertainty network frame planning method for distribution network with electric vehicles[J]. *Electric Power*, 2020, 53(4): 139-146(in Chinese).
- [23] XIAO N. An algorithm for solving stochastic chance-constrained programming problem[J]. *Advanced Materials Research*, 2014, 912: 1138-1141.
- [24] HUANG D, XIE L, WU Z. Dynamic economic dispatch for micro-grid based on the chance-constrained programming[J]. *Journal of Electrical Engineering & Technology*, 2017, 12(3): 1064-1072.
- [25] 李鹏, 蔡永青, 韩肖清, 等. 计及随机模糊双重不确定性的交直流混合微网优化运行[J]. *高电压技术*, 2020, 46(7): 2269-2279.  
LI P, CAI Y Q, HAN X Q, et al. Optimization operation of AC/DC hybrid microgrid considering random fuzzy double uncertainties[J]. *High Voltage Engineering*, 2020, 46(7): 2269-2279(in Chinese).
- [26] 德格吉日夫, 谭忠富, 李梦露, 等. 考虑不确定性的风储电站参与电力现货市场竞价策略[J]. *电网技术*, 2019, 43(8): 2799-2807.  
DE G J R F, TAN Z F, LI M L, et al. Bidding strategy of wind-storage power plant participation in electricity spot market considering uncertainty[J]. *Power System Technology*, 2019, 43(8): 2799-2807(in Chinese).
- [27] 马丽叶, 王志强, 陆肖宇, 等. 基于机会约束规划的风-火-蓄联合系统优化调度[J]. *电网技术*, 2019, 43(9): 3311-3320.  
MA L Y, WANG Z Q, LU X Y, et al. Optimal scheduling of combined wind-thermo-storage system based on chance constrained programming[J]. *Power System Technology*, 2019, 43(9): 3311-3320(in Chinese).
- [28] ZHAO Q, YANG R, DUAN F. An immune clonal hybrid algorithm for solving stochastic chance-constrained programming[J]. *Journal of Computational Information Systems*, 2012, 8(20): 8295-8302.
- [29] 韩畅, 梁博森, 林振智, 等. 防灾应急电源优化调度的机会约束规划方法[J]. *电力自动化设备*, 2018, 38(3): 147-154.  
HAN C, LIANG B M, LIN Z Z, et al. Chance-constrained programming method for optimal scheduling of emergency power source[J]. *Electric Power Automation Equipment*, 2018, 38(3): 147-154(in Chinese).
- [30] 贾伯岩, 马天祥, 张智远, 等. 计及不确定性的含分布式发电电网的配电网故障区段定位方法[J]. *电力系统保护与控制*, 2020, 48(14): 35-42.  
JIA B Y, MA T X, ZHANG Z Y, et al. Fault section location method for distribution networks with distributed generation and grid connection considering uncertainty[J]. *Power System Protection and Control*, 2020, 48(14): 35-42(in Chinese).
- [31] 茆诗松, 程依明, 濮晓龙, 等. 概率论与数理统计教程[M]. 北京: 高等教育出版社, 2004: 226-237.  
MAO S S, CHENG Y M, PU X L, et al. Probability theory and mathematical statistics course[M]. Beijing: Higher Education Press, 2004: 226-237(Chinese).
- [32] MATZINGER P. Tolerance, danger, and the extended family[J]. *Annual Review of Immunology*, 1994, 12(1): 991-1045.
- [33] POOJARI C A, VARGHESE B. Genetic algorithm based technique for solving chance constrained problems[J]. *European Journal of Operational Research*, 2008, 185(3): 1128-1154.
- [34] WANG Y, WANG B C, LI H X, et al. Incorporating objective function information into the feasibility rule for constrained evolutionary optimization[J]. *IEEE Transactions on Cybernetics*, 2016, 46(12): 2938-2952.
- [35] WANG B C, LI H X, LI J P, et al. Composite differential evolution for constrained evolutionary optimization[J]. *IEEE Transactions on Systems Man and Cybernetics Systems*, 2018, 49(7): 1482-1495.
- [36] SALEHINEJAD H, RAHNAMAYAN S, TIZHOOSH H R. Micro-differential evolution: diversity enhancement and a comparative study[J]. *Applied Soft Computing*, 2017, 52: 812-833.
- [37] YASUI T, SUGISAKA J, HIRAYAMA K. Structural optimization of silica-based 2x2 multimode interference coupler using a real-coded micro-genetic algorithm[J]. *Progress in Electromagnetics Research*, 2017, 55: 169-178.
- [38] CAMERO A, ARELLANO-VERDEJO J, ALBA E. Road map partitioning for routing by using a micro steady state evolutionary algorithm[J]. *Engineering Applications of Artificial Intelligence*, 2018, 71: 155-165.
- [39] VARGHESE B, POOJARI C A. Genetic algorithm based technique

for solving chance constrained problems arising in risk management[R]. CARISMA Technical Report, 2004: 1-49.

[40] 程晓娟, 韩庆兰, 全春光. 不确定条件下机械产品设计方案费效权衡优化[J]. *计算机集成制造系统*, 2015, 21(8): 1988-1994.

CHENG X J, HAN Q L, QUAN C G. Cost-effective trade-off of mechanical product design scheme under uncertainty[J]. *Computer Integrated Manufacturing System*, 2015, 21(8): 1988-1994(in Chinese).

## Micro immune optimization algorithm for single objective probabilistic constrained programming

LI Jing<sup>1</sup>, ZHANG Renchong<sup>1,\*</sup>, PAN Chunyan<sup>2</sup>, YANG Kai<sup>3</sup>

- (1. Computer and Information Engineering College, Guizhou University of Commerce, Guiyang 550014, China;
2. School of Mathematics and Statistics, Qiannan Normal University for Nationalities, Duyun 558000, China;
3. Guizhou Big Data Industry Research Institute, Guizhou University, Guiyang 550025, China)

**Abstract:** An immune optimization algorithm with a small population is proposed to solve a single objective probability constrained programming with no prior random distribution information. In the design of the algorithm, we develop an evolutionary framework with a micro population inspired by danger theory. Based on the amplitude of error of the estimated value, two approaches are proposed to estimate the individual's objective values and each chance constraint's probability respectively. According to the superior and inferior relationships among individuals, the population was divided into three types of sub-population for co-evolution. A version of individual life cycle is constructed, while adaptive crossover probability, adaptive mutation probability and adaptive mutation strategy as well as crossover strategy are designed to promote effective information exchange among the above sub-populations to co-evolve individuals in different directions. The results of numerical experiments show that the proposed algorithm has good search efficiency, search effect and noise reduction ability, and has certain competitiveness and application potential.

**Keywords:** probabilistic constrained programming; immune optimization; danger theory; adaptive sampling; micro population

**Received:** 2021-06-02; **Accepted:** 2021-09-24; **Published Online:** 2021-10-12 09:05

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20211011.1811.002.html](https://kns.cnki.net/kcms/detail/11.2625.V.20211011.1811.002.html)

**Foundation items:** Science and Technology Program of Guizhou Province of China (QKHJC [2020] 1Y423, QKHJC [2019] 1178); Project of Guizhou Big Data Application Engineering Research Center in Guizhou Province (QJHKY Zi [2017] 022); Youth Science and Technology Talent Development Project of Education Department in Guizhou Province (QJHKY Zi [2018] 276, QJHKY Zi [2018] 429)

\* **Corresponding author.** E-mail: [zhangrenchong1990@163.com](mailto:zhangrenchong1990@163.com)

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2022.0338

# 基于 MSST 及 HOG 特征提取的雷达辐射源信号识别

全大英<sup>1,\*</sup>, 唐泽雨<sup>1</sup>, 陈赞<sup>1</sup>, 楼维中<sup>2</sup>, 汪晓锋<sup>1</sup>, 章东平<sup>1</sup>

(1. 中国计量大学 信息工程学院 浙江省电磁波信息技术与计量检测重点实验室, 杭州 310018;

2. 中国电子科技集团有限公司第五十二研究所, 杭州 310000)

**摘 要:** 针对传统雷达信号识别算法在低信噪比下识别准确率低的问题, 提出了基于多重同步压缩 (MSST) 时频变换及方向梯度直方图 (HOG) 特征提取的雷达辐射源信号识别算法。所提算法在雷达时域信号短时傅里叶变换 (STFT) 基础上进行多重同步压缩处理获得信号时频分布图, 通过 HOG 算子对信号时频分布图进行 HOG 特征提取, 将提取的 HOG 特征通过主成分分析法 (PCA) 进行降维, 将降维后的特征参数送入支持向量机 (SVM) 对雷达信号进行分类与识别。实验结果表明: 所提算法具有较低的复杂度, 当信噪比为 -8 dB 时, 仿真实验与半实物仿真实验针对 9 种典型雷达信号的识别准确率达到 90% 以上。

**关键词:** 雷达信号识别; 方向梯度直方图; 多重同步压缩; 支持向量机; 主成分分析法  
**中图分类号:** TN974

**文献标志码:** A **文章编号:** 1001-5965(2023)03-0538-10

电子侦察在电子对抗领域发挥着至关重要的作用<sup>[1]</sup>, 而低截获概率 (low probability of intercept, LPI) 雷达信号识别是电子侦察的关键因素<sup>[2-3]</sup>, 有效识别 LPI 雷达信号可以帮助我方掌握战场主动权, 进而影响进一步的作战决策<sup>[4]</sup>。传统雷达信号识别主要基于脉冲宽度、脉冲幅度、载波频率、到达时间和到达角等常规参数特征。基于信号间参数的识别技术在信号流密度低、信号形式简单、辐射源数量少、电磁环境简单的情况下可以达到较好的效果, 然而随着雷达技术的进步, 战场电磁环境日益复杂, LPI 雷达具有低截获、高探测、大时宽和强干扰等特点, 传统方法已不能满足识别需求。

由于新型雷达信号具有丰富的脉内信息, 可通过分析信号脉内特征参数来提升识别的准确性<sup>[5]</sup>。近年来, 国内外众多学者通过分析雷达信号时域、频域和时频域特征对雷达信号识别问题展开了多项研究。文献 [6] 根据雷达信号频谱特点, 利用复杂度计算思想从几何形状上描述信号特征, 提出基

于频谱复杂度的雷达信号调制方式识别方法, 该方法在信噪比为 6 dB 时整体识别率可以达到 90% 以上, 但随着频率分辨率降低, 无法区分非线性调频信号和线性调频信号频谱的几何图形, 导致识别结果不理想。文献 [7] 通过分数阶傅里叶变换对信号进行 Chirp 基分解, 按照 Chirp 基载频与调频率的不同组合对信号划分类别, 并设置对应的分类特征参数, 计算信号的伪 Wigner-Ville 时频分布, 并提取 Zernike 矩特征, 将上述特征参数组成信号特征矢量, 使用残差神经网络分类器实现雷达信号识别, 当信噪比为 -2 dB 时, 该算法识别准确率平均可达到 93%, 但对 BPSK 信号识别率较低。文献 [8] 在 Choi-Williams 时频分布 (Choi-Williams distribution, CWD) 的基础上, 对图像进行伪 Zernike 矩和中心矩特征提取, 从而实现对信号的分类识别, 但在低信噪比情况下, 该算法识别率较低。文献 [7-8] 基于 Wigner-Ville 分布 (Wigner-Viller distribution, WVD) 和 CWD 时频变换提取信号的图像特征, 但由于 WVD 和 CWD

收稿日期: 2022-05-09; 录用日期: 2022-07-29; 网络出版时间: 2022-08-05 14:18

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20220805.1340.002.html](https://kns.cnki.net/kcms/detail/11.2625.V.20220805.1340.002.html)

基金项目: 浙江省自然科学基金 (LQ20F020021); 浙江省电磁波信息技术与计量检测重点实验室开放式项目 (2019KF0003)

\* 通信作者. E-mail: [qdy@cjl.edu.cn](mailto:qdy@cjl.edu.cn)

**引用格式:** 全大英, 唐泽雨, 陈赞, 等. 基于 MSST 及 HOG 特征提取的雷达辐射源信号识别 [J]. 北京航空航天大学学报, 2023, 49 (3): 538-547.  
QUAN D Y, TANG Z Y, CHEN Y, et al. Radar emitter signal recognition based on MSST and HOG feature extraction [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 538-547 (in Chinese).

时频分辨率有限, 对细节的描述不够, 不能体现不同信号之间的细微差别, 从而影响识别的准确率。针对文献 [7-8] 中时频分辨率低的问题, 文献 [9] 提出了基于短时傅里叶变换 (short-time Fourier transform, STFT) 的同步压缩 (synchrosqueezing transform, SST) 方法, 即在 STFT 的基础上, 将信号频率进行压缩, 以达到“挤压”的目的<sup>[10]</sup>。相比于 WVD 和 CWD 时频变换, 经同步压缩之后的时频分布图分辨率更高, 有利于雷达信号的准确识别。然而, 同步压缩在处理复杂时变信号时, 时频聚集性会大大降低, 导致时频模糊, 随着信号非平稳性的增强, 瞬时频率估计偏差越来越大。

为了解决 WVD 和 CWD 时频分辨率不足及时变信号同步压缩后时频模糊的问题, 本文首先采用多重同步压缩 (multi-synchrosqueezing transform, MSST) 的方法对信号进行时频变换, 即在原同步压缩的基础上进行多次迭代计算, 实现频率方向上的时频系数重排, 在保证时频聚集性的同时缓解了同步压缩时频特征模糊和能量发散问题。然后, 通过图像预处理算法, 进一步抑制噪声干扰, 并减少特征提取过程中的运算量。预处理过程中, 采用全国二值法选择灰度图中的最佳阈值, 有利于方向梯度直方图 (histogram of oriented gradient, HOG) 特征算法提取时频图像中信号轮廓的方向梯度值作为特征参数。同时, 为了去除高维 HOG 特征中的冗余特征量, 采用主成分分析法 (principal component analysis, PCA) 进行降维。最后, 利用支持向量机 (support vector machines, SVM) 实现雷达信号的分类识别。

## 1 MSST 时频变换与 HOG 特征提取

由于雷达信号种类繁多, 难以全面遍历研究, 选取包括 CW、LFM、NLFM、2FSK、BPSK、QPSK、LFM/FSK、LFM/BPSK 及 FSK/BPSK 在内的 9 种典型的雷达信号作为研究对象。为了实现在低信噪比下对这 9 个雷达信号的识别, 提出了一种基于 MSST 时频变换的 HOG 特征提取的新方法。

### 1.1 MSST 时频变换

MSST 是一种新的时频分析方法<sup>[11-12]</sup>, 通过在 STFT 获得的时频谱基础上进行多重同步压缩, 可有效提高时频分布图的分辨率, 并利用迭代算子优化算法流程, 减轻计算负担。同时, MSST 作为一种线性时频工具, 不存在交叉项的干扰。对于给定的雷达信号  $s(u)$ , 其 STFT 可定义为<sup>[13]</sup>

$$G(t, w) = \int_{-\infty}^{+\infty} g(u-t)s(u)e^{-jw(u-t)} du \quad (1)$$

式中:  $g(u)$  为窗函数;  $w$  为频率。

信号  $s(u)$  可进一步展开为

$$s(u) = A(u)e^{j\varphi(u)} \quad (2)$$

式中:  $A(u)$  为幅度;  $\varphi(u)$  为相位。

将信号的幅度和相位进行一阶泰勒展开:

$$\begin{cases} A(u) = A(t) \\ \varphi(u) = \varphi(t) + \varphi'(t)(u-t) \end{cases} \quad (3)$$

则信号  $s(u)$  可表示为

$$s(u) = A(u)e^{j(\varphi(t) + \varphi'(t)(u-t))} \quad (4)$$

于是信号  $s(u)$  的 STFT 时频谱可表示为

$$\begin{aligned} G(t, w) &= \int_{-\infty}^{+\infty} g(u-t)A(t)e^{j(\varphi(t) + \varphi'(t)(u-t))} e^{-jw(u-t)} du = \\ &A(t)e^{j\varphi(t)} \int_{-\infty}^{+\infty} g(u-t)e^{j(\varphi'(t)(u-t) - jw(u-t))} d(u-t) = \\ &A(t)e^{j\varphi(t)} \hat{g}(w - \varphi'(t)) \end{aligned} \quad (5)$$

对式 (5) 求偏导得

$$\partial_t G(t, w) = \partial_t (A(t)e^{j\varphi(t)} \hat{g}(w - \varphi'(t))) = G(t, w) j \varphi'(t) \quad (6)$$

当  $G(t, w) \neq 0$  时, 瞬时频率估计  $\hat{w}(t, w)$  可以表示为

$$\hat{w}(t, w) = \frac{\partial_t G(t, w)}{j G(t, w)} \quad (7)$$

再对时频谱执行同步压缩, 表达式如下:

$$T_s(t, \eta) = \int_{-\infty}^{+\infty} G(t, w) \delta(\eta - \hat{w}(t, w)) dw \quad (8)$$

式中:  $\delta(t)$  为脉冲函数。通过执行同步压缩可从频率方向上对 STFT 时频分布图进行压缩。

然而, SST 在处理复杂时变信号时, 信号的瞬时频率估计偏差增大, 时频聚集性大大降低, 导致时频模糊。而 MSST 通过迭代计算, 实现频率方向时频系数的多次重排, 可有效处理强调制信号和强时变信号, 同时, 在一定程度上能缓解同步压缩时频特征模糊和能量发散等问题。MSST 的具体迭代过程如下:

$$\begin{cases} T_s^{[2]}(t, \eta) = \int_{-\infty}^{+\infty} T_s^{[1]}(t, w) \delta(\eta - \hat{w}(t, w)) dw \\ T_s^{[N]}(t, \eta) = \int_{-\infty}^{+\infty} T_s^{[N-1]}(t, w) \delta(\eta - \hat{w}(t, w)) dw \end{cases} \quad (9)$$

式中:  $T_s^{[1]}(t, w)$  即  $T_s(t, w)$ ;  $N$  为迭代次数, 且  $N \geq 2$ 。

在  $T_s^{[1]}(t, w)$  的基础上再进行同步压缩变换, 即可得到二重同步压缩变换  $T_s^{[2]}(t, w)$ , 具体推导如下:

$$\begin{aligned} T_s^{[2]}(t, \eta) &= \int_{-\infty}^{+\infty} T_s(t, \xi) \delta(\eta - \hat{w}(t, w)) d\xi = \\ &\int_{-\infty}^{+\infty} G(t, w) \delta(\eta - \hat{w}(t, \hat{w}(t, w))) dw \end{aligned} \quad (10)$$

由式 (10) 可知, 当  $N = 2$  时, 得到一个新的瞬时频率估计  $\hat{w}(t, \hat{w}(t, w))$ , 为了证明第二次同步压缩的有效性, 需证明新瞬时频率估计比  $\hat{w}(t, w)$  更接近信号的真实瞬时频率。首先, 给定一个条件, 假设  $\varepsilon$  够小, 对于  $\forall t$ , 有  $|A'(t)| \leq \varepsilon$  和  $|\varphi'''(t)| \leq \varepsilon$ 。根据泰勒展式,  $A(u)$ 、 $\varphi(u)$  在  $t$  点展开有<sup>[14]</sup>

$$\begin{cases} A(u) = A(t) \\ \varphi(u) = \varphi(t) + \varphi'(t)(u-t) + \frac{1}{2}\varphi''(t)(u-t)^2 \end{cases} \quad (11)$$

故信号  $s(u)$  可展开为

$$s(u) = A(t)e^{j[\varphi(t) + \varphi'(t)(u-t) + 0.5\varphi''(t)(u-t)^2]} \quad (12)$$

若将窗函数指定成高斯窗  $g(u) = e^{-0.5u^2}$ , 则上述信号的 STFT 表达式为

$$G(t, w) = \int_{-\infty}^{+\infty} e^{-0.5(u-t)^2} A(t)e^{j[\varphi(t) + \varphi'(t)(u-t) + 0.5\varphi''(t)(u-t)^2]} e^{-jw(u-t)} du = A(t)e^{j\varphi(t)} \frac{1}{\sqrt{1-j\varphi''(t)}} e^{2\sqrt{1-j\varphi''(t)} \frac{(w-\varphi'(t))^2}{2}} \quad (13)$$

将式(13)代入式(7)中, 可得瞬时频率估计取实部后为

$$\hat{w}(t, w) = \varphi'(t) + \frac{(\varphi''(t))^2}{(\varphi''(t))^2 + 1} (w - \varphi'(t)) \quad (14)$$

进一步可得, 当  $N = 2$  时, 瞬时频率估计为

$$\begin{aligned} \hat{w}(t, \hat{w}(t, w)) &= \varphi'(t) + \frac{(\varphi''(t))^2}{(\varphi''(t))^2 + 1} (\hat{w}(t, w) - \varphi'(t)) = \\ \varphi'(t) + \left( \frac{(\varphi''(t))^2}{(\varphi''(t))^2 + 1} \right)^2 (w - \varphi'(t)) \end{aligned} \quad (15)$$

可求得

$$\begin{aligned} |\hat{w}(t, \hat{w}(t, w)) - \varphi'(t)| &= \left| \frac{(\varphi''(t))^2}{(\varphi''(t))^2 + 1} (\hat{w}(t, w) - \varphi'(t)) \right| \leq \\ |\hat{w}(t, w) - \varphi'(t)| \end{aligned} \quad (16)$$

式(16)表明了第二次同步压缩的瞬时频率  $\hat{w}(t, \hat{w}(t, w))$  比只压缩一次的瞬时频率  $\hat{w}(t, w)$  更加接近信号的真实瞬时频率, 即当  $N = 2$  时比  $N = 1$  能得到能量更加集中的时频谱。

同理, 可得  $N = 3$  时有

$$\begin{aligned} T_s^{[3]}(t, \eta) &= \int_{-\infty}^{+\infty} T_s^{[2]}(t, \xi) \delta(\eta - \hat{w}(t, w)) d\xi = \\ \int_{-\infty}^{+\infty} G(t, w) \delta(\eta - \hat{w}(t, \hat{w}(t, w))) dw \end{aligned} \quad (17)$$

由式(17)可以得到一个新的瞬时频率估计  $\hat{w}(t, \hat{w}(t, \hat{w}(t, w)))$ , 继续以此迭代, 最终得到多重瞬时频率估计, 令  $\hat{w}^{[N]}(t, w)$  表示 MSST 的瞬时频率估计, 则有

$$\hat{w}^{[N]}(t, w) = \varphi'(t) + \left( \frac{(\varphi''(t))^2}{(\varphi''(t))^2 + 1} \right)^N (w - \varphi'(t)) \quad (18)$$

因此, MSST 的表达式为

$$T_s^{[N]}(t, \eta) = \int_{-\infty}^{+\infty} T_s^{[N-1]}(t, w) \delta(\eta - \hat{w}(t, w)) dw \quad (19)$$

$$(t, w) \rightarrow (t, \hat{w}(t, w)) \rightarrow (t, \hat{w}(t, \hat{w}(t, w))) \rightarrow \dots \rightarrow (t, \hat{w}^{[N]}(t, w)) \quad (20)$$

由式(20)可以看出, 原始时频平面上的轨迹点从  $(t, w)$  变换为  $(t, \hat{w}^{[N]}(t, w))$ , 信号瞬时频率逐渐向真

实瞬时频率靠近, 信号瞬时频率得到有效压缩。

针对上述 9 种典型雷达信号, 图 1 给出了信噪比为 10 dB 时的 MSST 时频分布图像。可以看出, 9 种雷达信号的时频图表现出了明显差异。

## 1.2 时频图预处理

由于战场电磁环境复杂, 各种设备的存在会对接收机产生噪声干扰, 雷达时域信号时频化虽能有效降低噪声干扰, 但时频图像中仍存在大量干扰信息, 在特征提取前需先对时频图进行预处理。图 2 给出了该预处理流程及时频图变换的示意图。

时频图预处理过程可具体描述如下:

1) 将  $656 \times 875$  像素的原始图像剪裁成  $535 \times 679$  像素的图像, 去掉图像中的无效信息, 可有效降低运算量, 提高运行速度。

2) 对剪裁后的图像进行灰度化处理得到灰度图。

3) 通过维纳滤波抑制灰度图中的噪声干扰。

4) 对维纳滤波后的图像进行双三次插值, 使得图像边缘更加平滑。

5) 通过二值化处理将时频图像转换成黑白二值图像, 凸显信号轮廓的同时降低图像矩阵数据量, 最终得到  $224 \times 224$  像素的图像。

本节使用全阈值二值法来处理图像<sup>[15]</sup>。全局二值化阈值选择步骤如下:

1) 获取图像中每个灰度值的出现概率。例如,  $P_{ix}(0, i)$  为图像中第  $i$  个像素的灰度值,  $P_{ix}(1, i)$  为第  $i$  个灰度值出现的概率。

2) 计算灰度值的离散函数分布  $H(i)$ 。

$$H(i) = \sum_{j=0}^i P_{ix}(0, j) P_{ix}(1, j) \quad (21)$$

式中:  $i$  为图像中灰度值种类的总数。

3) 求前  $i$  种灰度值出现的概率之和  $P_{\text{Sum}}(i)$ 。

4) 求整体图像的灰度平均值  $\text{AGray}$ , 即图像中所有像素的灰度值之和再除以像素总数。

5) 计算不同灰度值下的阈值权重  $W_{\text{valve}}(i)$ 。

$$W_{\text{valve}}(i) = \frac{\text{AGray} \cdot H(i) - P_{\text{Sum}}(i)}{\sqrt{H(i)(1-H(i))}} \quad (22)$$

6) 将最大  $W_{\text{valve}}(i)$  值中  $i$  对应的灰度值作为全局二值化阈值。

## 1.3 HOG 特征提取

在图像处理中, HOG 特征参数一般通过统计图像中的梯度信息来描述物体的局部纹理特征。HOG 算子将图像分为若干个细胞单元 (cell), 其中每个 cell 中包含像素, 计算像素的梯度和边缘的方向直方图, 统计直方图并组合构成 HOG 特征参数。由于梯度主要存在于边缘地区, 该算法计算的局部区域梯度或方向边缘密度能够很好地描述图像中的线条边缘、拐点、形状及变化规律<sup>[16]</sup>。基于该思想, 将 HOG

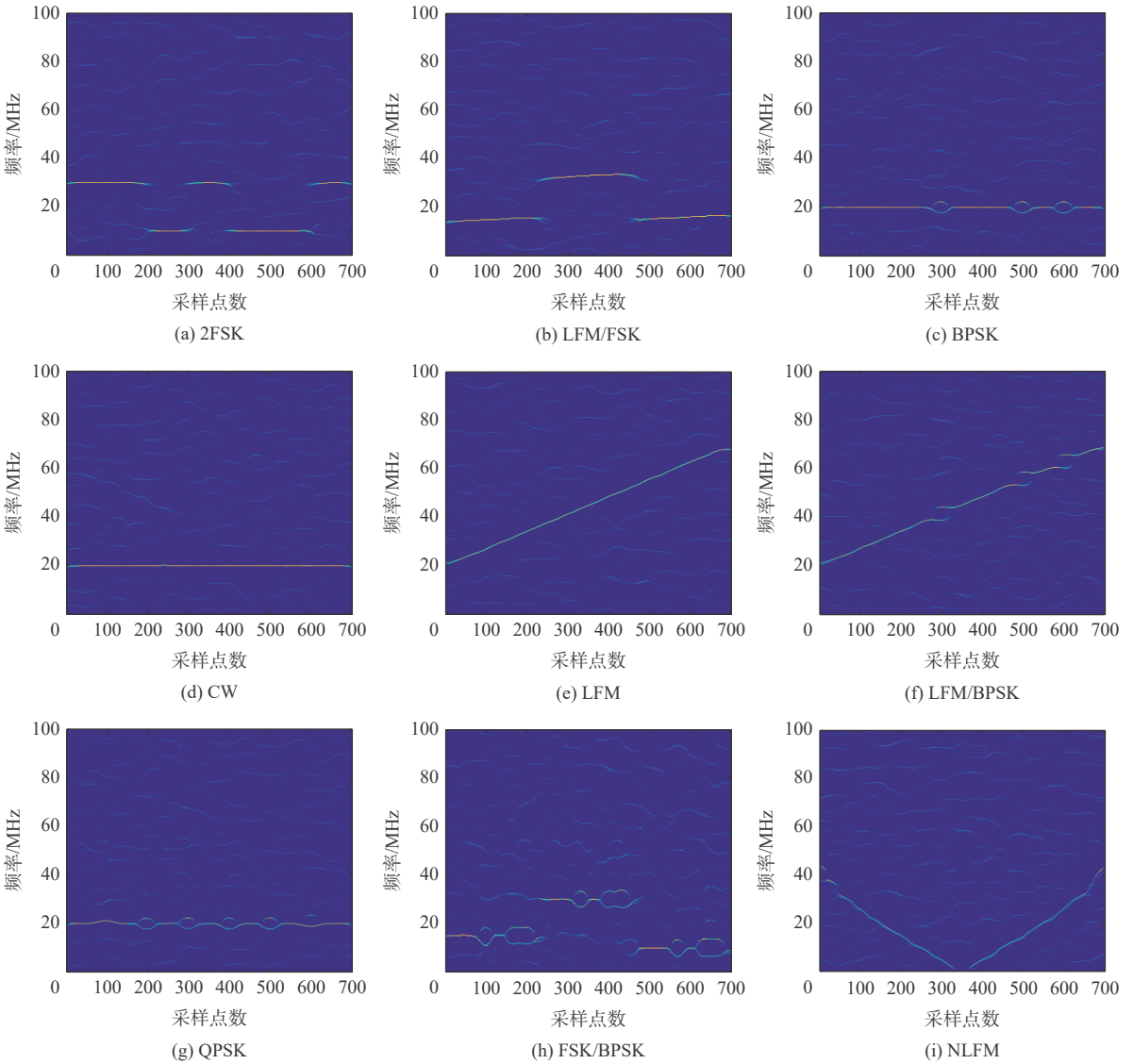


图 1 九种雷达信号的 MSST 时频图

Fig. 1 MSST time-frequency images of 9 kinds of radar signals

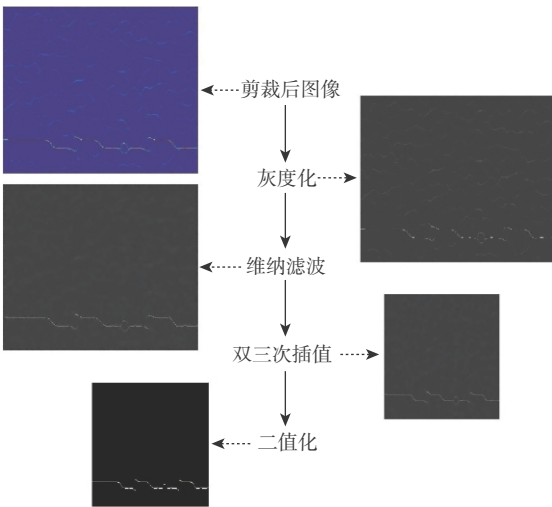


图 2 图像预处理流程

Fig. 2 Flow of image preprocessing

经过 1.1 节所述的 MSST 时频变换和二值化后, 时频聚集性更高, 线条轮廓更加清晰, 拐点更加明显, 增强了 HOG 算子提取有效梯度值作为特征参数的能力。具体的 HOG 特征提取算法实现过程如下:

1) 梯度的计算。在  $x$  方向上, 通过梯度算子  $[-1, 0, 1]$  对原图像做卷积运算, 得到在  $x$  轴方向上(向右为正方向)的梯度分量  $G_x(x, y)$ , 再在  $y$  轴上用梯度算子  $[-1, 0, 1]^T$  做同样的运算, 可得  $y$  轴方向(向上为正方向)的梯度分量  $G_y(x, y)$ ,  $H(x, y)$  为图像像素值, 如下:

$$G_x(x, y) = H(x + 1, y) - H(x - 1, y) \quad (23)$$

$$G_y(x, y) = H(x, y + 1) - H(x, y - 1) \quad (24)$$

像素点  $(x, y)$  处的梯度幅值和梯度方向如下:

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \quad (25)$$

$$\alpha(x, y) = \arctan\left(\frac{G_y(x, y)}{G_x(x, y)}\right) \quad (26)$$

特征提取引入到时频图像分析中, 以实现按时频域信号的边缘、形状和变化规律的检测。此外, 信号

2) 梯度方向直方图构建。将梯度方向  $180^\circ$  分成 9 个方向块, 如图 3 所示, 这 9 个方向块对应的角度区域作为图 4 直方图的横坐标, 再统计每一个角度区域对应的梯度信息, 得到梯度方向直方图。以表 1 中梯度方向和梯度值为例, 将像素的梯度值对应梯度方向角度映射到直方图中, 将梯度值的投影权值作为梯度方向直方图的幅值, 得到该 cell 的梯度方向直方图, 如图 4 所示, 最终统计直方图的 9 个幅值作为对应 cell 的 9 维特征向量。其中, cell 的边长为 8, 图像大小为  $224 \times 224$ , 则图像每一行有 28 个 cell, block 是由 4 个 cell 组成的正方形, 通过 block 在图像上移动进行特征提取, 每次移动步长为 1, 则一幅图共需要移动  $(224/8-1) \times (224/8-1)$  次, 也就是需要移动  $27 \times 27$  次, 因此可得一幅图的特征数为  $4 \times 9 \times 27 \times 27 = 26\,244$  个。

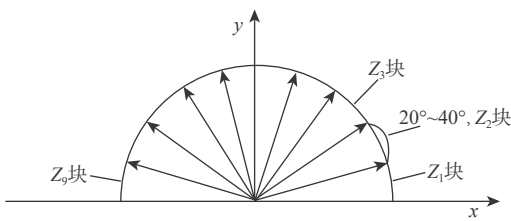


图 3 梯度方向分块

Fig. 3 Gradient direction block

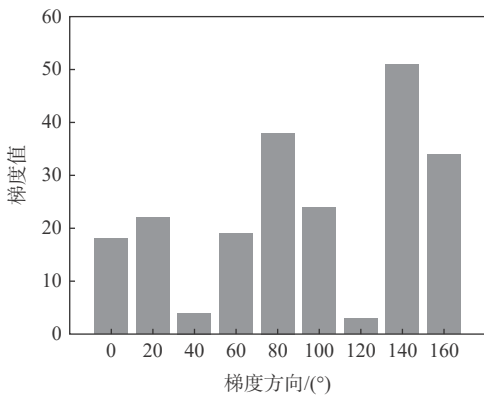


图 4 方向梯度直方图

Fig. 4 Histogram of oriented gradient

表 1 梯度方向和梯度值

Table 1 Gradient direction and gradient values

梯度方向/(°)	90	150	80	70	10	105	140	30
梯度值	30	68	4	38	36	12	17	8

#### 1.4 PCA 降维

HOG 算子提取的特征向量维度较高, 且存在大量的冗余信息, 这些冗余信息会降低分类速度和识别精度。因此, 可以通过式(27)所示的 PCA 法对 HOG 特征向量进行降维处理。

$$Y_{PCA} = W^T Z_{HOG} \quad (27)$$

式中:  $Z_{HOG}$  表示提取训练正样本 HOG 特征向量;  $Y_{PCA}$  表示降维后的特征向量。因此, 最终特征向量为协方差矩阵  $W$  的前  $K$  个特征值  $\lambda_1, \lambda_2, \dots, \lambda_K$  所对应的特征向量  $(Y_{PCA1}, Y_{PCA2}, \dots, Y_{PCA K})$ 。

## 2 识别算法流程

基于第 1 节所述, 雷达信号在经过 MSST 后得到信号时频分布图, 再在经过预处理后进行 HOG 特征提取, 将提取的 HOG 特征进行 PCA 降维, 最终采用 SVM 分类器对雷达信号的特征参数进行分类。算法整体流程如图 5 所示。

算法具体步骤如下:

- 1) 对接收到的时域雷达信号进行时频变换得到 MSST 时频图, 并构建数据集。
- 2) 将变换后时频图进行裁剪、灰度化、维纳滤波、双三次插值、二值化等图像预处理, 减少噪声干扰和冗余信息, 同时也降低数据计算量, 并凸显信号轮廓。
- 3) 对预处理后的图像进行 HOG 特征提取, 得到一组 26 244 维的特征向量。
- 4) 采用 PCA 法对 HOG 特征向量进行降维, 去除冗余信息, 提高识别速度。
- 5) 通过 SVM 分类器实现对特征向量的分类。

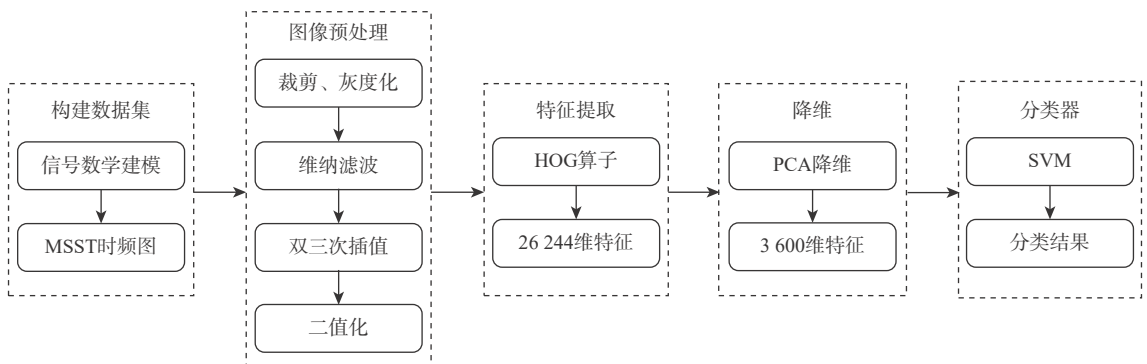


图 5 算法整体流程

Fig. 5 Overall flow chart of the proposed algorithm

### 3 实验结果与分析

针对本文提出的雷达信号识别算法, 从时频图对比、识别准确率、对比实验、鲁棒性分析 4 个方面展开仿真实验。仿真实验的计算机硬件条件为: Intel(R) Core(TM) i7-10875H CPU @ 2.30 GHz 和 16.0 GB RAM, GPU: NVIDIA GeForce RTX 2060, 仿真工具为 MATLAB。本文还搭建了半实物仿真实验平台, 针对实测数据进行了信号识别实验, 以验证算法的有效性。

#### 3.1 数据集设计

本文选取 CW、LFM、BPSK、QPSK、2FSK、NL-FM、LFM/BPSK、LFM/FSK 和 FSK/BPSK 共 9 种典型雷达信号进行实验。参数设置如下: 采样频率为 200 MHz, 载波频率为 20 MHz, 脉宽为 3.5  $\mu$ s, 带宽为 15 MHz。其中, BPSK 信号采用 7 位 Barker 码; 2FSK 有 2 个频率点 10 MHz 和 30 MHz, 复合调制参数设置与上一致。实验选取信噪比范围为 -14~10 dB, 步长为 2 dB, 每个信噪比下每种信号生成 600 张样本, 其中 400 张样本用于训练, 剩下 200 张样本用于测试, 9 种信号共生成 3 600 张训练集和 1 800 张测试集, 实验中附加的噪声为高斯白噪声。

#### 3.2 时频图对比

为了验证 MSST 相比 STFT 与 SST 具有较好的时频聚集性, 图 6 给出了 -4 dB 信噪比下 BPSK 信号的 STFT 时频图、SST 时频图及 MSST 时频图。从图 6(a) 中可以看出, BPSK 信号相位跳变处信号

发散, 信号分辨率低, 信号特性不明显, 容易与 CW 信号混淆。从图 6(b) 中可看出, SST 变换后的信号分辨率相比于 STFT 变换有一定的聚集性, 但在相位跳变处发散仍然严重, 而图 6(c) 中的信号经过 MSST 变换信号时频聚集度高, 信号特征更加明显, 相比于 STFT 变换和 SST 变换有更好的分辨率, 有利于后续 HOG 算子提取信号轮廓特征。

#### 3.3 识别准确率验证

根据第 1 节和第 2 节介绍的方法, 通过对雷达辐射源信号进行时频变换、特征提取及分类识别实验来验证本文算法的识别准确率。图 7(a) 为本文算法在不同信噪比下的识别准确率折线图。可以看出, 在信噪比不低于 -8 dB 时, 9 种雷达信号的整体识别准确率均在 96% 以上, 当信噪比为 -6 dB 时, 识别准确率接近 100%。还可以发现, 即使信噪比在较低的低 -10 dB 时, 本文算法仍可以获得 90% 以上的较高识别准确率。这是因为信号经过 MSST 后时频得到压缩, 信号线条更加明显, 再经过图像二值化进一步凸显了图像信号轮廓信息, 增强了 HOG 算子提取有效梯度值作为特征参数的能力, 从而提高了识别准确率。

从图 7(b) 可以看出, 当信噪比较低时, LFM/BPSK 信号、BPSK 信号和 LFM 信号识别准确率较低, 通过

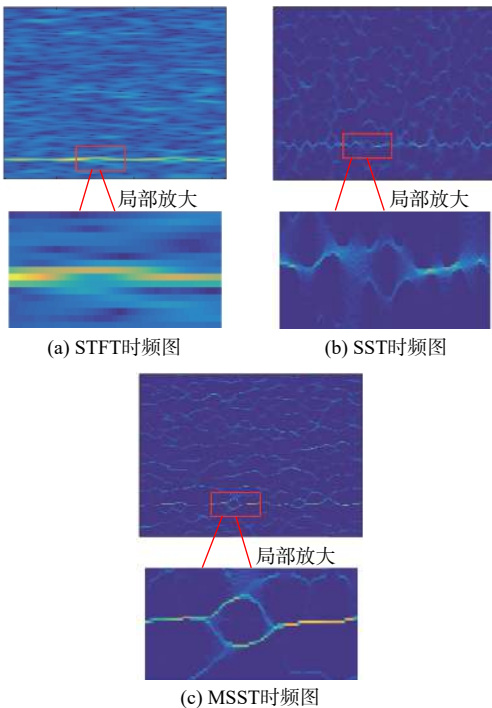
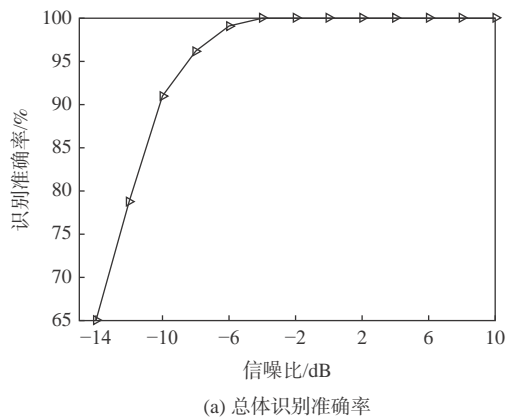
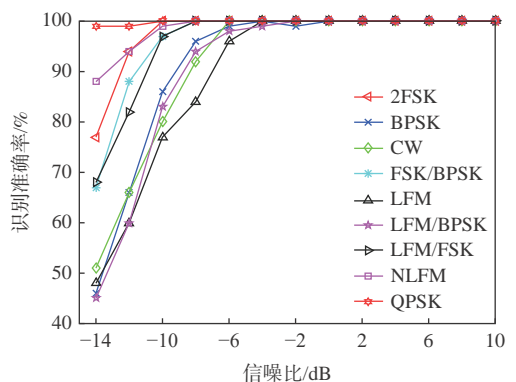


图 6 STFT、SST 和 MSST 时频图

Fig. 6 Time-frequency images of STFT, SST and MSST



(a) 总体识别准确率



(b) 9种雷达信号识别准确率

图 7 识别准确率

Fig. 7 Recognition accuracy

分析图 8 的混淆矩阵可知,信噪比为-10 dB 时,BPSK 和 CW、LFM 和 LFM/BPSK 信号之间容易引起混淆,4 种信号识别准确率在 80% 左右。这是因为随着信噪比降低,噪声对信号干扰加强,使得时频图的有效信号特征减少,从而识别准确率下降。但随着信噪比增加,当信噪比为-4 dB 时,从图 8 的混淆矩阵中可以看出,9 种信号识别准确率均达到 99% 以上。

2FSK	99.50	0	0	0	0	0	0	0	0.50
BPSK	0	83.00	16.00	1.00	0	0	0	0	0
CW	-0.50	18.50	80.50	0	0	0.50	0	0	0
FSK/BPSK	-1.00	1.00	0.50	96.50	0	0.50	0.50	0	0
LFM	0	0	0	0	75.50	24.50	0	0	0
LFM/BPSK	0	0	0	0	16.00	84.00	0	0	0
LFM/FSK	0	0	0	1.00	0	1.00	98.00	0	0
NLFM	0	0	0	0	0	0	0	98.50	1.50
QPSK	0	0	0	0	0	0	0	0	100

(a) 信噪比-10 dB, 识别准确率90.611 1%

2FSK	99.50	0	0	0	0	0	0	0	0.50
BPSK	0	99.50	0	0	0	0	0	0	0.50
CW	0	0	99.50	0	0	0	0	0	0.50
FSK/BPSK	0	0	0	100	0	0	0	0	0
LFM	0	0	0	0	100	0	0	0	0
LFM/BPSK	0	0	0	0	0	100	0	0	0
LFM/FSK	0	0	0	0	0	0	100	0	0
NLFM	0	0	0	0	0	0	0	100	0
QPSK	0	0	0	0	0	0	0	0	100

(b) 信噪比-4 dB, 识别准确率99.833 3%

图 8 九种雷达信号识别准确率混淆矩阵

Fig. 8 Confusion matrix of recognition rate of 9 kinds of radar signals

### 3.4 算法对比实验

#### 1) 算法识别准确率对比

为了验证本文算法的良好性能表现,在同等条件下与对称 Holder 系数法<sup>[13]</sup>、LBPV 法<sup>[17]</sup>及 AlexNet 模型<sup>[18]</sup>进行实验对比,实验结果如图 9 所示。

由图 9 可知,对称 Holder 系数法抗噪性较差,在信噪比低于 0 dB 时,分类器使用线性核或高斯核,识别准确率均低于 75%,主要是时域信号受噪声影响较大,从而导致可提取的有效特征参数减少。本文使用的时频图经过图像预处理后,噪声更少,二维平面可提取的有效信息更多,识别准确率相对于对称 Holder 系数法更高。LBPV 法使用的

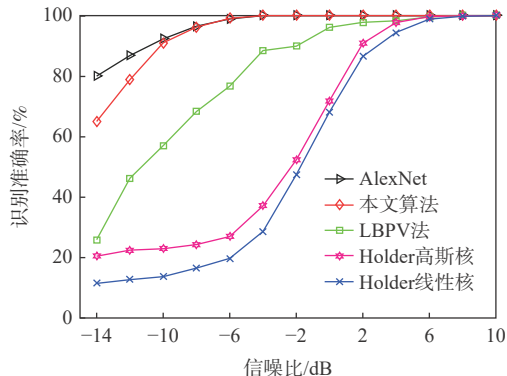


图 9 四种算法平均识别准确率对比

Fig. 9 Comparison of average recognition rates of four algorithms

LBPV 算子随着时频图成像质量下降,特征区分度不强,识别准确率下降严重,在信噪比低于-6 dB 时,识别准确率低于 80%。而本文基于 MSST 得到的时频图,频谱能量更加集中,时频图像更加精细,HOG 算子提取的特征也能更完整地表述信号,抗噪性能较好,因此在信噪比为-8 dB 时,9 种信号的整体识别准确率达到 96%,即使在更低信噪比的情况下,算法仍然能保持较高的识别准确率。AlexNet 模型在低信噪比下识别准确率相比于本文算法较好,但复杂度较高,且随着信噪比增加,识别准确率增长速率降低,当信噪比大于-6 dB 时,本文算法略高于 AlexNet 模型。

#### 2) 算法训练与测试时间对比

SVM 相比于深度学习具有复杂度低和样本需求量少的优势。由表 2 可知, AlexNet 模型训练时间较长,即使通过 GPU 并行计算,模型训练时间依旧长于本文的 SVM 模型。虽然本文使用的 SVM 模型训练样本较少,但在低信噪下依然有较高的识别准确率。较短的训练时间和良好的识别准确率使得 SVM 广泛应用于工程应用中。

同时,为了验证本文算法具有较好的识别速度,实验选取-2 dB 下的测试集进行识别,重复多次运行本文算法、对称 Holder 系数法<sup>[13]</sup>、LBPV 法<sup>[17]</sup>,求取每个算法的平均运行时间。从表 3 可知,本文算法相比于 LBPV 法识别速度更快,相比于对称 Holder 系数法运行速度略慢,但在低信噪比下,本文算法识别准确率高于对称 Holder 系数法。

表 2 算法训练时间

Table 2 Algorithm training time

算法	训练样本	训练时间/s				
		信噪比 -8 dB	信噪比 -6 dB	信噪比 -4 dB	信噪比 -2 dB	信噪比 0 dB
AlexNet	3 600	867	865	862	861	853
SVM	3 600	67	66	61	57	54

表 3 算法识别时间

Table 3 Algorithm recognition time

算法	HOG+SVM	LBPV+SVM	Holder+SVM
识别时间/s	39	48	20

### 3.5 识别鲁棒性证明

本节进一步通过测试算法在混合信噪比下的识别准确率来验证算法的鲁棒性。在不同的信噪比下进行实验, 训练集选取信噪比分别在-14 dB、2 dB、-2 dB、10 dB 条件下各生成 3 600 个样本, 测试集选取在-6 dB、10 dB、-10 dB、2 dB 下各生成 1 800 个样本。混合信噪比实验信噪比从-14 dB 到 10 dB 变化, 步长为 2 dB, 训练集针对每种信噪比选取 180 个样本, 共为 2 340 个样本, 测试集针对各信噪比选取 90 个样本, 共 1 170 个样本, 实验结果如表 4 所示。

由实验可知, 本文算法无论是在不同信噪比下还是混合信噪比下, 信号的整体识别准确率都在 90% 以上, 识别准确率较高, 具有良好的鲁棒性。

表 4 训练集和测试集在不同信噪比下的整体识别准确率

Table 4 Overall recognition rate of training set and test set under different SNR

训练集信噪比/dB	测试集信噪比/dB	总体识别准确率/%
-14	-6	93.06
2	10	100
-2	-10	90.22
10	2	97.61
-14~10	-14~10	92.02

### 3.6 半实物仿真实验

为了进一步验证本文算法的有效性, 在实验室搭建了如图 10 所示的雷达辐射源信号识别半实物仿真实验平台, 以开展信号识别实验。半实物仿真实验平台基本参数为: 支持最大收发信号带宽 200 MHz, 发送波形位宽 14 bit, 接收波形位宽 16 bit, 中频采样率 3 GSa/s, C 波段。

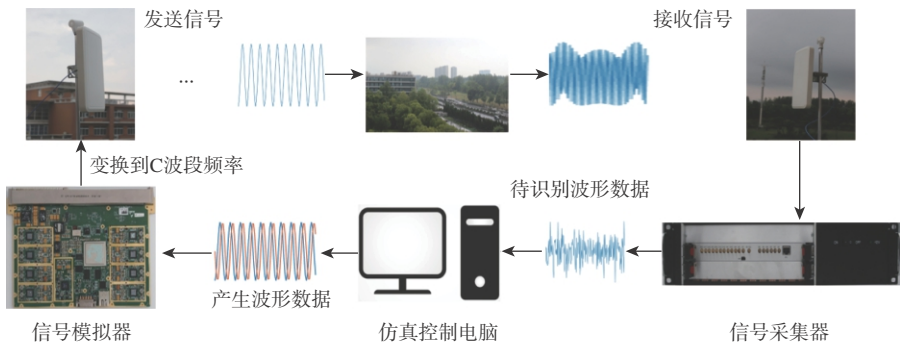


图 10 雷达辐射源信号识别半实物仿真实验平台

Fig. 10 Hardware-in-the-loop simulation experiment platform for radar radiation source signal recognition

所搭建半实物仿真平台由仿真控制电脑、信号模拟器和信号采集器构成。仿真控制电脑生成 5 400 个波形, 信号调制类别随机分布, 并对每个波形进行标注, 再将这些波形数据按照一定间隔发送到信号模拟器, 完成数模变换, 变换到 C 波段频率, 并通过天线发射到空旷区域; 同时仿真控制电脑控制信号采集器, 采集天线接收到的信号, 并且经过下变频、模数变换、脉冲提取后, 生成待识别波形数据集; 最终按照本文算法进行识别。图 11 给出了半实物仿真实验得到的识别准确率曲线, 并与 3.3 节实验结果做了对比。

从图 11 可看出, 半实物仿真实验整体识别准确率略低于仿真实验识别准确率。其原因在于: 半实物仿真实验所采集到的数据由于受到电波传播信道的影响, 信噪比会存在一定波动, 并且经传播的信号存在衰落, 造成 MSST 时频图像的变形, 最终影响了识别准确率。但在较低信噪比下, 半实物

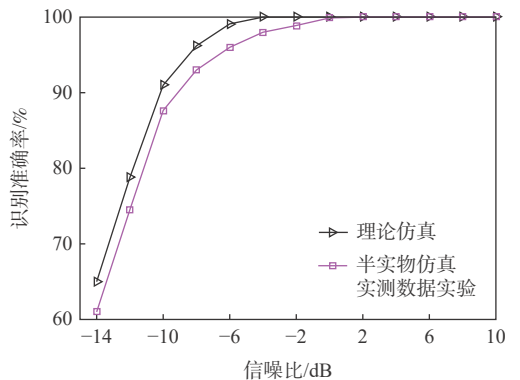


图 11 基于半实物仿真实测数据的识别准确率

Fig. 11 Recognition accuracy based on hardware-in-the-loop simulation data

仿真依然具有较高的识别准确率。当信噪比为-8 dB 时, 半实物仿真实验的识别准确率能达到 94%, 当信噪比为-2 dB 时, 识别准确率接近 100%, 表明本文算法在较为逼真的工程应用环境中也具有较好的识别准确率。

## 4 结 论

1) 针对低信噪比下雷达辐射源信号识别任务, 采用 HOG 特征对雷达信号 MSST 时频图进行特征提取, 并使用 PCA 进行降维, 最终通过 SVM 实现有效分类。

2) 通过对比实验可知, 本文算法识别性能较好, 计算复杂度低。

3) 采用半实物仿真实验进行验证。实验结果表明, 本文算法在低信噪比下对典型雷达信号具有较高的识别准确率, 为复杂电磁环境下的雷达信号识别提供了一种可行方法。

### 参考文献 (References)

- [1] KISHORE T R, RAO K D. Automatic intrapulse modulation classification of advanced LPI radar waveforms[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2017, 53(2): 901-914.
- [2] IGLESIAS V, GRAJAL J, ROYER P, et al. Real-time low complexity automatic modulation classifier for pulsed radar signals[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2015, 51(1): 108-126.
- [3] SCHLEHER D C. LPI radar: Fact or fiction[J]. *IEEE Aerospace and Electronic Systems Magazine*, 2006, 21(5): 3-6.
- [4] 黄颖坤, 金炜东, 葛鹏, 等. 基于多尺度信息熵的雷达辐射源信号识别[J]. *电子与信息学报*, 2019, 41(5): 1084-1091.
- HUANG Y K, JIN W D, GE P, et al. Radar emitter signal identification based on multi-scale information entropy[J]. *Journal of Electronics & Information Technology*, 2019, 41(5): 1084-1091(in Chinese).
- [5] 徐卓君, 杨雯婷, 杨承志, 等. 雷达脉内调制识别的改进残差神经网络算法[J]. *吉林大学学报(工学版)*, 2021, 51(4): 1454-1460.
- XU Z J, YANG W T, YANG C Z, et al. Improved residual neural network algorithm for radar intrapulse modulation recognition[J]. *Journal of Jilin University (Engineering Edition)*, 2021, 51(4): 1454-1460(in Chinese).
- [6] 刘鲁涛, 戴亮军, 陈涛. 基于频谱复杂度的雷达信号调制方式识别[J]. *哈尔滨工程大学学报*, 2018, 39(6): 1081-1086.
- LIU L T, DAI L J, CHEN T. Radar signal modulation recognition based on spectrum complexity[J]. *Journal of Harbin Engineering University*, 2018, 39(6): 1081-1086(in Chinese).
- [7] 谢存祥, 张立民, 钟兆根. 基于时频特征提取和残差神经网络的雷达信号识别[J]. *系统工程与电子技术*, 2021, 43(4): 917-926.
- XIE C X, ZHANG L M, ZHONG Z G. Radar signal recognition based on time-frequency feature extraction and residual neural network[J]. *Systems Engineering and Electronics*, 2021, 43(4): 917-926(in Chinese).
- [8] 白航, 赵拥军, 胡德秀, 等. 基于Choi-Williams时频图像特征的雷达辐射源识别[J]. *数据采集与处理*, 2012, 27(4): 480-485.
- BAI H, ZHAO Y J, HU D X, et al. Radar emitter recognition based on image feature of Choi-Williams time-frequency distribution[J]. *Data Acquisition and Processing*, 2012, 27(4): 480-485(in Chinese).
- [9] OBERLIN T, MEIGNEN S, PERRIER V. The Fourier-based synchrosqueezing transform[C]//2014 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway: IEEE Press, 2014: 315-319.
- [10] 倪雪, 王华力, 徐志军, 等. 基于STFT-SST和深度卷积网络的多相码雷达信号识别[J]. *数据采集与处理*, 2020, 35(6): 1090-1096.
- NI X, WANG H L, XU Z J, et al. Polyphase codes radar signal recognition based on STFT-SST and deep convolutional network[J]. *Data Acquisition and Processing*, 2020, 35(6): 1090-1096(in Chinese).
- [11] YU G, WANG Z, ZHAO P. Multisynchrosqueezing transform[J]. *IEEE Transactions on Industrial Electronics*, 2019, 66(7): 5441-5455.
- [12] LI D X, JIA H Y, YE Y C, et al. High power microwave signal detection based on second order multisynchrosqueezing transform[J]. *Journal of Physics:Conference Series*, 2020, 1617: 012049.
- [13] 王功明. 雷达信号脉内特征分析与识别关键技术研究[D]. 郑州: 战略支援部队信息工程大学, 2019.
- WANG G M. Research on key technology of radar signal intrapulse features analysis and recognition[D]. Zhengzhou: Strategic Support Force Information Engineering University, 2019(in Chinese).
- [14] 钱红艳. 多重同步压缩变换的提升算法及其在地震信号处理中的应用[D]. 成都: 成都理工大学, 2020.
- QIAN H Y. Lifting algorithm of multiple synchronous compression transform and its application in seismic signal processing[D]. Chengdu: Chengdu University of Technology, 2020(in Chinese).
- [15] ATAIE R, ZARANDI A A E, MEHRABANI Y S. An efficient inexact full adder cell design in CNFET technology with high-PSNR for image processing[J]. *International Journal of Electronics*, 2019, 106(6): 928-944.
- [16] 申彤, 庄建军, 黎文斯, 等. 基于HOG特征提取和支持向量机的东巴文识别[J]. *南京大学学报(自然科学)*, 2020, 56(6): 870-876.
- SHEN T, ZHUANG J J, LI W S, et al. Research on recognition of Dongba script by a combination of HOG feature extraction and support vector machine[J]. *Journal of Nanjing University (Natural Sciences)*, 2020, 56(6): 870-876(in Chinese).
- [17] 孟凡杰, 唐宏, 王义哲, 等. 基于时频图像纹理特征的雷达辐射源信号识别[J]. *弹箭与制导学报*, 2017, 37(3): 152-156.
- MENG F J, TANG H, WANG Y Z, et al. Radar emitter signal recognition based on time-frequency image texture feature[J]. *Journal of Projectiles, Rockets, Missiles and Guidance*, 2017, 37(3): 152-156(in Chinese).
- [18] 郭立民, 陈鑫, 陈涛. 基于AlexNet模型的雷达信号调制类型识别[J]. *吉林大学学报(工学版)*, 2019, 49(3): 1000-1008.
- GUO L M, CHEN X, CHEN T. Recognition of radar signal modulation type based on AlexNet model[J]. *Journal of Jilin University (Engineering Edition)*, 2019, 49(3): 1000-1008(in Chinese).

# Radar emitter signal recognition based on MSST and HOG feature extraction

QUAN Daying<sup>1,\*</sup>, TANG Zeyu<sup>1</sup>, CHEN Yun<sup>1</sup>, LOU Weizhong<sup>2</sup>, WANG Xiaofeng<sup>1</sup>, ZHANG Dongping<sup>1</sup>

(1. Key Laboratory of Electromagnetic Wave Information Technology and Metrology of Zhejiang Province,  
College of Information Engineering, China Jiliang University, Hangzhou 310018, China;

2. The 52nd Research Institute of China Electronics Technology Group Corporation, Hangzhou 310000, China)

**Abstract:** Aiming at the problem of low recognition accuracy of traditional radar signal recognition algorithms under low signal-to-noise ratio, a radar emitter recognition algorithm based on multi-synchrosqueezing transform (MSST) time-frequency transformation and histogram of direction gradient (HOG) feature extraction is proposed. The algorithm performs multiple synchronous compression processing on the basis of the short-time Fourier transform (STFT) of the radar time domain signal to obtain the signal time-frequency distribution image, then uses the HOG operator to extract the HOG feature of the signal time-frequency distribution image. The HOG features are dimensionally reduced by principal component analysis (PCA), and finally the feature parameters after dimension reduction are fed into the support vector machine (SVM) to classify and identify the radar signal. The experimental results show that the algorithm has low complexity, and when the signal-to-noise ratio is  $-8$  dB, the recognition accuracy of the simulation experiments and hardware-in-the-loop simulation experiments for 9 typical radar signals can reach more than 90%.

**Keywords:** radar signal recognition; histogram of direction gradient; multi-synchrosqueezing transform; support vector machine; principal component analysis

**Received:** 2022-05-09; **Accepted:** 2022-07-29; **Published Online:** 2022-08-05 14:18

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20220805.1340.002.html](https://kns.cnki.net/kcms/detail/11.2625.V.20220805.1340.002.html)

**Foundation items:** Zhejiang Provincial Natural Science Foundation of China (LQ20F020021); Open Project Funding of the Key Laboratory of Electromagnetic Wave Information Technology and Metrology of Zhejiang Province (2019KF0003)

\* **Corresponding author.** E-mail: [qdy@cjlu.edu.cn](mailto:qdy@cjlu.edu.cn)

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0265

# 超低轨卫星气动舵机辅助姿态控制方法设计

王涛<sup>1,2</sup>, 焦洪臣<sup>2,\*</sup>, 刘杰<sup>2</sup>, 陈乐宇<sup>3</sup>, 张迎春<sup>1</sup>

(1. 哈尔滨工业大学 航天学院, 哈尔滨 150001; 2. 中国空间技术研究院 遥感卫星总体部, 北京 100094;

3. 北京航空航天大学 宇航学院, 北京 100191)

**摘 要:** 针对超低轨卫星姿态控制差异化需求, 开展了基于气动舵机辅助的姿态控制策略研究。完成了超低轨道稀薄大气下卫星气动舵机布局设计与气动特性研究, 理论气动力可达  $10^{-1}$  N 量级, 气动力矩可达  $10^{-1}$  N·m 量级。在此基础上, 完成了基于气动舵机辅助的姿态控制策略研究。通过仿真验证, 在  $x$  轴采用动量轮控制、 $y$  轴和  $z$  轴采用气动舵机辅助控制情况下, 可实现优于  $0.004^\circ$  的三轴指向精度和优于  $0.0007(^\circ)/s$  的三轴姿态稳定度。所设计气动舵机辅助姿态控制策略对超低轨卫星技术应用与发展具有重要技术价值和工程意义。

**关键词:** 超低轨卫星; 气动舵机; 空气动力学; 姿态控制; 卫星构型

**中图分类号:** V474.2; TB553

**文献标志码:** A

**文章编号:** 1001-5965(2023)03-0548-11

超低轨卫星存在传统卫星不可比拟的优势。超低轨卫星轨道高度低, 能够快速到达预定轨道, 展开工作; 搭载的空间相机或雷达等设备指标无需太高, 就能够达到传统卫星的效果, 甚至更好, 有效载荷这一项就能够大幅度减少卫星成本; 由于体积、质量和成本相对传统遥感卫星可大幅缩减, 超低轨卫星的部署和发射方式极为灵活, 并且可通过多星组网的方式实现对特定区域优异的时间、空间分辨率<sup>[1-4]</sup>。

在早期的研究中, 气动力和力矩通常被视为干扰项, 需要进行控制或低效, 在实际的工程项目中也是如此应对。美国研发的 NanoEye 卫星是一种用于对地观测的图像卫星, 其运行的轨道高度范围为  $200 \sim 300$  km, 外形设计考虑到超低轨卫星所受气动力的影响, 主结构为圆柱形, 而在前端装有 2 块太阳能电池板, 构成一个楔子的形状, 减小作用在卫星上的气动力<sup>[5]</sup>。俄罗斯的 Yantar 系列及 Orlets 系列卫星均运行在超低轨道上, 近地点高度约为 200 km, 远地点高度约为 300 km。此外, 俄罗斯的

部分“琥珀”系列卫星也运行在超低轨道上, 为了抵抗气动力导致的轨道衰减, 第四代中的“琥珀-2K”卫星定期提升轨道, 从而使近地点高度保持在  $170 \sim 180$  km, 远地点高度保持在  $320 \sim 350$  km<sup>[6]</sup>。受气动阻力影响, 这些卫星具有较强的对地观测能力, 但工作寿命较短。

随着航天技术的发展, 研究人员开始进行利用气动力辅助超低轨卫星姿态控制方面的研究<sup>[7-11]</sup>。Kumar 等<sup>[12-13]</sup>研究了气动被动稳定控制, 对气动稳定和磁阻尼稳定进行仿真, 证明了气动被动稳定控制的可行性。Psiaki<sup>[14-15]</sup>研究微纳卫星的控制时, 利用气动力矩和磁力矩进行被动稳定控制, 仿真结果表明, 气动力矩和磁力矩控制系统对于微纳卫星姿态控制效果良好。Guettler<sup>[16]</sup>在 2007 年提出利用卫星受到的气动力矩实现卫星姿态主动控制, 通过线性模型对卫星姿态运动建模仿真, 证明了在较低轨道上才具有通过气动力矩进行姿态控制的可行性。此外, 研究人员发现低轨下的气动力还可用于卫星动量轮角动量卸载<sup>[17-19]</sup>。欧洲航天局于 2009 年

收稿日期: 2021-05-20; 录用日期: 2021-10-22; 网络出版时间: 2021-11-02 13:56

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20211101.1713.015.html](https://kns.cnki.net/kcms/detail/11.2625.V.20211101.1713.015.html)

基金项目: 国家自然科学基金(62005015)

\*通信作者. E-mail: [jhccast@163.com](mailto:jhccast@163.com)

**引用格式:** 王涛, 焦洪臣, 刘杰, 等. 超低轨卫星气动舵机辅助姿态控制方法设计 [J]. 北京航空航天大学学报, 2023, 49 (3): 548-558.

WANG T, JIAO H C, LIU J, et al. Design of attitude control method for ultra-low-orbit satellite with pneumatic steering gear [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 548-558 (in Chinese).

3月17日发射 GOCE 卫星,该卫星发射高度为 275 km,在轨运行高度为 250~260 km,主要用于重力场和海洋环流探测。卫星机身呈八边形棱柱体,长约 5.3 m,横截面为 1.1 m<sup>2</sup>。在卫星的后部安装有 2 对尾翼,当卫星姿态有偏移时,尾翼可以产生气动力矩以修正姿态,但由于超低轨道空气密度较低,这种被动控制方式调整缓慢,只是被用于辅助主动控制<sup>[20]</sup>。日本于 2017 年 12 月 23 日发射 SLATS 卫星,该卫星首先保持 271.5 km 的轨道高度,随后利用气动力制动,最终在 167.4 km 高度实现短期运行<sup>[21]</sup>。

针对满足高分辨率、高重访、全天时和快速响应需求的超低轨遥感卫星,研究满足其空间环境和任务需求差异化特点的姿态控制策略。不同于以往将空间大气环境对卫星的影响视为摄动干扰,通过补偿技术对其进行修正,本文引入航空空气动力学相关理论对控制手段和方法进行必要的补充和改进,在卫星姿态控制策略方面进行创新性研究,以气动舵机作为辅助手段,有效降低卫星传统控制机构压力,实现超低轨环境下卫星姿态的高精度保持。

## 1 超低轨卫星气动力和气动力矩特性

### 1.1 大气密度模型

本文采用航天领域常用的大气模型——COSPAR 国际标准大气模型(CIRA)来确定大气密度,这是国际上搜集来的,通过测量成千上万颗卫星由于大气阻力造成的轨道衰减的效应的数据,最终加以确定,高度在 500 km 以下的大气密度可以采用指数衰减模型进行计算。指数衰减模型为

$$\rho = \rho_0 e^{-\frac{(h-h_0)}{H}} \quad (1)$$

式中:  $h$  为实际高度;  $h_0$  为基准高度;  $\rho_0$  为基准面  $h=h_0$  处的大气密度;  $H$  为密度标高,由于密度的变化随高度的变化越到高层变化越慢,  $H$  应随高度缓慢放大。一个较好的近似是假定  $H$  随高度线性变化,具体为

$$H = H(r) = H_0 + \frac{\mu}{2}(h - h_0) \quad (2)$$

式中:  $\mu \approx 0.1$ ;  $H_0$  为基准密度标高。

若引用 CIRA-72 大气模式,在高度为 150 km 处,各基准数据取值为:  $\rho_0 = 2.07 \times 10^{-9} \text{ kg/m}^3$ ,  $H_0 = 24.6 \text{ km}$ ,  $\mu = 0.1$ 。

后续仿真分析中,大气密度模型参照上述模型在程序中预先设定,依据卫星轨道高度实时计算当前大气密度,并应用到气动计算模型中,以求得较为精确的气动力和力矩。

### 1.2 气动计算模型

超低轨卫星运行环境属于高层大气,气体稀

薄,分子平均自由程远大于卫星特征长度,因此,对于卫星受到的气动力与气动力矩模型按照自由分子流的理论进行分析。此时,卫星受到的气动力和力矩依赖于多种因素,如大气密度、卫星构型、表面材料特性、来流速度及矢量方向等。

在稀薄气体动力学的 3 个基本假设的前提下,基于气体分子与表面相互作用的麦克斯韦模型,可得到表面压力与剪切应力如下<sup>[22-23]</sup>:

$$p = \frac{\rho_\infty V_\infty^2}{2S^2} \left\{ \left( \frac{2 - \sigma_n}{\sqrt{\pi}} S \sin \theta_c + \frac{\sigma_n}{2} \sqrt{\frac{T_w}{T_\infty}} \right) e^{-(S \sin \theta_c)^2} + \left[ (2 - \sigma_n) \left( \frac{1}{2} + (S \sin \theta_c)^2 \right) + \frac{\sigma_n}{2} \sqrt{\frac{\pi T_w}{T_\infty}} (S \sin \theta_c) \right] [1 + \text{erf}(S \sin \theta_c)] \right\} \quad (3)$$

$$\tau = \frac{-\sigma_\tau \rho_\infty V_\infty^2 \cos \theta_c}{2\sqrt{\pi} S} \left\{ e^{-(S \sin \theta_c)^2} + \sqrt{\pi} (S \sin \theta_c) [1 + \text{erf}(S \sin \theta_c)] \right\} \quad (4)$$

式中:  $\rho_\infty$  为来流密度;  $V_\infty$  为来流速度;  $T_w/T_\infty$  为壁面温度和来流温度比;  $S$  为分子速度比,定义为  $S = V_\infty / (2kT_\infty/m)^{1/2}$ ,  $k$  为玻尔兹曼常数,  $m$  为单个分子质量;  $\sigma_n$  和  $\sigma_\tau$  分别为法向和切向动量适应系数;  $\theta_c$  为来流和表面的夹角;  $\text{erf}(S \sin \theta)$  为误差函数,定义为

$$\text{erf}(S \sin \theta_c) = (2/\sqrt{\pi}) \int_0^{S \sin \theta_c} e^{-x^2} dx \quad (5)$$

气动力计算采用当地化方法,基本原理与稀薄气体动力学相同,只是定义来流与表面夹角  $\theta_c$  是来流方向与  $x$  轴的夹角<sup>[24]</sup>。因此,当地压力系数和剪切力系数分别为

$$C_n = P_{1l} + P_{2l} \cos \theta_c + P_{3l} \cos^2 \theta_c \quad (6)$$

$$C_\tau = \sin \theta_c (P_{4l} + P_{5l} \cos \theta_c) \quad (7)$$

式中:系数  $P_{1l} \sim P_{5l}$  分别为

$$\begin{cases} P_{1l} = \frac{2 - \sigma_n}{2S^2} [1 + \text{erf}(S \cos \theta_c)] + \frac{\sigma_n}{2S^2} \sqrt{\frac{T_w}{T_\infty}} e^{-S^2 \cos^2 \theta_c} \\ P_{2l} = \frac{\sigma_n}{2S} \sqrt{\frac{\pi T_w}{T_\infty}} [1 + \text{erf}(S \cos \theta_c)] + \frac{2 - \sigma_n}{\sqrt{\pi} S} e^{-S^2 \cos^2 \theta_c} \\ P_{3l} = (2 - \sigma_n) [1 + \text{erf}(S \cos \theta_c)] \\ P_{4l} = \frac{\sigma_\tau}{\sqrt{\pi} S} e^{-S^2 \cos^2 \theta_c} \\ P_{5l} = \sigma_\tau [1 + \text{erf}(S \cos \theta_c)] \end{cases} \quad (8)$$

各符号取值如下:  $\theta_c$  满足  $\cos \theta_c > 0$ ;  $\sigma_n$ 、 $\sigma_\tau$  约为 0.8;  $S$  约为 3~14, 180~200 km 高度上,可取  $S=8$ 。

采用当地化方法得到气动力的压力系数和剪切力系数公式,对应气动力和气动力矩为

$$\begin{cases} dF_n = C_n \frac{1}{2} \rho V^2 dA_n \\ dF_\tau = C_\tau \frac{1}{2} \rho V^2 dA_\tau \end{cases} \quad (9)$$

当地微元总的气动力为其相加之和:

$$dF = dF_n n + dF_\tau t = (C_n n + C_\tau t) \frac{1}{2} \rho V^2 dA \quad (10)$$

而整星受到的总气动力,需要对式(10)进行曲面积分:

$$F = \iint_{\text{Surf}} (C_n n + C_\tau t) \frac{1}{2} \rho V^2 dA \quad (11)$$

式中: Surf 表示卫星表面曲面。

有了气动力,将其与每个微元相对质心的矢径  $r$  相乘,再进行曲面积分,就能够得到整星受到的气动力矩:

$$M = \iint_{\text{Surf}} [r \times (C_n n + C_\tau t)] \frac{1}{2} \rho V^2 dA \quad (12)$$

## 2 气动舵机辅助控制的超低轨卫星构型设计

根据整星初步构型和主要设备的分布方式,建立卫星模型如图1所示。主体部分横截面为平行四边形,且在其正  $x$  端面中心位置有中心空腔,为设想采用的吸气式电推进进气道,同时具有降低前向气动阻力的作用。在星上载荷与整星平台间采用隔震连接,从而降低超低轨气动环境及推进系统引入的振动影响。共4幅气动舵机分布于整星结构四边位置,每块舵机翼板由主翼和副翼组成,主翼转轴过整星质心,副翼转轴设置于主翼上,卫星质心位于其几何中心。整星初步设计尺寸如图2所示。

图2中,垂直舵机在星体坐标系  $xz$  平面内沿  $x$  轴对称分布,其型心连线过卫星质心且相对质心对称分布,主翼转动角度  $\alpha$  与副翼转动角度  $\beta$  的定义如图3所示。

根据图3构型可知,当主翼带动副翼同时转动

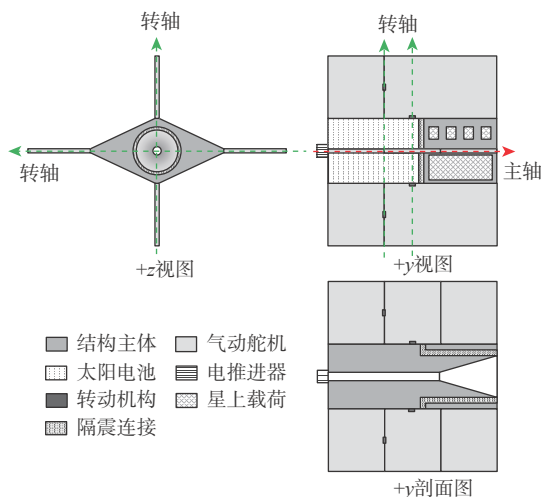
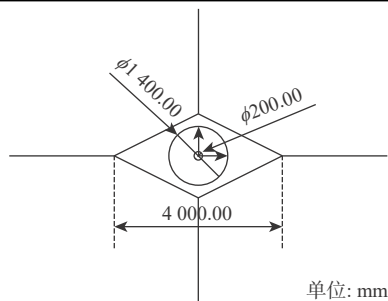
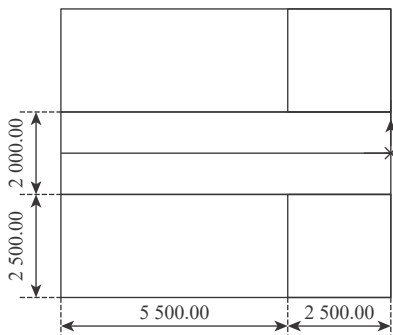


图1 超低轨卫星气动构型示意图

Fig. 1 Pneumatic structure of ultra-low-orbit satellite



(a) 卫星+z视图构型尺寸示意



(b) 卫星-y视图构型尺寸示意

图2 超低轨卫星初步设计尺寸

Fig. 2 Preliminary design dimension of ultra-low-orbit satellite

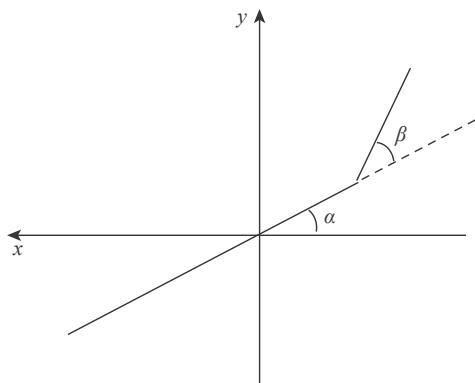


图3 超低轨卫星主副翼转角定义示意图

Fig. 3 Definition of rotation angles of main wings and ailerons of ultra-low-orbit satellite

时(即  $\beta=0$ ),仅产生过卫星质心的推力,而不会造成星体旋转;与之相对,当副翼存在转动角度时(即  $\beta \neq 0$ ),翼面引入的气动合力将偏离质心位置,从而形成转动力矩,继而改变卫星姿态。

## 3 气动特性的仿真结果及分析

气动力和力矩的计算结果受到很多因素的影响,结合卫星在轨运行过程中实际状况及姿轨控制的需要,针对以下几种情况进行计算分析:①在姿态控制过程中,卫星的姿态会有较大变化,因此计算气动力和力矩随卫星姿态角的变化情况;②利用气动力和力矩辅助卫星的姿轨控制是通过气动翼的偏转实现的,因此计算气动力和力矩随气动翼转

角的变化情况;③卫星在轨运行过程中,来流方向不断改变,导致卫星所受气动力和力矩发生变化,因此计算气动力和力矩随来流方向的变化情况。其他因素对气动力和力矩的影响较为单一且变化范围较小,因此设置为定值。仿真设计过程中初始状态参数设置如表 1 所示。

表 1 仿真参数设置

Table 1 Simulation parameters setting

参数	实际高度 $h/\text{km}$	来流速度 $V_\infty/(\text{m}\cdot\text{s}^{-1})$	法向动量 适应系数 $\sigma_n$	切向动量 适应系数 $\sigma_t$
数值	200	7 500	0.99	0.99

1) 姿态角变化

与一般工程实际的应用方式一致,卫星本体坐标系相对轨道坐标系的坐标转换矩阵,采用欧拉角进行描述,按照  $z$ - $y$ - $x$ (3-2-1)顺序旋转;绕  $x_b$  轴转动角度  $\varphi$ ,定义为滚转角;绕  $y_b$  轴转动角度  $\theta$ ,定义为俯仰角;绕  $z_b$  轴转动角度  $\Psi$ ,定义为偏航角。

先计算在各气动力转角均为零、来流方向指向  $x$  轴负方向的情况下,气动力和力矩随卫星俯仰角和偏航角的变化。由于在 本体坐标系下,当来流向量沿  $x$  轴负方向时,单独滚转角的变化对卫星气动受力没有影响,计算的只是气动力和力矩随偏航角及俯仰角的变化情况。而实际上,由于来流方向与  $x$  轴存在夹角,滚转角变化也会带来气动力的微弱变化,但卫星的构型沿  $x$  轴对称,而且来流方向与  $x$  轴方向的夹角很小,滚转角变化导致的气动力变化较小。

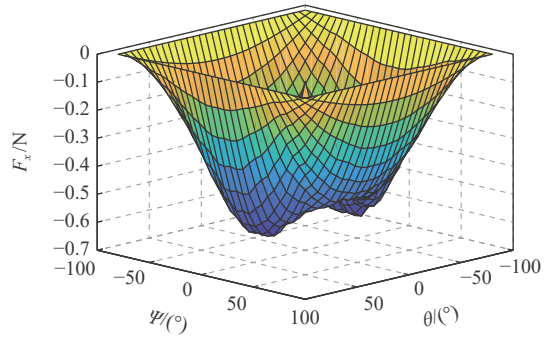
仿真沿各轴向气动力和气动力矩随姿态角度的变化情况,如图 4~图 6 所示。在气动力方面,如图 4(a)所示,  $F_x$  恒小于零;如图 5(a)所示,  $F_y$  主要受偏航角影响,且与偏航角呈正相关,受俯仰角影响较小;如图 6(a)所示,  $F_z$  主要受俯仰角影响,当偏航角较小时,与俯仰角呈负相关。三轴的气动力量级均为  $10^{-1}$  N。气动力矩方面,如图 4(b)所示,当姿态角较大时,  $M_x$  变化较大;如图 5(b)所示,  $M_y$  主要受俯仰角影响,符号与俯仰角一致;如图 6(b)所示,  $M_z$  符号与俯仰角一致,且当各角较大时变化较大。  $M_x$  量级为  $10^{-2}$  N·m,  $M_y$ 、 $M_z$  量级为  $10^{-1}$  N·m。

为了对气动力矩受姿态角影响进行更准确地分析,截取部分仿真切面,分别在偏航角和俯仰角为零时,变动另一姿态角,对受到的气动力矩进行分析,结果如图 7 所示。

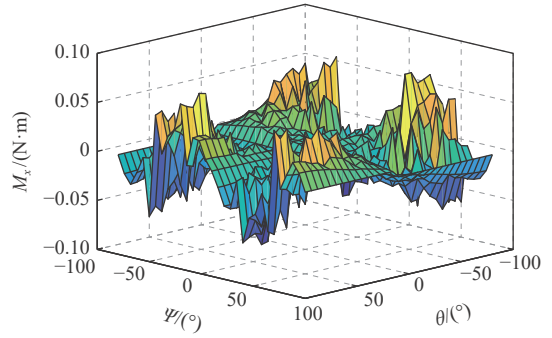
从分析结果可知,偏航角变化影响的主要是  $M_z$ ,其他轴力矩在零附近变化。俯仰角变化影响的主要是  $M_y$ ,其他轴力矩在零附近变化。上述力矩变化量级在  $10^{-3}\sim 10^{-2}$  N·m。

2) 来流方向变化

为简化计算过程,仿真中以极地轨道为参考,



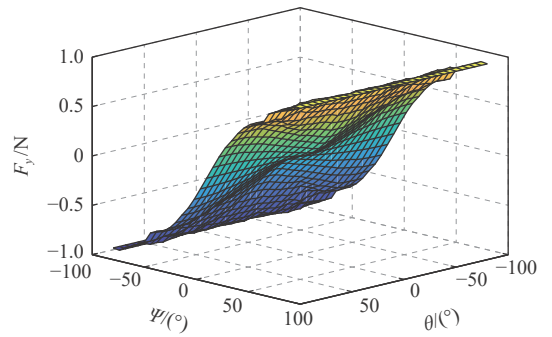
(a)  $F_x$  随姿态角变化



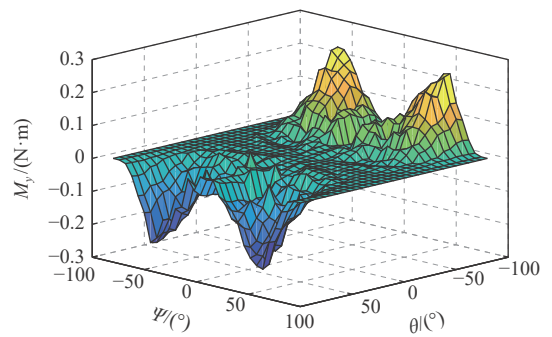
(b)  $M_x$  随姿态角变化

图 4  $x$  轴气动力及力矩随姿态角变化

Fig. 4 Change of aerodynamic force and moment of  $x$ -axis with satellite attitude



(a)  $F_y$  随姿态角变化



(b)  $M_y$  随姿态角变化

图 5  $y$  轴气动力及力矩随姿态角变化

Fig. 5 Change of aerodynamic force and moment of  $y$ -axis with satellite attitude

此时来流向量  $x$  轴和  $z$  轴分量始终不变,  $y$  轴分量会随着卫星运动呈现周期性变化。来流向量的  $y$

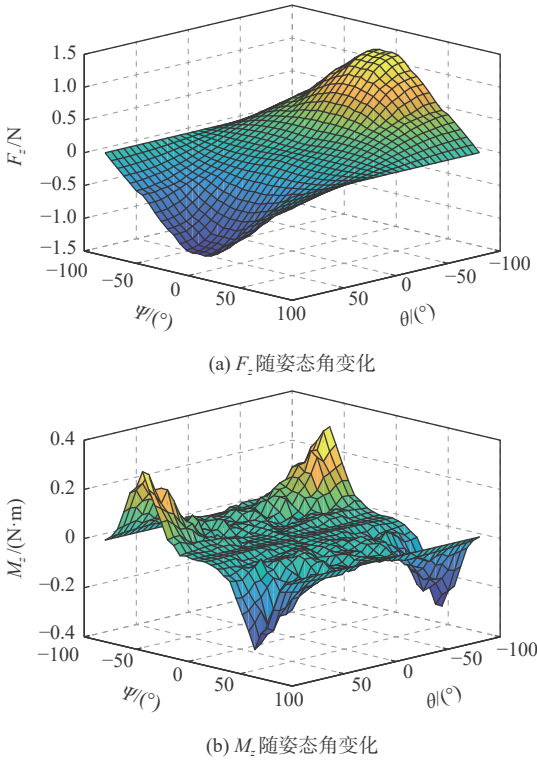


图6 z轴气动力及力矩随姿态角变化

Fig. 6 Change of aerodynamic force and moment of z-axis with satellite attitude

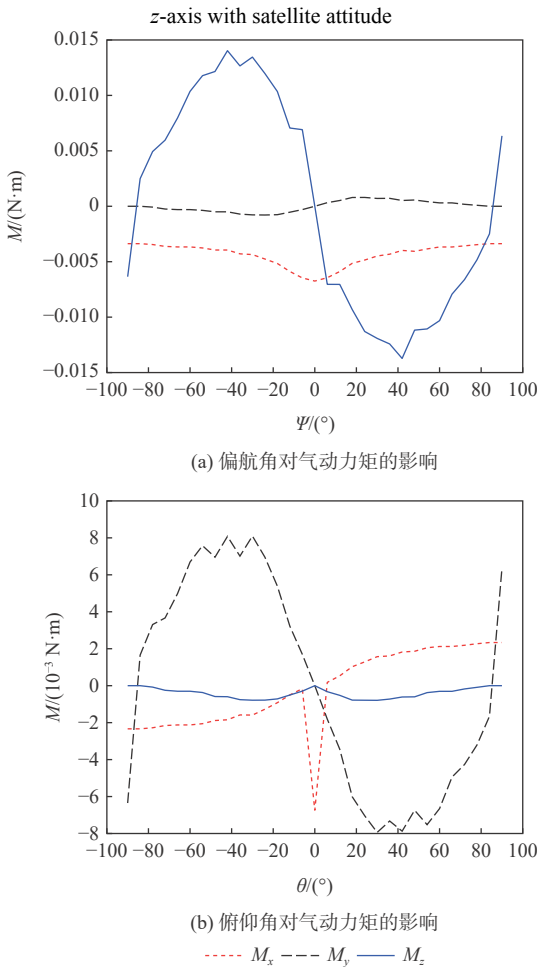


图7 星体姿态角度变化对气动力矩的影响

Fig. 7 Change of aerodynamic moment with satellite attitude

轴分量在 $-0.09\sim 0.09$ 范围内变化,对应来流向量与轨道坐标系 $x$ 轴负方向的夹角范围为 $-5.14^\circ\sim 5.14^\circ$ 。设置3个姿态角为零,各气动翼转角为零,根据来流向量的变化计算气动力和力矩的变化规律,结果如图8所示。

气动力方面, $F_x$ 始终为负,来流偏转越大,其值越大,量级为 $10^{-1}$  N; $F_y$ 与来流 $y$ 轴分量大致呈线性正相关,量级为 $10^{-2}$  N; $F_z$ 为零。气动力矩方面, $M_x$ 为负且变化很小, $M_y$ 在零附近轻微浮动, $M_z$ 正负与来流向量 $y$ 轴分量相反,量级为 $10^{-3}\sim 10^{-2}$  N·m。

### 3) 气动舵机偏转

在实际运用到气动力和力矩时,更多是各姿态角为零而气动翼转角不为零,因此设置3个姿态角为零,令来流指向 $x$ 轴负方向,计算气动力和力矩随气动翼转角的变化情况。偏转主翼面,计算气动力随主翼面转角的变化规律,结果如图9所示;偏转副翼面,计算气动力矩随副翼面转角的变化规律,结果如图10所示。

当主翼偏转时,气动阻力 $F_x$ 始终为负,且随其绝对值偏转角增大而增大,量级为 $10^{-1}$  N;侧向力量

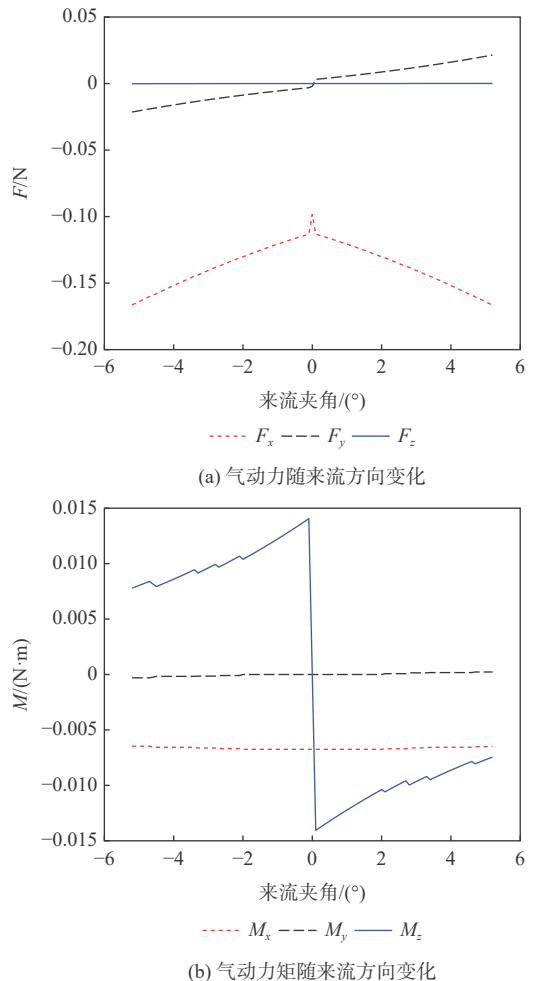
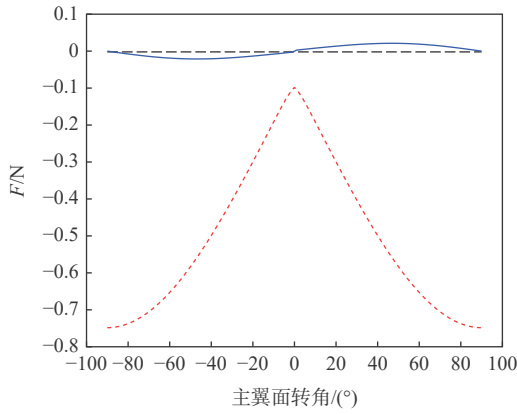
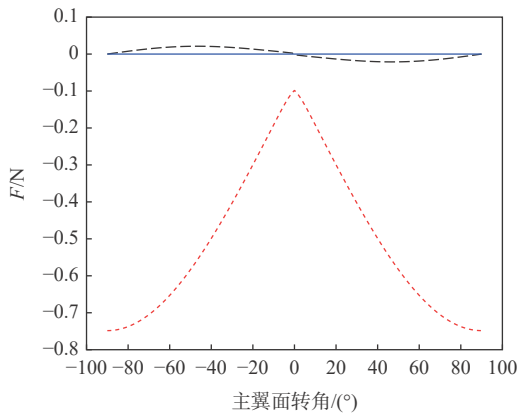


图8 不同来流方向下气动力与气动力矩

Fig. 8 Change of aerodynamic force and moment with air flow direction



(a) 不同水平主翼面转角对应的气动力

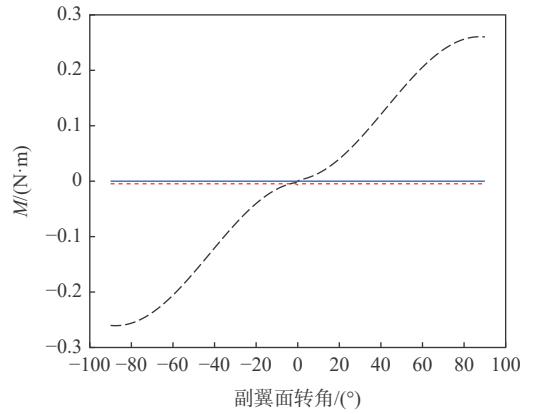


(b) 不同垂直主翼面转角对应的气动力

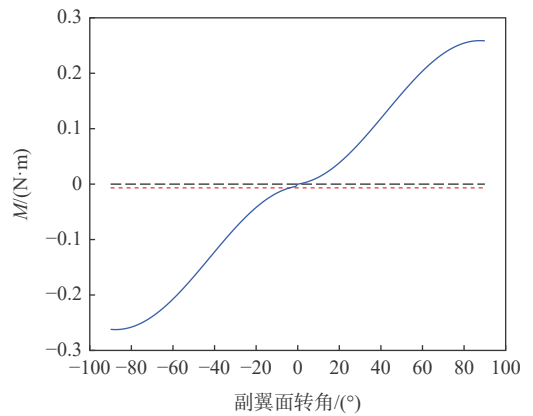
---  $F_x$  ---  $F_y$  —  $F_z$

图 9 气动力随主翼面转角的变化

Fig. 9 Change of aerodynamic force with rotation angle of main wings



(a) 不同水平副翼面转角对应的气动力矩



(b) 不同垂直副翼面转角对应的气动力矩

---  $M_x$  ---  $M_y$  —  $M_z$

图 10 气动力矩随副翼面转角的变化

Fig. 10 Change of aerodynamic moment with rotation angle of ailerons

级为  $10^{-2}$  N, 水平主翼面转角为正时  $F_z$  为正, 转角为负时  $F_z$  为负; 垂直主翼面转角为正时  $F_y$  为负, 转角为负时  $F_y$  为正, 可以利用这一特点通过偏转主翼产生所需气动力以辅助相对轨道位置维持。

当副翼偏转时, 相应轴的气动力矩随之变化, 可以达到  $10^{-1}$  N·m 量级, 而通常动量轮组只能产生  $10^{-2} \sim 10^{-1}$  N·m 的力矩, 气动力矩可以对姿态运动产生很大的影响, 其姿态调整能力与传统动量轮调姿方式相当, 因此完全可将其应用于姿态稳定控制中。水平副翼偏转导致  $M_y$  变化, 二者大致线性正相关; 垂直副翼偏转导致  $M_z$  变化, 二者大致线性正相关, 可以利用这一特点通过偏转副翼产生所需气动力矩以辅助姿态控制。

由上述分析可知, 分别偏转水平副翼和垂直副翼, 产生  $y$  轴和  $z$  轴气动力矩, 由于气动力矩较大, 可以为这两轴姿态控制提供控制力矩, 相应动量轮则不工作, 因此在此过程中  $y$ 、 $z$  轴动量轮角动量积累为零。由于  $x$  轴气动力矩是由侧向力提供的, 相较于另外两轴气动力矩有数量级的差距, 本文不考虑利用  $x$  轴气动力矩辅助姿态控制。

## 4 气动舵机辅助的姿态控制策略设计与仿真

### 4.1 气动舵机辅助的姿态控制策略设计

由第 3 节气动特性仿真分析结果可知, 当副翼偏转时, 相应轴的气动力矩随之变化, 而且在一定偏转范围内二者大致线性正相关。可利用这一关系, 通过副翼偏转提供所需气动力矩, 辅助动量轮进行姿态稳定控制。

基于气动舵机辅助的姿态控制框图如图 11 所示。具体控制过程如下:

1) 通过 PID 控制律, 根据期望姿态与实际姿态的偏差求得整星所需总的控制力矩  $T_{c\_total}$ , 并通过控制分配得到需要副翼提供的气动力矩  $T_{ca}$  和动量轮提供的控制力矩  $T_c$ 。其中, 动量轮控制部分采用传统方式, 不再赘述。

2) 利用所需的气动力矩  $T_{ca}$  求得副翼所需偏转角度  $\beta = \beta_a T_{ca} / T_a$ , 实际控制过程中通过控制副翼偏转角度进行气动舵机辅助姿态控制。其中,  $\beta_a$  为气动力矩和副翼转角大致线性相关的边界角度, 根据实际气动力矩曲线确定,  $T_a$  为  $\beta_a$  对应的气动力矩。

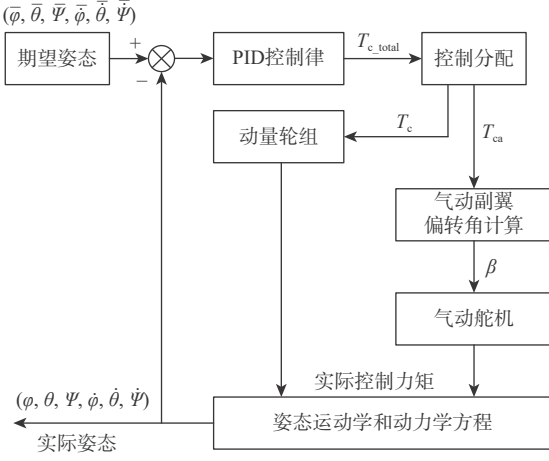


图 11 采用气动舵机辅助的姿态控制框图

Fig. 11 Block diagram of attitude control with pneumatic steering gear

3) 令相应轴的副翼偏转角度为  $\beta$ , 代入到气动计算函数中, 可得到副翼偏转角度为  $\beta$  时卫星受到的实际气动力矩。

假设卫星为地对地三轴稳定航天器, 姿态动力学可用如下方程描述:

$$J\dot{\omega}_{bl}^b + \omega_{bl}^b \times J\omega_{bl}^b = T_{c\_total} + T_e \quad (13)$$

式中:  $J$  为整星惯量矩阵;  $\omega_{bl}^b$  为卫星本体相对惯性坐标系的角速度在本体坐标系的分量;  $T_{c\_total}$  为整星控制力矩;  $T_e$  为其他干扰力矩。

经过小角度简化后, 采用欧拉角描述的姿态动力学方程为

$$\begin{cases} \ddot{\varphi} = \frac{T_x - (J_z - J_y)(\dot{\theta} - \omega_0)(\dot{\Psi} + \omega_0\varphi) + J_x\omega_0\dot{\Psi}}{J_x} \\ \ddot{\theta} = \frac{T_y - (J_x - J_z)(\dot{\Psi} + \omega_0\varphi)(\dot{\varphi} - \omega_0\Psi)}{J_y} \\ \ddot{\Psi} = \frac{T_z - (J_y - J_x)(\dot{\theta} - \omega_0)(\dot{\varphi} - \omega_0\Psi) + J_z\omega_0\dot{\varphi}}{J_z} \end{cases} \quad (14)$$

式中:  $T_x$ 、 $T_y$ 、 $T_z$  为包括控制力矩及卫星受到的其他干扰力;  $\omega_0$  为卫星轨道角速度。

为使姿态运动进行解耦控制, 设计 PID 控制律为

$$\begin{cases} T_{xc\_total} = -K_{px}\Delta\varphi - K_{ix}\int_0^t \varphi dt - K_{dx}\Delta\dot{\varphi} + \\ \quad (J_z - J_y)(\dot{\theta} - \omega_0)(\dot{\Psi} + \omega_0\varphi) - J_x\omega_0\dot{\Psi} \\ T_{yc\_total} = -K_{py}\Delta\theta - K_{iy}\int_0^t \theta dt - K_{dy}\Delta\dot{\theta} + \\ \quad (J_x - J_z)(\dot{\Psi} + \omega_0\varphi)(\dot{\varphi} - \omega_0\Psi) \\ T_{zc\_total} = -K_{pz}\Delta\Psi - K_{iz}\int_0^t \Psi dt - K_{dz}\Delta\dot{\Psi} + \\ \quad (J_y - J_x)(\dot{\theta} - \omega_0)(\dot{\varphi} - \omega_0\Psi) \end{cases} \quad (15)$$

式中:  $K_p$ 、 $K_i$ 、 $K_d$  分别为比例系数、积分系数、微分系数;  $\Delta\varphi$ 、 $\Delta\theta$ 、 $\Delta\Psi$  分别为当前姿态角与目标姿态角

的差值;  $J_x$ 、 $J_y$ 、 $J_z$  分别为各轴惯量。

通过式(15)计算得到期望控制力矩  $T_{c\_total}$ , 之后  $x$  轴通过动量轮组响应,  $y$ 、 $z$  轴通过舵面偏转响应, 产生各轴实际控制力矩。调整选取合适的控制参数, 分别取  $K_p = \{5, 2.8, 5\}$ ,  $K_d = \{140, 148, 210\}$ ,  $K_i = \{0.06, 0.06, 0.06\}$ 。

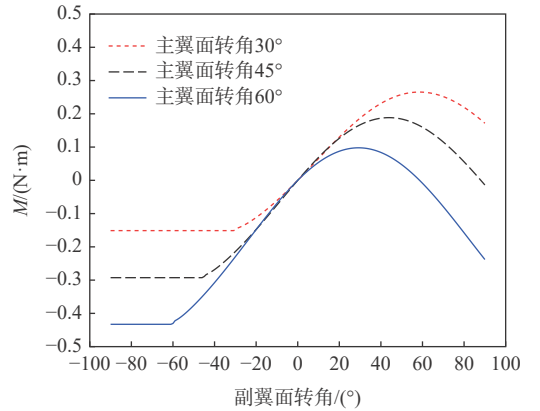
在充分考虑系统误差和环境干扰的前提下, 可实现卫星姿态高精度、高动态范围的控制。结合传统高分辨率中低轨遥感卫星姿态控制精度要求, 本文所设计基于气动舵机辅助的超低轨卫星满足如下姿态控制精度: ①控制指向精度:  $\leq 0.005^\circ$  (三轴  $3\sigma$ ); ②姿态稳定度:  $\leq 0.0005^\circ/s$  (三轴  $3\sigma$ )。

下面对气动力矩随副翼面转角的变化规律进行综合分析, 实现利用所需的气动力矩  $T_{ca}$  求得副翼所需偏转角度  $\beta$ 。

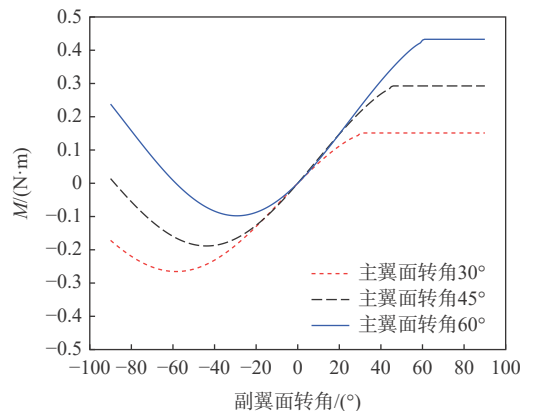
计算主翼偏转不同角度时气动力矩随副翼面转角的变化情况, 分别针对主翼面转角为正和为负作图, 如图 12 所示。

结合主翼面转角为零时气动力矩随副翼面转角的变化规律, 得到表 2。

上述分析是针对垂直翼面进行计算得到的, 由此前分析可知, 水平和垂直翼面的偏转导致的气动



(a) 主翼面转角为正时不同副翼面转角对应的气动力矩



(b) 主翼面转角为负时不同副翼面转角对应的气动力矩

图 12 主翼偏转不同角度时气动力矩随副翼面转角的变化  
Fig. 12 Change of aerodynamic moment with rotation angle of ailerons for different rotation angles of main wings

表 2 气动力矩与副翼偏转对照关系

Table 2 Relationship between aerodynamic moment and rotation angle of ailerons

主翼面转角	所需气动力矩为正	所需气动力矩为负
主翼面转角为正	(20°, 0.11 N·m)	(-20°, -0.13 N·m)
主翼面转角为零	(80°, 0.256 N·m)	(-80°, -0.256 N·m)
主翼面转角为负	(20°, 0.13 N·m)	(-20°, -0.11 N·m)

变化规律基本一致, 因此利用水平翼面偏转辅助姿态稳定控制时各项参数选取与此一致。

### 4.2 气动舵机辅助的姿态控制策略仿真结果及分析

开展仿真分析工作的仿真环境由 SolidWorks、ANSYS、MATLAB 等软件组成, 主要仿真步骤如下:

**步骤 1** 基于设计的卫星构型, 通过 SolidWorks 软件进行三维建模。

**步骤 2** 将三维模型导入 ANSYS 软件进行有限元网格划分, 进行微元特征计算。

**步骤 3** 将微元特征数据文件导入 MATLAB 软件进行气动力和气动力矩计算分析, 并依此开展基于气动辅助的姿态控制策略设计。

**步骤 4** 通过 MATLAB 软件建立卫星姿态控制数学模型, 本节主要开展此部分工作, 仿真工况设置为通过副翼偏转产生的气动力矩进行 y 轴和 z 轴姿态控制, 通过动量轮进行 x 轴姿态控制。

**步骤 5** 通过对仿真结果进行分析, 对所设计控制策略的有效性进行验证。

开展气动舵机辅助姿态控制策略仿真分析的主要流程如图 13 所示。

卫星运动的数学模型包含大量的微分方程, 因此总的思路为采用四阶龙格库塔法求解微分方程。

根据以往卫星设计经验, 初步设置仿真参数及初始值如表 3 所示。

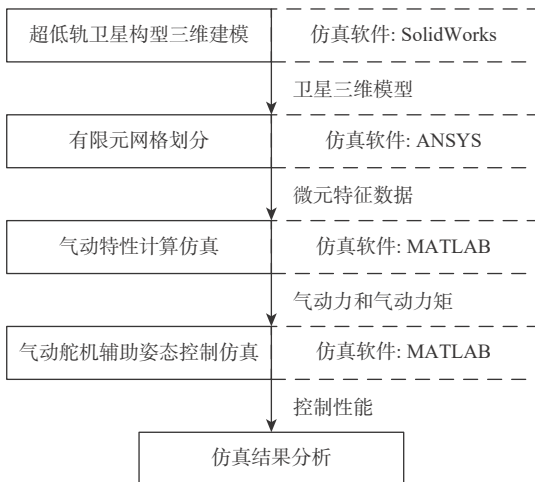


图 13 仿真流程

Fig. 13 Flowchart of simulation

动量轮控制周期及舵机控制周期均为 0.2 s。y 轴和 z 轴通过舵机偏转产生气动力矩进行姿态控制, 设置舵机偏转角度上限为 34.5°, x 轴由动量轮进行姿态控制。

由于在卫星实际运行过程中, 还存在各类测量误差及控制偏差, 为使仿真过程尽可能接近实际情况, 在可能出现误差的地方设置随机干扰, 如表 4 所示。

在上述情况下, 通过对姿态角控制的仿真, 获得气动舵机辅助下卫星姿态角和角速度变化曲线分别如图 14 和图 15 所示。

计算得到三轴的控制指向精度为 [0.000 7°, 0.003 3°,

表 3 仿真参数说明

Table 3 Description of simulation parameters

参数	数值
初始姿态角/(°)	[-10,20,10]
初始姿态角速度/(°·s <sup>-1</sup> )	[0,0,0]
x轴转动惯量J <sub>x</sub> /(kg·m <sup>2</sup> )	600
y轴转动惯量J <sub>y</sub> /(kg·m <sup>2</sup> )	2 000
z轴转动惯量J <sub>z</sub> /(kg·m <sup>2</sup> )	2 000
动量轮转动惯量/(kg·m <sup>2</sup> )	0.008
动量轮组每根轴输出力矩上限/(N·m)	0.2
动量轮组每根轴角动量上限/(N·m·s)	3
动量轮初始动量/(N·m·s)	[0,0,0]
期望姿态角/(°)	[0,0,0]
期望姿态角速度/(°·s <sup>-1</sup> )	[0,0,0]

表 4 仿真加入的随机干扰

Table 4 Random disturbance in simulations

误差类型	随机干扰的3σ
姿态角测量偏差/(°)	[0.01,0.01,0.01]
姿态角速度测量偏差/(°·s <sup>-1</sup> )	[0.001,0.001,0.001]
动量轮组角加速度控制偏差/(°·s <sup>-2</sup> )	0.001
气动翼转角控制偏差/(°)	0.01

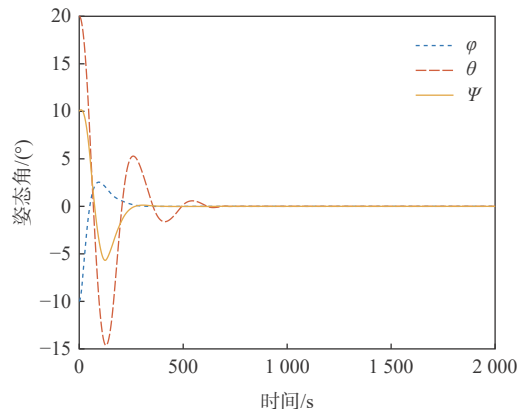


图 14 气动舵机辅助下的卫星姿态稳定过程

Fig. 14 Stabilization process of satellite attitude under assistance of pneumatic steering gear

$0.0019^\circ$ ];姿态稳定度为 $[0.00006, 0.00066, 0.00043](^\circ)/s$ ,与预想的姿态控制要求相差不大。

气动舵机偏转角度仿真结果如图16所示,舵机气动转矩、气动干扰力矩和动量轮输出力矩的仿真结果如图17~图19所示。

动量轮输出力矩在 $y$ 、 $z$ 轴分量均为0,只有 $x$ 轴分量不为0,舵机偏转分别在 $y$ 、 $z$ 轴产生气动控制力矩,上述结果表明成功实现 $x$ 轴动量轮组控制,  $y$ 、 $z$ 轴气动控制的功能。

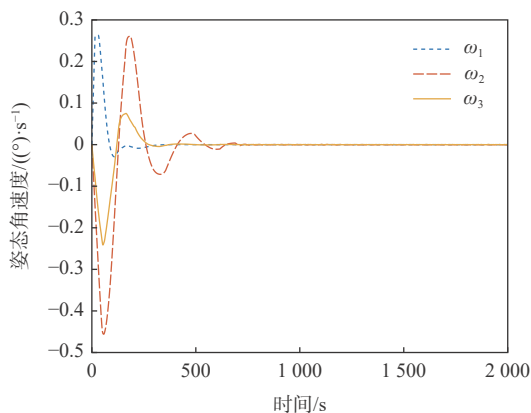


图15 气动舵机辅助下的卫星姿态角速度稳定过程

Fig. 15 Stabilization process of satellite angular velocity under assistance of pneumatic steering gear

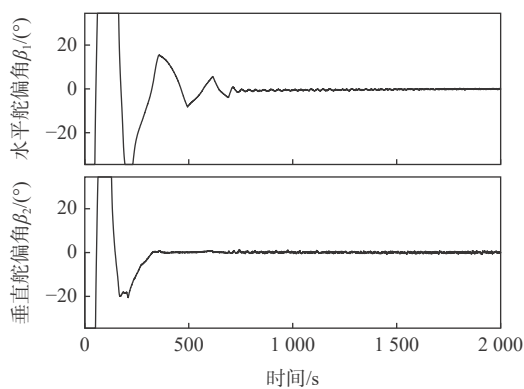


图16 气动舵机偏转角度

Fig. 16 Deflection angle of pneumatic steering gear

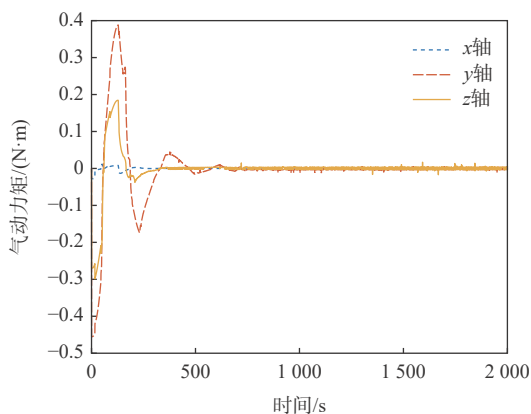


图17 舵机气动力矩

Fig. 17 Aerodynamic torque of steering gear

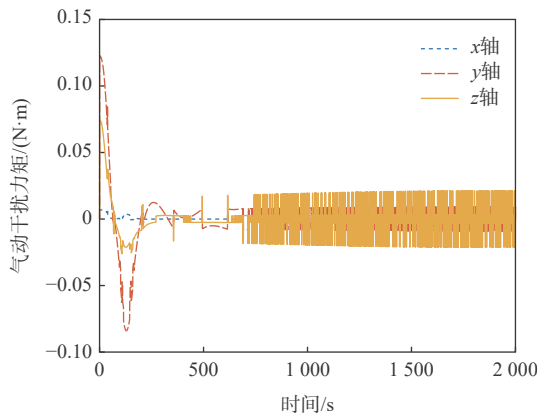


图18 舵机气动干扰力矩

Fig. 18 Aerodynamic disturbance torque of steering gear

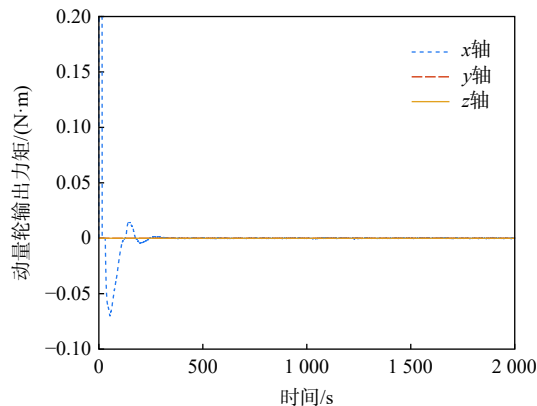


图19 动量轮输出力矩

Fig. 19 Output torque of momentum wheel

## 5 结论

1) 本文研究了基于气动舵机辅助的超低轨卫星姿态控制策略,并进行了计算仿真。利用稀薄气体动力学中的自由分子流理论对气动力与气动力矩进行了建模;结合超低轨大气环境特点和气动舵机辅助需求,设计构建了超低轨气动卫星构型;在此基础上,针对不同姿态角、来流方向和气动翼偏转角度下的气动力和气动转矩进行了数值计算和仿真分析,确定了在当前设计下气动力可达 $10^{-1}$  N量级、气动力矩可达 $10^{-1}$  N·m量级,其姿态调整能力与传统动量轮调姿方式相当,充分验证了气动舵机辅助超低轨卫星姿态控制策略的可行性。

2) 开展了基于气动力辅助的姿态控制策略设计及仿真分析,数值结果表明,超低轨卫星的三轴指向精度可达 $[0.0007^\circ, 0.0033^\circ, 0.0019^\circ]$ ,姿态稳定度可达 $[0.00006, 0.00066, 0.00043](^\circ)/s$ 。上述结果已满足传统高分辨率中低轨遥感卫星的姿态控制精度要求,同时也验证了基于气动舵机辅助的超低轨卫星姿态控制策略具有差异化的技术先进性和广阔的应用前景。

在未来研究中,将基于已有研究成果,深化论证方案的工程可实现方法,同时进一步探索基于气动舵机辅助的超低轨卫星轨道机动方法,将超低轨卫星姿轨控全流程与气动舵机辅助紧耦合,从而为实现高分辨新体制遥感卫星提供有力的技术支撑。

### 参考文献 (References)

- [1] BACON A, OLIVIER B S. Bringing down the cost of earth observation[C]//Proceedings of the 12th Reinventing Space Conference. Berlin: Springer, 2017: 1-7.
- [2] ROBERT P C E, ROMANO F, HERDRICH G, et al. Keynote: Discoverer-making commercial satellite operations in very low earth orbit a reality[C]//70th International Astronautical Congress (IAC). Reston: AIAA, 2019: 21-25.
- [3] DE FLORIO S, D'AMICO S, RADICE G. Virtual formation method for precise autonomous absolute orbit control[J]. *Journal of Guidance, Control, and Dynamics*, 2014, 37(2): 425-438.
- [4] DE FLORIO S, D'AMICO S, RADICE G. Precise autonomous orbit control in low earth orbit[C]//Astrodynamics Specialist Conference. Reston: AIAA, 2012: 4811.
- [5] WERTZ J R, SHAO A, TAYLOR C, et al. Moderately elliptical very low orbits (MEVLOs) as a long-term solution to orbital debris[C]//26th Annual AIAA/USU Conference on Small Satellites. Reston: AIAA, 2012: SSC12-IV-6.
- [6] 吴勤. 透视俄罗斯军用卫星发展现状[J]. *太空探索*, 2008(12): 46-49.
- WU Q. Perspective of Russian military satellite development status[J]. *Space Exploration*, 2008(12): 46-49(in Chinese).
- [7] 曾其铤. 气动力矩在超低轨道卫星姿态控制方面的应用研究[D]. 哈尔滨: 哈尔滨工业大学, 2009: 41-56.
- ZENG Q J. Applications of aerodynamic torque to ultra-low-orbit satellite attitude control[D]. Harbin: Harbin Institute of Technology, 2009: 41-56(in Chinese).
- [8] 田春华, 马广富, 李传江, 等. 三轴稳定卫星姿态控制系统的一般性问题[J]. *自动化技术与应用*, 2001(1): 9-12.
- TIAN C H, MA G F, LI C J, et al. General problems of satellite attitude control[J]. *Techniques of Automation and Applications*, 2001(1): 9-12(in Chinese).
- [9] 黄静. 三轴稳定航天器姿态最优控制方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2010: 75-78.
- HUANG J. Optimal attitude control for three-axis stabilized spacecrafts[D]. Harbin: Harbin Institute of Technology, 2010: 75-78(in Chinese).
- [10] 段广仁, 钟震, 姜苍华. 航天器的一种无源自适应姿态控制方法[J]. *哈尔滨工业大学学报*, 2011, 43(5): 1-7.
- DUAN G R, ZHONG Z, JIANG C H. One scheme of passive adaptive attitude control for spacecraft[J]. *Journal of Harbin Institute of Technology*, 2011, 43(5): 1-7(in Chinese).
- [11] 邵汉斌. 基于滑模控制的超低轨道航天器姿态控制方法研究[D]. 长沙: 国防科学技术大学, 2014: 12-16.
- SHAO H B. Ultra-low-orbit spacecraft attitude control based on sliding mode control[D]. Changsha: National University of Defense Technology, 2014: 12-16(in Chinese).
- [12] KUMAR R R, MAZANEK D D, HECK M L. Simulation and Shuttle Hitchhiker validation of passive satellite aerostabilization[J]. *Journal of Spacecraft and Rockets*, 1995, 32(5): 806-811.
- [13] KUMAR R R, MAZANEK D D, HECK M L. Parametric and classical resonance in passive satellite aerostabilization[J]. *Journal of Spacecraft and Rockets*, 1996, 33(2): 228-234.
- [14] PSIAKI M L. Spacecraft attitude stabilization using passive aerodynamics and acting magnetic torquing: AIAA-2003-5420[R]. Reston: AIAA, 2003.
- [15] PSIAKI M L. Magnetic torquer attitude control via asymptotic periodic linear quadratic regulation[J]. *Journal of Guidance, Control, and Dynamics*, 2001, 24(2): 386-394.
- [16] GUETTLER D B. Satellite attitude control using atmospheric drag [D]. Dayton: Air Force Institute of Technology, 2007: 23-33.
- [17] 刘辉, 伍斯宾斯基. 利用喷气装置卸载航天器积累角动量的最小工质损耗控制[J]. *航天控制*, 2004, 22(5): 32-35.
- LIU H, USPENSKY V B. Minimize propellant consumption during gyro system unloading process of spacecraft[J]. *Aerospace Control*, 2004, 22(5): 32-35(in Chinese).
- [18] 张利宾. 基于磁控和轮控的微小卫星姿态控制算法研究[D]. 哈尔滨: 哈尔滨工业大学, 2007: 26-51.
- ZHANG L B. Study on attitude control algorithm for micro-satellite using magnetotorquers and reaction wheels[D]. Harbin: Harbin Institute of Technology, 2007: 26-51(in Chinese).
- [19] 李太玉, 张育林. 基于能量最优解析解的飞轮磁卸载方法[J]. *上海航天*, 2006, 23(6): 1-9.
- LI T Y, ZHANG Y L. Momentum magnetic unloading basing on analytic equation of energy optimization[J]. *Aerospace Shanghai*, 2006, 23(6): 1-9(in Chinese).
- [20] STEIGER C, ROMANAZZO M, EMANUELLI P P. The deorbiting of ESA's gravity mission GOCE-Spacecraft operations in extreme drag conditions[C]//International Conference on Space Operations. Reston: AIAA, 2014: 1934.
- [21] 何慧东. 日本“超低轨道技术试验卫星”任务及应用[J]. *国际太空*, 2018(9): 50-53.
- HE H D. Japan's super low altitude test satellite mission and application[J]. *Space International*, 2018(9): 50-53(in Chinese).
- [22] 温生林. 超低轨道卫星动力学建模与控制方法研究[D]. 长沙: 国防科学技术大学, 2016: 20-35.
- WEN S L. Dynamic modeling and flight control for super low altitude satellite[D]. Changsha: National University of Defense Technology, 2016: 20-35(in Chinese).
- [23] 周伟勇, 张育林, 刘昆. 超低轨航天器气动力分析与减阻设计[J]. *宇航学报*, 2010, 31(2): 342-348.
- ZHOU W Y, ZHANG Y L, LIU K. Aerodynamics analysis and reduced drag design for the lower LEO spacecraft[J]. *Journal of Astronautics*, 2010, 31(2): 342-348(in Chinese).
- [24] 胡凌云, 张立华, 程晓丽, 等. 超低轨航天器气动设计与计算方法探讨[J]. *航天器工程*, 2016, 25(1): 10-18.
- HU L Y, ZHANG L H, CHENG X L, et al. Method of aerodynamic design and calculation for ultra-LEO spacecraft[J]. *Spacecraft Engineering*, 2016, 25(1): 10-18(in Chinese).

# Design of attitude control method for ultra-low-orbit satellite with pneumatic steering gear

WANG Tao<sup>1,2</sup>, JIAO Hongchen<sup>2,\*</sup>, LIU Jie<sup>2</sup>, CHEN Leyu<sup>3</sup>, ZHANG Yingchun<sup>1</sup>

(1. School of Astronautics, Harbin Institute of Technology, Harbin 150001, China;

2. Institute of Remote Sensing Satellite, China Academy of Space Technology, Beijing 100094, China;

3. School of Astronautics, Beihang University, Beijing 100191, China)

**Abstract:** To meet the differentiated needs of the attitude control of ultra-low orbit satellites, this study investigates the attitude control strategy with the assistance of pneumatic steering gears. The layout of the gear is designed and its aerodynamic characteristics are analyzed under the thin atmosphere of the ultra-low orbit, with the theoretical aerodynamic force up to the order of  $10^{-1}$  N, and the aerodynamic torque the order of  $10^{-1}$  N·m. On this basis, an attitude control strategy assisted by pneumatic steering gears is designed. Simulation results show that when the  $x$ -axis is controlled by the momentum wheel and the  $y$ -axis and  $z$ -axis are controlled by pneumatic steering gears, the three-axis pointing accuracy larger than  $0.004^\circ$  and the three-axis attitude stability larger than  $0.0007(^\circ)/s$  can be achieved. The attitude control strategy designed in this paper has important technical and engineering value for the application and development of ultra-low orbit satellites.

**Keywords:** ultra-low-orbit satellite; pneumatic steering gear; aerodynamics; attitude control; satellite structure

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0270

# 航天遥感器用泵驱两相流体回路热真空试验研究

孟庆亮<sup>1,2,\*</sup>, 赵振明<sup>1,2</sup>, 陈祥贵<sup>1,2</sup>, 朱许<sup>1,2</sup>

(1. 北京空间机电研究所, 北京 100094; 2. 先进光学遥感技术北京市重点实验室, 北京 100094)

**摘 要:** 针对航天遥感器核心组件的高精度与高稳定度的控温需求, 设计并搭建了一套泵驱两相流体回路 (MPTL) 试验装置, 该装置使用了一套具有被动冷却能力的两相控温型储液器。为验证 MPTL 系统在高真空、极低温与变化外热流条件下的工作能力, 在真空罐内对 MPTL 系统在不同工况点下的散热与控温能力进行了测试, 并通过温度和压力等数据研究了主回路的运行特性、储液器内热力学变化特性及两者之间的传热传质过程。结果表明: MPTL 系统的控温点可通过储液器进行快速调整, 蒸发器温度的变化受外热流与热源开关影响较小; 进入毛细管中的过冷液与储液器中的液相形成的温差保证了储液器冷量的供应; 主回路发生相态转变时, 储液器与主回路工质交换特性引起了系统压力降脉动。

**关键词:** 泵驱两相流体回路; 两相控温型储液器; 热控产品; 热真空试验; 航天遥感器

**中图分类号:** V416; TK124

**文献标志码:** A **文章编号:** 1001-5965(2023)03-0559-10

随着航天遥感技术的不断发展, 光学遥感器核心组件——光电转换器件对温度的要求越来越高, 这些器件需要满足较低的温度变化和较高的温度稳定度的条件<sup>[1-2]</sup>。光电转换器件对工作环境温度的要求非常严格, 温度波动过大会增大器件的暗电流和热噪声, 降低信噪比, 从而降低分辨率。为满足空间遥感相机高精度成像需要, 光电转换器件的温度在全寿命周期内要满足 $\pm 2.0\text{ }^{\circ}\text{C}$ 稳定性的要求。泵驱两相流体回路 (mechanically pumped two-phase loop, MPTL) 技术是一项先进的两相流热量管理技术, 具有控温精度高、传热距离远、传热量大等优点, 其在航天热控领域具有广阔的应用前景, 特别适用于航天相机、活动天线、大功率激光器等对控温要求高和传热量大的载荷或单机的热控<sup>[3]</sup>。近年来, MPTL 技术逐渐在航天领域开始搭载或应用, 已经在相关应用领域取得了较大的进展<sup>[4-7]</sup>。

为了验证 MPTL 的工作特性, 国内外研究者开

展了一系列的研究工作, 主要包括系统级运行特性和部组件与系统之间的相互耦合特性。在系统级工作特性方面, 刘杰<sup>[8]</sup>在实验室条件下对 MPTL 系统的启动特性、稳态和瞬态运行特性进行了研究, 指出 MPTL 技术相较于毛细泵驱动流体回路技术的优势; van Es 等<sup>[5]</sup>对阿拉法磁谱仪探测器 (AMS02) 用 MPTL 产品在轨启动、主备切换、转速调整的工作特性进行了介绍, 结果表明产品在轨运行稳定; 孙西辉<sup>[9]</sup>对 AMS02 用 MPTL 系统的并联蒸发器进行了热源启动、瞬态和稳态运行稳定性试验研究, 指出 MPTL 系统在保证长距离传热和高精度控温的同时, 能通过调节自身参数避免两相流不稳定性; 刘长鑫等<sup>[10]</sup>基于试验对 MPTL 系统的流量漂移现象进行了研究, 表明流量漂移量与系统阻力高度相关; 赵振明等<sup>[11]</sup>通过仿真与试验研究了 MPTL 系统用于 CCD 器件散热时启动特性、并联运行状态及内部工质流动与传热特性, 研究表明, MPTL 系

收稿日期: 2021-05-25; 录用日期: 2022-03-14; 网络出版时间: 2022-03-24 15:48

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20220324.0906.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20220324.0906.001.html)

基金项目: 国家自然科学基金 (51806010)

\* 通信作者. E-mail: [qlmeng@mail.ustc.edu.cn](mailto:qlmeng@mail.ustc.edu.cn)

**引用格式:** 孟庆亮, 赵振明, 陈祥贵, 等. 航天遥感器用泵驱两相流体回路热真空试验研究 [J]. 北京航空航天大学学报, 2023, 49 (3): 559-568.  
MENG Q L, ZHAO Z M, CHEN X G, et al. Thermal vacuum test study of mechanically pumped two-phase loop for space remote sensor [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 559-568 (in Chinese).

统可以很好地满足航天遥感器的散热与控温。

在部组件与系统之间的相互耦合特性方面,研究者主要关注了储液器与主回路之间的相互影响。黄臻成<sup>[12]</sup>研究了MPTL系统受到扰动(温度边界扰动、热源功耗阶梯变化)时储液器和主回路的响应特性,指出二者之间耦合特性对于设计MPTL系统的重要性;袁俊飞等<sup>[13]</sup>对储液器的响应特性与系统的启动和稳定运行之间的响应关系进行了研究,指出蒸发器工作点的控制与储液器内参数密切相关;van Gerner等<sup>[14-15]</sup>利用数值模拟分别对工质为CO<sub>2</sub>和R134a的MPTL系统在热源功率变化时的储液器及其他组件的工作特性进行了研究,指出储液器和回路的参数对系统运行稳定性的影响;孟庆亮等<sup>[16]</sup>通过系统级仿真模型研究了热源功率变化时,两相控温型储液器与系统级主回路之间的动态传热传质特性,指出储液器内气液两相参数、主回路参数会受到热源功率变化的影响。

然而,目前的研究缺少对MPTL系统工作特性在热真空环境下(真空和变化的轨道外热流)的验证与测试,以及对在此环境下热源频繁开关机时系统稳定性的研究。热真空环境下开展试验是航天器热控产品研制过程中重要的测试项目。由于地面上不存在也无法创造热控产品实际运行环境,热控产品的工作性能只能在发射后处于实际的空间环境和工作状态下才能得到真正的检验。热真空环境下的试验是在模拟的空间热环境中使MPTL产品按照运行状态吸收和排散热流,通过测量热控产品的各个参数对其性能进行验证,这个试验是在地面上最接近实际空间热环境条件下进行的,是验证MPTL产品传热性能的最有效手段。基于此,本文搭建了航天遥感器用MPTL试验系统,系统使用了具有被动冷却能力的两相控温型储液器。本文在热真空环境下,对MPTL系统在不同工况下的启动、控温和散热等工作特性进行了研究,并对新型储液器的工作性能进行了验证。

## 1 试验系统

### 1.1 MPTL系统

#### 1.1.1 系统组成

图1给出了航天遥感器用MPTL系统组成原理,主要由机械泵、储液器、换热器、预热器、蒸发器组件、辐射冷凝器及管路组成。MPTL系统内部工质为丙烷,该工质与系统各组件具有良好的相容性,适合长期在轨使用。此外,为监测系统的绝对压力及温度的变化趋势,在储液器附近布置了一套绝对压力传感器,各组件壁面和进出口布置了若干热电偶。图2为不包含辐射冷凝器的系统实物图。

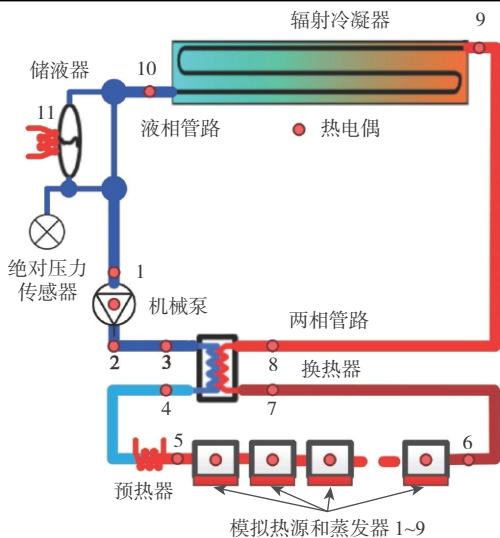


图1 MPTL系统组成

Fig. 1 Schematic diagram of MPTL system

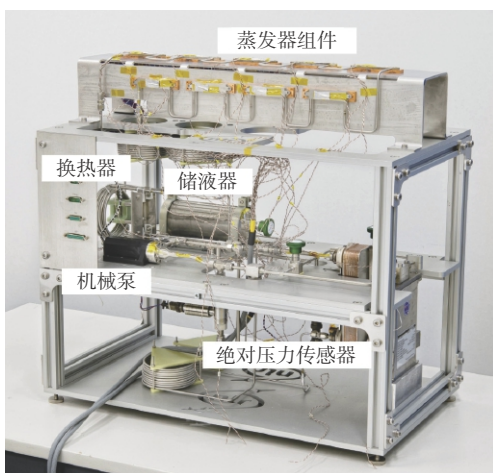


图2 MPTL系统实物图(不含辐射冷凝器)

Fig. 2 Photograph of MPTL system (not including radiator)

#### 1.1.2 系统工作原理

MPTL系统的工作原理可以通过工质在系统中循环的压焓变化过程进行分析。图3给出了系统内工质压焓变化示意图,图中的数字与图1相对应。系统的变化过程包括:①升压过程,机械泵的输入功率部分转化为工质的动能,工质从状态点1变为状态点2;②升温过程,由于管路的阻力,压力下降,工质状态点变为3,经过换热器的加热,变为换热器出口的状态点4,再经过预热器的加热,变为两相饱和态点5;③汽化过程,进入到蒸发器的工质吸收热源的热量,干度不断增大,受管路阻力的影响,工质饱和温度有所下降,蒸发器出口的状态点变为6;④换热过程,由于管路的阻力,变为换热器热端入口的状态点7,再经过换热器冷却,变为换热器热端出口的状态点8;⑤冷凝过程,受两相管路流阻的影响,工质状态点变为冷凝器入口处状态点9,

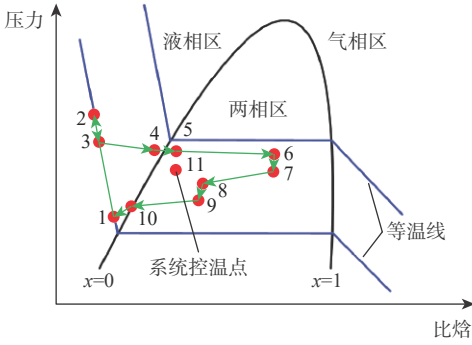


图 3 MPTL 系统内工质压焓变化示意图

Fig. 3 Pressure-enthalpy diagram of fluid in MPTL system

在辐射冷凝器的作用下, 工质变为干度为 0 的饱和态点 10, 再进一步冷却变为过冷态 1, 随后进入机械泵进行继续循环。状态点 11 位于储液器内部, 为系统控温点。

1.1.3 系统主要组件及参数

为保证 MPTL 系统在轨长期稳定的运行, 系统选用的机械泵为北京空间机电研究所自研的微型屏蔽式离心泵, 图 4 给出了屏蔽式离心泵实物图。屏蔽式离心泵只有静密封, 无动密封, 具有耐摩擦、长寿命、低漏率、免维护等特点, 特别适合对介质输送寿命与可靠性要求都很高的航天热控领域。离心泵的流量范围为 0~6 mL/s, 扬程范围为 0~20 m, 驱动功率为 0~15.0 W。此外, 由于离心泵的体积和质量均非常小, 产品工作时微振动的范围仅为 0~60.0 mg, 对遥感器在轨成像影响较小。

图 5 为两相控温型储液器的实物图。其内部为气液两相饱和态工质, 通过控制工质的饱和温度, 即可实现压力的控制。图中: Acc1~Acc3 为温度测点, 分别对应了储液器气相侧、液相侧和出口管路壁面温度。区别于传统储液器的冷却方案, 本文提出了一种具有被动冷却设计在两相控温型储液器, 该储液器通过储液器回流的过冷液与其内部

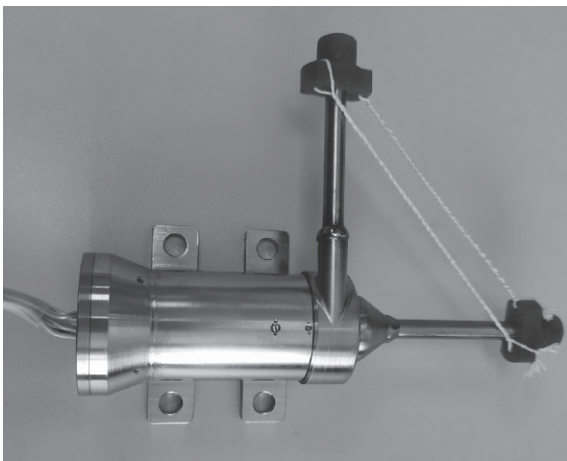


图 4 屏蔽式离心泵实物图

Fig. 4 Photograph of shield centrifugal micropump

工质的温差实现自动冷却。图 6 给出了冷却过程示意图。从辐射冷凝器回流的液体, 在进入机械泵之前被分为 2 条支路, 一路进入到图 5 所示的毛细管中, 通过毛细管与储液器内部流体进行冷量传递, 随后该路流体与另一路汇合共同流入机械泵中。进入储液器中的支路与储液器中的液相换热量的计算表达式为

$$Q_{acc} = hA_{tube}\Delta T \tag{1}$$

式中:  $Q_{acc}$  为换热量;  $h$  为对流换热系数;  $A_{tube}$  为支路管内壁表面积;  $\Delta T$  为工质与支路内壁面的温差。

$h$  的计算公式为

$$h = \frac{Nu \cdot k_f}{d} \tag{2}$$

式中:  $Nu$  为努塞尔数;  $k_f$  为工质的导热系数;  $d$  为支路内径。

进入到支路内工质为纯液相,  $Nu$  的计算公式为

$$Nu = \begin{cases} 4.36 & Re < 1960 \\ \frac{(f/8)(Re - 1000)Pr}{1 + 12.7(f/8)^{1/2}(Pr^{2/3} - 1)} & 1960 \leq Re < 6420 \\ 0.023Re^{0.8}Pr^n & Re \geq 6420 \end{cases} \tag{3}$$

式中:  $Re$  为雷诺数;  $f$  为摩擦系数;  $Pr$  为普朗特数;  $n=0.3$ 。

通过代入式(1)~式(3)中的参数, 得冷量为 5.0~10.0 W, 可以有效保证储液器的控温需求。本文的设计省去了传统冷却使用的半导体制冷装置, 降低了设计的复杂度, 减少了功耗需求, 提升了设计的可靠性。

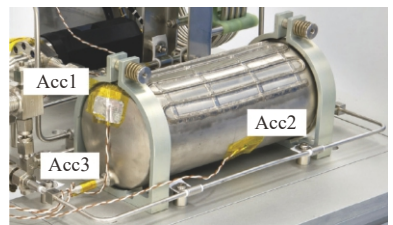


图 5 两相控温型储液器实物图

Fig. 5 Photograph of two-phase thermal-controlled accumulator

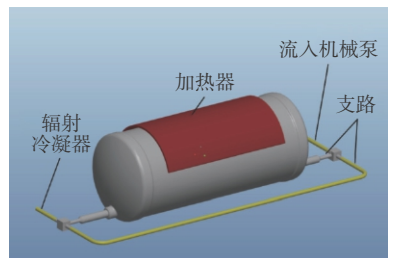


图 6 两相控温型储液器冷却过程示意图

Fig. 6 Schematic diagram of cooling process in two-phase thermal-controlled accumulator

蒸发器通过3D打印技术进行制造(见图7),材质为316L不锈钢,尺寸为90.0 mm×17.5 mm×4.0 mm。管路内部为直径2.9 mm的蛇形流道,用于两相工质换热与流动的空间,如图8所示。相比于传统的机加和焊接装配技术,3D打印技术降低了蒸发器的研制难度,节省了原材料,提高了生产效率。图9给出了蒸发器和模拟热源在空间的布局。模拟热源采用与探测器器件相同热容的紫铜块加工而成,并通过在模拟热源上方加载一定的功率模拟探测器器件发热,蒸发器和模拟热源通过导热胶和螺钉连接。系统包含了9个分布式热源,用于模拟遥感器的核心探测器组件,每个热源的发热量为5.0 W,共45.0 W。为了实现分散式热源的散热与控温,9片蒸发器串联使用。

表1给出了MPTL系统的主要参数。换热器为套管式逆流换热器,管长6.0 m,内部和外部管路分别为3.0 mm和6.0 mm直径的不锈钢管路。辐射冷凝器由带有安装槽的铝合金薄板和3.0 mm直径的不锈钢管路组成,不锈钢管路通过导热胶和卡箍与铝合金板进行导热连接。辐射冷凝器面积为



图7 3D打印蒸发器实物图

Fig. 7 Photograph of evaporator by 3D printing

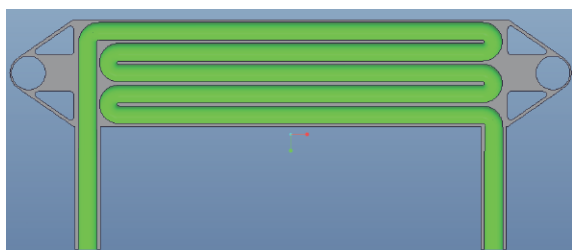


图8 蒸发器内部流道

Fig. 8 Internal flow channel of evaporator

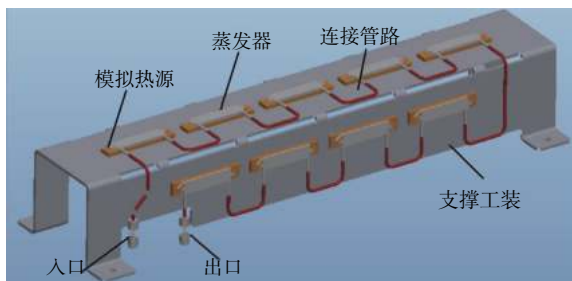


图9 蒸发器及模拟热源空间布置

Fig. 9 Spatial layout of evaporators and simulated heat sources

表1 MPTL系统主要参数

Table 1 Main parameters of MPTL system

参数	数值
两相管路直径/mm	外径4.0, 内径3.0
单相管路直径/mm	外径3.175, 内径2.06
管路总长度/m	20.0
辐射冷凝器	面积0.32 m <sup>2</sup> ; 发射率0.86
储液器体积/L	0.53
冷凝器控温功率/W	20.0
预热器功率/W	40.0
蒸发器功率/W	45.0

0.32 m<sup>2</sup>, 辐射面喷涂发射率为0.86的热控涂层,用于热量的高效排散。

## 1.2 试验参数

### 1.2.1 试验条件和工况

图10给出了MPTL产品在真空罐内放置图。真空罐内压力不高于 $1.3 \times 10^{-3}$  Pa, 热沉温度不高于100 K, 温度分布不均匀性不大于 $\pm 5.0$  K, 热沉表面发射率不小于0.9; 真空罐内表面发射率不低于0.9, 吸收比不小于0.95, 用于实现与MPTL辐射器的高效热交换。除辐射冷凝器外, MPTL系统的其他组件均放置在控温小舱内, 控温小舱各舱板的控温精度为 $\pm 1.0$  K, 用于模拟产品在轨工作环境温度。辐射冷凝器背面包覆20单元多层组件, 用于削弱冷凝器与其他组件真空条件下的辐射换热。为降低各组件的相互干扰, MPTL系统的各组件均包覆多层组件。MPTL系统通过隔热垫放置在真空罐平台上。此外, MPTL产品在轨工作时, 辐射冷凝器除受到真空低温的影响, 还受到轨道外热流的影响。表2给出了外热流变化值, 轨道周期为5 671.0 s, 试验过程中将按照表2周期加载。热真空试验时, 外热流通过在辐射冷凝器背面加载一定的热功率实现。

试验共进行了2个工况, 试验内容如表3所示。

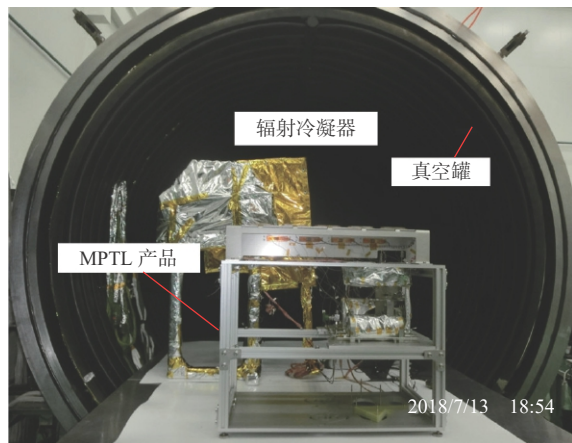


图10 MPTL产品真空罐放置图

Fig. 10 Layout of MPTL in vacuum tank

表 2 外热流随时间周期变化值

Table 2 Temporal variation of external heat flux

加载时间/s	功率/W
0~634	10.8
635~2 079	24.1
2 080~2 646	50.7
2 647~4 218	29.8
4 219~5 671	10.8

表 3 试验工况

Table 3 Test items

工况	系统工作点	试验内容
1	20 ℃	测试流体回路启动特性, 储液器控温稳定性及蒸发器温度稳定性
2	10 ℃	测试储液器升降温速率, 系统运行的稳定性

1.2.2 微重力影响分析

MPTL 系统的储液器和主回路内存在气液两相流体, 由于气液介质存在了较大的密度差, 重力效应可能会影响系统的流动和传热特性, 微重力下系统会呈现出与地面条件不同的工装特性。储液器可以通过内部设置的毛细结构有效实现流体的管控和温度控制<sup>[17]</sup>。杜王芳等<sup>[18]</sup>详细评述了关于重力无关性的研究进展, 对于分析多相热流体系统重力无关性具有重要的指导意义。本文主要根据文献<sup>[18]</sup>介绍的相关准则, 对 MPTL 系统中的流动和传热特性进行重力效应影响分析。

为研究重力效应对流动的影响, 分别计算 Bond 数和气相表观 Froude 数:

$$Bo = \frac{(\rho_L - \rho_G)gd^2}{\sigma} \quad (4)$$

$$Fr_{SG} = \frac{U_{SG}}{\sqrt{(\rho_L - \rho_G)gd/\rho_G}} \quad (5)$$

式中:  $\rho$ 、 $g$ 、 $d$ 、 $\sigma$  和  $U$  分别为密度、重力加速度、管路内径、表面张力系数和特征速度;  $U_{SG}$  为气相表观速度; 下标 L 和 G 分别表示液相和气相。

将系统的参数代入到式 (4) 和式 (5) 中, 可以得到  $Bo = 2.64$ ,  $Fr_{SG} = 2.63 \sim 4.6$ , 2 个数值满足式 (6) 和式 (7) 所述的主导作用力的重力无关性判据:

$$Bo \leq Bo_{cr} \approx 1.5 \sim 6 \quad (6)$$

$$Fr_{SG} \geq Fr_{SG,cr} \approx 0.54 \sim 2.2 \quad (7)$$

重力效应对系统沸腾换热的影响主要包括: ①垂直于预热器和蒸发器表面的重力分量引起的动力学不稳定性; ②波长过大阻止液体与预热器和蒸发器内壁面的接触。

对于影响①, 当界面波长满足式 (8) 所述的准则时, 则说明该特性与重力是无关的, 即

$$\left| \frac{(\rho_L - \rho_G)(\rho_L + \rho_G)\sigma g \cos \theta}{\rho_L^2 \rho_G^2 (U_G - U_L)^4} \right| \ll \frac{1}{4} \quad (8)$$

式中:  $\theta$  为热源和流动方向的夹角, 本文为 0。将相应数值代入到式 (8) 中, 得到式 (8) 左边项数值在 0.000 5~0.018 之间, 因此式 (8) 成立。

对于影响②, 需要满足:

$$We = \frac{\rho_L \rho_G (U_G - U_L)^2 L}{\sigma (\rho_L + \rho_G)} \geq 2\pi \quad (9)$$

式中:  $L$  为蒸发器内管路长度。通过代入数值到式 (9) 中, 得到左边项的数值在 1 940~10 215 之间, 式 (9) 成立。

因此, 本文搭建系统的流动和传热特性均满足重力无关性准则, 系统的设计参数可以满足航天遥感器在轨应用需求。

2 试验结果分析

2.1 工况 1 结果分析

图 11 为 MPTL 首次启动时储液器上测点变化曲线。试验前, MPTL 各组件温度长时间放置于真空低温环境下, 各组件的温度在 -5.0~0 ℃ 之间。试验开始后, 将储液器气相侧壁面的温度提升到 20.0 ℃, 储液器的升温速率为 1.2 ℃/min, 与此同时, 液相侧壁面和出口的温度分别上升至 4.9 ℃ 和 0.1 ℃。5 586 s 后, 机械泵开机, 在机械泵的作用下, 辐射冷凝器中过冷液体开始循环, 使得储液器出口温度开始降低。6 130 s 后, 预热器开启。9 646 s 后, 将辐射冷凝器的控温点从 -30.0 ℃ 调高至 -25.0 ℃, 随后储液器出口温度升高至 -2.0 ℃。10 591 s 后, 储液器出口温度开始出现振荡, 波动范围在 -6.0~-2.1 ℃ 之间, 与此同时, 储液器气相侧壁和蒸发器各处的温度也出现了小幅温度振荡, 温度振荡现象的出现意味着 MPTL 系统在预热器下游达到了两相状态, 两相态的出现引起了储液器与主回路工质的交换, 导致温

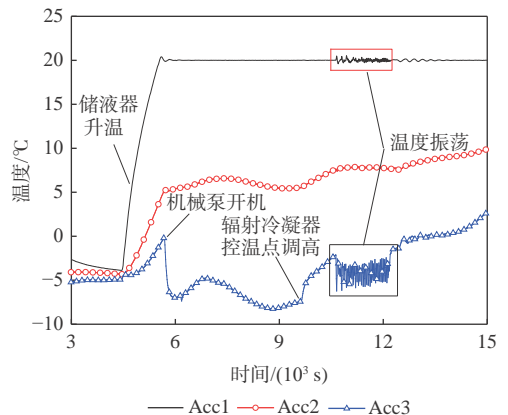


图 11 MPTL 启动时储液器上测点温度变化曲线  
Fig. 11 Temperature variation in measuring points of accumulator during MPTL start up

度振荡现象。温度振荡过程持续约 1 600 s, 随后储液器气相侧壁面温度稳定在 20 °C 附近。

图 12 和图 13 为系统绝对压力和根据压力拟合得到温度值的变化曲线。系统绝对压力的变化范围为 0.42~0.85 MPa, 压力增加、振荡的时刻与储液器气相侧壁面温度发生相应变化过程的时刻相同, 意味着对 MPTL 系统进行的操作同步影响了系统的压力和储液器内气相的温度。饱和温度压力和数据来源于 NIST 标准数据库<sup>[9]</sup>。饱和温度和压力的拟合公式为

$$T_{\text{sat}} = -30.1P_{\text{sat}}^2 + 94.4P_{\text{sat}} - 38.0 \quad (10)$$

式中:  $T_{\text{sat}}$  为饱和温度;  $P_{\text{sat}}$  为饱和压力。

通过将压力数据代入到式 (10) 中得到的温度 (计算温度) 变化数据, 并将该温度与实测温度 (储液器气相侧壁面温度) 对比, 发现变化趋势一致性较好, 表明储液器内饱和温度和压力是由其内部的气相所决定的。受加热功率和热阻的影响, 两者的温差在 0.2~0.3 °C 之间。图 12 和图 13 中温度和压力的振荡现象在 2.3 节中进一步解释。

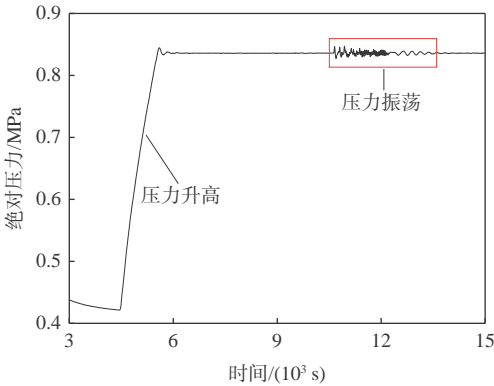


图 12 MPTL 启动时系统绝对压力变化曲线

Fig. 12 Variations in absolute system pressure during MPTL start up

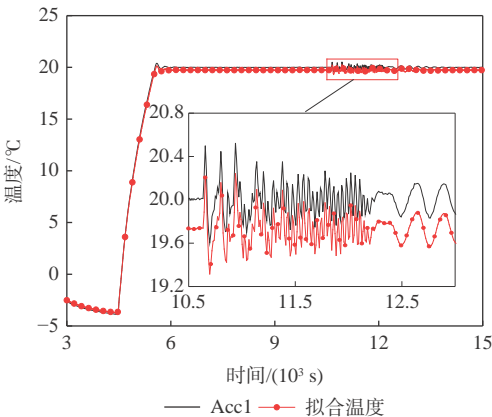


图 13 MPTL 启动时储液器气相温度与压力拟合温度对比

Fig. 13 Comparison between vapor temperature and fitting temperature from saturated pressure during MPTL start up

图 14 和图 15 为模拟热源开关机时蒸发器和换热器测点温度的变化曲线。从图 14 中可以看出, 模拟热源关机时, 9 个蒸发器的温度在 21.5~22.5 °C 之间, 模拟热源开机后, 蒸发器的温度升高, 温度在 22.0~23.0 °C 之间。蒸发器温度升高主要受到热阻的影响。模拟热源开关机过程中, 储液器气相温度保持稳定, 从而维持了蒸发器温度的稳定。此外, 蒸发器的温度比储液器气相的温度高 1.5~3 °C, 主要是受储液器与蒸发器的压差影响。从图 15 可以看出, 通过在 MPTL 中设置换热器可以将冷端入口的温度提高 7~12 °C。此外, 换热器冷端入口处的温度受外热流影响而发生波动, 但蒸发器温度并未发生波动, 这是由于蒸发器内工质为两相态, 其温度主要是由储液器所决定的, 储液器温度保持稳定, 蒸发器温度即可保持稳定, 储液器为模拟热源的工作提供了稳定的边界条件。

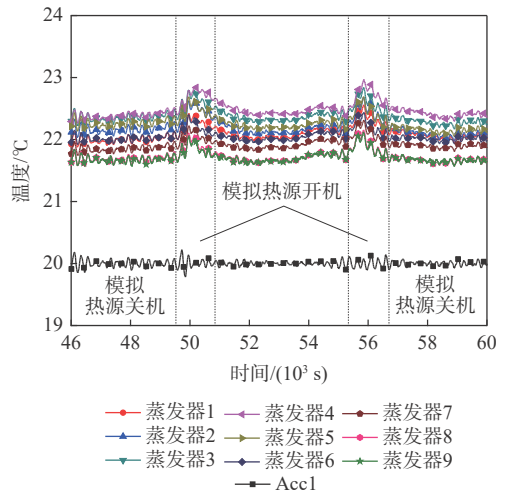


图 14 模拟热源开关机时蒸发器测点温度变化曲线

Fig. 14 Temperature variations in measuring points of evaporators during heat loads start up and power off

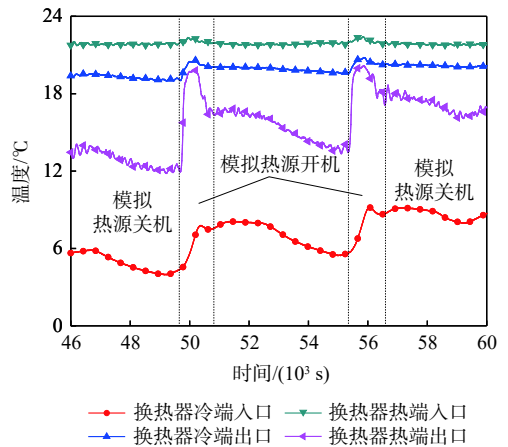


图 15 模拟热源开关机时换热器测点温度变化曲线

Fig. 15 Temperature variations in measure points of heat-exchanger during heat loads start up and power off

### 2.2 工况2 结果分析

图 16 和图 17 为工况 2 储液器和蒸发器测点温度变化曲线。从图 16 可以看出,系统达到两相态后,储液器气相壁面测点温度均出现小幅振荡,储液器出口管路壁面温度快速下降,意味着 MPTL 主回路的冷工质进入了储液器中。图 17 中,储液器及主回路温度在 14~16 °C 之间,首先将控温点提升至 20 °C,随后启动机械泵,再开启预热器,系统到达两相态后再将储液器降温,通过内部毛细管路与储液器内工质的换热作用,控温点降温至 10 °C 附近,并保持稳定。蒸发器与储液器温度变化保持一致,蒸发器温度维持在 11.3~12.7 °C。

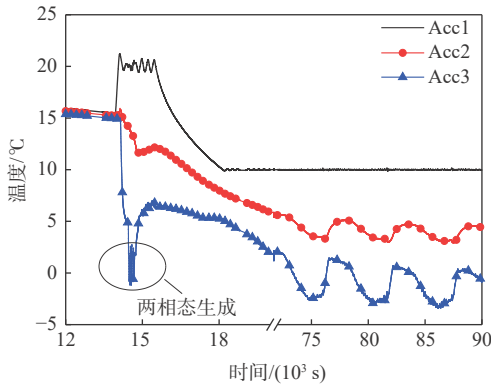


图 16 工况 2 储液器测点温度变化曲线

Fig. 16 Temperature variations in measuring points of accumulator for test condition 2

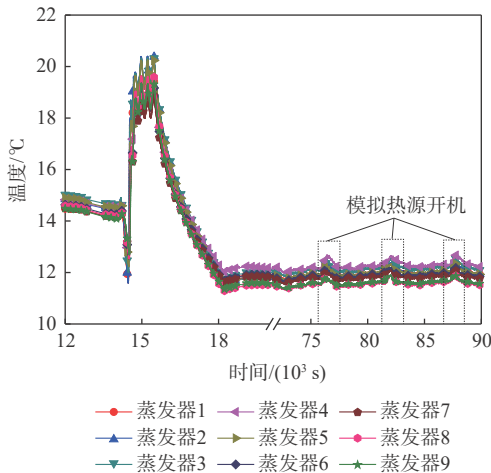


图 17 工况 2 蒸发器测点温度变化曲线

Fig. 17 Temperature variations in measuring points of evaporators for test condition 2

### 2.3 系统振荡过程分析

图 18~图 20 为工况 1 中两相段生成后,储液器测点温度、系统压力和储液器控温功率的变化曲线。从图 18 中可以看出,储液器气相侧壁面温度波动过程与储液器出口温度的变化趋势相反,即储液器气相侧壁面温度升高时,出口温度下降。这是

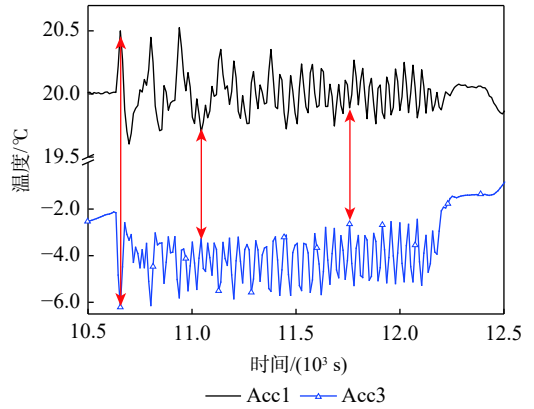


图 18 两相段生成后储液器测点温度变化趋势

Fig. 18 Temperature variations in measuring points of accumulator after two phase generation

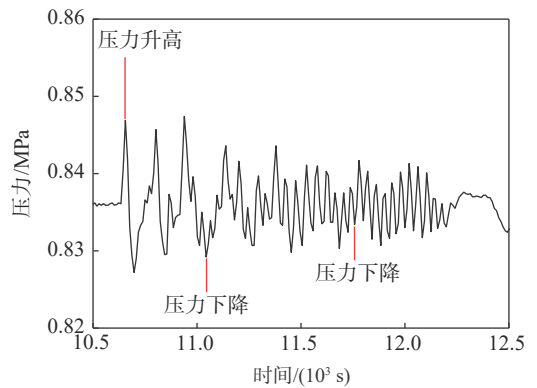


图 19 两相段生成后系统压力变化趋势

Fig. 19 Pressure variations of system after two phase generation

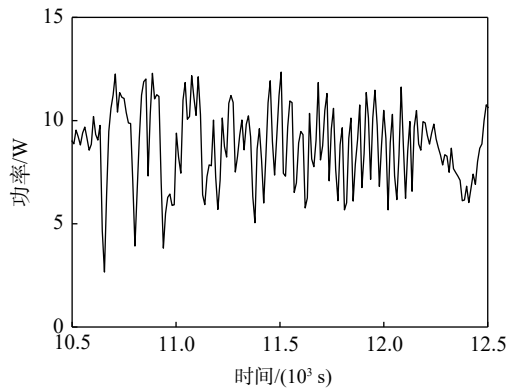


图 20 两相段生成后储液器功率变化趋势

Fig. 20 Variations in power applied on accumulator after two phase generation

由于两相段生成后, MPTL 主回路内多余的工质被推入到了储液器中,储液器出口温度下降,气相空间被快速压缩,气相温度和压力随之升高;反之,储液器内工质流入到主回路中时,储液器出口温度升高,气相空间膨胀,气相温度和压力下降。为了维持储液器的控温点,储液器加载的功率也会因此发生变化(见图 20),加载功率变化导致了 MPTL 主

回路与储液器工质交换过程(即温度和压力振荡过程)维持了一段时间。从图18~图20可以发现,振荡幅度逐渐减小,并在约12200s后,振荡过程停止,MPTL系统控温点趋于稳定。

图21给出了MPTL系统压力降脉动简化模型。单相条件下,主回路与储液器无工质交换过程,储液器前流量 $\dot{m}_1$ 和离心泵后流量 $\dot{m}_2$ 满足:

$$\dot{m}_1 = \dot{m}_2 \quad (11)$$

MPTL系统特性曲线与机械泵特性曲线的交汇点为 $P$ 。系统压降 $\Delta P$ 和系统流量的关系式为

$$\Delta P = K_f \dot{m}_2^2 \quad (12)$$

式中: $K_f$ 为阻力系数。当预热器生成两相后,系统主回路中的气相体积逐渐增加,导致 $\dot{m}_1$ 增加,且 $\dot{m}_1 > \dot{m}_2$ ,回路内多余的流量 $(\dot{m}_1 - \dot{m}_2)$ 将流入储液器中。由于两相段的增加,式(12)中阻力系数将增加, $\dot{m}_2$ 将减少,系统主回路的工作点将沿着MPTL系统特性曲线移动至压降最大值 $B$ 点。如果压降增加,工作点将漂移到 $C$ 点,即通过提供更大的流量来维持压力的增加。而在 $C$ 处,由于进入预热器的流量 $\dot{m}_2$ 突然增大,超过了系统流量 $\dot{m}_1$ ,两相段减少,储液器内多余的流量 $(\dot{m}_2 - \dot{m}_1)$ 将流入主回路中,系统工作点将再次移动至最小值 $D$ 。如果压降减少,工作点将漂移到 $A$ 点。由于 $\dot{m}_2$ 比 $\dot{m}_1$ 小,系统工作点将沿正斜率单相区增大至 $B$ 点,开始新一轮的循环。因此,对负斜率两相区内部任何点的扰动都会产生上述流量在 $ABCD$ A区域内循环。当储液器与主回路完成工质交换过程后,系统趋于稳定,振荡停止。

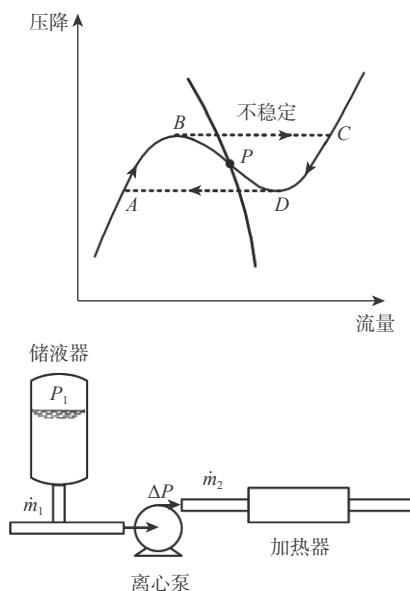


图21 MPTL系统压力降脉动简化模型

Fig. 21 Schematic diagram of pressure-drop oscillations for MPTL system

### 3 结论

本文在真空条件下对航天遥感器用MPTL系统进行了试验测试,主要考察了高真空与空间外热流条件下系统的散热与控温能力。MPTL系统使用了具有被动冷却能力的两相控温型储液器,无需额外的制冷功率即可实现储液器的精确控温。通过温度和压力等数据,研究了热真空条件下MPTL系统在不同工况时的工作性能及储液器与主回路之间的传热传质过程。主要结论如下:

1) 通过对不同系统控温点的测试,验证了MPTL系统可以根据需要进行控温点的调整,外热流变化与热源开关均对蒸发器温度影响较小,系统在这2种工况下均能维持稳定。MPTL系统灵活性较高,可以满足在轨不同工作模式下的使用需求。

2) 通过回流过冷液与储液器中液相冷量的交换,保证了储液器冷量的供应,实现了被动冷却能力。储液器内气相温度与系统绝对压力变化趋势保持较高的一致性,表明两相控温型储液器内的饱和状态点是由其内部的气相所决定的。

3) MPTL系统主回路的相态由液相到气液两相态转变时,系统出现了不稳定现象,主要体现在储液器内压力和温度、蒸发器温度出现振荡现象。MPTL系统出现的不稳定特性为压力降脉动,其主要是由储液器与主回路进行工质交换所引起的。

通过本文的研究表明,MPTL系统具有控温点可调、稳定性好、自适应能力强等特点,可以满足航天遥感器高稳定度的控温需求,进而可以保证遥感器的高精度成像。

### 参考文献 (References)

- [1] GREGORY T H, ABEL J, MANDI J. Thermal design considerations of the hubble space telescope (HST) science instrument control and data handler (SI C&DH-2): AIAA 2010-6055[R]. Reston: AIAA, 2010.
- [2] 李春林. 空间光学遥感器热控技术研究[J]. 宇航学报, 2014, 35(8): 863-870.  
LI C L. Research on space optical remote sensor thermal control technique[J]. Journal of Astronautics, 2014, 35(8): 863-870(in Chinese).
- [3] VAN ES J, VAN GERNER H J, VAN BENTHEM R C. Component developments in Europe for mechanically pumped loop systems (MPLs) for cooling applications in space[C]//46th International Conference on Environmental Systems, 2016: ICES-2016-196.
- [4] VAN ES J, PAUW A, VAN DONK G, et al. AMS02 tracker thermal control cooling system test results of the AMS02 thermal vacuum test in the ISS at ESA ESTEC: AIAA-2012-3577[R]. Reston: AIAA, 2012.
- [5] VAN ES J, PAUW A, VAN GERNER H J, et al. AMS02 tracker thermal control cooling system commissioning and operational res-

- ults: AIAA-2013-3389[R]. Reston: AIAA, 2013.
- [6] 于新刚,徐侃,苗建印,等.高热流散热泵驱两相流体回路设计及飞行验证[J].宇航学报,2017,38(2):192-197.  
YU X G, XU K, MIAO J Y, et al. Design and on-board validation of pumped two-phase fluid loop for high heat flux removal[J]. *Journal of Astronautics*, 2017, 38(2): 192-197(in Chinese).
- [7] 王镇锐,张兴斌,温世喆,等.结合TEC的泵驱两相温控系统的空间应用[J].宇航学报,2018,39(10):1176-1184.  
WANG Z R, ZHANG X B, WEN S Z, et al. Space applications of pumped two-phase temperature control system combined with TEC [J]. *Journal of Astronautics*, 2018, 39(10): 1176-1184(in Chinese).
- [8] 刘杰.航天机械泵驱动两相冷却环路循环特性的研究[D].上海:上海交通大学,2008.  
LIU J. Investigations on running characteristics of the mechanically pumped cooling loop for space applications[D]. Shanghai: Shanghai Jiaotong University, 2008(in Chinese).
- [9] 孙西辉.机械泵驱动CO<sub>2</sub>两相流体回路稳定性研究[D].广州:中山大学,2010.  
SUN X H. Stabilities research of the mechanically pumped two-phase CO<sub>2</sub> loop[D]. Guangzhou: Sun Yat-sen University, 2010(in Chinese).
- [10] 刘长鑫,张济民,徐涛,等.泵驱动两相流体回路流量漂移现象的试验研究[J].上海航天,2017,34(4):125-132.  
LIU C X, ZHANG J M, XU T, et al. Experimental research on flow excursion of mechanically pumped two-phase loop[J]. *Aerospace Shanghai*, 2017, 34(4): 125-132(in Chinese).
- [11] 赵振明,孟庆亮,张焕冬,等.CCD器件用机械泵驱动两相流体回路仿真与试验[J].北京航空航天大学学报,2019,45(5):893-901.  
ZHAO Z M, MENG Q L, ZHANG H D, et al. Simulation and experimental study of mechanically pumped two-phase loop for CCD[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2019, 45(5): 893-901(in Chinese).
- [12] 黄臻成.航天热控用机械驱动两相系统的数值模拟及控温特性分析[D].广州:中山大学,2008.  
HUANG Z C. Numerical simulation and thermal control characteristic of the space mechanically pumped two-phase CO<sub>2</sub> loop[D]. Guangzhou: Sun Yat-sen University, 2008(in Chinese).
- [13] 袁俊飞,唐大伟,曹宏章.泵驱动相变冷却系统中储液器压力响应特性[J].航空动力学报,2015,30(10):2384-2390.  
YUAN J F, TANG D W, CAO H Z. Response characteristics of accumulator pressure in pumped phase-change cooling system[J]. *Journal of Aerospace Power*, 2015, 30(10): 2384-2390(in Chinese).
- [14] VAN GERNER H J, BRAAKSMA N. Transient modelling of pumped two-phase cooling systems: comparison between experiment and simulation[C]//46th International Conference on Environmental Systems, 2016: ICES-2016-004.
- [15] VAN GERNER H J, BOLDER R, VAN ES J. Transient modelling of pumped two-phase cooling systems: comparison between experiment and simulation with R134a[C]//47th International Conference on Environmental Systems, 2017: ICES-2017-037.
- [16] 孟庆亮,张焕冬,赵振明.两相控温型储液器进出流量的瞬态数值模拟[J].北京航空航天大学学报,2019,45(11):2160-2169.  
MENG Q L, ZHANG H D, ZHAO Z M. Transient numerical simulations of flow rate into and out of two-phase temperature control accumulator[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2019, 45(11): 2160-2169(in Chinese).
- [17] 孟庆亮,周振华,赵振明.用于微重力下两相控温型储液器性能研究[J].工程热物理学报,2021,42(7):1770-1776.  
MENG Q L, ZHOU Z H, ZHAO Z M. Performance study of two-phase thermal-controlled accumulator for microgravity environment[J]. *Journal of Engineering Thermophysics*, 2021, 42(7): 1770-1776(in Chinese).
- [18] 杜王芳,乐述文,赵建福,等.多相热流体系统中的重力无关性准则[J].河北水利电力学院学报,2019,29(2):1-7.  
DU W F, YUE S W, ZHAO J F, et al. Criteria of gravity independence in multiphase thermal fluid system[J]. *Journal of Hebei University of Water Resources and Electric Engineering*, 2019, 29(2): 1-7(in Chinese).
- [19] LEMMON E W, HUBER M L, MCLINDEN M O. NIST standard reference database 23. NIST reference fluid thermodynamic and transport properties database (REFPROP). version 8.0. Standard reference data[Z]. Gaithersburg: NIST, 2007.

# Thermal vacuum test study of mechanically pumped two-phase loop for space remote sensor

MENG Qingliang<sup>1,2,\*</sup>, ZHAO Zhenming<sup>1,2</sup>, CHEN Xianggui<sup>1,2</sup>, ZHU Xu<sup>1,2</sup>

(1. Beijing Institute of Space Mechanics & Electricity, Beijing 100094, China;

2. Beijing Key Laboratory of Advanced Optical Remote Sensing Technology, Beijing 100094, China)

**Abstract:** In this paper, a test setup of mechanically pumped two-phase loop (MPTL) was constructed in response to the demand for high precision and high stability temperature control of the core components in space remote sensors. In this setup, a two-phase thermal-controlled accumulator with passive cooling was adopted. For the purpose of verifying the working performance of the MPTL system under the conditions of high vacuum, ultra-low temperature and varied external heat flux, the heat dissipation and temperature control ability of the MPTL system under different test conditions were tested in the vacuum chamber. Then the obtained test data, including temperature and pressure, were used to analyze the operating characteristics of the main loop, the thermodynamic behaviors in the accumulator and the heat and mass transfer between the main loop and the accumulator. The test results showed that the temperature control points of the MPTL system could be quickly adjusted by the accumulator, and the external heat flux and the action of starting up and powering off of the heat source had little influence on the temperature of the evaporators. The cooling capacity was provided by the temperature difference between the subcooling liquid entering the capillary tube and the liquid phase in the accumulator. During the phase transition in the main loop, the mass transfer behaviors between accumulator and the main loop gave rise to the pressure-drop oscillations.

**Keywords:** mechanically pumped two-phase loop; two-phase thermal-controlled accumulator; thermal control products; thermal vacuum test; space remote sensor

**Received:** 2021-05-25; **Accepted:** 2022-03-14; **Published Online:** 2022-03-24 15:48

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20220324.0906.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20220324.0906.001.html)

**Foundation item:** National Natural Science Foundation of China (51806010)

\* **Corresponding author.** E-mail: [qlmeng@mail.ustc.edu.cn](mailto:qlmeng@mail.ustc.edu.cn)

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0313

# 基于修正似然滤波的无人机编队相对导航方法

苏炳志<sup>1,\*</sup>, 王磊<sup>2</sup>, 张红伟<sup>1</sup>, 汪海涵<sup>1</sup>, 石璐璐<sup>1</sup>

(1. 中国直升机设计研究所, 天津 300300; 2. 陆军装备部驻天津地区航空军代室, 天津 300384)

**摘 要:** 针对无人机编队相对导航系统中视觉导航传感器量测数据存在随机时延问题, 提出一种能够处理多步随机延迟量测的修正似然容积卡尔曼滤波 (ML-CKF) 算法。用多个伯努利随机变量对量测模型进行修正以描述随机延迟; 通过边缘化延迟变量来计算滤波的似然函数以从延迟量测中提取准确的信息; 采用三阶球面-径向容积准则计算高斯加权积分以解决系统的非线性。滤波中的加权因子根据接收量测的特性进行调整, 因此, 所提修正似然滤波具有自适应卡尔曼滤波属性。利用罗德里格斯参数表示姿态误差, 设计了基于修正似然容积卡尔曼滤波的相对导航滤波器。仿真结果表明: 所提算法可以准确地估计出长机和僚机之间的相对位置、速度和姿态, 且估计精度高于容积卡尔曼滤波和传统随机时延滤波。

**关键词:** 无人机编队; 相对导航; 似然函数; 容积卡尔曼滤波; 随机延迟量测

**中图分类号:** V249.3

**文献标志码:** A

**文章编号:** 1001-5965(2023)03-0569-11

无人机采用编队形式开展协同侦查、联合攻击和战场毁伤评估等任务, 可以有效提高作战效率, 是未来体系化作战的主要模式<sup>[1]</sup>。高精度、高可靠的相对导航系统是无人机编队协同完成既定任务的关键技术。基于 INS/GPS 的相对导航方式是初期最常用的一种导航机制, 然而从外界接收到的 GPS 信号存在极易受到故意干扰和电子欺骗的缺陷<sup>[2-3]</sup>。因此, 学者们研究了一种不依赖于外部传感器的相对导航系统——视觉辅助惯性的相对导航<sup>[4]</sup>。基于 INS/Vision 的相对导航精度很大程度上取决于视觉传感器和惯性传感器数据融合方法。

由于形式简单和易于实现, 扩展卡尔曼滤波 (extended Kalman filtering, EKF) 是相对导航滤波器中最常用的一种滤波算法<sup>[5]</sup>。由于只利用一阶泰勒级数对非线性系统进行线性化, EKF 在强非线性系统中表现出不令人满意的性能甚至使滤波器发散<sup>[6]</sup>。学者们使用采样方法近似非线性分布解决非

线性问题的思维, 提出一系列确定性采样非线性高斯滤波。文献 [7-8] 从直观的统计信息转换视角开发了基于无迹变换的卡尔曼滤波 (unscented Kalman filtering, UKF)。然而, 在处理高维系统时, UKF 中存在负权重的 Sigma 点, 致使状态估计协方差出现非正定, 这将影响滤波器的稳定性。Arasaratnam 等将笛卡儿坐标系下的积分转换成球面和径向积分推导了基于三阶球面-径向容积准则的容积卡尔曼滤波 (cubature Kalman filtering, CKF)<sup>[9-10]</sup>。与 UKF 不同, CKF 的权重全为正值, 因此, 其具有更好的数值稳定性。

文献 [5-10] 中各种滤波算法都是基于量测是实时可用的假设开发得到的。但实际上, 由于通讯或网络堵塞等因素的存在, 到达滤波数据处理中心的量测数据往往存在随机时延<sup>[11]</sup>。为了处理量测随机时延问题, 学者们基于向后逆推和向前正推提出了无序 Sigma 点卡尔曼滤波<sup>[12]</sup> 和乱序粒子滤波<sup>[13]</sup>。

收稿日期: 2021-06-08; 录用日期: 2021-09-08; 网络出版时间: 2021-09-16 17:02

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210916.1133.004.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210916.1133.004.html)

\* 通信作者. E-mail: [subingzhi\\_hit@163.com](mailto:subingzhi_hit@163.com)

**引用格式:** 苏炳志, 王磊, 张红伟, 等. 基于修正似然滤波的无人机编队相对导航方法 [J]. 北京航空航天大学学报, 2023, 49 (3): 569-579.  
SU B Z, WANG L, ZHANG H W, et al. Relative navigation method based on modified likelihood filtering for unmanned aerial vehicle formation [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 569-579 (in Chinese).

尽管这些滤波算法能够很好地处理随机延迟量测,但需要时间戳来计算延迟时间,并且只能处理延迟时间不变的量测数据。为了摆脱这种限制和克服不足,学者们从延迟概率的角度出发,利用伯努利随机变量建立实际接收量测与理想量测之间的关系,对随机时延系统重新建模,提出了多种滤波体制以解决量测数据具有一步或多步延迟问题。Hermoso-Carazo 和 Linares-Perez<sup>[14-15]</sup>针对一步随机延迟量测提出改进型的 EKF 和 UKF,通过对状态和量测噪声的扩增,将其推广到多步随机时延系统中。Wang 等<sup>[16]</sup>概括性地推导了具有一般普遍性的单步随机时延非线性高斯滤波算法,并给出了单步随机时延容积卡尔曼滤波 (one-step randomly delayed CKF, ORD-CKF) 算法的详细过程。张勇刚等<sup>[17]</sup>提出一种带多步随机延迟量测高斯滤波器的一般框架,并给出多步随机时延容积卡尔曼滤波 (multiple-step randomly delayed CKF, MRD-CKF) 算法的具体计算流程。Esmzad 和 Esfanjani<sup>[18-19]</sup>通过修正常规卡尔曼滤波的似然函数,提出一种能够处理多步随机时延的修正似然卡尔曼滤波,该算法利用量测残差调节权重因子,使得滤波算法具有更高的估计精度;之后将其推广到非线性高斯系统中。

在文献 [18-19] 的研究基础上,采用估计精度和稳定性较好的三阶球面-径向容积准则计算非线性系统中的高斯加权积分,提出一种修正似然容积卡尔曼滤波 (modified likelihood CKF, ML-CKF) 算法,解决无人机编队相对导航系统中视觉量测存在多步随机时延问题。所提算法的递归过程包括状态预测和量测更新 2 个步骤。状态预测是基于三阶球面-径向容积准则和多步随机时延系统得到的。由于存在时延,接收到的量测与系统过去状态相关,在量测更新中通过边缘化延迟变量来计算滤波的似然函数以从延迟量测中准确地提取信息。建立无人机编队相对导航系统模型;采用一种无约束的三参数罗德里格斯参数表示姿态误差来传播和更新姿态四元数,设计基于 ML-CKF 的相对导航滤波器。最后,开展仿真验证,对比了本文所提算法 ML-CKF 与 CKF、ORD-CKF 和 MRD-CKF 的性能。

## 1 修正似然容积卡尔曼滤波

### 1.1 问题阐述及预处理

考虑式 (1) 所示的离散系统:

$$\begin{cases} \mathbf{x}_k = f(\mathbf{x}_{k-1}) + \mathbf{w}_{k-1} \\ \mathbf{z}_k = h(\mathbf{x}_k) + \mathbf{v}_k \end{cases} \quad (1)$$

式中:  $\mathbf{x}_k \in \mathbf{R}^n$  为动力学系统在  $k$  时刻的状态向量;  $\mathbf{z}_k \in \mathbf{R}^m$  为  $k$  时刻的量测向量;  $f(\cdot)$  和  $h(\cdot)$  分别为非线性系统状态转移方程和量测方程;  $\mathbf{w}_{k-1} \in \mathbf{R}^n$  和  $\mathbf{v}_k \in \mathbf{R}^m$  分别为满足  $E[\mathbf{w}_{k-1}\mathbf{w}_{j-1}^T] = \mathbf{Q}_{k-1}\delta_{k-1-j-1}$  和  $E[\mathbf{v}_k\mathbf{v}_j^T] = \mathbf{R}_k\delta_{kj}$  的零均值高斯白噪声,  $\mathbf{Q}_{k-1}$  为过程噪声协方差,  $\mathbf{R}_k$  为量测噪声协方差。

实际中,由于网络拥挤或通讯堵塞等原因,量测往往存在时延。假设最大延迟步数为  $d$ ,那么  $k$  时刻滤波器实际接收到的量测量可能是理论量测序列  $\{\mathbf{z}_{k-s}\}_{s=0}^d$  中的任意一个。因此,  $d$  步时延系统的量测方程可表述为

$$\begin{aligned} \mathbf{y}_k = & (1 - \tau_1)\mathbf{z}_k + \tau_1(1 - \tau_2)\mathbf{z}_{k-1} + \tau_1\tau_2(1 - \tau_3)\mathbf{z}_{k-2} + \cdots + \\ & \left( \prod_{s=1}^{d-1} \tau_s \right) (1 - \tau_d)\mathbf{z}_{k-d+1} + \left[ 1 - (1 - \tau_1) - \right. \\ & \left. \tau_1(1 - \tau_2) - \cdots - \left( \prod_{s=1}^{d-1} \tau_s \right) (1 - \tau_d) \right] \mathbf{z}_{k-d} = \\ & (1 - \tau_1)\mathbf{z}_k + \tau_1(1 - \tau_2)\mathbf{z}_{k-1} + \tau_1\tau_2(1 - \tau_3)\mathbf{z}_{k-2} + \cdots + \\ & \left( \prod_{s=0}^{d-1} \tau_s \right) (1 - \tau_d)\mathbf{z}_{k-d+1} + \sum_{s=0}^d \tau_s \mathbf{z}_{k-d} = \sum_{s=0}^d \tau^{(s,j)} \mathbf{z}_{k-s} \end{aligned} \quad (2)$$

式中:  $\mathbf{y}_k$  为实际接收到的量测量;  $\tau_0 = 1$ ,  $\tau_s (1 \leq s \leq d)$  为相互独立的伯努利随机变量,  $\tau^{(s,j)}$  为

$$\tau^{(s,j)} = \begin{cases} \left( \prod_{j=0}^s \tau_j \right) (1 - \tau_{s+1}) & 0 \leq s \leq d-1 \\ \prod_{j=0}^d \tau_j & s = d \end{cases} \quad (3)$$

式中:  $\tau^{(s,j)}$  ( $s = 0, 1, \dots, d$ ) 中只有一个系数取 1, 其他均为 0; 这意味着每个时刻接收到的实际量测  $\mathbf{y}_k$  是当前时刻或者过去时刻量测量  $\mathbf{z}_{k-s}$  ( $s = 0, 1, \dots, d$ ) 中的一个。

量测延迟  $s$  ( $0 \leq s \leq d$ ) 步的概率为

$$p^{(s,j)} = \begin{cases} \left( \prod_{j=0}^s p_j \right) (1 - p_{s+1}) & 0 \leq s \leq d-1 \\ \prod_{j=0}^d p_j & s = d \end{cases} \quad (4)$$

式中:  $p_j$  为伯努利随机变量  $\tau_j$  ( $0 \leq j \leq d$ ) 的期望, 即  $\tau_j$  取 1 的概率。

### 1.2 修正似然非线性高斯滤波框架

由式 (2) 可知滤波器实际接收到的量测量  $\mathbf{y}_k$  是  $\{\mathbf{z}_{k-s}\}_{s=0}^d$  中的一个, 即  $\mathbf{y}_k$  的值与  $\mathbf{x}_k, \mathbf{x}_{k-1}, \dots, \mathbf{x}_{k-d}$  相关。因此, 需要将  $\mathbf{x}_{k-1}, \mathbf{x}_{k-2}, \dots, \mathbf{x}_{k-d}$  扩增到当前状态当中, 得到扩增状态:

$$\mathbf{X}_k = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{x}_{k-1} \\ \vdots \\ \mathbf{x}_{k-d} \end{bmatrix}_{n_a \times 1} \quad (5)$$

式中:  $n_a = (d+1)n$  为扩增状态向量的维数。

相应的扩增系统为

$$\begin{cases} \mathbf{X}_k = \mathbf{F}(\mathbf{X}_{k-1}) + \mathbf{C}\mathbf{w}_{k-1} \\ \mathbf{y}_k = \mathbf{z}_{k-s} = \mathbf{h}(\mathbf{D}_s \mathbf{X}_k) + \mathbf{v}_{k-s} \quad s = 0, 1, \dots, d \end{cases} \quad (6)$$

式中:  $\mathbf{F}(\mathbf{X}_{k-1})$ 、 $\mathbf{C}$  和  $\mathbf{D}_s \mathbf{X}_k$  分别为

$$\begin{cases} \mathbf{F}(\mathbf{X}_{k-1}) = [\mathbf{f}^T(\mathbf{x}_{k-1}), \mathbf{x}_{k-1}^T, \dots, \mathbf{x}_{k-d}^T]^T \\ \mathbf{C} = [\mathbf{I}_{n \times n}, \mathbf{0}, \dots, \mathbf{0}]^T \\ \mathbf{D}_s \mathbf{X}_k = \mathbf{x}_{k-s} \end{cases} \quad (7)$$

滤波的递归过程为收到新的量测数据后, 求取扩增系统后验分布:

$$p(\mathbf{X}_k | \mathbf{y}_{1:k}) = \frac{p(\mathbf{y}_k | \mathbf{X}_k)}{p(\mathbf{y}_k | \mathbf{y}_{1:k-1})} p(\mathbf{X}_k | \mathbf{y}_{1:k-1}) \quad (8)$$

式中:  $p(\mathbf{y}_k | \mathbf{y}_{1:k-1})$  为归一化参数;  $p(\mathbf{X}_k | \mathbf{y}_{1:k-1})$  为预测分布, 按传统非线性高斯滤波算法计算获得;  $p(\mathbf{y}_k | \mathbf{X}_k)$  为似然函数, 与传统高斯滤波不同, 此处的似然函数不仅与当前状态  $\mathbf{x}_k$  相关, 还与过去时刻状态  $\mathbf{x}_{k-1}, \mathbf{x}_{k-2}, \dots, \mathbf{x}_{k-d}$  相关, 因此, 似然函数求取是滤波递推过程的关键。

### 1.2.1 求取似然函数

似然函数  $p(\mathbf{y}_k | \mathbf{X}_k)$  表示量测数据  $\mathbf{y}_k$  和扩增状态  $\mathbf{X}_k$  之间的关系。从式 (6) 可知, 接收到量测数据是  $\{\mathbf{z}_{k-s}\}_{s=0}^d$  中的一个, 将这个事件定义为  $\eta_k^s$ ,  $\eta_k^s$  共有  $d+1$  种可能取值。通过边缘化  $\eta_k^s$ , 可得似然函数:

$$p(\mathbf{y}_k | \mathbf{X}_k) = \sum_{s=0}^d p(\mathbf{y}_k, \eta_k^s | \mathbf{X}_k) = \sum_{s=0}^d (p(\eta_k^s) p(\mathbf{y}_k | \mathbf{X}_k, \eta_k^s)) \quad (9)$$

其中, 事件  $\eta_k^s$  发生的概率为

$$p(\eta_k^s) = p(\mathbf{y}_k = \mathbf{z}_{k-s}) = p^{(s,j)} \quad (10)$$

不同于常规贝叶斯滤波中的单峰概率密度函数, 式 (9) 似然函数是一多峰概率密度函数。由于式 (9) 包含更多关于量测与状态之间的信息, 因此, 与单峰概率密度函数相比, 式 (9) 所表示的似然函数更加精确。在量测噪声服从高斯分布时, 似然函数式 (9) 可以写成混合高斯形式:

$$p(\mathbf{y}_k | \mathbf{X}_k) = \sum_{s=0}^d p^{(s,j)} N(\mathbf{y}_k; \mathbf{h}(\mathbf{D}_s \mathbf{X}_k), \mathbf{R}_{k-s}) \quad (11)$$

式中:  $N(\mathbf{y}_k; \mathbf{h}(\mathbf{D}_s \mathbf{X}_k), \mathbf{R}_{k-s})$  为  $\mathbf{y}_k$  服从均值为  $\mathbf{h}(\mathbf{D}_s \mathbf{X}_k)$ 、方差为  $\mathbf{R}_{k-s}$  的正态分布。

### 1.2.2 传播预测分布

假设先验分布  $p(\mathbf{X}_{k-1} | \mathbf{y}_{1:k-1})$  和状态传递分布  $p(\mathbf{X}_k | \mathbf{X}_{k-1})$  均为高斯分布, 根据 Chapman-Kolmogorov 方程, 可得预测分布:

$$p(\mathbf{X}_k | \mathbf{y}_{1:k-1}) = \int_{\mathbf{R}^{n_a}} p(\mathbf{X}_k | \mathbf{X}_{k-1}) p(\mathbf{X}_{k-1} | \mathbf{y}_{1:k-1}) d\mathbf{X}_{k-1} = N(\mathbf{X}_k; \hat{\mathbf{X}}_{k/k-1}, \mathbf{P}_{k/k-1}) \quad (12)$$

其中:

$$\hat{\mathbf{X}}_{k/k-1} = \int_{\mathbf{R}^{n_a}} \mathbf{F}(\mathbf{X}_{k-1}) N(\mathbf{X}_{k-1}; \hat{\mathbf{X}}_{k-1/k-1}, \mathbf{P}_{k-1/k-1}) d\mathbf{X}_{k-1} \quad (13)$$

$$\mathbf{P}_{k/k-1} = \int_{\mathbf{R}^{n_a}} (\mathbf{F}(\mathbf{X}_{k-1}) - \hat{\mathbf{X}}_{k/k-1}) (\mathbf{F}(\mathbf{X}_{k-1}) - \hat{\mathbf{X}}_{k/k-1})^T N(\mathbf{X}_{k-1}; \hat{\mathbf{X}}_{k-1/k-1}, \mathbf{P}_{k-1/k-1}) d\mathbf{X}_{k-1} + \mathbf{C}\mathbf{Q}_{k-1}\mathbf{C}^T \quad (14)$$

式中:  $p(\mathbf{X}_{k-1} | \mathbf{y}_{1:k-1}) = N(\mathbf{X}_{k-1}; \hat{\mathbf{X}}_{k-1/k-1}, \mathbf{P}_{k-1/k-1})$ 。

### 1.2.3 更新后验分布

将式 (11) 和式 (12) 代入式 (8), 可得到后验分布:

$$p(\mathbf{X}_k | \mathbf{y}_k, \mathbf{y}_{1:k-1}) \approx \frac{N(\mathbf{X}_k; \hat{\mathbf{X}}_{k/k-1}, \mathbf{P}_{k/k-1})}{c_k} \sum_{s=0}^d p^{(s,j)} N(\mathbf{y}_k; \mathbf{h}(\mathbf{D}_s \mathbf{X}_k), \mathbf{R}_{k-s}) \quad (15)$$

式中:  $c_k = p(\mathbf{y}_k | \mathbf{y}_{1:k-1})$  为归一化参数, 根据全概率法则可得

$$c_k = \int_{\mathbf{R}^{n_a}} p(\mathbf{y}_k, \mathbf{X}_k | \mathbf{y}_{1:k-1}) d\mathbf{X}_k = \sum_{s=0}^d (p^{(s,j)} N(\mathbf{y}_k; \hat{\mathbf{y}}_{k/k-1}^s, \mathbf{P}_{k/k-1}^{yy,s} + \mathbf{R}_{k-s})) \quad (16)$$

其中:

$$\hat{\mathbf{y}}_{k/k-1}^s = \int_{\mathbf{R}^{n_a}} \mathbf{h}(\mathbf{D}_s \mathbf{X}_k) N(\mathbf{X}_k; \hat{\mathbf{X}}_{k/k-1}, \mathbf{P}_{k/k-1}) d\mathbf{X}_k \quad (17)$$

$$\mathbf{P}_{k/k-1}^{yy,s} = \int_{\mathbf{R}^{n_a}} (\mathbf{h}(\mathbf{D}_s \mathbf{X}_k) - \hat{\mathbf{y}}_{k/k-1}^s) (\mathbf{h}(\mathbf{D}_s \mathbf{X}_k) - \hat{\mathbf{y}}_{k/k-1}^s)^T N(\mathbf{X}_k; \hat{\mathbf{X}}_{k/k-1}, \mathbf{P}_{k/k-1}) d\mathbf{X}_k \quad (18)$$

### 1.3 修正似然容积卡尔曼滤波算法

采用三阶球面-径向容积准则计算修正似然非线性高斯滤波中的高斯加权积分, 详细给出修正似然容积卡尔曼滤波算法的递推计算更新过程。

#### 1.3.1 状态预测

根据式 (5), 可得  $k-1$  时刻的扩增状态估计  $\hat{\mathbf{X}}_{k-1/k-1}$  及其协方差  $\mathbf{P}_{k-1/k-1}$ :

$$\begin{cases} \hat{\mathbf{X}}_{k-1/k-1} = \begin{bmatrix} \hat{\mathbf{x}}_{k-1/k-1} \\ \vdots \\ \hat{\mathbf{x}}_{k-d/k-1} \\ \hat{\mathbf{x}}_{k-d-1/k-1} \end{bmatrix} \\ \mathbf{P}_{k-1/k-1} = \begin{bmatrix} \mathbf{P}_{k-1,k-1/k-1}^{xx} & \cdots & \mathbf{P}_{k-1,k-d/k-1}^{xx} & \mathbf{P}_{k-1,k-d-1/k-1}^{xx} \\ \vdots & & \vdots & \vdots \\ \mathbf{P}_{k-d,k-1/k-1}^{xx} & \cdots & \mathbf{P}_{k-d,k-d/k-1}^{xx} & \mathbf{P}_{k-d,k-d-1/k-1}^{xx} \\ \mathbf{P}_{k-d-1,k-1/k-1}^{xx} & \cdots & \mathbf{P}_{k-d-1,k-d/k-1}^{xx} & \mathbf{P}_{k-d-1,k-d-1/k-1}^{xx} \end{bmatrix} \end{cases} \quad (19)$$

扩增状态预测  $\hat{\mathbf{X}}_{k/k-1}$  及其协方差  $\mathbf{P}_{k/k-1}$ :

$$\begin{cases} \hat{\mathbf{X}}_{k/k-1} = \begin{bmatrix} \hat{\mathbf{x}}_{k/k-1} \\ \hat{\mathbf{x}}_{k-1/k-1} \\ \vdots \\ \hat{\mathbf{x}}_{k-d/k-1} \end{bmatrix} \\ \mathbf{P}_{k/k-1} = \begin{bmatrix} \mathbf{P}_{k,k/k-1}^{xx} & \mathbf{P}_{k,k-1/k-1}^{xx} & \cdots & \mathbf{P}_{k,k-d/k-1}^{xx} \\ \mathbf{P}_{k-1,k/k-1}^{xx} & \mathbf{P}_{k-1,k-1/k-1}^{xx} & \cdots & \mathbf{P}_{k-1,k-d/k-1}^{xx} \\ \vdots & \vdots & & \vdots \\ \mathbf{P}_{k-d,k/k-1}^{xx} & \mathbf{P}_{k-d,k-1/k-1}^{xx} & \cdots & \mathbf{P}_{k-d,k-d/k-1}^{xx} \end{bmatrix} \end{cases} \quad (20)$$

式 (20) 中除了  $\hat{\mathbf{x}}_{k/k-1}$ 、 $\mathbf{P}_{k,k/k-1}^{xx}$ 、 $\{\mathbf{P}_{k,k-s/k-1}^{xx}\}_{s=1}^d$  和  $\{\mathbf{P}_{k-s,k/k-1}^{xx}\}_{s=1}^d$  是新出现的之外,其余分块矩阵可从式 (19) 中获得,本节将计算获得这些新出现的分块矩阵。

基于  $\hat{\mathbf{X}}_{k-1/k-1}$ 、 $\mathbf{P}_{k-1/k-1}$  和三阶球面-径向容积点序列  $\chi_i = [1]_i$  ( $[1]_i$  为式 (22) 中  $[1]$  的第  $i$  个  $n_a$  行向量) 生成状态预测所需容积点:

$$\chi_{i,k-1/k-1} = \begin{bmatrix} \chi_{i,k-1/k-1}^x \\ \vdots \\ \chi_{i,k-d/k-1}^x \\ \chi_{i,k-d-1/k-1}^x \end{bmatrix} = \sqrt{\mathbf{P}_{k-1/k-1}} \chi_i + \hat{\mathbf{X}}_{k-1/k-1} \quad i = 1, 2, \dots, 2n_a \quad (21)$$

$$[1] = \sqrt{n_a} \underbrace{\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \right\}}_{\text{共 } 2n_a \text{ 个}} \quad (22)$$

状态预测  $\hat{\mathbf{x}}_{k/k-1}$ 、协方差  $\mathbf{P}_{k,k/k-1}^{xx}$ 、互协方差  $\mathbf{P}_{k,k-s/k-1}^{xx}$  和  $\mathbf{P}_{k-s,k/k-1}^{xx}$  计算如下:

$$\hat{\mathbf{x}}_{k/k-1} = \frac{1}{2n_a} \sum_{i=1}^{2n_a} f(\chi_{i,k-1/k-1}^x) \quad (23)$$

$$\mathbf{P}_{k,k/k-1}^{xx} = \frac{1}{2n_a} \sum_{i=1}^{2n_a} (f(\chi_{i,k-1/k-1}^x) - \hat{\mathbf{x}}_{k/k-1}) \cdot (f(\chi_{i,k-1/k-1}^x) - \hat{\mathbf{x}}_{k/k-1})^T + \mathbf{Q}_{k-1} \quad (24)$$

$$\mathbf{P}_{k,k-s/k-1}^{xx} = \frac{1}{2n_a} \sum_{i=1}^{2n_a} (f(\chi_{i,k-1/k-1}^x) - \hat{\mathbf{x}}_{k/k-1}) \cdot (f(\chi_{i,k-s/k-1}^x) - \hat{\mathbf{x}}_{k-s/k-1})^T \quad s = 1, 2, \dots, d \quad (25)$$

$$\mathbf{P}_{k-s,k/k-1}^{xx} = (\mathbf{P}_{k,k-s/k-1}^{xx})^T \quad s = 1, 2, \dots, d \quad (26)$$

将式 (23)~式 (26) 代入式 (20), 可得到扩增状态预测  $\hat{\mathbf{X}}_{k/k-1}$  及其协方差  $\mathbf{P}_{k/k-1}$ 。

### 1.3.2 量测更新

1) 生成量测更新所需容积点。基于扩增状态预测  $\hat{\mathbf{X}}_{k/k-1}$  及其协方差  $\mathbf{P}_{k/k-1}$  和三阶球面-径向容积点序列  $\chi_i$  构造量测更新所需容积点:

$$\chi_{i,k/k-1} = \begin{bmatrix} \chi_{i,k/k-1}^x \\ \vdots \\ \chi_{i,k-d+1/k-1}^x \\ \chi_{i,k-d/k-1}^x \end{bmatrix} = \sqrt{\mathbf{P}_{k/k-1}} \chi_i + \hat{\mathbf{X}}_{k/k-1} \quad i = 1, 2, \dots, 2n_a \quad (27)$$

2) 计算第  $s$  个子更新所需量。计算延迟  $s$  步量测预测均值  $\hat{y}_{k/k-1}^s$ 、协方差  $\mathbf{P}_{k/k-1}^{yy,s}$ 、互协方差  $\mathbf{P}_{k/k-1}^{xy,s}$  和滤波增益  $\mathbf{K}_k^s$ :

$$\hat{y}_{k/k-1}^s = \frac{1}{2n_a} \sum_{i=1}^{2n_a} h(\mathbf{D}_s \chi_{i,k/k-1}) \quad (28)$$

$$\mathbf{P}_{k/k-1}^{yy,s} = \frac{1}{2n_a} \sum_{i=1}^{2n_a} (h(\mathbf{D}_s \chi_{i,k/k-1}) - \hat{y}_{k/k-1}^s) \cdot (h(\mathbf{D}_s \chi_{i,k/k-1}) - \hat{y}_{k/k-1}^s)^T \quad (29)$$

$$\mathbf{P}_{k/k-1}^{xy,s} = \frac{1}{2n_a} \sum_{i=1}^{2n_a} (\chi_{i,k/k-1} - \hat{\mathbf{X}}_{k/k-1}) \cdot (h(\mathbf{D}_s \chi_{i,k/k-1}) - \hat{y}_{k/k-1}^s)^T \quad (30)$$

$$\mathbf{K}_k^s = \mathbf{P}_{k/k-1}^{xy,s} (\mathbf{P}_{k/k-1}^{yy,s} + \mathbf{R}_{k-s})^{-1} \quad (31)$$

3) 第  $s$  个子更新。用新接收到的量测  $y_k$  和延迟  $s$  步量测预测  $\hat{y}_{k/k-1}^s$  ( $0 \leq s \leq d$ ) 更新得到  $\hat{\mathbf{X}}_{k/k}^s$  和  $\mathbf{P}_{k/k}^s$ :

$$\hat{\mathbf{X}}_{k/k}^s = \hat{\mathbf{X}}_{k/k-1} + \mathbf{K}_k^s (y_k - \hat{y}_{k/k-1}^s) \quad s = 0, 1, \dots, d \quad (32)$$

$$\mathbf{P}_{k/k}^s = \mathbf{P}_{k/k-1} - \mathbf{K}_k^s (\mathbf{P}_{k/k-1}^{xy,s})^T \quad s = 0, 1, \dots, d \quad (33)$$

4)  $d+1$  个子更新融合结果。融合  $d+1$  个子更新的结果  $\hat{\mathbf{X}}_{k/k}^s$  和  $\mathbf{P}_{k/k}^s$  得到扩增状态估计及其协方差:

$$\hat{\mathbf{X}}_{k/k} = \sum_{s=0}^d \mu_k^s \hat{\mathbf{X}}_{k/k}^s \quad (34)$$

$$\mathbf{P}_{k/k} = \sum_{s=0}^d \mu_k^s \left[ \mathbf{P}_{k/k}^s + (\hat{\mathbf{X}}_{k/k} - \hat{\mathbf{X}}_{k/k}^s) (\hat{\mathbf{X}}_{k/k} - \hat{\mathbf{X}}_{k/k}^s)^T \right] \quad (35)$$

式中:  $\mu_k^s$  为每个子更新的加权因子, 可根据残差和延迟概率实时调节:

$$\mu_k^s = \frac{p^{(s,j)} N(y_k; \hat{y}_{k/k-1}^s, \mathbf{P}_{k/k-1}^{yy,s} + \mathbf{R}_{k-s})}{\sum_{i=0}^d [p^{(i,j)} N(y_k; \hat{y}_{k/k-1}^i, \mathbf{P}_{k/k-1}^{yy,i} + \mathbf{R}_{k-i})]} \quad (36)$$

修正似然容积卡尔曼滤波算法的具体流程如图 1 所示。

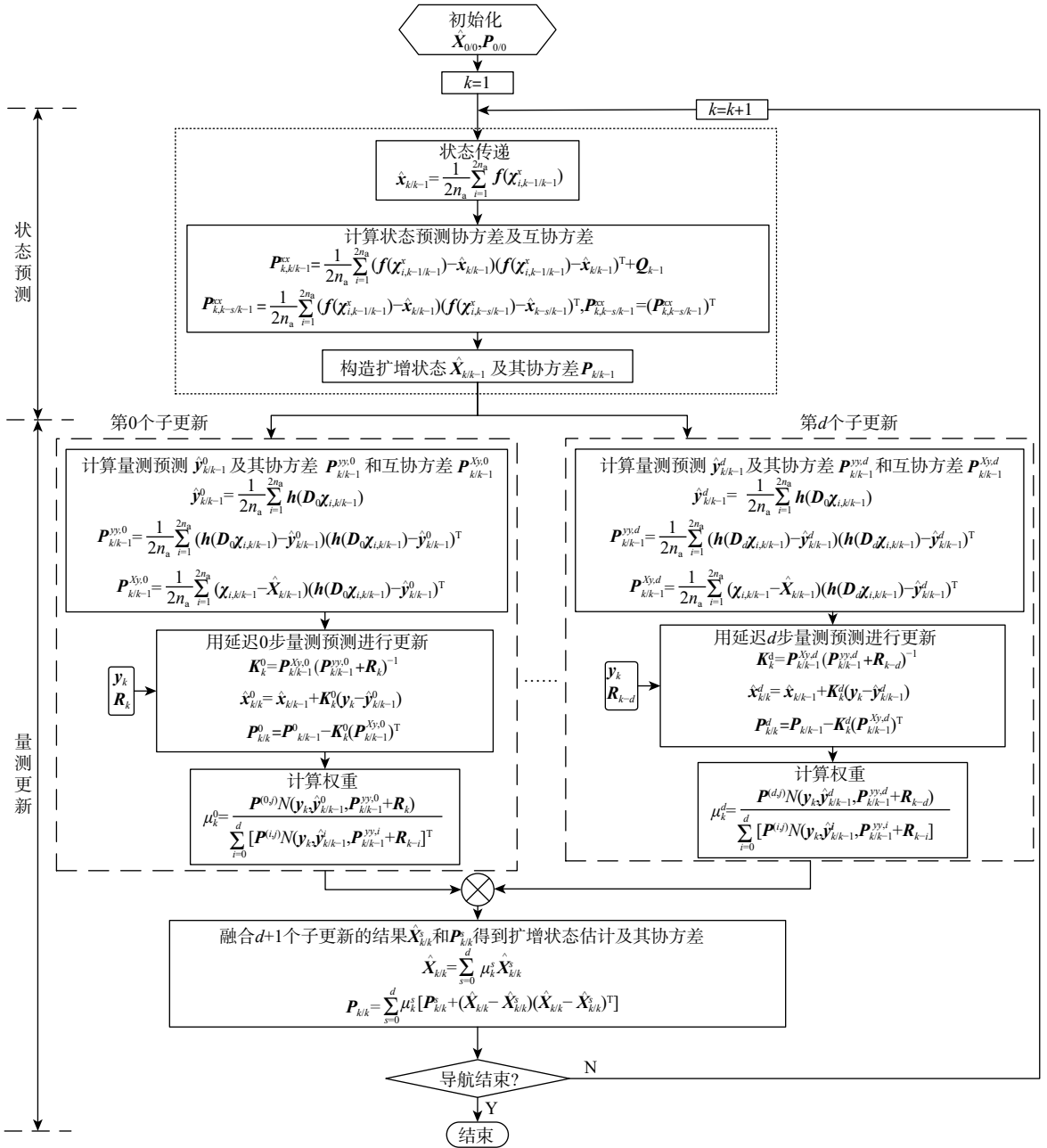


图1 修正似然容积卡尔曼滤波流程

Fig. 1 Flowchart of modified likelihood cubature Kalman filtering

## 2 相对导航系统建模

以僚机本体系为基准建立僚机和长机之间的相对姿态、位置和速度方程,以长机上的导航相机与僚机上特征光点之间的视线矢量作为量测量,将长机上的惯性和视觉量测信息传输给僚机进行融合估计得到僚机和长机之间的相对运动信息。

### 2.1 相对导航方程

#### 2.1.1 相对姿态运动方程

采用工程实用性强的四元数表示姿态,即

$$q = \begin{bmatrix} q_{13} \\ q_4 \end{bmatrix} = \begin{bmatrix} e \sin \frac{\theta}{2} \\ \cos \frac{\theta}{2} \end{bmatrix} \quad (37)$$

式中:  $q_{13} = [q_1 \ q_2 \ q_3]^T$  为姿态四元数的矢量部分;  $q_4$  为姿态四元数的标量部分;  $e$  和  $\theta$  分别为欧拉旋转轴和旋转角。

相对姿态四元数  $q_{fl}$  定义为

$$q_{fl} = q_f \otimes q_l^{-1} \quad (38)$$

式中:  $q_f$  和  $q_l$  分别为僚机和长机的姿态四元数;  $q^{-1} = [-q_{13} \ q_4]^T$  为四元数  $q$  的共轭四元数。

符号  $\otimes$  表示四元数乘法,定义为

$$p \otimes q = [\Psi(p) \ p] \ q = [\Xi(q) \ q] \ p \quad (39)$$

其中:

$$\Xi(q) = \begin{bmatrix} q_4 I_{3 \times 3} + [q_{13} \times] \\ -q_{13}^T \end{bmatrix} \quad (40)$$

$$\Psi(p) = \begin{bmatrix} p_4 \mathbf{I}_{3 \times 3} - [p_{13} \times] \\ -p_{13}^T \end{bmatrix} \quad (41)$$

相对姿态运动学方程为

$$\dot{q}_{\text{fl}} = \frac{1}{2} \Xi(q_{\text{fl}}) \omega_{\text{fl}}^f \quad (42)$$

式中:  $\omega_{\text{fl}}^f$  为相对角速度, 计算如下:

$$\omega_{\text{fl}}^f = \omega_{\text{fl}}^f - \mathbf{C}_1^f \omega_{\text{ll}}^l \quad (43)$$

相对姿态矩阵(长机本体系至僚机本体系的姿态转换矩阵)  $\mathbf{C}_1^f$  定义如下:

$$\mathbf{C}_1^f = \Xi^T(q_{\text{fl}}) \Psi(q_{\text{fl}}) \quad (44)$$

### 2.1.2 相对质心运动方程

相对位置定义如下:

$$\mathbf{r}_{\text{fl}}^f = \mathbf{r}_{\text{fl}}^f - \mathbf{C}_1^f \mathbf{r}_{\text{ll}}^l = \mathbf{C}_1^f (\mathbf{r}_{\text{fl}}^l - \mathbf{r}_{\text{ll}}^l) \quad (45)$$

式中:  $\mathbf{r}_{\text{fl}}^f$  为僚机相对于长机的位置矢量在僚机本体系下的表示;  $\mathbf{r}_{\text{fl}}^l$  和  $\mathbf{r}_{\text{ll}}^l$  分别为僚机和长机在惯性坐标系中的位置矢量在各自本体系下的表示;  $\mathbf{C}_1^f$  为惯性坐标系至僚机本体系的姿态转换矩阵。

求式(45)的二阶导, 得到

$$\dot{\mathbf{r}}_{\text{fl}}^f = \frac{d^2}{dt^2} (\mathbf{C}_1^f) (\mathbf{r}_{\text{fl}}^l - \mathbf{r}_{\text{ll}}^l) + 2 \frac{d}{dt} (\mathbf{C}_1^f) \frac{d}{dt} (\mathbf{r}_{\text{fl}}^l - \mathbf{r}_{\text{ll}}^l) +$$

$$\mathbf{C}_1^f \frac{d^2}{dt^2} (\mathbf{r}_{\text{fl}}^l - \mathbf{r}_{\text{ll}}^l) = -\dot{\omega}_{\text{fl}}^f \times \mathbf{r}_{\text{fl}}^f - [\omega_{\text{fl}}^f \times] \cdot [\omega_{\text{fl}}^f \times] \mathbf{r}_{\text{fl}}^f - 2\dot{\omega}_{\text{fl}}^f \times \dot{\mathbf{r}}_{\text{fl}}^f + \mathbf{C}_1^f (\ddot{\mathbf{r}}_{\text{fl}}^l - \ddot{\mathbf{r}}_{\text{ll}}^l) \quad (46)$$

式中:  $[\omega_{\text{fl}}^f \times]$  为  $\omega_{\text{fl}}^f$  的反对称阵。

根据牛顿法则给出长机和僚机的惯性加速度:

$$\begin{cases} \ddot{\mathbf{r}}_{\text{ll}}^l = \mathbf{a}_{\text{ll}}^l + \mathbf{g}_{\text{ll}}^l \\ \ddot{\mathbf{r}}_{\text{fl}}^l = \mathbf{a}_{\text{fl}}^l + \mathbf{g}_{\text{fl}}^l \end{cases} \quad (47)$$

将式(47)代入式(46), 可得

$$\begin{aligned} \ddot{\mathbf{r}}_{\text{fl}}^f = & -\dot{\omega}_{\text{fl}}^f \times \dot{\mathbf{r}}_{\text{fl}}^f - [\omega_{\text{fl}}^f \times] [\omega_{\text{fl}}^f \times] \mathbf{r}_{\text{fl}}^f - \\ & 2\dot{\omega}_{\text{fl}}^f \times \dot{\mathbf{r}}_{\text{fl}}^f + \mathbf{a}_{\text{fl}}^f - \mathbf{C}_1^f \mathbf{a}_{\text{ll}}^l + \mathbf{C}_1^f (\mathbf{g}_{\text{fl}}^l - \mathbf{g}_{\text{ll}}^l) \end{aligned} \quad (48)$$

## 2.2 相对导航传感器测量模型

### 2.2.1 惯性导航系统测量模型

加速度计和陀螺仪是惯性导航系统中重要的传感器, 其测量模型分别为

$$\begin{cases} \hat{\mathbf{a}}_{\text{fl}}^f = \mathbf{a}_{\text{fl}}^f + \mathbf{b}_{\text{fa}} + \boldsymbol{\eta}_{\text{fa}} \\ \hat{\mathbf{b}}_{\text{fa}} = \boldsymbol{\eta}_{\text{fa}} \end{cases} \quad (49)$$

$$\begin{cases} \hat{\boldsymbol{\omega}}_{\text{fl}}^f = \boldsymbol{\omega}_{\text{fl}}^f + \mathbf{b}_{\text{fg}} + \boldsymbol{\eta}_{\text{fg}} \\ \hat{\mathbf{b}}_{\text{fg}} = \boldsymbol{\eta}_{\text{fg}} \end{cases} \quad (50)$$

式中:  $\mathbf{b}_{\text{fa}}$  和  $\mathbf{b}_{\text{fg}}$  分别为加速度计和陀螺仪的漂移;  $\boldsymbol{\eta}_{\text{fa}}$  和  $\boldsymbol{\eta}_{\text{fg}}$  分别为加速度计和陀螺仪的噪声;  $\boldsymbol{\eta}_{\text{fa}}$  和  $\boldsymbol{\eta}_{\text{fg}}$  分别为加速度计和陀螺仪的随机游走。

### 2.2.2 视觉导航系统测量模型

视觉导航系统的量测量是长机上导航相机与僚机上特征光点之间的视线矢量, 其测量原理如图2

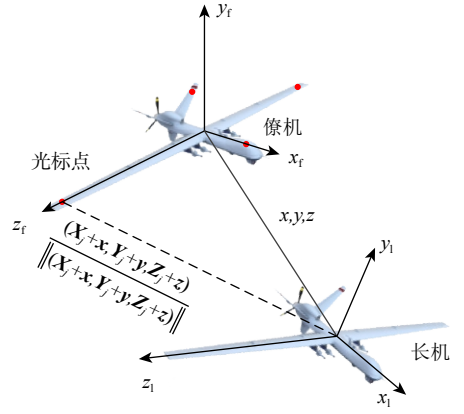


图2 视觉传感器量测

Fig. 2 Measurement of visual sensor

所示, 量测量为

$$\mathbf{b}_j = (\mathbf{C}_1^f)^T \frac{1}{\sqrt{(X_j+x)^2 + (Y_j+y)^2 + (Z_j+z)^2}} \begin{bmatrix} X_j+x \\ Y_j+y \\ Z_j+z \end{bmatrix} \quad (51)$$

式中:  $[X_j \ Y_j \ Z_j]^T$  为第  $j$  个特征光点在僚机本体系上的位置;  $x, y, z$  为相对位置向量  $\mathbf{r}_{\text{fl}}^f$  的三轴分量。

## 3 相对导航滤波器设计

### 3.1 状态向量

若直接把受归一化约束的四元数作为状态分量, 存在四元数姿态协方差阵奇异性问题, 因此, 采用一种无约束的三参数罗德里格斯参数表示姿态误差来传播和更新姿态四元数, 进行滤波器设计。

定义状态矢量:

$$\mathbf{x}_k = \left[ \delta \mathbf{s}_k^T \ (\mathbf{r}_{\text{fl},k}^f)^T \ (\mathbf{v}_{\text{fl},k}^f)^T \ \mathbf{b}_{\text{fa},k}^T \ \mathbf{b}_{\text{fg},k}^T \right]^T \quad (52)$$

式中:  $\delta \mathbf{s}_k$  为罗德里格斯参数;  $\mathbf{r}_{\text{fl},k}^f$  为相对位置矢量;  $\mathbf{v}_{\text{fl},k}^f$  为相对速度矢量;  $\mathbf{b}_{\text{fa},k}$  为加速度计漂移;  $\mathbf{b}_{\text{fg},k}$  为陀螺仪漂移。

### 3.2 状态预测

$k-1$ 时刻相对导航滤波器的扩增状态为

$$\left\{ \begin{aligned} \hat{\mathbf{x}}_{k-1/k-1} &= \begin{bmatrix} \hat{\mathbf{x}}_{k-1/k-1} \\ \vdots \\ \hat{\mathbf{x}}_{k-d/k-1} \\ \hat{\mathbf{x}}_{k-d-1/k-1} \end{bmatrix} \\ \hat{\mathbf{x}}_{k-1/k-1} &= \begin{bmatrix} \delta \hat{\mathbf{s}}_{k-1/k-1} \\ \hat{\mathbf{r}}_{\text{fl},k-1/k-1}^f \\ \hat{\mathbf{v}}_{\text{fl},k-1/k-1}^f \\ \mathbf{b}_{\text{fa},k-1/k-1} \\ \mathbf{b}_{\text{fg},k-1/k-1} \end{bmatrix} \end{aligned} \right. \quad (53)$$

根据罗德里格斯参数获得姿态四元数误差为

$$\delta \mathbf{q}_{\Pi, i, k-1/k-1} = \begin{bmatrix} \delta \mathbf{q}_{13, i, k-1/k-1}^T & \delta q_{4, i, k-1/k-1} \end{bmatrix}^T \quad (54)$$

$$\delta q_{4, i, k-1/k-1} = \left[ -\alpha \|\boldsymbol{\chi}_{i, k-1/k-1}^{\delta s}\|^2 + \lambda \sqrt{\lambda^2 + (1 - \alpha^2) \|\boldsymbol{\chi}_{i, k-1/k-1}^{\delta s}\|^2} \right] / \left( \lambda^2 + \|\boldsymbol{\chi}_{i, k-1/k-1}^{\delta s}\|^2 \right) \quad (55)$$

$$\delta \mathbf{q}_{13, i, k-1/k-1} = \lambda^{-1} [\alpha + \delta q_{4, i, k-1/k-1}] \boldsymbol{\chi}_{i, k-1/k-1}^{\delta s} \quad (56)$$

式中:  $\alpha$  和  $\lambda$  为尺度参数, 具体选择为  $\alpha = 0, \lambda = 2$ ; 容积点  $\boldsymbol{\chi}_{i, k-1/k-1}^{\delta s}$  按式 (21) 计算得到:

$$\boldsymbol{\chi}_{i, k-1/k-1}^x = \begin{bmatrix} \boldsymbol{\chi}_{i, k-1/k-1}^{\delta s} \\ \delta \boldsymbol{\chi}_{i, k-1/k-1}^r \\ \delta \boldsymbol{\chi}_{i, k-1/k-1}^v \\ \delta \boldsymbol{\chi}_{i, k-1/k-1}^{b_{fa}} \\ \delta \boldsymbol{\chi}_{i, k-1/k-1}^{b_{fg}} \end{bmatrix} \quad (57)$$

根据  $k-1$  时刻估计的姿态四元数均值和姿态四元数误差得到:

$$\hat{\mathbf{q}}_{\Pi, 0, k-1/k-1} = \hat{\mathbf{q}}_{\Pi, k-1/k-1} \quad (58)$$

$$\hat{\mathbf{q}}_{\Pi, i, k-1/k-1} = \delta \mathbf{q}_{\Pi, i, k-1/k-1} \otimes \hat{\mathbf{q}}_{\Pi, k-1/k-1} \quad i = 1, 2, \dots, 2n_a \quad (59)$$

按式 (42) 相对姿态运动传播得到新的姿态四元数:

$$\hat{\mathbf{q}}_{\Pi, i, k/k-1} = \hat{\mathbf{q}}_{\Pi, i, k-1/k-1} + \hat{\mathbf{q}}_{\Pi, i, k-1/k-1} \Delta t \quad (60)$$

式 (60) 采用欧拉积分对姿态进行递推, 对于高动态情况再考虑采用其他积分方法。  $\hat{\mathbf{q}}_{\Pi, i, k-1/k-1}$  计算如下:

$$\dot{\hat{\mathbf{q}}}_{\Pi, i, k-1/k-1} = \frac{1}{2} \boldsymbol{\Xi} (\hat{\mathbf{q}}_{\Pi, i, k-1/k-1}) \hat{\boldsymbol{\omega}}_{\Pi}^f \quad (61)$$

$$\hat{\boldsymbol{\omega}}_{\Pi}^f = \hat{\boldsymbol{\omega}}_{\Pi}^f - \boldsymbol{\chi}_{i, k-1/k-1}^{b_{fa}} - \mathbf{C}_1^f (\hat{\mathbf{q}}_{\Pi, i, k-1/k-1}) \boldsymbol{\omega}_{\Pi}^1 \quad (62)$$

传播后的姿态四元数误差为

$$\delta \hat{\mathbf{q}}_{\Pi, i, k/k-1} = \hat{\mathbf{q}}_{\Pi, i, k/k-1} \otimes [\hat{\mathbf{q}}_{\Pi, 0, k/k-1}]^{-1} \quad (63)$$

由传播后的姿态四元数误差计算传播后的罗德里格斯参数:

$$\boldsymbol{\chi}_{i, k/k-1}^{\delta s} = \lambda \frac{\delta \mathbf{q}_{13, i, k/k-1}}{\alpha + \delta q_{4, i, k/k-1}} \quad (64)$$

式中:  $\delta \mathbf{q}_{\Pi, i, k/k-1} = [\delta \mathbf{q}_{13, i, k/k-1}^T \quad \delta q_{4, i, k/k-1}]^T$ 。

计算传播后的相对位置和速度

$$\begin{cases} \boldsymbol{\chi}_{i, k/k-1}^r = \boldsymbol{\chi}_{i, k-1/k-1}^r + \boldsymbol{\chi}_{i, k-1/k-1}^v \Delta t \\ \boldsymbol{\chi}_{i, k/k-1}^v = \boldsymbol{\chi}_{i, k-1/k-1}^v + \boldsymbol{\chi}_{i, k-1/k-1}^a \Delta t \end{cases} \quad (65)$$

式中:  $\boldsymbol{\chi}_{i, k-1/k-1}^a$  按式 (48) 计算得到。

传播后的加速度计和陀螺仪漂移为

$$\begin{cases} \boldsymbol{\chi}_{i, k/k-1}^{b_{fa}} = \boldsymbol{\chi}_{i, k-1/k-1}^{b_{fa}} \\ \boldsymbol{\chi}_{i, k/k-1}^{b_{fg}} = \boldsymbol{\chi}_{i, k-1/k-1}^{b_{fg}} \end{cases} \quad (66)$$

将式 (64)~式 (66) 和式 (23)~式 (26) 代入

式 (20) 中可得到扩增状态预测及其协方差。

### 3.3 量测更新

延迟  $s(0 \leq d)$  步量测预测:

$$\hat{\mathbf{y}}_{i, k/k-1}^s = \begin{bmatrix} (\mathbf{b}_{i, k/k-1}^{s,1})^T & (\mathbf{b}_{i, k/k-1}^{s,2})^T & \dots & (\mathbf{b}_{i, k/k-1}^{s,N})^T \end{bmatrix}^T \quad (67)$$

式中:  $N$  为特征光点数量; 第  $j$  个特征光点量测

$\mathbf{b}_{i, k/k-1}^{s,j}$ :

$$\mathbf{b}_{i, k/k-1}^{s,j} = \frac{[\mathbf{C}_1^f (\hat{\mathbf{q}}_{\Pi, i, k-s/k-1})]^T \begin{bmatrix} \mathbf{X}_j + \boldsymbol{\chi}_{i, k-s/k-1}^{rx} \\ \mathbf{Y}_j + \boldsymbol{\chi}_{i, k-s/k-1}^{ry} \\ \mathbf{Z}_j + \boldsymbol{\chi}_{i, k-s/k-1}^{rz} \end{bmatrix}}{\sqrt{(\mathbf{X}_j + \boldsymbol{\chi}_{i, k-s/k-1}^{rx})^2 + (\mathbf{Y}_j + \boldsymbol{\chi}_{i, k-s/k-1}^{ry})^2 + (\mathbf{Z}_j + \boldsymbol{\chi}_{i, k-s/k-1}^{rz})^2}} \quad (68)$$

$$\boldsymbol{\chi}_{i, k-s/k-1}^r = [\boldsymbol{\chi}_{i, k-s/k-1}^{rx} \quad \boldsymbol{\chi}_{i, k-s/k-1}^{ry} \quad \boldsymbol{\chi}_{i, k-s/k-1}^{rz}]^T \quad (69)$$

将式 (67) 和式 (28)~式 (33) 代入式 (34) 可得到量测更新后的扩增状态:

$$\hat{\mathbf{x}}_{k/k} = \begin{cases} \begin{bmatrix} \hat{\mathbf{x}}_{k/k} \\ \hat{\mathbf{x}}_{k-1/k} \\ \vdots \\ \hat{\mathbf{x}}_{k-d/k} \end{bmatrix} \\ \hat{\mathbf{x}}_{k/k} = \begin{bmatrix} \delta \hat{\mathbf{s}}_{k/k} \\ \hat{\mathbf{r}}_{\Pi, k/k}^f \\ \hat{\mathbf{v}}_{\Pi, k/k}^f \\ \mathbf{b}_{fa, k/k} \\ \mathbf{b}_{fg, k/k} \end{bmatrix} \end{cases} \quad (70)$$

相对导航的位置  $\hat{\mathbf{r}}_{\Pi, k/k}^f$  和速度  $\hat{\mathbf{v}}_{\Pi, k/k}^f$  从式 (70) 获得, 相对姿态计算如下:

$$\hat{\mathbf{q}}_{\Pi, k/k} = \delta \mathbf{q}_{\Pi, k/k} \otimes \hat{\mathbf{q}}_{\Pi, 0, k/k-1} \quad (71)$$

式中:  $\delta \mathbf{q}_{\Pi, k/k}$  由罗德里格斯参数估计值  $\delta \hat{\mathbf{s}}_{k/k}$  按式 (54)~式 (56) 计算得到。

融合式 (19)~式 (36) 修正似然容积卡尔曼滤波算法和式 (52)~式 (71) 相对导航滤波器设计过程, 基于 ML-CKF 的相对导航滤波器流程如下。

1) 状态预测。

- ① 根据式 (21) 计算状态预测容积点;
- ② 根据式 (54)~式 (57) 计算姿态四元数误差;
- ③ 根据式 (58) 和式 (59) 计算状态预测所需的姿态四元数容积点;
- ④ 根据式 (60)、式 (65) 和式 (66) 进行相对姿态、相对位置、相对速度、加速度计漂移和陀螺仪漂移递推;

⑤ 根据式 (64) 计算经过状态递推后的罗德里格斯参数估计值;

⑥ 根据式 (23)~式 (26) 和式 (20) 计算扩增状态及其协方差。

2) 量测更新。

- ① 根据式 (27) 计算量测更新所需容积点;
- ② 根据式 (67) 计算延迟  $s(0 \leq s \leq d)$  步量测预测

容积点;

③ 根据式 (28)~式 (31) 计算延迟  $s(0 \leq s \leq d)$  步量测预测均值、协方差、互协方差和滤波增益;

④ 根据式 (32)~式 (33) 计算第  $s$  个子更新结果;

⑤ 根据式 (34)~式 (35) 进行扩增状态及其协方差更新;

⑥ 根据式 (71) 计算量测更新后的罗德里格斯参数, 根据式 (70) 可见相对导航位置和速度可从扩增状态中获取。

## 4 仿真分析

为了验证所提算法在处理随机时延量测上的优越性, 将开展数值仿真验证, 将 ML-CKF 与现有 CKF<sup>[9]</sup>、ORD-CKF<sup>[16]</sup> 和 MRD-CKF<sup>[17]</sup> 进行比较。

### 4.1 仿真初始条件

惯性器件、视觉导航系统的采样时间及滤波步长均为 0.1 s, 传感器的偏差相关参数如表 1 所示。

表 1 惯性和视觉传感器偏差参数

Table 1 Deviation parameter of inertial and visual sensors

参数	数值
加速度计初始漂移/mg	0.2
加速度计随机游走/(mg·s <sup>-1/2</sup> )	0.002
加速度计噪声/(mg·s <sup>1/2</sup> )	0.02
陀螺仪初始漂移/(°·h <sup>-1</sup> )	0.1
陀螺仪随机游走/(°·h <sup>-3/2</sup> )	0.06
陀螺仪噪声/(°·h <sup>-1/2</sup> )	0.01
视觉传感器噪声/μrad	80

3 个特征光点即可保证惯性/视觉相对导航系统完全可观测, 为了提高相对导航精度, 将特征光点数提高一倍, 6 个特征光点位置分布如表 2 所示。

表 2 特征光点位置列表

Table 2 List of beacon locations

特征光点标号	X <sub>j</sub> /m	Y <sub>j</sub> /m	Z <sub>j</sub> /m
1	1.5	0	0
2	-2.5	0	0
3	0	2.5	0
4	0	-2.5	0
5	0	3.5	-0.5
6	0	-3.5	0.5

假设长机运动在北-天-东坐标系内, 初始时刻从经度 120°、纬度 40°和高度 1 000 m 位置开始运动, 僚机与长机形成编队飞行, 参考文献 [20] 规划典型轨迹, 长机和僚机的位置变化方程如式 (72) 和式 (73) 所示, 相应的运动轨迹如图 3 所示。

$$\mathbf{r}_{f/N}^N = \begin{bmatrix} 100t + 4\cos(0.02t) \\ 1\,000\sin(0.04t) \\ 1\,000 + 0.5t \end{bmatrix} \quad (72)$$

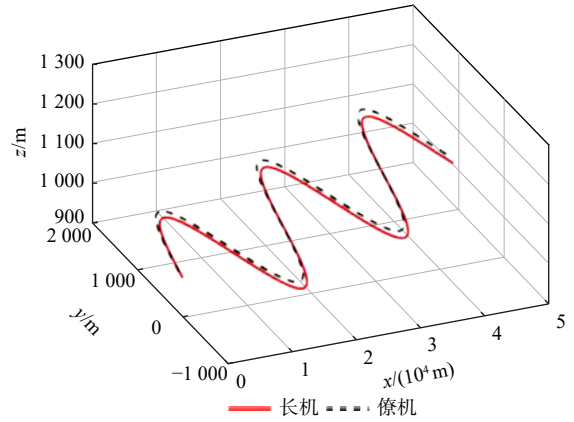


图 3 长机和僚机飞行轨迹

Fig. 3 Flight path of leader and follower

$$\mathbf{r}_{f/N}^N = \begin{bmatrix} 100t + 4\cos(0.02t) + 3\cos(0.01t) + 100 + 0.01t \\ 1\,000\sin(0.04t) + 2\cos(0.001t) + 100 + 0.01t \\ 1\,000 + 0.5t + 4\cos(0.001t) + 0.01t \end{bmatrix} \quad (73)$$

假设视觉量测最大延迟步数为 3 步, 无延迟概率  $p^{(0,j)}$ 、延迟 1 步概率  $p^{(1,j)}$ 、延迟 2 步概率  $p^{(2,j)}$  和延迟 3 步概率  $p^{(3,j)}$  分别为

$$\begin{cases} p^{(0,j)} = 1 - p_d \\ p^{(1,j)} = \frac{p_d}{2} \\ p^{(2,j)} = \frac{p_d}{4} \\ p^{(3,j)} = \frac{p_d}{4} \end{cases} \quad (74)$$

式中:  $p_d$  为总延迟概率。

### 4.2 仿真结果与分析

取延迟概率  $p_d = 0.3$  进行仿真, 得到仿真结果如图 4~图 9 所示。

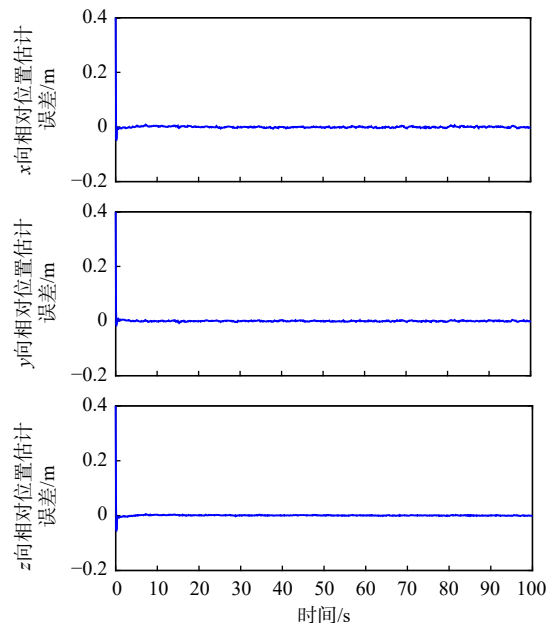


图 4 相对位置估计误差

Fig. 4 Estimation error of relative position

图 4~图 6 分别为 ML-CKF 估计得到的相对位置、相对速度和相对姿态误差, 从图 4~图 6 中曲线可知基于 ML-CKF 的相对导航误差曲线能够快速收敛。图 7~图 9 是 CKF、ORD-CKF、MRD-CKF 和 ML-CKF 算法 100 次蒙特卡罗打靶仿真三维相对位置、速度、姿态误差模值的比较结果。仿真结果表明, CKF 的估计精度最差, 这是由于其未对延迟量测进行处理; MRD-CKF 和 ML-CKF 的估计精度高于 ORD-CKF, 这是因为 MRD-CKF 和 ML-CKF 能够处理量测存在多步随机时延问题, 而 ORD-CKF 只能处理单步随机延迟量测。此外, 与 MRD-CKF 相比, ML-CKF 具有更高的估计精度, 这归功于

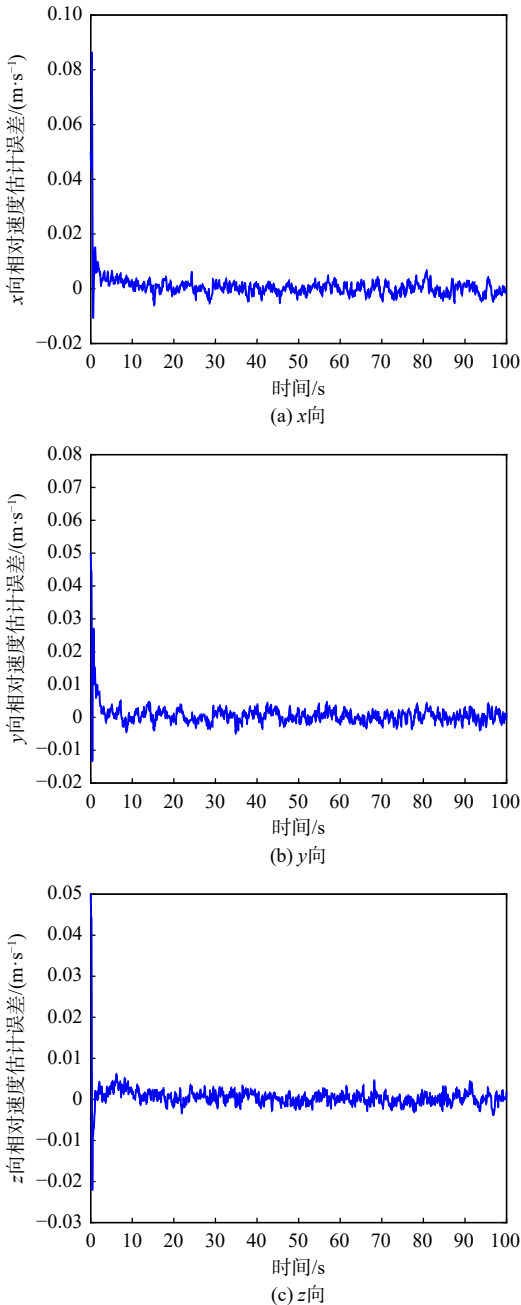


图 5 相对速度估计误差

Fig. 5 Estimation error of relative velocity

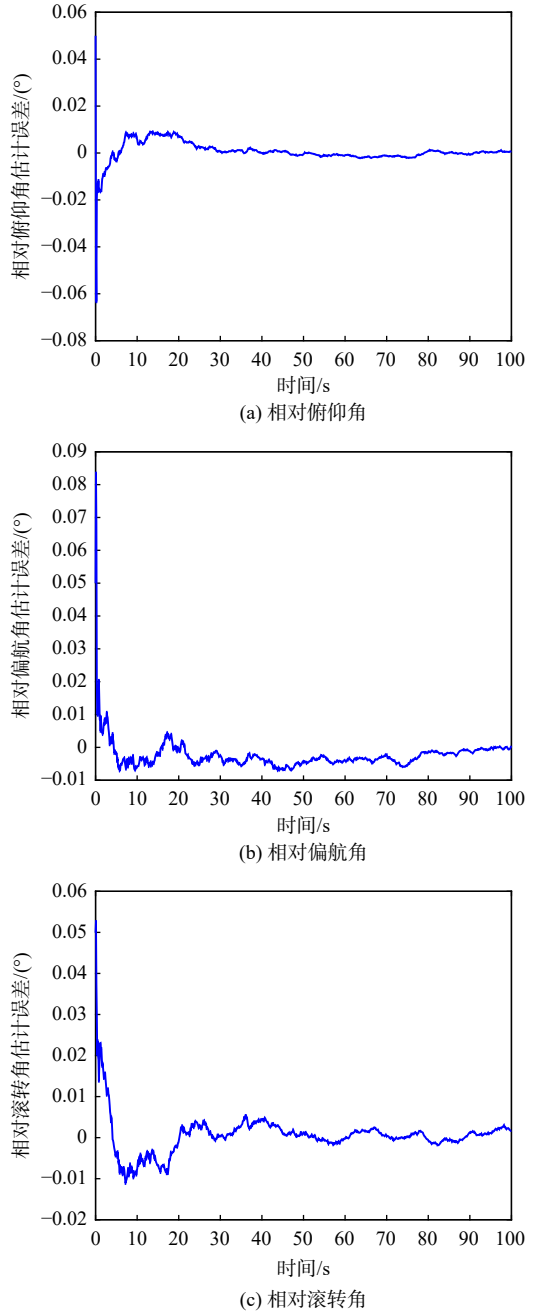


图 6 相对姿态估计误差

Fig. 6 Estimation error of relative attitude

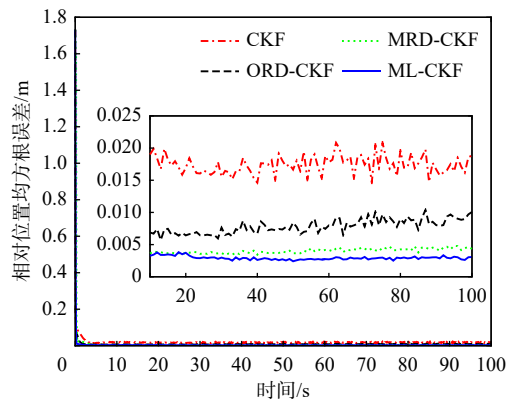


图 7 相对位置估计精度对比

Fig. 7 Comparison of estimation accuracies of relative positions

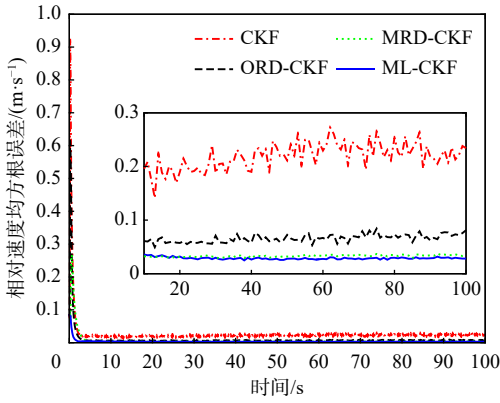


图8 相对速度估计精度对比

Fig. 8 Comparison of estimation accuracies of relative velocities

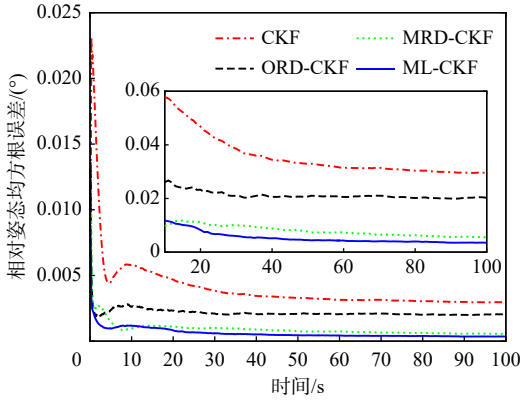


图9 相对姿态估计精度对比

Fig. 9 Comparison of estimation accuracies of relative attitudes

于ML-CKF中的加权因子是在延迟概率的基础上根据残差实时调节的。如表3所示4种滤波算法单次计算耗时,从表3中可见本文所提算法ML-CKF计算耗时低于MRD-CKF,高于CKF和ORD-CKF;但ML-CKF的估计精度远高于CKF和ORD-CKF;从估计精度和计算耗时综合考虑角度出发,在工程应用中ML-CKF是一种折中的选择。

表3 不同滤波算法的计算耗时

Table 3 Single computation time of different filtering algorithms

滤波算法	计算耗时/ms
CKF	0.60
ORD-CKF	1.88
MRD-CKF	12.72
ML-CKF	5.49

## 5 结论

1) 仿真结果表明,采用无约束三参数罗德里格斯参数表示姿态误差来传播和更新姿态,设计基于修正似然容积卡尔曼滤波的无人机编队,相对导航滤波器在视觉延迟量测存在多步随机延迟的情况

下,也能精确估计出相对位置、速度和姿态。

2) 通过蒙特卡罗打靶仿真统计多次仿真试验的均方根误差,基于ML-CKF的相对导航精度高于CKF、ORD-CKF和MRD-CKF,验证了所提算法的优越性。

3) 提出的基于ML-CKF的视觉辅助惯性相对导航算法还可应用于无人机空中加油、卫星编队相对导航和航天器交会对接等航空航天领域。

## 参考文献 (References)

- [1] ZHU Y F, SUN Y R, ZHAO W, et al. A novel relative navigation algorithm for formation flight[J]. *Journal of Aerospace Engineering*, 2020, 234(2): 308-318.
- [2] ELLINGSON G, BRINK K, MCLAIN T. Relative navigation of fixed-wing aircraft in GPS-denied environments[J]. *Navigation*, 2020, 67: 255-273.
- [3] 万九卿, 布少聪, 钟丽萍. 基于混合动态信念传播的多无人机协同定位算法[J]. *北京航空航天大学学报*, 2016, 42(5): 934-944.  
WAN J Q, BU S C, ZHONG L P. Cooperative localization algorithm of multi-UAVs based on dynamics hybrid belief propagation[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2016, 42(5): 934-944(in Chinese).
- [4] FOSBURY A M, CRASSIDIS J L. Relative navigation of air vehicles[J]. *Journal of Guidance, Control and Dynamics*, 2008, 31(4): 824-834.
- [5] JEONG J, KIM S, SUK J. Parametric study of sensor placement for vision-based relative navigation system of multiple spacecraft[J]. *Acta Astronautica*, 2017, 141: 36-49.
- [6] XU Z, QI N, CHEN Y. Parameter estimation of a three-axis spacecraft simulator using recursive least-squares approach with tracking differentiator and extended Kalman filter[J]. *Acta Astronautica*, 2015, 117: 254-262.
- [7] JULIER S J, UHLMANN J K, DURRANT-WHYTE H F. A new method for the nonlinear transformation of means and covariances in filters and estimators[J]. *IEEE Transaction Automatica Control*, 2000, 45(3): 77-82.
- [8] 卢道华, 付怀达, 王佳, 等. 基于IMU与UKF的船舶升沉运动信息测量方法[J]. *北京航空航天大学学报*, 2021, 47(7): 1323-1331.  
LU D H, FU H D, WANG J, et al. Measurement of ship's heave motion information based on IMU and UKF algorithm[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2021, 47(7): 1323-1331(in Chinese).
- [9] ARASARATNAM I, HAYKIN S. Cubature Kalman filters[J]. *IEEE Transaction Automatica Control*, 2009, 54(6): 1254-1269.
- [10] 李兆铭, 杨文革, 丁丹, 等. 多星对合作目标的分布式协同导航滤波算法[J]. *北京航空航天大学学报*, 2018, 44(3): 462-469.  
LI Z M, YANG W G, DING D, et al. Distributed coordinated navigation filtering algorithm for cooperative target by multi-satellite[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2018, 44(3): 462-469(in Chinese).
- [11] 崔乃刚, 王小刚, 郭继峰. 基于Sigma-point卡尔曼滤波的INS/Vision相对导航方法研究[J]. *宇航学报*, 2009, 30(6): 2220-2225.  
CUI N G, WANG X G, GUO J F. Reserch on relative navigation method based on INS/Vision using Sigma-point Kalman filter[J].

- Journal of Astronautics*, 2009, 30(6): 2220-2225(in Chinese).
- [12] PLETT G L, ZARZHITSKY D, PACK D J. Out-of-order sigma-point Kalman filtering for target localization using cooperating unmanned aerial vehicles[M]. Berlin: Advances in Cooperative Control and Optimization, 2007: 21-43.
- [13] KIM Y, HONG K, BANG H. Utilizing out-of-sequence measurement for ambiguous update in particle filtering[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2018, 54(1): 493-501.
- [14] HERMOSO-CARAZO A, LINARES-PEREZ J. Extended and unscented filtering algorithms using one-step randomly delayed observations[J]. *Applied Mathematics and Computation*, 2007, 190(2): 1375-1393.
- [15] HERMOSO-CARAZO A, LINARES-PEREZ J. Unscented filtering algorithm using two-step randomly delayed observations in nonlinear systems[J]. *Applied Mathematical Modelling*, 2009, 33: 3705-3717.
- [16] WANG X X, LIANG Y, PAN Q, et al. Gaussian filter for nonlinear systems with one-step randomly delayed measurement[J]. *Automatica*, 2013, 49: 976-986.
- [17] 张勇刚, 黄玉龙, 赵琳. 一种带多步随机延迟量测高斯滤波器的一般框架解[J]. *自动化学报*, 2015, 41(1): 122-135.
- ZHANG Y G, HUANG Y L, ZHAO L. A general framework solution to Gaussian filter with multiple-step randomly-delayed measurements[J]. *Acta Automatica Sinica*, 2015, 41(1): 122-135(in Chinese).
- [18] ESMZAD R, ESFANJANI M E. Modified likelihood Kalman filter for systems with incomplete, delayed and lost measurements[J]. *System and Control Letters*, 2018, 120: 23-28.
- [19] ESMZAD R, ESFANJANI M E. Bayesian filter for nonlinear systems with randomly delayed and lost measurements[J]. *Automatica*, 2019, 107: 36-42.
- [20] 张旭. 基于鲁棒自适应滤波的无人机编队相对导航方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2017: 56-58.
- ZHANG X. Research on relative navigation method of UAV formation based on robust adaptive filtering [D]. Harbin: Harbin Institute of Technology, 2017: 56-58(in Chinese).

## Relative navigation method based on modified likelihood filtering for unmanned aerial vehicle formation

SU Bingzhi<sup>1,\*</sup>, WANG Lei<sup>2</sup>, ZHANG Hongwei<sup>1</sup>, WANG Haihan<sup>1</sup>, SHI Lulu<sup>1</sup>

(1. China Helicopter Research and Development Institute, Tianjin 300300, China;

2. Aviation Military Representation Office of Army Armament Department in Tianjin Region, Tianjin 300384, China)

**Abstract:** A modified likelihood cubature Kalman filtering (ML-CKF) is proposed to solve the problem that the measurements of vision-based relative navigation sensor for unmanned aerial vehicle formation are randomly delayed by multiple steps. The measurement model is modified by the Bernoulli random variables to describe the random delay. The likelihood function of the filtering is calculated by marginalizing out the delay variable to extract accurate information from the delayed measurements. The third-degree spherical-radial rule is utilized to compute the Gaussian-weighted integrals for the nonlinear system. The proposed modified likelihood filtering has the property of adaptive filtering because the weighting factors of the filtering are tuned based on the characteristics of the received measurements. By utilizing the Rodrigues parameters to denote the attitude errors, the relative navigation filter of unmanned aerial vehicle formation is designed based on the ML-CKF. Simulation results indicate that the proposed filtering algorithm could accurately estimate the relative position, velocity and attitude between the leader and follower. Moreover, the estimation accuracy of ML-CKF is superior to cubature Kalman filtering and conventional randomly delayed filtering.

**Keywords:** unmanned aerial vehicle formation; relative navigation; likelihood function; cubature Kalman filtering; randomly delayed measurements

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0314

# 面向基准约束与余量约束的配准算法

朱燊<sup>1,2,3,\*</sup>, 肖世宏<sup>1,2,3</sup>, 陈志同<sup>4</sup>

(1. 中国航空制造技术研究院, 北京 100024; 2. 复杂构件数控加工工艺及装备北京市重点实验室, 北京 100024;

3. 数字化制造技术航空科技重点实验室, 北京 100024; 4. 北京航空航天大学 机械工程及自动化学院, 北京 100191)

**摘 要:** 为了保证待加工型面加工余量的均匀性及其相对于基准之间的位置关系, 提出一种同时基于基准约束与余量约束的配准算法。所提算法在配准模型中增加基准约束与余量约束, 建立在基准约束与余量约束下的配准模型, 引入局部坐标系进行配准计算, 并将计算结果反算至全局坐标系中。通过引入局部坐标系进行配准计算的算法, 降低了测量点变换过程中测量点的运动自由度, 从而降低了计算过程中变量的维度, 提高了计算效率, 保证了计算结果满足基准面约束与余量约束。所提算法在典型的飞机零件“襟翼滑轨”的简化模型上进行了应用验证, 相关结果表明: 在平面基准约束与余量约束下的配准算法, 计算时间为仅在余量约束下配准算法计算时间的 33.6%, 配准后基准面上测量点的最大偏差小于 0.04 mm, 待加工面上余量的波动小于 0.03 mm, 适用于该类零件的自适应精加工过程。

**关键词:** 数控加工; 自适应加工; 配准; 基准约束; 余量约束

**中图分类号:** TP391.7; V262.3

**文献标志码:** A **文章编号:** 1001-5965(2023)03-0580-08

飞机的襟翼滑轨是一类重要零件, 其加工工序繁杂、具有较大的尺寸与质量, 其导轨面与基准之间必须保持相应的位置关系。精加工前零件的基准面已加工到位、导轨面已完成粗加工, 因此, 该零件在数控精加工前必须经过反复地找正, 并根据零件的状态进行加工坐标系的调整<sup>[1]</sup>。

通过配准的方式来提高零件的找正效率、分配合适的加工余量是几何自适应加工技术的一个重要研究内容。在配准算法方面, 主要算法有迭代最近点 (iterative closest point, ICP) 算法、鲁棒点匹配 (robust point matching, RPM) 算法、核相关 (kernel correlation, KC) 算法、混合高斯模型 (Gaussian mixture models, GMM) 算法等<sup>[2]</sup>, 用于解决点集之间的配准问题。其中 ICP 算法是一种较为有效且快速的算法, 在 ICP 算法的基础上, Zhu 等<sup>[3]</sup>、Béarée 等<sup>[4]</sup>、Shi

等<sup>[5]</sup>、Bergström 等<sup>[6]</sup> 分别提出不同的方法对 ICP 算法的计算效率进行提升, Zhou 等<sup>[7]</sup>、Du 等<sup>[8]</sup> 针对 ICP 算法进行了改善, 可解决带有噪声的数据点的匹配问题。传统的 ICP 算法是一类在全局坐标系下计算旋转矩阵与平移矢量的配准算法。

在应用相关配准算法时, 需要根据工程背景增加相应的约束<sup>[9-10]</sup>。为了将配准算法与数控加工的应用相结合, 冯亚洲等<sup>[11]</sup> 提出基于公差约束的配准优化算法, Chen 等<sup>[12]</sup> 提出测量点向偏置面的配准算法。此外, 相关的配准算法还可以用于机器人加工<sup>[13]</sup>、航空零部件的检测<sup>[14]</sup>、零件修复<sup>[15]</sup>、整体叶盘的自适应铣削<sup>[16]</sup>、整体叶盘的自适应磨削<sup>[17]</sup> 等多个与数字化制造相关的方向。

配准技术通常用在零件的精加工工序中。在绝大多数的航空零部件的实际加工过程中, 精加工

收稿日期: 2021-06-08; 录用日期: 2021-08-27; 网络出版时间: 2021-09-16 15:01

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210916.0907.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210916.0907.001.html)

基金项目: 中国航空制造技术研究院基金 (KS912053115)

\* 通信作者. E-mail: [zhuyork@sina.com](mailto:zhuyork@sina.com)

**引用格式:** 朱燊, 肖世宏, 陈志同. 面向基准约束与余量约束的配准算法 [J]. 北京航空航天大学学报, 2023, 49 (3): 580-587.

ZHU Y, XIAO S H, CHEN Z T. A registration algorithm with datum constraints and allowance constraints [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 580-587 (in Chinese).

前的零件并不是毛坯状态。在绝大多数情况下, 零件待加工面的配准结果必须保证待加工面与基准面或基准孔的相对位置关系, 从而确保某些关键的形位尺寸。这种情况下的配准就需要保证几何体的变换方式是沿某个特定方向的平移或绕某个特定轴的旋转, 并且配准后的结果需要保证待加工面余量的波动较小。这就需要根据特定的矢量或平面建立局部坐标系进行旋转平移变换, 并将结果反算至全局坐标系中。本文即针对复杂零件的数控精加工过程, 提出面向基准约束与余量约束的配准算法, 开发相应的软件模块, 用于实现零件精加工中的自适应调整。

## 1 余量约束配准与运动变换关系

### 1.1 面向余量约束的配准算法

为了建立数学模型描述带余量约束的配准问题, 首先定义 2 种距离, 绝对距离与有向距离。检测点  $p_i$  到理论点  $q_i$  之间的绝对距离  $d_i$  如式 (1) 所示, 该距离为绝对坐标系下 2 点之间的范数:

$$d_i = \|p_i - q_i\| \quad (1)$$

检测点  $p_i$  到理论点  $q_i$  之间的有向距离  $d_i^{cs}$  如式 (2) 所示, 矢量  $n_i$  为理论点  $q_i$  的法矢, 法矢指向零件外侧。

$$d_i^{cs} = (p_i - q_i) \cdot n_i \quad (2)$$

配准的过程为: 通过旋转矩阵  $R$  和平移矢量  $T$  将检测点  $p_i$  变换至理论点  $q_i$  附近的过程。该旋转矩阵  $R$  和平移矢量  $T$  即为约束模型求解的结果。通过旋转矩阵与平移矢量的变换, 检测点  $p_i (i=1, 2, \dots, m)$  到理论点  $q_i$  之间的绝对距离  $d_i$  和有向距离  $d_i^{cs}$  为

$$\begin{cases} d_i = \|R \cdot p_i + T - q_i\| \\ d_i^{cs} = (R \cdot p_i + T - q_i) \cdot n_i \end{cases} \quad (3)$$

对于有余量约束下的配准问题, 可建立模型如式 (4) 所示。该优化模型表明, 通过旋转矩阵  $R$  和平移矢量  $T$  的变换, 在第  $i$  个检测点所对应的余量大于给定值  $\varepsilon_i$  的条件下, 检测点与其对应的理论点之间的偏差的平方和最小:

$$\begin{cases} \min \sum d_i^2 \\ \text{s.t. } d_i^{cs} \geq \varepsilon_i \end{cases} \quad (4)$$

### 1.2 局部坐标系与全局坐标系下的旋转、平移变换之间的数学关系

本文所提基于基准约束的配准算法最大的特点是不在全局坐标系 ( $O_{wcs}xyz$ ) 内进行配准, 而是通过坐标系的转换, 将点转换至局部坐标系 ( $Oe_1e_2e_3$ ) 下, 再进行配准, 最后, 根据局部坐标系与全局坐标系之间的关系计算出对应的全局坐标系下的旋转

矩阵  $R$  与平移矢量  $T$ 。因此, 必须建立全局坐标系下的旋转矩阵  $R^{wcs}$ 、平移矢量  $T^{wcs}$  与局部坐标系下的旋转矩阵  $R^{lcs}$ 、平移矢量  $T^{lcs}$  之间的关系。

如图 1 所示, 全局坐标系下局部坐标系的 3 个方向矢量为  $e_1$ 、 $e_2$ 、 $e_3$ , 坐标系原点为  $O$ , 令矩阵  $M=[e_1 \ e_2 \ e_3]$ , 任意一点  $P$  在局部坐标系内为  $P^{lcs}$ , 在全局坐标系内为  $P^{wcs}$ 。在局部坐标系内经过运动的点在全局坐标系内的位置如式 (5) 所示, 而在全局坐标系内经过运动的点的位置如式 (6) 所示。

$$P^{wcs} = M \cdot (R^{lcs} \cdot P^{lcs} + T^{lcs}) + O \quad (5)$$

$$P^{wcs} = R^{wcs} \cdot (M \cdot P^{lcs} + O) + T^{wcs} \quad (6)$$

若已知局部坐标系下的运动  $R^{lcs}$ 、 $T^{lcs}$ , 则全局坐标系下的运动  $R^{wcs}$ 、 $T^{wcs}$  如式 (7) 所示, 式 (7) 将用于算法中将局部坐标系下的运动关系转换至全局坐标系。

$$\begin{cases} R^{wcs} = M \cdot R^{lcs} \cdot M^{-1} \\ T^{wcs} = M \cdot T^{lcs} + O - R^{wcs} \cdot O \end{cases} \quad (7)$$

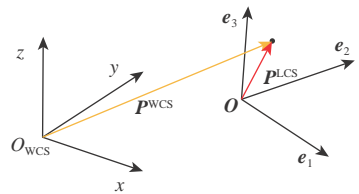


图 1 不同坐标系之间的关系

Fig. 1 Relationship between different coordinate

## 2 面向基准与余量约束的配准算法

在实际加工过程中, 精加工前的零件毛坯上通常已具有 1 个或 2 个基准, 在零件的自适应加工中, 必须保证待加工面与基准面的相对位置关系, 在此约束下的配准称为面向基准约束的配准。常用的基准为平面基准或圆柱面基准, 因此, 本节将结合式 (4) 的余量约束条件, 分别针对这 2 种基准约束建立相应的配准模型。

### 2.1 面向基准平面约束的配准算法

平面基准是最常见的一类基准, 在滑轨精加工之前, 其导轨面的下侧面即为基准平面。该基准面在粗加工阶段已完成加工, 且与零件内型面的位置直接相关。所提算法的基本思路是: 首先, 配准基准平面; 其次, 将所有点按平面配准后得到的旋转矩阵与平移矢量进行变换; 然后, 在基准平面上建立局部坐标系 (局部坐标系的  $x$  轴与  $y$  轴平行于基准平面); 再次, 将待加工面上的检测点在局部坐标系下进行配准, 配准过程中检测点仅能够沿局部坐标系的  $x$  轴与  $y$  轴平移与绕局部坐标系的  $z$  轴旋

转;最后,将局部坐标系下的配准结果转换至全局坐标系中。

式(8)为任意平面的数学表达,参数 $A$ 、 $B$ 、 $C$ 构成垂直于该平面的单位法矢:

$$Ax + By + Cz + D = 0 \quad (8)$$

对于式(8)所示的平面,若基准平面的检测点为 $\mathbf{p}_{Di}$ ( $p_{Di1}$ ,  $p_{Di2}$ ,  $p_{Di3}$ ),待加工面的检测点为 $\mathbf{p}_i$ ( $p_{i1}$ ,  $p_{i2}$ ,  $p_{i3}$ )。空间上的点 $\mathbf{p}_{Di}$ 可向空间内的任意平面求绝对距离如式(9)所示。通过式(10)求解点 $\mathbf{p}_{Di}$ 到平面的最近距离的最小二乘解,可解得检测点 $\mathbf{p}_{Di}$ 对应的平面为 $\Pi_{DP}$ : $A_1x + B_1y + C_1z + D_1 = 0$ ,相应的理论基准平面为 $\Pi$ : $A_0x + B_0y + C_0z + D_0 = 0$ 。

$$\mathbf{R}_G = \begin{pmatrix} \cos \varphi + (1 - \cos \varphi) \gamma_1^2 & (1 - \cos \varphi) \gamma_1 \gamma_2 + \gamma_3 \sin \varphi & (1 - \cos \varphi) \gamma_1 \gamma_3 - \gamma_2 \sin \varphi \\ (1 - \cos \varphi) \gamma_1 \gamma_2 + \gamma_3 \sin \varphi & \cos \varphi + (1 - \cos \varphi) \gamma_2^2 & (1 - \cos \varphi) \gamma_2 \gamma_3 + \gamma_1 \sin \varphi \\ (1 - \cos \varphi) \gamma_1 \gamma_3 - \gamma_2 \sin \varphi & (1 - \cos \varphi) \gamma_2 \gamma_3 + \gamma_1 \sin \varphi & \cos \varphi + (1 - \cos \varphi) \gamma_3^2 \end{pmatrix} \quad (13)$$

$$\mathbf{T}_G = \frac{|D_1 - D_0|}{\sqrt{A_0^2 + B_0^2 + C_0^2}} \begin{pmatrix} A_0 \\ B_0 \\ C_0 \end{pmatrix} \quad (14)$$

将待加工面的检测点 $\mathbf{p}_i$ 进行变换,得到点 $\mathbf{p}_i^1$ :

$$\mathbf{p}_i^1 = \mathbf{R}_G \mathbf{p}_i + \mathbf{T}_G \quad (15)$$

基准平面约束建立后,根据理论基准平面 $\Pi$ 建立局部坐标系LCS,将全部的检测点 $\mathbf{p}_i^1$ 及对应的理论点 $\mathbf{q}_i$ 变换至局部坐标系下,分别为 $\mathbf{p}_i^{1LCS}$ 与 $\mathbf{q}_i^{LCS}$ :

$$\begin{cases} \mathbf{p}_i^1 \rightarrow \mathbf{p}_i^{1LCS} \\ \mathbf{q}_i \rightarrow \mathbf{q}_i^{LCS} \end{cases} \quad (16)$$

局部坐标系内,在平面约束下的一般配准模型如式(17)所示,式(17)中矩阵 $\mathbf{R}^{LCS}$ 对应的旋转运动仅具有一个自由度,通常为绕局部坐标系的 $\mathbf{e}_3$ 轴的旋转运动,矢量 $\mathbf{T}^{LCS}$ 对应的运动为沿 $\mathbf{e}_1$ 轴与 $\mathbf{e}_2$ 轴的平移运动。

$$\min \|f(\mathbf{x})\| = \min \sum (\mathbf{R}^{LCS} \mathbf{p}_i^{1LCS} + \mathbf{T}^{LCS} - \mathbf{q}_i^{LCS})^2 \quad (17)$$

平面基准约束下的带余量的配准优化模型如式(18)所示,式(18)表示在平面基准的约束下,经过第一次变换测量点的平移方向需平行于基准平面,测量点的旋转轴平行于平面基准的法矢。

$$\begin{cases} \min \|f(\mathbf{x})\|^2 = \min \sum (\mathbf{R}^{LCS} \mathbf{p}_i^{1LCS} + \mathbf{T}^{LCS} - \mathbf{q}_i^{LCS})^2 \\ \text{s.t. } d_j^{cs} \geq \varepsilon_j, j \in [1, m] \end{cases} \quad (18)$$

式(18)计算完成后,通过式(7)求解全局坐标系下对应的旋转矩阵与平移矢量。

## 2.2 面向基准圆柱面约束的配准算法

基准圆柱面是指在实际零件的加工中存在着一个基准孔或基准轴的情况。为了保证加工后的

$$d_i = \frac{|Ap_{i1} + Bp_{i2} + Cp_{i3} + D|}{\sqrt{A^2 + B^2 + C^2}} \quad (9)$$

$$\min \sum d_i^2 = \min \sum \frac{(Ap_{i1} + Bp_{i2} + Cp_{i3} + D)^2}{A^2 + B^2 + C^2} \quad (10)$$

式(10)的优化模型可化简为式(11),式(11)中 $\mathbf{x}=(A, B, C, D)$ ,求解后可得最佳拟合的平面 $\Pi_{DP}$ 。

$$\min \|h(\mathbf{x})\|^2 = \min \sum (Ap_{i1} + Bp_{i2} + Cp_{i3} + D)^2 \quad (11)$$

令 $\mathbf{n}_1=(A_1, B_1, C_1)$ , $\mathbf{n}_0=(A_0, B_0, C_0)$ ,根据式(12)及Goldman矩阵,可得到平面 $\Pi_{DP}$ 向平面 $\Pi$ 变换的旋转矩阵 $\mathbf{R}_G$ 如式(13)所示,平移矢量 $\mathbf{T}_G$ 如式(14)所示。

$$\begin{cases} \boldsymbol{\gamma} = [\gamma_1, \gamma_2, \gamma_3]^T = \mathbf{n}_0 \times \mathbf{n}_1 \\ \boldsymbol{\varphi} = (\mathbf{n}_0, \mathbf{n}_1) \end{cases} \quad (12)$$

曲面相对于基准孔或基准轴的位置,需要在检测基准圆柱面的基础上保证已加工的基准圆柱面与理论基准圆柱面同轴,在此基础上再进行待加工区域的配准。空间内任意圆柱面为

$$\begin{cases} x(s, t) = x_0 + r \cos(t) \cos(\beta') + r \sin(t) \sin(\alpha') \sin(\beta') + s \cos(\alpha) \\ y(s, t) = y_0 + r \sin(t) \cos(\alpha') + s \cos(\beta) \\ z(s, t) = z_0 + r \sin(t) \sin(\alpha') \cos(\beta') - r \cos(t) \sin(\beta') + s \cos(\gamma) \\ 0 \leq t \leq 2\pi, 0 \leq s \leq h \end{cases} \quad (19)$$

式中: $x$ 、 $y$ 、 $z$ 为空间圆柱面上点的位置; $t$ 与 $s$ 为圆柱面上点的自变量; $\alpha'$ 、 $\beta'$ 、 $\gamma'$ 为圆柱体中轴所对应的空间角; $\alpha$ 、 $\beta$ 、 $\gamma$ 为对应的计算角,通过 $\alpha'$ 、 $\beta'$ 、 $\gamma'$ 进行计算得到; $\mathbf{p}_{CO}=(x_0, y_0, z_0)$ 为圆柱面底面圆心; $r$ 为圆柱面半径<sup>[18]</sup>。

基准圆柱面的检测点为 $\mathbf{p}_{Ci}$ ( $p_{Ci1}$ ,  $p_{Ci2}$ ,  $p_{Ci3}$ ),其至圆柱面的绝对距离为

$$d_i = |(\|\mathbf{p}_{Ci} - \lambda \mathbf{e}\| - r)| \quad (20)$$

式中: $\mathbf{e}=(\cos(\alpha'), \cos(\beta'), \cos(\gamma'))$ 为圆柱面中心轴对应的单位矢量; $\lambda = \|(\mathbf{p}_{Ci} - \mathbf{p}_{CO}) \times \mathbf{e}\|$ 。

建立最小二乘模型如式(21)所示,求解后可得实际圆柱面基准的中心轴矢量 $\mathbf{e}$ 与实际的圆柱面底面圆心 $\mathbf{p}_{CO}$ 。若理论圆柱面基准的中心轴矢量为 $\mathbf{e}_0$ ,令式(12)中 $\mathbf{n}_1=\mathbf{e}$ 、 $\mathbf{n}_0=\mathbf{e}_0$ ,则通过式(12)与式(13)可求解对应的变换矩阵 $\mathbf{R}_G$ 。若理论圆柱面底面圆心为 $\mathbf{p}_{COL}$ ,则 $\mathbf{T}_G=(\mathbf{p}_{COL} - \mathbf{R}_G \mathbf{p}_{CO})$ 。

$$\min \sum d_i^2 = \min \sum (\|\mathbf{p}_i - \lambda \mathbf{e}\| - r)^2 \quad (21)$$

同样利用式(15)将待加工面的检测点 $\mathbf{p}_i$ 进行变换,并得到点 $\mathbf{p}_i^1$ ,该变换在全局坐标系下进行。

随后, 以圆柱基准中心轴矢量  $e_0$  与  $p_{COL}$  点为基础建立局部坐标系 LCS, 使用式 (16) 将全部第 1 次变换后的检测点  $p_i^1$  及对应的理论点  $q_i$  变换至局部坐标系下。

式 (18) 可用于表示基准圆柱面约束下带余量的配准优化模型。此时, 式 (18) 中矩阵  $R^{LCS}$  对应的旋转运动同样仅具有一个自由度, 通常为绕局部坐标系的  $e_3$  轴的旋转运动(此时  $e_3$  与圆柱基准中心轴矢量  $e_0$  相同), 矢量  $T^{LCS}$  对应的运动则改变为沿  $e_3$  轴的平移运动。式 (18) 变换模型说明, 在圆柱面基准的约束下, 经过变换的测量点  $p_i^{LCS}$  的平移方向与圆柱面中心轴平行, 测量点可绕圆柱面中心轴进行转动, 配准后的毛坯姿态需满足各测量点对应的加工余量大于限定值。求解式 (18) 后, 通过式 (7) 求解对应的全局坐标系下的旋转矩阵与平移矢量。

### 2.3 面向基准平面约束与基准圆柱面约束的配准算法

在某些零件的配准过程中存在着同时使用基准面与基准孔进行定位的情况, 这种情况必须保证基准平面与基准孔互相垂直。基于基准平面约束与基准圆柱面约束的旋转矩阵  $R_G$  与平移矢量  $T_G$  的求解流程如图 2 所示。

首先, 输入基准平面与基准圆柱面的测量点与理论曲面; 随后, 根据基准平面的测量点拟合基准平面, 并根据式 (13) 与式 (14) 求解将拟合基准平面变换至理论基准平面的旋转矩阵  $R_{G1}$  与平移矢量

$T_{G1}$ ; 根据式 (15) 将基准圆柱面的测量点进行变换, 并拟合基准圆柱面; 根据拟合基准圆柱面的中心轴方向与拟合基准平面的法矢方向, 判断基准平面与基准圆柱面是否满足垂直度要求, 若不满足则判断此基准圆柱面需要在加工后进行修整, 若垂直度满足要求则根据基准圆柱面的底面圆心的理论值与实测值求解平移矢量  $T_{G2}$  及其投影矢量  $T_{G3}$ ; 随后输出旋转矩阵  $R_G$  与平移矢量  $T_G$  用于变换待加工面上的检测点  $p_i$ 。本过程始终在全局坐标系下进行数据处理, 处理过程中可对零件的状态进行过程检测, 避免关键尺寸在终检中出现不合格再返修的情况。

同样利用式 (15) 将待加工面的检测点  $p_i$  进行变换, 得到点  $p_i^1$ 。以理论圆柱基准中心轴与理论平面基准  $\Pi$  建立局部坐标系 LCS, 使用式 (16) 将全部的检测点  $p_i^1$  及对应的理论点  $q_i$  变换至局部坐标系下。

在平面基准与圆柱面基准的双重约束下带余量的配准模型如式 (22) 所示, 矩阵  $R^{LCS}$  对应的旋转运动仅具有一个自由度, 通常为绕局部坐标系  $e_3$  轴的旋转运动, 无平移运动。

$$\begin{cases} \min \|f(x)\|^2 = \min \sum (R^{LCS} p_i^{LCS} - q_i^{LCS})^2 \\ \text{s.t. } d_j^{cs} \geq \varepsilon_j, j \in [1, m] \end{cases} \quad (22)$$

式 (22) 计算完成后, 通过式 (7) 求解全局坐标系下的旋转矩阵与平移矢量。

### 2.4 面向基准约束的配准模型求解算法

通过在机测量技术得到基准面上的检测点与待加工面上的检测点后, 进行面向基准约束配准模型的求解流程如图 3 所示。首先, 将基准面上的检

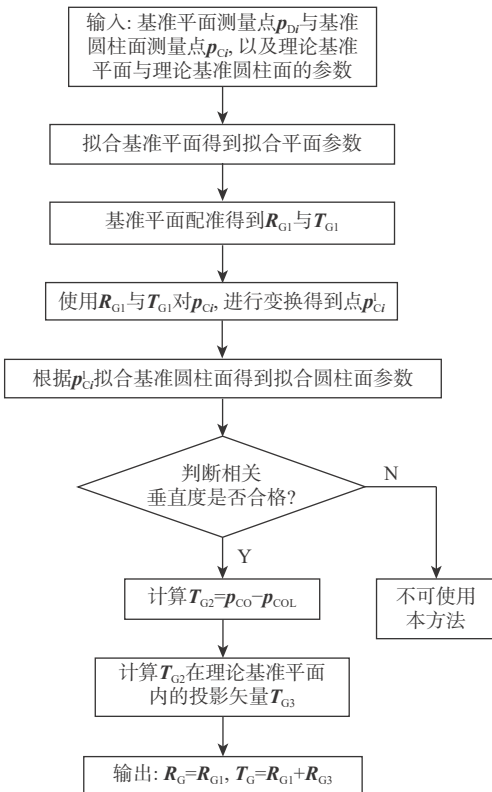


图 2 基于基准约束下的变换矩阵与平移矢量求解流程  
Fig. 2 Process of solving transformation matrix and translation vector based on datum constraints

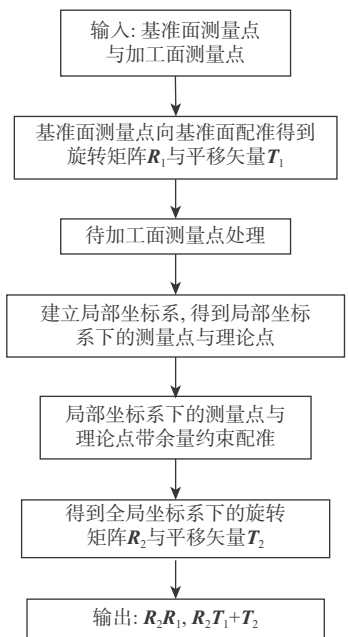


图 3 面向基准约束的配准模型求解流程  
Fig. 3 Process of solving registration model based on datum constraints

测点向基准面进行配准,得到矩阵  $R_G$  与矢量  $T_G$ ,并建立相应的局部坐标系;然后,对待加工面上的点进行处理,将实测点与理论点变换至局部坐标系下;根据零件的余量情况进行带余量、少自由度的配准;最后,将局部坐标系下计算得到的旋转矩阵与平移矢量进行处理,得到全局坐标系下的旋转矩阵  $R_2$  与平移矢量  $T_2$ ,输出使用的旋转矩阵  $(R_2R_1)$  与平移矢量  $(R_2T_1+T_2)$ 。通过基准约束降低了求解配准模型过程中变量的数量,可以保证配准后的零件基准面满足工艺需求。

### 3 算例分析

通过 vs2010 对本文所提算法进行开发,并在此基础上针对某典型零件“襟翼滑轨”的简化模型进行计算仿真,该简化模型如图 4 所示。

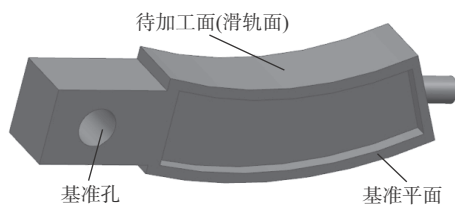


图 4 具有典型特征的零件模型

Fig. 4 Design model with typical features

该零件在精加工阶段最重要的加工工序是加工襟翼滑轨的导轨面,即图 4 中标识的待加工面。待加工面在精加工前具有一定的加工余量,在完成加工后必须与图 4 中的基准平面、基准孔(即基准圆柱面)之间保持较高的位置精度,且精加工前基准平面已完成加工。

为了测试本文所提算法,可如图 5 所示规划测量点,测量点分布在基准平面、基准圆柱面与待加工面上。使用基准平面作为约束,进行基于平面基准约束与余量约束的配准,解得旋转矩阵与平移矢量后将实测点向理论点进行变换,得到变换后的结果如图 6 所示。对应点之间的误差如图 7 所示,由图 7 可知,在平面基准的约束下,可保证平面上测量点误差较小且平均,并保持待加工面具有较为平均的余量,但基准孔存在较大的偏差,具体数据如表 1 所示。其中,方法 1 为基准平面约束与余量约束下的配准结果,方法 2 为基准平面、基准孔与余量约束下的配准结果,方法 3 为仅余量约束下的配准结果。所提算法适用于基准面加工合格、基

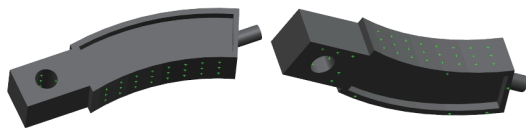


图 5 零件表面检测点分布

Fig. 5 Measuring points on surfaces of part

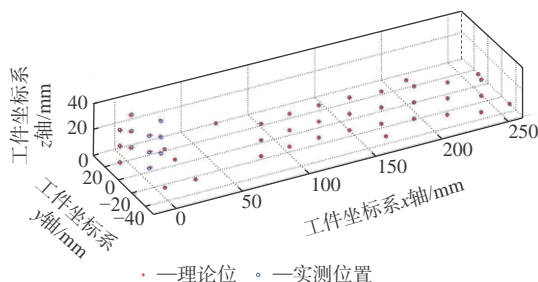


图 6 在基准平面约束下的配准结果

Fig. 6 Registration result of method based on constraint of a plane

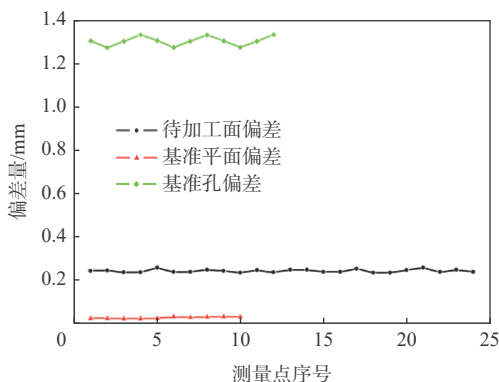


图 7 在基准平面约束下配准后偏差曲线

Fig. 7 Deviations of different points after applying registration method based on constraint of a plane

准孔留有足够余量的毛坯精加工,可在加工完待加工面后重制基准孔,从而保证曲面相对于基准的位置度。

使用基准平面与基准圆柱面作为约束,并在余量约束下进行加工面的配准,得到变换后的结果如图 8 所示,相关点之间的误差如图 9 所示。由图 9 可知,在基准平面与基准孔的双重约束下,可保证基准平面与基准孔上的测量点误差较小且平均,并保持待加工面具有较为平均的余量,待加工面上余量的波动小于 0.03 mm,具体数据如表 1 所示。所提算法适用于基准面与基准孔加工精度都合格的

表 1 不同方法配准结果比较

Table 1 Comparison of different registration method

方法	基准平面最大偏差/mm	基准孔最大偏差/mm	待加工面最小偏差/mm	待加工面最大偏差/mm	偏差值波动量/mm	计算时间/s
方法 1	0.038 436	1.337 247	0.241 043	0.263 811	0.022 768	0.146
方法 2	0.033 438	0.054 354	0.186 074	0.219 661	0.033 587	0.254
方法 3	0.149 557	1.320 907	0.213 938	0.247 241	0.033 303	0.434

毛坯精加工, 保证了滑轨面加工后相对于 2 个基准的位置度。

在无基准约束的情况下, 仅通过余量约束进行加工面的配准, 得到变换后的结果如图 10 所示, 相关点之间的误差如图 11 所示。由图 10 可知, 在无基准约束的情况下, 待加工面虽可以保持较为平均的余量分布, 但基准位置已出现偏差较大或偏差波动较大的情况, 具体数据如表 1 所示。无基准约束的情况下, 为了保证滑轨面加工后相对于基准的位置度, 需在待加工面加工完成后重制基准面与基准孔。

由表 1 可知本文所提算法对保证零件待加工面

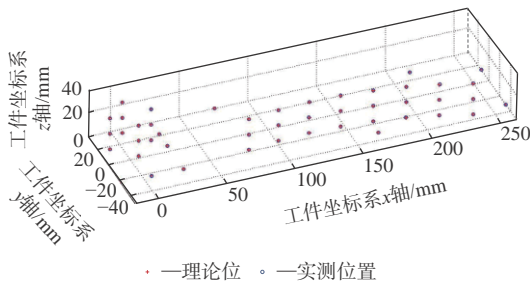


图 8 在基准平面与基准圆柱面约束下的配准结果

Fig. 8 Registration result of method based on constraint of a plane and a cylinder

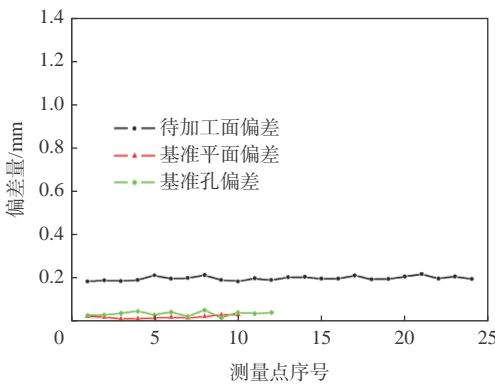


图 9 在基准平面约束与基准孔约束配准后偏差曲线

Fig. 9 Deviations of different points after applying registration method based on constraint of a plane and a cylinder

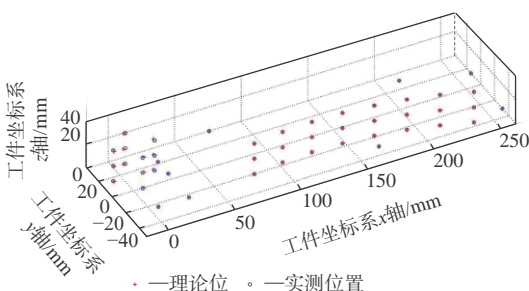


图 10 余量约束下配准结果

Fig. 10 Registration result of method based on allowance constraints

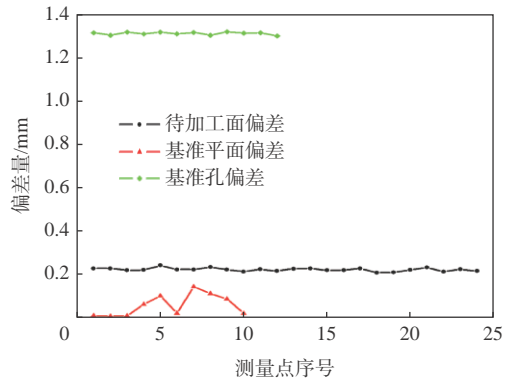


图 11 仅余量约束下配准后偏差曲线

Fig. 11 Deviations of different points after applying registration method based on allowance constraints

的余量分布及加工后待加工面的位置精度有重要的意义。3 种方法的计算时间说明选用带基准约束的配准算法计算时间要短于无基准约束条件下的计算时间, 方法 1 的计算时间是方法 3 计算时间的 33.6%, 偏差值波动量也更小。

### 4 结 论

1) 本文提出一种基于基准约束与余量约束的几何自适应配准算法, 进行了相应的软件开发, 并通过某滑轨零件的简化模型进行验证。

2) 本文所提算法分别在全局坐标系和局部坐标系内配准基准面和待加工面, 可有效地保证待加工面相对于已加工基准的位置精度, 提高产品加工的合格率。

3) 对于该襟翼滑轨零件来说, 当同时使用基准平面约束与余量约束的配准方法时, 得到的待加工面的余量最大、余量波动最小。因此, 对于襟翼滑轨零件来说, 基准平面加工完成、基准孔留有余量的毛坯最适合用于精加工。

4) 使用基准平面约束与余量约束的配准方法时, 程序的计算时间最短, 其计算时间是仅基于余量约束配准方法计算时间的 33.6%。

### 参考文献 (References)

[1] 王文理, 袁士平. 自适应加工技术在数控加工领域的分类与应用[J]. *航空制造技术*, 2013(6): 26-29.  
WAGN W L, YUAN S P. Category and application of adaptive machining technology in NC machining field[J]. *Aeronautical Manufacturing Technology*, 2013(6): 26-29(in Chinese).

[2] MAISELI B, GU Y, GAO H. Recent developments and trends in point set registration methods[J]. *Journal of Visual Communication and Image Representation*, 2017, 46: 95-106.

[3] ZHU L, BARHAK J, SRIVATSAN V, et al. Efficient registration for precision inspection of free-form surface[J]. *International Journal of Advanced Manufacturing Technology*, 2007,

- 32(5): 505-515.
- [4] BÉARÉE R, DIEULOT J Y, RABATÉ P. An innovative subdivision-ICP registration method for tool-path correction applied to deformed aircraft parts machining[J]. *International Journal of Advanced Manufacturing Technology*, 2011, 53(5): 463-471.
- [5] SHI X Y, PENG J J, LI J P, et al. The iterative closest point registration algorithm based on the normal distribution transformation[J]. *Procedia Computer Science*, 2019, 147: 181-190.
- [6] BERGSTRÖM P, EDLUND O, SÖDERKVIST I. Repeated surface registration for on-line use[J]. *International Journal of Advanced Manufacturing Technology*, 2011, 54(5): 677-689.
- [7] ZHOU Y Q, XU S Y, JIN C C, et al. Multiple point sets registration based on expectation maximization algorithm[J]. *Computers and Electrical Engineering*, 2018, 70: 1-11.
- [8] DU S Y, XU G L, ZHANG S R, et al. Robust rigid registration algorithm based on pointwise correspondence and correntropy[J]. *Pattern Recognition Letters*, 2020, 132: 91-98.
- [9] 郑德华, 岳东杰, 岳建平. 基于几何特征约束的建筑物点云配准算法[J]. *测绘学报*, 2018, 31(4): 464-468.  
ZHENG D H, YUE D J, YUE J P. Geometric feature constraint based algorithm for building scanning point cloud registration[J]. *Acta Geodaetica et Cartog Raphica Sinica*, 2018, 31(4): 464-468(in Chinese).
- [10] 俞慈君, 李江雄, 余锋杰, 等. 带工程约束的点匹配算法[J]. *机械工程学报*, 2010, 46(5): 183-190.  
YU C J, LI J X, YU F J, et al. 3D points registration algorithm with engineering constraints[J]. *Journal of Mechanical Engineering*, 2010, 46(5): 183-190(in Chinese).
- [11] 冯亚洲, 任军学, 梁永收, 等. 多目标约束的精锻叶片几何重构优化算法[J]. *航空学报*, 2018, 39(7): 421844.  
FENG Y Z, REN J X, LIANG Y S, et al. Multi-objective optimization algorithm for geometric reconstruction of precision forged blade[J]. *Acta Aeronautica et Astronautica Sinica*, 2018, 39(7): 421844(in Chinese).
- [12] CHEN S Y, WU C C, XUE S, et al. Fast registration of 3D point clouds with offset surfaces in precision grinding of free-form surfaces[J]. *The International Journal of Advanced Manufacturing Technology*, 2018, 97(9): 3595-3606.
- [13] XIE S Q, CHENG D, WONG S. Three-dimensional object recognition system for enhancing the intelligence of a KUKA robot[J]. *The International Journal of Advanced Manufacturing Technology*, 2008, 38(7): 822-839.
- [14] HE W T, LI Z W, GUO Y Y, et al. A robust and accurate automated registration method for turbine blade precision metrology[J]. *The International Journal of Advanced Manufacturing Technology*, 2018, 97(9): 3711-3721.
- [15] FENG C, LIANG J, GONG C Y, et al. Repair volume extraction method for damaged parts in remanufacturing repair[J]. *The International Journal of Advanced Manufacturing Technology*, 2018, 98(5): 1523-1536.
- [16] 张定华, 张莹, 吴宝海, 等. 自适应加工技术在整体叶盘制造中的应用[J]. *航空制造技术*, 2013(13): 51-55.  
ZHANG D H, ZHANG Y, WU B H, et al. Application of adaptive machining technology in blisk manufacturing[J]. *Aeronautical Manufacturing Technology*, 2013(13): 51-55(in Chinese).
- [17] 陈志同, 朱熡, 张云, 等. 整体叶盘超硬磨料砂轮数控磨削加工技术[J]. *航空制造技术*, 2018, 61(19): 64-72.  
CHEN Z T, ZHU Y, ZHANG Y, et al. CNC grinding technology with super-abrasive grinding wheels for blisk[J]. *Aeronautical Manufacturing Technology*, 2018, 61(19): 64-72(in Chinese).
- [18] 王兴波. 空间任意方向圆柱面的参数方程[J]. *佛山科学技术学院院报(自然科学版)*, 2010, 28(6): 58-63.  
WANG X B. Parametric equation of cylinder in arbitrary direction[J]. *Journal of Foshan University(Natural Science Edition)*, 2010, 28(6): 58-63(in Chinese).

## A registration algorithm with datum constraints and allowance constraints

ZHU Yu<sup>1,2,3,\*</sup>, XIAO Shihong<sup>1,2,3</sup>, CHEN Zhitong<sup>4</sup>

(1. AVIC Manufacturing Technology Institute, Beijing 100024, China;

2. Beijing Key Laboratory of NC Machining Technology and Equipment of Aero-Component, Beijing 100024, China;

3. Aeronautical Key Laboratory for Digital Manufacturing Technology, Beijing 100024, China;

4. School of Mechanical Engineering and Automation, Beihang University, Beijing 100191, China)

**Abstract:** This paper presents a registration algorithm based on datum constraints and allowance constraints, which allows the surface to be machined with a uniform allowance and to maintain accurate position relations. First, a local coordinate system was introduced to compute the registration results. Then a conversion method would be employed to solve the corresponding results in the global coordinate system. The introduction of the local coordinate system reduced the degree of freedom of the point set during registration, thus decreasing the number of dimensions and improving the computational efficiency while the algorithm was solved. The results can satisfy both the datum constraints and the allowance constraints. The algorithm was applied on a simplified model of a typical aircraft part “flap track”. The computational results showed that the computing time used by this method was only 33.6 percent of that used by the registration method with the allowance constraints alone, the maximal error for measure points was less than 0.04 mm, and the fluctuating of the allowance was less than 0.03 mm. Therefore, the method is suitable for finishing the surface of similar aircraft parts adaptively.

**Keywords:** computer numerical control machining; adaptive machining; registration; datum constraint; allowance constraint

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0310

# 直反信号协同的 GNSS-R BSAR 距离多普勒成像算法

吴世玉, 杨东凯, 王峰\*, 苗铎

(北京航空航天大学 电子信息工程学院, 北京 100191)

**摘 要:** 针对目前基于全球导航卫星系统反射信号的双基地合成孔径雷达 (GNSS-R BSAR) 在一站固定模式下的大斜视, 斜距历程复杂, 回波信号方位空变导致回波信号难以处理的问题, 提出改进的距离多普勒成像新算法。所提算法采用 GNSS 信号作为辐射源, 根据一站固定模式下 GNSS-R BSAR 合成孔径时间长特点, 引入高阶等效斜视距离模型, 得到导航卫星与目标斜距相对时间变化的精确描述。先通过直射信号与回波信号时域对消进行距离徙动校正, 实现全场景目标距离徙动的精确校正; 再通过方位向分块混合相关处理来克服回波信号方位向的移变性质, 实现全场景高效精确成像。所提算法的成像效率优于传统后向投影时域 (BP) 算法, 成像精度与 BP 算法相当, 且可根据需要通过调整方位分块的宽度来提升聚焦效果。最后, 用 GPS-L5 信号进行仿真和实验, 仿真和实验结果验证了所提算法的可行性和高效性。

**关键词:** 全球导航卫星系统反射信号; 双基地合成孔径雷达; 方位空变处理; 距离多普勒算法; 距离徙动校正

**中图分类号:** V19; TN958.97

**文献标志码:** A **文章编号:** 1001-5965(2023)03-0588-09

基于全球导航卫星系统反射信号的双基地合成孔径雷达 (global navigation satellite system-reflectometry bistatic synthetic aperture radar, GNSS-R BSAR) 以全球导航卫星系统 (global navigation satellite system, GNSS) 卫星作为非合作辐射源, 使用合成孔径雷达技术获取地球表面图像<sup>[1-2]</sup>, 在遥感应用<sup>[3-6]</sup>、目标探测<sup>[7-8]</sup>、地物识别及地表形变监测<sup>[9-10]</sup>中变得越来越重要。与 GNSS-R 的遥感应用相比, GNSS-R BSAR 可以对地貌及地物纹理特征有很好的展现, 因此, 可以提取到更加丰富的特征信息提高 GNSS-R 遥感反演的精度。

与传统合成孔径雷达 (synthetic aperture radar, SAR) 相比, GNSS-R BSAR 具有明显优势。首先, 由于无需额外发射机, 使得该技术的设备重量轻、功耗低、成本低; 其次, 140 多颗在轨导航卫星使得地球任何区域可同时接收超过 20 颗卫星的信号,

为不同角度的多辐射源融合提供了可行性, 可获取观测区域更加丰富的地表特征信息; 最后, GNSS 可以提供精确的授时服务, 保证了该技术的同步性能。

一站固定模式<sup>[11]</sup>是 GNSS-R BSAR 的主要工作模式之一, 通过将接收机安装在地面某高点, 或悬停的无人机和空间浮空器上, 对固定区域进行合成孔径成像, 实施长时间观测。尽管后向投影时域 (back projection, BP) 算法可处理任意几何构型且满足长时间合成孔径<sup>[1,12]</sup>的需求, 但 BP 算法计算量大, 限制了其被广泛应用。相比于 BP 算法, 距离多普勒 (range-Doppler, RD) 算法<sup>[13]</sup>和线性调频变标 (chirp scaling, CS) 算法<sup>[14-15]</sup>等频域算法可提高成像效率, 但为获得高方位向分辨率, 一站固定式通常需几百秒甚至上千秒的合成孔径, 回波信号在方位向上不再满足位移不变性, 因此文献<sup>[13-15]</sup>频域算法在长时间合成孔径情况下会产生散焦现象。

收稿日期: 2021-06-08; 录用日期: 2021-06-21; 网络出版时间: 2021-08-03 09:30

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210802.1445.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210802.1445.001.html)

基金项目: 国家自然科学基金 (41774028); 中国博士后科学基金 (BX20200039)

\*通信作者. E-mail: [wangfeng\\_buaa@buaa.edu.cn](mailto:wangfeng_buaa@buaa.edu.cn)

**引用格式:** 吴世玉, 杨东凯, 王峰, 等. 直反信号协同的 GNSS-R BSAR 距离多普勒成像算法 [J]. 北京航空航天大学学报, 2023, 49 (3): 588-596. WU S Y, YANG D K, WANG F, et al. GNSS-R BSAR range-Doppler imaging algorithm based on synchronization of direct and echo signal [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 588-596 (in Chinese).

为了实现一站固定模式下长时间合成孔径 GNSS-R BSAR 图像的聚焦, 提出一种改进的 RD 成像算法, 在文献 [16] 提出的高阶等效斜距模型的基础上, 利用直达信号的伪码和载波相位对回波信号的距离徙动和相位误差进行补偿, 提高距离徙动校正精度, 并采用方位向分块混合相关处理方法进行方位向相位补偿, 解决方位向移变导致随方位向位置变化的多普勒调频率问题。

## 1 GNSS-R BSAR 信号建模

### 1.1 高阶等效斜视距离模型

导航卫星的椭圆运动轨迹导致导航卫星与目标的精确斜距表达式极其复杂。通常采用二阶等效距离模型、等效斜视距离模型<sup>[17]</sup>等近似描述卫星与目标的相对运动。二阶等效距离模型是将卫星与目标斜距表达式在多普勒中心时刻进行泰勒展开并保留至二次项。即

$$R(t) = R_0 - \lambda f_d t - \frac{\lambda}{2} f_t t^2 \quad (1)$$

式中:  $t$  为方位时间, 在雷达领域也称慢时间;  $R_0$  为多普勒中心时刻卫星与目标的斜距;  $\lambda$  为载波的波长;  $f_d$  为卫星相对目标的多普勒中心频率;  $f_t$  为卫星相对目标的多普勒调频率。等效斜视距离模型将卫星的开普勒椭圆轨道在局部近似为匀速直线运动, 等效斜视距离模型的表达式为

$$R(t) = \sqrt{R_0^2 + v^2 t^2 + 2R_0 v t \cos \varphi} \quad (2)$$

式中:  $v$  为卫星的等效地面速度;  $\varphi$  为卫星的等效斜视角。

尽管理论上等效斜视距离模型较二阶等效距离模型更加精确, 但随着合成孔径时间增加, 等效斜视距离模型的斜距误差将超过斜距误差阈值, 导致方位向散焦, 无法满足 GNSS-R BSAR 在一站固定模式下长时间的合成孔径需求。为更精确描述导航卫星的运动, 本文采用文献 [16] 提出的高阶等效斜视距离模型计算 GNSS-R BSAR 的斜距。该模型表示为

$$R(t) = \sqrt{R_0^2 + \left(vt + \frac{At^2}{2}\right)^2 + 2R_0 \left(vt + \frac{At^2}{2}\right) \cos \varphi} \quad (3)$$

式中:  $A$  为导航卫星的等效加速度。如图 1 与图 2 所示为等效斜视距离模型和本文采用的高阶等效斜视距离模型的斜距补偿误差随合成孔径时间的关系。由图 1 可知等效斜距模型的有效合成孔径时长约为 50 s, 本文采用模型极大提高了斜距精度, 如图 2 所示, 且合成孔径时长为 873 s(约 14.5 min)

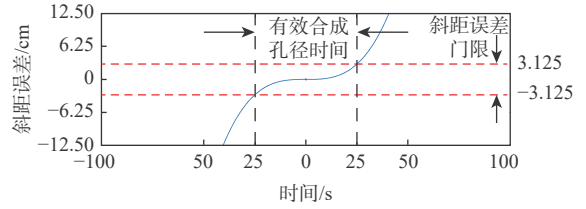


图 1 等效斜视模型-斜距误差仿真

Fig. 1 Equivalent squint range model-slope range error simulation

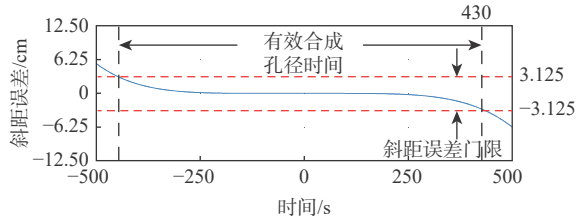


图 2 高阶等效斜视距离模型-斜距误差仿真

Fig. 2 Improved equivalent squint range model-slope range error simulation

以内精度仍满足方位向聚焦要求。

### 1.2 GNSS-R BSAR 信号模型

如图 3 所示 GNSS-R BSAR 的一站固定模式几何构型, 其中  $R_B(t)$  为接收机到卫星的瞬时斜距,  $R_T(t)$  为目标到导航卫星的瞬时斜距,  $R_R$  为目标中心到接收机的距离。总斜距为  $R(t) = R_T(t) + R_R$ 。GNSS-R BSAR 的接收系统包含 1 个直达通道和 1 个回波通道。直达通道通过右旋圆极化(right-hand circular polarization, RHCP)全向型天线接收 GNSS 直射信号, 并获得准确的载波相位、码相位和定位信息, 为反射通道的信号同步提供精确的参考信息。直射信号经正交解调和 SAR 数据格式化后的二维形式<sup>[18]</sup>可表示为

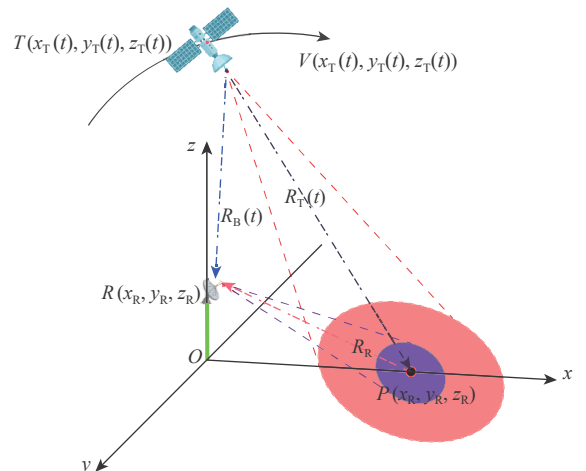


图 3 GNSS-R BSAR 一站固定模式的几何构型

Fig. 3 GNSS-R BSAR one station fixed pattern geometric configuration

$$s_a(\eta, \tau) = W(\eta)N(\eta)C \left[ \tau - \frac{R_B(\eta)}{c} \right] \exp \left[ -j2\pi f_c \frac{R_B(\eta)}{c} \right] \quad (4)$$

式中： $\eta$ 为方位向慢速时间； $\tau$ 为距离向快速时间； $W(\eta)$ 为接收信号在方位向的矩形包络； $N(\eta)$ 为导航电文； $C(\tau)$ 为伪随机码； $f_c$ 为发射信号的载波中心频率； $c$ 为电磁波传播速度。回波通道通过高增益左旋圆极化(left-hand circular polarization, LHCP)天线接收目标区域反射的回波信号，经剥离导航电文，且省略对成像算法不产生影响的回波信号的幅度和初始相位后可表示为

$$s_r(\eta, \tau) = W(\eta)C \left[ \tau - \frac{R(\eta)}{c} \right] \exp \left[ -j2\pi f_c \frac{R(\eta)}{c} \right] \quad (5)$$

## 2 直反信号协同的RD成像算法

基于高阶等效斜视模型，本文提出一种如图4所示的直反信号协同的一站固定模式下GNSS-R BSAR距离多普勒成像算法。首先，对直射信号进行捕获、跟踪、定位获得导航卫星的位置、速度、载波相位及码相位；然后，通过检测直射信号的多普勒相位提取接收机本振频率漂移等引起的相位误差，并将提取出的相位误差在反射通道信号成像的过程中予以补偿，提高成像质量。其中，IFFT为反傅里叶变换，FFT为傅里叶变换。

### 2.1 基于直反协同的距离向压缩

与传统SAR不同，GNSS-R BSAR距离压缩通过回波信号和伪随机码自相关实现。以周期为1 ms和码速率为10.23 MHz的GPS-L5信号伪随机码为例，其自相关函数的峰值旁瓣比(peak side lobe ratio, PSLR)高于35 dB，且在高斯白噪声条件下，自相关增益为40 dB，使得GNSS接收机可从噪声中检测微弱的导航信号。经距离向自相关后的回波信号为

$$s(\eta, \tau) = W(\eta)P \left[ \tau - \frac{R(\eta)}{c} \right] \exp \left[ -j2\pi f_c \frac{R(\eta)}{c} \right] \quad (6)$$

式中： $P(\tau)$ 为伪随机码的自相关函数。

由于一站固定模式下GNSS-R BSAR等效斜视角和导航卫星的多普勒中心频率通常较大，因此，回波信号距离徙动轨迹具有很大的线性项(距离走动)和二次项(距离弯曲)。如果直接进行方位向傅里叶变换，则由于距离向与方位向耦合使得距离向压缩的信号在距离向发生散焦。在传统单站SAR中，通常引入二次距离压缩因子进行距离徙动校正将距离向信号重新聚焦。在GNSS-R BSAR中，由于导航信号调制了无显式表达式的伪随机码，而非线性调频信号，因此，距离向与方位向耦合的形式与传统单站合成孔径雷达不同，二次距离压缩滤波器的解析式难以获得。本文在方位时域以直达天线的相位中心为参考，通过直射信号的距离徙动信息，对回波场景中的距离徙动对消来进行距离徙动的校正。可以根据卫星位置、目标区域、接收机3点位置信息计算导航卫星直射信号与回波信号传输的时延差 $T_{\text{delay}}$ ，通过补偿该时延差，可得到导航卫星同一时刻发出的直射信号及反射信号。将距离徙动校正后，只剩下目标到接收机的固定距离项，及微小的残差项。此时再进行方位向傅里叶变换，由于距离向与方位向耦合引起的距离向信号散焦已经消除。经距离徙动校正后的回波信号表达式为

$$s(\eta, \tau) = W(\eta)P \left[ \tau - \frac{R_T(\eta) + R_R - R_B(\eta)}{c} \right] \cdot \exp \left[ -j2\pi \frac{R_T(\eta) + R_R}{\lambda} \right] \quad (7)$$

其中，直达信号的距离徙动可表示为

$$R_B(\eta) = \sqrt{R_{B_0}^2 + \left( v\eta + \frac{A\eta^2}{2} \right)^2} + 2R_{B_0} \left( v\eta + \frac{A\eta^2}{2} \right) \cos \varphi_1 \quad (8)$$

式中： $\varphi_1$ 为卫星到直射信号天线相位中心的等效斜视角。回波信号的距离徙动为

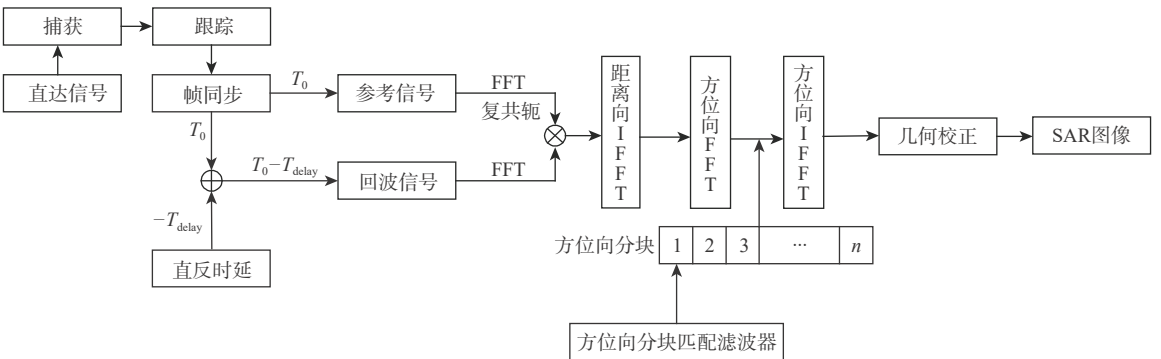


图4 所提算法流程

Fig. 4 Flowchart of the proposed algorithm

$$R_T(\eta) = \sqrt{R_{T_0}^2 + \left(v\eta + \frac{A\eta^2}{2}\right)^2} + 2R_{T_0} \left(v\eta + \frac{A\eta^2}{2}\right) \cos\varphi_2 \quad (9)$$

式中:  $\varphi_2$  为卫星到成像区域中心的等效斜视角。直射信号的距离徙动与回波信号的距离徙动的残差项表达式为

$$\Delta R(\eta) = R_T(\eta) - R_B(\eta) \quad (10)$$

由式 (10) 可知, 残差项主要取决于卫星、场景中心、回波天线相位中心的位置及卫星速度。若残差值的变化在距离向中不超过一个距离门, 则不会影响到成像结果。为验证所提算法的边界条件, 进行场景参数如表 1 所示的仿真分析。仿真中选取真实卫星星历数据, 并在笛卡儿坐标系原点架设接收机, 目标点设置在  $x$  轴上, 目标点间隔为 1 m, 沿  $x$  轴的正向遍历。

如图 5 所示, 随着合成孔径时间及接收机与目标点距离的增加, 残差项误差变大。当回波天线相位中心与目标点距离在 5 km 内时, 距离徙动校正方法满足长时间合成孔径处理; 当回波天线相位中心与目标距离较远时, 合成孔径时间 60 s 内也可满足距离向压缩要求。

由于 GNSS-R BSAR 一站固定式的天线相位与目标点距离通常小于 5 km, 因此可满足距离徙动校正需求。经距离徙动校正后的回波形式为

表 1 残差项仿真参数

Table 1 Residual simulation parameters

坐标	卫星位置/km	卫星速度/(m·s <sup>-1</sup> )	接收机位置/m	目标点/m
$x$	11 769	500.96	0	1~5 000
$y$	1 124.8	2 891.8	0	0
$z$	12 482	-369.24	1 000	0

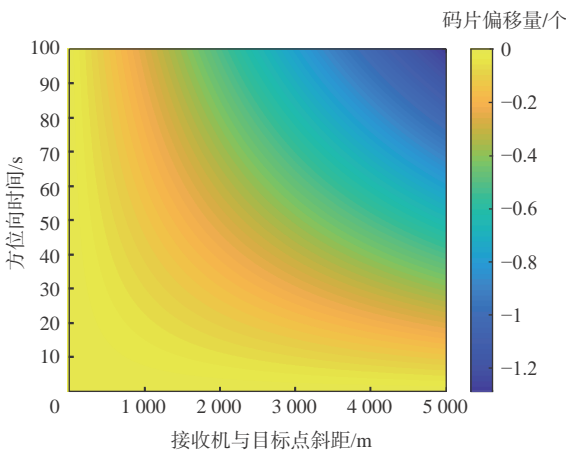


图 5 残差与斜距及合成孔径时间的关系

Fig. 5 Relationship between residual error and slant range and synthetic aperture time

$$s(\eta, \tau) = W(\eta)P \left[ \tau - \frac{R_R - \Delta R(\eta)}{c} \right] \exp \left[ -j2\pi \frac{R_T(\eta) + R_R}{\lambda} \right] \quad (11)$$

校正后的信号通过方位向傅里叶变换至距离多普勒域。在距离多普勒域距离徙动经过一致校正后的信号形式为

$$s(f, \tau) = W[\eta(f)]P(\tau) \exp \left[ -j2\pi \frac{R(f)}{\lambda} \right] \quad (12)$$

式中:  $W[\eta(f)]$  为距离多普勒域的回波信号方位向包络。

### 2.2 基于分块混合相关的方位向压缩

由于一站固定模式下 GNSS-R BSAR 的合成孔径时间过长, 高阶等效斜视模型中存在的等效加速度, 导致随方位向时间变化的多普勒调频率, 因此, 回波信号已不再具有方位移不变的性质, 同一距离门的信号无法再通过相同的方位向匹配滤波器进行一致压缩。本文采用了方位向分块混合相关的处理方法, 动态更新匹配滤波器的调频率参数, 解决回波信号方位向移动问题。每一子块的回波信号使用子块中心位置的调频率参数作为子块整体混合匹配滤波器的调频率使得子块中心的方位向相位得到完全补偿。为解决边缘点部分产生的散焦, 在处理效率和处理精度之间做一个权衡, 设定子块边缘点方位相位误差上限为

$$\text{error}_{\max} \leq 2\pi \frac{4v + \lambda f_a \sin\varphi_1}{8\lambda v R_{T_0}} \cdot \frac{l}{2} A \left( \frac{T_p}{2} \right)^2 \quad (13)$$

式中:  $l$  为方位向子块的宽度;  $T_p$  为合成孔径的长度。通过式 (13) 根据所需图像聚焦的质量和效率选择分块宽度和合成孔径长度。混合相关匹配滤波器为

$$h(f, \tau) = p(\tau) \exp \left( -j2\pi \frac{R(f)}{\lambda} \right) \quad (14)$$

经方位向傅里叶变换的信号在距离多普勒域的聚焦为

$$s(\eta, \tau) = s_a(B_a\eta)p(\tau) \quad (15)$$

式中:  $s_a(\eta) = \sin\eta/\eta$ ;  $B_a$  为回波信号的方位向带宽。式 (15) 聚焦为矩形窗的傅里叶变换, 为方位向剖面形式。最终通过几何校正得到 GNSS-R BSAR 的雷达图像。

## 3 仿真分析

通过成像仿真进行所提算法有效性的验证。仿真选用信号为 GPS PRN03 卫星发射的 GPS-L5 信号, 参数设置如表 2 和表 3 所示, 场景布置如图 6 所示。方位向混合相关处理时子块的宽度设置为 2 km。

表2 仿真参数

Table 2 Simulation parameters

参数	距离向 采样频率/ MHz	载波 频率/ MHz	信号 带宽/ MHz	成像区域 大小/ (km×km)	合成 孔径 时间/s	脉冲 重复 频率/Hz
数值	62	1 176.45	20.46	6.5×6.5	300	1 000

表3 场景参数

Table 3 Scene parameters

坐标	接收机位置/ m	场景中心位置/ km	卫星位置/ km	卫星速度/ (m·s <sup>-1</sup> )
x	0	12.5	20 133.725 8	1 392.706 8
y	0	0	10 697.303 2	-2 766.685 6
z	100	0	728.029 1	138.306 3

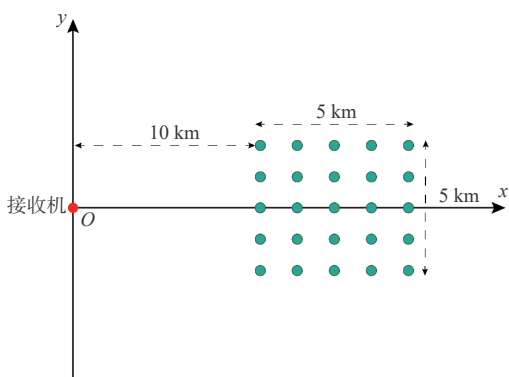


图6 场景点目标分布

Fig. 6 Scene point target distribution map

如图7所示为所提算法对仿真场景的成像结果。场景中的25个点目标都被聚焦到了正确的位置。如图8和表4所示为点目标13、25的距离向和方位向成像横截面及其评估结果。点目标13位于方位向子块中心点,其方位向和距离横截面与表4所示的理论结果一致,在距离向和方位向均未出现失真。点目标25位于方位向子块的边缘位置,尽管该处目标有最大的残差相位,方位向横截面出现成像的最大程度失真,但从方位向横截面上可看到仅为一个微小失真,主要体现在主瓣展宽,距离向峰值旁瓣比和积分旁瓣比均出现0.20 dB失真,方位向的PSLR和积分旁瓣比(ISLR)有0.21 dB的失真。对称的左右旁瓣说明该点残差的线性相位已被补偿掉,残差的相位误差主要是高阶项。

## 4 实验验证

为进一步验证所提算法的有效性,本文开展了如图9所示基于GPS-L5信号一站固定模式成像实验。实验场景在如图10和表5所示的北京航空航天大学体育场东边的建筑物群。GNSS-R BSAR成像数据采集系统同时采集直射和目标反射的GPS-

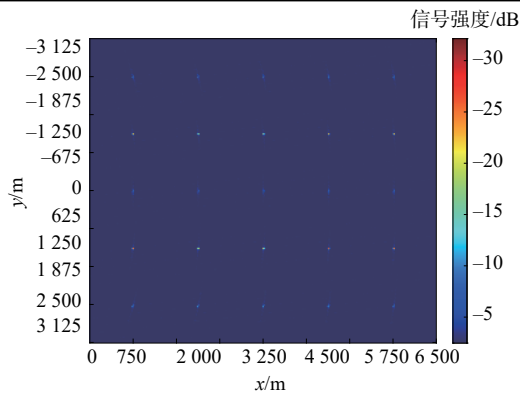


图7 所提算法的成像结果

Fig. 7 The proposed algorithm imaging results

L5信号,其中直达信号用增益为3 dB的全向右旋圆极化天线接收,用于实现导航信号的捕获跟踪定位及与回波信号的同步;目标反射信号用增益为19 dB的左旋圆极化天线接收,用于目标成像。具体的系统参数及成像参数如表6所示。

由于GPS PRN03卫星与地面固定接收机构成较好的双基构型,可使GNSS-R BSAR图像具有较好的距离向分辨率,因此,本次实验选用GPS PRN03卫星信号作为辐射源。合成孔径中心时刻GPS PRN03卫星在站心坐标系下的位置和速度信息如表7所示。

对采集数据用所提算法和传统BP算法进行合成孔径时长为1 800 s的成像。成像结果如图11与图12所示,可看出2种算法获得图像非常相似,说明2种算法的成像能力相似。尽管经过长时间的合成孔径,GNSS-R BSAR的方位向分辨率达到了亚米级,但是GPS L5信号的带宽为20.46 MHz,其准单站模式的距离向分辨率约为15 m,在双站模式下,受几何构型的影响,比准单站的距离向分辨率更差。因此,现有的GNSS-R BSAR图像无法达到光学图像一样的直观解析能力,需要进行雷达图像的解释。为了表征其分辨能力这里将获得雷达图像与光学图像进行了基于位置信息的匹配,结果如图13所示。由于回波天线与目标场景的高度近似相同,且GPS PRN03卫星和接收机都位于目标区域的西侧,导致场景中建筑物主要的回波信号来自于西部边缘部分。因此,在图13中雷达图像的强散射回波都聚焦到了建筑物的西侧。图10中的9个建筑物都被聚焦到了雷达图像正确的位置,链球围栏回波是由线的旁瓣接收,因此在图像中聚焦的能量较弱。其他建筑物为天线的主瓣接收均显示了良好的聚焦效果。

为了进一步评估所提算法获得成像质量和计算复杂度,将所提算法与BP算法进行了对比分析。本文分别对2种算法获得的图像进行了交叉

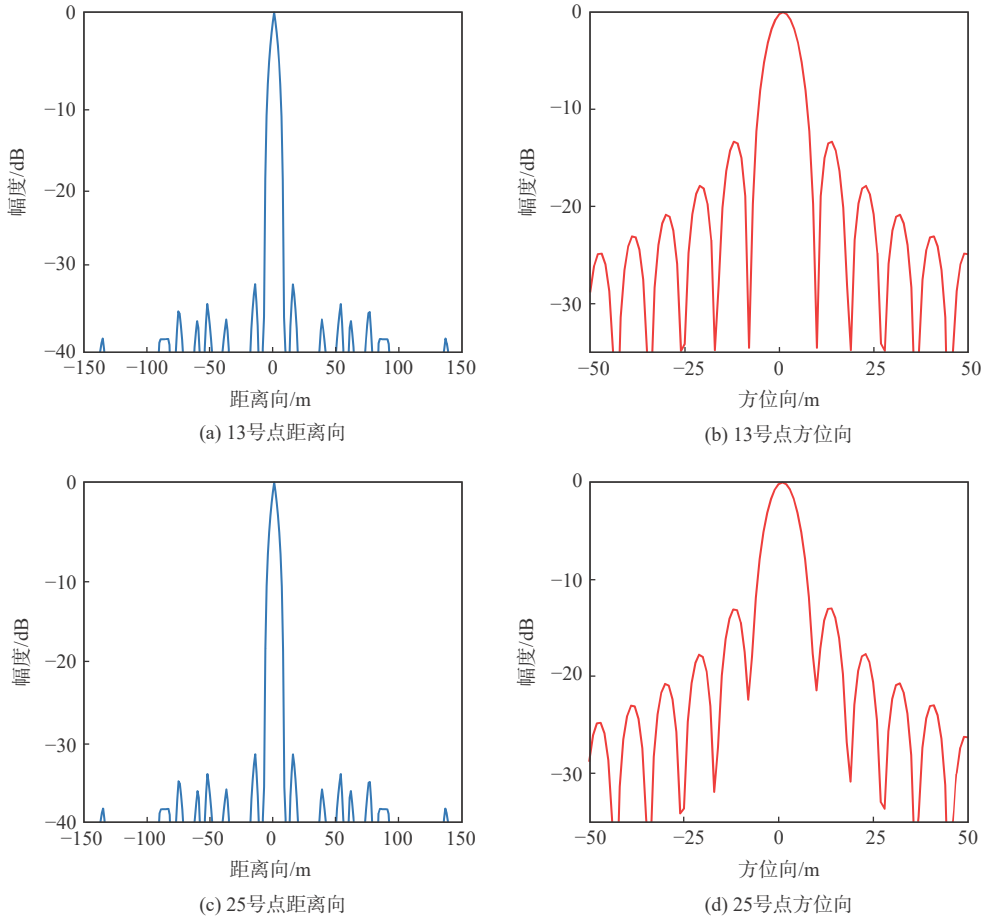


图 8 13 号与 25 号点目标仿真横截面分析

Fig. 8 Cross-section analysis of target simulation at No. 13 and No. 25

表 4 所选点目标的评估参数

Table 4 Evaluation parameters of selected point target

参数	距离向			方位向		
	PSLR/dB	ISLR/dB	分辨率/m	PSLR/dB	ISLR/dB	分辨率/m
目标13	-34.8	-12.8	16.8	-13.3	-9.94	5.625
目标25	-34.8	-12.6	16.8	-13.11	-9.76	5.626
理论值	-35	-12.8	16.3	-13.3	-9.95	5.625

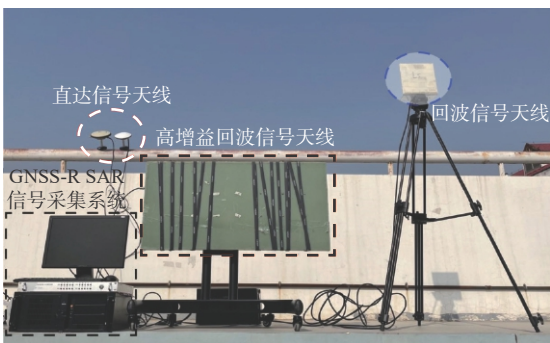


图 9 GNSS-R BSAR 数据采集系统

Fig. 9 GNSS-R BSAR data collection system

无论在距离向还是方位向的聚焦效果都是相当的。以体育馆边沿处沿距离向进行了测量分辨率约为 16.8 m, 与理论值一致。在 1 800 s 的合成孔径下, 方位向分辨率达到了亚米级, 实验中无法直接对方位向的分辨率进行测量, 这里通过测量体育馆的物理长度进行评估, 由图 15 可以看出, 方位向的测量值为 40 m 与光学测量值一致。实验结果说明了所提算法的成像性能与时域算法成像性能相当。

本文所提算法的整个成像过程执行了 4 次距离向 FFT, 2 次方位向 FFT, 1 次复数乘法, 整个所提算法的计算复杂度可以表示为

$$O(RD) = 4N_r N_a \log_2 N_r + 2N_r N_a + N_r N_a \quad (16)$$

式中:  $N_r$  和  $N_a$  分别为距离向和方位向的采样点数。BP 算法的距离向压缩过程中需要执行 4 次距离向 FFT 和 1 次复数乘法, 其计算量为  $4N_r N_a \log_2 N_r$ , 后向投影的过程中计算量可以表示为  $N_i N_x N_y N_a$ , 其中,  $N_x$  和  $N_y$  为成像区域的 2 个维度的像素点个数,  $N_i$  为像素点间隔。整个 BP 算法的计算复杂度可以表示为

$$O(BP) = 4N_r N_a \log_2 N_r + N_r N_a + N_i N_x N_y N_a \quad (17)$$

由式(17)可以看出, 随着成像区域和合成孔径时间的增加, BP 算法的计算量会急剧增加。在本

横截面的分析。选择了以体育馆西部边沿部分进行剖析, 因为该区域的回波信号能量较强且连续性较好。如图 14 和图 15 所示, 所提算法和 BP 算法

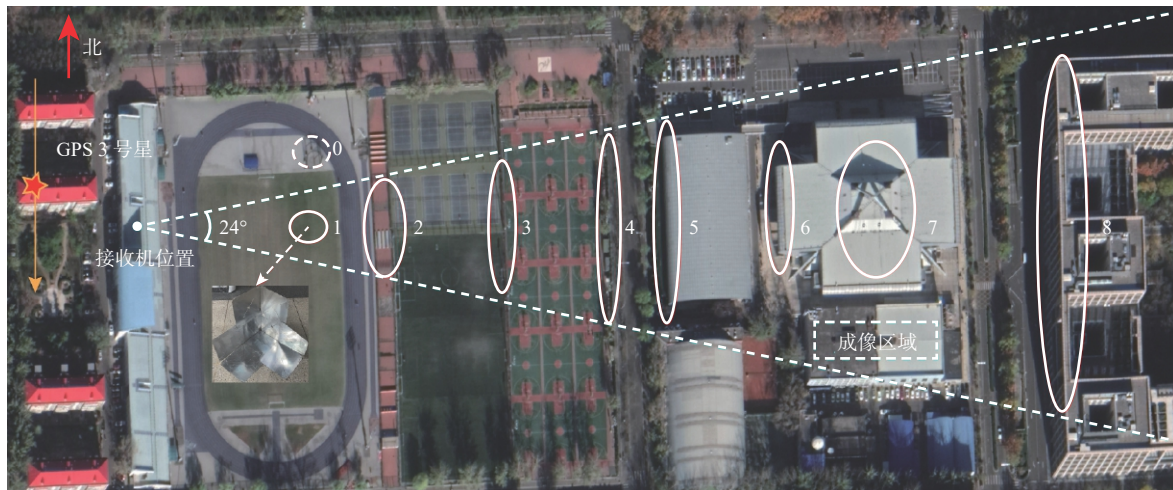


图 10 北京航空航天大学体育场周边的光学图像(谷歌地图)

Fig. 10 Optical image around Beihang University stadium (Google map)

表 5 实验场景主要回波目标

Table 5 Main echo target of experimental scene

编号	建筑物
目标0	链球围栏
目标1	角反射器
目标2	两道铁栅栏
目标3	篮球场铁栅栏
目标4	篮球场铁栅栏
目标5	游泳馆
目标6	体育馆
目标7	体育馆顶部
目标8	新主楼

表 6 数据采集系统及成像参数

Table 6 Data acquisition system and imaging parameters

参数	数值
采样频率/MHz	62
量化比特/bit	14
载频/MHz	1 176.45
信号带宽/MHz	20.46
合成孔径时间/s	1 800
成像区域大小/(m×m)	600×600
脉冲重复频率/Hz	1 000
回波天线海拔高度/m	60.52

表 7 GPS PRN03 卫星的位置和速度信息

Table 7 GPS PRN03 satellite position and speed information

坐标	卫星位置/km	速度/(m·s <sup>-1</sup> )
x	20 133.725 8	1 392.706 8
y	10 697.303 2	-2 766.685 6
z	728.029 1	138.306 3

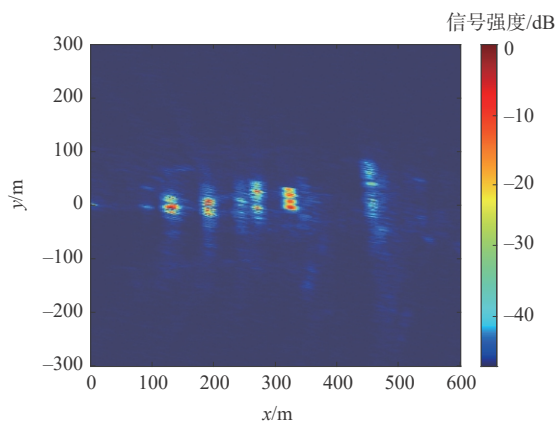


图 11 BP 算法成像结果

Fig. 11 BP algorithm imaging results

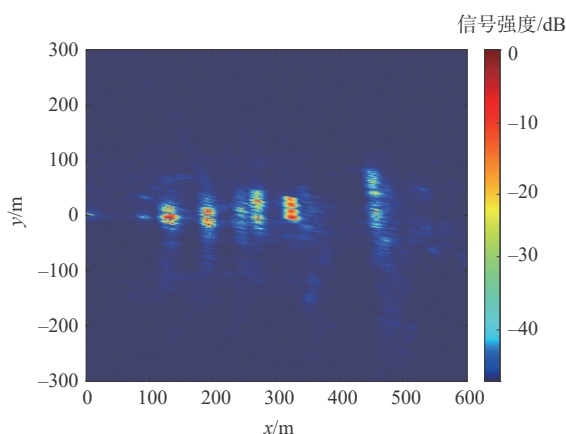


图 12 所提算法成像结果

Fig. 12 The proposed algorithm imaging results

次实验中,相同的计算平台(处理器为 AMD-3800 X@3.8 GHz, 8 核心 16 线程,内存为 64 GHz),所提算法的计算时间为 350 s 左右, BP 算法的计算时间为 19 880 s 左右,计算效率提升了 56.8 倍。实验结果验证了所提算法的可行性和高效性。

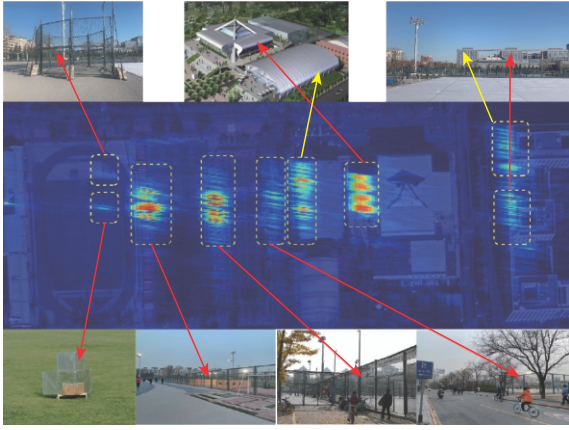


图 13 所提算法成像结果光学匹配图

Fig. 13 The proposed algorithm imaging result optical matching map

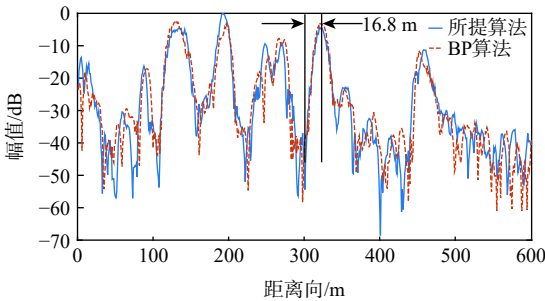


图 14 成像结果交叉横截面的分析-距离向剖面

Fig. 14 Analysis of cross-section of imaging results-range profile

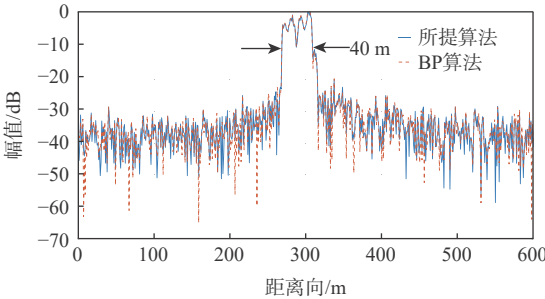


图 15 成像结果交叉横截面的分析-方位向剖面

Fig. 15 Analysis of cross-section for imaging results-azimuth profile

### 5 结论

1) 针对 GNSS-R BSAR 的一站固定模式, 提出了一种改进的距离多普勒算法。首先, 通过引入改进的等效斜视距离模型, 改善了长时间合成孔径导致的方位向散焦的问题; 其次, 利用导航信号良好的同步性能, 通过直射与回波信号时域对消的方式, 完成了距离徙动校正; 最后, 通过方位向分块混合相关的算法, 提高了方位向的聚焦精度。

2) 基于系统空间分辨率的考量, 在仿真与实验

验证中采用 GPS PRN03 卫星发射的 GPS-L5 信号。仿真结果表明, 所提算法在成像质量和效率方面均表现良好。在 1 800 s 合成孔径的实测中, 距离向的分辨率为 16.8 m, 方位向分辨率达到了亚米级, 该分辨率可以满足大部分遥感测量的需求。所提算法与传统的 BP 成像算法相比, 在一站固定模式下成像质量相当, 在效率方面远优于传统的 BP 算法。

### 参考文献 (References)

[ 1 ] ANTONIOU M, CHERNIAKOV M. GNSS-based bistatic SAR: A signal processing view[J]. EURASIP Journal on Advances in Signal Processing, 2013(1): 1-16.

[ 2 ] ANTONIOU M, STOVE A, SAYIN A, et al. Passive SAR satellite constellation for near-persistent earth observation: Prospects and issues[J]. IEEE Aerospace and Electronic Systems Magazine, 2018, 33: 4-15.

[ 3 ] WANG F, YANG D, YANG L. Feasibility of wind direction observation using low-altitude global navigation satellite system-reflectometry[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(12): 5063-5075.

[ 4 ] MARTIN-NEIRA M. A passive reflectometry and interferometry system (PARIS): Application to ocean altimetry[J]. ESA Journal, 1993, 17: 331-355.

[ 5 ] RODRIGUEZ-ALVAREZ N, BOSCH-LLUIS X, CAMPS A, et al. Soil moisture retrieval using GNSS-R techniques: Experimental results over a bare soil field[J]. IEEE Transactions on Geoscience and Remote Sensing, 2009, 47(11): 3616-3624.

[ 6 ] 洪学宝, 张波, 阮宏梁, 等. 基于相关功率修正的地基GNSS-R土壤湿度反演[J]. 北京航空航天大学学报, 2021, 47(8): 1558-1564. HONG X B, ZHANG B, RUAN H L, et al. Ground-based GNSS-R soil moisture retrieval based on correlation power correction[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(8): 1558-1564(in Chinese).

[ 7 ] ZHOU X, WANG P, CHEN J, et al. A modified radon fourier transform for GNSS-based bistatic radar target detection[J]. IEEE Geoscience and Remote Sensing Letters, 2020(99): 1-5.

[ 8 ] GAO C, YANG D, HONG X, et al. Experimental results about traffic flow detection by using GPS reflected signals[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(12): 5076-5087.

[ 9 ] LIU F, FAN X, ZHANG T, et al. GNSS-Based SAR interferometry for 3-D deformation retrieval: Algorithms and feasibility study[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(10): 5736-5748.

[ 10 ] LIU F, ANTONIOU M, ZENG Z, et al. Coherent change detection using passive GNSS-Based BSAR: Experimental proof of concept[J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 51(8): 4544-4555.

[ 11 ] 仇晓兰, 丁赤飏, 胡东辉. 双站SAR成像处理技术[M]. 北京: 科学技术出版社, 2010: 147-148. QIU X L, DING C B, HU D H. Double-station SAR imaging processing technology[M]. Beijing: Science and Technology Press,

- 2010: 147-148 (in chinese) .
- [12] SHAO Y F, WANG R, DENG Y K, et al. Fast backprojection algorithm for bistatic SAR imaging[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, 10(5): 1080-1084.
- [13] ANTONIOU M, SAINI R, CHERNIAKOV M. Results of a space-surface bistatic SAR image formation algorithm[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2007, 45(11): 3359-3371.
- [14] ZENG T, WANG R, LI F, et al. A modified nonlinear chirp scaling algorithm for spaceborne/stationary bistatic SAR based on series reversion[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, 51(5): 3108-3118.
- [15] ZHOU X K, CHEN J, WANG P B, et al. An efficient imaging algorithm for GNSS-R bi-static SAR[J]. *Remote Sensing*, 2019, 11(24): 2945.
- [16] WANG P, LIU W, CHEN J, et al. A high-order imaging algorithm for high-resolution spaceborne SAR based on a modified equivalent squint range model[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(3): 1225-1235.
- [17] 黄岩, 李春升, 陈杰, 等. 高分辨星载SAR改进ChirpScaling成像算法[J]. *电子学报*, 2000, 28(3): 35-38.
- HUANG Y, LI C S, CHEN J, et al. Improved ChirpScaling imaging algorithm for high resolution spaceborne SAR[J]. *Acta Electronics*, 2000, 28(3): 35-38(in Chinese).
- [18] ZENG H C, WANG P B, CHEN J, et al. A novel general imaging formation algorithm for GNSS-based bistatic SAR[J]. *Sensors*, 2016, 16(3): s16030294.

## GNSS-R BSAR range-Doppler imaging algorithm based on synchronization of direct and echo signal

WU Shiyu, YANG Dongkai, WANG Feng\*, MIAO Duo

(School of Electronic and Information Engineering, Beihang University, Beijing 100191, China)

**Abstract:** Aiming at the problems that the current global navigation satellite system-reflectometry bistatic synthetic aperture radar (GNSS-R BSAR) has a large squint in the fixed mode of one station, the slant range history is complicated, and the echo signal's azimuth is changed, the echo signal is difficult to process, an improved Range Doppler imaging algorithm is proposed. The method uses GNSS signal as the radiation source, and introduces a high-order squint range model based on the long GNSS-R BSAR synthetic aperture time in the one-stop fixed mode to obtain an accurate description of the relative time variation of the squint range between the navigation satellite and the target. Based on this model, firstly, the range migration is corrected by the time-domain cancellation of the direct signal and the echo signal to realize the accurate correction of the target range migration in the whole scene; By azimuth-block hybrid correlation processing, the azimuth shifting nature of the echo signal is overcome, and efficient and accurate imaging of the whole scene is realized. The imaging efficiency of the proposed algorithm is better than that of the traditional BP algorithm, the imaging accuracy is comparable to that of the back projection (BP) algorithm, and the focusing effect can be improved by adjusting the width of the orientation bins as needed. Finally, to validate the proposed algorithm, we conducted simulations and experiments with GPS-L5 signals, the simulation and experimental results verified the feasibility and efficiency of the proposed algorithm.

**Keywords:** global navigation satellite system-reflectometry; bistatic synthetic aperture radar; azimuth change processing; range Doppler algorithm; range cell migration correction

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0273

# 涡轴发动机时延鲁棒串级 PI 控制器设计

陈义峰, 郭迎清\*, 毛皓天

(西北工业大学 动力与能源学院, 西安 710072)

**摘 要:** 针对涡轴发动机分布式控制系统中存在时延导致系统性能退化的问题, 利用线性矩阵不等式 (LMI) 方法设计了时延鲁棒串级 PI 控制器。先利用内模控制 (IMC) 方法得到涡轴发动机串级控制器内、外环的 PI 控制器结构; 再利用频域回路成形的方法给出保证系统具有期望性能的 LMI 形式约束条件; 利用梯度近似的方法通过劳斯-赫尔维茨判据得到保证系统稳定的控制器参数约束条件; 最后, 在基于 TrueTime 的涡轴发动机分布式非线性仿真平台上进行数字仿真。仿真结果表明: 在 0.04 s 时延条件下, 当功率需求下降 5% 时, 系统的调节时间小于 5 s, 功率涡轮转速超调不超过 0.5%, 且其最大燃油变化率只有传统串级 PI 控制系统的 67%; 设计的控制器能有效应对对涡轴发动机分布式控制系统中存在的时延问题, 同时能够以更小的代价保证系统具有期望的性能。

**关键词:** 涡轴发动机分布式控制系统; 时延; 鲁棒串级 PI 控制器; 线性矩阵不等式; 频域回路成形; TrueTime 仿真

**中图分类号:** V233.7

**文献标志码:** A

**文章编号:** 1001-5965(2023)03-0597-09

当前, 涡轴发动机在军用与民用领域广泛的应用于直升机、发电机和地面车辆的辅助动力装置等。直升机是一个高度集成的系统, 其升力通过总矩杆角度进行控制, 因此, 当前涡轴发动机的控制方式为功率涡轮定转速控制。涡轴发动机主要由燃气发生器和功率涡轮两部件组成, 而功率涡轮与主旋翼连接并通过燃气发生器产生的燃气进行驱动。由于旋翼系统是一个大惯性系统, 当前涡轴发动机多采用串级控制系统结构, 以内环控制器保证燃气发生器的快速响应而以外环控制器保证转子系统的转速恒定<sup>[1-3]</sup>。

分布式控制系统具有系统重量轻、可靠性高、模块化、低生命周期成本及利于新控制技术(如主动控制技术等)的整合等优点。随着人们对发动机性能需求的不断提高, 这些优点引起了人们的持续

关注, 分布式控制系统被认为是未来发动机控制系统的发展方向之一<sup>[4-6]</sup>。分布式控制系统中, 传统的传感器/执行机构将被相应的智能传感器/执行机构所取代, 这些节点通过通信网络进行相互连接。控制系统中的所有信号都在数据总线上进行传输<sup>[5]</sup>。

对于涡轴发动机串级控制系统, 当其所有节点通过总线实现通信时, 该系统被称为网络串级控制系统 (networked cascade control system, NCCS)<sup>[2]</sup>。由于总线带宽约束或节点瞬时失效等原因, 数据传输过程中不可避免的会存在时延问题, 而这些时延会导致系统的性能降低甚至导致系统失稳<sup>[3, 7]</sup>。因此, 为了保证系统的控制性能, 控制系统必须对时延具有鲁棒性。

针对 NCCS, Huang 等<sup>[7]</sup>开展的相关研究中, 提出了 4 种 NCCS 结构并对其时延进行了分析, 针对

收稿日期: 2021-05-26; 录用日期: 2021-08-20; 网络出版时间: 2021-09-13 10:16

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210911.1227.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210911.1227.001.html)

基金项目: 国家科技重大专项 (2017-V-0015-0067)

\*通信作者. E-mail: [yqguo@nwpu.edu.cn](mailto:yqguo@nwpu.edu.cn)

**引用格式:** 陈义峰, 郭迎清, 毛皓天. 涡轴发动机时延鲁棒串级 PI 控制器设计 [J]. 北京航空航天大学学报, 2023, 49 (3): 597-605.

CHEN Y F, GUO Y Q, MAO H T. Design of time-delay robust cascade PI controller for turboshaft engine [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 597-605 (in Chinese).

系统中的固定时延,文献[8]设计了串级比例-积分-微分(proportional-integral-derivative, PID)控制器,文献[9]针对一个控制周期内的短时延设计了串级 $H_\infty$ 控制器,文献[10-11]则针对固定时延提出了内模控制(internal model control, IMC)方法和二自由度方法等鲁棒串级PID控制器设计方法。对于涡轴发动机分布式控制系统, Belapurkar 和 Yedavalli<sup>[2]</sup>提出了2种串级控制系统结构,并采用LQR方法设计了长时延下的串级状态反馈控制器; Chen等<sup>[12]</sup>采用Lyapunov-Krasovskii方法给出了静态串级控制器设计方法,并利用区域极点配置方法来保证系统的动态性能; Liu等<sup>[3,13]</sup>针对涡轴发动机分布式控制系统中存在的时延与随机丢包等问题,提出一种新的分析时延系统稳定性及求解最大允许时延的方法,以及利用Lyapunov定理给出了存在随机丢包的系统控制器设计方法。文献[2,12-13]针对涡轴发动机分布式控制系统的研究都是直接从静态控制器的角度进行控制器设计。

在当前的涡轴发动机控制中, PID控制仍是一种非常重要的控制方法。本文针对涡轴发动机分布式控制系统中存在的时延设计了鲁棒串级PI控制器。首先,分析了涡轴发动机分布式串级控制系统的结构及时延组成,利用IMC方法处理串级控制系统内外环同时存在的时延,并确定了此串级控制系统的控制器为串级PI控制器结构;然后,利用频

域回路成形的办法,确定了保证系统具有期望的动态性能约束,并将控制器设计问题转化为具有线性矩阵不等式(linear matrix inequality, LMI)约束的广义特征值求解问题;接着,为了保证系统稳定,利用梯度近似的方法将劳斯-赫尔维茨判据转化为LMI形式的约束,将所有的约束进行综合并求解优化问题获得期望的鲁棒串级PI控制器的内外环参数;最后,利用基于TrueTime的涡轴发动机分布式仿真平台对控制器进行了数字仿真验证。

## 1 鲁棒串级PI控制器设计

对于涡轴发动机分布式控制系统的时延鲁棒串级PI控制器设计方法包括三部分:IMC-PI控制器的设计, LMI形式的频域回路成形约束条件及LMI形式的稳定性约束条件获取。

### 1.1 串级IMC-PI控制器的设计

当前的涡轴发动机控制系统,针对转子系统的大惯性特性,为了保证转子转速恒定,同时,燃气发生器具有较快的响应速度,采用的串级控制系统结构如图1所示<sup>[2]</sup>。其内环为燃气发生器的转速 $N_g$ 控制系统,而外环为转子系统的转速 $N_p$ 控制系统,系统的输入为参考转速 $r$ ,控制量为燃油流量 $W_f$ ,燃气发生器输出轴力矩 $Q_s$ 用于驱动转子系统,而旋翼系统的功率需求LDL是涡轴发动机控制系统的扰动量。

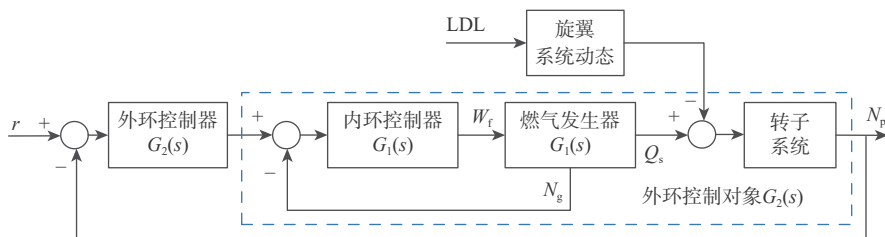


图1 涡轴发动机串级控制系统原理<sup>[2]</sup>

Fig. 1 Block of cascade control system of turboshaft engine<sup>[2]</sup>

根据内外环传感器、内外环控制器、执行机构和控制对象之间的通信方式, NCCS被分为4类<sup>[7]</sup>。对于涡轴发动机分布式控制系统,当系统的所有控制算法仍在全权限数字电子控制器(full authority digital electronic controller, FADEC)中计算时,涡轴发动机分布式控制系统属于第三类NCCS<sup>[14]</sup>,即传感器与控制器之间、控制器与执行机构之间都通过网络连接,而内外环控制器是同一个节点,其结构如图2所示。

涡轴发动机分布式控制系统内的时延包括3个:内环传感器与内环控制器之间的时延 $\tau_{s1}$ ,外环传感器与外环控制器之间的时延 $\tau_{s2}$ 和控制器与

执行机构之间的时延 $\tau_a$ 。

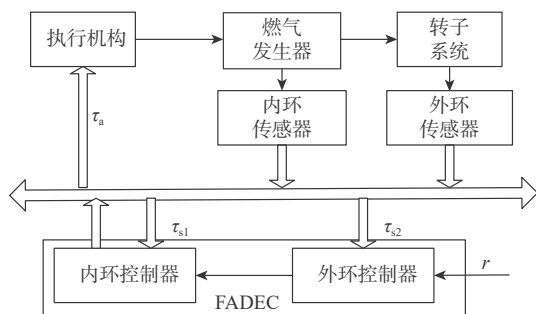


图2 涡轴发动机分布式控制系统第III类NCCS结构示意图  
Fig. 2 Type III NCCS configuration of distributed control system of turboshaft engine

当采用传统传递函数方法直接设计串级 PI 控制器时, 内环时延会对外环控制对象产生复杂的影响从而导致外环控制器设计复杂化。采用 IMC 方法可以明确串级控制系统内外环时延的关系<sup>[10]</sup>: 内环时延为所有内环节之间时延之和, 而外环时延为内环总时延与外环所有节点之间时延之和。因此, 对于本节涡轴发动机 NCCS, 可以利用 IMC 方法从内环到外环依次设计 PI 控制器, 且此时系统的内环时延为  $\tau_1 = \tau_{s1} + \tau_a$ , 而外环时延为  $\tau_2 = \tau_1 + \tau_{s2}$ 。

以单环控制系统为例说明采用 IMC 方法设计控制器的步骤, 对于串级控制系统则从内环到外环逐层进行设计<sup>[10]</sup>。对于一个单位负反馈的闭环 IMC 控制系统, 其结构如图 3 所示,  $G(s)$  为控制对象的传递函数, 根据 IMC 控制器设计方法<sup>[15-16]</sup>, 控制对象的模型  $\hat{G}(s)$  为

$$\hat{G}(s) = G_A(s)G_M(s) \quad (1)$$

式中:  $G_A(s)$  为由模型的时滞环节和正零点项构成的项;  $G_M(s)$  为由模型的增益系数、负零点项和负极点项构成的项。单位负反馈控制器  $C(s)$  的表达式为

$$C(s) = \frac{q(s)}{1 - q(s)\hat{G}(s)} \quad (2)$$

式中:  $q(s) = G_M^{-1}(s)f(s)$  为内模控制器,  $f(s) = 1/(\lambda s + 1)^n$  为 IMC 滤波器, 其阶数  $n$  在选择时必须保证控制器具有适当的阶数, 而常数  $\lambda$  为待设计参数。

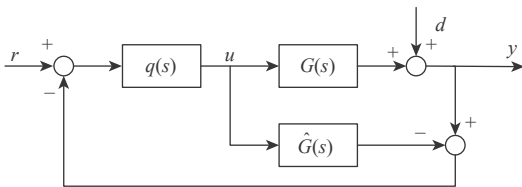


图 3 IMC 控制系统结构示意图

Fig. 3 IMC control system

针对某型民用涡轴发动机, 本文利用 GasTurb 软件和 MATLAB 环境构建了该型涡轴发动机的部件级模型及仿真平台<sup>[17]</sup>。通过系统辨识的方法可以得到该型涡轴发动机在地面静止状态下, 燃气发生器转速  $N_g$  和转子系统转速  $N_p$  均为 100% 状态时燃气发生器增量模型  $G_1(s)$  为带有正零点的一阶传递函数, 并且将内环时延作为时滞环节, 其表达式为

$$G_1(s) = \frac{K_1(\beta_{11}s + 1)e^{-\tau_1 s}}{s + \alpha_{11}} \quad \beta_{11} < 0 \quad (3)$$

式中:  $s$  为传递函数自变量;  $K_1$ 、 $\beta_{11}$  和  $\alpha_{11}$  为模型的常数系数;  $e$  为自然常数。模型的输入为燃油流量增量  $\Delta W_f$ , kg/s; 输出为燃气发生器的转速增量  $\Delta N_g$ ,

r/min。

针对此控制对象利用 IMC 方法设计控制器, 用一阶 Pade 近似时滞环节, 此时可得

$$q_1(s) = \frac{(s + \alpha_{11})(1 + 0.5\tau_1 s)}{K_1(\lambda_1 s + 1)^n} \quad (4)$$

式中:  $q_1(s)$  为内环的内模控制器;  $\lambda_1$  为内环控制器待设计参数。取 IMC 滤波器的阶数为 2 阶, 可以得到内环控制器  $C_1(s)$  为

$$C_1(s) = \frac{q_1(s)}{1 - q_1(s)G_1(s)} = \frac{1}{K_1(2\lambda_1 + 0.5\tau_1 - \beta_{11})} \left(1 + \frac{\alpha_{11}}{s}\right) \frac{(1 + 0.5\tau_1 s)}{\frac{\lambda_1^2 + 0.5\tau_1\beta_{11}}{2\lambda_1 + 0.5\tau_1 - \beta_{11}}s + 1} \quad (5)$$

内环控制器为带有一阶滤波器的 PI 控制器。

同样, 利用系统辨识的方法, 可以得到该工作点处外环控制对象 (包含了内环闭环及转子系统) 的模型  $G_2(s)$  为带有正零点的二阶传递函数, 时延作为时滞环节, 其表达式为

$$G_2(s) = \frac{K_2(\beta_{21}s + 1)(\beta_{22}s + 1)e^{-\tau_2 s}}{(s + \alpha_{21})(s + \alpha_{22})} \quad \beta_{21} < 0 \quad (6)$$

式中:  $K_2$ 、 $\beta_{21}$ 、 $\beta_{22}$ 、 $\alpha_{21}$ 、 $\alpha_{22}$  为模型的常数系数。模型输入为燃气发生器转速增量  $\Delta N_g$ , 输出为功率涡轮转速增量  $\Delta N_p$ , r/min。IMC 滤波器取为 2 阶, 可得外环控制器  $C_2(s)$  为

$$C_2(s) = \frac{1}{K_2\varphi} \left(1 + \frac{\alpha_{21}}{s}\right) \frac{(s + \alpha_{22})(1 + 0.5\tau_2 s)}{(\beta_{22}s + 1) \left(\frac{\lambda_2^2 + 0.5\tau_2\beta_{21}}{\varphi}s + 1\right)} \quad (7)$$

式中:  $\varphi$  为  $2\lambda_2 + 0.5\tau_2 - \beta_{21}$  的缩写形式;  $\lambda_2$  为外环控制器待设计参数, 所以外环控制器为带有双二阶滤波器的 PI 控制器。

### 1.2 鲁棒 PI 参数约束

在确定 PI 控制器结构后就需要对控制器的参数进行求解。对于一个单位负反馈 PI 控制系统, 假设其控制器为  $C(s)$ , 而系统的控制对象仍为  $G(s)$ 。PI 控制器设计目标是求解合适的控制器参数, 使得闭环系统具有期望的性能及足够的鲁棒性。而系统的性能与开环系统频域幅相曲线的形状具有高度相关性。因此, 利用频域回路成形的方法可知期望的控制器需要同时满足以下的约束条件<sup>[18-20]</sup>:

1) 为了保证系统具有期望的穿越频率及相位裕度, 开环传递函数需满足:

$$\begin{cases} |C(j\omega_c)G(j\omega_c)| = 0 \\ \arg\{C(j\omega_c)G(j\omega_c)\} = -\pi + \phi_m \end{cases} \quad (8)$$

式中:  $\omega_c$  为期望的穿越频率, rad/s;  $\phi_m$  为期望的相

位裕度, rad。

2) 为了克服低频干扰的影响,同时为了保证系统的跟踪能力,开环系统应满足:

$$|C(j\omega)G(j\omega)| \geq \xi \quad \omega \leq \omega_s \quad (9)$$

式中:  $\xi$  为开环系统在低频域的期望幅值下限, dB;  $\omega_s$  为低频干扰的频率, rad/s。

3) 为了抑制高频噪声,开环系统应满足:

$$|C(j\omega)G(j\omega)| \leq \rho \quad \omega \geq \omega_t \quad (10)$$

式中:  $\rho$  为开环系统在高频域的期望幅值上限, dB;  $\omega_t$  为高频噪声的频率, rad/s。

4) 为了限制控制量的幅值,控制器应满足:

$$|C(j\omega)| \leq \delta \quad \omega \geq \omega_r \quad (11)$$

式中:  $\delta$  为控制器在高频域的期望幅值上限, dB;  $\omega_r$  为高频频率, rad/s。

对于线性的 PI 控制器可以表达为

$$C(s) = K_p + \frac{K_I}{s} = \mathbf{W}(s)\mathbf{X} \quad (12)$$

式中:  $\mathbf{X} = [K_p \quad K_I]^T$  为待整定参数构成的向量,  $K_p$  为 PI 控制器的比例项系数,  $K_I$  为 PI 控制器的积分项系数;  $\mathbf{W}(s) = [1 \quad 1/s]$ 。根据幅值与相位裕度的定义、Schur 补引理及复矩阵与实矩阵的转换定理<sup>[18,21]</sup>, 可将式(8)~式(12)的控制器参数求解问题转换成带约束式(13)~式(16)的广义特征值求解问题, 4条约束式(8)~式(11)与4个矩阵不等式约束式(13)、式(16)、式(14)和式(15)一一对应。

$$\min_{\mathbf{X}} \quad \text{s.t.} \quad \begin{bmatrix} 0 & \Gamma_1 & 0 & \Gamma_2 \\ * & 0 & \Gamma_3 & 0 \\ * & * & 0 & \Gamma_4 \\ * & * & * & 0 \end{bmatrix} < \gamma \mathbf{I} \quad (13)$$

$$\begin{bmatrix} -\Phi & C_R(j\omega_t) & 0 & C_I(j\omega_t) \\ * & -|G(j\omega_t)|^{-2} & -C_I(j\omega_t) & 0 \\ * & * & -\Phi & C_R(j\omega_t) \\ * & * & * & -|G(j\omega_t)|^{-2} \end{bmatrix} < \mathbf{0} \quad (14)$$

$$\begin{bmatrix} -\Psi & C_R(j\omega_r) & 0 & C_I(j\omega_r) \\ * & -1 & -C_I(j\omega_r) & 0 \\ * & * & -\Psi & C_R(j\omega_r) \\ * & * & * & -1 \end{bmatrix} < \mathbf{0} \quad (15)$$

$$\sqrt{(C_R(j\omega_s))^2 + (C_I(j\omega_s))^2} > R \quad (16)$$

式中:  $(\cdot)_R = \text{Re}(\cdot)$ 、 $(\cdot)_I = \text{Im}(\cdot)$  分别为复数的实部与虚部; \* 为对称矩阵中的对称项;  $\mathbf{I}$  为单位矩阵;  $\Gamma_1 = D_R - C_R(j\omega_c)G_R(j\omega_c) + C_I(j\omega_c)G_I(j\omega_c)$ ;  $\Gamma_2 = D_I - C_R(j\omega_c)G_I(j\omega_c) - C_I(j\omega_c)G_R(j\omega_c)$ ;  $\Gamma_3 = -D_I + C_R(j\omega_c)G_I(j\omega_c) + C_I(j\omega_c)G_R(j\omega_c)$ ;  $\Gamma_4 = \Gamma_1$ ;  $D = e^{j(\phi_m - \pi)}$ ;

$$\Phi = 10^{\rho/10}; \Psi = 10^{\delta/10}; R = 10^{\xi/20}/|G(j\omega_s)|。$$

对于带约束式(13)~式(16)的广义特征值问题, 约束式(13)~式(15)都是 LMI 约束, 但约束式(16)是非线性约束。约束式(16)的可行域为半径为  $R$  的圆以外的区域, 但该区域为非凸区域, 无法利用 LMI 工具箱求解。为了对优化问题进行求解, 利用圆的切线将可行域分成4个凸可行域进行求解, 如图4所示, 4个凸可行域可表示为

$$[\mathbf{W}_R(j\omega_s) + \mathbf{W}_I(j\omega_s)]\mathbf{X} < -\sqrt{2}R \quad (17)$$

$$[\mathbf{W}_I(j\omega_s) - \mathbf{W}_R(j\omega_s)]\mathbf{X} < -\sqrt{2}R \quad (18)$$

$$[\mathbf{W}_R(j\omega_s) + \mathbf{W}_I(j\omega_s)]\mathbf{X} > \sqrt{2}R \quad (19)$$

$$[\mathbf{W}_I(j\omega_s) - \mathbf{W}_R(j\omega_s)]\mathbf{X} > \sqrt{2}R \quad (20)$$

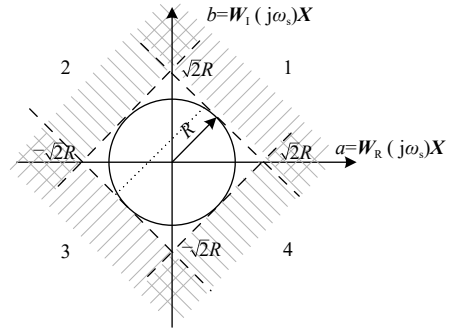


图4 由4条切线构成的凸可行域

Fig. 4 Feasible convex region divided by four tangents

注: ① 由于将圆外可行域变成4条切线分割成的可行域, 因此缺失了一部分可行域, 这导致此方法具有一定的保守性, 可以通过增加切线的数量来降低该方法的保守性。对于圆形不可行域, 其面积为  $\pi R^2$ , 采用正外切四边形时, 其不可行域面积为  $4R^2$ , 为圆域的1.273倍; 当采用正外切八边形时, 不可行域面积为  $3.314R^2$ , 为圆域的1.013倍。以增加计算量为代价来降低方法的保守性。② 对于4个可行域, 只需要保证优化问题的解落在其中1个可行域即可, 并不需要4个条件同时得到满足。

### 1.3 涡轴发动机串级 PI 控制器的稳定性约束

1.2节给出的 PI 控制器参数整定约束条件式(8)~式(11)主要是保证系统具有期望的性能, 但系统的稳定性还需要额外的条件进行约束。可通过劳斯-赫尔维茨判据来对控制器参数进行约束, 该方法在文献[18]中已给出。

对于涡轴发动机分布式控制系统串级 PI 控制器, 首先, 确定内环 PI 控制器参数的约束条件, 然后, 再确定外环 PI 控制器参数的约束条件。

对于内环燃气发生器, 其传递函数为式(3), 时滞环节采用一阶 Pade 近似, 内环控制器参数为  $\mathbf{X}_I = [K_{PI} \quad K_{II}]^T$  可以得到内环系统的特征方程为

$$A_1 s^3 + B_1 s^2 + D_1 s + E_1 = 0 \quad (21)$$

式中:  $A_1 = 0.5\tau_1 - 0.5K_1\tau_1\beta_{11}K_{P1}$ ;  $B_1 = 0.5\tau_1\alpha_{11} + 1 - 0.5K_1\tau_1\beta_{11}K_{I1} - 0.5K_1\tau_1K_{P1} + K_1\beta_{11}K_{P1}$ ;  $D_1 = \alpha_{11} - 0.5K_1\tau_1K_{I1} + K_1\beta_{11}K_{I1} + K_1K_{P1}$ ;  $E_1 = K_1K_{I1}$ 。利用梯度近似, 其中当  $X_{10} = [1 \ 1]^T$  时得到特征方程系数记为  $A_{10}$ 、 $B_{10}$ 、 $D_{10}$ 、 $E_{10}$ , 可以得到内环 LMI 形式的稳定性约束条件为式(23)。

同样方法, 外环控制对象的传递函数为式(6), 时滞采用一阶 Pade 近似, 外环控制器的参数为  $X_2 = [K_{P2} \ K_{I2}]^T$ , 可以得到外环控制系统的特征方程为

$$A_2 s^4 + B_2 s^3 + D_2 s^2 + E_2 s + F_2 = 0 \quad (22)$$

式中:  $A_2$ 、 $B_2$ 、 $D_2$ 、 $E_2$ 、 $F_2$  的表达式分别如式(24)~式(28)所示。再通过梯度近似的方法来对约束进行线性化,  $X_{20} = [1 \ 1]^T$  时得到系数记为  $A_{20}$ 、 $B_{20}$ 、 $D_{20}$ 、 $E_{20}$ 、 $F_{20}$ , 最后得到外环 LMI 形式的稳定性约束如式(31)所示。

$$\begin{cases} [0.5K_1\tau_1\beta_{11} \ 0] X_1 < 0.5\tau_1 \\ [0.5K_1\tau_1 - K_1 \ 0.5K_1\tau_1\beta_{11}] X_1 < 1 + 0.5\alpha_{11}\tau_1 \\ [0 \ K_1] X_1 > 0 \\ \left[ \begin{matrix} (-0.5K_1\tau_1 + K_1\beta_{11})D_{10} + K_1B_{10} \\ -0.5K_1\tau_1\beta_{11}(\alpha_{11} - 0.5K_1\tau_1 + K_1\beta_{11}) + (-0.5K_1\tau_1 + K_1\beta_{11})B_{10} - 0.5K_1\tau_1 \\ -\alpha_{11}(0.5\tau_1\alpha_{11} + 1) + (-0.5K_1\tau_1 + K_1\beta_{11}) \cdot (K_1 - 0.5K_1\tau_1\beta_{11} - 0.5K_1\tau_1 + K_1\beta_{11}) \end{matrix} \right]^T X_1 > \end{cases} \quad (23)$$

$$A_2 = 0.5\tau_2 - 0.5K_2\tau_2\beta_{21}\beta_{22}K_{P2} \quad (24)$$

$$B_2 = 0.5\tau_2(\alpha_{21} + \alpha_{22}) + 1 - 0.5K_2\tau_2(\beta_{21} + \beta_{22})K_{P2} - 0.5K_2\tau_2\beta_{21}\beta_{22}K_{I2} + K_2\beta_{21}\beta_{22}K_{P2} \quad (25)$$

$$D_2 = 0.5\tau_2\alpha_{21}\alpha_{22} + \alpha_{21} + \alpha_{22} - 0.5K_2\tau_2K_{P2} - 0.5K_2\tau_2(\beta_{21} + \beta_{22})K_{I2} + K_2(\beta_{21} + \beta_{22})K_{P2} + K_2\beta_{21}\beta_{22}K_{I2} \quad (26)$$

$$E_2 = \alpha_{21}\alpha_{22} - 0.5K_2\tau_2K_{I2} + K_2K_{P2} + K_2(\beta_{21} + \beta_{22})K_{I2} \quad (27)$$

$$F_2 = K_2K_{I2} \quad (28)$$

同时定义 2 个梯度矩阵:

$$\nabla f_{21}(X_{20}) = \begin{bmatrix} (-0.5K_2\tau_2(\beta_{21} + \beta_{22}) + K_2\beta_{21}\beta_{22})D_{20} + (-0.5K_2\tau_2 + K_2(\beta_{21} + \beta_{22}))B_{20} + 0.5K_2\tau_2\beta_{21}\beta_{22}E_{20} - K_2A_{20} \\ -0.5K_2\tau_2\beta_{21}\beta_{22}D_{20} + (-0.5K_2\tau_2(\beta_{21} + \beta_{22}) + K_2\beta_{21}\beta_{22})B_{20} - (-0.5K_2\tau_2 + K_2(\beta_{21} + \beta_{22}))A_{20} \end{bmatrix} \quad (29)$$

$$\nabla f_{22}(X_{20}) = \begin{bmatrix} [(K_2\beta_{21}\beta_{22} - 0.5K_2\tau_2(\beta_{21} + \beta_{22}))D_{20}E_{20} + (K_2(\beta_{21} + \beta_{22}) - 0.5K_2\tau_2)B_{20}E_{20} + K_2B_{20}D_{20} + 0.5K_2\tau_2\beta_{21}\beta_{22}E_{20}^2 - 2K_2A_{20}E_{20} - 2(K_2\beta_{21}\beta_{22} - 0.5K_2\tau_2(\beta_{21} + \beta_{22}))B_{20}F_{20}] \\ [-0.5K_2\tau_2\beta_{21}\beta_{22}D_{20}E_{20} + (K_2\beta_{21}\beta_{22} - 0.5K_2\tau_2(\beta_{21} + \beta_{22}))B_{20}E_{20} + (K_2(\beta_{21} + \beta_{22}) - 0.5K_2\tau_2)B_{20}D_{20} - 2(K_2(\beta_{21} + \beta_{22}) - 0.5K_2\tau_2)A_{20}E_{20} + K_2\tau_2\beta_{21}\beta_{22}B_{20}F_{20} - B_{20}^2F_{20}] \end{bmatrix} \quad (30)$$

$$\begin{cases} [0.5K_2\tau_2\beta_{21}\beta_{22} \ 0] X_2 < 0.5\tau_2 \\ [0.5K_2\tau_2(\beta_{21} + \beta_{22}) - K_2 \ 0.5K_2\tau_2\beta_{21}\beta_{22}] X_2 < 0.5\tau_2(\alpha_{21} + \alpha_{22}) + 1 \\ [0 \ K_2] X_2 > 0 \\ \nabla f_{21}(X_{20})^T X_2 > -(0.5\tau_2\alpha_{21}\alpha_{22} + \alpha_{21} + \alpha_{22})B_{20} + \alpha_{21}\alpha_{22}A_{20} - (0.5\tau_2(\alpha_{21} + \alpha_{22}) + 1 - B_{20})D_{20} + 0.5K_2\tau_2\beta_{21}\beta_{22}E_{20} \\ \nabla f_{22}(X_{20})^T X_2 > -[(0.5\tau_2\alpha_{21}\alpha_{22} + \alpha_{21} + \alpha_{22})B_{20}E_{20} + (E_{20} - 2\alpha_{21}\alpha_{22})A_{20}E_{20} + (0.5\tau_2(\alpha_{21} + \alpha_{22}) + 1 - B_{20})D_{20}E_{20} + 2(B_{20} - 0.5\tau_2(\alpha_{21} + \alpha_{22}) + 1)B_{20}F_{20} + (\alpha_{21}\alpha_{22} - E_{20})B_{20}D_{20} + (-0.5K_2\tau_2\beta_{21}\beta_{22})E_{20}^2] \end{cases} \quad (31)$$

至此, 关于时延鲁棒串级 PI 控制器设计的所有 LMI 形式约束已经得到, 然后利用 LMI 工具箱 YALMIP 对其进行求解<sup>[22]</sup>, 便可以得到期望的涡轴发动机分布式控制系统内外环 PI 控制器的相关参数, 最后通过式(5)和式(7)分别得到内外环的 IMC 滤波器参数  $\lambda_1$  和  $\lambda_2$ , 从而得到串级控制系统内环一阶滤波器和外环双二阶滤波器的相关参数。

## 2 基于 TrueTime 的非线性仿真

本节主要通过基于 TrueTime 工具箱搭建的某

型涡轴发动机分布式控制系统非线性仿真平台来验证所设计的时延鲁棒串级 PI 控制器的有效性及性能。

对于某型涡轴发动机, 通过系统辨识的方法可以得到燃气发生器加时延的模型式(3)的表达式为

$$G_1(s) = \frac{2.991 \times 10^5 (-0.01s + 1)e^{-\tau_1 s}}{s + 4.001} \quad (32)$$

涡轴发动机燃气发生器的期望控制带宽为 10 rad/s 左右, 而发动机控制系统的高频为 40 rad/s 以上, 低频为 0.1 rad/s 以下, 并且发动机控制系统幅

值的理想斜率为 $-20\text{ dB/dec}^{[1,23]}$ 。因此,约束式(8)~式(11)中的参数分别取 $\omega_c = 9\text{ rad/s}$ 、 $\phi_m = 83^\circ$ 、 $\xi = 20\text{ dB}$ 、 $\omega_s = 0.1\text{ rad/s}$ 、 $\rho = -10\text{ dB}$ 、 $\omega_l = 50\text{ rad/s}$ 、 $\delta = -10\text{ dB}$ 、 $\omega_r = 50\text{ rad/s}$ 。同时,本文假设图2所示的涡轴发动机分布式控制系统中,内环传感器到控制器的时延 $\tau_{s1}$ 和内环控制器到执行机构的时延 $\tau_a$ 均为半个控制周期,即 $0.01\text{ s}$ 。因此,内环控制系统的时延为 $\tau_1 = \tau_{s1} + \tau_a = 0.02\text{ s}$ ,即时滞环节为 $e^{-0.02s}$ 。然后利用1.2节与1.3节联合的控制器设计方法可以得到内环PI控制器参数 $\mathbf{X}_1 = 10^{-5} \times [2.96\ 11.86]^T$ ,内环IMC滤波器参数 $\lambda_1 = 0.0464$ 。

根据内环控制器的参数及系统辨识的方法,可以得到涡轴发动机外环控制对象加时延的模型式(6)的表达式为

$$G_2(s) = \frac{3.783(-0.01s+1)(0.2847s+1)e^{-\tau_2 s}}{(s+10.6651)(s+1.2649)} \quad (33)$$

涡轴发动机整机的期望带宽在 $0.1 \sim 3.5\text{ rad/s}$ ,因此,约束式(8)~式(11)中的参数分别取 $\omega_c = 3\text{ rad/s}$ 、 $\phi_m = 85^\circ$ 、 $\xi = 10\text{ dB}$ 、 $\omega_s = 0.1\text{ rad/s}$ 、 $\rho = -10\text{ dB}$ 、 $\omega_l = 50\text{ rad/s}$ 、 $\delta = -10\text{ dB}$ 、 $\omega_r = 50\text{ rad/s}$ 。假设外环传感器到控制器的时延 $\tau_{s2} = 0.02\text{ s}$ ,则外环控制系统的时延为 $\tau_2 = \tau_1 + \tau_{s2} = 0.04\text{ s}$ ,即模型的时滞环节为 $e^{-0.04s}$ 。最后利用1.2节与1.3节联合的控制器设计方法可以得到外环PI控制器的参数 $\mathbf{X}_2 = [0.3093\ 3.2935]^T$ ,外环IMC滤波器参数 $\lambda_2 = 0.3333$ 。在得到内外环的控制器后,内外环的开环幅相曲线如图5和图6所示。通过幅相曲线可以看出,带有控制器的开环系统具有足够大的低频增益(大于 $30\text{ dB}$ )和足够小的高频增益(小于 $-14\text{ dB}$ ),内环穿越频率为期望频率( $9\text{ rad/s}$ ),外环穿越频率为 $1\text{ rad/s}$ 。

在MATLAB环境中利用Simulink/TrueTime工具箱<sup>[24]</sup>搭建的涡轴发动机分布式控制系统非线性仿真平台采用部分分布式结构<sup>[14]</sup>,其原理如图7所示。

在图7中,采用网络模块来仿真总线特性;智

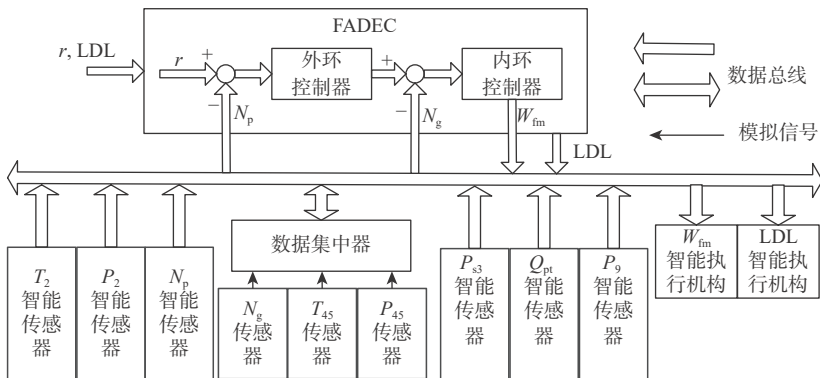


图7 基于TrueTime的涡轴发动机分布式控制系统仿真平台

Fig. 7 Simulation platform of TrueTime based distributed control system of turboshaft engine

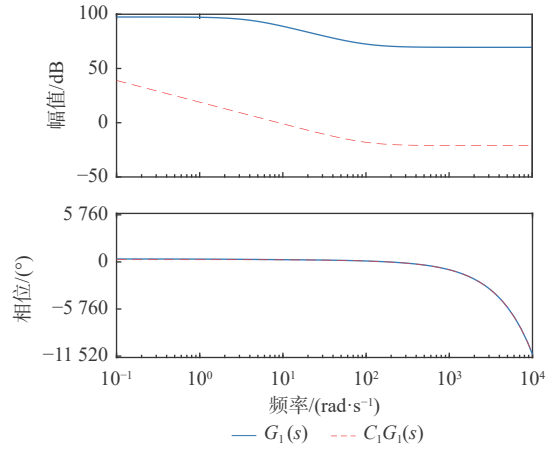


图5 内环开环系统的Bode图

Fig. 5 Bode plot of inner open loop

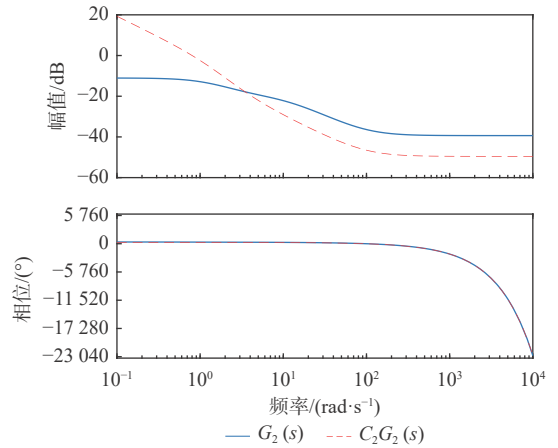


图6 外环开环系统的Bode图

Fig. 6 Bode plot of outer open loop

能传感器、数据集中器、智能执行机构及FADEC节点都采用计算模块来仿真,最上部为串级PI控制器结构。网络模块可以仿真9种总线通信协议,此平台选择CAN总线通信协议,并通过参考消息的方式实现TTCAN总线通信,带宽选择最高 $1\text{ Mbps}$ ,由于不考虑丢包问题,丢包率设置为0。仿真过程中,节点的固定时延通过在节点初始化程序中用ttSetNetworkParameter函数进行设置,随机时延通

过 Simulink 中的可变传输时延模块加入智能执行机构中。

在仿真中, 对比鲁棒串级 PI 控制器与利用 PID Tuner 工具箱设计得到的传统串级 PI 控制器的性能。利用该工具箱设计串级 PI 控制器时内外环的幅值与相位裕度保持一致, 得到其内环参数为  $K_{P1} = 3.16 \times 10^{-5}$ ,  $K_{I1} = 7.9 \times 10^{-5}$ , 外环控制器参数为  $K_{P2} = 4.817$ ,  $K_{I2} = 10.44$ , 涡轴发动机的控制周期为 0.02 s, 控制器的离散化采用 Tustin 方法。首先, 仿真微小时延 (仿真系统中信号传输时延为 0.001 5 s, 即时延在一个控制周期内); 然后, 对控制器设计点的时延 (总时延为 0.04 s, 即 2 个控制周期内) 进行仿真; 最后, 为了比较 2 个控制器对于时延的鲁棒性, 将仿真总时延增加到 3 个控制周期, 即  $\tau_{s1} = \tau_{s2} = \tau_a = 0.02$  s, 总矩杆角度由 100% 下降到 95%, 分别在 2 个串级控制器作用下, 系统的参数相对变化量如图 8 所示。

从图 8 可以看出, 当时延分别在 1 个控制周期内和 2 个控制周期内时, 2 个串级 PI 控制器都能保持系统的稳定, 且功率涡轮的转速超调小于 0.5%, 调节时间小于 5 s (涡轴发动机小扰动下的性能指标为<sup>[25]</sup>:  $N_p$  超调不超过 5%~7%, 稳态精度为 0.2% 左右, 调节时间为 8~10 s), 当时延增加到 3 个控制周期时,

传统的串级 PI 控制器无法保证系统稳定, 而鲁棒串级 PI 控制器仍能保证系统稳定。因此, 鲁棒串级 PI 控制器对时延具有更强的鲁棒性, 能够容忍更大的时延, 且通过仿真可知其保证系统稳定的时延边界为 0.097 s, 传统串级 PI 控制器的时延边界为 0.057 s, 即控制系统稳定时延容忍边界由原来不足 3 个控制周期 (0.06 s) 扩大到了接近 5 个控制周期 (0.1 s)。从燃油流量可以看出时延条件下鲁棒串级 PI 控制器下最大燃油流量变化率为 0.259 6 kg/s<sup>2</sup>, 而传统串级 PI 控制器下最大燃油流量变化率为 0.386 5 kg/s<sup>2</sup>, 说明鲁棒串级 PI 控制器对执行机构的快速性要求降低。

从图 6 可知, 设计控制器时鲁棒串级 PI 控制器外环穿越频率实际为 1 rad/s, 将 2 种方法下外环穿越频率都取为 1 rad/s, 而其他条件不变再重新设计外环控制器, 鲁棒串级 PI 控制器的外环控制器参数  $\bar{X}_2 = [0.2807 \quad 2.989]^T$ , 利用 PID Tuner 得到的外环控制器参数  $K_{P2} = 1.632$ ,  $K_{I2} = 2.862$ 。将最大时延增加为 0.08 s (4 个控制周期), 其他条件不变再进行仿真, 得到的系统响应曲线如图 9 所示。

从图 9 可以看出, 分别在 0.001 5 s (1 个周期内) 和 0.04 s (2 个控制周期) 时延下, 2 个控制器都能满足控制需求, 而当时延增加到 4 个控制周期时, 传

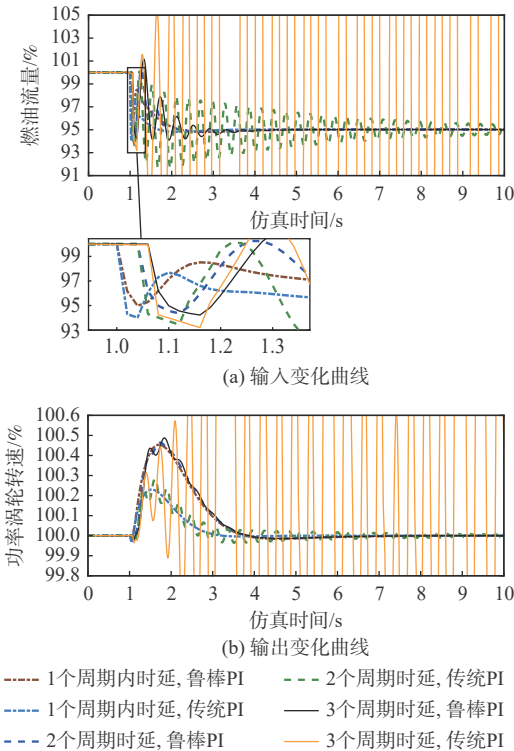


图 8 外环穿越频率为 3 rad/s, 不同时延时 2 个控制器作用下系统的性能比较

Fig. 8 Comparison of control performance of two controllers with different time delays when crossing frequency of outer loop is 3 rad/s

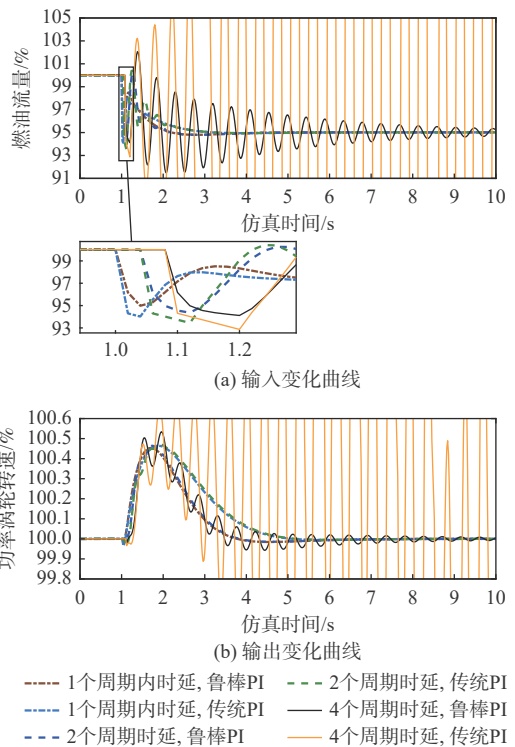


图 9 外环穿越频率为 1 rad/s, 不同时延时 2 个控制器作用下系统的性能比较

Fig. 9 Comparison of control performance of two controllers with different time delays when crossing frequency of outer loop is 1 rad/s

统控制器便不能保证系统稳定而在鲁棒串级 PI 控制器作用下系统仍能保持稳定且系统性能满足要求,且传统控制器的时延边界只有 0.077 s(不足 4 个周期),小于鲁棒串级 PI 控制器的时延边界。

同时,从图 8 与图 9 对应的仿真结果可以得到,2 个传统串级 PI 控制器的时延边界存在差异,主要是由其外环期望穿越频率的差异导致的,外环期望穿越频率越大,系统的响应速度越快而系统的鲁棒性就越差。所以,鲁棒串级 PI 控制器相比于传统串级 PI 控制器,能够保证涡轴发动机在时延条件下的性能与鲁棒性有一个更好的协调性且实现该控制所需代价更小。

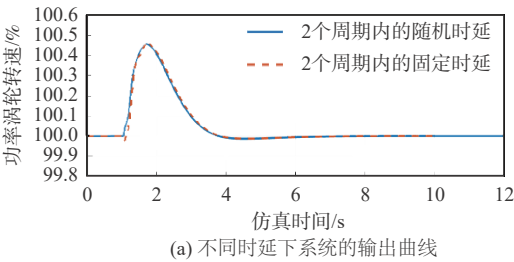
图 8 与图 9 对应的计算与仿真针对的都是固定时延,但该时延鲁棒串级 PI 控制器对于上限不超过计算设定时延(0.04 s)的随机时延也是有效的。在固定时延与随机时延下,系统的响应曲线及随机时延的大小如图 10 所示。从图 10 可以看出,在随机时延下,功率涡轮转速  $N_p$  的超调不超过 0.5%,调节时间小于 5 s。因此,随机时延不超过上限时鲁棒串级 PI 控制器仍是有效的。

明通过频域回路成形方法与 LMI 方法相结合,可以为存在时延的涡轴发动机分布式控制系统设计出鲁棒性更好的控制器以应对时延对涡轴发动机分布式控制系统性能的影响,并实现鲁棒性与动态性能更好的协调。

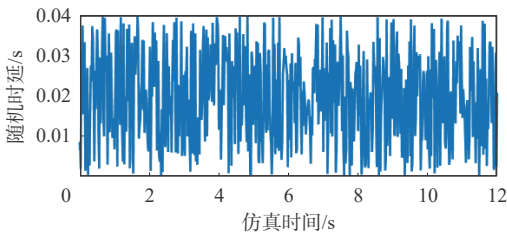
但在控制器设计过程中,约束条件的转换采用了一些保守性方法,如何解决求解方法的保守性是未来值得研究的问题。

## 参考文献 (References)

- [1] DELOSREYES G, GOUCHOE D R. The design of a turboshaft speed governor using modern control techniques: 19880000957[R]. Cleveland: NASA Lewis Research Center, 1986: 40-50.
- [2] BELAPURKAR R K, YEDAVALLI R K. LQR control design of discrete-time networked cascade control systems with time delay[C]//Proceedings of the ASME 2011 Dynamic Systems and Control Conference. New York: ASME Press, 2011: 299-304.
- [3] LIU X F, LUO C S. Stability analysis for GE T700 turboshaft distributed engine control systems[J]. *IEEE Access*, 2017, 5: 22485-22491.
- [4] CULLEY D E, PALUSZEWSKI P J, STOREY W, et al. The case for distributed engine control in turboshaft engine systems [C]//65th Annual Forum and Technology Display. Reston: AIAA, 2009, 2: 1732-1745.
- [5] CULLEY D, THOMAS R, SAUS J. Concepts for distributed engine control[C]//43rd AIAA/ASME/SAE/ ASEE Joint Propulsion Conference & Exhibit. Reston: AIAA, 2007, 7: 6910-6920.
- [6] PAKMEHR M, CHAKRAVARTHY MURALIDHAR R S, SUTTER L F, et al. Developing concepts for optimal scheduling in distributed turbine engine control systems[C]//AIAA Propulsion and Energy 2019 Forum. Reston: AIAA, 2019: 1-10.
- [7] HUANG C Z, BAI Y, LI X L. Fundamental issues in networked cascade control systems[C]// Proceedings of the IEEE International Conference on Automation and Logistics. Piscataway: IEEE Press, 2008: 3014-3018.
- [8] HUANG C Z, BAI Y, QIU Z C. PID tuning of networked cascade control system based on desired closed-loop system responses[J]. *Control & Instruments in Chemical Industry*, 2010, 37(3): 19-24.
- [9] HUANG C Z, BAI Y, LIU X J.  $H_\infty$  state feedback control for a class of networked cascade control systems with uncertain delay[J]. *IEEE Transactions on Industrial Informatics*, 2010, 6(1): 62-72.
- [10] AZAR A T, SERRANO F E. Robust IMC-PID tuning for cascade control systems with gain and phase margin specifications[J]. *Neural Computing and Applications*, 2014, 25(5): 983-995.
- [11] ALFARO V M, VILANOVA R, ARRIETA O. Robust tuning of Two-Degree-of-Freedom (2-DoF) PI/PID based cascade control systems[J]. *Journal of Process Control*, 2009, 19(10): 1658-1670.
- [12] CHEN Y F, GUO Y Q, LI R C. State feedback control for partially distributed turboshaft engine with time delay and packet dropouts [C]//Proceedings of the 37th Chinese Control Conference. Piscataway: IEEE Press, 2018: 9854-9859.
- [13] LIU X F, SUN X.  $H_\infty$  networked cascade control system design for



(a) 不同延迟下系统的输出曲线



(b) 系统的输出曲线

图 10 在固定时延与随机时延下系统性能对比

Fig. 10 Performance comparison with constant and random time delays

## 3 结论

对涡轴发动机分布式控制系统中存在时延下的鲁棒串级 PI 控制器设计方法进行了研究。得到以下结论:

1) 对于内外环同时存在时延的涡轴发动机分布式控制系统,IMC 控制器设计方法是一种有效的设计 PI 控制器的方法。

2) 在鲁棒串级 PI 控制器作用下,系统能够满足控制需求且能够容忍的时延边界扩大为 0.097 s(将近 5 个控制周期)大于传统串级 PI 控制器。说

- turboshaft engines with random packet dropouts[J]. *International Journal of Aerospace Engineering*, 2017, 2017: 1-12.
- [14] 陈义峰, 郭迎清, 李睿超, 等. 涡轴发动机分布式控制系统架构设计[J]. *航空计算技术*, 2019, 49(5): 21-26.  
CHEN Y F, GUO Y Q, LI R C, et al. Design of turboshaft engine distributed control system architecture[J]. *Aeronautical Computing Technique*, 2019, 49(5): 21-26(in Chinese).
- [15] WANG Q G, HANG C C, YANG X P. Single-loop controller design via IMC principles[J]. *Automatica*, 2001, 37(12): 2041-2048.
- [16] 潘笑, 钟祎勃. 基于IMC的PID控制器的设计实现[J]. *计算机仿真*, 2005, 22(8): 80-82.  
PAN X, ZHONG Y Q. Design and realization of PID control based on IMC [J]. *Computer Simulation*, 2005, 22(8): 80-82(in Chinese).
- [17] MAO H T, GUO Y Q, LI R C, et al. Versatile simulation platform for turboshaft engine control system[C]//*Proceedings of the 38th Chinese Control Conference*. Piscataway: IEEE Press, 2019: 7211-7216.
- [18] NAJAFIZADEGAN H, MERRIKH-BAYAT F, JALILVAND A. IMC-PID controller design based on loop shaping via LMI approach[J]. *Chemical Engineering Research and Design*, 2017, 124: 170-180.
- [19] 王海泉. 航空发动机 $H_\infty$ 鲁棒控制器设计及仿真验证[D]. 西安: 西北工业大学, 2009: 11-14.  
WANG H Q,  $H_\infty$  robust controller design for aero-engine and simulation[D]. Xi'an: Northwestern Polytechnical University, 2009: 11-14(in Chinese).
- [20] DAMEN A, WWEILAND S. *Robust control*[M]. Eindhoven: Eindhoven University of Technology Press, 2002: 61-68.
- [21] 俞立. 鲁棒控制: 线性矩阵不等式处理方法[M]. 北京: 清华大学出版社, 2002: 8-10.  
YU L. *Robust control: Linear matrix inequalities processing method*[M]. Beijing: Tsinghua University Press, 2002: 8-10(in Chinese).
- [22] LOFBERG J. YALMIP: A toolbox for modeling and optimization in MATLAB[C]//2014 IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2005: 284-289.
- [23] BALLIN M G. A high fidelity real-time simulation of a small turboshaft engine: 19880016994[R]. Moffett Field: NASA AMES Research Center, 1988: 1-2.
- [24] CERVIN A, HENRIKSSON D, LINCOLN B, et al. How does control timing affect performance? Analysis and simulation of timing using Jitterbug and TrueTime[J]. *IEEE Control Systems*, 2003, 23(3): 16-30.
- [25] 杨征山, 李胜泉, 章霖官. 涡轴发动机动力涡轮转速控制回路方案研究[J]. *航空发动机*, 2005, 31(2): 46-50.  
YANG Z S, LI S Q, ZHANG L G. Speed control loop schemes for power turbines of turboshaft engine[J]. *Aeroengine*, 2005, 31(2): 46-50(in Chinese).

## Design of time-delay robust cascade PI controller for turboshaft engine

CHEN Yifeng, GUO Yingqing\*, MAO Haotian

(School of Power and Energy, Northwestern Polytechnical University, Xi'an 710072, China)

**Abstract:** To address the problem of system performance degradation caused by time delay in the distributed control system of turboshaft engines, this study proposes a method for designing a time-delay robust cascade PI controller based on linear matrix inequalities (LMIs) method. Firstly, the PI controller structure of the inner and outer loops of the cascade controller of the turboshaft engine is obtained according to the internal model control (IMC) method. The constraints in the form of LMI are given by the frequency loop shaping method, which ensures that the system has the desired performance. Then, gradient approximation method is used to derive stability constraints of the system based on the Routh-Hurwitz criterion. Finally, the digital simulation is carried out on the distributed nonlinear simulation platform for the turboshaft engine. The simulation results show that when the power demand drops by 5% with 0.04 s time delay, the adjustment time of the system is less than 5 s, and the power turbine speed overshoot does not exceed 0.5%, with the maximum fuel rate being 67% of the legacy control loop. This shows that the proposed method can effectively deal with the time delay in the turboshaft engine with distributed control system, and ensure the desired system performance at a lower cost.

**Keywords:** distributed control system of turboshaft engine; time delay; robust cascade PI controller; linear matrix inequalities; frequency loop shaping; TrueTime simulation

Received: 2021-05-26; Accepted: 2021-08-20; Published Online: 2021-09-13 10:16

URL: [kns.cnki.net/kcms/detail/11.2625.V.20210911.1227.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210911.1227.001.html)

Foundation item: National Science and Technology Major Project (2017-V-0015-0067)

\* Corresponding author. E-mail: [yqguo@nwpu.edu.cn](mailto:yqguo@nwpu.edu.cn)

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0274

# 多种群合作学习的多模态多目标路径规划算法

赵萌, 路辉\*, 王诗琪, 杨思旖, 王赞

(北京航空航天大学 电子信息工程学院, 北京 100191)

**摘要:** 为同时规划出满足多种目标需求的多条可行路径, 提高规划路径的鲁棒性与实用性, 提出一种基于多种群合作学习的路径规划算法。基于粒子群算法的基本思想, 先针对单一种群在多维目标空间内搜索时容易陷入局优的问题, 提出基于多目标分解的子种群划分策略, 平衡算法在目标空间内各个维度上的搜索能力。再依据地图中栅格点的出入度信息提取关键路径点。在编码阶段, 根据关键路径点提供的维度信息, 利用实数编码的方式初始化种群, 降低解空间大小; 在解码阶段, 提出利用精英解的解码经验指导可行解的快速搜索, 使解码经验能够被有效传递, 降低解码的不确定性, 提高了算法的寻优能力。最后, 将多个种群的搜索结果进行非支配排序, 得到满足优化目标的所有路径。实验结果表明: 与标准粒子群算法相比, 基于解码经验表指导的多种群合作学习算法具有更强的搜索能力和寻优能力, 能够解决多模态多目标路径规划问题。

**关键词:** 多模态多目标优化; 路径规划; 粒子群算法; 子种群划分; 解码经验表

**中图分类号:** TP181

**文献标志码:** A

**文章编号:** 1001-5965(2023)03-0606-11

多模态优化问题旨在找到多个全局或局部最优解, 为决策者提供多样的选择, 增加决策的鲁棒性; 多目标优化问题是指通过权衡多个目标的最优化, 为决策者提供更好的选择, 提高决策的实用性, 与单目标的全局最优点不同, 多目标优化的最优解通常用 Pareto 解集表示, 集合中的解是相互非支配的。多模态多目标优化 (multimodal multi-objective optimization, MMO) 问题指具有多个全局或局部最优 Pareto 解集的优化问题<sup>[1]</sup>, 在实际应用中, 每个解集包含多种备选方案。丰富的决策方案不仅可以保障决策的有效性, 还可以减少突发事件导致的经济损失。

在路径规划问题中, 决策者希望可以同时获得满足目标需求的多条路径, 从而保障决策的稳定性。目前大多数研究针对最小化路径长度的规划需求, 在满足约束条件的前提下同时规划出多条最短路径<sup>[2-3]</sup>, 这是一种单目标多模态的优化问题。然

而, 单一目标通常难以满足人们的实际需求, 考虑到路网环境的不确定性, 人们希望可以同时获得多条路径长度、堵塞路段数量、交叉路口数量等目标值最优且相等的路径。因此, 在路径规划问题中, 多模态优化与多目标优化具有紧耦合的关系。虽然已经有学者对多目标路径规划问题进行了深入研究<sup>[4-7]</sup>, 但是由于问题建模复杂, 求解难度大, 关于多模态多目标路径规划问题 (multimodal multi-objective optimization path planning, MMOPP) 的研究成果却寥寥无几。尽管多式联运问题的研究正处于白热化阶段, 但是这些研究往往只考虑运输成本和时间成本<sup>[8-9]</sup>, 缺少在多模态优化需求条件下, 针对复杂路网中超多目标优化问题的研究。

从应用需求的角度分析, MMOPP 问题的研究通常分为 2 种。一种是针对复杂任务中的路径规划<sup>[10-12]</sup>, 决策者需要同时考虑路径长度短、有途经点等规划需求。另外一种考虑物流配送需求的

收稿日期: 2021-05-27; 录用日期: 2021-08-22; 网络出版时间: 2021-09-29 16:45

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210929.1105.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210929.1105.001.html)

基金项目: 国家自然科学基金 (61671041); 陕西省组合与智能导航重点实验室开放基金 (SKLIIN-20190201)

\*通信作者. E-mail: [mluhui@buaa.edu.cn](mailto:mluhui@buaa.edu.cn)

**引用格式:** 赵萌, 路辉, 王诗琪, 等. 多种群合作学习的多模态多目标路径规划算法 [J]. 北京航空航天大学学报, 2023, 49 (3): 606-616.  
ZHAO M, LU H, WANG S Q, et al. A multimodal multi-objective path planning algorithm based on multi-swarm cooperative learning [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 606-616 (in Chinese).

车辆路径规划<sup>[13-14]</sup>, 这种问题的研究通常不需要考虑路径的多样性, 主要考虑配送车辆的装载量约束与配送网点的时间约束。本文针对复杂路网中 MMOPP 问题的研究需求, 在多模态优化需求条件下, 研究超多目标路径规划问题。相关研究成果可以为车辆或无人载体在复杂路网中的路径规划提供理论与技术支撑。

经典的路径规划算法, 如 A\* 算法、Dijkstra 算法等仅考虑路径长度的优化, 难以实现对 MMOPP 问题的求解。因此, 大多数学者引入具有自组织、自适应、自学习特性的群体智能算法实现多目标路径规划。文献 [15] 将蚁群算法与 Dijkstra 算法相结合, 通过丰富路径搜索方向, 解决了路径长度短、路径平滑度高、机器人行进安全性高 3 种路径规划需求。文献 [16] 提出用强化学习算法优化蚁群算法的参数, 提高了算法的全局搜索能力。除蚁群算法外, 粒子群优化 (particle swarm optimization, PSO) 算法<sup>[17]</sup>、灰狼优化算法<sup>[18]</sup>、萤火虫算法<sup>[19]</sup>、天牛群算法<sup>[20]</sup> 等也可用于解决多目标路径规划问题。针对无人机三维作业场景的特殊性, 文献 [21] 通过在 NSGA-II 算法中引入增加、删除算子, 同时解决了路径长度短与躲避威胁区域 2 种路径规划需求。此外, 无人机燃料损耗的最小化也通常是路径规划问题中需要考虑的优化目标之一<sup>[22]</sup>。

大多数算法只考虑到了路径规划中的多目标需求, 尽管这类算法的搜索能力强, 收敛性好, 但是由于算法设计过程中没有考虑多模态的规划需求, 决策空间内解的多样性难以得到保持, 决策的鲁棒性受限。因此, 为提升算法的搜索能力与寻优能力, 提高解的多样性, 本文在 MMOPP 问题的研究

中侧重解决路径规划问题中的多模态需求, 在 PSO 算法框架下, 提出多种群合作学习 (multi-swarm cooperative learning, MSCL) 算法。利用多个种群的合作搜索, 将多目标优化问题转化为多个单目标问题进行求解, 丰富决策空间内解的多样性, 同时, 兼顾算法在目标空间不同维度上的搜索能力。此外, 为了缓解编解码匹配的不唯一性引起的搜索效率低的问题, 提出利用解码经验表传递有效的解码经验, 以提高算法的寻优能力。算法验证阶段使用的 MMOPP 测试集包括对路径长度、道路宽度、道路拥挤度、堵塞路段等待优化目标的数值量化, 并且包括必经点问题的求解约束。因此, 本文的仿真测试实验与实际应用场景具有紧耦合的关系。仿真测试结果验证了多种群合作学习算法可以并行解决多模态和多目标的路径规划需求, 与其他算法的比较结果说明了本文所提算法具有更强的搜索能力与寻优能力, 并且解的多样性更好。

### 1 多种群合作学习算法

多种群合作学习算法是利用多个种群的无交叉搜索并行解决多模态和多目标的路径规划需求。如图 1 所示, 其中  $f_1, f_2, \dots, f_n$  为不同的优化目标。算法实现的过程分为 3 个阶段: 首先, 依据地图中关键路径点的数量对路径码进行编码; 然后, 将初始化后的种群划分为多个子种群, 明确每个子种群的优化目标; 最后, 依据子种群的优化目标对个体的路径码进行解码, 计算相应的适应度值, 并保留精英解对应的路径。整个过程持续循环, 直到满足最大迭代次数。

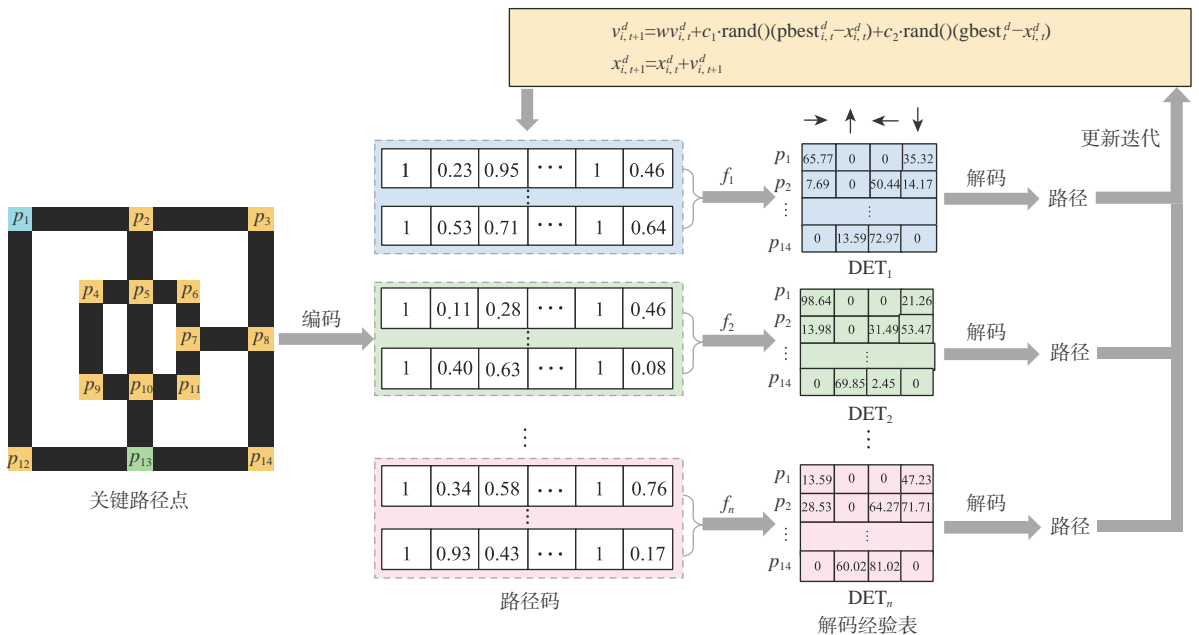


图 1 多种群合作学习算法框架

Fig. 1 Framework of the proposed algorithm

### 1.1 基于关键点的路径编码

为了降低可行解的搜索难度,利用地图中的关键路径点表征路径信息。如图2所示,用1、2、3、4分别表示路径节点与其右、上、左、下的节点具有邻接关系。假设任意路径节点 $r_j$ 处的可行前进方向集合 $W_j$ ,如果 $r_j$ 属于关键路径点集合 $P = \{p_1, p_2, \dots, p_K\}$ ,则一定满足以下条件之一:

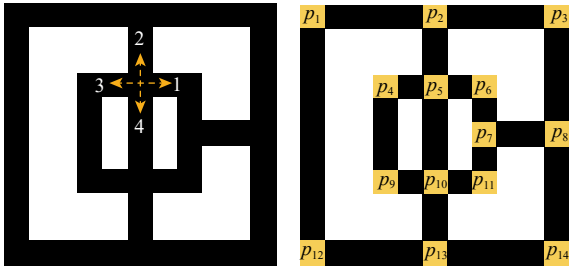


图2 关键路径点的提取

Fig. 2 Extraction of key path points

条件1  $\text{card}(W_j) > 2$ ;

条件2 如果 $\text{card}(W_j) = 2$ ,则 $W_j \neq \{1, 3\}$ 且 $W_j \neq \{2, 4\}$ ;

条件3  $r_j \in \{S, T\}$ 。

其中 $j = 1, 2, \dots, R$ ,  $R$ 为地图所有路径节点的数量,  $K$ 为关键路径点的数量,且 $K < R$ 。  $S$ 和 $T$ 分别为路径的起点和终点。

显然,关键路径点的提取不会影响可行路径的搜索,而且可以通过减小解空间的维度提高可行解的搜索效率。因此,采用地图中关键路径点数量

$K$ 定义粒子编码的维度,即粒子个体是 $K$ 维的实数码,记为路径码。

### 1.2 基于多目标分解的子种群划分

在多种群合作学习算法中,种群的初始化采用随机数生成器的方式,即 $v_{i1}^d = \text{rand}(0, 1)$ ,  $x_{i1}^d = \text{rand}(0, 1)$ ,其中, $d = 1, 2, \dots, K$ 。然而,如图3所示,当有3个优化目标(最小化)时,由于目标空间内的最优解在同一区域,所以,种群中所有粒子的搜索方向是一致的,从任意目标空间的切面看,受全局最优解的牵引,粒子均朝向左下角(最小值)移动。然而,多模态路径规划的目标不是找到一个全局最优解,而是需要综合多个目标,搜索到多个最优解。当优化目标较多时,种群中的粒子具有朝向同一目标区域的搜索趋势,容易出现早熟现象,即单一种群有限的搜索能力会影响解的多样性。因此,为了提高算法的搜索能力与寻优能力,本文提出基于多目标分解的子种群划分策略。

根据 Pareto 解集的定义,某个解在单一目标维度上最优,即有机会成为 Pareto 最优解。因此,在多种群合作学习算法中,种群会依据优化目标的个数进行均匀划分,每个子种群在迭代寻优的过程中只负责优化一个目标,如图4所示,所有粒子具有相同的优化目标,其在目标空间内的运动具有一致性。当有3个优化目标(最小化)时,每个种群分别负责在目标1、目标2、目标3的维度内搜索最优解,从而避免由于粒子在目标空间内的一致性移动

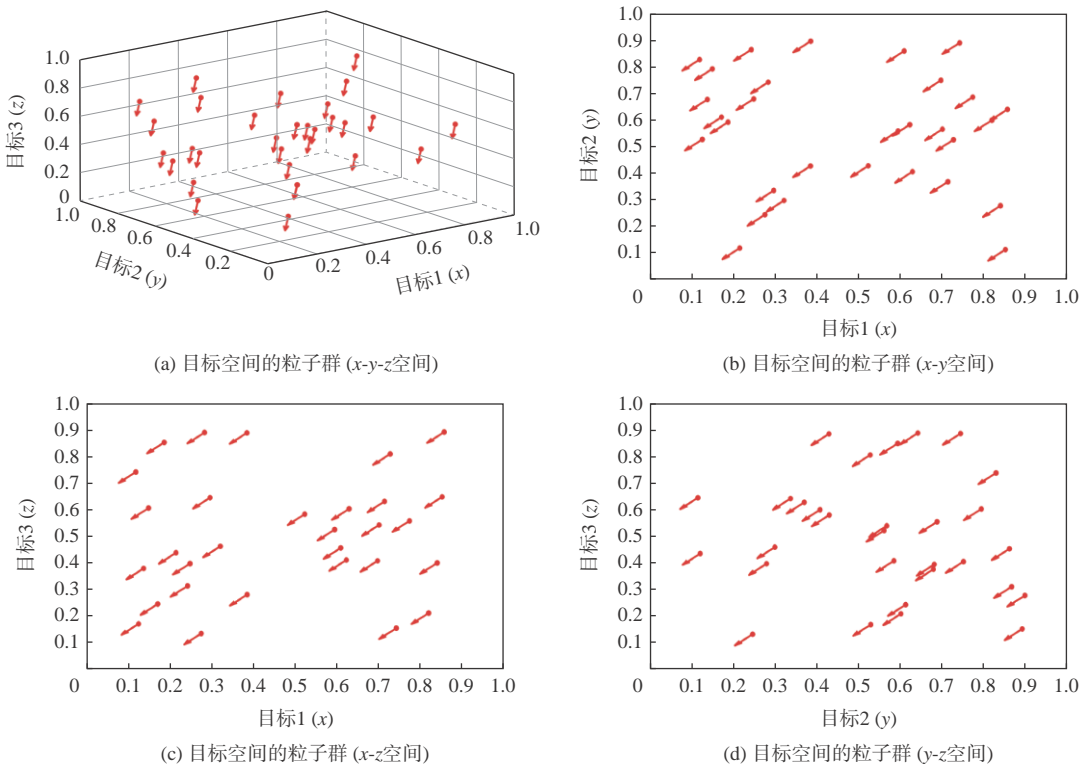


图3 有3个优化目标时单一种群的粒子群寻优过程

Fig. 3 Optimization process of particles when there are three objectives to be optimized

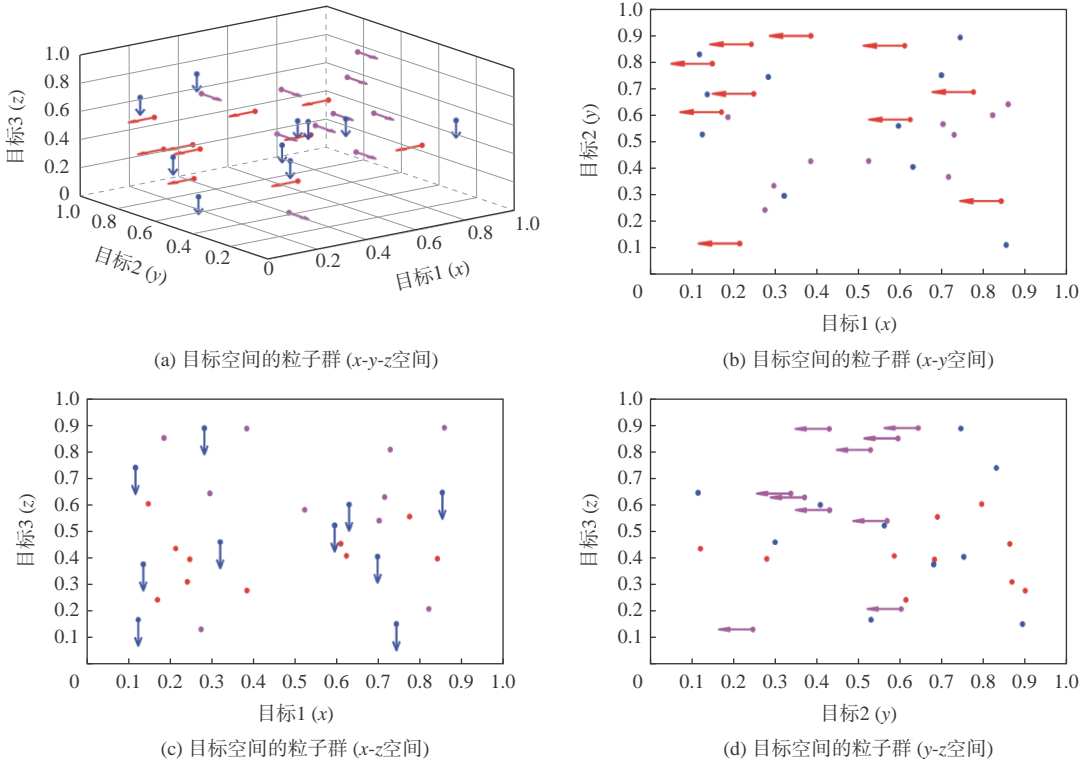


图4 有3个优化目标时基于子种群划分策略的粒子群寻优过程

Fig. 4 Optimization process of particles when there are three objectives to be optimized

导致陷入局优的问题。多种群分工协作的机制使得粒子可以在目标空间内分散地搜索,趋向不同的方向寻找最优解,进而实现多目标的分维寻优,提高解的多样性。因此,在多模态多目标路径规划问题的求解中具有更强的搜索能力与寻优能力。

### 1.3 基于粒子群的迭代过程

在多种群合作学习算法中,种群中的每个个体根据标准粒子群优化算法更新迭代自身的速度与位置<sup>[23]</sup>。粒子群优化算法是 Eberhart 和 Kennedy<sup>[24]</sup>于 1995 年提出的一种基于群体协作的随机搜索算法。该算法的基本思想是模拟鸟群的觅食行为,通过粒子群在解空间内的搜索实现对优化问题的求解。如式 (1) 和式 (2) 所示,种群中粒子的状态用速度和位置 2 个向量描述:

$$v_{i,t} = [v_{i,t}^1, v_{i,t}^2, \dots, v_{i,t}^K] \quad i = 1, 2, \dots, N \quad (1)$$

$$x_{i,t} = [x_{i,t}^1, x_{i,t}^2, \dots, x_{i,t}^K] \quad i = 1, 2, \dots, N \quad (2)$$

式中:  $N$  和  $D$  别为种群中粒子的数量和解空间的维度;  $t$  为粒子的迭代次数。每个粒子根据自己的历史最优解和全局最优解的位置调整移动速度,如图 5 所示,速度与位置的更新过程可以由式 (3) 和式 (4) 进行描述:

$$v_{i,t+1}^d = wv_{i,t}^d + c_1 \cdot \text{rand}() (\text{pbest}_{i,t}^d - x_{i,t}^d) + c_2 \cdot \text{rand}() (\text{gbest}_{i,t}^d - x_{i,t}^d) \quad (3)$$

$$x_{i,t+1}^d = x_{i,t}^d + v_{i,t+1}^d \quad (4)$$

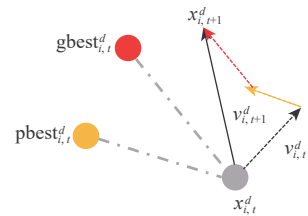


图5 粒子的寻优过程

Fig. 5 The optimization process of the particles.

式中:  $v_{i,t}^d$  和  $x_{i,t}^d$  分别为第  $i$  个粒子在第  $t$  次迭代时第  $d$  维度的速度和位置;  $\text{pbest}_{i,t}^d$  和  $\text{gbest}_{i,t}^d$  分别为粒子历史最优解和种群全局最优解的位置;  $c_1$  和  $c_2$  分别为认知加速系数和社会加速系数,反映粒子朝向  $\text{pbest}_{i,t}^d$  和  $\text{gbest}_{i,t}^d$  移动的能力;  $w$  为惯性因子,用于反映粒子保持原速不变的能力,是提升算法搜索能力的关键参数<sup>[23]</sup>;  $\text{rand}()$  为 0 到 1 之间的随机数。

### 1.4 基于解码经验表的路径解码

根据路径规划的需求,从路径起点对应的维度开始解码,然后依据每个关键路径点处的可行进方向依次选择下一个路径点,直至当前的解码维度对应于终点,或重复到达某一路径节点,则停止对该粒子个体(路径码)的解码。如果可连续解码至终点,则得到一条可行路径,计算路径对应的各个目标值,经过非支配排序得到精英解。然而,地图中每个关键路径点都有多个可行进方向,路径码的解码空间不能一一映射地回溯至编码空间。针对这

种编解码匹配的不唯一性,提出采用解码经验表传递有效的解码经验,降低解码的不确定性。如图6所示,解码经验表 $D$ 是 $T \times M$ 的二维矩阵,其中, $T$ 为路径码的维度, $M$ 为关键路径点的最大可行进方向数量,即 $M = \max(\text{card}(W_v))$ ,其中 $v = 1, 2, \dots, T$ 。在本文使用的测试问题(test problem, TP)中, $M = 4$ 。解码经验表中的初始数值为0,在粒子解码过程中,表中数值的更新依据式(5)所示的Bellman方程:

$$D(p_t, a_t) = (1 - \alpha)D(p_t, a_t) + \alpha [R(p_t, a_t) + \gamma \max_a (D(p_{t+1}, a))] \quad (5)$$

式中: $\alpha$ 和 $\gamma$ 分别为学习率和折扣率;在路径解码问题中, $D(p_t, a_t)$ 和 $R(p_t, a_t)$ 分别为在当前解码维度对应的关键路径点 $p_t$ 处选择动作 $a_t$ 得到的经验值与奖励值。奖励值 $R(p_t, a_t)$ 的设计与优化目标有关,具体可以分为如下三类:

$$1) f = \min(L)$$

$$R(p_t, a_t) = \begin{cases} 500 & p_{t+1} = T \\ 10h(p_{t+1}) & p_{t+1} \in E \\ -1 & \text{其他} \end{cases} \quad (6)$$

式中: $L$ 为路径长度; $h(p_{t+1})$ 为 $p_{t+1}$ 在精英解得到的可行路径中出现的频次; $E$ 为所有精英解对应的路径中出现的关键路径点集合。

$$2) f = \min \left( \sum_{i=1}^q F_1^{p_i} \right)$$

$$R(p_t, a_t) = \beta \frac{1}{F_1^{p_{t+1}}} \quad (7)$$

式中: $q$ 为从起点到终点的路径点总数; $F_1^{p_i}$ 为路径点 $p_i$ 的 $F_1$ 值(表示路径宽度、拥挤度等信息); $\beta$ 为用于调节奖励值大小的常数。

$$3) f = \min(q)$$

$$R(p_t, a_t) = \begin{cases} -1 & Q(p_{t+1}) = 0 \\ \delta \frac{1}{Q(p_{t+1})} & \text{其他} \end{cases} \quad (8)$$

式中: $q$ 为从起点到终点的路径点总数; $Q(p_{t+1})$ 为经过 $p_{t+1}$ 的所有可行路径中最小路径点总数,如果所有路径都没有通过 $p_{t+1}$ ,则 $Q(p_{t+1}) = 0$ ;  $\delta$ 为调节奖励值大小的常数。

	1(右)	2(上)	3(左)	4(下)
$p_1$	$D(p_1, a_1)$	$D(p_1, a_2)$	$D(p_1, a_3)$	$D(p_1, a_4)$
$p_2$	$D(p_2, a_1)$	$D(p_2, a_2)$	$D(p_2, a_3)$	$D(p_2, a_4)$
$p_3$	$D(p_3, a_1)$	$D(p_3, a_2)$	$D(p_3, a_3)$	$D(p_3, a_4)$
$p_4$	$D(p_4, a_1)$	$D(p_4, a_2)$	$D(p_4, a_3)$	$D(p_4, a_4)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$p_D$	$D(p_D, a_1)$	$D(p_D, a_2)$	$D(p_D, a_3)$	$D(p_D, a_4)$

图6 解码经验表

Fig. 6 Decoding experience table

此外,为了保障算法的寻优能力,路径码的解码过程不仅参考解码经验表中的经验值,也会依概率 $\varepsilon$ 随机选择下一个关键路径点。

1.5 算法流程

多种群合作学习算法流程如图7所示。

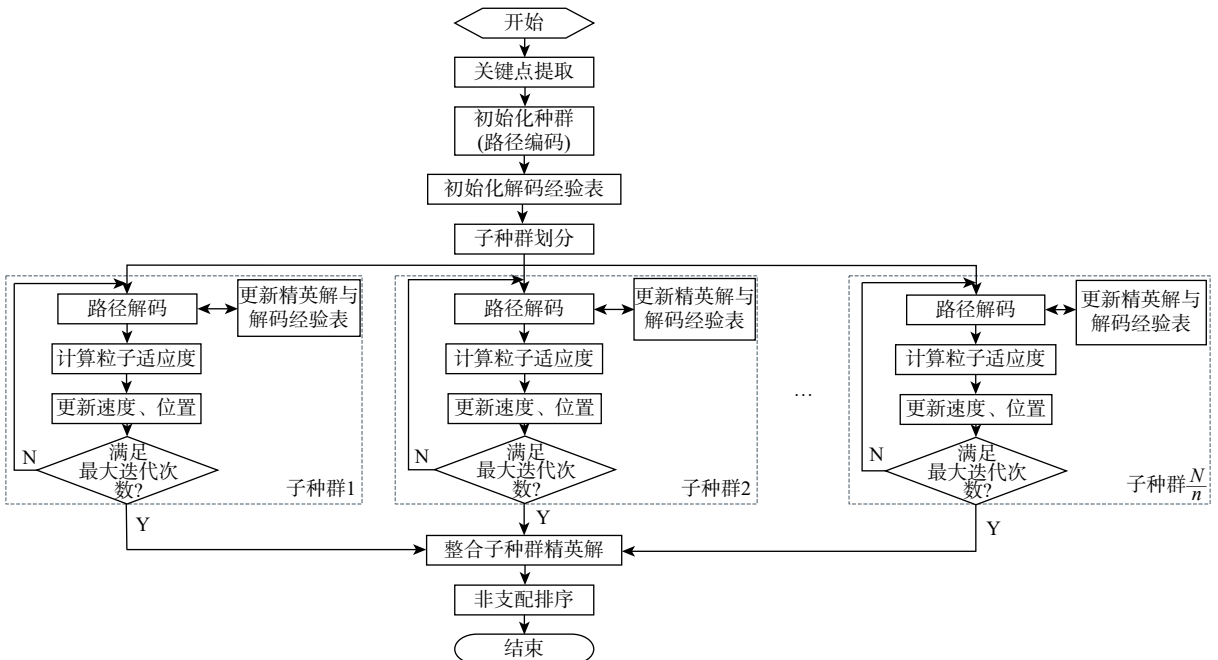


图7 一个多模态多目标路径规划问题的应用实例

Fig. 7 An application example of MMOPP problem

## 2 多种群合作学习算法应用实例

为了进一步说明多种群合作学习算法的作用机理,以图8所示面积大小为11×11的栅格地图为例,展示粒子在多模态多目标路径规划问题中的编码与解码过程。其中,路径规划的起点与终点分别为 $p_1$ 和 $p_{13}$ ,图8中红色区域模拟路网中的拥挤堵塞路段,则优化问题可描述为

$$\min y = (f_1(x), f_2(x), f_3(x)) \quad (9)$$

式中: $x$ 为可行路径; $f_1(x)$ 、 $f_2(x)$ 和 $f_3(x)$ 分别为路径的长度、经过的红色区域数量和经过的交叉点数量。

### 2.1 路径编码

根据地图中关键点数量提供的决策变量维度,粒子个体将在种群初始化阶段完成编码,如图9所示,种群中每个粒子个体都是一个14维的实数码。

	$p_1$	$p_2$	$p_3$	$p_4$	$p_5$	$p_6$	$p_7$	$p_8$	$p_9$	$p_{10}$	$p_{11}$	$p_{12}$	$p_{13}$	$p_{14}$
种群1	0.28	0.18	0.80	0.08	0.84	0.70	0.77	0.54	0.17	0.55	0.23	0.05	0.63	0.96
种群2	1.00	0.59	0.33	0.08	0.06	0.27	0.37	0.17	0.25	0.15	0.71	0.70	0.35	0.54
种群3	0.46	0.79	0.58	0.15	0.13	0.14	0.63	0.08	0.06	0.53	0.37	0.51	0.91	0.22
	0.52	0.20	0.78	0.52	0.24	0.83	0.27	0.80	0.01	0.83	0.28	0.96	0.60	0.62
	0.42	0.59	0.30	0.79	0.56	0.53	0.45	0.94	0.33	0.19	0.18	0.20	0.78	0.91

图9 种群(路径码)初始化结果

Fig. 9 Result of initialing particle swarm (path codes)

### 2.2 路径解码

在路径解码时,从起点开始基于解码经验表中的经验信息进行解码,图10为以种群1为例的解码过程。种群中每个个体的路径码解码都是从起

点 $p_1$ 开始,依概率 $\varepsilon(0 < \varepsilon < 1)$ 选择下一个路径点,即有 $\varepsilon$ 的概率在邻接点中随机选择下一个路径点,有 $(1-\varepsilon)$ 的概率依据解码经验选择下一个路径点,如果解码经验表中邻接关键点的经验值相同,则随机

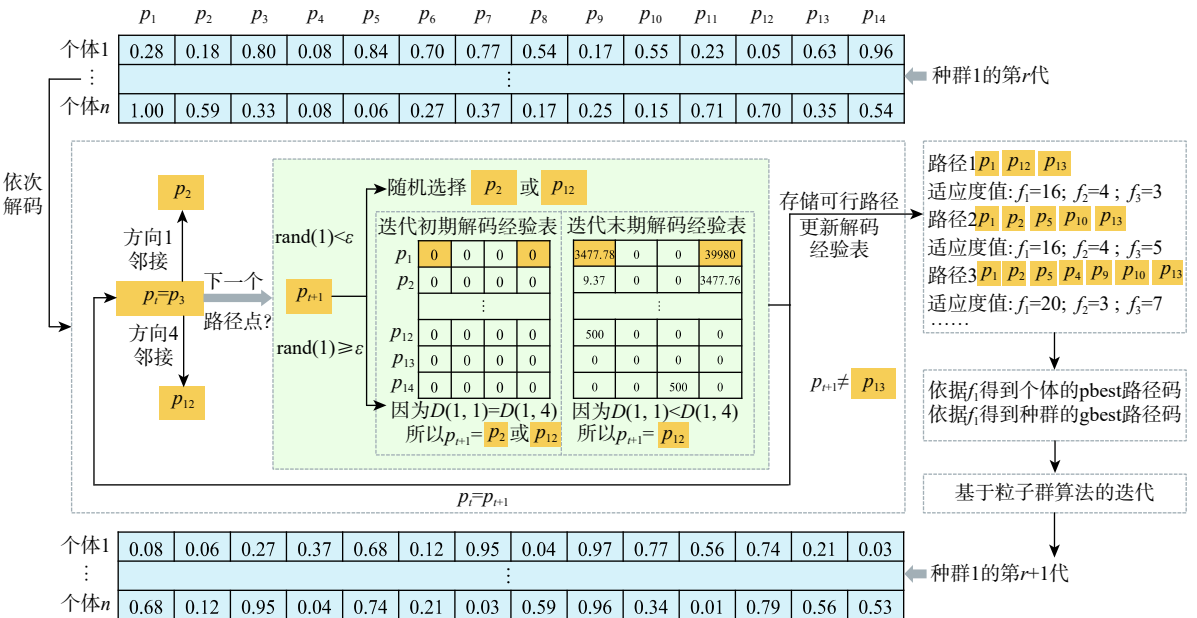


图10 以种群1为例的路径解码过程

Fig. 10 Path decoding process of the first sub-swarm that is responsible for optimizing path length

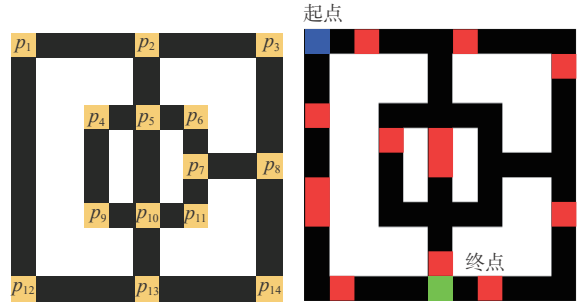


图8 多种群合作学习算法流程

Fig. 8 Flowchart of multi-swarm cooperative learning algorithm

考虑到路径规划时有3个待优化目标,因此,种群被划分为3个子种群,每个种群分别负责搜索最短路径长度、最少红色区域数量、最少交叉点数量对应的路径。

选择一个作为下一个路径点,如果经验值不同,则选择经验值最大的一个作为下一个路径点。依照此策略持续解码并更新解码经验表,直至下一个路径点是终点,存储可行路径。当个体解码完成时,更新个体的  $pbest_{i,t}$ ; 当种群中所有个体解码完成时,更新种群的  $gbest_t$ 。由于种群 1 的优化目标是 minimized 路径长度,因此,  $pbest_{i,t}$  与  $gbest_t$  的更新仅依据目标值  $f_1(x)$ 。最后,基于粒子群算法的速度更新公式更新个体位置。整个过程循环迭代直至最大迭代次数。

### 3 仿真测试与验证分析

为了验证多种群合作学习算法在 MMOPP 问题上的有效性,本文采用 Liang 等<sup>[25]</sup> 构建的 MMOPP 测试问题集检测算法的性能。此测试集将实际路网地图中的道路特征建模成规则的栅格地图,利用黑色栅格表示可行路段,根据优化目标种类的不同分为三类,共计 12 个测试问题。如图 11 所示,第 1 类测试问题模拟实际路网地图中的拥挤路段,红色栅格表示拥堵路段,其优化目标为最小化路径长度、拥挤路段数量、交叉点数量,根据问题规模的不同,这类问题包含 5 个测试问题。在第 2 类测试问题中,如图 12 所示,其有多个表征不同类型道路信息的  $F$  值( $F_1, F_2, \dots$ ),用不同的  $F$  值模拟道路拥挤度或道路宽度等信息,优化目标为最小化路径长度和各个  $F$  值,根据优化目标数量的不同,这类问题包含 5 个测试问题。图 13 为第 3 类测试问题模拟了有必经点约束的路径规划需求,而且允许折返路径存在,黄色栅格表示路径规划过程中的必经路段,与第 2 类问题类似,每个栅格都有多种不同的  $F$  值( $F_1, F_2, \dots$ ),其优化目标为最小化路径长度和各个  $F$  值,根据问题规模的不同,这类问题包含

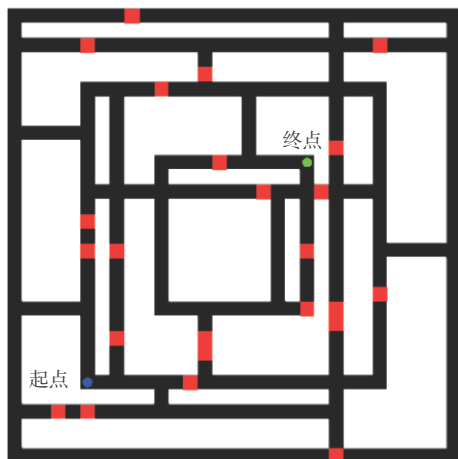


图 11 第 1 类 MMOPP 测试问题实例

Fig. 11 An example of the first kind of MMOPP test problem

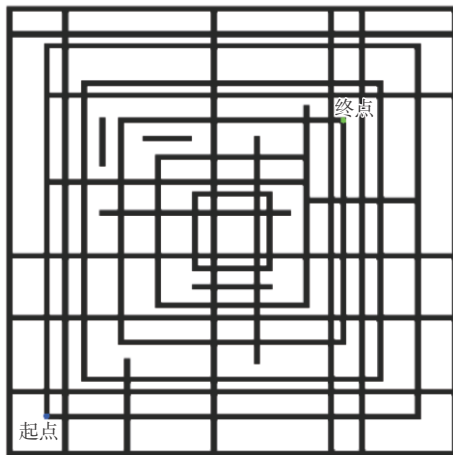


图 12 第 2 类 MMOPP 测试问题实例

Fig. 12 An example of the second kind of MMOPP test problem

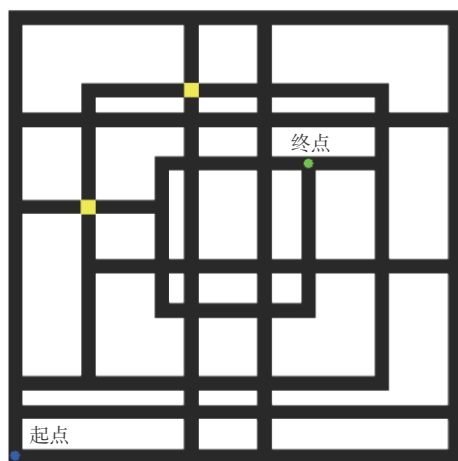


图 13 第 3 类 MMOPP 测试问题实例

Fig. 13 An example of the third kind of MMOPP test problem

2 个测试问题。因此,本文基于本节 12 个 MMOPP 测试问题,从求解时间、搜索能力、寻优能力三方面对算法的性能进行评估,并将实验结果与 PSO 算法、融合子种群划分策略的粒子群算法(PSOK)、融合解码经验表的粒子群算法(PSOD)进行对比,进一步分析验证子种群划分、解码经验表策略的有效性。

本文的所有实验结果都是 50 次实验的平均值,针对同一个测试问题,不同算法的参数、种群大小、迭代次数完全相同。然而,考虑到迭代次数与问题规模大小相关,在本文的仿真测试实验中,针对不同测试问题搜索最优路径所需要的迭代次数是不同的,迭代次数的设置依据问题的规模大小与算法的稳定性表现。具体实验参数设置如表 1 所示。

表 2 为从不同算法搜索时间的对比结果,可以看出,由于在 MSCL 算法中,每个种群内部自主选择个体最优解与群体最优解,并且在粒子个体解码过程中有效经验的总结与利用需要耗费额外的时间,因此,算法的求解时间比其他算法长。尽管

表 1 实验参数  
Table 1 Experimental parameters

参数	数值	测试问题	迭代次数
$c_1$	2	TP1	5
$c_2$	2	TP2	10
$v_{\min}$	0.4	TP3	15
$v_{\max}$	0.9	TP4	10
$x_{\min}$	0	TP5	150
$x_{\max}$	1	TP6	10
$\alpha$	0.9	TP7	50
$\gamma$	0.1	TP8	300
$\varepsilon$	0.5	TP9	300
$\delta$	100	TP10	300
$\beta$	10	TP11	10
$n$	50	TP12	300

表 2 不同算法的求解时间

Table 2 Solving time of different algorithms

测试问题	求解时间/s			
	PSO	PSOK	PSOD	MSCL
TP1	<b>0.07</b>	0.07	0.11	0.13
TP2	<b>0.16</b>	0.18	0.27	0.36
TP3	<b>0.26</b>	0.29	0.43	0.63
TP4	<b>0.18</b>	0.20	0.30	0.41
TP5	<b>2.89</b>	3.26	6.34	8.09
TP6	<b>0.14</b>	0.15	0.21	0.25
TP7	<b>0.79</b>	0.90	1.52	1.47
TP8	<b>8.21</b>	9.05	21.99	16.60
TP9	<b>11.30</b>	11.81	35.35	21.66
TP10	<b>24.32</b>	25.63	60.52	40.94
TP11	<b>0.18</b>	0.20	0.31	0.33
TP12	<b>5.41</b>	6.21	10.88	12.05

表 3 不同算法搜索到的路径数量

Table 3 Number of paths searched by different algorithms

测试问题	路径数量			
	PSO	PSOK	PSOD	MSCL
TP1	7.08	7.24	8.56	<b>8.66</b>
TP2	8.14	7.86	10.44	<b>12.02</b>
TP3	3.52	3.48	5.86	<b>3.56</b>
TP4	5.38	5.56	7.28	<b>7.78</b>
TP5	5.38	5.1	11.26	<b>16.86</b>
TP6	4.5	4.54	<b>4.62</b>	<b>4.62</b>
TP7	15.32	15.32	12.98	<b>15.58</b>
TP8	44.34	45.02	33.66	<b>44.1</b>
TP9	73.32	74.54	51.64	<b>81.68</b>
TP10	111	108.8	150.02	<b>215.42</b>
TP11	3.98	<b>4</b>	<b>4</b>	<b>4</b>
TP12	7.88	7.62	10.5	<b>12</b>

表 4 不同算法的搜索到的最优路径数量对比

Table 4 Number of optimal paths searched by different algorithms

测试问题	最优路径数量			
	PSO	PSOK	PSOD	MSCL
TP1	5.72	5.82	7.48	<b>7.78</b>
TP2	6.74	6.82	9.82	<b>11.62</b>
TP3	0.54	0.58	2.14	<b>2.58</b>
TP4	3.2	3.46	7.08	<b>7.34</b>
TP5	3.4	2.76	10.48	<b>15.28</b>
TP6	4.4	4.34	<b>4.56</b>	4.5
TP7	15.1	15.16	12.08	<b>15.38</b>
TP8	30.52	31.48	22.32	<b>37.74</b>
TP9	68.48	69.26	48.54	<b>79.98</b>
TP10	73.82	72.78	109.16	<b>179.56</b>
TP11	3.96	<b>4</b>	<b>4</b>	<b>4</b>
TP12	5.64	5.08	9.8	<b>11.94</b>

MSCL 算法的搜索时间较长,但是,不同算法的搜索时间依旧在同一量级。此外,为了分析比较算法的搜索能力与寻优能力,在表 3 和表 4 中分别记录了不同算法完成最大迭代次数时搜索到满足优化需求的路径总数和最优路径数量。这 2 个指标可以反映算法的搜索能力与寻优能力。其中,最优路径是 50 次实验的所有结果经过非支配排序后得到的。在每个测试问题中,性能最好的算法结果加粗表示。从表 3 和表 4 所示结果可以看出,当实验条件完全相同时,在测试问题 11 中,PSOD 算法和 MSCL 算法性能相同。除测试问题 6(TP6)外,本文提出的 MSCL 算法可以搜索到更多满足多目标规划需求的路径,并且最优路径的数量明显多于其他算法的规划结果。这是由于 MSCL 算法采用了子种群划分策略,平衡了粒子在各个目标维度上的搜索能力,并且经验解码表使粒子在搜索过程中可以有效地利用精英解提供的解码经验。因此,多个种群的合作学习具有更强的寻优能力。此外,本文对

比最优路径在搜索到的总路径中所占的比例,并在表 5 中展示了实验结果。从最优路径占比的结果可以看出除测试问题 6(TP6)外,MSCL 算法的路径规划结果中最优路径占比最多。

测试问题 11(TP11)只有 2 个优化目标,并且在 40×40 栅格大小的决策空间内最优路径数量很少。因此,问题规模小,求解难度小,不同算法的性能差别较小。然而,对于测试问题 6(TP6),PSOD 算法与 MSCL 算法搜索到的路径数量相等,但是 PSOD 搜索到的路径中最优路径更多,其最优路径所占的比例比 MSCL 算法多 1.2%。为了分析此现象产生的原因,对不同测试问题的求解难度从优化目标数量、地图变长、编码维度、最优路径数量 4 个维度进行分析,通过非支配排序的方式,得到如表 6 所示的问题难度等级,等级越低,问题求解越容易。由于测试问题 11 和问题 12 具有必经点约束,且规划要求不同,为保证分析结果的公平性,未分析这

表5 不同算法搜索到的最优路径占比

Table 5 Proportion of optimal paths in paths searched by different algorithms

测试问题	最优路径占比/%			
	PSO	PSOK	PSOD	MSCL
TP1	63.56	64.67	83.11	<b>86.44</b>
TP2	51.85	52.46	75.54	<b>89.38</b>
TP3	13.50	14.50	53.50	<b>64.50</b>
TP4	40.00	43.25	88.50	<b>91.75</b>
TP5	17.89	14.53	55.16	<b>80.42</b>
TP6	88.00	86.80	<b>91.20</b>	90.00
TP7	94.38	94.75	75.50	<b>96.13</b>
TP8	63.58	65.58	46.50	<b>78.63</b>
TP9	65.22	65.96	46.23	<b>76.17</b>
TP10	4.96	4.89	7.34	<b>12.08</b>
TP11	99.00	<b>100.00</b>	<b>100.00</b>	<b>100.00</b>
TP12	40.29	36.29	70.00	<b>85.29</b>

表6 测试问题难度等级分析

Table 6 Analysis of difficulty level of test problems

测试问题	目标数量	地图边长/格	编码维度	最优路径数量	求解难度等级
TP3	3	50	71	4	1
TP6	2	40	45	5	1
TP1	2	40	45	9	2
TP4	3	50	63	8	2
TP2	3	40	46	13	3
TP7	3	40	46	16	4
TP5	3	84	139	19	5
TP8	4	50	71	48	5
TP9	5	50	63	105	5
TP10	7	84	139	1487	6

2个问题的求解难度。

显然,测试问题3(TP3)与测试问题6(TP6)的求解难度低,算法的性能对比结果不明显,由于实验结果是50次实验的平均值,少数实验规划路径数量少1条会导致路径数量的指标相差 $10^{-2}$ 的数量级,最终导致最优路径占比结果的细微差距。因此,为了进一步验证子种群划分与解码经验表的引入提升了MSCL算法的寻优能力,基于问题难度等级的增加,不同算法在最优路径数量与搜索时间指标上的对比结果如图14和图15所示,横轴为测试问题,其难度等级依次递增。

从图14和图15可以看出,对于简单的MMOPP问题,不同算法的寻优能力区别不大。但是,相较于PSO算法,子种群划分与解码经验表的引入可以有效地提升算法性能。更重要的是,随着问题求解难度等级的依次递增,MSCL算法与其他算法在寻优能力方面的差距逐渐增大。并且,对于测试问题

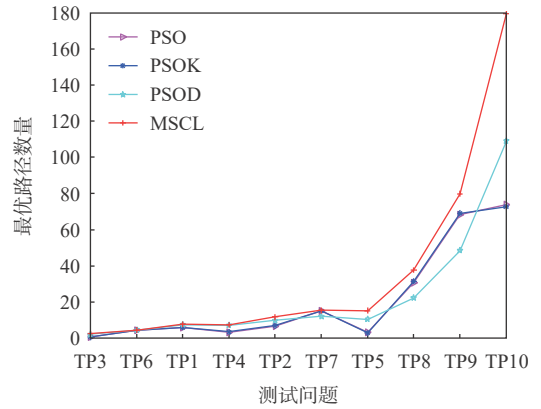


图14 不同算法搜索到的最优路径数量

Fig. 14 Number of the optimal paths searched by different algorithms

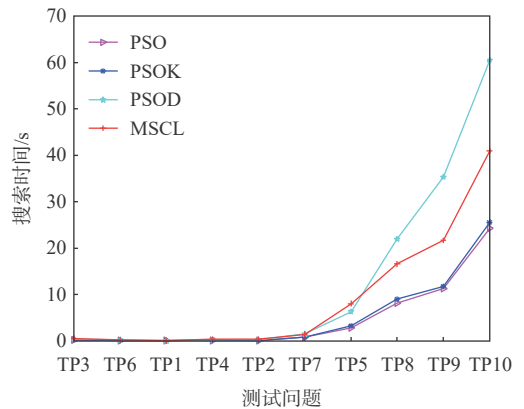


图15 不同算法规划路径所需时间

Fig. 15 Time used in path planning for different algorithms

8~10这类优化目标多、地图尺寸大的复杂MMOPP问题,MSCL算法的求解效率、搜索能力、寻优能力都极大地优于其他算法。这进一步说明了多种群合作学习算法能够有效地解决多模态多目标路径规划问题,尤其是在超多目标的多模态路径规划问题中具有很好的应用前景。

## 4 结论

1) 针对路径规划研究中存在解空间大、编码困难的问题,采用基于关键点的路径编码,该策略可以降低编码维度,提高算法的搜索效率。

2) 针对多模态多目标的路径规划需求,考虑在保障算法寻优能力的同时,平衡粒子群在目标空间内不同维度上的搜索能力,提出基于多目标分解的子种群划分。

3) 为解决路径规划问题中编解码匹配的不唯一性,提出基于解码经验表的路径解码,通过传递有效的解码经验,降低解码的不确定性,进一步提高算法的寻优能力。

4) 基于MMOPP测试集的实验结果验证了多

种群合作学习算法能够有效地解决 MMOPP 问题,并且,在复杂的 MMOPP 问题中,所提算法的寻优能力更强,最优解的多样性更好,搜索效率受问题规模大小的影响较小。

本文提出的多种群合作学习算法说明了解码经验的有效传递可以提升算法的搜索能力,提高优化问题的求解效率。因此,在后续的研究工作中笔者将继续优化与改进解码经验的传递策略,解决具有多模态多目标优化需求的实际问题。

### 参考文献 (References)

- [1] DEB K. Multi-objective genetic algorithms: Problem difficulties and construction of test problems[J]. *Evolutionary Computation*, 1999, 7(3): 205-230.
- [2] FENG G, KORKMAZ T. Finding multi-constrained multiple shortest paths[J]. *IEEE Transactions on Computers*, 2015, 64(9): 2559-2572.
- [3] 王树西, 李安渝. Dijkstra算法中的多邻接点与多条最短路径问题[J]. *计算机科学*, 2014, 41(6): 217-224.  
WANG S X, LI A Y. Multi-adjacent-vertexes and multi-shortest-paths problem of Dijkstra algorithm[J]. *Computer Science*, 2014, 41(6): 217-224(in Chinese).
- [4] HAYAT S, YANMAZ E, BROWN T X, et al. Multi-objective UAV path planning for search and rescue[C]//IEEE International Conference on Robotics and Automation. Piscataway: IEEE Press, 2017, 5569-5574.
- [5] 马小铭, 靳伍银. 基于改进蚁群算法的多目标路径规划研究[J]. *计算技术与自动化*, 2020, 39(4): 100-105.  
MA X M, JIN W Y. Multi-objective path planning based on improved and colony algorithm[J]. *Computing Technology and Automation*, 2020, 39(4): 100-105(in Chinese).
- [6] DENG Y T, RONG D C, SHANGGUAN W, et al. Multi-objective path optimization method in terminal building based on improved genetic algorithm[C]//Chinese Automation Congress. 2020, 3181-3186.
- [7] NAZARAHARI M, KHANMIRZA E, DOOSTIE S. Multi-objective multi-robot path planning in continuous environment using an enhanced genetic algorithm[J]. *Expert Systems with Applications*, 2018, 115: 106-120.
- [8] LI Z J, LIU Y, YANG Z. An effective kernel search and dynamic programming hybrid heuristic for a multimodal transportation planning problem with order consolidation[J]. *Transportation Research Part E: Logistics and Transportation Review*, 2021, 152: 102408.
- [9] WANG Z Z, ZHANG M H, CHU R J, et al. Modeling and planning multimodal transport paths for risk and energy efficiency using AND/OR graphs and discrete ant colony optimization[J]. *IEEE Access*, 2020, 8: 132642-132654.
- [10] 张启钱, 许卫卫, 张洪海, 等. 复杂低空物流无人机路径规划[J]. *北京航空航天大学学报*, 2020, 46(7): 1275-1286.  
ZHANG Q Q, XU W W, ZHANG H H, et al. Plan planning for logistics UAV in complex low-altitude airspace[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2020, 46(7): 1275-1286(in Chinese).
- [11] MIAO C, GHEN G, YAN C, et al. Path planning optimization of indoor mobile robot based on adaptive ant colony algorithm[J]. *Computers & Industrial Engineering*, 2021, 156: 107230.
- [12] ZHOU X, WANG X, GU X. Welding robot path planning problem based on discrete MOEA/D with hybrid environment selection[J]. *Neural Computing and Applications*, 2021, 4: 12881-12903.
- [13] 肖博. 基于多目标优化的震后应急物流路径规划研究[D]. 西安: 长安大学, 2019: 10-14.  
XIAO B. Study on post-earthquake emergency logistics path planning based on multi-objective optimization[D]. Xi'an: Chang'an University, 2019: 10-14(in Chinese).
- [14] CHOU Y T, HSIA S Y, LAN C H. A hybrid approach on multi-objective route planning and assignment optimization for urban lorry transportation[C]//International Conference on Applied System Innovation. Piscataway: IEEE Press, 2017: 1006-1009.
- [15] 蒋强, 易春林, 张伟, 等. 基于蚁群算法的移动机器人多目标路径规划[J]. *计算机仿真*, 2021, 38(2): 318-325.  
JIANG Q, YI C L, ZHANG W, et al. The multi-objective path planning for mobile robot based on ant colony algorithm[J]. *Computer Simulation*, 2021, 38(2): 318-325(in Chinese).
- [16] 王晨宇. 基于智能优化算法的多目标路径规划方法研究[D]. 桂林: 桂林电子科技大学, 2020: 18-27.  
WANG C Y. Research on multi-objective path planning method based on intelligent optimization algorithm[D]. Guilin: Guilin University of Electronic Technology, 2020: 18-27(in Chinese).
- [17] 熊昕霞, 何利力. 基于混合粒子群算法的移动机器人路径规划[J]. *计算机系统应用*, 2021, 30(4): 153-159.  
XIONG X X, HE L L. Path planning for mobile robot based on improved particle swarm optimization algorithm[J]. *Computer Systems & Applications*, 2021, 30(4): 153-159(in Chinese).
- [18] GUL F, RAHIMAN W, ALHADY S S N, et al. Meta-heuristic approach for solving multi-objective path planning for autonomous guided robot using PSO-GWO optimization algorithm with evolutionary programming[J]. *Journal of Ambient Intelligence and Humanized Computing*, 2021, 12: 7873-7890.
- [19] HIDALGO-PANIAGUA A, VEGA-RODRÍGUEZ M A, FERRUZ J, et al. Solving the multi-objective path planning problem in mobile robotics with a firefly-based approach[J]. *Soft Computing*, 2017, 21(4): 1-16.
- [20] 段益琴. 基于多目标优化的多机器人路径规划研究[D]. 重庆: 重庆邮电大学, 2020: 23-34.  
DUAN Y Q. Research on multi-robot path planning based on multi-objective optimization[D]. Chongqing: Chongqing University of Posts and Telecommunications, 2020: 23-34(in Chinese).
- [21] 樊娇, 雷涛, 董南江, 等. 基于改进NSGA-II算法的多目标无人机路径规划[J]. *火力与指挥控制*, 2022, 47(2): 43-48.  
FAN J, LEI T, DONG N J, et al. Multi-objective UAV path planning based on an improved NSGA-II[J]. *Fire Control & Command Control*, 2022, 47(2): 43-48(in Chinese).
- [22] DHIKARI D, KIM E, REZA H. A fuzzy adaptive differential evolution for multi-objective 3D UAV path optimization[C]//IEEE Congress on Evolutionary Computation. Piscataway: IEEE Press, 2017: 2258-2265.
- [23] SHI Y, EBERHART R. A modified particle swarm optimizer

- [C]/IEEE World Congress on Computational Intelligence. Piscataway: IEEE Press, 1998: 69-73.
- [24] EBERHART R, KENNEDY J. A new optimizer using particle swarm theory[C]//Proceedings of the Sixth International Symposium on Micro Machine and Human Science. Piscataway: IEEE Press, 1995: 39-43.
- [25] LIANG J, YUE C, LI G, et al. Problem definitions and evaluation criteria for the CEC 2021 on multimodal multiobjective path planning optimization[EB/OL]. (2020-12)[2021-05-24]. <https://cec2021.mini.pw.edu.pl/en/program/competitions-C-8>.

## A multimodal multi-objective path planning algorithm based on multi-swarm cooperative learning

ZHAO Meng, LU Hui\*, WANG Shiqi, YANG Siyi, WANG Zan

(School of Electronic and Information Engineering, Beihang University, Beijing 100191, China)

**Abstract:** An algorithm based on multi-swarm cooperative learning was proposed to plan multiple optimal paths to meet multiple objectives, which can improve the robustness and practicability of the planned paths. The concept of the particle swarm optimization algorithm served as the algorithm's guidance. First, to address the issue that a single population is easy to trap in local optimum in the multi-dimensional target space, a strategy of sub-swarm division was proposed. The population was divided into many sub-swarms according to the number of objectives, balancing the searching ability of the algorithm in each dimension of the target space. Second, key path points were extracted according to the in-degree and out-degree of the path points in the map. In the coding process, real coding was used to initialize the population. The dimension of the path code was equal to the number of key path points, reducing the size of the solution space. In the decoding process, the decoding experience of the elite solutions guided the fast search for feasible solutions. This method can transfer the decoding experience efficiently and reduce the uncertainty of decoding, which improved the optimization ability of the algorithm. Finally, the search results of all sub-swarms were sorted by the non-dominated sorting method to obtain the paths satisfying the planning objectives. The path planning algorithm based on the multi-swarm cooperative learning outperforms the standard particle swarm optimization algorithm in terms of search and optimization ability and is capable of solving the multimodal multi-objective path planning problem.

**Keywords:** multimodal multi-objective optimization; path planning; particle swarm optimization algorithm; sub-swarm division; decoding experience table

**Received:** 2021-05-27; **Accepted:** 2021-08-22; **Published Online:** 2021-09-29 16:45

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20210929.1105.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210929.1105.001.html)

**Foundation items:** National Natural Science Foundation of China (61671041); The Foundation of Shaanxi Key Laboratory of Integrated and Intelligent Navigation (SKLIIN-20190201)

\* **Corresponding author.** E-mail: [mluhui@buaa.edu.cn](mailto:mluhui@buaa.edu.cn)

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0280

# 基于线性约束最小方差的稳健波束形成算法

吕岩<sup>1,2</sup>, 曹菲<sup>1,\*</sup>

(1. 火箭军工程大学核工程学院, 西安 710025; 2. 中国人民解放军 96746 部队, 库尔勒 841000)

**摘要:** 针对平面阵列天线波束形成过程中的波达方向 (DOA) 估计失配问题, 提出在期望信号 (SOI) 方向附近增加线性约束的算法, 有效提升了平面阵列波束形成的稳健性; 此外, 针对增加线性约束会导致波束形成算法自由度降低的问题, 以均匀线阵为例, 提出在广义旁瓣相消 (GSC) 算法模型中添加阻塞矩阵预选环节的算法, 有效解决了添加线性约束所致的自由度损失问题, 从而使算法在提升稳健性的同时保持了原有的自由度。最后, 通过计算机仿真实验验证了所提算法的有效性。

**关键词:** 稳健波束形成; 波达方向估计失配; 线性约束最小方差; 广义旁瓣相消; 阻塞矩阵

**中图分类号:** TN957.2

**文献标志码:** A

**文章编号:** 1001-5965(2023)03-0617-08

阵列信号处理作为一种空间滤波技术, 可通过调整阵元权值控制波束形成, 使其高增益主瓣波束指向期望信号 (signal of interest, SOI), 同时将干扰信号方向的增益置零, 与波达方向估计<sup>[1]</sup>(direction of arrival, DOA) 配合达到接收 SOI 和抑制干扰的目的<sup>[2]</sup>。目前, 波束形成已被广泛应用于雷达、声呐、地震波监测、现代通讯和医学成像等诸多领域<sup>[3-5]</sup>。

波束形成效果与 DOA 的估计精度密不可分, 因此, 当 DOA 估计出现偏差时, 波束形成算法性能会严重下降<sup>[6-7]</sup>。对此, 国内外学者开展了大量的研究工作, 提出了许多稳健波束形成的算法, 主要包含线性约束最小方差<sup>[8]</sup>(linearly constrained minimum variance, LCMV) 算法、对角加载算法<sup>[9]</sup>和最差性能优化波束形成算法<sup>[10]</sup>等。其中对角加载算法通过设置惩罚权值作为加载因子, 有效提高了波束形成的稳健性, 但加载因子的选取目前并不存在严格的标准, 使算法的应用受到限制。最差性能优化算法通过求解优化问题确定加载因子大小, 与对角加载算法相比具备一定优势, 但在信噪比<sup>[11]</sup>(signal noise ratio, SNR) 较高时, 性能会有明显的下降。广义旁

瓣相消 (generalized sidelobe canceller, GSC) 算法是 LCMV 的一种等效实现算法<sup>[12-13]</sup>, 能够将 LCMV 转变为无约束优化问题, GSC 模型分为两支路, 其中主支路允许期望信号通过, 辅助支路允许干扰信号和噪声通过, 并利用两支路所含信号的差异进行自适应相消。

本文基于 LCMV 算法, 分别以平面阵列天线和均匀线阵为对象, 针对平面阵俯仰角 DOA 估计偏差, 采用在 SOI 方向附近增加线性约束的算法提高波束形成稳健性; 而后, 针对文献 [7] 以损失算法自由度为代价的不足, 通过添加阻塞矩阵预选环节改进了 GSC 模型, 避免了在 SOI 方向附近添加线性约束, 因此, 所提算法可在提升波束形成稳健性的同时使其保持原有的自由度。

## 1 信号模型

### 1.1 均匀线阵

假设  $W$  个具有任意方向性的阵元沿直线均匀排列, 间距为窄带入射信号半波长。对于  $K$  个远场入射信号, 则第  $i$  个阵元的输出为

收稿日期: 2021-05-28; 录用日期: 2021-08-29; 网络出版时间: 2021-09-14 14:25

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210914.1125.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210914.1125.001.html)

\* 通信作者. E-mail: [caofei101@126.com](mailto:caofei101@126.com)

**引用格式:** 吕岩, 曹菲. 基于线性约束最小方差的稳健波束形成算法 [J]. 北京航空航天大学学报, 2023, 49 (3): 617-624.

LYU Y, CAO F. Robust beamforming based on linear constraint minimum variance algorithm [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 617-624 (in Chinese).

$$y_i(n) = \sum_{l=1}^K e^{-j2(i-1)\pi \sin \theta_l} s_l(n) + v_i(n) \quad (1)$$

式中:  $j$  为虚数,  $j = \sqrt{-1}$ ;  $s_l(n)$  为阵元从第  $l$  个信号源接收到的信号, 信号之间互不相关;  $\theta_l$  为入射角;  $v_i(n)$  为第  $i$  个阵元的高斯白噪声。将式 (1) 表达成矩阵的形式:

$$\mathbf{y}(n) = \mathbf{A}\mathbf{s}(n) + \mathbf{v}(n) \quad (2)$$

式中:

$$\mathbf{y}(n) = [y_1(n), y_2(n), \dots, y_w(n)]^T \quad (3)$$

包含各个信号源 DOA 的导向矩阵为

$$\mathbf{A} = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_K)] \quad (4)$$

式中:  $\mathbf{a}(\theta_l) = [1, e^{-j\phi_l}, e^{-j2\phi_l}, \dots, e^{-j(M-1)\phi_l}]^T$ , 其中  $\phi_l = 2\pi \sin \theta_l$ 。远场信号源所组成的矩阵  $\mathbf{s}(n) = [s_1(n), s_2(n), \dots, s_K(n)]^T$ , 高斯白噪声组成的矩阵  $\mathbf{v}(n) = [v_1(n), v_2(n), \dots, v_w(n)]^T$ 。

根据式 (2), 可知阵列接收信号的相关矩阵 (含噪声) 为

$$\mathbf{R}_y = E[\mathbf{y}(n)\mathbf{y}^H(n)] \quad (5)$$

$$\mathbf{R}_y = \mathbf{A}\mathbf{R}_s\mathbf{A}^H + \mathbf{R}_v \quad (6)$$

式中:  $\mathbf{R}_s = E[\mathbf{s}(n)\mathbf{s}^H(n)]$  为  $K$  个信号源的相关矩阵;  $\mathbf{R}_v$  为高斯白噪声相关矩阵, 可表示为

$$\mathbf{R}_v = E[\mathbf{v}(n)\mathbf{v}^H(n)] = \sigma_v^2 \mathbf{I}_w \quad (7)$$

式中:  $\sigma_v^2$  为噪声功率。

## 1.2 矩形平面阵

平面阵列是线性阵列的一种推广, 假设  $M$  行  $N$  列的均匀矩形平面阵列, 阵元总数为  $Q = MN$ , 如图 1 所示。

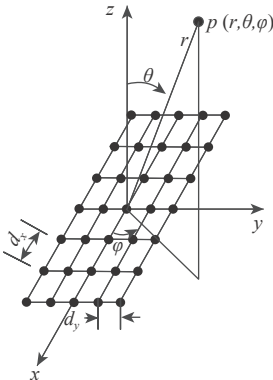


图 1 平面阵列

Fig. 1 Planar array

图 1 中黑色网格点位置为独立天线阵元, 各阵元间距同样为入射波半波长, 即  $d_x = d_y = \lambda/2$ 。对于

$K$  个远场窄带入射信号, 第  $q$  个阵元接收到的信号为

$$y_q(n) = \sum_{i=1}^K a(\theta_i, \varphi_i) s_i(n) + v_q(n) \quad (8)$$

将式 (8) 写成矩阵形式:

$$\mathbf{y}(n) = \mathbf{A}\mathbf{s}(n) + \mathbf{v}(n) \quad (9)$$

式中:

$$\begin{cases} \mathbf{A} = [\mathbf{a}(\theta_1, \varphi_1), \mathbf{a}(\theta_2, \varphi_2), \dots, \mathbf{a}(\theta_K, \varphi_K)] \\ \mathbf{a}(\theta_i, \varphi_i) = [e^{-j2\pi/\lambda\alpha_{1,i}}, e^{-j2\pi/\lambda\alpha_{q,i}}, \dots, e^{-j2\pi/\lambda\alpha_{Q,i}}] \\ \alpha_{q,i} = X_q\mu_i + Y_q\eta_i \\ \mu_i = \sin \theta_i \cos \varphi_i, \\ \eta_i = \sin \theta_i \sin \varphi_i \end{cases} \quad (10)$$

其中:  $\mathbf{a}(\theta_i, \varphi_i)$  为在方位角为  $\theta_i$ 、俯仰角为  $\varphi_i$  时的导向矢量;  $X_q$  为第  $q$  个阵元的横坐标;  $Y_q$  为第  $q$  个阵元的纵坐标。

## 2 波束形成算法

### 2.1 LCMV 算法

LCMV 算法通过最小化输出功率, 同时满足若干线性约束条件, 从而确定阵元的权值矢量, 问题的表达式为

$$\mathbf{w} = \arg \min_{\mathbf{w}} \mathbf{w}^H \mathbf{R}_y \mathbf{w} \quad \text{s.t. } \mathbf{C}^H \mathbf{w} = \mathbf{f} \quad (11)$$

式中:  $\mathbf{C}$  为  $Q \times (J+1)$  维约束矩阵,  $Q$  为阵元数,  $J$  为干扰信号源个数;  $\mathbf{f}$  为  $J+1$  维约束向量; 根据 Lagrange 乘法, 解得式 (11) 的最优解为

$$\mathbf{w} = \mathbf{R}_y^{-1} \mathbf{C} (\mathbf{C}^H \mathbf{R}_y^{-1} \mathbf{C})^{-1} \mathbf{f} \quad (12)$$

### 2.2 GSC 算法

如图 2 所示, GSC 算法的权值矢量分为  $\mathbf{w}_q$  和  $\mathbf{w}_a$  两部分, 分别代表主支路权向量和辅助支路权向量。

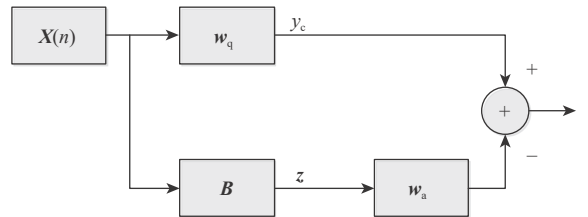


图 2 GSC 模型

Fig. 2 Model of GSC

图 2 中阻塞矩阵  $\mathbf{B}$  为  $Q \times (Q-J-1)$  维, 起阻塞 SOI 的作用, 满足  $\mathbf{B}^H \mathbf{C} = \mathbf{0}$ 。GSC 的权向量可表示为

$$\mathbf{w} = \mathbf{w}_q - \mathbf{B}\mathbf{w}_a \quad (13)$$

式中:

$$\begin{cases} \mathbf{w}_q = (\mathbf{C}\mathbf{C}^H)^{-1}\mathbf{C}\mathbf{f} \\ \mathbf{w}_a = (\mathbf{B}^H\mathbf{R}_y\mathbf{B})^{-1}\mathbf{B}^H\mathbf{R}_y\mathbf{w}_q \end{cases} \quad (14)$$

### 2.3 平面阵稳健波束形成算法和改进 GSC 模型

#### 2.3.1 平面阵稳健波束形成算法

假定 $(\theta_s, \varphi_s)$ 为平面阵列 SOI 的 DOA 估计值, 此时为保证 SOI 损失最小, 设置 $\mathbf{C} = [\mathbf{a}(\theta_s, \varphi_s)]$ 和约束向量 $\mathbf{f} = 1$ , 可得

$$\mathbf{w}_{\text{ref}} = (\mathbf{C}\mathbf{C}^H)^{-1}\mathbf{C}\mathbf{f} \quad (15)$$

由于 DOA 估计值和真实值之间可能存在误差, 并且波束图对 DOA 估计值和期望值之间的误差比较敏感, 较小的误差也会对波束形成产生较大影响, 所以需对 LCMV 算法添加额外的线性约束, 增强其稳健性。以平面阵天线的俯仰角 $\varphi_s$ 存在误差为例, 定义角度增量为 $\delta_\varphi$ , 则与期望角度 $\varphi_i$ 邻近的俯仰角为

$$\boldsymbol{\varphi} = [\varphi_s - \varepsilon\delta_\varphi, \dots, \varphi_s, \dots, \varphi_s + \varepsilon\delta_\varphi] \quad (16)$$

式中:  $\varepsilon \in [1, \text{floor}((MN - 1)/2)]$ , 其中 floor 表示向下取整运算, 可得

$$\tilde{\mathbf{C}} = [\mathbf{a}(\theta_s, \varphi_s - \varepsilon\delta_\varphi), \dots, \mathbf{a}(\theta_s, \varphi_s), \dots, \mathbf{a}(\theta_s, \varphi_s + \varepsilon\delta_\varphi)] \quad (17)$$

采用式 (17) 作为约束矩阵相当于在期望方向附近添加了一组“字典”, 真实 DOA 可利用该字典进行稳健匹配, 示意如图 3 所示。

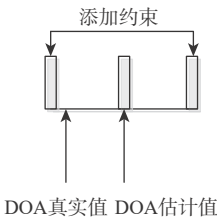


图 3 字典示意图

Fig. 3 Sign for dictionary

采用 $\mathbf{f}$ 为全 1 向量进行约束会产生过约束现象, 导致波束形成产生畸变, 取:

$$\tilde{\mathbf{f}} = (\mathbf{w}_{\text{ref}}^H\mathbf{C})^T \quad (18)$$

此时, 线性约束问题变为

$$\mathbf{w} = \arg \min_{\mathbf{w}} \mathbf{w}^H\mathbf{R}_y\mathbf{w} \quad \text{s.t.} \quad \tilde{\mathbf{C}}^H\mathbf{w} = \tilde{\mathbf{f}} \quad (19)$$

采用 Lagrange 乘子法构建目标函数:

$$J = \mathbf{w}^H\mathbf{R}_y\mathbf{w} + \lambda_1 (\mathbf{w}^H\mathbf{a}(\theta_s, \varphi_s - \varepsilon\delta_\varphi) - f_1) + \dots + \lambda_k (\mathbf{w}^H\mathbf{a}(\theta_s, \varphi_s + \varepsilon\delta_\varphi) - f_k) \quad (20)$$

式中:  $\lambda_1, \lambda_2, \dots, \lambda_k$  为拉格朗日乘子, 令式 (20) 导数

为 0 并求解, 可得

$$\mathbf{w} = \mathbf{R}_y^{-1}\tilde{\mathbf{C}}(\tilde{\mathbf{C}}^H\mathbf{R}_y^{-1}\tilde{\mathbf{C}})^{-1}\tilde{\mathbf{f}} \quad (21)$$

#### 2.3.2 改进 GSC 模型

文献 [7] 通过添加线性约束的方式提高了 LCMV 算法的稳健程度, 但以消耗算法自由度为代价, 且仅利用了 GSC 算法的主支路权向量, 未充分利用辅助支路。GSC 算法主支路和辅助支路的主要区别在于是否含有 SOI, 理想情况下主支路中 SOI 含量较高, 辅助支路不含 SOI, 所以两支路信号的相关程度代表着阻塞矩阵的阻塞效果。当主支路和辅助支路相关程度较高时, 表示阻塞矩阵效果降低, DOA 的真实值和估计值之间存在误差; 反之, 相关程度较低时, 表示阻塞效果较好, 基于该项性质, 本文在 GSC 模型中添加阻塞矩阵预选环节, 如图 4 所示。

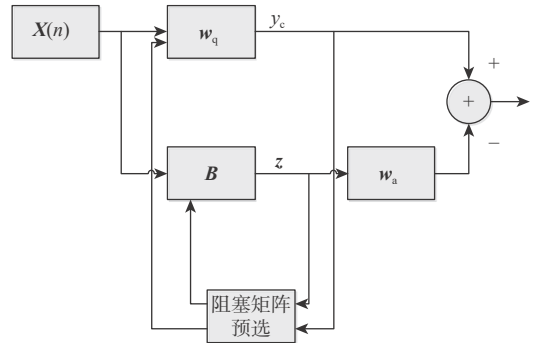


图 4 改进 GSC 模型

Fig. 4 Model of improved GSC

根据图 4, 在 SOI 方向附近选取不同角度计算阻塞矩阵, 利用不同阻塞矩阵计算两支路 $y_c$ 和 $z$ 的互相关向量, 并计算互相关向量的 $l_1$ 范数, 选择较小范数时的阻塞矩阵作为最佳矩阵。所提算法即为对阻塞矩阵进行预选, 未在 SOI 的 DOA 附近添加额外约束, 因此未对算法自由度产生影响, 改进 GSC 算法的流程如下:

**步骤 1** 选取 SOI 的 DOA 附近角度计算阻塞矩阵。

**步骤 3** 计算 $y_c$ 和 $z$ 的互相关向量, 并计算相关向量的 $l_1$ 范数。

**步骤 3** 执行 $\zeta$ 次步骤 2, 选择最小 $l_1$ 范数时的阻塞矩阵作为 $\mathbf{B}_{\text{best}}$ 。

**步骤 4** 利用 $\mathbf{B}_{\text{best}}$ 计算权值 $\mathbf{w}$ 。

可知, 改进 GSC 算法和原算法的主要差别在于矩阵预选部分, 经算法复杂度分析可得

$$O'_{\text{GSC}} = O_{\text{GSC}} + 4W^3\zeta \quad (22)$$

式中:  $W$ 和 $\zeta$ 分别为天线阵元数量和预选阻塞矩阵

次数。

### 3 仿真分析

#### 3.1 平面阵稳健波束形成算法仿真

仿真基于  $8 \times 8$  的矩形平面阵, 为便于比较俯仰角出现 DOA 失配时对波束形成性能的影响, 先将方位角设定为统一值, 设置 SOI 的 DOA 为  $(\theta_s, \varphi_s) = (15^\circ, 10^\circ)$ , 2 个干扰的 DOA 分别为  $(15^\circ, -40^\circ)$  和  $(15^\circ, 25^\circ)$ , 信噪比为 0 dB, 干噪比 (interference to noise ratio, INR) 为 10 dB。

图 5~图 7 分别为采用 LCMV 算法生成的空间增益三维图、等高线图和方位角等于  $15^\circ$  时俯仰角与波束增益的截面图, 可知 SOI 方向波束增益为峰值, 2 个干扰均被有效抑制。

假设 SOI 方向出现 DOA 失配, 实际俯仰方向波达角为  $11^\circ$ , 此时俯仰角和增益的截面图如图 8 所示。可知 DOA 失配时, 波束图出现较大畸变, 无法在 SOI 方向形成高增益波束, 干扰抑制能力也随之变差, 算法性能严重下降。

设置角度增量  $\delta_\varphi = 1^\circ$ , 在 SOI 方向附近额外添加 2 个约束, 则可得

$$\tilde{\mathbf{C}} = [\mathbf{a}(15^\circ, 9^\circ), \mathbf{a}(15^\circ, 10^\circ), \mathbf{a}(15^\circ, 11^\circ)] \quad (23)$$

此时俯仰角和波束增益的截面图如图 9 所示, 可知 SOI 方向波束未发生畸变, 虽然对 2 个干扰的抑制能力略有下降, 但仍能达到良好的抑制效果。

为进一步验证所提算法的性能, 将方位角设置为不同数值, SOI 的 DOA 设置为  $(\theta_s, \varphi_s) = (15^\circ, 10^\circ)$ ,

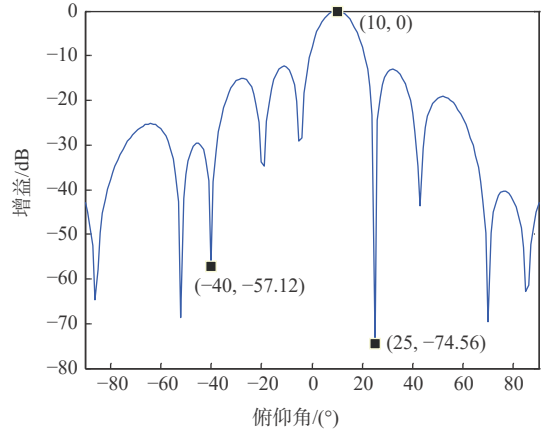


图 7 方位角  $15^\circ$  波束截面

Fig. 7 Sectional drawing of  $15^\circ$

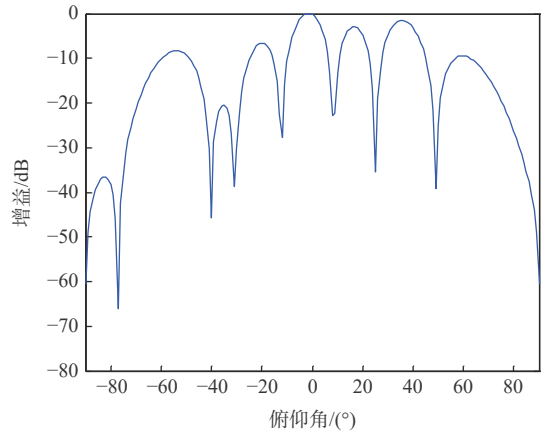


图 8 波束畸变

Fig. 8 Pattern synthesis distortion

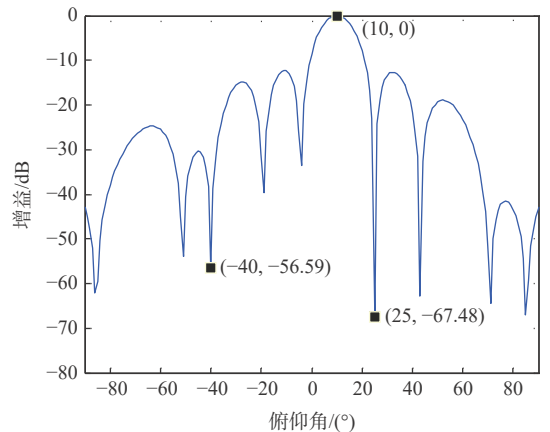


图 9 DOA 失配增益

Fig. 9 Pattern synthesis of DOA mismatch

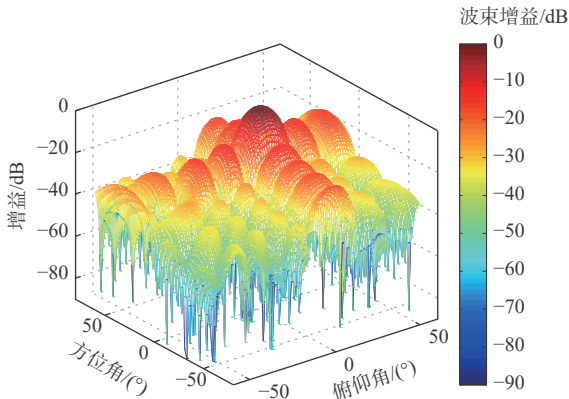


图 5 波束增益

Fig. 5 Pattern synthesis

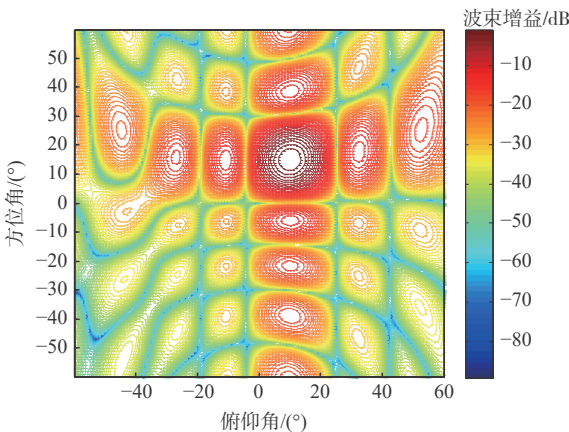


图 6 等高线

Fig. 6 Contour plot

2 个干扰信号的 DOA 分别设置为  $(10^\circ, -40^\circ)$  和  $(20^\circ, 25^\circ)$ , 实际波达角仍为  $(\theta_s, \varphi_s) = (15^\circ, 11^\circ)$ , 此时阵列的空间增益和等高线如图 10 和图 11 所示。

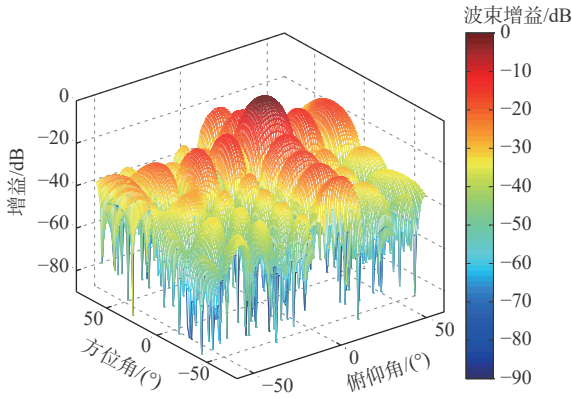


图 10 平面阵波束增益

Fig. 10 Pattern synthesis of planar array

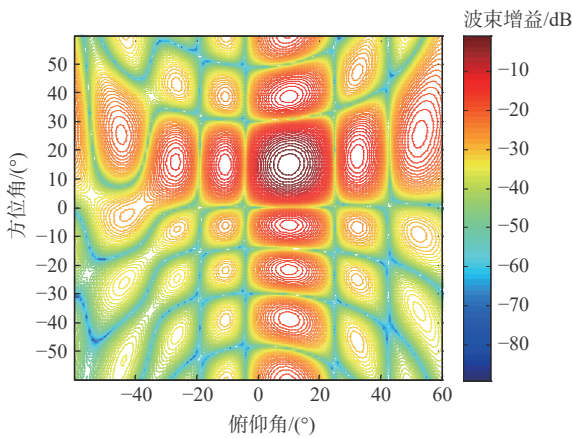


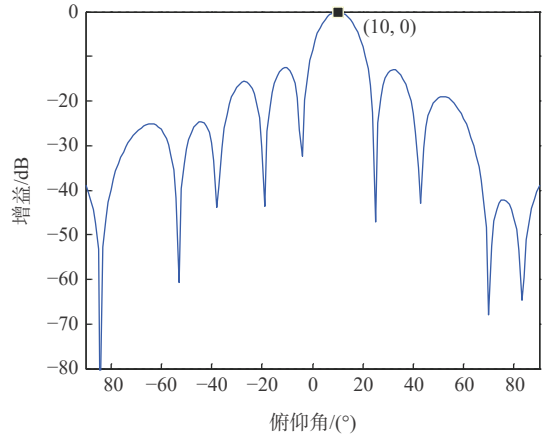
图 11 平面阵等高线

Fig. 11 Contour plot of planar array

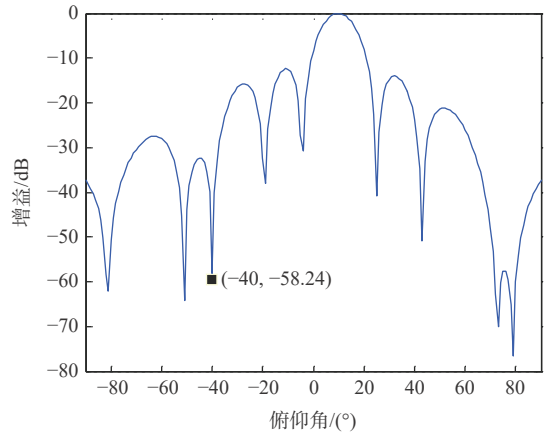
为进一步表明测试结果, 分别生成方位角为  $15^\circ$ 、 $10^\circ$  和  $20^\circ$  时的俯仰角与波束增益的截面如图 12 所示。根据图 12 可知, 主瓣波束未发生畸变, 干扰抑制效果良好, 添加约束取得了提升波束形成稳健性的作用。

### 3.2 改进 GSC 模型仿真

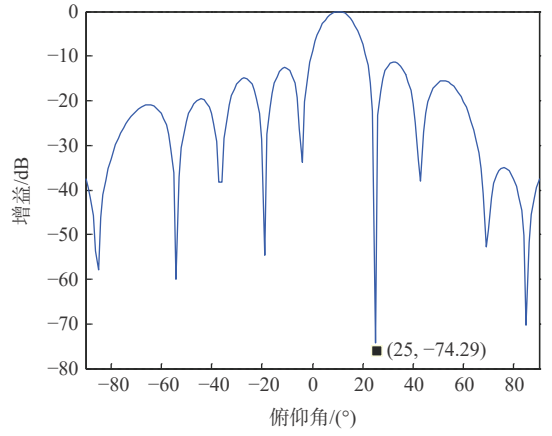
仿真基于 16 阵元的均匀线阵, 设置 SOI 的 DOA 为  $30^\circ$ , 3 个干扰 DOA 分别为  $40^\circ$ 、 $-25^\circ$  和  $-60^\circ$ , INR 为 30 dB。假设信号实际的 DOA 为  $31^\circ$ , 为验证改进 GSC 模型的有效性, 分别选取  $29^\circ$ 、 $30^\circ$ 、 $31^\circ$  这 3 个 DOA 生成阻塞矩阵, 并计算  $y_c$  和  $z$  互相关向量的  $l_1$  范数, 图 13(a) 为  $l_1$  范数随 SNR 从  $-20 \sim 20$  dB 之间的变化趋势 (已归一化), 图 13(a) 中蓝、绿和红色曲线分别对应  $31^\circ$ 、 $30^\circ$  和  $29^\circ$ 。进一步令 INR 与随机数 rand 相乘, 即 INR 处于  $0 \sim 30$  dB 之间随机变化。图 13(b) 为该状态下 200 次蒙特卡罗 [14-15] 实验的  $l_1$  范数平均值随 SNR 的变化趋势, 可知蓝色曲线的  $l_1$  范数最小, 说明其更接近真实 DOA, 应选择  $31^\circ$



(a) 方位角为  $15^\circ$



(b) 方位角为  $10^\circ$



(c) 方位角为  $20^\circ$

图 12 不同方位角的波束截面

Fig. 12 Sectional drawings of pattern syntheses for different azimuths

进行后续计算。

选定 DOA 为  $31^\circ$  构造阻塞矩阵, 并求取权值  $w$ , 图 14 为 SNR 分别为  $-10$  dB 和  $10$  dB 时, 改进 GSC 模型和文献 [7] 算法的波束形成对比, 从图 14 可知 2 种算法对  $-25^\circ$  方向干扰的抑制能力相近, 但改进 GSC 模型在  $40^\circ$  和  $-60^\circ$  方向上的增益要比文献 [7] 更低, 干扰抑制效果更优。

图 15(a) 为干扰方向不变 ( $40^\circ$ 、 $-25^\circ$  和  $-60^\circ$ ), SOI

的实际 DOA 不同时,  $l_i$  范数随 SNR 从  $-20 \sim 20$  dB 之间的变化趋势(已归一化的 200 次蒙特卡罗实验平均值)。图 15(b) 为实际 DOA 不变, 干扰信号方向不同时,  $l_i$  范数随 SNR 从  $-20 \sim 20$  dB 之间的变化

趋势。从图 15 中可以看出, 与实际 DOA 偏差越大, 则  $l_i$  范数越大, 表示  $y_c$  和  $z$  之间的相关程度较高, 阻塞矩阵对于 SOI 的阻塞效果不理想。随着与实际 DOA 偏差的逐渐降低,  $y_c$  和  $z$  之间的相关程度降

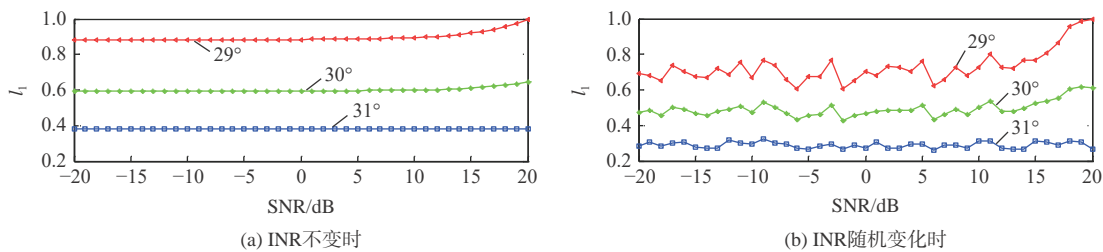


图 13 范数变化趋势

Fig. 13 Trends of norms

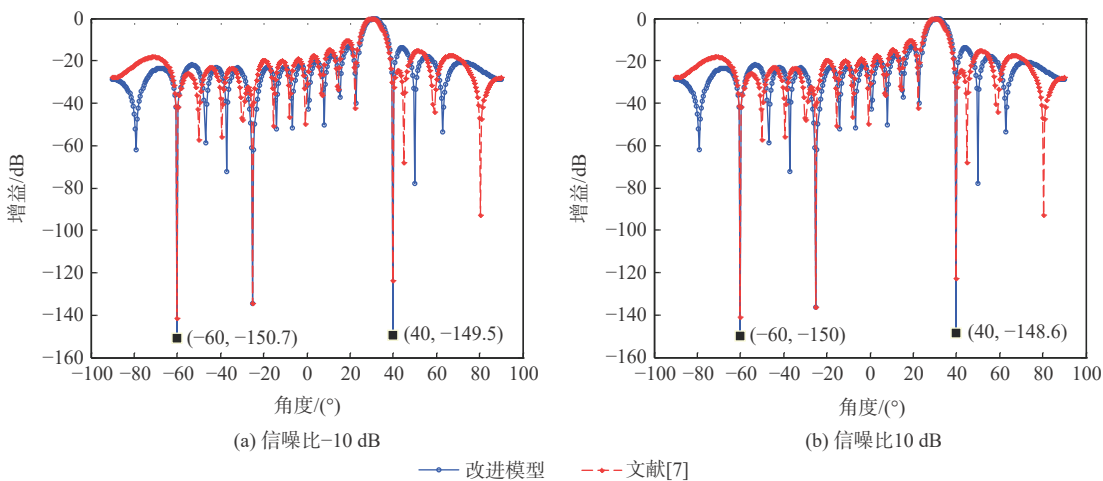


图 14 改进 GSC 和文献 [7] 波束形成对比

Fig. 14 Comparison of pattern synthesis between improved GSC and Ref. [7]

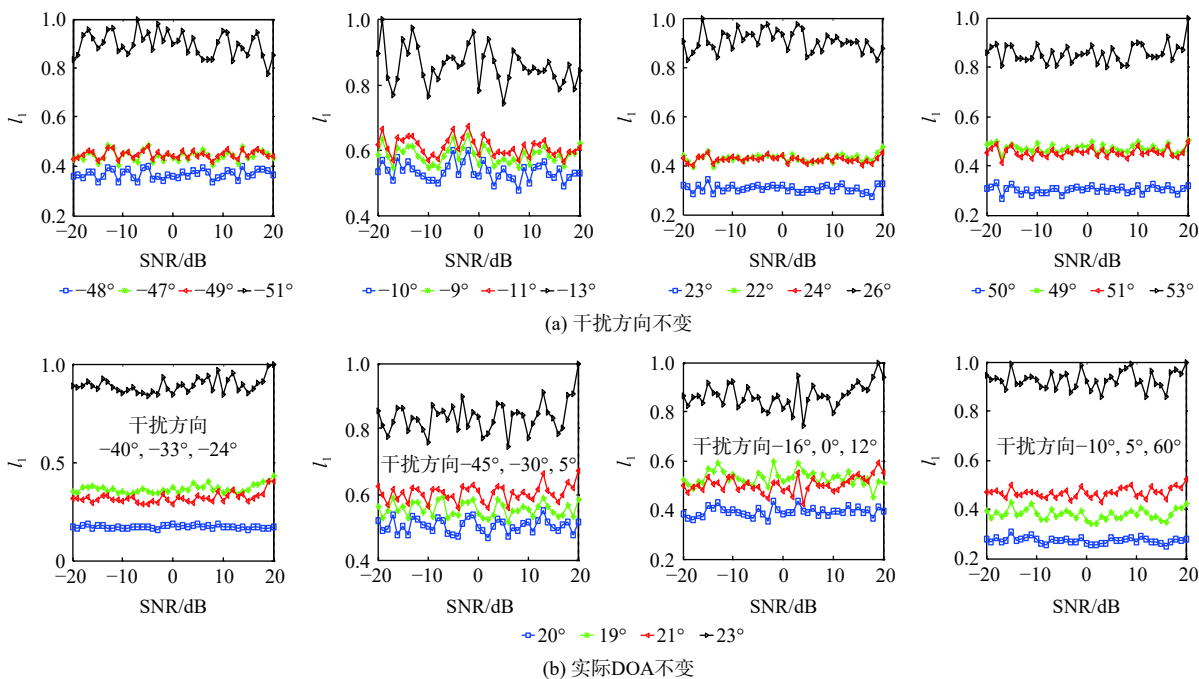


图 15 不同参数时的范数变化趋势

Fig. 15 Trends of norms under different parameters

低, 阻塞矩阵对于 SOI 的阻塞效果逐步提升。

表 1 和图 16 为 GSC、文献 [7] 和改进 GSC 算法运行耗时的对比, 仿真平台为 MATLAB2014, CPU3.40 GHz、RAM16 GB。根据结果可知当阵元数量较少时, 3 种算法运行耗时相差不大。随着阵元数量增多, 3 种算法的运行时间逐渐增加, 其中改进 GSC 算法增加的最快, 当阵元数量为 100 时, 改进 GSC 算法约比文献 [7] 算法耗时多 2.4 倍。

表 1 不同阵元数量耗时

Table 1 Time consumed versus different elements

算法	不同阵元数量耗时/ms				
	10	30	60	80	100
GSC	0.180 6	0.389 4	0.677 5	0.902 9	1.206 0
文献[7]	0.212 7	0.458 8	0.754 8	1.150 5	1.356 3
改进GSC	0.275 5	0.902 6	1.663 4	2.400 4	3.278 6

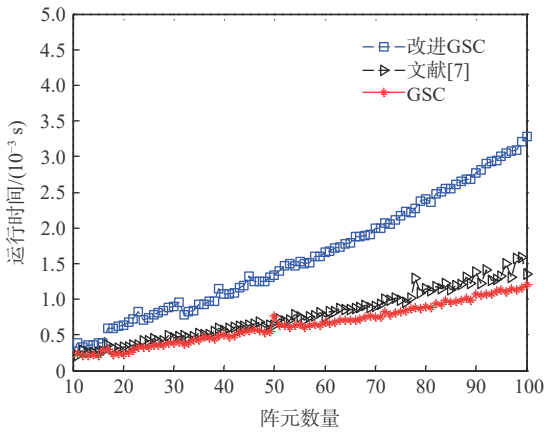


图 16 运行时间对比

Fig. 16 Comparison of running times

## 4 结 论

1) 本文针对平面阵列天线波束形成过程中的 DOA 估计失配问题, 以俯仰角存在误差为例, 采用在参考方向附近生成“字典”, 增加线性约束的算法, 起到增强波束形成稳健性的作用。

2) 针对增加线性约束会导致算法自由度降低的问题, 采用增加阻塞矩阵预选环节改进了 GSC 算法, 通过  $l_1$  范数平均值变化趋势、波束形成增益和算法运行时间三方面的对比, 验证了改进 GSC 算法虽然会增加运行时间, 但可以在干扰方向产生更深的零陷, 并可使算法保持原有自由度;

下一步将研究平面阵俯仰、方位角双失配时的线性约束添加算法和将改进 GSC 算法应用于平面阵。

### 参考文献 (References)

[1] WEI Z, LI X, WANG B, et al. An efficient super-resolution DOA estimator based on grid learning[J]. *Radio Engineering*, 2019,

28(4): 785-792.

[2] YANG X P, LI S, SUN Y, et al. Robust wideband adaptive beamforming with null broadening and constant beamwidth[J]. *IEEE Transactions on Antennas and Propagation*, 2019, 67(8): 5380-5389.

[3] AKBAR S, RAJA M A Z, ZAMAN F, et al. Design of bio-inspired heuristic techniques hybridized with sequential quadratic programming for joint parameters estimation of electromagnetic plane waves[J]. *Wireless Personal Communications*, 2017, 96: 1475-1494.

[4] TALISA S H, OHAVER K W, COMBERIATE T M, et al. Benefits of digital phased array radars[J]. *Proceedings of the IEEE*, 2016, 104(3): 530-543.

[5] SCHULDT C. Trigonometric interpolation beamforming for a circular microphone array[C]//*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Piscataway: IEEE Press, 2019: 431-435.

[6] SHERSON T, KLEIJN W B, HEUSDENS R. A distributed algorithm for robust LCMV beamforming[C]//*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Piscataway: IEEE Press, 2016: 101-105.

[7] KHAN M, MALIK A N, ZAMAN F, et al. Robust LCMV beamformer for direction of arrival mismatch without beam broadening[J]. *Wireless Personal Communications*, 2019, 104: 21-36.

[8] CAMELLIA D, SIEH K T. Performance comparison of FA, PSO and CS application in SINR optimisation for LCMV beamforming technique[J]. *Wireless Personal Communications*, 2018, 103: 2177-2195.

[9] WANG W, WU R. A novel diagonal loading method for robust adaptive beamforming[J]. *Progress in Electromagnetics Research C*, 2011, 18: 245-255.

[10] VOROBYOV S A, GERSHMAN A B, LUO Z Q. Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem[J]. *IEEE Transactions on Signal Processing*, 2003, 51(2): 313-324.

[11] 王博, 谢军伟, 张晶, 等. 基于SD-LCMV算法的FDA平台外干扰抑制[J]. *北京航空航天大学学报*, 2019, 45(11): 2247-2256.

WANG B, XIE J W, ZHANG J, et al. FDA platform external interference suppression based on SD-LCMV algorithm[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2019, 45(11): 2247-2256(in Chinese).

[12] 王永良, 丁前军, 李荣峰. 自适应阵列处理[M]. 北京: 清华大学出版社, 2009: 44-47.

WANG Y L, DING Q J, LI R F. Adaptive array processing[M]. Beijing: Tsinghua University Press, 2009: 44-47(in Chinese).

[13] WANG P, SHI Y, JIANG J, et al. Generalized sidelobe canceller for ultrasound imaging based on eigenvalue decomposition[J]. *Acoustical Physics*, 2019, 65(1): 123-131.

[14] QIN Y H, LIU Y M, LIU J Y, et al. Underdetermined wideband doa estimation for off-grid sources with coprime array using sparse bayesian learning[J]. *Sensors*, 2018, 18(1): 253-263.

[15] YUAN J W, ZHANG G, LIU W B, et al. Off-grid sparse DOA estimation based iterative reweighted linear interpolation in spatial coloured noise[J]. *Electronics Letters*, 2020, 56(11): 573-575.

# Robust beamforming based on linear constraint minimum variance algorithm

LYU Yan<sup>1,2</sup>, CAO Fei<sup>1,\*</sup>

(1. Nuclear Engineering College, Rocket Force University of Engineering, Xi'an 710025, China;

2. Unit 96746 of PLA, Korla 841000, China)

**Abstract:** In order to solve the problem of estimation mismatch of the direction of arrival (DOA) in the beamforming of planar antenna arrays, a method of adding linear constraints near the direction of the signal of interest (SOI) is used to improve the robustness of beamforming. In addition, to address the problem that adding linear constraints will lead to a reduction in the degree of freedom of the algorithm, taking the uniform linear array as an example, a blocking matrix pre-selection link is added to the generalized sidelobe cancellation (GSC) model. By properly addressing the degrees of freedom loss brought on by the addition of linear constraints, the new approach increases resilience while preserving the original degree of freedom. Finally, the effectiveness of the proposed algorithm is verified by experimental results.

**Keywords:** robust beamforming; direction of arrival mismatch; linearly constrained minimum variance; generalized sidelobe canceller; blocking matrix

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0267

# 火箭整流罩锥壳夹层结构不确定性轻量化设计

董欣心, 岳振江\*, 王志, 刘莉

(北京理工大学 宇航学院, 北京 100081)

**摘 要:** 为分析运载火箭整流罩锥壳夹层结构不确定性对结构热稳定性的影响并指导结构的轻量化设计, 建立整流罩前锥段夹层圆锥壳模型, 并建立温度场模型, 据此对圆锥壳开展热稳定性分析, 推导力热联合载荷作用下整流罩前锥段夹层结构失稳临界轴压。在此基础上, 针对主要不确定性因素, 开展灵敏度分析并建立区间不确定性优化模型, 采用区间可能度方法将其转化为确定性问题, 并采用遗传算法-区间分析算法 (GA-CIAM) 实现结构优化设计。计算结果表明: 考虑气动力/热载荷及材料参数不确定性影响, 对整流罩前锥段结构开展优化设计, 在满足设计要求的前提下, 有效实现结构轻量化。

**关键词:** 整流罩锥壳; 夹层结构; 热稳定性分析; 区间不确定性优化; 轻量化设计

**中图分类号:** V421.1

**文献标志码:** A

**文章编号:** 1001-5965(2023)03-0625-11

随着航天技术不断进步, 火箭的飞行速度也在不断提高, 这对整流罩的综合性能提出了更高要求。火箭高速飞行时, 整流罩表面驻点温度近似正比于飞行速度的平方<sup>[1]</sup>, 摩擦带来的气动热引起整流罩壁的快速升温, 由于外部环境对结构的约束及各部分之间的相互作用, 结构不能完全自由的热胀冷缩, 引起很高的热应力。

针对这一问题, 国内外许多学者开展了圆柱壳及圆锥壳的热稳定性分析相关研究。Meyers等<sup>[2]</sup>分析了对称铺层复合材料平板的热屈曲响应, 并讨论了不同方向铺层及不同边界条件下的热屈曲情况。此后, 许多国外学者基于一阶剪应变理论对功能分级材料 (functionally gradient materials, FGM) 开展了一系列研究<sup>[3-6]</sup>, 对不同隔热材料、不同加筋形式的圆柱壳、圆锥壳结构在力载荷作用下的屈曲临界载荷及在温度载荷作用下的屈曲临界温度进行分析, 推导得到了其理论解。

在国内, 也有许多学者针对薄壳的热稳定性问题开展研究。张东等<sup>[7]</sup>对轴压和温度载荷作用下

圆柱加筋壳的热屈曲问题进行分析, 通过推导给出了临界屈曲载荷和临界温度载荷的解析表达。整流罩多采用复合材料, 复合材料能够有效提升结构的隔热性能及热载荷承受能力, 针对复合材料的热稳定性问题, 唐统帅<sup>[8]</sup>分析了复合材料格栅加筋结构的热应力分布规律, 并对加筋板结构尺寸开展设计以提升其散热性能和承载性能。徐腾飞等<sup>[9]</sup>基于经典薄板理论, 利用分离变量法对柔性支撑 FGM 矩形薄板的热屈曲进行分析, 推导得到了其封闭式解。

整流罩锥段需要同时承受轴压载荷和气动热引起的热载荷, 对于圆锥壳在力热联合载荷作用下的屈曲分析, 相关研究依然较少。并且, 飞行过程中所受载荷及结构材料的力学性能参数、热物性参数均存在一定的不确定性。现有的分析及基于此开展的结构设计并未充分考虑参数不确定引起的结构承载能力不确定性, 给定的安全系数通常存在较大冗余。近年来, 不确定性优化设计<sup>[10-11]</sup>的概念逐渐在设计领域受到重视, 其前提是不确定性量化

收稿日期: 2021-05-24; 录用日期: 2021-10-19; 网络出版时间: 2021-11-09 09:58

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20211108.1605.002.html](https://kns.cnki.net/kcms/detail/11.2625.V.20211108.1605.002.html)

\*通信作者. E-mail: [mountain\\_yue@bit.edu.cn](mailto:mountain_yue@bit.edu.cn)

**引用格式:** 董欣心, 岳振江, 王志, 等. 火箭整流罩锥壳夹层结构不确定性轻量化设计 [J]. 北京航空航天大学学报, 2023, 49 (3): 625-635.  
DONG X X, YUE Z J, WANG Z, et al. Uncertainty lightweight design of sandwich structure of rocket fairing cone [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 625-635 (in Chinese).

(uncertainty quantification, UQ), 主要思想是在设计过程中通过对模型输出响应的不确定性定量描述提升模型的可信度。不确定性优化设计可以通过定量分析承载能力波动范围, 使设计结果更加精确, 在降低结构质量的同时保证可靠性。

不确定性优化设计方法旨在通过稳健性设计获取可行域内对于设计变量变化不敏感的最优值, 按处理方法可以分为概率方法、非概率方法和概率-非概率混合方法。通常, 如果不确定性因素的概率分布已知可以采用概率方法, 若分布未知而变化范围已知可以采用区间方法, 若部分参数分布已知而部分只能确定变化区间则可以采用混合方法。其中区间方法对于数据的要求较低, 设计简单, 因此在工程上使用更加广泛。

本文考虑整流罩部分承受较强的气动加热, 存在热应力, 建立整流罩锥段的夹层结构模型, 分析轴压载荷和气动热载荷联合作用下锥壳的载荷承受能力。建立区间不确定性优化模型, 采用遗传算法嵌套区间分析方法求解, 对整流罩前锥段夹层结构开展结构轻量化设计, 并对优化结果进行对比分析。

# 1 圆锥壳热稳定性分析方法

## 1.1 整流罩前锥段圆锥壳模型建立

整流罩前锥段常采用复合材料结构, 由上下2层面板及中间的夹心芯子组成, 还包括一些加强框、加强板及支撑桁、爆炸螺栓等连接和加强结构。为对其进行参数化表征, 作出一定假设, 认为前锥段为完整的复合材料结构, 不考虑局部连接和加强结构, 建立参数化模型如图1所示。沿壳长度为  $x$  方向,  $x$  方向的位移为  $u$ , 垂直于  $x$  方向为壳的厚度方向  $z$  方向,  $z$  方向的位移为  $w$ 。

为计算热稳定性, 需对面板及芯子的材料性能参数进行定义。性能参数主要包括材料的弹性模量、泊松比, 以及描述物体在温度改变时膨胀收缩能力的热膨胀系数。热膨胀系数的定义为在等压条件下, 单位温度变化所导致的体积变化。本文为

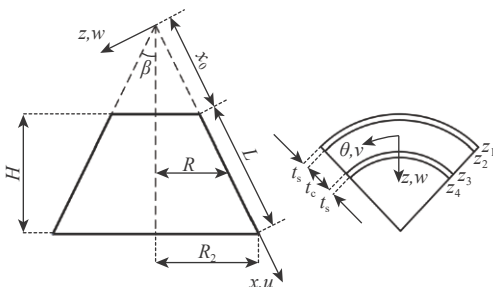


图1 前锥段夹层结构参数化模型

Fig. 1 Parameterized model of sandwich structure of the front cone section

简化计算, 设面板和芯子泊松比相等。

## 1.2 温度场模型建立

锥壳所受热应力与结构内部温度分布相关, 因此, 需要建立温度场模型, 分析在热流作用下结构内部温度分布情况。壁面厚度方向的尺寸远小于另外2个方向, 此时可以把锥壳体面近似视为平板进行考虑。在壁面内, 可以假设热传导只沿壁厚方向进行, 将原本的三维传热问题简化为一维传热问题<sup>[12]</sup>。

传热边界条件为: 结构层外表面净吸收热流密度为常值  $q_0$ , 内壁为绝热条件。初始时刻, 结构层沿厚度方向不存在温差, 即  $t = 0$  时刻有  $T_0 = T_n = T_{init}$ 。由此初始条件计算一段时间后结构层沿厚度方向的温度分布, 采用有限差分法, 将结构层分为  $N$  层, 以每层中点处的温度代表该层温度, 如图2所示, 通过热传导傅里叶方程描述温度空间分布随时间变化。

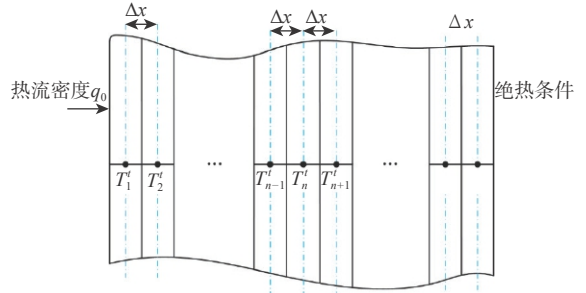


图2 沿壁厚方向温度分布

Fig. 2 Temperature distribution along thickness direction

热传导傅里叶方程为

$$\rho c \frac{\partial T}{\partial t} = \frac{\partial}{\partial x} \left( k \frac{\partial T}{\partial x} \right) \tag{1}$$

式中:  $k$  为材料的热导率;  $\rho$  为材料密度;  $c$  为材料比热容。温度对时间的导数可以表示为

$$\frac{\partial T}{\partial t} = \frac{T_n^i - T_n^{i-1}}{\Delta t} \tag{2}$$

式中:  $\Delta t$  为时间差分的时间间隔;  $T_n^j$  为  $j$  时刻第  $n$  层位置处的温度。温度梯度可以表示为

$$\frac{\partial}{\partial x} \left( k \frac{\partial T}{\partial x} \right) = \frac{k_{n+1/2}^{i-1} \frac{T_{n+1}^i - T_n^i}{\Delta x} - k_{n-1/2}^{i-1} \frac{T_n^i - T_{n-1}^i}{\Delta x}}{\Delta x} \tag{3}$$

式中:  $k_i^j$  为  $j$  时刻第  $n$  层位置处的热导率;  $\Delta x$  为单层厚度;  $k_{n+1/2}^{i-1}$  和  $k_{n-1/2}^{i-1}$  分别由其相邻时刻  $k$  值的调和平均求得, 满足关系:

$$\begin{cases} \frac{2}{k_{n+1/2}^{i-1}} = \frac{1}{k_n^{i-1}} + \frac{1}{k_{n+1}^{i-1}} \\ \frac{2}{k_{n-1/2}^{i-1}} = \frac{1}{k_{n-1}^{i-1}} + \frac{1}{k_n^{i-1}} \end{cases} \tag{4}$$

引入传热边界条件, 在外表面处给定净吸收热

流密度  $q_0$ , 在表面外与第一个节点  $x_1$  距离  $\Delta x$  处构造虚拟节点  $x_0$ , 在内壁内侧构造虚拟节点  $x_{N+1}$ , 则有:

$$\begin{cases} -k_1 \frac{T_2 - T_0}{2\Delta x} = q_0 \\ -k_N \frac{T_{N+1} - T_{N-1}}{2\Delta x} = 0 \end{cases} \quad (5)$$

此时可以将式 (3) 改写为矩阵形式, 并代入边界条件, 得到:

$$\mathbf{A} \begin{bmatrix} T'_1 \\ T'_2 \\ \vdots \\ T'_{N-1} \\ T'_N \end{bmatrix} = \begin{bmatrix} T'^{-1}_1 + 2 \frac{q_0 \Delta t}{\rho c \Delta x} \\ T'^{-1}_2 \\ \vdots \\ T'^{-1}_{N-1} \\ T'^{-1}_N \end{bmatrix} \quad (6)$$

式中: 矩阵  $\mathbf{A}$  为

$$\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} & 0 & \cdots & 0 \\ A_{21} & A_{22} & A_{23} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & A_{N-1,N-2} & A_{N-1,N-1} & A_{N-1,N} \\ 0 & \cdots & 0 & A_{N,N-1} & A_{N,N} \end{bmatrix} \quad (7)$$

矩阵中系数表达式为  $A_{N,N} = 1 + \frac{\Delta t}{\rho c \Delta x^2} (k'^{-1}_{N-1/2} + k'^{-1}_N)$ ,  $A_{11} = 1 + \frac{\Delta t}{\rho c \Delta x^2} (k'^{-1}_1 + k'^{-1}_{1+1/2})$ ,  $A_{i,i+1} = -\frac{\Delta t}{\rho c \Delta x^2} k'^{-1}_{i+1/2}$ ,  $A_{ii} = 1 + \frac{\Delta t}{\rho c \Delta x^2} (k'^{-1}_{i-1/2} + k'^{-1}_{i+1/2})$ ,  $A_{i+1,i} = -\frac{\Delta t}{\rho c \Delta x^2} k'^{-1}_{i+1/2}$ ,  $i = 2, 3, \dots, N-1$ 。

由此得到了由  $t-1$  时刻推至  $t$  时刻结构温度分布的表达式, 结合初始时刻温度分布即可求得任一时刻温度分布。

### 1.3 圆锥壳热稳定性分析

依据文献 [5], 将应力-应变方程进行整合并在壳体厚度方向积分取矩, 可以得到锥壳的力、力矩合力及横向力合力表达式为

$$\begin{cases} N_x = A_{11}\varepsilon_{xm} + A_{12}\varepsilon_{\theta m} + B_{11}k_x + B_{12}k_\theta + \Phi_a \\ N_\theta = A_{12}\varepsilon_{xm} + A_{22}\varepsilon_{\theta m} + B_{12}k_x + B_{22}k_\theta + \Phi_a \\ N_{x\theta} = A_{66}\gamma_{x\theta m} + 2B_{66}k_{x\theta} \\ M_x = B_{11}\varepsilon_{xm} + B_{12}\varepsilon_{\theta m} + D_{11}k_x + D_{12}k_\theta + \Phi_b \\ M_\theta = B_{12}\varepsilon_{xm} + B_{22}\varepsilon_{\theta m} + D_{12}k_x + D_{22}k_\theta + \Phi_b \\ M_{x\theta} = B_{66}\gamma_{x\theta m} + 2D_{66}k_{x\theta} \\ Q_x = A_{44}\gamma_{xz m} = A_{44}(w_x + \phi_x) \\ Q_\theta = A_{55}\gamma_{\theta z m} = A_{55}\left(\frac{1}{x \sin \beta} w_\theta + \phi_\theta\right) \end{cases} \quad (8)$$

式中: 系数  $A_{ij}$ 、 $B_{ij}$ 、 $D_{ij}$  及  $\Phi_a$ 、 $\Phi_b$  表达式为

$$\begin{cases} A_{11} = A_{22} = \frac{E_1}{1-\nu^2}, A_{12} = \frac{E_1\nu}{1-\nu^2} \\ A_{44} = A_{55} = \frac{5E_1}{12(1+\nu)}, A_{66} = \frac{E_1}{2(1+\nu)} \\ B_{11} = B_{22} = \frac{E_2}{1-\nu^2}, B_{12} = \frac{\nu E_2}{1-\nu^2}, B_{66} = \frac{E_2}{2(1+\nu)} \\ D_{11} = D_{22} = \frac{E_3}{1-\nu^2}, D_{12} = \frac{\nu E_3}{1-\nu^2}, D_{66} = \frac{E_3}{2(1+\nu)} \\ \Phi_a = -\int_{-\frac{h}{2}}^{\frac{h}{2}} \frac{E_{sh}\alpha_{sh}\Delta T(z)}{1-\nu} dz \\ \Phi_b = -\int_{-\frac{h}{2}}^{\frac{h}{2}} \frac{E_{sh}\alpha_{sh}\Delta T(z)z}{1-\nu} dz \end{cases} \quad (9)$$

式中:  $\nu$  为材料泊松比。

$$\begin{cases} E_1 = E_s(z_4 - z_3) + E_c(z_3 - z_2) + E_s(z_2 - z_1) = 2t_s E_s + t_c E_c \\ E_2 = 0 \\ E_3 = \frac{E_s(z_4 - z_1)^3}{12} + \frac{(E_c - E_s)(z_3 - z_2)^3}{12} = \frac{E_s(2t_s + t_c)^3}{12} + \frac{(E_c - E_s)t_c^3}{12} \end{cases} \quad (10)$$

式中:  $E_s$  为面板弹性模量;  $E_c$  为芯子弹性模量。

$$\begin{cases} E_{sh} = \begin{cases} E_s & z_1 \leq z \leq z_2 \\ E_c & z_2 \leq z \leq z_3 \\ E_s & z_3 \leq z \leq z_4 \end{cases} \\ \alpha_{sh} = \begin{cases} \alpha_s & z_1 \leq z \leq z_2 \\ \alpha_c & z_2 \leq z \leq z_3 \\ \alpha_s & z_3 \leq z \leq z_4 \end{cases} \end{cases} \quad (11)$$

式中:  $\alpha_s$  为面板热膨胀系数;  $\alpha_c$  为芯子热膨胀系数。

依据相邻平衡准则, 平衡路径上的每个位移分量都会受到固有扰动, 以建立新的平衡。因此, 屈曲状态下结构位移场可以写作屈曲前平衡状态位移加扰动的形式<sup>[13]</sup>:

$$\begin{cases} u = u_0 + u_1 \\ v = v_0 + v_1 \\ w = w_0 + w_1 \\ \phi_x = \phi_{x0} + \phi_{x1} \\ \phi_\theta = \phi_{\theta0} + \phi_{\theta1} \end{cases} \quad (12)$$

式中:  $u_0$ 、 $v_0$ 、 $w_0$  和  $\phi_{x0}$ 、 $\phi_{\theta0}$  分别为屈曲前平衡状态位移、转角场;  $u_1$ 、 $v_1$ 、 $w_1$  和  $\phi_{x1}$ 、 $\phi_{\theta1}$  分别为扰动位移、转角场。

由此, 屈曲状态下的力和力矩的合力也可表示为平衡状态量加扰动量:

$$\begin{cases} N_x = N_{x0} + N_{x1} \\ N_\theta = N_{\theta0} + N_{\theta1} \\ N_{x\theta} = N_{x\theta0} + N_{x\theta1} \\ M_x = M_{x0} + M_{x1} \\ M_\theta = M_{\theta0} + M_{\theta1} \\ M_{x\theta} = M_{x\theta0} + M_{x\theta1} \\ Q_x = Q_{x0} + Q_{x1} \\ Q_\theta = Q_{\theta0} + Q_{\theta1} \end{cases} \quad (13)$$

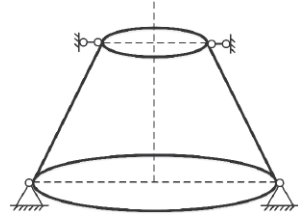


图3 圆锥壳边界条件

Fig. 3 Boundary conditions of conical shell

式中:下标0表示平衡状态;下标1表示扰动场。已知下标为0的项为平衡状态项,本身即满足平衡方程,因此将其略去。对于扰动项,由于高阶项远小于一阶项,因此将高阶项略去,仅保留一阶项,得到平衡方程:

$$\begin{cases} xN_{x1,x} + \frac{1}{\sin\beta}N_{x\theta1,x} + N_{x1} - N_{\theta1} = 0 \\ \frac{1}{\sin\beta}N_{\theta1,\theta} + xN_{x\theta1,x} + 2N_{x\theta1} = 0 \\ xM_{x1,xx} + 2M_{x1,x} + \frac{2}{\sin\beta}M_{x\theta1,x\theta} + \frac{2}{x\sin\beta}M_{x\theta1,\theta} + \\ \frac{2}{x\sin^2\beta}M_{\theta1,\theta\theta} - M_{\theta1,x} - N_{\theta1}\cot\beta + \\ \left(xN_{x0}w_{1,x} + \frac{1}{\sin\beta}N_{x\theta0}w_{1,\theta}\right)_{,x} + \frac{1}{\sin\beta} \cdot \\ \left(N_{x\theta0}w_{1,x} + \frac{1}{x\sin\beta}N_{\theta0}w_{1,\theta}\right)_{,\theta} = 0 \\ (x\sin\beta M_{x1})_{,x} + M_{x\theta1,\theta} - M_{\theta1}\sin\beta - x\sin\beta Q_{x1} = 0 \\ (x\sin\beta M_{x\theta1})_{,x} + M_{\theta1,\theta} + M_{x\theta1}\sin\beta - x\sin\beta Q_{\theta1} = 0 \end{cases} \quad (14)$$

采用伽辽金方法对式(14)进行求解。认为圆锥壳两端为简支边界条件,如图3所示。

由此可知在壳两端 $x = x_0$ 及 $x = x_0 + L$ 处满足:

$$N_{x1} = v_1 = w_1 = \phi_{\theta1} = M_{x1} = 0 \quad (15)$$

满足边界条件的近似解可以写为如下形式:

$$\begin{cases} u_1 = A \cos \frac{m\pi(x-x_0)}{L} \sin n\theta \\ v_1 = B \sin \frac{m\pi(x-x_0)}{L} \cos n\theta \\ w_1 = C \sin \frac{m\pi(x-x_0)}{L} \sin n\theta \\ \phi_{x1} = \Phi_1 \cos \frac{m\pi(x-x_0)}{L} \sin n\theta \\ \phi_{\theta1} = \Phi_2 \frac{1}{x\sin\alpha} \sin \frac{m\pi(x-x_0)}{L} \cos n\theta \end{cases} \quad (16)$$

式中: $m$ 为轴向失稳半波数; $n$ 为环向失稳波数; $A$ 、 $B$ 、 $C$ 、 $\Phi_1$ 、 $\Phi_2$ 均为常系数; $L$ 为锥段长度。这一形式已经满足了在壳两端 $v_1 = w_1 = \phi_{\theta1} = 0$ 的边界条件,为满足 $N_{x1} = M_{x1} = 0$ ,还需要有:

$$\begin{cases} \int_0^{2\pi} \left( \frac{\partial N_{x1}}{\partial x} \right)_{x=x_0, x_0+L} d\theta = 0 \\ \int_0^{2\pi} \left( \frac{\partial M_{x1}}{\partial x} \right)_{x=x_0, x_0+L} d\theta = 0 \end{cases} \quad (17)$$

将式(16)代入微分方程,得到变形后的方程如下:

$$\begin{cases} L_{11}A + L_{12}B + L_{13}C + L_{14}\Phi_1 + L_{15}\Phi_2 = 0 \\ L_{21}A + L_{22}B + L_{23}C + L_{24}\Phi_1 + L_{25}\Phi_2 = 0 \\ L_{31}A + L_{32}B + (L_{33} + \bar{N}_{x0}L_{36} + N_{\theta0}L_{37})C + \\ L_{34}\Phi_1 + L_{35}\Phi_2 = 0 \\ L_{41}A + L_{42}B + L_{43}C + L_{44}\Phi_1 + L_{45}\Phi_2 = 0 \\ L_{51}A + L_{52}B + L_{53}C + L_{54}\Phi_1 + L_{55}\Phi_2 = 0 \end{cases} \quad (18)$$

式中: $\bar{N}_{x0} = xN_{x0}$ ;系数 $L_{ij}$ 见附录A。解此方程可得临界热屈曲载荷为

$$\bar{N}_{x0} = -\frac{L_{31}}{L_{36}} \cdot \frac{D_1}{D_3} + \frac{L_{32}}{L_{36}} \cdot \frac{D_2}{D_3} + \frac{L_{34}}{L_{36}} \cdot \frac{D_4}{D_3} - \frac{L_{35}}{L_{36}} \cdot \frac{D_5}{D_3} - \frac{L_{37}}{L_{36}} N_{\theta0} - \frac{L_{33}}{L_{36}} \quad (19)$$

式中:系数 $D_i$ 分别为

$$D_1 = \begin{vmatrix} L_{12} & L_{13} & L_{14} & L_{15} \\ L_{22} & L_{23} & L_{24} & L_{25} \\ L_{42} & L_{43} & L_{44} & L_{45} \\ L_{52} & L_{53} & L_{54} & L_{55} \end{vmatrix}, D_2 = \begin{vmatrix} L_{11} & L_{13} & L_{14} & L_{15} \\ L_{21} & L_{23} & L_{24} & L_{25} \\ L_{41} & L_{43} & L_{44} & L_{45} \\ L_{51} & L_{53} & L_{54} & L_{55} \end{vmatrix}$$

$$D_3 = \begin{vmatrix} L_{11} & L_{12} & L_{14} & L_{15} \\ L_{21} & L_{22} & L_{24} & L_{25} \\ L_{41} & L_{42} & L_{44} & L_{45} \\ L_{51} & L_{52} & L_{54} & L_{55} \end{vmatrix}, D_4 = \begin{vmatrix} L_{11} & L_{12} & L_{13} & L_{15} \\ L_{21} & L_{22} & L_{23} & L_{25} \\ L_{41} & L_{42} & L_{43} & L_{45} \\ L_{51} & L_{52} & L_{53} & L_{55} \end{vmatrix}$$

$$D_5 = \begin{vmatrix} L_{11} & L_{12} & L_{13} & L_{14} \\ L_{21} & L_{22} & L_{23} & L_{24} \\ L_{41} & L_{42} & L_{43} & L_{44} \\ L_{51} & L_{52} & L_{53} & L_{54} \end{vmatrix}$$

在火箭飞行的真实情况下,锥壳轴向允许膨胀,因此,温度场仅引起轴向应变而不引起轴向热应力。由此可得力热联合作用条件下屈曲前稳定平衡状态的薄膜内力为

$$\begin{cases} N_{x0} = -\frac{P}{\pi x \sin 2\beta} \\ \bar{N}_{x0} = xN_{x0} = -\frac{P}{\pi \sin 2\beta} \\ N_{\theta0} = -\frac{x_0 + L}{x} \int_{-h/2}^{h/2} \frac{T(z)E(z)\alpha(z)}{1-\nu} dz \\ N_{x\theta0} = 0 \end{cases} \quad (20)$$

式中: $P$ 为火箭推力; $h$ 为壳段厚度。

将式(19)代入式(20),通过循环比较确定 $m$ 和 $n$ 对应的最小屈曲载荷即为圆锥壳的临界屈曲载

荷, 其表达式为

$$P_{cr} = -\frac{\pi \sin 2\beta}{L_{36}} \left( -L_{31} \frac{D_1}{D_3} + L_{32} \frac{D_2}{D_3} + L_{34} \frac{D_4}{D_3} - L_{35} \frac{D_5}{D_3} - L_{33} + L_{37} \frac{x_0 + L}{x_0} \int_{-h/2}^{h/2} \frac{T(z)E(z)\alpha(z)}{1-\nu} dz \right) \quad (21)$$

## 2 区间不确定性优化方法

### 2.1 区间不确定性优化模型

区间不确定性优化问题在确定性优化问题的基础上引入不确定性变量  $U$ , 此时目标及约束函数同时为关于设计变量  $X$  及不确定性变量  $U$  的函数, 其优化模型表示为

$$\begin{cases} \min_X f(X, U) \\ \text{s.t. } g_i(X) \leq (=, \geq) b_i^l = [b_i^l, b_i^R] \\ i = 1, 2, \dots, p; X \in \Omega^x \\ U \in U^l = [U^l, U^R]; U_j \in U_j^l = [U_j^l, U_j^R] \\ j = 1, 2, \dots, q \end{cases} \quad (22)$$

式中:  $X$  为  $n$  维设计变量, 设计空间为  $\Omega^x$ ;  $U$  为  $q$  维不确定性变量,  $U^l$  为其下边界,  $U^R$  为其上边界;  $f$  为目标函数,  $g_i$  为第  $i$  个约束函数, 二者均为  $X$  和  $U$  的连续函数;  $b_i^l$  为第  $i$  个不确定性约束区间。

对于一个确定的设计点, 不确定性优化问题中的目标函数为不确定性变量  $U$  的函数, 可将其视为一个区间数。优化过程中需要通过目标函数值比较设计点的优劣进行决策, 采用区间序方法对区间数进行比较, 对于一个区间  $A$ , 则可以通过定义区间中点及半径的方式将其定义为  $A = \langle a^c, a^w \rangle = \{x | a^c - a^w \leq x \leq a^c + a^w\}$ , 其中  $a^c$ 、 $a^w$  分别为区间的中点和半径。针对最小化问题的区间序关系定义为:  $A \leq_{cw} B$  当且仅当  $a^c \leq b^c$  且  $a^w \leq b^w$ ;  $A <_{cw} B$  当且仅当  $A \leq_{cw} B$  且  $A \neq B$ , 即在设计过程中更偏好于均值更优且不确定性更低的结果。

在处理不确定性约束时, 引入区间可能度方法<sup>[14]</sup>, 定量比较区间两区间大小关系。认为区间数  $A$  和  $B$  在区间内均匀分布, 区间  $A$  和  $B$  所有可能的相对关系可分为以下 6 种, 如图 4 所示。

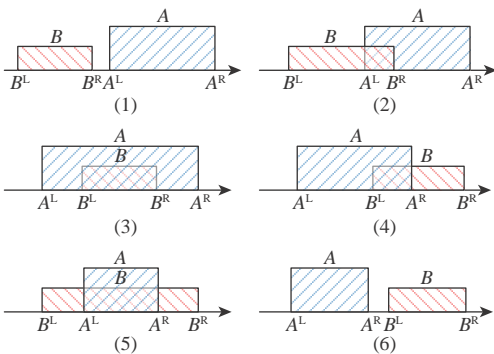


图 4 区间  $A$  与区间  $B$  可能的相对关系

Fig. 4 Possible relationship between interval  $A$  and  $B$

对应不同相对关系, 基于概率方法给出区间可能度模型:

$$P(A \leq B) = \begin{cases} 0 & A^L \geq B^R \\ \frac{B^R - A^L}{2(A^R - A^L)} \frac{B^R - A^L}{B^R - B^L} & B^L \leq A^L < B^R \leq A^R \\ \frac{B^L - A^L}{A^R - A^L} + \frac{B^R - B^L}{2(A^R - A^L)} & A^L < B^L < B^R \leq A^R \\ \frac{B^L - A^L}{A^R - A^L} + \frac{A^R - B^L}{A^R - A^L} \frac{B^R - A^R}{B^R - B^L} + \frac{A^R - B^L}{2(A^R - A^L)} \frac{A^R - B^L}{B^R - B^L} & A^L < B^L \leq A^R < B^R \\ \frac{B^R - A^R}{B^R - B^L} + \frac{A^R - A^L}{2(B^R - B^L)} & B^L \leq A^L < A^R < B^R \\ 1 & A^R < B^L \end{cases} \quad (23)$$

依据区间可能度方法, 可以将不确定性约束转化为确定性约束, 定义  $\lambda_i$  为第  $i$  个约束预定的区间可能度水平, 将原约束  $g_i(X, U) \leq b_i^l$  转化为

$$P(g_i(X, U) \leq b_i^l) \geq \lambda_i \quad (24)$$

式中:  $g_i(X) = [g_i^L(X), g_i^R(X)]$ ,  $g_i^L(X)$  及  $g_i^R(X)$  可通过对不确定性变量  $U$  进行优化或由区间分析方法获得。

将原不确定性问题的目标函数及约束函数分别转换为确定形式后, 得到新的优化模型:

$$\begin{cases} \min_X w f^c(X) + (1-w) f^w(X) \\ \text{s.t. } P(g_i(X, U) \leq b_i^l) \geq \lambda_i; i = 1, 2, \dots, p \\ X \in \Omega^x \end{cases} \quad (25)$$

式中:  $w \in [0, 1]$  为权重。

### 2.2 区间分析方法

传统的区间不确定性优化常采用遗传算法等智能优化算法确定目标及约束函数的上下界, 结合外层的设计变量寻优, 双层嵌套优化导致计算量巨大。采用区间分析方法取代内层优化可以有效提升计算效率, 常用的区间分析方法包括基于 Taylor 展开的区间分析方法 (Taylor interval analysis method, TIAM) 和由此发展出来的配点型区间分析方法 (collocation interval analysis method, CIAM)<sup>[15]</sup>。配点型区间分析方法由基于 Taylor 展开的区间分析方法发展而来, 具有更高的计算精度, 其思想为采用第 1 类 Chebyshev 多项式逼近响应函数, 获得响应函数的区间上下界。

对于一个有  $o$  个不确定性变量的问题, 对其中第  $i$  个不确定性变量进行逼近, 仅将其取为区间变量, 其他量均取中点值, 得到向量  $U_i = (U_1^c, U_2^c, \dots, U_{i-1}^c, U_i, U_{i+1}^c, \dots, U_o^c)^T$ 。在此引入第一类 Chebyshev 多

项式:

$$T_n(x) = \cos(n \arccos x)$$

$$x \in [-1, 1]; n = 0, 1, 2, \dots \quad (26)$$

将原区间变量变换至 $[-1, 1]$ 区间, 将 $U_i$ 表示为

$$U_i = U_i^c + x \Delta U_i^l \quad x \in [-1, 1] \quad (27)$$

将式(27)代入向量 $U_i$ , 得到 $U_i$ 的表达式为 $U_i = (U_1^c, U_2^c, \dots, U_{i-1}^c, U_i^c + x \Delta U_i^l, U_{i+1}^c, \dots, U_o^c)^T$ , 仅含第 $i$ 个不确定性变量的约束函数可以表示为

$$g(\mathbf{X}^*, U_i) = g[\mathbf{X}^*, (U_1^c, U_2^c, \dots, U_i^c + x \Delta U_i^l, \dots, U_o^c)^T]$$

$$x \in [-1, 1] \quad (28)$$

此函数仅包含一个变量 $x$ , 故可以将其简记为 $g(x)$ ,  $x \in [-1, 1]$ 。用 $r$ 阶多项式 $P_r(x)$ 逼近原函数,  $P_r(x)$ 的矩阵形式写作:

$$P_r(x) = \frac{2}{q} \begin{bmatrix} g(x_1) \\ g(x_2) \\ \vdots \\ g(x_q) \end{bmatrix}^T \cdot T(x) \quad (29)$$

$$T(x) = \begin{bmatrix} T_0(x) \\ T_1(x) \\ \vdots \\ T_r(x) \end{bmatrix}$$

式中: 矩阵 $T(x)$ 表示为

$$T(x) = \begin{bmatrix} 1/2 & T_1(x_1) & T_2(x_1) & \dots & T_r(x_1) \\ 1/2 & T_1(x_2) & T_2(x_2) & \dots & T_r(x_2) \\ 1/2 & \vdots & \vdots & & \vdots \\ 1/2 & T_1(x_q) & T_2(x_q) & \dots & T_r(x_q) \end{bmatrix} \quad (30)$$

此时 $P_r(x)$ 写为了关于变量 $x$ 的 $j$ 阶表达式, 可以依据多项式性质求得其极值点位置, 得到约束函数 $g(\mathbf{X}^*, U_i)$ 在第 $i$ 个不确定性变量区间上的最大值和最小值点, 此时对应的 $U_i$ 分别记为 $U_i^{\max}$ 和 $U_i^{\min}$ , 对每个不确定性变量依次通过上述过程求得极值点位置, 最终得到约束 $g(\mathbf{X}^*, U_i)$ 的最大值点和最小值点分别为

$$\begin{cases} U^{\max} = [U_1^{\max}, U_2^{\max}, \dots, U_o^{\max}] \\ U^{\min} = [U_1^{\min}, U_2^{\min}, \dots, U_o^{\min}] \end{cases} \quad (31)$$

将式(31)代入式(28)求得其最大值和最小值分别为

$$\begin{cases} g^{\max}(\mathbf{X}^*, U) = g(\mathbf{X}^*, U^{\max}) \\ g^{\min}(\mathbf{X}^*, U) = g(\mathbf{X}^*, U^{\min}) \end{cases} \quad (32)$$

当优化模型中有多个约束时, 对每个约束依次计算上下界, 然后采用区间可能度方法计算。

### 2.3 不确定性优化流程

采用遗传算法-区间分析方法(GA-CIAM)求解不确定性优化问题。优化流程如图5所示。

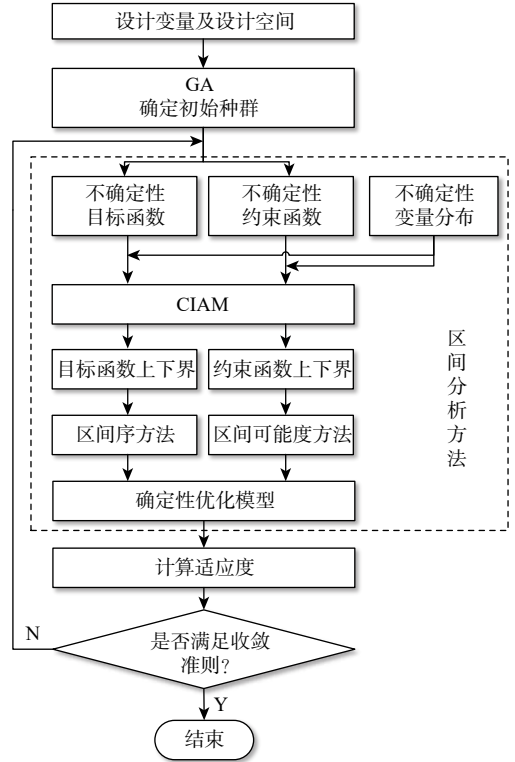


图5 不确定性优化流程

Fig. 5 Flow diagram of uncertainty optimization

## 3 优化模型建立及结果分析

### 3.1 优化模型建立

由于在结构设计阶段, 整体外形尺寸往往已经由总体设计确定, 因此, 给定其整体外形参数, 仅对夹层结构厚度展开不确定性优化设计。锥壳锥角为 $17^\circ$ , 壳长 $4.2\text{ m}$ , 小端半径为 $0.63\text{ m}$ , 大端半径为 $1.85\text{ m}$ 。面板、芯子泊松比均为 $0.3$ , 玻璃板热膨胀系数为 $7.1 \times 10^{-6} (1/^\circ\text{C})$ , 泡沫芯子热膨胀系数为 $8 \times 10^{-5} (1/^\circ\text{C})$ , 结构初始温度为均匀分布的 $20\text{ }^\circ\text{C}$ , 净吸收热流密度为 $20\text{ kW/m}^2$ , 作用时间 $150\text{ s}$ 。为计算结构质量, 需给定材料密度, 玻璃板密度为 $1.8 \times 10^{-6}\text{ kg/mm}^3$ , 泡沫夹层密度为 $3 \times 10^{-7}\text{ kg/mm}^3$ 。设计变量为面板及芯子厚度, 设计空间上下界和初始值如表1所示。不确定性变量为需用轴压、热流参数及材料的力热性能参数。不确定性变量上下界如表2所示。

为描述不同因素对响应不确定性的影响程度, 通常进行灵敏度分析, 方差灵敏度分析法又称为Sobol方法, 通过输入变量的方差在系统响应中的

表1 设计变量设计空间及初值

Table 1 Design space and initial value of design variables

参数	面板厚度 $t_s/\text{mm}$	芯子厚度 $t_c/\text{mm}$
设计空间	[0.5, 3]	[2, 50]
初始值	1.5	18

表 2 不确定性变量上下界

Table 2 upper and lower bounds of uncertainty variables

不确定性变量	变量上下界
需用轴压载荷 $F/\text{MN}$	[14.25, 15.75]
净吸收热流 $q_0/(\text{kW}\cdot\text{m}^{-2})$	[19, 21]
热流作用时间 $t/\text{s}$	[142.5, 157.5]
玻璃板弹性模量 $E_s/\text{MPa}$	[68 400, 75 600]
泡沫夹层弹性模量 $E_c/\text{MPa}$	[66.5, 73.5]
玻璃板比热容 $c_s/(\text{kJ}\cdot(\text{kg}\cdot\text{K})^{-1})$	[1.425, 1.575]
泡沫夹层比热容 $c_c/(\text{kJ}\cdot(\text{kg}\cdot\text{K})^{-1})$	[1.188, 1.313]
玻璃板导热率 $k_s/(\text{kW}\cdot(\text{m}\cdot\text{K})^{-1})$	$[3.04\times 10^{-4}, 3.36\times 10^{-4}]$
泡沫夹层导热率 $k_c/(\text{kW}\cdot(\text{m}\cdot\text{K})^{-1})$	$[2.85\times 10^{-5}, 3.15\times 10^{-5}]$

贡献度来衡量该变量的重要程度, 能充分表征变量对响应的影响程度, 因此, 采用 Sobol 法分析灵敏度, Sobol 指数的取值在 [0, 1] 之间, 取值越大说明响应对输入的变化更加灵敏。本文对热流参数, 分别是净吸收热流  $q_0$  和热流作用时间  $t$ , 以及材料热物性参数, 分别是玻璃板比热容、泡沫夹层比热容、玻璃板导热率、泡沫夹层导热率进行灵敏度分析。结果如图 6 和图 7 所示。

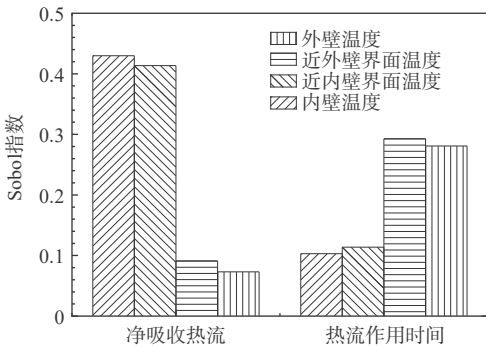


图 6 热流不确定性因素 Sobol 灵敏度指数

Fig. 6 Sobol sensitivity index of heat flux uncertainties

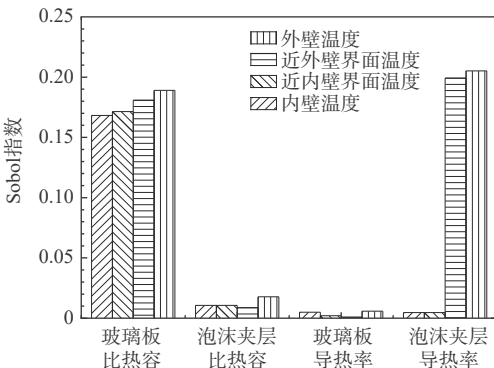


图 7 材料不确定性因素 Sobol 灵敏度指数

Fig. 7 Sobol sensitivity index of material uncertainties

从灵敏度分析结果图 6、图 7 可知, 热流不确定性对温度分布影响较大, 其中净吸收热流密度主要影响外壁及近外壁界面温度, 内壁及近内壁界面温度对热流作用时间更敏感。材料不确定性因素对响应影响差别较大, 泡沫层比热容和玻璃板导热率

对响应灵敏度很小, 玻璃板比热容对各层的影响相似, 泡沫夹层导热率对外壁温度基本无影响, 对内壁温度起主要作用。

以结构质量作为目标函数, 约束为结构不失稳且内壁温度满足温度限制。结构质量为结构尺寸参数的函数, 表示为

$$M = (\pi R_1 L + \pi R_2 L)(2t_s \rho_s + t_c \rho_c) \quad (33)$$

式中:  $R_1$  为锥壳小端半径;  $R_2$  为锥壳大端半径;  $L$  为锥段长度;  $\rho_s$  为面板材料密度;  $\rho_c$  为芯子材料密度。

约束条件首先考虑结构在力热联合条件下的稳定性问题, 要求结构在轴压和热应力作用下失稳临界轴压不大于需用轴压载荷, 同时, 要求结构内壁温度低于临界温度, 本算例设定为  $T_{cr} = 40\text{ }^\circ\text{C}$ 。计算过程中, 对某一设计点, 需首先由热传导求得带有不确定性的结构内外温度上下边界, 再将温度及材料参数的不确定性共同作为稳定性分析的输入, 求得结构临界轴压载荷的上下边界。由此可以得到前锥段夹层结构优化问题的不确定性优化模型为

$$\begin{cases} \min M(X) \\ \text{s.t. } F^l \leq F_{cr}^l(X; E_s^l, E_c^l, q_0^l, t^l, c_s^l, c_c^l, k_s^l, k_c^l) \\ T_{in}^l(X; q_0^l, t^l, c_s^l, c_c^l, k_s^l, k_c^l) \leq T_{cr} \\ X \in \Omega^n \end{cases} \quad (34)$$

式中: 设计变量  $X = [t_s, t_c]$ ;  $\Omega^n$  为  $X$  的设计空间; 上标  $l$  表示区间;  $F_{cr}$  为力热载荷联合作用下的失稳临界轴压,  $T_{in}$  为结构内壁温度, 均为设计变量和区间不确定性变量的函数, 其余符号同前文所述。通过区间序及区间可能度方法将其转换为确定性优化模型如下:

$$\begin{cases} \min M(X) \\ \text{s.t. } P(F^l \leq F_{cr}^l(X; E_s^l, E_c^l, q_0^l, t^l, c_s^l, c_c^l, k_s^l, k_c^l)) \geq \lambda_1 \\ P(T_{in}^l(X; q_0^l, t^l, c_s^l, c_c^l, k_s^l, k_c^l) \leq T_{cr}) \geq \lambda_2 \\ X \in \Omega^n \end{cases} \quad (35)$$

式中:  $\lambda_i$  为不同约束的区间可能度。

### 3.2 优化结果及分析

优化方法选用 GA-CIAM 方法, 外层采用 GA 对设计变量寻优, 内层通过 CIAM 方法确定约束上下界。本问题仅有 2 个设计变量, 故将遗传算法的种群规模设为 25, 最大代数 100, 交叉概率 0.8, 变异概率 0.1, 罚函数法的初始罚函数为 10, 罚因子为 100。CIAM 方法多项式阶数为 5 阶, 高斯积分点数为 10, 将不同约束的区间可能度都选为 0.8 进行优化。得到设计结果为面板厚度  $t_s = 1.600\text{ mm}$ , 芯子厚度  $t_c = 16.500\text{ mm}$ , 最终结构质量为  $352.674\text{ kg}$ 。

将所得优化结果的目标函数及约束的性能参数与优化前进行对比, 结果如表 3 所示。

表3 优化结果对比

Table 3 Comparison of optimization results

优化	设计变量/mm	结构质量/kg	临界轴压/MN	内壁温度/℃	外壁温度/℃
优化前	[2/18]	414.206	[21.676,24.826]	[20.002,20.003]	[572.055,611.647]
优化后	[1.600/16.500]	352.674	[14.230,18.087]	[20.002,20.019]	[603.321,769.212]

从结果表3可以看出,优化后相比优化前失稳临界轴压有一定下降,满足稳定性约束,优化后面板及芯子厚度变薄,内外壁温差增大,结构质量减轻了14.86%。本算例中,优化前后结果均满足内壁温度约束,如果作用在结构上的热流增大或加热时间变长,该约束可能会变为起作用约束,因此,需要对其进行限制。

保持其余参数不变,分别将区间可能度水平改变为0.2、0.5、1.0重新进行优化,将所得结果进行对比,结果如表4所示。

表4 不同可能度水平下不确定性优化结果

Table 4 Uncertainty optimization results under different possibility levels

可能度水平 $\lambda$	最优设计变量 $[t_s, t_c]$ /mm	优化后目标函数值 $M$ /kg
0.2	[1.520/15.759]	335.288
0.5	[1.387/18.314]	344.726
0.8	[1.600/16.500]	352.674
1.0	[1.765/15.765]	364.333

当可能度水平取为0.2时,结构质量相比优化前减重19.05%;当可能度水平取为1.0时,结构质量相比优化前减重12.04%。可能度水平由1.0减小至0.2,质量下降约7.97%,可以看出对问题而言,在需用载荷变化范围内结构重量变化不大,可以通过适当的质量代价追求更高的设计可靠性。

## 4 结论

1) 建立了火箭整流罩圆锥壳夹层结构参数化模型,建立一维热传导模型获取热流作用下结构温度场,推导了考虑热应力情况下圆锥壳的稳定性平衡方程,采用伽辽金法求解方程,给出了失稳临界轴压的表达式。

2) 建立了区间不确定性优化模型,采用区间可能度方法将不确定性优化模型转换为确定性优化模型,采用GA-CIAM方法代替原本的GA嵌套寻优方法,降低了计算量。

3) 对整流罩前锥段夹层结构圆锥壳开展了考虑气动力/热载荷及材料参数不确定性的结构优化设计,获得了满足设计要求且质量更轻的设计结果,研究了改变可能度水平对设计结果的影响,为不同可靠性要求的设计过程提供指导。

## 参考文献 (References)

- [1] 杨灿. 耐高温复合材料整流罩性能及热力学分析[D]. 哈尔滨: 哈尔滨工程大学, 2014.  
YANG C. Study on the high temperature-resistant composites fairing and its mechanical properties under high service temperature [D]. Harbin: Harbin Engineering University, 2014(in Chinese).
- [2] MEYERS C A, HYER M W, SHUART M J. Thermal buckling and postbuckling of symmetrically laminated composite plates[J]. Journal of Thermal Stresses, 1992, 14(4): 519-540.
- [3] NAJ R, BOROJERDY M S, ESLAMI M R. Thermal and mechanical instability of functionally graded truncated conical shells[J]. Thin-Walled Structures, 2008, 46(1): 65-78.
- [4] DUNG D V, CHAN D Q. Analytical investigation on mechanical buckling of FGM truncated conical shells reinforced by orthogonal stiffeners based on FSDT[J]. Composite Structures, 2017, 159(1): 827-841.
- [5] DUC N D, SEUNG-EOCK K, CHAN D Q. Thermal buckling analysis of FGM sandwich truncated conical shells reinforced by FGM stiffeners resting on elastic foundations using FSDT[J]. Journal of Thermal Stresses, 2017, 41(3): 1-35.
- [6] TORABI J, KIANI Y, ESLAMI M R. Linear thermal buckling analysis of truncated hybrid FGM conical shells[J]. Composites Part B Engineering, 2013, 50(7): 265-72.
- [7] 张东, 穆京京, 蒋亦轲, 等. 轴压下圆柱加筋壳的热屈曲[J]. 强度与环境, 2019, 46(6): 32-37.  
ZHANG D, MU J J, JIANG Y M, et al. Thermal buckling of stiffened cylinders under axial compression[J]. Structure & Environment Engineering, 2019, 46(6): 32-37(in Chinese).
- [8] 唐统帅. 基于散热和承载功能的格栅加筋板研究[D]. 哈尔滨: 哈尔滨工业大学, 2014.  
TANG T S. Research on grid-stiffened panels base on heat transfer and load-carrying capacity [D]. Harbin: Harbin Institute of Technology, 2014(in Chinese).
- [9] 徐腾飞, 辛健强, 董永朋, 等. 柔性支撑上FGM矩形薄板热屈曲的封闭形式解[J]. 强度与环境, 2018, 45(3): 20-29.  
XU T F, XIN J Q, DONG Y P, et al. Closed form solution for thermal buckling of rectangular FGM thin plates with flexible support[J]. Structure & Environment Engineering, 2018, 45(3): 20-29(in Chinese).
- [10] BELLMAN R E, ZADEH L A. Application series decision-making in a fuzzy environment[J]. Management Science, 1970, 17(4): 141-164.
- [11] FAES M. Interval methods for the identification and quantification of inhomogeneous uncertainty in finite element models [D]. Belgium: KU Leuven, 2017.
- [12] 尹莲花. 高速飞行器热防护结构分析与热结构优化方法研究[D].

北京: 北京理工大学, 2008.

YIN L H. Analysis of thermal protection structure and study of thermal structural optimization for high speed aircraft [D]. Beijing: Beijing Institute of Technology, 2008(in Chinese).

[13] BRUSH D O, ALMROTH B O, HUTCHINSON J W. Buckling of bars, plates, and shells[J]. Journal of Applied Mechanics, 1975, 42(4): 911.

[14] JIANG C, HAN X, LIU G R, et al. A nonlinear interval number programming method for uncertain optimization problems[J]. European Journal of Operational Research, 2008, 188(1): 1-13.

[15] 祁武超, 邱志平. 基于区间分析的结构非概率可靠性优化设计[J]. 中国科学:物理学 力学 天文学, 2013, 43(1): 85-93.

QI W C, QIU Z P. Non-probabilistic reliability-based structural design optimization based on the interval analysis method[J]. Science China Physics, Mechanics & Astronomy, 2013, 43(1): 85-93(in Chinese).

### 附录 A:

力热载荷联合作用下, 稳定平衡方程式 (18) 中系数  $L_{ij}$  的表达式如下:

$$L_{11} = -\frac{m^2\pi^3}{L^2} A_{11} \sin\beta \left[ \frac{(x_0+L)^4 - x_0^4}{8} + \frac{3L^3(2x_0+L)}{8m^2\pi^2} \right] - \frac{\pi}{4} L(2x_0+L) \sin\beta \left[ A_{22} + \frac{n^2}{\sin^2\beta} A_{66} \right] + \frac{\pi}{4} L(2x_0+L) A_{11} \sin\beta \quad (A1)$$

$$L_{12} = -\frac{mn\pi^2}{L} (A_{11} + A_{66}) \left[ \frac{(x_0+L)^3 - x_0^3}{6} + \frac{L^3}{4m^2\pi^2} \right] - \frac{nL^2}{4m} (A_{22} + A_{66}) \quad (A2)$$

$$L_{13} = \frac{m\pi^2}{L} \cot\beta \sin\beta A_{12} \left[ \frac{(x_0+L)^3 - x_0^3}{6} + \frac{L^3}{4m^2\pi^2} \right] + \cot\beta \sin\beta \frac{L^2}{4m} A_{22} \quad (A3)$$

$$L_{14} = -\frac{m^2\pi^3}{L^2} B_{11} \sin\beta \left[ \frac{(x_0+L)^4 - x_0^4}{8} + \frac{3L^3(2x_0+L)}{8m^2\pi^2} \right] - \frac{\pi}{4} L(2x_0+L) \sin\beta \left[ B_{22} + \frac{n^2}{\sin^2\beta} B_{66} \right] + \frac{\pi}{4} L(2x_0+L) B_{11} \sin\beta \quad (A4)$$

$$L_{15} = -\frac{mn\pi^2}{L} \frac{1}{\sin\beta} (2x_0+L)(B_{12} + B_{66}) \quad (A5)$$

$$L_{21} = -\frac{mn\pi^2}{L} (A_{12} + A_{66}) \left[ \frac{(x_0+L)^3 - x_0^3}{6} - \frac{L^3}{4m^2\pi^2} \right] - \frac{nL^2}{4m} (A_{22} + A_{66}) \quad (A6)$$

$$L_{22} = -\frac{m^2\pi^3}{L^2} A_{66} \sin\beta \left[ \frac{(x_0+L)^4 - x_0^4}{8} - \frac{3L^3(2x_0+L)}{8m^2\pi^2} \right] - \frac{\pi}{4} L(2x_0+L) A_{66} \sin\beta - \frac{\pi}{4} L(2x_0+L) \sin\beta \cdot \left[ \frac{n^2}{\sin^2\beta} A_{22} + A_{66} \right] \quad (A7)$$

$$L_{23} = \frac{n\pi}{4} L(2x_0+L) \cot\beta A_{22} \quad (A8)$$

$$L_{24} = -\frac{mn\pi^2}{L} (B_{12} + B_{66}) \left[ \frac{(x_0+L)^3 - x_0^3}{6} - \frac{L^3}{4m^2\pi^2} \right] - \frac{nL^2}{4m} (B_{22} + B_{66}) \quad (A9)$$

$$L_{25} = -\frac{m^2\pi^3}{L^2} B_{66} \left[ \frac{(x_0+L)^3 - x_0^3}{8} - \frac{L^3}{4m^2\pi^2} \right] - \frac{n^2\pi L}{2\sin^2\beta} B_{22} + \frac{\pi}{4} L B_{66} \quad (A10)$$

$$L_{31} = -\frac{m^3\pi^4}{L^3} B_{11} \sin\beta \left\{ \frac{(x_0+L)^5 - x_0^5}{10} + \frac{L^2 [x_0^3 - (x_0+L)^3]}{2m^2\pi^2} + \frac{3L^5}{4m^4\pi^4} \right\} + \frac{m\pi^2}{L} \cot\beta \sin\beta A_{12} \left[ \frac{(x_0+L)^4 - x_0^4}{8} - \frac{3L^3(2x_0+L)}{8m^2\pi^2} \right] + \cot\beta \sin\beta A_{22} \frac{L^2(2x_0+L)}{4m} + \left[ \frac{m\pi^2}{L} \sin\beta (3B_{11} + B_{12}) + \frac{mn^2\pi^2}{L \sin\beta} (B_{11} + 2B_{66}) \right] \cdot \left[ \frac{(x_0+L)^3 - x_0^3}{6} - \frac{L^3}{4m^2\pi^2} \right] + \frac{L^2}{4m} \sin\beta B_{22} \left( \frac{n^2}{\sin^2\beta} - 1 \right) \quad (A11)$$

$$L_{32} = -\frac{m^2n\pi^3}{L^2} (B_{12} + B_{66}) \left[ \frac{(x_0+L)^4 - x_0^4}{8} - \frac{3L^3(2x_0+L)}{8m^2\pi^2} \right] + n\pi \cot\beta A_{22} \left[ \frac{(x_0+L)^3 - x_0^3}{6} - \frac{L^3}{4m^2\pi^2} \right] + \frac{n\pi L(2x_0+L)}{4} B_{22} \left( \frac{n^2}{\sin^2\beta} - 2 \right) \quad (A12)$$

$$L_{33} = -\frac{m^2\pi^3}{L^2} \cot\beta \sin\beta B_{12} \left[ \frac{(x_0+L)^4 - x_0^4}{8} - \frac{3L^3(2x_0+L)}{8m^2\pi^2} \right] - \pi \cot^2\beta \sin\beta A_{22} \left[ \frac{(x_0+L)^3 - x_0^3}{6} - \frac{L^3}{4m^2\pi^2} \right] - \frac{\pi L(2x_0+L)}{4} \cot\beta \sin\beta B_{22} \left( \frac{n^2}{\sin^2\beta} - 2 \right) \quad (A13)$$

$$L_{34} = \frac{m^3\pi^4}{L^3} D_{11} \sin\beta \left\{ \frac{(x_0+L)^5 - x_0^5}{10} + \frac{L^2 [x_0^3 - (x_0+L)^3]}{2m^2\pi^2} + \frac{3L^5}{4m^4\pi^4} \right\} + \frac{m\pi^2}{L} \cot\beta \sin\beta B_{12} \left[ \frac{(x_0+L)^4 - x_0^4}{8} - \frac{3L^3(2x_0+L)}{8m^2\pi^2} \right] + \cot\beta \sin\beta B_{22} \frac{L^2(2x_0+L)}{4m} + \left[ \frac{m\pi^2}{L} \sin\beta(3D_{11} + D_{22}) + \frac{mn^2\pi^2}{L \sin\beta} (D_{12} + 2D_{66}) \right] \cdot \left[ \frac{(x_0+L)^3 - x_0^3}{6} - \frac{L^3}{4m^2\pi^2} \right] + \frac{L^2}{4m} \sin\beta D_{22} \left( \frac{n^2}{\sin^2\beta} - 1 \right) \quad (\text{A14})$$

$$L_{35} = \frac{m^3n\pi^3}{L^2} \frac{1}{\sin\beta} (D_{12} + 2D_{66}) \left[ \frac{(x_0+L)^3 - x_0^3}{6} - \frac{L^3}{4m^2\pi^2} \right] + \frac{n\pi}{4\sin\beta} L(2x_0+L) \cot\beta B_{22} - \frac{n\pi}{\sin\beta} L(D_{12} + 2D_{66}) + \frac{n\pi L}{2\sin\beta} D_{22} \left( \frac{n^2}{\sin^2\beta} - 2 \right) - \frac{n\pi}{4\sin\beta} L(D_{12} + 2D_{66} + 4D_{66}) \quad (\text{A15})$$

$$L_{36} = -\frac{m^2\pi^3}{L^2} \sin\beta \left[ \frac{(x_0+L)^4 - x_0^4}{8} - \frac{3L^3(2x_0+L)}{8m^2\pi^2} \right] \quad (\text{A16})$$

$$L_{37} = -\frac{n^2\pi}{x_0 \sin\beta} \left[ \frac{(x_0+L)^4 - x_0^4}{8} - \frac{3L^3(2x_0+L)}{8m^2\pi^2} \right] \quad (\text{A17})$$

$$L_{41} = -\frac{m^2\pi^3}{L^2} B_{11} \sin^2\beta \left[ \frac{(x_0+L)^4 - x_0^4}{8} + \frac{3L^3(2x_0+L)}{8m^2\pi^2} \right] + \frac{\pi}{4} L(2x_0+L) [\sin^2\beta(B_{11} - B_{22}) - n^2 B_{66}] \quad (\text{A18})$$

$$L_{42} = -\frac{mn\pi^2}{L} \sin\beta(B_{12} + B_{66}) \left[ \frac{(x_0+L)^3 - x_0^3}{6} + \frac{L^3}{4m^2\pi^2} \right] - \frac{nL^2}{4m} \sin\beta(B_{22} + B_{66}) \quad (\text{A19})$$

$$L_{43} = -\frac{m\pi^2}{L} A_{44} \sin^2\beta \left[ \frac{(x_0+L)^4 - x_0^4}{8} + \frac{3L^3(2x_0+L)}{8m^2\pi^2} \right] + \frac{m\pi^2}{L} \sin\beta \cos\beta B_{12} \left[ \frac{(x_0+L)^3 - x_0^3}{6} + \frac{L^3}{4m^2\pi^2} \right] + \frac{L^2}{4m} \sin\beta \cos\beta B_{22} \quad (\text{A20})$$

$$L_{44} = -\left( \frac{m^2\pi^3}{L^2} D_{11} \sin^2\beta + \pi A_{44} \sin^2\beta \right) \cdot \left[ \frac{(x_0+L)^4 - x_0^4}{8} + \frac{3L^3(2x_0+L)}{8m^2\pi^2} \right] + \frac{\pi L}{4} (2x_0+L) [\sin^2\beta(D_{11} - D_{22}) - n^2 D_{66}] \quad (\text{A21})$$

$$L_{45} = -\frac{mn\pi^2}{4} (2x_0+L)(D_{12} + D_{66}) \quad (\text{A22})$$

$$L_{51} = -\frac{mn\pi^2}{4} (B_{12} + B_{66}) \left[ \frac{(x_0+L)^3 - x_0^3}{6} + \frac{L^3}{4m^2\pi^2} \right] - \frac{nL^2}{4m} (B_{12} + B_{66}) \quad (\text{A23})$$

$$L_{52} = -\frac{m^2\pi^3}{L^2} B_{66} \sin\beta \left[ \frac{(x_0+L)^4 - x_0^4}{8} + \frac{3L^3(2x_0+L)}{8m^2\pi^2} \right] - \frac{\pi}{2} L(2x_0+L) \left( \frac{n^2 B_{22}}{2\sin\beta} + B_{66} \sin\beta \right) \quad (\text{A24})$$

$$L_{53} = -n\pi A_{55} \frac{m^2\pi^3}{L^2} \left[ \frac{(x_0+L)^3 - x_0^3}{6} - \frac{L^3}{4m^2\pi^2} \right] + \frac{n\pi L}{4} (2x_0+L) \cot\beta B_{22} \quad (\text{A25})$$

$$L_{54} = -\frac{mn\pi^2}{L} (D_{12} + D_{66}) \left[ \frac{(x_0+L)^3 - x_0^3}{6} - \frac{L^3}{4m^2\pi^2} \right] - \frac{nL^2}{4m} (D_{22} + D_{66}) \quad (\text{A26})$$

$$L_{55} = -\pi \left( A_{55} + \frac{m^2\pi^2}{L^2} D_{66} \right) \left[ \frac{(x_0+L)^3 - x_0^3}{6} - \frac{L^3}{4m^2\pi^2} \right] + \frac{\pi L}{4} \left( D_{66} - \frac{2n^2}{\sin^2\beta} D_{22} \right) \quad (\text{A27})$$

# Uncertainty lightweight design of sandwich structure of rocket fairing cone

DONG Xinxin, YUE Zhenjiang\*, WANG Zhi, LIU Li

(School of Aerospace Engineering, Beijing Institute of Technology, Beijing 100081, China)

**Abstract:** In order to analyze the influence of uncertainty on the thermal stability of the launch vehicle fairing cone shell sandwich structure and to guide the lightweight design of the structure, a model of the fairing front cone section sandwich shell is established and a temperature field model is built, based on which the thermal stability analysis of the cone shell is carried out and the critical axial pressure under the combined force-thermal load is derived. For the primary uncertainty factors, interval uncertainty optimization models and sensitivity analyses are also produced. The interval probability approach is then used to convert these models into deterministic optimization problems, which are then resolved using the genetic algorithm-collocation interval analysis method (GA-CIAM) method. The calculation results show that considering the influence of aerodynamic/thermal load and material parameter uncertainty, the optimization design of the front cone section of the fairing can effectively realize the structural lightweight on the premise of meeting the design requirements.

**Keywords:** cone shell of fairing; sandwich structure; thermal stability analysis; interval uncertainty optimization; lightweight design

<http://bhxb.buaa.edu.cn>     [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0268

# 基于 Logistic 回归麻雀算法的图像分割

陈刚<sup>1</sup>, 林东<sup>1,\*</sup>, 陈飞<sup>2</sup>, 陈祥宇<sup>1</sup>

(1. 福州大学 物理与信息工程学院, 福州 350108; 2. 福州大学 计算机与大数据学院/软件学院, 福州 350108)

**摘 要:** 针对麻雀搜索算法后期种群多样性减少、易陷入局部最优解等问题, 提出一种新的改进麻雀搜索算法。所提算法先引入小孔成像反向学习策略对发现者的位置进行更新, 提升寻优位置的多样性; 其次受 Logistic 模型的启发, 提出一种新的自适应因子对安全阈值进行动态控制, 平衡所提算法的全局搜索与局部开发的能力。通过与其他算法在 6 个基准函数上进行仿真对比, 结果表明: 所提算法的收敛精度与速度均优于其他算法。在工程应用上, 用所提算法优化 K-means 聚类算法进行图像分割, 峰值信噪比 (PSNR)、结构相似性 (SSIM) 及特征相似性 (FSIM) 3 种度量指标验证了其良好的分割性能。

**关键词:** 麻雀搜索算法; 图像分割; 小孔成像反向学习; Logistic 模型; K-means 聚类算法  
**中图分类号:** TP391.41

**文献标志码:** A     **文章编号:** 1001-5965(2023)03-0636-11

近几年, 智能优化算法广泛应用于机器学习、组合优化、图像处理等领域, 由于其良好的寻优能力, 很多学者将其应用到图像处理领域, 并取得不错的成效。Capor 等<sup>[1]</sup>以萤火虫算法为基础, 对 K-means 聚类算法进行改进, 能够在医学图像上实现更好的分割平均误差、峰值信噪比和结构相似性指标; Li 等<sup>[2]</sup>提出了一种基于动态粒子群优化 K-means 聚类算法, 在图像分割中比传统 K-means 聚类算法具有更好的视觉效果, 在提高图像分割质量和效率方面也具有明显的优势; Kapoor 等<sup>[3]</sup>提出灰狼优化 (grey wolf optimizer, GWO) 算法在卫星图像分割中的应用, 对 GWO 进行适当的修改, 以用作自动聚类算法; Khriissi 等<sup>[4]</sup>利用正余弦算法 (sine cosine algorithm, SCA) 解决了传统聚类算法存在的聚类中心初始化和局部最优收敛等缺点。文献 [1-4] 证实了一些智能优化算法通过优化 K-means 聚类算法的初始聚类中心, 可以取得不错的图像分割效果。

麻雀搜索算法 (sparrow search algorithm, SSA) 是 2020 年提出的一种新型智能优化算法, 与 GWO 算法、粒子群优化 (particle swarm optimization, PSO)

算法、引力搜索算法 (gravitational search algorithm, GSA) 等传统智能优化算法相比, 该算法在搜索精度、收敛速度和稳定性等方面具有优越性<sup>[5]</sup>, 但也像大多数智能优化算法一样存在收敛早熟的现象。因此, 学者们陆续提出一系列改进的算法来避免这个现象的发生, 并提升 SSA 的性能。吕鑫等<sup>[6-7]</sup>结合鸟群算法中飞行行为的思想来优化 SSA, 此外, 还通过引入高斯变异和 Tent 混沌扰动形成一种混沌麻雀搜索优化算法; Liu 等<sup>[8]</sup>引入混沌策略对算法的总体进行优化, 并采用自适应惯性权重来平衡算法的收敛性、算法速率和探测能力; Yuan 等<sup>[9]</sup>利用重心反向学习机制初始化种群, 并在发现者位置更新部分引入学习系数, 以提高算法的全局搜索能力; Zhang 和 Ding<sup>[10]</sup>主要利用 Logistic 映射、自适应超参数、变异算子来增强 SSA 的全局寻优能力; Zhu 和 Yousefi<sup>[11]</sup>提出了一种新的自适应 SSA, 并通过采用 DE/best/1 变异策略和动态缩放因子来避免算法陷入局部最优; Liu 等<sup>[12]</sup>提出平衡 SSA, 在 Lévy 飞行机制中, 使用随机游走算法对局部搜索进行适当的调整; Wang 等<sup>[13]</sup>构造了一种基于伯努利

收稿日期: 2021-05-24; 录用日期: 2021-08-27; 网络出版时间: 2021-09-13 08:48

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210910.1545.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210910.1545.001.html)

\* 通信作者. E-mail: [lindong@fzu.edu.cn](mailto:lindong@fzu.edu.cn)

**引用格式:** 陈刚, 林东, 陈飞, 等. 基于 Logistic 回归麻雀算法的图像分割 [J]. 北京航空航天大学学报, 2023, 49 (3): 636-646.

CHEN G, LIN D, CHEN F, et al. Image segmentation based on Logistic regression sparrow algorithm [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 636-646 (in Chinese).

混沌映射、动态自适应加权、Cauchy 变异和反向学习的混沌麻雀搜索算法,并将该算法用于优化微网簇系统的运行成本;Ouyang 等<sup>[14]</sup>在发现者阶段引入基于透镜原理的反向学习策略,提高麻雀个体的搜索范围,然后提出可变螺旋搜索策略,使跟随者的搜索更加仔细和灵活。还有一些学者将改进 SSA 应用于储能系统的成本最小化<sup>[15]</sup>、故障诊断<sup>[16]</sup>、机器人路径规划<sup>[17]</sup>等问题上,并取得不错的成效。

文献 [6-17] 是学者们关于 SSA 的一些改进麻雀搜索算法,虽然在一定程度下能减少算法收敛早熟现象的产生,但依然存在以下 2 个问题:①大多数文献直接替换了原 SSA 中发现者的位置更新机制,缺乏一定的学习能力与择优能力;②仅注重全局搜索能力的提升,忽略了局部开发能力与全局搜索能力之间的平衡性。为此,本文提出一种新的改进麻雀搜索算法(modified sparrow search algorithm, MSSA)。一方面,文献 [14] 采用基于透镜原理的反向学习策略优化发现者的位置,受此思路,选择同样由相似三角形原理得到的基于小孔成像的反向学习策略进行优化。本文与文献 [14] 最大的不同是, MSSA 保留了 SSA 中发现者原有的位置更新机制并进行比较,选择更优的位置参与种群迭代,因此,在发现者位置更新阶段利用小孔成像反向学习的策略,丰富了麻雀种群寻优位置的多样性,提高了算法全局搜索能力;另一方面,本文与文献 [10] 中利用 Logistic 映射的性质初始化种群不同的是, Logistic 回归是围绕一个 Logistic 曲线展开的, Logistic 曲线是一条严格单调递增的 S 型曲线,可进行伸缩平移,因此,受 Logistic 回归的启发,本文提出一种新的自适应因子对安全阈值进行动态控制,进一步平衡所提算法的全局搜索能力与局部开发能力。相比于 4 种经典算法和文献 [6] 中的改进麻雀搜索算法(improved sparrow search algorithm, ISSA), MSSA 在 6 个标准测试函数上表现出良好的寻优性能。在实际应用中,利用 MSSA 优化 K-means 聚类算法的初始聚类中心进行图像分割实验,并与传统 K-means 聚类算法、PSO、SSA 及 ISSA 进行对比。此外,选取峰值信噪比(peak signal to noise ratio, PSNR)、结构相似性(structural similarity, SSIM)和特征相似性(feature similarity, FSIM)这 3 种图像分割的度量指标来评估算法的分割性能,实验结果表明 MSSA 的分割性能更好,具有较强的鲁棒性。

本文介绍了基于小孔成像的反向学习策略来丰富发现者的寻优位置,一种基于 Logistic 回归的

自适应因子来平衡算法的全局搜索与局部开发能力,提出一种 MSSA 优化 K-means 聚类算法进行图像分割的算法,并用 PSNR、SSIM 和 FSIM 这 3 种度量指标验证其良好的分割性能。

## 1 改进麻雀搜索算法

### 1.1 麻雀搜索算法

SSA 主要是根据麻雀的觅食行为与反捕食行为而提出的新型群智能优化算法。麻雀的觅食行为对应着 SSA 中发现者与跟随者,即每次迭代中选取位置较好的几只麻雀作为发现者,负责全局搜索食物并为所有跟随者提供觅食的区域和方向,剩下的麻雀作为跟随者,跟随发现者争夺食物;而麻雀的反捕食行为对应着 SSA 中侦察预警机制,即部分种群的麻雀进行侦察预警,如果发现危险则放弃食物飞向新的位置。在  $d$  维解空间中,麻雀个体的位置代表搜索空间的一组有效解,麻雀个体的能量储备代表着适应度值。

SSA 中发现者的位置更新为

$$\mathbf{x}_i^{t+1} = \begin{cases} \mathbf{x}_i^t \exp\left(\frac{-i}{\alpha \text{iter}_{\max}}\right) & R_2 < \text{ST} \\ \mathbf{x}_i^t + QL & R_2 \geq \text{ST} \end{cases} \quad (1)$$

式中:  $\mathbf{x}_i^t$  为种群中第  $t$  代第  $i$  个个体位置;  $\text{iter}_{\max}$  为最大迭代次数;  $\alpha$  为  $(0,1]$  中的均匀随机数;  $Q$  为服从标准正态分布的随机数;  $L$  为每个元素为 1 的  $1 \times d$  维的矩阵; ST 为安全阈值,取值范围为  $[0.5,1]$ ;  $R_2$  为警报信号,取值范围为  $[0,1]$ 。

跟随者的位置更新为

$$\mathbf{x}_i^{t+1} = \begin{cases} Q \exp\left(\frac{\mathbf{x}_{\text{worst}}^t - \mathbf{x}_i^t}{i^2}\right) & i > \frac{n}{2} \\ \mathbf{x}_p^{t+1} + |\mathbf{x}_i^t - \mathbf{x}_p^{t+1}| A^+ L & i \leq \frac{n}{2} \end{cases} \quad (2)$$

式中:  $\mathbf{x}_{\text{worst}}^t$  为当前种群最差的位置;  $\mathbf{x}_p^{t+1}$  为当前种群最优的位置;  $A^+ = A^T(AA^T)^{-1}$ , 其中  $A$  为  $1 \times d$  维的矩阵,矩阵中每个元素随机赋值 1 或 -1;  $n$  为麻雀种群大小。

侦察者的位置更新为

$$\mathbf{x}_i^{t+1} = \begin{cases} \mathbf{x}_{\text{best}}^t + \beta |\mathbf{x}_i^t - \mathbf{x}_{\text{best}}^t| & f_i > f_g \\ \mathbf{x}_i^t + K \left( \frac{|\mathbf{x}_i^t - \mathbf{x}_{\text{worst}}^t|}{f_i - f_w + \varepsilon} \right) & f_i = f_g \end{cases} \quad (3)$$

式中:  $\mathbf{x}_{\text{best}}^t$  为当前全局最优位置;  $\beta$  为服从标准正态分布的随机数;  $K$  为  $[-1,1]$  之间的均匀随机数;  $f_i$ 、 $f_g$  和  $f_w$  分别为当前种群的适应度值、全局最优适应度值和全局最差适应度值;  $\varepsilon$  为最小的常量,以避免零除误差。

## 1.2 基于 Logistic 回归麻雀算法

### 1.2.1 逐维小孔成像反向学习优化发现者位置

在式 (1) 发现者更新位置阶段, 当  $R_2 < ST$  时麻雀个体的位置随着迭代次数的增加而整体呈递减的趋势, 意味着迭代后期种群的多样性逐渐减少, 增大了算法陷入局部最优解的概率。

受小孔成像的启发, 将反向原理简化到学习策略中, 可以帮助算法寻得更广的反向位置<sup>[18]</sup>。假设每一维解空间的区间为  $[a_j, b_j]$ , 高度为  $h$  的火焰在  $x$  轴上的投影为麻雀种群中发现者的位置  $x_{i,j}^t$ , 然后火焰通过小孔屏在接收屏上得到一个高度为  $h'$  的倒像, 同时在  $x$  轴上的投影为  $x_{i,j}'$  ( $x_{i,j}^t$  的反向点), 其中  $j$  为维数,  $j=1, 2, \dots, d$ , 如图 1 所示。所以由三角形相似原理可得

$$\frac{(a_j + b_j)/2 - x_{i,j}^t}{x_{i,j}^t - (a_j + b_j)/2} = \frac{h}{h'} \quad (4)$$

令  $\frac{h}{h'} = k$ , 式 (4) 可变换为

$$x_{i,j}' = \frac{(a_j + b_j)}{2} + \frac{(a_j + b_j)}{2k} - \frac{x_{i,j}^t}{k} \quad (5)$$

当  $k=1$  时, 可得

$$x_{i,j}' = (a_j + b_j) - x_{i,j}^t \quad (6)$$

由式 (6) 可以看出, 当  $k=1$  时的小孔成像反向学习就是常见的反向学习策略, 而常见的反向学习策略得到的候选解一般是固定的, 但小孔成像反向学习可通过改变接收屏与小孔屏之间的距离来调整比例系数  $k$ , 从而得到更广的反向位置(解)。

利用逐维小孔成像反向学习策略对发现者位置进行更新, 将 SSA 中  $R_2 < ST$  时发现者的位置通过式 (5) 映射到空间中得到反向位置(解), 然后将发现者位置的适应度值与反向学习后的适应度值进行比较筛选, 若反向学习后的适应度值优于原适应度值, 则反向位置(解)取代原发现者的位置; 反之, 则保留原发现者位置至下一代。因此, 引入小孔成像反向学习策略不仅丰富了发现者的寻优位置, 使麻雀个体每一维可大可小, 也在一定

程度上增加种群的多样性, 减少收敛早熟现象的产生。

### 1.2.2 基于 Logistic 模型的自适应因子

在麻雀自身的觅食过程中, 一旦发现捕食者, 它们通过唧唧喳喳的叫声作为警报信号。当  $R_2$ (警报信号)  $< ST$ (安全阈值), 发现者在觅食区域内未发现捕食者的存在, 可继续向其他区域开展大范围搜索, 代表算法的全局搜索能力; 当  $R_2$ (警报信号)  $\geq ST$ (安全阈值), 发现者意识到危险的存在, 向其他麻雀发出逃离信号, 所有麻雀迅速飞往其他安全区域觅食, 代表算法的局部开发能力。而安全阈值在 SSA 中往往是固定参数, 虽然能发挥算法的全局搜索能力, 但忽略了算法的局部开发能力, 可能会导致算法收敛早熟现象的发生。

为了平衡算法的全局搜索能力与局部开发能力, 提出一种基于 Logistic 模型的自适应因子。Sigmoid 函数是生物学中一种常见的 Logistic 曲线, 也常应用于神经网络中, 其可将变量映射到区间  $[0, 1]$  之间, 表达式为

$$S(x) = \frac{1}{1 + e^{-x}} \quad (7)$$

显而易见, Sigmoid 函数是一个严格单调递增、连续光滑的阈值函数, 对式 (7) 进行伸缩平移, 并引入幅值、伸缩与平移因子, 可得

$$S(x) = Ag \left( \frac{1}{1 + e^{ax+b}} + \omega \right) \quad (8)$$

式中:  $Ag$  为幅值增益;  $a$  和  $b$  分别为伸缩与平移因子;  $\omega$  为一个很小的正数, 可动态设定函数的上下界。图 2 为不同参数下对应的函数曲线, 其中  $\omega=0.05$ , 可根据实际应用设定。在算法迭代前期的发现者阶段, 可保证安全阈值长时间处于较大值的范围内, 发现者对安全区域进行大范围的搜索, 指引种群往最优位置附近收敛; 随着迭代次数的增加, 安全阈值开始非线性减小; 当到迭代后期时, 安全阈值往往处于较小值的范围内, 发现者在最优位置的附近进行局部开发, 逼近理论最优解, 提高算法的收敛精度。

### 1.2.3 算法伪代码

**输入:** 最大迭代次数  $G$ , 种群大小  $P$ , 发现者的数量  $PD$ , 侦察者的数量  $SD$ , 安全阈值  $ST$ , 警报值  $R_2$ 。初始化麻雀种群并定义其相关参数。

**输出:**  $x_{best}, f_g$ 。

1: While( $t < G$ )

2: 对麻雀种群的适应度值进行排序, 找出当前最优与最差的麻雀个体。

3:  $R_2 = \text{rand}(1)$ ;

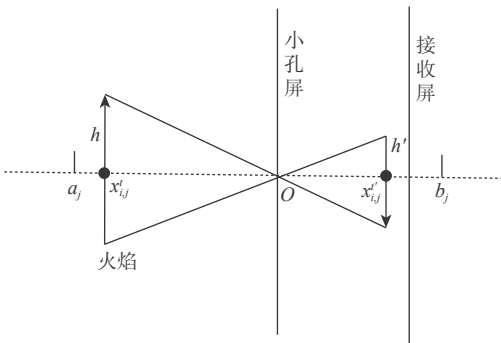


图 1 小孔成像反向学习原理

Fig. 1 Principle of reverse learning for small hole imaging

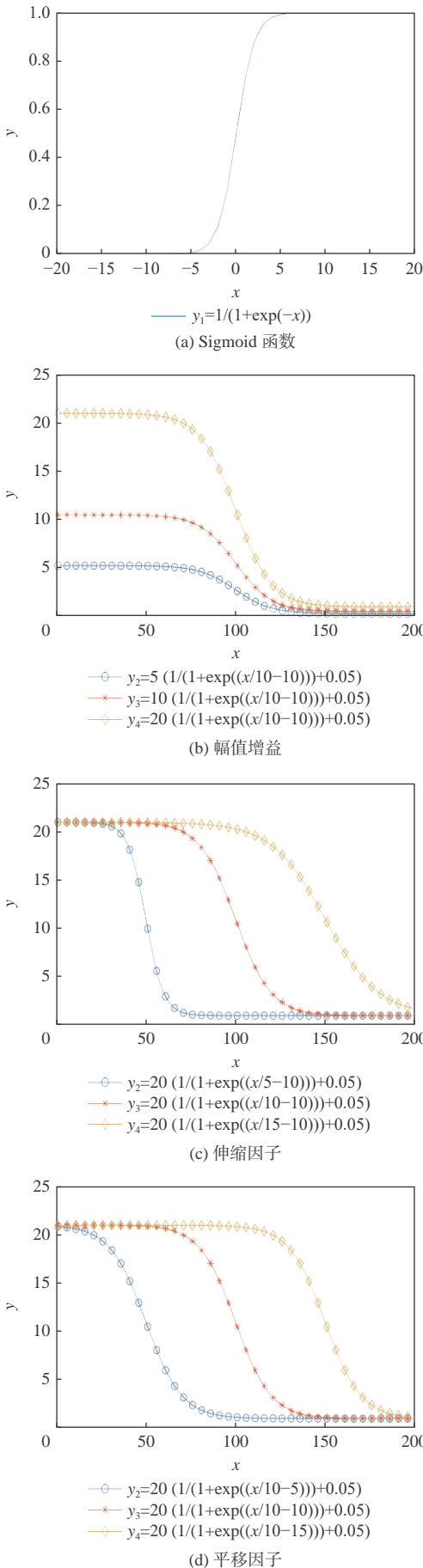


图 2 不同参数下对应的函数曲线

Fig. 2 Curves of functions corresponding to different parameters

- 4: If  $R_2 < ST \left( \frac{1}{1 + e^{ax+b}} + \omega \right)$
- 5: For  $i=1: PD$
- 6: 根据式 (1) 更新发现者的位置;
- 7: 根据式 (5) 得到反向位置;
- 8: 若式 (5) 的适应度值优于式 (1) 的适应度值, 则取代该位置; 反之则保留至下一代。
- 9: End for
- 10: Else
- 11: For  $i=1: PD$
- 12: 根据式 (1) 更新发现者的位置;
- 13: End for
- 14: End if
- 15: For  $i=(PD+1): P$
- 16: 根据式 (2) 更新跟随者的位置;
- 17: End for
- 18: For  $l=1: SD$
- 19: 根据式 (3) 更新麻雀个体位置;
- 20: End for
- 21: 得到新的最优个体的位置。
- 22: 若新个体的位置优于之前的个体位置, 则更新最优位置。
- 23:  $t=t+1$
- 24: End while
- 25: Return  $x_{best}, f_g$

## 2 基于MSSA 的K-means 图像分割算法

传统 K-means 算法通常是随机选择  $K$  个聚类中心, 但若选取的聚类中心不恰当, 可能会得到次优解或者陷入局部最优解, 因此, 需要选择合适的初始聚类中心。文献 [19-22] 通过相关算法优化 K-means 算法以解决初始聚类中心敏感等问题, 借鉴此思路, 利用 MSSA 对 K-means 算法的初始聚类中心进行优化, 其关键在于结合 K-means 算法建立 MSSA 的目标函数, 如式 (9), 即可理解为单目标优化问题, 利用 MSSA 求得最佳初始聚类中心点使得目标函数适应度值最小。

$$f = \sum_{n=1}^p \sum_{m=1}^q |X_n - Y_m|^2 \quad (9)$$

式中:  $X_n$  为图像中第  $n$  个像素灰度值;  $Y_m$  为第  $m$  个聚类中心;  $p$  为像素的数量;  $q$  为聚类中心的个数。

基于 MSSA 的 K-means 图像分割算法主要包含 2 部分: ①利用 MSSA 的全局搜索能力在图像点集里找到最佳初始聚类中心; ②将 MSSA 输出的最佳初始聚类中心用在 K-means 算法上进行图像分割, 有效地避免易陷入局部最优解的问题, 并且得到更好的分割效果。基于 MSSA 的 K-means 图像分割算法的流程如图 3 所示。

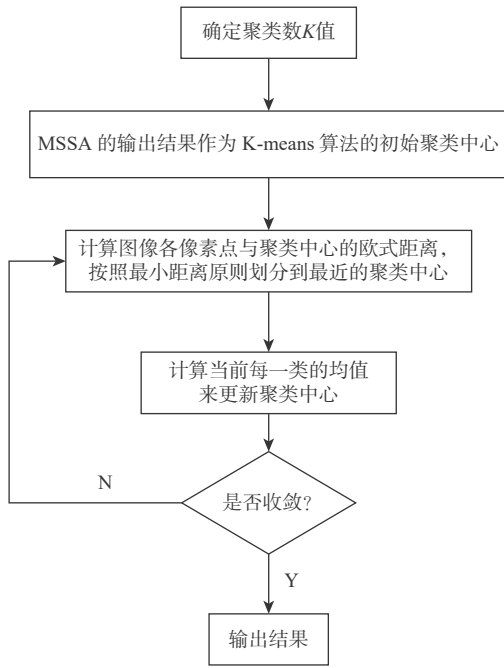


图3 基于MSSA的K-means图像分割

Fig. 3 K-means image segmentation based on MSSA

### 3 实验结果与分析

#### 3.1 仿真实验环境与测试函数

为了验证MSSA在求解实际问题中的性能,选取6个标准测试函数进行仿真实验,如表1所示。同时,将MSSA与PSO、SCA、人工蜂群<sup>[23]</sup>(artificial bee colony, ABC)算法、SSA及ISSA进行性能对比。

所有的仿真实验均在CPU为AMD锐龙5 4600 H、16 GB内存、Windows10 64位的操作系统的计算机下实现,并采用MATLAB2018A软件进行编译运行。算法的通用参数设置为:种群大小 $P=100$ ,最大迭代次数 $G=200$ ,其他参数设置如表2所示,并记录各个算法运行30次的平均值与标准差,算法的运行结果如表3所示。

从表3中的数据可以分析出,MSSA与传统算法PSO、SCA、ABC相比,MSSA有更强的全局寻优能力和更好的稳定性。另一方面,与传统算法SSA,以及其他学者改进的算法相比,在 $F_1 \sim F_3$ 高维

表1 测试函数

Table 1 Test functions

编号	函数名	函数表达式	取值区间	最小值
$F_1$	Zakharov	$f(x) = \sum_{i=1}^d x_i^2 + \left( \sum_{i=1}^d 0.5ix_i \right)^2 + \left( \sum_{i=1}^d 0.5ix_i \right)^4$	$[-5,10]$	0
$F_2$	Schwefel2.22	$f(x) = \sum_{i=1}^d  x_i  + \prod_{i=1}^d  x_i $	$[-10,10]$	0
$F_3$	Rotated Hyper-Ellipsoid	$f(x) = \sum_{i=1}^d \left( \sum_{j=1}^i x_j \right)^2$	$[-100,100]$	0
$F_4$	Rastrigin	$f(x) = \sum_{i=1}^d [x_i^2 - 10 \cos(2\pi x_i) + 10]$	$[-5.12,5.12]$	0
$F_5$	Schaffer N.2	$f(x_1, x_2) = 0.5 + \frac{\sin^2(x_1^2 - x_2^2) - 0.5}{[1 + 0.001(x_1^2 + x_2^2)]^2}$	$[-100,100]$	0
$F_6$	Drop-Wave	$f(x_1, x_2) = -\frac{1 + \cos(12\sqrt{x_1^2 + x_2^2})}{0.5(x_1^2 + x_2^2) + 2}$	$[-5.12,5.12]$	-1

表2 算法参数设置

Table 2 Parameter settings for algorithms

算法	参数设置
PSO	$w_{\max} = 0.9, w_{\min} = 0.2, c_1 = 2, c_2 = 2$
SCA	$a = 2, r_2 \in [0, 2\pi], r_3 \in [-2, 2], r_4 \in [0, 1]$
ABC	$L = \text{round}(0.6dP), a=1$
SSA	$PD = 35, ST = 0.6, SD = 70$
ISSA	$PD = 35, ST = 0.6, SD = 70$
MSSA	$PD = 35, ST = 0.6 \left( \frac{1}{1 + e^{t/10-20}} \right), SD = 70$

表 3 算法性能对比  
Table 3 Performance comparison of algorithms

函数	维度	PSO		SCA		ABC		SSA		ISSA		MSSA	
		均值	标准差	均值	标准差	均值	标准差	均值	标准差	均值	标准差	均值	标准差
$F_1$	30	$2.52 \times 10^2$	$9.35 \times 10$	$4.19 \times 10$	$1.40 \times 10$	$1.34 \times 10^3$	$4.47 \times 10^2$	$4.23 \times 10^{-40}$	$2.32 \times 10^{-39}$	$5.36 \times 10^4$	$1.80 \times 10^3$	$5.07 \times 10^{-46}$	$2.73 \times 10^{-45}$
$F_2$	30	$1.29 \times 10$	$2.23 \times 10$	$6.22 \times 10^{-1}$	$5.48 \times 10^{-1}$	$8.58 \times 10$	$1.58 \times 10$	$1.83 \times 10^{-35}$	$9.69 \times 10^{-35}$	$1.32 \times 10^{-5}$	$1.21 \times 10^{-5}$	$2.71 \times 10^{-38}$	$1.47 \times 10^{-37}$
$F_3$	30	$6.11 \times 10^2$	$2.14 \times 10^2$	$1.18 \times 10^4$	$5.52 \times 10^3$	$6.32 \times 10^4$	$1.03 \times 10^4$	$6.68 \times 10^{-46}$	$3.66 \times 10^{-45}$	$2.97 \times 10^{-4}$	$7.07 \times 10^{-4}$	$9.46 \times 10^{-65}$	$4.81 \times 10^{-64}$
$F_4$	30	$2.17 \times 10^2$	$2.42 \times 10$	$8.13 \times 10$	$3.61 \times 10$	$2.60 \times 10^2$	$1.27 \times 10$	0	0	$1.02 \times 10^{-7}$	$2.15 \times 10^{-7}$	0	0
$F_5$	2	$3.36 \times 10^{-1}$	$1.72 \times 10^{-1}$	0	0	$3.61 \times 10^{-8}$	$7.15 \times 10^{-8}$	0	0	0	0	0	0
$F_6$	2	-0.88	$1.76 \times 10^{-1}$	-1	0	-0.99	$1.57 \times 10^{-5}$	-1	0	-1	0	-1	0

单峰问题中, MSSA 的收敛精度比 SSA、ISSA 更高, 算法稳定性更好; 在函数  $F_4 \sim F_6$  多峰问题中 SSA、MSSA 均能找到理论最优值, 方差也为 0, 而其他学者改进的算法 ISSA 在  $F_4$  上未能找到理论最优值。

为了更加直观地反映算法的收敛特性, 图 4 为 6 种算法在  $F_1 \sim F_6$  测试函数上的收敛曲线。从图 4(a)~图 4(c) 中可以更加明显地看出 MSSA 的收敛精度均优于其他算法; 在图 4(d) 中, MSSA 在与 SSA 都能找到理论最优值的同等情况下, 收敛速度最快; 在图 4(e) 和图 4(f) 中, 除 PSO 和 ABC 外, 其他算法均接近甚至达到理论最优值, MSSA 的收敛速度均优于其他算法。从图 4 的实验结果可以得出, MSSA 的寻优能力较强, 性能稳定, 能够有效地跳出局部最优解。

### 3.2 时间复杂度分析

从 3.1 节的实验结果可以看出, 使用不同的算法得到的结果可能一样, 但未考虑算法的运行效率问题, 并不能说明两者性能相同。一个算法的运行效率可以用算法的时间复杂度与空间复杂度来衡量, 而在大多数情况下是用算法的时间复杂度来比较。为了充分验证 MSSA 的突出性能, 计算它与 SSA、ISSA 的时间复杂度, 如表 4 所示, 其中,  $O(\cdot)$  为时间复杂度。根据 1.1 节 SSA 的原理介绍可知, 算法可分为 5 个阶段组成: 初始化种群阶段、发现者位置更新阶段、跟随者位置更新阶段、侦察预警机制阶段、全局最优位置更新阶段。参数设置与 3.1 节保持一致, 其中  $d$  为维度, 求解适应度函数的时间为  $f(d)$ 。

由表 4 的对比结果可知, 改进算法 ISSA、MSSA 与 SSA 在 5 个阶段的时间复杂度相同, 即意味着提出的改进算法 MSSA 并不是以牺牲运行效率为代价来换取性能的提升。

### 3.3 基于 MSSA 的 K-means 图像分割实验

为了验证基于 MSSA 的 K-means 图像分割算

法性能的显著性, 选择 8 组测试图像进行实验, 并选取 PSO、SSA、ISSA 及 MSSA 这 4 种算法优化 K-means 聚类算法和传统 K-means 聚类算法进行图像分割效果的对比。由于 K-means 聚类算法对  $K$  值的依赖性比较强,  $K$  值选择不当会对结果产生很大的影响, 因此  $K$  值设为 5, 避免无关因素的干扰。算法通用参数: 种群大小  $P=30$ , 最大迭代次数  $G=200$ , 各算法的其他参数与 3.1 节保持一致, 其中  $PD=9$ ,  $SD=21$ 。

从主观视觉效果上看, K-means 的分割效果过于粗糙, PSO、SSA 与 ISSA 的分割效果比 K-means 略好, 但一些重要细节没有分割出来, 而 MSSA 的分割效果显然要比前 4 种算法好。例如, 如图 5 所示: 第 1 行(实验 1)和第 7 行(实验 7)分割效果中 MSSA 能够将部分细微的特征信息分割出来, 虽然分割效果与 ISSA 一样, 但分割性能仍然是最优的; 第 2 行(实验 2)分割效果图中 MSSA 分割完整性较好, 颜色显著; 第 3 行(实验 3)分割效果图中传统 K-means 聚类算法能分割出图像的大致轮廓, 但无法分割出原始图像中的细节信息, 而 MSSA 表现出更加清晰的纹理信息和显著特征; 第 4~6 行(实验 4~6)分割效果图中 MSSA 比其他算法保留了更多的细节信息, 轮廓更加清晰, 表现出算法更强的鲁棒性与适应性; 第 8 行(实验 8)分割效果图中 MSSA 很大程度上接近于原图。综合分析可知, MSSA 的分割效果比传统 K-means 聚类算法、PSO、SSA 及 ISSA 好, 图像细节信息保留得更加完整, 鲁棒性更强。

为了能够进一步定量评估各优化算法的分割性能, 选取了 3 种常见的图像分割的度量指标 PSNR、SSIM 及 FSIM 对本节 5 种算法的分割结果进行分析, 如表 5~表 12 所示。

PSNR 用于测量分割后的图像与原始图像之间的差值<sup>[24]</sup>, PSNR 定义可由式 (10) 和式 (11) 表示, PSNR 值越大, 代表图像分割质量越好。

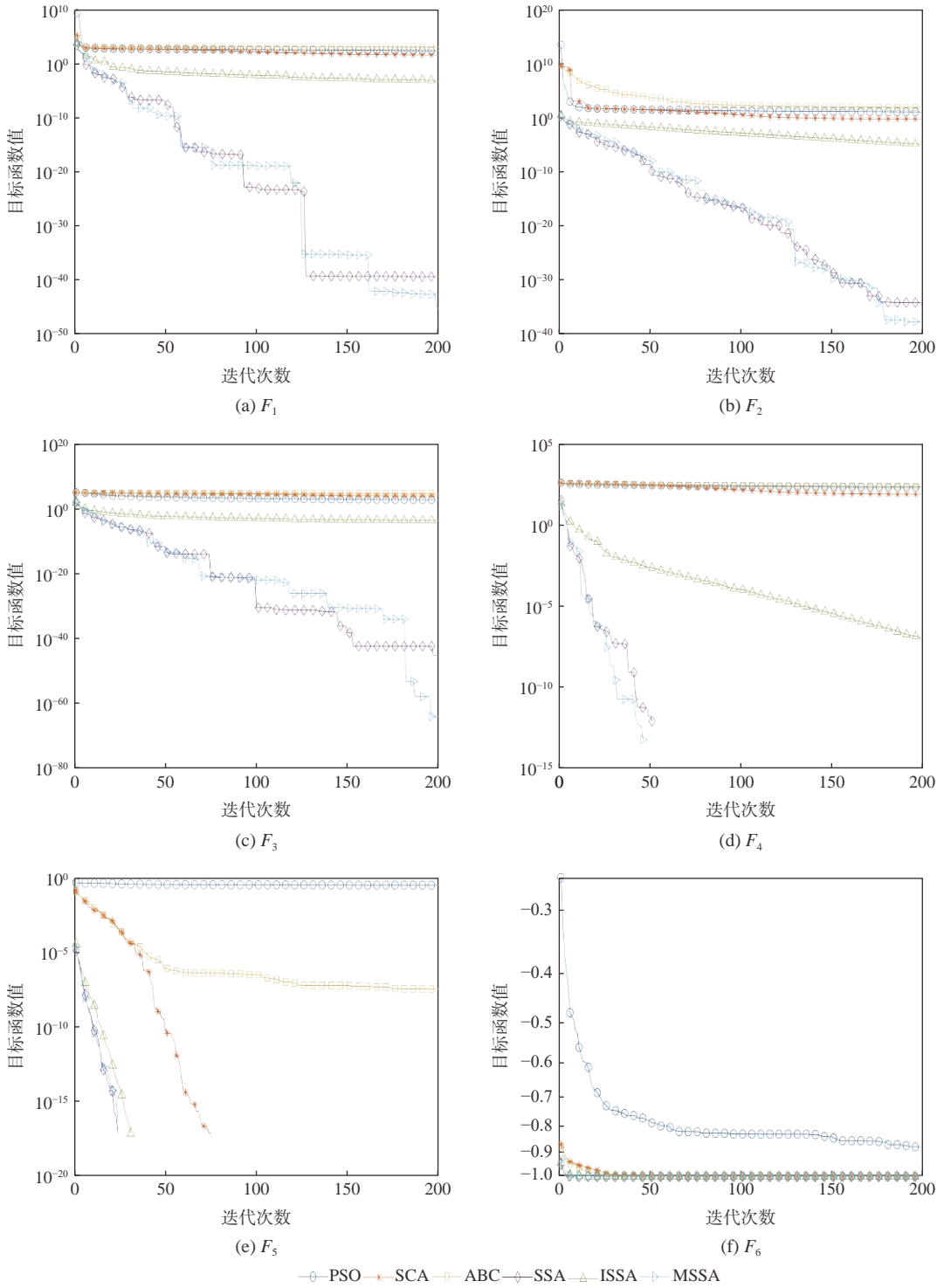


图4 算法的收敛曲线对比

Fig. 4 Comparison of convergence curves of algorithms

表4 时间复杂度对比

Table 4 Comparison of time complexity

算法	初始化种群阶段	发现者位置更新阶段	跟随者位置更新阶段	侦察预警机制阶段	全局最优位置更新阶段
SSA	$O(d+f(d))$	$O(d)$	$O(d)$	$O(d)$	$O(d)$
ISSA	$O(d+f(d))$	$O(d)$	$O(d)$	$O(d)$	$O(d)$
MSSA	$O(d+f(d))$	$O(d)$	$O(d)$	$O(d)$	$O(d)$

$$PSNR = 20\lg\left(\frac{255}{RMSE}\right) \tag{10}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^M \sum_{j=1}^Q (I(i, j) - Seg(i, j))^2}{M \times Q}} \tag{11}$$

式中: RMSE 为像素的均方根误差;  $M \times Q$  表示图像的大小;  $I(i, j)$  为原始图像的像素灰度值;  $Seg(i, j)$  为分割图像的像素灰度值。

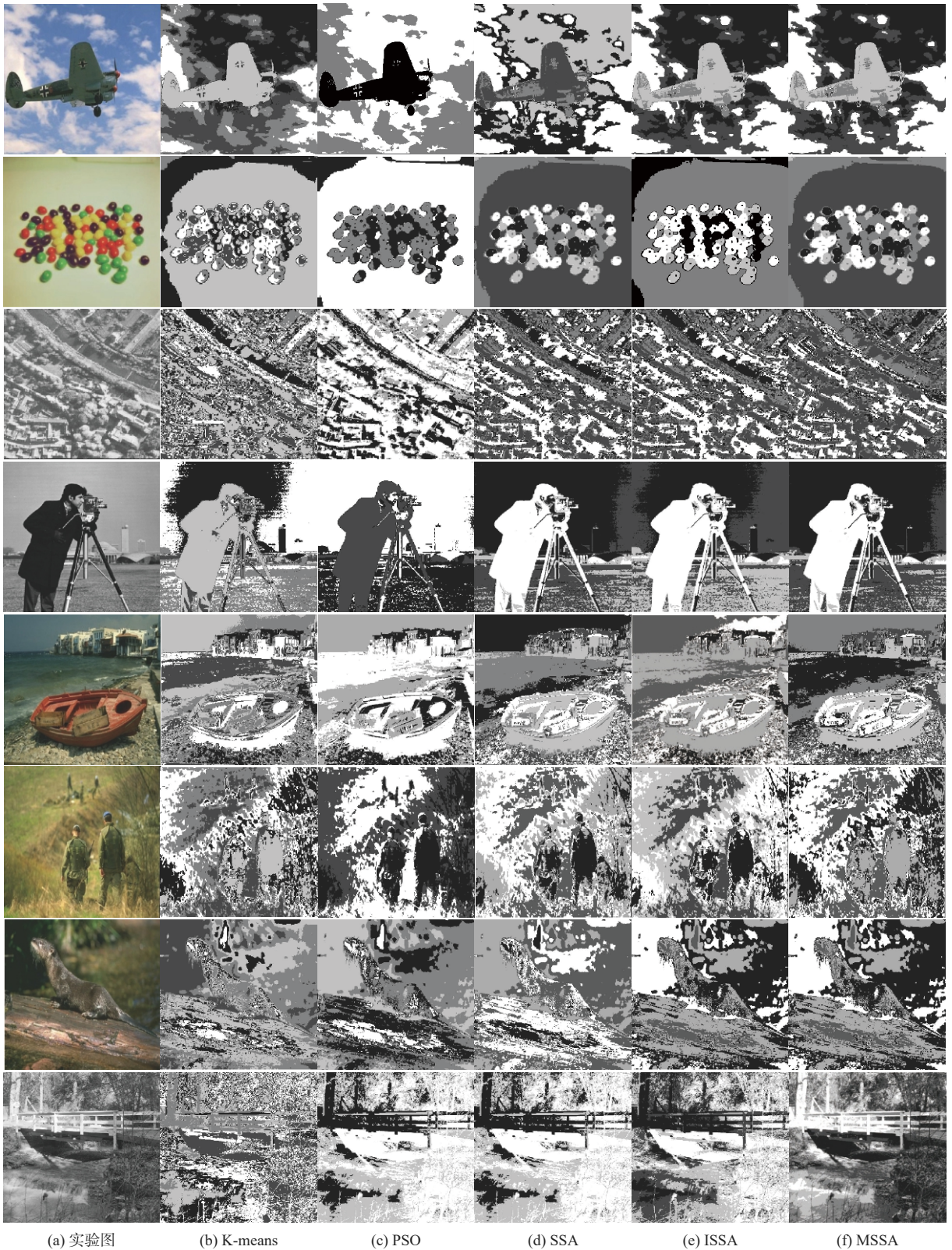


图 5 分割结果对比

Fig. 5 Comparison of segmentation results

SSIM 用于测量原始图像与分割图像之间的相似度。SSIM 值越大, 分割效果越好。SSIM 定义为

$$SSIM = \frac{(2u_1u_{seg} + c_1)(2\sigma_{1,seg} + c_2)}{(u_1^2 + u_{seg}^2 + c_1)(\sigma_1^2 + \sigma_{seg}^2 + c_2)} \quad (12)$$

式中:  $u_1$  和  $u_{seg}$  分别为原始图像与分割图像的平均

值;  $\sigma_1$ 和 $\sigma_{seg}$ 分别为原始图像与分割图像的标准差;  $\sigma_{1,seg}$ 为原始图像与分割图像的协方差;  $c_1$ 和 $c_2$ 为确保稳定性的常量。

FSIM 是反映原始图像和分割图像之间特征相似性的度量指标, 用于评价局部结构和提供对

表5 实验1的分割结果评估

Table 5 Evaluation of segmentation results for experiment 1

算法	PSNR	SSIM	FSIM
K-means	4.596 70	0.027 19	0.649 77
PSO	37.519 20	0.568 16	0.945 10
SSA	38.362 90	0.618 93	0.920 49
ISSA	39.625 70	0.716 13	0.939 37
MSSA	39.625 70	0.716 13	0.939 37

表6 实验2的分割结果评估

Table 6 Evaluation of segmentation results for experiment 2

算法	PSNR	SSIM	FSIM
K-means	4.012 60	0.025 74	0.616 62
PSO	36.238 70	0.462 39	0.935 54
SSA	41.532 60	0.750 25	0.949 16
ISSA	39.730 70	0.657 58	0.942 55
MSSA	41.553 40	0.751 00	0.949 76

表7 实验3的分割结果评估

Table 7 Evaluation of segmentation results for experiment 3

算法	PSNR	SSIM	FSIM
K-means	5.253 40	0.008 13	0.335 02
PSO	39.316 30	0.627 81	0.979 12
SSA	39.007 10	0.643 16	0.908 03
ISSA	39.108 80	0.648 45	0.906 52
MSSA	40.300 30	0.682 29	0.928 95

表8 实验4的分割结果评估

Table 8 Evaluation of segmentation results for experiment 4

算法	PSNR	SSIM	FSIM
K-means	6.097 30	0.101 49	0.522 70
PSO	37.963 20	0.598 25	0.939 10
SSA	39.386 30	0.689 16	0.969 29
ISSA	38.571 20	0.627 14	0.966 76
MSSA	39.550 90	0.699 45	0.972 00

表9 实验5的分割结果评估

Table 9 Evaluation of segmentation results for experiment 5

算法	PSNR	SSIM	FSIM
K-means	8.211 90	0.041 97	0.416 87
PSO	38.800 00	0.548 37	0.961 89
SSA	38.850 60	0.583 58	0.897 83
ISSA	37.288 00	0.450 27	0.908 25
MSSA	39.180 70	0.600 85	0.897 22

表10 实验6的分割结果评估

Table 10 Evaluation of segmentation results for experiment 6

算法	PSNR	SSIM	FSIM
K-means	6.707 90	0.027 50	0.506 15
PSO	39.117 20	0.661 99	0.944 26
SSA	39.829 90	0.627 55	0.934 91
ISSA	37.795 00	0.548 57	0.932 13
MSSA	40.383 20	0.671 80	0.908 82

表11 实验7的分割结果评估

Table 11 Evaluation of segmentation results for experiment 7

算法	PSNR	SSIM	FSIM
K-means	9.086 30	0.050 00	0.504 98
PSO	39.300 20	0.596 69	0.935 19
SSA	37.144 40	0.439 53	0.917 22
ISSA	40.737 10	0.680 29	0.916 86
MSSA	40.737 10	0.680 29	0.916 86

表12 实验8的分割结果评估

Table 12 Evaluation of segmentation results for experiment 8

算法	PSNR	SSIM	FSIM
K-means	6.522 40	0.014 55	0.345 18
PSO	36.683 60	0.490 72	0.965 19
SSA	36.683 60	0.490 72	0.965 19
ISSA	37.742 20	0.551 24	0.978 44
MSSA	39.945 10	0.646 52	0.988 83

$$S_L(x) = S_{PC}(x)S_G(x) \quad (14)$$

$$S_{PC}(x) = \frac{2PC_1(x)PC_2(x) + T_1}{PC_1^2(x)PC_2^2(x) + T_1} \quad (15)$$

$$S_G(x) = \frac{2G_1(x)G_2(x) + T_2}{G_1^2(x)G_2^2(x) + T_2} \quad (16)$$

$$G = \sqrt{G_x^2 + G_y^2} \quad (17)$$

$$PC(x) = \frac{E(x)}{\varepsilon + \sum_m A_s(x)} \quad (18)$$

$$PC_m(x) = \max(PC_1(x), PC_2(x)) \quad (19)$$

式中:  $\Omega$ 为原始图像的所有像素区域;  $S_L(x)$ 为相似性得分;  $PC_m(x)$ 为相位一致性度量;  $T_1$ 和 $T_2$ 为常量;  $G$ 为梯度下降;  $E(x)$ 为在位置  $x$  上的响应矢量大小, 并且尺度为  $s$ ;  $\varepsilon$  为一个极小的数值;  $A_s(x)$ 为尺度  $s$  的局部大小。

从表5~表12的度量指标上看, K-means 分割结果较差, PSNR值、SSIM值及FSIM值均远低于 PSO、SSA、ISSA及MSSA。MSSA在图5的实验1和实验7分割结果中PSNR值、SSIM值及FSIM

比度信息<sup>[25]</sup>。FSIM的取值范围为[0,1], 其值越接近1, 代表分割效果越好。FSIM的定义如式(13)~式(18)所示:

$$FSIM = \frac{\sum_{l \in \Omega} S_L(x)PC_m(x)}{\sum_{l \in \Omega} PC_m(x)} \quad (13)$$

值与ISSA一致,仅在FSIM值上略小于PSO,但是整体分割性能要优于其他算法;MSSA在图5的实验2、实验4和实验8的分割结果中PSNR值、SSIM值及FSIM值均高于其他算法,并在图5的实验8中FSIM值接近于1,表明分割图像与原始图像具有较高的特征相似性;MSSA在剩余测试图像的分割结果中PSNR值、SSIM值均要高于其他算法,仅在FSIM值上要低于PSO。从整体的分割结果评估中可以看出MSSA无论是在主观视觉效果还是度量指标,其分割性能要优于其他算法,可以产生良好的分割效果,也间接证明了MSSA良好的寻优性能,解决了SSA易陷入局部最优解的问题和K-means聚类算法对初始聚类中心敏感的缺点,从而得到最佳的初始聚类中心,进一步提升图像的分割质量。

## 4 结论

针对SSA存在的不足,提出一种新的改进算法MSSA,并通过6个基准函数与其他算法进行仿真对比,最后利用MSSA优化K-means算法进行图像分割,得出以下结论:

1) MSSA的全局搜索与局部开发能力得到提升,能够有效地跳出局部最优解,证实了提出的2种策略进行算法改进的有效性。首先,采用小孔成像反向学习策略丰富发现者的寻优位置,帮助算法跳出局部最优解;其次,对Sigmoid函数进行伸缩平移,提出一种新的自适应因子对安全阈值进行动态控制,平衡算法的全局搜索能力与局部开发能力。

2) 通过在6个基准函数上的仿真实验结果及改进算法的时间复杂度对比,可以看出MSSA的寻优性能要优于PSO、SCA、ABC和SSA,并在相同的算法时间复杂度的情况下比其他学者提出的改进算法ISSA的收敛精度更高,收敛速度更快。

3) 3个度量指标PSNR、SSIM及FSIM验证了MSSA算法突出的分割性能,解决了SSA易陷入局部最优解的问题和K-means聚类算法对初始聚类中心敏感的缺点,其分割效果均优于传统K-means算法、PSO算法、SSA和ISSA。

由于聚类数 $K$ 值的不确定性,将成为接下来改进的工作重点。

## 参考文献 (References)

- [1] CAPOR HROSIK R, TUBA E, DOLICANIN E, et al. Brain image segmentation based on firefly algorithm combined with K-means clustering[J]. *Studies in Informatics and Control*, 2019, 28(2): 167-176.
- [2] LI H Y, HE H Z, WEN Y G. Dynamic particle swarm optimization and K-means clustering algorithm for image segmentation[J]. *Optik*, 2015, 126(24): 4817-4822.
- [3] KAPOOR S, ZEYA I, SINGHAL C, et al. A grey wolf optimizer based automatic clustering algorithm for satellite image segmentation[J]. *Procedia Computer Science*, 2017, 115: 415-422.
- [4] KHRISSI L, EL AKKAD N, SATORI H, et al. Clustering method and sine cosine algorithm for image segmentation[J]. *Evolutionary Intelligence*, 2022, 15(1): 669-682.
- [5] XUE J K, SHEN B. A novel swarm intelligence optimization approach: sparrow search algorithm[J]. *Systems Science & Control Engineering*, 2020, 8(1): 22-34.
- [6] 吕鑫,慕晓冬,张钧.基于改进麻雀搜索算法的多阈值图像分割[J].*系统工程与电子技术*, 2021, 43(2): 318-327.  
LYU X, MU X D, ZHANG J. Multi-threshold image segmentation based on improved sparrow search algorithm[J]. *Systems Engineering and Electronics*, 2021, 43(2): 318-327(in Chinese).
- [7] 吕鑫,慕晓冬,张钧,等.混沌麻雀搜索优化算法[J].*北京航空航天大学学报*, 2021, 47(8): 1712-1720.  
LYU X, MU X D, ZAHNG J, et al. Chaos sparrow search optimization algorithm[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2021, 47(8): 1712-1720(in Chinese).
- [8] LIU G Y, SHU C, LIANG Z W, et al. A modified sparrow search algorithm with application in 3d route planning for UAV[J]. *Sensors*, 2021, 21(4): 1224.
- [9] YUAN J H, ZHAO Z W, LIU Y P, et al. DMPPT control of photovoltaic microgrid based on improved sparrow search algorithm[J]. *IEEE Access*, 2021, 9: 16623-16629.
- [10] ZHANG C L, DING S F. A stochastic configuration network based on chaotic sparrow search algorithm[J]. *Knowledge-Based Systems*, 2021, 220: 106924.
- [11] ZHU Y L, YOUSEFI N. Optimal parameter identification of PEMFC stacks using adaptive sparrow search algorithm[J]. *International Journal of Hydrogen Energy*, 2021, 46(14): 9541-9552.
- [12] LIU T T, YUAN Z, WU L, et al. Optimal brain tumor diagnosis based on deep learning and balanced sparrow search algorithm[J]. *International Journal of Imaging Systems and Technology*, 2021, 31(4): 1921-1935.
- [13] WANG P, ZHANG Y, YANG H W. Research on economic optimization of microgrid cluster based on chaos sparrow search algorithm[J]. *Computational Intelligence and Neuroscience*, 2021, 2021: 1-18.
- [14] OUYANG C T, ZHU D L, WANG F Q. A learning sparrow search algorithm[J]. *Computational Intelligence and Neuroscience*, 2021, 2021: 3946958.
- [15] TIAN H, WANG K Q, YU B, et al. Hybrid improved sparrow search algorithm and sequential quadratic programming for solving the cost minimization of a hybrid photovoltaic, diesel generator, and battery energy storage system[J]. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 2021, 2021: 1-17.
- [16] XING Z, YI C, LIN J H, et al. Multi-component fault diagnosis of wheelset-bearing using shift-invariant impulsive dictionary matching pursuit and sparrow search algorithm[J]. *Measurement*, 2021, 178: 109375.
- [17] LI J. Robot path planning based on improved sparrow algorithm[J]. *Journal of Physics:Conference Series*, 2021, 1861(1): 012017.
- [18] 徐航,张达敏,王依柔,等.基于高斯映射和小孔成像学习策略的

- 鲸鱼优化算法[J]. *计算机应用研究*, 2020, 37(11): 3271-3275.
- XU H, ZHANG D M, WANG Y R, et al. Whale optimization algorithm based on gauss map and small hole imaging learning strategy[J]. *Application Research of Computers*, 2020, 37(11): 3271-3275(in Chinese).
- [19] ERISOGLU M, CALIS N, SAKALLIOGLU S. A new algorithm for initial cluster centers in k-means algorithm[J]. *Pattern Recognition Letters*, 2011, 32(14): 1701-1705.
- [20] XIONG C Q, HUA Z, LV K, et al. An improved K-means text clustering algorithm by optimizing initial cluster centers[C]//International Conference on Cloud Computing and Big Data (CCBD). Piscataway: IEEE Press, 2016 : 265-268.
- [21] RANA S, JASOLA S, KUMAR R. A hybrid sequential approach for data clustering using K-means and particle swarm optimization algorithm[J]. *International Journal of Engineering, Science and Technology*, 2010, 2(6): 167-176.
- [22] SHI H B, XU M. A data classification method using genetic algorithm and K-means algorithm with optimizing initial cluster center[C]//IEEE International Conference on Computer and Communication Engineering Technology. Piscataway: IEEE Press, 2018 : 224-228.
- [23] HUSSAIN S F, PERVEZ A, HUSSAIN M. Co-clustering optimization using artificial bee colony(ABC) algorithm[J]. *Applied Soft Computing*, 2020, 97: 106725.
- [24] ABD EL AZIZ M, EWEES A A, HASSANIEN A E. Whale optimization algorithm and moth-flame optimization for multilevel thresholding image segmentation[J]. *Expert Systems with Applications*, 2017, 83: 242-256.
- [25] ZHAO D, LIU L, YU F H, et al. Ant colony optimization with horizontal and vertical crossover search: Fundamental visions for multi-threshold image segmentation[J]. *Expert Systems with Applications*, 2021, 167: 114122.

## Image segmentation based on Logistic regression sparrow algorithm

CHEN Gang<sup>1</sup>, LIN Dong<sup>1,\*</sup>, CHEN Fei<sup>2</sup>, CHEN Xiangyu<sup>1</sup>

(1. College of Physics and Information Engineering, Fuzhou University, Fuzhou 350108, China;

2. College of Computer and Data Science/College of Software, Fuzhou University, Fuzhou 350108, China)

**Abstract:** The sparrow search algorithm is improved to address its decrease of population diversity in the later stage and its easy fall into the local optimal solution. The improved algorithm introduces the oppositional learning strategy based small hole imaging to update the discoverer's position, enhancing the diversity of the optimal position. Then, inspired by the Logistic model, a new adaptive factor is proposed to dynamically control the safety threshold, thus balancing the global search and local development capabilities of the algorithm. Simulations of comparison with other algorithms in six benchmark functions are conducted, and experimental results show higher convergence accuracy and speed of the improved algorithm than those of the other algorithms. In engineering applications, the proposed algorithm optimizes the K-means clustering algorithm for image segmentation with satisfactory segmentation performance in terms of peak signal to noise ratio (PSNR), structural similarity (SSIM) and feature similarity (FSIM).

**Keywords:** sparrow search algorithm; image segmentation; oppositional learning based small hole imaging; Logistic model; K-means clustering algorithm

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0269

# 基于电感特征的开关磁阻电机电流斩波控制策略

陈越<sup>1</sup>, 蒋启龙<sup>1,\*</sup>, 王金锁<sup>1</sup>, 姚卫丰<sup>2</sup>

(1. 西南交通大学 电气工程学院, 成都 611756; 2. 深圳职业技术学院 机电工程学院, 深圳 518055)

**摘要:** 开关磁阻电机 (SRM) 运行在基速以下时常采用电流斩波控制 (CCC), 针对传统电流斩波控制中电流动态跟踪能力弱、换相区内转矩脉动大及功率器件开关频率不固定的问题, 根据电感曲线的变化特征进行区间分段, 提出一种基于参考电流补偿的电流斩波控制策略。在低电感区段内根据电机转速、负载的大小对参考电流进行补偿, 提高相绕组在换相过程中的转矩输出能力和动态响应能力; 在电感曲线的线性上升阶段, 采用固定频率的脉冲宽度调制 (PWM) 波进行控制, 使输出转矩更平滑。搭建三相 12/8 极开关磁阻电机仿真模型及硬件在环实验平台, 考虑电机在不同转速、负载下的运行工况, 选取转矩脉动指标进行对比。仿真与实验结果表明: 所提控制策略能有效减小开关磁阻电机的转矩脉动, 提高电机的运行性能。

**关键词:** 开关磁阻电机; 电流斩波控制; 转矩脉动; 参考电流; 硬件在环

**中图分类号:** TM352

**文献标志码:** A

**文章编号:** 1001-5965(2023)03-0647-10

开关磁阻电机 (switched reluctance motor, SRM) 具有结构简单、容错能力强、调速范围宽等优点, 自问世以来便得到了快速的发展。但 SRM 特殊的双凸极结构和运行时磁路的高度饱和特性, 使得其电磁转矩是转子位置角和定子相电流的非线性函数, SRM 的这种特点决定了其具有瞬时输出转矩脉动大的问题。如何降低 SRM 的转矩脉动成为电气传动及电机控制领域的研究热点。

目前, 国内外学者在 SRM 的转矩脉动抑制策略方面做了大量研究。文献 [1] 针对 SRM 电感非线性的特点, 通过优化开通角和关断角来提高 SRM 的运行性能。文献 [2] 提出一种基于相电感特征的 SRM 无位置传感器控制方法, 用以提高电机控制中转子位置角度的计算精度。文献 [3] 提出一种适用于 SRM 宽速域调速的离线转矩分配函数 (torque sharing function, TSF), 该离线 TSF 在不增加铜耗的情况下能显著降低转矩脉动。文献 [4] 在传统转矩分配函数控制策略下, 通过对导通角和重叠角进行

优化, 进一步拓宽了 SRM 的调速范围。文献 [5] 在直接转矩控制下选取一种通用全桥变换器直接与 SRM 各相绕组连接, 该方法具有良好的动、稳态性能。文献 [6] 以提高转矩电流比为目标对传统直接转矩控制进行改进, 在降低转矩脉动的同时提高了效率。文献 [7] 在直接转矩的思想提出直接瞬时转矩控制, 该方法以电机瞬时转矩为控制对象, 对降低 SRM 的转矩脉动有较好的效果。文献 [8] 采用一种新型四电平功率变换器来改善传统直接瞬时转矩控制电压矢量少的问题, 使绕组具有更快的磁化和退磁特性。文献 [9] 提出一种基于脉冲宽度调制 (pulse width modulation, PWM) 的 SRM 直接瞬时转矩控制方法, 该方法在运行时能自适应地调整滞环大小。文献 [10] 设计了一种比例和积分 (proportional integral, PI) 参数自适应的电流调节器, 能有效改善 SRM 调速系统的运行性能。文献 [11-12] 提出了基于预测电流控制算法的 SRM 转矩脉动抑制策略, 以降低电机在换相时的转矩脉动。另外,

收稿日期: 2021-05-24; 录用日期: 2022-01-16; 网络出版时间: 2022-03-11 11:57

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20220309.2059.006.html](https://kns.cnki.net/kcms/detail/11.2625.V.20220309.2059.006.html)

\* 通信作者. E-mail: [double\\_long@126.com](mailto:double_long@126.com)

**引用格式:** 陈越, 蒋启龙, 王金锁, 等. 基于电感特征的开关磁阻电机电流斩波控制策略 [J]. 北京航空航天大学学报, 2023, 49 (3): 647-656. CHEN Y, JIANG Q L, WANG J S, et al. Current chopping control strategy of switched reluctance motor based on inductance characteristics [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 647-656 (in Chinese).

文献 [13-15] 中基于模糊控制、神经网络控制及基于迭代学习的控制等智能控制理论也被广泛应用于 SRM 的转矩脉动抑制研究。

SRM 在基速下常采用电流斩波控制 (current chop control, CCC), 但是传统的 CCC 难以使实际电流跟踪参考电流, 存在相转矩的欠补偿问题。在传统 CCC 的基础上, 本文提出一种基于电感特征分段控制的 CCC 策略 (改进 CCC)。通过对参考电流进行补偿, 增强绕组在换相区间内的转矩输出能力。最后, 对所提控制策略进行了仿真和硬件在环实验验证。

## 1 开关磁阻电机的传统电流斩波控制

### 1.1 数学模型

在线性模型下, SRM 的第  $k$  相电压平衡方程式可以表示为

$$u_k = R_k i_k + \frac{d\psi_k}{dt} \quad (1)$$

式中:  $u_k$  为第  $k$  相电机绕组电压;  $R_k$  为绕组电阻;  $i_k$  为绕组电流;  $t$  为时间;  $\psi_k$  为绕组磁链, 表示关于  $i_k$  和转子位置  $\theta$  的函数, 也是  $L_k(\theta, i_k)$  和电流  $i_k$  的乘积, 表示为

$$\psi_k(\theta, i_k) = L_k(\theta, i_k) i_k \quad (2)$$

式中:  $L_k$  为第  $k$  相绕组电感;  $\theta$  为转子位置角。

将式 (2) 代入式 (1), 电压方程可以改写为

$$u_k = R_k i_k + L_k(\theta, i_k) \frac{di_k}{dt} + i_k \frac{\partial L_k(\theta, i_k)}{\partial \theta} \omega_r = R_k i_k + L_k(\theta, i_k) \frac{di_k}{dt} + e(\theta, i_k) \quad (3)$$

式中:  $\omega_r = d\theta/dt$ , 为转子角速度;  $e(\theta, i_k)$  为第  $k$  相绕组电感变化产生的反电动势。

SRM 转矩的产生遵循磁阻最小原理。在磁路饱和状态下运行的 SRM 是一种高度非线性的机电装置, 忽略各绕组间互感影响, 在磁路线性区域, 第  $k$  相所产生的转矩  $T_k$  可以近似表示为

$$T_k = \frac{1}{2} i_k^2 \frac{\partial L_k(\theta, i_k)}{\partial \theta} \quad (4)$$

SRM 在运行过程中的总瞬时转矩为各相瞬时转矩之和, 即合成总瞬时转矩为

$$T_{\text{total}}(\theta, i) = \sum_k^m \frac{1}{2} i_k^2 \frac{\partial L_k(\theta, i_k)}{\partial \theta} = T_L + B\omega_r + J \frac{d\omega_r}{dt} \quad (5)$$

式中:  $T_{\text{total}}$  为总瞬时转矩;  $m$  为电机相数;  $T_L$  为负载转矩;  $B$  为摩擦系数;  $J$  为转动惯量。

由式 (5) 可知, 线性模式下总电磁转矩是相电

流和转子位置的函数, 相电流的微小变化将引起相转矩按电流的平方倍变化。考虑实际电机中磁路饱和的影响, 相转矩与其电流不完全按照平方的关系变化, 但仍随电流的增大而增大。因此, 对 SRM 运行时的相电流进行精确控制是提高运行性能的关键。

### 1.2 电流斩波控制逻辑

SRM 运行过程中电流波形的形态、电流峰值及其出现的位置都会对电机的运行性能造成影响。传统 CCC 下的电流瞬时波形如图 1 所示, 在绕组导通区间, 电流滞环控制器对绕组电流进行斩波控制, 通过对滞环带的合理设置, 可以将绕组电流  $i_k$  控制在参考电流  $i_{\text{ref}}$  附近波动, 得到的电流波形趋近于平顶电流波形。图中:  $\theta_0$  为不对齐位置,  $\theta_1$  为定子齿开始重叠的位置,  $\theta_2$  为定子齿全部重合的起始位置,  $\theta_{\text{on}}$  和  $\theta_{\text{off}}$  为相绕组开通角和关断角。

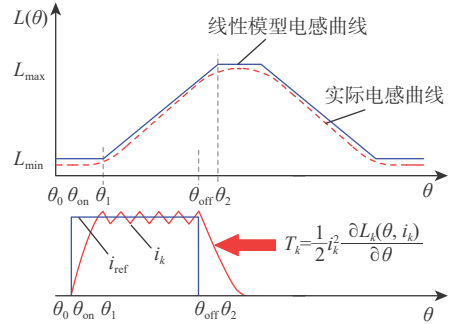


图 1 斩波控制下瞬时相电流波形

Fig. 1 Instantaneous current waveform controlled by chopper

CCC 中通常将一个开关管参与斩波的方式称为软斩波, 软斩波采用自然续流的方式对绕组进行退磁, 电流的变化速率相对缓慢, 不易造成转矩频繁的波动, 图 2 为 SRM 的 A 相绕组在电动状态下采用 CCC 的示意图,  $u_s$  为电源电压;  $v_L$  为绕组  $L$  的电压; S1 和 S2 为开关管; D1 和 D2 为续流二极管。参考电流  $i_{\text{ref}}$  由速度环给出, 将其与绕组电流  $i_k$  比较后送入滞环控制器, 从而输出开关管的导通逻辑。图 2 中设定电流上限  $i_h = i_{\text{ref}} + 0.5h$ , 电流下限  $i_l = i_{\text{ref}} - 0.5h$ ,  $h$  为滞环宽度且  $h = i_h - i_l$ 。当电机某相绕组导通后, 下开关管闭合, 上开关管参与斩波, 绕组电流  $i_k$  从 0 开始上升。当电流瞬时值达到滞环上限值  $i_h$  时, 上开关管关断斩波, 电流  $i_k$  自然续流; 当绕组电流  $i_k$  瞬时值达到滞环带下限值  $i_l$  时, 上开关管又重新导通, 电流  $i_k$  开始重新上升。在整个绕组导通区间, 上开关管的反复通断使得绕组电流在参考电流附近反复波动。本文功率变换桥选用不对称半桥结构的功率电路。根据其工作过程, 定义一相绕组中励磁、续流和退磁 3 种开关状态分别为

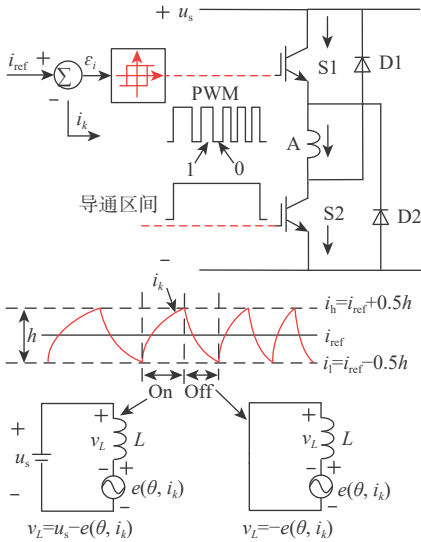


图 2 电流软斩波示意图

Fig. 2 Schematic diagram of soft chopping current control

“1”、“0”和“-1”。电流软斩波时,在绕组导通区间内电路工作在“1”和“0”状态。

### 1.3 转矩变化特征

从式(4)可知,相转矩含有电感  $L_k$  关于  $\theta$  的偏导,以三相 12/8 极 SRM 为研究对象,通过有限元仿真得出样机在恒定电流时的电感曲线及其偏导数  $\partial L/\partial\theta$ ,以  $I=6\text{ A}$  为例,如图 3 所示。

从图 3 可知,在小电感区  $\partial L/\partial\theta$  的值比较小;在电感转折区,  $\partial L/\partial\theta$  急剧增大。一般为了增加励磁相的输出转矩,在小电感区绕组电流就已达至斩波幅值,假设斩波过程电流平均值大小为  $I_0$ ,电机相电感在  $\partial L/\partial\theta$  突变区前期和后期偏导数分别为  $k_{L1}$  和  $k_{L2}$ ,显然  $k_{L1} \ll k_{L2}$ ,在该区域内相转矩有:

$$T_e = \frac{1}{2} I_0^2 k_{L1} \ll T_e' = \frac{1}{2} I_0^2 k_{L2} \quad (6)$$

在小电感区内,  $\partial L/\partial\theta$  的值比较小,此时电流幅值被斩波带所限,导致在该区间内产生的相转矩很小,此时退磁相产生的转矩通常已进入下降状态,容易造成合成转矩的不足。在电感转折区,  $\partial L/\partial\theta$  的值急剧增大,从式(6)可以看出,电机相转矩从一个很小值急剧增大,然而控制器的响应速度往往难以快速跟踪输出,当退磁相转矩的减少量不能抵消励磁相转矩的增加量时,合成转矩又会表现出过补偿现象。之后转子进入电感线性上升区,  $\partial L/\partial\theta$  的值相对稳定,此时 CCC 输出转矩也相对稳定。在小电感区及电感转折区,传统 CCC 难以保持输出瞬时转矩的稳定,需要采取额外的控制措施。

为了在小电感区使电机的输出转矩能够达到期望值,要求该相绕组在该区间的电流也要达到一定的幅值,这样才能增强电机的转矩输出能力。而

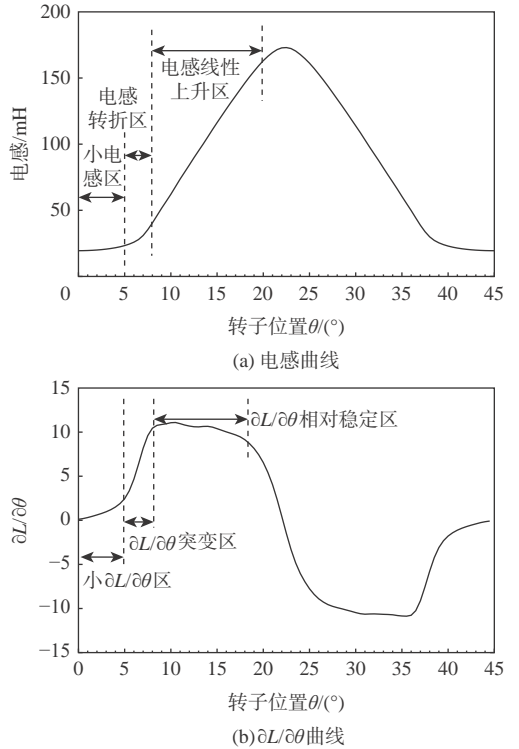


图 3 样机  $I=6\text{ A}$  时电感曲线及  $\partial L/\partial\theta$  曲线

Fig. 3 Curves of inductance and  $\partial L/\partial\theta$  for prototype  $I=6\text{ A}$

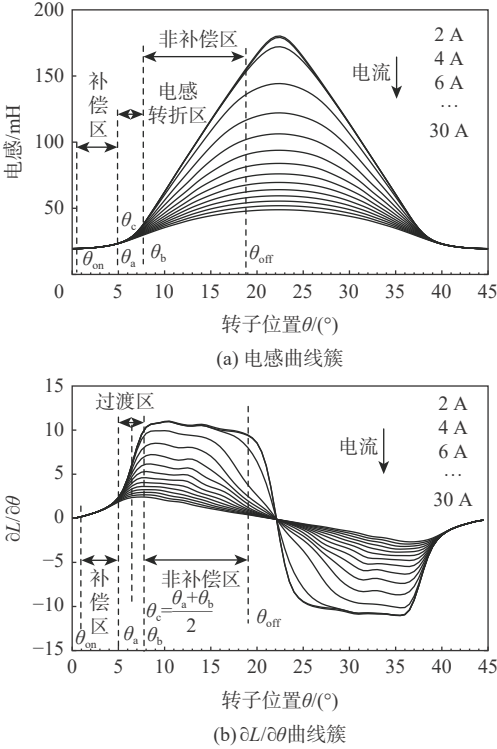
传统 CCC 的峰值电流被限,在换相的初始阶段,绕组电流难以满足该区间内的转矩输出要求,不仅容易造成大的转矩波动,而且系统在扰动下的动态响应也十分缓慢。为了增强电机的转矩输出能力,提出一种基于电感特征的 CCC 策略。

## 2 基于电感特征的电流斩波控制策略

### 2.1 斩波电流补偿策略

由第 1 节可知,在相绕组导通区间内,小电感区及电感上升区应当有不同的斩波电流。因此,可以根据电感曲线的变化特点进行分段控制。通过有限元仿真所得样机的电感曲线簇及其  $\partial L/\partial\theta$  曲线簇如图 4 所示。

由图 4(b)可知,电感关于  $\theta$  的偏导数  $\partial L/\partial\theta$  在  $\theta_a$  和  $\theta_b$  附近发生明显变化。因此,以  $\theta_a$  和  $\theta_b$  为分界点,将绕组导通区间分为参考电流补偿区、过渡区和参考电流非补偿区,并取  $\theta_a=5^\circ$ ,  $\theta_b=7.8^\circ$ 。在参考电流补偿区,有  $\theta_{on} < \theta < \theta_a$ ,此时  $\partial L/\partial\theta$  的值很小,在原参考电流下相绕组的转矩输出能力不足,通过预存的“补偿电流—转速—负载”函数计算出补偿电流  $\Delta i_{ref}$  的大小,然后将补偿电流叠加到原参考电流之上,再送入电流滞环控制器进行斩波控制,即实际斩波的参考电流为  $i_{ref}' = i_{ref} + \Delta i_{ref}$ ;过渡区间即为  $\partial L/\partial\theta$  突变区,有  $\theta_a < \theta < \theta_b$ ,在该阶段内由于  $\partial L/\partial\theta$  快速上升而使得相转矩也快速上升,为了防止合成转

图4 样机电感曲线簇与 $\partial L/\partial\theta$ 曲线簇Fig. 4 Curve clusters of inductance and  $\partial L/\partial\theta$  of prototype

矩在该区间内有大的波动,需要将在补偿区内叠加的补偿电流减小甚至去掉。 $\theta_c$ 为 $\theta_a$ 和 $\theta_b$ 的中点,在 $\theta_c$ 处提前将补偿电流减小到0;在参考电流非补偿区,即 $\theta_b < \theta < \theta_{off}$ 时, $\partial L/\partial\theta$ 的值相对稳定,相转矩在该区间段内也相对稳定,保持原参考电流 $i_{ref}$ 不变,无需进行电流补偿。

参考电流补偿值的大小根据转速和负载大小进行调节。轻载时,由绕组电感变化而产生的旋转电动势对绕组电流影响比较大,此时补偿电流的大小主要由电机转速决定;当电机负载较重时,由相绕组电流变化产生的变压器电动势增加,此时补偿电流的大小应主要由负载的大小决定。根据仿真或实验数据拟合出一个“补偿电流—转速—负载”函数,控制器根据该函数实时计算补偿电流值的大小。

开通角 $\theta_{on}$ 对换相电流有很大影响,采用在线调节开通角的方式,根据文献[16]计算开通角大小,计算式为

$$\theta_{on} = \theta_1 - \frac{L_{min} i_{ref} \omega_r}{u_s} \quad (7)$$

开通角在线调节流程如图5所示。

## 2.2 电流软斩波下的PWM控制策略

在传统CCC下,开关管的通断完全由电流滞环控制器决定,控制器根据斩波逻辑输出频率变化的PWM波,存在开关频率不固定的问题。并且随

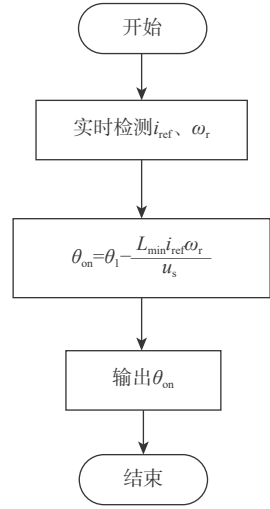


图5 开通角在线调节流程

Fig. 5 Flow chart of online adjustment of turn-on angle

着转速的升高,电流的变化速率加快,当电流采样速率较低时,存在下一时刻的采样值超出滞环带的情况,控制器无法及时将电流控制在滞环宽度以内,从而也造成了转矩脉动的产生。为了进一步控制相电流,对电流控制器进行改进,输出频率固定的PWM波来控制开关管的通断。

固定频率斩波控制的实现过程如图6所示, $T_s$ 为原电流软斩波时控制周期; $T'_s$ 为固定频率控制时的周期。当电机运行于低速状态且滞环宽度不是太小时,电流的变化速率较慢,电流斩波的频率较低,此时开关管在开通状态“on”或关断状态“off”任意一个状态上的保持时间比较长。在整个“on”阶段,绕组两端承受正压,处于正向励磁状态,此时可将“on”阶段视为一占空比 $D=1$ 的PWM波;在整个“off”阶段绕组都处于零压状态,绕组自然续流,可视为占空比 $D=0$ 的PWM波。此时可以在“on”或者“off”阶段用频率固定、占空比可调的PWM波来代替原电流软斩波输出的PWM波形,实现定频控制。

定频控制过程仍然按照电感特征的区间分段进行控制,在电流补偿区,需要使相电流快速达到目标参考值,而实际电流通常需要经过多个周期才能达到目标,因此,此阶段内PWM的占空比设置为1,即控制器按照参考电流进行补偿后的斩波策略输出控制逻辑。在电流非补偿区,当绕组电流上升至 $i_b$ 或下降至 $i_1$ 时达到滞环带的边界点,按电流斩波方式关断或开通开关管,而在滞环宽度内按照固定频率的PWM进行控制,并在滞环带的边界点处进行PWM占空比的更新。图6中 $i_k$ 和 $i'_k$ 分别为原电流斩波控制和固定频率PWM控制时的电流波形。在电流的上升阶段,施加占空比略小于1的

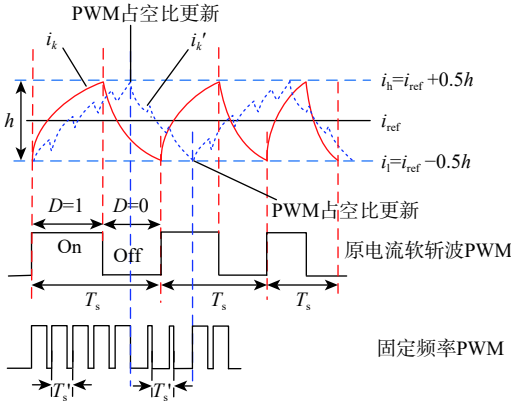


图 6 电流控制器 PWM 调制图

Fig. 6 PWM diagram of current controller

PWM 波, 经过几个 PWM 周期后, 绕组电流  $i_k$  到达滞环带的上斩波点并更新 PWM 的占空比; 在电流下降阶段并不完全关断主开关管, 而是施加占空比略大于 0 的 PWM 波。之后在每个电流斩波周期内施加相同的控制, 单个控制周期内电流变化的幅值变小, 不易出现超调, 从而产生的转矩也更加平滑。

所提电流软斩波控制方法的系统框图如图 7 所示, 主要由速度环控制器、参考电流补偿器、电流控制器、功率变换器、位置传感器及开关磁阻电机等组成, 图中:  $\omega_{ref}$  为角速度参考值,  $\omega_{act}$  为角速度实际反馈值,  $\varepsilon_\omega$  为角速度误差值,  $\omega$  为角速度, S1~S6 为功率开关。

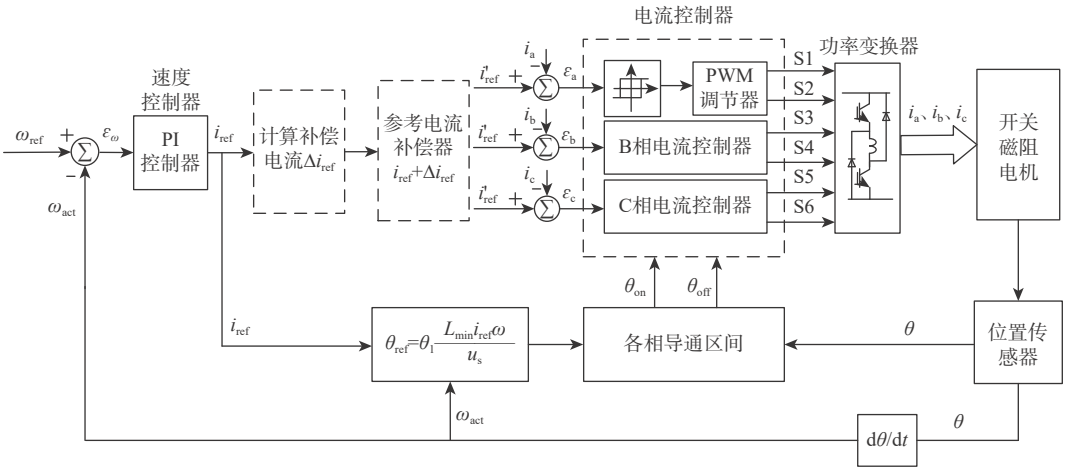


图 7 电流软斩波控制系统框图

Fig. 7 Block diagram of soft chopping current control system

### 3 仿真与实验

基于第 2 节的 CCC 方法, 搭建了 12/8 极三相 SRM 控制系统仿真模型进行实验, 电机模型根据查表法建立。仿真参数如下: 额定电压为 540 V, 额定功率为 7.5 kW, 额定转速为 1 000 r/min, 定子相绕组电阻  $R=1.1 \Omega$ , 电机系统转动惯量  $J=0.02 \text{ kg}\cdot\text{m}^2$ , 摩擦系数  $B=0.001 \text{ Ns/m}^2$ , 仿真中在线调节开通角, 关断角  $\theta_{off}=19^\circ$ 。电流斩波滞环宽度  $h=1 \text{ A}$ , PWM 频率为 10 kHz, 设定电流上升阶段占空比  $D=0.9$ , 下降阶段  $D=0.1$ 。为了验证本文所提策略对于提升电机运行性能的有效性, 采用转矩脉动系数指标进行对比, 转矩脉动系数  $K_T$  可以表示为

$$K_T = \frac{T_{\max} - T_{\min}}{T_{\text{av}}} \times 100\% \quad (8)$$

式中:  $T_{\max}$ 、 $T_{\min}$  和  $T_{\text{av}}$  分别为稳态时最大转矩、最小转矩和平均转矩。

通过仿真获得不同转速和负载下的最优参考电流补偿值, 仿真结果如图 8 所示。

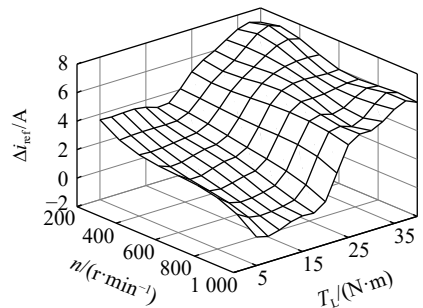


图 8 参考电流补偿值曲面

Fig. 8 Compensation surface of reference current

根据仿真结果拟合出的“ $\Delta i_{ref} - n - T_L$ ”的函数表达式为

$$\begin{cases} \Delta i_{ref} = 3.597 - 0.003 6n + 0.036 8T_L & 0 < T_L \leq 20, \quad 0 < n \leq 1000 \\ \Delta i_{ref} = 3.104 - 0.001 3n + 0.106 4T_L & T_L > 20, \quad 0 < n \leq 1000 \end{cases} \quad (9)$$

图 9 为电机负载转矩为 30 N·m, 转速为 300 r/min 时传统 CCC 与改进 CCC 后的电压、电流及瞬时转矩的仿真波形。图 10 为两相间的典型换相转矩波

形,可以看出,传统 CCC 初始导通阶段转矩上升缓慢,然后又急剧上升,合成转矩呈现出从过小到过大的变化;而改进 CCC 明显提升了初始导通阶段励磁相绕组的输出转矩,励磁绕组在小电感区有更强的转矩输出能力,相转矩的上升过程更为平稳,没有出现过大的相转矩峰值,有效避免了换相过程中大的转矩波动。

图 11 和图 12 负载转矩为 30 N·m,转速为 500 r/min 和 1 000 r/min 时的仿真波形。可以看出,在励磁相绕组导通后,励磁相的电流得到补偿,能够有效增强换相初期的转矩输出能力,在非补偿区采用固定频率的 PWM 波控制,使得输出转矩更为平缓。

电机转速分别为 300, 500, 1 000 r/min 时,传统 CCC 转矩脉动分别为 84.4%、70.1% 和 56.2%,改进 CCC 后的转矩脉动分别为 28.8%、26.9% 和 36.8%。在额定转速下运行时,电机的运动电动势增加,转矩脉动系数的减小幅度变小,但改进 CCC 的转矩脉动系数相比于传统 CCC 仍然能够得到有效降低,仿真结果表明了所提策略的有效性。

为了进一步验证所提策略的有效性,基于 RT BOX 硬件在环实验平台,对所提策略进行实验验证。控制算法通过以 DSP 为核心的控制器实现,电机及功率回路通过 RT BOX 仿真盒子模拟实现。硬件在环实验参数与仿真一致。图 13~图 16 为电

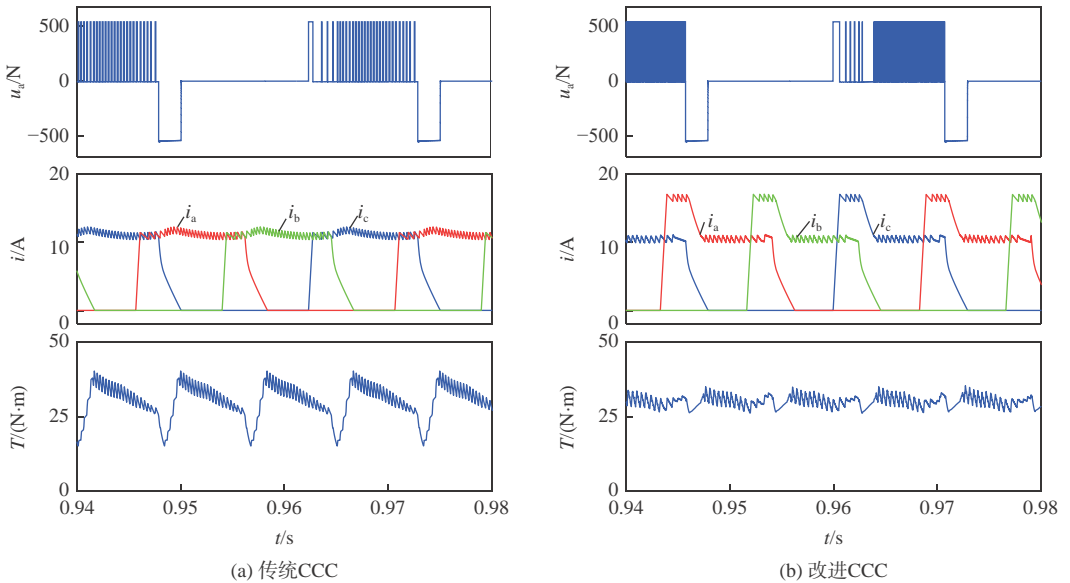


图 9  $n=300$  r/min,  $T_L=30$  N·m 仿真结果

Fig. 9 Simulation results at  $n=300$  r/min and  $T_L=30$  N·m

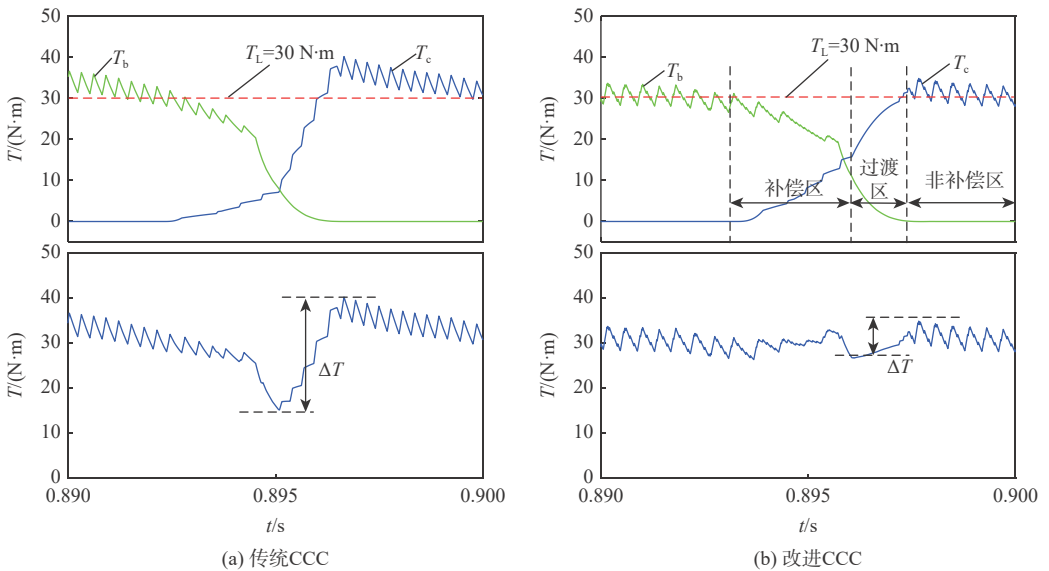


图 10 典型换相转矩波形

Fig. 10 Typical commutation torque waveform

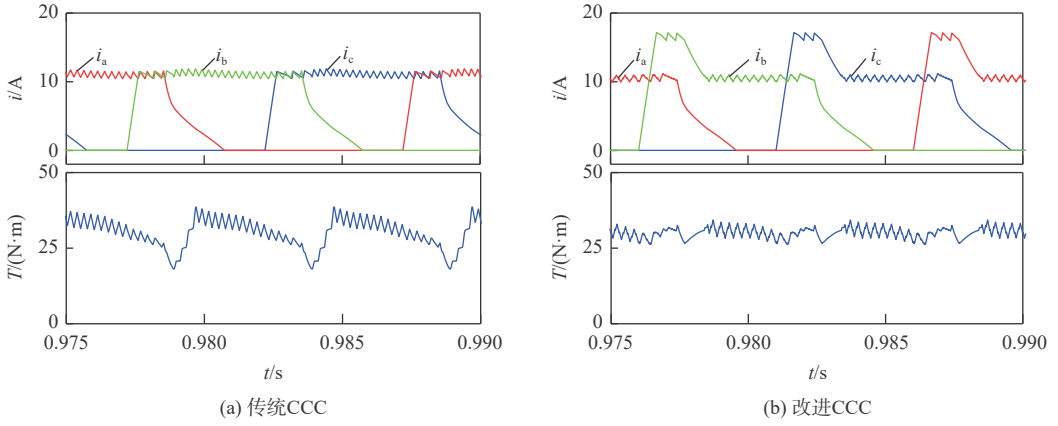


图 11  $n=500$  r/min,  $T_L=30$  N·m 仿真结果

Fig. 11 Simulation results at  $n=500$  r/min and  $T_L=30$  N·m

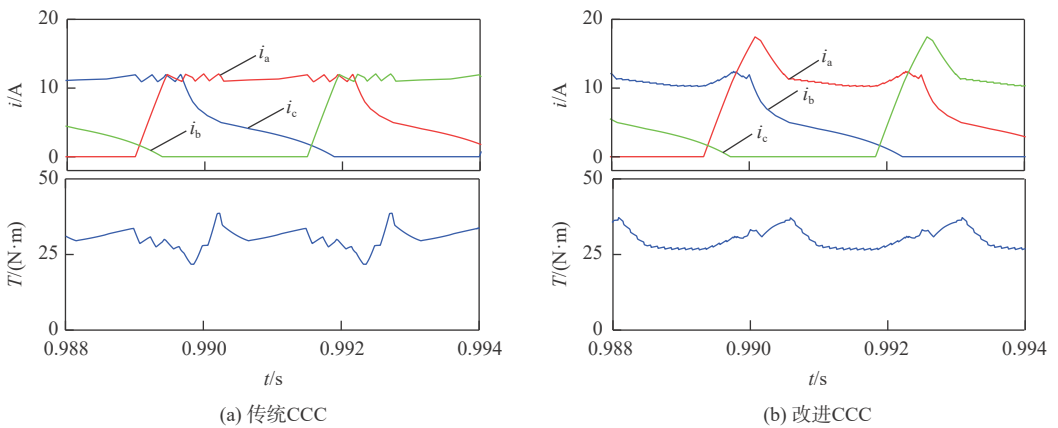


图 12  $n=1000$  r/min,  $T_L=30$  N·m 仿真结果

Fig. 12 Simulation results at  $n=1000$  r/min and  $T_L=30$  N·m

机转速为 300, 500 r/min, 负载转矩为 10, 30 N·m 时稳态相电流和瞬时转矩波形。其中, 图 13~图 14 中波器的时间刻度为每格 10 ms, 图 15~图 16 中波器的时间刻度为每格 5 ms。

从图 13~图 16 可以看出, 在传统 CCC 下, 由

于电流滞环控制的作用使得相电流的幅值被限, 造成换相过程中励磁相绕组的转矩输出不足, 随着负载的加重, 转矩脉动也越大。采用改进 CCC 后, 换相区间的参考电流得到补偿, 绕组电流明显提高, 增加了励磁相绕组的转矩输出能力, 能够对换相区

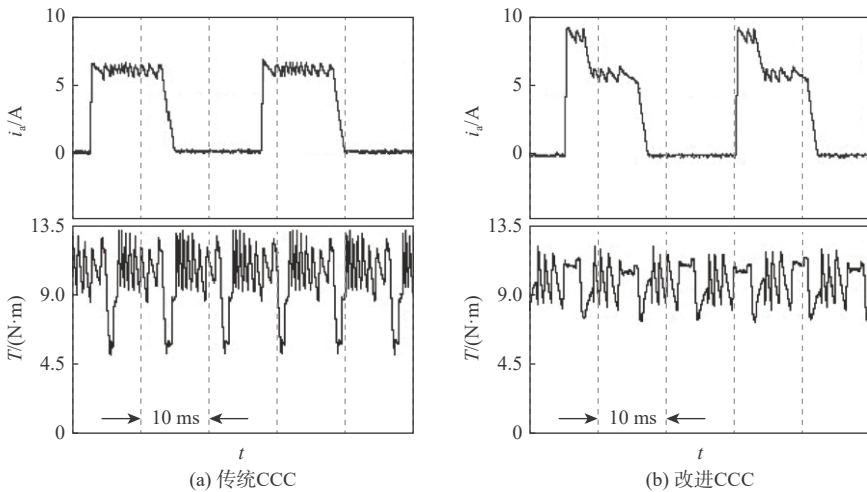
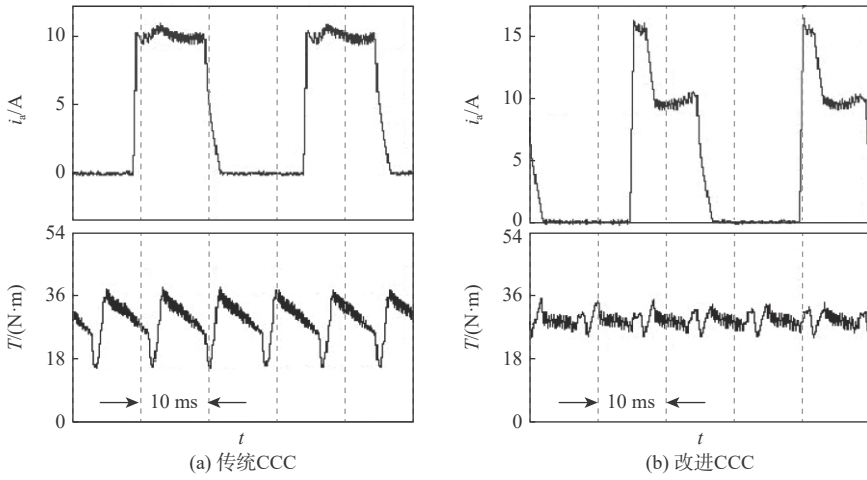
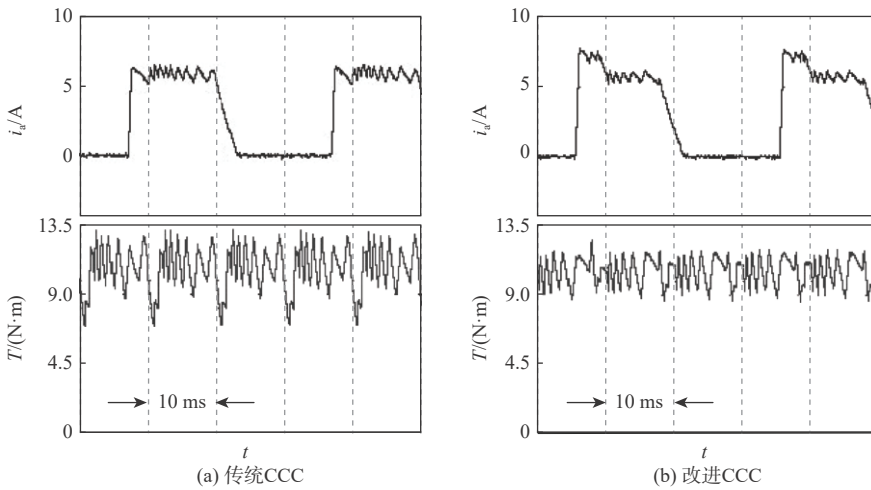
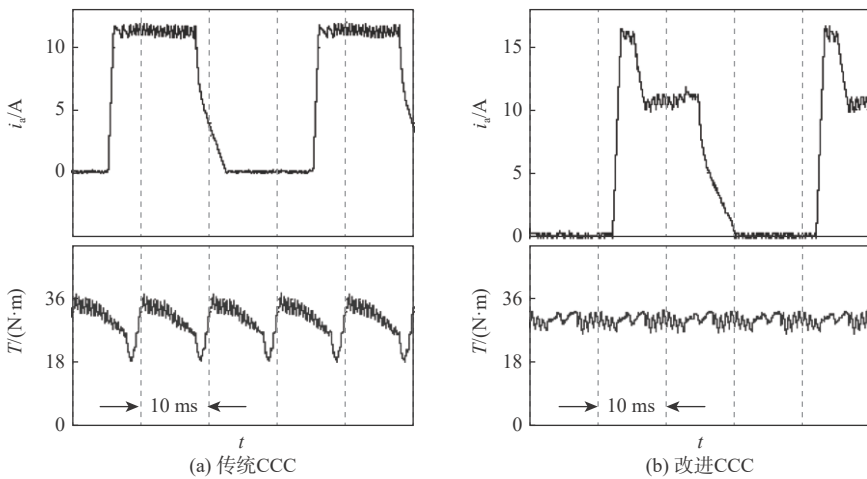


图 13  $n=300$  r/min,  $T_L=10$  N·m 实验结果

Fig. 13 Experimental results at  $n=300$  r/min and  $T_L=10$  N·m

图 14  $n=300$  r/min,  $T_L=30$  N·m 实验结果Fig. 14 Experimental results at  $n=300$  r/min and  $T_L=30$  N·m图 15  $n=500$  r/min,  $T_L=10$  N·m 实验结果Fig. 15 Experimental results at  $n=500$  r/min and  $T_L=10$  N·m图 16  $n=500$  r/min,  $T_L=30$  N·m 实验结果Fig. 16 Experimental results at  $n=500$  r/min and  $T_L=30$  N·m

间的转矩起到补偿作用。在负载转矩为 30 N·m 时,励磁相绕组的补偿电流值更大,对换相区间的转矩脉动抑制作用也更明显。

图 17 和图 18 分别为负载转矩为 10 N·m 和

30 N·m 时 2 种方法的转矩脉动系数对比图。可以看出,改进 CCC 在不同转速下都能有效减小转矩脉动,并且在负载较重时,转矩脉动抑制效果越明显,验证了所提策略的有效性。

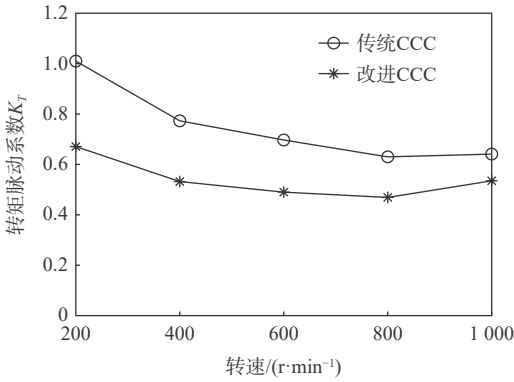


图 17  $T_L=10 \text{ N}\cdot\text{m}$  时转矩脉动系数对比

Fig. 17 Comparison of torque ripple coefficient at  $T_L=10 \text{ N}\cdot\text{m}$

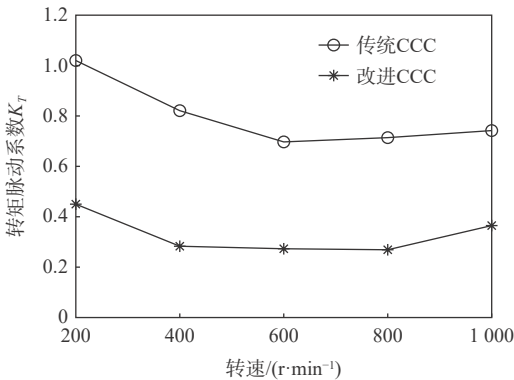


图 18  $T_L=30 \text{ N}\cdot\text{m}$  时转矩脉动系数对比

Fig. 18 Comparison of torque ripple coefficient at  $T_L=30 \text{ N}\cdot\text{m}$

### 4 结 论

以三相 12/8 极开关磁阻电机为研究对象, 对 SRM 的 CCC 技术进行了研究。通过仿真和硬件在环实验分析, 得出以下结论:

- 1) 传统 CCC 在电感的线性上升区能获得较好的控制效果, 但在低电感区及电感转折区难以保持合成瞬时转矩的稳定。
- 2) 为了增强低电感区间的转矩输出能力, 采用了一种基于电感曲线分段, 对参考电流进行补偿的控制策略, 并在电感曲线的线性上升阶段, 采用固定频率的 PWM 波进行控制, 使得输出转矩更平滑。
- 3) 相比传统 CCC, 本文改进 CCC 能有效减小换相区间的转矩脉动。

### 参考文献 (References)

[ 1 ] ZUO S G, LIU M T, HU S L, et al. Torque ripple reduction of switched reluctance motor by optimising switch angle based on analytical modelling[J]. *IET Electric Power Applications*, 2020, 14(8): 1488-1495.

[ 2 ] 匡斯建, 张小平, 刘苹, 等. 基于相电感非饱和区定位的开关磁阻电机无位置传感器控制方法[J]. *电工技术学报*, 2020, 35(20): 4296-4305.

KUANG S J, ZHANG X P, LIU P, et al. Sensorless control method for switched reluctance motors based on locations of phase inductance characteristic points[J]. *Transactions of China Electrotechnical Society*, 2020, 35(20): 4296-4305(in Chinese).

[ 3 ] YE J, BILGIN B, EMADI A. An offline torque sharing function for torque ripple reduction in switched reluctance motor drives[J]. *IEEE Transactions on Energy Conversion*, 2015, 30(2): 726-735.

[ 4 ] XUE X D, CHENG K W E, HO S L. Optimization and evaluation of torque-sharing functions for torque ripple minimization in switched reluctance motor drives[J]. *IEEE Transactions on Power Electronics*, 2009, 24(9): 2076-2090.

[ 5 ] 颜宁, 曹鑫, 邓智泉. 基于全桥变换器的开关磁阻电机直接转矩控制[J]. *中国电机工程学报*, 2018, 38(S1): 235-242.

YAN N, CAO X, DENG Z Q. Direct torque control of switched reluctance motors based on full-bridge converter[J]. *Proceedings of the CSEE*, 2018, 38(S1): 235-242 (in Chinese).

[ 6 ] YAN N, CAO X, DENG Z Q. Direct torque control for switched reluctance motor to obtain high torque –ampere ratio[J]. *IEEE Transactions on Industrial Electronics*, 2019, 66(7): 5144-5152.

[ 7 ] 程勇, 曹晓晓, 张怡龙. 开关磁阻电机滞环-脉宽调制直接瞬时转矩控制[J]. *电机与控制学报*, 2020, 24(8): 74-82.

CHENG Y, CAO X X, ZHANG Y L. Hysteresis-PWM direct instantaneous torque control of switched reluctance motor[J]. *Electric Machines and Control*, 2020, 24(8): 74-82(in Chinese).

[ 8 ] LIANG J, LEE D H, AHN J W. Direct instantaneous torque control of switched reluctance machines using 4-level converters[J]. *IET Electric Power Applications*, 2009, 3(4): 313-323.

[ 9 ] 朱叶盛, 章国宝, 黄永明. 基于PWM的开关磁阻电机直接瞬时转矩控制[J]. *电工技术学报*, 2017, 32(7): 31-39.

ZHU Y S, ZHANG G B, HUANG Y M. PWM-based direct instantaneous torque control of switched reluctance machine[J]. *Transactions of China Electrotechnical Society*, 2017, 32(7): 31-39(in Chinese).

[ 10 ] 王喜莲, 许振亮. 基于PI参数自适应的开关磁阻电机调速控制研究[J]. *中国电机工程学报*, 2015, 35(16): 4215-4223.

WANG X L, XU Z L. Speed regulation control of switched reluctance motors based on PI parameter self-adaptation[J]. *Proceedings of the CSEE*, 2015, 35(16): 4215-4223(in Chinese).

[ 11 ] DÚBRAVKA P, RAFAJDUS P, MAKYŠ P, et al. Control of switched reluctance motor by current profiling under normal and open phase operating condition[J]. *IET Electric Power Applications*, 2017, 11(4): 548-556.

[ 12 ] 蔡辉, 王辉, 李孟秋, 等. 基于预测电流控制算法的开关磁阻电机转矩脉动抑制策略[J]. *中国电机工程学报*, 2019, 39(16): 4899-4909.

CAI H, WANG H, LI M Q, et al. Predictive current control of switched reluctance motor for torque ripples minimization[J]. *Proceedings of the CSEE*, 2019, 39(16): 4899-4909(in Chinese).

[ 13 ] LIN Z Y, REAY D S, WILLIAMS B W, et al. Torque ripple reduction in switched reluctance motor drives using B-spline neural networks[J]. *IEEE Transactions on Industry Applications*, 2006, 42(6): 1445-1453.

[ 14 ] LI C H, WANG G F, LI Y, et al. Robust adaptive neural network control for switched reluctance motor drives[J]. *Automatika*, 2018, 59(1): 24-34.

- [15] 潘再平, 罗星宝. 基于迭代学习控制的开关磁阻电机转矩脉动抑制[J]. *电工技术学报*, 2010, 25(7): 51-55.  
PAN Z P, LUO X B. Torque ripple minimization of switched reluctance motor based on iterative learning control[J]. *Transactions of China Electrotechnical Society*, 2010, 25(7): 51-55 (in Chinese).
- [16] XU Y Z, ZHONG R, CHEN L, et al. Analytical method to optimise turn-on angle and turn-off angle for switched reluctance motor drives[J]. *IET Electric Power Applications*, 2012, 6(9): 593-603.

## Current chopping control strategy of switched reluctance motor based on inductance characteristics

CHEN Yue<sup>1</sup>, JIANG Qilong<sup>1,\*</sup>, WANG Jinsuo<sup>1</sup>, YAO Weifeng<sup>2</sup>

(1. School of Electrical Engineering, Southwest Jiaotong University, Chengdu 611756, China;

2. School of Mechanical and Electrical Engineering, Shenzhen Polytechnic, Shenzhen 518055, China)

**Abstract:** Switched reluctance motors (SRMs) often use current chopping control (CCC) when running below the base speed. In the traditional control, however, the dynamic tracking ability of current is weak, the torque ripple in the commutation area is large, and the switching frequency of power devices is not fixed. A soft chopping current control strategy is proposed based on reference current compensation and the interval segmented according to the characteristics of inductance curve changes. In the low-inductance section, the reference current is compensated according to motor speed and load to improve the torque output capability and dynamic response capability of the phase winding during commutation. In the linear rising stage of the inductance curve, a fixed frequency pulse width modulation (PWM) wave is used to control the output torque to make it smoother. Then this study sets up a three-phase 12/8 pole switched reluctance motor simulation model and a semi-physical experiment platform. Considering the operating conditions of the motor with different speeds and loads, the torque ripple index is chosen for comparison. Simulation and experimental results show that the control method proposed in this paper can effectively reduce the torque ripple of SRMS and improve the performance of the motor.

**Keywords:** switched reluctance motor; current chopping control; torque ripple; reference current; hardware in the loop

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0271

# GNSS 外辐射源雷达低慢小目标探测概率

苗铎, 杨东凯, 许志超, 王峰\*, 吴世玉

(北京航空航天大学 电子信息工程学院, 北京 100191)

**摘 要:** 尽管全球导航卫星系统 (GNSS) 外辐射源雷达具有信号源广泛、覆盖率高、容易进行时间同步等特点, 受到了国内外研究机构的广泛关注, 但由于卫星位置变化和单颗卫星的目标探测性能有限, 难以满足实际探测需求。根据几何构型给出 GNSS 外辐射源雷达双基地角计算方式, 仿真研究双基地角与目标雷达散射截面 (RCS) 的关系, 分析探测时间与目标最大探测距离的关系, 得到目标探测概率的理论表达式, 并据此评估基于 GPS L5 信号的外辐射源雷达在单星、多源融合及前后向协同探测模式的目标探测概率。仿真结果表明: 单星前向和后向探测模式的有效探测时间覆盖率不足 1%, 采用前后向协同及多源融合的探测方式, 可有效提升 GNSS 外辐射源雷达的目标探测性能至 25%; 通过采用连续扫描检测的方式实时改变接收天线的照射方向进行目标探测, 在前后向协同的多源融合探测模式下, 有效探测时间覆盖率达到 98.96%, 基本满足全天时有有效探测需求。

**关键词:** 全球导航卫星系统; 外辐射源雷达; 目标探测; 探测概率; 多源融合

**中图分类号:** V19; TN958.97

**文献标志码:** A **文章编号:** 1001-5965(2023)03-0657-08

近年来, 随着科学技术的不断发展, 传统单基地雷达面临着电磁干扰、低空/超低空突防、高速反辐射导弹、隐身飞机等威胁, 无法有效应对新技术和新方案的挑战<sup>[1]</sup>。外辐射源雷达作为一种双/多基地无源雷达, 具有隐蔽性强、成本低、可用信号源丰富等优点, 已成为雷达领域的研究热点之一。20 世纪 90 年代末, 美国洛克希德·马丁公司以调频广播 (frequency modulation, FM) 信号作为外辐射源, 推出了成熟适用的商用被动雷达“沉默哨兵”<sup>[2-3]</sup>。目前外辐射源雷达所使用的非合作外辐射源主要有 FM 信号<sup>[4]</sup>、数字电视信号<sup>[5]</sup>、全球移动通信信号<sup>[6]</sup>及导航卫星信号<sup>[7]</sup>等。在各类外辐射源中, 全球导航卫星系统 (global navigation satellite system, GNSS) 因信号源广泛、覆盖率高、易进行时间同步等特点, 使得针对该技术的研究日益增多。

双基地雷达的几何构型复杂多变, 主要分后向和前向探测模式。后向探测模式与传统单基地雷达探测模式类似, 利用反射回波中携带的目标时延及多普勒信息检测目标。相比于后向探测模式, 前向散射雷达 (forward scattering radar, FSR) 可以提高目标的雷达散射截面 (radar cross section, RCS), 在反隐身探测领域具有重要意义, 成为了一项新的研究方向<sup>[8]</sup>。英国伯明翰大学微波集成系统实验室 (microwave integrated systems laboratory, MISL) 近年对前向散射雷达开展了一系列研究, 完善了地面、海洋及空中目标探测的理论基础<sup>[9-15]</sup>。北京理工大学自 2006 年起与伯明翰大学 MISL 开展联合研究, 在前向散射雷达的杂波分析、运动目标参数估计及多目标分辨特性等问题上取得了一定的研究进展<sup>[16-18]</sup>。尽管基于外辐射源雷达目标探测的基

收稿日期: 2021-05-25; 录用日期: 2021-08-22; 网络出版时间: 2021-09-14 15:18

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210914.1234.004.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210914.1234.004.html)

基金项目: 中国博士后创新人才支持计划 (BX20200039)

\* 通信作者. E-mail: [wangf.19@163.com](mailto:wangf.19@163.com)

**引用格式:** 苗铎, 杨东凯, 许志超, 等. GNSS 外辐射源雷达低慢小目标探测概率 [J]. 北京航空航天大学学报, 2023, 49 (3): 657-664.

MIAO D, YANG D K, XU Z C, et al. Low-altitude, slow speed and small target detection probability of passive radar based on GNSS signals [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 657-664 (in Chinese).

基础理论已逐渐成熟,但探测模式单一且存在一定的局限性,限制了外辐射源雷达的探测性能和使用环境,尤其由于 GNSS 外辐射源雷达的信号源并非固定辐射源,使得某些时刻可用的辐射源较少,单一模式的探测难以满足实际的探测需求,漏警率较高。

本文以“低慢小”飞行器作为典型的探测目标,统计分析 GNSS 无源雷达在不同单一探测模式下目标探测概率,针对在单一模式下探测概率较低的问题,提出多辐射源融合及前后向协同探测模式,并分析协同融合的外辐射源雷达目标探测性能。

## 1 双基地角与目标雷达散射截面积的关系

### 1.1 GNSS 无源雷达双基地角

在实际探测场景中,双基地角 $\beta$ 是一个受多因素影响的复杂三维空间变量。GNSS 无源雷达的空间几何构型如图 1 所示,其中选取地面基站的接收天线作为原点  $O$ ,天线照射方向在地平面上的投影作为  $y$  轴建立三维空间坐标系,图中: $S'$  为卫星  $S$  在  $yOz$  平面的投影, $\phi_T$  为卫星与  $yOz$  平面的夹角, $R_T$  为卫星辐射源到目标的距离, $R_R$  为目标到地面基站的距离, $L$  为卫星辐射源到地面基站的距离, $\beta$  为双基地角, $\theta_T$  为卫星仰角, $\theta_R$  为地面基站接收天线在  $yOz$  平面的照射方向, $\beta'$  为双基地角 $\beta$ 在  $yOz$  平面的投影。

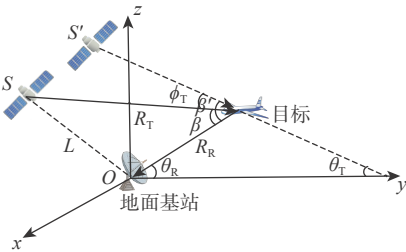


图 1 双基地雷达三维空间几何构型

Fig. 1 Three-dimensional geometric configuration of bistatic radar

根据几何关系可知

$$\cos \beta' = \frac{\mathbf{S'T} \cdot \mathbf{RT}}{\|\mathbf{S'T}\| \cdot \|\mathbf{RT}\|} \quad (1)$$

$$\cos \beta = \frac{\mathbf{ST} \cdot \mathbf{RT}}{\|\mathbf{ST}\| \cdot \|\mathbf{RT}\|} \quad (2)$$

$$\cos \phi_T = \frac{\mathbf{ST} \cdot \mathbf{S'T}}{\|\mathbf{ST}\| \cdot \|\mathbf{S'T}\|} \quad (3)$$

$$\beta' = \theta_T + \theta_R \quad (4)$$

根据式 (1)~式 (4) 可以得到三维空间下双基地角 $\beta$ 的表达式为

$$\beta = \arccos(\cos \beta' \cos \phi_T) = \arccos[\cos(\theta_T + \theta_R) \cos \phi_T] \quad (5)$$

当双基地角 $\beta < 135^\circ$ 时,卫星-目标-接收机之间构成后向散射结构;当双基地角 $\beta > 135^\circ$ 时,卫星-目标-接收机之间构成前向散射结构。

### 1.2 目标雷达散射截面积

目标 RCS 是衡量目标接收辐射源所发射的电磁波后,对入射波进行再辐射能力的物理量,目标 RCS 在后向模式中定义为

$$\sigma_{BS} = \lim_{R_d \rightarrow \infty} \left( 4\pi R_d^2 \frac{\|\mathbf{E}_s\|^2}{\|\mathbf{E}_i\|^2} \right) \quad (6)$$

式中: $R_d$  为目标到接收天线的距离, $R_d \rightarrow \infty$  保证接收天线位于远场区域; $\mathbf{E}_s$  为接收天线处的电场强度,表征接收天线处目标散射功率密度; $\mathbf{E}_i$  为目标处入射的电场强度。在前向模式中,可根据巴比涅原理通过等效理论计算目标 RCS,计算式为<sup>[19]</sup>

$$\sigma_{FS} = \frac{4\pi}{\lambda^2} \left| \iint \exp(\mathbf{j}\mathbf{k} \cdot \boldsymbol{\rho}) dA \right|^2 \quad (7)$$

式中: $\mathbf{j} = \sqrt{-1}$ ;  $\lambda$  为波长; $\mathbf{k}$  为波数矢量; $\boldsymbol{\rho}$  为不透明屏幕平面上的坐标矢量; $dA$  为微分表面积。

由于通过电磁理论精确计算目标 RCS 十分困难,因此在仿真过程中采用近似计算的方法来估计复杂目标 RCS。常用近似计算方法有物理光学法 (physical optics, PO)、几何光学法 (geometrical optics, GO)、物理绕射理论 (physical theory of diffraction, PTD)、几何绕射理论 (geometrical theory of diffraction, GTD)、矩量法 (method of moments, MOM) 等<sup>[20]</sup>。矩量法将被探测目标分割成多个小部分,用标准的矩阵代数求解麦克斯韦积分方程组,实现目标 RCS 的解算。相比于其他方法,由于该方法求解积分方程的使用限制更小,可应用在多种情况的近似计算中,因此,本文采用矩量法进行“低慢小”目标 RCS 的计算。为简化仿真计算的复杂度,选取如图 2 所示的金属圆柱体作为“低慢小”飞行器的近似模型,其中圆柱体半径  $r=0.27$  m,高度  $h=0.3$  m;入射波为频率 1 176.45 MHz 的 GPS L5 信号,沿  $z$  轴负半轴方向进行照射;双基地角 $\beta$ 取值范围为  $0^\circ \sim 360^\circ$ 。仿真过程中仅考虑入射波和散射波形成的双基地角对双基地 RCS 的影响,忽略因目标姿态角变化引起的 RCS 变化。

图 3 为仿真的 GNSS 信号左右旋极化的目标 RCS 随双基角的变化曲线。目标再辐射的 GNSS 散射信号发生了极化变化,由右旋圆极化波转变成椭圆极化波。随着双基地角 $\beta$ 的变化,椭圆极化波的轴比不断改变。在后向散射区域散射波以左旋圆极化波为主,应采用左旋天线接收,且随着双基

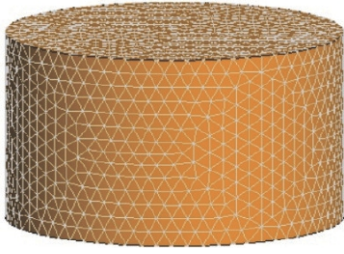
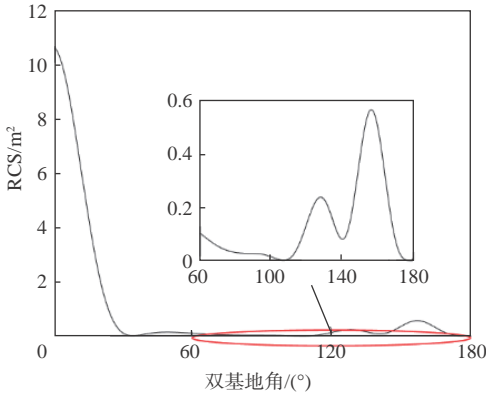
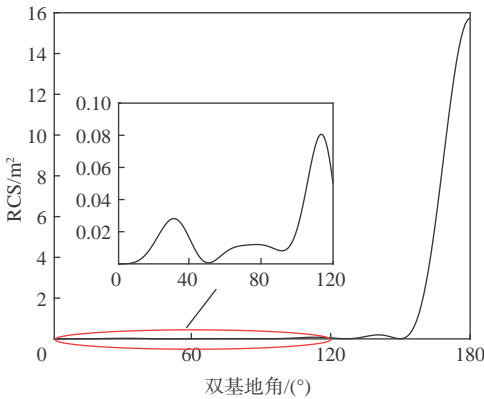


图 2 金属圆柱体模型

Fig. 2 Metal cylinder model



(a) 左旋圆极化接收天线



(b) 右旋圆极化接收天线

图 3 不同极化情况下的 RCS 仿真结果

Fig. 3 RCS simulation results under different polarization conditions

地角增大, 目标 RCS 呈下降趋势, 即在后向模式中应尽量选取双基地角较小的几何构型进行目标探测。在前向散射区域散射波以右旋圆极化波为主, 应采用右旋天线接收, 且随着双基地角的增大, 目标 RCS 呈上升趋势, 在双基地角接近 180° 时达到最大。此外, 通过对比可看出前向模式下目标 RCS 较后向模式高。

## 2 GNSS 外辐射源雷达探测性能

相比于 GPS L1 信号, GPS 卫星发射的新型民用信号 GPS L5 信号的发射功率更高、频带宽度更宽且引入了不含数据调制的导频通道, 使其更适合作为目标探测的外辐射源<sup>[21]</sup>。在不考虑传播损耗的

情况下, 地面基站接收天线所接收的信号功率为<sup>[22]</sup>

$$P_r = \frac{P_t G_r \sigma_b(\beta)}{4\pi R_{ig}^2} \quad (8)$$

式中:  $P_t$  为照射到地球表面的信号功率, 对于 GPS L5 信号, 为 -157 dBW;  $G_r$  为接收天线的增益;  $\sigma_b(\beta)$  为双基地角为  $\beta$  的目标的双基地雷达散射截面;  $R_{ig}$  为目标到地面基站接收天线的距离。假设接收信号中的噪声为高斯白噪声, 其功率  $N_r$  可表示为

$$N_r = k T_s B \quad (9)$$

式中:  $k=1.38 \times 10^{-23}$  J/K, 为玻尔兹曼常数;  $T_s$  为等效噪声温度, 通常取室温 290 K;  $B$  为信号带宽, 假设为 GPS L5 信号的信号带宽 20.46 MHz。计算可得噪声功率约为 -131 dBW。由式 (8) 和式 (9) 可计算得到接收到信号的信噪比为

$$SNR_r = \frac{P_r}{N} = \frac{P_t G_r \sigma_b(\beta)}{4\pi R_{ig}^2 N_r} \quad (10)$$

考虑相干积分时间为  $T_{int}$  和非相干累加次数为  $N$  的相干积分和非相干累加增益, 接收信号的信噪比为

$$SNR_{det} = \frac{P_t G_r \sigma_b(\beta) G_p N^{0.8}}{4\pi R_{ig}^2 N_r} \quad (11)$$

式中:  $G_p = BT_{int}$  为扩频增益,  $T_{int}$  为相干积分时间。最大探测距离与检测到目标所需最低信噪比之间的关系为

$$R_{ig,max} = \sqrt{\frac{P_t G_r \sigma_b(\beta) G_p N^{0.8}}{4\pi N_r SNR_{det,min}}} = \sqrt{\frac{P_t G_r \sigma_b(\beta) BT_{int} N^{0.8}}{4\pi N_r SNR_{det,min}}} \quad (12)$$

本文选取“低慢小”飞行器作为典型目标, 假设 RCS 为 5 m<sup>2</sup>, 接收天线增益为 25 dB。当目标检测所需最低信噪比为 8 dB 时, 目标最大检测距离与积分时间及非相干累加次数的关系如图 4 所示。从图 4 可知, 随着相干积分时间及非相干累加次数的增加, 最大探测距离不断提高。由于目标出现在探测区域的时间有限, 积分时间并不能无限增大。通常将该目标的出现时间作为最大检测时间, 以限制数据处理过程中的相干积分时间与非相干累加次数。最大检测时间可表示为相干积分时间与非相干累加次数的乘积  $T_{det} = T_{int} N$ 。受目标飞行速度、飞行姿态、接收天线波束角范围等因素约束, 通常以约 1 s 的检测时间对大多数民用交通为最佳<sup>[23]</sup>。选取 L5 信号导频通道的 Neuman-Hoffman 同步码 (NH 码) 周期 (20 ms) 作为相干积分时间, 最大处理时间为 1 s, 则非相干累加次数为 50。选取

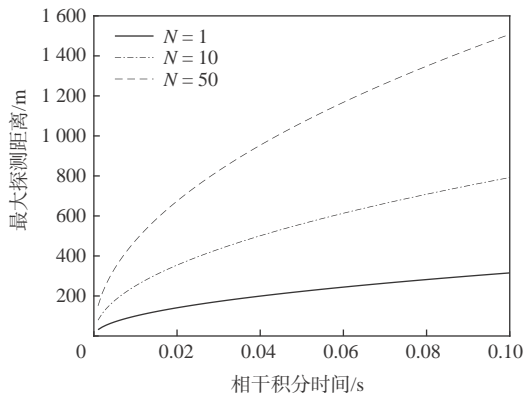


图4 最大探测距离随积分时间及非相干累加次数的变化

Fig. 4 Variation of the maximum detection distance at different integration time and incoherent accumulation times

“低慢小”飞行器作为典型目标,代入式(12)计算可得最大探测距离为674.29 m。

### 3 “低慢小”目标的探测概率

#### 3.1 目标的探测概率

假设接收机的输入信号 $r(t)$ 由目标散射的回波信号 $s(t)$ 和均值为0、方差为 $\psi^2$ 的加性高斯白噪声 $n(t)$ 组成。假设输入噪声与信号彼此独立,则二元检测表达式可表示为

$$r(t) = \begin{cases} H_0 : n(t) \\ H_1 : s(t) + n(t) \end{cases} \quad (13)$$

式中: $H_0$ 为信号中仅存在噪声; $H_1$ 为信号中同时包含目标信号和噪声。在雷达目标检测过程中,常用恒虚警率(constant false alarm rate, CFAR)检测对是否存在目标进行判决,即在保持虚警概率 $P_{fa}$ 不变的条件下,依据奈曼-皮尔逊(Neyman-Pearson, NP)准则确定使检测概率 $P_d$ 最大的门限值 $V_T$ ,且当输入信号 $r(t)$ 超过门限值 $V_T$ 时判定目标出现。如式(14)所示,虚警概率 $P_{fa}$ 定义为当接收信号中只存在噪声时,信号 $r(t)$ 的样本超过门限值 $V_T$ 的概率<sup>[24]</sup>。

$$P_{fa} = \int_{V_T}^{\infty} \frac{r}{\psi^2} \exp\left(-\frac{r^2}{2\psi^2}\right) dr = \exp\left(-\frac{V_T^2}{2\psi^2}\right) \quad (14)$$

在虚警概率一定的条件下门限值 $V_T$ 可以表示为

$$V_T = \sqrt{2\psi^2 \ln\left(\frac{1}{P_{fa}}\right)} \quad (15)$$

检测概率为噪声与信号共同存在的情况下,信号样本超过门限值 $V_T$ 的概率。对于复杂目标通常根据目标的起伏特性划分为四类Swerling模型计算检测概率。由于“低慢小”飞行器可认为是由很多RCS值较小的散射体组成,因此,可采用Swerling I型起伏模型描述其探测概率。Swerling I型目标的检测概率可表示为

$$P_d = 1 - \Gamma_1(V_T, n_p - 1) + \left(1 + \frac{1}{n_p \text{SNR}}\right)^{n_p - 1} \cdot \Gamma_1\left(\frac{V_T}{1 + \frac{1}{n_p \text{SNR}}}, n_p - 1\right) e^{-V_T/(1 + n_p \text{SNR})} \quad (16)$$

式中: $n_p$ 为非相干累加次数; $\Gamma_1(x, N)$ 为不完全的伽马函数,可近似表示为

$$\Gamma_1(x, N) = \int_0^x \frac{e^{-v} v^{N-1}}{(N-1)!} dv \quad (17)$$

式中: $v$ 为不完全伽马函数的积分变量。

图5为虚警概率 $P_{fa} = 1 \times 10^{-6}$ ,非相干累加次数为50时,目标检测概率随SNR的变化。从图5中可看出,在累加次数一定的情况下,目标检测概率与接收信号SNR成正比例关系。通常认为当目标检测概率达到90%时可实现目标的有效探测<sup>[24]</sup>,此时所需的接收信号的SNR约为8 dB。

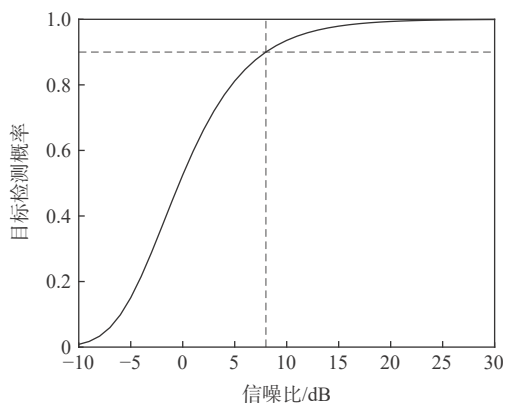


图5 目标检测概率随SNR的变化

Fig. 5 Variation of target detection probability with SNR

实际探测过程中,信噪比难以被直接观测,由式(11)可知,在一定条件下SNR是目标最大探测距离的函数。假定所选探测目标RCS为 $5 \text{ m}^2$ ,接收天线增益为25 dB,虚警概率 $P_{fa} = 1 \times 10^{-6}$ ,图6为目标检测概率随最大探测距离的变化。随着最大探测距离的增加,目标的检测概率不断下降,当检测概率低于90%时可以认为在该距离下无法有效检测到目标。对比不同累加情况下的变化曲线,不进行累加情况下,最大探测距离不足100 m;单独进行20 ms的相干积分和50次的非相干累加时,最大探测距离可以提升到200 m左右;2种累加方法共同作用时,目标的最大探测距离超过600 m,探测性能得到了有效提升。

另一方面,将式(10)代入式(16)可知,目标检测概率是关于RCS的函数。当其他条件固定时,目标检测的概率将随双基地角的变化而改变,即几何构型直接影响到双基地雷达的目标检测性能。以图2所示仿真模型作为“低慢小”目标,接收天线增益为25 dB,相干积分时间为20 ms,非相干累加次

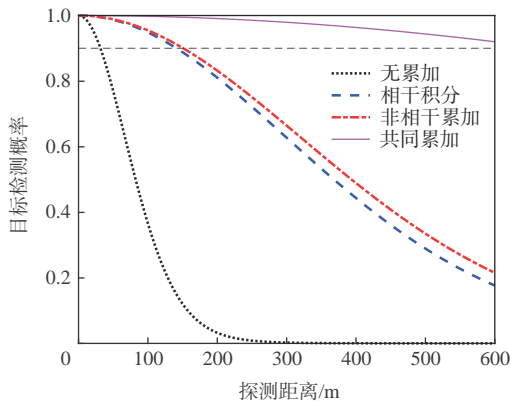


图 6 目标检测概率随探测距离的变化

Fig. 6 Variation of target detection probability with detection distance

数为 50, 计算不同目标探测距离下目标检测概率随双基地角的变化情况, 结果如图 7 所示。随着探测距离的提升, 目标探测概率减小。当目标探测距离相同时, 检测概率随双基地角改变呈现出先下降后上升的变化趋势。在后向散射情况下, 较小的双基地角可实现对目标的有效探测, 随着双基地角的提高, 检测概率逐渐减小; 当形成前向散射的几何构型后, 目标检测概率随着双基地角的增加呈递增趋势, 在双基地角为 180°时取得最大值。当目标探测距离为 500 m 时, 后向散射模式满足有效探测的双基地角范围为 0°~20°, 前向散射情况下对应的双基地角范围是 161°~180°, 即约 21.67% 的双基地角条件可满足有效探测需求, 实际探测中需要准确配置几何构型以提高雷达探测性能。

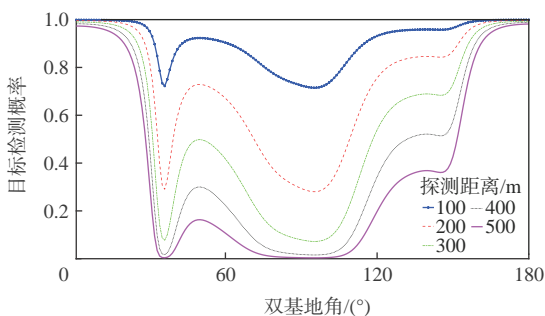


图 7 目标检测概率随双基地角的变化

Fig. 7 Variation of target detection probability under different bistatic angle

### 3.2 单颗卫星的有效探测时间覆盖率

为模拟卫星在一天的双基地角变化情况, 需设定模拟实验地点及实验日期, 通过仿真实验区域的坐标信息, 结合精密星历计算得到卫星在任意时刻的双基地角大小。截止到 2020 年 12 月 31 日, 发射 GPS L5 信号的在轨卫星共有 16 颗。选取模拟观测地点为北京首都国际机场, 经、纬和高分别

为 116.6143 °E、40.0554 °N 和 50 m。模拟观测日期为 2021 年 3 月 16 日。根据 GPS 官方提供的精密星历计算每颗卫星在 24 h 内的双基角变化。选取模拟探测目标为图 2 所示的仿真目标, 接收天线增益为 25 dB, 相干积分时间为 20 ms, 非相干累加次数为 50, 目标探测距离为 500 m。一天内前后向散射模式下单颗卫星的目标探测概率统计如图 8 所示。统计结果显示在后向散射模式下单颗卫星能实现有效探测的情况仅占全天时间的 0.26%, 而前向散射模式下有效探测的时间覆盖率为 0.52%。这说明采用单颗卫星进行目标探测时, 一天内仅有极少数几何构型满足实际探测需求, 须严格设计接收天线的位置及朝向获得最优探测效果。在后向散射模式下大部分时间内 L5 卫星同目标和接收天线构成的双基地角较大, 使得目标双基地 RCS 较小, 不足以实现有效探测。前向散射模式对几何构型要求更严格, 据统计在单星情况下全天仅有 4.95% 时间可满足前向散射模式双基地角要求, 无法充分发挥前向散射高目标 RCS 优势, 使得整体探测概率较低。

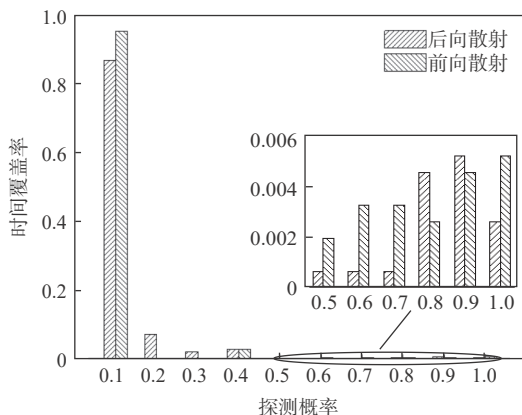


图 8 单颗卫星的目标探测概率

Fig. 8 Target detection probability of a single satellite

### 3.3 多辐射源融合的有效探测时间覆盖率

由 3.2 节结果可知单颗卫星的目标检测性能较差, 无法满足“低慢小”目标的探测需求。多辐射源融合是提升目标探测性能的可靠方案。在目标探测过程中将每颗卫星看作独立的辐射源进行探测, 不同卫星的信号可通过其伪随机码 (pseudo random noise, PRN) 进行区分, 多辐射源的探测过程可假设为一个独立分布的随机过程。由于每颗星的位置不同, 同一时刻其构成的双基地角不同, 导致不同辐射源下目标的双基地 RCS 不同, 探测结果存在差异。多辐射源融合指的是将多颗卫星单独的探测结果进行叠加, 取并集得到探测结果, 即对探测结果进行了逻辑或运算, 其探测概率可表示为

$$P_d = 1 - \prod_i (1 - P_{di}) \quad (18)$$

式中:  $P_{di}$ 为单颗卫星的目标检测概率。在后向散射模式下进行多源融合的目标探测,其探测概率统计如图9所示。相比于图8的单星检测结果,多辐射源融合后的目标探测概率明显得到提升。多源融合的后向散射模式在全天内有8.33%的时间可实现有效探测,而前向散射模式的有效探测的覆盖率可达到11.46%,极大地提高了目标检测概率,有效提升了目标探测性能。

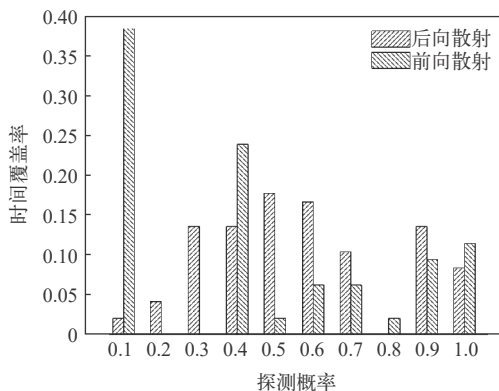


图9 多源融合的目标探测概率

Fig. 9 Target detection probability with multi-satellites

### 3.4 前后向协同的有效探测时间覆盖率

在单一探测方式下后向模式由于目标本身的双基地 RCS 较小,探测概率受到了较大程度限制。前向模式的目标探测概率受限于前向散射构型的覆盖率。如图9所示,多源融合后一天内可以形成前向散射结构的时间仅占61%,使得有效探测的时间覆盖率仅达到11.46%,整体探测性能有限。前后向协同探测方式可充分发挥GNSS外辐射源雷达的探测性能。前后向协同的探测模式的多源融合目标探测概率统计如图10所示。利用前后向协同的多源融合探测方式,目标探测性能进一步提升,

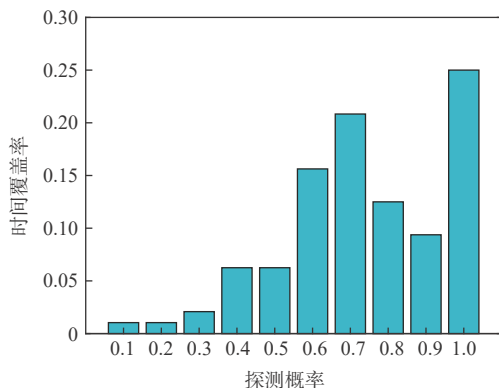


图10 前后向协同情况下多源融合的目标探测概率

Fig. 10 Target detection probability under backscattering and forward scattering condition with multi-satellites

有效探测的时间覆盖率达25%,极大程度减小了几何构型对目标探测场景的制约,使GNSS外辐射源雷达的性能得到充分发挥。

### 3.5 连续扫描式检测的有效探测时间覆盖率

根据3.2节~3.4节的结论可以看出,当天线照射方向固定时,进行有效探测的时间覆盖率并不理想。为了实现目标的连续检测,可以根据卫星所在的方位,实时改变天线照射方向以提高目标的检测概率,即在后向散射模式下使卫星与天线照射方向所成的方位角为 $0^\circ$ ,前向散射模式下使方位角为 $180^\circ$ 。图11为单星单基情况下,采用连续检测方式对全天的目标探测概率统计结果。在该情况下,后向散射模式有效探测的时间覆盖率提高到7.23%,前向散射模式下时间覆盖率提高到8.27%。可以看出实时改变天线照射方向确保在最优几何构型下进行目标探测时,检测性能得到了有效提升。

图12为在多源融合的情况下实时改变天线位置进行连续扫描检测时有效探测的时间覆盖率。结果显示在后向散射模式下多源融合

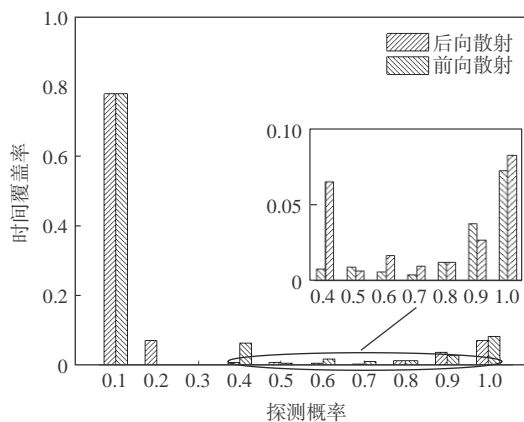


图11 连续检测情况下单颗卫星的目标探测概率

Fig. 11 Target detection probability under continuous detection of a single satellite

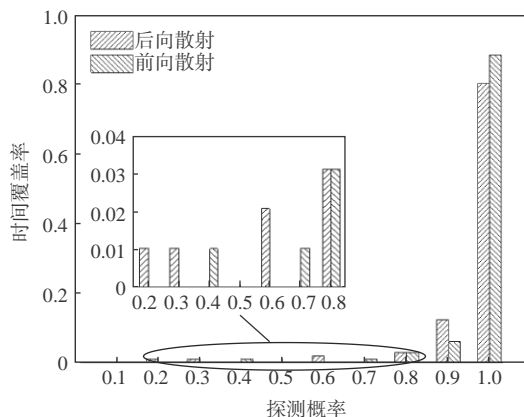


图12 连续检测情况下多源融合的目标探测概率

Fig. 12 Target detection probability under continuous detection with multi-satellites

检测目标的有效探测时间覆盖率可达80.21%,前向模式下达88.54%;通过前后向协同有效探测的时间覆盖率可达98.96%,基本实现了全天时的有效探测。值得注意的是连续扫描检测实现多源融合需要依靠多接收天线协同工作,无疑增加了系统和算法复杂度。

## 4 结论

1) 后向散射模式以左旋圆极化散射波为主,而前向以右旋圆极化波为主。在建立的近似模型基础上,利用Swerling起伏模型,计算了不同情况下GNSS外辐射源雷达的目标检测概率,研究了多辐射源融合及前后向协同探测对于目标检测概率的提升作用。单星情况下目标检测性能较低,有效探测的时间覆盖率不足1%;通过采用多星融合及前后向协同可将有效探测的时间覆盖率提高至25%。通过卫星位置对天线照射方向进行实时调整,使系统以连续扫描的模式进行工作,单星及多源融合下有效探测的时间覆盖率得到了极大提升,在前后向协同的多源融合探测模式下,有效探测的时间覆盖率可达98.96%,基本可满足全天时检测的需求。

2) 由于实际目标的双基地RCS计算复杂,除外辐射源的照射方向引起的双基地角变化之外,目标的具体形状、表面材料及运动姿态均会影响双基地RCS。

本文旨在研究多源融合对目标探测概率的提升效果,仅采取简易模型进行了近似分析。在后续研究中将针对具体目标进行更精确仿真建模及分析。此外本文选用的卫星辐射源为GPS L5信号,仅16颗卫星在轨。随着GPS现代化和北斗-3系统的全面建成,在轨高带宽卫星信号的增大,GNSS外辐射源的目标探测性能将进一步提升。

## 参考文献 (References)

- [1] 李刚. 新体制雷达及其关键技术[J]. 电子技术与软件工程, 2019(15): 60-62.
- LI G. New system radar and its key technology[J]. Electronic Technology & Software Engineering, 2019(15): 60-62(in Chinese).
- [2] 朱庆明, 吴曼青. 一种新型无源探测与跟踪雷达系统——“沉默哨兵”[J]. 现代电子, 2000(1): 1-6.
- ZHU Q M, WU M Q. Silent sentry: A novel passive detecting and tracking radar system[J]. Modern Electronics, 2000(1): 1-6(in Chinese).
- [3] 万显荣, 易建新, 占伟杰, 等. 基于多照射源的被动雷达研究进展与发展趋势[J]. 雷达学报, 2020, 9(6): 939-958.
- WAN X R, YI J X, ZHAN W J, et al. Research progress and development trend of the multi-illuminator-based passive radar[J]. Journal of Radars, 2020, 9(6): 939-958(in Chinese).
- [4] ZAIMBASHI A. Broadband target detection algorithm in FM-based passive bistatic radar systems[J]. IET Radar, Sonar & Navigation, 2016, 10(8): 1485-1499.
- [5] RAJA ABDULLAH R S A, ALHAJI MUSA S, ABDUL RASHID N E, et al. Passive forward-scattering radar using digital video broadcasting satellite signal for drone detection[J]. Remote Sensing, 2020, 12(18): 3075.
- [6] ZEMMARI R, DAUN M, FELDMANN M, et al. Maritime surveillance with GSM passive radar: detection and tracking of small agile targets[C]//International Radar Symposium. Piscataway: IEEE Press, 2013: 245-251.
- [7] ILIOUDIS C, CLEMENTE C, SORAGHAN J. GNSS-based passive UAV monitoring: A feasibility study[J]. IET Radar, Sonar & Navigation, 2020, 14(4): 516-524.
- [8] 胡程, 刘长江, 曾涛. 双基地前向散射雷达探测与成像[J]. 雷达学报, 2016, 5(3): 229-243.
- HU C, LIU C J, ZENG T. Bistatic forward scattering radar detection and imaging[J]. Journal of Radars, 2016, 5(3): 229-243(in Chinese).
- [9] CHERNIAKOV M, ABDULLAH R S A R, JANČOVIČ P, et al. Automatic ground target classification using forward scattering radar[J]. IEE Proceedings: Radar, Sonar and Navigation, 2006, 153(5): 427-437.
- [10] SIZOV V, CHERNIAKOV M, ANTONIOU M. Forward scattering radar power budget analysis for ground targets[J]. IET Radar, Sonar & Navigation, 2007, 1(6): 437-446.
- [11] HRISTOV S, DANIEL L, HOARE E, et al. Target shadow profile reconstruction in ground-based forward scatter radar[C]//IEEE Radar Conference. Piscataway: IEEE Press, 2015: 846-851.
- [12] GASHINOVA M, DANIEL L, SIZOV V, et al. Phenomenology of Doppler forward scatter radar for surface targets observation[J]. IET Radar, Sonar & Navigation, 2013, 7(4): 422-432.
- [13] MARRA M, DE LUCA A, HRISTOV S, et al. New algorithm for signal detection in passive FSR[C]//IEEE Radar Conference. Piscataway: IEEE Press, 2015: 218-223.
- [14] CONTU M, DE LUCA A, HRISTOV S, et al. Passive multifrequency forward-scatter radar measurements of airborne targets using broadcasting signals[J]. IEEE Transactions on Aerospace and Electronic Systems, 2017, 53(3): 1067-1087.
- [15] DE LUCA A, DANIEL L, GASHINOVA M, et al. Target parameter estimation in moving transmitter moving receiver forward scatter radar[C]//International Radar Symposium. Piscataway: IEEE Press, 2017: 1-7.
- [16] 胡程, 龙腾, 曾涛, 等. 前向散射雷地杂波物理建模及频谱扩展分析[J]. 中国科学:信息科学, 2010, 40(12): 1646-1659.
- HU C, LONG T, ZENG T, et al. Physical modeling and spectrum spread analysis of ground clutter in forward scattering radar[J]. Scientia Sinica(Informationis), 2010, 40(12): 1646-1659(in Chinese).
- [17] 胡程, 刘长江, 曾涛, 等. 双基地前向散射雷地杂波分析与模拟方法[J]. 信号处理, 2013, 29(3): 293-303.
- HU C, LIU C J, ZENG T, et al. Statistical analysis and simulation method of forward scattering clutter in bistatic radar[J]. Journal of Signal Processing, 2013, 29(3): 293-303(in Chinese).
- [18] 李文海, 董锡超, 胡程. 前向散射雷地海面目标探测信号建模与分析[J]. 信号处理, 2019, 35(6): 994-1001.

- LI W H, DONG X C, HU C. Signal modeling and analysis of forward scatter radar in sea surface target detection[J]. *Journal of Signal Processing*, 2019, 35(6): 994-1001 (in Chinese).
- [19] GLASER J I. Bistatic RCS of complex objects near forward scatter[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 1985, 21(1): 70-78.
- [20] 马焱. 雷达散射截面积的特征提取算法研究与电路设计[D]. 西安: 西安电子科技大学, 2017: 8-10.
- MA B. Algorithm research and circuit design on feature extraction of radar cross section[D]. Xi'an: Xidian University, 2017: 8-10(in Chinese).
- [21] 伍岳, 罗和平, 邱蕾. L5载波的信号质量分析[J]. *测绘工程*, 2014, 23(12): 8-11.
- WU Y, LUO H P, QIU L. Quality analysis of GPS L5 signal[J]. *Engineering of Surveying and Mapping*, 2014, 23(12): 8-11(in Chinese).
- [22] BEHAR V, KABAKCHIEV C. Detectability of air targets using bistatic radar based on GPS L5 signals[C]//International Radar Symposium. Piscataway: IEEE Press, 2011: 212-217.
- [23] MIKHAIL C. 双基地雷达: 一项新兴技术[M]. 陈筠力, 译. 北京: 中国宇航出版社, 2015: 385.
- MIKHAIL C. Bistatic radar: Emerging technology[M]. CHEN J L, translated. Beijing: China Astronautic Publishing House, 2015: 385(in Chinese).
- [24] BASSEM R M, ATEF Z E. 雷达系统设计MATLAB仿真[M]. 朱国富, 黄晓涛, 黎向阳, 等译. 北京: 电子工业出版社, 2016: 49-65.
- BASSEM R M, ATEF Z E. MATLAB simulations for radar systems design[M]. ZHU G F, HUANG X T, LI X Y, et al, translated. Beijing: Publishing House of Electronics Industry, 2016: 49-65(in Chinese).

## Low-altitude, slow speed and small target detection probability of passive radar based on GNSS signals

MIAO Duo, YANG Dongkai, XU Zhichao, WANG Feng\*, WU Shiyu

(School of Electronic and Information Engineering, Beihang University, Beijing 100191, China)

**Abstract:** Although passive radar based on global navigation satellite system (GNSS) signals has received domestic and international attention owing to its advantages such as the availability of multiple satellites, global coverage, and convenience of time synchronization, it is difficult to meet the actual detection requirements due to the movement of satellites and the limited target detection performance of a single satellite. The bistatic angle calculation procedure is described in accordance with the geometric configuration, and the theoretical expression of target detection probability is derived after analyzing the relationship between target radar cross section (RCS) and bistatic angle, and the connection between detection time and target maximum detection range. On these grounds, the target detection probability of passive radar based on GPS L5 signal under the detection pattern of single-satellite, multi-satellites fusion and forward-backward cooperation is evaluated. The simulation results show that the effective probability of detection under single-satellite forward or backward mode is less than 1%, which has an effective promotion to 25% under the detection pattern of multi-satellites fusion with forward-backward cooperation. Furthermore, the continuous detection method is adopted to change the irradiation direction of the receiving antenna to detect the target. In the forward and backward cooperative multi-source fusion detection mode, the effective detection time coverage is up to 98.96%, and the all-time effective detection is basically realized.

**Keywords:** global navigation satellite system; passive radar; target detection; target detection probability; multiple illuminators fusion

<http://bhxb.buaa.edu.cn>     [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0272

# 大行程柔性微定位平台的伴生转动分析

孟刚<sup>1</sup>, 黄河<sup>1</sup>, 吴伟官<sup>1</sup>, 居勇健<sup>1</sup>, 曹毅<sup>1,2,\*</sup>

(1. 江南大学机械工程学院, 无锡 214122; 2. 江苏省食品先进制造装备技术重点实验室, 无锡 214122)

**摘 要:** 大行程柔性微定位平台在运动过程中不可避免地产生伴生转动现象, 并对其定位精度造成消极影响。为降低伴生转动对平台定位精度的影响, 提出一种基于柔性杆的三移一转(3-PPPR)型大行程柔性微定位平台, 基于线弹性梁理论模型并考虑柔性杆轴向形变, 对两移一转(PPR)柔性运动副伴生转角进行了理论建模, 并基于此完成了对所提平台在单轴、双轴及三轴驱动时产生伴生转角的理论分析; 再采用有限元分析对理论模型进行验证。最后探究了柔性杆尺寸参数与平台伴生转角之间的灵敏度关系, 为所提平台性能提升奠定了基础, 并据此提出了改善所提平台运动性能的优化方案。结果表明: 3种驱动条件下平台伴生转角理论值与仿真值最大相对误差为2.46%。

**关键词:** 柔性微定位平台; 伴生转动; 多轴驱动; 有限元仿真; 灵敏度分析

**中图分类号:** V414.5; TH122

**文献标志码:** A     **文章编号:** 1001-5965(2023)03-0665-09

柔性机构是一类依靠材料弹性形变实现力和运动传递的新型机构<sup>[1-2]</sup>, 具有无间隙、无摩擦磨损、定位精度高、可靠性强、适应性广等优点<sup>[3]</sup>。柔性微定位平台<sup>[4-5]</sup>是一类典型的能实现稳定、精确定位的柔性机构, 是实现精密定位的主要载体。因此, 被广泛应用于航空航天<sup>[6-7]</sup>、微机电系统<sup>[8]</sup>、生物医药<sup>[9]</sup>及精密工程<sup>[10]</sup>等领域。

随着精密定位技术的迅速发展, 对柔性微定位平台的性能提出了更高的要求, 兼具大行程、高定位精度及良好动态特性已成为微定位平台设计的基本准则<sup>[11]</sup>。目前, 对具有大行程的柔性微定位平台研究虽已取得一定成果, 但运动过程中产生的伴生转动对平台定位精度造成的不利影响, 导致其难以满足较高水平的定位需求, 因此, 对大行程柔性微定位平台伴生转动的研究是实现其精密定位的关键所在。牟新明等<sup>[12]</sup>提出了对平行簧片机构伴生转动的理论建模方法, 有效降低了机构的导向误差; Luo和Liu<sup>[13]</sup>结合伪刚体模型对大行程分布柔

度式导向机构的伴生转角进行了理论分析, 并基于此进一步提高了其伴生位移的理论精度; Hao等<sup>[14-15]</sup>采用分布柔度式柔性杆设计了一种空间平动柔性微定位平台, 并对其伴生转角、耦合位移进行分析, 并基于刚度中心重合的结构优化方法提出了一种伴生转动最小化的柔性微定位平台; Su等<sup>[16]</sup>通过对柔性机构寄生运动的研究, 提出了基于瞬心的伴生转动分析方法; 此外, Awatar等<sup>[17]</sup>基于约束设计法, 提出了一种大行程柔性微定位平台, 并对其伴生转动进行了有限元仿真分析; Lin等<sup>[18]</sup>基于对称结构提出了一种低伴生转动的微定位平台, 并完成了相关实验测定。

综合文献<sup>[12-18]</sup>不难发现, 目前对伴生转动的研究多集中于对基本运动副的理论建模、对平台伴生转动的定性分析及仿真实验测定, 而对微定位平台整体伴生转动的系统性理论分析相对欠缺。基于此, 本文针对大行程柔性微定位平台伴生转动问题进行深入研究。首先, 提出一种基于柔性杆的大

收稿日期: 2021-05-25; 录用日期: 2021-08-20; 网络出版时间: 2021-08-30 14:49

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210827.1522.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210827.1522.001.html)

基金项目: 江苏省“六大人才高峰”计划(ZBZZ-012); 高等学校学科创新引智计划(B18027); 江苏省研究生创新计划(KYCX20-1925)

\*通信作者. E-mail: [caoyi@jiangnan.edu.cn](mailto:caoyi@jiangnan.edu.cn)

**引用格式:** 孟刚, 黄河, 吴伟官, 等. 大行程柔性微定位平台的伴生转动分析[J]. 北京航空航天大学学报, 2023, 49(3): 665-673.

MENG G, HUANG H, WU W G, et al. Parasitic rotation of large stroke compliant micro-positioning platform[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(3): 665-673 (in Chinese).

行程柔性微定位平台结构;其次,基于线弹性梁理论模型建立了柔性运动副伴生转角的理论模型;然后,分别对所提平台在单轴、双轴及三轴驱动条件下伴生转动进行理论分析,并基于有限元仿真验证了理论模型的正确性;最后,探究了柔性杆结构参数与平台伴生转动之间的灵敏度关系,并据此提出了改善所提平台运动性能的优化方案。

## 1 微定位平台的结构设计

柔性微定位平台与刚性并联机构类似,为保证其具有良好动态特性,常采用并联式结构<sup>[19]</sup>,同时采用分布柔度式<sup>[20]</sup>柔性单元满足平台大行程的需求。为便于实现高精度定位,平台应具有良好的解耦特性<sup>[21]</sup>及较低的伴生转动。

图1为具有两移一转(PPR)自由度的柔性运动副,其中P(prismatic joint)代表移动副,R(revolute joint)代表转动副,其结构为2组四杆结构通过反向串联连接组成,该连接方式在增加运动行程的同时,抵消了2组四杆结构运动时产生的寄生运动,提高了平台运动的解耦特性。图中: $L$ 为连接模块十字型杆长度, $D$ 为刚性连接模块高度, $s$ 为外四杆结构自由端刚性块高度, $l$ 为柔性杆长度, $t$ 为柔性杆厚度, $d$ 为外四杆结构柔性杆杆间距。基于自由度与约束空间拓扑法<sup>[22-23]</sup>及螺旋理论<sup>[24]</sup>,其运动自由度可表示为

$$\{x, y, \theta_z\}_{in} \cup \{x, y, \theta_z\}_{out} = \{x, y, \theta_z\} \quad (1)$$

式中: $x$ 、 $y$ 为运动副沿 $x$ 、 $y$ 轴的移动自由度; $\theta_z$ 为运动副绕 $z$ 轴的转动自由度。

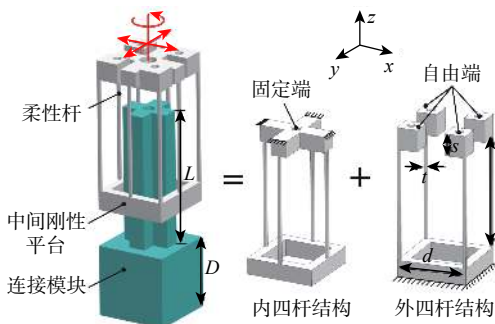


图1 PPR型柔性运动副

Fig. 1 PPR compliant kinematic joint

基于图1的PPR型柔性运动副所提的三移一转(3-PPPR)型柔性支链结构如图2所示,4个相同的两移一转型柔性运动副并联连接组成其驱动副,并在正交方向通过与被动副串联连接,组成支链整体。支链驱动副及其整体运动自由度可分别表示为

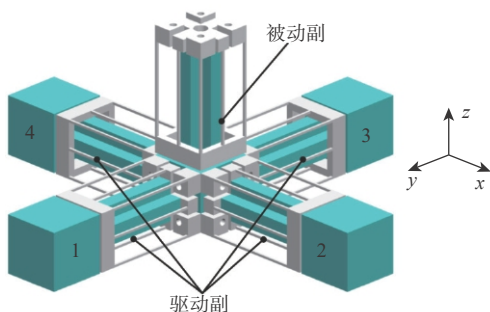
$$\{x, z, \theta_y\}_{1,3} \cup \{y, z, \theta_x\}_{2,4} = \{z\} \quad (2)$$


图2 PPPR支链结构

Fig. 2 Structure of PPPR branch chain

$$\{x, y, \theta_z\} \cap \{z\} = \{x, y, z, \theta_z\} \quad (3)$$

支链驱动副的单移动自由度满足驱动器不可承受横向载荷的特点,支链整体自由度数不少于微定位平台预期可实现的自由度数,满足空间平动柔性微定位平台支链的自由度要求。

基于图2的柔性支链,所提的3-PPPR型大行程柔性微定位平台结构如图3所示,3条相同的PPPR型柔性支链相对动平台呈空间正交分布,其结构的完全各向同性保证了平台各方向具有相同的静、动态特性,且无冗余驱动现象,便于对其进行分析。平台运动自由度可表示为

$$\{x, y, z, \theta_x\} \cup \{x, y, z, \theta_y\} \cup \{x, y, z, \theta_z\} = \{x, y, z\} \quad (4)$$

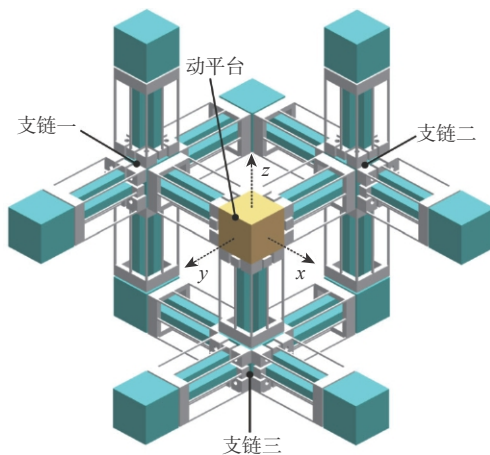


图3 3-PPPR平台结构示意图

Fig. 3 Structure of 3-PPPR micro-positioning platform

## 2 微定位平台伴生转动分析

空间平动柔性微定位平台在理想状态下,运动副内部结构受力变形完全相等,输出动平台只有平动没有转动。但实际工作过程中,运动副内部柔性杆受驱动力作用产生的弹性变形并不完全相等,进而导致输出动平台在平动的同时,产生微小转动,称之为伴生转动<sup>[25]</sup>,其转角为伴生转角。伴生转动的存在将影响平台的定位精度,因此,有必要对平台伴生转动进行理论分析。

### 2.1 PPR 运动副伴生转动分析

图 4 为一端固定的柔性导向梁, 基于梁的弯曲变形理论, 柔性梁挠度曲线近似的微分方程为

$$EI\omega'' = M(x_i) = F(l - x_i) - M - F_N\omega \quad (5)$$

式中:  $E$  为材料弹性模量;  $I$  为柔性杆截面惯性矩;  $l$  为杆长度;  $\omega$  为其弯曲变形量;  $F$  和  $M$  分别为其末端所受力与力矩;  $F_N$  为其轴向拉/压力。

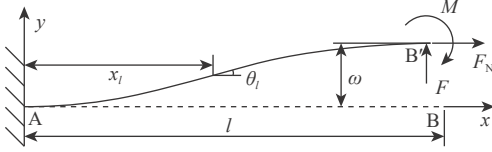


图 4 柔性固定导向梁受力分析

Fig. 4 Force analysis of compliant fixed guide beam

考虑到 PPR 柔性运动副反向串联结构, 柔性杆不会产生轴向形变及拉/压力, 则有  $F_N=0$ 。

其挠度及端截面转角方程分别表示为

$$\omega'' = \frac{F(l - x_i) - M}{EI} \quad (6)$$

$$\theta_i = \omega' = \frac{Fx_i(2l - x_i) - 2Mx_i}{2EI} \quad (7)$$

式中:  $\omega''$  为柔性杆挠度;  $\omega'$  为柔性杆端截面转角。

如图 5 所示, 基于柔性固定导向梁组成的 PPR 柔性运动副在承受水平力  $F$  时产生伴生转动, 其转动效果等效于外四杆结构分别受力  $F_a$ 、力偶  $M_a$  及内四杆结构受力  $F_b$ 、力偶  $M_b$  作用的累加, 其中:  $F_a = F_b = F$ 、 $M_a = F(s + 0.5D)$ 、 $M_b = F(s + l + 0.5D)$ ,  $s$ 、 $l$  和  $D$  均为柔性运动副结构参数(见图 1)。

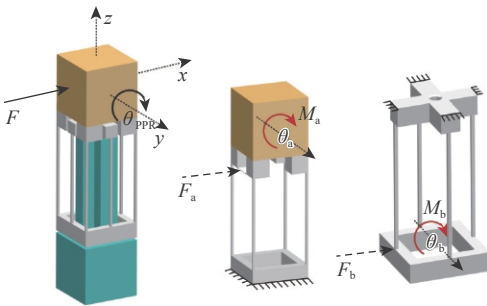


图 5 PPR 运动副伴生转动示意图

Fig. 5 Schematic diagram of parasitic rotation of PPR joint

假设运动副中刚性部件具有绝对刚度, 运动过程中不产生自身形变, 则 4 种载荷情况下的形变示意图分别如图 6~图 9 所示, 其伴生转动将在本节进行逐一分析。

图 6 为外四杆结构受力形变示意图, 在水平力  $F_a$  作用下, 两侧柔性杆产生不同形变状态。左侧柔性杆受拉力  $F_{N1}$ , 产生轴向拉伸形变量为  $\Delta l_1$ , 右侧柔性杆受压力  $F_{N2}$ , 其压缩量为  $\Delta l_2$ , 二者导致平台

产生偏转角  $\theta_{F_a}$ , 基于几何关系, 则有

$$d \sin \theta_{F_a} = \Delta l_1 + \Delta l_2 \quad (8)$$

考虑其对称结构, 两侧柔性杆末端受力、力偶关系及整体平衡方程可表示为

$$\begin{cases} F_1 = F_2 = \frac{F_a}{4} \\ F_{N1} = F_{N2} = F_N \\ \Delta l_1 = \Delta l_2 = \frac{F_N l}{EA} \\ M_1 = M_2 \\ M_1 + M_2 = (F_{N1} + F_{N2}) \frac{d}{2} \cos \theta_{F_a} \end{cases} \quad (9)$$

式中:  $F_1$  和  $F_2$  为柔性杆所受横向力;  $A$  为柔性杆截面积。

基于式 (7) 端截面转角方程, 柔性杆末端转角可表示为

$$\theta_{F_a} = \theta|_{x_i=l} = \frac{F l^2 - 2Ml}{2EI} \quad (10)$$

联立式 (8)~式 (10) 可得

$$\theta_{F_a} = \frac{F_a l^2}{2E(4I + Ad^2)} \quad (11)$$

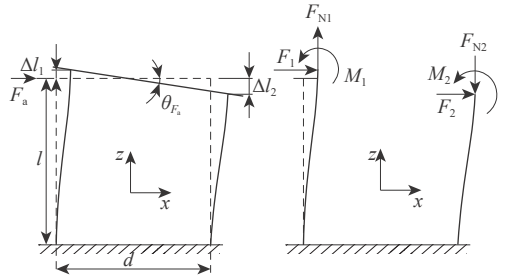


图 6 外四杆结构受力形变示意图

Fig. 6 Deformation of external beams subjected to horizontal force

图 7 为外四杆结构受力矩形变简化示意图, 在附加力矩  $M_a$  作用下, 左、右两侧柔性杆分别承受拉力  $F_{N1}$ 、压力  $F_{N2}$ , 并产生轴向形变  $\Delta l_1$ 、 $\Delta l_2$ , 设平台产生偏转角为  $\theta_{M_a}$ , 根据几何关系有

$$d \sin \theta_{M_a} = \Delta l_1 + \Delta l_2 \quad (12)$$

其力平衡方程可表示为

$$\begin{cases} M_1 = M_2 = \frac{M_a}{4} \\ F_{N1} = F_{N2} = F_N \\ \Delta l_1 = \Delta l_2 = \frac{F_N l}{EA} \\ M_1 + M_2 = (F_{N1} + F_{N2}) \frac{d}{2} \cos \theta_{M_a} \end{cases} \quad (13)$$

联立式 (12) 和式 (13) 可得偏转角  $\theta_{M_a}$  为

$$\theta_{M_a} = \frac{M_a l}{EA d^2} \quad (14)$$

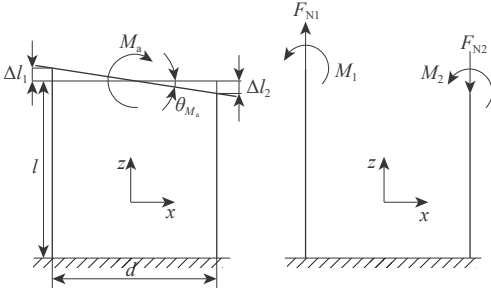


图7 外四杆结构受力矩形变形示意图

Fig. 7 Deformation of external beams subjected to moment

图8为内四杆结构受力 $F_b$ 作用示意图,为便于分析,可将该力分解为沿 $x'$ 、 $y'$ 轴方向的两分力 $F_{b1}$ 、 $F_{b2}$ ,并分别产生附加力偶 $M_{b1}$ 、 $M_{b2}$ 。力 $F_{b1}$ 作用下,内四杆结构产生绕 $x'$ 轴方向的偏转角 $\theta_1$ ,其偏转原理与上述力 $F_a$ 作用于外四杆结构导致其产生偏转的原理一致,基于式(11)可得

$$\begin{cases} F_{b1} = \frac{\sqrt{2}}{2} F \\ d' = \frac{\sqrt{2}}{2} d \\ \theta_1 = \frac{F_{b1} l^2}{2E(4I + A(d')^2)} \end{cases} \quad (15)$$

同理,可求得力 $F_{b2}$ 作用导致内四杆结构产生绕 $y'$ 轴方向的偏转角 $\theta_2$ 为

$$\theta_2 = \theta_1 \quad (16)$$

根据其几何关系,二者综合作用导致内四杆结构产生的绕 $y$ 轴伴生转角 $\theta_{F_b}$ 为

$$\theta_{F_b} = \frac{(\theta_1 + \theta_2)d'}{d} = \sqrt{2}\theta_1 \quad (17)$$

附加力偶 $M_{b1}$ 、 $M_{b2}$ 大小相等,方向相反,二者共同作用引起的结构绕 $z$ 轴方向扭转角互相抵消,故在此不予考虑。

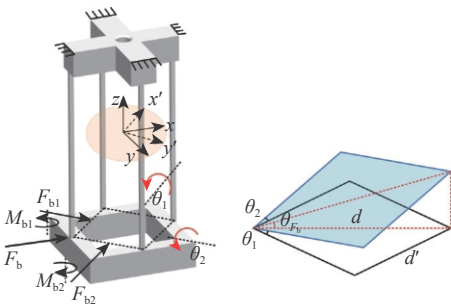


图8 内四杆结构受力形变示意图

Fig. 8 Deformation of internal beams subjected to horizontal force

图9为内四杆结构受力矩 $M_b$ 作用下产生的伴生转动示意图,左、右两侧柔性杆受力分别产生轴向形变,中间柔性杆不产生形变,设其偏转角为 $\theta_{M_b}$ ,则其形变及力偶平衡关系为

$$\begin{cases} M_1 = M_2 = \frac{M_b}{2} \\ F_{N1} = F_{N2} = F_N \\ d \sin \theta_{M_b} = \Delta l_1 + \Delta l_2 = \frac{2F_N l}{EA} \\ M_1 + M_2 = (F_{N1} + F_{N2}) \frac{d}{2} \cos \theta_{M_b} \end{cases} \quad (18)$$

求得

$$\theta_{M_b} = \frac{2M_b l}{EA d^2} \quad (19)$$

综合内外四杆分别受力、力矩的4种载荷情况下结构的伴生转动分析,PPR柔性运动副受水平作用力产生的伴生转角为

$$\theta_{PPR} = \theta_{F_a} + \theta_{M_a} - \theta_{F_b} + \theta_{M_b} \quad (20)$$

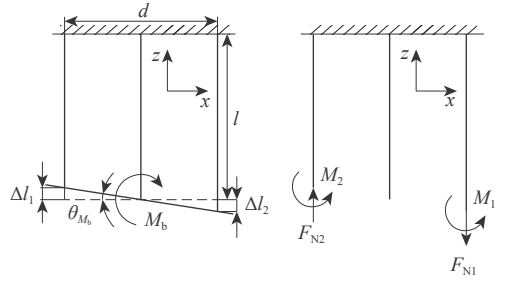


图9 内四杆结构受力偶形变示意图

Fig. 9 Deformation of internal beams subjected to moment

## 2.2 PPR支链伴生转动分析

图10(a)为基于PPR型柔性运动副组成的平台支链伴生转动示意图,在水平力 $F$ 作用下,支链产生的伴生转角由两部分组成:①被动PPR副受力 $F$ 产生的伴生转角 $\theta_{PPR}$ ,已在2.1节分析;②驱动P副受力矩 $M$ 产生的伴生扭转角 $\alpha$ 、偏转角 $\beta$ ,示意如图10(b)所示。

支链受水平力 $F$ 作用时,在驱动副中心产生附加力矩 $M = F(s + 2l + L + D)$ ,如图10(b)所示。在力矩 $M$ 作用下, $x$ 轴方向运动副发生绕 $y$ 轴方向偏转,偏转角为 $\alpha$ , $y$ 轴方向运动副发生绕 $y$ 轴方向扭转,其扭转角为 $\beta$ ,且 $\alpha = \beta$ 。假设 $M_d$ 、 $M_t$ 分别为使运动副发生偏转、扭转的力矩分量,则

$$\begin{cases} M = 2(M_d + M_t) l \\ \alpha = \frac{M_d}{K_d} = \frac{M_t}{K_t} = \beta \\ K_d = M / \theta_{PPR} \end{cases} \quad (21)$$

式中: $K_d$ 和 $K_t$ 分别为运动副偏转及扭转刚度, $K_t$ 的具体分析可参阅文献[26-28]。

基于式(21),支链驱动副产生的伴生扭转角和偏转角为

$$\alpha = \beta = \frac{M}{2(K_d + K_t)} \quad (22)$$

则支链整体受 $x$ 方向水平力 $F$ 作用产生绕 $y$ 轴

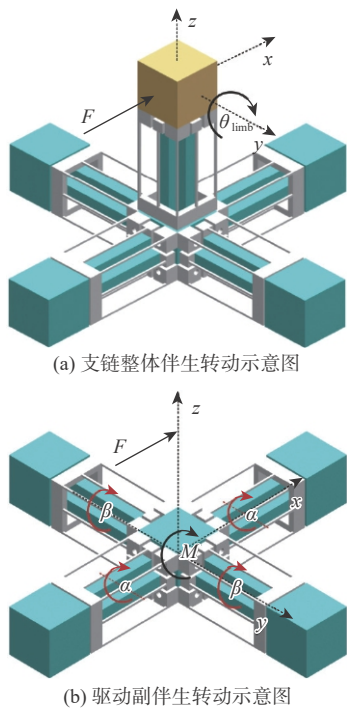


图 10 PPR 型柔性支链伴生转动示意图

Fig. 10 Schematic diagram of parasitic rotation of PPR limb

方向的伴生转角为

$$\theta_{limb} = \theta_{PPR} + \beta \quad (23)$$

### 2.3 3-PPR 平台伴生转动分析

微定位平台运动包括单轴、双轴及三轴驱动, 分别对应不同的运动状态, 故其伴生转动也不尽相同, 因此, 有必要分别对 3 种运动状态下伴生转动进行分析。本文考虑  $x$  单轴驱动,  $x$ 、 $y$  双轴驱动及  $x$ 、 $y$ 、 $z$  三轴同时驱动情况下平台的伴生转动。

#### 2.3.1 单轴驱动平台伴生转动分析

以  $x$  轴方向单轴驱动为例, 其伴生转动示意如图 11 所示, 平台  $y$ 、 $z$  方向支链均会产生伴生转动, 其中  $y$  方向支链产生绕  $z$  轴的伴生转角  $\theta_z^y$ ,  $z$  方向支链产生绕  $y$  轴的伴生转角  $\theta_y^z$ 。

图 12 为  $xz$  平面支链伴生转动原理图, 平台沿  $x$  轴方向单轴驱动时,  $z$  方向支链及动平台受力产生绕  $y$  轴转角  $\theta_y^z$ , 由于结构的整体性, 继而导致  $x$  支

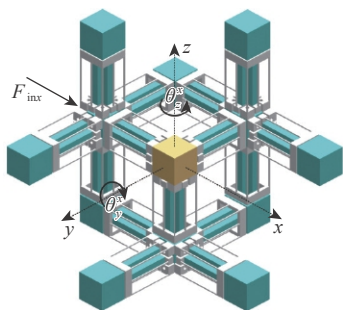


图 11 单轴驱动平台伴生转动示意图

Fig. 11 Parasitic rotations of uni-axis driven platform

链产生同样伴生转角, 则有

$$\begin{cases} F_z = F_{inx} - 4k_{PPR}\delta_{inx} \\ k_{PPR} = \frac{2Et^4}{l^3} \\ \delta_{inx} = \frac{F_{inx}}{6k_{PPR}} \\ K_x = K_z = F_z/\theta_{limb} \\ \theta_y^z = F_z/(K_x + K_z) \end{cases} \quad (24)$$

式中:  $F_z$  为  $z$  方向支链的有效水平力;  $k_{PPR}$  为 PPR 柔性运动副的平移刚度;  $\delta_{inx}$  为  $x$  方向驱动力作用下产生的驱动位移;  $K_x$  和  $K_z$  为输入力条件下平台支链的广义伴生转动刚度。

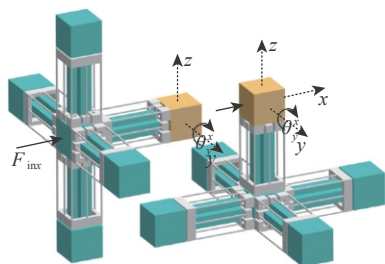


图 12 平台伴生转动原理图

Fig. 12 Schematic diagram of parasitic rotation of platform

基于式(24),  $x$  轴驱动时平台绕  $y$  轴伴生转角为

$$\theta_y^z = 0.5\theta_{limb} \quad (25)$$

同理,  $yz$  平面内平台绕  $z$  轴伴生转角  $\theta_z^y = \theta_y^z$ 。

同理, 当微定位平台沿  $y$ 、 $z$  轴方向单轴驱动时, 输出动平台各方向产生的伴生转角为

$$\begin{cases} \theta_x^y = \theta_z^y = 0.5\theta_{limb} \\ \theta_x^z = \theta_y^z = 0.5\theta_{limb} \end{cases} \quad (26)$$

式中:  $\theta_j^i$  为沿  $i$  轴方向作用力, 平台产生的绕  $j$  方向的伴生转角。

#### 2.3.2 双轴驱动平台伴生转动分析

微定位平台双轴驱动时, 以  $x$ 、 $y$  轴同时驱动为例, 其产生的伴生转动示意如图 13 所示, 该驱动条件下产生伴生转动效果为  $x$  单轴驱动、 $y$  单轴驱动伴生转动效果的叠加, 即

$$\begin{cases} \theta_x^{xy} = \theta_x^x \\ \theta_y^{xy} = \theta_y^y \\ \theta_z^{xy} = \theta_z^x + \theta_z^y = 0 \end{cases} \quad (27)$$

式中:  $x$ 、 $y$  单轴驱动时, 动平台分别产生的绕  $z$  轴的伴生转角  $\theta_z^x$ 、 $\theta_z^y$  大小相等、方向相反, 二者互相抵消, 故动平台在该情况下不产生绕  $z$  轴的伴生转角。

同理, 微定位平台沿  $x$ 、 $z$  方向及  $y$ 、 $z$  方向双轴驱动时, 输出动平台各方向产生的伴生转角分别为

$$\begin{cases} \theta_x^{xz} = \theta_x^x \\ \theta_y^{xz} = \theta_y^x + \theta_y^z = 0 \\ \theta_z^{xz} = \theta_z^x \end{cases} \quad (28)$$

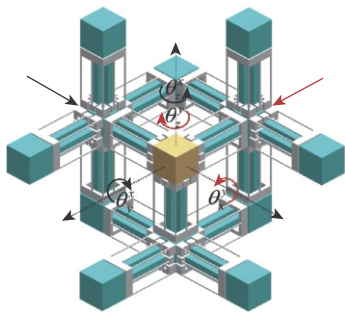


图 13 双轴驱动平台伴生转动示意图

Fig. 13 Parasitic rotations of bi-axis driven platform

$$\begin{cases} \theta_x^{yz} = \theta_x^y + \theta_x^z = 0 \\ \theta_y^{yz} = \theta_y^z \\ \theta_z^{yz} = \theta_z^y \end{cases} \quad (29)$$

### 2.3.3 三轴驱动平台伴生转动分析

平台  $x$ 、 $y$ 、 $z$  三轴同时驱动时, 其伴生转动示意图如图 14 所示, 动平台产生伴生转动效果为  $x$  单轴驱动、 $y$  单轴驱动、 $z$  单轴驱动伴生转动效果的叠加, 即

$$\begin{cases} \theta_x^{yz} = \theta_x^y + \theta_x^z = 0 \\ \theta_y^{yz} = \theta_y^x + \theta_y^z = 0 \\ \theta_z^{yz} = \theta_z^x + \theta_z^y = 0 \end{cases} \quad (30)$$

式 (30) 表明该微定位平台三轴同时驱动时, 动平台在各方向均不会产生伴生转动, 其定位精度不受伴生转动影响。

通过对柔性运动副、支链及微定位平台 3 种驱动状态下伴生转动的分析, 明确了其变形机理, 为微定位平台伴生转动的灵敏度分析及优化奠定了理论基础。

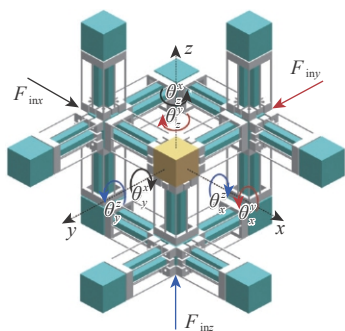


图 14 三轴驱动平台伴生转动示意图

Fig. 14 Parasitic rotations of tri-axis driven platform

## 3 仿真验证

为验证第 2 节伴生转动理论模型的正确性, 基于 ABAQUS 对平台进行有限元仿真验证, 平台选用材料 7075 铝合金及材料特性、尺寸数值如表 1 所示, 为尽可能提高仿真精度, 设定柔性单元部分网格为 1 mm, 其他部分为 4 mm, 选用四面体单元进行网格划分以提高分析速度, 经检测其整体网格化误差为 1.30%, 符合仿真要求, 在各支链被动副连

接模块处施加固定约束。为分析平台不同驱动条件下的伴生转动, 对平台分别施加单轴、双轴及三轴驱动力, 如图 15 所示。

对平台各方向分别施加 20 N 驱动力, 各驱动状态下动平台产生绕各方向伴生转角的理论值、仿真值及其相对误差如表 2 所示。

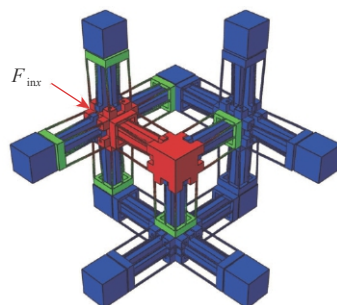
以  $x$  轴方向单轴驱动为例, 在 0~20 N 驱动力作用下, 支链及平台产生绕  $y$  轴伴生转角的理论、仿真值曲线如图 16 所示。

基于表 2 及图 16 可以得出以下结论:

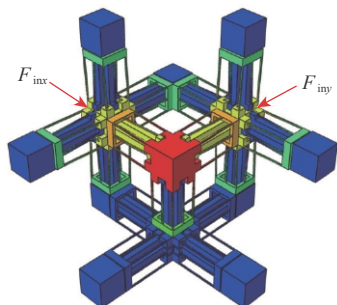
表 1 平台材料及尺寸参数

Table 1 Material and dimension parameters of platform

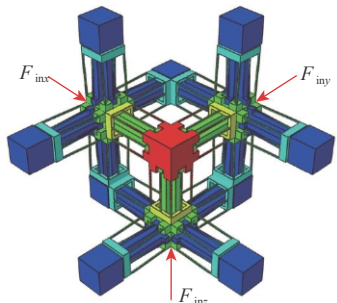
参数	数值	参数	数值
$l/\text{mm}$	40	弹性模量 $E/\text{GPa}$	71.7
$t/\text{mm}$	1	剪切模量 $G/\text{GPa}$	26.9
$d/\text{mm}$	24	泊松比 $\nu$	0.33
$s/\text{mm}$	6	屈服强度 $\sigma/\text{MPa}$	503
$L/\text{mm}$	47	密度 $\rho/(\text{kg}\cdot\text{m}^{-3})$	2 810
$D/\text{mm}$	25		



(a) 单轴驱动



(b) 双轴驱动



(c) 三轴驱动

图 15 不同驱动条件下平台的仿真分析

Fig. 15 Simulation analysis of platform under different driving conditions

表 2 平台伴生转角理论、仿真值及其相对误差

Table 2 Theoretical values, simulation values and relative errors of parasitic rotational angles

运动状态	伴生转角	理论值/rad	仿真值/rad	相对误差/%
单支链	$\theta_{limb}$	$8.96 \times 10^{-4}$	$8.77 \times 10^{-4}$	2.12
	$\theta_y^x$	$4.48 \times 10^{-4}$	$4.39 \times 10^{-4}$	2.01
	$\theta_z^x$	$4.48 \times 10^{-4}$	$4.40 \times 10^{-4}$	1.79
x、y双轴	$\theta_x^{xy}$	$4.48 \times 10^{-4}$	$4.37 \times 10^{-4}$	2.46
	$\theta_y^{xy}$	$4.48 \times 10^{-4}$	$4.42 \times 10^{-4}$	1.34
	$\theta_z^{xy}$	0	$1.2 \times 10^{-7}$	
x、y、z三轴	$\theta_x^{xyz}$	0	$4.0 \times 10^{-7}$	
	$\theta_y^{xyz}$	0	$8.8 \times 10^{-7}$	
	$\theta_z^{xyz}$	0	$4.0 \times 10^{-7}$	

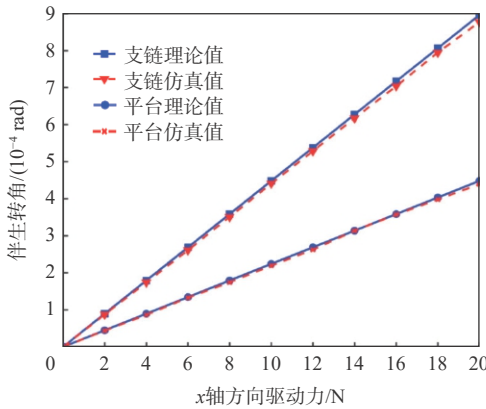


图 16 支链及平台伴生转角

Fig. 16 Parasitic rotational angles of limb and platform

1) 在 0~20 N 驱动力作用下, 支链产生最大伴生转角的理论及仿真值分别为  $8.96 \times 10^{-4}$  rad 和  $8.77 \times 10^{-4}$  rad, 最大相对误差不超过 2.12%; 平台产生最大伴生转角的理论及仿真值分别为  $4.48 \times 10^{-4}$  rad 和  $4.39 \times 10^{-4}$  rad, 最大相对误差不超过 2.01%, 二者结果的高度吻合, 验证了伴生转动理论模型的正确性。

2) 驱动条件相同的情况下, 平台单轴驱动产生的伴生转角为支链伴生转角的一半, 这与 2.3.1 节理论分析模型相吻合。

3) 平台在多轴驱动时, 可利用结构特性抵消部分伴生转动, 提高了该方向定位精度; 且平台三轴同时驱动时, 动平台各方向均不存在伴生转动, 运动效果最佳。

4) 平台运动过程中, 驱动力与伴生转角之间呈现明显的线性关系, 表明该平台具有良好的线性特性; 同时, 也有利于通过线性误差补偿的方式来进一步提高平台定位精度, 简化了平台的控制系统设计。

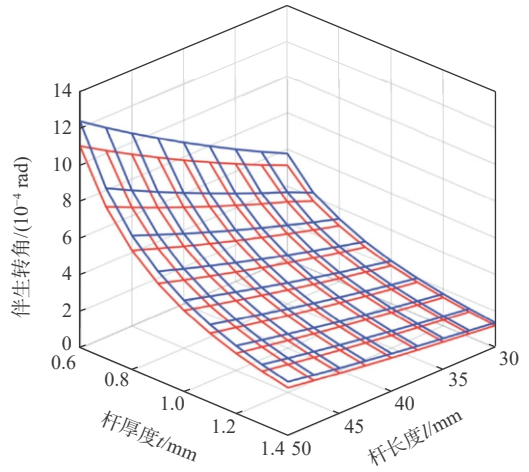
### 4 灵敏度分析

为保证大行程柔性微定位平台具有较高的定位精度, 需尽量减小平台运动过程中产生的伴生转角, 因此, 有必要对其进行灵敏度分析, 进一步明确

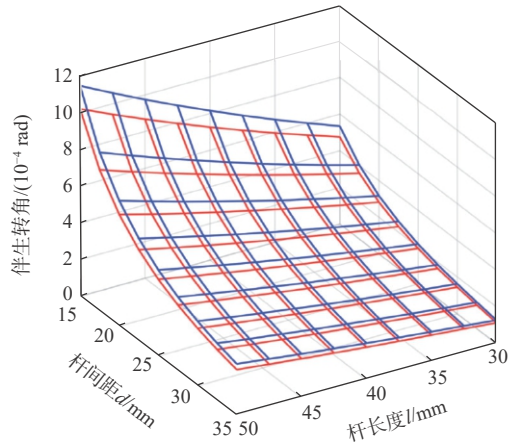
柔性杆结构参数与平台伴生转角之间的关系, 以便设计满足性能要求的微定位平台。

基于第 2 节理论模型及有限元仿真, 分别对平台伴生转角进行灵敏度分析, 得出平台伴生转角与柔性杆尺寸参数(杆长度  $l$ 、杆厚度  $t$ 、杆间距  $d$ )之间的关系如图 17 所示。

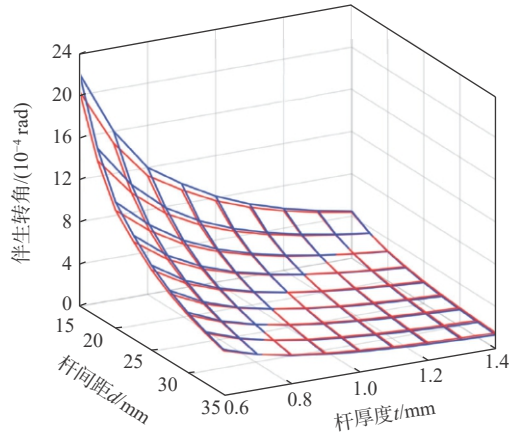
由图 17 可知伴生转角理论模型与仿真结果具



(a) 平台伴生转角与柔性杆长度、厚度之间的关系



(b) 平台伴生转角与柔性杆长度、间距之间的关系



(c) 平台伴生转角与柔性杆厚度、间距之间的关系

□ 理论值 □ 仿真值

图 17 平台伴生转角与柔性杆尺寸参数之间的关系

Fig. 17 Relationship between beam parameters and parasitic rotation of platform

有较高的吻合度及相同的变化趋势,表明了柔性杆尺寸参数在一定范围内变化对理论模型具有良好的适应性。同时,灵敏度分析结果显示,通过减小柔性杆的长度 $l$ 、增大柔性杆的厚度 $t$ 及杆间距 $d$ 均可有效减小平台运动过程中产生的伴生转角,且柔性杆厚度 $t$ 对伴生转角的影响最为显著,杆长度 $l$ 次之,杆间距 $d$ 影响最小。

值得指出的是,通过调节柔性杆长度 $l$ 来减小伴生转角的同时,亦会限制微定位平台的有效运动行程,难以同时满足大行程、高定位精度的性能要求。

因此,在保证平台运动性能的同时,可通过以下方面降低其伴生转动的影响:

1) 调节柔性杆尺寸参数,通过增大柔性杆厚度 $t$ 及杆间距 $d$ 减小伴生转角。

2) 改善驱动方式,采用多轴同时驱动的方式可有效减小或抵消部分方向伴生转角。

3) 优化结构布局,采用支链对称分布的五、六支链结构,利用结构对称性抵消平台单轴驱动产生的部分伴生转角。

## 5 结论

1) 提出一种基于柔性杆的3-PPPR型大行程柔性微定位平台。

2) 基于线弹性梁变形理论,考虑柔性杆轴向微小形变,对柔性运动副及平台在单轴、双轴及三轴同时驱动条件下产生的伴生转角进行系统性分析,并完成理论建模。

3) 通过有限元仿真对柔性运动副及平台在3种驱动条件下产生的伴生转角进行分析,其最大相对误差不超过2.46%,验证了理论模型的正确性。

4) 基于理论模型,对平台伴生转角进行灵敏度分析,明确了各尺寸参数与伴生转角之间的关系,为平台性能进一步提升奠定了基础。

后续将搭建微定位平台的物理样机及实验系统,对平台的相关性能及理论模型做进一步的测试与验证。

## 参考文献 (References)

[1] 于靖军,郝广波,陈贵敏,等. 柔性机构及其应用研究进展[J]. *机械工程学报*, 2015, 51(13): 53-68.  
YU J J, HAO G B, CHEN G M, et al. State-of-art of compliant mechanisms and their applications[J]. *Journal of Mechanical Engineering*, 2015, 51(13): 53-68(in Chinese).

[2] HOWELL L L. *Compliant mechanisms*[M]. New York: John Wiley and Sons, 2001: 2-14.

[3] 贾晓辉,刘今越,田延岭. 空间全柔性并联机构动力学分析[J]. *农业机械学报*, 2012, 43(8): 210-214.  
JIA X H, LIU J Y, TIAN Y L. Dynamics analysis of spatial com-

pliant parallel mechanism[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2012, 43(8): 210-214(in Chinese).

[4] 周睿,周辉,桂和利,等. 基于柔性铰链的二自由度微动平台分析及优化[J]. *北京航空航天大学学报*, 2018, 44(9): 1982-1990.  
ZHOU R, ZHOU H, GUI H L, et al. Analysis and optimization of 2-DoF micro-positioning stage based on flexible hinges[J]. *Journal of Beijing University of Aeronautics and Astronautics*. 2018, 44(9): 1982-1990 (in Chinese).

[5] 曹毅,王保兴,孟刚,等. 大行程三平动柔性微定位平台的设计分析及优化[J]. *机械工程学报*, 2020, 56(17): 71-81.  
CAO Y, WANG B X, MENG G, et al. Design analysis and optimization of large range spatial translational compliant micro-positioning stage[J]. *Journal of Mechanical Engineering*, 2020, 56(17): 71-81 (in Chinese).

[6] 王萍萍,刘磊. 柔性航天器高精度隔振与定向研究[J]. *宇航学报*, 2012, 33(9): 1195-1202.  
WANG P P, LIU L. Research on high accuracy pointing of the flexible spacecraft with Stewart platform[J]. *Journal of Astronautics*, 2012, 33(9): 1195-1202(in Chinese).

[7] 刘璟龙,张崇峰,邹怀武,等. 基于干扰观测器的柔性空间机器人在轨精细操作控制方法[J]. *航空学报*, 2021, 42(1): 523899.  
LIU J L, ZHANG C F, ZOU H W, et al. On-orbit precise operation control method for flexible joint space robots based on disturbance observer[J]. *Acta Aeronautica et Astronautica Sinica*, 2021, 42(1): 523899(in Chinese).

[8] DI GIAMBERARDINO P, BAGOLINI A, BELLUTTI P, et al. New MEMS tweezers for the viscoelastic characterization of soft materials at the microscale[J]. *Micromachines*, 2017, 9(1): 15.

[9] 李杨民,汤晖,徐青松,等. 面向生物医学应用的微操作机器人技术发展态势[J]. *机械工程学报*, 2011, 47(23): 1-13.  
LI Y M, TANG H, XU Q S, et al. Development status of micromanipulator technology for biomedical applications[J]. *Journal of Mechanical Engineering*, 2011, 47(23): 1-13(in Chinese).

[10] WANG R Z, ZHANG X M. A planar 3-DOF nanomanipulation platform with large magnification[J]. *Precision Engineering*, 2016, 46: 221-231.

[11] HAO G B, KONG X W. Design and modeling of a large-range modular XYZ compliant parallel manipulator using identical spatial modules[J]. *Journal of Mechanisms and Robotics*, 2012, 4(2): 1-10.

[12] 牟新明,王建华,杨密. 平行簧片机构力学分析与计算[J]. *纳米技术与精密工程*, 2005, 3(4): 278-282.  
MU X M, WANG J H, YANG M. Mechanical analysis and computation of parallel springs[J]. *Nanotechnology and Precision Engineering*, 2005, 3(4): 278-282(in Chinese).

[13] LUO Y Q, LIU W Q. Analysis of the displacement of distributed compliant parallel-guiding mechanism considering parasitic rotation and deflection on the guiding plate[J]. *Mechanism and Machine Theory*, 2014, 80: 151-165.

[14] HAO G B. Towards the design of monolithic decoupled XYZ compliant parallel mechanisms for multi-function applications[J]. *Mechanical Sciences*, 2013, 4(2): 291-302.

[15] HAO G B, LI H Y. Design of 3-legged XYZ compliant parallel manipulators with minimised parasitic rotations[J]. *Robotica*, 2015, 33(4): 787-806.

[16] SU H J, SHI H, YU J J. A symbolic formulation for analytical com-

- pliance analysis and synthesis of flexure mechanisms[J]. *Journal of Mechanical Design*, 2012, 134(5): 051009.
- [17] AWTAR S, USTICK J, SEN S. An XYZ parallel kinematic flexure mechanism with geometrically decoupled degrees of freedom[J]. *Journal of Mechanisms and Robotics*, 2013, 5(1): 015001.
- [18] LIN H R, CHENG C H, HUNG S K. Design and quasi-static characteristics study on a planar piezoelectric nanopositioner with ultra-low parasitic rotation[J]. *Mechatronics*, 2015, 31(1): 180-188.
- [19] DONG W, SUN L N, DU Z J. Design of a precision compliant parallel positioner driven by dual piezoelectric actuators[J]. *Sensors and Actuators A:Physical*, 2007, 135(1): 250-256.
- [20] HERPE X, WALKER R, DUNNIGAN M, et al. On a simplified nonlinear analytical model for the characterisation and design optimisation of a compliant XY micro-motion stage[J]. *Robotics and Computer-Integrated Manufacturing*, 2018, 49: 66-76.
- [21] 李海洋, 郝广波, 于靖军, 等. 空间平动柔性并联机构的系统设计方法研究[J]. *机械工程学报*, 2018, 54(13): 57-65.  
LI H Y, HAO G B, YU J J, et al. Systematic approach to the design of spatial translational compliant parallel mechanisms[J]. *Journal of Mechanical Engineering*, 2018, 54(13): 57-65 (in Chinese).
- [22] HOPKINS J B, CULPEPPER M L. Synthesis of multi-degree of freedom, parallel flexure system concepts via freedom and constraint topology (FACT) Part I: Principles[J]. *Precision Engineering*, 2010, 34(2): 259-270.
- [23] HOPKINS J B, CULPEPPER M L. Synthesis of multi-degree of freedom, parallel flexure system concepts via Freedom and constraint topology (FACT) Part II: Principles[J]. *Precision Engineering*, 2010, 34(2): 271-278.
- [24] 黄真, 刘婧芳, 曾达幸. 基于约束螺旋理论的机构自由度分析的普遍方法[J]. *中国科学*, 2009, 39(1): 84-93.  
HUANG Z, LIU J F, ZENG D X. A general methodology for mobility analysis of mechanism based on the constraint screw theory[J]. *Science in China*, 2009, 39(1): 84-93 (in Chinese).
- [25] MURANAKA Y, INABA M, ASANO T, et al. Parasitic rotation in parallel spring movements[J]. *International Journal of The Japan Society for Precision Engineering*, 1991, 25(3): 208-213.
- [26] KOSEKI Y, TANIKAWA T, KOYACHI N, et al. Kinematic analysis of translational 3-DOF micro parallel mechanism using matrix method[C]// *IEEE/RSJ International Conference on Intelligent Robots and Systems*. Piscataway: IEEE Press, 2000: 786-792.
- [27] 王保兴, 孟刚, 林苗, 等. 3-PPP型柔性并联微定位平台的设计与分析[J]. *北京航空航天大学学报*, 2020, 46(4): 798-807.  
WANG B X, MENG G, LIN M, et al. Design and analysis of a 3-PPP compliant parallel micro-positioning stage[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2020, 46(4): 798-807 (in Chinese).
- [28] TANG H, LI Y M. Design analysis and test of a novel 2-DOF nanopositioning system driven by dual mode[J]. *IEEE Transactions on Robotics*, 2013, 29(3): 650-662.

## Parasitic rotation of large stroke compliant micro-positioning platform

MENG Gang<sup>1</sup>, HUANG He<sup>1</sup>, WU Weiguan<sup>1</sup>, JU Yongjian<sup>1</sup>, CAO Yi<sup>1, 2, \*</sup>

(1. School of Mechanical Engineering, Jiangnan University, Wuxi 214122, China;

2. Jiangsu Key Laboratory of Advanced Food Manufacturing Equipment and Technology, Wuxi 214122, China)

**Abstract:** Parasitic rotation is inevitable during the movement of large stroke compliant micro-positioning platforms, causing a negative impact on their positioning precision. To reduce this effect, a 3-PPPR compliant micro-positioning platform with large stroke is proposed based on compliant beams. Then, based on linear elastic beam theory, the theoretical parasitic rotational angle of the PPR compliant kinematic joint is modeled considering the axial deformations of the beams. The parasitic rotational angles of the platform are also analyzed theoretically in uniaxial, biaxial and triaxial cases. Furthermore, the theoretical models are verified by finite element analysis. Finally, the sensitivity between the dimension parameters of the compliant beams and the parasitic rotational angle of the platform is analyzed, laying a foundation for the improvement of the platform. On this basis, the optimization schemes are proposed to improve the motion performance of the platform. results show that the maximum relative errors of the theoretical and simulated values of the parasitic rotational angle is 2.46% in three driven cases.

**Keywords:** compliant micro-positioning platform; parasitic rotation; multiple axis drive; finite element analysis; sensitivity analysis

Received: 2021-05-25; Accepted: 2021-08-20; Published Online: 2021-08-30 14:49

URL: [kns.cnki.net/kcms/detail/11.2625.V.20210827.1522.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210827.1522.001.html)

Foundation items: The Six Talent Peaks in Jiangsu Province (ZBZZ-012); Overseas Expertise Introduction Project for Discipline Innovation (B18027); Postgraduate Research and Practice Innovation of Jiangsu Province (KYCX20-1925)

\* Corresponding author. E-mail: [caoyi@jiangnan.edu.cn](mailto:caoyi@jiangnan.edu.cn)

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0276

# 基于 CNN 机翼气动系数预测

吕召阳<sup>1,2</sup>, 聂雪媛<sup>1,2,\*</sup>, 赵奥博<sup>1,2</sup>

(1. 中国科学院力学研究所, 北京 100190; 2. 中国科学院大学工程科学学院, 北京 100049)

**摘 要:** 随着机器学习的快速发展和其突出的非线性映射能力, 越来越多的学者将机器学习方法应用到流体力学领域。为克服传统数学拟合不能很好的解决系统非线性问题, 以及现有文献中所提及的一些基于神经网络的气动参数预测方法, 需要进行参数化处理而带来的不便, 同时为实现多变量多输出气动参数快速预测的目的, 基于卷积神经网络考虑机翼迎角和浮沉建立了一种多变量多输出的机翼气动参数预测模型, 实现了机翼气动参数的快速预测。结果表明: 所建模型具有较高且稳定的预测精度, 并且计算效率较计算流体力学 (CFD) 提高了 40 倍。

**关键词:** 卷积神经网络; 机器学习; 气动参数预测; 气动降阶; 深度学习

**中图分类号:** V221+.3; TB553

**文献标志码:** A **文章编号:** 1001-5965(2023)03-0674-07

传统机翼气动参数的获得方法可以归结为 2 种: 采用风洞试验测量和计算流体力学建模计算。但是这些方法存在高成本、计算量大和测试周期长等问题<sup>[1-2]</sup>。所以, 如何实现机翼状态改变后气动参数的快速预测成为了实现颤振主动控制的一大难点。

传统快速预测方法的基本思想是根据气动参数的统计规律, 利用数学拟合的方法得到几何参数与飞行状态和气动参数的数学关系式, 从而实现气动参数的快速计算, 但这种方法一般适用范围较小, 并且对存在很强非线性问题的系统不能很好保证预测精度<sup>[3]</sup>。随着最近几年神经网络的快速发展, 其突出的非线性映射能力吸引越来越多的学者将此方法运用到气动参数预测上。2003 年 Suresh 等<sup>[4]</sup> 基于递归神经网络建模, 预测大迎角下旋翼的升力系数并与试验数据进行比较, 证明了其方法的可行性。2011 年 Carpenter 等<sup>[5]</sup> 提出一种单隐含层的神经网络用于导弹气动参数的预测。刘昕<sup>[6]</sup> 提出一种基于径向基函数的神经网络 (radial basis function neural network, RBFNN) 模型, 并将其成功运用

到机翼沉浮振动的升阻力预测中。原智杰等<sup>[3]</sup> 基于优化的反向传播 (back propagation, BP) 神经网络进行了导弹的气动参数预测, 证明了其方法有良好的泛化和拟合能力。Balla 等<sup>[7]</sup> 提出了一种多输出的神经网络来进行二维和三维机翼的气动系数预测, 并与本征正交分解 (proper orthogonal decomposition, POD) 方法进行了比较, 结果表明神经网络具有更好的性能, 特别是在预测包含激波的流场中。

神经网络虽然在气动参数预测中表现出可喜的结果, 相比于传统拟合方法具有较好的非线性拟合能力, 但对于文献 [3-7] 所提到的研究, 在进行模型训练时都需要先对机翼进行参数化处理, 这无疑给模型的建立和训练增加了困难。如果模型能直接将机翼状态图片作为训练对象, 则将大大降低模型的建立和训练难度。在众多机器学习方法中, 卷积神经网络可以将图片直接作为训练对象, 并且具有强大的非线性映射能力和自动提取流场特征的功能<sup>[8]</sup>, 这无疑给本文研究提供了方向。陈海等<sup>[9]</sup> 基于卷积神经网络建立了单输入单输出的模型, 对不同翼型的升力系数进行了预测并通过对预测结果

收稿日期: 2021-05-27; 录用日期: 2021-07-09; 网络出版时间: 2021-08-16 16:44

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210816.1551.002.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210816.1551.002.html)

\* 通信作者. E-mail: [niexueyuan@imech.ac.cn](mailto:niexueyuan@imech.ac.cn)

**引用格式:** 吕召阳, 聂雪媛, 赵奥博. 基于 CNN 机翼气动系数预测 [J]. 北京航空航天大学学报, 2023, 49 (3): 674-680.

LYU Z Y, NIE X Y, ZHAO A B. Prediction of wing aerodynamic coefficient based on CNN [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 674-680 (in Chinese).

的误差分析证明了其方法的有效性。然而对于机翼颤振控制一类的问题而言, 其翼型并不改变, 结构相对其平衡位置的空间状态发生改变, 如何基于卷积神经网络对这类问题进行多变量输出的气动参数预测的研究尚未开展。

针对文献 [3-9] 研究现状, 本文基于卷积神经网络(convolutional neural network, CNN), 以二维翼型为对象, 建立一种基于卷积神经网络的多变量多输出预测模型, 用于实现对机翼气动参数的快速预测, 最后通过测试集上的预测结果和计算流体力学(computational fluid dynamics, CFD)的计算结果进行误差分析, 对本文所提气动参数预测模型的有效性和准确性进行验证。

### 1 CNN 理论与建模

随着深度学习的快速发展, 以及其突出的非线性映射能力, 越来越多的学者将机器学习方法运用到流体力学领域。其中由于卷积神经网络具有强大的图像特征提取和非线性映射能力, 因此, 卷积神经网络成为当下流场预测及流场特征研究中不可忽视的工具。卷积神经网络主要组成部分包括: 输入层、卷积层、池化层、激励层、全连接层及输出层。其中具有代表性的卷积神经网络如 LeNet, 如图 1 所示。

目前虽然提出了许多经典的卷积神经网络框架, 但是卷积神经网络的框架并不是一成不变的, 可以根据具体问题加以设计。本文针对建立气动参数快速预测模型, 为颤振主动控制过程服务这一

具体问题, 基于卷积神经网络原理建立了一种多变量多输出的预测模型, 其具体结构为

1) 输入层。目的是实现机翼颤振控制过程中机翼的不同状态下(包括迎角和沉浮)的气动参数快速预测, 为了保留机翼的状态信息和简化数据处理过程, 模型的输入数据采用机翼状态改变后的原始图像, 同时将在这一状态下经过 CFD 计算出的升力和力矩系数作为标签。具体形式如图 2 所示, 图下方数字分别为图的编号, 升力系数, 阻力系数。

2) 卷积层。卷积层是卷积神经网络的核心层。虽然卷积层数越多、网络深度越深训练出的网络模型预测精度越高, 但考虑实际情况和训练成本, 所建立的网络采用 3 层卷积层。

3) 池化层。卷积神经网络池化层的设置不仅可以减少网络的计算量, 从而很大程度上节约训练的时间成本, 而且还能在一定程度上防止模型过拟合的发生<sup>[10]</sup>。池化操作一般包括均值池化(对滤波器大小区域不为零像素点取均值)和最大池化(对滤波器大小区域所有像素点取最大值), 其中均值池化算法得到的特征数据对背景信息比较敏感, 最大池化会对数据纹理特征信息更加敏感。结合本

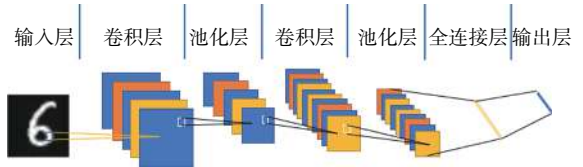


图 1 LeNet 卷积神经网络经典结构

Fig. 1 Classic structure of LeNet convolutional neural network

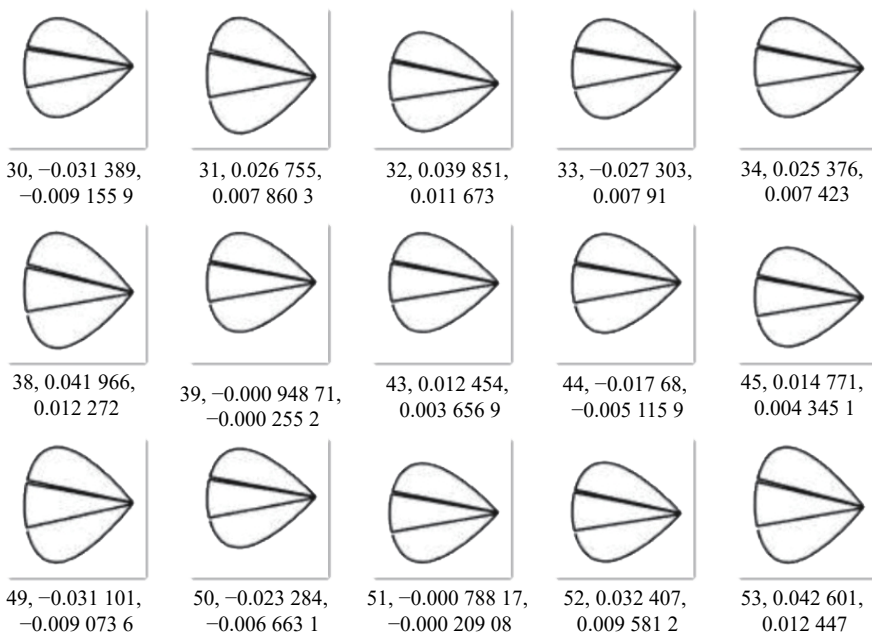


图 2 基于 RBF 网格变形机翼状态图及升力和力矩系数

Fig. 2 Deformed wing state diagram and lift drag coefficient based on RBF mesh

文要解决的问题,模型的池化层采用最大池化的方法和下采样的方式,即在卷积层后接一个池化层,以达到减小计算量和防止过拟合的目的。

4) 激励层。激励层的设置是神经网络能更好实现非线性映射的基础。饱和的非线性激励函数是传统神经网络经常使用的函数,包括 sigmoid、tanh 和 softsign 函数<sup>[11-12]</sup>,但近几年研究发现,不饱和非线性激励函数(ReLU)相比传统饱和非线性激励函数具有更为出色的收敛速度,所以不饱和非线性激励函数逐渐得到学者的青睐<sup>[13-14]</sup>。为了能得到更好的收敛速度,模型将选用 ReLU 作为激励函数,ReLU 函数表达式为

$$f(x) = \max(0, x) \quad (1)$$

式中: $x$  为自由变量。

为了能更好地拟合机翼状态与机翼升力及力矩系数的映射关系,整个网络的输出层采用 tanh 非线性激励函数,其函数表达式为

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2)$$

5) 全连接层。全连接层是卷积神经网络能进行后续处理和采用梯度下降法进行网络参数训练的基础。由于涉及到的问题需要同时预测出 2 个未知量,所以建立的卷积神经网络采用并列 2 个全连接层分别实现升力系数和力矩系数的预测。

6) 输出层。建立卷积神经网络的目的是实现机翼升力和力矩系数的同时快速预测。因此,模型设置 2 个输出层。

本节分别介绍了卷积神经网络每层的作用和所建立的卷积神经网络每层的设置细节。除此之外,为了能使所建立的卷积神经网络更好的避免过拟合问题,除了在每层卷积层后面接池化层外,还在全连接层后设置了 dropout 层(其中在训练过程中 keep\_prob=0.5,测试过程中 keep\_prob=1)。因此,建立的卷积神经网络具体结构和参数设置如表 1 所示。

表 1 卷积神经网络具体结构和参数设置

Table 1 Specific structure and parameter setting of CNN

层名称	参数
输入层	机翼状态改变后的原始图像(200×200×3)
卷积层1	卷积核(5×5×96)+relu
池化层1	滤波器大小(3×3)strides(步长)=2
卷积层2	卷积核(3×3×96)+relu
池化层2	滤波器大小(3×3)strides(步长)=2
卷积层3	卷积核(3×3×128)+relu
池化层3	滤波器大小(3×3)strides(步长)=2
全连接层(并列2个)	relu+dropout
输出层(并列2个)	激活函数tanh

## 2 气动弹性系统模型

本文研究的目的是利用卷积神经网络理论建立气动参数快速预测模型,为颤振主动控制过程服务,所以计算训练样本的研究模型选取颤振研究中常使用的二元翼型,如图 3 所示<sup>[15]</sup>。

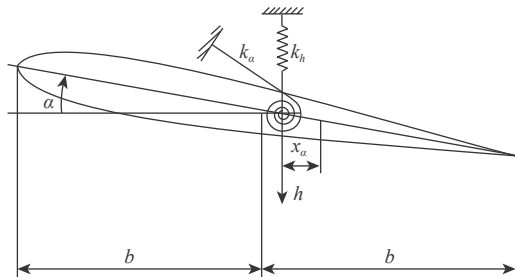


图 3 二元翼型<sup>[15]</sup>

Fig. 3 Two dimensional airfoil<sup>[15]</sup>

对二元翼型进行受力分析,其非线性控制方程可以表示为

$$\begin{bmatrix} m & mx_a b \\ mx_a b & I_a \end{bmatrix} \begin{bmatrix} \ddot{h} \\ \ddot{\alpha} \end{bmatrix} + \begin{bmatrix} k_h & 0 \\ 0 & k_a(\alpha) \end{bmatrix} \begin{bmatrix} h \\ \alpha \end{bmatrix} = \begin{bmatrix} -L \\ M \end{bmatrix} \quad (3)$$

式中: $h$  为沉浮位移; $\alpha$  为俯仰角; $k_h$  为沉浮方向刚度系数; $k_a$  为俯仰方向刚度系数; $I_a$  为机翼惯性矩; $m$  为机翼质量; $x_a$  为弹性轴到机翼重心的无量纲距离; $b$  为机翼半弦长; $L$  和  $M$  分别为气动力简化到机翼质心上的力和力矩。

从式(3)可知,求解方程只需获得机翼所受的气动力便可,又因气动力可以化简为气动升力  $L$ , 和力矩  $M$ , 因此,通过获得机翼升力系数  $C_L$  和力矩系数  $C_M$ (单位体积流体在机翼面上产生的升力和力矩)也可以实现机翼气动力的计算。因此,可将如何获得气动力来求解控制方程的问题,进一步转变成如何获得机翼升力和力矩系数来求解控制方程的问题。

为了给第 1 节所建立的卷积神经网络模型提供训练数据,基于 RBF 网格变形理论通过 CFD 计算获得机翼不同迎角和沉浮下的升力和力矩系数和几何外形,分别作为后续卷积神经网络的训练输入和标签。其中气动计算采用的是基于雷诺平均二维 Navier-Stokes 方程(Reynolds equation, RANS),其守恒的流动方程可以表示为

$$\frac{\partial}{\partial t} \int_S Q dS + \oint_l (G_c - G_v) \cdot n dl = 0 \quad (4)$$

式中: $S$  为控制体体积; $Q$  为守恒通量; $l$  为控体体边界; $G_c$  为对流通量; $G_v$  为粘性通量; $n$  为表面法向量。

### 3 机翼升力系数预测及模型性能测试

#### 3.1 模型的训练

对于卷积神经网络来讲,所取的训练数据在所运用工况范围分布越均匀其训练出的模型精度越高、稳定性越好、适用性越强。因此,本文基于拉丁超立方抽样方法在所涉及的工况范围内随机抽取了1100个工况条件,以保证训练数据在工况范围内的随机性和均匀性,其工况分布如图4所示。并利用编写的基于RBF网格变形程序结合Fluent对所抽取的工况进行计算,其计算结果作为训练和预测验证的数据。为了能更充分的体现所建立的卷积神经网络的性能,在这些计算工况中又随机选取70个工况数据做为测试集,剩余的1030个工况数据作为训练集。模型训练方法采用TensorFlow中提供的Adam算法优化器,其中学习率设置为 $1 \times 10^{-5}$ ,为了提高训练效率,数据分10批输入,每批103个样本。卷积核的初始化采用标准差为0.01的正态分布随机数,偏执初始化采用标准差为0.1的正态分布随机数。模型误差判断标准采用均方根误差RMSE,其具体表达形式为

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (\text{observed}_i - \text{predicted}_i)^2} \quad (5)$$

式中: $N$ 为每批训练样本的数量; $\text{observed}_i$ 为CFD计算出的真值; $\text{predicted}_i$ 为卷积神经网络的预测值。

神经网络模型训练结束的条件可设置为达到最大迭代次数或达到精度要求。为了能获得更好的模型预测精度,同时避免模型训练时间过长,采取固定迭代次数的方法来控制模型训练的结束,其中迭代次数设置2000次。所用到的建模和训练的工作站配置为Intel(R) Core(TM) i9-10 900 K CPU @

3.70 GHz(3 696 MHz), 128 GB 内存, 显卡 NVIDIA GeForce RTX 3 090 (24 576 MB)。采用GPU的方法对模型进行训练,因为其训练速度比CPU训练大约快12倍。模型训练具体流程如图5所示。

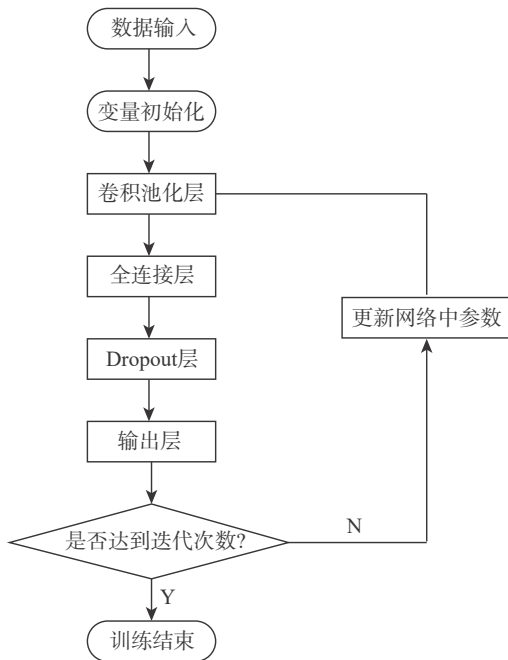


图5 CNN训练流程

Fig. 5 CNN training process

#### 3.2 结果分析

通过利用训练集进行模型的训练,其模型训练过程中升力和力矩系数的RMSE都是逐渐降低的,图6(a)和图6(b)分别为训练过程升力和力矩系数最后一批前500次迭代过程中的误差变化曲线,从图6中可以看出误差变化是随迭代次数的增加而逐渐降低,虽模型在训练到500步误差降低速度逐渐变缓,但也是呈下降趋势。

模型完全训练结束后,为测试模型训练的预测效果和预测精度,将未参与训练的工况数据输入到训练好的模型中进行机翼的升力和力矩系数的预测,预测结果如图7(a)和图7(b)所示,由图7可知,预测值和CFD的计算值几乎重合。但为了更好地分析其模型的预测精度,本文计算了预测值和真值的预测误差,如图8所示,从图8中可以看出不管是升力系数还是力矩系数,其预测误差都不超过 $2.5 \times 10^{-3}$ 。

#### 3.3 模型性能测试

本文目的是建立一种对气动参数快速预测的模型,所以模型的计算效率是评价本模型的重要指标之一。为了测试训练好的模型的计算效率,选择在相同的工作站上,分别统计基于第2节提出的CFD仿真方法的计算时间(假设迭代600步收敛)

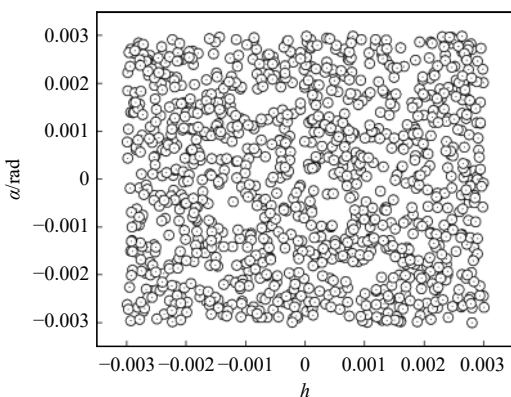


图4 样本工况分布点

Fig. 4 Distribution points of sample working conditions

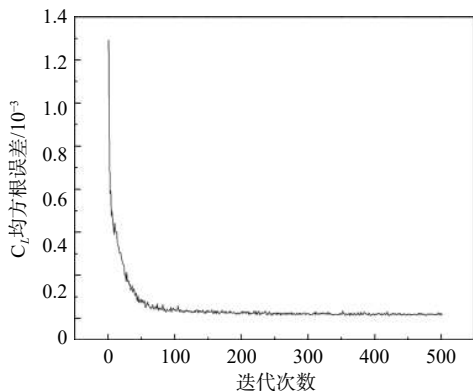
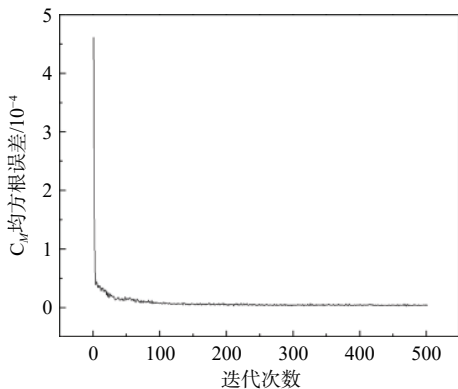
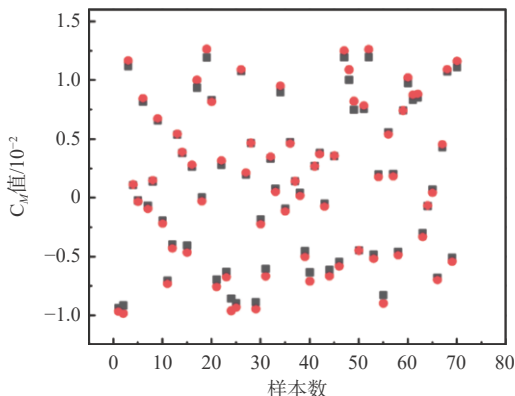
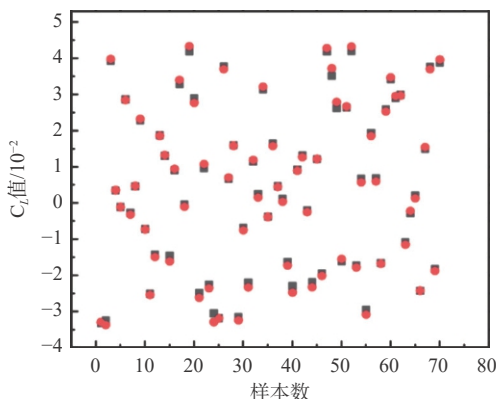
(a)  $C_L$ 均方根误差(b)  $C_M$ 均方根误差

图6 训练过程中误差变化曲线

Fig. 6 Error variation curve during training

(a)  $C_M$ 预测结果对比(b)  $C_L$ 预测结果对比

■ CNN预测值 ● CFD计算值

图7 预测结果

Fig. 7 Prediction results

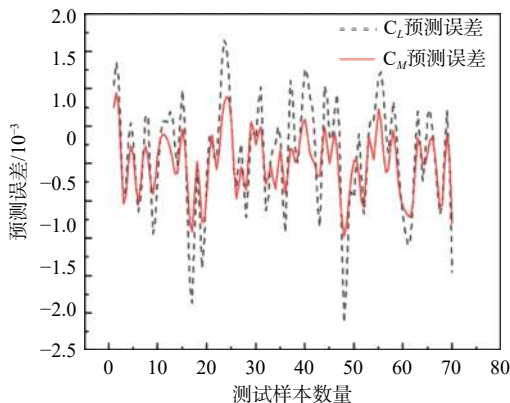


图8 预测误差曲线

Fig. 8 Prediction error curve

和训练好的 CNN 模型预测 70 个测试集的时间。其中 CFD 仿真计算一种工况需要 20 min, 所建立的 CNN 模型计算 70 个工况只需要 0.11 s 左右。可以看出本文所建立的机翼气动参数快速预测模型大大提高了机翼气动参数的计算效率。

模型的稳定性也是衡量模型好坏的重要标准之一, 为了验证本节训练结果是偶然还是必然的。本文设计了一种模型稳定性实验。为避免其他因素的影响, 采用同一工作站, 同一参数, 相同训练数据, 相同迭代次数对模型进行训练, 并统计每次训练好的模型在测试集上的误差绝对值的平均值, 将其作为模型稳定性的衡量标准。本文总共进行了 5 次实验, 为节约时间成本, 每次实验迭代设置为 1 000 步, 其结果如图 9 所示。可以看出, 不管检测  $C_L$  端口的输出值, 还是  $C_M$  端口的输出值, 其平均误差都存在波动, 但从数据上看, 5 次计算的平均误差的数量级相同, 所以可以认为建立的预测模型具有较好的稳定性。

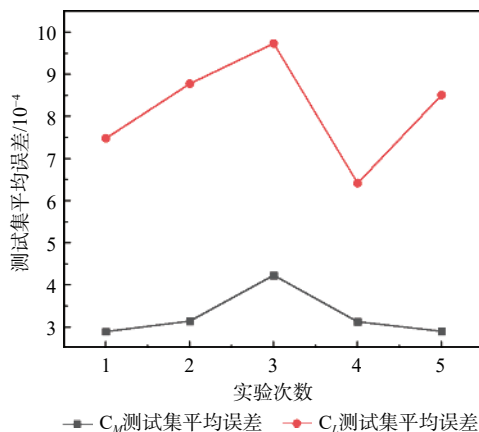


图9 测试集平均误差

Fig. 9 Average error of test set

## 4 结论

1) 基于卷积神经网络设计了一种快速预测机

翼在不同攻角和浮沉状态下升阻力系数的模型,大大缩减了机翼气动参数计算的时间,为机翼颤振主动抑制研究提供了一种气动降阶模型。

2) 通过实验证明了基于卷积神经网络建立的气动参数预测模型有很好的稳定性,同时,还发现可以通过减小模型的学习率和增大模型的迭代次数来进一步提高模型预测精度的稳定性。

### 参考文献 (References)

- [1] 安效民,徐敏,陈士楷.多场耦合求解非线性气动弹性的研究综述[J]. *力学进展*, 2009, 39(3): 284-298.  
AN X M, XU M, CHEN S L. An overview of CFD/CSD coupled solution for nonlinear aeroelasticity[J]. *Advances in Mechanics*, 2009, 39(3): 284-298(in Chinese).
- [2] 李亚东,张子军,张钧尧,等.电动飞机气动焦点辨识及飞行试验研究[J]. *航空工程进展*, 2021, 12(3): 78-84.  
LI Y D, ZHANG Z J, ZHANG J Y, et al. Research on aerodynamic center identification and flight test of electric aircraft[J]. *Advances in Aeronautical Science and Engineering*, 2021, 12(3): 78-84(in Chinese).
- [3] 原智杰,张公平,崔茅,等.基于神经网络的导弹气动参数预测[J]. *航空兵器*, 2020, 27(5): 28-32.  
YUAN Z J, ZHANG G P, CUI M, et al. Prediction of missile's aerodynamic parameters based on neural network[J]. *Aero Weaponry*, 2020, 27(5): 28-32(in Chinese).
- [4] SURESH S, OMKAR S N, MANI V, et al. Lift coefficient prediction at high angle of attack using recurrent neural network[J]. *Aerospace Science and Technology*, 2003, 7(8): 595-602.
- [5] CARPENTER M, HARTFIELD R, BURKHALTER J. A comprehensive approach to cataloging missile aerodynamic performance using surrogate modeling techniques and statistical learning[C]//29th AIAA Applied Aerodynamics Conference. Reston: AIAA, 2011: 3029.
- [6] 刘昕.基于神经网络的机翼气动参数预测仿真研究[J]. *计算机仿真*, 2015, 32(12): 67-71.  
LIU X. Simulation of airfoil plunging aerodynamic parameter prediction based on neural network[J]. *Computer Simulation*, 2015, 32(12): 67-71(in Chinese).
- [7] BALLA K, RUBEN S, OUBAY H, et al. An application of neural networks to the prediction of aerodynamic coefficients of aerofoils and wings[J]. *Applied Mathematical Modelling*, 2021, 96: 456-479.
- [8] 叶舒然,张珍,王一伟,等.基于卷积神经网络的深度学习流场特征识别及应用进展[J]. *航空学报*, 2021, 42(4): 524736.  
YE S R, ZHANG Z, WANG Y W, et al. Progress in deep convolutional neural network based flow field recognition and its applications[J]. *Acta Aeronautica et Astronautica Sinica*, 2021, 42(4): 524736(in Chinese).
- [9] 陈海,钱炜祺,何磊.基于深度学习的翼型气动系数预测[J]. *空气动力学学报*, 2018, 36(2): 294-299.  
CHEN H, QIAN W Q, HE L. Aerodynamic coefficient prediction of airfoils based on deep learning[J]. *Acta Aerodynamica Sinica*, 2018, 36(2): 294-299(in Chinese).
- [10] 吴正文.卷积神经网络在图像分类中的应用研究[D].成都:电子科技大学,2015:8-12.  
WU Z W. Application research of convolution neural network in image classification[D]. Chengdu: University of Electronic Science and Technology of China, 2015: 8-12(in Chinese).
- [11] 李宏伟,吴庆祥.智能传感器中神经网络激活函数的实现方案[J]. *传感器与微系统*, 2014, 33(1): 46-48.  
LI H W, WU Q X. Implementation scheme for activated function of neural networks in intelligent sensors[J]. *Transducer and Microsystem Technologies*, 2014, 33(1): 46-48(in Chinese).
- [12] 尹宝才,王文通,王立春.深度学习研究综述[J]. *北京工业大学学报*, 2015, 41(1): 48-59.  
YIN B C, WANG W T, WANG L C. Review of deep learning[J]. *Journal of Beijing University of Technology*, 2015, 41(1): 48-59(in Chinese).
- [13] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. *Communications of the ACM*, 2017, 60(6): 84-90.
- [14] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines[C]//Proceedings of the 27th International Conference on Machine Learning, 2010: 807-814.
- [15] 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).

# Prediction of wing aerodynamic coefficient based on CNN

LYU Zhaoyang<sup>1,2</sup>, NIE Xueyuan<sup>1,2,\*</sup>, ZHAO Aobo<sup>1,2</sup>

(1. Institute of Mechanics, Chinese Academy of Sciences, Beijing 100190, China;

2. School of Engineering Science, University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** With the rapid development of machine learning and its outstanding nonlinear mapping ability, more and more scholars apply machine learning methods to the field of fluid mechanics. To overcome the obstacle that the traditional mathematical fitting cannot well present the system nonlinearity and the inconvenience of some neural network-based aerodynamic parameter prediction methods due to the need of parametric processing, and to achieve the multi-variable and multi-output aerodynamic parameters, this paper establishes a multi-variable and multi-output model based on convolutional neural network considering the variable angle of attack and the heave of the wing to realize the rapid prediction of the aerodynamic coefficient of the wing. The results show that this model has high prediction accuracy and its computational efficiency is 40 times higher than computational fluid dynamics (CFD). Moreover, the designed stability experiment results show that the proposed model has good stability.

**Keywords:** convolutional neural network; machine learning; prediction of aerodynamic parameters; pneumatic order reduction; deep learning

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0282

# 自适应变异麻雀搜索优化算法

唐延强<sup>1,2</sup>, 李成海<sup>2,\*</sup>, 宋亚飞<sup>2</sup>, 陈晨<sup>3</sup>, 曹波<sup>1</sup>

(1. 空军工程大学 研究生院, 西安 710051; 2. 空军工程大学 防空反导学院, 西安 710051; 3. 西安卫星测控中心, 西安 710043)

**摘 要:** 针对麻雀搜索算法前期易陷入局部极值点、后期寻优精度不高等问题, 提出一种自适应变异麻雀搜索算法 (AMSSA)。先通过猫映射混沌序列初始化种群, 增强初始种群的随机性、遍历性, 提高算法的全局搜索能力; 再引入柯西变异和 Tent 混沌扰动, 拓展局部搜索能力, 使陷入局部极值点的个体跳出限制继续搜索; 最后, 提出探索者-跟随者数量自适应调整策略, 利用各阶段探索者和跟随者数量的改变增强算法前期的全局搜索能力和后期的局部深度挖掘能力, 提高算法的寻优精度。选取 16 个基准函数和 Wilcoxon 检验进行验证, 实验结果表明: 所提算法与其他算法相比, 寻优精度、收敛速度和稳定性都取得较大提升。

**关键词:** 麻雀搜索算法; 猫映射混沌; 柯西变异; Tent 混沌; 自适应调整策略

**中图分类号:** TP301.6

**文献标志码:** A **文章编号:** 1001-5965(2023)03-0681-12

元启发式算法又称群智能优化算法, 是一种模拟自然界中某些生物的行为或受到某种物理现象启发提出的一类仿生学算法, 其中心思想是在某个空间平衡全局搜索和局部搜索寻找最优解。由于元启发式算法操作简单、求解高效, 吸引了众多学者的关注, 目前这些算法被广泛应用于图像处理<sup>[1]</sup>、训练神经网络<sup>[2]</sup>、信号处理<sup>[3]</sup>和特征选择<sup>[4]</sup>等领域。

近年来相继出现了越来越多的元启发式算法, 常见的算法有: 粒子群优化 (particle swarm optimization, PSO) 算法<sup>[5]</sup>、灰狼优化 (grey wolf optimization, GWO) 算法<sup>[6]</sup>、鲸鱼优化算法 (whale optimization algorithm, WOA)<sup>[7]</sup>、樽海鞘群算法 (salp swarm algorithm, SSA1)<sup>[8]</sup>、正弦余弦算法 (sine cosine algorithm, SCA)<sup>[9]</sup>和麻雀搜索算法 (sparrow search algorithm, SSA2)<sup>[10]</sup>等, 由于樽海鞘群算法和麻雀搜索算法的英文缩略词相同, 本文以 SSA1 表示樽海鞘群算法, SSA2 表示麻雀搜索算法。虽然元启发式算法越来越多, 但都离不开 2 个阶段: 探索 and 开发。探索阶段即算法以高度随机性进行全局搜索, 随着迭代次数的增加, 算法开始在已探索的区域进行深度开发。因

此, 保证探索和开发之间的平衡显得尤为重要。若探索阶段比重过大, 将会导致收敛速度变慢, 相反, 将会导致过早收敛, 易陷入局部极值点。

针对这些问题, 众多学者已经提出一些有效的改进方法: 陈忠云等<sup>[11]</sup>提出一种多子群的共生非均匀高斯变异 SSA1, 较好的平衡了樽海鞘群的探索和开发能力; 刘景森等<sup>[12]</sup>针对 SSA1 求解结果不稳定问题, 引入自适应领导者-跟随者调整策略, 增强了算法的稳定性; 周娇等<sup>[13]</sup>为避免 WOA 早熟收敛的缺陷, 采用猫映射混沌序列结合反向解方法取代随机产生初始种群, 使 WOA 在初始种群多样性及寻解遍历性上有所增强; 周璟<sup>[14]</sup>采用 Tent 混沌映射改进了狼群初始化方法, 使狼群初始分布更加均匀, 加强了算法的全局搜索能力; 文献<sup>[15]</sup>在 WOA 的惯性权重中引入随机因子, 然后将 Levy 飞行的随机搜索模式引入算法中, 进一步增强了 WOA 的全局和局部搜索能力; 文献<sup>[16]</sup>对烟花算法进行基于对立的学习初始化种群, 对非最优个体采用自适应  $t$  分布变异, 对最优个体采用精英对立学习, 减少了算法的运行时间; 何庆等<sup>[17]</sup>采用柯西变异提高算

收稿日期: 2021-05-31; 录用日期: 2021-08-29; 网络出版时间: 2021-09-15 14:46

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210915.0903.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210915.0903.001.html)

基金项目: 国家自然科学基金 (61703426); 中国博士后科学基金 (2018M633680); 陕西省高校科协青年人才托举计划 (20190108)

\*通信作者. E-mail: [lichenghai\\_ns@163.com](mailto:lichenghai_ns@163.com)

**引用格式:** 唐延强, 李成海, 宋亚飞, 等. 自适应变异麻雀搜索优化算法 [J]. 北京航空航天大学学报, 2023, 49 (3): 681-692.  
TANG Y Q, LI C H, SONG Y F, et al. Adaptive mutation sparrow search optimization algorithm [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 681-692 (in Chinese).

法跳出局部极值点的能力。

SSA2是2020年新提出的一种元启发式算法,该算法将搜索群体分为探索者、跟随者和预警者3部分,其相互分工寻找最优值。虽然易于实现,但如何调节各部分之间的控制参数,以及如何保证3部分之间可以较好的相互配合是一个必须考虑的问题。目前有一些对SSA2的改进:吕鑫等<sup>[18]</sup>将麻雀种群通过Tent混沌序列初始化,同时通过高斯变异和Tent混沌扰动对个体进行变异和扰动,克服了SSA2易陷入局部极值点的缺陷;文献<sup>[19]</sup>利用重心反向学习机制初始化种群,使种群具有更好的空间解分布,其次在探索者的位置更新部分引入学习系数,提高算法的全局搜索能力,最后使用变异算子来改善加入者的位置更新;文献<sup>[20]</sup>引入混沌策略来增强算法种群的多样性,并采用自适应惯性权重来平衡算法的收敛速度和探索能力。这些改进措施在一定程度上提高了算法的寻优性能,但都没有更加深入考虑探索者和跟随者之间的关系。基于此,本文提出一种自适应变异麻雀搜索算法(adaptive mutation sparrow search optimization algorithm, AMSSA),首先通过猫映射混沌序列对种群初始化,提高麻雀个体的随机性及遍历性,再通过柯西变异和Tent混沌扰动对个体进行调整,避免种群过于“集中”或“分散”,最后引入自适应探索者-跟随者数量调整式来平衡寻找全局和局部最优的能力。通过对16个基准函数进行测试,并进行Wilcoxon秩和检验分析,验证了本文所提算法的有效性和可行性。

## 1 麻雀搜索算法

SSA2的思想来自于麻雀种群的觅食行为和反捕食行为,可抽象为探索者-跟随者-预警者模型。探索者具有较高的能量储备,适应度值高,主要为跟随者提供觅食区域和方向。跟随者跟随适应度值最优的探索者寻找食物以获得自己的能量储备,增加自身的适应度值。部分跟随者也可能不断地监视探索者,从而争夺食物。预警者在意识到危险时会发出警报,同时迅速向安全区域移动来获得更好的位置,处于种群中间的麻雀则随机行走靠近别的麻雀,即反捕食行为。同时若警报值大于安全阈值时,探索者需要将所有的跟随者带离危险区域。

在SSA2中,假设有 $N$ 只麻雀在一个 $D$ 维搜索空间中,则每只麻雀的位置为

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_{1,1} & \mathbf{x}_{1,2} & \cdots & \mathbf{x}_{1,d} & \cdots & \mathbf{x}_{1,D} \\ \mathbf{x}_{2,1} & \mathbf{x}_{2,2} & \cdots & \mathbf{x}_{2,d} & \cdots & \mathbf{x}_{2,D} \\ \vdots & \vdots & & \vdots & & \vdots \\ \mathbf{x}_{i,1} & \mathbf{x}_{i,2} & \cdots & \mathbf{x}_{i,d} & \cdots & \mathbf{x}_{i,D} \\ \vdots & \vdots & & \vdots & & \vdots \\ \mathbf{x}_{N,1} & \mathbf{x}_{N,2} & \cdots & \mathbf{x}_{N,d} & \cdots & \mathbf{x}_{N,D} \end{bmatrix} \quad (1)$$

式中: $i=1,2,\dots,N$ ;  $d=1,2,\dots,D$ ;  $\mathbf{x}_{i,d}$ 为第 $i$ 只麻雀在第 $d$ 维的位置。

由于探索者引导着整个麻雀群体的流动,可在任何地方寻找食物,所以其位置更新为

$$\mathbf{x}_{i,d}^{t+1} = \begin{cases} \mathbf{x}_{i,d}^t \exp\left(\frac{-i}{\alpha \cdot \text{iter}_{\max}}\right) & R_2 < \text{ST} \\ \mathbf{x}_{i,d}^t + QL & R_2 \geq \text{ST} \end{cases} \quad (2)$$

式中: $t$ 为当前迭代次数; $\text{iter}_{\max}$ 为迭代次数的最大值; $\alpha$ 为区间(0,1)的随机数; $Q$ 为服从正态分布的随机数; $L$ 为 $1 \times d$ 的矩阵,其中每个元素均为1; $R_2 \in [0,1]$ ,表示警报值; $\text{ST} \in [0.5,1]$ ,表示安全阈值。当 $R_2 < \text{ST}$ 时,意味着周围没有捕食者,探索者将进入宽搜索模式;如果 $R_2 \geq \text{ST}$ ,则意味着一些麻雀已经发现了天敌,所有麻雀都需要迅速飞到其他安全区域。

跟随者跟随探索者寻找食物,并可能与探索者竞争争夺食物,增加自己的捕食率,其位置更新为

$$\mathbf{x}_{i,d}^{t+1} = \begin{cases} Q \exp\left(\frac{\mathbf{x}_{\text{worst}_d^t} - \mathbf{x}_{i,d}^t}{i^2}\right) & i > N/2 \\ \mathbf{x}_{\text{best}_d^{t+1}} + \frac{1}{D} \sum_{d=1}^D (|\mathbf{x}_{i,d}^t - \mathbf{x}_{\text{best}_d^{t+1}}| \cdot \text{rand}\{-1,1\}) & i \leq N/2 \end{cases} \quad (3)$$

式中: $\mathbf{x}_{\text{worst}_d^t}$ 为第 $t$ 次迭代时麻雀在第 $d$ 维的全局最差位置; $\mathbf{x}_{\text{best}_d^{t+1}}$ 为第 $t+1$ 次迭代时探索者在第 $d$ 维的最佳位置; $\text{rand}$ 为给定范围内随机赋值。当 $i > N/2$ 时,表示适应度较差的第 $i$ 个跟随者最有可能饿死;否则第 $i$ 个跟随者在探索者所处的最佳位置附近随机找一个位置觅食。

假设预警者麻雀约占麻雀种群的10%~20%,其初始位置随机确定,数学模型可表示为

$$\mathbf{x}_{i,d}^{t+1} = \begin{cases} \mathbf{x}_{\text{best}_d^t} + \beta |\mathbf{x}_{i,d}^t - \mathbf{x}_{\text{best}_d^t}| & f_i \neq f_g \\ \mathbf{x}_{i,d}^t + K \left( \frac{|\mathbf{x}_{i,d}^t - \mathbf{x}_{\text{worst}_d^t}|}{|f_i - f_w| + \varepsilon} \right) & f_i = f_g \end{cases} \quad (4)$$

式中: $\beta$ 为均值为0、方差为1的随机正态分布,表示步长控制参数; $K$ 为[-1,1]之间的随机数; $f_i$ 、 $f_g$ 和 $f_w$ 分别为当前麻雀的适应度值、当前全局最佳适应度值和最差适应度值; $\varepsilon$ 为最小常数,以避免零分误差。 $f_i \neq f_g$ 表示麻雀位于种群的边缘;当 $f_i = f_g$ 时,表明处于种群中间的麻雀意识到危险,需要移动到其他地方。

## 2 改进麻雀搜索算法

### 2.1 猫映射混沌初始化种群

混沌作为一种非线性的自然现象,以其混沌序列具有遍历性、随机性等优点,被广泛用于优化搜索问题。为了保持种群的多样性,使个体之间尽可

能均匀分布, 本文采用混沌序列初始化策略代替 SSA2 算法中随机生成种群的方法。目前在优化领域形成了多种混沌映射, 包括 Logistic 映射、Tent 映射和猫映射等。Logistic 映射作为一种典型的混沌系统, 其序列的概率密度函数由于服从切比雪夫分布, 映射点呈现出中间密度低, 两边密度高的特点, 导致遍历不均匀性, 影响算法的搜索效率。文献 [21] 提出 Tent 映射的遍历均匀性和收敛速度皆优于 Logistic 映射, 然而 Tent 映射易在小循环周期和不动点上出现问题, 只有最优解仅为边缘值时才可得最优解。针对 Logistic 映射和 Tent 映射的缺点, 本文通过猫映射来生成 SSA2 算法的初始种群。猫映射表达式为

$$\begin{bmatrix} y_{i+1} \\ w_{i+1} \end{bmatrix} = \begin{bmatrix} 1 & a_1 \\ b_1 & a_1 b_1 + 1 \end{bmatrix} \begin{bmatrix} y_i \\ w_i \end{bmatrix} \bmod 1 \quad (5)$$

式中:  $a_1$  和  $b_1$  为任意实数;  $\bmod 1$  为求  $a_1$  小数部分。由于猫映射结构简单, 不易陷入小循环周期和不动点, 通过该映射产生初始种群具有更好的遍历均匀性。

根据猫映射的特点, 在可行域中产生混沌序列并结合反向解初始化的方法步骤如下。

随机产生当前种群的一个可行解, 记为

$$\{Y_i = [y_{i1}, y_{i2}, \dots, y_{id}, \dots, y_{iD}]; y_{id} \in [lb_{id}, ub_{id}]\} \quad (6)$$

则反向解为

$$\begin{cases} Y' = [Y'_1, Y'_2, \dots, Y'_d, \dots, Y'_D] \\ y'_{id} = q(lb_{id} - ub_{id}) - y_{id} \end{cases} \quad (7)$$

式中:  $q$  为区间  $[0, 1]$  上均匀分布实数;  $lb_{id}$ 、 $ub_{id}$  分别为可行解的上、下界。

## 2.2 Tent 混沌和柯西变异扰动策略

### 2.2.1 Tent 混沌扰动

在 2.1 节中提到, Tent 混沌映射存在小周期和不稳定周期点, 为避免落入小周期点和不稳定周期点, 文献 [22] 在原有的 Tent 映射表达式上引入随机变量  $\text{rand}(0, 1) \frac{1}{N}$ , 改进后的 Tent 混沌映射位置表达式为

$$z_{i+1} = \begin{cases} \frac{2z_i + \text{rand}(0, 1)}{N} & 0 \leq z_i \leq \frac{1}{2} \\ \frac{2(1 - z_i) + \text{rand}(0, 1)}{N} & \frac{1}{2} < z_i \leq 1 \end{cases} \quad (8)$$

贝努利移位变换后的表达式为

$$z_{i+1} = 2z_i \bmod 1 + \frac{\text{rand}(0, 1)}{K_p} \quad (9)$$

式中:  $K_p$  为序列内粒子个数。

### 2.2.2 柯西变异

柯西变异来源于连续型概率分布的柯西分布<sup>[23]</sup>, 主要特点为 0 处峰值较小, 从峰值到 0 下降缓慢, 使变异范围更均匀, 变异式为

$$\text{mutation}(x) = x(1 + \tan(\pi(u - 0.5))) \quad (10)$$

式中:  $x$  为原来个体位置;  $\text{mutation}(x)$  为经过柯西变异后的个体位置;  $u$  为  $(0, 1)$  区间内的随机数。

## 2.3 改进探索者位置更新公式

在 SSA2 算法中, 麻雀探索者只受上一代探索者位置的影响,  $\exp\left(\frac{-i}{\alpha \cdot \text{iter}_{\max}}\right)$  的值在不断迭代过程中自适应降低, 其值较大时, 探索者进入广泛搜索模式, 随着其值降低, 探索者主要进行窄搜索模式, 即在最优解附件深度挖掘, 提高算法的收敛精度。因此,  $\exp\left(\frac{-i}{\alpha \cdot \text{iter}_{\max}}\right)$  的值尤为重要。由于  $\exp\left(\frac{-i}{\alpha \cdot \text{iter}_{\max}}\right)$  值的微小改变对探索者的影响较大, 因此, 本文在式 (2) 的基础上, 将探索者更新式改为

$$x'_{i,d} = \begin{cases} x'_{i,d} \cdot \frac{2}{\exp\left(\frac{4i}{\alpha \cdot \text{iter}_{\max}}\right)^m} & R_2 < ST \\ x'_{i,d} + QL & R_2 \geq ST \end{cases} \quad (11)$$

令自适应参数  $c = \frac{2}{\exp\left(\frac{4i}{\alpha \cdot \text{iter}_{\max}}\right)^m}$ ,  $m \in [1, 4]$ , 分

析参数  $m$  对探索者性能的影响, 如图 1 所示。

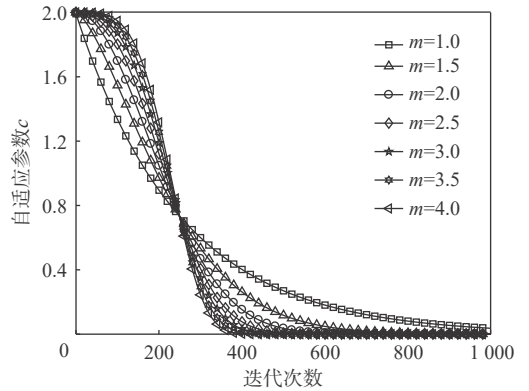


图 1 c 变化曲线

Fig. 1 Variation curves of  $c$

$m$  值的选取影响着麻雀探索者对于全局和局部搜索之间的平衡关系。从图 1 可以看出, 每条曲线都是前期收敛速度快, 后期收敛速度慢, 为了选取最合适的  $m$  值, 选取 Schaffer 函数对该参数进行测试, Schaffer 函数的特点是在全局最优点 3.14 范围左右存在大量的局部极小点, 函数强烈振荡。经过 7 次改变  $m$  值, 每次进行 50 次寻优计算取平均值, 得出  $m$  值对 SSA2 算法的影响如表 1 所示。可以看出, 由于 SSA2 算法本身良好的寻优性能, 可以避过 Schaffer 函数的局部极小点, 寻找到最佳值 0, 因此,  $m$  值对 SSA2 算法的影响主要体现在收敛次数上, 当  $m$  值增加时, 平局收敛次数先减少后增加,  $m$  值为 2 时平均收敛迭代次数最小。由表 1 可知,

$m=2$  时, 参数  $c$  可以使 SSA2 算法的探索能力和觅食能力达到较好的平衡, 此时探索者可在前期充分广泛搜索目标, 后期着重对最优位置进行挖掘。

表 1 参数  $m$  对 SSA2 的影响

Table 1 Influence of parameter  $m$  on SSA2

$m$	最佳值	平均值	标准差	平均收敛次数
1.0	0	$5.509 \times 10^{-72}$	$3.120 \times 10^{-71}$	891
1.5	0	0	0	723
2.0	0	0	0	151
2.5	0	$7.505 \times 10^{-35}$	$4.747 \times 10^{-34}$	289
3.0	0	0	0	224
3.5	0	0	0	557
4.0	0	0	0	292

2.4 探索者-跟随者自适应调整策略

在 SSA2 算法中, 探索者和跟随者的数目比例保持不变, 这会导致在迭代前期, 探索者的数目相对较少, 无法对全局进行充分的搜索, 在迭代后期, 探索者的数目又相对较多, 此时已不需要更多的探索者进行全局搜索, 而需要增加跟随者的数量进行

精确的局部搜索。为解决这个问题, 本文提出探索者-跟随者自适应调整策略, 该策略在迭代前期, 探索者可以占种群数目的多数, 随着迭代次数的增加, 探索者的数目自适应减少, 跟随者的数目自适应增加, 逐步从全局搜索转为局部精确搜索, 从整体上提高算法的收敛精度。探索者和跟随者数目调整为

$$r = b \left( \tan \left( -\frac{\pi t}{4 \text{iter}_{\max}} + \frac{\pi}{4} \right) - k \cdot \text{rand}(0,1) \right) \quad (12)$$

$$p\text{Num} = rN \quad (13)$$

$$s\text{Num} = (1 - r)N \quad (14)$$

式中:  $p\text{Num}$  为探索者数目;  $s\text{Num}$  为跟随者数目;  $b$  为比例系数, 用于控制探索者和跟随者之间的数目;  $k$  为扰动偏离因子, 对非线性递减值  $r$  进行扰动。

2.5 改进后麻雀搜索算法

AMSSA 流程如图 2 所示。

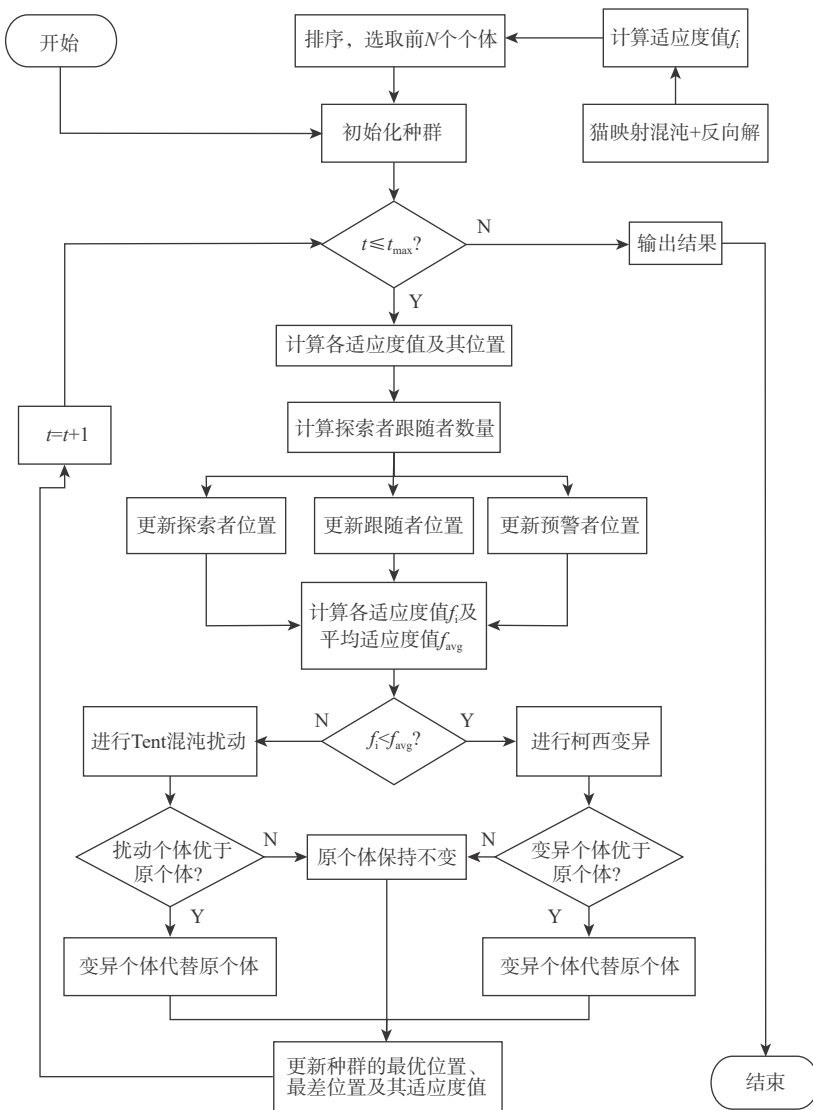


图 2 AMSSA 流程

Fig. 2 Flow chart of AMSSA

### 3 本文所提算法性能测试

为验证本文所提算法的寻优能力及可行性, 将 AMSSA 与 PSO、GWO、WOA、SSA1、SSA2 算法同时在 16 个基准函数上进行对比测试。

#### 3.1 测试函数

具体基准函数如表 2 所示。其中  $f_1 \sim f_6$  为高维单峰函数,  $f_7 \sim f_{11}$  为高维多峰函数,  $f_{12} \sim f_{16}$  为低维多峰函数。高维单峰函数只有一个全局最优点, 没有局部极值点, 主要为了测试函数的收敛速度。多峰函数有多个局部极值点, 从高维和低维 2 种角度出发, 用来观察函数在不同维度跳出局部极值点的性能。

### 3.2 结果分析

#### 3.2.1 收敛精度及稳定性分析

在 MATLAB R2016b 环境下对 6 种算法进行仿真对比实验, 为了避免偶然误差过大, 在实验中选择各基准函数独立运行 40 次, 取最佳值、平均值和标准差作为评价指标, 实验中设定种群规模为 30, 最大迭代次数为 500, 最终得到表 3。

首先, 在 6 个高维单峰函数中, AMSSA 在求解  $f_1$ 、 $f_2$  和  $f_3$  函数时寻找到理论最优值 0,  $f_1$  和  $f_3$  函数标准差也为 0, 除此之外只有 SSA2 在求解  $f_1$  时得到最优值 0; 其次, 在求解  $f_4$  和  $f_5$  函数时, AMSSA 得出的最优值、平均值和标准差比其他 5 种算法提升了至少 2 个数量级; 在求解  $f_6$  函数时, 虽然提升程

表 2 基准函数

Table 2 Benchmark function

编号	函数名称	函数式	维度	定义域	最佳值
$f_1$	Sphere	$f_1(x) = \sum_{i=1}^N x_i^2$	30	$[-100,100]$	0
$f_2$	Schwefel's	$f_2(x) = \sum_{i=1}^N  x_i  + \prod_{i=1}^N  x_i $	30	$[-10,10]$	0
$f_3$	Quadric	$f_3(x) = \sum_{i=1}^N \left( \sum_{j=1}^N x_j \right)^2$	30	$[-100,100]$	0
$f_4$	Rosenbrock	$f_4(x) = \sum_{i=1}^{N-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$	30	$[-30,30]$	0
$f_5$	Step	$f_5(x) = \sum_{i=1}^N ( x_i + 0.5 )^2$	30	$[-100,100]$	0
$f_6$	Quartic	$f_6(x) = \sum_{i=1}^N ix_i^4 + \text{random}[0,1]$	30	$[-1.28,1.28]$	0
$f_7$	Schwefel	$f_7(x) = \sum_{i=1}^N -x_i \sin(\sqrt{ x_i })$	30	$[-500,500]$	-418.982 9N
$f_8$	Rastrigrin	$f_8(x) = \sum_{i=1}^N [x_i^2 - 10 \cos(2\pi x_i) + 10]$	30	$[-5.12,5.12]$	0
$f_9$	Ackley	$f_9(x) = -20 \exp\left(-0.2 \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}\right) - \exp\left(\frac{1}{N} \sum_{i=1}^N \cos(2\pi x_i)\right) + 20 + e$	30	$[-32,32]$	0
$f_{10}$	Griewing	$f_{10}(x) = \frac{1}{4000} \sum_{i=1}^N x_i^2 - \prod_{i=1}^N \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	30	$[-600,600]$	0
$f_{11}$	Generalized penalized	$f_{11}(x) = \frac{\pi}{N} \left\{ 10 \sin^2(\pi y_1) + \sum_{i=1}^{N-1} (y_i - 1)^2 [1 + 10 \sin^2(\pi y_{i+1})] \right\} + (y_N - 1)^2$	30	$[-50,50]$	0
$f_{12}$	Foxholes	$f_{12}(x) = \left( \frac{1}{500} + \sum_{j=1}^{25} \frac{1}{j + \sum_{i=1}^2 (x_i - a_{ij})^6} \right)^{-1}$	2	$[-65,65]$	1
$f_{13}$	Hartmann 6-D	$f_{13}(x) = -\sum_{i=1}^4 c_i \exp\left(-\sum_{j=1}^6 a_{ij} (x_j - p_{ij})^2\right)$	6	$[0,1]$	-3.322 37
$f_{14}$	Schkel	$f_{14}(x) = -\sum_{i=1}^{10} [(X - a_i)(X - a_i)^T + c_i]^{-1}$	4	$[0,10]$	-10.536 3
$f_{15}$	Six-Hump Camel	$f_{15}(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 - 4x_2^4$	2	$[-5,5]$	-1.031 6
$f_{16}$	Kowalik	$f_{16}(x) = \sum_{i=1}^{11} \left[ a_i - \frac{x_1 (b_i^2 + b_i x_2)}{b_i^2 + b_i x_3 + x_4} \right]^2$	4	$[-5,5]$	0.000 307

度不高, 但最优值和稳定性仍高于其他算法。在 5 个高维多峰函数中, 对于  $f_7$  的求解, 6 种算法的寻优能力基本相同, AMSSA 并未表现出算法的优越性; 对于  $f_8$  和  $f_{10}$  的求解, 在 WOA、SSA2 中已经可

以达到最优值, 改进后 AMSSA 中保持了原算法的性能; 对于  $f_9$  的求解, AMSSA 并未突出其优越性, 同 SSA2 一样, 在达到某个极值点后就跳不出该极值点, 但标准差为 0, 稳定性强; 高维多峰函数中只

表3 基准函数测试结果对比

Table 3 Comparison of benchmark function test results

编号	算法	最佳值	平均值	标准值	编号	算法	最佳值	平均值	标准值
$f_1$	PSO	$1.057 \times 10^{-5}$	$1.422 \times 10^{-4}$	$2.013 \times 10^{-4}$	$f_9$	PSO	$2.200 \times 10^{-3}$	$1.381 \times 10^{-1}$	$3.759 \times 10^{-1}$
	GWO	$2.026 \times 10^{-29}$	$1.522 \times 10^{-27}$	$1.998 \times 10^{-27}$		GWO	$7.550 \times 10^{-14}$	$1.021 \times 10^{-13}$	$1.668 \times 10^{-14}$
	WOA	$2.579 \times 10^{-87}$	$6.403 \times 10^{-73}$	$2.573 \times 10^{-72}$		WOA	$8.882 \times 10^{-16}$	$4.530 \times 10^{-15}$	$2.955 \times 10^{-15}$
	SSA1	$2.628 \times 10^{-8}$	$1.302 \times 10^{-7}$	$1.163 \times 10^{-7}$		SSA1	$9.313 \times 10^{-1}$	2.648	1.245
	SSA2	0	$1.227 \times 10^{-51}$	$7.761 \times 10^{-51}$		SSA2	$8.882 \times 10^{-16}$	$8.882 \times 10^{-16}$	0
	AMSSA	0	$1.010 \times 10^{-182}$	0		AMSSA	$8.882 \times 10^{-16}$	$8.882 \times 10^{-16}$	0
$f_2$	PSO	$5.100 \times 10^{-3}$	$4.260 \times 10^{-2}$	$6.160 \times 10^{-2}$	$f_{10}$	PSO	$1.354 \times 10^{-6}$	$1.030 \times 10^{-1}$	$9.200 \times 10^{-3}$
	GWO	$2.312 \times 10^{-17}$	$8.104 \times 10^{-17}$	$3.908 \times 10^{-17}$		GWO	0	$3.200 \times 10^{-3}$	$8.000 \times 10^{-3}$
	WOA	$4.803 \times 10^{-58}$	$1.678 \times 10^{-51}$	$7.650 \times 10^{-51}$		WOA	0	0	0
	SSA1	$4.260 \times 10^{-2}$	2.371	1.860		SSA1	$6.545 \times 10^{-4}$	$1.330 \times 10^{-2}$	$1.050 \times 10^{-2}$
	SSA2	$1.129 \times 10^{-118}$	$5.065 \times 10^{-31}$	$3.185 \times 10^{-30}$		SSA2	0	0	0
	AMSSA	0	$4.831 \times 10^{-35}$	$2.656 \times 10^{-34}$		AMSSA	0	0	0
$f_3$	PSO	$2.569 \times 10$	$7.363 \times 10$	$3.642 \times 10$	$f_{11}$	PSO	$9.907 \times 10^{-8}$	$1.040 \times 10^{-2}$	$3.930 \times 10^{-2}$
	GWO	$2.630 \times 10^{-5}$	$7.351 \times 10^{-5}$	$1.307 \times 10^{-5}$		GWO	$1.280 \times 10^{-2}$	$4.300 \times 10^{-2}$	$1.470 \times 10^{-2}$
	WOA	$3.610 \times 10^3$	$4.822 \times 10^4$	$1.793 \times 10^4$		WOA	$4.500 \times 10^{-3}$	$2.920 \times 10^{-2}$	$4.220 \times 10^{-2}$
	SSA1	$2.330 \times 10^2$	$1.467 \times 10^3$	$7.677 \times 10^2$		SSA1	2.115	6.832	3.754
	SSA2	$2.774 \times 10^{-5}$	$4.129 \times 10^{-5}$	$1.628 \times 10^{-5}$		SSA2	$2.785 \times 10^{-16}$	$9.587 \times 10^{-13}$	$3.536 \times 10^{-12}$
	AMSSA	0	$1.168 \times 10^{-167}$	0		AMSSA	$1.335 \times 10^{-22}$	$3.992 \times 10^{-15}$	$8.391 \times 10^{-15}$
$f_4$	PSO	$1.563 \times 10$	$9.915 \times 10$	$5.719 \times 10$	$f_{12}$	PSO	$9.980 \times 10^{-1}$	3.635	2.619
	GWO	$2.610 \times 10$	$2.693 \times 10$	$6.457 \times 10^{-1}$		GWO	$9.980 \times 10^{-1}$	3.792	3.810
	WOA	$2.707 \times 10$	$2.798 \times 10$	$4.642 \times 10^{-1}$		WOA	$9.980 \times 10^{-1}$	2.838	3.215
	SSA1	$2.430 \times 10$	$1.659 \times 10^2$	$2.593 \times 10^2$		SSA1	$9.980 \times 10^{-1}$	1.097	$3.762 \times 10^{-1}$
	SSA2	$6.711 \times 10^{-9}$	$3.419 \times 10^{-5}$	$1.144 \times 10^{-4}$		SSA2	$9.980 \times 10^{-1}$	4.847	5.242
	AMSSA	$1.658 \times 10^{-13}$	$8.871 \times 10^{-7}$	$3.077 \times 10^{-6}$		AMSSA	$9.980 \times 10^{-1}$	9.611	4.962
$f_5$	PSO	$1.008 \times 10^{-5}$	$1.894 \times 10^{-4}$	$3.878 \times 10^{-4}$	$f_{13}$	PSO	-3.322	-3.274	$5.900 \times 10^{-2}$
	GWO	$7.462 \times 10^{-5}$	$7.011 \times 10^{-1}$	$3.678 \times 10^{-1}$		GWO	-3.322	-3.254	$8.450 \times 10^{-2}$
	WOA	$7.440 \times 10^{-2}$	$4.005 \times 10^{-1}$	$2.182 \times 10^{-1}$		WOA	-3.322	-3.209	$1.158 \times 10^{-1}$
	SSA1	$2.344 \times 10^{-8}$	$2.011 \times 10^{-7}$	$2.792 \times 10^{-7}$		SSA1	-3.322	-3.220	$6.150 \times 10^{-2}$
	SSA2	$1.179 \times 10^{-14}$	$1.538 \times 10^{-11}$	$3.735 \times 10^{-11}$		SSA2	-3.322	-3.280	$5.740 \times 10^{-2}$
	AMSSA	$2.101 \times 10^{-23}$	$7.744 \times 10^{-15}$	$1.533 \times 10^{-14}$		AMSSA	-3.322	-3.216	$3.114 \times 10^{-2}$
$f_6$	PSO	$4.340 \times 10^{-2}$	$1.793 \times 10^{-1}$	$5.150 \times 10^{-2}$	$f_{14}$	PSO	-1.054 × 10	-9.165	2.805
	GWO	$3.836 \times 10^{-4}$	$1.900 \times 10^{-3}$	$8.773 \times 10^{-4}$		GWO	-1.054 × 10	-1.033 × 10	1.283
	WOA	$3.416 \times 10^{-5}$	$2.700 \times 10^{-3}$	$3.400 \times 10^{-3}$		WOA	-1.054 × 10	-7.119	3.386
	SSA1	$6.170 \times 10^{-2}$	$1.766 \times 10^{-1}$	$6.360 \times 10^{-2}$		SSA1	-1.054 × 10	-8.422	3.343
	SSA2	$8.384 \times 10^{-5}$	$1.700 \times 10^{-3}$	$1.400 \times 10^{-3}$		SSA2	-1.054 × 10	-8.508	2.652
	AMSSA	$2.468 \times 10^{-6}$	$8.001 \times 10^{-4}$	$7.674 \times 10^{-4}$		AMSSA	-1.054 × 10	-1.006 × 10	$6.803 \times 10^{-15}$
$f_7$	PSO	$-7.082 \times 10^3$	$-4.601 \times 10^3$	$1.108 \times 10^3$	$f_{15}$	PSO	-1.032	-1.032	$2.043 \times 10^{-16}$
	GWO	$-7.586 \times 10^3$	$-5.865 \times 10^3$	$9.074 \times 10^2$		GWO	-1.032	-1.032	$2.281 \times 10^{-8}$
	WOA	$-1.257 \times 10^4$	$-1.057 \times 10^4$	$1.769 \times 10^3$		WOA	-1.032	-1.032	$5.898 \times 10^{-10}$
	SSA1	$-9.017 \times 10^3$	$-7.584 \times 10^3$	$6.607 \times 10^2$		SSA1	-1.032	-1.032	$3.233 \times 10^{-14}$
	SSA2	$-9.618 \times 10^3$	$-8.525 \times 10^3$	$5.415 \times 10^2$		SSA2	-1.032	-1.032	$1.067 \times 10^{-16}$
	AMSSA	$-8.839 \times 10^3$	$-6.541 \times 10^3$	$6.724 \times 10^2$		AMSSA	-1.032	-1.032	$2.073 \times 10^{-18}$
$f_8$	PSO	$3.609 \times 10$	$6.026 \times 10$	$1.460 \times 10$	$f_{16}$	PSO	$3.275 \times 10^{-4}$	$8.612 \times 10^{-4}$	$1.552 \times 10^{-4}$
	GWO	0	3.174	4.412		GWO	$3.075 \times 10^{-4}$	$3.900 \times 10^{-3}$	$7.700 \times 10^{-3}$
	WOA	0	0	0		WOA	$3.229 \times 10^{-4}$	$7.261 \times 10^{-4}$	$4.383 \times 10^{-4}$
	SSA1	$2.388 \times 10$	$5.313 \times 10$	$2.077 \times 10$		SSA1	$4.024 \times 10^{-4}$	$1.400 \times 10^{-3}$	$3.100 \times 10^{-3}$
	SSA2	0	0	0		SSA2	$3.075 \times 10^{-4}$	$3.219 \times 10^{-4}$	$5.463 \times 10^{-5}$
	AMSSA	0	0	0		AMSSA	$3.075 \times 10^{-4}$	$3.075 \times 10^{-4}$	$8.314 \times 10^{-10}$

有对  $f_{11}$  的求解突出体现了 AMSSA 算法寻优性能的提高。对于 5 个低维多峰函数,除  $f_{16}$  外,  $f_{12} \sim f_{15}$  在 6 种算法的求解下均可以达到最优值,说明在低维条件下算法可以更好的寻优;在  $f_{12}$  中,SSA2、AMSSA

相较于其他 4 种算法稳定性不足,但 AMSSA 比 SSA2 稳定性高; $f_{13}$  中 AMSSA 的稳定性仅次于 WOA; $f_{14}$  和  $f_{15}$  中 AMSSA 的稳定性最好;在  $f_{16}$  中,只有 GWO、SSA2 和 AMSSA 这 3 种算法找到到最优解,稳定性

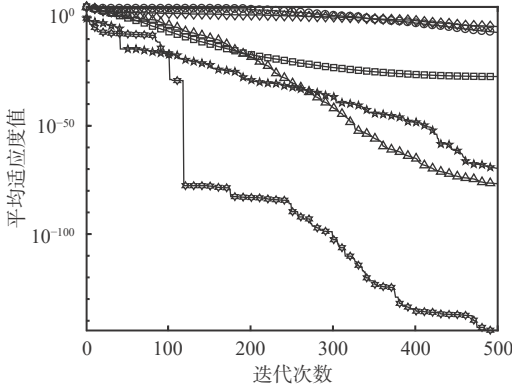
依次增加。

通过分析说明,对于每一种算法,单峰函数的求解精度比多峰函数要高,低维函数比高维函数的求解精度要高,但不论单峰多峰、高维低维函数,AMSSA 相比于其他 5 种算法,不仅在寻优准度上

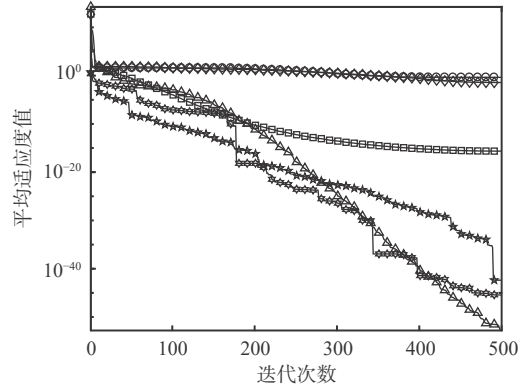
有相对应的提升,并且稳定性更优。

### 3.2.2 收敛曲线分析

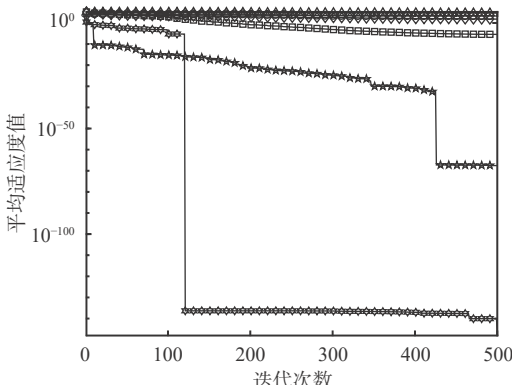
为了直观地比较 6 种算法的收敛速度,再分别用 6 种算法对 16 个基准函数进行实验,结果如图 3 所示。



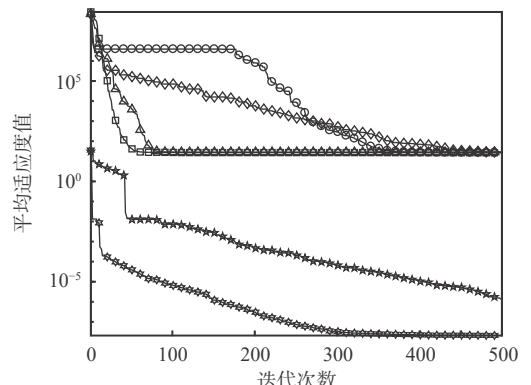
(a)  $f_1$  函数



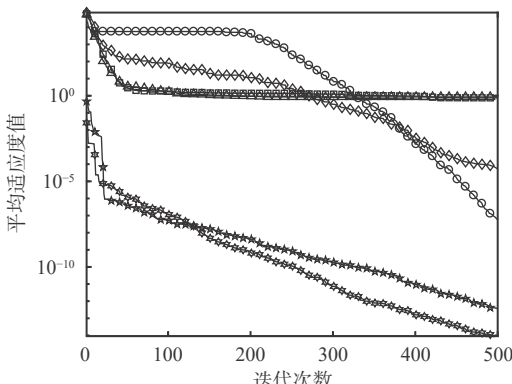
(b)  $f_2$  函数



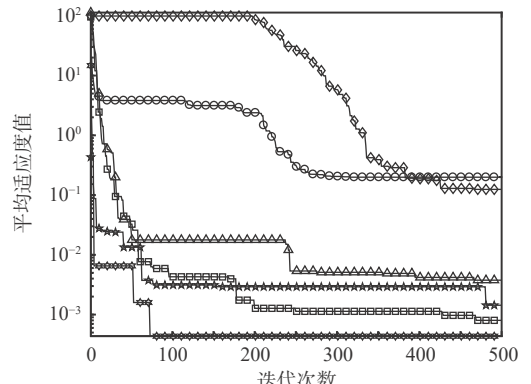
(c)  $f_3$  函数



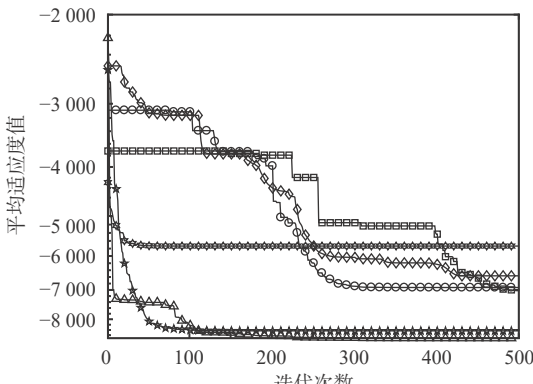
(d)  $f_4$  函数



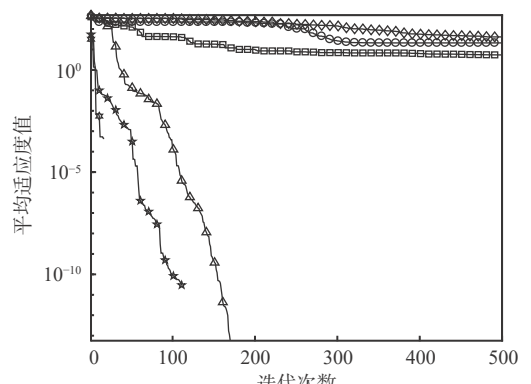
(e)  $f_5$  函数



(f)  $f_6$  函数



(g)  $f_7$  函数



(h)  $f_8$  函数

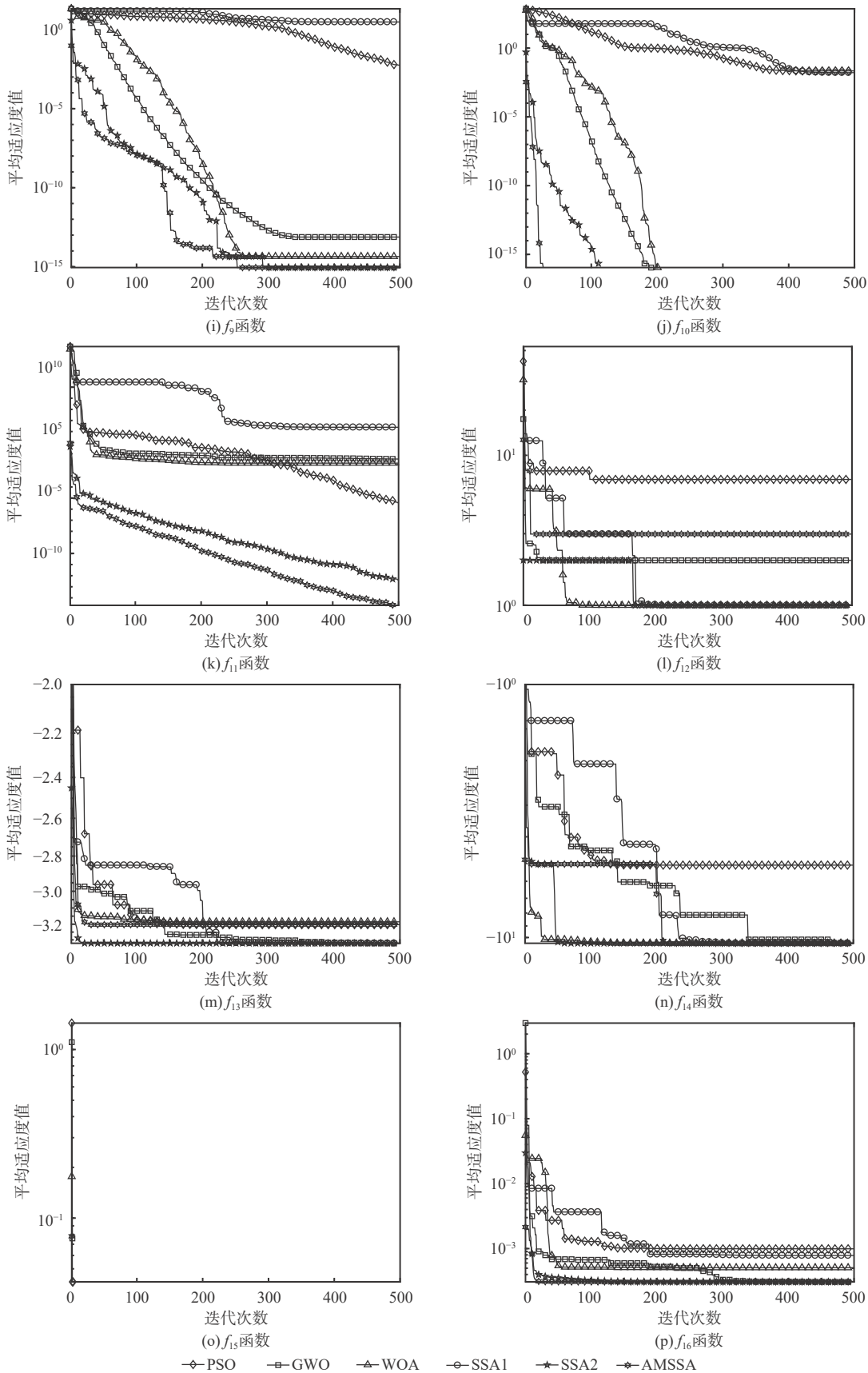


图3 基准函数平均收敛曲线

Fig. 3 Average convergence curve of benchmark function

图3所示可以清晰地表达各算法在寻优过程中适应度值的变化情况。图3(a)~图3(f)中

AMSSA的收敛速度较快,这是因为加入猫映射混沌初始化的结果,并且除图3(b)外,AMSSA的收敛

精度远远高于其他算法。图 3(b) 中 AMSSA 的收敛精度仅次于 WOA, 这与表 3 中的数值刚好对应。其次, 图 3(a)~图 3(f) 中可以清晰的看出有断层现象和 AMSSA 跳出局部极值点的特征, 而 PSO、GWO、WOA、SSA1 基本陷入局部极值点就再未能跳出。在图 3(g)~图 3(k) 高维多峰函数的 5 幅图中, 图 3(g) 中 AMSSA 陷入局部极值点后未跳出, 这也与表格中  $f_7$  函数相对应, 图 3(h)、图 3(j) 和图 3(k) 中表现出 AMSSA 寻优精度和收敛性能明显优于其他算法。同时, 图 3(h) 和图 3(j) 中 AMSSA 的曲线未显示完全, 这是因为已经达到最佳值 0, 算法性能优异。在图 3(l)~图 3(p) 低维多峰的 5 幅图中, 图 3(l)、图 3(m)、图 3(n) 和图 3(o) 中每种算法皆能够寻找到最优解, 图 3(p) 只有 AMSSA、SSA2 和 GWO 寻找到最优解, 并且 AMSSA 收敛速度最快。图 3(o) 中由于  $f_{15}$  函数对每种算法来说都易寻找到最优点, 所以迭代开始每种算法都立刻找到最优点, 图形只有一条竖直的线。

### 3.2.3 Wilcoxon 秩和检验分析

文献 [24] 提出, 对于改进算法性能的评估, 应该进行统计检验。换言之, 仅基于平均值和标准偏差值来比较算法优劣还不够, 需要进行统计检验以证明所提算法比其他算法具有显著的改进优势。通过对每一次的结果进行独立比较, 来体现算法的稳定性和公平性。本文采用 Wilcoxon 秩和检验在 5% 的显著性水平下进行判断 AMSSA 的每次结果是否统计上显著与其他 5 种算法的最佳结果不同, 当  $p$  值小于 5% 时, 拒绝假设, 说明对比算法

具有显著性差异; 否则接受假设, 表明对比算法的寻优能力整体上相同。表 4 给出了 16 个基准函数下 AMSSA 与其他 5 种算法之间的秩和检验  $p$  值, 因为当 2 个对比算法都达到最佳值时, 无法进行比较, 所以表 4 中 NaN 表示“不适用”, 即不能进行显著性判断,  $R$  为显著性判断的结果, “+”、“-”、“=” 分别表示 AMSSA 性能优于、劣于和相当于对比算法。

从表 4 中可以得出, 大部分  $p$  值远远小于 5%, 这说明 AMSSA 的优越性比其他 5 种算法在统计上是显著的。在 AMSSA 与 SSA2 对比,  $f_{13} \sim f_{16}$  函数中  $R$  值为“-”, 这是因为本身 SSA2 寻优性能较好, AMSSA 与 SSA2 皆能寻找到最佳值, 只是在平均值的体现上有所不同, AMSSA 在低维多峰函数上的寻优性能有一定的提升, 但提升空间不大。

### 3.2.4 模型消融实验

由于 AMSSA 的各种添加机制导致其寻优性能比较好, 但具体哪一种添加机制起作用, 是否有某一种添加机制不起作用还未知, 为了表明各种机制的作用, 使实验更具说服力, 本文进行了 AMSSA 模型的消融实验。分别以 SSA、自适应 SSA(ASSA)、变异 SSA(MSSA) 与 AMSSA 在 16 个基准函数上做对比实验, 最终结果如表 5 所示, 其中 Mean 为平均值, Std 为标准值。

从表 5 中可以看出, 在高维单峰函数  $f_1 \sim f_6$  中: 除  $f_6$  外, 其他函数中的 ASSA 与 MSSA 皆对寻优性能起到了一定作用; 在  $f_6$  中, 虽然 ASSA 与 MSSA

表 4 Wilcoxon 秩和检验  $p$  值

Table 4  $p$ -value for Wilcoxon's rank-sum test

函数	PSO		GWO		WOA		SSA1		SSA2	
	$p$	$R$	$p$	$R$	$p$	$R$	$p$	$R$	$p$	$R$
$f_1$	$8.25 \times 10^{-15}$	+	$8.25 \times 10^{-15}$	+	$8.64 \times 10^{-11}$	+	$9.55 \times 10^{-15}$	+	$5.57 \times 10^{-10}$	+
$f_2$	$1.17 \times 10^{-14}$	+	$1.26 \times 10^{-14}$	+	$1.85 \times 10^{-5}$	+	$1.17 \times 10^{-14}$	+	$1.98 \times 10^{-7}$	+
$f_3$	$1.29 \times 10^{-14}$	+	$9.55 \times 10^{-15}$	+	$9.55 \times 10^{-15}$	+	$9.55 \times 10^{-15}$	+	$2.69 \times 10^{-7}$	+
$f_4$	$1.43 \times 10^{-14}$	+	$1.43 \times 10^{-14}$	+	$1.43 \times 10^{-14}$	+	$1.43 \times 10^{-14}$	+	$1.94 \times 10^{-10}$	+
$f_5$	$1.43 \times 10^{-14}$	+	$1.43 \times 10^{-14}$	+	$1.43 \times 10^{-14}$	+	$1.43 \times 10^{-14}$	+	$8.40 \times 10^{-13}$	+
$f_6$	$1.43 \times 10^{-14}$	+	$6.54 \times 10^{-2}$	-	$6.94 \times 10^{-4}$	+	$1.43 \times 10^{-14}$	+	$9.57 \times 10^{-1}$	-
$f_7$	$3.46 \times 10^{-11}$	+	$2.40 \times 10^{-3}$	+	$2.83 \times 10^{-10}$	+	$9.92 \times 10^{-7}$	+	$2.89 \times 10^{-13}$	+
$f_8$	$1.96 \times 10^{-16}$	+	$1.90 \times 10^{-16}$	+	NaN	=	$1.96 \times 10^{-16}$	+	NaN	=
$f_9$	$1.96 \times 10^{-16}$	+	$1.77 \times 10^{-16}$	+	$1.90 \times 10^{-11}$	+	$1.96 \times 10^{-16}$	+	NaN	=
$f_{10}$	$1.96 \times 10^{-16}$	+	$4.18 \times 10^{-4}$	+	NaN	=	$1.96 \times 10^{-16}$	+	NaN	=
$f_{11}$	$1.43 \times 10^{-14}$	+	$1.43 \times 10^{-14}$	+	$1.43 \times 10^{-14}$	+	$1.43 \times 10^{-14}$	+	$5.12 \times 10^{-11}$	+
$f_{12}$	$1.17 \times 10^{-5}$	+	$1.12 \times 10^{-1}$	-	$5.10 \times 10^{-3}$	+	$2.90 \times 10^{-8}$	+	$7.00 \times 10^{-3}$	+
$f_{13}$	$3.77 \times 10^{-2}$	+	$2.27 \times 10^{-8}$	+	$9.88 \times 10^{-12}$	+	$6.32 \times 10^{-13}$	+	$3.40 \times 10^{-1}$	-
$f_{14}$	$1.62 \times 10^{-2}$	+	$7.18 \times 10^{-5}$	+	$2.61 \times 10^{-6}$	+	$9.02 \times 10^{-7}$	+	$5.68 \times 10^{-1}$	-
$f_{15}$	$9.60 \times 10^{-8}$	+	$2.93 \times 10^{-15}$	+	$2.93 \times 10^{-15}$	+	$6.88 \times 10^{-15}$	+	$6.32 \times 10^{-1}$	-
$f_{16}$	$1.43 \times 10^{-14}$	+	$8.66 \times 10^{-12}$	+	$1.05 \times 10^{-13}$	+	$1.79 \times 10^{-14}$	+	$6.68 \times 10^{-1}$	-

表5 模型消融实验结果  
Table 5 Experimental results of model ablation

算法	$f_1$		$f_2$		$f_3$			
	Mean	Std	Mean	Std	Mean	Std		
SSA	$2.526 \times 10^{-60}$	$1.050 \times 10^{-59}$	$1.212 \times 10^{-26}$	$7.105 \times 10^{-26}$	$7.076 \times 10^{-26}$	$3.363 \times 10^{-25}$		
ASSA	$1.690 \times 10^{-75}$	$1.069 \times 10^{-74}$	$9.746 \times 10^{-29}$	$6.164 \times 10^{-28}$	$5.576 \times 10^{-29}$	$1.763 \times 10^{-28}$		
MSSA	$1.679 \times 10^{-67}$	$1.032 \times 10^{-66}$	$4.333 \times 10^{-37}$	$1.978 \times 10^{-36}$	$5.084 \times 10^{-29}$	$1.605 \times 10^{-28}$		
AMSSA	$1.034 \times 10^{-181}$	0	$1.966 \times 10^{-37}$	$1.243 \times 10^{-36}$	$1.026 \times 10^{-155}$	$3.243 \times 10^{-155}$		
算法	$f_4$		$f_5$		$f_6$			
	Mean	Std	Mean	Std	Mean	Std		
SSA	$9.633 \times 10^{-5}$	$2.729 \times 10^{-4}$	$5.003 \times 10^{-11}$	$1.924 \times 10^{-10}$	$1.600 \times 10^{-3}$	$1.300 \times 10^{-3}$		
ASSA	$3.436 \times 10^{-6}$	$8.914 \times 10^{-6}$	$2.327 \times 10^{-14}$	$3.221 \times 10^{-14}$	$1.600 \times 10^{-3}$	$1.800 \times 10^{-3}$		
MSSA	$3.527 \times 10^{-6}$	$7.135 \times 10^{-6}$	$2.300 \times 10^{-14}$	$3.999 \times 10^{-14}$	$2.000 \times 10^{-3}$	$1.800 \times 10^{-3}$		
AMSSA	$4.318 \times 10^{-7}$	$8.616 \times 10^{-7}$	$2.029 \times 10^{-14}$	$3.859 \times 10^{-14}$	$1.400 \times 10^{-3}$	$8.627 \times 10^{-4}$		
算法	$f_7$		$f_8$		$f_9$		$f_{10}$	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std
SSA	$-8.463 \times 10^3$	$5.314 \times 10^2$	0	0	$8.882 \times 10^{-16}$	0	0	0
ASSA	$-8.492 \times 10^3$	$5.921 \times 10^2$	0	0	$8.882 \times 10^{-16}$	0	0	0
MSSA	$-8.726 \times 10^3$	$9.075 \times 10^2$	0	0	$8.882 \times 10^{-16}$	0	0	0
AMSSA	$-8.759 \times 10^3$	$4.052 \times 10^2$	0	0	$8.882 \times 10^{-16}$	0	0	0
算法	$f_{11}$		$f_{12}$		$f_{13}$			
	Mean	Std	Mean	Std	Mean	Std		
SSA	$3.474 \times 10^{-12}$	$1.105 \times 10^{-11}$	6.260	5.454	-3.263	$6.270 \times 10^{-2}$		
ASSA	$6.636 \times 10^{-15}$	$1.377 \times 10^{-14}$	7.620	5.746	-3.274	$6.140 \times 10^{-2}$		
MSSA	$1.356 \times 10^{-15}$	$2.427 \times 10^{-15}$	$9.980 \times 10^{-1}$	$1.480 \times 10^{-16}$	-3.298	$5.010 \times 10^{-2}$		
AMSSA	$1.045 \times 10^{-15}$	$8.414 \times 10^{-16}$	$9.980 \times 10^{-1}$	$1.655 \times 10^{-16}$	-3.274	$6.140 \times 10^{-2}$		
算法	$f_{14}$		$f_{15}$		$f_{16}$			
	Mean	Std	Mean	Std	Mean	Std		
SSA	-8.373	2.793	-1.032	$1.958 \times 10^{-16}$	$3.456 \times 10^{-4}$	$8.718 \times 10^{-5}$		
ASSA	-8.914	2.612	-1.032	$1.958 \times 10^{-16}$	$3.404 \times 10^{-4}$	$1.040 \times 10^{-4}$		
MSSA	-8.914	2.612	-1.032	$2.094 \times 10^{-16}$	$3.215 \times 10^{-4}$	$6.452 \times 10^{-5}$		
AMSSA	-9.455	2.280	-1.032	$1.958 \times 10^{-16}$	$3.075 \times 10^{-4}$	$8.442 \times 10^{-10}$		

相较于 SSA 没有提升效果,但二者结合后,AMSSA 的稳定性最高。在高维多峰函数  $f_7 \sim f_{11}$  中:函数  $f_7$  的测试结果与  $f_6$  类似,稳定性提升,但寻优性能提升效果不大;函数  $f_8 \sim f_{10}$  中所有算法都可以找到最优值;函数  $f_{11}$  中,ASSA 和 MSSA 的效果都比 SSA2 好,同时 AMSSA 的效果最好。在低维多峰函数  $f_{12} \sim f_{16}$  中,由于每种算法可以无限接近最优值,因此 AMSSA 性能提升不大,而且在  $f_{13}$  中,ASSA、MSSA 和 AMSSA 的性能不如 SSA2。

### 3.2.5 AMSSA 时间复杂度分析

假设算法中种群规模为  $N$ ,维度为  $D$ ,最大迭代次数为  $iter_{max}$ ,随机初始化种群参数时间为  $s_1$ ,求个体适应度值的时间为  $j(D)$ ;探索者数量为  $pNum$ ,每一维更新时间为  $s_2$ ;跟随者数量为  $sNum$ ,每一维更新时间为  $s_3$ ;预警者每一维更新时间为  $s_4$ 。因此初

始阶段时间复杂度为:  $T_1 = O(s_1 + N(j(D) + Ds_1))$ ;探索者更新时间复杂度为:  $T_2 = O(pNum \times s_2 D)$ ;跟随者更新时间复杂度为:  $T_3 = O(sNum \times s_3 D)$ ;预警者更新时间复杂度为:  $T_4 = O((N - pNum - sNum)s_4 D)$ 。综上,SSA 总的时间复杂度为:  $T = T_1 + (T_2 + T_3 + T_4)iter_{max} = O(D + j(D))$ 。

在 AMSSA 中,假设猫映射混沌所需时间为  $u_1$ ,排序选择时间为  $u_2$ ,因此,初始阶段时间复杂度为:  $T_{11} = O(s_1 + N(u_1 + j(D) + Ds_1) + u_2)$ ;探索者和跟随者个数更新式为  $u_3$ ,则探索者更新时间复杂度为:  $T_{22} = O(pNum \times s_2 D + iter_{max} u_3)$ ;跟随者更新时间复杂度为:  $T_{33} = O(sNum \times s_3 D + iter_{max} u_3)$ ;预警者更新时间复杂度为:  $T_{44} = O((N - pNum - sNum)s_4 D)$ 。在柯西变异和 Tent 混沌扰动过程中:设求解平均适应度值  $f_{avg}$  时间为  $u_4$ ,计算扰动式时间和柯西变异

式时间分别为  $u_5$ 、 $u_6$ 。麻雀适应度值与平均适应度值比较和择优更新目标位置的时间分别为  $u_7$ 、 $u_8$ ，则这个阶段的时间复杂度为： $T_{55} = O(u_4 + u_5 + u_6 + N(j(D) + u_7) + u_8)$ 。综上，AMSSA的时间复杂度为： $T_1 = T_{11} + (T_{22} + T_{33} + T_{44} + T_{55}) \text{iter}_{\max} = O(D + j(D))$ 。由于  $T_1 = T$ ，表明 AMSSA 的时间复杂度没有增加。

综和本节分析，通过模型消融实验证明所提的每一种策略都对 AMSSA 算法最终的性能有所帮助，通过时间复杂度分析说明 AMSSA 的时间复杂度与 SSA2 等价，通过 AMSSA 在 3 类函数上的实验得出，虽然 AMSSA 相较于 SSA2 在低维多峰函数中的寻优性能不明显，但在高维单峰和高维多峰函数中寻优性能比其他 5 种算法高出至少 2 个数量级，收敛速度也有明显的提升，稳定性增强，充分说明 AMSSA 算法的各种性能总体上优于其他 5 种算法，表明了 AMSSA 的优越性和可行性。

## 4 结论

本文在标准麻雀搜索算法的基础上，引入了混沌映射和柯西变异及探索者跟随者数量的自适应调整策略，提出一种自适应变异麻雀搜索优化算法 AMSSA。

1) 首先，猫映射混沌加快了种群初始化的收敛速度；其次，自适应探索者-跟随者调整策略使算法前期着重寻找全局最优，后期着重局部寻优整体上提高了收敛精度；最后，Tent 混动扰动和柯西变异使算法跳出局部极值点，这使得 AMSSA 相较于 SSA2 寻优精度更高，搜索能力更强。

2) 选取 16 个基准测试函数与其他 5 种算法进行比较，并进行 Wilcoxon 秩和检验对算法显著性水平进行验证。实验表明，AMSSA 寻优性能提升较明显，具有良好的开拓能力，鲁棒性强，表现出良好的实时能力。

下一步考虑将 AMSSA 与其他改进的 SSA 做对比，并将其应用到图像分割、人脸识别等实际工程中，用于解决工程优化领域中的约束优化、多目标优化等问题，从而验证 AMSSA 在实际工程中的优越性。

## 参考文献 (References)

- [1] YANG L N, SUN X, LI Z L. An efficient framework for remote sensing parallel processing: Integrating the artificial bee colony algorithm and multiagent technology[J]. *Remote Sensing*, 2019, 11(2): 152.
- [2] SHAFFIEE HAGHSHEENAS S, PIROUZ B, SHAFFIEE HAGHSHEENAS S, et al. Prioritizing and analyzing the role of climate and urban parameters in the confirmed cases of COVID-19 based on artificial intelligence applications[J]. *International Journal of Environmental Research and Public Health*, 2020, 17(10): 3730.
- [3] DASH J, DAM B, SWAIN R. Implementation of narrow-width automatic digital tuner using popular swarm intelligence technique[J]. *Engineering Applications of Artificial Intelligence*, 2019, 79: 87-99.
- [4] HU P, PAN J S, CHU S C. Improved binary grey wolf optimizer and its application for feature selection[J]. *Knowledge-Based Systems*, 2020, 195: 105746.
- [5] KENNEDY J, EBERHART R. Particle swarm optimization[C]// Proceedings of International Conference on Neural Networks. Piscataway: IEEE Press, 1995: 1942-1948.
- [6] MIRJALILI S, MIRJALILI S M, LEWIS A. Grey wolf optimizer[J]. *Advances in Engineering Software*, 2014, 69: 46-61.
- [7] MIRJALILI S, LEWIS A. The whale optimization algorithm[J]. *Advances in Engineering Software*, 2016, 95: 51-67.
- [8] MIRJALILI S, GANDOMI A H, MIRJALILI S Z. Salp swarm algorithm: A bio-inspired optimizer for engineering design problems[J]. *Advances in Engineering Software*, 2017, 114: 163-191.
- [9] MIRJALILI S. SCA: A sine cosine algorithm for solving optimization problems[J]. *Knowledge-Based Systems*, 2016, 96: 120-133.
- [10] XUE J K, SHEN B. A novel swarm intelligence optimization approach: Sparrow search algorithm[J]. *Systems Science & Control Engineering*, 2020, 8(1): 22-34.
- [11] 陈忠云, 张达敏, 辛梓芸. 多子群的共生非均匀高斯变异樽海鞘群算法[J]. *自动化学报*, 2022, 48(5): 1307-1317.  
CHEN Z Y, ZHANG D M, XIN Z Y. Multi-subpopulation based symbiosis and non-uniform Gaussian mutation salp swarm algorithm[J]. *Acta Automatica Sinica*, 2022, 48(5): 1307-1317(in Chinese).
- [12] 刘景森, 袁蒙蒙, 左方. 面向全局搜索的自适应领导者樽海鞘群算法[J]. *控制与决策*, 2021, 36(9): 2152-2160.  
LIU J S, YUAN M M, ZUO F. Global search-oriented adaptive leader salp swarm algorithm[J]. *Control and Decision*, 2021, 36(9): 2152-2160(in Chinese).
- [13] 周娇, 王力, 陈小青. 基于改进鲸鱼优化算法的最大2维熵图像分割[J]. *激光技术*, 2021, 45(3): 378-385.  
ZHOU J, WANG L, CHEN X Q. Image segmentation of 2-D maximum entropy based on the improved whale optimization algorithm[J]. *Laser Technology*, 2021, 45(3): 378-385(in Chinese).
- [14] 周璟. 混沌狼群围捕算法的车间机器人导航路径规划[J]. *机械设计与制造*, 2020(1): 251-255.  
ZHOU J. Workshop used robot navigation path planning method based on chaotic wolf pack besieging algorithm[J]. *Machinery Design & Manufacture*, 2020(1): 251-255(in Chinese).
- [15] YAN C, LI M X, LIU W. Prediction of bank telephone marketing results based on improved whale algorithms optimizing S\_Kohonen network[J]. *Applied Soft Computing*, 2020, 92: 106259.
- [16] LI X G, HAN S F, GONG C Q. Analysis and improvement of fireworks algorithm[J]. *Algorithms*, 2017, 10(1): 26.
- [17] 何庆, 林杰, 徐航. 混合柯西变异和均匀分布的蝗虫优化算法[J]. *控制与决策*, 2021, 36(7): 1558-1568.  
HE Q, LIN J, XU H. Hybrid Cauchy mutation and uniform distribution of grasshopper optimization algorithm[J]. *Control and Decision*, 2021, 36(7): 1558-1568(in Chinese).

- [18] 吕鑫, 慕晓冬, 张钧, 等. 混沌麻雀搜索优化算法[J]. *北京航空航天大学学报*, 2021, 47(8): 1712-1720.  
LYU X, MU X D, ZHANG J, et al. Chaos sparrow search optimization algorithm[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2021, 47(8): 1712-1720(in Chinese).
- [19] LIU G Y, SHU C, LIANG Z W, et al. A modified sparrow search algorithm with application in 3d route planning for UAV[J]. *Sensors*, 2021, 21(4): 1224.
- [20] YUAN J H, ZHAO Z W, LIU Y P, et al. DMPPT control of photovoltaic microgrid based on improved sparrow search algorithm[J]. *IEEE Access*, 2021, 9: 16623-16629.
- [21] CHEN S R, WANG S Y. An optimization method for an integrated energy system scheduling process based on NSGA-II improved by tent mapping chaotic algorithms[J]. *Processes*, 2020, 8(4): 426.
- [22] 张娜, 赵泽丹, 包晓安, 等. 基于改进的Tent混沌万有引力搜索算法[J]. *控制与决策*, 2020, 35(4): 893-900.  
ZHANG N, ZHAO Z D, BAO X A, et al. Gravitational search algorithm based on improved Tent chaos[J]. *Control and Decision*, 2020, 35(4): 893-900(in Chinese).
- [23] 郭振洲, 王平, 马云峰, 等. 基于自适应权重和柯西变异的鲸鱼优化算法[J]. *微电子学与计算机*, 2017, 34(9): 20-25.  
GUO Z Z, WANG P, MA Y F, et al. Whale optimization algorithm based on adaptive weight and cauchy mutation[J]. *Microelectronics & Computer*, 2017, 34(9): 20-25(in Chinese).
- [24] DERRAC J, GARCIA S, MOLINA D, et al. A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms[J]. *Swarm and Evolutionary Computation*, 2011, 1(1): 3-18.

## Adaptive mutation sparrow search optimization algorithm

TANG Yanqiang<sup>1,2</sup>, LI Chenghai<sup>2,\*</sup>, SONG Yafei<sup>2</sup>, CHEN Chen<sup>3</sup>, CAO Bo<sup>1</sup>

(1. School of Graduate, Air Force Engineering University, Xi'an 710051, China;

2. Air and Missile Defense College, Air Force Engineering University, Xi'an 710051, China;

3. Xi'an Satellite Control Center, Xi'an 710043, China)

**Abstract:** To address the problems that the sparrow search algorithm is prone to fall into local extremum points in the early stage and not high accuracy in the later stage of the search, an adaptive variational sparrow search algorithm (AMSSA) is proposed. Firstly, the initial population is initialized by cat mapping chaotic sequences to enhance the randomness and ergodicity of the initial population and improve the global search ability of the algorithm; Secondly, the Cauchy mutation and Tent chaos disturbance are introduced to expand the local search ability, so that the individuals caught in the local extremum can jump out of the limit and continue the search. Finally, the explorer-follower number adaptive adjustment strategy the adaptive adjustment strategy of explorer-follower number is proposed to enhance the global search ability in the early stage and the local depth mining ability in the later stage of the algorithm by using the change of the explorer and follower numbers in each stage to improve the optimization-seeking accuracy of the algorithm. Sixteen benchmark functions and the Wilcoxon test are selected for validation, and the experimental results show that the AMSSA achieves greater improvement in search accuracy, convergence speed and stability compared with other algorithms.

**Keywords:** sparrow search algorithm; cat map chaos; Cauchy mutation; Tent chaos; adaptive adjustment strategy

Received: 2021-05-31; Accepted: 2021-08-29; Published Online: 2021-09-15 14:46

URL: [kns.cnki.net/kcms/detail/11.2625.V.20210915.0903.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210915.0903.001.html)

Foundation items: National Natural Science Foundation of China (61703426); China Postdoctoral Science Foundation (2018M633680); Young Talent Fund of University Association for Science and Technology in Shaanxi Province (20190108)

\* Corresponding author. E-mail: [lichenghai\\_ns@163.com](mailto:lichenghai_ns@163.com)

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0296

# 基于 M-AFSA 的 MPRM 逻辑电路面积优化

邵艺璇<sup>1,2</sup>, 何振学<sup>1,2,\*</sup>, 周宇豪<sup>1,2</sup>, 霍志胜<sup>3</sup>, 肖利民<sup>3</sup>, 王翔<sup>4</sup>

(1. 河北农业大学 河北省农业大数据重点实验室, 保定 071001; 2. 河北农业大学 信息科学与技术学院, 保定 071001;  
3. 北京航空航天大学 计算机学院, 北京 100191; 4. 北京航空航天大学 电子信息工程学院, 北京 100191)

**摘 要:** 现有基于传统智能优化算法的 MPRM 电路面积优化算法存在效果差的问题。由于 MPRM 电路面积优化属于组合优化问题, 先提出一种多策略协同进化人工鱼群算法 (M-AFSA), 该算法引入基于反向学习的种群初始化策略, 以提高种群多样性及初始种群解的质量; 引入觅食与追尾交互性策略, 以加强人工鱼个体之间的信息交流、提高所提算法的收敛速度; 引入自适应扰动策略, 以增加人工鱼个体位置变异的随机性、避免所提算法陷入局部最优。此外, 提出一种 MPRM 逻辑电路面积优化方法, 利用所提算法来搜索电路面积最小的最佳极性。基于北卡罗莱纳州微电子中心 (MCNC) Benchmark 电路的实验结果表明: 与遗传算法相比, 所提算法优化电路平均面积百分比最高为 57.24%, 平均为 39.57%; 与人工鱼群算法相比, 所提算法优化电路平均面积百分比最高为 33.53%, 平均为 14.54%; 与改进的人工鱼群算法相比, 所提算法优化电路平均面积百分比最高为 30.25%, 平均为 13.86%。

**关键词:** 混合极性 Reed-Muller 电路; 面积优化; 组合优化; 人工鱼群算法; 反向学习

**中图分类号:** V443; TP391.72

**文献标志码:** A **文章编号:** 1001-5965(2023)03-0693-09

数字逻辑电路既可以采用基于“与/或/非”运算的传统布尔逻辑实现, 也可以采用基于“与/异或”运算的 Reed-Muller(RM) 逻辑来实现。基于“与/异或”运算的 RM 逻辑是区别于布尔逻辑的另一种电路表达方式。研究表明, 与传统的布尔逻辑电路相比, 利用 RM 逻辑实现部分电路 (如运算电路、奇偶校验电路等功能电路) 具有更紧凑的结构和良好的可测性<sup>[1-3]</sup>。RM 逻辑分为混合极性 RM (mixed polarity Reed-Muller, MPRM) 逻辑和固定极性 RM (fixed polarity Reed-Muller, FPRM) 逻辑。其中, MPRM 逻辑表达式是 RM 逻辑中常见的一种规范表达形式。对于输入变量数为  $n$  的 Boolean 逻辑电路来说, 具有  $2^n$  个固定极性, 对应  $2^n$  个不同 FPRM

表达式; 具有  $3^n$  个不同的混合极性, 对应  $3^n$  个不同的 MPRM 表达式。显然, 对于同一电路来说, MPRM 逻辑比 FPRM 逻辑具有更大的优化空间。

极性是 RM 电路的重要属性, 决定着 RM 电路表达式的繁简, 进而影响 RM 电路的性能。MPRM 电路面积优化是在极性优化空间中搜索对应某一/组电路性能最好的最佳极性。因此, MPRM 电路面积优化属于组合优化问题。对于中大规模 MPRM 电路来说, 其极性优化空间随输入变量数的增加呈指数型增长, 穷尽搜索法无法在有限的时间内获得最优解。为满足中大规模 MPRM 逻辑电路快速、有效搜索最佳极性的需求, 需结合并行搜索能力更强、优化效率更高的智能优化算法。

收稿日期: 2021-06-03; 录用日期: 2021-09-23; 网络出版时间: 2021-10-29 17:05

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20211029.1450.002.html](https://kns.cnki.net/kcms/detail/11.2625.V.20211029.1450.002.html)

基金项目: 国家自然科学基金青年科学基金 (62102130); 河北省自然科学基金 (F2020204003); 河北省高等学校科学技术研究项目-青年项目 (BJ2019008); 河北农业大学引进人才科研专项 (YJ201829); 中央引导地方科技发展资助项目 (226Z0201G)

\* 通信作者. E-mail: [hezhenxue@buaa.edu.cn](mailto:hezhenxue@buaa.edu.cn)

**引用格式:** 邵艺璇, 何振学, 周宇豪, 等. 基于 M-AFSA 的 MPRM 逻辑电路面积优化 [J]. 北京航空航天大学学报, 2023, 49 (3): 693-701.  
SHAO Y X, HE Z X, ZHOU Y H, et al. Area optimization of MPRM circuits based on M-AFSA [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 693-701 (in Chinese).

智能优化算法在 MPRM 电路面积优化领域得到了广泛应用。文献 [4] 将基于相异度的局部改善策略融入传统遗传算法, 提出了一种基于混合遗传算法的 MPRM 最小化算法, 以避免算法陷于局部最优, 增强了全局搜索能力。文献 [5] 将模拟退火算法与粒子群算法相结合提出一种模拟退火离散粒子群优化 (simulated annealing and discrete particle swarm optimization, SADPSO) 算法, 基于该算法对 MPRM 电路进行面积优化。文献 [6] 提出三值多样性粒子群算法并应用于 MPRM 电路的综合优化, 表现出了一定的性能优势。然而, 由于受传统智能优化算法种群多样性差、收敛速度慢、易陷入局部最优等问题的影响, 使得现有 MPRM 逻辑电路面积优化方法的优化效果有待提高。

人工鱼群算法 (artificial fish swarms algorithm, AFSA) 是一种全新的群智能优化算法, 通过鱼群中人工鱼的觅食、聚群、追尾等行为以搜索全局最优解, 具有鲁棒性强和全局寻优能力强等优点, 在组合优化领域得到了广泛应用<sup>[7-14]</sup>。研究表明, 在求解 NP-Hard 优化问题上, 人工鱼群算法效果明显优于遗传算法和粒子群算法<sup>[15-16]</sup>。

本文针对现有 MPRM 电路面积优化效果差的问题, 提出一种有效的 MPRM 逻辑电路面积优化方法。与现有 MPRM 电路面积优化方法相比, 主要贡献如下:

1) 提出一种多策略协同进化人工鱼群算法 (multi-strategies artificial fish swarms algorithm, M-AFSA), 以用于求解三变量组合优化问题。该算法引入基于反向学习的种群初始化策略以提高种群多样性及初始种群解的质量; 引入提出觅食与追尾交互性策略, 以加强人工鱼个体之间的信息交流提高收敛速度; 引入自适应扰动策略, 以增大算法进化后期人工鱼位置变异的随机性、避免陷入局部最优、增强算法的全局寻优能力。

2) 提出一种 MPRM 逻辑电路面积优化方法, 该方法利用提出的 M-AFSA 算法搜索电路面积最小的最佳极性。

3) 基于北卡罗莱纳州微电子中心 (Microelectronics Center of North Carolina, MCNC) Benchmark 基准测试电路, 所提算法分别与基于遗传算法 (genetic algorithm, GA) 的面积优化方法、基于 AFSA 的面积优化方法、文献 [4] 所提的混合遗传算法 (hybrid genetic algorithm, HGA) 及文献 [17] 中烟花进化人工鱼群算法 (fireworks evolution AFSA, FEAFSA) 的面积优化方法进行了实验对比, 实验结果验证了所提算法的有效性。

## 1 相关知识

### 1.1 MPRM 逻辑表达式

一个输入变量数为  $n$  的 MPRM 逻辑电路中, 对应的逻辑函数表达式具有  $3^n$  个不同的极性。若极性为  $p$ , 则该极性对应的 MPRM 逻辑电路表达式可表示为

$$f^p(x_{n-1}, x_{n-2}, \dots, x_0) = \bigoplus_{j=0}^{2^n-1} b_j \pi_j \quad (1)$$

式中:  $p(0 \leq p < 3^n - 1)$  为 MPRM 逻辑电路的极性, 其三进制形式为  $p_{n-1}p_{n-2} \dots p_1 \dots p_0$ , 其中  $p_i \in \{0, 1, 2\}$ ;  $\bigoplus$  为 XOR 运算,  $\bigoplus$  为异或门;  $\pi_j$  为与项, 可表示为  $\bar{x}_{n-1}\bar{x}_{n-2} \dots \bar{x}_0$ ;  $b_j$  为与项系数, 其中  $b_j \in \{0, 1\}$ , 判断表达式中是否出现该与项; 下标  $j$  为与项序数, 其二进制形式为  $j_{n-1}j_{n-2} \dots j_0$ ; 变量  $\bar{x}_i$  在与项中的表达形式由与项序数和极性确定, 具体如下:

$$\bar{x}_i = \begin{cases} x_i & j_i = 1, p_i = 0 \\ \bar{x}_i & j_i = 1, p_i = 1 \\ x_i, \bar{x}_i & j_i = 1, p_i = 2 \\ 1 & j_i = 0 \end{cases} \quad (2)$$

极性是 MPRM 电路的重要属性, 不同的极性对应着不同的 MPRM 电路表达式。因此, 极性决定着 MPRM 电路表达式的繁简程度, 进而影响 MPRM 电路的性能。

**例 1** 以 3 输入变量的 Boolean 表达式  $f(x_0, x_1, x_2) = x_0x_1\bar{x}_2 + \bar{x}_0x_1\bar{x}_2 + \bar{x}_0\bar{x}_1x_2$  为例, 其对应极性为 6 (三进制表示为 020)、11 (三进制表示为 102) 和 17 (三进制表示为 122) 的 MPRM 电路表达式分别为

$$f_6(x_0, x_1, x_2) = x_1 \oplus x_2 \oplus x_0\bar{x}_1x_2 \oplus x_1x_2 \quad (3)$$

$$f_{11}(x_0, x_1, x_2) = x_1\bar{x}_2 \oplus \bar{x}_0x_2 \oplus \bar{x}_0x_1x_2 \quad (4)$$

$$f_{17}(x_0, x_1, x_2) = x_1\bar{x}_2 \oplus \bar{x}_0\bar{x}_1x_2 \quad (5)$$

可以明显看出, 不同极性对应的 MPRM 表达式的繁简程度差别很大, 其中, 极性为 17 的 MPRM 电路表达式最简洁, 具有潜在的功耗和面积等性能优势。

### 1.2 人工鱼群算法

人工鱼群算法由浙江大学李晓磊和钱积新<sup>[18]</sup>于 2003 年提出, 是一种全新的群智能优化算法。鱼个体数目最多的地方一般就是本水域中富含营养物质最多的地方, 依据这一现象及鱼的特性, 模仿鱼的觅食、聚群、追尾等行为, 从而实现全局寻优。人工鱼的初始种群随机生成, 人工鱼个体  $w = 0, 1, \dots, n-1$ ; 设种群规模为  $\text{popSize}$ ,  $p_{\text{Best}}$  为个体最优值,  $g_{\text{Best}}$  为全局最优值。人工鱼群算法对应的伪代码如下所示, 其中,  $\text{AF\_follow}()$  表示追尾行为,

AF\_swarm() 表示聚群行为, AF\_prey() 表示觅食行为。

输入: 进化参数。

输出: 最优解。

1. Initialize the artificial fish population;
2. FOR  $w \leftarrow 1$  to popSize DO
3. Evaluate each individual and update  $p_{\text{Best}}$ ;
4. Update  $g_{\text{Best}}$ ;
5. END FOR
6. WHILE terminal condition not met
7. FOR each individual  $w(w = 1, 2, \dots, \text{popSize})$
8. Perform AF\_follow();
9. Perform AF\_swarm();
10. Perform AF\_prey();
11. Evaluate individual  $w$ ;
12. Update  $p_{\text{Best}}$  and  $g_{\text{Best}}$ ;
13. END FOR
14. END WHILE
15. RETURN  $g_{\text{Best}}$

### 1.2.1 觅食行为

鱼在水中随机移动时, 会向食物方向快速移动。设  $X_a$  为其中某条鱼  $a$  当前的位置, 若在视野范围内存在另外一条人工鱼  $b$ , 该位置  $X_b$  的食物浓度为  $Y_b$ , 如果  $Y_b > Y_a$ , 则当前人工鱼向  $b$  移动; 否则, 在视野内随机寻找其他人工鱼, 与当前人工鱼的食物浓度进行比较, 确定下一代; 当觅食次数达到最大尝试次数时, 当前人工鱼通过随机行为产生下一代。人工鱼的位置更新为

$$X_a^{t+1} = \begin{cases} X_a + \frac{X_b - X_a}{|X_b - X_a|} \cdot \text{step} \cdot \text{rand}() & Y_b > Y_a \\ X_a + \text{visual} \cdot \text{rand}() & \text{其他} \end{cases} \quad (6)$$

式中: step 为步长; visual 为视野; rand() 为产生区间  $[0, 1]$  之间的随机数;  $t$  为当前迭代次数。

### 1.2.2 追尾行为

人工鱼游动过程中, 朝着食物浓度高的位置移动, 周围的人工鱼追随着向食物方向移动。设  $X_{\text{best}}$  是当前人工鱼  $a$  视野范围内食物浓度最高的位置, 则人工鱼  $a$  向  $X_{\text{best}}$  移动步长位置, 否则进行觅食行为。人工鱼的位置更新为

$$X_a^{t+1} = \begin{cases} X_a + \frac{X_{\text{best}} - X_a}{|X_{\text{best}} - X_a|} \cdot \text{step} \cdot \text{rand}() & Y_{\text{best}} > Y_a \\ \text{prey}(a) & \text{其他} \end{cases} \quad (7)$$

式中: prey( $a$ ) 为人工鱼  $a$  的觅食行为。

### 1.2.3 聚群行为

人工鱼游动过程中, 会自然聚集成群。设人工鱼  $X_c$  所在的位置是人工鱼群的中心位置, 则人工鱼  $X_a$  向  $X_c$  移动步长位置, 否则进行觅食行为。人工鱼的位置更新为

$$X_a^{t+1} = \begin{cases} X_a + \frac{X_c - X_a}{|X_c - X_a|} \cdot \text{step} \cdot \text{rand}() & Y_{\text{best}} > Y_a \\ \text{prey}(a) & \text{其他} \end{cases} \quad (8)$$

## 2 多策略协同进化人工鱼群算法

传统人工鱼群算法的提出是为了解决连续型优化问题, 并且在算法迭代后期, 人工鱼搜索的盲目性较大, 影响算法的收敛性及解的质量。由于 MPRM 电路面积优化属于三值组合优化问题, 所以本节基于人工鱼群算法提出 M-AFSA。所提算法的主要创新点为: 在种群初始化阶段加入反向学习机制, 提出基于反向学习的种群初始化策略, 以提高种群多样性及初始种群解的质量; 为提高算法的收敛速度, 提出觅食与追尾交互性策略, 加强人工鱼个体之间的信息交流, 避免因视野和步长为固定值, 导致算法迭代次数过多; 提出自适应扰动策略, 增大算法迭代后期人工鱼位置变异的随机性, 防止算法过早收敛, 增强算法的全局寻优能力。

### 2.1 距离计算

$X_a$  与  $X_b$  之间的距离是指极性之间的差异程度, 记为  $\text{Dis}(X_a, X_b)$ , 计算式为

$$\text{Dis}(X_a, X_b) = \sum_{k=1}^{k=n} \text{sign}(|X_a^k - X_b^k|) \quad (9)$$

式中:

$$\text{sign}(e) = \begin{cases} 0 & e = 0 \\ 1 & e > 0 \\ -1 & e < 0 \end{cases} \quad (10)$$

**例 2** 以 5 输入变量为例, 2 条人工鱼的位置分别为 00 020、01 210, 则 2 条鱼之间的距离为 3。

### 2.2 基于反向学习的种群初始化策略

初始种群的分布在一定程度上影响算法的收敛速度。为提高初始种群解的质量, 提出基于反向学习的种群初始化策略, 在随机初始化种群时, 加入反向学习机制, 保留最优个体, 产生初始种群, 以扩大种群多样性、提高初始种群解的质量、增强算法收敛性。

反向学习概念由学者 Rahnamayan 等<sup>[19]</sup> 提出, 其主要目的是得到当前可行解的反向解, 并对当前可行解及其反向解进行评估, 选出最优解。该方法可有效提高种群多样性, 增强群智能算法的全局寻优能力。

设  $s$  为定义在区间  $[d, f]$  内的某一实数, 则  $s$  的反向解为

$$\tilde{s} = d + f - s \quad (11)$$

如果  $d = 0$  且  $f = 1$ , 则  $s$  的反向解为

$$\tilde{s} = 1 - s \quad (12)$$

设  $h(s_1, s_2, \dots, s_z, \dots, s_N)$  为  $n$  维搜索空间内的某一位置, 则有

$$\tilde{s}_z = d_z + f_z - s_z \quad z = 1, 2, \dots, n \quad (13)$$

故  $h(s_1, s_2, \dots, s_z, \dots, s_n)$  的反向解为  $\tilde{h}(\tilde{s}_1, \tilde{s}_2, \dots, \tilde{s}_z, \dots, \tilde{s}_n)$ 。

### 2.3 觅食与追尾交互性策略

在传统人工鱼群算法中, 觅食行为是人工鱼不断地在视野范围内随机寻找其他较优位置, 直至找到适应度值优于当前人工鱼的位置。因此, 觅食行为可以使人工鱼更全面地搜索整个解空间。但该行为易使算法复杂度过高, 影响算法的收敛速度。并且在该行为中, 参数的设定对行为的性能有很大影响。参数设置过大, 算法初期收敛速度很快, 但易陷入局部最优; 设置过小, 会使算法的复杂程度过高, 影响算法的运行效率。为解决该问题, 更有效地平衡算法的开发和探索能力, 本文在基于人工鱼群算法的基础上, 对人工鱼群寻优策略进行了优化, 在追尾、觅食阶段引入了觅食与追尾交互性策略。在追尾行为中, 鱼群不断地向局部最优值移动, 直至找到优于当前值的位置。在觅食行为中, 鱼群不断地随机移动, 直至找到优于当前值的位置。本文采用交互性追尾和觅食操作, 让表现较好地个体执行先追尾后觅食的过程, 以加快收敛速度。让表现较差的个体执行先觅食后追尾的过程, 以跳出局部最优解, 增强全局搜索能力。觅食与追尾交互性策略增加了鱼群间信息的交流, 更有效地平衡算法的开发和探索能力。

### 2.4 自适应扰动策略

随着迭代次数的增加, 人工鱼群算法的种群多样性降低, 易使算法陷入局部最优。因此, 在原有行为基础上引入遗传算法中的交叉和变异策略以提高种群多样性。当鱼群完成追尾和觅食行为之后对其实施扰动, 以增加跳出局部最优解的概率, 缩短寻找全局最优解的时间。概率过小, 扰动行为会失去作用, 不利于跳出局部最优解; 概率过大, 位置改变的随机性会偏大, 破坏最优解收敛方向。因此, 本文引入自适应扰动行为, 随着迭代次数的增加, 计算得出的概率值随之变化。迭代前期, 扰动概率较小, 加快收敛速度, 迭代后期, 扰动概率增大, 增强算法跳出局部最优的能力。自适应扰动策略可增加人工鱼位置变异的随机性, 防止人工鱼群算法过早收敛。自适应扰动策略的数学模型为

$$P = \frac{1}{\log_2(\alpha + \alpha \cdot \text{gen})} + \beta \quad (14)$$

式中:  $\text{gen}$  为迭代次数;  $\alpha$  和  $\beta$  为固定常数, 经过大

量实验对比可得,  $\alpha = 2, \beta = 0.3$  时, 得到的实验数据最优。

### 2.5 算法描述

根据 2.1 节~2.4 节所述, M-AFSA 流程如图 1 所示。

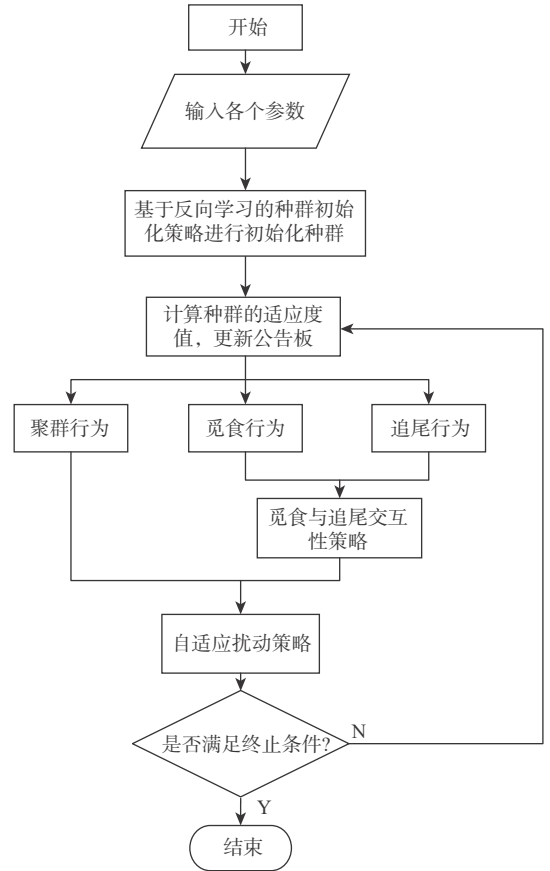


图 1 M-AFSA 流程

Fig. 1 Flow chart of M-AFSA

具体步骤如下:

**步骤 1** 利用基于反向学习的种群初始化策略进行初始化种群。各项参数如下:  $N$  为种群规模,  $\text{visual}$  为人工鱼的视野范围,  $\delta$  为人工鱼群的拥挤度因子,  $\text{Try\_num}$  为人工鱼移动的最大尝试次数,  $\sigma$  为自适应扰动策略中的扰动因子,  $T$  为当代迭代次数,  $T_{\max}$  为最大迭代次数。

**步骤 2** 公告板的初始化。比较初始种群的适应度值, 将最优适应度值对应的人工鱼个体的位置赋值给公告板。

**步骤 3** 进行聚群行为, 寻找当代种群的中心位置。若该位置的适应度值优于当前人工鱼的位置并且不拥挤, 则向该位置方向移动, 产生新的位置。

**步骤 4** 进行觅食与追尾交互性策略, 判断个体适应度值。适应度值高的人工鱼个体执行先追尾后觅食的过程, 适应度值低的人工鱼个体执行先觅食后追尾的过程。

**步骤 5** 计算扰动因子, 若  $\text{rand}() < \sigma$ , 则执行自

适应扰动策略, 否则进行下一步。

**步骤 6** 判断算法是否满足终止条件, 若满足, 结束循环; 否则, 重复步骤 3~步骤 5。

### 2.6 时间复杂度分析

设 M-AFSA 的种群规模为  $N$ , 迭代次数为  $Gen$ 。

1) 初始化种群阶段。由于引入反向学习概念, 需要进行  $2N$  次操作, 因此该阶段的时间复杂度为  $O(2N)$ 。

2) 公告板的初始化阶段。进行  $N$  次比较, 找到初始种群的最优解, 因此, 该阶段的时间复杂度为  $O(N)$ 。

进行聚群行为与进行觅食与追尾交互性策略阶段。聚群行为中每个个体需要进行一次对比, 一次移动,  $N$  次拥挤度计算。在觅食行为和追尾行为阶段, 由于引入了觅食与追尾交互性策略, 每个个体需要进行一次对比, 一次移动,  $N$  次寻找视野范围内最优解,  $Try\_num$  次试探次数, 因此, 该阶段的时间复杂度为  $O(2N + N^2) + O(2N + N^2 + NTry\_num)$ 。

3) 自适应扰动策略阶段。每个个体需要进行交叉操作一次, 变异操作一次, 因此, 该阶段的时间复杂度为  $O(2N)$ 。

因此, M-AFSA 在  $Gen$  次迭代后的时间复杂度为  $O(2N) + O(Gen(2N^2 + 7N + NTry\_num))$ 。

### 2.7 空间复杂度分析

M-AFSA 中人工鱼群的种群规模为  $N$ , 在算法运行过程中, 需要  $N$  个结构体  $A\_fish$  存储人工鱼的各项参数及长度为  $l$  的一维数组, 所需的空间为  $Nl$ ; 需要一个结构体存储种群最优解, 所需空间为  $l$ 。因此, 该算法需要的存储空间为  $O(Nl + l)$ 。

## 3 MPRM 逻辑电路面积优化方法

MPRM 电路极性优化空间随电路输入数的增加呈指数型增长。较小规模的电路面积优化可以通过穷尽搜索方法实现, 但是对于中大规模 MPRM 逻辑电路的面积优化, 穷尽搜索方法无法在有限的时间内得到最优解。为满足中大规模 MPRM 逻辑电路快速、有效面积优化的需要, 本文提出一种 MPRM 逻辑电路面积优化方法, 该方法利用 M-AFSA 搜索电路面积最小的最佳极性, 从而实现电路面积优化。

### 3.1 编码方案

人工鱼的位置表示 MPRM 逻辑电路中的极性, 通过三进制编码, 将极性的三进制形式表示人工鱼个体所在的位置。

**例 3** 以 5 输入变量的电路为例: 极性为 6 的三进制表示为 00020, 则代表该极性的人工鱼的位置为 00020。

### 3.2 面积估算模型

在电路逻辑综合中, 面积模型以二输入逻辑门作为计算单位。因此, 电路中的多输入逻辑门需要

分解为二输入逻辑门。本文将二输入与门和二输入异或门的总数用于衡量 MPRM 电路面积。

其中, 二输入与门的个数  $num\_and$  为

$$num\_and = \sum_{i=0}^{2^n-1} \left[ b_j \left( \sum_{k=0}^{n-1} j_i - 1 \right) \right] \quad (15)$$

二输入异或门的个数  $num\_xor$  为

$$num\_xor = \sum_{i=0}^{2^n-1} b_j - 1 \quad (16)$$

因此, MPRM 逻辑电路面积估算模型为

$$A(p) = \sum_{i=0}^{2^n-1} b_j - 1 + \sum_{i=0}^{2^n-1} \left[ b_j \left( \sum_{k=0}^{n-1} j_i - 1 \right) \right] \quad (17)$$

式中:  $b_j$  为与项系数,  $b_j \in \{0, 1\}$  用于判断表达式中是否出现该与项; 下标  $j$  表示与项序号, 其二进制形式为  $j_{n-1}j_{n-2}\dots j_0$ 。

**例 4** 以 3 输入变量的 MPRM 逻辑电路  $f(x_0, x_1, x_2) = x_1\bar{x}_2 \oplus \bar{x}_0\bar{x}_1x_2$  为例, 其二输入与门的个数为 3, 二输入的异或门个数为 1, 因此, 其电路面积为 4。

### 3.3 适应度函数

适应度值越高, 则人工鱼个体的质量越好。因此, 将面积计算函数的倒数作为适应度函数:

$$fitness(w) = \frac{1}{Area(w)} \quad (18)$$

式中:  $fitness(w)$  为个体  $w$  的适应度值;  $Area(w)$  为个体  $w$  的电路面积。

### 3.4 MPRM 逻辑电路面积优化方法描述

图 2 为 MPRM 电路面积优化方法流程, 首先, 通过列表法将 Boolean 逻辑表达式转化为 MPRM

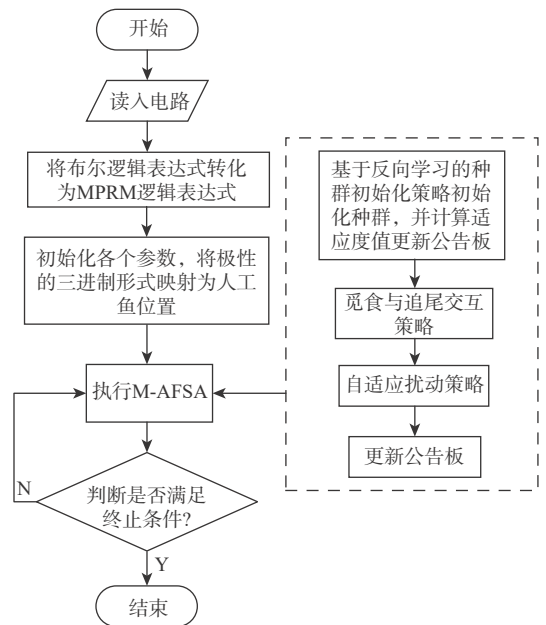


图 2 MPRM 电路面积优化方法流程

Fig. 2 Flow chart of MPRM circuit area optimization method

逻辑表达式;然后,通过三进制编码方式对人工鱼的位置进行编码,并利用汉明距离表示人工鱼之间的距离;最后,通过M-AFSA搜索电路面积最小的最佳极性。

## 4 实验结果与分析

本文所提方法使用C语言实现,软件环境为:Windows 10操作系统、VS 2019编译器,硬件环境为: Intel (R) Core (TM)i5-10210U CPU (2.11GHZ) 8GB RAM。测试电路为MCNC Benchmark基准电路。

为了验证所提方法的有效性,将所提算法与GA、AFSA、文献[4]所提的HGA及文献[17]所提的FEAFSA进行比较。从MCNC Benchmark基准电路测试集中随机选取11个电路作为实验电路,其中,输入变量个数为5~8的为小规模电路,输入变量个数为9~17的为中大规模电路。由于5种算法都具有随机性,本实验将每个算法在每个测试电路上运行20次,并将实验结果的最优值和平均值作为实验数据。

### 4.1 电路面积对比

5种算法的电路面积优化对比数据如表1、表2所示,其中MCNC表示电路名称,为验证所提方法的有效性,从小、中、大规模电路中,各随机选取若干电路,In表示测试电路的输入变量个数,A\_ave表示20次运行中求得的电路面积的平均值,T\_ave表示对应的平均时间。 $S_1$ 、 $S_2$ 、 $S_3$ 、 $S_4$ 分别表示M-AFSA与GA、HGA、AFSA及FEAFSA相比较,所优化的最小面积百分比; $S_5$ 、 $S_6$ 、 $S_7$ 、 $S_8$ 分别表示M-AFSA与GA、HGA、AFSA及FEAFSA比较,所优化的面积平均值百分比。

表1中数据为5种算法求得的电路的最小面积,表2中数据为5种算法求得的电路的平均面积以及对应的平均时间。由表1可知,对于小规模电路来说,5种算法得到的最小面积一致,但是M-AFSA得到的电路面积的平均值要优于其他4种算法。对于中大规模电路来说,M-AFSA得到的最小面积优于其他4种算法,且平均值明显优于其他4种算法。M-AFSA比GA在平均面积上最高节省了57.24%、

表1 最优面积数据

Table 1 Optimal area data

MCNC	In	GA	HGA	AFSA	FEAFSA	M-AFSA	$S_1$ /%	$S_2$ /%	$S_3$ /%	$S_4$ /%
p82	5	2	2	2	2	2	0	0	0	0
pope	6	37	37	37	37	37	0	0	0	0
inc	7	18	18	18	18	18	0	0	0	0
luc	8	6	6	6	6	6	0	0	0	0
exps	8	223	204	204	204	204	8.52	0	0	0
newxcpla1	9	29	28	28	28	28	3.45	0	0	0
prom2	9	84	36	36	36	36	57.14	0	0	0
br1	12	175	156	117	117	117	33.14	25.00	0	0
misex3	14	2 215	2 079	2 415	2 079	2 079	6.14	0	13.91	0
table3	14	3 873	3 522	2 538	2 784	2 538	34.47	27.94	0	8.84
table5	17	201	201	2	158	145	27.86	27.86	0	8.23

表2 平均面积与平均时间数据

Table 2 Average area and average time data

MCNC	In	GA		HGA		AFSA		FEAFSA		M-AFSA		$S_5$ /%	$S_6$ /%	$S_7$ /%	$S_8$ /%
		A_ave	T_ave/s	A_ave	T_ave/s	A_ave	T_ave/s	A_ave	T_ave/s	A_ave	T_ave/s				
p82	5	3.00	0.45	2.00	0.43	2.00	0.19	2.00	0.04	2.00	0.02	33.00	0	0	0
pope	6	42.20	0.51	37.00	0.54	37.20	0.08	37.65	0.21	37.00	0.18	12.32	0	0.54	1.73
inc	7	29.30	0.51	18.33	0.51	23.90	0.04	19.15	0.10	18.35	0.10	37.37	-0.11	23.22	4.18
luc	8	7.80	0.45	6.00	0.47	8.15	0.02	7.60	0.03	6.00	0.02	23.08	0	26.38	21.05
exps	8	275.50	0.91	207.87	0.90	234.75	0.14	241.80	0.17	206.45	0.57	25.06	0.68	12.06	14.62
newxcpla1	9	57.00	1.05	29.87	1.04	33.2	0.52	37.75	1.98	30.05	1.39	47.28	-0.60	9.49	20.40
prom2	9	130.85	0.84	44.00	0.88	68.80	0.23	67.50	0.17	55.95	0.65	57.24	-27.1	18.68	17.11
br1	12	277.40	0.81	176.33	0.81	168.05	0.19	197.75	0.28	142.00	0.38	48.81	19.47	15.50	28.19
misex3	14	3 110.45	13.82	2 079.00	16.20	2 886.55	16.51	2 495.90	27.32	2 145.35	24.51	31.03	-3.19	25.68	14.05
table3	14	5 446.05	675.63	3 522.00	1 057.66	3 092.00	98.47	3 392.00	152.21	3 038.65	889.99	44.20	13.72	1.73	10.42
table5	17	347.85	24.78	228.93	61.92	198.90	1.90	192.15	3.12	183.75	23.54	47.18	19.74	7.62	4.37

平均节省了 39.57%; M-AFSA 比 HGA 在平均面积上最高节省了 19.74%、平均节省了 2.06%; 比 AFSA 在平均面积上最高节省了 33.53%、平均节省了 14.54%; 比 FEAFSA 在平均面积上最高节省了 30.25%、平均节省了 13.86%。

### 4.2 收敛性对比

为验证 M-AFSA 算法的收敛性, 随机选取 8 个 MCNC Benchmark 基准电路。4 种算法的收敛性如

图 3 所示, 其中, 横坐标为迭代次数, 纵坐标为运行 20 次测试电路面积的平均值。可以看出, 与 GA、AFSA 及 FEAFSA 相比, M-AFSA 能够在迭代次数最小的情况下, 找到最小面积。

从图 3 中可以直观地看出, 通过基于反向学习的种群初始化产生的初始种群的平均面积, 均小于其他对比算法, 验证了基于反向学习的种群初始化策略的有效性; 此外, 由 M-AFSA 曲线与 AFSA 曲

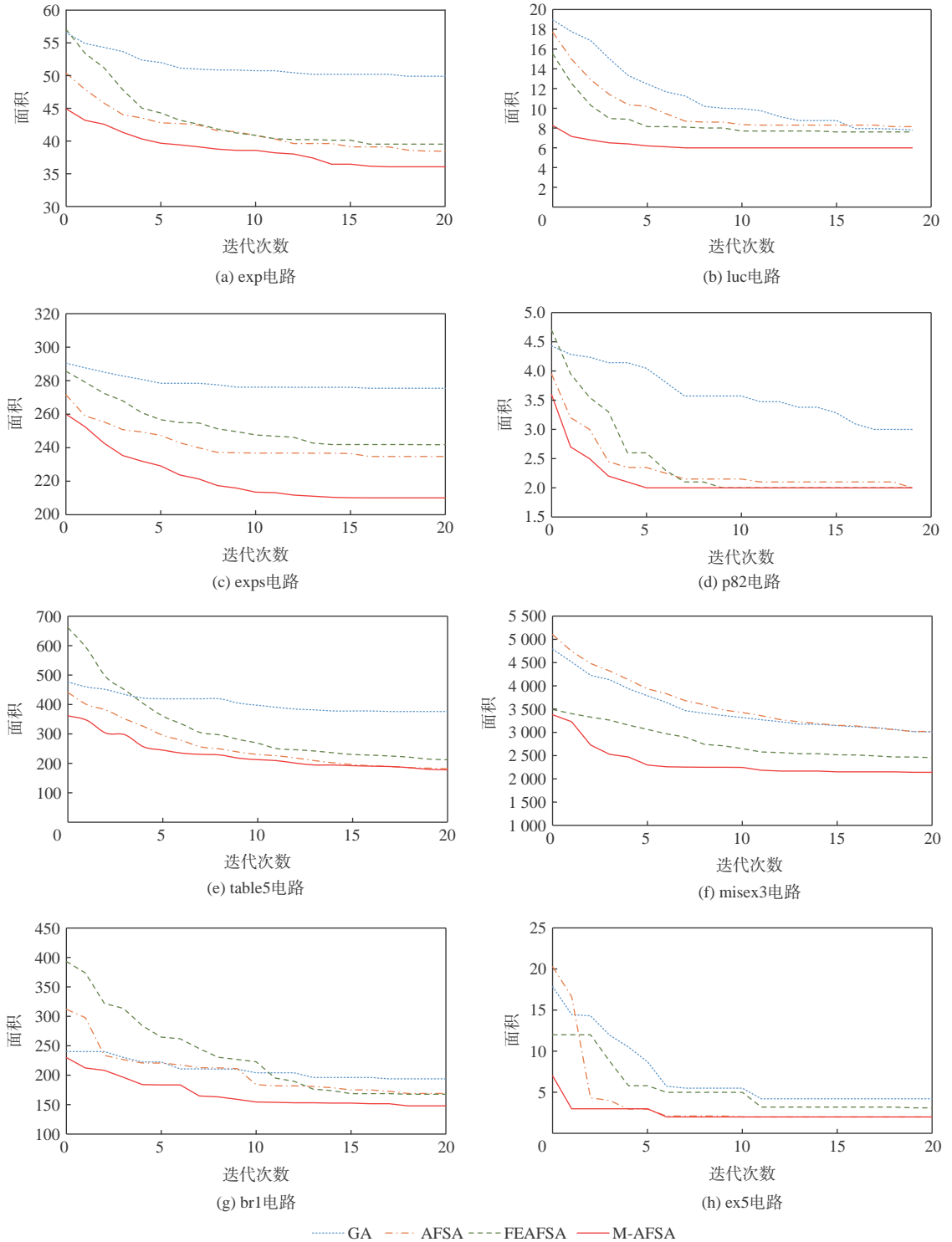


图 3 8 种基准电路最小面积优化曲线

Fig. 3 Optimization curves of minimum area of eight reference circuit

线相比,可看出 M-AFSA 收敛速度快,验证了觅食与追尾交互性策略可有效地提高算法的收敛速度;最后,从图 3 中可看出 M-AFSA 找到的面积绝大多数小于其他 3 种算法,验证了自适应扰动策略可有效地避免 M-AFSA 陷入局部最优。

## 5 结论

本文所提算法用于解决 MPRM 电路面积优化问题。主要结论如下:

1) 提出的 M-AFSA 在初始化种群方面加入反向学习机制,提高了种群的多样性、初始种群解的质量和算法的收敛速度;M-AFSA 在追尾和觅食这 2 种基本行为中加入觅食与追尾交互性策略,增加个体之间的信息交流,提高局部搜索能力,加快算法的收敛速度;传统人工鱼算法随着迭代次数的增加,人工鱼群算法的种群多样性降低,为改善这一缺点,M-AFSA 加入自适应扰动策略,增加人工鱼的位置变异随机性,防止人工鱼群算法过早收敛。

2) 提出一种 MPRM 电路面积优化方法,该方法利用 M-AFSA 来搜索 MPRM 逻辑电路面积的最优解。

3) 通过大量 MCNC 测试电路的实验结果可表明,M-AFSA 在 MPRM 逻辑电路优化方面具有较好的收敛性和寻优性能。基于 MCNC Benchmark 电路的实验结果表明,与 GA 相比,所提算法优化电路平均面积百分比最高为 57.24%;与 HGA 相比,所提算法优化电路平均面积百分比最高为 19.74%;与 AFSA 相比,所提算法优化电路平均面积百分比最高为 33.53%;与 FEAFSA 相比,所提算法优化电路平均面积百分比最高为 30.25%。

## 参考文献 (References)

- [1] HE Z X, XIAO L M, GU F, et al. EDOA: An efficient delay optimization approach for mixed-polarity Reed-Muller logic circuits under the unit delay model[J]. *Frontiers of Computer Science*, 2019, 13(5): 1102-1115.
- [2] HE Z X, PAN Y H, WANG K J, et al. Area optimization for MPRM logic circuits based on improved multiple disturbances fireworks algorithm[J]. *Applied Mathematics and Computation*, 2021, 399: 126008.
- [3] HE Z X, LIU J, XIAO L M, et al. A polarity optimization algorithm taking into account polarity conversion sequence[J]. *IEEE Access*, 2019, 7: 54809-54818.
- [4] 卜登立. 基于混合遗传算法的MPRM最小化[J]. 浙江大学学报(理学版), 2016, 43(2): 184-189.  
BU D L. Hybrid genetic algorithm for MPRM minimization[J]. *Journal of Zhejiang University (Science Edition)*, 2016, 43(2): 184-189(in Chinese).
- [5] CHEN C D, LIN B, ZHU M, et al. Verification method for area optimization of mixed-polarity reed-muller logic circuits[J]. *Journal of Engineering Science and Technology Review*, 2018, 11(1): 28-34.
- [6] 俞海珍, 汪鹏君, 张会红, 等. 基于三值多样性粒子群算法的MPRM电路综合优化[J]. 电子学报, 2017, 45(7): 1601-1607.  
YU H Z, WANG P J, ZHANG H H, et al. Optimization of MPRM circuits based on ternary diversity particle swarm optimization[J]. *Acta Electronica Sinica*, 2017, 45(7): 1601-1607(in Chinese).
- [7] ZONG Y S, HUANG G Y. Application of artificial fish swarm optimization semi-supervised kernel fuzzy clustering algorithm in network intrusion[J]. *Journal of Intelligent & Fuzzy Systems*, 2020, 39(2): 1619-1626.
- [8] ZHOU X B, YU X, ZHANG Y M, et al. Trajectory planning and tracking strategy applied to an unmanned ground vehicle in the presence of obstacles[J]. *IEEE Transactions on Automation Science and Engineering*, 2021, 18(4): 1575-1589.
- [9] LIU Q, ODAKA T, KUROIWA J, et al. An artificial fish swarm algorithm for the multicast routing problem[J]. *IEICE Transactions on Communications*, 2014, 97(5): 996-1011.
- [10] FEI T, ZHANG L Y. Application of BFO-AFSA to location of distribution centre[J]. *Cluster Computing*, 2017, 20(4): 3459-3474.
- [11] LI M, XU J. An AFSA-inspired vector energy routing algorithm based on fluid mechanics[J]. *Tehnicki Vjesnik*, 2020, 27(1): 290-296.
- [12] LIU Y, WANG J, SHAHBAZZADE S. The improved AFSA algorithm for the berth allocation and quay crane assignment problem[J]. *Cluster Computing*, 2019, 22(2): 3665-3672.
- [13] CHEN Y G, ZHU J Y, WAN L, et al. ACOA-AFSA fusion dynamic coded cooperation routing for different scale multi-hop underwater acoustic sensor networks[J]. *IEEE Access*, 2020, 8: 186773-186788.
- [14] WANG S H, ZHANG Z S, REN Y P, et al. UAV photogrammetry and AFSA-Elman neural network in slopes displacement monitoring and forecasting[J]. *KSCE Journal of Civil Engineering*, 2020, 24(1): 19-29.
- [15] LI J P, DONG P W, ZHENG C, et al. PSO-AFSA global maximum power point tracking algorithm with adaptive evolutionary strategy for PV system[C]// *Recent Advances in Intelligent Information Hiding and Multimedia Signal Processing*. Berlin: Springer, 2019: 60-67.
- [16] 李迎, 张璟, 刘庆, 等. 求解大规模多背包问题的高级人工鱼群算法[J]. 系统工程与电子技术, 2018, 40(3): 710-716.  
LI Y, ZHANG J, LIU Q, et al. Advanced artificial fish swarm algorithm for large scale multiple knapsack problem[J]. *Systems Engineering and Electronics*, 2018, 40(3): 710-716(in Chinese).
- [17] 夏平凡, 倪志伟, 朱旭辉. 基于烟花进化人工鱼群算法和多重分形的属性选择方法在空气质量预测中应用[J]. 系统科学与数学, 2020, 40(7): 1157-1177.  
XIA P F, NI Z W, ZHU X H. Attribute selection method based on fireworks evolution artificial fish swarm algorithm and multi-fractal dimension with its application in air quality prediction[J]. *Journal of Systems Science and Mathematical Sciences*, 2020, 40(7): 1157-1177(in Chinese).
- [18] 李晓磊, 钱积新. 基于分解协调的人工鱼群优化算法研究[J]. 电路与系统学报, 2003(1): 1-6.

LI X L, QIAN J X. Studies on artificial fish swarm optimization algorithm based on decomposition and coordination techniques[J]. *Journal of Circuits and Systems*, 2003(1): 1-6(in Chinese).

[19] RAHNAMAYAN S, TIZHOOSH H R, SALAMA M M A. Opposition-based differential evolution[J]. *IEEE Transactions on Evolutionary Computation*, 2008, 12(1): 64-79.

## Area optimization of MPRM circuits based on M-AFSA

SHAO Yixuan<sup>1,2</sup>, HE Zhenxue<sup>1,2,\*</sup>, ZHOU Yuhao<sup>1,2</sup>, HUO Zhisheng<sup>3</sup>, XIAO Limin<sup>3</sup>, WANG Xiang<sup>4</sup>

(1. Key Laboratory of Agricultural Big Data of Hebei Province, Hebei University of Agricultural, Baoding 071001, China;

2. Information Science and Technology, Hebei University of Agricultural, Baoding 071001, China;

3. School of Computer, Beihang University, Beijing 100191, China;

4. School of Electronic and Information Engineering, Beihang University, Beijing 100191, China)

**Abstract:** The existing mixed polarity Reed-Muller (MPRM) circuit area optimization algorithms based on the traditional intelligent optimization algorithms have the problem of poor performance. The MPRM circuit's area optimization is a combinatorial optimization issue, hence an artificial fish swarm algorithm with many strategies (M-AFSA) is initially suggested. In this algorithm, a population initialization strategy based on reverse learning is introduced to improve the population diversity and the quality of the initial population solution; the interactive strategies of foraging and rearing were introduced to enhance the information exchange between the artificial fish individuals and improve the convergence speed of the algorithm; Adaptive perturbation strategy is introduced to increase the randomness of location variation of artificial fish and avoid the algorithm falling into local optimum. Moreover, we present an area optimization method for MPRM logic circuits, which uses the proposed multi-strategy coevolutionary artificial fish swarm algorithm to search for the optimal polarity with the minimum circuit area. The experimental results based on the MCNC Benchmark circuit show that compared with the genetic algorithm, the maximum area saving percentage obtained by this algorithm is 57.24%, and the average area save percentage obtained by this algorithm is 39.57%. Compared with the artificial fish swarm algorithm, the maximum and average area saving percentages obtained by this algorithm are 33.53% and 14.54%, respectively. Compared with the improved artificial fish swarm algorithm, the maximum and average area saving percentages obtained by this algorithm are 30.25% and 13.86%, respectively.

**Keywords:** mixed polarity Reed-Muller circuit; area optimization; combinatorial optimization; artificial fish swarms algorithm; opposition-based learn

**Received:** 2021-06-03; **Accepted:** 2021-09-23; **Published Online:** 2021-10-29 17:05

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20211029.1450.002.html](https://kns.cnki.net/kcms/detail/11.2625.V.20211029.1450.002.html)

**Foundation items:** National Natural Science Foundation for Youth (62102130); Natural Science Foundation of Hebei Province (F2020204003); Hebei Youth Talents Support Project (BJ2019008); Introducing Talent Research Project of Hebei Agricultural University (YJ201829); Central Government Guides Local Science and Technology Development Fund Project (226Z0201G)

\* **Corresponding author.** E-mail: [hezhenxue@buaa.edu.cn](mailto:hezhenxue@buaa.edu.cn)

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0300

# 基于动态特征选择的遥感图像目标检测算法

陈超, 赵巍\*

(北京航空航天大学 电子信息工程学院, 北京 100191)

**摘要:** 在遥感图像目标检测领域内, 旋转物体的检测存在挑战, 卷积神经网络在提取信息时会受制于固定的空间结构, 采样点无法聚焦于目标; 遥感图像尺度变化大, 不同物体需要具有不同尺度感受野的特征映射, 具有单一尺度感受野的特征映射无法包含所有有效信息。基于此, 提出了可变形对齐卷积, 根据候选边框调节采样点, 并根据特征映射学习采样点的细微偏移, 使采样点聚焦于目标, 从而实现动态特征选择; 同时提出了基于可变形对齐卷积的感受野自适应模块, 对具有不同尺度感受野的特征映射进行融合, 自适应地调整神经元的感受野。在公开数据集上的大量实验验证了所提算法可以提高遥感图像目标检测的精度。

**关键词:** 深度学习; 卷积神经网络; 遥感图像目标检测; 特征选择; 感受野

**中图分类号:** TP391.4

**文献标志码:** A

**文章编号:** 1001-5965(2023)03-0702-08

在遥感图像研究中, 目标检测是理解地面信息的重要手段, 目的是在遥感图像中自动搜索出感兴趣目标, 并标定其位置、判别其类别。近年来, 随着深度学习的发展<sup>[1]</sup>, 通用图像的目标检测取得了巨大的发展和进步, 同时随着高质量的遥感数据集(如 DOTA 数据集<sup>[2]</sup>)的建立, 遥感图像目标检测也逐渐成为研究热点。与一般图像不同, 地理空间图像通常来自鸟瞰视角, 这就造成遥感图像中目标尺寸变化巨大, 目标姿态不一, 目标长宽比变化极端, 且遥感成像还受到光照、环境、季节等多种因素影响。上述变化均为遥感图像目标检测带来了额外的挑战。

为实现更好的检测性能, 目前大部分先进的遥感图像目标检测算法都基于域卷积神经网络(region-convolutional neural networks, R-CNN)的系列网络<sup>[3-6]</sup>。R-CNN 系列网络生成大量的水平检测框(horizontal bounding box, HBB)作为感兴趣区域, 通过复杂的感兴趣区域池化操作(如感兴趣区域池化<sup>[4]</sup>、感兴趣区域对齐<sup>[6]</sup>、感兴趣区域可变形池化<sup>[7]</sup>)

提取感兴趣区域的特征, 基于这些区域的特征进行分类与定位。在遥感图像中, 水平的感兴趣区域会造成预测边框与旋转物体之间的不匹配问题<sup>[8]</sup>。例如, 遥感图像中的物体通常具有一定的朝向且可能呈密集排列(如在港口停靠的船只与在停车场停放的车辆), 导致水平感兴趣区域包含若干个物体。自然的解决方法<sup>[9-10]</sup>是利用定向检测框(oriented bounding box, OBB)作为锚(anchor)来处理旋转的物体, 这些方法需要使用大量不同角度、比例、尺度的锚, 导致网络的效率较低。近期, RoI Transformer<sup>[8]</sup>算法使用将水平感兴趣区域转换为定向感兴趣区域的方式来解决效率低的问题, 但仍需要进行复杂的感兴趣区域池化操作。

与基于 R-CNN 的网络相比, 单阶段网络不进行感兴趣区域池化操作, 直接对感兴趣区域进行分类与回归, 通常效率较高但精度较差<sup>[2]</sup>。单阶段网络在遥感图像目标检测上性能较差的原因主要是由于单阶段网络没有感兴趣区域池化操作, 单纯依赖卷积提取特征。而卷积神经网络在提取信息时

收稿日期: 2021-06-04; 录用日期: 2021-08-13; 网络出版时间: 2021-09-09 09:18

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210908.1932.002.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210908.1932.002.html)

\*通信作者. E-mail: [zhao29wei@263.net](mailto:zhao29wei@263.net)

**引用格式:** 陈超, 赵巍. 基于动态特征选择的遥感图像目标检测算法[J]. 北京航空航天大学学报, 2023, 49(3): 702-709.

CHEN C, ZHAO W. Remote sensing target detection based on dynamic feature selection[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(3): 702-709 (in Chinese).

受制于固定的空间结构, 采样的位置始终是固定的, 无法实现灵活的特征选择, 而遥感目标朝向、尺度、形状不一, 采样点无法聚焦于目标。针对卷积神经网络受制于固定空间结构的问题, 可变形卷积 (deformable convolution)<sup>[7]</sup> 根据特征映射在原本的采样点处学习一个偏移, 使采样点聚焦于目标, 但其要求堆叠大量的可变形卷积, 造成速度的瓶颈; 对齐卷积 (alignment convolution)<sup>[11]</sup> 根据物体的方向与尺度使采样点跟随旋转目标动态调整, 但其采样点仍然维持矩形的形状, 无法根据物体的具体形状实现聚焦。

除了考虑局部的特征聚焦外, 遥感图像目标尺度变化巨大, 还需对网络的多尺度进行更深入的研究。而目前多尺度的研究主要基于特征金字塔结构<sup>[12]</sup>, 其不同层的特征映射检测不同尺度的物体。特征金字塔的出发点是模型对于不同尺度的物体应具有不同的感受野, 具有不同尺度感受野的特征映射包含了不同的有效信息, 对于同一尺度的不同物体, 感受野也需要有细微的差别。

针对上述卷积神经网络在提取目标信息时受

受制于固定的空间结构, 单阶段遥感图像目标检测网络无法实现特征聚焦的问题, 本文提出了可变形对齐卷积, 根据定向边框的长、宽和角度改变卷积的采样点, 根据特征动态给采样点细微的偏移, 实现了动态的特征选择, 以适应不同尺度、角度、形状的物体; 受启发于同一尺度的不同类别物体需要的感受野的细微差异, 提出了基于可变形对齐卷积的感受野自适应模块, 融合具有不同尺度感受野的特征映射, 自适应地改变神经元的感受野。以上述 2 个创新点为基础, 提出了基于动态特征选择的单阶段遥感图像精确目标检测算法 DFSNet。

### 1 DFSNet

DFSNet 使用 RetinaNet<sup>[13]</sup> 作为基础模型, 首先使用骨干网络提取多尺度的特征, 再使用感兴趣区域转换模块将水平检测框转换为定向检测框, 然后使用感受野自适应模块得到包含不同感受野信息的特征映射, 感受野自适应模块中使用动态特征选择层实现特征的聚焦, 最后经分类与回归子网络得到检测的结果。网络的总体结构如图 1 所示。

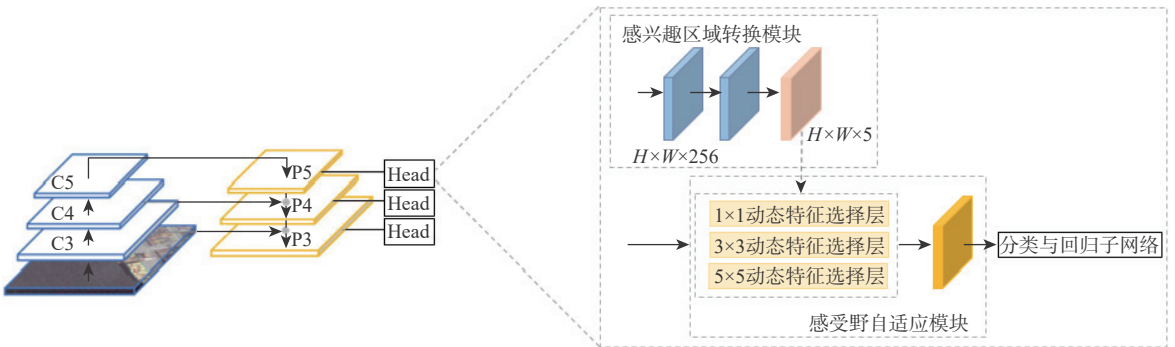


图 1 DFSNet 总体结构

Fig. 1 Overall structure of DFSNet

#### 1.1 基础网络 RetinaNet

RetinaNet 的具体结构如图 2 所示, 主要包括骨干网络和预测分支。骨干网络包括特征提取网络和特征金字塔模块。其中, 特征提取网络的作用是提取图像特征, RetinaNet 主要使用 ResNet<sup>[14]</sup> 作为特征提取网络; 特征金字塔模块<sup>[12]</sup> 的作用是融合多尺度的特征映射。预测分支包括分类子网络与回归子网络, 均由 1 层卷积构成, 分别实现对包围目标的类别识别与边框回归。

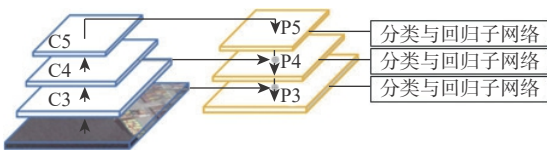


图 2 RetinaNet 总体结构

Fig. 2 Overall structure of RetinaNet

#### 1.2 感兴趣区域转换模块

如图 3 所示, 图中的实线边框表示一个定向检测框  $(x, y, w, h, \theta)$ , 其中,  $(x, y)$  为边框的中心点坐标,  $w$  和  $h$  分别为边框的宽和长,  $\theta \in [-\pi/4, 3\pi/4]$  为边框与  $x$  轴的夹角。由于设置多个不同角度、比例、尺度的锚将给网络造成极大的计算负担, 而不在锚中加入角度参数直接预测角度精度较低, 因此 DFSNet 中的 RetinaNet 不直接对定向旋转框进行预测, 而是先预测定向旋转框的最小外接水平矩形 (见图 3 虚线矩形), 再由感兴趣区域转换模块回归出定向旋转框 (见图 3 实线矩形), 降低了网络的学习难度。

感兴趣区域转换模块由 RoI Transformer<sup>[8]</sup> 提出, 是一个有物体分类分支和边框回归分支的轻量级网络, 分类分支与回归分支均由 2 层卷积构成, 完成了从水平检测框到定向检测框的转换。

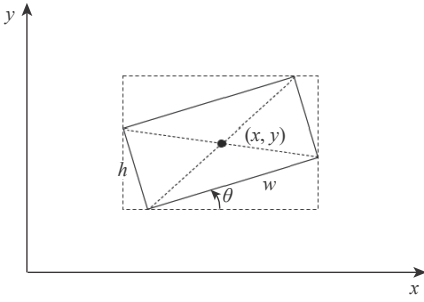


图3 两种输出边框

Fig. 3 Two output bounding boxes

### 1.3 可变形对齐卷积

标准的二维卷积(见图4(a))利用一个网格  $\mathbf{R}=\{(r_x, r_y)\}$  (如  $\mathbf{R}=\{(-1, -1), (-1, 0), \dots, (0, 1), (1, 1)\}$ ) 在特征映射  $\mathbf{X}$  上进行采样, 然后将采样的值乘以权重  $\mathbf{W}_{\text{weight}}$  再求和, 得到特征映射  $\mathbf{Y}$ 。  $\mathbf{Y}$  在每一个位置  $p$  的特征计算公式为

$$\mathbf{Y}(p) = \sum_{r_n \in \mathbf{R}} \mathbf{W}_{\text{weight}}(r_n) \cdot \mathbf{X}(p + r_n) \quad (1)$$

式中:  $r_n$  为网格  $\mathbf{R}$  中的第  $n$  个采样点。

如图4(b)所示, 与标准卷积相比, 可变形卷积<sup>[7]</sup> 根据特征在采样点上学习一系列偏移量  $\mathbf{O}_{\text{DC}}$ 。  $\mathbf{Y}$  在每一个位置  $p$  的特征计算公式变为

$$\mathbf{Y}(p) = \sum_{r_n \in \mathbf{R}, o_{r_n} \in \mathbf{O}_{\text{DC}}} \mathbf{W}_{\text{weight}}(r_n) \cdot \mathbf{X}(p + r_n + o_{r_n}) \quad (2)$$

式中:  $o_{r_n}$  为偏移量  $\mathbf{O}_{\text{DC}}$  中在采样点  $r_n$  的偏移量。

如图4(c)所示, 与标准卷积相比, 对齐卷积<sup>[11]</sup> 根据感兴趣区域给采样点增加偏移量  $\mathbf{O}_{\text{AC}}$ ,  $\mathbf{O}_{\text{AC}}$  的计算方式如式(3)所示。  $\mathbf{Y}$  在每一个位置  $p$  的特征计算公式变为

$$\mathbf{O}_{\text{AC}} = \frac{1}{S} \left( \frac{1}{k} (w, h) \cdot \mathbf{R}(\theta) \right) \cdot \mathbf{R}^T(\theta) \quad (3)$$

式中:  $w$  和  $h$  分别为在  $p$  的感兴趣区域  $(x, y, w, h, \theta)$  的宽和高;  $k$  为卷积核的尺寸;  $S$  为特征映射的步长;  $\mathbf{R}(\theta) = (\cos \theta, -\sin \theta; \sin \theta, \cos \theta)^T$  为旋转矩阵。

$$\mathbf{Y}(p) = \sum_{r_n \in \mathbf{R}, o_{r_n} \in \mathbf{O}_{\text{AC}}} \mathbf{W}_{\text{weight}}(r_n) \cdot \mathbf{X}(p + r_n + o_{r_n}) \quad (4)$$

式中:  $o_{r_n}$  为偏移量  $\mathbf{O}_{\text{AC}}$  中在采样点  $r_n$  的偏移量。

可变形卷积在提取特征时, 可以更多地把注意力放在与目标有关的位置上, 更好地处理不同形状的目标; 对齐卷积在提取旋转物体的特征时, 可以使采样点的分布贴近定向边框的形状, 更好地提取旋转物体的特征。因此, 将可变形卷积与对齐卷积结合起来, 设计了可变形对齐卷积, 如图4(d)所示。可变形对齐卷积能够先根据候选边框调整采样点, 使采样点的分布(图4(d)中矩形点)贴近定向边框的形状, 再通过特征学习偏移, 聚集于与目标有关的位置上(图4(d)中三角形点), 从而实现更好的特征提取。可变形对齐卷积的采样点偏移量  $\mathbf{O}_{\text{DAC}}$  为

$$\mathbf{O}_{\text{DAC}} = \mathbf{O}_{\text{AC}} + \mathbf{O}_{\text{DC}} \quad (5)$$

图5为核尺寸为  $3 \times 3$  的不同卷积的采样点对比, 可以更形象地说明不同卷积对于采样点的影响。如图5(a)所示, 标准二维卷积受制于固定的空间结构, 采样点为图中圆点, 即通常的采样点, 无法将采样点集中在舰船上; 如图5(b)所示, 经过可变形卷积, 采样点为图中的三角形点, 可以集中于舰船之上, 但这需要大量的可变形卷积的堆叠, 严重影响模型的速度; 如图5(c)所示, 经过对齐卷积后, 采样点为图中的矩形点, 呈定向框分布, 初步实现对舰船的聚焦; 如图5(d)所示, 经过可变形对齐卷积后, 网络可以先根据舰船的方向、尺度、长宽比大致实现对舰船的聚焦, 再根据特征图对采样点进

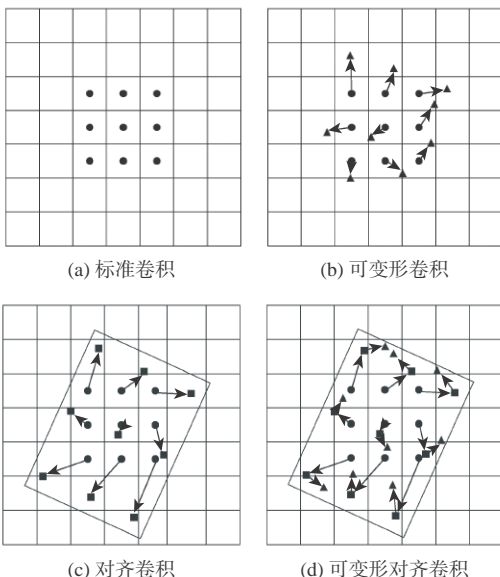
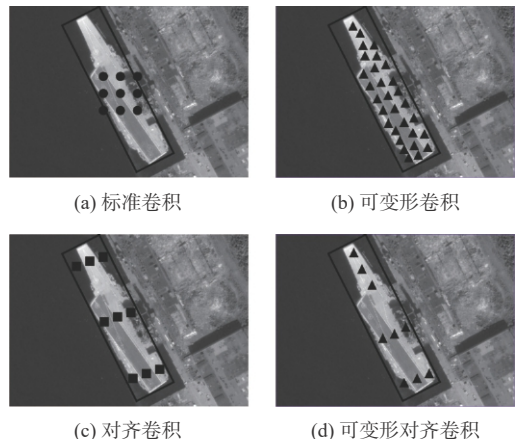
图4 核尺寸为  $3 \times 3$  的不同卷积的采样点对比Fig. 4 Comparison of sampling locations of different convolutions with kernel size of  $3 \times 3$ 

图5 经过不同卷积的舰船采样点

Fig. 5 Ship's sampling locations after different convolutions

行细微的调整, 得到与物体形状更为匹配的三角形采样点。可变形对齐卷积能够改变采样点动态地选择特征, 实现对舰船的聚焦。

### 1.4 动态特征选择层

基于可变形对齐卷积, 构建了动态特征选择层, 如图 6 所示。对于感兴趣区域转换模块输出的  $H \times W \times 5$  的特征映射, 首先将其解码为定向边框  $(x, y, w, h, \theta)$  的形式, 然后利用式 (3) 得到对齐的偏移量; 同时特征映射经过卷积得到进一步的偏移量。2 种偏移量输入到可变形对齐卷积, 特征映射经可变形对齐卷积后得到采样点聚焦于目标的特征映射, 实现了特征的动态选择。

### 1.5 感受野自适应模块

具有同一尺度的不同类别的物体, 所需要的感受野也应有细微的差别。例如, 同一个尺度的车辆与足球场, 足球场的边缘、纹理特征不明显, 需要的感受野更大以便提取更多背景信息, 而车辆的边缘、纹理特征明显, 感受野可以相对稍小。可变形卷积层输出的特征映射的感受野受到感兴趣区域边框参数与核大小控制, 为了融合不同尺度感受野的信息, 受 TridentNet<sup>[15]</sup> 与 DRN<sup>[16]</sup> 启发, 提出了基于可变形对齐卷积的感受野自适应模块, 融合不同核尺寸的可变形卷积层输出的特征映射。如图 7 所示, 输入的特征映射  $X$  经过具有不同核尺寸 ( $1 \times 1$ 、

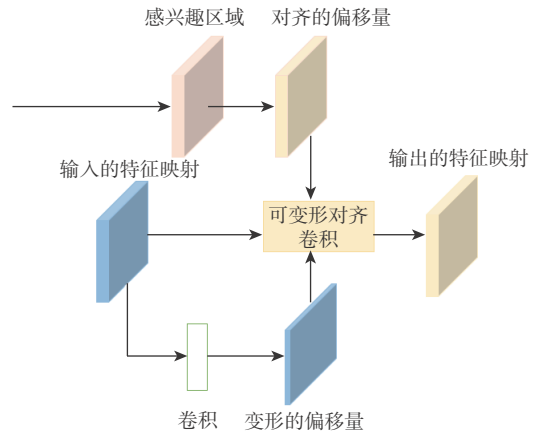


图 6 动态特征选择层

Fig. 6 Dynamic feature selection layer

$3 \times 3$ 、 $5 \times 5$ ) 的动态特征选择层, 分别得到特征映射  $X_i \in \mathbf{R}^{H \times W \times C}$  ( $i \in 1, 2, 3$ ), 经过  $1 \times 1$  卷积、Batch Normalization<sup>[17]</sup> 和 ReLU 激活函数得到特征映射  $A_i \in \mathbf{R}^{H \times W \times 1}$  ( $i \in 1, 2, 3$ ), 如式 (6) 所示, 通过 fast normalization<sup>[18]</sup> 得到不同特征映射的权重  $W_i \in \mathbf{R}^{H \times W \times 1}$  ( $i \in 1, 2, 3$ ), 最终如式 (7) 所示得到融合的特征  $Y$ 。

$$W_i = \frac{A_i}{\sum_i A_i} \quad (6)$$

$$Y = \sum_i W_i \cdot X_i \quad (7)$$

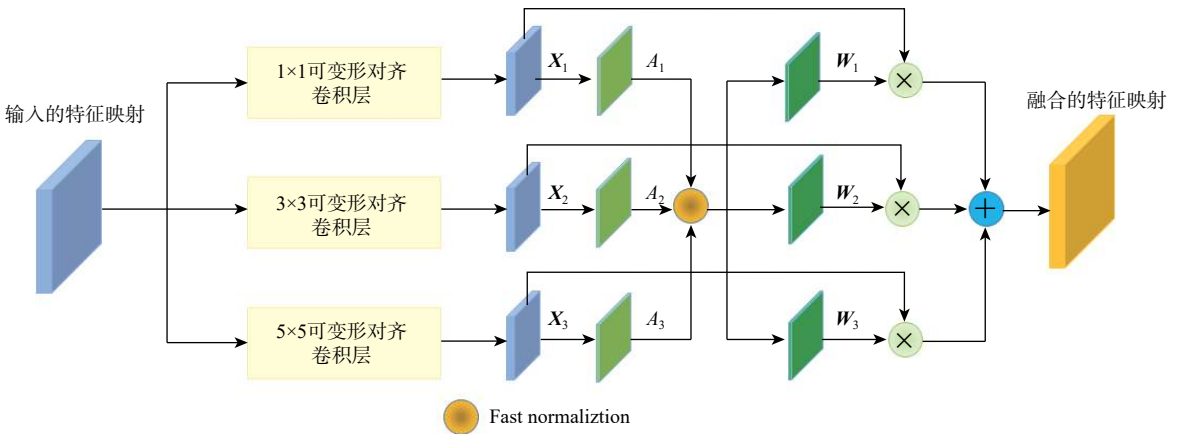


图 7 感受野自适应模块

Fig. 7 Receptive field adaptive module

## 1.6 网络损失函数

### 1.6.1 基本的回归目标

针对旋转检测框, 回归的目标主要有

$$\begin{cases} \Delta x_g = ((x_g - x) \cos \theta, -(y_g - y) \sin \theta) \cdot \left(\frac{1}{w}, \frac{1}{h}\right)^T \\ \Delta y_g = ((x_g - x) \sin \theta, (y_g - y) \cos \theta) \cdot \left(\frac{1}{w}, \frac{1}{h}\right)^T \\ (\Delta w_g, \Delta h_g) = \lg(w_g, h_g) - \lg(w, h) \\ \Delta \theta_g = \frac{1}{\pi}(\theta_g - \theta + k\pi) \end{cases} \quad (8)$$

式中:  $(x_g, y_g, w_g, h_g, \theta_g)$  和  $(x, y, w, h, \theta)$  分别代表真实边框和锚;  $k$  为一个确保  $(\theta_g - \theta + k\pi) \in [-\pi/4, 3\pi/4]$  的整数。在 DFSNet 的动态特征选择层中, 锚是提前设置的水平边框,  $\theta$  设为 0; 在感受野自适应模块中, 锚是动态特征选择层输出的旋转边框结果。

### 1.6.2 匹配策略

使用交并比 (intersection over union, IoU) 作为边框匹配的指标, 如图 8 所示, IoU 代表 2 个边框 (在训练阶段匹配正负样本时, 使用锚与真实边框

计算 IoU, 在推理阶段使用不同的预测边框计算 IoU) 的重叠程度, 计算公式为

$$\text{IoU} = \frac{S_{A \cap B}}{S_{A \cup B}} \quad (9)$$

式中:  $A$  和  $B$  代表 2 个边框。

将与真实框的 IoU 大于 0.5 的锚设为正样本, 小于 0.4 的设为负样本。动态特征选择层中, 锚是水平框, 真实边框是旋转边框, 此时根据锚与旋转边框的最小外接矩形计算 IoU。感受野自适应模块中, 直接计算动态特征选择层输出的旋转边框与真实的旋转边框的 IoU。

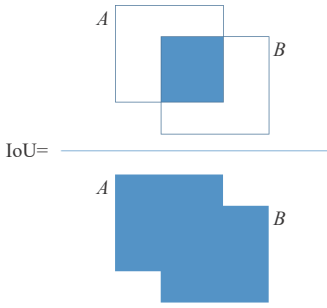


图 8 IoU 的解释

Fig. 8 Explanation of IoU

### 1.6.3 损失函数

DFSNet 的损失函数由动态特征选择层的损失与感受野自适应模块的损失构成, 每一部分损失由分类的损失与回归的损失构成, 总的损失函数如下:

$$L = L_c(c^T, l^*) + L_r(x^T, g^*) + L_c(c^R, l^*) + L_r(x^R, g^*) \quad (10)$$

式中:  $L_c$  代表类别的损失函数, 采用聚焦损失函数<sup>[1,13]</sup>;  $L_r$  代表边框回归的损失函数, 采用 smooth L1 损失函数;  $l^*$  和  $g^*$  分别代表真实的类别和回归的目标;  $c^T$  和  $x^T$  代表感受野转换模块输出的类别和边框;  $c^R$  和  $x^R$  代表感受野自适应模块输出的类别和边框。

## 2 实验结果

实验采用的硬件环境为: CPU 配置为 Intel(R) Xeon(R) E5-2680 V3@2.50 GHz, GPU 配置为 NVIDIA 1080 Ti。实验的软件环境为: Ubuntu16.04.6 LTS、CUDA10.0、PyTorch<sup>[19]</sup> 1.3.1 和 Python3.6.9。

### 2.1 数据集

DOTA 数据集<sup>[2]</sup>是目前最通用的遥感图像目标检测数据集, 包含 2 806 张图片, 图片的大小在 800×800 到 4 000×4 000 像素范围内, 包含 188 282 个不同尺度、不同方向、不同形状的实例。DOTA 数据集包含 15 个类别的物体, 分别为飞机 (PL)、棒球场 (BD)、桥梁 (BG)、田径场 (GTF)、小型车辆 (SV)、大型车辆 (LV)、船只 (SH)、网球场 (TC)、篮球场

(BC)、油罐 (ST)、足球场 (SBF)、环岛 (RA)、港口 (HA)、游泳池 (SP) 及直升机 (HC)。

### 2.2 实验细节

采用 ResNet50 FPN 作为骨干网络。对于特征金字塔 (如 P3~P7) 的每一层特征映射, 只在一个点上生成一个正方形的锚, 锚的尺度是特征映射步长的 4 倍 (如 32、64、128、256、512)。实验均在 DOTA 数据集上训练 24 个轮次, 使用 SGD 作为优化器, 初始学习率设置为 0.01, 在第 18 和第 22 个轮次学习率下降 10 倍。实验输入图片尺寸为 1 024×1 024, 使用 4 张 1080 Ti 显卡进行单尺度的训练, 每次迭代处理 16 张图片, 数据增强仅使用随机的镜像翻转, 测试时使用一张显卡, 非极大值抑制 (non-maximum suppression, NMS) 的阈值为 0.1。

### 2.3 评价指标

平均精度 (average precision, AP) 是精度-召回曲线下的面积, DOTA 数据集使用 IoU=0.5 时的 AP 作为评估指标。mAP 代表对不同类别 AP 求平均的结果。

### 2.4 实验结果及分析

#### 2.4.1 基础模型 RetinaNet 结果

在基础模型 RetinaNet 中, 将锚的长宽比分别设置为 0.5、1、1.5, 尺度设置为特征映射步长的 2、4、8 倍, 在特征映射的一个点设置 9 个锚。最终可以在 DOTA 数据集上达到 68.05% 的 mAP, 验证了基础模型的可靠性。

#### 2.4.2 可变形对齐卷积的作用

为验证可变形对齐卷积的有效性, 将可变形对齐卷积与其他卷积进行对比。实验中, 将 DFSNet 的可变形对齐卷积替换为其他卷积, 并且不使用感受野自适应模块, 而是在感受野转换模块后经过 3×3 的不同卷积得到修正的特征映射, 再进行类别识别与边框回归, 其他网络结构与设置维持不变。特别的, 可变形卷积需要一定的堆叠才会起作用, 因此实验中使用了 2 层可变形卷积。

如表 1 所示, 可变形卷积与对齐卷积都能使网络的识别精度得以提高, 可变形对齐卷积能够结合两者的优点, 实现对目标的聚焦, 在标准卷积的基础上 mAP 提升了 2.01%。

#### 2.4.3 感兴趣区域转换模块、动态特征选择层、感受野自适应模块的作用

为了验证感兴趣区域转换模块、动态特征选择层、感受野自适应模块的作用, 对其进行了消融实验, 逐步添加感兴趣区域转换模块、动态特征选择层与感受野自适应模块。如表 2 所示, 通过添加感

表 1 可变形对齐卷积与其他卷积对比

Table 1 Comparison between deformable alignment convolution and other convolutions

方法	mAP/%
标准卷积	71.17
可变形卷积	71.68
对齐卷积	72.45
可变形对齐卷积	73.18

表 2 DFSNet 的消融实验对比

Table 2 Ablation studies of DFSNet

方法	感兴趣区域 转换模块	动态特征 选择层	感受野自 适应模块	mAP
基线				68.05
DFSNet的 不同设置	√			71.17
	√	√		73.18
	√	√	√	74.04

感兴趣区域转换模块、动态特征选择层、感受野自适应模块, 模型的性能持续提高, 表明了三者可以相互兼容, 并且可以联合使用来获得最佳性能。

2.4.4 与其他模型对比及可视化

如表 3 所示, DFSNet 在使用 ResNet50 FPN 作骨干网络时, 在 DOTA 数据集上可以达到 74.04%

的 mAP, 远远优于其他单阶段或双阶段模型, 有着优异的性能, 说明 DFSNet 能够有效提高不同尺度、不同类别物体的检测精度。与基础模型 RetinaNet 相比, DFSNet 在小型车辆 (SV)、大型车辆 (LV)、船只 (SH)、油罐 (ST) 提升很大, 证明 DFSNet 能够对密集排列的物体实现聚焦。DFSNet 在 DOTA 数据集的可视化结果如图 9 所示。可以看到, DFSNet 相比基础模型有着优异的检测效果。

表 3 DFSNet 与其他模型在 DOTA 数据集上的对比结果

Table 3 Comparison of DFSNet and other methods on DOTA

双/单阶段	模型	AP/%					mAP/%
		PL	BD	BR	GTF	SV	
双阶段	FR-O <sup>[2]</sup>	79.42	77.13	17.70	64.05	35.30	54.72
	RoI Trans former <sup>[8]</sup>	88.64	78.52	43.44	75.92	68.81	71.10
	CAD-Net <sup>[20]</sup>	87.80	82.40	49.40	73.50	71.10	72.84
	SCRDet <sup>[21]</sup>	89.98	80.65	52.09	68.36	68.36	71.89
单阶段	RetinaNet	88.82	81.74	44.44	65.72	67.11	69.57
	DRN <sup>[16]</sup>	88.91	80.22	43.52	63.35	73.48	69.90
	R3Det <sup>[22]</sup>	89.54	81.99	48.46	62.52	70.48	70.60
	DFSNet	89.12	77.40	52.05	73.47	78.02	74.01
双/单阶段	模型	AP/%					mAP/%
		LV	SH	TC	BC	ST	
双阶段	FR-O <sup>[2]</sup>	38.02	37.16	89.41	69.64	59.28	58.70
	RoI Trans former <sup>[8]</sup>	73.68	83.59	90.74	77.27	81.46	81.35
	CAD-Net <sup>[20]</sup>	63.50	76.60	90.90	79.20	73.30	76.70
	SCRDet <sup>[21]</sup>	60.32	72.41	90.85	87.94	86.86	79.68
单阶段	RetinaNet	55.82	72.77	90.55	82.83	76.30	75.65
	DRN <sup>[16]</sup>	70.69	89.94	90.14	83.85	84.11	83.75
	R3Det <sup>[22]</sup>	74.29	77.54	90.80	81.39	83.54	81.51
	DFSNet	79.31	87.35	90.90	85.13	84.90	85.52
双/单阶段	模型	AP/%					mAP/%
		SBF	RA	HA	SP	HC	
双阶段	FR-O <sup>[2]</sup>	50.30	52.91	47.89	47.40	46.30	48.96
	RoI Trans former <sup>[8]</sup>	58.39	53.54	62.83	58.93	47.67	56.27
	CAD-Net <sup>[20]</sup>	48.40	60.90	62.00	67.00	62.20	60.10
	SCRDet <sup>[21]</sup>	65.02	66.68	66.25	68.24	65.21	66.28
单阶段	RetinaNet	54.19	63.64	63.71	69.73	53.37	60.93
	DRN <sup>[16]</sup>	50.12	58.41	67.62	68.60	52.50	59.45
	R3Det <sup>[22]</sup>	61.97	59.82	65.44	67.46	60.05	62.95
	DFSNet	60.90	63.83	67.31	67.56	53.35	62.59

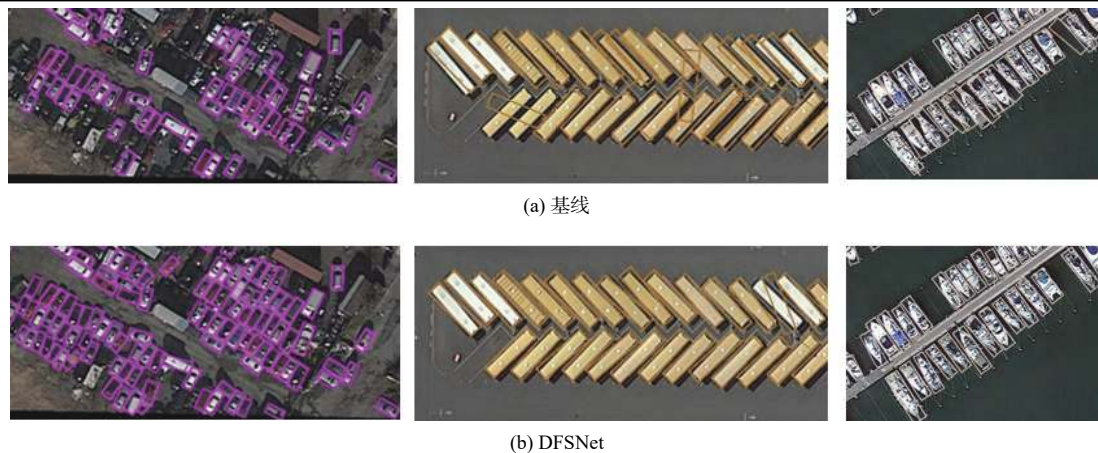


图9 DOTA数据集可视化结果

Fig. 9 Visual results on DOTA

### 3 结论

在RetinaNet的基础上,增加了感兴趣区域转换模块、动态特征选择层、感受野自适应模块,提出了基于动态特征选择的遥感图像目标检测算法DFSNet。

1) 引入感兴趣区域转换模块,使得DFSNet算法在定向框检测过程中不需要使用大量不同角度、比例、尺度的锚,算法在DOTA数据集上mAP达到71.17%,比基础模型RetinaNet提高了3.12%。

2) 使用可变形对齐卷积解决了卷积神经网络受制于固定的空间结构的问题,使得DSFNet算法可以针对不同尺度、不同长宽比、不同形状的物体动态的选择特征,实现了对物体的聚焦。使用可变形对齐卷积,算法在DOTA数据集上mAP达到73.18%,性能进一步提高了2.01%。

3) 使用感受野自适应模块,动态调整神经元的感受野,算法在DOTA数据集上mAP达到74.04%,性能进一步提高了0.86%。

#### 参考文献 (References)

[1] 王彦情, 马雷, 田原. 光学遥感图像舰船目标检测与识别综述[J]. 自动化学报, 2011, 37(9): 1029-1039.  
WANG Y Q, MA L, TIAN Y. State-of-the-art of ship detection and recognition in optical remotely sensed imagery[J]. Acta Automatica Sinica, 2011, 37(9): 1029-1039(in Chinese).

[2] XIA G S, BAI X, DING J, et al. DOTA: A large-scale dataset for object detection in aerial images[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 3974-3983.

[3] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2014: 580-587.

[4] GIRSHICK R. Fast R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2015: 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] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2961-2969.

[7] DAI J F, QI H Z, XIONG Y W, et al. Deformable convolutional networks[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 764-773.

[8] DING J, XUE N, LONG Y, et al. Learning RoI Transformer for oriented object detection in aerial images[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2019: 2849-2858.

[9] LIU Z K, HU J G, WENG L B, et al. Rotated region based CNN for ship detection[C]//2017 IEEE International Conference on Image Processing(ICIP). Piscataway: IEEE Press, 2017: 900-904.

[10] MA J Q, SHAO W Y, YE H, et al. Arbitrary-oriented scene text detection via rotation proposals[J]. IEEE Transactions on Multimedia, 2018, 20(11): 3111-3122.

[11] HAN J M, DING J, LI J, et al. Align deep features for oriented object detection[EB/OL]. (2021-07-12)[2021-07-12].<https://arxiv.org/abs/2008.09397>.

[12] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 2117-2125.

[13] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2980-2988.

[14] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 770-778.

[15] LI Y H, CHEN Y T, WANG N Y, et al. Scale-aware trident networks for object detection[C]//Proceedings of the IEEE/CVF Inter-

- national Conference on Computer Vision. Piscataway: IEEE Press, 2019: 6054-6063.
- [16] PAN X J, REN Y Q, SHENG K K, et al. Dynamic refinement network for oriented and densely packed object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 11207-11216.
- [17] IOFFE S, SZEGEDY C. Batch Normalization: Accelerating deep network training by reducing internal covariate shift[C]//International Conference on Machine Learning, 2015: 448-456.
- [18] TAN M X, PANG R M, LE Q V. EfficientDet: Scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 10781-10790.
- [19] PASZKE A, GROSS S, MASSA F, et al. PyTorch: An imperative style, high-performance deep learning library[EB/OL]. (2019-12-03)[2021-06-01].<https://arxiv.org/abs/11912.01703>.
- [20] ZHANG G J, LU S J, ZHANG W. CAD-Net: A context-aware detection network for objects in remote sensing imagery[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(12): 10015-10024.
- [21] YANG X, YANG J R, YAN J C, et al. SCRDet: Towards more robust detection for small, cluttered and rotated objects[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE Press, 2019: 8232-8241.
- [22] YANG X, LIU Q Q, YAN J C, et al. R3Det: Refined single-stage detector with feature refinement for rotating object[EB/OL]. (2020-12-08)[2021-06-01].<https://arxiv.org/abs/1908.05612v1>.

## Remote sensing target detection based on dynamic feature selection

CHEN Chao, ZHAO Wei\*

(School of Electronic and Information Engineering, Beihang University, Beijing 100191, China)

**Abstract:** In the field of remote sensing image target detection, there still are challenges in oriented object detection. Convolutional neural network is subject to a fixed spatial structure when extracting information, and sampling locations cannot focus on objects. The scale of the remote sensing image varies greatly, and different objects require receptive fields of different scales to obtain feature map. Meanwhile, feature map with a single-scale receptive field cannot contain all effective information. In response to the first problem, deformable alignment convolution was proposed, which can first adjust the sampling locations according to the region of interest, and further learn slight offsets according to feature map, so that sampling locations can focus on objects and realize dynamic feature selection. For the second question, receptive field adaptive module based on deformable alignment convolution was proposed to fuse feature map with receptive fields of different scales and adaptively adjust the receptive field of neurons. Extensive experiments on public datasets showed that this method can improve the accuracy of remote sensing image target detection.

**Keywords:** deep learning; convolutional neural network; remote sensing image target detection; feature selection; receptive field

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0301

# 改进 YOLOv4 的表面缺陷检测算法

李彬, 汪诚\*, 丁相玉, 巨海娟, 郭振平, 李卓越

(空军工程大学 基础部, 西安 710038)

**摘要:** 为解决航空发动机部件表面缺陷检测精度低、检测速度慢的问题,提出一种改进的 YOLOv4 算法进行智能检测。在路径聚合网络 (PANet) 结构中融合浅层特征与深层特征, 增大特征检测尺度, 同时去除自下而上的路径增强结构, 提高小目标检测精度和整体检测速度; 根据各类缺陷数量不同的情况, 优化聚焦损失中的平衡参数, 增加权重因子调节各类缺陷的损失权重, 将改进后的聚焦损失代替分类误差中的交叉熵损失函数, 降低样本不平衡和难易样本对检测精度的影响。实验表明: 相比于原始 YOLOv4 算法, 改进后的 YOLOv4 算法在测试集上的平均精度均值 (mAP) 为 90.10%, 提高了 2.17%; 检测速度为 24.82 fps, 提高了 1.58 fps, 检测精度也高于单发多框检测 (SSD) 算法、EfficientDet 算法、YOLOv3 算法和 YOLOv4-Tiny 算法。

**关键词:** YOLOv4; 表面缺陷检测; 航空发动机; 小目标检测; 聚焦损失

**中图分类号:** V239; TP181

**文献标志码:** A **文章编号:** 1001-5965(2023)03-0710-08

飞机在飞行过程中, 航空发动机长时间处于高温、高压和高转速的条件下, 工作环境十分恶劣, 极易出现发动机叶片打伤、燃烧室裂纹等故障<sup>[1]</sup>。另外, 飞行过程中发生的外物撞击(如鸟类撞击、冰雹撞击等)也会造成发动机零部件的结构损伤<sup>[2]</sup>, 这些损伤给飞机的飞行安全构成了严重威胁。因此, 针对航空发动机部件表面进行有效的检测对保障飞行安全具有重要意义。

目前, 常规的检测方法主要有磁粉检测、渗透检测、涡流检测、孔探法检测、超声红外热成像检测等, 其中孔探法检测是发动机外场检测应用最多的技术手段<sup>[3]</sup>。但这些方法对操作人员专业素质要求很高, 并且检测效率较低。随着深度学习卷积神经网络的不断发展, 以单发多框检测 (single shot multibox detector, SSD) 算法<sup>[4]</sup>、YOLO 系列<sup>[5-8]</sup>等算法为代表的单阶段目标检测算法和以 Faster R-CNN 算法<sup>[9]</sup>、区域全卷积网络 (region-based fully convolutional networks, R-FCN)<sup>[10]</sup>、Mask R-CNN<sup>[11]</sup> 等算法为

代表的两阶段目标检测算法在很多领域得到了广泛应用。文献 [12] 在 YOLOv3-Tiny 算法基础上增加特征映射模块, 同时在特征提取网络中引用残差网络, 能有效用于遥感图像舰船检测。文献 [13] 提出了一种残差网络与膨胀卷积相融合的 SSD 改进算法, 用于航空发动机安装工位时的定位检测。此外, 不少学者也将卷积神经网络应用到航空发动机表面缺陷检测领域。文献 [14] 使用 Faster R-CNN 和 SSD 模型对航空发动机缺陷样本进行训练, 为发动机缺陷检测提供了检测平台。文献 [15] 将改进先验框后的 R-FCN 用于航空发动机叶片裂纹检测, 并将改进特征金字塔后的 YOLOv3 用于裂纹识别。文献 [16] 对 SSD 模型进行了先验框聚类方法和使用底层卷积特征提取的改进, 用于航空发动机凸台缺陷的检测。以上文献虽然使用了深度学习的方法来检测航空发动机表面缺陷, 但仍存在检测精度不足、检测速度不快的问题。

对此, 本文将航空发动机部件表面常见的 4 类

收稿日期: 2021-06-04; 录用日期: 2021-10-11; 网络出版时间: 2023-03-10 09:47

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20211113.1325.003.html](https://kns.cnki.net/kcms/detail/11.2625.V.20211113.1325.003.html)

\*通信作者. E-mail: [warrant\\_74@126.com](mailto:warrant_74@126.com)

**引用格式:** 李彬, 汪诚, 丁相玉, 等. 改进 YOLOv4 的表面缺陷检测算法 [J]. 北京航空航天大学学报, 2023, 49 (3): 710-717.

LI B, WANG C, DING X Y, et al. Surface defect detection algorithm based on improved YOLOv4 [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 710-717 (in Chinese).

缺陷(裂纹、缺口、凹坑、划痕)作为检测目标, 采用 YOLOv4 算法作为基础模型, 在路径聚合网络(path aggregation network, PANet)结构中融合浅层特征与深层特征, 增大特征检测尺度, 同时去除自下而上的路径增强结构, 提高小目标检测精度和整体检测速度, 再将参数优化后的聚焦损失(focal loss)代替分类误差中的交叉熵损失函数, 降低样本不平衡和难易样本对检测精度的影响, 提高整体检测精度, 从而有效检测航空发动机部件表面的缺陷。

# 1 YOLOv4 算法

## 1.1 网络结构

YOLOv4 算法属于单阶段目标检测算法, 其将目标分类和目标定位集成在同一个网络中进行预测, 实现了“端到端”的检测。YOLOv4 算法的网络结构如图 1 所示。主干特征提取网络采用 CSPDarknet53<sup>[17]</sup>, CSPDarknet53 是在 Darknet53<sup>[7]</sup> 基础上

改进的, 其在残差单元结构中使用了跨阶段局部网络(cross stage partial network, CSPNet)<sup>[17]</sup>, 进一步增强网络的学习能力, 在保证精度的前提下降低内存成本和计算瓶颈, 同时也使用了 Mish<sup>[18]</sup> 激活函数代替原有的 Leaky ReLU<sup>[19]</sup> 激活函数。Mish 激活函数在负值区比 Leaky ReLU 更为平滑, 有利于梯度的计算和更新, 能获得更好的准确性和泛化能力。主干特征提取网络之后引入了空间金字塔池化(spatial pyramid pooling, SPP)结构<sup>[20]</sup>, SPP 结构利用不同的最大池化层结合进行处理, 池化核大小分别为 1×1、5×5、9×9、13×13, 能有效增加感受野, 分离出显著的上下文特征。PANet<sup>[21]</sup> 是在特征金字塔网络(feature pyramid network, FPN)<sup>[22]</sup> 基础上增加自下而上的路径增强(bottom-up path augmentation)结构, 缩短低层特征到高层特征的信息路径, 加快低层信息流通。检测头部 YOLO Head 沿用了 YOLOv3 中的 Head, 由 3×3 和 1×1 卷积层组成, 对 PANet 处理后输出的特征层进行结果预测。

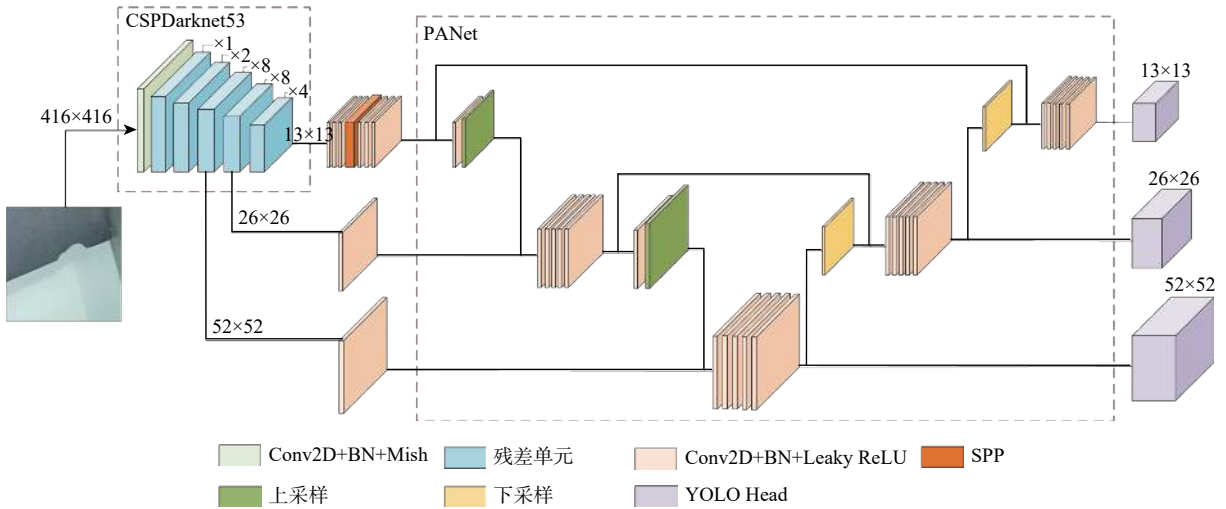


图 1 YOLOv4 网络结构

Fig. 1 YOLOv4 network structure

## 1.2 损失函数

YOLOv4 算法的损失误差包括置信度误差  $L_{conf}$ 、分类误差  $L_{cls}$  和回归框预测误差  $L_{loc}$ 。其中, 置信度误差和分类误差沿用了 YOLOv3 算法中的损失函数, 具体如式 (1) 和式 (2) 所示。而在回归框预测误差中采用完整交并比(complete intersection over union, CIoU)损失, CIoU 损失是在交并比(intersection over union, IoU)损失<sup>[23]</sup>、泛化交并比(generalized intersection over union, GIoU)损失<sup>[24]</sup>、距离交并比(distance-intersection over union, DIoU)损失<sup>[25]</sup> 上进一步改进的, 既考虑到检测框和目标重叠面积, 解决了边界框不重合时的问题, 也考虑到边界框中心点距离和宽高比的尺度信息, 提高了预测回归的精度和速度, 具体如式 (3) 和式 (4) 所示。

$$L_{conf} = - \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} [\bar{C}_i^j \ln C_i^j + (1 - \bar{C}_i^j) \ln (1 - C_i^j)] - \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{noobj} [\bar{C}_i^j \ln C_i^j + (1 - \bar{C}_i^j) \ln (1 - C_i^j)] \quad (1)$$

$$L_{cls} = - \sum_{i=0}^{S^2} I_{ij}^{obj} \sum_{c \in \text{classes}} \{ \bar{P}_i^j(c) \ln (P_i^j(c)) + (1 - \bar{P}_i^j(c)) \ln (1 - P_i^j(c)) \} \quad (2)$$

$$L_{loc} = \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} (2 - w^{gt} h^{gt}) [1 - CIoU(X, Y)] \quad (3)$$

$$CIoU(X,Y) = IoU(X,Y) - \frac{\rho^2(X_{ctr}, Y_{ctr})}{m^2} - uv \quad (4)$$

式中:  $S^2$  为网格数;  $B$  为每个网格中的预测框个数;  $I_{ij}^{obj}$  和  $I_{ij}^{noobj}$  分别为预测框包含目标和不含目标的指示值;  $\bar{c}$  为置信度真实值;  $C$  为预测置信度;  $\lambda_{noobj}$  为惩罚权重系数;  $\bar{P}(c)$  为单元格内目标属于类别  $c$  的实际概率;  $P(c)$  为预测类别  $c$  的概率;  $w^{gt}$ 、 $h^{gt}$  分别为真实框的宽、高;  $IoU(X,Y)$  为预测框  $X$  与真实框  $Y$  的交并比;  $\rho^2(X_{ctr}, Y_{ctr})$  为预测框与真实框的中心点欧氏距离;  $m$  为同时包含预测框和真实框的最小封闭区域的对角线距离;  $u$  为平衡调整参数;  $v$  为衡量长宽比一致性的参数。

## 2 改进的 YOLOv4 算法

### 2.1 改进的 PANet 结构

相比之前版本的 YOLO 系列算法, YOLOv4 使用的 PANet 结构可对不同特征层的语义信息进行融合补充, 适合检测不同尺寸的目标。通过 PANet 处理, 输出的特征层大小分别为原始输入图像的

1/32、1/16、1/8。例如, 输入图像尺寸大小为  $416 \times 416$ , 获得的特征层大小分别为  $13 \times 13$ 、 $26 \times 26$  和  $52 \times 52$ , 这在一定程度上可以满足不同尺寸目标的检测。但由于航空发动机部件表面存在许多微小缺陷, 缺陷目标在图像中像素占比较少, 但数量占比较多, 原有的 PANet 结构对微小缺陷目标仍缺乏有效的检测。因此, 本文在原有特征层基础上继续融合浅层特征与深层特征, 增大特征检测尺度, 形成一个新的特征检测层。特征层大小为输入图像的 1/4, 即  $104 \times 104$ 。增加新的特征检测层会导致网络结构参数数量的增加, 且自下而上的路径增强结构在小目标居多的情况下发挥的作用有限, 因此为减少参数冗余, 保证足够的检测速度, 去除 PANet 中自下而上的路径增强结构。改进后的 PANet 结构如图 2 蓝色虚线框所示。改进后的 PANet 在结构上与 FPN 相似, 但在特征融合过程中使用相邻层拼接 (concat) 代替相邻层相加 (add), 可达到更好的融合效果。改进后的结构不仅继承了原有结构的特征融合效果, 而且能在降低网络参数的同时获取更浅层的特征, 提高小目标的特征提取能力。

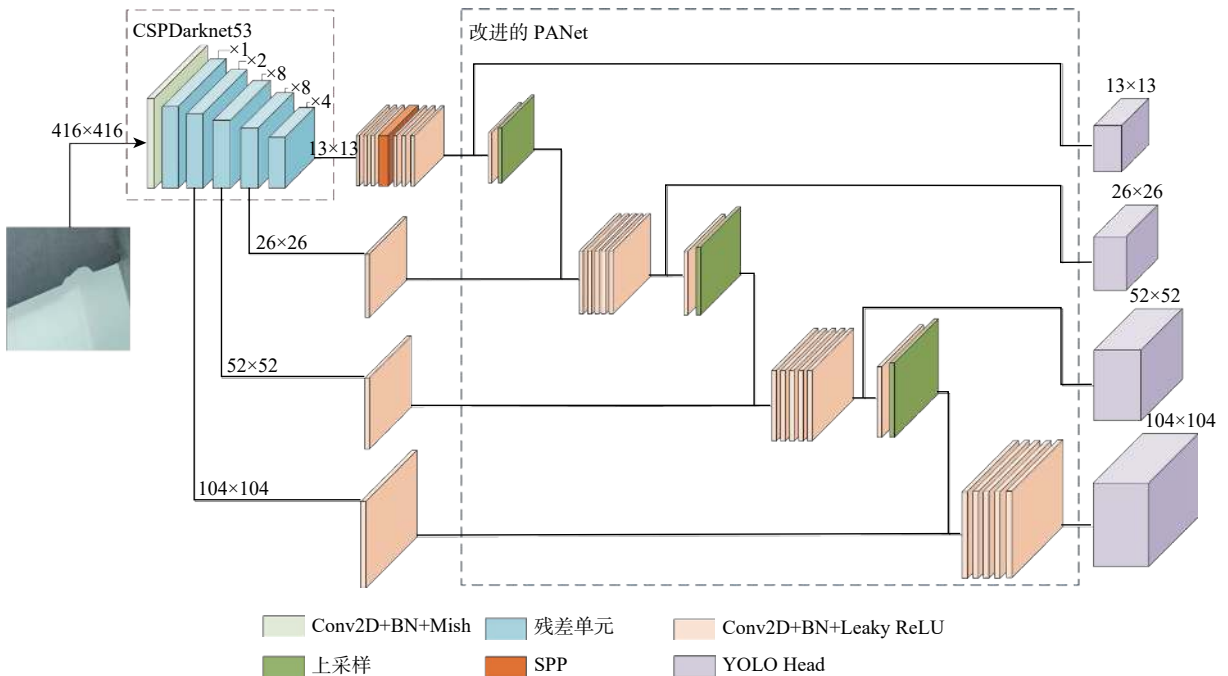


图 2 改进的 YOLOv4 结构

Fig. 2 Improved YOLOv4 structure

### 2.2 改进的聚焦损失

在 YOLOv4 中, 当输入图像大小为  $416 \times 416$  时, 3 个特征层共会生成 10647 个先验框。本文训练集中每张图像约包含 2.3 个目标, 绝大多数先验框中不包含目标, 容易导致训练时正负样本不平衡。YOLOv4 分类误差中采用的交叉熵损失函数难以解决样本不平衡的问题, 而聚焦损失<sup>[26]</sup>具有调节样本不平衡和挖掘困难样本(即难以学习的样

本)的作用, 聚焦损失函数如下:

$$FL(p_t) = -\alpha_t(1-p_t)^y \ln p_t \quad (5)$$

$$\alpha_t = \begin{cases} \alpha & y = 1 \\ 1 - \alpha & \text{其他} \end{cases} \quad (6)$$

$$p_t = \begin{cases} p & y = 1 \\ 1 - p & \text{其他} \end{cases} \quad (7)$$

式中:  $p \in [0,1]$  表示样本预测的概率;  $\alpha$  为用于控制

正负样本对总损失影响的平衡参数, 可调节样本不平衡对损失的影响,  $\alpha \in (0, 1)$ , 当正样本占比更少时, 可减小  $\alpha$  的取值, 加大样本预测错误时的惩罚;  $\gamma$  为用于控制难易样本对总损失的聚焦参数, 当  $\gamma > 1$  时, 可降低简单样本的损失, 使得模型更关注学习难分辨的样本。

虽然聚焦损失可以起到一定的平衡正负样本和难易样本对总损失影响的作用, 但在航空发动机部件表面缺陷训练集中, 各类缺陷数量不同, 检测难度也不同, 仅设定恒定不变的参数, 难以做到对各类缺陷进行针对性训练。为提升模型训练效果和检测精度, 本文采用聚焦损失代替交叉熵损失函数应用于分类误差, 并在其基础上根据各类缺陷数量对平衡参数  $\alpha'$  进行优化, 同时增加权重因子  $\beta_i$ , 改进方法如下:

$$FL(p_i) = -\alpha_i \beta_i (1 - p_i)^\gamma \ln p_i \quad (8)$$

$$\alpha' = \alpha_0 \left( \frac{n_i}{N} + 1 \right) \quad (9)$$

$$\alpha_i = \begin{cases} \alpha' & y = 1 \\ 1 - \alpha' & \text{其他} \end{cases} \quad (10)$$

$$\beta_i = \beta_0 e^{-\frac{n_i}{N}} \quad (11)$$

式中:  $\alpha_0$  和  $\beta_0$  分别为平衡参数和权重因子的初始化系数;  $n_i$  为第  $i$  类缺陷的目标数量;  $N$  为所有缺陷的目标总量。  $\alpha'$  可根据缺陷数量不同而动态调整缺陷预测正确和错误时的损失权重。

当某类缺陷数量较少, 其预测正确时损失权重会较低, 但预测错误时损失权重则会较高, 从而增大对该类缺陷预测错误的惩罚。  $\beta_i$  可根据缺陷数量不同而动态调整其整体损失在总损失中的权重, 当某类缺陷数量较少, 其整体损失权重就会更大, 这样能使得模型更好地学习小数量样本的特征。  $\gamma$  可用于调整难易样本对总损失的影响, 本文  $\gamma$  设为 3, 可放大低置信度样本的损失值, 使得模型更关注难分辨的样本。综合使用  $\alpha'$ 、 $\beta_i$  和  $\gamma$  三种损失调节参数, 可有效提高模型的检测精度。

### 3 实验与结果分析

#### 3.1 实验环境及参数设定

本文在 Windows 10 系统中采用 PyTorch 深度学习框架搭建网络, Python 版本: 3.6.10; Torch 版本: 1.2.0; CPU: Inter(R) Xeon(R) Gold 5218 CPU @ 2.30 GHz; GPU: NVIDIA GeForce RTX 2080 Ti。

实验设置的输入图像大小为 416×416, 采用 COCO 预训练模型进行参数初始化。整个训练过程为 200 轮 epoch, num\_workers 设置为 8, 权重衰减系数为 0.000 5。前 100 轮 epoch 的学习率设置为 0.001,

batch size 设置为 8, 后 100 轮 epoch 的学习率设置为 0.000 1, batch size 设置为 4。同时, 在每一轮 epoch 训练过程中同时开启 mosaic 数据增强和 random 数据增强, 提高网络的训练效果。

#### 3.2 实验数据处理

本文采用的实验数据是航空发动机部件表面缺陷图像, 包括裂纹、缺口、凹坑、划痕 4 种缺陷, 如图 3 所示。

由于采集的图像数量有限, 在训练前采用目标裁剪、水平翻转、曝光度调整和添加噪声(高斯噪声、椒盐噪声)的方式进行数据增强。数据增强后的图像数量为 3 700 张, 随机抽取图像分配给训练集、验证集和测试集, 比例为 6:2:2, 即训练集 2 220 张, 验证集 740 张, 测试集 740 张。表 1 显示了各数据集的目标数量, 其中, 裂纹、划痕的数量较多, 缺口、凹坑的数量较少。

采用 K-means 算法对目标框进行聚类, 图 4 为聚类中心为 12 的聚类结果。可知, 大多数缺陷都分布在尺寸较小的区域, 获得的先验框可覆盖大部

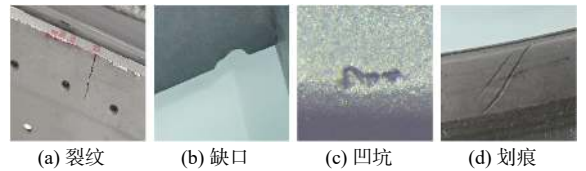


图 3 缺陷类型

Fig. 3 Defect types

表 1 数据集的目标数量

Table 1 Number of targets in dataset

缺陷类型	目标数量			总和
	训练集	验证集	测试集	
裂纹	1 579	580	595	2 754
缺口	848	272	270	1 390
凹坑	931	313	349	1 593
划痕	1 782	628	649	3 059

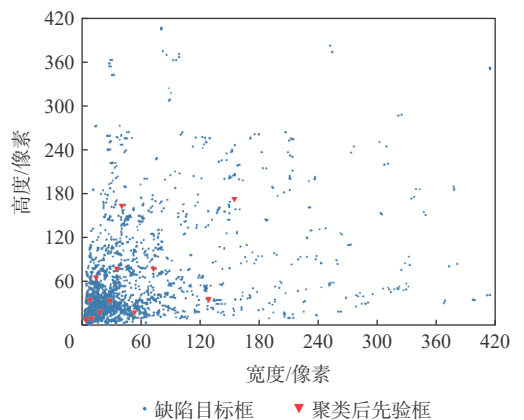


图 4 聚类中心为 12 的聚类结果

Fig. 4 Clustering results when cluster center is 12

分的缺陷目标,能满足不同特征层的检测。

### 3.3 结果分析

#### 3.3.1 评价准则

本文以预测框与目标框的 IoU 大于 0.5 作为评判成功检测到目标的依据。采用平均精度均值 (mean average precision, mAP) 和加权调和平均 ( $F_\beta$ ) 作为模型检测精度的评估指标, mAP 和  $F_\beta$  值越大,表明检测精度越高;采用单位时间内检测的图像数量 (Speed) 作为模型检测速度的评估指标, Speed 越大,表明检测速度越快。定义如下:

$$F_\beta = \frac{(1 + \beta^2)PR}{\beta^2 P + R} \quad (12)$$

$$AP = \int_0^1 P(R)dR \quad (13)$$

$$mAP = \frac{\sum AP}{K} \quad (14)$$

$$Speed = \frac{NumFigure}{TotalTime} \quad (15)$$

式中:  $P$  为准确率;  $R$  为召回率;为减少出现缺陷漏检的情况,保证召回率的重要程度,本文的  $\beta$  值设定为 1.2; AP 为平均精准度;  $K$  为缺陷类别数; TotalTime 为检测的总时间; NumFigure 为检测的图像数量。

#### 3.3.2 性能对比

为验证改进 PANet 结构和改进聚焦损失的检测效果,本文在实验中设置原始 PANet 结构+交叉熵损失函数 (YOLOv4)、改进 PANet 结构+交叉熵损失函数 (YOLOv4-A)、改进 PANet 结构后+改进的聚焦损失 (YOLOv4-B) 3 组实验进行对比,每种模型训练结束后选取最后 3 个 epoch 的训练权重进行测试,并将平均结果作为最终检测结果。

在 YOLOv4-B 中,选取不同的平衡参数  $\alpha_0$  和权重因子  $\beta_0$  会带来不同的训练效果,进而影响最终的检测精度。表 2 为不同参数组合下的检测精度对比情况。可知,当  $\alpha_0$  为 0.20 时,  $\beta_0$  越大, mAP 值就越大;当  $\alpha_0$  为 0.25 时,  $\beta_0$  越大, mAP 值先增大后减小;当  $\alpha_0$  为 0.30 时,  $\beta_0$  越大, mAP 值就越小。这是因为  $\alpha_0$  和  $\beta_0$  直接影响改进后的聚焦损失对不同类型缺陷的损失权重,选取合适的参数组合才能有效提高其调节损失的效果。对比不同参数组合下的 mAP 值可知,选取  $\alpha_0$  为 0.25、 $\beta_0$  为 3.00 的参数组合时检测精度最高。

本文选取最佳参数组合 ( $\alpha_0=0.25$ ,  $\beta_0=3.00$ ) 的 YOLOv4-B 作为标准与其他模型进行检测性能对比,对比结果如表 3 所示。

使用改进的 PANet 结构后, mAP 值由 87.93% 增加到 89.61%,增加了 1.68%。其中凹坑和划痕的

表 2 不同参数组合下的检测精度对比

Table 2 Comparison of detection accuracy under different parameter combinations

$\alpha_0$	$\beta_0$	mAP/%	AP/%			
			裂纹	缺口	凹坑	划痕
0.20	2.50	88.02	85.04	86.91	91.73	88.38
0.20	3.00	88.24	85.94	87.74	90.66	88.60
0.20	3.50	89.01	84.90	90.62	92.18	88.32
0.25	2.50	89.09	86.40	90.92	90.07	88.95
0.25	3.00	90.10	87.09	91.21	92.65	89.45
0.25	3.50	88.49	84.55	90.09	90.27	89.06
0.30	2.50	89.00	86.62	87.66	92.29	89.41
0.30	3.00	88.62	86.86	90.65	89.30	87.68
0.30	3.50	88.34	85.98	89.53	89.91	87.94

表 3 不同模型的检测性能对比

Table 3 Comparison of detection performance of different models

网络结构	mAP/%	Speed/fps	AP/%			
			裂纹	缺口	凹坑	划痕
YOLOv4	87.93	23.24	86.64	92.29	86.02	86.78
YOLOv4-A	89.61	24.94	86.36	93.12	89.10	89.86
YOLOv4-B	90.10	24.82	87.09	91.21	92.65	89.45

注: fps 表示帧/s。

检测效果提升较大,凹坑的 AP 值从 86.02% 上升到 89.10%,划痕的 AP 值从 86.78% 上升到 89.86%,二者 AP 值提升均超过 3%。缺口的 AP 值从 92.29% 上升到 93.12%,提升了 0.83%。这是因为使用的航空发动机部件表面缺陷数据集中存在较多像素占比小的缺陷,特别是发动机叶片容易产生小凹坑,机匣表面容易出现小划痕,使用改进的 PANet 结构提升了模型对小目标的检测精度。而裂纹这类缺陷在所有缺陷中数量占比较多,且形状多为细长型,部分目标标注时目标框较大,增加新特征层带来的参数冗余导致这类缺陷检测精度略有下降。各类模型的  $F_\beta$  结果对比如图 5 所示。图 5 中的  $F_\beta$  值变化规律与 AP 值类似,说明改进 PANet 结构整体上提高了检测精度,且缺陷遗漏率也有所降低。

改进 PANet 结构后检测速度由 23.24 fps 提升到 24.94 fps,提升了 1.70 fps。去除自下而上的路径增强结构降低了网络结构的参数量,从而提高了检测速度。具体的参数变化如表 4 所示。原始 YOLOv4 网络中总参数量为 63 953 841,参数容量为 243.96 MB;改进后总参数量为 48 971 468,参数容量为 186.81 MB,总参数量相比减少了 23.43%,参数容量相比也降低了 23.43%。

对比 YOLOv4-A 和 YOLOv4-B 的结果可知,采用改进的聚焦损失作为分类误差的损失函数后,进一步提升了 mAP 值,从 89.61% 提升至 90.10%,提

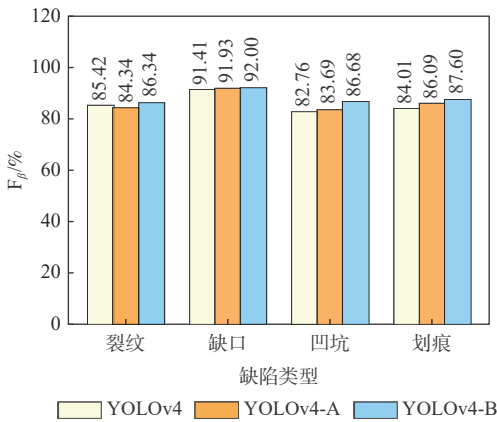


图 5 不同模型的  $F_\beta$  对比

Fig. 5 Comparison of  $F_\beta$  of different models

表 4 不同模型的网络结构参数数量对比

Table 4 Comparison of network structure parameters of different models

网络结构	总参数量	参数容量/MB
YOLOv4	63 953 841	243.96
YOLOv4-A	48 971 468	186.81
YOLOv4-B	48 971 468	186.81

升了 0.49%。其中, 裂纹、凹坑这 2 类缺陷 AP 值均有提高, 裂纹 AP 值提升了 0.73%, 凹坑 AP 值提升较大, 提升了 3.55%。由于裂纹这类缺陷中包含部分像素模糊的目标, 检测较为困难, 聚焦参数  $\gamma$  会增大这些困难样本的损失值, 使得模型更专注于学习这些困难样本; 凹坑这类缺陷目标较小, 不易分辨, 而且数量相对也较少, 平衡参数  $\alpha'$  会增大其预测错误时的损失, 同时权重因子  $\beta_i$  会进一步增大其整体损失占总损失的比重, 二者综合提高了这类缺陷的检测精度。划痕的 AP 值略有下降, 缺口的 AP 值有一定的下降, 主要是因为采用改进的聚焦损失后, 各类缺陷的损失在总损失中的占比进行了调整, 而数据集中划痕数量最多, 缺口的特征较易分辨, 使得其损失权重相比交叉熵损失函数有所下降, 进而对 AP 值有一定的影响。总体而言, 采用改进的聚焦损失后, 各类缺陷的 AP 值相对更为平衡。同时, 各类缺陷的  $F_\beta$  均有所提升, 说明模型的总体检测精度进一步提升, 各类缺陷漏检量也有所减少。在检测速度方面, YOLOv4-A 和 YOLOv4-B 检测速度基本一致。如表 4 所示, 使用改进的聚焦损失后并不影响网络结构的总参数量和参数容量。图 6 为各类缺陷样本的检测结果。其中, 测试图片是测试集中选取的缺陷图像, 测试结果包括缺陷的位置、类型和置信度。

表 5 为不同算法在航空发动机部件表面缺陷数据集下的检测性能对比。SSD、YOLOv3 是常见的单阶段目标检测算法, EfficientDet<sup>[27]</sup> 是谷歌大脑

于 2019 年提出的目标检测算法, 采用 EfficientNet<sup>[28]</sup> 作为主干特征提取网络, 包括 D0 至 D7 共 8 种结构, 本文采用 D0、D1、D2 三种结构进行对比。YOLOv4-Tiny 是 YOLOv4 的简化版, 仅使用 2 个特征层进行分类与回归预测。

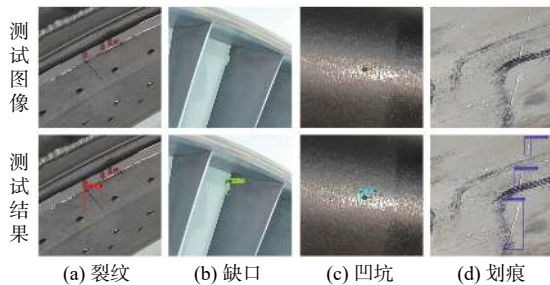


图 6 各类缺陷样本的检测结果

Fig. 6 Test results of various defect samples

表 5 不同算法的检测性能对比

Table 5 Comparison of detection performance of different algorithms

网络结构	mAP/%	Speed/fps
SSD	60.07	33.07
EfficientDet-D0	49.39	13.07
EfficientDet-D1	57.71	10.52
EfficientDet-D2	63.40	9.45
YOLOv3	86.25	30.37
YOLOv4-Tiny	54.25	64.55
YOLOv4	87.93	23.24
YOLOv4-B	90.10	24.82

由表 5 可知, 在检测精度方面, YOLOv4-B 算法的 mAP 值均高于其他算法, 比 SSD 高 30.03%, 比 EfficientDet-D0、EfficientDet-D1、EfficientDet-D2 分别高 40.71%、32.39%、26.70%, 比 YOLOv3 高 3.85%, 比 YOLOv4-Tiny 高 35.85%, 比 YOLOv4 高 2.17%。在检测速度方面, YOLOv4-B 算法的检测速度比 EfficientDet-D0、EfficientDet-D1、EfficientDet-D2、YOLOv4 更快, 但低于 SSD、YOLOv3、YOLOv4-Tiny。主要原因是: SSD 和 YOLOv3 的主干特征提取网络分别为 VGG 和 Darknet53, 其参数量均低于 YOLOv4, YOLOv4-Tiny 是 YOLOv4 的简化版, 其参数量也远低于 YOLOv4, 因此在检测速度上更有优势。但考虑到航空发动机部件缺陷检测对精度的要求, 综合来说, YOLOv4-B 检测性能要优于其他算法。

## 4 结 论

本文提出了一种改进 YOLOv4 的表面缺陷检测算法, 以航空发动机部件表面缺陷作为检测对象, 并与其他常用算法进行性能对比, 验证了该算法的有效性。结论如下:

- 1) 通过在 PANet 结构中融合浅层特征与深层

特征,增大特征检测尺度,同时去除自下而上的路径增强结构,mAP值由87.93%提升至89.61%,提升了1.68%。其中,凹坑和划痕的检测精度提升最大,二者AP值提升均超过3%。检测速度由23.24 fps提升至24.94 fps,提升了1.70 fps。

2)采用参数优化后的聚焦损失代替分类误差中的交叉熵损失函数后,降低了样本不平衡和难易样本对检测精度的影响,mAP由89.61%提升至90.10%。其中,凹坑AP值提升了3.55%,裂纹AP值提升了0.73%。

3)在检测精度方面,改进后的YOLOv4算法检测的mAP值高于SSD、EfficientDet、YOLOv3和YOLOv4-Tiny算法,更适合用于检测航空发动机部件表面缺陷。

### 参考文献 (References)

- [1] 李华. 基于孔探图像分析的航空发动机故障诊断专家系统研究[D]. 南京: 南京航空航天大学, 2015.  
LI H. Research on aeroengine fault diagnosis expert system based on endoscopic image analysis[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2015(in Chinese).
- [2] 关玉璞, 陈伟, 高德平. 航空发动机叶片外物损伤研究现状[J]. *航空学报*, 2007, 28(4): 851-857.  
GUAN Y P, CHEN W, GAO D P. Present status of investigation of foreign object damage to blade in aeroengine[J]. *Acta Aeronautica et Astronautica Sinica*, 2007, 28(4): 851-857(in Chinese).
- [3] 何嘉辉, 张栋善, 赵成, 等. 航空发动机叶片裂纹检测技术及应用分析[J]. *内燃机与配件*, 2020(15): 151-152.  
HE J H, ZHANG D S, ZHAO C, et al. Detection technology and application analysis of aero-engine blade crack[J]. *Internal Combustion Engine & Parts*, 2020(15): 151-152(in Chinese).
- [4] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//Proceedings of European Conference on Computer Vision. Berlin: Springer, 2016: 21-37.
- [5] 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.
- [6] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 6517-6525.
- [7] REDMON J, FARHADI A. YOLOv3: An incremental improvement[C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 1804-0276.
- [8] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 10934.
- [9] 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.
- [10] DAI J F, LI Y, HE K M, et al. R-FCN: Object detection via region-based fully convolutional networks[C]//Proceedings of the 30th International Conference on Neural Information Processing Systems. New York: ACM, 2016: 379-387.
- [11] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN [C]//IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 2980-2988.
- [12] 王玺坤, 姜宏旭, 林珂玉. 基于改进型YOLO算法的遥感图像舰船检测[J]. *北京航空航天大学学报*, 2020, 46(6): 1184-1191.  
WANG X K, JIANG H X, LIN K Y. Remote sensing image ship detection based on modified YOLO algorithm[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2020, 46(6): 1184-1191(in Chinese).
- [13] 陈科山, 郝宇, 何泓波, 等. 基于R-D SSD模型航空发动机安装工位检测算法[J]. *北京航空航天大学学报*, 2021, 47(4): 682-689.  
CHEN K S, HAO Y, HE H B, et al. Detection algorithm of aero-engine installation station based on R-D SSD model[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2021, 47(4): 682-689(in Chinese).
- [14] 旷可嘉. 深度学习及其在航空发动机缺陷检测中的应用研究[D]. 广州: 华南理工大学, 2017.  
KUANG K J. Research on deep learning and its application on the defects detection for aero engine[D]. Guangzhou: South China University of Technology, 2017(in Chinese).
- [15] 李浩. 基于图像识别的航空发动机叶片裂纹检测研究[D]. 成都: 电子科技大学, 2019.  
LI H. Research on the blade crack detection of aero-engine based on image recognition[D]. Chengdu: University of Electronic Science and Technology of China, 2019(in Chinese).
- [16] 陈为, 梁晨红. 基于改进SSD的航空发动机目标缺陷检测[J]. *控制工程*, 2021, 28(12): 2329-2335.  
CHEN W, LIANG C H. Aeroengine target defect detection based on improved SSD[J]. *Control Engineering of China*, 2021, 28(12): 2329-2335(in Chinese).
- [17] WANG C Y, MARK L H Y, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Piscataway: IEEE Press, 2020: 1571-1580.
- [18] MISRA D. Mish: A self regularized non-monotonic neural activation function[EB/OL]. (2020-08-13) [2021-06-01]. <https://arxiv.org/abs/1908.08681>.
- [19] MAAS A L, HANNUN A Y, NG A Y. Rectifier nonlinearities improve neural network acoustic models[C]//Proceedings of International Conference on Machine learning, 2013: 1-6.
- [20] 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.
- [21] 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.
- [22] 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: 2117-2125.

- [23] YU J H, JIANG Y N, WANG Z Y, et al. UnitBox: An advanced object detection network[C]//Proceedings of the 24th ACM International Conference on Multimedia. New York: ACM, 2016: 516-520.
- [24] REZATOFIGHI H, TSOI N, GWAK J, et al. Generalized intersection over union: A metric and a loss for bounding box regression [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2019: 658-666.
- [25] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2020, 34(7): 12993-13000.
- [26] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [27] TAN M X, PANG R M, LE Q V. EfficientDet: Scalable and efficient object detection[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2020: 10778-10787.
- [28] TAN M X, LE Q V. EfficientNet: Rethinking model scaling for convolutional neural networks [EB/OL].(2020-09-11) [2021-06-01]. <https://arxiv.org/abs/1905.11946>.

## Surface defect detection algorithm based on improved YOLOv4

LI Bin, WANG Cheng\*, DING Xiangyu, JU Haijuan, GUO Zhenping, LI Zhuoyue

(Fundamentals Department, Air Force Engineering University, Xi'an 710038, China)

**Abstract:** In order to enhance the accuracy and speed of surface defect detection of aeroengine components, an improved YOLOv4 algorithm is proposed for intelligent detection. Firstly, shallow features and deep features were integrated into the path aggregation network (PANet) to improve the feature detection scale, and the bottom-up path augmentation structure was removed to increase the accuracy of small target detection and the overall detection speed. Then, according to the numbers of various defects, the balance parameter of the focal loss was optimized, and a weight factor was added to adjust the loss weight of various defects. The improved focal loss was used to replace the cross-entropy loss function in the classification error, thus reducing the impact of imbalanced samples and hard and easy samples on the detection accuracy. The experimental results show that the mean average precision (mAP) of the improved YOLOv4 on the test set is 90.10%, which is 2.17% higher than that of the traditional YOLOv4, and the detection speed is 24.82 fps, which is increased by 1.58 fps. The detection accuracy is also higher than other algorithms including single shot multibox detector (SSD), EfficientDet, YOLOv3 and YOLOv4-Tiny.

**Keywords:** YOLOv4; surface defect detection; aeroengine; small target detection; focal loss

<http://bhxb.buaa.edu.cn> [jbuaa@buaa.edu.cn](mailto:jbuaa@buaa.edu.cn)

DOI: 10.13700/j.bh.1001-5965.2021.0303

# 一种基于组合赋权法的小波去噪质量评价方法

李晋斐, 赵冬青\*, 王栋民, 蔡聪聪, 贾晓雪, 张乐添

(信息工程大学 地理空间信息学院, 郑州 450001)

**摘 要:** 针对传统质量评价指标在小波阈值去噪中理论依据不足的问题, 提出了一种基于组合赋权法的小波去噪质量评价方法, 能够为小波去噪参数的选择提供有效评价。通过分析在真值未知情况下均方根误差 (RMSE)、信噪比 (SNR)、平滑度等单项指标的特点, 选取 RMSE 与平滑度作为小波去噪指标, 对其进行归一化处理, 采用信息熵权与变异系数的方法进行组合赋权, 将归一化指标与对应权值线性组合, 得到一种新的指标即为复合评价指标, 其值越小, 说明去噪效果越好, 所选参数越优。仿真实验表明, 在真值已知情况下, 该评价指标具有更高的准确性, 能够适用于不同的分解层数与小波基函数, 具有比传统方法更好的适用性; 实测数据表明, 所提方法得出的小波去噪峰值域更加光滑, 波形更加平稳, 去噪效果更佳。

**关键词:** 小波阈值去噪; 质量评价指标; 均方根误差; 平滑度; 组合赋权; 惯性元件随机误差

中图分类号: V249.3; P227.9

文献标志码: A

文章编号: 1001-5965(2023)03-0718-08

在多传感器融合定位中, 由于惯性导航系统能够持续输出姿态和位置信息, 一般多以惯性导航误差模型建立状态方程。惯性元件随机误差对惯性导航系统产生的影响较大, 因此需有效降低惯性元件的随机误差, 达到提高惯性导航系统测量精度的目的。小波变换具有在时域和频域同时表征局部特征的能力<sup>[1]</sup>, 采用多分辨率特性, 使得去除噪声的过程中保持原始信息, 能够很好地刻画信号的非平稳特征, 且不需要建立精确的随机误差模型, 即可分析处理非平稳信号, 故非常适用于处理陀螺的状态信息<sup>[2]</sup>。

所有小波去噪算法都有共同的结构, 但其有效性非常依赖相关参数的选择, 主要涉及小波基函数、阈值选取规则、阈值函数、分解尺度等<sup>[3]</sup>。该过程中, 很多学者关注了如何改进阈值函数来提高去噪能力, 并取得了一定效果, 而对于其他参数的选择却十分困难<sup>[4]</sup>, 更多的是以经验或直接给定参数

来得到最终结果, 很少分析这些参数对最终效果的影响, 也没有明确的小波基函数与分解层数等最优参数的确定方法。例如, 对于惯性元件随机误差的小波去噪参数, 文献 [5] 直接选定 db4 小波基, 3 尺度分解; 文献 [6] 直接选定 db6 小波基, 5 尺度分解。然而, 根据多次实验对比发现, 采用小波去噪时, 去噪效果随着每个尺度下所选择的小波变化而变化, 很难直接确定分解层数; 而分解层数与小波基函数的选取对去噪效果的影响不可忽略。要使去噪效果最好, 必须选择最佳的小波参数, 通常选择信噪比 (signal-noise ratio, SNR)、相关系数  $R$ 、均方根误差 (root mean square error, RMSE)、平滑度  $r$  等参数作为指标来衡量小波去噪效果<sup>[7]</sup>。在实际应用中, 由于信号真值是未知的, 上述单一的评价标准通常无法准确判断去噪效果, 具有很大的局限性。因此, 陶珂和朱建军<sup>[8]</sup>通过融合去噪过程中信号 RMSE、SNR 和平滑度的变化量构造具有收敛特

收稿日期: 2021-06-07; 录用日期: 2021-07-18; 网络出版时间: 2021-09-24 17:55

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210924.1200.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210924.1200.001.html)

基金项目: 国家自然科学基金 (41774037)

\* 通信作者. E-mail: [dongqing.zhao@hotmail.com](mailto:dongqing.zhao@hotmail.com)

**引用格式:** 李晋斐, 赵冬青, 王栋民, 等. 一种基于组合赋权法的小波去噪质量评价方法 [J]. 北京航空航天大学学报, 2023, 49 (3): 718-725. LI J F, ZHAO D Q, WANG D M, et al. A quality evaluation method for wavelet denoising based on combinatorial weighting method [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 718-725 (in Chinese).

性的新统计算,通过识别拐点判断最佳分解尺度。朱建军等<sup>[9]</sup>通过变异系数法进行定权,选择信号的平滑度与RMSE 2种指标进行组合,进而指导小波去噪相关参数的选择。王旭和王昶<sup>[10]</sup>利用熵权法将RMSE与平滑度变化量线性组合得到复合评价指标,利用其收敛特性确定最优分解层数。实际计算表明,在真值未知时,这些方法都有较好效果,但在准确度上存在不同程度的缺陷,且在判断收敛性方面缺乏相应的理论依据。因此,本文构建了一种新的评价指标,并通过惯性元件随机误差的分析,验证其在小波降噪过程中对相关参数选择的有效性。

## 1 小波阈值去噪质量分析

小波阈值去噪算法的实质是抑制信号的无用部分,最大限度保留信号的有用部分。利用小波分析的多尺度特性,把含有噪声的信号分解到多个尺度上,再在每一尺度上分别设定对应的阈值,其中,将噪声信号归纳为小于该阈值的小波系数,将有用信号归纳为大于该阈值的小波系数,分别去除各尺度上的噪声信号,同时保留有用信号,即可重构出去噪后的信号<sup>[11]</sup>。可以看出,同阈值函数一样,小波基函数和分解层数的选择对小波阈值的去噪结果也有很大影响。

### 1.1 小波阈值去噪模型

一个含噪声的一维信号可表示为<sup>[12]</sup>

$$f(t) = s(t) + \sigma w(t) \quad (1)$$

式中: $f(t)$ 为含有噪声的信号; $s(t)$ 为纯净信号; $\sigma$ 为噪声标准差; $w(t)$ 为噪声信号。

有用信号一般为低频信号,而噪声通常为高频信号。例如,对原始信号进行3层分解:

$$S = cA_1 + cD_1 = (cA_2 + cD_2) + cD_1 = [(cA_3 + cD_3) + cD_2] + cD_1 \quad (2)$$

式中: $cA_i$ 为分解的近似部分, $cD_i$ 为分解的细节部分, $i = 1, 2, 3$ ,噪声部分通常包含在 $cD_3$ 、 $cD_2$ 、 $cD_1$ 中。

总结去噪过程,主要包括以下3个步骤<sup>[13]</sup>:

1) 对原始信号进行小波分解。选择最佳的小波基和分解层数,将原始信号进行多尺度分解得到相对应的小波系数。

2) 小波分解阈值量化。选择相对应的阈值及阈值函数,对小波系数进行阈值量化处理。

3) 小波系数重构。将阈值化处理后的小波系数进行重构,得到去噪后的信号。

从上可知,最优的小波阈值去噪方案,不仅与阈值函数的选取有关,小波基函数、分解层数的选择都直接影响信号去噪的质量,而对于小波基函数

与分解层数的选择大多基于经验或直接给定得到,未充分分析其对去噪效果的影响,如果能选择最优的小波基函数和分解层数,对于提高去噪效果将会有明显的帮助<sup>[14]</sup>。

### 1.2 小波去噪质量评价指标

小波基函数及分解层数都直接影响小波去噪效果,而使用不同的小波基与分解层数,去噪后的效果也不相同。如果分解尺度过小,则信号中仍存在较多的噪声数据,无法获得理想的去噪结果;如果分解尺度过大,则会将信号中的部分细节信息当做噪声删除,容易造成信号失真,同时增加计算复杂度,而去噪效果却没有显著改善<sup>[15]</sup>。因此,在去噪效果和分解层数之间找到平衡非常重要,需要构建一种全面的质量评价指标来衡量去噪效果。

当前, RMSE、SNR、平滑度 $r$ 等指标常被用于分析小波去噪的效果<sup>[16]</sup>。RMSE指去噪信号与原始信号之间方差的平方根,其值越小说明去噪效果越优;SNR是原始信号功率与噪声信号功率之间的比值,其值越大说明去噪效果越优<sup>[17]</sup>;平滑度是去噪信号一阶差分与原始信号一阶差分之间方差根比值,代表信号整体的变异信息,其值越小说明去噪效果越优。

为了评估这些指标的有效性,本文采用一组真值已知、小波去噪专用的测试数据即多普勒信号和一组真值未知的陀螺输出数据来进行小波分析,同陀螺信号类似,多普勒信号是一种非平稳信号,在原始信号上加入8 dB的高斯白噪声,以及系数为0.1的随机游走,信号长度为1 024个点,采样频率为1 Hz。由于已知未加噪声的纯净信号,可求得重构信号与理论纯净信号的RMSE及SNR,当RMSE最小且SNR最大时,所配参数即为最佳的去噪参数。真实信号为采集的加拿大诺瓦泰公司SPAN-ISA-100C型高精度光纤惯导的静态陀螺数据,时长为5 h,采样频率为200 Hz。利用上述评价指标来判断最优分解层数,对于2类信号,都以sym 4作为小波基函数,分解层数为2~10。图1为真值已知情况下多普勒信号小波去噪的质量评价结果,从RMSE和SNR结果中可以看出,最优分解层数为6层,平滑度无法判定。图2为真值未知情况下陀螺输出数据小波去噪的质量评价结果,此时无法找出最优分解层数,这是因为所测信号的真值是未知的,参与计算的是含有噪声的原始信号与去噪信号,如果未去掉任何噪声或去噪效果差,但得到的RMSE值却很小,SNR值很大,这与评价依据不一致,纵然平滑度是减小的,但却无法判断最优层数。通过分析可得在真值未知情况下3种评价指标特点如表1所示。

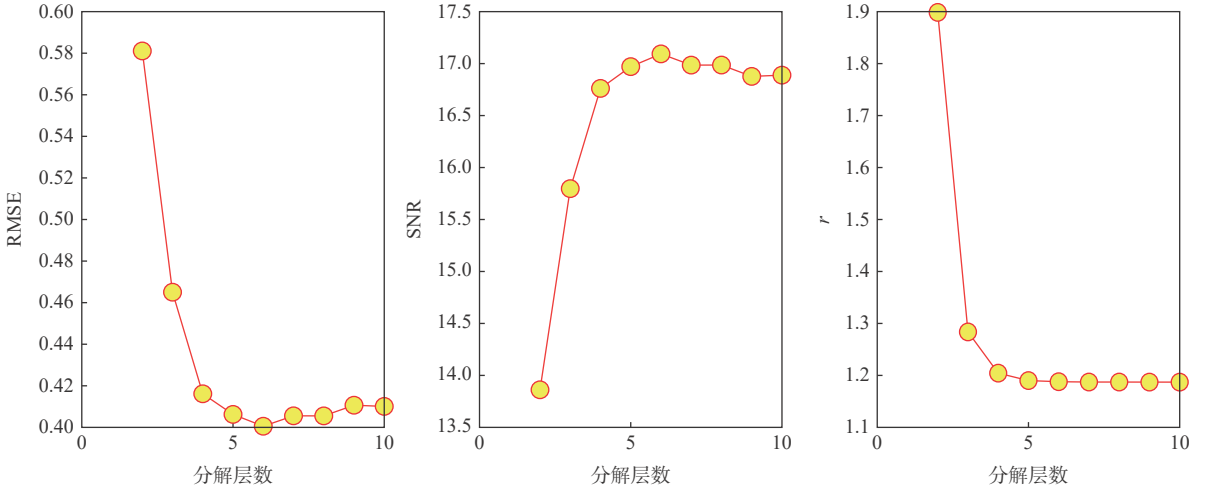


图1 真值已知的多普勒信号的单一评价指标趋势

Fig. 1 Trends of single evaluation index of Doppler signal with known truth value

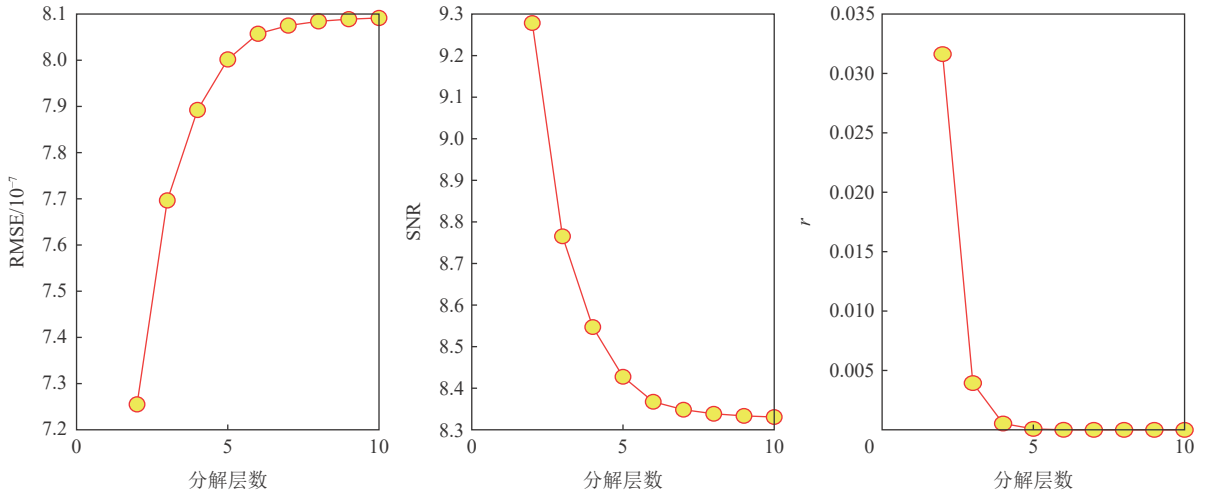


图2 真值未知的陀螺输出数据的单一评价指标趋势

Fig. 2 Trends of single evaluation index of gyro output data with unknown truth value

表1 真值未知时评价指标特点

Table 1 Characteristics of evaluation indexes when truth value is unknown

评价指标	关注信息	与分解层数相关性
RMSE	细节信息	正相关
SNR	细节信息	负相关
平滑度	近似信息	负相关

因此,单一的评价指标具有一定的欺骗性,难以满足惯性元件误差分析的需求。若将多类评价指标进行赋权融合,以更加全面的视角来构建复合评价指标,从而满足在真值未知情况下小波去噪参数的选择。

## 2 复合评价指标构建

考虑到模型需简单且具有可行性,根据 RMSE、SNR 及平滑度的几何与物理意义,选取 RMSE 与平滑度 2 个指标来构建复合评价指标。主要原因在

于: RMSE 和 SNR 都是关注信号的细节信息,而平滑度关注的是信号近似信息,为反映去噪信号的整体特征,应使用平滑度指标,同时,考虑到第 1 节实验中 SNR 值与平滑度值随分解层数的增加都减小,而 RMSE 值随分解层数的增加而增大,故用平滑度与 RMSE 构建复合评价指标。随着分解层数不断增加,必定出现一个极值,即去噪信号的近似信息与细节信息达到了最优的组合,对应的是最优分解层数。

由于 RMSE 与平滑度值变化的范围不同,先将其归一化<sup>[18]</sup>,使 2 个指标都处于 [0,1] 区间。归一化算法为

$$\text{PRMSE} = \frac{\text{RMSE} - \min(\text{RMSE})}{\max(\text{RMSE}) - \min(\text{RMSE})} \quad (3)$$

$$\text{Pr} = \frac{r - \min(r)}{\max(r) - \min(r)} \quad (4)$$

式中: PRMSE 和 Pr 分别为归一化后 RMSE 和平滑度

的值;  $\min$ 和 $\max$ 分别为最小值和最大值函数。

由于这2个指标对信号特征的描述程度不同,在复合过程中权重也不一致,需要进行赋权处理。目前,有关评价指标权系数的确定方法很多,主要分为主观赋权法、客观赋权法、主客观集成赋权法3类<sup>[19]</sup>,其中客观赋权法主要以其强大的数学理论做支撑,相比其他方法更加可靠。本文基于信息熵权法和变异系数法2种客观赋权思想,提出了一种基于熵权与变异系数组合赋权的复合评价指标。

首先,以RMSE为例,给出熵权法定权的计算公式为

$$P_i^{\text{PRMSE}} = \text{PRMSE}(i) / \sum_{i=1}^N \text{PRMSE}(i) \quad (5)$$

$$e_{\text{PRMSE}} = - \left( \frac{1}{\ln N} \right) \sum_{i=1}^N P_i^{\text{PRMSE}} \ln P_i^{\text{PRMSE}} \quad (6)$$

$$w_{\text{PRMSE}} = \frac{1 - e_{\text{PRMSE}}}{(1 - e_{\text{PRMSE}}) + (1 - e_{\text{Pr}})} \quad (7)$$

式中:  $P_i^{\text{PRMSE}}$ 为第*i*个尺度的均方根误差占总均方根误差的比重;  $e_{\text{PRMSE}}$ 和 $e_{\text{Pr}}$ 分别为RMSE和平滑度的熵值;  $N$ 为序列长度;  $w_{\text{PRMSE}}$ 为RMSE的权值。同样,平滑度的权值计算原理同上。

其次,给出变异系数法定权的计算公式:

$$\text{CV}_{\text{PRMSE}} = \frac{\sigma_{\text{PRMSE}}}{\mu_{\text{PRMSE}}} \quad (8)$$

$$\text{CV}_{\text{Pr}} = \frac{\sigma_{\text{Pr}}}{\mu_{\text{Pr}}} \quad (9)$$

$$W_{\text{PRMSE}} = \frac{\text{CV}_{\text{Pr}}}{\text{CV}_{\text{PRMSE}} + \text{CV}_{\text{Pr}}} \quad (10)$$

$$W_{\text{Pr}} = \frac{\text{CV}_{\text{PRMSE}}}{\text{CV}_{\text{PRMSE}} + \text{CV}_{\text{Pr}}} \quad (11)$$

式中:  $\sigma_{\text{PRMSE}}$ 和 $\mu_{\text{PRMSE}}$ 分别为RMSE的标准差和均值;  $\sigma_{\text{Pr}}$ 和 $\mu_{\text{Pr}}$ 分别为平滑度的标准差和均值;  $\text{CV}_{\text{PRMSE}}$ 和 $\text{CV}_{\text{Pr}}$ 分别为RMSE变异系数和平滑度变异系数;  $W_{\text{PRMSE}}$ 和 $W_{\text{Pr}}$ 分别为RMSE和平滑度的权值。

得到上述2种赋权方法所计算的权值后,为了弥补各自不足,最大限度地减少有用信息的损失,将熵权法与变异系数法相结合,构建组合赋权法,从而使指标的权值更加客观合理,得到最终权值,则有

$$\widehat{W}_{\text{PRMSE}} = \lambda W_{\text{PRMSE}} + (1 - \lambda) W_{\text{PRMSE}} \quad (12)$$

$$\widehat{W}_{\text{Pr}} = \lambda W_{\text{Pr}} + (1 - \lambda) W_{\text{Pr}} \quad (13)$$

式中:  $\lambda$ 为常值系数;  $\widehat{W}_{\text{PRMSE}}$ 和 $\widehat{W}_{\text{Pr}}$ 分别为组合赋权法得到的RMSE和平滑度的权值。

在本文中,由于熵权法和变异系数法没有明确的可靠性比较,设为0.5将组合赋权得到的RMSE与平滑度的权值分别与归一化得到的RMSE与平滑度

值相乘并作和,得到复合评价指标*H*,其表达式为

$$H = \widehat{W}_{\text{PRMSE}} \cdot \text{PRMSE} + \widehat{W}_{\text{Pr}} \cdot \text{Pr} \quad (14)$$

根据RMSE与平滑度的定义及定权过程可知,随着分解层数的增加,复合评价指标*H*将会出现一个极值,且这个极值为最小值,所对应的分解层数即为最优分解层数。

### 3 实验分析

#### 3.1 仿真实验

为了能够准确比较验证本文所提出的复合评价指标在判断最优参数方面的有效性,采用第1节真值已知的仿真信号进行分析。

分析时采用控制变量法,即在小波基不变的情况下,采用不同分解层数对含有噪声的信号进行处理,并选用不同的小波基分别对含噪信号进行处理。其他参数保持不变,阈值估计方法采用基于史坦的无偏似然估计原理(rigrsure)及软阈值处理函数。为验证方法的适用性,本文选取了文献[9]和文献[10]2种较为成熟的方法进行对比。

以sym4小波基及db5小波基为例,先计算多普勒信号在真值已知情况下(原始信号为无噪声的纯净信号)的RMSE及SNR,再计算其在真值未知情况下(原始信号为加噪声的信号)的RMSE与平滑度,并计算本文的复合评价指标*H*、采用文献[9]的指标*T*和文献[10]的指标*S*,结果如表2和表3所示。

从表2可以看出,选用sym4小波基进行去噪处理时,在真值已知情况下,当分解层数为6时, RMSE最小,同时SNR最大,说明最优分解层数为6;在真值未知情况下,随着分解层数的增加, RMSE不断增加,平滑度*r*不断降低,这与前文判断一致,无法判断最优分解层数。根据复合评价指标*H*,当分解层数为6时出现最小值,说明最优分解层数为6,这与实际情况相符;文献[9]的指标*T*也有类似的结论;而文献[10]的评价指标*S*从第5层开始变化率趋于平缓,判定最优分解层数为5,与实际情况不符,同时对于判断变化率是否趋于平缓具有一定程度的主观性。

从表3可以看出,选用db5小波基进行去噪处理时,在真值已知情况下最优分解层数为7,而在真值未知情况下,利用指标*H*得到的最优分解层数为7,与实际情况相符;指标*T*得到的最优分解层数为6,与实际情况不符;指标*S*得到的最优分解层数为7,与实际情况相符。由此可以看到,复合评价指标*H*相比现有方法精度更高,具有更高的可信度。

为了更加全面地分析本文方法的可靠性,避免单一小波基函数对实验结果的误导,在真值未知情况下分别选择不同类型的小波基进行去噪分析,得

表2 不同分解层数利用 sym4 小波基处理的评价指标

Table 2 Evaluation indexes of sym4 wavelet basis processing for different decomposition layers

分解层数	真值已知		真值未知		$T$	$S$	$H$
	RMSE	SNR	RMSE	$r$			
2	0.581 1	13.861 7	0.778 1	0.103 0	0.866 5	0.890 3	0.899 5
3	0.465 0	15.796 9	0.844 7	0.069 7	0.212 3	0.256 6	0.193 3
4	0.416 2	16.761 3	0.856 8	0.065 4	0.132 8	0.233 3	0.105 9
5	0.406 2	16.970 8	0.866 2	0.064 6	0.128 8	0.122 3	0.097 9
6	0.400 6	17.093 4	0.867 2	0.064 5	0.127 5	0.146 2	0.096 2
7	0.405 6	16.985 4	0.869 9	0.064 4	0.130 6	0.109 7	0.098 3
8	0.405 5	16.986 0	0.869 9	0.064 4	0.130 6	0.116 4	0.098 3
9	0.418 7	16.876 8	0.870 4	0.064 4	0.131 2	0.132 0	0.098 7
10	0.410 1	16.888 8	0.872 1	0.064 3	0.133 5		0.100 5

表3 不同分解层数利用 db5 小波基处理的评价指标

Table 3 Evaluation indexes of db5 wavelet basis processing for different decomposition layers

分解层数	真值已知		真值未知		$T$	$S$	$H$
	RMSE	SNR	RMSE	$r$			
2	0.541 1	14.480 5	0.821 9	0.080 8	0.859 3	0.892 5	0.893 3
3	0.456 1	15.965 5	0.883 3	0.050 0	0.236 9	0.267 8	0.219 0
4	0.414 9	16.788 0	0.895 5	0.045 7	0.156 1	0.248 9	0.129 6
5	0.382 5	17.492 8	0.905 7	0.044 4	0.140 3	0.157 4	0.108 7
6	0.375 8	17.648 3	0.909 4	0.044 1	0.139 7	0.119 5	0.106 5
7	0.375 2	17.662 2	0.910 3	0.044 0	0.140 0	0.109 5	0.106 27
8	0.379 5	17.562 9	0.910 7	0.044 0	0.140 1	0.108 7	0.106 3
9	0.379 7	17.558 6	0.910 9	0.044 0	0.140 7	0.107 5	0.106 5
10	0.380 2	17.546 3	0.911 0	0.044 0	0.144 5		0.106 7

到复合评价指标,从而判断最优分解层数,并以真值已知情况下得到的最优分解层数为参考,对比本文方法与现有方法,结果如表4所示。

表4 不同小波基对应的最优分解层数

Table 4 Number of optimal decomposition layers corresponding to different wavelet bases

小波基	最优分解层数			
	真值已知	本文方法	文献[9]	文献[10]
db3	9	9	9	6
db4	6	6	6	7
db5	7	7	6	7
db6	6	6	5	7
db7	6	6	5	6
db8	6	8	6	6
sym3	9	9	9	6
sym4	6	6	6	5
Sym5	6	5	5	7
Sym6	6	6	6	5
Sym7	7	7	6	7
Sym8	6	6	5	6
coif3	6	6	6	5
coif4	6	6	6	6
coif5	5	6	6	5

从表4可以得到,本文方法的正确率为80%,文献[9]和文献[10]所用方法的正确率分别为53%和47%,由此可看出,相比于现有方法,本文所提方法具有更高的准确率及可信度,而且可广泛应用于各种类型的小波基,是一种可靠的小波去噪质量评价方法。

### 3.2 实测惯性数据处理

为验证本文方法对惯性元件随机误差降噪处理的适用性,分别利用加拿大诺瓦泰公司 SPAN-ISA-100C 型高精度光纤惯导和迈普时空研发的 MP-M39 型 MEMS 惯性测量单元进行数据的采集。将2款不同精度惯导设备安装在静态环境中,并预热30 min 后开始工作。为了使2款设备的微小误差能够表现出来,需尽可能采集多的数据,为了达到光纤惯导采样时间,将2款设备同样采集5 h 实验数据<sup>[20]</sup>,记录得到本次实验测试数据样本,采样频率设为200 Hz。

首先,利用Allan方差分析误差信号的特点,Allan方差的最大特点就是能够对各种误差源进行单独辨识分析。2款陀螺仪的误差分析结果如表5和表6所示。

表5 SPAN-ISA-100C 陀螺仪 Allan 方差分析结果

Table 5 Allan variance analysis results of SPAN-ISA-100C gyroscope

坐标轴	角度随机游走/ ( $^{\circ}$ ) $\cdot$ $h^{-\frac{1}{2}}$	零偏不稳定性/ ( $^{\circ}$ ) $\cdot$ $h^{-1}$	角速率随机游走/ ( $^{\circ}$ ) $\cdot$ $h^{-\frac{3}{2}}$
x轴	0.030 4	0.045 5	0.054 2
y轴	0.026 9	0.120 8	0.185 7
z轴	0.025 4	0.043 0	0.057 9

表6 MP-M39 陀螺仪 Allan 方差分析结果

Table 6 Allan variance analysis results of MP-M39 gyroscope

坐标轴	角度随机游走/ ( $^{\circ}$ ) $\cdot$ $h^{-\frac{1}{2}}$	零偏不稳定性/ ( $^{\circ}$ ) $\cdot$ $h^{-1}$	角速率随机游走/ ( $^{\circ}$ ) $\cdot$ $h^{-\frac{3}{2}}$
x轴	0.199 3	9.504 1	28.599 4
y轴	0.136 3	4.198 7	14.920 2
z轴	0.186 8	7.591 2	29.389 7

由表5和表6可知,陀螺仪的主要误差来自于角度随机游走、零偏不稳定性及角速率随机游走,随着陀螺精度的降低,3轴各项误差源也随之变得不稳定,同时误差系数变大。

在小波去噪之前,确保其他参数一致,采用rigsure 阈值选取准则、软阈值处理函数,选取不同的小波基函数,分解层数为2~10层,采用控制变量法,选定 sym6 小波基函数,分别利用本文复合评价指标H及文献[9]的指标T进行最优分解层数的判定。2款惯导设备的3轴陀螺数据所对应的最优分解层数列于表7。其中,x1、y1、z1分别表示SPAN-ISA-100C陀螺仪的3轴数据,x2、y2、z2分别表示MP-M39陀螺仪的3轴数据。

在实验中同时发现,对于这2款惯导设备共6组数据,每组数据选取不同小波基函数得到的最优分解层数是相同的,SPAN-ISA-100C的x轴输出数据最优分解层数为4,其余数据得到的最优分解层数为5。由此,可以将对应最优分解层数设置为表7中的分解层数,其他参数保持不变,根据复合评价指标H判定最优小波基函数,得到对应最优小波基函数如表8所示。

可知,复合评价指标H可以将小波阈值去噪参数中的分解层数与小波基函数的选择统一起来,对SPAN-ISA-100C的x轴输出数据选用sym6小波基

表7 陀螺数据对应的最优分解层数

Table 7 Number of optimal decomposition layers corresponding to gyro data

方法	陀螺数据					
	x1	y1	z1	x2	y2	z2
本文	4	5	5	5	5	5
文献[9]	4	4	4	4	4	4

表8 陀螺数据对应的最佳小波基函数

Table 8 Optimal wavelet basis functions corresponding to gyro data

陀螺数据	x1	y1	z1	x2	y2	z2
小波基	sym6	sym4	db8	db7	db7	db7

函数,4层分解层数;y轴输出数据选用sym4小波基函数,5层分解层数;z轴输出数据选用db8小波基函数,5层分解层数,取得最优的去噪效果。对MP-M39的3轴输出数据皆选用db7小波基函数,5层分解层数,取得最好的去噪效果。另外,对于不同的传感器所选择的最优参数往往不同,且对于同一传感器,由于其所处环境不同,所选用小波去噪参数亦不一致。基于经验所选择的参数缺乏理论指导,并不一定准确可靠。

同时,从表7可以看出,利用2种方法进行小波去噪时,SPAN-ISA-100C的x轴数据对应分解层数是相同的,但其他最优分解层数有所不同。为了直观判定2种方法所得结果的准确性,以SPAN-ISA-100C的y轴输出为例,分别利用2种方法得到降噪信号的时域曲线与频谱曲线,结果如图3和图4所示。

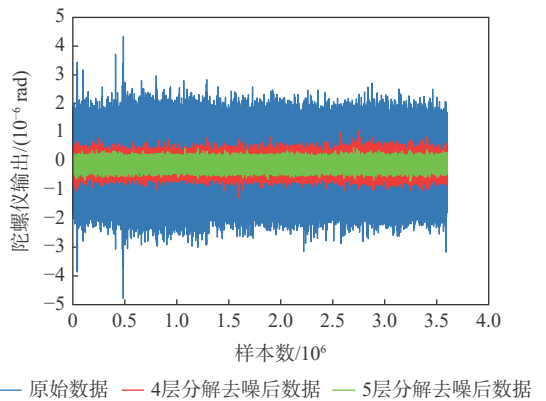


图3 原始信号与降噪信号时域曲线

Fig. 3 Time-domain curves of original signal and denoising signal

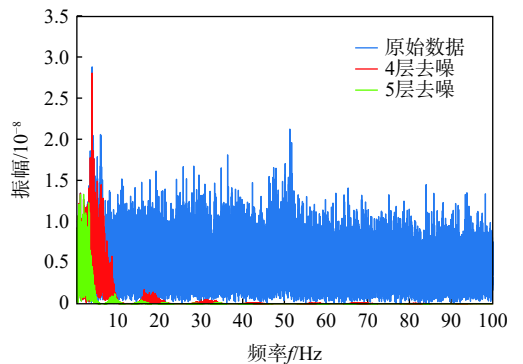


图4 原始信号与降噪信号频谱曲线

Fig. 4 Spectrum curves of original signal and denoising signal

图3中,小波去噪后的信号都保留了原始信号的变化趋势,相比于4层分解,5层分解降噪信号峰值曲线更加光滑,波形图更加平稳,消除噪声更加彻底。图4为信号的单边振幅谱图,去噪主要剔除高频无用细节信息,同时保留了低频有用近似信息,这也与小波去噪原理相符。5层分解信号比4层分解信号更有效地剔除了高频噪声,同时也说明5层分解降噪信号的SNR更高。由此可得,选取5层分解降噪在此状态下效果最优,这也从另一个侧面说明了复合评价指标 $H$ 的有效性。

为了更加清晰地判断小波去噪效果,图5给出了原始信号与分解层数为4和5的去噪信号的功率谱密度估计。可以看出,保留了低频有用信息,剔除了高频无用信息,同时,对中间的过渡信息适当进行了保留,相比于4层分解降噪信号,5层分解降噪信号剔除了较为微弱的噪声信号,也证明了5层分解层数是最优分解层数。由于篇幅所限,本文不再分析其他输出数据去噪效果对比结果。

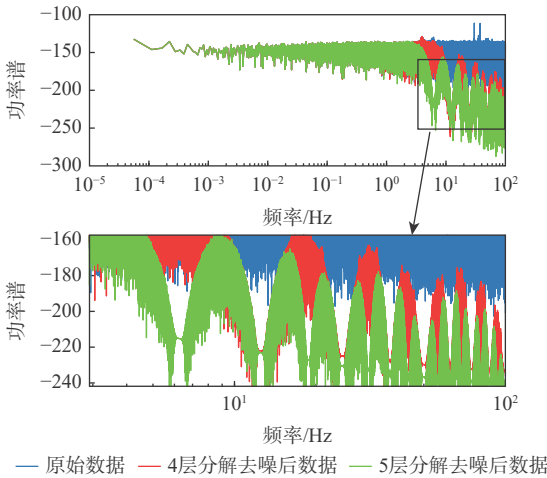


图5 原始信号与降噪信号功率谱密度图

Fig. 5 Power spectral density diagram of original signal and denoising signal

## 4 结论

针对传统方法中经验模型缺乏理论依据的不足,本文提出了一种基于组合赋权法的小波去噪质量评价方法,得出以下结论:

1) 通过实验发现,在真值未知情况下,单一评价指标无法指导小波阈值去噪过程中相关参数的选择,需构建更加全面的质量评价指标。

2) 选取RMSE、平滑度2个指标作为评价依据,对其进行初始化,通过引入信息熵权与变异系数2种客观赋权法进行组合赋权,将归一化值与权重线性组合形成了一种新的小波去噪质量评价指标,并得出了相关计算公式。

3) 通过对真值已知的模拟信号和真值未知的实测信号进行处理,证明该方法能够指导小波阈值去噪过程中相关参数的选择,具有更高的准确性与广泛适用性。

## 参考文献 (References)

- [1] SHI Y B, ZHANG J J, JIAO J J, et al. Calibration analysis of high-G MEMS accelerometer sensor based on wavelet and wavelet packet denoising[J]. *Sensors*, 2021, 21(4): 1231.
- [2] 吴保峰, 夏林元, 吴东金, 等. 基于改进小波阈值的MEMS陀螺去噪算法[J]. *电子器件*, 2020, 43(5): 1104-1107.  
WU B F, XIA L Y, WU D J, et al. Denoising algorithm of MEMS gyroscope based on improved threshold function[J]. *Chinese Journal of Electron Devices*, 2020, 43(5): 1104-1107(in Chinese).
- [3] JANG Y I, SIM J Y, YANG J R, et al. The optimal selection of mother wavelet function and decomposition level for denoising of DCG signal[J]. *Sensors*, 2021, 21(5): 1851.
- [4] 李文华, 汪立新, 沈强, 等. 基于EMD的MEMS陀螺仪随机漂移分析方法[J]. *北京航空航天大学学报*, 2021, 47(9): 1927-1932.  
LI W H, WANG L X, SHEN Q, et al. Random drift analysis method of MEMS gyroscope based on EMD[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2021, 47(9): 1927-1932(in Chinese).
- [5] 景冰洁, 韩跃平, 张鹏, 等. 改进提升小波阈值法在MEMS陀螺仪误差分析中的应用[J]. *仪表技术与传感器*, 2020(12): 95-100.  
JING B J, HAN Y P, ZHANG P, et al. Application of improved lifting wavelet threshold method in error analysis of MEMS gyroscope[J]. *Instrument Technique and Sensor*, 2020(12): 95-100(in Chinese).
- [6] 刘晓光, 胡静涛, 高雷, 等. 基于改进小波阈值的微机械陀螺去噪方法[J]. *中国惯性技术学报*, 2014, 22(2): 233-236.  
LIU X G, HU J T, GAO L, et al. Micro mechanical gyro denoising method based on improved wavelet threshold[J]. *Journal of Chinese Inertial Technology*, 2014, 22(2): 233-236(in Chinese).
- [7] 邵旋, 康兴无, 王旭平, 等. 基于多元信息融合的小波降噪质量综合评估指标[J]. *兵器装备工程学报*, 2020, 41(12): 155-160.  
SHAO X, KANG X W, WANG X P, et al. Comprehensive evaluation index of wavelet denoising quality based on multi-information fusion[J]. *Journal of Ordnance Equipment Engineering*, 2020, 41(12): 155-160(in Chinese).
- [8] 陶珂, 朱建军. 多指标融合的小波去噪最佳分解尺度选择方法[J]. *测绘学报*, 2012, 41(5): 749-755.  
TAO K, ZHU J J. A hybrid indicator for determining the best decomposition scale of wavelet denoising[J]. *Acta Geodaetica et Cartographica Sinica*, 2012, 41(5): 749-755(in Chinese).
- [9] 朱建军, 章浙涛, 匡翠林, 等. 一种可靠的小波去噪质量评价指标[J]. *武汉大学学报(信息科学版)*, 2015, 40(5): 688-694.  
ZHU J J, ZHANG Z T, KUANG C L, et al. A reliable evaluation indicator of wavelet de-noising[J]. *Geomatics and Information Science of Wuhan University*, 2015, 40(5): 688-694(in Chinese).
- [10] 王旭, 王昶. 一种基于熵权法的小波去噪复合评价指标[J]. *大地测量与地球动力学*, 2018, 38(7): 698-702.  
WANG X, WANG C. A kind of wavelet de-noising composite evaluation index based on entropy method[J]. *Journal of Geodesy*

- and *Geodynamics*, 2018, 38(7): 698-702(in Chinese).
- [11] 冯明,周程瑜,张坤,等. 回转误差测试中系统噪声分离技术[J]. *北京航空航天大学学报*, 2020, 46(4): 666-673.  
FENG M, ZHOU C Y, ZHANG K, et al. Separation technology of system noise in error motion test[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2020, 46(4): 666-673(in Chinese).
- [12] 孔玲军. MATLAB小波分析超级学习手册[M]. 北京: 人民邮电出版社, 2014.  
KONG L J. MATLAB super Learning handbook of wavelet analysis[M]. Beijing: Post&Telecom Press, 2014(in Chinese).
- [13] 赵宣懿,孔雪博,熊智,等. 基于低成本MEMS陀螺的小波阈值去噪应用研究[J]. *传感器与微系统*, 2017, 36(12): 54-56.  
ZHAO X Y, KONG X B, XIONG Z, et al. Research of application of wavelet threshold de-noising based on low-cost MEMS gyro[J]. *Transducer and Microsystem Technologies*, 2017, 36(12): 54-56(in Chinese).
- [14] SOLTANI A A, SHAHRTASH S M. Decision tree-based method for optimum decomposition level determination in wavelet transform for noise reduction of partial discharge signals[J]. *IET Science, Measurement & Technology*, 2020, 14(1): 9-16.
- [15] LI Y N, LI Z H. Application of a novel wavelet shrinkage scheme to partial discharge signal denoising of large generators[J]. *Applied Sciences*, 2020, 10(6): 2162.
- [16] BALDAZZI G, SOLINAS G, DEL VALLE J, et al. Systematic analysis of wavelet denoising methods for neural signal processing[J]. *Journal of Neural Engineering*, 2020, 17(6): 11-27.
- [17] ROUIS M, OUAFI A, SBAA S. Optimal level and order detection in wavelet decomposition for PCG signal denoising[J]. *Biomedizinische Technik Biomedical Engineering*, 2019, 64(2): 163-176.
- [18] 谈渊,甘学辉,张东剑,等. 基于小波去噪的激光多普勒振动信号处理[J]. *激光技术*, 2022, 46(1): 129-133.  
TAN Y, GAN X H, ZHANG D J, et al. Laser Doppler vibration signal processing based on wavelet denoising[J]. *Laser Technology*, 2022, 46(1): 129-133(in Chinese).
- [19] 王永弟,许承权,范千. 熵权、变异系数及模糊多准则决策在测量平差中的综合应用[J]. *工程勘察*, 2012, 40(9): 58-61.  
WANG Y D, XU C Q, FAN Q. Integrated application of entropy theory, variation coefficient and fuzzy multi-criteria decision making in surveying adjustment[J]. *Geotechnical Investigation & Surveying*, 2012, 40(9): 58-61(in Chinese).
- [20] 严恭敏,李四海,秦永元. 惯性仪器测试与数据分析[M]. 北京: 国防工业出版社, 2012: 143-146.  
YAN G M, LI S H, QIN Y Y. Inertial instrument testing and data analysis[M]. Beijing: National Defense Industry Press, 2012: 143-146 (in Chinese).

## A quality evaluation method for wavelet denoising based on combinatorial weighting method

LI Jinfei, ZHAO Dongqing\*, WANG Dongmin, CAI Congcong, JIA Xiaoxue, ZHANG Letian

(Institute of Geospatial Information, Information Engineering University, Zhengzhou 450001, China)

**Abstract:** Addressing such a problem with the traditional indicator system for quality evaluation as an insufficient theoretical basis for wavelet threshold denoising, a combination weighting approach-based method for evaluation of wavelet denoising quality is proposed with the expectation of effectively evaluating the selection of wavelet denoising parameters. Through analysis of characteristics of individual indicators such as root-mean-square error (RMSE), signal-noise ratio (SNR) and smoothness with the truth-value unknown, RMSE and smoothness are selected as wavelet denoising indicators. They are first normalized, then processed with information entropy and coefficient of variation for combination weighting, and, in the end, linearly combined with the corresponding weights to produce a new indicator, i.e., the composite index. A smaller composite index indicates better denoising effect and better parameters selected. According to a simulated experiment, the index outperforms the conventional approach in terms of accuracy given the truth-value and is applicable to various decomposition levels and wavelet base functions. According to experimental data, this method achieves smoother wavelet denoising peak regions, steadier waveforms, and a better denoising effect.

**Keywords:** wavelet threshold denoising; quality evaluation index; root mean square error; smoothness; combination weighting; inertial element random error

Received: 2021-06-07; Accepted: 2021-07-18; Published Online: 2021-09-24 17:55

URL: [kns.cnki.net/kcms/detail/11.2625.V.20210924.1200.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210924.1200.001.html)

Foundation item: National Natural Science Foundation of China (41774037)

\* Corresponding author. E-mail: [dongqing.zhao@hotmail.com](mailto:dongqing.zhao@hotmail.com)

# 柔性涡流阵列传感器孔边裂纹监测技术

樊祥洪, 缙百勇\*, 陈涛, 何宇廷, 崔荣洪, 喻健

(空军工程大学 航空工程学院, 西安 710038)

**摘 要:** 针对这一问题, 提出了一种带双面补强的柔性涡流阵列传感器并用于孔边裂纹监测。通过 COMSOL 有限元软件建立传感器和被测试件的有限元模型, 分析提高距离、垫片磁导率变化和裂纹扩展对传感器输出信号的影响; 制备带补强和不带补强的传感器, 并开展挤压实验和在线疲劳裂纹监测实验; 根据实验结果和仿真结果之间的差异进行误差分析。结果显示: 随着提高距离和垫片磁导率的增加, 传感器输出感应电压逐渐增大; 传感器裂纹监测灵敏度随着提高距离的增加而逐渐减小; 带补强的传感器可以在螺栓拧紧扭矩为 63 N·m 的条件下工作, 而不带补强的传感器在拧紧力矩为 50 N·m 时完全失效; 通过在线疲劳裂纹监测实验验证了带补强的传感器对裂纹具有定量监测能力, 裂纹监测精度与激励线圈之间的间距一致, 可达 1 mm; 实验结果与仿真结果之间的差异主要由提高距离引入。

**关键词:** 柔性涡流阵列传感器; 孔边裂纹监测; 灵敏度; 双面补强; 提高距离

**中图分类号:** V215.6; TP212.1

**文献标志码:** A **文章编号:** 1001-5965(2023)03-0726-09

金属结构作为主要承力结构, 在飞机、高铁等重型设备中发挥着中流砥柱的作用。然而金属结构在复杂恶劣的环境下服役, 极易引入损伤(如疲劳裂纹、腐蚀等)。关键部位若存在损伤未及时发现极易引发灾难性的后果。为确保结构的安全, 需要定期对结构进行无损检测并对损伤部位进行修理。提高维修的经济性, 则需要采用视情维修策略。作为视情维修技术关键核心在于对结构进行结构健康监测。

结构健康监测技术就是通过结构上布置先进传感器来监测表征结构健康状态的参数(如温度、应变、腐蚀、裂纹和振动等), 通过数据采集和处理分析结构的健康状态, 给出结构的维修计划<sup>[1-2]</sup>。由此, 可以看出传感器技术起着重要的作用。传感器性能的优劣决定着整个结构健康监测系统功能的好坏。目前, 结构健康监测技术采用的传感器有智能涂层传感器<sup>[3]</sup>、比较真空度传感器<sup>[4]</sup>、光纤传感器<sup>[5]</sup>、应变传感器<sup>[6]</sup>、超声导波传感器<sup>[7]</sup>

和涡流传感器<sup>[8]</sup>等。各传感器有着自己的优缺点和适用范围。智能涂层传感器可以监测金属和非金属表面裂纹的萌生和扩展, 但是涂层工艺要求较高, 如物理气相沉积(physical vapor deposition, PVD)薄膜传感器, 可以定量监测裂纹, 裂纹监测精度可小于 1 mm, 但是该传感器用于带螺栓孔板具有较大的挑战<sup>[9]</sup>。比较真空度传感器可以定量监测表面裂纹, 裂纹监测精度可达 1 mm, 但是对于工作在高压应力状态下具有较大的挑战性。光纤传感器具有很好的形状适应性, 可以监测复杂结构, 但是该传感器受环境干扰较大, 容易引入“虚警”和“误警”。应变传感器可以根据应变的变化来判断裂纹的产生, 该传感器价格便宜、安装方便, 但是对于微小裂纹的萌生和扩展不灵敏。超声导波传感器具有监测范围大、监测效率高的优良特点, 在复合材料监测上应用广泛, 但是该传感器对于复杂结构, 尤其是多孔结构具有较大的挑战性。

涡流传感器作为一种常规的无损检测传感器

收稿日期: 2021-06-07; 录用日期: 2021-09-26; 网络出版时间: 2021-11-01 13:24

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20211101.0924.001.html](https://kns.cnki.net/kcms/detail/11.2625.V.20211101.0924.001.html)

基金项目: 国家杰出青年科学基金(52007197)

\*通信作者. E-mail: [goubaiyong789@126.com](mailto:goubaiyong789@126.com)

**引用格式:** 樊祥洪, 缙百勇, 陈涛, 等. 柔性涡流阵列传感器孔边裂纹监测技术[J]. 北京航空航天大学学报, 2023, 49(3): 726-734.

FAN X H., GOU B Y., CHEN T., et al. Hole edge crack monitoring technology of flexible eddy current array sensor[J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49(3): 726-734 (in Chinese).

具有众多优良特点,既可以作为无损检测传感器使用,又可以用于结构健康监测。其中,柔性涡流阵列传感器具有质量轻、柔性可弯曲的优良特点,在结构健康监测领域具有广泛的应用前景。不少研究者对其开展了深入的研究,如美国 JENTEK Sensor 公司的 Goldfine 等<sup>[10]</sup>提出了一种蜿蜒状绕组磁力计(meandering winding magnetometer, MWM)传感器,开展了大量的研究, Schlicker<sup>[11]</sup>提出了含有多个 MWM 阵列的新型传感器,研究了基于该传感器的结构扫描成像技术。

螺栓连接结构作为结构的重要承力结构,孔边作为应力集中部位,极易产生疲劳裂纹。“斯坦福大学的”删除。Rakow 和 Chang<sup>[12]</sup>开发了一种 SHM 螺栓,其由一个涡流传感器薄膜和螺栓集成而成,主要用于螺栓孔壁的裂纹监测,虽然 SHM 螺栓能承受较大的载荷,但是该型传感器的灵敏度不高,容易受到环境因素的影响。何宇廷等<sup>[13-15]</sup>开展了大量柔性涡流阵列传感器的研究,针对孔边裂纹监测问题,提出了一种花萼状的柔性涡流阵列传感器并建立传感器解析模型和半解析模型,文献<sup>[16]</sup>提出了一种用于孔边裂纹监测的同向激励布局柔性涡流阵列传感器,文献<sup>[17]</sup>提出了提高裂纹监测灵敏度的方法,但是未考虑对带有螺栓的孔边进行监测。

综上,本文从提高柔性涡流阵列传感器抗压力性能的角度,提出了一种采用双面补强并含参考通道的柔性涡流阵列传感器。首先,通过 COMSOL 有限元软件建立传感器与被测结构的有限元模型,分析提高距离、垫片磁导率对传感器输出信号的影响规律及传感器对裂纹识别的灵敏度。然后,根据传感器仿真模型尺寸制备带补强和不带补强 2 种传感器,开展挤压实验和在线疲劳裂纹监测实验,验证该传感器具有螺栓孔边裂纹定量监测能力。

## 1 裂纹监测原理

### 1.1 带双面补强的柔性涡流阵列传感器

常规柔性涡流阵列传感器由激励线圈和感应线圈阵列组成,采用柔性印刷技术将激励线圈和感应线圈印制在柔性基底上,如图 1 所示。

本文在此基础上,通过在前端监测区域上下添加补强片,提高传感器的抗压能力,如图 2 所示。

根据连接结构孔边的受力特性可知,与载荷方向垂直的左右两侧区域应力较为集中,该部位是产生疲劳裂纹的危险部位。因此,在对疲劳裂纹监测时,只需要监测该部位。本文设计的双面补强的柔性涡流阵列传感器共有 7 个感应通道,左右两侧各 3 个感应通道,上面有 1 个感应通道,该通道是参考通道,用于消除环境对传感器输出信号的影响,具体示意图如图 3 所示。激励线圈按从内向外的顺序分别定义为激励线圈 1、激励线圈 2、激励线圈 3 和激励线圈 4。

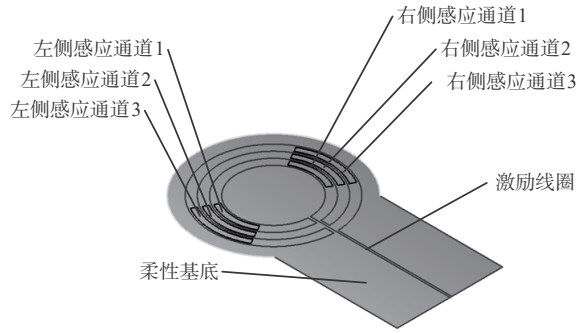


图 1 柔性涡流阵列传感器示意图

Fig. 1 Schematic diagram of flexible eddy current array sensor

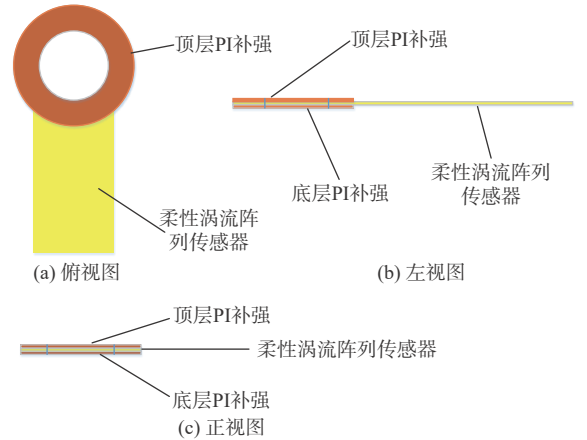


图 2 带双面补强的柔性涡流阵列传感器

Fig. 2 Flexible eddy current array sensor with double-sided reinforcement

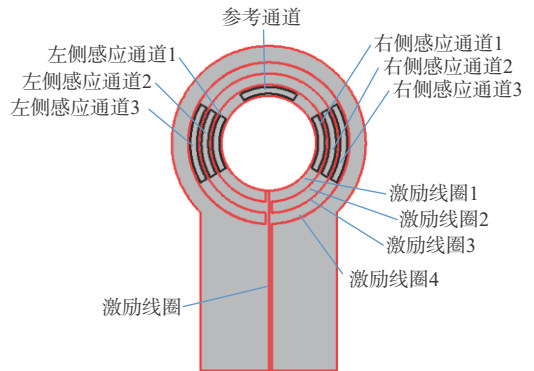


图 3 带参考通道的柔性涡流阵列传感器

Fig. 3 Flexible eddy current array sensor with reference channel

### 1.2 传感器工作原理

柔性涡流阵列传感器的工作原理基于电磁感应定律。通过给激励线圈施加交流电时,激励线圈会在空间中产生交变磁场。当传感器靠近金属导体时,根据电磁感应定律,在导体上会产生与激励电流方向相反的涡流。本文采用的激励电流频率较高,达到 1 MHz。可以近似认为涡流分布在被测试件表面,且涡流流动方向与激励电流方向相反。

当孔边有裂纹产生时,由于介质的不连续性,位于激励线圈下方的涡流会沿着裂纹尖端流动,进而形成一个扰动磁场,而感应线圈感应的是激励磁场、涡流磁场和扰动磁场三者之间的矢量和,当裂纹产生时,传感器的感应通道信号就会发生变化。因此,通过信号的变化来判断裂纹的萌生及扩展情况。根据图4可以看出,裂纹产生在左侧区域,同时,裂纹扩展穿过了左侧感应通道1,到达了左侧感应通道2。此时,左侧感应通道1和左侧感应通道2的信号会发生变化。

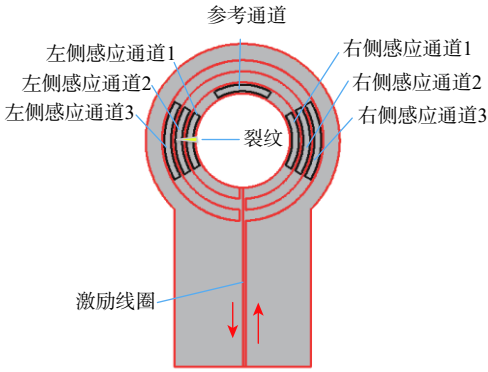


图4 裂纹扩展示意图

Fig. 4 Schematic diagram of crack propagation

## 2 仿真分析

带双面补强的柔性涡流阵列传感器用于螺栓孔边进行裂纹监测时,传感器的提高距离和螺栓的电磁特性(电导率、磁导率)都会影响传感器的输出信号。本节建立传感器的有限元模型,分析提高距离、垫片磁导率和垫片电导率对传感器输出信号的影响。

### 2.1 二维有限元模型

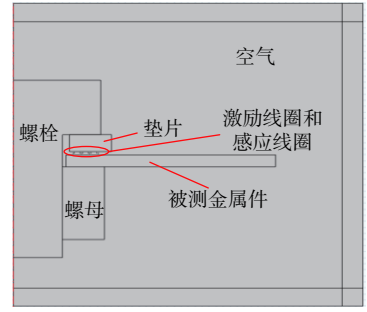
本文采用COMSOL软件有限元,通过二维轴对称建立传感器与被测试件之间的有限元模型,如图5所示。

该模型主要由激励线圈、感应线圈、被测金属件、螺栓、螺母、垫片和空气组成。螺栓、螺母及垫片采用的铁磁性材料,相对磁导率设置为2000,激励线圈与被测金属件之间的间距定义为提高距离 lift-off,考虑到传感器双面PI补强片的厚度一样,因此,设感应线圈与垫片表面之间的间距为 lift-off。主要仿真参数如表1所示。

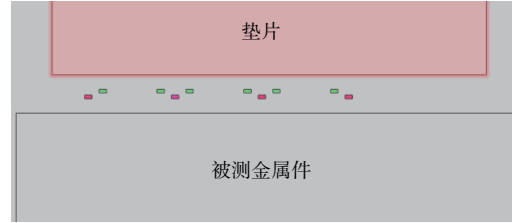
### 2.2 仿真结果与分析

在COMSOL有限元软件的后处理模块中作出传感器附近区域涡流分布图,如图6所示。

根据图6可以看出,由于激励频率较高,金属导体表面产生的涡流密度受趋肤效应的影响,涡流集中在激励线圈下方。将传感器的提高距离作为扫描参数,作出感应通道输出电压随提高距离变化的趋势图,如图7所示。



(a) 仿真模型示意图



(b) 局部放大图

图5 传感器仿真模型

Fig. 5 Sensor simulation model

表1 仿真参数

Table 1 Simulation parameters

参数	数值
$f$ 激励频率/MHz	1
$\sigma_{AL}$ 2024铝合金电导率/( $S \cdot m^{-1}$ )	$1.74 \times 10^7$
$\mu_{AL}$ 2024铝合金相对磁导率	1
$\sigma_b$ 螺栓、垫片、螺母电导率/( $S \cdot m^{-1}$ )	$8.41 \times 10^6$
$\mu_b$ 螺栓、垫片、螺母相对磁导率	2000
$H_{AL}$ 被测金属件厚度/mm	2
$C_c$ 激励线圈、感应线圈宽度/mm	0.1
$C_t$ 激励线圈、感应线圈厚度/mm	0.03
$J_t$ 柔性基底厚度/mm	0.03
$B_t$ 补强片厚度/mm	0.1
$D_c$ 激励线圈与感应线圈之间的间距/mm	0.2
$R$ 圆孔半径/mm	6
$R_1, R_2, R_3, R_4$ 激励线圈1、2、3、4半径/mm	7、8、9、10
lift-off提高距离/mm	0.1

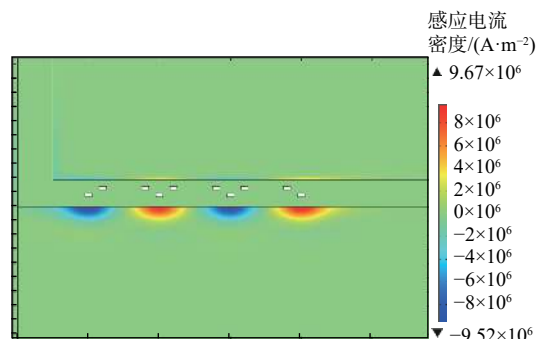


图6 试验件剖面涡流分布图

Fig. 6 Eddy-current profile of test piece

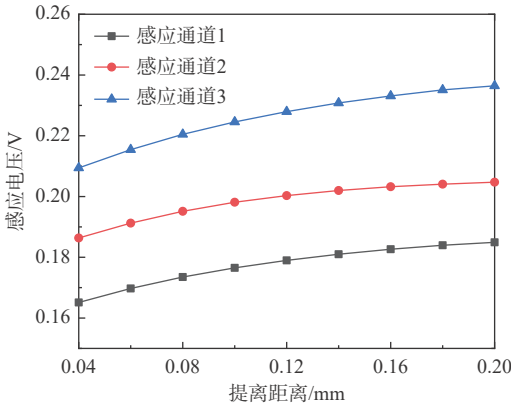


图 7 感应电压随提高距离变化趋势

Fig. 7 Trend of induced voltage with lift-off

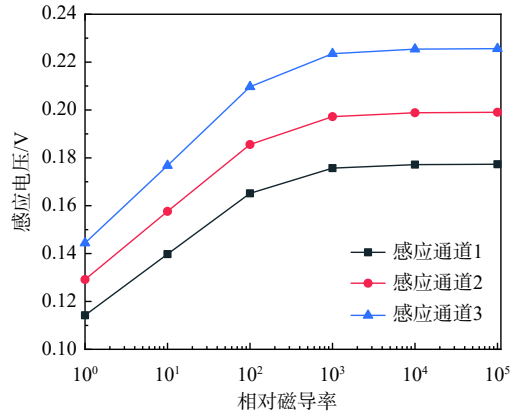


图 8 感应电压随相对磁导率的变化趋势

Fig. 8 Trend of induced voltage with relative permeability

根据图 7 可知, 3 个感应通道中, 感应通道 3 的感应电压最大, 其次是感应通道 2, 最小的是感应通道 1。这主要是由于感应通道 3 位于传感器最外通道, 所围成的面积最大, 穿过的磁通量最大。因此, 在相同频率变化条件下, 感应通道 3 的感应电压最大。3 个通道随着提高距离的增加, 感应电压逐渐增大。以螺栓、垫片和螺母的相对磁导率为扫描参数, 作出传感器感应通道输出电压随相对磁导率的变化趋势图, 如图 8 所示。

根据图 8 可知, 随着相对磁导率的增加, 感应电压逐渐增加, 当相对磁导率增加到 1 000 时, 感应电压近似保持不变。

### 2.3 裂纹扩展仿真分析

通过建立传感器的二维仿真模型, 分析感应电压和提高距离对传感器感应通道输出电压的影响。本节建立传感器与裂纹扩展的耦合模型, 分析裂纹扩展过程中对传感器输出信号的影响, 具体仿真模型如图 9 所示。

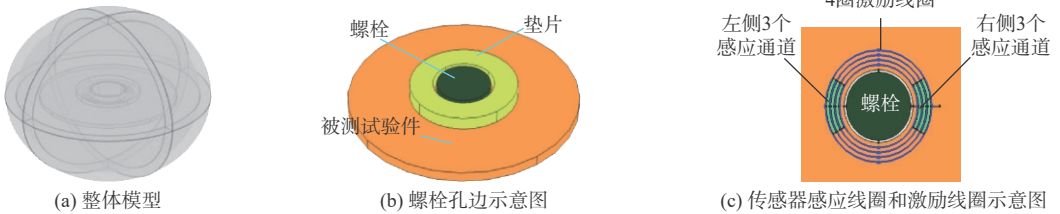


图 9 传感器三维仿真模型

Fig. 9 Sensor simulation model

考虑到传感器的激励频率较高, 涡流的趋肤深度较小。为提高有限元模型计算效率, 将螺栓简化为圆柱体, 在激励线圈和感应单元上方加上垫片。因此, 该仿真模型主要由激励线圈、感应线圈、被测试件、简化后的螺栓及垫片组成。通过在孔边预制一条宽度为 0.2 mm 的裂纹, 传感器的提高距离设置为 0.1 mm。将裂纹的长度作为扫描参数, 通过控制裂纹的长度来模拟裂纹的扩展过程。

### 2.4 传感器裂纹识别灵敏度

通过 COMSOL 有限元软件的后处理, 计算穿过感应通道磁通量对时间的导数, 即感应电压为

$$V_i = - \oint_{S_i} \frac{\partial \mathbf{B}}{\partial t} \times d\mathbf{S} \quad i = 1, 2, 3 \quad (1)$$

式中:  $\mathbf{B}$  为感应通道中的磁感应强度;  $S_i$  为第  $i$  ( $i = 1, 2, 3$ ) 个感应通道围起来的面积。

本文定义传感器的跨阻抗为感应电压与激励电流之比:

$$T_i = \frac{V_i}{I_e} = A e^{j\theta} \quad (2)$$

式中:  $T_i$  为跨阻抗;  $A$  为跨阻抗幅值;  $\theta$  为感应电压与激励电流之间的相位差;  $I_e$  为激励电流。

传感器对裂纹识别的灵敏度  $S_c$  定义为跨阻抗幅值的变化率, 如下:

$$S_c = \frac{|A - A_0|}{A_0} \times 100\% \quad (3)$$

式中:  $A_0$  为传感器在初始时刻裂纹未产生时的跨阻抗幅值。

当裂纹向右扩展依次传感器 3 个感应通道时, 3 个感应通道对裂纹识别的灵敏度变化趋势如图 10 所示。

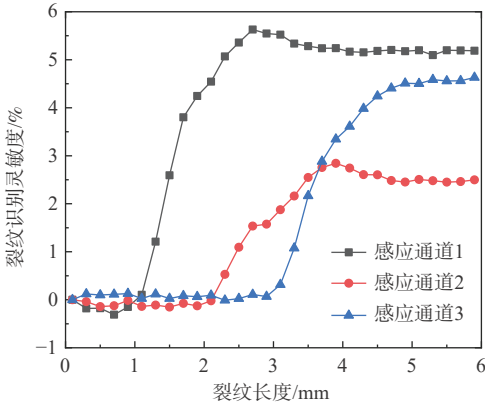


图 10 裂纹识别灵敏度变化趋势图

Fig. 10 Variation trend of crack identification sensitivity

根据图 10 可知,当裂纹未扩展至激励通道 1 时,传感器 3 个感应通道对裂纹识别的灵敏度保持不变,当裂纹扩展穿过感应通道 1 时(裂纹长度为 1 mm),其灵敏度开始呈现上升趋势,此时,感应通道 2 和感应通道 3 对裂纹识别的灵敏度没有开始发生变化。当裂纹继续扩展穿过激励线圈 2 时(裂纹长度为 2 mm),感应通道 2 的灵敏度开始发生变化,此时感应通道 3 的灵敏度未发生变化,当裂纹继续扩展穿过激励线圈 3 时,感应通道 3 的灵

敏度开始发生变化。同时,感应通道 1 对裂纹识别的灵敏度最大,灵敏度为 5.6%;其次是感应通道 3,灵敏度为 4.6%;最小的是感应通道 2,灵敏度为 2.9%。且在裂纹扩展过程中,传感器各个感应通道之间相互独立,未出现一个感应通道的变化引起另一个感应通道裂纹识别灵敏度的变化。因此,可以根据各个感应通道灵敏度开始变化的先后顺序对裂纹的扩展情况进行定量监测。同时,裂纹识别的精度等于传感器激励线圈之间的间距,即等于 1 mm。

通过 COMSOL 有限元软件的后处理模块作出裂纹扩展过程中,被测金属件表面涡流的流动状态,如图 11 所示,裂纹长度用  $a$  表示。

根据图 11 可知,在激励线圈下方的被测试验件会产生与激励电流方向相反的涡流,激励线圈 1 和激励线圈 3 下方的涡流沿着逆时针方向流动,激励线圈 2 和激励线圈 4 下方的涡流沿着顺时针方向流动。当裂纹未产生时,涡流集中在激励线圈下方区域流动,当裂纹扩展至激励线圈 1 时,如图 11(b)所示,激励线圈左侧的涡流发生扰动,而激励线圈 1 和激励线圈 2 之间的区域没有产生扰动涡流。随着裂纹继续扩展将有裂纹产生,因此,感应通道 1 对裂纹识别的灵敏度会发生变化,而激励

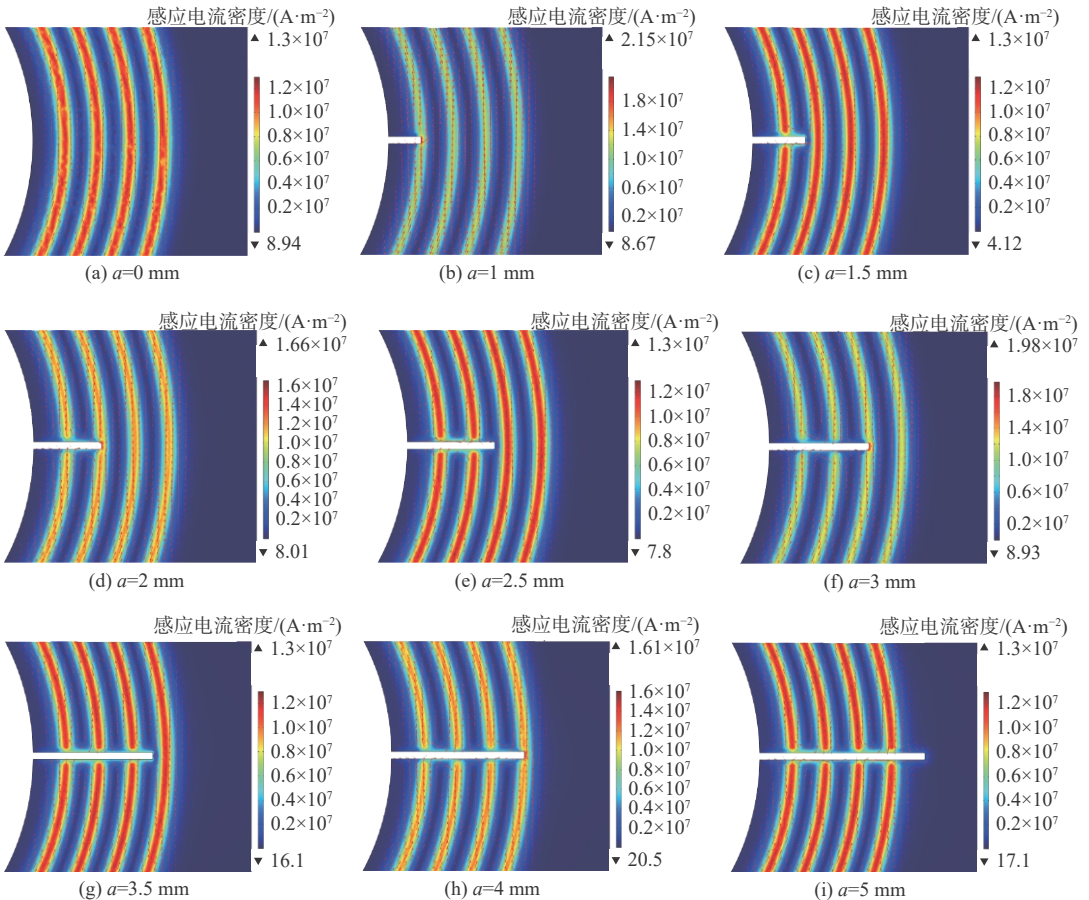


图 11 裂纹扩展对涡流的扰动作用

Fig. 11 Effects of crack growth on eddy currents

线圈 2 和激励线圈 3 之间、激励线圈 3 和激励线圈 4 之间没有扰动涡流产生, 因此, 感应通道 2 和感应通道 3 的灵敏度不会发生变化。当裂纹扩展依次穿过了激励线圈 2 和激励线圈 3, 感应通道 2 和感应通道 3 之间的灵敏度依次发生变化。根据图 11(i) 可以看出, 激励线圈 1 与激励线圈 2、激励线圈 3 和激励线圈 4 之间的扰动涡流明显大于激励线圈 2 与激励线圈 3 之间的涡流。因此, 感应通道 1 和感应通道 3 对裂纹识别的灵敏度大于感应通道 2 之间的灵敏度。

根据仿真结果可知, 将传感器置于垫片下方可以定量监测裂纹的扩展, 且裂纹监测精度与激励线圈之间的间距一致。

### 3 实验验证

为验证本文传感器具有孔边裂纹监测能力, 依据仿真尺寸制备了带双面补强的柔性涡流阵列传感器, 通过 MTS 疲劳试验机加载, 开展在线疲劳裂纹监测实验, 加载应力为 180 MPa, 频率为 20 Hz, 螺栓的直径为 12 mm。根据《航空制造工程手册》<sup>[18]</sup>, 该直径的螺栓施加的最大拧紧扭矩值为 63 N·m。为了验证该传感器在结构的极限状态下工作情况, 先开展了预紧力大小对带补强传感器和不带补强传感器耐挤压对比实验。施加范围为 20~63 N·m, 步长为 5 N·m, 实验现场如图 12 所示, 实验结果如表 2 所示。



(a) 带补强 (b) 不带补强

图 12 传感器挤压实验

Fig. 12 Sensor extrusion experiment

表 2 挤压实验结果

Table 2 Results of extrusion experiment

拧紧扭矩/(N·m)	带补强的传感器	不带补强的传感器
20	正常工作	正常工作
25	正常工作	正常工作
30	正常工作	感应通道5坏
35	正常工作	感应通道3、4坏
40	正常工作	感应通道2坏
45	正常工作	感应通道6坏
50	正常工作	感应通道1、7坏
55	正常工作	
60	正常工作	
63	正常工作	

根据实验结果可知, 随着螺栓拧紧扭矩增大, 不带补强的传感器各个感应通道发生破坏(线圈断裂), 而采用补强的传感器在 63 N·m 的螺栓拧紧扭矩下依然可以正常工作。因此, 在对螺栓孔孔边裂纹进行监测时, 需要采用带补强的柔性涡流阵列传感器。

#### 3.1 实验系统

该系统由带双面补强的柔性涡流阵列传感器、信号源、功放模块、运放模块及信号采集与处理系统组成, 具体连接方式如图 13 所示, 实验现场如图 14 所示。为考察传感器在极限状态下是否正常工作, 螺栓的拧紧扭矩施加为 63 N·m。

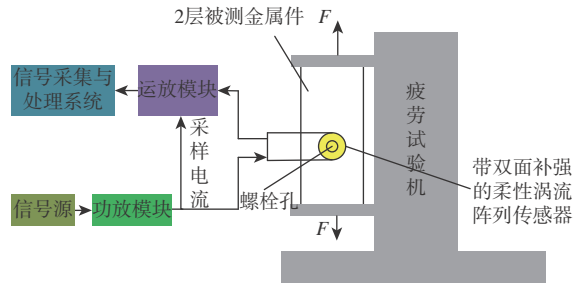


图 13 疲劳裂纹在线监测系统

Fig. 13 On-line fatigue crack monitoring system

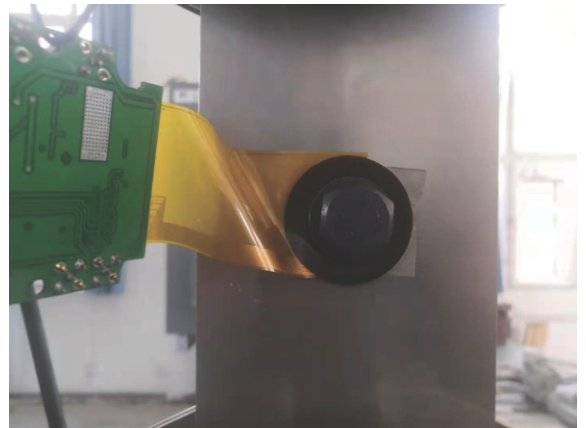


图 14 在线疲劳裂纹监测实验现场

Fig. 14 Experiment site of online crack monitoring

#### 3.2 实验结果

当裂纹依次穿过各感应通道时, 传感器的感应通道信号发生变化。作出传感器各感应通道随裂纹扩展的灵敏度变化趋势图, 如图 15 所示。

根据图 15 可知, 当裂纹扩展至感应通道 1 时, 感应通道的信号开始发生明显变化, 感应通道 2 和感应通道 3 的信号未发生变化。随着裂纹扩展至感应通道 2 时, 感应通道 2 的信号开始变化, 此时感应通道 1 信号变化趋势变缓, 感应通道 3 信号未发生变化。当裂纹扩展至感应通道 3 时, 感应通道 3 的信号开始发生变化。可以看出, 感应通道 1 的裂纹识别灵敏度最高, 约为 2.8%, 其次感应通道

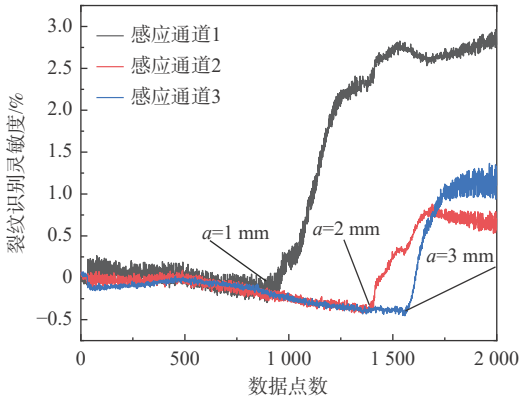


图 15 裂纹监测结果

Fig. 15 Crack monitoring results

3的裂纹识别灵敏度为1.2%，最小为感应通道2，约为0.8%。同时，整个灵敏度变化曲线有下降的趋势，这是由于在加载过程中，螺栓孔边受摩擦的作用，温度有所升高，造成被测结构电导率和垫片相对磁导率变化，因此传感器的输出信号会发生微弱变化，该传感器最大裂纹识别灵敏度为2.8%，而温度对其影响为0.4%，因此该部分的影响不容忽视。

本文针对这一问题设计了一个参考通道，该参考通道位于传感器的正上方，如图3所示。由于该通道位于非应力集中区域，该处一般不会产生疲劳裂纹，该参考通道可以感应环境因素对其影响。因此，为消除环境因素对各感应通道输出信号的影响，可以采用各感应通道灵敏度变化曲线与参考参考通道的灵敏度变化曲线相减求得。由此，可以得到经参考通道修正后的裂纹识别灵敏度变化，如图16所示。

根据图16可知，通过修正后的裂纹监测结果不影响传感器对裂纹扩展的监测性能。感应通道1、感应通道2和感应通道3裂纹识别灵敏度分别为3.4%、1.5%和1.8%。各感应通道间的灵敏度大小关系相比于未修正的裂纹监测结果偏大。

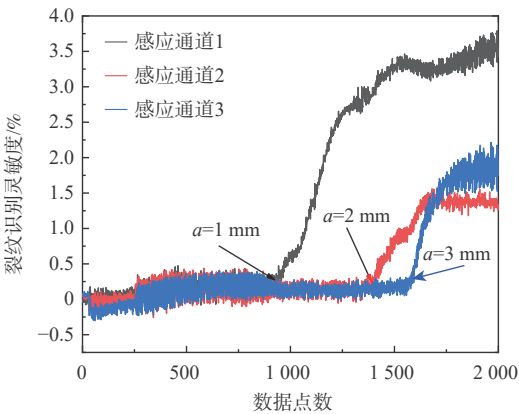


图 16 通过参考通道修正的裂纹监测结果

Fig. 16 Crack monitoring results modified by reference channel

### 3.3 实验结果分析

通过开展在线疲劳裂纹监测实验，验证了该传感器可以通过安装在垫片下方实现对孔边裂纹的扩展进行实时监测。在监测过程中，传感器未发生损坏，验证了带双面补强的柔性涡流阵列传感器具有抗压能力，且能够监测裂纹的萌生和扩展。当裂纹刚好扩展到激励线圈1时，感应通道1的输出信号开始发生变化，当裂纹扩展至激励线圈2时，感应通道2的输出信号发生变化，随着裂纹继续扩展至激励线圈3时，感应通道3的输出信号开始发生变化。因此，可以根据传感器各感应通道输出信号开始变化的点实现对裂纹扩展的定量监测，且裂纹监测精度与传感器激励线圈之间的间距一致，为1 mm。根据二维仿真模型可知，随着提高距离的增加，传感器感应线圈的感应电压增加。在二维仿真模型后处理模块中，计算激励线圈下方区域涡流的大小随提高距离变化趋势，区域划分如图17所示，计算结果如图18所示。



图 17 激励线圈下方区域划分

Fig. 17 Division of area below excitation coil

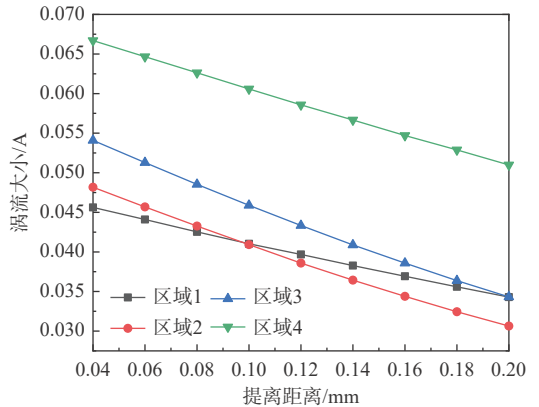


图 18 不同区域涡流大小随提高距离变化趋势

Fig. 18 Variation trend of eddy current size with lifting distance in different areas with lift-off

根据图18可以看出，随着传感器提高距离增加，激励线圈下方4个区域的涡流大小均呈现下降趋势。由于提高距离较大，被测试验件表面形成的涡流较小，当裂纹扩展穿过激励线圈下方区域时，所形成的扰动涡流较小。因此，由扰动涡流形成的扰动磁场较小。综合传感器感应通道感应电压增大及扰动涡流减小两者因素，传感器对裂纹识别的灵敏度较小，但是也不影响传感器对孔边裂纹的定量监测。

### 4 误差分析

根据实验结果和仿真结果对裂纹监测之间的差异, 本节开展误差分析。一方面, 考虑到传感器在安装过程中受到垫片的挤压, 传感器会发生变形, 传感器与金属表面之间的间距。另一方面, 考虑到传感器还有保护层、PI 补强板的厚度在安装过程中需要采用 PI 胶进行黏接, 因此传感器实际的提高距离会大于 0.1 mm。根据图 7 和图 18 可知, 随着提高距离的增加, 感应通道的感应电压会增加, 被测试件表面的涡流大小会减小。随着裂纹的扩展, 传感器对裂纹识别的灵敏度会降低。因此, 本文为了证明该结论, 开展了仿真分析, 分析提高距离在 0.15 mm 和 0.2 mm 下, 传感器对裂纹监测的灵敏度变化趋势, 如图 19 所示。根据仿真结果可以看出, 随着提高距离的增加, 传感器各感应通道对裂纹识别的灵敏度均呈现下降趋势, 对比实验结果和仿真结果可以看出, 在提高距离为 0.15 mm 时, 传感器对裂纹识别的实验结果和仿真结果较为相近。因此, 可以看出仿真结果和实验结果的差异是由提高距离引入。但是, 在实验过程中, 带补强片的传感器在裂纹扩展过程中, 可以明显引起传感

器感应通道信号的变化。因此, 可以判断由补强片引起的提高距离的增加并不严重影响传感器对裂纹的识别和定量监测。

### 5 结论

- 1) 带双面补强的柔性涡流阵列传感器具有较好的抗压性能, 传感器在疲劳裂纹监测实验过程中不发生破坏。
- 2) 传感器的参考通道可以有效消除外界环境对传感器输出信号的影响。
- 3) 传感器对裂纹识别的精度与激励线圈之间的间距相等。
- 4) 传感器随提高距离的增加, 裂纹监测灵敏度降低, 实验结果与仿真结果之间的差异由提高距离引入。

### 参考文献 (References)

[1] CHEN T, HE Y T, DU J Q. A high-sensitivity flexible eddy current array sensor for crack monitoring of welded structures under varying environment[J]. *Sensors*, 2018, 18(6): 1780.

[2] CHEN T, DU J Q, HE Y T. A structural crack monitoring gasket for aircraft bolt-jointed structures with temperature compensation[J]. *Smart Materials and Structures*, 2018, 27(11): 115004.

[3] LIU M B, LI B B, LI J T, et al. Smart coating sensor applied in crack detection for aircraft[J]. *Applied Mechanics and Materials*, 2013, 330: 383-388.

[4] ROACH D. Real time crack detection using mountable comparative vacuum monitoring sensors[J]. *Smart Structures and Systems*, 2009, 5(4): 317-328.

[5] MAJUMDER M, GANGOPADHYAY T K, CHAKRABORTY A K, et al. Fibre bragg gratings in structural health monitoring—Present status and applications[J]. *Sensors and Actuators A: Physical*, 2008, 147(1): 150-164.

[6] HUNT S R, HEBDEN I G. Validation of the eurofighter typhoon structural health and usage monitoring system[J]. *Smart Materials and Structures*, 2001, 10(3): 497.

[7] WILLBERG C, DUCZEK S, PEREZ J M V, et al. Comparison of different higher order finite element schemes for the simulation of Lamb waves[J]. *Computer Methods in Applied Mechanics and Engineering*, 2012, 241-244: 246-261.

[8] YAMAGUCHI T, UEDA M. An active sensor for monitoring bearing wear by means of an eddy current displacement sensor[J]. *Measurement Science and Technology*, 2006, 18(1): 311-317.

[9] 刘凯, 崔荣洪, 侯波, 等. PVD薄膜传感器裂纹检测概率测定与分析[J]. *材料工程*, 2019, 40(9): 160-166.

LIU K, CUI R H, HOU B, et al. Estimation and analysis of probability of PVD film sensor for crack detection[J]. *Journal of Materials Engineering*, 2019, 40(9): 160-166(in Chinese).

[10] GOLDFINE N J, WASHBAUGH A P, DEARLOVE J V, et al. Imposed  $\omega$ -k magnetometer and dielectrometer applications[M]. Berlin: Springer, 1993.

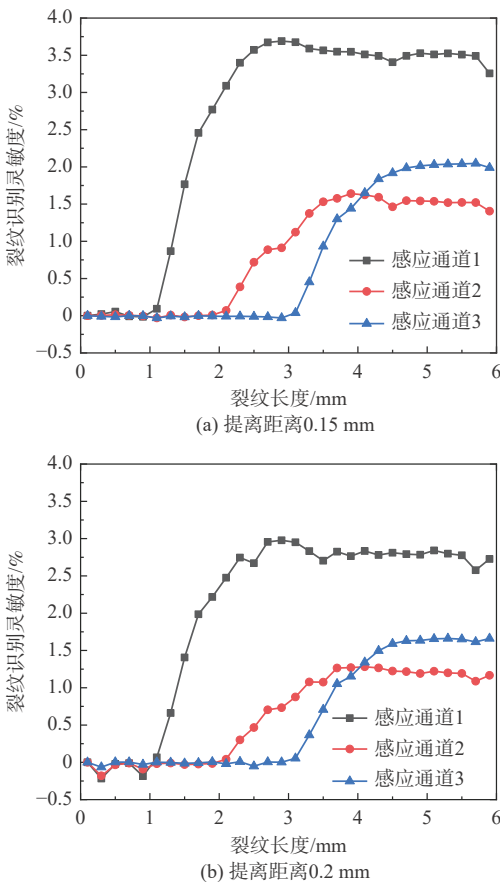


图 19 不同提高距离下裂纹识别灵敏度

Fig. 19 Sensitivity of crack monitoring at different lift-off

- [11] SCHLICKER D E. Imaging of absolute electrical properties using electroquasistatic and magnetoquasistatic sensor arrays[D]. Cambridge: Massachusetts Institute of Technology, 2006.
- [12] RAKOW A, CHANG F K. A structural health monitoring fastener for tracking fatigue crack growth in bolted metallic joints[J]. *Structural Health Monitoring*, 2012, 11(3): 253-267.
- [13] 丁华, 焦胜博, 何宇廷, 等. 花萼状涡流阵列传感器裂纹扰动半解析模型构建[J]. *中国电机工程学报*, 2014, 34(3): 495-502.  
DING H, JIAO S B, HE Y T, et al. Semi-analytical crack perturbation model construction of rosette eddy current sensors[J]. *Proceedings of the CSEE*, 2014, 34(3): 495-502(in Chinese).
- [14] 李培源, 何宇廷, 杜金强, 等. 基于柔性涡流传感器疲劳裂纹监测试验研究[J]. *传感器与微系统*, 2015, 34(1): 24-31.  
LI P Y, HE Y T, DU J Q, et al. Research on fatigue crack monitoring experimental based on flexible eddy current sensor[J]. *Transducer and Microsystem Technologies*, 2015, 34(1): 24-31(in Chinese).
- [15] 杜金强, 何宇廷, 李培源. 矩形柔性涡流阵列传感器裂纹检测研究[J]. *传感器与微系统*, 2014, 33(5): 12-17.  
DU J Q, HE Y T, LI P Y. Research on crack inspecting of rectangular flexible eddy current array sensor[J]. *Transducer and Microsystem Technologies*, 2014, 33(5): 12-17(in Chinese).
- [16] FAN X H, CHEN T, HE Y T, et al. An excitation coil layout method for improving the sensitivity of a rosette flexible eddy current array sensor[J]. *Smart Materials and Structures*, 2020, 29(1): 015020.
- [17] FAN X H, CHEN T, DU J Q, et al. Methods for improving sensitivity of crack quantitative monitoring of flexible eddy current array sensor[J]. *Smart Materials and Structures*, 2020, 29(8): 085033.
- [18] 《航空制造工程手册》总编委会. 航空制造工程手册: 飞机装配[M]. 2版. 北京: 航空工业出版社, 2010: 391-392.  
Aeronautical manufacturing engineering handbook editorial board. *Aeronautical manufacturing engineering manual: Aircraft assembly*[M]. 2nd ed. Beijing: Aviation Industry Press, 2010: 391-392(in Chinese).

## Hole edge crack monitoring technology of flexible eddy current array sensor

FAN Xianghong, GOU Baiyong<sup>\*</sup>, CHEN Tao, HE Yuting, CUI Ronghong, YU Jian

(Aviation Engineering School, Air Force Engineering University, Xi'an 710038, China)

**Abstract:** A flexible eddy current array sensor with double-side reinforcement is proposed for hole edge crack monitoring. First, with COMSOL finite element method, a finite element model of the sensor and the test piece was established, through which the effects of lift-off, permeability of the gasket and crack growth on the output signals of the sensor were analyzed. Then, sensors with and without reinforcement were prepared and the extrusion experiment and on-line fatigue crack monitoring experiment were carried out, followed by error analysis based on the difference between the experimental results and the simulation results. The results showed that with the increase of lift-off and permeability of the gasket, the output inductive voltage of the sensor increased gradually, while the sensitivity of the crack detection sensor decreased with the mounting lift-off. The sensor with reinforcement can work when the tightening torque was 63 N·m, while the sensor without reinforcement completely failed when the tightening torque was 50 N·m. The online crack monitoring experiments also verified that the sensor with reinforcement has quantitative crack monitoring ability, and the crack monitoring accuracy is consistent with the distance between the excitation coils, which is up to 1 mm. The difference between the experimental results and the simulation results was mainly introduced by the lift-off.

**Keywords:** flexible eddy current array sensor; hole edge crack monitoring; sensitivity; double-sided reinforcement; lift-off

# 基于循环神经网络的 SPMA 协议信道状态智能检测改进算法

张彦晖<sup>1</sup>, 吕娜<sup>1,\*</sup>, 缪竞成<sup>1</sup>, 高旗<sup>1</sup>, 王翔<sup>1</sup>, 陈卓<sup>2</sup>

(1. 空军工程大学信息与导航学院, 西安 710077; 2. 中国人民解放军 94619 部队, 六安 237000)

**摘 要:** 为实现对时敏目标的快速探测、定位和打击, 战术瞄准网络技术 (TTNT) 对战术信息接入信道、交互传输的实时性、可靠性提出高要求。TTNT 采用基于统计优先的多址接入 (SPMA) 协议, 通过周期性计算统计平均的思想, 估计当前信道状态, 控制战术信息接入信道的时机。该思想仅适用于流量相对平稳的情况, 在流量非平稳时会导致较大的信道状态检测误差。针对此问题, 引入流量预测技术, 提出基于循环神经网络的 SPMA 协议信道状态智能检测改进算法。利用循环神经网络的学习特点学习历史流量数据的隐含特征, 构建流量预测器对瞬时时刻的流量脉冲到达数进行实时预测, 从而准确获取当前信道状态。实验结果表明: 所提算法对信道状态的检测结果更接近真实值, 显著降低了信道忙闲状态的误判率。

**关键词:** 统计优先的多址接入协议; 信道状态检测; 流量预测; 循环神经网络; 战术瞄准网络技术

中图分类号: TN929.5; V243

文献标志码: A

文章编号: 1001-5965(2023)03-0735-10

作为数字化战场的“黏合剂”和“倍增器”, 数据链战术通信系统有机链接不同作战任务单元, 显著提升作战效能<sup>[1]</sup>。21 世纪初研制的战术瞄准网络技术 (tactical target network technology, TTNT) 数据链, 有机连通侦察单元和火控单元, 实现对时敏目标的协同探测、协同定位和精确打击<sup>[2-3]</sup>。

精确火力打击对 TTNT 数据链战术信息的时效性、可靠性提出了极高要求, 如时延小于 2 ms, 高优先级信息可靠性 99%<sup>[4-6]</sup>。TTNT 数据链采用基于统计优先的多址接入 (statistical priority-based multiple access, SPMA) 协议, 通过随机竞争接入、多信道并行传输、信道状态统计检测、多优先级分级处理等机制<sup>[7-15]</sup>, 确保战术信息的低时延、高可靠传输。

SPMA 协议的信道状态检测算法通过周期性统计网络中传输的流量脉冲数目, 对信道占用状态进

行判断, 以一个周期内信道占用率的平均值作为当前发送时刻信道状态的检测值。TTNT 数据链在协同探测、定位和打击过程中, 生成和传输的信息多样, 不仅有目标探测、态势感知、任务决策、战术协同等信息, 也有火力协同、精确制导和威胁告警等信息。这些信息流量的带宽需求变化大, 如窄带的态势、任务分发等信息, 宽带的目标图像、视频等信息; 另外, 信息流量的生成规律多样, 有周期性、随机性、阶段性等; 信息流量的信道接入需求变化较大, 高优先级信息需要立即接入信道、低优先级信息排队接入信道等。可见, TTNT 数据链中交互传输的信息流量变化复杂, 具有阶段突发性和强波动性。对于这些流量特点, 基于统计平均的信道状态检测存在较大判决误差, 无法准确反映波动特点。而目前对 SPMA 协议的研究主要集中在多优

收稿日期: 2021-06-08; 录用日期: 2021-08-29; 网络出版时间: 2021-09-14 17:41

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20210914.1500.007.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210914.1500.007.html)

基金项目: 国家自然科学基金 (61703427); 陕西省自然科学基金 (2020JQ-483)

\* 通信作者. E-mail: [1105406986@qq.com](mailto:1105406986@qq.com)

**引用格式:** 张彦晖, 吕娜, 缪竞成, 等. 基于循环神经网络的 SPMA 协议信道状态智能检测改进算法 [J]. 北京航空航天大学学报, 2023, 49 (3): 735-744. ZHANG Y H, LYU N, MIAO J C, et al. Improved intelligent detection algorithm for SPMA protocol channel state based on recurrent neural network [J]. Journal of Beijing University of Aeronautics and Astronautics, 2023, 49 (3): 735-744 (in Chinese).

先级业务的排队性能理论分析<sup>[13]</sup>、信道状态阈值动态设置<sup>[10]</sup>等,信道状态检测算法的改进研究较少。因此,为减小检测误差,更加准确地判断信道瞬时状态,需要设计一种能够适应 TTNT 数据链流量变化的信道状态检测算法。

流量预测通常采用预测模型按照特定的规律对未来某段时间的流量变化进行预测<sup>[16]</sup>。传统的流量预测模型如泊松过程、马尔可夫、自回归、卡尔曼滤波模型等,将网络中流量建模成某种数学模型,理论基础较为成熟,评价易于表达,但 TTNT 数据链流量特点复杂,将其理想化为某种数据模型会带来较大误差。考虑到流量变化的本质来源于用户的行为,产生了基于用户行为建模的流量预测方法,通过分析用户的行为特点,计算用户在不同空间和时间的通信概率,从而实现流量预测,但是这种方法结构复杂,且在 TTNT 数据链应用背景下,对用户行为的学习难度较大,计算效率和准确性偏低,难以大规模推广。

随着人工智能的发展,产生了基于深度学习的流量预测方法并不断发展。文献 [17] 分析对比了人工神经网络 (artificial neural networks, ANN) 和几种传统的平稳时间预测模型的性能,得出 ANN 模型的综合性能要优于其他几种模型的结论。文献 [18] 提出了基于 BP 神经网络的非线性网络流量预测方法,并证明了其有效性。文献 [19] 将 ANN 与多层感知器、堆叠自编码器模型进行对比,证明 ANN 的计算复杂度较小且准确性相对较高,并指出循环神经网络 (recurrent neural networks, RNN) 与其他类型的神经网络相比,具有更高的预测性能。文献 [20] 进一步将几种常见的 RNN 模型进行对比,发现长短期记忆网络 (long short-term memory network, LSTM) 和门控循环单元 (gated recurrent unit, GRU) 的性能好于其他 ANN 模型。文献 [21] 详细总结了各种类型神经网络流量预测模型的优缺点和适应场景。文献 [22-23] 提出在 RNN 中引入注意力机制,进一步降低了计算复杂度。

本文结合应用人工智能的流量预测算法与 SPMA 协议信道检测机制的关联性,针对 SPMA 协议信道状态检测算法对 TTNT 数据链流量的突发性和强波动性特点适应性不足问题,提出一种基于 RNN 的 SPMA 协议信道状态智能检测改进算法。使用 RNN 预测当前到达的流量脉冲数,从而较为准确地得到当前的信道状态。

## 1 问题描述

### 1.1 SPMA 协议及其信道状态检测机制

SPMA 协议作为 TTNT 数据链的信道接入协议,通过控制用户接入信道发送数据的时机,保证

了数据传输的实时性和可靠性。SPMA 协议按照不同消息的服务质量 (quality of service, QoS) 需求把消息划分成不同的优先级,并分别设置不同优先级的信道占用率阈值。消息到达后进入不同的优先级队列,高优先级消息优先进入待发送状态。其信道接入控制流程如图 1 所示。当有数据包进入待发送状态时,先进行信道状态检测,将当前信道占用率与对应待发送数据包的阈值对比,若小于阈值,信道状态为闲,发送数据;若大于阈值,信道状态为忙,数据包回退一段时间,若在回退时间内有更高数据包到达,则更新待发送状态的数据包为更高优先级数据包,再次进行信道状态检测,若回退时间内无更高优先级数据包到达,则回退时间结束后再次进行信道状态检测。直到信道状态检测结果为闲,发送数据。

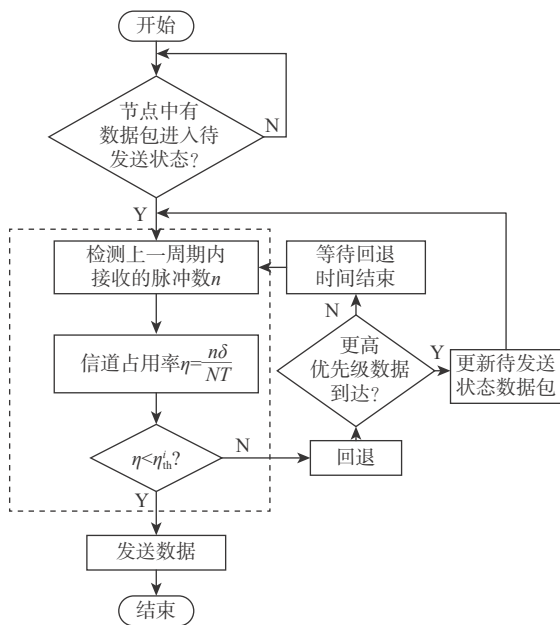


图 1 SPMA 协议信道接入控制流程

Fig. 1 Flow chart of SPMA protocol channel access control

SPMA 协议传统的信道状态检测算法是通过周期性地检测系统在该周期内传输的数据脉冲总数计算出信道占用率,进而与阈值对比,判断信道忙闲状态。

信道占用率的计算公式为

$$\eta = \frac{n\delta}{NT} \quad (1)$$

式中:  $n$  为一个周期内到达的流量脉冲数;  $\delta$  为传输每个流量脉冲需要的时间;  $T$  为周期长度;  $N$  为子信道数量。

进而可对信道状态检测机制进行数学建模。信道检测机制的输出共有 2 种情况,即信道闲和信道忙,如下:

$$\begin{cases} H_0: \eta \geq \eta_{th} \\ H_1: \eta < \eta_{th} \end{cases} \quad (2)$$

式中:  $H_0$  为信道忙;  $H_1$  为信道闲;  $\eta_{th}$  为信道占用率阈值。

将式 (1) 代入式 (2) 可得

$$\begin{cases} H_0: \eta = \frac{n\delta}{NT} \geq \eta_{th} & n \geq \frac{\eta_{th}NT}{\delta} = \gamma \\ H_1: \eta = \frac{n\delta}{NT} < \eta_{th} & n < \frac{\eta_{th}NT}{\delta} = \gamma \end{cases} \quad (3)$$

$$\begin{cases} H_0: n \geq \gamma \\ H_1: n < \gamma \end{cases} \quad (4)$$

由式 (4) 可判断出根据周期内到达的流量脉冲数, 即可判断信道忙闲状态。

### 1.2 问题分析

1.1 节对 SPMA 协议的信道状态检测算法进行了数学描述, 用周期内的平均信道占用率作为依据判断信道忙闲状态, 而理论上只有周期  $T$  无限趋近于 0 时才能计算出当前时刻信道状态的瞬时值, 本节对理想条件下的信道状态检测机制进行了理论分析, 并与实际的检测结果进行对比, 总结了 SPMA 协议信道状态检测算法存在的问题。

假设流量脉冲数的到达率  $\alpha(t)$  为一个随时间变化的函数, 则周期内到达的流量脉冲数  $n$  为

$$n = \int_{t_0-T}^{t_0} \alpha(t) dt \quad (5)$$

结合式 (1) 和式 (5) 可知, 当  $T \rightarrow 0$  时, 信道占用率瞬时值为

$$\eta = \lim_{T \rightarrow 0} \frac{n\delta}{NT} = \lim_{T \rightarrow 0} \frac{\left( \int_{t_0-T}^{t_0} \alpha(t) dt \right) \delta}{NT} = \frac{\alpha(t_0-T)\delta}{N} \approx \frac{\alpha(t_0)\delta}{N} \quad (6)$$

将式 (6) 代入式 (2) 可得

$$\begin{cases} H_0: \eta = \frac{\alpha(t_0)\delta}{N} \geq \eta_{th} & \alpha(t_0)\delta \geq \eta_{th}N \\ n(t_0) \geq \eta_{th}N = \frac{\gamma\delta}{T} = \gamma' \\ H_1: \eta = \frac{\alpha(t_0)\delta}{N} < \eta_{th} \\ \alpha(t_0)\delta < \eta_{th}N & n(t_0) < \eta_{th}N = \frac{\gamma\delta}{T} = \gamma' \end{cases} \quad (7)$$

$$\begin{cases} H_0: n(t_0) \geq \gamma' \\ H_1: n(t_0) < \gamma' \end{cases} \quad (8)$$

式中:  $n(t_0)$  为从  $t_0$  时刻开始, 一个流量脉冲传输时间  $\delta$  内流量脉冲到达数量。

通过上述分析可以得出, SPMA 协议信道状态检测算法的信道占用率检测值  $\eta_{detect}$  和理论值  $\eta_{theory}$  分别为

$$\begin{cases} \eta_{detect} = \frac{n\delta}{NT} = \frac{n\delta/T}{N} = \frac{\bar{\alpha}\delta}{N} \\ \bar{\alpha} = n/T \\ \eta_{theory} = \frac{\alpha(t_0)\delta}{N} \end{cases} \quad (9)$$

由式 (4)、式 (8) 和式 (9) 的结果可知, SPMA 协议的信道状态检测算法关注的是一个周期内网络中传输的流量脉冲数来判断信道忙闲状态, 而理论上必须已知当前时刻待传输的流量脉冲数才能判断信道忙闲状态, 由式 (9) 可知, 信道状态检测算法更关注平均值, 而真正需要检测的是瞬时值, 信道状态检测算法的检测结果较理论上的真实值存在一定误差。不同样本情况下信道状态检测算法检测情况如图 2 所示。当周期内流量脉冲数变化较小时, 信道状态检测算法能够较准确地表示当前信道状态; 而当周期内流量脉冲数变化较大时, 信道状态检测算法将会带来较大的误差, 影响信道状态判断结果。综上所述, 为降低信道状态检测的误差, 需要对历史周期数据进行处理, 得到一个尽量接近当前时刻待传输流量脉冲数, 从而相对准确地判断出信道忙闲状态。流量预测技术可以通过历史的流量数据来对下一刻的流量进行预测, 恰好与 SPMA 的信道状态检测问题契合。

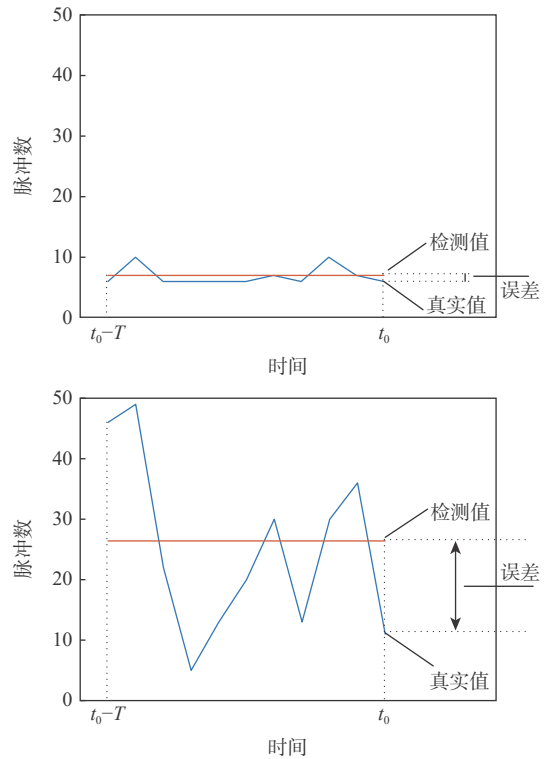


图 2 不同样本下信道状态检测算法误差  
Fig. 2 Error of channel state detection algorithm under different samples

## 2 信道状态智能检测改进算法

针对第 1 节中分析的 SPMA 协议信道状态检测算法存在的问题, 提出一种基于 RNN 的 SPMA 协议信道状态智能检测改进算法。为提高信道状态检测的准确性, 引入流量预测技术, 采用基于 RNN 的流量预测算法, 通过对历史流量数据的学

习,较为准确地预测当前时刻到达的流量脉冲数,设计流量脉冲数阈值,进而对比判断出信道忙闲状态。图3为改进后SPMA协议信道接入控制流

程。为叙述方便,将SPMA协议信道状态检测算法简称为传统算法,将SPMA协议信道状态智能检测改进算法简称为改进算法。

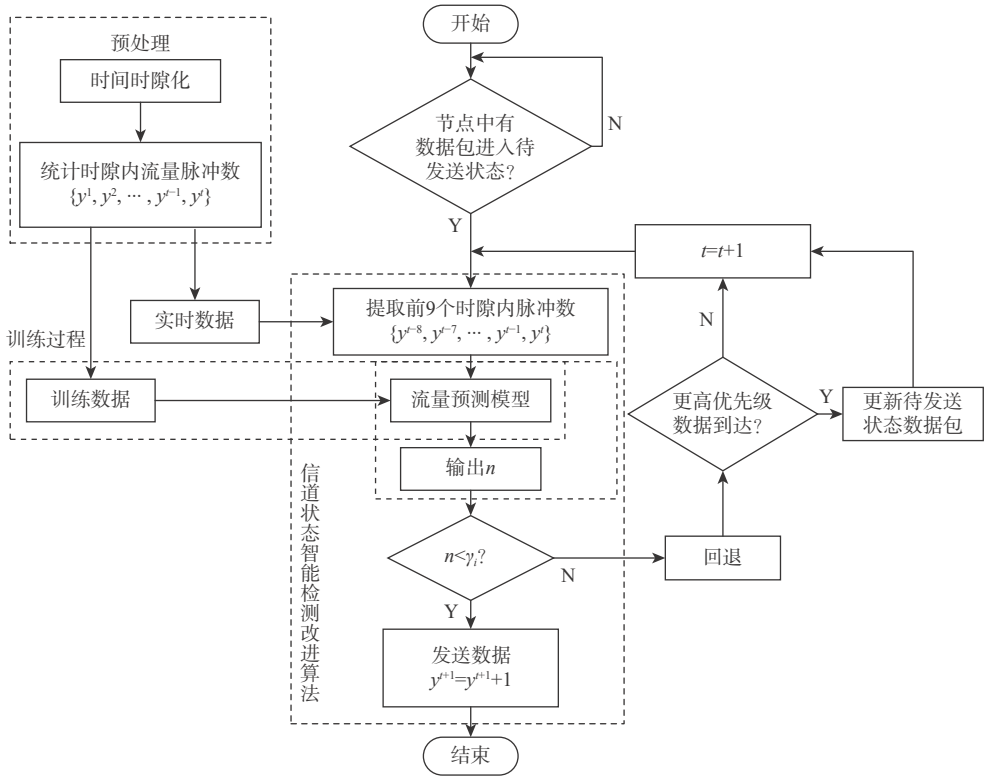


图3 改进后SPMA协议信道接入控制流程

Fig. 3 Flow chart of improved SPMA protocol channel access control

## 2.1 预处理

预处理主要根据原始数据特点和流量预测模型的输入格式,将原始数据处理成流量预测模型可以接收且易于学习的数据。

### 2.1.1 处理思路

上层数据是流量变化曲线,流量预测模型主体是循环神经网络,适合处理序列数据,因此预处理要把流量曲线转化成序列。

进一步分析序列的构建方法。传统算法将时间划分成周期,每周期仅计数一次,作为信道的估计,为能反映各信道的情况,周期通常不能选的过小。与传统算法不同,采用基于RNN的流量预测算法需要更多的数据来反映流量的变化规律,且根据1.2节的理论分析,采用小时隙块来表示当前信道状态更接近真实值,因此要将时间划分成小的时隙。计算出每个时隙内到达的流量脉冲数,则将流量数据转化成RNN擅长处理的时间序列。

### 2.1.2 时隙块大小

关于时隙块大小的选择,理论上选择单个流量脉冲发送时间 $\delta$ 能够最为准确地表示信道状态,但在四信道情况下,容易导致序列出现长0或者长

4的情况,不能较好地反映出流量变化趋势,综合流量脉冲长度和信道带宽,最终选择10ms为一个小时隙块,能够显著反映流量变化。则一段时间内的流量脉冲数可表示为

$$\{y^1, y^2, y^3, \dots, y^{n-1}, y^n\}$$

借鉴文献[10]中传统算法的周期100ms,不增加计数时间,选择前9个时隙数据作为RNN的输入,下一时隙数据作为输出,并向后滑动,构建训练集:

$$\begin{aligned} \{y^1, y^2, y^3, \dots, y^9\} &\rightarrow y^{10} \\ \{y^2, y^3, y^4, \dots, y^{10}\} &\rightarrow y^{11} \\ \{y^3, y^4, y^5, \dots, y^{11}\} &\rightarrow y^{12} \\ \{y^{n-9}, y^{n-8}, y^{n-7}, \dots, y^{n-1}\} &\rightarrow y^n \end{aligned}$$

## 2.2 流量预测模型

序列预测通常采用RNN的2个变体:LSTM和GRU。LSTM包含2个隐态和3个门,参数较多,网络较为复杂,相比之下GRU在性能上与LSTM相差不大,但其网络结构更为简单,因此采用GRU搭建流量预测模型,进行流量预测。GRU最早在文献[24]中提出,并在应用中不断改进发展,本文模型在经典GRU的基础上,在门函数内增

加偏置以提高拟合能力。GRU 包含 2 个门和 1 个隐态, 分别为更新门、重置门和状态信息, 结构如图 4 所示。

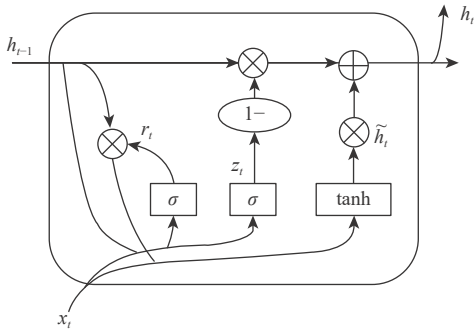


图 4 门控循环单元结构

Fig. 4 Structure diagram of gated recurrent unit

$r_t$  为更新门, 控制上一时刻信息输入到当前状态的待更新信息  $\tilde{h}_t$  的程度:

$$\begin{cases} r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \\ \tilde{h}_t = \tanh(W_{\tilde{h}} \cdot [r_t \times h_{t-1}, x_t]) \end{cases} \quad (10)$$

$z_t$  为重置门, 对当前时刻待更新信息进行有选择的更新:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad (11)$$

则更新后的状态信息为

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times \tilde{h}_t \quad (12)$$

式中:  $W$  和  $b$  分别为各个门的连接权重和偏置;  $x$  为 GRU 输入;  $\sigma$  为 sigmoid 函数;  $\cdot$  与  $x$  分别为矩阵乘法与矩阵内按元素乘积。

具体的流量预测神经网络结构如图 5 所示。

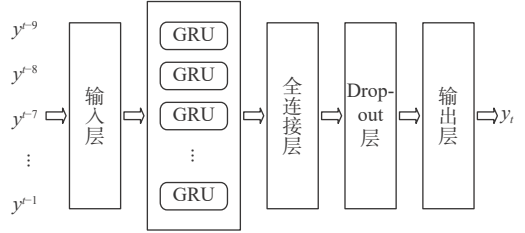


图 5 流量预测神经网络结构

Fig. 5 Structure diagram of traffic prediction neural network

### 2.3 训练及预测

训练过程是神经网络进行线下学习数据特征规律, 预测是完成训练的神经网络在线上实现预测功能。通过分析数据特点, 合理设计训练及预测过程有助于提升模型的预测性能。

如图 6 所示, 空战背景下, TTNT 数据链用户在完成不同任务或者完成任务的不同阶段产生的业务种类及其 QoS 需求不同, 而不同时间下的各种类型业务在网络中传输所产生的流量不同, 导致流量具有周期性、突发性、强波动性的特点, 即用户行为特征决定业务特征, 业务特征决定流量脉冲特征。

在同样或者相似的作战背景中, 历史的用户行为和实时的用户行为是具有相关性的, 其传输的业务即历史统计业务和实时业务, 也是具有相关性的, 而作为底层的流量脉冲也必然会承载着用户行为特征信息和业务特征信息, 通过线下对历史流量脉冲的学习, 能够学习到隐藏的序列特征和脉冲序列上层的特征, 即用户行为特征和业务特征, 从而在线上实现对实时情况的预测。

图 7 为流量预测模型的结构及其训练和预测过程。由图 5 可知, 神经网络能够实现式 (13) 所示的映射关系:

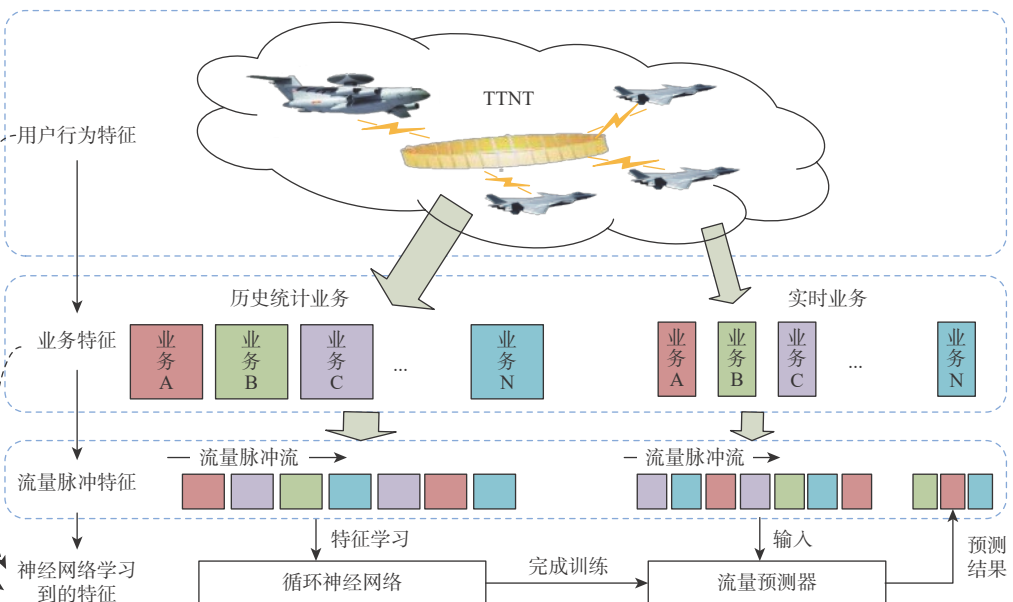


图 6 循环神经网络特征学习过程

Fig. 6 Feature learning process of recurrent neural network

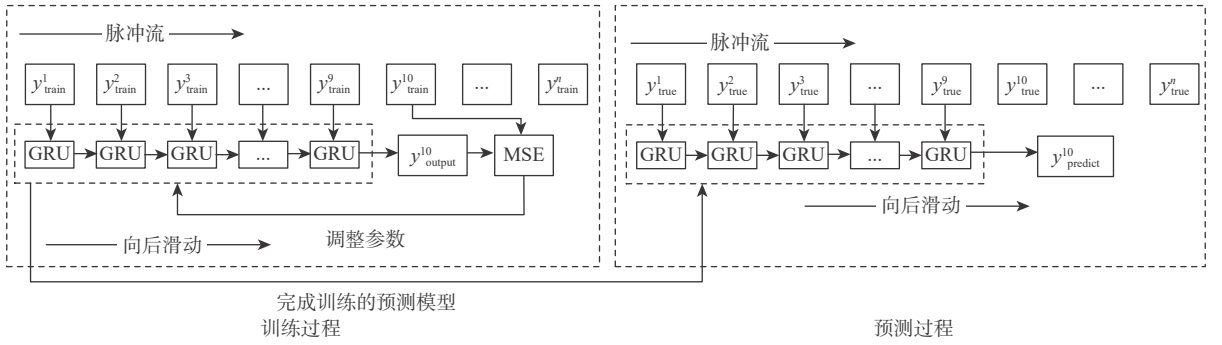


图7 流量预测模型的训练和预测过程

Fig. 7 Training and prediction process of traffic prediction model

$$y^t = F(\{y^{t-9}, y^{t-8}, y^{t-7}, \dots, y^{t-1}\}) \quad (13)$$

则将图7所示的训练集输入可得

$$y_{\text{output}}^{10} = F(\{y_{\text{train}}^1, y_{\text{train}}^2, y_{\text{train}}^3, \dots, y_{\text{train}}^9\}) \quad (14)$$

$$\text{MSE} = (y_{\text{output}}^{10} - y_{\text{train}}^{10})^2 \quad (15)$$

MSE为误差,作为损失函数来调整神经网络的参数。在实验中通常采用批处理,当一个批次有 $m$ 个样本时,MSE调整为

$$\text{MSE} = \frac{1}{m} \sum_{i=1}^m (y_{\text{output}}^i - y_{\text{train}}^i)^2 \quad (16)$$

通过调整参数,完成训练后能够得到图7右侧所示适应TTNT数据链流量特点的神经网络映射关系 $G(\cdot)$ ,如式(17)所示,实现预测功能:

$$y_{\text{predict}}^{10} = G(\{y_{\text{true}}^1, y_{\text{true}}^2, y_{\text{true}}^3, \dots, y_{\text{true}}^9\}) \quad (17)$$

## 3 仿真验证

### 3.1 数据集

本文实验数据利用wireshark软件在TTNT数据链仿真平台进行抓包构建,用TTNT数据链仿真平台模拟实际作战场景中TTNT数据链从发现目标到摧毁目标到最后的评估过程,为TTNT数据链设置各种类型的业务,模拟真实作战中产生的各种业务,每次任务模拟5 min,共模拟5次任务。

对5次模拟作战任务的全过程进行抓包,以10 ms为间隔统计流量脉冲数,生成流量脉冲序列,每个序列含30 000个数据。前4次抓包产生的数据作为训练集,最后1次产生的数据作为测试集进行实验,并在测试集中抽取具有代表性的连续的700个数据进行分析。

### 3.2 流量预测准确性分析

算法预测准确性采用4个常用指标评判,即均方误差(mean squared error, MSE)、均方根误差(root mean squared error, RMSE)、平均绝对误差(mean absolute error, MAE)、 $R^2$ 值:

$$\text{MSE} = \frac{1}{m} \sum_{i=1}^m (y_{\text{predict}}^i - y_{\text{true}}^i)^2$$

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_{\text{predict}}^i - y_{\text{true}}^i)^2}$$

$$\text{MAE} = \frac{1}{m} \sum_{i=1}^m |y_{\text{predict}}^i - y_{\text{true}}^i|$$

$$R^2 = 1 - \frac{\sum_{i=1}^m (y_{\text{predict}}^i - y_{\text{true}}^i)^2}{\sum_{i=1}^m (\bar{y} - y_{\text{true}}^i)^2}$$

式中:MSE、RMSE和MAE表示检测值的均方误差、均方根误差和平均绝对误差,误差越小检测值越准确; $R^2$ 值表示检测算法对数据的拟合能力, $R^2$ 值越大表示算法的拟合能力越强,更适合处理该类型数据。

图8和图9分别为传统算法和改进算法的流量预测结果。可以看出,当流量脉冲数波动较小、较为平滑时,2种算法均能达到较好的预测结果,而当出现突发流量或者流量波动性较强的情况,传统算法的预测结果存在显著的误差,而改进算法仍能够实现较为准确的预测。

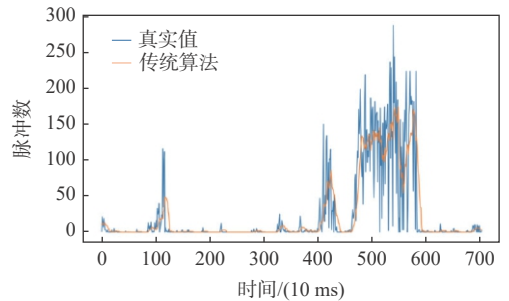


图8 传统算法的检测结果

Fig. 8 Detection results of traditional algorithms

表1为2种算法预测结果的评价指标对比。改进算法在4个评价指标均大幅度优于传统算法,证明了改进算法的检测准确性更强,对TTNT数据链流量数据的拟合能力更好,更适合检测TTNT数据链的信道状态。

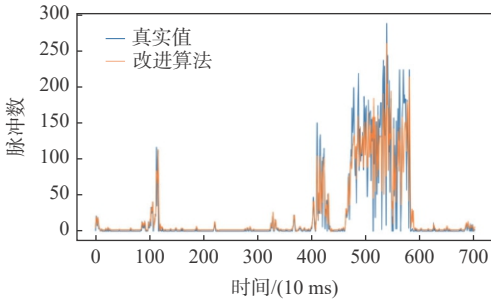


图 9 改进算法的检测结果

Fig. 9 Detection results of improved algorithms

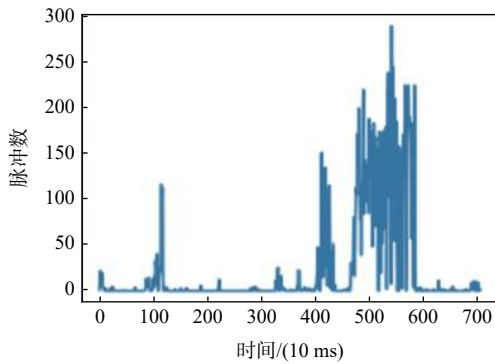
表 1 两种算法检测准确性对比

Table 1 Comparison of detection accuracy of two algorithms

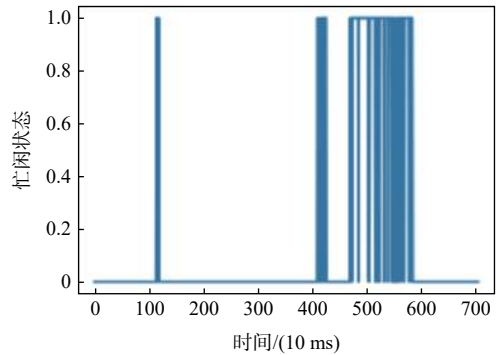
算法	MSE	RMSE	MAE	$R^2$
传统算法	931.53	30.52	13.71	0.67
改进算法	313.56	17.70	7.78	0.89

### 3.3 忙闲状态误判率对比

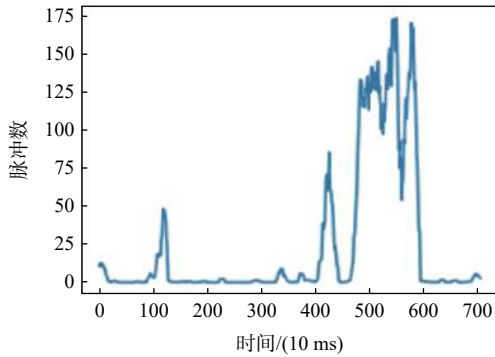
3.2 节中对比了流量预测的准确性, 但是 SPMA 的信道状态检测机制最终是要判断出信道的忙闲状态, 从而决定是否发送数据。为进一步证明本文算法的有效性, 本节对改进算法与传统算法的信道忙闲状态的误判率进行对比。0 为信道闲, 数据可以发送; 1 为信道忙, 数据不能发送。为对比改进算法对不同优先级消息的信道忙闲状态的检测情况, 本节设置高、中、低 3 个优先级, 为保证高优先级消息优先传输, 高优先级消息阈值大于低优先级消息阈值, 将高优先级消息阈值设置为数据的 3 倍均方根 174、中优先级消息阈值设置为 2 倍均方根 116、低优先级消息阈值设置为 1 倍均方根 58。若当前信道状态小于阈值, 信道闲, 输出为 0; 信道状态大于阈值, 信道忙, 输出为 1。图 10 对比了真实



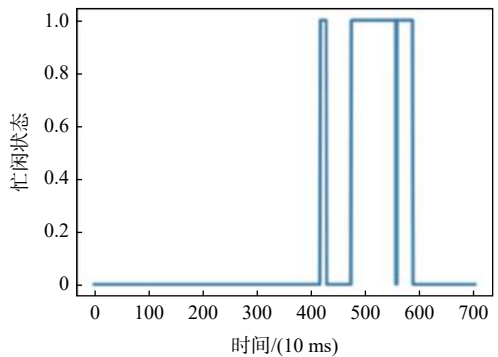
(a) 真实流量脉冲分布图



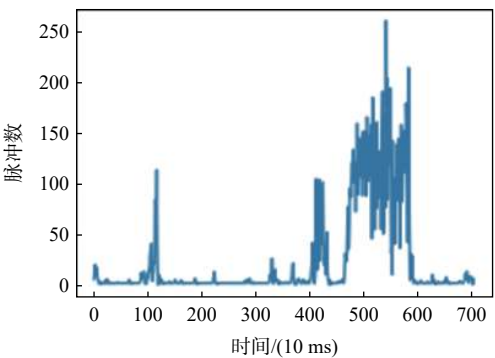
(b) 真实信道忙闲状态



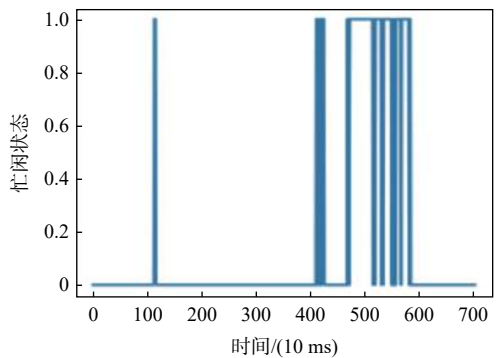
(c) 传统算法检测的流量脉冲分布



(d) 传统算法检测的信道忙闲状态



(e) 改进算法检测的流量脉冲分布



(f) 改进算法检测的信道忙闲状态

图 10 两种算法检测的流量脉冲及信道忙闲状态与真实数据的对比

Fig. 10 Comparison of traffic pulses and channel busy/idle status detected by two algorithms with real data

数据流量和真实信道忙闲状态与传统算法和改进算法检测流量和信道忙闲状态,图中所选阈值为58。

从图10中可以看出,改进算法的误判率明显低于传统算法,但在流量波动较小的情况下,2种算法的误判率差异不大,而在流量波动较大的情况下,传统算法难以对信道状态进行准确判断。

表2为不同类型流量下对低中高优先级消息的忙闲状态误判率。突发流量情况下,改进算法对低优先级消息的误判率明显低于传统算法;对于中高优先级消息,2种算法均不会出现误判。通过分析突发流量与各个优先级消息的阈值可以得出,当突发流量脉冲数低于阈值时,突发流量不会导致信道忙闲状态误判;当突发流量脉冲数高于阈值时,则会导致信道忙闲状态误判,且改进算法的误判率较传统算法显著降低。流量平稳情况下,2种算法基本上不会产生误判,这是因为一方面该区间真实值均稳定在阈值以下,另一方面2种算法均能较好地拟合流量波动小的情况,不会产生过大的误差导致检测值超过阈值。流量具有强波动性的情况,改进算法相对于传统算法误判率明显降低,在各个优先级阈值下的误判率降低了4.91%~15.23%。

表2 两种算法在不同类型流量下对低中高优先级消息的忙闲状态误判率

Table 2 Misjudgment probability of busy and idle status of two algorithms for low, medium and high priority messages under different types of traffic

流量类型	所用算法	低优先级消息误判率/%	中优先级消息误判率/%	高优先级消息误判率/%
突发流量[100,120]	传统算法	20	0	0
	改进算法	5	0	0
稳定流量[200,400]	传统算法	0	0	0
	改进算法	0	0	0
强波动流量[400,600]	传统算法	22.61	25.55	13.76
	改进算法	7.38	12.28	8.85

### 3.4 处理时延对比

改进算法在提高信道检测精度的同时也增加了一定的算法复杂度,导致处理时延增加,为验证改进算法是否满足 TTNT 数据链实时性需求,对改进算法的处理时延进行分析。

实验选用一台内存为 16 GB、搭载 2.5 GHz 主频处理器的 64 位 Win 10 操作系统 PC 代替 TTNT 数据链端机,由于算法单次运算时延过小,测量单次数据误差较大,实验中取算法运行 1 000 次的时延进行分析,如表3所示。

从实验结果看,改进算法的时延较传统算法增加了 0.044 1 ms,但 SPMA 协议的信道检测和信道

表3 两种算法时延对比

Table 3 Time Delay comparison of two algorithms

算法	1 000次时延/ms	平均单次时延/ms
传统算法	4.24	0.004 2
改进算法	48.30	0.048 3

接入是并行处理的,信道检测机制实时地检测信道状态,每隔固定的周期更新信道状态,改进算法多带来的 0.044 1 ms 对于信道状态更新周期 10 ms 来说几乎可以忽略,而信道接入时只需要将当前信道状态与阈值对比决定是否发送信息,因此改进算法的处理时延不会对通信时延造成影响,不会影响 TTNT 数据链的实时性。

## 4 结论

1) 对 SPMA 协议信道检测机制进行了完整的数学分析,指出 SPMA 协议信道检测机制对 TTNT 数据链流量突发性和强波动性特点适应性不足的问题。

2) 引入流量预测技术,提出基于循环神经网络的 SPMA 信道状态检测智能改进算法。

3) 改进算法预测结果的 MSE 为 313.56、RMSE 为 17.70、MAE 为 7.78、 $R^2$  为 0.89。预测性能相比于传统算法有明显提高,尤其是存在突发流量和强波动性流量数据。

4) 不同阈值下,改进算法的忙闲状态误判率较传统算法有所降低。分析验证集各段(模拟不同流量特点)的误判率,在存在突发流量和流量波动大的情况下,改进算法的误判率较传统算法降低了 4.91%~15.23%。

5) 本文算法在提升检测精度的同时也导致了有一定的处理时延,通过实验计算改进算法带来的平均时延为 0.048 3 ms,但改进算法带来的时延不会影响 TTNT 数据链通信的实时性。

本文算法虽在检测准确性和忙闲状态误判率 2 个指标上有明显提升,但预测误差和忙闲状态误判率仍存在降低空间,可以进一步调整优化算法以达到更高的性能。

## 参考文献 (References)

- [1] 吕娜,张岳彤,陈柯帆,等.数据链理论与系统[M].2版.北京:电子工业出版社,2018:5-6.  
LV N, ZHANG Y T, CHEN K F, et al. Data link theory and system[M]. 2nd ed. Beijing: Publishing House of Electronics Industry, 2018: 5-6(in Chinese).
- [2] 黄振,周永将.美军网络中心战的重要元素——协同数据链[J].现代导航,2017,8(1):70-73.  
HUANG Z, ZHOU Y J. Element of US army network centric war-

- fare—Cooperative sata link[J]. *Modern Navigation*, 2017, 8(1): 70-73 (in Chinese).
- [3] 戴辉. 武器协同数据链发展需求[J]. *指挥信息系统与技术*, 2011, 2(5): 11-14.
- DAI H. Development requirement of weapon cooperative data link[J]. *Command Information System and Technology*, 2011, 2(5): 11-14(in Chinese).
- [4] 金荣, 张衡阳. 美军TTNT数据链发展应用现状[J]. *现代导航*, 2014, 5(2): 154-156.
- JIN R, ZHANG H Y. Development situation of tactical targeting networks for US force[J]. *Modern Navigation*, 2014, 5(2): 154-156(in Chinese).
- [5] 陈志辉, 李大双. 对美军下一代数据链TTNT数据链技术的分析与探讨[J]. *信息安全与通信保密*, 2011(5): 76-79.
- CHEN Z H, LI D S. Analysis and exploration on tactical targeting network technology for the next generation tactical data link[J]. *Information Security and Communications Privacy*, 2011(5): 76-79(in Chinese).
- [6] HERDER J C, STEVENS J A. Method and architecture for TTNT symbol rate scaling modes: U. S. Patent 7, 839, 900[P]. 2010-11-23.
- [7] 周礼明. 统计优先级无线多址接入访问技术研究[D]. 成都: 电子科技大学, 2019: 10-18.
- ZHOU L M. Research on statistical priority based multiple access[D]. Chengdu: University of Electronic Science and Technology of China, 2019: 10-18(in Chinese).
- [8] 李志林. 面向高时敏业务的统计优先级多址接入技术研究[D]. 南京: 南京航空航天大学, 2019: 17-20, 45-66.
- LI Z L. Research on statistical priority-based multiple access technology for high time-sensitive services[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2019: 17-20, 45-66(in Chinese).
- [9] LIU J, PENG T, QUAN Q Y, et al. Performance analysis of the statistical priority-based multiple access[C]//3rd IEEE International Conference on Computer and Communications. Piscataway: IEEE Press, 2017: 30-35.
- [10] 刘健. 基于优先级概率统计的多址接入协议研究[D]. 北京: 北京邮电大学, 2018: 19-23.
- LIU J. Performance analysis of the statistical priority-based multiple access[D]. Beijing: Beijing University of Posts and Telecommunication, 2018: 19-23(in Chinese).
- [11] 王希洋. 基于统计优先级的数据链MAC协议研究[D]. 西安: 西安电子科技大学, 2017: 21-31.
- WANG X Y. Research on statistical priority-based MAC protocol for data link[D]. Xi'an: Xidian University, 2017: 21-31(in Chinese).
- [12] 吴琼. 基于统计优先级的链路层协议性能仿真与研究[D]. 北京: 北京理工大学, 2014: 17-22.
- WU Q. Performance simulation and research of the statistical priority-based multiple access[D]. Beijing: Beijing Institute of Technology, 2014: 17-22(in Chinese).
- [13] 王玉龙. TTNT数据链战术数据链媒体接入控制协议关键技术研究[D]. 北京: 北京交通大学, 2018: 21-36.
- WANG Y L. Research on key technologies of MAC protocol in TTNT tactical data link[D]. Beijing: Beijing Jiaotong University, 2018: 21-36(in Chinese).
- [14] 郑文庆, 金虎, 郭建蓬, 等. 一种新型数据链MAC协议及其信道占用研究[J]. *计算机仿真*, 2019, 36(7): 148-152.
- ZHENG W Q, JIN H, GUO J P, et al. Research on a new data link MAC protocol and its channel occupancy[J]. *Computer Simulation*, 2019, 36(7): 148-152(in Chinese).
- [15] 程文驰, 张施施. 一种数据链动态门限统计优先级多址接入协议[J]. *指挥与控制学报*, 2020, 6(1): 75-80.
- CHENG W C, ZHANG S S. A data link-oriented dynamic threshold statistical priority multiple access protocol[J]. *Journal of Command and Control*, 2020, 6(1): 75-80(in Chinese).
- [16] 臧运娟. 无线通信网络流量预测综述[J]. *电子设计工程*, 2017, 25(4): 150-153.
- ZANG Y J. Traffic prediction in wireless communication network: A survey[J]. *Electronic Design Engineering*, 2017, 25(4): 150-153 (in Chinese).
- [17] FENG H F, SHU Y T. Study on network traffic prediction techniques[C]//International Conference on Wireless Communications. Piscataway: IEEE Press, 2005: 1041-1044.
- [18] 刘杰, 黄亚楼. 基于BP神经网络的非线性网络流量预测[J]. *计算机应用*, 2007, 27(7): 1770-1772.
- LIU J, HUANG Y L. Nonlinear network traffic prediction based on BPneuralnetwork[J]. *Computer Applications*, 2007, 27(7): 1770-1772 (in Chinese).
- [19] OLIVERIRA T P, BARBAR J S, SOARES A S. Computer network traffic prediction: a comparison between traditional and deep learning neural networks[J]. *International Journal of Big Data Intelligence*, 2016, 3(1): 28-37.
- [20] VINAYAKUMAR R, SOMAN K P, POORNACHANDRAN P. Applying deep learning approaches for network traffic prediction[C]//International Conference on Advances in Computing, Communications and Informatics. Piscataway: IEEE Press, 2017: 2353-2358.
- [21] 杜爽, 徐展琦, 马涛, 等. 基于神经网络模型的网络流量预测综述[J]. *无线电通信技术*, 2020, 46(2): 216-222.
- DU S, XU Z Q, MA T, et al. A survey of network traffic prediction based on neural network models[J]. *Radio Communications Technology*, 2020, 46(2): 216-222(in Chinese).
- [22] 郭佳, 余永斌, 杨晨阳. 基于全注意力机制的多步网络流量预测[J]. *信号处理*, 2019, 35(5): 758-767.
- GUO J, YU Y B, YANG C Y. Multi-step prediction of traffic load with all-attention mechanism[J]. *Journal of Signal Processing*, 2019, 35(5): 758-767(in Chinese).
- [23] LI M, WANG Y W, WANG Z W, et al. A deep learning method based on an attention mechanism for wireless network traffic prediction[J]. *Ad Hoc networks*, 2020, 107: 102258.
- [24] CHO K, MERRIENBOER B V, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]//Proceedings of the 2014 Conference on Empirical Methods in natural Language Processing, 2014: 1724-1734.

# Improved intelligent detection algorithm for SPMA protocol channel state based on recurrent neural network

ZHANG Yanhui<sup>1</sup>, LYU Na<sup>1,\*</sup>, MIAO Jingcheng<sup>1</sup>, GAO Qi<sup>1</sup>, WANG Xiang<sup>1</sup>, CHEN Zhuo<sup>2</sup>

(1. Information and Navigation College, Air Force Engineering University, Xi'an 710077, China;

2. The 94619 Unit of the PLA, Lu'an 237000, China)

**Abstract:** By connecting sensors and shooters, tactical target network technology (TTNT) can realize rapid detection, positioning and strike of time-sensitive targets in denial environments. To that end, real-time and reliable channel access is highly required for tactical information transmission. TTNT uses the statistical priority-based multiple access (SPMA) protocol, which periodically calculates the statistical average number of arrival traffic pulses, to estimate the current channel state and thus control the timing of tactical information access. However, methods based on statistical average are merely suitable for stationary traffic, and will lead to large error in channel state detection when the traffic is non-stationary. To solve this problem, the traffic prediction technology was adopted and an improved detection algorithm for SPMA protocol channel state based on recurrent neural network was proposed. Meanwhile, in order to accurately obtain the current channel state, the recurrent neural network was employed to learn the hidden characteristics of historical traffic data, and a traffic predictor was constructed to timely predict the number of traffic pulses arriving at an instant. Experiments showed that the results of communication state detection with our algorithm is more realistic, which can significantly reduce the false judgment rate of the channel state.

**Keywords:** statistical priority-based multiple access protocol; channel state detection; traffic prediction; recurrent neural network; tactical target network technology

**Received:** 2021-06-08; **Accepted:** 2021-08-29; **Published Online:** 2021-09-14 17:41

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20210914.1500.007.html](https://kns.cnki.net/kcms/detail/11.2625.V.20210914.1500.007.html)

**Foundation items:** National Natural Science Foundation of China (61703427); Natural Science Foundation of Shaanxi Province (2020JQ-483)

\* **Corresponding author.** E-mail: [1105406986@qq.com](mailto:1105406986@qq.com)

# 《北京航空航天大学学报》征稿简则

《北京航空航天大学学报》是北京航空航天大学主办的以航空航天科学技术为特色的综合性自然科学学术期刊(月刊)。本刊以反映航空航天领域研究成果与动态、促进学术交流、培养科技人才和推动科技成果向社会生产力转化为办刊宗旨。本刊为中国自然科学技术核心期刊,并被 Ei Compendex 等国内外权威文献检索系统收录。本刊向国内外公开发行人,为进一步提高办刊质量和所刊出文章的学术水平,特制定本简则。

## 1 论文作者及内容

1.1 本刊面向海内外所有学者。

1.2 主要刊载与航空航天科学技术有关的材料科学与工程、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、编委讨论后决定录用与否。

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3期 2023年3月

JOURNAL OF BEIJING UNIVERSITY OF  
AERONAUTICS AND ASTRONAUTICS  
(JBUAA)

(Monthly, Started in 1956)

Vol.49 No.3 March 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 Qinqing  
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



9 771001 596236