

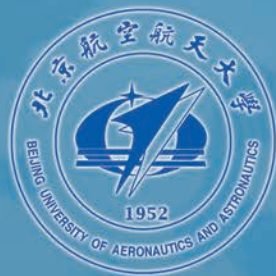


北航学报
赠阅

ISSN 1001-5965
CODEN BHHDE8

北京航空航天大学 学报

JOURNAL OF BEIJING UNIVERSITY OF
AERONAUTICS AND ASTRONAUTICS



2018-05

Vol.44 No.5

目 次

- 基于统计 α 算法的过程挖掘 余建波, 董晨阳, 李传锋, 程辉, 孙习武 (895)
- 基于形态学的自动驾驶仪振动信号基线漂移去噪 张景元, 何玉珠 (907)
- 微小空间内电磁干扰要素检测与辨识方法 凌波, 林贇, 孙宏涛, 苏东林 (914)
- 强噪声环境下自适应 CRPF 故障诊断方法 王进花, 曹洁, 李伟, 黄玲 (923)
- 基于 CPS 方法的工业机器人系统 陈友东, 常石磊, 冯强国 (931)
- 航空燃油柱塞泵滑靴副混合润滑特性数值仿真 符江锋, 李华聪, 曾夏青, 刘显为 (939)
- 采用 FRM 的窄过渡带奇型排列结构改进方法 张文旭, 赵文童, 陈涛, 陈亚静 (951)
- 基于欧拉壁面液膜模型的三维热气防冰腔数值仿真 李延, 郭涛, 常红亮 (959)
- 随机和区间变量共存条件下的模型确认指标 赵录峰, 吕震宙, 阚丽娟 (967)
- 基于系统 PSN 曲线的齿轮箱疲劳可靠度评估 马洪义, 谢里阳 (975)
- 基于自适应强跟踪 CQKF 的目标跟踪算法 刘畅, 杨锁昌, 汪连栋, 张宽桥 (982)
- 基于混合地图的护理机器人室内导航方法 张立志, 陈殿生, 刘维惠 (991)
- 基于 RS-TOPSIS 的空中目标威胁评估 杨远志, 于雷, 周中良, 阮铖巍 (1001)
- 两种典型低空风切变对火箭弹弹道特性的影响 陈健伟, 王良明, 李子杰 (1008)
- 脉冲多普勒引信抗周期调制干扰性能研究 刘少坤, 闫晓鹏, 栗革, 于洪海 (1018)
- 基于标校源辅助的不相交多目标到达时差定位 秦兆涛, 王俊, 陶磊岩, 魏少明 (1026)
- 直线驱动电静液作动器的匹配设计规则 王子蒙, 焦宗夏, 李兴鲁 (1037)
- 导航卫星双频复用信号模拟失真的仿真 蔚小龙, 寇艳红 (1048)
- 基于 PCE 的谐波减速器动态精度不确定性分析 张金洋, 张建国, 彭文胜, 刘育强, 汪龙 (1056)
- 硒掺杂锗碲相变存储材料的第一性原理研究 方治乾, 缪奶华, 周健 (1066)
- 一种基于深度学习的交互式电话号码识别方法 韩京冶, 许福, 陈志泊, 刘辉 (1074)
- 基于 Wiener 过程的发动机多阶段剩余寿命预测 黄亮, 刘君强, 贡英杰 (1081)
- 一种考虑共因失效的 PMS 可靠性建模分析方法 吴欢, 焦健, 赵廷弟 (1088)
- 多维力加载装置动力学建模及加载试验 侯立果, 王丹, 安大卫, 郭江真, 陈五一, 樊锐 (1095)
- 来流导致的高超声速气动热不确定度量化分析 张伟, 王小永, 于剑, 阎超 (1102)
- 龙卷风维持特性的探索 邸浩宇, 徐晶磊, 高歌 (1110)

期刊基本参数: CN 11-2625/V * 1956 * m * A4 * 222 * zh * P * ¥50.00 * 900 * 26 * 2018-05

(编辑 姜嘉 张嵘 李晶 张欣蔚 孙芳 王艳梅)

CONTENTS

Process mining based on statistical α -algorithm
..... YU Jianbo, DONG Chenyang, LI Chuanfeng, CHENG Hui, SUN Xiwu (895)

Removing baseline drift in vibration signal of autopilot based on morphology
..... ZHANG Jingyuan, HE Yuzhu (907)

Detection and identification method of electromagnetic interference elements in limited-space
..... LING Bo, LIN Yun, SUN Hongtao, SU Donglin (914)

An adaptive CRPF fault diagnosis method under strong noise condition
..... WANG Jinhua, CAO Jie, LI Wei, HUANG Ling (923)

Industrial robot system based on CPS approach
..... CHEN Youdong, CHANG Shilei, FENG Qiangguo (931)

Numerical simulation of hybrid lubrication characteristics of slipper pair of aviation fuel piston pump
..... FU Jiangfeng, LI Huacong, ZENG Xiaqing, LIU Xianwei (939)

An improved approach for odd-arrangement structure with narrow transition band based on FRM
..... ZHANG Wenxu, ZHAO Wentong, CHEN Tao, CHEN Yajing (951)

Numerical simulation of 3D hot-air anti-icing chamber based on Eulerian wall film model
..... LI Yan, GUO Tao, CHANG Hongliang (959)

A validation metric for model with mixture of random and interval variables
..... ZHAO Lufeng, LYU Zhenzhou, KAN Lijuan (967)

Assessment of gearbox fatigue reliability based on system PSN curve
..... MA Hongyi, XIE Liyang (975)

Target tracking algorithm based on adaptive strong tracking CQKF
..... LIU Chang, YANG Suochang, WANG Liandong, ZHANG Kuanqiao (982)

Care robot indoor navigation method based on hybrid map
..... ZHANG Lizhi, CHEN Diansheng, LIU Weihui (991)

Air target threat evaluation based on RS-TOPSIS
..... YANG Yuanzhi, YU Lei, ZHOU Zhongliang, RUAN Chengwei (1001)

Influence of two typical kinds of low-level wind shear on ballistic performance of rockets
..... CHEN Jianwei, WANG Liangming, LI Zijie (1008)

Anti-periodic modulation jamming performance of pulse Doppler fuze
..... LIU Shaokun, YAN Xiaopeng, LI Ping, YU Honghai (1018)

TDOA localization of multiple disjoint sources based on a calibration emitter
..... QIN Zhaotao, WANG Jun, TAO Leiyang, WEI Shaoming (1026)

Matching design rules of linear-driven electro-hydrostatic actuator
..... WANG Zimeng, JIAO Zongxia, LI Xinglu (1037)

Simulation of analog distortion of dual-frequency multiplexing signal generated by navigation satellite
..... YU Xiaolong, KOU Yanhong (1048)

Dynamic accuracy uncertainty analysis of harmonic reducer based on PCE
..... ZHANG Jinyang, ZHANG Jianguo, PENG Wensheng, LIU Yuqiang, WANG Long (1056)

First-principles study of Se doped GeTe phase-change material
..... FANG Zhiqian, MIAO Naihua, ZHOU Jian (1066)

A deep learning based interactive recognition method for telephone numbers
..... HAN Jingye, XU Fu, CHEN Zhibo, LIU Hui (1074)

Multi-phase residual life prediction of engines based on Wiener process
..... HUANG Liang, LIU Junqiang, GONG Yingjie (1081)

A reliability modeling and analysis method for PMS considering common cause failure
..... WU Huan, JIAO Jian, ZHAO Tingdi (1088)

Dynamic modeling and loading experiment of multi-dimensional loading device
..... HOU Liguang, WANG Dan, AN Dawei, GUO Jiangzhen, CHEN Wuyi, FAN Rui (1095)

Uncertainty quantification analysis in hypersonic aerothermodynamics due to freestream
..... ZHANG Wei, WANG Xiaoyong, YU Jian, YAN Chao (1102)

Exploration of tornado maintenance features
..... DI Haoyu, XU Jinglei, GAO Ge (1110)

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

DOI: 10.13700/j.bh.1001-5965.2017.0320

基于统计 α 算法的过程挖掘



余建波^{1,*}, 董晨阳¹, 李传锋¹, 程辉², 孙习武²

(1. 同济大学 机械与能源工程学院, 上海 201804; 2. 上海航天设备制造总厂, 上海 200245)

摘 要: 过程挖掘算法是从管理信息系统产生的事件日志中提取信息、发现知识并实现 workflow 建模的工具,也是目前 workflow 最主要的建模工具。然而现有的过程挖掘算法存在准确度较低、运行时间长和拟合度过高等问题,影响最终 workflow 模型的准确率。提出了一种基于统计 α 算法的过程挖掘算法,在保证算法较高的准确率和合适的拟合度的同时,降低算法运行时间,保证了算法的效率。首先,提出了重名活动识别算法,作为过程挖掘的预处理活动,提高了算法的准确性;其次,提出了统计 α 算法作为过程挖掘核心算法,有效消除了事件日志中噪声的影响;最后,提出了新的非自由选择结构识别算法,进一步提高了算法的鲁棒性和准确率。通过仿真实验和真实案例验证了该算法在准确率和运行时间上的优越性。

关 键 词: workflow 建模; 过程挖掘; 统计 α 算法; 重名活动; 非自由选择结构

中图分类号: TP311.5

文献标识码: A **文章编号:** 1001-5965(2018)05-0895-12

workflow 技术逐渐成为大型企业管理业务流
程的主要工具,其中 workflow 建模是 workflow 管理的基础,也是决定业务流程管理是否科学的先决条件,因此需要利用过程挖掘思想进行 workflow 建模,确保业务流程模型的有效性和准确性。过程挖掘是从事件日志中发现过程、控制、数据、组织和社会结构等信息的重要工具,该工具有助于分析和改进已经实施的业务过程,是实现 workflow 建模的重要工具^[1]。同时过程挖掘也是实现 workflow 过程建模的基础,是 workflow 管理的最基本工作,也是 workflow 分析的前提条件^[2]。

过程挖掘的基本思想由 Cook 和 Wolf^[3] 提出,目标是从软件过程的事件日志中自动发现过程模型。Agrawal 等^[4] 最早将过程挖掘算法应用到 workflow 管理中,假设活动之间相互独立,利用活动之间的先后关系构造活动之间的依赖性。Pinter 和 Golani^[5] 对 Agrawal 等^[4] 的算法进行了拓

展,将活动的开始时间和完成时间分开,从而检测任务之间的并行关系。Herbst 和 Karagiannis^[6] 进一步考虑了事件日志中存在的重名活动对活动依赖关系识别的影响,在前人的基础上加入了重名活动识别的预处理过程,提出了 MerfeSeq、SplitSeq 和 SplitPar 3 个算法。Schimm^[7] 提出了一种以获取完备且最小过程模型为目标的算法,该模型需要产生日志中记录的全部活动,但是该模型在出现隐藏活动或者非自由选择结构时会出错。Greco 等^[8] 以挖掘过程模型的层次树为目标,这些模型从不同的抽象层次描述事件日志。目前,对过程挖掘算法研究贡献最大的是 van der Aalst 教授及其团队。2004 年, van der Aalst 等^[9] 首先提出了基于 workflow 网络 (Workflow net, WF-net) 行为推理的 α 算法,可以在事件日志完备的情况下,发现合理的 SWF-net (Structural Workflow net)。为了解决 workflow 中存在的各种特殊结构的

收稿日期: 2017-05-15; 录用日期: 2017-09-15; 网络出版时间: 2017-10-26 09:13

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20171026.0913.001.html

基金项目: 国家自然科学基金 (51375290, 71777173); 上海市航天科技创新基金 (SAST2015054); 中央高校基本科研业务费专项资金 (22120180068)

* 通信作者. E-mail: jbyu@tongji.edu.cn

引用格式: 余建波, 董晨阳, 李传锋, 等. 基于统计 α 算法的过程挖掘[J]. 北京航空航天大学学报, 2018, 44(5): 895-906.

YU J B, DONG C Y, LI C F, et al. Process mining based on statistical α -algorithm[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 895-906 (in Chinese).

问题, α 算法发展成为一系列的算法—— α 系列算法。 $\alpha +$ 算法^[10] 在 α 算法的基础上, 考虑了短循环结构的挖掘, 可以成功发现事件日志中的循环结构, 得到一个合理的 SWF-net。tsinghua- α 算法^[11] 考虑了事件日志中的活动不为原子活动, 即每个活动都由一个 START 和一个 COMPLETION 活动组成, 用活动之间的重叠关系修正了 α 算法中的规则, 使得整个算法更加严谨, 但是该算法对事件日志的要求较高, 实用性并不高。 $\alpha + +$ 算法^[12] 是在 $\alpha +$ 算法基础上的拓展, 考虑了事件日志中的非自由选择结构, 该算法提出了 3 类活动间接依赖关系的发掘方法, 可以发现日志中存在的所有非自由选择结构, 不足的是该算法会挖掘太多冗余关系, 出现较严重的过拟合情况。 $\alpha \#$ 算法^[13] 是在 $\alpha +$ 算法上的另一拓展, 从挖掘隐藏活动入手, 用隐藏活动和不可见任务作为非自由选择结构的构成要素。 $\alpha *$ 算法^[14] 是在经典 α 算法上的拓展, 提出了可以发现事件日志中重名活动的规则, 该算法作为 α 算法的预处理步骤, 可以识别日志中命名不同但内容重复的活动, 使得 α 算法的准确率明显提高。 α 系列算法在工作流模型基本结构和特殊结构的识别上有较高的准确率, 但是其无法消除噪声的干扰, 在对含有噪声的事件日志进行挖掘时, 准确率会明显下降。

为了消除事件日志中的噪声影响, van der Aalst 等^[15] 提出了 Petri 网模型同活动因果矩阵的关系, 将形式化的 Petri 网转化成数学化的矩阵模型, 进而引入遗传算法对工作流模型进行迭代和进化, 从而可以消除日志中的噪声干扰, 同时得到较为完善的结论。李莉等^[16] 沿用活动因果矩阵的思想, 将变异粒子群算法与过程挖掘融合起来, 增强算法跳出局部最优的能力, 提高了算法的运行速度。宋炜和刘强^[17] 同样在因果矩阵的启发下, 采用模拟退火算法实现过程挖掘, 实现了对工作流中基本结构和特殊结构的挖掘。黄黎等^[18] 在李莉等^[16] 的基础上定义基于时序行为的老化因子, 并引入高斯变异的多种群协作的自适应策略, 对粒子群算法进行改进, 提高了算法的稳定性和收敛速度。以上都是基于进化算法的过程挖掘算法, 这些算法在准确率上有很好的表现, 但是当初始因果矩阵比较复杂, 即工作流长度较大时, 迭代算子太大导致算法运行时间很长, 算法效率较低。

本文提出了一种基于统计 α 算法的过程挖掘算法, 旨在同时满足高准确性和高效率 2 个目标, 将 α 系列算法和进化算法的优势结合起来。本文主要贡献点如下: ①提出了一种重名活动识

别和处理方法, 解决了 α 系列算法目前没有解决的重名活动处理, 将重名活动识别和处理作为 α 算法的预处理活动, 消除重名活动对于依赖关系识别的影响, 显著提高了算法的准确率。②提出了一种统计 α 算法, 通过概率计算有效消除了事件日志中的噪声影响, 明显提高了算法的准确率, 并保证了算法的运行效率。③考虑了工作流中常常出现的非自由选择结构, 提出一个新的识别非自由选择结构的算法, 在尽量少的添加间接依赖关系的前提下, 保证了非自由选择结构的识别率和准确性。

1 统计 α 算法体系

本文提出的基于统计 α 算法的过程挖掘算法方案如图 1 所示, 主要包含 2 个模块: 预处理模块和过程挖掘模块。其中, 预处理模块主要包括完备事件日志的提取、循环结构识别与处理、重名活动识别与处理 3 个步骤。首先根据完备事件日志的定义从数据库中挑选符合条件的事件日志作为过程挖掘的基础; 然后采用文献[19]中对于循环结构的识别规则, 挖掘出长度不同的循环结构, 并将循环结构体用一个活动集合来代替, 从而消除循环结构对于重名活动的影响; 最后利用本文提出的重名活动识别算法, 将事件日志中的重名活动进行分类识别和处理, 以消除重名活动对于活动依赖关系判断的影响。在过程挖掘模块中, 首先给出活动依赖关系定义, 并说明直接依赖与间接依赖之间的差异性; 然后在此基础上给出统计 α 算法的详细算法思路; 最后在统计 α 算法得到的直接依赖关系的基础上, 进一步对非自由选择结构中产生的间接依赖进行挖掘, 从而完善活动间的依赖关系得到活动关系矩阵。在得到最终的活动关系矩阵之后, 根据因果矩阵的构造规则, 可以将活动关系矩阵转化成果因果矩阵进而得到对应的 WF-net^[15]。

1.1 依赖关系与完备事件日志

事件日志是工作流执行过程中产生的重要信息, 在事件日志中往往会记录工作流中每个活动的名称、开始时间和执行人等信息。考虑到事件日志中往往会缺失活动的结束时间, 每个活动的持续时间并不是一个可以直接从事件日志中得到的信息, 很难在 WF-net 中引入持续时间的限定, 赋时 WF-net 无法从目前的事件日志中得到。因此, 本文提出的过程挖掘算法重点在于活动之间依赖关系的挖掘, 通过构造活动关系矩阵, 间接得到因果矩阵, 进而生成 WF-net。

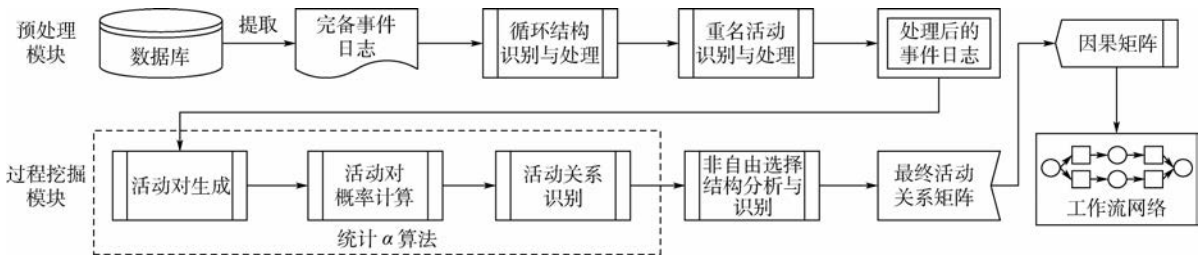


图1 基于统计 α 算法的过程挖掘方案

Fig.1 Scheme of process mining based on statistical α -algorithm

本文挖掘的活动依赖关系分为直接依赖和间接依赖2个方面。若有活动集合 T ,存在事件日志 $W \subseteq T$,有 $a, b \in T$,若 a 和 b 有依赖关系:

1) 若存在 $a \cdot \cap \cdot b \neq \emptyset \vee a \cap b \cdot \neq \emptyset$,则称 a 和 b 存在直接依赖。直接依赖关系主要有表1中4种关系。

2) 若不存在 $a \cdot \cap \cdot b \neq \emptyset \vee a \cap b \cdot \neq \emptyset$,则称 a 和 b 存在间接依赖。间接依赖关系主要指间接因果关系,用记号 $a \mapsto_w b$ 表示。

表1 直接依赖关系分类与定义

Table 1 Classification and definition of direct relations

直接依赖关系分类	符号	定义
顺序关系	$a >_w b$	$a, b \in \sigma = \{\dots, a, b, \dots\}$
因果关系	$a \rightarrow_w b$	$a >_w b \wedge b \not>_w a$
选择关系	$a \#_w b$	$a \not>_w b \wedge b \not>_w a$
并行关系	$a \parallel_w b$	$a >_w b \wedge b >_w a$

由直接依赖关系定义可以看到,过程挖掘算法的挖掘结果同事件日志的完备性有着重要关系,不完备的事件日志往往得到错误的结果。因此,在使用过程挖掘算法之前需要给出完备事件日志的定义。

定义1(完备事件日志) 令 T 为活动集合,有事件日志 $W \subseteq T$, W 为完备事件日志当且仅当:
①对于任意日志 $W' \subseteq T$,满足 $>_{w'} \subseteq >_w \wedge \rightarrow_{w'} \subseteq \rightarrow_w \wedge \#_{w'} \subseteq \#_w \wedge \parallel_{w'} \subseteq \parallel_w$;②对于任意 $t \in T$,满足 $\exists \sigma \in W: t \in \sigma$ 。

完备事件日志是判断活动间依赖关系的前提条件,日志的完备性要求每种可能出现的流程轨迹至少出现一次。本文中所有算法都是基于完备事件日志基础之上的,在当今大数据的环境中,由于数据量足够大,可以认为拿到的事件日志都是完备的。

1.2 重名活动识别算法

重名活动是过程挖掘中一种常见的特殊结构,表2给出了某临床路径工作流的事件日志部分内容,活动 t_{25} 和 t_{45} 记录着同样的活动“医患交流”,记 t_{25} 和 t_{45} 为重名活动。为了区分同一条

表2 某临床路径事件日志部分内容

Table 2 Event log of a clinical pathway (part)

活动序号	活动名称
:	:
25	医患交流
26	手术室准备
27	药物准备
28	手术工具准备
29	最后确认
:	手术阶段
41	术后整理
42	送病人回病房
43	送水送药
44	病人休息
45	医患交流
46	病情讨论
:	:

流程轨迹中的重名活动,用记号 $\delta_i(A, n_i)$ 表示在流程轨迹 δ_i 中第 n_i 个活动 A 。然而通过对于事件日志内容的具体分析可以发现, t_{25} 所表示的“医患交流”指的是手术前医患对于手术的沟通, t_{45} 显然代表了术后医生对于患者的询问。虽然 t_{25} 和 t_{45} 具有相同的文字描述,但是两者的实际意义并不相同。如果将两者当作重名活动来判断会导致活动依赖关系判别出现错误,因此需要将重名活动判别作为过程挖掘的预处理步骤。

循环结构同样是工作中常见的特殊结构,多出现在发生资源等待或者返工的情况下。当循环结构出现时,必然会出现大量重名活动接连出现,为了简化重名活动判别过程,本文使用文献[19]中给出的对于循环结构的识别作为预处理,以消除循环结构对于重名活动判别的影响。

本文对于重名活动的判别基于重名活动之上,考虑那些命名相同的活动是否具有含义也相同。将重名活动根据其具体活动内容是否相同分为2种:活动内容相同的称为重复性重名活动,简称重复活动(Duplicate Activities, DA);活动内容不同的称为非重复性重名活动(Homononyms Activities, HA)。为了探索重名活动在何时为重复活

动,本文以活动组(activities group)为单位,给出基于活动组的重复活动的定义。

定义2(活动组) 对于任意流程轨迹 $\delta_i \in W$, 如果 δ_i 中活动 $A \in \delta_i$, A 的2个前驱活动 T_p, T_{pp} 和2个后继活动 T_s, T_{ss} , 则将 $\{T_{pp}, T_p, A, T_s, T_{ss}\}$ 这样一个有序集合记作一个活动组, 记作 G_A 。如果 δ_i 中存在重名活动 $\delta_i(A, n), \delta_i(A, m) \in \delta_i$, 且两者对应的活动组分别为 G_A 和 G'_A , 若 G_A 和 G'_A 中包含的除了 A 之外的元素相同, 但元素排列顺序不同, 则称为2个活动组相等, 记作 $G_A = G'_A$ 。

定义3(重复活动) 对于任意流程轨迹 $\delta_i \in W$, 如果 δ_i 中存在重名活动 $\delta_i(A, n), \delta_i(A, m) \in \delta_i$, 且两者对应的活动组分别为 G_A 和 G'_A , 若 G_A 和 G'_A 中除了活动 A 之外的其他活动对应相等, 则活动 $\delta_i(A, n)$ 和活动 $\delta_i(A, m)$ 可以定义为重复活动。

对于事件日志 W 中的某条流程轨迹 $\delta_i \in W$, 如果在 δ_i 中存在重名活动 $\delta_i(A, n), \delta_i(A, m) \in \delta_i$, T_p 和 T_s 分别为活动 $\delta_i(A, n)$ 的前驱任务和后继任务, T_{pp} 为 T_p 的前驱活动, T_{ss} 为 T_s 的后继活动; T'_p 和 T'_s 分别为活动 $\delta_i(A, m)$ 的前驱任务和后继任务, T'_{pp} 为 T'_p 的前驱活动, T'_{ss} 为 T'_s 的后继活动。活动 $\delta_i(A, n)$ 和活动 $\delta_i(A, m)$ 对应的活动组分别为 G_A 和 G'_A , U 为重复活动集合。结合上述定义2和定义3, 给出重复活动判别规则如下:

- 1) 若 $G_A = G'_A, T_p = T'_p$ 且 $T_s = T'_s$, 则 $\langle \delta_i(A, n), \delta_i(A, m) \rangle \in U$ 。
- 2) 若 $G_A = G'_A, T_p = T'_p, T_s \neq T'_s, T_s = T'_{ss}$, 且 $T_{ss} = T'_s$, 则 $\langle \delta_i(A, n), \delta_i(A, m) \rangle \in U$ 。
- 3) 若 $G_A = G'_A, T_p \neq T'_p, T_s = T'_s, T_p = T'_{pp}$, 且 $T_{pp} = T'_p$, 则 $\langle \delta_i(A, n), \delta_i(A, m) \rangle \in U$ 。
- 4) 若 $G_A = G'_A, T_p \neq T'_p, T_s \neq T'_s, T_p = T'_{pp}, T_{pp} = T'_p, T_s = T'_{ss}$, 且 $T_{ss} = T'_s$, 则 $\langle \delta_i(A, n), \delta_i(A, m) \rangle \in U$ 。

5) 若 $G_A \neq G'_A, T_p = T'_p, T_s \neq T'_s$ 且 $T_s \#_W T'_s$, 则 $\langle \delta_i(A, n), \delta_i(A, m) \rangle \in U$ 。

6) 若 $G_A \neq G'_A, T_p \neq T'_p, T_s = T'_s$ 且 $T_p \#_W T'_p$, 则 $\langle \delta_i(A, n), \delta_i(A, m) \rangle \in U$ 。

7) 若 $G_A \neq G'_A, T_p \neq T'_p, T_p \#_W T'_p, T_s \neq T'_s$ 且 $T_s \#_W T'_s$, 则 $\langle \delta_i(A, n), \delta_i(A, m) \rangle \in U$ 。

8) 若 $G_A \neq G'_A, T_p \neq T'_p, T_p \#_W T'_p, T_s \neq T'_s, T_s = T'_{ss}$, 且 $T_{ss} = T'_s$, 则 $\langle \delta_i(A, n), \delta_i(A, m) \rangle \in U$ 。

9) 若 $G_A \neq G'_A, T_s \neq T'_s, T_s \#_W T'_s, T_p \neq T'_p, T_p = T'_{pp}$, 且 $T_{pp} = T'_p$, 则 $\langle \delta_i(A, n), \delta_i(A, m) \rangle \in U$ 。

1.3 非自由选择结构识别与分析

非自由选择结构是 workflow 中又一种常见的特殊结构, 通常情况下非自由选择结构由间接因果依赖关系导致。不考虑循环结构的影响, 非自由选择结构可分为如图2所示的三大类: 局部非自由选择、全局非自由选择和并行非自由选择。在图2(a)中, 对于活动 E 和 F 的选择取决于其之前对活动 B 和 C 的选择, 2个选择包含在同一个并行活动中, 选择的结果总能最终到达库所 O , 不影响后续工作流, 因此称为局部非自由选择。在图2(b)中, 对于活动 D 和 E 的选择取决于对于活动 A 和 B 的选择, 这2个选择结构并不在同一个结构之下, 更可能中间隔着若干个活动, 因此称为全局非自由选择。在图2(c)中, 若不考虑间接依赖, 很容易忽略活动 D 和 G 之间的依赖关系, 而此时用于并行关系的存在, 如果在活动 B 和 D 之间选择了 D , 则会发生 workflow 网在 F 处发生死锁, 此情况不同于图2(a), 称为并行非自由选择。

通过以上分析可以发现: 非自由选择结构的产生总是由另一个选择结构导致的, 前一个选择结构对于活动的选择往往决定了后面某个选择结构的结果。因此为了探索非自由选择结构, 需要考察各个选择结构之间的关联, 探索两者之间可

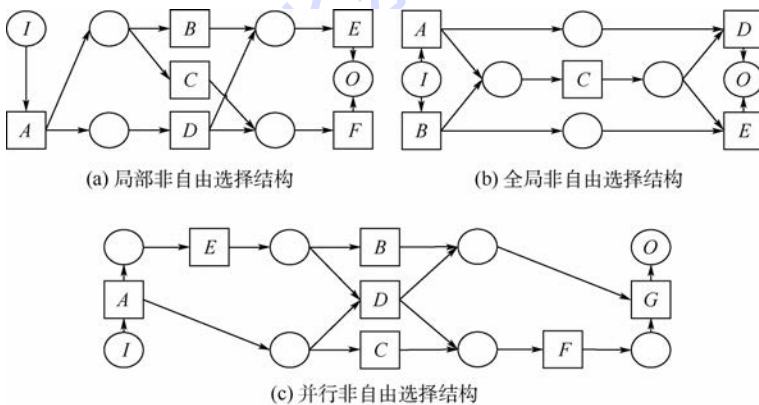


图2 非自由选择结构分类

Fig. 2 Classification of non-free-choice constructs

能存在的依赖关系,为此本文首先给出几个活动关系定义。

定义 4(活动关系补充) 令 $N = (P, T, F)$ 为一个不包含循环结构的工作流网络, F 为流关系, W 为工作流事件日志, T 为活动集合, 其中 $a, b \in T$:

- 1) $a \triangleleft_w b$, 当且仅当 $a \#_w b$, 且 $\exists c \in T, c \rightarrow_w a, c \rightarrow_w b$ 。
- 2) $a \triangleright_w b$, 当且仅当 $a \#_w b$, 且 $\exists c \in T, a \rightarrow_w c, b \rightarrow_w c$ 。
- 3) $a \gg_w b$, 当且仅当 $a \not\#_w b$, 对于流程轨迹 $\sigma = \{t_1, t_2, \dots, t_n\}$, 有 $i, j \in \{1, 2, \dots, n\}$ 满足 $i < j, t_i = a, t_j = b$, 对于任意 $k \in \{i + 1, \dots, j - 1\}$, 满足 $t_k \neq a \wedge t_k \neq b \wedge \neg (t_k \triangleleft_w a \text{ or } t_k \triangleright_w a)$ 。

基于定义 4, 在不包含循环结构的情况下, 给出间接依赖 (\mapsto_w) 判断规则: 令 W 为完备事件日志, $N = (P, T, F) = \alpha(W)$ 是为利用统计 α 算法对事件日志 W 进行挖掘得到的 WF-net, 在 N 中有 2 次选择关系 $a \#_w b, x \#_w y$, 且选择关系 $a \#_w b$ 不发生在 $x \#_w y$ 之后:

规则 1 若 $a \gg_w x \wedge b \gg_w y, \neg (a \gg_w y \vee b \gg_w x)$, 则 $a \mapsto_w x \wedge b \mapsto_w y$ 。

规则 2 若 $\neg (a \gg_w x) \wedge (a \gg_w y) \wedge \neg (b \gg_w x) \wedge \neg (b \gg_w y), a \triangleright_w x \wedge b \triangleright_w y, \neg (a \triangleright_w y) \wedge \neg (b \triangleright_w x)$, 则 $a \mapsto_w x \wedge b \mapsto_w y$ 。

规则 3 若 $\neg (a \gg_w x) \wedge \neg (a \gg_w y) \wedge \neg (b \gg_w x) \wedge \neg (b \gg_w y), \exists p, t, p', t': a \gg_w p, b \gg_w t$ 且 $x \gg_w p', y \gg_w t'$, 有 $p \triangleright_w p' \wedge t \triangleright_w t' \wedge \neg (p \triangleright_w t') \wedge \neg (t \triangleright_w p')$, 则 $a \mapsto_w x \wedge b \mapsto_w y$ 。

在上述 3 条规则中, 规则 1 针对全局非自由选择结构, 如图 3(a) 所示, 选择结构 $a \#_w b, x \#_w y$ 在时间上存在着一定的前后顺序, 即存在 $a \gg_w x \wedge b \gg_w y$, 但是在事件日志却不存在 $a \gg_w y \vee b \gg_w x$, 说明对于 x, y 的选择由选择结构 $a \#_w b$ 决定, 因此需要加入间接依赖关系 $a \mapsto_w x \wedge b \mapsto_w y$ (见图 3(b))。规则 2、规则 3 针对局部和并行非自由选择结构, 如图 3(c)、(e) 所示, 2 个选择结构 $a \#_w b, x \#_w y$ 在时间上不存在前后顺序, 4 个活动的发生次序不确定, 此时考察选择关系之后的活动。若存在 $a \triangleright_w x \wedge b \triangleright_w y$, 且 $\neg (a \triangleright_w y) \wedge \neg (b \triangleright_w x)$, 此时 2 个选择结构均为非自由选择, 活动 a 和 x 必须并行发生, 或者活动 b 和 y 并行发生, 因此活动 a 和 x, b 和 y 相互之间存在间接依赖关系 (见图 3(d))。若不存在 $a \triangleright_w x \vee b \triangleright_w y$, 考察 2 个并行关系的后续活动, 若有活动 $p, t, p', t': a \gg_w p, b \gg_w t, x \gg_w p', y \gg_w t'$, 满足规则 2 中的条件: $p \triangleright_w p' \wedge t \triangleright_w t' \wedge \neg (p \triangleright_w t') \wedge \neg (t \triangleright_w p')$, 则同样有 $a \mapsto_w x \wedge b \mapsto_w y$ 。

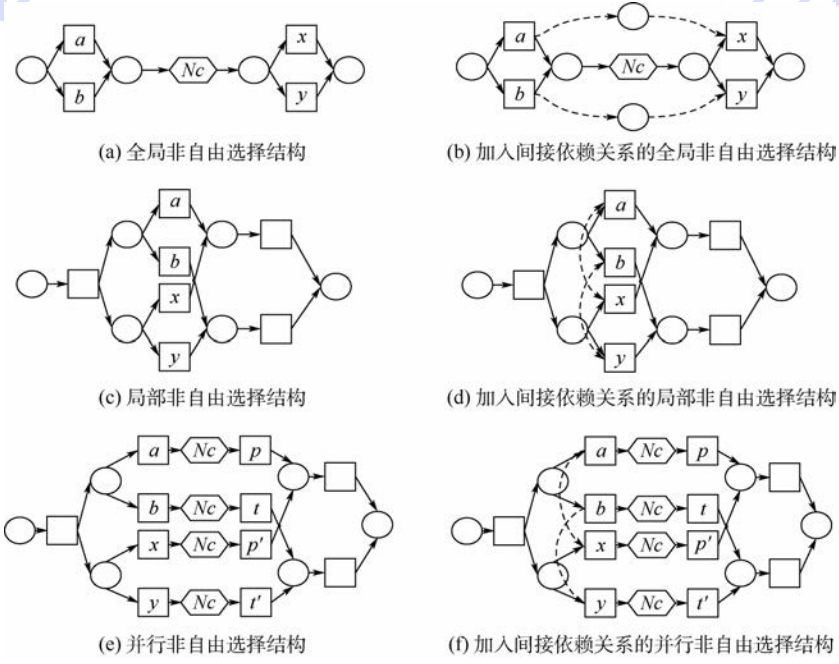


图 3 间接依赖判断规则

Fig. 3 Rules to judge indirect relations

1.4 统计 α 算法

由于经典 α 系列算法要求事件日志中不包含噪声, 同时要求事件日志完备。而在如今数据

量日益增大的条件下, 完备事件日志不再是一个难以达到的条件, 而事件日志中的噪声却逐渐成为无法控制的干扰因素。为了解决噪声的干

扰,有学者提出将 α 算法和遗传算法等进化算法结合起来,将因果矩阵作为迭代算法的基本算子,尽管该方法在消除噪声方面有较好的效果,但是却由于变异、交叉次数过多导致算法运行时间较长,算法效率较低。本文提出的统计 α 算法借用了概率计算的思想来消除噪声干扰提高算法准确率,同时算法的运行时间也可以维持在较低的水平。本文提出的基于统计 α 算法的过程挖掘具体方案如图4所示。

本文提出的统计 α 算法以活动对为单位,通过每2个活动之间的活动关系进一步分析整个工作流程中所有活动间的相关关系。首先给出活动对的定义。

定义5(活动对) 对于任意一个事件日志 W ,假设有流程轨迹 $\delta_i = \{\dots, T_i, T_{i+1}, \dots\} \subseteq W$ ($i=1,2,\dots$),流程轨迹中的元素按照执行时间的先后顺序排列,则活动对(Activities Couple, AC)可定义如下:

- 活动对 {
- FA,第1个活动名称(T_i);
- SA,第2个活动名称(T_{i+1});
- f ,发生概率;
- R ,活动关系=顺序关系(默认)
- }

活动对是一个具有多个属性的结构体。

定义6(同前活动对) 对于任意活动对 $AC(A,B)$,若存在一个活动对和它有相同的第1个活动 A ,第2个活动却不相同,则称之为 $AC(A,B)$ 的同前活动对,记作 $\#AC(A,B)$ 。

定义7(转置活动对) 对于任意活动对 $AC(A,B)$,若存在1个活动对的第1个活动等于它的第2个活动,第2个活动等于它的第1个活动,则称之为 $AC(A,B)$ 的转置活动对,记作 $-AC(A,B)$ 。

活动对中的发生概率 f 根据该活动对在事件日志中的出现次数计算得到。对于活动对

$AC(T_i, T_{i+1})$ 的概率:

$$f = \frac{AC(T_i, T_{i+1}) \cdot F}{N + n_r}$$

式中: $f=AC(T_i, T_{i+1}) \cdot F$ 表示该活动对的出现次数,需要遍历所有事件日志获得; N 为流程轨迹数; $n_r = n_r(T_i) + n_r(T_{i+1})$ 为活动 T_i 和活动 T_{i+1} 重名活动的数目总和。

在计算活动对概率之后,可以在此基础上进一步判别活动对的活动关系 R ,具体算法步骤如图5所示。

步骤1 单个活动对关系判别。本文对于活动关系的判别建立在每个活动对的关系之上,而对于单个活动对的关系判别则是基于经典 α 算法中的判别规则加以改进。具体算法如下。

输入:活动对集合 $Array(AC)$;显著性水平 α 。

输出:更新活动关系后的活动对集合。

```

Foreach(AC(a,b) in Array(AC))
    If(AC(a,b).F < alpha) {
        AC(a,b)为噪声数据;
        Array(AC).Remove(AC(a,b))
    }
    Else {
        - AC(a,b) = AC(b,a) // 转置活动对
        If(AC(b,a).F > alpha) {
            a ||_w b
        }
        Else {
            #AC(a,b) = AC(a,c) // 同前活动对
            If(AC(a,c).F < alpha) {
                a ->_w b
            }
            Else {
                a >_w b; a >_w c
                If(AC(b,c).F > alpha && AC(c,b).F > alpha) {
                    b ||_w c
                }
            }
            Else {
                b #_w c
            }
        }
    }
    
```

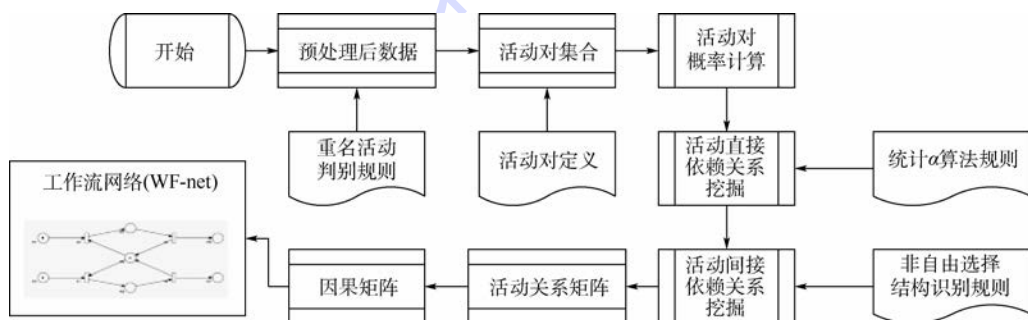


图4 基于统计 α 算法的过程挖掘步骤

Fig.4 Steps of process mining based on statistical α -algorithm

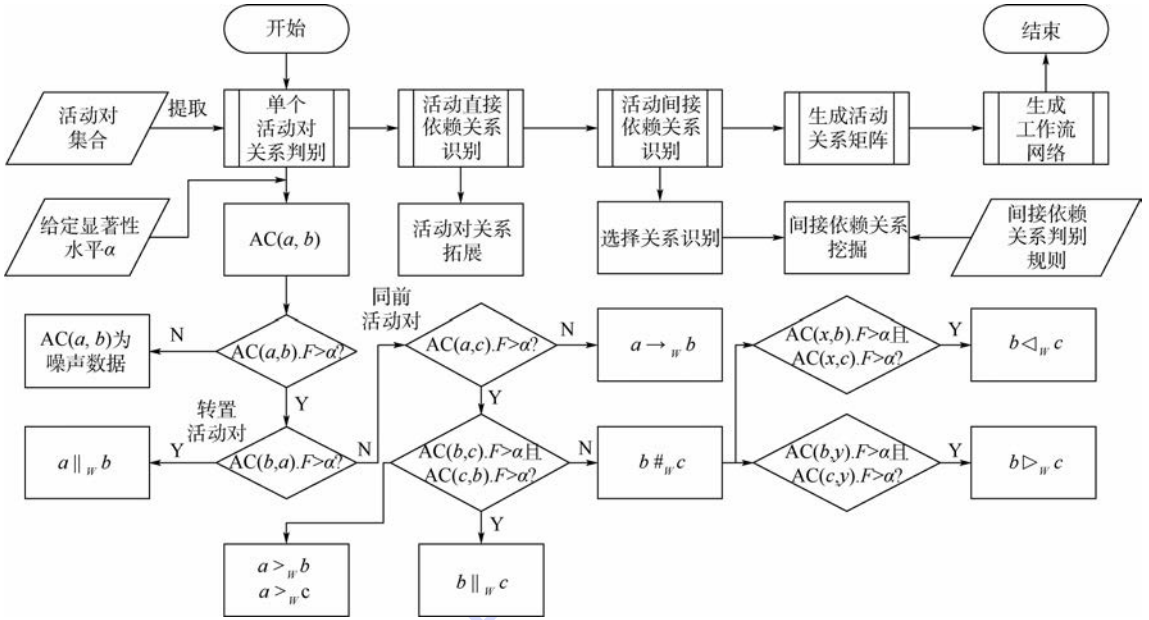


图5 统计 α 算法算法步骤

Fig. 5 Steps of statistical α -algorithm

$$\{ \text{If}(AC(x,b).F > \alpha \ \&\& \ AC(x,c).F > \alpha) \{$$

$$b < \#_w c \}$$

$$\text{If}(AC(b,y).F > \alpha \ \&\& \ AC(c,y).F > \alpha) \{$$

$$b > \#_w c \}$$

$$\} \}$$

作流网络。

2 实验与结果验证

本文采用仿真数据和真实案例分别对提出的算法进行有效性验证。由于本文提出的算法基于大量数据的基础之上,因此为了实验数据满足数量上的要求,本文以仿真数据进行实验为主,真实数据则是为了进一步对算法进行验证。仿真数据的产生原理在文献[20]的基础上加以改进,主要分为如下4个步骤:

步骤1 确定流程轨迹长度。流程轨迹长度是最影响算法运行效率的指标,而为了尽量多得出出现包含各种不同的活动关系,需要轨迹达到一定的长度。为了均衡两者之间的矛盾,本文将仿真数据分为3类,分别用于对重名活动、非自由选择结构和统计 α 算法进行验证。

步骤2 确定事件日志规模。事件日志规模主要指流程轨迹的数目,该数目需要保证足够多的数量以确保事件日志的完备性。本文考虑5种事件日志规模:200、400、600、800和1000条流程轨迹。

步骤3 确定噪声规模。噪声类型主要分为5种:①删除流程轨迹的某个活动;②调换流程轨迹中某2个活动;③随机替换流程轨迹中的某个活动;④随机的在流程轨迹添加一个随机活动;⑤人为地生成重名活动,并随机替换已有的活动。噪声规模按照文献[20]中的表述,噪声最少有一

步骤2 活动直接依赖关系识别。在得到了单个活动对关系之后,需要将活动间的关系进行整合以得到多个活动间的关系,这里给出几个活动关系整合规则:

1) 因果关系的传递性:若有 $a \rightarrow_w b \wedge b \rightarrow_w c$, 则 $a \rightarrow_w c$ 。

2) 并行关系的不确定性:当 $a \parallel_w b \wedge b \parallel_w c$ 时,若有 $a \parallel_w c$, 则 $a \parallel_w b \parallel_w c$; 若 $a \#_w c$, 则活动 $(a \parallel_w b) \#_w (b \parallel_w c)$, 2个选择关系共用同一个活动 b (或者重复活动 b)。

3) 选择关系的不确定性:当 $a \#_w b \wedge b \#_w c$ 时,若有 $a \#_w c$, 则 $a \#_w b \#_w c$; 若 $a \parallel_w c$, 则活动 $(a \#_w b) \parallel_w (b \#_w c)$, 2个选择关系共用同一个活动 b (或者重复活动 b), 这种情况经常出现非自由选择结构,需要进一步判别。

步骤3 活动间接依赖关系识别。在完成活动直接关系的识别后,按照上文的叙述,需要重新对选择关系活动进行考察,挖掘选择关系之间可能存在的间接依赖关系,以完善活动关系。

步骤4 生成工作流网络。将直接活动关系和间接活动关系进行整合,得到活动关系矩阵。将活动关系矩阵转化为因果矩阵,并据此生成工

个,最多不可超过轨迹长度的1/3,本文将噪声规模主要分为5%、10%和20%3种规模,以检测不同情况下统计 α 算法对于噪声的抗干扰性。

步骤4 确定活动执行的优先度。为了得到非自由选择结构,本文将对选择结构中的活动人为设置执行优先度,优先度越高,选择结构中该活动的发生几率越高。执行优先度等级设置为如下4种:

1) Level1:无优先级,所有活动有同样的执行优先度,设为0.5。

2) Level2:低优先级,活动执行优先度在0.4~0.6之间。

3) Level3:中优先级,活动执行优先度在0.25~0.75之间。

4) Level4:高优先级,活动执行优先度在0.01~0.09之间。

2.1 仿真实验与算法性能分析

本节为了多方面验证本文提出算法的有效性,将实验分为3个部分:重名活动识别算法验证实验、统计 α 算法验证实验和非自由选择结构识别算法验证实验。3个实验有各自的侧重点,因此对于实验数据的安排同样有所区别。

2.1.1 重名活动识别算法的验证实验

由于重名活动识别是在统计 α 算法之前,因此在重名活动识别中暂时不考虑活动执行的优先度等级,同时为了保证重名活动不被其他噪声

所干扰,噪声类型也只选择类型⑤,噪声规模按照5%、10%和20%3种来进行实验,实验数据安排如表3所示。该实验的评价指标有3个:识别率(是否能准确捕捉每一个重名活动)、准确率(重名活动中重复活动识别准确率)和运行效率(算法运行时间)。该实验的对比算法为Herbst和Karagiannis^[6]提出的重名活动识别算法。

得到的实验结果如表4所示。2个算法在识别率和准确率上一致,可能由于重名活动数量并不多的关系,两者在这2个方面没有任何差异。在运行时间上,本文提出的算法要优于文献[6]算法,因为本文在重名活动识别时,以活动组作为最小单位,减少了算法迭代和比较的内容,算法时间上有明显减少。

表3 重名活动识别算法验证实验数据安排

Table 3 Experimental data arrangement for cognominal activity identification algorithm verification

数据编号	流程轨迹长度	事件日志规模	噪声类型	噪声规模/%	优先度等级
C1	22	200	⑤	5	Level1
C2	22	200	⑤	10	Level1
C3	22	200	⑤	20	Level1
C4	42	200	⑤	5	Level1
C5	42	200	⑤	10	Level1
C6	42	200	⑤	20	Level1

表4 重名活动识别算法对比实验结果

Table 4 Comparison of experimental results of cognominal activity identification algorithm

数据编号	重名活动数量	本文算法			文献[6]算法		
		识别率/%	准确率/%	运行时间/ms	识别率/%	准确率/%	运行时间/ms
C1	2对	100	100	10.12	100	100	15.12
C2	3对	100	100	10.14	100	100	15.36
C3	5对	100	100	10.80	100	100	15.66
C4	3对	100	100	21.33	100	100	30.54
C5	5对	100	100	21.35	100	100	30.87
C6	9对	100	100	21.55	100	100	31.03

2.1.2 统计 α 算法的验证实验

由于非自由选择结构识别是在统计 α 算法之后进行,因此本节暂时不考虑非自由选择结构的影响,优先度等级统一设置为Level1。由于统计 α 算法在重名活动识别的基础之上,因此这里噪声类型设置为①~④。主要考虑的实验因素为事件日志规模和噪声规模,即一因素三水平和另一因素五水平的实验,具体实验数据安排如表5所示。该实验的衡量指标有3个:泛化性 g (本文用数据量的2/3作为算法训练样本,用剩下的1/3数据进行拟合,泛化性 $g = N_d / N_m + \text{noise}$, N_d 为

表5 统计 α 算法验证实验数据安排

Table 5 Experimental data arrangement for statistical α -algorithm verification

数据编号	流程轨迹长度	事件日志规模	噪声类型	噪声规模/%	优先度等级
S1~S5	42	200~1000	①~④	5	Level1
S6~S10	42	200~1000	①~④	10	Level1
S11~S15	42	200~1000	①~④	20	Level1

拟合数据中活动数目, N_m 为算法得到模型中的活动数目,noise表示噪声规模。 g 的值越接近1越好, g 的值大于1表示算法得到的模型中噪声过

多, g 的值太小表明算法中遗漏的活动太多)、准确率 a (活动关系准确率) 和运行时间 t 。该部分的对比算法为经典 α 算法^[9]、基于变异粒子群算法的过程挖掘算法 (HAC-PSO)^[16]。

实验结果如表 6 所示。为了分别比较事件日志规模和噪声规模对最终结果的影响, 本文采用控制变量的方法分别考察这 2 个因素对 3 个评价指标的影响。图 6 为 3 种算法在噪声规模同样为 5% 的情况下, 泛化性、准确率和运行时间 3 个指标的对比情况。可以看到, 在噪声规模相同的情况下, 统计 α 算法和 HAC-PSO 的泛化性和准确率表现差不多, 两者都是比较接近 100%。而经典 α 算法在泛化性上出现了明显的超出 100% 的现象, 说明受到了噪声的较大影响, 同样经典 α 算法的准确率也是低于另外 2 个算法 5% 左右, 正好是噪声规模的大小。而运行时间上, 经典 α 算法和统计 α 算法都维持在一个较低的水平, 但

是 HAC-PSO 的运行时间却远远高于这 2 种算法。在噪声规模为 10% 和 15% 时有着类似的结果。

图 7 给出了事件日志规模为 600, 噪声规模不同情况下, 泛化性、准确率和运行时间 3 个指标的对比情况。可以看到, 在噪声规模增加的情况下, 统计 α 算法和 HAC-PSO 可以在泛化性和准确率上保持较好的水平, 而经典 α 算法随着噪声规模的增加, 泛化性快速增加, 并且超过 100%, 准确率直线下降, 可见经典 α 算法受噪声的影响十分大。而随着噪声规模增加, 统计 α 算法运行时间也有小幅增加, 而经典 α 算法则基本保持不变, HAC-PSO 的运行时间远超过前 2 种算法。

总结来说, 统计 α 算法在泛化性和准确率上要明显优于经典 α 算法, 与 HAC-PSO 基本持平。但是在算法运行时间上却明显优于 HAC-PSO, 基本与经典 α 算法持平, 兼具了两者的优点, 实现

表 6 统计 α 算法对比实验结果

Table 6 Comparison of experimental results of statistical α -algorithm

数据编号	本文算法 ($\alpha = 0.05$)			经典 α 算法 ^[9]			HAC-PSO ^[16]		
	泛化性/%	准确率/%	运行时间/ms	泛化性/%	准确率/%	运行时间/ms	泛化性/%	准确率/%	运行时间/ms
S1	99.10	97.98	3 350	105.00	92.98	2 289	99.65	98.56	12 651
S2	99.16	98.88	5 830	104.90	93.86	4 820	99.58	98.96	33 624
S3	99.52	98.82	7 316	104.88	93.82	9 213	99.58	99.19	75 623
S4	99.36	99.25	9 833	104.95	94.22	15 632	99.69	99.69	136 362
S5	99.88	99.65	11 369	104.36	94.65	30 205	99.78	99.96	206 534
S6	101.02	97.23	4 069	111.02	86.88	2 441	99.16	98.03	12 543
S7	100.56	97.79	6 151	110.56	87.25	4 922	99.23	98.15	34 123
S8	100.23	98.56	8 536	110.23	88.36	9 365	99.26	98.45	76 245
S9	100.20	98.33	10 623	110.20	89.12	15 664	99.32	98.32	132 013
S10	99.93	99.21	12 988	109.93	89.35	30 157	99.49	98.98	214 756
S11	100.98	97.56	4 419	114.98	82.36	2 674	98.52	97.23	13 025
S12	100.56	97.88	6 496	114.56	83.65	5 155	98.62	97.53	34 216
S13	100.23	98.65	9 441	114.43	82.98	9 998	99.15	98.68	77 487
S14	100.59	98.98	14 635	115.09	84.98	17 897	99.48	99.32	140 259
S15	100.20	99.52	19 062	114.35	83.66	32 390	99.60	99.56	219 874

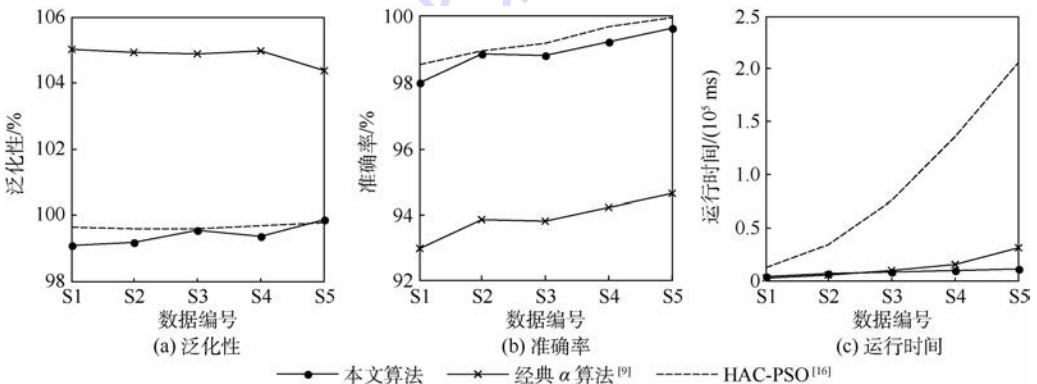


图 6 相同噪声规模下的算法结果对比

Fig. 6 Comparison of algorithm results under the same noise scale

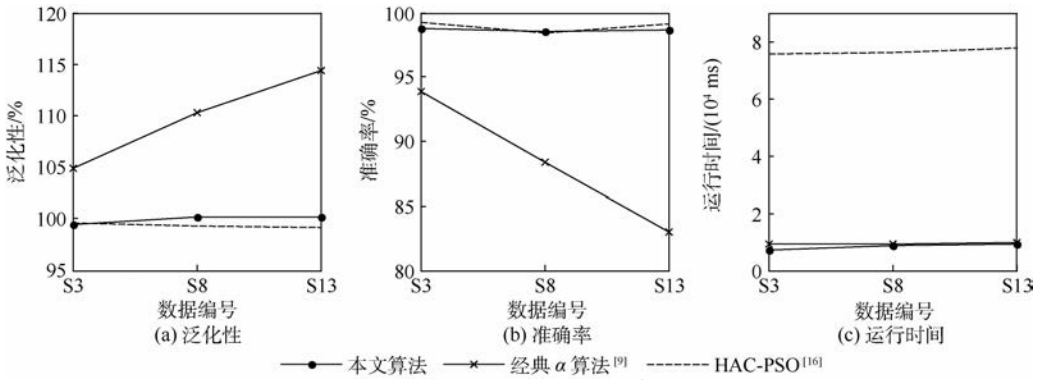


图7 相同事件日志规模下的算法结果对比

Fig. 7 Comparison of algorithm results under the same event log scale

了算法的高准确率和高效率。

2.1.3 非自由选择结构识别算法的验证实验

非自由选择结构在统计 α 算法的基础是在噪声消除的基础之上进行的,因此这里不再考虑噪声对算法的影响。为了使非自由选择结构发生几率足够大,活动执行优先级等级直接设置为Level4,具体实验数据安排如表7所示。该算法的衡量指标有3个:识别率(是否能识别每一个非自由选择结构,本文利用识别数目来表达识别率)、准确率(识别出来的非自由选择结构是否准确)、间接依赖数目(间接依赖数目过少会导致非自由选择结构的遗漏,过多会导致模型过于冗余, workflow运行效率低。当准确率和识别率达到较高水平时,间接依赖数目越少越好)和运行时间(不包括统计 α 算法的时间)。该算法的对比算法为 $\alpha++$ 算法。

实验结果如表8所示。可以看到,本文提出的算法在识别率上偶尔会低于 $\alpha++$ 算法,但准

准确率能达到较好的水平。从2种算法挖掘出的间接依赖数目来看, $\alpha++$ 算法^[12]挖掘出更多的间接依赖,但是2个算法的准确率表现却没有差异,可见 $\alpha++$ 算法中产生了许多冗余依赖关系。在运行时间方面,本文算法相较于 $\alpha++$ 算法略有优势。综合来看,虽然本文算法在识别率上表现不如 $\alpha++$ 算法,但是该算法能在尽量少的间接依赖数目下得到同样准确率的结果,因此本文算法同样具有足够的竞争力。

表7 非自由选择结构识别算法验证实验数据安排
Table 7 Experimental data arrangement for non-free-choice construct identification algorithm verification

数据编号	流程轨迹长度	事件日志规模	优先级等级
F1	50	600	Level4
F2	60	600	Level4
F3	70	600	Level4
F4	80	600	Level4
F5	90	600	Level4

表8 非自由选择结构识别算法对比实验结果

Table 8 Comparison of experimental results of non-free-choice construct identification algorithm

数据编号	非自由选择结构数量	本文算法				$\alpha++$ 算法 ^[12]			
		识别数目	准确率/%	间接依赖数目	运行时间/ms	识别数目	准确率/%	间接依赖数目	运行时间/ms
F1	5	5	100	14	1025	5	100	20	1562
F2	6	6	100	18	1365	6	100	32	1845
F3	7	7	100	20	1852	7	100	36	2214
F4	8	7	100	22	1955	8	100	42	2456
F5	9	9	100	28	2123	9	100	50	2814

2.2 真实数据实验与分析

本文真实数据来自于内蒙古某三甲医院临床路径系统中产生的事件日志,这里选用急性阑尾炎的事件日志作为实验数据,表9给出该病种事件日志分析后的基本信息。真实数据实验主要选取如下衡量指标:重名活动识别率 cr、重名活动准确率 ca、模型泛化性 mg、模型准确率 ma、非自由选择结构识别率 fr、非自由选择结构准确率 fa 和

运行时间 t 。实验对比算法为 $\alpha++$ 算法^[12]、基于遗传算法的过程挖掘(GA)^[15]、文献[6]算法和 HAC-PSO^[16]。

实验结果如表10所示。可以看到,本文提出的算法有着如下优势:①可以处理更多的特殊结构,如重名活动、非自由选择结构等;②在识别率和准确率上与现有算法相比并没有什么差距;③在保证准确率的情况下同样保证了运

表 9 急性阑尾炎病种事件日志信息

Table 9 Description for event logs of acute appendicitis

病种名称	流程轨迹长度	事件日志规模	重名活动数量	噪声类型	噪声规模/%	非自由选择结构数量
急性阑尾炎	53	200	10 对	① ~ ④	6	2

表 10 真实数据实验结果

Table 10 Experimental results for real data

算 法	重名活动识别率/%	重名活动准确率/%	模型泛化性/%	模型准确率/%	非自由选择结构识别率/%	非自由选择结构准确率/%	运行时间/ms
本文算法	100	100	100.56	99.36	100	100	4 036
α + 算法 ^[12]			106.88	94.50	100	100	3 855
GA ^[15]			99.89	99.36	100	100	39 558
文献[6]算法	100	100	103.33	89.89			2 011
HAC-PSO ^[16]			98.78	96.85	50	50	10 254

行效率,运行时间远优于 GA 和 HAC-PSO。

73-82.

3 结 论

本文提出了一种基于统计 α 算法的过程挖掘算法,从 workflow 事件日志中挖掘有用的信息,实现了 workflow 模型的建立。

1) 在经典 α 系列算法的基础上增加了重名活动识别与处理,并将该过程作为 α 算法的预处理活动,以消除重名活动对活动依赖关系识别的干扰。

2) 运用统计学的思想结合经典 α 算法的规则,形成统计 α 算法,消除了事件日志中噪声的干扰问题,同时保证了算法较高的运行效率。

3) 提出了一种新的非自由选择结构识别算法,在保证较高的识别率和准确率的情况下,尽量少得添加间接依赖关系,消除了冗余的活动依赖关系。

通过仿真实验和真实数据实验,验证了本文算法的有效性和优越性。后续工作主要是非自由选择结构识别率的进一步提高以及隐藏活动的挖掘,进一步完善该算法。

参考文献 (References)

[1] 闻立杰. 基于工作流的过程挖掘算法研究[D]. 北京:清华大学,2007.
WEN L J. Studies on algorithms for process mining based on WF-net[D]. Beijing: Tsinghua University, 2007 (in Chinese).

[2] 曾庆田. 过程挖掘的研究现状与问题综述[J]. 系统仿真学报,2007,19(增刊1):275-280.
ZENG Q T. A survey of research issues and approaches on process mining [J]. Journal of System Simulation, 2007, 19 (s1): 275-280 (in Chinese).

[3] COOK J E, WOLF A L. Automating process discovery through event-data analysis[C] // Proceedings of the 17th International Conference on Software Engineering. New York: ACM, 1995:

[4] AGRAWAL R, GUNOPULOS D, LEYMAN F. Mining process models from workflow logs [C] // International Conference on Extending Database Technology. Berlin: Springer, 1998: 467-483.

[5] PINTER S S, GOLANI M. Discovering workflow models from activities' lifespans [J]. Computers in Industry, 2004, 53 (3) : 283-296.

[6] HERBST J, KARAGIANNIS D. Workflow mining with InWoLvE [J]. Computers in Industry, 2004, 53 (3) : 245-264.

[7] SCHIMM G. Mining exact models of concurrent workflows [J]. Computers in Industry, 2004, 53 (3) : 265-281.

[8] GRECO G, GUZZO A, PONTIERI L, et al. Discovering expressive process models by clustering log traces [J]. IEEE Transactions on Knowledge and Data Engineering, 2006, 18 (8) : 1010-1027.

[9] VAN DER AALST W M P, WEIJTERS T, MARUSTER L. Workflow mining: Discovering process models from event logs [J]. IEEE Transactions on Knowledge and Data Engineering, 2004, 16 (9) : 1128-1142.

[10] DEMEDEIROS A K A, DONGEN B F, VAN DER AALST W M P, et al. Process mining for ubiquitous mobile systems: An overview and a concrete algorithm [C] // International Workshop on Ubiquitous Mobile Information and Collaboration Systems. Berlin: Springer, 2004: 151-165.

[11] WEN L, WANG J, VAN DER AALST W M P, et al. A novel approach for process mining based on event types [J]. Journal of Intelligent Information Systems, 2009, 32 (2) : 163-190.

[12] WEN L, VAN DER AALST W M P, WANG J, et al. Mining process models with non-free-choice constructs [J]. Data Mining and Knowledge Discovery, 2007, 15 (2) : 145-180.

[13] WEN L, WANG J, SUN J. Mining invisible tasks from event logs [M] // LI Q, FENG L, PEI J, et al. Advances in data and web management. Berlin: Springer, 2007: 358-365.

[14] LI J, LIU D, YANG B. Process mining: Extending α -algorithm to mine duplicate tasks in process logs [M] // SUI Q, YANG D, WANG T. Advances in web and network technologies, and information management. Berlin: Springer, 2007: 396-407.

[15] DEMEDEIROS A K A, WEIJTERS A J M M, VAN DER AALST

- W M P. Using genetic algorithms to mine process models: Representation, operators and results [R]. Eindhoven: Eindhoven University of Technology, 2005.
- [16] 李莉, 李洪奇, 谢绍龙. 基于变异粒子群算法的过程挖掘[J]. 计算机集成制造系统, 2012, 18(3): 634-638.
LI L, LI H Q, XIE S L. Process mining based on mutation-particle swarm optimization[J]. Computer Integrated Manufacturing System, 2012, 18(3): 634-638 (in Chinese).
- [17] 宋炜, 刘强. 基于模拟退火算法的过程挖掘研究[J]. 电子学报, 2009, 37(增刊1): 135-139.
SONG W, LIU Q. Business process mining based on simulated annealing[J]. Acta Electronica Sinica, 2009, 37(s1): 135-139 (in Chinese).
- [18] 黄黎, 谭文安, 许小媛. 一种基于时序行为的流过程协同重构算法[J]. 计算机工程与科学, 2017, 39(5): 897-903.
HUANG L, TAN W A, XU X Y. Cooperative streaming process reengineering based on sequential behaviors[J]. Computer Engineering & Science, 2017, 39(5): 897-903 (in Chinese).
- [19] DEMEDEIROS A K A, DONGEN B F, VAN DER AALST W M P, et al. Process mining: Extending the α -algorithm to mine short loops [R]. Eindhoven: Eindhoven University of Technology, 2004.
- [20] VAN DER AALST W M P, DONGEN B F, HERBST J, et al. Workflow mining: A survey of issues and approaches [J]. Data & Knowledge Engineering, 2003, 47(2): 237-267.

作者简介:

余建波 男, 博士, 教授, 博士生导师。主要研究方向: workflow 管理、信号处理、质量控制等。

董晨阳 男, 硕士研究生。主要研究方向: workflow 管理、过程挖掘等。

李传锋 男, 硕士研究生。主要研究方向: workflow 管理、过程挖掘等。

Process mining based on statistical α -algorithm

YU Jianbo^{1,*}, DONG Chenyang¹, LI Chuanfeng¹, CHENG Hui², SUN Xiwu²

(1. School of Mechanical Engineering, Tongji University, Shanghai 201804, China;

2. Shanghai Aerospace Equipment Manufacturing Factory, Shanghai 200245, China)

Abstract: Workflow technology is widely used in business process management. However, there are still many problems during the execution of business process because of the imperfect workflow model. Process mining is the most useful tool of workflow modeling, which can obtain objective and valuable information from event logs and build process model. Nevertheless, the existing process mining algorithms still have some problems, such as low accuracy, long operation time and overfitting, which will decrease the accuracy of the workflow model. This paper proposed a new process mining algorithm based on statistical α -algorithm, which can not only ensure the accuracy and suitable fitness, but also decrease the operation time. First, cognominal activity identification rules were proposed to be the pre-treated process of process mining, which could improve the accuracy of algorithm. Second, statistical α -algorithm was proposed as the core algorithm of process mining to eliminate the influence of noise in event logs. Moreover, a new algorithm was proposed to identify non-free-choice constructs, which improved the robustness and accuracy of the algorithm. The accuracy and efficiency of the algorithm are verified by simulation and real case.

Keywords: workflow modeling; process mining; statistical α -algorithm; cognominal activities; non-free-choice constructs

Received: 2017-05-15; Accepted: 2017-09-15; Published online: 2017-10-26 09:13

URL: kns.cnki.net/kcms/detail/11.2625.V.20171026.0913.001.html

Foundation items: National Natural Science Foundation of China (51375290, 71777173); Shanghai Aerospace Science and Technology Innovation Fund (SAST2015054); the Fundamental Research Funds for the Central Universities (22120180068)

* Corresponding author. E-mail: jbyu@tongji.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2017.0371

基于形态学的自动驾驶仪振动信号基线漂移去噪



张景元, 何玉珠*

(北京航空航天大学 仪器科学与光电工程学院, 北京 100083)

摘 要: 导弹自动驾驶仪在振动测试过程中存在信号基线漂移且污染严重的问题, 而传统的时频处理方法难以达到去噪要求, 因此基于形态学基本原理提出了一种用于解决振动信号基线漂移的滤波方法。该滤波方法由3级结构组成, 前2级结构均是基于形态学基本原理, 第3级进行相消与平滑处理, 通过相互级联, 可以有效抑制基线漂移。此外, 通过引入粒子群优化(PSO)算法使得该滤波方法更具适应性。对比实验利用该滤波方法和对比方法对自动驾驶仪实测振动信号与标准 ECG 信号进行了处理, 结果表明: 该滤波方法在抑制基线漂移方面要优于小波阈值去噪和传统的形态学去噪。

关键词: 形态学滤波; 振动信号; 基线漂移; 阈值去噪; 粒子群优化(PSO)

中图分类号: TP391

文献标识码: A

文章编号: 1001-5965(2018)05-0907-07

导弹自动驾驶仪包含陀螺仪、加速度计等敏感元件, 其对调整和稳定弹体姿态有着重要的作用。为了更好地检测自动驾驶仪性能, 测试中通常要进行环境振动试验。然而振动测试中, 由于硬件不稳定, 使采集到的信号污染严重^[1], 再加之由于敏感元件零偏不稳定, 造成积分器输出的信号存在基线漂移, 这对后续信号特征的提取产生了很大影响。因此, 有必要对振动测试信号进行有效的滤波处理。

为了抑制基线漂移, 最常用的方法是通过一个高通滤波器去除采样数据中的漂移噪声^[2-3]。但是当基线漂移非常严重时, 这种方法的去噪效果并不理想。潘超等^[4]将传统的多项式算法进行改进, 用于校正长周期加速度信号中的基线漂移, 但是在进行降维时, 这种方法容易丢失信号的有用成分。Morita 和 Kitagawa^[5]利用一系列模拟光谱研究了扰动相关移动窗口二维基线漂移的影响。王远等^[6]利用改进的小波变换对信号进行10层分解, 然后通过预测相消的方法有效抑制了

AEM (Array Transient Electromagnetic) 数据中的基线漂移, 虽然这种方法的去噪效果优于插值算法, 但是需要对被测信号的细节信息十分了解。庞宇等^[7]利用改进的形态学算法抑制了 ECG 信号中的基线漂移。此外, 罗玉荣等^[8]利用陷波滤波器和盲源分离法实现了对微弱信号基线漂移的抑制。在光谱研究方面, 很多研究者根据实际的应用背景和信号特征, 也提出了许多适用于抑制光谱基线漂移的算法^[9-12]。本文则是根据振动信号的特点, 将形态学去噪思想应用到了具体的工程实践中。

形态学滤波器是一种非线性信号滤波器, 有着严格的数学理论基础。这种方法的局部修正能力较好, 去噪过程相对简单, 不需要信号的频域信息, 只需通过简单的闭-开、开-闭运算即可达到提取信号、抑制噪声的目的^[13]。近年来, 这种方法被广泛应用于图像处理^[14-15]、电力系统^[16]以及振动信号处理^[17]领域。因此, 通过比较上述方法的优缺点, 本文基于形态学基本原理和粒子群优化

收稿日期: 2017-06-05; 录用日期: 2017-08-01; 网络出版时间: 2017-12-15 16:58

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20171215.1130.004.html

* 通信作者: E-mail: heyuzhuhe@buaa.edu.cn

引用格式: 张景元, 何玉珠. 基于形态学的自动驾驶仪振动信号基线漂移去噪[J]. 北京航空航天大学学报, 2018, 44(5): 907-913.
ZHANG J Y, HE Y Z. Removing baseline drift in vibration signal of autopilot based on morphology[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 907-913 (in Chinese).

(PSO)算法,提出了一种用于振动信号去噪处理的3级形态学滤波方法,并对3种自动驾驶仪的实测振动信号进行了处理,效果较好。

1 基础理论

数学形态学最初是由数学家 Matheron 和 Serra 创立的一种信号分析方法,其基本思想主要包括2部分:①目标信号由集合来描述;②通过预先设计的结构元素来局部地修正被测信号的几何结构。这种方法在结构上较为简单,其基本变换仅有2种:腐蚀和膨胀。其他运算(如形态开-闭和形态闭-开)都是这2种基本变换的线性组合。以一维离散信号 $f(n)$ 为例,数学描述如下。

设 $g(n)$ 为形态学滤波中所用结构元素,其定义域为 $G = \{0, 1, \dots, M-1\}$, $F = \{0, 1, \dots, N-1\}$ 则是 $f(n)$ 的定义域,且 $N \gg M$, 则 $f(n)$ 关于 $g(n)$ 的腐蚀和膨胀运算定义如下^[18]:

$$(f \ominus g)(n) = \min_{m=0,1,\dots,M-1} \{f(n+m) - g(m)\} \quad (1)$$

$$n = 0, 1, \dots, N + M - 2$$

$$(f \oplus g)(n) = \max_{m=0,1,\dots,M-1} \{f(n-m) + g(m)\} \quad (2)$$

$$n = 0, 1, \dots, N - M$$

式中: \ominus 与 \oplus 分别代表腐蚀、膨胀运算。腐蚀运算和膨胀运算分别代表了一个收缩过程和一个膨胀过程。收缩过程对于减少信号峰值、加宽谷域方面效果良好;而膨胀过程则起到了相反的作用。形态学滤波中,通常是将这2种运算进行级联得到开、闭运算,其定义如下:

$$(f \circ g)(n) = [(f \ominus g) \oplus g](n) \quad (3)$$

$$(f \cdot g)(n) = [(f \oplus g) \ominus g](n) \quad (4)$$

式中:符号“ \circ ”表示开运算,“ \cdot ”表示闭运算。开运算和闭运算是形态学滤波中最基本的滤波方式,分别用于抑制信号的正、负脉冲噪声。但是单独使用的去噪效果不好,通常将两者级联使用,形成形态开-闭(O_c)和形态闭-开(C_o)运算,其定义如下^[18]:

$$O_c(f(n)) = f \circ g \circ g \quad (5)$$

$$C_o(f(n)) = f \cdot g \cdot g \quad (6)$$

据上述定义,开-闭、闭-开运算可以同时起到

抑制信号中正、负脉冲干扰的作用。实际应用中,为了抑制形态开的反扩展性和形态闭的扩展性,尽可能降低滤波过程中统计偏移所造成的影响,通常将这2种滤波器组合使用,组合滤波器的输出形式为

$$R(n) = (O_c(f(n)) + C_o(f(n)))/2 \quad (7)$$

式中: $R(n)$ 为组合滤波器的输出信号; $f(n)$ 为输入信号,即待滤波信号。

2 滤波器设计

2.1 滤波器实现方式

对于实测信号 $y(n)$ 而言,其信号成分可以表示为

$$y(n) = x(n) + h(n) + l(n) + \sigma(n) \quad (8)$$

式中: $x(n)$ 为有用信号; $h(n)$ 为高频干扰; $l(n)$ 为基线漂移噪声; $\sigma(n)$ 为随机干扰。 $h(n)$ 在时域上表现出很窄的波形,利用形态学去噪时,可以通过较窄的形态滤波器予以滤除;基线漂移噪声 $l(n)$ 变化缓慢,其特征接近于低频线性变化,可以通过较宽的形态滤波器予以滤除; $\sigma(n)$ 则可以通过适当的平滑处理进行解决。

基于上述分析,振动信号形态学级联滤波器的数学模型可以表示为如图1所示的基本结构。根据图1,去噪步骤共有3步。

步骤1 利用式(5)、式(6)对输入信号 $x(n)$ 进行去噪,其目的是滤除背景杂波及高频脉冲干扰,滤波器数学模型为

$$f_{1,1}(n) = O_c x(n, k_1) \quad (9)$$

$$f_{1,2}(n) = C_o x(n, k_1) \quad (10)$$

式中: $f_{1,1}(n)$ 和 $f_{1,2}(n)$ 分别为2种运算方式步骤1滤波后的输出信号; $x(n, k_1)$ 为待滤波信号 $x(n)$ 通过结构元素 k_1 进行形态滤波。步骤1的目的是去除 $h(n)$ 。滤波后所剩信号包含 $x(n)$ 、 $l(n)$ 和 $\sigma(n)$ 。

步骤2 将步骤1滤波后的信号再次进行处理。步骤2选择较宽的结构元素滤除有用信号 $x(n)$,该级滤波器的结构形式为

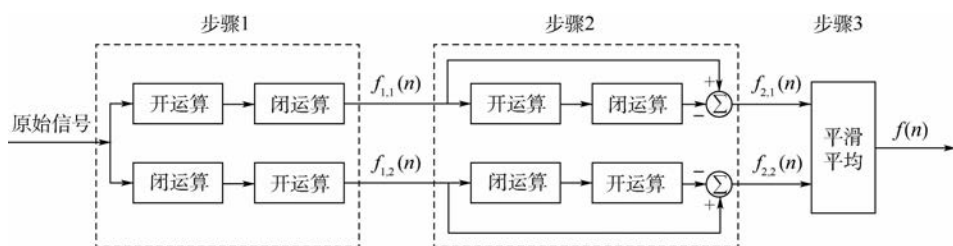


图1 级联滤波器结构模型

Fig. 1 Structure model of cascading filter

$$f_{2,1}(n) = O_c f_{1,1}(n, k_2) \quad (11)$$

$$f_{2,2}(n) = C_o f_{1,2}(n, k_2) \quad (12)$$

式中: $f_{2,1}(n)$ 和 $f_{2,2}(n)$ 分别为采用 2 种运算方式进行步骤 2 滤波后的输出信号。步骤 2 滤波后, 所剩信号只有 $l(n)$ 和 $\sigma(n)$ 。

步骤 3 将步骤 1 和步骤 2 滤波所得信号进行相消与平滑处理, 得到最终的去噪信号, 其表达式为

$$f(n) = \frac{1}{2T_0} \sum_{n-T_0}^n (f_{1,1}(n) - f_{2,1}(n) + f_{1,2}(n) - f_{2,2}(n)) \quad (13)$$

式中: T_0 为滑动滤波周期。

2.2 滤波器关键元素选取

$h(n)$ 和 $l(n)$ 滤除的关键在于形态学滤波器形状的选择, 而该形状可由形态学结构元素来决定。结构元素种类繁多, 常见的有直线形、矩形、菱形、抛物线形等规则或不规则曲线, 去噪时往往根据需要进行选择。滤波结构元素 k_1 、 k_2 的选择对于去噪结果有很大影响, 其形状、尺寸决定了去噪的效果^[11]。但是, 越复杂的结构元素, 如圆盘结构元素和抛物线结构元素, 计算量较大, 直接影响信号处理速度, 难以适应实时性要求较高的自动驾驶仪振动测试。因此, 本文选用简单实用的直线形结构元素。

对于直线形结构元素而言, 决定其形状的元素有 2 个, 即宽度 L 和角度 θ , L 决定了滤波器的宽度, θ 决定了滤波器的方向。图 2 为直线形结构元素的示意图, 其中 $L = \{3, 5\}$, $\theta = \{0, 45, 90\}$ 。这 2 个元素的选取会对去噪效果产生一定的影响, 如果选取不当, 甚至可能使去噪失败。

因此, 考虑到 L 和 θ 对去噪的影响, 并将其代入式(13), 得到最终的输出信号表达式为

$$f_o(n) = \frac{1}{2T_0} \sum_{n-T_0}^n (f_{1,1}(n | (L_1, \theta_1)) - f_{2,1}(n | (L_2, \theta_2)) + f_{1,2}(n | (L_1, \theta_1)) - f_{2,2}(n | (L_2, \theta_2))) \quad (14)$$

式中: L_1 和 θ_1 分别为第 1 级滤波器中所使用结构元素的宽度和角度; L_2 和 θ_2 分别为第 2 级滤

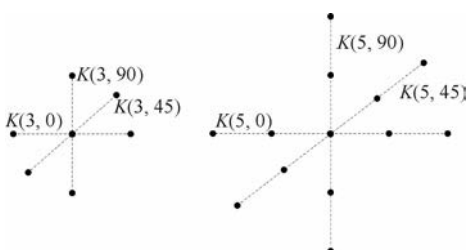


图 2 直线形结构元素基本模型

Fig. 2 Basic model of linear structural elements

波器中所使用结构元素的宽度和角度。

为了合理选择直线形结构元素 $K(L, \theta)$ 的 2 个参数 L 和 θ , 本文采取以下方式: 如果考虑结构元素所有可能的情况, 那么由其组成的集合可以称为全方位结构元素, 即

$$K_{L,\theta} = \{K(L + i, \theta + j) | i \in \mathbf{N}^+, j \in [0, 2\pi]\}$$

但是在实际的去噪过程中, 不同的样本信号需要采用不同的结构元素才能实现较好的去噪效果。因此, 为了提高本节所述滤波器的自适应能力, 本文又进一步引入了 PSO 算法来实现对结构元素的最优化选择。

PSO 算法是一种解决工程优化问题的有效方法, 经典的 PSO 算法可以表示为

$$v_{ij}(t+1) = v_{ij} + c_1 r_1 (p_{ij} - x_{ij}(t)) + c_2 r_2 (g_j - x_{ij}(t)) \quad (15)$$

式中: c_1 和 c_2 为学习因子; r_1 和 r_2 为随机变量; x_{ij} 为个体最佳位置, $i = 1, 2, \dots, NP, j = 1, 2, \dots, N, NP$ 为种群大小, N 为节点数; v_{ij} 为用于更新个体的最佳位置 (p_{ij}) 和最佳位置的全局自适应值 (g_j)。

为了优化滤波器的参数, 目标函数定义为

$$Q_{\text{SNR}}(L_1, \theta_1, L_2, \theta_2) = 10 \lg \left(\frac{\sum_{n=1}^S (x(n))^2}{\sum_{n=1}^S (x(n) - y(n, L_1, \theta_1, L_2, \theta_2))^2} \right) \quad (16)$$

式中: $x(n)$ 表示原始信号, $y(n)$ 表示滤波后的信号; S 为信号长度, 且 $n = 1, 2, \dots, S$ 。

3 应用实验

为了验证本文方法在去除自动驾驶仪振动信号基线漂移噪声方面的优越性, 将现场采集到的信号作为去噪样本, 利用小波阈值去噪、传统形态学去噪和本文方法分别进行去噪处理。图 3(a) 为现场采集到 A 型号导弹自动驾驶仪在静止状态时的反馈信号, 图 3(b) 为振动测试时的反馈信号。由图 3 可知, 在静止状态下, 信号虽然也受到了一定的干扰, 但是不存在基线漂移噪声, 经过振动试验, 同一反馈信号出现了漂移噪声。

3.1 传统形态学去噪

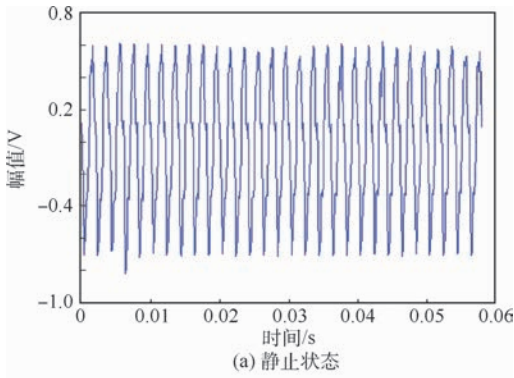
为了对图 3(b) 所示的信号进行处理, 首先采用式(7)所示的传统形态学去噪方法进行处理。消噪后的信号波形如图 4 所示。

利用传统形态学去噪后, 信噪比 (Signal to Noise Ratio, SNR) 提高到了 15.17, 均方差 (Root

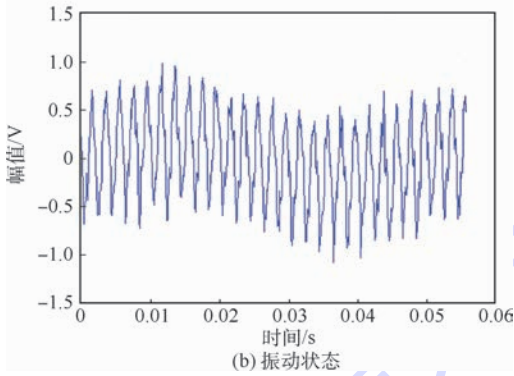
表1 不同分解层数下小波阈值去噪结果

Table 1 Denoising results of wavelet transformation with different wavelet-bases

小波基	分解层数	均方差	信噪比	波形相似比
sym8	2	0.151	20.80	0.94
sym8	3	0.202	13.82	0.89
sym8	4	0.428	-21.84	0.32
db3	2	0.165	18.71	0.93
db3	3	0.254	7.76	0.83
db3	4	0.428	-21.42	0.32



(a) 静止状态



(b) 振动状态

图3 静止和振动状态实测信号

Fig. 3 Measured signals in static and vibration states

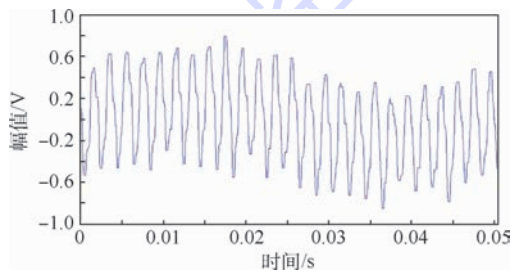


图4 传统形态学方法去噪结果

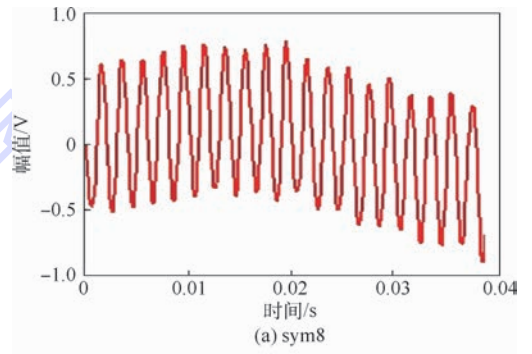
Fig. 4 Denoising results of traditional morphological method

Mean Square Error, RMSE) 为 0.164, 对振动信号的噪声起到了一定的抑制作用, 但是根据图 4 所示的去噪结果来看, 基线漂移噪声仍然存在。

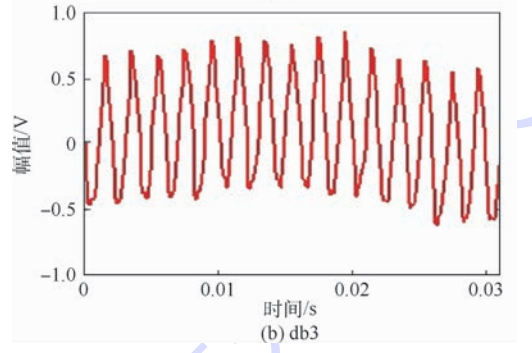
3.2 小波阈值去噪

本文利用小波变换同样对该实测振动信号进行了去噪处理。选择常用的“db”小波和“sym”小波作为小波基, 分别进行去噪处理。表 1 为 2 种小波基经过不同分解层数(2,3,4)后的去噪指标。

由表 1 所示的去噪指标可知, 当分解层数大于 3 时, 信噪比变为负值, 而且波形相似比急速下降至 0.3 左右, 说明有用信号丢失严重, 因此分解层数不能大于 3 层。此外, 比较表 1 中数据可知, 当分解层数为 2 时, 去噪效果较好, 此时去噪后的波形如图 5 所示。可知, “sym8”小波去噪后的信号光滑度较好, 但 2 种小波基去噪后都存在一个问题, 即基线漂移噪声仍然存在。



(a) sym8



(b) db3

图5 小波阈值方法去噪结果

Fig. 5 Denoising results of wavelet threshold method

3.3 本文方法去噪

在利用本文方法进行去噪之前, 需确定第 1 级和第 2 级滤波器的结构元素。根据 2.2 节对全方位结构元素的定义, 由式(16)可得 L_1 、 L_2 和 θ_1 、 θ_2 的取值对去噪效果的影响。由图 6 可知, 当信噪比达到最大时, $L_1 < 5$, $L_2 > 28$, $\theta_1 < 22$, $\theta_2 > 40$ 。

利用 PSO 算法进行 6 次优化实验, 优化结果如表 2 所示。信噪比最终达到 28 左右, 此时优化所得结构元素的具体取值如表 2 所示。

由表 2 的优化结果可知, 6 次优化实验所得结果符合图 6 所得结构元素的取值范围, 进一步说明 PSO 算法的引入增加了本文方法的健壮性和鲁棒性。

结构元素确定之后, 根据式(9)~式(13), 得到本文方法滤波后的结果, 如图 7 所示。由滤波结果可知, 本文方法能够对基线漂移进行校正。

3.4 第 2 种振动信号去噪

为了进一步验证本文方法适用于处理导弹自

动驾驶仪振动测试信号,将 B 型号导弹自动驾驶仪振动过程中的某实测信号进行滤波处理,处理后的结果如图 8 所示。

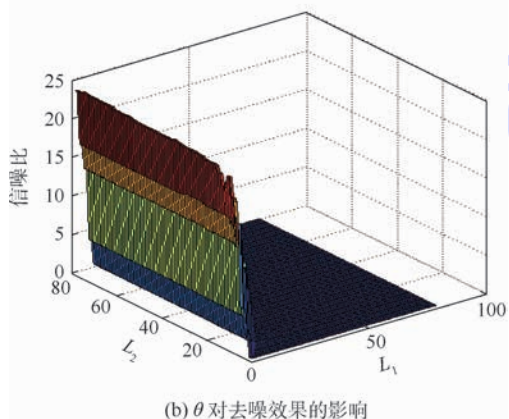
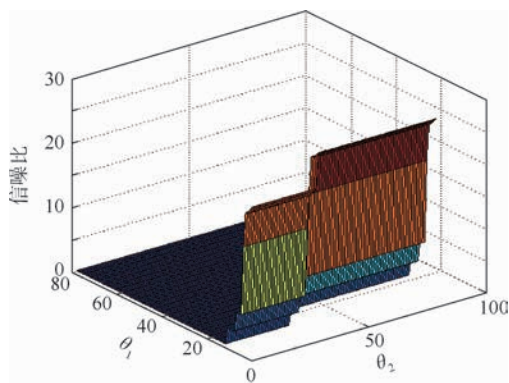


图 6 L 和 θ 对去噪效果的影响

Fig. 6 Influence of L and θ on denoising

表 2 PSO 算法优化结果

Table 2 Optimization results using PSO algorithm

实验次数	L_1	θ_1	L_2	θ_2	信噪比
1	2	2	29	40	28.7
2	3	3	30	80	28.7
3	3	2	78	72	28.7
4	3	5	68	65	27.6
5	3	5	76	40	27.6
6	2	3	36	72	28.1

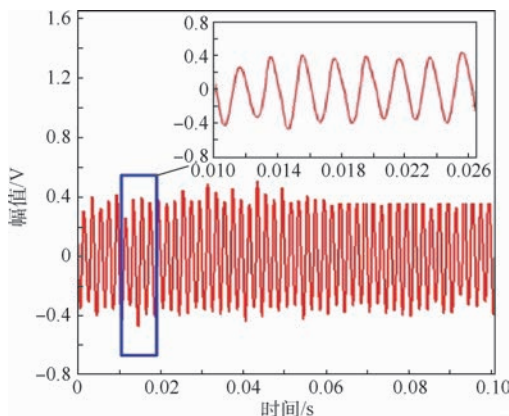
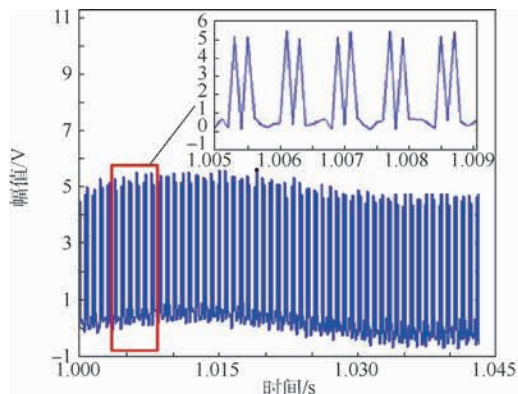
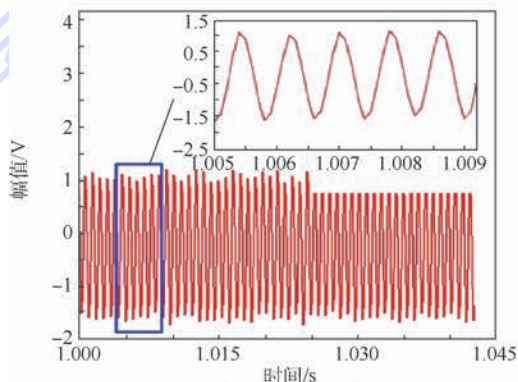


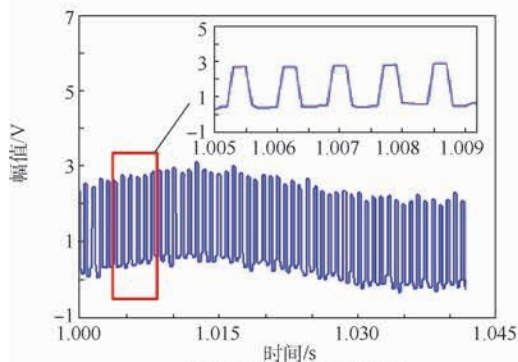
图 7 本文方法对自动驾驶仪实测振动信号的去噪结果
Fig. 7 Denoising results of measured vibration signal of autopilot using proposed method



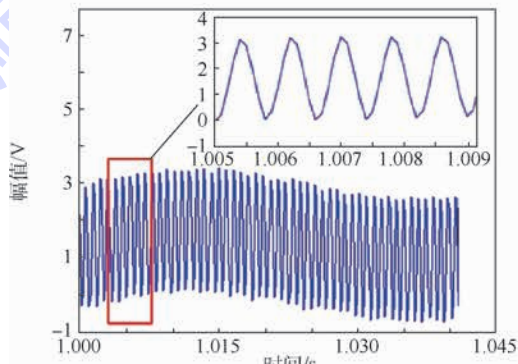
(a) B 型号导弹自动驾驶仪振动信号原始样本



(b) 本文方法去噪结果



(c) 传统形态学方法去噪结果



(d) 小波阈值方法去噪结果

图 8 本文方法和对比方法对 B 型号导弹自动驾驶仪实测振动信号的去噪结果

Fig. 8 Denoising results of measured vibration signal of Type B missile autopilot by proposed method and reference method

由图8(a)可知,理论上被测信号应当为正/余弦形式的波形,但是振动过程中测得的原始信号波形失真严重,且含有基线漂移噪声,此时无法有效提取信号特征,直接影响对自动驾驶仪相关性能的分析。利用本文方法去噪后,根据图8(b)可明显看出,此时基线漂移噪声被有效抑制,而且背景杂散噪声也被消除,有效信号得到了恢复。而另外2种对比方法虽然对噪声有一定的抑制作用,但是无法消除根本性的基线漂移噪声。

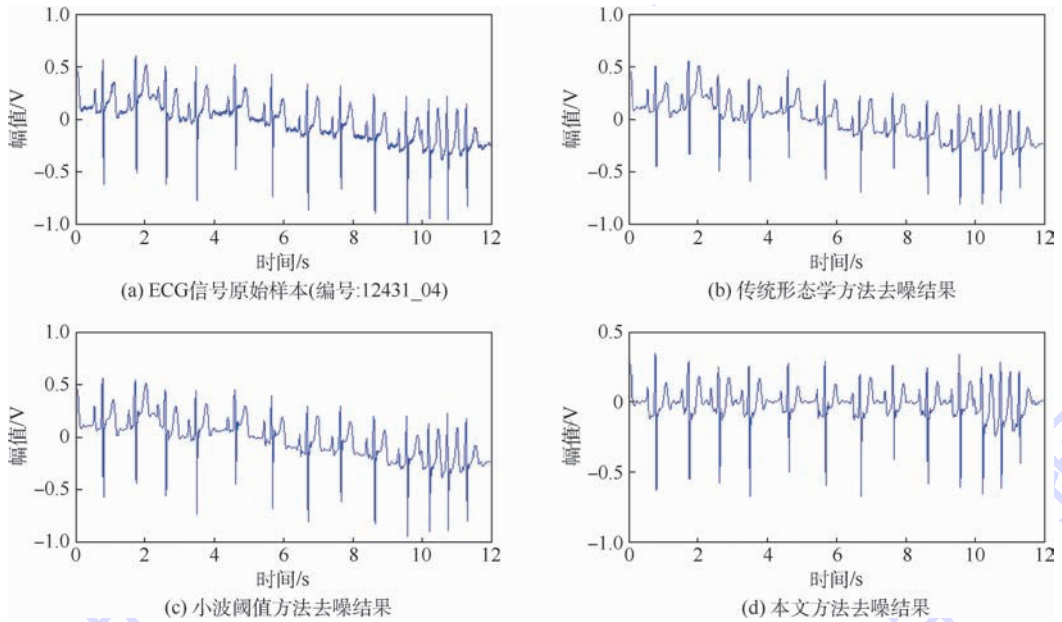


图9 含有基线漂移噪声的 ECG 信号去噪结果

Fig.9 Denoising results of ECG signals containing baseline drift noise

5 结论

本文分析了工程应用背景,根据形态学基本原理提出了适用于自动驾驶仪振动信号去噪处理的具体方法,通过对实测数据进行处理,得到如下结论:

1) 不同的振动信号,其基线漂移噪声不同,需选择不同的结构元素进行基线漂移校正。线性结构元素的2个参数,即宽度和角度的取值对去噪结果有很大影响。

2) 与传统的小波阈值去噪和形态学去噪相比,本文方法能够有效抑制基线漂移,降低样本信号的均方差,提高信噪比。

3) 通过对典型的含有基线漂移的 ECG 信号进行处理,进一步验证了本文方法在去除同类型振动信号时具有较强的自适应能力。

参考文献 (References)

[1] EVANS J W, KUNDU P, HOROVITZ S G, et al. Separating

4 对比实验

医学心电信号(ECG)是一种常见的非周期振动信号,该类信号经常伴有基线漂移。为了验证本文方法在矫正基线漂移时的自适应能力,本文从 MIT-BIH 心率失常数据库中随机选择了一种含有基线漂移成分的心电信号,其编号为 12431_04,利用本文方法和对比方法分别对该样本信号进行了去噪分析,结果如图9所示。

slow BOLD from non-BOLD baseline drifts using multi-echo fMRI[J]. NeuroImage, 2015, 105:189-197.

[2] CHIU H C. Stable baseline correction of digital strong-motion data[J]. Bulletin of Seismological Society of America, 1997, 87(4):932-944.

[3] BOORE D M, BOMMER J J. Processing of strong-motion accelerograms: Needs, options and consequences[J]. Soil Dynamics and Earthquake Engineering, 2005, 25(2):93-115.

[4] PAN C, ZHANG R F, LUO H, et al. Baseline correction of vibration acceleration signals with inconsistent initial velocity and displacement[J]. Advances in Mechanical Engineering, 2016, 8(10):1-11.

[5] MORITA S, KITAGAWA K. Effect of baseline drift on perturbation-correlation moving-window two-dimensional correlation spectroscopy[J]. Vibrational Spectroscopy, 2012, 60:217-219.

[6] WANG Y, JI Y J, LI S Y. A wavelet-based baseline drift correction method for grounded electrical source airborne transient electromagnetic signals[J]. Exploration Geophysics, 2013, 44(4):229-237.

[7] 邓璐, 庞宇, 赵艳霞, 等. 基于形态学的心电信号基线漂移矫正方法[J]. 数字通信, 2013, 40(3):14-16.

DENG L, PANG Y, ZHAO Y X, et al. Removal method of baseline drift from ECG signals based on morphology filter[J]. Dig-

- ital Communication, 2013, 40(3):14-16 (in Chinese).
- [8] LUO Y R, HARGRAVES R H, BELLE A, et al. A hierarchical method for removal of baseline drift from biomedical signals: Application in ECG analysis [J]. Scientific World Journal, 2013, 2013: 896056.
- [9] DU Y G, CHAMBERS S A. Etalon-induced baseline drift and correction in atom flux sensors based on atomic absorption spectroscopy [J]. Applied Physics Letters, 2014, 105(16):163113.
- [10] LOPATKA M, BARCARU A, SJERPS M J, et al. Leveraging probabilistic peak detection to estimate baseline drift in complex chromatographic samples [J]. Journal of Chromatography A, 2016, 1431: 122-130.
- [11] LIU G F, LUO X L, YANG J. Baseline drift effect on the performance of neutron and γ ray discrimination using frequency gradient analysis [J]. Chinese Physics C, 2013, 37(6):63-69.
- [12] ZHU F, QIN B J, FENG W Y, et al. Reducing Poisson noise and baseline drift in x-ray spectral images with bootstrap Poisson regression and robust nonparametric regression [J]. Physics in Medicine and Biology, 2013, 58(6):1739-1758.
- [13] WANG R Q, LI Q, ZHANG M. Application of multi-scaled morphology in denoising seismic data [J]. Applied Geophysics, 2008, 5(3):197-203.
- [14] SALEMBIER P, WILKINSON M H F. Connected operators: A review of region-based morphological image processing techniques [J]. IEEE Signal Processing Magazine, 2009, 26(6):136-157.
- [15] 赵于前, 王小芳, 李桂源. 基于多尺度多结构元素的肝脏图像分割 [J]. 光电子·激光, 2009, 20(4):563-566.
- ZHAO Y Q, WANG X F, LI G Y. Liver image segmentation based on multi-scale and multi-structure elements [J]. Journal of Optoelectronics · Laser, 2009, 20(4):563-566 (in Chinese).
- [16] GAUTAM S, BRAHMA S M. Overview of mathematical morphology in power systems-A tutorial approach [C] // 2009 IEEE Power & Energy Society General Meeting. Piscataway, NJ: IEEE Press, 2009: 3523-3529.
- [17] KOZUMPLÍK J, PROVAZNÍK I. Fast timevarying linear filters for suppression of baseline drift in electrocardiographic signals [J]. Biomedical Engineering Online, 2017, 16:24.
- [18] SERRA J. Image analysis and mathematical morphology [M]. New York: Academic Press, 1982.

作者简介:

张景元 男, 硕士研究生。主要研究方向: 自动测试技术与故障诊断。

何玉珠 男, 博士, 教授。主要研究方向: 测试系统通用性技术、故障诊断、定位技术。

Removing baseline drift in vibration signal of autopilot based on morphology

ZHANG Jingyuan, HE Yuzhu *

(School of Instrumentation Science and Opto-electronics Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: The baseline drift and heavy pollution for vibration test of missile autopilot are still problems. The requirement of denoising is difficult to be achieved by traditional time-frequency method. In this paper, to filter out the baseline drift noise, a new morphological filtering method based on the basic principle of generalized morphology is proposed. The proposed method is composed of three-level structure: the former two are based on the morphological principle, and the third level is designed for cancellation and smoothing. Thus, baseline drift can be effectively suppressed by cascading. In addition, the proposed method is more adaptive by introducing particle swarm optimization (PSO). In the final experiments, the real signals of autopilot and ECG signals are denoised by the proposed method and reference methods. The experimental results show that the proposed method is better than wavelet denoising and traditional morphological denoising in suppressing baseline drift.

Keywords: morphological filtering; vibration signal; baseline drift; threshold denoising; particle swarm optimization (PSO)

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

DOI: 10.13700/j.bh.1001-5965.2017.0347

微小空间内电磁干扰要素检测与辨识方法



凌波^{1,2}, 林赞², 孙宏涛², 苏东林^{1,*}

(1. 北京航空航天大学 电子信息工程学院, 北京 100083; 2. 中国航空工业集团公司成都飞机设计研究所, 成都 610091)

摘 要: 针对射频传感器综合一体化带来的电磁兼容性接口由机箱、连接器和线缆变为背板-模块,使常规电磁干扰(EMI)检测要求和检测手段无法直接适用于模块的问题,基于电磁干扰要素理论,着重研究了综合射频机架内部微小空间内电磁干扰要素检测方法及原理,探索了综合射频机架微小空间下模块电磁发射特征信息的测试方法,研究了综合射频机架工作状态及功能线程的激励控制方法,以及综合射频机架远场电磁发射特性数据、功能模块近场电磁发射特性数据和综合射频机架接口线缆束上的电流传导发射特性数据等电磁干扰要素提取和识别方法。为验证电磁干扰要素检测与辨识方法的可行性,以某机载综合射频机架为例,开展了大量试验研究与分析,包括综合射频机架远场电磁发射特性测试、功能模块近场电磁发射特性测试、综合射频机架互连线缆束电流传导发射特性测试,并将多种工作模式下模块测试结果和综合射频机架电磁发射测试结果进行比较,验证了本文方法的可行性和准确性。

关键词: 综合射频; 微小空间; 射频模块; 电磁兼容; 电磁干扰(EMI)

中图分类号: TN011.4; V219

文献标识码: A **文章编号:** 1001-5965(2018)05-0914-09

机载射频功能综合一体化已成为新一代机载电子设备的必然发展趋势,也带来“基于模块”的电磁兼容性要求、设计方法、检测评估方法等新的电磁兼容性问题^[1-5]。

传统机载电子设备内部架构形式大多为板卡形式,机箱内部的背板、母板在开敞区,可使用电流环、电场探头等近场测试手段直接对机箱内板卡上的电路器件、模块进行电磁干扰(EMI)故障定位测试。相比之下,标准模块内部电路完全封闭不可见,模块封装后对外无显性接口,且模块安装于机架内,操作空间狭小,使得综合机架内部功能模块的电磁干扰识别、定位问题变得非常困难,传统电磁干扰测试方法使用受限。

针对综合射频机架的多样化、复杂化电磁兼

容性问题,本文研究了综合射频机架内部微小空间电磁干扰要素^[6-9]检测方法及原理;对综合射频机架开展了电磁干扰要素测试研究,提出了机架内电磁干扰问题的诊断和识别定位新方法和技术,试验结果验证了该方法的有效性和正确性。

1 电磁干扰要素检测方法

1.1 传统电磁干扰检测

传统电磁干扰检测方法直接检测设备整体的电磁辐射(见图1),其测试数据包含了被测设备所有的辐射特征,是设备内部所有辐射参量在时域和频域的叠加结果^[10-12]。该方法虽能发现辐射超标情况,却难以对产生辐射发射的具体位置、模块、线束等进行精准定位,仅仅是一种“通过”

收稿日期: 2017-05-22; 录用日期: 2017-05-26; 网络出版时间: 2017-06-15 11:03

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20170615.1103.001.html

基金项目: 国家自然科学基金(61427803)

*通信作者: E-mail: sdl@buaa.edu.cn

引用格式: 凌波, 林赞, 孙宏涛, 等. 微小空间内电磁干扰要素检测与辨识方法[J]. 北京航空航天大学学报, 2018, 44(5): 914-922.
LING B, LIN Y, SUN H T, et al. Detection and identification method of electromagnetic interference elements in limited-space[J].
Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 914-922 (in Chinese).

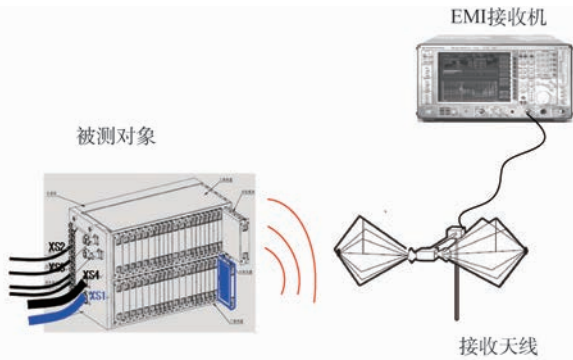


图 1 传统电磁干扰检测方法

Fig. 1 Traditional EMI detection method

或“不通过”评估式的测试方法。

通常电子设备设计和制造单位最为关心的并非产品的测试结果,而是在出现电磁干扰后可以快速对辐射超标点定位及进行有效改进。因为测试状态单一、精细化程度低、问题定位耗时长,所以传统测试方法已经无法适应射频综合技术需求。

1.2 微小空间内模块电磁干扰检测

针对上述问题,提出了对局部器件、模块和电缆等实施近距离精确测量的研究需求^[13-16]。尤其是综合射频机架出现后,这一问题更为急迫。

综合射频机架设计采用标准封装模块化设计,其内部由多个功能模块组成,模块间通过背后的互连总线高速背板进行数据交换,封装模块与高速背板之间通过快插接口进行连接。模块化设计结构紧凑,内部空间几合线度狭小(见图 2),模

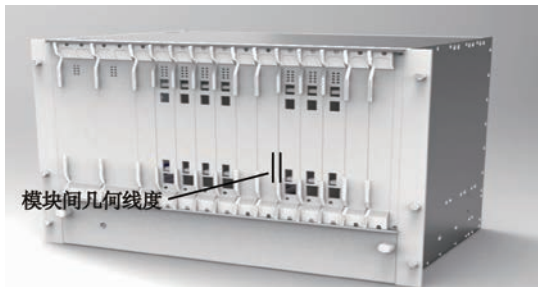


图 2 模块化设计的电子设备

Fig. 2 Modular electronic equipment

块两侧均采用金属壁板封装,电磁干扰通常存在于前端和背部接口的电磁泄漏以及模块与模块之间的互耦。

为解决综合射频机架内的电磁干扰,必须研究和设计一种新的适用于综合射频机架内微小空间电磁干扰检测方法。

2 电磁干扰要素检测与辨识方法

与传统被试品不同,综合射频机架内综合了众多的功能线程,不同功能线程所需的功能模块不同。因此在综合传感器软件调度下,不同功能线程将呈现不同的电磁发射和电磁敏感特性,检测综合射频机架的电磁兼容性时,必须首先实现功能线程激励和控制。

2.1 机架工作状态及功能线程激励控制方法

射频功能综合化后,之前各独立射频功能以功能线程的形式被综合化到综合射频机架内部,即功能线程单元分别由机架内部的独立功能模块或者多个功能模块组合实现。因此要实现综合射频机架的工作状态及模式的控制和切换,须研究综合射频机架内部功能线程的激励设置与控制方法。以 V/UHF 通信功能线程为例,其主要由机架内部的接收激励模块、信号处理模块和前端控制管理以及其他辅助模块完成,如图 3 所示。

V/UHF 通信功能线程的语音激励通过音频采集系统输入和控制,经过前端控制管理、信号处理和接收激励模块以及外部的天线射频接口单元等实现 V/UHF 通信功能。该功能线程的状态、模式控制由航电座舱显示控制激励接口实现。

在研究分析了 V/UHF 通信功能线程的逻辑硬件组成和工作原理后,设计了 V/UHF 通信功能线程不同模式(AM、FM、DS、FH、D/H)和状态(待机、发射、接收)下的测试激励控制方法,如图 4 所示。

2.2 干扰检测与辨识流程

为实现综合射频机架微小空间内电磁干扰

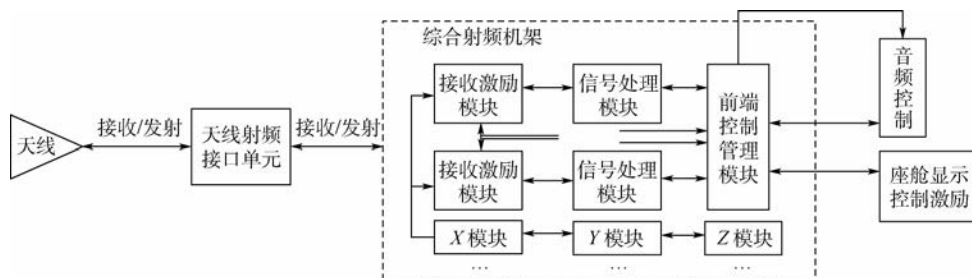


图 3 V/UHF 通信功能线程物理逻辑组成

Fig. 3 Physical logic composition of V/UHF communication functional thread

检测及对干扰源的准确辨识,本文研究并提出了一种新的电磁干扰检测与辨识流程。首先,获取综合射频机架整体对外的远场电磁发射特性数据;

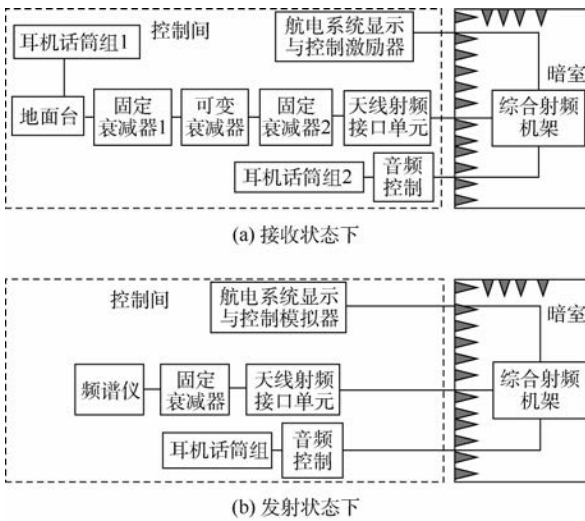


图4 V/UHF通信功能线程激励控制方法设计
Fig.4 Incentive and control method design of V/UHF communication functional thread

然后,分别使用微小空间近场探头、电流检测探头获取机架内功能模块的近场电磁发射特性数据和机架对外互连接口线缆束上的电流传导发射特性数据;最后,利用接口与模块拓扑关系以及近场探头的空间位置分辨能力,进行机架内功能模块的电磁发射要素分析、提取和识别,实现将机架电磁发射超标问题准确定位至内部某一功能模块上。微小空间内电磁干扰检测与辨识流程如图5所示。

2.3 综合射频机架远场电磁发射特性获取方法

与传统 GJB 151B—2013^[17] 标准测试方法略有不同,本测试应覆盖综合射频机架的典型任务设计剖面 and 模式,即综合射频机架的工作状态需按照内部功能线程、航电系统任务模式进行设定和激励。可以分别实施超短波通信、塔康导航和无线电高度表等功能线程的激励,也可同时实施上述功能线程的激励。

为有效获取确保电磁兼容性的实际数据,测试时应使被测对象工作于最大发射工作状态。

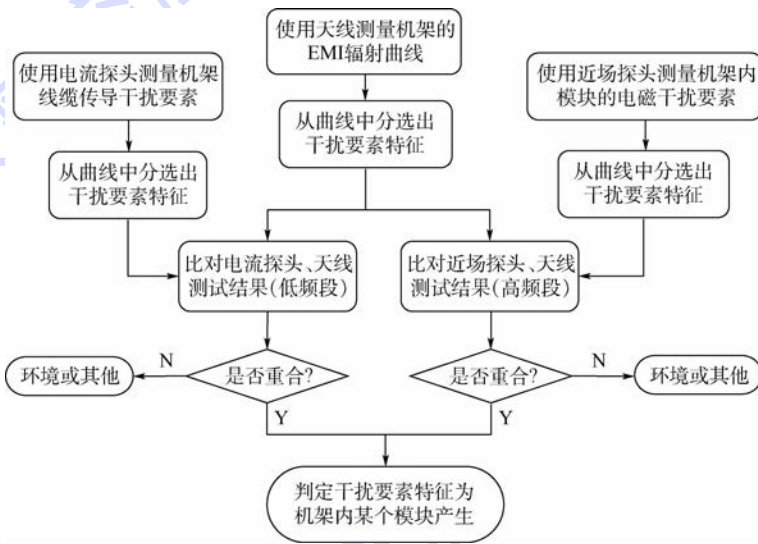


图5 微小空间内电磁干扰检测与辨识流程
Fig.5 Detection and identification procedure of EMI in limited-space

2.4 功能模块近场电磁发射特性获取方法

由于功能模块布局密集、连接管脚复杂,常规测试天线和测试探头难以区分模块的发射特性。为此笔者团队开发研制了微小空间近场电磁和磁场探头,使用该探头可以有效实施综合射频机架的伴随式近场电磁发射要素测量,获取综合射频机架在线动态工作状态下内部功能模块的近场发射特性数据,包括频率和近场幅度。

由于笔者团队研制的微小空间探头对测试场的扰动很小,因此可使探头到被测功能模块测试

部位的间距控制在接近 2 mm。图6给出了逐一采集机架内所有功能模块 30 MHz ~ 18 GHz 频段内的近场发射特性数据的示意图。同样,对不同功能模块测试时,应对其进行功能线程激励,并使其处于最大工作状态。

2.5 综合射频机架互连线缆束电流传导发射特性获取方法

对综合射频机架上每一个对外互连接口线缆束使用传统电磁发射检测探头进行传导电流发射测量,获取机架接口线缆束上的电流传导发射特

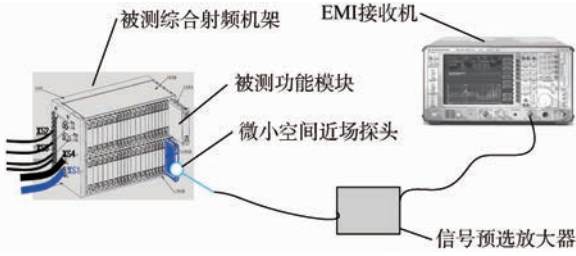


图 6 功能模块近场电磁发射特性测试方法示意图
Fig. 6 Schematic of near-field electromagnetic emission test method of functional module

性数据,包括频率和近场幅度。测试时,应使被测对象工作于最大发射工作状态,对每根线缆束需检测 2 ~ 500 MHz 频段内的电流传导发射数据。

2.6 干扰特征提取及匹配识别方法

1) 数据预处理

对测试数据的幅度进行滑动平均预处理,得到电磁发射测量数据的噪声背景基准值 E_{noise} , 滑动平均处理方式见式(1),其中滑动平均点数 N 一般取值为 9,倒数 1 ~ 8 位数据进行滑动平均时做补零处理。

$$E_{noise} = \frac{1}{N} \sum_{i=1}^N E(i) \quad (1)$$

2) 要素特征提取

将电磁发射原始数据和滑动平均处理得到的噪声基准值按照测试频率点从低至高对应做相减处理,比较差值;若原始数据与滑动平均得到的噪声基准差值大于一定门限(6 ~ 12 dB),认为该发射频率及幅度为关注的电磁发射要素对象,否则舍弃该点测试数据。处理完毕后,实现从综合射频机架远场电磁发射特性数据、功能模块近场电磁发射特性数据和综合射频机架接口线缆束电流传导发射特性数据中分别识别、提取出需要关注和有用的电磁发射要素集合。

式(2)表征了识别、提取出的综合射频机架远场电磁发射要素集合。 M_{far} 为远场电磁发射要素矩阵, $F_{far}(m)$ 和 $E_{far}(m)$ 分别为第 m 个远场发射要素的频率和幅度。

$$M_{far} = [F_{far}(1), F_{far}(2), \dots, F_{far}(m); E_{far}(1), E_{far}(2), \dots, E_{far}(m)] \quad (2)$$

式(3)表征了识别、提取出的某一功能模块近场电磁发射要素集合。 M_{near} 为功能模块近场电磁发射要素矩阵, $F_{near}(n)$ 和 $E_{near}(n)$ 分别为第 n 个发射要素的频率和幅度。

$$M_{near} = [F_{near}(1), F_{near}(2), \dots, F_{near}(n); E_{near}(1), E_{near}(2), \dots, E_{near}(n)] \quad (3)$$

式(4)表征了识别、提取出的综合射频机架接口线缆束电流传导发射要素集合。 M_{cable} 为线缆电流传导发射要素矩阵, $F_{cable}(k)$ 和 $E_{cable}(k)$ 分别为第 k 个电流传导发射要素的频率和幅度。

$$M_{cable} = [F_{cable}(1), F_{cable}(2), \dots, F_{cable}(k); E_{cable}(1), E_{cable}(2), \dots, E_{cable}(k)] \quad (4)$$

3) 干扰识别定位

分别将识别和提取出的互连接口线缆束电流传导发射要素矩阵 M_{cable} 与综合射频机架远场电磁发射要素矩阵 M_{far} 、功能模块近场电磁发射要素矩阵 M_{near} 与综合射频机架远场电磁发射要素矩阵 M_{far} 进行干扰要素匹配分析。

设置频率相对百分比误差阈值 P ,若满足式(5)的准则要求,则认为干扰要素匹配识别成功,即确认该电磁干扰要素来自被测综合射频机架内部。

$$\begin{cases} \min_{i=1}^n \{ |F_{far}(j) - F_{near}(i)| / F_{far}(j) \times 100\% \} \leq P \\ \min_{i=1}^k \{ |F_{far}(j) - F_{cable}(i)| / F_{far}(j) \times 100\% \} \leq P \end{cases} \quad (5)$$

根据微小空间近场探头的空间分辨能力、机架接口线缆拓扑交联关系,进一步定位该干扰要素来自综合射频机架内部的具体哪个功能模块。电磁干扰要素的识别、定位具体流程如图 7 所示。

以功能模块近场电磁发射要素矩阵 M_{near} 与综合射频机架远场电磁发射要素矩阵 M_{far} 进行匹配分析为例,最终的干扰要素识别定位结果形式见表 1。

表 1 干扰要素识别定位结果

Table 1 Results of interference element identification and positioning

电磁干扰要素	干扰要素识别定位结果			
	模块 1	模块 2	...	模块 n
$F_{far}/E_{far}(1)$				
$F_{far}/E_{far}(2)$				
\vdots				
$F_{far}/E_{far}(j)$	√	√		
\vdots				
$F_{far}/E_{far}(m-1)$		√		
$F_{far}/E_{far}(m)$				

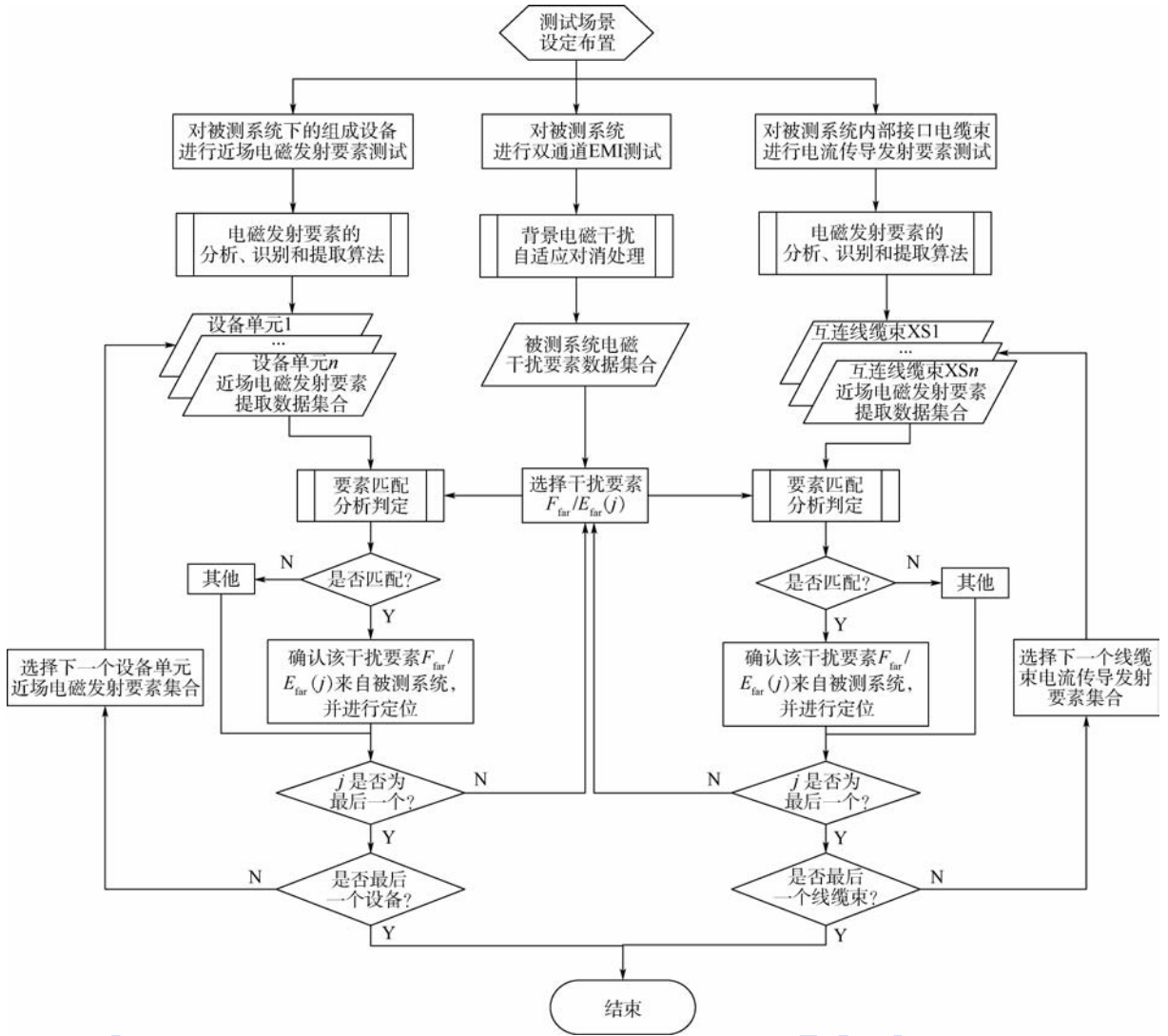


图7 干扰特征提取及识别方法

Fig. 7 Method of extraction and identification of interference features

3 方法验证及应用

为验证电磁干扰要素检测与辨识方法的可行性和识别结果,本文以某机载综合射频机架为例,开展了大量试验研究与分析。具体包括:综合射频机架远场电磁发射特性测试、功能模块近场电磁发射特性测试和综合射频机架互连线缆束电流传导发射特性测试。

3.1 综合射频机架远场电磁发射特性测试

本节分别测试了综合射频机架在静默和工作模式1~5下的远场电磁发射结果。综合射频机架各工作模式设置如表2所示,部分测试结果见图8,图中红线为GJB 151B—2013^[17]标准规定的发射限值要求。可见,不同模式和功能线程工作时架的电磁发射频谱和量值是不同的。

由于不同工作状态下综合射频机架的辐射测试结果存在较大差异,因此如要对不同频点的辐

表2 综合射频机架远场电磁发射特性测试状态设置
Table 2 State setting of electromagnetic emission characteristic test of RF integrated rack in far-field

功能线程	综合射频机架工作模式/功能线程设置					
	静默	模式1	模式2	模式3	模式4	模式5
短波	静默	静默	静默	静默	1波道	1波道
超短波	静默	43波道	19波道	10波道	静默	静默
线程1	静默	静默	空空	静默	静默	空地
线程2	静默	A模式	静默	静默	A模式	静默
线程3	静默	入网	静默	静默	入网	静默
线程4	静默	5波道	5波道	5波道	5波道	5波道
线程5	静默	1波道	1波道	1波道	1波道	1波道
线程6	静默	开启	开启	开启	开启	开启
线程7	静默	传密	传密	传密	传密	传密

射源进行定位和识别,必须基于近场电磁发射测试结果进行进一步分析,并同时与远场测试结果进行相互验证。

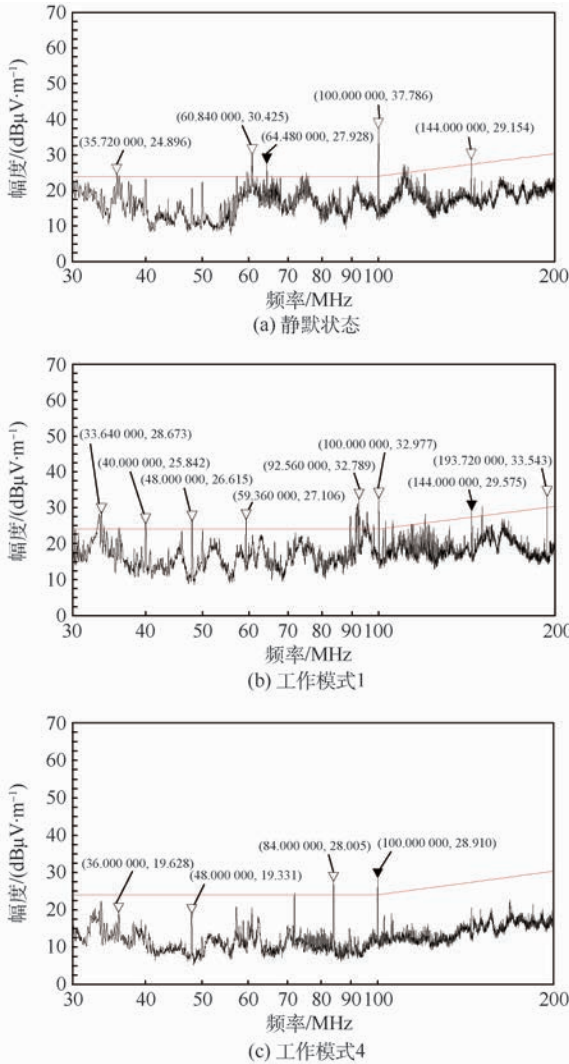


图8 综合射频机架不同工作状态下的测试结果

Fig. 8 Test results of RF integrated rack under different modes

3.2 综合射频机架互连线缆束电流传导发射特性测试

采用2.5节的方法在综合射频机架接口XS4、XS5、XS6的线缆处进行线缆束电流传导发射测试,将测试结果与综合射频机架的远场电磁发射结果(RE102)进行对比,结果如图9所示。

通过图9可知,综合射频机架的外部电磁发

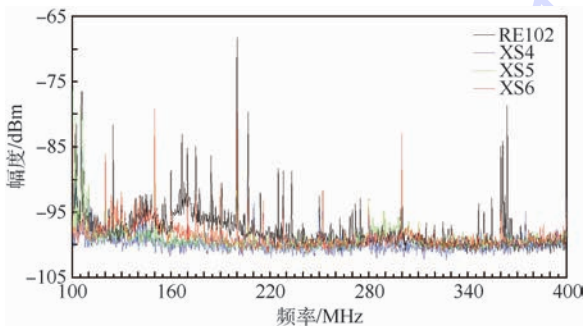


图9 综合射频机架互连线缆束电流传导发射测试结果

Fig. 9 Test results of cable current of RF integrated rack

射,其中一部分是通过线束直接传导发射形成的,还有一部分频点并未包含在线束发射频谱中,通过对机架内部电磁发射源的分析可以得到。线缆作为传导辐射源,仅能反映机架内部的部分辐射状态,而对电磁发射状态的精确分析,需要对模块进行基于工作模式的精准分析和测试。

3.3 功能模块近场电磁发射特性测试

结合3.1节和3.2节的测试结果,为了进一步验证2.6节中的干扰特征提取及匹配识别方法,需对综合射频机架内的功能模块进行电磁发射测试。测试采用高灵敏、高空间分辨能力微型近场探头、预选放大器组件和频谱分析仪,测量机架工作时各模块接插件接口以及壳体缝隙处的电磁泄漏。具体测试模块如表3所示。

由于机架的发射状态与工作模式相关,因此为了清晰地对试验结果进行对比,功能模块测试时,将综合射频机架调整为工作模式1。通过对此工作模式下的模块测试结果和机架远场电磁发射测试结果进行比较,验证方法的可行性和准确性。部分模块的测试结果如图10所示。

图10中, Environment noise (蓝色曲线)表示近场探头在远离被测模块背景空间下的测试噪底,互连接口、壳体缝隙和RF接口等表示近场探头在靠近被测功能模块不同位置处的干扰信号测试结果。

应用以上测试方法,完成了综合射频机架电磁辐射超标问题的识别和定位,结果如表4所示。综合射频机架远场测试发现,机架在35.64、40、48、59.36、100、145、191.7 MHz频点辐射超标;经过对综合射频机架进行模块的近场测试,并与机架的远场辐射测试结果进行比较分析,最终将超标频点的辐射源头定位至机架内部的具体功能模块上。表4中第1列数据为机架的远场电磁发射超标频点,其余为各模块的辐射发射较强的频点;表格内颜色一致的单元表示模块的发射测试结果

表3 功能模块列表

Table 3 List of functional modules

模块序号	模块名称	测试结果
模块1	信号处理模块	图10(a)
模块2	控制管理模块	图10(b)
模块3	告警模块	图10(c)
模块4	窄带接收模块	图10(d)
模块5	宽带接收模块	图10(e)
模块6	开关模块	图10(f)
模块7	频率源模块	图10(g)
模块8	电源模块	图10(h)
模块9	时频模块	图10(i)

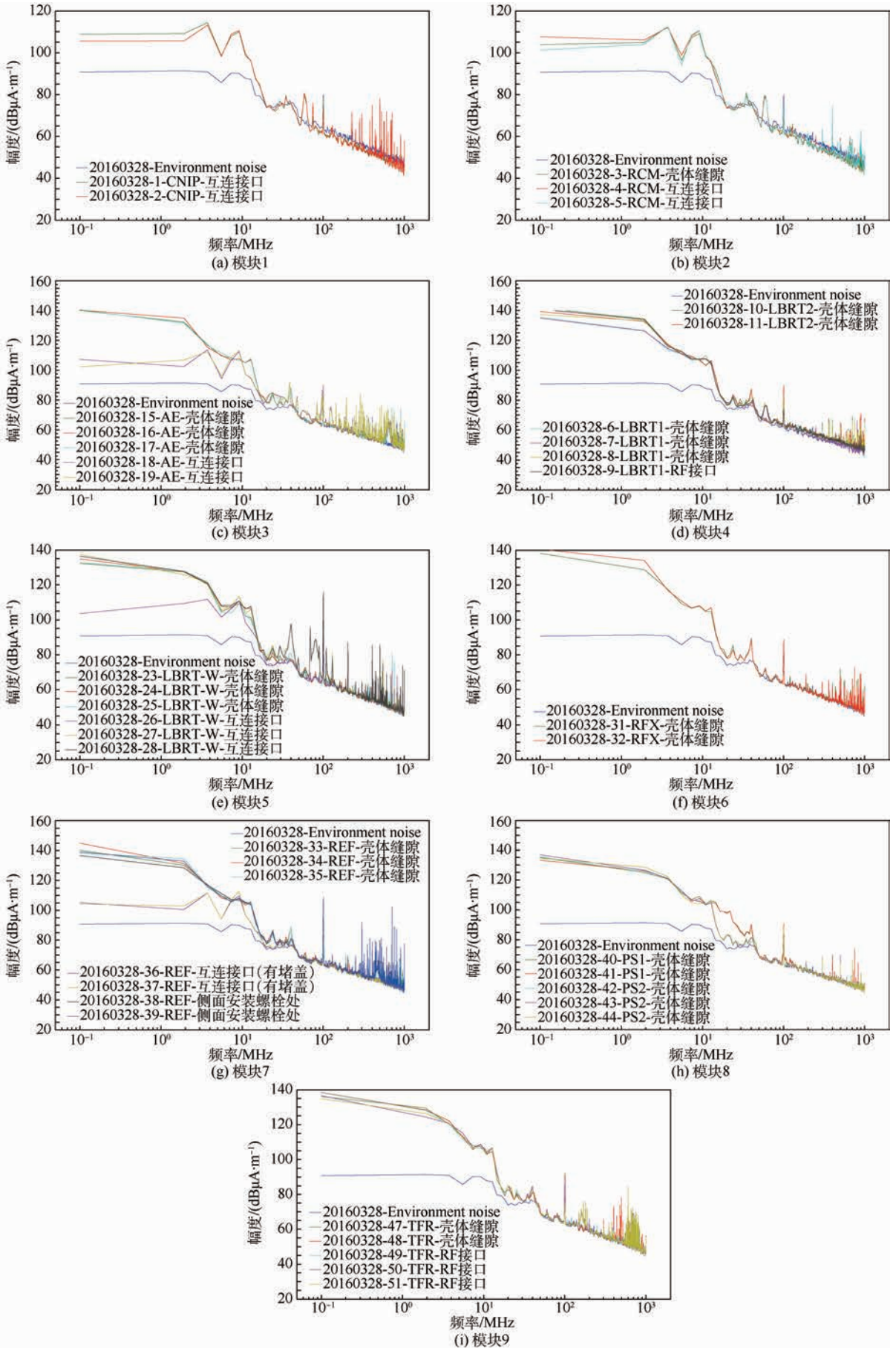


图 10 模块 1 ~ 9 的测试结果

Fig. 10 Test results of Module 1-9

表 4 综合射频机架电磁干扰要素测试定位结果

Table 4 Test and positioning results of electromagnetic interference elements of RF integrated rack MHz

超标频点	模块发射峰值频点及识别结果								
	模块 1	模块 2	模块 3	模块 4	模块 5	模块 6	模块 7	模块 8	模块 9
35.64	58.3	34.6	23.7	23.7	23.7	23.7	23.7	25.5	12.8
40	100	60	29.2	40	29.2	40	40	41.9	20.1
48		100	38.3	60	40	60	60	100	25.5
59.36			40	100	49.2	100	100	120	40
100			47.4		60	120	129.2	189.2	60
145			54 ~ 65		69.2 ~ 89.2	189.2	189.2		100
191.7			100		100		200		149.2
			143.7						
			192.8						

果与机架的发射测试结果一致吻合,被成功识别。例如,对于 59.36 MHz 超标频点,发射源头最终定位至机架内模块 2、模块 4、模块 5、模块 6、模块 7、模块 9。机架的 7 个干扰超标频点,其中有 6 个被准确识别和定位,识别定位成功率大于 80%。

4 结 论

本文依据射频综合环境下可更换模块的特点及其电磁兼容特性开展研究,主要取得了以下结论:

1) 研究了射频综合环境下功能模块干扰要素检测与识别技术,形成射频综合系统下机架和功能模块的电磁干扰要素辨识方法。

2) 实现了对所需测试的功能模块以及要素集进行独立工作状态控制和接口信号激励,构建了射频综合模块的在线动态测试环境。

3) 通过对比射频综合机架在不同功能线程状态下的 RE102 电磁发射测试数据和机架内部各模块的近场电磁发射测试数据,得到了电磁干扰要素,使得综合机架所有 RE102 测试超标频点均可在模块测试数据中找到对应信息。

本研究为综合射频机架可更换功能模块以及干扰敏感耦合要素集的电磁兼容特性测试和诊断提供了技术支撑,保证了各功能模块形成的射频综合系统最终满足电磁兼容指标要求。通过某型号实装综合射频机架的测试应用,验证了本文提出的检测识别方法的有效性。

参考文献 (References)

[1] 苏东林,谢树果,戴飞,等. 系统级电磁兼容量化设计理论与方法[M]. 北京:国防工业出版社,2015:115.
SU D L, XIE S G, DAI F, et al. The theory and method of quantification design on system-level electromagnetic compatibility [M]. Beijing: National Defense Industry Press, 2015: 115 (in Chinese).

[2] 苏东林,王冰切,金德琨,等. 电子战特种飞机电磁兼容预设

计技术[J]. 北京航空航天大学学报, 2006, 32(10): 1239-1243.

SU D L, WANG B Q, JIN D K, et al. The prediction techniques of electromagnetic compatibility for EW-warfare aircraft [J]. Journal of Beijing University of Aeronautics and Astronautics, 2006, 32(10): 1239-1243 (in Chinese).

[3] Department of Defense, U. S. Electromagnetic environmental effects and spectrum certification guidance for the acquisition process: MIL-HDBK-237C[S]. Washington, D. C.: Department of Defense, U. S., 2001.

[4] CHEN W Q, SU D L, HE X M, et al. A behavioral simulation and analysis method for communication system[C]//2009 20th International Zurich Symposium on Electromagnetic Compatibility. Piscataway, NJ: IEEE Press, 2009: 357-360.

[5] NIU M, SU D L, WEI Y, et al. Characterizing harmonic components of electromagnetic interference spectrum based on grey system model[C]//2014 XXXIth URSI General Assembly and Scientific Symposium (URSI GASS). Piscataway, NJ: IEEE Press, 2014: 1-4.

[6] ZHU K, SU D L, XU H, et al. A parasitic extraction method of PIN diode based on analysing ringing in SMPS[C]//2016 Asia-Pacific International Symposium on Electromagnetic Compatibility (APEMC). Piscataway, NJ: IEEE Press, 2016: 785-787.

[7] SHANG X, SU D L, XU H, et al. A noise source impedance extraction method for operating SMPS using modified LISN and simplified calibration procedure [J]. IEEE Transactions on Power Electronics, 2017, 32(6): 4132-4139.

[8] SU D L, ZHU K, NIU M, et al. A quantified method for characterizing harmonic components from EMI spectrum[C]//2015 Asia-Pacific International Symposium on Electromagnetic Compatibility (APEMC). Piscataway, NJ: IEEE Press, 2015: 561-564.

[9] SONG D W, SU D L, SHANG X F, et al. Harmonic component extraction algorithm of conducted emission spectrum based on EMD and periodic extensional model[C]//2016 Asia-Pacific International Symposium on Electromagnetic Compatibility (APEMC). Piscataway, NJ: IEEE Press, 2016: 797-799.

[10] 王冰切,苏东林,张晓雷. 飞机表面绕射射线的寻迹方法[J]. 北京航空航天大学学报, 2007, 33(7): 785-788.
WANG B Q, SU D L, ZHANG X L. Discrete ray path tracing on aircraft[J]. Journal of Beijing University of Aeronautics and

- Astronautics, 2007, 33(7): 785-788 (in Chinese).
- [11] CHEN X, HE S Y, YU D F, et al. Ray-tracing method for creeping waves on arbitrarily shaped nonuniform rational B-splines surfaces [J]. Journal of the Optical Society of America A, 2013, 30(4): 663-670.
- [12] RUAN Y C, ZHOU X Y, JESSIE Y C, et al. The UTD analysis to EM scattering by arbitrarily convex objects using ray tracing of creeping waves on numerical meshes [C] // IEEE Antennas and Propagation Society International Symposium. Piscataway, NJ: IEEE Press, 2008: 1-4.
- [13] SURAZHISKY V, SURAZHISKY T, KIRSANOV D, et al. Fast exact and approximate geodesics on meshes [J]. ACM Transactions on Graphics, 2005, 24(3): 553-560.
- [14] FU S, ZHANG Y H, HE S Y, et al. Creeping ray tracing algorithm for arbitrary NURBS surfaces based on adaptive variable step Euler method [J]. International Journal of Antennas and Propagation, 2015, 2015: 1-12.
- [15] JONATHAN R P, ERIC L S. Exact geodesics and shortest paths on polyhedral surfaces [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(6): 1006-1015.
- [16] MARTINEZ D, VELHO L, CARVALHO P C. Computing geodesics on triangular meshes [J]. Computers & Graphics, 2005, 29(5): 667-675.
- [17] 中国人民解放军总装备部. 军用设备和分系统电磁发射和敏感度要求与测量: GJB 151B—2013 [S]. 北京: 中国人民解放军总装备部, 2013.
- The General Armament Department of PLA. Electromagnetic emission and susceptibility requirements and measurements for military equipment and subsystems: GJB 151B—2013 [S]. Beijing: The General Armament Department of PLA, 2013 (in Chinese).

作者简介:

凌波 男, 博士研究生, 研究员。主要研究方向: 航空电子武器综合设计、电磁兼容与电磁环境。

林赞 男, 研究员。主要研究方向: 射频综合系统设计、电磁兼容防护设计。

孙宏涛 男, 硕士, 工程师。主要研究方向: 系统级电磁兼容设计、飞行器电磁环境效应测试技术。

苏东林 女, 博士, 教授, 博士生导师。主要研究方向: 电磁兼容、电磁环境效应、计算电磁学、射频微波电路与系统、新型飞行器机载共形/共用/小型化天线等。

Detection and identification method of electromagnetic interference elements in limited-space

LING Bo^{1,2}, LIN Yun², SUN Hongtao², SU Donglin^{1,*}

(1. School of Electronic and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;

2. AVIC Chengdu Aircraft Design & Research Institute, Chengdu 610091, China)

Abstract: It is an inevitable trend that more and more divided RF systems are integrated into one rack for the new generation airborne electronic equipment development. However, the “module-based” electromagnetic compatibility requirements, evaluation methods and other new electromagnetic compatibility issues have been inevitably brought out by the RF integrated technology’s change from “device-based” design to “module-based” design. According to the principals of electromagnetic interference (EMI) element, the EMI detection theory and testing techniques for emission features of RF module in limited-space of integrated racks are researched and discussed, as well as the simulating measures of RF function thread. A novel methodology is proposed for EMI detection and identification in limited-space from the far-field radiation features of integrated racks, near-field emission features of function module and the conducted current emission features of integrated racks’ interfacing cables. Many experimental tests are carried out upon one real airborne RF integrated racks by the research team, including far-field radiation tests, near-field emission tests of function module as well as cable current emission tests of the RF integrated racks. This method finally realizes the electromagnetic compatibility test of the function module under the integrated environment, and also is proved to be feasible and accurate.

Keywords: integrated RF; limited-space; RF module; electromagnetic compatibility; electromagnetic interference (EMI)

Received: 2017-05-22; **Accepted:** 2017-05-26; **Published online:** 2017-06-15 11:03

URL: kns.cnki.net/kcms/detail/11.2625.V.20170615.1103.001.html

Foundation item: National Natural Science Foundation of China (61427803)

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

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

DOI: 10.13700/j.bh.1001-5965.2017.0353

强噪声环境下自适应 CRPF 故障诊断方法

王进花^{1,2,3}, 曹洁^{1,2,3,*}, 李伟¹, 黄玲^{1,2,3}

(1. 兰州理工大学 电气工程与信息工程学院, 兰州 730050; 2. 甘肃省工业过程先进控制重点实验室, 兰州 730050;

3. 兰州理工大学 电气与控制工程国家级实验教学示范中心, 兰州 730050)

摘 要: 针对非线性非高斯系统在实际工作环境中受强噪声干扰影响导致的故障诊断精度低的问题,提出了一种状态转移密度方差自适应更新的代价评估粒子滤波(CRPF)故障诊断方法。通过设计观测值与先验状态之间的相关性判别函数,根据噪声和误差的大小实时自适应调整状态转移密度方差,增强算法对强噪声干扰的适应能力;研究了残差自适应阈值的设计方法,通过引入滑动窗求区间均值代替基于参数置信区间自适应阈值的均值和方差,在保证故障诊断准确性的前提下减少计算时间。以 160 MW 燃油机组为例,通过对不同强噪声环境下的汽包水位传感器故障诊断实例分析,结果表明该方法在复杂噪声环境下故障诊断的准确性得到了明显提高,同时减少了计算时间。

关键词: 故障诊断; 强噪声; 代价评估粒子滤波(CRPF); 自适应阈值; 漏诊率; 误诊率

中图分类号: TP277

文献标识码: A

文章编号: 1001-5965(2018)05-0923-08

现代工业过程规模越来越庞大,系统内部各部件联系非常紧密,随着对现代复杂设备故障诊断精度的要求提高,设备的可靠性及安全性备受学术界的广泛关注^[1-2]。基于模型的故障诊断方法是一种能够有效反映物理系统故障本质特征的方法,可充分利用系统内部的深层信息,有利于系统的故障隔离和辨识,获得了广泛研究。但因现代工业系统的复杂性及其部件工作特性的耦合影响,建立精确的系统模型极为困难,模型偏移误差与参数偏差是必然存在,同时受环境与人为干扰因素的影响,大大降低了故障诊断的准确性。

粒子滤波(Particle Filter, PF)是基于蒙特卡罗采样思想来实现状态的近似估计,理论上可以适用于任何非线性非高斯系统,因此在故障诊断领域获得了广泛的研究和应用^[3-4]。但是目前存

在的 PF 算法及其改进算法,在应用中假设噪声统计特性均为已知的高斯分布^[5-6]。事实上,噪声统计特性在实际应用中很可能是未知的,且可能随系统本身或者量测环境的变化而变化,现有的 PF 算法难以满足诊断系统精度的要求。代价评估粒子滤波(Cost Reference Particle Filter, CRPF)不需要已知系统噪声的统计特性,便可实现状态估计^[7],因此在一定程度上解决了复杂噪声背景下的状态估计问题,该方法在故障诊断和目标跟踪等领域得到了良好的应用。胡振涛和潘泉等^[8]提出了一种基于 CRPF 的残差似然比检验故障诊断算法,克服了外界随机扰动对于滤波精度的不利影响。卢锦等^[9]针对天波雷达背景噪声强度大和统计特性未知等特点,采用 CRPF 估计天波超视距雷达目标状态。Lim^[10]采用 CRPF

收稿日期: 2017-05-24; 录用日期: 2017-06-30; 网络出版时间: 2017-09-11 10:43

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20170911.1043.004.html

基金项目: 国家自然科学基金(61763028); 甘肃省自然科学基金(1506RJZA105, 1606RJZA145)

* 通信作者. E-mail: caoj@lut.cn

引用格式: 王进花, 曹洁, 李伟, 等. 强噪声环境下自适应 CRPF 故障诊断方法[J]. 北京航空航天大学学报, 2018, 44(5): 923-930.
WANG J H, CAO J, LI W, et al. An adaptive CRPF fault diagnosis method under strong noise condition[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 923-930 (in Chinese).

实现未知噪声统计特性的非线性系统动态状态估计。Yu^[11]提出了一种 H_∞ 和 CRPF 结合的滤波算法,提高了未知非高斯噪声的非线性动力学系统的状态估计精度。

CRPF 是以粒子的代价或风险为参考来实现重采样及状态估计,代价和风险的计算不需要已知噪声统计特性^[9],因此,在解决非线性、噪声统计特性未知问题中得到了良好应用。但在实际应用中仍然存在不足,CRPF 以当前时刻的状态预测值为均值,在一定的高斯分布范围内随机选取粒子完成状态更新^[10,12],但随着时间的不断推移,状态转移密度协方差逐渐趋于恒定值,对强噪声扰动或时变噪声的修正能力将减弱甚至消失,大大降低了算法对噪声环境的适应能力,将导致较大的误诊和漏诊。对复杂噪声环境下的故障诊断方法的研究是现代复杂系统故障诊断与故障预测研究中必须解决的基础科学问题,因此,本文研究对提高诊断系统的精确性,进一步提高现代复杂设备故障诊断的可靠性具有重要意义。

本文根据噪声的变化情况自适应调整状态转移密度方差,确定状态更新的合理区域,增强算法对时变噪声的适应能力;同时,利用统计方法设计残差判别函数的自适应阈值,提高故障检测的准确性。在此基础上,对强噪声背景下的非线性非高斯系统进行了故障诊断仿真分析。

1 非线性系统状态空间模型

考虑非线性非高斯随机状态空间模型如下:

$$\begin{cases} \mathbf{x}_k = g(\mathbf{x}_{k-1}, \boldsymbol{\theta}, \mathbf{u}_{k-1}) + \mathbf{v}_k \\ \mathbf{y}_k = h(\mathbf{x}_k, \boldsymbol{\theta}, \mathbf{u}_k) + \mathbf{w}_k \end{cases} \quad (1)$$

式中: k 为采样时刻; \mathbf{x}_k 为 k 时刻系统状态向量; \mathbf{u}_k 为 k 时刻系统输入; $\boldsymbol{\theta}$ 为系统参数变量; \mathbf{y}_k 为 k 时刻系统状态的量测向量; $g(\cdot)$ 和 $h(\cdot)$ 分别为系统状态转移函数和量测函数,它们都可以是非线性函数; \mathbf{v}_k 和 \mathbf{w}_k 分别为系统状态噪声和量测噪声,其统计特性未知。

假设系统可能存在 S 种故障模式,则共有 $S+1$ 个系统模型,其中 1 个模型为无故障模式。采用多模型方法表示的非线性系统状态空间模型如下:

$$\begin{cases} \mathbf{x}_k^i = g^i(\mathbf{x}_{k-1}^i, \boldsymbol{\theta}^i, \mathbf{u}_{k-1}^i) + \mathbf{v}_k^i \\ \mathbf{y}_k^i = h^i(\mathbf{x}_k^i, \boldsymbol{\theta}^i, \mathbf{u}_k^i) + \mathbf{w}_k^i \end{cases} \quad i = 0, 1, \dots, S \quad (2)$$

系统量测的预测输出为

$$\hat{\mathbf{y}}_k^i = h^i(\mathbf{x}_{k|k-1}^i, \boldsymbol{\theta}^i, \mathbf{u}_k^i) \quad (3)$$

式中: $\mathbf{x}_{k|k-1}^i$ 为系统状态的一步预测值; $\hat{\mathbf{y}}_k^i$ 为系统

量测的预测输出值。

系统残差可表示为

$$\mathbf{r}_k^i = \mathbf{y}_k^i - \hat{\mathbf{y}}_k^i \quad (4)$$

残差是系统实际输出和依据模型预测输出的差值,反映了实际系统和数学模型之间的不一致程度,主要由系统噪声和故障确定,当系统无故障时,残差的变化反映了系统中的噪声和干扰因素对系统输出的影响。当系统处于强噪声干扰环境,噪声会引起残差较明显的变化,可能引起故障误报或漏报。可见,获得准确的残差是提高故障诊断准确性的关键。因此,利用 CRPF 对未知噪声系统的状态估计优势,克服强噪声的影响而得到准确的系统状态的估计值,当估计值越接近真实值,则会减小噪声对残差信号的影响,提高故障诊断的准确性。

2 代价评估粒子滤波

在 CRPF 算法中用代价函数和风险函数表示粒子性能质量,同时引入遗忘因子,并基于代价最小化原则给出一种噪声统计特性未知情况下的粒子权重评价方法。

代价函数定义为

$$c(\mathbf{x}_{0:k}^{i_p} | \mathbf{y}_{1:k}, \lambda) = \lambda c(\mathbf{x}_{0:k-1}^{i_p} | \mathbf{y}_{1:k-1}, \lambda) + \Delta c(\mathbf{x}_k^{i_p} | \mathbf{y}_k) \quad (5)$$

将式(5)简写为: $c_k^{i_p} = \lambda c_{k-1}^{i_p} + \Delta c_k^{i_p}$ 。式中: i_p 为粒子索引; λ ($0 \leq \lambda \leq 1$) 为遗忘因子; $\Delta c_k^{i_p} = \|\mathbf{y}_k - h(\mathbf{x}_k^{i_p})\|^q$ ($q \geq 1$) 为代价增量。

风险函数定义为

$$R_k^{i_p} = \lambda c_{k-1}^{i_p} + \|\mathbf{y}_k - h(g(\mathbf{x}_{k-1}^{i_p}))\|^q \quad q \geq 1 \quad (6)$$

概率质量函数(PMF)也称类权值,计算如下:

$$\tilde{\pi}_k^{i_p} \propto \mu_1(R_k^{i_p}) = \frac{1}{(R_k^{i_p} - \min\{R_k^{i_p}\}_{i_p=1}^N + \delta)^\beta} \quad (7)$$

式中: $\delta, \beta > 0$, δ 是为了保证分母不为 0; N 为粒子数。

按照以上参数定义,CRPF 算法通过风险估计、选择、粒子传递和代价更新步骤,递推计算获得状态估计。算法步骤如下:

步骤 1 在 $k=0$ 时刻,从先验分布中获取 N 个样本, $\mathbf{x}_0^{i_p} \sim P_0(\mathbf{x}_0)$, 设定初始时刻粒子的代价 $c_0^{i_p}$ 和状态转移密度方差 σ_0^{2,i_p} , 则 $k=0$ 时刻样本和代价集合为 $\{\mathbf{x}_0^{i_p}, c_0^{i_p}\}_{i_p=1}^N$, $i_p = 1, 2, \dots, N$ 。

步骤 2 递归更新。

1) 分别按照式(6)和式(7)计算风险函数 $R_k^{i_p}$ 和概率质量函数 $\tilde{\pi}_k^{i_p}$ 。

2) 重采样。根据 $\tilde{\pi}_k^{i_p}$ 的大小,随机选择 N 个

粒子,生成粒子代价集 $\{\tilde{\mathbf{x}}_{k-1}^{i_p}, \tilde{\mathbf{c}}_{k-1}^{i_p}\}_{i_p=1}^N$ 。

3) 粒子更新。

$$\mathbf{x}_k^{i_p} \sim p_k(\mathbf{x}_k | \tilde{\mathbf{x}}_{k-1}^{i_p}) = N(g(\tilde{\mathbf{x}}_{k-1}^{i_p}), \sigma_k^{2,i_p} \mathbf{I}) \quad (8)$$

式中: \mathbf{I} 为维数与 \mathbf{x} 相同的单位矩阵。

状态转移密度的方差为

$$\sigma_k^{2,i_p} = \frac{k-1}{k} \sigma_{k-1}^{2,i_p} + \frac{\|\tilde{\mathbf{x}}_k^{i_p} - g(\tilde{\mathbf{x}}_{k-1}^{i_p})\|^2}{k \cdot \dim[\mathbf{x}]} \quad (9)$$

式中: $\dim[\mathbf{x}]$ 表示取 \mathbf{x} 的维数。

可见,在 CRPF 算法中, σ_k^{2,i_p} 是在线调整的。

4) 按照式(5)计算代价函数 $\mathbf{c}_k^{i_p}$,按式(10)计算 $\pi_k^{i_p}$,并归一化 $\pi_k^{i_p}$ 。

$$\pi_k^{i_p} \sim \mu_2(\mathbf{c}_k^{i_p}) = \frac{1}{(\mathbf{c}_k^{i_p} - \min\{\mathbf{c}_k^{i_p}\}_{i_p=1}^N + \delta)^\beta} \quad (10)$$

步骤3 状态估计。

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^{\text{CRPF}} = \sum_{i_p=1}^N \pi_k^{i_p} \mathbf{x}_k^{i_p}$$

从以上算法步骤可见,在 CRPF 算法的递推估计中不涉及对过程噪声与量测噪声的计算,因此在计算中不需要已知噪声统计特性,改善了 PF 中由于外界随机干扰对依据量测似然度评价的不利影响。从式(8)和式(9)的粒子更新过程可知,粒子的时间更新是在确定的概率区域中随机选取新粒子,此区域的方差 σ_k^{2,i_p} 是根据当前时刻的状态估计值与前一时刻的方差递推更新,当 $k \rightarrow \infty$ 时, $\frac{k-1}{k} \rightarrow 1, \frac{1}{k} \rightarrow 0$,这时 $\sigma_k^{2,i_p} \rightarrow \sigma_{k-1}^{2,i_p}$,可见随着时间的推移,系统状态的当前值失去对方差的调节作用,其方差趋于恒定,对于时变噪声或噪声波动较大的情况下,会大大降低算法的准确度。

3 自适应 CRPF 故障诊断方法

3.1 状态转移密度的方差自适应更新

由对 CRPF 的递推估计过程的分析可知,状态转移密度方差直接影响着状态预测的准确性,但在 CRPF 算法中,方差的递归更新随着时间的推移最终趋于恒定值,对于时变噪声出现随机波动的情况下,在确定的小范围中很有可能无法获得准确的状态预测。因此,本文考虑当前观测值与先验状态之间的互相关系数的大小,自适应调整状态转移密度方差,增强算法对时变噪声的适应能力。

测量值和先验状态之间的相关系数表示为

$$\gamma_k = E(g(\hat{\mathbf{x}}_{k-1}), \mathbf{y}_k) \quad (11)$$

状态转移密度的方差按照式(12)进行修正:

$$\sigma_k^{2,i_p} = \omega_{1,k} \sigma_{k-1}^{2,i_p} + \omega_{2,k} \frac{\|\tilde{\mathbf{x}}_k^{i_p} - g(\tilde{\mathbf{x}}_{k-1}^{i_p})\|^2}{\dim[\mathbf{x}]} \quad (12)$$

式中: $\omega_{1,k}$ 和 $\omega_{2,k}$ 为方差的自适应调节系数, $\omega_{1,k} = \gamma_k, \omega_{2,k} = 1 - \gamma_k$ 。相关系数大,说明状态噪声小,则 $\omega_{1,k} > \omega_{2,k}$,方差主要由前一时刻的方差 σ_{k-1}^{2,i_p} 决定,这样可适当增强粒子的扰动;相关系数小,说明噪声扰动较大,这时 $\omega_{1,k} < \omega_{2,k}$,方差主要由 $\|\tilde{\mathbf{x}}_k^{i_p} - g(\tilde{\mathbf{x}}_{k-1}^{i_p})\|^2$ 确定,利用当前时刻的噪声扰动对方差进行修正,提高算法对时变噪声背景下状态估计的准确性。

3.2 自适应阈值设计

残差是判断是否发生故障的主要依据,若 k 时刻的残差 $r_k > r_{th}$, r_{th} 为所设定的阈值,则说明系统发生了故障。最简单的阈值选择方法是将阈值设定为一个固定的常数,但是实际系统不可避免地存在建模误差、复杂噪声及干扰等不确定性因素,在这种情况下若仍采用传统的固定阈值来判定故障是否发生,很容易出现故障的误报和漏报。因此,依据系统不确定性的变化,采用自适应阈值可以避免上述情况的发生。文献[13-14]将统计学中参数置信区间的思想应用于自适应阈值的设计,假设残差的统计特性符合正态分布,残差的均值和方差求取方法分别为

$$\eta(u_j, t_k) = \frac{1}{n} \sum_{i=1}^n r_i(t_k) \Big|_{u_j} \quad (13)$$

$$\sigma_v^2(u_j, t_k) = \frac{1}{n-1} \sum_{i=1}^n (r_i(t_k) - \eta(u_j, t_k))^2 \Big|_{u_j} \quad (14)$$

式中: $\eta(u_j, t_k)$ 为输入为 u_j 时 t_k 时刻残差的均值; $\sigma_v^2(u_j, t_k)$ 为输入为 u_j 时 t_k 时刻残差的方差; n 为系统重复运行并测量系统残差的次数。

置信度为 $1 - \alpha$ 的残差的置信区间表示为

$$P\{\bar{\eta} - z\alpha < \eta < \bar{\eta} + z\alpha\} = 1 - \alpha \quad (15)$$

式中: α 为置信水平,通常设定置信度为 95% ~ 98%。若设 $\alpha = 0.03$,则置信度为 97%, $z = 2.17$,由式(16)求得自适应阈值为

$$r_{th,k} = \eta(u_j, t_k) \pm 2.17 \sigma_v(u_j, t_k) \quad (16)$$

但采用这种方法计算自适应阈值,为了得到当前时刻残差的均值和方差,在算法迭代的每一步都需要系统重复运行 n 次,这样计算量近似成指数增长,很难满足实际系统的实时性要求。本文受 Hashemi 和 Pisu^[15] 的启发,引入滑动窗通过求区间均值来解决上述问题。将残差的标准差表示为

$$(\text{STD}_{k+1})^2 = \sum_{i=k+2-M}^{k+1} \frac{(r_i - \mu_i)^2}{M} = (\text{STD}_k)^2 + \frac{(r_{k+1} - \mu_{k+1})^2}{M} - \frac{(r_{k+1-M} - \mu_{k+1-M})^2}{M} \quad (17)$$

$$\mu_k = \sum_{i=k+1-M}^k \frac{r_i}{M} = \mu_{k-1} + \frac{r_k}{M} - \frac{r_{k-M}}{M} \quad (18)$$

式中： μ_k 为残差的均值； M 为滑动窗宽度。

由式(17)和式(18)可见，在每一个采样周期，只需要在前一时刻的基础上通过当前时刻的残差对均值与方差进行更新，无需重复多次运行系统，大大减少运行时间。同时，对于随机噪声来说，在一定的时间内对某一个采样值重复执行多次，与对不同时间点的多个采样值重复执行多次，在该时间范围内噪声的分布趋于一致，因此，令 $\sigma_v^2(u_j, t_k) = (\text{STD}_k)^2$ ， $\eta(u_j, t_k) = \mu_k$ ，采用式(17)和式(18)分别计算残差的方差与均值，再由式(16)求得每一时刻的残差自适应阈值。

3.3 算法步骤

步骤1 初始化 $\mathbf{x}_0^{i_p} \sim P_0(\mathbf{x}_0)$ 、 $\mathbf{c}_0^{i_p}$ 和 σ_0^{2,i_p} ，生成初始样本代价集合 $\{\mathbf{x}_0^{i_p}, \mathbf{c}_0^{i_p}\}_{i_p=1}^N$ 。

步骤2 递归更新： $k=1, 2, \dots, K$ (总时间步数)。

1) for $i_p = 1, 2, \dots, N$ ，计算 $R_k^{i_p}$ 和 $\tilde{\pi}_k^{i_p}$ 。

2) 重采样。根据 $\tilde{\pi}_k^{i_p}$ 的大小，随机选择 N 个

粒子，生成粒子代价集 $\{\tilde{\mathbf{x}}_{k-1}^{i_p}, \tilde{\mathbf{c}}_{k-1}^{i_p}\}_{i_p=1}^N$ 。

3) 根据式(8)更新状态，由式(11)计算互相关系数，然后根据式(12)修正状态转移密度协方差。

4) 计算代价函数： $\mathbf{c}_k^{i_p} = \lambda \tilde{\mathbf{c}}_{k-1}^{i_p} + \|\mathbf{y}_k - h(\mathbf{x}_k^{i_p})\|^q$ ，按式(10)计算并归一化 $\pi_k^{i_p}$ 。

步骤3 估计状态 $\hat{\mathbf{x}}_k$ 。

步骤4 故障检测和隔离。将获得的状态估计值 $\hat{\mathbf{x}}_k$ 代入量测方程，计算得到 $\hat{\mathbf{y}}_k$ ，由式(4)计算残差，由式(16)~式(18)计算残差自适应阈值，若 $r_k > r_{th,k}$ ，则有故障发生；否则，无故障发生。然后，进一步进行故障隔离。

4 实例分析与讨论

采用一个 160 MW 燃油机组的动态模型^[16]为对象，以燃油调节阀开度、汽轮机调节阀开度和给水调节阀开度为输入，选取汽包压力、汽包液体密度为状态变量，汽包水位作为观测变量，分别在伽马噪声和高斯混合噪声影响下对汽包水位传感器的 3 种故障进行实例分析，验证本文方法的性能。该燃油机组模型的离散方程为

$$\begin{cases} x_{1,k} = x_{1,k-1} - A_1 \Delta t + v_{1,k} \\ x_{2,k} = x_{2,k-1} + A_2 \Delta t + v_{2,k} \\ y_k = 0.05B + w_k \end{cases}$$

式中：

$$A_1 = a_{11}u_{2,k}x_{1,k-1}^{9/8} - a_{12}u_{1,k} + a_{13}u_{3,k}$$

$$A_2 = [a_{21}u_{3,k} - (a_{22}u_{2,k} - a_{23})x_{1,k-1}]/85$$

$$B = b_1x_{2,k} + 100a_{cs} + q_c/9 - b_2$$

$$a_{cs} = \frac{b_3x_{2,k}(b_4x_{1,k} - b_5)}{x_{2,k}(b_6 - b_7x_{1,k})}$$

$$q_c = (b_8u_{2,k} - b_9)x_{1,k} + b_{10}u_{1,k} - b_{11}u_{3,k} - b_{12}$$

其中： x_1 为汽包压力； x_2 为汽包液体密度； u_1 为燃油调节阀开度； u_2 为汽轮机调节阀开度； u_3 为给水调节阀开度； y 为汽包水位； v_k 为状态噪声； w_k 为量测噪声，它们都是统计特性未知的非高斯噪声。

针对汽包水位传感器故障进行实验，假设传感器有 3 种故障模式：

故障模式 1 恒偏差故障， $y_k = 0.05B + 0.5 + w_k$ 。

故障模式 2 恒增益故障， $y_k = 0.08B + w_k$ 。

故障模式 3 卡死故障， $y_k = w_k$ 。

设置初始状态 $\mathbf{x}_0 = [108 \quad 428]^T$ ，离散步长 $\Delta t = 0.1$ s，粒子数 $N = 500$ ，传感器采样频率为 1 Hz，其他参数： $a_{11} = 0.0018$ ， $a_{12} = 0.9$ ， $a_{13} = 0.15$ ， $a_{21} = 141$ ， $a_{22} = 1.1$ ， $a_{23} = 0.19$ ， $b_1 = 0.131$ ， $b_2 = 0.068$ ， $b_3 = 0.00154$ ， $b_4 = 0.8$ ， $b_5 = 25.6$ ， $b_6 = 1.0394$ ， $b_7 = 0.00123$ ， $b_8 = 0.854$ ， $b_9 = 0.147$ ， $b_{10} = 45.59$ ， $b_{11} = 2.514$ ， $b_{12} = 2.096$ ， $\mathbf{u}_i = [0.3 \quad 0.4 \quad 0.5]^T$ 。状态初始先验分布 $\mathbf{x}_0^{i_p} \sim N(\mathbf{x}_0, \Sigma_0)$ ， $\Sigma_0 = \text{diag}(0.01, 0.01)$ ， $\mathbf{c}_0^{i_p} = \mathbf{0}$ ， $\sigma_0^{2,i_p} = [0.01 \quad 0.01]^T$ ，滑动窗宽度 $M = 20$ 。

本文方法的评价指标选择平均绝对误差、故障误报率和漏报率，评价指标定义如下：

平均绝对误差 MAE 为

$$\text{MAE} = \frac{1}{N_s T} \sum_{s=1}^{N_s} \sum_{k=1}^T |\mathbf{x}_{sk} - \hat{\mathbf{x}}_{sk}|$$

式中： \mathbf{x}_{sk} 和 $\hat{\mathbf{x}}_{sk}$ 分别为第 s 次仿真第 k 步状态的实际值和估计值；总的仿真次数为 $N_s = 50$ ；一次仿真中的时间步数 $T = 3000$ 。

故障漏报率 p_m 和误报率 p_f 为

$$\begin{cases} p_m = \frac{E}{FG} \\ p_f = \frac{C}{FD} \end{cases}$$

式中： F 为系统总共的运行次数； E 为发生故障时残差小于阈值的采样点总数； G 为系统一次运行的总采样点数； C 为系统无故障发生且残差值大于阈值的采样点总数； D 为系统无故障时的采样点总数。

4.1 实验 1

本文所设定的偏差故障为 0.5 的小偏差，增

益故障的增益系数变化仅为0.03,噪声对这种小故障的检测和诊断结果影响更为严重,采用伽马噪声和高斯混合噪声来模拟工程实际中的复杂噪声,采用CRPF算法和自适应代价评估粒子滤波(ACRPF)算法对汽包压力 x_1 和汽包液体密度 x_2 进行估计,验证本文方法在强噪声环境下的状态估计精度。

采样周期为0.1s,步长为3000,总仿真时长为300s,粒子数为500。设定 $0 \leq t \leq 100$ s时,系统处于正常模态; $100 \text{ s} < t \leq 200$ s时,系统处于故障模态1; $200 \text{ s} < t \leq 300$ s时,系统处于故障模态2。

实验中所采用的噪声如下:

伽马噪声:

$$\begin{cases} v_k \sim 0.09 \times 10^{-2} \Gamma(0.25, 0.5) \\ w_k \sim 0.16 \times 10^{-2} \Gamma(0.25, 0.5) \end{cases}$$

高斯混合噪声:

$$\begin{cases} v_k \sim 10^{-1} (0.8N(0.2, 0.5) + 0.5N(0.02, 0.08) + 0.4N(0.01, 0.1)) \\ w_k \sim 10^{-1} (0.01N(0.02, 0.5) + 0.05N(0.02, 0.08) + 0.04N(0.01, 0.1)) \end{cases}$$

图1和图2分别为伽马噪声和高斯混合噪声背景下,CRPF算法和改进的ACRPF算法对汽包压力 x_1 和汽包液体密度 x_2 的估计误差, $0 \sim 100$ s系统处于正常工作状态,2种算法的估计误差都比较小。可见,由于ACRPF算法能够根据状态估计误差自适应调整状态转移密度方差,增强了对强噪声扰动的自适应修正能力,因此,对 x_1 和 x_2 的估计精度均高于改进前的CRPF算法。尤其对于 x_2 ,在100s系统出现故障,采用原有CRPF算法估计误差急剧增加,而本文改进算法的估计

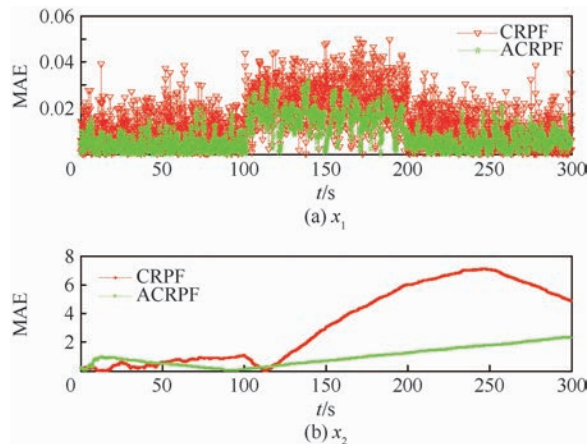


图1 伽马噪声背景下2种算法对 x_1 和 x_2 的跟踪误差对比

Fig.1 Tracking error comparison of x_1 and x_2 between two algorithms under condition of Gamma noise

精度得到了大幅提高。表1为图1和图2中 x_1 和 x_2 的平均绝对误差,数据同样说明了本文算法的状态估计准确性均得到了明显提高。

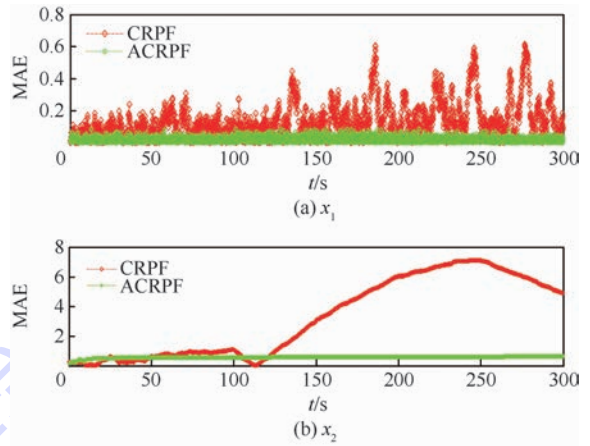


图2 高斯混合噪声背景下2种算法对 x_1 和 x_2 的跟踪误差对比

Fig.2 Tracking error comparison of x_1 and x_2 between two algorithms under condition of Gaussian mixture noise

表1 不同噪声背景下 x_1, x_2 的平均绝对误差

Table 1 Mean absolute errors of x_1 and x_2 under different noise conditions

状 态	伽马噪声		高斯混合噪声	
	改进前	改进后	改进前	改进后
x_1	0.0111	0.0082	0.1952	0.0095
x_2	3.5750	1.7071	4.3659	0.8860

4.2 实验2

分别在伽马噪声和高斯混合噪声影响下对上述系统进行故障诊断,对本文ACRPF故障诊断方法的性能进行验证。设残差 $r_0 = |y - \hat{y}_{\text{正常}}|$,若 $r_0 > r_{th}$,说明有故障发生;然后分别计算 $r_1 = |y - \hat{y}_{\text{故障1}}|$, $r_2 = |y - \hat{y}_{\text{故障2}}|$, $r_3 = |y - \hat{y}_{\text{故障3}}|$,若 $r_i = 0 (i=1,2,3)$,则说明发生了模态为 i 的故障。系统工作过程中的故障状态设定和噪声形式同实验1。

图3为文献[13-14]的自适应阈值与本文自适应阈值分别在伽马噪声和高斯混合噪声下的对比。对比2种自适应阈值曲线,其变化趋势和范围均非常接近;表2中,第2列和第3列、第4列和第5列相应的漏报率非常接近。可见,2种自适应阈值的漏报率基本相同,而误报率近似值均为0。因此,采用本文改进的自适应阈值完全可等价文献[13-14]中的阈值,而本文自适应阈值的时间计算复杂度要比文献[13-14]小一个数量级,大大节约了计算时间。

图4和图5分别为伽马噪声和高斯混合噪声背景下对160MW燃油机组的2种故障进行检测和隔离的结果。由残差 r_0 可判断,系统在100s

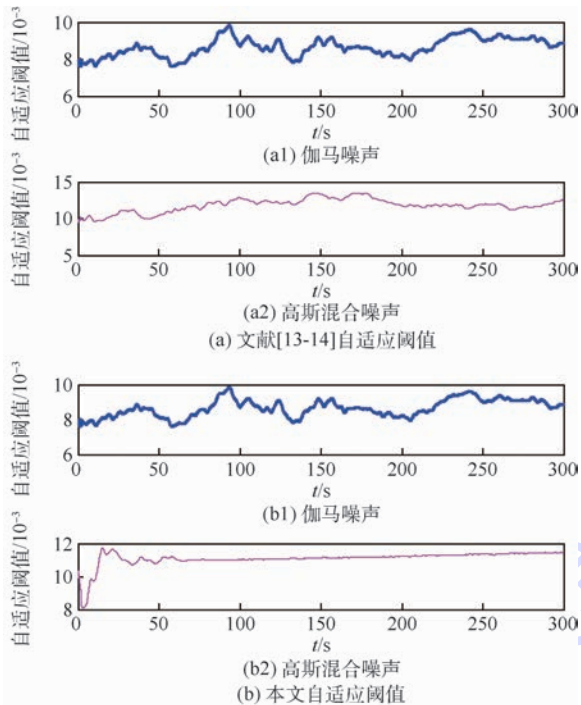


图3 自适应阈值对比

Fig. 3 Comparison of adaptive thresholds

表2 故障漏报率和时间计算复杂度比较

Table 2 Comparison of missed diagnosis rates and time computation complexity

参数	CRPF 故障诊断方法		ACRPF 故障诊断方法	
	文献[13-14]	本文	文献[13-14]	本文
伽马噪声下漏报率	0.0047	0.0045	0.0038	0.0037
高斯混合噪声下漏报率	0.0063	0.0061	0.0045	0.0046
时间计算复杂度	$O(n^3)$	$O(n^2)$	$O(n^3)$	$O(n^2)$

以前处于正常工作状态,100 s 以后均处于故障状态;为了充分说明故障检测的有效性,在图 4 和图 5 中 0 ~ 100 s 时间段内,选择残差和自适应阈值的极小值点分别标注了残差 r_0 的坐标值以及对应的自适应阈值的坐标值。图 4 中,在时刻 17.3 s 对应的残差 r_0 的值为 0.002 63,自适应阈值为 0.008 872;在时刻 98.1 s 对应的 r_0 为 0.002 21,自适应阈值为 0.01。图 5 中,在时刻 3.4 s 对应的 r_0 为 0.002 876,自适应阈值为 0.006 455。可见, r_0 的值均小于对应时刻的自适应阈值,在图 5 中的 25 s 和 84.3 s 满足同样的结论,说明系统处于正常工作状态。在时间段 100 ~ 200 s,比较 $r_1 \sim r_3$ 可见, r_1 更接近于 0,说明系统发生了模态 1 故障,即恒偏差故障。为了更清楚比较残差大小,在图 4 的 199.3 s 和图 5 的 180、199.6 s 处分别标注了图中最接近的残差 r_1 和 r_2 的值,均满足 $r_1 < r_2$;在时间段 200 ~ 300 s, r_2 最接

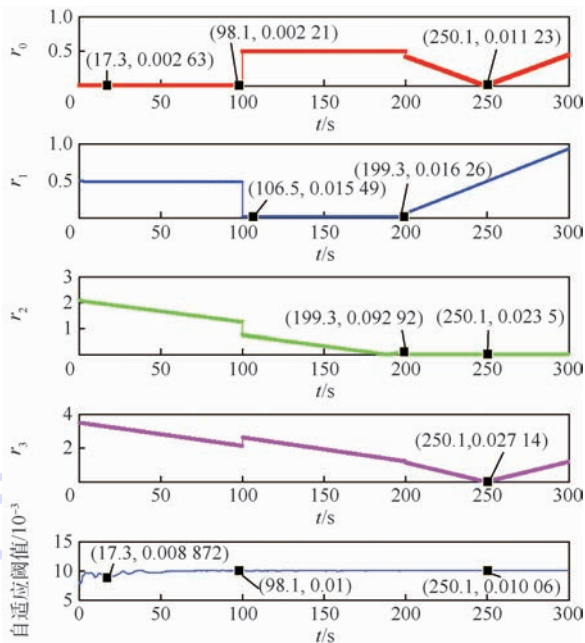


图4 伽马噪声背景下故障检测及隔离

Fig. 4 Fault detection and isolation under condition of Gamma noise

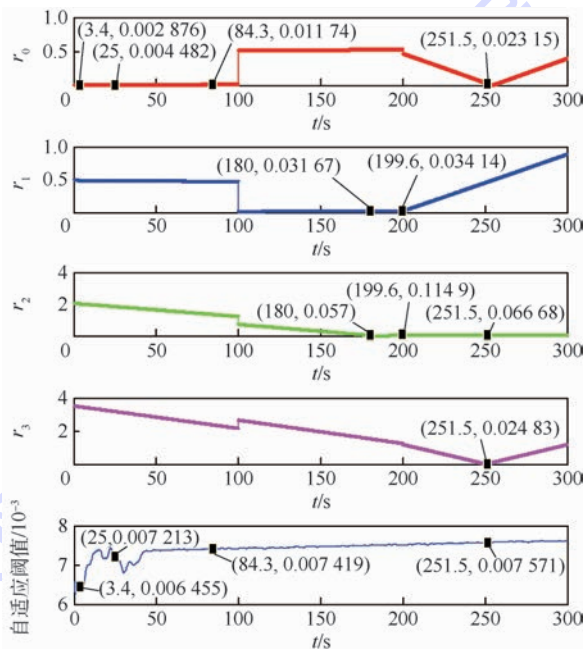


图5 高斯混合噪声背景下故障检测及隔离

Fig. 5 Fault detection and isolation under condition of Gaussian mixture noise

近 0,说明系统发生了模态 2 故障,即恒增益故障。虽然在图 4 的 250.1 s 和图 5 的 251.5 s 处 r_0 的值较小,但依然大于自适应阈值。可见,ACRPF 故障诊断方法在不同噪声下均可实现故障的准确检测和隔离。

表 2 对比了 CRPF 故障诊断方法和 ACRPF 故障诊断方法的漏报率,实验中系统运行 50 次,

每次运行的采样点数为 3 000,通过对所有采样点在 50 次运行中出现漏报的采样点数进行统计,并利用漏报率公式计算得到漏报率。总体来看,ACRPF 故障诊断方法在不同噪声下的漏报率均小于 CRPF 故障诊断方法的漏报率,分别对比表 2 的前 2 列和后 2 列的数据,可见本文的改进自适应阈值和文献[13-14]所得到的故障漏报率基本相同,但本文改进自适应阈值的时间计算复杂度低。结果表明,ACRPF 故障诊断方法在强噪声背景下,可提高故障诊断准确性,减少系统的运行时间。

5 结 论

1) CRPF 算法可在未知噪声统计特性的条件下实现对非线性非高斯系统状态的递推估计,为复杂非线性系统的故障诊断提供了一种可借鉴的方法。

2) 本文通过分析 CRPF 算法对强噪声干扰下状态估计的性能,设计相关性判别函数实现对 CRPF 的状态转移密度方差的自适应调整,克服了 CRPF 随着时间的递推,对强噪声及时变噪声的修正能力逐渐消失而导致的误差变大甚至发散的问题。

3) 引入滑动窗利用求区间均值的方法对残差的自适应阈值进行了改进,解决了在计算残差均值和方差时系统重复运行而导致的耗时间问题。

4) 应用本文方法,在不同的强噪声背景下对非线性非高斯系统进行了故障检测和隔离,对漏报率和误报率进行了分析,提高了故障诊断准确性并可减少运行时间,具有重要的理论意义和工程应用价值。

参考文献 (References)

- [1] JIN X H, QIAO W, PENG Y Y, et al. Quantitative evaluation of wind turbine faults under variable operational conditions[J]. IEEE Transactions on Industry Applications, 2016, 52(3): 2061-2069.
- [2] PARK J, HA J M, OH H, et al. Model-based fault diagnosis of a planetary gear: A novel approach using transmission error[J]. IEEE Transactions on Reliability, 2016, 65(4): 1830-1841.
- [3] SUN Y S, RAN X R, LI Y M, et al. Thruster fault diagnosis method based on Gaussian particle filter for autonomous underwater vehicles[J]. International Journal of Naval Architecture and Ocean Engineering, 2016, 8(3): 243-251.
- [4] 万磊, 杨勇, 李岳明. 水下机器人执行器的高斯粒子滤波故障诊断方法[J]. 上海交通大学学报, 2013, 47(7): 1072-1076.
WAN L, YANG Y, LI Y M. Actuator fault diagnosis of automatic under water vehicle using Gaussian particle filter[J]. Journal of Shanghai Jiaotong University, 2013, 47(7): 1072-1076 (in Chinese).
- [5] 曹洁, 李伟, 李军, 等. 强噪声背景下鲁棒的说话人跟踪方法[J]. 华中科技大学学报(自然科学版), 2015, 43(10): 363-366.
CAO J, LI W, LI J, et al. Robust speaker tracking under the background of strong noise[J]. Journal of Huazhong University of Science and Technology(Natural Science Edition), 2015, 43(10): 363-366(in Chinese).
- [6] 王强, 刘永葆, 贺星, 等. 噪声方差自适应修正的混合系统故障诊断方法[J]. 振动与冲击, 2016, 35(8): 14-20.
WANG Q, LIU Y B, HE X, et al. A hybrid-system fault-diagnosis method based on noise-variance adaptive correction[J]. Journal of Vibration and Shock, 2016, 35(8): 14-20 (in Chinese).
- [7] MÍGUEZ J, BUGALLO M F, DJURIC' P M. A new class of particle filters for random dynamical systems with unknown statistics[J]. EURASIP Journal on Applied Signal Processing, 2004(15): 2287-2294.
- [8] 胡振涛, 潘泉, 杨峰, 等. 基于 CRPF 的残差似然比检验故障诊断算法[J]. 系统工程与电子技术, 2009, 31(12): 3022-3025.
HU Z T, PAN Q, YANG F, et al. Residual likelihood ratio test for fault diagnosis based on cost reference particle filter[J]. Systems Engineering and Electronics, 2009, 31(12): 3022-3025(in Chinese).
- [9] 卢锦, 苏洪涛, 水鹏朗. 采用代价参考粒子滤波器估计天波雷达目标状态[J]. 西安电子科技大学学报, 2013, 40(5): 20-25.
LU J, SU H T, SHUI P L. Target state estimation of the over-the-horizon radar using the cost reference particle filter[J]. Journal of Xidian University, 2013, 40(5): 20-25 (in Chinese).
- [10] LIM J. Particle filtering for nonlinear dynamic state systems with unknown noise statistics[J]. Nonlinear Dynamics, 2014, 78(2): 1369-1388.
- [11] YU Y H. Combining H_{∞} filter and cost-reference particle filter for conditionally linear dynamic systems in unknown non-Gaussian noises[J]. Signal Processing, 2013, 93(7): 1871-1878.
- [12] LIM J, KIM T, HONG D. Estimating the number of competing terminals by cost-reference particle filtering in non-saturated wireless-LAN[J]. Telecommunication System, 2016, 62(3): 519-527.
- [13] SHI Z, GU F, LENNOX B, et al. The development of an adaptive threshold for model-based fault detection of a nonlinear electro-hydraulic system[J]. Control Engineering Practice, 2005, 13(11): 1357-1367.
- [14] 蒋栋年, 李炜. 基于自适应阈值的粒子滤波非线性系统故障诊断[J]. 北京航空航天大学学报, 2016, 42(10): 2099-2106.
JIANG D N, LI W. Fault diagnosis of particle filter nonlinear systems based on adaptive threshold[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(10): 2099-2106(in Chinese).

- [15] HASHEMI A, PISU P. Adaptive threshold-based fault detection and isolation for automotive electrical systems[C] // 2011 9th World Congress on Intelligent Control and Automation. Piscataway, NJ: IEEE Press, 2011: 1013-1018.
- [16] 郭健彬, 纪丁菲, 王鑫, 等. 混杂系统粒子滤波混合状态估计及故障诊断算法[J]. 系统工程与电子技术, 2015, 37(8): 1936-1942.
- GUO J B, JI D F, WANG X, et al. Hybrid state estimation and fault diagnosis algorithm of hybrid systems using particle filter [J]. Systems Engineering and Electronics, 2015, 37(8): 1936-1942 (in Chinese).
- 作者简介:**
王进花 女, 博士研究生, 副教授。主要研究方向: 故障诊断、非线性滤波方法及应用。
曹洁 女, 教授, 博士生导师。主要研究方向: 智能信息处理、非线性理论及应用。

An adaptive CRPF fault diagnosis method under strong noise condition

WANG Jinhua^{1,2,3}, CAO Jie^{1,2,3,*}, LI Wei¹, HUANG Ling^{1,2,3}

(1. College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China;

2. Key Laboratory of Gansu Advanced Control for Industrial Processes, Lanzhou 730050, China;

3. National Experimental Teaching Center of Electrical and Control Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: Aimed at the problem of low precision in fault diagnosis of nonlinear non-Gaussian system due to serious noise interference under the actual working condition, this paper puts forward a new fault diagnosis method, which can adaptively update the state transition density variance of a cost reference particle filter (CRPF). By designing the correlation discriminant function between the measurement value and the prior state, the variance of the state transition density was adjusted adaptively according to the magnitudes of noise and error, and the adaptability of the algorithm to strong noise interference is dramatically enhanced. Furthermore, the method for designing adaptive threshold of residual was studied, and the sliding window was also introduced to calculate the mean of interval instead of the mean and variance of the adaptive threshold based on parameter confidence interval, which was expected to reduce the calculation time under the premise of ensuring the accuracy of fault diagnosis. Taking 160 MW fuel unit as an example, drum level sensor fault diagnoses under different strong noise conditions were analyzed. From the results, it is found that the accuracy of fault diagnosis in the complex noise environment is obviously improved and the computation time is greatly reduced.

Keywords: fault diagnosis; strong noise; cost reference particle filter (CRPF); adaptive threshold; missed diagnosis rate; misdiagnosis rate

Received: 2017-05-24; **Accepted:** 2017-06-30; **Published online:** 2017-09-11 10:43

URL: kns.cnki.net/kcms/detail/11.2625.V.20170911.1043.004.html

Foundation items: National Natural Science Foundation of China (61763028); Natural Science Foundation of Gansu Province, China (1506RJZA105, 1606RJZA145)

* **Corresponding author.** E-mail: caoj@lut.cn

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

DOI: 10.13700/j.bh.1001-5965.2017.0374

基于 CPS 方法的工业机器人系统



陈友东*, 常石磊, 冯强国

(北京航空航天大学 机械工程及自动化学院, 北京 100083)

摘 要: 基于信息物理融合系统(CPS)的工业机器人系统是智能制造的使能技术和设备,提出了一种基于CPS方法的工业机器人系统,实现了物理世界和信息世界的融合。系统分为物理层、网络层、控制层和应用层。物理层的数据通过网络层上传给控制层,更新物理世界在信息世界的信息;应用层对其进行优化、决策;控制层将决策转变为设备的控制信息,通过网络层实现物理层的控制。采用该方法实现的工业机器人系统,在ER3A-C60型工业机器人上,采用自制的控制系统和MICRO-6013CM型工业相机进行验证,证实了该方法的可行性。

关键词: 信息物理融合系统(CPS); 工业机器人; 控制系统; Modbus; 仿真

中图分类号: TP242.2

文献标识码: A

文章编号: 1001-5965(2018)05-0931-08

目前,生产越来越趋向于小批量和个性化,制造周期越来越短,制造成本越来越低。对柔性生产来说,生产过程根据加工需要调整的需求越来越迫切。信息物理融合系统(Cyber-Physical System, CPS)深度融合了计算、控制和通信技术^[1-2],能够满足生产柔性化和开放化的要求^[3]。CPS是计算和物理过程的集成,嵌入式计算机和网络监视控制的物理过程,通过物理过程反馈的数据影响计算,反之亦然^[4]。通过这样的融合,可以实现诸如智能制造、智能汽车、智能建筑和智能电网等应用场景^[3-5]。工业机器人作为智能制造的使能设备,基于CPS方法的工业机器人系统可以提高工业机器人的柔性和开放性。

关于工业机器人CPS的研究,Michniewicz和Reinhart^[6]将模块化机器人单元灵活地用在工厂级别的任务规划上,但没有系统地分析机器人单元的具体实现,没有描述工业机器人控制方式。Chen等^[7]将指令作为重要的变量,将制造资源以及工作任务映射到操作状态上,设计了一种机床

设备上的CPS架构。Chen等^[8]在研究现有机器人结构的基础上设计了一种CPS结构,该设计仅停留在概要设计上,并没有实现。Luo和Kuo^[9]开展了7自由度机器人物理信息融合的研究,根据功能分为运动服务、障碍检测、目标抓取和人类协作检测等几大模块,建立起设备、虚拟、云端及人的系统,该系统可以灵活智能地实现7自由度机器人控制,提高了机器人控制的智能化,但没有将控制系统作为整体进行讨论。Colombo和Karnouskos^[10]将基于服务的系统引入工业自动化环境中,将工业机器人等工业设备在云端映射为服务模块,通过各个服务模块的相互调用,优化得到最优方案。

CPS系统一般采用分层结构^[11]。Tan等^[12]提出一种三层架构,传感器和执行单元直接作用于物理世界,中间采用下一代网络进行信息传输,由CPS单元进行信息处理。王小乐等^[13]提出一种面向服务的CPS体系架构,包含节点层、网络层、资源层和服务层。根据工业机器人的应用场

收稿日期: 2017-06-05; 录用日期: 2017-07-07; 网络出版时间: 2017-10-19 11:12

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20171019.1112.002.html

基金项目: 国家科技支撑计划(2015BAF01B04); 北京市科技计划(D161100003116002)

* 通信作者. E-mail: chenyd@buaa.edu.cn

引用格式: 陈友东, 常石磊, 冯强国. 基于CPS方法的工业机器人系统[J]. 北京航空航天大学学报, 2018, 44(5): 931-938.

CHEN Y D, CHANG S L, FENG Q G. Industrial robot system based on CPS approach[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 931-938 (in Chinese).

景^[14],本文提出了一种基于CPS方法的工业机器人系统,该系统分为物理层、网络层、控制层和应用层,实现了信息世界与物理世界的融合。物理层的数据通过网络层上传给控制层,更新物理世界在信息世界的信息。应用层根据此信息,对其进行分析、优化决策,控制层将决策转变为设备的控制信息,通过网络层,实现物理层的控制。

1 系统组成

CPS是信息世界和物理世界的融合。基于CPS方法的工业机器人系统分为物理层、网络层、控制层和应用层,如图1所示。物理层通过网络层将信息传给控制层,控制层实现对物理世界的控制。优化、故障诊断和决策在应用层实现,提

高了系统效率和智能化水平。

物理层由机器人单元、传感器、辅助设备、操作对象和人组成。机器人单元包括机器人本体、伺服驱动、控制器系统、控制PC和工具等;传感器在物理层采集数据传递给控制层,应用层对这些数据进行优化和决策;辅助设备包含其他与机器人协同工作的设备,如CNC、AGV和传送带等。机器人不再单独控制,而是与其周围设备统一控制。

网络层连接物理世界和信息世界,起到通信作用。信息世界包括控制层和应用层。控制层是设备的具体信息存储及设备控制部分,包括系统内设备的控制模块和传感器的处理模块。这些模块可以通过网络快速添加到系统中,提供给应用层使用。应用层提供系统的仿真、故障诊断、任务优化和决策等。



图1 基于CPS方法的工业机器人系统

Fig. 1 Industrial robot system based on CPS approach

2 系统实现

基于CPS方法的工业机器人系统具体组成如图2所示。应用层包括仿真优化和交互界面等。仿真优化将任务进行动态模拟,检查任务的可行性;在机器人实际运行时根据采集到的信息实时显示机器人当前的运行状态,监测机器人的运行过程。任务路径可以手工输入、传感器输入、CAD/CAM导入,通过任务解析模块提取任务轨迹,任务路径在仿真模块中仿真,并根据仿真结果

优化任务,调整轨迹;指令生成模块根据调整的轨迹生成机器人的运动指令,供机器人执行。仿真模块在机器人运行过程中,通过采集的机器人位姿和状态信息,呈现机器人运动和状态,实现现实世界与虚拟世界的同步。

交互界面是人与系统交互的接口,显示机器人状态,给人提供操作的界面,是实现人与系统交互的重要部分。人的参与是CPS系统的重要环节,通过交互界面对系统下达指令、对系统进行监督;系统也可以通过传感器得到人的运动、位置等

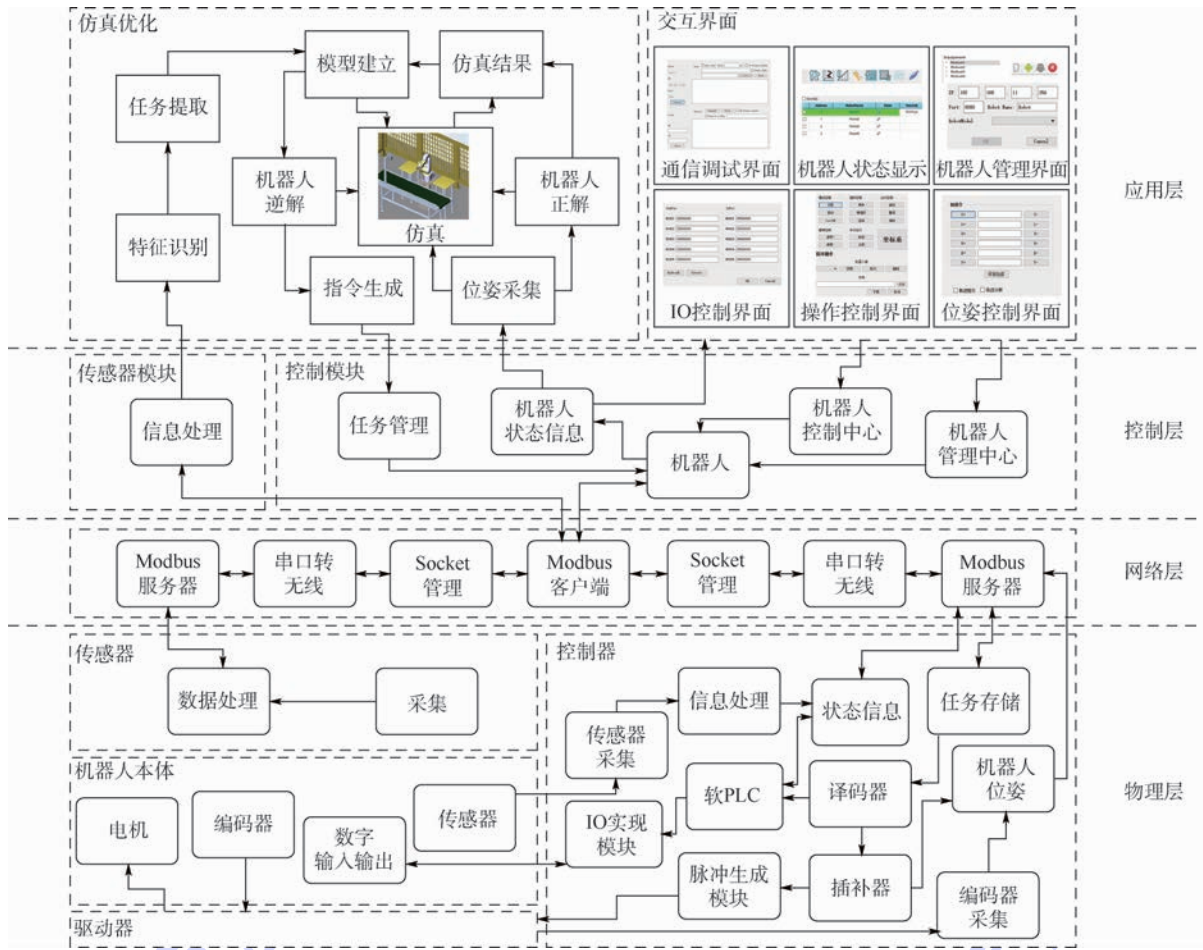


图 2 基于 CPS 方法的机器人系统的具体组成

Fig. 2 Specific components of industrial robot system based on CPS approach

信息,实现与人的互动。

控制层实现传感器采集信息的处理,提供给应用层使用;应用层实现真实机器人在信息世界的表示,并管理信息世界的机器人,管理仿真生成的任务,为交互界面提供机器人状态信息以及实现交互界面的控制指令。

网络层实现控制层与硬件层之间的通信,实现从 TCP/IP (Modbus) 通信协议到串口 (Modbus) 协议的转换,以及 Modbus 数据帧的打包和解析。控制器实现机器人的动作和操作。控制器对运动指令进行译码,对轨迹插补,最终生成脉冲发送到驱动器,控制机器人运动。采集编码器的值,获得机器人的位姿;机器人本体内部传感器采集信息传递给控制器,进行控制器内部的处理,得到相应的机器人状态。

在实际执行过程中,物理层需要控制物理设备,如工业机器人,对实时性要求较高,采用实时操作系统实现其强实时的要求;信息世界对信息进行优化、决策、发布控制指令,对实时性要求不高,采用 Windows 操作系统满足弱实时的要求。

2.1 物理层

物理层的硬件实现如图 3 所示。机器人控制的核心是机器人控制器。控制器控制机器人本体,接收反馈信息,并与控制层通信。为了实现高性能低成本的控制,机器人控制器采用嵌入式架构。设计时将硬件与软件模块化,提高了系统可维护性、扩展性。控制器硬件由 ARM (TI 的 AM3517) 和 FPGA (Altera 的 Cyclone IV 系列的 EP4CE6F17C8) 组成。ARM 完成计算和逻辑控制,FPGA 完成脉冲实现以及 IO 接口。根据功能,将控制器集成在核心板、通信板、IO 板、电源板以及底板 5 个独立的模块上。

采用开源的实时操作系统 $\mu\text{C}/\text{OS}-\text{II}$ 作为控

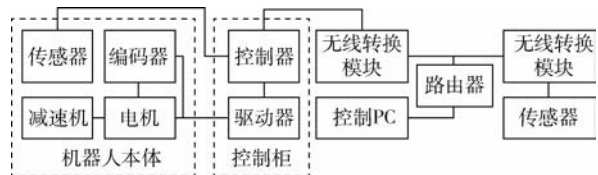


图 3 机器人控制系统物理层

Fig. 3 Physical layer of robot control system

制器的软件平台,以满足硬件设备对实时性的要求。软件系统分为译码器、插补器、软 PLC、传感器处理和通信 5 个模块,采用实时组件实现了系统的开放性,如图 4 所示。译码器将任务文档解析出控制指令;插补器根据不同的插补方式插补出机器人的控制位姿规划;软 PLC 实现内置 PLC 功能,读取输入,控制输出;通信模块采用 Modbus 协议;传感器处理模块处理机器人内部传感器上传信息,得到机器人状态信息。

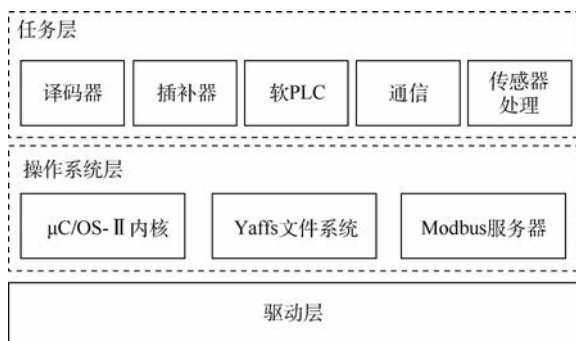


图 4 控制器软件结构

Fig. 4 Software structure of controller

2.2 网络层

网络层是物理层和控制层、应用层之间的桥梁,实现它们之间的通信和联系。控制器、传感器通过串口无线转换模块与控制 PC 进行通信,它们之间采用 TCP/IP 通信协议,应用协议采用 Modbus;控制器、传感器到无线转换模块间采用串口通信,通信协议采用 Modbus 协议。控制器和传感器具有 Modbus 服务器,控制 PC 具有 Modbus 客户端。客户端与服务器的通信流程如图 5 所示。

Modbus 客户端将数据和操作,以及与之对应的功能码打包成 Modbus 数据帧,数据帧通过

无线 TCP/IP 通信协议发送至串口无线转换模块。串口无线转换模块将 TCP/IP 数据包转换为 RS232 的数据包,服务器对数据包进行解析,实现数据读取和操作,并生成响应包,发回到客户端,客户端对 Modbus 数据解析,得到所需的数据,或者确认相关的操作。Modbus 协议的主要功能码如表 1 所示^[15]。

表 1 Modbus 协议的主要功能码^[15]

Table 1 Function code mainly used in Modbus process^[15]

功能码	备注
1	读取线圈状态取得一组逻辑线圈的当前状态
2	读取输入状态取得一组开关输入的当前状态
3	一个或多个保持寄存器中取得当前的二进制值
4	在一个或多个输入寄存器中取得当前的二进制值
5	强置一个逻辑线圈的通断状态
6	把具体二进制装入一个保持寄存器
15	强置一串连续逻辑线圈的通断
16	把具体二进制装入一串连续的保持寄存器支持浮点数
20	读通用寄存器
21	写通用寄存器

2.3 控制层

控制层采用 Qt 4.8.6 实现。控制层与物理层通信,给应用层提供基础,同步机器人的信息和状态,处理传感器提取的数据信息。控制层采用面向对象的方法进行设计。类的关系如图 6 所示。

控制层接收应用层的控制信号、管理信号,通过网络层发送控制和管理数据给硬件层,控制和管理机器人。机器人控制包括对机器人的启动、伺服使能、任务更换、手动速度设置等。机器人管理包括对机器人的增加、删除、连接以及断开。机器人执行的操作和数据,在 Modbus 客户机打包成 Modbus 数据帧后,通过 QTcpSocket 将数据帧发送出去。

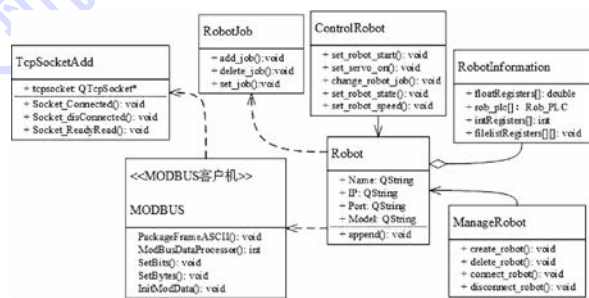


图 6 控制层类的关系

Fig. 6 Class diagram of control layer

2.4 应用层

2.4.1 仿真优化

仿真模块在 Qt4.8.6 和 Coin3D 3.1.3 开源库的基础上开发,类关系如图 7 所示。仿真模块包括 2 部分功能:对机器人进行离线仿真;实时显

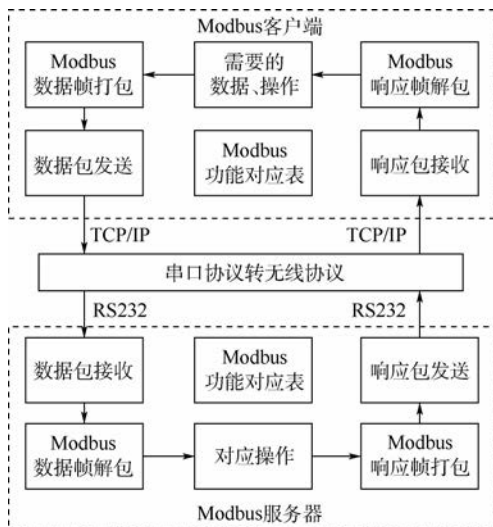


图 5 Modbus 通信流程

Fig. 5 Process of Modbus communication

示机器人当前的状态和信息。离线仿真对当前的任务进行仿真,验证路径的合理性,在此基础上对轨迹进行优化,生成机器人的运动指令。信息世界对机器人轨迹跟踪、分析,当物理机器人动

作时,信息世界的机器人也在实时同步动作。仿真模块的模型是物理世界机器人的三维模型、环境模型、工具模型。

2.4.2 交互界面

交互界面主要实现人与设备的交互,将信息通过更直接的方式表示。交互界面包括机器人管理、状态、位置信息、设置、IO 控制等。交互界面采用 Qt4.8.6 实现,类关系如图 8 所示。每个界面都由一个控制类和一个边界类组成。边界类是具体的 UI 界面,控制类继承边界类实现具体的操作响应、数据流。将所有的 UI 界面都作为主界面的包含项,需要查看某些界面就可以点开独立的界面,查看数据或者进行操作。

人是 CPS 中重要的一环,友好地交互体现人的意图。在操作时,系统通过传感器感知人类,通过交互界面与人实时交互。

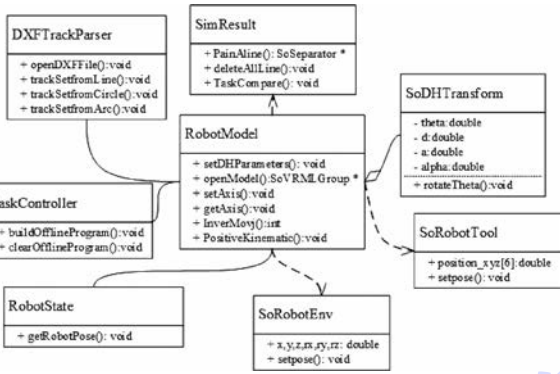


图 7 仿真环境类的关系

Fig.7 Class diagram of simulation environment

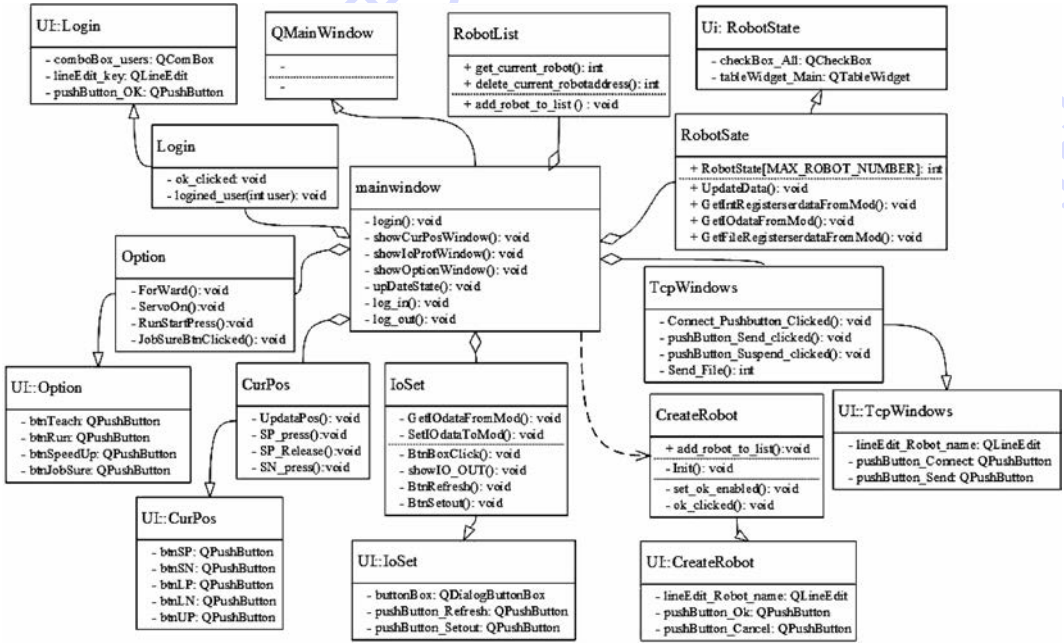


图 8 交互界面类的关系

Fig.8 Class diagram of interaction interface

3 实验

实验系统由埃夫特的 ER3A-C60 型机器人、自行研发的嵌入式控制器、MICRO-6013CM 型工业相机和控制 PC 组成,如图 9 所示。机器人、工业相机通过无线网络和系统通信。控制器和工业相机通过模块 HLK-RM04 将串口转化为无线通信和控制 PC 通信。通过相机采集、提取工件信息获得机器人的运动轨迹并仿真,得到机器人运动指令,控制机器人运动。机器人的运动信息实时上传给应用层,更新交互界面的显示,并在仿真

环境中实时更新机器人位姿,对机器人进行监控。

相机采集图像信息,提取轮廓,得到工件的加工信息,输出加工路径,将路径通过串口无线上传给应用层。应用层对其仿真优化,获得机器人最终的运动轨迹,如图 10 所示。实验模拟对工件去毛刺,刀具在工件正上方 z 坐标为 200 mm 的平面内运动,工件为长为 84.057 mm、宽为 74.316 mm 的矩形,矩形端点坐标为 $A(445.353, 32.085, 200)$ mm、 $B(529.41, 32.085, 200)$ mm、 $C(529.41, -42.231, 200)$ mm、 $D(445.353, -42.231, 200)$ mm。

实验操作流程如图 11 所示。创建无线环境,

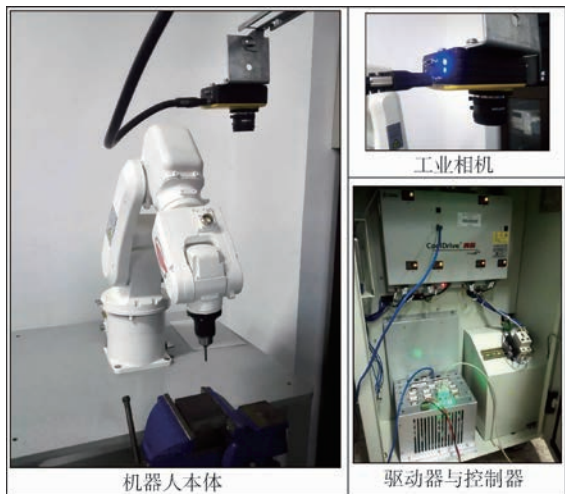
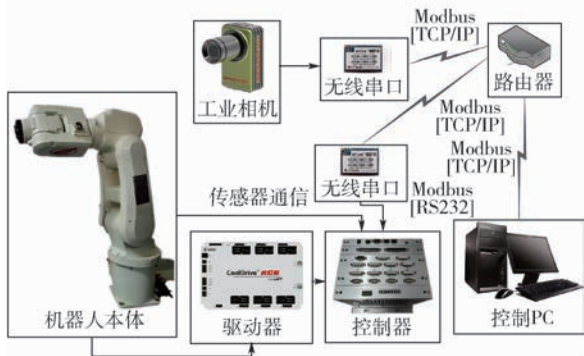


图9 实验系统

Fig. 9 Experimental system



启动机器人和工业相机,在应用中创建机器人,输入机器人的IP地址、DH参数、模型文件等。根据环境、工具模型,建立机器人的工作环境。连接之后,可以看到机器人的当前插补方式、操作坐标系、速度等级、循环模式、任务状态等信息。根据工业相机采集的工件轮廓信息得到机器人运动轨迹,仿真得到机器人的运动指令。运动指令下载到机器人,控制机器人运动。机器人的信息通过网络层传输给控制层,控制层将其提供给应用层仿真,并在交互界面显示。交互界面显示机器人的运行状态、IO状态等,仿真实时同步机器人的位姿。

实验中,实际机器人与仿真机器人对比如图12所示,仿真机器人与实际机器人运行轨迹一致。

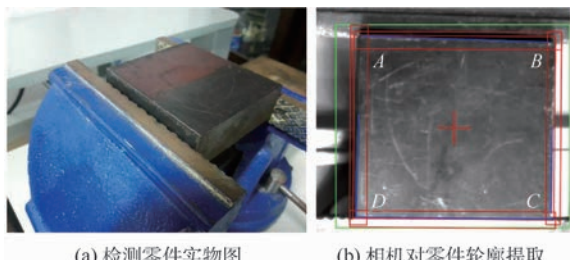


图10 任务轨迹相对位置

Fig. 10 Relative position of task trajectory

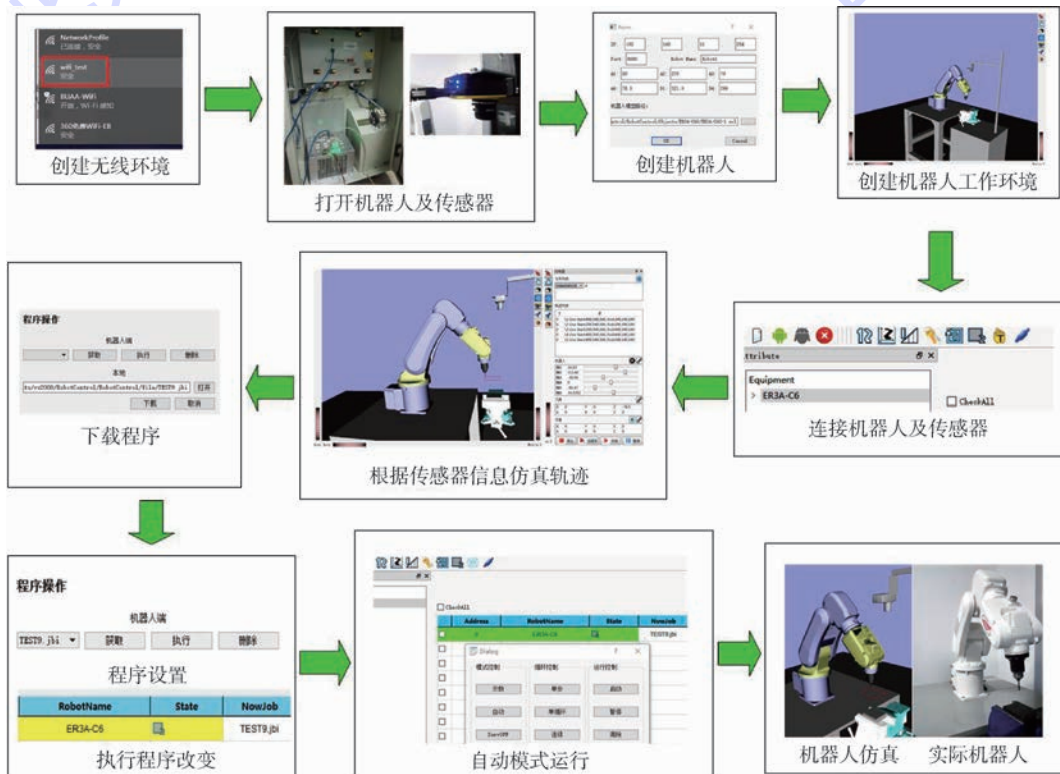


图11 操作流程

Fig. 11 Operation process

加工一个零件的周期为 5.5 s,实际轨迹与工件轮廓的标准方差为

$$S = \sqrt{\frac{\sum_{i=1}^n (s_i - \bar{s})^2}{n}} = 0.6071 \quad (1)$$

式中: s_i 为实际记录轨迹点坐标; \bar{s} 为理论轨迹点坐标。

图 13(a)为机器人轨迹对比,图 13(b)为机器人各轴角度的变化曲线。图 13(a)中红色线为机器人实际运行的轨迹,蓝色线为工件的实际轮廓线,两线基本重合。

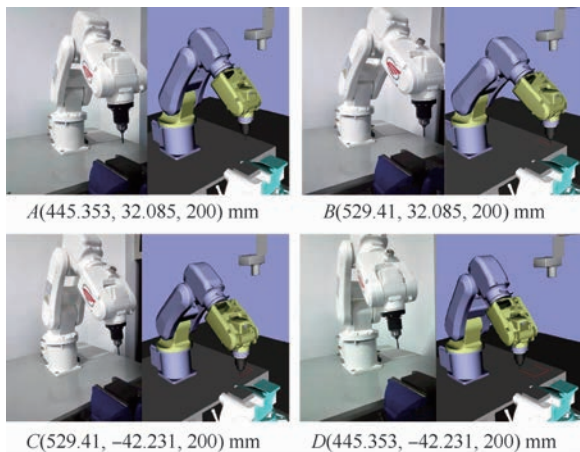


图 12 运行中的机器人与仿真模型

Fig. 12 Running robot and simulation model

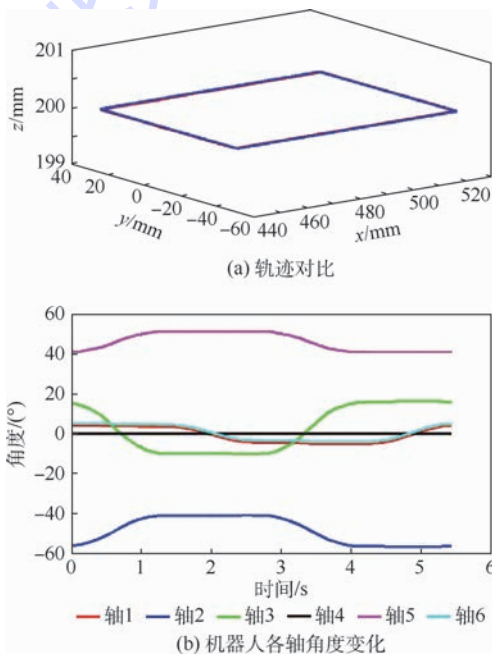


图 13 任务轨迹分析

Fig. 13 Analysis of task trajectory

机器人系统,该系统分为物理层、网络层、控制层和应用层。

2) 实现了基于 CPS 方法的工业机器人系统,该系统能够实现设备数据采集、通信和控制,机器人任务提取、仿真、优化和监控。

3) 通过实验验证了该系统的可行性。

参考文献 (References)

[1] LEITÃO P, COLOMBO A W, KARNOUSKOS S. Industrial automation based on cyber-physical systems technologies: Prototype implementations and challenges [J]. Computers in Industry, 2016, 81 (s1) :11-25.

[2] 何积丰. Cyber-physical systems [J]. 中国计算机学会通讯, 2010, 6(1) :25-30.

HE J F. Cyber-physical systems [J]. Communications of the China Computer Federation, 2010, 6(1) :25-30 (in Chinese).

[3] WANG L, TÖRNGREN M, ONORI M. Current status and advancement of cyber-physical systems in manufacturing [J]. Journal of Manufacturing Systems, 2015, 37(2) :517-527.

[4] LEE E A. Cyber physical systems: Design challenges [C] // 11th IEEE Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing. Piscataway, NJ: IEEE Press, 2008:363-369.

[5] U. S. Government, National Institute of Standards and Technology. Foundations for innovation: Strategic R&D opportunities for 21st century cyber-physical systems-connecting computer and information systems with the physical world, robots, autonomous vehicles [R]. Washington, D. C. :NIST, 2015.

[6] MICHNIEWICZ J, REINHART G. Cyber-physical-robotics—Modelling of modular robot cells for automated planning and execution of assembly tasks [J]. Mechatronics, 2016, 34 (s1) : 170-180.

[7] CHEN J, YANG J, ZHOU H, et al. CPS modeling of CNC machine tool work processes using an instruction-domain based approach [J]. Engineering, 2015, 1(2) :247-260.

[8] CHEN W L, CAO R M, HAO L N, et al. Researches on robot system architecture in CPS [C] // 2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems. Piscataway, NJ: IEEE Press, 2015:603-607.

[9] LUO R C, KUO C W. Intelligent seven-DoF robot with dynamic obstacle avoidance and 3-D object recognition for industrial cyber-physical systems in manufacturing automation [J]. Proceedings of the IEEE, 2016, 104(5) :1102-1113.

[10] COLOMBO A W, KARNOSKOS S. Towards the factory of the future: A service-oriented cross-layer infrastructure [J]. ICT Shaping the World: A Scientific View, 2009, 2009 :65-81.

[11] 温景容, 武穆清, 宿景芳. 信息物理融合系统 [J]. 自动化学报, 2012, 38(4) :507-517.

WEN J R, WU M Q, SU J F. Cyber-physical system [J]. Acta Automatica Sinica, 2012, 38(4) :507-517 (in Chinese).

[12] TAN Y, GODDARD S, REZ L C. A prototype architecture for cyber-physical systems [J]. ACM Sigbed Review, 2008, 5(1) : 1-2.

4 结 论

1) 本文提出了一种基于 CPS 方法的工业机

- [13] 王小乐,陈丽娜,黄宏斌,等.一种面向服务的CPS体系框架[J].计算机研究与发展,2010,47(s2):299-303.
WANG X Y, CHEN L N, HUANG H B, et al. A service-oriented architecture framework of cyber-physical systems[J]. Journal of Computer Research and Development, 2010, 47(s2): 299-303 (in Chinese).
- [14] 王田苗,陶永.我国工业机器人技术现状与产业化发展战略[J].机械工程学报,2014,50(9):1-13.
WANG T M, TAO Y. The present situation of industrial robot technology in China and the strategy of industrialization development[J]. Chinese Journal of Mechanical Engineering, 2014, 50(9): 1-13 (in Chinese).
- [15] Modbus IDA. Modbus application protocol specification v1. 1a [EB/OL]. (2004-06-04) [2017-07-01]. <http://www.Modbus-IDA.org>.

作者简介:

陈友东 男,博士,副教授。主要研究方向:工业机器人、人机协作、运动控制、数控系统。

常石磊 男,硕士研究生。主要研究方向:工业机器人。

冯强国 男,硕士研究生。主要研究方向:工业机器人。

Industrial robot system based on CPS approach

CHEN Youdong*, CHANG Shilei, FENG Qiangguo

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

Abstract: Industrial robot system based on cyber-physical system (CPS) is a key and enabling technology for intelligent manufacturing. This paper presents an industrial robot system based on CPS approach, which is the fusion of physical world and information world. The system is divided into physical layer, network layer, control layer and application layer. Data from the physical layer are sent to the control layer via the communication layer to update the data in the information world. The application layer analyzes and optimizes these data to make decisions. In the control layer, the decisions are transformed into the device control data. The control data are sent to the physical layer to control the devices through the communication layer. A machining experiment is conducted to validate the feasibility of the method by using the ER3A-C60 industrial robot, MI-CRO-6013CM smart camera and in-house industrial control system.

Keywords: cyber-physical system (CPS); industrial robot; control system; Modbus; simulation

Received: 2017-06-05; **Accepted:** 2017-07-07; **Published online:** 2017-10-19 11:12

URL: kns.cnki.net/kcms/detail/11.2625.V.20171019.1112.002.html

Foundation items: National Key Technology Research and Development Program of the Ministry of Science and Technology of China (2015BAF01B04); Beijing Municipal Science and Technology Project (D161100003116002)

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

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

DOI: 10.13700/j.bh.1001-5965.2017.0309

航空燃油柱塞泵滑靴副混合润滑特性数值仿真



符江锋*, 李华聪, 曾夏青, 刘显为

(西北工业大学 动力与能源学院, 西安 710072)

摘 要: 针对航空燃油柱塞泵滑靴副的动静压混合支承下的润滑问题,在滑靴副的运动学和动力学模型基础上,考虑静压支承与滑靴非规则的空间曲线运动特征所产生的动压效应,建立了动静压效应下的滑靴副混合润滑数学模型。基于有限体积法进行了滑靴副的润滑特性仿真计算研究,分别对混合润滑机理下的油膜厚度变化规律、油膜压力分布影响因素和滑靴抗倾覆能力进行了仿真分析研究。仿真结果表明:动静压混合支承所得到的油膜厚度变化趋势更符合滑靴副实际的润滑状态;中心油膜厚度、滑靴最大倾斜角和转子转速主要对动压效应产生影响,而滑靴副进口压力即柱塞泵供油压力主要影响油膜的静压作用;提高滑靴的抗倾覆能力可通过增大滑靴底面工作半径或者减小滑靴中心油池半径来增强油膜的动压效应,抵消滑靴受到的倾覆力矩。

关键词: 燃油柱塞泵; 滑靴副; 动压效应; 润滑特性; 静压支承

中图分类号: V228.1

文献标识码: A **文章编号:** 1001-5965(2018)05-0939-12

航空燃油柱塞泵作为发动机喷口油源泵,其工作在发动机机匣外的高温复杂环境中,先进发动机的发展需求使得燃油柱塞泵进一步向高转速、大增压比和强功率负荷的方向发展。滑靴作为柱塞泵的关键摩擦副,长期工作在高速、重载的状态下,其油膜润滑性能是否良好直接制约燃油柱塞泵高压化、高速化并影响柱塞泵性能、寿命及可靠性^[1]。

目前,提高摩擦副性能及寿命研究除了发展抗磨材料外,保证其最佳的润滑状态是降低摩擦和减少磨损最普遍、最有效的办法。为此,国内外相关学者针对柱塞泵滑靴副的润滑特性开展了长期而广泛的研究,并取得了显著的和试验成果。早期,国外对于轴向柱塞泵滑靴副的相关研究主要集中在滑靴受力、姿态和静压支承润滑特性、滑靴的动力学建模等理论及试验研究方面。

近年来,随着数值分析方法和计算机仿真技术的兴起和发展,国外学者越来越多地采取数值计算和软件仿真相结合的方法对柱塞泵关键摩擦副进行研究,典型代表有美国的 Purdue 大学和德国的亚琛工业大学。Purdue 大学的 Ivantysynova 和 Wieczorek 等^[2-4]基于所设计的 CASPAR 软件对轴向柱塞泵的配流副、柱塞副和滑靴副 3 处油膜进行了全面的润滑特性仿真研究;随后又通过 CFD 软件建立了滑靴副的热弹流动态仿真模型,对高压下的滑靴变形、油膜的温度场进行了数学建模和仿真研究。Ivantysynova 和 Schenk^[5]于 2015 年在之前研究工作的基础上,建立了滑靴副瞬态热弹性动压润滑模型,该模型涵盖了滑靴的动力学特性、压力场模型、滑靴和斜盘的变形以及滑靴副的热传递模型,是迄今为止最贴近实际情况的一个润滑模型。德国亚琛工业大学 IFAS 开发了一

收稿日期: 2017-05-15; 录用日期: 2017-08-11; 网络出版时间: 2017-09-27 18:21

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20170927.1821.003.html

基金项目: 航空科学基金(20150653006); 中央高校基本科研业务费专项资金(G2017KY0003)

* 通信作者. E-mail: fjf@nwpu.edu.cn

引用格式: 符江锋, 李华聪, 曾夏青, 等. 航空燃油柱塞泵滑靴副混合润滑特性数值仿真[J]. 北京航空航天大学学报, 2018, 44(5): 939-950. FU J F, LI H C, ZENG X Q, et al. Numerical simulation of hybrid lubrication characteristics of slipper pair of aviation fuel piston pump[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 939-950 (in Chinese).

种可用于建立轴向柱塞泵完整仿真模型的软件 DSHplus, 该软件可以与多体运动学分析软件 ADMAS 进行联合仿真来实现滑靴副的动静压仿真。Deeken^[6]、Murrenhoff 和 Wohlers^[7] 通过 ADMAS 计算出滑靴副主要部件的动力学参数, 与油膜计算模型联合仿真分析滑靴副的瞬态特性。2014年, Deeken 等^[8] 在对滑靴副进行多体运动学仿真的基础上, 考虑温度场对滑靴润滑性能的影响并进行了仿真验证。国内对于柱塞泵滑靴副的研究起步较晚, 2010—2011年, 徐兵和李迎兵等^[9-10] 用滑靴3点处油膜厚度值对整个滑靴副的油膜厚度场进行了描述, 搭建了楔形油膜润滑模型。2011年, 刘洪等^[11] 对轴向柱塞泵滑靴副的油膜的动态规律进行了数学建模, 分析了滑靴润滑油膜与转子转速及柱塞腔压力之间的变化规律。2013年, 于思淼^[12] 重点研究了滑靴副油膜的动压效应和挤压效应, 应用 MATLAB 编制了滑靴副流体动压支承分析软件。2014年, 王亚军^[13] 对求解油膜压力场的有限体积法进行了优化, 基于 WENO 重构计算模型建立了滑靴副静压支承自适应分析模型。2013年, 魏超等^[14] 研究了表面微结构对滑靴副油膜承载力和抗倾覆力矩等润滑特性的影响。总体而言, 国内对滑靴副的研究还处于初级阶段, 在理论研究时对滑靴副做了很多的理想假设, 故不能完全真实反映滑靴副的润滑情况。

航空燃油柱塞泵与其他普通轴向柱塞泵相比, 其空间结构更加复杂, 耦合件精密度要求高, 各个组成元件之间的运动和受力关系也要复杂得多。此外, 燃油柱塞泵采用低介质黏度的航空3号煤油作为润滑介质, 并长期工作在高温、高压、高转速恶劣环境中, 因此对燃油柱塞泵的摩擦副润滑性能提出了更高的要求, 而国内针对航空燃油柱塞泵滑靴副仅开展了优化设计和静压支承下的油膜计算分析研究^[15-17], 尚未开展动静压效应下的滑靴副内部实际润滑分布规律研究。因航空燃油柱塞泵滑靴副多以静压支承原理进行设计, 滑靴在斜盘上的运动轨迹是非规则的空间曲线, 实际中滑靴将发生倾覆现象, 滑靴底面的楔形油膜因为相对运动将产生动压支承力, 使滑靴副的压力场分布变得更加复杂。为精确反映滑靴副的内部实际润滑情况, 本文在建立航空燃油柱塞泵运动学及动力学模型的基础上, 考虑油膜的静压支承及油膜的动压效应, 开展了航空燃油柱塞泵滑靴副的润滑特性仿真研究。

1 航空燃油柱塞泵滑靴副动力学与油膜动静压润滑模型

1.1 滑靴副运动学及动力学模型

航空燃油柱塞泵滑靴通常采用阻尼孔型静压支承结构(见图1), 即在柱塞腔和滑靴底部油池之间加设一个阻尼小孔(常称进口阻尼)来起调节作用, 使滑靴和斜盘在一定的油膜厚度下达到受力平衡。因斜盘工作面为一球面, 为保证良好贴合, 滑靴底面通常也为一近似球面。由于斜盘表面的曲率半径远大于滑靴副的相关尺寸, 为研究方便, 后续研究中假设滑靴底面和斜盘表面均为圆形平面。图1中: r_1 为滑靴底面油池的半径(直径为 d_1), 油池通过直径为 d_0 的阻尼小孔与柱塞腔相通, 滑靴工作面外径为 d_2 , h 为油膜厚度, F_A 为沿着轴向方向的正压紧力, p_0 为中心油池压力, p_r 为柱塞泵转子腔压力, d_z 为柱塞直径。

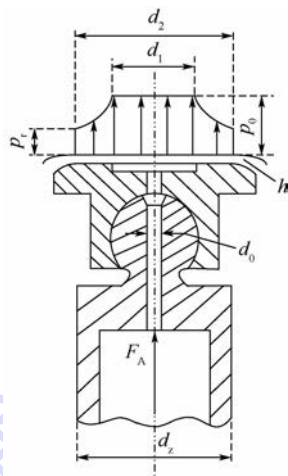


图1 航空燃油柱塞泵滑靴副结构

Fig. 1 Slipper pair structure of aviation fuel piston pump

航空燃油柱塞泵滑靴通过与柱塞头部形成的球铰副随着柱塞在斜盘表面进行周期运动, 其运行轨迹是一复杂的空间曲线。若将滑靴看成一个质点, 且同时忽略滑靴运动过程中自身的旋转, 则滑靴与柱塞球头中心 C 的运动情况完全相同。故可以通过分析柱塞的运动来体现滑靴的运动情况, 本文中只给出结果, 中间详细推导过程略过。滑靴的速度及加速度模型分别如下:

$$v = -L\omega \left[\lambda B \sin \alpha + \frac{\lambda^2 B (A + B \cos \alpha) \sin \alpha}{\sqrt{\lambda^2 (A + B \cos \alpha)^2 - \lambda^2 + 1}} \right] \quad (1)$$

$$a = L\omega^2 \left\{ \lambda B \cos \alpha + \frac{\lambda^2 B [(A + B \cos \alpha) \cos \alpha - B \sin^2 \alpha]}{\sqrt{\lambda^2 (A + B \cos \alpha)^2 - \lambda^2 + 1}} \right\} +$$

$$\left. \frac{\lambda^4 B^2 (A + B \cos \alpha)^2 \sin^2 \alpha}{[\lambda^2 (A + B \cos \alpha)^2 - \lambda^2 + 1]^{1.5}} \right\} \quad (2)$$

式中: $A = \cos \theta \sin \tau$; $B = -\cos \tau \sin \theta$; θ 为柱塞锥顶角; $\lambda = R_0/L$, R_0 为斜盘球面中心 O_1 至柱塞锥顶点 O 的距离; α 为转子转角; τ 为 R_0 与负 X 坐标轴的夹角; L 为斜盘球面中心 O_1 至柱塞球头中心 C 的距离; ω 为角速度。

柱塞泵高速运转过程中,滑靴受力可从柱塞对滑靴的合力、油膜承载力以及滑靴所受离心力 3 个方面进行分析,本文中只给出结果,中间详细推导过程略过。柱塞诸多力中沿柱塞轴线和沿滑靴轴线的力将柱塞压向滑靴,使滑靴受到一个沿轴线方向的正压紧力 F_A ,故有

$$F_A = (F_{L1} + F_{s1} + F_i + F_{cz}) \frac{1}{\cos \psi} + F_{L2} + F_{s2i} \quad (3)$$

式中: F_{L1} 为柱塞腔内的油压力; F_{s1} 为柱塞弹簧力; F_i 为相对运动惯性力; F_{cz} 为离心力沿柱塞轴线的分量; F_{L2} 为转子腔内的油压力; F_{s2i} 为保持架弹簧力; ψ 为柱塞轴线与滑靴轴线的夹角。

由式(3)可知,由于 F_A 表达式中的各个分量都是转子转角 α 的函数,故 F_A 也随 α 作周期变化。

滑靴随柱塞一起绕转子旋转,在斜盘表面的运动轨迹是一条复杂的空间曲线,和柱塞一样,滑靴也受到离心力的作用,若近似认为滑靴离心力作用在滑靴重心 C_s 上,滑靴在旋转过程中受到的离心力 F_{cs} 为

$$F_{cs} = m_s \omega^2 (\rho + l_c \cos \psi) \sin \theta \quad (4)$$

$$F_f = \mu_0 \frac{R_s}{h} \pi (r_2^2 - r_1^2) \quad (5)$$

式中: F_f 为滑靴副间油膜的黏性摩擦力,也是转子转角 α 的单变量函数, F_f 垂直于柱塞平面,与滑靴在斜盘上的运动方向相反; m_s 为滑靴质量; l_c 为滑靴重心到球窝球心 C (即柱塞球头中心) 的距离; R_s 为滑靴的旋转半径; ρ 和 ψ 为转子转角 α 的函数; μ_0 为油液动力黏度; r_1 为中心油池半径; r_2 为滑靴底面工作半径。

滑靴离心力 F_{cs} 随着角度 α 而变化,其作用在柱塞平面 W 上,其向垂直于转子轴线向外。

以典型 RZB-3 型航空燃油柱塞泵为对象,该泵的最大转速 $n = 4620 \text{ r/min}$,柱塞个数 $Z = 9$,柱塞直径 $d_z = 16 \text{ mm}$,柱塞锥顶角 $\theta = 14^\circ$,斜盘角度 $\varphi = 15.5^\circ$,其他结构参数略。通过上述方法对滑靴一个周期内 (α 从 0° 到 360°) 各个受力进行仿真分析,如图 2 所示。

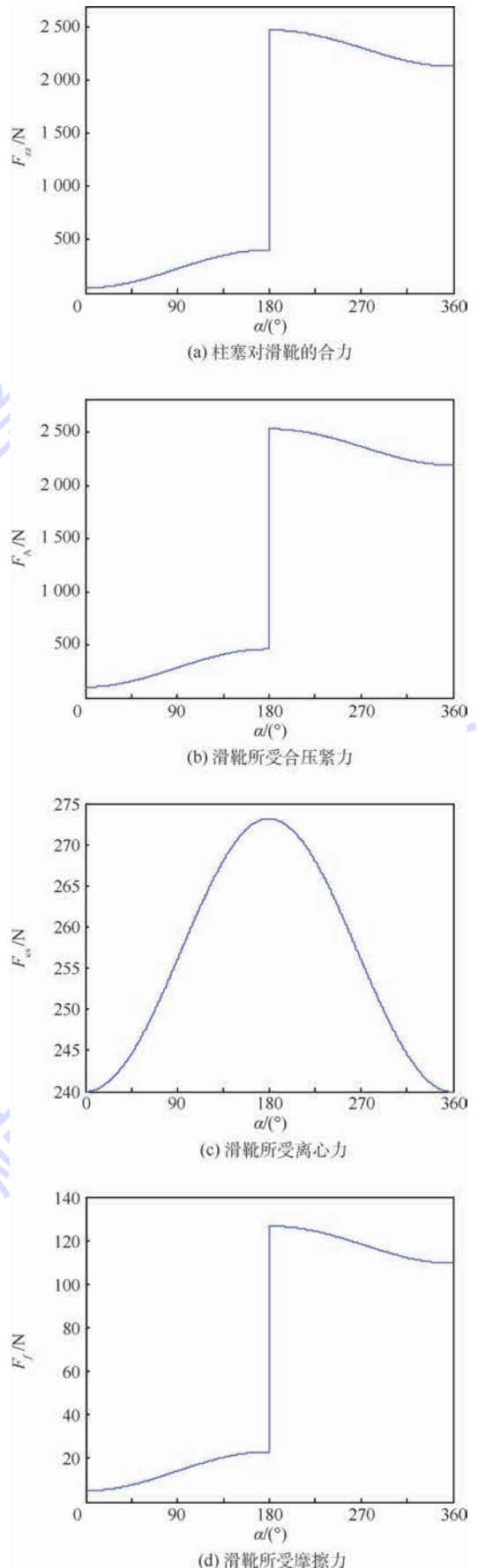


图 2 滑靴副受力仿真

Fig. 2 Force simulation of slipper pair

图2(a)、(b)分别为柱塞对滑靴的合力 F_z 以及滑靴所受合压紧力 F_A ,可知滑靴在吸油区受到的合压紧力较小,而在排油区受到的合压紧力可达2000 N以上。在柱塞泵吸排油切换点($\alpha = 180^\circ$),柱塞对滑靴的合力和滑靴所受的合压紧力将发生巨变,且在切换点滑靴所受合压紧力最大值为2532 N。

图2(c)、(d)分别为滑靴所受离心力 F_{cs} 和摩擦力 F_f 的仿真结果,可知滑靴受到的离心力和摩擦力相对于其他力而言较大,对滑靴的运动产生重要影响,在研究滑靴副的润滑特性时需要予以考虑。

滑靴底面的油膜支承力可以分为2个部分。第1部分为中心油池压力 p_0 ($0 \leq r \leq r_1$)产生的支承力,其值为

$$F_{hd1} = \pi r_1^2 p_0 \quad (6)$$

第2部分为环形密封带内压力 p_h ($r_1 \leq r \leq r_2$)产生的支承力,通过对密封带内的压力积分可得

$$F_{hd2} = \int_{r_1}^{r_2} p_h \cdot 2\pi r dr \quad (7)$$

整个滑靴底面的油膜支承力为

$$F_{hd} = F_{hd1} + F_{hd2} = \pi r_1^2 p_0 + \frac{\pi}{2} (p_0 - p_r) \frac{r_2^2 - r_1^2}{\ln(r_2/r_1)} \quad (8)$$

1.2 油膜动静压润滑模型

通过滑靴副的力平衡和流量守恒建立相应的平衡方程,分别计算滑靴副间油膜的压力分布、滑靴底面环形密封带的燃油泄漏量、滑靴底面油膜支承力和滑靴底面中心油池压力,得出静压支承下的油膜厚度计算模型如下:

$$h = \left(\frac{3}{2} d_0^2 \mu \mu_0 \frac{\ln(r_2/r_1)}{p_0 - p_r} \right)^{\frac{1}{3}} \left[\frac{2}{\rho_R} (p_h - p_0) \right]^{\frac{1}{6}} \quad (9)$$

式中: μ 为小孔流量系数; ρ_R 为燃油的密度。

由流体力学动压润滑理论^[18]可知,滑靴与斜盘之间形成的楔形油膜由于相对运动会产生附加压力场,这是因为流体通过收敛间隙速度增大,产生额外的动压力,从而改变滑靴副油膜静压支承下的压力场分布,并使整个滑靴副的润滑情况变得更为复杂。由于滑靴副油膜的厚度非常小,设定滑靴副油膜为层流,流体黏性和压力在厚度方向保持不变等假设是合理的。在层流状态下,选取柱坐标系为润滑油膜参考坐标系,由于滑靴倾斜方向与运动速度方向并不一致,所以在滑靴径向和周向都有动压产生。根据 Navier-Stokes 方程和连续性方程,可推导出柱坐标系下的适合油膜描述的控制方程为

$$\begin{aligned} & \frac{1}{12r} \cdot \frac{\partial}{\partial r} \left(\frac{rh^3}{\eta} \cdot \frac{\partial p}{\partial r} \right) + \frac{1}{12r} \cdot \frac{\partial}{\partial \theta^*} \left(\frac{h^3}{\eta} \cdot \frac{\partial p}{\partial \theta^*} \right) = \\ & \frac{1}{r} \cdot \frac{\partial}{\partial r} \left(\frac{rhv_r}{2} \right) + \frac{1}{r} \cdot \frac{\partial}{\partial \theta^*} \left(\frac{h v_{\theta^*}}{2} \right) - v_r \frac{\partial h_i}{\partial r} - \\ & \frac{v_{\theta^*}}{r} \cdot \frac{\partial h_i}{\partial \theta^*} + \frac{\partial h_i}{\partial t} - \frac{\partial h_b}{\partial t} \end{aligned} \quad (10)$$

式中: r 为滑靴半径; p 为油膜压力; θ^* 为滑靴坐标系中的角度; η 为燃油黏度; v_r 和 v_{θ^*} 分别为滑靴与斜盘相对滑动速度的径向和周向分量; h_i 和 h_b 分别为滑靴顶部和斜盘底部壁面的厚度。

式(10)即为滑靴副油膜等温不可压黏性定常 Reynolds 方程。式(10)右边前2项表示油膜在滑靴与斜盘楔形间隙下运动产生的动压效应;中间2项表示由滑靴表面的边界速度梯度引起的变化;后2项表示由滑靴和斜盘的弹性变形引起的油液挤压效应。

2 滑靴副动静压混合润滑特性计算

2.1 基本假设

动压效应的存在使得滑靴副油膜的润滑状况变得错综复杂,其压力场和厚度场的求解也变得更加困难。由于滑靴副间的流动是错综复杂的空间流动,基于滑靴副基本润滑性质分析,研究中对润滑油膜作如下假设:

- 1) 与黏性力相比,忽略油膜受到的体积力和惯性力,如重力和离心力。
- 2) 燃油为 Newton 流体,流动为层流,不考虑油膜中的湍流和涡流。
- 3) 油液在滑靴副壁面无滑移,即油膜在滑靴底面和斜盘表面的油层速度与壁面速度相同。
- 4) 在润滑油膜厚度方向上,流体黏性和压力保持不变。
- 5) 与油膜厚度相比,固体表面的曲率半径很大,因而忽略油膜曲率引起的速度方向的变化。
- 6) 滑靴副属于窄面密封,假设油膜温度场均匀分布。

2.2 边界条件

边界条件是进行数值求解的重要前提,为求解滑靴副油膜压力分布情况,首先需明确油膜的压力边界条件和速度边界条件。

滑靴副工作时,燃油从中心油池流到柱塞泵转子腔在密封带内形成润滑油膜,故滑靴油膜的内压力边界 p_1 为中心油池压力,外压力边界 p_2 为转子腔压力,即

$$\begin{cases} r = r_1, p_1 = p_0 \\ r = r_2, p_2 = p_r \end{cases} \quad (11)$$

又由于油液在滑靴副壁面无滑移,故可知油膜在厚度方向的速度边界条件为

$$\begin{cases} z = 0, u = 0, v = 0 \\ z = h, u = v_r, v = v_\theta \end{cases} \quad (12)$$

式中: u 为油膜速度的径向分量; v 为油膜速度的周向分量。

滑靴副油膜的控制方程为椭圆型偏微分方程,对于这一类方程的求解在现今阶段主要采用有限差分法和有限体积法进行离散,这 2 种方法在本质上都是基于流体的流量守恒。由相关研究可知,后者在收敛性和求解精度上优于前者,故选取有限体积法求解油膜压力场。有限体积法求解滑靴副油膜压力场的基本过程是由简化的 Navier-Stokes 方程积分得到油膜的流速方程,沿油膜厚度方向对流速方程进行积分得到单位宽度的截面流量,乘以控制体积的宽度得到整个体积流量,代入流量守恒方程进行求解。具体过程如下:

$$\begin{cases} \frac{dV_x}{dt} = f_x - \frac{1}{\rho_R} \cdot \frac{\partial p}{\partial x} + \eta \left(\frac{\partial^2 V_x}{\partial x^2} + \frac{\partial^2 V_x}{\partial y^2} + \frac{\partial^2 V_x}{\partial z^2} \right) \\ \frac{dV_y}{dt} = f_y - \frac{1}{\rho_R} \cdot \frac{\partial p}{\partial y} + \eta \left(\frac{\partial^2 V_y}{\partial x^2} + \frac{\partial^2 V_y}{\partial y^2} + \frac{\partial^2 V_y}{\partial z^2} \right) \\ \frac{dV_z}{dt} = f_z - \frac{1}{\rho_R} \cdot \frac{\partial p}{\partial z} + \eta \left(\frac{\partial^2 V_z}{\partial x^2} + \frac{\partial^2 V_z}{\partial y^2} + \frac{\partial^2 V_z}{\partial z^2} \right) \end{cases} \quad (13)$$

式(13)即为不可压缩流体的简化 Navier-Stokes 方程,将其转化成圆柱坐标形式并在平面上积分可得油膜的流速为

$$\begin{cases} u = \frac{1}{2\eta} \cdot \frac{\partial p}{\partial r} [z(z-h)] + \frac{z}{h} v_r \\ v = \frac{1}{2\eta} \cdot \frac{\partial p}{r \partial \theta} [z(z-h)] + \frac{z}{h} v_\theta \end{cases} \quad (14)$$

将式(14)沿油膜厚度方向积分则可以得到单位宽度截面上通过的体积流量为

$$\begin{cases} q_r = \int_0^h u dz = -\frac{1}{12\eta} \cdot \frac{\partial p}{\partial r} h^3 + \frac{h}{2} v_r \\ q_\theta = \int_0^h v dz = -\frac{1}{12\eta} \cdot \frac{\partial p}{r \partial \theta} h^3 + \frac{h}{2} v_\theta \end{cases} \quad (15)$$

式中: q_r 为体积分量的径向分量; q_θ 为体积分量的周向分量。

通过对滑靴副求解域进行径向和周向网格划分,即将滑靴副润滑区域划分成有限个控制体积,并使每一个体积单元充分得微小以保证后续求解过程中的迭代精度。以每一个网格节点为中心,选取半网格范围空间区域作为离散控制体积。根据流量守恒原理,从各边界流入和流出的流量相等,整个体积内的流量增量为零,即

$$Q_{i-\frac{1}{2},j} + Q_{i,j-\frac{1}{2}} - Q_{i+\frac{1}{2},j} - Q_{i,j+\frac{1}{2}} = 0 \quad (16)$$

利用式(15)即可求出式(16)中的各项流量,如下:

$$Q_{i-\frac{1}{2},j} = -\frac{r_{i-\frac{1}{2},j} \Delta \theta}{12\eta} \cdot \frac{\partial p}{\partial r} h^3 + \frac{r_{i-\frac{1}{2},j} \Delta \theta h}{2} v_r \quad (17)$$

$$Q_{i,j-\frac{1}{2}} = -\frac{\Delta r}{12\eta r_{i,j-\frac{1}{2}}} \cdot \frac{\partial p}{\partial \theta} h^3 + \frac{\Delta r h}{2} v_\theta \quad (18)$$

$$Q_{i+\frac{1}{2},j} = -\frac{r_{i+\frac{1}{2},j} \Delta \theta}{12\eta} \cdot \frac{\partial p}{\partial r} h^3 + \frac{r_{i+\frac{1}{2},j} \Delta \theta h}{2} v_r \quad (19)$$

$$Q_{i,j+\frac{1}{2}} = -\frac{\Delta r}{12\eta r_{i,j+\frac{1}{2}}} \cdot \frac{\partial p}{\partial \theta} h^3 + \frac{\Delta r h}{2} v_\theta \quad (20)$$

将式(17)~式(20)代入式(16),并对方程中压力 p 的偏微分项进行向后差分处理,由此可得到各个节点 $p_{(i,j)}$ 的数值计算表达式。各网格节点处的离散差分表达式构成了一个非线性方程组,而方程组中的系数都是各个节点待求油膜厚度或油膜压力的函数。本文中采取数值迭代的方法求解上述非线性方程组,从而得到整个油膜的的压力分布。

实际求解过程中,先给定初始的中心油膜厚度 h_c 和滑靴最大倾斜角 β_{\max} ,采用 Gauss-Seidel 超松弛迭代算法进行求解,并设定数值收敛条件为

$$\varepsilon = \sum_{i=2}^m \sum_{j=2}^n \left| \frac{p_{i,j}^{k+1} - p_{i,j}^k}{p_{i,j}^k} \right| \leq [\varepsilon] \quad (21)$$

式中: m 和 n 分别为径向和周向网格节点数; ε 为从 k 次迭代到 $k+1$ 次的相对误差,取 $[\varepsilon] = 1 \times 10^{-6}$ 。若迭代误差满足要求,油膜压力场求解完毕,通过积分可求得滑靴副油膜支承载力、动压力矩和泄漏量等润滑特性参数。

3 数值仿真分析

3.1 中心油膜厚度和最大倾斜角

基于动静压混合润滑特性求解方法,通过 MATLAB 编程对滑靴副的非线性动力学模型进行数值求解,可以得到滑靴副一个周期内的中心油膜厚度和滑靴最大倾斜角,其中最大倾斜角取的是绝对值,没有考虑其与坐标系位置的关系,结果如图 3 和图 4 所示。

图 3 中,在考虑油膜动压效应的情况下,滑靴副油膜厚度的变化趋势与静压润滑下的变化趋势大致吻合,但油膜厚度整体上是减小的,且吸排油区切换时油膜厚度不再发生跳变,这正是动压效应的作用,同时也说明只考虑静压支承载力求得的油膜厚度是不符合实际的。

由图 4 可知,滑靴将在吸排油区切换点($\alpha = 180^\circ$)发生最大倾斜,由动压润滑原理可推出,滑

靴在该点将产生最大动压支承力。

若已知滑靴运动周期中某一点的中心油膜厚度和最大倾斜角,在 $\alpha = 180^\circ$ 条件下得到该状况下整个滑靴副的油膜厚度分布,如图5所示。

由图5可知, $\alpha = 180^\circ$ 时油膜厚度最大值为 $16.9 \mu\text{m}$,最小值为 $3.1 \mu\text{m}$ 。通过这种求解方式,可得到滑靴旋转周期内任意位置的油膜厚度分布,以及这一位置滑靴底面任意点上的油膜厚度值。

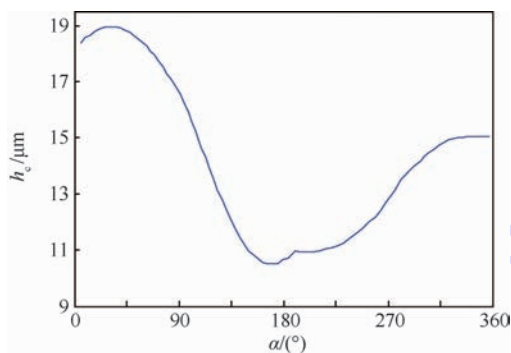


图3 滑靴副中心油膜厚度

Fig. 3 Central oil film thickness of slipper pair

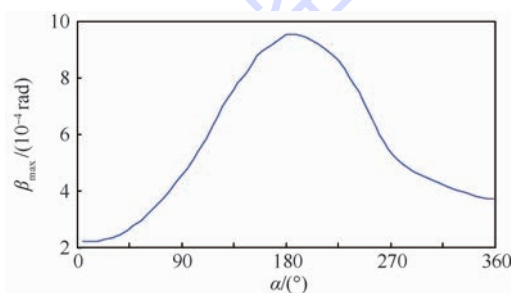


图4 滑靴最大倾斜角

Fig. 4 Maximum inclination angle of slipper

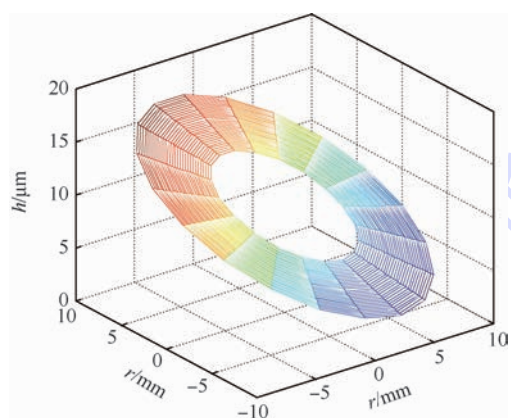


图5 $\alpha = 180^\circ$ 时的油膜厚度分布

Fig. 5 Distribution of oil film thickness with $\alpha = 180^\circ$

3.2 影响油膜压力分布的因素

1) 动压效应的仿真验证

取 $\alpha = 180^\circ$ 即吸排油区切换点进行数值模

拟,由图3和图4可知此时中心油膜厚度为 $10.6 \mu\text{m}$,最大倾斜角为 $9.5 \times 10^{-4} \text{ rad}$ 。按此条件数值模拟得到的结果如图6所示。

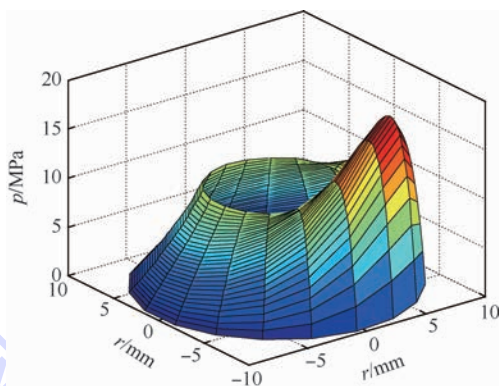


图6 实际情况下的油膜压力分布

Fig. 6 Distribution of oil film pressure in actual situation

为分析该条件下油膜动静压混合支承力中油膜静压支承力和动压支承力的大小,对该情况下的油膜动静压混合支承力进行积分求解,混合支承力为

$$F_{\text{hd}} = \pi r_1^2 p_0 + \int_{r_1}^{r_2} \int_0^{2\pi} p r dr d\theta \quad (22)$$

其中,油膜的静压支承力为

$$F_j = \pi r_2^2 p_r + \frac{\pi}{2} (p_0 - p_r) \frac{r_2^2 - r_1^2}{\ln(r_2/r_1)} \quad (23)$$

油膜的动压支承力为

$$F_d = F_{\text{hd}} - F_j \quad (24)$$

通过式(22)~式(24),得到该时刻的油膜动静压混合支承力即全部承载力为 2509.5 N ,其中静压支承力为 1790 N ,动压支承力为 719.5 N 。

为验证动静压混合支承数值模拟计算结果的准确性,将数值计算结果与在滑靴受力计算结果(见图2(b))进行对比。可知在同等条件下,滑靴所受合压紧力为 2532.5 N ,动静压混合支承力为 2509.5 N ,动静压混合支承仿真计算的支承力较滑靴压紧力相比,相对误差仅为 0.9% ,表明燃油柱塞泵滑靴副动静压混合润滑分析方法是正确的,且所建立的动静压混合数学模型具有高的计算精度。此外还可以发现,燃油柱塞泵滑靴副动静压混合润滑机理中,油膜的承载力主要由静压支承力提供,但动压支承力同样不可忽略。

2) 中心油膜厚度

滑靴副油膜形状主要由中心油膜厚度和滑靴最大倾斜角确定。在滑靴最大倾斜角下对不同中心油膜厚度的滑靴副油膜压力分布进行了数值模拟。取 $\beta_{\text{max}} = 7.5 \times 10^{-4} \text{ rad}$, $h_c = 10, 11, 12 \mu\text{m}$,其他仿真参数保持不变,仿真结果如图7所示。

的中心油膜厚度变化可能造成数十兆帕的动压力差值。

3) 滑靴最大倾斜角

取滑靴中心油膜厚度 h_c 为固定值 $11 \mu\text{m}$, 滑靴最大倾斜角 β_{\max} 分别取为 6.0×10^{-4} 、 7.5×10^{-4} rad, 其他工作参数保持不变, 仿真结果如图 8 所示。

由图 8 可知, 滑靴中心油膜厚度固定时, 随着滑靴最大倾斜角的增加, 楔形油膜的收敛性增强, 从而促进动压力的产生, 使收敛间隙的动压效应更为显著。通过计算动压支承力可知, 相比于中心油膜厚度, 动压效应对于最大倾斜角的变化更为敏感。同时可以发现, 在这一过程中, 油膜的静压作用区域压力分布几乎没有发生变化。

对图 7 和图 8 分析可知, 滑靴副油膜的动压力和静压力是相对独立的, 它们之间相互影响较小, 且静压作用在滑靴工作过程中更为稳定, 在油膜承载力中占主导地位; 动压作用虽然贡献更小, 但对于滑靴抗倾覆的作用更大。

4) 转子转速

柱塞泵工作时, 滑靴随着转子在斜盘表面进行旋转运动, 现改变转子转速进行数值模拟。取 $h_c = 11 \mu\text{m}$, $\beta_{\max} = 7.5 \times 10^{-4}$ rad, 转子转速分别为 3000、4000、5000 r/min, 其他仿真参数保持不变, 计算结果如图 9 所示。

由图 9 的仿真结果可知, 在滑靴相关参数保持不变的情况下, 随着柱塞泵转子转速的增加, 滑靴在斜盘表面的相对运动速度相应增大, 根据动压润滑原理, 在相同的收敛间隙下, 流体相对于收敛间隙的流速越大, 动压效应越明显, 产生的动压力也越大。观察各转子转速下油膜的静压作用区域可知, 油膜的静压力几乎维持不变。由图 9(d) 可知, 燃油柱塞泵滑靴副在高转子转速下的油膜总承载力大于低转子转速下总承载力, 具有更好的润滑效果。

5) 供油压力

滑靴副流体计算域的出口压力即为柱塞泵的供油压力, 其在柱塞泵工作时并不是保持恒定的, 故有必要对不同供油压力下的滑靴油膜压力分布进行分析。仿真参数为 $h_c = 11 \mu\text{m}$, $\beta_{\max} = 7.5 \times 10^{-4}$ rad, 供油压力分别取 10、20、30 MPa, 仿真结果如图 10 所示。

由图 10(a)、(b)、(c) 可以直观发现, 供油压力的增加使滑靴副油膜压力得到了整体提升, 无论是体现静压作用的中心压力还是体现动压效应的峰值压力都是随之增大的, 但静压支承力随供

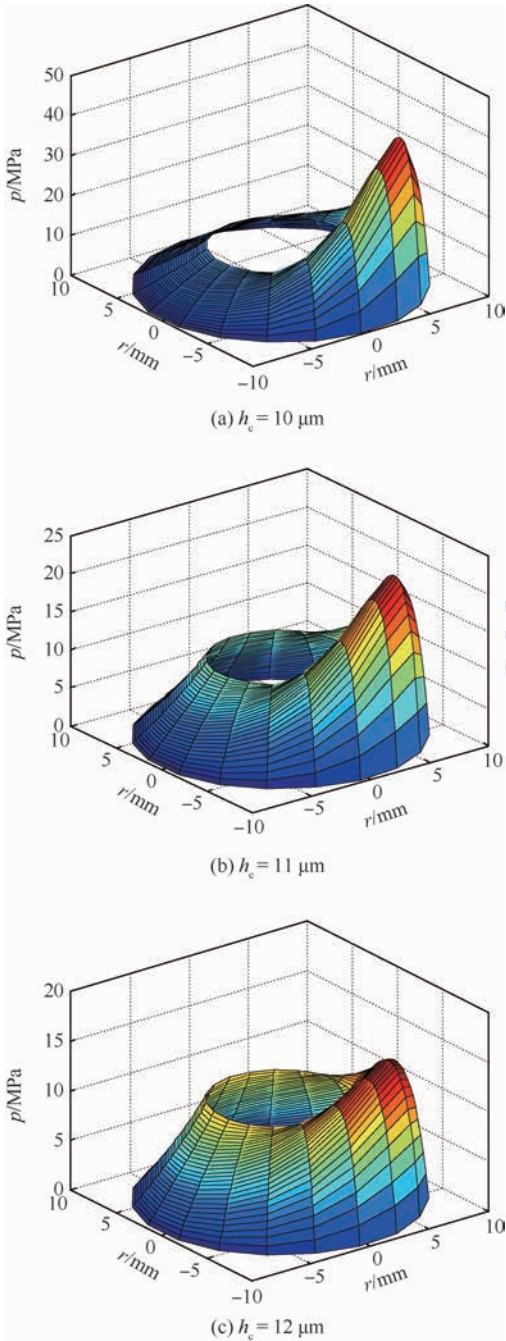
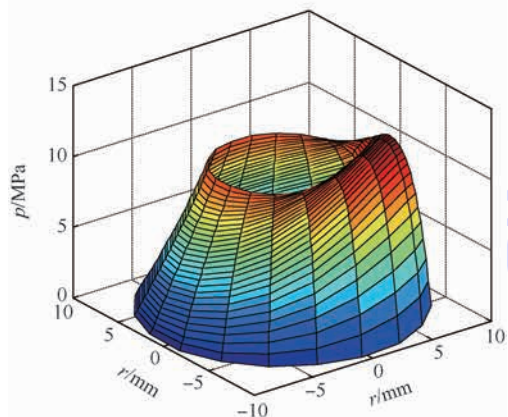
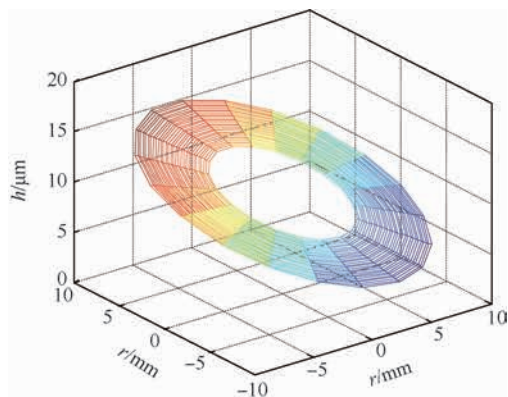
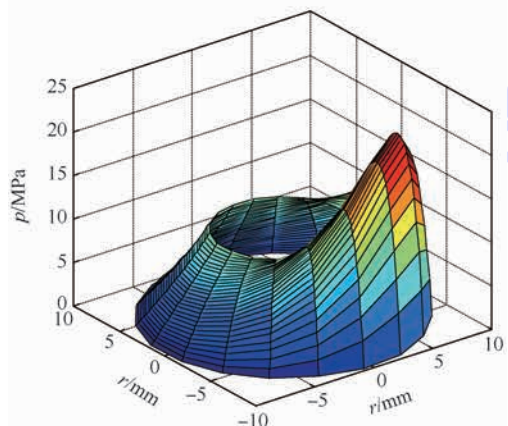
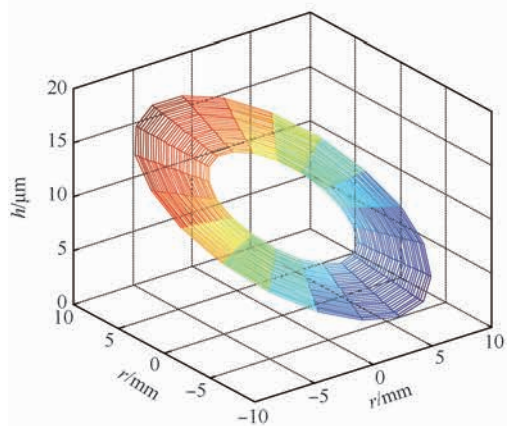
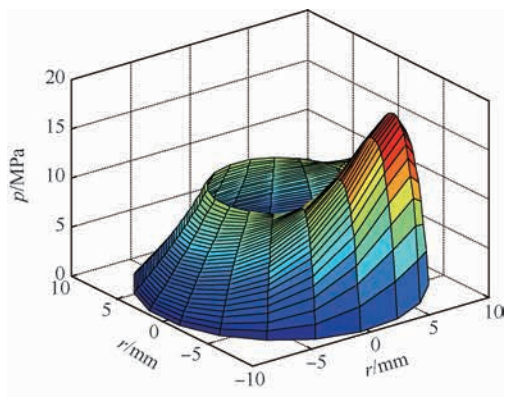
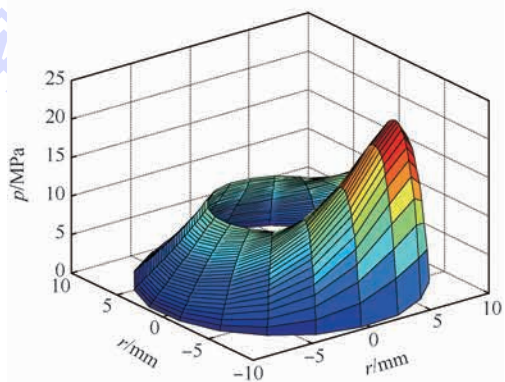


图 7 不同中心油膜厚度下的油膜压力分布
Fig.7 Oil film pressure distribution under different central oil film thickness

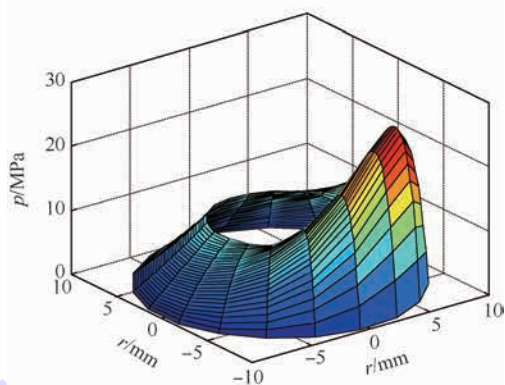
由图 7(a)、(b)、(c) 对比可知, 当滑靴最大倾斜角保持不变时, 不同的中心油膜厚度下, 滑靴油膜的静压作用区域压力分布几乎不发生改变, 动压力峰值随着中心油膜厚度的增加而减小。通过计算可知, 油膜整体的静压支承力保持不变, 动压支承力则是减小的, 但静压力分布随着中心油膜厚度的增加其压力梯度向边缘发散。同时可以直观地发现, 油膜的动压效应对于中心油膜厚度的变化十分敏感, 在最大倾斜半径上, $10 \mu\text{m}$ 中心油膜厚度下的最大压力可达 40 MPa 以上, $1 \mu\text{m}$

(a) $\beta_{\max} = 6.0 \times 10^{-4}$ rad(b) $\beta_{\max} = 7.5 \times 10^{-4}$ rad

(a) 转子转速为3000 r/min



(b) 转子转速为4000 r/min



(c) 转子转速为5000 r/min

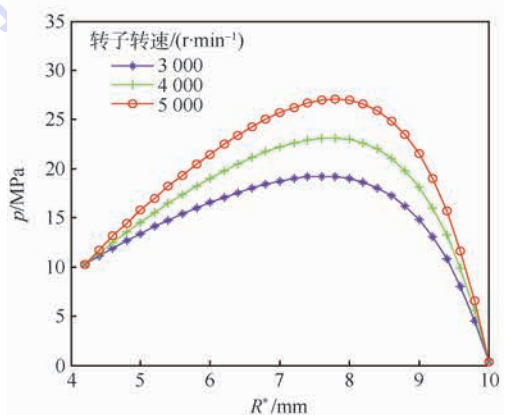
(d) 最大倾斜半径 R' 下不同转子转速的油膜压力变化

图8 不同最大倾斜角下的油膜厚度及压力分布

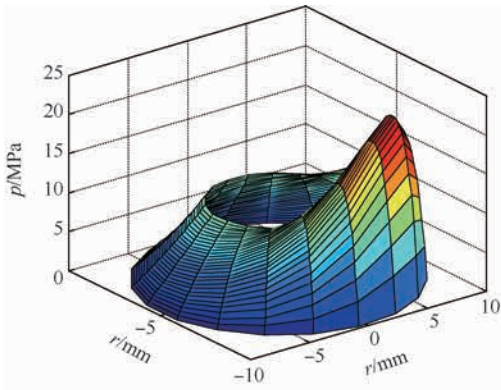
Fig. 8 Oil film thickness and pressure distribution under different maximum inclination angles

图9 不同转子转速下的油膜压力分布

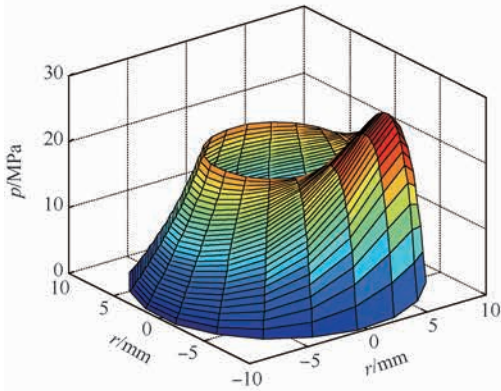
Fig. 9 Distribution of oil film pressure under different rotor speeds

油压力的增加效果更为显著,油膜的整体承载力随供油压力的增大而提高。由图 10(d)中最大倾斜半径为 4 处即滑靴油池中心的压力值可知,随着供油压力的增加,中心油池压力增加十分明显,且几乎与供油压力大小一致。对上述 3 种供油压力下的油膜动静压支承力进行计算,结果如图 11 所示。

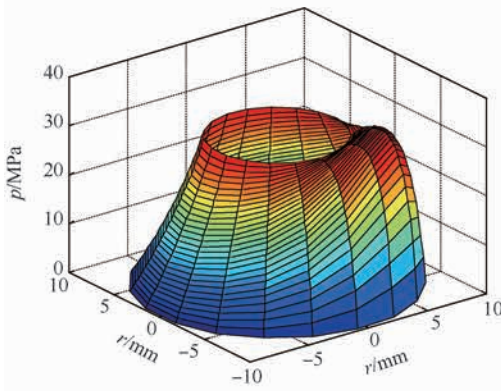
图 11 中,供油压力的变化对滑靴副油膜的总承载力影响显著,其中主要是对油膜的静压作用影响较大,而动压效应相对比较稳定,动压支承力变化较小。



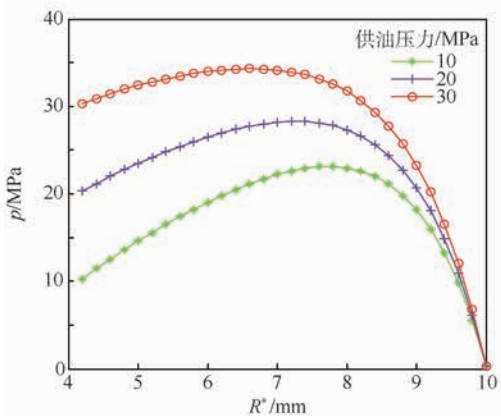
(a) 供油压力为 10 MPa



(b) 供油压力为 20 MPa



(c) 供油压力为 30 MPa



(d) 最大倾斜半径 R' 下不同供油压力的油膜压力变化

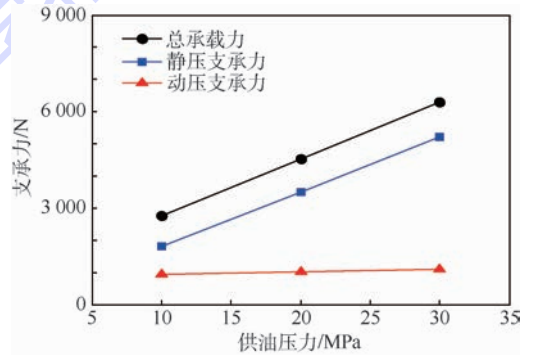


图 11 不同供油压力下的油膜动静压支承力

Fig. 11 Dynamic and static bearing force of oil film under different oil supply pressure

3.3 结构参数对抗倾覆能力的影响

数值模拟取滑靴底面工作半径 r_2 分别为 9、10、11 mm,其他仿真参数与 RZB-3 型燃油柱塞泵滑靴副仿真时的参数保持一致,取 $h_c = 11 \mu\text{m}$, $\beta_{\text{max}} = 7.5 \times 10^{-4} \text{ rad}$,出口压力为 10.29 MPa,进口压力为 0.3 MPa,转子转速为 4 620 r/min,中心油池半径固定为 5 mm,仿真结果如图 12 所示。

由图 12 可知,滑靴底面工作半径对滑靴副油膜的压力分布影响很大,随着滑靴底面工作半径的增大,油膜的静压作用区和动压效应区的压力梯度都随之增大,整个油膜的承载力有了很大的提升。通过计算得到不同滑靴底面工作半径下的动静压支承力,如图 13 所示。

图 13 可知,静压作用随着滑靴底面工作半径的增大其增强趋势比较缓慢,而动压效应对于滑靴底面工作半径的变化则十分敏感,这一点也可以从图 12 的最大压力峰值变化得到验证。故可以得到以下结论:通过增大滑靴底面工作半径可以增强油膜的动压效应,从而提高滑靴的抗倾覆能力。

取滑靴底面工作半径为固定值 10 mm,中心油池半径 r_1 分别取 4、5、6 mm,其他参数不变,数

图 10 不同供油压力下的油膜压力分布

Fig. 10 Distribution of oil film pressure under different oil supply pressure

值模拟计算结果如图 14 所示。

从图 14 的仿真结果可知,中心油池半径对滑靴副油膜压力分布的影响与滑靴底面工作半径类似,不同中心油池半径下油膜的静压作用和动压作用都会发生改变,且动压作用变化更为敏感。同样对该情况下的动静压支承力进行计算,结果如图 15 所示。

由图 15 可知,随着中心油池半径的增大,油膜总的承载力有缓慢减小的趋势,这是因为虽然油膜的静压支承力增大了,但动压支承力快速减小,这说明中心油池半径越小,油膜的动压效应越强。

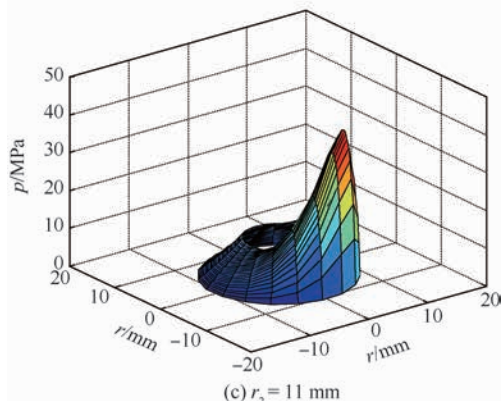
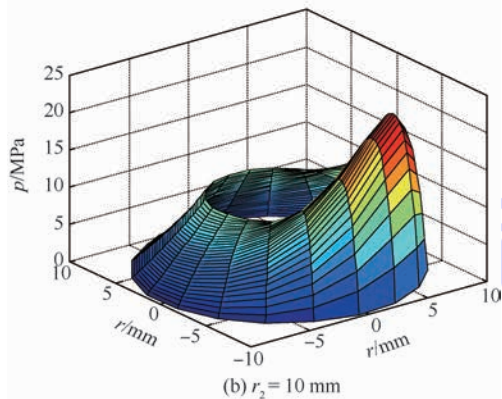
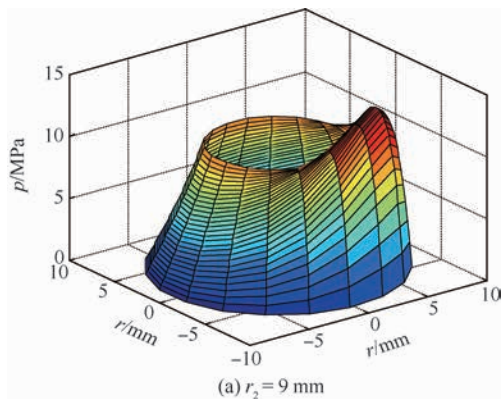


图 12 不同滑靴底面工作半径下的油膜压力分布

Fig. 12 Distribution of oil film pressure under different working radius of slipper bottom

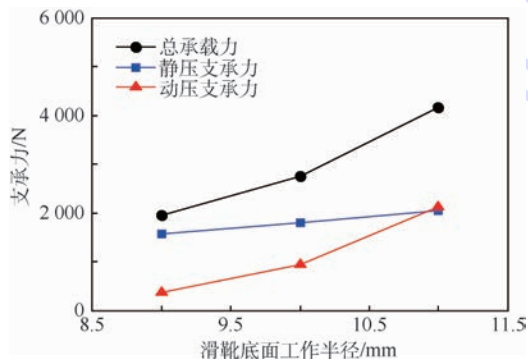


图 13 不同滑靴底面工作半径下的油膜动静压支承力
Fig. 13 Dynamic and static bearing force of oil film under different working radiuses of slipper bottom

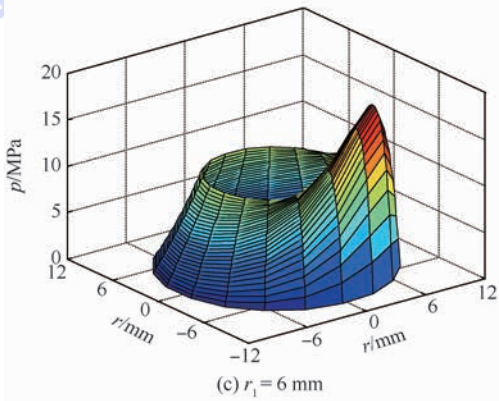
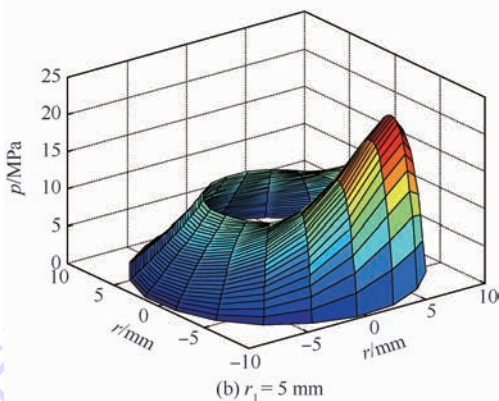
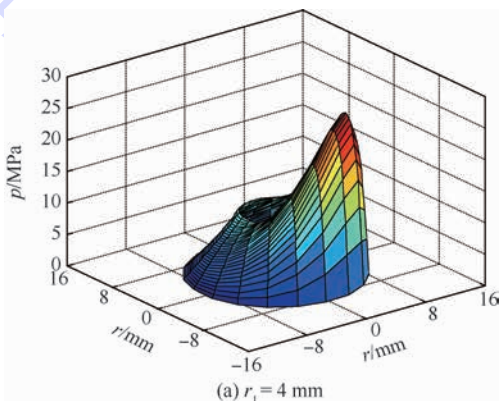


图 14 不同中心油池半径下的油膜压力分布
Fig. 14 Distribution of oil film pressure under different central oil pool radius

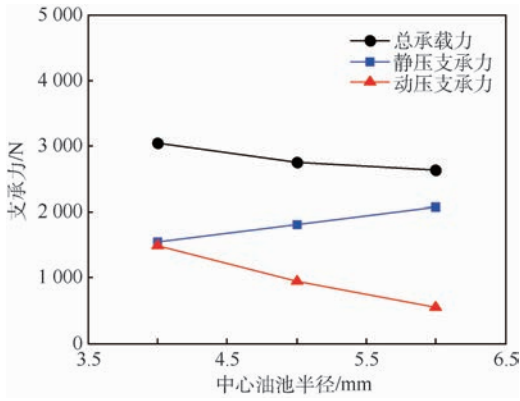


图15 不同中心油池半径下的油膜动静压支承力

Fig. 15 Dynamic and static bearing forces of oil film under different central oil pool radii

4 结论

本文考虑动静压效应对航空燃油柱塞泵滑靴副润滑特性进行了仿真研究,得到以下结论:

1) 滑靴合力的变化对于滑靴副油膜的润滑至关重要,建立滑靴副混合润滑模型时必须考虑滑靴在旋转运动过程中所受到的离心力和摩擦力;对比分析静压润滑和混合润滑机理下的油膜厚度仿真结果可知,动压效应在整个支承力中占28%,因此不可忽略,因动压效应的产生使得动静压混合支承所得到的油膜厚度整体偏小。对比滑靴在 $\alpha = 180^\circ$ 时所受合压紧力与动静压混合支承力计算结果,动静压混合支承力较滑靴合压紧力的计算误差仅为0.9%,表明动静压混合润滑现象更符合滑靴副实际的润滑状态,且所建立的动静压混合润滑数学模型是可信的。

2) 中心油膜厚度、滑靴最大倾斜角和转子转速是动压效应的主要影响因素,其产生的动压的主要作用在于保障滑靴抗倾覆能力;而滑靴副进口压力即柱塞泵供油压力是油膜的静压作用的主要影响因素,对油膜的静压作用影响占比为62%~85%,但进口压力对动压效应的变化影响甚微。滑靴底面工作半径和中心油池半径对油膜的动压效应影响非常大,其中滑靴底面工作半径增加2 mm,动压效应可增加5.59倍,中心油池半径减小2 mm,动压效应可增加2.67倍。若要提升滑靴的抗倾覆能力,可通过优化以上参数保障滑靴的抗倾覆能力。

参考文献 (References)

[1] 陈永琴. 航空燃油柱塞泵运动学与动力学特性分析研究[D]. 西安:西安电子科技大学,2012.
CHEN Y Q. Analysis of kinematics and dynamics characteristics

of aviation fuel piston pump[D]. Xi'an:Xi'an Electronic and Science University,2012(in Chinese).

[2] IVANTYSNOVA M. A new approach to the design of sealing and bearing gaps of displacement machines[C]//Proceedings of the 4th JFPS International Symposium on Fluid Power. Tokyo:JFPS,1999:45-50.

[3] WIECZOREK U,IVANTYSNOVA M. Computer aided optimization of bearing and sealing gaps in hydrostatic machine—The simulation tool CASPAR[J]. International Journal of Fluid Power,2002,3(1):7-20.

[4] PELOSI M,IVANTYSNOVA M. Heat transfer and thermal elastic deformation analysis on the piston/cylinder interface of axial piston machines[J]. Journal of Tribology,2012,134(4):119-128.

[5] SCHENK A,IVANTYSNOVA M. A transient thermoelasto-hydrodynamic lubrication model for the slipper/washplate in axial piston machines[J]. Journal of Tribology,2015,137(3):031701.

[6] DEEKEN M. Simulation of the tribological contacts in an axial piston machine[C]//ASME 2004 International Mechanical Engineering Congress and Exposition. New York:ASME,2004:71-75.

[7] WOHLERS A,MURENHOFF H. Tribological simulation of hydrostatic swash plate bearing in an axial piston pump[C]//Power Transmission and Motion Control Symposium,2007:129-144.

[8] SCHLEIHS C,VIENNET E,DEEKEN M. 3D-CFD simulation of an axial piston displacement unit[C]//9th International Fluid Power Conference,2014:332-343.

[9] 徐兵,李迎兵,张斌,等. 轴向柱塞泵滑靴副倾覆现象数值分析[J]. 机械工程学报,2010,46(20):161-168.
XU B,LI Y B,ZHANG B,et al. Numerical analysis of the phenomenon of sliding pair of axial piston pump[J]. Journal of Mechanical Engineering,2010,46(20):161-168(in Chinese).

[10] 李迎兵. 轴向柱塞泵滑靴副油膜特性研究[D]. 杭州:浙江大学,2011.
LI Y B. Axial piston pump slipper oil film characteristics of slide[D]. Hangzhou:Zhejiang University,2011(in Chinese).

[11] 刘洪,苑士华,彭增雄. 轴向柱塞泵滑靴油膜动态仿真[J]. 北京理工大学学报,2011,31(11):1282-1286.
LIU H,YUAN S H,PENG Z X. Dynamic simulation of oil film in sliding piston pump[J]. Journal of Beijing Institute of Technology,2011,31(11):1282-1286(in Chinese).

[12] 于思淼. 轴向柱塞泵用滑靴流体静动压支撑的特性分析及结构优选[D]. 哈尔滨:哈尔滨工业大学,2013.
YU S M. Characteristics analysis and structure optimization of Slipper with static and dynamic pressure support used in the axial piston pump[D]. Harbin:Harbin Institute of Technology,2013(in Chinese).

[13] 王亚军. 高压高速轴向柱塞泵滑靴性能研究[D]. 北京:北京理工大学,2014.
WANG Y J. Research on the performance of high speed axial piston pump slipper[D]. Beijing:Beijing Institute of Technology,2014(in Chinese).

[14] 魏超,胡纪滨,薛冰,等. 表面微结构对轴向柱塞泵滑靴润滑

- 特性的影响[C]//第十一届全国摩擦学大会,2013:1-7.
- WEI C, HU J B, XUE B, et al. Effect of surface microstructure on lubrication characteristics of sliding shoe of axial piston pump[C]//The 11th of Tribology Conference, 2013: 1-7 (in Chinese).
- [15] 何必海,孙健国,叶志锋.航空燃油柱塞泵静压润滑油膜计算分析[J].航空动力学报,2009,24(12):191-197.
- HE B H, SUN J G, YE Z F. Calculation and analysis of lubricating oil film of sliding piston for aviation fuel piston pump[J]. Journal of Aerospace Power, 2009, 24(12): 191-197 (in Chinese).
- [16] 何必海,孙健国,叶志锋.燃油柱塞泵滑靴副和配流副油膜计算研究[J].航空动力学报,2010,25(6):1437-1442.
- HE B H, SUN J G, YE Z F. Study on the oil film of the slipper pair and the oil distribution pair of fuel injection pump[J]. Journal of Aerospace Power, 2010, 25(6): 1437-1442 (in Chinese).
- [17] 徐佩佩,叶志锋,王彬.航空燃油柱塞泵滑靴油膜的多目标优化设计[J].航空动力学报,2014,29(8):1981-1986.
- XU P P, YE Z F, WANG B. Multi objective optimization design of oil film for sliding piston oil pump[J]. Journal of Aerospace Power, 2014, 29(8): 1981-1986 (in Chinese).
- [18] 林硕,苑士华,刘洪.考虑油膜非均匀性的滑靴润滑特性研究[J].北京理工大学学报,2014,34(4):358-362.
- LIN S, YUAN S H, LIU H. Study on the lubrication characteristics of sliding shoes considering the heterogeneity of oil film [J]. Journal of Beijing Institute of Technology, 2014, 34(4): 358-362 (in Chinese).

作者简介:

符江锋 男,博士,助理研究员。主要研究方向:航空发动机控制、航空燃油泵建模与仿真。

Numerical simulation of hybrid lubrication characteristics of slipper pair of aviation fuel piston pump

FU Jiangfeng*, LI Huacong, ZENG Xiaqing, LIU Xianwei

(College of Power and Energy, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: The aim of this paper was to solve the slipper pair lubrication problem of the aero fuel piston pump under the mixed support of dynamic and static pressure. Based on the mathematical model of kinematics and dynamics of slipper pair, the mathematical model of the hybrid lubrication of the slipper pair was established based on the dynamic pressure effect caused by the static pressure support and the irregular spatial curve of sliding shoe movement. Then the lubrication characteristics of slipper pair were simulated based on the finite volume method, and the variation law of the oil film thickness, the influence factors of the oil film pressure distribution and the anti-sliding performance of the sliding shoe were analyzed. The simulation results show that the variation trend of the oil film thickness obtained by the hybrid lubrication is consistent with the actual lubrication state, the thickness of the oil film, the maximum inclination angle of the slipper and the rotor speed have an important influence on the dynamic pressure effect, and the inlet pressure of the slipper pair mainly affects the static pressure of the oil film. The anti-overturning ability of the sliding shoe can be enhanced by increasing the working radius of the bottom surface of the sliding shoe or reducing the radius of the oil pool of the sliding shoe center, so as to counteract the overturning moment of the sliding shoe.

Keywords: fuel piston pump; slipper pair; dynamic pressure effect; lubrication characteristics; static pressure support

Received: 2017-05-15; Accepted: 2017-08-11; Published online: 2017-09-27 18:21

URL: kns.cnki.net/kcms/detail/11.2625.V.20170927.1821.003.html

Foundation items: Aeronautical Science Foundation of China (20150653006); the Fundamental Research Funds for the Central Universities (G2017KY0003)

* Corresponding author. E-mail: fjj@nwpu.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2017.0365

采用 FRM 的窄过渡带奇型排列结构改进方法



张文旭*, 赵文童, 陈涛, 陈亚静

(哈尔滨工程大学 信息与通信工程学院, 哈尔滨 150001)

摘 要: 为了解决数字信道化中窄过渡带滤波器组计算复杂度较高的问题,提出了一种采用频率响应屏蔽(FRM)的窄过渡带奇型排列非最大抽取滤波器组高效结构。在设计该结构的原型滤波器时,首先通过对原型滤波器进行插值得到 FRM 上支路的滤波器,然后将信号延时与插值后的滤波器相减得到位于下支路的互补滤波器,再利用屏蔽滤波器对插值产生的多余镜像进行屏蔽,最终得到所设计的 FRM 滤波器。对该改进结构进行了 MATLAB 仿真,仿真结果验证了该改进结构的正确性。与多相滤波器组结构相比,采用该结构可节省约 80% 的乘法器资源。由于该改进结构中将抽取提前,降低了系统采样率的限制条件,可以直接应用于高速采样系统中。

关键词: 窄过渡带; 原型滤波器; 频率响应屏蔽(FRM); 滤波器组; 屏蔽滤波器

中图分类号: TN971. *1

文献标识码: A **文章编号:** 1001-5965(2018)05-0951-08

在数字信道化系统设计中,数字信道化的过程实质上就是利用滤波器组进行均匀或非均匀设计的过程。滤波器组技术作为信号处理的一种常见方法和手段,其应用范围很广,如通信、图像处理、语音、雷达等^[1-3]。其中,具有窄过渡带特点的滤波器组计算复杂度高的问题业已成为该研究方向亟待解决的问题之一^[4]。频率响应屏蔽(FRM)技术因其在解决窄过渡带滤波器组设计中具有明显优势而得以推广^[5]。

文献[6]运用凸规划算法设计 FRM 结构中的滤波器,并提出了 3 种优化算法;文献[7]针对工程应用方面在经典 FRM 滤波器结构的基础上优化设计二级 FRM 滤波器,这种优化方法虽然可以消除内插因子的约束条件,但是屏蔽滤波器的设计较为复杂;文献[8-9]利用 FRM 技术设计任意带宽的 FIR 数字滤波器,降低了计算的复杂度,

但是因为滤波器都在这 2 种结构的前置位,所以这 2 种结构都会受到采样率的限制;文献[10]提出了一种基于 FRM 的 DFT 滤波器组优化算法,相比于传统的 DFT 滤波器组在一定程度上降低了计算复杂度。重构滤波器组也是滤波器组研究的重要组成部分,文献[11]通过窗函数设计 FRM 滤波器,虽然设计的综合滤波器组计算复杂度很高,但是为重构滤波器设计提供了一种新的思路;文献[12]将屏蔽技术应用到非均匀滤波器组的结构构造中,根据多速率理论提出了一种低复杂度的综合滤波器设计方法;文献[13]运用 FRM 结构设计两通道窄过渡带的完全重构滤波器组,缺点是在实现时计算量比较大。文献[13]提出了一种基于 FRM 的偶型排列信道化改进结构,缺点是偶型排列结构的第 0 个信道输出 IQ 量不正交,需要对输出信号单独进行希尔伯特变换,这样

收稿日期: 2017-05-31; 录用日期: 2017-11-23; 网络出版时间: 2017-12-20 10:01

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20171219.1632.006.html

基金项目: 国家自然科学基金(61571146)

*通信作者. E-mail: zhangwenxu@hrbeu.edu.cn

引用格式: 张文旭, 赵文童, 陈涛, 等. 采用 FRM 的窄过渡带奇型排列结构改进方法[J]. 北京航空航天大学学报, 2018, 44(5): 951-958. ZHANG W X, ZHAO W T, CHEN T, et al. An improved approach for odd-arrangement structure with narrow transition band based on FRM[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 951-958 (in Chinese).

第0个信道的延时会与其他信道不同。

本文将FRM技术在窄过渡带滤波器组设计中的优势推广到分析滤波器组高效结构设计中,由于奇型排列结构所有信道输出特性完全一致,不像偶型排列结构中第0个信道输出IQ量不正交,故而本文主要以奇型排列结构为研究重点。本文提出了一种采用FRM的窄过渡带奇型排列非最大抽取改进结构,该改进结构针对非最大抽取系统进行设计,不受信道数和抽取数的限制,适用性更加广泛。

1 频率响应屏蔽技术

FRM技术适用于窄过渡带滤波器设计,因为其在降低滤波器阶数和减少计算复杂度方面有很好的效果。FRM滤波器的基本框架如图1所示,其传递函数的z变换可表示为^[14-15]

$$H(z) = H_a(z^L)H_{Ma}(z) + (z^{-L(N_a-1)/2} - H_a(z^L))H_{Mc}(z) \quad (1)$$

式中: $H_a(z)$ 为原型滤波器; $H_{Ma}(z)$ 和 $H_{Mc}(z)$ 为屏蔽滤波器; N_a 为原型滤波器 $H_a(z)$ 的长度; L 为插值倍数。

滤波器插值过程如图2所示,上图为滤波器插值之前的幅频特性,下图为滤波器插值之后的幅频特性。屏蔽滤波器的作用是对插值所产生的滤波器多余镜像进行屏蔽。

$$H_a(z^L) = \sum_{n=0}^{N_a-1} h_a(n)z^{-nL} \quad (2)$$

$$H_{Ma}(z) = z^{-M_1} \sum_{n=0}^{N_1} h_{Ma}(n)z^{-n} \quad (3)$$

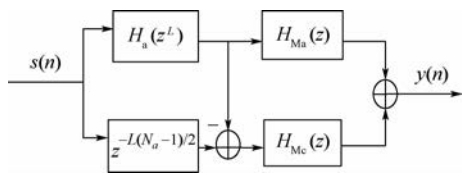


图1 FRM滤波器的基本框架
Fig.1 Basic framework of FRM filter

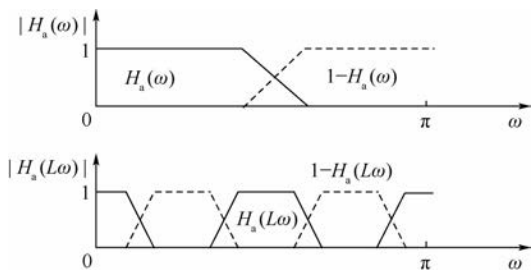


图2 滤波器插值过程(L=4)

Fig.2 Process of filter interpolation for L=4

$$H_{Mc}(z) = z^{-M_2} \sum_{n=0}^{N_2} h_{Mc}(n)z^{-n} \quad (4)$$

式中: $h_a(n)$ 、 $h_{Ma}(n)$ 和 $h_{Mc}(n)$ 的系数具有偶对称性; N_1 和 N_2 分别为屏蔽滤波器 $H_{Ma}(z)$ 和 $H_{Mc}(z)$ 的长度; M_1 和 M_2 的取值情况如下:

$$\begin{cases} M_1 = 0, M_2 = (N_1 - N_2)/2 & N_1 \geq N_2 \\ M_1 = (N_1 - N_2)/2, M_2 = 0 & N_1 < N_2 \end{cases} \quad (5)$$

将信号延时 $L(N_a - 1)/2$ 个点后与 $H_a(z^L)$ 相减可以得到下支路的输出。考虑到插值 L 的数值不被限制,滤波器 $H_a(z)$ 长度 N_a 应为奇数。在对2个屏蔽滤波器的输出进行合成时,要求滤波器 $H_{Ma}(z)$ 和 $H_{Mc}(z)$ 长度同为奇数或者同为偶数。

令零相位 $H(\omega)$ 的频率响应为

$$H(\omega) = H_1(\omega) + H_2(\omega) \quad (6)$$

又因为 $z^{-L(N_a-1)/2} - H_a(z^L)$ 可看作 $H_a(z)$ 进行 L 倍插值后的互补滤波器,则

$$H_1(\omega) = H_a(L\omega)H_{Ma}(\omega) \quad (7)$$

$$H_2(\omega) = (1 - H_a(L\omega))H_{Mc}(\omega) \quad (8)$$

式中: $H_a(\omega)$ 、 $H_{Ma}(\omega)$ 和 $H_{Mc}(\omega)$ 分别为 $H_a(z)$ 、 $H_{Ma}(z)$ 和 $H_{Mc}(z)$ 的零相位频率响应。

FRM滤波器频带合成过程图^[12,16]如图3所示,从上至下第1幅图中实线为原型滤波器 $H_a(\omega)$ 插值后的幅频特性,虚线为屏蔽滤波器 $H_{Ma}(\omega)$ 的幅频特性;第2幅图中实线为原型滤波器 $H_a(\omega)$ 的互补滤波器插值后的幅频特性,虚线为屏蔽滤波器 $H_{Mc}(\omega)$ 的幅频特性;第3幅图中实线为对插值后的原型滤波器进行屏蔽得到的 $H_1(\omega)$ 幅频特性,虚线为对插值后的原型滤波器的互补滤波器进行屏蔽得到的 $H_2(\omega)$ 幅频特性;第4幅图为最终得到的FRM滤波器的幅频特性。

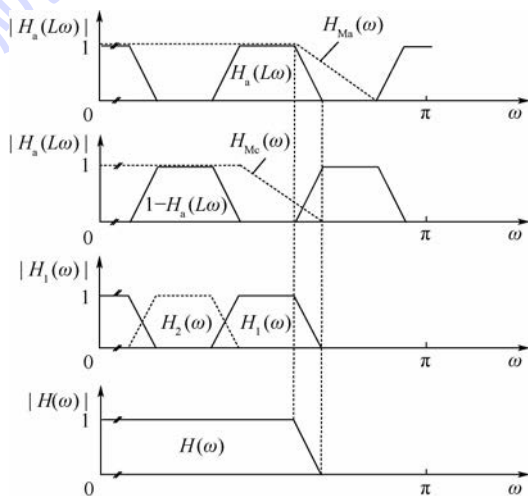


图3 采用FRM技术设计低通滤波器^[12,16]

Fig.3 Design of low-pass filter based on FRM^[12,16]

2 调制滤波器组

2.1 滤波器组频带划分

调制滤波器组在调制类型划分方面分为 2 种类型,分别为余弦调制和复指数调制。余弦调制可以看作复指数调制的特例。调制滤波器组在频带划分结构设计上也分为 2 种类型,分别为奇型排列和偶型排列。针对实际问题,笔者研究了调制滤波器组的奇型排列方式,并且对滤波器组中的 M 个滤波器进行均匀调制,滤波器组频带划分方式如图 4 所示。

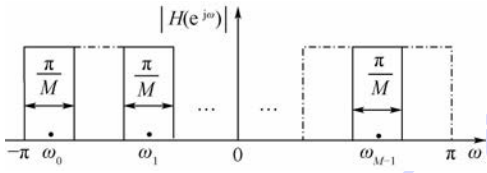


图 4 滤波器组频带划分

Fig. 4 Band division of filter bank

图 4 所示的奇型排列的划分方式避免了因为实信号造成的信道冗余问题,而且能够实现信号的无盲区无混叠接收。上述排列方式的虚线频带为对应的镜像,第 k 个信道对应的中心频率 ω_k 满足:

$$\omega_k = \left(k - \frac{2M-1}{4} \right) \frac{2\pi}{M} \quad k = 0, 1, \dots, M-1 \quad (9)$$

2.2 高效信道化结构

均匀滤波器组低通实现结构如图 5 所示。假设信道数为 M ,抽取倍数为 K ,令 $M = FK$,令 F 取正整数(当 $F = 1$ 时为临界抽取情况),输入信号为 $s(n)$,复指数调制因子为 $e^{j\omega_k n}$,低通滤波器为 $g(n)$,原型滤波器阶数为 N ,多相滤波器阶数为 $T = N/M$ 。输入信号经过信道化后的输出信号表示为 $y_k(m)$, $k = 0, 1, \dots, M-1, m = 0, 1, \dots$ 。

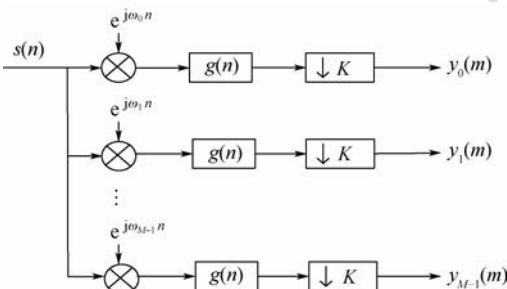


图 5 滤波器组低通实现结构

Fig. 5 Low-pass structure of filter bank

如图 5 所示,第 k 个信道的输出信号可以表示为

$$y_k(m) = (s(n) e^{j\omega_k n}) \otimes g(n) \Big|_{n=mK} = \sum_{i=0}^{N-1} s(n-i) e^{j\omega_k(n-i)} g(i) \Big|_{n=mK} = \sum_{i=0}^{N-1} s(mK-i) e^{j\omega_k(mK-i)} g(i) \quad (10)$$

式中: \otimes 表示卷积运算; $k = 0, 1, \dots, M-1$ 。

令信号 K 倍抽取后的多相结构表达式为 $s_p(m) = s(mK-p)$,滤波器多相分量表达式为 $g_p(m) = g(mM+p)$,令 $i = iM+p, p = 0, 1, \dots, M-1$,则式(10)可以表示为

$$y_k(m) = \sum_{p=0}^{M-1} \sum_{i=0}^{T-1} s(mK-iM-p) e^{j\omega_k(mK-iM-p)} g(iM+p) = \sum_{p=0}^{M-1} \sum_{i=0}^{T-1} s_p(m-iF) e^{j\omega_k(m-iF)K} g_p(i) e^{-j\omega_k p} \quad (11)$$

令 $l = iF, h_p(i) = g_p(i/F)$,则 $h_p(i)$ 是 $g_p(i)$ 的 F 倍内插,第 k 个信道的输出信号表示为

$$y_k(m) = \sum_{p=0}^{M-1} \sum_{l=0}^{(T-1)F} s_p(m-l) e^{j\omega_k(m-l)K} h_p(l/F) e^{-j\omega_k p} = \sum_{p=0}^{M-1} \sum_{l=0}^{(T-1)F} s_p(m-l) e^{j\omega_k(m-l)K} h_p(l) e^{-j\omega_k p} \quad (12)$$

则式(12)的 z 变换可以表示为

$$y_k(z) = \sum_{m=0}^{+\infty} y_k(m) z^{-m} = \sum_{p=0}^{M-1} s_p(z e^{-j\omega_k K}) h_p(z) e^{-j\omega_k p} \quad (13)$$

滤波器组中的 M 个滤波器通带带宽减少为之前的 $1/M$,所以在对滤波器 $K(K \leq M)$ 倍抽取时不会产生混叠现象。将中心频率 ω_k 代入式(13)中, K 倍抽取提前,可得对应的奇型排列信道化高效结构,如图 6 所示。该高效结构将抽取模块提到信道化结构最前端,使之适应于高速系统中。

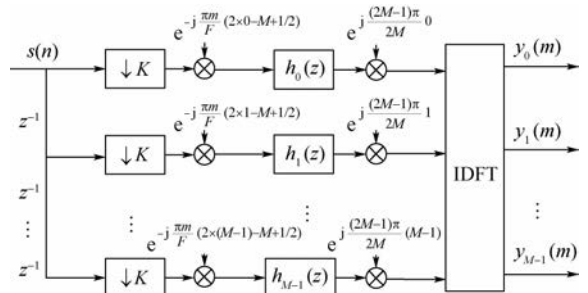


图 6 滤波器组信道化高效结构

Fig. 6 Efficient channelization structure of filter bank

3 窄过渡带奇型排列高效结构

3.1 FRM 高效结构

设 FRM 原型滤波器 z 变换表达式 $P(z)$ 为

$$P(z) = H_{ap}(z^L)H_{Map}(z) + H_{cp}(z^L)H_{Mcp}(z) \quad (14)$$

式中: $H_{ap}(z)$ 、 $H_{Map}(z)$ 和 $H_{Mcp}(z)$ 分别为长度为 N_{ao} 、 N_{Ma} 和 N_{Mc} 的零相位对称滤波器; $H_{cp}(z) = 1 - H_{ap}(z)$ 。

可以得到滤波器的因果线性相位形式为

$$H_{ao}(z) = z^{-L(N_{ao}-1)/2} H_{ap}(z^L) \quad (15)$$

$$H_{co}(z) = z^{-L(N_{co}-1)/2} H_{cp}(z^L) \quad (16)$$

$$H_{Ma}(z) = z^{-(N_{Mac}-1)/2} H_{Map}(z) \quad (17)$$

$$H_{Mc}(z) = z^{-(N_{Mac}-1)/2} H_{Mcp}(z) \quad (18)$$

式中: $N_{Mac} = \max\{N_{Ma}, N_{Mc}\}$ 。

将式(15)~式(18)代入式(14),则可以得到 FRM 滤波器组的各通道传递函数表达式为

$$H_k(z) = W_M^{ak} H_{ao}(zW_M^k) H_{Ma}(zW_M^k) + W_M^{ak} H_{co}(zW_M^k) H_{Mc}(zW_M^k) \quad (19)$$

式中: $k = 0, 1, \dots, M-1$; $a = (N_p - 1)/2$, $N_p = L(N_{ao} - 1) + N_{Mac}$; $W_M = e^{-j2\pi/M}$ 。

将滤波器的因果线性相位形式进行多相分解,可得如下表达式:

$$H_{ao}(z) = \sum_{l=0}^{M-1} z^{-l} H_{ao,l}(z^M) \quad (20)$$

$$H_{co}(z) = \sum_{l=0}^{M-1} z^{-l} H_{co,l}(z^M) \quad (21)$$

$$H_{Ma}(z) = \sum_{p=0}^{M-1} z^{-p} H_{Ma,p}(z^M) \quad (22)$$

$$H_{Mc}(z) = \sum_{p=0}^{M-1} z^{-p} H_{Mc,p}(z^M) \quad (23)$$

式中: $l = 0, 1, \dots, M-1$; $p = 0, 1, \dots, M-1$ 。

将式(20)~式(23)代入式(19),即可得到 FRM 滤波器组各通道传递函数的多相结构表达式为

$$H_k(z) = W_M^{ak} \sum_{l=0}^{M-1} (zW_M^k)^{-l} H_{ao,l}(z^M) \sum_{p=0}^{M-1} (zW_M^k)^{-p} \cdot H_{Ma,p}(z^M) + W_M^{ak} \sum_{l=0}^{M-1} (zW_M^k)^{-l} \cdot H_{co,l}(z^M) \sum_{p=0}^{M-1} (zW_M^k)^{-p} H_{Mc,p}(z^M) \quad (24)$$

在对 L 倍插值之后的滤波器进行屏蔽时,需要 M 通道屏蔽滤波器组的正好可以覆盖插值产生的多余镜像,即 $L = IM, I = 1, 2, \dots$ 。原型滤波器和屏蔽滤波器需要满足:

$$\omega_{map} = \omega_{ap}/L + \pi(L/M' - 1)/L \quad (25)$$

$$\omega_{mas} = \pi/L + \omega_{ap}/L + \pi(L/M' - 1)/L \quad (26)$$

$$\omega_{mcp} = \omega_{as}/L + \pi(L/M' - 1)/L \quad (27)$$

$$\omega_{mcs} = \pi/L + \omega_{as}/L + \pi(L/M' - 1)/L \quad (28)$$

式中: M' 为信道数; ω_{ap} 和 ω_{as} 分别为 H_{ap} 的通带和阻带截止频率; ω_{map} 、 ω_{mas} 、 ω_{mcp} 、 ω_{mcs} 分别为屏蔽滤波器 $H_{Map}(z)$ 和 $H_{Mcp}(z)$ 的通带截止频率和阻带截止频率。

又由于实信号的频谱应为正值,且关于 $k\pi(k=0,1,\dots)$ 对称,只需考虑 $0 \sim \pi$ 的信道信号,则数值上 $M' = 2M$, M 表示实际应用信道数数目。

假设上半个支路的屏蔽滤波器为原型屏蔽滤波器,那么对原型屏蔽滤波器进行 π/L 频移就能够得到下半支路的屏蔽滤波器,然后再把上下 2 个支路进行合并就能够得到带宽为 π/M 的子带。FRM 滤波器频率响应的合成过程如图 7 所示,从上至下第 1 幅图中实线为原型滤波器 $H_a(\omega)$ 的幅频特性,虚线为原型滤波器 $H_a(\omega)$ 的互补滤波器的幅频特性;第 2 幅图中实线为插值后的原型滤波器 $H_a(\omega)$ 的幅频特性,虚线为插值后的原型滤波器 $H_a(\omega)$ 的互补滤波器的幅频特性;第 3 幅图为屏蔽滤波器 $H_{Ma}(\omega)$ 的幅频特性;第 4 幅图为屏蔽滤波器 $H_{Mc}(\omega)$ 的幅频特性;第 5 幅图为最终合成的 FRM 滤波器 $H(\omega)$ 的幅频特性。

将 FRM 技术引入信道化高效结构中的滤波器设计中。用式(24)中 FRM 滤波器组的 M 个滤波器代替式(13)中滤波器组中的 M 个滤波器,可以得到第 k 个信道输出信号的表达式为

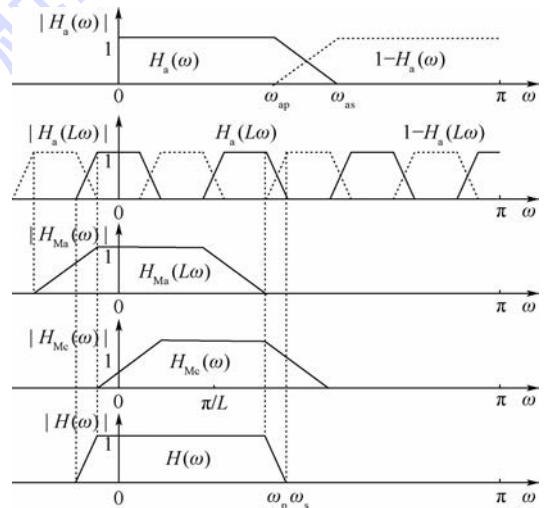


图7 FRM 滤波器的频率响应合成过程

Fig. 7 Frequency response synthesis process of FRM filter

$$y_k(z) = W_M^{ak} \sum_{l=0}^{M-1} s_l(z e^{-j\omega_l K}) z^{-l} H_{ao,l}(z^M) \cdot \sum_{p=0}^{M-1} (z W_M^k)^{-p} H_{Ma,p}(z^M) e^{j \frac{(2M-1)\pi}{2M} p} + W_M^{ak} \sum_{l=0}^{M-1} s_l(z e^{-j\omega_l K}) z^{-l} H_{co,l}(z^M) \cdot \sum_{p=0}^{M-1} (z W_M^k)^{-p} H_{Mc,p}(z^M) e^{j \frac{(2M-1)\pi}{2M} p} \quad (29)$$

又因为 $H_{co}(z) = z^{-L(N_{ao}-1)/2} - H_{ao}(z)$, 满足:

$$H_{co,l}(z^M) = \begin{cases} z^{-L(N_{ao}-1)/2+q} - H_{ao,q}(z^M) & l = q \\ -H_{ao,l}(z^M) & \text{其他} \end{cases} \quad (30)$$

式中: q 为 $L(N_{ao}-1)/(2M)$ 的余数, $q = 0, 1, \dots, M-1$ 。

由式(29)、式(30)结合图 6, 将 K 倍抽取提前, 可得基于 FRM 的窄过渡带滤波器组高效结构, 如图 8 所示。

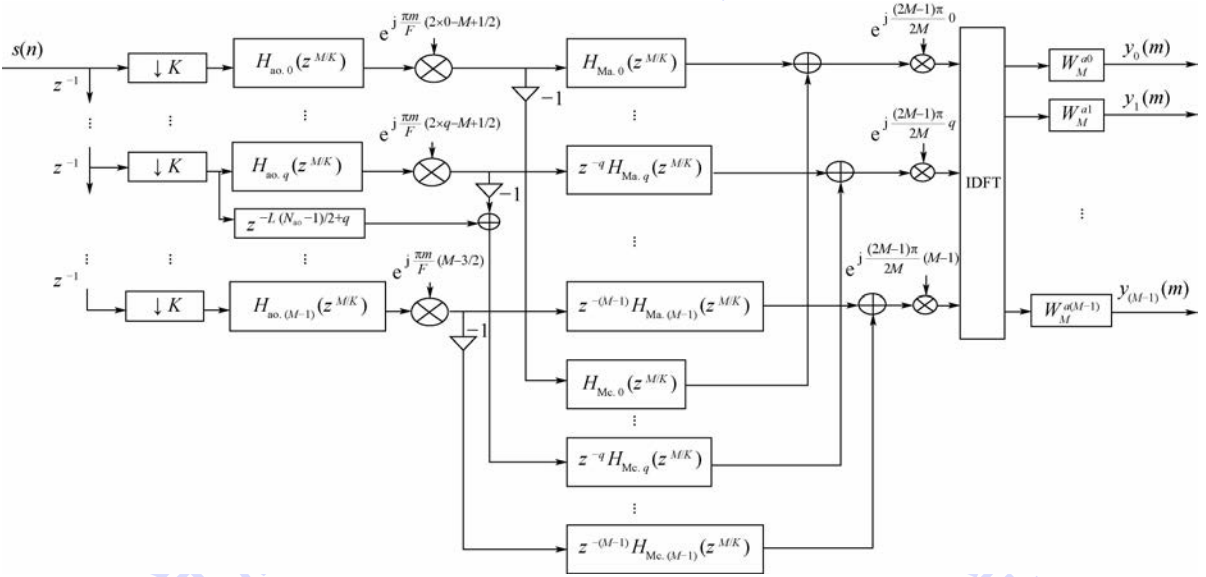


图 8 采用 FRM 的窄过渡带奇型排列高效结构

Fig. 8 Odd-arrangement efficient structure with narrow transition band based on FRM

3.2 奇型排列高效结构复杂度分析

高效结构的计算复杂度主要受到硬件实现所需的乘法器数量影响, 因此复杂度数值用于表征高效结构中的乘法器数量。滤波器中所用的乘法器数目与滤波器长度成反比, 即复杂度数值与过渡带宽成反比。令实现滤波器组的代价函数为 C 。

实现一个均匀滤波器组所需要的乘法器数目为

$$C_r = M\gamma/\Delta f \quad (31)$$

式中: γ 为比例系数; M 为信道数; Δf 为窄过渡带滤波器的归一化过渡带宽。

同理, 实现一个滤波器组的多相结构的复杂度为

$$C_p = \gamma/\Delta f + M \ln M + 2M \quad (32)$$

本文设计方法实现的代价函数为

$$C_{FRM} = \gamma/\Delta H_a + 2\gamma/\Delta H_m + 2M \ln M + 2M \quad (33)$$

式中: ΔH_a 为原型滤波器的归一化过渡带宽; ΔH_m 为屏蔽滤波器的归一化过渡带宽。

在 FRM 滤波器中, 假设 ω_p 为通带截止频率, ω_s 为阻带起始频率, 其归一化过渡带宽为 $\Delta f =$

$\omega_s - \omega_p$; θ 为原型滤波器的通带截止频率, ϕ 为原型滤波器的阻带截止频率; 插值倍数为 L , 则原型滤波器的过渡带宽可以表示为 $\Delta H_a = L\Delta f$ 。通带和阻带纹波对滤波器设计复杂度影响并不敏感, 故而暂不考虑纹波的影响。

原型滤波器采用半带滤波器。由图 7 可知, 屏蔽滤波器的过渡带宽为 $\Delta H_m = 1/L$ 。将所有滤波器的过渡带宽表达式代入式(33)可得到

$$C_{FRM} = \gamma/(L\Delta f) + 2\gamma L + 2M \ln M + 2M \quad (34)$$

内插因子 L 数值的选择会影响整个原型滤波器和屏蔽滤波器的设计复杂度。根据复杂度最低准则, 可以得到最优 L 值的表达式为

$$L = \frac{1}{\sqrt{2\Delta f}} \quad (35)$$

因为 L 为正整数, 取

$$L = \text{round}\left(\frac{1}{\sqrt{2\Delta f}}\right) \quad (36)$$

式中: $\text{round}(X)$ 表示取距离 X 最近的整数。当要求 L 为 M 的整数倍时, L 取值为 M' 的最小倍数。 L 值代入式(33)可得

$$C_{FRM} = \frac{\gamma}{\text{round}\left(\frac{1}{\sqrt{2}\Delta f}\right)\Delta f} + 2M\text{lb } M + 2\gamma\text{round}\left(\frac{1}{\sqrt{2}\Delta f}\right) + 2M \quad (37)$$

根据式(31)~式(33)可知,本文 FRM 改进方法可以降低滤波器组的计算复杂度。在设计原型低通滤波器时采用半带滤波器,因为半带滤波器有近一半的值为零,因此可以进一步有效减少计算量。在设计 2 个屏蔽滤波器时,可以先设计其中一个屏蔽滤波器,再通过调制这个屏蔽滤波器得到另一个屏蔽滤波器。通过这种设计方法得到的 2 个屏蔽滤波器的通带宽度和阻带宽度都是相等的,所以可以使设计复杂度再次降低。

4 系统仿真及分析

4.1 FRM 滤波器组仿真与复杂度分析

设置滤波器组通道数 $M = 8$, 内插因子 $L = 16$, 窄过渡带归一化过渡带宽为 0.004。设计原型滤波器时采用半带滤波器,利用 MATLAB 中的 firpm 函数进行设计,通带和阻带截止频率分别为 0.468 和 0.532,长度 $N_{ao} = 79$,滤波器的幅频响应如图 9 所示。由式(25)、式(26)可得原型屏蔽滤波器的通带截止频率为 0.029,阻带截止频率为 0.092,长度 $N_{Mac} = 82$,幅频响应如图 10 所示。合成之后的 FRM 滤波器幅频响应如图 11 所示。所设计的 FRM 滤波器组幅频响应如图 12 所示。

在设计得到的滤波器过渡带相同的情况下,直接实现这样的滤波器组结构需要乘法器数量为 $1280 \times 8 = 10240$,滤波器组的多相实现结构需要乘法器数量为 $1280 + 8\text{lb } 8 + 16 = 1307$ 。在采用了 79 阶数半带滤波器的设计中,实际非零系数为 40,则本文提出的窄过渡带滤波器组改进结构所需的乘法器数为 $40 + 86 \times 2 + 2 \times 8\text{lb } 8 + 16 = 251$ 。而文献[9]需要乘法器数目为 $125 + 150 +$

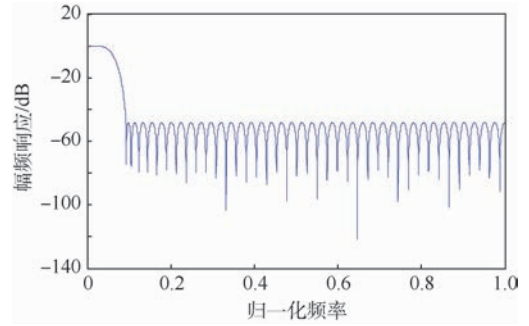


图 10 原型屏蔽滤波器幅频响应

Fig. 10 Amplitude-frequency response of prototype masking filter

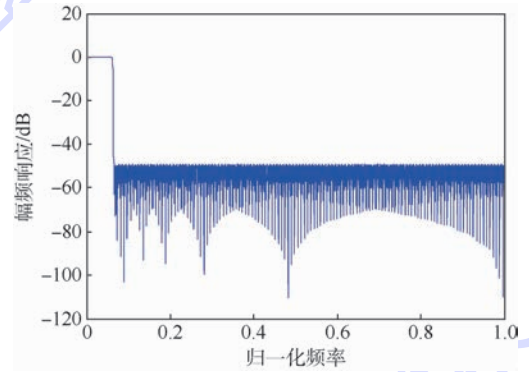


图 11 FRM 滤波器幅频响应

Fig. 11 Amplitude-frequency response of FRM filter

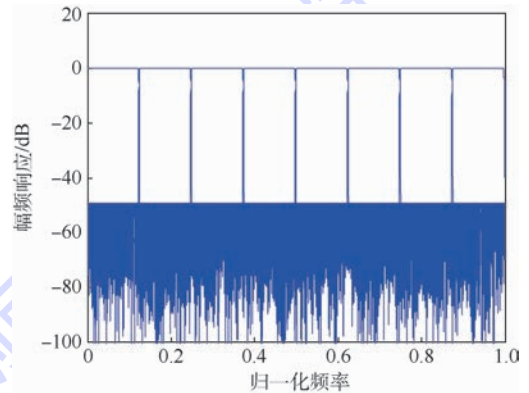


图 12 FRM 滤波器组幅频响应

Fig. 12 Amplitude-frequency response of FRM filter bank

$146 + 2 \times 8\text{lb } 8 + 16 = 460$ 。通过对比本文提出的窄过渡带滤波器组改进结构,可以发现该改进结构比直接实现结构节省了 97.5% 的乘法器资源,比多相滤波器组高效结构节省了 80.8% 的乘法器资源,比文献[9]提出的高效结构节省了 45.4% 的乘法器资源。因此,本文中所提出的窄过渡带滤波器组改进结构实现了计算复杂度的降低,更加易于工程应用实现。

4.2 FRM 高效结构性能仿真与分析

用 MATLAB 进行仿真时,设置采样频率为

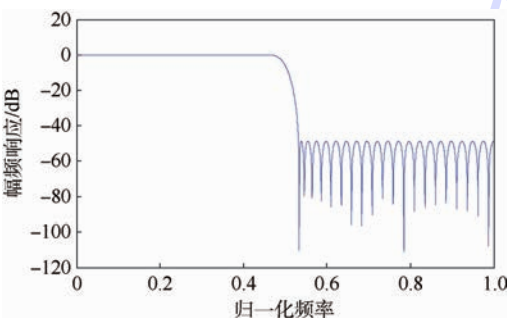


图 9 原型滤波器幅频响应

Fig. 9 Amplitude-frequency response of prototype filter

960 MHz,信道数为8。同时输入7个信号,信号各项参数设置如表1所示。

图13为FRM高效结构的仿真输出信号频谱,LFM1跨信道3、4输出;AM1在信道5输出;LFM2在信道2输出;DSB1在信道6输出;LFM3在信道1输出;DSB2在信道7输出;AM2在信道0输出。通过将得到的信号频谱图与仿真输入信号进行对比,可以验证出FRM高效结构的正确性。

LFM1的起始频率60 MHz处于3、4信道的中心处,截止频率为80 MHz,故LFM1出现了跨信道的现象。LFM3的起始频率为320 MHz,截止频率为358 MHz,其频谱位于信道1和7的中心附近,由于滤波器组过渡带窄的原因并未出现跨信道现象。采用了FRM窄过渡带设计结构没有产生信号跨信道的现象。所以采用FRM窄过渡带的信道化结构可以减少信号跨信道现象,但是想要完全消除这种现象还是无法做到的。

表1 输入信号参数设置

Table 1 Parameter setting of input signals

输入信号	载波频率/ MHz	调制频率/ MHz	理论输出 信道编号
LFM1	70	-10 ~ +10	3, 4
AM1	140	1	5
LFM2	205	-15 ~ +15	2
DSB1	285	0.7	6
LFM3	339	-19 ~ +19	1
DSB2	400	0.8	7
AM2	440	2	0

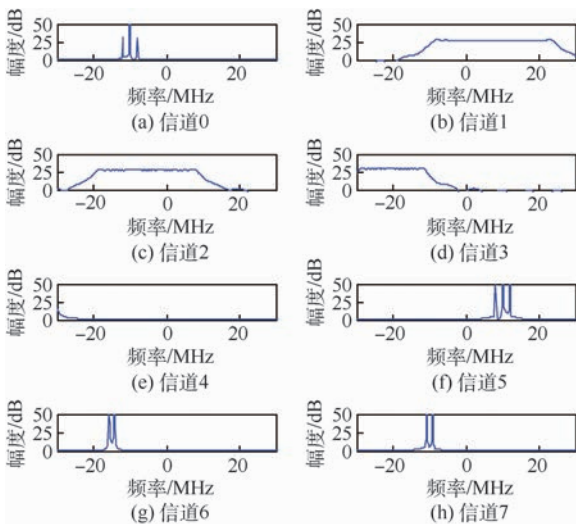


图13 FRM滤波器组输出频域波形

Fig. 13 Output frequency domain waveform of FRM filter bank

5 结论

1) 本文提出的一种采用FRM的窄过渡带奇

型排列改进结构以通用性较强的非最大抽取系统为设计出发点,具有广泛的适用性。

2) 结构中下采样模块位于该高效结构的前置位,解决了由于采样率的限制不能应用于高速信号处理系统的问题。

3) 该改进结构相比于多相滤波高效结构具有更低的计算复杂度,为工程化实现奠定了理论与技术基础。

参考文献 (References)

- [1] HARRIS F J, DICK C, RICE M. Digital receivers and transmitters using polyphase filter banks for wireless communication [J]. IEEE Transactions on Microwave Theory and Techniques, 2003, 51(4): 1395-1412.
- [2] BARROS A K, RUTKOWSKI T, LTAKURA F. Estimation of speech embedded in a reverberant and noisy environment by independent component analysis and wavelets [J]. IEEE Transactions on Neural Networks, 2002, 13(4): 888-893.
- [3] DO M N, VETTERLI M. The contourlet transform: An efficient directional multiresolution image representation [J]. IEEE Transactions on Image Processing, 2005, 14(12): 2091-2106.
- [4] HONG Y, LIAN Y. Continuous-time FIR filters based on frequency-response masking technique [C] // 2015 IEEE International Conference on Digital Signal Processing. Piscataway, NJ: IEEE Press, 2015: 191-195.
- [5] ROMERO D E T. High-speed multiplierless frequency response masking (FRM) FIR filters with reduced usage of hardware resources [C] // 2015 IEEE 58th International Midwest Symposium on Circuits and Systems. Piscataway, NJ: IEEE Press, 2015: 1-4.
- [6] LU W, HINAMOTO T. A unified approach to the design of interpolated and frequency-response-masking FIR filters [J]. IEEE Transactions on Circuits and Systems I: Regular Papers, 2016, 63(12): 2257-2266.
- [7] WEI Y, HUANG S G, MA X J. A novel approach to design low-cost two-stage frequency response masking filters [J]. IEEE Transactions on Circuits and Systems, 2015, 62(10): 982-986.
- [8] SHEN T, LIM Y C. Low complexity frequency-response masking filters using modified structure based on serial masking [C] // 19th European Signal Processing Conference. Kessariani: EURASIP, 2011: 1400-1404.
- [9] 陈涛, 王莹, 刘勇. 基于频率响应屏蔽的窄过渡带信道化接收机 [J]. 吉林大学学报, 2015, 45(1): 335-340.
CHEN T, WANG Y, LIU Y. Digital channelized receiver with narrow transition band based on FRM filter [J]. Journal of Jilin University, 2015, 45(1): 335-340 (in Chinese).
- [10] LI N, NOWROUZIAN B. Application of frequency-response masking technique to the design of a novel modified-DFT filter banks [C] // IEEE ISCAS, 2006: 3293-3296.
- [11] KUMAR A, SUNKARIA R K. Design of prototype filter using windowing and linear optimization technique for the non-uniform filter banks [C] // 2013 International Conference on Recent Trends in Information Technology. Piscataway, NJ: IEEE

- Press, 2014:352-358.
- [12] SUMEDH D, SMITHA K G, VINOD A P. A low complexity reconfigurable channel filter based on decimation, interpolation and frequency response masking[C]//2013 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway, NJ: IEEE Press, 2013:5583-5587.
- [13] 张文旭, 陈亚静, 陈涛, 等. 基于 FRM 的偶型排列信道化改进结构[J]. 北京邮电大学学报, 2016, 39(3):27-33.
ZHANG W X, CHEN Y J, CHEN T, et al. Improved structure of even-arrangement channelized based on FRM [J]. Journal of Beijing University of Posts and Telecommunications, 2016, 39(3):27-33 (in Chinese).
- [14] TSUI K M, CHAN S C, LIM Y C. Design of multiplet perfect reconstruction filter banks using frequency response masking technique[J]. IEEE Transactions on Circuits and Systems I: Regular Papers, 2008, 55(9):2707-2715.
- [15] BINDIYA T S, ELIAS E. Metaheuristic algorithms for the design of multiplier-less non-uniform filter banks based on frequency response masking [J]. Soft Computing, 2014, 18(8):1529-1547.
- [16] DHABU S, VINOD A P. Design and FPGA implementation of reconfigurable linear phase digital filter with wide cutoff frequency range and narrow transition bandwidth[J]. IEEE Transactions on Circuits and Systems II: Express Briefs, 2016, 63(2):181-185.

作者简介:

张文旭 男, 博士, 讲师。主要研究方向: 宽带信号检测、处理与识别。

赵文童 男, 硕士研究生。主要研究方向: 数字信道化技术。

陈涛 男, 博士, 教授, 博士生导师。主要研究方向: 宽带信号检测、处理与识别。

陈亚静 女, 硕士。主要研究方向: 数字信道化技术。

An improved approach for odd-arrangement structure with narrow transition band based on FRM

ZHANG Wenxu*, ZHAO Wentong, CHEN Tao, CHEN Yajing

(College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China)

Abstract: In order to solve the problem of high computation complexity of narrow transition band filter bank, an odd-arrangement non-maximally decimated efficient structure with narrow transition band filter bank based on frequency response masking (FRM) is proposed. Firstly, when the prototype filter of this structure is designed, the FRM filter of the upper arm is obtained through the interpolated prototype filter. Secondly, the complementary filter of the lower arm is obtained by the delay signal subtracted from the interpolated prototype filter. Then the masking filter is used to shield the mirror caused by the interpolated filter. Finally, the desired FRM filter is obtained. The improved structure is simulated by MATLAB, and the simulation results verify the correctness of this improved structure. This efficient structure saves 80% of the multiplier resources compared with the polyphase filter bank architecture. Because the decimation module is located at the front of this improved structure, the limited condition of the sampling rate of the system is reduced and it can be directly used in the high-speed sampling system.

Keywords: narrow transition band; prototype filter; frequency response masking (FRM); filter bank; masking filter

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

DOI: 10.13700/j.bh.1001-5965.2017.0362

基于欧拉壁面液膜模型的三维热气防冰腔数值仿真



李延*, 郭涛, 常红亮

(中国航空工业集团公司 第一飞机设计研究院, 西安 710089)

摘 要: 提出了一种基于欧拉壁面液膜(EWF)模型的热气防冰腔性能仿真计算的新方法。通过 FLUENT 软件用户自定义标量(UDS)框架求解水滴控制方程获取三维表面水滴撞击特性。通过对各微元的水收集率、水膜蒸发率等进行质量平衡分析得到了通过该微元的质量流量,并以此作为 EWF 模型质量流量边界条件进行空气驱动下三维水膜厚度分布的计算,进而建立了防冰表面水膜流动动态模型。在此基础上建立了适用于三维防冰表面的耦合换热模型,通过引入亚松驰因子实现了内外流场、水膜流动及蒙皮导热的松散耦合求解。通过对某发动机短舱模型三维算例计算结果的分析 and 对比,结果表明所采用的计算方法是合理可信的,可以用于三维防冰腔性能的计算。

关键词: 欧拉壁面液膜(EWF)模型; 热气防冰; 松散耦合; 防冰热载荷; 水膜; 数值仿真; 三维

中图分类号: V244.1⁺.5

文献标识码: A **文章编号:** 1001-5965(2018)05-0959-08

飞机结冰会对飞行安全产生严重影响,对于亚声速飞机,由于气动加热效果不明显,因此一般都会重要迎风部件安装有防除冰系统。现有的防除冰系统中,热气防冰系统是使用最为广泛,系统可靠性最高的系统^[1]。系统所用热气引自发动机压气机,高温热气通过分配管路进入各防护区域,加热蒙皮,以将撞击到蒙皮表面的过冷水蒸发,从而达到防冰的效果。

评估热气防冰系统性能的有效方法包括冰风洞试验、自然结冰试验和数值仿真计算,其中冰风洞试验耗资较大,且冰风洞所能达到的环境条件是有限的,不能完全满足系统设计的需求;而自然结冰试验属于危险科目,必须在系统经过了较为充分的验证后方可进行,因此在设计阶段不能通过自然结冰试验实现对系统性能的验证。在热气防冰系统的设计和试验验证中,数值仿真方法可发挥重要

的作用,系统参数的确定、系统的性能仿真以及系统设计优化等均可借助数值仿真的方法来进行。

在众多参数中,防冰表面温度是评价防冰系统性能的最重要指标。而防冰表面温度的计算过程复杂,涉及热气防冰腔内外流场的流固耦合计算。其中,防冰表面流动更是涉及气、固、液三相复杂传热流动计算,建立合理准确的数学模型是计算结果准确性的关键。目前国内外关于防冰表面水膜的流动模型大都建立在 Messinger 结冰热力学模型的基础上^[2-4],国外 FENSAP-ICE 软件^[5-6]基于计算得到的空气-壁面剪切力,建立水膜流动模型,并通过内外流场的松散耦合计算得到了防冰表面温度,并可用于三维计算;李延和郭涛^[7]将等直机翼防冰腔部流动简化为二维流动,防冰腔内部仍为三维流动,通过内外的松散耦合计算得到了防冰表面温度信息。卜雪琴、侯盼雪

收稿日期: 2017-05-31; 录用日期: 2017-08-11; 网络出版时间: 2017-10-13 14:05

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20171013.1405.005.html

* 通信作者: E-mail: 276637284@qq.com

引用格式: 李延, 郭涛, 常红亮. 基于欧拉壁面液膜模型的三维热气防冰腔数值仿真[J]. 北京航空航天大学学报, 2018, 44(5): 959-966. LI Y, GUO T, CHANG H L. Numerical simulation of 3D hot-air anti-icing chamber based on Eulerian wall film model [J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 959-966 (in Chinese).

等^[8-9]基于 FLUENT 的用户自定义函数(UDF)功能建立了等直机翼和后掠翼的防冰表面温度计算的强耦合模型;郁嘉等^[10]采用类似 FENSAP-ICE 软件的方法,求解空气流场获取防冰表面剪切力,并计算剪切力驱动下的水膜流动,从而进行内外流场的耦合。

本文提出了一种计算防冰表面温度的新方法,基于 FLUENT 软件的欧拉壁面液膜(EWF)模型^[11]对水膜的动态流动过程进行模拟,通过求解空气驱动下的水滴流动场得到水滴收集特性,二者结合建立防冰表面热力学模型,进而采用松散耦合的方式进行内外流场的耦合计算,最终计算得到防冰表面温度分布和水膜厚度分布等防冰性能的重要评价指标。

1 数学模型

1.1 内外空气流场及蒙皮导热计算

空气流场的计算通过求解 Navier-Stokes 方程进行,湍流模型选用 S-A 模型,用于内外流场的初始化计算。在外流场计算中,模型壁面设置为温度边界条件;内流场和蒙皮导热同时计算,其中蒙皮表面边界条件为对流边界条件。

1.2 水滴撞击特性的计算

水滴收集特性通过求解空气驱动下的水滴场实现,控制方程为^[12-15]

$$\frac{\partial(\rho\alpha)}{\partial t} + \nabla \cdot (\rho\alpha\mathbf{u}) = 0 \quad (1)$$

$$\frac{\partial(\rho\alpha\mathbf{u})}{\partial t} + \nabla \cdot (\rho\alpha\mathbf{u}\mathbf{u}) = \rho\alpha K(\mathbf{u}_a - \mathbf{u}) + \rho\alpha\mathbf{F} \quad (2)$$

式中: t 为时间; ρ 为水滴密度; α 为水滴体积分数; \mathbf{u} 为水滴速度矢量; \mathbf{u}_a 为空气速度矢量; \mathbf{F} 为作用在水滴上除阻力以外的外力; K 为空气和水滴的动量交换系数,其表达式为

$$K = \frac{18\mu_a f}{\rho d^2} \quad (3)$$

其中: μ_a 为空气的动力黏度; d 为水滴直径; f 为水滴的阻力系数,可按球形且不变形水滴的阻力系数计算,表达式为

$$f = 1 + 0.15Re_r^{0.687} + 0.0175(1 + 4.25 \times 10^4 Re_r^{-1.16})^{-1} \quad (4)$$

式中: Re_r 为空气和水滴的相对雷诺数,可用式(5)计算:

$$Re_r = \frac{\rho_a d |\mathbf{u}_a - \mathbf{u}|}{\mu_a} \quad (5)$$

其中: ρ_a 为空气密度。

在水滴撞击特性的计算中,忽略方程中的非稳态项,方程的求解通过 FLUENT 的用户自定义

标量(UDS)功能实现方程的求解,通过在连续性方程右侧添加数值扩散项消除局部计算结果的奇异现象,添加的数值扩散项为^[15]

$$\Gamma = a(e^b - 1.0) \quad (6)$$

式中: a 根据经验取为 0.001^[15]; b 为当前网格与其周围相邻网格水滴体积分数的偏离程度,

$$b = \sqrt{\frac{1}{k} \sum_{i=1}^k (\alpha_p - \alpha_{Ni})^2} \quad (7)$$

式中: k 为相邻网格的个数;下标 P 表示当前网格;下标 Ni 表示当前网格相邻的网格。

求解增加了数值扩散相 Γ 的式(1)和式(2)即可得到水滴体积分数在计算域内的分布,进而利用式(8)计算得到防冰表面的局部收集系数为

$$\beta = \frac{\alpha(\mathbf{u} \cdot \mathbf{n})}{V_\infty} \quad (8)$$

式中: \mathbf{n} 为表面微元的外法向量; V_∞ 为空气的来流速度。

在进行防冰腔水膜流动计算时,得到局部收集系数后,进而可计算得到各微元的水收集率 \dot{m}_{imp} ,作为 EWF 模型边界条件的一项:

$$\dot{m}_{imp} = LWC \cdot V_\infty \beta ds \quad (9)$$

式中:LWC 为液态水含量; ds 为微元的底面积。

1.3 EWF 模型及防冰表面热载荷的计算

1.3.1 EWF 模型

如图 1 所示,云层中的过冷水滴会撞击到防冰表面,而在表面形成水膜,如果水滴较大,则会发生水滴的破碎、飞溅等现象,水膜在流动过程中受壁面剪切力、空气压力以及表面张力和自身黏性的影响,在特定的条件下,水膜的流动会由于表面张力的作用而发生分离,从而形成细流。本文所使用的 EWF 模型可以计算以上各种条件下的水膜流动,而对于本文所研究的防冰表面性能耦合计算的问题,暂不考虑过冷大水滴问题才会涉及的水滴的破碎、飞溅^[16],同时假设水膜流动是连续的,不考虑水膜形成细流的问题。

对于防冰表面的水膜流动而言,水膜厚度相对蒙皮表面的曲率半径而言是一个微小量,沿水膜厚度方向的物理属性可认为不发生变化,同时

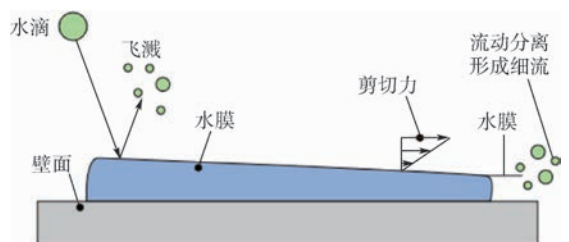


图 1 防冰表面水的流动

Fig. 1 Water flow on anti-icing surface

水膜的流动可认为平行于蒙皮表面,而无垂直方向的流动。在此基础上建立水膜流动的质量、动量、能量方程^[11]:

$$\frac{\partial h}{\partial t} + \nabla_s \cdot [h\mathbf{V}_1] = \frac{\dot{m}_s}{\rho_1} \quad (10)$$

$$\frac{\partial h\mathbf{V}_1}{\partial t} + \nabla_s \cdot (h\mathbf{V}_1\mathbf{V}_1) = -\frac{h\nabla_s P_L}{\rho_1} + g_\tau h + \frac{3}{2\rho_1}\tau_{fs} - \frac{3\nu_1}{h}\mathbf{V}_1 + \frac{\mathbf{q}}{\rho_1} \quad (11)$$

$$\frac{\partial (hT_f)}{\partial t} + \nabla_s \cdot (\mathbf{V}_1 h T_f) = \frac{1}{\rho c_p} \left[k_f \left(\frac{T_s - T_f}{h/2} - \frac{T_f - T_w}{h/2} \right) + \dot{q}_{imp} + \dot{m}_{evap} L(T_s) \right] \quad (12)$$

式中: h 为水膜厚度; ∇_s 为表面梯度算子; \mathbf{V}_1 水膜速度矢量; ρ_1 为水膜密度; \dot{m}_s 为单位面积的液体质量流量; g_τ 为重力在平行于壁面的分量; τ_{fs} 为气液交界面的剪切应力; ν_1 为水膜的运动黏度; \mathbf{q} 为水滴撞击或水膜分离引起的表面压强变化; T_f 为水膜沿厚度方向的平均温度; \mathbf{V}_1 为水膜平均速度; c_p 为水的定压比热容; T_s 为气液界面处的水膜温度; T_w 为壁面处的水膜温度; k_f 为水膜的导热系数; \dot{q}_{imp} 为撞击水携带的能量; \dot{m}_{evap} 为水膜蒸发率; L 为水膜的蒸发潜热; P_L 为空气在壁面的压力,重力沿垂直于壁面的分量以及表面张力的合力。

容易看出,欧拉水膜模型可同时计算在壁面压力梯度、重力、表面张力、剪切力、液体黏性力以及外界其他因素作用下的水膜流动,相比文献[7-9]中的算法,考虑的因素更加全面,更加接近实际物理过程。

对于防冰表面的水膜流动而言,可忽略重力、黏性力、液滴撞击等因素的影响,由此动量方程可简化为

$$\frac{\partial h\mathbf{V}_1}{\partial t} + \nabla_s \cdot (h\mathbf{V}_1\mathbf{V}_1) = -\frac{h\nabla_s P_L}{\rho_1} + \frac{3}{2\rho_1}\tau_{fs} \quad (13)$$

在本文的计算模型中,由于水膜的温度同时由防冰腔内部和外部的换热决定,因此欧拉水膜的能量方程不能直接用于系统的耦合计算,需要根据防冰表面的换热分析建立内外流场的耦合换热模型,具体见1.4节。

1.3.2 防冰表面热载荷的计算

根据经典的 Messinger 结冰热力学模型^[17],防冰表面热流 Q_{anti} 包括对流换热热流 Q_{conv} 、水滴撞击热流 Q_{cap} 、水蒸发吸收的热量 Q_{evap} 以及从微元中流入和流出的水携带的能量,其中流入流出微元的液态水携带的能量差值取决于不同微元间的温度差,一般而言,相邻微元间的温度差别并不明显,

因此相比较于前4项热流,流入流出控制体的能量变化可以忽略,因此防冰表面的热载荷可表示为

$$Q_{anti} = Q_{conv} + Q_{cap} + Q_{evap} \quad (14)$$

对流换热热流的计算可直接通过求解 Navier-Stokes 方程的能量方程获得,水滴撞击热流包括水滴的动能和加热水滴至水膜温度的能量:

$$Q_{cap} = \dot{m}_{imp} \left[\frac{1}{2} V_\infty^2 + c_p (T_\infty - T_w) \right] \quad (15)$$

式中: T_∞ 为来流温度。

水的蒸发吸收的热量取决于水膜蒸发率 \dot{m}_{evap} 和水的蒸发潜热,水膜蒸发率可用式(16)计算:

$$\dot{m}_{evap} = 0.694 \frac{h_{conv}}{c_{pa}} \left[\frac{e(T_w)}{P_e} \cdot \frac{T_e}{T_w} - \frac{e(T_\infty)}{P_\infty} \right] \cdot ds \quad (16)$$

式中: c_{pa} 为空气的定压比热容; P_∞ 为环境静压; T_e 为附面层厚度处的空气温度; P_e 为附面层厚度处的静压; e 为饱和水蒸汽压; h_{conv} 为表面对流换热系数,可根据求解 Navier-Stokes 方程的能量方程获得的表面对流换热热流计算得到:

$$h_{conv} = Q_{conv} / (T_w - T_{rec}) \quad (17)$$

其中: T_{rec} 为附面层恢复温度。

$e(T)$ 的计算式为^[18]

$$e(T) = 6894.7 \times \exp \left(20.15247167 - \frac{11097.16963}{1.8T} \right) \quad (18)$$

式中: T 为温度。

计算得到防冰表面热流后,采用等效对流换热系数 h_{eff} 作为内流场和蒙皮导热的边界条件:

$$h_{eff} = Q_{anti} / (T_w - T_{ref}) \quad (19)$$

式中: T_{ref} 为参考温度,可设置为环境温度,此处参考温度应与内流场对流边界条件设置的参考温度一致。

1.4 耦合换热模型的建立

1.4.1 水膜流动过程中的热载荷计算

本文水膜流动的质量和动量守恒方程的求解借助 FLUENT 软件的 EWF 模型进行,将水滴收集率、蒸发率作为水膜流动的边界条件,并设置合理的时间步长,捕捉水膜在不同时间的厚度分布,同时通过编写 UDF 在每个时间步长更新水膜流动的边界条件,并以 CFL 数的变化量小于设定值作为计算收敛的标志,其中 CFL 数的定义为

$$CFL = \frac{u_i \Delta t}{\Delta x} + \frac{v_i \Delta t}{\Delta y} + \frac{w_i \Delta t}{\Delta z} \quad (20)$$

式中: u_i 、 v_i 、 w_i 分别为水膜的在 x 、 y 、 z 方向的流速; Δt 为时间步长; Δx 、 Δy 、 Δz 分别为相邻微元中

心点距离在 x, y, z 方向的分量。

在每个时间步长内,对于每个微元进行如下计算:

1) 获取当前微元的水膜厚度和微元面积,计算得到当前微元上水膜的质量为

$$M_f = h \cdot ds \quad (21)$$

2) 根据水滴收集特性计算结果,计算当前时间步长 Δt 内微元收集的水量为

$$M_{imp} = LWC \cdot V_\infty \beta \cdot ds \cdot \Delta t \quad (22)$$

3) 计算当前微元在时间步长内的理论蒸发量 M_{evap} 。

4) 计算微元的质量源流量:

$$M_{source} = M_{imp} - M_{evap} \quad (23)$$

5) 若 $M_{source} < 0$,则进一步比较 $|M_{source}|$ 和 M_f 的大小:

$$M_{source} = \begin{cases} M_{source} & |M_{source}| < M_f \\ -M_f & |M_{source}| > M_f \end{cases} \quad (24)$$

6) 计算得到质量源流量后,进一步根据 1.3.2 节的方法计算每个微元的等效对流换热系数,并设置用户自定义存储(UDM)将等效对流换热系数存储在计算文件中。

7) 计算 CFL 数,并将其与前一步长的 CFL 数进行对比,当残差小于设定值时,即可认为水膜流动达到了稳定状态,以此时计算得到的等效对流换热系数作为防冰表面热载荷的计算结果,将其作为内流场和蒙皮导热计算的边界条件。

1.4.2 界面数据的交换

为便于工程计算,内外流场交界面处的网格节点并非一一对应,因此进行内外流场的数据交换时,需采用合理的插值方法进行,本文借助 FLUENT 软件的 profiles 功能,可选用 2 种不同的插值方法,分别为最小距离法和反距离加权插值^[19],二者的基本原理如下:最小距离法是将已知网格中与待定节点距离最小的节点数据作为待定节点的数据;反距离加权插值是将待定节点与已知节点间距离乘方的倒数作为权值,计算待定节点的数据信息。

本文在计算时,采用反距离加权插值方法。

1.4.3 耦合收敛的控制

对于内外流场的耦合,若直接将计算得到的蒙皮温度分布或等效对流换热系数分布作为内

外侧流场计算的边界条件,会导致数值的波动而无法达到收敛,因此在实际计算时引入亚松弛因子。

在本文的计算中,仅对等效对流换热系数引入亚松弛因子 ω :

$$h_{eff} = h_{eff}^{i-1} + (h_{eff}^i - h_{eff}^{i-1}) \cdot \omega \quad (25)$$

式中: h_{eff}^i 和 h_{eff}^{i-1} 分别为当前和前一时间步的等效对流换热系数。

2 算例分析

2.1 计算模型和计算状态

为验证本文计算方法的准确性,计算了某民机发动机短舱防冰系统的防冰性能。其中防冰腔采用旋流喷嘴的形式,热空气从喷嘴中喷出,在防冰腔的环形通道中循环流动,充分换热后从防冰腔后壁板上的开孔排出机外。根据几何模型分别对外流场和内流场(含蒙皮)划分网格,并计算流场作为耦合计算的初值。

短舱几何外形和防冰腔分别见图 2 和图 3。

计算采用的外部环境和供气参数见表 1。

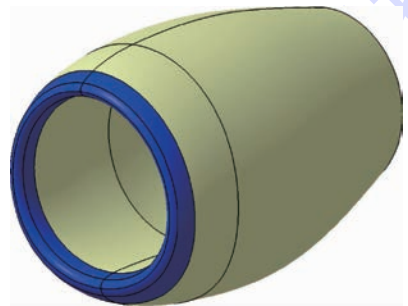


图 2 短舱几何外形及其防护区域

Fig. 2 Nacelle's geometric shape and protection area

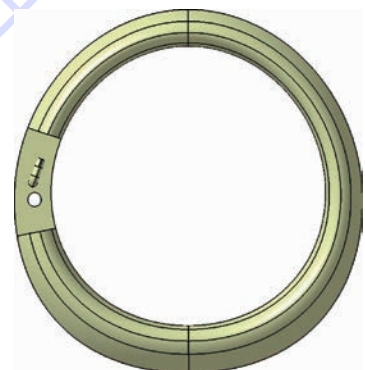


图 3 短舱防冰腔

Fig. 3 Nacelle anti-icing chamber

表 1 计算条件

Table 1 Calculation conditions

参数	马赫数	环境压力/Pa	环境温度/℃	水滴直径/ μm	液态水含量/ $(\text{g} \cdot \text{m}^{-3})$	供气压力/kPa	供气温度/℃
数值	0.3	70 124	-9.4	20	0.5	350	150

2.2 水滴撞击特性

在空气流场的基础上,采用 1.2 节中方法求解水滴流动的 UDS 方程,其中离散格式采用二阶迎风格式以保证足够的计算精度,将局部收集系数存储在 FLUENT 软件 UDM 中,作为水膜流动的边界条件之一。求解水滴流场得到的局部收集系数结果和在相同条件下用 FENSAP-ICE 软件计算得到的局部水滴收集系数计算结果见图 4。

从二者的对比结果可以明显看出,本文与 FENSAP-ICE 软件的计算结果基本相同,因此本文采用的基于 FLUENT 软件 UDS 的水滴场求解方法是正确的,此方法可以用于任意三维部件水滴撞击特性的计算。

2.3 水膜流动的动态过程

在计算水膜流动时,根据 1.4.1 节编写 UDF 文件,将求得的水质量流量源项加载至壁面(含防护区域和非防护区域),计算不同时间步内水膜流动的动态过程,其中内流场和蒙皮导热计算得到的蒙皮表面温度作为温度边界条件,时间步长为 0.03 s,图 5 为部分时刻水膜厚度的计算结果。

从不同时刻的水膜厚度分布可以看出如下规律,在初始阶段,由于水滴的撞击水膜厚度逐渐增长,但由于收集量较少,水膜仅分布在前缘水滴撞击区域;随着时间的累积,收集量逐渐增大,

但撞击区域的加热热流并不能使收集到的水蒸发,因此水膜沿壁面向后溢流,由于撞击区域后

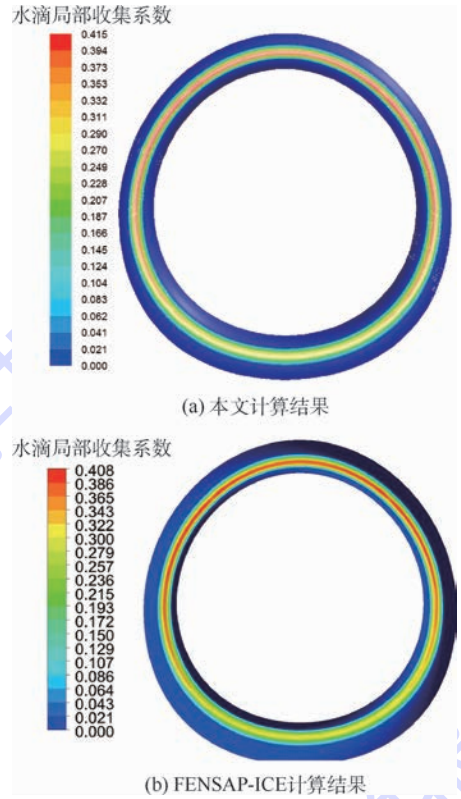


图 4 水滴局部收集系数计算结果的对比
Fig. 4 Comparison of water droplet local collection coefficient

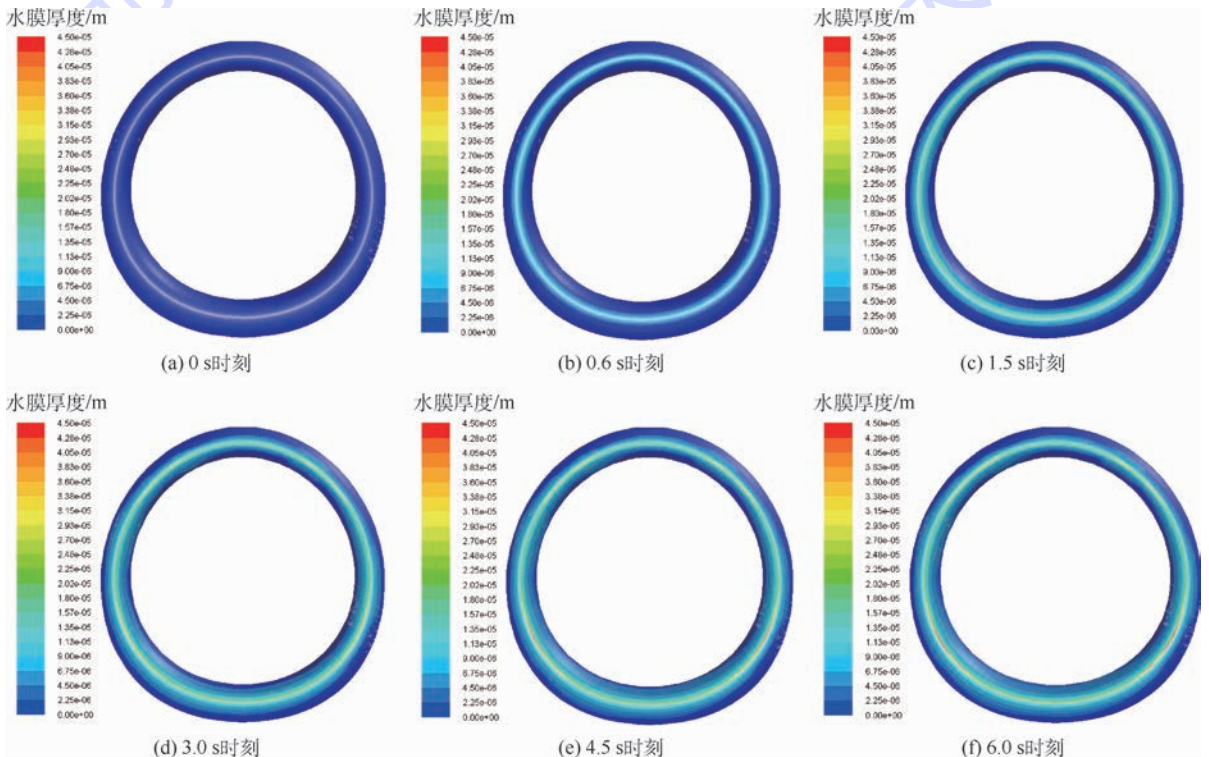


图 5 不同时刻的水膜厚度分布
Fig. 5 Water film thickness distribution at different time

不再有水滴的撞击,但加热热流仍会使流入该区域的水膜蒸发;最终在一定时间后,整个防护区域的水膜厚度在水滴撞击、加热蒸发、水膜流动等因素的作用下趋于稳定。

此外,若加热热流不足以使所收集的液态水在防护区域内蒸发,则水膜会流出防护区域外,在此区域没有加热热流,因此水膜会迅速冻结,形成溢流冰。在本文的计算条件下,水膜有极少量流出了防护区域,如图6所示,因此会在形成溢流冰,关于溢流冰外形及厚度的计算,将在后续研究工作中开展。

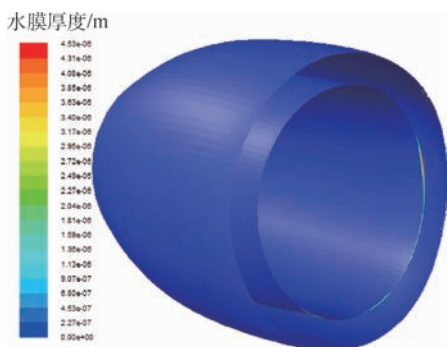


图6 防护区域外的水膜厚度分布

Fig.6 Water film thickness distribution in non-protection area

2.4 防冰表面温度及热载荷

防冰表面温度是评价热气防冰系统性能的重要参数,本文通过内外流场、水膜流动及防冰

热载荷等参数的耦合计算,得到了收敛后的短舱防护区域蒙皮表面温度分布,同时用 FENSAP-ICE 软件计算了相同模型下的防冰表面温度,二者的对比见图7,同时沿周向截取了4个截面进行温度分布的对比,见图8。

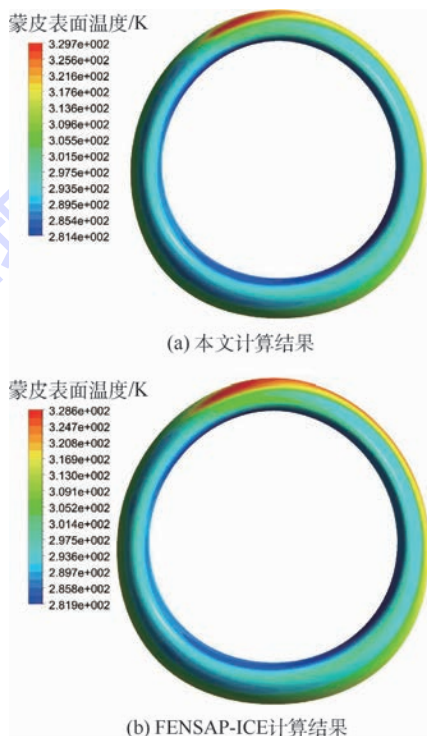


图7 加热区域蒙皮表面温度分布对比
Fig.7 Comparison of skin temperature distribution of heated area

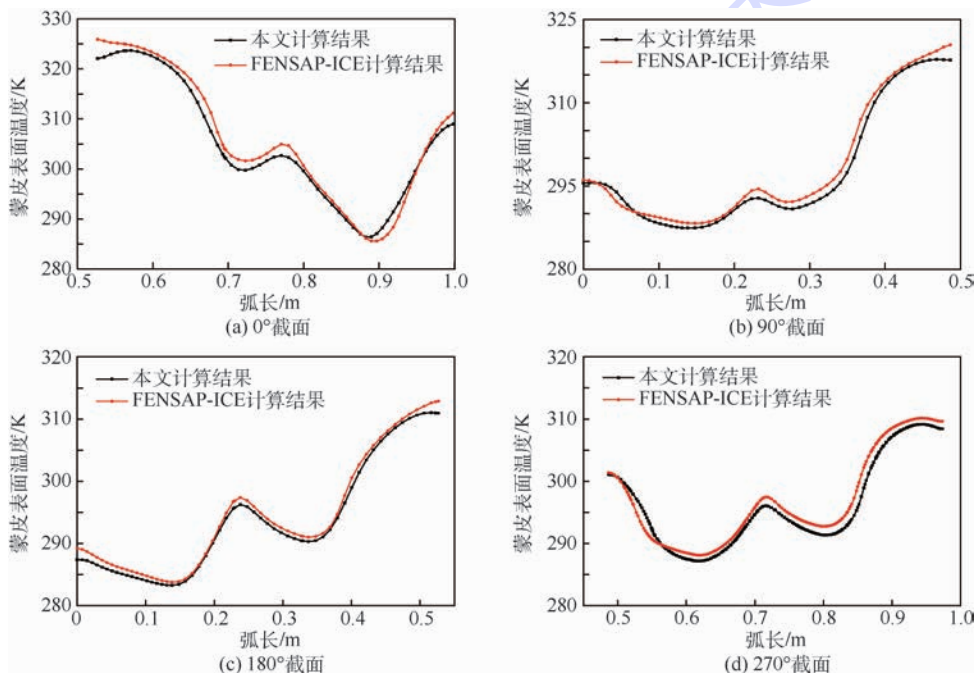


图8 不同截面蒙皮表面温度分布对比

Fig.8 Comparison of skin temperature distribution on different sections

从蒙皮表面温度计算结果可以看出,计算收敛后的最高温度为329.7 K,最低温度为281.4 K,最低温度大于冰点,表明在整个防护区域内未出现结冰现象。而最高温度出现在外侧上部区域,这是由于气流喷射引起的局部高温;整个防护区域内低温出现的位置为远离前缘的区域,在此区域存在气流加速引起的对流换热系数的增加,从而导致散热量增大而造成此区域温度较低。

从 FENSAP-ICE 软件的计算结果与本文计算结果的对比可以明显看出二者具有较高的一致性,可以验证本文计算方法的正确性。

计算收敛后的等效对流换热系数分布见图9。可以看出,在前缘驻点区域,由于流动为层流导致对流换热系数较小,进而导致水的蒸发量很小。由于对流换热和蒸发散热是防冰热载荷的决定因素,因此在前缘驻点区域,等效对流换热系数很小。而在驻点往后的区域,等效对流换热系数受空气对流换热、蒸发散热、水滴撞击携带热量以及水膜流动换热的综合作用。此外由于水膜溢流区域以外无水膜的蒸发散热,导致此区域等效对流换热系数迅速减小。

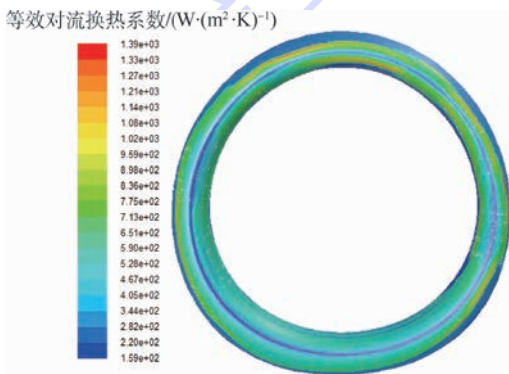


图9 等效对流换热系数分布

Fig.9 Distribution of equivalent convective heat transfer coefficient

3 结论

本文基于 FLUENT 软件 EWF 模型提出了一种热气防冰腔性能仿真计算的新方法,通过对算例的对比和分析表明:

1) 基于 UDS 的水滴撞击特性计算方法的计算结果与 FENSAP-ICE 软件计算结果基本相同,可满足三维表面水滴撞击特性的计算。

2) 提出了基于 EWF 模型结合 UDF 自定义质量流量边界条件求解水膜在防冰表面的动态流动过程方法,实现了三维水膜的流动及传热计算,结果真实可信。

3) 建立了适应于三维外形的防冰性能换热计算模型,并通过算例计算得到了某民机发动机短舱防冰系统在典型外部环境和供气参数下的三维防冰表面温度、防冰表面水膜厚度分布以及表征防冰热载荷的等效对流换热系数。通过对算例计算结果的分析 and 对比可知,本文提出的基于 EWF 模型的热气防冰腔性能仿真计算方法是切实有效的,其计算结果也是合理可信的。

后续将在本文研究内容的基础上开展计算结果的试验验证以及溢流冰仿真研究工作。

参考文献 (References)

- [1] 林贵平,卜雪琴,申晓斌,等.飞机结冰与防冰技术[M].北京:北京航空航天大学出版社,2016:30-46.
LIN G P, BU X Q, SHEN X B, et al. Aircraft icing and anti-icing technology [M]. Beijing: Beihang University Press, 2016: 30-46 (in Chinese).
- [2] DONG W, ZHU J, MIN X H. Calculation of the heat transfer and temperature on the aircraft anti-icing surface [C] // 27th International Congress of the Aeronautical Sciences, 2010: 1-9.
- [3] DOMINGOS R H, PAPANAKIS M, ZAMORA A O. Computational methodology for bleed air ice protection system parametric analysis: AIAA-2010-7834 [R]. Reston: AIAA, 2010.
- [4] MORENCY F, BRAHIMI M T, TEZOK F, et al. Anti-icing system simulation using CANICE [J]. Journal of Aircraft, 1999, 36 (6): 999-1006.
- [5] PELLISSIER M, HABASHI W G, PUEYO A. Design optimization of hot-air anti-icing systems by FENSAP-ICE: AIAA-2010-1238 [R]. Reston: AIAA, 2010.
- [6] WANG H Z, TRAN P, HABASHI W G. Anti-icing simulation in wet air of a piccolo system using FE NSAP-ICE: SAE 2007-01-3357 [R]. Warrendale: SAE International, 2007.
- [7] 李延,郭涛.基于松散耦合的三维热气防冰腔数值仿真 [C] // 第六届中国航空学会青年科技论坛文集.北京:航空工业出版社,2014:1245-1250.
LI Y, GUO T. Numerical simulation of three dimensional thermal anti-icing chamber by loose couple method [C] // CSAA the sixth Youth Science and Technology Forum. Beijing: Aviation Industry Press, 2014: 1245-1250 (in Chinese).
- [8] 卜雪琴,林贵平,郁嘉.三维内外热耦合计算热气防冰系统表面温度 [J]. 航空动力学报, 2009, 24 (11): 2495-2500.
BU X Q, LIN G P, YU J. Three dimensional conjugate heat transfer for the surface temperature of wing hot-air anti-icing system [J]. Journal of Aerospace Power, 2009, 24 (11): 2495-2500 (in Chinese).
- [9] 侯盼雪,林贵平,卜雪琴,等.后掠翼热气防冰系统数值仿真 [J]. 航空学报, 2012, 33 (5): 809-817.
HOU P X, LIN G P, BU X Q, et al. Numerical simulation of a swept wing hot-air anti-icing system [J]. Acta Aeronautica et Astronautica Sinica, 2012, 33 (5): 809-817 (in Chinese).
- [10] 郁嘉,赵柏阳,卜雪琴,等.某型飞机发动机短舱热气防冰系统性能数值模拟 [J]. 空气动力学学报, 2016, 34 (3):

302-307.

YU J, ZHAO B Y, BU X Q, et al. Numerical simulation of the performance of an engine nacelle hot-air anti-icing system[J]. Acta Aerodynamica Sinica, 2016, 34(3): 302-307 (in Chinese).

[11] Ansys Inc. ANSYS FLUENT 14 User Guide [Z]. New Hampshire: ANSYS Inc., 2011.

[12] 毛晓东, 卜雪琴, 赵国昌, 等. 基于 UDS 框架的水撞击特性数值计算方法[J]. 沈阳航空航天大学学报, 2016, 33(1): 8-12.

MAO X D, BU X Q, ZHAO G C, et al. Calculation of water droplet impingement characteristics based on UDS frame[J]. Journal of Shenyang Aerospace University, 2016, 33(1): 8-12 (in Chinese).

[13] WIROGO S, SRIRAMBHATLA S. An eulerian method to calculate the collect efficiency on two and three dimensional bodies; AIAA-2003-1073[R]. Reston: AIAA, 2003.

[14] TONG X L, LUKE E A. Eulerian simulations of icing collection efficiency using a singularity diffusion model; AIAA-2005-1246[R]. Reston: AIAA, 2005.

[15] 杨胜华, 林贵平, 申晓斌. 三维复杂表面水滴撞击特性计算[J]. 航空动力学报, 2010, 25(2): 284-290.

YANG S H, LIN G P, SHEN X B. Water droplet impingement prediction for three dimensional complex surfaces[J]. Journal of Aerospace Power, 2010, 25(2): 284-290 (in Chinese).

[16] 王超, 常士楠, 吴梦龙, 等. 过冷大水滴飞溅特性数值分析

[J]. 航空学报, 2014, 35(4): 1004-1011.

WANG C, CHANG S N, WU M L, et al. Numerical investigation of splashing characteristics in super-cooled large droplet regime [J]. Acta Aeronautica et Astronautica Sinica, 2014, 35(4): 1004-1011 (in Chinese).

[17] MESSINGER B L. Equilibrium temperature of an unheated icing surface as a function of air speed[J]. Journal of the Aeronautical Sciences, 1953, 20(1): 29-42.

[18] VELAZQUEZ M T, HANSMAN R J. Implementation of combined feather and surface-normal ice growth models in LEWICE/X; AIAA-1995-0735[R]. Reston: AIAA, 1995.

[19] 陈鹏, 邓飞, 刘思亭. 三维空间属性插值方法的研究[J]. 电脑知识与技术, 2015, 11(7): 235-239.

CHEN P, DENG F, LIU S T. Research of three dimensional space attribute interpolation methods[J]. Computer Knowledge and Technology, 2015, 11(7): 235-239 (in Chinese).

作者简介:

李延 男, 硕士, 工程师。主要研究方向: 飞机防除冰系统设计与仿真计算。

郭涛 男, 博士, 高级工程师。主要研究方向: 飞机防除冰系统设计。

常红亮 男, 学士, 研究员。主要研究方向: 飞机防除冰系统设计。

Numerical simulation of 3D hot-air anti-icing chamber based on Eulerian wall film model

LI Yan*, GUO Tao, CHANG Hongliang

(The First Aircraft Institute, Aviation Industry Corporation of China, Xi'an 710089, China)

Abstract: A new computation method of hot-air anti-icing chamber performance based on Eulerian wall film (EWF) model was presented in this paper. User defined scalar (UDS) of FLUENT software was used to calculate water droplet impingement efficiency by solving droplet governing equation. Mass balance analysis of water collection rate and film evaporation rate was performed to get the mass flow rate of each micro unit, which was used as boundary condition in calculating 3D film thickness distribution driven by air, and then the dynamic model of film flow on anti-icing surface was set up. Based on the above work, conjugate heat transfer model for 3D anti-icing surface was built, and under-relaxation factor was used in solving the loose coupling of inner/outer flow field, water film flow and skin heat conduction. The method of this paper was used in the calculation of a nacelle anti-icing chamber performance, and the results show good compliance with the physical phenomenon. The method in this paper can be used in 3D hot-air anti-icing chamber performance calculation.

Keywords: Eulerian wall film (EWF) model; hot-air anti-icing; loose couple; anti-icing heat load; water film; numerical simulation; three-dimensional

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

DOI: 10.13700/j.bh.1001-5965.2017.0345

随机和区间变量共存条件下的模型确认指标

赵录峰¹, 吕震宙^{1,*}, 阚丽娟²

(1. 西北工业大学 航空学院, 西安 710072; 2. 空军工程大学 装备管理与安全工程学院, 西安 710051)



摘 要: 现有的不确定性模型确认方法建立在概率理论基础之上, 仅仅适用于随机不确定性因素影响下的模型确认, 而不适合随机和区间变量共存条件下的模型确认问题。针对这一问题, 研究了随机和区间变量共存条件下的模型确认方法。首先, 分析了随机和区间变量共存条件下数学模型的特点; 然后, 运用概率方法和区间理论, 提出了一种新的模型确认指标, 通过模型响应量的上下界分布函数(CDF)与实验响应量的上下界经验 CDF 之间差异, 来度量随机和区间输入变量共存条件下模型预测与实际物理实验结果之间的一致性; 讨论了所提指标的数学性质, 给出了指标的计算方法和步骤; 最后, 采用一个数字算例和一个工程算例验证了所提指标在随机和区间输入变量共存条件下进行模型确认的可行性和有效性。

关键词: 模型确认; 指标; 随机变量; 区间变量; 混合模型

中图分类号: O212.1; TP391.9

文献标识码: A **文章编号:** 1001-5965(2018)05-0967-08

在航空、航天、机械、土木工程和武器系统的研制过程中, 为了节约实验费用, 提高工作效率, 计算机仿真已经成为产品设计工作的一种重要手段。由于描述真实物理世界的数学模型是计算机仿真的核心, 它的准确性直接决定了产品的设计与验证质量, 因此模型确认工作在工程设计中发挥着越来越重要的作用。模型确认是指依据模型用途对该模型在描述真实物理世界时的准确程度进行客观评估的过程^[1-3]。开展模型确认工作, 首先需要建立一套科学合理的模型确认指标, 来对模型的准确性进行量化评估, 它是模型确认工作的基础和依据。模型确认指标是指用于度量模型输出响应量与对应的实验结果之间差异程度的方法^[4-5]。

在实际工程研制过程中, 模型确认往往会受到各种不确定性因素的影响。文献[6-9]将模型确认过程中的不确定性主要归纳为4个方面: 模型参数不确定性、数学模型不确定性、实验结果不

确定性和实验数据有限而导致的信息不确定性等。根据不确定性的描述方法不同, 所建模型具有不同的形式。当模型的输入变量数据信息比较充分, 且它的不确定性采用概率的方法进行描述时, 这种模型称为随机不确定性模型。当数据信息量较少, 或受认知水平的限制, 模型输入变量不确定性难以采用概率方法描述时, 通常采用分布未知、边界可知的区间变量来描述, 所建的这种数学模型称为区间不确定性模型。然而, 在许多工程实际中, 所建模型往往既含有随机变量, 又含区间变量, 这类模型称为随机-区间混合不确定性模型。在随机-区间混合不确定性模型中, 随机输入变量和区间输入变量之间可能相互独立, 也可能存在相关性, 本文所研究的随机-区间混合不确定性模型就是指随机输入变量和区间输入变量共存且相互独立的单输出模型。

近几年, 学者对不确定情况下的模型确认方

收稿日期: 2017-05-22; 录用日期: 2017-06-30; 网络出版时间: 2017-08-04 16:24

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20170804.1624.008.html

基金项目: 国家自然科学基金(51475370); 中央高校基本科研业务费专项资金(3102015BJ(II)CG009)

* 通信作者. E-mail: zhenzhou@nwpu.edu.cn

引用格式: 赵录峰, 吕震宙, 阚丽娟. 随机和区间变量共存条件下的模型确认指标[J]. 北京航空航天大学学报, 2018, 44(5): 967-974. ZHAO L F, LYU Z Z, KAN L J. A validation metric for model with mixture of random and interval variables[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 967-974 (in Chinese).

法进行了大量研究。文献[10]对现有方法进行了分类和对比分析,给出了各类方法在模型确认时的优点、局限性和适应范围。比如:假设检验和贝叶斯方法^[10-11]都是建立在古典假设检验的基础,是在一定置信水平下根据实验数据对原假设的正确与否进行判断,只能给出接受原假设,或者拒绝原假设的结论,而没有定量地度量模型预测与实验结果之间的差异程度。因此,这2类方法与模型确认的定义不完全一致。虽然根据模型确认的定义,文献[2]提出了频率指标法,但由于该方法仅考虑了模型预测与实验结果均值之间的差异,而没有考虑它们的离散性,因此仅适合响应量分布中心发生偏离情况下的模型确认。文献[2, 12-13]综合考虑了模型和实验的所有不确定性信息,根据模型输出响应量的分布函数曲线与实验结果的经验分布函数(Cumulative Distribution Function, CDF)曲线之间的面积差异,提出了基于直接面积法的模型确认指标和 u-pooling 方法,这2种方法虽然在模型确认方面具有很大的先进性,但它们不适合多维相关输出情况下的模型确认。为此,在这2个指标的基础上,文献[14]通过多输出响应量的数学期望列阵和协方差矩阵等数字特征,提出了多输出模型确认的局部混合矩指标和全局混合矩指标;文献[15]运用多维概率积分转换(Probability Integral Transformations, PIT),定义了 PIT 面积指标和 t-pooling 方法;文献[16]运用马氏距离(Mahalanobis Distance, MD),建立了基于 MD 的模型确认指标和 MD-pooling 方法;这些方法分别适合单一位置和多个位置的多维相关响应量模型问题。此外,文献[17]将核主成分分析与面积法的思想相结合,提出了基于核主成分分析的多输出模型确认方法。

已有方法虽然大大地推进了不确定性条件下模型确认技术的发展,但它们都是建立在概率描述不确定性的论理论框架之下,仅适合于随机不确定性影响下的模型确认,而不适合随机和区间变量共存条件下的模型确认问题。关于随机和区间变量共存条件下模型确认方面的文献,目前还非常少见。因此,本文以随机输入变量和区间输入变量共存且相互独立的单输出模型为研究对象,分析了随机-区间混合不确定性模型的特点,定义了一种新的随机和区间变量共存条件下模型确认指标。

1 混合不确定性模型及其特点

在工程实际中,随机输入变量和区间输入变

量经常出现在同一个数学模型中,定义这种随机-区间混合不确定性模型的表达式为

$$Z^m = g(\mathbf{X}, \mathbf{Y}) \quad (1)$$

式中: $\mathbf{X} = (X_1, X_2, \dots, X_{n_r})$ 为 n_r 维随机向量。 $f_x(\mathbf{x})$ 为随机向量 \mathbf{X} 的联合概率密度函数, $f_{x_i}(x)$ ($i=1, 2, \dots, n_r$)为第 i 维随机变量的边缘概率密度函数。 $\mathbf{Y} = (Y_1, Y_2, \dots, Y_{n_l})$ 为 n_l 维区间向量,它的取值规律由相应的区间 $[Y_j^l, Y_j^u]$ ($j=1, 2, \dots, n_l$)描述, Y_j^l 和 Y_j^u 分别为第 j 维区间变量的下上界。且 \mathbf{X} 和 \mathbf{Y} 相互独立。

当随机向量 \mathbf{X} 取任意一个实现值 \mathbf{x}^* 时, \mathbf{X} 的不确定性对输出响应量的影响将会被消除,则模型的输出响应量为区间变量。当随机向量 \mathbf{X} 按照它的联合概率密度 $f_x(\mathbf{x})$ 在其取值范围内取全部实现值时,与其对应的模型的输出响应量则为一个随机区间变量。图1为某随机-区间混合不确定性模型输出响应量的区间分布函数示意图。其中 z 为模型输出响应量, $F_z(z)$ 为模型输出响应量对应的分布函数。由于随机向量 \mathbf{X} 取实现值 \mathbf{x}^* 时所对应的模型输出响应量为区间变量,所以当随机向量 \mathbf{X} 按照它的联合概率密度 $f_x(\mathbf{x})$ 在其取值范围内取全部实现值时,模型输出响应量 $Z^m | \mathbf{X}$ 为随机区间变量,它的上界 $Z^{m,u} | \mathbf{X}$ 和下界 $Z^{m,l} | \mathbf{X}$ 都为随机变量。图1中的2条曲线分别为随机向量 \mathbf{X} 取所有实现值时,模型输出响应量 $Z^m | \mathbf{X}$ 的上下界分布函数 $F_{z^u | \mathbf{X}}^m(z^u)$ 和 $F_{z^l | \mathbf{X}}^m(z^l)$ 。

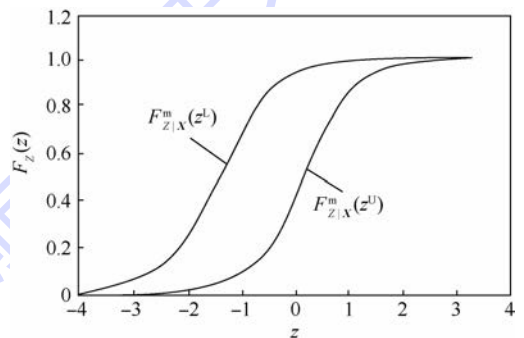


图1 模型响应量的上下界分布函数

Fig. 1 CDFs of upper and lower bounds of model responses

2 混合不确定性模型确认新指标

2.1 指标的构建

由第1节分析可知,当随机向量 \mathbf{X} 按照它的联合概率密度 $f_x(\mathbf{x})$ 在其取值范围内取全部实现值时,随机-区间混合不确定性模型输出响应量 $Z^m | \mathbf{X}$ 的上下界皆为随机变量,其分布函数分别为 $F_{z^u | \mathbf{X}}^m(z^u)$ 和 $F_{z^l | \mathbf{X}}^m(z^l)$ 。同理,可以得到相应的实验响应量 $Z^e | \mathbf{X}$ 的上下界经验分布函数分别为

$S_{z|x}^e(z^u)$ 和 $S_{z|x}^e(z^l)$, 如图 2 中阶梯线所示。

由图 2 可推断出, 当所建模型与物理过程完全一致时, 模型输出响应量的上下界分布函数与实验响应量的上下界经验分布函数相同, 即 2 条曲线与 2 条阶梯线之间阴影部分的面积应该趋于零。如果模型输出响应量的上下界分布函数与实验响应量的上下界经验分布函数之间的差异越大, 则可推断出所建模型与物理过程的不一致性越大, 即 2 条曲线与 2 条阶梯线之间阴影部分的面积越大, 则所建模型与物理过程之间差异越大, 所模型的准确性越差。因此, 通过 2 条曲线与 2 条阶梯线之间的阴影部分面积, 可以定量地评估所建模型的准确性程度。为此, 本文定义的随机-区间混合不确定性模型确认指标的表达式为

$$d(F^m, S^e) = \int_{-\infty}^{+\infty} (|F_{z|x}^m(z^u) - S_{z|x}^e(z^u)| + |F_{z|x}^m(z^l) - S_{z|x}^e(z^l)|) dz \quad (2)$$

对于本文所建指标, 还需要明确以下 3 点:

1) 该指标仅仅是用来度量随机-区间混合不确定性模型响应量与相应的实验结果之间差异程度, 适合于随机输入变量和区间输入变量共存且相互独立情况下的单输出模型确认问题。

2) 该指标既可用于量化评判不同备选模型的优劣, 又可用于模型预测能力评估, 该指标是制定模型确认要求(标准)的基础。

3) 模型确认要求的制定非常复杂, 需要结合实际情况依据模型确认指标具体制定, 不同领域不同问题的模型确认要求是不同的。模型确认指标取值模型可接受的问题, 既涉及到模型确认要求(标准), 又需要考虑模型确认指标的计算结果, 仅仅根据模型确认指标计算结果无法给出模型是否能够接受的结论。

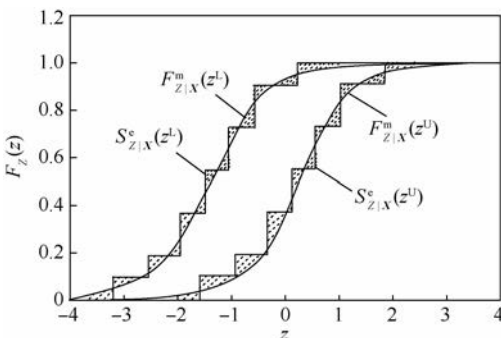


图 2 模型和实验响应量上下界分布函数和经验分布函数

Fig. 2 CDFs and empirical CDFs of upper and lower bounds of model and experimental responses

2.2 指标的数学性质

进一步分析可以得到, 新定义的模型确认指

标具有下列数学性质:

1) 非负性

由于式(2)的被积函数为非负函数, 因此模型确认指标 $d(F^m, S^e)$ 为非负数。

2) 对称性

由于

$$d(F^m, S^e) = \int_{-\infty}^{+\infty} (|F_{z|x}^m(z^u) - S_{z|x}^e(z^u)| + |F_{z|x}^m(z^l) - S_{z|x}^e(z^l)|) dz = \int_{-\infty}^{+\infty} (|S_{z|x}^e(z^u) - F_{z|x}^m(z^u)| + |S_{z|x}^e(z^l) - F_{z|x}^m(z^l)|) dz = d(S^e, F^m)$$

所以随机-区间混合不确定性模型确认指标 $d(F^m, S^e)$ 具有对称性。

3) 三角不等性

设 F^{m1} 和 F^{m2} 分别为模型 1 和模型 2 输出响应量对应的分布函数。

由于

$$d(F^{m1}, F^{m2}) = \int_{-\infty}^{+\infty} (|F_{z|x}^{m1}(z^u) - F_{z|x}^{m2}(z^u)| + |F_{z|x}^{m1}(z^l) - F_{z|x}^{m2}(z^l)|) dz = \int_{-\infty}^{+\infty} [|(F_{z|x}^{m1}(z^u) - S_{z|x}^e(z^u)) - (F_{z|x}^{m2}(z^u) - S_{z|x}^e(z^u))| + |(F_{z|x}^{m1}(z^l) - S_{z|x}^e(z^l)) - (F_{z|x}^{m2}(z^l) - S_{z|x}^e(z^l))|] dz \leq \int_{-\infty}^{+\infty} (|F_{z|x}^{m1}(z^u) - S_{z|x}^e(z^u)| + |F_{z|x}^{m2}(z^u) - S_{z|x}^e(z^u)| + |F_{z|x}^{m1}(z^l) - S_{z|x}^e(z^l)| + |F_{z|x}^{m2}(z^l) - S_{z|x}^e(z^l)|) dz = \int_{-\infty}^{+\infty} (|F_{z|x}^{m1}(z^u) - S_{z|x}^e(z^u)| + |F_{z|x}^{m1}(z^l) - S_{z|x}^e(z^l)|) dz + \int_{-\infty}^{+\infty} (|F_{z|x}^{m2}(z^u) - S_{z|x}^e(z^u)| + |F_{z|x}^{m2}(z^l) - S_{z|x}^e(z^l)|) dz = d(F^{m1}, S^e) + d(F^{m2}, S^e)$$

所以随机-区间混合不确定性模型确认指标 $d(F^m, S^e)$ 具有三角不等性。

4) 理想模型指标趋零性

理想情况下, 当且仅当数学模型与物理实验过程完全一致时, 如果样本量趋于无穷, 则模型输出响应量的上下界分布函数曲线应该与实验响应量的上下界经验分布函数曲线重合, 即 $d(F^m, S^e)$ 收敛于零。

2.3 指标的求解

本文依据大数定理, 运用 Monte Carlo 数字模拟法, 提出了所建模型确认指标的求解方法, 具体

过程可概括为7个步骤。

步骤1 获取模型随机输入向量的实现值。根据随机输入向量 \mathbf{X} 的联合概率密度 $f_{\mathbf{X}}(\mathbf{x})$, 产生样本容量为 M_R^m 的模型输入随机向量实现值。 $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{in_R})$, $i = 1, 2, \dots, M_R^m$ 。

步骤2 离散模型区间输入向量。将模型的区间输入向量 $\mathbf{Y} = (Y_1, Y_2, \dots, Y_{n_1})$ 在其取值区间内均匀的离散化为样本容量为 N_1 的样本 $\mathbf{y}_k = (y_{k1}, y_{k2}, \dots, y_{kn_1})$, $k = 1, 2, \dots, N_1$ 。

步骤3 求解模型输出响应量的区间集。在模型随机输入向量 \mathbf{X} 的每一个实现值 $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{in_R})$ ($i = 1, 2, \dots, M_R^m$) 处, 根据区间输入向量 \mathbf{Y} 的 N_1 个离散化样本 $\mathbf{y}_k = (y_{k1}, y_{k2}, \dots, y_{kn_1})$ ($k = 1, 2, \dots, N_1$), 计算对应的模型响应量 $z_k^m | \mathbf{x}_i$, 进而求得相应于 $\mathbf{X} = \mathbf{x}_i$ 的模型输出响应量的区间集 $[z^{mL} | \mathbf{x}_i, z^{mU} | \mathbf{x}_i]$ ($i = 1, 2, \dots, M_R^m$)。

步骤4 求解模型输出响应量的上、下界分布函数。依据步骤3求得的 M_R^m 个模型输出响应量不同区间的上下界 $z^{mU} | \mathbf{x}_i$ 和 $z^{mL} | \mathbf{x}_i$, 求得模型输出响应量的上下界分布函数 $F_{z^U | \mathbf{X}}(z^U)$ 和 $F_{z^L | \mathbf{X}}(z^L)$ 。

步骤5 求解实验输出响应量的区间集。对于按照 \mathbf{X} 的联合概率密度 $f_{\mathbf{X}}(\mathbf{x})$ 选取 M_R^e 个随机输入向量 \mathbf{X} 的实验样本值 $\mathbf{x}_j = (x_{j1}, x_{j2}, \dots, x_{jn_R})$ ($j = 1, 2, \dots, M_R^e$), 分别固定每个实验样本 \mathbf{x}_j , 通过实验得到与区间变量向量 \mathbf{Y} 的 N_1 个离散化样本 $\mathbf{y}_k = (y_{k1}, y_{k2}, \dots, y_{kn_1})$ 对应的输出响应量观测值 $z_k^e | \mathbf{x}_j$, 最后求得相应于 $\mathbf{X} = \mathbf{x}_j$ 的实验输出响应量的区间集 $[z^{eL} | \mathbf{x}_j, z^{eU} | \mathbf{x}_j]$ ($j = 1, 2, \dots, M_R^e$)。

步骤6 求解实验输出响应量的上下界分布函数。依据步骤5求得的 M_R^e 个实验输出响应量的不同区间的上下界 $z^{eU} | \mathbf{x}_j$ 和 $z^{eL} | \mathbf{x}_j$, 求得实验输出响应量的上下界经验分布函数 $S_{z^U | \mathbf{X}}(z^U)$ 和 $S_{z^L | \mathbf{X}}(z^L)$ 。

步骤7 通过式(2)计算随机-区间混合不确定性模型确认指标 $d(F^m, S^e)$ 。

3 算例分析

为了进一步验证所提指标的可行性和有效性, 下面通过2个算例来进行说明。

3.1 数字算例

在本算例中, 实验数据通过式(3)来获取。

$$y^e(x, \theta) = \sin(2x - 0.5\pi\theta) + \cos(\theta + 0.25\pi) + 0.2x + \varepsilon_e \quad (3)$$

式中: $y^e(x, \theta)$ 为实验响应量, $\theta = 1.5$ 为输入参数, $x \in [3, 6]$ 为输入区间变量; $\varepsilon_e \sim N(0, 0.2^2)$ 为实验响应量的测量误差。选取与物理实验对应的6个备选数学模型, 分为2组进行讨论。各个备选模型输入变量的参数或分布函数存在一定差异, 具体情况见表1。设模型的随机输入向量的样本容量用 M_R^m 表示, 物理实验随机输入向量的样本容量用 M_R^e 表示, 区间输入向量的离散点数量用 N_1 表示。

1) 第1测试组

在第1测试组的3个模型中, 由于模型1与实验模型相同, 模型2的参数 θ 大于实验模型的参数 θ ; 模型3的参数 θ 大于模型2的参数 θ , 所以可以定性地判断出模型1最准确, 其次是模型2, 最后是模型3。这组测试是为了验证在模型参数不同的情况下, 所建指标是否可行和有效。

本组测试选取模型的随机输入变量的样本容量 $M_R^m = 10\ 000$, 实验随机输入变量的样本容量 $M_R^e = 1\ 000$, 区间输入变量的离散点数量 $N_1 = 300$, 按照2.3节提出的指标求解方法步骤, 表1中第1测试组的3个模型确认指标的计算结果见表2。

由表2可以得到, 模型1的指标值最小(0.0076), 模型2次之(0.1551), 而模型3的指标值最大(0.3320), 由此可以定量地判断出: 在这3个模型中模型1最优, 其次是模型2, 最后是模型3, 这一判断结论正好与定性分析的结论相同。此外, 按照2.2节讨论的该指标的性质4), 当且仅当模型与实验过程完全一致时, 对应的模

表1 2组测试计算模型

Table 1 Computation models in two test cases

测试组	模型编号	模型公式
第1组	1	$y^{m1}(x) = y^e(x, \theta = 1.5)$
	2	$y^{m2}(x) = y^e(x, \theta = 1.6)$
	3	$y^{m2}(x) = y^e(x, \theta = 1.7)$
第2组	4	$y^{m4}(x) = y^e(x, \theta \sim N(1.5, 0.15^2))$
	5	$y^{m5}(x) = y^e(x, \theta \sim N(1.5, 0.3^2))$
	6	$y^{m6}(x) = y^e(x, \theta \sim N(1.6, 0.3^2))$

注: $y^{mi}(x)$ ($i = 1, 2, \dots, 6$) 表示第 i 个模型的输出响应量。

表2 数字算例的指标计算结果

Table 2 Metric computation results of numerical test case

测试组	模型编号	指标值
第1组	1	0.0076
	2	0.1551
	3	0.3320
第2组	4	0.0502
	5	0.1635
	6	0.2767

型确认指标值将会收敛于零。但从表2可以得到,模型1的指标值(0.0076)并不为零。这一问题的产生原因是在指标求解过程中, $F_{z|x}^m(z)$ 和 $S_{z|x}^e(z)$ 的上下界均是通过有限样本估计所得,这样必然会产生一定的误差。但如果进一步增大样本容量,或提高模型计算精度,那么 $F_{z|y}^m(y,z)$ 和 $F_{z|y}^e(y,z)$ 上下界的计算结果将会趋于理论值,同时指标的计算结果也将会趋于零。由此可见,在模型参数存在差异的情况下,不同模型与物理实验之间的差异程度能够运用所建指标进行有效度量,并能够正确判断它们之间的优劣。

2) 第2测试组

第2测试组也包括3个数学模型。与实验模型相比,模型4、模型5和模型6均将参数 θ 作为随机变量。模型4和模型5中参数 θ 的均值同为1.5,但模型4中 θ 的方差比模型5的小;模型5和模型6中参数 θ 的方差相同,但模型6中 θ 的均值大于模型5中 θ 的均值。显然在这3个模型中,模型4相对最优,其次是模型5,而模型6最差。第2测试组是为了验证在模型参数的分布函数具有很小差异的情况下所建指标是否依然可行有效。

同理,在第2测试组中,选取模型的随机输入变量的样本容量 $M_R^m = 10000$,实验随机输入变量的样本容量 $M_R^e = 1000$,区间输入变量的离散点 $N_I = 300$,按照2.3节提出的指标求解方法步骤,表1中第2测试组的3个模型确认指标的计算结果见表2。

由表2可知,模型4的指标值最小(0.0502),模型5次之(0.1635),而模型6的指标值最大(0.2767),即通过模型确认指标的结果可以判断出模型4相对最优,其次是模型5,而模型6最差,这一结论与定性分析结果完全一致。因此,所建指标能够对分布函数参数存在很小差异的不同模型的优劣进行正确地判断。

3.2 工程算例

对航空发动机而言,最为关键的一个转动部件就是涡轮盘,发动机在起动和加速过程中,涡轮盘既要承受非常大的离心力,又要承受很大的热应力。由于涡轮盘的结构复杂,工作环境恶劣,因此工作一段时间后,一些部位(如榫槽槽底、销钉孔等)容易出现应力集中而出现裂纹。图3为某型航空发动机涡轮盘的榫槽槽底裂纹示意图。

设该型航空发动机在最大转速工作时,它的涡轮盘的榫槽槽底所受最大载荷的实验数据由解析方程式(4)产生:

$$F = \frac{C\omega^2}{2\pi} + 2\rho\omega^2 J + \varepsilon_e \tag{4}$$

式中: ρ 、 C 、 ω 和 J 分别为发动机涡轮盘的质量密度、系数、转动角速度和截面惯性矩, ρ 、 C 和 J 皆为随机变量,它们的分布参数见表3。转动角速度 $\omega = 2n\pi$ 为区间变量,其中 $n \in [180, 220]$ 为最大转动频率。测量误差 $\varepsilon_e \sim N(0, 1.1 \times 10^{10})$ 。

假设工业部门建立了3个不同的航空发动机涡轮盘预测模型。与实验模型相比,这3个模型中的随机变量存在一定差异,具体情况见表4。

由表4可以得到,与实验表达式相比,模型1是一个完全正确的模型,模型2中的随机变量的变异系数发生了变化,模型3中随机变量不仅变异系数发生了变化,而且均值也发生了变化。因此,可以定性判断出模型1优于模型2,模型2优于模型3。

同数学算例相似,在本测试中选取模型的随机输入变量的样本容量 $M_R^m = 10000$,实验随机



图3 航空发动机涡轮盘模型裂纹示意图

Fig.3 Schematic diagram of crack of an aero engine turbo blade model

表3 发动机涡轮盘输入变量分布参数

Table 3 Distribution parameters of input variables of engine turbo blade

随机变量	分布类型	均值	变异系数
ρ	对数正态	8 240	0.1
C	对数正态	5.67	0.1
J	正态	1.22×10^{-4}	0.1

表4 航空发动机涡轮盘模型输入变量分布参数

Table 4 Distribution parameters of input variables of models of aero engine turbo blade

随机变量	分布类型	模型1		模型2		模型3	
		均值	变异系数	均值	变异系数	均值	变异系数
ρ	对数正态	8 240	0.1	8 240	0.15	8 240	0.15
C	对数正态	5.67	0.1	5.67	0.15	5.67	0.15
J	正态	1.22×10^{-4}	0.1	1.22×10^{-4}	0.15	1.31×10^{-4}	0.15

输入变量的样本容量 $M_R^c = 1\ 000$, 区间输入变量的离散点 $N_1 = 40$, 按照 2.3 节提出的指标求解方法步骤, 模型确认指标的计算结果见表 5。

由表 5 可以得到, 运用本文提出的模型确认指标能够清楚地判断出模型 1 优于模型 2, 模型 2 优于模型 3, 这一结论与定性结论吻合, 从而再次验证了本文所提指标的可行性和有效性。

表 5 航空发动机涡轮盘模型确认指标计算结果

Table 5 Model validation metric computation results of aero engine turbo blade

模型	模型 1	模型 2	模型 3
指标值	0.0123	0.1369	0.2396

3.3 模型确认指标的风险分析

虽然前文的理论和算例分析的结果都表明, 本文所提指标对于解决随机和区间变量共存条件下的模型确认问题可行有效, 但这一结论是以模型和实验样本比较充足为前提条件。一旦模型或实验的样本量比较匮乏时, 运用该指标进行模型确认就会存在一定的风险。下面以算例 2 为例, 分 2 种情况对其存在风险进行讨论。

1) M_R^c 固定, M_R^m 变化

保持区间输入变量离散点 $N_1 = 40$ 和实验随机输入变量的样本容量 $M_R^c = 1\ 000$ 不变, 当模型随机输入变量的样本容量 M_R^m 变化时, 模型确认指标计算结果随的变化情况如图 4 所示。

由图 4 可以得到, 随着模型随机输入变量的样本容量 M_R^m 的增大, 模型确认指标值 d 迅速减小, 直到 $M_R^m = 100$ 时, 指标值 d 开始趋于稳定, 并收敛到一个固定值。这一现象表明随着 M_R^m 的不断增大, 模型响应量的区间上界的分布函数的计算结果趋于其理论解。但当 $M_R^m < 30$ 时, 由于模型响应量的区间上界的分布函数的估计误差比较大, 在这种情况下, 运用模型确认指标对不同的模型的优劣进行判断就会存在错判误判的风险。因此, 运用本文方法进行模型确认时, 要尽量获得充足的模型随机输入变量样本。

2) M_R^m 固定, M_R^c 变化

保持区间输入变量的离散点 $N_1 = 40$ 和模型随机输入变量 $M_R^m = 10\ 000$ 不变, 当实验随机输入变量的样本容量 M_R^c 变化时, 模型确认指标的计算结果随 M_R^c 的变化情况如图 5 所示。

由图 5 可以得到, 随着实验随机输入变量样本容量 M_R^c 的增大, 模型确认指标 d 同样迅速减小并趋于收敛。当 $M_R^c < 30$ 时, 模型确认指标 d 的计算结果存在相互交叉重叠现象, 这表明在实

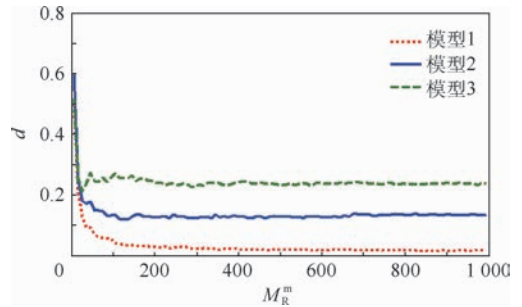


图 4 模型确认指标 d 随 M_R^m 的变化

Fig. 4 Model validation metric d versus M_R^m

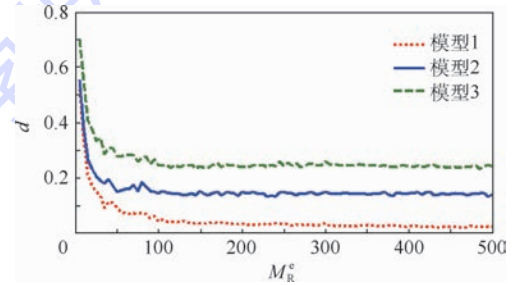


图 5 模型确认指标 d 随 M_R^c 的变化

Fig. 5 Model validation metric d versus M_R^c

验随机输入变量的样本小于 30 时, 通过模型确认指标对不同模型的优劣进行判断, 会存在一定的错判误判风险。产生这一问题的原因是由于实验随机输入变量的样本过少时, 物理实验输出响应量上下界的经验分布函数难以准确估计所致。但当实验随机输入变量的样本容量大于 100 时, 该指标能够客观地评估不同数学模型与实验的一致性。

4 结 论

1) 分析了随机和区间变量共存条件下数学模型的特点。对于随机-区间混合不确定性因素影响下的数学模型而言, 当它的随机输入变量取任意一个实现值时, 模型输出响应量则变为一个区间变量; 当随机输入变量按照它的概率密度在其取值范围内取所有实现值时, 模型的输出响应量则为一个随机区间变量。因此, 随机-区间混合不确定性模型确认的实质上就是比较模型输出响应量随机区间变量与实验输出响应量随机区间变量之间的差异程度。

2) 提出了随机和区间变量共存条件下的模型确认指标。根据随机-区间混合不确定性模型的特点, 运用概率方法和区间理论, 将模型输出响应量的上下界分布函数与实验输出响应量的上下界经验分布函数之间的面积差异, 定义为模型确认指标, 解决了随机和区间变量共存条件下, 数学

模型在描述真实物理实验时的准确性评估方法问题。

3) 分析了本文所提指标可能面临的风险。虽然理论和算例分析的结果都表明,本文所提指标在随机和区间变量共存条件下的模型确认方面可行有效,但该指标的有效性是建立在模型和实验样本比较充足的前提条件之下。一旦模型或实验的样本量比较匮乏时,运用新指标对不同的数学模型与实验过程的一致性进行判断时,会存在一定的错判误判风险。

参考文献 (References)

- [1] OBERKAMPF W L, ROY C J. Verification and validation in scientific computing [M]. New York: Cambridge University Press, 2010; 21-25.
- [2] OBERKAMPF W L, BARONE M F. Measures of agreement between computation and experiment: Validation metrics [J]. Journal of Computational Physics, 2006, 217; 5-36.
- [3] OBERKAMPF W L, SINDIR M, CONLISK A. Guide for the verification and validation of computational fluid dynamics simulations [M]. Reston: AIAA, 1998; 2-4.
- [4] XIONG Y, CHEN W, TSUI K L, et al. A better understanding of model updating strategies in validating engineering models [J]. Compute Methods in Applied Mechanics and Engineering, 2009, 189(15-16); 1327-1337.
- [5] MESSER M, PANCHAL J H, KRISHNAMURTHY V, et al. Model selection under limited information using a value of information based indicator [J]. Journal of Mechanical Design, 2010, 132(12); 121008.
- [6] KENNEDY M C, O'HAGAN A. Bayesian calibration of computer models [J]. Journal of the Royal Statistical Society, 2001, 63(3); 425-464.
- [7] BAYARRI M J, BERGER J O, PAULO R, et al. A framework for validation of computer models [J]. Technometrics, 2007, 49(2); 138-154.
- [8] ARENDT P D, APLEY D W, CHEN W, et al. Improving identifiability in model calibration using multiple responses [J]. Journal of Mechanical Design, 2012, 134(10); 100909.
- [9] ARENDT P D, APLEY D W, CHEN W. Quantification of model uncertainty: Calibration, model discrepancy, and identifiability [J]. Journal of Mechanical Design, 2012, 134(10); 100908.
- [10] LIU Y, CHEN W, ARENDT P, et al. Toward a better understanding of model validation metrics [J]. Journal of Mechanical Design, 2011, 133(7); 071005.
- [11] REBBA R, MAHADEVAN S. Validation of models with multivariate output [J]. Reliability Engineering and System Safety, 2006, 91(8); 861-871.
- [12] FERSON S, OBERKAMPF W L. Validation of imprecise probability models [J]. International Journal of Reliability and Safety, 2009, 3; 3-22.
- [13] FERSON S, OBERKAMPF W L, GINZBURG L. Model validation and predictive capability for the thermal challenge problem [J]. Computer Methods in Applied Mechanics and Engineering, 2008, 197(29-32); 2408-2430.
- [14] 赵录峰, 吕震宙, 张磊刚, 等. 多输出模型确认中的混合矩指标 [J]. 国防科技大学学报, 2015, 37(6); 61-68.
- ZHAO L F, LYU Z Z, ZHANG L G, et al. Mixed moment validation metric for models with multivariate output [J]. Journal of National University of Defense Technology, 2015, 37(6); 61-68 (in Chinese).
- [15] LI W, CHEN W, JIANG Z, et al. New validation metrics for models with multiple correlated responses [J]. Reliability Engineering and System Safety, 2014, 127; 1-11.
- [16] ZHAO L F, LU Z Z, YUN W Y, et al. Validation metric based on Mahalanobis distance for models with multiple correlated responses [J]. Reliability Engineering and System Safety, 2017, 159; 80-89.
- [17] 胡嘉蕊, 吕震宙. 基于核主成分分析的多输出模型确认方法 [J]. 北京航空航天大学学报, 2017, 43(7); 1470-1480.
- HU J R, LYU Z Z. Model validation model with multivariate output based on kernel principal component analysis [J]. Journal of Beijing University of Aeronautics and Astronautics, 2017, 43(7); 1470-1480 (in Chinese).

作者简介:

赵录峰 男, 博士研究生。主要研究方向: 可靠性工程、模型确认。

吕震宙 女, 教授, 博士生导师。主要研究方向: 可靠性工程、灵敏度分析、模型确认和多学科优化。

阚丽娟 女, 博士研究生。主要研究方向: 安全性工程。

A validation metric for model with mixture of random and interval variables

ZHAO Lufeng¹, LYU Zhenzhou^{1,*}, KAN Lijuan²

(1. School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, China;

2. Equipment Management and Safety Engineering College, Air Force Engineering University, Xi'an 710051, China)

Abstract: The existing model validation methods under uncertainty based on theory of probability are only applicable to validate model with random variables, but inapplicable to validate model with the mixture of random and interval variables. To address this issue, the validation method for model with the mixture of random and interval variables is studied in this paper. First, the characteristics of the mathematical model with the mixture of random and interval variables are analyzed. Second, a new validation metric is proposed by using interval theory and probability method. This metric provides a comparison between the cumulative distribution functions (CDFs) of the upper and the lower bounds of the model responses and the empirical CDFs of the upper and the lower bounds of the experimental responses to show the disagreement between the quantitative predictions from a model and the physical observations. The mathematical properties of the new metric are discussed, and its estimation method and procedures are presented. Finally, the feasibility and effectiveness of the proposed validation metric are illustrated by a numerical test case and an engineering application with mixture of random and interval variables.

Keywords: model validation; metric; random variables; interval variables; mixed model

Received: 2017-05-22; **Accepted:** 2017-06-30; **Published online:** 2017-08-04 16:24

URL: kns.cnki.net/kcms/detail/11.2625.V.20170804.1624.008.html

Foundation items: National Natural Science Foundation of China (51475370); the Fundamental Research Funds for the Central Universities (3102015BJ(II)CG009)

* **Corresponding author.** E-mail: zhenzhoulu@nwpu.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2017.0378

基于系统 PSN 曲线的齿轮箱疲劳可靠度评估



马洪义, 谢里阳*

(东北大学 现代设计与分析研究所, 沈阳 110819)

摘 要: 蒙特卡罗方法可以准确评估复杂机械系统疲劳共因失效概率,但效率偏低,因此提出系统 PSN 曲线的概念和基于此概念的系统可靠度蒙特卡罗评估方法。在给定的恒幅载荷下,基于同一零件的疲劳寿命在不同应力水平下的概率分位点具有一致性的原则,对系统中零件 PSN 曲线进行随机抽取;根据线性累积损伤法则和相应的系统可靠度模型,得到齿轮传动的恒幅载荷下的疲劳寿命分布,拟合恒幅载荷与寿命分布之间的关系得到系统 PSN 曲线。将系统视为一个零件,完成“零件”—“系统”—“零件”的寿命分析过程。通过损伤等效原则,将随机载荷下的复杂串联系统可靠度评估问题转化为恒幅载荷下零件的可靠度评估问题。

关键词: 疲劳可靠度; 系统 PSN 曲线; 串联系统; 共因失效; 蒙特卡罗

中图分类号: V215.5⁺2; TH132.4

文献标识码: A **文章编号:** 1001-5965(2018)05-0975-07

齿轮传动系统在可靠度意义上可视为串联系统,系统中任意轮齿疲劳失效则系统失效。齿轮传动系统作为设备中造价昂贵的关键部件,其疲劳可靠度评估十分重要。

目前,科研人员针对齿轮传动系统可靠度评估问题进行了许多研究,研究方法有通过对产品寿命直接分析,通过状态监测手段评估,以及基于应力-强度干涉模型的评估方法。

通过对产品寿命直接分析^[1]得到寿命分布,通过寿命分布预测可靠度。该方法适用于已经投入使用的齿轮产品,可以更全面地体现环境随机性和复杂性。但是对处于设计、研发或只有少量生产的产品,往往难以获得足够的失效数据。

通过状态监测手段评估系统可靠度。例如对齿轮箱润滑油的监测以及振动情况的监测^[2-3],在线监测信号可以很好地反映当前系统的磨损状态,但齿轮属于无初始裂纹零件,其疲劳失效具有

突然性,多数疲劳失效在破坏前并不能发出异常信号。

基于应力-强度干涉模型^[4]评估系统可靠度。Dong 等^[5]基于裂纹扩展理论研究了齿轮轮齿的齿面接触疲劳寿命。Place 等^[4]以扩展的应力-强度干涉模型研究了直升机齿轮箱系统的可靠度,并以 2 个零件的串联系统为例,通过多元累积分布函数的轮廓线方法计算了考虑失效相关的系统可靠度。Nejad 等^[6]以经验公式方法计算了风力发电机齿轮箱内各轮齿对输入轴扭矩的应力响应,并使用线性累积损伤方法计算了随机载荷下各轮齿的寿命。笔者团队^[7]运用全概率思想建立了考虑失效相关的风力发电机齿轮箱串联系统可靠度函数。此外,经典的结构可靠度评估方法^[8-10]和较新的系统可靠度评估方法^[11-13]都是在应力-强度干涉模型基础上的不同扩展。然而齿轮系统存在多部件,多失效模式且失效相关,同

收稿日期: 2017-06-05; 录用日期: 2017-07-07; 网络出版时间: 2017-09-15 17:09

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20170915.1709.003.html

基金项目: 国家自然科学基金(51335003)

* 通信作者. E-mail: lyxie@mail.neu.edu.cn

引用格式: 马洪义, 谢里阳. 基于系统 PSN 曲线的齿轮箱疲劳可靠度评估[J]. 北京航空航天大学学报, 2018, 44(5): 975-981.
MA H Y, XIE L Y. Assessment of gearbox fatigue reliability based on system PSN curve [J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 975-981 (in Chinese).

时具有变构特性,现有的基于应力-强度干涉模型的可靠度评估方法^[4-13]尚未完全考虑影响可靠度的因素(失效相关、系统变构特性)。

针对上述问题,本文提出系统 PSN 曲线的思想,在此基础上通过蒙特卡罗仿真方法建立齿轮串联系统的寿命评估模型,以提高仿真效率且兼顾齿轮传动系统共因失效。

1 齿轮箱结构及零件 PSN 曲线

1.1 齿轮箱结构

齿轮系统结构选用 750 kW 级风力发电机常用增速齿轮箱,该齿轮箱由一级行星轮系与两级固定轴斜齿轮副组成,如图 1 所示。

图 1 中 T_{in} 为输入轴, r 为固定的内齿圈, s 为太阳轮, c 为行星架, p 为行星轮, 4、5 为互相啮合的中速级斜齿轮, 6、7 为互相啮合的高速级斜齿轮, T_{out} 为输出轴。

由文献[14]获知风力发电机常用齿轮箱各齿轮轮齿几何参数如表 1 所示。

表 1 中 Z 为齿数, M_n 为齿轮模数, α_n 为齿轮法

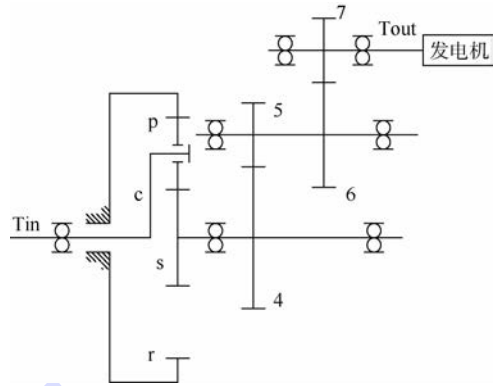


图 1 齿轮箱传动原理图

Fig.1 Schematic of gearbox transmission

向压力角, β 为齿轮螺旋角, α'_i 为端面压力角。

根据齿轮传动系统结构和各齿轮轮齿参数,按照机械设计手册中斜齿轮齿根强度校核相关理论,计算各轮齿所受尺面接触应力 σ_H 与输入轴扭矩关系如表 2 所示。表中 $\sqrt{K_1} = \sqrt{K_2} = \sqrt{K_3} = \sqrt{K_4} = \sqrt{K_A K_V K_{H\beta} K_{H\alpha}} = 2.03$, K_A 为使用系数, K_V 为动载系数; $K_{H\beta}$ 为纵向载荷分布系数; $K_{H\alpha}$ 为横向载荷分布系数; T_{in} 为输入轴扭矩。

表 1 齿轮箱零件参数

Table 1 Gearbox component parameters

参数	太阳轮	行星轮	内齿圈	齿轮 4	齿轮 5	齿轮 6	齿轮 7
Z	21	38	99	84	23	92	24
M_n	10	10	10	8	8	5	5
$\alpha_n / (^\circ)$	20	20	20	20	20	20	20
$\beta / (^\circ)$	7.5	7.5	7.5	14	14	14	14
$\alpha'_i / (^\circ)$	22.6	22.6	22.6	21.5	21.5	21.5	21.5

表 2 齿面接触应力

Table 2 Gear surface contact stress

啮合齿轮	太阳轮/行星轮	行星轮/内齿圈	齿轮 1/齿轮 2	齿轮 3/齿轮 4
σ_H / MPa	$0.0521 \sqrt{K_1 T_{in}}$	$0.0225 \sqrt{K_2 T_{in}}$	$0.0309 \sqrt{K_3 T_{in}}$	$0.0316 \sqrt{K_4 T_{in}}$

1.2 零件 PSN 曲线

文献[15]中 PSN 曲线不包含寿命在 $10^4 \sim 10^7$ 循环之外的寿命所对应的载荷,因此大载荷造成的损伤不能根据该曲线定量。近些年超高周疲劳研究发现很多小的载荷仍能对材料造成损伤,且大、小载荷均会出现在随机载荷历程中。因此,本文在文献[15]的基础上重新构建了零件 PSN 曲线。零件在恒幅循环载荷下的寿命服从对数正态分布^[15],其对数寿命中值 PSN 曲线为 $\mu(s) = \ln(c) + m \cdot \ln(s)$,其中 $c = 5.6322 \times 10^{31}$, $m = -8.4517$, c 和 m 为参数。本文将不同应力下的对数寿命的标准差 σ 均设为其对数寿命均值的 3.5%,即 $\sigma(s) = 0.035\mu(s)$ 。通过计算多级恒幅载荷下相同失效概率对应的寿命值,然后

拟合应力与寿命值关系,得到该失效概率下的零件 PSN 曲线如图 2 所示。

经过比较,该曲线在寿命 $10^4 \sim 10^7$ cycle 部

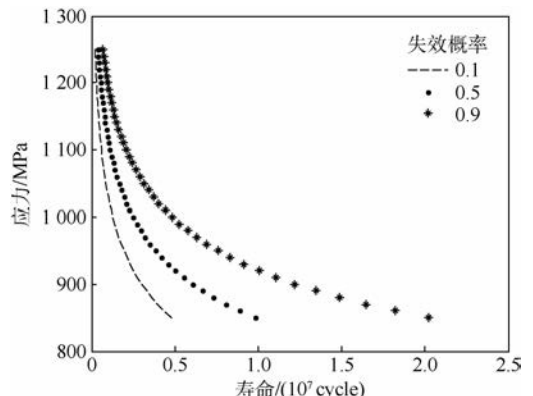


图 2 零件 PSN 曲线

Fig.2 PSN curves of components

分接近文献[15]的曲线,其中 cycle 为循环次数。

$$N_s = c_1 s_i^{-m_1} \quad (1)$$

2 系统寿命评估的蒙特卡罗仿真

蒙特卡罗仿真计算寿命的关键是零件所受载荷的抽样是否正确,零件疲劳性能的抽样是否合理,以及零件受载后的损伤计算。假设零件所受载荷为假设零件所受载荷为服从某种分布的随机变量,当载荷作用一次时,即从分布中随机抽取一个载荷值。零件的疲劳性能抽样是靠对零件的PSN曲线的随机生成实现的:首先对1 000 MPa恒幅应力下的寿命进行随机抽样得到一个寿命样本,然后计算该寿命样本对应的累积概率 P_1 (P_1 为第1次失效概率 P 的计算样本值),则 P_1 所对应的PSN曲线抽样完为常数 c_1 和 m_1 , s 为确定性应力, N_s 为确定性寿命变量:

当第 i 次载荷作用时,产生确定的应力,该载荷产生的损伤 $1/N_s(s_i)$ 为确定值,即 $1/(c_1 s_i^{-m_1})$ 。在多次载荷作用后,应用线性累积损伤法则计算总的轮齿损伤,当任意轮齿的损伤大于1时,系统失效。

每个轮齿都相当于串联系统的一个零件,表1显示该系统共有457个零件。齿轮系统每次啮合时都有若干轮齿参与啮合,而参与啮合的轮齿所受应力均来源于相同的外部输入载荷——输入轴扭矩,参与啮合的轮齿损伤为共因损伤。齿轮系统每次啮合时参与啮合的轮齿不同,所以从抵抗疲劳载荷的能力上看,齿轮系统是时变的。本文编制蒙特卡罗数值仿真流程(见图3)计算齿轮传动系统可靠度。

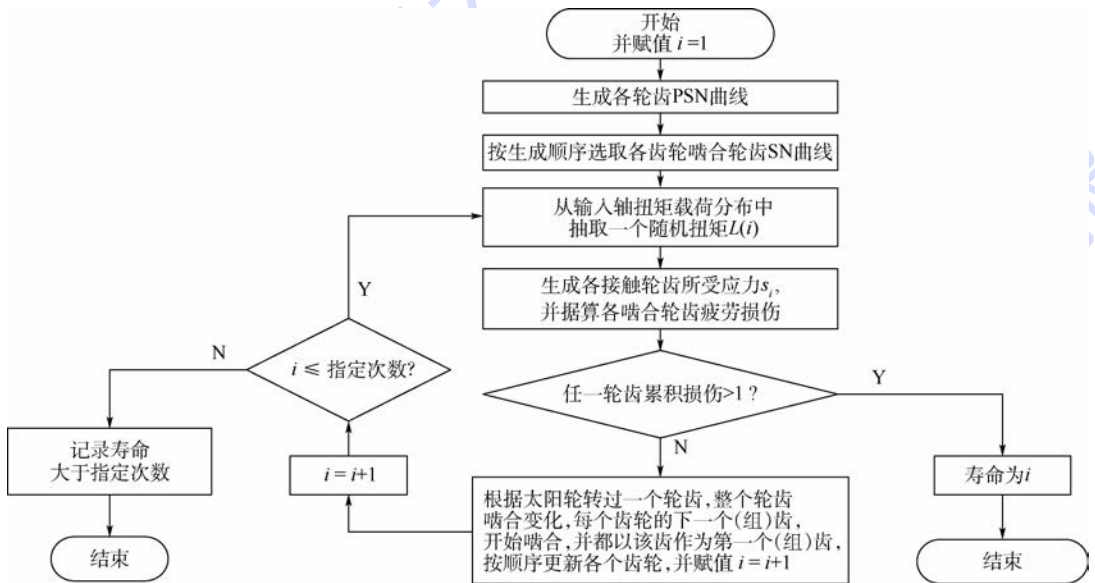


图3 齿轮传动系统可靠度仿真流程

Fig. 3 Gear transmission system reliability simulation process

假设输入轴扭矩载荷为随机变量且服从威布尔分布 $W(1.25 \times 10^8, 1.63)$,50次仿真得到50个系统寿命数据,以对数正态分布对寿命数据进行拟合,得到对数正态分布参数为(11.4857, 0.3480),其统计分布如图4所示。

使用MATLAB拟合优度KS检验模块进行假设检验,检验统计量的观测值ksstat为0.0628。由于ksstat小于1,证明在显著水平0.05下接受系统寿命服从参数为(11.4857, 0.3480)的对数正态分布的假设。

当串联系统寿命较短或受载次数较少时,仿真次数较少,上述方法是简单有效的。当设计寿命较长时(例如风力发电机齿轮箱中值设计寿

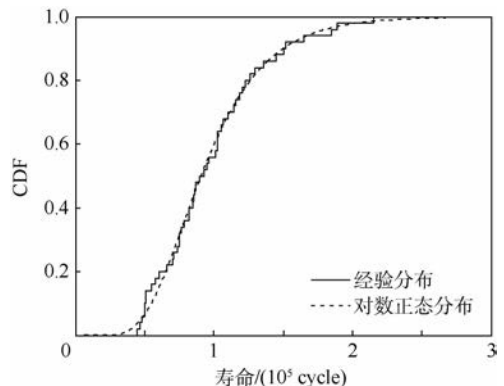


图4 基于零件PSN,随机载荷下齿轮系统寿命概率分布
Fig. 4 Gear system life probability distribution under random load based on component PSN

命通常为 20 年),例如 90% 可靠度下的寿命为 1×10^{10} 次以上时,采用上述的直接的蒙特卡罗仿真方法需要耗费大量时间。

为此本文在串联系统寿命的直接仿真基础上,进行了提升效率的研究。定义系统的恒幅循环下的疲劳特性(串联系统 PSN 曲线),这样就能够避免仿真时频繁使用多个零件的 PSN 曲线,从而达到把系统看成一个零件的目的,把复杂的系统问题变成简单的零件可靠度评估问题。

3 齿轮传动系统 PSN 曲线

将已经建立的蒙特卡罗仿真程序中随机载荷变为指定的确定性载荷,用于计算系统在恒幅载荷下的寿命。本文系统 PSN 曲线为 PSN 曲线的扩展,其中的“S”不是应力而是扭矩,亦可以写为系统 PTN 曲线(T 代表扭矩),表示不同失效概率 P 下寿命 N 与扭矩 S 的函数关系。在确定性的 3~4 级扭矩或者更多级恒幅扭矩下,将得到的寿命进行统计分析,可以得到各恒幅扭矩下的寿命拟合分布,进而获得串联系统 PSN 曲线。

3.1 齿轮传动系统恒幅应力下寿命生成

根据国标 GB/T 24176—2009^[16] 所述,为获得材料的 PSN 曲线,一般选定 3~5 级应力,且每级应力不低于 8 个试样。

齿轮系统内部的不同轮齿受应力不同,而输入轴扭矩一致,所以针对齿轮串联系统的蒙特卡罗仿真实验设计为在输入轴扭矩为 $2 \times 10^8, 3 \times 10^8, 5 \times 10^8, 6 \times 10^8 \text{ N} \cdot \text{mm}$ 下,各做 50 次仿真实验,寿命计算结果如表 3 所示。

表 3 各级恒幅扭矩下齿轮系统寿命统计及统计矩

Table 3 Gear system life and statistics moment under various constant torque

输入轴扭矩/ ($10^8 \text{ N} \cdot \text{mm}$)	寿命 标准差/cycle	寿命 均值/cycle
2	3.3×10^5	6.6×10^5
3	6.2×10^4	1.1×10^5
5	4.2×10^3	1.2×10^4
6	2.2×10^3	5.9×10^3

3.2 串联系统恒幅应力下寿命统计矩分析

将表 3 中齿轮串联系统的寿命均值和标准差与输入轴扭矩的关系画成散点图,从散点图中可以看出寿命均值和标准差与载荷呈现幂函数规律,所以选用形式为 $y = ax^b$ 的方程进行非线性回归,其散点图与回归曲线如图 5 所示。

齿轮串联系统的恒幅输入轴扭矩下标准差与

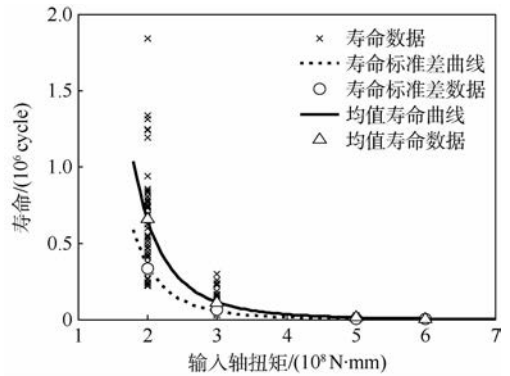


图 5 各级恒幅扭矩下齿轮系统寿命及统计矩

Fig. 5 Gear system life and statistical moment under various constant torque

应力函数关系、寿命均值与输入轴扭矩函数关系分别为

$$\text{var}(X_g/T) = (2.68 \times 10^{44} T^{-4.68})^2 \quad (2)$$

$$E(X_g/T) = 3.77 \times 10^{41} T^{-4.31} \quad (3)$$

式中: X_g 为齿轮系统寿命随机变量; T 为恒幅输入轴扭矩; $\text{var}(X_g/T)$ 为恒幅输入轴扭矩 T 下的寿命 X_g 的方差; $E(X_g/T)$ 为恒幅输入轴扭矩 T 下的寿命 X_g 的均值。

3.3 齿轮传动系统恒幅扭矩下寿命曲线

一般情况下,产品的寿命服从对数正态分布。本文按照对数正态分布对齿轮传动串联系统的恒幅载荷下的寿命进行分布拟合,拟合结果通过对数正态分布的假设检验。多级输入轴扭矩下的经验分布与对数正态分布曲线如图 6 所示。

从图 6 可以看出,齿轮串联系统在不同恒幅循环输入扭矩下的寿命的对数值的经验分布与取对数后的正态拟合分布之间的差异,可见齿轮系统拟合很好。

如果随机变量服从对数正态分布,则其概

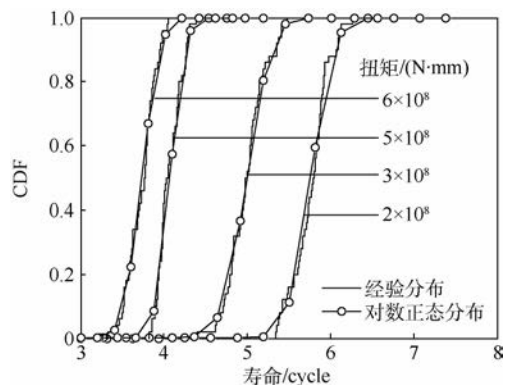


图 6 齿轮系统各级扭矩下对数寿命分布正态拟合

Fig. 6 Normal fitting of gear system logarithmic life distribution under various torque

率密度函数可以表达如式(4)所示,且其分布参数 μ 和 σ 与随机变量的数学期望和方差满足式(5)和式(6)的函数关系。

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{(2\sigma^2)}\right] \quad (4)$$

$$\mu = \ln(E(X)) - 0.5 \ln\left(1 + \frac{\text{var}(X)}{E(X)^2}\right) \quad (5)$$

$$\sigma^2 = \ln\left(1 + \frac{\text{var}(X)}{E(X)^2}\right) \quad (6)$$

结合式(2)~式(6)推导出齿轮串联系统恒幅扭矩下寿命分布参数与扭矩 T 的关系为

$$\mu_g(T) = \ln(3.77 \times 10^{41} T^{-4.31}) - 0.5 \ln(1 + 5.05 \times 10^5 T^{-0.14}) \quad (7)$$

$$\sigma_g(T) = \sqrt{\ln(1 + 5.05 \times 10^5 T^{-0.14})} \quad (8)$$

由此可得任意恒幅载荷下链条系统的寿命分布。通过各级载荷下的相同寿命分位点对载荷关系的拟合可以得到系统 PSN 曲线,其拟合结果如图 7 所示。

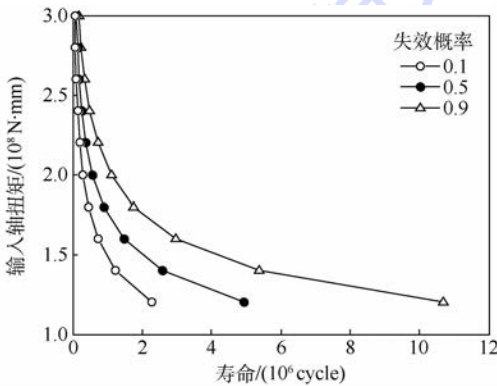


图 7 齿轮系统 PSN 曲线

Fig. 7 Gear system PSN curves

4 系统寿命评估方法

以系统 PSN 曲线代表系统的疲劳性能进行蒙特卡罗仿真时,系统寿命计算就简化为从已知分布中抽取一个载荷,从 PSN 曲线中随机抽取一条 SN 曲线代表系统疲劳性能,基于该 SN 曲线,每抽取一个载荷,系统产生 Miner 累积损伤,当疲劳损伤之和为 1 时,系统破坏,此时的载荷抽样次数即为系统寿命。

通过上述方法,求得齿轮传动系统寿命服从对数正态分布,其分布参数为 11.4014 和 0.3584。

对于变幅载荷下的零件疲劳寿命的计算,一个方便、可行的方法是根据疲劳累积损伤等效原则,把复杂变幅载荷历程转换为损伤等效的恒幅循环载荷^[17-18],在工程中可以认为在该等效恒幅载荷下的寿命分布与变幅载荷下的寿命分布近似。以

齿轮串联系统为例,在得到齿轮串联系统的 PSN 曲线之后,将该系统视为一个“零件”,进一步得到失效概率为 0.5 的中值 PSN 曲线为

$$N_{0.5} = 4.6634 \times 10^9 T^{-0.2499} \quad (9)$$

基于该曲线,“零件”在任意载荷 T 一次作用下将产生损伤 $1/N_{0.5}(T)$ 。当 T 服从某分布,密度函数为 $g(T)$ 时,其作用一次产生损伤的期望为

$$D = \int_0^{+\infty} g(T) \frac{1}{4.6634 \times 10^9 T^{-0.2499}} dT \quad (10)$$

由此可以计算出在该载荷分布下的平均寿命,该平均寿命对应于中值系统 PSN 曲线中的扭矩 T 即为等效载荷。通过该等效载荷求取“零件”恒幅寿命分布(见图 8)。

从图 8 中可以看出,在 3 种方法下,无论是链条串联系统还是齿轮串联系统的寿命分布都很相近。而以效率来看,无疑基于“串联系统 PSN”和等效载荷的模型最高,在针对设计寿命较长的系统时,可以通过在几级大载荷下获得其 PSN 曲线,将其等效为“零件”;而根据中值 SN 曲线可以很轻易的获得等效载荷,在等效载荷下求“零件”的寿命分布非常简单。

采用 3 种方法获得的寿命特征值及相对误差如表 4 所示。

方法 b 寿命分布与方法 a 寿命分布的误差为引入系统 PSN 曲线代替复杂的齿轮系统结构所产生;方法 c 的寿命分布与方法 a 寿命分布的相

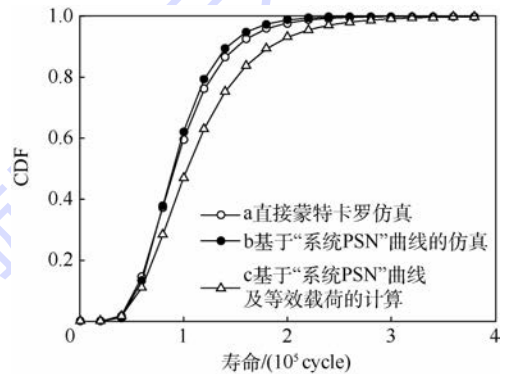


图 8 齿轮串联系统寿命分布

Fig. 8 Life distribution of gear series system

表 4 3 种方法寿命特征值比较

Table 4 Comparison of characteristic life of three methods

方 法	系统 寿命均值/ (10^4 cycle)	系统 寿命标准差/ (10^4 cycle)	以 a 为参照的 均值相对 误差/%	以 a 为参照的 标准差相对 误差/%
a	9.8	4.0	0	0
b	9.5	3.5	3.1	12.5
c	11	5.3	12.2	32.5

对误差产生原因在于系统 PSN 曲线和等效载荷代替了随机载荷的双重误差。

从图 8 和表 4 中可以看出,基于系统 PSN 曲线的和随机载荷的方法可以大大降低结构复杂程度对计算效率的影响,且能保持一个数量级左右的精度。可见系统 PSN 曲线的概念对于复杂结构的共因失效可靠度评估有一定工程意义。

基于系统 PSN 及等效载荷的系统可靠度(寿命分布)计算以损失一定计算精度获得计算效率显著提升,尤其在低设计寿命产品评估中,直接仿真方法由于过低的运算效率没有实用价值,提升效率往往会带来精度的降低符合一般科学规律。基于系统 PSN 及等效载荷的计算方法效率不受系统结构大小和载荷分布的影响,对于长设计寿命和短设计寿命的可靠度评估都是一种稳定的工程计算方法。

5 结 论

1) 根据寿命统计矩与载荷的函数关系,得到了寿命分布参数-载荷之间的幂函数回归方程,即在任意给定恒幅载荷下可以得到串联系统寿命的对数分布参数,通过对多级载荷下的寿命分布计算得到了系统 PSN 曲线。

2) 提出了基于系统 PSN 曲线和等效载荷的串联系统寿命评估方法。通过齿轮传动系统的 3 种寿命计算方法(在随机载荷下基于零件 PSN 曲线的直接仿真方法,在随机载荷下基于系统 PSN 的仿真方法,在等效载荷下基于系统 PSN 曲线的计算方法)的比较,证明了本文方法对齿轮传动系统寿命评估简单高效,具有普遍适用性。

参考文献 (References)

- [1] ASTRIDGE D G. Helicopter transmissions—design for safety and reliability[J]. Proceedings of the Institution of Mechanical Engineers Part G; Journal of Aerospace Engineering, 1989, 203 (27): 123-138.
- [2] SHENG S W. Investigation of oil conditioning, real-time monitoring and oil sample analysis for wind turbine gearboxes; NREL/PR-5000-50301 [R]. Golden: National Renewable Energy Laboratory, 2011.
- [3] SHENG S, OYAGUE F, BUTTERFIELD S. Investigation of various wind turbine drive train condition monitoring techniques; NREL/CP-500-46160 [R]. Golden: National Renewable Energy Laboratory, 2010.
- [4] PLACE C S, STRUTT J E, ALLSOPP K, et al. Reliability prediction of helicopter transmission systems using stress-strength interference with underlying damage accumulation[J]. Quality & Reliability Engineering International, 1999, 15 (2): 69-78.
- [5] DONG W, XING Y, MOAN T, et al. Time domain-based gear contact fatigue analysis of a wind turbine drivetrain under dynamic conditions[J]. International Journal of Fatigue, 2013, 48 (1): 133-146.
- [6] NEJAD A R, GAO Z, MOAN T. On long-term fatigue damage and reliability analysis of gears under wind loads in offshore wind turbine drivetrains[J]. International Journal of Fatigue, 2014, 61 (2): 116-128.
- [7] 谢里阳, 周金宇, 李翠玲, 等. 系统共因失效分析及其概率预测的离散化建模方法[J]. 机械工程学报, 2006, 42 (1): 62-68.
- [8] XIE L Y, ZHOU J Y, LI C L. Common cause failure analysis and discretely modeling for system probability prediction[J]. Journal of Mechanical Engineering, 2006, 42 (1): 62-68 (in Chinese).
- [9] DITLEVSEN O, MADSEN H O. Structural reliability methods [M]. New York: John Wiley & Sons, 1996.
- [10] ZHAO Y G, ONO T. Moment method for structural reliability [J]. Structural Safety, 2001, 23 (6): 47-75.
- [11] NAESS A, LEIRA B J, BATSEVYCH O. Reliability analysis of large structural systems[J]. Probabilistic Engineering Mechanics, 2012, 28 (3): 164-168.
- [12] XIE L Y, ZHOU J Y, HAO C Z. System-level load-strength interference based reliability modeling of k -out-of- n system[J]. Reliability Engineering & System Safety, 2004, 84 (3): 311-317.
- [13] 谢里阳, 王正. 随机恒幅循环载荷疲劳可靠度异量纲干涉模型[J]. 机械工程学报, 2008, 44 (1): 1-6.
- [14] XIE L Y, WANG Z. Dissimilar-dimension interference model of fatigue reliability under uncertain cyclic load [J]. Journal of Mechanical Engineering, 2008, 44 (1): 1-6 (in Chinese).
- [15] 谢里阳, 刘建中, 吴宁祥, 等. 风电装备传动系统及零部件疲劳可靠度评估方法[J]. 机械工程学报, 2014, 50 (11): 1-8.
- [16] XIE L Y, LIU J Z, WU N X. Fatigue reliability evaluation method for gear component and system of wind turbine[J]. Journal of Mechanical Engineering, 2014, 50 (11): 1-8 (in Chinese).
- [17] OYAGUE F. Gearbox reliability collaborative (GRC) description and loading; NREL/TP-5000-47773 [R]. Golden: Office of Scientific & Technical Information, 2011: 9.
- [18] 朱孝录. 调质钢齿轮接触疲劳强度可靠度试验研究[J]. 齿轮, 1983, 7 (3): 1-9.
- [19] ZHU X L. The study of quenched and tempered steel gear contact fatigue strength reliability tests[J]. Gear, 1983, 7 (3): 1-9 (in Chinese).
- [20] 中华人民共和国质量监督检验检疫总局. 金属材料 疲劳试验 数据统计方案与分析方法: GB/T 24176—2009 [S]. 北京: 中国标准出版社, 2009.
- [21] General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China. Metallic materials-fatigue testing-statistical planning and analysis of data: GB/T 24176—2009 [S]. Beijing: Standards Press of China, 2009 (in Chinese).

- [17] GARDNER E D. Reliability of components subject to cumulative fatigue[D]. Arizona:University of Arizona,1971:35-45.
- [18] CHEN D. New approaches to the estimation of cumulative fatigue reliability[J]. Reliability Engineering and System Safety, 1991,33(2):231-247.

作者简介:

马洪义 男,博士研究生。主要研究方向:机械系统可靠度评估。

谢里阳 男,博士,教授,博士生导师。主要研究方向:机械系统可靠度评估、疲劳与断裂失效分析。

Assessment of gearbox fatigue reliability based on system PSN curve

MA Hongyi, XIE Liyang*

(Institute of Modern Design and Analysis, Northeast University, Shenyang 110819, China)

Abstract: To improve the efficiency of Monte Carlo simulation applied to evaluating the common cause failure probability of complex mechanical system, the concept of system fatigue life PSN curve was set up first, and then a Monte Carlo method for system reliability assessment based on this concept was proposed. With a given constant load, based on the correspondence of probability percentiles between specimen fatigue lives associated with different cyclic stress levels, every single PSN curve of component can be extracted stochastically. According to the linear cumulative damage rule and the corresponding system reliability model, the fatigue life distribution of gear-series-system was acquired. System PSN curve can be obtained by fitting constant load and life distribution. The system can be treated as a component, and the life analysis process of component-system-component has been completed. By means of damage equivalence principle between random load and constant load, the problem of reliability assessment of a complicated series system under random load can be converted to the problem of reliability assessment of component under constant load.

Keywords: fatigue reliability; system PSN curve; series system; common cause failure; Monte Carlo

Received: 2017-06-05; **Accepted:** 2017-07-07; **Published online:** 2017-09-15 17:09

URL: kns.cnki.net/kcms/detail/11.2625.V.20170915.1709.003.html

Foundation item: National Natural Science Foundation of China (51335003)

* **Corresponding author.** E-mail: lyxie@mail.neu.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2017.0312

基于自适应强跟踪 CQKF 的目标跟踪算法

刘畅¹, 杨锁昌^{1,*}, 汪连栋², 张宽桥¹

(1. 军械工程学院 导弹工程系, 石家庄 050000; 2. 电子信息系统复杂电磁环境效应国家重点实验室, 洛阳 471003)

摘 要: 针对容积积分卡尔曼滤波(CQKF)受模型不确定性影响较大及需要精确已知噪声统计特性的缺点,提出了一种自适应强跟踪 CQKF 算法。该算法根据强跟踪滤波原理,引入渐消因子调整状态预测协方差矩阵,强迫残差序列正交,有效抑制了模型不确定性引起的滤波发散。在滤波过程中,利用 Sage-Husa 时变噪声统计估值器对过程噪声及量测噪声实时估计,提高了算法在未知时变噪声环境下的滤波精度。目标跟踪仿真实验验证了算法的有效性和鲁棒性。

关键词: 目标跟踪; 容积积分卡尔曼滤波(CQKF); 强跟踪滤波; 噪声统计估值器; 自适应滤波

中图分类号: TN953

文献标识码: A

文章编号: 1001-5965(2018)05-0982-09

目标跟踪是一类典型的非线性滤波问题,贝叶斯估计理论为非线性滤波提供了严谨的解决框架。对于线性高斯系统,贝叶斯估计的最优解为卡尔曼滤波^[1]。但对于非线性高斯系统,很难得到高斯加权积分的解析解,因此科研人员提出了许多次优非线性滤波方法。其中,扩展卡尔曼滤波(Extended Kalman Filter, EKF)^[2]由于简单高效得到广泛应用,但 EKF 采用非线性函数的一阶泰勒展开近似非线性函数,对于强非线性系统会产生较大的近似误差,且其需要计算雅可比矩阵,这既增加了计算复杂度也要求非线性函数连续可微。

为了克服 EKF 的缺点,随机采样型滤波器和确定采样型滤波器分别被提出,其核心是用一组随机或确定加权采样点逼近状态的后验分布,以采样点的加权和近似“非线性函数×高斯函数”的积分。随机采样型滤波器的主要应用形式——粒子滤波(Particle Filter, PF)^[3],以蒙特卡罗随机采样得到的粒子近似状态后验分布,理论上适用于解决任意非线性滤波问题,但存在权值退化、粒

子多样性匮乏、实时性差等严重缺陷。依据采样点选取策略的不同,确定采样型滤波器主要分为中心差分卡尔曼滤波(Central Difference Kalman Filter, CDKF)^[4]、无迹卡尔曼滤波(Unscented Kalman Filter, UKF)^[5]和容积卡尔曼滤波(Cubature Kalman Filter, CKF)^[6]等。CDKF 以多项式插值拟合逼近状态后验分布,计算简单但滤波精度较低且易受参数取值影响。UKF 以 UT(Unscented Transformation)变换逼近状态后验分布,能够以二阶泰勒精度逼近非线性状态而计算量与 EKF 同阶,但系统状态维数大于 3 时,中心采样点权值为负,从而引起协方差矩阵的非正定和滤波发散。CKF 以三阶球面-径向容积准则逼近状态后验分布,其采样点权值始终为正,且有严格的数学证明,与 UKF 相比滤波精度更高而计算量相近,因此,近年来得到了广泛的应用。

为进一步提高 CKF 的精度,容积积分卡尔曼滤波(Cubature Quadrature Kalman Filter, CQKF)^[7]以容积准则和高斯-拉盖尔积分准则近似高斯加

收稿日期: 2017-05-15; 录用日期: 2017-06-26; 网络出版时间: 2017-10-13 14:05

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20171013.1405.004.html

* 通信作者. E-mail: yangsuochang_jx@sina.com

引用格式: 刘畅, 杨锁昌, 汪连栋, 等. 基于自适应强跟踪 CQKF 的目标跟踪算法[J]. 北京航空航天大学学报, 2018, 44(5): 982-990. LIU C, YANG S C, WANG L D, et al. Target tracking algorithm based on adaptive strong tracking CQKF[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 982-990 (in Chinese).

权积分,滤波精度与切比雪夫-拉盖尔多项式的阶数 m 成正相关关系。CQKF 是 CKF 的拓展形式,当 $m \geq 2$ 时,其精度高于 CKF。但 CQKF 要求系统模型和噪声精确已知,而现代战场复杂对抗环境使得系统模型具有很大的不确定性,系统状态突变,过程噪声及量测噪声未知时变,从而造成 CQKF 滤波精度下降甚至发散。为解决噪声统计特性未知情况下的滤波问题,贝叶斯法^[8]、相关法^[9-11]、协方差匹配法^[12]和极大似然估计法^[13]等多种噪声估计算法分别被提出。贝叶斯法涉及到多重积分,计算量大且未必能得到最优封闭解,多限于理论研究。相关法主要应用于线性系统,其非线性系统中的拓展尚未无偏性证明^[14]。协方差匹配法是一种有偏估计方法,存在稳态估计误差且估计精度相对较低。极大似然估计法由于能够直接构造含有待估计参数的概率密度函数,且计算量适中而得到广泛关注,其中以基于极大后验估计(Maximum A Posterior, MAP)准则提出的 Sage-Husa 时变噪声统计估值器^[15]应用最为广泛。

基于上述考虑,本文提出了一种自适应强跟踪 CQKF(AST-CQKF)算法。AST-CQKF 借鉴强跟踪滤波(Strong Tracking Filter, STF)的思想,引入时变渐消因子对预测协方差矩阵在线调整,同时利用 Sage-Husa 时变噪声统计估值器对噪声统计特性实时估计,并将估计结果引入滤波迭代过程中,从而增加了 CQKF 算法的鲁棒性和自适应性。通过仿真实验验证了本文算法的有效性和可行性。

1 容积积分卡尔曼滤波

1.1 贝叶斯滤波框架

考虑如下非线性系统:

$$\begin{cases} \mathbf{x}_k = f(\mathbf{x}_{k-1}) + \mathbf{v}_{k-1} \\ \mathbf{z}_k = h(\mathbf{x}_k) + \boldsymbol{\eta}_k \end{cases} \quad (1)$$

式中: $\mathbf{x}_k \in \mathbf{R}^n$ 为 k 时刻的状态向量; $\mathbf{z}_k \in \mathbf{R}^p$ 为 k 时刻的量测向量; $f: \mathbf{R}^n \rightarrow \mathbf{R}^n$ 和 $h: \mathbf{R}^n \rightarrow \mathbf{R}^p$ 为已知非线性函数; $\mathbf{v}_{k-1} \in \mathbf{R}^n$ 和 $\boldsymbol{\eta}_k \in \mathbf{R}^p$ 为不相关的高斯白噪声,且 $\mathbf{v}_{k-1} \sim N(0, \mathbf{Q}_{k-1})$, $\boldsymbol{\eta}_k \sim N(0, \mathbf{R}_k)$; \mathbf{Q}_{k-1} 和 \mathbf{R}_k 为噪声协方差矩阵。

假设 $k-1$ 时刻的后验概率密度函数为高斯分布,即

$$p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) = N(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) \quad (2)$$

$$p(\mathbf{z}_k | \mathbf{z}_{1:k-1}) = N(\mathbf{z}_k; \hat{\mathbf{z}}_{k|k-1}, \mathbf{P}_{k|k-1}^{zz}) \quad (3)$$

其中状态一步预测及协方差为

$$\hat{\mathbf{x}}_{k|k-1} = \int f(\mathbf{x}_{k-1}) N(\mathbf{x}_{k-1}; \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1}) d\mathbf{x}_{k-1} \quad (4)$$

$$\mathbf{P}_{k|k-1} = \int f(\mathbf{x}_{k-1}) f^T(\mathbf{x}_{k-1}) N(\mathbf{x}_{k-1}; \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1}) d\mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k|k-1} \hat{\mathbf{x}}_{k-1|k-1}^T + \mathbf{Q}_{k-1} \quad (5)$$

输出预测、自协方差及互协方差分别为

$$\hat{\mathbf{z}}_{k|k-1} = \int h(\mathbf{x}_k) N(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) d\mathbf{x}_k \quad (6)$$

$$\mathbf{P}_{k|k-1}^{zz} = \int h(\mathbf{x}_k) h^T(\mathbf{x}_k) N(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) d\mathbf{x}_k - \hat{\mathbf{z}}_{k|k-1} \hat{\mathbf{z}}_{k|k-1}^T + \mathbf{R}_k \quad (7)$$

$$\mathbf{P}_{k|k-1}^{xz} = \int \mathbf{x}_k h^T(\mathbf{x}_k) N(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) d\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1} \hat{\mathbf{z}}_{k|k-1}^T \quad (8)$$

非线性系统式(1)在最小方差估计准则下的最优高斯滤波器为^[16]

$$p(\mathbf{x}_k | \mathbf{z}_{1:k}) = N(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}) \quad (9)$$

式中:

$$\begin{cases} \hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}) \\ \mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{P}_{k|k-1}^{zz} \mathbf{K}_k^T \end{cases} \quad (10)$$

其中: $\mathbf{K}_k = \mathbf{P}_{k|k-1}^{xz} \mathbf{P}_{k|k-1}^{zz-1}$

因此,贝叶斯滤波的核心问题转化为计算形式如式(11)的积分表达式:

$$I(f) = \int f(\mathbf{x}) g(\mathbf{x}) d\mathbf{x} \quad (11)$$

式中: $f(\mathbf{x})$ 为任意非线性函数; $g(\mathbf{x})$ 为高斯密度函数。通常式(11)难以得到解析表达式,主要采取一系列采样点 ξ_j 及其权重 w_j 的加权和进行数值近似,即

$$I(f) \approx \sum_{j=1}^m w_j f(\xi_j) \quad (12)$$

1.2 容积积分准则

对于任意函数 $f(\mathbf{x})$, $\mathbf{X} \in \mathbf{R}^n$, 式(11)的一般形式为

$$I(f) = \frac{1}{\sqrt{|\boldsymbol{\Sigma}|} (2\pi)^n} \int_{\mathbf{R}^n} f(\mathbf{x}) e^{-(1/2)(\mathbf{X}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{X}-\boldsymbol{\mu})} d\mathbf{X} \quad (13)$$

可以在球面坐标系中表示为^[7]

$$I(f) = \frac{1}{\sqrt{(2\pi)^n}} \int_{r=0}^{\infty} \int_{U_n} [f(\mathbf{C}r\mathbf{Z} + \boldsymbol{\mu})] r^{n-1} e^{-r^2/2} dr \quad (14)$$

式中: $\mathbf{X} = \mathbf{C}r\mathbf{Z} + \boldsymbol{\mu}$, 且 $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\mu}$ 为高斯分布的均值, \mathbf{C} 为协方差矩阵 $\boldsymbol{\Sigma}$ 的 Cholesky 分解, 即 $\boldsymbol{\Sigma} = \mathbf{C}\mathbf{C}^T$; r 和 \mathbf{Z} 为积分变量; $U_n = \{\mathbf{Z} \in \mathbf{R}^n | \mathbf{Z}\mathbf{Z}^T = \mathbf{I}\}$ 为单位超球面; $ds(\cdot)$ 为 U_n 上的面积元素。

若 $\boldsymbol{\mu} = \mathbf{0}$ 且 $\boldsymbol{\Sigma}$ 为单位矩阵, 则式(14)中的积分为

$$\int_{U_n} f(\mathbf{C}r\mathbf{Z} + \boldsymbol{\mu}) ds(\mathbf{Z}) \quad (15)$$

由三阶球面-径向容积准则可以近似为

$$\int_{U_n} f(r\mathbf{Z} + \boldsymbol{\mu}) ds(\mathbf{Z}) \approx \frac{2\sqrt{\pi^n}}{2n\Gamma(n/2)} \sum_{i=1}^{2n} f[ru_i] \quad (16)$$

式中: $\Gamma(\cdot)$ 为伽马函数; $[u_i]$ ($i=1, 2, \dots, 2n$) 为位于单位超球面与坐标轴交点的容积点, 即完全对称点集[1]的第 i 个列向量

$$[1] = \left\{ \underbrace{\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_n, \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_n, \dots, \underbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}}_n, \underbrace{\begin{bmatrix} -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_n, \underbrace{\begin{bmatrix} 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_n, \dots, \underbrace{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ -1 \end{bmatrix}}_n \right\}_{n \times 2n} \quad (17)$$

将式(16)代入式(14), 令 $\lambda = r^2/2$, 则

$$I(f) = \frac{1}{2n\Gamma(n/2)} \int_{\lambda=0}^{+\infty} \left(\sum_{i=1}^{2n} g(\sqrt{2\lambda}u_i) \right) \lambda^{(n/2-1)} e^{-\lambda} d\lambda \quad (18)$$

根据高斯-拉盖尔积分准则, 积分可以近似为

$$\int_{\lambda=0}^{+\infty} f(\lambda) \lambda^\alpha e^{-\lambda} d\lambda \approx \sum_{j=1}^m A_j \lambda_j \quad (19)$$

式中: 积分点 λ_j ($j=1, 2, \dots, m$) 为 m 阶切比雪夫-拉盖尔多项式的 m 个根, 即 λ_j 满足

$$L_m^\alpha(\lambda) = \lambda^m - \frac{m}{1!}(m+\alpha)\lambda^{m-1} + \frac{m(m-1)}{2!} \dots - \dots = 0 \quad (20)$$

式中: $\alpha = n/2 - 1$ 。

对应的权值 A_j 为

$$A_j = \frac{m!\Gamma(\alpha+m+1)}{\lambda_j [L_m^\alpha(\lambda_j)]^2} \quad (21)$$

将式(19)~式(21)代入式(18), 则得到容积积分准则为

$$I(f) = \frac{1}{2n\Gamma(n/2)} \left[\sum_{i=1}^{2n} \sum_{j=1}^m A_j f(\sqrt{2\lambda_j}) [u_i] \right] \quad (22)$$

由式(22)可知, 对于如式(1)所示的 n 维状态估计问题, 需要计算满足

$$I(f) \approx \sum_{l=1}^{2mn} w_l f(\xi_l) \quad (23)$$

的 $2mn$ 个采样点 ξ_l (本文称其为 CQ 点) 及对应权值 w_l , 其计算方式如下:

$$\xi_l = \sqrt{2\lambda_j} [u_i] \quad i=1, 2, \dots, 2n; j=1, 2, \dots, m; l=1, 2, \dots, 2mn \quad (24)$$

$$w_l = \frac{1}{2n\Gamma(n/2)} \cdot \frac{m!\Gamma(\alpha+m+1)}{\lambda_j [L_m^\alpha(\lambda_j)]^2} \quad (25)$$

1.3 CQKF 算法

步骤 1 初始化。

1) 令

$$\hat{\mathbf{x}}_{0|0} = E(\mathbf{x}_0) \\ \mathbf{P}_{0|0} = E[(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|0})(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|0})^T]$$

2) 根据式(24)和式(25)计算 CQ 点 ξ_l 及对应权值 w_l 。

步骤 2 预测更新。

1) 假设 $p(\mathbf{x}_{k-1} | \mathbf{z}_{k-1}) = N(\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1})$ 已知(由上一时刻估计得到), 分解协方差矩阵:

$$\mathbf{P}_{k-1|k-1} = \mathbf{S}_{k-1|k-1} \mathbf{S}_{k-1|k-1}^T$$

2) 计算 CQ 点:

$$\boldsymbol{\chi}_{l,k-1|k-1} = \mathbf{S}_{k-1|k-1} \xi_l + \hat{\mathbf{x}}_{k-1|k-1}$$

3) 更新 CQ 点随状态方程的转移:

$$\boldsymbol{\chi}_{l,k|k-1} = f(\boldsymbol{\chi}_{l,k-1|k-1})$$

4) 计算一步状态预测及预测误差协方差矩阵:

$$\hat{\mathbf{x}}_{k|k-1} = \sum_{l=1}^{2mn} w_l \boldsymbol{\chi}_{l,k|k-1} \\ \mathbf{P}_{k|k-1} = \sum_{l=1}^{2mn} w_l [\boldsymbol{\chi}_{l,k|k-1} - \hat{\mathbf{x}}_{k|k-1}] [\boldsymbol{\chi}_{l,k|k-1} - \hat{\mathbf{x}}_{k|k-1}]^T + \mathbf{Q}_{k-1}$$

步骤 3 量测更新。

1) 分解协方差矩阵:

$$\mathbf{P}_{k|k-1} = \mathbf{S}_{k|k-1} \mathbf{S}_{k|k-1}^T$$

2) 计算 CQ 点:

$$\boldsymbol{\chi}_{l,k|k-1} = \mathbf{S}_{k|k-1} \xi_l + \hat{\mathbf{x}}_{k|k-1}$$

3) 更新 CQ 点随量测方程的转移:

$$\mathbf{Z}_{l,k|k-1} = h(\boldsymbol{\chi}_{l,k|k-1})$$

4) 计算量测预测值:

$$\hat{\mathbf{z}}_{k|k-1} = \sum_{l=1}^{2mn} w_l \mathbf{Z}_{l,k|k-1}$$

5) 计算自协方差矩阵及互协方差矩阵:

$$\mathbf{P}_{k|k-1}^{zz} = \sum_{l=1}^{2mn} w_l [\mathbf{Z}_{l,k|k-1} - \hat{\mathbf{z}}_{k|k-1}] [\mathbf{Z}_{l,k|k-1} - \hat{\mathbf{z}}_{k|k-1}]^T + \mathbf{R}_k$$

$$\mathbf{P}_{k|k-1}^{xz} = \sum_{l=1}^{2mn} w_l [\boldsymbol{\chi}_{l,k|k-1} - \hat{\mathbf{x}}_{k|k-1}] [\mathbf{Z}_{l,k|k-1} - \hat{\mathbf{z}}_{k|k-1}]^T$$

6) 计算卡尔曼增益:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}^{xz} (\mathbf{P}_{k|k-1}^{zz})^{-1}$$

7) 估计状态:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})$$

8) 估计状态预测误差协方差矩阵:

$$P_{k|k} = P_{k|k-1} - K_k P_{k|k-1}^{zz} K_k^T$$

2 强跟踪滤波理论

为了提高 EKF 对于系统模型不确定性及状态突变的鲁棒性,周东华等^[17]提出的 STF 利用衰减记忆滤波思想,在计算预测误差协方差矩阵时引入渐消因子,强迫残差序列正交,从而保证滤波器对系统实际状态的跟踪效果。

非线性系统式(1)的 STF 方程为^[18]

$$\begin{cases} \hat{x}_{k|k-1} = f(\hat{x}_{k-1|k-1}) \\ P_{k|k-1} = \lambda_k F_{k|k-1} P_{k-1|k-1} F_{k|k-1}^T + Q_{k-1} \\ \hat{z}_{k|k-1} = h(\hat{x}_{k|k-1}) \\ K_k = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1} \\ \hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k (z_k - \hat{z}_{k|k-1}) \\ P_{k|k} = (I_{n \times n} - K_k H_k) P_{k|k-1} \end{cases} \quad (26)$$

式中:

$$\begin{cases} F_{k|k-1} = \left. \frac{\partial f(x_{k-1})}{\partial x_{k-1}} \right|_{x_{k-1} = \hat{x}_{k-1|k-1}} \\ H_k = \left. \frac{\partial h(x_k)}{\partial x_k} \right|_{x_k = \hat{x}_{k|k-1}} \end{cases}$$

$I_{n \times n}$ 为 n 阶单位矩阵; λ_k 为渐消因子,其计算式为

$$\lambda_k = \max\left(1, \frac{\text{tr}[N_k]}{\text{tr}[M_k]}\right) \quad (27)$$

式中:

$$N_k = V_k - H_k Q_{k-1} H_k^T - \beta R_k \quad (28)$$

$$M_k = H_k F_{k|k-1} P_{k-1|k-1} F_{k|k-1}^T H_k^T \quad (29)$$

$$V_k = \begin{cases} \varepsilon_1 \varepsilon_1^T & k = 1 \\ \frac{\rho V_{k-1} + \varepsilon_k \varepsilon_k^T}{1 + \rho} & k \geq 2 \end{cases} \quad (30)$$

$$\varepsilon_k = z_k - \hat{z}_{k|k-1} \quad (31)$$

其中: $\text{tr}(\cdot)$ 为矩阵求迹运算; $\beta \geq 1$ 为弱化因子; $0 < \rho \leq 1$ 为遗忘因子,通常取 $\rho = 0.95$ 。由于 STF 是 EKF 的改进形式,在计算渐消因子时仍需要求解雅可比矩阵,因此不能直接将其引入 CQKF,需要研究不利用 H_k 计算渐消因子的等价表述方式。文献[19]给出了 H_k 、 N_k 和 M_k 的等效表述形式,即

$$H_k = (P_{k|k-1}^{zz(e)})^T [(P_{k|k-1}^{(e)})^{-1}]^T \quad (32)$$

$$N_k = V_k - (P_{k|k-1}^{zz(e)})^T [(P_{k|k-1}^{(e)})^{-1}]^T \cdot$$

$$Q_{k-1} (P_{k|k-1}^{(e)})^{-1} P_{k|k-1}^{zz(e)} - \beta R_k \quad (33)$$

$$M_k = P_{k|k-1}^{zz(e)} + N_k - V_k + (\beta - 1) R_k \quad (34)$$

其中: $P_{k|k-1}^{(e)}$ 、 $P_{k|k-1}^{zz(e)}$ 和 $P_{k|k-1}^{zz(e)}$ 分别为未引入渐消

因子的一步状态预测误差协方差阵、输出预测互协方差阵和输出预测自协方差阵。用式(33)和式(34)替换式(28)和式(29),即得到强跟踪 CQKF 中渐消因子的计算方式,即

$$\begin{cases} \lambda_k = \max\left(1, \frac{\text{tr}[N_k]}{\text{tr}[M_k]}\right) \\ N_k = V_k - (P_{k|k-1}^{zz(e)})^T [(P_{k|k-1}^{(e)})^{-1}]^T \cdot \\ \quad Q_{k-1} (P_{k|k-1}^{(e)})^{-1} P_{k|k-1}^{zz(e)} - \beta R_k \\ M_k = P_{k|k-1}^{zz(e)} + N_k - V_k + (\beta - 1) R_k \\ V_k = \begin{cases} \varepsilon_1 \varepsilon_1^T & k = 1 \\ \frac{\rho V_{k-1} + \varepsilon_k \varepsilon_k^T}{1 + \rho} & k \geq 2 \end{cases} \\ \varepsilon_k = z_k - \hat{z}_{k|k-1} \end{cases} \quad (35)$$

渐消因子的具体引入方式见第4节。

3 时变噪声统计估值器

考虑噪声的一般形式,即 $v_{k-1} \in \mathbf{R}^n$ 和 $\eta_k \in \mathbf{R}^p$ 是带时变均值和协方差且线性无关的高斯白噪声,且 $v_{k-1} \sim N(q_{k-1}, Q_{k-1})$, $\eta_k \sim N(r_k, R_k)$ 。针对 q_k 、 Q_k 、 r_k 和 R_k 等噪声参数的估计问题,文献[15]基于 MAP 准则得到了 Sage-Husa 噪声统计估值器,并给出了最优 MAP 估值器、次优 MAP 估值器、次优无偏 MAP 估值器和时变噪声统计估值器等多种形式。文献[20]和文献[21]分别将其拓展到 UKF 和 CKF,给出了适用于 UKF 和 CKF 的递推算法。由于 UKF、CKF 和 CQKF 同属确定采样型滤波器,借鉴 UKF 和 CKF 的递推形式,得到如下适用于 CQKF 的次优无偏 MAP 常值噪声统计估值器:

$$\hat{q}_k = \frac{1}{k} \left[(k-1) \hat{q}_{k-1} + \hat{x}_{k|k} - \sum_{l=1}^{2mn} w_l f(\chi_{l,k-1|k-1}) \right] \quad (36)$$

$$\hat{Q}_k = \frac{1}{k} \left[(k-1) \hat{Q}_{k-1} + K_k \varepsilon_k \varepsilon_k^T K_k^T + P_{k|k} - \sum_{l=1}^{2mn} w_l [\chi_{l,k|k-1}^* - \hat{x}_{k|k-1}] [\chi_{l,k|k-1}^* - \hat{x}_{k|k-1}]^T \right] \quad (37)$$

$$\hat{r}_k = \frac{1}{k} \left[(k-1) \hat{r}_{k-1} + z_k - \sum_{l=1}^{2mn} w_l h(\chi_{l,k|k-1}) \right] \quad (38)$$

$$\hat{R}_k = \frac{1}{k} \left[(k-1) \hat{R}_{k-1} + \varepsilon_k \varepsilon_k^T - \sum_{l=1}^{2mn} w_l [Z_{l,k|k-1} - \hat{z}_{k|k-1}] [Z_{l,k|k-1} - \hat{z}_{k|k-1}]^T \right] \quad (39)$$

对于时变噪声统计而言,应当强调新近数据的作用,逐渐遗忘陈旧数据。选取加权系数

$\{\gamma_i\}$,使之满足

$$\gamma_i = \gamma_{i-1} b \quad 0 < b < 1, \sum_{i=1}^k \gamma_i = 1 \quad (40)$$

于是有

$$\begin{cases} \gamma_i = d_k b^{i-1} & i = 1, 2, \dots, k \\ d_k = \frac{1-b}{1-b^k} \end{cases} \quad (41)$$

式中: b 为遗忘因子。将次优无偏 MAP 常值噪声统计估计器中的加权和系数 $1/k$ 替换为 $\{\gamma_i\}$, 即得到适用于 CQKF 的时变噪声统计估值器:

$$\hat{\mathbf{q}}_k = (1 - d_k) \hat{\mathbf{q}}_{k-1} + d_k \left[\hat{\mathbf{x}}_{k|k} - \sum_{l=1}^{2mn} w_l f(\boldsymbol{\chi}_{l,k-1|k-1}) \right] \quad (42)$$

$$\hat{\mathbf{Q}}_k = (1 - d_k) \hat{\mathbf{Q}}_{k-1} + d_k \left[\mathbf{K}_k \boldsymbol{\varepsilon}_k \boldsymbol{\varepsilon}_k^T \mathbf{K}_k^T + \mathbf{P}_{k|k} - \sum_{l=1}^{2mn} w_l [\boldsymbol{\chi}_{l,k|k-1}^* - \hat{\mathbf{x}}_{k|k-1}] [\boldsymbol{\chi}_{l,k|k-1}^* - \hat{\mathbf{x}}_{k|k-1}]^T \right] \quad (43)$$

$$\hat{\mathbf{r}}_k = (1 - d_k) \hat{\mathbf{r}}_{k-1} + d_k \left[\mathbf{z}_k - \sum_{l=1}^{2mn} w_l h(\boldsymbol{\chi}_{l,k|k-1}) \right] \quad (44)$$

$$\hat{\mathbf{R}}_k = (1 - d_k) \hat{\mathbf{R}}_{k-1} + d_k \left[\boldsymbol{\varepsilon}_k \boldsymbol{\varepsilon}_k^T - \sum_{l=1}^{2mn} w_l [\mathbf{Z}_{l,k|k-1} - \hat{\mathbf{z}}_{k|k-1}] [\mathbf{Z}_{l,k|k-1} - \hat{\mathbf{z}}_{k|k-1}]^T \right] \quad (45)$$

4 自适应强跟踪 CQKF

将时变渐消因子和时变噪声统计估值器嵌入标准 CQKF, 并将其拓展到噪声均值非零的情形, 即可得到 AST-CQKF 算法, 流程如图 1 所示。

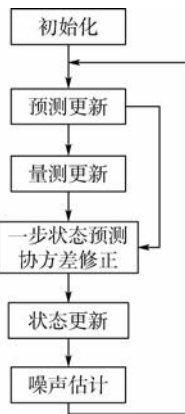


图 1 AST-CQKF 算法流程图

Fig. 1 AST-CQKF algorithm flowchart

算法具体流程如下:

步骤 1 初始化。

1) 令

$$\begin{aligned} \hat{\mathbf{x}}_{0|0} &= E(\mathbf{x}_0) \\ \mathbf{P}_{0|0} &= E[(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|0})(\mathbf{x}_0 - \hat{\mathbf{x}}_{0|0})^T] \end{aligned}$$

$$\begin{aligned} \hat{\mathbf{q}}_0 &= \mathbf{q}_0 \\ \hat{\mathbf{Q}}_0 &= \mathbf{Q}_0 \\ \hat{\mathbf{r}}_0 &= \mathbf{r}_0 \\ \hat{\mathbf{R}}_0 &= \mathbf{R}_0 \end{aligned}$$

2) 根据式(24)和式(25)计算 CQ 点 ξ_l 及对应权值 w_l 。

步骤 2 预测更新。

1) 假设 $p(\mathbf{x}_{k-1} | \mathbf{z}_{k-1}) = N(\hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1})$ 已知(由上一时刻估计得到), 分解协方差矩阵:

$$\mathbf{P}_{k-1|k-1} = \mathbf{S}_{k-1|k-1} \mathbf{S}_{k-1|k-1}^T$$

2) 计算 CQ 点:

$$\boldsymbol{\chi}_{l,k-1|k-1} = \mathbf{S}_{k-1|k-1} \boldsymbol{\xi}_l + \hat{\mathbf{x}}_{k-1|k-1}$$

3) 计算 CQ 点随状态方程的转移:

$$\boldsymbol{\chi}_{l,k|k-1}^* = f(\boldsymbol{\chi}_{l,k-1|k-1}) + \hat{\mathbf{q}}_{k-1}$$

4) 计算一步状态预测及预测误差协方差矩阵:

$$\hat{\mathbf{x}}_{k|k-1} = \sum_{l=1}^{2mn} w_l \boldsymbol{\chi}_{l,k|k-1}^* + \hat{\mathbf{q}}_{k-1}$$

$$\mathbf{P}_{k|k-1}^{(e)} = \sum_{l=1}^{2mn} w_l [\boldsymbol{\chi}_{l,k|k-1}^* - \hat{\mathbf{x}}_{k|k-1}] [\boldsymbol{\chi}_{l,k|k-1}^* - \hat{\mathbf{x}}_{k|k-1}]^T + \hat{\mathbf{Q}}_{k-1}$$

步骤 3 量测更新。

1) 分解协方差矩阵:

$$\mathbf{P}_{k|k-1}^{(e)} = \mathbf{S}_{k|k-1}^{(e)} (\mathbf{S}_{k|k-1}^{(e)})^T$$

2) 计算 CQ 点:

$$\boldsymbol{\chi}_{l,k|k-1}^{(e)} = \mathbf{S}_{k|k-1}^{(e)} \boldsymbol{\xi}_l + \hat{\mathbf{x}}_{k|k-1}$$

3) 计算 CQ 点随量测方程的转移:

$$\mathbf{Z}_{l,k|k-1}^{(e)} = h(\boldsymbol{\chi}_{l,k|k-1}^{(e)}) + \hat{\mathbf{r}}_k$$

4) 计算量测预测值:

$$\hat{\mathbf{z}}_{k|k-1}^{(e)} = \sum_{l=1}^{2mn} w_l \mathbf{Z}_{l,k|k-1}^{(e)} + \hat{\mathbf{r}}_k$$

5) 计算自协方差矩阵和互协方差矩阵:

$$\begin{aligned} \mathbf{P}_{k|k}^{2z(e)} &= \sum_{l=1}^{2mn} w_l [\mathbf{Z}_{l,k|k-1}^{(e)} - \hat{\mathbf{z}}_{k|k-1}^{(e)}] [\mathbf{Z}_{l,k|k-1}^{(e)} - \hat{\mathbf{z}}_{k|k-1}^{(e)}]^T + \hat{\mathbf{R}}_k \\ \hat{\mathbf{z}}_{k|k-1}^{(e)} &= \sum_{l=1}^{2mn} w_l [\boldsymbol{\chi}_{l,k|k-1}^{(e)} - \hat{\mathbf{x}}_{k|k-1}] [\mathbf{Z}_{l,k|k-1}^{(e)} - \hat{\mathbf{z}}_{k|k-1}^{(e)}]^T \end{aligned}$$

步骤 4 一步状态预测协方差修正。

1) 根据式(35)计算渐消因子 λ_k 。

2) 利用 λ_k 修正一步状态预测协方差矩阵:

$$\begin{aligned} \mathbf{P}_{k|k-1} &= \lambda_k \sum_{l=1}^{2mn} w_l [\boldsymbol{\chi}_{l,k|k-1} - \hat{\mathbf{x}}_{k|k-1}] [\boldsymbol{\chi}_{l,k|k-1} - \hat{\mathbf{x}}_{k|k-1}]^T + \hat{\mathbf{Q}}_{k-1} \end{aligned}$$

3) 分解修正后的协方差矩阵:

$$\mathbf{P}_{k|k-1} = \mathbf{S}_{k|k-1} \mathbf{S}_{k|k-1}^T$$

4) 计算 CQ 点:

$$\mathbf{X}_{l,k|k-1} = \mathbf{S}_{k|k-1} \xi_l + \hat{\mathbf{x}}_{k|k-1}$$

5) 计算 CQ 点经量测方程的转移:

$$\mathbf{Z}_{l,k|k-1} = h(\mathbf{X}_{l,k|k-1}) + \hat{\mathbf{r}}_k$$

6) 计算量测预测值:

$$\hat{\mathbf{z}}_{k|k-1} = \sum_{l=1}^{2mn} w_l \mathbf{Z}_{l,k|k-1} + \hat{\mathbf{r}}_k$$

7) 计算自协方差矩阵和互协方差矩阵:

$$\mathbf{P}_{k|k}^{zz} = \sum_{l=1}^{2mn} w_l [\mathbf{Z}_{l,k|k-1} - \hat{\mathbf{z}}_{k|k-1}] [\mathbf{Z}_{l,k|k-1} - \hat{\mathbf{z}}_{k|k-1}]^T + \hat{\mathbf{R}}_k$$

$$\mathbf{P}_{k|k}^{xz} = \sum_{l=1}^{2mn} w_l [\mathbf{X}_{l,k|k-1} - \hat{\mathbf{x}}_{k|k-1}] [\mathbf{Z}_{l,k|k-1} - \hat{\mathbf{z}}_{k|k-1}]^T$$

步骤 5 状态更新。

1) 计算卡尔曼增益:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}^{xz} (\mathbf{P}_{k|k-1}^{zz})^{-1}$$

2) 估计状态:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1})$$

3) 估计状态预测误差协方差矩阵:

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k \mathbf{P}_{k|k-1} \mathbf{K}_k^T$$

步骤 6 噪声估计。

根据式(42)~式(45)对噪声统计特性递推估计。

5 仿真验证

考虑一个典型的二维平面雷达跟踪问题^[7],目标以固定未知转弯速率 Ω 做圆周运动,其状态方程为

$$\mathbf{X}_k = \begin{pmatrix} 1 & \frac{\sin(\Omega T)}{\Omega} & 0 & -\left(\frac{1 - \cos(\Omega T)}{\Omega}\right) & 0 \\ 0 & \cos(\Omega T) & 0 & -\sin(\Omega T) & 0 \\ 0 & \frac{1 - \cos(\Omega T)}{\Omega} & 1 & \frac{\sin(\Omega T)}{\Omega} & 0 \\ 0 & \sin(\Omega T) & 0 & \cos(\Omega T) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \mathbf{X}_{k-1} + \mathbf{v}_{k-1} \quad (46)$$

式中: $\mathbf{X}_k = [x_k \quad \dot{x}_k \quad y_k \quad \dot{y}_k \quad \Omega]^T$ 为目标状态向量, x_k 和 y_k 为目标在笛卡儿坐标系中的位置, \dot{x}_k 和 \dot{y}_k 分别为目标在 x 和 y 方向的速度; T 为采样周期; 过程噪声 $\mathbf{v}_{k-1} \sim N(0, \mathbf{Q})$, $\mathbf{Q} = \text{diag}[q_1 \mathbf{M} \quad q_2 T]$, q_1 和 q_2 为与过程噪声强度成正比的比例系数, \mathbf{M} 表达式为

$$\mathbf{M} = \begin{bmatrix} T^3 & T^2 \\ 3 & 2 \\ T^2 & T \end{bmatrix} \quad (47)$$

雷达固定于坐标原点,对目标的距离 r_k 及方位角 θ_k 进行测量,测量方程为

$$\mathbf{Z}_k = \begin{pmatrix} \sqrt{x_k^2 + y_k^2} \\ \arctan\left(\frac{y_k}{x_k}\right) \end{pmatrix} + \boldsymbol{\eta}_k \quad (48)$$

式中: $\mathbf{Z}_k = (r_k \quad \theta_k)^T$ 为量测向量, r_k 为目标与原点的距离, θ_k 为目标和原点的连线与 x 轴所成角度。量测噪声 $\boldsymbol{\eta}_k \sim N(0, \mathbf{R})$, $\mathbf{R} = \text{diag}[\sigma_r^2 \quad \sigma_\theta^2]$ 。

仿真参数设置如下: $T = 1 \text{ s}$, $\Omega = -3 (\text{ }^\circ)/\text{s}$, $q_1 = 0.1 \text{ m}^2/\text{s}^3$, $q_2 = 1.75 \times 10^{-4} \text{ s}^{-3}$, $\beta = 3.5$, $\rho = 0.95$, $b = 0.95$, $\sigma_r = 10 \text{ m}$, $\sigma_\theta = \sqrt{10} \text{ m} \cdot \text{rad}$, $m = 2$ 。

初始状态真实值为

$$\mathbf{X}_0 = [1000 \text{ m} \quad 300 \text{ m/s} \quad 1000 \text{ m} \quad 0 \quad -3 (\text{ }^\circ)/\text{s}]^T$$

对应的协方差矩阵为

$$\mathbf{P}_{0|0} = \text{diag}[100 \text{ m}^2 \quad 10 \text{ m}^2/\text{s}^2 \quad 100 \text{ m}^2 \quad 10 \text{ m}^2/\text{s}^2 \quad 100 \text{ m} \cdot \text{rad}^2/\text{s}^2]^T$$

为保证仿真效果,采取 200 次独立的蒙特卡罗仿真,仿真步数为 100。每次仿真开始前,初始状态的估计值 $\hat{\mathbf{X}}_{0|0}$ 从高斯分布 $N(\mathbf{X}_0, \mathbf{P}_{0|0})$ 中随机选取。选取位置、速度和转弯率的均方根误差 (Root-Mean Square Error, RMSE) 为比较不同算法性能的评估指标。以 k 时刻的 RMSE 为例,其定义分别为

$$\text{RMSE}_{\text{pos}}(k) = \sqrt{\frac{1}{N} \sum_{n=1}^N [(x_k^n - \hat{x}_k^n)^2 + (y_k^n - \hat{y}_k^n)^2]} \quad (49)$$

$$\text{RMSE}_{\text{vel}}(k) = \sqrt{\frac{1}{N} \sum_{n=1}^N [(\dot{x}_k^n - \hat{\dot{x}}_k^n)^2 + (\dot{y}_k^n - \hat{\dot{y}}_k^n)^2]} \quad (50)$$

$$\text{RMSE}_{\text{ome}}(k) = \sqrt{\frac{1}{N} \sum_{n=1}^N (\Omega_k^n - \hat{\Omega}_k^n)^2} \quad (51)$$

式中: (x_k^n, y_k^n) 和 $(\hat{x}_k^n, \hat{y}_k^n)$ 分别为第 n 次仿真中的位置真实值和估计值; $(\dot{x}_k^n, \dot{y}_k^n)$ 和 $(\hat{\dot{x}}_k^n, \hat{\dot{y}}_k^n)$ 分别为第 n 次仿真中的速度真实值和估计值; Ω_k^n 和 $\hat{\Omega}_k^n$ 分别为第 n 次仿真中的转弯率真实值和估计值。

为验证算法在系统状态模型不准确及噪声统计特性不准确下的有效性,设置以下 2 种仿真情形。

5.1 系统量测噪声的统计特性不准确

假设系统过程噪声 $\mathbf{v}_{k-1} \sim N(0, \mathbf{Q})$ 精确已知,而量测噪声的协方差矩阵与真实值不相符,选取其初始协方差矩阵为 $\hat{\mathbf{R}}_0 = \text{diag}[1000\sigma_r^2 \quad 1000\sigma_\theta^2]$ 。仿真过程中,量测真值由量测方程及协方差矩阵 \mathbf{R} 得到,而滤波器使用矩阵 $\hat{\mathbf{R}}_0$ 进行状态估计,以此比较 AST-CQKF 与 CQKF 算法在已知噪声特性与实

际噪声特性存在偏差时的滤波精度。经过 200 次独立的蒙特卡罗仿真,得到位置、速度和转弯率的 RMSE 结果分别如图 2~图 3 所示(情形 1)。

由图 2 可知,通过 AST-CQKF 算法得到的位置、速度均方根误差明显小于 CQKF 算法,而转弯率均方根误差基本相同,可见 AST-CQKF 算法能够有效减小噪声统计特性不准确带来的估计误差,提高估计精度。此外,噪声统计特性的改变对位置、速度、转弯率的影响程度不同,对位置影响最大,速度次之,转弯率几乎无影响。

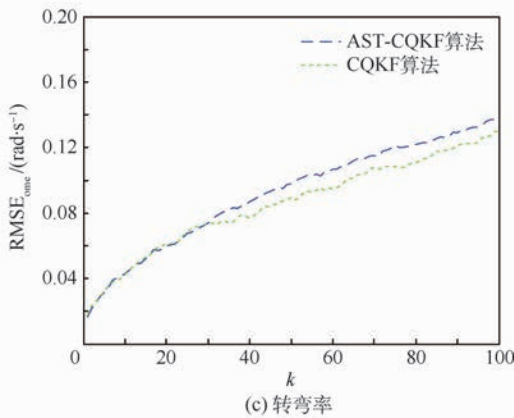
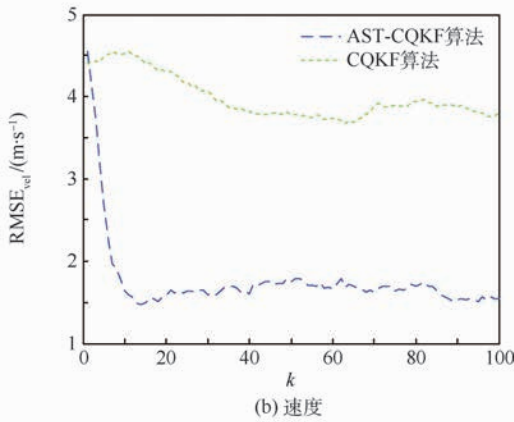
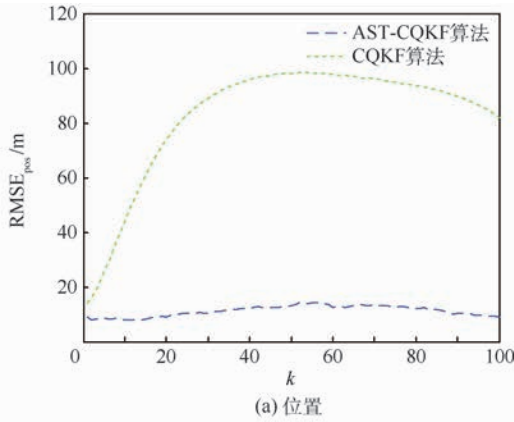


图 2 跟踪均方根误差(情形 1)
Fig.2 RMSE tracking(Case 1)

5.2 系统状态模型不准确

为设置系统模型的不确定性,将式(46)变形得

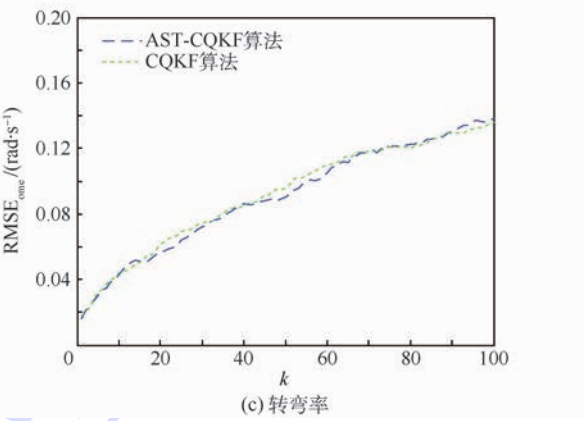
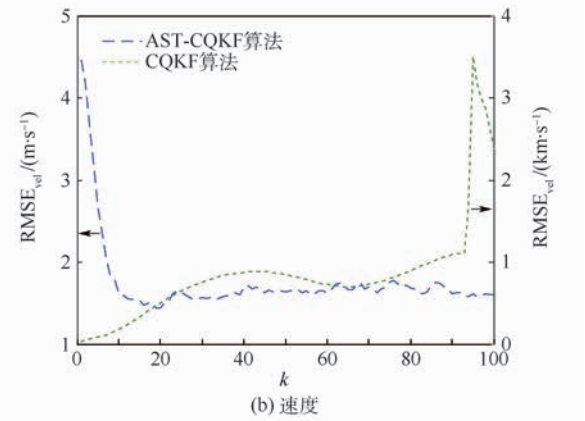
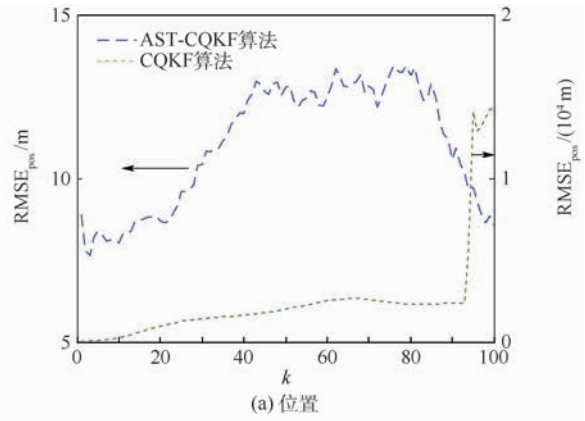


图 3 跟踪均方根误差(情形 2)
Fig.3 RMSE tracking(Case 2)

$$\begin{aligned}
 & \mathbf{X}_k = \\
 & \begin{pmatrix} 1 & a \frac{\sin(\Omega T)}{\Omega} & 0 & -c \left(\frac{1 - \cos(\Omega T)}{\Omega} \right) & 0 \\ 0 & a \cos(\Omega T) & 0 & c - \sin(\Omega T) & 0 \\ 0 & a \frac{1 - \cos(\Omega T)}{\Omega} & 1 & c \frac{\sin(\Omega T)}{\Omega} & 0 \\ 0 & a \sin(\Omega T) & 0 & c \cos(\Omega T) & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \\
 & \mathbf{X}_{k-1} + \mathbf{v}_{k-1} \quad (52)
 \end{aligned}$$

式中: a 和 c 为可调节的参数,模型正常时两者均为 1,通过设置不同的系数以模拟系统状态模型的不确定性,本文设定 $a = 1.1, c = 1.2$ 。仿真过程中,状态真值由精确的状态方程得到,而滤波器使用不精确的状态方程进行状态估计,以此比较

AST-CQKF 与 CQKF 算法在系统状态模型不准确时的滤波精度。经过 200 次独立的蒙特卡罗仿真,得到位置、速度、转弯率的 RMSE 结果分别如图 3 所示(情形 2)。

由图 3 可知,系统状态模型不准确时,CQKF 在仿真开始后迅速发散,无法对目标保持跟踪,而 AST-CQKF 算法虽然不能完全消除模型不确定性的影响,但能够防止滤波发散,保持滤波收敛性,对于模型不确定型具有鲁棒性。

设置不同的仿真步数与仿真次数,得到 CQKF 算法与 AST-CQKF 算法的仿真计算时间如表 1 所示。由表 1 可以看出,与前者相比,后者的仿真计算时间增加了一倍左右,尤其是仿真次数较多时,计算时间的增加更为明显,这是由于 AST-CQKF 算法步骤更多、复杂度更高引起的。但单次仿真时,2 种算法的时间均小于 0.1 s,因此与滤波精度的提高相比,AST-CQKF 算法的时间增加处于可接受的范围内。

表 1 CQKF 与 AST-CQKF 算法蒙特卡罗
仿真计算时间

Table 1 Calculation time of Monte Carlo simulation in
CQKF and AST-CQKF algorithm

仿真步数	仿真次数	仿真计算时间/s	
		CQKF	AST-CQKF
100	200	3.3508	6.9334
100	100	1.7423	3.4171
100	1	0.0344	0.0829
50	200	1.7366	3.3809
50	100	0.9034	1.6991
50	1	0.0237	0.0622

6 结论

CQKF 算法在复杂对抗环境下的目标跟踪主要存在 2 方面问题:①系统状态空间模型发生突变导致滤波发散;②噪声统计特性发生改变或不精确造成滤波精度下降。针对该问题,本文提出了一种新的自适应强跟踪 CQKF (AST-CQKF) 算法应用到目标跟踪:

1) 将强跟踪滤波器与 CQKF 相结合,给出了适用于 CQKF 的渐消因子计算方法,利用渐消因子修正一步状态预测协方差矩阵,进而用于状态估计,克服了模型不准确影响滤波精度和滤波稳定性的问题,改善了 CQKF 的跟踪性能。

2) 利用 Sage-Husa 时变噪声统计估值器对噪声实时估计,得到了 AST-CQKF 算法,增强了滤波器对于噪声统计特性变化的自适应能力,有效地提高了滤波精度。

3) 仿真结果表明,在系统量测噪声统计特性不准确或状态空间模型不准确的情况下,CQKF 算法的滤波精度急剧下降甚至滤波发散,AST-CQKF 算法则能够实现系统状态的快速准确跟踪,有效地克服了 CQKF 算法的局限性。

参考文献 (References)

- [1] KALMAN R E. A new approach to linear filtering and prediction theory[J]. Transactions on ASME Journal of Basic Engineering, 1960, 82(D):35-46.
- [2] JAZWINSKI A H. Stochastic processes and filtering theory [M]. New York: Academic Press, 1970:235-237.
- [3] 李天成,范红旗,孙树栋. 粒子滤波理论、方法及其在多目标跟踪中的应用[J]. 自动化学报, 2015, 41(12):1981-2002.
- [4] LI T C, FAN H Q, SUN S D. Particle filtering: Theory, approach, and application for multitar-get tracking[J]. Acta Automatica Sinica, 2015, 41(12):1981-2002 (in Chinese).
- [5] 韩萍,干浩亮,何炜琨,等. 基于迭代中心差分卡尔曼滤波的飞机姿态估计[J]. 仪器仪表学报, 2015, 36(1):187-193.
- [6] HAN P, GAN H L, HE W K, et al. Iterated central difference Kalman filter based aircraft attitude estimation [J]. Chinese Journal of Scientific Instrument, 2015, 36(1):187-193 (in Chinese).
- [7] 王宝宝,吴盘龙. 基于平方根无迹卡尔曼滤波平滑算法的水下纯方位目标跟踪[J]. 中国惯性技术学报, 2016, 24(2):180-184.
- [8] WANG B B, WU P L. Underwater bearing-only tracking based on square-root unscented Kalman filter smoothing algorithm [J]. Journal of Chinese Inertial Technology, 2016, 24(2):180-184 (in Chinese).
- [9] 张龙,崔乃刚,杨峰,等. 高阶容积卡尔曼滤波及其在目标跟踪中的应用[J]. 哈尔滨工程大学学报, 2016, 37(4):573-578.
- [10] ZHANG L, CUI N G, YANG F, et al. High-degree cubature Kalman filter and its application in target tracking [J]. Journal of Harbin Engineering University, 2016, 37(4):573-578 (in Chinese).
- [11] BHAUMIK S, WATI S. Cubature quarature Kalman filter [J]. IET Signal Processing, 2013, 7(7):533-541.
- [12] LAINIOTIS D G. Optimal adaptive estimation: Structure and parameters adaption [J]. IEEE Transactions on Automatic Control, 1971, 16(2):160-170.
- [13] MEHRA R K. On the identification of variances and adaptive filtering [J]. IEEE Transactions on Automatic Control, 1970, 15(2):175-184.
- [14] ODELSON B J, RAJAMANI M R, RAWLINGS J B. A new autocovariance least-squares method for estimating noise covariances [J]. Automatica, 2006, 42(2):303-308.
- [15] AKESSON B M, JORGENSON J B, POULSEN N K, et al. A generalized autocovariance least-squares method for Kalman filter tuning [J]. Journal of Process Control, 2008, 18(7-8):769-779.
- [16] MYERS K A, TAPLEY B D. Adaptive sequential estimation

- with unknown noise statistics[J]. IEEE Transactions on Automatic Control, 1976, 21(8): 520-523.
- [13] KASHYAP R L. Maximum likelihood identification of stochastic linear systems[J]. IEEE Transactions on Automatic Control, 1970, 15(1): 25-34.
- [14] LIMA F V, RAJAMANI M R, SODERSTROM T A, et al. Covariance and state estimation of weakly observable systems: Application to polymerization processes[J]. IEEE Transactions on Control Systems Technology, 2013, 21(4): 1249-1257.
- [15] 李宁, 祝瑞辉, 张勇刚. 基于 Sage-Husa 算法的自适应平方根 CKF 目标跟踪方法[J]. 系统工程与电子技术, 2014, 36(10): 1899-1905.
- LI N, ZHU R H, ZHANG Y G. Adaptive square CKF method for target tracking based on Sage-Husa algorithm[J]. Systems Engineering and Electronics, 2014, 36(10): 1899-1905 (in Chinese).
- [16] 王小旭, 潘泉, 黄鹤, 等. 非线性系统确定采样型滤波算法综述[J]. 控制与决策, 2012, 27(6): 801-812.
- WANG X X, PAN Q, HUANG H, et al. Overview of deterministic sampling filtering algorithms for nonlinear system[J]. Control and Decision, 2012, 27(6): 801-812 (in Chinese).
- [17] 周东华, 席裕庚, 张钟俊. 一种带多重次优渐消因子的扩展卡尔曼滤波器[J]. 自动化学报, 1991, 17(6): 689-695.
- ZHANG D H, XI Y G, ZHANG Z J. A suboptimal multiple fading extended Kalman filter[J]. Acta Automatica Sinica, 1991, 17(6): 689-695 (in Chinese).
- [18] 方君, 戴邵武, 许文明, 等. 基于 ST-SRCKF 的超高速强机动目标跟踪算法[J]. 北京航空航天大学学报, 2016, 42(8): 1698-1708.
- FANG J, DAI S W, XU W M, et al. Highly maneuvering hyper-velocity-target tracking algorithm based on ST-SRCKF[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(8): 1698-1708 (in Chinese).
- [19] 张龙, 崔乃刚, 王小刚, 等. 强跟踪-容积卡尔曼滤波在弹道式再入目标跟踪中的应用[J]. 中国惯性技术学报, 2015, 23(2): 211-218.
- ZHANG L, CUI N G, WANG X G, et al. Strong tracking-cubature Kalman filter for tracking ballistic reentry target[J]. Journal of Chinese Inertial Technology, 2015, 23(2): 211-218 (in Chinese).
- [20] 赵琳, 王小旭, 孙明, 等. 基于极大后验估计和指数加权的自适应 UKF 滤波算法[J]. 自动化学报, 2010, 36(7): 1007-1019.
- ZHAO L, WANG X X, SUN M, et al. Adaptive UKF filtering algorithm based on maximum a posterior estimation and exponential weighting[J]. Acta Automatica Sinica, 2010, 36(7): 1007-1019 (in Chinese).
- [21] 丁家琳, 肖建, 赵涛. 自适应 CKF 强跟踪滤波器及其应用[J]. 电机与控制学报, 2015, 19(11): 111-120.
- DING J L, XIAO J, ZHAO T. Adaptive CKF strong tracking filter and application[J]. Electric Machines and Control, 2015, 19(11): 111-120 (in Chinese).

作者简介:

刘畅 男, 博士研究生。主要研究方向: 非线性滤波、目标跟踪。

杨锁昌 男, 博士, 教授, 博士生导师。主要研究方向: 导弹制导。

Target tracking algorithm based on adaptive strong tracking CQKF

LIU Chang¹, YANG Suochang^{1,*}, WANG Liandong², ZHANG Kuanqiao¹

(1. Department of Missile Engineering, Ordnance Engineering College, Shijiazhuang 050000, China;

2. State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System, Luoyang 471003, China)

Abstract: As cubature quadrature Kalman filter (CQKF) is easily influenced by uncertainty of state-space model and need to know exactly noise statistics, a new type of adaptive CQKF algorithm with strong tracking behavior is proposed. Based on the theory of strong tracking filter, the new algorithm introduces fading factor to adapt to covariance matrix and reinforces residual sequence to be orthogonal, which effectively suppresses the filtering divergence caused by the model uncertainty. In the process of filtering, processing noise and measurement noise should be estimated online by the Sage-Husa noise statistics estimator, which will improve the filter precision under the circumstance of unknown time-varying noise. Simulations of target tracking demonstrate the efficiency and robustness of the algorithm.

Keywords: target tracking; cubature quadrature Kalman filter (CQKF); strong tracking filter; noise statistics estimators; adaptive filter

Received: 2017-05-15; Accepted: 2017-06-26; Published online: 2017-10-13 14:05

URL: kns.cnki.net/kcms/detail/11.2625.V.20171013.1405.004.html

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

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

DOI: 10.13700/j.bh.1001-5965.2017.0325

基于混合地图的护理机器人室内导航方法

张立志, 陈殿生*, 刘维惠

(北京航空航天大学 机械工程及自动化学院, 北京 100083)



摘 要: 护理机器人在室内三维结构化环境下进行导航时, 面临着三维建图计算量大且地图中缺乏语义信息的缺点。提出了基于点和平面特征的混合地图构建方法, 结合点和平面在地图构建中的优势, 并基于该混合地图搭建室内导航系统。首先, 快速地提取特征点和特征平面, 使用解释树的方法进行数据关联, 并使用平滑建图工具构建因子图, 进行机器人位姿和路标的联合优化, 改进并更新混合地图。然后, 搭建室内导航系统, 实现了三维障碍物检测、路径规划与运动控制。最后, 在走廊环境下进行了室内导航实验, 并以由激光雷达构建的二维栅格地图为参考, 分析了地图构建效果和机器人定位精度, 证明了基于混合地图的室内导航系统在室内结构化环境下的优势。

关键词: 护理机器人; 室内导航; 同步定位与建图; 数据关联; 障碍物检测; 路径规划

中图分类号: TP242

文献标识码: A 文章编号: 1001-5965(2018)05-0991-10

同步定位与建图 (Simultaneous Localization and Mapping, SLAM) 作为移动机器人技术的核心, 决定了移动机器人是否能够实现完全自主的关键。近年来, 由于在二维平面内的 SLAM 技术的成熟及其应用限制, 针对三维环境的 SLAM 研究越来越多。三维 SLAM 构建的地图模型接近真实环境, 也包含了更多的环境信息, 能够让机器人在三维环境中进行定位与路径规划, 使机器人在三维环境中的自主移动成为可能。随着越来越多的三维 SLAM 算法被提出, 三维 SLAM 面临着地图构建精度、实时性和缺乏语义模型等难题, 也出现了许多解决方案。

室内环境多为结构化环境, 具有许多平面区域, 许多研究者据此提出了基于平面特征的 SLAM 算法, 建立基于平面特征的地图。平面地

图结构紧凑且具有语义模型的优点, 形象地反映了室内环境结构化的特点, 不仅可以用来提取语义地图, 进行基于语义信息的导航, 还可以用来辅助抓取等任务的实现。

研究者们提出了许多基于平面特征的 SLAM 算法。Weingarten 和 Siegwart^[1] 提取了基于扩展卡尔曼滤波器 (Extended Kalman Filter, EKF) 的平面特征 SLAM。使用移动机器人上的旋转激光雷达获取点云数据, 采用区域增长法分割平面, 用 SP-model^[2] 表示平面特征, 使用基于 EKF 的 SLAM 框架建立三维平面地图。Lee 等^[3] 提出了一种在室内环境下从有噪声的 RGB-D 传感器中提取平面特征并建立三维地图的方法。该方法中的平面特征参数使用球坐标表示, 同时使用点云中点的位置和误差估计平面的空间误差, 显示了

收稿日期: 2017-05-17; 录用日期: 2017-07-13; 网络出版时间: 2017-09-22 15:07

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20170922.1507.007.html

基金项目: 北京市科技计划重大项目课题 (D141100003614002)

* 通信作者. E-mail: chends@163.com

引用格式: 张立志, 陈殿生, 刘维惠. 基于混合地图的护理机器人室内导航方法[J]. 北京航空航天大学学报, 2018, 44(5): 991-1000. ZHANG L Z, CHEN D S, LIU W H. Care robot indoor navigation method based on hybrid map[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 991-1000 (in Chinese).

优于其他平面表现形式的建图效果。Kaess^[4]提出了一种无限平面的最少参数表示形式,通过将齐次化的参数映射到四元数来表示无限平面,该表示方法非常适合进行最小二乘优化,如使用ISAM 递增优化算法^[5]对位姿图进行优化求解,分别在仿真环境、真实环境中使用该表示方法进行了平面建图实验,实现了快速稳定的平面地图创建。但平面特征提取和匹配耗费了大量的计算时间,可以进一步进行改进。

尽管平面特征广泛存在于室内环境下,但由于传感器测距范围和视角的限制,往往获取很少的平面特征,同时获取的平面特征中往往存在平行的平面,仅仅使用平面来进行位姿估计常常出现失败的情况。因此,结合平面特征和其他特征的位姿估计引起研究者的兴趣,如平面与线结合、平面与点结合等。

为了解决仅使用平面特征进行 SLAM 带来的问题,Trevor 等^[6]使用 RGB-D 传感器和激光雷达,基于平面特征和直线特征进行位姿估计,利用激光雷达的常测距距离和大视角来弥补 RGB-D 传感器小视角短距离的缺陷,构建基于直线和平面特征的地图。Taguchi 等^[7]提出了基于点和平面特征的 SLAM 系统,使用点和平面的组合来进行位姿估计,一个基于 RANSAC 算法的配准算法用于估计相机的运动。但是该系统的运算量过大,同时不能有效地对重复纹理环境进行建图。在 Taguchi 等^[7]的基础上,Ataer-Cansizoglu 等^[8]提出了一种基于预测的相机位姿跟踪算法,实现特征的快速匹配,提高了构建地图的效率,并能够很好地对重复纹理环境进行建图,构建基于点和平面特征的混合地图。Salas-Moreno 等^[9]提出了一种 Dense Planar SLAM 算法,同时使用带有边界的平面点和非平面点,建立密集的环境地图。该算法可以检测、拼接、压缩室内结构环境下的平面,而且通过使用 GPU (Graphics Processing Unit),能够达到实时建图效果。Salas-Moreno 等^[9]同时展示了该平面区域 SLAM 算法在人工智能领域的潜在应用前景。

因此,目前多数基于平面特征的 SLAM 系统存在着不能构建大范围地图、可靠性差和计算量大的问题。在前期研究^[10]的基础上,本文同样使用点和平面特征进行位姿估计,构建大范围地图,提高 SLAM 的可靠性和实用性。但是不同的是,构建的是基于点和平面特征的混合地图。以所提出的 SLAM 算法为核心,搭建室内导航系统,并进行

室内导航实验,分析所提出的 SLAM 算法的性能。

1 护理机器人样机

护理机器人的概念设计如图 1 所示。为了使护理机器人样机具备移动能力,设计了全向移动底盘。机器人具有里程计、激光雷达、IMU、Kinect 和超声波等传感器。通过里程计可以使用航迹推算算法估计机器人的运动。Kinect 固定在机器人顶部,作为 SLAM 的传感器。同时,为了能够实现抓取等对环境物体的操作任务,在服务机器人上集成了 7 自由度模块化机械臂。护理机器人样机如图 2 所示,表 1 列举出了样机指标,其他传感器包括里程计、IMU、激光雷达、超声波。

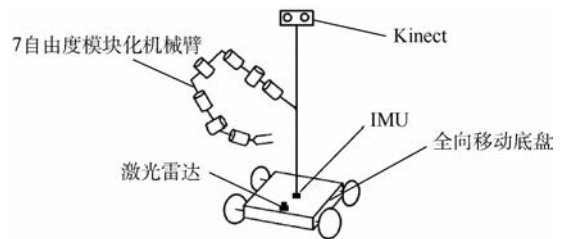


图 1 护理机器人概念设计图

Fig. 1 Care robot conceptual design sketch



图 2 护理机器人样机

Fig. 2 Care robot prototype

表 1 护理机器人样机指标

Table 1 Care robot prototype index

指标	数值
整机质量/kg	120
外形尺寸/(m × m × m)	0.9 × 0.65 × 1.35
机械臂自由度	7
Kinect 安装高度/m	1.2

2 室内导航系统总体架构

采用 ROS^[11] 导航程序包框架,搭建护理机器人的室内导航系统。针对不同的移动机器人平台,需要配置里程计信息、传感器输入和底盘控制器,同时需要提供定位和地图。室内导航系统的架构如图 3 所示,由 EKF 位姿估计、定位与建图、障碍物地图、路径规划与运动控制几个部分组成。

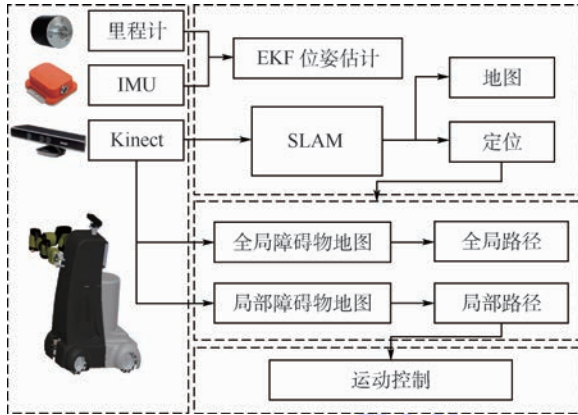


图 3 室内导航系统架构图

Fig. 3 Indoor navigation system architecture diagram

3 定位与建图

3.1 总体框架

本文室内导航系统的定位与建图部分基于乔治亚理工学院的平滑建图工具 (Georgia Tech Smoothing and Mapping, GTSAM)^[12]。图 4 概述了 SLAM 系统的框架,主要由特征提取、数据关联、因子图的建立和优化、地图的更新与改进等部分组成。

SLAM 系统基于点和平面 2 种特征,特征点提取使用 ORB (Oriented FAST and Rotated BRIEF) 特征点检测^[13],特征平面提取使用基于直线元素的快速平面提取方法^[14]。每个数据帧中包含机器人当前的姿态估计、特征点信息、特征平面信息。特征点的信息有特征描述符、三维位置及其协方差,即

$$P_i = \{d_i, I_i(u, v), p_i(x, y, z), \Sigma_i\} \quad (1)$$

式中: P_i 为所提取的特征点; d_i 为特征点的描述符; $I_i(u, v)$ 为特征点在图像中的位置, u, v 分别为水平、垂直位置; $p_i(x, y, z)$ 为特征点在相机坐标系下的三维位置; Σ_i 为三维位置的协方差。

特征平面的信息有平面参数、平面曲率和表示平面区域的内点点云,即

$$L_j = \{\xi_j(n, d), \kappa_j, c_j\} \quad (2)$$

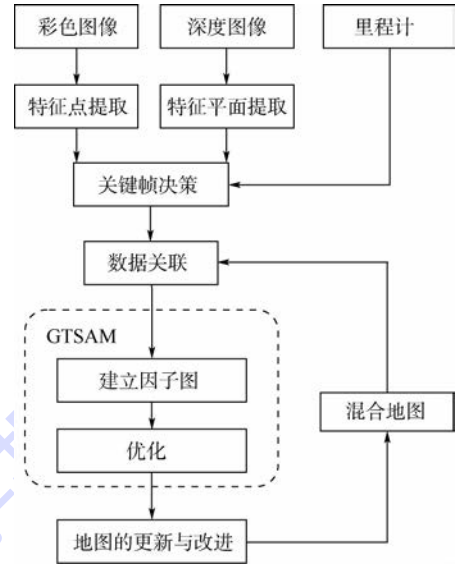


图 4 SLAM 系统框架图

Fig. 4 SLAM system framework diagram

式中: L_j 为所提取的特征平面; $\xi_j(n, d)$ 为平面参数, n, d 分别为平面法向量、原点到平面的距离; κ_j 为平面曲率; c_j 为包含平面的所有内点的点云。

由于平面内点的数量比较多,使用点云库 (Point Cloud Library, PCL)^[15] 中的体素滤波器 (voxel grid filter) 对内点进行下采样,控制平面内点的数目。

在基于图优化的 SLAM 系统中,需要根据某一准则确定关键帧,以减少数据帧的数量,降低内存使用量和系统的计算量,提高定位与建图的效率。关键帧的确定通常根据相对位姿^[16] 或观测信息变化量^[17] 进行决策。所提出的 SLAM 系统的关键帧由机器人当前数据帧与上一关键帧的相对位姿进行判断,设定相对平移阈值和相对旋转阈值,当相对平移大于阈值或相对旋转大于阈值时,将当前数据帧作为关键帧。在实验验证中,选取平移阈值为 0.1 m,旋转阈值为 10°。

在 SLAM 过程中,需要找到当前观测值与地图中路标的一一对应关系,即数据关联问题^[18]。数据关联的结果决定着 SLAM 系统定位和建图的效果,也影响 SLAM 系统的运算时间。为了实现观测特征与地图路标快速可靠的匹配,所提出的 SLAM 系统使用基于解释树^[19] 的数据关联方法。

使用 GTSAM 工具中的 iSAM2^[20] 库进行图优化,根据优化后的结果,更新定位信息和地图信息。构建的环境地图模型为基于点和平面特征的混合地图。由于传感器视角和距离的限制,在地图构建过程中,往往会有属于同一环境平面区域的平面特征被重复地添加到地图中,导致地图中存在重复的平面路标,不仅破坏了 SLAM 系统的

建图的效果,同时也影响定位精度。因此,需要检测地图中的重复平面路标,并将重复的平面路标合并,对地图进行改进。

3.2 数据关联方法

本文所使用的数据关联方法基于解释树(interpretation tree),其主要思想是考虑所有的匹配可能性,建立解释树,再使用特征之间的约束来剔除错误的子树。如图5所示, $\{F_1, F_2, \dots, F_m\}$ 为观测特征, $\{L_1, L_2, \dots, L_n\}$ 为地图中的路标,用“*”代表没有在地图中找到对应的匹配路标。根据特征点或特征平面的所有可能数据关联结果建立解释树,再根据特征之间的约束关系,逐级向下删除错误的子树。约束关系可以是一元约束,即观测特征与地图路标之间的约束关系,如几何参数相似性、位置相关性等;也可以为二元约束,即2个特征匹配对中,2个观测特征的空间相对位姿与2个地图路标的空间相对位姿的约束关系,如平面的夹角约束、点的距离约束等。

对于特征点的匹配,使用描述符的相似性、三维位置的相关性和点到点的距离约束来剔除错误的子树。

1) 描述符的相似性。ORB特征点的描述符为256维的向量。计算描述符间的汉明距离(Hamming distance),设定最大汉明距离阈值,如果汉明距离小于阈值,则特征点和路标点符合描述符相似性约束。

2) 三维位置的相关性。根据机器人位置和相机的内参数,将地图中路标点的三维位置传递到相机坐标系下,对路标点的观测值进行预测。再计算特征点观测值与预测值间的马氏距离(Mahalanobis distance),如果马氏距离小于阈值,则特征点和路标点的三维位置相关。

3) 点到点的距离约束。该约束属于二元约束,如图6所示, f_1 和 f_2 为观测值, l_1 和 l_2 为路

标的预测值。假设 f_1 与 l_1 的关联成立,判断 f_2 与 l_2 的关联是否成立。计算特征点 f_1 与 f_2 之间的距离 d_f ,路标点 l_1 与 l_2 之间的距离 d_l ,如果满足 $|d_f - d_l| < \delta$,其中 δ 为允许偏差,则特征点和路标点之间符合点到点的距离约束, f_2 与 l_2 的关联成立。

对于特征平面的匹配,使用平面参数的相似性、平面区域的重叠和平面与平面间的夹角约束来剔除错误的子树。

1) 平面参数的相似性。将地图中的路标平面传递到机器人坐标系下,得到平面观测的预测值。计算预测值与观测值之间的角度偏差与距离偏差,如果角度偏差和距离偏差同时小于对应阈值,则特征平面和路标平面的平面参数相似。

2) 平面区域的重叠。匹配的平面应该在平面所表示的平面区域上存在着一定的重叠。采用八叉树(octree)^[21]的数据结构,使用平面的内点,计算2个平面的重叠率。首先将路标平面的所有内点传递到机器人坐标系下,建立八叉树;再使用特征平面的每一个内点,判断是否与八叉树相交,计算相交的内点数;如果相交的内点数大于阈值,则特征平面与路标平面所代表的平面区域存在着重叠。

3) 平面与平面间的夹角约束。该约束属于二元约束,如图7所示,假设观测值 f_1 与预测值 l_1 的关联成立,判断 f_2 与 l_2 是否关联。计算观测值 f_1 与 f_2 之间的夹角 α ,预测值 l_1 与 l_2 之间的夹角 β ,如果 $|\alpha - \beta| < \varepsilon$,其中 ε 为允许偏差,则特征平面和路标平面符合平面间的夹角约束,即 f_2 与 l_2 的关联成立。

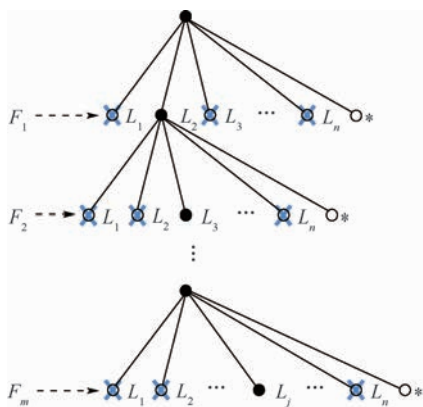


图5 基于解释树的数据关联

Fig.5 Data association based on interpretation tree

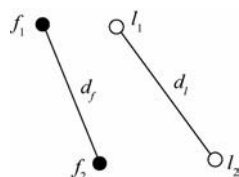


图6 点到点的距离约束

Fig.6 Point to point distance constraint

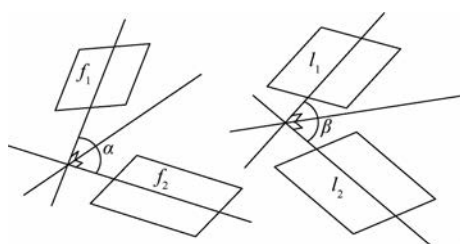


图7 平面与平面间的夹角约束

Fig.7 Angle constraint between plane and plane

3.3 地图的更新与改进

根据优化结果,更新基于点和平面的混合地图。对于每一个点路标,地图中存储的信息有特征点的描述符、点的三维位置,即

$$P_i^{(w)} = \{d_i^{(w)}, p_i^{(w)}\} \quad (3)$$

式中: $P_i^{(w)}$ 为在世界坐标系下的点路标; $d_i^{(w)}$ 为点路标的描述符; $p_i^{(w)}$ 为点路标的三维位置。

而对于每一个平面路标,地图中存储的信息有平面在世界坐标系下的平面参数、平面内点点云,即

$$L_j^{(w)} = \{\xi_j^{(w)}, c_j^{(w)}\} \quad (4)$$

式中: $L_j^{(w)}$ 为在世界坐标系下的平面路标; $\xi_j^{(w)}$ 为在世界坐标系下的平面参数; $c_j^{(w)}$ 为在世界坐标系下的包含平面的所有内点的点云。点路标的三维位置 $p_i^{(w)}$ 和平面路标的平面参数 $\xi_j^{(w)}$ 可以从优化结果中直接获取。表示平面区域的内点点云可以将多个观测平面特征的内点传递到世界坐标系下再相加而得到。但是随着 SLAM 进行,内点点云会变的非常多,为了控制平面内点的数量,使用点云库中的体素网格滤波器对相加后的内点点云进行下采样。不仅减少内点的数量,也会提高平面数据关联的效率。经过体素网格滤波器下采样后的平面内点点云如图 8 所示,椭圆标记部分为新观测到的平面内点。

由于深度传感器视角和测距范围的限制,会有属于同一平面区域的重复路标被添加到地图中。如图 9 所示为建立的一个房间内的平面地图,平面 1 被先观测到,并被添加入地图中;随着地图构建过程的进行,平面 2 被从另外一个视角观测到,也被添加入地图中。但是,平面 1 和平面 2 属于同一个环境平面区域,应当将平面 1 和平面 2 进行合并。

重复的平面路标满足 2 个条件:

- 1) 平面参数相似。
- 2) 平面区域重叠。

判断平面参数的相似性、平面区域是否重叠采用与数据关联中平面匹配相似的方法。

将重复的平面路标合并得到一个平面路标,应该优先保留平面区域大的平面路标。在合并过程中,不仅需要 2 个平面的内点合并,还需要在因子图中用合并平面路标变量替换原有路标变量,并替换相关因子。假设 2 个重复的平面路标中内点数较多的路标为 L_1 ,另一个平面路标为 L_2 ,它们在因子图中的识别符分别为 S_1 和 S_2 ,重复平面的合并过程如下:

- 1) 从因子图中删除平面路标 L_2 对应的

变量。

- 2) 对因子图中的所有因子,用 S_1 替换 S_2 。
- 3) 平面路标 L_1 和 L_2 的内点相加,经过体素网格滤波器下采样后作为合并平面的内点。
- 4) 进行地图优化和更新。

图 9 所示的重复平面路标合并后的平面地图见图 10,平面 2 被合并到平面 1 中。不仅正确地描述了环境平面区域,也正确地处理了变量间的约束关系,减少了变量优化误差,提高了建图效果。

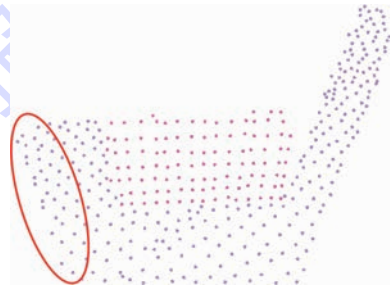


图 8 平面路标点云更新

Fig. 8 Plane landmark point cloud update

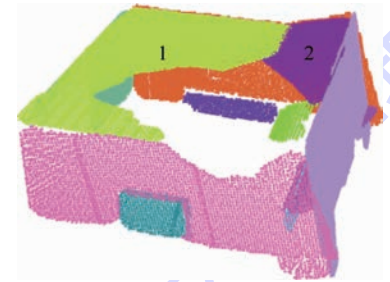


图 9 存在重复平面路标的平面地图

Fig. 9 Plane map with reduplicated plane landmarks

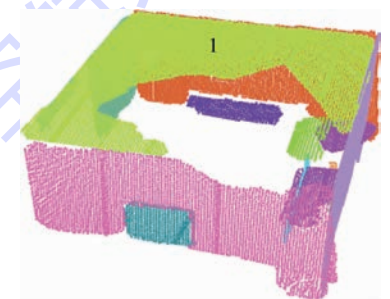


图 10 重复平面路标合并后的平面地图

Fig. 10 Plane map after reduplicated plane landmarks merging

4 障碍物检测

室内导航系统采用 2 个障碍物地图来存储环境障碍物信息,即全局障碍物地图和局部障碍物地图。全局障碍物地图用来存储整个环境的障碍物信息,为了降低计算量,全局障碍物地图的更新频率很低;局部障碍物地图用来存储机器人附近

一定区域内的障碍物信息,更新频率比较高,可以实时反应出环境信息,用于动态避障。

障碍物地图采用三维占栅格地图,如图 11 所示,每个栅格有未知、占用、空闲 3 个状态。输入数据为 Kinect 传感器得到的深度点云测量值,根据深度测量值,对三维占栅格进行障碍物标记和清除操作。为了便于进行路径规划,通常将三维占栅格地图可以投影到二维平面上,生成二维栅格地图,如图 12 所示。

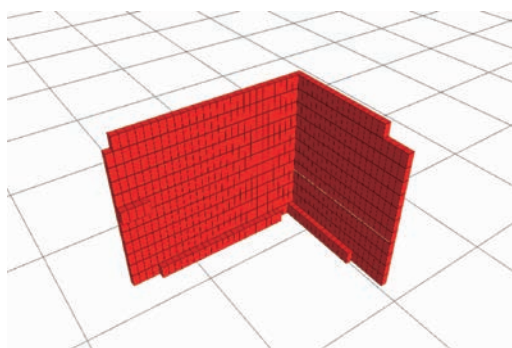


图 11 三维占栅格地图

Fig. 11 3D occupancy grid map

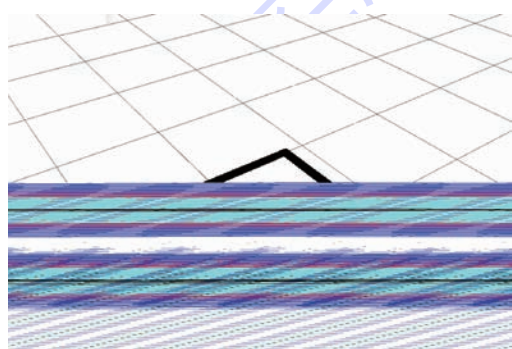


图 12 投影后的二维栅格地图

Fig. 12 2D occupancy grid map after projection

5 路径规划与运动控制

路径规划采用全局和局部相结合的方式。全局路径在全局障碍物地图上规划出一条全局路径,局部路径的目标点跟随全局路径,在局部障碍物地图上进行规划,实现机器人的动态避障。

由于机器人的投影足迹具有一定的形状,不能把机器人看作一个点来进行路径规划。因此,在使用代价地图进行路径规划之前,需要根据机器人的足迹形状,对代价地图进行膨胀操作。

如图 13 所示,红色多边形为机器人足迹,指定了内切圆半径 R_i 和外接圆半径 R_o 。对于代价地图中的每一个栅格,假设将机器人中心放在该栅格中心位置,栅格到障碍物的最小距离为 d_g 。显然, d_g 越大,机器人距离障碍物越远,碰撞的可

能越小。

在进行膨胀操作时,随着 d_g 的增大赋予递减的代价值,对于 $d_g < R_i$ 的栅格,赋予一个致命的代价值;对于 $R_i < d_g < R_o$ 的栅格,赋予较高的代价值;对 $d_g > R_o$ 的栅格,采用指数函数递减的代价值。将图 12 中的障碍物地图膨胀后的代价地图如图 14 所示。显然,在进行路径规划时,应尽量选择代价值小的栅格,更能保证机器人的避障性能。

全局路径规划使用 A* 算法^[22],局部路径规划使用动态窗口法(Dynamic Window Approach, DWA)^[23]。图 15 显示了一个路径规划的结果,红色多边形为护理机器人投影到二维平面上的足迹,绿色曲线为全局规划路径,蓝色曲线为局部规划路径。



图 13 膨胀半径示意图

Fig. 13 Schematic diagram of inflated radius

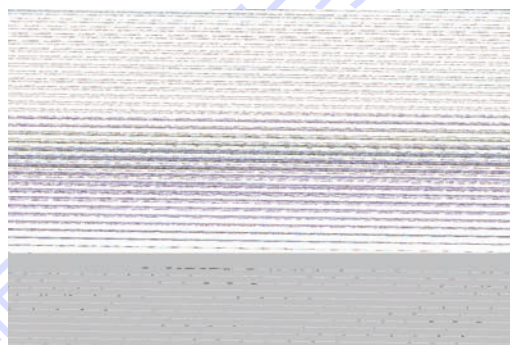


图 14 膨胀操作后的代价地图

Fig. 14 Costmap after inflation operation

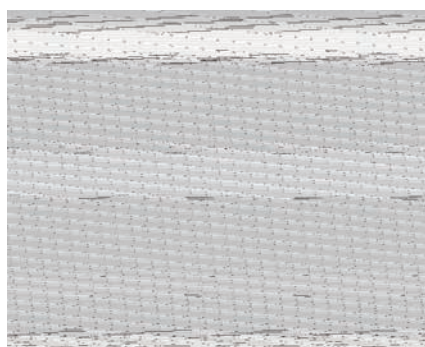


图 15 全局与局部路径规划

Fig. 15 Global and local path planning

6 室内导航实验

在室内走廊环境下进行基于点和平面特征的 SLAM 实验。走廊环境为典型的室内结构化环境,有许多的平面区域,但环境纹理比较少,提取的特征点会比较少。为了分析基于点和平面特征的 SLAM 的定位精度和建图效果,使用激光雷达传感器,应用 GMapping^[24] 建立二维占栅格地图,将 SLAM 轨迹与 GMapping 轨迹进行比较,并将 SLAM 构建的混合地图与 GMapping 构建的占栅格地图进行对比。激光雷达的有效测距范围为 30 m,测量角度范围为 270°,构建的二维占栅格地图非常精确,常被作为评估其他 SLAM 算法的参考值。从建图效果、定位精度和运算时间 3 个方面分析实验结果。

所有的 SLAM 实验运行在单个线程内,CUP 主频为 2.9 GHz,内存为 12 GB,没有使用 GPU。

6.1 建图效果

SLAM 构建的基于点和平面特征的混合地图如图 16 所示,不同颜色的平面区域代表不同的平面路标,不同的球形点代表不同的点路标。地图中含有 45 个平面路标和 266 个点路标。平面路标正确地描述了环境平面区域,三维地图与真实环境基本吻合。

GMapping^[24] 构建的二维占栅格地图如图 17 所示,基本符合建筑物地图。将 SLAM 地图与二维占栅格进行比较,如图 18 所示。可以看到,SLAM 构建的混合地图很好地与栅格地图重叠,尤其是混合地图中的平面路标,与栅格地图可以很好地重合。

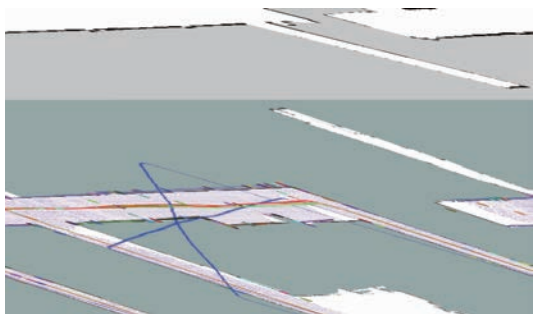


图 16 基于点和平面特征的混合地图

Fig. 16 Hybrid map based on point and plane features

6.2 定位精度

以 GMapping 轨迹作为参考值,将 SLAM 轨迹与 GMapping 轨迹进行对比(见图 19),来估计所提出 SLAM 算法的定位精度,并计算位姿的均方根误差(Root Mean Square Error, RMSE)^[25],

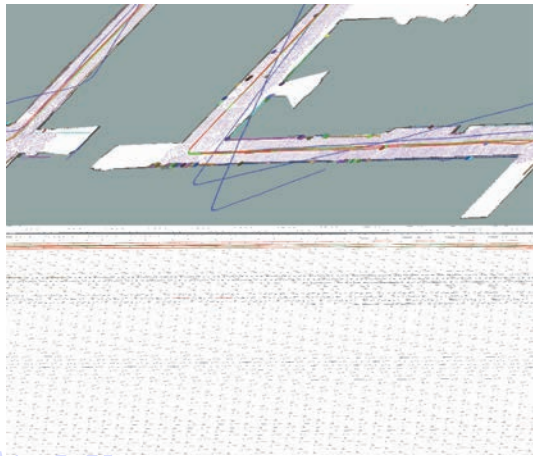


图 17 GMapping 生成的占栅格地图

Fig. 17 Occupancy grid map generated by GMapping

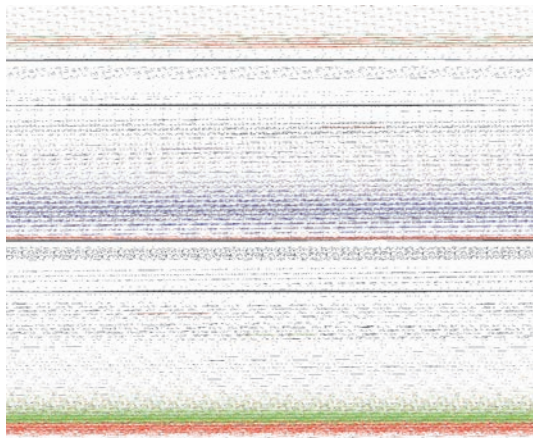


图 18 SLAM 地图与占栅格地图对比

Fig. 18 Comparison of SLAM map and occupancy grid map

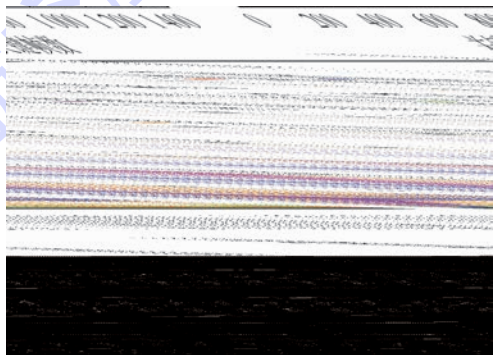


图 19 SLAM 轨迹与 GMapping 轨迹对比

Fig. 19 Comparison of SLAM trajectory and GMapping trajectory

SLAM 实验的详细结果如表 2 所示。可以看到,在长度达 180 m 的室内实验中,SLAM 轨迹与 GMapping 轨迹基本吻合,SLAM 轨迹的平移 RMSE 为 0.106 3 m、旋转 RMSE 为 1.194 1°。

表 2 SLAM 实验结果

Table 2 SLAM experimental results

实验结果	数值
轨迹长度/m	179.74
持续时间/s	617.53
数据帧数	6176
关键帧数	1495
平移误差均值/m	0.0970
平移 RMSE/m	0.1063
旋转误差均值/(°)	1.0840
旋转 RMSE/(°)	1.1941

6.3 运算时间

SLAM 过程中各步骤的运算时间如图 20 所示,特征点提取和特征平面提取的详细运算时间如图 21 所示。数据帧处理的时间保持稳定,约为 30 ms。建地图的时间随着 SLAM 的进行而不断增加。进一步分析建图过程中各个步骤的运算时间,图 22 描述了建图过程中平面匹配、点匹配、图优化、地图改进、路标更新的运算时间。

从图 22 可以看出,平面匹配和点匹配的运算时间基本保持稳定,但会随着 SLAM 过程的进行而略有增加。图优化的运算时间存在着很大的波动性,而地图改进的运算时间中存在着稳定上升部分和跳变部分。稳定上升部分为检测重复平面路标所花费的时间,平面路标越多,运算时间越长;跳变部分为重复路标合并所花费的时间,只有在检测到重复平面时才进行,因此出现向上的跳变。路标更新的运算时间随着 SLAM 过程的进行而显著增加,主要由于平面路标的内点点云更新需要许多的计算量。SLAM 过程中各个步骤的平均运算时间如表 3 所示。

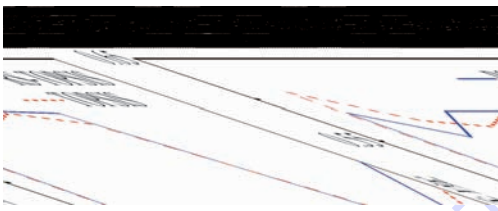


图 20 SLAM 各步骤运算时间

Fig. 20 Each SLAM step runtime



图 21 特征提取运算时间

Fig. 21 Feature extraction runtime



图 22 建图各步骤运算时间

Fig. 22 Each mapping step runtime

表 3 SLAM 各步骤平均运算时间

Table 3 Each SLAM step average runtime

步骤	运算时间/ms
特征平面提取	17.39
特征点提取	27.11
特征平面数据关联	9.66
特征点数据关联	22.78
因子图优化	116.07
地图改进	100.35
路标更新	193.50
总计	499.36

7 结论

本文提出了基于混合地图的护理机器人室内导航方法,具有如下特点:

1) 基于点和平面特征,构建混合地图,结合单独使用点或平面特征进行导航的优点,以提高导航系统的可靠性和实用性。

2) 以混合地图为核心,搭建护理机器人的室内导航系统,进行了室内导航实验,验证了所提出的基于混合地图的导航系统的可靠性和实用性。

在将来的研究中,可以加入闭环检测功能,提高建图效果;也可以对建图过程中的地图改进和更新算法进行改进,减少计算量。

参考文献 (References)

- [1] WEINGARTEN J, SIEGWART R. 3D SLAM using planar segments[C]//2006 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, NJ: IEEE Press, 2006: 3062-3067.
- [2] CASTELLANOS J A, TARDOS J D. Mobile robot localization and map building[M]. Berlin: Springer, 1999: 9-30.
- [3] LEE T K, LIM S, LEE S, et al. Indoor mapping using planes extracted from noisy RGB-D sensors[C]//2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, NJ: IEEE Press, 2012: 1727-1733.
- [4] KAESS M. Simultaneous localization and mapping with infinite planes[C]//2015 IEEE International Conference on Robotics and Automation. Piscataway, NJ: IEEE Press, 2015: 4605-4611.
- [5] KAESS M, RANGANATHAN A, DELLAERT F. iSAM: Incremental smoothing and mapping[J]. IEEE Transactions on Ro-

- botics,2008,24(6):1365-1378.
- [6] TREVOR A J B, ROGERS J G, CHRISTENSEN H I. Planar surface SLAM with 3D and 2D sensors[C]//2012 IEEE International Conference on Robotics and Automation. Piscataway, NJ:IEEE Press,2012:3041-3048.
- [7] TAGUCHI Y, JIAN Y D, RAMALINGAM S, et al. Point-plane SLAM for hand-held 3D sensors[C]//2013 IEEE International Conference on Robotics and Automation. Piscataway, NJ:IEEE Press,2013:5182-5189.
- [8] ATAER-CANSIZOGLU E, TAGUCHI Y, RAMALINGAM S, et al. Tracking an RGB-D camera using points and planes[C]//2013 IEEE International Conference on Computer Vision Workshops. Piscataway, NJ:IEEE Press,2013:51-58.
- [9] SALAS-MORENO R F, GLOCKEN B, KELLY P H J, et al. Dense Planar SLAM[C]//2014 IEEE International Symposium on Mixed and Augmented Reality. Piscataway, NJ:IEEE Press,2014:157-164.
- [10] ZHANG L Z, CHEN D S, LIU W H. Point-plane SLAM based on line-based plane segmentation approach[C]//2016 IEEE International Conference on Robotics and Biomimetics. Piscataway, NJ:IEEE Press,2016:1287-1292.
- [11] QUIGLEY M, CONLEY K, GERKEY B P, et al. ROS: An open-source robot operating system[C]//ICRA Workshop on Open Source Software. Piscataway, NJ:IEEE Press,2009:1-6.
- [12] DELLAERT F. Factor graphs and GTSAM: A hands-on introduction;GT-RIM-CP&R-2012-02[R]. Atlanta:Georgia Institute of Technology,2012:1-26.
- [13] RUBLEE E, RABAUD V, KONOLIGE K, et al. ORB: An efficient alternative to SIFT or SURF[C]//2011 IEEE International Conference on Computer Vision. Piscataway, NJ:IEEE Press,2011:2564-2571.
- [14] ZHANG L Z, CHEN D S, LIU W H. Fast plane segmentation with line primitives for RGB-D sensor[J]. International Journal of Advanced Robotic Systems,2016,13(6):1-8.
- [15] RUSU R B, COUSINS S. 3D is here: Point cloud library (PCL) [C]//2011 IEEE International Conference on Robotics and Automation. Piscataway, NJ:IEEE Press,2011:1-4.
- [16] KLEIN G, MURRAY D. Parallel tracking and mapping for small AR workspaces[C]//IEEE and ACM International Symposium on Mixed and Augmented Reality. Piscataway, NJ:IEEE Press,2008:225-234.
- [17] MUR-ARTAL R, TARDOS J D. ORB-SLAM2: An open-source SLAM system for monocular, stereo and RGB-D cameras[J]. IEEE Transactions on Robotics,2017,33(5):1255-1262.
- [18] GAD A, MAJDI F, FAROOQ M. A comparison of data association techniques for target tracking in clutter[C]//2002 IEEE International Conference on Information Fusion. Piscataway, NJ:IEEE Press,2002:1126-1133.
- [19] POPPINGA J, VASKEVICIUS N, BIRK A, et al. Fast plane detection and polygonalization in noisy 3D range images[C]//2008 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, NJ:IEEE Press,2008:3378-3383.
- [20] KAESS M, JOHANSSON H, ROBERTS R, et al. iSAM2: Incremental smoothing and mapping using the Bayes tree[J]. International Journal of Robotics Research,2012,31(2):216-235.
- [21] KAMMERL J, BLODOW N, RUSU R B, et al. Real-time compression of point cloud streams[C]//2012 IEEE International Conference on Robotics and Automation. Piscataway, NJ:IEEE Press,2012:778-785.
- [22] HART P E, NILSSON N J, RAPHAEL B. A formal basis for the heuristic determination of minimum cost paths[J]. IEEE Transactions on Systems Science & Cybernetics,1968,4(2):100-107.
- [23] FOX D, BURGARD W, THRUN S. The dynamic window approach to collision avoidance[J]. IEEE Robotics & Automation Magazine,1997,4(1):23-33.
- [24] GRISETTIY G, STACHNISS C, BURGARD W. Improving grid-based SLAM with Rao-Blackwellized particle filters by adaptive proposals and selective resampling[C]//2005 IEEE International Conference on Robotics and Automation. Piscataway, NJ:IEEE Press,2005:2432-2437.
- [25] STURM J, ENGELHARD N, ENDRES F, et al. A benchmark for the evaluation of RGB-D SLAM systems[C]//2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, NJ:IEEE Press,2012:573-580.

作者简介:

张立志 男,博士研究生。主要研究方向:机器人移动导航。

陈殿生 男,博士,教授,博士生导师。主要研究方向:助老助残机器人、仿生机器人、机电智能控制技术。

刘维惠 女,博士研究生。主要研究方向:机械臂运动控制。

Care robot indoor navigation method based on hybrid map

ZHANG Lizhi, CHEN Diansheng*, LIU Weihui

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

Abstract: When the care robot is navigating in the indoor 3D structured environment, it is faced with the disadvantage of the large computational cost for map building and the lack of semantic information in the map. This paper presents a hybrid map building method based on point and plane features, which combines the advantages of point and plane features in the map building. Furthermore, an indoor navigation system is built based on the proposed hybrid map. First, point and plane features are fast extracted, and then data association is achieved using the interpretation tree approach. The smoothing and mapping tool is utilized to construct the factor graph and jointly optimize robot poses and landmarks, and the hybrid map is refined and updated. Second, the indoor navigation system is built, which implements the 3D obstacle detection, path planning and motion control. Finally, the indoor navigation experiments were carried out in a corridor environment. With the 2D occupancy grid map constructed by laser as the reference, the performance of map building and robot localization accuracy were analyzed, which proves that the indoor navigation system based on hybrid map shows its advantages in indoor structured environments.

Keywords: care robot; indoor navigation; simultaneous localization and mapping; data association; obstacle detection; path planning

Received: 2017-05-17; **Accepted:** 2017-07-13; **Published online:** 2017-09-22 15:07

URL: kns.cnki.net/kcms/detail/11.2625.V.20170922.1507.007.html

Foundation item: Major Subject of Beijing Science and Technology Program (D141100003614002)

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

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

DOI: 10.13700/j.bh.1001-5965.2017.0342

基于 RS-TOPSIS 的空中目标威胁评估

杨远志, 于雷*, 周中良, 阮铖巍

(空军工程大学 航空航天工程学院, 西安 710038)



摘 要: 空中目标威胁评估是地面防空系统武器配置和资源管理的基础。针对威胁评估的实时性和人为主观性等问题,在综合粗糙集(RS)理论和逼近理想解排序法(TOPSIS)的基础上,建立了RS-TOPSIS空中目标威胁评估模型。通过RS理论对目标属性赋权,减少人为主观因素的影响与对先验信息的需求,进而结合TOPSIS分析贴近度并计算得到目标威胁程度。模型基于数据驱动,易于实现并具备良好的实时性。仿真结果表明该方法能有效评估目标威胁程度,为空中目标威胁程度的实时评估提供了一种新的工程决策方法。

关键词: 粗糙集(RS); 逼近理想解排序法(TOPSIS); 威胁评估; 地面防空; 空中目标

中图分类号: TP391

文献标识码: A

文章编号: 1001-5965(2018)05-1001-07

威胁评估作为美国三军实验室理事联席会(JDL)提出的信息融合模型^[1]中的第三级组成部分,对于地面防空系统的武器配置和资源管理具有极其重要的意义,能够显著提高多目标攻击作战效率和自身生存概率^[2]。

空中目标威胁评估是典型的多属性决策问题,对其的研究已经取得了丰硕的成果:文献[3]提出灰色聚类决策方法,分析指标变化对威胁程度的影响并确定特征指标的选取,实现对空中目标的威胁排序;文献[4]在多层次的评估模型中引入模糊认知图理论,并考虑协同目标可能不具有相同评价机制的事实,构建模糊结构实现对协同空战目标的威胁评估;文献[5-6]针对威胁评估问题中的不确定信息处理,提出直觉模糊集理论,基于熵权构建多属性威胁评估模型,建立威胁特征的直觉模糊矩阵,实现对目标威胁程度的评估。上述方法可以实现对目标的威胁评估,但灰色评估和模糊理论均涉及人为主观因素参与,会由于

决策者不同的知识结构、经验集成和个人偏好,使最终评估结果受到主观因素影响。文献[7-8]针对评估中无法动态调整因素权值的问题,结合BP神经网络和萤火虫算法,构建动态的威胁评估模型;文献[9-10]针对目标威胁评估中权重确定的不足,提出利用回归型支持向量机方法获取威胁度与指标的非线性量化关系,对目标威胁程度进行量化评估。神经网络和支持向量机方法均具有良好的容错性和评估能力,但需要大量的先验信息作为训练样本,且输入训练样本过大会导致算法复杂度增加,难以满足防空作战需求。

针对上述方法存在的不足,本文采用粗糙集-逼近理想解排序法(Rough Set-Technique for Order Preference by Similarity to Ideal Solution, RS-TOPSIS)实现对空中目标的威胁评估。RS理论基于数据驱动,通过对不确定、大样本数据的处理,消除数据冗余,进行约简属性,提取决策规则并计算属性权重。TOPSIS作为一种多属性决策的常用

收稿日期: 2017-05-22; 录用日期: 2017-06-16; 网络出版时间: 2017-08-04 16:24

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20170804.1624.007.html

基金项目: 国家自然科学基金(61472443)

* 通信作者. E-mail: yl_panda@163.com

引用格式: 杨远志, 于雷, 周中良, 等. 基于RS-TOPSIS的空中目标威胁评估[J]. 北京航空航天大学学报, 2018, 44(5): 1001-1007. YANG Y Z, YU L, ZHOU Z L, et al. Air target threat evaluation based on RS-TOPSIS[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 1001-1007 (in Chinese).

方法,能综合处理多个相互制约的决策准则,实现精细量化评估,但该方法在构造归一化矩阵时,权重系数一般由专家系统给定,影响了评估结果的准确性,限制了该方法的适用范围。因此本文提出将 RS 与 TOPSIS 相结合,由 RS 确定属性权重并代入 TOPSIS,建立基于 RS-TOPSIS 的空中目标威胁评估模型,实现对目标的威胁排序,拓宽 TOPSIS 在评估中的适用范围。

1 基于 RS 的目标属性权重计算

RS 理论的概念是由波兰数学家 Pawlak^[11]于 1982 年提出来的,基本思想是通过等价关系将信息系统进行分类,实现数据挖掘和知识发现,适用于处理不确定、不完备信息系统,具有较高的实时性,操作简单且易于实现。采用 RS 理论确定 TOPSIS 中的权重向量,建立影响空中目标威胁程度的属性集,基于样本数据构建决策信息系统,并通过数据离散和属性约简等步骤构建属性权重计算流程。

1.1 目标属性的确定

依据空中目标的进攻态势建立威胁评估模型,主要考虑目标类型、目标速度、目标航向角、目标干扰能力、目标高度、目标距离 6 个主要性能指标^[10,12]。

目标类型是指大型目标(如轰炸机)、小型目标(如巡航导弹)和直升机等空中目标;目标速度直接关系到地面防空系统的反应时间和毁伤概率;目标航向角是指目标速度在水平面的投影相对于目标和地面防空系统连线的夹角;目标干扰能力是指目标具备的破坏、削弱地面防空系统电子对抗的能力;目标高度是指目标相对于防空系统所在水平面的垂直距离;目标距离是指目标和地面防空系统的直线距离,反映了目标的攻击企图和达成攻击的可能性。

1.2 RS 基本概念

定义 1^[13] 称 $\{U, A, F, D\}$ 是决策信息系统,其中 $U = \{x_1, x_2, \dots, x_n\}$ 为对象的集合,称为论域。 $A = \{a_1, a_2, \dots, a_m\}$ 为属性的非空有限集合。 $F = \{f_l: U \rightarrow V_l (l \leq m)\}$ 为 U 与 A 之间的关系集,其中 V_l 为 $a_l (l \leq m)$ 的值域。 $D: U \rightarrow V_D$ 为决策, V_D 取有限值。每个属性子集 $a \subseteq A$ 决定了一个不可区分的关系 $i(A)$

$$i(A) = \{(x, y) \in U * U \mid \forall a \in A, a(x) = a(y)\} \quad (1)$$

关系 $i(a)$ 构成了 U 的一个分类,用 $U/i(a)$

表示。

定义 2^[14] 称 $\{U, A, F, D\}$ 是决策信息系统,对于任意 $B \subseteq A$, 记

$$R_B = \{(x_i, x_j) \mid f_l(x_i) = f_l(x_j) (a_l \in B)\} \quad (2)$$

R_B 为 U 上的等价关系,记

$$[x_i]_B = \{x_j \mid (x_i, x_j) \in R_B\} \quad (3)$$

则 $U/R_B = \{[x_i]_B \mid x_i \in U\}$ 是 U 上的划分。

同理

$$R_D = \{(x_i, x_j) \mid D(x_i) = D(x_j)\} \quad (4)$$

1.3 目标属性权重确定流程

步骤 1 将已知决策的目标各属性信息填充到决策表中,区别序号、决策与属性,形成初始决策表。

步骤 2 决策表数据离散化处理。不同目标作战意图的信息差别较大,采用等间隔法对数据进行离散。

步骤 3 构建决策辨识矩阵^[15]:

$$D_d = (D_d([x_i]_A, [x_j]_A) \mid [x_i]_A, [x_j]_A \in U/R_A) \quad (5)$$

步骤 4 求取决策约简集。 B 为决策协调集,当且仅当对于任意的 $D_d([x_i]_A, [x_j]_A) \neq \emptyset$, 有 $B \cap D_d([x_i]_A, [x_j]_A) \neq \emptyset$ (6)

且 B 的任何真子集均不为决策协调集时,称 B 为决策约简集,即可以得到核心属性及其信息系统。

步骤 5 构建条件属性 A 与决策属性 D 关于论域的分类,得到 $U/i(A)$ 与 $U/i(D)$ 。

步骤 6 依次去掉各核心条件属性,得到新的分类 $U/i(A - a_i), i = 1, 2, \dots, n$ 。

步骤 7 计算决策属性 D 对条件属性 A 的支持度^[16]

$$K = I_A(D) = \frac{|P_A(D)|}{|U|} \quad (7)$$

式中: $|U|$ 、 $|P_A(D)|$ 分别为论域 U 和正域 $P_A(D)$ 的基数,即其包含元素的个数。

正域 $P_A(D)$ 表示由那些根据属性知识判定肯定属于 x 的 U 中的元素所组成的最大集合:

$$R = \{x \in U \mid [x]_R \subseteq U\} \quad (8)$$

步骤 8 计算条件属性关于决策属性的重要性:

$$S_{A-a_i} = I_A(D) - I_{A-a_i}(D) \quad (9)$$

式中: $I_{A-a_i}(D)$ 为决策属性 D 对去除 a_i 属性后的分类的支持度。

步骤 9 计算属性的权重:

$$\omega(a_i) = S_{A-a_i} / \sum_{i=1}^n S_{A-a_i} \quad (10)$$

2 基于 RS-TOPSIS 的空中目标威胁评估模型

TOPSIS 是多属性决策问题的一种常用方法。其中心思想是先计算得到一个正理想解和一个负理想解,寻找与正理想解越近且与负理想解越远的解。将 RS 方法得到的属性权重代入 TOPSIS,构建基于 RS-TOPSIS 的空中目标威胁评估模型。具体模型构建步骤如下^[15,17]:

步骤 1 建立初始评价矩阵。

构建初始评价矩阵 $H = (h_{ij})_{m \times n}$, 其中 h_{ij} 为第 m 个评价对象的第 n 个目标属性 ($i=1, 2, \dots, m; j=1, 2, \dots, n$)。

$$H = (h_{ij})_{m \times n} = \begin{bmatrix} h_{11} & \dots & h_{1n} \\ \vdots & & \vdots \\ h_{m1} & \dots & h_{mn} \end{bmatrix}_{m \times n} \quad (11)$$

步骤 2 建立标准化决策矩阵。

对 H 矩阵中目标属性进行量纲归一化处理,消除被评价对象不同属性之间的量纲差异。

该步骤需要判定 H 矩阵中各目标属性的类型,一类是效益型属性,其属性值越大,目标函数越大,可用式(12)进行计算:

$$y_{ij} = \frac{h_{ij} - \wedge h_{ij}}{\sqrt{h_{ij} - \wedge h_{ij}}} \quad (12)$$

式中:“ \wedge ”表示合取运算;“ \vee ”表示析取运算。

另一类是成本型属性,其属性值越大,目标函数越小,可用式(13)进行计算:

$$y_{ij} = \frac{\vee h_{ij} - h_{ij}}{\sqrt{\vee h_{ij} - \wedge h_{ij}}} \quad (13)$$

通过式(12)、式(13)的计算即可得到标准化决策矩阵 $Y = (y_{ij})_{m \times n}$ 。

步骤 3 建立加权标准化决策矩阵。

对标准化决策矩阵 Y 进行加权处理,使矩阵 Y 的每一行与 RS 方法求得的权重系数 ω_j 相乘,得到加权标准化决策矩阵:

$$C = (c_{ij})_{m \times n} = \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \vdots & & \vdots \\ c_{m1} & \dots & c_{mn} \end{bmatrix} = (\omega_j y_{ij})_{m \times n} \quad (14)$$

步骤 4 计算各评价对象相对贴近度。

确定加权标准化决策矩阵 C 的各评价属性下的最大和最小的样本值,分别记为正理想解 C^+ 与负理想解 C^- :

$$C^+ = \{ \max_{1 \leq i \leq m} c_{ij} \} = \{ c_1^+, c_2^+, \dots, c_n^+ \} \quad (15)$$

$$C^- = \{ \min_{1 \leq i \leq m} c_{ij} \} = \{ c_1^-, c_2^-, \dots, c_n^- \} \quad (16)$$

计算各个评价对象到正、负理想解的距离 S_i^+ 、 S_i^- :

$$S_i^+ = \sqrt{\sum_{j=1}^n (c_{ij} - c_j^+)^2} \quad (17)$$

$$S_i^- = \sqrt{\sum_{j=1}^n (c_{ij} - c_j^-)^2} \quad (18)$$

计算各个评价对象的相对贴近度 L_i :

$$L_i = S_i^- / (S_i^- + S_i^+) \quad 0 \leq L_i \leq 1 \quad (19)$$

步骤 5 威胁等级排序。

相对贴近度即评价对象的威胁度量值,形成空中目标威胁队列矢量 T_e , 可以表示为

$$T_e = [L_1, L_2, \dots, L_i, \dots, L_m] \quad (20)$$

根据威胁等级数值大小,对威胁队列 T_e 进行降序排列,则得到威胁评估结果 T'_e :

$$T'_e = \text{sort}[T_e] = \text{sort}([L_1, L_2, \dots, L_i, \dots, L_m]) = [T'_{e1}, T'_{e2}, \dots, T'_{ej}, \dots, T'_{em}] \quad (21)$$

3 威胁评估处理流程

将 RS 理论与 TOPSIS 相结合,构建空中目标威胁评估模型,其威胁评估处理流程如图 1 所示。

基于地面防空系统侦收到的空中目标信息,建立 TOPSIS 的初始评价矩阵和 RS 理论的初始决策表,分别同时进行两路并行处理。一路输入 RS 处理模块,进行数据离散化和属性约简处理,得到核心属性并计算属性依赖度,从而得到属性权重向量。另一路输入 TOPSIS 处理模块,对矩阵进行量纲归一化处理后,与 RS 处理模块得到的权重向量相乘,得到加权标准化决策矩阵,进一步计算正、负理想解及其距离,并计算相对贴近度,最后对相对贴近度进行排序得到空中目标威胁排序结果。

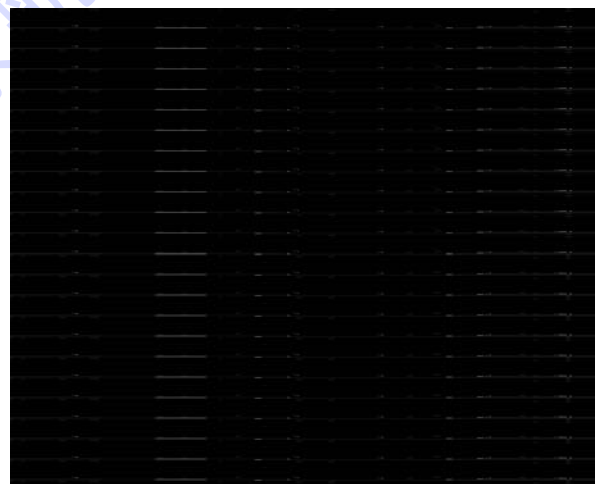


图 1 威胁评估处理流程

Fig. 1 Processing flow of threat evaluation

4 仿真分析

本节给出 RS-TOPSIS 在空中目标威胁评估中的应用实例。从文献[12]所采用的 15 组目标威胁数据库数据中选取 9 组典型数据作为本文仿真分析样本数据(见表 1)。

从表 1 可以看出,目标类型、目标干扰能力以及目标高度 3 项指标数据是定性表述,在将数据代入模型运算之前,需要先经过量化处理,以满足 RS 的数据处理方式。在文献[11,18]的基础上,充分考虑 3 项指标的不同值对于威胁程度的影响,对各指标进行分析和量化处理。

1) 目标类型。根据目标特征及其威胁程度,以 Miller^[19]的人类认知理论为量化依据,对各目标类型的定性描述语言分别作如下量化处理:小型目标(如巡航导弹)、大型目标(如轰炸机)、直升机依次量化为 3、2、1。

2) 目标干扰能力。目标具有的干扰能力越强,地面防空系统的电子对抗能力越弱,越容易受到目标的干扰,导致武器系统对目标的命中概率越低,因此将干扰能力的强、中、弱、无依次量化为 4、3、2、1。

3) 目标高度。空中目标越高,距离地面防空系统越远,可以采取对抗措施的时间越长,其威胁程度越小,因此将目标高度的高、中、低、超低依次量化为 4、3、2、1。

量化之后,根据等间隔离散的方法对数据进行处理,可以得到离散化后的数据如表 2 所示。为表示方便,将 6 个目标属性及目标威胁值分别依次记为 $a_1, a_2, a_3, a_4, a_5, a_6, D$ 。

在数据离散化的基础上,根据式(5)构造如表 3 所示的决策辨识矩阵,从而约简属性。

按照式(6)对表 3 进行约简,可以得到核心属性为 a_1, a_5, a_6 ,即目标类型、目标高度和目标距离。

条件属性以及决策对论域的分类分别为

$$U/i(A) = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

$$U/i(D) = \{1, (2, 3, 6), (4, 5), (7, 8, 9)\}$$

依次去掉属性 a_1, a_5, a_6 得到新的分类为

$$U/i(A - a_1) = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

$$U/i(A - a_5) = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

$$U/i(A - a_6) = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

由此可得 A 的 a_4 正域为

$$P_A(a_4) = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

同理:

$$P_{A-a_1}(a_4) = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

$$P_{A-a_5}(a_4) = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

$$P_{A-a_6}(a_4) = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

根据式(7)~式(10)可以计算得到核心属性的权重为

$$\omega(a_1) = \omega(a_5) = \omega(a_6) = 0.3333$$

其余属性由于不是核心属性,对其赋权重 0。从而得到各属性权重向量:

$$W = [0.3333, 0, 0, 0, 0.3333, 0.3333]$$

将表 1 的原始数据用以构建式(11)所表述的初始评价矩阵 H ,利用式(12)、式(13)对 H 矩阵进行量纲归一化处理,得到标准化决策矩阵 Y ,然后利用式(14)将矩阵 Y 与权重向量 W 相乘,得到加权标准化决策矩阵 C (见表 4)。

由式(15)、式(16)计算得到正、负理想解分别为

$$C^+ = \{0.3333, 0, 0, 0, 0.3333, 0.3333\}$$

$$C^- = \{0, 0, 0, 0, 0, 0\}$$

通过式(17)、式(18)计算正、负理想解的距离 S_i^+, S_i^- ,由式(19)得到各目标相对于 C^+ 的相对贴适度 L_i ,并根据式(20)、式(21)对相对贴适度进行排序,得到模型预测结果(见表 5)。

根据相对贴适度得到仿真结果可能分类,将可能分类与原始分类进行对比,发现除去第一组数据以外其他完全一致,这可能是由于原始样本数据本身或者其数据分类存在错误,仿真结果的正确率在 88.9% 以上,验证了本文方法的有效性。

表 1 目标威胁数据库数据^[12]

Table 1 Data of target threat database^[12]

目标	目标类型	目标速度/(m·s ⁻¹)	目标航向角/(°)	目标干扰能力	目标高度	目标距离/m	目标威胁值
t_1	大型目标	500	130	强	高	360	0.5212
t_2	大型目标	550	90	中	中	160	0.5828
t_3	大型目标	650	110	强	低	280	0.6465
t_4	小型目标	600	50	中	高	160	0.6853
t_5	小型目标	750	150	中	超低	400	0.7541
t_6	小型目标	640	180	强	中	280	0.6764
t_7	直升机	80	6	弱	高	180	0.3471
t_8	直升机	88	140	无	超低	320	0.3592
t_9	直升机	90	180	弱	低	170	0.3474

表 2 离散化数据

Table 2 Data after discretization

目 标	a_1	a_2	a_3	a_4	a_5	a_6	D
t_1	2	3	3	4	4	4	2
t_2	2	3	2	3	3	2	3
t_3	2	4	3	4	2	3	3
t_4	3	3	1	3	4	2	4
t_5	3	4	3	3	1	4	4
t_6	3	4	4	4	3	3	3
t_7	1	1	1	2	4	2	1
t_8	1	1	3	1	1	4	1
t_9	1	1	4	2	2	2	1

表 3 决策辨识矩阵

Table 3 Decision discernibility matrix

目 标	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9
t_1	\emptyset	$a_3 a_4 a_5 a_6$	$a_2 a_5 a_6$	$a_1 a_3 a_4 a_6$	$a_1 a_2 a_4 a_5$	$a_1 a_2 a_3 a_5 a_6$	$a_1 a_2 a_3 a_4 a_6$	$a_1 a_2 a_4 a_5$	$a_1 a_2 a_3 a_4 a_5 a_6$
t_2	$a_3 a_4 a_5 a_6$	\emptyset	\emptyset	$a_1 a_3 a_5$	$a_1 a_2 a_3 a_5 a_6$	\emptyset	$a_1 a_2 a_3 a_4 a_5$	$a_1 a_2 a_3 a_4 a_5 a_6$	$a_1 a_2 a_3 a_4 a_5$
t_3	$a_2 a_5 a_6$	\emptyset	\emptyset	$a_1 a_2 a_3 a_4 a_5 a_6$	$a_1 a_4 a_5 a_6$	\emptyset	$a_1 a_2 a_3 a_4 a_5 a_6$	$a_1 a_2 a_4 a_5 a_6$	$a_1 a_2 a_3 a_4 a_6$
t_4	$a_1 a_3 a_4 a_6$	$a_1 a_3 a_5$	$a_1 a_2 a_3 a_4 a_5 a_6$	\emptyset	\emptyset	$a_2 a_3 a_4 a_5 a_6$	$a_1 a_2 a_4$	$a_1 a_2 a_3 a_4 a_5 a_6$	$a_1 a_2 a_3 a_4 a_5$
t_5	$a_1 a_2 a_4 a_5$	$a_1 a_2 a_3 a_5 a_6$	$a_1 a_4 a_5 a_6$	\emptyset	\emptyset	$a_3 a_4 a_5 a_6$	$a_1 a_2 a_3 a_4 a_5 a_6$	$a_1 a_2 a_4$	$a_1 a_2 a_3 a_4 a_5 a_6$
t_6	$a_1 a_2 a_3 a_5 a_6$	\emptyset	\emptyset	$a_2 a_3 a_4 a_5 a_6$	$a_3 a_4 a_5 a_6$	\emptyset	$a_1 a_2 a_3 a_4 a_5 a_6$	$a_1 a_2 a_3 a_4 a_5 a_6$	$a_1 a_2 a_4 a_5 a_6$
t_7	$a_1 a_2 a_3 a_4 a_6$	$a_1 a_2 a_3 a_4 a_5$	$a_1 a_2 a_3 a_4 a_5 a_6$	$a_1 a_2 a_4$	$a_1 a_2 a_3 a_4 a_5 a_6$	$a_1 a_2 a_3 a_4 a_5 a_6$	\emptyset	\emptyset	\emptyset
t_8	$a_1 a_2 a_4 a_5$	$a_1 a_2 a_3 a_4 a_5 a_6$	$a_1 a_2 a_4 a_5 a_6$	$a_1 a_2 a_4 a_5 a_6$	$a_1 a_2 a_4$	$a_1 a_2 a_3 a_4 a_5 a_6$	\emptyset	\emptyset	\emptyset
t_9	$a_1 a_2 a_3 a_4 a_5 a_6$	$a_1 a_2 a_3 a_4 a_5$	$a_1 a_2 a_3 a_4 a_6$	$a_1 a_2 a_3 a_4 a_5$	$a_1 a_2 a_3 a_4 a_5 a_6$	$a_1 a_2 a_4 a_5 a_6$	\emptyset	\emptyset	\emptyset

表 4 各属性的加权标准化决策矩阵

Table 4 Weighted standard decision matrix of each attribute

目 标	a_1	a_2	a_3	a_4	a_5	a_6
t_1	0.1667	0	0	0	0	0.0556
t_2	0.1667	0	0	0	0.1111	0.3333
t_3	0.1667	0	0	0	0.2222	0.1667
t_4	0.3333	0	0	0	0	0.3333
t_5	0.3333	0	0	0	0.3333	0
t_6	0.3333	0	0	0	0.1111	0.1667
t_7	0	0	0	0	0	0.3056
t_8	0	0	0	0	0.3333	0.1111
t_9	0	0	0	0	0.2222	0.3194

表 5 模型预测结果

Table 5 Prediction results of model

目 标	正理想解距离	负理想解距离	相对贴 近度	可能 分类	原始 分类
t_1	0.4647	0.1757	0.2774	1	2
t_2	0.2777	0.3889	0.5834	3	3
t_3	0.2605	0.3240	0.5543	3	3
t_4	0.3333	0.4714	0.5858	4	4
t_5	0.3333	0.4714	0.5858	4	4
t_6	0.2777	0.3889	0.5834	3	3
t_7	0.4722	0.3056	0.3929	1	1
t_8	0.4006	0.3513	0.4672	1	1
t_9	0.3516	0.3891	0.5253	1	1

根据当前主流威胁评估算法与工程实现的参数设定^[16]来设置权重系数,将传统 TOPSIS 相对贴 近度与本文方法进行对比,可以得到如图 2 所示的结果。

对图 2 进行分析,将 RS-TOPSIS 算法与传统 TOPSIS 相比较,其相对贴 近度变化规律趋于一致,2 条曲线基本拟合,这说明采用 RS 方法得到的权重向量在一定程度上可以取代专家赋权,得到更加客观的评估结果,减少人为主观因素的影响。同时,RS 基于数据驱动,降低对先验信息的需求,拓宽 TOPSIS 的使用范围,构建更加科



图 2 RS-TOPSIS 与 TOPSIS 相对贴 近度对比

Fig.2 Comparison of relative similarity scale between RS-TOPSIS and TOPSIS

学、合理的空中目标威胁评估模型,从而提升方法的客观性和通用性。

5 结 论

1) 本文采用 RS 理论确定目标属性权重,规避人为主观因素的影响,在一定程度上可以取代专家赋权,结合 TOPSIS 构建威胁评估模型,得到空中目标的威胁度量值,实现对目标的定量评估,拓宽了 TOPSIS 的适用范围。

2) RS 理论基于数据驱动,减少人为主观因

素的影响及对先验信息的需求,且对数据具有一定的容错性,对于战场环境复杂和信息不完备的实际情况具有极强的适用性。模型可为后续的武器配置和资源管理提供支撑,可作为后续使用电子对抗或者采用火力打击决策时的理论依据。

3) 模型运行流程固化,无需人为干预,仿真结果与实际情况相符,验证了本文模型的有效性。且该模型具有较好的实时性,过程简单易于操作,对于指导空中目标威胁评估的工程实践具有极其重要的意义。

参考文献 (References)

- [1] HALL D L, LLINAS J. An introduction to multisensory data fusion[J]. Proceeding of IEEE, 1997, 85(1): 6-23.
- [2] 张堃, 王雪, 张才坤, 等. 基于 IFE 动态直觉模糊法的空战目标威胁评估[J]. 系统工程与电子技术, 2014, 36(4): 697-701.
- ZHANG K, WANG X, ZHANG C K, et al. Evaluating and sequencing of air target based on IFE and dynamic intuitionistic fuzzy sets[J]. Systems Engineering and Electronics, 2014, 36(4): 697-701 (in Chinese).
- [3] 申卯兴, 许进, 王帅. 空中目标威胁排序的灰色聚类决策方法[J]. 系统工程与电子技术, 2008, 30(9): 1721-1723.
- SHEN M X, XU J, WANG S. Grey cluster decision making way for the threat ordering of aerial targets[J]. Systems Engineering and Electronics, 2008, 30(9): 1721-1723 (in Chinese).
- [4] CHEN J, YU G H, GAO X G. Cooperative threat assessment of multi-aircrafts based on synthetic fuzzy cognitive map[J]. Journal of Shanghai Jiaotong University (Science), 2012, 17(2): 228-232.
- [5] 潘红华, 王建国, 朱森, 等. 目标威胁判断的模糊模式识别模型[J]. 兵工学报, 2004, 25(5): 576-580.
- PAN H H, WANG J M, ZHU S, et al. A fuzzy pattern recognition model for the threat evaluation of targets[J]. Acta Armamentarii, 2004, 25(5): 576-580 (in Chinese).
- [6] XU Y J, WANG Y C, MIU X D. Multi-attribute decision making method for air target threat evaluation based on intuitionistic fuzzy sets[J]. Journal of Systems Engineering and Electronics, 2012, 23(6): 891-897.
- [7] 陈洁钰, 姚佩阳, 王勃, 等. 基于结构熵和 IGSO-BP 算法的动态威胁评估[J]. 系统工程与电子技术, 2015, 37(5): 1076-1083.
- CHEN J Y, YAO P Y, WANG B, et al. Dynamic threat assessment based on structure entropy and IGSO-BP algorithm[J]. Systems Engineering and Electronics, 2015, 37(5): 1076-1083 (in Chinese).
- [8] 王改革, 郭立红, 段红, 等. 基于萤火虫算法优化 BP 神经网络的目标威胁估计[J]. 吉林大学学报(工学版), 2013, 43(4): 1064-1069.
- WANG G G, GUO L H, DUAN H, et al. Target threat assessment using glowworm swarm optimization and BP neural network[J]. Journal of Jilin University (Engineering and Technology Edition), 2013, 43(4): 1064-1069 (in Chinese).
- [9] 郭辉, 徐浩军, 刘凌. 基于回归型支持向量机的空战目标威胁评估[J]. 北京航空航天大学学报, 2010, 36(1): 123-126.
- GUO H, XU H J, LIU L. Target threat assessment of air combat based on support vector machines for regression[J]. Journal of Beijing University of Aeronautics and Astronautics, 2010, 36(1): 123-126 (in Chinese).
- [10] 李姜, 郭立红. 基于改进支持向量机的目标威胁估计[J]. 光学精密工程, 2014, 22(5): 1354-1362.
- LI J, GUO L H. Target threat assessment using improved SVM[J]. Optics and Precision Engineering, 2014, 22(5): 1354-1362 (in Chinese).
- [11] PAWLAK Z. Rough sets[J]. International Journal of Information and Computer Science, 1982, 11(5): 341-356.
- [12] 刘海波, 王和平, 沈立项. 基于 SAPSO 优化灰色神经网络的空中目标威胁估计[J]. 西北工业大学学报, 2016, 34(1): 25-32.
- LIU H B, WANG H P, SHEN L X. Target threat assessment using SAPSO and grey neural network[J]. Journal of Northwestern Polytechnical University, 2016, 34(1): 25-32 (in Chinese).
- [13] 索中英, 程嗣怡, 袁修久, 等. 优势决策信息系统规则获取方法及应用[J]. 兵工学报, 2015, 36(3): 539-545.
- SUO Z Y, CHENG S Y, YUAN X J, et al. Rule acquisition method and application of dominance decision-making information system[J]. Acta Armamentarii, 2015, 36(3): 539-545 (in Chinese).
- [14] 张文修, 仇国芳. 基于粗糙集的不确定决策[M]. 北京: 清华大学出版社, 2006: 26-27.
- ZHANG W X, QIU G F. Uncertain decision making based on rough sets[M]. Beijing: Tsinghua University Press, 2006: 26-27 (in Chinese).
- [15] 杨远志, 王红卫, 索中英, 等. 基于粗糙集-逼近理想解排序的辐射源威胁排序方法[J]. 兵工学报, 2016, 37(5): 945-952.
- YANG Y Z, WANG H W, SUO Z Y, et al. An emitter threat evaluating method based on rough set and TOPSIS[J]. Acta Armamentarii, 2016, 37(5): 945-952 (in Chinese).
- [16] ZHANG H Y, LI Y J, ZHANG K, et al. Hazard degree identification of goafs based on scale effect of structure by RS-TOPSIS method[J]. Journal of Central South University, 2015, 22: 684-692.
- [17] LIU H Y, ZHAO J M, LI D A. Single guide league ventricular late potentials extraction based on signal correlation[J]. Journal of Information and Computational Science, 2013, 10(14): 4673-4680.
- [18] 武华, 苏秀琴. 基于群广义直觉模糊软集的空袭目标威胁评估方法[J]. 控制与决策, 2015, 30(8): 1462-1467.
- WU H, SU X Q. Threat assessment of aerial targets based on group generalized intuitionistic fuzzy soft sets[J]. Control and Decision, 2015, 30(8): 1462-1467 (in Chinese).
- [19] MILLER G A. The magical number seven plus or minus two: Some limits on our capacity for processing information[J]. Psychological Review, 1994, 101(2): 343-352.

作者简介:

杨远志 男,博士研究生。主要研究方向:武器系统总体技术与作战运用。

于雷 男,博士,教授,博士生导师。主要研究方向:武器系统总体技术与作战运用。

周中良 男,博士,副教授,硕士生导师。主要研究方向:武器系统总体技术与作战运用。

阮铖巍 男,博士研究生。主要研究方向:武器系统总体技术与作战运用。

Air target threat evaluation based on RS-TOPSIS

YANG Yuanzhi, YU Lei*, ZHOU Zhongliang, RUAN Chengwei

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

Abstract: Air target threat evaluation is the foundation for weapon allocation and resource management within the ground surface air defense system. Aimed at the problems of real-time and human subjectivity for threat evaluation, an air target threat evaluation model based on RS-TOPSIS is established according to combining rough set (RS) theory and technique for order preference by similarity to ideal solution (TOPSIS). RS theory which can avoid the influence of subjective factors and the requirement for prior information is used to determine the weight of target attribute, then close degree is analyzed with TOPSIS, and threat degree of air target is obtained. The model driven with data is easy to implement and has good real-time performance. The simulation results show that this method can effectively evaluate the threat degree and thus provides a new engineering decision-making method for real-time evaluation of air target threat degree.

Keywords: rough set (RS); technique for order preference by similarity to ideal solution (TOPSIS); threat evaluation; ground surface air defense; air target

Received: 2017-05-22; Accepted: 2017-06-16; Published online: 2017-08-04 16:24

URL: kns.cnki.net/kcms/detail/11.2625.V.20170804.1624.007.html

Foundation item: National Natural Science Foundation of China (61472443)

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

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

DOI: 10.13700/j.bh.1001-5965.2017.0358

两种典型低空风切变对火箭弹弹道特性的影响



陈健伟, 王良明*, 李子杰

(南京理工大学 能源与动力工程学院, 南京 210094)

摘 要: 针对低空风切变风场特点, 基于流体力学基本原理, 建立微下击暴流和低空急流 2 种典型低空风切变的工程化模型, 以外弹道理论为基础, 将风场模型与火箭弹 6 自由度刚体弹道模型相结合。以某型尾翼火箭弹为例, 研究分析了火箭弹在主动段分别受到 2 种低空风切变影响下的弹道特性变化。仿真结果表明, 微下击暴流和低空急流对火箭弹的飞行时间、射程、侧偏、落速及攻角特性均有影响; 相比于微下击暴流, 低空急流对火箭弹弹道特性的影响更为显著; 提高风切变风场的强度和尺度, 均会增大对火箭弹弹道特性的影响程度, 且风场强度是决定风切变对火箭弹弹道特性影响的主要因素。

关键词: 火箭弹; 低空风切变; 微下击暴流; 低空急流; 弹道特性

中图分类号: TJ012.3

文献标识码: A **文章编号:** 1001-5965(2018)05-1008-10

风切变是一种广泛存在的大气现象, 其表现为大气中任意两点间风速和风向的突然变化, 包括水平风切变(同一高度)和垂直风切变(不同高度)。高度 600 m 以内发生的风切变定义为低空风切变。低空风切变现象的形成通常伴随着强烈对流性或锋面活动频繁的天气, 同时也与周边环境、地理地形等有着密切联系^[1]。若飞行器运动途经低空风切变区域, 则该区域内风速的剧烈变化会改变飞行器的受力状态, 从而影响其飞行稳定性和飞行轨迹。火箭弹在发射初始阶段, 尤其是刚脱离定向器时, 运动速度小, 抗扰动力弱, 变化剧烈的风场可引起较大的距离和方向偏差^[2]。如何考虑和较为真实地反映低空风切变对火箭弹飞行过程和弹道特性的影响, 对于火箭弹的发射控制有着重要意义。

变化风场对火箭弹飞行过程的影响一直以来是弹道学研究中的一个重要问题, 而对其中的风切变现象, 所展开的研究则较为有限。文献[3]

基于外弹道学及气象学理论建立了不同风场下弹丸的飞行动力学模型, 并分析了不同风场对弹丸弹道特性的影响。在风切变风场建模方面, 文献[4]提出了 2 种低空风切变的简化模型, 并研究了低空风切变对飞行器掠地飞行的影响。文献[5]基于 BRAMS(Brazilian developments on the Regional Atmospheric Modeling System)模型建立了垂直风切变的风廓线模型。文献[6-7]采用涡环法构建了微下击暴流的风场模型。文献[8]在文献[6-7]所建模型的基础上考虑了微下击暴流风场中湍流效应。文献[9]以欧洲天气预报中心 30 年数据为基础, 提出了基于统计数据的低空急流识别模型。文献[10]提出了一种适用于沿海地区的新型中尺度低空急流预报模型。文献[11-13]对中国低空急流风切变的识别探测及数值仿真研究进展进行了综述。

目前对于低空风切变的相关研究主要集中在低空风切变的检测识别以及危害防范方面, 而对

收稿日期: 2017-05-27; 录用日期: 2017-08-21; 网络出版时间: 2017-10-18 16:46

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20171018.1646.003.html

基金项目: 国家自然科学基金(61603191)

* 通信作者. E-mail: lmwang802@163.com

引用格式: 陈健伟, 王良明, 李子杰. 两种典型低空风切变对火箭弹弹道特性的影响[J]. 北京航空航天大学学报, 2018, 44(5): 1008-1017. CHEN J W, WANG L M, LI Z J. Influence of two typical kinds of low-level wind shear on ballistic performance of rockets[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 1008-1017 (in Chinese).

于低空风切变这一大气现象对火箭弹飞行过程的影响这一问题的探究较少。基于此,本文采用数值仿真的方法研究分析 2 种典型的低空风切变作用下,火箭弹弹道特性的变化情况,并分析不同风场特征参数对其弹道特性的影响规律,以期为火箭弹的弹道研究及发射控制提供一定的参考。

1 低空风切变风场模型

不同形式的风切变具有不同的时间和空间特征^[14],对火箭弹飞行的影响严重程度也不同。综合考虑多项因素,本文选取 2 种形成概率高、危害尺度大的典型风切变形式:微下击暴流和低空急流进行建模分析。

1.1 微下击暴流模型

作为一种常见的低空风切变现象,微下击暴流的形成通常与对流风暴有关,其形成于强对流云团内,从云底部产生一股局部性垂直向下气流,下沉后与地面碰撞变为辐散或直线型地面气流^[14],形成过程如图 1 所示。目前针对微下击暴流的工程化模型主要有偶极子模型和涡环模型^[15],为简化仿真流程、降低模型计算复杂度,在保证风场空间特性的基础上,选取涡环模型来构建微下击暴流风场模型。

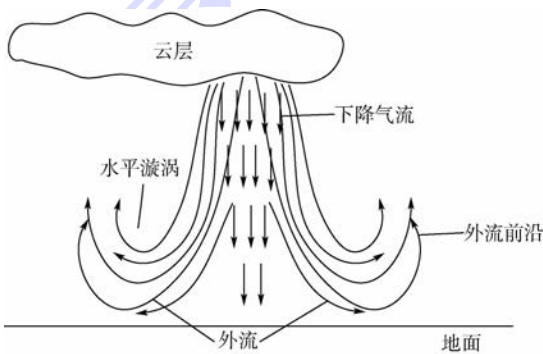


图 1 微下击暴流的形成过程

Fig. 1 Forming process of microburst

以地面坐标系为参考系,配置涡环模型,其中 $Oxyz$ 为地面坐标系,以地面上方点 O_p 为中心,建立半径为 R 的主涡环,如图 2 所示,主涡环曲线方程为

$$\begin{cases} (x - x_p)^2 + (y - y_p)^2 = R^2 \\ z = z_p \end{cases} \quad (1)$$

式中: (x_p, y_p, z_p) 为点 O_p 在地面坐标系中的坐标。

主涡环流线方程为^[6]

$$\psi_p = \frac{\Gamma}{2\pi}(r_{\max} + r_{\min})F(k) \quad (2)$$

式中: Γ 为涡环强度,由涡环中心垂直速度 $V_z(0)$

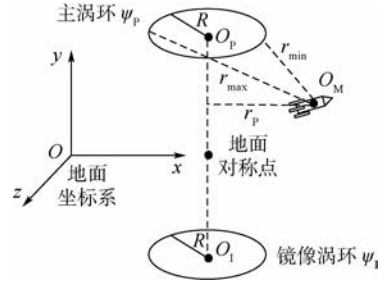


图 2 涡环模型示意图

Fig. 2 Schematic diagram of vortex ring model

和涡环半径 R 确定^[6]

$$\Gamma = 2RV_z(0) \quad (3)$$

r_{\max} 和 r_{\min} 分别为空间中任意一点 $O_M(x_M, y_M, z_M)$ 到主涡环的最大和最小距离; $F(k)$ 为椭圆积分函数, k 定义为^[6]

$$k = \left| \frac{r_{\max} - r_{\min}}{r_{\max} + r_{\min}} \right| \quad (4)$$

当 $0 \leq k \leq 1$ 时, $F(k)$ 可近似为^[6]

$$F(k) \approx \frac{0.788k^2}{0.25 + 0.75\sqrt{1 - k^2}} \quad (5)$$

由涡环流线方程可以求出涡环径向(平行于水平面)和轴向(平行于 Oz 轴)的诱导速度分别为

$$\begin{cases} v_r^p = \frac{1}{r_p} \cdot \frac{\partial \psi_p}{\partial z_p} \\ v_z^p = -\frac{1}{r_p} \cdot \frac{\partial \psi_p}{\partial r_p} \end{cases} \quad (6)$$

式中: r_p 为点 O_M 到涡环轴线的距离, $r_p = \sqrt{(x_M - x_p)^2 + (y_M - y_p)^2}$ 。由涡环径向诱导速度可得地面系下沿 Ox 轴和 Oz 轴的速度分量分别为

$$\begin{cases} v_x^p = \frac{x_M - x_p}{r_p} v_r^p \\ v_z^p = \frac{y_M - y_p}{r_p} v_z^p \end{cases} \quad (7)$$

图 2 中,在主涡环关于水平面(xOz 平面)对称的位置设置一镜像涡环,2 个涡环的中心 O_p 和 $O_1(x_p, y_p, -z_p)$ 连线垂直于 xOz 平面。这一设置是因为实际情况下,气流到达地面后沿水平方向扩散,其垂直方向风速应为 0,由于镜像涡环流线方程 ψ_1 与主涡环流线方程形式相同、符号相反: $\psi_1 = -\psi_p$,因此两涡环产生的诱导速度在地面处同值反向,相互抵消。根据镜像涡环的流线方程,推导出其在空间任意点的诱导风速,根据叠加原理结合式(6)和式(7)即可计算出点 O_M 处的合速度:

$$\begin{pmatrix} V_{wx} \\ V_{wy} \\ V_{wz} \end{pmatrix} = \begin{pmatrix} v_x^p \\ v_y^p \\ v_z^p \end{pmatrix} + \begin{pmatrix} v_x^1 \\ v_y^1 \\ v_z^1 \end{pmatrix} \quad (8)$$

则空间内任一点 O_M 的流线方程可表示为^[6]

$$\psi = \psi_p + \psi_1 = \frac{\Gamma}{2\pi} \left[\frac{0.788k^2(r_{\max} + r_{\min})}{0.25 + 0.75\sqrt{1 - k^2}} - \frac{0.788k'^2(r'_{\max} + r'_{\min})}{0.25 + 0.75\sqrt{1 - k'^2}} \right] \quad (9)$$

式中: r'_{\max} 和 r'_{\min} 分别为点 O_M 到镜像涡环的最大和最小距离; $k' = \left| \frac{r'_{\max} - r'_{\min}}{r'_{\max} + r'_{\min}} \right|$ 。

在上述建模过程中,需考虑以下情形:

1) $O_p O_1$ 连线上的点(符合 $r_p = 0$), 根据流线方程计算得到的诱导速度趋近于 $+\infty$, 不符合实际, 因此该直线上的诱导速度由涡环位函数求偏导获得

$$\begin{cases} V_{wx} = V_{wy} = 0 \\ V_{wz} = \frac{\Gamma}{2R} \cdot \frac{1}{\left[1 + \left(\frac{z_M - z_p}{R} \right)^2 \right]^{1.5}} \end{cases} \quad (10)$$

2) 涡环线上的点(符合 $r_p = R, z_M = z_p$), 根据流线方程计算得到的诱导速度也趋近于 $+\infty$, 不符合实际。因此, 文献[16]提出的方法: 以涡环线上的点为圆心, 构造半径为 r 的封闭环形圆柱, 涡量在该环形圆柱中均匀分布, 涡环线上的点诱导速度为 0, 环形圆柱外侧的点满足流线方程 ψ_p , 由涡环线至环形圆柱外侧, 诱导速度呈线性分布。流速分布情况见图 3。

当点 O_M 位于环形圆柱内时, 联合 $(x - x_p)^2 + (y - y_p)^2 = R^2$ 和直线 $O_p O_M$ 在水平面内的投影直线方程, 即可推导计算出点 M 的坐标 P_M , 进一步, 环形圆柱外侧点 N 的坐标 P_N 为

$$P_N = P_M + \frac{r}{|\overrightarrow{MO_M}|} (P_{O_M} - P_M) \quad (11)$$

式中: $\overrightarrow{MO_M}$ 为点 M 指向点 O_M 的向量, $|\overrightarrow{MO_M}|$ 为向量 $\overrightarrow{MO_M}$ 的模。已知点 N 的坐标后便可通过式(6)和式(7)计算点 N 处矢量风速 V_{wN} 。则 O_M 点风矢量为

$$V_w = \frac{|\overrightarrow{MO_M}|}{r} V_{wN} \quad (12)$$

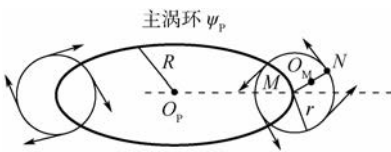


图 3 环形圆柱内风矢量示意图

Fig.3 Schematic diagram of wind vector in annular cylinder

1.2 低空急流模型

边界层急流主要指对流层下层中的强风带, 是一种在稳定的地面边界层内发生的地面逆温现象^[17]。基于流体力学中平面壁面射流原理^[18],

结合低空平均风剖面建立边界层急流模型。

根据流体力学相关理论, 某一射流在水平方向的分速度 $u(x, H)$ 和该射流的速度幅值 $u_m(x)$ 满足如下条件:

$$\frac{u(x, H)}{u_m(x)} = 1 - \tanh^2\left(k_1 \frac{H}{x}\right) \quad (13)$$

式中: H 为高度。

假设沿水平方向 x 的速度分布是均匀的, 又可写成

$$\frac{u(x, H)}{u_m(x)} = 1 - \tanh^2\left(C_s \frac{H - H_s}{H_s}\right) \quad (14)$$

式中: H_s 为表征对称分布的自由射流的最大速度高度; C_s 为射流形状参数, 表征 H_s 与射流垂直方向厚度 B 的关系。其中, 自由射流的厚度定义为垂直高度内, 大于射流速度幅值 7% 的速度分布范围, 表达式为

$$B = 4 \frac{H_s}{C_s} \quad (15)$$

在边界层的平均风廓线(指数风廓线)基础上, 加上已推导的射流速度分布关系式, 整理合并得垂直方向上低空急流速度分布的表达式为

$$u(x, H) = u_R \left(\frac{H}{H_R}\right)^{m_p} + u_s \left[1 - \tanh^2\left(C_s \frac{H - H_s}{H_s}\right) \right] \quad (16)$$

式中: u_R 为参考高度 H_R 上对应的风速; 文献[19]中根据中国风资料确定的指数 m_p 的计算公式为

$$m_p = 1/\ln\left(\left(\frac{Z_2 Z_1}{Z_0}\right)^{0.5} - 0.0403 \ln(u_R/6)\right) \quad (17)$$

式中: Z_0 为地面粗糙度; $Z_1 = H_R$; Z_2 为实际高度。

对于地面边界层中急流的矢量切变, 按照风速模型中类似的叠加方式进行建模, 可得高度 H 相对高度 H_0 处的风向偏差近似为

$$\alpha(H) = \alpha_{H_0} + \arctan\left(\frac{H - H_0}{H_T - H_0} \tan(\alpha_{H_T} - \alpha_{H_0})\right) + \alpha_{H_L} \left(1 - \tanh^2\left(C_L \frac{H - H_L}{H_L}\right) \right) \quad (18)$$

式中: H_0, H_L, H_T 分别为基准高度、急流层内风向偏差最大处高度、急流层顶高度; $\alpha_{H_0}, \alpha_{H_L}, \alpha_{H_T}$ 分别为 3 个高度对应处的风向与地转风之间的夹角; C_L 为射流形状参数, 表征 H_L 与射流垂直方向厚度 B 的关系, 其计算方法可类比于 C_s 。

1.3 仿真验证

本节通过数值仿真的方法, 获取风场模型计算结果, 并与相关资料中实测风切变风场分布情况进行对比。设置 2 种风切变风场的模型参数如

表1、表2所示。其中,表2中风向均指对应高度处风向与地转风风向的夹角,同时,仿真时假设急流层内,最大风速与最大风向偏差出现在同一高度。

按表1和表2设置的参数进行风场计算仿真,微下击暴流风场分布情况如图4、图5所示。

由图4可以看出,涡环中心轴附近区域,风

表1 微下击暴流模型参数

Table 1 Model parameters of microburst

参数	数值
涡环中心坐标/m	(800,0,600)
涡环半径/m	1000
涡核半径/m	400
中心垂直诱导风速/($m \cdot s^{-1}$)	-10

表2 低空急流模型参数

Table 2 Model parameters of low-level jet

参数	数值
地面粗糙度/m	2
基准高度/m	3.5
基准高度风速/($m \cdot s^{-1}$)	4
基准高度风向/($^{\circ}$)	0
急流层最大风速/($m \cdot s^{-1}$)	10
最大风速处高度/m	180
最大风速处风向/($^{\circ}$)	30
急流层顶高度/m	500
急流层顶风向/($^{\circ}$)	60
风速分布形状参数	0.8
风向分布形状参数	0.3

速垂直分量较大,沿水平方向,距离涡环中心轴越远风速越小。由图5可以看出。水平剖面内,涡环中心附近为风速极值区,风速沿涡环中心向四周呈辐射状衰减分布。

通过图4、图5与图6^[20]、图7^[21]实测风场数据的比较可以看出,基于涡环法的风场模型能够较为真实地反映微下击暴流的实际风速分布情况,模型空间特征合理,满足弹道仿真的要求。

图8为低空急流风场仿真结果,通过图8与图9^[17]某次低空急流实测风场数据的比较可以看出,基于平面壁面射流原理的风场模型仿真

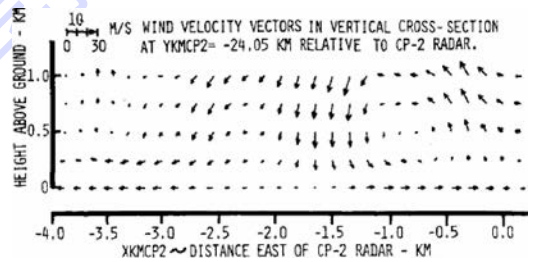


图6 1982年JAWS计划实测微下击暴流风切变风场特征图^[20]

Fig. 6 Wind field features of 1982 JAWS wind shear data showing a microburst flow pattern^[20]

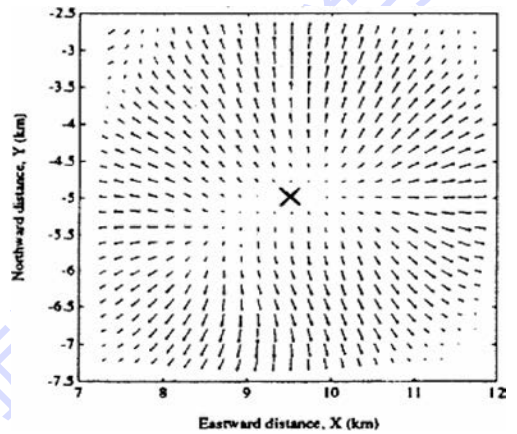


图7 NASA 1988年美国丹佛微下击暴流事故风场图^[21]

Fig. 7 Wind field of 1988 microburst event in Denver of US by NASA^[21]

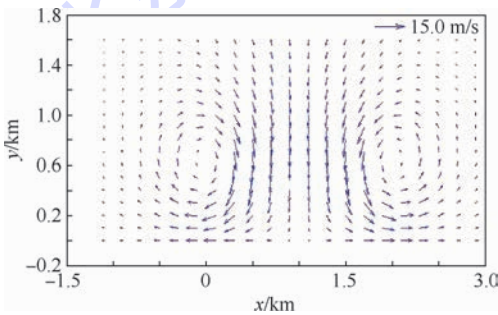


图4 涡环中心垂直剖面风矢量图

Fig. 4 Wind vector diagram of vertical section of vortex ring center

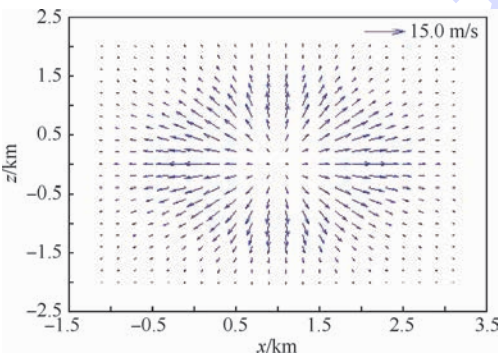


图5 $y = 100\text{ m}$ 水平剖面风矢量图

Fig. 5 Wind vector diagram of horizontal section ($y = 100\text{ m}$)

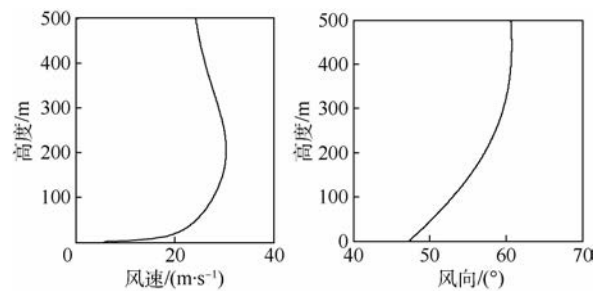


图8 低空急流风场仿真结果

Fig. 8 Wind field simulation results of low-level jet

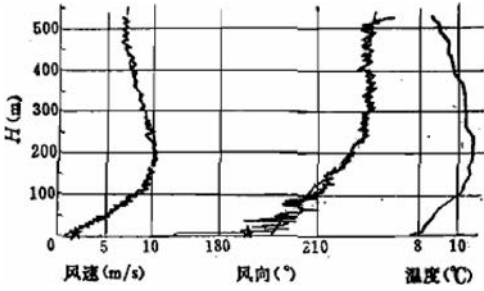


图9 低空急流实际测量数据^[17]

Fig.9 Actual measurement data of low-level jet^[17]

得出的风速风向随高度分布与实测风场在趋势上基本一致,该模型能够在一定程度上反映低空急流风切变的风场特性。

2 火箭弹刚体弹道模型

火箭弹作为自由刚体的,其运动过程由质心运动和绕质心运动两部分组成。根据外弹道学理论^[22],选取合理的参考坐标系对火箭弹进行受力分析,在建立火箭弹运动方程时,将其当作常质量物体来处理。同时,忽略燃气惯性力及其力矩,不考虑因火药燃烧造成的质心移动加速度和转动惯量变化率的影响。

忽略地球自转等次要因素的影响,在弹道坐标系内分解火箭弹质心运动矢量方程,在弹轴坐标系内分解投影得到标量化的火箭弹绕心运动方程组。以惯性坐标系为参考,火箭弹的质心运动规律满足质心运动定理:

$$m \frac{d\mathbf{v}}{dt} = \mathbf{F} \quad (19)$$

式中: t 为弹丸飞行时间; m 为弹丸质量; \mathbf{v} 为火箭弹质心速度矢量;火箭弹质心所受合力 \mathbf{F} 在弹道坐标系下的分量依次为 F_{x_2} 、 F_{y_2} 、 F_{z_2} 。将式(19)向弹道坐标系投影,则可得到质心运动方程在弹道坐标系中的标量方程,即

$$\begin{cases} \frac{d|\mathbf{v}|}{dt} = \frac{1}{m} F_{x_2} \\ \frac{d\theta_1}{dt} = \frac{1}{m v \cos \psi_2} F_{y_2} \\ \frac{d\psi_2}{dt} = \frac{1}{m |\mathbf{v}|} F_{z_2} \end{cases} \quad (20)$$

式中: θ_1 为速度高低角; ψ_2 为速度方位角。

式(20)确定了火箭弹质心速度的大小与方向变化和作用在弹上的力之间的关系。描述火箭弹质心位置坐标变化的质心运动方程如下:

$$\begin{cases} \frac{dX}{dt} = |\mathbf{v}| \cos \psi_2 \cos \theta_1 \\ \frac{dY}{dt} = |\mathbf{v}| \cos \psi_2 \sin \theta_1 \\ \frac{dZ}{dt} = |\mathbf{v}| \sin \psi_2 \end{cases} \quad (21)$$

式中: (X, Y, Z) 为火箭弹质心在地面坐标系下的坐标。弹丸的绕心运动由动量矩定理确定,即

$$\frac{d\mathbf{G}}{dt} = \mathbf{M} \quad (22)$$

式中: \mathbf{G} 为火箭弹相对于其质量中心的角动量; \mathbf{M} 为火箭弹受到合外力相对其质心的力矩。将式(22)向弹轴坐标系投影,略去部分小量,得到火箭弹绕质心运动的动力学方程和运动学方程分别为

$$\begin{cases} \frac{d\omega_\xi}{dt} = \frac{1}{C} M_\xi \\ \frac{d\omega_\eta}{dt} = \frac{1}{A} M_\eta - \frac{C}{A} \omega_\xi \omega_\zeta + \omega_\zeta^2 \tan \varphi_2 \\ \frac{d\omega_\zeta}{dt} = \frac{1}{A} M_\zeta + \frac{C}{A} \omega_\xi \omega_\eta - \omega_\zeta \omega_\eta \tan \varphi_2 \\ \frac{d\varphi_a}{dt} = \frac{1}{\cos \varphi_2} \omega_\zeta \\ \frac{d\varphi_2}{dt} = -\omega_\eta \\ \frac{d\gamma}{dt} = \omega_\xi - \omega_\zeta \tan \varphi_2 \end{cases} \quad (23)$$

式中: ω_ξ 、 ω_η 、 ω_ζ 为火箭弹转动角速度在弹轴坐标系下分量; C 和 A 分别为火箭弹的极转动惯量和赤道转动惯量; M_ξ 、 M_η 、 M_ζ 为火箭弹所受外力矩在弹轴坐标系下分量; φ_a 和 φ_2 分别为弹轴高低角和弹轴方位角; γ 为火箭弹自转角。

考虑风的影响时,速度坐标系下作用在火箭弹上的力 \mathbf{F}_2 可以表示为火箭弹相对空气的速度的函数:

$$\mathbf{F}_2 = f(\mathbf{v} - \boldsymbol{\omega}_2) \quad (25)$$

式中: $\boldsymbol{\omega}_2$ 为速度坐标系内的风速矢量,当获得地面坐标系下风速分量大小 ω_x 、 ω_z 时,可通过式(26)求得风速矢量在速度坐标系内的分量大小 ω_{x_2} 、 ω_{y_2} 、 ω_{z_2} ,进而代入式(21)进行弹道解算。

$$\begin{cases} \omega_{x_2} = \omega_x \cos \psi_2 \cos \theta_1 + \omega_z \sin \psi_2 \\ \omega_{y_2} = -\omega_x \sin \theta_1 \\ \omega_{z_2} = -\omega_x \sin \psi_2 \cos \theta_1 + \omega_z \cos \psi_2 \end{cases} \quad (26)$$

实际情况下,火箭弹主动段火药燃烧使得弹体总质量不断减小,因此按火药匀速率燃烧的情况来考虑火箭弹质量变化,即主动段内,火箭弹总质量变化微分方程为

$$\frac{dm}{dt} = m_b \quad (27)$$

式中: m_b 为火箭弹火药质量燃烧速率,主动段结束后 $m_b=0$ 。

将式(20)、式(21)和式(23)、式(24)、式(27)联立,即可得到描述火箭弹运动过程的6自由度刚体弹道方程,通过积分的方法求得火箭弹在任意时刻的运动状态和对应的弹道诸元^[23]。

3 仿真分析

将火箭弹的6自由度刚体弹道模型与2种典型风切变的风场模型相结合,探究不同风场条件下火箭弹飞行时间、射程、侧偏和落速等弹道特性的变化规律。

以某122 mm尾翼火箭弹为例进行仿真计算。其弹体基本参数和发射初始条件分别如表3、表4所示,发动机工作的时间内火箭弹质量按匀速率减小考虑,转动惯量和质心位置均相应随时间匀速率变化。仿真中,除风场分布外的其他气象条件均按照炮兵标准气象条件^[22]进行设置。

表3 火箭弹弹体参数

Table 3 Body parameters of rocket

参数	数值
弹径/m	0.122
弹长/m	2.9
发动机比冲/s	250
发动机工作时间/s	3.0
火箭弹起始质量/kg	70.0
初始赤道转动惯量/($\text{kg} \cdot \text{m}^2$)	40.05
初始极转动惯量/($\text{kg} \cdot \text{m}^2$)	0.147
火箭弹初始质心位置/m	1.533
火药质量燃烧速率/($\text{kg} \cdot \text{s}^{-1}$)	8

表4 火箭弹发射条件

Table 4 Launching conditions of rocket

参数	数值
初速/($\text{m} \cdot \text{s}^{-1}$)	40
射角/($^\circ$)	50
射向/($^\circ$)	0

3.1 微下击暴流对火箭弹弹道特性影响

为了简化模型、便于分析,本文所建立的涡环模型未考虑风场强度的时变衰减以及涡环不对称等情形。基于涡环法的风场模型中,中心垂直诱导风速反映了微下击暴流的强度,而涡环半径大小则是风场空间尺度的决定性因素,因此,选取涡环中心垂直诱导风速和涡环半径作为研究参数,进行弹道仿真。

表5、表6分别给出了不同涡环半径和中心

表5 不同涡环半径下弹道仿真结果

Table 5 Ballistic simulation results with different radii of vortex ring

涡环半径/m	飞行时间/s	射程/m	侧偏/m	落速/($\text{m} \cdot \text{s}^{-1}$)
800	101.0	33 793	-8.6	360
900	100.5	33 708	-8.4	358
1 000	100.2	33 672	-8.2	356
1 100	99.8	33 662	-8.2	355

表6 不同中心垂直诱导风速下弹道仿真结果

Table 6 Ballistic simulation results with different central vertical induction wind velocities

中心垂直诱导风速/($\text{m} \cdot \text{s}^{-1}$)	飞行时间/s	射程/m	侧偏/m	落速/($\text{m} \cdot \text{s}^{-1}$)
5	103.0	34 102	-8.6	360
10	100.2	33 672	-8.2	356
15	97.4	33 203	-8.1	354
20	94.5	32 724	-8.0	350

垂直诱导风速下火箭弹的弹道特征参数,其中:无风情况下,飞行时间为105.8 s,射程为34 478 m,侧偏为-9.3 m,落速为367 m/s。可以看出,和弹道无风情况下相比,弹丸的飞行时间缩短,射程和侧偏均减小,落速也有所减小。由表5可以看出,随着涡环半径的增大,上述弹道特征参数的减小幅度相应提高,但逐渐趋于平缓。一方面是由于火箭弹在推力作用下会在很短的时间内冲过微下击暴流的水平影响区域而继续上升;另一方面,在风场中心垂直诱导风速不变的情况下,涡环半径的增加会降低单位空间内的风切变强度。由表6可以看出,当微下击暴流风场中心垂直诱导风速增大时,风场内风切变强度增大,因而弹丸的飞行时间、射程、侧偏和落速的减小幅度也相应提高。

图10、图11分别为不同涡环半径和垂直诱导风速下火箭弹高低攻角和方向攻角的变化曲线。图中: t 为弹丸飞行时间, V 为微下击暴流风场中心垂直诱导风速, δ_1 和 δ_2 分别为高低攻角和方向攻角。

由图10、图11可以看出,由于受到中心垂直气流的作用,火箭弹穿越微下击暴流风场区域时,弹轴迅速向气流作用方向摆动,高低攻角由正变负,与无风情况下相比,弹丸高低攻角和方向攻角的震荡幅值增大,恢复稳定时间增加。增大涡环半径或中心垂直诱导风速,弹丸的攻角幅值和恢复稳定时间均相应增加。由于微下击暴流风场区域相对于弹丸射向对称,且中心气流垂直向下,因此风场对弹丸的高低攻角的影响大于对方向攻角的影响。

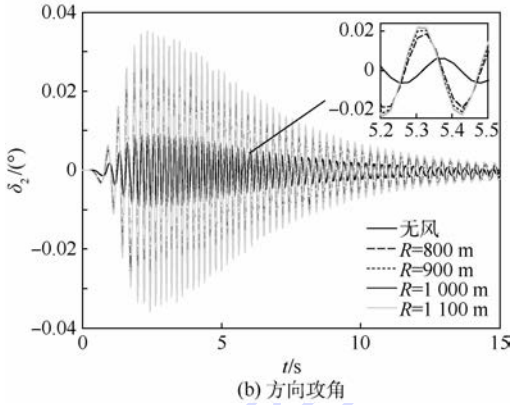
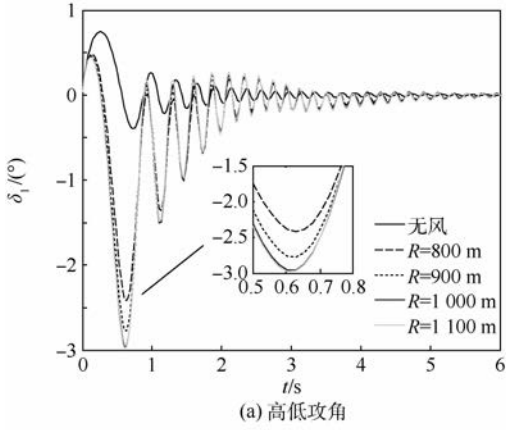


图 10 不同涡环半径下高低攻角和方向攻角曲线
Fig. 10 Pitch attack angle and yaw attack angle curves with different radii of vortex ring

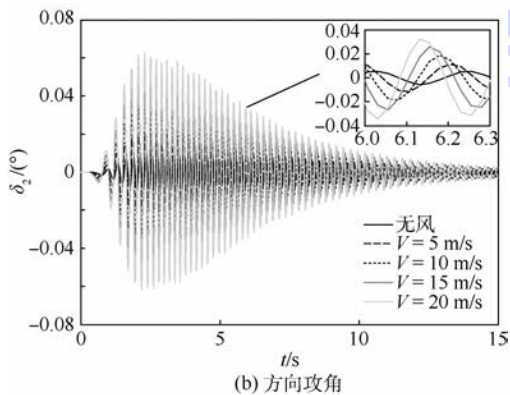
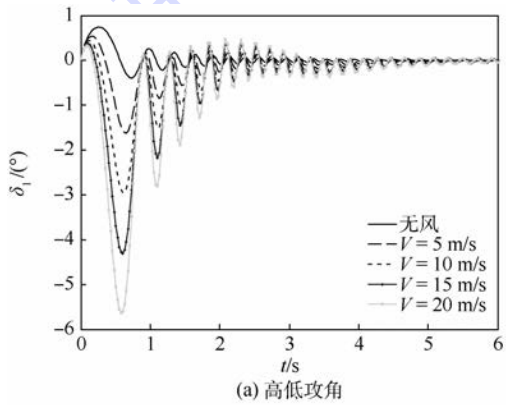


图 11 不同中心垂直诱导风速下高低攻角和方向攻角曲线
Fig. 11 Pitch attack angle and yaw attack angle curves with different central vertical induction wind velocities

由表 5、表 6 和图 10、图 11 可以看出,相比于微下击暴流风切变的尺度(涡环半径)变化,其强度(中心垂直诱导风速)变化对火箭弹弹道特性的影响更为显著。

3.2 低空急流对火箭弹弹道特性影响

低空急流是一种非定常流动,但考虑到火箭弹能进场或离场穿越低空急流风切变风场的时间很短,近似可按定常流动处理,基于 1.2 节建立的模型进行低空急流环境下的弹道仿真。

表 7、表 8 分别给出了不同急流层最大风速和急流层顶高度下火箭弹的弹道特征参数,无风情况下参数与表 5、表 6 处一致。

由表 7、表 8 可以看出,与无风条件下相比,低空急流风切变对火箭弹弹道产生的影响表现在:弹丸飞行时间缩短 10 s 以上,射程和侧偏均减小,落速降低,其中侧偏相比无风条件下增加了 2.5 km 以上。当急流层最大风速增大时,弹丸飞行时间、射程、侧偏及落速的变化量相应增大,且上述弹道参数的变化量与急流层最大风速的增量基本线性相关;而当急流层顶高度逐渐增大时,上述弹道参数变化量虽然也随之增大,但增幅逐渐减小,即火箭弹弹道参数的变化与急流层顶高度的变化不呈线性相关的关系。

由图 12、图 13 可以看出,火箭弹穿越低空急流风切变区域时飞行姿态发生了较大改变,具体表现在:由于受到初始扰动,弹丸高低攻角和方向攻角幅值增大且由正变负,在穿越风场区域约 2.5 s 的时间内,高低攻角和方向攻角保持负值,

表 7 不同急流层最大风速下弹道仿真结果

Table 7 Ballistic simulation results with different maximum wind velocities of low-level jet

最大风速/ ($m \cdot s^{-1}$)	飞行 时间/s	射程/ m	侧偏/ m	落速/ ($m \cdot s^{-1}$)
6	92.8	32 277	-2 612.1	348
10	90.9	31 914	-2 849.9	346
14	89.0	31 535	-3 078.2	344
18	87.2	31 144	-3 293.2	341

表 8 不同急流层顶高度下弹道仿真结果

Table 8 Ballistic simulation results with different height of low-level jet

层顶 高度/m	飞行 时间/s	射程/ m	侧偏/ m	落速/ ($m \cdot s^{-1}$)
400	91.7	32 052	-3 005.2	348
500	90.9	31 914	-2 849.9	346
600	90.2	31 807	-2 703.5	345
700	89.7	31 721	-2 572.9	344

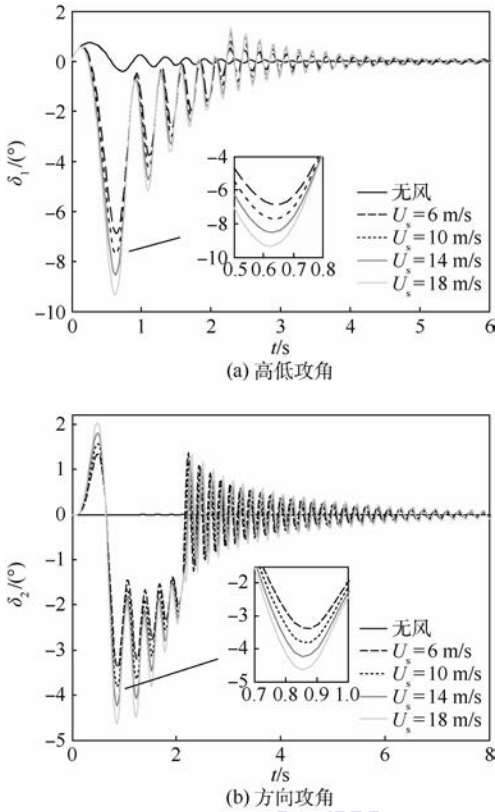


图 12 不同急流层最大风速下高低攻角和方向攻角曲线

Fig. 12 Pitch attack angle and yaw angle curves with different maximum wind velocities of low-level jet

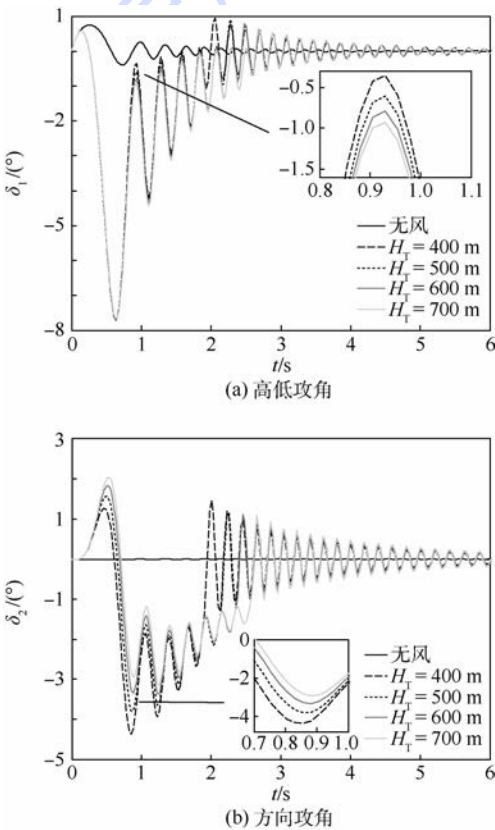


图 13 不同急流层顶高度下方向高低攻角和方向攻角曲线

Fig. 13 Pitch attack angle and yaw angle curves with different height of low-level jet

即弹轴轴线始终低于火箭弹质心的速度矢量线,离开风切变区域后,在尾翼稳定作用下,攻角振荡衰减并逐渐趋于0, U_s 为急流层最大风速。改变低空急流风切变的风场特征参数后可以看出,急流层最大风速越大,则风切变强度越高,火箭弹攻角变化幅值越高,恢复稳定的时间越长;急流层顶高度越高,代表低空急流风切变区域越大,但由于急流层最大风速不变,因此实际上单位空间内的风切变强度变弱,所以弹丸攻角变化幅值相应减小,攻角恢复稳定的时间变长。

通过上述仿真探究可以看出,低空急流风场中,一定高度范围内风速风向发生较大变化,相比微下击暴流风切变而言,低空急流风切变不仅降低了火箭弹的射程,对弹丸落点侧偏也有较大的影响,且在弹丸飞行姿态方面,低空急流对弹丸攻角变化的影响要更为显著。

4 结论

1) 基于涡环原理的微下击暴流模型和基于平面壁面射流原理的低空急流模型具有良好的三维特性,能够在一定程度上反应2种典型低空风切变的风场特性。

2) ①2种低空风切变均会缩短火箭弹的飞行时间,降低弹丸射程,减小弹丸落速,增大弹丸飞行过程中攻角幅值,增加攻角恢复稳定的时间。②相对于微下击暴流,低空急流对火箭弹各项弹道特征参数的影响更为明显,尤其在弹丸的落点侧偏方面,火箭弹在主动段穿越低空急流区域后侧偏会发生较大变化。③增加风切变的风场强度(微下击暴流的中心垂直诱导风速、低空急流的急流层最大风速)和空间尺度(微下击暴流涡环半径、低空急流层顶高度)均会提高对火箭弹弹道特性的影响,决定风切变风场对火箭弹弹道特性影响程度的主要因素为风场强度。

当火箭弹受控时,其飞行过程中穿越风切变区域时的控制动态特性变化,以及如何对火箭弹遭遇低空风切变所形成的落点偏差进行修正,都是后续需要进一步研究的内容。

参考文献 (References)

[1] 张冉,肖刚,徐悦. TAWS 低空风切变告警功能仿真验证平台设计[J]. 光电与控制,2012,19(7):51-56.
ZHANG R, XIAO G, XU Y. A simulation and verification platform for low-level windshear warning function of TAWS[J]. Electronics Optics and Control, 2012, 19(7): 51-56 (in Chinese).
[2] 王建明,潘红华,许涛. 舰载火箭炮弹道风确定探讨[J]. 弹

- 道学报,2004,16(1):11-15.
- WANG J M, PAN H H, XU T. Research on ballistic wind of ship-borne rocket [J]. *Journal of Ballistics*, 2004, 16(1): 11-15 (in Chinese).
- [3] 吴汉洲, 宋卫东, 张磊, 等. 低空风场建模与对弹丸弹道特性影响的研究 [J]. *军械工程学院学报*, 2015, 27(4): 38-42.
- WU H Z, SONG W D, ZHANG L, et al. Low-level wind model building and its influence on trajectory characteristic of projectiles [J]. *Journal of Ordnance Engineering College*, 2015, 27(4): 38-42 (in Chinese).
- [4] HUANG Z, LI X, DENG Y. The low-altitude wind shear and its influence upon hedgehopping [C] // *World Aviation Congress & Exposition*. Warrendale: SAE International, 2000: 2000-01-5570.
- [5] DIAS JR C Q, DIAS N L, FUENTES J D, et al. Convective storms and non-classical low-level jets during high ozone level episodes in the Amazon region: An ARM/GOAMAZON case study [J]. *Atmospheric Environment*, 2017, 155: 199-209.
- [6] IVAN M. A ring-vortex downburst model for flight simulations [J]. *Journal of Aircraft*, 1986, 23(3): 232-236.
- [7] 陶杨, 韩维, 吴亮. 虚拟时变微下击暴流的数值仿真方法 [J]. *应用力学学报*, 2014, 31(6): 847-852.
- TAO Y, HAN W, WU L. Numerical simulation method of virtual time-varying microburst [J]. *Chinese Journal of Applied Mechanics*, 2014, 31(6): 847-852 (in Chinese).
- [8] 刘时杰, 徐浩军, 薛源, 等. 具有湍流效应的不规则微下击暴流构建方法研究 [J]. *飞行力学*, 2014, 32(4): 303-307.
- LIU S J, XU H J, XUE Y, et al. Construction-method of irregular microburst with turbulence effect [J]. *Flight Dynamics*, 2014, 32(4): 303-307 (in Chinese).
- [9] FIEDLER S, SCHEPANSKI K, HEINOLD B, et al. A climatology of nocturnal low-level jets over North-Africa and implications for simulating mineral dust emission [C] // *EGU General Assembly Conference*, 2012: 4529.
- [10] NUNALEE C G, BASU S. Mesoscale modeling of coastal low-level jets: Implications for offshore wind resource estimation [J]. *Wind Energy*, 2014, 17(8): 1199-1216.
- [11] 赛瀚, 苗峻峰. 中国地区低空急流研究进展 [J]. *气象科技*, 2012, 40(5): 766-771.
- SAI H, MIAO J F. Research progress of low altitude jet in China [J]. *Meteorological Science and Technology*, 2012, 40(5): 766-771 (in Chinese).
- [12] 刘鸿波, 何明洋, 王斌, 等. 低空急流的研究进展与展望 [J]. *气象学报*, 2014(2): 191-206.
- LIU H B, HE M Y, WANG B, et al. Advances in low-level jet research and future prospects [J]. *Acta Meteorologica Sinica*, 2014(2): 191-206 (in Chinese).
- [13] GREENE R A. Airborne detection of low-level wind shear [J]. *Journal of Aircraft*, 1979, 16(12): 823-827.
- [14] 谷润平, 崔锋. 飞机运行低空气象环境仿真建模分析 [J]. *航空计算技术*, 2015, 45(3): 1-3.
- GU R P, CUI F. Modeling and simulation of low altitude meteorological environment for aircraft operations [J]. *Aeronautical Computing Technique*, 2015, 45(3): 1-3 (in Chinese).
- [15] 张冉. 低空风切变下大型飞机建模、危险探测与控制律研究 [D]. 上海: 上海交通大学, 2012.
- ZHANG R. Research on large aircraft modeling, hazards detection and control under low-level windshear [D]. Shanghai: Shanghai Jiao Tong University, 2012 (in Chinese).
- [16] 高振兴, 顾宏斌. 用于飞行实时仿真的微下击暴流建模研究 [J]. *系统仿真学报*, 2008, 20(23): 6524-6528.
- GAO Z X, GU H B. Research on modeling of microburst for real time flight simulation [J]. *Journal of System Simulation*, 2008, 20(23): 6524-6528 (in Chinese).
- [17] 肖业伦, 金长江. 大气扰动中的飞行原理 [M]. 北京: 国防工业出版社, 1993: 27-47.
- XIAO Y L, JIN C J. Flight principle in atmospheric turbulence [M]. Beijing: National Defense Industry Press, 1993: 27-47 (in Chinese).
- [18] 丁祖荣. 工程流体力学 [M]. 北京: 机械工业出版社, 2013: 117-120.
- DING Z R. Engineering fluid mechanics [M]. Beijing: Machinery Industry Press, 2013: 117-120 (in Chinese).
- [19] 董双林. 军用设备气候极值 空中风速: GJB 1172. 14—1991 [S]. 北京: 国防科学技术工业委员会, 1991.
- DONG S L. Climate extremes of military equipment: Wind speed: GJB 1172. 14—1991 [S]. Beijing: National Defense Science and Technology Industry Committee, 1991 (in Chinese).
- [20] ZHANG Y, SARKAR P, HU H. An experimental study of microburst-wind induced loads on a wind turbine model [C] // *AIAA Applied Aerodynamics Conference*. Reston: AIAA, 2013: 23-42.
- [21] WANKE C R, HANSMAN J R. A data fusion algorithm for multi-sensor microburst hazard assessment [C] // *AIAA Atmospheric Flight Mechanics Conference*. Reston: AIAA, 1992: 17-29.
- [22] 徐明友. 火箭外弹道学 [M]. 哈尔滨: 哈尔滨工业大学出版社, 2004: 136-146.
- XU M Y. Exterior ballistics of rockets [J]. Harbin: Harbin Institute of Technology Press, 2004: 136-146 (in Chinese).
- [23] 韩子鹏. 火箭外弹道学 [M]. 北京: 北京理工大学出版社, 2008: 210-212.
- HAN Z P. Exterior ballistics of shells and rockets [J]. Beijing: Beijing University of Technology Press, 2008: 210-212 (in Chinese).

作者简介:

陈健伟 男, 博士研究生。主要研究方向: 弹箭飞行与控制。

王良明 男, 博士, 教授, 博士生导师。主要研究方向: 飞行器制导与控制技术、新型弹箭的弹道理论与技术。

Influence of two typical kinds of low-level wind shear on ballistic performance of rockets

CHEN Jianwei, WANG Liangming*, LI Zijie

(School of Energy and Power Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

Abstract: According to the wind field features of low-level wind shear, the engineering models of microburst and low-level jet are established based on the principle of fluid mechanics. The wind field models are combined with the six-degree-of-freedom ballistics model of rockets based on exterior ballistics theory. As an example, a kind of empennage-rocket was studied on its ballistic performance of prior trajectory influenced by these two kinds of wind shear. The simulation results show that both of these two kinds of wind shear have impact on the flight time, range, side-slip, falling velocity and attack angle of the rocket. Compared to the microburst, the low-level jet has a higher effect on the ballistic performance of rocket. Improving the strength and scale of these two kinds of wind shear makes a greater impact on the ballistic performance of rocket, and meanwhile, the strength of wind field is the determining factor for wind shear to affect the ballistic performance of rocket.

Keywords: rockets; low-level wind shear; microburst; low-level jet; ballistic performance

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

DOI: 10.13700/j.bh.1001-5965.2017.0369

脉冲多普勒引信抗周期调制干扰性能研究



刘少坤, 闫晓鹏*, 栗苹, 于洪海

(北京理工大学 机电动态控制重点实验室, 北京 100081)

摘 要: 为了定量分析和表征脉冲多普勒(PD)引信的抗干扰性能,以处理增益为参量,研究了周期调制干扰信号对脉冲多普勒引信的作用效果,明确了脉冲多普勒引信抗干扰性能的影响因素。理论推导了典型周期调制干扰信号作用下,脉冲多普勒引信从接收回波到相关检测这一过程的处理增益,获得了脉冲多普勒引信抗干扰性能与影响因素之间的量化关系,既为脉冲多普勒引信抗干扰设计提供了理论依据,又为脉冲多普勒引信抗干扰性能评估提供了技术支撑。理论推导、仿真和实测结果表明:脉冲多普勒引信抗周期性调幅干扰性能弱于抗周期性调频干扰性能;抗周期调制干扰性能主要受距离门脉冲宽度、多普勒滤波器类型和带宽影响。

关键词: 脉冲多普勒(PD)引信; 抗干扰性能; 处理增益; 周期调幅干扰; 周期调频干扰

中图分类号: TJ434.1

文献标识码: A

文章编号: 1001-5965(2018)05-1018-08

脉冲多普勒(PD)引信是应用多普勒效应工作的一种脉冲体制引信,它具有脉冲体制引信所具有的测距和连续波多普勒引信所具有的测速特性,在各类精确制导武器系统中得到了广泛应用^[1-3]。但是,随着引信干扰技术的发展,引信干扰机对脉冲多普勒引信构成了日益严重的威胁^[4]。

无线电引信抗干扰性能的客观科学评价是确保引信有效使用、弹药威力有效发挥的基础与前提^[5-6]。目前,针对无线电引信抗干扰性能的研究主要包括模糊函数切割法^[7-8]、干扰信号作用下引信启动概率^[9-11]和处理增益方法等^[12-16]。文献[8]提出了一种基于模糊函数切割法的无线电引信抗干扰性能测度方法,能够测度同一体制不同工作参数下引信的抗干扰性能,但无法获取不同干扰波形对引信的干扰效果。文献[9]提出了基于启动概率的压制性干扰效果评估方法,仅适

用于评估满足高斯随机过程的压制性干扰信号干扰效果。文献[12]定量研究了脉冲多普勒引信抗压制式有源噪声干扰的能力,缺乏对引信威胁更大的周期调制干扰的抗干扰性能研究。

本文以处理增益为表征参量,理论推导了典型周期调制干扰信号作用下脉冲多普勒引信处理增益量化表达式,定量分析了脉冲多普勒引信抗调幅和调频类信号干扰性能,明确了引信抗周期调制干扰的影响因素,为引信抗干扰性能设计提供了理论依据。

1 脉冲多普勒引信工作原理

本文以脉冲对连续波相干检测脉冲多普勒引信为例进行说明,其工作原理如图1所示。脉冲产生器产生调制脉冲对振荡器产生调制,已调脉冲信号经过功率放大器,由天线向空间辐射。目标回波信号由接收天线进入接收机,经过带通滤

收稿日期: 2017-06-01; 录用日期: 2017-08-31; 网络出版时间: 2017-10-19 18:02

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20171019.1802.005.html

基金项目: 国家自然科学基金(61673066); 国防“973”计划(613196)

* 通信作者. E-mail: yanxiaopeng@bit.edu.cn

引用格式: 刘少坤, 闫晓鹏, 栗苹, 等. 脉冲多普勒引信抗周期调制干扰性能研究[J]. 北京航空航天大学学报, 2018, 44(5): 1018-1025. LIU S K, YAN X P, LI P, et al. Anti-periodic modulation jamming performance of pulse Doppler fuze[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 1018-1025 (in Chinese).

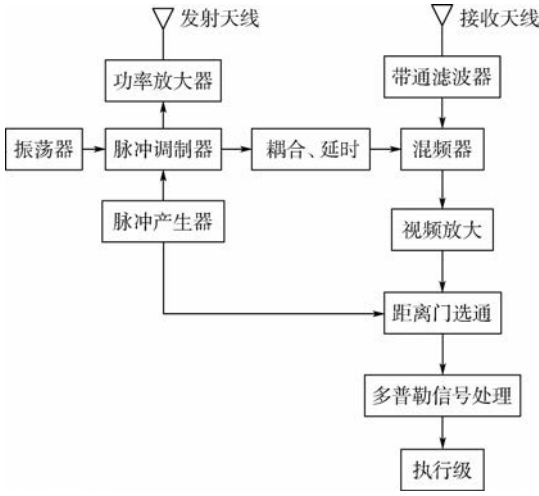


图1 脉冲多普勒引信原理框图

Fig. 1 Block diagram of PD fuze principle

波器滤除带外噪声后进入混频器,与振荡器产生的连续波本振信号进行混频,得到幅度受多普勒信号调制的视频脉冲信号,再经过视频放大器后,送入距离门选通电路。经距离门选通后的有用信号通过多普勒信号处理,提取目标信息,推动执行级输出点火信号。

脉冲多普勒引信发射脉冲信号为

$$u_a(t) = U_{Tm} \cos(\omega_0 t + \phi_0) \cdot (P_{\frac{\tau_M}{2}}(t) \otimes \sum_{N=-\infty}^{\infty} \delta(t - NT_M)) \quad (1)$$

式中: U_{Tm} 为射频脉冲幅度; ω_0 为载波频率; $P_{\frac{\tau_M}{2}}(t)$ 为宽度为 τ_M 的脉冲, t 为时间; $\delta(t)$ 为狄利克雷函数; T_M 为脉冲重复周期; \otimes 为卷积算子; ϕ_0 为脉冲多普勒引信发射信号初始相位; N 为脉冲个数。

发射信号遇目标后,部分能量被反射,目标回波信号为

$$u_c(t) = U_{Rm} \cos((\omega_0 + \omega_d)t + \phi_0) \cdot (P_{\frac{\tau_M}{2}}(t - \tau) \otimes \sum_{N=-\infty}^{\infty} \delta(t - NT_M)) \quad (2)$$

式中: U_{Rm} 为目标回波信号幅度; ω_d 为多普勒频率; τ 为回波信号延迟。

由式(2)可知,目标回波信号的平均功率为

$$P_r = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} u_c^2(t) dt = \frac{U_{Rm}^2 \tau_M}{2T_M} \quad (3)$$

回波信号经天线接收,通过等效带宽为 $[\omega_0 - \omega_{off}, \omega_0 + \omega_{off}]$ 的带通滤波器,输出信号为

$$U_{bpo} = \frac{U_{Rm}}{2} \sum_{n=-a}^a \frac{2 \sin(n\Omega_M \tau_M / 2)}{n} e^{-jn\Omega_M \tau} \cdot (\delta(\omega - n\Omega_M - \omega_0 - \omega_d) + \delta(\omega - n\Omega_M + \omega_0 + \omega_d)) \quad (4)$$

其中: ω 为角频率; n 为谐波次数; ω_{off} 为带通滤波

器截止频率,通常选定为 π/τ_M ; $\Omega_M = 2\pi/T_M$; a 为满足式(5)的最大整数,

$$\omega_0 - \omega_{off} \leq a\Omega_M + \omega_0 + \omega_d \leq \omega_0 + \omega_{off} \quad (5)$$

回波信号通过带通滤波器后,进入混频器与本振信号 $U_0 \cos(\omega_0 t + \phi_d)$ 混频,滤除高次频率分量,得到视频脉冲信号:

$$U_{ho} = \frac{KU_{Rm}U_0}{4} \sum_{n=-a}^a \frac{2 \sin(n\Omega_M \tau_M / 2)}{n} e^{-jn\Omega_M \tau} \cdot (\delta(\omega - n\Omega_M - \omega_d) e^{j\phi_0} + \delta(\omega - n\Omega_M + \omega_d) e^{-j\phi_0}) \quad (6)$$

式中: K 为混频器系数; U_0 为本振信号幅度。

视频脉冲信号经过放大处理后,送至距离门选通电路。设距离门脉宽与引信发射信号脉宽相等,则视频脉冲信号通过距离门后的频谱为

$$\frac{KU_{Rm}U_0U_f\Omega_M\Delta\tau}{4} e^{-jn\Omega_M(\tau_M+\tau_i)} \sum_{N=-a}^a \frac{\sin(N\Omega_M\Delta\tau/2)}{N\Omega_M\Delta\tau/2} \cdot (\delta(\omega - N\Omega_M - \omega_d) e^{j\phi_i} + \delta(\omega - N\Omega_M + \omega_d) e^{-j\phi_i}) \quad (7)$$

式中: ϕ_i 为距离门; U_f 为距离门信号幅值; $\Delta\tau = \tau_M + \tau_i - \tau$ 为距离门输出脉冲宽度, τ_i 为距离门信号延迟。

由式(7)可知,多普勒滤波器输出基带多普勒信号功率为

$$P_o = \frac{K^2 U_{Rm}^2 U_0^2 U_f^2 \tau_M^2}{8 T_M^2} \quad (8)$$

2 周期调制干扰信号作用下脉冲多普勒引信处理增益

干扰信号通过混频器后,信号载频 ω_j 被滤除,干扰信号载频变为原调制频率 ω_c ,为了使干扰信号能够完全通过多普勒滤波器,实际使用中通常将调制频率 ω_c 设置为近似等于引信多普勒频率 ω_d 。因此,干扰信号参数满足:

$$\begin{cases} \omega_j \approx \omega_0 \\ \omega_c = \omega_d \end{cases} \quad (9)$$

引信处理增益定义为引信输出信干比与输入信干比的比值,它是衡量引信抗干扰性能的重要指标,因此本文采用处理增益定量计算和表征脉冲多普勒引信抗周期调制干扰性能。

2.1 正弦波调幅信号干扰下引信处理增益

正弦波调幅干扰信号表达式为

$$U_j = A_j (1 + m_a \cos(\omega_c t)) \cos(\omega_j t + \phi_j) \quad (10)$$

式中: A_j 为干扰信号幅值; m_a 为调制度; ϕ_j 为干扰信号初始相位。

由式(10)得干扰信号的输入功率为

$$P_{ji} = \left(1 + \frac{m_a^2}{2}\right) \frac{A_j^2}{2} \quad (11)$$

干扰信号被引信天线接收后,首先通过一个等效带宽为 $\left[\omega_0 - \frac{\pi}{\tau_M}, \omega_0 + \frac{\pi}{\tau_M}\right]$ 的带通滤波器,滤除带外噪声。对式(10)进行傅里叶变换,得到其频谱为

$$U_j(\omega) = A_j \pi (\delta(\omega - \omega_j) e^{j\phi_j} + \delta(\omega + \omega_j) e^{-j\phi_j}) + \frac{m_a A_j \pi}{2} (\delta(\omega - \omega_j - \omega_c) e^{j\phi_j} + \delta(\omega + \omega_j + \omega_c) e^{-j\phi_j}) + \frac{m_a A_j \pi}{2} (\delta(\omega - \omega_j + \omega_c) e^{j\phi_j} + \delta(\omega + \omega_j - \omega_c) e^{-j\phi_j}) \quad (12)$$

干扰信号通过带通滤波器后,进入混频器,与本振信号 $U_0 \cos(\omega_0 t + \phi_0)$ 混频,得到中频信号:

$$U_{jd} = \frac{KA_j U_0}{2} \cos(\omega_j t - \omega_0 t + \phi_j - \phi_0) + \frac{KA_j U_0 m_a}{4} \cos(\omega_j t - \omega_0 t + \phi_j - \phi_0 + \omega_c t) + \frac{KA_j U_0 m_a}{4} \cos(\omega_j t - \omega_0 t + \phi_j - \phi_0 - \omega_c t) \quad (13)$$

式中:距离门选通脉冲表示为

$$u_{if} = U_f P_{\tau_M}^2(t - \tau_i) \otimes \delta(t - NT_M) \quad (14)$$

中频信号 U_{jd} 通过距离门后得到输出信号为

$$U_{jf} = \left(\frac{KA_j U_0}{2} \cos(\omega_j t - \omega_0 t + \phi_j - \phi_0) + \frac{KA_j U_0 m_a}{4} \cos(\omega_j t - \omega_0 t + \phi_j - \phi_0 + \omega_c t) + \frac{KA_j U_0 m_a}{4} \cos(\omega_j t - \omega_0 t + \phi_j - \phi_0 - \omega_c t) \right) \cdot (U_f P_{\tau_M}^2(t - \tau_i) \otimes \delta(t - NT_M)) \quad (15)$$

对式(15)进行傅里叶变换,并令 $\omega_j = \omega_0$ 得:

$$S_1(\omega) \otimes S_2(\omega) = \frac{2\pi\tau_M}{T_M} e^{j(\phi_j - \phi_0)} B \delta(\omega \pm n\Omega_M) + \frac{2\pi\tau_M}{T_M} e^{-j(\phi_j - \phi_0)} B \delta(\omega \pm n\Omega_M) + \frac{m_a}{2} \cdot \frac{2\pi\tau_M}{T_M} e^{j(\phi_j - \phi_0)} B \delta(\omega - \omega_c \pm n\Omega_M) + \frac{m_a}{2} \cdot \frac{2\pi\tau_M}{T_M} e^{-j(\phi_j - \phi_0)} B \delta(\omega + \omega_c \pm n\Omega_M) + \frac{m_a}{2} \cdot \frac{2\pi\tau_M}{T_M} e^{j(\phi_j - \phi_0)} B \delta(\omega + \omega_c \pm n\Omega_M) + \frac{m_a}{2} \cdot \frac{2\pi\tau_M}{T_M} e^{-j(\phi_j - \phi_0)} B \delta(\omega - \omega_c \pm n\Omega_M) \quad (16)$$

式中: $B = \frac{KA_j U_0 U_f}{4} \sum_{-\infty}^{\infty} \text{Sa}\left(\frac{n\Omega_M \tau_M}{2}\right) e^{-jn\Omega_M \tau_i}$; $S_1(\omega)$

为 u_{if} 的傅里叶变换; $S_2(\omega)$ 为 U_{jd} 的傅里叶变换。

当且仅当式(16)中 $n=0$ 时所有频率成分能够通过多普勒滤波器; $n \neq 0$ 时,对应的频率成分被多普勒滤波器滤除。因此,式(16)通过多普勒滤波器后,余项项为

$$U_{jo} = \frac{KA_j U_0 U_f \Omega_M \tau_M}{4} e^{j(\phi_j - \phi_0)} \delta(\omega) + \frac{KA_j U_0 U_f \Omega_M \tau_M}{4} e^{-j(\phi_j - \phi_0)} \delta(\omega) + \frac{KA_j U_0 U_f \Omega_M \tau_M m_a}{8} e^{j(\phi_j - \phi_0)} \delta(\omega - \omega_c) + \frac{KA_j U_0 U_f \Omega_M \tau_M m_a}{8} e^{-j(\phi_j - \phi_0)} \delta(\omega + \omega_c) + \frac{KA_j U_0 U_f \Omega_M \tau_M m_a}{8} e^{j(\phi_j - \phi_0)} \delta(\omega + \omega_c) + \frac{KA_j U_0 U_f \Omega_M \tau_M m_a}{8} e^{-j(\phi_j - \phi_0)} \delta(\omega - \omega_c) \quad (17)$$

多普勒滤波器输出信号功率为

$$P_{jo} = K^2 A_j^2 U_0^2 U_f^2 \tau_M^2 \frac{2 + m_a^2}{8T_M^2} \quad (18)$$

综上,正弦波调幅干扰信号作用下,脉冲多普勒引信处理增益为

$$G_{AS} = \frac{P_o}{P_{jo}} \cdot \frac{P_{ji}}{P_r} = \frac{T_M}{2\tau_M} \quad (19)$$

2.2 方波调幅信号干扰下引信处理增益

方波调幅干扰信号时域表达式为

$$U_j = A_j (1 + m_a \cdot m(t)) \cos(\omega_j t + \phi_j) \quad (20)$$

式中: $m(t)$ 为方波时域表达式,满足:

$$m(t) = P_{\tau_0}^2(t) \otimes \sum_{N=-\infty}^{\infty} \delta(t - NT_{j0}) \quad (21)$$

式中: T_0 为方波宽度; T_{j0} 为方波重复周期。

由式(20)和式(21)得干扰信号的输入功率为

$$P_{ji} = (1 + m_a^2) \frac{A_j^2}{2} \quad (22)$$

干扰信号通过带通滤波器后,进入混频器,与本振信号 $U_0 \cos(\omega_0 t + \phi_0)$ 进行混频,滤除低频分量,得到中频信号

$$U_{jd} = \frac{1}{2} \pi KA_j U_0 e^{j(\phi_j - \phi_0)} \delta(\omega - \omega_j + \omega_0) + \frac{1}{4} KA_j U_0 m_a \Omega_{j0} T_0 e^{j(\phi_j - \phi_0)} \delta(\omega - \omega_j + \omega_0) + \frac{1}{2} \pi KA_j U_0 e^{-j(\phi_j - \phi_0)} \delta(\omega + \omega_j - \omega_0) + \frac{1}{4} KA_j U_0 m_a \Omega_{j0} T_0 e^{-j(\phi_j - \phi_0)} \delta(\omega + \omega_j - \omega_0) +$$

$$\begin{aligned} & \frac{1}{2}K m_a A_j U_0 e^{j(\phi_j - \phi_0)} A \delta(\omega - \omega_j + \omega_0 - \Omega_{j0}) + \\ & \frac{1}{2}K m_a A_j U_0 e^{j(\phi_j - \phi_0)} A \delta(\omega - \omega_j + \omega_0 + \Omega_{j0}) + \\ & \frac{1}{2}K m_a A_j U_0 e^{-j(\phi_j - \phi_0)} A \delta(\omega + \omega_j - \omega_0 + \Omega_{j0}) + \\ & \frac{1}{2}K m_a A_j U_0 e^{-j(\phi_j - \phi_0)} A \delta(\omega + \omega_j - \omega_0 - \Omega_{j0}) \end{aligned} \quad (23)$$

式中: $A = \sin\left(\pi \frac{T_0}{T_{j0}}\right)$.

中频信号通过距离门后,进入多普勒滤波器,滤除带外频谱。令 $\frac{T_0}{T_{j0}} = \frac{1}{2}$,则通过多普勒滤波器后,干扰信号的功率为

$$P_{Jo} = A_j^2 U_0^2 U_f^2 \tau_M^2 K^2 \frac{4\pi^2 + 4\pi^2 m_a + m_a^2 \pi^2 + 8m_a^2}{16\pi^2 T_M^2} \quad (24)$$

综上,方波调幅干扰信号作用下,脉冲多普勒引信的处理增益为

$$G_{AQ} = \frac{2\pi^2(1 + m_a^2)}{4\pi^2 + 4\pi^2 m_a + \pi^2 m_a^2 + 8m_a^2} \cdot \frac{T_M}{\tau_M} \quad (25)$$

2.3 正弦波调频信号干扰下引信处理增益

正弦波调频干扰信号时域表达式为

$$U_j = A_j \cos(\omega_j t + m_f \sin \omega_m t) \quad (26)$$

式中: m_f 为调制指数; ω_m 为调制角频率。

由式(26)得正弦波调频干扰信号的功率为

$$P_{Ji} = \frac{A_j^2}{2} \sum_{n=-\infty}^{\infty} J_n^2(m_f) \quad (27)$$

式中: $J_n(m_f)$ 为第一类 n 阶贝塞尔函数。

根据贝塞尔函数的性质 $\sum_{n=-\infty}^{\infty} J_n^2(m_f) = 1$,

式(27)简化为

$$P_{Ji} = \frac{A_j^2}{2} \quad (28)$$

干扰信号通过多普勒滤波器后输出功率为

$$P_{Jo} = \frac{A_j^2 U_0^2 U_f^2 \tau_M^2 K^2}{8T_M^2} \quad (29)$$

综上,正弦波调频干扰信号作用下,脉冲多普勒引信的处理增益为

$$G_{FS} = \frac{P_o}{P_{Jo}} \cdot \frac{P_{Ji}}{P_r} = \frac{T_M}{\tau_M} \quad (30)$$

2.4 方波调频信号干扰下引信处理增益

方波调频干扰信号在时域上表现为两种频率的载波按调制周期交替出现,因此方波调频干扰信号可以用式(31)表示:

$$\begin{aligned} U_j = A_j & \left(\frac{P_{rj0}}{2} (t) \otimes \sum_{n=-\infty}^{\infty} \delta(\omega - nT_s) \cdot \cos(\omega_{j0}t) + \right. \\ & \left. \frac{P_{r_s - \tau_{j0}}}{2} (t - T_s/2) \otimes \sum_{n=-\infty}^{\infty} \delta(\omega - nT_s) \cdot \right. \\ & \left. \cos(\omega_0 t) \right) \end{aligned} \quad (31)$$

式中: T_s 为调制周期; τ_{j0} 为在一个调制周期中,频率成分 ω_{j0} 持续的时长,通常满足 $\tau_{j0} = T_s/2$ 。 $\Delta\omega = \omega_{j0} - \omega_0$ 为调制频偏。

由式(31)得方波调频信号的输入功率为

$$P_{Ji} = \frac{A_j^2}{2\pi^2} \left(\sum_{n=-\infty}^{\infty} \frac{\sin^2(n\pi\tau_{j0}/T_s)}{n^2} + \sum_{n=-\infty}^{\infty} \frac{\sin^2(n\pi(T_s - \tau_{j0})/T_s)}{n^2} \right) \quad (32)$$

当 $\tau_{j0} = T_s/2$ 时,式(32)简化为

$$P_{Ji} = \frac{8 + \pi^2}{4\pi^2} A_j^2 \quad (33)$$

干扰信号通过距离选通门,并进行多普勒滤波后,得到输出功率为

$$P_{Jo} = \frac{A_j^2 U_0^2 U_f^2 K^2 \tau_M^2}{16T_M^2} \quad (34)$$

综上,方波调频干扰信号作用下,脉冲多普勒引信的处理增益为

$$G_{FQ} = \frac{P_o}{P_{Jo}} \cdot \frac{P_{Ji}}{P_r} = \frac{\pi^2 + 8}{\pi^2} \cdot \frac{T_M}{\tau_M} \quad (35)$$

3 处理增益影响因素分析

以正弦波调幅干扰信号为例,研究影响处理增益的主要因素。正弦波调幅干扰信号由引信天线接收后,进入混频器,与本振信号混频得

$$\begin{aligned} U_{jd} = & \frac{KA_j U_0}{2} \cos(\phi_j - \phi_0) + \\ & \frac{KA_j U_0 m_a}{4} \cos(2\omega_0 t + \phi_j - \phi_0 + \omega_c t) + \\ & \frac{KA_j U_0}{2} \cos(2\omega_0 t + \phi_j - \phi_0) + \\ & \frac{KA_j U_0 m_a}{4} \cos(\phi_j - \phi_0 + \omega_c t) + \\ & \frac{KA_j U_0 m_a}{4} \cos(\phi_j - \phi_0 - \omega_c t) + \\ & \frac{KA_j U_0 m_a}{4} \cos(2\omega_0 t + \phi_j - \phi_0 - \omega_c t) \end{aligned} \quad (36)$$

滤除高频分量后的干扰信号为

$$\begin{aligned} U_{jd} = & \frac{KA_j U_0}{2} \cos(\omega_j t - \omega_0 t + \phi_j - \phi_0) + \\ & \frac{KA_j U_0 m_a}{4} \cos(\omega_j t - \omega_0 t + \phi_j - \phi_0 + \omega_c t) + \\ & \frac{KA_j U_0 m_a}{4} \cos(\omega_j t - \omega_0 t + \phi_j - \phi_0 - \omega_c t) \end{aligned} \quad (37)$$

上述信号经过距离门可得

$$S_1(\omega) \otimes S_1(\omega) = \frac{2\pi\tau_M}{T_M} e^{j(\phi_j - \phi_0)} B\delta(\omega \pm n\Omega_M) + \frac{2\pi\tau_M}{T_M} e^{-j(\phi_j - \phi_0)} B\delta(\omega \pm n\Omega_M) + \frac{m_a}{2} \cdot \frac{2\pi\tau_M}{T_M} e^{j(\phi_j - \phi_0)} B\delta(\omega - \omega_c \pm n\Omega_M) + \frac{m_a}{2} \cdot \frac{2\pi\tau_M}{T_M} e^{-j(\phi_j - \phi_0)} B\delta(\omega + \omega_c \pm n\Omega_M) + \frac{m_a}{2} \cdot \frac{2\pi\tau_M}{T_M} e^{j(\phi_j - \phi_0)} B\delta(\omega + \omega_c \pm n\Omega_M) + \frac{m_a}{2} \cdot \frac{2\pi\tau_M}{T_M} e^{j(\phi_j - \phi_0)} B\delta(\omega - \omega_c \pm n\Omega_M) \quad (38)$$

式(38)单边谱如图2所示。

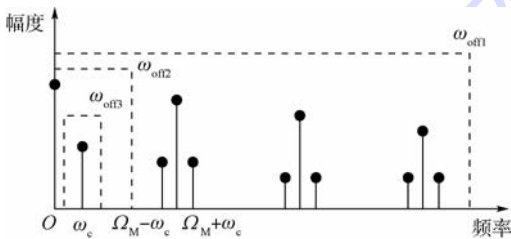


图2 距离门输出单边谱

Fig. 2 One-sided spectrum of range gate output

当多普勒滤波器截止频率为 ω_{om1} 时,干扰信号大量频率分量能够通过多普勒滤波器。随着多普勒滤波器截止频率降低至 ω_{om2} 时,仅有基带($n=0$)处的干扰信号和直流分量能够通过多普勒滤波器,干扰信号功率得到有效抑制。当多普勒滤波器由低通滤波器改变为带通滤波器,且带宽减小至 ω_{om3} 时,仅有基带干扰信号能够通过多普勒滤波器,直流分量被滤除,干扰信号功率进一步得到抑制。此时,干扰信号输出功率为

$$P_{J_0} = \frac{A_j^2 U_0^2 \tau_f^2 K^2 m_a^2}{32 T_M^2} \quad (39)$$

表1 周期调制干扰信号作用下脉冲多普勒引信处理增益仿真结果

Table 1 Simulation results of PD fuze processing gain under interference of periodic modulation jamming signal

干扰信号 样式	干扰波性参数			处理增益		表达式
	载波幅值/V	调制周期/kHz	调制频偏/MHz	理论值/dB	仿真值/dB	
正弦波调幅	0.8165	40		20.000	20.525	$\frac{T_M}{2\tau_M}$
方波调幅	0.7071	40		19.114	19.550	$\frac{2\pi^2(1+m_a^2)}{4\pi^2+4\pi^2m_a+\pi^2m_a^2+8m_a^2} \cdot \frac{T_M}{\tau_M}$
正弦波调频	1	40	20	23.010	23.319	$\frac{T_M}{\tau_M}$
方波调频	1.0508	40	20	25.590	24.901	$\frac{\pi^2+8}{\pi^2} \cdot \frac{T_M}{\tau_M}$

此时的引信处理增益为

$$G_{AS} = \frac{8+4m_a^2}{m_a^2} \cdot \frac{T_M}{\tau_M} \quad (40)$$

由上述理论分析可知,引信处理增益主要受多普勒滤波器类型和带宽、距离门脉冲宽度影响。通过替换多普勒滤波器为带通滤波器,同时减小多普勒滤波器带宽和距离门脉冲宽度,可以抑制干扰信号输出功率,提高引信处理增益和抗干扰性能。

脉冲多普勒引信距离门脉冲宽度通常与调制脉冲宽度相同,调制脉冲宽度要求取式(41)的2个分式中较小值^[17]:

$$\begin{cases} \tau_M \leq 2\sigma_R/c \\ \tau_M \leq 2R_{min}/c \end{cases} \quad (41)$$

式中: σ_R 为距离分辨率; R_{min} 为引信最小作用距离; c 为光速。调制脉冲宽度越窄,测距精度越高^[2],因此,减小距离门脉冲宽度在提升引信抗干扰性能的同时不会降低引信的探测性能。

4 仿真与讨论

根据脉冲多普勒引信工作原理,基于MATLAB建立了脉冲多普勒引信干扰模型,针对6种干扰源进行仿真,对理论推导结果进行了验证。仿真参数设置如下:脉冲重复周期 $T_M=4000$ ns;发射脉冲宽度 $\tau_M=20$ ns;多普勒频率 $f_d=40$ kHz;多普勒滤波器截止频率 $F_d=66.7$ kHz;仿真距离30 m;预定炸高15 m;距离门延迟 $\tau_i=100$ ns;距离门脉冲宽度与发射信号脉冲宽度相等;信号幅值1 V;干扰信号输入功率0.5 W,调制度为1。

在上述工作参数下,引信回波信号输入功率为0.0025 W,距离门选通后输出功率为 3.125×10^{-6} W。典型周期调制干扰信号作用下,脉冲多普勒引信处理增益,如表1所示。

图3、图4分别给出了干扰信号参数不变的情况下,引信处理增益与引信距离门脉冲宽度、多普勒低通滤波器带宽的量化关系。图5为将多普勒低通滤波器替换为多普勒带通滤波器后,滤波器带宽与处理增益量化关系。如图3所示,周期调制干扰信号作用下处理增益在距离门脉冲宽度由20 ns增加至100 ns的过程中均降低了约7 dB。如图4所示,引信处理增益受多普勒滤波器带宽变化影响较为明显。当滤波器带宽由50 kHz增加至350 kHz时,各类干扰信号作用下的引信处理增益减小值均在4.5 dB以上。如图5所示,在多普勒滤波器带宽不变,滤波器类型由低通滤波器变为带通滤波器的情况下,不同干扰信号作用下处理增益变化差距较大,其中方波调幅作用下处理增益增加值达到10 dB,而正弦波调频信号作用下处理增益仅增加0.1 dB。

综上,脉冲多普勒引信抗干扰性能的影响因素主要包括距离门选通和多普勒滤波。减小距离门脉冲宽度和多普勒滤波器带宽,并将多普勒低通滤波器替换为带通滤波器可以在不影响引信信号输出功率的情况下减小干扰信号通过距离门和滤波器的功率,进而提高引信抗干扰性能。

表2给出了不同距离门脉冲宽度条件下使引信启动的最小干扰功率实测值。由表2可以看出,随着距离门脉冲宽度的增加,引信启动的最小干扰功率逐渐减小。表3给出了不同滤波器带宽

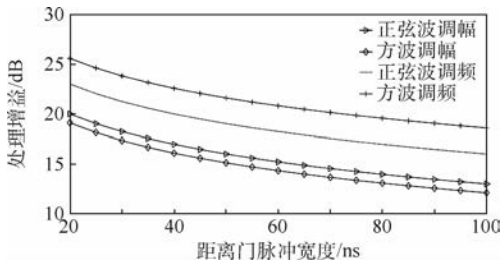


图3 距离门脉冲宽度与处理增益量化关系
Fig.3 Quantitative relationship between range gate pulse width and processing gain

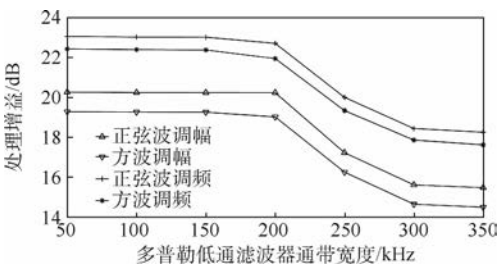


图4 多普勒低通滤波器通带宽度与处理增益量化关系
Fig.4 Quantitative relationship between Doppler low pass filter bandwidth and processing gain

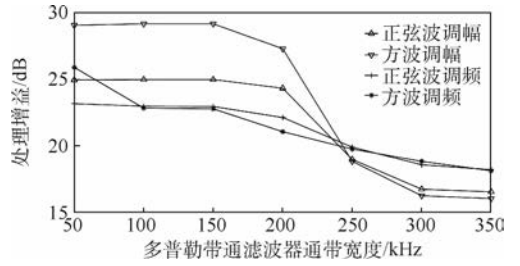


图5 多普勒带通滤波器通带宽度与处理增益量化关系
Fig.5 Quantitative relationship between Doppler band pass filter bandwidth and processing gain

表2 不同距离门脉冲宽度条件下使引信启动最小干扰功率实测值

Table 2 Measured value of minimum interference power of fuze starting at different range gate pulse width

干扰信号样式	最小干扰功率			dBm
	400 ns 脉冲宽度	800 ns 脉冲宽度	1200 ns 脉冲宽度	
方波调幅	-13	-16	-17	
方波调频	-12	-13	-15	

表3 不同多普勒滤波器通带宽度条件下使引信启动最小干扰功率实测值

Table 3 Measured value of minimum interference power of fuze starting at different Doppler filter bandwidth

干扰信号样式	最小干扰功率			dBm
	10 kHz 通带宽度	20 kHz 通带宽度	30 kHz 通带宽度	
方波调幅	-16	-23	-25	
正弦波调幅	-20	-21	-23	
方波调频	-13	-14	-14	
正弦波调频	-14	-14	-15	

条件下使引信启动的最小干扰功率实测值。由表3可以看出,随着多普勒滤波器带宽的增加,引信启动的最小干扰功率逐渐减小。在相同多普勒滤波器带宽条件下,调幅类干扰信号作用下,引信启动的最小干扰功率小于调频类干扰信号作用下,引信启动的最小干扰功率。

理论推导、仿真和实测结果表明:

1) 仿真和实测结果与理论推导结果吻合,验证了脉冲多普勒引信在周期调制干扰信号作用下,脉冲多普勒引信处理增益量化表达式的正确性。

2) 干扰信号功率相同的情况下,脉冲多普勒引信抗调幅信号干扰能力弱于抗调频干扰能力。

3) 干扰信号功率保持不变的情况下,随着距离门脉冲宽度增加,脉冲多普勒引信处理增益相应降低,不同干扰信号作用下,引信处理增益受距离门脉冲宽度变化影响的变化趋势一致。

4) 干扰信号功率保持不变的情况下,随着多普勒滤波器带宽增加,脉冲多普勒引信处理增益相应降低。不同干扰信号作用下,引信处理增益变化受多普勒滤波器带宽变化影响的程度不同,但变化趋势一致。

5) 干扰信号功率保持不变的情况下,将多普勒低通滤波器替换为带通滤波器,脉冲多普勒引信处理增益得到明显提升。随着带通滤波器带宽增加,脉冲多普勒引信处理增益降低。不同干扰信号作用下,引信处理增益变化受多普勒滤波器带宽变化影响的程度不同,但变化趋势一致。

5 结论

本文以引信处理增益为表征参量,定量研究了脉冲多普勒引信抗周期调制干扰性能。

1) 获取了脉冲多普勒引信抗干扰性能与其影响因素之间的量化关系,既为脉冲多普勒引信抗干扰性能的优化设计提供了理论依据,又为脉冲多普勒引信抗干扰性能评估提供了技术支撑。

2) 理论推导、仿真和实测结果表明脉冲多普勒引信具有较强的抗周期调制干扰性能,抗干扰性能与引信距离门脉冲宽度、多普勒滤波器类型和带宽有关。

参考文献 (References)

- [1] HARTLEY J D. A method for suppression of pulsed interference in a pulse Doppler radar [C] // 2008 International Conference on Radar. Piscataway, NJ: IEEE Press, 2008: 265-270.
- [2] 赵惠昌. 无线电引信设计原理与方法 [M]. 北京: 国防工业出版社, 2012: 147-161.
ZHAO H C. Fundamentals and methodology of radio fuze [M]. Beijing: National Defense Industry Press, 2012: 147-161 (in Chinese).
- [3] 隋鉴, 李国林. 针对脉冲多普勒引信的二维多假目标干扰 [J]. 电光与控制, 2016, 23 (3): 7-10.
SUI J, LI G L. Two-dimensional multiple false targets jamming against pulse Doppler fuze [J]. Electronic Optics & Control, 2016, 23 (3): 7-10 (in Chinese).
- [4] 张彪, 闫晓鹏, 栗苹, 等. 基于支持向量机的无线电引信抗扫频式干扰研究 [J]. 兵工学报, 2016, 37 (4): 635-640.
ZHANG B, YAN X P, LI P, et al. Research on anti-frequency sweeping jamming of radio fuze based on support vector machine [J]. Acta Armamentarii, 2016, 37 (4): 635-640 (in Chinese).
- [5] ZHANG C C, GE J J, SHENG J T. Analysis of ground clutter characteristics of airborne forward looking pulse Doppler radar in high PRF [C] // 2006 International Conference on Radar. Piscataway, NJ: IEEE Press, 2006: 1-3.
- [6] SKOLNIK M I. Radar handbook [M]. 2nd ed. New York: McGraw-Hill, Inc., 1990: 226-243.
- [7] 周新刚, 赵惠昌, 涂友超. 脉冲多普勒引信抗干扰性能评判方法和仿真 [J]. 系统仿真学报, 2011, 23 (1): 207-211.
ZHOU X G, ZHAO H C, TU Y C. ECCM evaluation and simulation of pulse Doppler fuze [J]. Journal of System Simulation, 2011, 23 (1): 207-211 (in Chinese).
- [8] 赵惠昌, 周新刚. 基于模糊函数切割法的线性调频连续波引信抗干扰性能测度 [J]. 兵工学报, 2009, 30 (12): 1591-1595.
ZHAO H C, ZHOU X G. Anti-jamming performance evaluation of linear frequency modulated continuous wave fuze based on ambiguity function incision [J]. Acta Armamentarii, 2009, 30 (12): 1591-1595 (in Chinese).
- [9] 涂友超. 典型伪码体制引信参数提取及干扰效果研究 [D]. 南京: 南京理工大学, 2010.
TU Y C. Study on parameter extraction and jamming effective of typical pseudo-random code fuze [D]. Nanjing: Nanjing University of Science and Technology, 2010 (in Chinese).
- [10] 熊金柱, 曹玉瑞, 涂友超. 基于启动概率的伪码调相引信干扰效果研究 [J]. 河南大学学报, 2010, 40 (5): 460-463.
XIONG J Z, CAO Y R, TU Y C. Study on jamming effect of pseudo-random binary-phase-coded fuze based on starting probability [J]. Journal of Henan University, 2010, 40 (5): 460-463 (in Chinese).
- [11] 涂友超, 赵惠昌, 周新刚. 噪声调频干扰下伪码调相引信启动概率分析 [J]. 南京理工大学学报, 2011, 35 (2): 252-256.
TU Y C, ZHAO H C, ZHOU X G. Analysis on starting probability of pseudo-random binary-phase-coded fuze under noise FM jamming [J]. Journal of Nanjing University of Science and Technology, 2011, 35 (2): 252-256 (in Chinese).
- [12] 李泽, 栗苹, 郝新红, 等. 脉冲多普勒引信抗有源噪声干扰性能研究 [J]. 兵工学报, 2015, 36 (6): 1001-1008.
LI Z, LI P, HAO X H, et al. Anti-active noise jamming performance of pulse Doppler fuze [J]. Acta Armamentarii, 2015, 36 (6): 1001-1008 (in Chinese).
- [13] 涂友超, 赵惠昌, 周新刚. 伪码体制引信抗瞄准式噪声调幅干扰性能研究 [J]. 南京理工大学学报, 2008, 32 (3): 350-355.
TU Y C, ZHAO H C, ZHOU X G. Performance of anti-noise AM spot jamming of pseudo-random code fuze [J]. Journal of Nanjing University of Science and Technology, 2008, 32 (3): 350-355 (in Chinese).
- [14] 刘己斌, 赵惠昌, 陆建伟. 几种伪码体制引信的抗噪声性能分析 [J]. 兵工学报, 2005, 26 (1): 20-24.
LIU J B, ZHAO H C, LU J W. Performance analysis concerning anti-noise for several pseudo-random code fuzes [J]. Acta Armamentarii, 2005, 26 (1): 20-24 (in Chinese).
- [15] 熊刚, 杨小牛, 赵惠昌. 伪码调相与正弦调频复合引信抗干扰性能分析 [J]. 现代雷达, 2007, 29 (12): 12-17.
XIONG G, YANG X N, ZHAO H C. Performance analysis of anti noise of pseudo-random PM and sine frequency modulation combined fuze [J]. Modern Radar, 2007, 29 (12): 12-17 (in Chinese).
- [16] 闫岩, 崔占忠. 超宽带无线电引信抗干扰性能研究 [J]. 兵工学报, 2010, 31 (1): 13-17.

YAN Y, CUI Z Z. Anti-jamming performance of ultra wideband radio fuze[J]. Acta Armanentarii, 2010, 31(1): 13-17 (in Chinese).

[17] 梁棠文. 防空导弹引信设计及仿真技术[M]. 北京: 宇航出版社, 1995: 233-235.

LIANG T W. Design and simulation technology of air defense missile fuze[M]. Beijing: China Astronautic Publishing House, 1995: 233-235 (in Chinese).

作者简介:

刘少坤 男, 博士研究生。主要研究方向: 无线电引信对抗

技术。

闫晓鹏 男, 博士, 副教授。主要研究方向: 引信信息对抗技术、引信电磁环境与效应、引信信号处理技术。

栗苹 女, 博士, 教授。主要研究方向: 近感探测技术、引信信息对抗技术、引信电磁环境与效应。

于洪海 男, 硕士研究生。主要研究方向: 无线电引信信号处理。

Anti-periodic modulation jamming performance of pulse Doppler fuze

LIU Shaokun, YAN Xiaopeng*, LI Ping, YU Honghai

(Science and Technology on Electromechanical Dynamic Control Laboratory, Beijing Institute of Technology, Beijing 100081, China)

Abstract: The jamming effect of periodic modulation signal on pulse Doppler (PD) fuze is studied based on the parameter of processing gain to quantitatively analyze and characterize the anti-jamming performance of PD fuze and obtain the factor that affects the anti-jamming performance of PD fuze. The processing gain of the procedure from echo reception to correlation detection under the interference of typical periodic modulation jamming style is derived. The quantitative relation between the anti-jamming performance of PD fuze and the characteristic parameter that affects the anti-jamming performance is obtained. It provides a theoretical basis for the improvement of interference performance of PD fuze and technical support for the interference performance assessment of PD fuze. Theoretical derivation, simulation and measurement results indicate that the anti-jamming performance of PD fuze to periodic amplitude modulation is worse than its anti-jamming performance to periodic frequency modulation, and the anti-periodic modulation jamming performance mainly depends on range gate pulse width, Doppler filter bandwidth and filter type.

Keywords: pulse Doppler (PD) fuze; anti-jamming performance; processing gain; periodical amplitude modulation jamming; periodical frequency modulation jamming

Received: 2017-06-01; **Accepted:** 2017-08-31; **Published online:** 2017-10-19 18:02

URL: kns.cnki.net/kcms/detail/11.2625.V.20171019.1802.005.html

Foundation items: National Natural Science Foundation of China (61673066); National Security Program on Key Basic Research Project of China (613196)

* **Corresponding author.** E-mail: yanxiaopeng@bit.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2017.0370

基于标校源辅助的不相交多目标到达时差定位

秦兆涛¹, 王俊¹, 陶磊岩², 魏少明^{1,*}

(1. 北京航空航天大学 电子信息工程学院, 北京 100083; 2. 北京遥感设备研究所, 北京 100854)



摘 要: 鉴于定位站位置误差会极大地降低多站无源定位的目标定位精度, 提出了一种标校源辅助的不相交多目标到达时差(TDOA)闭式定位算法。该算法首先使用标校源减小定位站位置误差, 并估计对应的误差统计特性, 然后使用更新的定位站位置, 利用两步加权最小二乘(TS-WLS)算法实现不相交多目标的高精度 TDOA 定位。通过克拉美罗界(CRLB)推导, 从理论上分析了该闭式定位算法的定位性能; 通过仿真实验, 验证了标校源校正技术可提高对多目标的定位精度, 并且在较小的 TDOA 观测误差和定位站位置误差下, 对多目标的定位性能可以达到 CRLB。该算法不需要初始值估计和迭代运算, 同时避免了定位站和目标位置的联合估计, 计算量较小。

关键词: 标校源; 不相交多目标; 到达时差(TDOA); 定位精度; 克拉美罗界(CRLB)

中图分类号: TN957.51

文献标识码: A 文章编号: 1001-5965(2018)05-1026-11

无源定位技术在信号处理领域应用广泛, 不仅在雷达^[1]、声呐^[2]和导航等传统领域持续发挥作用, 而且在无线通信、传感器网络^[3-4]等新兴商业领域中得到了应用。分布式定位站通过接收未知辐射源信号, 可提取与目标位置有关的观测量信息, 如到达角度(Angle of Arrival, AOA)、到达时差(Time Difference of Arrival, TDOA)^[5]、到达频差(Frequency Difference of Arrival, FDOA)^[6]和信号到达强度(Received Signal Strength, RSS)^[7]等信息, 利用一种或多种观测量信息^[8], 通过求解观测量信息与目标位置参数的几何定位方程, 可以获得目标的位置估计。

该几何定位方程对目标位置参数非线性, 而且非凸, 所以不易求解, 很多文献都对该问题进行了深入研究。迭代定位算法^[9-10]作为传统的无源定位算法, 通过泰勒级数展开, 在较小的观测量噪

声下可以达到克拉美罗界(Cramer-Rao Lower Bound, CRLB), 但需要较准确的初始值估计, 计算量大而且存在定位发散问题。球面内插(Spherical Interpolation, SI)算法^[11]和球面相交(Spherical Intersection, SX)算法^[12]是2种闭式算法, 不需迭代运算而且计算量小, 但定位性能达不到最优。Chan和Ho提出了经典的两步加权最小二乘(Two-Step Weighted Least Squares, TS-WLS)算法^[13], 该算法可以得到目标位置的闭式解析解, 不需初始值和迭代运算, 在较小的观测量噪声下能达到CRLB。Einemo和So^[14]通过改进TS-WLS算法, 将其应用在分布式MIMO^[14-16]雷达中, 借助外辐射源得到了非辐射目标的位置估计。然而这些研究都是建立在定位站位置准确已知的基础上, 实际中难以实现。文献[17]考虑了定位站位置误差的存在, 发现定位站位置误差会对目标

收稿日期: 2017-06-05; 录用日期: 2017-09-08; 网络出版时间: 2017-09-27 17:13

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20170927.1713.001.html

基金项目: 国家自然科学基金(61501012, 61671035)

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

引用格式: 秦兆涛, 王俊, 陶磊岩, 等. 基于标校源辅助的不相交多目标到达时差定位[J]. 北京航空航天大学学报, 2018, 44(5): 1026-1036. QIN Z T, WANG J, TAO L Y, et al. TDOA localization of multiple disjoint sources based on a calibration emitter[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 1026-1036 (in Chinese).

定位性能有较大的影响,通过将定位站位置统计分布信息引入到加权矩阵中,提出一种存在定位站位置误差下的时差定位算法,改善了目标定位精度。在定位站与目标具有相对运动时,又提出了一种存在定位站位置和速度误差下的时频差联合定位算法^[18],提高了目标的位置和速度估计精度。即使这样,存在定位站位置误差下的定位与不存在定位站位置误差下的定位相比,定位性能也相差甚大。标校源的使用有效缓解了这一问题,文献[19]使用位置已知的标校源,通过测量来自该标校源的时差观测量,结合预测时差值,可估计定位站的位置偏差,从而改善定位站位置,极大地提高了目标的定位精度。

上述文献只考虑了对单目标的定位研究,对多目标定位却研究很少。文献[20]在定位站位置不确定情况下,提出了一种渐进有效的定位不相交多目标时差定位算法,其中不相交性可以来自时间、频率或两者都有,从而消除模糊便于估计每个目标的观测量信息,该算法可以对多目标进行有效定位,但需要目标位置假设,而且需要对定位站和目标进行联合估计,增大了运算量。在此基础上,文献[21]提出了一种新的定位不相交多目标的时差定位算法,该算法不需要初始值估计,避免了定位站和目标的联合估计,降低了计算量,而且在较低的测量噪声和定位站位置误差下,对远、近场目标的定位性能均能达到 CRLB。如前所述,定位站位置误差对单目标的定位性能有较大影响,对多目标定位而言,也具有相同的结论。

本文研究不相交多目标的时差定位问题,通过使用位置已知的标校源改善定位站位置,减小定位站位置不确定性对多目标定位性能的影响,提高对多目标的定位精度。首先对标校源辅助的多目标时差定位进行 CRLB 性能分析,并与未使用标校源辅助的多目标时差定位 CRLB 对比,发现标校源的使用理论上可以大幅提高多目标的定位精度。因此,本文提出了一种使用标校源的多目标时差定位闭式算法,该算法基于加权最小二乘估计,分为 2 个阶段:第 1 阶段使用来自标校源的时差观测量,改善定位站位置,并估计对应的位置误差统计特性;第 2 阶段使用更新的定位站位置,利用 TS-WLS 算法实现对多目标的高精度定位。该算法不需要目标位置假设和迭代运算,没有局部收敛和定位发散问题,只对目标位置进行估计,避免了定位站和目标的联合估计,算法计算复杂度较小。而且相比于单目标定位算法,本文算法可考虑并利用来自不同目标的定位观测量误

差相关性信息,从而完善定位算法加权矩阵,不仅可同时估计多个目标的位置信息,而且提高了目标定位精度。算法定位性能分析和仿真结果表明,在较小的时差观测误差和定位站位置误差下,对多目标的定位性能可以达到 CRLB。

1 定位场景

本文使用多定位站对三维空间的不相交多目标进行时差定位,定位场景如图 1 所示。设有 K 个目标(图中“ Δ ”表示),位置分别为 $\mathbf{u}_k^o = [x_k^o, y_k^o, z_k^o]^T, k = 1, 2, \dots, K$, 则需估计的多目标位置参数可表示为 $\mathbf{u}^o = [\mathbf{u}_1^{oT}, \mathbf{u}_2^{oT}, \dots, \mathbf{u}_K^{oT}]^T$ 。设有 M 个定位站(图“ \circ ”表示真实定位站位置,“ \bullet ”表示已知带误差的定位站位置)接收辐射源和标校源信号,定位站真实位置设为 $\mathbf{s}_i^o = [x_i^o, y_i^o, z_i^o]^T, i = 1, 2, \dots, M$ 。 M 个定位站位置可表示为 $\mathbf{s}^o = [\mathbf{s}_1^{oT}, \mathbf{s}_2^{oT}, \dots, \mathbf{s}_M^{oT}]^T$ 。标校源(图中“ \square ”表示)位置准确已知,可表示为 $\mathbf{c} = [x_c, y_c, z_c]^T$ 。

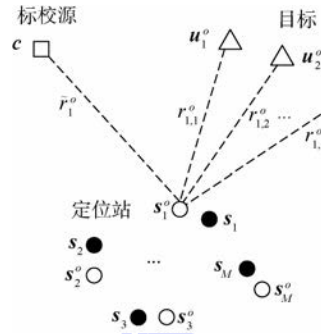


图 1 多目标定位场景

Fig. 1 Multi-source localization scenario

定位站接收来自目标和标校源的信号,通过与参考定位站(设为 s_1 定位站)的接收信号进行相关,可估计相应的 TDOA,乘以信号传播速度,得到信号到达距离差(Range Difference of Arrival, RDOA)。来自第 k 个目标的 RDOA 观测量可表示为

$$r_{i1,k} = r_{i1,k}^o + \Delta r_{i1,k} \quad (1)$$

式中: $i = 2, 3, \dots, M; k = 1, 2, \dots, K; r_{i1,k}$ 为 RDOA 观测量; $\Delta r_{i1,k}$ 为 RDOA 观测误差; $r_{i1,k}^o$ 为 RDOA 真实值,可表示为 $r_{i1,k}^o = r_{i,k}^o - r_{1,k}^o, r_{i,k}^o$ 为第 k 个目标和第 i 个定位站的真实距离,可表示为

$$r_{i,k}^o = \|\mathbf{u}_k^o - \mathbf{s}_i^o\| \quad (2)$$

式中: $\|\cdot\|$ 为 2-范数。

来自第 k 个目标的所有 RDOA 观测量可表示为 $\mathbf{r}_k = [r_{21,k}, r_{31,k}, \dots, r_{M1,k}]^T = \mathbf{r}_k^o + \Delta \mathbf{r}_k, \mathbf{r}_k^o$ 为 RDOA 真值向量, $\mathbf{r}_k^o = [r_{21,k}^o, r_{31,k}^o, \dots, r_{M1,k}^o]^T, \Delta \mathbf{r}_k$

为相应的 RDOA 观测误差向量, $\Delta \mathbf{r}_k = [\Delta r_{21,k}, \Delta r_{31,k}, \dots, \Delta r_{M1,k}]^T$ 。综上, 来自 K 个目标的 RDOA 观测可进一步表示为 $\mathbf{r} = [\mathbf{r}_1^T, \mathbf{r}_2^T, \dots, \mathbf{r}_k^T]^T = \mathbf{r}^o + \Delta \mathbf{r}$, $\mathbf{r}^o = [\mathbf{r}_{11}^o, \mathbf{r}_{21}^o, \dots, \mathbf{r}_{k1}^o]^T$, $\Delta \mathbf{r} = [\Delta \mathbf{r}_1^T, \Delta \mathbf{r}_2^T, \dots, \Delta \mathbf{r}_k^T]^T$, \mathbf{r} 为 $K(M-1) \times 1$ 维向量, 设 $\Delta \mathbf{r}$ 为零均值高斯随机矢量, 且协方差矩阵为 \mathbf{Q}_α 。

假设来自已知标校源的 RDOA 校正观测量可估计, 并有

$$\tilde{r}_{i1} = \tilde{r}_{i1}^o + \Delta \tilde{r}_{i1} \quad i = 2, 3, \dots, M \quad (3)$$

式中: \tilde{r}_{i1} 为校正 RDOA 观测值; $\Delta \tilde{r}_{i1}$ 为观测误差; \tilde{r}_{i1}^o 为相应真值, 可表示为

$$\tilde{r}_{i1}^o = \tilde{r}_i^o - \tilde{r}_1^o = \|\mathbf{c} - \mathbf{s}_i^o\| - \|\mathbf{c} - \mathbf{s}_1^o\| \quad (4)$$

式中: \tilde{r}_i^o 为标校源到第 i 个定位站的真实距离。

来自标校源的所有校正 RDOA 观测量可表示为 $\tilde{\mathbf{r}} = [\tilde{r}_{21}, \tilde{r}_{31}, \dots, \tilde{r}_{M1}]^T = \tilde{\mathbf{r}}^o + \Delta \tilde{\mathbf{r}}$, $\tilde{\mathbf{r}}^o$ 为校正 RDOA 真值向量, $\tilde{\mathbf{r}}^o = [\tilde{r}_{21}^o, \tilde{r}_{31}^o, \dots, \tilde{r}_{M1}^o]^T$, $\Delta \tilde{\mathbf{r}}$ 为相应的观测误差向量, $\Delta \tilde{\mathbf{r}} = [\Delta \tilde{r}_{21}, \Delta \tilde{r}_{31}, \dots, \Delta \tilde{r}_{M1}]^T$, 设 $\Delta \tilde{\mathbf{r}}$ 为零均值高斯随机矢量, 且协方差矩阵为 \mathbf{Q}_c 。

由于定位站存在位置误差, 所以真实位置 \mathbf{s}_i^o 未知, 含误差的定位站位置 \mathbf{s}_i 已知, 可表示为

$$\mathbf{s}_i = \mathbf{s}_i^o + \Delta \mathbf{s}_i \quad i = 1, 2, \dots, M \quad (5)$$

式中: $\Delta \mathbf{s}_i$ 为第 i 个定位站的位置误差向量。

M 个含有位置误差的定位站位置可用 $3M \times 1$ 维向量 \mathbf{s} 表示

$$\mathbf{s} = [\mathbf{s}_1^T, \mathbf{s}_2^T, \dots, \mathbf{s}_M^T]^T = \mathbf{s}^o + \Delta \mathbf{s} \quad (6)$$

式中: \mathbf{s}^o 为真实定位站位置, 可表示为 $\mathbf{s}^o = [\mathbf{s}_1^{oT}, \mathbf{s}_2^{oT}, \dots, \mathbf{s}_M^{oT}]^T$; $\Delta \mathbf{s}$ 为定位站位置误差向量, $\Delta \mathbf{s} = [\Delta \mathbf{s}_1^T, \Delta \mathbf{s}_2^T, \dots, \Delta \mathbf{s}_M^T]^T$, 设 $\Delta \mathbf{s}$ 为零均值高斯随机矢量, 且协方差矩阵为 \mathbf{Q}_β 。

假设误差项 $\Delta \mathbf{r}$ 、 $\Delta \tilde{\mathbf{r}}$ 、 $\Delta \mathbf{s}$ 相互独立^[19], 且来自多目标和标校源的 RDOA 观测量在相同的定位站位置下获得^[21]。本文利用来自标校源和目标的 RDOA 观测量 $\tilde{\mathbf{r}}$ 和 \mathbf{r} , 使用闭式定位算法获得多目标的位置估计。

2 克拉美罗界推导

CRLB 可衡量确定性参数估计的有效性, 由参数估计的无偏性推导而来, 是无偏估计所能达到的最小方差, 可由 Fisher 信息矩阵 \mathbf{J} 获得。在第 1 节描述的定位场景下, 令 $\mathbf{m} = [\mathbf{r}^T, \tilde{\mathbf{r}}^T, \mathbf{s}^T]^T$ 表示所有的量测, $\mathbf{n} = [\mathbf{u}^{oT}, \mathbf{s}^{oT}]^T$ 表示未知参量, \mathbf{r} 、 $\tilde{\mathbf{r}}$ 、 \mathbf{s} 均为高斯分布且相互独立, 则似然函数可表示为

$$p(\mathbf{m} | \mathbf{n}) = p(\mathbf{r}, \tilde{\mathbf{r}}, \mathbf{s} | \mathbf{u}^o, \mathbf{s}^o) =$$

$$p(\mathbf{r} | \mathbf{u}^o, \mathbf{s}^o) p(\tilde{\mathbf{r}} | \mathbf{s}^o) p(\mathbf{s} | \mathbf{s}^o) =$$

$$k_1 \exp \left[-\frac{1}{2} (\mathbf{r} - \mathbf{r}^o)^T \mathbf{Q}_\alpha^{-1} (\mathbf{r} - \mathbf{r}^o) \right] \cdot$$

$$k_2 \exp \left[-\frac{1}{2} (\tilde{\mathbf{r}} - \tilde{\mathbf{r}}^o)^T \mathbf{Q}_c^{-1} (\tilde{\mathbf{r}} - \tilde{\mathbf{r}}^o) \right] \cdot$$

$$k_3 \exp \left[-\frac{1}{2} (\mathbf{s} - \mathbf{s}^o)^T \mathbf{Q}_\beta^{-1} (\mathbf{s} - \mathbf{s}^o) \right] \quad (7)$$

式中: $k_1 = 1 / [(2\pi)^{K(M-1)/2} |\mathbf{Q}_\alpha|^{1/2}]$, $k_2 = 1 / [(2\pi)^{(M-1)/2} |\mathbf{Q}_c|^{1/2}]$ 和 $k_3 = 1 / [(2\pi)^{3M/2} |\mathbf{Q}_\beta|^{1/2}]$ 均为常数。对上述似然函数取对数, 然后求导, 并与自身共轭相乘, 最后求期望便可得到 Fisher 信息矩阵 \mathbf{J} , \mathbf{J}^{-1} 即为 CRLB。

$$\mathbf{J} = E \left[\left(\frac{\partial \ln p(\mathbf{m} | \mathbf{n})}{\partial \mathbf{n}} \right) \left(\frac{\partial \ln p(\mathbf{m} | \mathbf{n})}{\partial \mathbf{n}} \right)^T \right] = \begin{bmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{Y}^T & \mathbf{Z} \end{bmatrix} \quad (8)$$

式中:

$$\begin{cases} \mathbf{X} = (\partial \mathbf{r}^o / \partial \mathbf{u}^o)^T \mathbf{Q}_\alpha^{-1} (\partial \mathbf{r}^o / \partial \mathbf{u}^o) \\ \mathbf{Y} = (\partial \mathbf{r}^o / \partial \mathbf{u}^o)^T \mathbf{Q}_\alpha^{-1} (\partial \mathbf{r}^o / \partial \mathbf{s}^o) \\ \mathbf{Z} = (\partial \mathbf{r}^o / \partial \mathbf{s}^o)^T \mathbf{Q}_\alpha^{-1} (\partial \mathbf{r}^o / \partial \mathbf{s}^o) + \\ \quad \mathbf{Q}_\beta^{-1} + (\partial \tilde{\mathbf{r}}^o / \partial \mathbf{s}^o)^T \mathbf{Q}_c^{-1} (\partial \tilde{\mathbf{r}}^o / \partial \mathbf{s}^o) \end{cases} \quad (9)$$

其中: $\partial \mathbf{r}^o / \partial \mathbf{u}^o$ 为 $K(M-1) \times 3K$ 维矩阵, $\partial \mathbf{r}^o / \partial \mathbf{u}^o = \text{diag} \{ (\partial \mathbf{r}_{11}^o / \partial \mathbf{u}_1^o), (\partial \mathbf{r}_{21}^o / \partial \mathbf{u}_2^o), \dots, (\partial \mathbf{r}_{k1}^o / \partial \mathbf{u}_k^o) \}$, $\partial \mathbf{r}_k^o / \partial \mathbf{u}_k^o = [\partial r_{21,k}^o / \partial u_{21,k}^o, \partial r_{31,k}^o / \partial u_{31,k}^o, \dots, \partial r_{M1,k}^o / \partial u_{M1,k}^o]^T$ 。已知 $r_{i1,k}^o = r_{i,k}^o - r_{1,k}^o$, 且 $r_{i,k}^o = \|\mathbf{u}_k^o - \mathbf{s}_i^o\|$, 可进一步得, $\partial r_{i1,k}^o / \partial \mathbf{u}_k^o = \boldsymbol{\rho}_{\mathbf{u}_k^o, \mathbf{s}_i^o} - \boldsymbol{\rho}_{\mathbf{u}_k^o, \mathbf{s}_1^o}$, $i = 2, 3, \dots, M$, $\boldsymbol{\rho}_{\mathbf{u}_k^o, \mathbf{s}_i^o}$ 为第 k 个目标指向第 i 个定位站的单位矢量, 可表示为 $\boldsymbol{\rho}_{\mathbf{u}_k^o, \mathbf{s}_i^o} = (\mathbf{u}_k^o - \mathbf{s}_i^o) / \|\mathbf{u}_k^o - \mathbf{s}_i^o\|$ 。 $\partial \mathbf{r}^o / \partial \mathbf{s}^o$ 为 $K(M-1) \times 3M$ 维矩阵, $\partial \mathbf{r}^o / \partial \mathbf{s}^o = [(\partial \mathbf{r}_{11}^o / \partial \mathbf{s}^o)^T, (\partial \mathbf{r}_{21}^o / \partial \mathbf{s}^o)^T, \dots, (\partial \mathbf{r}_{k1}^o / \partial \mathbf{s}^o)^T]^T$, $\partial \mathbf{r}_k^o / \partial \mathbf{s}^o = [\partial r_{21,k}^o / \partial s^o, \partial r_{31,k}^o / \partial s^o, \dots, \partial r_{M1,k}^o / \partial s^o]^T$ 。可进一步得, $\partial r_{i1,k}^o / \partial \mathbf{s}^o = [\boldsymbol{\rho}_{\mathbf{u}_k^o, \mathbf{s}_i^o}^T, \mathbf{0}_{3(i-2)}^T, -\boldsymbol{\rho}_{\mathbf{u}_k^o, \mathbf{s}_1^o}^T, \mathbf{0}_{3(M-i)}^T]^T$, $\mathbf{0}$ 表示相应维数的全零列向量。 $\partial \tilde{\mathbf{r}}^o / \partial \mathbf{s}^o$ 为 $(M-1) \times 3K$ 维矩阵, $\partial \tilde{\mathbf{r}}^o / \partial \mathbf{s}^o = [\partial \tilde{r}_{21}^o / \partial s^o, \partial \tilde{r}_{31}^o / \partial s^o, \dots, \partial \tilde{r}_{M1}^o / \partial s^o]^T$, $\partial \tilde{r}_{i1}^o / \partial \mathbf{s}^o$ 计算过程同 $\partial r_{i1,k}^o / \partial \mathbf{s}^o$, 可得 $\partial \tilde{r}_{i1}^o / \partial \mathbf{s}^o = [\boldsymbol{\rho}_{\mathbf{c}, \mathbf{s}_i^o}^T, \mathbf{0}_{3(i-2)}^T, -\boldsymbol{\rho}_{\mathbf{c}, \mathbf{s}_1^o}^T, \mathbf{0}_{3(M-i)}^T]^T$, 式中 $\boldsymbol{\rho}_{\mathbf{c}, \mathbf{s}_i^o}$ 为标校源指向第 i 个定位站的单位矢量, 可表示为 $\boldsymbol{\rho}_{\mathbf{c}, \mathbf{s}_i^o} = (\mathbf{c} - \mathbf{s}_i^o) / \|\mathbf{c} - \mathbf{s}_i^o\|$ 。

令 $\mathbf{C}(\mathbf{n}) = \mathbf{J}^{-1}$, 则矩阵 $\mathbf{C}(\mathbf{n})$ 的左 $3K \times 3K$ 块矩阵即为不相交多目标位置的无偏估计 CRLB, 可用 $\mathbf{C}(\mathbf{u})$ 表示。对式(8)应用分割矩阵求逆公式^[20]可得

$$\mathbf{C}(\mathbf{u}) = (\mathbf{X} - \mathbf{Y} \mathbf{Z}^{-1} \mathbf{Y}^T)^{-1} = \mathbf{X}^{-1} + \mathbf{X}^{-1} \mathbf{Y} (\mathbf{Z} - \mathbf{Y}^T \mathbf{X}^{-1} \mathbf{Y})^{-1} \mathbf{Y}^T \mathbf{X}^{-1} \quad (10)$$

式中: \mathbf{X}^{-1} 为不存在定位站位置误差时的 CRLB;

由文献[19]的附录II推导可知, $\mathbf{X}^{-1} \mathbf{Y} (\mathbf{Z} - \mathbf{Y}^T \mathbf{X}^{-1} \cdot \mathbf{Y})^{-1} \mathbf{Y}^T \mathbf{X}^{-1}$ 在 $\mathbf{Q}_\beta \neq \mathbf{0}$ 时正定 ($\mathbf{0}$ 表示全零矩阵), 意味着即使使用标校源校正技术, 也不会将最优定位性能改善到没有定位站位置误差时的 CRLB。当不存在标校源时, 不相交多目标的 CRLB 可由文献[21]中式(43)表示, 与式(10)比较发现, 标校源的使用对 CRLB 的影响体现在 \mathbf{Z} 中引入了 $(\partial \bar{\mathbf{r}} / \partial \mathbf{s}^o)^T \mathbf{Q}_c^{-1} (\partial \bar{\mathbf{r}} / \partial \mathbf{s}^o)$ 项, 文献[19]进一步推导了使用标校源与未使用标校源的 CRLB 之间的关系, 如文献[19]中式(12)所示, 可用式(11)表示

$$\mathbf{C}(\mathbf{u})_2 = \mathbf{C}(\mathbf{u})_1 - \mathbf{\Gamma} \quad (11)$$

式中: $\mathbf{C}(\mathbf{u})_1$ 为未使用标校源时的 CRLB; $\mathbf{C}(\mathbf{u})_2$ 为使用标校源时的 CRLB; $\mathbf{\Gamma}$ 表示引入标校源后的定位性能改善程度, 经分析知 $\mathbf{\Gamma}$ 为半正定矩阵, 意味着标校源的引入会提升 (若不提升, 至少不会降低) 对多目标的无源定位性能。可通过举例直观说明这一结论。

以 2 个目标为例, 设目标 1 位置为 $[2\ 000, 2\ 500, 3\ 000]$ m, 目标 2 位置为 $[600, 650, 550]$ m, 标校源位置已知, 设为 $[1\ 500, 1\ 550, 1\ 500]$ m, 定位站实际位置如表 1 所示。来自两目标和已知标校源的 RDOA 观测量平均误差功率均设为 $\sigma_r^2 = 10^{-3}$ m², $\mathbf{Q}_\alpha, \mathbf{Q}_c, \mathbf{Q}_\beta$ 3 种协方差矩阵的构造方式可具体见第 4 节。

对两目标的位置估计 CRLB 随定位站位置误差的变化如图 2 所示, $f(\mathbf{C}(\mathbf{u})_1)$ 和 $f(\mathbf{C}(\mathbf{u})_2)$ 分别为 $\mathbf{C}(\mathbf{u})_1$ 和 $\mathbf{C}(\mathbf{u})_2$ 的主对角线元素之和的平方根, 用来表征 CRLB 的大小, σ_s 为定位站位置误差标准差。虚线表示目标 1 的 CRLB 曲线, 实线表示目标 2 的 CRLB 曲线, 其中带有“○”的曲线表示未使用标校源, 带有“□”的曲线表示使用标校源。从图中可以明显看出, 不管对目标 1 还是目标 2, 标校源的使用可以大幅提高定位性能: 对目标 1, 当定位站位置噪声功率 $\sigma_s^2 > -20$ dB 时, 定位性能至少可以改善 8 dB; 对目

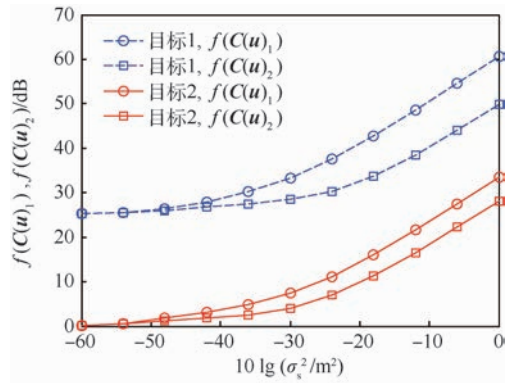


图 2 多目标定位 CRLB 比较

Fig. 2 CRLB comparison of multi-source localization

标 2, 当定位站位置噪声功率 $\sigma_s^2 > -20$ dB 时, 定位性能至少可以改善 4 dB。

3 定位算法及性能分析

本节为本文闭式定位算法的理论推导和定位性能分析。首先利用来自标校源的校正 RDOA 观测量估计定位站位置误差, 进而改善定位站位置; 然后基于校正的定位站位置, 引入冗余变量, 将 RDOA 定位方程伪线性化, 使用 TS-WLS 定位算法获得多目标的位置估计。最后通过定位性能分析评估该算法的有效性。

3.1 定位算法

本文定位算法分为 2 个阶段: 第 1 阶段借助标校源校正定位站位置, 通过一步闭式运算估计定位站位置误差, 不需迭代运算, 占整个算法计算量的比重较小; 第 2 阶段使用校正的定位站位置对多目标进行 TS-WLS 定位, 使用更多的时差观测量信息, 且需要进行两步 WLS 算法, 占整个算法计算量的比重较大。

1) 第 1 阶段: 校正定位站位置

由于标校源位置已知, 结合定位站位置可以预测 RDOA, 与校正 RDOA 观测值的差异可以提供定位站位置误差和时差观测误差的信息, 利用该信息可以校正定位站位置。由第 1 节可知, 校正 RDOA 观测量可表示为

$$\bar{r}_{i1} = \bar{r}_{i1}^o + \Delta \bar{r}_{i1} = \|\mathbf{c} - \mathbf{s}_i^o\| - \|\mathbf{c} - \mathbf{s}_i^o\| + \Delta \bar{r}_{i1} \quad (12)$$

式中: $i = 2, 3, \dots, M$ 。 $\mathbf{s}_i^o, \mathbf{s}_i^o$ 未知, $\mathbf{s}_i, \mathbf{s}_i$ 已知, 可对其进行泰勒展开处理, 忽略二阶及高阶项得 $\|\mathbf{c} - \mathbf{s}_i^o\| \approx \|\mathbf{c} - \mathbf{s}_i\| + \boldsymbol{\rho}_{e, \mathbf{s}_i}^T \cdot \Delta \mathbf{s}_i$ 。预测 RDOA 可表示为 $\bar{r}_{i1} = \|\mathbf{c} - \mathbf{s}_i\| - \|\mathbf{c} - \mathbf{s}_i\|$, 因此校正 RDOA 与预测 RDOA 存在如下关系:

$$\bar{r}_{i1} = \bar{r}_{i1} + \boldsymbol{\rho}_{e, \mathbf{s}_i}^T \cdot \Delta \mathbf{s}_i - \boldsymbol{\rho}_{e, \mathbf{s}_1}^T \cdot \Delta \mathbf{s}_1 + \Delta \bar{r}_{i1} \quad (13)$$

对所有的校正 RDOA 与预测 RDOA 可用以下矩阵形式表示:

表 1 定位站位置坐标

Table 1 Position coordinates of sensors

定位站	位置坐标		
	x	y	z
s_1	300	100	150
s_2	400	150	100
s_3	300	500	200
s_4	350	200	100
s_5	-100	-100	-100
s_6	200	-300	-200

$$\tilde{\mathbf{h}} = \tilde{\mathbf{G}} \cdot \Delta \mathbf{s} + \tilde{\Delta \mathbf{r}} \quad (14)$$

式中: $\Delta \mathbf{s}$ 为待估计的定位站位置误差项; $\tilde{\Delta \mathbf{r}} = [\Delta \tilde{r}_{21}, \Delta \tilde{r}_{31}, \dots, \Delta \tilde{r}_{M1}]^T$ 为校正 RDOA 观测量误差, 且协方差矩阵为 \mathbf{Q}_c 。

$$\tilde{\mathbf{h}} = \begin{bmatrix} \tilde{r}_{21} - \hat{r}_{21} \\ \tilde{r}_{31} - \hat{r}_{31} \\ \vdots \\ \tilde{r}_{M1} - \hat{r}_{M1} \end{bmatrix} \quad (15)$$

$$\tilde{\mathbf{G}} = \begin{bmatrix} -\boldsymbol{\rho}_{c,s_1}^T & \boldsymbol{\rho}_{c,s_2}^T & \mathbf{0} & \cdots & \mathbf{0} \\ -\boldsymbol{\rho}_{c,s_1}^T & \mathbf{0} & \boldsymbol{\rho}_{c,s_3}^T & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\boldsymbol{\rho}_{c,s_1}^T & \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\rho}_{c,s_M}^T \end{bmatrix}$$

其中: $\tilde{\mathbf{G}}$ 中 $\mathbf{0}$ 表示 3 维全零列向量。

由于 $\Delta \mathbf{s}$ 统计分布具有先验知识, $\Delta \mathbf{s}$ 为零均值高斯随机矢量, 且协方差矩阵为 \mathbf{Q}_β 。对非确定性参数或随机参数估计, 可使用贝叶斯高斯-马尔可夫理论^[19], $\Delta \mathbf{s}$ 的最小线性均方误差估计为

$$\Delta \hat{\mathbf{s}} = (\mathbf{Q}_\beta^{-1} + \tilde{\mathbf{G}}^T \mathbf{Q}_c^{-1} \tilde{\mathbf{G}})^{-1} \tilde{\mathbf{G}}^T \mathbf{Q}_c^{-1} \tilde{\mathbf{h}} \quad (16)$$

因此定位站位置可改善为

$$\hat{\mathbf{s}} = \mathbf{s} - \Delta \hat{\mathbf{s}} = \mathbf{s}^o + \Delta \mathbf{s} - \Delta \hat{\mathbf{s}} = \mathbf{s}^o + \tilde{\Delta \mathbf{s}} \quad (17)$$

式中: $\tilde{\Delta \mathbf{s}} = \Delta \mathbf{s} - \Delta \hat{\mathbf{s}}$ 为校正后的定位站位置误差。

由式(16)可推导 $\tilde{\Delta \mathbf{s}}$ 的协方差矩阵为

$$\text{cov}(\tilde{\Delta \mathbf{s}}) = (\mathbf{Q}_\beta^{-1} + \tilde{\mathbf{G}}^T \mathbf{Q}_c^{-1} \tilde{\mathbf{G}})^{-1} \quad (18)$$

定位站位置校正前后的协方差矩阵差值可表示为 $\Delta = \text{cov}(\Delta \mathbf{s}) - \text{cov}(\tilde{\Delta \mathbf{s}}) = \mathbf{Q}_\beta - (\mathbf{Q}_\beta^{-1} + \tilde{\mathbf{G}}^T \mathbf{Q}_c^{-1} \tilde{\mathbf{G}})^{-1}$, 经计算与分析可知 Δ 半正定, 即标校源减小了定位站位置误差的协方差矩阵, 因此可以改善多目标定位性能。

2) 第 2 阶段: TS-WLS 定位

定位站位置经校正后, 可使用 TS-WLS 算法进行多目标定位。TS-WLS 算法分为 2 步: 第 1 步引入冗余变量将 RDOA 定位方程伪线性化, 使用 WLS 算法估计目标位置和冗余变量; 第 2 步利用目标位置与冗余变量的约束关系, 结合第 1 步估计结果构建线性方程, 再次使用 WLS 算法, 估计最终多目标位置。

第 1 步: 由第 1 节可知, 对第 k 个目标满足 $r_{i1,k}^o = r_{i,k}^o - r_{1,k}^o$, 将 $r_{1,k}^o$ 移到等号左侧, 并对两边平方, 结合式(2)可得如下 RDOA 观测方程:

$$(r_{i1,k}^o)^2 + \mathbf{s}_1^{oT} \mathbf{s}_1^o - \mathbf{s}_i^{oT} \mathbf{s}_i^o = -2(\mathbf{s}_i^o - \mathbf{s}_1^o)^T \mathbf{u}_k^o - 2r_{i1,k}^o r_{1,k}^o \quad (19)$$

式中: 无误差项 $r_{i1,k}^o, \mathbf{s}_i^o$ 可由相应的误差项代替, 即 $r_{i1,k}^o = r_{i1,k} - \Delta r_{i1,k}, \mathbf{s}_i^o = \hat{\mathbf{s}}_i - \Delta \tilde{\mathbf{s}}_i$ 。对 $r_{1,k}^o, r_{1,k}^o =$

$\|\mathbf{u}_k^o - \mathbf{s}_1^o\|, \mathbf{s}_1^o$ 未知, 可在校正值 $\hat{\mathbf{s}}_1$ 处进行泰勒展开处理, 保留一阶线性项, 可得 $r_{1,k}^o \approx \|\mathbf{u}_k^o - \hat{\mathbf{s}}_1\| + \boldsymbol{\rho}_{\mathbf{u}_k^o, \hat{\mathbf{s}}_1}^T \cdot \tilde{\Delta \mathbf{s}}_1 = \hat{r}_{1,k}^o + \boldsymbol{\rho}_{\mathbf{u}_k^o, \hat{\mathbf{s}}_1}^T \cdot \tilde{\Delta \mathbf{s}}_1, \hat{r}_{1,k}^o$ 为第 k 个目标与校正定位站 $\hat{\mathbf{s}}_1$ 的距离, $\boldsymbol{\rho}_{\mathbf{u}_k^o, \hat{\mathbf{s}}_1}$ 为第 k 个目标指向校正定位站 $\hat{\mathbf{s}}_1$ 的单位矢量。在式(19)中用误差项代替所有的无误差项, 并忽略二阶及高阶误差项, 整理可得

$$\boldsymbol{\varepsilon}_{i1,k} = 2r_{i1,k}^o \Delta r_{i1,k} + 2(\mathbf{u}_k^o - \hat{\mathbf{s}}_i)^T \tilde{\Delta \mathbf{s}}_i - 2(\mathbf{u}_k^o - \hat{\mathbf{s}}_1 + r_{i1,k} \boldsymbol{\rho}_{\mathbf{u}_k^o, \hat{\mathbf{s}}_1})^T \tilde{\Delta \mathbf{s}}_1 = r_{i1,k}^2 + \hat{\mathbf{s}}_1^T \hat{\mathbf{s}}_1 - \hat{\mathbf{s}}_i^T \hat{\mathbf{s}}_i + 2(\hat{\mathbf{s}}_i - \hat{\mathbf{s}}_1)^T \mathbf{u}_k^o + 2r_{i1,k} \hat{r}_{1,k}^o \quad (20)$$

式中: 距离参量 $\hat{r}_{1,k}^o$ 与 \mathbf{u}_k^o 存在开方关系, 所以方程对 \mathbf{u}_k^o 非线性。可将 $\hat{r}_{1,k}^o$ 作为冗余变量, 假设与目标位置参数 \mathbf{u}_k^o 独立, 从而将式(20)伪线性化。令 $\boldsymbol{\varphi}_{1,k}^o = [\mathbf{u}_k^{oT}, \hat{r}_{1,k}^o]^T$, 对式(20)用矩阵形式可表示为

$$\boldsymbol{\varepsilon}_{1,k} = \mathbf{h}_{1,k} - \mathbf{G}_{1,k} \cdot \boldsymbol{\varphi}_{1,k}^o \quad (21)$$

式中: $\boldsymbol{\varepsilon}_{1,k} = [\boldsymbol{\varepsilon}_{21,k}, \boldsymbol{\varepsilon}_{31,k}, \dots, \boldsymbol{\varepsilon}_{M1,k}]^T = \mathbf{B}_{1,k} \cdot \Delta \mathbf{r}_k + \mathbf{D}_{1,k} \cdot \tilde{\Delta \mathbf{s}}, \mathbf{B}_{1,k} = 2 \text{diag}(r_{2,k}^o, r_{3,k}^o, \dots, r_{M,k}^o), \mathbf{D}_{1,k}$ 为 $(M-1) \times 3M$ 维矩阵, 第 $i-1$ 行可表示为 $2 \times [- (\mathbf{u}_k^o - \hat{\mathbf{s}}_1)^T - r_{i1,k} \boldsymbol{\rho}_{\mathbf{u}_k^o, \hat{\mathbf{s}}_1}^T, \mathbf{0}_{3(i-2)}^T, (\mathbf{u}_k^o - \hat{\mathbf{s}}_i)^T, \mathbf{0}_{3(M-i)}^T]$ 。

$$\mathbf{h}_{1,k} = \begin{bmatrix} r_{21,k}^2 + \hat{\mathbf{s}}_1^T \hat{\mathbf{s}}_1 - \hat{\mathbf{s}}_2^T \hat{\mathbf{s}}_2 \\ r_{31,k}^2 + \hat{\mathbf{s}}_1^T \hat{\mathbf{s}}_1 - \hat{\mathbf{s}}_3^T \hat{\mathbf{s}}_3 \\ \vdots \\ r_{M1,k}^2 + \hat{\mathbf{s}}_1^T \hat{\mathbf{s}}_1 - \hat{\mathbf{s}}_M^T \hat{\mathbf{s}}_M \end{bmatrix} \quad (22)$$

$$\mathbf{G}_{1,k} = -2 \begin{bmatrix} (\hat{\mathbf{s}}_2 - \hat{\mathbf{s}}_1)^T & r_{21,k} \\ (\hat{\mathbf{s}}_3 - \hat{\mathbf{s}}_1)^T & r_{31,k} \\ \vdots & \vdots \\ (\hat{\mathbf{s}}_M - \hat{\mathbf{s}}_1)^T & r_{M1,k} \end{bmatrix}$$

由于 K 个目标的 RDOA 观测量基于相同的定位站位置误差 $\tilde{\Delta \mathbf{s}}$, 所以可构建所有目标的矩阵定位方程

$$\boldsymbol{\varepsilon}_1 = \mathbf{h}_1 - \mathbf{G}_1 \cdot \boldsymbol{\varphi}_1^o \quad (23)$$

式中: $\mathbf{h}_1 = [\mathbf{h}_{1,1}^T, \mathbf{h}_{1,2}^T, \dots, \mathbf{h}_{1,K}^T]^T; \mathbf{G}_1$ 为块对角矩阵, $\mathbf{G}_1 = \text{diag}(\mathbf{G}_{1,1}, \mathbf{G}_{1,2}, \dots, \mathbf{G}_{1,K})$; $\boldsymbol{\varphi}_1^o = [\boldsymbol{\varphi}_{1,1}^{oT}, \boldsymbol{\varphi}_{1,2}^{oT}, \dots, \boldsymbol{\varphi}_{1,K}^{oT}]^T; \boldsymbol{\varepsilon}_1 = \mathbf{B}_1 \cdot \Delta \mathbf{r} + \mathbf{D}_1 \cdot \tilde{\Delta \mathbf{s}}, \boldsymbol{\varepsilon}_1 = [\boldsymbol{\varepsilon}_{1,1}^T, \boldsymbol{\varepsilon}_{1,2}^T, \dots, \boldsymbol{\varepsilon}_{1,K}^T]^T, \mathbf{B}_1$ 为块对角矩阵, $\mathbf{B}_1 = \text{diag}(\mathbf{B}_{1,1}, \mathbf{B}_{1,2}, \dots, \mathbf{B}_{1,K}), \mathbf{D}_1 = [\mathbf{D}_{1,1}^T, \mathbf{D}_{1,2}^T, \dots, \mathbf{D}_{1,K}^T]^T$ 。

使用 WLS 算法求解式(23), 其 WLS 解为

$$\boldsymbol{\varphi}_1 = (\mathbf{G}_1^T \mathbf{W}_1 \mathbf{G}_1)^{-1} \mathbf{G}_1^T \mathbf{W}_1 \mathbf{h}_1 \quad (24)$$

式中: \mathbf{W}_1 为加权矩阵, 若取 $\mathbf{W}_1 = \text{cov}(\boldsymbol{\varepsilon}_1)^{-1}$, 则 WLS 估计与近似 ML 估计等效。由于 $\Delta \mathbf{r}$ 与 $\tilde{\Delta \mathbf{s}}$ 相互独立, 结合 $\boldsymbol{\varepsilon}_1$ 的表达式, 可得

$$\mathbf{W}_1 = [\mathbf{B}_1 \mathbf{Q}_\alpha \mathbf{B}_1^T + \mathbf{D}_1 (\mathbf{Q}_\beta^{-1} + \tilde{\mathbf{G}}^T \mathbf{Q}_c^{-1} \tilde{\mathbf{G}})^{-1} \mathbf{D}_1^T]^{-1} \quad (25)$$

当 RDOA 观测误差和校正后的定位站位置误差足够小,采用扰动法^[13]可得 φ_1 的协方差矩阵

$$\text{cov}(\varphi_1) = (\mathbf{G}_1^T \mathbf{W}_1 \mathbf{G}_1)^{-1} \quad (26)$$

第 2 步:由于 φ_1 中含有冗余变量 $\hat{r}_{1,k}^o$,与目标位置参数存在约束关系 $\hat{r}_{1,k}^o = \|\mathbf{u}_k^o - \hat{\mathbf{s}}_1\|$,所以第 2 步 WLS 解并非最优。第 2 步 WLS 利用该约束关系构造误差向量方程,再次使用 WLS 估计,改善目标的定位精度。

在第 1 步 WLS 解中, $\varphi_{1,k}(1:3)$ 为第 k 个目标的位置估计, $\varphi_{1,k}(4)$ 为第 k 个目标到校正后参考定位站的距离估计。令 $\varphi_{1,k}(1:3) = \mathbf{u}_k^o + \Delta\varphi_{1,k}(1:3)$, $\varphi_{1,k}(4) = \hat{r}_{1,k}^o + \Delta\varphi_{1,k}(4)$,其中, $\Delta\varphi_{1,k}(1:3)$ 为第 k 个目标的位置估计误差, $\Delta\varphi_{1,k}(4)$ 为相应的距离估计误差。根据上述分析,满足:

$$\begin{aligned} &(\varphi_{1,k}(1:3) - \hat{\mathbf{s}}_1) \odot (\varphi_{1,k}(1:3) - \hat{\mathbf{s}}_1) = \\ &(\mathbf{u}_k^o - \hat{\mathbf{s}}_1) \odot (\mathbf{u}_k^o - \hat{\mathbf{s}}_1) + \\ &2(\mathbf{u}_k^o - \hat{\mathbf{s}}_1) \odot \Delta\varphi_{1,k}(1:3) + \\ &\Delta\varphi_{1,k}(1:3) \odot \Delta\varphi_{1,k}(1:3) \end{aligned} \quad (27)$$

$$\varphi_{1,k}^2(4) = (\hat{r}_{1,k}^o)^2 + 2\hat{r}_{1,k}^o \cdot \Delta\varphi_{1,k}(4) + (\Delta\varphi_{1,k}(4))^2 \quad (28)$$

式中: \odot 表示 Schur 积,即元素与元素相乘,且 $\hat{r}_{1,k}^o{}^2 = (\mathbf{u}_k^o - \hat{\mathbf{s}}_1)^T (\mathbf{u}_k^o - \hat{\mathbf{s}}_1)$ 。

令 $\varphi_{2,k}^o = (\mathbf{u}_k^o - \hat{\mathbf{s}}_1) \odot (\mathbf{u}_k^o - \hat{\mathbf{s}}_1)$,忽略二阶误差项,由式(27)与式(28)可得矩阵表达式

$$\boldsymbol{\varepsilon}_{2,k} = \mathbf{h}_{2,k} - \mathbf{G}_{2,k} \cdot \varphi_{2,k}^o \quad (29)$$

式中: $\boldsymbol{\varepsilon}_{2,k} = \mathbf{B}_{2,k} \cdot \Delta\varphi_{1,k}$, $\mathbf{B}_{2,k} = 2\text{diag}((\mathbf{u}_k^o - \hat{\mathbf{s}}_1)^T, \hat{r}_{1,k}^o)$ 。

$$\left\{ \begin{aligned} \mathbf{h}_{2,k} &= \begin{bmatrix} (\varphi_{1,k}(1:3) - \hat{\mathbf{s}}_1) \odot (\varphi_{1,k}(1:3) - \hat{\mathbf{s}}_1) \\ \varphi_{1,k}^2(4) \end{bmatrix} \\ \mathbf{G}_{2,k} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \end{aligned} \right. \quad (30)$$

扩展到 K 个目标,可用以下矩阵方程表示:

$$\boldsymbol{\varepsilon}_2 = \mathbf{h}_2 - \mathbf{G}_2 \cdot \boldsymbol{\varphi}_2^o \quad (31)$$

式中: $\mathbf{h}_2 = [\mathbf{h}_{2,1}^T, \mathbf{h}_{2,2}^T, \dots, \mathbf{h}_{2,K}^T]^T$; \mathbf{G}_2 为块对角矩阵, $\mathbf{G}_2 = \text{diag}(\mathbf{G}_{2,1}, \mathbf{G}_{2,2}, \dots, \mathbf{G}_{2,K})$; $\boldsymbol{\varphi}_2^o = [\varphi_{2,1}^o{}^T, \varphi_{2,2}^o{}^T, \dots, \varphi_{2,K}^o{}^T]^T$; $\boldsymbol{\varepsilon}_2 = [\boldsymbol{\varepsilon}_{2,1}^T, \boldsymbol{\varepsilon}_{2,2}^T, \dots, \boldsymbol{\varepsilon}_{2,K}^T]^T = \mathbf{B}_2 \cdot \Delta\boldsymbol{\varphi}_1$, \mathbf{B}_2 为块对角矩阵, $\mathbf{B}_2 = \text{diag}(\mathbf{B}_{2,1}, \mathbf{B}_{2,2}, \dots, \mathbf{B}_{2,K})$, $\Delta\boldsymbol{\varphi}_1 = [\Delta\varphi_{1,1}^T, \Delta\varphi_{1,2}^T, \dots, \Delta\varphi_{1,K}^T]^T$ 为第 1 步

WLS 解的误差向量。

再次使用 WLS 算法,式(31)的 WLS 解为

$$\boldsymbol{\varphi}_2 = (\mathbf{G}_2^T \mathbf{W}_2 \mathbf{G}_2)^{-1} \mathbf{G}_2^T \mathbf{W}_2 \mathbf{h}_2 \quad (32)$$

式中: \mathbf{W}_2 为加权矩阵,取 $\mathbf{W}_2 = \text{cov}(\boldsymbol{\varepsilon}_2)^{-1}$,根据式(26),并结合 $\boldsymbol{\varepsilon}_2$ 的表达式可得

$$\mathbf{W}_2 = [\mathbf{B}_2 (\mathbf{G}_1^T \mathbf{W}_1 \mathbf{G}_1)^{-1} \mathbf{B}_2^T]^{-1} \quad (33)$$

将式(31)代入式(32)并结合式(33),可得 $\boldsymbol{\varphi}_2$ 的协方差矩阵

$$\text{cov}(\boldsymbol{\varphi}_2) = (\mathbf{G}_2^T \mathbf{W}_2 \mathbf{G}_2)^{-1} \quad (34)$$

根据 $\boldsymbol{\varphi}_2^o$ 的定义,第 k 个目标的最终位置估计可表示为

$$\mathbf{u}_k = \text{diag}(\text{sgn}(\varphi_{1,k}(1:3) - \hat{\mathbf{s}}_1)) \sqrt{\varphi_{2,k}} + \hat{\mathbf{s}}_1 \quad (35)$$

式中: sgn 为符号函数,可消除平方根运算产生的符号模糊,使最终解与第 1 步 WLS 解的符号保持一致。

注意到,在求解加权矩阵 \mathbf{W}_1 和 \mathbf{W}_2 时,需要目标和定位站的真实位置,均不可获得。为便于算法实现,可令 $\mathbf{W}_1 = \mathbf{Q}_\alpha^{-1}$,代入式(24)获得目标位置的初始估计值,结合定位站位置校正值,由式(25)更新 \mathbf{W}_1 ,从而改善目标位置估计。使用第 1 步 WLS 解可得到加权矩阵 \mathbf{W}_2 。

综上所述,使用标校源对多目标的时差定位算法实现步骤如下:

步骤 1 根据校正 RDOA 观测量,通过式(16)估计定位站位置误差,通过式(17)校正定位站位置。

步骤 2 令 \mathbf{W}_1 为 RDOA 观测量协方差矩阵的逆,代入式(24)获得目标位置的初始估计值。

步骤 3 根据初始目标位置,通过式(25)更新 \mathbf{W}_1 ,代入式(24)获得第 1 步 WLS 的改善目标位置。

步骤 4 根据第 1 步 WLS 解,通过式(33)构建 \mathbf{W}_2 ,代入式(32)获得第 2 步 WLS 解。

步骤 5 由式(35)得到多目标的最终位置估计。

3.2 性能分析

本节对多目标定位算法的定位精度进行分析,并与 CRLB 比较,评估算法的有效性。

由式(35)可得 $\boldsymbol{\varphi}_{2,k} = (\mathbf{u}_k - \hat{\mathbf{s}}_1) \odot (\mathbf{u}_k - \hat{\mathbf{s}}_1)$,令 $\boldsymbol{\varphi}_{2,k} = \boldsymbol{\varphi}_{2,k}^o + \Delta\boldsymbol{\varphi}_{2,k}$, $\mathbf{u}_k = \mathbf{u}_k^o + \Delta\mathbf{u}_k$, $\Delta\boldsymbol{\varphi}_{2,k}$ 为第 2 步 WLS 解的估计误差, $\Delta\mathbf{u}_k$ 为第 k 个目标的定位误差,将含有误差项的 $\boldsymbol{\varphi}_{2,k}$ 与 \mathbf{u}_k 代入 $\boldsymbol{\varphi}_{2,k} = (\mathbf{u}_k - \hat{\mathbf{s}}_1) \odot (\mathbf{u}_k - \hat{\mathbf{s}}_1)$ 中,忽略二阶误差项 $\Delta\mathbf{u}_k \odot \Delta\mathbf{u}_k$,可得

$$\Delta \mathbf{u}_k = \mathbf{B}_{3,k}^{-1} \cdot \Delta \boldsymbol{\varphi}_{2,k} \quad (36)$$

式中: $\mathbf{B}_{3,k} = 2 \text{diag}(\mathbf{u}_k^o - \hat{s}_1)$ 。

扩展到 K 个目标, 则定位误差 $\Delta \mathbf{u}$ 与第 2 步 WLS 解的估计误差 $\Delta \boldsymbol{\varphi}_2$ 的关系可表示为

$$\Delta \mathbf{u} = \mathbf{B}_3^{-1} \cdot \Delta \boldsymbol{\varphi}_2 \quad (37)$$

式中: $\Delta \mathbf{u} = [\Delta \mathbf{u}_1^T, \Delta \mathbf{u}_2^T, \dots, \Delta \mathbf{u}_K^T]^T$; $\Delta \boldsymbol{\varphi}_2 = [\Delta \boldsymbol{\varphi}_{2,1}^T, \Delta \boldsymbol{\varphi}_{2,2}^T, \dots, \Delta \boldsymbol{\varphi}_{2,K}^T]^T$; \mathbf{B}_3 为块对角矩阵, $\mathbf{B}_3 = \text{diag}(\mathbf{B}_{3,1}, \mathbf{B}_{3,2}, \dots, \mathbf{B}_{3,K})$ 。

忽略 \mathbf{B}_3 中的定位站位置误差, 由式 (37) 可得多目标最终位置估计的协方差矩阵

$$\text{cov}(\mathbf{u}) = E[\Delta \mathbf{u} \cdot \Delta \mathbf{u}^T] = \mathbf{B}_3^{-1} \cdot \text{cov}(\boldsymbol{\varphi}_2) \cdot \mathbf{B}_3^{-T} \quad (38)$$

式中: $\text{cov}(\boldsymbol{\varphi}_2)$ 由式 (34) 获得。将式 (34)、式 (33) 和式 (25) 分别代入式 (38) 中, 经化简可得

$$\text{cov}(\mathbf{u}) = [\mathbf{P}_1^T (\mathbf{Q}_\alpha + \mathbf{P}_2 (\mathbf{Q}_\beta^{-1} + \tilde{\mathbf{G}}^T \mathbf{Q}_c^{-1} \tilde{\mathbf{G}})^{-1} \mathbf{P}_2^T)^{-1} \mathbf{P}_1]^{-1} \quad (39)$$

式中: $\mathbf{P}_1 = \mathbf{B}_1^{-1} \mathbf{G}_1 \mathbf{B}_2^{-1} \mathbf{G}_2 \mathbf{B}_3$; $\mathbf{P}_2 = \mathbf{B}_1^{-1} \mathbf{D}_1$ 。

根据矩阵求逆引理, 式 (39) 可整理得

$$\text{cov}(\mathbf{u}) = (\hat{\mathbf{X}} - \hat{\mathbf{Y}} \hat{\mathbf{Z}}^{-1} \hat{\mathbf{Y}}^T)^{-1} \quad (40)$$

式中: $\hat{\mathbf{X}} = \mathbf{P}_1^T \mathbf{Q}_\alpha^{-1} \mathbf{P}_1$; $\hat{\mathbf{Y}} = \mathbf{P}_1^T \mathbf{Q}_\alpha^{-1} \mathbf{P}_2$; $\hat{\mathbf{Z}} = \mathbf{Q}_\beta^{-1} + \tilde{\mathbf{G}}^T \mathbf{Q}_c^{-1} \tilde{\mathbf{G}} + \mathbf{P}_2^T \mathbf{Q}_\alpha^{-1} \mathbf{P}_2$ 。

将式 (40) 与式 (10) 比较可知, $\text{cov}(\mathbf{u})$ 与 $\mathbf{C}(\mathbf{u})$ 具有相同的结构, 而且表达式内的对应矩阵也具有相同的结构, 文献 [19, 21] 均对两者的关系进行了分析。设定 3 个条件: ①定位站位置误差相比于标校源与定位站距离足够小, 即 $\|\Delta \mathbf{s}_i\| \ll \tilde{r}_i^o$; ②定位站位置误差相比于目标与定位站距离足够小, 即 $\|\Delta \mathbf{s}_i\| \ll r_{i,k}^o$; ③RDOA 观测误差相比于目标与定位站距离足够小, 即 $|\Delta r_{i,k}| \ll r_{i,k}^o$ 。当满足条件 ①、②、③ 时, 文献 [19] 证明了使用标校源定位单目标的有效性; 当满足条件 ②、③ 时, 文献 [21] 证明了无标校源下多目标定位的有效性。本文在上述 3 个前提条件下, 借鉴文献 [19] 附录 V 和文献 [21] 附录 B 的证明过程 (由于证明过程较为复杂, 在此不再赘述), 可证明等式 $\mathbf{P}_1 \approx -\partial \mathbf{r}^o / \partial \mathbf{u}^o$, $\mathbf{P}_2 \approx -\partial \mathbf{r}^o / \partial \mathbf{s}^o$, $\tilde{\mathbf{G}} \approx -\partial \tilde{\mathbf{r}}^o / \partial \mathbf{s}^o$ 成立, 从而使得

$$\text{cov}(\mathbf{u}) \approx \mathbf{C}(\mathbf{u}) \quad (41)$$

综上, 在定位站位置误差和 RDOA 观测误差均足够小, 或标校源与目标距离定位站均足够远时, 本文定位算法可以实现有效估计, 定位性能理论上可以达到 CRLB。

4 仿真分析

本节将通过计算机仿真验证本文闭式算法的

定位性能。采用 6 定位站, 定位站真实位置坐标如表 1 所示, 带误差的定位站位置通过在定位站真实位置的基础上叠加零均值高斯白噪声获得, 定位站位置误差协方差矩阵结构设为 $\mathbf{Q}_\beta = \sigma_s^2 \mathbf{R}_s$, $\mathbf{R}_s = \text{diag}(10, 10, 10, 2, 2, 2, 10, 10, 10, 40, 40, 40, 20, 20, 20, 3, 3, 3)$, σ_s 为定位站位置误差标准差, 即各定位站具有不同的位置误差, 但每个定位站各位置坐标误差相同。来自目标和标校源的 RDOA 观测量通过在对真值的基础上叠加一定强度的零均值高斯白噪声产生, 设协方差矩阵结构分别为 $\mathbf{Q}_\alpha = \sigma_r^2 \mathbf{R}_l$, $\mathbf{Q}_c = \sigma_r^2 \mathbf{R}_c$, 其中 σ_r 为 RDOA 观测量误差标准差, $\sigma_r^2 = 10^{-3} \text{ m}^2$, $\mathbf{R}_l = \text{blkdiag}(\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_K)$, $\mathbf{R}_1 = \mathbf{R}_2 = \dots = \mathbf{R}_K = \mathbf{R}_c = \mathbf{R}$, \mathbf{R} 为 $(M-1) \times (M-1)$ 维矩阵, 其对角元素为 1, 其余元素为 0.5。假设不同目标产生的 RDOA 观测量互相独立。定位性能用均方误差 (Mean Square Error, MSE) 表征, 即 $\text{MSE}(\mathbf{u}_k) = \sum_{l=1}^L \|\mathbf{u}_{k,l} - \mathbf{u}_k^o\|^2 / L$, 其中 $\mathbf{u}_{k,l}$ 为第 l 次蒙特卡罗仿真的目标位置估计值, L 为蒙特卡罗仿真次数, $L = 5000$ 。

进行以下 3 项仿真: 仿真 1 为基于标校源辅助的单目标定位性能仿真; 仿真 2 为基于标校源辅助的多目标定位性能仿真; 仿真 3 探究标校源位置对多目标定位性能的影响。

1) 仿真 1

单目标定位几何构型如图 3 所示, 目标位置设为 $\mathbf{u}^o = [2000, 2500, 3000]^T \text{ m}$, 标校源位置设为 $\mathbf{c} = [1500, 1550, 1500]^T \text{ m}$ 。

在标校源辅助下, 使用本文算法对单目标的定位性能仿真如图 4 所示。“□”为单目标定位性能 (MSE) 随定位站位置误差的变化曲线, 虚线表示使用标校源校正定位的 CRLB, 实线表示没有定位站位置误差下的 CRLB ($f(\mathbf{C}(\mathbf{u}))_3$)。为便于比较, 同时进行了不存在标校源时的单目标定

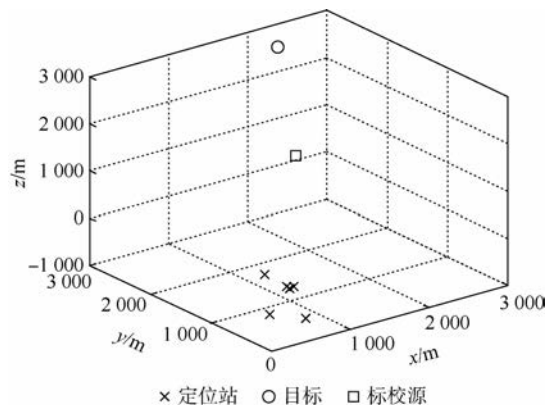


图 3 仿真 1 目标定位几何构型

Fig. 3 Geometry for source localization in Simulation 1

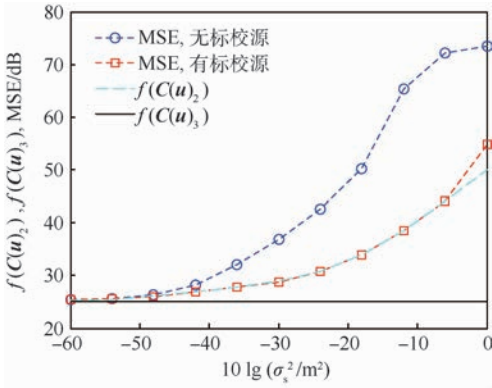


图 4 单目标定位性能比较

Fig. 4 Comparison of single source localization performance

位性能(MSE)仿真,用“○”表示。

从图 4 中可以看出,当定位站位置误差小于 -6 dB 时,本文算法的定位精度可以达到 CRLB,即可以对单目标进行有效定位;当定位站位置误差大于 -6 dB 时,发生门限效应,定位精度迅速偏离 CRLB;当定位站位置误差大于 -50 dB 时,采用标校源辅助的单目标定位性能要优于未采用标校源辅助的单目标定位性能,且随着定位站位置误差的增大,性能改善程度更加明显;使用标校源校正定位,定位性能仍达不到没有定位站位置误差下的定位精度。

2) 仿真 2

多目标定位几何构型如图 5 所示,符号含义同仿真 1。为简化处理,本仿真仅对 2 个不相交目标进行定位,目标数量可多于 2 个。设目标 1 为远场目标,位置为 $\mathbf{u}_1^0 = [2000, 2500, 3000]^T$ m,目标 2 为近场目标,位置为 $\mathbf{u}_2^0 = [600, 650, 550]^T$ m,标校源位置为 $\mathbf{c} = [750, 500, 400]^T$ m,定位站位置不变。

在标校源辅助下,使用本文算法对 2 个目标的定位性能仿真如图 6 所示。曲线含义同仿

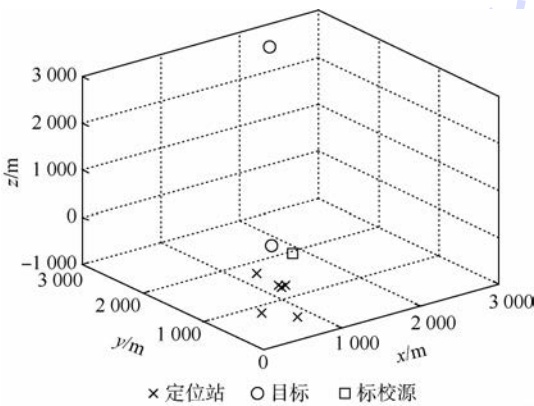


图 5 仿真 2 目标定位几何构型

Fig. 5 Geometry for source localization in Simulation 2

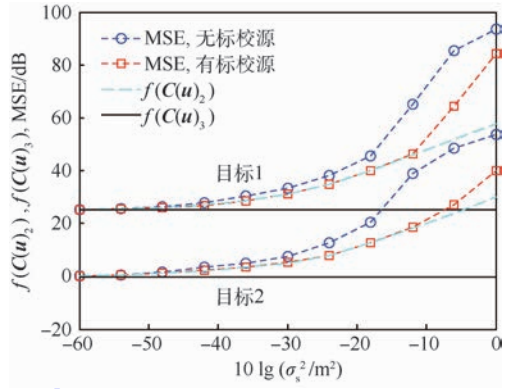


图 6 多目标定位性能比较

Fig. 6 Comparison of multi-source localization performance

真 1,在此不再赘述,其中定位性能较差的一组曲线为目标 1 的定位性能曲线,定位性能较好的一组曲线为目标 2 的定位性能曲线。

从图 6 中可以看出,不管是对近场目标还是远场目标,采用本文算法均可以完成有效定位。当定位站位置误差小于 -12 dB 时,对目标 1 的定位精度可以达到 CRLB;当定位站位置误差大于 -12 dB 时,发生门限效应;当定位站位置误差大于 -24 dB 时,本方算法的定位性能要明显优于未采用标校源辅助的定位性能,但达不到没有定位站位置误差下的定位精度。当定位站位置误差小于 -12 dB 时,对目标 2 的定位精度可以达到 CRLB;当定位站位置误差大于 -6 dB 时,发生门限效应;当定位站位置误差大于 -24 dB 时,本方算法的定位性能要明显优于未采用标校源辅助的定位性能。

3) 仿真 3

多目标定位几何构型如图 7 所示。对 2 个不相交目标进行仿真,设目标 1 位置为 $\mathbf{u}_1^0 = [314, 483, 209]^T$ m,目标 2 位置为 $\mathbf{u}_2^0 = [600, 650, 550]^T$ m,标校源 1 位置为 $\mathbf{c}_1 = [750, 500, 400]^T$ m,标校源 2 位置为 $\mathbf{c}_2 = [1500, 1550, 1500]^T$ m,定位站位置不变,同仿真 1。

改变标校源位置,采用本文算法对两目标的定位性能仿真如图 8 所示。“□”表示使用标校源 1 进行校正定位时目标位置估计性能随定位站位置误差的变化曲线,“×”表示使用标校源 2 校正定位时的定位性能曲线,“○”表示无标校源辅助下的定位性能曲线。其中定位性能较好的一组曲线为目标 1 的定位性能曲线,定位性能较差的一组曲线为目标 2 的定位性能曲线。

从图 8 中可以看出,无论标校源处于位置 1 还是位置 2,不管对目标 1 还是目标 2,本文算法的定位性能均优于未使用标校源时的定位性

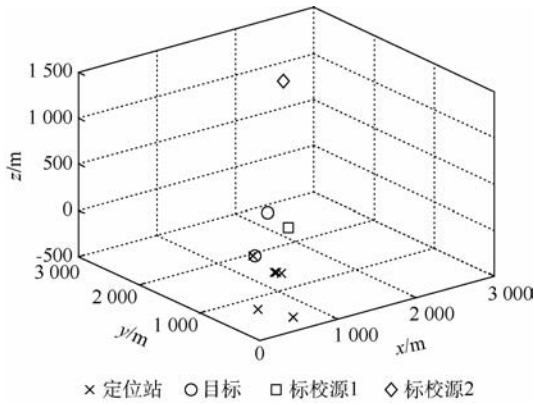


图7 仿真3目标定位几何构型

Fig.7 Geometry for source localization in Simulation 3

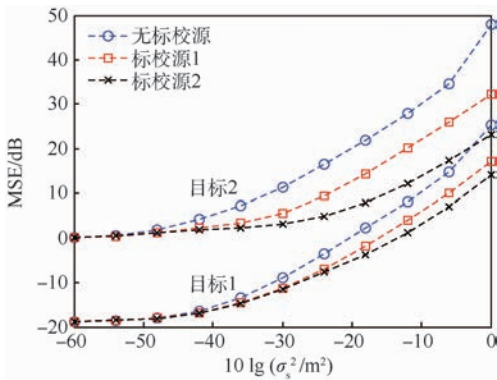


图8 改变标校源位置时多目标定位性能比较

Fig.8 Comparison of multi-source localization performance via changing calibration position

能。但标校源位置不同,对目标的定位性能改善程度也不同:对目标1,在定位站位置误差大于 -18 dB时,采用标校源1定位精度至少改善4 dB,采用标校源2定位精度至少改善5.8 dB;对目标2,在定位站位置误差大于 -24 dB时,采用标校源1定位精度至少改善7 dB,采用标校源2定位精度至少改善11.6 dB。

文献[19]表明,若使用差分校正(Differential Calibration, DC)技术,标校源离未知目标越近,定位性能改善程度越大。但仿真3结果表明,采用本文算法进行定位,离未知目标(目标1或目标2)更远的标校源反而提供了更好的定位性能。这是因为DC技术与本文算法对定位站位置的校正机理不同,至于如何设置标校源位置,使本文算法对定位精度的改善做出更大贡献仍有待研究。

5 结论

本文提出了一种基于标校源辅助的不相交多目标时差定位算法。

1) 首先利用标校源校正定位站位置,然后基

于校正的定位站位置,通过TS-WLS算法对多目标进行时差定位。

2) 该算法不需要初始值估计,不用对定位站和目标位置进行联合估计,通过闭式代数运算同时估计多目标位置,降低了计算量。

3) 标校源的引入虽增加了部分计算复杂度,但经CRLB分析表明,使用标校源理论上可以提高对多目标的定位精度。

4) 对该算法进行了定位性能分析和仿真验证,仿真结果表明:①对多目标,不管是近场目标还是远场目标,使用标校源校正技术均可以提高目标定位精度;②在定位站位置误差和时差观测误差较小时,算法定位性能可以达到CRLB;③标校源的位置不同,对多目标的定位精度改善程度不同。

参考文献 (References)

- [1] YANG H, CHUN J. An improved algebraic solution for moving target localization in non-coherent MIMO radar systems [J]. IEEE Transactions on Signal Processing, 2016, 64 (1): 258-270.
- [2] LI Y R, HO K C. Efficient closed-form estimators for multi-static sonar localization [J]. IEEE Transactions on Aerospace and Electronic Systems, 2015, 51 (1): 600-614.
- [3] SHI J, FAN L, ZHANG X L, et al. Multi-target positioning for passive sensor network via bistatic range space projection [J]. Science China Information Science Letter, 2016, 59: 019302: 1-019302: 3.
- [4] LIU C, FANG D Y, JIANG H B, et al. RSS distribution-based passive localization and its application in sensor networks [J]. IEEE Transactions on Wireless Communications, 2016, 15 (4): 2883-2895.
- [5] WANG Y, HO K C. TDOA positioning irrespective of source range [J]. IEEE Transactions on Signal Processing, 2017, 65 (6): 1447-1460.
- [6] NGUYEN N H, DOGANCA Y K. Single-platform passive emitter localization with bearing and Doppler-shift measurements using pseudolinear estimation techniques [J]. Signal Processing, 2016, 125: 336-348.
- [7] YIN F, ZHAO Y X, FREDRIK G, et al. Received-signal-strength threshold optimization using Gaussian processes [J]. IEEE Transactions on Signal Processing, 2017, 65 (8): 2164-2177.
- [8] NGUYEN N H, DOGANCA Y K. Multistatic pseudolinear target motion analysis using hybrid measurements [J]. Signal Processing, 2017, 130: 22-36.
- [9] FOY W H. Position-location solution by Taylor-series estimation [J]. IEEE Transactions on Aerospace and Electronic Systems, 1976, 12 (2): 187-194.
- [10] TORRIERI D J. Statistical theory of passive location systems [J]. IEEE Transactions on Aerospace and Electronic Systems,

1984,20(2):183-198.

- [11] SMITH J O, ABEL J S. Closed-form least-squares source location estimation from range-difference measurements [J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1987,35(12):1661-1669.
- [12] SCHAU H C, ROBINSON A Z. Passive source localization employing intersecting spherical surfaces from time-of-arrival differences [J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1987,35(8):1223-1225.
- [13] CHAN Y T, HO K C. A simple and efficient estimator for hyperbolic location [J]. IEEE Transactions on Signal Processing, 1994,42(8):1905-1915.
- [14] EINEMO M, SO H C. Weighted least squares algorithm for target localization in distributed MIMO radar [J]. Signal Processing, 2015,115:144-150.
- [15] YANG H, CHUN J, CHASE D. Hyperbolic localization in MIMO radar system [J]. IEEE Antennas Wireless Propagation Letter, 2015,14:618-621.
- [16] REZA M V, MICHAEL R B. On the CRLB of TDOA-FDOA estimation from MIMO signals [C] // IEEE/ION Position, Location & Navigation Symposium-plans. Piscataway, NJ: IEEE Press, 2016:772-778.
- [17] HO K C, KOVAVISARUCH L, PARIKH H. Source localization using TDOA with erroneous receiver positions [C] // Proceedings of the International Symposium on Circuits and Systems. Piscataway, NJ: IEEE Press, 2004:453-456.
- [18] HO K C, LU X N, KOVAVISARUCH L, et al. Source localization using TDOA and FDOA measurements in the presence of receiver location errors: Analysis and solution [J]. IEEE Transactions on Signal Processing, 2007,55(2):684-696.
- [19] HO K C, YANG L. On the use of a calibration emitter for source localization in the presence of sensor position uncertainty [J]. IEEE Transactions on Signal Processing, 2008,56(12):5758-5772.
- [20] YANG L, HO K C. An approximately efficient TDOA localization algorithm in closed-form for locating multiple disjoint sources with erroneous sensor positions [J]. IEEE Transactions on Signal Processing, 2009,57(1):4598-4615.
- [21] SUN M, HO K C. An asymptotically efficient estimator for TDOA and FDOA positioning of multiple disjoint sources in the presence of sensor location uncertainties [J]. IEEE Transactions on Signal Processing, 2011,59(7):3434-3440.

作者简介:

秦兆涛 男,博士研究生。主要研究方向:无源定位、雷达信号处理。

王俊 男,博士,教授,博士生导师。主要研究方向:数字信号实时处理、雷达信号处理。

陶磊岩 男,硕士,高级工程师。主要研究方向:实时信号处理等。

魏少明 男,博士,实验师。主要研究方向:目标跟踪、雷达散射中心提取、目标三维重构。

TDOA localization of multiple disjoint sources based on a calibration emitter

QIN Zhaotao¹, WANG Jun¹, TAO Leiyan², WEI Shaoming^{1,*}

(1. School of Electronic and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;

2. Beijing Remote Sensing Equipment Institute, Beijing 100854, China)

Abstract: The passive source localization accuracy can be greatly reduced due to the sensor position error. A localization algorithm using time difference of arrival (TDOA) for multiple disjoint sources was proposed, which improves the localization accuracy by means of a single calibration emitter. The sensor position error was first reduced by using the calibration emitter, and the corresponding error statistical knowledge was estimated. Then based on the updated sensor position, TDOA localization of multiple disjoint sources with high accuracy was realized by utilizing the algorithm of two-step weighted least squares (TS-WLS). The Cramer-Rao lower bound (CRLB) was theoretically derived to analyze the localization performance of the closed-form algorithm. And the theoretical derivation was validated by the simulations. The simulation results indicate that the localization accuracy of multiple disjoint sources is obviously improved by using calibration correction technique. Moreover, the solution performance is shown to reach the CRLB under small TDOA observation error and sensor position error. The developed estimator is computationally attractive because it does not require initial solution estimation and iterative computation. Furthermore, joint estimation between source positions and sensor positions is not needed, which reduces the calculation amount.

Keywords: calibration emitter; multiple disjoint sources; time difference of arrival (TDOA); localization accuracy; Cramer-Rao lower bound (CRLB)

Received: 2017-06-05; **Accepted:** 2017-09-08; **Published online:** 2017-09-27 17:13

URL: kns.cnki.net/kcms/detail/11.2625.V.20170927.1713.001.html

Foundation items: National Natural Science Foundation of China (61501012, 61671035)

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

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

DOI: 10.13700/j.bh.1001-5965.2017.0304

直线驱动电静液作动器的匹配设计规则

王子蒙, 焦宗夏*, 李兴鲁

(北京航空航天大学 自动化科学与电气工程学院, 北京 100083)



摘 要: 电静液作动器(EHA)是一种高度集成的泵控传动系统,A380和F35等机型已经利用EHA驱动主飞控舵面。但传统的以旋转电机和柱塞泵构成的EHA面临低频响的问题,以一种具有良好动态特性的新原理直线驱动电静液作动器(LEHA)为对象,研究其参数匹配设计问题。LEHA的关键特征在于新型的直线协同配流泵,以及采用直线电机直接驱动泵的吸排油和配流组件。首先,从静态指标匹配性上考虑LEHA的参数设计规则,得到最大空载速度约束条件、最大静态输出力约束条件、系统最高压力约束条件;然后,根据系统模型分析系统各参数对频宽的影响,得到LEHA的动态性能匹配设计规则,具体是指LEHA频宽指标对直线电机振荡频率的约束条件;最后,分析LEHA的功率约束条件,给出了LEHA对惯性负载、弹性负载和黏性阻尼负载在输出力-速度坐标下的功率包络条件,得到LEHA的负载匹配设计约束。LEHA参数设计的6项匹配设计约束条件能够为LEHA的设计过程提供理论依据。

关键词: 电静液作动器(EHA); 设计规则; 参数设计; 直线泵; 直线振荡电机

中图分类号: TH137.31

文献标识码: A

文章编号: 1001-5965(2018)05-1037-11

飞机作动系统由功率液传走向功率电传(power-by-wire)是当前的发展趋势^[1-3],高性能的功率电传作动器是功率电传系统的一项关键部件。目前飞机功率电传作动器的技术方案主要包括机电作动器(Electro-Mechanical Actuator, EMA)^[4]和电静液作动器(Electro-Hydrostatic Actuator, EHA)^[5]。其中EMA适合于小功率作动需求,而EHA面向大功率作动需求更有优势^[6-7]。

EHA本质上是一种泵控的静液传动系统,采用旋转电机驱动柱塞泵,然后由柱塞泵输出的流体介质来推动作动筒^[8]。EHA的变量控制方式有3种:变转速电机+定排量泵、定转速电机+变排量泵和变转速电机+变排量泵方案^[9-11]。由于变排量泵的结构复杂,并且双变量控制难度高,因

此目前一般不采用双变量控制结构。变转速电机+定排量泵的配置是主流的EHA方案,通过控制电机的旋转速度和方向来控制泵的输出流量大小和吸排油方向,从而控制液压作动筒的输出位移。对于变转速控制的EHA,受电机的驱动力矩和旋转部件的惯量比的限制,目前研制成功的EHA动态特性不如电液伺服作动器^[12-14]。

针对小功率驱动场合,Anderson和Lindler^[15]面向无人机舵面的作动需求开发了功能材料驱动的固液复合作动器,具有结构简单、体积小的优势。这种功能材料驱动的固液复合作动器的设计核心在于开发新原理的直线泵^[16],现有单向阀配流的直线泵^[17-19]及完全主动配流的直线泵^[20]。由于功能材料的变形率低,上述采用功能材料驱

收稿日期: 2017-05-12; 录用日期: 2017-11-19; 网络出版时间: 2017-12-19 14:43

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20171219.1026.003.html

基金项目: 国家“973”计划(2014CB046406); 国家自然科学基金(51235002)

* 通信作者. E-mail: zxjiao@buaa.edu.cn

引用格式: 王子蒙, 焦宗夏, 李兴鲁. 直线驱动电静液作动器的匹配设计规则[J]. 北京航空航天大学学报, 2018, 44(5): 1037-1047. WANG Z M, JIAO Z X, LI X L. Matching design rules of linear-driven electro-hydrostatic actuator[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 1037-1047 (in Chinese).

动的直线泵输出流量小,因此以其构成的固液复合作动器输出功率偏小,并且输出速度低。

针对旋转电机+柱塞泵驱动 EHA 的动态特性问题和功能材料驱动固液复合作动器的功率不足问题,文献[21]提出了一种基于直线驱动及互配流原理的新型 EHA,简称直线驱动电静液作动器 (Linear-driven Electro-Hdrostatic Actuator, LEHA),LEHA 的核心创新点在于提出了一种具有多个单元的主动配流直线泵原理,采用直线电磁元件直接驱动直线泵的吸排油和配流部件,构成 EHA。

文献[21]主要给出了直线泵的配流机理的证明,给出了直线泵的多单元拓扑连接方式,说明了协同配流直线泵相对于单向阀配流直线泵具有的优势。LEHA 是一种具有新原理的电静液作动器,还没有形成完善的设计方法,本文将研究 LEHA 的参数设计问题,给出 LEHA 中关键参数的匹配设计约束规则。

1 直线驱动电静液作动器原理

LEHA 与传统 EHA 的主要区别在于其所采用的直线驱动方式,LEHA 利用直线振荡电机驱动直线泵的吸排油和配流部件,在直线泵的输出口得到稳定的流量,进而驱动液压作动筒以构成静液传动作动器,其原理如图 1 所示,图中: x_c 为作动筒位移; Q_1 和 Q_2 分别为流入 2 个容腔的流量; p_1 和 p_2 分别为 2 个容腔的压力; F_1 为负载力。

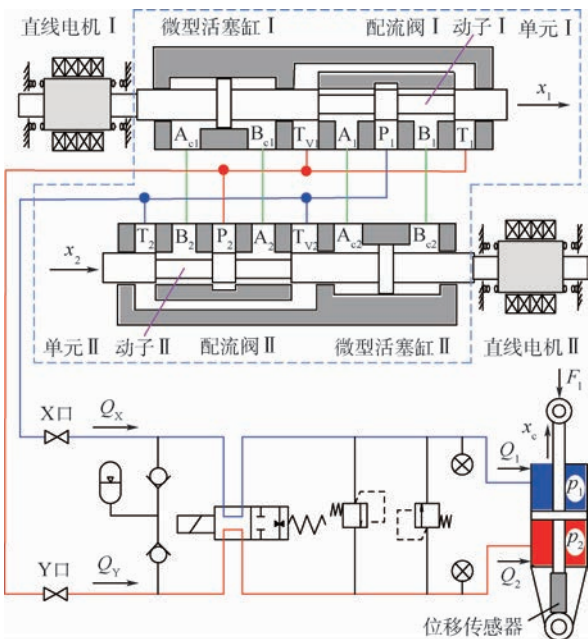


图 1 LEHA 原理
Fig. 1 Schematic of LEHA

LEHA 中的直线泵具有 2 个相似的单元,一个单元包含一个微型活塞缸和一个三位四通配流阀,微型液压缸的活塞与滑阀阀芯在结构上整合为一个轴状零件,微型液压缸的两腔分别与另一个单元配流阀的 A 口及 B 口相连、一个单元配流阀的 P 口与另一个单元配流阀的 T_V 口相连。根据上述油路连接规则,可以看出直线泵的单元之间具有协同配流的特征,即一个单元的吸排油部件产生的流量由另一个单元的滑阀进行配流,因此该直线泵定义为协同配流直线泵。

为描述直线泵的几何流量,定义符号函数

$$\text{sgn}(x) = \begin{cases} -1 & x < 0 \\ 0 & x = 0 \\ 1 & x > 0 \end{cases} \quad (1)$$

对于图 1 所示的由 2 个单元构成的直线泵,不考虑油液的压缩性、泵的泄漏等因素,其输出流量与 2 个单元的位移 x_1 、 x_2 的关系,即直线泵的几何流量为^[21]

$$\begin{cases} Q_X = \text{sgn}(x_2) \cdot A_p \dot{x}_1 - \text{sgn}(x_1) \cdot A_p \dot{x}_2 \\ Q_Y = -\text{sgn}(x_2) \cdot A_p \dot{x}_1 + \text{sgn}(x_1) \cdot A_p \dot{x}_2 \end{cases} \quad (2)$$

式中: Q_X 为直线泵 X 口的输出流量; Q_Y 为直线泵 Y 口的输出流量; A_p 为直线泵的吸排油活塞面积; \dot{x}_1 和 \dot{x}_2 分别为单元 I 和单元 II 的速度。

可以看出,在交互配流原理的作用下,一个单元的阀芯位置决定了由另一个单元吸排油活塞运动速度产生的流量的方向。为讨论在正弦函数作用下的双单元直线泵几何流量,定义 2 个单元的运动函数为

$$\begin{cases} x_1 = S_p \sin(\omega t) \\ x_2 = S_p \sin(\omega t + \varphi) \end{cases} \quad (3)$$

式中: φ 为 2 个单元的相位差; ω 为角频率; S_p 为吸排油活塞振幅。

则 2 个单元的运动速度为

$$\begin{cases} \dot{x}_1 = \omega S_p \cos(\omega t) \\ \dot{x}_2 = \omega S_p \cos(\omega t + \varphi) \end{cases} \quad (4)$$

将式(3)及式(4)代入式(2),得到直线泵的几何流量为

$$\begin{cases} Q_X = \omega A_p S_p (\text{sgn}(\sin(\omega t + \varphi)) \cos(\omega t) - \text{sgn}(\sin(\omega t)) \cos(\omega t + \varphi)) \\ Q_Y = -\omega A_p S_p (\text{sgn}(\sin(\omega t + \varphi)) \cos(\omega t) - \text{sgn}(\sin(\omega t)) \cos(\omega t + \varphi)) \end{cases} \quad (5)$$

对式(5)给出几何流量在时域上的表达式在一个周期 T 内进行积分,得到泵在一个运行周期内的排油体积为

$$V_T = \int_0^T Q_x dt = 8A_p S_p \sin \varphi \quad (6)$$

定义 f_m 为 2 个单元的往复运动频率,则单位时间内的排油体积为

$$V_s = f_m V_T = 8A_p S_p f_m \sin \varphi \quad (7)$$

从式(7)可以看出泵的排量与吸排油活塞面积和振幅成正比,与相位差有正弦运算的关系;要改变直线泵流量的方向,只能调节单元动子的相位差。

2 LEHA 的指标需求和参数设计

LEHA 作为一种高度集成的电静液作动器,其参数设计涉及到系统与外界交互作用的参数匹配问题,根据 LEHA 的系统性能指标条件得到其元部件的工作范围约束条件;同时也涉及到系统内部元部件之间的匹配问题,如直线电机输出力、振幅、频率和直线泵的吸排油活塞面积的匹配问题,定义其主要性能指标如表 1 所示。

根据 LEHA 的原理图可以得到 LEHA 的主要设计参数包括直线电机的输出力、振幅和谐振频率,还包括直线泵的吸排油活塞面积,以及作动筒的活塞面积和行程,定义 LEHA 中需要设计的参数如表 2 所示。

表 1 LEHA 性能指标

Table 1 Performance parameter of LEHA

参 数	符 号
作动器最大空载速度	$v_{act-max}$
作动器最大静态输出力	$F_{act-max}$
作动器最大角功率	$P_{act-max}$
-3 dB, -90° 频宽	f_{band}
系统最高压力	P_{max}

表 2 LEHA 的关键参数

Table 2 Key parameters of LEHA

参 数	符 号
直线泵吸排油活塞面积	A_p
吸排油活塞振幅	S_p
直线电机振荡频率	f_m
直线电机最大电磁力	F_{e-max}
作动筒活塞面积	A_{cyl}
作动筒行程	S_{cyl}

3 LEHA 的参数匹配设计规则

LEHA 的设计规则推导过程为:首先分析静态指标匹配设计问题,然后根据系统模型分析系统各参数对频宽的影响,得到 LEHA 的动态性能匹配设计规则,最后从系统功率的充足性和匹配性上讨论负载带来的设计约束条件。

3.1 静态指标设计规则

3.1.1 最大空载速度约束条件

考虑到相位差不理想带来的影响,直线泵在单位时间内的排油体积如式(7)所述。实际上由于间隙的存在,直线泵还必然存在一定的泄漏,将直线泵本身的内泄漏利用直线泵的容积效率 η_p 来表示,则可以定义直线泵的平均流量为

$$Q_{pa} = 8\eta_p A_p S_p f_m \sin \varphi \quad (8)$$

对于作动筒而言,若不考虑其外泄漏和两腔之间的内泄漏,则作动筒速度 v_{cyl} 与直线泵平均流量之间的关系为

$$v_{cyl} = Q_{pa} / A_{cyl} \quad (9)$$

实际上对于作动筒内泄漏的建模,部分文献将其内泄漏流量 Q_{leak} 简单建模为

$$Q_{leak} = C_{leak} \Delta p \quad (10)$$

式中: C_{leak} 为泄漏系数; Δp 为两腔压差。

定义作动筒的容积效率为 η_c , 则有

$$v_{act} = \eta_c Q_{pa} / A_{cyl} \quad (11)$$

式中: v_{act} 为作动筒速度。

则作动筒最大空载速度为

$$v_{act-max} = 8\eta_p \eta_c A_p S_p f_m / A_{cyl} \quad (12)$$

式(12)给出了 LEHA 的速度约束条件,即根据作动器最大空载速度得到 A_p 、 S_p 、 f_m 和 A_{cyl} 之间的约束关系。定义式(12)为 LEHA 参数设计的第 1 项约束条件。

3.1.2 系统最高压力约束条件

直线泵的出口压力与负载大小和直线泵本身的能力有关,在此只考虑负载充足的情况,即只考虑电机和泵组合系统的指标对直线泵输出压力的影响。直线电机的最大电磁力为 F_{e-max} , 若按照直线泵输出压力与电机输出力静态对应的关系,系统最高压力为

$$p_{max} = F_{e-max} / A_p \quad (13)$$

实际上由于直线泵的主动配流作用,直线电机的位置闭环控制效果会影响直线泵的输出性能,因此从一个周期内能量转换的角度来分析系统压力与设计变量的关系。考虑 $[0, T]$ 时间段的工作周期,若 2 个直线电机的振荡波形为三角函数曲线,定义电机位移为

$$\begin{cases} x_1 = S_p \sin(2\pi f_m t) \\ x_2 = S_p \sin(2\pi f_m t + \varphi) \end{cases} \quad (14)$$

则电机速度为

$$\begin{cases} v_1 = 2\pi f_m S_p \cos(2\pi f_m t) \\ v_2 = 2\pi f_m S_p \cos(2\pi f_m t + \varphi) \end{cases} \quad (15)$$

假设电机始终工作在第一象限,即电磁力与

动子速度始终同方向,则2个电机在一个周期内输出的机械能为

$$E_{\text{motor}} = \int_0^T (F_{e-\text{max}} - F_{\text{fric}}) |v_1| dt + \int_0^T (F_{e-\text{max}} - F_{\text{fric}}) |v_2| dt \quad (16)$$

将摩擦力 F_{fric} 的影响转换为效率 η_m 来考虑,定义电机输出力

$$F_m = F_{e-\text{max}} - F_{\text{fric}} = \eta_m F_{e-\text{max}} \quad (17)$$

当相位差 $\varphi = \pi/2$,结合式(14)~式(17),得到2个电机在一个周期内输出的机械能为

$$E_{\text{motor}} = \eta_m F_{e-\text{max}} \cdot 2\pi f_m S_p \left(\int_0^T |\cos(2\pi f_m t)| dt + \int_0^T |\cos(2\pi f_m t + \pi/2)| dt \right) = 2\pi \eta_m F_{e-\text{max}} S_p f_m \left(\int_0^T |\cos(2\pi f_m t) - \sin(2\pi f_m t)| dt \right) = 8\eta_m F_{e-\text{max}} S_p \quad (18)$$

假设直线泵的输出压力为恒定值 p_{max} ,结合式(8),在一个周期内直线泵输出的液压能为

$$E_{\text{pump}} = \int_0^T p_{\text{max}} Q_{\text{pa}} dt = 8\eta_p A_p S_p p_{\text{max}} \sin \varphi \quad (19)$$

当相位差 $\varphi = \pi/2$,有

$$E_{\text{pump}} = 8\eta_p A_p S_p p_{\text{max}} \quad (20)$$

定义直线泵输出液压能量与输入机械能量 $E_{\text{p-in}}$ 的关系为 $E_{\text{pump}} = \eta_{\text{m-p}} E_{\text{p-in}}$, $\eta_{\text{m-p}}$ 为直线泵效率系数,因此得

$$8\eta_p A_p S_p p_{\text{max}} = \eta_{\text{m-p}} \cdot 8\eta_m F_{e-\text{max}} S_p \quad (21)$$

则系统最高压力 p_{max} 可以表达为

$$p_{\text{max}} = \eta_{\text{m-p}} \eta_m F_{e-\text{max}} / (\eta_p A_p) \quad (22)$$

式(22)给出了 p_{max} 与 $F_{e-\text{max}}$ 和 A_p 的关系,可以看出直线泵的输出压力与电机驱动力、吸排油活塞面积与驱动及配流过程中的效率相关,效率系数受配合间隙、直线电机的位移控制效果影响。

定义系数 $K_\eta = \eta_{\text{m-p}} \eta_m / \eta_p$,显然根据物理意义可知其值域为 $[0, 1]$ 。定义式(22)为 LEHA 参数设计的第2项约束条件。

3.1.3 最大静态输出力约束条件

得到了直线泵的输出压力与设计参数之间的约束表达式,然后考虑作动器的最大静态输出力 $F_{\text{act-max}}$ 与设计参数之间的约束关系。在堵转情况下,作动筒输出力与系统最高压力的关系为

$$F_{\text{act-max}} = p_{\text{max}} A_{\text{cyl}} \quad (23)$$

结合式(22),有

$$F_{\text{act-max}} = \eta_{\text{m-p}} \eta_m F_{e-\text{max}} A_{\text{cyl}} / (\eta_p A_p) \quad (24)$$

式(24)描述了作动器最大静态输出力对 $F_{e-\text{max}}$ 、 A_p 和 A_{cyl} 的设计约束条件,定义式(23)为

LEHA 参数设计的第3项约束条件。

3.2 动态性能匹配设计规则

3.2.1 LEHA 的线性化模型

1) 直线振荡电机模型

根据电压平衡方程、电磁力与电流的关系和动子质量力平衡方程,直线振荡电机的模型为

$$\begin{cases} u = K_v \dot{x} + Ri + Ldi/dt \\ F_e = K_e i \\ m_e \ddot{x} + c\dot{x} + F_l + F_f + K_s x = F_e \end{cases} \quad (25)$$

式中: K_v 为电机的反电动势系数,由电机的几何结构及永磁体材料特性决定; R 为线圈电阻; L 为线圈电感; K_e 为电机的力系数; m_e 为电机动子质量; c 为电机的黏性摩擦系数; F_l 为除黏性摩擦力外的其他摩擦项; K_s 为弹簧刚度系数; u 为绕组电压; i 为绕组电流; F_e 为电磁力。

若不考虑除黏性摩擦力外的其他摩擦项,根据式(25)得到直线振荡电机输出位移 x_m 关于控制变量 u 和干扰变量 F_l 的传递函数为

$$x_m = K_e u / [Lm_e s^3 + (Lc + Rm_e) s^2 + (LK_s + Rc + K_e K_v) s + RK_s] + (Ls + R) F_l / [Lm_e s^3 + (Lc + Rm_e) s^2 + (LK_s + Rc + K_e K_v) s + RK_s] \quad (26)$$

首先不考虑负载的作用,即令 $F_l = 0$,以电压 u 为输入,以位移 x_m 为输出,得到电机的线性传递函数为

$$x_m = K_e u / [Lm_e s^3 + (Lc + Rm_e) s^2 + (LK_s + Rc + K_e K_v) s + RK_s] \quad (27)$$

实际上,由于直线电机的电感为毫亨级,因此传递函数分母的三次项系数很小,可以将三阶系统简化为二阶系统,简化传递函数为

$$x_m = \frac{K_e}{Rm_e s^2 + (Rc + K_e K_v) s + RK_s} u \quad (28)$$

式(28)二阶系统的标准形式表达式为

$$x_m = \frac{K_m \omega_m^2}{s^2 + 2\zeta_m \omega_m s + \omega_m^2} u \quad (29)$$

式中:自然角频率 $\omega_m = \sqrt{K_e/m_e}$;阻尼比 $\zeta_m = (Rc + K_e K_v) / (2R \sqrt{K_e m_e})$;增益系数 $K_m = 1/R$ 。

考虑实际的电机及其驱动系统模型,电机驱动器运行在电流模式,采用运算放大器搭建增益电路和积分电路来实现 PI 控制器,对电机进行流闭环控制,驱动器的输入指令为电压,电压指令到电流指令的增益为 K_{dr} 。

2) 液压作动筒模型

作动筒输出位移关于输入流量和负载力的传递函数为

$$x_{act} = \frac{4\beta A_{cyl}}{s(m_{cyl}V_t s^2 + B_{cyl}V_t s + 4\beta A_{cyl}^2)} Q_{cyl} - \frac{V_t}{m_{cyl}V_t s^2 + B_{cyl}V_t s + 4\beta A_{cyl}^2} F_{leyl} \quad (30)$$

式中: β 为油液弹性模量; B_{cyl} 为作动筒黏性摩擦系数; V_t 为两腔容积; Q_{cyl} 为作动筒油口流量; F_{leyl} 为作动筒活塞杆所受负载力; m_{cyl} 为作动筒动子质量。

为分析设计参数对 LEHA 的频率特性影响,先不考虑负载力的影响,则有

$$x_{act} = \frac{1}{s} \cdot \frac{K_{cyl} \omega_{cyl}^2}{s^2 + 2\zeta_{cyl} \omega_{cyl} s + \omega_{cyl}^2} Q_{cyl} \quad (31)$$

式中:自然角频率 $\omega_{cyl} = \sqrt{4\beta A_{cyl}^2 / (m_{cyl} V_t)}$;阻尼比

$\zeta_{cyl} = B_{cyl} \sqrt{V_t} / (4A_{cyl} \sqrt{m_{cyl} \beta})$;增益系数 $K_{cyl} = 1/A_{cyl \circ}$

3) LEHA 系统模型

对 LEHA 的控制框图而言,可以将直线泵简化描述为一个比例环节,其输入变量对输出变量的影响“瞬时”完成,不具有时延特性,实际上是将直线泵作为一项传动比环节。直线电机和作动筒的线性化模型都是三阶系统,输入变量对输出变量的影响需要经过储能环节的延迟作用。

综合直线电机的控制框图模型和作动筒的框图模型,并且将直线泵表示为变增益的比例环节,得到 LEHA 的控制框图模型如图 2 所示。

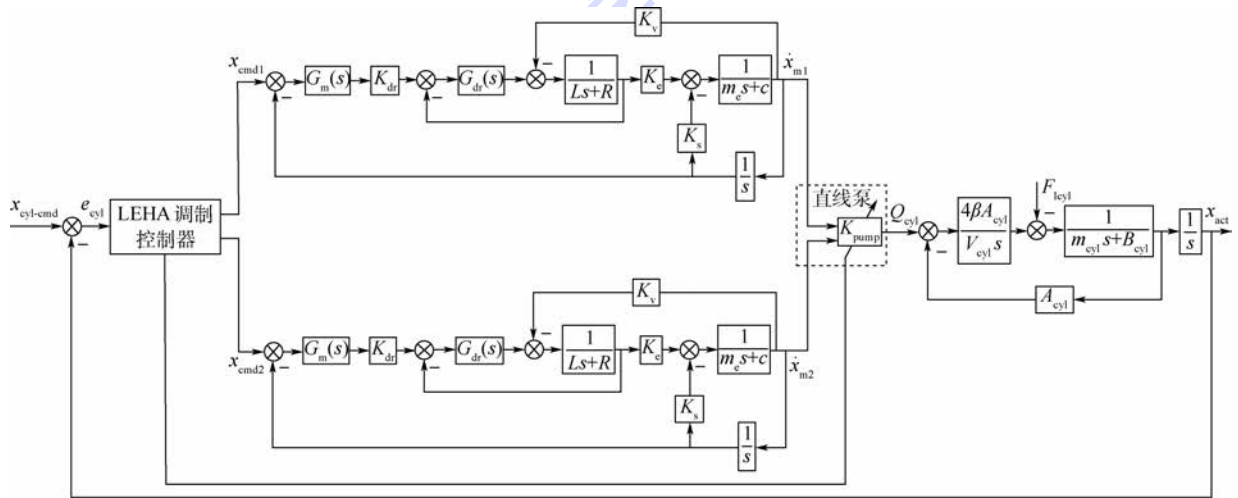


图 2 LEHA 控制框图

Fig. 2 Control block diagram of LEHA

3.2.2 LEHA 频宽与调制频率的关系

为分析调制频率对 LEHA 频率特性的影响,统计不同频率的结果得到作动器在该调制频率时的频率特性;为了对比不同电机振荡频率时的作动器频宽,以作动器行程的 5% 作为正弦位移指令的幅值,仿真计算电机振荡频率为 10 ~ 100 Hz 时的作动器频率特性。对于给定的 2 mm、5 Hz 正弦位移指令,在调制频率为 30、100 Hz 时,LEHA 的位移响应和直线电机位移响应分别如图 3 和图 4 所示。

从图 3 和图 4 可以看出,调制频率越高,在一个作动器运动周期内直线泵的吸排油周期数越多,从图 4(b)可以看出电机振荡曲线的幅值包络线为正弦曲线,相邻周期的振幅变化小,振荡电机的调幅控制效果好。根据 LEHA 的调制工作原理,调制频率越高,直线泵的流量调节速度越快,作动器的输出调节速度越快,LEHA 具有更好的位移响应特性。

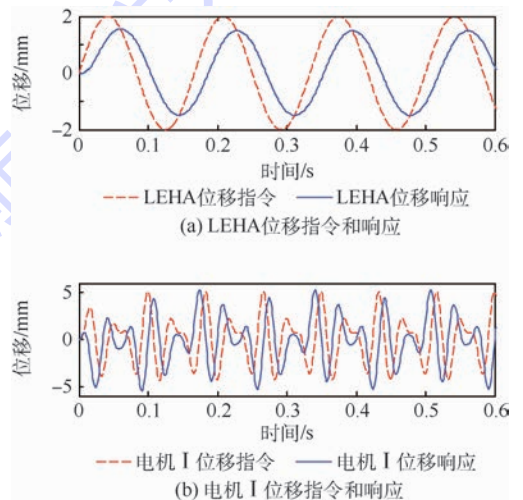


图 3 30 Hz 调制频率的 LEHA 位移响应

Fig. 3 Displacement response of LEHA when modulation frequency is 30 Hz

统计在 10 ~ 100 Hz 调制频率范围内 LEHA 的 -3 dB, -90° 频宽,在不同调制频率下 LEHA 的频宽如图 5 所示。可以看出调制频率越高,LEHA

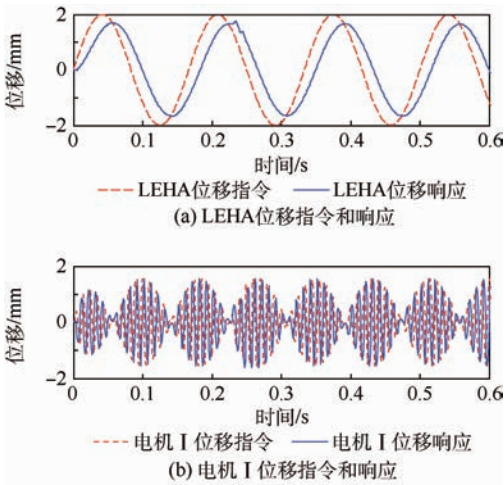


图4 100 Hz 调制频率的 LEHA 位移响应

Fig. 4 Displacement response of LEHA when modulation frequency is 100 Hz

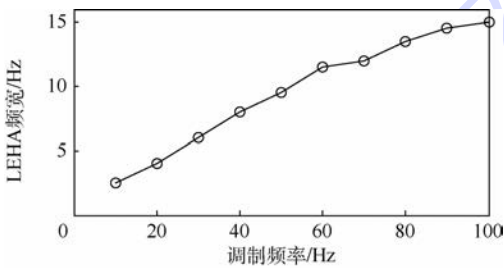


图5 LEHA 频宽与调制频率的关系

Fig. 5 Relationship between bandwidth of LEHA and modulation frequency

的频宽越高,但是增长速率随着调制频率的增加而下降(见图5)。

3.2.3 动态性能设计规则

本节将在适度简化的基础上,分析 LEHA 中各项参数对系统动态性能的影响,得到基于 LEHA 的频宽性能需求的直线电机振荡频率约束条件。

在图2所示的 LEHA 线性化框图模型中,直线泵被简化为变增益的比例环节,其作用是将直线电机的输出位移转换为泵的输出流量信号。由于泵内的容腔很小,因此不考虑油液的可压缩性,可以忽略油液弹性带来的直线泵模型中的时延特性,其对系统动态性能的影响在于增益系数会影响整个系统前向通道的增益,从而影响系统的响应特性。

对于 LEHA 而言,由于直线电机、作动筒等子系统的叠加作用,其频率特性比任何单独的子系统都差;因此要使得 LEHA 系统的频宽达到期望值 f_{band} ,则直线电机和液压作动筒的固有频率都应该明显高于 f_{band} 。引入增益系数 K_{frep} 和 K_{frec} ,因此有 $\omega_m \geq K_{frep} \omega_{band}$ 及 $\omega_{cyl} \geq K_{frec} \omega_{band}$ 。

$$\begin{cases} f_m \geq K_{frep} f_{band} \\ \sqrt{\frac{4\beta A_{cyl}^2}{m_{cyl} V_t}} \geq 2\pi K_{frec} f_{band} \end{cases} \quad (32)$$

式中:油液弹性模量 β 的取值与油液空气溶解量密切相关,按工程经验可以取 400 MPa。

实际上作动筒的液压固有频率一般很高,对于 LEHA 需求的 10 Hz 左右的频宽,液压固有频率不会成为限制作动器频宽的主要因素,因此式(32)只需要考虑直线电机的频率约束条件。

系数 K_{frep} 实际上反映了 LEHA 频宽指标对直线电机振荡频率的最低要求,关于直线电机振荡频率对 LEHA 性能的影响,与电机的控制器效果密切相关,式(32)中参数 K_{frep} 取值可以从图5所给出的 LEHA 频宽与调制频率的关系中得到。结合式(32)和图5可以得到 LEHA 频宽指标对直线电机振荡频率的约束条件,定义其为 LEHA 参数设计的第4项约束条件。

3.3 负载匹配设计规则

作动器作为一种驱动系统,会受到各种类型的负载,典型的负载包括:惯性负载、弹性负载和黏性阻尼负载等,或者以上典型负载的混合作用。作动器在驱动对象时,系统必须具有足够的功率,一般将负载力和负载速度的关系定义为负载特性,以负载力和负载速度为坐标所绘制的关系曲线称为负载轨迹。本节根据 LEHA 的角功率需求给出一项约束设计规程,然后针对惯性负载、弹性负载和黏性阻尼负载3种情况讨论 LEHA 的参数设计规则。

3.3.1 作动器最大角功率约束

根据一般舵机的设计规范,作动器的最大角功率定义为最大静态输出力与最大空载速度的积,即

$$P_{act-max} = F_{act-max} v_{act-max} \quad (33)$$

结合式(12)和式(24),有

$$P_{act-max} = \frac{\eta_{m-p} \eta_m F_{e-max} A_{cyl}}{\eta_p A_p} \cdot \frac{8\eta_p \eta_c A_p S_p f_m}{A_{cyl}} \quad (34)$$

即

$$P_{act-max} = 8\eta_m \eta_{m-p} \eta_c F_{e-max} S_p f_m \quad (35)$$

定义式(35)为 LEHA 参数设计的第5项约束条件。

3.3.2 惯性负载的参数设计约束

作动器驱动的对象惯量定义为 m_{obj} ,则惯性负载力为

$$F = m_{obj} \ddot{x}_{cyl} \quad (36)$$

讨论位移 x 为三角函数曲线的情况,即 $x = x_0 \sin(\omega_1 t)$,则负载轨迹方程为

$$\begin{cases} \dot{x} = \omega_1 x_0 \cos(\omega_1 t) \\ F = -m_{\text{obj}} x_0 \omega_1^2 \sin(\omega_1 t) \end{cases} \quad (37)$$

式(37)可以等效变形为椭圆方程

$$\left(\frac{\dot{x}}{\omega_1 x_0}\right)^2 + \left(\frac{F}{m_{\text{obj}} x_0 \omega_1^2}\right)^2 = 1 \quad (38)$$

作动器驱动上述惯性负载所需的功率为

$$F\dot{x} = -m_{\text{obj}} x_0^2 \omega_1^3 \cos(\omega_1 t) \sin(\omega_1 t) \quad (39)$$

作动器实际的功率特性曲线应该由实验测得,然后根据负载匹配条件:功率特性曲线包络负载曲线,可以得到作动器参数的设计约束条件。但是在设计阶段,无法确定在 F - v 坐标下 LEHA 具体的输出功率特性曲线,但是其范围不会超过最大输出力和最大速度的乘积而得到的等功率曲线。根据式(35)得到 LEHA 角功率对应的等功率曲线为

$$F\dot{x} = 8\eta_m \eta_{m-p} \eta_c F_{e-\text{max}} S_p f_m \quad (40)$$

显然式(40)所定义的 LEHA 输出功率不小于式(39)定义的驱动惯性负载的需求功率,因此只需求式(39)所定义函数的最大值。

定义 $\theta = \omega_1 t$, 函数 $f_1(\theta) = m_{\text{obj}} x_0^2 \omega_1^3 \cos \theta \sin \theta$, 则函数 $f_1(\theta)$ 在定义域 $[0, 2\pi]$ 内的最大值为

$$f_{1\text{max}}(\theta) = \frac{1}{2} m_{\text{obj}} x_0^2 \omega_1^3 \quad (41)$$

结合式(40)和式(41),得到 LEHA 对于式(36)所定义惯性负载的功率包络设计约束条件为

$$F_{e-\text{max}} S_p f_m \geq \frac{m_{\text{obj}} x_0^2 \omega_1^3}{16\eta_m \eta_{m-p} \eta_c} \quad (42)$$

3.3.3 弹性负载的参数设计约束

作动器驱动的弹性负载力为

$$F = Kx \quad F \leq F_{\text{max}} \quad (43)$$

式中: K 为负载弹性系数。

讨论位移 x 为三角函数曲线的情况,定义

$$x = x_0 \sin(\omega_1 t) \quad (44)$$

则负载方程为

$$\begin{cases} \dot{x} = \omega_1 x_0 \cos(\omega_1 t) \\ F = Kx_0 \sin(\omega_1 t) \end{cases} \quad (45)$$

式(45)实质上是一个椭圆方程,等价变换为

$$\left(\frac{F}{Kx_0}\right)^2 + \left(\frac{\dot{x}}{\omega_1 x_0}\right)^2 = 1 \quad (46)$$

根据相切条件,等功率线方程为

$$F\dot{x} = K\omega_1 x_0^2 \cos(\omega_1 t) \sin(\omega_1 t) \quad (47)$$

定义 $\theta = \omega_1 t$, 函数 $f_2(\theta) = K\omega_1 x_0^2 \cos \theta \sin \theta$, 则函数 $f_2(\theta)$ 在定义域 $[0, 2\pi]$ 内的最大值为

$$f_{2\text{max}}(\theta) = \frac{1}{2} K\omega_1 x_0^2 \quad (48)$$

结合式(40)和式(48),得到 LEHA 对于弹性负载的功率包络设计约束条件为

$$F_{e-\text{max}} S_p f_m \geq \frac{K\omega_1 x_0^2}{16\eta_m \eta_{m-p} \eta_c} \quad (49)$$

3.3.4 黏性阻尼负载的参数设计约束

黏性阻尼负载力为

$$F = B_1 \dot{x} \quad x \leq x_{\text{max}} \quad (50)$$

讨论位移 x 为三角函数曲线的情况,即

$$x = x_0 \sin(\omega_1 t) \quad (51)$$

则负载方程为

$$\begin{cases} \dot{x} = \omega_1 x_0 \cos(\omega_1 t) \\ F = B_1 \omega_1 x_0 \cos(\omega_1 t) \end{cases} \quad (52)$$

式中: B_1 为阻尼系数。

在 F - \dot{x} 坐标系下,黏性负载轨迹为直线 $\dot{x} = F/B_1$, 因此 LEHA 的等功率曲线只需包络该直线离远点最远的端点,则有

$$8\eta_m \eta_{m-p} \eta_c F_{e-\text{max}} S_p f_m \geq B_1 \dot{x}_{\text{max}}^2 \quad (53)$$

因此 LEHA 对黏性阻尼负载的参数设计约束条件为

$$F_{e-\text{max}} S_p f_m \geq \frac{B_1 \omega_1^2 x_0^2}{8\eta_m \eta_{m-p} \eta_c} \quad (54)$$

总结得到 LEHA 的 6 项参数设计约束条件如表 3 所示。

表 3 LEHA 参数设计的 6 项约束条件

Table 3 Six constraint conditions for LEHA parameter design

项 目	约束条件内容
条件 1: 速度约束	$v_{\text{act-max}} = 8\eta_p \eta_c A_p S_p f_m / A_{\text{cyl}}$
条件 2: 压力约束	$p_{\text{max}} = \eta_{m-p} \eta_m F_{e-\text{max}} / (\eta_p A_p)$
条件 3: 输出力约束	$F_{\text{act-max}} = p_{\text{max}} A_{\text{cyl}}$
条件 4: 频宽约束	$f_m \geq K_{\text{freq}} f_{\text{band}}, K_{\text{freq}} = f(f_{\text{band}})$
条件 5: 角功率约束	$P_{\text{act-max}} = 8\eta_m \eta_{m-p} \eta_c F_{e-\text{max}} S_p f_m$
条件 6: 负载约束	惯性负载 $F_{e-\text{max}} S_p f_m \geq \frac{m_{\text{obj}} x_0^2 \omega_1^3}{16\eta_m \eta_{m-p} \eta_c}$ 弹性负载 $F_{e-\text{max}} S_p f_m \geq \frac{K\omega_1 x_0^2}{16\eta_m \eta_{m-p} \eta_c}$ 黏性阻尼负载 $F_{e-\text{max}} S_p f_m \geq \frac{B_1 \omega_1^2 x_0^2}{8\eta_m \eta_{m-p} \eta_c}$

4 LEHA 参数设计算例

LEHA 的主要参数指标包括最大空载速度、最大静态输出力、系统功率、动态频宽指标和系统工作压力水平等,以小功率的 EHA 为例,定义如表 4 所示的 LEHA 指标需求,3 种负载条件如表 5 所示。效率系数取值分别为 $\eta_m = 0.95, \eta_{m-p} = 0.8, \eta_p = 0.9, \eta_c = 0.9$ 。

根据 Liang 等开展的高功率密度直线振荡电机研究^[22],根据经验值将吸排油活塞振幅 S_p

表4 LEHA性能指标需求

Table 4 Performance parameter requirement of LEHA

参数	数值
作动器最大空载速度 $v_{act-max}/(m \cdot s^{-1})$	0.06
作动器最大静态输出力 $F_{act-max}/kN$	6
作动器最大角功率 $P_{act-max}/W$	360
-3 dB, -90°频宽 f_{band}/Hz	6

表5 LEHA负载参数

Table 5 Parameters of LEHA load

负载类型	参数值	负载条件取值
惯性负载	$m_{obj} = 60 \text{ kg}$	
$F = m_{obj} \ddot{x}_{cyl}$	$x_0 = 0.002 \text{ m}$	$1.87 \text{ N} \cdot \text{m} \cdot \text{s}^{-1}$
$x = x_0 \sin(\omega_1 t)$	$\omega_1 = 14\pi$	
弹性负载	$K = 1.5 \times 10^5 \text{ N} \cdot \text{m}^{-1}$	
$F = Kx$	$x_0 = 0.04 \text{ m}$	$65.3 \text{ N} \cdot \text{m} \cdot \text{s}^{-1}$
$x = x_0 \sin(\omega_1 t)$	$\omega_1 = \pi$	
黏性阻尼负载	$B_1 = 600 \text{ N} \cdot \text{m}^{-1} \cdot \text{s}$	
$F = B_1 \dot{x}$	$x_0 = 0.04 \text{ m}$	$26.3 \text{ N} \cdot \text{m} \cdot \text{s}^{-1}$
$x = x_0 \sin(\omega_1 t)$	$\omega_1 = 4\pi$	

设计为 $\pm 5 \text{ mm}$, 作动筒行程 S_{cyl} 选取为 $\pm 40 \text{ mm}$, 系统最高压力 p_{max} 选取为 6 MPa 。其他需要设计的参数包括: 吸排油活塞面积 A_p 、电机振荡频率 f_m 、最大电磁力 F_{e-max} 、作动筒活塞面积 A_{cyl} 。

根据输出力约束条件式(23), 结合密封件的标准尺寸, 选取作动筒的活塞杆直径为 33 mm , 活塞直径为 48 mm , 则作动筒活塞面积 A_{cyl} 为 954 mm^2 。

根据角功率的定义, 约束条件1、2和5实际上有一项是重复条件, 可以得到设计参数 A_p 、 f_m 、 F_{e-max} 需要满足的约束条件为

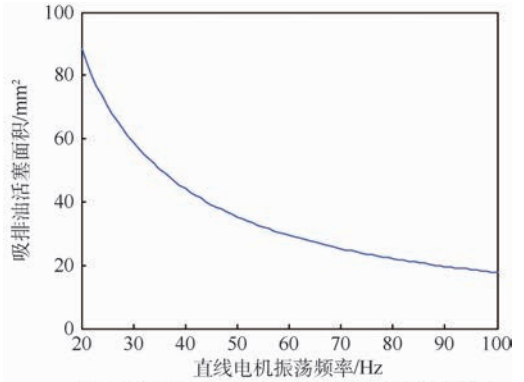
$$\begin{cases} A_p f_m = \frac{v_{act-max} A_{cyl}}{8 \eta_p \eta_c S_p} \\ F_{e-max} f_m = \frac{P_{act-max}}{8 \eta_m \eta_{m-p} \eta_c S_p} \end{cases}$$

代入参数值, 得到 A_p 、 f_m 、 F_{e-max} 需要满足的约束如图6所示。

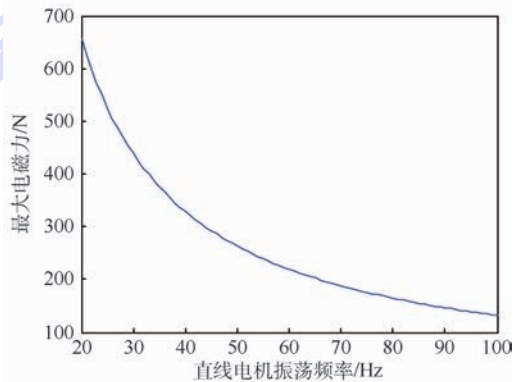
从图6可以看出, 提高直线电机的振荡频率, 可以在保证充足流量的条件下减小直线泵的吸排油活塞面积, 从而减小直线电机的最大电磁力需求, 进而减小电机的尺寸和质量, 提升直线电机的功率重量比, 同时也能提高LEHA的功率重量比。

根据LEHA频宽对直线电机振荡频率的约束条件, 当LEHA频宽指标为 6 Hz 时, 根据图5给出的分析结果, 直线振荡电机的谐振频宽不低于 30 Hz 。针对表4给出的指标需求, LEHA的关键参数设计结果如表6所示。

根据表6给出的设计参数, LEHA中直线电



(a) 直线电机振荡频率-吸排油活塞面积约束关系



(b) 直线电机振荡频率-最大电磁力约束关系

图6 直线电机参数约束条件

Fig. 6 Parameter constraint condition of linear motor

表6 LEHA的关键参数取值

Table 6 Key parameter values of LEHA

参数	数值
直线泵吸排油活塞面积 A_p/mm^2	59
吸排油活塞振幅 S_p/mm	± 5
直线电机振荡频率 F_m/Hz	30
直线电机最大电磁力 F_{e-max}/N	439
作动筒活塞面积 A_{cyl}/mm^2	954
作动筒行程 S_{cyl}/mm	± 40

机最大电磁力、电机振幅和振荡频率的乘积为

$$F_{e-max} S_p f_m = 65.85 \text{ N} \cdot \text{m} \cdot \text{s}^{-1}$$

显然表6给出的设计参数满足负载约束条件, 因此根据本文所述的6项参数设计规则, 能够针对具体的指标完成LEHA的参数设计。

5 LEHA设计指标实验验证

根据表6所示的设计参数加工LEHA的原理样机, 如图7所示。

LEHA原理样机中直线电机的振荡频率为 30 Hz , 直线电机力系数为 27.5 N/A 。实际LEHA样机中直线泵的参数与表6所给参数不同, 吸排油活塞直径为 10.5 mm , 活塞杆直径为 7 mm , 直线泵的吸排油活塞面积为 48 mm^2 。

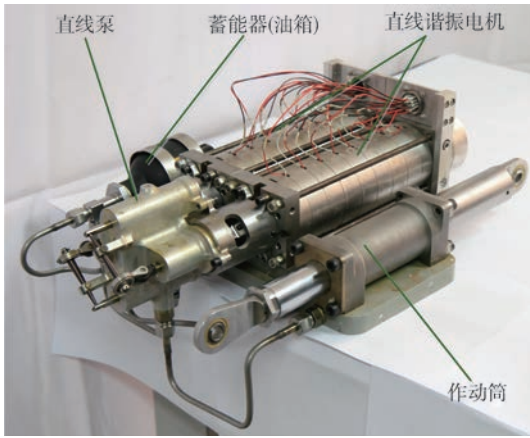


图7 LEHA 原理样机

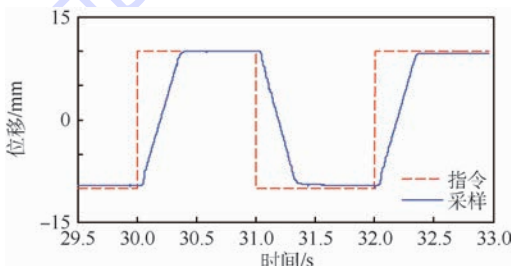
Fig. 7 LEHA prototype

5.1 作动筒最大空载速度测试

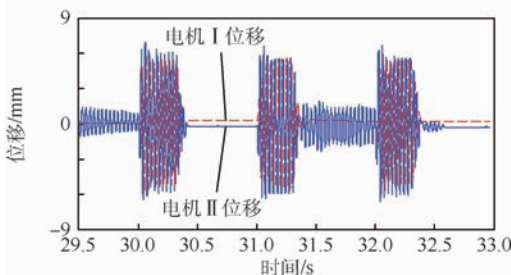
以幅值为 $\pm 10\text{ mm}$ 的低频方波信号为位移指令,测试 LEHA 的位移跟踪响应,如图 8 所示。在上升沿和下降沿时间段内,电机实际输出振幅为 6 mm ,LEHA 的速度为 0.057 m/s 。根据式(12),LEHA 速度与直线泵振幅正相关,因此若直线电机振幅为 5 mm ,按照直线泵样机 48 mm^2 吸排油活塞面积,LEHA 速度为 0.0475 m/s ;若按图 6(a)给出的约束条件,当直线电机振荡频率为 30 Hz 时,将直线泵吸排油活塞面积设计为 59 mm^2 ,则 LEHA 能达到 0.06 m/s 的最大空载速度。

绕组电流如图 9 所示。

直线泵输出压力的平均值为 5.6 MPa ,电机振荡波形在峰值处存在一定变形,但是过零点时刻 2 个电机的相位差接近 90° ,直线泵的配流效率高。图 9 中,直线电机电流峰值为 13 A ,对应电磁力为 358 N ,当直线泵的吸排油活塞面积为 48 mm^2 时,根据式(22)给出的压力约束条件,直线泵输出压力为 5.95 MPa ,与实测值 5.7 MPa 的误差为 4.2% ,实验结果与按照设计规则给出的参数算例相符合。



(a) LEHA位移指令和采样



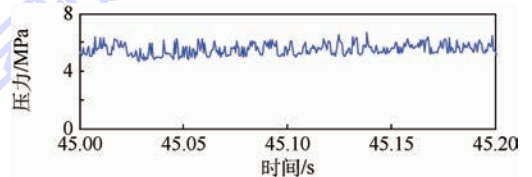
(b) 电机 I 和电机 II 位移采样

图8 LEHA 最大空载速度

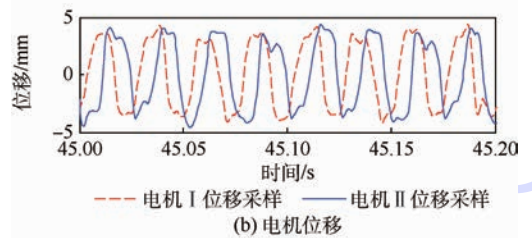
Fig. 8 Maximum no-load velocity of LEHA

5.2 直线泵最高输出压力测试

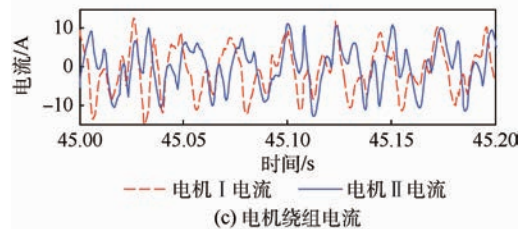
以节流阀作为直线泵负载,测量直线泵的最高输出压力,电机位移、直线泵的输出压力和电机



(a) 直线泵输出压力



(b) 电机位移



(c) 电机绕组电流

图9 直线泵性能测试结果

Fig. 9 Testing results of linear pump performance

5.3 LEHA 调制频率特性测试

目前 LEHA 原理样机中直线电机的振荡频率设计为 30 Hz ,实验测试调制频率在 $10\sim 30\text{ Hz}$ 范围内 LEHA 的频率特性,统计结果如图 10 所示, A 为输出幅值与指定幅值之比。

LEHA 的频宽随调制频率的增长而增加,如图 11 所示,实验结果与图 5 仿真结果符合,调制频率为 30 Hz 时,LEHA 频宽为 5.5 Hz 。

LEHA 的工作原理本质上是利用直线泵的互配流机制将 2 个电机的高频调制振荡位移解调成可控的作动筒位移。电机的振荡频率越高,单位时间内调制信号的波形周期数越多,在利用一个包络线进行调幅时,每 2 个相邻周期的幅值变化率越小,这样使得相邻周期的波形具有更好的相似性。对于不同的电机振荡频率,如果波形的畸变率能保证大致相当,则在高频下电机的响应速

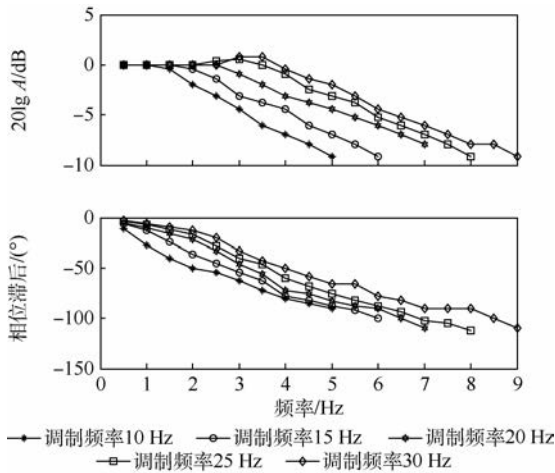


图10 不同调制频率时的LEHA频率特性

Fig. 10 LEHA frequency response to different modulation frequency

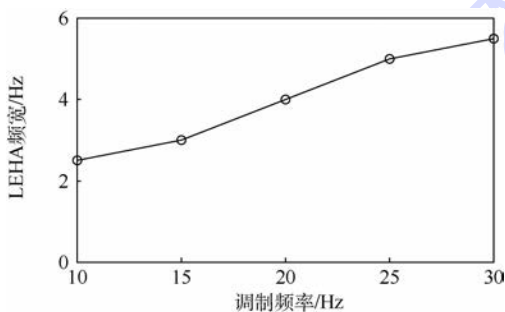


图11 LEHA频宽与调制频率的关系(实验结果)

Fig. 11 Relationship between bandwidth of LEHA and modulation frequency (experimental result)

度越快;并且调制频率越高,在由直线泵将多个通道的运动整合变换成可控流量时,具有更精密的调节效果;因此提高内部通道的调制频率,即提高直线电机的振荡频率,将使LEHA达到更高的频宽。

6 结论

LEHA作为一种新原理的泵控式电静液作动器,涉及到机械、电气和液压等多个学科,本文研究了LEHA的参数匹配设计规则,得到了LEHA参数设计的6项匹配规则,包括:最大空载速度约束条件、最大静态输出力约束条件、系统最高压力约束条件、动态频宽约束条件、角功率约束条件和负载匹配约束条件,上述设计规则能够为LEHA的设计过程提供充足的理论依据,指导LEHA的参数设计过程;并且通过LEHA原理样机的实验测试验证了其参数设计规则。

根据LEHA的参数设计约束条件可知:

1) 直线电机的振荡频率越高,LEHA频宽越

高,因此应该尽量提高直线电机的谐振频率。

2) 在同等的流量需求下,直线电机的振荡频率越高,直线泵的吸排油活塞面积越小;在保证系统压力的前提下,直线电机的最大电磁力需求越小,因此电机体积可以设计得越小,因此可以提升系统的功率重量比。

参考文献 (References)

- [1] WHEELER P, BOZHKO S. The more electric aircraft: Technology and challenges[J]. IEEE Electrification Magazine, 2014, 2(4): 6-12.
- [2] WHEELER P. Technology for the more and all electric aircraft of the future[C]// IEEE International Conference on Automatica. Piscataway, NJ: IEEE Press, 2016: 1-5.
- [3] SARILOGLU B, MORRIS C T. More electric aircraft: Review, challenges, and opportunities for commercial transport aircraft[J]. IEEE Transactions on Transportation Electrification, 2015, 1(1): 54-64.
- [4] DERRIEN J C, SECURITE S D. Electro-mechanical actuator (EMA) advanced technologies for flight controls[C]// 28th International Congress of the Aeronautical Sciences, 2012: 1-10.
- [5] ALLE N, HIREMATH S S, MAKARAM S, et al. Review on electro hydrostatic actuator for flight control[J]. International Journal of Fluid Power, 2016, 17(2): 125-145.
- [6] BOTTEN S L, WHITLEY C R, KING A D. Flight control actuation technology for next-generation all-electric aircraft[J]. Technology Review Journal, 2000, 23(6): 55-68.
- [7] CHAKRABORTY I, TRAWICK D R, JACKSON D, et al. Electric control surface actuator design optimization and allocation for the more electric aircraft[C]// 2013 Aviation Technology, Integration, and Operations Conference. Reston: AIAA, 2013: 4283.
- [8] VAN DEN BOSSCHE D. The A380 flight control electro-hydrostatic actuators, achievements and lessons learnt[C]// 25th International Congress of the Aeronautical Sciences, 2006: 1-8.
- [9] ANDERSON J A. Variable displacement electro-hydrostatic actuator[C]// IEEE Aerospace and Electronics Conference. Piscataway, NJ: IEEE Press, 1991: 529-534.
- [10] SHEN W, MAI Y, SU X, et al. A new electric hydraulic actuator adopted the variable displacement pump[J]. Asian Journal of Control, 2016, 18(1): 178-191.
- [11] JIANG J H, LIU Q, LIU J L, et al. Performance comparison of displacement and speed control in electro-hydrostatic actuator[J]. Applied Mechanics and Materials, 2015, 779: 205-211.
- [12] HANKAN C, TUNA B, BULENT E P. A complete analysis for single rod electro hydrostatic actuators[C]// 10th International Fluid Power Conference. Dresden: IFAS, 2016: 119-132.
- [13] NAVARRO R. Performance of an electro-hydrostatic actuator on the F-18 systems research aircraft; NASA/TM-97-206224[R]. Edwards; Dryden Flight Research Center, 1997: 1-37.
- [14] MCCULLOUGH. Design and characterization of a dual electro-hydrostatic actuator[D]. Hamilton: McMaster University, 2011: 4-11.

- [15] ANDERSON E, LINDLER J. Smart material actuator with long stroke and high power output [C] // 43rd Structures, Structural Dynamics, and Materials Conference. Reston: AIAA, 2002:1354.
- [16] JOHN S, SIROHI J, WANG G, et al. Comparison of piezoelectric, magnetostrictive, and electrostrictive hybrid hydraulic actuators [J]. Journal of Intelligent Material Systems and Structures, 2007, 18(10):1035-1048.
- [17] KAN J, TANG K, REN Y, et al. Study on a piezohydraulic pump for linear actuators [J]. Sensors and Actuators A: Physical, 2009, 149(2):331-339.
- [18] CHAUDHURI A, WERELEY N M. Experimental validation of a hybrid electrostrictive hydraulic actuator analysis [J]. Journal of Vibration and Acoustics, 2010, 132(2):021006.
- [19] KIM G W, WANG K W. Helmholtz resonance in a piezoelectric-hydraulic pump-based hybrid actuator [J]. Smart Materials and Structures, 2010, 20(1):015010.
- [20] TAN H, HURST W, LEO D. Performance modeling of a piezohydraulic actuation system with active valves [J]. Smart Materials and Structures, 2004, 14(1):91-110.
- [21] LI Y, JIAO Z, YAN L, et al. Conceptual design and composition principles analysis of a novel collaborative rectification structure pump [J]. Journal of Dynamic Systems, Measurement, and Control, 2014, 136(5):054507.
- [22] LIANG H, JIAO Z, YAN L, et al. Design and analysis of a tubular linear oscillating motor for directly-driven EHA pump [J]. Sensors and Actuators A: Physical, 2014, 210:107-118.

作者简介:

王子蒙 男, 博士研究生。主要研究方向: 机电/电液作动系统仿真、设计与控制方法。

焦宗夏 男, 教授, 博士生导师。主要研究方向: 复杂机电装备设计与控制、飞控作动系统与余度舵机、流体传输管路振动机理的主动控制技术、机电系统的智能化设计仿真。

李兴鲁 男, 博士研究生。主要研究方向: 机电/电液作动系统仿真、设计与控制方法。

Matching design rules of linear-driven electro-hydrostatic actuator

WANG Zimeng, JIAO Zongxia*, LI Xinglu

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

Abstract: Electro-hydrostatic actuator (EHA) is a highly integrated pump controlled transmission system, and it has been successfully applied to A380 and F35 to drive main flight control surface. However, the traditional EHA composed by rotary motor and axial piston pump has limited dynamic response. This paper introduces a novel linear-driven electro-hydrostatic actuator (LEHA) with fast dynamic response, and studies the parameter design rules of LEHA. The key feature of LEHA is the novel linear pump with collaborative rectification principle. Pistons and valve spools in collaborative rectification pump are directly driven by linear motor. The parameter design rules of LEHA are firstly considered from the static index matching rules, and the constraint conditions of the maximum no-load speed, the constraint condition of the maximum static output force and the constraint condition of the maximum pressure of the system are deduced. According to LEHA's mathematic model, the influence of each parameter on the bandwidth of the LEHA is analyzed, and the dynamic performance design rule of LEHA is obtained, which is the constraint condition of the LEHA bandwidth index to the resonance frequency of the linear motor. The power constraint condition of LEHA is analyzed finally, and the power envelope condition of LEHA for inertial load, elastic load and viscous damping load is given in the output force-velocity coordinates, from which the power matching design rule of the LEHA is obtained. In conclusion, the 6 matching design constraints of LEHA parameter design rule can provide a theoretical basis for the design process of LEHA.

Keywords: electro-hydrostatic actuator (EHA); design rule; parameter design; linear pump; linear oscillating motor

Received: 2017-05-12; Accepted: 2017-11-19; Published online: 2017-12-19 14:43

URL: kns.cnki.net/kcms/detail/11.2625.V.20171219.1026.003.html

Foundation items: National Basic Research Program of China (2014CB046406); National Natural Science Foundation of China (51235002)

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

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

DOI: 10.13700/j.bh.1001-5965.2017.0334

导航卫星双频复用信号模拟失真的仿真



蔚小龙, 寇艳红*

(北京航空航天大学 电子信息工程学院, 北京 100083)

摘 要: 导航卫星有效载荷信号生成链路组成部件的非理想特性不可避免会引起发射导航信号的失真, 目前对于这种失真带来的信号测距性能损失并未有适用的解析模型。针对宽带双频复用信号的模拟失真, 采用相关损耗、S 曲线偏差和载波相位偏差作为评估指标, 仿真评估了载荷模拟组件中高功率放大器(HPA)、HPA 前置滤波器、多路复用器(OMUX)对 AltBOC、TD-AltBOC、ACEBOC 和 AltLOC 不同调制方式测距性能的影响。其中对于 HPA 使用了有记忆 HPA 行为模型, 考虑了不同 HPA 前置滤波器带宽和 HPA 工作点的影响, 并评估了失真信号相对于理想信号在宽带导频和单边带导频 2 种跟踪方式下的性能损失。仿真结果可以为信号调制方式比较评估、有效载荷信号生成链路的优化设计提供参考。

关键词: 恒包络复用; 模拟失真; 相关损耗; S 曲线偏差; 有记忆高功率放大器(HPA)

中图分类号: TN967.1

文献标识码: A **文章编号:** 1001-5965(2018)05-1048-08

卫星有效载荷是资源与功率受限的系统。为了提高功率效率, 载荷高功率放大器(High Power Amplifier, HPA)通常工作在饱和状态, 这种情况下非恒包络的导航信号经过 HPA 放大后会产生较大的非线性失真; 因而现代导航信号生成时在 HPA 放大前需要将同一频带的多个信号分量在数字基带进行恒包络复用调制; 将复用后的基带信号调制在单频载波上发射也可以节省有效载荷发射通道^[1]。虽然恒包络调制的基带信号在无限带宽的情况下包络恒定, 然而实际卫星 HPA 前置滤波器的带宽限制会破坏复用信号的恒包络特性, 导致 HPA 输出信号产生非线性失真; HPA 前置滤波器和 HPA 后继的多路复用器(Output Multiplexer, OMUX)二者的非理想滤波效应也会引起信号的线性失真, 导致卫星最终发射信号的接收性能变差^[2-3]。

北斗全球系统计划在 B2 频带 B2a 和 B2b 频

点分别播发 2 个不同服务的信号, 信号体制设计初期曾提出的 4 分量 AltLOC 调制方式因其非恒包络特性而被放弃^[1], 双频恒包络调制方式 AltBOC(15, 10)被采纳为 B2 频带下行信号的基线调制方式。另外, 中国的信号体制研究人员针对 B2 信号的双频恒包络复用也提出了 TD-AltBOC 和 ACEBOC 调制方式^[4-5]。目前为止, 公开发表的文献在分析不同调制方式下的信号测距精度时集中于研究热噪声、干扰和多径引起的跟踪误差标准差^[6-8], 很少提及卫星有效载荷非理想特性对上述几种双频复用调制方式的影响, 但是事实上有效载荷的非理想特性不可避免, 对其引起的相关损耗、S 曲线偏差和载波相位偏差等性能损失的分析目前也并未有适用的解析模型。目前关于载荷非理想特性影响分析的相关文献中对于信号接收性能的分析没有考虑不同的接收处理方式, 并且对于 HPA 行为模型一般只考虑了无记忆

收稿日期: 2017-05-19; 录用日期: 2017-06-09; 网络出版时间: 2017-06-12 10:55

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20170612.1055.001.html

基金项目: 国家自然科学基金(61271197)

* 通信作者. E-mail: kouy@buaa.edu.cn

引用格式: 蔚小龙, 寇艳红. 导航卫星双频复用信号模拟失真的仿真[J]. 北京航空航天大学学报, 2018, 44(5): 1048-1055.

YU X L, KOU Y H. Simulation of analog distortion of dual-frequency multiplexing signal generated by navigation satellite [J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 1048-1055 (in Chinese).

的 Saleh 模型和 Rapp 模型^[9-12]。但是当输入为宽带信号时,HPA 的记忆效应不容忽视^[2],特别 B2 信号又是典型的宽带信号。文献[13]采用有记忆 HPA 模型中的 Wiener 模型分析了载荷非理想特性对 AltBOC 信号相位噪声、相关峰形状和码/载波跟踪误差的影响,但在其仿真模型中没有考虑 HPA 功率回退的情况,也没有分析不同的 HPA 前置带宽及不同的接收机接收处理下载荷非理想特性对信号接收性能的影响。

由于目前并没有解析表达式能够简化载荷非理想特性引起的模拟失真对信号性能影响的分析,应通过针对具体调制方式和接收处理方式,并且逼近真实信道特性的仿真来初步评估这种影响,在条件成熟时再对实际载荷设备所发射的信号进行测试评估,从而为信号调制方式的比较评估和载荷电路的优化设计提供参考。本文首先基于实际工程器件特性建立包含 HPA 前置滤波器、HPA、OMUX 的有效载荷模拟组件非理想特性模型;其中针对 B2 宽带信号采用了有记忆 HPA 行为模型而非简单的无记忆模型,参考了文献[14]中根据实际 HPA 测试结果提取的模型参数;滤波器模型采用了具有群时延非线性的 IIR 模型而非简单的 FIR 滤波器,更接近真实信道射频滤波器的特性。然后选择相关损耗、S 曲线偏差和载波相位偏差这 3 个直接反映信道非理想特性对信号测距性能的影响,同时又无法通过解析算式简单估算的重要指标,通过仿真来评估不同 HPA 前置滤波器双边带宽及 HPA 工作点情况下非理想特性对 AltBOC、TD-AltBOC、ACEBOC 和 AltLOC 信号接收性能的影响,包括对宽带导频跟踪和下边带导频跟踪 2 种接收方式的影响。研究结果表明:同样的载荷模拟通道特性对不同调制方式信号性能的影响有很大差异,而对同一种调制方式在不同跟踪方式下的影响也有很大差异;并不

存在一组设计参数使得上述 3 种性能同时达到最优;载荷电路的优化设计应综合考虑其各个组成部件的联合影响而不能仅考察其中部分组件的影响,需要根据仿真以至实际设备试验结果和服务需求进行折中考虑。

1 模拟失真评估的仿真模型

AltBOC、TD-AltBOC、ACEBOC 和 AltLOC 4 种调制方式的差异主要体现在信号分量调制复用方式、副载波构造方式以及信号分量间的功率比例方面。在接收机端的信号跟踪方式上,上下边带联合宽带导频跟踪理论上可以提供更高的精度,但相应的接收机成本、复杂度较高,一般更常用的是仅利用单边带导频分量进行跟踪。为了评估不同信号通过卫星有效载荷导航信号生成链路模拟通道后的失真及其在不同接收处理方式下对测距性能的影响,本文建立了如图 1 所示的仿真模型。首先生成待评估调制方式的理想基带复用信号 S_{BB} ,然后将 S_{BB} 送入等效到基带的模拟失真通道中。其中 HPA 前置滤波器的带宽限制会破坏理想复用信号的恒包络特性,造成 HPA 引入非线性失真;此外 HPA 前置滤波器和 HPA 后继的 OMUX 的非理想幅频/相频响应也会造成信号的线性失真。接下来,若分析宽带导频跟踪方式,此时考察的是接收带宽为复用信号发射带宽的情况,所以直接将模拟失真通道输出的失真基带信号 $S_{BB-Distortion}$ 输入相关器;但是若分析单边带导频跟踪方式,还需要首先根据所跟踪边带的中心频点将该边带信号的主瓣变频到基带,再经过一个信号分量主瓣带宽的线性相位砖墙滤波器,经过这些预处理后的单边带失真基带信号 $S_{BB-PreProc-SSB}$ 再送入相关器。在相关器中,针对所分析的信号跟踪方式,选取复用信号 S_{BB} 中相应信号分量作为参考信号 S_{Ref} ,最终计算 S_{Ref} 与 $S_{BB-Distortion}$ 或

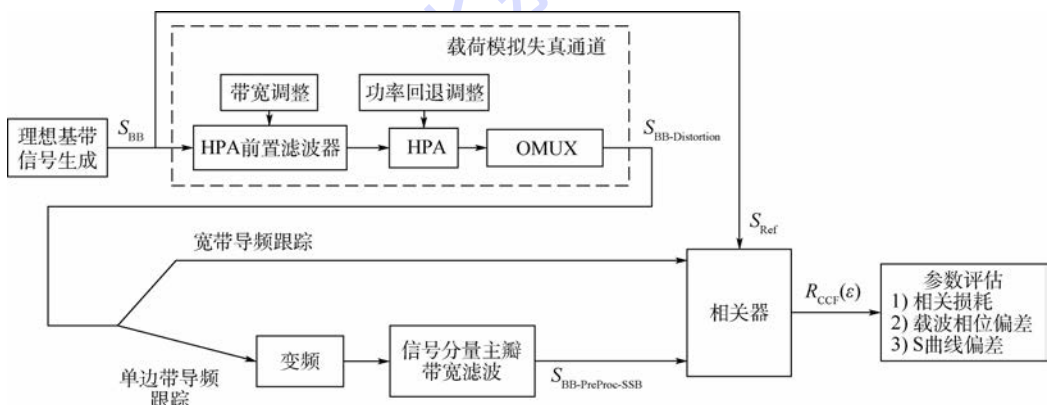


图 1 载荷模拟失真评估的仿真模型

Fig. 1 Simulation model for evaluation of payload analog distortion

$S_{\text{BB-PreProc-SSB}}$ 的互相关函数 $R_{\text{CCF}}(\varepsilon)$, 其中 ε 为码相位误差, 并评估载荷模拟失真引起的性能下降。

需说明的是, 载荷中模拟组件一般还包括上变频调制单元, 但是根据文献[13]分析, 正常情况下, 上变频调制单元引入的相位噪声对相关峰形状几乎没有影响, 仅仅会引起宽带导航信号的相位出现微小波动, 因此本文将忽略上变频调制单元对导航信号测距性能的影响。

1.1 HPA 前置滤波器和 OMUX

OMUX 的信道特性可近似等效为双边带宽为信号发射带宽的滤波器。在几种常用的射频滤波器类型中, 巴特沃斯滤波器的通带最为平坦, 实际应用最为广泛。与文献[13]相同, 本文将 HPA 前置滤波器和 OMUX 都建模为 6 阶巴特沃斯滤波器。

1.2 高功率放大器

在各种有记忆 HPA 行为模型中, 记忆多项式模型由于结构简单和参数易于求解而在功率放大器行为模型和预失真方面得到了广泛应用。记忆多项式模型离散形式表达式为^[14]

$$y(n) = \sum_{k=0}^K \sum_{q=0}^Q c_{kq} x(n-q) |x(n-q)|^k$$

式中: $x(n)$ 和 $y(n)$ 分别为 HPA 的输入和输出信号; K 和 Q 分别为记忆多项式模型的阶数和记忆深度; 由于 k 为偶数时产生的偶次项在通过后继的带通滤波器时会被滤除, 所以这里 k 只取奇数; c_{kq} 为记忆多项式模型的系数, 仿真中可采用文献[14]给出的一组提取自实际 AB 类功放的系数。

1.3 HPA 输出功率回退调整

为了改善非线性失真, 可以将 HPA 的工作点从饱和点回退一定功率值, 其代价是功率效率的降低, 不同功率回退点对非线性失真的影响不同。因此在将 HPA 前置滤波器滤波后的复用信号送入 HPA 之前, 需要进行幅度归一化, 然后根据 HPA 工作点的输出功率回退 (Output Back Off, OBO) 值 x_{OBO} 调整幅度^[10]。

需要注意的是, 根据有记忆 HPA 模型仿真得到的幅度失真 (Amplitude Modulation-Amplitude Modulation, AM-AM) 特性是一簇分散的点, 可以通过多项式拟合得到单一的 AM-AM 特性曲线。拟合曲线输出信号幅度最大值即为输出信号饱和幅度 $U_{\text{o,sat}}$, 由 HPA 工作点的 x_{OBO} 可以算得对应的输出信号幅度 $U_{\text{o,OBO}} = U_{\text{o,sat}} \sqrt{10^{x_{\text{OBO}}/10}}$, 再由拟合曲线插值得到输入信号幅度 $U_{\text{i,OBO}}$, 将幅度归一化后的滤波信号幅度乘以 $U_{\text{i,OBO}}$ 后再输入到 HPA。

2 评估指标

本文选取的相关损耗、S 曲线偏差和载波相位偏差 3 个评估指标都是基于如下的失真基带信号和本地参考信号的归一化互相关函数而定义的^[15]:

$$R_{\text{CCF}}(\varepsilon) = \frac{\int_0^{T_p} S_{\text{BB-PreProc}}(t) S_{\text{Ref}}^*(t - \varepsilon) dt}{\sqrt{\int_0^{T_p} |S_{\text{BB-PreProc}}(t)|^2 dt \int_0^{T_p} |S_{\text{Ref}}(t)|^2 dt}}$$

式中: $S_{\text{BB-PreProc}}$ 为待评估的射频失真信号经下变频以及载波多普勒去除等预处理后得到的基带信号; 参考信号 S_{Ref} 为接收机本地产生的理想基带信号; T_p 通常取为参考信号的主码周期。

2.1 相关损耗

失真信号的相关损耗值等于在所分析带宽内失真信号和理想信号相关功率的差值, 其值可以用来表征失真造成的有用信号分量的功率损耗。

相关功率计算式为

$$P_{\text{CCF}} = 20 \lg(\max(|R_{\text{CCF}}(\varepsilon)|))$$

式中: $\max(|R_{\text{CCF}}(\varepsilon)|)$ 为复数互相关函数 $R_{\text{CCF}}(\varepsilon)$ 幅值的最大值。

2.2 S 曲线偏差

设 δ 为相关器的超前减滞后间距, 则非相干超前减滞后功率型鉴相器的 S 曲线计算式为^[15]

$$f_{\text{SCurve}}(\varepsilon, \delta) = \left| R_{\text{CCF}}\left(\varepsilon - \frac{\delta}{2}\right) \right|^2 - \left| R_{\text{CCF}}\left(\varepsilon + \frac{\delta}{2}\right) \right|^2$$

隐式方程 $f_{\text{SCurve}}(\varepsilon_{\text{bias}}(\delta), \delta) = 0$ 的解 $\varepsilon_{\text{bias}}(\delta)$ 定义为锁定点偏差。当 S 曲线存在多个过零点时, 应选取相关峰附近所对应的过零点; 而失真信号的 S 曲线偏差值定义为在所分析的 δ 取值范围内, 最大和最小锁定点偏差值的差值。

2.3 载波相位偏差

对于采用反正切型鉴相器的载波相位锁定环路, 载波相位偏差 $\Delta_{\phi}(\varepsilon)$ 的计算公式如下^[16]:

$$\Delta_{\phi}(\varepsilon) = \arctan\left(\frac{\text{Im}(R_{\text{CCF}}(\varepsilon))}{\text{Re}(R_{\text{CCF}}(\varepsilon))}\right)$$

式中: $\text{Im}()$ 和 $\text{Re}()$ 分别为求复数的虚部和实部的函数。

3 仿真结果分析

本节给出在不同 HPA 工作点和 HPA 前置滤波器带宽设置下, 第 1 节的 4 种复用调制信号的模拟失真在宽带导频跟踪、下边带导频跟踪 2 种不同跟踪方式下对测距性能影响的仿真结果。

所选取的 3 个测距性能指标即相关损耗、S 曲线偏差和载波相位偏差指标均与接收带宽有关。本文设定对于宽带导频跟踪方式,统一考察接收带宽为复用信号发射带宽的情况,即双边带宽 92 MHz,参考 AltBOC 信号的发射带宽^[2];对于下边带导频跟踪,统一考察下边带导频分量主瓣带宽的情况,即双边带宽 20.46 MHz。

3.1 理想基带信号性能评估

首先通过对理想无失真信号进行仿真评估来考察评估系统的正确性和精度。将所生成的图 1 所示的理想基带信号不经过虚框中所示的载荷模拟失真通道而直接送入后面的预处理、相关运算和性能参数评估环节。其中 TD-AltBOC 和 AltLOC 的理想基带信号可以基于文献[1,4]中的基带信号表达式生成,而 AltBOC 和 ACEBOC 的理想基带信号可以基于文献[5,17]中的查找表来生成。

仿真得到的理想无失真信号相关损耗均为 0,表 1 给出了 S 曲线偏差和载波相位偏差指标的测量结果。对比后文中有失真情况下的测量结果可见,评估系统引入的系统误差很小,足以满足评估要求。

表 1 理想无失真信号的评估指标测量结果

Table 1 Evaluation index measurement results of ideal signal without distortion

调制方式	宽带导频跟踪		下边带导频跟踪	
	载波相位偏差 / (°)	S 曲线偏差/m	载波相位偏差 / (°)	S 曲线偏差/m
AltBOC	0.413 7	0.053 4	0.633 1	0.027 6
TD-AltBOC	0.002 8	0.008 1	0.151 9	0.020 5
ACEBOC	0.436 0	0.014 8	0.141 1	0.013 7
AltLOC			0.487 6	0.004 7

3.2 HPA 工作点对不同调制方式的影响

本节仿真评估在固定带宽下,当 HPA 输出功率回退值在 -10 ~ 0 dB 区间变化时,各调制方式在不同跟踪方式下的性能。其中 HPA 前置滤波器和 OMUX 的双边带宽均设为发射带宽 92 MHz。对于 AltLOC 信号,只评估下边带导频跟踪方式的性能。图 2 和图 3 分别给出了宽带导频跟踪和下边带导频跟踪的仿真结果。

由图 2 可见,在整个 -10 ~ 0 dB 的工作点范围内,对于宽带导频跟踪而言 TD-AltBOC 信号的相关损耗值最大,高于 ACEBOC 约 0.19 ~ 0.25 dB,高于 AltBOC 约 0.32 ~ 0.41 dB;TD-AltBOC 的 S 曲

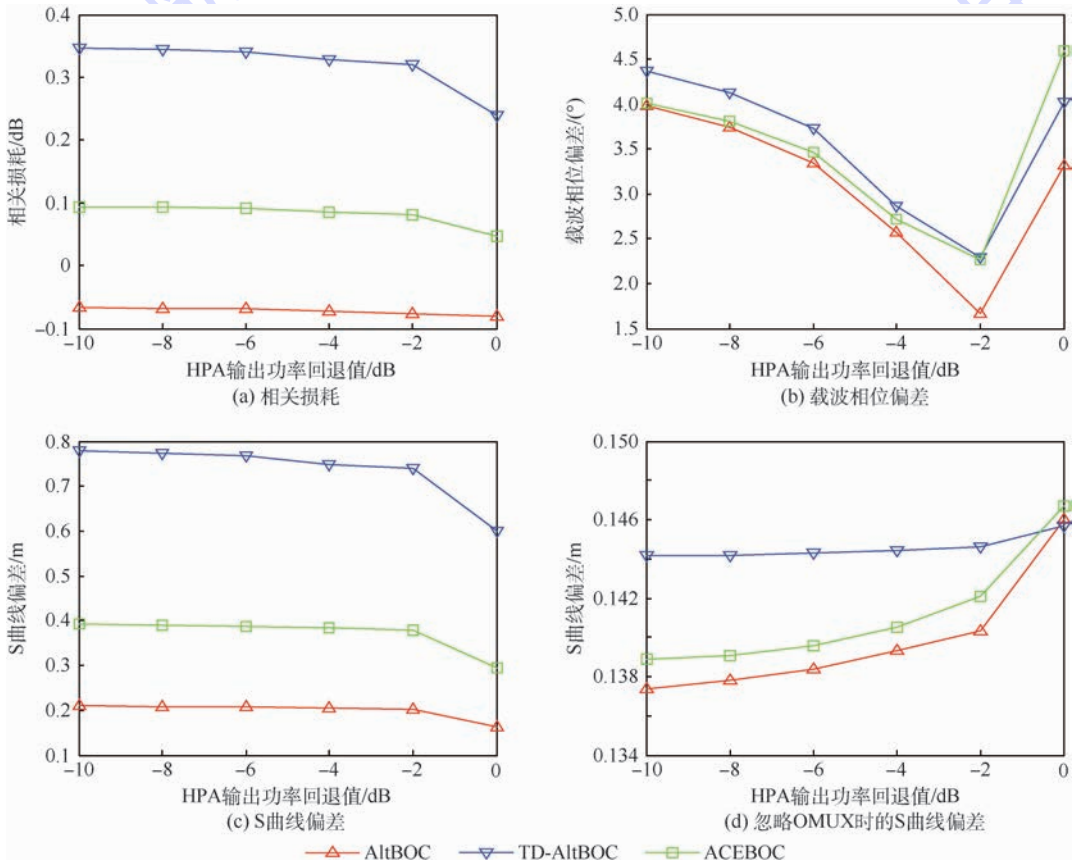


图 2 HPA 工作点对各调制方式宽带导频跟踪性能的影响

Fig. 2 Impact of HPA operating points on performance of wideband pilot tracking for different modulation schemes

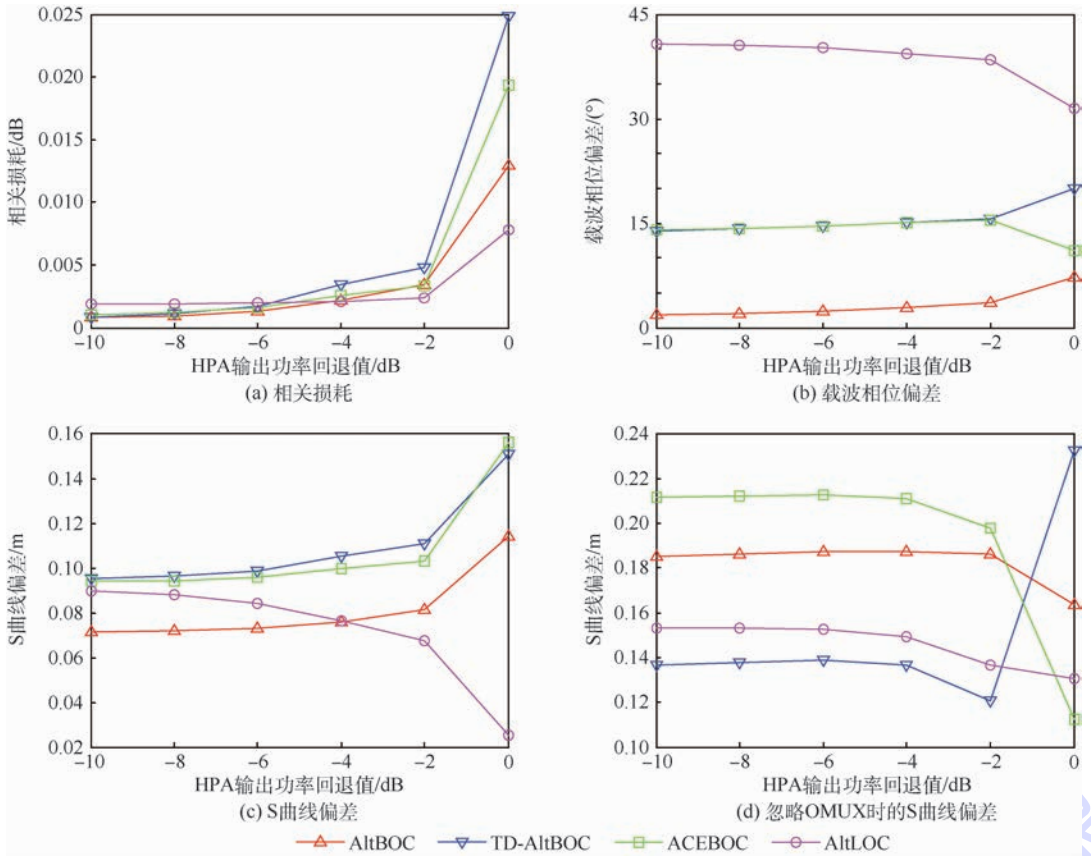


图3 HPA工作点对各调制方式下边带导频跟踪性能的影响

Fig.3 Impact of HPA operating points on performance of lower sideband pilot tracking for different modulation schemes

线偏差最大,大于 ACEBOC 约 0.30 ~ 0.38 m,大于 AltBOC 约 0.43 ~ 0.46 m; AltBOC 的载波相位偏差最小,当 HPA 输出功率回退约 2 dB 时,3 种调制方式均达到载波相位偏差最小值。

另外,如果不考虑 OMUX 的滤波效应,则 HPA 输出功率回退越多则信号失真引起的 S 曲线偏差越小。然而当综合考虑 HPA 前置滤波器、HPA 以及 OMUX 三者的联合影响时,随着 HPA 输出功率回退值增加,3 种调制方式的相关损耗和 S 曲线偏差反而单调增加。而对于载波相位偏差而言则存在一个最佳的功率回退值。

由图 3 可见,在单边带导频跟踪方式下,AltBOC、TD-AltBOC、ACEBOC 3 种信号的相关损耗和 S 曲线偏差较之宽带导频跟踪方式均显著减小,相关损耗基本可以忽略不计;而 TD-AltBOC 和 ACEBOC 的载波相位偏差则明显增加。随着 HPA 输出功率回退越多,4 种调制方式的相关损耗均趋近于 0,AltBOC、ACEBOC 和 TD-AltBOC 的 S 曲线偏差减小,而 AltLOC 的 S 曲线偏差反而增大。不同调制方式的载波相位偏差则表现出较大的差异,在整个 -10 ~ 0 dB 范围内 AltBOC 最小可降至 2°,AltLOC 最大可增至 41.5°,ACEBOC 和 TD-AltBOC 在 -10 ~ -2 dB 范围内维持在 15° 附近。

在越临近饱和点处功率回退对载波相位偏差的影响越大。另外,如果不考虑 OMUX 的滤波效应,AltBOC、ACEBOC 和 AltLOC 的 S 曲线偏差均在饱和点处达到最小值,而 TD-AltBOC 在临近饱和点时 S 曲线偏差显著增大。

3.3 HPA 前置滤波器带宽对不同调制方式的影响

本节仿真评估在 HPA 工作于饱和点的条件下,当 HPA 前置滤波器双边带宽在 50 ~ 110 MHz 范围内变化时,各调制方式的接收性能。其中 OMUX 的双边带宽仍固定为信号发射带宽 92 MHz。对于 AltLOC 信号,仍只分析下边带导频跟踪方式的性能。图 4 和图 5 分别给出了宽带导频跟踪和下边带导频跟踪的仿真结果。

由图 4 可知,在 HPA 前置滤波器带宽为 50 ~ 110 MHz 的范围内,对于宽带导频跟踪而言,TD-AltBOC 信号的相关损耗和 S 曲线偏差性能最差,ACEBOC 次之,AltBOC 最好。随着带宽的减小,TD-AltBOC 信号的相关损耗增加,另外 2 种信号相关损耗的变化则不明显。3 种调制的 S 曲线偏差均在最窄的 50 MHz 带宽上达到其最小值。而 3 种调制的载波相位偏差随带宽呈非单调的变化特性,需要根据特定条件下的仿真结果选择最佳

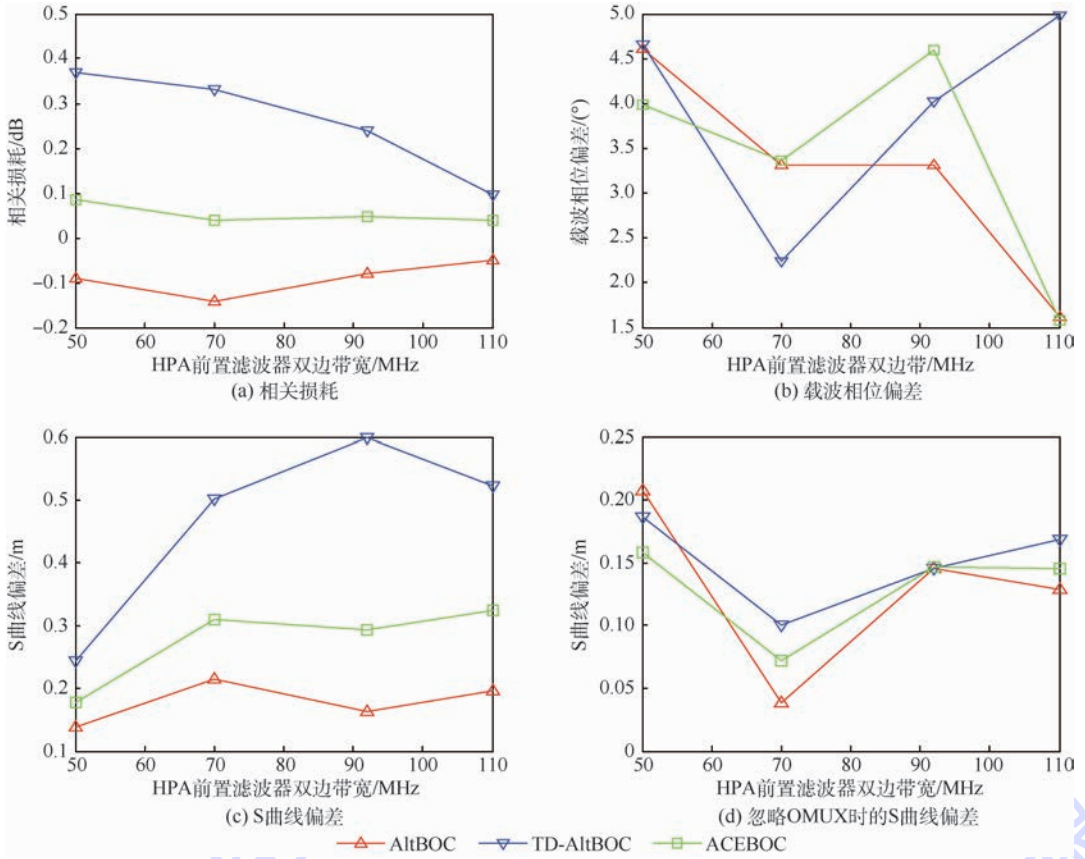


图 4 HPA 前置滤波器带宽对各调制方式宽带导频跟踪性能的影响

Fig. 4 Impact of pre-HPA filter bandwidth on performance of wideband pilot tracking for different modulation schemes

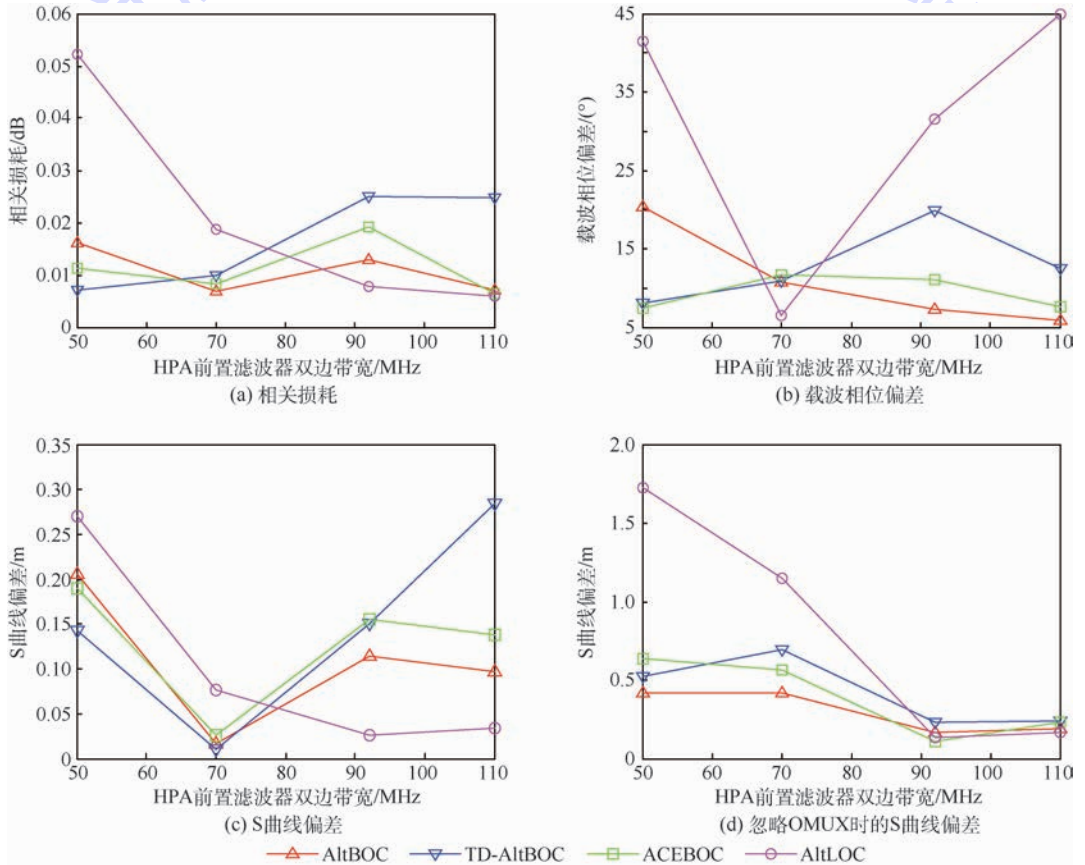


图 5 HPA 前置滤波器带宽对各调制方式下边带导频跟踪性能的影响

Fig. 5 Impact of pre-HPA filter bandwidth on performance of lower sideband pilot tracking for different modulation schemes

的 HPA 前置滤波器带宽。另外,如果不考虑 OMUX 的滤波效应,3 种调制的 S 曲线偏差均在 70 MHz 带宽上达到其最小值。

由图 5 可知,对于单边带导频跟踪而言,Alt-LOC 相关损耗随着带宽的增加而单调下降;而其他 3 种调制方式的相关损耗值均小于 0.025,基本可以忽略。而除了 AltBOC 信号的载波相位偏差随带宽增加而下降之外,其他信号的载波相位偏差以及 4 种调制的 S 曲线偏差均呈现随带宽非单调变化的特性,因而需要根据特定条件下的仿真结果和服务需求选择最佳的 HPA 前置滤波器带宽。另外,如果不考虑 OMUX 的滤波效应,4 种调制的 S 曲线偏差均在 90 MHz 带宽上达到其最小值。

4 结 论

1) 有效载荷信号生成链路的模拟失真对不同调制方式信号性能的影响有很大差异,而对同一种调制方式在不同接收处理方式下的影响也有很大差异;通过针对具体调制方式和接收处理方式、并且逼近真实信道特性的仿真来初步评估各种信道非理想特性对所关心的信号性能指标的影响,可以为信号调制方式的比较评估和载荷电路的优化设计提供参考。

2) 载荷 HPA 功率回退值、HPA 前置滤波器带宽等关键设计参数的特定变化对信号的相关损耗、S 曲线偏差和载波相位跟踪偏差不同指标的影响有好有坏,并不存在一组设计参数使得这 3 种性能同时达到最优,HPA 前置滤波器的带宽并非越宽越好,HPA 功率回退值也并非越大越好,载荷电路的优化设计需要根据逼近真实信道特性的仿真以至实际设备试验结果和服务需求进行折中考虑。

3) 对于 AltBOC、TD-AltBOC 和 ACEBOC 3 种调制方式,在宽带导频跟踪时载荷模拟失真引起的相关损耗和 S 曲线偏差 TD-AltBOC 最严重,ACEBOC 次之,AltBOC 最好;而在单边带导频跟踪时 3 种信号间的差异减小,其中相关损耗则可忽略不计。然而 3 种调制载波相位偏差的大小排序则随着载荷射频通道多个设计参数的变化而不同。

4) 载荷电路的优化设计应综合考虑其所有组成部件的联合影响,而不能仅考察其中部分部件的影响,否则可能得到南辕北辙的设计结果。

参考文献 (References)

- [1] LESTARQUIT L, ARTAUD G, ISSLER J L. AltBOC for dummies or everything you always wanted to know about AltBOC [C] // Proceedings of the 21st International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2008). Manassas: ION, 2008: 961-970.
- [2] REBEYROL E, MACABIAU C, JULIEN O, et al. Signal distortions at GNSS payload level [C] // Proceedings of the 19th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2006). Manassas: ION, 2006: 1595-1605.
- [3] RAPISARDA M, ANGELETTI P, CASINI E. A simulation framework for the assessment of navigation payload non-idealities [C] // 2nd Workshop on GNSS Signal and Signal Processing, 2007: 24-25.
- [4] TANG Z P, ZHOU H W, WEI J L, et al. TD-AltBOC: A new COMPASS B2 modulation [J]. Science China Physics, Mechanics & Astronomy, 2011, 54(6): 1014-1021.
- [5] YAO Z, ZHANG J, LU M. ACE-BOC: Dual-frequency constant envelope multiplexing for satellite navigation [J]. IEEE Transactions on Aerospace and Electronic Systems, 2016, 52(1): 466-485.
- [6] 唐祖平. GNSS 信号设计与评估若干理论研究 [D]. 武汉: 华中科技大学, 2009: 79-83.
- [7] TAN Z P. Research on relevant theory for GNSS signal design and evaluation [D]. Wuhan: Huazhong University of Science and Technology, 2009: 79-83 (in Chinese).
- [8] ZHU Y, YAO Z, CUI X, et al. Comparative analysis of dual-frequency constant envelope multiplexing techniques for ARNS band [C] // International Technical Meeting of the Institute of Navigation (ITM 2014). Manassas: ION, 2014: 792-801.
- [9] GAO Y, LI C, FU L, et al. Overall performance comparison of three dual-frequency constant envelope modulation schemes for GNSS [C] // Proceedings of 2016 China Satellite Navigation Conference (CSNC), 2016: 47-56.
- [10] 李星, 耿淑敏, 欧钢, 等. 非线性放大器对伪码测距误差的影响分析 [J]. 国防科技大学学报, 2008, 30(5): 49-55.
- [11] LI X, GENG S M, OU G, et al. A study on PN code ranging error caused by nonlinear power amplifier [J]. Journal of National University of Defense Technology, 2008, 30(5): 49-55 (in Chinese).
- [12] 黄旭方, 胡修林, 唐祖平. 星上高功率放大器对导航信号功率谱和伪码跟踪精度的影响 [J]. 电子学报, 2009, 37(3): 640-645.
- [13] HUANG X F, HU X L, TANG Z P. Impact of satellite's high power amplifiers on spectrum of navigation signals and PRN tracking accuracy [J]. Acta Electronica Sinica, 2009, 37(3): 640-645 (in Chinese).
- [14] 吴勇敢, 刘亚欢, 程成, 等. 高功率放大器和带宽对恒包络导航信号的影响分析 [J]. 遥测遥控, 2011, 32(3): 14-20.
- [15] WU Y G, LIU Y H, CHENG C, et al. Analysis of the impact of high power amplifier and bandwidth on the constant envelope navigation signals [J]. Journal of Telemetry, Tracking and Command, 2011, 32(3): 14-20 (in Chinese).
- [16] 周鸿伟, 魏蛟龙, 张小清, 等. 导航卫星有效载荷非理想特性研究 [J]. 华中科技大学学报 (自然科学版), 2014, 42(7): 118-123.

- ZHOU H W, WEI J L, ZHANG X Q, et al. Research on non-ideal property of payload core device on navigation satellite [J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2014, 42 (7): 118-123 (in Chinese).
- [13] CHEN Y, KOU Y, ZHANG Z. Analog distortion of wideband signal in satellite navigation payload [C] // Proceedings of 2012 China Satellite Navigation Conference (CSNC), 2012: 89-100.
- [14] DING L, ZHOU G T, MORGAN D R, et al. A robust digital baseband predistorter constructed using memory polynomials [J]. IEEE Transactions on Communications, 2004, 52 (1): 159-165.
- [15] SOELLNER M, KURZHALS C, HECHENBLAIKNER G, et al. GNSS offline signal quality assessment [C] // Proceedings of the 21st International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS 2008). Manassas: ION, 2008: 16-19.
- [16] 杨东旭. GNSS 空间信号模拟失真及其预失真方法研究 [D]. 北京: 北京航空航天大学, 2015: 39-40.
- YANG D X. Research on simulated distortion of GNSS spatial signal and its predistortion method [D]. Beijing: Beihang University, 2015: 39-40 (in Chinese).
- [17] European Union. Galileo OS SIS ICD (Issue 1.3) [EB/OL]. [2017-05-10]. http://www.gseuropa.eu/system/files/galileo_documents/Galileo_OS_SIS_ICD.pdf.

作者简介:

蔚小龙 男, 硕士研究生。主要研究方向: 卫星导航。

寇艳红 女, 博士, 副教授。主要研究方向: 卫星导航、无线通信和数字信号处理技术。

Simulation of analog distortion of dual-frequency multiplexing signal generated by navigation satellite

YU Xiaolong, KOU Yanhong*

(School of Electronic and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: The non-ideal components of the signal generator of a navigation satellite payload can distort the transmitted signal. On the one hand, no applicable analytical model about the ranging performance loss caused by such signal distortions has been established by far. On the other hand, the existing simulation studies about the analog distortions of wide-band dual-frequency multiplexing signals are not quite comprehensive. Taking the correlation loss, S-curve bias, and carrier phase bias as evaluation index, this paper simulates and evaluates the effects of the payload analog components including the high power amplifier (HPA), pre-HPA filter, and output multiplexer (OMUX) on the ranging performance of 4 different dual-frequency multiplexing modulations including AltBOC, TD-AltBOC, ACEBOC and AltLOC. In the simulation, we employ a memory behavioral model of HPA instead of a memoryless model, taking the effects of different pre-HPA filter bandwidth and HPA operating points into account. In addition, the performance loss of the distorted signal relative to the ideal signal is evaluated for both the double sideband and the single sideband pilot tracking methods. The study provides a useful reference for not only the comparison and evaluation of different signal modulation schemes but also the optimization design of the payload signal generator.

Keywords: constant envelope multiplexing; analog distortion; correlation loss; S-curve bias; memory high power amplifier (HPA)

Received: 2017-05-19; Accepted: 2017-06-09; Published online: 2017-06-12 10:55

URL: kns.cnki.net/kcms/detail/11.2625.V.20170612.1055.001.html

Foundation item: National Natural Science Foundation of China (61271197)

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

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

DOI: 10.13700/j.bh.1001-5965.2017.0305

基于 PCE 的谐波减速器动态精度不确定性分析

张金洋^{1,2}, 张建国^{1,2,*}, 彭文胜^{1,3}, 刘育强⁴, 汪龙⁵

(1. 北京航空航天大学 可靠性与系统工程学院, 北京 100083; 2. 北京航空航天大学 可靠性与环境工程重点实验室, 北京 100083; 3. 中国航空综合技术研究所, 北京 100028; 4. 中国空间技术研究院 总体部, 北京 100029; 5. 北京卫星制造厂, 北京 100094)

摘 要: 谐波减速器的动态精度不仅与其各个部件的制造公差和装配间隙有关, 还必须考虑谐波减速器柔性和摩擦的影响。目前谐波减速器精度问题研究大多只考虑单一因素, 在进行精度分析时没有考虑到模型参数的不确定性对精度的影响。本文研究了谐波减速器在静态因素(加工、装配)和动力学因素(柔性、摩擦)综合作用下的动态精度问题; 建立了考虑静态误差、柔性的非线性动力学模型; 利用多项式混沌展开(PCE)方法进行参数灵敏度分析和不确定性分析, 并和 Monte Carlo 方法作了比较, 结果表明 PCE 方法效率更高。并基于动态精度 PCE 进行可靠性分析, 得到动态精度可靠度。

关键词: 谐波减速器; 动态精度; 多项式混沌展开(PCE); 灵敏度分析; 可靠性分析
中图分类号: V442

文献标识码: A **文章编号:** 1001-5965(2018)05-1056-10

谐波减速器是一种建立在弹性变形理论上的新型机械传动方式, 在航空航天、机器人等精密定位领域广泛应用。相比于其他传统的传动方式, 谐波传动具有精度高、传动比大、传动平稳和传动效率高等优点。为了提高谐波减速器的传动精度, 国内外学者对谐波传动误差进行了深入的研究^[1-4], 主要从谐波减速器的加工、装配因素等传动机理方面考虑, 并没有考虑谐波减速器的柔性和非线性摩擦等动态特性对其传动误差的影响。另外, 在对谐波减速器进行静态误差分析时, 并没有考虑相关参数不确定性的影响以及静态误差和动态误差之间的耦合关系。Hsia^[5]提出了谐波传动误差主要是由波发生器带动柔轮运动时柔轮变形引起的, 从设计的角度研究柔轮柔性变形对动态误差的影响, 但未考虑其他非线性因素的影响。Tuttle 和 Seering^[6]考虑了谐波减速器柔性以及静

态误差等非线性因素, 同时深入研究了刚柔轮齿啮合的非线性机理, 建立了非线性动力学模型, 并通过试验验证了所建模型的正确性。游斌弟和赵阳^[7]研究了谐波减速器动态误差在考虑齿轮啮合摩擦和扭转刚度非线性因素时的动力学响应, 利用拉格朗日方程建立谐波减速器的误差动力学方程, 研究了不同频率谐波强迫激励下动态误差影响。Preissner 等^[8]建立了综合考虑柔轮非线性扭转、滞回特性和运动误差的谐波传动模型, 着重研究了柔轮的滞回特性对动态精度的影响。这些研究大多只考虑谐波减速器的柔轮动力学特性对动态精度的影响, 没有考虑动力学模型参数的不确定性。

本文从静态误差和动态误差产生的机理方面着手, 综合考虑了加工、装配误差、柔轮柔性以及刚柔轮摩擦综合作用下的谐波减速器的动态精度

收稿日期: 2017-05-12; 录用日期: 2017-06-16; 网络出版时间: 2017-07-11 11:46

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20170711.1146.003.html

基金项目: 国家“973”计划(2013CB733000)

* 通信作者: E-mail: zjg@buaa.edu.cn

引用格式: 张金洋, 张建国, 彭文胜, 等. 基于 PCE 的谐波减速器动态精度不确定性分析[J]. 北京航空航天大学学报, 2018, 44(5): 1056-1065. ZHANG J Y, ZHANG J G, PENG W S, et al. Dynamic accuracy uncertainty analysis of harmonic reducer based on PCE [J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 1056-1065 (in Chinese).

问题;在此基础上考虑相关参数的不确定性对动态精度的影响,利用多项式混沌展开(Polynomial Chaos Expansion, PCE)方法进行动态精度不确定性分析,构建动态精度可靠度模型计算动态精度可靠度。结果表明,利用 PCE 方法能够很好地处理谐波减速器非线性以及多变量的动态精度不确定性问题。

1 谐波减速器动态精度分析

谐波减速器主要由 3 部分组成:波发生器、柔轮和刚轮,其结构简图如图 1 所示。电机输出一定转速驱动波发生器旋转,波发生器带动柔轮发生柔性变形与刚轮啮合传递运动。其中波发生器是输入端,柔轮是输出端。

谐波减速器的动态误差定义为在一定转速条件下柔轮理想输出位置和实际输出位置之差 $\tilde{\theta}$ 为

$$\tilde{\theta} = \frac{\theta_m}{N} - \theta_1 \quad (1)$$

式中: θ_m 为波发生器输入角位置; θ_1 为柔轮输出角位置; N 为谐波传动比。

在一定转速下谐波减速器柔轮会发生柔性变形,当其做正反往复运动时,就会出现如图 2 所示的滞回误差。

谐波减速器动态误差由 2 部分组成^[9]:一部分是由谐波减速器基本部件加工、装配引起的

静态误差 $\tilde{\theta}_p$;另一部分是由柔轮的柔性和摩擦引起的波发生器与柔轮产生相对运动产生的滞回误差 $\tilde{\theta}_s$ 。因此动态误差也可以表示为

$$\tilde{\theta} = \tilde{\theta}_p + \tilde{\theta}_s \quad (2)$$

静态误差^[10]主要由谐波减速器各个部件的加工、装配误差所产生。其中刚轮加工误差产生的运动误差 Δ_{c1} 为

$$\Delta_{c1} = \frac{1}{2} \Delta F_{p2} \sin(2\omega_b t) + \frac{1}{2} \Delta f'_{r2} \sin(2Z_2 \omega_b t) \quad (3)$$

式中: ΔF_{p2} 为刚轮周节累积误差; $\Delta f'_{r2}$ 为刚轮切向相邻齿综合误差; ω_b 为波发生器角频率; t 为时间; Z_2 为刚轮齿数。

柔轮加工误差引起的运动误差 Δ_{c1} 为

$$\Delta_{c1} = \frac{1}{2} \Delta F_{p1} \sin\left(2\omega_b \frac{Z_2}{Z_1} t\right) + \frac{1}{2} \Delta f'_{r1} \sin(2Z_2 \omega_b t) \quad (4)$$

式中: ΔF_{p1} 为柔轮周节累积误差; $\Delta f'_{r1}$ 为柔轮切向相邻齿综合误差; Z_1 为柔轮齿数。

刚轮、柔轮装配误差,不考虑装配误差初相角的影响,由刚轮安装偏心误差 E_c 产生的运动误差为

$$\Delta_{c2} = E_c \sin(2\omega_b t) / \cos \alpha_n \quad (5)$$

式中: α_n 为啮合角。由柔轮安装偏心误差 E_r 产生的运动误差为

$$\Delta_{c2} = E_r \sin\left(2\omega_b \frac{Z_2}{Z_1} t\right) / \cos \alpha_n \quad (6)$$

由波发生器安装偏心误差 E_b 产生的运动误差 Δ_b 为

$$\Delta_b = E_b \sin(\omega_b t) / \cos \alpha_n \quad (7)$$

综上,可以得到总的静态误差为

$$\tilde{\theta}_p = (\Delta_{r1} + \Delta_{r2} + \Delta_{c1} + \Delta_{c2} + \Delta_b) \cdot \frac{0.9}{\sqrt{0.1(Z_1 + Z_2)}} \cdot \frac{412.8}{d} \quad (8)$$

式中: d 为刚轮分度圆直径。

动态精度包括机构的静态误差和柔轮柔性和机构摩擦引起的滞回误差,分析谐波减速器的动态精度,应该在静态误差模型基础之上建立机构的动力学模型。本文研究的谐波减速器物理简化模型如图 3 所示,图中: J_m 和 J_1 分别为输入和输出端转动惯量; B_m 、 B_1 和 B_{sp} 分别为谐波减速器输入端、输出端阻尼和刚柔啮合处阻尼。

将谐波减速器的柔轮等效为一个非线性扭转弹簧,柔轮扭转力矩特性用滞回误差角 $\tilde{\theta}_s$ 的三次多项式表示为

$$T_k = k_1 \tilde{\theta}_s + k_2 \tilde{\theta}_s^3 \quad (9)$$

式中: T_k 为柔轮的非线性扭矩; k_1 和 k_2 为谐波减

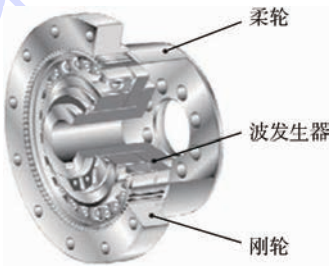


图 1 谐波减速器结构简图

Fig. 1 Structure diagram of harmonic reducer

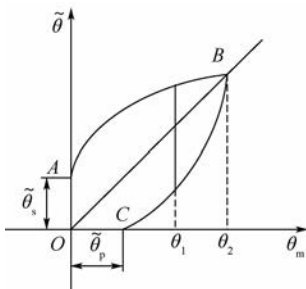


图 2 谐波减速器往复运动动态误差曲线

Fig. 2 Dynamic error curve of harmonic reducer reciprocation

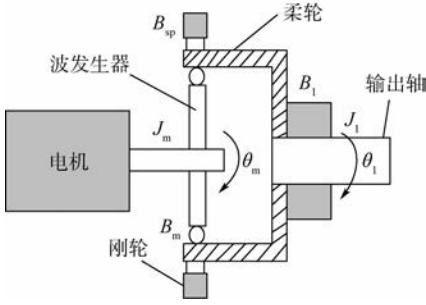


图3 谐波减速器物理简化模型

Fig. 3 A physical simplified model of harmonic reducer

速器柔轮等效扭转刚度。由图3所示的谐波减速器各部件之间的动力学关系,系统的动能 T 为

$$T = \frac{1}{2} J_m \dot{\theta}_m^2 + \frac{1}{2} J_1 \dot{\theta}_1^2 \quad (10)$$

系统的势能 V 为

$$V = \int_0^{\theta_1 - \theta_m / N - \bar{\theta}_p} T_k d\theta_s \quad (11)$$

系统的瑞利耗散函数为

$$D = \frac{1}{2} B_m \dot{\theta}_m^2 + \frac{1}{2} B_1 \dot{\theta}_1^2 + \frac{1}{2} B_{sp} \dot{\theta}_s^2 \quad (12)$$

因此,谐波减速器拉格朗日动力学方程^[11]为

$$J_m \ddot{\theta}_m + (k_1 + k_2 \bar{\theta}_s^2) \bar{\theta}_s \left(-\frac{1}{N} + \frac{d\bar{\theta}_p}{d\theta_m} \right) + B_m \dot{\theta}_m + B_{sp} \dot{\theta}_s \left(-\frac{1}{N} + \frac{d\bar{\theta}_p}{d\theta_m} \right) = T_m \quad (13)$$

$$J_1 \ddot{\theta}_1 + (k_1 + k_2 \bar{\theta}_s^2) \bar{\theta}_s + B_1 \dot{\theta}_1 + B_{sp} \dot{\theta}_s = 0 \quad (14)$$

$$\dot{\theta}_m = \omega_b \quad (15)$$

式中: T_m 为输入转矩。

根据式(1)和式(13)~式(15)可以得到谐波减速器动力学模型,动力学模型含有非线性静态误差项 $\bar{\theta}_p$ 和非线性扭转刚度 k_2 ,同时动力学模型含有参数较多,因此谐波减速器动力学模型具有高维非线性的特点,采用数值仿真方法求解。

2 基于PCE动态精度不确定性模型

2.1 PCE基本理论

PCE基本理论是用一个属于某个对应分析类型的正交多项式混沌之和(含有一个或多个随机变量)来近似地表示一个随机过程。

假设 $\{\xi_i\}_{i=1}^n$ 表示机构的随机参数,根据Wiener^[12]提出的高斯随机过程的均匀 Hermite 多项式展开模型,随机过程输出响应可以表示为

$$y = c_0 + \sum_{i_1=1}^n c_{i_1} \Pi_1(\xi_{i_1}) + \sum_{i_1=1}^n \sum_{i_2=1}^{i_1} c_{i_1 i_2} \Pi_2(\xi_{i_1}, \xi_{i_2}) + \sum_{i_1=1}^n \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} c_{i_1 i_2 i_3} \Pi_3(\xi_{i_1}, \xi_{i_2}, \xi_{i_3}) + \dots \quad (16)$$

$c = (c_0, c_1, \dots)$ 为待定系数矢量; $\xi = [\xi_1, \xi_2, \dots, \xi_n]$ 为服从标准正态分布的随机变量矢量; $\Pi_d(\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_d})$ 为 d 次多维 Hermite 多项式。将式(16)截断用 s 项来近似精度则可简化^[12]为

$$y = \sum_{j=0}^{s-1} c_j \Pi_j(\xi(\eta)) \quad (17)$$

式中: η 为随机事件; c_j 为待求解的确定性系数; $\Pi_j(\xi_1, \xi_2, \dots, \xi_n)$ 为广义 Wiener-Askey 多项式混沌。对于一个 n 维 Hermite 多项式,可表示为

$$H_n(\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_n}) = e^{\frac{1}{2}\xi^T \xi} (-1)^n \frac{\partial^n}{\partial \xi_{i_1} \partial \xi_{i_2} \dots \partial \xi_{i_n}} e^{-\frac{1}{2}\xi^T \xi} \quad (18)$$

Hermite 多项式的随机变量如果是标准正态分布,则满足

$$\langle \Pi_p(\xi) \Pi_q(\xi) \rangle = 0 \quad p \neq q \quad (19)$$

式中: $\langle \cdot \cdot \rangle$ 为希尔伯特空间上的内积,此处定义为

$$\langle \Pi_p(\xi) \Pi_q(\xi) \rangle = \int \Pi_p(\xi) \Pi_q(\xi) W(\xi) d\xi \quad (20)$$

式中: 多项式基对应的权重函数 $W(\xi)$ 为

$$W(\xi) = \frac{\exp\left(-\frac{1}{2}\xi^T \xi\right)}{\sqrt{(2\pi)^n}} \quad (21)$$

如果机构输入的随机变量的个数为 n , 又机构输出响应的多项式展开式最高阶次为 p , 则待定系数的个数 P 可以用式(22)求得:

$$P = \frac{(n+p)!}{n!p!} \quad (22)$$

2.2 动态精度不确定性模型

谐波减速器的动态精度不仅跟机构部件的制造公差、装配间隙等参数有关,而且还受机构动力学参数(如等效刚度、阻尼和转动惯量)的影响。上述参数在制造、测量过程中必然会存在不确定性而不是一个固定的数值,参数的不确定性通过机构的动力学模型影响动态精度响应的不确定性。

设谐波减速器的动力学模型用 M 表示,机构部件的制造公差、间隙、等效刚度和转动惯量等不确定参数可以用向量 ψ 表示,即

$$\psi = [\psi_1, \psi_2, \dots, \psi_n] \quad (23)$$

式中: ψ_i 为各个具体的不确定参数, n 为不确定参数总的个数。谐波减速器动态精度 $\bar{\theta}$ 在任一时刻 t 可以表示为向量 ψ 的确定性函数:

$$\bar{\theta}[t, \psi] = M(\psi_1, \psi_2, \dots, \psi_n) \quad (24)$$

设 $\psi_i \sim N(\mu_i, \sigma_i)$, 通过标准变换 $\xi_i = \frac{\psi_i - \mu_i}{\sigma_i}$,

$\xi = [\xi_1, \xi_2, \dots, \xi_n]$, 则动态精度用一阶、二阶、三阶 Hermite 多项式展开如式(25)~式(27)

所示^[13]:

$$\bar{\theta}[t, \xi] = c_{0,1} + \sum_{i=1}^n c_{i,1} \xi_i \quad (25)$$

$$\bar{\theta}[t, \xi] = c_{0,2} + \sum_{i=1}^n c_{i,2} \xi_i + \sum_{i=1}^n c_{ii,2} (\xi_i^2 - 1) + \sum_{i=1}^{n-1} \sum_{j>i}^n c_{ij,2} \xi_i \xi_j \quad (26)$$

$$\bar{\theta}[t, \xi] = c_{0,3} + \sum_{i=1}^n c_{i,3} \xi_i + \sum_{i=1}^n c_{ii,3} (\xi_i^2 - 1) + \sum_{i=1}^n c_{iii,3} (\xi_i^3 - 3\xi_i) + \sum_{i=1}^{n-1} \sum_{j>i}^n c_{ij,3} \xi_i \xi_j + \sum_{i=1}^n \sum_{j=1}^n c_{ijj,3} (\xi_i \xi_j^2 - \xi_i) + \sum_{i=1}^{n-2} \sum_{j>i}^{n-1} \sum_{k>j}^n c_{ijk,3} \xi_i \xi_j \xi_k \quad (27)$$

式中: $c_{i,m}$ 、 $c_{ij,m}$ 、 $c_{ijj,m}$ 和 $c_{ijk,m}$ 为展开式中的待定系数。

本文采用随机响应面配点法^[14]计算多项式混沌的展开系数。在随机向量展成的空间中,每一组样本 $\{\xi_1, \xi_2, \dots\}$ 都会对应一个点,这些点称为配点,对于 Wiener 的多项式混沌,其展开式基函数为 Hermite 多项式,若 Hermite 多项式的最高阶数为 p ,则相应的配点通常取 $p+1$ 阶 Hermite 多项式的根。未知系数可以通过式(28)和式(29)计算:

$$[c_0(t) \quad c_1(t) \quad \dots \quad c_{p-1}(t)]^T = (\Pi(\xi) \Pi^T(\xi))^{-1} \cdot \Pi(\xi)^T [\theta(t, \xi^0) \quad \theta(t, \xi^1) \quad \dots \quad \theta(t, \xi^k)]^T \quad (28)$$

$$\Pi(\xi) = \begin{pmatrix} \Pi_0(\xi^1) & \Pi_1(\xi^1) & \dots & \Pi_{p-1}(\xi^1) \\ \Pi_0(\xi^2) & \Pi_1(\xi^2) & \dots & \Pi_{p-1}(\xi^2) \\ \vdots & \vdots & & \vdots \\ \Pi_0(\xi^k) & \Pi_1(\xi^k) & \dots & \Pi_{p-1}(\xi^k) \end{pmatrix} \quad (29)$$

式中: $\xi^1, \xi^2, \dots, \xi^k$ 为采样点; k 为采样点数。

根据 Hermite 多项式正交性,动态精度的均值可以通过式(30)求得^[15]:

$$\begin{aligned} \bar{\theta}[t, \xi] &= E[\bar{\theta}[t, \xi]] = \int_{\mathbf{R}} \left(\sum_{j=0}^{p-1} c_j \Pi_j(\xi) \right) W(\xi) d\xi = \\ &= c_0 \int_{\mathbf{R}} \Pi_0(\xi) W(\xi) d\xi + \sum_{j=1}^{p-1} c_j \int_{\mathbf{R}} \Pi_j(\xi) W(\xi) d\xi = \\ &= c_0 + \sum_{j=1}^{p-1} c_j \int_{\mathbf{R}} \Pi_j(\xi) W(\xi) d\xi = c_0 \end{aligned} \quad (30)$$

由式(30)可以看出,动态精度均值是多项式混沌展开式的 0 阶项,同样的方法得到动态精度 $\bar{\theta}[t, \xi]$ 的均方差为

$$\sigma^2(\bar{\theta}[t, \xi]) = \text{var}(\bar{\theta}[t, \xi]) = E((\Pi(\xi) - \bar{\Pi}(\xi))^2) = \sum_{j=0}^{p-1} (c_j^2 \langle \Pi_j^2 \rangle) \quad (31)$$

3 不确定性及可靠性求解分析

3.1 谐波减速器 Dymola 仿真模型

根据谐波减速器动力学模型,本文针对 XB1-50 型号谐波减速器采用 Modelica 语言在 Dymola^[16] 编译环境下建立谐波减速器的动力学仿真模型(见图 4),主要包括直 PID 控制模块、电机模块和谐波减速器模块。此模型能够较好处理谐波减速器的机、电耦合问题,在求解非线性微分方程时精度较高。

采用如图 5 所示的谐波减速器精度测试平台对所建立的非线性谐波减速动力学仿真模型进行验证。采用恒温箱封闭,电机输入端和谐波减速器输出端分别连接有角速度和角度传感器,测定输入、输出端角位置以及角速度。实验条件下,电机输入转速为 100 r/min。

根据实测和设计参数数据确定谐波减速器动力学模型各参数如表 1 所示,静态误差模型参数如表 2 所示。

在 Dymola 仿真模型中通过 PID 调节器控制电机输出转速为 100 r/min,仿真时间为 10 s,待电机输入转速及误差波动曲线稳定后,通过采样输入角在 3 个周期均匀变化下动态误差值,实验条件下,通过角位移传感测得同样周期内谐波减速器输入端角位移 θ_m 以及输出端角位移 θ_l ,求得动态误差的测量值。将动态误差仿真值与测量值对比如表 3 所示。

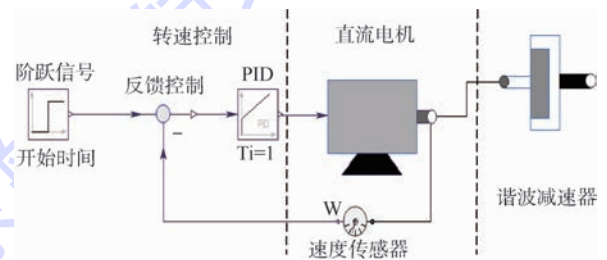


图 4 谐波减速器 Dymola 仿真模型
Fig. 4 Dymola simulation model of harmonic reducer

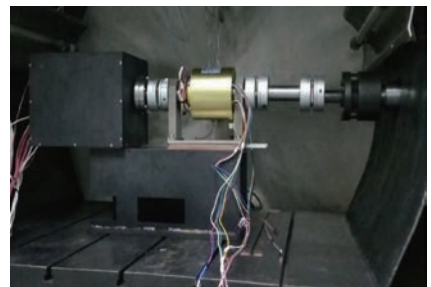


图 5 谐波减速器精度测试平台
Fig. 5 Precision test platform of harmonic reducer

表1 谐波减速器动力学模型参数

Table 1 Parameters of harmonic reducer dynamic model

参数	数值
$J_m / (\text{kg} \cdot \text{m}^2)$	3.2×10^{-4}
$J_1 / (\text{kg} \cdot \text{m}^2)$	8.5×10^{-4}
$B_m / (\text{N} \cdot \text{m} \cdot \text{s} \cdot \text{rad}^{-1})$	1.7×10^{-4}
$B_1 / (\text{N} \cdot \text{m} \cdot \text{s} \cdot \text{rad}^{-1})$	5.0×10^{-4}
$B_{sp} / (\text{N} \cdot \text{m} \cdot \text{s} \cdot \text{rad}^{-1})$	2.8×10^{-4}
$k_1 / (\text{N} \cdot \text{m} \cdot \text{rad}^{-1})$	7160
$k_2 / (\text{N} \cdot \text{m} \cdot \text{rad}^{-3})$	21576
N	90

表2 静态误差模型参数

Table 2 Parameters of static error model

参数	数值
Z_2	182
Z_1	180
$\alpha_n / (^\circ)$	20
E_i / m	3.1×10^{-5}
E_c / m	1.78×10^{-4}
E_b / m	2.05×10^{-4}
$\Delta F_{p2} / \text{m}$	3.6×10^{-5}
$\Delta f_{r2}^* / \text{m}$	1.14×10^{-4}
$\Delta F_{p1} / \text{m}$	3.6×10^{-5}
$\Delta f_{r1}^* / \text{m}$	1.14×10^{-4}

表3 不同电机输入角下动态误差实验值与仿真值

Table 3 Experimental values and simulation values of dynamic errors under different motor input corners

电机输入角/ $(^\circ)$	动态误差实验值/ $(^\circ)$	动态误差仿真值/ $(^\circ)$	真实误差/ $(^\circ)$	相对误差/%
50	0.023473	0.021908	0.001565	4.7
100	0.018379	0.017472	0.000907	4.9
200	0.016727	0.015486	0.001241	7.4
300	0.018503	0.017256	0.001247	6.7
400	0.016205	0.017369	0.001164	7.1
500	0.019181	0.018688	0.000493	2.5
700	0.017800	0.019320	0.002820	8.5
800	0.019794	0.017913	0.001881	9.5
1000	0.020950	0.022243	0.001293	4.1

通过表3中数据可以看出,实验得到的动态误差值与仿真值比较接近,相对误差在2.5%~9.5%范围内,考虑到相关参数不确定性,相对误差在接受范围之内,由此可见所建立的谐波减速器Dymola仿真模型能够很好地模拟考虑间隙和柔性非线性因素综合作用下的动态精度。

仿真得到谐波减速器输入、输出转速,动态误差和静态误差曲线如图6~图9所示,图6为谐波减速器输入转速曲线,转速经过短暂波动后稳定后在10.499 rad/s。由于谐波减速器柔性和控

制器惯性的作用,图7所示谐波减速器输出转速先出现大范围波动,然后稳定在0.1 rad/s附近波动。图8为动态误差随输入角 θ_m 波动曲线,开始波动程度较大,稳定后,呈周期性波动。由图9可以看出,综合考虑静态误差、柔性和摩擦作用的动态误差曲线比只考虑静态误差曲线波动幅值要大 0.005° ,相比增加25.4%,因此在对谐波减速器进行动态误差分析时,有必要考虑柔性和摩擦的影响。

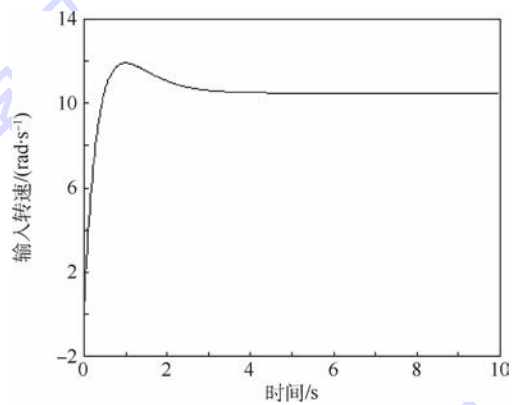


图6 谐波减速器输入转速

Fig. 6 Speed of harmonic reducer input

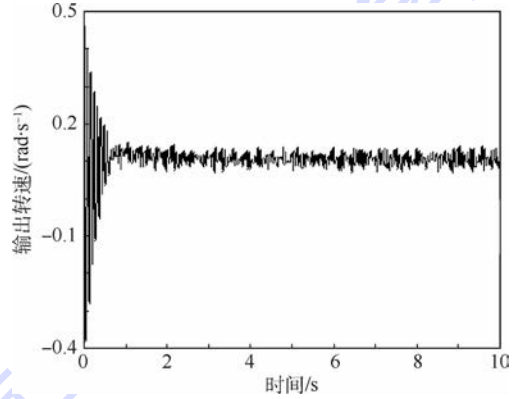


图7 谐波减速器输出转速

Fig. 7 Speed of harmonic reducer output

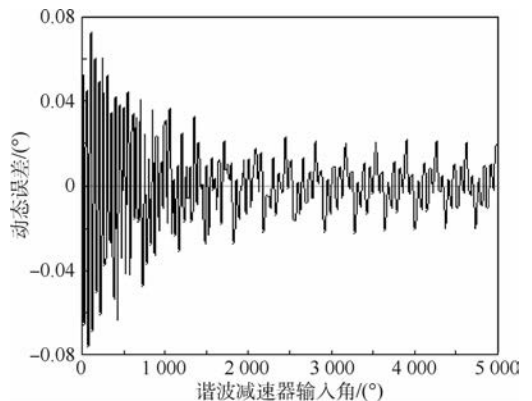


图8 谐波减速器动态误差

Fig. 8 Dynamic error of harmonic reducer

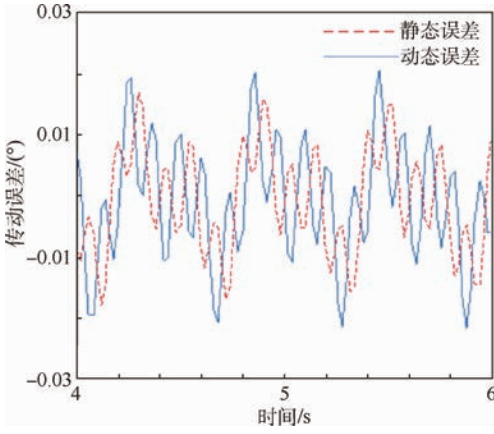


图 9 静态误差和动态误差曲线

Fig.9 Static error and dynamic error curves

3.2 动态精度灵敏度及不确定性分析

Sobol 敏感度分析方法^[17]是一种基于方差分解的 Monte Carlo 方法, Sobol 方法考虑了随机输入参量在整个取值范围内对输出响应的贡献以及随机参数的交互作用, Sobol 敏感度指标可以直接从 PCE 式的系数得到。

动态精度 PCE 式(26)和式(27)可以改写为

$$\tilde{\theta}[t, \xi] = \sum_{\alpha \in I_{i_1 i_2 \dots i_s}} c_\alpha \psi_\alpha(\xi_{i_1}, \xi_{i_2}, \dots, \xi_{i_s}) \quad (32)$$

式中: c_α 为 PCE 多项式系数; ψ_α 为 PCE 多项式的项; $I_{i_1 i_2 \dots i_s} = \{\alpha \in (\alpha_1, \alpha_2, \dots, \alpha_N) : \alpha_k = 0 \Leftrightarrow k \notin (i_1 i_2 \dots i_s), \forall k = 1, 2, \dots, N\}$, 根据多项式混沌的正交性, 可得

$$D_{i_1 i_2 \dots i_s} = \sum_{\alpha \in I_{i_1 i_2 \dots i_s}} c_\alpha^2 \quad (33)$$

$$D = \sum_{i=1}^N D_i + \sum_{1 \leq i < j \leq N} D_{ij} + \dots + D_{1,2,\dots,N} \quad (34)$$

由式(34)可以看出分解项表征不同随机变量及其相互作用对谐波减速器动态精度输出响应方差的贡献, 因此可以定义 Sobol 敏感度指标为

$$S_{i_1, i_2, \dots, i_s} = \frac{D_{i_1, i_2, \dots, i_s}}{D} \quad (35)$$

其满足

$$\sum_{i=1}^N S_i + \sum_{1 \leq i < j \leq N} S_{i,j} + \dots + S_{1,2,\dots,N} = 1 \quad (36)$$

式中: S_i 为主效应敏感度指标, 表征各个随机变量对动态精度响应方差的贡献。

考虑到当不确定参数较多时, PCE 式的待定系数呈指数增长, 为了简化分析, 对谐波减速器动力学模型参数进行灵敏度分析。考虑到输出轴转速相对输出轴转速小很多, 输出轴转动惯量相对输入轴转动惯量大很多, 故不考虑输出端阻尼和输入端转动惯量不确定参数的影响。静态误差模型中柔轮周节累积误差与柔轮安装偏心向量、刚

轮周节累积误差与刚轮安装偏心向量、刚柔轮切向相邻齿综合误差在静态误差模型中形式上相同, 为了简化不确定性分析, 故只考虑刚轮、柔轮、波发生器安装偏心向量和柔轮切向相邻齿综合误差不确定性参数的影响。取谐波减速器不确定随机向量 $\xi = [\xi_1, \xi_2, \dots, \xi_9] = [B_m, J_1, k_1, k_2, B_{sp}, E_f, E_c, E_b, \Delta f'_n]$, 对应的分布如表 4 所示。多项式混沌展开的响应为时刻 $t_0 = 5.7$ s 谐波减速器动态误差值, 取二阶混沌多项式为最高阶多项式基, $n = 9, p = 2$ 多项式混沌展开的待定系数为 55 个, 三阶 Hermite 多项式的根为 $t = [\tau_1, \tau_2, \tau_3] = [\sqrt{-1/3}, 0, \sqrt{1/3}]$, 则 PCE 各个随机变量的配点如表 5 所示。

在随机向量空间 $\xi = [\xi_1, \xi_2, \dots, \xi_9]$ 关于中心点对称进行 110 次抽样仿真, 求得动态精度 $\tilde{\theta}[t_0, \xi]$ 二阶多项式混沌展开一次项系数如表 6 所示。

根据展开式一次项系数可以得到各随机变量对动态精度 $\tilde{\theta}[t_0, \xi]$ 影响大小顺序依次为 $k_1 > J_1 > \Delta f'_n > E_c > B_m > E_f > k_2 > E_b > B_{sp}$

可以得出动态误差主要与谐波减速器的动力学参数有关。在仿真模型中, 分别将扭转刚度 k_1 , 输出轴转动惯量 J_1 , 柔轮切向相邻齿综合误

表 4 谐波减速器不确定性参数及分布

Table 4 Uncertainty parameters and distribution of harmonic reducer

参数	分布	均值	标准差
B_m	正态分布	$1.7 \times 10^{-4} \text{ N} \cdot \text{m} \cdot \text{s} \cdot \text{rad}^{-1}$	0.000 03
J_1	正态分布	$8.5 \times 10^{-4} \text{ kg} \cdot \text{m}^2$	0.000 05
k_1	正态分布	$7.16 \times 10^3 \text{ N} \cdot \text{m} \cdot \text{rad}^{-1}$	200
k_2	正态分布	$2.1576 \times 10^4 \text{ N} \cdot \text{m} \cdot \text{rad}^{-3}$	500
B_{sp}	正态分布	$2.8 \times 10^{-4} \text{ N} \cdot \text{m} \cdot \text{s} \cdot \text{rad}^{-1}$	0.000 04
E_f	正态分布	$3.1 \times 10^{-5} \text{ m}$	0.000 005
E_c	正态分布	$1.78 \times 10^{-4} \text{ m}$	0.000 04
E_b	正态分布	$2.05 \times 10^{-4} \text{ m}$	0.000 05
$\Delta f'_n$	正态分布	$1.14 \times 10^{-4} \text{ m}$	0.000 04

表 5 动态精度多项式混沌展开式配点

Table 5 Collocation of dynamic accuracy polynomial chaos expansion

参数	多项式混沌展开配点
B_m	(0.000 187, 0.000 17, 0.000 152)
J_1	(0.000 879, 0.000 85, 0.000 821)
k_1	(0.000 721 8, 0.000 716 0, 0.000 710 2)
k_2	(0.020 999, 0.021 576, 0.022 153)
B_{sp}	(0.000 303, 0.000 28, 0.000 257)
E_f	(0.000 033 8, 0.000 031, 0.000 028 1)
E_c	(0.000 201, 0.000 178, 0.000 154)
E_b	(0.000 176 3, 0.000 205, 0.000 233 7)
$\Delta f'_n$	(0.000 116 9, 0.000 14, 0.000 163 1)

差 $\Delta f'_n$ 的取值增加 20%，得到动态误差曲线如图 10~图 12 所示，对应的 $\bar{\theta}[t_0, \xi]$ 分别增加 48.3%，减少 41.7%，增加 22.5%，这与 PCE 灵敏度分析结果基本一致。在动力学模型参数中，对动态精度影响最大的是柔轮的等效扭转刚度，当扭转刚度 k_1 增大时，柔轮刚性增大，动态误差值增大。其次是输出轴转动惯量，输出轴转动惯量增加，动态误差波动幅值减小。静态误差模型中对动态误差影响较大的因素是刚柔轮的切向相邻齿综合误差。可见刚柔轮的齿加工精度对谐波减速器的动态精度影响较大，是整个谐波减速器设计中的关键。

为了减少参数不确定性分析的计算量，选取主要的不确定性随机参数向量 $\zeta = [J_1, \Delta f'_n, k_1, E_c, B_m]$ ，PCE 响应为动态精度在 $t = 5 \sim 6$ s 内等间距 50 个时间点的仿真值，每个时间点对称抽取 42 个样本点，图 13 给出了 PCE 方法和 1 000 次 Monte Carlo 方法所得到的谐波减速器动态误差的均值，可以看出这 2 种方法所求的结果十分吻合。图 14 给出了 2 种方法的动态误差的均方差，这 2 种方法求得的均方差之间的误差很小，PCE 方法精度接近于 Monte Carlo 方法的精度，而 PCE 方法仅通过 42 次仿真计算就得到了 Monte Carlo 方法 1 000 次仿真的计算结果，效率明显优于 Monte Carlo 方法。

表 6 动态精度 $\bar{\theta}[t_0, \xi]$ 多项式混沌展开式系数

Table 6 Polynomial chaotic expansion coefficient of dynamic accuracy $\bar{\theta}[t_0, \xi]$

项	系数	项	系数
1	0.0125	ξ_5	0.0009
ξ_1	0.0328	ξ_6	0.0140
ξ_2	-0.1540	ξ_7	0.0749
ξ_3	0.3576	ξ_8	0.0072
ξ_4	-0.0093	ξ_9	0.1023

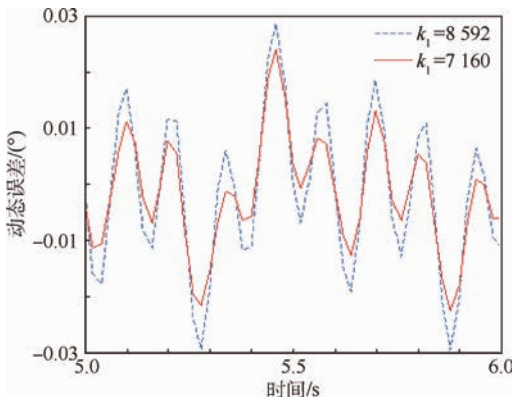


图 10 不同扭转刚度 k_1 时动态误差变化曲线
Fig. 10 Dynamic error curves at different torsional stiffness k_1

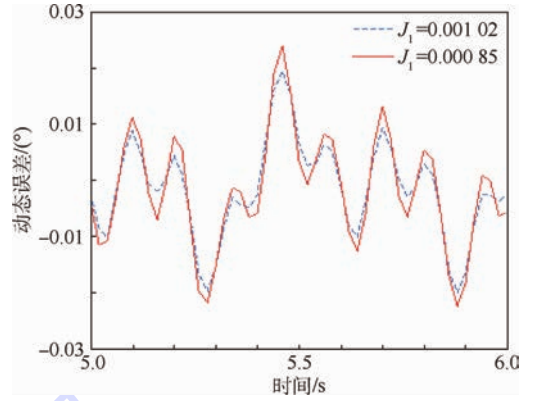


图 11 不同输出轴转动惯量 J_1 时动态误差变化曲线
Fig. 11 Dynamic error curves at different output shaft moment of inertia J_1

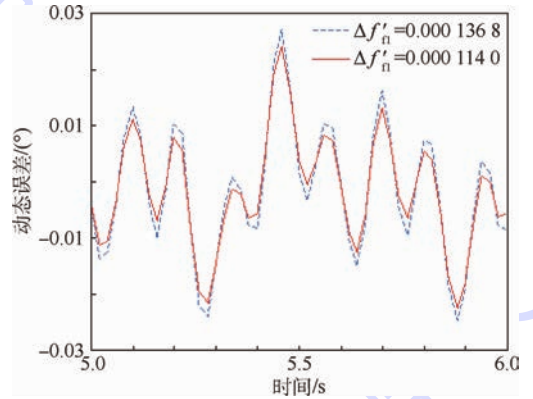


图 12 不同柔轮切向相邻齿综合误差 $\Delta f'_n$ 时动态误差变化曲线

Fig. 12 Dynamic error curves at different flexible wheel tangential adjacent gear comprehensive error $\Delta f'_n$

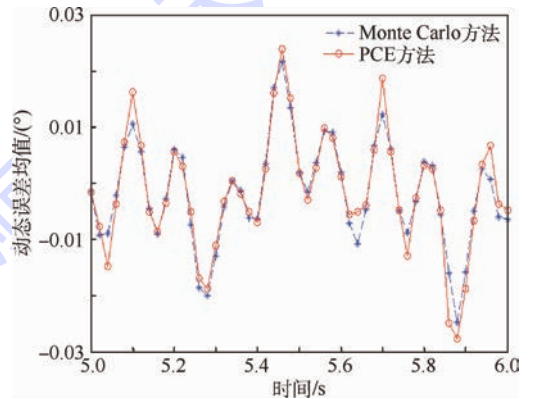


图 13 2 种方法动态误差均值比较
Fig. 13 Dynamic error mean comparison between two methods

3.3 动态精度可靠性分析

本文定义谐波减速器动态精度可靠度为谐波减速器在一定初始条件和时域内，动态误差绝对值的最大值不超过一定的阈值的概率，则动态精度可靠性功能函数可表示为

$$G(\zeta_1, \zeta_2, \dots, \zeta_5) = \bar{\theta}_{thr} - \bar{\theta}[t_m, \zeta] \quad (37)$$

式中: $\bar{\theta}_{thr}$ 为动态精度的阈值; $\bar{\theta}[t_m, \zeta]$ 为一定时域内动态误差绝对值的最大值; t_m 为最大值时间。不确定参数 $\zeta = [\zeta_1, \zeta_2, \dots, \zeta_5] = [J_1, \Delta f_{in}, k_1, E_c, B_m]$, 取二阶多项式混沌展开, 时域 $t_1 \sim [5, 6]$, 响应为 $\bar{\theta}[t_m, \zeta]$, 抽取 42 个样本点, 求得各项系数如表 7 所示。

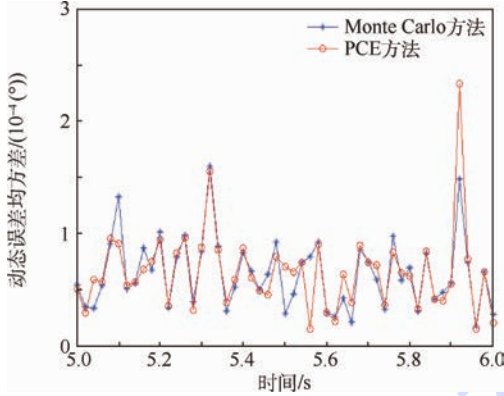


图 14 2 种方法动态误差均方差比较

Fig. 14 Dynamic error's mean square error comparison between two methods

表 7 动态误差 $\bar{\theta}[t_m, \zeta]$ 多项式混沌展开式系数
Table 7 Polynomial chaos expansion coefficients of dynamic error $\bar{\theta}[t_m, \zeta]$

项	系数	项	系数	项	系数
1	0.024 151	$\zeta_2 - 1$	-0.000 8	$\zeta_1 \zeta_5$	-0.005 3
ζ_1	-0.236 1	$\zeta_3 - 1$	0.049 0	$\zeta_2 \zeta_3$	-0.005 4
ζ_2	0.036 6	$\zeta_4 - 1$	-0.000 2	$\zeta_2 \zeta_4$	0.000 1
ζ_3	0.720 9	$\zeta_5 - 1$	-0.001 5	$\zeta_2 \zeta_5$	0.001 5
ζ_4	-0.023 6	$\zeta_1 \zeta_2$	0.003 7	$\zeta_3 \zeta_4$	0.003 9
ζ_5	-0.021 8	$\zeta_1 \zeta_3$	-0.036 5	$\zeta_3 \zeta_5$	0.004 2
$\zeta_1 - 1$	0.036 5	$\zeta_1 \zeta_4$	0.000 6	$\zeta_4 \zeta_5$	-0.000 5

设动态误差的阈值为 $\bar{\theta}_{thr} = 0.027^\circ$, 根据一次二阶矩 FORM 法^[18], 通过式 (38) 求得动态精度可靠度系数 β 为

$$\beta = \frac{G(\zeta_1^*, \zeta_2^*, \dots, \zeta_5^*) + \sum_{i=1}^5 \left[-\frac{\partial G}{\partial \zeta_i} \Big|_{\zeta^*} (\mu_{\zeta_i} - \zeta_i^*) \right]}{\left[\sum_{i=1}^5 \left(\frac{\partial G}{\partial \zeta_i} \Big|_{\zeta^*} \sigma_{\zeta_i} \right)^2 \right]^{\frac{1}{2}}} \quad (38)$$

式中: ζ_i^* 为设计验算点; $\zeta^* = [\zeta_1^*, \zeta_2^*, \dots, \zeta_5^*]$, 初始点取均值点。 $\mu_{\zeta_i}, \sigma_{\zeta_i}$ 分别为随机变量 ζ_i 的均值和标准差。 ζ_i^* 可通过式 (39) 求得:

$$\zeta_i^* = \mu_{\zeta_i} + \beta \sigma_{\zeta_i} \frac{-\frac{\partial G}{\partial \zeta_i} \Big|_{\zeta^*} \sigma_{\zeta_i}}{\left[\sum_{i=1}^5 \left(\frac{\partial G}{\partial \zeta_i} \Big|_{\zeta^*} \sigma_{\zeta_i} \right)^2 \right]^{\frac{1}{2}}} \quad (39)$$

通过迭代计算得到动态精度可靠度计算结果如表 8 所示。

在 Dymola 中进行 500 次 Monte Carlo 抽样仿

真实验, 得到动态误差仿真曲线如图 15 所示。

根据式 $P_f = n_f/n_i$ 求动态精度在一定时域内的失效概率, n_f 为动态误差绝对值最大值超过阈值的次数, n_i 为抽样仿真次数, 表 9 给出了 $n_i = 500 \sim 10\,000$ 次的模拟结果。

从表 9 中可以看出, 当 $n = 8\,000$ 时, P_f 收敛, Monte Carlo 仿真实验得到动态精度可靠度为 0.9619, 利用 PCE 方法和 FORM 法得到的可靠度 0.9584, 两者求得的结果相近, 而 PCE 方法仅用到 42 次抽样仿真, 效率明显优于 Monte Carlo 方法。

表 8 FORM 法可靠度计算结果

Table 8 Reliability calculation results based on FORM method

迭代次数	β	可靠度
1	2.783	0.997 3
2	1.782	0.962 6
3	1.737	0.958 8
4	1.732	0.958 4
5	1.732	0.958 4

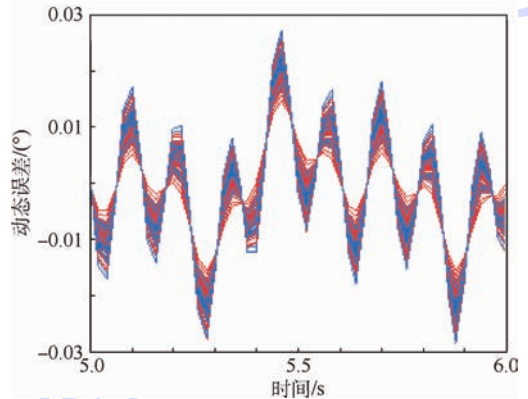


图 15 动态误差的 Monte Carlo 仿真曲线

Fig. 15 Monte Carlo simulation curves of dynamic error

表 9 Monte Carlo 仿真实验估算失效概率

Table 9 Failure probability estimation under Monte Carlo simulation experiment

n_i	n_f	P_f
500	14	0.028 0
2000	61	0.030 5
4000	126	0.031 5
6000	205	0.034 2
7000	254	0.036 3
8000	305	0.038 1
10000	381	0.038 1

4 结 论

本文考虑了静态误差、柔性和摩擦等因素综

合作用下的谐波减速器动态精度问题,建立了谐波减速器非线性动力学模型,利用多项式混沌展开(PCE)方法处理动态精度的随机不确定性。解决了以下问题:

1) 将静态误差与动态误差结合起来,建立了含有静态误差项的非线性动力学模型,通过实验验证了考虑动力学因素和静态误差综合作用下的动态精度更接近实际情况。

2) 基于PCE方法,通过灵敏度分析得到影响动态精度的主要参数是柔轮等效刚度、输出轴转动惯量和柔轮切向相邻齿综合误差。可以看出,柔轮的加工精度对动态精度影响较大。通过灵敏度分析选取影响较大的参数进行不确定性分析,可以减少不确定性分析的计算量。

3) 利用PCE方法得到动态精度的随机统计特性,并将PCE方法与传统的Monte Carlo方法比较,效率更高。利用PCE方法和FORM法计算一定时域内的动态精度可靠度,比传统的Monte Carlo方法计算可靠度效率更高。

参考文献 (References)

- [1] NYE T, KRAML R. Harmonic drive gear error: Characterization and compensation for precision pointing and tracking [C] // Processing of the 25th Aerospace Mechanics, Symposium. Washington, D. C. : NASA, 1991 : 237-252.
- [2] EMELYANOV A F. Calculation of the kinematic error of a harmonic gear transmission taking into account the compliance of elements [J]. Soviet Engineering Research, 1983, 3(7) : 7-10.
- [3] GRAVAGNO F, MUCINO V H, PENNESSTRI E. Influence of wave generator profile on pure kinematic error and centrodes of harmonic drive [J]. Mechanism and Machine Theory, 2016, 104 : 100-107.
- [4] 沙晓晨, 范元勋. 谐波减速器传动误差的研究 [J]. 机械制造, 2015, 44(5) : 50-54.
SHA X C, FAN Y X. Study of transmission error of harmonic drive reducer [J]. Machine Building Automation, 2015, 44(5) : 50-54 (in Chinese).
- [5] HSIA L M. The analysis and design of harmonic gear drives [C] // Processing of the 1988 IEEE International Conference on Systems, Man, and Cybernetics. Piscataway, NJ : IEEE Press, 1988 : 616-620.
- [6] TUTTLE T, SEERING W. Kinematic error, compliance, and friction in a harmonic drive gear transmission [C] // ASME Design Technical Conferences 19th Design Automation. New York : ASME, 1993 : 319-324.
- [7] 游斌弟, 赵阳. 考虑非线性因素的谐波齿轮传动动态误差研究 [J]. 宇航学报, 2010, 31(5) : 1297-1282.
YOU B D, ZHAO Y. Study on dynamic error of harmonic drive with nonlinear factors [J]. Journal of Astronautics, 2010, 31(5) : 1297-1282 (in Chinese).
- [8] PREISSNER C, ROYSTON T J, SHU D. A high-fidelity harmonic drive model [J]. Journal of Dynamic Systems, Measurement, and Control, 2012, 134(1) : 011002.
- [9] KIRCANSKI N, GOLDENBERG A, ANGELS J. Nonlinear modeling and parameter identification of harmonic drive gear transmissions [C] // Processing of the 32nd IEEE Conference on Robotics and Automatic. Piscataway, NJ : IEEE Press, 1995 : 3027-3032.
- [10] 王爱东. 机器人用谐波齿轮传动装置的运动精度分析 [D]. 北京 : 中国科学院, 2001 : 11-15.
WANG A D. Kinematic accuracy analysis of gear transmission for harmonic reducer of robot [D]. Beijing : Chinese Academy of Sciences, 2001 : 11-15 (in Chinese).
- [11] FATHI H, PRASANNA S, FRIEDHELM A. On the kinematic error in harmonic drive gears [J]. Journal of Mechanical Design, 2001, 123(1) : 90-97.
- [12] WIENER N. The homogeneous chaos [J]. American Journal of Mathematics, 1938, 60(4) : 897-936.
- [13] 赵珂, 高正红, 黄江涛, 等. 基于PCE方法的翼型不确定性分析及稳健设计 [J]. 力学学报, 2014, 46(1) : 11-19.
ZHAO K, GAO Z H, HUANG J T, et al. Uncertainty quantification and robust design of airfoil based on polynomial chaos technique [J]. Chinese Journal of Theoretical and Applied Mechanics, 2014, 46(1) : 11-19 (in Chinese).
- [14] LOEVETT T, PONCI F, MONTI A. A polynomial chaos approach to measurement uncertainty [J]. IEEE Transactions on Instrumentation and Measurement, 2006, 55(3) : 729-736.
- [15] BENJAMIN L, HOSAM K, JEFFREY L. Recursive maximum likelihood parameter estimation for state space systems using polynomial chaos theory [J]. Automatica, 2011, 47(11) : 2420-2424.
- [16] 陶海川, 来新民. 基于Dymola的无刷直流电机仿真模型 [J]. 计算机仿真, 2005, 22(5) : 63-65.
TAO H C, LAI X M. Computer simulation of brushless DC motor system based on Dymola [J]. Computer Simulation, 2005, 22(5) : 63-65 (in Chinese).
- [17] SURET B. Global sensitivity analysis using polynomial chaos expansion [J]. Reliability Engineering and System Safety, 2008, 93(7) : 964-979.
- [18] PENG W S, ZHANG J G, ZHU D T. ABCLS methods for high-reliability aerospace mechanism with truncated random uncertainties [J]. Chinese Journal of Aeronautics, 2015, 28(4) : 1066-1075.

作者简介:

张金洋 男, 硕士研究生。主要研究方向: 机械产品可靠性。

张建国 男, 博士, 教授, 博士生导师。主要研究方向: 机械产品可靠性、机电产品可靠性。

彭文胜 男, 博士研究生。主要研究方向: 机械设计、机电产品可靠性设计。

Dynamic accuracy uncertainty analysis of harmonic reducer based on PCE

ZHANG Jinyang^{1,2}, ZHANG Jianguo^{1,2,*}, PENG Wensheng^{1,3}, LIU Yuqiang⁴, WANG Long⁵

(1. School of Reliability and Systems Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China; 2. Science and Technology on Reliability and Engineering Laboratory, Beijing University of Aeronautics and Astronautics, Beijing 100083, China; 3. AVIC China Aero-poly Technology Establishment, Beijing 100028, China; 4. Institute of Spacecraft System Engineering CAST, Beijing 100029, China; 5. Beijing Satellite Manufacturer, Beijing 100094, China)

Abstract: The dynamic accuracy of the harmonic reducer is related to the parts' tolerance and assembly clearance as well as the flexibility and friction of the harmonic reducer. Most of the published literatures consider only a single factor and do not take into account the influence of model parameter uncertainty. In this paper, the dynamic accuracy is researched considering static factors (machining and assembly) and dynamic characteristics (flexibility and friction). The nonlinear dynamic model is established, which contains static error and flexibility term. Polynomial chaos expansion (PCE) is used to handle the parameter sensitivity and uncertainty. By comparison, PCE is more efficient than Monte Carlo. Dynamic accuracy reliability is finally obtained through reliability analysis based on dynamic accuracy PCE.

Keywords: harmonic reducer; dynamic accuracy; polynomial chaos expansion (PCE); sensitivity analysis; reliability analysis

Received: 2017-05-12; **Accepted:** 2017-06-16; **Published online:** 2017-07-11 11:46

URL: kns.cnki.net/kcms/detail/11.2625.V.20170711.1146.003.html

Foundation item: National Basic Research Program of China (2013CB733000)

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

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

DOI: 10.13700/j.bh.1001-5965.2017.0332

硒掺杂锗碲相变存储材料的第一性原理研究

方治乾¹, 缪奶华^{1,2}, 周健^{1,2,*}

(1. 北京航空航天大学 材料科学与工程学院, 北京 100083;

2. 北京航空航天大学 国际交叉科学研究院 集成计算材料工程中心, 北京 100083)



摘 要: 硒(Se)掺杂可以大幅提高锗碲(GeTe)相变存储材料的再结晶温度,使其具有更高的服役温度和更好的数据保持力,然而Se掺杂对GeTe微观结构和电学性质的影响机制尚不清楚。采用第一性原理计算方法,对Se掺杂GeTe相变存储材料的几何构型、成键性质和电子性质进行了理论研究。结果表明,对于GeTe完美晶体,掺杂的Se原子优先取代Te原子。而对含本征Ge空位的GeTe体系,Se倾向于取代与Ge空位最近邻的Te原子。Se原子与Ge空位具有吸引作用,抑制了Ge空位的移动,从而提高其再结晶温度。Se掺杂导致含Ge空位的菱方相体积收缩,带隙减小,而使含Ge空位的面心立方相体积膨胀,带隙增大。Se掺杂减小了GeTe两晶相的体积差异。计算结果为解释实验中Se掺杂导致的奇特相变性质提供了重要线索。

关键词: Se掺杂GeTe; Ge空位; 相变存储材料; 第一性原理计算; 元素掺杂

中图分类号: TB34; O474

文献标识码: A **文章编号:** 1001-5965(2018)05-1066-08

硫系化合物半导体长期以来被用作光学信息存储介质,如DVD-RW和DVD-RAM等^[1]。近年来,其也被广泛研究用于最有前景的下一代非易失性电子信息存储器,相变存储器(Phase Change Memories, PCM)^[2-3]。相变存储器利用相变存储材料在低电导、低反射率的非晶态与高电导、高反射率的晶态之间快速相变的特性来实现信息存储。常用的相变存储材料主要集中于GeTe-Sb₂Te₃伪二相线上,包括位于两端的母体材料GeTe、Sb₂Te₃以及三元的Ge₁Sb₂Te₄、Ge₁Sb₄Te₇和Ge₂Te₂Sb₅等^[4-6]。其中,GeTe是一种性能优异的相变存储材料^[7-8]。相较于Ge₂Sb₂Te₃(GST),GeTe具有较高的再结晶温度(189℃)和较高的熔点(700℃)^[9],其数据能在105℃左右长期保持,可用于较高温度的信息存储。然而,随

着现代科技的进一步发展,汽车电子和航空航天领域对信息存储提出了更高的挑战。汽车电子领域要求存储器工作温度在120℃左右,而航空航天领域则要求其能在150℃下工作10年。GeTe材料已无法满足这些行业的需求。亟需对GeTe相变存储材料进行性能改进,以满足信息存储领域日益提升的高温应用要求。

元素掺杂是一种有效的半导体改性方式。科研人员对此进行了研究,实验表明Se掺杂GeTe相变存储材料具有更优异的性质。Vinod等制备了不同Se掺杂量的(GeTe)_{1-x}Se_x薄膜,指出Se掺杂能大幅提高GeTe的再结晶温度,增强其数据保持力^[10];Se掺杂同时能提高GeTe非晶的光学带隙,10%的Se掺杂可大幅提高GeTe的非晶电阻,使晶相与非晶相的电阻差异增加2个数量

收稿日期: 2017-05-18; 录用日期: 2017-06-12; 网络出版时间: 2017-07-14 17:27

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20170714.1727.001.html

基金项目: 国家自然科学基金(61274005)

*通信作者. E-mail: jzhou@buaa.edu.cn

引用格式: 方治乾, 缪奶华, 周健. 硒掺杂锗碲相变存储材料的第一性原理研究[J]. 北京航空航天大学学报, 2018, 44(5): 1066-1073. FANG Z Q, MIAO N H, ZHOU J. First-principles study of Se doped GeTe phase-change material[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 1066-1073 (in Chinese).

级^[11]。由于 GeTe 中可以存在高达 8% ~ 10% 的本征 Ge 空位^[12-13], Vinod 和 Sangunni 推测掺入的 Se 原子可能占据 Ge 空位的位置,从而抑制了 GeTe 的结晶^[11]。然而,这些实验研究主要针对 Se 掺杂后的物相组成、电学性质、光学性质和热动态性质等。Se 掺杂对于体系的原子构型、成键情况等微观性质的影响以及性能增强的作用机制尚不明晰。这些微观性质需要通过第一性原理计算模拟来进一步理解,从而更好地为未来实验提供理论指导。此外,与 GeTe 同系列的 $\text{Ge}_2\text{Te}_2\text{Sb}_5$ 的 Se 掺杂实验表明,当 $\text{GST}_{1-x}\text{Se}_x$ 中 Se 含量 x 超过 10% 时,出现了奇特的相转变性质:掺杂后的 GST 材料在受热时直接由无定型态结晶为稳态六方,而跳过中间的亚稳面心立方相(face-centered cubic, fcc)^[14]。而 Wang 等进行了 Sb_2Se 掺杂 $\text{Ge}_2\text{Sb}_2\text{Te}_5$ 的研究,指出当 Se 原子含量超过 9% 时,亚稳面心立方相转变为稳定六方相的过程被抑制^[15]。Se 掺杂 PCM 对相转变过程的影响较大。对硫化化合物相变存储材料的 Se 掺杂进行计算研究,也可以为解释相变存储材料晶相演变机制提供线索。

本文采用第一性原理计算方法,对 Se 掺杂完美 GeTe 和含本征 Ge 空位的 GeTe 的几何构型、成键情况和电子性质进行了研究。结果表明,对于完美 GeTe,Se 优先取代 Te 原子。而对于存在本征 Ge 空位的 GeTe 晶体,Se 掺杂原子不会占据 Ge 空位,而倾向于取代 Ge 空位最近邻 Te 原子。Se 原子与 Ge 空位的相互吸引,形成了较强的 Ge—Se 键,抑制了 GeTe 相变存储材料的结晶,从而提高了其数据保持能力和服役温度。Se 原子与本征 Ge 空位最近邻时,使两相的电子局域函数(Electron Localization Function, ELF)分布和体积更加接近,可能是导致奇特相转变性质的重要原因。

1 计算方法

本文计算基于维也纳大学开发的从头算模拟软件包(Vienna Ab-initio Simulation Package, VASP)^[16]。体系采用了 $\text{Ge}(4s^24p^2)$ 、 $\text{Te}(5s^25p^4)$ 的标准 PAW 赝势。与先前的 GeTe 计算一致,本文使用了广义梯度近似 GGA-PBE^[17-19] 来描述交换关联泛函。截断能设置为 350 eV, k 点设置为 $3 \times 3 \times 3$ 的 Gamma 中心 k 点网格。截断能和 k 点的设置都进行了收敛性测试。自洽计算的收敛判据设置为电子步能量差小于 10^{-5} eV, 而结构弛豫收敛判据则为每个原子上的残余应力均小于 0.1 eV/nm。

GeTe 包含 2 个晶相,分别为稳态菱方相(rhombohedral)和高温亚稳面心立方相^[20]。信息存储主要通过 GeTe 非晶态和稳态菱方相之间的快速转变来实现^[21]。因此,本文主要研究 Se 掺杂对于菱方相的影响,并与 Se 掺杂的面心立方相进行了类比分析。GeTe 菱方相是一种扭曲的氯化钠结构,可以看做由 GeTe 面心立方相中 Ge 原子亚结构和 Te 原子亚结构沿 $[111]$ 方向错动一定距离形成的^[22]。这种结构变形属于派尔斯变形(Peierls distortion)^[23]。GeTe 菱方相中形成了交替的 3 个 Ge—Te 长键和 3 个 Ge—Te 短键。从 $[111]$ 方向观察,GeTe 也可以被当做层状结构;其中,由 Ge—Te 短键连接的原子可视为处在同一层。基于 GeTe 菱方结构的惯用胞,构建了 $2 \times 2 \times 2$ 的超胞,包含 64 个原子,结构如图 1(a) 所示。对 GeTe 面心立方相,同样采用了 $2 \times 2 \times 2$ 的超胞,如图 1(b) 所示。所有构型均进行了结构弛豫。计算优化得到的单胞晶格数据如表 1 所示。对于无缺陷的菱方相和面心立方相,计算得到晶格常数和晶胞角度分别为 $a_0 = 0.6086 \text{ nm}$, $\alpha = 88.14^\circ$ 和 $a_0 = 0.6019 \text{ nm}$, $\alpha = 90^\circ$ 与实验测量

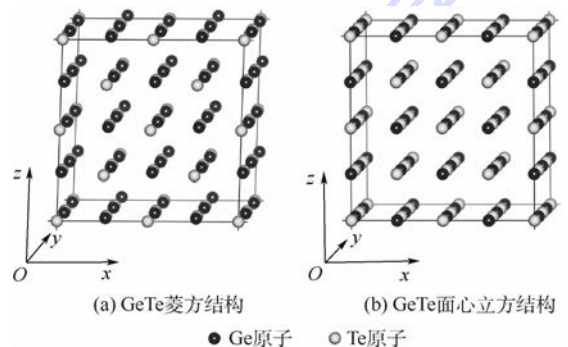


图1 GeTe 晶体结构

Fig. 1 Crystalline structure of GeTe

表1 计算得到的单胞晶格数据

结构	a_0/nm	$\alpha/(\circ)$	V/nm^3	E_g/eV
菱方相 GeTe	0.6086	88.14	0.22505	0.62
菱方相 $\text{Ge}_{32}\text{Te}_{31}\text{Se}_1$	0.6083	87.94	0.22462	0.58
菱方相 $\text{Ge}_{31}\text{Te}_{31}\text{Se}_1$	0.6058	87.77	0.22188	0.58
菱方相 $\text{Ge}_{31}\text{Te}_{32}$	0.6062	87.82	0.22234	0.59
菱方相 $\text{Ge}_{31}\text{Se}_1\text{Te}_{32}$	0.6084	88.12	0.22486	—
面心立方相 GeTe	0.6019	90	0.21808	0.38
面心立方相 $\text{Ge}_{32}\text{Te}_{31}\text{Se}_1$	0.6003	90	0.21634	0.40
面心立方相 $\text{Ge}_{31}\text{Te}_{31}\text{Se}_1$	0.6024	90	0.21852	0.52
面心立方相 $\text{Ge}_{31}\text{Te}_{32}$	0.5994	90	0.21537	0.31
面心立方相 $\text{Ge}_{31}\text{Se}_1\text{Te}_{32}$	0.6010	90	0.21704	—

注:晶格常数 a_0 、晶胞角度 α 、惯用胞体积 V 以及带隙 E_g 。

值及前人的计算值吻合^[20,24]。由于 GeTe 中存在大量的本征 Ge 空位,在掺杂分析中必须考虑 Ge 空位的影响。对于存在本征 Ge 空位的 GeTe 体系的 Se 掺杂计算,分别构造了 Ge₃₁Te₃₂(共 63 个原子)的菱方结构和面心立方结构,并对所有构型均进行了结构优化。

掺杂形成能的计算式^[25-26]为

$$\Delta E_{\text{form}} = E_{\text{final}} - E_{\text{initio}} - \sum_{i=1}^{n\text{-type}} (N_i \mu_i) \quad (1)$$

式中: E_{final} 和 E_{initio} 分别为末态体系和初始体系的总能量; μ_i 为第 i 种原子的化学势,通常是其最稳定单质中单个原子的能量; N_i 为缺陷形成后与形成前第 i 种原子的数量差。

本文考虑了缺陷生长环境对于掺杂形成能的影响。当 Se 取代 Te 时,在富 Te 情况下,被取代的 Te 原子视为与体相 Te 单质中的一个 Te 原子结合;而富 Ge 情况下,被取代的 Te 原子则视为与体相 Ge 单质中的一个 Ge 原子结合,重新形成新的 Ge—Te 单元。故而在富 Te 情况下,Te 原子化学势即为 Te 体相单质中单个原子的能量,而 Ge 原子化学势则为 $\mu_{\text{Ge}} = \mu_{\text{GeTe}} - \mu_{\text{Te}}$ 。其中 μ_{GeTe} 为 GeTe 晶相中 Ge、Te 原子对的能量, μ_{Te} 为 Te 原子化学势。反之,在富 Ge 情况下,Ge 原子化学势为 Ge 体相单质中单个原子能量,而 Te 原子化学势为 $\mu_{\text{Te}} = \mu_{\text{GeTe}} - \mu_{\text{Ge}}$ 。

2 结果与讨论

2.1 GeTe 完美晶体的 Se 掺杂

本节研究了 Se 掺杂 GeTe 完美晶体。当掺杂 Se 原子时,可能的占位情况有:取代 Ge 原子、取代 Te 原子或占据间隙位置。由于 Se 间隙原子造成较大的晶格畸变,具有较高形成能,故不作比较。计算得到不同取代位置的形成能 E_f 如表 2 所示。

从表 2 可以看出,对于 GeTe 的菱方相和面心立方相,无论在富 Ge 或富 Te 的情况下,掺杂的 Se 原子均优先取代 Te 原子。Se 掺杂的形成能为负,表明 Se 原子容易掺杂到 GeTe 中。从表 1 中还可以看出,少量 Se 掺杂对于 GeTe 完美晶体菱方

结构和面心立方结构的体积影响很小,造成体积的微量坍塌。主要由于 Se 的原子半径小于 Te。此外,较低的形成能和较小的晶格变化,表明在 GeTe 中能够实现高浓度的 Se 掺杂。这与实验观测一致,实验中 Se 掺杂量达到 20% 时,依然没有造成 GeTe 结构的破坏^[11]。

对 GeTe 完美晶体的 Se 掺杂进行了电子局域函数分析,如图 2 所示。由于 Se、Te 属于同主族元素,Se 原子与 Te 原子具有十分相似的电子局域函数分布。对于 GeTe 菱方相,由于派尔斯形变,形成的 Ge—Te 短键键长 $R_{\text{Ge1-Te}} = 0.286 \text{ nm}$, Ge—Te 长键键长 $R_{\text{Ge2-Te}} = 0.325 \text{ nm}$,与实验值十分接近^[20]。当 Se 取代 Te 原子时,也呈现出长短键交替的情况,然而,其中 Ge—Se 短键变得更短 ($R_{\text{Ge1-Se}} = 0.266 \text{ nm}$),而 Ge—Se 长键变得更长 ($R_{\text{Ge2-Se}} = 0.334 \text{ nm}$)。这是由于 Se 原子具有较强的电负性,与其同层最近邻 Ge 原子(Ge1)形成了较强的 Ge—Se 键。同时 Se 原子较小的原子半径,使得其与邻层最近邻 Ge 原子(Ge2)的键合变弱。Se 掺杂使得体系中原子分层排布的现象更加明显。

此外,GeTe 面心立方结构则具有 6 个等同的 Ge—Te 键, $R_{\text{Ge-Te}} = 0.301 \text{ nm}$,如图 2(c) 所示。Se 取代 Te 原子之后,Se 原子的电子局域函数依然呈球状分布,形成 6 个等同的 Ge—Se 键, $R_{\text{Ge-Se}} = 0.292 \text{ nm}$,键长较 Ge—Te 键变短。两相中,Se 掺杂均形成了较强的 Ge—Se 键,对 Ge 原子的移动起到抑制作用。Ge—Se 原子主要为共价键结合,

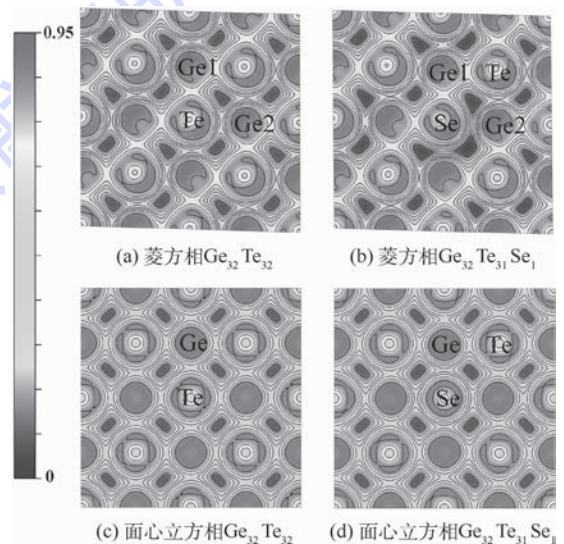


图 2 完美 GeTe 晶体掺杂前后在 (100) 面上的 ELF 截面图(等高线间隔为 0.14)

Fig. 2 ELF contour plots on (100) plane for Se doped ideal GeTe crystal (interval is 0.14)

表 2 Se 掺杂完美 GeTe 不同取代位置的形成能

Table 2 Formation energy of Se doping at different substitution positions in ideal GeTe eV

Se 掺杂	菱方相		面心立方相	
	富 Ge	富 Te	富 Ge	富 Te
取代 Ge	1.05	0.88	0.67	0.50
取代 Te	-0.41	-0.24	-0.54	-0.37

详细解释见后续的电子态密度分析。

2.2 含本征 Ge 空位的 GeTe 晶体的 Se 掺杂

本文 2.1 节计算基于 GeTe 的完美晶体模型。然而,实际上 GeTe 是一种非化学计量比的半导体,GeTe 中存在高浓度的本征 Ge 空位。对 GeTe 进行掺杂研究时,必须考虑其中大量存在的本征 Ge 空位的影响。由于 Ge、Se 具有相近的离子半径(分别为 67 pm 和 64 pm),Vinod 和 Sangunni^[11] 推测 Se 原子可能占据 Ge 空位的位置。因此,本文研究了 Se 掺杂含 Ge 空位的 GeTe 体系(Ge₃₁Te₃₂,共63 个原子)。对 Se 占据 Ge 空位和 Se 取代 Te 原子的情况作了重点比较。

根据掺入的 Se 原子与 Ge 空位距离的不同,构建了多种模型。计算得到的形成能如图 3 所示。当 Se 占据 Ge 空位时,Se 原子与 Ge 空位距离即为 0。对于菱方相,Ge 空位最近邻 Te 原子有 2 种,分别为同层最近邻和邻层最近邻,用 A1 和 A2 表示。点间连线以便于观察。从图 3 可以看出,无论对于菱方相还是面心立方相,当存在 Ge 空位时,掺杂的 Se 原子依然不会占据 Ge 空位位置,而是倾向于取代 Ge 空位最近邻的 Te 原子。对于菱方相,则优先取代 Ge 空位同层最近邻的 Te 原子。从另一个角度来看,Se 杂质原子与 Ge 空位具有一定互相吸引作用。这种吸引作用会抑制 Ge 空位的移动。

图 4 为菱方相中 Se 取代 Ge 空位同层最近邻 Te 原子时(A1 构型)和邻层最近邻 Te 原子时(A2 构型)的原子结构,虚线环表示 Ge 空位,深色球、浅色球分别表示 Ge、Te 原子,1 号原子为 Se。形成能计算结果表明,Se 原子更倾向于占据 Ge 空位同层最近邻的位置。A1 与 A2 之间形成能差异在 50 meV/supercell 左右。由图 4 可以发现,Ge 空位使体系中形成了较强的层间 Ge—Te 键,部分抵消了 Ge 空位带来的影响。A1 构型的成键情况为:3 个 Ge—Se 键($R_{1,2} = R_{1,3} = 0.256 \text{ nm}$, $R_{1,4} = 0.297 \text{ nm}$),2 个层间 Ge—Te 较强键($R_{5,6} = R_{7,8} = 0.312 \text{ nm}$)。A2 构型成键情况为:3 个 Ge—Se 键($R_{1,2} = 0.282 \text{ nm}$, $R_{1,3} = R_{1,4} = 0.260 \text{ nm}$),3 个层间 Ge—Te 键($R_{5,6} = R_{7,8} = 0.312 \text{ nm}$, $R_{9,10} = 0.309 \text{ nm}$)。由于 Se 取代同层的最近邻 Te 原子(A1)相较于邻层的最近邻 Te 原子(A2)所引起的晶格畸变更小,因此更容易形成。

进一步对含 Ge 空位的菱方相和面心立方相的 Se 掺杂进行了电子局域函数分析,如图 5 所示。其中图 5(a)~(c)为菱方相,图 5(d)~(f)为面心立方相。从图 5(a)中可以看出,对于菱方相,Ge 空位对 ELF 的影响十分局域。Ge 空位同层最近邻 Te 原子(Te1),与邻层的 Ge 原子(Ge1)形成较强的层间 Ge—Te 结合($R_{\text{Ge1-Te1}} = 0.311 \text{ nm}$)。

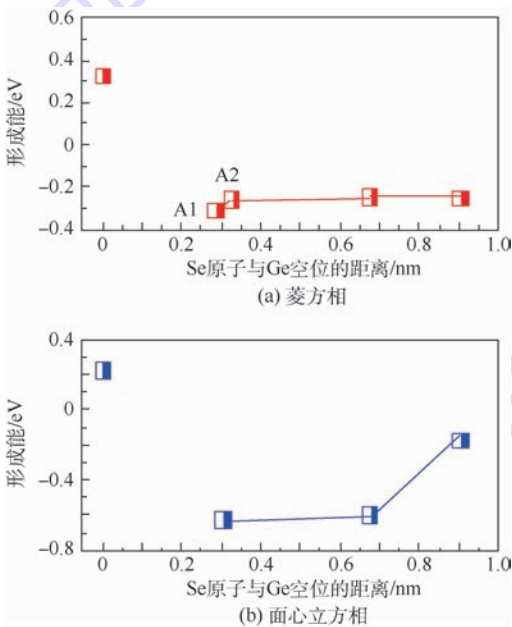


图 3 Se 掺杂 Ge₃₁Te₃₂ 菱方相和 Ge₃₁Te₃₂ 面心立方相的形成能随 Se 原子与 Ge 空位之间距离的变化
Fig. 3 Formation energy of Se doping in rhombohedral Ge₃₁Te₃₂ and fcc Ge₃₁Te₃₂ versus distance between doped Se atom and Ge vacancy

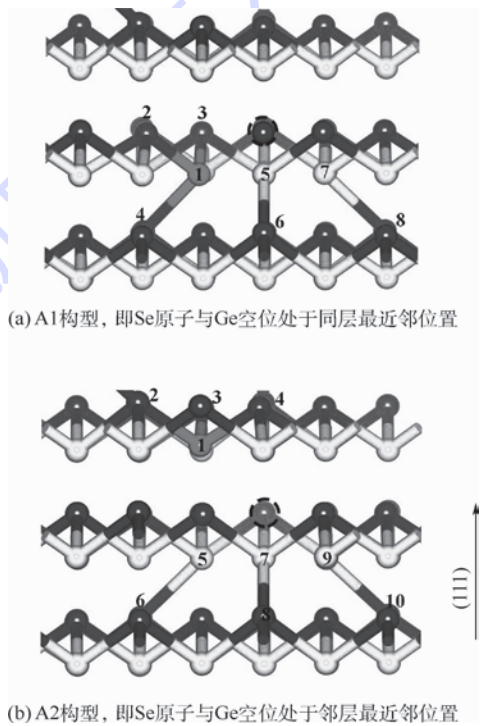


图 4 Se 掺杂 GeTe 的 2 种构型
Fig. 4 Two configurations of Se doped GeTe

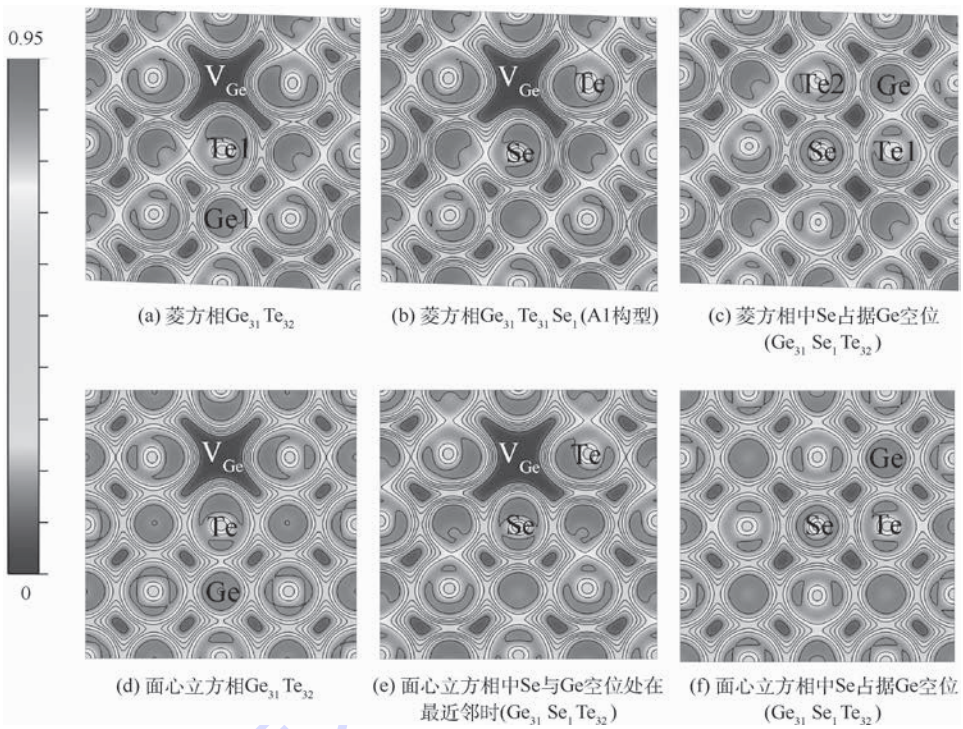


图5 Se掺杂含Ge空位的GeTe体系的在(100)面上的ELF截面图(等高线间隔为0.12)

Fig. 5 ELF contour plots on (100) plane for Se doped GeTe systems with intrinsic Ge vacancy (interval is 0.12)

图5(b)为Se占据Ge空位同层最近邻时的情况,ELF分布与未掺杂体系十分接近,因此Se掺杂对GeTe菱方相电子分布的作用较小。而当Se占据Ge空位时,Se原子周围的电子分布类似于球形,形成近乎等同的6个Se—Te键, ($R_{\text{Se-Te1}} = 0.309 \text{ nm}$, $R_{\text{Se-Te2}} = 0.316 \text{ nm}$)。此时Se杂质原子对周围Te原子的电子局域函数分布造成较大的影响,导致体系能量升高,从而具有较高的形成能。掺杂的Se原子不会占据Ge空位。

对于GeTe面心立方相,Ge空位对电子局域函数的影响也较小。与众不同的是,当掺入的Se原子与Ge空位最近邻时,造成电子局域函数的较大变化。此时GeTe体系中的电子局域函数分布甚至接近于菱方相中的分布,部分Ge—Te键出现长短键分化,从而降低了体系的能量,形成能较低。Se掺杂模糊了GeTe面心立方相和菱方相的界线。而Se占据Ge空位时,对周围Te原子的影响较大,形成能较高,不易存在。

以上分析得出,对于GeTe实际晶体,Se掺杂元素并不会占据Ge空位的位置。掺入的Se原子更倾向于靠近Ge空位,二者具有一定吸引作用,从而限制了Ge空位的移动。GeTe相变存储材料的相转变与电场下Ge空位的定向移动关系密切^[27-28]。Se掺杂对于Ge空位移动的抑制作用,是导致其再结晶温度升高的重要原因。此外,实

验和计算指出,GeTe和 $\text{Ge}_2\text{Sb}_2\text{Te}_5$ 中大量存在的本征Ge空位并非以单一Ge空位的形式存在,而是形成多空位群或者小孔洞^[29-30]。Se杂质原子可能聚集在空位群或孔洞处,从而对相变存储材料的动态相变过程起到更大的影响。Se掺杂含有Ge空位的GeTe体系,使得两晶相的相似度增加。这种性质也可能是实验中观察到Se掺杂PCM出现奇特相变现象的原因。

另外,从表1晶胞数据中可以明显观察到,对于含Ge空位的菱方相,Se掺杂导致晶胞体积减小。而反常的是,Se掺杂含Ge空位的面心立方相导致晶格常数和体积增大。主要是Se掺杂后电子局域函数变化所造成的。Se掺杂后,GeTe菱方相和面心立方相将具有更相近的晶格常数和体积。这对于减小GeTe相变过程中的体积变化十分有益,有助于提高相变存储材料的稳定性。

不同体系的总态密度如图6所示,虚线表示费米能级位置,设置为0 eV处。其带隙变化值如表1所示。Ge空位导致费米能级向价带内移动,呈现出p型半导体的性质。对于GeTe菱方相,Se掺杂对态密度的影响较小,造成带隙的少量降低。3% Se掺杂使菱方结构 $\text{Ge}_{31}\text{Te}_{32}$ 的带隙仅仅下降了1.7%。Se掺杂可使GeTe菱方晶体保持良好的导电性质。此外,对于含Ge空位的GeTe面心

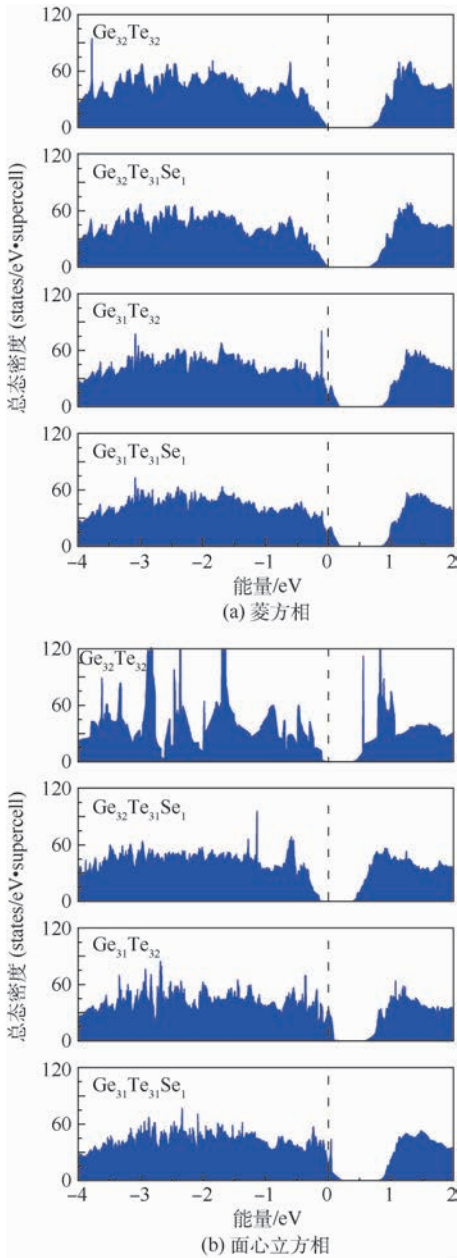


图6 菱方相和面心立方相总态密度

Fig. 6 Total density of states for rhombohedral phase and fcc

立方晶体,Se掺杂则会使导带底向能量更高处移动,显著增加其带隙。面心立方结构的 $Ge_{31}Te_{31}Se_1$ 相较于 $Ge_{31}Te_{32}$,带隙增加了67.7%。这主要是由掺杂后晶格常数增大所引起的。

图7(a)、(b)分别为 $Ge_{32}Te_{31}Se_1$ 菱方相和面心立方相中,掺杂的Se原子与其最近邻Ge原子和Te原子的分波态密度图(PDOS),虚线表示费米能级位置,设置为0 eV处。可以看出,Ge 4p电子和Te 4p电子具有很强的p-p耦合,形成Ge—Te共价键。而Se原子的PDOS分布与Te原子十分相似。掺杂的Se原子与Ge原子形成类似于Ge—Te键的Ge—Se共价键。Se原子的3p电子态密度分布离价带顶较远,处在较低能量位置。

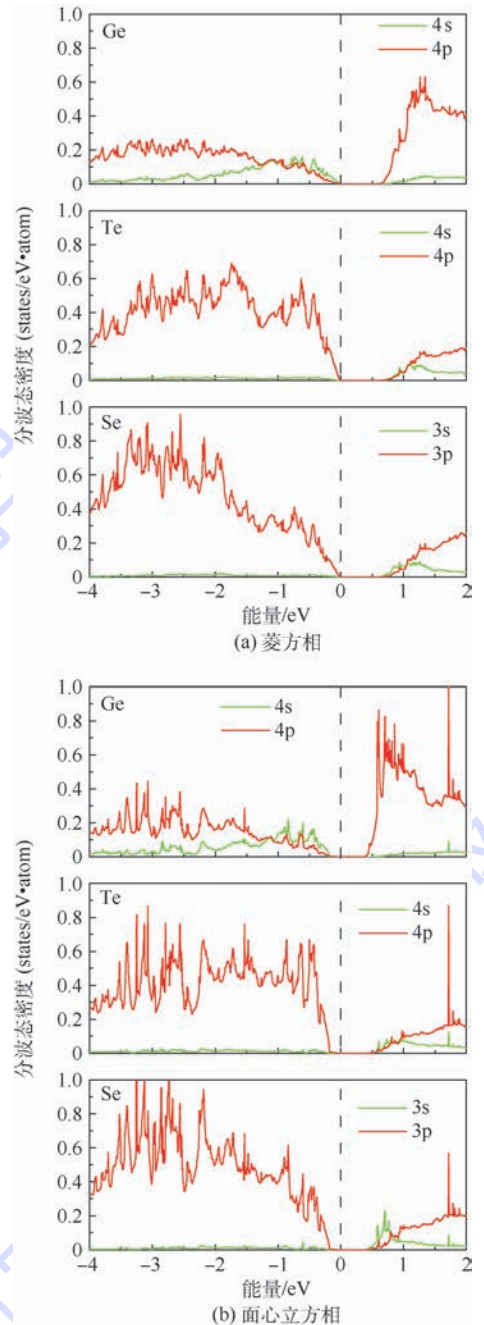


图7 菱方相 $Ge_{32}Te_{31}Se_1$ 和面心立方相 $Ge_{32}Te_{31}Se_1$ Ge原子、Te原子和Se原子的分波态密度

Fig. 7 Partial density of states for Ge atom, Te atom and Se atom in rhombohedral $Ge_{32}Te_{31}Se_1$ and fcc $Ge_{32}Te_{31}Se_1$ Se原子与Ge原子的p电子态密度耦合更加显著,因而Ge—Se共价键较Ge—Te共价键更强。

3 结论

1) 对于GeTe完美晶体,Se掺杂优先取代Te原子。而对于含Ge空位的GeTe晶体,Se原子不会占据Ge空位,而更倾向于取代Ge空位最近邻的Te原子。对于菱方相,则倾向于取代Ge空位同层最近邻的Te原子。Se杂质原子与Ge空位存在一定的吸引作用,阻碍了Ge空位的移动,可

以提高 GeTe 相变存储材料的再结晶温度,使其具有更好的数据保持力和更高的服役温度。

2) Se 掺杂导致菱方相的体积收缩,而面心立方相的体积膨胀,减小了由菱方相到面心立方相转变时的体积差异,提高了相变存储材料的稳定性。

3) Se 原子掺杂对 GeTe 菱方相的电子结构影响较小,使其带隙略微减小。故少量 Se 掺杂不会影响 GeTe 菱方相的导电性质。此外,Se 掺杂使面心立方相的带隙增大。Se 掺杂含 Ge 空位的面心立方晶体呈现类似于菱方相的电子局域函数分布,减小了两相的差异。为解释实验中 Se 掺杂 PCM 出现的奇特相转变性质提供了重要线索。

综上,Se 掺杂的 GeTe 是一种性质优异的相变存储材料,具有较高的再结晶温度,适用于航空航天、汽车电子等较高温度的信息存储领域。对 Se 掺杂后的非晶态性质需要后期进一步探索。

致谢 感谢北京航空航天大学司晨老师在研究过程中给予的帮助和指导。

参考文献 (References)

- [1] OVSHINSKY S R. Reversible electrical switching phenomena in disordered structures [J]. *Physical Review Letters*, 1968, 21 (20): 1450-1453.
- [2] RAOUX S, WELNIC W, IELMINI D. Phase change materials and their application to nonvolatile memories [J]. *Chemical Reviews*, 2010, 110 (1): 240-267.
- [3] WONG H S P, SALAHUDDIN S. Memory leads the way to better computing [J]. *Nature Nanotechnology*, 2015, 10 (3): 191-194.
- [4] LENCER D, SALINGA M, GRABOWSKI B, et al. A map for phase-change materials [J]. *Nature Materials*, 2008, 7 (12): 972-977.
- [5] LENCER D, SALINGA M, WUTTIG M. Design rules for phase-change materials in data storage applications [J]. *Advanced Materials*, 2011, 23 (18): 2030-2058.
- [6] RAOUX S. Phase change materials [J]. *Annual Review of Materials Research*, 2009, 39 (1): 25-48.
- [7] PERNIOLA L, SOUSA V, FANTINI A, et al. Electrical behavior of phase-change memory cells based on GeTe [J]. *IEEE Electron Device Letters*, 2010, 31 (5): 488-490.
- [8] BRUNS G, MERKELBACH P, SCHLOCKERMANN C, et al. Nanosecond switching in GeTe phase change memory cells [J]. *Applied Physics Letters*, 2009, 95 (4): 043108.
- [9] CHONG T, SHI L, ZHAO R, et al. Phase change random access memory cell with superlattice-like structure [J]. *Applied Physics Letters*, 2006, 88 (12): 122114.
- [10] VINOD E M, SINGH A K, GANESAN R, et al. Effect of selenium addition on the GeTe phase change memory alloys [J]. *Journal of Alloys and Compounds*, 2012, 537 (19): 127-132.
- [11] VINOD E M, SANGUNNI K S. The effect of Se doping on spectroscopic and electrical properties of GeTe [J]. *Thin Solid Films*, 2014, 550 (1): 569-574.
- [12] KOLOBOV A V, TOMINAGA J, FONS P, et al. Local structure of crystallized GeTe films [J]. *Applied Physics Letters*, 2003, 82 (3): 382-384.
- [13] TONG F, MIAO X S, WU Y, et al. Effective method to identify the vacancies in crystalline GeTe [J]. *Applied Physics Letters*, 2010, 97 (26): 261904.
- [14] VINOD E M, RAMESH K, SANGUNNI K S. Structural transition and enhanced phase transition properties of Se doped Ge₂Sb₂Te₅ alloys [J]. *Scientific Reports*, 2015, 5: 8050.
- [15] WANG M, LU Y, SHEN X, et al. Effect of Sb₂Se on phase change characteristics of Ge₂Sb₂Te₅ [J]. *CrystEngComm*, 2015, 17 (26): 4871-4876.
- [16] HAFNER J. Ab-initio simulations of materials using VASP: Density-functional theory and beyond [J]. *Journal of Computational Chemistry*, 2008, 29 (13): 2044-2078.
- [17] BLÖCHL P E. Projector augmented-wave method [J]. *Physical Review B*, 1994, 50 (24): 17953-17979.
- [18] PERDEW J P, WANG Y. Pair-distribution function and its coupling-constant average for the spin-polarized electron gas [J]. *Physical Review B*, 1992, 46 (20): 12947-12954.
- [19] MIAO N, SA B, ZHOU J, et al. Investigation on Ge_{5-x}Sb_xTe₅ phase-change materials by first-principles method [J]. *Applied Physics A*, 2010, 99 (4): 961-964.
- [20] GOLDAK J, BARRETT C S, INNES D, et al. Structure of alpha GeTe [J]. *The Journal of Chemical Physics*, 1966, 44 (9): 3323-3325.
- [21] LEVIN E M, BESSER M F, HANUS R. Electronic and thermal transport in GeTe: A versatile base for thermoelectric materials [J]. *Journal of Applied Physics*, 2013, 114 (8): 083713.
- [22] CHATTOPADHYAY T, BOUCHERLE J. Neutron diffraction study on the structural phase transition in GeTe [J]. *Journal of Physics C: Solid State Physics*, 1987, 20 (10): 1431.
- [23] PEIERLS R E. Quantum theory of solids [M]. Oxford: Oxford University Press, 1955.
- [24] SUN Z, TIAN S, SA B. Investigation of the structure and properties of rhombohedral Cu-Ge-Te alloys by ab initio calculations [J]. *Intermetallics*, 2013, 32 (2): 292-296.
- [25] ZHANG S, WEI S H, ZUNGER A, et al. Defect physics of the CuInSe₂ chalcopyrite semiconductor [J]. *Physical Review B*, 1998, 57 (16): 9642.
- [26] WEI S H. Overcoming the doping bottleneck in semiconductors [J]. *Computational Materials Science*, 2004, 30 (3): 337-348.
- [27] NAM S W, CHUNG H S, LO Y C, et al. Electrical field driven and dislocation-templated amorphization in phase-change nanowires [J]. *Science*, 2012, 336 (6088): 1561-1566.
- [28] NUKALA P, AGARWAL R, QIAN X, et al. Direct observation of metal-insulator transition in single-crystalline germanium telluride nanowire memory devices prior to amorphization [J]. *Nano Letters*, 2014, 14 (4): 2201-2209.
- [29] SUN Z, ZHOU J, BLOMQVIST A, et al. Formation of large voids in the amorphous phase-change memory Ge₂Sb₂Te₅ alloy [J].

Physical Review Letters, 2009, 102:075504.

- [30] UPADHYAY M, ABHAYA S, MURUGAVEL S, et al. Experimental evidence for presence of voids in phase change memory material[J]. RSC Advances, 2014, 4(8):3659-3668.

作者简介:

方治乾 男, 硕士研究生。主要研究方向: 相变存储材料。

缪奶华 男, 博士, 副教授, 硕士生导师。主要研究方向: 计算材料学、相变存储材料、热电材料。

周健 男, 博士, 副教授, 硕士生导师。主要研究方向: 计算材料学、相变存储材料、热电材料。

First-principles study of Se doped GeTe phase-change material

FANG Zhiqian¹, MIAO Naihua^{1,2}, ZHOU Jian^{1,2,*}

(1. School of Material Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;

2. Center for Integrated Computational Materials Engineering, International Research Institute for Multidisciplinary Science, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: Doping Se can significantly improve the recrystallization temperature of GeTe phase-change material according to recent experiments, endowing GeTe with a higher working temperature and better data retention. However, the impact of Se on the structure and electrical properties of GeTe is not clear. In this paper, we investigated the effect of Se on the microstructure, bonding characters and electrical properties of crystalline GeTe using first-principles calculation. The results show that the doping Se atom prefers to replace Te in ideal GeTe, while for GeTe systems with intrinsic Ge vacancies, Se tends to replace the Te atoms which are the nearest neighbors of Ge vacancy. The attraction between Se atom and Ge vacancies hinders the movement of Ge vacancies, and thus increases the recrystallization temperature. Furthermore, a shrink of lattice volume and a small reduction of band gap are found in rhombohedral GeTe with Ge vacancies through doping Se, while in face-centered cubic GeTe with Ge vacancies, Se doping causes an expansion in lattice volume and an increase in band gap. Doping Se reduces the volume discrepancy between the two crystalline phases. The calculation results provide clues for explaining the unique phase transformation phenomena of Se doped phase-change materials.

Keywords: Se doped GeTe; Ge vacancy; phase-change materials; first-principles calculation; elemental doping

Received: 2017-05-18; Accepted: 2017-06-12; Published online: 2017-07-14 17:27

URL: kns.cnki.net/kcms/detail/11.2625.V.20170714.1727.001.html

Foundation item: National Natural Science Foundation of China (61274005)

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

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

DOI: 10.13700/j.bh.1001-5965.2017.0357

一种基于深度学习的交互式电话号码识别方法



韩京冶¹, 许福^{1,*}, 陈志泊¹, 刘辉²

(1. 北京林业大学 信息学院, 北京 100083; 2. 北京航空航天大学 计算机学院, 北京 100083)

摘 要: 物流、保险和中介服务等行业需要频繁地拨打电话, 而人工拨打电话效率较低, 高效的电话号码识别技术具有重要的应用价值。传统的印刷体数字识别方法存在人工设计特征过程复杂、识别字体单一等不足, 难以满足实际应用需求。本文提出了一种基于深度学习的交互式的电话号码识别方法, 通过鼠标双击图像中的电话号码, 自动截取出包含此号码的目标区域, 并进行灰度化、二值化、目标区域定位、字符分割和图片补白等预处理操作, 在此基础上利用改进的 LeNet-5 卷积神经网络 (CNN) 自动学习图像特征, 支持多种字体、字形和字号的印刷体数字识别, 并利用交互式识别和内存池等方法提高识别速度。实验结果表明, 单一字符的识别率为 99.86%, 整个号码的识别率为 99.50%, 整个号码平均识别时间为 91 ms。本文方法识别精度高、识别速度快, 具有较为广泛的应用前景。

关键词: 深度学习; 卷积神经网络 (CNN); 电话号码识别; 交互式识别; 目标区域定位

中图分类号: TP391.43

文献标识码: A

文章编号: 1001-5965(2018)05-1074-07

物流、保险、中介服务等行业需要频繁拨打电话联系客户, 日均通话量上百个, 希望以最快速度拨出电话。传统的人工拨打方式存在易拨错、效率低等问题, 采用电话号码识别技术拨打电话是一种有效的解决途径。

电话号码识别的核心技术是印刷体数字识别, 这是模式识别的重要研究领域, 具有较高的学术研究和广泛的商业应用前景。现有的印刷体数字识别方法主要包括 2 类: 一类为基于结构特征的方法, 如罗佳和王玲^[1]提出了一种基于凹凸特性笔顺编码的识别方法。倪桂博和梁晓尊^[2]通过分析印刷体数字的形状, 提出了一种基于结构形状的印刷体数字识别方法。陈爱斌和陆丽娜^[3]提出了基于多特征的印刷体数字识别方法, 采用提取数字区域像素、水平过线和垂直过线

的特征进行识别。另一类为基于统计特征的方法, 如曾志军和孙国强^[4]提出了一种基于改进的 BP 网络实现字符识别的方法。刘春丽和吕淑静^[5]提出了方向特征和密度特征相混合的识别方法。上述 2 类方法均需手工设计特征模板, 依赖设计者的先验知识和手工参数调配, 因而只能支持相对有限的特征, 且不具备可移植性。手工设计特征工程量大、耗时长, 难以应对多种字体、字形的识别。虽然市面上的光学字符识别 (Optical Character Recognition, OCR) 软件如汉王、AB-BYY 和云脉 OCR 等能够识别多种字体、字形的数字, 但需要对整个图像进行处理和识别, 不能只针对图像中的某个号码进行识别, 执行速度较慢, 难以满足实际应用需求。

本文采用深度学习 (deep learning) 和交互式

收稿日期: 2017-05-27; 录用日期: 2017-06-23; 网络出版时间: 2017-09-19 14:41

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20170919.1441.001.html

基金项目: 国家自然科学基金 (61772078); 北京市重点研发计划 (D171100001817003); 中央高校基本科研业务费专项资金 (YX2014-17)

* 通信作者. E-mail: xufu@bjfu.edu.cn

引用格式: 韩京冶, 许福, 陈志泊, 等. 一种基于深度学习的交互式电话号码识别方法[J]. 北京航空航天大学学报, 2018, 44(5): 1074-1080. HAN J Y, XU F, CHEN Z B, et al. A deep learning based interactive recognition method for telephone numbers[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 1074-1080 (in Chinese).

方法解决上述传统印刷体数字识别方法存在的问题。深度学习由 Hinton 和 Salakhutdinov^[6] 在 2006 年提出,是机器学习的一个新领域。深度学习与传统模式识别方法的最大不同在于,其能从数据中自动学习特征表示,可以包含成千上万的参数。相比于传统的模式识别方法,识别精度更高、抗干扰能力更强。目前深度学习已成功应用于人脸识别^[7-9]、行人检测^[10]、机器翻译^[11]、图像分类^[12-13]、场景识别^[14]和门牌识别^[15]等领域。采用深度学习方法可以自动提取并学习印刷体数字的特征,可以更有效地支持多种字体、字形的印刷体数字识别。利用交互式识别方法,通过鼠标双击图像中的电话号码,并只针对包含该号码的目标区域进行处理和识别,有助于提高识别速度,能够满足实际应用需求。

1 电话号码图像预处理

图像预处理是电话号码识别的重要步骤,预处理的效果将直接影响最终识别结果,具体流程如图 1 所示。

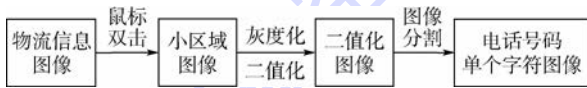


图 1 预处理流程图

Fig. 1 Preprocessing flowchart

1.1 鼠标双击截取目标区域

如图 2 所示,物流配送软件中包含大量的电话号码,在实际应用中只需对单一号码进行拨号。传统的光学字符识别需要对整个图像进行预处理和识别,执行速度较慢。本文采用交互式的识别方法,通过鼠标双击图像中的电话号码,根据点击的坐标截取出包含该电话号码的目标区域(见

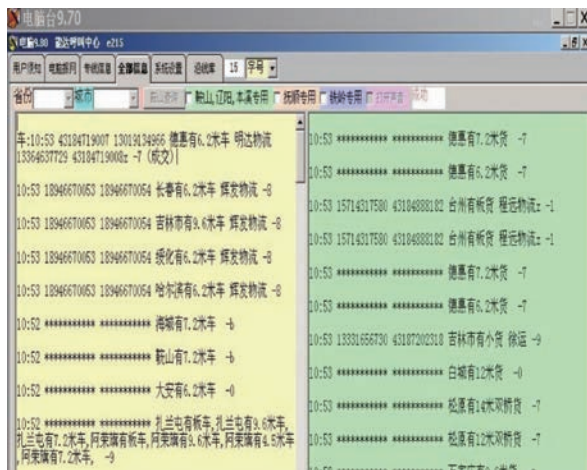


图 2 原始图像

Fig. 2 Original image

图 3), 然后对目标区域进行后续的预处理和识别,有效减少了程序的执行时间。



图 3 电话号码目标区域图像

Fig. 3 Image of target area containing telephone numbers

1.2 灰度化

将彩色图像转换成灰度图像,减少后续图像处理的计算量,且灰度图像仍然能反应图像整体及局部的亮度、色度等重要特征,不影响最后的识别效果。本文采用加权平均值法,将 R、G 和 B 3 个分量赋予合适的权值进行加权平均,得到灰度化图像(见图 4)。



图 4 灰度化图像

Fig. 4 Grayscale image

1.3 二值化

本文采用最大类间方差 (OTSU) 算法^[16],通过阈值来计算图像中目标区域和背景区域的类间方差,将原图像区分出前景和背景,经过 OTSU 算法,得到二值化图像(见图 5)。

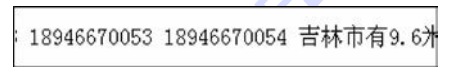


图 5 二值化图像

Fig. 5 Binary image

根据背景区域和目标区域的颜色差异,有可能生成白底黑字或黑底白字 2 类二值化图像。本文的训练集均为白底黑字图像,为提高识别精度,需要将黑底白字二值化图像转成白底黑字图像。方法是计算二值化图像中黑色像素占总像素的比例,当该值大于 50% 时,进行颜色反转操作,将其转化为白底黑字图像。

1.4 目标区域定位

目标区域定位是本文提出的交互式识别方法的关键步骤。截取到的电话号码目标区域图像(见图 3)可能包含多个号码,相邻电话号码间通常会以空格符区分。因此,需要区分出多个电话号码间的空格(见图 3 中的①),以及同一电话号码内部的连续数字之间的空白(见图 3 中的②),由此定位出鼠标双击的号码。随着字号的增大,空格符的大小和连续数字的间隔大小也会增大。本文采用字符高度与其间距的比值作为划分电话号码边界的特征值。为此,选取了宋体、仿宋、楷

体、黑体、微软雅黑、隶书、幼圆、Times New Roman、Cambria 和 Calibri 10 种常见字体, 9 号、10 号、11 号、12 号、14 号和 16 号 6 种常见字号, 计算其数字横向宽度, 找出横向宽度最小和最大的数字。显然, 横向宽度最小的 2 个数字间空白最大, 横向宽度最大的 2 个数字间的空白最小。因此, 计算区分连续数字、数字间存在空格 2 种情况的阈值 T , 只需计算横向宽度最大的 2 个数字间存在空格的情况下, 字符高度与其间距的比值 t_{down} , 即可确定出下界。然后, 计算横向宽度最小的 2 个连续数字(二者间无空格)情况下, 数字高度与其间距的比值 t_{up} , 即可确定出上界。最终, 阈值 T 的取值范围是 $t_{\text{down}} < T < t_{\text{up}}$ 。以宋体和黑体为例, 当数字为宋体时, 横向宽度最小的是 1, 而其余 9 个数字宽度相等; 当数字为黑体时, 横向宽度最小的数字是 1, 最大的数字是 4, 计算结果如表 1 和表 2 所示。

由表 1 可知, 宋体 9 号字时, 阈值范围为 1.33 ~ 2.67; 10 号字时, 阈值范围为 1.13 ~ 2.25; 11 号字时, 阈值范围为 1.00 ~ 3.33; 12 号字时, 阈值范围为 1.10 ~ 3.67; 14 号字时, 阈值范围为 1.30 ~ 4.33; 16 号字时, 阈值范围为 1.67 ~ 3.50。选取所有下界中的最大值和所有上界中的最小值, 即阈值范围为 1.67 ~ 2.25 时, 区分连续数字和数字间存在空格 2 种情况的效果最好。最后, 统计 10 种字体的阈值范围如表 3 所示, 取最大的下界 t_{down} 为 2.00, 最小的上界 t_{up} 为 2.25, 得到阈

表 1 宋体数字高度与数字间距离的比值

Table 1 Ratio of number's height to distance in SimSun font

字 号	下界(t_{down})	上界(t_{up})
9 号	1.33	2.67
10 号	1.13	2.25
11 号	1.00	3.33
12 号	1.10	3.67
14 号	1.30	4.33
16 号	1.67	3.50

表 2 黑体数字高度与数字间距离的比值

Table 2 Ratio of number's height to distance in SimHei font

字 号	下界(t_{down})	上界(t_{up})
9 号	1.33	4.00
10 号	1.29	4.50
11 号	1.25	2.50
12 号	1.57	3.67
14 号	1.30	2.60
16 号	1.36	3.00

表 3 10 种字体的阈值范围

Table 3 Threshold ranges of ten types of fonts

字 体	阈值范围
宋体	1.67 ~ 2.25
黑体	1.57 ~ 2.50
仿宋	1.33 ~ 2.25
楷体	1.33 ~ 2.25
微软雅黑	2.00 ~ 3.00
隶书	1.38 ~ 2.50
幼圆	1.10 ~ 2.25
Times New Roman	1.71 ~ 2.50
Cambria	2.00 ~ 3.00
Calibri	1.86 ~ 4.00

值范围为 2.00 ~ 2.25, 然后, 取上界和下界的平均值, 得到 $T = 2.125$, 此时区分效果最好。

目标区域定位主要方法是, 获取电话号码所在位置的坐标, 并以该坐标为基准上下移动扫描, 当扫描结果全为白色像素线时, 则确定其为上下边界。以该坐标为基准纵向扫描, 记录连续白色像素的列数即 2 个数字间的距离, 然后计算数字高度与数字间距离的比值 X 。当 X 小于阈值 T 时, 为数字间存在空格的情况, 即可确定出号码的左右边界; 反之, 为连续数字情况, 应归类为同一个号码。目标区域定位后截取图像如图 6 所示。

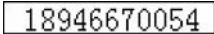


图 6 截取后图像

Fig. 6 Cropped image

1.5 字符分割与图片补白

电话号码中包含多个字符, 识别时将其转化为单个字符进行判断, 因此需要进行分割操作, 将电话号码分割为单个字符。屏幕中电话号码的常见字号为 6 磅(8 像素) ~ 20 磅(26 像素), 因此本文的训练集采用 28 像素 \times 28 像素的图像。因字号大小不同, 分割后的图像大小也会不同(见图 7(a))。为保证其特征不变, 需要进行归一化操作, 将分割后长宽大于 28 像素的字符图像进行等比例缩放, 再通过补白操作, 将其调整为 28 像素 \times 28 像素(见图 7(b))。



图 7 补白前后图像对比

Fig. 7 Comparison of image before and after padding

2 基于改进 LeNet-5 网络的识别方法

卷积神经网络(Convolutional Neural Network,

CNN)的 LeNet-5 模型,最初是由 Lecun 等^[17]提出的用于二维图像识别。原始的 LeNet-5 模型采用 Sigmoid 激活函数为

$$f(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

但 Sigmoid 函数容易产生梯度消失,导致训练出现问题。因此本文对原始 LeNet-5 模型进行改进,将 Sigmoid 激活函数替换为 ReLU 函数为

$$f(x) = \max(0, x) \quad (2)$$

和 Sigmoid 函数相比,使用 ReLU 可以大幅加快随机梯度下降算法的收敛速度,并通过简单阈值就可以得到激活值,减少了复杂的计算。

如图 8 所示,本文改进的 LeNet-5 模型由 7 层组成(不含输入层),以 28 像素 × 28 像素的字符图像作为输入,其中网络 C1 是由 20 个特征图组成的卷积层,每层神经元与输入图像的一个 5 像素 × 5 像素的邻域相连接,每个特征图的大小是 24 像素 × 24 像素。本文在 C1 ~ S2 层中加入

ReLU 激活函数,在训练的初始化阶段对神经元进行抑制和激活,稀疏后的模型能够更好地挖掘数据相关特征。网络层 S2 是由 20 个 12 像素 × 12 像素的特征图组成的降采样层,特征图的每个神经元与 C1 层的一个大小为 2 像素 × 2 像素的邻域连接。网络层 C3 是由 50 个 8 像素 × 8 像素的特征图组成的卷积层,特征图的每个神经元与 S2 层的若干个特征图的 5 像素 × 5 像素的邻域连接。网络层 S4 是由 50 个 4 像素 × 4 像素的特征图组成的降采样层,特征图的每个神经元与 C3 层的一个 2 像素 × 2 像素的邻域连接。网络层 C5 是由 500 个 1 像素 × 1 像素的特征图组成的卷积层,每个神经元与 S4 层的所有特征图的 4 像素 × 4 像素的邻域相连。网络层 F6 包括 500 个神经元,与 C5 层进行全连接,本文将这里原始的 Sigmoid 激活函数替换为 ReLU 函数。最后,根据本文数据集的类别个数,将输出层调整为 13 个神经元,数据的分类判断在这一层完成。

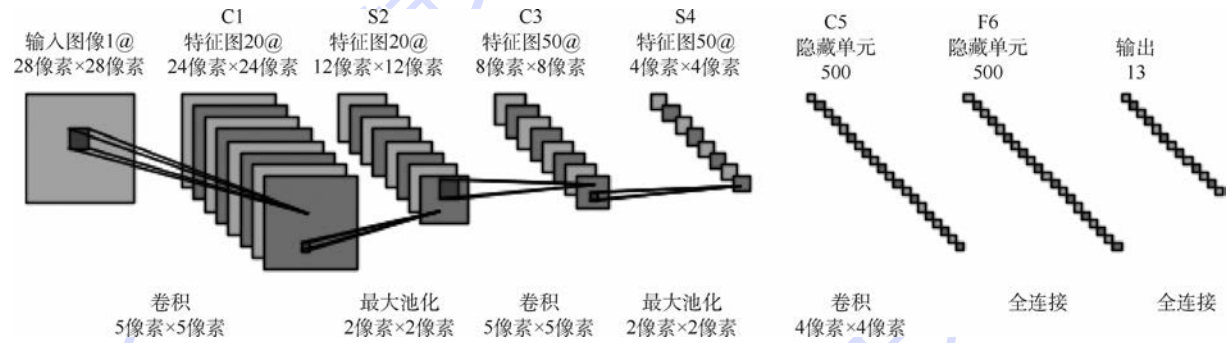


图 8 改进的 LeNet-5 CNN 结构

Fig. 8 Improved LeNet-5 CNN structure

3 实验

本文实验流程如图 9 所示,主要包括训练和测试 2 部分组成。训练部分包括数据集生成、改进的 LeNet-5 神经网络搭建、模型训练与生成;测试部分包括电话号码图像预处理,以及调用生成的模型进行号码识别。

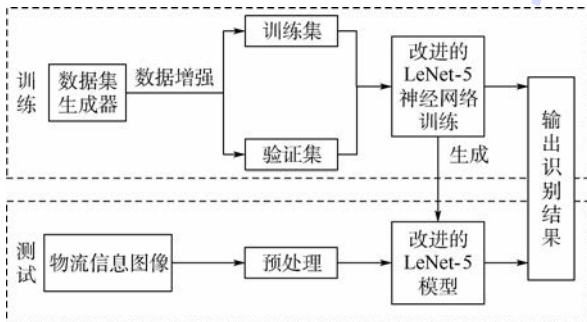


图 9 实验流程图

Fig. 9 Experimental flowchart

3.1 数据集建立

传统模式识别任务中,构造适用于多种字体的特征模板是一个耗时和繁琐的工作。本文采用自动生成图像数据集的方法,设计一款数据集生成器。原理是通过读取系统字库文件,自动生成不同的字体、字号和字形的字符图像,这样既缩短数据集的准备时间,又达到数据增强的目的,使数据集包含更多的不同字体、字形和字号特征的字符(见图 10)。

电话号码一般包括 0~9 十个阿拉伯数字、冒号、横线和汉字(电话:010-1234567)。为提高识别精度,在 0~9 十个数字类的数据集基础上,添加冒号类、横线类和汉字类,在识别过程中将这 3 类字符区分出来,从而更好地识别数字字符。通过本文的数据集生成器生成宋体、仿宋、楷体、黑体、幼圆、微软雅黑、隶书、Times New Roman、Cambria 和 Calibri 10 种常见字体,9 号、10 号、

11号、12号、14号和16号6种常见字号,以及常规、倾斜、粗体和粗体倾斜4种常见字形的字符图像,这样每类字符有240张图片。为了进一步增加数据集大小和一定程度上避免过拟合,对生成图片进行任意角度旋转、上下左右平移操作,使每类数据增加到2700张。

本文数据集共13类,每类2700张,总共35100张字符图像。随机选取75%的数据用作训练集,另外25%的数据用作验证集。

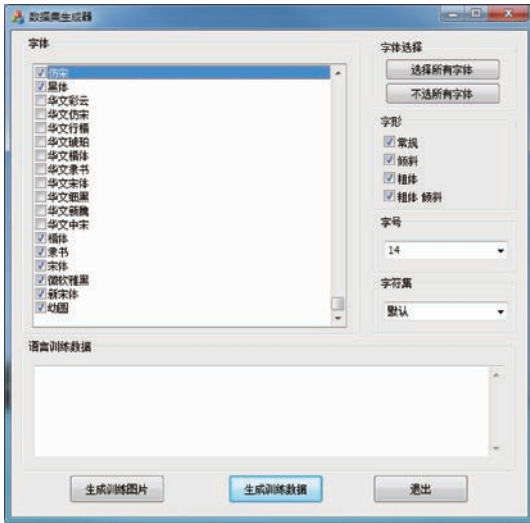


图10 数据集生成器

Fig. 10 Dataset generator

3.2 模型训练

本文在Linux环境下,利用深度学习框架Caffe搭建改进的LeNet-5神经网络模型,并利用NVIDIA GTX 970 GPU加速模型训练。如图11所示,验证集上的识别精度曲线(黄色)识别精度稳定在99%以上,具有较好的识别效果。

目前Windows平台在桌面领域占据垄断地位,为避免网络延迟、服务器过载等问题,本文没有采用传统的服务器远程识别方法,而是将Caffe

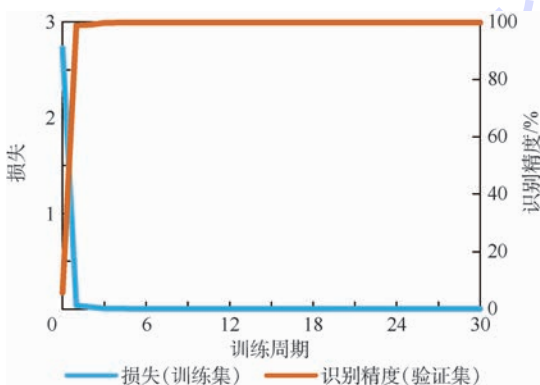


图11 模型训练的损失和识别精度曲线

Fig. 11 Loss and recognition accuracy curves of model training

框架移植到Windows平台,构造模型进行本地识别,这样不依赖网络环境和第三方服务,能更好地优化识别速度。

3.3 识别速度优化

为了提高电话号码的识别速度,主要在5个方面进行优化。第一,整个识别任务在用户本地电脑进行,不依赖网络环境和第三方服务,避免网络延迟导致识别速度降低。第二,与传统OCR识别相比,通过交互式的识别方法,不需要对整个版面进行处理,只对鼠标双击的号码区域进行处理,缩短了预处理时间;第三,生成的训练模型文件较大,将模型文件常驻内存,避免重复加载模型;第四,在内存中完成图像预处理和字符识别操作,减少IO操作时间;第五,在字符分割时会产生多个字符图像,由于所申请字符图像内存块的大小不一,频繁使用时会产生大量的内存碎片,进而降低性能。因此,采用内存池分配方式进行优化,提高内存分配效率,进而提高识别速度。

4 实验结果与分析

为了检验本文方法的识别效果,实现了一个基于深度学习的交互式电话号码识别系统。随机选取物流配送信息中1000个包含多种字体、字号和字形的电话号码进行测试。实验机配置:CPU Intel Core i5-4590、8G内存、64-bit Windows 7旗舰版。

4.1 识别精度

在上述1000个号码的测试结果中,单个字符的识别率为99.86%,整个电话号码的识别率为99.50%。文献[3]中报告的单个字符最好的识别率为99.70%,本文方法的单个字符识别精度略高于文献[3]中提及的最优方法,同时其方法所选用的测试集均为白底黑字的印刷体数字图片,不能处理复杂背景的印刷体数字,而本文选用测试集如图2所示,包含不同背景颜色和多种文本信息的物流信息图片,复杂度更高,识别难度更大。

4.2 识别速度

在验证识别精度的同时,也进行了识别速度的验证,图12给出了其中一次测试结果。本文选用了3款主流商用OCR软件:“ABBYY FineReader 12”、“汉王PDF OCR 8.0”和开源OCR引擎“Tesseract-OCR v3.05”。利用上述软件,对1000个电话号码图像进行识别。对比结果如表4所示,可以看出,3款软件识别1000个电话号码平均识别时间均在200ms以上,而本文方法的平均识

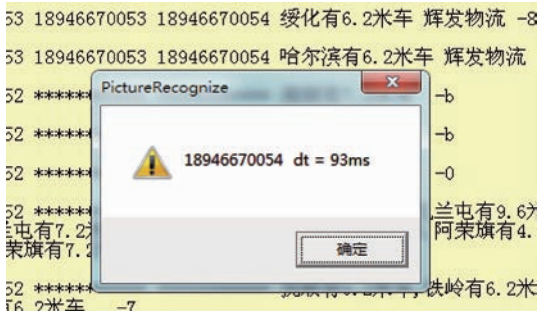


图 12 识别速度测试实例

Fig. 12 An example of recognition speed test

表 4 本文方法与 3 种软件方法识别速度对比

Table 4 Comparison of recognition speed between proposed method and three software methods ms

识别方法	单个电话号码平均识别时间
本文方法	91
Tesseract-OCR v3.05	225
汉王 PDF OCR 8.0	383
ABBYY FineReader 12	433

别时间为 91 ms,速度明显优于上述 3 种主流 OCR 软件。本文采用交互式识别方法,只对图像中感兴趣的号码区域进行处理和识别,缩短了预处理时间。同时通过本地化识别、内存中完成预处理、内存池优化等操作,进一步提高识别速度。

5 结 论

本文提出了一种基于深度学习的交互式电话号码识别方法,主要结论如下:

1) 本文方法可达到较为优异的识别效果,单个字符的识别率 99.86%,整个电话号码的识别率 99.50%,略高于传统方法中最好的识别率。本文方法避免了传统手工构造特征模板的不足,利用卷积神经网络自动提取和学习字符图像特征,识别精度更高、适应能力更强。

2) 本文提出的交互式识别方法,通过鼠标双击图像中的号码,截取包含该号码的目标区域进行预处理,并通过目标区域定位方法提取号码,避免了传统识别方法需要对整个版面进行处理的不足,大幅提高了识别速度。同时,将模型文件常驻内存,避免了重复加载,将图像处理和字符识别操作在内存中完成,并利用高效的内存池分配方式等方法优化了识别速度,相比文中提及 3 种主流的 OCR 软件,速度提升了 2 倍以上。

3) 统计了常见的字体、字号的印刷体数字高度与数字间距离的比值。实验结果表明,该值能作为划分电话号码边界的特征,是实现交互式识别方法的重要特征。

4) 改进了原始的 LeNet-5 卷积神经网络方法,使用 ReLU 激活函数代替 Sigmoid 函数,防止梯度消失,增加了稀疏激活性,加快了计算速度。

5) 通过自动读取系统字库文件自动生成不同字体、字形和字号的字符图片,解决了字符识别中数据集准备繁琐、耗时和字符特征单一的问题。

参考文献 (References)

- [1] 罗佳,王玲.基于凹凸特性笔顺编码的手写体数字识别方法[J].计算机工程与科学,2010,29(5):69-70.
LUO J,WANG L.A new method for the off-line recognition of handwritten digits based on convex-concave coding[J].Computer Engineering & Science,2010,29(5):69-70(in Chinese).
- [2] 倪桂博,梁晓尊.基于结构形状的印刷体数字识别方法[J].软件导刊,2010,9(5):67-68.
NI G B,LIANG X Z.The method of printed figures based on structure[J].Software Guide,2010,9(5):67-68(in Chinese).
- [3] 陈爱斌,陆丽娜.基于多特征的印刷体数字识别[J].计算技术与自动化,2011,30(3):105-108.
CHEN A B,LU L N.The printed number character recognition based on feature[J].Computing Technology and Automation,2011,30(3):105-108(in Chinese).
- [4] 曾志军,孙国强.基于改进的 BP 网络数字字符识别[J].上海理工大学学报,2008,30(2):201-204.
ZENG Z J,SUN G Q.Number character recognition based on improved BP neural network[J].Journal of University of Shanghai for Science and Technology,2008,30(2):201-204(in Chinese).
- [5] 刘春丽,吕淑静.基于混合特征的孟加拉手写体数字识别[J].计算机工程与应用,2007,43(20):214-215.
LIU C L,LV S J.Bangla handwritten numeral recognition based on blend features[J].Computer Engineering & Applications,2007,43(20):214-215(in Chinese).
- [6] HINTON G E,SALAKHUTDINOV R R.Reducing the dimensionality of data with neural networks[J].Science,2006,313(5786):504-507.
- [7] SUN Y,WANG X,TANG X.Deep learning face representation from predicting 10,000 classes[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE Press,2014:1891-1898.
- [8] TAIGMAN Y,YANG M,RANZATO M,et al.DeepFace:Closing the gap to human-level performance in face verification[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE Press,2014:1701-1708.
- [9] SUN Y,CHEN Y,WANG X,et al.Deep learning face representation by joint identification-verification[C]//International Conference on Neural Information Processing Systems. London:MIT Press,2014:1988-1996.
- [10] ZHANG L,LIN L,LIANG X,et al.Is faster R-CNN doing well for pedestrian detection?[C]//European Conference on Computer Vision. Berlin:Springer,2016:443-457.
- [11] SINGH S P,KUMAR A,DARBARI H,et al.Machine transla-

- tion using deep learning: An overview [C] // International Conference on Computer, Communications and Electronics. Piscataway, NJ: IEEE Press, 2017: 162-167.
- [12] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2016: 770-778.
- [13] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks [C] // Advances in Neural Information Processing Systems. London: MIT Press, 2012: 1097-1105.
- [14] CORDTS M, OMRAN M, RAMOS S, et al. The cityscapes dataset for semantic urban scene understanding [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2016: 3213-3223.
- [15] FARABET C, COUPRIE C, NAJMAN L, et al. Learning hierarchical features for scene labeling [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35 (8): 1915-1929.
- [16] OTSU N. A threshold selection method from gray-level histograms [J]. IEEE Transactions on Systems, Man, and Cybernetics, 1979, 9 (1): 62-66.
- [17] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86 (11): 2278-2324.

作者简介:

韩京冶 男, 硕士研究生。主要研究方向: 图像处理、深度学习。

许福 男, 博士, 副教授。主要研究方向: 图像处理、编译技术、软件工程。

陈志泊 男, 博士, 教授, 博士生导师。主要研究方向: 数据库技术、林业信息工程。

刘辉 男, 博士。主要研究方向: 软件工程。

A deep learning based interactive recognition method for telephone numbers

HAN Jingye¹, XU Fu^{1,*}, CHEN Zhibo¹, LIU Hui²

(1. School of Information Science and Technology, Beijing Forestry University, Beijing 100083, China;

2. School of Computer Science and Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: Some sectors such as logistics, insurance and intermediary agents need to make calls frequently. Manually callings lead to low efficiency, so that telephone number recognition has important practical values. The traditional methods for printed number recognition involve complicated templates designing, which cannot meet the requirements of practical applications. An interactive method based on deep learning is proposed to recognize telephone numbers. Through double-clicking the phone number in an image, this method automatically crops the target area which contains the number and performs preprocessing operations such as grayscale, binarization, target area localization, character segmentation and image padding. An improved LeNet-5 convolutional neural network (CNN) is utilized to make image recognition, which supports the recognition of printed numbers in a variety of fonts, glyphs and font sizes. The recognition speed is optimized through multiple means such as interactive recognition and memory pool. Experimental results show that the accuracy of recognition for a single character is 99.86%, the accuracy for a telephone number is 99.50%, and the average recognition time of a telephone number is 91 ms. Comparing with the traditional methods, the new method has relatively higher accuracy and faster speed in recognition, which can be widely used in many sectors.

Keywords: deep learning; convolutional neural network (CNN); telephone number recognition; interactive recognition; target area localization

Received: 2017-05-27; **Accepted:** 2017-06-23; **Published online:** 2017-09-19 14:41

URL: kns.cnki.net/kcms/detail/11.2625.V.20170919.1441.001.html

Foundation items: National Natural Science Foundation of China (61772078); Key R&D Program of Beijing (D171100001817003); the Fundamental Research Funds for the Central Universities (YX2014-17)

* **Corresponding author.** E-mail: xufu@bjfu.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2017.0383

基于 Wiener 过程的发动机多阶段剩余寿命预测

黄亮, 刘君强*, 贡英杰

(南京航空航天大学 民航学院, 南京 211106)



摘 要: 针对现阶段发动机的寿命预测研究没有考虑到非线性与多阶段的问题, 提出了基于多阶段非线性 Wiener 过程的航空发动机实时剩余寿命预测的方法。该方法融合了同类型发动机的历史性能退化监测数据与个体发动机的实时监测数据。首先, 考虑了发动机性能退化非线性的特点, 并采用多阶段 Wiener 过程建立发动机的性能退化模型。然后, 根据发动机的历史性能监测数据, 利用极大似然估计和一维搜索方法进行参数先验分布的估计。再次, 在先验分布和个体发动机的退化数据的基础上, 用贝叶斯方法对参数分布更新。最后, 得到个体发动机剩余寿命的实时预测值。通过实例验证本文方法预测的准确性。

关键词: 剩余寿命; 多阶段; 非线性; Wiener 过程; 实时预测; 贝叶斯方法

中图分类号: V263

文献标识码: A

文章编号: 1001-5965(2018)05-1081-07

发动机的失效是材料处于恶劣的运行环境下缓慢退化的结果^[1-2]。随着系统复杂程度的提高, 传统基于失效机理模型的方法很难构建出可靠的故障模型来对应产品的失效过程^[3-4]。基于统计性能退化监测数据的方法只需要运行过程中产生的退化数据, 不必建立精确的数学模型, 具有较高的计算优势。

目前, 基于统计性能退化监测数据的寿命预测是研究的热门^[5-6], 并且已经研究出多种性能退化模型, 例如随机系数模型^[7]、随机过程模型^[8-9]和随机滤波模型^[10]等。基于统计性能退化监测数据的寿命预测方法, 利用退化过程的随机性求出剩余寿命的概率分布情况, 便于反映预测结果的随机性。在基于随机过程模型中, 最常采用 Wiener 过程。Wiener 过程便于描述系统因受到外部环境的影响、内部的状态改变以及负载情况, 而具有随机的非单调独立增量过程^[11]。由于 Wiener 过程可以用来描述随机的非单调独立增

量过程, 目前大量采用 Wiener 过程对复杂系统进行性能退化建模^[12]。通常所说的 Wiener 过程是指具有线性漂移系数的一类随机过程^[13-14]。Wang^[15]提出了基于期望最大化 (Expectation-Maximization, EM) 算法的性能退化模型参数估计方法。

上述研究丰富了 Wiener 过程在剩余寿命预测领域中的应用, 但还存在着如下几个问题:

1) 假设系统的性能退化过程是关于时间的线性函数, 或运用某些尺度转换方法将非线性关系转化为线性关系^[16]。例如 Gebraeel^[17]采用取对数方法将指数退化模型转变为一般的线性退化模型。然而, 还有很多非线性退化过程无法转换为线性的, 例如航空发动机性能退化过程具有明显的非线性特点, 直接建立非线性退化模型更满足实际情况。司小胜等^[18]采用非线性 Wiener 过程对复杂系统的剩余寿命进行研究。

2) 在目前的剩余寿命预测研究中, 鲜有研究

收稿日期: 2017-06-06; 录用日期: 2017-08-31; 网络出版时间: 2017-10-30 17:39

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20171030.1739.002.html

基金项目: 民航联合研究基金 (U1533128)

* 通信作者。E-mail: liujunqiang@nuaa.edu.cn

引用格式: 黄亮, 刘君强, 贡英杰. 基于 Wiener 过程的发动机多阶段剩余寿命预测[J]. 北京航空航天大学学报, 2018, 44(5): 1081-1087. HUANG L, LIU J Q, GONG Y J. Multi-phase residual life prediction of engines based on Wiener process[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 1081-1087 (in Chinese).

多阶段退化的情况。刘君强等^[19]在发动机的寿命预测中研究了多阶段的情况,实验结果表明发动机的退化过程具有多阶段性。

3) 如何体现个体性能退化的差异性。由于环境、材料和误差等的影响,同类型产品的性能退化过程存在差异性,在产品的剩余寿命预测中需要考虑个体性能退化的差异性。刘君强等^[19]提出性能退化模型参数具有随机性,服从高斯-伽马分布,从而反映出个体退化的差异性。

针对以上问题,本文根据航空发动机性能退化具有非线性、多阶段的性质,用多阶段非线性 Wiener 过程构建出性能退化模型。根据多个同类产品历史性能退化监测数据,利用极大似然估计与一维搜索方法,进行参数先验分布的估计,在先验分布和实时获取的个体发动机退化数据的基础上,用贝叶斯方法求得参数的后验分布,从而实现个体发动机剩余寿命的实时精确预测。

1 剩余寿命预测模型

1.1 模型分析与假设

通过分析航空发动机具有复杂多阶段非线性的性能退化特点,本文将采用多阶段非线性的 Wiener 过程建立性能退化模型。在该非线性 Wiener 过程中漂移系数不再与时间无关,而是关于时间的非线性函数。基于非线性 Wiener 过程的性能退化模型可以表示为

$$X(t) = \int_0^t u(t; \theta) dt + \sigma B(t) \quad (1)$$

式中: $X(t)$ 为在时刻 t 的性能退化量; $u(t; \theta)$ 为非线性 Wiener 过程的漂移系数, θ 为漂移系数的非线性函数中的未知参数; σ 为扩散系数; $B(t)$ 为标准布朗运动。当 $u(t; \theta) = u$ 时,上述非线性性能退化模型转化为一般线性 Wiener 退化模型。

为了通过实际发动机监测数据进行实例的验证,根据文献[8],本文假设 $u(t; \theta) = urt^{r-1}$, 此时的非线性退化模型变为

$$X(t) = ut^r + \sigma B(t) \quad (2)$$

式中: r 为漂移系数的非线性函数中的未知参数。

利用 Wiener 过程进行多阶段非线性的退化建模需要做出以下假设:

假设 1 当性能退化量没有超过失效阈值时,发动机处于正常状态;超过失效阈值时,则发动机处于失效状态。

假设 2 为描述航空发动机性能退化的个体差异性,将模型参数 u, σ^2 视为随机变量,设 $w = 1/\sigma^2$, 可得 u, w 的联合分布是高斯-伽马分布:

$$f(u, w) = \frac{w^{1/2}}{\sqrt{2\pi c}} \exp\left[-\frac{w(u-d)^2}{2c}\right] \cdot \frac{b^a}{\Gamma(a)} w^{a-1} \exp(-bw) \quad (3)$$

式中: a, b, c, d 为 u 和 w 组成的联合分布中的参数。

1.2 多阶段性能退化模型

多阶段的性能退化模型可以表示为

$$X(t) = (X(0) + X_1(t))I_{(0, t_1)}(t) + (X(t_1) + X_2(t - t_1))I_{(t_1, t_2)}(t) + \dots + (X(t_{n-1}) + X_n(t - t_{n-1}))I_{(t_{n-1}, \infty)}(t) \quad (4)$$

式中: $I(t)$ 为示性函数; $X(0)$ 为性能退化模型的初始值; $X(t_i)$ 为 t_i 时的性能退化值, t_i 为到达每个性能退化分界值的时间; $X_n(t - t_{n-1})$ 为第 n 阶段的退化函数。

根据 1.1 节分析与假设,本文将使用多阶段非线性的 Wiener 随机过程构建出性能退化模型。此时, $X(t)$ 表示为

$$X(t) = (u_1 t^{r_1} + \sigma_1 B(t))I_{(0, t_1)}(t) + [W_1 + u_2(t - t_1)^{r_2} + \sigma_2 B(t - t_1)]I_{(t_1, t_2)}(t) + \dots + [W_{n-1} + u_n(t - t_{n-1}) + \sigma_n B(t - t_{n-1})]I_{(t_{n-1}, \infty)}(t) \quad (5)$$

式中: W_{n-1} 为第 $n-1$ 个性能退化阶段的分界值。

1.3 寿命分布

剩余寿命 L_r 定义为产品自 t' 时的退化量 $X(t')$ 到其第一次超出失效阈值 W_n 所用的时间, 剩余寿命 L_r 表达式为

$$L_r = \inf\{X(t+t') \geq W_n \mid X(t') \leq W_n\} = \inf\{X(t+t') - X(t') \geq W_n - X(t')\} \quad (6)$$

式中: $X(t')$ 和 $X(t+t')$ 分别为在 t' 和 $t'+t$ 时刻的性能退化量; W_n 为失效阈值。

因为本文构建的退化模型考虑了非线性的性质,很难计算求得 L_r 的准确分布。为此, Si 等^[8]给出了剩余寿命概率密度函数的一个近似表达式。航空发动机的剩余寿命 L_r 的概率密度函数为

$$f(t|t') = \frac{W_n - X(t') - u(t+t')^r + ut'^r + urt(t+t')^{r-1}}{\sqrt{2\pi\sigma^2 t^3}} \exp\left[-\frac{(W_n - X(t') - u(t+t')^r + ut'^r)^2}{2\sigma^2 t}\right] \quad (7)$$

1.4 剩余寿命预测的多阶段性

令 ξ_k 为航空发动机性能退化量 $X(t)$ 从初始时刻到其第一次超过第 k 个性能退化阶段分界值 W_k 的时间, ξ_k 为

$$\xi_k = \inf\{t; X(t) \geq W_k, t > 0\} \quad (8)$$

根据 Wiener 过程的齐次马尔可夫性质,则 ξ_k

可以改写为

$$\xi_k = \xi_{k-1} + \inf\{t; X'(t) \geq W_k - W_{k-1}, t > 0\} \quad (9)$$

式中: $X'(t)$ 为第 k 个阶段的性能退化模型; $X'(t) = u_k t^{r_k} + \sigma_k B(t)$, 其中 u_k 、 r_k 和 σ_k 表示第 k 个阶段性能退化模型的参数。

令 $\Delta\xi_k = \xi_k - \xi_{k-1}$, $\Delta W_k = W_k - W_{k-1}$, 则式(9)

可以改写为

$$\Delta\xi_k = \inf\{t; X'(t) \geq \Delta W_k, t > 0\} \quad (10)$$

根据式(7)可得第 k 个性能退化阶段的寿命分布函数为

$$f(\Delta\xi_k | u_k, r_k, \sigma_k) = \frac{\Delta W_k - u_k \Delta\xi_k^{r_k} + u_k r_k \Delta\xi_k^{r_k}}{\sqrt{2\pi\sigma_k^2 \Delta\xi_k^3}} \cdot \exp\left[-\frac{(\Delta W_k - u_k \Delta\xi_k^{r_k})^2}{2\sigma_k^2 \Delta\xi_k^3}\right] \quad (11)$$

由式(11)可以推导出航空发动机寿命 T 为

$$T = \Delta\xi_1 + \Delta\xi_2 + \dots + \Delta\xi_n \quad (12)$$

由于 $E(\Delta\xi_k) = \left(\frac{\Delta W_k}{u_k}\right)^{\frac{1}{r_k}}$, 航空发动机寿命 T

的期望为

$$E(T) = \left(\frac{W_1 - X(0)}{u_1}\right)^{\frac{1}{r_1}} + \left(\frac{\Delta W_2}{u_2}\right)^{\frac{1}{r_2}} + \dots + \left(\frac{\Delta W_n}{u_n}\right)^{\frac{1}{r_n}} \quad (13)$$

当 t' 时刻的性能退化量 $X(t')$ 满足要求

$X(\xi_{k-1}) \leq X(t') \leq X(\xi_k)$, $L_{t'}$ 表达式为

$$L_{t'} = (\xi_k - t') + \Delta\xi_{k+1} + \dots + \Delta\xi_n \quad (14)$$

2 参数估计

由于很难直接估算出未知超参数 a 、 b 、 c 和 d 的值, 本文将采用两阶段方法进行估计。首先, 根据现有的航空发动机的历史退化监测数据, 利用极大似然估计与一维搜索方法估算出一组 u 、 σ^2 的值。然后, 根据估算出的一组 u 、 σ^2 的值, 再利用极大似然估计方法估算出 a 、 b 、 c 和 d 的值。

2.1 第 1 阶段参数估计

假设现在共有 m 个同类型的航空发动机的性能退化监测数据, 其中第 i 个发动机的第 k 个阶段的监测时刻记为 $t_{ik|1}, t_{ik|2}, \dots, t_{ik|n}$, 与此时刻相对应的性能退化监测数据记为 $y_{ik|1}, y_{ik|2}, \dots, y_{ik|n}$ 。

根据多阶段非线性 Wiener 过程的性质, 可知 $X_{ik} \sim N(u_{ik}t^{r_k}, \sigma_{ik}^2 t)$, 其中 u_{ik} 、 r_k 和 σ_{ik} 为第 i 个发动机的第 k 个阶段的 Wiener 过程的未知参数, 可求得参数 $\Theta = [u_{ik}, r_k, \sigma_{ik}]$ 基于第 i 个发动机的第 k 个阶段的退化监测数据的对数似然函数为

$$l(\Theta | X) = -\frac{n_{ik}}{2} \ln(2\pi) - n_{ik} \ln \sigma_{ik} - \frac{1}{2} \sum_{j=1}^{n_{ik}} \ln t_{ik|j} - \sum_{j=1}^{n_{ik}} \left\{ - (y_{ik|j} - X(t_{i|k-1}) - u_{ik}(t_{ik|j} - t_{i|k-1})^{r_k})^2 / [2\sigma_{ik}^2 (t_{ik|j} - t_{i|k-1})] \right\} \quad (15)$$

式中: $y_{ik|j}$ 为第 i 个发动机在第 k 个阶段的第 j 次监测值; $t_{i|k-1}$ 为第 i 个发动机到达第 $k-1$ 个阶段分界值的时间; n_{ik} 为第 i 个发动机在第 k 个阶段的监测次数; $t_{ik|j}$ 为第 i 个发动机在第 k 个阶段的第 j 次监测时间。

先固定未知参数 r_k , 对 u_{ik} 、 σ_{ik}^2 ($i = 1, 2, \dots, m$) 求偏导, 计算结果为

$$\begin{cases} u_{ik} = \frac{\sum_{j=1}^{n_{ik}} (y_{ik|j} - X(t_{i|k-1}))}{\sum_{j=1}^{n_{ik}} (t_{ik|j} - t_{i|k-1})^{r_k}} \\ \sigma_{ik}^2 = \frac{1}{n_{ik}} \sum_{j=1}^{n_{ik}} \left\{ [y_{ik|j} - X(t_{i|k-1}) - u_{ik}(t_{ik|j} - t_{i|k-1})^{r_k}]^2 / (t_{ik|j} - t_{i|k-1}) \right\} \end{cases} \quad (16)$$

根据式(15)进一步可求得 m 个航空发动机的退化模型参数 $\Theta = [u_{ik}, r_{ik}, \sigma_{ik}]$ 的完全对数似然函数为

$$l(\Theta | X, m) = \sum_{i=1}^m \left(-\frac{n_{ik}}{2} \ln(2\pi) - n_{ik} \ln \sigma_{ik} - \frac{1}{2} \sum_{j=1}^{n_{ik}} \ln t_{ik|j} - \sum_{j=1}^{n_{ik}} \left(- (y_{ik|j} - X(t_{i|k-1}) - u_{ik}(t_{ik|j} - t_{i|k-1})^{r_k})^2 / [2\sigma_{ik}^2 (t_{ik|j} - t_{i|k-1})] \right) \right) \quad (17)$$

将式(16)求得的 u_{ik} 、 σ_{ik} ($i = 1, 2, \dots, m$) 代入式(17), 可求得关于 r_k 的轮廓似然函数, 然后利用一维搜索方法可得 r_k 的估计值, 然后将 r_k 再代入式(16), 可得 u_{ik} 、 σ_{ik}^2 ($i = 1, 2, \dots, m$) 的一组估计值。

2.2 第 2 阶段参数估计

利用第 1 阶段求得的一组 $(u_{1k}, w_{1k}), (u_{2k}, w_{2k}), \dots, (u_{mk}, w_{mk})$ 来估计参数 a_k, b_k, c_k, d_k , 其中 $w_{ik} = 1/\sigma_{ik}^2$ 。由式(3)可得 a_k, b_k, c_k, d_k 的似然函数为

$$l(a_k, b_k, c_k, d_k | u_{ik}, w_{ik}) = \prod_{i=1}^m \frac{w_{ik}^{1/2}}{\sqrt{2\pi c_k}} \exp\left[-\frac{w_{ik}(u_{ik} - d_k)^2}{2c_k}\right] \cdot \frac{b_k^{a_k}}{\Gamma(a_k)} w_{ik}^{a_k-1} \exp(-b_k w_{ik}) \quad (18)$$

极大化 $l(a_k, b_k, c_k, d_k | u_{ik}, w_{ik})$, 求得估计值为

$$\left\{ \begin{aligned} \psi(a_k) - \ln a_k &= \frac{\sum_{i=1}^m \ln w_{ik}}{m} + \ln m - \ln \left(\sum_{i=1}^m w_{ik} \right) \\ b_k &= \frac{m a_k}{\sum_{i=1}^m w_{ik}} \\ c_k &= \frac{\sum_{i=1}^m (w_{ik} u_{ik}^2 - 2 d_k w_{ik} u_{ik} + d_k^2 w_{ik} u_{ik})}{m} \\ d_k &= \frac{\sum_{i=1}^m w_{ik} u_{ik}}{\sum_{i=1}^m w_{ik}} \end{aligned} \right. \quad (19)$$

式中： $\psi(\hat{a})$ 为对数伽马函数。

3 基于贝叶斯方法的模型参数更新

由于材料、环境等外界因素的影响,同一型号发动机的性能退化路径存在着个体的差异性。 $\Delta Y_k = (\Delta y_{1,k}, \Delta y_{2,k}, \dots, \Delta y_{n,k})$ 为个体发动机处于第 k 个退化过程的监测数据。

当获得个体发动机实时性能退化数据后,根据贝叶斯方法可得第 k 个性能退化阶段的参数 u_k 和 w_k 的后验分布为

$$f(u_k, w_k | \Delta Y_k) \propto l(\Delta Y_k | u_k, w_k) f(u_k, w_k) \quad (20)$$

式中: $f(u_k, w_k | \Delta Y_k)$ 为个体发动机第 k 性能退化阶段模型参数的后验分布函数; $f(u_k, w_k)$ 为第 k 性能退化阶段的模型参数 u_k 和 w_k 的先验分布函数; $l(\Delta Y_k | u_k, w_k)$ 为似然函数。

根据文献[20]中共轭先验分布的性质,个体发动机第 k 个性能退化阶段中的模型参数 u_k 和 w_k 更新后的后验分布表达式为

$$f(u_k, w_k | \Delta Y_k) = \frac{w_k^{1/2}}{\sqrt{2\pi c_k'}} \exp\left[-\frac{w_k(u_k - d_k')^2}{2c_k'}\right] \cdot \frac{b_k^{a_k'}}{\Gamma(a_k')} w_k^{a_k'-1} \exp(-b_k' w_k) \quad (21)$$

式中:

$$\left\{ \begin{aligned} a_k' &= n_k/2 + a_k \\ b_k' &= C - B^2/(4A) \\ c_k' &= 1/(2A) \\ d_k' &= B/(2A) \\ A &= 1/(2c_k) + \sum_{j=1}^n \{ [(t_j - t_{k-1})^r - (t_{j-1} - t_{k-1})^r]^2 / (2\Delta t_j) \} \\ B &= d_k/c_k + \sum_{j=1}^n \{ \Delta x_j [(t_j - t_{k-1})^r - (t_{j-1} - t_{k-1})^r]^2 / \Delta t_j \} \\ C &= b_k + d_k^2/(2c_k) + \sum_{j=1}^n \Delta x_j^2 / (2\Delta t_j) \\ \Delta t_j &= t_j - t_{j-1} \\ \Delta x_j &= X(t_j) - X(t_{j-1}) \end{aligned} \right. \quad (22)$$

其中: a_k' 、 b_k' 、 c_k' 、 d_k' 分别为 a_k 、 b_k 、 c_k 、 d_k 更新后的值; t_{k-1} 为发动机到达第 $k-1$ 个性能退化阶段分界值的时间; A 、 B 和 C 为多项式。

由于参数 u 和 w 的联合分布是高斯-伽马分布,更新后的 u_k 和 w_k 的后验分布期望为

$$\begin{cases} u_k = E(u_k | \Delta Y_k) = d_k' \\ w_k = E(w_k | \Delta Y_k) = a_k'/b_k' \end{cases} \quad (23)$$

此时 $\sigma_k^2 = b_k'/a_k'$ 。

4 个体发动机的实时剩余寿命预测

根据发动机的历史性能监测数据和单台发动机的实时监测数据,建立多阶段航空发动机性能退化模型。首先利用极大似然估计和一维搜索方法,进行参数先验分布的估计。在先验分布和实时获取的个体发动机退化数据的基础上,采用贝叶斯方法对参数分布更新。从而实现对个体发动机剩余寿命的实时精确预测。剩余寿命预测的具体步骤如图1所示。

步骤1 基于历史性能退化数据,根据式(14)建立退化模型的对数似然函数,利用极大似然与一维搜索方法估算出模型参数的先验分布。

步骤2 当获取到个体发动机的实时性能退化监测数据后,运用贝叶斯方法,根据式(20)~式(22)对发动机性能退化模型参数进行实时更新。

步骤3 根据式(23)可以求得发动机性能

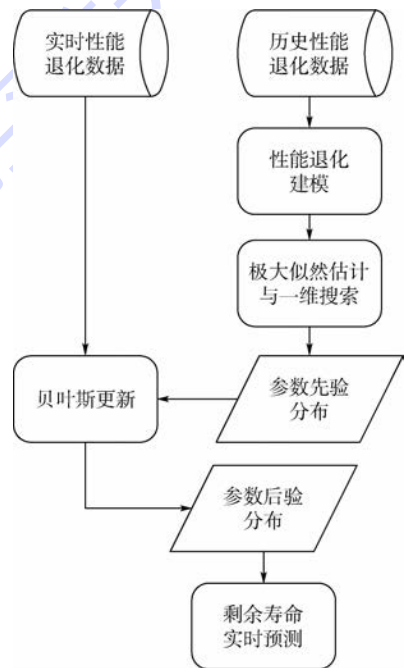


图1 剩余寿命的预测流程

Fig.1 Prediction process of residual life

退化模型中的参数 u, σ^2 的后验分布的期望估计值。

步骤 4 根据式 (11) ~ 式 (14) 多阶段非线性 Wiener 过程的寿命预测方法,求得个体发动机实时的剩余寿命。

5 实例验证

排气温度裕度 (Exhaust Gas Temperature Margin, EGTM) 是监测发动机性能的关键指标。7 台同类型的航空发动机在全寿命周期内监测的 EGTM 数据如表 1 所示。为了验证多阶段非线性 Wiener 过程在航空发动机剩余寿命预测中的准确性,选择表 1 中前 5 台发动机的 EGTM 历史监测数据用作先验参数估计,后 2 台发动机的 EGTM 监测数据用作实时更新。根据经验,EGTM 的失效阈值定义为 0°C 。

图 2 给出了该 7 台发动机的性能退化路径。从图中可知,在大于 35°C 时,EGTM 的上升速度较慢, $0 \sim 35^\circ\text{C}$ 时,EGTM 的退化速度较快。将发动机的退化分界值定义为 $\xi = 35^\circ\text{C}$ 。此时,可采用两阶段非线性的 Wiener 过程进行退化建模。

根据表 1 中给出的航空发动机历史 EGTM 监测数据,利用极大似然估计方法和一维搜索方法,即可得到发动机性能退化模型各阶段参数先验分布的估计值,超参数的估计结果如表 2 所示。

表 1 某型号发动机 EGTM 监测数据

Table 1 EGTM monitoring data of a certain model of engine

当前时间/ cycle	发动机编号						
	1	2	3	4	5	6	7
100	69	70.8	74	67	67	61	69
200	63.3	56.2	70	65	61.4	65.3	63
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
3600	21.6	19	8	-0.3	15	13.9	19.4
3700	18.7	17.1	5.6		10.4	12.5	16.5
3800	9.5	16.4	-0.1		12.8	6.8	13.6
3900	1.2	10.4				-0.9	10.7
4000	-0.8						6.8

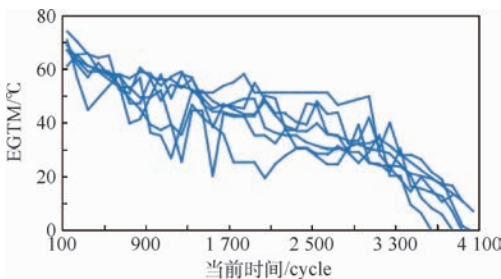


图 2 航空发动机退化路径

Fig. 2 Degradation path of aeroengines

从图 2 中可知,该型号发动机在不同阶段的退化速率存在着明显的差异,进一步说明了发动机的退化过程具有多阶段的特点。

在计算求得航空发动机各阶段性能退化模型参数的先验分布后,可以运用贝叶斯方法,基于实时性能退化数据对模型参数的后验分布进行更新。以编号为 6 的发动机为例,在获得实时性能退化监测数据后,经贝叶斯更新后的模型参数第 1 阶段和第 2 阶段估计结果分别如表 3 和表 4 所示。

在得到了参数更新后的估计值后,可以对发动机性能退化各阶段的剩余寿命进行实时预测。

图 3 和图 4 分别给出了编号为 6 的航空发动机在第 1 和第 2 阶段的实时剩余寿命概率密度分布图。

为直观反映出基于多阶段非线性 Wiener 的退化模型与传统的单阶段线性 Wiener 随机过程在寿命预测问题中的精确性,定义相对误差 (Relative Error, RE) 指标为

$$RE = \frac{|L - T|}{L} \quad (24)$$

式中: L 为预测剩余寿命; T 为实际剩余寿命。

2 种模型的相对误差计算结果如表 5 所示。

根据表 5 的相对误差对比结果,可计算得到单阶段线性与多阶段非线性的平均误差分别为 0.3728 和 0.2506。基于多阶段非线性 Wiener 过程

表 2 超参数的先验估值

Table 2 Prior estimates of hyper parameter

阶段	a	b	c	d	r
1	1.8758	0.0457	0.5613	-1.0007	0.4438
2	0.9711	0.0196	0.0698	-0.1665	0.7919

表 3 贝叶斯更新后的参数第 1 阶段估计结果

Table 3 Bayes updated parameter estimation results at the first stage

当前 时间/ cycle	超参数				估计	
	a	b	c	d	u	σ^2
1000	6.8758	0.8303	0.3825	-1.1262	-1.1262	0.1208
1500	9.3758	6.9966	0.3773	-1.2123	-1.2123	0.7562
2000	11.876	7.2086	0.3774	-1.2244	-1.2244	0.6070

表 4 贝叶斯更新后的参数第 2 阶段估计结果

Table 4 Bayes updated parameter estimation results at the second stage

当前 时间/ cycle	超参数				估计	
	a	b	c	d	u	σ^2
2500	3.4711	0.4523	0.0221	-0.0803	-0.0803	0.1303
3000	5.9711	1.9884	0.0153	-0.0834	-0.0834	0.3330
3500	8.4711	3.2986	0.0122	-0.0958	-0.0958	0.3893

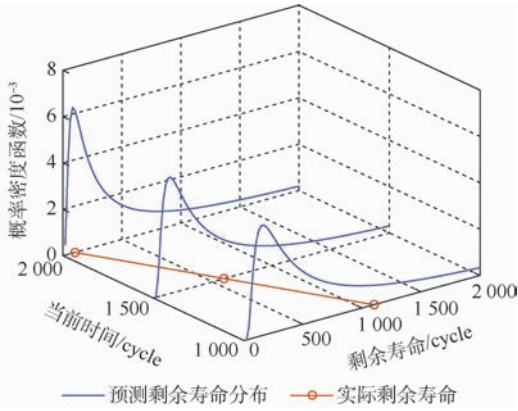


图3 第1阶段剩余寿命概率密度分布

Fig.3 Probability density distribution of residual life at the first stage

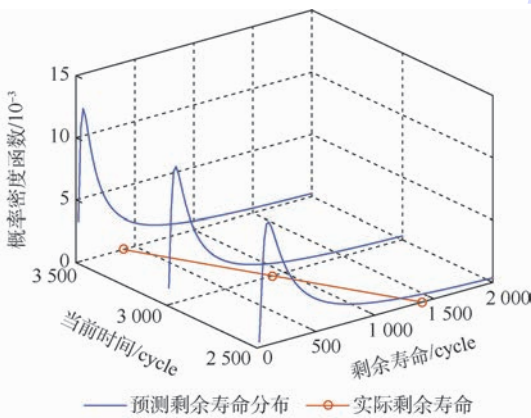


图4 第2阶段剩余寿命概率密度分布

Fig.4 Probability density distribution of residual life at the second stage

表5 2种模型相对误差对比结果

Table 5 Relative error comparison results of two models

当前时间/ cycle	相对误差	
	单阶段线性	多阶段非线性
1 000	0.046 6	0.018 6
1 500	0.072 1	0.037 9
2 000	0.076 8	0.264 2
2 500	0.182 1	0.175
3 000	0.624 4	0.49
3 500	1.235 0	0.518

的剩余寿命预测平均误差值小于传统的预测,证实本文方法用于航空发动机剩余寿命预测的结果更精确。产生以上误差对比结果的原因是航空发动机性能退化具有多阶段非线性特点,基于多阶段非线性 Wiener 过程构建的性能退化模型更加符合实际航空发动机的性能退化路径。此外,同类型发动机由于环境和材料等因素的影响会产生个体的差异性,考虑个体差异性的航空发动机剩余寿命预测结果更加精确。

6 结 论

1) 本文采用了基于多阶段非线性的 Wiener 过程来建立航空发动机的性能退化模型,能够较好地满足航空发动机的实际退化路径。

2) 与不考虑个体差异的 Wiener 过程的性能退化模型相比,考虑个体差异性的 Wiener 过程更能描述复杂系统的退化过程,剩余寿命的估算结果更加的准确。

3) 当获取个体发动机的实时监测数据后,能够结合历史性能退化监测数据,运用贝叶斯方法对模型参数的后验分布进行实时更新,实现对个体发动机剩余寿命的实时精确预测。

参考文献 (References)

- [1] PECHT M G. Prognostics and health management of electronics [M]. Hoboken: Wiley Online Library, 2008: 1-19.
- [2] GUAN Q, TANG Y, XU A. Objective Bayesian analysis accelerated degradation test based on Wiener process models [J]. Applied Mathematical Modelling, 2016, 40(4): 2743-2755.
- [3] 谢吉伟, 刘君强, 王小磊. 应用交互式多模型算法的设备剩余寿命预测 [J]. 空军工程大学学报(自然科学版), 2016, 17(2): 98-102.
XIE J W, LIU J Q, WANG X L. A residual useful lifetime prediction based on interacting multiple model algorithm [J]. Journal of Air Force Engineering University (Natural Science Edition), 2016, 17(2): 98-102 (in Chinese).
- [4] SI X S, WANG W, HU C H, et al. Remaining useful life estimation—A review of the statistical data driven approaches [J]. European Journal of Operational Research, 2011, 213(1): 1-14.
- [5] FENG L, WANG H L, SI X S, et al. A state-space-based prognostic model for hidden and age-dependent nonlinear diffusion degradation process [J]. IEEE Transactions on Automation Science and Engineering, 2013, 10(4): 1072-1086.
- [6] 彭宝华, 周经伦, 孙权. 基于退化与寿命数据融合的产品剩余寿命预测 [J]. 系统工程与电子技术, 2011, 33(5): 1073-1078.
PENG B H, ZHOU J L, SUN Q. Residual lifetime prediction of products based on fusion of degradation data and lifetime data [J]. Systems Engineering and Electronic, 2011, 33(5): 1073-1078 (in Chinese).
- [7] GEBRAEEL N Z, ELWANY A H, JING P. Residual life predictions in the absence of prior degradation knowledge [J]. IEEE Transactions on Reliability, 2009, 58(1): 106-117.
- [8] SI X S, WANG W, HU C H, et al. Remaining useful life estimation based on nonlinear diffusion degradation process [J]. IEEE Transactions on Reliability, 2012, 61(1): 50-67.
- [9] PENG W, LI Y F, YANG Y J, et al. Inverse Gaussian process models for degradation analysis: A Bayesian perspective [J]. Reliability Engineering & System Safety, 2014, 130(1): 175-189.

- [10] WANG W. A two-stage prognosis model in condition based maintenance[J]. *European Journal of Operational Research*, 2007, 182(3):1177-1187.
- [11] 王小林,程志君,郭波. 基于维纳过程金属化膜电容器的剩余寿命预测[J]. *国防科技大学学报*, 2011, 33(4):146-151. WANG X L, CHENG Z J, GUO B. Residual life forecasting of metallized film capacitor based on Wiener process[J]. *Journal of National University of Defense Technology*, 2011, 33(4):146-151 (in Chinese).
- [12] 王小林,郭波,程志君. 融合多源信息的维纳过程性能退化产品的可靠性评估[J]. *电子学报*, 2012, 40(5):977-982. WANG X L, GUO B, CHENG Z J. Reliability assessment of products with Wiener process degradation by fusing multiple information[J]. *Acta Electronica Sinica*, 2012, 40(5):977-982 (in Chinese).
- [13] WANG W, CARR M, XU W, et al. A model for residual life prediction based on Brownian motion with an adaptive drift[J]. *Microelectronic Reliability*, 2011, 51(2):285-293.
- [14] SI X S, WANG W, HU C H, et al. Estimating remaining useful life with three-source variability in degradation modelling[J]. *IEEE Transactions on Reliability*, 2014, 63(1):167-190.
- [15] WANG X. Wiener processes with random effects for degradation data[J]. *Journal of Multivariate Analysis*, 2010, 101(2):340-351.
- [16] PARK C, PADGETT W J. Accelerated degradation models for failure based on geometric Brownian motion and Gamma processes[J]. *Life Time Data Analysis*, 2005, 11(4):511-527.
- [17] GEBRAEEL N Z. Sensory-updated residual life distributions for components with exponential degradation patterns[J]. *IEEE Transactions on Automation Science and Engineering*, 2006, 3(4):382-392.
- [18] 司小胜,胡昌华,周东华. 带测量误差的非线性退化过程建模与剩余寿命估计[J]. *自动化学报*, 2013, 39(5):530-541. SI X S, HU C H, ZHOU D H. Nonlinear degradation process modeling and remaining useful life estimation subject to measurement error[J]. *Acta Automatica Sinica*, 2013, 39(5):530-541 (in Chinese).
- [19] 刘君强,谢吉伟,左洪福,等. 基于随机 Wiener 过程的航空发动机剩余寿命预测[J]. *航空学报*, 2015, 36(2):564-574. LIU J Q, XIE J W, ZUO H F, et al. Residual lifetime prediction for aeroengines based on Wiener process with random effects[J]. *Acta Aeronautica et Astronautica Sinica*, 2015, 36(2):564-574 (in Chinese).
- [20] 茆诗松,汤银才. 贝叶斯统计[M]. 北京:统计出版社, 2012:10-16. MAO S S, TANG Y C. *The Bayesian statistics* [M]. Beijing: China Statistics Press, 2012:10-16 (in Chinese).

作者简介:

黄亮 男,硕士研究生。主要研究方向:发动机健康管理。

刘君强 男,博士,副教授。主要研究方向:民航信息管理、健康管理。

Multi-phase residual life prediction of engines based on Wiener process

HUANG Liang, LIU Junqiang*, GONG Yingjie

(College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

Abstract: Due to the fact that the research on the life prediction of the engine does not take into account both the nonlinearity and the multi-stage problems at the present stage, a method for forecasting the residual life of aeroengines in real time based on multi-phase nonlinear Wiener process is proposed. This method can effectively fuse the historical performance degradation monitoring data of the same type of aeroengines with the real-time monitoring data of the individual aeroengine. Firstly, the nonlinearity of performance degradation is considered, and the multi-stage Wiener process was used to establish the performance degradation model of engine. Secondly, according to the historical performance monitoring data of the aeroengines, the prior distribution of parameters is estimated by using maximum likelihood estimation and one-dimensional search method. Thirdly, according to the real-time performance degradation data of individual aeroengine and the prior distribution, the Bayesian method is used to update the model parameters. Finally, the real-time predicted values of the residual life of the individual aeroengine are obtained. By the test of the actual data, the results show that the proposed method is accurate.

Keywords: residual useful life; multi-phase; nonlinear; Wiener process; real-time prediction; Bayesian method

Received: 2017-06-06; **Accepted:** 2017-08-31; **Published online:** 2017-10-30 17:39

URL: kns.cnki.net/kcms/detail/11.2625.V.20171030.1739.002.html

Foundation item: Joint Research Foundation for Civil Aviation (U1533128)

* **Corresponding author.** E-mail: liujunqiang@nuaa.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2017.0386

一种考虑共因失效的 PMS 可靠性建模分析方法



吴欢, 焦健*, 赵廷弟

(北京航空航天大学 可靠性与系统工程学院, 北京 100083)

摘 要: 共因失效(CCF)打破了系统内组件失效的独立性假设,会对系统特别是多阶段任务系统(PMS)的可靠性评估产生显著影响。针对多阶段任务系统中随机共因失效(PCCF)对任务可靠性的影响问题,对共因事件之间的关系进行分析,利用贝叶斯理论扩展了共因事件的概率模型,使其适用于互斥、相互独立和统计相关等多种统计关系。在此基础上提出了综合应用二元决策图(BDD)和马尔可夫(Markov)模型的模块化建模分析方法。首先,利用故障树对任务过程建模;然后,在考虑共因失效的情况下采用 BDD 和 Markov 模型分别计算系统中静态模块和动态模块;再次,由全概率公式计算任务可靠性;最后,以卫星首次转轨过程为对象,验证了方法的有效性,并通过与已有案例的对比,分析了共因失效对任务可靠性的影响。

关键词: 多阶段任务系统(PMS); 随机共因失效(PCCF); 二元决策图(BDD); Markov 模型; 动态性

中图分类号: V474.2⁺5

文献标识码: A

文章编号: 1001-5965(2018)05-1088-07

复杂工程系统通常包含许多不同功能的子系统,并且系统的任务成功要求系统成功完成多个连续非重叠的子阶段^[1]。在每个阶段中,同一组件的工作方式、环境条件、系统的配置以及失效判据等会随着阶段的改变而改变。通常把在这种具有多个任务的系统称为多阶段任务系统(Phased-Mission System, PMS)。PMS 的可靠性评估目前已经得到广泛的研究,大致可概括为解析法和仿真法 2 种形式。其中解析法又可以细分为 3 种:①组合法^[2],通常使用二元决策图(Binary Decision Diagram, BDD)解决不可修复的 PMS;②状态空间法^[3-4],主要基于 Markov 模型,解决组件的可维修性,但是其规模会随系统组件数量的增长呈指数增长趋势;③模块化方法^[5-8],结合了组合法和状态空间法的优点。

PMS 可靠性评估中常常假设各个单元的失

效是相互独立的,这种假设能够为具体求解带来很大方便,但是不能完全真实地反映实际情况。PMS 中相似单元在同一阶段以及同一单元在多个阶段具有一定的相关性,若简单地在系统各单元失效相互独立的假设下进行系统可靠性分析与计算,常常会导致过大误差。PMS 中往往存在由于某种共同的原因,简称共因(Common Cause, CC),造成多个组件失效,从而导致系统失效,即共因失效(Common Cause Failure, CCF)。

根据共因导致多个组件的失效概率相同与否,共因失效可分为对称性共因失效(各组件失效的概率相同)和非对称性共因失效(各组件的失效概率既可能因组件而异,也可能随时间变化)。在非对称性共因失效研究方面,学者们普遍认为共因对受其影响的单元的作用存在不同概率。文献[9]首次提出了随机共因失效(Probabi-

收稿日期: 2017-06-07; 录用日期: 2017-07-13; 网络出版时间: 2017-09-22 15:07

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20170922.1507.010.html

* 通信作者。E-mail: jiaojian@buaa.edu.cn

引用格式: 吴欢, 焦健, 赵廷弟. 一种考虑共因失效的 PMS 可靠性建模分析方法[J]. 北京航空航天大学学报, 2018, 44(5): 1088-1094. WU H, JIAO J, ZHAO T D. A reliability modeling and analysis method for PMS considering common cause failure[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 1088-1094 (in Chinese).

listic CCF, PCCF)的概念及其逻辑门。文献[10]在此基础上分析了静态与动态容错系统的可靠性。文献[11]提出了改进的显性和隐性方法,分析在内部与外部多个共因影响下系统的可靠性。文献[12]又进一步分析了 PCCF 影响下的多阶段任务系统可靠性。

现有共因失效研究主要限于单阶段任务系统,然而 PMS 的共因失效(特别是随机共因失效)更复杂,尤其是在系统和组件具有动态行为的情况下,现有面向 PMS 的随机共因失效算法不适用于系统的动态性。例如,系统中出现的诸如顺序相关、功能相关等动态逻辑,BDD 方法无法解决此类问题;此外,现有的研究通常假设不同的共因之间统计独立,这与实际系统也存在一定差距。针对这些问题,本文提出一种新的模块化方法,首先,利用 BDD 方法对系统静态模块进行求解,利用 Markov 模型对系统动态模块进行求解;其次,计算扩展共因之间的概率,使其适用于互斥、相互独立和统计相关等不同统计关系,从而使分析结果更符合实际情况。

1 基本假设条件

本文在考虑随机共因失效情况下研究 PMS 可靠性评估问题,遵循以下基本假设条件:

- 1) 组件是不可修复的。
- 2) 系统存在随机共因失效,且不同共因之间存在相关性。
- 3) 系统不同组件之间的失效概率都是相互独立的。
- 4) 组件失效服从指数分布。
- 5) C_{ij} 表示第 i 个阶段出现的第 j 个共因,其出现在阶段 i 开始时刻且在 i 阶段结束时结束,即该共因对组件影响时间为阶段持续时间。

2 共因事件模型

由于共因的存在,PMS 中的多个组件可能会在一个或多个阶段内同时发生失效,但在 PMS 工作运行过程中共因并不是必然发生,因此根据共因是否发生可以将系统的样本空间划分为不同事件。假设系统中有 m 个阶段,存在 L 个共因,根据这 L 个共因发生与否可以将系统的样本空间划分为 2^L 个不相交的事件:若阶段 1 中有 L_1 个共因,阶段 2 中有 L_2 个共因,以此类推阶段 i 有 L_i 个共因,则有 $i = 1, 2, \dots, m; j = 1, 2, \dots, L_n (n = 1, 2, \dots, m)$ 。每个事件称为一个随机共因事件,使

用 E_n 表示,则有 E_1 表示 L 个共因都不发生, E_2 表示 L 个共因中仅有阶段 1 中的第一个共因(即 C_{11})发生,以此类推 E_{2^L} 表示 L 个共因都发生。 E_n 逻辑表达式为

$$\begin{cases} E_1 = \overline{C_{11}} \cap \dots \cap \overline{C_{1L_1}} \cap \dots \cap \overline{C_{i1}} \cap \dots \cap \overline{C_{iL_i}} \cap \dots \cap \overline{C_{m1}} \cap \dots \cap \overline{C_{mL_m}} \\ E_2 = C_{11} \cap \dots \cap \overline{C_{1L_1}} \cap \dots \cap \overline{C_{i1}} \cap \dots \cap \overline{C_{iL_i}} \cap \dots \cap \overline{C_{m1}} \cap \dots \cap \overline{C_{mL_m}} \\ \vdots \\ E_{2^L} = C_{11} \cap \dots \cap C_{1L_1} \cap \dots \cap C_{i1} \cap \dots \cap C_{iL_i} \cap \dots \cap C_{m1} \cap \dots \cap C_{mL_m} \end{cases} \quad (1)$$

式中: $L_1 + L_2 + \dots + L_m = L$ 。

$P_{(E_n)}$ 为 E_n 的发生概率,有 $\sum_{n=1}^{2^L} P_{(E_n)} = 1$,且 $E_i \cap E_j = \emptyset (i \neq j)$ 。

利用共因的发生概率就可以由式(1)计算各个 E_n 的发生概率。目前,研究通常假设不同共因之间是相互独立的,从而可以很方便地计算每个 E_n 的概率。但现实中多个共因之间的关系要复杂得多,利用贝叶斯理论可以对 E_n 的概率计算模型进行扩展^[8],使其包含互斥、相互独立和统计相关等关系,支持可靠性分析评估。限于篇幅,本文以 2 个共因为例,给出各种情况下每个 E_n 的概率计算公式,多个共因时的计算公式可以此类推。

假设存在 2 个共因 C_{11} 和 C_{12} ,其发生概率分别为 P_1 和 P_2 。则共因发生相关性的表达式如下所示:

- 1) 当共因之间为互斥关系时,有

$$\begin{cases} P_{(E_1)} = 1 - P_1 - P_2 \\ P_{(E_2)} = P_1 \\ P_{(E_3)} = P_2 \\ P_{(E_4)} = 0 \end{cases} \quad (2)$$

- 2) 当共因之间相互独立时,有

$$\begin{cases} P_{(E_1)} = (1 - P_1)(1 - P_2) \\ P_{(E_2)} = P_1(1 - P_2) \\ P_{(E_3)} = (1 - P_1)P_2 \\ P_{(E_4)} = P_1P_2 \end{cases} \quad (3)$$

- 3) 当共因之间统计相关时,若

$$\begin{cases} P\{C_{11} | \overline{C_{12}}\} = P_h \\ P\{C_{11} | C_{12}\} = P_{\bar{h}} \end{cases}$$
 则

$$\begin{cases} P_{(E_1)} = (1 - P_2)(1 - P_{\bar{h}}) \\ P_{(E_2)} = (1 - P_2)P_{\bar{h}} \\ P_{(E_3)} = P_2(1 - P_h) \\ P_{(E_4)} = P_2P_h \end{cases} \quad (4)$$

3 考虑共因的 PMS 可靠性评估方法

本文采用模块化建模方法,综合利用 BDD 和 Markov 模型对 PMS 进行建模,使用隐式替代法考虑随机共因失效的影响,从而实现对 PMS 的可靠性进行评估^[13]。

3.1 基础方法理论

1) BDD 模型

基于香农定理,BDD 方法可以表示为 if-then-else 的形式^[14-15]。BDD 主要使用布尔函数表达,故障树在转换为 BDD 时,由于不同底事件的排序会得到不一样的 BDD,相应的计算量也有所不同。本文采用结构式中的从上至下、从左到右的排序方法^[16-17]。

2) Markov 模型

使用 Markov 模型对动态模块进行求解的分析步骤如下:

步骤 1 定义动态模块 M 的状态空间。

步骤 2 描绘状态转移图。根据系统特征以及步骤 1 总结的状态空间来绘制状态转移图。

步骤 3 定义模块 M 在时刻 t 的状态为 $S_i (S_i \in W)$, 其概率用 $P_{i(t)}$ 表示。

步骤 4 建立状态转移方程。使用步骤 3 的定义以及步骤 2 的状态转移图建立状态转移方程。

步骤 5 求解状态转移方程,得出阶段 i 结束时刻 T_i 模块 M 的各个状态 S_i 的状态概率。

步骤 6 计算系统总的失效概率。根据系统功能需求,将每个失效模式的概率相加即为系统失效概率。

3.2 引入共因失效的基本思路

如前所述,系统中若不考虑共因失效,常常会使系统可靠性过于理想化,因此在 PMS 的可靠性评估中应考虑共因失效的影响,且应尽可能考虑共因之间的各种统计关系(如第 2 节所述)。共因失效影响的引入分为以下步骤:

步骤 1 建立所有随机共因的事件空间 E_n , 其中包含所有随机共因发生和不发生的组合。

步骤 2 在不考虑随机共因的影响下,构建系统的可靠性模型,分析模型得出所需要的状态空间。

步骤 3 由步骤 2 的状态空间,计算每个状态空间下每个静态模块所涉及组件在每个随机共因事件影响下的条件故障概率。

步骤 4 评估系统在每个随机共因事件影响下的条件故障概率。

步骤 5 用一个全概率公式将随机共因失效的影响加入整个系统的故障概率中。

3.3 PMS 可靠性评估步骤

在引入共因失效的基本思路指导下,考虑随机共因失效的 PMS 可靠性评估主要分为以下 5 步:

步骤 1 按第 2 节给出的模型及算法构建 PMS 工作运行过程事件空间,并计算各中随机共因事件的概率。

步骤 2 在不考虑随机共因失效的情况下,构建系统的可靠性模型。其主要步骤如下:

1) 首先,将每个阶段用故障树表示,再根据多阶段任务系统的特点得到系统级的故障树;其次,根据 Rauzy 方法^[18]将系统级故障树进行整合得到系统级相互独立的模块,再将每个独立模块作为底事件组成系统级故障树;最后,将系统故障树模型转化并简化为系统级最简 BDD 模型。

2) 根据其逻辑结构特点将上述的独立模块进行静态模块和动态模块归类。使用与、或、表决等静态逻辑门的为静态模块;包含至少一个功能相关、冷备份等动态逻辑门的为动态模块。

步骤 3 计算步骤 2 所得到的系统级失效 BDD 模型中每个静态模块组件的条件失效概率。此时应考虑随机共因失效对组件的影响,即计算条件失效概率。若随机共因事件 E_n 发生时组件 x 受 k 个共因影响,则静态模块中组件 x 在随机共因事件 E_n 发生时的随机失效概率为

$$P_{(x|E_n)} = 1 - \exp\left(-\sum_{i=1}^i \lambda_{ix} T_i\right) \exp\left(-\sum_{i=1}^i \sum_{j=1}^j \lambda_{ijx} T_{ij}\right) \quad (5)$$

式中: λ_{ix} 为组件 x 在阶段 i 的内部失效率; λ_{ijx} 为组件 x 受共因 C_{ij} 影响的失效率; T_{ij} 为组件 x 受共因 C_{ij} 影响的持续时间; $\exp\left(-\sum_{i=1}^i \lambda_{ix} T_i\right)$ 为内部失效影响下组件的可靠性概率; $\exp\left(-\sum_{i=1}^i \sum_{j=1}^j \lambda_{ijx} T_{ij}\right)$ 为随机共因失效影响下组件的可靠性概率。

步骤 4 评估系统在随机共因失效影响下的条件失效概率 $P_{(PMS\text{fails} | E_n)}$ 。对于步骤 2 得到的失效模型,使用 BDD 方法对静态模块进行求解,使用 Markov 模型对动态模块进行求解;进而再计算整个系统的失效概率。其中静态模块中的组件失效概率已由步骤 3 得到,根据实际的逻辑关系可以很容易求得静态模块的条件失效概率;使用 M_{dy} 表示动态模块,则动态模块在随机共因失效影响下的条件失效概率为

$$P_{(M_{dy}, fails | E_n)} = 1 - (1 - P_M) \prod_{i=1}^k (1 - P_{c_i}) \quad (6)$$

式中: P_M 为动态模块的内部失效概率, 通过将组件的内部失效率作为动态模块对应 Markov 链的转移率计算而得; 同理 P_{c_i} 为动态模块受随机共因失效影响下的失效概率, 通过将组件由共因导致的失效率作为动态模块对应 Markov 链的转移率计算而得。

步骤5 评估整个系统的可靠性概率。

根据全概率公式得出系统由于某个随机共因事件影响的失效概率为

$$P_{(PMSfail \cdot E_n)} = P_{(PMSfails | E_n)} P_{(E_n)}$$

则系统受所有随机共因事件影响下的失效概率为

$$P_{(PMSfail \cdot PCCF)} = \sum_{n=1}^{2L} P_{(PMSfails | E_n)} P_{(E_n)}$$

因此系统受随机失效影响下的可靠性概率为

$$R = 1 - \sum_{n=1}^{2L} P_{(PMSfails | E_n)} P_{(E_n)} \quad (7)$$

4 案例分析

4.1 分析对象说明

卫星是一个典型的 PMS, 本文以文献 [16] 给出的地球同步轨道卫星为例, 分析其在首次变轨中受 PCCF 影响所得到的任务系统可靠性概率。卫星首次变轨经历太阳捕获、地球捕获、地球指向、远地点点火准备和远地点点火等 5 个阶段, 其所涉及的组件及分系统介绍详见表 1。

系统所涉及的具体参数描述如下:

1) 系统中存在 3 个共因: C_{41} 、 C_{51} 和 C_{52} 。对应失效概率分别为 $P_{C_{41}} = 0.7$, $P_{C_{51}} = 0.6$, $P_{C_{52}} = 0.4$ 。 C_{41} 分别与 C_{51} 、 C_{52} 统计独立; C_{51} 与 C_{52} 统计相关。其中:

$$P\{C_{52} | C_{51}\} = P'_h = 0.6$$

$$P\{C_{52} | \overline{C_{51}}\} = P'_h = 0.5$$

2) 组件的失效率描述。组件的内部失效率

表 1 组件及分系统介绍

Table 1 Introduction of components and subsystems

分系统	单机	符号	简介
姿态控制	计算机	A	含 A_a 和 A_b 共 2 台, 冷备份
	陀螺	B	含 B_a 、 B_b 和 B_c 共 3 台, 3 取 2
姿轨控	数字太阳敏感器	C	含 C_a 、 C_b 、 C_c 和 C_d 共 4 台, 其中仅 C_a 和 C_b 热备份
	红外地球敏感器	D	含 D_a 和 D_b 共 2 台, 热备份
星敏感器	星敏感器	E	含 E_a 、 E_b 和 E_c 共 3 台, 3 取 2
	490 N 发动机	F	1 台 F
推进	10 N 推力器	G	含 G_a 和 G_b 共 2 套, 热备份

分别为: $\lambda_A = 2.44$, $\lambda_B = 1.22$, $\lambda_C = 6.10$, $\lambda_D = 2.44$, $\lambda_F = 1.72$, $\lambda_G = 1.22$ (单位均为 10^{-8} min^{-1}); 组件由于共因所导致的失效率分别为: $\lambda_{41}(A) = 2$, $\lambda_{41}(C) = 3$, $\lambda_{41}(E) = 6$, $\lambda_{41}(G) = 7$, $\lambda_{51}(B) = 1$, $\lambda_{51}(D) = 2$, $\lambda_{51}(F) = 3$, $\lambda_{52}(A) = 5$, $\lambda_{52}(C) = 6$, $\lambda_{52}(E) = 2$, $\lambda_{52}(G) = 4$ (单位均为 10^{-4} min^{-1})。

3) 阶段持续时间。每个阶段的持续时间分别为: $T_1 = 45$, $T_2 = 698$, $T_3 = 35$, $T_4 = 120$, $T_5 = 57$ (单位均为 min)。

4.2 可靠性建模与评估

构建卫星首次变轨过程中的随机共因事件空间如式 (8) 所示:

$$\begin{cases} E_1 = \overline{C_{41}} \cap \overline{C_{51}} \cap \overline{C_{52}} \\ E_2 = C_{41} \cap \overline{C_{51}} \cap \overline{C_{52}} \\ E_3 = \overline{C_{41}} \cap C_{51} \cap \overline{C_{52}} \\ E_4 = \overline{C_{41}} \cap \overline{C_{51}} \cap C_{52} \\ E_5 = C_{41} \cap C_{51} \cap \overline{C_{52}} \\ E_6 = C_{41} \cap \overline{C_{51}} \cap C_{52} \\ E_7 = \overline{C_{41}} \cap C_{51} \cap C_{52} \\ E_8 = C_{41} \cap C_{51} \cap C_{52} \end{cases} \quad (8)$$

计算各共因失效事件的失效概率:

$$\begin{cases} P_{(E_1)} = (1 - P_{C_{41}})(1 - P'_h)(1 - P_{C_{52}}) = 0.09 \\ P_{(E_2)} = P_{C_{41}}(1 - P'_h)(1 - P_{C_{52}}) = 0.21 \\ P_{(E_3)} = (1 - P_{C_{41}})(1 - P_{C_{52}})P'_h = 0.09 \\ P_{(E_4)} = (1 - P_{C_{41}})P_{C_{52}}(1 - P'_h) = 0.048 \\ P_{(E_5)} = P_{C_{41}}(1 - P_{C_{52}})P'_h = 0.21 \\ P_{(E_6)} = P_{C_{41}}P_{C_{52}}(1 - P'_h) = 0.112 \\ P_{(E_7)} = (1 - P_{C_{41}})P_{C_{52}}P'_h = 0.072 \\ P_{(E_8)} = P_{C_{41}}P_{C_{52}}P'_h = 0.168 \end{cases}$$

首先, 采用故障树对所涉及的 5 个任务阶段建模 (详见文献 [16])。根据模块化思想, 将整个系统总结为 8 个独立的子模块, 分别为

$$M_1 = \{A_a, A_b\}$$

$$M_2 = \{B_a, B_b, B_c\}$$

$$M_3 = \{C_c\}$$

$$M_4 = \{C_d\}$$

$$M_5 = \{G_a, G_b\}$$

$$M_6 = \{C_a, C_b\}$$

$$M_7 = \{D_a, D_b\}$$

$$M_8 = \{F\}$$

其中: M_1 为动态模块, 其他为静态模块。各阶段的 BDD 模型详见文献 [16], 并将其简化为系统级的 BDD 模型如图 1 所示, 图中 M_{ji} 表示在第 i 阶

段的 M_j 。

由式(6)计算组件条件的失效概率如表2所示。由系统 BDD 模型可知系统失效所涉及的模块为 M_{24} 、 M_{34} 、 M_{44} 、 M_{54} 、 M_{63} 、 M_{73} 和 M_{85} 。

评估系统在随机共因失效影响下的条件失效概率。其中 $M_2 \sim M_8$ 采用 BDD 求解所得失效表达式如式(9)所示; M_1 的 Markov 模型如图2所示,求解 Markov 模型所得微分方程为式(10),其中 P_{S_i} 代表状态 S_i 的失效概率;系统失效表达式为式(11)。由式(7)、式(9)和式(10)可求得每个模块在共因失效时间影响下的失效概率,如表3所示。

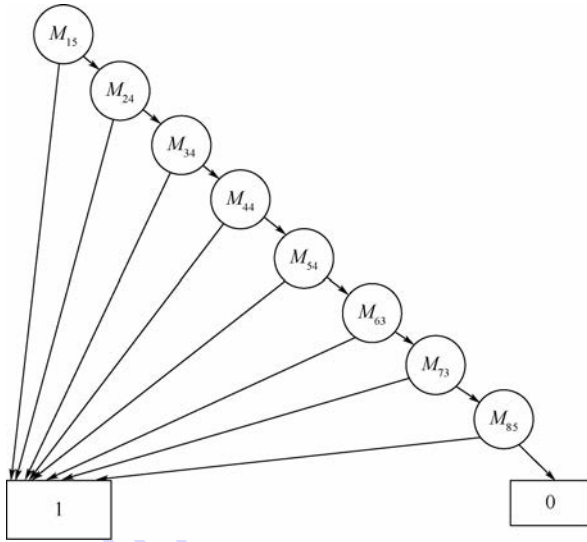


图1 系统 BDD 模型

Fig. 1 BDD model of system

$$\begin{cases} M_2 = B_a B_b + B_b B_c + B_a B_c \\ M_3 = C_c \\ M_4 = C_d \\ M_5 = G_a G_b \\ M_6 = C_a C_b \\ M_7 = D_a D_b \\ M_8 = F \end{cases} \quad (9)$$

$$\begin{cases} \frac{dP_{S_1}(t)}{dt} = -\lambda_{A_a} P_{S_1}(t) \\ \frac{dP_{S_2}(t)}{dt} = \lambda_{A_a} P_{S_1}(t) - \lambda_{A_b} P_{S_2}(t) \\ \frac{dP_{S_3}(t)}{dt} = \lambda_{A_b} P_{S_2}(t) \end{cases} \quad (10)$$

$$P_{(PMS\text{fails} | E_2)} = 1 - \overline{M_{15}} \cdot \overline{M_{24}} \cdot \overline{M_{34}} \cdot \overline{M_{44}} \cdot \overline{M_{54}} \cdot \overline{M_{63}} \cdot \overline{M_{73}} \cdot \overline{M_{85}} \quad (11)$$

由式(11),参考表3的数据可得系统在随机共因失效事件影响下的条件失效概率(保留10位有效数字):

$$P_{(PMS\text{fails} | E_1)} = 0.0001259982$$

$$P_{(PMS\text{fails} | E_2)} = 0.0847337058$$

$$P_{(PMS\text{fails} | E_3)} = 0.0171077073$$

$$P_{(PMS\text{fails} | E_4)} = 0.0097647737$$

$$P_{(PMS\text{fails} | E_5)} = 0.1002784504$$

$$P_{(PMS\text{fails} | E_6)} = 0.0795109022$$

$$P_{(PMS\text{fails} | E_7)} = 0.0265827793$$

$$P_{(PMS\text{fails} | E_8)} = 0.0951443501$$

表2 组件条件失效概率

Table 2 Conditional failure probability of components

模块	组件	$E_1/10^{-5}$	E_2	E_3	$E_4/10^{-5}$	E_5	E_6	E_7	E_8
M_{24}	B_{a4}	1.086	1.086×10^{-5}	1.086×10^{-5}	1.086	1.086×10^{-5}	1.086×10^{-5}	1.086×10^{-5}	1.086×10^{-5}
	B_{b4}	1.068	1.068×10^{-5}	1.068×10^{-5}	1.068	1.068×10^{-5}	1.068×10^{-5}	1.068×10^{-5}	1.068×10^{-5}
	B_{c4}	1.068	1.068×10^{-5}	1.068×10^{-5}	1.068	1.068×10^{-5}	1.068×10^{-5}	1.068×10^{-5}	1.068×10^{-5}
M_{63}	C_{a3}	4.746	4.746×10^{-5}	4.746×10^{-5}	4.746	4.746×10^{-5}	4.746×10^{-5}	4.746×10^{-5}	4.746×10^{-5}
	C_{b3}	4.746	4.746×10^{-5}	4.746×10^{-5}	4.746	4.746×10^{-5}	4.746×10^{-5}	4.746×10^{-5}	4.746×10^{-5}
M_{34}	C_{c4}	5.478	0.0354	5.478×10^{-5}	5.478	0.0354	0.0354	5.478×10^{-5}	0.0354
M_{44}	C_{d4}	5.478	0.0354	5.478×10^{-5}	5.478	0.0354	0.0354	5.478×10^{-5}	0.0354
M_{54}	G_{a4}	1.068	0.0806	1.068×10^{-5}	1.068	0.0806	0.0806	1.068×10^{-5}	0.0806
	G_{b4}	1.068	0.0806	1.068×10^{-5}	1.068	0.0806	0.0806	1.068×10^{-5}	0.0806
M_{73}	D_{a3}	1.891	1.891×10^{-5}	1.891×10^{-5}	1.891	1.891×10^{-5}	1.891×10^{-5}	1.891×10^{-5}	1.891×10^{-5}
	D_{b3}	1.891	1.891×10^{-5}	1.891×10^{-5}	1.891	1.891×10^{-5}	1.891×10^{-5}	1.891×10^{-5}	1.891×10^{-5}
M_{85}	F_5	1.643	1.643×10^{-5}	0.0170	1.643	0.0170	1.643×10^{-5}	0.0170	0.0170

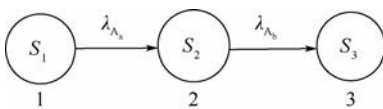


图2 M_1 对应的的 Markov 模型

Fig. 2 Markov model of M_1

由式(11)计算系统可靠性概率为 $R = 0.9323242726$

在不考虑共因影响的情况下,卫星首次变轨的任务成功概率为 0.99988149^[16],但由于卫星在执行任务过程中存在共因事件,尽管随机共因

表3 模块条件失效概率

Table 3 Conditional failure probability of modules

模块	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8
M_{15}	1.000×10^{-8}	0.00987	1.000×10^{-8}	0.00964	0.00987	0.00422	0.00964	0.00422
M_{24}	3.422×10^{-10}	3.422×10^{-10}	3.422×10^{-10}	3.422×10^{-10}	3.422×10^{-10}	3.422×10^{-10}	3.422×10^{-10}	3.422×10^{-10}
M_{63}	2.252×10^{-9}	2.252×10^{-9}	2.252×10^{-9}	2.252×10^{-9}	2.252×10^{-9}	2.252×10^{-9}	2.252×10^{-9}	2.252×10^{-9}
M_{34}	5.478×10^{-5}	0.0354	5.478×10^{-5}	5.478×10^{-5}	0.0354	0.0354	5.478×10^{-5}	0.0354
M_{44}	5.478×10^{-5}	0.0354	5.478×10^{-5}	5.478×10^{-5}	0.0354	0.0354	5.478×10^{-5}	0.0354
M_{54}	1.141×10^{-10}	0.0065	1.141×10^{-10}	1.141×10^{-10}	0.0065	0.0065	1.141×10^{-10}	0.0065
M_{73}	3.576×10^{-10}	3.576×10^{-10}	3.576×10^{-10}	3.576×10^{-10}	3.576×10^{-10}	3.576×10^{-10}	3.576×10^{-10}	3.576×10^{-10}
M_{85}	1.643×10^{-5}	1.643×10^{-5}	0.0170	1.643×10^{-5}	0.0170	1.643×10^{-5}	0.0170	0.0170

失效的概率并不高,但是仍然使得系统的任务成功概率降低了约6.8%,且随着工作时间的累积,其影响效果还会逐渐增大。因此,对于卫星等PMS来说,随机共因事件是一个不容忽视的可靠性影响因素。

5 结论

- 1) 利用贝叶斯理论扩展了PMS中随机共因失效的概率模型,使其具有更大的适用范围。
- 2) 提出了将随机共因失效引入PMS可靠性评估的基本思路,并给出了详细的可靠性评估步骤。
- 3) 给出了分别利用BDD和Markov模型以及全概率公式进行求解的概率模型。
- 4) 通过案例分析,一方面验证了方法的可行性和有效性,另一方面确认了如果不考虑共因失效的影响,可靠性评估结论将过于乐观。

在后续研究中,笔者将在考虑系统组件存在维修活动以及级联失效关系情况下,研究如何对PMS可靠性进行评估,并进一步优化本文提出的建模分析方法,使其更加有效、便捷。

参考文献 (References)

- [1] XING L, AMARI S V. Reliability of phased-mission systems [M]. Berlin: Springer, 2008: 349-368.
- [2] XING L, LEVITIN G. BDD-based reliability evaluation of phased-mission systems with internal/external common-cause failures[J]. Reliability Engineering and System Safety, 2013, 112: 145-153.
- [3] WU X Y, YAN H, LI L. Numerical method for reliability analysis of phased mission system using Markov chains[J]. Communication Statistics-Theory and Method, 2012, 41 (21): 3960-3973.
- [4] WU X Y, WU X Y. Extended object-oriented Petri net model for mission reliability simulation of repairable PMS with common cause failures[J]. Reliability Engineering and System Safety, 2015, 136: 109-119.
- [5] SHRESTHA A, XING L, DAI Y. Reliability analysis of multi-state phased-mission systems with unordered and ordered states [J]. IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans, 2011, 41 (4): 625-636.
- [6] LU J M, WU X Y. Reliability evaluation of generalized phased-mission systems with repairable components [J]. Reliability Engineering and System Safety, 2014, 121: 136-145.
- [7] MO Y, XING L, AMARI S. A multiple-valued decision diagram based method for efficient reliability analysis of non-repairable phased-mission systems[J]. IEEE Transactions on Reliability, 2014, 63 (1): 320-330.
- [8] WANG D, TRIVEDI K S. Reliability analysis of phased-mission system with independent component repairs[J]. IEEE Transactions on Reliability, 2007, 56 (3): 540-551.
- [9] XING L, WANG W. Probabilistic common-cause failures analysis[C]// Proceeding of the 2008 Annual Reliability and Maintainability Symposium. Piscataway, NJ: IEEE Press, 2009: 354-358.
- [10] XING L, BODDU P, SUN Y, et al. Reliability analysis of static and dynamic fault-tolerant systems subject to probabilistic common-cause failures[J]. Journal of Risk and Reliability, 2010, 224 (1): 43-53.
- [11] WANG C, XING L, LEVITIN G. Explicit and implicit methods for probabilistic common-cause failure analysis[J]. Reliability Engineering and System Safety, 2014, 131 (3): 175-184.
- [12] WANG C, XING L, LEVITIN G. Probabilistic common cause failures in phased-mission systems[J]. Reliability Engineering and System Safety, 2015, 144: 53-60.
- [13] OU Y, DUGAN J B. Modular solution of dynamic multi-phase systems[J]. IEEE Transaction on Reliability, 2004, 53 (4): 499-508.
- [14] XING L, AMARI S V. Binary decision diagrams and extensions for system reliability analysis [M]. Boston: John Wiley & Sons, 2015.
- [15] LI S, SI S, DUI H, et al. A novel decision diagrams extension method[J]. Reliability Engineering and System Safety, 2014, 126: 107-115.
- [16] 张华, 宗益燕, 韦锡峰, 等. 地球同步轨道卫星多阶段任务可靠性建模[J]. 航天器环境工程, 2016, 33 (4): 439-445.
ZHANG H, ZONG Y Y, WEI X F, et al. Phased-mission system reliability modeling of geostationary satellite based on mission profile[J]. Spacecraft Environment Engineering, 2016, 33 (4): 439-445 (in Chinese).
- [17] 朱海鹏. 基于BDD的多阶段任务系统可靠性建模分析

[D]. 成都:电子科技大学,2010:37-38.

ZHU H P. Reliability modeling and analysis method for PMS based on BDD[D]. Chengdu: University of Electronic Science and Technology of China,2010:37-38(in Chinese).

[18] RAUZY A. New algorithms for fault tree analysis[J]. Reliability Engineering and System Safety,1993,40(3):203-211.

作者简介:

吴欢 女,硕士研究生。主要研究方向:可靠性/安全性建模、

分析。

焦健 男,博士,讲师。主要研究方向:可靠性/安全性分析、建模、仿真等。

赵廷弟 男,博士,研究员,博士生导师。主要研究方向:系统安全与可靠性工程、系统安全与可靠性设计分析与验证、安全预测与风险评估技术、复杂系统安全与可靠性理论等。

A reliability modeling and analysis method for PMS considering common cause failure

WU Huan, JIAO Jian*, ZHAO Tingdi

(School of Reliability and Systems Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Abstract: Common cause failures (CCFs) in a system destroy the hypothesis that the failures are independent, which may significantly impact the reliability evaluation of the system, especially the phased-mission system (PMS). Aimed at the impact of probabilistic common cause failure (PCCF) on reliability of mission in PMS, this paper discussed the relationship between common cause events and extended the probabilistic model of common cause events using Bayesian theory to make the model fit for different statistical relations including mutually exclusive, s-independent and s-dependent. Moreover, a module-based modeling and analysis method using binary decision diagram (BDD) and Markov model was proposed. First, the fault tree of each phase was constructed. Then, considering CCF, BDD and Markov model were used to deal with the static and dynamic module in PMS respectively. Third, mission reliability was evaluated using total probability law. Finally, a case study of satellite for its orbit transfer was supplied to verify the effectiveness of the method. In addition, the result of this paper was compared with the existing case to analyze the influence of CCFs on mission reliability.

Keywords: phased-mission system (PMS); probabilistic common cause failure (PCCF); binary decision diagram (BDD); Markov model; dynamic

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

DOI: 10.13700/j.bh.1001-5965.2017.0349

多维力加载装置动力学建模及加载试验



侯立果¹, 王丹¹, 安大卫², 郭江真^{1,*}, 陈五一¹, 樊锐¹

(1. 北京航空航天大学 机械工程及自动化学院, 北京 100083;

2. 航空工业成都飞机工业(集团)有限责任公司 技术中心, 成都 610092)

摘 要: 设计了一种基于 6-PUS 并联机构的多维力加载装置, 可对目标施加多维时变载荷。首先, 根据向量叠加原理和牛顿迭代法, 分别推导了加载装置的运动学逆解与正解。其次, 基于旋量理论, 推导加载装置的速度和静力雅可比矩阵, 建立动平台与各支链的运动学和静力学映射。再次, 根据虚功原理, 建立加载装置的动力学模型, 并且通过对比 Adams 和 MATLAB 模型的仿真结果, 验证动力学模型的正确性。最后, 设计并完成多维力加载试验, 验证加载装置的多维力加载能力, 为后续的数控机床加载试验提供新的加载手段和理论基础。

关键词: 并联机构; 运动学; 动力学; 多维力加载装置; 数控机床

中图分类号: TH122

文献标识码: A **文章编号:** 1001-5965(2018)05-1095-07

数控机床是装备制造业的工作母机, 其可靠性技术目前已成为制约行业发展的关键共性技术^[1]。现阶段检测数控机床可靠性的主要方法是“经时试验”, 即使用该机床长时间切削金属坯料, 使机床主轴承受切削力, 以此发现机床长期运行时所发生的故障。但是, 该方法耗费大量的原料和刀具, 显著增加了测试成本。所以, 需要研制一种可向机床主轴施加多维载荷的加载装置, 以改进传统测试方法。

并联机构由于其结构紧凑、承载能力强、积累误差小和动态性能好等特点在工业领域被广泛使用^[2]。其动平台具备空间多自由度, 故可以将其应用于多维力加载领域。关于并联加载装置的研究, 最早可追溯到 1962 年 Gough 和 Whitehall 发明的 6 自由度并联机器人^[3], 将其应用于对轮胎的加载中, 测试轮胎在不同载荷下的性能。徐彬^[4]采用 Stewart 平台结构的形式, 对大型旋挖钻

机动臂加载性能参数进行了分析。刘少欣等^[5]对 6-UPS 并联加载装置的动力学模型进行了简化。王博和黄其涛^[6]提出了一种正交并联机构的多轴加载试验系统, 用以模拟对接机构组合体实际应用中受到的复杂载荷, 是并联机构一种新的应用形式。Stokes 等^[7]基于 6 自由度并联机构设计一种针对脊柱的加载装置, 在末端安装多维力传感器, 对动物脊柱进行了刚度测量。Nierenberger 和 Flohic 等^[8-9]将 Stewart 平台用作多维力材料加载试验机, 并对混凝土试验件进行拉力-剪力-剪切力二维加载试验。Guo 等^[10-12]同样基于 Stewart 平台, 设计多维力加载试验机, 对试验机的输出载荷进行检测。但是, 上述研究采用的 Stewart 平台的驱动电机位于运动支链中, 随连杆一起运动, 使机构运动部件的质量增大, 对机构的动态特性产生不利影响。相比之下, 6-PSS 并联机构由于其驱动电机固定, 可有效减小运动部件

收稿日期: 2017-05-23; 录用日期: 2017-06-30; 网络出版时间: 2017-09-22 15:07

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20170922.1507.008.html

基金项目: 一院高校联合创新基金 (CALT201608)

* 通信作者: E-mail: jzguo@buaa.edu.cn

引用格式: 侯立果, 王丹, 安大卫, 等. 多维力加载装置动力学建模及加载试验[J]. 北京航空航天大学学报, 2018, 44(5): 1095-1101. HOU L G, WANG D, AN D W, et al. Dynamic modeling and loading experiment of multi-dimensional loading device[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 1095-1101 (in Chinese).

的质量,机构的动态特性好,便于在多维加载过程中实时跟随目标的空间运动。王洪瑞等^[13]建立了6-PSS并联机器人的动力学模型;孙小勇等^[14]对6-PSS并联机器人的参数进行了优化设计;Wang等^[15]基于Hexaglide型并联机构设计多维力加载装置,并对其刚度分布进行分析,但是并未对加载装置进行多维力加载试验。

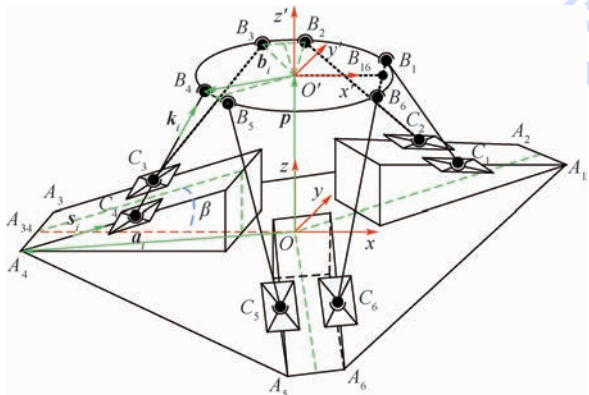
本文提出一种基于6-PUS并联机构的多维力加载装置,可对目标提供多维载荷。首先根据向量叠加原理和牛顿迭代法分别介绍了6-PUS并联机构的逆、正解运算,并由旋量理论建立了机构的运动学和静力学模型;然后根据虚功原理建立了机构的动力学模型并进行仿真验证;最后对多维力加载装置进行多维力加载试验。

1 6-PUS 并联机构的运动学建模

1.1 运动学逆解

如图1所示,6-PUS并联机构由动平台、静平台及6根完全相同的支链组成。其中,动平台复合球铰中心 $B_i (i = 1, 2, \dots, 6)$ 的分布半径为 r_p , 6个球铰链分为3组,沿圆周均布,每组之间夹角为 120° ; 6个直线驱动及静平台复合虎克铰 $C_i (i = 1, 2, \dots, 6)$ 同样分为3组在静平台上沿圆周均布,每组驱动的安装平面倾角为 β 。将滑块置于零位,每组两驱动的中点的分布半径为 r_b , 各支链中的连杆长度为 L 。

选任一支链,使用向量叠加描述该运动链得 $a_i + d_i s_i + L_i k_i = p + R b_i'$ (1) 式中: $a_i = \overrightarrow{OA_i}$; d_i 为各运动支链中移动副的驱动量; L_i 为定长连杆的长度; s_i 为静坐标系中移动副的方向向量; k_i 为静坐标系中沿连杆方向的单位向量; p 和 R 分别为动坐标系 $O' \{x', y', z'\}$ 相



对于静坐标系 $O \{x, y, z\}$ 的原点位置和坐标系旋转矩阵;在动坐标系中定义 $b_i' = \overrightarrow{O'B_i}$ 。

令 $p + R b_i' - a_i = h_i$, 根据机构构型和运动的连续性求解式(1)得

$$d_i = h_i s_i - \sqrt{(h_i \cdot s_i)^2 - (h_i^2 - L_i^2)} \quad (2)$$

1.2 运动学正解

与运动学逆解相反,多维力加载装置的运动学正解是已知6个运动支链的电机驱动量,求解机构动平台位姿的过程,即已知驱动量 $d_i (i = 1, 2, \dots, 6)$, 求解动平台中心点 O' 的位置向量 $p = [p_x, p_y, p_z]^T$ 和姿态向量 $\Theta = [\theta_x, \theta_y, \theta_z]^T$, 其中向量 p 和 Θ 的3个分量分别是动平台沿 x, y, z 轴的位移和绕其轴线的转角。

将动平台的位姿表示为 $Q = [p^T \ \Theta^T]^T$, 可以根据姿态向量求解出旋转矩阵 $R = f(\Theta)$, 由运动学反解可知

$$f_i(Q) = (p + R b_i' - a_i - d_i s_i)^2 - L_i^2 = 0 \quad (3)$$

式中:旋转矩阵 R 用四元数表示,相比欧拉角,可有效避免奇异问题。通常,可将四元数定义为 $A = a_0 + a_1 i + a_2 j + a_3 k$, ($a_i \in \mathbf{R}, i = 0, 1, 2, 3$), 且

$$f_7(Q) = a_0^2 + a_1^2 + a_2^2 + a_3^2 - 1 = 0 \quad (4)$$

则 $Q = [p \ A]^T, A = [a_0 \ a_1 \ a_2 \ a_3]^T$ 。

将 R 代入式(3),再与式(4)相加可得7个方程,最后利用牛顿迭代法即可求解得出位置向量 Q 的7个分量未知数。求解过程如图2所示,首先确定动平台初始位姿 Q_0 , 然后用牛顿迭代法求出迭代位姿 $Q_n (n$ 为迭代次数, $n \geq 0)$, 当由末端迭代位姿 Q_n 反解计算出的杆长与已知杆长的差值向量 $\Delta f = [f_1 \ f_2 \ \dots \ f_7]$ 的模满足 $|f(Q_n)| < \xi = 10^{-10}$ m 时,即可认为此时的迭代位姿 Q_n 可以作为当前动平台的实际位姿向量 Q , 迭代结束。

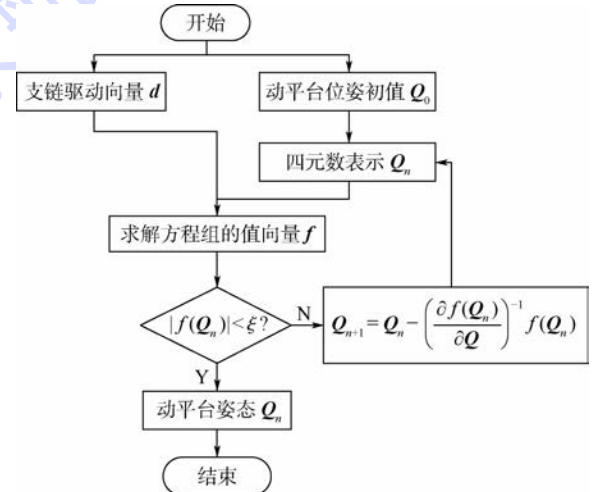


图1 多维力加载装置结构及坐标系定义示意图
Fig.1 Schematic of structure and coordinate system definition of tested multi-dimensional loading device

图2 机构运动学正解流程图
Fig.2 Flowchart of forward kinematics of mechanism

2 运动学与静力学分析

2.1 速度及静力雅可比矩阵分析

任取一条运动支链作为研究对象,如图3所示。将动平台中心点的速度旋量用 Plücker 坐标表示为 $T = [v^T \ \omega^T]^T$, 其中 v 和 ω 分别表示动平台质心点相对静坐标系的平移速度和旋转角速度。

由旋量法可得

$$T = \dot{d}_i \mathcal{S}_{i1} + \dot{\theta}_{i2} \mathcal{S}_{i2} + \dot{\theta}_{i3} \mathcal{S}_{i3} + \dot{\theta}_{i4} \mathcal{S}_{i4} + \dot{\theta}_{i5} \mathcal{S}_{i5} + \dot{\theta}_{i6} \mathcal{S}_{i6} \quad (5)$$

式中: $\mathcal{S}_{i1} = [0^T \ s_i^T]^T$; $\mathcal{S}_{ij} = [k_{ij}^T \ (c_i \times k_{ij})^T]^T$ ($j = 2, 3$), $\mathcal{S}_{ij} = [k_{ij}^T \ (b_i \times k_{ij})^T]^T$ ($j = 4, 5, 6$); $c_i = \overrightarrow{O'C_i}$; $b_i = \overrightarrow{O'B_i}$, k_{ij} 和 θ_{ij} ($j = 2, 3, \dots, 6$) 分别为转动副旋转轴方向向量和旋转速率, 且 $k_{i6} = k_i$, s_i 和 d_i 分别为移动副的方向向量和移动速率。

将式(5)的两端同时与单位速度旋量 $\mathcal{S}_i = [k_i^T \ (c_i \times k_i)^T]^T$ 取互易积, 且 $\mathcal{S}_i \circ \mathcal{S}_{ij} = 0$ ($j = 2, 3, \dots, 6$), 整理为矩阵形式可得

$$q = J_v x_p \quad (6)$$

式中: 动平台质心速度 $x_p = [v_p^T \ \omega_p^T]^T$; 滑块移动速度 $q = [d_1 \ d_2 \ \dots \ d_6]^T$; 速度雅可比矩阵 $J_v = [J_{v1}^T \ J_{v2}^T \ \dots \ J_{v6}^T]^T_{6 \times 6}$, $J_{vi} = \left[\begin{matrix} k_i^T & (c_i \times k_i)^T \\ k_i^T s_i & k_i^T s_i \end{matrix} \right]$ ($i = 1, 2, \dots, 6$)。

由速度和静力学对偶关系可知:

$$F = J_f \tau = J_v^T \tau \quad (7)$$

式中: $F = [f^T \ m^T]^T$ 为施加在动平台质心的外力载荷 f 为外力, m 为力矩; $\tau = [\tau_1 \ \tau_2 \ \dots \ \tau_6]^T$ 为移动副滑块驱动力向量; J_f 为力雅可比矩阵。

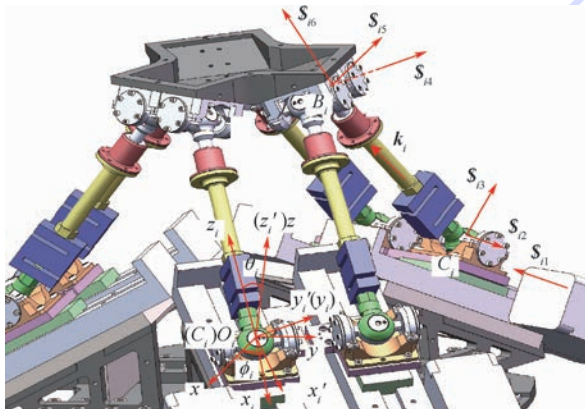


图3 支链坐标系及速度旋量分布示意图

Fig.3 Schematic of branched-chain coordinate system and velocity screw distribution

2.2 速度分析

选取任一定长连杆为研究对象, 以 C_i 为原点, 建立如图3所示的支链坐标系 $O_i \{x_i, y_i, z_i\}$, 其初始状态 x 、 y 和 z 轴分别与静坐标系的3个坐标轴平行, 故连杆方向向量可使用 ZYZ 欧拉角表示, 即支链坐标系 O_i 先绕静坐标系 z 轴旋转角 ϕ_i , 再绕新坐标系 O_i' 的 y_i' 轴旋转角度 θ_i , 此时 $k_i = z_i$ 。

对式(1)的右侧求导可知:

$$v_{bi} = \frac{d}{dt}(p + Rb_i) = v_p + \omega_p \times b_i = J_{bi} \dot{x}_p \quad (8)$$

式中: $J_{bi} = [E_{3 \times 3} \ -b_i \times]$ 。

将 v_{bi} 通过坐标变换转换至支链坐标系得

$${}^i v_{bi} = {}^i R_A v_{bi} = {}^i R_A J_{bi} \dot{x}_p = {}^i J_{bi} \dot{x}_p \quad (9)$$

式中: ${}^i R_A = {}^A R_i$, ${}^A R_i$ 为支链坐标系相对静坐标系的旋转矩阵; ${}^i J_{bi} = [{}^i J_{bix}^T \ {}^i J_{biy}^T \ {}^i J_{biz}^T]^T$ 。

对式(1)等号左侧求导且将其转化至支链坐标系中得

$${}^i v_{bi} = \dot{d}_i {}^i s_i + L_i {}^i \omega_i \times {}^i k_i \quad (10)$$

对式(10)的两端同时点乘 ${}^i k_i$, 可得

$$\dot{d}_i = \frac{{}^i v_{biz}}{{}^i s_i \cdot {}^i k_i} = \frac{{}^i v_{biz}}{{}^i s_i^T {}^i k_i} = \frac{{}^i J_{biz}^T \dot{x}_p}{{}^i s_i^T {}^i k_i}$$

将驱动力整理为矩阵形式为

$$q = J_p \dot{x}_p \quad (11)$$

$$\text{式中: } J_p = \left[\begin{matrix} {}^1 J_{b1z}^T & {}^2 J_{b2z}^T & \dots & {}^6 J_{b6z}^T \\ \frac{1}{s_1^T k_1} & \frac{2}{s_2^T k_2} & \dots & \frac{6}{s_6^T k_6} \end{matrix} \right]^T$$

对式(10)两端同时又乘 ${}^i k_i$, 由 ${}^i \omega_i \cdot {}^i k_i = 0$ 可得连杆角速度在支链坐标系中的表示为

$${}^i \omega_i = \frac{{}^i k_i \times {}^i v_{bi} - \dot{d}_i {}^i k_i \times {}^i s_i}{L_i} = \frac{\left({}^i \Omega_{ki} {}^i J_{bi} - \frac{({}^i k_i \times {}^i s_i) \cdot {}^i J_{biz}}{s_i^T k_i} \right) \dot{x}_p}{L_i}$$

式中: ${}^i \Omega_{ki} = {}^i k_i \circ$ 。

由滑块的质心向量 $r_{ci} = a_i + d_i s_i$ 可得其质心速度为

$$v_{ci} = \dot{d}_i s_i = \frac{s_i \cdot J_{biz}}{s_i^T k_i} \dot{x}_p \quad (12)$$

同理, 由连杆的质心向量 $r_{li} = a_i + d_i s_i + L_i k_i$ 可得连杆的质心速度, 将其转换至支链坐标系:

$${}^i v_{li} = \dot{d}_i s_i + \frac{L_i}{2} {}^i \omega_i \times {}^i k_i = \left\{ \frac{{}^i s_i \cdot J_{biz}}{s_i^T k_i} - \frac{{}^i \Omega_{ki}^2 {}^i J_{bi} - [{}^i k_i ({}^i k_i \cdot {}^i s_i) - {}^i s_i] \cdot J_{biz}}{2 s_i^T k_i} \right\} \dot{x}_p \quad (13)$$

使用 Plucker 坐标表示滑块的质心速度 $\dot{\mathbf{x}}_{ci}$ 和连杆质心速度 $\dot{\mathbf{x}}_{li}$,

$$\dot{\mathbf{x}}_{ci} = [\mathbf{v}_{ci}^T \quad \mathbf{0}^T]^T = \mathbf{J}_{ci} \dot{\mathbf{x}}_p \quad (14)$$

$$\dot{\mathbf{x}}_{li} = [{}^i \mathbf{v}_{li}^T \quad {}^i \boldsymbol{\omega}_i^T]^T = {}^i \mathbf{J}_{li} \dot{\mathbf{x}}_p = [{}^i \mathbf{J}_{li1}^T \quad {}^i \mathbf{J}_{li2}^T]^T \dot{\mathbf{x}}_p \quad (15)$$

式中:

$${}^i \mathbf{J}_{li1} = \frac{{}^i \mathbf{s}_i {}^i \mathbf{J}_{biz}}{{}^i \mathbf{s}_i^T {}^i \mathbf{k}_i} - \frac{{}^i \boldsymbol{\Omega}_{ki}^2 {}^i \mathbf{J}_{bi} - [{}^i \mathbf{k}_i ({}^i \mathbf{k}_i \cdot {}^i \mathbf{s}_i) - {}^i \mathbf{s}_i] \frac{{}^i \mathbf{J}_{biz}}{{}^i \mathbf{s}_i^T {}^i \mathbf{k}_i}}{2}$$

$$\mathbf{J}_{li2} = \frac{{}^i \boldsymbol{\Omega}_{ki} {}^i \mathbf{J}_{bi} - \frac{({}^i \mathbf{k}_i \times {}^i \mathbf{s}_i) {}^i \mathbf{J}_{biz}}{{}^i \mathbf{s}_i^T {}^i \mathbf{k}_i}}{L_i}$$

$$\mathbf{J}_{ci} = \begin{bmatrix} {}^i \mathbf{s}_i {}^i \mathbf{J}_{biz} & \mathbf{0}_{6 \times 3} \\ {}^i \mathbf{s}_i^T {}^i \mathbf{k}_i & \end{bmatrix}^T$$

2.3 加速度分析

对式(10)求导且两端同时点乘 ${}^i \mathbf{k}_i$ 可得滑块的加速度在支链坐标系中为

$$\ddot{d}_i = \frac{{}^i \dot{\mathbf{v}}_{bi} \cdot {}^i \mathbf{k}_i + L_i {}^i \boldsymbol{\omega}_i^2}{{}^i \mathbf{s}_i \cdot {}^i \mathbf{k}_i} \quad (16)$$

对式(10)求导且两端同时又乘 ${}^i \mathbf{k}_i$ 可得连杆角加速度在支链坐标系中为

$${}^i \dot{\boldsymbol{\omega}}_i = \frac{{}^i \mathbf{k}_i \times {}^i \dot{\mathbf{v}}_{bi} - \ddot{d}_i {}^i \mathbf{k}_i \times {}^i \mathbf{s}_i}{L_i} \quad (17)$$

对式(12)求导得滑块质心加速度向量为

$$\ddot{\mathbf{v}}_{ci} = \ddot{d}_i {}^i \mathbf{s}_i \quad (18)$$

对式(13)求导得连杆质心加速度向量,并将其转换至支链坐标系得

$${}^i \dot{\mathbf{v}}_{li} = \ddot{d}_i {}^i \mathbf{s}_i + \frac{L_i}{2} {}^i \dot{\boldsymbol{\omega}}_i \times {}^i \mathbf{k}_i + \frac{L_i}{2} {}^i \boldsymbol{\omega}_i \times ({}^i \boldsymbol{\omega}_i \times {}^i \mathbf{k}_i) \quad (19)$$

3 动力学分析

3.1 动力学方程

由虚功原理可知,多维力加载装置各部件所受广义主动力及广义惯性力合力所做的功为零。

静坐标系下动平台中心点广义主动力和广义惯性力的合力(包含力 \mathbf{f}_p 及力矩 \mathbf{n}_p)为

$$\mathbf{F}_p = \begin{bmatrix} \mathbf{f}_p \\ \mathbf{n}_p \end{bmatrix} = \begin{bmatrix} \mathbf{f}_e + m_p \mathbf{g} - m_p \dot{\mathbf{v}}_p \\ \mathbf{n}_e - {}^A \mathbf{I}_p \dot{\boldsymbol{\omega}}_p - \boldsymbol{\omega}_p \times ({}^A \mathbf{I}_p \boldsymbol{\omega}_p) \end{bmatrix}$$

式中: \mathbf{f}_e 和 \mathbf{n}_e 分别为动平台中心点外力和外力矩; m_p 为动平台质量; ${}^A \mathbf{I}_p$ 为动平台在静坐标系下的转动惯量。

同理,静坐标系下各滑块质心点的广义主动力和广义惯性力合力(包含力 \mathbf{f}_{ci} 及力矩 \mathbf{n}_{ci})为

$$\mathbf{F}_{ci} = \begin{bmatrix} \mathbf{f}_{ci} \\ \mathbf{n}_{ci} \end{bmatrix} = \begin{bmatrix} m_{ci} \mathbf{g} - m_{ci} \dot{\mathbf{v}}_{ci} \\ \mathbf{0} \end{bmatrix}$$

式中: m_{ci} 为滑块质量。同理,支链坐标系下各连杆质心点的广义主动力和广义惯性力的合力(包含力 \mathbf{f}_{li} 及力矩 \mathbf{n}_{li})表示为

$$\mathbf{F}_{li} = \begin{bmatrix} \mathbf{f}_{li} \\ \mathbf{n}_{li} \end{bmatrix} = \begin{bmatrix} m_{li} \mathbf{R}_A \mathbf{g} - m_{li} \dot{\mathbf{v}}_{li} \\ -{}^i \mathbf{I}_{li} \dot{\boldsymbol{\omega}}_i - {}^i \boldsymbol{\omega}_i \times ({}^i \mathbf{I}_{li} \boldsymbol{\omega}_i) \end{bmatrix}$$

式中: m_{li} 为连杆质量; ${}^i \mathbf{I}_{li}$ 为支链坐标系下连杆的转动惯量。

由虚功原理得

$$\delta \mathbf{q}^T \boldsymbol{\tau} + \delta \mathbf{x}_p^T \mathbf{F}_p + \sum_{i=1}^6 \delta \mathbf{x}_{ci}^T \mathbf{F}_{ci} + \sum_{i=1}^6 \delta {}^i \mathbf{x}_{li}^T \mathbf{F}_{li} = 0 \quad (20)$$

当各构件运动时间相同时, $\delta \mathbf{q} = \mathbf{J}_p \delta \mathbf{x}_p$, $\delta \mathbf{x}_{ci} = \mathbf{J}_{ci} \delta \mathbf{x}_p$, $\delta {}^i \mathbf{x}_{li} = {}^i \mathbf{J}_{li} \delta \mathbf{x}_p$, 将此式(3)代入式(20)得

$$\boldsymbol{\tau} = -\mathbf{J}_p^{-T} \left(\mathbf{F}_p + \sum_{i=1}^6 \mathbf{J}_{ci}^T \mathbf{F}_{ci} + \sum_{i=1}^6 {}^i \mathbf{J}_{li}^T \mathbf{F}_{li} \right) \quad (21)$$

3.2 动力学仿真验证

首先,建立 Adams 理论模型,并结合模型的质量参数在 MATLAB 中推导多维力加载装置的动力学方程,建立动力学方程;其次,同时设定 Adams 和 MATLAB 中的机构模型的动平台质心运动轨迹如式(22)所示:

$$\mathbf{x}_p = [40 \cos(\pi t), 40 \sin(\pi t), 5t, 20 \cos(\pi t), 20 \sin(\pi t), 30 \sin(\pi t)]^T \quad (22)$$

则其质心速度和加速度向量也可求出;再次,同时设定 Adams 和 MATLAB 中的机构模型的动平台质心点所受外力为沿 z 轴负方向 1 000 N;最后,在 Adams 中测量 6 个滑块的驱动力,与 MATLAB 计算得出的驱动力进行比较,并计算二者之间的误差,如图 4 所示。由图 4 可知,Adams 机构理论模型与 MATLAB 动力学模型误差维持在 2 N 以内,小于驱动力的 0.5%,支链 2、3 和 5 在 0.04 s 时误差达到 4 N,但仍小于驱动力的 1%,验证了本文运动学及动力学分析的正确性。

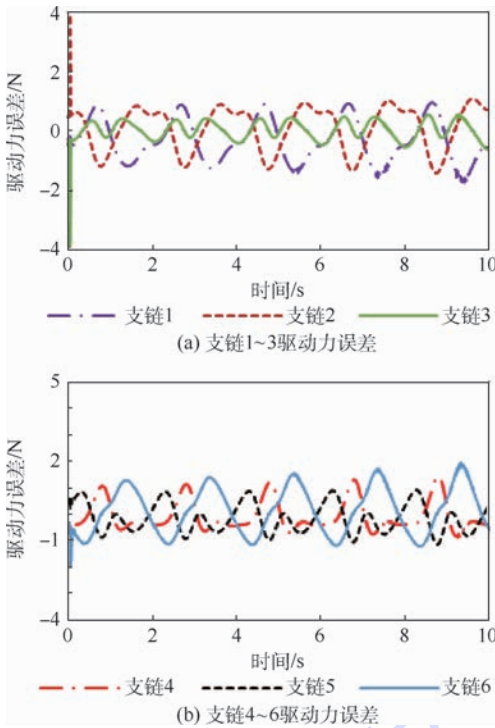


图4 Adams 与 MATLAB 模型所得各支链驱动力的误差

Fig.4 Deviation of branched-chain actuation forces derived from Adams and MATLAB model

4 多维力加载试验

进行多维力加载试验时,将多维力加载装置放置在三轴联动龙门式机床试验平台上,动平台中心与机床主轴连接,主轴处于静止状态,改变动平台的位姿,其与主轴的相对运动会对应产生作用力,实现对主轴的多维力加载。具体加载过程如下:第一,设定期望动平台输出的多维力值;第二,由光栅尺反馈各支链的驱动量,正解得机构动平台位姿;第三,根据动平台位姿和动力学方程求解得出施加期望多维力所需要的各支链驱动力,并由此推导出各传感器期望力;第四,使用模糊PID控制算法^[16]对各支链进行力闭环控制,使动平台输出期望多维力。

多维力加载试验分为2部分:六维三角波载荷和六维正弦载荷,分别代表对机床主轴匀速加载和变加速度加载2种形式。试验采用“双十指标”评估加载性能,即实际加载力与参考加载力误差小于10%,相位误差小于 10° 。通过比较各方向实际载荷和参考载荷间的跟随误差,分析和评估多维力加载装置的实时性和准确性。

1) 六维三角波载荷加载

同时在主轴上施加幅值为200 N的三维三角

波力和幅值为 $10 \text{ N} \cdot \text{m}$ 三维三角波力矩,得到理论力、理论力矩及加载误差如图5所示。

观察图5可知,多维力加载装置对主轴施加的六维三角波载荷中,沿 x 、 y 和 z 轴的加载力误差最大值分别为4.680、2.636和5.264 N,基本处于参考加载力最大幅值的5%以内。另外,沿三轴加载力矩误差最大值分别为0.652、0.515和 $0.321 \text{ N} \cdot \text{m}$,小于参考加载力矩最大幅值的10%,满足加载需要。

2) 六维正弦载荷加载

同时在主轴上施加幅值为200 N的三维正弦力和幅值为 $10 \text{ N} \cdot \text{m}$ 的三维正弦力矩,得到理论力、理论力矩及加载误差如图6所示。

由图6可知,多维力加载装置对主轴施加的六维正弦载荷中,沿 x 、 y 和 z 轴加载力误差最大值分别为18.291、22.417和19.201 N,基本维持在参考载荷的10%以内。另外,沿三轴的加载力矩误差最大值分别为1.583、2.062和 $1.607 \text{ N} \cdot \text{m}$,在参考加载力矩的10%以内,满足加载需要。

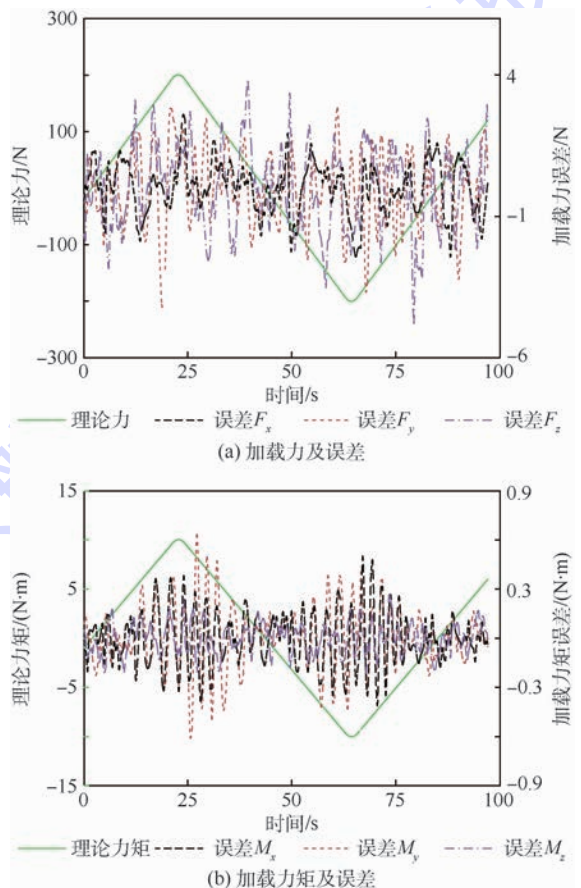


图5 六维三角波载荷加载误差

Fig.5 Loading errors along six axes with triangular wave load

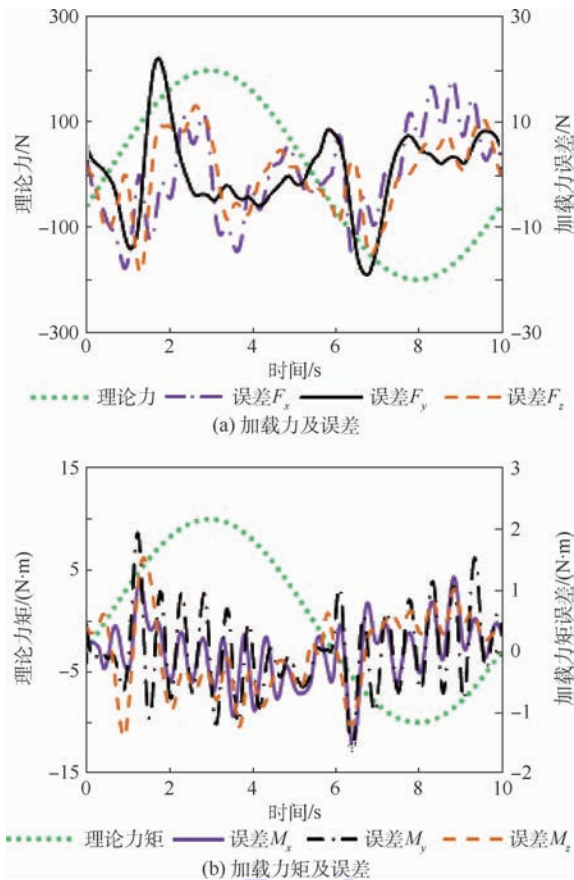


图6 六维正弦载荷加载误差

Fig. 6 Loading errors along six axes with sine load

5 结论

本文提出一种基于6-PUS并联机构的多维力加载装置,可对机床主轴施加多维载荷。建立了加载装置的运动学、静力学以及动力学模型,并进行了动力学仿真验证和多维力加载试验,得出以下结论:

1) 采用虚功原理可对6-PUS并联机构进行准确的动力学建模。用MATLAB建立动力学模型,并在Adams中建立结构相同的虚拟样机,通过给定相同的动平台运动轨迹和外载荷,对比Adams和MATLAB仿真所得的驱动力,误差小于1%,验证了动力学建模的正确性。

2) 6-PUS并联机构具备对目标进行多维力加载的能力,加载试验结果表明,该机构可对目标施加三维三角波和正弦载荷信号,最大跟随误差均小于10%,证明其可以对目标进行匀速和变加速多维力加载,为数控机床加载试验提供新的加载手段和理论基础。

参考文献 (References)

[1] 杨兆军,陈传海,陈菲,等. 数控机床可靠性技术的研究进展

[J]. 机械工程学报,2013,49(20):130-139.

YANF Z J, CHEN C H, CHEN F, et al. NC machine tool reliability technology database applications[J]. Journal of Mechanical Engineering, 2013, 49(20):130-139 (in Chinese).

[2] 郭江真,王丹,樊锐,等. 3PRS/UPS冗余驱动并联机器人刚度特性分布[J]. 北京航空航天大学学报,2014,40(4):500-506.

GUO J Z, WANG D, FAN R, et al. Stiffness characteristics distribution of 3PRS/UPS parallel manipulator with actuation redundancy[J]. Journal of Beijing University of Aeronautics and Astronautics, 2014, 40(4):500-506 (in Chinese).

[3] GOUGH V E, WHITEHALL S G. Universal tyre test machine [C]//Proceedings of the FISITA Ninth International Technical Congress. London: Institution of Mechanical Engineers, 1962: 117-137.

[4] 徐彬. 基于并联机构的多维力加载试验系统研究[D]. 哈尔滨: 哈尔滨工业大学, 2013: 10-40.

XU B. Research on multi-DOF loading system based on parallel mechanism[D]. Harbin: Harbin Institute of Technology, 2013: 10-40 (in Chinese).

[5] 刘少欣,王丹,陈五一. 6-UPS并联多维力加载装置的动力学建模及简化[J]. 机床与液压,2014,42(19):5-8.

LIU S X, WANG D, CHEN W Y. Dynamics modeling and simplification of a 6-UPS parallel multi-dimensional loading device [J]. Machine Tool and Hydraulics, 2014, 42(19):5-8 (in Chinese).

[6] 王博,黄其涛. 正交并联多自由度加载试验系统研究[J]. 机电工程,2016,33(11):1334-1338.

WANG B, HUANG Q T. Loading test system of orthogonal parallel multi degree of freedom [J]. Mechanical and Electrical Engineering, 2016, 33(11):1334-1338 (in Chinese).

[7] STOKES I A, GARDMEMORSE M, CHURCHILL D, et al. Measurement of a spinal motion segment stiffness matrix [J]. Journal of Biomechanics, 2002, 35(4):517-521.

[8] NIERENBERGER M, PONCELET M, PATTOFATTO S, et al. Multiaxial testing of materials using a Stewart platform: Case study of the Nooru-Mohamed test [J]. Experimental Techniques, 2012, 38(2):74-83.

[9] FLOHIC J L, PARPOIL V, BOUISSOU S, et al. A 3D displacement control by digital image correlation for the multiaxial testing of materials with a Stewart platform [J]. Experimental Mechanics, 2014, 54(5):817-828.

[10] GUO J Z, WANG D, FAN R, et al. Development of a material testing machine with multi-dimensional loading capability [J]. Journal of Advanced Mechanical Design Systems & Manufacturing, 2016, 10(2):JAMDSM0017.

[11] GUO J Z, FAN R, WANG D, et al. Development of multidimensional loading material testing machine [J]. Materials Science Forum, 2013, 770:164-169.

[12] 樊锐,李茜,王丹. 6PUS并联机构的运动学整机标定[J]. 北京航空航天大学学报,2016,42(5):871-877.

FAN R, LI X, WANG D. Integral kinematics calibration of 6PUS parallel mechanism [J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(5):871-877 (in Chinese).

- [13] 王洪瑞,马丽,肖金壮.一种 6-PSS 并联机器人动力学模型的建立[J].中国高新技术企业,2010(10):19-21.
WANG H R, MA L, XIAO J Z. A 6-PSS parallel robot dynamic model[J]. China High-Tech Enterprises, 2010(10):19-21 (in Chinese).
- [14] 孙小勇,郑彬,鲍捷,等.高速 6-PSS 并联机器人参数优化设计[J].农业机械学报,2015,46(5):372-378.
SUN X Y, ZHENG B, BAO J, et al. Parameter optimization design of high speed 6-PSS parallel robot[J]. Transactions of the Chinese Society of Agricultural Machinery, 2015, 46(5):372-378 (in Chinese).
- [15] WANG D, FAN R, CHEN W. Stiffness analysis of a Hexaglide parallel loading mechanism[J]. Mechanism & Machine Theory, 2013, 70(6):454-473.
- [16] 樊锐,刘力军,王丹,等.6-PUS 并联加载机构模糊 PID 力控制系统设计[J].航空制造技术,2016,513(18):34-39.
FAN R, LIU L J, WANG D, et al. Design of fuzzy PID force control system based on 6-PUS parallel machine[J]. Aeronautical Manufacturing Technology, 2016, 513(18):34-39 (in Chinese).

作者简介:

侯立果 男,硕士研究生。主要研究方向:并联机构。

郭江真 男,博士。主要研究方向:多维力加载技术。

Dynamic modeling and loading experiment of multi-dimensional loading device

HOU Ligu¹, WANG Dan¹, AN Dawei², GUO Jiangzhen^{1,*}, CHEN Wuyi¹, FAN Rui¹

(1. School of Mechanical Engineering and Automation, Beijing University of Aeronautics and Astronautics, Beijing 100083, China;

2. Technical Center, AVIC Chengdu Aircraft Industrial (Group) Co., Ltd., Chengdu 610092, China)

Abstract: A multi-dimensional loading device based on a 6-PUS parallel mechanism was designed to apply time-varying multi-dimensional loads to target. Firstly, according to the principle of vector superposition and Newton-Raphson method, the inverse and forward kinematics of loading device were analyzed. Secondly, based on the screw theory, the velocity and static force Jacobian matrix, which represent the kinematic and static mapping between the moving platform and actuators, were derived. Thirdly, according to the principle of virtual work, the dynamic model of loading device was established and validated by comparing the actuation forces derived from Adams and MATLAB simulation models. Finally, the multi-dimensional loading tests were designed and completed to verify the multi-dimensional loading ability of the device. The loading device provides a new loading method and theoretical basis for the loading tests of numerical control machine tools.

Keywords: parallel mechanism; kinematics; dynamics; multi-dimensional loading device; numerical control machine tool

Received: 2017-05-23; Accepted: 2017-06-30; Published online: 2017-09-22 15:07

URL: kns.cnki.net/kcms/detail/11.2625.V.20170922.1507.008.html

Foundation item: CALT University Innovation Fund (CALT201608)

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

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

DOI: 10.13700/j.bh.1001-5965.2017.0303

来流导致的高超声速气动热不确定度量化分析



张伟, 王小永, 于剑, 阎超*

(北京航空航天大学 航空科学与工程学院, 北京 100083)

摘 要: 通常的 CFD 计算都是确定性的, 然而复杂工程数值模拟中必然存在误差与不确定度, 分析与辨识其不确定度来源, 对不确定度进行量化分析, 对数值模拟可信度评估有重要意义。在高超声速飞行器气动热计算中, 为获得更加可靠的气动热数据和鉴定影响气动热预测的关键因素, 对返回舱开展了气动热不确定度量化分析和敏感性分析。首先选取来流速度、来流温度、壁面温度和来流密度 4 个不确定性输入变量, 并且假定来流速度变化范围为 $\pm 120 \text{ m/s}$ ($\pm 2\%$), 来流温度、壁面温度和来流密度变化范围为 $\pm 10\%$ 。然后采用拉丁超立方抽样法生成样本, 再通过热化学非平衡数值模拟方法进行气动热计算, 最后分别运用基于非嵌入式多项式混沌 (NIPC) 的方法和基于 Sobol 指数的方法开展不确定度量化和敏感性分析。结果表明, 在给定的输入变量不确定度的条件下, 壁面热流不确定度不小于 15.9%, 在驻点和肩部存在峰值分别约为 19.8% (0.087 MW/m^2) 和 17.3% (0.076 MW/m^2); 相比而言, 在给定变化范围内壁面热流对来流密度和来流速度更为敏感, 来流温度和壁面温度对热流变化不产生明显影响。

关键词: 气动热; 热化学非平衡; 多项式混沌 (PC); 不确定度量化; 敏感性分析

中图分类号: V211.3

文献标识码: A

文章编号: 1001-5965(2018)05-1102-08

随着计算机技术和数值计算方法的发展, 计算流体力学^[1]在工程中起到了越来越重要的作用。目前大部分的数值模拟都是确定性的, 即在计算中需要给出确定的来流条件、几何尺寸、边界条件以及精确的物理模型。然而高度复杂的航天飞行器的分析和设计中不确定性因素是普遍存在的, 例如几何模型在载荷作用下的变形和大气来流条件变化等。不确定性因素可分为偶然不确定度 (aleatory uncertainty) 和认知不确定度 (epistemic uncertainty) 两大类^[2]。偶然不确定度是由于事物固有的随机性导致的不确定度; 而认知不确定度是由于缺乏理论知识导致的不确定度。不确定度量化 (uncertainty quantification) 分析对复杂

航天系统的鲁棒设计和可靠性分析是非常重要的, 并对飞船返回舱的热防护系统 (Thermal Protection System, TPS) 设计具有重要参考意义^[3]。

飞船返回舱在再入大气层过程中飞行高度逐渐降低, 大气来流等条件也在不断发生变化。Champion^[4]的研究表明高度为 48 ~ 69 km 的大气密度波动约为 5%。因此为得到可靠的气动热数值预测, 必须考虑来流参数的不确定性和由此带来的气动热预测的不确定度。目前已经发展出很多不确定度量化方法, 如已经在工程中应用较多的蒙特卡罗 (Monte Carlo) 方法^[5-6], 这种方法的主要缺陷是需要大量样本, 对计算资源消耗较大。近年来, 采用多项式混沌 (Polynomial Chaos, PC)

收稿日期: 2017-05-12; 录用日期: 2017-08-11; 网络出版时间: 2017-10-16 10:13

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20171016.1013.001.html

* 通信作者。E-mail: yanchao@buaa.edu.cn

引用格式: 张伟, 王小永, 于剑, 等. 来流导致的高超声速气动热不确定度量化分析[J]. 北京航空航天大学学报, 2018, 44(5): 1102-1109. ZHANG W, WANG X Y, YU J, et al. Uncertainty quantification analysis in hypersonic aerothermodynamics due to freestream[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 1102-1109 (in Chinese).

方法在 CFD 不确定分析得到了广泛应用。多项式混沌方法又可分为嵌入式和非嵌入式^[7]。本文采用的是非嵌入式多项式混沌(Non-Intrusive Polynomial Chaos, NIPC)方法,即选择一组正交多项式作为空间的无限基函数,将变量分解为确定和随机两部分从而构造出多项式混沌展开式。该方法相比于嵌入式多项式混沌方法只需把 CFD 求解器当做黑盒子,基于确定性的解来近似多项式混沌展开式的系数。在 NIPC 方法的基础上,可采用 Sobol 指数来衡量输入变量的敏感性。

目前,NIPC 方法被广泛应用到 CFD 不确定性分析中。Hosder 等^[7-8]将 NIPC 方法应用于求解楔角变化的随机斜激波问题和来流马赫数与迎角变化的三维 NACA65A004 机翼气动力问题,研究了单个随机输入变量和多维不确定性输入变量的情况,结果表明 NIPC 方法具有较高的效率和准确性。Loeven 和 Bijl^[9]开展了 NACA5412 翼型数值模拟的几何不确定度量化分析,研究了翼型最大弯度、最大弯度弦向位置以及厚度 3 个不确定性因素对升力系数的影响。Weaver 和 Alexenko^[10]对 FIRE2 飞行器进行了高超声速数值模拟,研究了来流速度、来流密度、碰撞系数和来流温度对气动热和流场的影响。Bettis 和 Hosder^[2]对高超声速再入流动的不确定度问题进行了分析,研究混合不确定性输入变量对气动热的影响。相比于国外,国内类似的工作还未开展,相关的研究还未见公开文献报道。

本文以 Apollo 飞船返回舱为研究对象开展气动热数值预测,选取了来流速度、来流温度、壁面温度和来流密度 4 个不确定性输入变量,其中来流速度变化范围为 ± 120 m/s ($\pm 2\%$),来流温度、壁面温度和来流密度变化范围为 $\pm 10\%$,利用 NIPC 方法对气动热进行不确定度量化和敏感性分析,为飞船返回舱气动热防护设计提供参考依据,进一步用于高超声速飞行器的鲁棒优化设计和可靠性分析。

1 计算方法

1.1 控制方程

采用有限体积法求解层流 Navier-Stokes 方程组,曲线坐标系 (τ, ξ, η, ζ) 下,三维无量纲守恒形式的热化学非平衡控制方程为

$$\frac{\partial \tilde{Q}}{\partial \tau} + \frac{\partial \tilde{F}}{\partial \xi} + \frac{\partial \tilde{G}}{\partial \eta} + \frac{\partial \tilde{H}}{\partial \zeta} = \frac{Ma}{Re} \left(\frac{\partial \tilde{F}_v}{\partial \xi} + \frac{\partial \tilde{G}_v}{\partial \eta} + \frac{\partial \tilde{H}_v}{\partial \zeta} \right) + S \quad (1)$$

$$\tilde{Q} = 1/J(\rho_i, \rho u, \rho v, \rho w, \rho E, \rho E_v) \quad (2)$$

$$S = 1/J(\omega_i, 0, 0, 0, 0, S_v) \quad (3)$$

式中: \tilde{Q} 为守恒变量; \tilde{F} 、 \tilde{G} 和 \tilde{H} 为无黏通量; \tilde{F}_v 、 \tilde{G}_v 和 \tilde{H}_v 为黏性通量; J 为雅可比矩阵行列式; u 、 v 和 w 为速度分量; ρ 为密度; ρ_i 为组元 i 的密度; E 为总能量; E_v 为振动能量; S 为源项; S_v 为振动能量源项; ω_i 为组元 i 的化学反应质量生成率; τ 为时间; ξ 、 η 和 ζ 为广义坐标。以上各式具体表达式参考文献[11]。

1.2 离散方法

时间离散采用隐式 LUSGS 方法;无黏通量空间离散采用 Roe 格式,二阶 MUSCL 重构,采用 minmod 限制器;黏性通量离散采用二阶中心格式。

1.3 物理模型

真实气体效应计算过程中单一组分的黏性系数、平动温度热传导系数和振动温度热传导系数分别由 Blotter 曲线拟合和 Eucken 关系式计算。混合气体的黏性系数、热传导系数由 Wilke 半经验公式计算。

1.4 化学反应模型

化学非平衡源项主要反应了组分间化学反应的影响。根据组分选取的不同,空气的化学反应模型一般分为 2、5、7 和 11 组分化学反应模型。对于 n_s 个组分的混合气体,设化学反应总数为 n_r ,其化学反应方程为如下通式^[12]:



式中: X_i 为单位体积摩尔数的化学组分或催化剂, $X_i = \rho_i / M_i$, M_i 为气体组分的摩尔质量, α_r 和 β_r 分别为化学反应式中反应物和生成物的计量系数。

则 ω_i 具体形式为

$$\omega_i = M_i \sum_{r=1}^{n_r} (\beta_{r_i} - \alpha_{r_i}) (R_r^f - R_r^b) \quad (5)$$

整理后源项表达式为

$$\omega_i = M_i \sum_{r=1}^{n_r} (\beta_{r_i} - \alpha_{r_i}) \left[k_r^f \prod_{i=1}^{n_s} \left(\frac{\rho_i}{M_i} \right)^{\alpha_{r_i}} - k_r^b \prod_{i=1}^{n_s} \left(\frac{\rho_i}{M_i} \right)^{\beta_{r_i}} \right] \quad (6)$$

式中: α_{r_i} 和 β_{r_i} 分别为化学反应式中反应物和生成物的计量系数; k_r^f 和 k_r^b 分别为化学反应式的正向和逆向反应速率系数。

1.5 不确定度分析方法

NIPC 方法主要基于混沌多项式展开 (Polynomial Chaos Expansions, PCE) 理论, 具有坚实的数学基础, 是一种非常有效的基于随机展开的不确定度分析方法。PCE 方法的一个重要特性是可以将任一随机变量分解为确定和随机 2 个独立部分。在不确定性的流体力学问题中, 任意随机变量 α^* (如压力、温度和壁面热流等) 可以表示为

$$\alpha^*(\mathbf{x}, \boldsymbol{\lambda}) \approx \sum_{j=0}^P \alpha_j(\mathbf{x}) \Psi_j(\boldsymbol{\lambda}) \quad (7)$$

式中: $\alpha_j(\mathbf{x})$ 为确定部分的耦合系数; $\Psi_j(\boldsymbol{\lambda})$ 为以 n 维随机变量 $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_n)$ 为自变量的 j 阶随机基函数。

在式(7)中, $\Psi_j(\boldsymbol{\lambda})$ 的项数是无穷的, 出于计算量的考虑, PCE 模型通常在某阶 p 被截断。相应地, p 阶 PCE 模型的项数可以表示为阶数 p 和随机变量的维数 n 的函数^[13] 为

$$N_t = P + 1 = \frac{(n + p)!}{n!p!} \quad (8)$$

根据随机输入参数的分布类型, 需要选取不同的多项式基函数。当输入参数服从正态分布时选取 Hermite 正交多项式基函数; 当输入参数服从均匀分布时选取 Legendre 正交多项式基函数。

目前, 主要有 2 种形式的基于非嵌入式混沌多项式的不确定度分析方法: 随机响应面法 (Stochastic Response Surface Method, SRSM) 和基于 Galerkin 投影法。本文采用随机响应面法来求解 PCE 系数, 基本流程如图 1 所示。

为了计算 PCE 模型中的未知 PCE 系数 $\alpha_j (j=0, 1, \dots, P)$, 首先需要选取有效的样本点, 这里将回归中用到的样本点数表示为 N 。通常采用过采样 (over sampling) 策略, 文献[8]中推荐用两倍于未知 PCE 系数个数的样本 (即 $N = 2N_t = 2(P + 1)$) 可以得到比较满意的结果。采用最小二次回归来求解 PCE 系数, 将样本 $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \dots, \lambda_j, \lambda_{j+1}, \dots, \lambda_N]^T$ 和相应的函数响应值 $\mathbf{G} = [g(X_1), g(X_2), \dots, g(X_N)]^T$ 分别代入到 PCE 模型右端和左端得

$$\begin{bmatrix} \Psi_0(\lambda_1) & \Psi_1(\lambda_1) & \dots & \Psi_p(\lambda_1) \\ \Psi_0(\lambda_2) & \Psi_1(\lambda_2) & \dots & \Psi_p(\lambda_2) \\ \vdots & \vdots & & \vdots \\ \Psi_0(\lambda_N) & \Psi_1(\lambda_N) & \dots & \Psi_p(\lambda_N) \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_p \end{bmatrix} = \begin{bmatrix} g(X_1) \\ g(X_2) \\ \vdots \\ g(X_N) \end{bmatrix} \quad (9)$$

可以简写为

$$A\boldsymbol{\alpha} = \mathbf{G}$$

定义:

$$K(\boldsymbol{\alpha}) = \sum_{j=1}^N \varepsilon_j^2 = \sum_{j=1}^N [g(X_j) - \hat{g}(\lambda_j)]^2 \quad (10)$$

式中: $\hat{g}(\lambda_j) = \sum_{j=0}^P \alpha_j(\mathbf{x}) \Psi_j(\boldsymbol{\lambda})$, 也就是基于 PCE 模型估算出的各样本点上的函数响应值。

利用线性最小二次回归, 使 $K(\boldsymbol{\alpha})$ 最小, 可以求得 PCE 系数为

$$\begin{cases} \mu_R = \alpha_0 \\ \sigma_R = \sqrt{\sum_{j=1}^P \alpha_j \langle \Psi_j^2(\boldsymbol{\lambda}) \rangle} \end{cases} \quad (11)$$

敏感性是表征随机物理问题中每个输入变量对输出变量不确定度的相对贡献大小。基于方差的 Sobol 分解方法可以用于敏感性分析, PCE 系数确定之后, Sobol 指数可以通过以下推导获得, 输出变量总方差可以表示为^[14]

$$D = \sigma_R^2 = \sum_{j=1}^P \alpha_j \langle \Psi_j^2(\boldsymbol{\lambda}) \rangle \quad (12)$$

总方差可以分解为

$$D = \sum_{i=1}^{i=n} D_i + \sum_{1 \leq i < j \leq n} D_{i,j} + \sum_{1 \leq i < j < k \leq n} D_{i,j,k} + \dots + D_{1,2,\dots,n} \quad (13)$$

部分方差表示为

$$D_{i_1, i_2, \dots, i_s} = \sum_{\beta \in \{i_1, i_2, \dots, i_s\}} \alpha_\beta \langle \Psi_\beta^2(\boldsymbol{\lambda}) \rangle \quad 1 \leq i_1 < \dots < i_s \leq n \quad (14)$$

Sobol 指数定义为

$$S_{i_1, i_2, \dots, i_s} = \frac{D_{i_1, i_2, \dots, i_s}}{D} \quad (15)$$

输入变量 i 的 Sobol 指数 (S_{T_i}) 定义为包含变量 i 的所有部分 Sobol 指数的和, 即

$$S_{T_i} = \sum_{L_i} \frac{D_{i_1, i_2, \dots, i_s}}{D} \quad L_i = \{(i_1, i_2, \dots, i_s) : \exists k, 1 \leq k \leq s, i_k = i\} \quad (16)$$

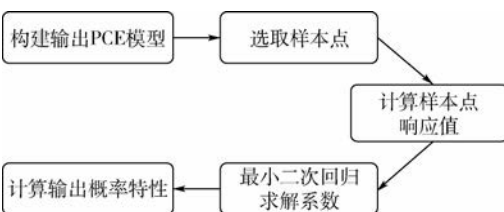


图 1 不确定度分析方法流程

Fig. 1 Flowchart of uncertainty analysis method

2 算例验证与结果分析

2.1 算例验证

本文首先选取 ELECTRE 标模^[15]对本文的热化学非平衡程序进行验证。ELECTRE 球锥标模几何外形如图 2(a)所示。来流速度 $u_\infty = 4230 \text{ m/s}$ ($Ma = 12.9$), 来流密度 $\rho_\infty = 6.944 \times 10^{-4} \text{ kg/m}^3$, 来流温度 $T_\infty = 265 \text{ K}$, 空气质量分数组成为 $Y_N = 0.77$, $Y_O = 0.23$, Y_N 和 Y_O 分别为 N_2 和 O_2 的质量分数, 采用双温度热力学模型和 Park 的 11 组分 21 反应化学反应动力学模型, 壁面为等温壁 ($T_w = 343 \text{ K}$), 并且设为完全催化壁。网格示意图如图 2(b)所示, 网格数为 80×90 , 第一层网格法向间距取为 $1 \times 10^{-5} \text{ m}$ 。

图 3 为本文章序的热流计算结果与飞行试验的对比。计算结果与飞行试验数据相吻合。在头

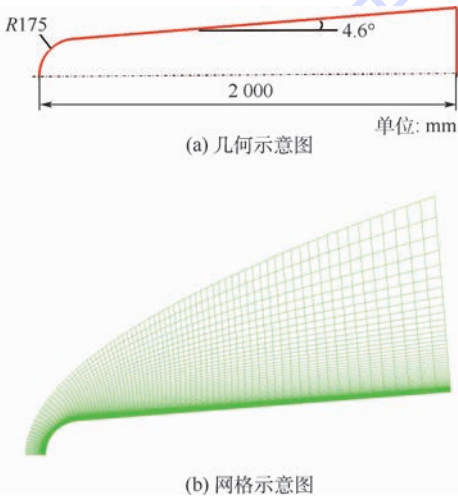


图 2 ELECTRE 标模几何与网格示意图
Fig.2 Schematic of geometry and computational mesh of ELECTRE vehicle

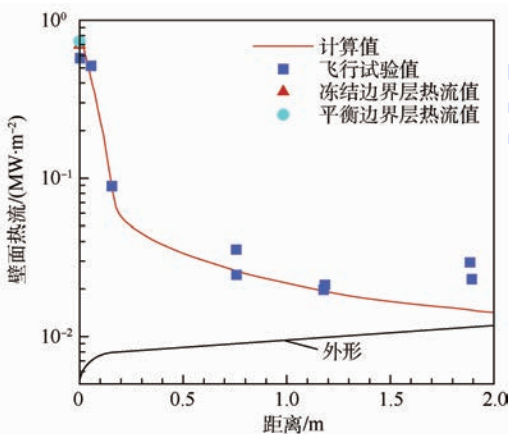


图 3 ELECTRE 标模壁面热流分布
Fig.3 Heat flux distribution on wall surface of ELECTRE vehicle

部驻点区域, 本文章序计算的热流与 Fay-Riddell 热流公式的结果吻合, 当略高于飞行试验结果。产生差异的原因是在实际飞行条件下 ELECTRE 标模的热防护设备减小了壁面催化效果^[16]。

图 4 对比了本文与文献^[15]中 ELECTRE 标模的驻点线温度分布。由图 4 可知, 本文章序的计算结果与文献^[15]的计算结果相吻合, 图中 T 为平动温度, T_v 为振动温度。

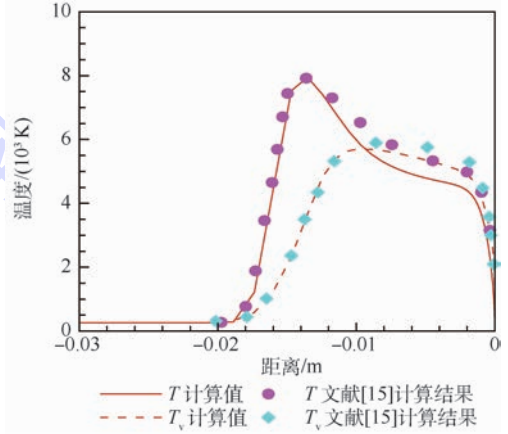


图 4 ELECTRE 标模驻点线温度分布
Fig.4 Temperature distribution along stagnation line of ELECTRE vehicle

2.2 结果分析

2.2.1 网格无关性验证与基准状态分析

计算模型为钝头体的 Apollo 飞船返回舱前体^[15], Apollo 飞船返回舱几何外形如图 5(a)所示, 基准状态流动条件如表 1 所示。计算中采用双温度热力学模型, 选用 Park 的 11 组分 21 反应化学反应动力学模型。壁面温度取为 800 K , 并且设为非催化壁面条件。网格如图 5(b)所示, 第一层网格高度为 $2 \times 10^{-4} \text{ m}$ 。

为了保证计算在一定网格密度下不受网格数量的影响, 本文选取了 3 个水平的网格数量来开展网格无关性研究: A (100×140), B (75×110), C (50×80)。图 6 给出了 3 个网格水平下壁面热流结果对比。可以看出网格 A 与网格 B 结果非常接近, 而网格 C 与其有一定区别, 故网格 B 已经可以满足网格无关性要求, 本文采用网格 B (75×110) 计算。

图 7 为基准状态下平动温度和压力的等值线图, 图 8 为平动温度及振动温度沿驻点线变化。激波过后, 平动温度首先被激发而升高。随后平动能和振动能之间产生能量松弛过程, 振动温度也随之增加。一段距离后, 平动温度和振动温度达到平衡。

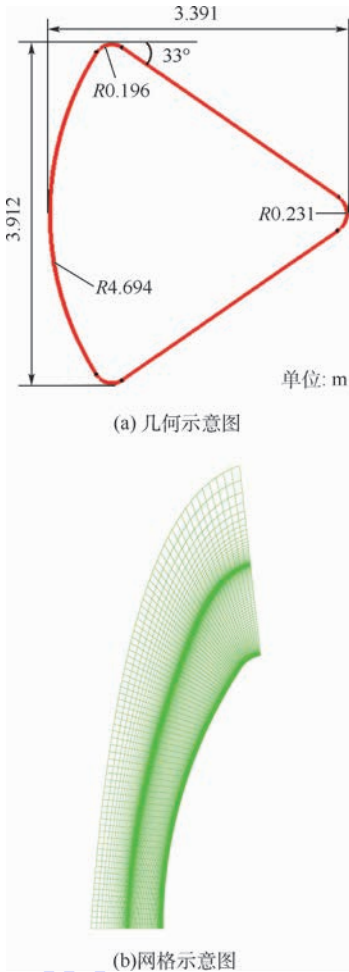


图5 Apollo 返回舱几何与网格示意图
Fig.5 Schematic of geometry and computational mesh of Apollo returning capsule

表1 基准状态流动条件

Table 1 Flow condition of baseline state

参数	高度/km	$u_\infty / (\text{m} \cdot \text{s}^{-1})$	T_∞ / K	$\rho_\infty / (\text{kg} \cdot \text{m}^{-3})$	Y_N / Y_O
数值	67.3	6040	225	9.8×10^{-5}	0.77/0.23

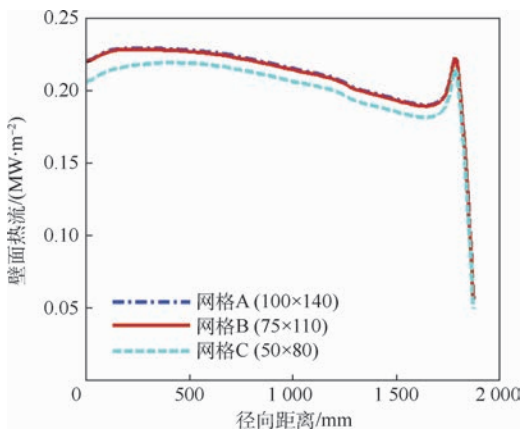


图6 不同网格水平下壁面热流对比

Fig.6 Wall heat flux comparison under different grid levels

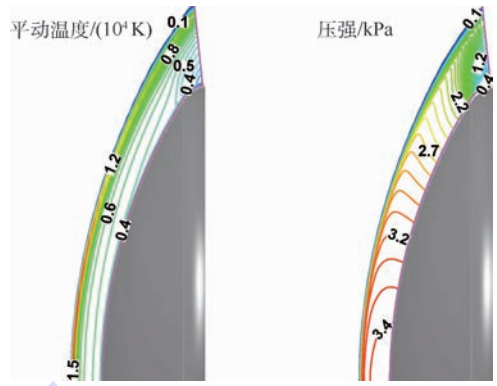


图7 基准状态平动温度与压力分布

Fig.7 Translational temperature and pressure contours for baseline case

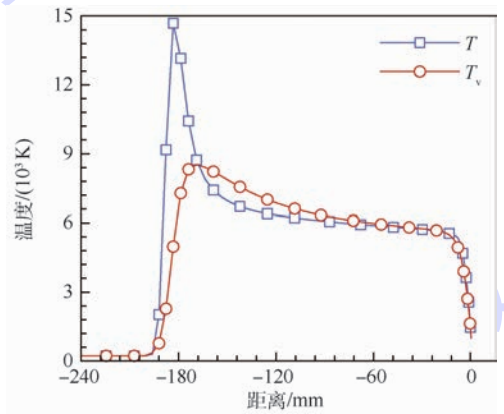


图8 Apollo 返回舱驻点线温度分布

Fig.8 Temperature distribution along stagnation line of Apollo returning capsule

2.2.2 不确定度量化

选取大气来流条件(来流速度、来流温度与来流密度)和边界条件(壁面温度)共4个变量作为随机输入来研究输出响应的不确定度传播和敏感性分析。4个变量的不确定度大小如表2所示,取值范围参考了文献[10],来流速度变化范围为 $\pm 120 \text{ m/s}$ ($\pm 2\%$),来流温度、壁面温度和来流密度变化范围为 $\pm 10\%$,并假设均服从均匀分布。

PCE模型在二阶截断,由式(8)计算得到样本数为30,采用拉丁超立方(Latin Hypercube Design, LHD)^[12]方法进行抽样。95%置信度条件下,定义输出响应壁面热流的不确定度为 $UQ\% =$

表2 输入变量不确定度

Table 2 Uncertainties of input variable

参数	不确定度
u_∞	$\pm 120 \text{ m/s}$ ($\pm 2\%$)
T_∞	$\pm 10\%$
T_w	$\pm 10\%$
ρ_∞	$\pm 10\%$

$$100 \times 1.96\sigma_R/\mu_R。$$

图 9 为基准状态下的壁面热流和 NIPC 方法计算的壁面均值热流。壁面热流沿径向距离总体呈先减小后增大再减小的趋势,在驻点和肩部附近出现峰值,在肩部附近,曲率突然减小且膨胀波气流加速换热效应起主导作用,故其壁面热流较高。基准状态下壁面热流和 NIPC 方法计算的壁面均值热流在驻点处差异为 3.2%,在肩部峰值处差异约为 0.78%。

图 10 为相应的壁面热流不确定度和输入变量敏感性指数分布。壁面热流的不确定度沿径向距离先减小后增大再减小,在驻点和肩部存在峰值分别约为 19.8% 和 17.3%,相应的热流不确定度区间大小分别约为 0.087 MW/m² 和 0.076 MW/m²。

驻点和肩部热流密度的平均值、标准差、95% 置信不确定区间和不确定度如表 3 所示。

不确定度量化结果表明,不确定性输入变量会引起壁面热流明显的变化,其不确定度范围

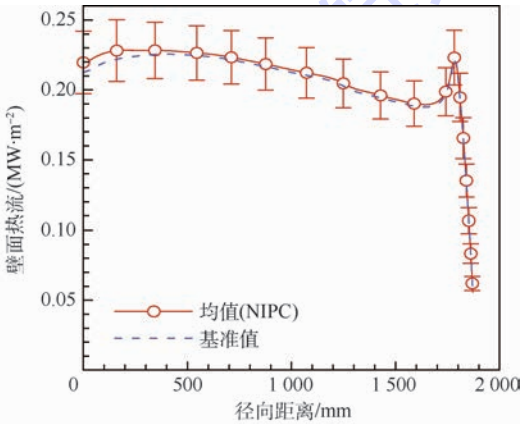


图 9 壁面热流对比

Fig. 9 Wall heat flux comparison

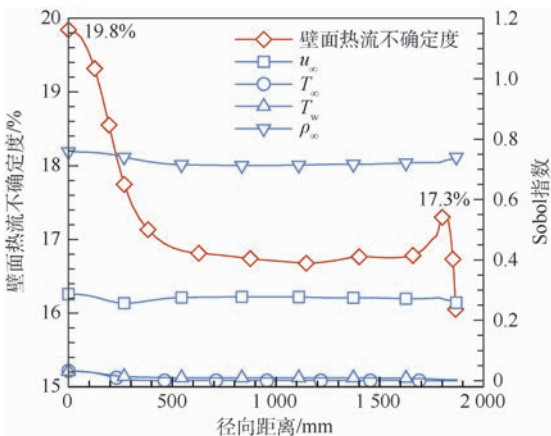


图 10 壁面热流不确定度和输入变量 Sobol 指数分布

Fig. 10 Uncertainty of wall heat flux and Sobol index distribution of input variable

表 3 不确定度量化结果

Table 3 Results of uncertainty quantification

量化结果	驻点 $S_R = 0$	肩部 $S_R = 1817$ mm
平均值 $\mu_R / (\text{MW} \cdot \text{m}^{-2})$	0.2201	0.2230
标准差 $\sigma_R / (\text{MW} \cdot \text{m}^{-2})$	0.0223	0.0195
2- σ 区间	(0.1764, 0.2638)	(0.1848, 0.2612)
不确定度/%	19.8	17.3

约为 15.9% ~ 19.8%,并且在驻点(19.8%)和肩部(17.3%)达到峰值。

2.2.3 敏感性分析

各输入变量的 Sobol 指数沿径向距离变化如图 9 所示,在来流速度、来流温度、壁面温度和来流密度 4 个不确定性变量中,各变量的 Sobol 指数在驻点处稍大,沿径向距离变化较为平缓,故各变量对不同位置的壁面热流影响规律较为一致。来流密度的不确定度对壁面热流变化贡献最大,其 Sobol 指数约为 0.72。其次是来流速度,其 Sobol 指数约为 0.27。相比而言壁面热流几乎不受来流温度和壁面温度影响。

驻点和肩部各输入变量的 Sobol 指数如表 4 所示。

各输入不确定度对输出不确定度相对贡献如图 11 所示。在驻点处,来流密度和来流速度对壁面热流不确定度贡献分别为 62.29% 和 23.56%,相应的不确定度和不确定度区间大小分别为 12.3%、4.8% 和 0.054、0.02 MW/m²。在肩部,来流密度和来流速度对壁面热流不确定度贡献分别为 72.76% 和 26.74%,相应的不确定度和不确定度区间大小分别为 12.6%、4.7% 和 0.056、0.02 MW/m²。来流温度、壁面温度及其他交互影响项几乎不对壁面热流变化起作用。

敏感性分析表明,在高超声速飞行器气动热数值模拟和试验中,需要重点关注来流密度和来流速度的变化。来流密度变化(10%)和来流速度变化(2%)对壁面热流的不确定度贡献最大分别约为 70% 和 25%,引起的不确定度分别约为 12% 和 5%,故来流密度和来流速度的变化会对气动热数值预测产生不可忽略甚至明显的影响,而来流温度和壁面温度几乎不产生影响。

表 4 驻点和肩部各输入变量 Sobol 指数

Table 4 Sobol index of each input variable at stagnation and shoulder

Sobol 指数	驻点 $S_R = 0$	肩部 $S_R = 1817$ mm
u_∞	0.2867	0.2681
T_∞	0.0325	0.0003
T_w	0.0312	0.0034
ρ_∞	0.7580	0.7295

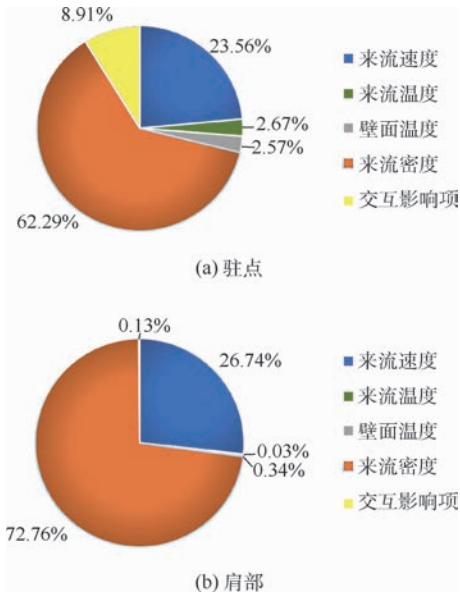


图 11 驻点和肩部各输入变量对总体不确定度的相对贡献

Fig. 11 Relative importance of each input variable on overall uncertainty at stagnation and shoulder

3 结论

本文对高超声速地球再入飞行器进行了热化学非平衡气动热数值预测,选取了来流速度、来流温度、壁面温度和来流密度 4 个不确定性输入变量,其中来流速度变化范围为 ± 120 m/s ($\pm 2\%$),来流温度、壁面温度和来流密度变化范围为 $\pm 10\%$,采用非嵌入式多项式混沌(NIPC)方法对气动热进行了不确定度量化分析和敏感性分析,得到的结论如下:

1) ELECTRE 标模热流计算结果与飞行试验值相吻合,表明热化学非平衡计算程序具有较高的气动热预测精度。Apollo 飞船返回舱再入时由于高温高马赫数,流场表现出典型的热化学非平衡特性,壁面热流在驻点和肩部附近存在峰值。

2) 采用基于 NIPC 方法的不确定性分析方法,得到了壁面各点热流的均值、标准差和不确定度。来流条件的变化会引起壁面热流明显变化,壁面热流的不确定度沿壁面先减小后增大再减小,在驻点和肩部存在 2 个峰值,其不确定度和不确定度区间大小分别约为 19.8%、17.3% 和 0.087、0.076 MW/m²。

3) 采用基于 Sobol 指数的敏感性方法,对计算结果进行了敏感性分析。各变量对不同位置的壁面热流影响规律一致,来流密度变化和来流速度变化对热流不确定度贡献最大,而来流温度变

化和壁面温度变化几乎不产生影响,在高超声速气动热预测中应重点关注来流密度和来流速度的变化。

4) 不确定度量化和敏感性分析对数值模拟可信度评估有重要意义,为高超声速气动热数值预测提供了可靠的参考,可进一步用于高超声速飞行器的鲁棒性优化设计和可靠性分析。后续工作将对热化学非平衡数值模拟中的化学反应模型进行不确定度分析。

参考文献 (References)

- [1] 阎超. 计算流体力学方法及应用[M]. 北京:北京航空航天大学出版社,2006:2-6.
YAN C. Computational fluid dynamic's methods and applications[M]. Beijing:Beihang University Press,2006:2-6 (in Chinese).
- [2] BETTIS B, HOSDER S. Uncertainty quantification in hypersonic reentry flows due to aleatory and epistemic uncertainties[C]//49th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition. Reston: AIAA, 2011.
- [3] BRUNE A J, WEST T K, HOSDER S, et al. A review of uncertainty analysis for hypersonic inflatable aerodynamic decelerator design[C]//21st AIAA International Space Planes and Hypersonics Technologies Conference. Reston: AIAA, 2017:2373.
- [4] CHAMPION K S W. Middle atmosphere density data and comparison with models[J]. Advances in Space Research, 1990, 10(6):17-26.
- [5] ECKHARDT R. Stan Ulam, John Von Neumann, and the Monte Carlo method[J]. Los Alamos Science Special Issue, 1987(15Special):131-137.
- [6] ZAREMBA S K. The mathematical basis of Monte Carlo and quasi-Monte Carlo methods[J]. SIAM Review, 1968, 10(3):303-314.
- [7] HOSDER S, WALTERS R, PEREZ R. A non-intrusive polynomial chaos method for uncertainty propagation in CFD simulations[C]//44th AIAA Aerospace Sciences Meeting and Exhibit. Reston: AIAA, 2006:891.
- [8] HOSDER S, WALTERS R, BALCH M. Efficient sampling for non-intrusive polynomial chaos applications with multiple uncertain input variables[C]//Non-Deterministic Approaches Conference. Reston: AIAA, 2007:1939.
- [9] LOEVEN A, BIJL H. Airfoil analysis with uncertain geometry using the probabilistic collocation method[C]//49th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference. Reston: AIAA, 2008:2070.
- [10] WEAVER A B, ALEXEENKO A. Flowfield uncertainty analysis for hypersonic CFD simulations[C]//48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition. Reston: AIAA, 2010:1180.
- [11] 周禹. 高超声速热化学非平衡流场数值模拟研究[D]. 北京:北京航空航天大学,2009:13-19.
ZHOU Y. Numerical simulation of hypersonic thermal and

- chemical nonequilibrium flows[D]. Beijing: Beihang University, 2009;13-19 (in Chinese).
- [12] WANG X, YAN C, ZHENG W L, et al. Laminar and turbulent heating predictions for mars entry vehicles[J]. *Acta Astronautica*, 2016, 128: 217-228.
- [13] 熊芬芬, 杨树兴, 刘宇, 等. 工程概率不确定性分析方法[M]. 北京: 科学出版社, 2015: 115-117.
- XIONG F F, YANG S X, LIU Y, et al. Engineering probability uncertainty analysis method[M]. Beijing: Science Press, 2015: 115-117 (in Chinese).
- [14] SOBOL I M. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates[J]. *Mathematics & Computers in Simulation*, 2001, 55(1-3): 271-280.
- [15] MUYLEAERT J, WALPOT L, HAUSER J. Standard model testing in the European high enthalpy facility F4 and extrapolation to flight[C] // Joint Propulsion Conference and Exhibit. Reston: AIAA, 1992: S396.
- [16] HAO J A, WANG J Y, LI C H. Numerical study of hypersonic flows over reentry configurations with different chemical non-equilibrium models[J]. *Acta Astronautica*, 2016, 126: 1-10.

作者简介:

张伟 男, 硕士研究生。主要研究方向: 计算流体力学及飞行器设计。

阎超 男, 博士, 教授, 博士生导师。主要研究方向: 计算流体力学。

Uncertainty quantification analysis in hypersonic aerothermodynamics due to freestream

ZHANG Wei, WANG Xiaoyong, YU Jian, YAN Chao*

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

Abstract: The CFD calculation is usually deterministic. However, errors and uncertainties always exist in the numerical simulation of complex engineering. Analysis and identification of source of uncertainty, and quantification of uncertainty play important roles in assessing the credibility of simulation results. Uncertainties are generally ubiquitous in highly complex aerospace systems. To obtain more reliable aerothermodynamics prediction, the uncertainty quantification and sensitivity analysis were carried out for the returning capsule hypersonic reentry flight. First, four uncertainty input variables (freestream velocity, freestream temperature, wall temperature and freestream density) were selected, whose variation ranges are ± 120 m/s ($\pm 2\%$), $\pm 10\%$, $\pm 10\%$ and $\pm 10\%$ respectively. And the samples were generated by Latin hypercube Design. Then, the thermochemical non-equilibrium numerical simulation method was used to calculate the aerodynamic heat. Finally, methods based on non-intrusive polynomial chaos (NIPC) and Sobol index were applied to uncertainty quantification and sensitivity analysis. The results show that the wall heat flux is not less than 15.9% under the given condition of uncertainty input variable uncertainty, and the peak value of the stagnation and shoulder are about 19.8% (0.087 MW/m²) and 17.3% (0.076 MW/m²) respectively. The uncertainty of heat flux is more sensitive to freestream density and velocity, and meanwhile freestream temperature and wall temperature variations almost have no impact on heat flux.

Keywords: aerothermodynamics; thermochemical non-equilibrium; polynomial chaos (PC); uncertainty quantification; sensitivity analysis

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

DOI: 10.13700/j.bh.1001-5965.2017.0387

龙卷风维持特性的探索

邸浩宇, 徐晶磊*, 高歌

(北京航空航天大学 能源与动力工程学院, 北京 100083)



摘 要: 利用计算流体力学(CFD)方法建立了模拟龙卷风的装置模型。基于龙卷风在平面上的速度型,拟合函数关系式,在三维CFD程序中设置初始速度场以及相关的边界条件,匹配不同的初始温度场进行了关于龙卷风维持和发展的一系列数值模拟。重现并利用漩涡管的 Ranque-Hilsch 效应以及汇流换热原理分析了本文2种不同的温度型对龙卷风内部流场变量随时间推移所发生的变化,同时揭示了龙卷风产生并维持下去的一种可能性,即冷热气流汇流换热生成温度场可能是产生和维持龙卷风的直接原因,并在此基础上提出使龙卷风削弱甚至消亡的方法,即破坏温度场。

关键词: 计算流体力学(CFD); 龙卷风; 数值模拟; 汇流换热原理; 温度场

中图分类号: O354.1; V211.1⁺⁴

文献标识码: A **文章编号:** 1001-5965(2018)05-1110-07

在自然界中,龙卷风可以产生摧毁性的能量,但龙卷风是如何生成并获得如此巨大的动能,特别是如何在较长时间维持旋转一直是个谜。对于龙卷风的研究,国内外学者主要通过现场实测、实验研究和数值模拟这3种手段。

在早期的现场实测中,由于龙卷风发生的地点和时间的不可预测性,以及测量技术的限制,导致实测研究发展很缓慢。Hoecker^[1]针对1957年发生的Dallas龙卷风,通过走访目击者以及察看龙卷风的图片和视频资料,指出涡核区域存在下沉气流并给出了龙卷风内部分位置的速度结构。21世纪初,多普勒雷达等的使用大大丰富了数据上的空缺,比较典型的是Alexander和Wurman^[2,3]利用移动式多普勒雷达实测了1998年5月30日发生在南达科他州Spencer的龙卷风,并提供了该龙卷风不同高度和时间的速度分布。

相比现场实测,实验研究和数值模拟更加方便易于收集数据。在龙卷风的实验研究中,

Chang^[4]于1971年率先建立了龙卷风发生装置并测定了龙卷风切向和径向的速度分布特征。基于Chang^[4]的研究,Ward^[5]于1972年利用蜂窝板消除龙卷风竖向的涡量,建立了著名的Ward模型。此后,许多科研人员又在Ward模型的基础上进行实验并改进建立了诸多模型,其中普渡大学的Church等^[6]引入旋转网格改进了Ward模型并进行了实验,发现龙卷风的涡可分裂呈现出多涡状态。Jischke和Parang^[7]利用Ward模拟器发现了核心半径这一参数的重要性,并提出将涡流比、高宽比和涡流雷诺数作为龙卷风模拟设计的主要参数。上述实验均未考虑龙卷风的横向移动,直到爱荷华州立大学开创了此先河,于2003年建立了可以使龙卷风进行平移运动的模拟发生器。之后Sarkar等^[8]利用其对龙卷风进行了风场特性的相关实验研究,并取得了和实测基本一致的结果。

近年来,随着计算机技术和计算流体力学

收稿日期: 2017-06-07; 录用日期: 2017-10-13; 网络出版时间: 2017-12-04 17:16

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20171204.1358.002.html

基金项目: 国防科技重点实验室基金(9140C410505150C41192)

*通信作者. E-mail: xujl@buaa.edu.cn

引用格式: 邸浩宇, 徐晶磊, 高歌. 龙卷风维持特性的探索[J]. 北京航空航天大学学报, 2018, 44(5): 1110-1116.

DI H Y, XU J L, GAO G. Exploration of tornado maintenance features[J]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44(5): 1110-1116 (in Chinese).

(CFD)相关模拟软件的迅猛发展,对龙卷风的研究更多地转入了数值模拟方面。Hassenzahl^[9]采用限制涡流的方式对龙卷风的涡流特性进行了研究。Maruyama^[10-11]于2009年应用大涡模拟研究了龙卷风的涡流特性;并于2011年在龙卷风模型中加入碎片,分析了碎片的运动特性,且利用其验证了龙卷风的涡核特性。Natarajan等^[12-13]则研究了建筑物表面的粗糙度对龙卷风风场的影响,并于2011年通过数值模拟了Ward-type TVC型、WinDEEE Dome型和AVE型这3种龙卷风发生器形成的龙卷风风场的涡流特性,与实验对比发现Ward-type TVC型与之更相符。徐枫等^[14]针对单涡龙卷风的风场特性,分析了切向风速在径向和高度的速度分布规律,并与Rankin涡模型和参数化气旋模型相对比,验证了模拟结果的合理性。

在以往对龙卷风的数值模拟研究中大多都以不可压缩假设处理,忽略了这种极端大气现象中气流可压缩性的影响。与此同时,忽略了温度场内气流之间的热交换。但龙卷风多发于气温很高的夏季时节,而且其内外的温度差很大,气体温度越高其包含的热能越多,在某些情况下,气体所储存的热能可转化为机械能的形式而输出能量做功。北京航空航天大学的高歌教授指出,目前热机做功基本上都是通过膨胀做功的原理来实现,例如拉瓦尔喷管等。但在自然界中所发生的某些现象虽然没有相应的做功机构,但仍然可以达到做功的目的,典型的例子就是龙卷风。显然龙卷风气体自身所具有的巨大热能是龙卷风能够做功的主要动力来源。故本文通过开发计算流体力学软件CFL3D(cfl3d v5 manual)^[15]给定了龙卷风初始速度场和边界条件,匹配不同的初始温度场以及龙卷风外围大气温度情况,研究龙卷风风场的发展演化与温度场的关系。

1 数值模拟设置

1.1 数值模拟几何模型

本文龙卷风的初场是给定的,且主要研究的是龙卷风的发展演化,因此其发生区域给定在一尺寸为1000 m×1000 m×5 m的长方体内。模型的网格是采用软件Pointwise生成的结构网格,因为参数变化主要集中在核心半径区域,所以在中心加密,从内到外网格由密变疏。网格维数为193×193×5,网格俯视图如图1所示。

1.2 控制方程及计算方法

本文选用的控制方程为连续性方程、Navier-Stokes方程以及能量方程。三维非定常流体的连

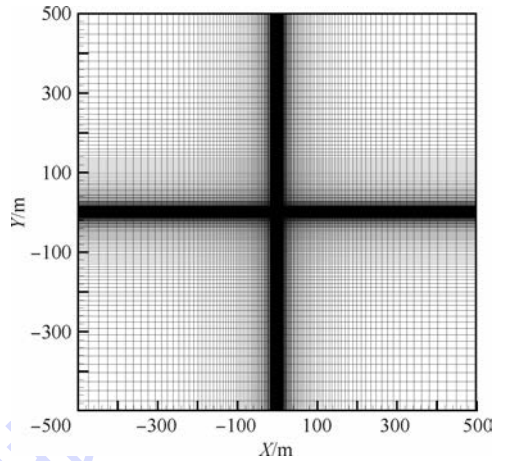


图1 龙卷风数值模拟的网格划分

Fig.1 Mesh generation for numerical simulation of tornado

续性方程为

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{V}) = 0$$

式中: ρ 为密度; t 为时间; ∇ 为Nabla算子; \mathbf{V} 为流体运动速度。

本次模拟采用可压缩流动模拟,故应用于本文的Navier-Stokes方程和能量方程分别为

$$\rho \mathbf{R} - \nabla p + \nabla \cdot \boldsymbol{\tau}_{ij} = \rho \frac{d\mathbf{V}}{dt}$$

$$\rho \frac{de}{dt} + p \nabla \cdot \mathbf{V} = \nabla \cdot (\lambda \nabla T) +$$

$$\nabla \cdot (\mathbf{V} \cdot \boldsymbol{\tau}_{ij}) + \rho q$$

式中: $\rho \mathbf{R}$ 为单位体积流体所受的质量力, \mathbf{R} 为质量力; p 为黏性流体的动压; $\boldsymbol{\tau}_{ij}$ 为黏性应力张量; e 为单位质量流体的能量; λ 为导热系数; T 为流体温度; q 为由于热辐射或其他原因在单位时间内传入的热量。

计算方法方面,无黏项空间离散格式采用三阶Roe格式+MUSCL插值,黏性通量空间离散格式为中心差分,时间推进方式采用近似因子分解算法,双时间步推进实现精确地实时计算。

1.3 初始速度场及边界条件的设置

龙卷风的初始速度场是根据文献[16]中所提供的某种情况下龙卷风的速度剖面图并查找龙卷风相关风强级别后加以修改的,如图2所示。 r_c 为初始速度场的涡核半径(速度最大处的半径)。在本文的模型中,初始速度场的半径 r 为140 m,涡核半径 r_c 为50 m,所取最大速度为60 m/s,属于F2级龙卷风速度范围。

本文忽略了龙卷风在水平方向上的移动,输入文件中每一步的时间步长 dt 为0.1 s,每个步长内迭代5次,计算步数NTSTEP为100、1000、2000、4000和8000。在数值模拟中边界条件的

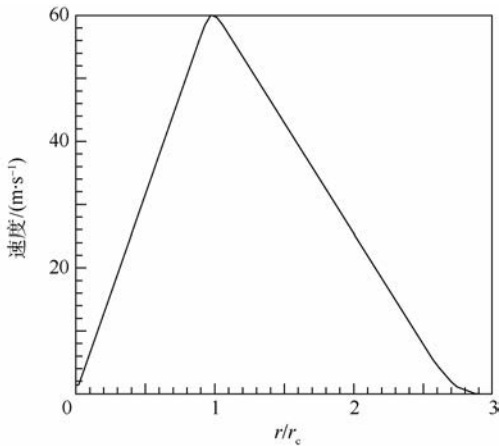


图2 初始速度场的速度分布

Fig. 2 Velocity distribution of initial velocity field

选择直接关系到模拟的可信度甚至成败,由于本文研究的是龙卷风内部某一截平面上的变化,相当于是水平截取龙卷风中的一部分,故设置6个面的边界条件均为无黏面(inviscid surface)。

1.4 不同初始温度场的设置

1) 算例1初始温度场为均一的温度场,即场域中初始温度均一样,为288 K。

2) 算例2的初始温度场采用类似于算例1相对稳定后的形式。

2 模拟结果与分析

在龙卷风的发展过程中,气流的旋转运动必然会改变局部气流的参数变化乃至影响到全场。在这些参数中,本文着重关注速度场、温度场、压力场和密度场这4个参数场,通过对比分析来解析龙卷风的发展过程。速度场直观地展现了场中机械动能的大致分布,显示出了龙卷风破坏力最强的地方;温度场的变化则是明确了内能转化为动能的方向以及全场能量变化的趋势;压力场最大的作用是提供足够的压力来抵消气流旋转产生的离心力来维持龙卷风旋转,而且压力主要依托温度和密度存在;密度场的变化主要体现了气流的运动方向,是向某一方向汇聚还是发散,可以说明该状态下龙卷风是被加强还是被削弱。

2.1 算例1

龙卷风的旋转外形是漏斗状的,其风速包含水平运动速度、竖直方向速度、径向速度和切向速度。本文忽略龙卷风水平运动速度,又由于发展演化过程中主要是内外气流的混合过程,所以主要考虑切向速度和径向速度的合速度的变化,即龙卷风在平面上的速度变化。图3给出了算例1随时间推移在平面上的速度变化曲线。由于在均

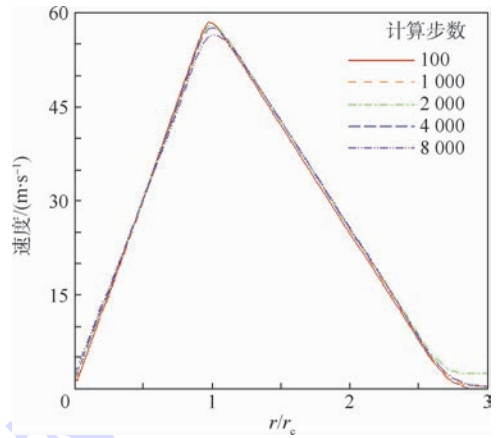


图3 不同计算步数下的速度变化曲线(算例1)

Fig. 3 Velocity variation curves under different calculation steps (Case 1)

匀温度场情况下,场内的变量变化并不剧烈,主要集中在初始速度型区域,所以只给出半径为 $3r_c$ 内的变量曲线。由图3可以看出,速度型随时间虽然是呈整体降低的形式,但降低的程度很小,而且扩散不厉害,也就是说必然有额外的能量提供来抵消动能的耗散。在整个场域中,气流除了有以动能为主的机械能外,就只有气流本身所储存的内能。而气体的内能主要是以温度的形式来体现的。

图4给出了算例1在不同计算步数下的温度变化曲线。从图中可以直观地看出,随着计算步数的增加,即随着时间的推移,中心温度开始降低最终形成中心低温区,而速度型边缘气流温度的变化则是先升高降低到保持为288 K。整个温度场从均一温度场变为有一定梯度的不均一温度场,显然是能量的转化导致了温度的变化,气流的内能在温度降低时将部分能量释放出来转化成了动能。这是气流以一定速度旋转的结果,可以用漩涡管的Ranque-Hilsch效应来解释,即当气流在向心旋流器中,通过向心螺旋通道送入旋涡管产生旋流时,会发生温度分离现象,旋涡管中外围的气流温度比中心的气流温度高且该温差随初始气流温度的升高而升高。

图5的压力变化曲线和图6的密度变化曲线应该结合图4的温度变化曲线分析,因为这3个参数必须满足气体状态方程 $P = \rho RT$ 。从图5和图6中参数的变化情况看,3个参数是符合该状态方程的。从图5中可以看出,外围高压区和中心低压区的形成向气流提供了足够的向心力来阻止扩散。从图6中可以看出,中心密度是先急剧下降之后又回升,是因为刚开始是压力场提供的向心力并不能够抵消气流旋转所产生的离心

力,因而中心区域气流会先向外扩散,通过和外围气流挤压换热产生足够强的压力并回流之后,再形成比较稳定的密度场。

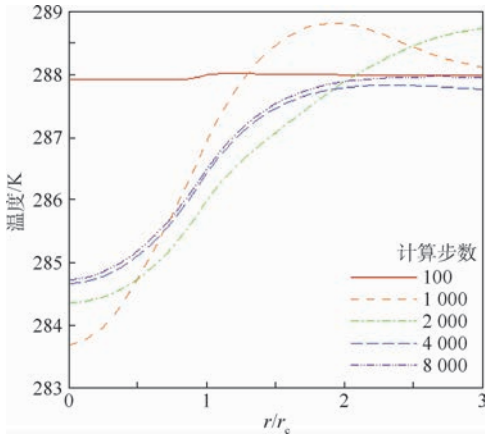


图 4 不同计算步数下的温度变化曲线(算例 1)
Fig. 4 Temperature variation curves under different calculation steps (Case 1)

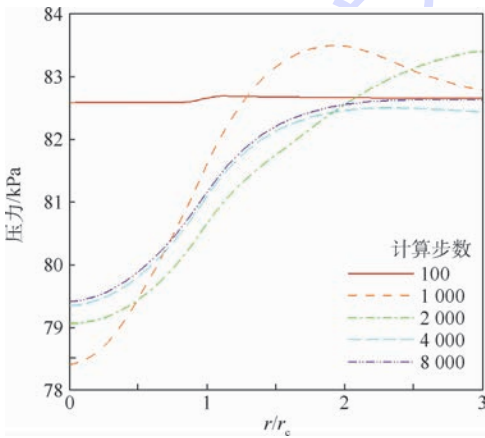


图 5 不同计算步数下的压力变化曲线(算例 1)
Fig. 5 Pressure variation curves under different calculation steps (Case 1)

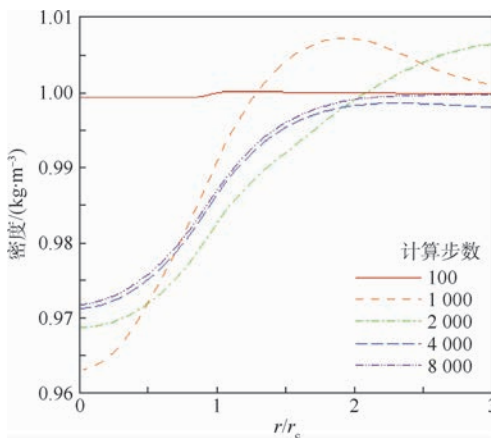


图 6 不同计算步数下的密度变化曲线(算例 1)
Fig. 6 Density variation curves under different calculation steps (Case 1)

最终,气流流场内的变量经过一系列的变化形成了一个各个参数互相匹配的新变量场,由此得到的新的旋涡场所具有各个参数场的形式即为所有稳定的旋涡场所具有的共有特性。

2.2 算例 2

由 2.1 节可知,算例 1 的数值模拟的理论基础是 Ranque-Hilsch 效应,通过验证可知漩涡场的旋转会自发的产生一个与之匹配的温度场来维持。而龙卷风也是一个比较特殊的漩涡场,故算例 2 在算例 1 的基础上将均一的温度场改为类似于算例 1 相对稳定后的温度场形式,将其放入原初始速度场中计算并分析结果。这样做的目的是观察加大温度梯度之后,龙卷风风场的变化。

算例 2 初始温度场为中心低温 253 K,到初始速度场边缘处线性升高为 303 K,初始速度场外气流温度也为 303 K。温度场函数关系式为

$$T = \begin{cases} 253 + \frac{5}{14}r & r \leq 140 \text{ m} \\ 303 & r > 140 \text{ m} \end{cases}$$

在有温度梯度的温度场情况中,场内的变量变化会延伸到部分外围气流,所以给出的都是半径为 $5r_c$ 内的变量曲线图。

图 7 给出了算例 2 随时间推移的速度变化曲线,可以看出在 2000 步的时候速度达到最大值,约为 62.7 m/s。图 8 将算例 1 中最大速度曲线同算例 2 中的速度曲线对比,可以发现算例 2 中速度曲线的各个速度最大值均大于算例 1 中出现的最大值,说明在算例 2 中速度在被加强。

从图 9 中可以看出,龙卷风在旋转过程中,中心温度升高减小了内外气流之间的温度差,在 2000 步时温差达到最小,同时速度也达到了最

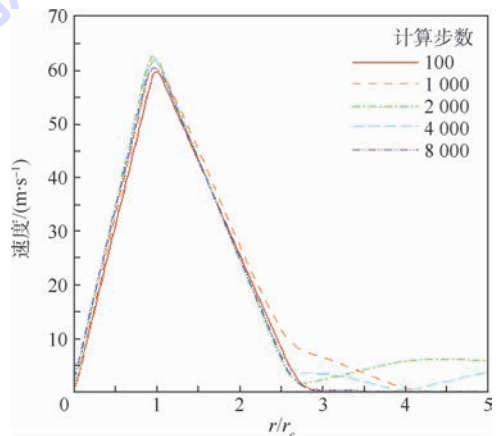


图 7 不同计算步数下的速度变化曲线(算例 2)
Fig. 7 Velocity variation curves under different calculation steps (Case 2)

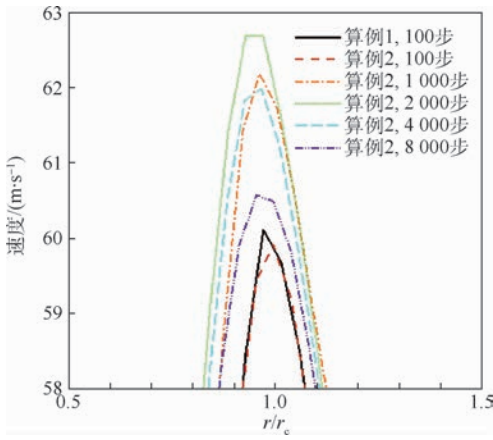


图8 速度变化曲线对比

Fig. 8 Comparison of velocity variation curve

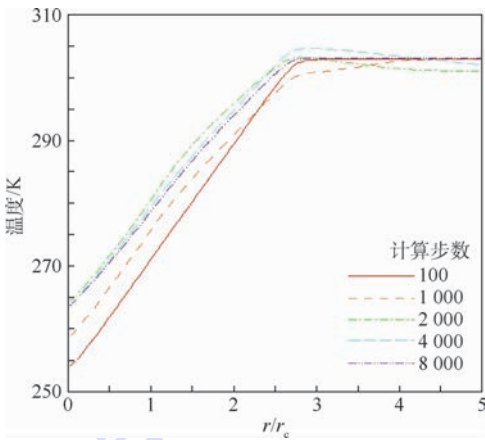


图9 不同计算步数下的温度变化曲线(算例2)

Fig. 9 Temperature variation curves under different calculation steps (Case 2)

大值。在整个变化过程中,温度型都基本保持了相似的形态,仅仅在温差上有变化,说明温度型的选择是正确的。在内外气流之间汇流换热的时候,外部高温气流的热能输运到内部低温气流使其温度升高。此时,内部气流的旋转速度增大则表明外部高温气流提供给内部低温气流的热能并非全部被用来增加内能,而是有部分热能转化为了动能加强了气流的旋转。这是由汇流换热效应所引起的变化。在汇流换热中,根据总压静压的关系:

$$P^* = P \left(1 + \frac{k-1}{2} Ma^2 \right)^{\frac{k}{k-1}} \quad (1)$$

由式(1)以及动量方程

$$dP = -\rho V dV$$

可推导出总温和总压的关系式为

$$\frac{dP^*}{P^*} = -\frac{kMa^2}{2} \cdot \frac{dT^*}{T^*} \quad (2)$$

式中: Ma 为气流马赫数; k 为绝热指数; P^* 和 T^* 分别为总压和总温,总温代表了热能的大小,总压

代表的做功能力的大小。

由式(2)可知,总温的降低会增大总压,而且在此基础上气流马赫数(速度)越高,其温度变化带来的总压变化越大,由热能储存成的机械能增多。高速气流降温、低速气流升温和降温带来的总压增强远大于升温带来的总压损失。显然,图8中速度的增大与图9中温度的变化验证了这一点。

因此可以认为,大气中的温差及内外对流是龙卷旋涡形成的前提条件,而使龙卷旋涡得以持续甚至增强的能量则来源于周围的热能。龙卷风外围的热气流与旋涡中心的冷气流形成的有一定温度梯度的温度场,成为把大气热能转化为旋涡流动动能的主要因素。自然界中,龙卷风自转的同时还发生沿地表的平移,那么龙卷风就源源不断地获得环境的热能,从而维持自身旋转。

从图10可以看出,此时算例2中提供的压力要大于算例1中所匹配出来的压力。从图11中可以更加直观的看出这一结果,对比图11与图6

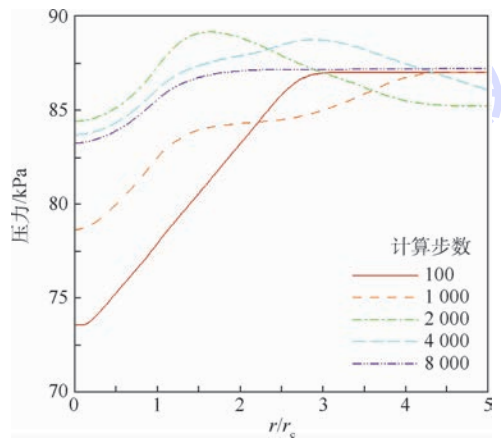


图10 不同计算步数下的压力变化曲线(算例2)

Fig. 10 Pressure variation curves under different calculation steps (Case 2)

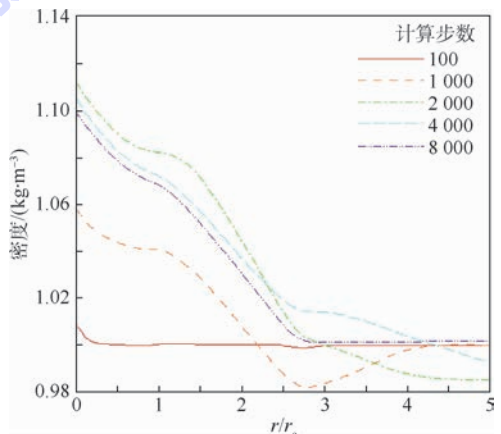


图11 不同计算步数下的密度变化曲线(算例2)

Fig. 11 Density variation curves under different calculation steps (Case 2)

发现在相同半径内,算例1中的密度变化曲线所对应的值大都在1以下,而算例2的密度变化曲线所对应的值则大都在1以上,显然在相同半径内算例2的密度积分要远大于算例1的密度积分,也就是说在这一半径内,算例2中包含有更多质量的气流,且这些气流的速度也很高,需要更大的压力来抵制离心力。这些集中在一定范围内的高速高质量的气流所具有的高动能表明该龙卷风的强度被大大加强,进一步说明了算例2的初始温度场加强了龙卷风的破坏能力。由此可以认为以这一温度型为基础,温度梯度越大对龙卷风风场的加强作用越大。

3 结论

1) 均一温度场的旋涡在旋转过程中会发生Ranque-Hilsch效应,产生一个不均一的温度场,通过这个温度场产生过程中气流内能所转化的动能来维持旋转,与此同时引起压力场和密度场的变化,最终这些变量互相匹配成一个新的旋涡变量场。由于新旋涡变量场是自适应生成的,所以其具有的变量场形式可代表所有稳定旋涡的变量场。

2) 特征如算例2的不均一的温度场更加贴近于实际龙卷风的温度场,其通过将气流汇聚到核心区域的加大旋转气流质量的同时加强气流旋转速度,来强化龙卷风旋涡并加剧其破坏力,其是由汇流换热原理所引起的。

3) 通过算例1的引申和算例2的验证,得到龙卷风的生成和强化主要依托于其温度场的变化,温度场梯度的大小在一定程度上决定了该龙卷风的破坏能力。因此,为使龙卷风的强度迅速降低甚至消亡,破坏龙卷风的温度场也许是一种行之有效的方法。

参考文献 (References)

- [1] HOECKER W H. Wind speed and air flow patterns in the Dallas tornado of April 2, 1957 [J]. Monthly Weather Review, 1960, 88(5): 167-180.
- [2] ALEXANDER C R, WURMAN J. The 30 May 1998 Spencer, South Dakota, storm. Part I: The structural evolution and environment of the tornadoes [J]. Monthly Weather Review, 2005, 133(1): 72-96.
- [3] WURMAN J, ALEXANDER C R. The 30 May 1998 Spencer, South Dakota, storm. Part II: Comparison of observed damage and radar-derived winds in the tornadoes [J]. Monthly Weather Review, 2005, 133(1): 97-119.
- [4] CHANG C C. Tornado wind effects on buildings and structures with laboratory simulation [C] // Proceedings of the Third International Conference on Wind Effects on Buildings and Structures, 1971: 231-240.
- [5] WARD N B. The exploration of certain features of tornado dynamics using a laboratory model [J]. Journal of the Atmospheric Sciences, 1972, 29(6): 1194-1204.
- [6] CHURCH C R, SNOW J T, AGEE E M. Tornado vortex simulation at Purdue University [J]. Bulletin of the American Meteorological Society, 1977, 58(9): 900-908.
- [7] JISCHKE M C, PARANG M. Properties of simulated tornado-like vortices [J]. Journal of the Atmospheric Sciences, 1974, 31(2): 506-512.
- [8] SARKAR P P, HAAN F L, GALLUS JR W A, et al. A laboratory tornado simulator: Comparison of laboratory, numerical and full-scale measurements [C] // 10th Americas Conference on Wind Engineering. Baton Rouge, LA: American Association for Wind Engineering, 2005.
- [9] HASSENZAHL H C. Numerical investigations of a tornado vortex using vorticity confinement [D]. Madison: University of Wisconsin Madison, 2007.
- [10] MARUYAMA T. A numerically generated tornado-like vortex by large eddy simulation [C] // Proceedings of 7th Asia Pacific Conference on Wind Engineering. Taipei: Wind Engineering, 2009.
- [11] MARUYAMA T. Simulation of flying debris using a numerically generated tornado-like vortex [J]. Journal of Wind Engineering and Industrial Aerodynamics, 2011, 99(4): 249-256.
- [12] NATARAJAN D. Numerical simulation of tornado-like vortices [D]. London: The University of Western Ontario, 2011.
- [13] NATARAJAN D, HANGAN H. Numerical study on the effects of surface roughness on tornado-like flows [C] // 11th Americas Conference on Wind Engineering (11ACWE). Baton Rouge, LA: American Association for Wind Engineering, 2009.
- [14] 徐枫, 肖仪清, 李波, 等. 龙卷风风场特性的 CFD 数值模拟 [J]. 空气动力学学报, 2013, 31(3): 350-356.
XU F, XIAO Y Q, LI B, et al. Study of the adaptive optimal design method based on design variables space [J]. Acta Aerodynamica Sinica, 2013, 31(3): 350-356 (in Chinese).
- [15] KRIST S L, BIEDRON R T, RUMSEY C L. CFL3D user's manual (version 5.0): NASA-TM-1998-208444 [R]. Washington, D. C.: NASA, 1998.
- [16] 王锦, 周强, 曹曙阳, 等. 龙卷风风场的试验模拟 [J]. 同济大学学报(自然科学版), 2014, 42(11): 1654-1659.
WANG J, ZHOU Q, CAO S Y, et al. Physical study on tornado-like flow based on tornado vortex simulator [J]. Journal of Tongji University (Natural Science), 2014, 42(11): 1654-1659 (in Chinese).

作者简介:

邸浩宇 男, 硕士研究生。主要研究方向: CFD 数值模拟、湍流模式理论。

徐磊 男, 博士, 讲师。主要研究方向: 计算流体力学方法、湍流模式理论。

高歌 男, 博士, 教授。主要研究方向: 发动机气动热力。

Exploration of tornado maintenance features

DI Haoyu, XU Jinglei*, GAO Ge

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

Abstract: A numerical model of the tornado was established by using computational fluid dynamics (CFD) method. Based on the speed type of the tornado in the plane, which was fitted to be the function relation, the initial velocity field and the related boundary condition were set in the 3D CFD program. Different initial temperature fields were carried out on a series of numerical simulations of tornado maintenance and development. The Ranque-Hilsch effect of the whirlpool and the convergence heat transfer principle were used to analyze the changes of the internal flow field variables of the two different temperature types in the tornado. It reveals the possibility of how the tornado was produced and continued, that is, the hot and cold air flow mixing and heat transfer which generated a temperature field may be the direct cause for producing and continuing a tornado. And a method to make tornado weak or even disappear was proposed, that is, destroying the temperature field.

Keywords: computational fluid dynamics (CFD); tornado; numerical simulation; convergence heat transfer principle; temperature field

Received: 2017-06-07; **Accepted:** 2017-10-13; **Published online:** 2017-12-04 17:16

URL: kns.cnki.net/kcms/detail/11.2625.V.20171204.1358.002.html

Foundation item: National Defense Science and Technology Key Laboratory Foundation of China (9140C410505150C41192)

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

《北京航空航天大学学报》征稿简则

《北京航空航天大学学报》是北京航空航天大学主办的以航空航天科学技术为特色的综合性自然科学学术期刊(月刊)。本刊以反映航空航天领域研究成果与动态、促进学术交流、培养科技人才和推动科技成果向社会生产力转化为办刊宗旨。本刊为中国自然科学技术核心期刊,并被 Ei Compendex 等国内外权威文献检索系统收录。本刊向国内外公开发行人,为进一步提高办刊质量和所刊出文章的学术水平,特制定本简则。

1 论文作者及内容

1.1 本刊面向海内外所有学者。

1.2 主要刊载与航空航天科学技术有关的材料科学及工程、飞行器设计与制造、宇航科学与工程、信息与电子技术、控制技术和自动化工程、流体力学和动力工程、计算机科学及应用技术、可靠性工程与失效分析等领域的研究文章。航空航天科学技术民用方面以及具有航空航天工程背景的应用数学、应用物理、应用力学和工程管理等方面的文章也在本刊优先考虑之列。

2 来稿要求

2.1 论文应具有创新性、科学性、学术性和可读性。

2.2 论文为原创作品,尚未公开发表过,并且不涉及泄密问题。若发生侵权或泄密问题,一切责任由作者承担。

2.3 主题明确,数据可靠,图表清晰,逻辑严谨,文字精练,标点符号正确。

2.4 文稿撰写顺序:中文题名(一般不超过 20 个汉字),作者中文姓名、单位、所在城市、邮政编码(作者顺序始终不得更改),中文摘要(包括目的、方法、结果及结论),中文关键词(5~8 个),中图分类号,引言,正文,参考文献,英文题名,作者英文姓名、单位、所在城市、邮政编码、国别,英文摘要,英文关键词。首、末页下角注明基金项目名称及编号,作者信息。

2.5 作者请登录本刊网页进行在线投稿。

3 稿件的审核、录用与版权

3.1 来稿须经专家两审和主编、编委讨论后决定刊用与否。

3.2 若来稿经审查后认定不宜在本刊发表,将及时告知作者。如果在投稿满 3 个月后仍未收到本刊任何通知,作者有权改投它刊。在此之前,请勿一稿多投,否则一切后果自负。

3.3 来稿一经刊登,即赠送单行本。

3.4 来稿一经作者签字并在本刊刊出,即表明所有作者都已经认可其版权转至本刊编辑部。本刊在与国内外文献数据库或检索系统进行交流及合作时,不再征询作者意见。

邮寄地址:100083 北京市海淀区学院路 37 号 北京航空航天大学学报编辑部

办公地点:北京航空航天大学办公楼 405,407,409 房间

电 话:(010)82315594,82338922,82314839,82315426

E-mail: jbuua@buaa.edu.cn

http://bhxb.buaa.edu.cn

http://www.buaa.edu.cn

《北京航空航天大学学报》
第五届编辑委员会



主任(主编): 赵沁平

(以下按姓氏笔划为序)

副主任(副主编): 丁希仑 王少萍 孙志梅 李秋实 李焕喜 杨嘉陵
苗俊刚 相艳 徐立军 钱德沛 曹晋滨

编委: 马殿富 王琪 王聪 邓小燕 王青云 王荣明 刘宇
刘红 江洁 刘强 闫鹏 朱天乐 刘铁钢 齐铂金
陈万春 邹正平 苏东林 杨世春 沈成平 邱志平 宋知人
杨树斌 张晓林 杨晓奕 杨继萍 李惠峰 吴新开 张瑞丰
杨照华 宋凝芳 周锐 林宇震 林贵平 战强 姚仰平
胡庆雷 赵秋红 段海滨 赵巍胜 席平 郭宏 徐洁
徐世杰 郭洪波 康锐 翟锦 熊华钢

北京航空航天大学学报

Beijing Hangkong Hangtian Daxue Xuebao

(原《北京航空学院学报》)

(月刊 1956年创刊)

第44卷第5期 2018年5月

JOURNAL OF BEIJING UNIVERSITY OF
AERONAUTICS AND ASTRONAUTICS
(JBUA)

(Monthly, Started in 1956)

Vol.44 No.5 May 2018

主管单位 中华人民共和国工业和信息化部

主办单位 北京航空航天大学

主编 赵沁平

编辑出版 《北京航空航天大学学报》
编辑部

邮编 100083

地址 北京市海淀区学院路37号

印刷 北京科信印刷有限公司

发行 北航文化传媒集团

发行范围 国内外发行

联系电话 (010) 82315594 82338922
82314839

电子信箱 jbuua@buaa.edu.cn

Administrated by Ministry of Industry and Information
Technology of the People's Republic of China

Sponsored by Beijing University of Aeronautics
and Astronautics (BUAA)
(Beijing 100083, P. R. China)

Chief Editor ZHAO Qinping

Edited and Published by Editorial Board of JBUA

Printed by Beijing Kexin Printing Co., Ltd.

Distributed by BUAA Culture Media Group Limited

Telephone (010) 82315594 82338922
82314839

E-mail jbuua@buaa.edu.cn

http://bhxb.buaa.edu.cn

刊号 ISSN 1001-5965
CN 11-2625/V

国内定价 50.00元/期

ISSN 1001-5965

