



北航学报  
赠阅

ISSN 1001-5965  
CODEN BHHDE8

# 北京航空航天大学 学报

JOURNAL OF BEIJING UNIVERSITY OF  
AERONAUTICS AND ASTRONAUTICS



**2019-09**

Vol.45 No.9

目 次

基于铱星机会信号的定位技术 ..... 秦红磊, 谭滋中, 丛丽, 赵超 (1691)

含内热源的多孔方腔流热耦合非正交 MRT-LBM 数值模拟 .....  
..... 张莹, 黄逸宸, 陈岳, 马明, 李培生, 王昭太 (1700)

机动发射条件下空间飞行器上升段弹道设计 ..... 鲜勇, 任乐亮, 郭玮林, 张大巧, 李冰 (1713)

TC4 喷丸强化仿真与试验 ..... 王延忠, 李菲, 陈燕燕, 张亚萍, 吴泽刚, 王成 (1723)

基于出行链的电动汽车充电行为影响因素分析 ..... 于海洋, 张路, 任毅龙 (1732)

基于有限状态机的交会对接飞行任务规划方法 ..... 杨胜, 王鑫哲, 李蒙 (1741)

不确定容量下时隙分配问题两阶段规划模型 ..... 元尧, 王瑛, 梁颖, 姚岷 (1747)

基于可变阶变分模型的医用低剂量 CT 图像去噪 ..... 王娜, 张权, 刘祎, 贾丽娜, 桂志国 (1757)

空间目标的 ISAR 成像及轮廓特征提取 ..... 杨虹, 张雅声, 尹灿斌 (1765)

舰载机安全逃逸复飞的参数适配包线 ..... 林佳铭, 吴光辉, 王立新, 刘海良, 王云 (1777)

主动防御飞行器的范数型微分对策制导律 ..... 郭志强, 孙启龙, 周绍磊, 闫实 (1787)

贮箱爆炸碎片初始速度及影响因素 ..... 赵蓓蕾, 赵继广, 崔村燕, 王岩, 孙武华 (1797)

Dirichlet 混合样本的 EM 算法与动态聚类算法比较 ..... 夏棒, EMILION Richard, 王惠文 (1805)

基于高阶 LADRC 的 V/STOL 飞机悬停/平移模式鲁棒协调解耦控制 .....  
..... 高阳, 吴文海, 嵇绍康, 郑毅 (1812)

基于相位条纹的高精度 GPS 码相位测量方法 ..... 傅圣友, 王兆瑞, 金声震, 艾国祥 (1824)

基于 DWNN 的机电作动器渐变非线性故障诊断 ..... 王剑, 王新民, 谢蓉, 曹宇燕, 李婷 (1831)

冲突证据决策新方法及应用 ..... 赵静, 关欣, 刘海桥 (1838)

基于小波域清晰度评价的光场全聚焦图像融合 ..... 谢颖贤, 武迎春, 王玉梅, 赵贤凌, 王安红 (1848)

基于深度学习的无人机数据链信噪比估计算法 ..... 孙宇航, 曾国奇, 刘春辉, 张多纳 (1855)

基于深度卷积的残差三生网络研究与应用 ..... 厉铮泽, 杨小远, 朱日东, 王敬凯 (1864)

一种考虑 GPS 信号中断的导航滤波算法 ..... 何康辉, 董朝阳, 王青 (1874)

多向气动驱动器软体仿生舌弯曲状态的研究 ..... 董虎, 林苗, 顾苏程, 曹毅, 李巍 (1882)

基于强化学习的时间触发通信调度方法 ..... 李浩若, 何锋, 郑重, 李二帅, 熊华钢 (1894)

共用支承-转子结构系统振动耦合特性分析 ..... 章健, 张大义, 王永锋, 马艳红, 洪杰 (1902)

期刊基本参数: CN 11-2625/V \* 1956 \* m \* A4 \* 220 \* zh \* P \* ¥50.00 \* 900 \* 24 \* 2019-09

(编辑 张 嵘 李 晶 张欣蔚 孙 芳 王艳梅 贺 伟)

CONTENTS

Positioning technology based on IRIDIUM signals of opportunity  
..... QIN Honglei, TAN Zizhong, CONG Li, ZHAO Chao (1691)

Non-orthogonal multiple-relaxation-time lattice Boltzmann method for numerical simulation of thermal coupling with porous square cavity flow containing internal heat source  
..... ZHANG Ying, HUANG Yichen, CHEN Yue, MA Ming, LI Peisheng, WANG Zhaotai (1700)

Design of ascent trajectory of space vehicle in mobile launch condition  
..... XIAN Yong, REN Leliang, GUO Weilin, ZHANG Daqiao, LI Bing (1713)

TC4 shot peening simulation and experiment  
..... WANG Yanzhong, LI Fei, CHEN Yanyan, ZHANG Yaping, WU Zegang, WANG Cheng (1723)

Influential factors analysis of electric vehicle charging behavior based on trip chain  
..... YU Haiyang, ZHANG Lu, REN Yilong (1732)

RVD flight mission planning and scheduling method based on finite state machine  
..... YANG Sheng, WANG Xinzhe, LI Meng (1741)

Two-stage programming model for time slot allocation problem under uncertain capacity  
..... QI Yao, WANG Ying, LIANG Ying, YAO Di (1747)

Medical low-dose CT image denoising based on variable order variational model  
..... WANG Na, ZHANG Quan, LIU Yi, JIA Lina, GUI Zhiguo (1757)

ISAR imaging and contour feature extraction of space targets  
..... YANG Hong, ZHANG Yasheng, YIN Canbin (1765)

Parameter suitability envelope for safety bolter of a carrier-based aircraft  
..... LIN Jiaming, WU Guanghui, WANG Lixin, LIU Hailiang, WANG Yun (1777)

Norm differential game guidance law for active defense aircraft  
..... GUO Zhiqiang, SUN Qilong, ZHOU Shaolei, YAN Shi (1787)

Initial velocity and influence factors of tank explosion fragments  
..... ZHAO Beilei, ZHAO Jiguang, CUI Cunyan, WANG Yan, SUN Wuhua (1797)

Comparison between EM algorithm and dynamical clustering algorithm for Dirichlet mixture samples  
..... XIA Bang, EMILION Richard, WANG Huiwen (1805)

High-order LADRC based robust coordinated decoupling control for V/STOL aircraft in hover/translation mode  
..... GAO Yang, WU Wenhai, JI Shaokang, ZHENG Yi (1812)

High-precision GPS code phase measurement method based on phase stripe  
..... FU Shengyou, WANG Zhaorui, JIN Shengzhen, AI Guoxiang (1824)

Gradual fault diagnosis for electromechanical actuator based on DWNN  
..... WANG Jian, WANG Xinmin, XIE Rong, CAO Yuyan, LI Ting (1831)

A new conflict evidence decision method and its application  
..... ZHAO Jing, GUAN Xin, LIU Haiqiao (1838)

Light field all-in-focus image fusion based on wavelet domain sharpness evaluation  
..... XIE Yingxian, WU Yingchun, WANG Yumei, ZHAO Xianling, WANG Anhong (1848)

SNR estimation algorithm for UAV data link based on deep learning  
..... SUN Yuhang, ZENG Guoqi, LIU Chunhui, ZHANG Duona (1855)

Research and application of residual triplet network based on deep convolution  
..... LI Zhengze, YANG Xiaoyuan, ZHU Ridong, WANG Jingkai (1864)

A navigation filtering algorithm considering GPS signal interruption  
..... HE Kanghui, DONG Chaoyang, WANG Qing (1874)

Motion characteristics of soft bionic tongue based on multi-directional pneumatic actuator  
..... DONG Hu, LIN Miao, GU Sucheng, CAO Yi, LI Wei (1882)

Time-triggered communication scheduling method based on reinforcement learning  
..... LI Haoruo, HE Feng, ZHENG Zhong, LI Ershuai, XIONG Huagang (1894)

Coupling vibration characteristics analysis of shared support-rotors system  
..... ZHANG Jian, ZHANG Dayi, WANG Yongfeng, MA Yanhong, HONG Jie (1902)

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

DOI: 10.13700/j.bh.1001-5965.2018.0717

# 基于铱星机会信号的定位技术

秦红磊, 谭滋中, 丛丽\*, 赵超

(北京航空航天大学 电子信息工程学院, 北京 100083)



**摘 要:** 全球导航卫星系统(GNSS)到达地球表面的信号弱,易受干扰,并且需要投入大量的基础设施,而机会信号可有效弥补这方面的不足。针对目前陆基机会信号(广播、数字电视和移动基站等)存在的覆盖性、可用性限制,提出了利用铱星系统实现天基机会信号定位。通过对铱星信号的通信体制进行深入研究,建立了瞬时多普勒定位数学模型,提出了利用铱星信号的单音信号获取多普勒频率信息,并结合轨道预测模型计算的卫星轨道信息实现铱星定位的方法。实测数据验证结果表明,利用实际铱星信号能够实现精度优于200 m的定位。研究成果对基于机会信号的定位技术的理论研究及应用具有参考意义。

**关键词:** 机会信号; 铱星; 天基; 通信; 瞬时多普勒定位

**中图分类号:** P228.1; TN967.1

**文献标识码:** A      **文章编号:** 1001-5965(2019)09-1691-09

随着位置服务需求的范围越来越广,全球导航卫星系统(Global Navigation Satellite System, GNSS)得到了快速发展和广泛应用,然而其缺点也逐渐被认识到。GNSS信号较弱,容易受到有意或无意的干扰,且建立完备的卫星导航定位系统基础设施往往需要较大的资金投入,机会信号可有效弥补这方面的不足。目前典型机会信号如调频广播(FM)、数字电视(DTV)<sup>[1]</sup>等设施主要集中在人口密集的城市地区,在沙漠、海洋及偏远地区覆盖性不足。而铱星卫星信号具有覆盖范围广、信号强度高及抗干扰性强等优点,因此采用铱星系统进行定位可以有效克服陆基机会信号覆盖性不足的问题。

铱星通信系统是由66颗低轨卫星组成的全球卫星移动通信系统。2009年,美国铱星通信公司宣布开始设计第二代铱星系统——Iridium NEXT,截至到2018年底,已发射65颗第二代卫星。2016年,美国铱星通信公司表示,该公司的

卫星授时与定位(Satellite Timing and Location, STL)系统已经具备作为全球定位系统(Global Positioning System, GPS)备份或GPS补充的能力,STL系统将使用廉价的芯片为地球任意位置的用户提供精确的位置、导航和授时服务<sup>[2]</sup>,可作为GPS备份或核验GPS信号,然而该系统具体技术尚未公布。Joerger等<sup>[3]</sup>提出了利用铱星信号增强GPS系统实现浮点载波相位的定位技术,主要分析了利用铱星实现载波相位快速整周模糊的求解。

本文利用铱星系统信号作为机会信号实现定位。首先,深入研究铱星信号通信体制,主要分析了铱星信号的结构,为定位信息观测的获取提供了基础。其次,介绍了瞬时多普勒的定位技术,并建立了数学模型。然后,提出了利用铱星信号的单音信号提取瞬时多普勒频率测量信息,结合轨道预测模型计算卫星轨道实现多历元定位的方法。最后,利用实际接收的铱星信号进行定位验证。

收稿日期: 2018-12-03; 录用日期: 2019-03-30; 网络出版时间: 2019-04-08 11:21

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20190403.2044.002.html

基金项目: 国防科技创新特区项目

\* 通信作者. E-mail: congli\_bh@buaa.edu.cn

**引用格式:** 秦红磊, 谭滋中, 丛丽, 等. 基于铱星机会信号的定位技术[J]. 北京航空航天大学学报, 2019, 45(9): 1691-1699.

QIN H L, TAN Z Z, CONG L, et al. Positioning technology based on IRIDIUM signals of opportunity [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1691-1699 (in Chinese).

## 1 铱星系统及通信体制

铱星星座分布决定了接收机可见卫星数量,而定位观测信息提取方法与信号通信体制有关,本节主要分析了铱星系统的星座结构和信号通信体制。

### 1.1 铱星系统星座基本结构

1987年, Bertiger 首次提出铱星系统的概念<sup>[4]</sup>。铱星系统空间星座部分由 66 颗近极圆轨道卫星组成,共 6 个轨道面,每个轨道共有 11 颗卫星和 1 颗备用卫星。轨道倾角为  $86.4^\circ$ ,轨道高度为 780 km,轨道周期为 100.13 min。5 条同向轨道间隔为  $31.6^\circ$ ,第 6 条反向轨道与第 1 条轨道间隔为  $22^\circ$ ,可实现包括南北两极的全球覆盖<sup>[5]</sup>。Iridium NEXT 星座将包括 81 颗卫星,其中 66 颗运行的低轨道卫星将替换现有卫星星座,还有 6 颗在轨备份卫星及 9 颗地面备份卫星。

用户段经由 L 频段用户链路与空间段进行互联,其相控阵天线分成 3 块安装在卫星上,每块天线板产生 16 个点波束,每颗卫星共 48 个点波束,整个星座共有 3 168 个点波束,实际覆盖全球只需要 2 150 个点波束,其余点波束在卫星向高纬度地区移动时逐渐关闭,从而节省功率和减少高纬度地区具有多星可见性特点造成的干扰<sup>[6]</sup>。

星间链路是铱星有别于其他卫星移动通信系统的一大特点,更为地面站对星钟的校正提供了方便。卫星覆盖区内的任何用户通过星间链路就可以与其他覆盖区域内的任何用户进行通信,无需地面中继。用户终端可拥有多条路径与系统信关站通信,从而具有抗干扰和抗摧毁能力。

### 1.2 铱星信号基本结构

用户链路采用 FDMA/TDMA/SDMA/TDD 多

址方式,将每颗卫星 48 个点波束中的相邻 12 个点波束分为一组对总可用频带进行空分频率复用,每个点波束内再把频带按 FDMA 方式分为多条 TDMA 通道。在每个 TDMA 通道内,对同一用户的上行和下行链路采用 TDD 复用,即上下行链路分别处于同一条 TDMA 载波的不同帧的同一帧内。

分配给铱星的带宽为 1 616.0 ~ 1 626.5 MHz。1 616.0 ~ 1 626.0 MHz 为双工信道,作为业务信道使用,将 10 MHz 频带分为 30 个子带,每个子带又分为 8 个信道,共 240 个信道,每个子带频宽为 333.333 kHz,每个信道的频宽为 41.67 kHz。1 626.0 ~ 1 626.5 MHz 为单工信道,作为信令信道使用,将 0.5 MHz 频带分为 12 个信道,每个信道频宽为 41.67 kHz。上述双工信道和单工信道各自的频宽 41.67 kHz 又分为工作频宽 31.50 kHz 和保护频宽 10.17 kHz<sup>[7]</sup>。1 617.775 ~ 1 618.725 MHz 为铱星专属频率,1 618.725 ~ 1 626.5 MHz 为共享频率。铱星系统用户链路 FDMA 频率分配如图 1 所示。

铱星信号 TDMA 帧长为 90 ms。其中,单工信道分配时隙 20.32 ms,实际 Iridium NEXT 信号中下行用户链路的单工信道信号为 6.5 ~ 20.32 ms。双工信道占用 4 个 8.28 ms 的上行链路时隙和 4 个 8.28 ms 的下行链路时隙。每条 TDMA 载波的速率为 50 Kbit/s,每条信道的速率为 4.8 Kbit/s。帧结构如图 2 所示。

10 MHz 业务频带共有 240 个信道频带,分到 12 个点波束,每个点波束为 20 个信道频带。0.5 MHz 信令频带共有 12 个信道频带,分配到每个点波束。因此,48 个点波束中的每个点波束的频率包括 32 个信道频带。每个信道频带通过时

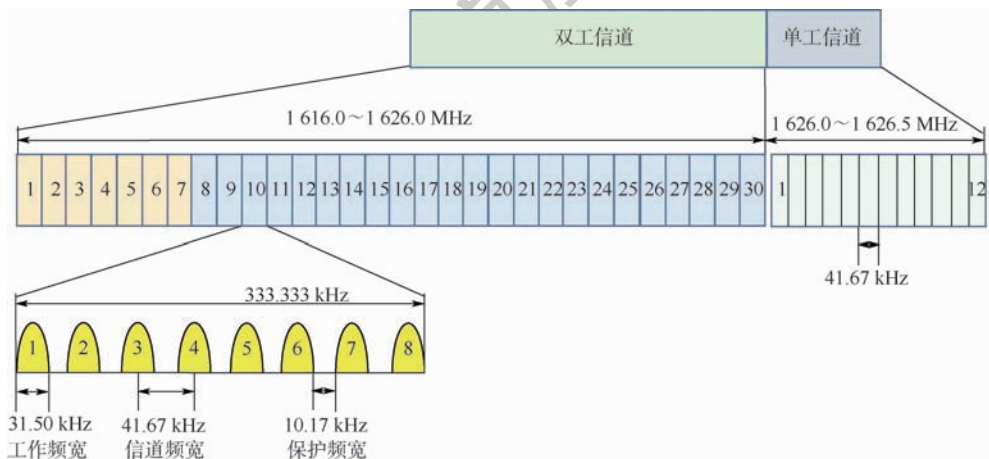


图 1 铱星系统用户链路 FDMA 频率分配

Fig. 1 IRIDIUM user links FDMA frequency assignment

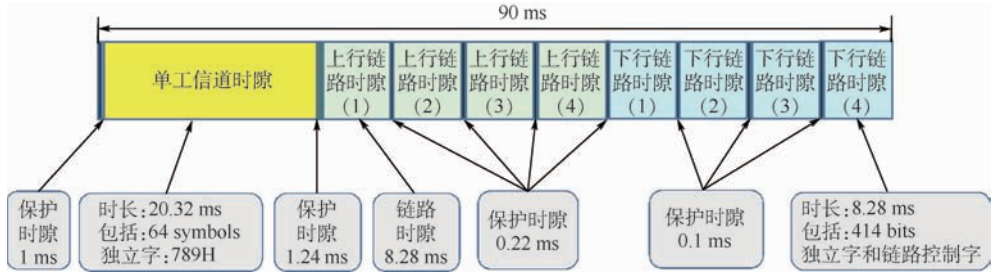


图 2 铱星系统用户链路 TDMA 帧结构

Fig. 2 IRIDIUM user links TDMA frame structure

分复用可得 4 个双工信道,每个点波束可得 80 个双工信道。整个卫星有 48 个点波束,最多可得 3 840 个 4.8 Kbit/s 的双工信道<sup>[8]</sup>。

用户下行链路传输 90 ms 突发信号,每个突发信号由 3 部分组成,分别为单音信号(tone)、独立字和数据信息。图 3 给出了铱星突发信号的生成技术及结构。

单工或双工信道数据信息首先采用 DQPSK 调制,数据率为 50 Kbit/s,然后将经 BPSK 调制后的独立字放置在数据信息前,该独立字主要用于修正铱星信号在传播过程中及滤波器处理过程中引入的相位旋转,独立字与数据信息组合之后,采用根升余弦滤波器进行脉冲整形,并将 2.6 ms 的单音信号放置在符号信息前,最后进行上变频得到突发信号<sup>[9]</sup>。

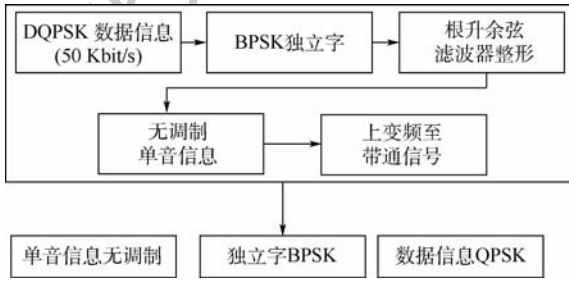


图 3 铱星系统用户链路信号结构

Fig. 3 IRIDIUM user links signal structure

## 2 基于瞬时多普勒的铱星定位技术

本节介绍了瞬时多普勒定位技术并给出了数学模型,同时根据铱星信号特性给出了定位观测信息获取及铱星机会信号定位方法。

### 2.1 瞬时多普勒定位算法

卫星与接收机之间的相对运动产生多普勒现象,多普勒频率能够反映卫星位置及速度与接收机位置及速度之间的关系,因此可采用测频信息实现定位。早期的子午仪系统(transit)是一种典型的基于多普勒的卫星定位系统,由于空间星座的局限性,可见卫星数目较少,因此采用积分多普

勒作为观测信息,定位解算时将测量信息转换为距离差双曲面,经过多个不同时刻的累积,多个双曲面交叉获得位置。在同时可见多颗卫星情况下,可采用瞬时多普勒频率实现即时定位,即利用多颗卫星的瞬时等多普勒圆锥曲面交叉获得接收机位置。

伪距定位线性导航状态更新方程为

$$\partial z = H \partial x + \epsilon \quad (1)$$

式中: $\partial z = z - \bar{z}$ , $\partial z$  为先验伪距测量偏差向量, $z$  为已测量的伪距向量, $\bar{z}$  为预测的伪距向量; $H$  为雅可比矩阵; $\epsilon$  为测量和线性化误差向量。

令  $\delta x = [\delta_x, \delta_y, \delta_z, \delta_b]^T$  表示状态更新向量, $\epsilon'$  表示误差变化量。对式(1)进行求导得到

$$\frac{\partial z}{\partial t} - \frac{\partial \bar{z}}{\partial t} = H \frac{\partial(\delta x)}{\partial t} + \frac{\partial(H)}{\partial t} \delta x + \epsilon' \quad (2)$$

根据接收机测速经典线性方程有

$$H \frac{\partial(\delta x)}{\partial t} = H \begin{bmatrix} \delta_x \\ \delta_y \\ \delta_z \\ \delta_b \end{bmatrix} \quad (3)$$

式中: $\delta_x, \delta_y, \delta_z$  为接收机先验速度的更新状态量; $\delta_b$  为接收机频偏的更新状态量。

根据多普勒频率与接收机位置之间的关系有

$$\frac{\partial(H)}{\partial t} \delta x = \begin{bmatrix} -\frac{\partial e^{(1)}}{\partial t} \\ -\frac{\partial e^{(2)}}{\partial t} \\ \vdots \\ -\frac{\partial e^{(k)}}{\partial t} \end{bmatrix} \begin{bmatrix} \delta_x \\ \delta_y \\ \delta_z \end{bmatrix} \quad (4)$$

式中: $\frac{\partial e^{(k)}}{\partial t} = \frac{\partial}{\partial t} \left( \frac{x^{(k)} - x_{xy=0}}{|x^{(k)} - x_{xy=0}|} \right)$ , $x^{(k)}$  为卫星  $k$  的位置, $x_{xy=0}$  为接收机先验位置信息, $e^{(k)}$  为卫星  $k$  的视距单位向量。

令  $d = \frac{\partial z}{\partial t}$  表示测量的多普勒向量, $\tilde{d} = \frac{\partial \bar{z}}{\partial t}$  表示预测的多普勒测量向量,综合式(1)~式(4),瞬时多普勒定位线性方程为

$$\partial \mathbf{d} = \mathbf{H} \begin{bmatrix} \delta_x \\ \delta_y \\ \delta_z \\ \delta_b \end{bmatrix} + \begin{bmatrix} -\frac{\partial \mathbf{e}^{(1)}}{\partial t} \\ -\frac{\partial \mathbf{e}^{(2)}}{\partial t} \\ \vdots \\ -\frac{\partial \mathbf{e}^{(k)}}{\partial t} \end{bmatrix} \begin{bmatrix} \delta_x \\ \delta_y \\ \delta_z \end{bmatrix} + \boldsymbol{\varepsilon}' \quad (5)$$

式中:  $\frac{\partial \mathbf{e}^{(k)}}{\partial t} = \left( \frac{\partial(\mathbf{x}^{(k)})}{\partial t} - \mathbf{e}^{(k)} \cdot (\mathbf{e}^{(k)} \cdot \mathbf{v}^{(k)}) \right) \frac{1}{r^{(k)}}$

表示卫星速度减去卫星视距方向的速度分量比上视距几何距离,  $\mathbf{v}^{(k)}$  为卫星  $k$  的速度向量。

式(5)即为瞬时多普勒定位的数学模型, 估计量包括接收机位置、速度及频偏等 7 个状态量。当位置状态量已知时, 式(5)即为接收机测速经典数学模型<sup>[10]</sup>。对于静态定位, 利用 4 颗卫星即可实现定位, 即

$$\partial \mathbf{d} = \begin{bmatrix} \left[ -\left( \frac{\partial(\mathbf{x}^{(1)})}{\partial t} - \mathbf{e}^{(1)} \cdot (\mathbf{e}^{(1)} \cdot \mathbf{v}^{(1)}) \right) \frac{1}{r^{(1)}} \right]^T & 1 \\ \left[ -\left( \frac{\partial(\mathbf{x}^{(2)})}{\partial t} - \mathbf{e}^{(2)} \cdot (\mathbf{e}^{(2)} \cdot \mathbf{v}^{(2)}) \right) \frac{1}{r^{(2)}} \right]^T & 1 \\ \vdots & \vdots \\ \left[ -\left( \frac{\partial(\mathbf{x}^{(k)})}{\partial t} - \mathbf{e}^{(k)} \cdot (\mathbf{e}^{(k)} \cdot \mathbf{v}^{(k)}) \right) \frac{1}{r^{(k)}} \right]^T & 1 \end{bmatrix} \cdot \begin{bmatrix} \delta_x \\ \delta_y \\ \delta_z \\ \delta_b \end{bmatrix} + \boldsymbol{\varepsilon}' \quad (6)$$

## 2.2 铱星定位技术

Iridium NEXT 提供的 STL 服务可实现室内定位精度 20 ~ 35 m (67%), 授时精度达  $1 \mu\text{s}$ <sup>[11]</sup>。这主要得益于在突发信号中的 DQPSK 数据中插入的伪随机序列, STL 接收机可通过与本地码相关实现高精度测距。Iridium NEXT 卫星采用星载 GPS 接收机进行精密定轨与授时的策略也为 STL 提供高精度定位与授时服务提供了保障。2006 年以前, 铱星任务计划与轨道分析组织提供铱星实时精确轨道数据库信息, 然而美国国防部与美国铱星通信公司合作后, 该组织不再提供铱星精确轨道信息<sup>[12]</sup>。本文利用 Iridium NEXT 的多普勒频移作为定位观测量, 采用公开的 TLE 数据及轨道预测模型计算 Iridium NEXT 轨道, 定位算法采用瞬时多普勒定位技术, 最终实现基于 Iridium NEXT 卫星的多历元接收机静态定位。

铱星单工下行用户链路的 12 个频带中, 持续存在突发信号的频带为 7 信道和 11 信道, 这里采用突发信号中的首段连续波单音信号测量多普勒

频移。相比于 GNSS 信号, 铱星信号到达地面的功率明显高于噪声信号。因此, 可采用快速傅里叶变换 (Fast Fourier Transform, FFT) 对多普勒频率进行粗测量, 在相邻频率带宽内采用最大似然估计 (MLE) 方法<sup>[13]</sup>进行多普勒精确测量, 最优估计函数为

$$I(f) = \frac{1}{N} \left| \sum_{n=0}^{N-1} x[n] \exp(-j2\pi fn) \right|^2 \quad (7)$$

式中:  $I(f)$  为最优估计函数;  $f$  为载波频率估计;  $x[n]$  为信号。

信号相位估计为

$$\hat{\phi} = \arctan \left( \frac{-\sum_{n=0}^{N-1} x[n] \sin(2\pi \hat{f}_0 n)}{\sum_{n=0}^{N-1} x[n] \cos(2\pi \hat{f}_0 n)} \right) \quad (8)$$

式中:  $\hat{f}_0$  为信号频率的最优估计值。

铱星采用极地轨道, 相比于高纬度地区,  $40^\circ$  纬度地区平均可见卫星个数为 1 ~ 2 颗, 接收机在待定点可持续搜集 10 ~ 20 min 卫星信号集中处理实现定位结算。因此, 当观测  $m$  个历元将会有  $m$  个接收机钟偏参数, 可根据接收机稳定度将钟偏参数用三阶或高阶多项式描述, 从而简化解算, 如下:

$$b'_i = b'_0 + b'_1(t_i - t_0) + b'_2(t_i - t_0)^2 \quad (9)$$

式中:  $b'_i$  为  $t_i$  时刻的接收机频偏参数;  $t_0$  为选定的参考时刻, 如可采用中央时刻作为参考时刻<sup>[14]</sup>。当采用外接高精度原子钟时, 接收机频偏影响可忽略。类似的, 对卫星频偏影响也可采用此种方法进行有效补偿, 即与接收机位置信息一并求解。

## 3 实验验证

### 3.1 铱星信号时频分析及解调

本节主要在频域和时域分析铱星信号, 并对信号进行解调处理。利用搭建的硬件采集系统对实际 Iridium NEXT 的用户链路单工下行 0.5 MHz 频带信号进行采集, 中心频率设置为 1 626.25 MHz, 采集的中频数据频率为 28.25 MHz, 数据长度为 20 s。分别对 7 信道信号和 11 信道信号进行处理。

#### 1) 7 信道信号

铱星 7 信道信号是 Ring Alert 信号, 所在的子带的中心频率为 1 626.270 833 MHz, 其实时频谱如图 4 所示, 经下变频后为 28.270 833 MHz。该信道信息每 48 帧发送一次, 即同一卫星的同一点波束重复发送该信号的时间间隔为 4.32 s。对

20 s内的某一突发信号进行处理,其时域信号如图5所示。图6和图7分别给出了信号解调后的IQ矢量图及相位图。

通过时域和频域分析结果可以看到,实际采集的信号中此突发信号带宽约为26.6667 kHz,持续时间为6.8 ms。采用最大似然估计方法捕获的信号中心频率为28.244 935 683 MHz,即卫星信

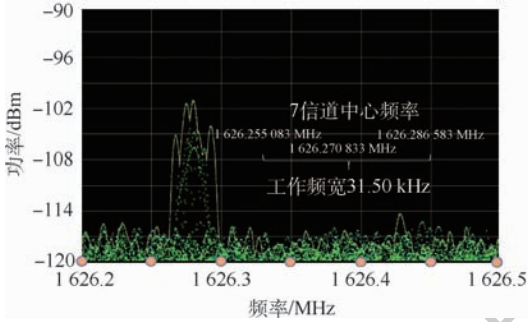


图4 铱星系统用户链路单工7信道信号实时频谱  
Fig.4 Real-time frequency spectrum of IRIDIUM user links simplex 7 channel signal

号的多普勒频率为25 897.317 Hz。通过图7可以看出,铱星突发信号主要由无调制、BPSK调制及QPSK调制3部分组成,其中单音连续波信号无任何信息,独立字采用BPSK调制,而数据信息采用QPSK调制。

2) 11信道信号

铱星11信道信号为Primary Message信号,所在的子带中心频率为1626.437 50 MHz,其实时频谱如图8所示,经下变频后为28.437 5 MHz。对20 s内的某一突发信号进行处理,其时域信号如图9所示。图10和图11分别给出了信号解调后的IQ矢量图及相位图。

通过时域和频域分析结果可以看到,实际采集的信号中此突发信号带宽约为29.1905 kHz,持续时间约为7.5 ms。采用最大似然估计方法捕获的信号中心频率为28.403 872 765 MHz,即卫星信号的多普勒频率为33 627.235 Hz。通过图11可以看出,11信道突发同样由无调制单音连续波、BPSK调制独立字及QPSK调制数据信息3部分组成。

以上给出了20 s数据中的2个通道的突发信

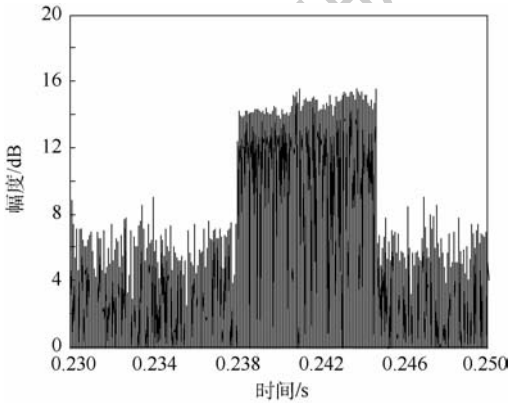


图5 铱星系统用户链路单工7信道时域信号  
Fig.5 Time domain signal of IRIDIUM user links simplex 7 channel

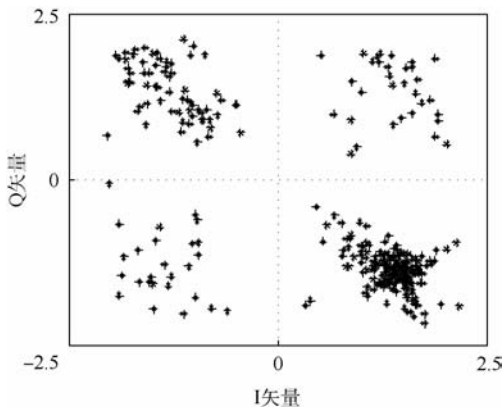


图6 铱星系统用户链路单工7信道基带信号IQ矢量图  
Fig.6 IQ vector of IRIDIUM user links simplex 7 channel baseband signal

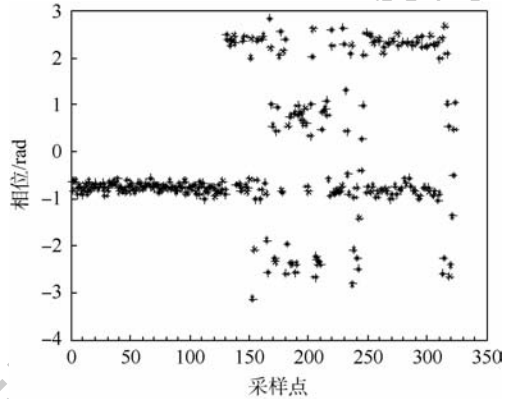


图7 铱星系统用户链路单工7信道基带信号相位图  
Fig.7 Phase of IRIDIUM user links simplex 7 channel baseband signal

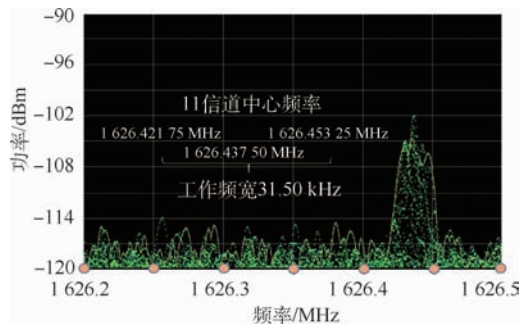


图8 铱星系统用户链路单工11信道信号实时频谱  
Fig.8 Real-time frequency spectrum of IRIDIUM user links simplex 11 channel signal

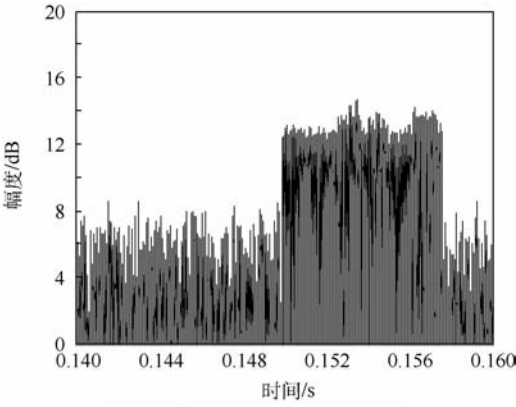


图9 铱星系统用户链路单工11信道时域信号

Fig.9 Time domain signal of IRIDIUM user links simplex 11 channel

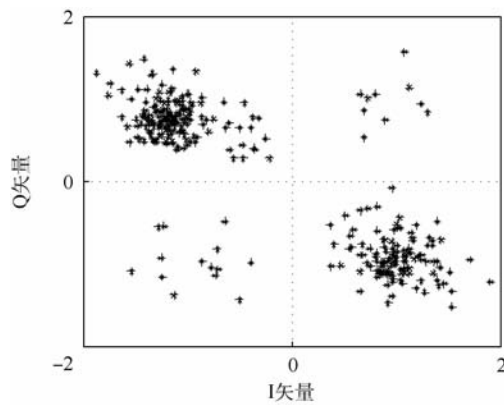


图10 铱星系统用户链路单工11信道基带信号IQ矢量图

Fig.10 IQ vector of IRIDIUM user links simplex 11 channel baseband signal

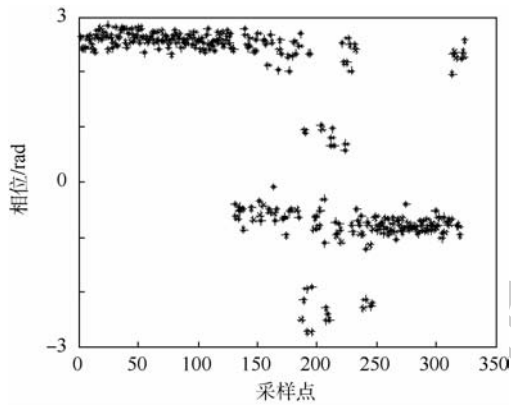


图11 铱星系统用户链路单工11信道基带信号相位图

Fig.11 Phase of IRIDIUM user links simplex 11 channel baseband signal

号分析结果,20 s 内其他突发信号持续时间在 6.5 ~ 20.32 ms 之间,其中单音信号持续时间为 2.6 ms 左右。另一方面,对于 Ring Alert 信号,实际数据中存在严格 90 ms 间隔的 2 个连续突发信号,即此时接收机位于同一卫星的 2 个相邻点波束覆盖范围内,出现的严格 90 ms 整倍数间隔的

2 个突发信号则来自于同一颗卫星的非相邻点波束的信号,而出现的并非严格 90 ms 整倍数间隔的 2 个突发信号表示此时接收机接收 2 个不同卫星的信号。

### 3.2 铱星多普勒定位

本节主要利用硬件采集系统接收 Iridium NEXT 信号,对卫星数据进行多普勒信息提取,并结合预测的轨道信息实现接收机定位。

持续接收信号 30 min,采用信号时域检测技术对每个 360 ms 数据块进行检测,当存在信号时提取多普勒信息。7 信道和 11 信道信号多普勒捕获结果如图 12 和图 13 所示。

可以看出,30 min 内总共可见卫星数量为 7 颗,同一时刻最多可见卫星数量为 2 颗。第 1 颗卫星为刚过天线头顶,第 3、5、7 颗卫星与第 1 颗卫星属于同一轨道,且该 3 颗卫星星下点轨迹与天线位置距离较近。第 2、4、6 颗卫星属于同一轨道卫星,且该 3 颗卫星星下点轨迹距离接收天线位置较远。由于同一卫星用户下行链路 7 信

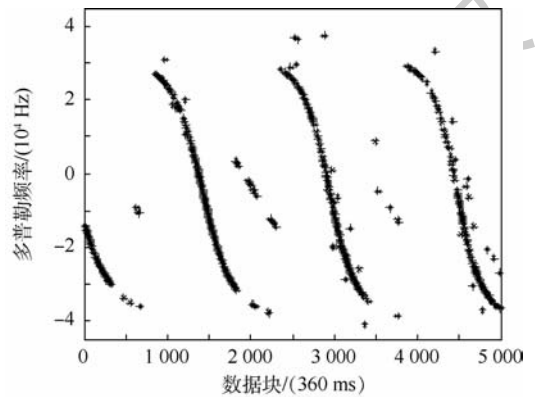


图12 铱星系统用户链路单工7信道信号多普勒曲线

Fig.12 Doppler curves of IRIDIUM user links simplex 7 channel signal

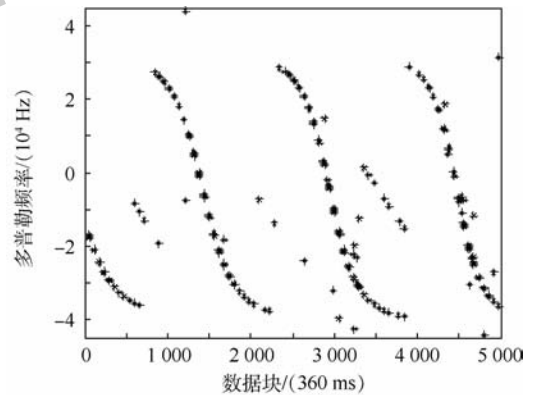


图13 铱星系统用户链路单工11信道信号多普勒曲线

Fig.13 Doppler curves of IRIDIUM user links simplex 11 channel signal

道和 11 信道均存在信号,因此两信道信号多普勒曲线大体形状相同,唯一区别是两信道信号并非同时发射,且 11 信道内信号发射时间间隔具有一定集中性。

本文采用 6 颗卫星的 7 信道信号多普勒频移作为铱星定位观测信息,共 25 个观测历元,利用高程辅助<sup>[15]</sup>瞬时多普勒定位进行定位解算。真实接收天线位置采用 GPS 记录仪进行标定(精度为米级)。定位误差分析采用不同 25 个观测历元多普勒信息进行组合 800 次,每 50 次组合方式定位结果进行定位误差均值和误差均方根值(RMS)统计。

表 1 给出了铱星定位误差的统计结果。可以看出,相比于北向和天顶方向,东向具有较大的定位误差,最大误差达到 404 m。图 14 和图 15 给出了高程辅助情况下的水平定位误差统计结果。可以看出,高程辅助情况下定位误差具有较小的均值及 RMS 值。东向误差波动相比于北向误差波动较大,但具有较小的定位误差均值。这是由于铱星轨道的倾角较大,因此南北方向运行的卫星的等多普勒圆锥面在受到误差影响时,定位误差分布在东西方向上较大。

图 16 给出了瞬时多普勒定位及高程辅助瞬时多普勒定位的误差 RMS 值。可以看到,高程辅助可有效改善定位性能,水平定位精度为 200 m 左右。

表 1 铱星多普勒定位误差统计结果

Table 1 Error statistic results of IRIDIUM Doppler positioning

统计参数	东向误差/m	北向误差/m	天顶误差/m
均值(最大)	213	49	-124
均值(最小)	106	16	-61
RMS(最大)	404	109	208
RMS(最小)	307	73	164

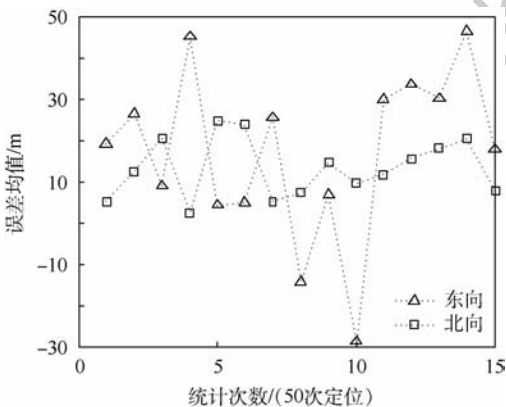


图 14 水平定位误差均值

Fig. 14 Error mean value of horizontal positioning

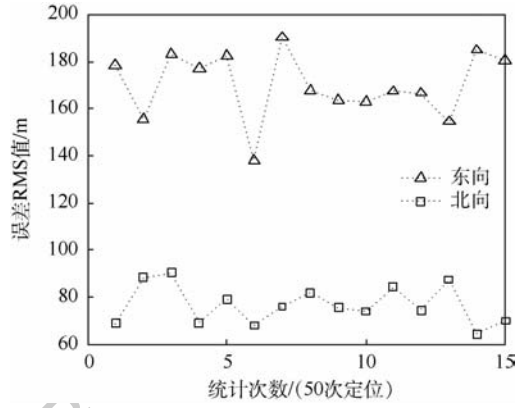


图 15 水平定位误差 RMS 值

Fig. 15 Error RMS value of horizontal positioning

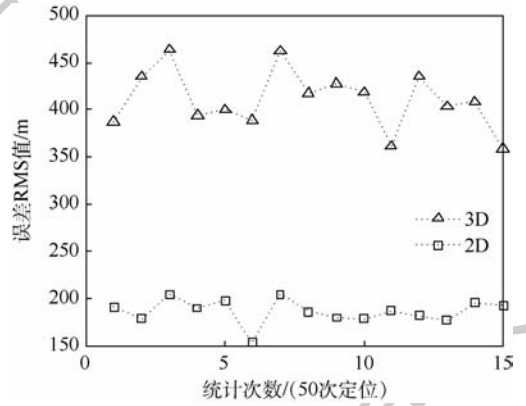


图 16 瞬时多普勒定位及高程辅助瞬时多普勒定位误差 RMS 值

Fig. 16 Error RMS value of instantaneous Doppler positioning and height aided instantaneous Doppler positioning

## 4 结 论

本文提出了利用铱星机会信号实现定位的方法,并利用实际信号进行实验验证。

1) 铱星信号主要由无调制单音信号、BPSK 调制独立字和 QPSK 调制数据信息组成。实际铱星用户链路下行单工信道持续存在突发信号的有 7 信道及 11 信道,实验验证结果表明,此突发信号持续时间在 6.5 ~ 20.32 ms 之间,解调实际信号获得基带信息的 IQ 矢量图和相位图符合信号结构分析结果。

2) 采用 FFT 和最大似然估计方法对铱星单音信号能够有效多普勒频移测量,同时给出多历元静态定位接收机频偏补偿的方法。

3) 通过持续接收铱星信号并提取多普勒测量信息,并结合轨道预测模型计算的卫星轨道实现接收机定位,结果表明,在高程辅助情况下,接收机水平定位精度达到 200 m。

本文成果为 GNSS 受到强干扰导致不能工作

的环境下,提供一种有效的定位手段。

### 参考文献 (References)

- [ 1 ] YANG C, NGUYEN T. Self-calibrating position location using signals of opportunities [ C ] // Aerospace and Electronics Conference, 2009; 1055-1063.
- [ 2 ] LAWRENCE D, COBB H S, GUTT G, et al. Test results from a LEO-satellite-based assured time and location [ C ] // Proceedings of the 2016 International Technical Meeting of the Institute of Navigation, 2016; 125-129.
- [ 3 ] JOERGER M, GRATTON L, PERVAN B, et al. Analysis of IRIDIUM-augmented GPS for floating carrier phase positioning [ J ]. Journal of the Institute of Navigation, 2010, 57 ( 2 ) : 137-160.
- [ 4 ] BLOOM J. Eccentric orbits: The Iridium story [ M ]. Atlanta: Atlantic Monthly Press, 2016; 18-20.
- [ 5 ] PRATT S R, RAINES R A, FOSSA C E, et al. An operational and performance overview of the IRIDIUM low earth orbit satellite system [ J ]. IEEE Communications Surveys and Tutorials, 1999, 2(2) : 2-10.
- [ 6 ] PRATT S R, RAINES R A, GUNSCH G H, et al. An overview of the IRIDIUM ( R ) low earth orbit ( LEO ) satellite system [ C ] // National Aerospace and Electronics Conference. Piscataway, NJ: IEEE Press, 1998; 152-159.
- [ 7 ] ICAO. Manual for ICAO aeronautical mobile satellite (ROUTE) service. Part2-Iridium. Draft v4. 0 [ Z ]. Montreal: ICAO, 2007; 9-12.
- [ 8 ] 闵士权. 卫星通信系统工程设计与应用 [ M ]. 北京: 电子工业出版社, 2015; 352-361.
- MIN S Q. Satellite communication system engineering design and application [ M ]. Beijing: Electronics Industry Press, 2015; 352-361 ( in Chinese ).
- [ 9 ] SHAHRIAR C M R. A scheme to mitigate interference from Iridium satellite downlink signal captured by omnidirectional antenna array [ C ] // IEEE Antennas and Propagation Society International Symposium. Piscataway, NJ: IEEE Press, 2008; 1-4.
- [ 10 ] VAN DIGGELEN F. A-GPS: Assisted GPS, GNSS, and SBAS [ M ]. Boston: Artech House, 2009; 261-264.
- [ 11 ] COBB S, LAWRENCE D, GUTT G, et al. Differential and rubidium disciplined test results from an Iridium-based secure timing solution [ C ] // International Technical Meeting of the Institute of Navigation. Manassas: ION, 2017, 2; 1111-1116.
- [ 12 ] BOYCE III W H. Examination of NOARD TLE accuracy using the Iridium constellation [ J ]. Advances in the Astronautical Sciences, 2005, 119(2) : 2133-2141.
- [ 13 ] KAY S M. Fundamentals of statistical signal processing: Estimation theory [ M ]. Upper Saddle River: Prentice-Hall Inc., 1993; 193-197.
- [ 14 ] 张守信. GPS 卫星测量定位理论与应用 [ M ]. 长沙: 国防科技大学出版社, 1996; 86-87.
- ZHANG S X. Theory and application of GPS satellite measurement and positioning [ M ]. Changsha: National University of Defense Technology Press, 1996; 86-87 ( in Chinese ).
- [ 15 ] 王萌, 马利华, 张丽荣, 等. 区域定位系统中高程辅助三星定位算法 [ J ]. 上海交通大学学报, 2012, 46(10) : 1647-1652.
- WANG M, MA L H, ZHANG L R, et al. Three-satellite positioning algorithm with altitude aiding for regional navigation satellite system [ J ]. Journal of Shanghai Jiaotong University, 2012, 46(10) : 1647-1652 ( in Chinese ).

#### 作者简介:

秦红磊 男, 教授, 博士生导师。主要研究方向: 机会信号定位、相对导航、室内定位及组合导航等。

谭滋中 男, 博士研究生。主要研究方向: 卫星导航定位、低轨卫星定位及机会信号定位等。

丛丽 女, 博士, 副教授, 硕士生导师。主要研究方向: 组合导航定位技术、机会信号定位及室内定位等。

## Positioning technology based on IRIDIUM signals of opportunity

QIN Honglei, TAN Zizhong, CONG Li\*, ZHAO Chao

(School of Electronic and Information Engineering, Beihang University, Beijing 100083, China)

**Abstract:** The global navigation satellite system (GNSS) has a number of disadvantages such as weak signal strength on the ground, high infrastructure requirement and easy to be interfered, while signals of opportunity can make up these shortages. Space-based signals of opportunity positioning based on IRIDIUM is implemented for insufficiency of the coverage and availability of ground-based signals of opportunity such as FM, DTV and mobile base station. This paper realizes the function of positioning based on IRIDIUM through combining Doppler shift information which is acquired with the tone signal of IRIDIUM signal and orbital information of satellite calculated from orbit prediction model after establishing mathematical model of instantaneous Doppler positioning by analyzing the communication system of IRIDIUM. The experimental results based on actual signal show that the positioning accuracy can reach better than 200 m. The research of this paper is significant in theoretical research and practice application of positioning technology based on signals of opportunity.

**Keywords:** signals of opportunity; IRIDIUM; space-based; communication; instantaneous Doppler positioning

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

DOI: 10.13700/j.bh.1001-5965.2018.0781

# 含内热源的多孔方腔流热耦合非正交 MRT-LBM 数值模拟

张莹<sup>1</sup>, 黄逸宸<sup>1</sup>, 陈岳<sup>1</sup>, 马明<sup>2</sup>, 李培生<sup>1,\*</sup>, 王昭太<sup>1</sup>

(1. 南昌大学 机电工程学院, 南昌 330031; 2. 美国圣母大学 航空机械系, 南本德 46556)



**摘 要:** 针对含内热源的多孔方腔内自然对流现象问题, 采用非正交多弛豫时间 (MRT) 格子 Boltzmann 方法进行了研究。分析了 Rayleigh 数 ( $10^4 \leq Ra \leq 10^6$ )、内热源布局方式 (水平、垂直及对角布局)、内热源几何尺寸大小 ( $A = 1/16, 1/8, 3/16, 1/4$ ) 及两内热源间的间距 ( $S = 5/64, 13/64, 21/64$ ) 对流动传热的影响。结果表明: 在  $Ra = 10^4, 10^5$  和  $S = 5/64$  的情况下, 任意内热源几何尺寸, 内热源采用对角布局方式可获得更好的对流换热效果; 在  $Ra = 10^5, 10^6$  和  $S = 13/64, 21/64$  的情况下, 水平布局方式更优; 在内热源采用水平布局,  $Ra = 10^4$  的情况下, 任意内热源几何尺寸, 对流换热效果均随着内热源间距的增大而增强; 而随着  $Ra$  增大, 内热源几何尺寸减小, 对流换热效果随着内热源间距的增大先增大后减小, 而后随着内热源间距增大其对流换热效果减弱; 对角布局也有相似规律, 在其他条件一致的情况下, 随着内热源几何尺寸的增加, 其对流换热效果增强。

**关键词:** 多孔方腔; 内热源; 自然对流; Nusselt 数; 多弛豫时间 (MRT); Boltzmann 模型

**中图分类号:** TB61<sup>+</sup>1

**文献标识码:** A

**文章编号:** 1001-5965(2019)09-1700-13

多孔介质中的流体流动与传热在化工、机械、能源、环境等诸多科学与工程领域有着广泛的应用, 如化工分离过程、化学污染物在饱和土壤中的分散、地热储层流体流动、热交换器的优化设计、电子器件及设备的冷却、房间通风和太阳能集热器等<sup>[1-2]</sup>。因此, 对多孔方腔内的流体流动和传热特性进行深入研究具有重要意义。近年来, 国内外学者采用数值方法对多孔方腔内的对流传热问题进行了研究。Saeid<sup>[3]</sup>采用有限容积法研究了受热壁面温度在空间上呈正弦变化的多孔腔内的自然对流传热; Roy 和 Basak<sup>[4]</sup>采用有限元法研究了壁面均匀和非均匀加热边界的方腔内自然对流传热, 表明壁面均匀和非均匀加热对方腔壁

面的传热速率有较大的影响。

近年来, 许多学者采用的都是传统的数值模拟方法, 该方法受数值稳定性、计算精度等条件的制约, 难以满足微尺度传热流动、多孔介质等领域研究的要求。格子 Boltzmann 方法 (LBM) 是近年来新兴发展起来的介观模拟方法, 其基于分子运动理论, 与传统方法相比, 具有程序结构简单、并行计算特性良好、边界条件处理简单等特点<sup>[5-6]</sup>, 因此在求解多孔介质领域内的传热流动问题方面具有很好的应用。目前, 已有许多学者采用 LBM 方法研究多孔腔内的自然对流传热。Zhao 等<sup>[7]</sup>采用双种群热格子 BGK 模型模拟多孔介质中的二维自然对流传热, 研究了孔隙密度、孔隙度及多

收稿日期: 2019-01-02; 录用日期: 2019-03-22; 网络出版时间: 2019-06-10 15:45

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190606.1552.005.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190606.1552.005.html)

基金项目: 国家自然科学基金 (51566012, 11562011); 江西省自然科学基金 (20181BAB206031); 江西省研究生创新专项基金 (YC2019-S016)

\* 通信作者. E-mail: nucdns1995z@163.com

**引用格式:** 张莹, 黄逸宸, 陈岳, 等. 含内热源的多孔方腔流热耦合非正交 MRT-LBM 数值模拟[J]. 北京航空航天大学学报, 2019, 45(9): 1700-1712. ZHANG Y, HUANG Y C, CHEN Y, et al. Non-orthogonal multiple-relaxation-time lattice Boltzmann method for numerical simulation of thermal coupling with porous square cavity flow containing internal heat source [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1700-1712 (in Chinese).

孔介质结构(形状)对自然对流的影响;Dixit 和 Babu<sup>[8]</sup>采用基于 BGK 模型的 LBM 方法研究了高 Rayleigh 数下的方腔内自然对流传热;Liu 等<sup>[9]</sup>采用多弛豫时间(MRT)模型模拟了垂直受热壁面多孔方腔内自然对流现象;Liu 等<sup>[10]</sup>基于广义非达西(Darcy)模型提出了一种非正交 MRT-LBM 方法,研究了多孔介质在典型初等体积尺度下的对流换热,结果表明,非正交 MRT-LBM 模型能够正确模拟对流换热过程,且比格子 BGK(LBGK)模型具有更好的数值稳定性。

对于方腔内自然对流的研究,大多数学者集中在处理垂直壁面加热的情况,而部分学者研究了方腔内含内热源、壁面冷却的情况。Lam 和 Prakash<sup>[11]</sup>采用有限元法对多孔腔体中顶部及底部放置热源的腔内自然对流现象进行了模拟研究;李培生等<sup>[12]</sup>采用 LBM 方法研究了内置高温体倾斜多孔腔体中的流体流动与传热机理,结果表明,孔隙率、达西数、Raleigh 数及倾角对流体与热壁面之间的自然对流传热能力有较大影响;Sivavashi 等<sup>[13]</sup>采用有限容积法研究了两相流在内热源与溶质源作用下通过充满流体饱和多孔介质的方形封体的稳态双扩散自然对流,分析了腔体倾角,内热源的不同形状、位置和排列对流动传热传质特性的影响,并从热力学第二定律出发,进行了熵产分析,确定了最佳的内热源构型;Selimefendigil<sup>[14]</sup>采用 CFD 方法研究了方形腔内两内热源的自然对流冷却过程,分析了在腔体顶部、中部和底部放置内热源,以及两内热源之间的距离对传热特性的影响。基于前人的工作,本文采用非正交 MRT 方法研究了多孔方腔内含双内热源的自然对流冷却过程,讨论分析了在固定孔隙率  $\varepsilon = 0.4$ ,普朗特数  $Pr = 0.71$ ,达西数  $Da = 10^{-2}$ ,不同 Raleigh 数情况下,内热源在多孔方腔水平中心、垂直中心、对角位置布局对冷、热壁面平均努塞尔数的影响,并在水平及对角布局的方

式下讨论分析了内热源几何尺寸大小及两内热源之间的间距对传热特性的影响,从而得出内热源布局方式、内热源几何尺寸大小及间距对换热效果的影响规律,为电子器件及设备的冷却等实际工程应用领域提供帮助。

## 1 问题描述及计算模型

物理模型如图 1 所示。二维封闭方腔内放置布局方式为水平(Case 1)、垂直(Case 2)、对角(Case 3)的两发热方块,方块表面温度为  $T_h$ ,边长为  $H$ ,两方块间的距离为  $d$ 。外部方腔左右边壁温度为  $T_c$  ( $T_h > T_c$ ),上下边壁绝热,边长为  $L$ ,各壁面边界上的速度梯度均为零。内置方块与方腔的长度比定义为  $A = H/L$ ,两方块间的距离与方腔的长度比定义为  $S = d/L$ 。发热方块与腔体之间填充了各向同性、均质的多孔材料。本文不对内热源赋值,仅采用放置发热方块来代替内热源,在此前提下研究方腔内的流体流动及传热现象。

假设流体不可压缩,考虑温差引起的浮升力项,并忽略黏性耗散,引入 Boussinesq 假设,采用修正的 Brinkman-Darcy-Forchheimer 渗流模型描述腔体内的复杂流动,结合能量方程,二维多孔方腔内流体流动及传热的广义控制方程为

$$\nabla \cdot \mathbf{u} = 0 \quad (1)$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \left( \frac{\mathbf{u}}{\varepsilon} \right) = -\frac{1}{\rho} \nabla(\varepsilon p) + \nu_c \nabla^2 \mathbf{u} + \mathbf{F} \quad (2)$$

$$\sigma \frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T = \nabla(\alpha_c \nabla T) \quad (3)$$

式中: $\rho$ 为流体密度; $\mathbf{u}$ 、 $p$ 和  $T$ 分别为流体的平均速度、压力和温度; $\varepsilon$ 为腔体的孔隙率; $\nu_c$ 为流体的有效运动黏度; $\alpha_c$ 为有效热扩散系数; $\sigma$ 为多孔腔体内固相与液相之间的热容比值; $\mathbf{F}$ 为外力项,是由多孔介质和其他外力所引起的合外力,其可以表示为

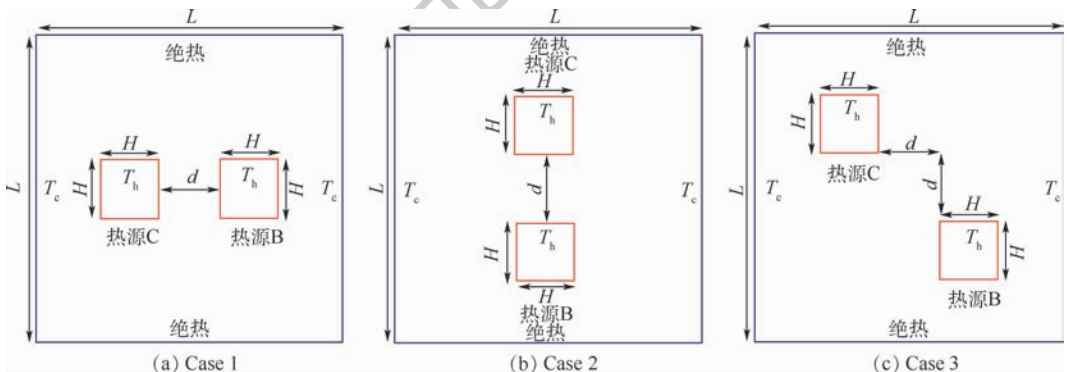


图1 物理模型

Fig. 1 Physical model

$$\mathbf{F} = -\frac{\varepsilon \nu}{K} \mathbf{u} - \frac{\varepsilon F_\varepsilon}{\sqrt{K}} |\mathbf{u}| \mathbf{u} + \varepsilon \cdot \mathbf{G} \quad (4)$$

其中:  $\nu$  为流体的运动黏度;  $K$  为多孔介质的渗透率;  $F_\varepsilon$  为几何形状因子;  $\mathbf{G}$  为浮升力,  $\mathbf{G} = g\beta(T - T_0)\mathbf{j}$ ,  $g$  为重力加速度,  $\beta$  为流体的热膨胀系数,  $T_0$  为方腔内的平均温度,  $\mathbf{j}$  为  $y$  轴的单位矢量。

为了描述多孔方腔内的对流传热特性, 采用的无量纲参数有: 达西数  $Da = K/L^2$ , Raleigh 数  $Ra = g\beta\Delta TL^3/(v\alpha)$ ,  $\Delta T$  为特征温度, 雷诺数  $Re = LU/v$ , 普朗特数  $Pr = v/\alpha$ , 黏度比  $J = \nu_e/\nu$ , 热扩散率比  $\gamma = \alpha_e/\alpha$ 。其中,  $\alpha$  为热扩散率,  $U$  为特征速度。为了量化传热效率, 本文采用平均努塞尔数来表达, 垂直壁面与水平壁面的平均努塞尔数为

$$\begin{cases} Nu_{\text{avg}} = -\frac{1}{\Delta T} \int_0^L \left(\frac{\partial T}{\partial x}\right) dy \\ Nu_{\text{havg}} = -\frac{1}{\Delta T} \int_0^L \left(\frac{\partial T}{\partial y}\right) dx \end{cases} \quad (5)$$

式中:  $Nu_{\text{avg}}$  和  $Nu_{\text{havg}}$  分别为垂直和水平壁面的平均努塞尔数。

## 2 耦合双分布非正交 MRT-LBM 模型

### 2.1 流场

LBM 方法是在粒子间的相互运动下, 通过构建粒子分布函数来描述每个粒子处于某一状态下的概率, 并通过统计方法得出系统的宏观参数。非正交 MRT-LBM 模型速度空间的密度分布演化方程为

$$f_i(x + \mathbf{e}\delta_t, t + \delta_t) - f_i(x, t) = -\mathbf{M}^{-1} \cdot \left\{ \mathbf{S} \cdot [\mathbf{m} - \mathbf{m}^{\text{eq}}]_{x,t} - \delta_t \left( \mathbf{I} - \frac{\mathbf{S}}{2} \right) \cdot \mathbf{F}_{x,t} \right\}_i \quad (6)$$

式中:  $f_i$  为密度分布函数;  $\delta_t$  为单位时间步长;  $\mathbf{e}$  为离散速度的方向矢量;  $\mathbf{M}$  为速度空间的非正交转换矩阵;  $\mathbf{S}$  为速度空间的对角松弛系数矩阵,  $\mathbf{S} = \text{diag}(s_0, s_1, s_2, s_3, s_4, s_5, s_6, s_7, s_8)$ ;  $\mathbf{m}$  和  $\mathbf{m}^{\text{eq}}$  分别为  $f$  和  $f^{\text{eq}}$  对应的矩空间分布函数,  $\mathbf{m} = \mathbf{M} \cdot f$ ,  $\mathbf{m}^{\text{eq}} = \mathbf{M} \cdot f^{\text{eq}}$ ;  $\mathbf{I}$  为单位矩阵;  $\mathbf{F}_{x,t}$  为合外力;  $f$  和  $f^{\text{eq}}$  分别为密度分布函数和密度平衡态分布函数。

速度空间采用 D2Q9 离散速度模型, 各离散速度的方向矢量为

$$\mathbf{e}_i = \begin{cases} (0, 0) & i = 0 \\ (\cos[(i-1)\pi/2], \sin[(i-1)\pi/2])c & i = 1, 2, 3, 4 \\ (\cos[(2i-1)\pi/2], \sin[(2i-1)\pi/2])c & i = 5, 6, 7, 8 \end{cases} \quad (7)$$

式中:  $c$  为格子单位速度。

非正交转换矩阵  $\mathbf{M}$  为

$$\mathbf{M} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & -1 & 0 & 1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 & -1 & 1 & 1 & -1 & -1 \\ 0 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 \\ 0 & 1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (8)$$

矩空间平衡态分布函数  $\mathbf{m}^{\text{eq}}$  为

$$\begin{cases} m_0^{\text{eq}} = \rho, m_1^{\text{eq}} = \rho u_x, m_2^{\text{eq}} = \rho u_y \\ m_3^{\text{eq}} = \frac{2}{3}\rho + \frac{\rho(u_x^2 + u_y^2)}{\varepsilon} \\ m_4^{\text{eq}} = \frac{\rho(u_x^2 - u_y^2)}{\varepsilon}, m_5^{\text{eq}} = \frac{\rho u_x u_y}{\varepsilon} \\ m_6^{\text{eq}} = \frac{\rho u_y}{3}, m_7^{\text{eq}} = \frac{\rho u_x}{3} \\ m_8^{\text{eq}} = \frac{\rho}{9} + \frac{\rho(u_x^2 + u_y^2)}{3\varepsilon} \end{cases} \quad (9)$$

式中:  $u_x$  和  $u_y$  分别为节点的水平 and 垂直方向速度。

速度空间的密度平衡态分布函数  $f_i^{\text{eq}}$  为

$$f_i^{\text{eq}} = \omega_i \rho \left[ 1 + \frac{\mathbf{e}_i \cdot \mathbf{u}}{C_s^2} + \frac{\mathbf{u}\mathbf{u} : (\mathbf{e}_i \mathbf{e}_i - C_s^2 \mathbf{I})}{2\varepsilon C_s^4} \right] \quad (10)$$

式中:  $\omega_i$  为权重系数,  $\omega_0 = 4/9$ ,  $\omega_1 = \omega_2 = \omega_3 = \omega_4 = 1/9$ ,  $\omega_5 = \omega_6 = \omega_7 = \omega_8 = 1/36$ ;  $C_s$  为格子声速,  $C_s = c/\sqrt{3}$ 。

合外力  $\mathbf{F}$  为

$$\begin{cases} F_0 = 0, F_1 = \rho F_x, F_2 = \rho F_y \\ F_3 = \frac{2\rho(u_x F_x + u_y F_y)}{\varepsilon} \\ F_4 = \frac{2\rho(u_x F_x - u_y F_y)}{\varepsilon}, F_5 = \frac{\rho(u_x F_y + u_y F_x)}{\varepsilon} \\ F_6 = \frac{\rho F_y}{3}, F_7 = \frac{\rho F_x}{3}, F_8 = \frac{2\rho(u_x F_y + u_y F_x)}{3\varepsilon} \end{cases} \quad (11)$$

式中:  $F_x$ 、 $F_y$  分别为合外力  $\mathbf{F}$  的  $x$ 、 $y$  方向分力。

有效运动黏度  $\nu_e$  可以表示为

$$\nu_e = C_s^2(\tau_v - 0.5)\delta_t \quad (12)$$

式中:  $\tau_v$  为计算密度场分布函数的弛豫时间。

非正交 MRT-LBM 方法同常规的 LBM 方法一样, 分布函数的演化主要由 2 部分组成: 碰撞和迁移。矩空间进行的碰撞过程为

$$\mathbf{m}^*(x, t) = \mathbf{m}(x, t) - \mathbf{S} \cdot$$

$$(\mathbf{m} - \mathbf{m}^{eq})_{(x,t)} + \delta_i \left( \mathbf{I} - \frac{\mathbf{S}}{2} \right) \mathbf{F} \quad (13)$$

式中:  $\mathbf{m}^*$  为碰撞后的矩空间分布函数。

完成碰撞后,通过  $f^* = \mathbf{M}^{-1} \cdot \mathbf{m}^*$  变换回速度空间,然后执行迁移过程:

$$f_i(x + \mathbf{e}_i \delta_i, t + \delta_i) = f_i^*(x, t) \quad (14)$$

得到更新的分布函数后,流体的宏观密度  $\rho$  和速度  $\mathbf{u}$  可以通过引入临时速度  $\mathbf{v}$  求解得到

$$\begin{cases} \mathbf{v} = \sum_{i=0}^8 \mathbf{e}_i f_i / \rho + \frac{\delta_i}{2} \varepsilon \mathbf{G} \\ \mathbf{u} = \frac{\mathbf{v}}{l_0 + \sqrt{l_0^2 + l_1 |\mathbf{v}|}} \end{cases} \quad (15)$$

$$\begin{cases} \rho = \sum_{i=0}^8 f_i \\ l_1 = \frac{\varepsilon \delta_i F_e}{2\sqrt{K}} \\ l_0 = \frac{1 + \varepsilon \delta_i \mathbf{v} / (2K)}{2} \end{cases} \quad (16)$$

## 2.2 温度场

二维温度场可以采用多种离散速度模型计算,本文采用计算量相对较小的 D2Q5 速度模型,该模型已足够对二维温度场进行准确模拟。二维温度场分布函数的演化方程为

$$g_i(x + \mathbf{e}_i \delta_i, t + \delta_i) - g_i(x, t) = -(\mathbf{N}^{-1} \mathbf{Q} \mathbf{N})_{ij} \cdot (g_i - g_i^{eq})_{(x,t)} \quad (17)$$

式中:  $g_i$  为温度分布函数;  $g_i^{eq}$  为温度平衡态分布函数;  $\mathbf{Q}$  为温度场对角松弛系数矩阵,  $\mathbf{Q} = \text{diag}(\zeta_0, \zeta_1, \zeta_2, \zeta_3, \zeta_4)$ ;  $\mathbf{N}$  为温度场非正交转换矩阵;  $\mathbf{e}_i$  为温度场各离散速度的方向矢量,表示为

$$\mathbf{e}_i = \begin{cases} (0, 0) & i = 0 \\ \left( \cos\left[ (i-1) \frac{\pi}{2} \right], \sin\left[ (i-1) \frac{\pi}{2} \right] \right) c & i = 1, 2, 3, 4 \end{cases} \quad (18)$$

求解温度场的 MRT-LBM 模型的演化过程与求解速度场类似,其宏观温度  $T$  可以表示为:  $\sigma T =$

$$\sum_{i=0}^4 g_i$$

非正交转换矩阵  $\mathbf{N}$  为

$$\mathbf{N} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 1 & -1 \end{bmatrix} \quad (19)$$

矩空间平衡态分布函数  $\mathbf{n}^{eq}$  为

$$\begin{cases} n_0^{eq} = \sigma T, n_1^{eq} = u_x T, n_2^{eq} = u_y T \\ n_3^{eq} = \varphi \sigma T, n_4^{eq} = 0 \end{cases} \quad (20)$$

式中:  $\varphi$  为常值。

速度空间的温度平衡态分布函数  $g_i^{eq}$  为

$$g_i^{eq} = \begin{cases} (1 - \varphi) \sigma T & i = 0 \\ \frac{\varphi \sigma T}{4} + \frac{(\mathbf{e}_i \cdot \mathbf{u}) T}{2} & i = 1, 2, 3, 4 \end{cases} \quad (21)$$

有效热扩散系数可表示为

$$\alpha_e = \frac{\sigma(\tau_T - 0.5)\delta_i}{4} \quad (22)$$

式中:  $\tau_T$  为计算温度场的分布函数的弛豫时间。

## 2.3 数值设定

在模拟中,设置了  $\varphi = 1/2$ ,  $C_{sT}^2 = c^2 \varphi / 2 = 1/4$ ,  $J = 1$ ,  $\sigma = 1$ ,  $\gamma = 1$ , 其中,  $C_{sT}^2$  为 D2Q5 模型的格子声速,非正交 MRT-LBM 模型的矩空间对角松弛系数为

$$\begin{cases} s_0 = s_1 = s_2 = 1, s_3 = 1.6, s_4 = s_5 = 1/\tau_v \\ s_6 = s_7 = 1.2, s_8 = 1.8 \\ \zeta_0 = 1, \zeta_1 = \zeta_2 = 1/\tau_T, \zeta_3 = \zeta_4 = 1.5 \end{cases} \quad (23)$$

式中:弛豫时间分别为

$$\begin{cases} \tau_v = 0.5 + \frac{MaJL}{c\delta_i} \sqrt{\frac{3Pr}{Ra}} \\ \tau_T = 0.5 + \frac{\gamma C_{sT}^2 (\tau_v - 0.5)}{\sigma J C_{sT}^2 Pr} \end{cases} \quad (24)$$

## 3 数值模拟结果与分析

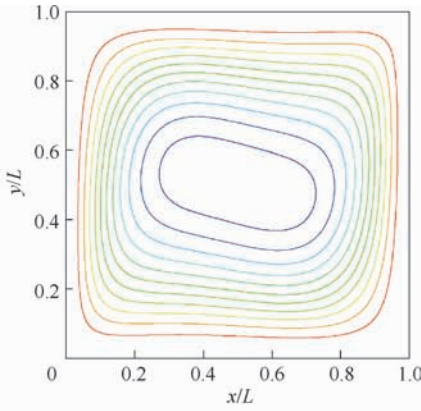
### 3.1 模型验证

为了验证数值结果的准确性,模拟了填充有饱和多孔介质的方腔内自然对流传热,方腔内无放置方块,左壁面为高温,右壁面为低温。通过计算热壁面平均努塞尔数 ( $Nu_{avg}$ ) 与文献[15]的结果进行比较,结果如表 1 所示。在孔隙率  $\varepsilon = 0.4$ ,  $Ra = 10^5$ ,  $Da = 10^{-2}$ ,  $Pr = 1.0$  的情况下,与文献[10]比较了流线和等温线图,结果如图 2 所示。由表 1 可以得出,本文非正交 MRT-LBM 方法

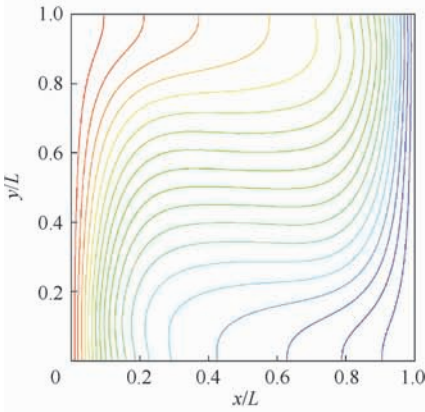
表 1 热壁面平均努塞尔数比较

Table 1 Comparison of average Nusselt number at hot sidewall

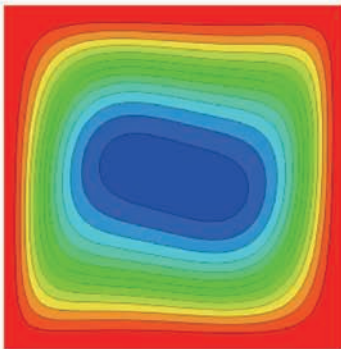
Da	Ra	Pr	$\varepsilon$	$Nu_{avg}$		相对误差/%
				本文	文献[15]	
$10^{-2}$	$10^3$	1.0	0.4	1.0332	1.01	2.30
$10^{-2}$	$10^4$	1.0	0.4	1.4302	1.408	1.58
$10^{-2}$	$10^5$	1.0	0.4	3.0334	2.983	1.69
$10^{-4}$	$10^6$	1.0	0.6	2.7476	2.725	0.83
$10^{-4}$	$10^7$	1.0	0.9	9.2579	9.202	0.61



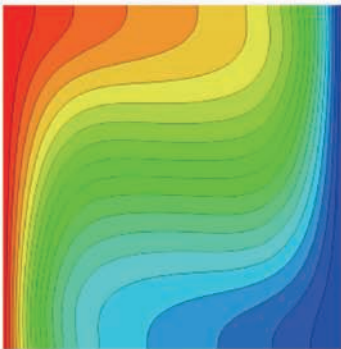
(a) 文献[10]流线图



(b) 文献[10]等温线图



(c) 本文流线图



(d) 本文等温线图

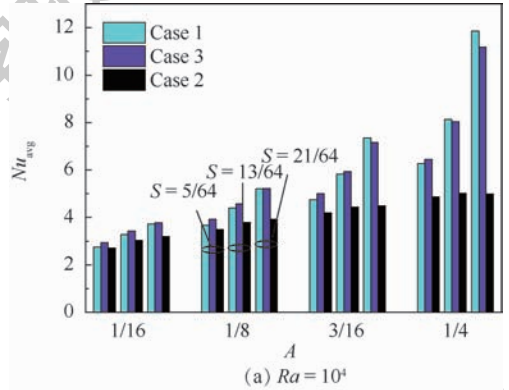
图2 流线图与等温线图比较 ( $Ra = 10^5, Da = 10^{-2}, Pr = 1.0, \varepsilon = 0.4$ )

Fig. 2 Comparison of streamlines and isotherms ( $Ra = 10^5, Da = 10^{-2}, Pr = 1.0, \varepsilon = 0.4$ )

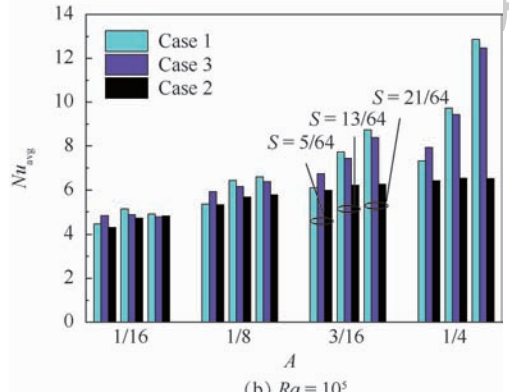
模拟的数值结果与文献[15]相近, 相对误差都在2.5%以内, 充分验证了非正交MRT-LBM方法模拟多孔方腔内自然对流传热的可靠性。

### 3.2 内热源布局方式对流动及传热的影响

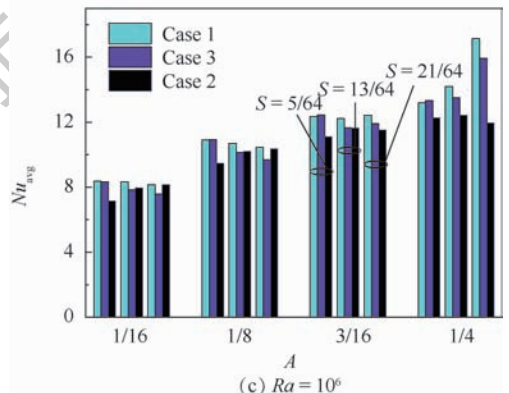
图3和图4显示了在不同内热源几何尺寸大小A及间距S下, 布局方式对 $Ra = 10^4, 10^5, 10^6$ 下热、冷壁面平均努塞尔数的影响。图5~图7的流线图与等温线图给出了最大流函数值与平均努塞尔数, 分别为 $\Psi_{\max\text{right}}$  (右漩涡)、 $\Psi_{\max\text{left}}$  (左漩涡)、 $\Psi_{\max\text{down}}$  (下漩涡)、 $\Psi_{\max\text{up}}$  (上漩涡)、 $\Psi_{\max\text{side}}$



(a)  $Ra = 10^4$



(b)  $Ra = 10^5$



(c)  $Ra = 10^6$

图3 不同Raleigh数、内热源几何尺寸大小及间距下布局方式对热壁面平均努塞尔数的影响

Fig. 3 Influence of layout mode on heated wall average Nusselt number with different Rayleigh numbers, heat source sizes and spacing

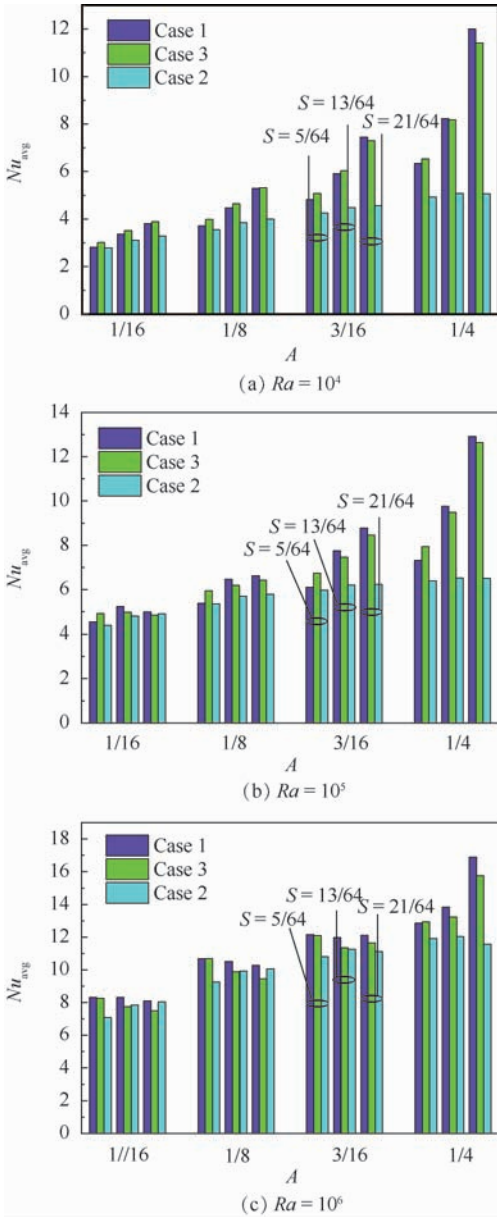


图4 不同 Rayleigh 数、内热源几何尺寸大小及间距下布局方式对冷壁面平均努塞尔数的影响

Fig. 4 Influence of layout mode on cold wall average Nusselt number with different Rayleigh numbers, heat source sizes and spacing

(侧漩涡)最大流函数值及  $Nu_{avghot}$  (热壁面)、 $Nu_{avgcool}$  (冷壁面)平均努塞尔数。由于热壁面平均努塞尔数受多个因素的综合影响,因此本文中对主要的影响因素做出了分析。从图3和图4中可以发现,在  $Ra = 10^4$ , 两内热源间距  $S = 5/64$  时,任意内热源几何尺寸大小,内热源采用对角布局都要比水平及垂直布局具有更好的换热效果。这是由于水平与垂直布局时,其相间的对流换热区域相比于对角布局的换热区域小,在其相间的等温线分布疏散,表明其相间的方块表面温度梯度低,且在其表面发展起来的热边界层较厚,因此导

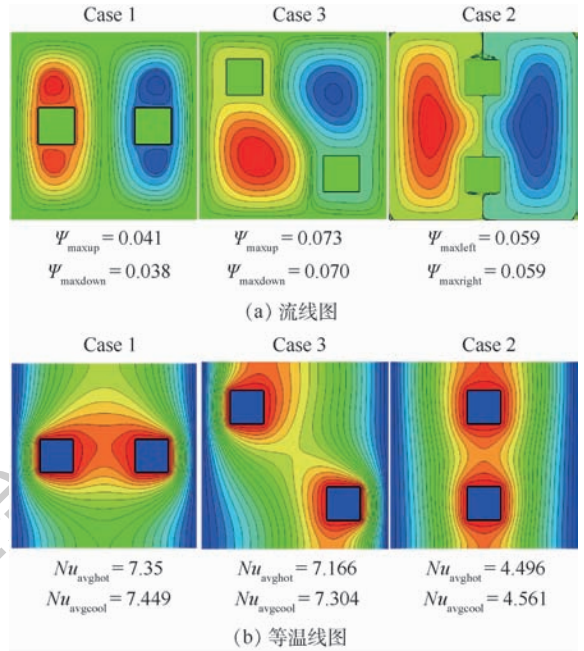


图5  $Ra = 10^4, A = 3/16, S = 21/64$  下3种布局方式的流线图与等温线图

Fig. 5 Streamlines and isotherms for three layout modes at  $Ra = 10^4, A = 3/16, S = 21/64$

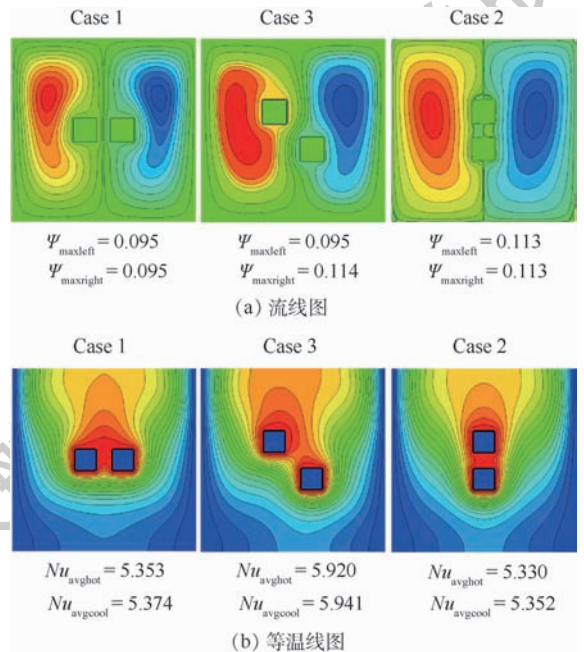


图6  $Ra = 10^5, A = 1/8, S = 5/64$  下3种布局方式的流线图与等温线图

Fig. 6 Streamlines and isotherms for three layout modes at  $Ra = 10^5, A = 1/8, S = 5/64$

致换热不充分,换热效果减弱。同理在  $S = 13/64, A = 1/16, 1/8, 3/16$  情况下,仍然采用对角布局的换热效果更好。但随着内热源几何尺寸大小及间距的增加,其相间的对流换热区域增大,因此对热壁面平均努塞尔数的主要影响转变为在热源

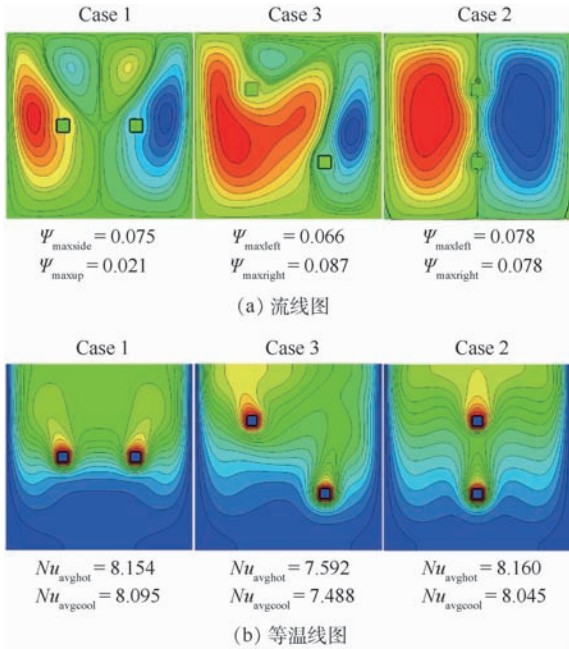


图7  $Ra = 10^6, A = 1/16, S = 21/64$  下3种布局方式的流线图与等温线图

Fig. 7 Streamlines and isotherms for three layout modes at  $Ra = 10^6, A = 1/16, S = 21/64$

B、C周围形成的漩涡对其表面努塞尔数的影响。如图5所示,在 $A = 3/16$ 且 $S = 21/64$ 时,水平布局下内热源B、C的周围产生了4个漩涡,且均铺展其上下表面,因此在相应的热源表面形成的热边界层较薄,温度梯度较陡,聚集了较密集的等温线;而在垂直及对角布局时,仅在周围产生了2个漩涡,且其位置相对于热源的表面比之水平布局时较远,尽管漩涡强度较之水平布局时更强,但综合影响即此时采用水平布局换热效果更好。

当 $Ra = 10^5$ 时,对流换热效果增强,因此在热源B、C周围形成的漩涡对其表面努塞尔数的影响对总的平均努塞尔数的影响增强,而在水平与垂直布局时,热源B、C之间相间区域大小的影响减弱,但在内热源间的间距 $S$ 小的情况下,其仍然是主要影响,如图6所示。分析过程与上述一致,因此在 $Ra = 10^5, S = 5/64$ 时,任意内热源几何尺寸大小,采用对角布局可获得最佳换热效果,在 $S = 13/64, 21/64$ 时,则采用水平布局更优。

当 $Ra$ 增加为 $10^6$ 时,腔内以对流换热为主导机制。此时在热源B、C周围形成的漩涡对其表面努塞尔数的影响对总的平均努塞尔数的影响显著增强,而水平与垂直布局时,热源B、C之间相间区域大小的影响显著减弱。因此,如图3、图4所示,在两热源间的间距 $S = 5/64$ 时,任意内热源几何尺寸大小下,水平及对角布局的换热效果相差无几,因此均可采用这2种布局方式;在 $S =$

$13/64, 21/64$ 时,任意内热源几何尺寸大小下,则采用水平布局能够获得最佳换热效果,其中垂直布局在 $A = 1/16, S = 21/64$ 时,其换热效果与水平布局相差无几,因此在此情况下宜采用水平或垂直布局。如图7所示,原因在于:热源在 $A = 1/16, S = 21/64$ 的情况下,垂直布局时在热源B、C周围产生的2个漩涡,因对流换热强度的增加,其漩涡范围增大并将热源表面覆盖,因此在热源B、C的左右表面形成较薄的热边界层,聚集较密集的等温线,温度梯度较陡;而水平布局时,虽然在热源B、C周围产生了4个漩涡,但靠近上壁面的两漩涡的位置距离其表面较远,因此对热源B、C表面的努塞尔数影响不大,并且靠近方腔左右壁面的两漩涡相比垂直布局时漩涡强度更小,且没有较好的覆盖热源B、C表面,产生的热边界层较厚,温度梯度较低。因此综合影响,两者布局方式的换热效果差不多,而对角布局时,从图7中可以看出,虽然热源周围产生了3个漩涡,但因热源C周围的漩涡明显偏离其表面,且向右上角迁移,这在一定程度上削弱了方腔右边壁的漩涡强度,造成相应的热源表面等温线聚集较疏,温度梯度较低,因此综合影响导致了此时的对角布局换热效果最差。

### 3.3 内热源间距对流动及传热的影响

#### 3.3.1 水平布局

如图8所示,在 $Ra = 10^4$ 的情况下,任意内热源几何尺寸大小,热源B、C的对流换热效果均随着热源间间距的增加而增强。这是因为方腔内的传热机制以导热为主,因此随着热源间的间距 $S$ 增大,热源C的左壁面及热源B的右壁面与方腔左右边壁的距离减小,其表面形成的热边界层变薄,温度梯度变得更陡,因此换热效果增强,如图9所示。而热源B、C的上壁面及下壁面因为热源间间距的增大,在热源B、C周围产生的漩涡逐渐向热源B、C上下两表面迁移,且随着内热源几何尺寸大小的增加,漩涡范围增大,如图10所示。在热源B、C上下两表面的漩涡显著增强了对流换热,因此换热效果增强。对于热源B、C相间的区域,其随着热源间间距的增加,对流换热区域增大,对流换热得以充分发展,其相间的热源壁面热边界层变薄,但随着继续增加,热边界层变得稳定,因此其对流换热区域增大的影响减弱,且增大到一定值时,热源B、C相间区域的壁面主要受到热源B、C下壁面漩涡的影响。当 $Ra$ 增加为 $10^5$ 时,其对流换热强度增强,但内热源几何尺寸大小使得随着热源间间距的增大,热源C的左壁

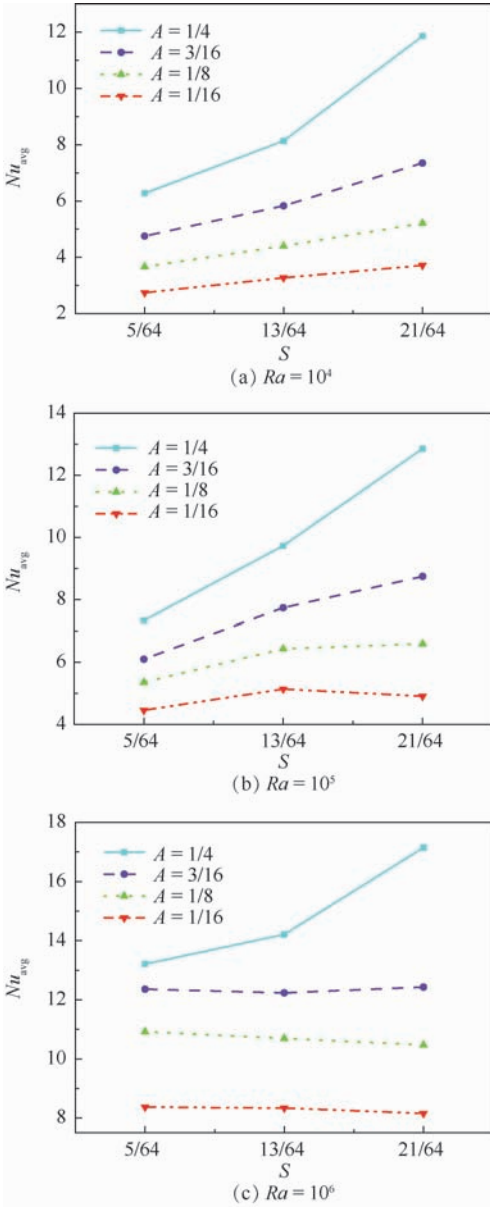


图8 水平布局时,不同 Rayleigh 数、内热源几何尺寸大小及间距对热壁面平均努塞尔数的影响

Fig.8 Influence of different Rayleigh numbers, heat source sizes and spacing on heated wall average Nusselt number in horizontal layout

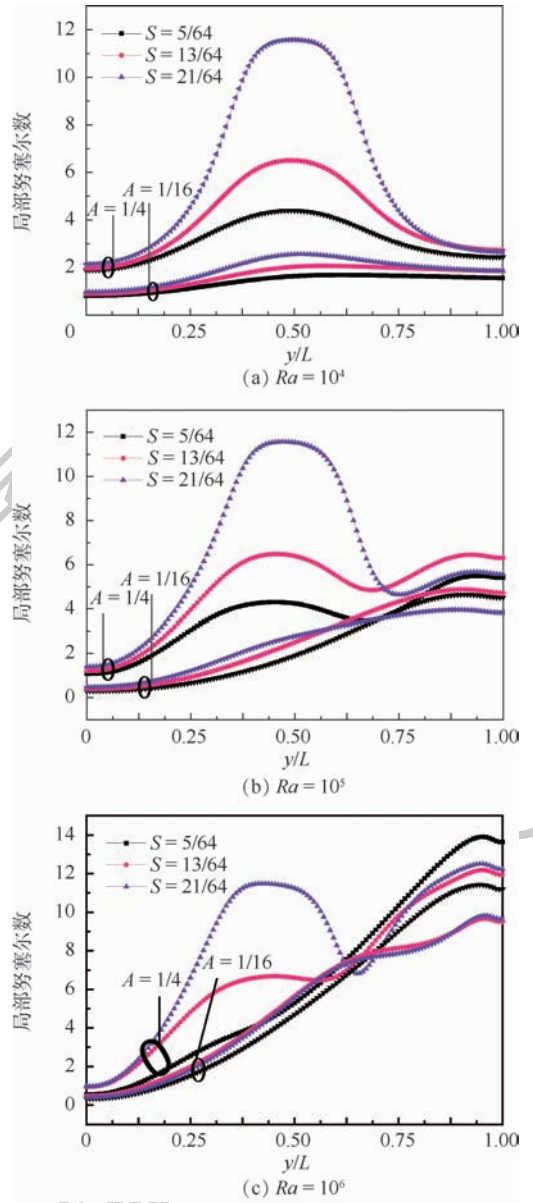


图9 水平布局时,不同 Rayleigh 数、内热源几何尺寸大小及间距下冷壁面局部努塞尔数的分布

Fig.9 Distribution of local Nusselt number on cold wall with different Rayleigh numbers, heat source sizes and spacing in horizontal layout

面及热源 B 的右壁面与方腔左右边壁的距离足够小时,引起的换热效果增强仍是影响总体换热强度的主要因素,因为此时对应的热源表面温度梯度大,如图 9 所示。但随着内热源几何尺寸大小的减小,其影响逐渐减弱,而热源 B、C 周围产生的漩涡的影响转化为主导影响。对于热源 B、C 的上壁面及下壁面因为热源间间距的增大,在其周围产生的漩涡逐渐向其上下两表面迁移,但随着热源间间距的继续增加,漩涡逐渐脱离热源 B、C 的下表面。因此,其上表面形成较薄的热边界层,温度梯度增大,换热增强,而下表面换热减弱。

对于热源 B、C 相间的区域,其分析过程与  $Ra = 10^4$  时一致。因此,综合影响表现为如图 8 所示。在  $Ra = 10^5, A = 1/4, 3/16, 1/8$  时,热源 B、C 的对流换热效果随着热源间间距的增加而增强;在  $A = 1/16$  时,其对流换热效果随着热源间间距的增大先增强后减弱。

当  $Ra$  增加为  $10^6$  时,方腔内的对流换热强度继续增强,并且以对流为传热的主导机制。但当内热源为大尺寸,随着热源间间距的增大使得其左右表面足够接近方腔壁面时,此时热源 B、C 的左右壁面的平均努塞尔数仍然是影响总体换热强

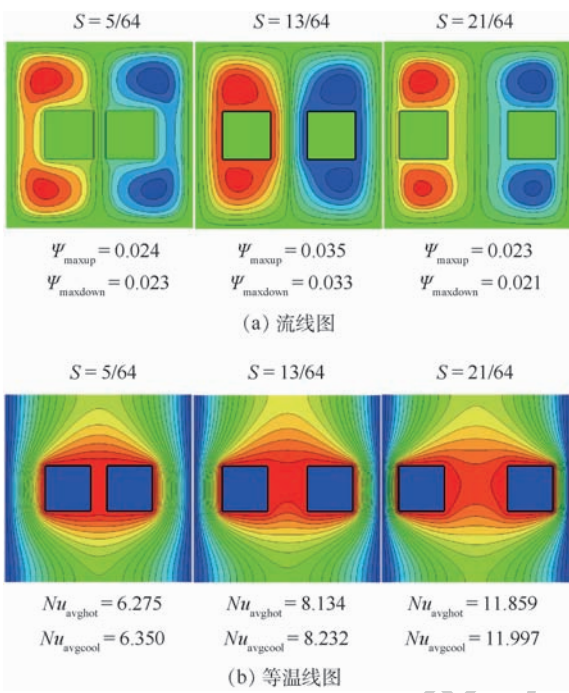


图 10 水平布局时,  $Ra = 10^4$ ,  $A = 1/4$ ,  $S = 5/64$ ,  $13/64, 21/64$  下的流线图与等温线图

Fig. 10 Streamlines and isotherms for horizontal layout at  $Ra = 10^4$ ,  $A = 1/4$ ,  $S = 5/64$ ,  $S = 13/64$ ,  $S = 21/64$

度的主要因素,如图 9 所示,分析过程与上述一致。但随着内热源几何尺寸大小的减小,其周围产生的涡流的影响迅速成为主导影响,随着热源间间距的增大,在其表面形成的涡流逐渐向方腔左右两壁的下表面迁移,且在迁移过程中在方腔上壁表面分裂出 2 个小涡流,小涡流随着间距的增大充分发展,如图 11 所示,因此对于热源 C 的左壁面及热源 B 的右壁面,因涡流迁移导致其涡流强度减弱,表面的热边界层增厚,温度梯度减小,所以在热源间间距小时其热边界层较薄,换热效果更好,并且随着热源间间距的增加,从图 11 可以看出,涡流逐渐脱离热源 B、C 下表面,而对于热源 B、C 相间的区域,其过程与  $Ra = 10^4$  时一致。因此,综合影响表现为如图 8 所示。在  $Ra = 10^6$ ,  $A = 1/4$  时,热源 B、C 的对流换热效果随着热源间间距的增加而增强;在  $A = 3/16$  时,其对流换热效果随着热源间间距的增加先减弱后增强;在  $A = 1/8, 1/16$  时,其对流换热效果随着热源间间距的增大而减弱。

### 3.3.2 对角布局

通过图 12 可以看出,在  $Ra = 10^4$  时,任意内热源几何尺寸大小下,热源 B、C 的对流换热效果均随着热源间间距的增加而增强。这是因为低 Rayleigh 数下,方腔内的传热机制以导热为主,随着热源间间距的增大,热源 B、C 之间相间的区域

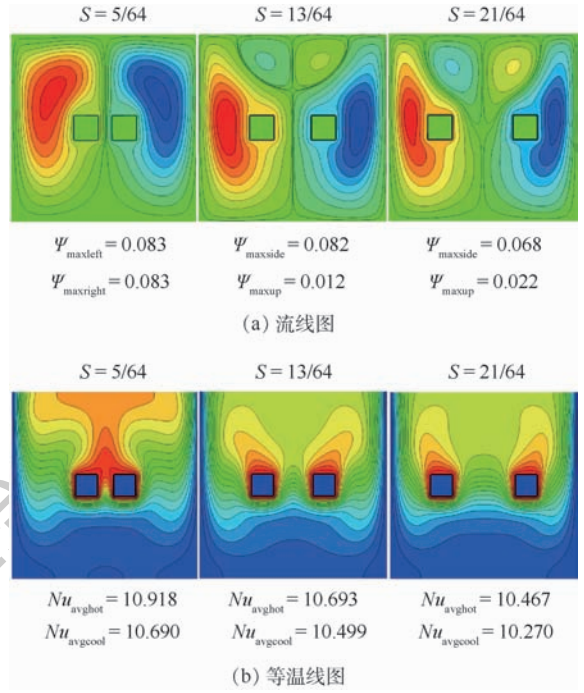


图 11 水平布局时,  $Ra = 10^6$ ,  $A = 1/8$ ,  $S = 5/64$ ,  $13/64, 21/64$  下的流线图与等温线图

Fig. 11 Streamlines and isotherms for horizontal layout at  $Ra = 10^6$ ,  $A = 1/8$ ,  $S = 5/64$ ,  $S = 13/64$ ,  $S = 21/64$

范围增大,对流换热得以充分发展,因此相应表面发展的热边界层变薄,温度梯度变陡;并且热源 C 的左壁面及热源 B 的右壁面与方腔左右边壁的距离减小,温度梯度增大,两壁面的换热增强。如图 13 所示,因而此情况下,换热效果随着热源间间距的增加而增强。

当  $Ra = 10^5$  时,腔内对流换热效果增强,此时从图 12 中可以看出,内热源几何尺寸大小  $A = 1/4, 3/16, 1/8$  时,其表面对流换热效果随着热源间间距的增加而增强;在  $A = 1/16$  时,热源 B、C 的对流换热效果随着  $S$  的增大而减弱。这是因为,在  $A = 1/4, 3/16, 1/8$  时,随着热源间间距的增大,热源 C 的左壁面及热源 B 的右壁面与方腔左右边壁的距离减小,温度梯度增大,两壁面的换热增强,但随着内热源几何尺寸的减小,此影响减弱,而随之增强的是热源 B、C 周围形成的涡流对表面平均努塞尔数的影响,随着热源间间距的增大,热源 C 周围形成的涡流逐渐向方腔左下角迁移因而逐渐脱离其左表面和下表面,而热源 B 周围形成的涡流则逐渐向方腔右上角迁移而逐渐脱离其下表面和右表面,导致这些表面形成的热边界层变厚,温度梯度减小。因而综合影响出现上述现象。

当  $Ra = 10^6$  时,腔内以对流为传热的主导机制,此时从图 12 中可以看出,内热源几何尺寸大

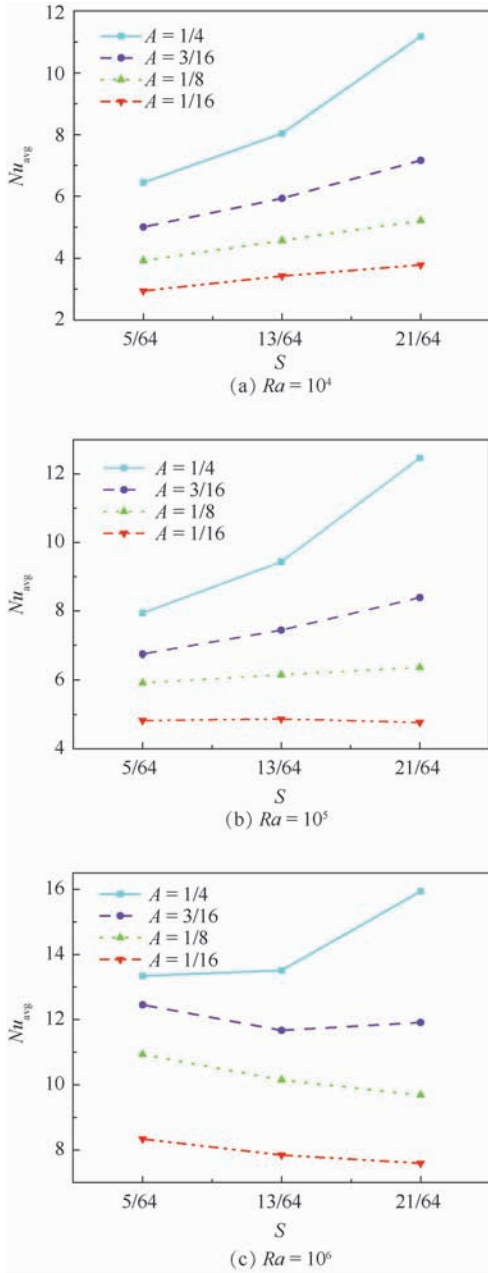


图 12 对角布局时,不同 Rayleigh 数、内热源几何尺寸大小及间距下对热壁面平均努塞尔数的影响

Fig. 12 Influence of different Rayleigh numbers, heat source sizes and spacing on heated wall average Nusselt number in diagonal layout

小  $A = 1/4$  时,其对流换热效果随着热源间间距的增加而增强;在  $A = 3/16$  时,其对流换热效果随着热源间间距的增加先减弱后增强;在  $A = 1/8, 1/16$  时,其对流换热效果随着热源间间距的增大而减弱。这是由于在以对流为传热主导的因素下,只有当内热源几何尺寸大小足够大且使得热源间间距增加时,热源 C 的左表面与热源 B 的右表面足够接近方腔壁面,此时引起的换热增强为对热源总体平均努塞尔数大小的主要影响,如图 13 所示。其分析过程与上述一致。而随着内热源几何

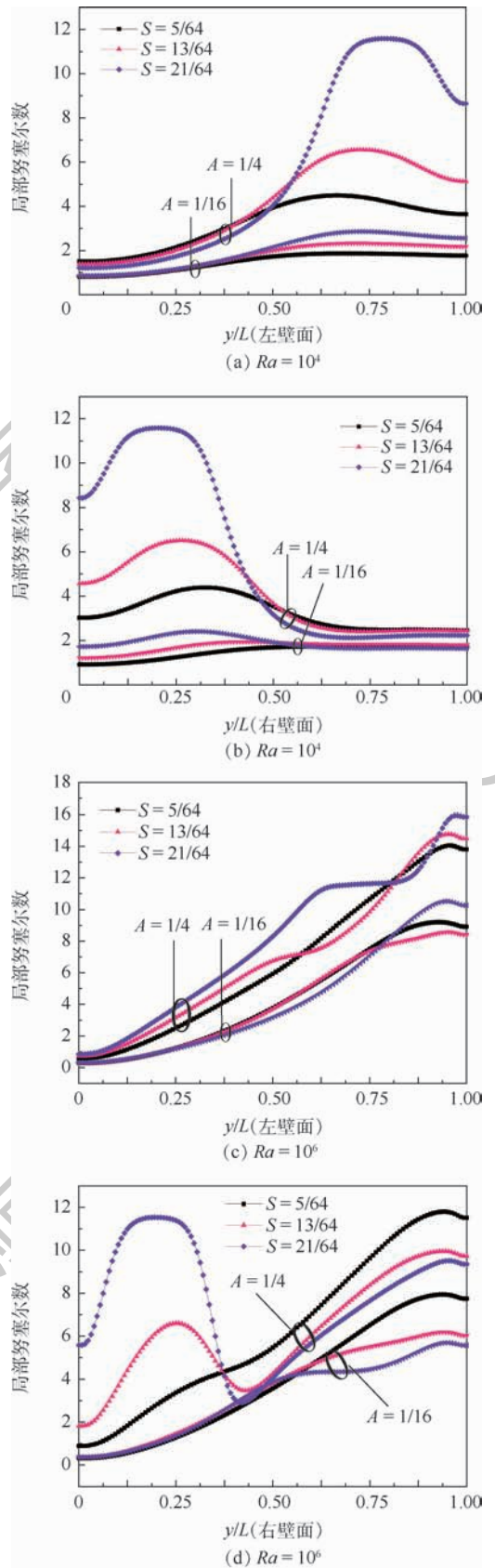


图 13 对角布局时,不同 Rayleigh 数、热源几何尺寸大小及间距下冷壁面局部努塞尔数的分布

Fig. 13 Distribution of local Nusselt number on cold wall with different Rayleigh numbers, heat source sizes and spacing in diagonal layout

尺寸大小的逐渐减小,热源 B、C 周围产生的漩涡对平均努塞尔数的影响增强且转变为主要影响。如图 14 所示,在  $A = 1/16$  时,随着热源间间距的增加,热源 C 周围产生的漩涡逐渐向方腔左下角迁移因而逐渐脱离其左表面和下表面,且随着热源间间距继续增加,其漩涡开始向方腔右上角迁移,一定程度上阻碍了热源 B 周围产生的漩涡的充分发展,并随着热源间间距的增加使得其在方腔上壁面分裂出一个小漩涡,且小漩涡得到充分发展,因而在一定程度上随着热源间间距的增加削弱了方腔右表面上的漩涡强度,并且使得相应表面上形成的热边界层变厚,温度梯度减小,虽然小漩涡得到充分发展,但因其位置距离热源 B、C 的上表面较远,因而对其表面努塞尔数的影响不大,因此其综合影响导致热源在此尺寸下,随着热源间间距的增加,其换热效果逐渐减弱。

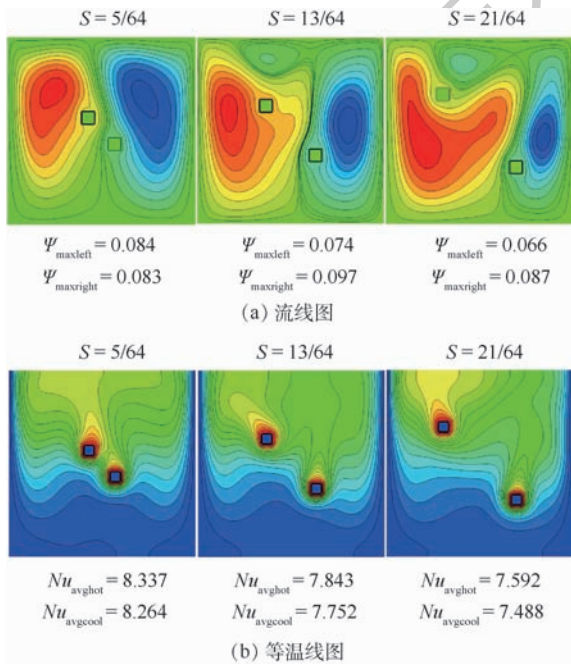


图 14 对角布局时,  $Ra = 10^6, A = 1/16, S = 5/64, 13/64, 21/64$  下的流线图与等温线图

Fig. 14 Streamlines and isotherms for diagonal layout at  $Ra = 10^6, A = 1/16, S = 5/64, 13/64, 21/64$

### 3.4 内热源几何尺寸大小对流动及传热的影响

通过图 3、图 4 可以看出,在  $Ra = 10^4, 10^5, 10^6$ , 热源间间距  $S = 5/64, 13/64, 21/64$ , 内热源几何尺寸大小  $A = 1/16, 1/8, 3/16, 1/4$  情况下,随着内热源几何尺寸大小的增加,方腔内的对流换热效果显著增强。这是因为在固定  $Ra$  数、热源间间距的情况下,热源表面距离方腔边壁的距离远近为对其总平均努塞尔数的主要影响,因此随着内热源几何尺寸大小的增大,其表面距离方腔

边壁的距离减小,导致热边界层变薄,温度梯度变陡,因此换热效果增强。

## 4 结论

本文采用非正交 MRT-LBM 方法,对多孔方腔内含双加热块的自然对流冷却过程进行了数值模拟研究,研究过程通过改变  $Ra$  数、内热源几何尺寸大小、布局方式、热源间间距来研究方腔内加热块壁面和冷壁面平均努塞尔数的变化规律,从而得出最佳换热模型。研究结果如下:

1) 布局方式对冷、热壁面平均努塞尔数的影响。在  $Ra = 10^4, S = 5/64$  时,任意内热源几何尺寸大小及  $S = 13/64, A = 1/16, 1/8, 3/16$  情况下,宜采用对角布局方式;在  $S = 21/64$  时,  $A = 3/16, 1/4$  情况下,宜采用水平布局方式;在  $Ra = 10^5$  情况下,  $S = 5/64$  时,任意内热源几何尺寸大小,采用对角布局方式可获得更好的换热效果;在其他热源间间距及内热源几何尺寸大小下,采用水平布局方式更优;在  $Ra = 10^6$  情况下,  $S = 5/64$  时,宜采用水平或者对角布局;在  $A = 1/16, S = 21/64$  时,宜采用水平或者垂直布局,其换热效果更好;但在其他内热源几何尺寸大小及热源间间距下,则采用水平布局方式更优。

2) 相同的布局方式下,热源间间距对冷、热壁面平均努塞尔数的影响。在热源水平及对角布局方式下,  $Ra = 10^4$  时,任意内热源几何尺寸大小,  $Ra = 10^5$  时,  $A = 1/8, 3/16, 1/4$  情况下,其换热效果均随着热源间间距的增大而增强;而在  $Ra = 10^5, A = 1/16$  情况下,随着热源间间距的增大,其换热效果先增强后减弱;在  $Ra = 10^6, A = 1/4$  时,其换热效果随着热源间间距的增大而增强;在  $A = 3/16$  时,随着热源间间距增大,其换热效果先减弱后增强;而在  $A = 1/8, 1/16$  情况下时,随着热源间间距增大,其换热效果减弱。

3) 热源大小对冷、热壁面平均努塞尔数的影响。在相同的热源间间距  $S = 5/64, 13/64, 21/64, Ra = 10^4, 10^5, 10^6$  情况下,任意内热源几何尺寸大小,随着内热源几何尺寸大小增加,其对流换热效果增强,因此结合上述规律,发现采用水平布局方式,内热源几何尺寸大小选择为  $1/4$ , 热源间间距选择为  $21/64$  情况下,可获得最佳的换热效果。

### 参考文献 (References)

[ 1 ] BASAK T, ROY S, PAUL T, et al. Natural convection in a square cavity filled with a porous medium: Effects of various

- thermal boundary conditions [J]. *International Journal of Heat and Mass Transfer*, 2006, 49(7-8): 1430-1441.
- [2] KHASHAN S A, AL-AMIRI A M, POP I. Numerical simulation of natural convection heat transfer in a porous cavity heated from below using a non-Darcian and thermal non-equilibrium model [J]. *International Journal of Heat and Mass Transfer*, 2006, 49(5-6): 1039-1049.
- [3] SAEID N H. Natural convection in porous cavity with sinusoidal bottom wall temperature variation [J]. *International Communications in Heat and Mass Transfer*, 2005, 32(3-4): 454-463.
- [4] ROY S, BASAK T. Finite element analysis of natural convection flows in a square cavity with non-uniformly heated wall(s) [J]. *International Journal of Engineering Science*, 2005, 43(8): 668-680.
- [5] 何雅玲, 王勇, 李庆. 格子 Boltzmann 方法的理论及应用 [M]. 北京: 科学出版社, 2009.
- HE Y L, WANG Y, LI Q. Lattice Boltzmann method: Theory and applications [M]. Beijing: Science Press, 2009 (in Chinese).
- [6] 郭照立. 格子 Boltzmann 方法的原理及应用 [M]. 北京: 科学出版社, 2009.
- GUO Z L. Theory and applications of lattice Boltzmann method [M]. Beijing: Science Press, 2009 (in Chinese).
- [7] ZHAO C Y, DAI L N, TANG G H, et al. Numerical study of natural convection in porous media (metals) using lattice Boltzmann method (LBM) [J]. *International Journal of Heat and Fluid Flow*, 2010, 31(5): 925-934.
- [8] DIXIT H N, BABU V. Simulation of high Rayleigh number natural convection in a square cavity using the lattice Boltzmann method [J]. *International Journal of Heat and Mass Transfer*, 2006, 49(3): 727-739.
- [9] LIU Q, HE Y L, LI Q, et al. A multiple-relaxation-time lattice Boltzmann model for convection heat transfer in porous media [J]. *International Journal of Heat and Mass Transfer*, 2014, 73: 761-775.
- [10] LIU Q, HE Y L, DAWSON K A, et al. Lattice Boltzmann simulations of convection heat transfer in porous media [J]. *Physica A: Statistical Mechanics and Its Applications*, 2017, 465: 742-753.
- [11] LAM P A K, PRAKASH K A. A numerical study on natural convection and entropy generation in a porous enclosure with heat sources [J]. *International Journal of Heat and Mass Transfer*, 2014, 69(1): 390-407.
- [12] 李培生, 孙金丛, 张莹, 等. 内置高温体倾斜多孔腔体中自然对流的 LBM 模拟 [J]. *哈尔滨工程大学学报*, 2018, 39(6): 1073-1080.
- LI P S, SUN J C, ZHANG Y, et al. Lattice Boltzmann simulation of natural convection in an inclined porous cavity with a hot square obstacle [J]. *Journal of Harbin Engineering University*, 2018, 39(6): 1073-1080 (in Chinese).
- [13] SIAVASHI M, BORDBAR V, RAHNAMA P. Heat transfer and entropy generation study of non-Darcy double-diffusive natural convection in inclined porous enclosures with different source configurations [J]. *Applied Thermal Engineering*, 2017, 110: 1462-1475.
- [14] SELIMEFENDIGIL F. Numerical investigation and POD-based interpolation of natural convection cooling of two heating blocks in a square cavity [J]. *Arabian Journal for Science and Engineering*, 2014, 39(3): 2235-2250.
- [15] NITHIARASU P, SEETHARAMU K N, SUNDARARAJAN T. Natural convective heat transfer in a fluid saturated variable porosity medium [J]. *International Journal of Heat and Mass Transfer*, 1997, 40(16): 3955-3967.

#### 作者简介:

张莹 女, 博士, 教授, 博士生导师。主要研究方向: 复杂热流场模拟。

李培生 男, 博士, 教授, 博士生导师。主要研究方向: 固体废物、生物质、煤的清洁燃烧及资源化利用。

# Non-orthogonal multiple-relaxation-time lattice Boltzmann method for numerical simulation of thermal coupling with porous square cavity flow containing internal heat source

ZHANG Ying<sup>1</sup>, HUANG Yichen<sup>1</sup>, CHEN Yue<sup>1</sup>, MA Ming<sup>2</sup>, LI Peisheng<sup>1,\*</sup>, WANG Zhaotai<sup>1</sup>

(1. School of Mechanical & Electrical Engineering, Nanchang University, Nanchang 330031, China;

2. Department of Aerospace and Mechanical Engineering, University of Notre Dame, South Bend 46556, USA)

**Abstract:** In this paper, in order to solve the problem of natural convection in a porous square cavity containing an internal heat source, the non-orthogonal multiple-relaxation-time (MRT) lattice Boltzmann method was used. The influence of the value of Rayleigh number ( $10^4 \leq Ra \leq 10^6$ ), internal heat source layout (horizontal, vertical and diagonal layout), internal heat source size ( $A = 1/16, 1/8, 3/16, 1/4$ ), and spacing ( $S = 5/64, 13/64, 21/64$ ) between two internal heat sources on convective heat transfer was analyzed. The results indicate that in the case of  $Ra = 10^4, 10^5$  and  $S = 5/64$ , and the internal heat source is of any size, it can obtain better heat transfer by adopting the layout of diagonal; when  $Ra = 10^5, 10^6$  and  $S = 13/64, 21/64$ , horizontal is better. In horizontal layout of the internal heat source, at  $Ra = 10^4$ , the convection heat transfer effect in any internal heat source size is enhanced as the internal heat source spacing increases. However, as  $Ra$  increases, and internal heat source size decreases, the convective heat transfer effect first increases and then decreases with the increase of internal heat source space; then its effect decreases as internal heat source space increases. The layout of diagonal is in a similar situation. When other conditions are the same, the convective heat transfer effect increases with the increase of internal heat source size.

**Keywords:** porous square cavity; internal heat source; natural convection; Nusselt number; multiple-relaxation-time (MRT); Boltzmann model

**Received:** 2019-01-02; **Accepted:** 2019-03-22; **Published online:** 2019-06-10 15:45

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20190606.1552.005.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190606.1552.005.html)

**Foundation items:** National Natural Science Foundation of China (51566012, 11562011); Natural Science Foundation of Jiangxi Province (20181BAB206031); Graduate Innovation Special Foundation of Jiangxi Province (YC2019-S016)

\* **Corresponding author.** E-mail: [nucdns1995z@163.com](mailto:nucdns1995z@163.com)

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

DOI: 10.13700/j.bh.1001-5965.2018.0782

# 机动发射条件下空间飞行器上升段弹道设计

鲜勇\*, 任乐亮, 郭玮林, 张大巧, 李冰

(火箭军工程大学 作战保障学院, 西安 710025)



**摘 要:** 为提高空间飞行器机动发射能力,在飞行中段轨迹确定情况下,以入轨点位置、高度、速度、速度方位角、弹道倾角等作为终端约束,设计上升段弹道,实现以基准发射点为中心,一定范围内任意发射点上升段与飞行中段高精度交班。考虑到上升段终端入轨点约束条件多、精度要求高,且上升段弹道具有非线性、强耦合的特点,研究设计了二级、三级能量管理模型和变射面横向机动模型,并采用加入混合扰动算子的梯度粒子群算法对上升段弹道进行求解。仿真结果表明:优化设计的变射面横向机动弹道能够实现与飞行中段的高精度交班,上升段终端入轨点位置、高度、速度、速度方位角和弹道倾角平均偏差分别为 27.506 2 m、2.125 4 m、1.652 2 m/s、0.072 8°和 0.029 0°。

**关键词:** 机动发射; 变射面横向机动; 能量管理; 梯度粒子群算法; 混合扰动算子

**中图分类号:** V412.1

**文献标识码:** A **文章编号:** 1001-5965(2019)09-1713-10

在现代战争条件下,精确制导武器发展迅速,可以实现“点对点”打击,导致提高武器系统的生存能力变得十分重要。与传统的固定阵地发射方式相比,机动发射在提高生存能力的同时带来了打击的突然性,可以有效提高武器系统作战效能。同样,在应急状态下,能够实现卫星的机动发射,可以迅速对报废卫星进行补充,有效支持作战行动。为方便表述,将卫星、无人机、导弹等称为空间飞行器。

对卫星而言,其运行轨道根据任务需求提前确定;对导弹和无人机等飞行器而言,随着相关技术的发展,飞行中段往往过程复杂,不确定因素多<sup>[1-2]</sup>,需要考虑路径约束、状态约束、必经点、禁飞区等诸多约束条件<sup>[3-4]</sup>,将制约机动发射水平。目前,针对传统弹道导弹机动发射问题的研究相对成熟,而对于飞行中段复杂的空间飞行器机动发射研究较少<sup>[5]</sup>。文献[1]研究了基于 Gauss 伪谱法的滑翔段弹道快速规划方法,但是采用的弹

道规划模型有所简化,并且未考虑对制导控制系统的影响;文献[5]对全程弹道诸元的快速解算问题进行了研究,能够做到全程弹道的快速生成,但是该方法得到的入轨点高度、速度有一定的不确定性,对制导控制系统而言存在一定隐患;在机动发射背景下,文献[6]联合 BP 神经网络和 L-M 算法研究了上升段弹道快速计算方法,可以实现上升段弹道的快速设计,但是没有将入轨点位置和速度方位角作为终端约束,仍需研究飞行中段弹道快速规划方法。

本文从另外一个思路出发,在机动发射条件下,不改变飞行中段弹道,创新性地将研究方向转为对上升段弹道进行设计,并实现与飞行中段的高精度交班。目前,对上升段弹道的研究较为广泛,如文献[7-8]对空间飞行器入轨问题进行了研究,但是未对入轨点位置和速度方位角进行约束。在轨迹优化方面,文献[9]采用改进的粒子群算法对上升段轨迹进行优化,在收敛速度及精度上

收稿日期: 2019-01-02; 录用日期: 2019-02-02; 网络出版时间: 2019-02-28 11:52

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190227.1142.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190227.1142.001.html)

\* 通信作者。E-mail: xy603xy@163.com

**引用格式:** 鲜勇, 任乐亮, 郭玮林, 等. 机动发射条件下空间飞行器上升段弹道设计[J]. 北京航空航天大学学报, 2019, 45(9): 1713-1722. XIAN Y, REN L L, GUO W L, et al. Design of ascent trajectory of space vehicle in mobile launch condition[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1713-1722 (in Chinese).

有很大提高;文献[10]采用梯度粒子群算法对上升段交会弹道进行优化,显著提高了全局寻优能力和收敛速度。本文在此基础上,参考文献[11-13],加入混合扰动算子(BDO),减小搜索后期陷入局部最优解的概率。

机动发射条件下,传统的射面内飞行弹道将无法满足不同空间飞行器苛刻的入轨点交班精度要求,因此引入变射面横向机动模型。文献[14-16]对变射面横向机动弹道进行了深入研究,为本文研究提供了理论基础。

基于上述分析,本文以三级固体火箭为研究对象,通过设计二级和三级能量管理模型对入轨点速度大小进行调整,设计变射面横向机动模型来满足入轨点位置、高度、速度、速度方位角、弹道倾角等约束。通过仿真实验,对机动发射上升段弹道设计方法的有效性进行了验证。

## 1 机动发射

按照预定发射任务,根据发射点大地坐标 $(B_0, L_0, H_0)$ 和目标点大地坐标 $(B_m, L_m, H_m)$ 设计上升段及飞行中段弹道,将其称为基准弹道。其中, $B$ 为纬度, $L$ 为经度, $H$ 为高度。为增强空间飞行器机动发射能力,允许其在以 $(B_0, L_0, H_0)$ 为中心、半径为 $R$ 的发射区域内机动。当机动到 $(B_1, L_1, H_1)$ 后,保持基准弹道中的飞行中段弹道及入轨点状态不变,通过合理设计上升段弹道,使其同样满足入轨要求,完成与飞行中段的交班。文献[1]指出,飞行中段可通行区域小,并且综合考虑飞行中段各种约束条件,容易得到高精度交班对后续飞行十分重要,因此可将入轨点位置(经纬度)、高度、速度、速度方位角、弹道倾角作为入轨状态指标。

当空间飞行器机动至偏离基准弹道射面位置时,传统的射面内弹道设计方法将难以满足速度方位角要求,无法与基准弹道飞行中段实现交班,并最终导致严重偏离目标点。通过上升段横向机动可以完成对速度方位角的调整,并优化设计相关参数,形成变射面横向机动上升段弹道,实现与飞行中段高精度交班。变射面横向机动相关概念可以参考文献[14-16]。

为方便表述发射车机动后的位置,使用 $(r, A_d)$ 表示与基准弹道发射点的相对位置关系, $r$ 为两者之间的距离,取值范围为 $(0, R)$ , $A_d$ 为两者连线偏离基准弹道射击方位角的角度,取值范围为 $(0^\circ, 360^\circ)$ ,顺时针方向为正。

机动发射条件下,空间飞行器横向机动上升

段弹道示意图如图1所示。可知,空间飞行器机动位置可以分为3种情况:①机动后发射车仍在原瞄准方向;②机动后发射车在原瞄准方向偏北;③机动后发射车在原瞄准方向偏南。图1中,虚线为基准弹道与横向机动上升段弹道在地面投影, $OF$ 为基准弹道上升段, $AF$ 为横向机动弹道上升段, $F$ 为空间飞行器入轨点, $A$ 为发射点, $B$ 为一级关机点, $C$ 为二级关机点, $D$ 为三级关机点。

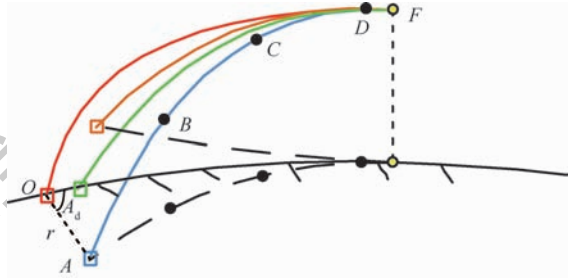


图1 空间飞行器机动发射示意图

Fig. 1 Schematic diagram of mobile launch of space vehicle

## 2 弹道设计

本文以三级固体火箭为例,主要由一、二、三级固体火箭及分导级组成。先需对基准弹道进行设计,实现空间飞行器入轨,由于本文主要研究对象为上升段,因此不对飞行中段进行建模;再以基准弹道入轨点状态作为发射车机动后空间飞行器入轨需要达到的状态,进行建模分析。传统弹道设计方法已经较为成熟,具体可参考文献[17],本文不再赘述。

### 2.1 基准弹道飞行程序角设计

#### 2.1.1 一级飞行程序角设计

为提高一级飞行程序对不同终端约束的适应性,在跨声速前后分别进行2次指数形式攻角转弯,具体形式可参考文献[18]。

#### 2.1.2 二级飞行程序角设计

火箭二级飞行环境影响因素相对较少,按照等斜率转弯方式设计二级飞行程序角,其表达式为

$$\varphi_{\text{ex}}(t) = \begin{cases} \varphi_{\text{ex}}(t_{g1}) & t_{g1} \leq t < t_{g1} + a \\ \varphi_{\text{ex}}(t_{g1}) + \dot{\varphi}_{\text{ex}1}(t - t_{g1} - a) & t_{g1} + a \leq t \leq t_{g2} - b \\ \varphi_{\text{ex}}(t_{g2} - b) & t_{g2} - b < t \leq t_{g2} \end{cases} \quad (1)$$

式中: $t_{g1}$ 为一级分离时刻; $t_{g2}$ 为二级分离时刻; $\dot{\varphi}_{\text{ex}1}$ 为二级转弯角速率; $a, b$ 为常数。

#### 2.1.3 三级飞行程序角设计

三级飞行程序与二级类似,表达式为

$$\varphi_{cx}(t) = \begin{cases} \varphi_{cx}(t_{g2}) & t_{g2} < t < t_{g2} + a \\ \varphi_{cx}(t_{g2}) + \dot{\varphi}_{cx2}(t - t_{g2} - a) & t_{g2} + a \leq t < t_{g3} - b \\ \varphi_{cx}(t_{g3} - b) & t_{g3} - b \leq t < t_{g3} \end{cases} \quad (2)$$

式中:  $t_{g3}$  为三级分离时刻;  $\dot{\varphi}_{cx2}$  为三级转弯角速率。

### 2.1.4 分导级飞行程序角设计

在基准弹道设计中,分导级姿态保持不变,按照最大工作时间进行飞行,即

$$\varphi_{cx}(t) = \varphi_{cx}(t_{g3}) \quad t_{g3} \leq t \leq t_{g4} \quad (3)$$

式中:  $t_{g4}$  为分导级关机时刻。

## 2.2 横向机动上升段弹道设计

火箭一级飞行高度低、空气稠密,为了减小控制系统压力,不适合做大幅度机动。因此,横向机动上升段弹道的一级飞行程序模型与基准弹道相同,对二、三级和分导级飞行程序模型进行优化。在基准弹道设计中,仅对入轨点高度、速度和弹道倾角进行约束,较为简单,而当发射车机动后,需要同时满足入轨点位置要求,约束更强。针对固体火箭通常采用耗尽关机方式,为能够对入轨点速度大小进行有效调整,本文设计了二、三级能量管理模型。

### 2.2.1 能量管理模型

入轨点能量是推力、引力和空气动力等共同做功的结果,而入轨点能量包括势能与动能。当入轨点高度一定时,引力做功为定值,因此,在火箭性能一定的情况下,必须调节推力对火箭做功大小,才能满足终端速度大小要求。本文通过控制火箭产生多余路径,从而对火箭能量进行消耗<sup>[19-20]</sup>。

以二级能量管理模型为例,给出俯仰程序角交变模型(见图2)。为了增大上升段能量管理可调整范围,同样对三级进行能量管理,其模型与二级类似。

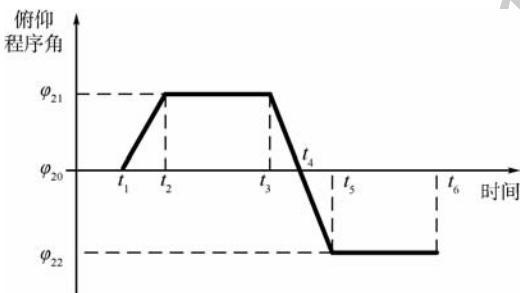


图2 俯仰程序角交变模型示意图

Fig.2 Schematic diagram of alternating model of pitching program angle

横向机动同样会产生多余的能量消耗,因此,最终的能量管理是纵横向飞行程序综合设计的结果。

### 2.2.2 横向机动上升段弹道飞行程序角模型

#### 1) 纵向飞行程序角设计

横向机动上升段弹道二、三级采用能量管理思想,一方面可以消耗能量,另一方面增加程序可调整范围,满足苛刻的多约束终端状态。纵向飞行程序设计如下:

二级飞行程序角为

$$\varphi_{cx}(t) = \begin{cases} \varphi_{cx}(t_{g1}) & t_{g1} \leq t < t_{g1} + a \\ \varphi_{cx}(t_{g1}) + \dot{\varphi}_{cx}^1(t - t_{g1} - a) & t_{g1} + a \leq t < t_{21} \\ \varphi_{cx}(t_{21}) & t_{21} \leq t < t_{21} + c \\ \varphi_{cx}(t_{21}) + \dot{\varphi}_{cx}^2(t - t_{21} - c) & t_{21} + c \leq t < t_{22} \\ \varphi_{cx}(t_{22}) & t_{22} \leq t < t_{g2} \end{cases} \quad (4)$$

式中:  $\dot{\varphi}_{cx}^1$ 、 $\dot{\varphi}_{cx}^2$  为转弯角速率;  $t_{21}$  为第1次转弯结束时刻;  $t_{22}$  为第2次转弯结束时刻;  $c$  为常数。

三级飞行程序角为

$$\varphi_{cx}(t) = \begin{cases} \varphi_{cx}(t_{g2}) & t_{g2} \leq t < t_{g2} + a \\ \varphi_{cx}(t_{g2}) + \dot{\varphi}_{cx}^3(t - t_{g2} - a) & t_{g2} + a \leq t < t_{31} \\ \varphi_{cx}(t_{31}) & t_{31} \leq t < t_{31} + c \\ \varphi_{cx}(t_{31}) + \dot{\varphi}_{cx}^4(t - t_{31} - c) & t_{31} + c \leq t < t_{32} \\ \varphi_{cx}(t_{32}) & t_{32} \leq t < t_{g3} \end{cases} \quad (5)$$

式中:  $\dot{\varphi}_{cx}^3$ 、 $\dot{\varphi}_{cx}^4$  为转弯角速率;  $t_{31}$  为第1次转弯结束时刻;  $t_{32}$  为第2次转弯结束时刻。

分导级飞行程序角为

$$\varphi_{cx}(t) = \begin{cases} \varphi_{cx}(t_{g3}) & t_{g3} \leq t < t_{g3} + t_{k3} \\ \dot{\varphi}_{cx}^5(t - t_{g3} - t_{k3}) + \varphi_{cx}(t_{g3}) & t_{g3} + t_{k3} \leq t < t_{g3} + t_{k3} + T_{fd} \end{cases} \quad (6)$$

式中:  $\dot{\varphi}_{cx}^5$  为转弯角速率;  $t_{k3}$  为三级与分导级过渡时间;  $T_{fd}$  为分导级工作时间。

#### 2) 横向飞行程序角设计

通过改变三级偏航程序角,实现横向机动,为防止偏航角出现突变,在按照反正切函数转弯前5s,先按照线性函数转弯至需要角度<sup>[16]</sup>,偏航程序角  $\psi_{cx}(t)$  为

$$\psi_{cx}(t) = \begin{cases} 0 & t_{g2} \leq t < t_{g2} + 3 \\ \frac{\psi}{900} \left( \frac{\pi}{2} + \arctan \left( \frac{t_{g2} + 8 - k_1}{k_2} \right) \right) (t - t_{g2} - 3) & t_{g2} + 3 \leq t < t_{g2} + 8 \\ \frac{\psi}{180} \left( \frac{\pi}{2} + \arctan \left( \frac{t - k_1}{k_2} \right) \right) & t_{g2} + 8 \leq t < t_{g3} - 3 \\ \psi_{cx}(t_{g3} - 3) & t \geq t_{g3} - 3 \end{cases} \quad (7)$$

式中： $\psi$  为最大偏航角，控制转弯幅度； $k_1$  为转弯时刻控制参数； $k_2$  为转弯快慢控制参数。

### 3 优化模型设计

#### 3.1 转换入轨点状态

为了方便描述入轨点状态与基准弹道入轨点状态的差别，以基准弹道入轨点在地面投影  $m$  为原点，建立东北坐标系  $m x_n y_n z_n$ ，具体定义见文献[17]。通常，入轨点位置速度是在地心大地直角坐标系下给出的，位置坐标可根据式(8)、式(9)将其转换到东北坐标系，速度坐标转换方法类似。

$$\begin{bmatrix} x_n \\ y_n \\ z_n \end{bmatrix} = \begin{bmatrix} n_{11} & n_{12} & n_{13} \\ n_{21} & n_{22} & n_{23} \\ n_{31} & n_{32} & n_{33} \end{bmatrix} \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix} - \begin{bmatrix} 0 \\ r_e \\ 0 \end{bmatrix} \quad (8)$$

式中： $x_n, y_n, z_n$  为东北坐标系下的坐标； $x_s, y_s, z_s$  为地心大地直角坐标系下的坐标； $r_e$  为  $\varphi_{sm}$  下的地球半径， $\varphi_{sm}$  为入轨点的地心纬度。

$$\begin{cases} n_{11} = -\sin \varphi_{sm} \cos \lambda_{sm} \\ n_{12} = -\sin \varphi_{sm} \sin \lambda_{sm} \\ n_{13} = \cos \varphi_{sm} \\ n_{21} = \cos \varphi_{sm} \cos \lambda_{sm} \\ n_{22} = \cos \varphi_{sm} \sin \lambda_{sm} \\ n_{23} = \sin \varphi_{sm} \\ n_{31} = -\sin \lambda_{sm} \\ n_{32} = \cos \lambda_{sm} \\ n_{33} = 0 \end{cases} \quad (9)$$

其中： $\lambda_{sm}$  为入轨点的地心经度。

为方便优化时加权处理，以  $m y_n$  为轴，逆时针

旋转  $\gamma' = \arccos \left( \frac{\tilde{V}_{zn}^f}{\|\tilde{V}_n^f\|} \right)$ ，得到  $m x'_n y'_n z'_n$ ， $\tilde{V}_n^f = (\tilde{V}_{xn}^f,$

$\tilde{V}_{yn}^f, \tilde{V}_{zn}^f)$  表示东北坐标系下基准弹道入轨点的速度矢量。将  $m x'_n y'_n z'_n$  坐标系下入轨点位置表示为

$\mathbf{X}_n^f = (x_n^f, y_n^f, z_n^f)$ ，速度表示为  $\mathbf{V}_n^f = (V_{xn}^f, V_{yn}^f, V_{zn}^f)$ 。

#### 3.2 横向机动上升段弹道约束条件

为保证上升段固体火箭结构稳定，仪器工作正常并减小控制系统控制难度，需要对上升段飞行加入一定的约束，主要包括过程约束、控制量约束及终端约束。

##### 1) 过程约束

为保证弹体结构不受损且弹载仪器工作正常，在火箭飞行过程中必须考虑动压及法向过载等过程约束。

$$\begin{cases} n_y = \left| \frac{X \sin \alpha + Y \cos \alpha}{m_r g} \right| \leq n_{y_{max}} \\ q = \frac{1}{2} \rho V^2 \leq q_{max} \end{cases} \quad (10)$$

式中： $n_y$  为法向过载； $n_{y_{max}}$  为弹体可承受最大法向过载； $X$  和  $Y$  分别为火箭所受阻力和升力； $\alpha$  为攻角； $m_r$  为火箭质量； $g$  为重力加速度； $q$  为动压； $\rho$  为大气密度； $V$  为速度大小； $q_{max}$  为最大动压值。本文中， $n_{y_{max}} = 3g, q_{max} = 90 \text{ kPa}$ 。

##### 2) 控制量约束

控制量约束包括攻角约束、纵横向转弯角速率约束，防止超出控制系统控制能力。

$$\begin{cases} \alpha \leq \alpha_{max} \\ |\dot{\varphi}_{cx}^i| \leq \dot{\varphi}_{max}^i \quad i = 1, 2, 3, 4, 5 \\ \dot{\psi}_{cx} \leq \dot{\psi}_{max} \end{cases} \quad (11)$$

式中： $\alpha_{max}$  为最大攻角； $\dot{\varphi}_{cx}^i$  为第  $i$  次等斜率转弯角速率； $\dot{\varphi}_{max}^i$  为第  $i$  次等斜率转弯最大角速率； $\dot{\psi}_{cx}$  为偏航角速率； $\dot{\psi}_{max}$  为最大偏航角速率。

##### 3) 终端约束

$$\begin{cases} \|\bar{\mathbf{L}}_f - \tilde{\mathbf{L}}_f\| \leq \delta_L \\ |H_f - \tilde{H}_f| \leq \delta_H \\ |V_f - \tilde{V}_f| \leq \delta_V \\ |\Theta_f - \tilde{\Theta}_f| \leq \delta_\Theta \\ |a_f - \tilde{a}_f| \leq \delta_a \end{cases} \quad (12)$$

式中： $\tilde{\mathbf{L}}_f$  为基准弹道入轨点经纬度； $\mathbf{L}_f$  为横向机动上升段弹道入轨点经纬度； $\tilde{H}_f, \tilde{V}_f, \tilde{\Theta}_f$  和  $\tilde{a}_f$  分别为基准弹道入轨点高度、速度大小、当地弹道倾角和速度方位角； $H_f, V_f, \Theta_f$  和  $a_f$  分别为横向机动上升段弹道入轨点高度、速度大小、当地弹道倾角和速度方位角； $\delta_L, \delta_H, \delta_V, \delta_\Theta, \delta_a$  分别表示对应量的允许最大偏差值。

#### 3.3 建立优化模型

##### 3.3.1 优化变量及目标函数

##### 1) 优化变量的确定

根据第2节分析，本文确定优化变量如下：

$$\mathbf{X} = (\theta_0, T_{cz}, T_{fd}, \tilde{\alpha}_1, a_1, \tilde{\alpha}_2, a_2, \varphi_{cx}^1, \varphi_{cx}^2, \varphi_{cx}^3, \varphi_{cx}^4, \varphi_{cx}^5, \psi, k_1, k_2) \quad (13)$$

式中:  $\theta_0$  为初始离面角;  $T_{cz}$  为垂直起飞时间;  $\tilde{\alpha}_1$ 、 $a_1$ 、 $\tilde{\alpha}_2$ 、 $a_2$  为一级两次指数形式攻角转弯参数。

### 2) 目标函数

在弹道优化方面,通常将不等式约束处理成罚函数形式。

$$G_i = \begin{cases} 0 & \text{满足第 } i \text{ 个约束} \\ M & \text{不满足第 } i \text{ 个约束} \end{cases} \quad (14)$$

式中:  $M$  为惩罚量,取很大的一个正数。

将终端状态表示在  $mx'_ny_nz'_n$  坐标系下,按照线性加权方式将终端约束处理成如下目标函数:

$$J = \min \left( a'_1 |x'_n - \tilde{x}'_n| + a'_2 |y'_n - \tilde{y}'_n| + a'_3 |z'_n - \tilde{z}'_n| + a'_4 |V_f - \tilde{V}_f| + a'_5 |\Theta_f - \tilde{\Theta}_f| + a'_6 |a_f - \tilde{a}_f| + \sum_{i=1}^N G_i \right) \quad (15)$$

式中:  $\tilde{x}'_n$ 、 $\tilde{y}'_n$ 、 $\tilde{z}'_n$  分别为  $mx'_ny_nz'_n$  坐标系下基准弹道入轨点 3 个方向位置坐标;  $N$  为约束条件总数;  $a'_i$  ( $i = 1, 2, \dots, 6$ ) 为加权系数;  $\Theta_f = \arcsin\left(\frac{V_{yn}^f}{\|V_n^f\|}\right)$ ;  $a_f = \arcsin\left(\frac{V_{xn}^f}{\|(V_{xn}^f, V_{zn}^f)\|}\right)$ 。

本文借鉴文献[3, 5-6]给出的终端状态,以式(16)为例,作为本文的终端约束。

$$\begin{cases} \tilde{x}'_n = 0 \text{ km} \\ \tilde{y}'_n = 90 \text{ km} \\ \tilde{z}'_n = 0 \text{ km} \\ \tilde{V}_f = 6200 \text{ m/s} \\ \tilde{\Theta}_f = 0^\circ \\ \tilde{a}_f = 0^\circ \end{cases} \quad (16)$$

### 3.3.2 优化算法

文献[10]采用梯度粒子群算法,已经很好地解决了弹道优化问题,本文借鉴该算法,将其用于横向机动上升段弹道相关参数的优化,并进行了相应改进。

文献[12]提到粒子速度趋于 0 时,粒子群算法趋于收敛,若未达到全局最优解,将陷入局部最优解。因此,本文在梯度粒子群算法后期,引入混合扰动算子<sup>[11]</sup>对粒子速度进行扰动,减小算法早熟概率。

#### 1) 混合扰动算子

文献[11]在优化算法中采用扰动操作,增大跳出局部最优的概率,并给出了混合扰动算子表达式。混合扰动算子能够综合 Gauss 分布及柯西

分布特点,产生范围更广的随机数,其表达式为  $\beta = b_1(C(0,1) + b_2N(0,1))$  (17)

式中:  $C(0,1)$  服从标准柯西分布;  $N(0,1)$  服从标准 Gauss 分布;  $b_1$ 、 $b_2$  为很小的正数。

#### 2) 改进速度更新公式

首先,随机选取第  $i$  个粒子的第  $j$  个变量作为增加采取扰动操作的对象。然后,按照式(18)对粒子速度进行更新。

$$\begin{cases} \mathbf{v}_{ij}^{k+1} = \omega \tilde{\mathbf{v}}_{ij}^k + c_1 r_1 (\mathbf{p}_{ij}^k - \mathbf{x}_{ij}^k) + c_2 r_2 (\mathbf{p}_{gj}^k - \mathbf{x}_{ij}^k) \\ \mathbf{v}_{ij}^{k+1} = \mathbf{v}_{ij}^{k+1} (1 + \beta) \end{cases} \quad (18)$$

式中:  $k$  为代数;  $\tilde{\mathbf{v}}_{ij}^k$  和  $\tilde{\mathbf{v}}_{ij}^{k+1}$  分别为扰动后的第  $k$  和  $k+1$  代粒子速度;  $\mathbf{x}_{ij}^k$  为第  $k$  代粒子位置;  $\mathbf{p}_{ij}^k$  和  $\mathbf{p}_{gj}^k$  分别为第  $k$  代个体最优解和全局最优解;  $r_1$ 、 $r_2$  为服从标准 Gauss 分布的随机数;  $\omega$  为惯性权重;  $c_1$ 、 $c_2$  为学习因子。

## 4 仿真计算与结果分析

本文借鉴文献[21]中固体火箭参数,并做适当改进,以适应本文对运载火箭的要求。假定 ( $N38^\circ 10'$ ,  $E117^\circ 37'$ ,  $1 \text{ km}$ ) 为基准发射点,根据基准弹道模型得到其入轨点位置为 ( $N44^\circ 15' 35''$ ,  $E119^\circ 24' 18''$ ,  $90 \text{ km}$ ),在地心大地直角坐标系下的速度矢量为 ( $669.9828$ ,  $-4423.4806$ ,  $4292.4251$ )  $\text{m/s}$ 。以发射点 ( $270^\circ$ ,  $25 \text{ km}$ ) 为例进行横向机动上升段弹道优化,其优化变量取值范围及优化结果如表 1 所示。同时以  $25 \text{ km}$  发射范围为例,对本文模型进行检验。

对发射坐标为 ( $270^\circ$ ,  $25 \text{ km}$ ) 的点位进行仿真,得到空间飞行器横向机动上升段弹道参数仿真结果,如图 3~图 8 所示。图 3~图 8 分别给出了横向机动上升段弹道高度、横向机动距离、速度、横向速度、绝对俯仰角、绝对偏航角、攻角和当地弹道倾角随时间的变化曲线。发射点 ( $270^\circ$ ,  $25 \text{ km}$ ) 是偏离基准弹道最远的位置,其所需横向机动距离将是最大的位置之一,横向机动距离越大,消耗能量越多,受火箭能量限制,将会制约可机动范围。仿真结果表明,在 ( $270^\circ$ ,  $25 \text{ km}$ ) 处入轨精度仍然较高,如表 2 所示,这也间接验证了  $25 \text{ km}$  范围任意位置高精度入轨的可行性。

在不加入变射面横向机动模型的情况下,本文采用相同的弹道设计方法,对发射坐标为 ( $270^\circ$ ,  $25 \text{ km}$ ) 的点位进行仿真,验证采用变射面横向机动模型的必要性。仿真结果表明,该情况

表1 优化变量取值范围及优化结果

Table 1 Range of value for variable optimization and results

变量	下限	上限	优化结果
$\theta_0/(\circ)$	-2.0	2.0	0.877 1
$T_{cx}/s$	3.0	9.0	3.872 8
$T_{fd}/s$	3.0	16.0	12.127 5
$\tilde{\alpha}_1/(\circ)$	3.9	8.0	5.467 1
$a_1$	0.5	0.6	0.570 0
$\tilde{\alpha}_2/(\circ)$	3.0	12.0	8.752 0
$a_2$	0.05	0.29	0.281 0
$\dot{\varphi}_{cx}^1/((\circ) \cdot s^{-1})$	-3.0	0	-3.000 0
$\dot{\varphi}_{cx}^2/((\circ) \cdot s^{-1})$	-3.0	2.0	0.151 3
$\dot{\varphi}_{cx}^3/((\circ) \cdot s^{-1})$	-2.0	4.0	0.983 6
$\dot{\varphi}_{cx}^4/((\circ) \cdot s^{-1})$	-3.0	3.0	1.151 3
$\dot{\varphi}_{cx}^5/((\circ) \cdot s^{-1})$	-4.0	4.0	-2.497 0
$\psi/(\circ)$	-25.0	25.0	24.205 7
$k_1$	143.0	183.0	182.113 3
$k_2$	3.0	12.0	4.602 8

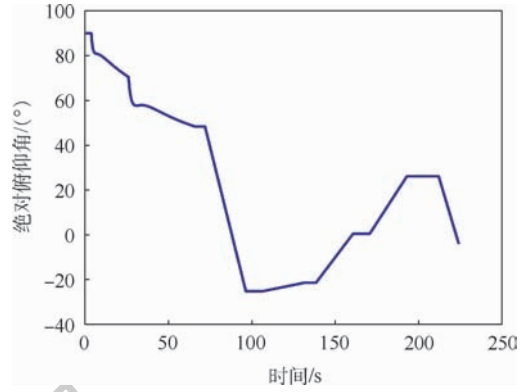


图5 绝对俯仰角随时间变化曲线  
Fig. 5 Variation curve of absolute pitch angle with time

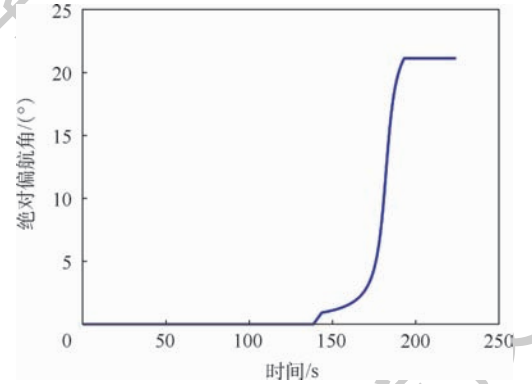


图6 绝对偏航角随时间变化曲线  
Fig. 6 Variation curve of absolute yaw angle with time

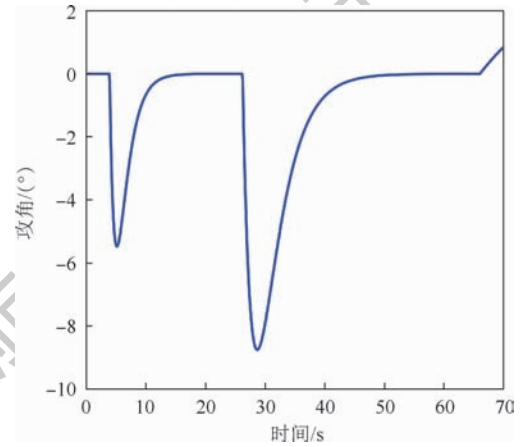


图7 前70s攻角随时间变化曲线  
Fig. 7 Variation curve of first 70 s' attack angle with time  
速度方位角偏差仅为0.028 8°,极大地降低了空间飞行器在飞行中段的修正难度。

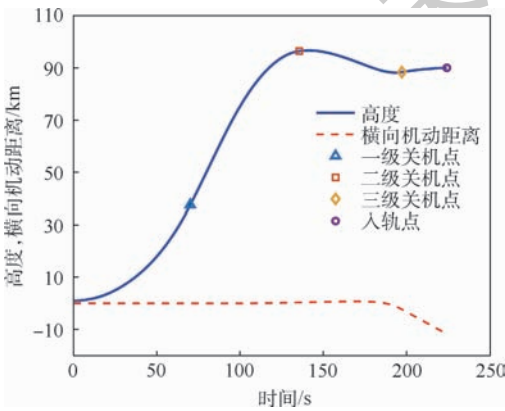


图3 高度、横向机动距离随时间变化曲线  
Fig. 3 Variation curves of height and transverse maneuvering distance with time

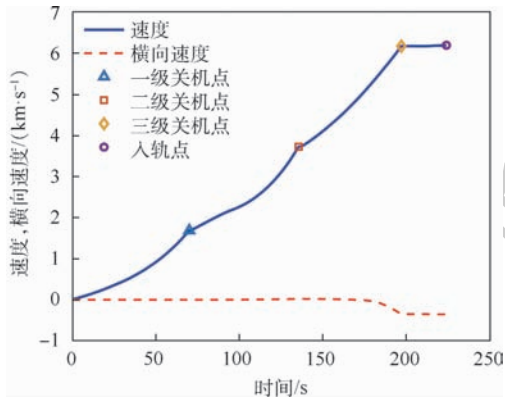


图4 速度、横向速度随时间变化曲线  
Fig. 4 Variation curves of velocity and transverse velocity with time

下能够在入轨点位置、高度、速度、弹道倾角方面实现高精度入轨,如表3所示,但是速度方位角偏差高达2.320 1°,而加入变射面横向机动模型后,

本文在此基础上,在25 km范围内随机产生100个发射点,如图9所示,以及(0°,25 km)、(90°,25 km)、(180°,25 km)、(270°,25 km)4个点,共仿真104个发射车机动后的位置,采用上述模型实现入轨,得到图10~图13所示结果。  
图10、图11分别给出了104个横向机动上升段弹道的高度随时间变化曲线及三维轨迹。

在前 220 s 内,每隔 5 s 记录一次该时刻对应的动压值,并统计 104 次仿真过程中  $t$  时刻的平均

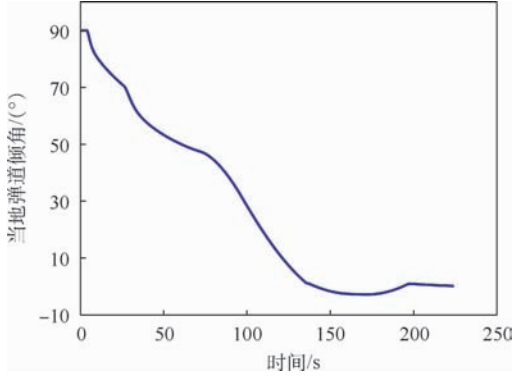


图 8 当地弹道倾角随时间变化曲线

Fig. 8 Variation curve of local ballistic inclination with time

表 2 横向机动上升段弹道入轨点参数偏差

Table 2 Parameter deviation of injection point of transverse maneuvering ascent trajectory

入轨点参数	与基准弹道入轨点偏差
位置/m	18.737 5
高度/m	0.516 4
速度/( $\text{m} \cdot \text{s}^{-1}$ )	0.775 6
速度方位角/( $^\circ$ )	0.028 8
弹道倾角/( $^\circ$ )	0.009 1

表 3 无变射面横向机动的上升段弹道入轨点参数偏差

Table 3 Parameter deviation of injection point of ascent trajectory without changeable-launching-plane transverse maneuvering

入轨点参数	与基准弹道入轨点偏差
位置/m	1.547 4
高度/m	0.181 3
速度/( $\text{m} \cdot \text{s}^{-1}$ )	0.271 0
速度方位角/( $^\circ$ )	2.320 1
弹道倾角/( $^\circ$ )	0.005 5

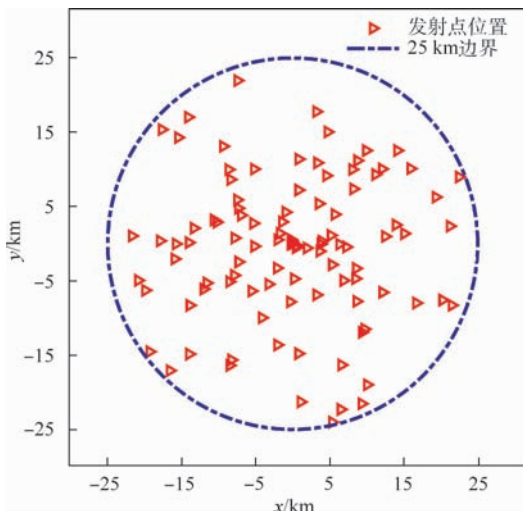


图 9 100 个随机发射点位置

Fig. 9 Position of 100 stochastic launch sites

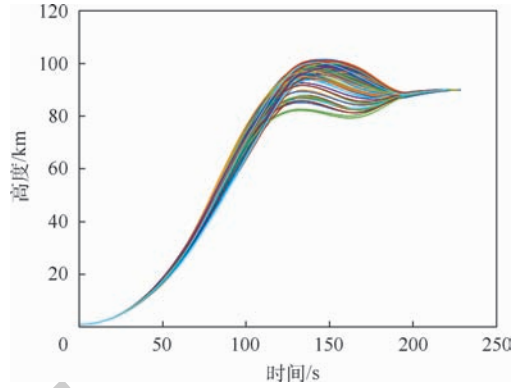


图 10 高度随时间变化仿真

Fig. 10 Variation simulation of height with time

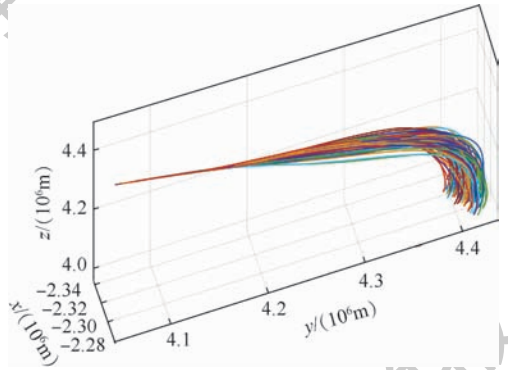


图 11 弹道三维仿真

Fig. 11 Three-dimensional simulation of ballistic

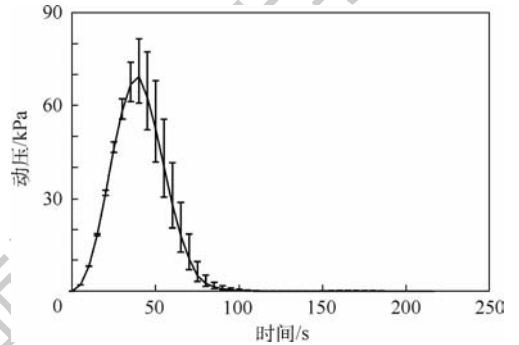


图 12 动压随时间变化仿真

Fig. 12 Variation simulation of dynamic pressure with time

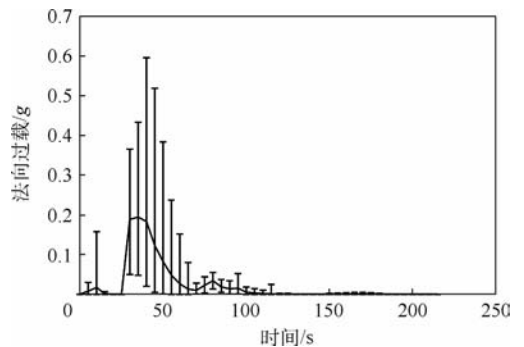


图 13 法向过载随时间变化仿真

Fig. 13 Variation simulation of normal overload with time

动压、最大动压和最小动压,得到图 12,可以看出,最大动压值小于 90 kPa,能够满足火箭飞行中的过程约束。

采用与图 12 类似的方法,图 13 给了法向过载随时间的变化曲线,最大法向过载值小于 0.6g。本文中一级火箭的工作时间为 70 s,从图 12、图 13 中可以看出,70 s 以后,法向过载和动压都很小,最大值分别不超过 0.1g 和 20 kPa,平均值分别不超过 0.1g 和 11 kPa,说明本文设计的二、三级能量管理模型并不会对弹体造成太大影响,具有一定的工程可行性。

表 4 给出了 104 条横向机动上升段弹道在入轨点的位置、高度、速度、速度方位角、弹道倾角与基准弹道对应状态量的平均偏差值及最大偏差值。可以看出,所有状态量的最大偏差值都在指标要求范围内。

表 4 104 条弹道入轨点参数偏差统计

Table 4 Statistics of parameter deviation of 104 ballistic injection points

入轨点参数	平均偏差	最大偏差
位置/m	27.5062	70.4319
高度/m	2.1254	7.0036
速度/(m·s <sup>-1</sup> )	1.6522	4.7691
速度方位角/(°)	0.0728	0.2896
弹道倾角/(°)	0.0290	0.1246

## 5 结论

本文针对空间飞行器机动发射上升段弹道设计问题进行了研究,总结如下:

1) 针对机动发射条件下上升段弹道规划难题,提出横向机动上升段弹道解决方案。以基准弹道入轨点状态为终端约束,合理设计上升段弹道,完成与基准弹道飞行中段的高精度交班,仿真结果表明,能够实现发射车在 25 km 范围内机动。

2) 对横向机动上升段弹道模型进行设计。以入轨点位置、高度、速度、速度方位角、弹道倾角等作为终端约束指标,通过设计二、三级能量管理模型及改变三级偏航角的变射面横向机动模型,满足了苛刻的多约束终端状态。

3) 针对横向机动上升段弹道非线性、强耦合、终端状态多约束问题,采用梯度粒子群算法对参数进行优化,并引入混合扰动算子,降低陷入局部最优解概率。仿真结果表明,能够解决本文面临的 15 个优化变量问题,精度较高。

本文建立的横向机动上升段弹道模型能够较好地解决空间飞行器机动发射问题,并且由于火

箭的性能将会影响入轨精度,本文模型也可以为火箭设计提供一定参考。同时,需要指出的是,本文尚未考虑弹道参数的快速计算问题,而要做到真正的机动发射,必须缩短诸元准备时间,这是后续需要进一步研究的方向。

## 参考文献 (References)

- [1] 陈法龙. 高超声速滑翔飞行器弹道快速规划研究[D]. 长沙:国防科学技术大学,2012.  
CHEN F L. Research of rapid trajectory planning for hypersonic glide vehicle [D]. Changsha: National University of Defense Technology, 2012 (in Chinese).
- [2] 谢愈,潘亮,谷学强,等. 高超声速飞行器多目标复杂约束滑翔弹道优化[J]. 国防科技大学学报,2017,39(2):9-17.  
XIE Y, PAN L, GU X Q, et al. Gliding trajectory optimization with multiple objectives and complicated constraints for hypersonic vehicles [J]. Journal of National University of Defense Technology, 2017, 39(2): 9-17 (in Chinese).
- [3] 陈铁彪,龚曼,王洪波,等. 临近空间可变形滑翔飞行器轨迹优化与性能分析[J]. 宇航学报,2018,39(9):944-952.  
CHEN T B, GONG M, WANG H B, et al. Trajectory optimization and performance analysis of the near-space morphing glide vehicles [J]. Journal of Astronautics, 2018, 39(9): 944-952 (in Chinese).
- [4] 王青,莫华东,吴振东,等. 基于能量的高超声速飞行器再入混合制导方法[J]. 北京航空航天大学学报,2014,40(5):579-584.  
WANG Q, MO H D, WU Z D, et al. Energy-based hybrid reentry guidance for hypersonic vehicles [J]. Journal of Beijing University of Aeronautics and Astronautics, 2014, 40(5): 579-584 (in Chinese).
- [5] 何睿智,刘鲁华,汤建国,等. 机动发射条件下助推滑翔导弹射击诸元快速解算[J]. 国防科技大学学报,2017,39(4):56-61.  
HE R Z, LIU L H, TANG G J, et al. Efficient method for firing data calculation of boost-glide missile in mobile condition [J]. Journal of National University of Defense Technology, 2017, 39(4): 56-61 (in Chinese).
- [6] 郭玮林,鲜勇,张大巧,等. 高超声速飞行器助推段弹道快速计算方法[J]. 中国惯性技术学报,2018,26(1):109-114.  
GUO W L, XIAN Y, ZHANG D Q, et al. Fast calculation method of booster trajectory for hypersonic vehicle [J]. Journal of Chinese Inertial Technology, 2018, 26(1): 109-114 (in Chinese).
- [7] 鲜勇,郭玮林,张大巧,等. 高超声速飞行器非连续点火助推增程弹道设计[J]. 固体火箭技术,2017,40(3):397-402.  
XIAN Y, GUO W L, ZHANG D Q, et al. Trajectory design of improving range for hypersonic vehicle based on discontinuous booster [J]. Journal of Solid Rocket Technology, 2017, 40(3): 397-402 (in Chinese).
- [8] 任京涛. 助推滑翔弹上升段多终端约束弹道设计及制导方法研究[D]. 哈尔滨:哈尔滨工业大学,2013.  
REN J T. Study on the multi-terminal constraint ballistic design

- and guidance method of the ascent segment of the boost-glide-missile [D]. Harbin: Harbin Institute of Technology, 2013 (in Chinese).
- [9] WU G, LIU L, WANG Y, et al. Ascent trajectory optimization of hypersonic vehicle based on improved particle swarm algorithm [C] // Chinese Automation Congress. Piscataway, NJ: IEEE Press, 2015: 115-120.
- [10] 李振华, 鲜勇, 雷刚, 等. 基于混合粒子群算法的上升段交会弹道快速优化设计 [J]. 航空动力学报, 2015, 30 (12): 3029-3034.
- LI Z H, XIAN Y, LEI G, et al. Rapid optimum design of ascent rendezvous trajectory for launch vehicles based on hybrid particle swarm algorithm [J]. Journal of Aerospace Power, 2015, 30 (12): 3029-3034 (in Chinese).
- [11] 陈道君, 龚庆武, 金朝意, 等. 基于自适应扰动量子粒子群算法参数优化的支持向量回归机短期风电功率预测 [J]. 电网技术, 2013, 37 (4): 974-980.
- CHEN D J, GONG Q W, JIN Z Y, et al. Short-term wind power prediction based on support vector regression machine optimized by adaptive disturbance quantum-behaved particle swarm optimization [J]. Power System Technology, 2013, 37 (4): 974-980 (in Chinese).
- [12] 王敏, 唐俊. 基于自适应扰动的粒子群优化算法 [J]. 计算机工程与应用, 2014, 50 (9): 111-115.
- WANG M, TANG J. Particle swarm optimization based on adaptive disturbance [J]. Computer Engineering and Applications, 2014, 50 (9): 111-115 (in Chinese).
- [13] BERGH F V D, ENGELBRECHT A P. A study of particle swarm optimization particle trajectories [J]. Information Sciences, 2006, 176 (8): 937-971.
- [14] 鲜勇, 刘炳琪, 李振华, 等. 一种横向机动弹道能量管理方法 [J]. 弹道学报, 2014, 26 (2): 43-47.
- XIAN Y, LIU B Q, LI Z H, et al. An energy management method of trajectory of transversely maneuvering missile [J]. Journal of Ballistics, 2014, 26 (2): 43-47 (in Chinese).
- [15] 鲜勇, 刘炳琪, 王安民, 等. 变射面弹道横向转弯飞行程序的设计与优化 [J]. 弹道学报, 2014, 26 (4): 30-35.
- XIAN Y, LIU B Q, WANG A M, et al. Design and optimization of lateral turning flight program on changeable launching plane trajectory [J]. Journal of Ballistics, 2014, 26 (4): 30-35 (in Chinese).
- [16] 鲜勇, 李少朋, 李振华, 等. 基于梯度粒子群算法的纵横向机动跳跃弹道设计及优化 [J]. 弹道学报, 2015, 27 (3): 1-6.
- XIAN Y, LI S P, LI Z H, et al. Design and optimization of vertical wavy and crosswise maneuvering trajectory based on gradient particle swarm algorithm [J]. Journal of Ballistics, 2015, 27 (3): 1-6 (in Chinese).
- [17] 张毅, 肖龙旭, 王顺宏. 弹道导弹弹道学 [M]. 长沙: 国防科技大学出版社, 1999.
- ZHANG Y, XIAO L X, WANG S H. Missile ballistics [M]. Changsha: National University of Defense Technology Press, 1999 (in Chinese).
- [18] 鲜勇, 刘凯, 凌王辉, 等. 基于二次点火的飞行器横向转移入轨弹道设计 [J]. 航空动力学报, 2018, 33 (2): 320-326.
- XIAN Y, LIU K, LING W H, et al. Trajectory design of launching-plane transforming orbit-injection of spacecraft based on engine twice ignition [J]. Journal of Aerospace Power, 2018, 33 (2): 320-326 (in Chinese).
- [19] 陈喆, 唐胜景, 郭杰. 固体火箭主动段终端多约束能量管理制导方法 [J]. 系统工程与电子技术, 2014, 36 (12): 2484-2489.
- CHEN Z, TANG S J, GUO J. Energy management based guidance of solid rocket with multi-constraints [J]. Systems Engineering and Electronics, 2014, 36 (12): 2484-2489 (in Chinese).
- [20] 陈克俊. 耗尽关机制导方法研究 [J]. 国防科技大学学报, 1996, 18 (3): 35-39.
- CHEN K J. An approach to depleted shutdown guidance [J]. Journal of National University of Defense Technology, 1996, 18 (3): 35-39 (in Chinese).
- [21] 杨希祥, 江振宇, 张为华. 基于粒子群算法的固体运载火箭上升段弹道优化设计研究 [J]. 宇航学报, 2010, 31 (5): 1304-1309.
- YANG X X, JIANG Z Y, ZHANG W H. A particle swarm optimization algorithm-based solid launch vehicle ascent trajectory optimum design [J]. Journal of Astronautics, 2010, 31 (5): 1304-1309 (in Chinese).

**作者简介:**

鲜勇 男, 博士, 教授, 博士生导师。主要研究方向: 飞行器设计、制导理论等。

任乐亮 男, 硕士研究生。主要研究方向: 飞行器导航与设计。

## Design of ascent trajectory of space vehicle in mobile launch condition

XIAN Yong\*, REN Leliang, GUO Weilin, ZHANG Daqiao, LI Bing

(College of War Support, Rocket Force University of Engineering, Xi'an 710025, China)

**Abstract:** In order to improve the mobile launch capability of space vehicle, the position, height, velocity magnitude, velocity azimuth and trajectory inclination of the injection point were used as terminal constraints to design the ascent trajectory under the condition that the mid-flight trajectory was determined. The space vehicle could be launched at any position in a certain range centered on the original launch point, and the high-precision shift with the mid-flight trajectory was completed. Considering many constraints and high-precision requirements of the injection point, and the nonlinear and strong coupling characteristics of the ascent trajectory, the two-stage and three-stage energy management model and the changeable-launching-plane transverse maneuvering model were designed, and then the gradient particle swarm algorithm with the blended disturbance operator was used to solve the ascent trajectory. The simulation results show that the high-precision shift with the mid-flight trajectory can be realized smoothly by optimized transverse maneuvering trajectory, and the average deviation of the position, height, velocity, velocity azimuth and trajectory inclination of the terminal injection point for ascent trajectory are 27.506 2 m, 2.125 4 m, 1.652 2 m/s, 0.072 8° and 0.029 0° respectively.

**Keywords:** mobile launch; changeable-launching-plane transverse maneuvering; energy management; gradient particle swarm algorithm; blended disturbance operator

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

DOI: 10.13700/j.bh.1001-5965.2018.0749

## TC4 喷丸强化仿真与试验

王延忠<sup>1</sup>, 李菲<sup>1</sup>, 陈燕燕<sup>1,\*</sup>, 张亚萍<sup>1</sup>, 吴泽刚<sup>2</sup>, 王成<sup>3</sup>

1. 北京航空航天大学 机械工程及自动化学院, 北京 100083; 2. 中国航发哈尔滨东安发动机有限公司, 哈尔滨 150066;

3. 中国北方车辆研究所, 北京 100072)

**摘 要:** 针对目前喷丸有限元仿真研究中弹丸规则排列、覆盖率无法精确计算的研究现状, 基于 ABAQUS 二次开发建立了弹丸空间位置随机分布、弹丸数量由覆盖率决定的多弹丸喷丸有限元模型, 实现了不同工况、不同喷丸覆盖率条件下所需弹丸个数的精确计算; 在此基础上研究了不同喷丸参数对 TC4 钛合金材料表面残余应力与表面粗糙度的影响规律。进行了 TC4 喷丸试验, 将得到的残余应力分布、表面粗糙度与仿真结果进行对比, 验证了该有限元模型的合理性。研究成果对喷丸工艺参数的确定具有一定的指导意义。

**关键词:** TC4 钛合金; 喷丸模型; 随机分布; 覆盖率; 残余压应力

**中图分类号:** TH162

**文献标识码:** A **文章编号:** 1001-5965(2019)09-1723-09

喷丸工艺能够在材料表面与表面层引入残余压应力场, 从而有效阻止疲劳裂纹的萌生与扩展, 提高结构件的疲劳寿命<sup>[1]</sup>, 具有效果显著、应用范围广等特点<sup>[2]</sup>。喷丸工艺参数的选取决定了材料表面残余应力的分布形式与分布状态, 对受喷件的疲劳寿命有至关重要的影响。然而喷丸参数的确定依赖于大量的喷丸试验, 试验周期长、成本高, 且喷丸效果与试验者的经验密切相关, 与此相比较, 喷丸仿真可以模拟各工况参数下的喷丸强化过程, 用时短, 成本低, 是研究喷丸强化机理与喷丸效果的主要手段。

为研究喷丸参数对工件表面完整性参数的影响规律, 国内外学者建立了不同的喷丸模型, 进行过大量的仿真研究。Mori 等<sup>[3]</sup> 基于有限元方法进行了喷丸残余应力场的研究, 较早建立了喷丸仿真模型。Guagliano<sup>[4]</sup> 研究了喷丸强度的拟合关系式。Meguid 等<sup>[5-8]</sup> 建立了不同的单弹丸撞击模型, 包括四分之一模型、轴对称模型等, 研究了不

同喷丸参数对不同材料表面完整性参数的影响规律。在此基础上发展出了在空间按一定位置分布的阵列弹丸模型<sup>[9-12]</sup>, 通过控制弹丸的位置分布与数量来提高覆盖率。国内夏明莉等<sup>[13]</sup> 也对 TC4 喷丸试验进行了研究。

总体来说, 国内外对喷丸仿真进行了大量的研究, 但研究对象以钢材料为主, 针对航空用钛合金材料的喷丸仿真研究相对较少, 且建立的仿真模型以单弹丸模型与阵列弹丸模型为主, 忽略了真实喷丸过程中弹丸位置的随机特性, 无法准确计算覆盖率, 影响了仿真的效果与真实性。

为了解决现有喷丸模型中弹丸位置数量固定、喷丸覆盖率无法精确计算的问题, 本文基于 ABAQUS 二次开发手段对航空常用材料 TC4 进行喷丸强化模拟, 建立弹丸空间位置随机分布、弹丸数量与喷丸覆盖率相关的多弹丸模型, 并通过喷丸试验验证有限元模型的合理性。在此基础上, 研究了主要喷丸参数对 TC4 表面残余应力分

收稿日期: 2018-12-20; 录用日期: 2019-01-18; 网络出版时间: 2019-04-1208: 36

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20190409.1609.004.html

基金项目: 基础产品创新科研项目(23709900000170006)

\* 通信作者. E-mail: besthouniao@foxmail.com

引用格式: 王延忠, 李菲, 陈燕燕, 等. TC4 喷丸强化仿真与试验[J]. 北京航空航天大学学报, 2019, 45(9): 1723-1731.

WANG Y Z, LI F, CHEN Y Y, et al. TC4 shot peening simulation and experiment[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1723-1731 (in Chinese).

布及表面粗糙度的影响规律,对 TC4 喷丸工艺的制定具有一定的指导意义。

## 1 喷丸有限元模型

本文研究对象为 TC4(Ti6Al4V),弹丸为钢丝切丸,在仿真过程中,由于主要的研究对象为受喷材料(即钛合金板材),且实际喷丸过程中弹丸的硬度、强度都要远高于目标板材,因此为提高运算效率,在仿真中将弹丸定义成刚体,材料为碳钢。受喷板材与弹丸材料的基本材料参数如表 1 所示。

在喷丸过程中,材料表层会发生高速塑性变形,材料的屈服应力和屈服极限在不同的应变速率下将发生改变,因而 TC4 材料的塑性参数采用 Johnson-Cook<sup>[14]</sup>模型(在材料定义时选择 Rate dependent)。该模型适用于大多数发生高应变速率变形的材料,材料的屈服极限表示为

$$\delta = (A + B\varepsilon^n)(1 + C \ln \dot{\varepsilon}^*) (1 - T^{*m}) \quad (1)$$

式中: $\delta$  为材料屈服极限; $A$  为材料屈服应力; $B$  为材料应变幂指数; $\varepsilon$  为材料等效塑性应变; $n$  为应变硬化指数; $C$  为应变率敏感系数; $\dot{\varepsilon}^*$  为应变影响因子; $T^*$  为温度影响因子; $m$  为温度敏感性系数。同时在 ABAQUS 中还需要设置材料熔点和参考温度(一般取室温),其他参数见表 2<sup>[15]</sup>,其中, $\dot{\varepsilon}_0$  为参考应变率。

考虑到计算时长,同时喷丸残余应力影响层深一般在 1 mm 之内,TC4 板料模型尺寸取为 1.5 mm × 1.5 mm × 1 mm。设定钛合金板材单元

表 1 基本材料参数

Table 1 Basic material parameters

参数	TC4	弹丸
密度/(t · mm <sup>-3</sup> )	4.428 × 10 <sup>-9</sup>	7.8 × 10 <sup>-9</sup>
弹性模量/GPa	110	200
泊松比	0.342	0.3
屈服强度/GPa	1.098	
抗拉强度/GPa	1.092	

表 2 TC4 材料本构模型参数<sup>[15]</sup>

Table 2 TC4 material constitutive model parameters<sup>[15]</sup>

参数	数值
A/MPa	1098
B/MPa	1092
n	0.93
m	1.1
熔点/K	1878
参考温度/K	298
C	0.014
$\dot{\varepsilon}_0$	1

为 C3D8R,在喷丸影响区采用局部细化网格方式划分单元,且网格尺寸小于丸粒直径的 1/10;弹丸单元类型为 C3D4。设置弹丸与板表面之间的摩擦模型为罚函数摩擦模型,摩擦系数取为 0.1,设置好边界条件与约束,改变弹丸的直径、速度与喷丸覆盖率等参数,在 ABAQUS 中运用 Dynamic-Explicit 分析步求解模拟弹丸对板材表面的撞击过程,求解时间设置为 5 × 10<sup>-4</sup> s。之后,将弹丸撞击脱离后的动态显式输出结果导入到 ABAQUS/Standard 隐式求解器中得到稳态结果。

后处理的研究对象主要是残余应力与表面粗糙度。根据粗糙度  $R_a$  的定义, $R_a$  为轮廓偏距绝对值的算术平均值,若撞击后材料表面层节点  $i$  沿层深方向的坐标值为  $U_i$ ,则粗糙度  $R_a$  可表示为

$$R_a = \sum_{i=1}^{num} \left| U_i - \sum_{i=1}^{num} U_i / num \right| / num \quad (2)$$

式中: num 为喷丸区表面节点总数; $i$  为节点编号; $U_i$  为节点  $i$  沿层深方向的坐标值。

每个节点的应力值可通过 Python 编写的后处理程序提取。根据以上方法最终可以得到材料表面粗糙度数值与喷丸影响区每一层单元体的残余应力均值。

## 2 弹丸位置随机化的实现

在实际喷丸过程中,弹丸的空间位置在喷射路径上是具有随机性的,随机分布的弹丸在高速运动下组成了丸粒流对工件表面进行冲击。在以往关于喷丸仿真的文献中,绝大部分研究忽略了这种位置随机性,而人为设定弹丸流在工件表面上方规则分布,使得材料表面在经喷丸后撞击产生的凹坑位置具有一定的规律性,这与事实情况是相悖的,进一步影响了仿真结果的准确性。为解决这一问题,本文基于 Python 生成了空间弹丸球心随机坐标,使弹丸在 TC4 板件碰撞研究区域正上方随机分布且互不干涉。任一弹丸的球心坐标  $(x_i, y_i, z_i)$  满足以下关系:

$$\begin{cases} x_{\min} \leq x_i \leq x_{\max} \\ y_{\min} \leq y_i \leq y_{\max} \\ z_{\min} \leq z_i \leq z_{\max} \end{cases} \quad (3)$$

式中: $(x_{\min}, y_{\min}, z_{\min})$ 、 $(x_{\max}, y_{\max}, z_{\max})$  为弹丸的空间分布范围,即 TC4 板材丸粒中心分布区的正上方范围。

任意两弹丸球心  $(x_i, y_i, z_i)$ 、 $(x_n, y_n, z_n)$  之间的距离满足以下关系:

$$(x_i - x_n)^2 + (y_i - y_n)^2 + (z_i - z_n)^2 \geq D^2 \quad (4)$$

式中: $D$ 为弹丸直径。

采用 `random.uniform()` 函数获得板材粒粒中心分布区正上方空间的随机弹丸球心坐标点,通过嵌套循环程序对任意两弹丸之间的距离进行判断,以确保任意两弹丸之间没有干涉现象。程序流程如图1所示,图2为生成的空间随机分布多弹丸模型。

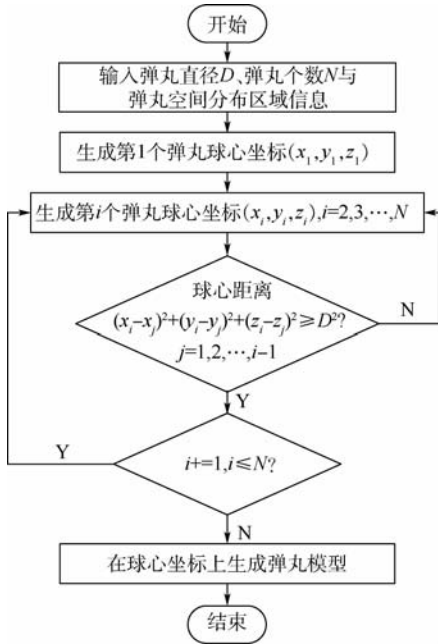


图1 弹丸位置随机坐标生成流程

Fig.1 Projectile random position coordinate generation flowchart

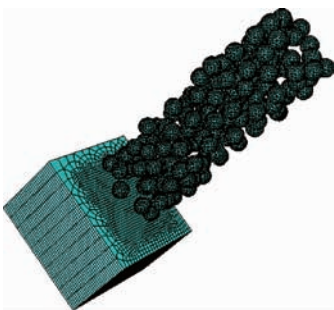


图2 空间随机分布多弹丸模型

Fig.2 Spatially randomized multi-projectile model

### 3 喷丸覆盖率计算

喷丸覆盖率指材料表面喷丸弹坑所占面积之和与零件表面总面积的比值。在以往的文献中,绝大部分研究通过增加固定数量的弹丸来改变覆盖率,这种方式无法定量地给出覆盖率的具体数值大小,且在弹丸直径改变时,弹丸数量一定的情况下喷丸覆盖率会发生变化,以往的研究中往往

忽略了这种变化,对仿真结果产生了影响。

本文在有限元仿真模型的基础上基于 Python 语言编写了覆盖率计算子程序,针对不同弹丸直径与弹坑直径计算喷丸覆盖率,以确定不同覆盖率下所需要的弹丸个数。在对覆盖率计算之前,首先需要计算单个弹坑的直径  $2a_p$ 。通过建立单弹丸的撞击模型,改变弹丸速度与直径,得到不同初始条件下板材撞击区域沿层深方向的位移,提取计算结果中的弹坑节点位移曲线,并将沿层深方向位移为零的两点距离取为弹坑直径的大小,如图3所示。

在得到单弹丸撞击弹坑直径  $2a_p$  后,为计算喷丸覆盖率,将 TC4 板材喷丸影响区表面分别沿  $x$  方向和  $y$  方向等分,生成  $N_x$  与  $N_y$  个节点,则共生成  $N_t = N_x \times N_y$  个节点之后计算每个节点到所有弹丸球心的距离  $S$ ,其中弹丸球心坐标由第2节中的程序生成,若对所有  $S$  都存在  $S > a_p$ ,则该节点处在未喷区域,如图4所示。将所有处于未喷区域的节点个数记为  $M$ ,则覆盖率  $Cov = (1 - M/N_t) \times 100\%$ 。开发的 Python 程序流程如图5所示。

在实际的喷丸强化过程中,若所形成的弹

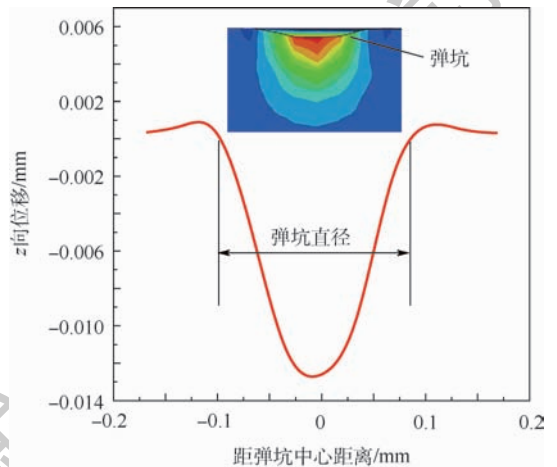


图3 弹坑直径提取方法

Fig.3 Crater diameter extraction method

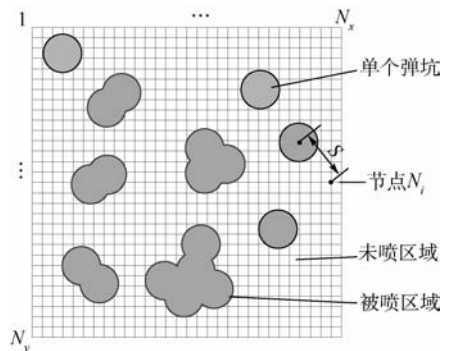


图4 喷丸覆盖率计算

Fig.4 Shot peening coverage calculation

坑面积占试样待喷区域总面积的98%以上,就可以认为覆盖率达到100%<sup>[12]</sup>。因此,在本文中,以覆盖率达到98%时的弹丸个数 $Z$ 为基准,200%覆盖率时所需弹丸个数则为 $2Z$ ,以此类推。不同工况下的弹坑半径及覆盖率达到98%时所需的弹丸个数如表3所示。

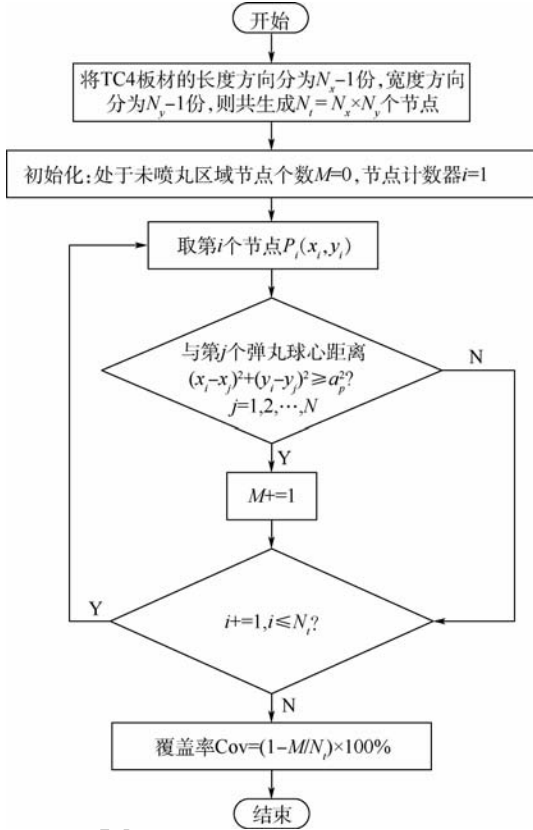


图5 喷丸覆盖率计算流程

Fig.5 Calculation flowchart of shot peening coverage

表3 不同喷丸工况下达到98%覆盖率时所需弹丸个数  
Table 3 Number of projectiles required to achieve 98% coverage under different shot peening conditions

工况		弹坑半径 $a_p/mm$	覆盖率达到98% 以上时所需 弹丸个数
弹丸粒度/mm	弹丸速度/ ( $m \cdot s^{-1}$ )		
0.3	65	0.084	110
0.4	65	0.101	90
0.5	65	0.135	65
0.6	65	0.152	45
0.7	65	0.169	30
0.5	70	0.135	65
0.5	75	0.135	65
0.5	80	0.135	65
0.5	85	0.135	65

## 4 仿真算例及结果分析

影响喷丸效果的因素有很多,在实际喷丸过程中,喷丸强度由喷丸压力、丸粒牌号、丸粒流量

等因素共同作用,在丸粒流量一定的情况下,主要通过调节喷丸压力与喷丸时间来达到控制表面质量的目的。对于航空结构件来说,由于大部分属于薄壁件,为使表面形成理想的残余压应力层,同时表面粗糙度不过分增大,喷丸参数需要进行严格地选取。同时,在喷丸强度一定的情况下,需要严格控制喷丸时间防止过喷。基于建立的仿真分析模型,本文研究了不同喷丸参数(弹丸速度、弹丸粒度、喷丸覆盖率)对TC4钛合金表面强化特征参数(残余应力与表面粗糙度)的影响规律。图6为弹丸速度65 m/s、弹丸粒度0.4 mm、喷丸覆盖率100%条件下的仿真结果云图。

### 1) 弹丸速度对表面强化特征参数的影响

弹丸速度是对喷丸强度影响最直接的因素。为研究不同弹丸速度对TC4材料表面残余应力场的影响规律,取弹丸直径为0.5 mm,喷丸覆盖率为100%,分别模拟65、70、75、80、85 m/s弹丸速度下的撞击情况。为了考虑喷丸后残余应力的均化过程,将撞击过程的动态显式输出结果导入到隐式求解器中得到稳态结果,得到沿层深的残余应力分布与不同喷丸速度下表面粗糙度,如图7与图8所示。

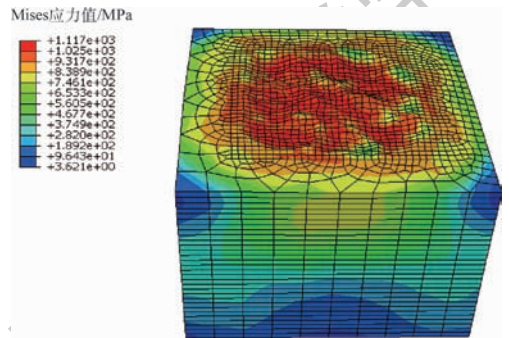


图6 喷丸仿真结果云图

Fig.6 Contour of shot peening simulation results

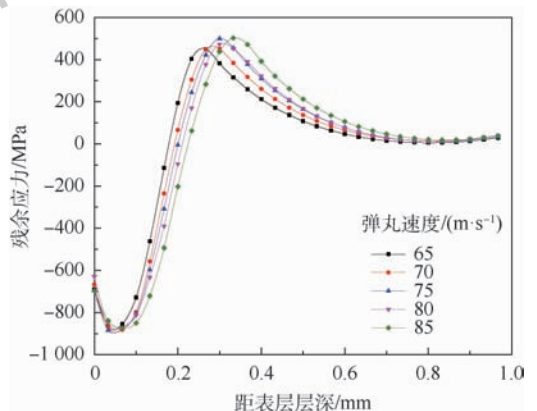


图7 不同弹丸速度对残余应力场的影响

Fig.7 Effect of different projectile velocities on residual stress field

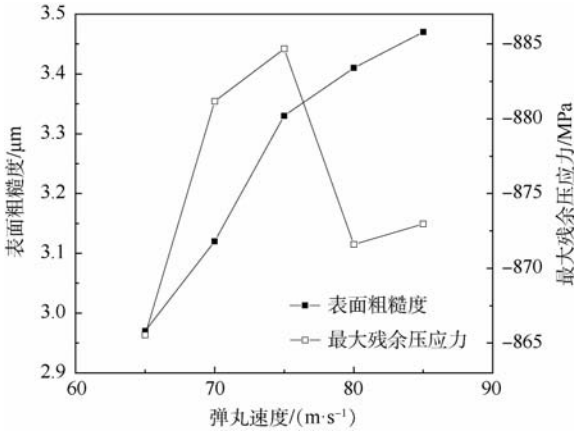


图8 不同弹丸速度对材料表面粗糙度与最大残余压应力的影响

Fig. 8 Effect of different projectile velocities on material surface roughness and maximum residual compressive stress

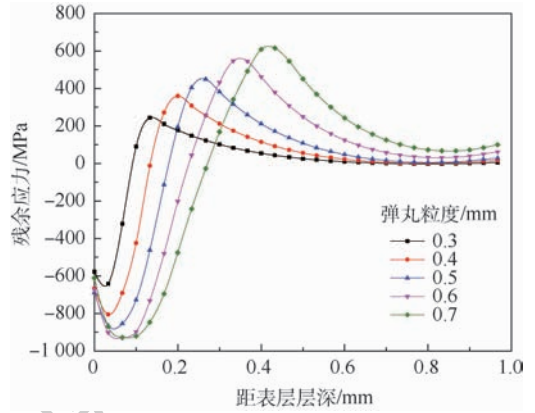


图9 不同弹丸粒度对残余应力场的影响

Fig. 9 Effect of different projectile particle sizes on residual stress field

由图7可以看出,在65~85 m/s的范围内,随着弹丸速度的增大,残余压应力层深增大,由65 m/s时的167 μm增至85 m/s时的200 μm,增幅为19.8%。由图8可以看到,弹丸速度从65 m/s增加至75 m/s范围内,最大残余压应力从-854.527 MPa上升到-881.164 MPa,增幅为3.12%,但之后随着弹丸速度的增加,最大残余压应力值开始下降。另一方面,表面粗糙度数值显著增加(见图8),从65 m/s下的2.97 μm到85 m/s下的3.47 μm,增加了16.8%。原因在于参与撞击的弹丸个数一定的情况下,速度越大,单个弹丸所带有的能量也就越大,撞击形成的凹坑深度也就越大,表面粗糙度也就随之增大。可见,弹丸速度增大可以显著增加残余压应力层深,但总体来说,弹丸速度增加对残余应力分布带来的增益影响相比表面质量的恶化来说并不显著,这对提高材料的疲劳寿命是非常不利的。在实际的喷丸过程中,在保证材料表面残余应力指标的前提下,应该严格控制喷丸强度(速度),防止由于喷丸速度过大带来的粗糙度增大、表面质量恶化、应力集中系数提高,反而减小受喷结构件的疲劳寿命。

## 2) 弹丸粒度对表面强化特征参数的影响

弹丸尺寸也对喷丸强度有着至关重要的影响。为研究不同弹丸粒度对TC4材料表面强化特征参数的影响规律,取弹丸速度为65 m/s,喷丸覆盖率为100%,分别模拟0.3、0.4、0.5、0.6、0.7 mm弹丸粒度的撞击情况,得到不同弹丸粒度下的残余应力场分布情况与表面粗糙度,如图9和图10所示。

由图9可以看出,当弹丸粒度由0.3 mm增加

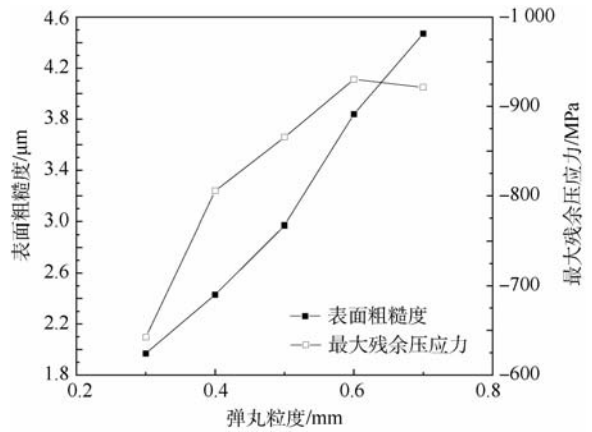


图10 不同弹丸粒度对材料表面粗糙度与最大残余压应力的影响

Fig. 10 Effect of different projectile particle sizes on material surface roughness and maximum residual compressive stress

至0.7 mm时,残余压应力层深由67 μm增加至267 μm,增幅显著,但同时材料次表层以下残余拉应力值显著增大。由图10可以看到,随着弹丸粒度的增大,残余压应力最大值由-642.628 MPa增加至-928.293 MPa,但当弹丸粒度超过0.6 mm时,残余应力最大值变化并不明显,同时TC4材料表面粗糙度随着弹丸粒度的增大而增大,且当弹丸粒度大于0.5 mm时,表面粗糙度的增大趋势更为明显。可见,在一定范围内,弹丸粒度的增大有利于改善残余应力场的分布,提高残余压应力的层深,增大最大残余压应力值,但当弹丸粒度达到一定值(0.5 mm)时,其值的继续增大一方面对残余压应力层深与最大残余压应力的提高不明显,另一方面材料表面粗糙度也急剧增大,这对材料的抗疲劳性能是非常不利的。

### 3) 喷丸覆盖率对表面强化特征参数的影响

喷丸覆盖率属于喷丸参数之一,喷丸覆盖率过低时,材料表面由于受喷不均匀容易产生残余拉应力,影响材料的疲劳寿命;另一方面,当喷丸覆盖率过高时,会增加喷丸时间,降低喷丸效率,同时对于各类航空薄壁件来说,喷丸覆盖率过大可能会造成表面的微变形,改变几何参数,影响工作效率。因此,喷丸覆盖率参数的控制对于喷丸效果具有重要意义。为研究喷丸覆盖率参数对TC4材料表面强化特征参数的影响规律,取弹丸速度为65 m/s,弹丸粒度为0.5 mm,分别模拟100%、150%、200%、250%和300%喷丸覆盖率下的撞击情况。在喷丸仿真之前,首先需要通过子程序确定不同喷丸覆盖率下所需生成的弹丸个数,如在弹丸速度65 m/s、弹丸粒度0.5 mm下,达到100%喷丸覆盖率所需的弹丸个数为65,因此达到150%、200%、250%、300%喷丸覆盖率所需的弹丸个数分别为98、130、165与195。得到不同喷丸覆盖率参数下的残余应力场分布情况与表面粗糙度,如图11和图12所示。

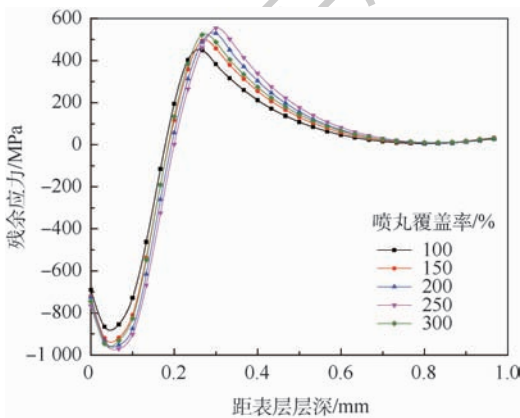


图11 不同喷丸覆盖率对残余应力场的影响

Fig. 11 Effect of different shot peening coverage on residual stress field

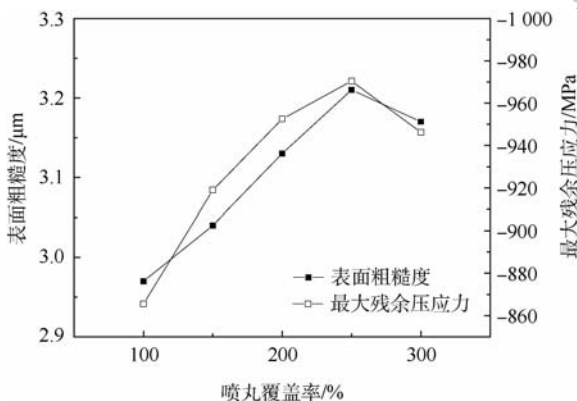


图12 不同喷丸覆盖率对材料表面粗糙度与最大残余压应力的影响

Fig. 12 Effect of different peening coverage on material surface roughness and maximum residual compressive stress

由图11可以看到,当喷丸覆盖率由100%增加到250%时,残余压应力层深度逐渐变大但变化不明显。由图12可以看到,最大残余压应力值由100%喷丸覆盖率时的-865.569 MPa上升至250%喷丸覆盖率时的-970.41 MPa,之后甚至略有下降。总体来说,喷丸覆盖率的变化对残余应力场的影响体现在一定范围以内,当超过该范围时影响并不显著。究其原因,当喷丸覆盖率超过一定值时,材料表层组织经过了充分撞击,并发生硬化现象,之后随着喷丸覆盖率的增加,弹丸的撞击不足以使材料产生更为剧烈的塑性变形,使得材料的残余应力分布特征基本保持不变,残余压应力层深及最大残余压应力值维持稳定,这与喷丸表面纳米化相关研究中当喷丸时间增加到一定值时,“变形层深度增加变缓而最终趋于稳定”<sup>[16]</sup>的试验现象是一致的。同时由图12可以看到,材料表面粗糙度随着喷丸覆盖率的增大而增大,这是由于随着残余撞击弹丸数量的增多,材料表面凹坑增多,表面粗糙度也就随之增大;但当喷丸覆盖率增大到一定程度(250%以上)时,表面粗糙度反而开始减小,原因在于当参与撞击的弹丸个数达到一定值时,材料表面的凹坑数量不再发生明显变化,但凹坑边缘的“凸起”在撞击作用下变得相对“平缓”,表面粗糙度反而减小。在实际的喷丸过程中,在一定范围内喷丸覆盖率的提高会增大材料表面粗糙度,超出该范围(250%)后尽管材料表面粗糙度会出现一定程度的减小,但考虑到时间成本的提高及喷丸效率的下降,同时最大残余压应力值增幅不大且残余压应力层深基本保持稳定,因此喷丸覆盖率应该严格控制在100%~200%之间。

## 5 喷丸分析模型试验验证

为了验证第1节所建立的喷丸分析模型,本节进行了TC4喷丸试验,对喷丸后样件进行了残余应力测量与表面粗糙度测量,并与仿真结果进行了对比。

实际喷丸过程中的工艺参数主要有喷射压力 $p_s$ 与弹丸流量 $\dot{m}$ ,弹丸粒度 $D_s$ 由所用的弹丸牌号确定。弹丸速度 $V_s$ 与喷射压力 $p_s$ 、弹丸流量、弹丸直径 $D_s$ 之间具有如下转换关系<sup>[17]</sup>:

$$V_s = 1.55 \frac{16.35 p_s}{1.53 \dot{m} + p_s} + \frac{29.5 p_s}{0.598 D_s + p_s} + 4.83 p_s \quad (5)$$

喷丸试验采用的是 ROSLER 气动式喷丸机。

喷丸设备如图 13 所示。受喷 TC4 试样如图 14 所示,且仅对试样中间 40 mm 进行喷丸处理。试验所用弹丸牌号为 CW14,喷丸工艺参数设置为 2.5 bar(1 MPa = 10 bar),流量为 3 kg/min,经过计算此时对应的弹丸速度为 70 m/s,因此将此工况下的试样残余应力检测结果与弹丸速度 70 m/s、弹丸粒度 0.36 mm、喷丸覆盖率 100% 的仿真结果进行对比。

为验证喷丸有限元模型的可行性,首先对一喷丸后试样进行  $x$  方向残余应力测量。检测电压为 20 kV,电流为 4 mA,靶材/射线为 V 靶材-K $\alpha$  射线,衍射角度为 139.7°,103 晶面,曝光时间为 40 s,回摆角度为 -18°~18°。在测量过程中,为获得试样沿层深的残余应力分布,需对试样进行电解剥层,单次的剥层深度由电解时间控制,之后利用千分表记录剥层深度,最终达到逐层抛光、逐层测量残余应力的目的。残余应力测量设备如图 15 所示。

对比结果如图 16 所示。对比试验与仿真结果可以看到,仿真与试验得到的残余应力分布曲线具有相同的变化趋势,即从表面开始沿层深的残余压应力值先增大后减小,且都在距表层约

30  $\mu\text{m}$  处达到最大;在残余应力影响层深方面,在 260  $\mu\text{m}$  处的残余应力仿真值为 -40.157 MPa,试验值为 -32.22 MPa,可以认为在 260  $\mu\text{m}$  处实际试样与仿真模型都达到了喷丸最大影响层深;同时可以看到,经喷丸工序后的样件残余应力试验测量值相比仿真值偏低,原因可能在于实际试验试样在制备过程中内部残留了残余拉应力,对喷丸效果起到了一定的抵消作用。

表面粗糙度的测量采用 VHX-6000 超景深显微镜如图 17 所示,测量标准为 GB/T 1031—2009<sup>[18]</sup>,取样长度为 2.5 mm,评定长度为 12.5 mm。最终测量值  $R_a = 2.89$ ,仿真值  $R_a = 3.16$ ,误差为 9.34%。

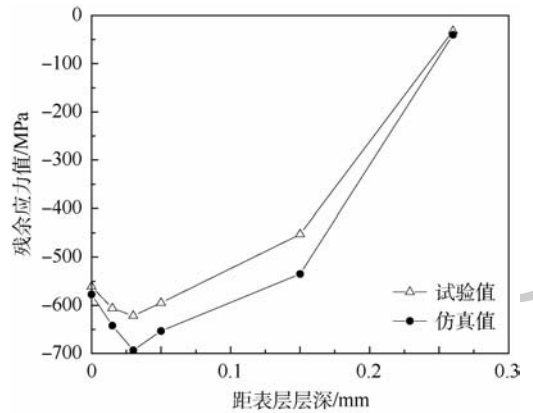


图 16 残余应力场试验值与仿真值对比

Fig. 16 Comparison between experimental values and simulated values of residual stress field



图 13 ROSLER 喷丸机

Fig. 13 ROSLER shot peening machine

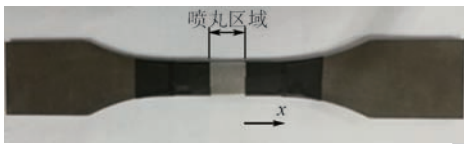


图 14 TC4 试样

Fig. 14 TC4 sample



图 15 Proto iXRD 型 X 射线衍射仪

Fig. 15 Proto iXRD X-ray diffractometer



图 17 VHX-6000 超景深显微镜

Fig. 17 VHX-6000 ultra-depth-of-field microscope

总体来说,仿真与试验残余压应力场具有相同变化趋势,且最大误差在 25% 之内,表面粗糙度仿真值相比实际值误差在 10% 以内。因此,可以认为本文建立的空间随机分布弹丸有限元模型很好地模拟了 TC4 的实际被喷过程。

## 6 结 论

本文基于 ABAQUS 二次开发建立了空间位置随机分布、与喷丸覆盖率相关的多弹丸喷丸模型,研究了不同喷丸参数对 TC4 材料表面残余应

力分布及表面粗糙度的影响规律,并开展相关TC4试样喷丸试验,得到以下结论:

1) 经喷丸试验的TC4试样相比仿真结果具有相同的沿层深残余应力变化趋势,且同一深度的残余压应力最大误差不超过25%,粗糙度误差不超过10%,验证了本文所建立的喷丸有限元仿真模型预测残余应力分布及表面粗糙度的可行性;除了TC4外,也可以将该模型推广到其他金属材料的喷丸仿真中。另一方面,TC4喷丸试验残余应力的测试结果相比仿真结果偏小,究其原因可能在于试样在加工过程中形成的残余拉应力还有残留,抵消了部分喷丸效果;而表面粗糙度的仿真值相比试验所得偏大,原因可能在于仿真模型中的弹丸为刚体,撞击的动能完全转化为了材料表面的变形能,使得表面粗糙度更大。

2) 弹丸速度增加,残余压应力层深随之增大,最大残余压应力也随之增大,但最大残余压应力的增幅并不明显,同时材料表面粗糙度却急剧增大;弹丸粒度的增加对提高残余压应力层深及最大残余压应力值具有非常明显的作用,但当弹丸粒度超过一定值时,增幅放缓,且材料表面粗糙度急剧增大;喷丸覆盖率提高,残余压应力层深及最大残余压应力值变化波动不大,表面粗糙度呈现先增大后趋于稳定。

3) 对于弹丸速度、弹丸粒度与喷丸覆盖率3个参数,弹丸粒度对于TC4材料残余压应力分布的影响最为显著,弹丸速度次之,喷丸覆盖率对残余压应力层深及最大残余压应力值影响最小,但另一方面,弹丸粒度的变化会对材料表面粗糙度产生非常显著的影响,基于此,在喷丸过程中,为得到理想的残余应力分布同时保证工件表面质量,应首先选取合适牌号的弹丸,在此基础上调整喷丸压力。同时,为保证工件的表面质量,喷丸覆盖率宜控制在100%~200%之间。

### 参考文献 (References)

- [1] MIAO H, ZUO D W, WANG M, et al. Numerical calculation and experimental research on residual stresses in precipitation-hardening layer of NAK80 steel for shot peening[J]. Chinese Journal of Mechanical Engineering, 2011, 24(3): 439-445.
- [2] 张洪伟, 张以都, 吴琼. 喷丸强化残余应力场三维数值分析[J]. 航空动力学报, 2010, 25(3): 603-609.  
ZHANG H W, ZHANG Y D, WU Q. Three-dimensional numerical analysis of residual stress field for shot-peening[J]. Journal of Aeronautical Power, 2010, 25(3): 603-609 (in Chinese).
- [3] MORI K, OSAKADA K, MATSUOKA N. Finite element analysis of peening process with plastically deforming shot[J]. Journal of Materials Processing Technology, 1994, 45(1-4): 607-612.
- [4] GUAGLIANO M. Relating Almen intensity to residual stresses induced by shot peening: A numerical approach[J]. Journal of Materials Processing Technology, 2001, 110(3): 277-286.
- [5] MEGUID S A, SHAGAL G, STRANART J C. 3D FE analysis of peening of strain-rate sensitive materials using multiple impingement model[J]. International Journal of Impact Engineering, 2002, 27(2): 119-134.
- [6] HONG T, OOI J Y, SHAW B A. A numerical study of the residual stress pattern from single shot impacting on a metallic component[J]. Advances in Engineering Software, 2008, 39(9): 743-756.
- [7] BARAGETTI S. Three-dimensional finite-element procedures for shot peening residual stress field prediction[J]. International Journal of Computer Applications in Technology, 2001, 14(1-3): 51-63.
- [8] KLEMENZ M, SCHULZE V, VÖHRINGER O, et al. Finite element simulation of the residual stress states after shot peening[C]// Materials Science Forum, 2006, 524-525: 349-354.
- [9] MAJZOBI G H, AZIZI R, NIA A A. A three-dimensional simulation of shot peening process using multiple shot impacts[J]. Journal of Materials Processing Technology, 2005, 164(20): 1226-1234.
- [10] WANG C, WANG L, WANG X G, et al. Numerical study of grain refinement induced by severe shot peening[J]. International Journal of Mechanical Sciences, 2018, 146-147: 280-294.
- [11] STRANART J C. Finite element modelling of shot-peening residual stresses[J]. Journal of Materials Processing Technology, 1999, 92-93(9): 401-404.
- [12] MEO M, VIGNJEVIC R. Finite element analysis of residual stress induced by shot peening process[J]. Advances in Engineering Software, 2003, 34(9): 569-575.
- [13] 夏明莉, 刘道新, 杜东兴, 等. 喷丸强化对TC4钛合金表面完整性及疲劳性能的影响[J]. 机械科学与技术, 2012, 31(8): 1349-1353.  
XIA M L, LIU D X, DU D X, et al. Effects of shot peening on surface integrity and fatigue properties of TC4 titanium alloy[J]. Mechanical Science and Technology for Aerospace Engineering, 2012, 31(8): 1349-1353 (in Chinese).
- [14] 鲁世红, 何宁. TC4钛合金动态本构模型与高速切削有限元模拟[J]. 兵器材料科学与工程, 2009, 32(1): 5-9.  
LU S H, HE N. Dynamic constitutive model of TC4 alloy material and finite element simulation of orthogonal high-speed cutting[J]. Ordnance Material Science and Engineering, 2009, 32(1): 5-9 (in Chinese).
- [15] LEUSE D. Experimental investigation of material models for Ti6Al4V and 202-T3 aluminum: DOT/FAA/AR-00/25[R]. Washington, D. C.: FAA, 2000.
- [16] 温爱玲, 陈春焕, 郑德有, 等. 高能喷丸表面纳米化对工业纯钛组织性能的影响[J]. 表面技术, 2003, 32(3): 16-18.  
WEN A L, CHEN C H, ZHEN D Y, et al. Effects of surface nanocrystallization induced by high energy shot peening on commercial pure titanium[J]. Surface Technology, 2003, 32(3): 16-18 (in Chinese).
- [17] 吕鹤婷. 喷丸对疲劳寿命影响的数值模拟研究[D]. 济南:

山东大学,2016:19.

LV H T. Numerical simulation research of the effect of shot-peening on fatigue life[D]. Jinan:Shandong University,2016:19 (in Chinese).

[18] 全国产品尺寸和几何技术规范标准化技术委员会. 产品几何技术规范(GPS)表面结构 轮廓法 表面粗糙度参数及其数值:GB/T 1031—2009[S]. 北京:中国标准出版社,2009:1-5.

National Technical Committee for Product Dimensions and Geometric Specifications. Geometrical product specifications (GPS)-Surface texture: Profile method-Surface roughness pa-

rameters and their values:GB/T 1031—2009[S]. Beijing:China Standard Press,2009:1-5(in Chinese).

#### 作者简介:

王延忠 男,博士,教授,博士生导师。主要研究方向:机械传动、摩擦润滑。

李菲 男,硕士研究生。主要研究方向:表面强化。

陈燕燕 女,博士。主要研究方向:抗疲劳加工。

## TC4 shot peening simulation and experiment

WANG Yanzhong<sup>1</sup>, LI Fei<sup>1</sup>, CHEN Yanyan<sup>1,\*</sup>, ZHANG Yaping<sup>1</sup>, WU Zegang<sup>2</sup>, WANG Cheng<sup>3</sup>

(1. School of Mechanical Engineering and Automation, Beihang University, Beijing 100083, China;

2. AECC Harbin Dongan Engine Co., Ltd., Harbin 150066, China;

3. China North Vehicle Research Institute, Beijing 100072, China)

**Abstract:** In view of the current research status on the finite element simulation study of shot peening, such as the projectiles' regular arrangement and the peening coverage cannot be calculated precisely, a finite element model with multiple randomly distributed shots whose number depends on the peening coverage was secondarily developed with ABAQUS, in which accurate number of required shots under different working conditions and different peening coverage is calculated. The influence of peening parameters on the surface residual stress and the surface roughness of TC4 titanium alloy material was studied. The TC4 shot peening experiment was carried out, residual stress distribution and surface roughness obtained by experiment were compared with those by simulation, and the results verified the reasonable prediction of residual stress distribution and surface roughness by the model, which has certain guiding significance for the determination of shot peening process parameters.

**Keywords:** TC4 titanium alloy; shot peening model; random distribution; coverage; residual compressive stress

**Received:** 2018-12-20; **Accepted:** 2019-01-18; **Published online:** 2019-04-12 08:36

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20190409.1609.004.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190409.1609.004.html)

**Foundation items:** National Defense Basic Research Program (JCKY2016220C002); Innovation Research Project of Basic Product (237099000000170006)

\* **Corresponding author.** E-mail: besthouniao@foxmail.com

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

DOI: 10.13700/j.bh.1001-5965.2018.0566

# 基于出行链的电动汽车充电行为影响因素分析

于海洋<sup>1,2,3</sup>, 张路<sup>1,2,3</sup>, 任毅龙<sup>1,2,3,\*</sup>

1. 北京航空航天大学 交通科学与工程学院, 北京 100083; 2. 北京航空航天大学 车路协同与安全控制北京市重点实验室, 北京 100083; 3. 北京航空航天大学 大数据科学与脑机智能高精尖创新中心, 北京 100083

**摘 要:** 随着电动汽车的快速发展,大规模电动汽车充电将给电力系统规划和运行带来不可忽视的影响,研究电动汽车的充电行为及其影响因素,并实时预测潜在的充电行为越发迫切。基于北京市私家电动汽车的历史车联网数据,引入出行链的概念,综合电动汽车充电过程和放电过程的数据,从实际出发考虑影响电动汽车充电行为的多种潜在因素,并通过logistic回归模型分析确定了显著影响充电行为的因素。分别基于单一和多个显著影响因素建立电动汽车充电行为的预测模型,预测结果表明,基于多个显著影响因素的模型准确率更高,且对晴天的预测效果更好。研究成果将有助于优化电动汽车的充电行为,进而提高电动汽车的充电效率。

**关键词:** 电动汽车; 充电行为; 出行链; 影响因素; 预测模型

**中图分类号:** U121; U469.72<sup>+2</sup>

**文献标识码:** A **文章编号:** 1001-5965(2019)09-1732-09

随着地球上资源的日益贫乏,绿色能源越来越受到人们的重视,节能环保也成为汽车技术发展的方向,电动汽车作为新能源汽车的代表,已被公认为是减少碳排放和解决石油短缺的重要技术<sup>[1-2]</sup>,且发展速度日益加快。但是由于电动汽车用户需求和行为的不确定性与相互差异,大规模电动汽车充电具有时间和空间上的随机性、间歇性、波动性等不确定特点,将给电力系统规划和运行带来不可忽视的影响<sup>[3-4]</sup>。因此,考虑多种潜在因素的影响,深入分析电动汽车的充电行为,能够为研究大规模电动汽车接入带来的影响与配电网优化调度策略打下良好的基础。

当前,国内外学者对电动汽车充电行为的研究很多。国内研究大多基于传统车辆行驶数据模拟电动汽车的运行规律,考虑不同情况下充电行为的差异,分析电动汽车充电负荷的计算方法。

文献[5]考虑电动汽车的空间动态转移特性下充电行为的差异,预测不同类型电动汽车的空间分布,并建立具体详细的负荷预测模型。文献[6]基于居民出行调研统计数据构建出行链,提出一种涉及电动汽车充电频率的负荷计算方法。文献[7]利用蒙特卡罗法模拟电动汽车的出行特性,提出一种考虑气温影响的电动汽车充电负荷计算方法。文献[8]考虑充电起始时间和日行驶里程2个影响因素,用蒙特卡罗仿真方法求得电动汽车功率需求,建立电动汽车功率需求的统计学模型。文献[9]基于调研数据分析不同类型电动汽车的充电行为,提出基于蒙特卡罗模拟的电动汽车充电负荷计算模型。文献[10]根据北京市机动车行驶规律,探讨对充电需求的影响因素(用户行驶开始时间、结束时间和行驶里程),采用蒙特卡罗法建立通勤用电动汽车充电需求模型。

收稿日期: 2018-09-28; 录用日期: 2019-04-21; 网络出版时间: 2019-05-08 10:00

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20190506.1423.006.html

基金项目: 国家重点研发计划(2018TFB1600702)

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

**引用格式:** 于海洋, 张路, 任毅龙. 基于出行链的电动汽车充电行为影响因素分析[J]. 北京航空航天大学学报, 2019, 45(9): 1732-1740. YU H Y, ZHANG L, REN Y L. Influential factors analysis of electric vehicle charging behavior based on trip chain[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1732-1740 (in Chinese).

国外对电动汽车充电行为的研究也非常重视,文献[11]发现消费者通常选择在晚上充电,但在高峰期,消费者更有可能使用快速充电桩。文献[12]以燃油汽车数据类比电动汽车数据,采用Logit模型进行分析,表明初始电荷状态、充电时间、起点与充电站的距离对电动汽车的路径选择有影响。文献[13]针对私家车GPS数据,考虑每个停车事件中电池的实际充电状态、停车持续时间、停车类型和车辆传动系,引入新的随机方法更精确地预测负荷。文献[14]基于对私家车GPS行驶数据的统计分析,分析电动汽车在不同场景下的充电时间分布特性,利用排队论模型预测电动汽车的负荷需求。文献[15]采用实际运营的插入式电动汽车的数据,提出基于回归和时间序列方法的组合模型,预测电动汽车充电时长。文献[16]利用3种不同的电动汽车充电数据来估算电动汽车的潜在电力需求,并模拟其对电网的影响。文献[17]基于排队理论分析电动汽车到站信息,提出电动汽车充电需求的随机模型,对电动汽车充电负载波动进行实时更清晰的短期预测。文献[18]研究电动汽车对低压配电网的影响,针对2种充电方案进行建模,评估电动汽车对变压器负载、馈线负载、电压偏差和总系统损耗的影响。

综上所述,由于缺少真实的电动汽车数据,现有研究大多采用蒙特卡罗模拟或人工调查的方法进行分析验证。但是,人工调查数据的准确率较低,而且数据的样本量不足;而模拟的数据多假定充电条件,忽视了电动汽车充电行为的自主选择性,很难反应真实情况,缺乏说服力;少部分采用的实际数据主要来自于美国的插入式电动汽车,由于中国交通和充电环境与美国不同,充电行为特征也是不同的。多数对充电行为的研究只分析了充电过程,忽略了放电过程对充电行为的影响。另外,部分考虑充电行为影响因素的文献只考虑了单一因素的影响,且着重研究电动汽车充电对于电网等其他问题的影响<sup>[19-20]</sup>,忽略了这些因素对充电行为本身的影响。

本文从实际出发,根据北京市私家电动汽车的历史车联网数据,针对电动汽车充电行为的不确定性,引入出行链的概念,以“链”的形式将电动汽车的时间维、空间维联系在一起,将电动汽车的充电过程和放电过程构建成一个整体,考虑电动汽车充电行为的多种潜在影响因素,统筹兼顾天气、旅程开始时间、出行链时长、出行链距离、旅程速度、起始电荷状态、上一次出行链状态等,通过

模型确定充电行为的显著影响因素,并基于显著影响因素对出行链是否充电进行建模和预测。

## 1 数据准备

本文采用电动汽车车联网数据。由于电动汽车充电行为与其出行规律紧密相关,为了更全面地分析电动汽车充电行为,引入出行链的概念。出行链是指用户为完成某项活动,按照时间顺序在不同区域间转移的连接形式<sup>[21]</sup>。本节首先对车联网数据进行简单描述,然后对数据集进行处理,得到基于出行链的充电行为数据,最后对充电行为数据进行总结分析。

### 1.1 电动汽车数据处理

本文所采用的电动汽车车联网数据由北京市130辆私家电动汽车行驶实况数据构成,共涵盖约5000万组行车数据信息,数据时间主要集中在2015年8月,平均每辆车约有15d的数据,每组行车数据平均时间间隔为10s;内容主要包括终端编号、充放电状态、续驶里程、动力电池SOC、放电电流、经度、纬度、里程和采集时间等。

首先对采集到的行车数据信息进行预处理,规范数据格式,填补缺失数据,剔除错误信息。然后根据动力电池电荷状态的变化情况,分别提取充电数据和放电数据。最后根据数据的结构组成特点及经验分析,规定车辆停止时间大于1h,则本次充放电过程或放电过程结束,实现单次充电过程和放电过程的识别。

在电动汽车的充放电过程中,上一次放电过程的结束时间、结束位置、荷电状态等对本次充电过程的开始时间、充电时长、充电位置等有很大的影响,因此将充电过程和放电过程结合起来可以更全面地分析充电行为。由于通常情况下出行链描述居民从家出发再返回家的一个过程<sup>[22]</sup>,过程中包含大量的时间、空间和活动类型信息,而且这些信息之间是相互联系和相互作用的。基于出行链的研究更能体现电动汽车充电过程和放电过程的关联性。

因此接下来通过经纬度在地图上匹配电动汽车的放电(即行驶)过程,然后根据经验将数据中晚上使用常规充电的充电地点视为家的所在地,根据时间和空点连续性对单次放电过程和充电过程进行匹配,以家为起讫点构建车辆出行链模型。同时记录单条出行链充电的次数、旅程起始点、总距离、总时间、旅行速度、起始电池核电量,以及每次快速充电的起始点、充电时长、充电量等。

## 1.2 充电行为数据总结

通过 1.1 节对电动汽车数据的处理可以看出,在出行链中的充电过程绝大多数都为快速充电,剔除 2 条使用常规充电的出行链数据,得到 120 辆电动汽车的 3 002 条出行链信息,按快速充电情况分为以下 4 种:无充电、快速充电 1 次、快速充电 2 次及快速充电 3 次,如表 1 所示。可以看到,发生快速充电行为的出行链远小于无快速充电行为,说明大部分情况下,电池电量能够满足旅程所需,用户习惯回到家后使用常规充电方式充电。

表 1 基于出行链的充电行为分布

Table 1 Distribution of charging behavior based on trip chains

充电次数	0	1	2	3
出行链分布	2 378	410	167	47

## 2 充电行为的潜在影响因素

本文更关心的是不同影响因素所引起的充电行为本身的差异。目前大多数对充电行为的研究多设定固定的充电模式或者假设在某个目的地充电,但实际用户也存在旅程中的充电行为。本文采用充电出行链频次(即指定情况下或一定时间内有充电行为的出行链数量)作为统计量,分析不同潜在影响因素所引起的基于出行链的充电频次差异情况。

对于充电行为的研究,最先考虑的便是天气问题,一方面当温度过高或过低时,司机会使用空调或者暖气,这将增加单位里程耗电量,增加电动汽车的充电需求。另一方面,遇到雨、雪等恶劣天气,会影响车辆的行驶速度,进一步影响电动汽车的充电行为。本文数据均发生在 2015 年 8 月,此时天气如表 2 所示。可见,气温差异并不明显,空调或暖气的使用并无差异,而且只存在部分小雨天气,未出现恶劣天气,因此天气的影响并不明显。

表 2 2015 年 8 月部分北京天气情况

Table 2 Some weather conditions in Beijing in August 2015

日期	最高气温/℃	最低气温/℃	天气
2015-08-01	31	23	阴转小到中雨
2015-08-02	30	23	多云转阴
2015-08-03	30	23	阵雨转多云
2015-08-04	33	22	晴转多云
2015-08-05	32	22	多云转雷阵雨
2015-08-06	31	22	多云

用户行驶过程中是否会充电往往取决于电动汽车电池现有剩余电量是否足以完成接下来的行程,起始电池核电量是非常重要的因素。由于本文所采用车辆充满电时的续航里程均在 145 ~ 150 km 之间,差异并不大,因此采用电池起始电荷状态对充电行为进行分析,电池电荷状态简称 SOC,主要用来反映电池的剩余容量,其数值上定义为剩余容量占电池容量的比值。随起始电荷状态的变化,充电出行链频次分布如图 1 所示。可以看出,起始电荷状态越高,有充电行为的出行链频次越低。

基于出行链的充电行为研究,旅程开始的时间被认为是必要的因素,电动汽车用户往往在晚上出行链结束时,到达住宅区后选择常规充电模式充电,因此越晚开始的旅程,中途选择快速充电的可能性越小。图 2 显示了一天 24 个时段出行的电动汽车充电频次变化情况。可以看出,上午开始的出行链充电的可能性更大。

在研究车辆充电行为时,出行链距离和时长的影响也很重要。对于出行链距离,以 50 km 为

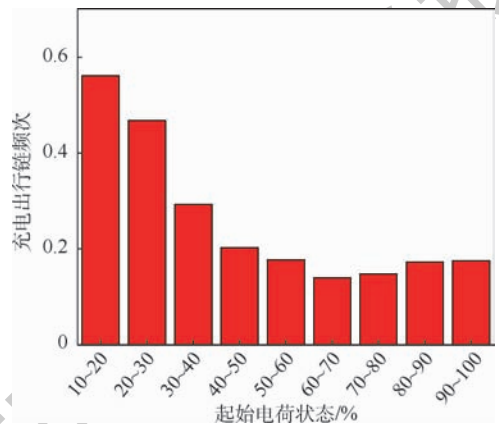


图 1 起始电荷状态的影响

Fig. 1 Influence of initial state of charge

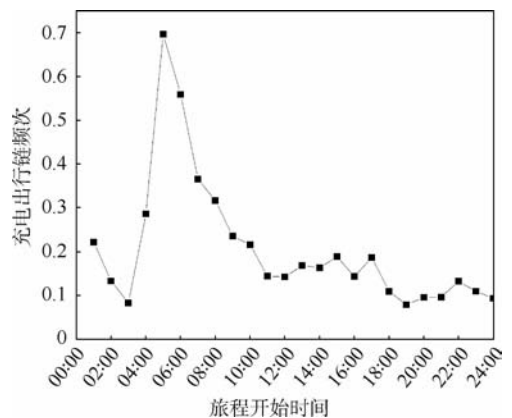


图 2 旅程开始时间的影响

Fig. 2 Influence of journey start time

出行链距离区间,计算出每个距离区间内的充电出行链频次,如图3所示,可以看出,出行链的充电频次和距离呈线性正相关关系。对于出行链时长的分析,将有充电行为和无充电行为的数据分成2个系列,绘制成折线图(见图4)。可以看到,有充电行为和无充电行为的出行链时长相差较大,有充电行为的出行链时长相对较大,平均在30 000~40 000 s,而无充电行为的出行链时长平均在5 000 s,相差一个数量级,因此出行链时间越长,充电的可能性就越大。

在研究车辆里程耗电量时,车辆速度是首先被考虑的因素,同样,对于出行链的充电行为,旅程平均速度也是一个重要的影响因素。首先计算各个出行链的平均速度,然后将平均速度从小到大归一化到 $[0,1]$ ,即将最大平均速度视为1,最小平均速度视为0,以0.05为一个区间绘制直方图(见图5)。可见,不同速度区间的出行链数量呈正态分布,当平均速度很小时,随着平均速度的增加,充电的出行链个数随之增加,当平均速度在0.30~0.35时达到最大,之后随着平均速度的增大,充电的出行链个数越来越少。

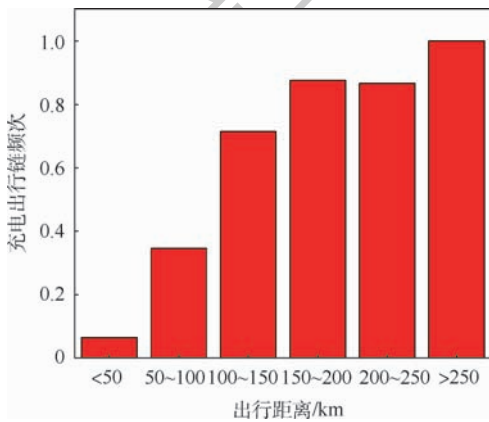


图3 出行链距离的影响

Fig.3 Influence of trip chain distance

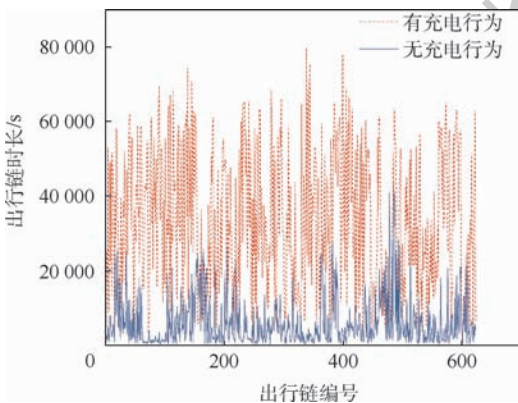


图4 出行链时长的影响

Fig.4 Influence of trip chain duration

将有充电行为的出行链状态计为1,无充电行为的出行链状态计为0,则上一次和本次的出行链的状态组合有:0-0、0-1、1-0、1-1,将这4种组合的出行链个数占比绘制成如图6所示的饼状图。可以看出,0-0即上一次和本次出行链均无充电行为,占77%,这说明大部分私家车用户选择常规充电模式充电,只有少部分用户使用快速充电方式;1-1即上一次和本次出行链均充电,占16%,说明有快速充电习惯的用户下次出行链也充电的概率更大。

基于上述分析,影响出行链充电行为的潜在影响因素主要包括天气、旅程开始时间、出行链时长、出行链距离、旅程速度、起始电荷状态及上一次出行链状态。

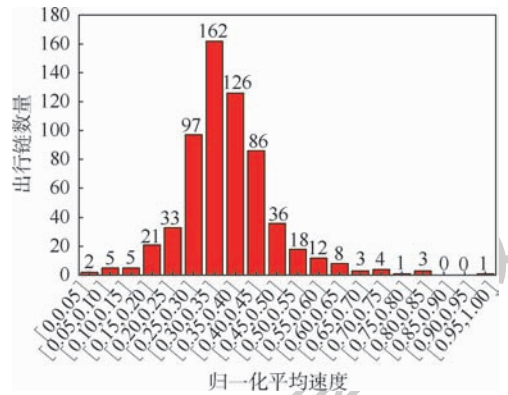


图5 旅程速度的影响

Fig.5 Influence of journey speed

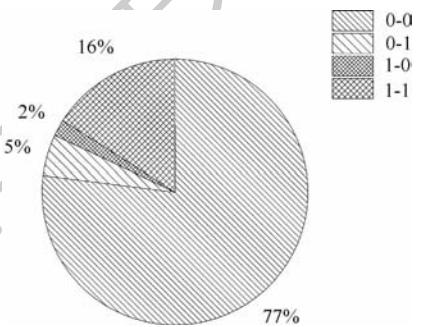


图6 上一次出行链状态的影响

Fig.6 Influence of last trip chain state

### 3 充电行为的显著影响因素

本文选取 logistic 回归模型分析确定充电行为的显著影响因素,该模型对自变量的分布要求不高,适用范围广,既能够很好地解决非线性问题,又可以解决多个自变量的问题,筛选出对因变量有显著影响的因素。

#### 3.1 logistic 回归模型

logistic 回归分析是一种广义的线性回归分

析模型,二元 logistic 回归模型则指因变量分为 2 类,取值 0 或 1。本文研究出行链的充电行为,将其视为二分类数据,用数字“0”代表没有充电行为的出行链,并将充电次数为 1、2、3 的旅程一起并称为有充电行为的出行链,用数字“1”代表。出行链的充电行为模型可以表示为

$$\text{logist}[\pi_i] = \text{lb} \left[ \frac{\pi_i}{1 - \pi_i} \right] = \hat{\alpha} + \hat{\beta}X_i \quad (1)$$

式中:  $\pi_i$  为第  $i$  条出行链有充电行为的概率,  $1 - \pi_i$  为第  $i$  条出行链没有充电行为的概率;  $\hat{\alpha}$  和  $\hat{\beta}$  分别为采用最大似然估计法得出的有充电行为概率的截距和系数矩阵;  $X_i = (x_{1,i}, x_{2,i}, \dots, x_{k,i})$  为影响第  $i$  条出行链是否充电的各类参数,  $x_{k,i}$  为影响充电行为的第  $k$  种潜在因素变量。将所有潜在影响因素变量放入此模型,可以筛选得到显著性水平在 0.05 以下的统计意义显著的参数。

### 3.2 显著影响因素的确定

对第 2 节提到的潜在影响因素进行汇总和简单描述,并将因素变量数字化,如表 3 所示。从表 2 可以看到,2015 年 8 月北京的天气呈现晴、晴转多云、晴转阴、多云、多云转阴、多云转雷阵雨、雷阵雨、阴转小到中雨、雷阵雨转中雨、阵雨等多种情况,为了便于分析,将其划分为晴天和雨天,视为二分类数据,晴天用 0 表示,雨天用 1 表示。将旅程开始时间按照早晚划分,只取开始时间对应的小时数,则旅程开始时间被分为 0-1, 1-2, ..., 23-24 等 24 个时段。对于出行链的时长,只

表 3 出行链充电行为的潜在影响因素

Table 3 Potential influential factors affecting charging behavior of travel chain

因素变量	描述	取值
天气	目标出行链所在日的天气情况	分晴雨天,晴天为 0,雨天为 1
旅程开始时间	目标出行链开始行驶的时间	24 h 制,只取出发对应的小时数
出行链时长	目标出行链中车辆行驶的总时长	只计放电过程时间,不计充电时间
出行链距离	目标出行链中车辆行驶的总距离	旅程中车辆行驶的距离,单位 km
旅程速度	目标出行链中车辆的平均行驶速度	出行链距离/时长,不计充电时间
起始电荷状态	目标出行链开始时电池的电荷状态	开始时剩余容量/电池容量,取值 0 ~ 100
上一次出行链状态	目标出行链之前一次出行链充电行为	取值 0 和 1, 1 为有充电, 0 为无充电

计入旅程放电时间,不计充电时间,并统一换算为以  $s$  为单位的数据。出行链距离则是整个出行链中车辆行驶的总距离,以  $\text{km}$  为单位。旅程速度为出行链距离与时长的比值,单位统一为  $\text{km/h}$ ,直接录入速度数据。起始电荷状态在数值上定义为剩余容量占电池容量的比值,其值的范围在 0 ~ 100 之间。上一次出行链状态是指目标出行链之前一次出行链的充电状态,将其视为二分类数据,数字 0 表示无充电行为,数字 1 表示有充电行为。

使用 SPSS 软件,把表 3 列出的所有潜在影响因素变量都放入模型进行回归,为了更好地拟合模型,根据回归结果剔除影响不显著的因素,保留显著性水平在 0.05 以下的统计意义上显著的参数,各因素的显著性水平如表 4 所示。表中:  $B$  为回归系数和截距(常数项),可以是负数(负相关时回归系数出现负值);  $SE$  为标准误差;  $Wals$  为卡方值,等于  $B$  除以其标准误差( $SE$ )的平方值;  $df$  为自由度;  $P$  为显著性;  $\text{Exp}(B)$  称为 OR,等于  $B$  值的返自然对数。

从表 4 中可以看出,旅程速度和天气是影响不显著的因素。旅程速度受用户自身影响较大,且有距离和时长的比值确定,相对来说对充电行为的影响并不显著;本文数据均发生在 2015 年 8 月,天气变化并不明显,只存在晴雨天差别,没有冰雪等恶劣天气,所以天气的影响并不显著。

旅程开始时间、出行链时长、出行链距离、起始电荷状态和上一次出行链状态这 5 个变量达到显著性水平,且在 0.01 水平上显著。不同时段开始的出行链,途中充电的概率也大不相同,电动汽车用户往往在晚上出行链结束时,到达住宅区后选择常规充电模式充电,则出行链越晚开始中途快速充电的可能性越小。出行链的距离和时长直接可以判断电动汽车的电量能否完成本次出行

表 4 基于潜在影响因素的模型分析结果

Table 4 Model analysis results based on potential influential factors

因素变量	$B$	$SE$	$Wals$	$df$	$P$	$\text{Exp}(B)$
旅程开始时间	0.032	0.017	3.485	1	<0.01	1.033
出行链时长	0	0	76.843	1	<0.01	1.000
出行链距离	0.012	0.004	7.817	1	<0.01	1.004
旅程速度	0.043	0.011	16.509	1	>0.05	1.044
起始电荷状态	-0.053	0.004	161.159	1	<0.01	0.949
上一次出行链状态	1.212	0.204	35.309	1	<0.01	3.359
天气	0.384	0.310	1.530	1	>0.05	1.468
常量	-2.651	0.428	38.438	1	<0.01	0.071

链,出行链的距离和时间越长,电动汽车充电的可能性越大。电动汽车电池起始电荷状态是否足以完成接下来的行程,是用户行驶过程中是否会充电的关键。根据电动汽车用户的规律,有快速充电习惯的用户会经常途中充电,因此上一次出行链充电的用户,目标出行链也充电的概率更大。

从回归系数( $B$ )可以看出,各自变量对因变量解释的变化方向,起始电荷状态的系数是负数,说明起始电荷状态越低越容易发生充电行为。综上,最终保留的显著影响因素为:旅程开始时间、出行链时长、出行链距离、起始电荷状态和上一次出行链状态。

## 4 充电行为预测模型

将3.2节确定的显著影响因素重新输入 logistic 回归模型,得到基于显著影响因素的各因素变量系数,根据系数确定充电行为的预测模型,将剩余数据集的显著影响因素变量输入充电行为的预测模型,并对模型实现的预测效果进行验证。

### 4.1 模型建立

将确定的显著影响因素重新输入 logistic 回归模型,确定各因素变量系数  $\hat{\alpha}$  与  $\hat{\beta}$ ,得到基于显著影响因素的充电行为回归模型。将式(1)进行变形,可以得到目标出行链有充电行为的概率计算方法。

$$p(Y_i = 1/X_i) = \pi_i = \frac{\exp(\hat{\alpha} + \hat{\beta}X_i)}{1 + \exp(\hat{\alpha} + \hat{\beta}X_i)} \quad (2)$$

式中: $Y_i$ 为第*i*条出行链的充电行为。如果车辆*i*有充电行为,则

$$y = \text{bool}(p_i > 0.5) \quad (3)$$

通过式(3)计算,最终判定  $p_i > 0.5$  为有充电行为的出行链。

### 4.2 预测结果

将确定的显著影响因素(旅程开始时间、出行链时长、出行链距离、起始电荷状态和上一次出行链状态)的数据重新输入 logistic 回归模型,并经过校正,确定各自变量对因变量的影响强度,得到基于显著影响因素的充电行为回归模型,具体结果如表5所示。可以看出,旅程开始时间、出行链时长、起始电荷状态和上一次出行链状态这4个变量在0.01水平上显著,出行链距离的影响在0.05水平上显著。

用3.1节得到的回归模型进行4.1节的变形,利用式(2)计算目标出行链有充电行为的概率,通过布尔函数公式(3)对剩余数据集的出行

链充电行为进行判定,并对预测模型实现效果进行验证,得到的验证结果如表6所示。充电行为整体的预测准确率达到84.77%,表示模型对出行链是否充电的行为识别度较高,预测效果较好。另外可明显看出,对无充电行为的出行链的识别率高于有充电行为。

为了比较单一和多个显著影响因素的预测模型,并将模型结果可视化,本文采用受试者工作特征曲线作图,该曲线简称 ROC 曲线,是根据一系列不同的二分类方式以真阳性率(灵敏度)为纵坐标,假阳性率(1-特异度)为横坐标绘制的曲线。ROC 曲线简单、直观,通过图示可观察分析方法的准确性,并可用肉眼作出判断。单一和多个显著影响因素的充电行为预测模型结果的 ROC 曲线如图7所示。结果表明,基于多个显著影响因素对充电行为预测的准确率更高,通过计算,基于多显著影响因素的曲线下面积(AUC)为0.893,AUC较大预测效果较好。

为了比较不同天气、不同情况下预测结果的差别,对不同情况下的预测结果分别绘制 ROC 曲线图,在对2种或2种以上预测情况进行比较时,以直观地鉴别优劣,靠近左上角的 ROC 曲线所代表的预测方法更准确,也可通过分别计算 ROC 曲线下的面积(AUC)进行比较,AUC越大预测效果越佳。

将预测结果中晴天和雨天2种数据划分开,分别绘制 ROC 曲线,如图8所示。可以看出,ROC曲线相差较大,通过计算得到晴天和雨天的

表5 基于显著影响因素的模型分析结果

Table 5 Model analysis results based on significant influential factors

因素变量	$B$	SE	Wals	df	$P$	Exp( $B$ )
旅程开始时间	0.030	0.018	2.919	1	<0.01	1.030
出行链时长	0	0	69.012	1	<0.01	1.000
出行链距离	0.022	0.004	38.200	1	<0.05	1.023
起始电荷状态	-0.052	0.004	163.219	1	<0.01	0.957
上一次出行链状态	1.120	0.201	31.071	1	<0.01	3.065
常量	-1.704	0.357	22.822	1	<0.01	0.182

表6 充电行为预测模型验证结果

Table 6 Model verification results for charging behavior prediction

有无充电出行链	实际个数	预测个数	预测准确个数	准确率/%	预测误差/%
充电出行链	126	120	92	76.67	23.33
无充电出行链	281	287	253	88.15	11.85
整体出行链	407	407	345	84.77	15.23

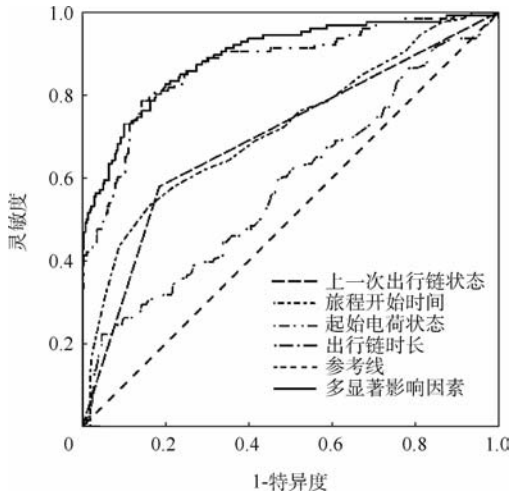


图7 充电行为预测模型的 ROC 曲线  
Fig. 7 ROC curves of charging behavior prediction model

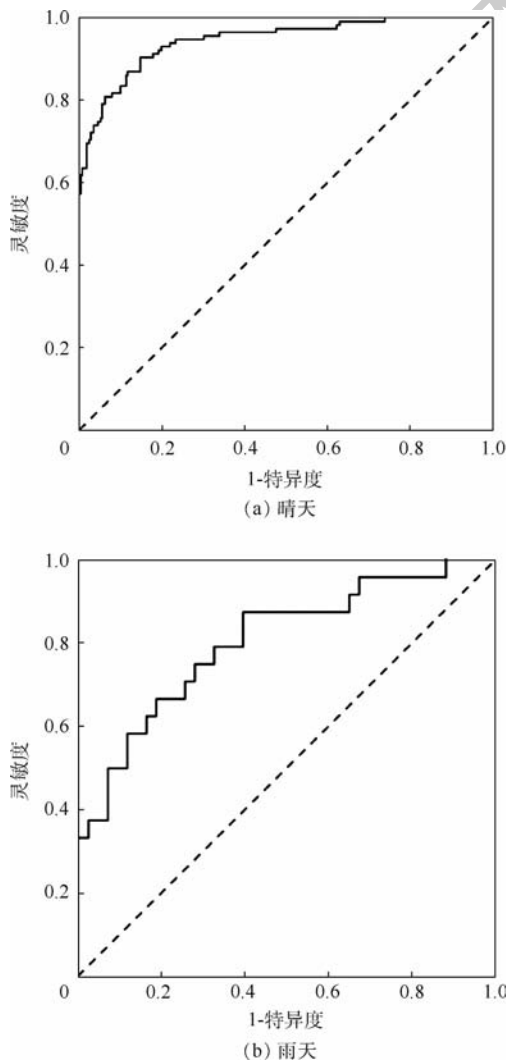


图8 不同天气情况下充电行为预测模型的 ROC 曲线  
Fig. 8 ROC curves of charging behavior prediction model under different weather conditions

ROC 曲线下面积分别为 0.911 和 0.805, 进一步说明该模型对晴天的预测效果更佳。

## 5 结论

本文基于出行链理论研究了电动汽车充电行为的影响因素,通过作图对各潜在影响因素做出了详细分析,并确定了电动汽车充电行为的显著影响因素,基于显著影响因素建立预测模型,对电动汽车的充电行为进行预测和验证。

1) 分析了天气、旅程开始时间、出行链时长、出行链距离、旅程速度、起始电荷状态、上一次出行链状态对电动汽车充电行为本身的影响。越晚开始的出行链中途快速充电的可能性越小;出行链的距离和时间越长,电动汽车充电的可能性越大;随着旅程速度的增加,充电频次的变化呈正态分布,电池起始电荷状态越低充电的频次越高;有快速充电的用户有惯性行为,上一次出行链充电的用户,目标出行链充电的概率更大。

2) 通过 logistic 回归模型分析,确定了本文中电动汽车充电行为的显著影响因素,包括旅程开始时间、出行链时长、出行链距离、起始电荷状态和上一次出行链状态。

3) 基于显著影响因素建立预测模型,基于多个影响因素对充电行为预测的准确率达到 84.77%;根据天气可以得出不同情况下的预测结果略有差别,对晴天的预测准确率更高。

本文虽然已经考虑了大多数因素,但由于数据的局限性,温度、天气、道路等级等因素尚未考虑;另外本文仅仅对单条出行链是否充电进行了预测,没有对具体的充电时间、充电时长进行建模和预测。为使本文提出的预测模型的准确率更高,接下来可以搜集更全面的数据,并将温度、恶劣天气、道路等级等因素考虑其中;在此基础上继续对充电次数、充电时间等进行预测,并进一步对出行链中的快速充电行为进行建模优化,为更准确地进行充电负荷预测、充电设施布设、电网规划建设奠定基础。

## 参考文献 (References)

- [1] BOULANGER A G, CHU A C, MAXX S, et al. Vehicle electrification: Status and issues [J]. Proceedings of the IEEE, 2011, 99(6): 1116-1138.
- [2] 张文亮, 武斌, 李武峰, 等. 我国纯电动汽车的发展方向及能源供给模式的探讨 [J]. 电网技术, 2009, 33(4): 1-5. ZHANG W L, WU B, LI W F, et al. Discussion on development trend of battery electric vehicles in China and its energy supply mode [J]. Power System Technology, 2009, 33(4): 1-5 (in Chinese).
- [3] 高赐威, 张亮. 电动汽车充电对电网影响的综述 [J]. 电网

- 技术,2011,35(2):127-131.
- GAO C W,ZHANG L. A survey of influence of electric vehicle charging on power grid[J]. Power System Technology,2011,35(2):127-131(in Chinese).
- [4] 胡泽春,宋永华,徐智威.电动汽车接入电网的影响与利用[J].中国电机工程学报,2012,32(4):1-10.
- HU Z C,SONG Y H,XU Z W. Impacts and utilization of electric vehicles integration into power systems[J]. Proceedings of the CSEE,2012,32(4):1-10(in Chinese).
- [5] 张艳娟,苏小林,闫晓霞,等.基于电动汽车时空特性的充电负荷预测[J].电力建设,2015,36(7):75-81.
- ZHANG Y J,SU X L,YAN X X, et al. A method of charging load forecast based on electric vehicle time-space characteristics[J]. Electric Power Construction,2015,36(7):75-81(in Chinese).
- [6] 陈丽丹,聂涌泉,钟庆.基于出行链的电动汽车充电负荷预测模型[J].电工技术学报,2015,30(4):216-225.
- CHEN L D,NIE Y Q,ZHONG Q. A model for electric vehicle charging load forecasting based on trip chains[J]. Transactions of China Electrotechnical Society,2015,30(4):216-225(in Chinese).
- [7] 王海玲,张美霞,杨秀.基于气温影响的电动汽车充电需求预测[J].电测与仪表,2017,54(23):123-128.
- WANG H L,ZHANG M X,YANG X. Electric vehicle charging demand forecasting based on influence of weather and temperature[J]. Electrical Measurement and Instrumentation,2017,54(23):123-128(in Chinese).
- [8] 田立亭,史双龙,贾卓.电动汽车充电功率需求的统计学建模方法[J].电网技术,2010,34(11):126-130.
- TIAN L T,SHI S L,JIA Z. A statistical model for charging power demand of electric vehicles[J]. Power System Technology,2010,34(11):126-130(in Chinese).
- [9] 罗卓伟,胡泽春,宋永华,等.电动汽车充电负荷计算方法[J].电力系统自动化,2011,35(14):36-42.
- LUO Z W,HU Z C,SONG Y H, et al. Study on plug-in electric vehicles charging load calculating[J]. Automation of Electric Power Systems,2011,35(14):36-42(in Chinese).
- [10] 杨冰,王丽芳,廖承林.大规模电动汽车充电需求及影响因素[J].电工技术学报,2013,28(2):22-27.
- YANG B,WANG L F,LIAO C L. Research on power charging demand of large-scale electric vehicles and its impacting factors[J]. Transactions of China Electrotechnical Society,2013,28(2):22-27(in Chinese).
- [11] MOON H B,PARK S Y,JEONG C, et al. Forecasting electricity demand of electric vehicles by analyzing consumers' charging patterns[J]. Transportation Research Part D:Transport and Environment,2018,62:64-79.
- [12] ASHTARI A,BIBEAU E,SHAHIDINEJAD S, et al. PEV charging profile prediction and analysis based on vehicle usage data[J]. IEEE Transactions on Smart Grid,2012,3(1):341-350.
- [13] STEEN D,CARLSON O,BERTLING L. Assessment of electric vehicle charging scenarios based on demographical data[J]. IEEE Transactions on Smart Grid,2012,3(3):1457-1468.
- [14] ALIZADEH M,SCAGLIONE A,DAVIES J, et al. A scalable stochastic model for the electricity demand of electric and plug-in hybrid vehicles[J]. IEEE Transactions on Smart Grid,2014,5(2):848-860.
- [15] WANG J,LIU K,YAMAMOTO T. Improving electricity consumption estimation for electric vehicles based on sparse GPS observations[J]. Energies,2017,10(1):129.
- [16] SCHÄUBLE J,KASCHUB T,ENSSLEN A, et al. Generating electric vehicle load profiles from empirical data of three EV fleets in Southwest Germany[J]. Journal of Cleaner Production,2017,150:253-266.
- [17] YANG Y,YAO E,YANG Z, et al. Modeling the charging and route choice behavior of BEV drivers[J]. Transportation Research Part C:Emerging Technologies,2016,65:190-204.
- [18] FERNANDEZ L P,ROMÁN T G S,COSENT R, et al. Assessment of the impact of plug-in electric vehicles on distribution networks[J]. IEEE Transactions on Power Systems,2011,26(1):206-213.
- [19] HÜBNER M,ZHAO L,MIRBACH T, et al. Impact of large-scale electric vehicle application on the power supply[C]//2009 IEEE Electrical Power & Energy Conference (EPEC). Piscataway,NJ:IEEE Press,2009:1-6.
- [20] IKEGAMI T,OGIMOTO K,YANO H, et al. Balancing power supply-demand by controlled charging of numerous electric vehicles[C]//2012 IEEE International Electric Vehicle Conference. Piscataway,NJ:IEEE Press,2012:1-8.
- [21] BOWMAN J L,BEN-AKIVA M E. Activity-based disaggregate travel demand model system with activity schedules[J]. Transportation Research Part A:Policy and Practice,2001,35(1):1-28.
- [22] 温剑锋,陶顺,肖湘宁,等.基于出行链随机模拟的电动汽车充电需求分析[J].电网技术,2015,39(6):1477-1484.
- WEN J F,TAO S,XIAO X N, et al. Analysis on charging demand of EV based on stochastic simulation of trip chain[J]. Power System Technology,2015,39(6):1477-1484(in Chinese).

#### 作者简介:

于海洋 男,博士,副教授,硕士生导师。主要研究方向:智能车路协同系统、交通大数据、交通控制与仿真。

张路 女,硕士研究生。主要研究方向:电动汽车的充电行为。

任毅龙 男,博士,讲师,硕士生导师。主要研究方向:智能车路协同系统、交通控制与仿真。

# Influential factors analysis of electric vehicle charging behavior based on trip chain

YU Haiyang<sup>1,2,3</sup>, ZHANG Lu<sup>1,2,3</sup>, REN Yilong<sup>1,2,3,\*</sup>

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

2. Beijing Key Laboratory for Cooperative Vehicle Infrastructure Systems and Safety Control, Beihang University, Beijing 100083, China;

3. Beijing Advanced Innovation Center for Big Data and Brain Computing, Beihang University, Beijing 100083, China)

**Abstract:** With the rapid development of electric vehicles, large-scale electric vehicle charging behavior will bring tremendous influence on the planning and operation of electric power systems. It is more and more urgent to study the charging behavior of electric vehicles and its influential factors, and predict the potential charging behavior in real time. Based on the historical data of private electric vehicles in Beijing, this paper introduces the concept of trip chain to comprehensively analyze the data of electric vehicle charging process and discharge process. This research considers the various potential influential factors on electric vehicles' charging behavior in the actual situation and determines the factors that significantly affect charging behavior through logistic regression analysis. Finally, the charging behavior forecasting model for electric vehicle is established based on the single and multiple significant influential factors. The results show that the model based on multiple significant influential factors has higher accuracy and better prediction effect in sunny days. This research will help optimize the charging behavior of electric vehicles, thus improving the charging efficiency of electric vehicles.

**Keywords:** electric vehicle; charging behavior; trip chain; influential factors; forecasting model

**Received:** 2018-09-28; **Accepted:** 2019-04-21; **Published online:** 2019-05-08 10:00

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20190506.1423.006.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190506.1423.006.html)

**Foundation item:** National Key R & D Program of China (2018TFB1600702)

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

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

DOI: 10.13700/j.bh.1001-5965.2019.0008

# 基于有限状态机的交会对接飞行任务规划方法



杨胜\*, 王鑫哲, 李蒙

(北京空间技术研制试验中心, 北京 100094)

**摘 要:**为实现多约束条件下载人航天器交会对接(RVD)飞行任务的快速、准确规划,提出了一种基于有限状态机(FSM)的飞行任务规划方法。通过飞行过程与有限状态机的映射关系建立图形化任务规划模型,以飞行事件为状态,以测控、光照等约束条件为输入信号,驱动有限状态机的状态转移,进行模型求解,实现交会对接飞行任务的自动规划。以中国神舟十号飞行任务中航天员手控交会对接试验为例,进行了规划验证。经对比,规划计算结果与任务实施结果一致,表明所提方法可以实现交会对接飞行任务的快速、准确规划。

**关键词:**交会对接(RVD); 任务规划; 有限状态机(FSM); 载人航天器; 图形化建模

中图分类号: V529

文献标识码: A 文章编号: 1001-5965(2019)09-1741-06

神舟八号、神舟九号、神舟十号与天宫一号交会对接(Rendezvous and Docking, RVD)任务的圆满成功,标志着中国突破和掌握了空间交会对接技术,成为世界上第3个自主掌握空间交会对接技术的国家<sup>[1]</sup>。空间交会对接是指追踪飞行器和目标飞行器在同一时间到达轨道上的同一位置并建立刚性连接形成组合体的过程。载人航天器交会对接飞行过程由一系列飞行事件组成,飞行事件的起始时间通常与测控条件、光照条件绑定,飞行事件之间的时间间隔还需要满足航天员准备、地面准备等条件。交会对接飞行任务规划就是以满足飞行时序、测控、光照等多约束条件为目标,合理安排各个飞行事件的执行时间,其是确保交会对接任务顺利实施的重要保证<sup>[2]</sup>。

现有的空间任务规划研究主要集中在对地观测卫星的对地观测任务规划,即根据用户观测需求和卫星约束条件确定卫星的观测动作和数据传输动作,主要思路是将规划问题抽象建立数学模型并精确求解<sup>[3-6]</sup>。载人航天器交会对接过程

中,飞行事件与飞行时序、测控、光照等约束条件强耦合,难以直接采用对地观测任务的规划模型和算法进行交会对接飞行任务规划,采用有限状态机(Finite State Machine, FSM)模型却可以直观地描述这一过程。

本文提出了一种基于有限状态机的交会对接飞行任务规划方法,将交会对接飞行事件作为状态,按照飞行时序建立Mealy型有限状态机模型。将光照条件、测控条件等作为输入信号,驱动有限状态机的状态转移,进行模型求解。以神舟十号与天宫一号载人交会对接任务中航天员手控交会对接为例验证了本文方法的有效性。

## 1 交会对接飞行任务规划

载人航天器交会对接飞行过程通常包括远程交会和近程交会2个阶段。远程交会段通过变轨缩小两飞行器间的相位差、轨道高度差和轨道面偏差<sup>[7]</sup>,近程交会段利用交会对接敏感器测量信息自动控制追踪飞行器逐步接近目标飞行器直至

收稿日期: 2019-01-10; 录用日期: 2019-04-05; 网络出版时间: 2019-05-08 09:25

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20190506.0836.001.html

\*通信作者。E-mail: buaa\_ys@tom.com

引用格式: 杨胜, 王鑫哲, 李蒙. 基于有限状态机的交会对接飞行任务规划方法[J]. 北京航空航天大学学报, 2019, 45(9): 1741-1746. YANG S, WANG X Z, LI M. RVD flight mission planning and scheduling method based on finite state machine [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1741-1746 (in Chinese).

对接<sup>[8]</sup>,手动控制作为自动控制的备份。近程交会段由接近、停泊、最后靠拢、对接等不同的飞行事件组成<sup>[9]</sup>,如图1所示。

载人航天器交会对接过程中,每个飞行事件均应满足一定的约束条件,主要包括轨道日照角、阳光抑制角和测控覆盖条件等<sup>[10-11]</sup>,详见附录A。同时,飞行事件还要满足飞行时序要求。通常飞行事件的顺序是确定的,但关键飞行事件的起始时间 $t$ 与测控条件、光照条件绑定,飞行事件之间的时间间隔 $\Delta t$ 还需要满足航天员操作、地面准备等条件。因此,交会对接飞行事件与飞行时序、测控、光照等约束条件强耦合,如图2所示。交会对接飞行任务规划就是在满足飞行时序、测控、光照等多约束条件下合理安排各个飞行事件的执行时间,确保交会对接任务顺利实施。

交会对接飞行任务规划问题本质上是把飞行事件的起始时间 $(t_1, t_2, \dots, t_n)$ 和时间间隔 $(\Delta t_1, \Delta t_2, \dots, \Delta t_{n-1})$ 作为规划变量的多目标、多变量、非线性优化问题,难以采用抽象的数学模型表达。从图2可以看出,交会对接过程中,当飞行事件的约束条件满足要求时,将按照飞行时序转入下一飞行事件,这一过程非常适合采用有限状态机进行描述。因此,本文提出了一种基于有限状态机的交会对接飞行任务规划方法。

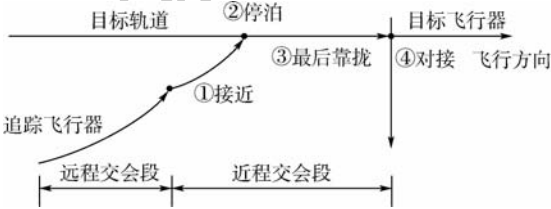


图1 交会对接主要阶段  
Fig.1 Main phases of RVD

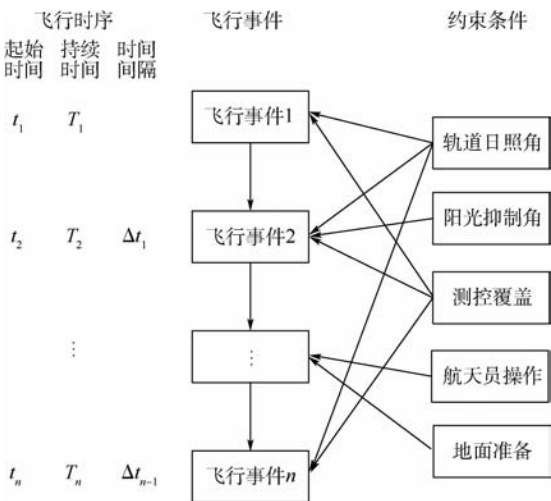


图2 交会对接飞行任务分析  
Fig.2 Analysis of RVD flight mission

## 2 基于有限状态机的规划方法

基于有限状态机的交会对接飞行任务规划主要包括任务分析、规划模型建立、模型求解和规划结果输出等4个部分,如图3所示。首先,对交会对接飞行过程中飞行事件的飞行时序和约束条件进行分析;然后,建立飞行过程与有限状态机的映射关系,构建飞行任务规划模型;其次,结合任务轨道数据,按照设定的起始时间和步长驱动有限状态机进行状态转移,实现模型求解;最后,以文本或图形的形式输出规划结果。

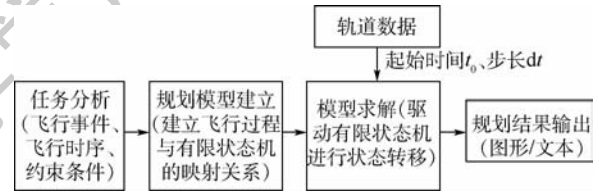


图3 基于有限状态机的交会对接飞行任务规划过程  
Fig.3 RVD flight mission planning and scheduling process based on FSM

### 2.1 有限状态机

有限状态机又称有限状态自动机,是表示有限个状态及其在这些状态之间的转移和动作等行为的数学模型,广泛应用于计算机、通信、数字逻辑设计、软件工程等领域<sup>[12-13]</sup>。在实际应用中,有限状态机根据是否使用输入信号分为 Moore 型和 Mealy 型 2 类。Moore 型有限状态机的输出信号仅与现态有关;Mealy 型有限状态机的输出信号不仅与现态有关,而且与所有输入信号有关,有利于减少模型中状态的数量<sup>[14]</sup>。

Mealy 型有限状态机是一个 6 元组,用  $M$  表示为

$$M = (S, s_0, \Sigma, \Lambda, T, G) \quad (1)$$

式中: $S$  为状态有限集合,  $S = \{s_0, s_1, \dots, s_n\}$ ;  $s_0$  为初始状态;  $\Sigma$  为输入信号的有限集合,  $\Sigma = \{\sigma_0, \sigma_1, \dots, \sigma_n\}$ ;  $\Lambda$  为输出信号的有限集合,  $\Lambda = \{\lambda_0, \lambda_1, \dots, \lambda_n\}$ ;  $T$  为状态转移函数的有限集合;  $G$  为输出函数的有限集合。

除了采用数学模型表示,有限状态机还可以采用图形的形式直观表达,如图4所示。

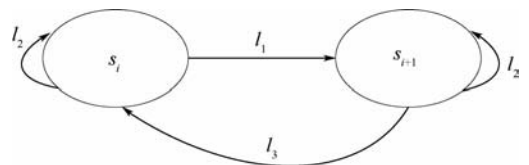


图4 状态转移示意图  
Fig.4 Schematic of state transition

状态转移由转移函数确定,与现态和输入信号有关。状态转移函数可记为

$$\delta(p, l) = q \quad (2)$$

式中: $p, q \in S$ ,表示在状态 $p$ 下,当输入信号 $\Sigma$ 和现态 $p$ 满足条件 $l$ 时,会转移到状态 $q$ ,简记为 $\delta(p, l, q)$ 。

状态转移函数包括现态转移至次态、现态转移至现态和现态转移至初态(或某特定状态)3种情况,图4直观地反映了任务规划的过程。

1)  $\delta_1(s_i, l_1, s_{i+1})$ :当输入信号和飞行时间满足飞行事件的约束条件和飞行时序时(条件 $l_1$ ),由现态转移至次态。

2)  $\delta_2(s_i, l_2, s_i)$ :当输入信号满足飞行事件的约束条件,但飞行时间尚不满足飞行时序要求时(条件 $l_2$ ),由现态转移至现态。

3)  $\delta_3(s_i, l_3, s_{j,j=0,1,\dots,n})$ :当输入信号不满足飞行事件的约束条件时(条件 $l_3$ ),由现态转移至初态(或某特定状态)。

## 2.2 映射关系建立

以交会对接飞行事件作为状态,以轨道日照角、阳光抑制角、测控覆盖条件作为输入信号,建立交会对接飞行过程与有限状态机的映射关系,如表1所示(以Mealy型有限状态机为例)。状态转移函数反映了交会对接飞行时序,输出符号和输出函数反映了某飞行事件(状态)下输出的指令和执行的动作,适用于在轨实时任务规划,本文暂不考虑。

通过建立映射关系可获得基于有限状态机的交会对接飞行任务规划模型 $M$ ,与文献[3-6]中数学建模方法相比,本文方法可以采用图形表达,能够直观地反映交会对接飞行过程。

表1 交会对接与有限状态机的映射关系

Table 1 Mapping relationship between RVD and FSM

模型参数	交会对接飞行过程
$S = \{s_0, s_1, \dots, s_n\}$	由交会对接飞行事件组成的状态集合,其中, $s_0$ 为飞行事件1, $s_1$ 为飞行事件2, $\dots$ , $s_n$ 为飞行事件 $n+1$
$s_0 \in S$	初始状态 $s_0$ (飞行事件1)
$\Sigma = \{\sigma_0, \sigma_1, \dots, \sigma_n\}$	由轨道日照角、阳光抑制角、测控覆盖条件等组成的输入信号集合
$A = \{\lambda_0, \lambda_1, \dots, \lambda_n\}$	输出信号集合
$T$	状态转移函数集合
$G$	输出函数集合

## 2.3 模型求解算法

基于有限状态机的交会对接飞行任务规划模型的求解思路是:按照步长 $dt$ 驱动有限状态机从初始状态 $s_0$ 开始运行,并根据状态转移函数进行

状态转移,当状态转移至 $s_n$ 时,模型求解完成。

具体步骤如下:

**步骤1** 获取任务轨道数据,并设定规划起始时间 $t_0$ 和步长 $dt$ 。

**步骤2** 在初始状态 $s_0$ 下,按照步长 $dt$ 驱动有限状态机开始运行。

1) 根据任务轨道计算轨道日照角 $\beta$ 、阳光抑制角 $\theta$ 、测控覆盖条件 $t_c$ ,作为有限状态机的输入信号。

2) 以现态和输入信号作为状态转移函数的输入,根据状态转移函数进行状态转移。当状态从现态 $s_i$ 转移至次态 $s_{i+1}$ 时,记录转移至 $s_{i+1}$ 的时间 $t_i$ 。返回步骤1)。

**步骤3** 当状态转移至 $s_n$ 时,终止有限状态机运行。

**步骤4** 以文本或图形的形式输出各状态的转入时间(飞行事件起始时间),任务规划完成。

**步骤5** 若在求解过程中出现始终无法从现态 $s_i$ 转移至次态 $s_{i+1}$ 的情况,说明飞行事件 $i$ (对应状态 $s_i$ )的约束条件不满足或者约束条件出现冲突。这种情况下,可以重新设定规划起始时间 $t_0$ ,再次进行规划求得最优解,也可以调整飞行事件 $i$ 的约束条件,从原规划起始时间 $t_0$ 开始,再次进行规划求得次优解。

模型求解可以通过C、C++、MATLAB等语言编制软件自动实现,当飞行任务临时调整或约束条件变化需要重新进行任务规划时,可实现快速、准确规划。

## 3 方法验证

以神舟十号与天宫一号载人交会对接任务中航天员手控交会对接试验为例,进行基于有限状态机的交会对接飞行任务规划方法验证。

### 3.1 任务分析

手控交会对接是指航天员利用飞船上的TV摄像机和操纵手柄,以及目标飞行器上的十字靶标,操纵飞船与目标飞行器精确交会和撤离。神舟十号手控交会对接飞行试验在飞船停靠期间进行,飞船与天宫一号分离,航天员手动控制载人飞船撤离至停泊点;在停泊点完成状态设置和确认后,航天员手动控制飞船接近天宫一号,并完成再次对接。

神舟十号手控交会对接过程飞行事件,按照执行顺序主要包括停靠、分离、撤离、停泊、接近、对接、停靠。经分析,分离和对接应满足轨道日照角约束和测控覆盖约束;撤离和接近应满足轨道

日照角约束、阳光抑制角约束和测控覆盖约束;停泊应满足轨道日照角约束,考虑到飞船撤离至停泊点后需重新进行状态设置和确认,持续时间  $T_3$  应不小于半个轨道周期,如表 2 所示。

表 2 手控交会对接试验飞行任务分析

Table 2 Flight mission analysis of manned RVD experiment

飞行事件	飞行时序		约束条件		
	起始时间	持续时间	轨道日照角	阳光抑制角	测控覆盖
停靠	$t_0$				
分离	$t_1$	$T_1$	$\beta \leq \min\{\beta_1, \beta_2\}$		$t_c \geq T_1$
撤离	$t_2$	$T_2$	$\beta \leq \min\{\beta_1, \beta_2\}$	$\theta \geq \max\{\theta_1, \theta_2, \dots, \theta_n\}$	$t_c \geq T_2$
停泊	$t_3$	$T_3$	$\beta \leq \min\{\beta_1, \beta_2\}$		
接近	$t_4$	$T_4$	$\beta \leq \min\{\beta_1, \beta_2\}$	$\theta \geq \max\{\theta_1, \theta_2, \dots, \theta_n\}$	$t_c \geq T_4$
对接	$t_5$	$T_5$	$\beta \leq \min\{\beta_1, \beta_2\}$		$t_c \geq T_5$
停靠	$t_6$				

### 3.2 模型实现

通过手控交会对接试验过程与有限状态机的映射关系(见表 1),构建飞行任务规划模型。为了直观反映飞行过程,采用图形的形式表示,如图 5 所示。其中,状态转移关系由每个状态的状态转移函数决定。图 5 中: $s_0$  的状态转移函数为  $\delta_1(s_0, l_1, s_1)$ ;  $s_1$  的状态转移函数为  $\{\delta_1(s_1, l_1, s_2), \delta_2(s_1, l_2, s_1), \delta_3(s_1, l_3, s_0)\}$ ;  $s_2$  的状态转移函数为  $\{\delta_1(s_2, l_1, s_3), \delta_2(s_2, l_2, s_2), \delta_3(s_2, l_3, s_0)\}$ ;  $s_3$  的状态转移函数为  $\{\delta_1(s_3, l_1, s_4), \delta_2(s_3, l_2, s_3), \delta_3(s_3, l_2, s_0)\}$ ;  $s_4$  的状态转移函数为  $\{\delta_1(s_4, l_1, s_5), \delta_2(s_4, l_2, s_4), \delta_3(s_4, l_3, s_3)\}$ ;  $s_5$  的状态转移函数为  $\{\delta_1(s_5, l_1, s_0), \delta_2(s_5, l_2, s_5), \delta_3(s_5, l_3, s_3)\}$ 。

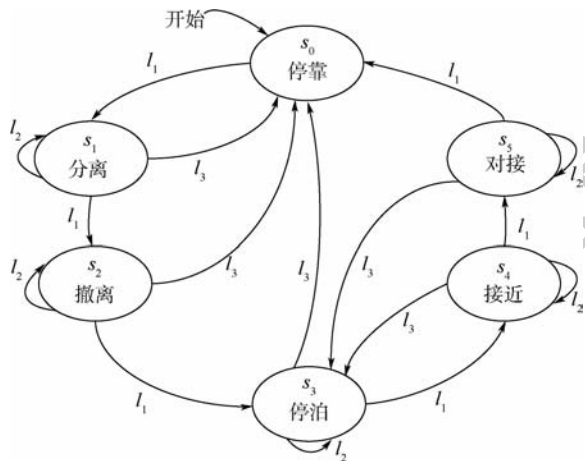


图 5 手控交会对接任务规划模型

Fig. 5 Mission planning and scheduling model of manned RVD

### 3.3 模型求解及规划结果对比

为实现模型自动求解,使用 MATLAB 实现本

文的规划方法,形成了交会对接飞行任务自动规划系统。设置规划的起始时间  $t_0$  为 2013 年 6 月 23 日 08:00:00(北京时),规划步长  $dt$  为 1 s,规划时间为 10 800 s。软件运行后,根据任务轨道按照步长计算轨道日照角、阳光抑制角和测控覆盖条件作为输入信号,驱动有限状态机从初始状态  $s_0$ “停靠”开始运行。当有限状态机按顺序再次运行至初始状态  $s_0$  时,任务规划结束。手控交会对接试验任务规划结果采用甘特图的形式表示,如图 6 所示。关键飞行事件的起始时间如表 3 所示。可以看出,本文的规划结果与神舟十号任务实际在轨执行结果<sup>[15]</sup>一致,验证了本文方法的有效性和正确性。

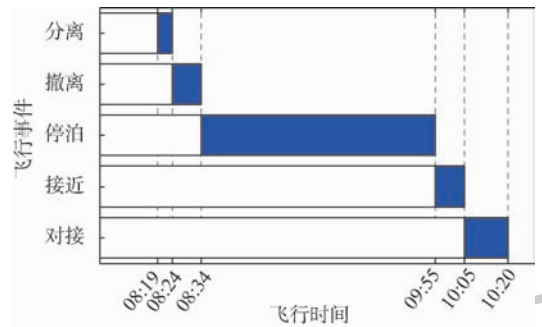


图 6 手控交会对接试验飞行任务规划结果

Fig. 6 Flight mission planning and scheduling results of manned RVD experiment

表 3 手控交会对接试验飞行任务规划结果对比

Table 3 Comparison of flight mission planning and scheduling results of manned RVD experiment

飞行事件	起始时间	
	本文规划结果	神舟十号在轨执行结果
撤离	08:24	08:26
接近	10:05	10:07

## 4 结论

本文提出了一种基于有限状态机的载人航天器交会对接飞行任务规划方法,主要结论如下:

1) 在分析载人航天器交会对接飞行任务规划的多约束、多变量、非线性特性的基础上,利用飞行过程与有限状态机的映射关系,建立了直观的、图形化的任务规划模型。

2) 利用 MATLAB 搭建了模型求解系统,通过驱动有限状态机的状态转移,进行自动求解,实现了交会对接飞行任务的快速、准确规划。

3) 以神舟十号与天宫一号载人交会对接任务中航天员手控交会对接试验为例,验证了本文方法的正确性和有效性。

本文方法可推广应用于其他航天器的飞行任务规划。后续通过在规划模型中建立飞行过程与“输出符号”和“输出函数”的映射关系,还可以实现飞行任务的实时规划和自动执行,进一步提高航天器的自主程度。

### 参考文献 (References)

- [1] MA X B, HUANG Z, YANG Q. Achievements and expectation of China's rendezvous and docking task [C] // Proceedings of the 64th International Astronautical Congress, 2013:1-7.
- [2] XING J J, LI J, ZHU H. Optimizing in advance: Shenzhou 8 space mission scheduling tryout architecture [C] // The 5th International Conference on Computer Science & Education. Piscataway, NJ: IEEE Press, 2010:152-155.
- [3] WOLE W J, SORENSEN S E. Three scheduling algorithms of the earth observing systems domain [J]. Management Science, 2000, 46(1):148-168.
- [4] GLOBUS A, CRAWFORD J, LOHN J, et al. A comparison of techniques for scheduling earth observing satellites [C] // The 16th Conference on Innovative Applications of Artificial Intelligence. Palo Alto: AAAI, 2004:836-843.
- [5] 冉承新, 王慧林, 熊纲要. 基于改进遗传算法的移动目标成像侦测任务规划问题研究 [J]. 宇航学报, 2010, 31(2):458-465.
- RAN C X, WANG H L, XIONG G Y. Research on mission-planning of ocean moving targets imaging reconnaissance based on improved genetic algorithm [J]. Journal of Astronautics, 2010, 31(2):458-465 (in Chinese).
- [6] 简平, 邹鹏, 熊伟. 基于 DPSO-SA 的低轨预警系统初始任务规划方法 [J]. 北京航空航天大学学报, 2013, 39(10):1381-1386.
- JIAN P, ZOU P, XIONG W. Original task planning method of early warning system of LEO based on DPSO-SA [J]. Journal of Beijing University of Aeronautics and Astronautics, 2013, 39(10):1381-1386 (in Chinese).
- [7] 张丽艳, 戚发劼, 李颐黎. 交会对接远距离导引精度分析 [J]. 北京航空航天大学学报, 2006, 32(6):667-670.
- ZHANG L Y, QI F R, LI Y L. Precision analysis of long distance navigation of rendezvous [J]. Journal of Beijing University of Aeronautics and Astronautics, 2006, 32(6):667-670 (in Chinese).
- [8] 朱仁璋, 林彦. 航天器交会最终逼近段相对姿态估计与控制 [J]. 北京航空航天大学学报, 2007, 33(5):544-548.
- ZHU R Z, LIN Y. Relative attitude estimation and control schemes for the final approach phase of spacecraft rendezvous [J]. Journal of Beijing University of Aeronautics and Astronautics, 2007, 33(5):544-548 (in Chinese).
- [9] WIGHERT F. Automated rendezvous and docking of spacecraft [M]. Cambridge: Cambridge University Press, 2003:267-282.
- [10] 李革非, 陈莉丹, 唐歌实, 等. 多约束交会对接发射窗口的分析和规划 [J]. 宇航学报, 2011, 32(11):2463-2470.
- LI G F, CHEN L D, TANG G S, et al. Analysis and programming of rendezvous launch window with multi-constraints [J]. Journal of Astronautics, 2011, 32(11):2463-2470 (in Chinese).
- [11] 周建平. 载人航天交会对接技术 [J]. 载人航天, 2011, 17(2):1-8.
- ZHOU J P. Rendezvous and docking technology of manned space flight [J]. Manned Spaceflight, 2011, 17(2):1-8 (in Chinese).
- [12] 王蓓, 赵廷弟, 焦健. 基于有限状态机的安全性仿真技术 [J]. 北京航空航天大学学报, 2011, 37(4):428-432.
- WANG B, ZHAO T D, JIAO J. Safety simulation technology based on finite state machine [J]. Journal of Beijing University of Aeronautics and Astronautics, 2011, 37(4):428-432 (in Chinese).
- [13] 王蓓, 赵廷弟. 应用 Stateflow 技术的安全性建模与仿真 [J]. 北京航空航天大学学报, 2011, 37(11):1415-1420.
- WANG B, ZHAO T D. Safety simulation applying Stateflow technology [J]. Journal of Beijing University of Aeronautics and Astronautics, 2011, 37(11):1415-1420 (in Chinese).
- [14] ANDERSON J A. Automata theory with modern applications [M]. Cambridge: Cambridge University Press, 2006:105-108.
- [15] 风云. 神舟十号任务取得圆满成功 [J]. 卫星应用, 2013(4):17.
- FENG Y. Shenzhou-10 mission achieve successful accomplishment [J]. Satellite Application, 2013(4):17 (in Chinese).

#### 作者简介:

杨胜 男, 博士, 高级工程师。主要研究方向: 载人航天器总体设计、飞行任务规划。

### 附录 A 交会对接过程约束条件分析

1) 轨道日照角约束。轨道日照角要求与航天器自身能源、热控等分系统的约束要求相关, 可综合为对航天器在一段飞行时间内太阳矢量与轨道平面夹角的要求。两航天器对接前后对轨道日照角要求较高<sup>[10]</sup>。假设追踪航天器轨道日照角要求为  $\beta_1$ , 目标航天器轨道日照角要求为  $\beta_2$ , 则交会对接过程中飞行事件均应满足轨道日照角约束  $\beta \leq \min\{\beta_1, \beta_2\}$ 。

2) 阳光抑制角约束。交会对接过程中, 追踪航天器用于进行相对状态测量的光学设备主要有激光雷达和 CCD 相机<sup>[1]</sup>, 手控交会还需要使用 TV 摄像机。为了确保相对导航光学设备正常工作, 要求太阳光不能直射光学测量设备的视场, 即太阳光不能进入以其视线方向为中心线、一定锥角的圆锥内。这种相对导航光学测量设备对太阳光照的要求也称为阳光抑制角要求。假设追踪航天器光学交会对接敏感器  $1, 2, \dots, n$  的阳光抑制角要求分别为  $\theta_1, \theta_2, \dots, \theta_n$ , 则接近、撤离过程中阳光抑制角约束为  $\theta \geq \max\{\theta_1, \theta_2, \dots, \theta_n\}$ 。

3) 测控覆盖约束。追踪航天器转入自主控

制后,要求地面飞控中心对对接、分离等交会对接关键飞行事件进行监视,并在必要时进行干预。假设关键飞行事件的持续时间分别为 $T_1, T_2, \dots$ ,

则交会对接过程中测控覆盖约束为: $\{t_c \geq T_1, t_c \geq T_2, \dots\}$ 。目前,利用陆海天基测控网可以满足交会对接测控覆盖和可靠性要求<sup>[11]</sup>。

## RVD flight mission planning and scheduling method based on finite state machine

YANG Sheng\*, WANG Xinzhe, LI Meng

(Beijing Space Technology Research and Test Center, Beijing 100094, China)

**Abstract:** The flight mission planning and scheduling method based on finite state machine (FSM) was proposed for planning and scheduling the rendezvous and docking (RVD) flight mission of manned spacecraft rapidly and exactly under multi-constraints. The graphical mission planning and scheduling model was established by mapping the RVD process with FSM. The flight events were taken as states and the constraints, such as sunlight condition, telemetry tracking and command coverage, were taken as input signals of the FSM. The model was solved automatically by driving the state machine transit from the initial state to the last one according to the flight sequence. Take the manned RVD experiment in China's Shenzhou-10 mission as an example, the proposed method was verified. The mission planning and scheduling result was consistent with the on-orbit execution. It shows that the FSM based modeling method can realize the RVD flight mission planning and scheduling rapidly and exactly.

**Keywords:** rendezvous and docking (RVD); mission planning and scheduling; finite state machine (FSM); manned spacecraft; graphical modeling

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

DOI: 10.13700/j.bh.1001-5965.2018.0757

# 不确定容量下时隙分配问题两阶段规划模型

亓尧<sup>1</sup>, 王瑛<sup>1,\*</sup>, 梁颖<sup>2</sup>, 姚颀<sup>3</sup>

(1. 空军工程大学 装备管理与无人机工程学院, 西安 710043; 2. 空军研究院, 北京 100089;

3. 国家空域管理中心, 北京 100091)



**摘 要:** 恶劣天气等不确定环境下,传统时隙分配方法易造成航班大量延误现象,为解决这一问题,分析了时隙分配过程,基于不确定理论,从权衡“请求时隙-计划时隙差”和“计划时隙-运行时隙差”的角度,提出了不确定容量下的时隙分配两阶段规划模型,分别构建了单机场模型和多机场模型。根据模型特点,设计了基于人工蜂群(ABC)算法的渐进二元启发式方法,提升了求解效率。通过算例分析,验证了所提模型和方法的有效性,同时对模型参数设置进行了分析。

**关键词:** 时隙分配; 两阶段规划; 不确定理论; 人工蜂群(ABC)算法; 启发式计算

**中图分类号:** V355

**文献标识码:** A

**文章编号:** 1001-5965(2019)09-1747-10

近年来,中国航空运输业实现蓬勃发展。根据中国民用航空局的统计数据<sup>[1]</sup>，“十二五”期间,中国航空运输总周转量增长 58.2%，旅客运输量增长 62.7%，货邮运输量增长 11.8%。2017年,中国航空运输总周转量达到 1 083.1 亿吨公里,旅客运输量 5.52 亿人次,货邮运输量 705.8 万吨。《中国民用航空发展第十三个五年规划》<sup>[2]</sup>提出,截止到 2020 年,中国航空运输总周转量将达到 1 420 亿吨公里,相较于“十二五”末的 2015 年,年均增长 10.8%;旅客运输量将达到 7.2 亿人次,年均增长 10.4%;货邮运输量将达到 850 万吨,年均增长 6.2%。急剧增长的空中交通流量使得空中交通网络日趋拥挤,尤其是在恶劣天气等不确定条件下,航班延误现象十分严重。据统计,2017 年,中国航空公司的平均航班正常率为 71.67%,在全部航班不正常的原因中,天气原因所占比例高达 51.28%<sup>[3]</sup>。

时隙分配(Time Slot Allocation, TSA)是进行

空中交通流量管理、实现空中交通容流匹配、解决航班延误的关键技术。时隙分配过程包括:航空公司提出时隙请求,流量管理部门按照一定的方式或依靠某种机制,将时隙分配给航班或航空公司,形成航班时隙计划,短期根据环境和条件变化进行时隙二次分配,调整航班计划,获得最终航班运行时刻。

时隙分配作为空中交通管理的重要手段,一直是国内外空管领域研究的热点。时隙分配方法方面,比较经典的有先到先服务(First Come First Serve, FCFS)<sup>[4]</sup>、按时刻分配(Ration By Schedule, RBS)<sup>[5]</sup>、压缩分配(compression)<sup>[5]</sup>、时隙信用交换(Slot Credit Substitutions, SCS)<sup>[6]</sup>、累积延误分配(Accrue Delay Based, ADB)<sup>[7]</sup>等。Madas 和 Zografos<sup>[8]</sup>对各时隙分配方法进行了对比分析,指出了各方法的适用性。时隙分配模型方面,1987 年,Odoni<sup>[9]</sup>首次提出通过航班时隙调整,将空中拥挤转化为地面等待问题(Ground Holding

收稿日期: 2018-12-25; 录用日期: 2019-02-02; 网络出版时间: 2019-02-26 09:26

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20190225.0927.002.html

基金项目: 国家自然科学基金(71601183)

\*通信作者. E-mail: yingwangkd@163.com

**引用格式:** 亓尧, 王瑛, 梁颖, 等. 不确定容量下时隙分配问题两阶段规划模型[J]. 北京航空航天大学学报, 2019, 45(9): 1747-1756. QI Y, WANG Y, LIANG Y, et al. Two-stage programming model for time slot allocation problem under uncertain capacity [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1747-1756 (in Chinese).

Problem, GHP), 开启了时隙分配问题研究的先河。Terrab 等<sup>[10]</sup>提出了网络流模型,并运用最小费用流理论来求解模型。Gilbo<sup>[11]</sup>将研究对象从单航班扩展到全部进离场航班,建立了多元受限下的时隙分配模型。Oussedik、Delahaye<sup>[12]</sup>和Pulugurtha、Nambisan<sup>[13]</sup>将智能算法引入时隙分配模型求解,提升了模型的求解速度,促进了模型在实际中的推广和应用。Madas 和 Zografos<sup>[14]</sup>对各时隙分配模型进行了对比分析。国内对时隙分配模型研究起步较晚,在时隙分配模型、时隙分配方法等方面比较有代表性的研究包括文献<sup>[15-20]</sup>。

容量等易受恶劣天气等不确定条件的影响,呈现不确定性特点。针对此现象,相关模型也得到了研究。1987年,Andreatta 和 Romanin-Jacur<sup>[21]</sup>第一次提出了单时间区间静态随机模型,并提出了动态规划方法进行模型求解。随后,随机型时隙分配模型及求解方法也得到了国内外学者的广泛研究<sup>[4,10,21-25]</sup>。随机型模型要求在大量历史数据基础上获得机场容量的概率分布,然而并非所有情况下均能够获取足量数据,此时需要有新的规划方法,解决不确定容量下的时隙分配问题。

从对时隙分配过程的描述可以看出,时隙分配要经历先后2次分配过程。第1次时隙分配输出的航班时隙计划是开展第2次时隙分配的输入。现有模型的问题在于:将2次时隙分配独立研究,第1次时隙分配采用名义容量约束,即给定时间段内能够服务的最大航空器数量;第2次时隙分配采用实际容量约束,在不确定条件影响下,容量名义值与实际值差距较大,造成2次分配结果差距过大,导致延误大面积产生,分配结果缺乏灵活性。为此,Corolli 等<sup>[26]</sup>首次将2次时隙分配纳入统一模型,针对不确定容量下的时隙分配问题,对“请求时隙-计划时隙差”和“计划时隙-运行时隙差”进行权衡,获得最优的时隙分配结果,但研究仅限于单机场模型,以航班数作为决策变量,限制了模型的实用性。岳仁田等<sup>[27]</sup>提出了带补偿的两阶段随机规划航班时隙分配模型,但同样仅限于单机场模型。

本文以不确定容量下的时隙分配问题为研究对象,针对时隙分配问题的现实需求和已有模型的不足,基于不确定理论<sup>[28-29]</sup>,以航班而非航班数为决策变量,提出了不确定容量下时隙分配两阶段规划模型,从单机场时隙分配模型扩展到多机场时隙分配模型。为了提升求解效率,本文基于人工蜂群(Artificial Bee Colony, ABC)算法设计

了一种渐进二元启发式方法,并验证了模型和方法的有效性,分析了模型参数设置。

## 1 问题描述与建模

### 1.1 问题定义

时隙分配过程可以分为2个阶段,如图1所示。第1个阶段,流量管理部门根据航空公司的时隙请求,进行初次分配,得到各航空公司的航班时隙计划;第2个阶段,在实际运行中,流量管理部门(或航空公司)根据短期(一般指24h以内)的实际情况(如天气变化、突发事件等)对航班时隙计划进行实时调控,完成时隙的再次分配,获得最终的航班运行时隙表。

根据时隙分配的两阶段过程,将机场容量作为唯一受限元,不涉及扇区、航路点容量,构建不确定容量下单机场时隙分配两阶段规划模型和多机场时隙分配两阶段规划模型。建模的基本思路为:第1个阶段,在已知航班时隙请求的条件下,将第2个阶段才能获得的容量实现值设置为不确定变量,取代传统模型中的名义容量,以最小化“请求时隙-计划时隙差”与“计划时隙-运行时隙差(即延误)”的加权和为目标函数,以不确定容量等条件为约束,构建第1阶段模型,求解航班时隙计划;第2个阶段,以最小化“计划时隙-运行时隙差”为目标,以容量实现值条件为约束,求解航班运行时隙。

为构建模型做以下几点基本假设:①航班实际运行时隙不早于计划时隙;②航班的飞行时间是固定的;③初次时隙分配时,机场容量未知,且难以获取足够样本以获取其概率分布。

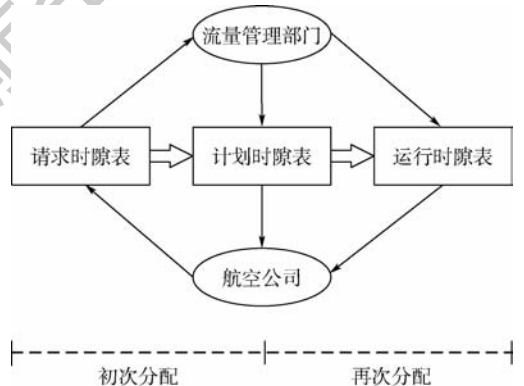


图1 时隙分配的2个阶段

Fig. 1 Two stages of time slot allocation

### 1.2 单机场模型

单机场时隙分配问题要求保证在该机场的每个时隙内,进场航班和离场航班在容量等约束范围内,否则要对一些航班的进离场时隙进行调整

以适应容量约束要求。

定义模型符号如下: $T$ 为离散化时隙集,其中时隙 $t \in \{0, 1, \dots, T-1\}$ ;  $F$ 为航班集,航班 $f \in F$ ;  $F_A$ 为进场航班集,进场航班 $f_A \in F_A$ ;  $F_D$ 为离场航班集,离场航班 $f_D \in F_D$ ;  $P$ 为联程航班集,联程航班对 $(f, f') \in P$ ;  $l_{ff'}$ 为联程航班最小周转时间;  $W$ 为权重向量;  $t_f^a$ 为航班 $f$ 的进场请求时隙;  $t_f^d$ 为航班 $f$ 的离场请求时隙;  $\alpha^i(\xi)$ 、 $\beta^i(\xi)$ 、 $\gamma^i(\xi)$ 分别为第 1 个阶段机场在时隙  $t$  的进场容量、离场容量、总容量,为不确定变量;  $\alpha^j(\xi)$ 、 $\beta^j(\xi)$ 、 $\gamma^j(\xi)$ 分别为第 2 个阶段机场在时隙  $t$  的进场容量、离场容量、总容量实现值。

构建单机场时隙分配第 1 阶段模型如下:

第 1 阶段决策变量定义为

$$x_f^i = \begin{cases} 1 & \text{航班 } f \text{ 被分配到时隙 } t \text{ 进场, } f \in F_A \\ 0 & \text{其他} \end{cases}$$

$$y_f^j = \begin{cases} 1 & \text{航班 } f \text{ 被分配到时隙 } t \text{ 离场, } f \in F_D \\ 0 & \text{其他} \end{cases}$$

第 1 阶段目标函数定义为

$$\min \left( \sum_{f \in F_A} \sum_{t \in T} (t - t_f^a) x_f^i + \sum_{f \in F_D} \sum_{t \in T} (t - t_f^d) y_f^j \right) + W \cdot E[f(x, y, \xi)] \quad (1)$$

第 1 阶段约束函数定义为

$$\sum_{t \in T} x_f^i = 1 \quad \forall f \in F_A \quad (2)$$

$$\sum_{t \in T} y_f^j = 1 \quad \forall f \in F_D \quad (3)$$

$$\sum_{f \in F_A} x_f^i \leq \alpha^i(\xi) \quad \forall t \in T \quad (4)$$

$$\sum_{f \in F_D} y_f^j \leq \beta^j(\xi) \quad \forall t \in T \quad (5)$$

$$\sum_{f \in F_A \cup F_D} (x_f^i + y_f^j) \leq \gamma^i(\xi) \quad \forall t \in T \quad (6)$$

$$\sum_{t \in [0, k]} y_f^j + \sum_{t \in [k - l_{ff'}, |T|]} x_{f'}^i \leq 1 \quad \forall (f, f') \in P, k \in [l_{ff'}, |T|] \quad (7)$$

$$x_f^i, y_f^j \in \{0, 1\} \quad \forall f \in F_A \cup F_D, t \in T \quad (8)$$

式(1)定义了第 1 阶段的目标函数,即最小化“请求时隙-计划时隙差”与“计划时隙-运行时隙差(即延误)”的加权和,前半部表示航班请求时隙与航班计划时隙的差,后半部表示航班延误(航班计划时隙与航班运行时隙的差)的期望值,  $\xi$  为定义在不确定空间  $(\Gamma, \mathcal{L}, \mathcal{M})$  上的不确定变量,  $f(x, y, \xi)$  为补偿函数。式(2)~式(8)定义了第 1 阶段模型的约束函数:式(2)和式(3)表示航班  $f$  的进场和离场能且仅能分配到 1 个时隙,式(4)~式(6)保证了在任意时隙中机场进场容量、离场容量和总容量限制总能满足,式(7)保证了联程航班周转时间的充足性,式(8)表示决策变量为 0-1 整数变量。

单机场时隙分配第 2 阶段模型如下:

第 2 阶段决策变量定义为

$$u_f^i(\xi) = \begin{cases} 1 & \text{航班 } f \text{ 在时隙 } t \text{ 进场, } f \in F_A \\ 0 & \text{其他} \end{cases}$$

$$v_f^j(\xi) = \begin{cases} 1 & \text{航班 } f \text{ 在时隙 } t \text{ 离场, } f \in F_D \\ 0 & \text{其他} \end{cases}$$

给定第 1 阶段决策变量  $(x_f^i, y_f^j)$  和不确定容量  $\alpha^i(\xi)$ 、 $\beta^j(\xi)$ 、 $\gamma^i(\xi)$  的实现值  $\alpha^i(\xi)$ 、 $\beta^j(\xi)$ 、 $\gamma^i(\xi)$ , 一定数量的航班不得不采取延误措施以平衡流量和容量。

第 2 阶段目标函数定义为

$$f(x, y, \xi) = \min \sum_{f \in F_A} \left( \sum_{t \in T} u_f^i(\xi) \cdot t - \sum_{t \in T} x_f^i \cdot t \right) + \sum_{f \in F_D} \left( \sum_{t \in T} v_f^j(\xi) \cdot t - \sum_{t \in T} y_f^j \cdot t \right) \quad (9)$$

第 2 阶段约束函数定义为

$$\sum_{t \in T} u_f^i(\xi) = 1 \quad \forall f \in F_A \quad (10)$$

$$\sum_{t \in T} v_f^j(\xi) = 1 \quad \forall f \in F_D \quad (11)$$

$$\sum_{f \in F_A} u_f^i(\xi) \leq \alpha^i(\xi) \quad \forall t \in T \quad (12)$$

$$\sum_{f \in F_D} v_f^j(\xi) \leq \beta^j(\xi) \quad \forall t \in T \quad (13)$$

$$\sum_{f \in F_A} u_f^i(\xi) + \sum_{f \in F_D} v_f^j(\xi) \leq \gamma^i(\xi) \quad \forall t \in T \quad (14)$$

$$\sum_{t \in [0, k]} u_f^i(\xi) + \sum_{t \in [k - l_{ff'}, |T|]} v_{f'}^j(\xi) \leq 1 \quad \forall (f, f') \in P, k \in [l_{ff'}, |T|] \quad (15)$$

$$u_f^i(\xi), v_f^j(\xi) \in \{0, 1\} \quad \forall f \in F_A \cup F_D, t \in T \quad (16)$$

式(9)定义了第 2 阶段的目标函数,即最小化“计划时隙-运行时隙差”,也就是最小化航班的总延误时间。式(10)~式(16)定义了第 2 阶段模型的约束函数:式(10)和式(11)表示航班  $f$  的进场和离场能且仅能分配到 1 个时隙,式(12)~式(14)保证了在任意时隙中机场进场容量、离场容量和总容量限制总能满足,式(15)保证了联程航班周转时间的充足性,式(16)表示决策变量为 0-1 整数变量。

从单机场两阶段规划模型可以看出,第 1 阶段模型和第 2 阶段模型的目标函数不同,第 1 阶段模型在不确定容量条件下,最小化“请求时隙-计划时隙差”与“计划时隙-运行时隙差(即延误成本)”加权和,而第 2 阶段模型是在容量确定条件下,最小化“计划时隙-运行时隙差(即延误成本)”。第 1 阶段容量约束与第 2 阶段容量约束不同,第 1 阶段模型中容量为不确定变量,第 2 阶段模型中容量为确定值,其他约束函数相同。

### 1.3 多机场模型

单机场模型实质上是假设各机场间的时隙分

配相互独立、互不影响。但实际上,在不更改航路和飞行速度的情况下,航班在起飞机场离场时隙的调整,肯定会导致航班在降落机场进场时隙的变化,加之机场容量的限制,时隙调整可能产生连锁反应,导致整个机场网络时隙的变化。此外,由于联程航班的存在,前序航班的时隙调整就会影响后续航班,造成时隙调整在机场网络上的传播。因此,有必要研究不确定容量下多机场时隙分配两阶段规划模型。

考虑相互连通的机场网络,构建不确定容量下多机场时隙分配两阶段规划模型。

定义模型符号如下: $I$ 为机场集,机场 $i \in I$ ;  $F_A^i$ 为机场 $i$ 的进场航班集,进场航班 $f \in F_A^i$ ;  $F_D^i$ 为机场 $i$ 的离场航班集,离场航班 $f \in F_D^i$ ;  $\text{etf}_f$ 为航班 $f$ 的飞行时间;  $\alpha_i^t(\xi)$ 、 $\beta_i^t(\xi)$ 、 $\gamma_i^t(\xi)$ 分别为第1阶段机场 $i$ 在时隙 $t$ 的进场容量、离场容量、总容量,为不确定变量;  $\alpha_i^t(\xi)$ 、 $\beta_i^t(\xi)$ 、 $\gamma_i^t(\xi)$ 分别为第2阶段机场 $i$ 在时隙 $t$ 的进场容量、离场容量、总容量实现值。

构建多机场时隙分配第1阶段模型如下:

第1阶段决策变量定义为

$$x_f^t = \begin{cases} 1 & \text{航班 } f \text{ 被分配到时隙 } t \text{ 进场, } f \in F_A^i \\ 0 & \text{其他} \end{cases}$$

$$y_f^t = \begin{cases} 1 & \text{航班 } f \text{ 被分配到时隙 } t \text{ 离场, } f \in F_D^i \\ 0 & \text{其他} \end{cases}$$

第1阶段目标函数定义为

$$\min \sum_{i \in I} \left( \sum_{f \in F_A^i} \sum_{t \in T} (t - t_f^a) x_f^t + \sum_{f \in F_D^i} \sum_{t \in T} (t - t_f^d) y_f^t \right) + W \cdot E[f(x, y, \xi)] \quad (17)$$

第1阶段约束函数定义为

$$\sum_{i \in T} x_f^t = 1 \quad \forall f \in F_A^i \quad (18)$$

$$\sum_{i \in T} y_f^t = 1 \quad \forall f \in F_D^i \quad (19)$$

$$\sum_{f \in F_A^i} x_f^t \leq \alpha_i^t(\xi) \quad \forall i \in I, t \in T \quad (20)$$

$$\sum_{f \in F_D^i} y_f^t \leq \beta_i^t(\xi) \quad \forall i \in I, t \in T \quad (21)$$

$$\sum_{f \in F_A^i \cup F_D^i} (x_f^t + y_f^t) \leq \gamma_i^t(\xi) \quad \forall i \in I, t \in T \quad (22)$$

$$x_f^{t+\text{etf}_f} - y_f^t = 0 \quad \forall f \in F, t \in T \quad (23)$$

$$\sum_{i \in [0, k]} y_f^t + \sum_{i \in [k - \text{etf}_f, |T|)} x_f^t \leq 1 \quad \forall (f, f') \in P, k \in [L_{ff'}, |T|) \quad (24)$$

$$x_f^t, y_f^t \in \{0, 1\} \quad \forall f \in F_A^i \cup F_D^i, t \in T, i \in I \quad (25)$$

式(17)定义了第1阶段的目标函数,即最小化“请求时隙-计划时隙差”与“计划时隙-运行时隙差(即延误)”的加权和,前半部表示航班请求

时隙与航班计划时隙的差,后半部表示航班延误(航班计划时隙与航班运行时隙的差)的期望值。式(18)~式(25)定义了第1阶段模型的约束函数:式(18)~式(19)表示航班 $f$ 的进场和离场能且仅能分配到一个时隙,式(20)~式(22)保证了在任意时隙中机场进场容量、离场容量和总容量限制总能满足,式(23)表示所有航班飞行时间都是固定的,式(24)保证了联程航班周转时间的充足性,式(25)表示决策变量为0-1整数变量。

多机场时隙分配第2阶段模型如下:

第2阶段决策变量定义为

$$u_f^t(\xi) = \begin{cases} 1 & \text{航班 } f \text{ 在时隙 } t \text{ 进场, } f \in F_A^i \\ 0 & \text{其他} \end{cases}$$

$$v_f^t(\xi) = \begin{cases} 1 & \text{航班 } f \text{ 在时隙 } t \text{ 离场, } f \in F_D^i \\ 0 & \text{其他} \end{cases}$$

给定第1阶段决策变量 $(x_f^t, y_f^t)$ 和不确定容量 $\alpha_i^t(\xi)$ 、 $\beta_i^t(\xi)$ 、 $\gamma_i^t(\xi)$ 的实现值 $\alpha_i^t(\xi)$ 、 $\beta_i^t(\xi)$ 、 $\gamma_i^t(\xi)$ ,一定数量的航班不得不采取延误措施以平衡流量和容量。

第2阶段目标函数定义为

$$f(x, y, \xi) = \min \sum_{i \in I} \left( \sum_{f \in F_A^i} \left( \sum_{t \in T} u_f^t(\xi) \cdot t - \sum_{i \in T} x_f^t \cdot t \right) + \sum_{f \in F_D^i} \left( \sum_{t \in T} v_f^t(\xi) \cdot t - \sum_{i \in T} y_f^t \cdot t \right) \right) \quad (26)$$

第2阶段约束函数定义为

$$\sum_{i \in T} u_f^t(\xi) = 1 \quad \forall f \in F_A^i \quad (27)$$

$$\sum_{i \in T} v_f^t(\xi) = 1 \quad \forall f \in F_D^i \quad (28)$$

$$\sum_{f \in F_A^i} u_f^t(\xi) \leq \alpha_i^t(\xi) \quad \forall i \in I, t \in T \quad (29)$$

$$\sum_{f \in F_D^i} v_f^t(\xi) \leq \beta_i^t(\xi) \quad \forall i \in I, t \in T \quad (30)$$

$$\sum_{f \in F_A^i \cup F_D^i} (u_f^t(\xi) + v_f^t(\xi)) \leq \gamma_i^t(\xi) \quad \forall i \in I, t \in T \quad (31)$$

$$u_f^{t+\text{etf}_f}(\xi) - v_f^t(\xi) = 0 \quad \forall f \in F, t \in T \quad (32)$$

$$\sum_{i \in [0, k]} v_f^t(\xi) + \sum_{i \in [k - \text{etf}_f, |T|)} u_{f'}^t(\xi) \leq 1 \quad \forall (f, f') \in P, k \in [L_{ff'}, |T|) \quad (33)$$

$$u_f^t(\xi), v_f^t(\xi) \in \{0, 1\} \quad \forall f \in F_A^i \cup F_D^i, t \in T, i \in I \quad (34)$$

式(26)定义了第2阶段的目标函数,即最小化“计划时隙-运行时隙差”,也就是最小化航班的总延误成本。式(27)~式(34)定义了第2阶段模型的约束函数:式(27)和式(28)表示航班 $f$ 的进场和离场能且仅能分配到一个时隙,式(29)~式(31)保证了在任意时隙中机场进场容量、离场容量和总容量限制总能满足,式(32)表示所有航班飞行时间都是固定的,式(33)保证了联程航班

周转时间的充足性,式(34)表示决策变量为0-1整数变量。

## 2 方法设计

针对模型规模大、二进制的特点,本文设计了基于ABC算法的渐进二元启发式方法对模型进行求解。

### 2.1 渐进二元启发式方法框架

渐进二元启发式方法主要针对模型二进制特点,其主要思想是:首先求解原问题的松弛问题,然后判断解的二元性,对不符合二进制约束的解渐进地增加二进制约束,最终求得符合模型约束的二进制最优解。渐进式二元启发式计算过程如下:

**步骤1** 读取原问题的松弛问题,并初始化迭代次数  $i = 1$ 。

**步骤2** 求解松弛问题,得到两阶段模型最优解  $(x^1, y^1)$  和  $(u^1, v^1)$ ,同时得到对应的第1阶段目标值  $V'_{RP}$  和第2阶段目标值  $V''_{RP}$ 。

**步骤3** 设置模型下界  $LB = V'_{RP} + V''_{RP}$ 。

**步骤4** 检查所有子问题是否均有二进制解,若结果为是,则转入步骤8,否则转入步骤5。

**步骤5** 对存在分数解的子问题增加二进制约束。

**步骤6** 更新迭代次数  $i = i + 1$ 。

**步骤7** 求解新问题,并更新最优解  $(x^i, y^i)$ 、 $(u^i, v^i)$  和目标函数值  $V'_{RP}$ 、 $V''_{RP}$ ,并转入步骤4。

**步骤8** 输出  $(x^i, y^i)$  和  $(u^i, v^i)$ ,判断  $V'_{RP} + V''_{RP} = LB$  是否成立,若是,则  $(x^i, y^i)$  和  $(u^i, v^i)$  为模型最优解,否则,  $(x^i, y^i)$  和  $(u^i, v^i)$  为模型  $\varepsilon$ -最优解。

### 2.2 基于ABC算法的求解器设计

2.1节提出了求解两阶段0-1整数规划模型的渐进二元启发式方法,从方法框架中可以看出,步骤2和步骤7都要进行问题求解,考虑到时隙分配问题特别是多机场时隙分配问题规模大的特点,本文将ABC算法引入到渐进二元启发式方法框架,作为求解器用于步骤2和步骤7的求解。

ABC算法是2005年Karaboga<sup>[30]</sup>通过模拟工蜂采蜜的整个过程提出的一种新型智能优化算法,与其他智能优化算法相比,其具有结构简单、易于实现、参数较少、鲁棒性强等优点<sup>[31]</sup>。

基于ABC算法的渐进二元启发式方法流程图如图2所示。

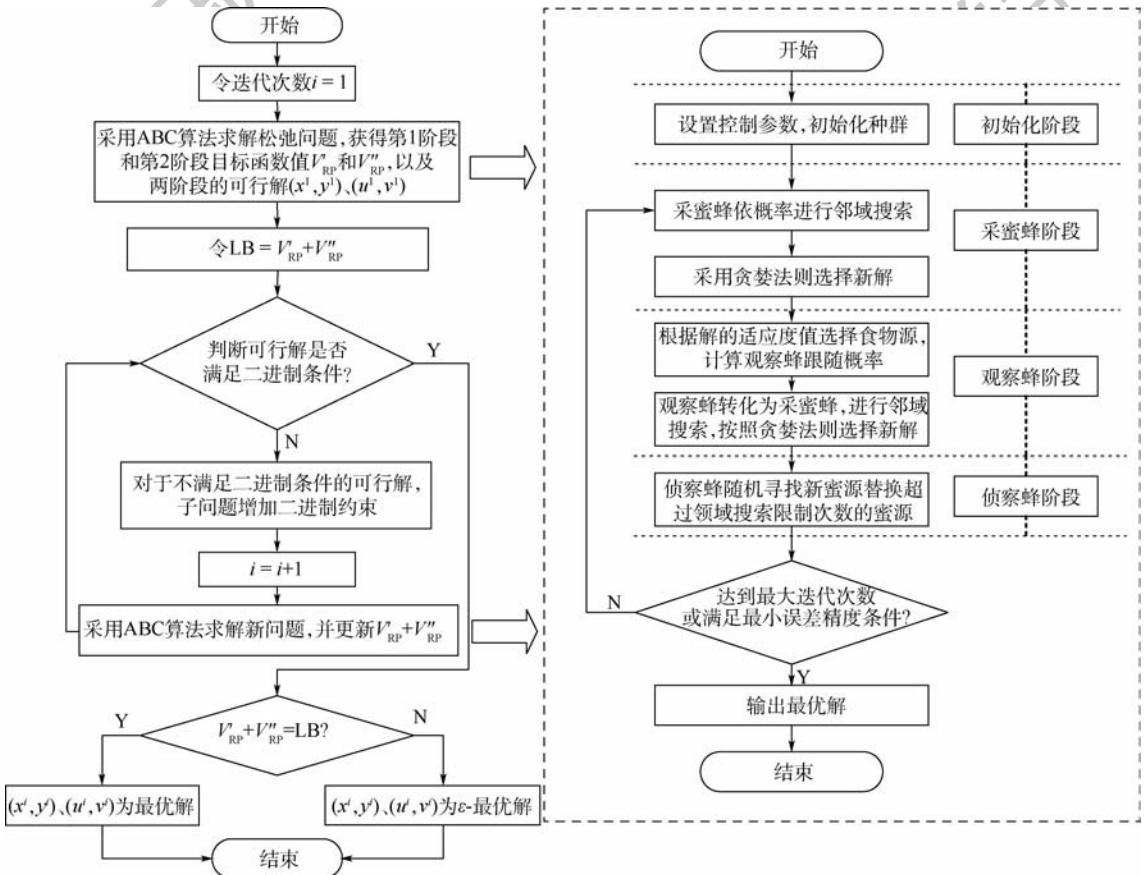


图2 基于ABC算法的渐进二元启发式方法计算框架

Fig.2 Progressive binary heuristic method computation framework based on ABC algorithm

表1 航班时隙请求

Table 1 Request of flight time slot

(a) 进场航班

航班	时隙	航班	时隙
A1	10:00	A16	11:05
A2	10:05	A17	11:10
A3	10:10	A18	11:15
A4	10:15	A19	11:20
A5	10:20	A20	11:20
A6	10:25	A21	11:25
A7	10:35	A22	11:30
A8	10:35	A23	11:35
A9	10:40	A24	11:40
A10	10:50	A25	11:40
A11	10:50	A26	11:40
A12	10:55	A27	11:40
A13	11:00	A28	11:50
A14	11:00	A29	11:55
A15	11:00	A30	11:55

(b) 离场航班

航班	时隙	航班	时隙
D1	10:00	D21	10:50
D2	10:00	D22	10:55
D3	10:05	D23	10:55
D4	10:05	D24	11:00
D5	10:10	D25	11:00
D6	10:10	D26	11:00
D7	10:15	D27	11:05
D8	10:15	D28	11:10
D9	10:15	D29	11:15
D10	10:20	D30	11:20
D11	10:20	D31	11:20
D12	10:20	D32	11:25
D13	10:25	D33	11:30
D14	10:25	D34	11:30
D15	10:30	D35	11:30
D16	10:30	D36	11:35
D17	10:30	D37	11:45
D18	10:35	D38	11:45
D19	10:40	D39	11:50
D20	10:45	D40	11:55

表2 联程航班

Table 2 Connecting flights

联程航班对	最小周转时间/min
A1, D12	25
A4, D18	25
A7, D20	30
A10, D27	30
A15, D28	20
A16, D30	20
A20, D38	30
A22, D40	25

### 3 算例验证

本节分别对本文提出的模型和算法进行算例验证,采用 MATLAB 进行仿真实验,仿真环境为 Intel Core i7 CPU@2.30 GHz,8 GB 内存。

#### 3.1 单机场模型算例验证

选取国内某机场终端区典型日运行数据为例,首先给出初始数据,然后采用本文提出的单机场时隙分配规划模型进行建模,并采用基于 ABC 算法的渐进二元启发式方法进行求解,最后进行分析说明。

##### 3.1.1 初始数据

该机场 10:00—12:00 的航班时隙请求如表 1 所示,其中时隙数量为 8,时隙长度为 15 min。联程航班需求如表 2 所示。表 3 给定了第 1 阶段不确定容量的分布及第 2 阶段容量实现值,其中第 1 阶段将容量设为不确定变量,并给出其不确定分布函数,如  $L(8,10)$  表示容量服从线性不确定分布,机场容量有 0% 的可能小于 8,100% 的可能小于等于 10。

##### 3.1.2 仿真结果

分别采用本文两阶段模型和传统模型进行仿真计算,其中两阶段模型的权重系数设为 1。算法参数设置如表 4 所示。仿真结果列于表 5,其中“请求时隙-计划时隙差”表示第 1 阶段初次时隙分配请求时隙和计划时隙的差值,“计划时隙-运行时隙差”表示第 2 阶段再次时隙分配计划时隙与运行时隙的差值。

##### 3.1.3 结果分析

由表 5 可以看出,本文模型所得出的“请求时隙-计划时隙差”为 13,即航班请求时隙和航班计划时隙有 13 个时隙的差异,“计划时隙-运行时隙差”为 7,即第 2 阶段航班运行延误时隙为 7,总体差异为 20;传统模型所得出的“请求时隙-计划时隙差”为 0,即满足所有航班时隙请求,“计划时隙-运行时隙差”为 24,即航班运行延误为 24。由此可以看出,本文模型虽然在初次时隙分配时增加了“请求时隙-计划时隙差”,但在第 2 阶段减少了运行延误,实现了总体目标的优化。从总体来看,时隙分配不确定两阶段模型较传统模型提升了 16.7%。

分析其原因,本文模型以牺牲航班时隙请求为代价,换取了更少的运行延误,从而实现时隙分配结果总体的优化。本质上,本文模型是对时隙资源的优化利用,模型在进行初次时隙分配时即考虑第 2 阶段时隙再次分配时容量的不确定性,

表3 机场两阶段容量

Table 3 Airport capacity in two stages

时段	第1阶段不确定容量分布			第2阶段容量实现值		
	进场	离场	机场	进场	离场	机场
10:00—10:30	$\mathcal{L}(3,5)$	$\mathcal{L}(6,8)$	$\mathcal{L}(9,11)$	4	7	10
10:30—11:00	$\mathcal{L}(2,4)$	$\mathcal{L}(5,7)$	$\mathcal{L}(7,9)$	3	6	8
11:00—11:30	$\mathcal{L}(2,4)$	$\mathcal{L}(4,6)$	$\mathcal{L}(6,8)$	2	5	6
11:30—12:00	$\mathcal{L}(3,5)$	$\mathcal{L}(5,7)$	$\mathcal{L}(8,10)$	4	6	9

表4 ABC算法控制参数设置

Table 4 Control parameter setting of ABC algorithm

控制参数	取值
种群规模	100
最大循环次数	1 000
最大限制搜索次数	50
观察蜂数量	种群规模的一半
采蜜蜂数量	种群规模的一半
侦察蜂数量	1

表5 时隙分配结果

Table 5 Results of time slot allocation

时段	两阶段模型			传统模型		
	请求时隙-计划时隙差	计划时隙-运行时隙差	合计	请求时隙-计划时隙差	计划时隙-运行时隙差	合计
10:00—10:15	0	0	0	0	0	0
10:15—10:30	1	0	1	0	1	1
10:30—10:45	0	0	0	0	1	1
10:45—11:00	0	0	0	0	0	0
11:00—11:15	3	1	4	0	4	4
11:15—11:30	3	2	5	0	6	6
11:30—11:45	4	2	6	0	7	7
11:45—12:00	2	2	4	0	5	5
总计	13	7	20	0	24	24

通过提前调整航班时隙请求,安排更合理的航班时隙计划,从而实现了时隙分配结果的整体优化。本节算例中,本文模型虽然仅比传统模型提升了16.7%,但在实际中,乘客所感受到的是第2阶段运行延误的减少,算例中,运行延误减少了70.8%,对乘客满意度的提升是巨大的。

### 3.2 多机场模型算例验证

京津冀、长三角、珠三角地区是中国在“十三五”期间着力打造的3个世界级机场群。本节选取3个区域的主要机场作为3个独立算例,对多机场时隙分配规划模型进行仿真验证,并对权重的影响进行分析。

首先列出3个算例中所含机场,以及国际民用航空组织(ICAO)给出的机场代码。

1) 京津冀。北京首都国际机场(ZBAA)、天津滨海国际机场(ZBTJ)、石家庄正定国际机场(ZBSJ)。

2) 长三角。上海浦东国际机场(ZSPD)、上海虹桥国际机场(ZSSS)、南京禄口国际机场(ZS-NJ)、杭州萧山国际机场(ZSHC)、合肥新桥国际机场(ZSOF)。

3) 珠三角。广州白云国际机场(ZGGG)、深圳宝安国际机场(ZGSZ)、珠海金湾机场(ZGSD)。

#### 3.2.1 初始数据及仿真结果

算例选取08:00—10:00三个机场群的典型运行数据,给定模型输入,包括各机场在第1个阶

段的不确定容量分布、第2个阶段的容量实现值,各航班请求时隙、起飞机场、降落机场、飞行时间,联程航班及最小周转时间、权重系数等初始数据。算例规模如表6所示。采用基于ABC算法的渐进二元启发式方法进行模型求解,计算过程与单机场模型相同,最终统计结果如表7、图3所示。

表6 算例规模

Table 6 Scale of computation samples

机场群	航班数量	进场航班数量	离场航班数量
京津冀	248	82	166
长三角	372	137	235
珠三角	309	84	225

表7 模型结果对比

Table 7 Model result comparison

机场群	权重系数	与传统模型对比		
		请求时隙-计划时隙差增加值	运行延误减少值	时隙分配优化百分比
京津冀	1	36	51	15.3
	3	67	176	66.9
长三角	1	63	91	10.3
	3	102	516	62.5
珠三角	1	75	139	62.7
	3	81	445	71.3

#### 3.2.2 结果分析

从表7和图3中可以看出,对比传统模型,本文模型在总体上体现了较大的优越性,提升比例

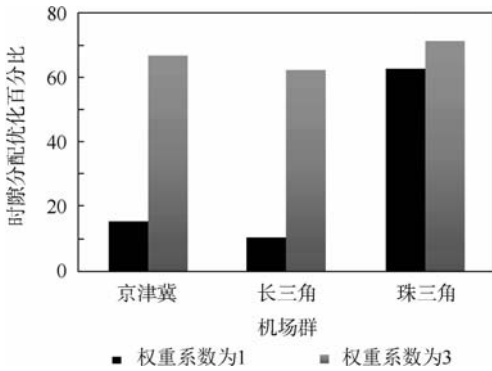


图3 时隙分配优化百分比示意图

Fig.3 Schematic diagram of time slot allocation optimization percentage

均大于10%,最大的提升是珠三角、权重系数为3的算例,提升比例达到了71.3%,提升最小的是长三角、权重系数为1的算例,提升比例达到了10.3%。这种提升的原因在3.1.3节中已经进行了分析,本文模型实现了对“请求时隙-计划时隙差”和“计划时隙-运行时隙差”的权衡,以牺牲时隙请求为代价,通过合理安排时隙计划,充分利用时隙资源,达到了总体优化的目标。

下面分析权重系数对模型结果的影响。如表7中所示,算例分别设定权重系数为1和3,可以看出,在其他输入数据相同的情况下,权重系数为3时本文模型比传统模型提升更大。具体来说,权重系数为3时,京津冀算例比权重系数为1时多提升了51.6%,长三角算例多提升了52.2%,珠三角算例多提升了8.6%。这是因为当设置权重系数为3时,就意味着把减少运行延误(“计划时隙-运行时隙差”)置于更加重要的位置,由于传统模型的特点,运行延误水平将被放大,而本文模型虽然可能牺牲更多的航班请求,但也会显著降低运行水平,因此显示出更大的优越性。

## 4 结 论

1) 本文提出的不确定容量下时隙分配两阶段规划单机场模型和多机场模型符合时隙分配两阶段过程,以航班而不是航班数为变量,将两阶段规划模型从单机场推广到多机场,与传统时隙分配模型相比,能够有效地对时隙分配结果进行优化。

2) 本文设计的基于ABC算法的渐进二元启发式方法,充分考虑了模型的二进制特点,从原问题的松弛问题出发,渐进增加二进制约束,降低了模型求解复杂度;利用ABC算法结构简单、易于

实现、参数较少、鲁棒性强的优点,将ABC算法设计为启发式计算求解器,算例证明,方法可以有效地对模型进行求解。

3) 算例对模型和方法进行了可行性和有效性验证。算例结果显示,在单机场和多机场情况下,本文模型均体现了较大优势;同时,模型权重系数越大,本文模型优势越大,对时隙分配结果提升效果越大。

未来应从模型完善及应用方面开展进一步研究。模型主要从考虑机型不同、优先级策略、航路变更等方面进行完善;应用方面主要考虑在更大规模的应用场景中对算法有效性进行验证。

## 参考文献 (References)

- [1] 中国民用航空局. 2015年民航行业发展统计公报[EB/OL]. (2016-05-30) [2018-12-24]. [http://www.caac.gov.cn/XXGK/XXGK/TJSJ/201605/t20160530\\_37643.html](http://www.caac.gov.cn/XXGK/XXGK/TJSJ/201605/t20160530_37643.html).  
Civil Aviation Administration of China. 2015 civil aviation industry development statistics bulletin[EB/OL]. (2016-05-30) [2018-12-24]. [http://www.caac.gov.cn/XXGK/XXGK/TJSJ/201605/t20160530\\_37643.html](http://www.caac.gov.cn/XXGK/XXGK/TJSJ/201605/t20160530_37643.html)(in Chinese).
- [2] 中国民用航空局,国家发展和改革委员会,交通运输部. 中国民用航空发展第十三个五年规划[EB/OL]. (2017-02-15) [2018-12-24]. [http://www.caac.gov.cn/XXGK/XXGK/FZGH/201704/t20170405\\_43502.html](http://www.caac.gov.cn/XXGK/XXGK/FZGH/201704/t20170405_43502.html).  
Civil Aviation Administration of China, National Development and Reform Commission, Ministry of Transport. The Thirteenth Five-Year Plan for the development of civil aviation in China [EB/OL]. (2017-02-15) [2018-12-24]. [http://www.caac.gov.cn/XXGK/XXGK/FZGH/201704/t20170405\\_43502.html](http://www.caac.gov.cn/XXGK/XXGK/FZGH/201704/t20170405_43502.html)(in Chinese).
- [3] 中国民用航空局运行监控中心. 2017年全国民航航班运行效率报告[EB/OL]. (2018-03-01) [2018-12-24]. [http://www.caac.gov.cn/XWZX/MHYW/201803/t20180328\\_56080.html](http://www.caac.gov.cn/XWZX/MHYW/201803/t20180328_56080.html).  
Operation Monitoring Center of Civil Aviation Administration of China. 2017 national civil aviation flight operation efficiency report[EB/OL]. (2018-03-01) [2018-12-24]. [http://www.caac.gov.cn/XWZX/MHYW/201803/t20180328\\_56080.html](http://www.caac.gov.cn/XWZX/MHYW/201803/t20180328_56080.html) (in Chinese).
- [4] HOFFMAN R L. Integer programming models for ground-holding in air traffic flow management[D]. City of College Park; University of Maryland, 1997.
- [5] HOFFMAN R L, HALL W, BALL M O, et al. Collaborative decision making in air traffic flow management[R]. Berkeley: University of California, Berkeley, 1999.
- [6] VOSSSEN T W. Fair allocation methods in air traffic management [D]. City of College Park; University of Maryland, 2002.
- [7] BALL M O, HOFFMAN R L, VOSSSEN T. An analysis of resource rationing methods for collaborative decision making [C]//Proceeding of ATM 2002-System Architectures and CNS Technologies Needed to Cope with the Air Traffic Capacity

- Problem and Related Evaluation Tools,2002:64-70.
- [ 8 ] MADAS M A,ZOGRAFOS K G. Airport slot allocation: From instruments to strategies [ J ]. Journal of Air Transport Management,2006,12(2):53-62.
- [ 9 ] ODONI A R. The flow management problem in air traffic control [ M ] // ODONI A R, BIANCO L, GIORGIO S. Flow control of congested networks. Berlin: Springer,1987:269-288.
- [ 10 ] TERRAB M, ODONI A, DEUTSCH O. Ground-holding strategies for ATC flow control [ C ] // AIAA Guidance, Navigation and Control Conference. Reston: AIAA,1989:1635-1646.
- [ 11 ] GILBO E P. Optimizing airport capacity utilization in air traffic flow management subject to constraints at arrival and departure fixes [ J ]. IEEE Transactions on Control Systems Technology,1997,5(5):490-503.
- [ 12 ] OUSSEDIK S, DELAHAYE D. Reduction of air traffic congestion by genetic algorithms [ C ] // International Conference on Parallel Problem Solving from Nature. Berlin: Springer,1998:855-864.
- [ 13 ] PULUGURTHA S S, NAMBISAN S S. Using genetic algorithms to evaluate aircraft ground holding policy in real time [ J ]. Journal of Transportation Engineering,2001,127(5):442-448.
- [ 14 ] MADAS M A, ZOGRAFOS K G. Airport slot allocation: A time for change? [ J ]. Transport Policy,2010,17(4):274-285.
- [ 15 ] 胡明华,徐肖豪. 空中交通流量控制的地面保持策略 [ J ]. 南京航空航天大学学报,1994,26(增刊):26-30.  
HU M H, XU X H. Ground-holding strategies for ATC flow control [ J ]. Journal of Nanjing University of Aeronautics and Astronautics,1994,26(S):26-30 (in Chinese).
- [ 16 ] 周茜,张学军,柳重堪. 时隙分配算法在 CDM GDP 程序中的应用 [ J ]. 北京航空航天大学学报,2006,32(9):1043-1045.  
ZHOU Q, ZHANG X J, LIU Z K. Slots allocation in CDM GDP [ J ]. Journal of Beijing University of Aeronautics and Astronautics,2006,32(9):1043-1045 (in Chinese).
- [ 17 ] 徐肖豪,王飞. 地面等待策略中的时隙分配模型与算法研究 [ J ]. 航空学报,2010,31(10):1993-2003.  
XU X H, WANG F. Research on slot allocation models and algorithms in ground holding policy [ J ]. Acta Aeronautica et Astronautica Sinica,2010,31(10):1993-2003 (in Chinese).
- [ 18 ] 张洪海,胡明华. CDM ADGDP 机场容量与时隙协同配置 [ J ]. 系统工程理论与实践,2010,30(10):1901-1908.  
ZHANG H H, HU M H. Collaborative allocation of capacity and slot in CDM ADGDP airport [ J ]. Systems Engineering-Theory & Practice,2010,30(10):1901-1908 (in Chinese).
- [ 19 ] 田勇,李永庆,万莉莉,等. 基于市场机制的地面等待时隙分配方法 [ J ]. 系统工程理论与实践,2014,34(6):1614-1619.  
TIAN Y, LI Y Q, WAN L L, et al. Slot allocation based on market mechanism in ground holding policy [ J ]. Systems Engineering-Theory & Practice,2014,34(6):1614-1619 (in Chinese).
- [ 20 ] 刘丽华. 市场机制下飞机推出时隙分配模型与算法研究 [ D ]. 哈尔滨:哈尔滨工业大学,2017.  
LIU L H. Research on model and algorithm for aircraft push-back slot allocation under market mechanism [ D ]. Harbin: Harbin Institute of Technology,2017 (in Chinese).
- [ 21 ] ANDREATTA G, ROMANIN-JACUR G. Aircraft flow management under congestion [ J ]. Transportation Science,1987,21(4):249-253.
- [ 22 ] RICETTA O, ODONI A R. Solving optimally the static ground-holding policy problem in air traffic control [ J ]. Transportation Science,1993,27(3):228-238.
- [ 23 ] MUKHERJEE A, HANSEN M M. A dynamic stochastic model for the single airport ground holding problem [ J ]. Transportation Science,2007,41(4):444-456.
- [ 24 ] 杨尚文,胡明华,张洪海. 随机型协同时隙分配模型 [ J ]. 系统工程理论与实践,2014,34(1):153-157.  
YANG S W, HU M H, ZHANG H H. Stochastic collaborative slot allocation models [ J ]. Systems Engineering-Theory & Practice,2014,34(1):153-157 (in Chinese).
- [ 25 ] 乐美龙,李星灿,高金敏. 机场到达时刻数量决策随机模型 [ J ]. 系统工程理论与实践,2017,37(11):2948-2954.  
LE M L, LI X C, GAO J M. Stochastic model of determining airport arrival slots number [ J ]. Systems Engineering-Theory & Practice,2017,37(11):2948-2954 (in Chinese).
- [ 26 ] COROLLI L, LULLI G, NTAIMO L. The time slot allocation problem under uncertain capacity [ J ]. Transportation Research Part C: Emerging Technologies,2014,46:16-29.
- [ 27 ] 岳仁田,赵胖胖,赵巍飞. 带补偿的两阶段随机规划航班时隙分配研究 [ J ]. 航空计算技术,2018,48(1):4-8.  
YUE R T, ZHAO P P, ZHAO Y F. Study on slot allocation based on two-stage stochastic programming with recourse [ J ]. Aeronautical Computing Technique,2018,48(1):4-8 (in Chinese).
- [ 28 ] LIU B. Uncertainty theory [ M ]. 4th ed. Berlin: Springer,2015:111-130.
- [ 29 ] ZHENG M F, YUAN Y, WANG Z T, et al. Study on two-stage uncertain programming based on uncertainty theory [ J ]. Journal of Intelligent Manufacturing,2017,28(3):633-642.
- [ 30 ] KARABOGA D. An idea based on honey bee swarm for numerical optimization [ R ]. Kayseri: Erciyes University,2005.
- [ 31 ] 江铭炎,袁东风. 人工蜂群算法及其应用 [ M ]. 北京:科学出版社,2014:47-55.  
JIANG M Y, YUAN D F. Artificial bee colony algorithm and its application [ M ]. Beijing: Science Press,2014:47-55 (in Chinese).

#### 作者简介:

亓尧 男,博士研究生。主要研究方向:不确定理论、空域资源规划。

王瑛 女,博士,教授,博士生导师。主要研究方向:装备系统工程、不确定理论。

## Two-stage programming model for time slot allocation problem under uncertain capacity

QI Yao<sup>1</sup>, WANG Ying<sup>1,\*</sup>, LIANG Ying<sup>2</sup>, YAO Di<sup>3</sup>

(1. Equipment Management and UAV Engineering College, Air Force Engineering University, Xi'an 710043, China;

2. Air Force Research Institute, Beijing 100089, China;

3. National Flight Flow Monitoring Center, Beijing 100091, China)

**Abstract:** In uncertain environment such as bad weather, it is easy to cause a large number of flight delays by the traditional time slot allocation method. To solve this problem, the time slot allocation process is first analyzed. Then a two-stage programming model for time slot allocation under uncertain capacity based on the uncertainty theory is proposed, including a single-airport model and a multi-airport model. The models highlight the tradeoff between the schedule slot/request slot discrepancies and operation slot/schedule slot discrepancies. According to the characteristics of the model, a progressive binary heuristic calculation method based on artificial bee colony (ABC) algorithm is designed to improve the efficiency of the solution. The validity of the model and algorithm is verified by the case study, and the model parameter setting is analyzed.

**Keywords:** time slot allocation; two-stage programming; uncertainty theory; artificial bee colony (ABC) algorithm; heuristic calculation

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

DOI: 10.13700/j.bh.1001-5965.2018.0775

# 基于可变阶变分模型的医用低剂量 CT 图像去噪

王娜<sup>1,2</sup>, 张权<sup>1,2</sup>, 刘祎<sup>1,2</sup>, 贾丽娜<sup>1,2</sup>, 桂志国<sup>1,2,\*</sup>

(1. 中北大学 生物医学成像与影像大数据山西省重点实验室, 太原 030051;

2. 中北大学 信息与通信工程学院, 太原 030051)



**摘 要:** 为了降低患者的辐射风险,低剂量 CT(LDCT)广泛用于临床诊断,但辐射剂量的减少在重建的 LDCT 图像中引入了斑点噪声和条纹伪影。为了提高 LDCT 图像的质量,提出了一种基于可变阶变分模型的后处理技术。所提出的变分模型使用边缘指示器控制变分阶数,根据图像的特征在一阶全变分(TV)正则项和二阶有界 Hessian(BH)正则项之间交替变换。采用基于快速傅里叶变换(FFT)的分裂 Bregman 算法求解所提出的变分模型。该模型在保留高剂量 CT(HDCT)图像相应结构的同时,有效抑制了斑点噪声和条纹伪影。重建的图像和实验数据表明,所提出的变分模型比现有的先进模型具有更好的质量。

**关键词:** 低剂量 CT(LDCT); 图像降噪; 边缘指示器; 全变分(TV); 有界 Hessian(BH); 快速傅里叶变换(FFT); 分裂 Bregman 算法

**中图分类号:** TP391

**文献标识码:** A **文章编号:** 1001-5965(2019)09-1757-08

X 射线 CT 为医学诊断和治疗提供了人体主要的解剖和病理信息<sup>[1]</sup>。近年来,越来越多的临床医生通过 CT 图像对患者的病症进行更准确的诊断,甚至借助 CT 图像的引导对病灶进行穿刺活检,对人体管道、体腔或器官组织内病理性积液、血肿、或脓肿进行穿刺引流达到治疗的效果。虽然 CT 技术的发展提高了人类的健康水平,但是高剂量和累积剂量的 X 射线照射会对人体造成不可逆转的放射性损害,甚至会诱发癌症和遗传性疾病<sup>[2-5]</sup>,这蕴藏着低剂量 CT(LDCT)技术研究和发展的新契机<sup>[6-8]</sup>。目前,在临床上降低 CT 诊断辐射剂量最行之有效的方法是通过调节 mA(毫安)或 mAs(毫安秒)的设置来降低 X 射线管电流,但是 X 射线管电流的降低势必导致重建后的 LDCT 图像被斑点噪声和条纹伪影所覆盖,严重退化的 LDCT 图像已不能满足临床诊断需求,

甚至会导致漏诊、误诊,这与医学 CT 技术发展的初衷相背而驰。

为使获取的 LDCT 图像满足临床诊断需求,有效提高 LDCT 图像的质量成为当前的研究热点<sup>[9-13]</sup>。学者们探索了许多策略<sup>[14-16]</sup>来提高 LDCT 图像的质量,目前直接对重建的 LDCT 图像实现消除噪声和抑制条形伪影的后处理方法是不依赖投影数据的,易于集成到大多数 CT 系统中,为 LDCT 成像技术的临床应用发展注入新的活力。鉴于难以对 LDCT 图像中分布不均匀的噪声和条形伪影进行统计建模,目前大多数后处理技术往往依赖于更多的启发式(heuristic)方法。近 20 年来,变分法在图像去噪领域独树一帜,具有良好的保边去噪性能<sup>[17-19]</sup>。基于此,研究者们将变分法应用到医学成像领域并取得新成效。受文献[17]启发,本文提出了一种基于可变阶变分模

收稿日期: 2019-01-02; 录用日期: 2019-04-05; 网络出版时间: 2019-05-07 16:53

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20190430.1707.002.html

基金项目: 国家自然科学基金(61671413,61801438); 国家重大科学仪器设备开发专项(2014YQ24044508); 山西省归国学者基金(2016-085)

\* 通信作者. E-mail: gzgtg@163.com

**引用格式:** 王娜, 张权, 刘祎, 等. 基于可变阶变分模型的医用低剂量 CT 图像去噪[J]. 北京航空航天大学学报, 2019, 45(9): 1757-1764. WANG N, ZHANG Q, LIU Y, et al. Medical low-dose CT image denoising based on variable order variational model [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1757-1764 (in Chinese).

型的后处理技术,用于改善 LDCT 图像的质量,该模型根据图像的局部特征自适应地选择正则化算子的阶数,在抑制斑点噪声和条形伪影的同时可以保留具有临床诊断意义的结构特征。

## 1 可变阶变分模型

基于变分法的图像降噪是求解关于图像的能量泛函最优解的问题,而图像降噪是一个不适定的反问题,所以常常运用正则化理论将不定问题转变为适定性问题来求最优解。

首先提出如下凸结合变分正则化模型:

$$\min_u E(u) = \frac{1}{2} \int_{\Omega} (u - u_0)^2 dx dy + \lambda \int_{\Omega} [\theta |\nabla u| + (1 - \theta) |\nabla^2 u|] dx dy \quad (1)$$

式中: $\Omega$ 为图像的定义域; $u$ 为降噪后的低剂量图像; $u_0$ 为带噪声和伪条纹的低剂量图像; $\nabla$ 为梯度算子; $\lambda$ 为正则化参数; $\theta \in [0, 1]$ 为变分阶数控制参数;等式右边第1项为保真项,第2项为对图像具有平滑作用的正则项。此外,本文模型中 Hessian 矩阵(即  $\nabla^2 u$ )记为:  $\nabla^2 u = \begin{bmatrix} \partial_x \partial_x u & \partial_x \partial_y u \\ \partial_x \partial_y u & \partial_y \partial_y u \end{bmatrix}$ 。显然,合理地选择  $\theta$ , 本文模型有以下几种情况:

1) 当  $\theta = 1$  时,本文模型可以改写为

$$\min_u E(u) = \frac{1}{2} \int_{\Omega} (u - u_0)^2 dx dy + \lambda \int_{\Omega} |\nabla u| dx dy \quad (2)$$

此时模型退化为正则化算子为一阶变分的总变分(TV)模型,该模型具有良好的保边性能,但会产生阶梯效应。

2) 当  $\theta = 0$  时,本文模型可以改写为

$$\min_u E(u) = \frac{1}{2} \int_{\Omega} (u - u_0)^2 dx dy + \lambda \int_{\Omega} |\nabla^2 u| dx dy \quad (3)$$

此时模型退化为正则化算子为二阶变分的有界 Hessian (BH) 模型,该模型具有良好的平滑性能,但是保边能力欠佳。

3) 当  $\theta \in (0, 1)$  时,本文模型类似于 TVBH 模型,兼顾了一阶变分和二阶变分,是对 TV 模型和 BH 模型的融合。

综上所述,参数  $\theta$  的选择决定着本文模型的滤波形式与滤波性能。由于参数常量  $\theta$  通常通过大量实验来试凑或凭经验取得最佳值,是对全局内容的粗略评估,忽略了图像的局部特征。

接下来考量图像的局部特征,用边缘扩散函

数  $\theta(z)$  来代替常量  $\theta$  改善本文模型的自适应性,提出了可变阶变分模型,优化的模型形式为

$$\min_u E(u) = \frac{1}{2} \int_{\Omega} (u - u_0)^2 dx dy + \lambda \int_{\Omega} [\theta(z) |\nabla u| + (1 - \theta(z)) |\nabla^2 u|] dx dy \quad (4)$$

$$\theta(z) = \exp\left[\frac{-3.31488}{(z/k)^8}\right] \quad (5)$$

式中: $k$ 为边缘阈值参数。

为了更好地保持医学图像中所包含的边缘细节信息,关于边缘扩散函数  $\theta(z)$  的特征检测因子  $z$  的构造,融合了梯度和局部熵 2 种特征检测算子。

1) 图像的梯度

梯度表征了图像灰度值的变化大小和方向,因此常用梯度模值来区分图像的边缘区域和非边缘区域。边缘区域梯度较大,平坦区域梯度较小,但是有些细节信息的梯度与平坦区域相差不大,噪声点处的梯度甚至比边缘的梯度还要大,这样梯度边缘检测算子可能会对图像中细节丰富的弱边缘区域和强噪声点造成误判,导致处理后的图像细节信息损失或者去噪不彻底。

2) 图像的局部熵

局部熵表征了图像中局部区域的像素灰度值变化的剧烈程度,从而可以反映图像中所包含信息的丰富程度。一副大小为  $M \times N$  的灰度图像  $f$  的熵值定义为

$$H = - \sum_{i=1}^M \sum_{j=1}^N p_{i,j} \lg p_{i,j} \quad (6)$$

$$p_{i,j} = \frac{f(i,j)}{\sum_{i=1}^M \sum_{j=1}^N f(i,j)} \quad (7)$$

式中: $f(i,j)$ 为位于图像  $(i,j)$  点处像素的灰度值; $p_{i,j}$ 表示  $(i,j)$  点处像素的灰度值在大小为  $M \times N$  的局部邻域中的分布概率; $H$ 为图像的局部熵。

根据局部熵可以有效地判定图像的局部特性,在灰度分布复杂的边缘细节区域,局部熵值较大;在灰度分布均匀的平坦区域,局部熵值较小。此外,局部熵具有较强的抗噪能力,独立噪声点对其影响很小。因此,局部熵可被广泛应用于图像处理中。

根据梯度和局部熵的特性,新的特征检测因子  $z$  设计为  $z = (|\nabla u_{\sigma}| + H)$ ,  $u_{\sigma}$  为对图像进行参数为  $\sigma$  (通常取为 0.5) 的高斯平滑。从式(5)可知,在给定阈值参数  $k$  的边缘扩散函数的引导下,本文模型可自适应地选择变分阶数。在强边缘

处,  $\theta(z) \approx 1$ , 此时正则化算子为一阶变分, 能够在滤除噪声时较好地保护边缘纹理等细节信息; 在平坦区域处,  $\theta(z) \approx 0$ , 此时正则化算子为二阶变分, 能够很好地平滑噪声; 在弱边缘处,  $0 < \theta(z) < 1$ , 此时正则化算子为混合阶变分, 实现较好的保边去噪效果。

## 2 可变阶变分模型的数值解法

### 2.1 离散化

为了提高本文提出的可变阶变分模型的运算速率, 类似文献 [17] 采用基于快速傅里叶变换 (FFT) 的分裂 Bregman 算法求解。

首先对连续的一阶、二阶和四阶微分算子及散度算子进行离散化。此外, 为了进一步提高分裂 Bregman 算法计算的速度, 采用周期性边界条件使 FFT 适用于分裂 Bregman 算法。设  $\Omega$  为尺寸大小为  $M \times N$  的二维灰度图像区域, 图像列和行方向的坐标分别用  $x$  和  $y$  表示。在像素点  $(i, j)$  处沿着坐标  $x$  和  $y$  方向的一阶向前差分记为

$$\partial_x^+ u_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j} & 1 \leq i \leq M, 1 \leq j < N \\ u_{i,1} - u_{i,j} & 1 \leq i \leq M, j = N \end{cases} \quad (8)$$

$$\partial_y^+ u_{i,j} = \begin{cases} u_{i+1,j} - u_{i,j} & 1 \leq i < M, 1 \leq j \leq N \\ u_{1,j} - u_{i,j} & i = M, 1 \leq j \leq N \end{cases} \quad (9)$$

一阶向后差分记为

$$\partial_x^- u_{i,j} = \begin{cases} u_{i,j} - u_{i,j-1} & 1 \leq i \leq M, 1 < j \leq N \\ u_{i,1} - u_{i,N} & 1 \leq i \leq M, j = 1 \end{cases} \quad (10)$$

$$\partial_y^- u_{i,j} = \begin{cases} u_{i,j} - u_{i-1,j} & 1 < i \leq M, 1 \leq j \leq N \\ u_{i,j} - u_{M,j} & i = 1, 1 \leq j \leq N \end{cases} \quad (11)$$

离散的二阶差分记为

$$\partial_x^+ \partial_x^- u_{i,j} = \partial_x^- \partial_x^+ u_{i,j} = \begin{cases} u_{i,N} - 2u_{i,j} + u_{i,j+1} & 1 \leq i \leq M, j = 1 \\ u_{i,j-1} - 2u_{i,j} + u_{i,j+1} & 1 \leq i \leq M, 1 < j < N \\ u_{i,j-1} - 2u_{i,j} + u_{i,1} & 1 \leq i \leq M, j = N \end{cases} \quad (12)$$

$$\partial_y^+ \partial_y^- u_{i,j} = \partial_y^- \partial_y^+ u_{i,j} = \begin{cases} u_{M,j} - 2u_{i,j} + u_{i+1,j} & i = 1, 1 \leq j \leq N \\ u_{i-1,j} - 2u_{i,j} + u_{i+1,j} & 1 < i < M, 1 \leq j \leq N \\ u_{i-1,j} - 2u_{i,j} + u_{1,j} & i = M, 1 \leq j \leq N \end{cases} \quad (13)$$

$$\partial_x^+ \partial_y^+ u_{i,j} = \partial_y^+ \partial_x^+ u_{i,j} = \begin{cases} u_{i,j} - u_{i+1,j} - u_{i,j+1} + u_{i+1,j+1} & 1 \leq i < M, 1 \leq j < N \\ u_{i,j} - u_{1,j} - u_{i,j+1} + u_{1,j+1} & i = M, 1 \leq j < N \\ u_{i,j} - u_{i+1,j} - u_{i,1} + u_{i+1,1} & 1 \leq i < M, j = N \\ u_{i,j} - u_{1,j} - u_{i,1} + u_{1,1} & i = M, j = N \end{cases} \quad (14)$$

$$\partial_x^- \partial_y^- u_{i,j} = \partial_y^- \partial_x^- u_{i,j} = \begin{cases} u_{i,j} - u_{i,N} - u_{M,j} + u_{M,N} & i = 1, j = 1 \\ u_{i,j} - u_{i,j-1} - u_{M,j} + u_{M,j-1} & i = 1, 1 < j \leq N \\ u_{i,j} - u_{i,N} - u_{i-1,j} + u_{i-1,N} & 1 < i \leq M, j = 1 \\ u_{i,j} - u_{i,j-1} - u_{i-1,j} + u_{i-1,j-1} & 1 < i \leq M, 1 < j \leq N \end{cases} \quad (15)$$

$$\partial_x^+ \partial_y^- u_{i,j} = \begin{cases} u_{i,j+1} - u_{i,j} - u_{M,j+1} + u_{M,j} & i = 1, 1 \leq j < N \\ u_{i,1} - u_{i,j} - u_{M,1} + u_{M,j} & i = 1, j = N \\ u_{i,j+1} - u_{i,j} - u_{i-1,j+1} + u_{i-1,j} & 1 < i \leq M, 1 \leq j < N \\ u_{i,1} - u_{i,j} - u_{i-1,j} + u_{i-1,j} & 1 < i \leq M, j = N \end{cases} \quad (16)$$

$$\partial_y^+ \partial_x^- u_{i,j} = \begin{cases} u_{i+1,j} - u_{i,j} - u_{i+1,N} + u_{i,N} & 1 \leq i < M, j = 1 \\ u_{1,j} - u_{i,j} - u_{1,N} + u_{i,N} & i = M, j = 1 \\ u_{i+1,j} - u_{i,j} - u_{i+1,j-1} + u_{i,j-1} & 1 \leq i < M, 1 < j \leq N \\ u_{1,j} - u_{i,j} - u_{1,j-1} + u_{i,j-1} & i = M, 1 < j \leq N \end{cases} \quad (17)$$

离散的四阶差分形式如式 (18) ~ 式 (20) 所示, 其中关于边界的有限差分的离散形式可参考式 (12) ~ 式 (15)。

$$\partial_x^+ \partial_x^- \partial_x^+ u_{i,j} = 6u_{i,j} + u_{i,j+2} + u_{i,j-2} - 4u_{i,j+1} - 4u_{i,j-1} \quad (18)$$

$$\partial_y^+ \partial_y^- \partial_y^+ u_{i,j} = 6u_{i,j} + u_{i+2,j} + u_{i-2,j} - 4u_{i+1,j} - 4u_{i-1,j} \quad (19)$$

$$\begin{aligned} \partial_x^- \partial_y^- \partial_x^+ \partial_y^+ u_{i,j} &= \partial_y^- \partial_x^- \partial_y^+ \partial_x^+ u_{i,j} = \\ &4u_{i,j} + u_{i+1,j+1} + u_{i-1,j+1} + u_{i+1,j-1} + u_{i-1,j-1} - \\ &2(u_{i,j+1} + u_{i,j-1} + u_{i+1,j} + u_{i-1,j}) \end{aligned} \quad (20)$$

离散化的梯度、Hessian 矩阵为

$$\nabla u = (\partial_x^+ u \quad \partial_y^+ u) \quad (21)$$

$$\nabla^2 u = \begin{bmatrix} \partial_x^- \partial_x^+ u & \partial_y^+ \partial_x^+ u \\ \partial_x^+ \partial_y^+ u & \partial_y^- \partial_y^+ u \end{bmatrix} \quad (22)$$

最后对于任意  $\mathbf{p} = (p_1 \quad p_2)$  的一阶散度  $\text{div}$ ,

任意  $\mathbf{q} = \begin{bmatrix} q_1 & q_2 \\ q_3 & q_4 \end{bmatrix}$  的二阶散度  $\text{div}^2$  的离散形式分

别记为

$$\text{div}(\mathbf{p})_{i,j} = \partial_x^-(p_1)_{i,j} + \partial_y^-(p_2)_{i,j} \quad (23)$$

$$\begin{aligned} \text{div}^2(\mathbf{q})_{i,j} &= \partial_x^+ \partial_x^-(q_1)_{i,j} + \partial_y^- \partial_x^-(q_2)_{i,j} + \\ &\partial_x^- \partial_y^-(q_3)_{i,j} + \partial_y^+ \partial_y^-(q_4)_{i,j} \end{aligned} \quad (24)$$

## 2.2 本文模型的分裂 Bregman 算法

为了求得本文提出的可变阶变分模型的能量泛函的最优值,首先引入辅助变量  $\mathbf{w}, \mathbf{v}$ , Bregman 迭代参数  $\mathbf{b}, \mathbf{d}$  和惩罚参数  $\theta_1, \theta_2$ 。将本文提出模型的能量泛函转换为以下形式:

$$\begin{aligned} E(u, \mathbf{w}, \mathbf{v}, \mathbf{b}, \mathbf{d}) &= \frac{1}{2} \int_{\Omega} (u - u_0)^2 dx dy + \\ &\lambda \int_{\Omega} [\theta(z) |\mathbf{w}| + (1 - \theta(z)) |\mathbf{v}|] dx dy + \\ &\frac{\theta_1}{2} \int_{\Omega} (\mathbf{w} - \nabla u - \mathbf{b})^2 dx dy + \\ &\frac{\theta_2}{2} \int_{\Omega} (\mathbf{v} - \nabla^2 u - \mathbf{d})^2 dx dy \end{aligned} \quad (25)$$

式中:  $\mathbf{w} = (w_1 \quad w_2)$  为一个与函数  $u$  的梯度相关

的二维函数且  $\mathbf{w} = \nabla u; \mathbf{v} = \begin{bmatrix} v_1 & v_2 \\ v_3 & v_4 \end{bmatrix}$  为一个与函数

$u$  的 Hessian 矩阵相关的四维向量函数且  $\mathbf{v} =$

$\nabla^2 u; \mathbf{b} = (b_1 \quad b_2)$  代表约束  $\mathbf{w} = \nabla u$  的 Bregman 迭

代参数;  $\mathbf{d} = \begin{bmatrix} d_1 & d_2 \\ d_3 & d_4 \end{bmatrix}$  代表约束  $\mathbf{v} = \nabla^2 u$  的 Breg-

man 迭代参数。

接下来固定变量  $(\mathbf{w}, \mathbf{v}; \mathbf{b}, \mathbf{d})$ , 得到关于  $u$  的欧拉方程为

$$\begin{aligned} u - \theta_1 \text{div}(\nabla u) + \theta_2 \text{div}^2(\nabla^2 u) &= \\ u_0 - \theta_1 \text{div}(\mathbf{w} - \mathbf{b}) + \theta_2 \text{div}^2(\mathbf{v} - \mathbf{d}) \end{aligned} \quad (26)$$

根据式(21)~式(24)可得式(26)的离散形式为

$$\begin{aligned} u_{i,j} - \theta_1 (\partial_x^- \partial_x^+ u_{i,j} + \partial_y^- \partial_y^+ u_{i,j}) + \\ \theta_2 (\partial_x^+ \partial_x^- \partial_x^- \partial_x^+ u_{i,j} + \partial_y^- \partial_x^- \partial_y^+ \partial_x^+ u_{i,j} + \\ \partial_x^- \partial_y^- \partial_x^+ \partial_y^+ u_{i,j} + \partial_y^+ \partial_y^- \partial_y^- \partial_y^+ u_{i,j}) = G_{i,j} \end{aligned} \quad (27)$$

式中:  $G_{i,j}$  为式(26)等号右侧部分, 离散形式为

$$G_{i,j} = u_0 - \theta_1 (\partial_x^-(w_{1i,j} - b_{1i,j}) + \partial_y^-(w_{2i,j} - b_{2i,j})) +$$

$$\begin{aligned} \theta_2 (\partial_x^+ \partial_x^-(v_{1i,j} - d_{1i,j}) + \partial_y^- \partial_x^-(v_{2i,j} - d_{2i,j}) + \\ \partial_x^- \partial_y^-(v_{3i,j} - d_{3i,j}) + \partial_y^+ \partial_y^-(v_{4i,j} - d_{4i,j})) \end{aligned} \quad (28)$$

在周期边界的条件下, 对等式(27) 两端分别进行离散傅里叶变换, 变换形式为

$$\begin{aligned} \mathcal{F}(u_{i,j} - \theta_1 (\partial_x^- \partial_x^+ u_{i,j} + \partial_y^- \partial_y^+ u_{i,j}) + \\ \theta_2 (\partial_x^+ \partial_x^- \partial_x^- \partial_x^+ u_{i,j} + \partial_y^- \partial_x^- \partial_y^+ \partial_x^+ u_{i,j} + \\ \partial_x^- \partial_y^- \partial_x^+ \partial_y^+ u_{i,j} + \partial_y^+ \partial_y^- \partial_y^- \partial_y^+ u_{i,j})) = \mathcal{F}(G_{i,j}) \end{aligned} \quad (29)$$

式中:  $\mathcal{F}$  为傅里叶变换符号。

式(29)关于离散频率  $r$  和  $s$  的等价形式为

$$\tau \mathcal{F}(u_{i,j}) = \mathcal{F}(G_{i,j}) \quad (30)$$

$$\begin{aligned} \tau = 1 - 2\theta_1 \left( \cos\left(\frac{2\pi s}{N}\right) + \cos\left(\frac{2\pi r}{M}\right) - 2 \right) + \\ 4\theta_2 \left( \cos\left(\frac{2\pi s}{N}\right) + \cos\left(\frac{2\pi r}{M}\right) - 2 \right)^2 \end{aligned} \quad (31)$$

通过对式(29)进行傅里叶逆变换, 求得  $u$  的解析解:

$$u_{i,j} = \mathcal{R} \left( \mathcal{F}^{-1} \left( \frac{\mathcal{F}(G_{i,j})}{\tau} \right) \right) \quad (32)$$

式中:  $\mathcal{F}^{-1}$  为傅里叶逆变换符号;  $\mathcal{R}$  为复数的实部。

此外, 关于变量  $\mathbf{w}$  的更新, 通过固定  $(u, \mathbf{w}; \mathbf{b}, \mathbf{d})$ , 获得与  $\mathbf{w}$  相关的欧拉方程:

$$\lambda \theta(z) \frac{\mathbf{w}}{|\mathbf{w}|} + \theta_1 (\mathbf{w} - \nabla u - \mathbf{b}) = \mathbf{0} \quad (33)$$

式(33)关于  $\mathbf{w}$  的解析解可以用如下二维广义软阈值公式求得:

$$\mathbf{w}_{i,j} = \max \left( \left| \nabla u_{i,j} + \mathbf{b}_{i,j} \right| - \frac{\lambda}{\theta_1} \theta(z), 0 \right) \frac{\nabla u_{i,j} + \mathbf{b}_{i,j}}{\left| \nabla u_{i,j} + \mathbf{b}_{i,j} \right|} \quad (34)$$

采用同样的方法更新变量  $\mathbf{v}$ , 其解析解为

$$\begin{aligned} \mathbf{v}_{i,j} = \max \left( \left| \nabla^2 u_{i,j} + \mathbf{b}_{i,j} \right| - \frac{\lambda}{\theta_2} (1 - \theta(z)), 0 \right) \cdot \\ \frac{\nabla^2 u_{i,j} + \mathbf{b}_{i,j}}{\left| \nabla^2 u_{i,j} + \mathbf{b}_{i,j} \right|} \end{aligned} \quad (35)$$

最后更新 Bregman 迭代参数  $\mathbf{b}, \mathbf{d}$ 。

综上所述, 本文提出的可变阶变分模型的基于 FFT 的分裂 Bregman 算法的求解流程可以归结如下:

**步骤 1** 初始化:  $u_0$  ( $u_0$  设定为待改善的 LDCT 图像)。

**步骤 2** 初始化:  $(\mathbf{w}, \mathbf{v}; \mathbf{b}, \mathbf{d}) = \mathbf{0}$ 。

**步骤 3** 设定参数  $\lambda, \theta_1$  及  $\theta_2$ 。

**步骤 4** 由  $(u, \mathbf{w}, \mathbf{v}) = \arg \min_{u, \mathbf{w}, \mathbf{v}} E(u, \mathbf{w}, \mathbf{v}; \mathbf{b}, \mathbf{d})$  更新变量  $u, \mathbf{w}$  及  $\mathbf{v}$ 。

**步骤 5** 更新 Bregman 迭代参数  $\mathbf{b}: \mathbf{b} \leftarrow \mathbf{b} +$

$\nabla u - w$ 。

步骤6 更新 Bregman 迭代参数  $d: d \leftarrow d + \nabla^2 u - v$ 。

步骤7 若满足迭代终止条件:  $\frac{\|u^n - u^{n-1}\|}{\|u^n\|} <$

$10^{-6}$ 时( $u^n$  和  $u^{n-1}$  分别为图像第  $n$  次和第  $n-1$  次迭代滤波后的结果),则停止迭代。反之,则转至步骤4,直到满足终止条件。

### 3 实验结果与分析

为了验证本文提出的可变阶变分模型对医用LDCT图像的去噪性能及在临床应用上的价值,选择实际CT扫描的胸腔体模、临床数据作为实验对象,其尺寸均为  $512 \text{ mm} \times 512 \text{ mm}$ 。图1(a)为胸腔体模HDCT图像,其通过一台多排探测器西门子16CT扫描仪在管电流为240 mAs、管电压为120 kVp(千伏峰值)的高剂量扫描协议下实际扫描一个人体胸腔的仿真体模,获得经反投影重建(FBP)算法重建后的图像,再利用大尺度邻域非局部伪影抑制(AS-LNLM)算法对重建后的图像质量进行进一步的改善而获得<sup>[20]</sup>。图1(b)为胸腔体模LDCT图像,其是通过将管电流降为30 mAs,在此低剂量的条件下扫描胸腔仿真体模的同一部位,经FBP重建后获得。图2(a)为胸腔体模HDCT图像的局部放大部分,图2(b)为胸腔体模LDCT图像的局部放大部分。图3(a)为临床胸腔LDCT数据,该数据是在管电流为60 mAs、管电压为120 kVp的低剂量扫描协议下得到的FBP重建后的DICOM格式的低剂量图像。图4(a)和图5(a)分别为临床腹

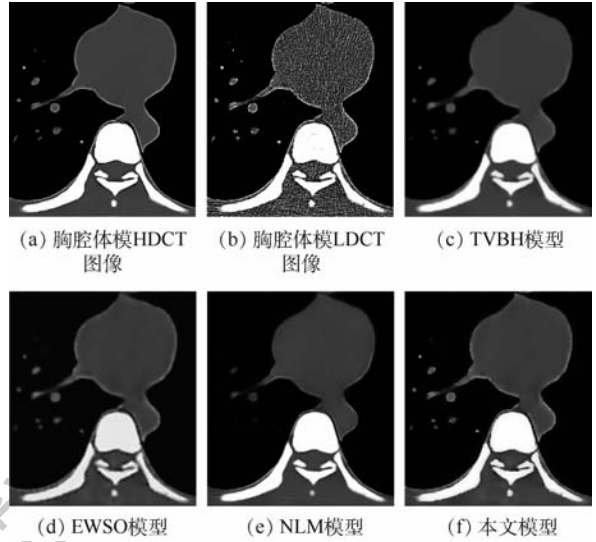


图2 真实胸腔体模局部放大部分的视觉比较  
Fig.2 Visual comparison of denoising results on the local enlarged drawing by the squares in Fig.1(a).

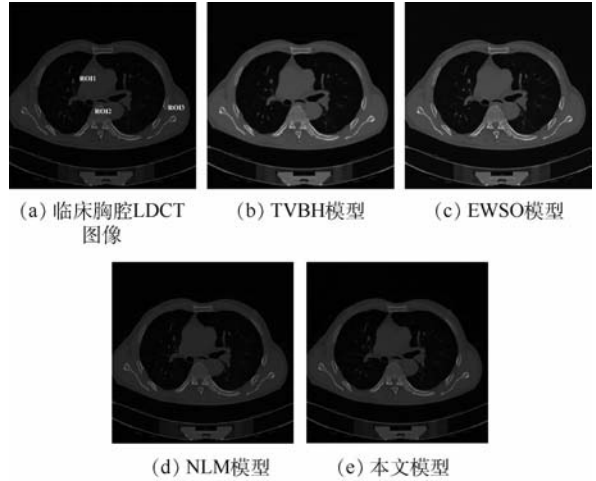


图3 临床胸腔LDCT图像降噪结果的视觉比较  
Fig.3 Visual comparison of denoising results on LDCT image of clinical thoracic

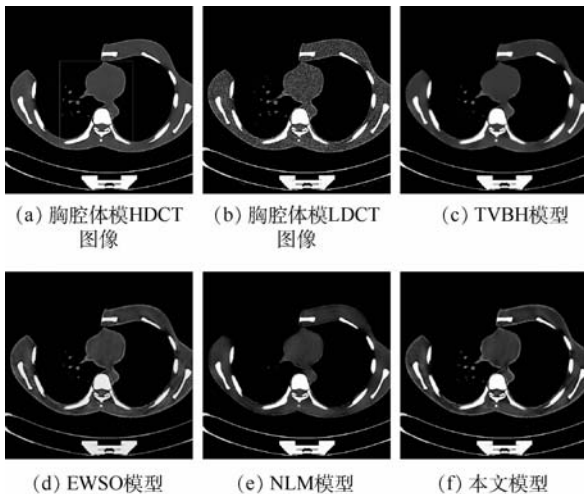


图1 真实胸腔体模LDCT图像降噪结果的视觉比较  
Fig.1 Visual comparison of denoising results on LDCT image of actual thoracic phantom

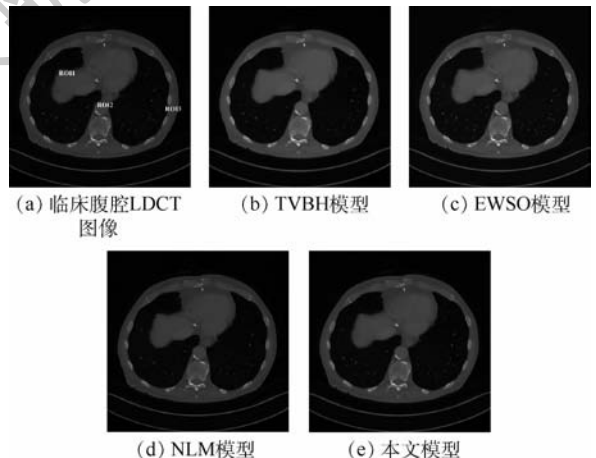


图4 临床腹腔LDCT图像降噪结果的视觉比较  
Fig.4 Visual comparison of denoising results on LDCT image of clinical abdominal

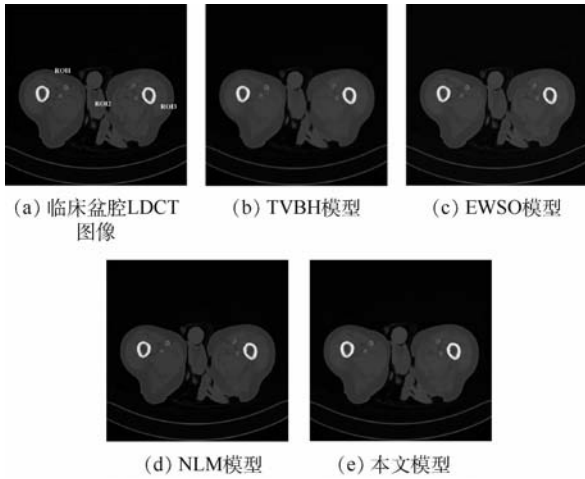


图5 临床盆腔 LDCT 图像降噪结果的视觉比较

Fig. 5 Visual comparison of denoising results on LDCT image of clinical pelvic

腔和盆腔 LDCT 数据,这 2 个数据是将辐射剂量降为标准剂量的 25% 的低剂量扫描协议下得到的 FBP 重建后的 DICOM 格式的低剂量图像。

实验中,分别采用 TVBH 模型<sup>[17]</sup>、EWSO 模型<sup>[18]</sup>、NLM 模型<sup>[21]</sup>和本文模型对胸腔体模 LDCT 图像和临床 LDCT 图像进行处理。本文模型的边缘扩散函数  $\theta(z)$  中,  $k$  是一个重要参数,其决定着要保持边缘的对比度并影响扩散过程。更具体地说,如果  $k$  值过大,图像的边缘被过度平滑;如果  $k$  值过小,图像中的一些噪声和伪影难以滤除。经大量实验验证,当  $k = 10$  时,本文模型降噪效果最佳。对于胸腔体模,本文模型的参数设定为:  $\lambda = 30, \theta_1 = \theta_2 = 5$ 。对于临床胸腔数据,本文模型的参数设定为:  $\lambda = 80, \theta_1 = \theta_2 = 10$ 。对于临床腹腔和盆腔数据,本文模型的参数设定为:  $\lambda = 50, \theta_1 = \theta_2 = 10$ 。为了客观说明本文模型的性能,其他对比模型的参数均按取得最优降噪效果设定。图 1 ~ 图 5 展示了各模型的降噪结果。

观察图 1 可知,本文模型对胸腔体模 LDCT 图像降噪效果在视觉上整体优于其他 3 种模型。为了更清晰地展示各模型的降噪效果,对图 1 进行了局部放大(见图 2)。图 2(c)为 TVBH 模型的降噪效果,该模型有效去除了噪声和条形伪影,但是过度平滑 LDCT 图像的一些细节信息会造成临床诊断的误诊。图 2(d)为 EWSO 模型的处理结果,该模型虽然可以抑制大部分噪声和条形伪影,但是把 LDCT 图像的组织结构附近的伪影误判为结构信息保留了下来,不能满足临床诊断需求。图 2(e)所示 NLM 模型的结果较好地保留了边缘信息,但是丢失了一些细小结构,同时也引入

了一些新的细小伪影,亦不能满足临床诊断需求。图 2(f)为本文模型的去噪结果,相对于其他 3 种对比模型,本文模型在抑制 LDCT 图像的条形伪影和边缘保持方面具有明显优势,其处理结果与 HDCT 图像最为相似,具有更加理想的视觉效果,为临床诊断带来便利。临床医学图像含有大量的弱边缘信息,相对仿真体模而言,临床医学 LDCT 图像的伪影去除难度更大。为了进一步验证本文模型在临床上的有效性,图 3 ~ 图 5 给出了各模型对临床数据 LDCT 图像的处理结果。根据对比结果可知,TVBH 模型很好地滤除了条形伪影,但是严重破坏了 LDCT 图像中组织结构;EWSO 模型和 NLM 模型的处理结果均含有部分伪影,视觉效果较差;本文模型可以有效滤除 LDCT 图像中的条形伪影,而且较好地保留了 LDCT 图像中的边缘组织结构信息,比其他 3 种模型具有更佳视觉效果。

此外,为了验证上述实验的有效性和合理性,针对图 1(a)所示的实际胸腔体模采用峰值信噪比(Peak Signal to Noise Ratio, PSNR)、平均结构相似度(Mean Structural Similarity Index Measure, MSSIM)和降噪时间作为衡量各模型去噪性能的客观性能评价指标,如表 1 所示。显然,本文模型相比于其他 3 种模型取得了较高的 PSNR 值和 MSSIM 值,这与上述实验的视觉效果对比图是吻合的,说明本文模型处理结果的降噪效果最优,且与对应的 HDCT 图像的相似度最高。另外,本文模型的降噪时间略高于 TVBH 模型和 EWSO 模型,明显低于 NLM 模型,总的来说耗时较短,便于应用于临床中。针对图 3 ~ 图 5 中的(a)图所示的临床数据缺乏高质量的参考图像,采用信噪比<sup>[22]</sup>(Signal to Noise Ratio, SNR)对局部感兴趣区域(ROI)进行质量评价,如表 2 所示。可以看出,本文模型对临床 LDCT 图像处理结果的定量评价均优于其他模型,进一步说明了本文模型处理临床数据的有效性。综上所述,本文模型是一个综合性能较优的去噪模型。

表 1 实际胸腔体模的定量比较

Table 1 Quantified comparison of actual thoracic phantom





降噪模型			
	PSNR/dB	MSSIM	降噪时间/s
TVBH 模型	25.259 3	0.886 4	19.45
EWSO 模型	24.213 9	0.914 9	10.83
NLM 模型	26.664 1	0.908 8	379.78
本文模型	27.205 1	0.926 6	27.88

表 2 临床数据的 SNR 值比较

Table 2 Comparison of SNR values of clinical data

LDCT 图像及 降噪模型									
	ROI1	ROI2	ROI3	ROI1	ROI2	ROI3	ROI1	ROI2	ROI3
LDCT 图像	0.5290	0.0497	0.0937	0.0733	0.2756	0.1200	0.0670	1.2014	0.0699
TVBH 模型	9.3063	0.0547	0.0983	0.1650	0.6978	0.2094	0.3417	2.4214	0.0780
EWSO 模型	14.3866	0.0546	0.0992	0.0935	0.4856	0.1624	0.1085	1.3202	0.0722
NLM 模型	1.3011	0.0580	0.0946	0.1731	0.8909	0.1594	0.1105	1.4467	0.0697
本文模型	26.2706	0.3608	2.0994	0.8750	1.2884	1.1297	0.6416	3.0495	0.0931

## 4 结 论

本文针对 LDCT 图像存在严重的斑点噪声和条纹伪影问题,提出了一种可变阶变分模型降噪模型。该模型兼顾了一阶变分模型和二阶变分模型的优点,根据图像的梯度模值和局部熵融合的特征检测因子将图像划分为平坦区域和细节区域,在不同的区域自适应地选择变分阶数,从而实现不同的扩散模式。同时采用基于 FFT 的分裂 Bregman 算法求解所提出的变分模型,提高了运算速度。仿真实验表明,本文所提出的可变阶变分模型不但有效地去除了 LDCT 图像中存在的噪声和条纹伪影,而且很好地保留了图像的边缘细节,明显改善了 LDCT 图像的质量,获得了与 HDCT 图相近的视觉效果,适用于临床诊断,具有实际应用价值。

### 参考文献 (References)

- [1] ZHU Y, ZHAO M, ZHAO Y, et al. Noise reduction with low dose CT data based on a modified ROF model[J]. *Optics Express*, 2012, 20(16): 17987-18004.
- [2] CHEN Y, YIN X, SHI L, et al. Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing[J]. *Physics in Medicine and Biology*, 2013, 58(16): 5803-5820.
- [3] ZHANG C, ZHANG T, LI M, et al. Low-dose CT reconstruction via  $L_1$  dictionary learning regularization using iteratively re-weighted least-squares[J]. *BioMedical Engineering OnLine*, 2016, 15(1): 66.
- [4] LEE D, LEE J, KIM H, et al. A feasibility study of low-dose single-scan dual-energy cone-beam CT in many-view under-sampling framework[J]. *IEEE Transactions on Medical Imaging*, 2017, 36(12): 2578-2587.
- [5] CHEN Y, LIU J, HU Y, et al. Discriminative feature representation: An effective postprocessing solution to low dose CT imaging[J]. *Physics in Medicine and Biology*, 2017, 62(6): 2103-2131.
- [6] CHEN Y, LIU J, XIE L, et al. Discriminative prior-prior image constrained compressed sensing reconstruction for low-dose CT imaging[J]. *Scientific Reports*, 2017, 7(1): 13868.
- [7] FRUSH D P, DONNELLY L F, ROSEN N S. Computed tomography and radiation risks: What pediatric health care providers should know[J]. *Pediatrics*, 2003, 112(4): 951-957.
- [8] BRENNER D J, HALL E J. Computed tomography-An increasing source of radiation exposure[J]. *New England Journal of Medicine*, 2007, 357(22): 2277-2284.
- [9] CHEN Y, SHI L, YANG J, et al. Radiation dose reduction with dictionary learning based processing for head CT[J]. *Australasian Physical and Engineering Sciences in Medicine*, 2014, 37(3): 483-493.
- [10] YANG Q, YAN P, ZHANG Y, et al. Low dose CT image denoising using a generative adversarial network with wasserstein distance and perceptual loss[J]. *IEEE Transactions on Medical Imaging*, 2018, 37(6): 1348-1357.
- [11] LIU J, MA J, ZHANG Y, et al. Discriminative feature representation to improve projection data inconsistency for low dose CT imaging[J]. *IEEE Transactions on Medical Imaging*, 2017, 36(12): 2499-2509.
- [12] HASAN A M, MELLI A, WAHID K A, et al. Denoising low-dose CT images using multi-frame blind source separation and block matching filter[J]. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 2018, 2(4): 279-287.
- [13] DIWAKAR M, KUMAR M. CT image denoising using NLM and correlation-based wavelet packet thresholding[J]. *IET Image Processing*, 2018, 12(5): 708-715.
- [14] YOU C, YANG Q, SHAN H, et al. Structure-sensitive multi-scale deep neural network for low-dose CT denoising[J]. *IEEE Access*, 2018, 6: 41839-41855.
- [15] LIU Y, ZHANG Y. Low-dose CT restoration via stacked sparse denoising autoencoders[J]. *Neurocomputing*, 2018, 284: 80-89.
- [16] 罗立民, 胡铁宁, 陈阳. 低剂量 CT 成像的研究现状与展望[J]. *数据采集与处理*, 2015, 30(1): 24-34.
- [17] LUO L M, HU Y N, CHEN Y. Research status and prospect for low-dose CT imaging[J]. *Data Acquisition and Processing*, 2015, 30(1): 24-34 (in Chinese).
- [18] LU W, DUAN J, QIU Z, et al. Implementation of high-order variational models made easy for image processing[J]. *Mathematical Methods in the Applied Sciences*, 2016, 39(14): 4208-4233.
- [19] DUAN J, QIU Z, LU W, et al. An edge-weighted second order variational model for image decomposition[J]. *Digital Signal*

Processing, 2016, 49: 162-181.

- [19] DUAN J, WARD W O C, SIBBETT L, et al. Introducing anisotropic tensor to high order variational model for image restoration[J]. Digital Signal Processing, 2017, 69: 323-336.
- [20] CHEN Y, YANG Z, HU Y, et al. Thoracic low-dose CT image processing using an artifact suppressed large-scale nonlocal means[J]. Physics in Medicine and Biology, 2012, 57(9): 2667-2688.
- [21] BUADES A, COLL B, MOREL J M. A non-local algorithm for image denoising[C] // IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:

IEEE Press, 2005, 2: 60-65.

- [22] WANG J, LU H, WEN J, et al. Multiscale penalized weighted least-squares sinogram restoration for low-dose X-ray computed tomography[J]. IEEE Transactions on Biomedical Engineering, 2008, 55(3): 1022-1031.

#### 作者简介:

王娜 女, 博士研究生。主要研究方向: 图像处理。

桂志国 男, 博士, 教授, 博士生导师。主要研究方向: 信号与信息处理、图像处理和识别、图像重建。

## Medical low-dose CT image denoising based on variable order variational model

WANG Na<sup>1,2</sup>, ZHANG Quan<sup>1,2</sup>, LIU Yi<sup>1,2</sup>, JIA Lina<sup>1,2</sup>, GUI Zhiguo<sup>1,2,\*</sup>

(1. Shanxi Provincial Key Laboratory for Biomedical Imaging and Big Data, North University of China, Taiyuan 030051, China;

2. School of Information and Communication Engineering, North University of China, Taiyuan 030051, China)

**Abstract:** Low-dose CT (LDCT) is widely used for clinical diagnosis to reduce radiation risk to patients.

However, the radiation dose reduction introduces mottle noise and streak artifacts into the reconstructed LDCT images. In this paper, a post-processing technique is proposed based on variable order variational model to improve the LDCT image quality. The proposed variational model employs the edge indicator to control the order of variation, which can alternate between the first order total variation (TV) regularizer and second order bounded Hessian (BH) regularizer based on the image feature. Moreover, the proposed model is solved by split Bregman algorithm based on fast Fourier transform (FFT). The proposed model effectively suppresses mottle noise and streak artifacts, meanwhile preserving structure in reference to high-dose CT (HDCT) images. The reconstructed images and experimental data indicate that the proposed model has better quality than some existing state-of-the-art models.

**Keywords:** low-dose CT (LDCT); image denoising; edge indicator; total variation (TV); bounded Hessian (BH); fast Fourier transform (FFT); split Bregman algorithm

**Received:** 2019-01-02; **Accepted:** 2019-04-05; **Published online:** 2019-05-07 16:53

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20190430.1707.002.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190430.1707.002.html)

**Foundation items:** National Natural Science Foundation of China (61671413, 61801438); National Key Scientific Instrument and Equipment Development Project of China (2014YQ24044508); Shanxi Provincial Foundation for Returned Scholars (Main Program), China (2016-085)

\* **Corresponding author.** E-mail: gzgtg@163.com

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

DOI: 10.13700/j.bh.1001-5965.2018.0680

# 空间目标的 ISAR 成像及轮廓特征提取



杨虹, 张雅声\*, 尹灿斌

(航天工程大学, 北京 101416)

**摘 要:** 空间目标的逆合成孔径雷达 (ISAR) 成像由于受自身遮挡及噪声干扰等影响, 导致生成的 ISAR 像难以直接进行图像分析及目标识别。由此, 以空间目标的 ISAR 成像建模为基础对 ISAR 像处理及特征提取展开了研究。首先, 分别建立了目标卫星的 ISAR 成像模型、ISAR 信号模型及 ISAR 图像函数提取模型, 并经过旁瓣抑制与相干斑滤波等初步处理得到了目标卫星的 ISAR 像。其次, 采用 Otsu 算法、canny 算子及 Hough 变换使卫星旋转至最长轴平行于像平面横轴, 通过闭运算填充卫星内部孔洞, 去除外部孤立噪声, 并基于连通域思想分割出卫星所在子区域, 实现了卫星的轮廓提取。所设计的图像处理算法能有效改善 ISAR 像质量, 提取的卫星轮廓线能较好地勾勒出目标卫星的外形结构, 为进一步展开卫星的识别工作奠定了重要基础。

**关键词:** 空间目标; 逆合成孔径雷达 (ISAR); ISAR 成像; 图像处理; 轮廓特征

**中图分类号:** TN951

**文献标识码:** A **文章编号:** 1001-5965(2019)09-1765-12

逆合成孔径雷达 (Inverse Synthetic Aperture Radar, ISAR) 成像具有全天时、全天候、远距离、高分辨等特点, 其能够提供丰富的目标结构信息。利用 ISAR 进行空间目标成像是空间态势感知的重要内容, 是夺取未来空天优势的重要保障。但是由于各种干扰因素及噪声的存在, 导致 ISAR 像的质量受损, 难以直接用于图像特征提取及空间目标识别等应用。为了解决该问题, 开展 ISAR 像的图像处理与改进显得尤为重要。文献[1]研究了 ISAR 像斑点噪声的去除方法, 并利用分水岭法进行 ISAR 像分割; 文献[2]针对 ISAR 像存在的低散斑噪声, 使用线性滤波器及中值滤波器提高图像质量, 并在图像质量严重受损时利用 Lee 滤波器替代中值滤波器; 文献[3]研究了利用色散相拉伸变换对数字图像进行边缘检测的方法; 文献[4]在数据不可用或严重受损的条件下提出利用梯度恢复算法对 ISAR 像进行分析与恢

复; 文献[5]针对雷达散射截面 (Radar Cross Section, RCS) 数据丢失的情况, 利用压缩感知算法生成目标的 ISAR 像, 并利用极化映射方法提取了目标特征; 文献[6]提出利用全局迭代阈值及局部散射点强度削减法分别对斑点噪声与横条纹干扰进行抑制, 并通过形态学、邻域平均法及图像归一化处理最终实现了 ISAR 像的预处理。上述研究成果均可用于 ISAR 像的处理与改进, 但是缺乏针对空间目标 ISAR 像实际问题 (如内部空洞存在、轮廓线损坏等) 的具体研究方案。

基于 ISAR 像的特征向量提取主要包括 ISAR 像数据生成、噪声抑制、边缘检测及轮廓特征提取等步骤。文献[7]研究了提取目标几何特征进行目标识别的方法; 文献[8]提出了侧抑制神经网络算法在目标识别中的应用; 文献[9]利用点目标仿真算法生成轨道目标的 ISAR 像, 并通过提取尺寸等参数进行了目标识别; 文献[10]研究了

收稿日期: 2018-11-22; 录用日期: 2019-04-26; 网络出版时间: 2019-06-10 09:30

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190606.0845.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190606.0845.001.html)

\* 通信作者。E-mail: lizhizys@139.com

**引用格式:** 杨虹, 张雅声, 尹灿斌. 空间目标的 ISAR 成像及轮廓特征提取[J]. 北京航空航天大学学报, 2019, 45(9): 1765-1776.  
YANG H, ZHANG Y S, YIN C B. ISAR imaging and contour feature extraction of space targets[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1765-1776 (in Chinese).

基于目标轮廓特征的分类器融合算法效果,并对4种融合算法的识别性能进行了对比分析;文献[11-12]提出了基于几何散列法的ISAR像自动识别方法及基于局部非负矩阵分解特征空间优化的ISAR像识别算法;文献[13]分别从目标微动特征、目标电磁特征等方面对ISAR像目标特征的提取方法展开了研究;文献[14]提出利用酉ESPRIT方法提取目标闭合轮廓线,为ISAR目标识别提供了更加稳定的特征;文献[15-17]针对ISAR像目标识别的特征提取与识别问题展开了深入研究。对于ISAR像特征向量尺度单一问题,文献[15]提出了一种基于直方图统计量的特征提取方法;文献[16]提出了一种基于分块双向二维投影梯度非负矩阵分解的特征向量提取方法,解决了特征提取过程中不允许负分解量存在的问题;文献[17]从融合角度出发,研究了ISAR像的Gabor特征,提出了一种基于Gabor小波变换幅值特征和相位特征相融合的ISAR像目标识别算法。

利用神经网络算法对空间目标进行识别时,需要确定目标在帧中的精确位置,解决该问题的方法之一在于利用目标轮廓线生成图像的矢量图,但由于实际情况中噪声干扰的存在,使得目标轮廓线模糊,增加了轮廓特征提取难度。对此,本文基于ISAR成像模型、ISAR信号模型及ISAR图像函数提取模型,设计了图像处理算法,该算法可以应用于不同目标的ISAR像。

## 1 ISAR 成像模型、坐标转换及方位分辨率

### 1.1 ISAR 成像模型

建立ISAR的卫星观测坐标系 $Oxyz$ ,卫星在观测坐标系 $Oxyz$ 中以速度 $\mathbf{V}$ 运动。以卫星的几何中心为原点,建立卫星的本地坐标系 $O'XYZ$ ,原点 $O'$ 与卫星几何中心重合。在本地坐标系中建立卫星的三维网格模型,利用三维网格模型上均匀分布的节点描述卫星的几何特性,如图1所示。

$N$ 为卫星三维网格模型上的任一节点,当前ISAR到节点 $N$ 的测量距离矢量为 $\mathbf{R}_N(p) = [x_N(p), y_N(p), z_N(p)]^T$ ,其可由观测坐标系 $Oxyz$ 的原点 $O$ 到节点 $N$ 的距离矢量表示:

$$\mathbf{R}_N(p) = \mathbf{R}_{O'}(p) + \mathbf{G}\mathbf{R}_{O'N} \quad (1)$$

式中: $\mathbf{R}_{O'}(p) = [x_{O'}(p), y_{O'}(p), z_{O'}(p)]$ 为第 $p$ 个脉冲时ISAR到卫星几何中心的距离矢量,由表达式 $\mathbf{R}_{O'}(p) = \mathbf{R}_{O'}(p-1) + \mathbf{V} \cdot T$ 确定, $p$ 为发射

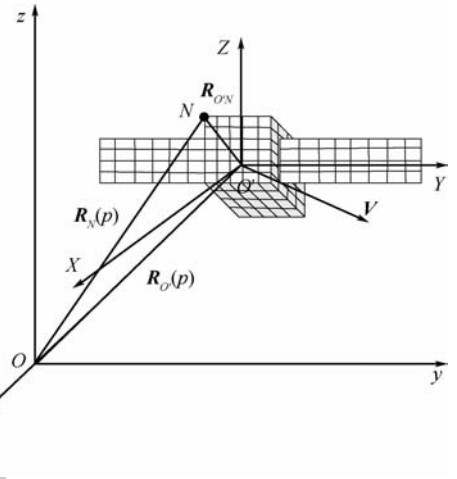


图1 ISAR 成像几何模型

Fig.1 ISAR imaging geometry model

的脉冲数, $p \in \{1, 2, \dots, M\}$ , $M$ 为发射的脉冲总数, $T$ 为脉冲重复周期, $\mathbf{R}_{O'}(0) = [x_{O'}(0), y_{O'}(0), z_{O'}(0)]$ 为ISAR在第1个脉冲 $p=1$ 时与卫星几何中心的距离矢量, $\mathbf{V} = [V \cos \alpha, V \cos \beta, V \cos \gamma]$ 为卫星的运动速度矢量, $\cos \alpha, \cos \beta, \cos \gamma$ 为速度矢量的方向余弦,且满足 $\cos \gamma = \sqrt{1 - \cos^2 \alpha - \sin^2 \beta}$ ; $\mathbf{R}_{O'N} = [X_{O'N}, Y_{O'N}, Z_{O'N}]$ 为本地坐标系 $O'XYZ$ 中卫星几何中心到节点 $N$ 的距离矢量, $X_{O'N} = m(\Delta X), Y_{O'N} = n(\Delta Y), Z_{O'N} = q(\Delta Z)$ , $m, n, q$ 分别为节点 $N$ 在本地坐标系 $O'XYZ$ 中的离散坐标, $\Delta X, \Delta Y, \Delta Z$ 为三维网格模型的网格单元尺寸; $\mathbf{G}$ 为本地坐标系 $O'XYZ$ 到观

测坐标系 $Oxyz$ 的转移矩阵,即:
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{G} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$
,转

移矩阵 $\mathbf{G}$ 可由欧拉表达式表示为

$$\begin{cases} a_{11} = \cos \theta_1 \cos \theta_3 - \sin \theta_1 \cos \theta_2 \sin \theta_3 \\ a_{12} = -\cos \theta_1 \sin \theta_3 - \sin \theta_1 \cos \theta_2 \cos \theta_3 \\ a_{13} = \sin \theta_1 \sin \theta_2 \\ a_{21} = \sin \theta_1 \cos \theta_3 + \cos \theta_1 \cos \theta_2 \sin \theta_3 \\ a_{22} = -\sin \theta_1 \sin \theta_3 + \cos \theta_1 \cos \theta_2 \cos \theta_3 \\ a_{23} = -\cos \theta_1 \sin \theta_2 \\ a_{31} = \sin \theta_2 \sin \theta_3 \\ a_{32} = \sin \theta_2 \cos \theta_3 \\ a_{33} = \cos \theta_2 \end{cases} \quad (2)$$

其中: $\theta_1, \theta_2, \theta_3$ 的表达式为

$$\theta_1 = \arctan\left(-\frac{A}{B}\right) \quad (3)$$

$$\theta_2 = \arccos\left(\frac{C}{(A^2 + B^2 + C^2)^{\frac{1}{2}}}\right) \quad (4)$$

$$\theta_3 = \arccos\left(\frac{V_x B - V_y A}{[(A^2 + B^2)(V_x^2 + V_y^2 + V_z^2)]^{\frac{1}{2}}}\right) \quad (5)$$

式中:  $A$ 、 $B$ 、 $C$  分别为基准平面  $Q$  的法向量分量;  $V_x$ 、 $V_y$ 、 $V_z$  分别为卫星速度矢量的分量。该基准平面  $Q$  的定义为: 过观测坐标系  $Oxyz$  原点  $O$ 、卫星几何中心在第  $p$  次脉冲时位置  $O'(x_{o'}(p), y_{o'}(p), z_{o'}(p))$  及第  $p$  次脉冲内卫星运动轨迹线方程所确定的平面。第  $p$  次脉冲内的卫星运动轨迹线方程可以表示为 (ISAR 的一个合成孔径时间为微秒级, 因此在一个合成孔径时间内将卫星运动轨迹近似为直线运动)

$$\frac{x - x_{o'}(p)}{V_x} = \frac{y - y_{o'}(p)}{V_y} = \frac{z - z_{o'}(p)}{V_z} \quad (6)$$

其矩阵方程可描述为

$$\det \begin{bmatrix} x & y & z \\ x_{o'}(p) & y_{o'}(p) & z_{o'}(p) \\ V_x & V_y & V_z \end{bmatrix} = 0 \quad (7)$$

式(7)也可以写为

$$Ax + By + Cz = 0 \quad (8)$$

式中:

$$\begin{cases} A = V_z y_{o'}(p) - V_y z_{o'}(p) \\ B = V_x z_{o'}(p) - V_z x_{o'}(p) \\ C = V_y x_{o'}(p) - V_x y_{o'}(p) \end{cases} \quad (9)$$

因此, 可以求得第  $N$  个散射点的位置矢量模型  $\mathbf{R}_N(p)$ , 在对 ISAR 线性调频信号建模时, 可以利用其计算从卫星散射体反射信号的时间延迟。

### 1.2 ISAR 成像几何转换

从卫星上反射回来的 ISAR 信号, 可以在二维跟踪坐标系中利用径向距离  $Y_v$  和横向距离  $X_u$  进行描述, 这就意味着卫星的三维几何被转换成在二维坐标系  $X_u$ - $Y_v$  中描述的二维几何。二维坐标系  $X_u$ - $Y_v$  与二维信号配准平面共面。二维信号配准平面是由卫星的几何中心、轨迹线及观测坐标系  $Oxyz$  的原点  $O$  所确定的平面。因此, 二维信号配准平面与式(8)定义的基准平面  $Q$  重合。

在二维信号配准平面中, 建立 2 个二维坐标系: 坐标系  $OX_u Y_v$  跟踪卫星的几何中心, 坐标系  $O'X_u Y_v$  与卫星本体固连 (见图 2)。图 2 中: 坐标系  $OX_u Y_v$  中,  $OY_v$  轴沿  $OO'$  连线方向,  $OX_u$  轴在基准平面  $Q$  内与  $OY_v$  相垂直; 坐标系  $O'X_u Y_v$  中,  $O'X_u$  轴与速度方向  $\mathbf{V}$  相同,  $O'Y_v$  轴在基准平面  $Q$  内与  $O'X_u$  轴垂直。

第  $N$  个散射点的坐标  $(X_{u'v'}, Y_{u'v'})$  可以由该点的位置矢量  $\mathbf{R}_{o'N}$  及坐标轴的单位方向矢量  $\mathbf{e}_x$ 、 $\mathbf{e}_y$  表示:

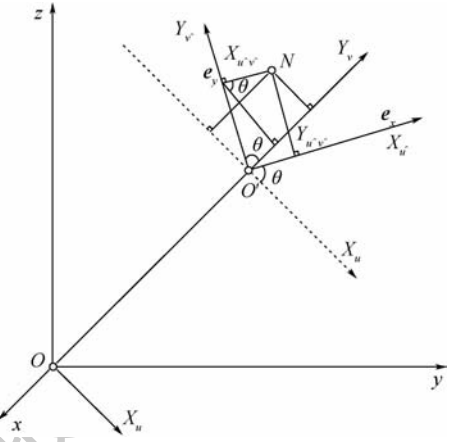


图 2 ISAR 信号平面上的两个二维坐标系  
Fig. 2 Two two-dimensional coordinate systems established on ISAR signal plane

$$X_{u'v'} = \frac{\mathbf{e}_x (\mathbf{G} \cdot \mathbf{R}_{o'N})}{|\mathbf{e}_x|} \quad (10)$$

$$Y_{u'v'} = \frac{\mathbf{e}_y (\mathbf{G} \cdot \mathbf{R}_{o'N})}{|\mathbf{e}_y|} \quad (11)$$

式中:  $\mathbf{e}_x = \mathbf{V}$ ;  $\mathbf{e}_y = [e_{y1}, e_{y2}, e_{y3}]^T$ ,  $e_{y1} = BV_z - CV_y$ ,  $e_{y2} = CV_x - AV_z$ ,  $e_{y3} = AV_y - BV_x$ 。

### 1.3 卫星运动分解

引入二维跟踪坐标系, 卫星的运动轨迹被分解为几何中心  $O'$  沿  $OY_v$  轴的平移及相对于几何中心  $O'$  的旋转。卫星几何中心平移后的位置矢量为  $\mathbf{R}_{o'}(p)$ , 卫星绕几何中心的旋转由本体坐标系  $O'X_u Y_v$  相对于跟踪坐标系  $OX_u Y_v$  的旋转定义。

坐标轴  $Y_v$  与  $Y_v$  之间的旋转角度  $\theta(p)$  可通过其方向矢量  $\mathbf{R}_{o'}(p)$  与  $\mathbf{e}_y$  内积的反余弦求得:

$$\theta(p) = \arccos\left(\frac{(x_{o'}(p)e_{y1} + y_{o'}(p)e_{y2} + z_{o'}(p)e_{y3}) / ((e_{y1}^2 + e_{y2}^2 + e_{y3}^2)^{\frac{1}{2}} \cdot (x_{o'}^2(p) + y_{o'}^2(p) + z_{o'}^2(p))^{\frac{1}{2}})}{1}\right) \quad (12)$$

在跟踪坐标系  $OX_u Y_v$  中的第  $u'v'$  个散射点的位置矢量可以表示为

$$\mathbf{R}_{u'v'}(p) = \mathbf{R}_{o'}(p) + \mathbf{H} \cdot \mathbf{R}_{u'v'} \quad (13)$$

式中:  $\mathbf{R}_{u'v'}(p) = [X_{u'v'}(p), Y_{u'v'}(p)]^T$  为第  $u'v'$  个点的位置矢量;  $\mathbf{R}_{o'}(p) = [0, \mathbf{R}_{o'}(p)]^T$  为卫星几何中心在坐标系  $OX_u Y_v$  中的距离矢量;  $\mathbf{H} = \begin{bmatrix} \cos \theta(p) & -\sin \theta(p) \\ \sin \theta(p) & \cos \theta(p) \end{bmatrix}$  为第  $u'v'$  个点相对于几何中心的旋转矩阵。

从式(13)中可以看出, 卫星的运动被分解为卫星几何中心的径向平移  $\mathbf{R}_{o'}(p) = [0, \mathbf{R}_{o'}(p)]^T$  和卫星绕其几何中心的旋转  $\mathbf{H} \cdot \mathbf{R}_{u'v'}$ 。

### 1.4 卫星的方位分辨率

为了确定 ISAR 像在帧中的位置,在基准平面  $Q$  内建立静止二维观测坐标系  $Ox'y'$ ,如图 3 所示。

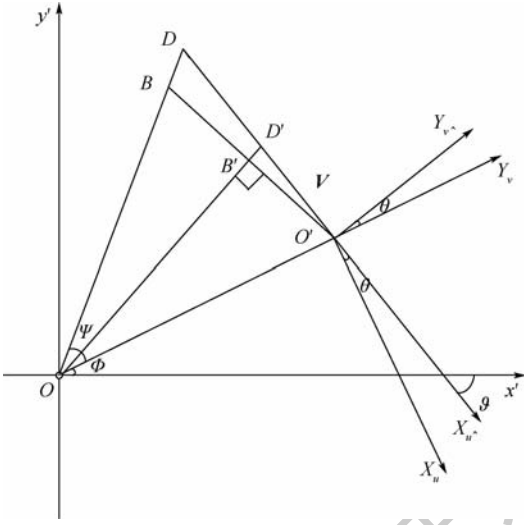


图3 卫星在相平面中的位置及方位角分辨率

Fig.3 Satellite position in imaging plane and azimuth resolution

$Ox'$  为基准平面  $Q$  与平面  $Oxy$  之间的交线,定义为:  $Ax + By + Cz = 0, z = 0$ ,也可以表示为:  $x/B = -y/A$ 。

根据坐标轴  $Oy'$  与  $Ox'$  及基准平面  $Q$  的正交性,可以求得  $Oy'$  的表达式为

$$\frac{x}{AC} = \frac{y}{BC} = \frac{-z}{A^2 + B^2} \quad (14)$$

坐标轴  $Ox'$  与速度矢量之间的夹角  $\vartheta$  定义为

$$\vartheta = \pi - \arccos\left(\frac{BV_x - AV_y}{\sqrt{A^2 + B^2}\sqrt{V_x^2 + V_y^2 + V_z^2}}\right) \quad (15)$$

由图 3 可知,在一个合成孔径时间内,卫星从点  $O'$  运动到点  $D$ 。雷达观测方向  $OO'$  与坐标轴  $Ox'$  之间的初始方位角  $\Phi(0)$  为

$$\Phi(0) = \arccos\left(\frac{x_{O'}(0)B - y_{O'}(0)A}{\sqrt{A^2 + B^2}\sqrt{x_{O'}^2(0) + y_{O'}^2(0) + z_{O'}^2(0)}}\right) \quad (16)$$

当卫星运动到点  $D$  时,方位角  $\Phi(p)$  为

$$\Phi(p) = \arccos\left(\frac{x_{O'}(p)B - y_{O'}(p)A}{\sqrt{A^2 + B^2}\sqrt{x_{O'}^2(p) + y_{O'}^2(p) + z_{O'}^2(p)}}\right) \quad (17)$$

点  $O'(0)$  为卫星几何中心在初始时刻的位置, $OO'$  为雷达到卫星几何中心的初始观测方向。在图 3 中, $O'D$  与  $Ox'$  之间的夹角大小为  $\vartheta$ ,定义

$O'D$  为一个合成孔径长度,其值由合成孔径时间内的发射脉冲数  $n_p$ 、脉冲发射周期及卫星运动速度确定:

$$O'D = 2O'D' = n_p VT \quad (18)$$

$OO'$  与  $OD$  之间的夹角  $\Psi$  表示一个合成孔径长度对应的角度大小:

$$\Psi = \Phi(p) - \Phi(0) \quad (19)$$

径向分辨率由发射脉冲的带宽决定,为使方位分辨率等于径向分辨率,有必要定义合成孔径的长度大小或者单位合成孔径时间内的脉冲周期数  $n_p$ 。方位分辨率大小由有效合成孔径长度计算得到,有效合成孔径长度  $O'B$  可以表示为

$$O'B = 2O'D'\sin(\vartheta + \Phi + \Psi) \quad (20)$$

基于衍射理论,发射波长为  $\lambda$  的雷达波的方位角分辨率大小为

$$\Delta\phi = \frac{\lambda}{O'B} = \frac{\lambda}{2O'D'\sin(\vartheta + \Phi + \Psi)} \quad (21)$$

由此,计算得到雷达方位分辨率  $\Delta L$  为

$$\Delta L = OO' \frac{\Delta\phi}{2} = \frac{OO' \cdot \lambda}{2n_p VT \sin(\vartheta + \Phi + \Psi)} \quad (22)$$

实际中,雷达与卫星间距离为百公里量级,在一个合成孔径时间内,方位分辨率可近似为

$$\Delta L = \frac{\lambda}{2\Psi} \quad (23)$$

为了实现特定的方位分辨率,由式(22)可以求得一个合成孔径周期内的脉冲数  $n_p$  为

$$n_p = \frac{OO' \cdot \lambda}{2\Delta L VT \sin(\vartheta + \Phi + \Psi)} \quad (24)$$

## 2 ISAR 信号模型及图像重构

在球面波的传播条件下,卫星反射的线性调频信号可以表示为

$$S(p, t) = \sum_N a_N \text{rect}\left(\frac{t - t_N(p)}{T_p}\right) \exp\{j[\omega(t - t_N(p)) + b(t - t_N(p))^2]\} \quad (25)$$

式中:

$$\text{rect}\left(\frac{t - t_N(p)}{T_p}\right) = \begin{cases} 1 & 0 \leq \frac{t - t_N(p)}{T_p} < 1 \\ 0 & \frac{t - t_N(p)}{T_p} < 0, \frac{t - t_N(p)}{T_p} \geq 1 \end{cases} \quad (26)$$

其中: $a_N$  为卫星上第  $N$  个散射点的后向散射系数; $T_p$  为线性调频信号脉冲宽度; $\omega = 2\pi c/\lambda$  为载波角频率,雷达波的传播速度为  $c = 3 \times 10^8$  m/s; $b = \pi\Delta F/T_p$ ,  $\Delta F/T_p$  线性调频信号的调频率,  $\Delta F$  为带宽; $t_N(p) = 2R_N(p)/c$  为信号从第  $N$  个散射

点返回的时间延迟,  $R_N(p)$  为雷达到散射点  $N$  的距离大小;  $t = t_{N\min}(p) + k\Delta T$  为径向测量的快时间,  $k \in \{1, 2, \dots, K\}$  为线性调频信号的采样数,  $K = T/\Delta T$  为信号样本的总数量,  $\Delta T$  为样本的采样周期,  $t_{N\min}(p) = 2R_{N\min}(p)/c$  为从卫星最近散射点反射 ISAR 信号的最小时间延迟。

在式(25)的两侧分别乘以发射波形的复共轭波形  $\exp[-j(\omega t + bt^2)]$ , 得到解调后的 ISAR 信号为

$$\hat{S}(k, p) = \sum_N a_N \text{rect}\left(\frac{t - (2R_N(p)/c)}{T_p}\right) \cdot \exp\left\{-j\left[2\omega_k \frac{R_N(p)}{c} - b\left(\frac{2R_N(p)}{c}\right)^2\right]\right\} \quad (27)$$

式中:  $\omega_k = \hat{\omega}(p) + \Delta\omega \cdot k$  为每次采样时刻的回波角频率,  $\Delta\omega \cdot k$  为  $k$  次采样角频率的变化值,  $\hat{\omega}(p) = \omega + 2b(2R_{N\min}(p)/c)$  为每个  $p$  值对应的恒定角频率, 随着脉冲发射数  $p$  缓慢变化。

从式(27)中可以看出, 三维图像函数  $a_N$  被转化成二维信号函数。通过对  $\hat{S}(k, p)$  进行逆变换, 可以实现图像函数  $a_N$  的提取。

$$a_N = \sum_{p=1}^M \sum_{k=1}^K \hat{S}(k, p) \cdot \exp\left\{j\left[2\omega_k \frac{R_N(p)}{c} - b\left(\frac{2R_N(p)}{c}\right)^2\right]\right\} \quad (28)$$

式(28)可以视为卫星所有散射点的相位补偿过程。因此, ISAR 成像属于一种全运动补偿。其中, 指数项  $\exp\left[-jb\left(\frac{2R_N(p)}{c}\right)^2\right]$  补偿  $\hat{S}(k, p)$  中的二次相位, 即实现二阶运动补偿。考虑到卫星常态运动时, 在一个合成孔径时间内的变化缓慢, 因此可以忽略二次项  $jb\left(\frac{2R_N(p)}{c}\right)^2$ 。

在将三维 ISAR 信号转换到二维信号平面并忽略掉二次项以后, 式(28)可以重写为

$$a_{u\hat{v}} = \sum_{p=1}^M \sum_{k=1}^K \hat{S}(k, p) \exp\left[j\left(2\omega_k \frac{R_{u\hat{v}}(p)}{c}\right)\right] \quad (29)$$

式中: 信号平面上散射点的离散坐标取值范围为:  $\hat{u} \in \{1, 2, \dots, M\}$ ,  $\hat{v} \in \{1, 2, \dots, K\}$ 。即在图像重构过程中, 方位向  $\hat{u}$  和径向  $\hat{v}$  的像素个数分别等于发射脉冲数  $p$  及每个脉冲内的采样数  $k$ 。

在雷达反射波近似为平面波的情况下, 雷达到任意散射点  $u\hat{v}$  的距离  $R_{u\hat{v}}(p)$  可以等效为径向距离, 因此  $R_{u\hat{v}}(p)$  可以表示为

$$R_{u\hat{v}}(p) = R_{O'}(p) + X_{u\hat{v}} \sin \theta(p) + Y_{u\hat{v}} \cos \theta(p) \quad (30)$$

表达式(29)可以重新写为

$$a_{u\hat{v}} = \sum_{p=1}^M \sum_{k=1}^K \hat{S}(k, p) \cdot \exp\left[j\left(\frac{2\omega_k}{c} X_{u\hat{v}} \sin \theta(p) + \frac{2\omega_k}{c} Y_{u\hat{v}} \cos \theta(p)\right)\right] \quad (31)$$

式中:  $\hat{S}(k, p) = \hat{S}(k, p) \exp\left[j\left(2\omega_k \frac{R_{O'}(p)}{c}\right)\right]$ 。

指数项  $\exp\left[j\left(2\omega_k \frac{R_{O'}(p)}{c}\right)\right]$  补偿卫星径向位移引起的相位, 补偿数据由 1.1 节中雷达到卫星几何中心的距离计算得到。在积分区间极小的情况下, 即  $\theta(p) \rightarrow 0$ , 可作如下近似:  $\cos \theta(p) \approx 1$ ,  $\sin \theta(p) = \theta(p) = p(\Delta\theta)$ ,  $\Delta\theta$  为  $\theta(p)$  的增量。此时, 式(31)也可以表示为

$$a_{u\hat{v}} = \sum_{p=1}^M \sum_{k=1}^K \hat{S}(k, p) \cdot \exp\left[j\left(\frac{2\omega_k}{c} X_{u\hat{v}} p(\Delta\theta) + \frac{2\omega_k}{c} Y_{u\hat{v}}\right)\right] \quad (32)$$

在补偿完卫星几何中心的二阶运动及径向位移之后, 可以通过二维傅里叶变换提取复成像函数。

$$\hat{a}_{u\hat{v}} = \sum_{p=1}^M \left[ \sum_{k=1}^K \hat{S}(k, p) \exp\left(j2\pi \frac{k\hat{v}}{K}\right) \right] \cdot \exp\left(j2\pi \frac{pu}{M}\right) \quad (33)$$

式中:  $\hat{v} = \frac{\Delta\omega}{\pi c} KY_{u\hat{v}}$  为第  $u\hat{v}$  个散射点在距离向上的位置坐标;  $\hat{u} = \frac{\omega_k(\Delta\theta)}{\pi c} MX_{u\hat{v}}$  为第  $u\hat{v}$  个散射点在方位向上的位置坐标。

本质上, 式(33)是一个自优化过程, 使位于离散坐标区间  $\hat{u} \in \{1, 2, \dots, M\}$ ,  $\hat{v} \in \{1, 2, \dots, K\}$  内的散射点强度值最大。脉内傅里叶变换实现距离压缩, 脉间傅里叶变换实现相位补偿。

### 3 ISAR 像处理

#### 3.1 ISAR 成像仿真

首先建立卫星的三维网格模型, 网格单元大小为  $\Delta X = 0.1 \text{ m}$ ,  $\Delta Y = 0.1 \text{ m}$ ,  $\Delta Z = 0.1 \text{ m}$ 。卫星在 788.9 km 的圆轨道运行, 轨道倾角为  $98.57^\circ$ , 升交点赤经为  $99.44^\circ$ , 近地点幅角为  $90^\circ$ 。ISAR 所在位置为东经  $119^\circ$ , 北纬  $29.5^\circ$ 。ISAR 发射线性调频信号照射卫星, 参数为: 载波频率  $f_c = 10 \text{ GHz}$ , 角频率  $\omega = 2\pi f_c$ , 周期  $T = 1/f_c$ , 波长  $\lambda = c/f_c$ , 频率带宽  $\Delta F = 1 \text{ GHz}$ , 脉冲宽度  $T_p = 1 \times$

$10^{-5}$  s, 调频率为  $\Delta F/T_p$ , 采样率  $f_s = 1 \times 10^7$  Hz, 采样周期  $\Delta T = 1/f_s$ 。利用式 (25) 生成卫星回波数据, 利用式 (33) 进行卫星图像提取, 并通过旁瓣抑制及相干斑滤波对图像进行初步处理, 得到卫星的 ISAR 仿真图像, 如图 4 所示。

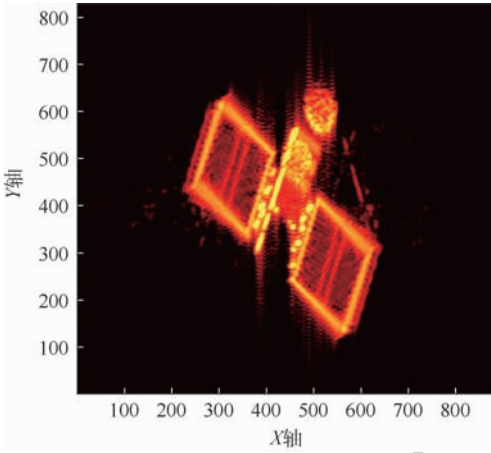


图 4 目标卫星的 ISAR 仿真图像

Fig. 4 Simulated ISAR image of target satellite

### 3.2 ISAR 像的旋转校正

实际过程中, 卫星在 ISAR 像中的位置取决于卫星的运动参数及其所处的空间位置, 这导致了卫星在 ISAR 像中的位置及指向是任意分布的, 增加了图像识别的难度。因此, 有必要对 ISAR 像进行旋转处理, 使目标卫星的纵轴 (最长轴) 与像平面的横轴平行。旋转校正首先需要将目标卫星从背景中分离, 其次检测出目标卫星最长轴, 最后将最长轴旋转至与像平面的横轴平行。

#### 3.2.1 ISAR 像的二值化处理

为了将卫星从背景中分离, 本文采用大津法 (Otsu) 进行最佳阈值确定, 该方法的原理是通过最大类间方差进行阈值的自适应确定。

将 ISAR 灰度图的灰度值大小表示为  $[0, 1, \dots, 255]$ , 且每个灰度值对应的像素个数为  $n_i$ , 总的像素个数为  $\mathcal{N} = n_0 + n_1 + \dots + n_{255}$ 。为了方便讨论, 将灰度直方图归一化, 转化成概率分布为

$$p_i = n_i / \mathcal{N} \quad p_i \geq 0, \quad \sum_{i=0}^{255} p_i = 1 \quad (34)$$

设灰度图阈值为  $\tau$ , 则图像的像素可以分为  $C_0$  与  $C_1$  两类 (卫星与背景),  $C_0$  表示灰度值大小为  $[0, 1, \dots, \tau]$  的像素,  $C_1$  表示灰度值大小为  $[\tau + 1, \tau + 2, \dots, 255]$  的像素。2 类像素的概率分布分别为

$$\varpi_0(\tau) = P(C_0) = \sum_{i=0}^{\tau} p_i \quad (35)$$

$$\varpi_1(\tau) = P(C_1) = \sum_{i=\tau+1}^{255} p_i \quad (36)$$

均值分别为

$$\mu_0(\tau) = \sum_{i=0}^{\tau} iP(i|C_0) = \sum_{i=0}^{\tau} ip_i / \varpi_0(\tau) \quad (37)$$

$$\mu_1(\tau) = \sum_{i=\tau+1}^{255} iP(i|C_1) = \sum_{i=\tau+1}^{255} ip_i / \varpi_1(\tau) \quad (38)$$

令灰度图的总平均值为  $\mu_T$ , 则有

$$\mu_T = \varpi_0(\tau)\mu_0 + \varpi_1(\tau)\mu_1 \quad (39)$$

式中:  $\varpi_0(\tau) + \varpi_1(\tau) = 1$ 。

$C_0$  与  $C_1$  之间的类间方差  $\sigma_B^2$  可以表示为

$$\sigma_B^2 = \varpi_0(\tau)(\mu_0(\tau) - \mu_T)^2 + \varpi_1(\tau)(\mu_1(\tau) - \mu_T)^2 \quad (40)$$

则最佳阈值  $\tau^*$  通过最大化类间方差  $\sigma_B^2$  得到, 即

$$\tau^* = \arg \max_{0 \leq i \leq 255} \sigma_B^2 \quad (41)$$

在算法实现过程中, 当  $\sigma_B^2$  取最大值时, 阈值  $\tau$  即为最佳阈值  $\tau^*$ , 从而将卫星从背景中分离, 得到图 5。

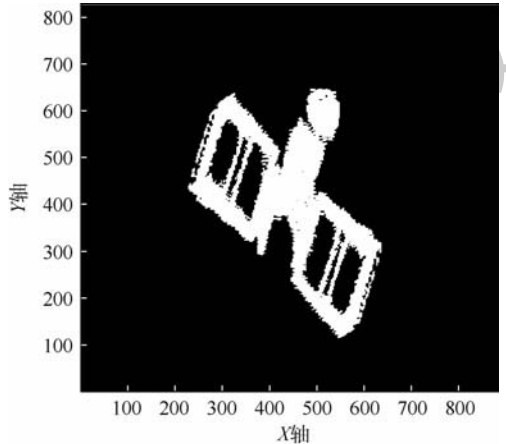


图 5 目标卫星 ISAR 像的二值化结果

Fig. 5 Binarization results of target satellite ISAR image

#### 3.2.2 卫星边缘检测

将卫星从背景中分离后, 为了对卫星进行旋转处理, 首先需要对卫星的边缘进行检测。边缘检测算法种类较多, 以 Roberts、Sobel、Kirsch、Prewitt 及高斯-拉普拉斯等为代表的边缘检测算子受噪声影响较大, 难以满足实际应用要求。因此, 本文采用最优边缘检测算法——canny 算子。canny 算子不易受噪声干扰, 同时能够检测出真正的弱边缘, 是目前应用最为广泛的边缘检测算子之一。

canny 算子在进行卫星边缘检测时, 主要分为以下 3 个步骤进行:

**步骤 1** 二维高斯平滑滤波。边缘检测算法对噪声十分敏感, 因此在进行 canny 算子边缘检

测前,有必要对图像进行去噪处理。通过采用二维高斯函数的一阶导数与图像进行卷积运算,从而得到滤掉高频噪声后的图像。其中,二维高斯函数的表达式为

$$G(x_i, y_i) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x_i^2 + y_i^2}{\sigma^2}\right) \quad (42)$$

式中:  $(x_i, y_i)$  为卫星 ISAR 图像中的每一点;  $\sigma$  由高斯滤波器的宽度决定,其取值大小影响图像的平滑程度。

$G(x_i, y_i)$  的一阶导数为

$$\nabla G = \left[ \frac{\partial G}{\partial x_i}, \frac{\partial G}{\partial y_i} \right] \quad (43)$$

在运算过程中,为了提高运算速度,将二维滤波器  $\nabla G$  分解成 2 个一维滤波器  $\frac{\partial G}{\partial x_i}$  与  $\frac{\partial G}{\partial y_i}$ , 分别表示为

$$\frac{\partial G}{\partial x_i} = \bar{\lambda} x_i \exp\left(-\frac{x_i^2}{2\sigma^2}\right) \exp\left(-\frac{y_i^2}{2\sigma^2}\right) \quad (44)$$

$$\frac{\partial G}{\partial y_i} = \bar{\lambda} y_i \exp\left(-\frac{x_i^2}{2\sigma^2}\right) \exp\left(-\frac{y_i^2}{2\sigma^2}\right) \quad (45)$$

式中:  $\bar{\lambda}$  为常量。

将 2 个一维滤波器  $\frac{\partial G}{\partial x_i}$  与  $\frac{\partial G}{\partial y_i}$  先后与图像  $f(x_i, y_i)$  进行卷积:

$$E_{x_i}(x_i, y_i) = \frac{\partial G}{\partial x_i} \otimes f(x_i, y_i) \quad (46)$$

$$I(x_i, y_i) = \frac{\partial G}{\partial y_i} \otimes E_{x_i}(x_i, y_i) \quad (47)$$

式中:  $I(x_i, y_i)$  为平滑滤波后所得图像。

**步骤 2** 梯度幅值与梯度方向计算,以及梯度幅值的非极大值抑制。利用  $2 \times 2$  邻域一阶偏导的有限差分方法,计算平滑滤波后图像  $I(x_i, y_i)$  的梯度幅值与梯度方向。将水平方向与垂直方向的偏导分别存储在  $P_{x_i}(i, j)$ 、 $P_{y_i}(i, j)$  2 个数组中。

$$P_{x_i}(i, j) = (I(i, j+1) - I(i, j) + I(i+1, j+1) - I(i+1, j))/2 \quad (48)$$

$$P_{y_i}(i, j) = (I(i, j) - I(i+1, j) + I(i, j+1) - I(i+1, j+1))/2 \quad (49)$$

利用式(48)、式(49),计算得到  $I(x_i, y_i)$  的梯度幅值  $M(i, j)$  与梯度方向  $H(i, j)$  分别为

$$M(i, j) = \sqrt{P_{x_i}^2(i, j) + P_{y_i}^2(i, j)} \quad (50)$$

$$H(i, j) = \arctan(P_{x_i}(i, j)/P_{y_i}(i, j)) \quad (51)$$

在计算过程中,若图像  $I(x_i, y_i)$  中像素点  $(i, j)$  的梯度幅值  $M(i, j)$  大于或等于沿  $H(i, j)$  梯度方向任意相邻 2 点像素之间的梯度幅值,则认为该点为可能的边缘点。

**步骤 3** 基于双阈值法的边缘确定。步骤 2 中确定了所有的可能边缘,为了进一步得到图像的真实边缘,采用高低双阈值方法,将所有边缘分为  $T_1(i, j)$  与  $T_2(i, j)$  两类。其中,  $T_1(i, j)$  由高阈值确定,不含假边缘,但有可能存在边缘的间断不连续。对此,当  $T_1(i, j)$  出现边缘间断现象时,在低阈值确定的  $T_2(i, j)$  中利用 8 邻域位置搜寻方法寻找可以连接到间断轮廓上的边缘,并通过递归跟踪算法,直至最终补充  $T_1(i, j)$  中所有间断边缘,从而实现图像的边缘检测。检测结果如图 6 所示。

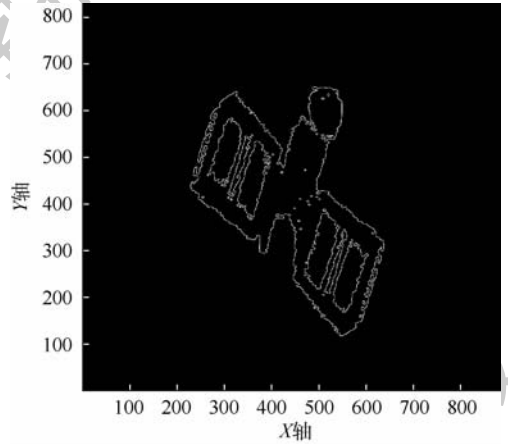


图 6 目标卫星 ISAR 像的边缘检测结果

Fig. 6 Edge detection results of target satellite ISAR images

### 3.2.3 旋转校正操作

实现卫星图像的边缘检测后,需要进一步对图像进行旋转校正。旋转校正需要确定卫星的纵轴(最长轴),对此本文采用 Hough 变换方法进行检测。

Hough 变换对图像局部缺损不敏感,且对噪声具有较强鲁棒性。其基于“点-线”对偶原理,将原有空间坐标系中的一条直线映射为参数空间中无数条曲线同时相交的一个点,进而将直线的检测问题转化为参数空间中点的检测问题,并通过参数空间中点的统计累加实现直线检测。具体过程如下:

基于方程  $\rho = x_i \cos \zeta + y_i \sin \zeta$  将图像中的每一点  $(x_i, y_i)$  映射为参数空间中的一条正弦曲线,  $\rho$  与  $\zeta$  为一对变量。遍历  $\zeta$  值,计算  $\rho$  值,并根据  $\zeta$  值与  $\rho$  值对二维累加器矩阵进行累加,得到共线点的个数。其中,  $\zeta$  的取值范围为  $[0^\circ, 180^\circ]$ ,  $\rho = \sqrt{x_i^2 + y_i^2} \sin(\zeta + \alpha)$ ,  $\alpha = \arctan(y_i/x_i)$ , 所以  $\rho$  的取值范围为  $\rho \in [-\sqrt{x_i^2 + y_i^2}, \sqrt{x_i^2 + y_i^2}]$ 。图 7 展示了将直角坐标系中的直线映射为参数空间中

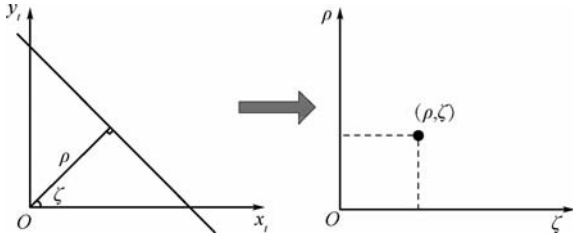


图7 直角坐标系中线到参数空间点的映射

Fig.7 Lines in a rectangular coordinate system mapped to points in parameter space

的一个点的示意图。图中： $\rho$ 表示原点到直线的垂直距离， $\zeta$ 表示垂线与 $x_i$ 轴的夹角。

利用 Hough 变换检测卫星纵轴(最长轴)的具体算法步骤如下：

**步骤1** 在 $\rho$ 与 $\zeta$ 合适的取值范围内建立一个离散化的参数空间。

**步骤2** 将参数空间量化为 $l_1 \times l_2$ 个单元( $l_1$ 为 $\zeta$ 的等分个数, $l_2$ 为 $\rho$ 的等分个数),并建立累加器矩阵。

**步骤3** 初始化 $l_1 \times l_2$ 个累加器,分配给参数空间中的每个单元。

**步骤4** 将卫星图像中的点 $(x_i, y_i)$ 代入方程 $\rho = x_i \cos \zeta + y_i \sin \zeta$ ,根据每个量化的 $\zeta$ 值计算 $\rho$ 值。

**步骤5** 将计算得到的 $\zeta$ 值与 $\rho$ 值代入相应的离散化参数空间中,并将对应位置的计数器值加1。

**步骤6** 当遍历完卫星图像中所有点后,检验离散空间中每个累加器的值,当累加器值为最大时,其所对应的 $\zeta$ 与 $\rho$ 值即为所求直线的参数,代入方程 $\rho = x_i \cos \zeta + y_i \sin \zeta$ 中即可求得的卫星的纵轴(最长轴)。

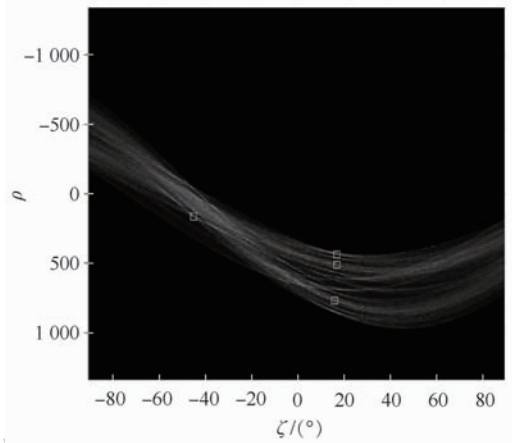
检测结果如图8所示。

根据所求 $\zeta$ 值,将卫星图像顺时针旋转 $\pi/2 - \zeta$ ,从而实现图像的纵轴(最长轴)与像平面的横轴平行,如图9所示。

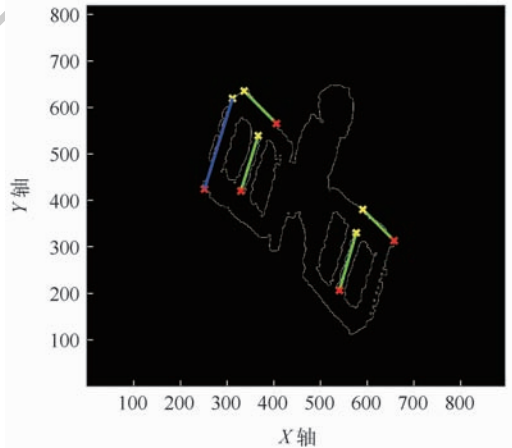
### 3.3 二值图像的闭运算

从卫星的二值图像中可以看到,卫星内部存在大量的孔洞,并且由于雷达脉冲旁瓣干扰的存在,卫星周围出现许多孤立的噪声区域。因此,为了提高图像质量,填充卫星内部孔洞,平滑卫星边界,且保持卫星形状、大小不变,提出利用二值形态学中的闭运算对3.2节中得到的卫星图像进行进一步处理。

二值形态学中的闭运算被定义为利用相同的“结构元素(structure element)”先后对二值图像进行膨胀运算与腐蚀运算的一种数学形态学计算



(a) 卫星边缘到参数空间的映射结果



(b) 卫星边缘的直线标识

图8 基于 Hough 变换的卫星边缘直线检测过程  
Fig.8 Satellite edge line detection process based on Hough transform

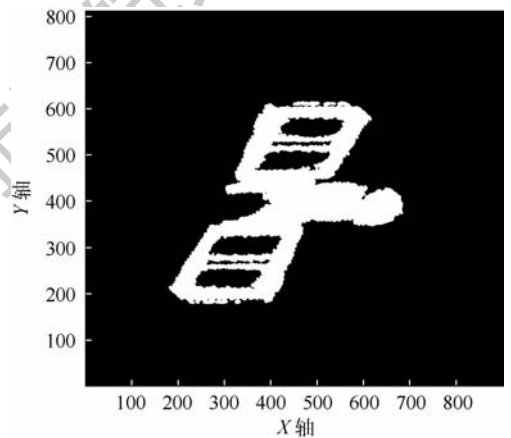


图9 目标卫星 ISAR 像的旋转校正结果  
Fig.9 Rotation correction results of target satellite ISAR images

子。其中,结构元素是二值形态学中的基本算子,通过与卫星图像中每个像素的周围区域进行特定的逻辑运算,得到每个像素位置新的像素值。像素值大小由结构元素的形状、大小及逻辑运算性

质决定。结构元素包括圆形、矩形、菱形、直线等多种形状,鉴于正方形结构元素具有各向同性且操作性较强的优点,采用正方形结构元素进行闭运算。

根据闭运算的操作顺序,先对卫星图像进行膨胀运算。膨胀运算可以填补卫星内部孔洞,并使卫星边界变得平滑。假设卫星图像为  $F$ , 结构元素为  $J$ , 令结构元素  $J$  在像平面上依次移动, 当经过每一个像素点时,  $J[s]$  有以下 3 种状态, 如图 10 所示: ①  $J[s_1] \subseteq F$ ; ②  $J[s_2] \subseteq F^c$ ; ③  $J[s_3] \cap F \neq \emptyset, J[s_3] \cap F^c \neq \emptyset$ 。

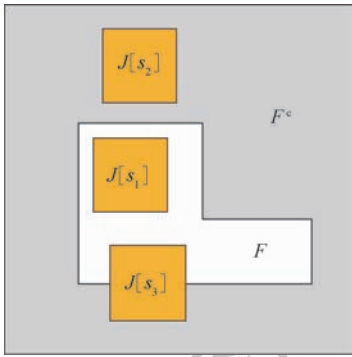


图 10  $J[s]$  的 3 种可能状态

Fig. 10 Three possible states of  $J[s]$

膨胀运算对应上述第 3 种状态,  $J[s]$  与  $F$  部分相关, 将  $F$  中的每一像素点扩展为结构元素  $J$ , 得到新的图像可以描述为

$$F \oplus J = \{s | J[s] \cap F \neq \emptyset\} \quad (52)$$

腐蚀运算可以去除孤立噪声区域, 同时恢复卫星尺寸大小。腐蚀运算对应上述第 1 种状态,  $J[s]$  与  $F$  最大相关, 通过腐蚀运算后的图像可以描述为

$$F \ominus J = \{s | J[s] \subseteq F\} \quad (53)$$

综上, 图像闭运算后得到的图像为

$$\text{CLOSE}(F) = (F \oplus J) \ominus J \quad (54)$$

闭运算算法步骤可以描述如下:

**步骤 1** 输入旋转后的二值图像, 大小为  $M_i \times N_i$ 。

**步骤 2** 选择正方形结构元素, 并设置结构元素大小为  $n_j \times n_j$ 。

**步骤 3** 膨胀运算。for  $i = 1$  to  $M_i$ ; for  $j = 1$  to  $N_i$ 。将  $(i, j)$  作为当前处理像素, 若当前像素值为 1, 直接操作下一个像素; 若当前像素值为 0, 判断周围  $n_j \times n_j - 1$  个像素是否全为 0, 若至少存在一个像素值为 1, 则将该像素值设置为 1, 否则操作下一个像素。

**步骤 4** 腐蚀运算。for  $i = 1$  to  $M_i$ ; for  $j = 1$  to

$N_i$ 。将  $(i, j)$  作为当前处理像素, 若当前像素值为 0, 直接操作下一个像素; 若当前像素值为 1, 判断周围  $n_j \times n_j - 1$  个像素是否全为 1, 若至少存在一个像素值为 0, 则将该像素值设置为 0, 否则操作下一个像素。

**步骤 5** 输出图像, 运算结束。

运算结果如图 11 所示。

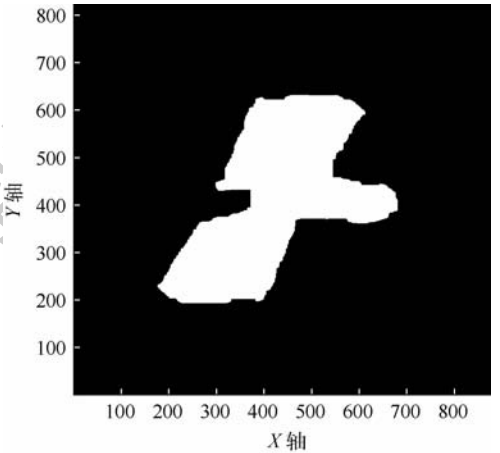


图 11 目标卫星二值图像的闭运算结果

Fig. 11 Closed operation results of target satellite binary image

### 3.4 卫星分割

在卫星识别过程中, 为了避免卫星在图像中出现位置不确定的状况, 同时提高识别速度, 本文利用形态学中的连通域思想, 将卫星所在区域分割出来。首先标记二值图中的所有连通域, 根据连通域几何特征, 求得表征卫星的最大连通域, 在此基础上, 计算连通域的几何中心, 并以该点为参考点分割出像素值大小为  $a \times a$  的子区域。具体的算法步骤如下:

**步骤 1** 输入二值图像  $F'(i, j)$ 。

**步骤 2** 采用 4 邻域网格进行逐像素扫描。若当前检测点  $(i, j)$  的像素值为 0, 则继续扫描下一个点。

**步骤 3** 若当前检测点的像素值为 1, 则继续检测其左侧点  $(i, j - 1)$  及上侧点  $(i - 1, j)$ , 如图 12 所示。并根据左侧点像素值  $I_{\text{Left}}$  及上侧点像素值  $I_{\text{Up}}$  的大小, 将检测结果分成 4 种情况: ①  $I_{\text{Left}}$  与  $I_{\text{Up}}$  均为 0, 则给该像素点一个新的标记 (新的连通域); ②  $I_{\text{Left}}$  与  $I_{\text{Up}}$  只有一个为 1, 则给该像素点与像素值为 1 的点相同的标记; ③  $I_{\text{Left}}$  与  $I_{\text{Up}}$  均为 1, 且具有相同的标记, 则给该像素点与其相同的标记; ④  $I_{\text{Left}}$  与  $I_{\text{Up}}$  均为 1, 但具有不同的标记, 将该像素点标记为较小的标记值。

**步骤 4** 继续回溯, 每次执行步骤 3 中的

①~④,直到回溯到连通域的开始元素。

**步骤5** 继续执行步骤2,直到二值图中所有的像素点被检测完。

**步骤6** 计算所有连通域的面积  $S_{connect}$  (连通域中的像素个数),保留最大连通域  $S_{MAX}$ ,计算  $S_{MAX}$  的几何中心  $(x_{tc}, y_{tc})$ 。

**步骤7** 定义矩阵  $F''_{a \times a}$ , 并为其赋值,令  $F''_{a \times a}(i - x_{tc} + \frac{a}{2} + 1, j - y_{tc} + \frac{a}{2} + 1) = F'(i, j)$ , 其中,  $i \in [x_{tc} - \frac{a}{2}, x_{tc} + \frac{a}{2}]$ ,  $j \in [y_{tc} - \frac{a}{2}, y_{tc} + \frac{a}{2}]$ , 遍历  $i, j$  所有值, 所得矩阵  $F''_{a \times a}$  即为卫星所在子区域图像, 从而实现卫星分割。

分割结果如图13所示。

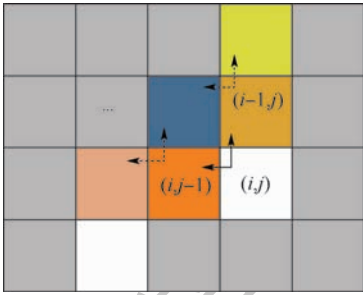


图12 邻域回溯法求连通域示意图

Fig. 12 Schematic diagram of neighborhood backtracking method to find connected domain

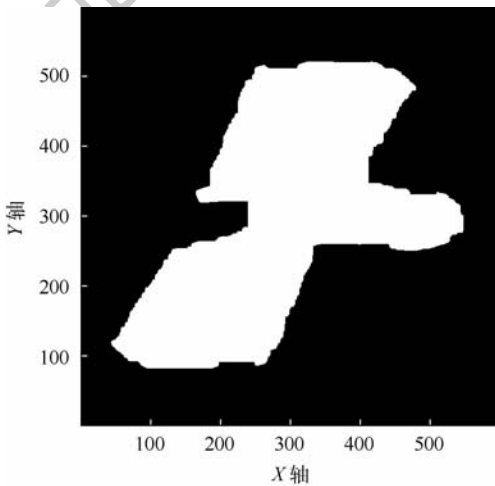


图13 目标卫星子区域分割结果

Fig. 13 Segmentation results of target satellite subregion

### 3.5 卫星轮廓提取

卫星轮廓是图像识别过程中最主要的特征之一,为此,继续对3.4节中生成的图像进行卫星轮廓提取。本文轮廓提取算法可以描述为

$$F''_{contour}(i, j) = \text{sign}\left[F''_{a \times a}(i, j) \text{floor}\left(\frac{3}{C}\right)\right] =$$

$$\begin{cases} 1 & F''_{a \times a}(i, j) \text{floor}\left(\frac{3}{C}\right) \geq 1 \\ 0 & F''_{a \times a}(i, j) \text{floor}\left(\frac{3}{C}\right) < 1 \end{cases} \quad (55)$$

式中:

$$C = \begin{bmatrix} F''_{a \times a}(i-1, j-1) & F''_{a \times a}(i-1, j) & F''_{a \times a}(i-1, j+1) \\ F''_{a \times a}(i, j-1) & F''_{a \times a}(i, j) & F''_{a \times a}(i, j+1) \\ F''_{a \times a}(i+1, j-1) & F''_{a \times a}(i+1, j) & F''_{a \times a}(i+1, j+1) \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (56)$$

其中:  $F''_{contour}(i, j)$  为提取的卫星轮廓; sign 为符号函数; floor 为向0取整函数; C 为像素点  $(i, j)$  在水平与垂直方向上4个相邻像素点的强度值加和。

卫星的轮廓提取结果如图14所示。可以看到,卫星轮廓线能够较好地勾勒出卫星的外形结构,该特征对于进一步展开卫星的识别工作具有重要意义。

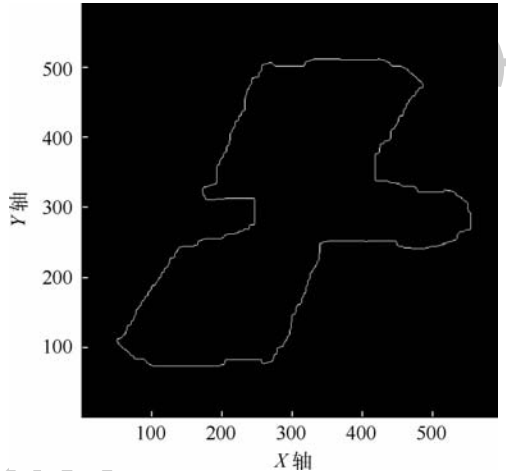


图14 目标卫星轮廓提取结果

Fig. 14 Extraction results of target satellite contour

## 4 结论

本文基于在轨卫星的三维几何结构建立了空间目标的 ISAR 成像模型,经过坐标转换、卫星运动轨迹分解、单位合成孔径时间内的脉冲周期数求解等步骤,建立了 ISAR 成像模型及图像函数提取模型,经过噪声抑制等初步处理得到了卫星的 ISAR 像;从 ISAR 像的自身特征出发,设计了空间目标 ISAR 像处理算法及轮廓提取算法,建立了较为连续的空间目标轮廓特征提取模型,计算结果显示,提取出的目标卫星的轮廓平滑且连续,为进一步展开空间目标识别奠定重要基础。

本文设计的图像处理算法的优点主要体现在

如下方面:

1) 算法结构简单,程序实现容易,对于受噪声干扰及轮廓线受损的 ISAR 像具有较好的处理效果。

2) 图像处理的每个步骤均具有较高的处理效率,处理后的 ISAR 像结构清晰,旋转校正效果良好,较好地填充了卫星内部孔洞,并去除了目标周围孤立的噪声区域。

3) 算法结构灵活,对于不同形状、帧中位置任意分布的 ISAR 像均可利用该算法进行处理。

### 参考文献 (References)

- [1] TOUMI A, HOELTZENER B, KHENCHAF A. Hierarchical segmentation on ISAR image for target recognition[J]. International Journal of Computational Intelligence Research, 2009, 5(1): 63-71.
- [2] BENEDETTO F, FULGINEI F, RLAUDANI A, et al. Automatic aircraft target recognition by ISAR image processing based on neural classifier[J]. International Journal of Advanced Computer Science and Applications, 2012, 3(8): 96-103.
- [3] ASGHARI M H, JALALI B. Edge detection in digital images using dispersive phase stretch transform[J]. International Journal of Biomedical Imaging, 2015, 2015: 687819.
- [4] STANKOVIC L. ISAR image analysis and recovery with unavailable or heavily corrupted data[J]. IEEE Transactions on Aerospace and Electronic Systems, 2015, 51(3): 2093-2106.
- [5] LEE S J, BAE J H, KANG B S, et al. Classification of ISAR images using sparse recovery algorithms[C]// IEEE Conference on Antenna Measurements & Applications. Piscataway, NJ: IEEE Press, 2014: 1-4.
- [6] 唐宁, 高勋章, 黎湘. ISAR 像自动识别中的预处理算法[J]. 系统工程与电子技术, 2011, 33(9): 2002-2006.  
TANG N, GAO X Z, LI X. Preprocessing algorithm for ISAR image automatic recognition[J]. Systems Engineering and Electronics Technology, 2011, 33(9): 2002-2006 (in Chinese).
- [7] MUSMAN S, KERR D, BACHMANN C. Automatic recognition of ISAR ship images[J]. IEEE Transactions of Aerospace and Electronic Systems, 1996, 32(4): 1392-1404.
- [8] BACHMANN C M, MUSMAN S A, SCHULTZ A. Classification of simulated radar imagery using lateral inhibition neural networks[C]// Neural Networks for Signal Processing. Piscataway, NJ: IEEE Press, 1992: 279-288.
- [9] ZELJKOVIC V, LI Q, VINCELETTE R, et al. Automatic algorithm for inverse synthetic aperture radar images recognition and classification[J]. IET Radar Sonar Navigation, 2010, 4(1): 96-109.
- [10] JDEY I, TOUMI A, DHIBI M, et al. The contribution of fusion techniques in the recognition systems of radar targets[C]// IET International Conference on Radar Systems. London: IET, 2012: 1-5.
- [11] 唐宁, 高勋章, 黎湘. 基于几何散列法的 ISAR 像自动目标识别[J]. 系统工程与电子技术, 2012, 34(4): 692-697.  
TANG N, GAO X Z, LI X. Automatic target recognition of ISAR images based on geometric hash[J]. Systems Engineering and Electronic Technology, 2012, 34(4): 692-697 (in Chinese).
- [12] TANG N, GAO X Z, LI X. Target classification of ISAR images based on feature space optimization of local non-negative matrix factorisation[J]. IET Signal Processing, 2012, 6(5): 494-502.
- [13] 李飞. 雷达图像目标特征提取方法研究[D]. 西安: 西安电子科技大学, 2014: 25-100.  
LI F. Study on target feature extraction based on radar image [D]. Xi'an: Xidian University, 2014: 25-100 (in Chinese).
- [14] 王欣. 基于酉 ESPRIT 的超分辨 ISAR 成像及其应用[D]. 西安: 西安电子科技大学, 2015: 31-74.  
WANG X. Unitary ESPRIT based on superresolution ISAR imaging and its application [D]. Xi'an: Xidian university, 2015: 31-74 (in Chinese).
- [15] 王芳, 盛卫星, 马晓峰, 等. 基于直方图统计量的逆合成孔径雷达目标识别[J]. 电波科学学报, 2012, 27(4): 92-98.  
WANG F, SHENG W X, MA X F, et al. Target clarification for ISAR image based on histogram statistics[J]. Journal of Radio Science, 2012, 27(4): 92-98 (in Chinese).
- [16] 王芳, 盛卫星, 马晓峰, 等. 基于 B(2D)2PGNMF 的 ISAR 像目标识别[J]. 南京理工大学学报(自然科学版), 2013, 37(6): 863-868.  
WANG F, SHENG W X, MA X F, et al. ISAR image target recognition based on B(2D)2PGNMF[J]. Journal of Nanjing University of Science and Technology (Natural Science), 2013, 37(6): 863-868 (in Chinese).
- [17] 王芳, 盛卫星, 马晓峰, 等. 基于 Gabor 幅值特征和相位特征相融合的 ISAR 像目标识别[J]. 电子与信息学报, 2013, 35(8): 1813-1819.  
WANG F, SHENG W X, MA X F, et al. ISAR image recognition with fusion of Gabor magnitude and phase feature[J]. Journal of Electronics & Information Technology, 2013, 35(8): 1813-1819 (in Chinese).

### 作者简介:

杨虹 女, 博士研究生。主要研究方向: 航天任务分析与设计。

张雅声 女, 教授, 博士生导师。主要研究方向: 航天任务分析与设计。

尹灿斌 男, 讲师。主要研究方向: 雷达成像。

## ISAR imaging and contour feature extraction of space targets

YANG Hong, ZHANG Yasheng\*, YIN Canbin

(Space Engineering University, Beijing 101416, China)

**Abstract:** The inverse synthetic aperture radar (ISAR) imaging of the space target is affected by target's own occlusion and noise interference, which makes the generated ISAR image difficult to be directly used for image analysis and target recognition. Therefore, the ISAR image processing and feature extraction are studied based on the ISAR imaging model of the space target. Firstly, the ISAR imaging model, ISAR signal model and ISAR image extraction model of the target satellite are established respectively, and the ISAR image of the target satellite is obtained through preliminary processing such as sidelobe suppression and speckle filtering. Secondly, the Otsu algorithm, the canny operator and the Hough transform are used to rotate the satellite to the longest axis parallel to the horizontal axis of the image plane. The closed operation method is used to fill the internal cavity of the satellite and remove the external isolated noise region. Based on the connected domain idea, the sub-area where the satellite is located is segmented, and the contour extraction of the satellite is realized. The designed image processing algorithm can effectively improve the ISAR image quality, and the extracted satellite contour can well outline the shape of the target satellite, which lays an important foundation for further satellite identification.

**Keywords:** space target; inverse synthetic aperture radar (ISAR); ISAR imaging; image processing; contour feature

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

DOI: 10.13700/j.bh.1001-5965.2018.0715

# 舰载机安全逃逸复飞的参数适配包线

林佳铭<sup>1</sup>, 吴光辉<sup>1</sup>, 王立新<sup>1,\*</sup>, 刘海良<sup>1</sup>, 王云<sup>2</sup>

(1. 北京航空航天大学 航空科学与工程学院, 北京 100083; 2. 中国舰船研究设计中心, 武汉 430064)

**摘 要:** 逃逸复飞特性对于确定舰载机的进舰速度、航母航速与着舰海况等级限制等均具有重要的影响。根据逃逸复飞的任务要求,提出了离舰滑跑距离、逃逸复飞段最大迎角相对进舰配平迎角的增量和着舰时的俯仰角作为安全性评定的3个参数。建立了数字飞行员模型,基于数字虚拟飞行仿真计算的方法,获得了满足逃逸复飞安全要求的参数适配包线。研究表明:在每一确定的着舰重量下,均对应存在一个最佳的进舰速度范围,能够使舰载机逃逸复飞的离舰滑跑距离最短;逃逸复飞段最大迎角相对进舰配平迎角的增量、着舰时舰载机的俯仰角分别决定了进舰速度大小的上、下边界;舰载机重心位置变化将使参数适配包线的边界位置和范围发生改变,但边界形状基本不变;航母航速减小将显著地缩小参数适配包线;适配包线内每一点的逃逸复飞成功率,可以为不同海况等级下的逃逸复飞安全性判断提供参考。

**关键词:** 舰载机; 逃逸复飞; 适配包线; 仿真模型; 甲板运动

**中图分类号:** V212.1

**文献标识码:** A **文章编号:** 1001-5965(2019)09-1777-10

舰载机着舰时,若尾钩未能勾住阻拦索,则需要在触舰后加大推力并进行俯仰操纵,经过三轮滑跑、抬前轮和两轮滑跑等一系列飞行阶段后,最终安全地飞离航母舰面,这一过程称为逃逸复飞<sup>[1-2]</sup>。逃逸复飞关系到舰载机的着舰安全性:一方面,逃逸复飞任务是进场着舰规范规定的评估内容之一,可用于预估舰载机合适的进舰速度范围<sup>[3-4]</sup>;另一方面,在开展机舰适配性研究时,也需要充分考虑着舰甲板风、航母运动等着舰环境特点<sup>[5]</sup>,确定满足逃逸复飞安全要求的作战使用条件,如航母航速要求、着舰海况等级限制等。

国外主要通过飞行员在环的地面模拟仿真试验和飞行试验的方法,来评估舰载机的逃逸复飞特性和飞行品质。目前可查阅到的公开文献十分有限,且未给出具体的安全性评定指标<sup>[2]</sup>。国内的研究主要是通过运动建模与仿真计算分析舰载机的逃逸复飞特性<sup>[6-9]</sup>,分别讨论了甲板运动、舰

面气流扰动、螺旋桨动力等因素对舰载机逃逸复飞特性的影响<sup>[10-11]</sup>。然而,目前较少见有针对逃逸复飞任务中机舰总体参数适配关系的研究。

在本文中,首先,针对逃逸复飞的任务要求,提出了离舰滑跑距离、逃逸复飞段最大迎角相对进舰配平迎角的增量(下文简称迎角增量)和着舰时的俯仰角3个安全性评定参数,用于研究和分析舰载机的逃逸复飞特性。其次,建立了数字飞行员模型用于开展数字虚拟飞行仿真计算,从而可以更准确地完成舰载机逃逸复飞的全过程数学仿真。在此基础上,计算分析了舰载机在不同着舰重量、进舰速度下的逃逸复飞特性,并获得了满足逃逸复飞安全要求的参数适配包线。最后,定量分析了舰载机重心位置、航母航速、不同海况等级等对参数适配包线的影响。研究结果可以为舰载机总体方案设计、机舰适配性研究、作战使用条件的确定等提供参考。

收稿日期: 2018-12-03; 录用日期: 2019-02-17; 网络出版时间: 2019-02-26 16:57

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190226.0946.002.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190226.0946.002.html)

\* 通信作者。E-mail: wlx\_c818@163.com

**引用格式:** 林佳铭, 吴光辉, 王立新, 等. 舰载机安全逃逸复飞的参数适配包线[J]. 北京航空航天大学学报, 2019, 45(9): 1777-1786. LIN J M, WU G H, WANG L X, et al. Parameter suitability envelope for safety bolter of a carrier-based aircraft[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1777-1786 (in Chinese).

# 1 逃逸复飞的安全性评定参数

美军联合军种规范指南(JSSG—2001A)的附录D给出了舰载机逃逸复飞任务的明确定义,并简述了逃逸复飞的飞行任务与操纵策略<sup>[1]</sup>。此外,国军标GJB 2753—96中也规定了舰载机着舰载荷的具体要求<sup>[12]</sup>,对于保证逃逸复飞时飞机起落架的结构强度安全具有指导意义。因此,本文基于上述文献,提出以下3个参数,用于评定舰载机逃逸复飞的安全性。

1) 舰载机从三轮着舰到逃逸复飞拉起这一过程的离舰滑跑距离( $X_{bo}$ ,见图1),是JSSG—2001A<sup>[1]</sup>明确指出可以表明舰载机逃逸复飞性能的一项参数,其能够反映飞机短距起飞的滑跑特性,以及机舰同向运动时相互位置变化的情况。逃逸复飞过程与正常舰面起飞不同,由于在此期间飞行员需进行一定的拉杆操纵,当飞机在到达甲板末端前就离舰爬升,不出现航迹下沉才是理想的逃逸复飞<sup>[3]</sup>。由此可知,通过计算使飞机到达甲板末端前就离舰爬升的逃逸复飞的滑跑距离,并与航母所能提供的甲板跑道长度进行比较,就能够判断逃逸复飞是否安全。

2) JSSG—2001A<sup>[1]</sup>指出舰载机在逃逸复飞过程中,其迎角应保持在 $\alpha_{pa}$ 和 $\alpha_{pa} + 3^\circ$ 之间( $\alpha_{pa}$ 为进舰配平迎角),且升力系数 $C_L$ 不大于 $0.9C_{Lmax}$ 。因此,本文提出逃逸复飞段最大迎角相对进舰配平迎角的增量( $\Delta\alpha_{bo}$ ,见图1)这一参数作为安全性评定参数,并且要求迎角增量不大于 $3^\circ$ ,原因如下:①舰载机在进舰下滑到逃逸复飞的全过程中,均处于低速大迎角飞行状态,此时既要保证飞机以足够大的迎角飞行,也要保证其在逃逸复飞段不出现因迎角过大而失速的危险,因此采用迎角增量这一参数对飞机的飞行迎角进行约束,从而保证其在合理、安全的范围内变化;②迎角过大还可能诱发飞机的横航向运动与操纵效能下降等问题,在逃逸复飞段会给飞行员的操纵造成不利影响;③在舰载机下滑变轨机动任务的规范要求中<sup>[1]</sup>,也对飞机迎角增量提出了类似

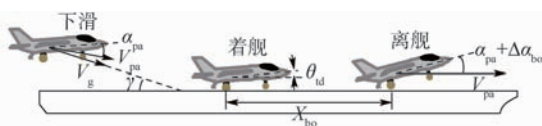


图1 逃逸复飞过程及安全性评定参数示意

Fig. 1 Illustration of bolter process and its safety evaluation parameters

的准则要求,即迎角增量不超过失速迎角与机动初始迎角两者之差的50%。

3) 基于飞机纵向地面操稳特性的要求,本文提出采用飞机着舰时的俯仰角( $\theta_{id}$ ,见图1)作为逃逸复飞安全性的评定参数,其原因在于:如果着舰时的俯仰角过小,飞机就可能出现着舰回跳、前起落架结构损坏等险情,因此需要限制飞机着舰时的俯仰角大小。本文确定 $\theta_{id}$ 的设计要求时,主要考虑了飞机的停机角和航母甲板运动的纵摇角,并且还预留了 $0.5^\circ$ 作为安全裕度。

图1表征了舰载机的逃逸复飞过程,并且简要地示意了一些主要的物理参数量,其中, $V_{pa}$ 为飞机进舰空速, $V_g$ 为飞机进舰地速, $\gamma$ 为飞机航迹倾角。

# 2 逃逸复飞的参数适配包线

逃逸复飞是典型的高风险试飞科目,飞行员需要具有较丰富的飞行经验和较高的操纵技术才能完成这一科目。本文采用数字虚拟飞行仿真计算的方法,从而较准确地分析了舰载机的逃逸复飞特性,并且计算相应的参数适配包线。

## 2.1 数字虚拟飞行仿真计算模型

数字虚拟飞行仿真计算方法<sup>[13]</sup>是指在设计阶段,由工程技术人员对飞机飞行动力学和飞行员操纵行为进行数学建模,并利用数字计算机对选定的考核飞行任务开展人机闭环仿真计算,依据所得结果考察飞机飞行和运营的安全性是否能够满足设计的要求。数字虚拟飞行仿真计算模型的完整结构如图2所示。

### 1) 舰载机运动模型

飞机的6自由度动力学方程如下:

$$\begin{cases} T_{IS} F_{aero}^S + T_{IB} (F_{eng}^B + F_{lg}^B) + mg^I = m[\ddot{x}, \ddot{y}, \ddot{z}]^T \\ T_{BS} M_{aero}^S + M_{eng}^B + M_{lg}^B = I[\dot{p}, \dot{q}, \dot{r}]^T \end{cases} \quad (1)$$

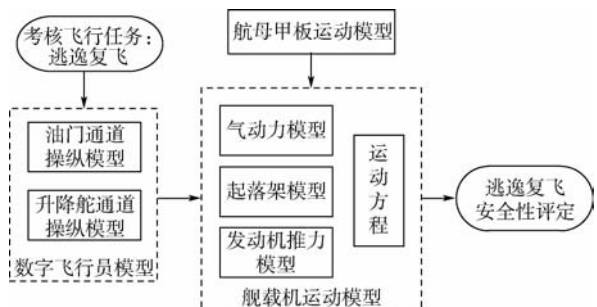


图2 数字虚拟飞行仿真计算模型结构

Fig. 2 Structure of digital virtual flight simulation model

式中: $F$ 和 $M$ 分别为力和力矩向量,其上标表示该力或力矩投影到某一坐标系下的分量(S、B分别指稳定轴系、机体轴系),其下标 aero、eng、lg 分别指气动力、推力、起落架作用力; $T_{IS}$ 、 $T_{IB}$ 、 $T_{BS}$ 为各个坐标系间的转换矩阵; $m$ 为飞机的着舰重量; $I$ 为飞机的转动惯量矩阵; $g$ 为重力加速度向量,上标 I 表示惯性参考系; $x$ 、 $y$ 、 $z$ 分别为飞机重心的 3 个坐标位置; $p$ 、 $q$ 、 $r$ 分别为飞机的 3 轴角速度。

通过建立起落架模型,可以计算前起与主起对于机体的作用力和力矩,从而准确地反映舰载机在触舰、三轮滑跑、抬前轮等一系列逃逸复飞过程中的动态响应特性。起落架支柱的缓冲力一般分为空气弹簧力和油液阻尼力,可由支柱压缩量及其变化率计算获得<sup>[14]</sup>。

舰载机的油门与升降舵作动器的动力学模型采用一阶环节表示,其相应的时间常数分别为 0.7 s 和 0.12 s。

### 2) 数字飞行员模型

数字飞行员模型是指模拟飞行员在逃逸复飞任务中的操纵行为,并给出操纵指令的数学模型,包括油门与升降舵 2 个通道,从而完成人机闭环数学仿真。

油门通道的操纵模型如式(2)所示。在飞机触舰前,油门指令 $\delta_{pc}$ 保持下滑配平时的指令 $\delta_{p,trim}$ 不变;在飞机触舰后,需要考虑飞行员的反应时间延迟,一般取为 0.5 s,在此期间油门指令仍然不变;在此之后,飞行员将给出最大油门指令 $\delta_{p,max}$ ,从而加大发动机推力进行逃逸复飞。 $t_{td}$ 为飞机触舰时刻。

$$\delta_{pc} = \begin{cases} \delta_{p,trim} & t < t_{td} + 0.5 \\ \delta_{p,max} & t \geq t_{td} + 0.5 \end{cases} \quad (2)$$

升降舵通道的操纵模型描述如下:在飞机触舰前,升降舵指令保持下滑配平时的指令不变;在飞机触舰后,飞行员的操纵反应时间延迟约为 1 s,在此期间升降舵指令同样不变;在此之后,要求飞行员通过操纵使飞机保持较为理想的离舰爬升迎角。对于技术较为熟练的飞行员,这一迎角误差补偿任务可以采用 McRuer 模型进行描述<sup>[15]</sup>:

$$H_{\alpha}(s) = k_{\alpha} \frac{1}{T_{\alpha}s + 1} e^{-\tau_{\alpha}s} \quad (3)$$

式中: $k_{\alpha}$ 为飞行员对迎角误差的操纵增益; $T_{\alpha}$ 为飞行员滞后补偿时间常数; $\tau_{\alpha}$ 为飞行员的信息反应延迟时间; $s$ 为复变量。

此外,在操纵模型中还需要加入俯仰角速度反馈,从而表征飞行员对角速度的感受和操纵<sup>[16-17]</sup>。图 3 给出了完整的升降舵通道飞行员

操纵模型。图中: $H_q(s)$ 为俯仰角速度感受回路; $H_{mm}(s)$ 为人体肌肉动作延迟模型,可用频率为 9 rad/s、阻尼比为 0.7 的二阶环节描述; $\alpha_{err}$ 为迎角误差; $\delta_{ec}$ 为升降舵指令。

$$H_q(s) = k_q e^{-\tau_q s} \quad (4)$$

式中: $k_q$ 为飞行员的运动感官增益; $\tau_q$ 为飞行员运动神经的延迟时间。

### 3) 航母甲板运动模型

航母的甲板运动较为复杂,既与航母自身的特征长度、排水量、航速航向等相关,也受到海浪和海面风等海况条件的影响,往往难以获得准确的解析模型,因此在工程研究时,多采用频谱模型或正弦模型加以描述<sup>[18]</sup>。本文采用简化和统一的正弦模型<sup>[19]</sup>对航母的甲板运动进行建模,如式(5)所示,主要考虑了垂荡、纵摇、横摇等 3 个运动,不同海况下的建模数据如表 1 所示。通过随机设置模型的初始相位,可以获得多组不同航母甲板运动的时域仿真曲线。

$$X_c = A_{X_c} \sin(\omega_{X_c} t + \varphi_{X_c}) \quad (5)$$

式中: $X_c$ 指航母的垂荡位移 $\Delta z_c$ 、纵摇角 $\theta_c$ 和横摇角 $\phi_c$ ,单位分别为 m、(°)和(°); $A$ 为甲板运动的幅值,单位为 m 或(°); $\omega$ 和 $\varphi$ 分别为甲板运动的频率和初始相位,单位分别为 rad/s、rad。

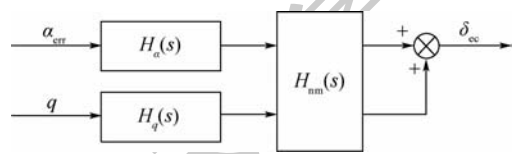


图 3 升降舵通道的飞行员操纵模型

Fig. 3 Control model of pilot in elevator channel

表 1 不同海况等级下的甲板运动建模数据

Table 1 Modeling data of deck motion under different sea states

参数	4 级海况	6 级海况
垂荡位移幅值/m	0.68	1.63
横摇角幅值/(°)	0.62	1.45
纵摇角幅值/(°)	0.52	1.24
垂荡频率/(rad · s <sup>-1</sup> )	0.35	0.35
横摇频率/(rad · s <sup>-1</sup> )	0.29	0.29
纵摇频率/(rad · s <sup>-1</sup> )	0.52	0.52

## 2.2 仿真算例

采用某舰载机的总体与气动数据来建立本文算例飞机的运动模型。飞机的典型着舰重量约为 10 × 10<sup>3</sup> kg,机翼面积约为 55 m,翼展约为 14 m,弦长约为 5.4 m,单发最大推力约为 50 kN,进舰下滑时的航迹倾角为 -3.5°。算例航母参考“尼米兹”级航母进行运动建模<sup>[19]</sup>,其斜角甲板跑道长

度约为 235 m, 典型航速约为 20 kn。

根据算例飞机与航母的有关参数, 确定逃逸复飞的安全准则如下: 按航母最后一根阻拦索到斜角甲板末端的跑道长度, 取  $X_{bo} < 155$  m; 基于 JSSG—2001A<sup>[1]</sup> 中的要求, 取  $\Delta\alpha_{bo} < 3^\circ$ ; 飞机的停机角约为  $2^\circ$ , 航母的最大纵摇角约为  $1.5^\circ$ , 再加上  $0.5^\circ$  作为安全裕度后, 取  $\theta_{id} > 4^\circ$ 。

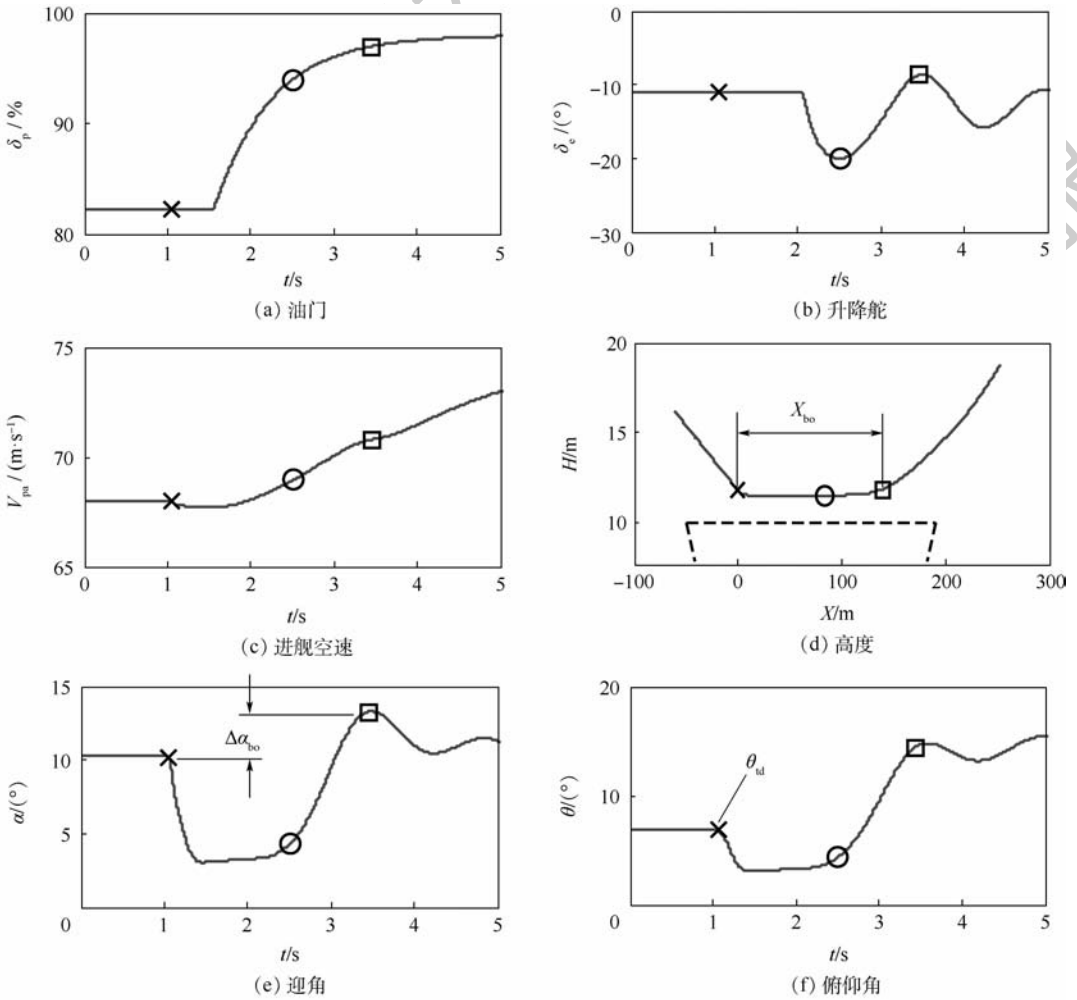
图 4 给出了一组逃逸复飞的仿真计算结果。此外, 离舰滑跑距离、迎角增量、着舰时的俯仰角等 3 个安全性评定参数也在对应的仿真曲线中作出了标识。

根据图 4 可知:

1) 仿真计算开始时, 舰载机位于舰尾上方, 净空高度约为 4.7 m, 在飞机触舰前飞行员保持油门和升降舵位于初始的配平下滑状态。仿真计

算 1 s 时舰载机触舰, 通常情况下飞行员做出逃逸复飞的决策判断时间约为 0.5 ~ 1 s, 因此飞行员在 1.5 s 左右给出了油门操纵指令, 在 2 s 左右给出了升降舵操纵指令。由于发动机推力变化的时间延迟, 故直至第 4 s 时发动机才达到最大推力状态。开始滑跑后, 飞机的迎角和俯仰角均减小, 因此飞行员首先给出纵向拉杆操纵指令。飞机抬前轮后开始两轮滑跑, 迎角和俯仰角有所增大, 此时飞行员为使飞机具有适当的飞离姿态, 减小了升降舵的拉杆量, 从而避免飞机过快抬头, 如图 4(f) 所示。

2) 舰载机着舰时的俯仰角约为  $6.8^\circ$ , 因此不会出现前起落架先触舰的情况; 迎角增量约为  $2.9^\circ$ ; 飞机的离舰滑跑距离约为 139 m, 故飞机能够在到达甲板末端前飞离舰面开始爬升。因此, 本次逃逸复飞是安全的。



× 飞机触舰 ○ 飞机抬前轮 □ 飞机离舰

图 4 逃逸复飞的仿真算例

Fig. 4 A simulation sample of bolter

### 2.3 评定参数的变化规律

以进舰速度为横坐标, 着舰重量为纵坐标, 重

复 2.1 节的数字虚拟飞行仿真计算, 可以获得舰载机不同着舰重量和进舰速度下的逃逸复飞性

能,并以图5所示的等高线形式进行描述。图5能够简洁、直观地表示着舰重量、进舰速度对于舰载机逃逸复飞特性的影响,从而为确定合适的设计着舰重量、进舰速度提供参考。

分析图5可知:

1) 在所有的等高线图中,其左侧直线段为下滑配平约束限制,即在该线段以左、以上的着舰重量、进舰速度组合下,舰载机无法建立稳定的下滑着舰飞行状态,因而也无法进一步分析其逃逸复飞特性。

2) 舰载机的着舰重量增大,则其在逃逸复飞时的离舰滑跑距离就随之增加。而在同一着舰重量下,离舰滑跑距离将随进舰速度的增加呈现出先减小后增大的特点,其原因在于:①增大进舰速度能够使舰载机更快地获得爬升拉起所需的升力,从而完成逃逸复飞并离舰,因此离舰滑跑距离有所缩短;②由于飞行员需要一定的反应和决策时间才能给出逃逸复飞操纵指令,因此在相同的滑跑时间内,离舰滑跑距离将随机舰相对速度增大而相应增加。

进一步,根据图5(a)所示的仿真结果还可以

看出:当着舰重量为  $10 \times 10^3$  kg 时,较为理想的进舰速度约为  $66 \sim 70$  m/s,此时的离舰滑跑距离最短,约为 140 m。而当着舰重量为  $12.5 \times 10^3$  kg 时,最短的离舰滑跑距离为 155 m,对应的进舰速度范围为  $73 \sim 77$  m/s。由此也就表明,在每一着舰重量下,均对应存在一个最佳的进舰速度范围,从而能够使逃逸复飞的离舰滑跑距离最短。

3) 如图5(b)所示,迎角增量的等高线近似于一簇相互平行的直线段,该值越小表明舰载机的俯仰姿态改变量就越小,因而其逃逸复飞特性也就越好。基于迎角增量的约束要求,可以确定各着舰重量下飞机的最小进舰速度。

4) 舰载机的着舰下滑角一般取为固定值  $-3.5^\circ$ ,因此飞机着舰时的俯仰角与下滑配平时的迎角一一对应,变化规律较为明确,即进舰速度越大、着舰重量越轻,则着舰时的俯仰角就越小,如图5(c)所示。从逃逸复飞安全性角度出发,为了避免前起落架先触舰的情况发生,应当对着舰时俯仰角的下限值进行约束,因此也就能够确定各着舰重量下飞机的最大进舰速度。

## 2.4 参数适配包线及其计算方法

本文所提出的参数适配包线,指的是满足一定舰载机飞行状态(如下滑角)、航母运动状态(如航母航速)和逃逸复飞安全性要求的舰载机着舰重量、进舰速度的组合情况,因而可以为机舰适配性研究和作战使用条件确定提供量化的评估参考。

参数适配包线的计算方法如下:

**步骤1** 按照2.2节和2.3节的方法,通过数字虚拟飞行仿真计算,获得舰载机在不同着舰重量和进舰速度下逃逸复飞性能评定参数的等高线图。

**步骤2** 对于每一项逃逸复飞性能的评定参数,确定其安全性判定标准,从而在性能等高线图中求出对应的边界约束。

**步骤3** 求取3条边界约束的交集后,就能够得到满足全部逃逸复飞安全要求的参数适配包线,如图6所示。图中:AB段为迎角增量的约束,

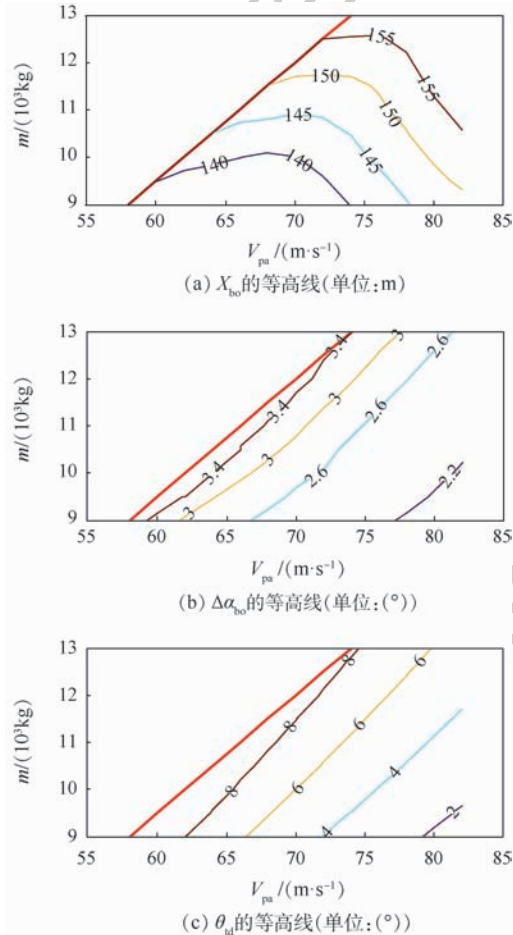


图5 不同着舰重量和进舰速度下的逃逸复飞性能

Fig. 5 Bolter performance with different landing weights and approach velocities

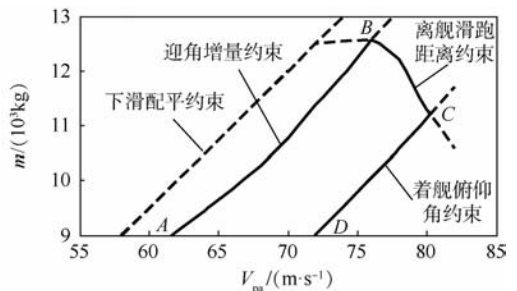


图6 逃逸复飞的参数适配包线

Fig. 6 Parameter suitability envelope for bolter

其决定了参数适配包线的左边界;BC段为离舰滑跑距离的约束,其决定了参数适配包线的上边界;CD段为着舰时俯仰角的约束,其决定了参数适配包线的右边界。

### 3 参数适配包线的影响因素

1) 舰载机在实际执行作战任务时,其燃油消耗和武器装载使用等均存在一定的不确定性,故其着舰时的重心位置等都可能偏差理想设计状态。重心位置的前后变化将直接影响飞机的纵向力矩特性,既会改变飞机进舰下滑时的配平飞行状态,也会影响其逃逸复飞滑跑时的抬前轮特性。美军研究中也指出,进行逃逸复飞飞行试验时应当特别关注飞机前后重心的配置问题<sup>[1]</sup>。

2) 航母航速是船舶设计部门最为关注的机舰适配参数之一,除了分析典型航速(20 kn)下的参数适配包线外,还十分关注零甲板风条件(即航速为零时)的机舰适配关系,因为它可以表征极限状态的使用条件。

3) 由于第2节建立参数适配包线时尚未考虑航母甲板运动的影响,只能反映平稳海况下的机舰参数适配关系。随着航母编队走向远海,高海况条件也是必然面临的作战使用环境,在舰载机逃逸复飞安全性研究时也必需考虑海况条件的影响。

综上所述,本节将进一步讨论分析舰载机重心位置、航母航速及海况等级等3个因素对参数适配包线的影响,便于发现舰载机总体方案设计和机舰适配特性中可能存在的问题,为方案的设计改进、作战使用条件的论证分析等提供参考。

#### 3.1 舰载机重心位置

通过改变仿真模型的初始计算条件,本文对比了重心位置分别为16%MAC和20%MAC时的逃逸复飞仿真曲线,并且计算获得这2个重心位置所对应的参数适配包线,仿真结果如图7和图8所示。舰载机的重心位置后移后,其进舰下滑时的配平迎角将有所减小,因此着舰俯仰角由6°减小至5.2°。同时,飞机的抬前轮特性有所提高,离舰滑跑距离由139 m缩短至136 m。重心位置后移对参数适配包线的影响如下:着舰俯仰角的约束边界向左移动,由CD段变为GH段;离舰滑跑距离的约束边界向上移动,由BC段变为FG段;迎角增量的约束边界略微向左移动,由AB段变为AE段。

#### 3.2 航母航速

对于舰载机的逃逸复飞任务而言,航母航

速减小后,机舰之间在水平方向上的相对速度增大,在同样的飞行员反应和操纵时间内,飞机的甲板滑跑距离就有所增大,因此满足离舰滑跑距离约束限制的参数适配范围相应减小,如图9所示。

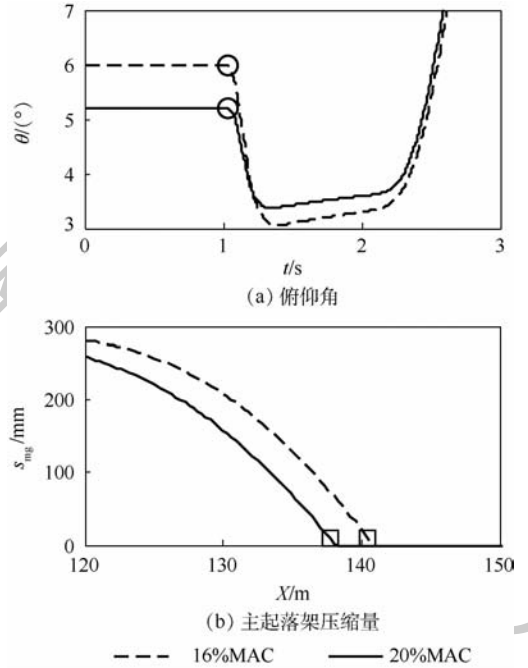


图7 不同重心位置时的仿真曲线  
Fig.7 Simulation curves with different positions of center of gravity

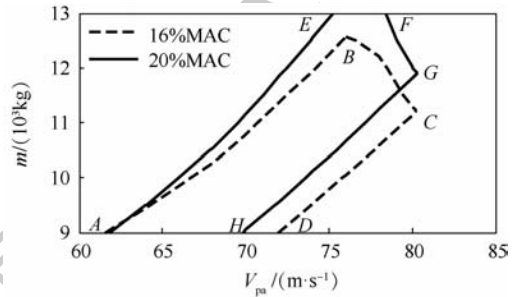


图8 重心位置对参数适配包线的影响  
Fig.8 Effect of position of center of gravity on parameter suitability envelope

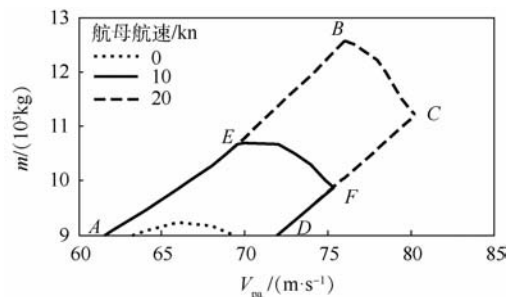


图9 航母航速对参数适配包线的影响  
Fig.9 Effect of carrier velocity on parameter suitability envelope

计算结果表明:航母航速由 20 kn 减小至 10 kn 后,适配的最大着舰重量由  $12.5 \times 10^3$  kg 减小为  $10.7 \times 10^3$  kg,同时适配的最大进舰速度由 80 m/s 减小为 75 m/s。零甲板风条件下,参数适配包线进一步缩小,此时应当严格约束飞机的进舰速度与着舰重量,提高逃逸复飞的安全。

### 3.3 着舰海况等级

在不同的着舰海况等级下,航母甲板的垂荡和俯仰运动的幅值不同,这将影响舰载机的逃逸复飞特性,主要体现在触舰时刻飞机在甲板上的位置、起落架载荷及逃逸复飞的轨迹等<sup>[10]</sup>。在现有的研究中,一般是根据甲板运动的数学模型,选取典型的初始相位值进行仿真计算,从而分析最有利和最不利的极限状态条件。然而,这样的定性处理方式只能适用于特定的甲板运动条件,无法体现出实际情况下初始相位具有随机性的特点。

因此,本文根据式(5)和表1所建立的甲板运动模型,随机选取100组不同的甲板运动初始相位组合,在图6所示的参数适配包线内进行逃

逸复飞的数字虚拟飞行仿真计算,并以逃逸复飞成功率  $p_{bo}$  作为评价指标,从而定量地建立起不同海况等级与逃逸复飞参数适配包线的联系,计算步骤如下:

**步骤1** 随机生成1组甲板运动的初始相位组合,例如  $(\varphi_{\phi_c}, \varphi_{\theta_c}, \varphi_{\Delta z_c}) = (18^\circ, 126^\circ, 144^\circ)$ ,那么按式(5)计算后可分别获得对应初始相位下,4级海况和6级海况条件的甲板运动参数的时间历程。重复这一过程,可以获得200组甲板运动参数的时间历程(4级海况和6级海况各100组),作为步骤2仿真时的计算输入。

**步骤2** 在图6所示的参数适配包线内,任意选取1种舰载机进舰速度和着舰重量的参数组合,例如  $(m, V_{pa}) = (11 \times 10^3 \text{ kg}, 72 \text{ m/s})$ ,再结合已有的200组甲板运动参数进行仿真,可以得到4级海况和6级海况下的仿真算例各100组。

**步骤3** 对仿真算例进行甄别,排除逃逸复飞的失败算例(见图10)后,可以计算逃逸复飞成功率:  $p_{bo}(11 \times 10^3 \text{ kg}, 72 \text{ m/s}, 4 \text{ 级海况}) = 70\%$ ,  $p_{bo}(11 \times 10^3 \text{ kg}, 72 \text{ m/s}, 6 \text{ 级海况}) = 55\%$ 。

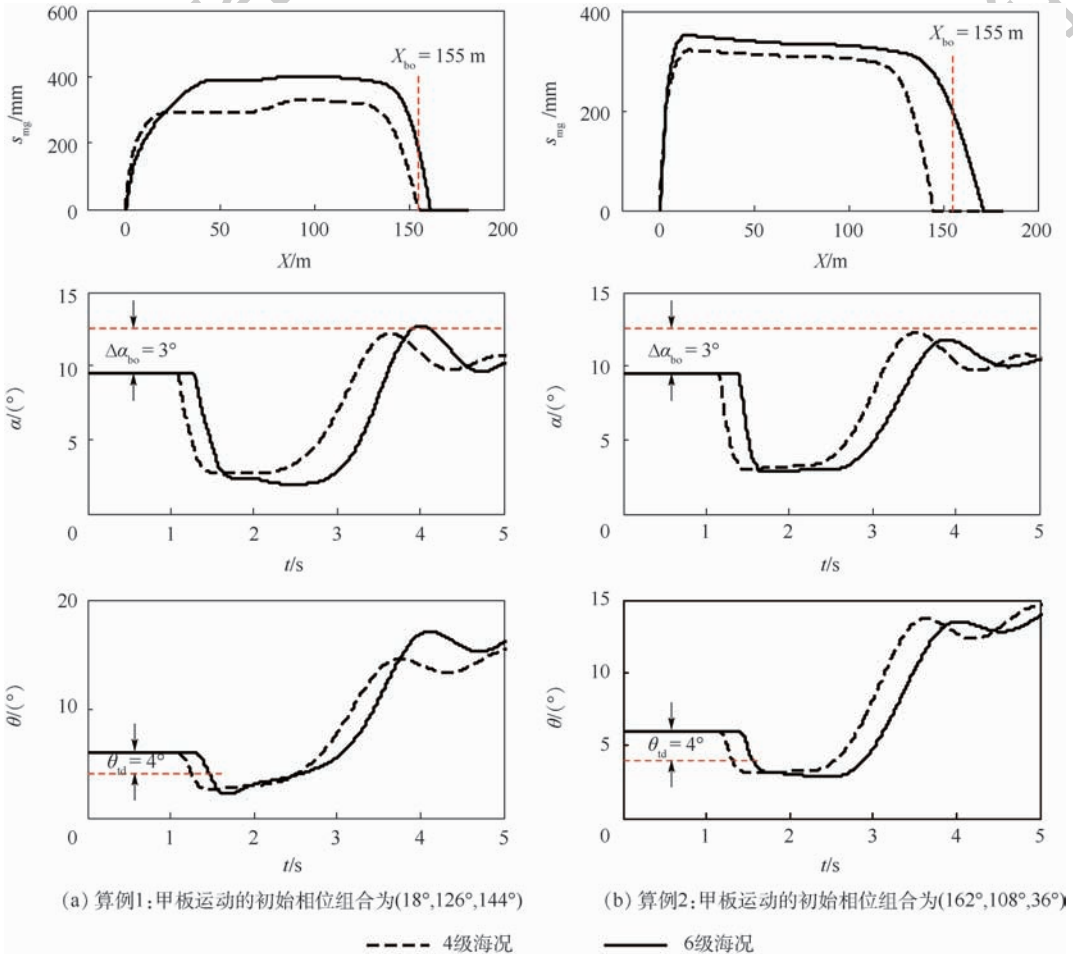


图10 不同海况等级下的逃逸复飞算例

Fig. 10 Bolter simulation samples under different sea states

**步骤4** 对于参数适配包线内的每一组进舰速度和着舰重量的参数组合,重复步骤2和步骤3,并且将参数适配包线内逃逸复飞成功率相同的点连线后,得到了不同海况等级下舰载机的逃逸复飞成功率,如图11所示。

图10给出了2组逃逸复飞仿真算例的对比,根据飞机的离舰滑跑距离(可由主起落架压缩量 $s_{mg}$ 进行计算, $s_{mg}$ 由零开始增大再减小至零的过程即对应了飞机着舰、滑跑和离舰的逃逸复飞过程)、迎角和俯仰角仿真曲线,可以判断逃逸复飞是否成功。图中标注的“ $X_{bo} = 155\text{ m}$ ”“ $\Delta\alpha_{bo} = 3^\circ$ ”“ $\theta_{td} = 4^\circ$ ”分别表示对应参数的安全性判定要求。在算例1中(见图10(a)),飞机离舰滑跑距离、迎角增量、着舰俯仰角3个参数在4级海况条件时分别为154 m、 $2.7^\circ$ 、 $6^\circ$ ,在6级海况条件时分别为160 m、 $3.1^\circ$ 、 $6^\circ$ 。基于离舰滑跑距离小于155 m、迎角增量小于 $3^\circ$ 、着舰俯仰角大于 $4^\circ$ 的安全性判定要求,因此判断4级海况下逃逸复飞成功,6级海况下逃逸复飞失败。同理可判断在算例2中(见图10(b)),由于飞机在6级海况下的离舰滑跑距离过大(约为170 m),因此飞机逃逸复飞失败。

一般来说,如果从飞机离舰后的航迹变化角度分析,甲板具有上升的运动趋势是有利的,因为此时飞机具有更大的初始爬升速度。但在逃逸复飞任务中,由于飞行员更关注飞机是否能够在到达甲板末端前就完成复飞拉起,因此只有当飞机

的爬升速度大于其所处位置的甲板上升速度,才能够表征飞机完成了逃逸复飞。从这一角度来看,甲板具有上升的运动趋势时未必会给逃逸复飞带来有利的影响,即飞机需要获得更大的升力和爬升速度才能克服甲板的上升运动,进而脱离甲板开始爬升。仿真计算结果表明,当甲板具有上升的运动趋势时,飞机逃逸复飞的离舰滑跑距离与迎角增量均有所增大。

如图11所示,当考虑了不同海况等级下的甲板运动影响后,原参数适配包线内的每一着舰重量、进舰速度组合均对应了一个逃逸复飞的成功率。仿真计算结果表明:由于高海况条件下,甲板运动的幅值增大,飞机更容易出现逃逸复飞失败的情况。与4级海况条件相比,6级海况条件下舰载机的逃逸复飞成功率平均下降了约15%。因此,为了提高舰载机的逃逸复飞安全,建议根据 $p_{bo}$ 进一步缩小参数适配包线,并作为设计使用条件时参考。

## 4 结论

1) 离舰滑跑距离、逃逸复飞段最大迎角相对进舰配平迎角的增量和着舰时的俯仰角等3个评定参数可用于分析舰载机的逃逸复飞特性。数字虚拟飞行的仿真计算结果表明,在每一设计着舰重量下均对应存在一个最佳的进舰速度范围,从而使逃逸复飞的离舰滑跑距离最短;同时,最大的迎角增量、着舰时的俯仰角分别决定了满足安全逃逸复飞要求时的最小、最大进舰速度。

2) 适配包线能够表征满足逃逸复飞安全要求的着舰重量、进舰速度的组合情况,其边界由对应的逃逸复飞评定参数确定。舰载机重心位置变化将使参数适配包线的边界位置和范围发生改变,但边界形状基本不变。航母航速减小将显著缩小参数适配包线。

3) 本文算例飞机具有零甲板风条件下的安全逃逸复飞能力。在6级海况条件下,其逃逸复飞的成功率较4级海况条件下下降了约15%。建议以定量计算获得的逃逸复飞成功率作为评价指标,进一步缩小参数适配包线作为设计使用条件,从而提高舰载机逃逸复飞的安全性。

## 参考文献 (References)

- [1] U. S. Department of Defense. Air vehicle joint service specification guide; JSSG—2001A [S]. Patuxent River; Naval Air Systems Command, 2002; D-40.
- [2] 杨一东,江驹,张宏涛,等. 着舰安全与复飞技术[M]. 北京:

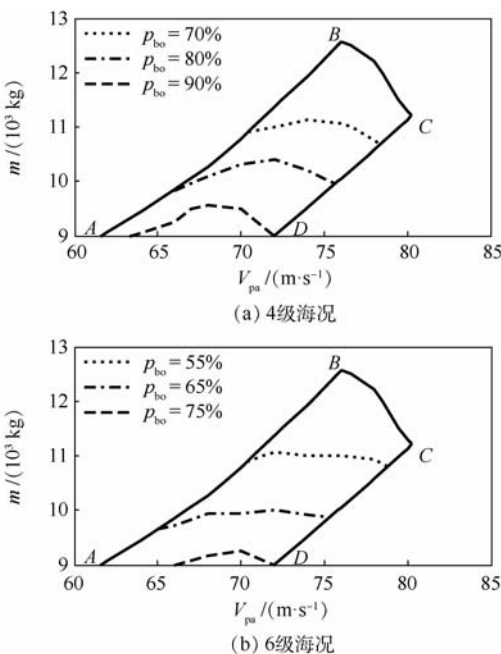


图11 不同海况等级下的逃逸复飞成功率

Fig. 11 Bolter success rate under different sea states

- 国防工业出版社,2013:116-150.
- YANG Y D,JIANG J,ZHANG H T, et al. Safety and waveoff technologies in carrier landing[M]. Beijing: National Defense Industry Press,2013:116-150 (in Chinese).
- [3] RUDOWSKY T,COOK S,HYNES M, et al. Review of the carrier approach criteria for carrier-based aircraft. Phase I: Final report[R]. Patuxent River: Naval Air Warfare Center Aircraft Division,2002.
- [4] 杨一栋. 舰载机进场着舰规范评估[M]. 北京: 国防工业出版社,2016:123-124.
- YANG Y D. Review of the carrier approach criteria[M]. Beijing: National Defense Industry Press,2006:123-124 (in Chinese).
- [5] 王永庆,罗云宝,王奇涛,等. 面向机舰适配的舰载飞机起降特性分析[J]. 航空学报,2016,37(1):269-277.
- WANG Y Q,LUO Y B,WANG Q T, et al. Carrier suitability-oriented launch and recovery characteristics of piloted carrier-based aircraft[J]. Acta Aeronautica et Astronautica Sinica,2016,37(1):269-277 (in Chinese).
- [6] 励纛,温玮,金长江. 舰载飞机逃逸复飞动力学特性研究[J]. 飞行力学,1994,12(2):1-9.
- LI Y,WEN W,JIN C J. The study of dynamic character of bolting and going-around for carrier aircraft[J]. Flight Dynamics,1994,12(2):1-9 (in Chinese).
- [7] 崔坤林,胡国才,罗云宝. 舰载飞机纵向逃逸性能仿真研究[J]. 海军航空工程学院学报,2009,24(4):435-439.
- CUI K L,HU G C,LUO Y B. Simulated investigation for carrier-based airplane longitudinal go-around performance[J]. Journal of Naval Aeronautical and Astronautical University,2009,24(4):435-439 (in Chinese).
- [8] 史卫民,韩维,李辉青. 基于 ADAMS/Aircraft 的舰载机逃逸性能分析[J]. 海军航空工程学院学报,2010,25(5):481-484.
- SHI W M,HAN W,LI H Q. The carrier-based aircraft bolter analysis based on ADAMS/Aircraft[J]. Journal of Naval Aeronautical and Astronautical University,2010,25(5):481-484 (in Chinese).
- [9] 陶杨,韩维,李军亮. 舰载机逃逸过程动力学研究[J]. 计算机仿真,2015,32(3):77-80.
- TAO Y,HAN W,LI J L. Dynamic research of carrier aircraft's escaping[J]. Computer Simulation,2015,32(3):77-80 (in Chinese).
- [10] 段萍萍,聂宏,魏小辉. 飞机触舰后逃逸复飞性能分析[J]. 中国机械工程,2014,25(9):1225-1231.
- DUAN P P,NIE H,WEI X H. Bolting and go-around performance analysis of carrier-based aircraft[J]. China Mechanical Engineering,2014,25(9):1225-1231 (in Chinese).
- [11] 张声伟,段卓毅,耿建中,等. 阻拦索断裂对螺旋桨舰载机着舰安全影响数值分析[J]. 航空学报,2019,40(4):622293.
- ZHANG S W,DUAN Z Y,GENG J Z, et al. Carrier suitability-oriented launch and recovery characteristics of piloted carrier-based aircraft[J]. Acta Aeronautica et Astronautica Sinica,2019,40(4):622293 (in Chinese).
- [12] 国防科学技术工业委员会. 舰载飞机强度和刚度规范地面载荷:GJB 2753—96[S]. 北京: 中国标准出版社,1997.
- The Commission of Science, Technology and Industry for National Defense. Carrier-based airplane strength and rigidity specification ground loads:GJB 2753—96[S]. Beijing: Standards Press of China,1997 (in Chinese).
- [13] SCHARL J,MAVRIS D N,BURDUN I Y. Use of flight simulation in early design: Formulation and application of the virtual testing and evaluation methodology: AIAA-2000-5590[R]. Reston: AIAA,2000.
- [14] 《飞机设计手册》总编委会. 飞机设计手册 第14册 起飞着陆系统设计[M]. 北京: 航空工业出版社,2002:96-98.
- The Chief Committee of Aircraft Design Manual. Aircraft design manual. Volume 14: Takeoff and landing system design[M]. Beijing: Aviation Industry Press,2002:96-98 (in Chinese).
- [15] 高金源,李陆豫,冯亚昌. 飞机飞行品质[M]. 北京: 国防工业出版社,2003:139-143.
- GAO J Y,LI L Y,FENG Y C. Aircraft handling qualities[M]. Beijing: National Defense Industry Press,2003:139-143 (in Chinese).
- [16] ZAAL P M,POOL D M,BRUIJN J, et al. Use of pitch and heave motion cues in a pitch control task[J]. Journal of Guidance, Control and Dynamics,2009,32(2):366-377.
- [17] 陈俊平,王立新. 低能量状态对飞行安全的危害及改出方法[J]. 航空学报,2017,38(8):61-71.
- CHEN J P,WANG L X. Hazards of low energy state to flight safety and recovery methods[J]. Acta Aeronautica et Astronautica Sinica,2017,38(8):61-71 (in Chinese).
- [18] 彭兢. 舰载飞机进舰着舰的自动引导和控制研究[D]. 北京: 北京航空航天大学,2001:33-41.
- PENG J. Research on the automatic guide and control of carrier-based airplane approach and landing[D]. Beijing: Beihang University,2001:33-41 (in Chinese).
- [19] DENISON N A. Automated carrier landing of an unmanned combat aerial vehicle using dynamics inversion[D]. Wright-Patterson Air Force Base: Air Force Institute of Technology,2007:42-43.

## 作者简介:

林佳铭 男,博士研究生。主要研究方向:舰载机起降安全与控制。

吴光辉 男,院士,博士生导师。主要研究方向:飞机总体设计。

王立新 男,教授,博士生导师。主要研究方向:飞机设计、飞行动力学与控制、飞行安全等。

刘海良 男,博士。主要研究方向:飞行动力学与控制。

王云 男,硕士,高级工程师。主要研究方向:船舶总体研究与设计。

## Parameter suitability envelope for safety bolter of a carrier-based aircraft

LIN Jiaming<sup>1</sup>, WU Guanghui<sup>1</sup>, WANG Lixin<sup>1,\*</sup>, LIU Hailiang<sup>1</sup>, WANG Yun<sup>2</sup>

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

2. China Ship Development and Design Center, Wuhan 430064, China)

**Abstract:** The bolter characteristics have significant effect on the approach airspeed determination of a carrier-based aircraft, the grade limits of ship speed and sea state during carrier landing. Three parameters are proposed as the safety indexes according to the bolter task requirements. They are taxiing distance of the bolter, angle of attack increment (the peak angle of attack during bolter minus the approach angle of attack), and pitch angle of touchdown. Numerical pilot model is formulated, and the parameter suitability envelope which satisfies the bolter safety requirements is obtained based on the digital virtual flight testing method. The results show that, for each landing weight, there is a best approach airspeed range to minimize the bolter taxiing distance. The angle of attack increment and the pitch angle of touchdown respectively determine the upper and lower boundary of approach airspeed. Movement of the center of gravity will change the position and range of the parameter suitability envelope, but the envelope shape is nearly unchanged. Reducing the ship speed will narrow the parameter suitability envelope. The success rate corresponding to each point in the parameter suitability envelope provides a reference for the bolter safety judgment under different grades of sea state.

**Keywords:** carrier-based aircraft; bolter; suitability envelope; simulation model; deck motion

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

DOI: 10.13700/j.bh.1001-5965.2018.0738

# 主动防御飞行器的范数型微分对策制导律

郭志强<sup>1</sup>, 孙启龙<sup>2</sup>, 周绍磊<sup>1,\*</sup>, 闫实<sup>1</sup>

(1. 海军航空大学 岸防兵学院, 烟台 264001; 2. 北京航天长征飞行器研究所, 北京 100076)

**摘 要:** 针对具有主动防御能力的飞行器受到攻击导弹的威胁后发射一枚导弹进行防御的制导问题, 基于微分对策理论对飞行器和防御弹的制导律进行了设计和分析。首先, 对于飞行器、防御弹和攻击弹的侧向控制均有界的情况, 基于一种范数型的性能指标推导得出了对策三方的最优制导策略。然后, 当攻击弹采用不同制导策略时, 对飞行器和防御弹能够对策成功的条件进行了分析, 给出了飞行器能够实现逃逸和防御弹能够完成拦截的最小机动条件。最后, 进行了非线性仿真, 结果表明了所提制导律的有效性, 并验证了飞行器若要逃脱攻击弹需满足其最小逃逸机动条件, 防御弹若要拦截攻击弹需满足其最小拦截机动条件。

**关键词:** 飞行器; 主动防御; 范数型; 微分对策; 制导律

**中图分类号:** V448.133

**文献标识码:** A      **文章编号:** 1001-5965(2019)09-1787-10

近年来, 一种被称为“三体 (triplet) 协同追逃”的制导问题在制导领域形成了一个关注的热点<sup>[1-3]</sup>。这里所谓的“三体”是指由一个目标 (target) 和一个可保护目标免受攻击的防御武器 (defender), 以及一个对目标进行拦截的攻击武器 (attacker) 所组成的攻-防-守三方<sup>[4-6]</sup>。由于微分对策理论的主要研究对象是一类具有对抗和竞争关系的群体, 所以攻-防-守三方之间的博弈关系就很自然地适合用微分对策理论加以分析和研究。

上述形式制导问题的研究首见于文献[7], 文中假设目标静止或低速移动, 根据初始弹目距离为目标和防御弹提出了一种闭环最优制导律。文献[8]基于微分对策和线性二次型理论, 为飞行器-防御弹-攻击弹提出了一种最优制导策略, 使得飞行器可以引诱攻击弹, 从而使攻击弹更易被防御弹拦截, 但从其仿真结果来看, 当飞行器速度小于攻击弹速度时, 飞行器总是会被攻击弹命

中, 它必须严格依靠防御弹提前将攻击弹拦截下来才可存活。文献[9-12]将基于微分对策的“三体”协同制导问题称为主动目标防御微分对策 (ATDDG), 并提出了一类基于最优航向的制导策略, 采用这种制导律时目标和防御弹的弹道十分平直且侧向加速度的需求为零。文献[13-14]基于微分对策理论, 将攻-防-守三方的角色进行了通俗化描述并对三方追逃问题进行分析。文献[15]对具有高阶自动驾驶仪动态的三方追逃问题进行了研究, 并得到了对策三方的最优制导策略。文献[16-17]考虑对策各方最大可用控制受约束的情况, 给攻击弹提出了一种先逃逸后追踪的制导策略。文献[18-19]也对控制有界条件下攻击弹的制导策略进行了分析, 并提出了2种改进型的攻击弹制导律。

本文针对飞行器-防御弹-攻击弹的三方追逃问题进行了研究, 但考虑到: ①文献[9-12]提出的制导律仅给目标和防御弹提供了最优航向, 不需

收稿日期: 2018-12-18; 录用日期: 2019-01-18; 网络出版时间: 2019-01-23 09:53

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190121.1340.004.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190121.1340.004.html)

基金项目: 国家自然科学基金 (61273058)

\* 通信作者. E-mail: zhouslsd@sina.com

**引用格式:** 郭志强, 孙启龙, 周绍磊, 等. 主动防御飞行器的范数型微分对策制导律[J]. 北京航空航天大学学报, 2019, 45(9): 1787-1796. GUO Z Q, SUN Q L, ZHOU S L, et al. Norm differential game guidance law for active defense aircraft[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1787-1796 (in Chinese).

采取侧向机动,因而弹道十分平直,但是这种方法忽略了实战当中无论飞行器、防御弹还是攻击弹都需要进行一定的机动<sup>[17]</sup>;②文献[8,13-15]提出的均是基于线性二次型形式的制导策略,这类制导策略没有考虑对策参与者的控制往往都是有界的这一实际问题,在应用中容易出现加速度饱和现象而使制导律失去最优性,产生较大的脱靶量;③文献[16-19]虽考虑了控制受有界约束的情况,但只从攻击弹的角度出发对制导问题进行研究,文献[16-17]只分析了攻击弹采取先逃逸后追踪策略时能够命中飞行器的条件,文献[18-19]也只讨论并设计了攻击弹的制导律,并未对飞行器和防御弹的对策条件进行分析。

鉴于上述情况,本文基于微分对策理论对控制有界情况下的三方追逃问题进行了分析,基于范数型的性能指标为对策各方提出了不会导致加速度饱和的最优制导策略,然后从防守方(飞行器和防御弹)的角度出发,对攻击弹采取不同策略时飞行器能够成功逃离攻击弹的机动条件(不严格依赖防御弹)和防御弹能够成功拦截攻击弹的机动条件分别进行了分析和推导。

## 1 问题描述与运动学建模

### 1.1 协同追逃对策问题

本文主要研究一个飞行器(目标) $T$ 受到一枚攻击弹 $M$ 的威胁后发射一枚防御弹 $D$ 进行防御的制导问题,考虑对策三方的侧向加速度均有界,并在制导律的推导过程中做出以下假设:

- 1) 对策三方在运动中均可视为质点,且速度均为常值。
- 2) 对策三方之间的相对运动可在初始视线方向附近线性化。
- 3) 重力因素可忽略,且制导问题可以在俯仰平面和偏航平面内解耦。

基于以上假设,本文主要在偏航平面内对制导问题进行研究,研究结论可以推广到三维情形。图1所示为笛卡儿惯性坐标系下末制导段三方的平面相对运动关系。图中: $V$ 、 $a$ 和 $\gamma$ 分别为速度、侧向加速度和航向角; $r_{MT}$ 和 $\lambda_{MT}$ 分别为飞行器与攻击弹之间的距离和视线角; $r_{MD}$ 和 $\lambda_{MD}$ 分别为防御弹与攻击弹之间的距离和视线角; $y_{MT}$ 为飞行器与攻击弹在垂直于初始视线 $LOS_{MTO}$ 方向上的距离; $y_{MD}$ 为防御弹与攻击弹在垂直于初始视线 $LOS_{MDO}$ 方向上的距离。下标 $T$ 、 $D$ 和 $M$ 分别表示飞行器、防御弹和攻击弹,数字0表示相应的初值。

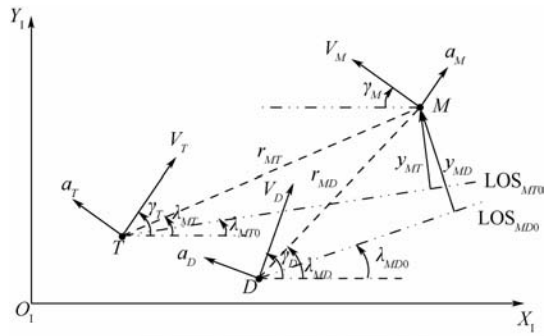


图1 对策三方的平面相对运动关系

Fig. 1 Relative in-plane movement relationship of three players

由图1可得极坐标下飞行器与攻击弹之间的相对运动方程为

$$\begin{cases} \dot{r}_{MT} = V_M \cos(\gamma_M + \lambda_{MT}) + V_T \cos(\gamma_T - \lambda_{MT}) \\ r_{MT} \dot{\lambda}_{MT} = V_M \sin(\gamma_M + \lambda_{MT}) - V_T \sin(\gamma_T - \lambda_{MT}) \end{cases} \quad (1)$$

类似可得防御弹与攻击弹之间的相对运动方程为

$$\begin{cases} \dot{r}_{MD} = V_M \cos(\gamma_M + \lambda_{MD}) + V_D \cos(\gamma_D - \lambda_{MD}) \\ r_{MD} \dot{\lambda}_{MD} = V_M \sin(\gamma_M + \lambda_{MD}) - V_D \sin(\gamma_D - \lambda_{MD}) \end{cases} \quad (2)$$

式中: $\dot{r}_{MT}$ 、 $\dot{r}_{MD}$ 分别为飞行器和攻击弹、防御弹和攻击弹的相对速度; $\dot{\lambda}_{MT}$ 、 $\dot{\lambda}_{MD}$ 分别为飞行器和攻击弹、防御弹和攻击弹的视线角速度。

飞行器、防御弹和攻击弹各自的线性动力学方程为

$$\begin{cases} \dot{x}_i = A_i x_i + B_i u'_i \\ a_i = C_i x_i + d_i u'_i \end{cases} \quad i = T, D, M \quad (3)$$

式中: $x_i$ 为三方各自的内部状态向量; $u'_i$ 为其控制指令。假设三方的控制系统均具有理想动态特性,则有 $A_i = 0$ 、 $B_i = 0$ 、 $C_i = 0$ 、 $d_i = 1$ ,相应的加速度指令 $u'_i = a_i$ 。

三方的航向角速度可表示为

$$\dot{\gamma}_i = \frac{a_i}{V_i} \quad i = T, D, M \quad (4)$$

将飞行器、防御弹和攻击弹的最大加速度指令记为 $u_i^{\max}$ ( $i = T, D, M$ ),则其控制指令应满足:

$$|u'_i| \leq u_i^{\max} \quad i = T, D, M \quad (5)$$

### 1.2 运动学建模与系统降阶

将飞行器在垂直于 $LOS_{MTO}$ 方向上的加速度记为 $u_{TN}$ ,攻击弹在垂直于 $LOS_{MTO}$ 方向上的加速度记为 $u_{MN}$ ,防御弹在垂直于 $LOS_{MDO}$ 方向上的加速度记为 $u_{DN}$ ,则可得

$$\begin{cases} u_{TN} = u'_T \cos(\gamma_{T0} - \lambda_{MTO}) \\ u_{MN} = u'_M \cos(\gamma_{M0} + \lambda_{MTO}) \\ u_{DN} = u'_D \cos(\gamma_{D0} - \lambda_{MDO}) \end{cases} \quad (6)$$

根据假设2),飞行器、防御弹和攻击弹的对策时间(即飞行时间)可近似表示为

$$\begin{cases} t_{IMT} = \frac{r_{MTO}}{\dot{r}_{MTO}} \\ t_{IMD} = \frac{r_{MDO}}{\dot{r}_{MDO}} \end{cases} \quad (7)$$

式中: $t_{IMT}$ 为飞行器与攻击弹的对策时间; $t_{IMD}$ 为防御弹与攻击弹的对策时间。在后面的非线性仿真中,飞行时间将根据式(7)进行实时更新。

由于飞行器希望自身被攻击弹命中之前防御弹便对攻击弹实施拦截,故对策时间应满足 $t_{IMT} > t_{IMD}$ ,可记 $\Delta t = t_{IMT} - t_{IMD}$ 。又由于防御弹和攻击弹的距离在 $t = t_{IMD}$ 时刻达到最小(此即防御弹的脱靶量),当 $t > t_{IMD}$ 时防御弹和攻击弹之间的距离开始变大,防御弹已无法再对攻击弹形成威胁,它也不能再对飞行器和攻击弹的对策局势产生影响。故可规定在 $t_{IMD} < t \leq t_{IMT}$ 时间内防御弹不再存在,其存在函数可用 $l(t)$ 表示:

$$l(t) = \begin{cases} 1 & t \leq t_{IMD} \\ 0 & t > t_{IMD} \end{cases} \quad (8)$$

则在对策中防御弹的加速度分量 $u_{DN}$ 可表示为

$$u_{DN} = u'_D \cos(\gamma_{D0} - \lambda_{MDO}) l(t) \quad (9)$$

将上述对策问题的状态向量取为

$$\mathbf{x} = [y_{MT} \quad y_{MD} \quad \dot{y}_{MT} \quad \dot{y}_{MD}]^T \quad (10)$$

记 $u_T = u_{TN}$ , $u_D = u_{DN}$ , $u_M = u_{MN}$ ,则系统在垂直于初始视线方向上的运动方程可写为

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}_{TD} + \mathbf{C}u_M \quad (11)$$

式中:

$$\begin{cases} \mathbf{u}_{TD} = [u_T \quad u_D]^T \\ \mathbf{A} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ \mathbf{B} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & -l(t) \end{bmatrix} \\ \mathbf{C} = [0 \quad 0 \quad 1 \quad \varepsilon]^T \\ \varepsilon = \cos(\lambda_{MTO} - \lambda_{MDO}) \end{cases}$$

因为防御弹为飞行器所发射,所以一般有 $\varepsilon = 1$ 。

利用类似于文献[8]提到的所谓终端投影的系统降阶方法,引入新的变量——零控脱靶量(Zero Effort Miss, ZEM):

$$\mathbf{Z}(t) = [Z_{MT}(t) \quad Z_{MD}(t)]^T = \mathbf{D}\Phi(t_f, t)\mathbf{x}(t) \quad (12)$$

式中: $\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$ 为系统状态提取矩阵; $\Phi(t_f, t) = L^{-1}[(s\mathbf{I} - \mathbf{A})^{-1}]$ 为系统方程式(11)的状态转移矩阵,且

$$\Phi(t_f, t) = \begin{bmatrix} 1 & 0 & t_{goMT} & 0 \\ 0 & 1 & 0 & t_{goMD} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (13)$$

式中: $t_{goMT}$ 为飞行器和攻击弹之间的剩余对策时间; $t_{goMD}$ 为防御弹和攻击弹之间的剩余对策时间。

经过式(12)的变换之后可得2个零控脱靶量:

$$\begin{cases} Z_{MT}(t) = y_{MT} + \dot{y}_{MT}t_{goMT} \\ Z_{MD}(t) = y_{MD} + \dot{y}_{MD}t_{goMD} \end{cases} \quad (14)$$

将零控脱靶量作为新的状态变量,并求其关于时间的导数,可得新的状态方程为

$$\begin{cases} \dot{Z}_{MT}(t) = (u_M - u_T)t_{goMT} \\ \dot{Z}_{MD}(t) = (u_M - u_D)t_{goMD} \end{cases} \quad (15)$$

## 2 范数型追逃对策

### 2.1 范数型的最优对策

在对策过程中,对于飞行器和防御弹组成的防守方来说,总是希望攻击弹的脱靶量 $y_{MT}(t_{IMT})$ 达到最大,同时希望防御弹的脱靶量 $y_{MD}(t_{IMD})$ 达到最小;而对策的另一方攻击弹其目的则完全相反,既希望 $y_{MD}(t_{IMD})$ 达到最大,又希望 $y_{MT}(t_{IMT})$ 达到最小。据此可选取2个脱靶量的绝对值之差作为对策问题的性能指标:

$$J = |y_{MD}(t_{IMD})| - |y_{MT}(t_{IMT})| \quad (16)$$

由于脱靶量的绝对值属于范数的范畴,此类微分对策问题又被称为范数型微分对策(Norm Differential Game, NDG)<sup>[20]</sup>。

对策三方所进行的最优化问题可描述为

$$\min_{u_T, u_D} \max_{u_M} J$$

利用零控脱靶量可将性能指标式(16)改写为

$$J = |Z_{MD}(t_{IMD})| - |Z_{MT}(t_{IMT})| \quad (17)$$

该最优问题的哈密顿函数为

$$H = \eta_{MT} \dot{Z}_{MT} + \eta_{MD} \dot{Z}_{MD} \quad (18)$$

式中： $\eta_{MT}$ 、 $\eta_{MD}$ 为伴随状态。

最优问题的伴随方程和横截条件分别为

$$\begin{cases} \dot{\eta}_{MT}(t) = -\frac{\partial H}{\partial Z_{MT}} = 0 \\ \dot{\eta}_{MD}(t) = -\frac{\partial H}{\partial Z_{MD}} = 0 \end{cases} \quad (19)$$

$$\begin{cases} \eta_{MT}(t_{IMT}) = \frac{\partial J}{\partial Z_{MT}} \Big|_{t=t_{IMT}} = -\text{sign}(Z_{MT}(t_{IMT})) \\ \eta_{MD}(t_{IMD}) = \frac{\partial J}{\partial Z_{MD}} \Big|_{t=t_{IMD}} = \text{sign}(Z_{MD}(t_{IMD})) \end{cases} \quad (20)$$

则进一步可得伴随状态为

$$\begin{cases} \eta_{MT}(t) = -\text{sign}(Z_{MT}(t_{IMT})) \\ \eta_{MD}(t) = \text{sign}(Z_{MD}(t_{IMD})) \end{cases} \quad (21)$$

根据最优控制的极大极小值原理,飞行器、防御弹和攻击弹的最优控制应满足:

$$\begin{cases} (u_T^*, u_D^*) = \arg \min_{(u_T, u_D)} H \\ u_M^* = \arg \max_{u_M} H \end{cases} \quad (22)$$

结合式(15)和式(18)可得最优制导策略为

$$\begin{cases} u_T^*(t) = u_T^{\max} \text{sign}(\eta_{MT} t_{goMT}) \\ u_D^*(t) = u_D^{\max} \text{sign}(\eta_{MD} t_{goMD}) l(t) \\ u_M^*(t) = u_M^{\max} \text{sign}(\eta_{MD} t_{goMD} l(t) + \eta_{MT} t_{goMT}) \end{cases} \quad (23)$$

由于  $t_{goMT} \geq 0, t_{goMD} \geq 0$ , 利用式(21)并将最优制导策略进一步写成闭环形式可得

$$\begin{cases} u_T^*(t) = -u_T^{\max} \text{sign}(Z_{MT}(t)) \\ u_D^*(t) = u_D^{\max} \text{sign}(Z_{MD}(t)) l(t) \\ u_M^*(t) = u_M^{\max} \text{sign}(\text{sign}(Z_{MD}(t)) t_{goMD} l(t) - \text{sign}(Z_{MT}(t)) t_{goMT}) \end{cases} \quad (24)$$

## 2.2 对策空间分析

### 2.2.1 飞行器-攻击弹的最优对策空间

飞行器与攻击弹进行一对一追逃对策时,易得其最优动力学为

$$\frac{dZ_{MT}^*}{dt} = (-u_M^{\max} + u_T^{\max}) t_{goMT} \text{sign}(Z_{MT}(t)) \quad (25)$$

对式(25)从0到t积分可得飞行器-攻击弹的最优对策轨迹为

$$|Z_{MT}^*(t)| = R_M + \frac{1}{2}(u_M^{\max} - u_T^{\max}) t_{goMT}^2 \quad (26)$$

式中： $R_M = |Z_{MT}^*(t_{IMT})| = |Z_{MT}^*(0)| - \frac{1}{2} t_{IMT}^2 (u_M^{\max} - u_T^{\max})$  为攻击弹的杀伤半径或其期望的最小脱靶量。

在本文所讨论的问题中,相较于攻击弹而言,一般有  $u_T^{\max} \leq u_M^{\max}$ , 故根据式(26)可得飞行器-攻

击弹的对策空间如图2所示。图中加粗的单调曲线  $Z_{MT}^*$  和  $-Z_{MT}^*$  表示攻击弹最终脱靶量为零的最优对策轨迹, 2条曲线之间的区域  $D_{MT0} (|Z_{MT}(t)| < Z_{MT}^*)$  为对策的奇异区, 在该区域内飞行器和攻击弹的最优制导策略是任意的, 如果初始时刻的零控脱靶量  $Z_{MT}(0)$  位于此区域内, 则攻击弹将对飞行器实现零脱靶量拦截(即  $Z_{MT}(t_{IMT}) = 0$ )。在奇异区  $D_{MT0}$  以外的2个区域  $D_{MT1}^+ (Z_{MT}(t) > Z_{MT}^*)$  和  $D_{MT1}^- (Z_{MT}(t) < -Z_{MT}^*)$ , 攻击弹不能对飞行器实现零脱靶量拦截, 双方均使用最优对策时的对策轨迹平行于曲线  $Z_{MT}^*$  和  $-Z_{MT}^*$ , 最终脱靶量  $Z_{MT}^*(t_{IMT}) \neq 0$ , 若飞行器能在某一时刻  $t (t < t_{IMT})$  使得  $|Z_{MT}(t)| > |Z_{MT}^*(t)|$ , 则攻击弹的最终脱靶量将大于  $R_M = Z_{MT}^*(t_{IMT})$  从而使飞行器实现逃逸。

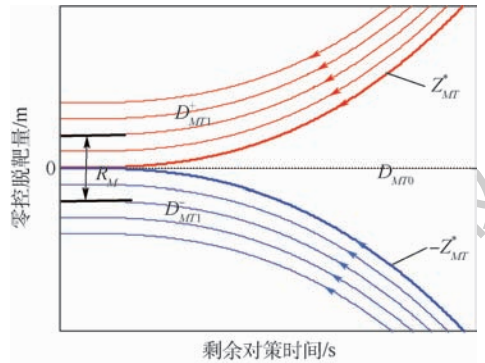


图2 飞行器-攻击弹的对策空间

Fig.2 Aircraft-attacking missile game space

### 2.2.2 防御弹-攻击弹的最优对策空间

防御弹与攻击弹进行一对一追逃对策时,易得其最优动力学为

$$\frac{dZ_{MD}^*}{dt} = (u_M^{\max} - u_D^{\max}) t_{goMD} \text{sign}(Z_{MD}(t)) \quad (27)$$

对式(27)从0到t积分可得防御弹-攻击弹的最优对策轨迹为

$$|Z_{MD}^*(t)| = R_D - \frac{1}{2}(u_M^{\max} - u_D^{\max}) t_{goMD}^2 \quad (28)$$

式中： $R_D = |Z_{MD}^*(t_{IMD})| = |Z_{MD}^*(0)| + \frac{1}{2} t_{IMD}^2 (u_M^{\max} - u_D^{\max})$  为防御弹的杀伤半径或其期望的最小脱靶量。

对于防御弹和攻击弹而言:①若  $u_D^{\max} \geq u_M^{\max}$ , 根据式(28)可得二者的对策空间与图2相似, 其分析也类似, 不再赘述。②若  $u_D^{\max} < u_M^{\max}$ , 则二者的对策空间如图3所示, 最优对策轨迹的值  $|Z_{MD}^*|$  随着时间t单调增加, 最终脱靶量趋于  $Z_{MD}^*(t_{IMD})$ , 若防御弹在某一时刻  $t (t < t_{IMD})$  使得

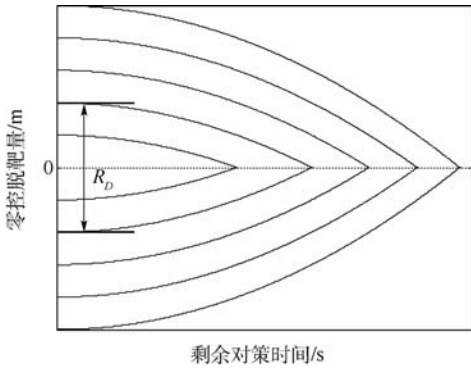


图3 防御弹-攻击弹的对策空间 ( $u_D^{\max} < u_M^{\max}$ )

Fig.3 Defender-attacking missile game space ( $u_D^{\max} < u_M^{\max}$ )

$|Z_{MD}(t)| < |Z_{MD}^*(t)|$ , 则防御弹的最终脱靶量将小于  $R_D = Z_{MD}^*(t_{IMD})$  从而将攻击弹拦截。

### 3 飞行器和防御弹的对策条件

本节根据 2.1 节推得的最优制导策略式(24)对三方的对策情况进行分析,并根据攻击弹的不同机动情况对飞行器实现逃逸和防御弹实现拦截的条件进行推导。这里假设对策三方的最大加速度指令满足不等式  $u_T^{\max} \leq u_D^{\max} \leq u_M^{\max}$ 。

将式(24)代入式(15)可得系统最优动力学方程为

$$\begin{cases} \frac{dZ_{MT}^*}{dt} = [u_M^{\max} \text{sign}(\text{sign}(Z_{MD}(t)))t_{goMD}l(t) - \text{sign}(Z_{MT}(t))t_{goMT} + u_T^{\max} \text{sign}(Z_{MT}(t))]t_{goMT} \\ \frac{dZ_{MD}^*}{dt} = [u_M^{\max} \text{sign}(\text{sign}(Z_{MD}(t)))t_{goMD}l(t) - \text{sign}(Z_{MT}(t))t_{goMT} - u_D^{\max} \text{sign}(Z_{MD}(t))]t_{goMD} \end{cases} \quad (29)$$

最优动力学方程(29)中均包含了 2 个零控脱靶量的符号函数,下面根据符号函数  $\text{sign}(Z_{MT}(t))$ 、 $\text{sign}(Z_{MD}(t))$  分 2 种情况进行分析。

#### 1) 情况 1

$\text{sign}(Z_{MD}(t)) = -\text{sign}(Z_{MT}(t))$ 。此时飞行器、防御弹与攻击弹的视线角速度  $\lambda_{MT}$  和  $\lambda_{MD}$  的符号相反,攻击弹的最优追踪策略(相对于飞行器)和最优逃逸策略(相对于防御弹)的符号也相反,攻击弹可在对飞行器进行追踪的同时对防御弹进行逃逸,即

$$u_{Me}^*(t) = u_M^{\max} \text{sign}(Z_{MD}(t)) = -u_M^{\max} \text{sign}(Z_{MT}(t)) = u_{Mp}^*(t) \quad (30)$$

式中:  $u_{Me}^*(t)$  表示攻击弹的最优逃逸策略;  $u_{Mp}^*(t)$  表示攻击弹的最优追踪策略。

将式(30)代入式(29),并分别从 0 到  $t$  积分,可得 2 个最优对策轨迹依然如式(26)和

式(28)所示。从式(30)也可看出,此时攻击弹的最优逃逸策略与最优追踪策略相同,这种情况下的最优对策只跟对策各方的初始条件有关,初始条件确定后各方的最优对策轨迹便随之确定,飞行器和防御弹只能采取各自的最优策略来应对,因而此种情况是非常简单的。根据式(26)和式(28)可得此时零控脱靶量如图 4 所示。

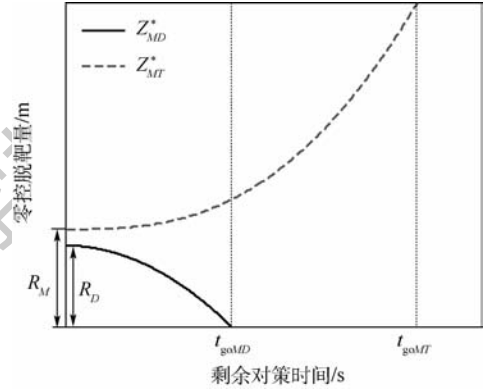


图4 2个零控脱靶量异号时的变化

Fig.4 Evolution of two ZEM in case of opposite signs

#### 2) 情况 2

$\text{sign}(Z_{MT}(t)) = \text{sign}(Z_{MD}(t))$ 。此时 2 个视线角速度  $\lambda_{MT}$  和  $\lambda_{MD}$  的符号相同,根据式(24)第 3 式可得  $t < t_{IMD}$  时攻击弹的最优对策为  $u_M^*(t) = 0$ , 但这并不意味着攻击弹不会采取机动,其意义是: ①攻击弹可选择最优逃逸策略  $u_{Me}^*(t) = u_M^{\max} \cdot \text{sign}(Z_{MD}(t))$  躲避防御弹,攻击弹与防御弹间的脱靶量将增大,但其与飞行器之间的脱靶量也将同时增大; ②攻击弹也可直接采取最优追踪策略  $u_{Mp}^*(t) = u_M^{\max} \text{sign}(Z_{MT}(t))$  来攻击飞行器,此时攻击弹与飞行器间的脱靶量将减小,但其与防御弹之间的脱靶量也将同时减小。下面根据攻击弹的对策情况分别进行分析。

①攻击弹首先采用最优逃逸策略而忽略对飞行器的拦截。此时攻击弹的制导策略为

$$u_{Me}^*(t) = u_M^{\max} \text{sign}(Z_{MD}(t)) = u_M^{\max} \text{sign}(Z_{MT}(t)) = -u_{Mp}^*(t) \quad (31)$$

此时攻击弹的制导策略与其最优追踪策略是完全相反的策略,若飞行器和防御弹均采用各自的最优策略,将式(31)代入式(29)并分别从 0 到  $t$  积分则可得三方的对策轨迹为

$$\begin{cases} |Z_{MT}^{**}(t)| = |Z_{MT}^{**}(t_{IMT})| - \frac{1}{2}(u_T^{\max} + u_M^{\max})t_{goMT}^2 \\ |Z_{MD}^*(t)| = R_D - \frac{1}{2}(u_M^{\max} - u_D^{\max})t_{goMD}^2 \end{cases} \quad (32)$$

根据式(32)第1式可得零控脱靶量同号时飞行器-攻击弹的对策轨迹如图5中的点线所示。由图5可以看出,由于攻击弹首先采用逃逸策略,其目的是逃脱防御弹的拦截而忽略了飞行器,导致其对飞行器的脱靶量将随着时间增大。如果攻击弹与飞行器之间的零控脱靶量在某一时刻 $t'(t' < t_{IMD})$ 满足 $|Z_{MT}^{**}(t')| > |Z_{MT}^*(t')|$ ,根据2.2.1节的分析可知,即便攻击弹再切换成最优追踪策略,飞行器已经不在飞行器-攻击弹最优对策的捕获区内,飞行器采用最优逃逸策略时仍能够以脱靶量 $|Z_{MT}^{**}(t_{IMT})| > R_M$ 结束与攻击弹的对抗,从而成功逃离攻击弹。

在攻击弹与防御弹进行最优逃逸-追踪对策的同时,飞行器以 $u_T(t) = -u_T^{max} \text{sign}(z_{MT}(t))$ 进行最优逃逸机动,若在推导式(32)第1式时进行一些代数运算则可得

$$|Z_{MT}^{**}(t)| = |Z_{MT}^{**}(0)| + \int_0^t (u_T^{max} + u_M^{max})(t_{IMT} - \xi) d\xi = |Z_{MT}^{**}(0)| + (u_T^{max} + u_M^{max})t_{IMT}t - \frac{1}{2}(u_T^{max} + u_M^{max})t^2 = |Z_{MT}^{**}(0)| + \frac{1}{2}(u_T^{max} + u_M^{max})t_{IMT}^2 - \frac{1}{2}(u_T^{max} + u_M^{max})t^2 + (u_T^{max} + u_M^{max})t_{IMT}t - \frac{1}{2}(u_T^{max} + u_M^{max})t^2 = |Z_{MT}^{**}(0)| + \frac{1}{2}(u_T^{max} + u_M^{max})t_{IMT}^2 - \frac{1}{2}(u_T^{max} + u_M^{max})t_{goMT}^2 \quad (33)$$

式中: $|Z_{MT}^{**}(0)|$ 为初始时刻( $t=0$ )的飞行器-攻击弹零控脱靶量。

再结合2.2.1节的分析可知,零控脱靶量式(33)与最优对策轨迹式(26)应有下列不等式关系成立:

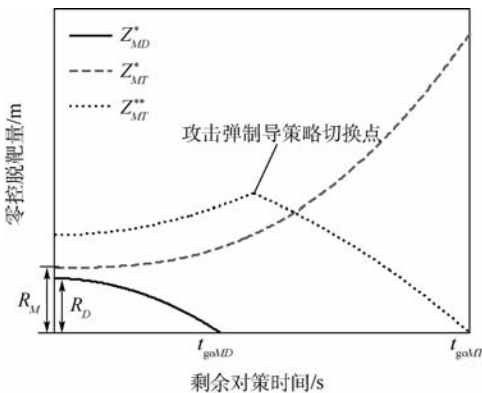


图5 攻击弹首先采用最优逃逸策略时的零控脱靶量

Fig. 5 Evolution of ZEM while attacking missile performs optimal evasion first

$$R_M + \frac{1}{2}(u_M^{max} - u_T^{max})t_{goMT}^2 \leq |Z_{MT}(0)| + \frac{1}{2}(u_T^{max} + u_M^{max})t_{IMT}^2 - \frac{1}{2}(u_T^{max} + u_M^{max})t_{goMT}^2 \quad (34)$$

整理可得

$$u_T^{max} \geq \frac{2R_M - 2|Z_{MT}^{**}(0)| + 2u_M^{max}t_{goMT}^2}{t_{IMT}^2} - u_M^{max} \quad (35)$$

又因为式(34)最晚应在 $t = t_{IMD}$ 时刻成立(此时 $t_{goMT} = t_{IMT} - t_{IMD} = \Delta t$ ),故 $u_T^{max}$ 应满足:

$$u_T^{max} \geq \frac{2R_M - 2|Z_{MT}^{**}(0)| + 2u_M^{max}(\Delta t)^2}{t_{IMT}^2} - u_M^{max} \quad (36)$$

式(36)即为飞行器的最大加速度指令应满足的最小逃逸机动条件。

②攻击弹直接采用最优追踪策略而忽略对防御弹的躲避。此时攻击弹的制导策略为

$$u_{Mp}^*(t) = -u_M^{max} \text{sign}(Z_{MT}(t)) = -u_M^{max} \text{sign}(Z_{MD}(t)) = -u_{Mc}^*(t) \quad (37)$$

此时攻击弹的制导策略与其最优逃逸策略是完全相反的策略,若飞行器和防御弹均采用其最优策略,将式(37)代入式(29)并分别从0到t积分可得三方的对策轨迹为

$$\begin{cases} |Z_{MT}^*(t)| = R_M + \frac{1}{2}(u_M^{max} - u_T^{max})t_{goMT}^2 \\ |Z_{MD}^*(t)| = |Z_{MD}^*(t_{IMD})| + \frac{1}{2}(u_M^{max} + u_D^{max})t_{goMD}^2 \end{cases} \quad (38)$$

根据式(38)第2式可得防御弹-攻击弹的最优对策轨迹如图6中的点线所示。由于攻击弹首先采用的是追踪策略,其目的是对飞行器进行拦截而忽略了对防御弹的躲避,此时其与防御弹之间的脱靶量也将随着时间而减小。如果攻击弹在捕获飞行器之前的某一时刻 $t'(t' < t_{IMD})$ 与防御弹之间的零控脱靶量满足 $|Z_{MD}^{**}(t')| < |Z_{MD}^*(t')|$ ,

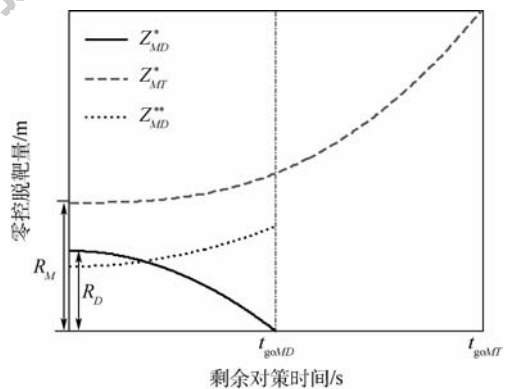


图6 攻击弹直接采用最优追踪策略时的零控脱靶量

Fig. 6 Evolution of ZEM while attacking missile performs optimal pursuit directly

则根据 2.2.2 节的分析可知,防御弹采用其最优追踪策略时就能够以脱靶量  $|Z_{MD}^{**}(t_{IMD})| < R_D$  结束与攻击弹的对抗,从而成功拦截攻击弹。

在攻击弹与飞行器进行最优追踪-逃逸对策的同时,防御弹以  $u_D(t) = u_D^{\max} \text{sign}(Z_{MD}(t))$  对攻击弹进行最优追踪机动,若在推导式(38)第2式时进行一些代数运算则可以得到:

$$\begin{aligned}
 |Z_{MD}^{**}(t)| &= |Z_{MD}^{**}(0)| - \int_0^t (u_M^{\max} + u_D^{\max})(t_{IMD} - \xi) d\xi = |Z_{MD}^{**}(0)| - (u_M^{\max} + u_D^{\max})t_{IMD}t + \frac{1}{2}(u_M^{\max} + u_D^{\max})t^2 = |Z_{MD}^{**}(0)| - \frac{1}{2}(u_M^{\max} + u_D^{\max})t_{IMD}^2 + \frac{1}{2}(u_M^{\max} + u_D^{\max})t_{goMD}^2 \\
 &= |Z_{MD}^{**}(0)| - \frac{1}{2}(u_M^{\max} + u_D^{\max})t_{IMD}^2 + \frac{1}{2}(u_M^{\max} + u_D^{\max})t_{goMD}^2 \quad (39)
 \end{aligned}$$

式中:  $|Z_{MD}^{**}(0)|$  为初始时刻 ( $t=0$ ) 的防御弹-攻击弹零控脱靶量。

再结合 2.2.2 节的分析可知,零控脱靶量式(39)与最优对策轨迹式(28)应有下列不等式关系成立:

$$\begin{aligned}
 R_D - \frac{1}{2}(u_M^{\max} - u_D^{\max})t_{goMD}^2 &\geq |Z_{MD}^{**}(0)| - \frac{1}{2}(u_M^{\max} + u_D^{\max})t_{IMD}^2 + \frac{1}{2}(u_M^{\max} + u_D^{\max})t_{goMD}^2 \quad (40)
 \end{aligned}$$

整理可得

$$u_D^{\max} \geq \frac{2|Z_{MD}^{**}(0)| - 2R_D + 2u_M^{\max}t_{goMD}^2}{t_{IMD}^2} - u_M^{\max} \quad (41)$$

又因为式(40)最晚应在  $t = t_{IMD}$  时刻成立(此时  $t_{goMD} = 0$ ),故可得

$$u_D^{\max} \geq \frac{2|Z_{MD}^{**}(0)| - 2R_D}{t_{IMD}^2} - u_M^{\max} \quad (42)$$

式(42)即为防御弹的最大加速度指令应满足的最小拦截机动条件。

## 4 仿真验证

本节基于非线性仿真对 2.1 节推导的制导律及第 3 节的分析进行验证。由于零控脱靶量  $Z_{MT}$  和  $Z_{MD}$  异号的情况对于飞行器和防御弹来说是最简单的情况,故仿真的主要内容为第 3 节中  $\text{sign}(Z_{MT}(t)) = \text{sign}(Z_{MD}(t))$  的情况。假设防御弹对攻击弹的脱靶量小于 10 m 时即认为攻击弹

被拦截 ( $R_D = 10$  m),攻击弹对飞行器的脱靶量小于 10 m 时即认为飞行器被命中 ( $R_M = 10$  m)。

### 4.1 攻击弹首先采用最优逃逸策略

此时攻击弹先采用式(31)所示的逃逸策略,逃脱之后再切换成式(37)所示的追踪策略拦截飞行器,而飞行器和防御弹的制导策略分别如式(24)第1式和第2式所示,其他相关仿真参数如表1所示。

为了验证飞行器能否逃离攻击弹,在仿真中应根据所给参数由式(36)进行实时判断。

1) 假设飞行器、防御弹和攻击弹的最大加速度指令分别为  $u_T^{\max} = 50 \text{ m/s}^2$ 、 $u_D^{\max} = 80 \text{ m/s}^2$  和  $u_M^{\max} = 180 \text{ m/s}^2$ 。基于上述参数,在仿真中可以发现飞行器的控制指令满足式(36)的约束,此时三方的运动轨迹如图7所示。从图7可以明显看出,当飞行器的最大加速度指令满足最小机动条件时可以成功逃脱攻击弹,飞行器并不严格依赖于防御弹提前将攻击弹拦截。这一点与文献[8]不同,该文中当飞行器速度小于攻击弹速度时,飞行器始终会被攻击弹命中,飞行器严格依赖防御弹的存在。图8为2个剩余对策时间随时间变化的示意图。可以看出,由于攻击弹先采用逃逸策略,剩余对策时间  $t_{goMT}$  与时间  $t$  已完全不是线性的关系,即使攻击弹切换成最优追踪策略,由于

表1 攻击弹首先逃逸时的仿真参数

Table 1 Simulation parameters while attacking missile evades first

仿真参数	飞行器	防御弹	攻击弹
初始位置/m	(0,0)	(0,0)	(9000, -200)
初始航向/(°)	60	20	0
速度/(m·s <sup>-1</sup> )	300	800	700

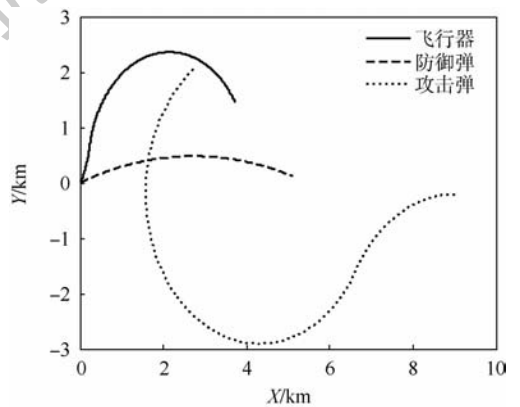


图7 攻击弹首先逃逸时的三方运动轨迹 ( $u_T^{\max}$  大于其最小逃逸机动)

Fig. 7 Trajectories of three players while attacking missile evades first ( $u_T^{\max}$  is larger than minimal evasion maneuver)

飞行器满足实现逃逸的最小机动条件,新的碰撞三角形无法建立, $t_{goMT}$ 与时间 $t$ 仍然不具有线性关系。

2) 假设飞行器、防御弹和攻击弹的最大加速度指令分别为  $u_T^{\max} = 50 \text{ m/s}^2$ 、 $u_D^{\max} = 80 \text{ m/s}^2$  和  $u_M^{\max} = 200 \text{ m/s}^2$ 。基于上述参数,在仿真过程中可以发现飞行器的控制指令不满足式(36)的约束,此时三方的运动轨迹如图9所示。从图9可以看出,由于攻击弹的最大加速度指令增大,导致飞行器的最大加速度不满足式(36)所示的最小机动条件,使得飞行器不能逃脱攻击弹的拦截(攻击弹对飞行器最终的脱靶量为  $y_{MT} = 4.481 \text{ m} < R_M$ )。

图10所示为2个剩余对策时间随时间的变化关系。由图中可以看出,由于飞行器不满足实现逃逸的最小机动条件,虽然在对策初期  $t_{goMT}$  与时间  $t$  不具有线性关系,但经过一段时间的过渡之后建立了新的碰撞三角形,从而有了近似的线性关系,最终飞行器逃逸失败。

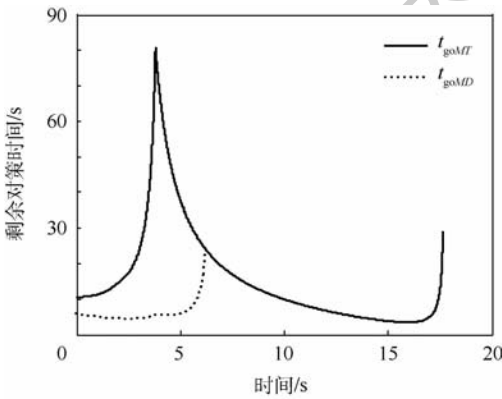


图8 攻击弹首先逃逸时2个剩余对策时间的变化 ( $u_T^{\max}$  大于其最小逃逸机动)

Fig. 8 Evolution of two time-to-go while attacking missile evades first ( $u_T^{\max}$  is larger than minimal evasion maneuver)

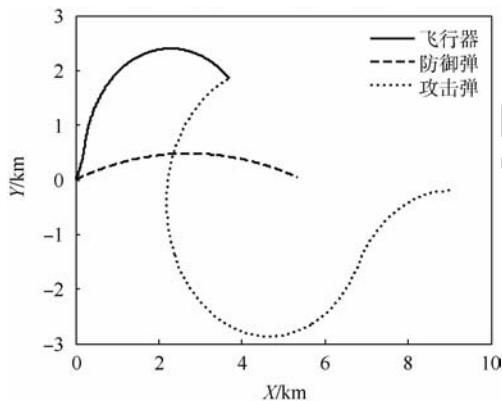


图9 攻击弹首先逃逸时的三方运动轨迹 ( $u_T^{\max}$  小于其最小逃逸机动)

Fig. 9 Trajectories of three players while attacking missile evades first ( $u_T^{\max}$  is smaller than minimal evasion maneuver)

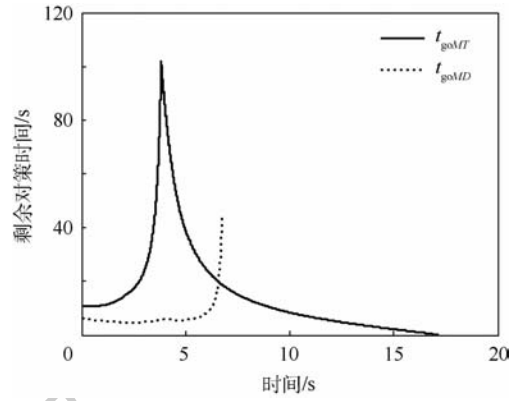


图10 攻击弹首先逃逸时2个剩余对策时间的变化 ( $u_T^{\max}$  小于其最小逃逸机动)

Fig. 10 Evolution of two time-to-go while attacking missile evades first ( $u_T^{\max}$  is smaller than minimal evasion maneuver)

### 4.2 攻击弹直接采用最优追踪策略

此时攻击弹只采用如式(37)所示的追踪策略,飞行器和防御弹的制导策略仍如式(24)所示,三方的其他相关参数如表2所示。

为了验证防御弹能否拦截攻击弹,在仿真中应根据所给参数由式(42)进行实时判断。

1) 假设飞行器、防御弹和攻击弹的最大加速度指令分别为  $u_T^{\max} = 50 \text{ m/s}^2$ 、 $u_D^{\max} = 80 \text{ m/s}^2$  和  $u_M^{\max} = 160 \text{ m/s}^2$ 。基于上述参数,在仿真过程中可以发现防御弹的加速度指令满足式(42)的约束,此时三方的运动轨迹如图11所示。从图11可以看出,防御弹在攻击弹命中飞行器之前成功将攻击弹拦截(脱靶量为  $y_{MD} = 5.902 \text{ m} < R_D$ )。图12为2个剩余对策时间随时间变化的示意图。可以看出,2个剩余对策时间与时间几乎呈线性关系,而剩余对策时间  $t_{goMD}$  小于  $t_{goMT}$ ,防御弹在攻击弹命中飞行器之前对其实施了拦截。

2) 假设飞行器、防御弹和攻击弹的最大加速度指令分别为  $u_T^{\max} = 40 \text{ m/s}^2$ 、 $u_D^{\max} = 40 \text{ m/s}^2$  和  $u_M^{\max} = 120 \text{ m/s}^2$ 。基于上述参数,在仿真过程中可以发现防御弹的加速度指令不满足式(42)的约束,此时三方的运动轨迹如图13所示。可以看出,防御弹未能在攻击弹命中飞行器之前将其拦截(防御弹对攻击弹脱靶量为  $y_{MD} = 15.199 \text{ m} > R_D$ )。图14为2个剩余对策时间与时间的变化关系。

表2 攻击弹直接追踪时的仿真参数

Table 2 Simulation parameters for attacking missile pursues directly

仿真参数	飞行器	防御弹	攻击弹
初始位置/m	(0,0)	(0,0)	(3000, -200)
初始航向/(°)	12	7	0
速度/(m·s <sup>-1</sup> )	300	800	700

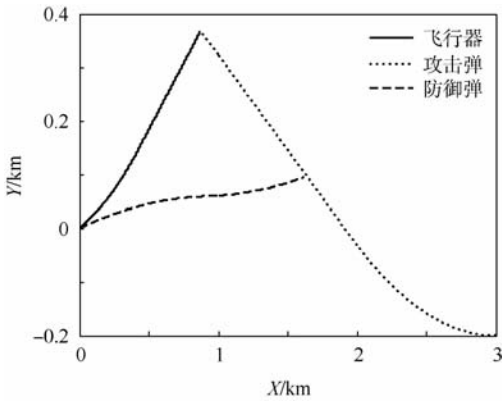


图 11 攻击弹直接追踪时的三方运动轨迹 ( $u_d^{\max}$  大于其最小拦截机动)

Fig.11 Trajectories of three players while attacking missile pursues directly ( $u_d^{\max}$  is larger than minimal interception maneuver)

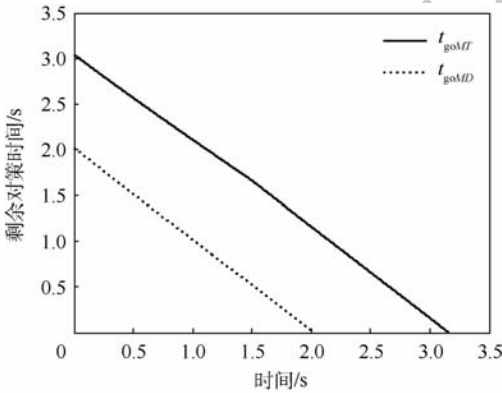


图 12 攻击弹直接追踪时 2 个剩余对策时间的变化 ( $u_d^{\max}$  大于其最小拦截机动)

Fig.12 Evolution of two time-to-go while attacking missile pursues directly ( $u_d^{\max}$  is larger than minimal interception maneuver)

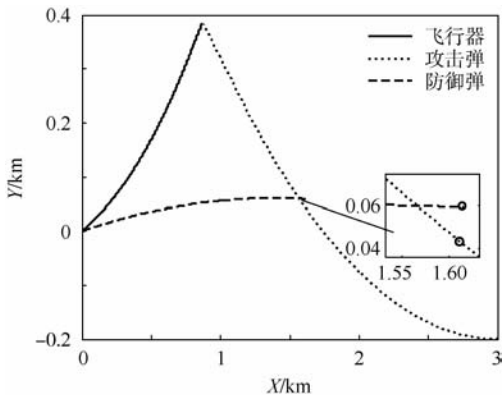


图 13 攻击弹直接追踪时的三方运动轨迹 ( $u_d^{\max}$  小于其最小拦截机动)

Fig.13 Trajectories of three players while attacking missile pursues directly ( $u_d^{\max}$  is smaller than minimal interception maneuver)

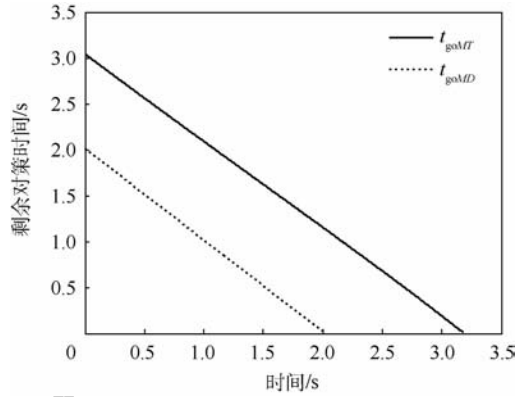


图 14 攻击弹直接追踪时 2 个剩余对策时间的变化 ( $u_d^{\max}$  小于其最小拦截机动)

Fig.14 Evolution of two time-to-go while attacking missile pursues directly ( $u_d^{\max}$  is smaller than minimal interception maneuver)

## 5 结 论

1) 针对具有主动防御能力的飞行器面对来袭导弹发射一枚防御弹进行防御的制导问题,给出了一种基于范数型性能指标的最优制导律。

2) 对攻击弹采用不同制导策略的情况,分别推导了飞行器能够逃离攻击弹和防御弹能够拦截攻击弹的条件。

3) 攻击弹采取先逃逸后追踪的策略时,飞行器若满足其最小逃逸机动条件则可逃脱攻击弹的拦截,否则飞行器将被攻击弹拦截。

4) 攻击弹直接采取追踪策略时,防御弹若满足其最小拦截机动条件则可拦截攻击弹,否则防御弹将无法在飞行器被命中前将攻击弹拦截。

## 参考文献 (References)

[ 1 ] SHAFERMAN V, SHIMA T. Cooperative multiple-model adaptive guidance for an aircraft defending missile [ J ]. Journal of Guidance, Control, and Dynamics, 2010, 33 ( 6 ) : 1801-1813.

[ 2 ] SHIMA T. Optimal cooperative pursuit and evasion strategies against a homing missile [ J ]. Journal of Guidance, Control, and Dynamics, 2011, 34 ( 2 ) : 414-425.

[ 3 ] PROKOPOV O, SHIMA T. Linear quadratic optimal cooperative strategies for active aircraft protection [ J ]. Journal of Guidance, Control, and Dynamics, 2013, 36 ( 3 ) : 753-764.

[ 4 ] RATNOO A, SHIMA T. Line-of-sight interceptor guidance for defending an aircraft [ J ]. Journal of Guidance, Control, and Dynamics, 2011, 34 ( 2 ) : 522-532.

[ 5 ] YAMASAKI T, BALAKRISHNAN S N, TAKANO H. Modified command to line-of-sight intercept guidance for aircraft defense [ J ]. Journal of Guidance, Control, and Dynamics, 2013, 36 ( 3 ) : 898-902.

[ 6 ] RATNOO A, SHIMA T. Guidance strategies against defended aerial targets [ J ]. Journal of Guidance, Control, and Dynamics, 2012, 35 ( 4 ) : 1059-1068.

- [ 7 ] BOYELL R L. Defending a moving target against missile or torpedo attack[J]. IEEE Transactions on Aerospace and Electronic Systems, 1976, 12(4): 522-526.
- [ 8 ] PERELMAN A, SHIMA T, RUSNAK I. Cooperative differential games strategies for active aircraft protection from a homing missile[J]. Journal of Guidance, Control, and Dynamics, 2011, 34(3): 761-773.
- [ 9 ] GARCIA E, CASBEER D W, PACTHER M. Cooperative strategies for optimal aircraft defense from an attacking missile[J]. Journal of Guidance, Control, and Dynamics, 2015, 38(8): 1510-1520.
- [10] GARCIA E, CASBEER D W, PACTHER M. Active target defense differential game; Fast defender case [ J ]. IET Control Theory and Applications, 2017, 11(17): 2985-2993.
- [11] GARCIA E, CASBEER D W, FUCHS C Z, et al. Cooperative missile guidance for active defense of air vehicles [ J ]. IEEE Transactions on Aerospace and Electronic Systems, 2018, 54(2): 706-721.
- [12] PACTHER M, GARCIA E, CASBEER D W. Active target defense differential game [ C ] // 52nd Annual Allerton Conference on Communication, Control, and Computing. Piscataway, NJ: IEEE Press, 2014: 46-53.
- [13] RUSNAK I. The lady, the bandits and the body-guard game [ C ] // The 44th Israel Annual Conference on Aerospace Sciences. Haifa: Israel Institute of Technology, 2004: 25-26.
- [14] RUSNAK I. The lady, the bandits and the body guards-A two team dynamic game [ C ] // Proceedings of the 16th World IFAC Congress. Prague: IFAC, 2005: 441-446.
- [15] RUSNAK I. Games based guidance in anti-missile defense for high order participants [ C ] // Proceeding of the 15th IEEE Mediterranean Electrotechnical Conference. Piscataway, NJ: IEEE Press, 2010: 812-817.
- [16] RUBINSKY S, GUTMAN S. Three body guaranteed pursuit and evasion; AIAA-2012-4910 [ R ]. Reston: AIAA, 2012.
- [17] RUBINSKY S, GUTMAN S. Three-player pursuit and evasion conflict [ J ]. Journal of Guidance, Control, and Dynamics, 2014, 37(1): 98-110.
- [18] 孙启龙, 齐乃明, 赵钧, 等. 攻击主动防御飞行器的微分对策制导律 [ J ]. 国防科技大学学报, 2018, 40(3): 7-14.  
SUN Q L, QI N M, ZHAO J, et al. Differential game guidance laws against active defense aircraft [ J ]. Journal of National University of Defense Technology, 2018, 40(3): 7-17 (in Chinese).
- [19] SUN Q L, QI N M, XIAO L X, et al. Differential game strategy in three-player evasion and pursuit scenarios [ J ]. Journal of Systems Engineering and Electronics, 2018, 29(2): 352-366.
- [20] TURETSKY V, SHINAR J. Missile guidance laws based on pursuit-evasion game formulations [ J ]. Automatica, 2003, 39(4): 607-618.

## 作者简介:

郭志强 男, 博士研究生。主要研究方向: 飞行器制导与控制。

孙启龙 男, 博士, 工程师。主要研究方向: 飞行器制导与控制。

周绍磊 男, 博士, 教授, 博士生导师。主要研究方向: 飞行器制导与控制。

闫实 男, 博士研究生。主要研究方向: 飞行器编队控制。

## Norm differential game guidance law for active defense aircraft

GUO Zhiqiang<sup>1</sup>, SUN Qilong<sup>2</sup>, ZHOU Shaolei<sup>1,\*</sup>, YAN Shi<sup>1</sup>

(1. Coastal Defense Academy, Naval Aviation University, Yantai 264001, China;

2. Beijing Institute of Space Long March Vehicle, Beijing 100076, China)

**Abstract:** For the pursuit-evasion problems between an active defense aircraft which can launch a defending missile from itself and an attacking missile, a guidance law for aircraft and defending missile is derived and analyzed based on the differential game theory. First, Optimal guidance strategies with bounded lateral controls of aircraft, defending missile and attacking missile are derived for the three players by using a norm performance index. Second, the conditions of a successful evasion for the aircraft and a successful interception for the defender are deduced, and the minimal evasion maneuver of the aircraft and the minimal interception maneuver of the defender are obtained. Finally, Nonlinear simulations are carried out to validate the guidance law proposed. It is verified that the aircraft can evade the attacking missile if its maneuver is equal or greater than the minimal evasion maneuver, and the defender can intercept the attacking missile if its maneuver is equal or greater than the minimal interception maneuver.

**Keywords:** aircraft; active defense; norm; differential game; guidance law

**Received:** 2018-12-18; **Accepted:** 2019-01-18; **Published online:** 2019-01-23 09:53

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20190121.1340.004.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190121.1340.004.html)

**Foundation item:** National Natural Science Foundation of China (61273058)

\* **Corresponding author.** E-mail: zhousld@sina.com

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

DOI: 10.13700/j.bh.1001-5965.2018.0746

# 贮箱爆炸碎片初始速度及影响因素

赵蓓蕾<sup>1</sup>, 赵继广<sup>2</sup>, 崔村燕<sup>1,\*</sup>, 王岩<sup>1</sup>, 孙武华<sup>3</sup>

(1. 航天工程大学 宇航科学与技术系, 北京 101416; 2. 航天工程大学 电子与光学工程系, 北京 101416;

3. 中国人民解放军 63618 部队, 库尔勒 841001)

**摘 要:** 为确定推进剂爆轰作用下贮箱爆炸碎片的初始速度, 基于能量守恒定律, 考虑爆炸碎片动能、爆轰产物动能和内能、贮箱壳体的破坏能及其膨胀做功所消耗的能量, 建立了贮箱爆炸碎片初始速度 (FIV) 模型。FIV 模型与典型经验公式计算结果、带壳炸药爆炸试验数据吻合较好, 验证了模型有效性。采用量纲分析法确定 FIV 模型中影响碎片初始速度的关键参量, 基于 AUTODYN 软件进行数值仿真, 分析贮箱壳体高径比、厚径比以及空气密度等参量对碎片初始速度的影响。结果表明: 爆炸碎片初始速度随着壳体高径比增大迅速减小, 当高径比大于 1.50 时, 速度衰减变缓; 碎片初始速度随着壳体厚径比增加近似呈线性减小; 当爆炸高度小于 20 km 时, 随着爆炸高度增大, 空气密度减小, 爆炸碎片的初始速度增大; 在爆炸高度大于 40 km 时, 空气非常稀薄, 可以忽略壳体膨胀做功对碎片初始速度的影响。

**关键词:** 贮箱; 爆炸碎片; 初始速度; 影响因素; 数值仿真

**中图分类号:** V423.4<sup>+</sup>1

**文献标识码:** A **文章编号:** 1001-5965(2019)09-1797-08

推进剂爆炸产生高温高压的爆轰产物, 爆轰产物迫使贮箱壳体向外膨胀, 膨胀到一定程度时, 壳体出现裂纹, 裂纹相互贯通后破裂成碎片。碎片在爆轰产物的冲击作用下开始运动, 此时碎片的速度称为初始速度<sup>[1]</sup>。爆炸碎片的初始速度是衡量碎片抛射距离及其对目标毁伤效果的重要参数, 它不仅影响碎片的飞散范围, 还影响碎片对周围设施设备和人员的损伤程度。

在化工安全领域, 爆炸碎片特性的研究主要利用 Monte-Carlo 方法仿真储罐爆炸碎片抛射轨迹<sup>[2]</sup>、构建爆炸碎片冲击失效模型<sup>[3]</sup>、预测沸腾液体膨胀蒸汽爆炸 (BLEVE) 事故中碎片的初始速度等<sup>[4]</sup>。在战斗部爆破领域, Gurney<sup>[5]</sup>以炸药质量和壳体质量的比值为基本参数提出了求解碎片飞散速度的 Gurney 公式, 该公式应用最广泛; 工程上, 相关学者基于炸药爆热提出了计算碎片

速度的半经验公式<sup>[6]</sup>, 该公式对于单质炸药的计算精度较好, 而混合炸药的计算误差较大; 印立魁和蒋建伟<sup>[7]</sup>基于 Gurney 公式建立了多层球形战斗部爆炸破片初始速度的计算模型, 给出了各层破片的平均初始速度。在航空航天领域, 王卫杰等<sup>[8]</sup>针对液体火箭芯一级贮箱爆炸, 推导了爆炸碎片初始速度 (Fragment Initial Velocity, FIV) 模型, 模型中忽略了贮箱壳体的破坏能和传给空气介质的能量; 美国国家航空航天局 (NASA) 基于图表法提出了求解柱形和球形高压容器爆炸碎片初始速度的方法<sup>[9]</sup>, 但该方法对容器内气体爆轰压强有一定的要求, 适用范围受到限制。

本文基于能量守恒定律, 考虑爆炸碎片动能、爆轰产物动能和内能、贮箱壳体的破坏能及其膨胀做功所消耗的能量, 构建了贮箱爆炸 FIV 模型。基于量纲分析法, 确定影响碎片初始速度的关键

收稿日期: 2018-12-18; 录用日期: 2019-03-15; 网络出版时间: 2019-04-03 16:50

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190330.1212.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190330.1212.001.html)

\* 通信作者。E-mail: cey6655@126.com

**引用格式:** 赵蓓蕾, 赵继广, 崔村燕, 等. 贮箱爆炸碎片初始速度及影响因素[J]. 北京航空航天大学学报, 2019, 45(9): 1797-1804. ZHAO B L, ZHAO J G, CUI C Y, et al. Initial velocity and influence factors of tank explosion fragments[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1797-1804 (in Chinese).

参量,利用 AUTODYN 软件中的光滑粒子流体动力学 (Smooth Particle Hydrodynamics, SPH) 算法进行数值仿真,分析贮箱壳体高径比、厚径比以及大气密度等对爆炸碎片初始速度的影响。

## 1 爆炸碎片初始速度模型

贮箱爆炸 FIV 模型基于以下假设:不考虑爆轰反应的过程,将化学反应区看作一个强间断面;将贮箱看作等壁厚的圆柱壳体,只考虑爆炸产物沿径向的飞散;推进剂全部参与爆轰反应,爆炸碎片初始速度的大小相同。

根据能量守恒定律,推进剂爆炸释放的总能量  $E_i$  主要用于以下几部分:

$$E_i = E_{dk} + E_k + E_i + E_d + E_a \quad (1)$$

式中:  $E_{dk}$  为贮箱爆炸碎片的动能;  $E_k$  为爆轰产物的动能;  $E_i$  为爆轰产物的内能;  $E_d$  为贮箱壳体的破坏能;  $E_a$  为贮箱壳体膨胀做功。

### 1) 推进剂爆炸释放的总能量

液体推进剂爆炸释放的总能量  $E_i$  为

$$E_i = m_p Q_p \quad (2)$$

式中:  $m_p$  为推进剂质量;  $Q_p$  为推进剂爆热。

### 2) 爆炸碎片的动能

设爆炸一共产生的碎片数量为  $N$ , 碎片初始速度为  $v_0$ , 由动能定律得

$$E_{dk} = \sum_{i=1}^N \frac{m_i}{2} v_0^2 = \frac{1}{2} M v_0^2 \quad (3)$$

式中:  $m_i$  为第  $i$  个碎片的质量;  $M$  为贮箱壳体的质量。

### 3) 爆轰产物的动能

若推进剂全部参与反应,根据质量守恒,爆轰产物的质量等于推进剂的质量,则爆轰产物动能  $E_k$  为

$$E_k = \frac{1}{2} m_p u_p^2 \quad (4)$$

式中:  $u_p$  为爆轰产物的质点速度。

爆轰物质点速度  $u_p$  与推进剂的爆速  $D$  之间的关系为<sup>[10]</sup>

$$u_p = \frac{D}{\gamma + 1} \quad (5)$$

式中:  $\gamma$  为爆轰产物的多方指数,  $\gamma = 3$ 。

将式(5)代入式(4)得

$$E_k = \frac{1}{32} m_p D^2 \quad (6)$$

### 4) 爆轰产物的内能

单位质量爆轰产物的比内能  $e_i$ 、爆热  $Q_p$  和爆速  $D$  之间的近似关系为<sup>[8]</sup>

$$e_i = Q_p - \left(\frac{D}{4}\right)^2 \quad (7)$$

则爆轰产物的内能  $E_i$  为

$$E_i = m_p e_i = m_p Q_p - \frac{1}{16} m_p D^2 \quad (8)$$

### 5) 贮箱壳体的破坏能

Hiroe<sup>[11]</sup>、胡永乐<sup>[12]</sup>、俞鑫炉<sup>[13]</sup> 等分别研究了内部爆炸载荷作用下,不同金属材料(铝、钢、钛合金)圆柱壳体的动态变形和破坏模式。结果表明,在冲击波加载下,不同材料的圆柱壳体因径向膨胀发生动态塑性变形,产生剪切带,主要呈现出剪切型断裂特征。

若只考虑贮箱壳体在爆炸载荷作用下的剪切断裂,则贮箱壳体的破坏能  $E_d$  等于剪切变形能:

$$E_d = \frac{1}{2} \tau \eta V_c \quad (9)$$

由剪切虎克定律得

$$\tau = G \eta \quad (10)$$

将式(10)代入式(9)得

$$E_d = \frac{\tau^2 V_c}{2G} \quad (11)$$

式中:  $\tau$  为剪切应力;  $\eta$  为剪应变,无量纲;  $G$  为剪切弹性模量;  $V_c$  为贮箱壳体的体积。

### 6) 贮箱壳体膨胀做功

爆炸瞬间,在空气中形成的冲击波阵面压力  $p_a$  为

$$p_a = \frac{k+1}{2} \rho_a v_a^2 \quad (12)$$

式中:  $k$  为空气的等熵绝热指数,对于强冲击波,  $k = 1.2$ ;  $\rho_a$  为未受冲击波扰动的空气密度;  $v_a$  为波阵面的空气质点速度。

空气冲击波的初始速度  $D_a$  接近于推进剂的爆速  $D$ , 而  $v_a$  与  $D_a$  的关系式为

$$v_a = \frac{2}{k+1} D_a = \frac{2}{k+1} D \quad (13)$$

将式(13)代入(12)得

$$p_a = \frac{10}{11} \rho_a D^2 \quad (14)$$

当地空气密度  $\rho_a$  随海平面高度  $y$  变化的一般表达式为

$$\rho_a = \rho_0 H(y) \quad (15)$$

式中:  $\rho_0$  为海平面的空气密度,  $\rho_0 = 1.293 \text{ kg/m}^3$ ;  $H(y)$  为空气密度修正系数<sup>[1]</sup>:

$$H(y) = \begin{cases} \left(1 - \frac{y}{44.308}\right)^{4.2553} & y \leq 11 \text{ km} \\ 0.297 e^{-\frac{y-11}{6.318}} & y > 11 \text{ km} \end{cases} \quad (16)$$

设爆轰瞬间作用于贮箱壳体上的空气冲击波阵面压力  $p_a$  保持不变,则贮箱壳体等压膨胀时克服空气冲击波压力所做的功  $E_a$ ,即

$$E_a = \int_{V_0}^{V_1} p_a dV = p_a (V_1 - V_0) \quad (17)$$

式中: $V_0$ 为贮箱的初始体积; $V_1$ 为贮箱完全膨胀后的体积。

根据剪切断裂判据<sup>[14]</sup>,壳体完全膨胀后的半径  $r_f$  和初始半径  $r_0$  的关系式为

$$r_f = r_0 \left( \frac{\sqrt{3} p_c}{\sigma} \right)^{\frac{1}{k_r}} \quad (18)$$

式中: $p_c$ 为爆轰产物的压力; $\sigma$ 为贮箱材料的屈服应力; $k_r$ 为修正系数,对于圆柱形壳体, $k_r = 2$ ,对于球形壳体, $k_r = 3$ 。

爆轰产物的压力  $p_c$  与爆速  $D$  之间的关系为

$$p_c = \frac{1}{\gamma + 1} \rho_p D^2 = \frac{1}{4} \rho_p D^2 \quad (19)$$

式中: $\rho_p$ 为推进剂的密度。

将式(14)、式(15)、式(18)和式(19)代入式(17)可得

$$E_a = \left( \frac{10}{11} \rho_0 H(y) D^2 V_0 \right) \left[ \left( \frac{\sqrt{3} \rho_p D^2}{4\sigma} \right)^{\frac{1}{3}} - 1 \right] \quad (20)$$

### 7) 爆炸碎片的初始速度

将式(2)、式(3)、式(6)、式(8)、式(11)和式(20)代入式(1),整理得爆炸碎片初始速度  $v_0$  为

$$v_0 = \left\{ \frac{1}{32} m_p D^2 - \left( \frac{10}{11} \rho_0 H(y) D^2 V_0 \right) \cdot \left[ \left( \frac{\sqrt{3} \rho_p D^2}{4\sigma} \right)^{\frac{1}{3}} - 1 \right] - \frac{\tau^2 V_c}{2G} \right\} / \frac{M}{2} \quad (21)$$

式(21)即为本文建立的贮箱爆炸 FIV 模型。

## 2 FIV 模型验证

### 2.1 计算碎片初始速度的经验公式

#### 1) Gurney 公式

Gurney<sup>[5]</sup>针对不同形状的战斗部提出了预测爆炸碎片初始速度的经验公式,对于圆柱形壳体来说,有

$$v_0 = \sqrt{2E_g} \sqrt{\frac{\beta}{1 + 0.5\beta}} \quad (22)$$

式中: $\sqrt{2E_g}$ 为 Gurney 常数,Gurney<sup>[5]</sup>通过试验得出  $\sqrt{2E_g}$  与爆速  $D$  成线性关系, $\sqrt{2E_g} = 520 + 0.28D$ ; $\beta = m_c/M$ ,表示固体炸药质量  $m_c$  与壳体质量  $M$  的比值。利用式(22)计算液体推进剂爆炸,需

要先将液体推进剂质量  $m_p$  换算为固体炸药质量  $m_c$ 。

#### 2) 工程半经验公式

工程上经常应用半经验公式计算碎片的初始速度为<sup>[6]</sup>

$$v_0 \approx 1.237 \sqrt{Q_p / \left( \frac{1}{\beta} + \frac{1}{2} \right)} \quad (23)$$

#### 3) 王卫杰模型

王卫杰等<sup>[8]</sup>将火箭贮箱看作圆柱形壳体,推导了爆炸碎片初始速度的计算公式:

$$v_0 = \frac{D}{4} \sqrt{\beta / \left( \frac{1}{2} + \frac{\beta}{4} \right)} \quad (24)$$

### 2.2 基于爆炸试验的模型验证

武汉理工大学孔祥韶等<sup>[15]</sup>开展了三硝基甲苯(TNT)炸药在 Q235 低碳钢圆柱壳体内部的爆炸试验,壳体内径为 110 mm,高度为 160 mm,厚度为 6 mm,装药量为 1.9 kg。对于 TNT 炸药, $\rho_p = 1630 \text{ kg/m}^3$ , $D = 6900 \text{ m/s}$ 。

北京理工大学张奇等<sup>[16]</sup>开展了 T-H-Al 炸药在钢制圆柱壳体内部的爆炸试验,壳体内填满炸药,壳体内径为 120 mm,厚度为 3.5 mm。对于 T-H-Al 炸药, $\rho_p = 1752 \text{ kg/m}^3$ , $D = 7054 \text{ m/s}$ 。

劳伦斯-利弗莫尔实验室开展了 Comp B 炸药在铜制圆柱壳体内部的爆炸试验<sup>[17]</sup>,壳体内填满炸药,壳体内径为 25.4 mm,厚度为 2.6 mm。对于 Comp B 炸药, $\rho_p = 1720 \text{ kg/m}^3$ , $D = 7980 \text{ m/s}$ 。

将文献[15-17]试验的相关参数分别代入 FIV 模型、Gurney 公式、工程半经验公式、王卫杰模型开展计算,计算结果、试验结果及两者之间相对误差  $\varepsilon$  如表 1、表 2 和表 3 所示。

表 1 文献[15]爆炸试验数据及其理论计算结果

Table 1 Explosion test data in Ref. [15] and theoretical calculation results

理论模型	$v_0 / (\text{m} \cdot \text{s}^{-1})$	$\varepsilon / \%$
FIV 模型	1356	12.3
Gurney 公式	1486	23.1
工程半经验公式	1621	34.3
王卫杰模型	1469	21.7

注:文献[15]中  $v_0 = 1207 \text{ m/s}$ 。

表 2 文献[16]爆炸试验数据及其理论计算结果

Table 2 Explosion test data in Ref. [16] and theoretical calculation results

理论模型	$v_0 / (\text{m} \cdot \text{s}^{-1})$	$\varepsilon / \%$
FIV 模型	1630	11.4
Gurney 公式	1608	9.9
工程半经验公式	1749	19.5
王卫杰模型	1647	12.6

注:文献[16]中  $v_0 = 1463 \text{ m/s}$ 。

表3 文献[17]爆炸试验数据及其理论计算结果  
Table 3 Explosion test data in Ref. [17] and theoretical calculation results

理论模型	$v_0/(m \cdot s^{-1})$	$\varepsilon/\%$
FIV 模型	1461	14.1
Gurney 公式	1482	15.8
工程半经验公式	1616	26.3
王卫杰模型	1490	16.4

注:文献[17]中  $v_0 = 1280 m/s$ 。

由表1、表2和表3可知,FIV模型计算的爆炸碎片初始速度与3组试验结果之间相对误差均小于15%,在误差允许范围内,与Gurney公式计算结果也吻合较好,验证了FIV模型的有效性。由于FIV模型考虑了贮箱的壳体变形和膨胀做功,因此计算结果比王卫杰模型更为准确。

### 3 爆炸碎片初始速度仿真分析

#### 3.1 碎片初始速度影响参量的量纲分析

由FIV模型可知,影响爆炸碎片初始速度 $v_0$ 的参量包括:①推进剂方面:推进剂质量 $m_p$ 、爆速 $D$ ;②贮箱壳体方面:壳体质量 $M$ 、内径 $d$ 、高度 $h$ 、厚度 $b$ 、屈服应力 $\sigma$ 、剪切应力 $\tau$ 、剪切弹性模量 $G$ ;③周围环境方面:空气密度 $\rho_a$ 。

影响爆炸碎片初始速度的参量较多,若依次对其分析,工作量巨大,有必要利用量纲分析法确定其中的关键影响参量,开展针对性分析。

爆炸碎片初始速度 $v_0$ 可写成函数表达式:

$$v_0 = f(m_p, D, M, d, h, b, \sigma, \tau, G, \rho_a) \quad (25)$$

选取 $D$ 、 $M$ 和 $d$ 为基本量,根据量纲分析 $\pi$ 定理对式(25)进行无量纲化,得

$$\frac{v_0}{D} = \varphi\left(\frac{m_p}{M}, \frac{h}{d}, \frac{b}{d}, \frac{\sigma}{M^{-1}d^{-1}D^2}, \frac{\tau}{M^{-1}d^{-1}D^2}, \frac{G}{M^{-1}d^{-1}D^2}, \frac{\rho_a}{Md^{-3}}\right) \quad (26)$$

若壳体的材料特性不变,式(26)可以简化为

$$v_0 = D\varphi\left(\frac{m_p}{M}, \frac{h}{d}, \frac{b}{d}, \frac{\rho_a}{Md^{-3}}\right) \quad (27)$$

将液体推进剂质量 $m_p$ 换算为固体炸药质量 $m_e$ ,式(27)可以写为

$$v_0 = D\varphi_1\left(\beta, \frac{h}{d}, \frac{b}{d}, \frac{\rho_a}{Md^{-3}}\right) \quad (28)$$

Gurney公式描述的是爆炸碎片初始速度 $v_0$ 与爆速 $D$ 、炸药质量与壳体质量比值 $\beta$ 的关系。但通过对FIV模型参量进行量纲分析可知,爆炸碎片初始速度 $v_0$ 还与贮箱壳体高径比 $h/d$ 、厚径

比 $b/d$ 以及空气密度 $\rho_a$ 等因素有关。

#### 3.2 碎片初始速度影响因素的分析方法

卫星作为火箭的有效载荷,位于火箭上面级。卫星贮箱内携带易燃易爆的液体推进剂,在火箭发射过程中任何意外都有可能引发贮箱爆炸。探究卫星贮箱壳体的高径比 $h/d$ 和厚径比 $b/d$ 对爆炸碎片初始速度 $v_0$ 的影响,可以为贮箱结构设计提供参考;探究空气密度 $\rho_a$ 对爆炸碎片初始速度 $v_0$ 的影响,有助于评估不同高度爆炸后碎片对周围设施设备和人员的危害。

若采用理论方法对爆炸碎片初始速度的影响因素开展分析,过程十分复杂;开展爆炸试验危险性较大;而数值仿真具有可重复性好、易操作、成本低等优点。基于传统网格法进行数值仿真,发生畸变的网格单元会被删除,而这些被删除的单元往往处于高应变率、大变形区域,导致无法准确统计爆炸碎片的平均速度。SPH算法是一种无网格法,避免了网格侵蚀的问题,因此,本文基于AUTODYN软件中的SPH算法研究贮箱壳体高径比 $h/d$ 、厚径比 $b/d$ 以及空气密度 $\rho_a$ 对碎片初始速度 $v_0$ 的影响。

#### 3.3 卫星贮箱爆炸仿真模型

液体推进剂爆炸要经历飞溅、雾化、混合,以及高温蒸发形成可燃气体的过程,液相爆炸和气相爆炸同时进行,数值仿真难以复现其复杂的爆炸反应过程,且爆轰瞬间碎片初始速度的大小只与爆炸源能量有关,通常利用TNT炸药的爆炸过程来模拟液体推进剂的爆炸过程。根据能量相似原理将液体推进剂总量 $m_p$ 换算成TNT炸药质量 $m_T$ ,即

$$m_T = Ym_p \quad (29)$$

式中: $Y$ 为液体推进剂的TNT当量系数。

以装载液态肼的某卫星贮箱为研究对象,将贮箱看作等壁厚圆柱壳体,壳体内径为 $d = 50 \text{ cm}$ ,高度为 $h = 100 \text{ cm}$ ,厚度 $b = 2 \text{ mm}$ 。贮箱满载状态下,液态肼质量 $m_p = 198.2 \text{ kg}$ ,TNT当量系数 $Y = 0.93^{[18]}$ ,则TNT炸药质量 $m_T = 184.3 \text{ kg}$ ,空气域为边长 $a = 120 \text{ cm}$ 的立方体。TNT炸药采用JWL(Jones-Wilkins-Lee)状态方程;壳体材料钛合金采用Johnson-Cook本构模型和Stochastic破坏模型,最大失效应变取为 $0.23^{[19]}$ ,其中 $A$ 、 $B$ 、 $C$ 、 $n$ 和 $m$ 为本构模型参数, $\rho$ 为材料密度;空气域采用POLYNOMIAL多项式状态方程,通过改变空气密度来建立贮箱不同高度爆炸的模型。各材料参数见表4。

表 4 TNT、钛和空气的材料参数  
Table 4 Material parameters of TNT, titanium and air

材料	$\rho/(\text{kg} \cdot \text{m}^{-3})$	$Q_p/(\text{kJ} \cdot \text{kg}^{-1})$	$D/(\text{m} \cdot \text{s}^{-1})$	$A/\text{MPa}$	$B/\text{MPa}$	$C$	$n$	$m$
TNT	1 630	4 225	6 900					
钛	4 510			1 077	845	0.025	0.58	0.75
空气	$1.293 \times 10^{-4} \sim 2.8 \times 10^{-4}$							

根据模型的几何对称性,建立卫星贮箱的 1/2 模型,如图 1 所示。贮箱壳体的 SPH 粒子大小为 1 mm,炸药的 SPH 粒子大小为 2 mm,空气的 SPH 粒子大小为 4 mm,在炸药中心设置起爆点。TNT 的爆炸反应大约在  $10^{-6}$  s 数量级时间内完成,经多次仿真计算得到,在  $t = 80 \mu\text{s}$  左右贮箱壳体完全破裂成碎片,因此将仿真计算时间设置为  $80 \mu\text{s}$ 。

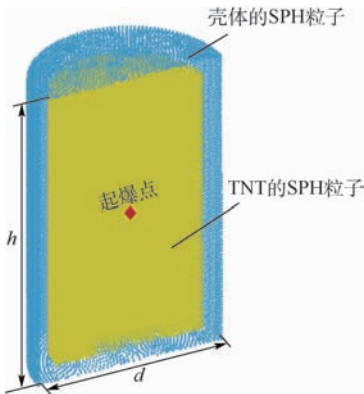


图 1 卫星贮箱的 1/2 仿真模型

Fig.1 1/2 simulation model of satellite tank

### 3.4 仿真结果与分析

为探究爆炸碎片初始速度与贮箱壳体高径比的关系,保持  $m_T = 184.3 \text{ kg}$ 、 $b = 2 \text{ mm}$ 、 $\rho_a = 1.293 \text{ kg/m}^3$  不变,考虑到工程实际,在  $0.41 \sim 5.83$  范围内改变壳体高径比  $h/d$ ,进行 10 组数值仿真。数值计算完成后,通过 AUTODYN 软件后处理统计生成爆炸碎片的平均速度  $\bar{v}_0$ ,将其看

作爆炸碎片的初始速度,结果如表 5 所示。表中,  $v_G$  为利用 Gurney 公式计算的碎片初始速度,  $\varepsilon_G$  为 Gurney 公式与仿真结果之间的相对误差;  $v_F$  为利用 FIV 模型计算的碎片初始速度;  $\varepsilon_F$  为 FIV 模型与仿真结果之间的相对误差。

表 5 中仿真结果均比 Gurney 公式计算结果要小,原因是仿真中爆轰反应的总能量部分用于壳体变形和对外膨胀做功,爆炸碎片动能相应减少;仿真结果与 FIV 模型的计算结果更接近,出现误差的原因是仿真中材料参数设置和状态方程选取与 FIV 模型并不完全一致。仿真结果与理论计算值之间的相对误差均小于 15%,在误差允许范围内,验证了仿真模型的有效性。绘制仿真得到的爆炸碎片初始速度  $\bar{v}_0$  随高径比  $h/d$  变化曲线,结果如图 2 所示。

由图 2 知,当  $h/d < 1.50$  时,爆炸碎片初始速度随着高径比增大而急剧衰减,较小高径比的贮箱内,爆轰产物在轴向上分布较为集中,爆炸瞬间爆轰产物对壳体的冲击作用较大,更多的爆轰产物内能转化为爆炸碎片动能。当  $h/d > 1.50$  时,初始速度的衰减幅度逐渐变缓。虽然贮箱高径比越大,发生意外爆炸后碎片初始速度越小,碎片对周围设施和人员的危害越小,但还要考虑卫星总体设计对贮箱尺寸的要求,合理选取高径比。

为探究爆炸碎片初始速度与贮箱壳体厚径比的关系,保持  $m_T = 184.3 \text{ kg}$ 、 $h = 100 \text{ cm}$ 、 $\rho_a =$

表 5 不同高径比下碎片初始速度的仿真值与理论值

Table 5 Simulation and theoretical values of fragment initial velocity under different height-diameter ratios

序号	$d/\text{cm}$	$h/\text{cm}$	$h/d$	$\bar{v}_0/(\text{m} \cdot \text{s}^{-1})$	$v_G/(\text{m} \cdot \text{s}^{-1})$	$\varepsilon_G/\%$	$v_F/(\text{m} \cdot \text{s}^{-1})$	$\varepsilon_F/\%$
1	35	204.1	5.83	2 804.5	3 138.2	11.9	2 674.7	-4.63
2	40	156.2	3.91	2 817.6	3 174.9	12.7	2 754.8	-2.23
3	45	123.5	2.74	2 831.4	3 203.6	13.1	2 882.3	1.80
4	50	100.0	2.00	2 848.5	3 227.6	13.3	2 951.8	3.63
5	55	82.6	1.50	2 867.3	3 247.6	13.3	3 036.2	5.89
6	60	69.4	1.16	2 898.4	3 264.3	12.6	3 134.9	8.16
7	65	59.2	0.91	2 925.0	3 278.7	12.1	3 179.1	8.69
8	70	51.0	0.73	2 943.9	3 291.4	11.8	3 249.6	10.4
9	75	44.4	0.59	2 955.6	3 302.3	11.7	3 286.6	11.2
10	80	39.1	0.49	2 962.2	3 311.7	11.8	3 296.5	11.3
11	85	34.6	0.41	2 966.8	3 320.4	11.9	3 301.4	11.3

1.293 kg/m<sup>3</sup> 不变,考虑到工程实际,在 0.001 ~ 0.01 范围内改变壳体厚径比  $b/d$  进行数值仿真。图 3 为爆炸碎片初始速度  $\bar{v}_0$  随壳体厚径比  $b/d$  的变化曲线。

由图 3 知,随着壳体厚径比  $b/d$  的增加,爆炸碎片的初始速度近似呈线性减小趋势。在爆炸总能量一定的情况下,壳体厚径比  $b/d$  越大,用于壳体变形的能量越多,爆炸碎片的动能相应地减小,爆炸碎片初始速度减小。工程实际中需要结合贮箱对壳体质量和耐压性能的要求,合理选取壳体厚度。

空气密度随着海拔高度升高而减小,因此贮箱在不同高度爆炸时,壳体膨胀做功不同,爆炸碎片的初始速度也不同。为探究爆炸碎片初始速度  $v_0$  与爆炸高度  $y$ 、空气密度  $\rho_a$  的关系,保持贮箱壳体内径  $d = 50$  cm,高度  $h = 100$  cm,厚度  $b = 2$  mm 不变,根据式 (15) 和式 (16),在  $y = 0 \sim 50$  km 范围内每间隔 5 km 改变爆炸高度,计算不同高度对

应的空气密度  $\rho_a$ ,对贮箱在不同密度的空气中爆炸进行数值仿真。图 4 为空气密度  $\rho_a$  和爆炸碎片初始速度  $\bar{v}_0$  随高度  $y$  的变化曲线。

由图 4 知,当  $y < 20$  km 时,随着爆炸高度增加,空气密度大幅衰减,爆炸碎片初始速度迅速增长,这是由于在爆炸总能量一定的情况下,爆炸高度越高,空气密度越小,贮箱壳体向外膨胀克服空气冲击波压力做功越少,爆炸碎片的动能增加,碎片初始速度增大。当  $y > 20$  km 时,碎片初始速度的增长趋于平缓。当  $y > 40$  km 时,空气已经非常稀薄,可以忽略贮箱壳体的膨胀做功对爆炸碎片初始速度的影响。

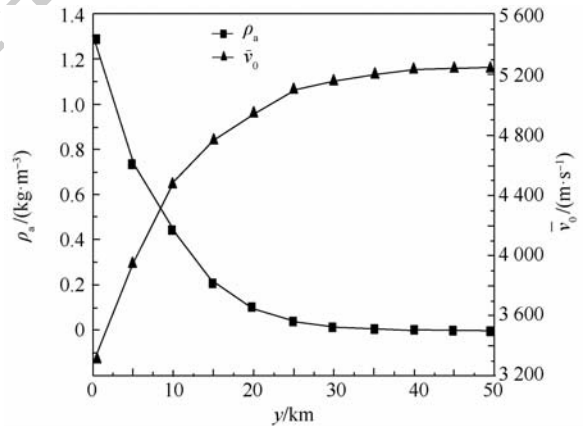


图 4 空气密度和碎片初始速度随高度的变化曲线  
Fig. 4 Curves of air density and fragment initial velocity varying with height

## 4 结论

1) 基于能量守恒定律,构建了贮箱爆炸 FIV 模型;采用量纲分析法确定 FIV 模型中影响碎片初始速度的关键参量为壳体高径比  $h/d$ 、厚径比  $b/d$  以及空气密度  $\rho_a$ 。

2) 爆炸碎片的初始速度随着壳体高径比  $h/d$  增大而减小。当  $h/d < 1.50$  时,初始速度迅速衰减,当  $h/d > 1.50$  时,衰减幅度减小。工程实际中要结合卫星总体设计对贮箱尺寸的要求,合理选取高径比。

3) 爆炸碎片的初始速度随着壳体厚径比  $b/d$  增加近似呈线性减小趋势。工程实际中要结合卫星贮箱对壳体质量和耐压性能的要求,合理选取厚径比。

4) 随着爆炸高度  $y$  增加,空气密度  $\rho_a$  大幅衰减,爆炸碎片的初始速度增大;当  $y > 20$  km 时,初始速度增长趋于稳定;当  $y > 40$  km 时,可以忽略壳体膨胀做功对初始速度的影响。

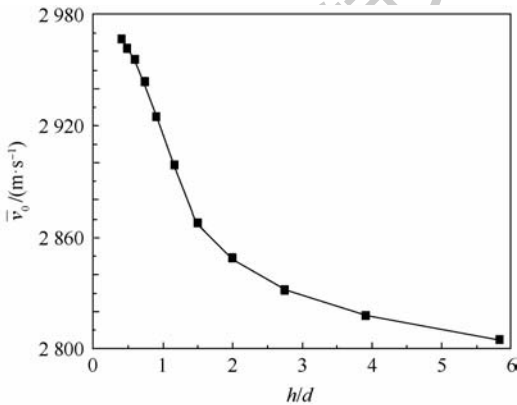


图 2 碎片初始速度随壳体高径比的变化曲线

Fig. 2 Curve of fragment initial velocity varying with height-diameter ratio of shell

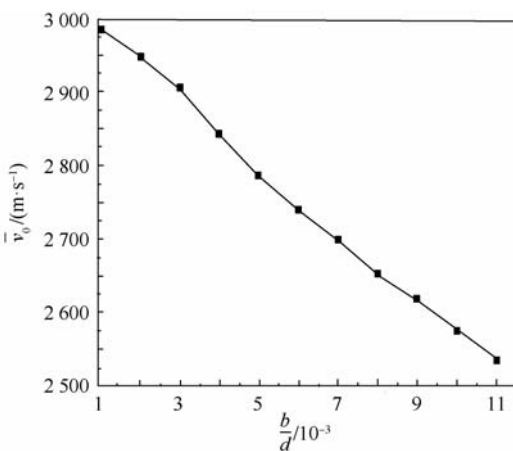


图 3 碎片初始速度随壳体厚径比的变化曲线

Fig. 3 Curve of fragment initial velocity varying with thickness-diameter ratio of shell

## 参考文献 (References)

- [1] 陈新华, 聂万胜. 液体推进剂爆炸危害性评估方法及应用 [M]. 北京: 国防工业出版社, 2005: 226-233.  
CHEN X H, NIE W S. Evaluation method and application of liquid propellant explosion harmfulness [M]. Beijing: National Defense Industry Press, 2005: 226-233 (in Chinese).
- [2] 邢志祥, 蒋军成, 赵晓芳. 液化石油气储罐爆炸碎片抛射的蒙特卡罗分析 [J]. 火灾科学, 2004, 13(1): 39-42.  
XING Z X, JIANG J C, ZHAO X F. Monte-Carlo analysis of the flight of missiles from LPG tank explosion [J]. Fire Safety Science, 2004, 13(1): 39-42 (in Chinese).
- [3] 左哲. 基于蒙特卡罗法的储罐爆炸碎片冲击失效模型研究 [J]. 中国安全科学学报, 2012, 22(3): 67-72.  
ZUO Z. Study on models of tank explosion fragments impact failure based on Monte Carlo method [J]. China Safety Science Journal, 2012, 22(3): 67-72 (in Chinese).
- [4] 潘科, 许开立. 储罐爆炸事故中抛射碎片初始速度预测 [J]. 东北大学学报 (自然科学版), 2014, 35(7): 1047-1050.  
PAN K, XU K L. Prediction of the fragments initial velocity in tank explosions accident [J]. Journal of Northeastern University (Natural Science), 2014, 35(7): 1047-1050 (in Chinese).
- [5] GURNEY R W. The initial velocities of fragments from bombs, shells and grenades; BRL report 405 [R]. Aberdeen: Ballistic Research Laboratory, 1943.
- [6] 张守中. 爆炸基本原理 [M]. 北京: 国防工业出版社, 1988: 503-504.  
ZHANG S Z. Basic principles of explosion [M]. Beijing: National Defense Industry Press, 1988: 503-504 (in Chinese).
- [7] 印立魁, 蒋建伟. 多层球形预制破片战斗部破片初速场的计算模型 [J]. 含能材料, 2014, 22(3): 300-305.  
YIN L K, JIANG J W. Calculation model of initial velocity field on multilayered spherical fragments warhead [J]. Chinese Journal of Energetic Materials, 2014, 22(3): 300-305 (in Chinese).
- [8] 王卫杰, 沈怀荣, 李怡勇, 等. 液体火箭爆炸碎片模型研究 [J]. 上海航天, 2013, 30(6): 35-38.  
WANG W J, SHEN H R, LI Y Y, et al. Study of liquid rocket explosion fragments model [J]. Aerospace Shanghai, 2013, 30(6): 35-38 (in Chinese).
- [9] BAKER W E, KULESZ J J, RICHKER R E, et al. Workbook for predicting pressure wave and fragment effects of exploding propellant tanks and gas storage vessels; NASA-CR-34906 [R]. Washington, D. C.: NASA, 1977.
- [10] 恽寿榕, 赵衡阳. 爆炸力学 [M]. 北京: 国防工业出版社, 2005: 18-20.  
YUN S R, ZHAO H Y. Explosion mechanics [M]. Beijing: National Defense Industry Press, 2005: 18-20 (in Chinese).
- [11] HIROE T, FUJIWARA K, HATA H, et al. Deformation and fragmentation behavior of exploded metal cylinders and the effects of wall materials, configuration, explosive energy and initiated locations [J]. International Journal of Impact Engineering, 2008, 35(12): 1578-1586.
- [12] 胡永乐, 陈子辰, 王峰超, 等. 内部爆炸加载条件下圆柱钢壳的动态断裂 [J]. 机械强度, 2010, 32(1): 99-104.  
HU Y L, CHEN Z C, WANG F C, et al. Dynamic fracture of cylindrical steel shell under inside-explosion loading [J]. Journal of Mechanical Strength, 2010, 32(1): 99-104 (in Chinese).
- [13] 俞鑫炉, 董新龙, 潘顺吉. 不同爆炸载荷下 TA2 钛合金圆管膨胀破坏过程 [J]. 爆炸与冲击, 2018, 38(1): 148-154.  
YU X L, DONG X L, PAN S J. Fracture behaviors of explosively driven TA2 alloy cylinders under different loadings [J]. Explosion and Shock Waves, 2018, 38(1): 148-154 (in Chinese).
- [14] HOGGATT C, RECHT R. Fracture behavior of tubular bombs [J]. Journal of Applied Physics, 1968, 39(3): 1856.
- [15] 孔祥韶, 吴卫国, 杜志鹏, 等. 圆柱形战斗部爆炸破片特性研究 [J]. 工程力学, 2014, 31(1): 243-249.  
KONG X S, WU W G, DU Z P, et al. Research on fragments characteristic of cylindrical warhead [J]. Engineering Mechanics, 2014, 31(1): 243-249 (in Chinese).
- [16] ZHANG Q, MIAO C Q, LIN D C, et al. Relation of fragment with air shock wave intensity for explosion in a shell [J]. International Journal of Impact Engineering, 2003, 28: 1129-1141.
- [17] KARPP R R, PREDEBON W W. Calculations of fragment velocities from naturally fragmenting munitions; ADB007377 [R]. Aberdeen: Ballistic Research Laboratory, 1979.
- [18] JULIEN H L, WOODS S S, RATHGEBER K, et al. Explosive events initiated by pyrovalves; AIAA-1999-2309 [R]. Reston: AIAA, 1999.
- [19] 刘旭阳. TC4 钛合金动态本构关系研究 [D]. 南京: 南京航空航天大学, 2010.  
LIU X Y. Dynamic constitutive relationship of TC4 titanium alloy [D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2010 (in Chinese).

## 作者简介:

赵蓓蕾 女, 博士研究生。主要研究方向: 航天任务分析与设计。

崔村燕 女, 博士, 副教授, 硕士生导师。主要研究方向: 航天任务分析与设计。

## Initial velocity and influence factors of tank explosion fragments

ZHAO Beilei<sup>1</sup>, ZHAO Jiguang<sup>2</sup>, CUI Cunyan<sup>1,\*</sup>, WANG Yan<sup>1</sup>, SUN Wuhua<sup>3</sup>

(1. Department of Aerospace Science and Technology, Space Engineering University, Beijing 101416, China;

2. Department of Electronic and Optical Engineering, Space Engineering University, Beijing 101416, China;

3. 63618 Troops of the PLA, Korla 841001, China)

**Abstract:** To determine the initial velocity of tank explosion fragments under the propellant detonation, the fragment initial velocity (FIV) model was established based on the energy conservation law, in which the kinetic energy of explosion fragments, the kinetic energy and internal energy of detonation products, the failure energy and the consumed energy for expansion work of tank shell were considered. The FIV model was in good agreement with the calculation results of typical empirical formulas and the experimental data, which verifies the effectiveness of the model. Based on the dimensional analysis method, the key parameters affecting the initial velocity were determined. Based on AUTODYN software, numerical simulation was conducted and the effects of height-diameter ratio, thickness-diameter ratio and air density on fragment initial velocity were analyzed. Results show that the initial velocity of explosion fragment decreases rapidly with the increase of height-diameter ratio, and the attenuation trend slows down when the height-diameter ratio exceeds 1.50. The initial velocity almost linearly decreases with the increase of thickness-diameter ratio. When the explosion height is less than 20 km, as the explosion height rises, the air density decreases, and the initial velocity increases. The air becomes very thin above 40 km, and the influence of shell expansion work on initial velocity can be neglected.

**Keywords:** tank; explosion fragment; initial velocity; influence factors; numerical simulation

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

DOI: 10.13700/j.bh.1001-5965.2018.0752

# Dirichlet 混合样本的 EM 算法与动态聚类算法比较

夏棒<sup>1,\*</sup>, EMILION Richard<sup>2</sup>, 王惠文<sup>3</sup>

(1. 中国工商银行 博士后科研工作站, 北京 100032; 2. 奥尔良大学 MAPMO 研究所, 奥尔良 45000;  
3. 北京航空航天大学 经济管理学院, 北京 100083)



**摘 要:** Dirichlet 分布是一类包含正参数向量的连续多元概率分布, 在比例结构问题中具有广泛的应用。针对 Dirichlet 混合样本的聚类问题, 进行了最大期望(EM)算法和动态聚类算法研究。首先, 推导其数学过程, 并给出算法迭代步骤。然后, 利用数字仿真实验, 比较了 EM 算法与动态聚类算法两种机器学习算法在 Dirichlet 混合样本中的聚类效果。最后, 计算对数似然函数值、程序运行时间、收敛迭代次数、聚类正确率、真正率(TPR)和假正率(FPR)6个评价指标。仿真实验结果表明, EM 算法聚类正确率更高但是运算效率相对较低, 而动态聚类算法运算效率较高但是损失了部分聚类正确率。因此, 实际应用中建议综合权衡聚类正确率与运算效率的相对需求后, 再选取合适算法进行 Dirichlet 混合样本聚类。

**关键词:** Dirichlet 分布; 混合样本; 最大期望(EM)算法; 动态聚类; 机器学习

**中图分类号:** O212

**文献标识码:** A      **文章编号:** 1001-5965(2019)09-1805-07

在概率统计中, Dirichlet 分布是一类包含正参数向量的连续多元概率分布。现实生活中很多常见的数据都满足 Dirichlet 分布, 例如三次产业比例结构数据, 如果各个产业比例之和为 1, 那么就服从一个 Dirichlet 分布。该分布被广泛应用在比例结构数据问题当中。

对 Dirichlet 混合分布进行估计, 并且对 Dirichlet 混合样本进行精确类别划分具有十分重要的意义。Fraley 和 Raftery<sup>[1]</sup>对一些经典的混合分布数据, 例如高斯混合分布, 总结了其混合模型的参数估计方法。在一系列的研究中, 主要有 2 种不同的方法被用于解决混合模型的分解问题。

最广泛运用的一种方法是把混合分布的分解问题转化为一个估计问题, 目标是估计出混合分布模型的参数, 通常利用极大似然估计方法。1977 年, Dempster<sup>[2]</sup>提出的最大期望(Expectation

Maximization, EM)算法便是解决此类问题的基础优化算法。在此基础上, 多种针对不同情形的 EM 算法的更新版本或适应版本被陆续提出, 包括 Celeux 和 Govaert<sup>[3-4]</sup>提出的随机 EM(Stochastic EM, SEM)算法、分类 EM(Classification EM, CEM)算法、Tanner 和 Wong<sup>[5]</sup>提出的蒙特卡罗 EM(Monte Carlo EM, MCEM)算法, 以及 Render<sup>[6]</sup>提出的一些其他版本, 更多具体内容介绍可以参考 Peel 和 McLachlan<sup>[7]</sup>的研究。

另一种方法则是基于聚类算法的思想。这些方法考虑将包含  $n$  个观测的混合样本划分为  $K$  个类别, 其中每个类别的观测都服从于一个共同的概率密度分布。这一类方法中最基础的为动态聚类算法<sup>[8-10]</sup>。Celeux 和 Govaert<sup>[4]</sup>正是将这类方法和 EM 算法的理念相融合才提出了 CEM 算法。

收稿日期: 2018-12-25; 录用日期: 2019-03-15; 网络出版时间: 2019-04-17 13:30

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190415.1822.005.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190415.1822.005.html)

基金项目: 国家自然科学基金(71420107025)

\* 通信作者: E-mail: 15011176214@163.com

**引用格式:** 夏棒, EMILION Richard, 王惠文. Dirichlet 混合样本的 EM 算法与动态聚类算法比较[J]. 北京航空航天大学学报, 2019, 45(9): 1805-1811. XIA B, EMILION R, WANG H W. Comparison between EM algorithm and dynamical clustering algorithm for Dirichlet mixture samples[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1805-1811 (in Chinese).

其他较为经典的方法还有迭代重定位算法<sup>[8,11]</sup>、层次分类算法<sup>[12]</sup>、神经网络算法<sup>[13-14]</sup>、重叠分类算法(例如可加聚类算法<sup>[15]</sup>、金字塔算法<sup>[16]</sup>和函数聚类模型<sup>[17-18]</sup>)。诸多近期研究还将 Copula 函数应用于混合模型之中<sup>[19-21]</sup>。

根据对文献的整体梳理,EM 算法和动态聚类算法是 2 种处理混合分布问题的核心方法。然而,已有研究大多针对的是高斯分布等经典分布类型,其分布形式较为简单。而对于 Dirichlet 混合分布的研究少有提及,然而这类分布在处理比例结构数据时具有重大价值。因此,本文提出 Dirichlet 混合样本的 EM 算法和动态聚类算法,并通过数字仿真实验对 2 种算法的效果进行了比较,从而总结出 2 种算法各自的优势及其适用的情形。

## 1 两种机器学习聚类算法

令随机向量  $\mathbf{p} = [p_1, p_2, \dots, p_K]$  的各个元素值都大于 0 且总和为 1,这样  $p_k (k=1, 2, \dots, K)$  就代表第  $k$  项的比例。在具有参数向量  $\alpha$  的 Dirichlet 分布中,  $\mathbf{p}$  的概率密度为

$$P(\mathbf{p}) \sim \text{Dir}(\alpha_1, \alpha_2, \dots, \alpha_K) = \frac{\Gamma\left(\sum_{k=1}^K \alpha_k\right)}{\prod_{k=1}^K \Gamma(\alpha_k)} \prod_{k=1}^K p_k^{\alpha_k-1} \quad (1)$$

式中:  $p_k > 0$ , 且  $\sum_{k=1}^K p_k = 1$ 。

### 1.1 Dirichlet 混合样本的 EM 算法

在统计领域中,EM 算法通常被用来寻找依赖于不可观察的隐性变量的概率模型中参数的最大似然估计,是一种迭代算法。其在高斯混合模型中已经得到了广泛的应用。由于 Dirichlet 分布较为复杂,EM 算法在 Dirichlet 混合分布中的应用较少。接下来,将介绍 Dirichlet 混合样本的 EM 算法推导过程。

设  $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$  为一个 Dirichlet 混合样本,该样本数据由  $m$  个 Dirichlet 分布产生,每个 Dirichlet 分布包含  $k$  个参数。令  $\mathbf{z} = [z_1, z_2, \dots, z_n]$  为隐藏类别标签变量,该标签决定了观测与 Dirichlet 分布的对应关系。因此,有

$$X_i | (z_i = c) \sim \text{Dir}(\alpha_{c1}, \alpha_{c2}, \dots, \alpha_{ck}) \quad (2)$$

式中:类别参数  $c = 1, 2, \dots, m, P(z_i = c) = \tau_c$  且  $\sum_{c=1}^m \tau_c = 1$ 。

本文目标是估计出 Dirichlet 混合分布的未知参数  $\theta = (\tau_c, \alpha_{cj}), c = 1, 2, \dots, m; j = 1, 2, \dots, k$ 。

对于该 Dirichlet 混合样本,其完整数据的似然函数为

$$L(\theta; \mathbf{x}, \mathbf{z}) = \prod_{i=1}^n \prod_{c=1}^m [\tau_c f(\mathbf{x}_i; \alpha_c)]^{I(z_i=c)} \quad (3)$$

式中:  $\alpha_c = [\alpha_{c1}, \alpha_{c2}, \dots, \alpha_{ck}]$  为第  $c$  类样本的 Dirichlet 参数向量;  $I$  为指示性函数。

结合 Dirichlet 分布的概率密度函数:

$$f(\mathbf{x}_i; \alpha_c) = \frac{\Gamma\left(\sum_{j=1}^k \alpha_{cj}\right)}{\prod_{j=1}^k \Gamma(\alpha_{cj})} \prod_{j=1}^k x_{ij}^{\alpha_{cj}-1} \quad (4)$$

可以得到全数据对数似然函数,即

$$\lg L(\theta; \mathbf{x}, \mathbf{z}) = \sum_{i=1}^n \sum_{c=1}^m I(z_i=c) \cdot \left[ \lg \tau_c + \lg \Gamma\left(\sum_{j=1}^k \alpha_{cj}\right) - \sum_{j=1}^k \lg \Gamma(\alpha_{cj}) + \sum_{j=1}^k (\alpha_{cj}-1) \lg x_{ij} \right] \quad (5)$$

在 E 步中,已知第  $t$  次迭代得到的参数估计值  $\theta^{(t)}$ ,  $z_j$  的条件分布可以按照贝叶斯定理得到,其计算公式为

$$T_{ci}^{(t)} = \frac{\tau_c^{(t)} f(\mathbf{x}_i, \theta_c^{(t)})}{\sum_{j=1}^m \tau_j^{(t)} f(\mathbf{x}_i, \theta_j^{(t)})} \quad (6)$$

此外,计算完整数据对数似然函数的期望  $Q$  为

$$Q(\theta | \theta^{(t)}) = E_{\mathbf{z} | \mathbf{x}, \theta^{(t)}} [\lg L(\theta; \mathbf{x}, \mathbf{z})] = \sum_{i=1}^n E_{z_i | \mathbf{x}_i, \theta^{(t)}} [\lg L(\theta; \mathbf{x}_i, z_i)] = \sum_{i=1}^n \sum_{c=1}^m T_{ci}^{(t)} \left[ \lg \tau_c + \lg \Gamma\left(\sum_{j=1}^k \alpha_{cj}\right) - \sum_{j=1}^k \lg \Gamma(\alpha_{cj}) + \sum_{j=1}^k (\alpha_{cj}-1) \lg x_{ij} \right] \quad (7)$$

可以看出,  $\tau_c$  和  $\alpha_{cj}$  出现在  $Q$  的不同线性项中,因此可以分别通过极大化求导,得出迭代估计值,这也是 M 步的主要内容。

在 M 步中,与其他混合分布模型相同,权重参数  $\tau_c$  的计算公式为

$$\tau_c^{(t+1)} = \frac{1}{n} \sum_{i=1}^n T_{ci}^{(t)} \quad (8)$$

而对于参数  $(\alpha_{c1}, \alpha_{c2}, \dots, \alpha_{ck})$ ,其估计值的更新公式为

$$(\alpha_{c1}^{(t+1)}, \alpha_{c2}^{(t+1)}, \dots, \alpha_{ck}^{(t+1)}) = \arg \max_{\alpha_{cj_0}} Q(\theta | \theta^{(t)}) = \arg \max_{\alpha_{cj_0}} \sum_{i=1}^n T_{ci}^{(t)} \left[ \lg \Gamma\left(\sum_{j=1}^k \alpha_{cj}\right) - \lg \Gamma(\alpha_{cj_0}) + \alpha_{cj_0} \lg x_{ij_0} \right] \quad (9)$$

对于混合模型中的每一个 Dirichlet 分布,可

以得到  $k$  个包含  $k$  个参数的方程,通过求解联立方程组,就可以得到每个 Dirichlet 分布参数  $(\alpha_{c_1}, \alpha_{c_2}, \dots, \alpha_{c_k}), c = 1, 2, \dots, m$  的估计值。

设  $\varepsilon$  为精度参数,当  $E_{z|x,\theta^{(t)}}[\lg L(\theta^{(t)}; \mathbf{x}, \mathbf{z})] - E_{z|x,\theta^{(t-1)}}[\lg L(\theta^{(t-1)}; \mathbf{x}, \mathbf{z})] \leq \varepsilon$  时,迭代过程终止。

Dirichlet 样本的 EM 算法的计算过程可以归纳为如下 4 个步骤:

**步骤 1** 随机初始化 Dirichlet 分布参数  $\alpha^{(0)}$  和权重参数  $\tau^{(0)}$ 。同时,设置精度参数  $\varepsilon$ 。

**步骤 2** 根据式(6)和式(8),计算出  $\tau^{(1)}$ 。

**步骤 3** 计算函数  $Q$ 。通过式(7)和式(9),可得到 Dirichlet 分布参数下一步估计  $\alpha^{(1)}$ 。计算  $E_{z|x,\theta^{(t)}}[\lg L(\theta^{(t)}; \mathbf{x}, \mathbf{z})] - E_{z|x,\theta^{(t-1)}}[\lg L(\theta^{(t-1)}; \mathbf{x}, \mathbf{z})]$ ,与  $\varepsilon$  进行比较。

**步骤 4** 重复进行步骤 2 和步骤 3,直到迭代到第  $t$  次时,2 次期望对数似然函数值之差小于精度参数值,即  $E_{z|x,\theta^{(t)}}[\lg L(\theta^{(t)}; \mathbf{x}, \mathbf{z})] - E_{z|x,\theta^{(t-1)}}[\lg L(\theta^{(t-1)}; \mathbf{x}, \mathbf{z})] < \varepsilon$ ,此时迭代达到了终止条件,计算过程结束。

## 1.2 Dirichlet 混合样本的动态聚类算法

动态聚类是一种快速估计出样本观测所属类别的常用方法。它的核心思想是在初始化观测所属类别后,通过计算似然值进行逐步迭代,不断更新观测类别和类分布参数,直到迭代收敛。接下来,本文将给出 Dirichlet 样本的动态聚类算法。

给定一个 Dirichlet 混合样本,样本包含  $n$  个独立观测  $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$ ,这些观测来自  $m$  个 Dirichlet 分布。每一步迭代过程的目的是估计出每个 Dirichlet 概率密度函数的参数,使得按照该类别划分计算得到的聚类准则参数例如似然值达到最优。

对于每个类别,需要估计其分布参数  $(\alpha_{c_1}, \alpha_{c_2}, \dots, \alpha_{c_k})$ 。得到  $m$  个 Dirichlet 概率密度函数后,可以计算出每个观测的  $m$  个概率密度值。通过极大化联合概率密度,每个观测被归为不同的类别。这个过程得到了一个新的类别划分。动态聚类则是重复该过程直至对数似然函数达到收敛临界值。对数似然函数方程为

$$LL = \sum_{i=1}^n \lg \sum_{c=1}^m p_c f_c(\mathbf{x}_i) \quad (10)$$

式中:  $p_c$  为每个类别的权重;  $f_c(\cdot)$  为每个类别的 Dirichlet 密度函数。

Dirichlet 样本的动态聚类算法的计算过程可以归纳为如下 4 个步骤:

**步骤 1** 随机初始化原始观测的类别划分。

将  $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$  划分为  $m$  个类别,得到类别划分  $\mathbf{C}^{(0)} = [C_1^{(0)}, C_2^{(0)}, \dots, C_m^{(0)}]$ 。

**步骤 2** 根据  $m$  个类别中每一类别的观测,分别估计出  $m$  个 Dirichlet 密度函数  $f_1^{(0)}(\cdot), f_2^{(0)}(\cdot), \dots, f_m^{(0)}(\cdot)$  的参数值。然后计算这一步的对数似然函数值。

**步骤 3** 对于每一个观测  $\mathbf{x}_i$ ,代入到  $m$  个 Dirichlet 密度函数中,得到  $f_1^{(0)}(\mathbf{x}_i), f_2^{(0)}(\mathbf{x}_i), \dots, f_m^{(0)}(\mathbf{x}_i)$ 。然后将观测划分到 Dirichlet 密度值最大的类别。遍历  $n$  个观测后,得到新的类别划分  $\mathbf{C}^{(1)} = [C_1^{(1)}, C_2^{(1)}, \dots, C_m^{(1)}]$ 。

**步骤 4** 重复步骤 2 和步骤 3,当 2 次迭代过程的对数似然函数值之差的绝对值  $|\text{LL}^{(s)} - \text{LL}^{(s-1)}| < \varepsilon$  时,动态聚类算法终止。其中,  $\varepsilon$  为预设的精度参数。

## 2 数字仿真实验

本节将生成 Dirichlet 混合样本,并运用提出的 Dirichlet EM 算法和动态聚类算法分别对混合样本进行聚类,然后比较 2 种算法聚类效果。仿真实验过程的实现全部基于 R 统计分析软件。

### 2.1 仿真数据的生成

为了便于计算,生成样本量为 1 000 的 Dirichlet 样本,样本中的每个观测为一个二维向量,向量元素值介于 0 到 1 之间,且各元素之和为 1。并且 1 000 个观测分别源自 2 个不同的类别  $C_1$  和  $C_2$ 。  $C_1$  包含 300 个观测,而  $C_2$  包含 700 个观测。此外,  $C_1$  和  $C_2$  分别满足不同的 Dirichlet 分布,其中  $\mathbf{x}_i \in C_1 \sim \text{Dir}(2, 3), \mathbf{x}_j \in C_2 \sim \text{Dir}(7, 4)$ 。精度参数  $\varepsilon = 0.001$ ,即当 2 次迭代的对数似然函数值之差小于 0.001 时,算法终止。

### 2.2 评价指标

共提出 6 个评价指标用于 2 种算法的比较。第 1 个指标为对数似然函数值,该指标衡量不同迭代次数间对原始样本拟合水平的差异。第 2 个指标为程序运行时间,第 3 个指标为收敛迭代次数,这 2 个指标共同衡量算法的运行效率。运行时间越短,迭代次数越少,则算法的运行效率越高。第 4 个指标为聚类正确率,用于衡量 2 种算法得到的类别划分结果与真实类别划分结果的一致程度,正确率越高,则与真实类别的一致程度越高。这里需要指出的是 EM 算法实质上是一种模糊聚类算法,它给出的是每一个观测属于各个类别的概率值,概率值大小反映观测属于该类别的倾向大小。因此,用 EM 算法进行类别划分,默认

的前提假设是将每个观测归到概率值最大的类别。最后 2 个指标为真正率 (True Positive Rate, TPR) 和假正率 (False Positive Rate, FPR), 用以检测非平衡二分类样本中正负样本的识别效率。TPR 为正确预测的正样本占有所有正样本的比例, 该指标值越大越优。FPR 为被错分为正样本的负样本占有所有负样本的比例, 该指标值越小越优。

### 2.3 仿真实验过程

一共设计 2 个数字仿真实验, 第 1 个仿真实验比较 2 种机器学习算法在单个 Dirichlet 混合样本中的聚类效果, 第 2 个仿真实验则通过多次重复实验, 削弱偶然因素影响, 进而比较 2 种算法的聚类效果。具体操作如下:

#### 1) 仿真实验一

按照 2.1 节的描述生成一个样本量为 1000 的 Dirichlet 混合样本, 其中 300 个观测满足 Dirichlet 分布  $\text{Dir}(2, 3)$ , 剩余 700 个观测满足 Dirichlet 分布  $\text{Dir}(7, 4)$ , 然后比较 EM 算法和动态聚类算法在 6 个评价指标上的效果。

#### 2) 仿真实验二

按照仿真实验一, 生成 100 个 Dirichlet 混合样本, 并重复仿真实验一 100 次, 然后计算 100 次实验得到的 6 个评价指标的平均值。对于每次仿真实验, 2 种算法都应用于同一个 Dirichlet 混合样本。

## 3 结果与讨论

首先, 比较仿真实验一各评价指标的结果。图 1 分别展示了 EM 算法和动态聚类算法的对数似然函数曲线随迭代次数的变化情况。

EM 算法与动态聚类 2 种算法的对数似然函数曲线都呈现单调递增趋势。300 次迭代终止时, 动态聚类算法对数似然函数值为 239.36, EM 算法的对数似然函数值为 304.15。此外, 二者的迭代次数也不同, EM 算法运行了 190 次后达到收敛精度, 而动态聚类算法只经过 13 次迭代后就达到所需要的收敛精度。

在程序运行时间上, EM 算法一次迭代平均耗时为 0.022 s, 而动态聚类一次迭代平均耗时为 0.068 s。但是, 由于达到预设精度所需要的迭代次数不同, 当达到迭代终止条件时, EM 算法需要 3.61 s 的运行时间, 而动态聚类算法具有更优的效果, 只需要 1.19 s 运行时间。

对于类别划分结果, 表 1 展示了 2 种算法的聚类结果。根据表中结果, 计算出 EM 算法的正确率为 79.9%, 而动态聚类算法的正确率为

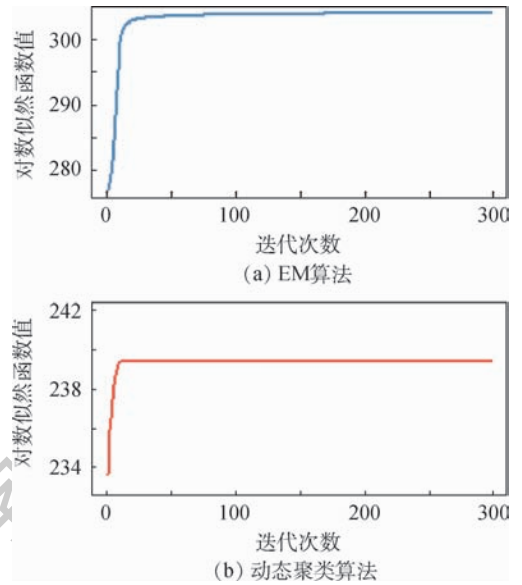


图 1 EM 算法与动态聚类算法在 300 次迭代过程中的对数似然函数曲线

Fig. 1 Log-likelihood function curves of EM algorithm and dynamical clustering algorithm during 300 times of iteration

表 1 2 种算法的聚类结果

Table 1 Clustering results of two algorithms

算法	真实类别	仿真类别	
		1	2
EM 算法	1	172	128
	2	73	627
动态聚类算法	1	235	65
	2	241	459

69.4%。这表明 EM 算法的整体预测准确性要优于动态聚类算法。

以  $C_1$  为正样本, 计算出 EM 算法的 TPR 和 FPR 分别为 57.3% 和 10.4%, 而动态聚类算法的 TPR 和 FPR 分别为 78.3% 和 34.4%。这表明 2 种算法呈现出不同的聚类特征。EM 算法的 FPR 更小, 对  $C_2$  具有相对更佳的识别效率, 而动态聚类的 TPR 更大, 对类别 1 具有相对更佳的识别效率。

接下来, 生成 100 个 Dirichlet 混合样本, 对于每个样本分别进行第 1 个仿真实验, 得到 100 组仿真实验结果。为了比较 2 种算法, 分别计算各评价指标的 100 次仿真平均值, 结果见表 2。

从表 2 可以看出, 在程序运行时间、收敛迭代次数和 TPR 3 个指标上, 动态聚类算法的表现要优于 EM 算法, 而对于聚类正确率和 FPR 2 个评价指标, EM 算法表现更优。这一结论表明, 针对 Dirichlet 混合样本, 动态聚类算法能够快速给出样本类别划分结果, 但是聚类正确率稍有牺牲, 而在时间要求较为宽松的情况下 EM 算法则能够给

出更加精确的类别划分结果。此外,EM算法的平均对数似然函数值为289.22,而动态聚类算法的平均对数似然函数值为214.83。

进一步比较2种算法的TPR与FPR指标,基本与仿真实验一的结果保持一致。EM算法的FPR较小,表明第2类样本被错误预测为第1类的情况越少,即对类别2的识别效率更优。而动态聚类算法的TPR较大,表明被正确预测的第1类样本更多,即对类别1的识别效率更优。

在达到收敛所需的迭代次数分布上,EM算法与动态聚类算法二者都呈现出正偏态特征,如图2所示。这表明在多次重复实验中,少数实验的随机参数初始值与最优值偏离较大,从而降低了算法优化效率,使得迭代次数增加,程序运行时间变长。根据仿真实验结果,以随机初始值与真实值的绝对距离为评价指标,比较迭代次数处于前25%与后25%的仿真实验在该指标上的差异。计算结果表明,在EM算法中前者的该项指标比后者要高出29.2%,在动态聚类算法中要高出

表2 仿真实验二两种算法的结果比较

Table 2 Comparison of two algorithms' results for the second simulation experiment

评价指标	EM算法	动态聚类算法
程序运行时间/s	2.106	1.390
收敛迭代次数	121.51	20.25
聚类正确率/%	80.5	71.3
TPR/%	59.9	77.3
FPR/%	8.6	31.1

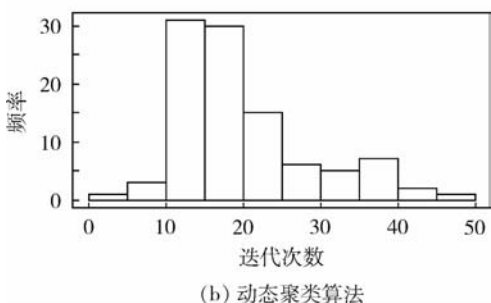
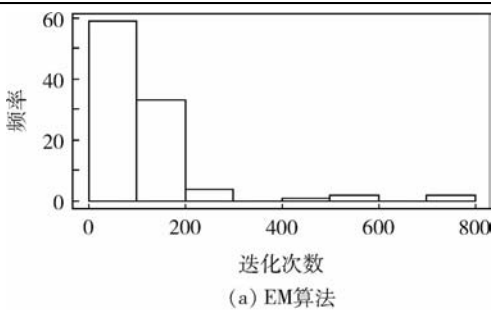


图2 EM算法与动态聚类算法的迭代次数直方图

Fig. 2 Histogram of iteration times for EM algorithm and dynamical clustering algorithm

10.9%。即迭代随机初始值离真实值偏差较大时,算法需要更多的迭代次数才能达到收敛状态。

但是,整体上EM算法的迭代次数集中在0~200,而动态聚类算法的迭代次数集中在10~20,后者的计算效率明显高于前者。需要说明的是,EM算法和动态聚类算法都是局部优化算法,最终聚类结果对各观测初始类别设定有较强的依赖,并且2种算法对各个参数的随机初始值也较为敏感。因此,在使用过程中,建议多次实验,每次进行不同的参数初始化设置,再根据轮廓系数等评价指标选取理想的聚类结果。

## 4 结论

本文针对分布形式较为复杂的Dirichlet混合样本,提出了2种聚类算法的应用,分别为EM算法和动态聚类算法,并给出EM算法的数学推导过程以及2种算法的迭代流程。

1) 仿真实验结果表明,EM算法聚类正确率更高但是运算效率相对较低,而动态聚类算法运算效率较高但是损失了部分正确率。因此在实际运用过程中,需要根据运算环境以及精度等客观需求,选取合适的算法对样本进行聚类。如果对结果的准确度要求较高,则使用EM算法;如果对运算速度要求较高,则使用动态聚类算法。

2) 需要指出的是,EM算法和动态聚类算法作为局部优化算法,对迭代初始值具有较强的敏感性,因此在算法运行过程中,建议多次运行选取效果较好的一次作为最终结果,聚类效果评价可参考轮廓系数、组间离差平方和、组内离差平方和等指标。

另外,现实情况下,不同观测之间往往具有一定的相关性,未来研究考虑将2种算法拓展到非独立Dirichlet分布情形之中。

## 参考文献 (References)

- [1] FRALEY C, RAFTERY A E. Model-based clustering, discriminant analysis, and density estimation[J]. Publications of the American Statistical Association, 2002, 97(458): 611-631.
- [2] DEMPSTER A P. Maximum likelihood from incomplete data via the EM algorithm[J]. Journal of Royal Statistical Society B, 1977, 39(1): 1-38.
- [3] CELEUX G, GOVAERT G. Stochastic algorithms for clustering[M]. Heidelberg: Physica-Verlag, 1990.
- [4] CELEUX G, GOVAERT G. A classification EM algorithm for clustering and two stochastic versions[J]. Computational Statistics & Data Analysis, 1992, 14(3): 315-332.
- [5] TANNER M A, WONG W H. The calculation of posterior distributions by data augmentation[J]. Publications of the American

- Statistical Association, 1987, 82(398):528-540.
- [6] RENDER E. Mixture densities, maximum likelihood and the EM algorithm[J]. SIAM Review, 1984, 26(2):195-239.
- [7] PEEL D, MCLACHLAN G J. Robust mixture modelling using the t distribution[J]. Statistics and Computing, 2000, 10(4):339-348.
- [8] DIDAY E, SCHROEDER A, OK Y. The dynamic clusters method in pattern recognition [C] // Proceedings of International Federation for Information Processing Congress, 1974:691-697.
- [9] SCOTT A J, SYMONS M J. Clustering methods based on likelihood ratio criteria[J]. Biometrics, 1971, 27(1):387-397.
- [10] SYMONS M J. Clustering criteria and multivariate normal mixtures[J]. Biometrics, 1981, 37(1):35-43.
- [11] WONG M A, LANE T. A kTH nearest neighbour clustering procedure[J]. Journal of the Royal Statistical Society, 1981, 45(3):362-368.
- [12] BROSSIER G. Piecewise hierarchical clustering[J]. Journal of Classification, 1990, 7(2):197-216.
- [13] BISHOP C M. Neural networks for pattern recognition[M]. Oxford:Oxford University Press, 1995.
- [14] BOCK H H. Clustering and neural networks[M] // RIZZI A, VICHI M, BOCK H H. Advances in data science and classification. Berlin:Springer, 1998:265-277.
- [15] ARABIE P, CARROLL J D. Mapelus: A mathematical programming approach to fitting the adclus model[J]. Psychometrika, 1980, 45(2):211-235.
- [16] DIDAY E. An intuitive form of expression:Pyramids[D]. Paris:INRIA, 1984.
- [17] JAMES G M, SUGAR C A. Clustering for sparsely sampled functional data[J]. Publications of the American Statistical Association, 2003, 98(462):397-408.
- [18] WINSBERG S, SOETE G D. Latent class models for time series analysis[J]. Applied Stochastic Models in Business & Industry, 1999, 15(15):183-194.
- [19] VRAC M, BILLARD L, DIDAY E. Copula analysis of mixture models[J]. Computational Statistics, 2012, 27(3):427-457.
- [20] KOSMIDIS I, KARLIS D. Model-based clustering using copulas with applications[J]. Statistics and Computing, 2016, 26(5):1079-1099.
- [21] MARBAC M, BIERNACKI C, VANDEWALLE V. Model-based clustering of Gaussian copulas for mixed data[J]. Communications in Statistics-Theory and Methods, 2017, 46(23):11635-11656.

#### 作者简介:

**夏棒** 男, 博士。主要研究方向:成分数据路径分析;人机交互、结构方程模型等。

**EMILION Richard** 男, 博士, 教授, 博士生导师。主要研究方向:Dirichlet 分布数据及其数学理论。

**王惠文** 女, 博士, 教授, 博士生导师。主要研究方向:复杂数据分析、投入产出表预测等。

# Comparison between EM algorithm and dynamical clustering algorithm for Dirichlet mixture samples

XIA Bang<sup>1,\*</sup>, EMILION Richard<sup>2</sup>, WANG Huiwen<sup>3</sup>

(1. Postdoctoral Workstation, Industrial and Commercial Bank of China, Beijing 100032, China;

2. MAPMO, University of Orleans, Orleans 45000, France;

3. School of Economics and Management, Beihang University, Beijing 100083, China)

**Abstract:** Dirichlet distribution is a kind of continuous multivariate probability distribution with positive parameter vectors, which is widely used in proportional structure problems. Expectation maximization (EM) algorithm and dynamical clustering algorithm of Dirichlet mixture samples are presented, their mathematical process is deduced, and the iteration steps of the algorithms are given. Then, using digital simulation experiments, the clustering effects of the two machine learning algorithms with Dirichlet samples are compared. By calculating six evaluation factors which are log-likelihood function value, program running time, convergence iteration times, clustering accuracy, true positive rate (TPR) and false positive rate (FPR), the simulation results show that EM algorithm has higher clustering accuracy but lower operational efficiency, while dynamical clustering algorithm has higher operational efficiency but loses some clustering accuracy. Therefore, in practical application, it is suggested to weigh the relative requirements of accuracy and operational efficiency before selecting a suitable algorithm to cluster Dirichlet samples.

**Keywords:** Dirichlet distribution; mixture sample; expectation maximization (EM) algorithm; dynamical clustering; machine learning

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

DOI: 10.13700/j.bh.1001-5965.2018.0758

# 基于高阶 LADRC 的 V/STOL 飞机悬停/平移模式鲁棒协调解耦控制

高阳<sup>1</sup>, 吴文海<sup>1,\*</sup>, 嵇绍康<sup>1</sup>, 郑毅<sup>2</sup>

(1. 海军航空大学 青岛校区, 青岛 266041; 2. 中国人民解放军 91599 部队 修理厂, 烟台 265200)



**摘 要:** 针对垂直/短距起降(V/STOL)飞机在悬停/平移模式下存在的动力学耦合、推力矢量控制冗余以及易受扰动风影响的问题,提出了一种基于高阶线性自抗扰控制(LADRC)的鲁棒协调解耦控制方法。首先根据V/STOL飞机的概念方案,建立了推力矢量模型和扰动风影响下的非线性悬停/平移运动模型。然后在此基础上,给出了该模式下位置和姿态的协调控制策略,据此通过控制量变换设计了六通道的自抗扰解耦控制律,其中利用LADRC对总扰动的实时估计补偿能力避免了多推力矢量的冗余控制。仿真比较结果验证了LADRC对悬停/平移模式控制的有效性以及对飞机内部参数摄动和外界突风干扰的鲁棒性。

**关键词:** 垂直/短距起降(V/STOL)飞机; 线性自抗扰控制(LADRC); 悬停/平移模式; 协调解耦控制; 突风扰动

**中图分类号:** V249.1

**文献标识码:** A

**文章编号:** 1001-5965(2019)09-1812-12

垂直/短距起降(Vertical and/or Short Take-Off and Landing, V/STOL)飞机是对垂直起降(Vertical Take-Off and Landing, VTOL)和短距起飞/垂直降落(Short Take-Off and Vertical Landing, STOVL)固定翼飞机的统称,它兼具旋翼和固定翼飞机的优势,既可以减少甚至摆脱对飞行跑道的依赖,又具备较大的飞行速度、航程、载荷和优异的机动性能。自20世纪40年代至今,诞生的各类V/STOL飞机型号达30余种<sup>[1-2]</sup>。特别是近年来,随着美国第5代战斗机F-35B的研发和服役,关于V/STOL飞机的研究越来越受到重视。但是相比常规起降飞机,V/STOL飞机具有更为复杂的动力学特性<sup>[3]</sup>和操控性<sup>[4]</sup>,这主要是由于其采用了诸多的推力矢量装置,如单发四转向喷管、引射增升器、升力风扇和矢量尾喷管等,从而使其诸

如垂直/短距起飞、过渡、悬停/平移、垂直着陆等非常规飞行模式的控制律设计难度大为增加。其中,悬停模式是实现其他模式的基础,平移模式是小范围地改变悬停的位置以实现定点着陆的前提,两者的动力学关系是一致的,因而实现悬停/平移模式的有效控制尤为关键。

对此,国内外学者从不同的角度进行了研究。较常见的一种是将VTOL飞机简化为平面VTOL(Planar VTOL, PVTOL)飞机模型,使其成为2输入3输出的欠驱动系统,然后分别采用近似反馈线性化<sup>[5]</sup>、最优控制<sup>[6]</sup>、基于观测器的输出反馈控制<sup>[7]</sup>和滑模控制<sup>[8]</sup>等方法设计控制器。然而,由于简化后的系统并不能有效地反映VTOL飞机的实际运动状态,选择推力和滚转力矩作为设计控制量,也难以将其转换为实际的多操纵面控制

收稿日期: 2018-12-25; 录用日期: 2019-02-02; 网络出版时间: 2019-02-21 15:10

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190221.0903.002.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190221.0903.002.html)

基金项目: 国家自然科学基金(60674090, 51505491)

\* 通信作者: E-mail: [hkdx\\_2017@126.com](mailto:hkdx_2017@126.com)

**引用格式:** 高阳, 吴文海, 嵇绍康, 等. 基于高阶LADRC的V/STOL飞机悬停/平移模式鲁棒协调解耦控制[J]. 北京航空航天大学学报, 2019, 45(9): 1812-1823. GAO Y, WU W H, JI S K, et al. High-order LADRC based robust coordinated decoupling control for V/STOL aircraft in hover/translation mode[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1812-1823 (in Chinese).

信号,因此这些研究更侧重于理论方法的探索而非工程问题的解决。另一种是直接针对 V/STOL 模型机/验证机建立相应的动力学模型,然后结合多操纵面分配方法设计相应的解耦控制器<sup>[9-10]</sup>。此外,文献[11]提出了一种基于新姿态解算方法的滞后切换策略,解决了尾座式推力矢量飞机的悬停控制问题。总的来说,现有研究大多是基于精确的飞机模型设计控制律,没有充分考虑系统的鲁棒性问题,且通过控制分配算法得到实际控制量,往往会破坏系统的稳定性。

由于 V/STOL 飞机在悬停/平移模式下飞行速度很小,使其极易受到局部扰动风的影响。特别是短时的突风,它会使飞机的空速、迎角和侧滑角等发生改变,进而影响飞行轨迹和姿态。因此,必须设计具有强抗风能力的鲁棒飞行控制律。对此,自抗扰控制(Active Disturbance Rejection Control, ADRC)应是一种行之有效的应对方案。大量理论<sup>[12-14]</sup>和应用<sup>[15-16]</sup>方面的研究表明,ADRC 不依赖被控对象的精确数学模型,对具有未建模动态、参数摄动和外界干扰的系统均能有效控制,具有“天生”的鲁棒性和抗干扰性。但受限于原始 ADRC 中非线性、非光滑的反馈结构,其理论分析十分困难,在应用中需要调节的控制器参数也较多。为简化 ADRC 的分析和实现,文献[17]利用带宽概念,提出了单参数调节的线性 ADRC(Linear

ADRC, LADRC)方法,并在文献[18]中直接假设系统不确定性的导数有界,证明了 LADRC 的收敛性,但由于该假设本质上是一个先验条件,因而很难得到验证。在此基础上,笔者基于单参数高增益观测器的思想,提出了一种高阶 LADRC 方法,并进一步放宽假设条件给出了收敛性证明,但缺少对该方法的实际应用<sup>[19]</sup>。

本文在上述研究的基础上,以先进的“矢量发动机+升力风扇”型 V/STOL 飞机为研究对象,首先构建推力矢量模型和基于扰动风影响的悬停/平移运动模型,然后在高阶 LADRC 的控制框架内对模型进行变换并选取适当控制量,来分别设计姿态和位置的自抗扰控制律。本文的意义在于展现如何应用 LADRC 实现 V/STOL 飞机在悬停/平移模式下的动力学解耦和多操纵量协调控制,同时验证 LADRC 对该模式控制的有效性和鲁棒性。

## 1 悬停/平移模式非线性模型

### 1.1 V/STOL 飞机概念模型

本文研究的 V/STOL 飞机为美国国家航空航天局(NASA)的某型 STOVL 验证机<sup>[20]</sup>,采用单发动机、三角翼、鸭翼、双垂尾的布局形式,如图 1(a)所示。推进系统采用类似 F-35B 的设计方案,由一个矢量发动机和一个轴驱动升力风扇组成,具体形式和参数如图 1(b)和表 1 所示

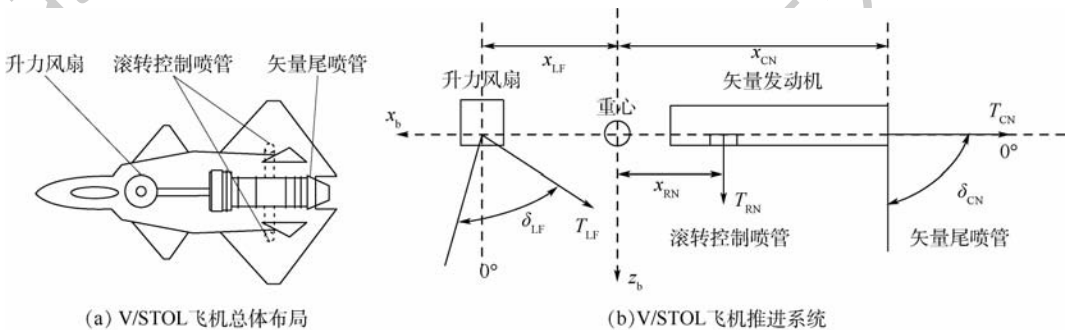


图 1 V/STOL 飞机概念方案

Fig. 1 Conceptual scheme of V/STOL aircraft

表 1 V/STOL 飞机推进系统的参数配置

Table 1 Configured parameters of V/STOL aircraft propulsion system

部件	推力/N (规定为正)	纵向偏转角/(°) (下偏或后偏为正)	侧向偏转角/(°) (左偏为正)	力臂/m (遵循机体坐标系)
矢量尾喷管	$T_{CNmax} = 80\ 000$	$\delta_{CN} \in [0, 90]$	$\delta_{CNy} \in [-12, 12]$	$x_{CN} = -6.01$
升力风扇	$T_{LFmax} = 89\ 000$	$\delta_{LF} \in [-20, 60]$	$\delta_{LFy} \in [-12, 12]$	$x_{LF} = 3.57$
滚转控制喷管	$T_{RNmax} = 17\ 330$			$x_{RN} = -1.91$ $y_{RN} = -y_{LRN} = 2$

( $y_{LRN}$ 、 $y_{RRN}$ 分别为左、右滚转控制喷管推力的侧向力臂)。矢量尾喷管在飞机对称平面内可向下偏转 $90^\circ$ ,同时可侧向偏转 $\pm 12^\circ$ ,能够产生 $80\,000\text{ N}$ 的悬停升力;升力风扇竖直安装于驾驶舱后方,可向前偏转 $20^\circ$ 、向后偏转 $60^\circ$ ,也可侧向偏转 $\pm 12^\circ$ ,能够产生高达 $89\,000\text{ N}$ 的悬停升力;由发动机外涵道引出的2个滚转控制喷管位于两翼根处,由压气机供气,共可产生占全机总推力 $9.3\%$ 的推力,用于滚转控制。

### 1.2 推力矢量模型

V/STOL飞机的推进系统可为其悬停/平移模式提供直接控制力和力矩。将各矢量推力在机体坐标系下分解后叠加,可得到由推进系统产生的纵向力 $T_{x_b}$ 、侧向力 $T_{y_b}$ 和法向力 $T_{z_b}$ 表达式为

$$\begin{cases} T_{x_b} = T_{CN} \cos \delta_{CN_y} \cos \delta_{CN} + T_{LF} \cos \delta_{LF_y} \sin \delta_{LF} \\ T_{y_b} = T_{CN} \sin \delta_{CN_y} + T_{LF} \sin \delta_{LF_y} \\ T_{z_b} = -T_{CN} \cos \delta_{CN_y} \sin \delta_{CN} - \\ \quad T_{LF} \cos \delta_{LF_y} \cos \delta_{LF} - T_{RN} \end{cases} \quad (1)$$

式中:滚转控制喷管的总推力占全机总推力的 $9.3\%$ ,有

$$\frac{T_{CN} + T_{LF}}{90.7\%} = \frac{T_{RN}}{9.3\%} \quad (2)$$

式中: $T_{RN} = -T_{RRN} - T_{LRN}$ , $T_{LRN}$ 和 $T_{RRN}$ 分别为左、右滚转控制喷管的推力(遵循机体坐标系)。

假设矢量尾喷口、升力风扇喷口和滚转控制喷口与飞机重心处于同一平面上,则由各轴向力产生的滚转力矩 $l_{prop}$ 、俯仰力矩 $m_{prop}$ 和偏航力矩 $n_{prop}$ 的表达式为

$$\begin{cases} l_{prop} = T_{RRN}y_{RRN} + T_{LRN}y_{LRN} \\ m_{prop} = T_{RN}x_{RN} + T_{CN}x_{CN} \cos \delta_{CN_y} \sin \delta_{CN} + \\ \quad T_{LF}x_{LF} \cos \delta_{LF_y} \cos \delta_{LF} \\ n_{prop} = T_{CN}x_{CN} \sin \delta_{CN_y} + T_{LF}x_{LF} \sin \delta_{LF_y} \end{cases} \quad (3)$$

此外,V/STOL飞机特有的大推力大范围偏转还会产生喷气诱导效应(Jet-Induced Effect, JIE),导致升力损失和抬头力矩。根据文献[21]的描述,可得JIE模型:

$$\begin{cases} L_{JIE} = (T_{CN} + T_{LF})C_{L_{JIE}} \\ m_{JIE} = (T_{CN} + T_{LF})d_e C_{m_{JIE}} \end{cases} \quad (4)$$

式中: $L_{JIE}$ 和 $m_{JIE}$ 分别为喷气诱导升力和喷气诱导俯仰力矩; $C_{L_{JIE}}$ 和 $C_{m_{JIE}}$ 分别为对应的增量系数; $d_e$ 为总的等效环流喷气直径。

当飞机处于悬停状态时,所有矢量喷管都竖

直向下,由力和力矩平衡可得

$$\begin{cases} T_{CN_0} + T_{LF_0} + T_{RN_0} = mg + |L_{JIE_0}| \\ T_{CN_0} |x_{CN}| + T_{RN_0} |x_{RN}| = T_{LF_0} x_{LF} + m_{JIE_0} \end{cases} \quad (5)$$

式中: $m$ 为飞机质量; $g$ 为重力加速度;次下标0代表悬停平衡状态。

这样,由式(2)、式(5)即可得到飞机悬停状态下的平衡推力 $T_{CN_0}$ 、 $T_{LF_0}$ 和 $T_{RN_0}$ 。

### 1.3 考虑扰动风影响的飞机运动模型

下面首先给出无风状态下的飞机运动模型。当V/STOL飞机处于悬停/平移模式时,飞行速度很小,气动力和力矩几乎为零,同时姿态角变化幅度也不大,可以将与姿态角相关的小值作为扰动项,从而得到仅由推进系统作用的简化六自由度非线性运动模型:

$$\begin{cases} \dot{x}_g = u_b + \Delta_{x_g} \\ \dot{y}_g = v_b + \Delta_{y_g} \\ \dot{z}_g = w_b + \Delta_{z_g} \\ \dot{u}_b = rv_b - qw_b + T_{x_b}/m + \Delta_{u_b} \\ \dot{v}_b = pw_b - ru_b + T_{y_b}/m + \Delta_{v_b} \\ \dot{w}_b = qu_b - pv_b + (T_{z_b} - L_{JIE} + mg)/m + \Delta_{w_b} \\ \dot{\phi} = p + \Delta_{\dot{\phi}} \\ \dot{\theta} = q + \Delta_{\dot{\theta}} \\ \dot{\psi} = r + \Delta_{\dot{\psi}} \\ \dot{p} = c_1 pq + c_2 qr + c_3 l_{prop} + c_4 n_{prop} \\ \dot{q} = c_5 pr + c_6 (r^2 - p^2) + c_7 (m_{prop} + m_{JIE}) \\ \dot{r} = c_8 pq - c_1 qr + c_4 l_{prop} + c_9 n_{prop} \end{cases} \quad (6)$$

式中:( $x_g, y_g, z_g$ )为飞机在地面坐标系下的位置,且有 $h = -z_g$ 为飞行高度;( $u_b, v_b, w_b$ )为飞机在机体坐标系下的纵向、侧向和法向速度;( $\phi, \theta, \psi$ )为飞机的姿态角;( $p, q, r$ )为绕机体轴转动的对应角速度; $c_1 \sim c_9$ 为与转动惯量相关的常系数; $\Delta$ 为与姿态角相关的小扰动项。

对式(6)中的位置、姿态方程求导并代入式(1)、式(3),可得纵向运动方程:

$$\begin{cases} \ddot{\theta} = \dot{f}_\theta + c_7 T_{CN} x_{CN} \cos \delta_{CN_y} \sin \delta_{CN} + \\ \quad c_7 T_{LF} x_{LF} \cos \delta_{LF_y} \cos \delta_{LF} + c_7 T_{RN} x_{RN} \\ \ddot{x}_g = \dot{f}_{x_g} + T_{CN} \cos \delta_{CN_y} \cos \delta_{CN}/m + \\ \quad T_{LF} \cos \delta_{LF_y} \sin \delta_{LF}/m \\ \ddot{z}_g = -\dot{h} = \dot{f}_{z_g} - T_{CN} \cos \delta_{CN_y} \sin \delta_{CN}/m - \\ \quad T_{LF} \cos \delta_{LF_y} \cos \delta_{LF}/m - T_{RN}/m \end{cases} \quad (7)$$

式中:

$$\begin{cases} \dot{f}_\theta = \dot{\Delta}_\theta + c_5 pr + c_6 (r^2 - p^2) + c_7 m_{JIE} \\ \dot{f}_{\dot{x}_g} = \dot{\Delta}_{\dot{x}_g} + \dot{\Delta}_{\dot{u}_b} + rv_b - qw_b \\ \dot{f}_{\dot{z}_g} = \dot{\Delta}_{\dot{z}_g} + \dot{\Delta}_{\dot{w}_b} + qu_b - pv_b + g - L_{JIE}/m \end{cases}$$

横侧向运动方程:

$$\begin{cases} \ddot{\phi} = f_\phi + c_3 T_{LRN} y_{LRN} + c_3 T_{RRN} y_{RRN} \\ \ddot{\psi} = f_\psi + c_9 T_{CN} x_{CN} \sin \delta_{CN_y} + c_9 T_{LF} x_{LF} \sin \delta_{LF_y} \\ \ddot{y}_g = f_{y_g} + T_{CN} \sin \delta_{CN_y} / m + T_{LF} \sin \delta_{LF_y} / m \end{cases} \quad (8)$$

式中:

$$\begin{cases} \dot{f}_\phi = \dot{\Delta}_\phi + c_1 pq + c_2 qr + c_4 n_{prop} \\ \dot{f}_\psi = \dot{\Delta}_\psi + c_8 pq - c_1 qr + c_4 l_{prop} \\ \dot{f}_{y_g} = \dot{\Delta}_{y_g} + f_{v_b} + pw_b - ru_b \end{cases}$$

当飞机受到扰动风影响时,就需要对上述运动模型进行修正。在有风影响的情况下,飞机的飞行速度矢量  $\mathbf{V}$ 、航迹速度矢量  $\mathbf{V}_k$  和风速矢量  $\mathbf{V}_w$  之间,具有如下的速度三角形关系:

$$\mathbf{V} = \mathbf{V}_k - \mathbf{V}_w \quad (9)$$

将式(9)分别在机体坐标系和地面坐标系下投影,可以得到如下的修正关系:

$$\begin{bmatrix} u'_b \\ v'_b \\ w'_b \end{bmatrix} = \begin{bmatrix} u_b \\ v_b \\ w_b \end{bmatrix} - \begin{bmatrix} u_{wb} \\ v_{wb} \\ w_{wb} \end{bmatrix} \quad (10)$$

$$\begin{bmatrix} \dot{x}'_g \\ \dot{y}'_g \\ \dot{z}'_g \end{bmatrix} = \begin{bmatrix} u_g \\ v_g \\ w_g \end{bmatrix} = \mathbf{B}_b^g \begin{bmatrix} u'_b \\ v'_b \\ w'_b \end{bmatrix} + \begin{bmatrix} u_{wg} \\ v_{wg} \\ w_{wg} \end{bmatrix} \quad (11)$$

式中:  $(u'_b, v'_b, w'_b)$  为飞机受扰动风影响并投影在机体坐标系下的修正速度;  $(u_{wg}, v_{wg}, w_{wg})$  为基于地面坐标系给出的轴向风速;  $(u_{wb}, v_{wb}, w_{wb})$  为转换到机体坐标系下的轴向风速;  $\mathbf{B}_b^g = (\mathbf{B}_b^g)^T$  为地面坐标系到机体坐标系的转换矩阵,因而有

$$\begin{bmatrix} u_{wb} & v_{wb} & w_{wb} \end{bmatrix}^T = \mathbf{B}_g^b \begin{bmatrix} u_{wg} & v_{wg} & w_{wg} \end{bmatrix}^T$$

由于机体尺度相比风场尺度很小,因而可忽略由扰动风速矢量沿翼展和机身方向的微小差异所造成的气动力矩变化。这样,将式(10)、式(11)代入式(6),可得到扰动风影响下的飞机质心运动模型:

$$\begin{cases} \dot{x}'_g = u'_b + u_{wg} + \Delta'_{x_g} \\ \dot{y}'_g = v'_b + v_{wg} + \Delta'_{y_g} \\ \dot{z}'_g = w'_b + w_{wg} + \Delta'_{z_g} \\ \dot{u}'_b = r(v'_b + v_{wb}) - q(w'_b + w_{wb}) - \dot{u}_{wb} + T_{x_b}/m + \Delta'_{u_b} \\ \dot{v}'_b = p(w'_b + w_{wb}) - r(u'_b + u_{wb}) - \dot{v}_{wb} + T_{y_b}/m + \Delta'_{v_b} \\ \dot{w}'_b = q(u'_b + u_{wb}) - p(v'_b + v_{wb}) - \dot{w}_{wb} + (T_{z_b} - L_{JIE} + mg)/m + \Delta'_{w_b} \end{cases} \quad (12)$$

式中:小扰动项  $\Delta'_{x_g}$ 、 $\Delta'_{y_g}$  和  $\Delta'_{z_g}$  作相应改变即可。

在式(12)的基础上仍然可以得到如式(7)、式(8)的飞机纵向和横侧向运动方程,只是其中的扰动项  $f_{\dot{x}_g}$ 、 $f_{\dot{y}_g}$  和  $f_{\dot{z}_g}$  变为

$$\begin{cases} f'_{\dot{x}_g} = \dot{\Delta}'_{x_g} + \dot{\Delta}'_{u_b} + \dot{u}_{wg} - \dot{u}_{wb} + r(v'_b + v_{wb}) - q(w'_b + w_{wb}) \\ f'_{\dot{y}_g} = \dot{\Delta}'_{y_g} + f_{v_b} + \dot{v}_{wg} - \dot{v}_{wb} + p(w'_b + w_{wb}) - r(u'_b + u_{wb}) \\ f'_{\dot{z}_g} = \dot{\Delta}'_{z_g} + \dot{\Delta}'_{w_b} + \dot{w}_{wg} - \dot{w}_{wb} + q(u'_b + u_{wb}) - p(v'_b + v_{wb}) + g - L_{JIE}/m \end{cases}$$

## 2 高阶LADRC方法

考虑如下单输入单输出不确定非线性系统:

$$\begin{cases} \dot{x}^{(n)}(t) = f(\cdot) + w(t) + bu(t) \\ y(t) = x(t) \end{cases} \quad (13)$$

式中:  $u$  为控制输入;  $y$  为系统输出;  $f(\cdot)$  为不确定的内部动态;  $w$  为未知的外部干扰;  $b$  为不确定的控制量系数,满足  $b_1 < b < b_2$ ,  $b_1$  和  $b_2$  为常数。

取状态变量  $x_1 = x, x_2 = \dot{x}, \dots, x_n = x^{(n-1)}$ , 将系统的所有不确定性当作总扰动并将其扩充为新的状态变量  $x_{n+1} = f(\cdot) + w + (b - b_0)u$ , 其中  $b_0 \in (b_1, b_2)$ , 可得系统(13)的扩张状态方程:

$$\begin{cases} \dot{x}_1(t) = x_2(t) \\ \vdots \\ \dot{x}_{n-1}(t) = x_n(t) \\ \dot{x}_n(t) = x_{n+1}(t) + b_0 u(t) \\ \dot{x}_{n+1}(t) = F(t) \\ y(t) = x_1(t) \end{cases} \quad (14)$$

为方便表示,在不引起歧义的前提下,后文将

省略时间变量  $t$ 。针对系统(14),可设计如下具有“扰动估计补偿”功能的 LADRC 控制器,它由线性跟踪微分器 (Linear Tracking Differentiator, LTD)、线性扩张状态观测器 (Linear Extended State Observer, LESO) 和线性状态误差反馈 (Linear State Error Feedback, LSEF) 3 部分组成。

### 2.1 线性跟踪微分器

LTD 在 LADRC 中相对独立,其作用在于跟踪给定的输入信号  $v_0$  并得到输入的各阶微分信号  $v_0^{(i)}$ ,即  $v_1 \rightarrow v_0, v_{i+1} \rightarrow v_0^{(i)} (i = 1, 2, \dots, n)$ 。高阶 LTD 的具体形式为

$$\begin{cases} \dot{v}_1 = v_2 \\ \vdots \\ \dot{v}_n = v_{n+1} \\ \dot{v}_{n+1} = R^{n+1} \left[ a_1 (v_1 - v_0) + a_2 \frac{v_2}{R} + \dots + a_{n+1} \frac{v_{n+1}}{R^n} \right] \end{cases} \quad (15)$$

式中:  $R$  为决定跟踪快慢的调节增益;系数  $a_j (j = 1, 2, \dots, n+1)$  需满足矩阵  $A$  是 Hurwitz 的。

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ a_1 & a_2 & a_3 & \dots & a_{n+1} \end{bmatrix}$$

关于高阶 LTD 式(15)的收敛性有如下结论<sup>[22]</sup>。

**引理 1** 若光滑函数  $v_0(t) : [0, \infty) \rightarrow \mathbf{R}$  满足  $B = \sup_{t \in [0, T_1]} |v_0^{(i)}| < \infty, i = 1, 2, \dots, n$ , 其中常数  $T_1 > 0, B > 0$ , 则对任意给定的 LTD 式(15)的初始值以及任意的  $\rho > 0, \tau_1 > 0$ , 存在常数  $R_0 > 0$ , 使得对任意的  $R > R_0, |v_j - v_0^{(j-1)}| < \rho (j = 1, 2, \dots, n+1)$  在  $t \in [\tau_1, T_1]$  上都成立。

### 2.2 线性扩张状态观测器

LESO 是 LADRC 的核心技术。可利用实时的输入  $u$ 、输出  $y$  信息构造如下的高阶 LESO:

$$\begin{cases} \dot{z}_1 = z_2 + \frac{k_1}{r}(y - z_1) \\ \vdots \\ \dot{z}_{n-1} = z_n + \frac{k_{n-1}}{r^{n-1}}(y - z_1) \\ \dot{z}_n = z_{n+1} + \frac{k_n}{r^n}(y - z_1) + b_0 u \\ \dot{z}_{n+1} = \frac{k_{n+1}}{r^{n+1}}(y - z_1) \end{cases} \quad (16)$$

式中:  $z_1, z_2, \dots, z_{n+1}$  分别为系统(14)状态  $x_1, x_2, \dots, x_{n+1}$  的估计;  $r$  为 LESO 调节增益的倒数;系

数  $k_j (j = 1, 2, \dots, n+1)$  需满足矩阵  $K$  是 Hurwitz 的。

$$K = \begin{bmatrix} -k_1 & 1 & 0 & \dots & 0 \\ -k_2 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -k_n & 0 & 0 & \dots & 1 \\ -k_{n+1} & 0 & 0 & \dots & 0 \end{bmatrix}$$

### 2.3 线性状态误差反馈

定义系统的状态误差  $e_1 = z_1 - v_1, \dots, e_n = z_n - v_n$ , 并考虑对总扰动  $x_{n+1}$  的实时补偿, 可设计如下的 LSEF:

$$u = (l_1 e_1 + \dots + l_n e_n + v_{n+1} - z_{n+1}) / b_0 \quad (17)$$

式中: 系数  $l_i (i = 1, 2, \dots, n)$  需满足矩阵  $L$  是 Hurwitz 的。

$$L = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ l_1 & l_2 & l_3 & \dots & l_n \end{bmatrix}$$

关于 LADRC 式(15)~式(17)的收敛性, 笔者在文献[19]中给出如下结论。

**引理 2** 对于不确定非线性系统(13), 若其内部动态  $f(\cdot)$  和外部干扰  $w$  有界且关于自变量的所有偏导数连续且有界, 设计如式(15)~式(17)的 LADRC 控制器, 则对任意给定的闭环系统初始值, 存在常数  $r_0 > 0, T_2 > 0$ , 使得对任意的  $r \in (0, r_0)$ , 满足  $\lim_{t \rightarrow \infty} |x_i - v_i| \leq T_2 r (i = 1, 2, \dots, n)$ 。

**注 1** 对于实际飞机系统, 由于飞行状态和控制输入总是有界的, 外界干扰亦是有限的, 因此总能满足引理 2 中关于不确定性及其偏导数有界的假设。

根据引理 1 和引理 2, 当  $R \rightarrow \infty, r \rightarrow 0$  时, 可得系统的跟踪误差满足:

$$\lim_{t \rightarrow \infty} |y - v_0| = \lim_{t \rightarrow \infty} |(x_1 - v_1) + (v_1 - v_0)| = 0 \quad (18)$$

## 3 悬停/平移模式控制器设计

### 3.1 控制策略

在稳定悬停状态下, 飞机的姿态和位置保持不变, 总推力与重力相平衡, 各推力矢量(大小和方向)亦保持不变。当飞机受扰使平衡状态被破坏, 或飞机需要改变悬停位置时, 根据各推力矢量的功能特点, 飞机姿态和位置的控制可由如下几种策略来实现, 具体如表 2 所示。

表 2 V/STOL 飞机悬停/平移模式的控制策略

Table 2 Control strategy for V/STOL aircraft in hover/translation mode

控制通道	控制策略	
姿态控制	滚转	① 左、右滚转控制喷管之间不同推力的转换
	俯仰	② 升力风扇与矢量尾喷管之间不同推力的转换
		③ 升力风扇与矢量尾喷管在对称平面内的偏转
位置控制	偏航	④ 升力风扇与矢量尾喷管之间不同推力的转换与各自的偏转同时作用
		⑤ 升力风扇与矢量尾喷管的侧向偏转
	纵向	③ 升力风扇与矢量尾喷管在对称平面内的偏转
	侧向	⑤ 升力风扇与矢量尾喷管的侧向偏转
	法向	⑥ 保持飞机俯仰力矩平衡的条件下,改变飞机的总推力

注:对于纵向通道,若仅改变推力的大小,无法实施控制;对于法向通道,若仅改变推力的方向,飞机会掉高度,无法进行有效控制。

对应 1.3 节的运动模型,将悬停/平移模式的控制分成纵向控制和横侧向控制。根据表 2 给出的控制策略,在纵向控制中,为了避免控制冗余,选择策略③来实现俯仰和纵向通道的耦合控制,利用策略⑥实现法向通道的控制;在横侧向控制中,利用策略①实现滚转通道的控制,利用策略⑤来实现偏航和侧向通道的耦合控制。

实际上,任意控制量的改变都会引起相应状态量的变化,进而通过动力学耦合影响其他状态量,但是这种影响反映在式(7)、式(8)中,是被当作系统的未知内部动态来处理的。接下来,将通过适当变换,利用 LESO 式(16)和 LSEF 式(17)对这种耦合不确定性连同其他扰动项进行实时的估计和补偿,进而实现各通道的动力学解耦以及多推力矢量的协调控制。

### 3.2 纵向运动控制器

#### 3.2.1 俯仰和纵向通道控制器

对于俯仰和纵向通道运动方程:

$$\ddot{\theta} = \ddot{f}_{\theta} + c_7 T_{CN} x_{CN} \cos \delta_{CN_y} \sin \delta_{CN} + c_7 T_{LF} x_{LF} \cos \delta_{LF_y} \cos \delta_{LF} + c_7 T_{RN} x_{RN} \quad (19)$$

$$\ddot{x}_g = \ddot{f}'_{x_g} + T_{CN} \cos \delta_{CN_y} \cos \delta_{CN} / m + T_{LF} \cos \delta_{LF_y} \sin \delta_{LF} / m \quad (20)$$

其包含关于控制量  $\delta_{CN}$  和  $\delta_{LF}$  的三角函数形式。

通常,当角度  $\delta$  在小范围内 ( $\pm 15^\circ$ ) 变化时,有如下近似表达式成立:  $\cos \delta \approx 1, \sin \delta \approx \delta$  ( $\delta$  以弧度计量)。但  $\delta_{CN}$  和  $\delta_{LF}$  的变化范围较大,使系统(19)、(20)难以转化为系统(14)的形式。对此,直接取  $u_1 = \sin \delta_{CN}, u_2 = \sin \delta_{LF}$  分别作为系统(19)、(20)的虚拟控制量。

由于  $\delta_{CN_y}$  和  $\delta_{LF_y}$  在其变化范围内有  $\cos \delta_{CN_y} \approx 1, \cos \delta_{LF_y} \approx 1$ , 因而可估计到虚拟控制量系数  $b_1 = c_7 x_{CN} T_{CN_0}, b_2 = T_{LF_0} / m$ , 而将其余部分作为系统的

总扰动,即

$$\ddot{\theta} = F_{\theta} + b_1 u_1 \quad (21)$$

$$\ddot{x}_g = F_{x_g} + b_2 u_2 \quad (22)$$

其中总扰动为

$$\begin{cases} F_{\theta} = \ddot{f}_{\theta} + c_7 x_{CN} (T_{CN} \cos \delta_{CN_y} - T_{CN_0}) \sin \delta_{CN} + \\ \quad c_7 T_{LF} x_{LF} \cos \delta_{LF_y} \cos \delta_{LF} + c_7 T_{RN} x_{RN} \\ F_{x_g} = \ddot{f}'_{x_g} + (T_{LF} \cos \delta_{LF_y} - T_{LF_0}) \sin \delta_{LF} / m + \\ \quad T_{CN} \cos \delta_{CN_y} \cos \delta_{CN} / m \end{cases}$$

这样,通过将系统的内部耦合动态、推力矢量冗余控制和外部干扰一起作为系统的总扰动,利用 LADRC 式(15)~式(17)对其进行实时估计和补偿,可分别对系统(21)、(22)设计如下的鲁棒协调解耦控制器:

$$\begin{cases} \dot{v}_{\theta,1} = v_{\theta,2} \\ \dot{v}_{\theta,2} = v_{\theta,3} \\ \dot{v}_{\theta,3} = R_{\theta}^3 \left[ a_{\theta,1} (v_{\theta,1} - \theta_c) + \frac{a_{\theta,2}}{R_{\theta}} v_{\theta,2} + \frac{a_{\theta,3}}{R_{\theta}^2} v_{\theta,3} \right] \\ \dot{z}_{\theta,1} = z_{\theta,2} + \frac{k_{\theta,1}}{r_{\theta}} (\theta - z_{\theta,1}) \\ \dot{z}_{\theta,2} = z_{\theta,3} + \frac{k_{\theta,2}}{r_{\theta}} (\theta - z_{\theta,1}) + b_1 u_1 \\ \dot{z}_{\theta,3} = \frac{k_{\theta,3}}{r_{\theta}} (\theta - z_{\theta,1}) \\ e_{\theta,1} = z_{\theta,1} - v_{\theta,1} \\ e_{\theta,2} = z_{\theta,2} - v_{\theta,2} \\ u_1 = (l_{\theta,1} e_{\theta,1} + l_{\theta,2} e_{\theta,2} + v_{\theta,3} - z_{\theta,3}) / b_1 \end{cases} \quad (23)$$

$$\begin{cases} \dot{v}_{x_g,1} = v_{x_g,2} \\ \dot{v}_{x_g,2} = v_{x_g,3} \\ \dot{v}_{x_g,3} = R_{x_g}^3 \left[ a_{x_g,1} (v_{x_g,1} - x_{gc}) + \frac{a_{x_g,2}}{R_{x_g}} v_{x_g,2} + \frac{a_{x_g,3}}{R_{x_g}^2} v_{x_g,3} \right] \\ \dot{z}_{x_g,1} = z_{x_g,2} + \frac{k_{x_g,1}}{r_{x_g}} (x_g - z_{x_g,1}) \\ \dot{z}_{x_g,2} = z_{x_g,3} + \frac{k_{x_g,2}}{r_{x_g}^2} (x_g - z_{x_g,1}) + b_2 u_2 \\ \dot{z}_{x_g,3} = \frac{k_{x_g,3}}{r_{x_g}^3} (x_g - z_{x_g,1}) \\ e_{x_g,1} = z_{x_g,1} - v_{x_g,1} \\ e_{x_g,2} = z_{x_g,2} - v_{x_g,2} \\ u_2 = (l_{x_g,1} e_{x_g,1} + l_{x_g,2} e_{x_g,2} + v_{x_g,3} - z_{x_g,3}) / b_2 \end{cases} \quad (24)$$

式中:  $\theta_c$ 、 $x_{gc}$  分别为俯仰和纵向通道的参考输入指令。进而可得  $\delta_{CN} = \arcsin u_1$ ,  $\delta_{LF} = \arcsin u_2$ 。

### 3.2.2 法向通道控制器

法向通道控制器设计的关键在于如何选择控制量,实现推力  $T_{CN}$ 、 $T_{LF}$  和  $T_{RN}$  控制的同时保持三者的协调。

对于法向通道运动方程:

$$\ddot{z}_g = -\ddot{h} = f'_{z_g} - T_{CN} \cos \delta_{CN_y} \sin \delta_{CN} / m - T_{LF} \cos \delta_{LF_y} \cos \delta_{LF} / m - T_{RN} / m \quad (25)$$

可将其转化为

$$\ddot{h} = F_{\ddot{h}} + (T_{CN} + T_{LF} + T_{RN}) / m \quad (26)$$

其中总扰动为

$$F_{\ddot{h}} = -f'_{z_g} + T_{CN} (\cos \delta_{CN_y} \sin \delta_{CN} - 1) / m + T_{LF} (\cos \delta_{LF_y} \cos \delta_{LF} - 1) / m$$

根据式(5)的力矩平衡和式(2)、式(4)可得

$$\begin{cases} T_{LF} = a T_{CN} \\ T_{RN} = \frac{9.3}{90.7} (1 + a) T_{CN} \end{cases} \quad (27)$$

式中:

$$a = \frac{|x_{CN}| + \frac{9.3}{90.7} |x_{RN}| - C_{m_{JIE0}} d_e}{x_{LF} - \frac{9.3}{90.7} |x_{RN}| + C_{m_{JIE0}} d_e}$$

将式(27)代入式(26),可得

$$\ddot{h} = F_{\ddot{h}} + b_3 u_3 \quad (28)$$

式中:  $u_3 = T_{CN}$ ;  $b_3 = 100(1 + a) / (90.7m)$ 。

对式(28),同样可设计如式(24)的法向通道控制器,只是为区别起见,需将  $x_g$ 、 $b_2$ 、 $u_2$  替换为  $h$ 、 $b_3$ 、 $u_3$ 。

## 3.3 横侧向运动控制器

### 3.3.1 滚转通道控制器

滚转控制喷管的总推力  $T_{RN}$  可通过  $u_3$  由式(27)得到,在此基础上,滚转通道控制器设计的目的在于,如何将  $T_{RN}$  合理分配给  $T_{LRN}$  和  $T_{RRN}$  以保持滚转姿态稳定。

对于滚转通道运动方程:

$$\ddot{\phi} = f_{\phi} + c_3 T_{LRN} y_{LRN} + c_3 T_{RRN} y_{RRN} \quad (29)$$

将  $T_{RN} = -T_{RRN} - T_{LRN}$  和  $y_{LRN} = -y_{RRN}$  代入可得

$$\ddot{\phi} = f_{\phi} + c_3 T_{RN} y_{RRN} + 2c_3 T_{RRN} y_{RRN}$$

从而有

$$\ddot{\phi} = F_{\ddot{\phi}} + b_4 u_4 \quad (30)$$

式中:  $F_{\ddot{\phi}} = f_{\phi} + c_3 T_{RN} y_{RRN}$ ;  $u_4 = T_{RRN}$ ;  $b_4 = 2c_3 y_{RRN}$ 。

这样,即可对式(30)设计如式(23)的滚转通道控制器,为区别起见,需将  $\theta$ 、 $b_1$ 、 $u_1$  替换为  $\phi$ 、 $b_4$ 、 $u_4$ 。

### 3.3.2 偏航和侧向通道控制器

对于偏航和侧向通道运动方程:

$$\ddot{\psi} = f_{\psi} + c_9 T_{CN} x_{CN} \sin \delta_{CN_y} + c_9 T_{LF} x_{LF} \sin \delta_{LF_y} \quad (31)$$

$$\ddot{y}_g = f'_{y_g} + T_{CN} \sin \delta_{CN_y} / m + T_{LF} \sin \delta_{LF_y} / m \quad (32)$$

由于  $\delta_{CN_y}$  和  $\delta_{LF_y}$  的变化范围小,有  $\sin \delta_{CN_y} \approx \delta_{CN_y}$ 、 $\sin \delta_{LF_y} \approx \delta_{LF_y}$ ,因而可将式(31)、式(32)分别变换为

$$\ddot{\psi} = F_{\ddot{\psi}} + c_9 T_{CN} x_{CN} \delta_{CN_y} + c_9 T_{LF} x_{LF} \delta_{LF_y} \quad (33)$$

$$\ddot{y}_g = F'_{y_g} + T_{CN} \delta_{CN_y} / m + T_{LF} \delta_{LF_y} / m \quad (34)$$

其中总扰动为

$$\begin{cases} F_{\ddot{\psi}} = f_{\psi} + c_9 T_{CN} x_{CN} (\sin \delta_{CN_y} - \delta_{CN_y}) + c_9 T_{LF} x_{LF} (\sin \delta_{LF_y} - \delta_{LF_y}) \\ F'_{y_g} = f'_{y_g} + T_{CN} (\sin \delta_{CN_y} - \delta_{CN_y}) / m + T_{LF} (\sin \delta_{LF_y} - \delta_{LF_y}) / m \end{cases}$$

对于如式(33)、式(34)的双输入单输出系统,可直接取虚拟控制量:

$$\begin{cases} u_5 = c_9 T_{CN} x_{CN} \delta_{CN_y} + c_9 T_{LF} x_{LF} \delta_{LF_y} \\ u_6 = T_{CN} \delta_{CN_y} / m + T_{LF} \delta_{LF_y} / m \end{cases}$$

从而得到

$$\ddot{\psi} = F_{\ddot{\psi}} + b_5 u_5 \quad (35)$$

$$\ddot{y}_g = F'_{y_g} + b_6 u_6 \quad (36)$$

式中:  $b_5 = 1, b_6 = 1$ 。

对式(35)、式(36),可分别设计如式(23)、式(24)的偏航和侧向通道控制器,其中需将  $\theta$ 、 $b_1$ 、 $u_1$  和  $x_g$ 、 $b_2$ 、 $u_2$  分别替换为  $\psi$ 、 $b_5$ 、 $u_5$  和  $y_g$ 、 $b_6$ 、 $u_6$ 。在此基础上,可得实际控制量:

$$\begin{bmatrix} \delta_{CN_y} \\ \delta_{LF_y} \end{bmatrix} = \mathbf{B}^{-1} \begin{bmatrix} u_5 \\ u_6 \end{bmatrix}$$

式中:矩阵  $\mathbf{B} = \begin{bmatrix} c_9 T_{CN} x_{CN} & c_9 T_{LF} x_{LF} \\ T_{CN}/m & T_{LF}/m \end{bmatrix}$  是非奇异的。

## 4 悬停/平移模式控制仿真

### 4.1 系统仿真参数

根据文献[20]提供的 STOVL 验证机数据,有  $m = 13\ 608\ \text{kg}$ ,  $c_1 = 0$ ,  $c_2 = -0.6853$ ,  $c_3 = 5.1578 \times 10^{-5}$ ,  $c_4 = 0$ ,  $c_5 = 0.9507$ ,  $c_6 = 0$ ,  $c_7 = 8.0874 \times 10^{-6}$ ,  $c_8 = -0.7614$ ,  $c_9 = 7.3027 \times 10^{-6}$ 。当 V/STOL 飞机处于稳定悬停的状态时,有  $C_{L_{JIE0}} = -0.0165$ ,  $C_{m_{JIE0}} = 0.0833$ , 另有  $d_c = 2.523$ 。

假设 V/STOL 飞机在某位置稳定悬停时受到外界扰动,飞行姿态发生改变,为此设定仿真初始状态:  $\phi_0 = 10^\circ$ ,  $\theta_0 = 15^\circ$ ,  $\psi_0 = 5^\circ$ ,  $u_{b0} = v_{b0} = w_{b0} = p_0 = q_0 = r_0 = 0$ ,  $x_{g0} = 50\ \text{m}$ ,  $y_{g0} = -20\ \text{m}$ ,  $h_0 = 100\ \text{m}$ 。控制目标是使飞行姿态恢复到平衡状态,同时进行平移运动以实现定点悬停,即  $\phi_c = \theta_c = \psi_c = 0$ ,  $x_{gc} = y_{gc} = 0\ \text{m}$ ,  $h_c = 100\ \text{m}$ 。

仿真过程中,考虑到 JIE 会随着喷管偏转角度、射流速度等因素的改变而改变,根据文献[22]的试验数据,设定  $C_{L_{JIE}}$  在  $0.7C_{L_{JIE0}} \sim 1.1C_{L_{JIE0}}$ 、 $C_{m_{JIE}}$  在  $0.4C_{m_{JIE0}} \sim 1.1C_{m_{JIE0}}$  范围内变化是合理的。

另外,考虑突风对飞机的干扰,采用典型的半波长(1-cosine)离散突风模型描述风场:

$$V_w = \begin{cases} 0 & x_w < 0 \\ \frac{V_{wm}}{2} \left( 1 - \cos \frac{\pi x_w}{|d_m|} \right) & 0 \leq x_w \leq |d_m| \\ V_{wm} & x_w > |d_m| \end{cases}$$

式中:  $d_m$  为突风尺度,  $V_{wm}$  为突风强度,分别代表突风最大时的位置和速度;  $x_w$  为进入风场后的飞行距离。仿真中,设定  $x_g$ 、 $y_g$ 、 $h$  方向的  $d_m$  分别为 80、-30、150 m,对应的  $V_{wm}$  分别为 -8、10、5 m/s。

### 4.2 控制器参数整定

基于控制器带宽的概念,LADRC 的参数选取

具有很强的规律性,其具体表现如下:

1) 选择  $a_j, k_j (j = 1, 2, \dots, n + 1)$  在满足矩阵  $\mathbf{A}, \mathbf{K}$  是 Hurwitz 的前提下,使  $h(\lambda_A) = (\lambda_A + w_A)^{n+1}$ ,  $h(\lambda_K) = (\lambda_K + w_K)^{n+1}$  往往比其他选择要好,其中  $h(\lambda_A), h(\lambda_K)$  分别代表矩阵  $\mathbf{A}, \mathbf{K}$  的特征多项式,通常使  $w_A = w_K = 1$  即可,这是因为还可通过  $R$  和  $r$  来任意调节 LTD 和 LESO。

2)  $l_i (i = 1, 2, \dots, n)$  决定 LADRC 反馈控制量的大小,一般可先使  $h(\lambda_L) = (\lambda_L + w_L)^n$  且  $w_L = 1$ ,其中  $h(\lambda_L)$  代表矩阵  $\mathbf{L}$  的特征多项式,然后调节  $R$  和  $r$ 。

3) 对于  $R$ ,由于 LTD 在 LADRC 中的相对独立性, $R$  的取值变化对系统的性能影响不大,无需进行精确调节,通常取  $R \in \{10, 20, 30, 40, 50\}$  可以满足需要。

4) 需要重点调节 LESO 参数  $r$ ,对此,可兼顾控制精度与控制量的要求,在  $[0.001, 1)$  之间进行调节。通常,系统的不确定性越大, $r$  取值越小;若要减少控制量,可适当增大  $r$ 。当调节  $r$  仍无法获得满意的控制性能(如超调量或稳态误差较大)时,可对  $l_i$  进行调整。

据此整定控制器参数如下:首先取  $a_{i,1} = -1, a_{i,2} = -3, a_{i,3} = -3, k_{i,1} = 3, k_{i,2} = 3, k_{i,3} = 1, R_i = 10, i = \phi, \theta, \psi, x_g, y_g, h$ ;然后通过调试,取  $l_{\phi,1} = -0.05, l_{\theta,1} = -1, l_{\psi,1} = -0.01, l_{x_g,1} = -1, l_{y_g,1} = -0.001, l_{h,1} = -1, l_{i,2} = -2$  以及  $r_\phi = 0.002, r_\theta = 0.005, r_\psi = 0.002, r_{x_g} = 0.001, r_{y_g} = 0.6, r_h = 0.1$ 。

### 4.3 仿真结果分析

为验证 LADRC 对 V/STOL 飞机悬停/平移模式控制的有效性,设定 3 种不同的情况进行仿真比较:情况 1 为不考虑 JIE 系数变化和突风影响,情况 2 为仅考虑 JIE 系数变化,情况 3 为既考虑 JIE 系数变化又考虑突风影响。

仿真时,考虑施加较为严重的干扰:即在 0 ~ 20 s 附加  $C_{L_{JIE}}$  和  $C_{m_{JIE}}$  的随机变化,变化规律服从均匀分布;在 1 ~ 6 s 加入水平的纵向和侧向突风干扰,10 ~ 15 s 加入垂直突风干扰。此外,还需具体考虑  $T_{CN}$  的约束范围:根据式(27)的限定可知,悬停/平移模式下矢量尾喷管的最大推力  $T_{CNmax}^{HT} = 53\ 200\ \text{N}$ ;同时要求矢量发动机的推力不能为零,将其最小推力限制在总推力的 5%,即  $T_{CNmin}^{HT} = 9\ 320\ \text{N}$ 。取仿真步长  $d = 0.001$ ,仿真结果如图 2、图 3 所示。从图中可以看出:

1) 在情况 1 下(无参数摄动 + 无突风),V/STOL 飞机的姿态角以及纵向、侧向位置均能平稳

响应指令信号,无振荡和超调;但由于升力风扇和矢量尾喷管的偏转,导致力平衡被打破,飞机法向净推力小于重力,使飞机出现小幅掉高度的现象,但在推力的补偿下,飞行高度很快恢复并保持在设定值;表明 LADRC 实现协调解耦控制的有效性。此外,各推力矢量亦能在约束范围内较为平滑的变化并最终稳定在初始值,其中,  $\delta_{CN_y}^{st} = \delta_{LF_y}^{st} = 0^\circ$ ,  $\delta_{CN}^{st} = 90^\circ$ ,  $\delta_{LF}^{st} = 0^\circ$ ,  $T_{CN}^{st} = 45\ 943\ N$ ,  $T_{LF}^{st} = 76\ 851\ N$ ,  $T_{LRN}^{st} = T_{RRN}^{st} = 6\ 295\ N$ , 上标 st 代表稳定状态。

2) 在情况 2 下(有参数摄动+无突风),各飞行状态依然能够保持原有的响应性能,这是源于 LESO 和 LSEF 对参数随机摄动的实时估计和补偿,表明 LADRC 对 V/STOL 飞机的内部参数变

化具有很好的鲁棒性。但由于设定的  $C_{LJIE}$  和  $C_{mJIE}$  是随  $d$  随机变化的,在补偿作用下不免地造成  $\delta_{CN}$  和  $\delta_{LF}$  也随  $d$  作相应变化而产生高频震颤;而实际情况下,  $C_{LJIE}$  和  $C_{mJIE}$  并不会在如此短的时间内作这样随机的大幅度变化,因而不必担心作动器的工作负担;其他控制量几乎不受参数变化的影响。

3) 在情况 3 下(有参数摄动+有突风),飞机遭遇强风干扰时依然能够保持姿态响应不变,水平轨迹受到的影响也很小,虽然飞行高度受向下突风的影响而出现小幅下降,但在可接受的范围内并很快恢复,这都得益于各控制量的即时调整。表明 LADRC 具有较强的抗干扰性,对存在参数摄动和突风干扰的悬停/平移模式具有良好的解耦控制性能。

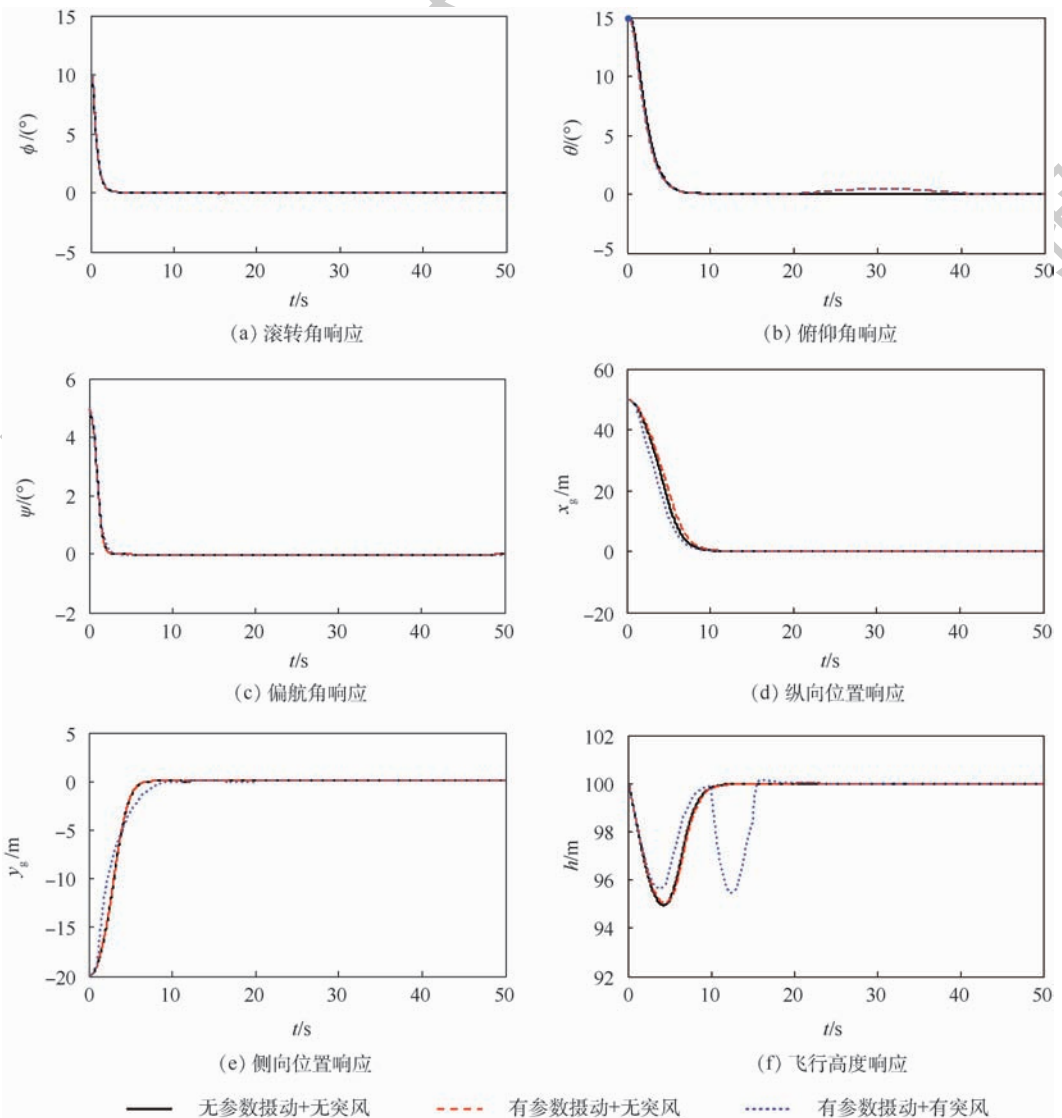


图 2 悬停/平移模式的飞行状态响应

Fig. 2 Flight state responses in hover/translation mode

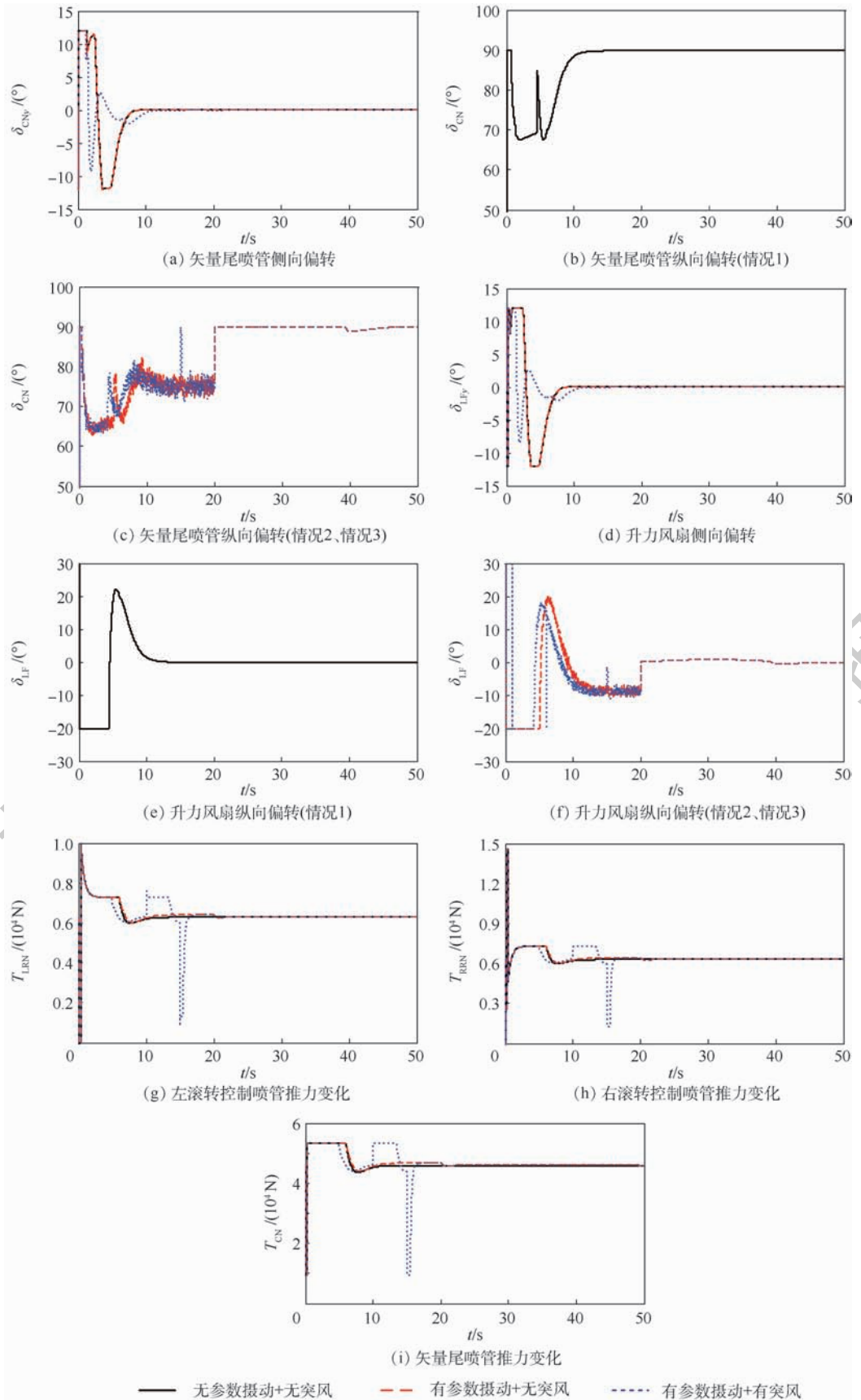


图 3 悬停/平移模式的推力矢量控制输入

Fig. 3 Thrust vector control inputs in hover/translation mode

## 5 结 论

针对固定翼 V/STOL 飞机在低速模式下受多推力矢量操纵而存在的动力学耦合、控制冗余和不确定干扰等问题,提出了一种直接式的鲁棒协调解耦飞行控制方法,有效避免了以往用多种控制方法糅合的方式来分别解决系统鲁棒性、冗余性和耦合性所产生的算法复杂、参数不易整定的控制难题。

1) 推导了一类“矢量发动机+升力风扇”型 V/STOL 飞机在低速悬停/平移模式下,受扰动风影响的全状态动力学模型,体现了多推力矢量控制下的运动特性。

2) 提出了一种基于高阶 LADRC 的鲁棒协调解耦控制策略,并在此框架内通过模型变换以提取适当控制量,而将其他控制项、动力学耦合项以及各类不确定干扰项当做总扰动,设计了 V/STOL 飞机的六通道解耦控制器。

3) 给出了控制器参数整定的一般准则,并在此基础上通过仿真比较,验证了所提出的飞行控制策略在实现动力学解耦和多操纵量分配方面的有效性,以及对参数摄动和外界干扰的鲁棒性。

### 参考文献 (References)

- [ 1 ] 索德军,梁春华,张世福,等. S/VTOL 战斗机及其推进系统的技术研究[J]. 航空发动机,2014,40(4):7-13.
- SUO D J, LIANG C H, ZHANG S F, et al. Technology of short/vertical takeoff and landing fighter and propulsion system[J]. Aeroengine, 2014, 40(4):7-13 (in Chinese).
- [ 2 ] MADDOCK I A, HIRSCHBERG M J. The quest for stable jet borne vertical lift: ASTOVL to F-35 STOVL: AIAA-2011-6999 [R]. Reston: AIAA, 2011.
- [ 3 ] WANG X Y, ZHU J H, ZHANG Y J. Dynamics modeling and analysis of thrust-vectoring V/STOL aircraft [C] // Proceedings of the 32nd Chinese Control Conference. Piscataway, NJ: IEEE Press, 2013: 1825-1830.
- [ 4 ] TIAN Y, HE Y, LI X Y, et al. Simulation testing method of V/STOL flight control strategy [C] // Proceedings of the 10th World Congress on Intelligent Control and Automation. Piscataway, NJ: IEEE Press, 2012: 1969-1975.
- [ 5 ] HAUSER J, SASTRY S, MEYER G. Nonlinear control design for slightly non-minimum phase systems: Application to V/STOL aircraft [J]. Automatica, 1992, 28(4): 665-679.
- [ 6 ] LIN F, ZHANG W, BRANDT R D. Robust hovering control of a PVTOL aircraft [J]. IEEE Transactions on Control Systems Technology, 1999, 7(3): 343-351.
- [ 7 ] CHWA D. Fuzzy adaptive output feedback tracking control of VTOL aircraft with uncertain input coupling and input-dependent disturbances [J]. IEEE Transactions on Fuzzy Systems, 2015, 23(5): 1505-1518.
- [ 8 ] CARLOS A I, MIGUEL S S C, JULIO M M, et al. Output feedback stabilization for the PVTOL aircraft system based on an exact differentiator [J]. Journal of Intelligent and Robotic Systems, 2018, 90(3-4): 443-454.
- [ 9 ] FAN Y, ZHU J H, MENG X Y, et al. Intelligent method based coordinated integrated flight control of a tailless STOVL [C] // Proceedings of the 8th World Congress on Intelligent Control and Automation. Piscataway, NJ: IEEE Press, 2010: 85-90.
- [ 10 ] ZHANG Y, ZUO J, ZHANG B. Modularized modeling and inverted dynamics control design for a lift-fan UAV in hover [C] // Proceedings of the IEEE International Conference on Information and Automation. Piscataway, NJ: IEEE Press, 2013: 570-575.
- [ 11 ] KUANG M C, ZHU J H. Hover control of a thrust-vectoring aircraft [J]. Science China Information Sciences, 2015, 58(7): 1-5.
- [ 12 ] GUO B Z, ZHAO Z L. On convergence of nonlinear active disturbance rejection for SISO systems [C] // Proceedings of the 24th Chinese Control and Decision Conference. Piscataway, NJ: IEEE Press, 2012: 3507-3512.
- [ 13 ] QI X H, LI J, XIA Y Q, et al. On the robust stability of active disturbance rejection control for SISO systems [J]. Circuits, Systems, and Signal Processing, 2017, 36(1): 65-81.
- [ 14 ] XUE W, HUANG Y. Performance analysis of 2-DOF tracking control for a class of nonlinear uncertain systems with discontinuous disturbances [J]. International Journal of Robust Nonlinear Control, 2018, 28(4): 1456-1473.
- [ 15 ] LI Z Y, LI X M, ZHOU Z Y. Active disturbance rejection controller for loitering unit with parameter uncertainty [C] // Proceedings of the 16th International Conference on Control, Automation and Systems. Piscataway, NJ: IEEE Press, 2016: 140-144.
- [ 16 ] LONG Y, DU Z J, CONG L, et al. Active disturbance rejection control based human gait tracking for lower extremity rehabilitation exoskeleton [J]. ISA Transactions, 2017, 67: 389-397.
- [ 17 ] GAO Z Q. Scaling and bandwidth-parameterization based controller tuning [C] // Proceedings of the American Control Conference. Piscataway, NJ: IEEE Press, 2003: 4989-4996.
- [ 18 ] ZHENG Q, GAO Z Q, GAO L Q. On stability analysis of active disturbance rejection control for nonlinear time-varying plants with unknown dynamics [C] // Proceedings of the 46th IEEE Conference on Decision and Control. Piscataway, NJ: IEEE Press, 2007: 3501-3506.
- [ 19 ] 高阳, 吴文海, 高丽. 高阶不确定非线性系统的线性自抗扰控制 [J/OL]. 控制与决策, (2018-10-12) [2018-11-02]. <http://kns.cnki.net/KCMS/detail/21.1124.TP.20181012.1727.010.html>.
- GAO Y, WU W H, GAO L. Linear active disturbance rejection control for high-order nonlinear systems with uncertainty [J/OL]. Control and Decision, (2018-10-12) [2018-11-02]. <http://kns.cnki.net/KCMS/detail/21.1124.TP.20181012.1727.010.html> (in Chinese).
- [ 20 ] FRANKLIN J A. Revised simulation model of the control system, displays, and propulsion system for a ASTOVL lift fan aircraft: NASA-TM-112208 [R]. Washington, D. C.: NASA,

1997.

- [21] BIRCKELBAW L G, MCNEILL W E, WARDWELL D A. Aerodynamics model for a generic ASTOVL lift-fan aircraft; NASA-TM-110347[R]. Washington, D. C. :NASA, 1995.
- [22] GUO B Z, ZHAO Z L. On convergence of tracking differentiator and application to frequency estimation of sinusoidal signals [C]// Proceedings of the 8th Asian Control Conference. Piscataway, NJ: IEEE Press, 2011: 1470-1475.

作者简介:

高阳 男,博士研究生。主要研究方向:推力矢量飞机飞行控制、非线性控制理论与应用。

吴文海 男,博士,教授,博士生导师。主要研究方向:飞行综合控制、战机攻击导引。

## High-order LADRC based robust coordinated decoupling control for V/STOL aircraft in hover/translation mode

GAO Yang<sup>1</sup>, WU Wenhai<sup>1,\*</sup>, JI Shaokang<sup>1</sup>, ZHENG Yi<sup>2</sup>

(1. Qingdao Branch, Naval Aeronautical University, Qingdao 266041, China;

2. Repair Depot, Troops 91599 of the PLA, Yantai 265200, China)

**Abstract:** To deal with the issue of dynamic coupling, thrust vector control redundancy and vulnerability to disturbed wind for vertical and/or short take-off and landing (V/STOL) aircraft in hover/translation mode, a robust coordinated decoupling flight control method based on high-order linear active disturbance rejection control (LADRC) is proposed. Firstly, the thrust vector model and the nonlinear hover/translation motion model under the disturbed wind are established according to the concept of V/STOL aircraft. On this basis, the coordinated control strategy of attitude and position is given in this mode. Then by transforming the control variables, six active disturbance rejection decoupling control laws are designed, which refrains from the redundant control of multiple thrust vectors using LADRC to compensate the total disturbances in real time. Simulation results show the effectiveness of the LADRC applied to hover/translation control and the robustness of the designed control laws for aircraft internal parameter perturbation and external gust disturbance.

**Keywords:** vertical and/or short take-off and landing (V/STOL) aircraft; linear active disturbance rejection control (LADRC); hover/translation mode; coordinated decoupling control; gust disturbance

Received: 2018-12-25; Accepted: 2019-02-02; Published online: 2019-02-21 15:10

URL: [kns.cnki.net/kcms/detail/11.2625.V.20190221.0903.002.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190221.0903.002.html)

Foundation items: National Natural Science Foundation of China (60674090, 51505491)

\* Corresponding author. E-mail: [hkdx\\_2017@126.com](mailto:hkdx_2017@126.com)

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

DOI: 10.13700/j.bh.1001-5965.2018.0767

# 基于相位条纹的高精度 GPS 码相位测量方法

傅圣友<sup>1,2</sup>, 王兆瑞<sup>2,\*</sup>, 金声震<sup>2</sup>, 艾国祥<sup>2</sup>

(1. 中国科学院大学 天文与空间科学学院, 北京 100049; 2. 中国科学院 国家天文台, 北京 100101)



**摘 要:** GPS 接收机在测量卫星到接收机的传播距离时, 通常能得到码相位和载波相位 2 个基本测量值。虽然载波相位测量值比码相位测量值精度高, 但存在整周模糊度的问题, 在实际应用中比采用码相位的技术付出的代价高很多。因此, 基于相位条纹技术, 提出了一种高精度的码相位测量方法。在传统码跟踪环的基础上, 通过提取互功率谱相位条纹的频率, 得到高精度的码相位测量值, 从而组装出高精度的码伪距。仿真实验结果表明: 在信噪比为 -15 dB 的情况下, 码相位测量误差均方差约 0.37 m, 优于传统延迟锁定环在相同条件下约 1.82 m 的跟踪精度。得到了比传统码跟踪环更高的码相位测量精度的同时, 不需要解算载波相位的整周模糊度, 对提高 GPS 定位精度具有研究意义和应用价值。

**关键词:** GPS; 互相关; 相位条纹; 码相位; 码跟踪环

**中图分类号:** P228.4; V249.3

**文献标识码:** A **文章编号:** 1001-5965(2019)09-1824-07

GPS 接收机对导航信号的跟踪主要是通过载波跟踪环和码跟踪环来完成的。其中, 码跟踪环通常是以延迟锁定环路 (Delay Locked Loop, DLL) 的形式实现的, 通过复制出 3 份不同相位的伪随机码: 超前 (early)、即时 (prompt) 和滞后 (late) 复制码, 利用伪码良好的自相关特性, 估算出即时码与接收码之间的相位差异。一方面获得码相位测量值, 另一方面同时输入给伪码数控振荡器 (NCO), 不断调节伪码发生器复制出的伪码相位, 使复制伪码和接收伪码时刻保持对齐。码相位测量值可以用于组装出信号发射时间, 进而得到码伪距测量值<sup>[1-2]</sup>。在不考虑多路径和大气延迟等接收机外部误差源的情况下, GPS 的一个 C/A 码码片长 293 m, 码伪距测量精度为 1% 码片的码环输出的码相位测量值包含大约 3 m 的误差<sup>[3]</sup>。

为了得到高精度的码伪距测量值, 国内外学者研究采用不同方法提高码跟踪环精度。文献[4]提出了采用载波多普勒的估计值辅助伪码环的环

路架构; 文献[5]提出一种自适应带宽的码跟踪环设计方案, 通过实时调整环路滤波器带宽提升码跟踪环性能; 文献[6]采用补偿迭代思想对码跟踪环进行了改造, 并改进提高了伪距测量方法; 文献[7]提出利用三点二次插值法处理相关谱谱峰数据在低采样率情况下提高码相位测量精度, 并与两点、三点线性内插法进行了对比; 文献[8]以相关峰值点为分界点对相关峰两侧分别做最小二乘拟合, 以两直线交点的横坐标与零相偏参考值的差值作为码相位的估值, 得到比最小二乘同步法和三点二次插值法更加稳定和精确的估值结果; 文献[9]提出最小二乘提高码相位测量精度; 文献[10]在其基础上, 给出了一种累加最小二乘法的实现结构, 使最小二乘法的适用范围更广、精度更高。但以上方法都是以两信号的相关函数及相关峰为基础, 需要两路或者更多路复制码信号进行相关计算, 存在一定局限性。

针对以上问题, 本文以互功率谱相位为研究

收稿日期: 2018-12-26; 录用日期: 2019-02-02; 网络出版时间: 2019-03-15 09:27

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190313.1703.002.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190313.1703.002.html)

基金项目: 国家重点研发计划 (2016YFB0501900); 国家自然科学基金 (11603041)

\* 通信作者. E-mail: zhaorui\_w@nao.cas.cn

**引用格式:** 傅圣友, 王兆瑞, 金声震, 等. 基于相位条纹的高精度 GPS 码相位测量方法[J]. 北京航空航天大学学报, 2019, 45(9): 1824-1830. FU S Y, WANG Z R, JIN S Z, et al. High-precision GPS code phase measurement method based on phase stripe[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1824-1830 (in Chinese).

对象,采用相位条纹法,在延迟锁定环的基础上,提出一种高精度的码相位测量方法。

### 1 相位条纹法原理

1981年,Piersol<sup>[11]</sup>提出了用相位参数估计延时的方法,通过测量互功率谱的相位谱回归直线斜率来估计延时。雷达技术中常用相位条纹法测量同一信号经反射后到达两相关接收机的时间延时。两路信号分别记为  $f_1(t)$  和  $f_2(t)$ , 其中,  $f_2(t) = f_1(t - \tau)$ ,  $\tau$  为信号到达两接收机的时延。对两路信号进行互相关计算,设  $\mathcal{F}[f_1(t)] = F_1(\omega)$ , 则  $\mathcal{F}[f_2(t)] = F_1(\omega)e^{-j\omega\tau}$ ,  $\mathcal{F}[\cdot]$  表示傅里叶变换。那么,两路信号的互功率谱密度为<sup>[12]</sup>  $S_{12}(\omega) = \mathcal{F}[R_{12}(t)] =$

$$F_1(\omega)F_2^*(\omega) = |F_1(\omega)|^2 e^{j\omega\tau} \quad (1)$$

式中: $R_{12}(t)$  为  $f_1(t)$  和  $f_2(t)$  的互相关函数。因此,互功率谱密度的相位为

$$\phi = \omega\tau = 2\pi\tau f \quad (2)$$

由式(2)可知,当时延  $\tau$  一定的情况下,两路信号互功率谱密度的相位  $\phi$  与频率  $f$  存在线性关系,且斜率只与信号时延  $\tau$  有关。考虑到计算出的互功率谱密度相位  $\phi$  的取值范围为  $[-\pi, \pi]$ , 所以  $\phi \sim f$  直线被切割为周期性的条纹,每条条纹斜率为  $2\pi\tau$ ,如图1所示。

图1中条纹间隔  $T_0$  为

$$T_0 = \frac{2\pi}{2\pi\tau} = \frac{1}{\tau} \quad (3)$$

条纹频率  $f_0$  (即每1 Hz 频带内条纹数)为

$$f_0 = \frac{1}{T_0} = \tau \quad (4)$$

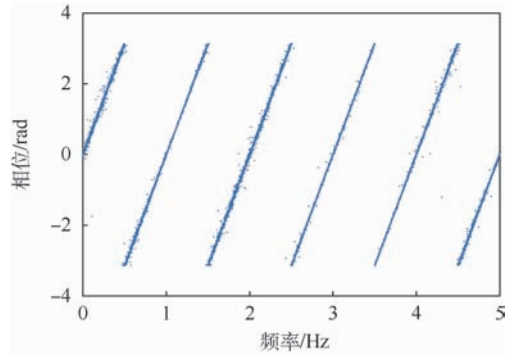


图1 延迟1s的两路相同信号互功率谱相位条纹  
Fig.1 Cross-correlation power spectrum phase stripes of two same signals with a delay of 1 s

因此,通过统计相位条纹的出现频率,即可测量出信号的时延。

### 2 相位条纹法测量码相位

在GPS接收机中,接收的卫星信号与本地NCO生成的伪码信号也可视为存在待测时延的两路相关信号,且伪码信号有足够宽的频带,有利于在频域中生成足够多的条纹,便于测量条纹出现的频率,测量出的条纹频率即是码相位测量值,可以用于组装码伪距。卫星信号经射频前端下变频之后,进入跟踪环路。当跟踪环路进入稳定跟踪状态后,采用I支路混频器输出的信号(已经消除载波相位误差和多普勒频移),与伪码NCO生成的即时支路伪随机码计算互功率谱密度,通过统计出现的相位条纹频率,即可获得码相位测量值。码跟踪环结构及相位条纹法结构(虚线部分)如图2所示。图中: $S_{IF}(n)$  为经过接收机射

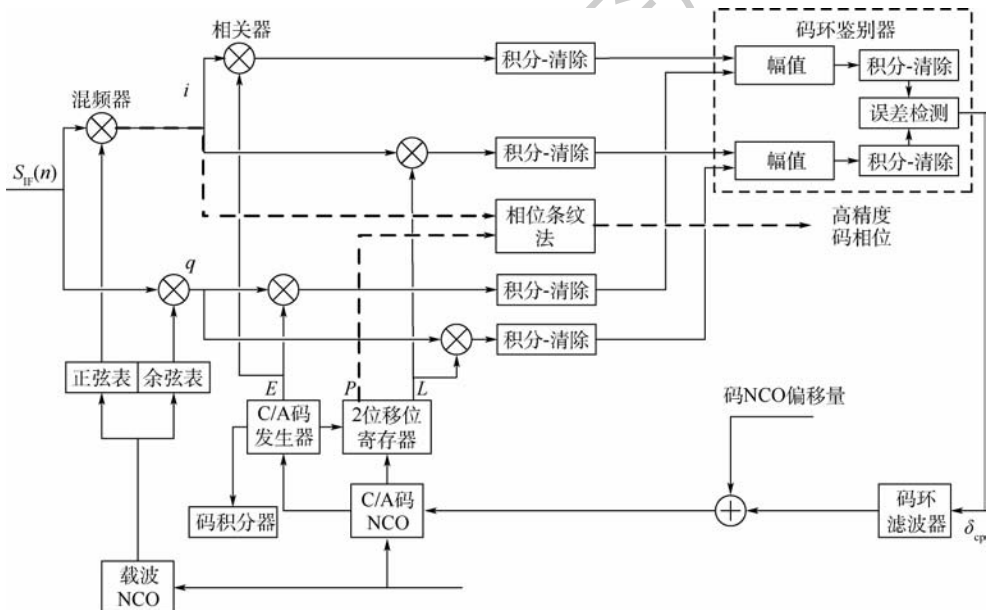


图2 典型接收机码跟踪环路及相位条纹法结构

Fig.2 Structure of typical receiver code tracking loop and phase stripe method

频前端处理采样后的中频数字信号;  $i$  和  $q$  分别表示混频后的两支路信号;  $E$ 、 $P$  和  $L$  分别为超前、即时和滞后复制码;  $\delta_{cp}$  为码环鉴别器输出的码相位测量结果。

### 2.1 互功率谱相位的计算

设经混频后的中频信号为  $f_1(t)$ , 本地生成的伪码信号为  $f_2(t)$ 。分别进行傅里叶变换, 得到  $F_1(\omega) = \mathcal{F}[f_1(t)]$ ,  $F_2(\omega) = \mathcal{F}[f_2(t)]$ 。则两路信号的互功率谱密度为

$$S_{12}(\omega) = \mathcal{F}[R_{12}(t)] = F_1(\omega)F_2^*(\omega) \quad (5)$$

其相位为复数  $S_{12}(\omega)$  的辐角主值:

$$\phi(\omega) = \arg[S_{12}(\omega)] \quad (6)$$

### 2.2 傅里叶变换时延提取和时延补偿法

计算出互功率谱的相位后, 获取码相位测量值的关键就是提取相位条纹的频率。经实验验证, 由于导航信号信噪比较低, 直接采用直线拟合效果较差, 傅里叶变换法提取条纹是更优的方式。该方法是对相位条纹  $\phi(f)$  再次进行傅里叶变换, 当条纹频率为  $f_0$  时, 条纹间隔  $T_0 = 1/f_0$ , 单个条纹  $\phi_1(f)$  可以表示为

$$\phi_1(f) = f \left[ u\left(f + \frac{1}{2f_0}\right) - u\left(f - \frac{1}{2f_0}\right) \right] \quad (7)$$

对式(7)进行傅里叶变换, 得到单个条纹的幅度谱为

$$|\Phi_1(\psi)| = |\mathcal{F}[\phi_1(f)]| = 2 \left( \frac{\sin \frac{\psi}{2f_0}}{\psi^2} - \frac{\cos \frac{\psi}{2f_0}}{\psi} \right) \quad (8)$$

式中:  $\psi$  为相位条纹函数  $\phi_1(f)$  的角频率。

当条纹以  $T_0$  为周期重复出现时, 周期性的多条相位条纹的傅里叶变换幅度谱为

$$|\Phi(\psi)| = |\mathcal{F}[\phi(f)]| = 2 \sum_{n=1}^{+\infty} \frac{\delta(\psi - 2\pi n f_0)}{n} \quad (9)$$

式中:  $\delta$  为冲激函数。

由式(9)可知, 两信号的互功率谱相位条纹  $\phi(f)$  经过再次傅里叶变换后得到幅度谱  $|\Phi(\psi)|$ , 当  $n = 1$  时,  $|\Phi(\psi)|$  的函数值为最大值, 此时  $\psi = 2\pi f_0$ , 即在  $\Phi \sim \frac{\psi}{2\pi}$  图中幅度最大值所对应的频率即为式(4)中的条纹频率  $f_0$ 。

考虑到接收机进入到稳定跟踪状态后, 本地码 NCO 生成的伪随机码已经经过了延迟锁定环的调整, 与接收信号之间的时延远小于一个码片长度。由式(4)得, 条纹频率  $f_0$  与时延  $\tau$  成正比, 时延  $\tau$  越大, 条纹数量越多。更多的条纹数使得

傅里叶变换法提取条纹的结果更接近式(9)计算的理论值。因此, 为了提高频带内条纹数量, 便于傅里叶变换法提取更精确的条纹频率, 应提前采用时延补偿法, 将本地信号额外延迟一个固定时间  $\tau_0$ , 使待测时延变为  $(\tau + \tau_0)$ , 其中  $\tau_0 \geq \tau$ 。用傅里叶变换法测出补偿后的两信号时延, 再减去补偿时延  $\tau_0$ , 即可获得原始时延即码相位测量值。

### 2.3 时域折叠法提高信噪比

接收信号中包含的噪声也会对条纹频率的提取造成一定程度的影响。当 GPS 卫星处于天顶方向(仰角  $90^\circ$ )时, 信号经过射频前端处理, 进入跟踪环路后, 噪声带宽为 C/A 码信号的带宽 2.046 MHz, 此时信噪比(SNR)约为  $-16.5 \text{ dB}^{[2]}$ 。而当卫星处于低仰角(约  $5^\circ$ )时, 中频信号信噪比约为  $-24.5 \text{ dB}$ 。较低的信噪比使得相位条纹淹没在噪声之中, 给提取条纹频率带来很大的困难。

时域折叠原理如图 3 所示。

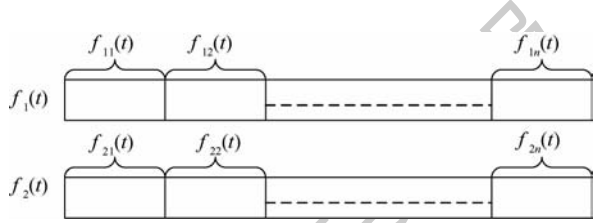


图 3 时域折叠法示意图

Fig. 3 Schematic diagram of time domain folding method

两段信号数据均被分为均匀的  $n$  段数据段, 分别表示为  $f_{11}(t), f_{12}(t), \dots, f_{1n}(t)$  和  $f_{21}(t), f_{22}(t), \dots, f_{2n}(t)$ 。经过分段后, 将每段数据叠加求和, 得到时域折叠后的信号数据  $f'_1(t)$  和  $f'_2(t)$  为

$$\begin{cases} f'_1(t) = \sum_{i=1}^n f_{1i}(t) \\ f'_2(t) = \sum_{i=1}^n f_{2i}(t) \end{cases} \quad (10)$$

GPS 信号中, 导航电文的数据码码长为 20 ms, 伪码周期仅为 1 ms。也就是在 20 ms 内, 可以忽略导航电文对伪码相位的影响。在此条件下, 对接收信号进行每段 1 ms 时长的时域折叠, 可以有效减小噪声功率, 提高信噪比。当折叠段数为 10 段时, 信噪比将提高约 10 dB。经时域折叠提高信噪比之后的两路信号  $f'_1(t)$  和  $f'_2(t)$  再计算互功率谱相位, 即可得到较为清晰的相位条纹。

### 2.4 相位条纹法实现步骤

相位条纹法测量码相位的完整步骤如图 4 所示,该方法需要接收机射频前端下变频的中频卫星信号,以及跟踪环路进入稳定跟踪状态后的本地复制载波和本地复制码。首先,将中频信号和本地复制载波经过混频器,滤除载波和多普勒频移,得到接收输入信号  $f_1(t)$ 。本地复制码作为本地输入信号  $f_2(t)$ ,经过时延补偿,得到  $f_2(t - \tau_0)$ ,与接收信号  $f_1(t)$  之间间隔一个较大的时延。其次,分别对

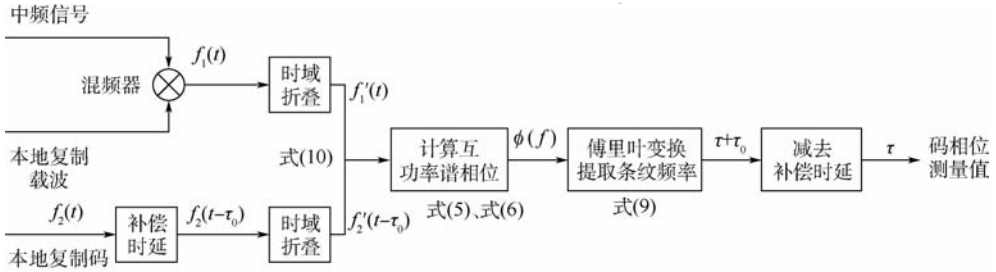


图 4 相位条纹法流程图

Fig. 4 Flowchart of phase stripe method

### 3 仿真实验

采用 MATLAB 软件编程对第 2 节方法进行仿真实验,仿真参数如表 1 所示。表中,码速率的单位 cps 表示码片每秒。通过软件编程,模拟生成卫星信号下变频之后的中频信号、本地载波信号和码信号。经 2.4 节的步骤后,仿真结果如图 5 ~ 图 8 所示。

图 5 为模拟经过射频前端处理之后进入跟踪环路的中频信号,有效信号已经完全淹没在噪声中;图 6 为中频信号经过与本地载波信号混频,并进行时域折叠后的信号,可以发现噪声幅度明显减小,有效地提高了信号信噪比。

图 7 为图 6 中的信号与本地码信号的互功率谱相位,低频段有明显的倾斜条纹,高频段的条纹被淹没在噪声中。其中部分集中在 0 相位的数据点是互功率谱幅值较小的数据,为了防止

表 1 仿真参数

Table 1 Simulation parameters

参数	数值
信号长度/ms	10
中频载波频率/MHz	4 <sup>[13]</sup>
伪码长度/bit	1 023
码速率/Mcps	1.023
采样频率/MHz	100
设定待测时延/ns	10
补偿时延/ns	1 000
时域折叠段数	10

接收信号和本地码信号进行式(10)所示的相同时间的信号时域折叠,得到高信噪比的信号  $f_1'(t)$  和  $f_2'(t - \tau_0)$ ,并保持两路信号的相关性。然后,通过式(5)和式(6),计算两信号的互功率谱相位,生成相位条纹。对相位条纹再次进行式(9)所示的傅里叶变换,可以得到二次变换幅度谱  $|\Phi(\psi)|$  最大值对应的条纹频率,即两信号的间隔时延为  $(\tau + \tau_0)$ 。最后,从测量时延中减去补偿的时延量,即可获得比延迟锁定环更高精度的码相位测量值  $\tau$ 。

由于 MATLAB 软件的舍入误差造成极大的相位偏差而被设为 0 相位。图 8 为采用傅里叶变换法提取条纹频率时,对图 7 的数据进行再次傅里叶变换后的功率谱幅度,最大值对应频率即为测得时延。将此结果减去补偿时间,即可得到码相位

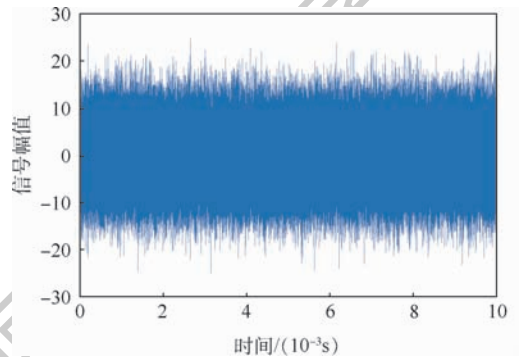


图 5 输入的中频信号

Fig. 5 Input IF signal

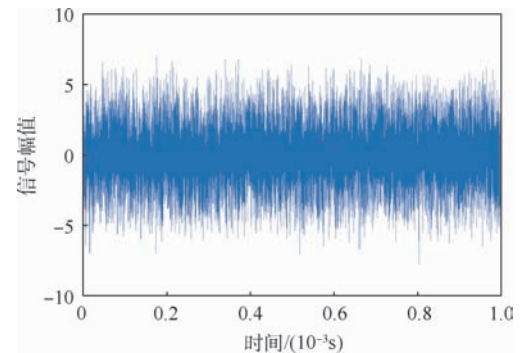


图 6 时域折叠后的信号

Fig. 6 Folded signal in time domain

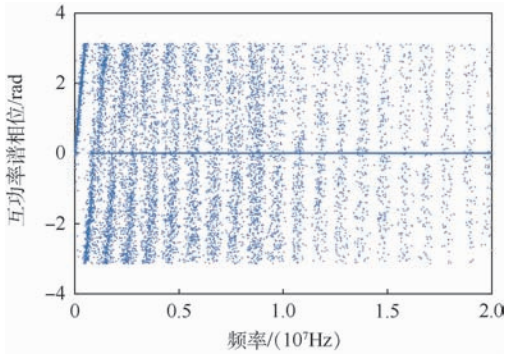


图7 接收信号与本地码信号的互功率谱相位条纹  
Fig.7 Cross-correlation power spectrum phase stripes between received signal and local code signal

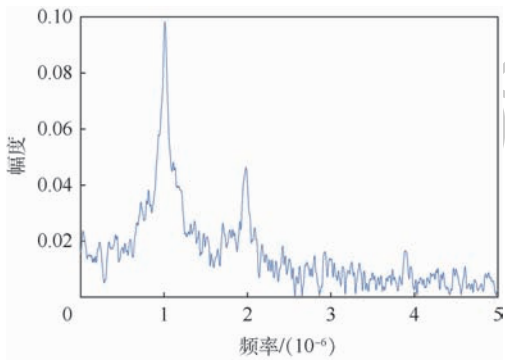


图8 条纹频率  
Fig.8 Stripe frequency

测量值。

采用以上程序,对于不同的信噪比情况,分别进行1000次蒙特卡罗仿真模拟,并与相同情况下延迟锁定环的测量误差进行对比。延迟锁定环常用的3种码鉴相器算法中,相比于非相干前减后幅值法和非相干前减后功率法,相干点积功率法的运算量较小,且得到的码相位测量值最为准确<sup>[14]</sup>。此处选取采用相干点积功率法鉴相器的延迟锁定环进行对比,其码相位测量误差均方差(RMSE)计算式为<sup>[2,15]</sup>

$$\sigma_{\Delta r} = cT_c \sqrt{\frac{d}{4TB_n \cdot \text{SNR}}} \quad (11)$$

式中: $c$ 为光速; $T_c$ 为伪码码宽,此处取值C/A码码宽,即 $\frac{1}{1023} \times 10^{-3}$ s; $d$ 为码环前后相关器间距,

按照目前大部分接收机制造商采用的相关器间距,此处取值为0.1码片<sup>[3]</sup>;  $T$ 为信号长度,即0.01s;  $B_n$ 为噪声带宽,此处取值为C/A码信号频谱主峰带宽2.046MHz; SNR为信号信噪比,与相位条纹法的仿真实验取值一致。对比结果如图9和表2所示。

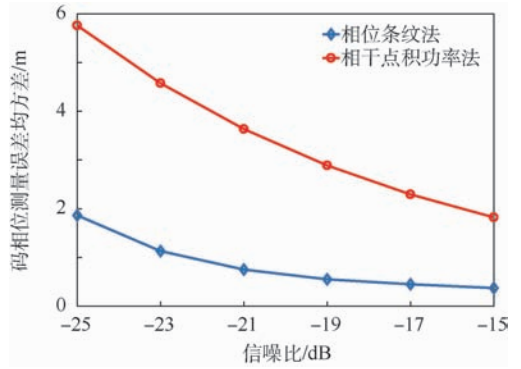


图9 码相位测量误差均方差对比  
Fig.9 RMSE of code phase measurement error comparison

表2 码相位测量误差均方差结果  
Table 2 RMSE results of code phase measurement error

信噪比/dB	码相位测量误差均方差/m	
	相位条纹法	相干点积功率法
-15	0.3711	1.8217
-17	0.4487	2.2933
-19	0.5497	2.8871
-21	0.7504	3.6347
-23	1.1292	4.5758
-25	1.8632	5.7606

## 4 结论

1) 根据本文的仿真实验结果,当卫星处于天顶方向,接收信号信噪比为-15dB时,码相位测量误差均方差可以达到约0.37m,约为一个码片宽度(293m)的0.13%,比延迟锁定环测量误差均方差1.82m提高了约80%。

2) 随着信噪比的下降,测量误差会逐渐增大。而当卫星处于低仰角(约5°)时,信噪比下降至约-25dB,测量误差均方差增大到1.86m,为码片宽度的0.63%,但仍比延迟锁定环测量误差均方差5.76m提高约68%。

3) 与传统跟踪环路相比,本文方法少量增加了接收机处理器的计算量和计算时间,显著提高了码伪距的测量精度,且不存在整周模糊度的问题。在接收机的实际应用中,可以用较低的成本提供高精度的码伪距测量值,具有十分广泛的应用价值。同时也能为高精度测距和定位提供一定的参考作用。

为使本文方法能更加准确、高效地实现码相位测量,仍需要优化部分参数和细节计算。

## 参考文献 (References)

[1] 谢钢. GPS原理与接收机设计[M]. 北京:电子工业出版社

- 社,2017.
- XIE G. Principles of GPS and receiver design [M]. Beijing: Publishing House of Electronics Industry,2017(in Chinese).
- [2] 鲁郁.北斗/GPS双模软件接收机原理与实现技术[M].北京:电子工业出版社,2016.
- LU Y. BeiDou/GPS dual-mode software receiver principle and implementation technology [M]. Beijing: Publishing House of Electronics Industry,2016(in Chinese).
- [3] MISRA P, ENGE P. Global positioning system: Signals, measurements, and performance [M]. 2nd ed. Warszawa: Ganga-Jamuna Press,2006.
- [4] 罗柳镇.高精度卫星测距技术研究及实现[D].南京:南京理工大学,2016:37-43.
- LUO L Z. Research and realization of high precision satellite ranging technology[D]. Nanjing:Nanjing University of Science and Technology,2016:37-43(in Chinese).
- [5] 刘盟超,赵丙风. GNSS接收机自适应带宽伪码跟踪环路设计[J].无线电通信技术,2017,43(4):64-66.
- LIU M C, ZHAO B F. Design of adaptive bandwidth DLL for GNSS receivers[J]. Radio Communications Technology,2017,43(4):64-66(in Chinese).
- [6] 邓中亮,赵洋,尹露,等.一种提高伪距测量精度方法研究[C]//第六届中国卫星导航学术年会论文集.北京:中国卫星导航系统管理办公室学术交流中心,2015:4.
- DENG Z L, ZHAO Y, YIN L, et al. An improved pseudo-range measurement method research [C]//6th China Satellite Navigation Conference (CSNC). Beijing: Academic Exchange Centre of China Satellite Navigation Office,2015:4(in Chinese).
- [7] 龚国辉,李思昆.提高DSSS信号PN码相位测量精度的三点二次插值法[J].通信学报,2007,28(2):130-133.
- GONG G H, LI S K. Improving DSSS signal PN code phase measurement precision by 3-points quadratic interpolation[J]. Journal on Communications, 2007, 28(2):130-133(in Chinese).
- [8] 罗海军,彭卫东,李明阳,等.基于分段直线拟合的伪随机码相位测量法[J].计算机测量与控制,2015,23(3):727-729.
- LUO H J, PENG W D, LI M Y, et al. Method of phase measurement of PN codes based on piecewise linear fitting[J]. Computer Measurement and Control, 2015, 23(3):727-729(in Chinese).
- [9] 胡修林,张俊,杨灵.直接序列扩频通信系统中PN码相位的精确测量方法[J].电讯技术,2005,45(3):128-131.
- HU X L, ZHANG J, YANG L. A method for accurately measuring the phase of PN code in DSSS systems[J]. Telecommunication Engineering,2005,45(3):128-131(in Chinese).
- [10] 张波,郭英,齐子森,等. PN码相位精确测量的累加最小二乘法[J].北京航空航天大学学报,2016,42(10):2265-2270.
- ZHANG B, GUO Y, QI Z S, et al. Precise measurement of PN code phase by accumulative least square method[J]. Journal of Beijing University of Aeronautics and Astronautics, 2016, 42(10):2265-2270(in Chinese).
- [11] PIERSOL A G. Time delay estimation using phase data[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing,1981,29(3):471-477.
- [12] 熊庆旭,刘锋,常青.信号与系统[M].北京:高等教育出版社,2011.
- XIONG Q X, LIU F, CHANG Q. Signals and systems[M]. Beijing: Higher Education Press,2011(in Chinese).
- [13] BRAASCH M S, VAN DIERENDONCK A J. GPS receiver architectures and measurements[J]. Proceedings of the IEEE, 1999,87(1):48-64.
- [14] 王松瑞.北斗二代接收机捕获跟踪算法研究与实现[D].西安:西安电子科技大学,2017:37-38.
- WANG S R. Research and implementation of the BD-II receiver acquisition and tracking algorithm[D]. Xi'an: Xidian University,2017:37-38(in Chinese).
- [15] VAN DIERENDONCK A J, FENTON P, FORD T. Theory and performance of narrow correlator spacing in a GPS receiver[J]. Navigation,1992,39(3):265-283.

#### 作者简介:

傅圣友 男,博士研究生。主要研究方向:卫星导航。

王兆瑞 男,博士,副研究员。主要研究方向:卫星导航、非线性信号处理。

金声震 男,博士,教授。主要研究方向:雷达信号处理。

艾国祥 男,院士,教授,博士生导师。主要研究方向:天体物理、卫星导航。

## High-precision GPS code phase measurement method based on phase stripe

FU Shengyou<sup>1,2</sup>, WANG Zhaorui<sup>2,\*</sup>, JIN Shengzhen<sup>2</sup>, AI Guoxiang<sup>2</sup>

(1. School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 100049, China;

2. National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China)

**Abstract:** When GPS receiver measures the propagation distance from satellites to the receiver, it can usually obtain two basic measurement values, code phase and carrier phase. Although the precision of carrier phase measurement is higher than that of code phase measurement, there exists the integer ambiguity problem, whose cost is much higher than that of using code phase technique in practical application. Therefore, based on the phase stripe technique, a high-precision code phase measurement method was proposed. On the basis of the traditional code tracking loop, the high-precision code phase measurement value is obtained by extracting the frequency of the phase stripes of the cross-correlation power spectrum, so as to assemble a high-precision code pseudo-range. Simulation results show that the precision of code phase measurement is about 0.37 m when the SNR is -15 dB, which is better than the tracking accuracy of 1.82 m under the same condition of traditional delay lock loop. At the same time of obtaining the higher-precision code phase measurements, the integer ambiguity of the carrier phase does not need to be solved, which has research significance and application value for improving the GPS positioning precision.

**Keywords:** GPS; cross-correlation; phase stripe; code phase; code tracking loop

**Received:** 2018-12-26; **Accepted:** 2019-02-02; **Published online:** 2019-03-15 09:27

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20190313.1703.002.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190313.1703.002.html)

**Foundation items:** National Key R & D Program of China (2016YFB0501900); National Natural Science Foundation of China (11603041)

\* **Corresponding author.** E-mail: zhaorui\_w@nao.cas.cn

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

DOI: 10.13700/j.bh.1001-5965.2018.0769

# 基于 DWNN 的机电作动器渐变性故障诊断

王剑, 王新民, 谢蓉\*, 曹宇燕, 李婷

(西北工业大学 自动化学院, 西安 710072)



**摘 要:** 飞机飞行控制系统机电作动器(EMA)的渐变性故障很难准确预判,若不能及早发现而任其发展就会影响到飞机的飞行安全性。针对EMA的渐变性故障,提出一种基于动态小波神经网络(DWNN)的故障诊断方法。首先,利用EMA在电机电枢绕组匝间短路、传动装置丝杆和滚珠磨损等多种渐变性故障状态下的运行数据来训练DWNN故障诊断模型;然后,利用训练好的DWNN模型对EMA渐变性故障进行诊断。创新之处在于DWNN模型利用小波分解算法去除了传感器测量信号中高频分量的影响,利用反馈神经网络的记忆能力融合了过去输入的信息和过去预测的信息,提高了对EMA渐变性故障诊断的准确性。通过对某型EMA进行故障诊断实验,仿真结果表明所提出的DWNN方法可以实现对EMA部件渐变性故障的准确诊断。

**关键词:** 机电作动器(EMA); 渐变性; 故障; 诊断; 动态小波神经网络(DWNN)

**中图分类号:** V242.5; TP277

**文献标识码:** A **文章编号:** 1001-5965(2019)09-1831-07

机电作动器(Electromechanical Actuator, EMA)是飞机飞行控制系统中的重要部件。EMA出现故障,不仅会影响飞行控制系统的工作,而且还有可能导致飞机存在失控的风险<sup>[1]</sup>。不同于对突发性故障的事后诊断,渐变性故障诊断要在系统部件失效发生的前期,准确预测出部件故障的演变趋势,及时地向飞机操作人员和维护人员提供预警和维护信号。渐变性故障的诊断难点在于,轴承磨损等渐变性故障的故障信号在退化初期变化非常缓慢,直到接近失效时才剧烈变化,因而在故障初期获得的训练数据变化也很缓慢,很难准确估计出渐变性故障经过多么长时间的演化将失效<sup>[2]</sup>。EMA渐变性故障的准确诊断是一项极具挑战性的任务<sup>[1,3]</sup>。

目前机电系统的故障诊断方法一般分为两大类:基于模型的故障诊断方法和基于数据的故障

诊断方法。基于模型的故障诊断方法依赖于获得较为精确的系统动态模型<sup>[4]</sup>。EMA的物理模型可以用一阶微分方程逼近,若利用EMA的物理模型确定各种故障模式对应的故障特征向量,结合基于数据的故障诊断方法将会获得更为准确的诊断结果<sup>[5]</sup>。递推最小二乘(Recursive Least Square, RLS)方法常被用来辨识模型中的故障特征参数,比如利用自回归滑动平均模型(Auto Regressive Moving Average model, ARMA)可以通过RLS方法来辨识EMA中的故障特征参数的实时取值<sup>[6]</sup>。RLS方法的不足之处在于,渐变性故障的初始阶段故障参数变化不明显,预测精度较差<sup>[7]</sup>。基于数据的故障诊断方法,无需精确的数学模型,但需要一个完备的数据库来训练和验证诊断算法。相对而言,基于数据的故障诊断方法具有自学习、自组织和自推理能力,使用起来更加

收稿日期: 2018-12-29; 录用日期: 2019-02-02; 网络出版时间: 2019-02-21 15:22

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190221.0934.003.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190221.0934.003.html)

基金项目: 国家自然科学基金(61703341)

\* 通信作者. E-mail: xierong@nwpu.edu.cn

**引用格式:** 王剑, 王新民, 谢蓉, 等. 基于DWNN的机电作动器渐变性故障诊断[J]. 北京航空航天大学学报, 2019, 45(9): 1831-1837. WANG J, WANG X M, XIE R, et al. Gradual fault diagnosis for electromechanical actuator based on DWNN[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1831-1837 (in Chinese).

灵活<sup>[8]</sup>。比如,基于特定的故障模式和属性,小波神经网络(Wavelet Neural Network, WNN)、多项式神经网络、概率神经网络、径向基神经网络、神经模糊结构等多种类型的模糊逻辑和神经网络工具已经被广泛应用于各种机电系统的故障诊断<sup>[9-12]</sup>。使用神经网络方法可以利用神经网络融合故障信号测量数据,建立故障特征与故障类别之间的相互关系,对渐变性故障进行预测<sup>[7]</sup>。但是,单纯使用神经网络方法难以满足EMA渐变性故障的预诊断的精度需求。

为此,本文提出一种基于动态小波神经网络(Dynamic Wavelet Neural Network, DWNN)的EMA渐变性故障诊断方法。利用小波多级分解算法来对传感器接收信号的频率分量进行取舍,利用动态神经网络的反馈机制来预测故障特征参数的演变趋势。与ARMA方法相比,DWNN方法利用小波分解算法去除了传感器测量信号中高频分量的影响,便于准确估计故障特征参数变化趋势,对EMA渐变性故障进行预诊断。

## 1 问题描述

### 1.1 EMA故障诊断特征向量

EMA是一种机电伺服系统,通过控制电机转速来带动减速传动机构,输出力矩驱动舵面负载。EMA的物理模型可以简化为一阶微分方程<sup>[13]</sup>:

$$\begin{cases} L_a \frac{di_a}{dt} = -R_a i_a - K_e \omega_m + v_a \\ J_m \frac{d\omega_m}{dt} = K_t i_a - B_m \omega_m - T_L \\ \frac{d\theta_m}{dt} = \omega_m \end{cases} \quad (1)$$

式中: $J_m$ 为EMA机电伺服系统的惯性; $B_m$ 为EMA机电伺服系统的阻尼; $L_a$ 为电机的电枢电感; $R_a$ 为电机的电枢电阻; $v_a$ 和 $i_a$ 分别为电机的电枢电压和电枢电流; $\omega_m$ 和 $\theta_m$ 分别为转速和电枢角度; $K_e$ 为电磁系数; $K_t$ 为转矩系数; $T_L$ 为转子的负载转矩。

下面分析EMA中部件存在的典型渐变性故障。首先,电枢绕组匝间短路故障是电机常见性缺陷,轻微匝间短路,EMA仍可继续运行;当匝间短路严重时,电枢电流显著增大,绕组温度升高,使电机振动加剧,甚至被迫停机。在EMA模型式(1)中,电枢绕组匝间短路表现为电机的电枢电感 $L_a$ 和电枢电阻 $R_a$ 发生变化。其次,EMA工作时其中机电传动装置的磨损是不可避免的,传动装置丝杆和滚珠磨损须经过长时间的演化,

最终演变为故障或失效,当滚珠丝杆副磨损严重时,在EMA模型式(1)中机电伺服系统的阻尼 $B_m$ 会发生变化。另外,EMA中控制面损失,将会导致EMA机电伺服系统的惯性 $J_m$ 发生变化。综上,EMA潜在故障可以通过实际的物理参数来反映<sup>[13]</sup>,即其故障特征向量可以选取为

$$\mathfrak{D}(t) = [J_m \quad B_m \quad L_a \quad R_a]^T$$

若能测得故障特征向量中各个参数的变化过程,将有助于判断EMA发生的电枢绕组匝间短路和传动机构过度磨损等典型部件的渐变性故障。但是这些故障参数不能由传感器直接测量。已有的研究表明使用RLS方法可以辨识出EMA的一阶差分方程,进而求出这些参数,使用RLS方法存在的困难是由于在渐变性故障的初始阶段故障参数变化不明显,预测精度较差,且不能同时辨识上述参数<sup>[14-15]</sup>。为了更准确地诊断EMA中发生的渐变性故障,并且能够处理同时发生的多个渐变性故障,本文使用DWNN模型具有的并行处理和自学习能力来辨识这些参数,对EMA的渐变性故障进行诊断。

### 1.2 DWNN模型

小波函数具有多分辨率特征,可以处理EMA故障中所需要的不同带宽数据,也可以处理某一传感器采集到信号中的多个频率分量。利用WNN模型可以将故障特征指示到相应的故障模式<sup>[16]</sup>。典型的WNN神经元函数 $f: \mathbf{R}^n \rightarrow \mathbf{R}$ 可以描述为

$$f(x) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} c_{j,k} \psi_{j,k}(x) \quad (2)$$

式中:系数 $c_{j,k}$ 为函数 $f$ 在小波函数 $\psi_{j,k}(x)$ 的投影, $j$ 和 $k$ 分别为伸缩因子和平移因子。 $\psi_{j,k}(x)$ 具有时间和频率变换内核,可描述为

$$\psi_{j,k}(x) = \sqrt{2^j} \psi(2^j x - k) \quad (3)$$

式中:小波函数 $\psi_{j,k}(x)$ 由其母函数 $\psi(x)$ 伸缩和平移得到。小波的多分辨率特性主要体现在 $j$ 和 $k$ 的取值。

利用小波的多分辨率特性,可以对采集到的信号在时域和频域权衡取舍。WNN是一种静态模型,只能反映输入与输出之间建立的一种静态关系<sup>[17]</sup>。为了对EMA系统故障的时间演变进行建模,需要在WNN的基础上引入反馈,形成DWNN模型。

一般反馈环节的记忆能力可以描述为

$$\mathbf{y}_{k+1} = \alpha \mathbf{y}_k + \mathbf{u}_k \quad (4)$$

式中: $\mathbf{u}_k$ 为 $k$ 时刻的控制输入; $\mathbf{y}_{k+1}$ 为 $k+1$ 时刻的输出值; $\alpha$ 为自反馈系数, $0 \leq \alpha \leq 1$ 。对式(4)

进行迭代计算可得

$$y_{k+1} = u_k + \alpha u_{k-1} + \alpha^2 u_{k-2} + \dots \quad (5)$$

可见,反馈单元的输出是其输入过去值的滑动平均和,因反馈的存在,提供了对输入信号的一个无限的记忆。对 EMA 中的故障特征参数而言,随采样点之间距离的加大,它们之间的互相影响应该越来越小,因此,对输入信号进行了截断处理,只保留了  $N$  个。另外,为了弥补由于截断输入信号所造成的相关信息损失,利用反馈网络的自身的记忆能力对输出信号时间序列进行了相空间重构,得到重构维数为  $M$  的预测信息。

将以上的反馈环节引入 WNN,形成 DWNN,其基本结构如图 1 所示。

图 1 所示的 DWNN 结构可以用公式描述为

$$y_k = f_{NN}(y_{k-1}, y_{k-2}, \dots, y_{k-M}, u_k, u_{k-1}, \dots, u_{k-N}) \quad (6)$$

式中: $f_{NN}$ 为神经网络函数; $N$ 为控制输入  $u_k$  保留的过去输入值的数目; $M$ 为输出  $y_k$  保留的过去输出值的数目。例如,本文  $u_k$  为电机在  $k$  时刻的测量电流, $y_k$  为电机在  $k$  时刻的转速或故障特征参数。式(6)可以抽象理解为差分方程,这与 EMA 的数学模型非常契合。

DWNN 模型遵循递归方式,能够将历史的和现在的数据映射到未来的数据,非常适合复杂系统故障的分类和预测<sup>[18]</sup>。EMA 是一个由电机、减速装置、电控单元等部件组合而成的复杂系统,其故障模式包括机械振动、卡死、阻塞等事件,也包括电气和电子故障事件,比如短路、传感器漂移、器件失效等。本文设计的 DWNN 模型能够从传感器数据中获得重要的故障特征,从而判断故障的位置、程度和演变趋势。

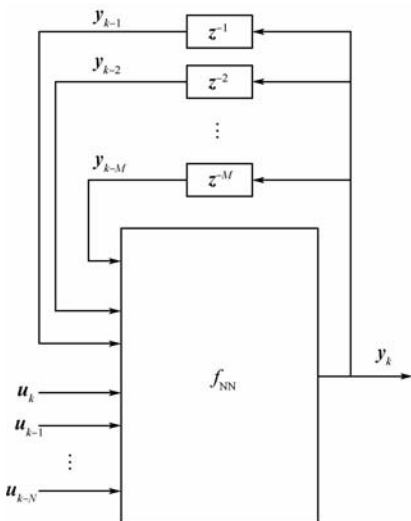


图 1 DWNN 基本结构

Fig. 1 Basic structure of DWNN

### 1.3 EMA 故障诊断问题

神经网络训练本质上是状态方程和测量方程的状态估计问题。在 DWNN 网络训练时,将权重  $w$  作为状态变量,并将输入和传感器的测量值  $\{u_i, z_i\}_{i=1}^{N_s}$  作为样本进行训练,其中  $N_s$  为训练样本总数。其状态方程和测量方程描述为

$$\begin{cases} w_{k+1} = w_k + \varepsilon_k \\ z_{k+1} = h(z_k, u_k, w_k) + v_k \end{cases} \quad (7)$$

式中: $h(\cdot)$ 为神经网络输入输出关系函数; $\varepsilon$ 为零均值的过程噪声向量,其协方差矩阵为  $\text{cov}(\varepsilon) = E(\varepsilon\varepsilon^T) = Q$ ; $v$ 为零均值的测量噪声向量,其协方差矩阵为  $\text{cov}(v) = E(vv^T) = R$ 。

设计 DWNN 模型需要确定神经网络的结构和参数,具体包括输入、输出的数目,以及小波节点数目和拓扑。根据式(2),神经元函数中包含无穷多个小波,工程上必须要对小波函数的数据进行截断。本文采用了  $k$  均值聚类来确定神经网络的节点数,每个聚类代表着一个节点。另外,记忆数据(即式(6)中  $M$  和  $N$  的取值)的选取会影响到故障的预测结果。若  $M$  和  $N$  的取值过大,可能会产生过拟合的预测结果;若  $M$  和  $N$  的取值偏小,会影响预测结果的准确性。本文根据 EMA 的物理方程来确定  $M$  和  $N$  的取值。

## 2 故障诊断算法

### 2.1 算法描述

给定训练样本  $\{u_i, z_i\}_{i=1}^{N_s}$ ,来训练 DWNN 的权重参数,是本文故障诊断算法的主要目的。本文使用了解耦无迹卡尔曼滤波器(Decoupling Unscented Kalman Filter, DUKF)来完成这一工作。采用 DUKF 的训练框架如图 2 所示。

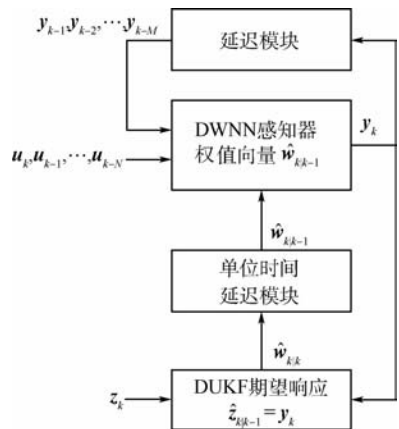


图 2 DUKF 训练框架

Fig. 2 Training framework of DUKF

图2描述了 DWNN 的监督学习过程。当 DWNN 的输出向量  $y_k$  与传感器实际测得的值  $z_k$  之间存在差异,就可以通过 DUKF 获得权重  $w_k$  的变化,进而预测 EMA 渐变故障发生的可能性。

### 2.2 DUKF 算法

在所设计 DWNN 中,采用了 DUKF 来跟踪模型中不同频率分量的信号状态。与 UKF 算法相比,DUKF 算法忽略了神经网络取值之间的耦合,将多个 0 值引入到协方差矩阵  $P_{k|k}$  中,降低了滤波器算法的计算复杂度。

采用 DUKF 算法对 DWNN 感知器输出信号进行处理。DUKF 通过在频率分量信号均值附近产生一组 Sigma 点,然后通过加权产生方差预测估计。其中,权重和状态的时域更新过程同式(7)。在滤波处理过程中,本文采用比例修正 Sigma 点的方法,将方差为  $P_{k|k+1}$  的  $n$  维随机向量  $\hat{w}_{k|k-1}$  扩展为  $2n+1$  个列向量<sup>[19]</sup>:

$$\begin{cases} \chi_0 = \hat{w}_{k|k-1} \\ \chi_i = \hat{w}_{k|k-1} - \sqrt{(n+\lambda)P_{k|k-1}} \Big|_i & i=1,2,\dots,n \\ \chi_i = \hat{w}_{k|k-1} + \sqrt{(n+\lambda)P_{k|k-1}} \Big|_{i-n} & i=n+1,n+2,\dots,2n \end{cases} \quad (8)$$

式中:  $\chi_i (i=0,1,\dots,2n)$  表示由随机向量  $\hat{w}_{k|k-1}$  扩展生产的  $2n+1$  个 Sigma 点;  $\lambda = \varpi^2(n+\kappa) - n$ ,  $\varpi (0 < \varpi < 1)$  为任意小的参数,参数  $\kappa = 0$ ;  $(\cdot) \Big|_i$  表示取矩阵  $(\cdot)$  的第  $i$  列元素。

#### 1) DUKF 测量值更新

产生测量方程的对应 Sigma 点,并计算预测估计值:

$$\begin{cases} Z_i = h(\chi_i, z_k, u_k) \\ \hat{z}_{k|k-1} = \sum_{i=0}^{2n} W_i Z_i \end{cases} \quad i=0,1,\dots,2n \quad (9)$$

计算卡尔曼增益  $K_k$ :

$$\begin{cases} K_k = (P_{xz})_k (P_{zz})_k^{-1} \\ (P_{zz})_k = R_k + \sum_{i=0}^{2n} (W_i [Z_i - \hat{z}_{k|k-1}] \cdot [Z_i - \hat{z}_{k|k-1}]^T) \\ (P_{xz})_k = \sum_{i=0}^{2n} (W_i [\chi_i - w_{k|k-1}] \cdot [Z_i - \hat{z}_{k|k-1}]^T) \end{cases} \quad (10)$$

式中:  $Z_i$  为 Sigma 点对应的观测向量;  $W_i$  为 Sigma 点  $\chi_i$  的加权因子,计算式为

$$W_i = \begin{cases} \lambda/(n+\lambda) & i=0 \\ 1/[2(n+\lambda)] & i=1,2,\dots,2n \end{cases} \quad (11)$$

#### 2) DUKF 估计值更新

$$\hat{w}_{k|k} = \hat{w}_{k|k-1} + K_k (z_k - \hat{z}_{k|k-1}) \quad (12)$$

误差协方差更新:

$$P_{k|k} = Q_k + \sum_{i=0}^{2n} (W_i [\chi_i - \hat{w}_{k|k-1}] [\chi_i - \hat{w}_{k|k-1}]^T) \quad (13)$$

利用 DUKF 算法对输入输出数据进行处理,对神经网络进行离线训练,然后在线预测新的测量值。

对于 DUKF 算法,WNN 的输入信号分量的输出数据为归一化数据。因为对这些分量的预测都是基于小波理论来实现的,满足正交特性。将各个分量的预测方差估计进行叠加后为原始信号的预测方差。

## 3 仿真算例与分析

以某型 EMA 的物理模型为对象,对本文所提出的基于 DWNN 的故障诊断方法进行了验证。选取 EMA 的工作电流  $i_a$  和电机转速  $\omega_m$  组成传感器测量向量  $z$ ;选取 EMA 的供电电压  $u_a$  和负载转矩  $T_L$  组成输入向量  $u$ ;选取电枢绕组电感  $L_a$  和电阻  $R_a$ 、机电伺服系统的惯性  $J_m$  和阻尼  $B_m$  组成故障特征向量  $\vartheta$ 。然后通过询问相关领域的专家或查找相关研究文献将选取的故障特征参数与相应的故障模式关联起来,建立如表 1 所示的故障特征表。

针对 EMA 电枢绕组匝间短路故障和传动机构过度磨损等典型部件的渐变故障,进行了故障诊断实例的仿真。分别通过 ARMA 和 DWNN 两种方法来检测故障特征向量  $\vartheta$  中相应特征参数的变化,从而判断正在发生的故障。

表 1 故障特征

Table 1 Fault features

故障模式	故障特征参数变化			
	$B_m$	$J_m$	$R_a$	$L_a$
传动机构过度磨损	+10%			
控制面损失		-10%		
电枢绕组匝间短路			-10%	-10%

注:“+”表示特征参数增大;“-”表示特征参数减小。

### 3.1 绕组匝间短路故障诊断

假设 EMA 电枢绕组匝间短路开始发生在第 100 s,电枢绕组匝间短路发生后会以指数函数的方式演变,在短期内造成电枢绕组更多的匝数短路。为了模拟电枢绕组匝间短路故障,这里通过在模型参数中加入与时间相关的指数权值来植入故障,加速了故障的演变过程。需要说明的是,

EMA 中的渐变性故障的发生通常是一个缓慢时变过程。

同时假设电枢绕组的电感值和电阻值与绕组的匝数是正相关的。即绕组匝数减少,电感值和电阻值会相应的减小。

分别应用 ARMA 和 DWNN 方法进行故障检测,诊断结果如图 3 所示。图 3 中的点划线为 EMA 的电枢绕组电感  $L_a$  和电阻  $R_a$  参数在正常情况下的不确定性边界,假设这些参数的摄动范围为  $\pm 5\%$ 。图 3 中的粗实线反映了使用 DWNN 方法对参数  $L_a$  和  $R_a$  的估计结果;图中的细实线反映了使用 ARMA 方法对参数  $L_a$  和  $R_a$  的估计结果。

从图 3 中可知,使用 DWNN 方法对参数  $L_a$  和  $R_a$  估计,获得的估计结果更平滑,这是因为采用小波分解算法(本文采用了 3 级小波分解算法),将被辨识信号分解成多个高频分量和低频分量,低频分量称为“近似”分量,反映信号总体趋势,高频信号则反映信号的逐次逼近过程,通过神经网络的学习,在参数估计时强调了低频分量的作用,而减弱了高频分量对估计结果的影响。考虑到本文主要针对 EMA 中的渐变性故障进行诊断,可以忽略高频分量的影响,以能够准确地获得信号的变化趋势。相对而言,使用 ARMA 方法对参数  $L_a$  和  $R_a$  估计,若要去掉信号中高频分量的影响,必须人为增加滤波算法;若滤波算法的使用不当,反而会影响辨识的准确性。本文使用 ARMA 方法时,对测量信号进行了高频滤波。

第 100 s 显现出来,即仿真过程中在第 100 s 对传动机构磨损故障特征参数  $B_m$  做了处理,改变了该故障特征参数的指数权值。

分别应用 ARMA 和 DWNN 方法进行故障检测,诊断结果如图 4 所示。图 4 中的点划线为 EMA 的传动系统阻尼  $B_m$  参数在正常情况下的不确定性边界,假设这些参数的摄动范围为  $\pm 5\%$ 。图 4 中的粗实线反映了使用 DWNN 方法对参数  $B_m$  的估计结果;图中的细实线反映了使用 ARMA 方法对参数  $B_m$  的估计结果。

从图 4 中可知,使用 DWNN 方法对参数  $B_m$  估计,获得的结果更稳定。相对而言,使用 ARMA 方法对参数  $B_m$  估计具有一定的不确定性。

根据表 1 给定的设定值,EMA 故障特征参数的变化大于 10% 时,可以认定 EMA 中相关部件发生了故障。若能提前发现 EMA 故障特征参数的变化大于 5%,就可以根据故障特征参数的变化趋势来预测可能发生的故障,在 EMA 故障发生之前对故障进行预诊断。结合图 3 和图 4 可以发现,使用 DWNN 方法对故障特征参数进行跟踪可以准确地判断故障特征参数的变化趋势,实现对 EMA 渐变性故障的诊断。

另外,由于使用了 DUKF 来完成训练故障特征参数,诊断匝间绕组短路故障,只需要 EMA 的电枢电压  $v_a$  和电枢电流  $i_a$  来训练故障特征参数电枢电感  $L_a$  和电枢电阻  $R_a$ ;诊断传动机构磨损故障,只需要 EMA 的转速  $\omega_m$  和负载转矩  $T_L$  来训练故障特征参数阻尼  $B_m$ 。因此,本文方法可以有效诊断同时存在的多渐变性故障。

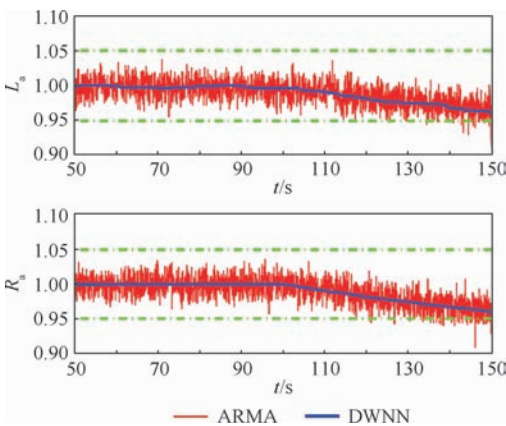


图 3 绕组匝间短路故障特征参数变化曲线

Fig. 3 Fault characteristic parameter variation curves of interturn short circuit of winding

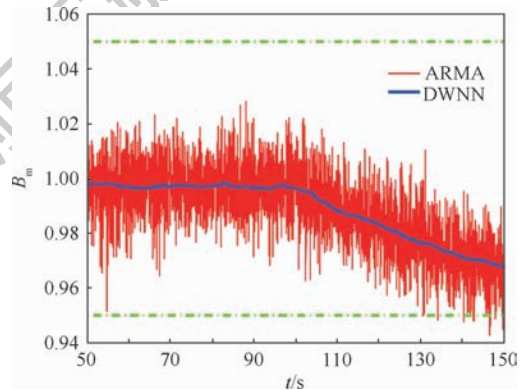


图 4 传动机构磨损故障特征参数变化曲线

Fig. 4 Fault characteristic parameter variation curves of wear of transmission device

### 3.2 传动机构磨损故障诊断

EMA 传动机构磨损是一个缓慢时变的过程,本文通过改变故障特征参数的指数权值,加速了该故障的演变过程。假设 EMA 传动机构磨损在

## 4 结 论

1) 针对 EMA 中的渐变性故障,使用 DWNN 方法可以利用小波分解算法去除信号中高频分量

的影响,获得更为平滑的故障特征参数估计值。

2) 使用 DWNN 方法可以准确地判断故障特征参数的变化趋势,从而实现对 EMA 中匝间绕组短路和传动机构磨损等渐变故障的预诊断。

本文提出的 DWNN 方法能处理类似于 EMA 的机电设备中各种类型的渐变故障,但是对于 EMA 的突变性故障,仍需要优化故障特征参数的估计过程,来减小故障特征参数突变后的检测时间。

### 参考文献 (References)

- [1] BALABAN E, SAXENA A, NARASIMHAN S, et al. Experimental validation of a prognostic health management system for electro-mechanical actuators; AIAA-2011-1518 [R]. Reston, AIAA, 2011.
- [2] QIAN Y, YAN R. Remaining useful life prediction of rolling bearings using an enhanced particle filter[J]. IEEE Transactions on Instrumentation & Measurement, 2015, 64(10): 2696-2707.
- [3] BALABAN E, SAXENA A, NARASIMHAN S, et al. Prognostic health-management system development for electromechanical actuators[J]. Journal of Aerospace Information Systems, 2015, 12(3): 329-344.
- [4] BYINGTON C, WATSON M, EDWARDS D, et al. A model-based approach to prognostics and health management for flight control actuators [C] // IEEE Aerospace Conference. Piscataway, NJ: IEEE Press, 2004: 3551-3562.
- [5] NARASIMHAN S, ROYCHOUDHURY I, BALABAN E, et al. Combining model-based and feature-driven diagnosis approaches—a case study on electromechanical actuators [C] // Proceedings of 21st International Workshop on Principles of Diagnosis. Washington, D. C. : NASA, 2010: 1-8.
- [6] DIXON R, PIKE A W. Application of condition monitoring to an electromechanical actuator: A parameter estimation based approach[J]. Computing and Control Engineering Journal, 2002, 13(2): 71-81.
- [7] DAUBECHIES I. The wavelet transform, time-frequency localization and signal analysis [J]. IEEE Transactions on Information Theory, 2015, 36(5): 961-1005.
- [8] BADIHI H, RAD J S, ZHANG Y M, et al. Data-driven model-based fault diagnosis in a wind turbine with actuator faults [C] // International Mechanical Engineering Congress and Exposition. New York: ASME, 2014: 1-6.
- [9] NAPOLITANO M R, AN Y, SEANOR B A. A fault tolerant flight control system for sensor and actuator failures using neural networks[J]. Aircraft Design, 2000, 3(2): 103-128.
- [10] YANG K, SHAN G, ZHAO L. Application of wavelet packet analysis and probabilistic neural networks in fault diagnosis [C] // Proceedings of 2006 6th World Congress on Intelligent Control and Automation. Piscataway, NJ: IEEE Press, 2006: 4378-4381.
- [11] JING J, LIU H, LU C. Fault diagnosis of electro-mechanical actuator based on WPD-STFT time-frequency entropy and PNN complexity[J]. Vibro Engineering PROCEDIA, 2017, 14(2): 130-135.
- [12] GUO P, SUN L, LI H, et al. Vibration fault diagnosis of hydro-power unit based on multi-fractal spectrum and improved BP neural network[J]. Journal of Hydroelectric Engineering, 2014, 33(3): 299-305.
- [13] WANG J, LI T, XIE R, et al. Fault feature extraction for multiple electrical faults of aviation electro-mechanical actuator based on symbolic dynamics entropy [C] // Proceedings of IEEE International Conference on Signal Processing Communications and Computing. Piscataway, NJ: IEEE Press, 2015: 421-426.
- [14] DELGADO M, GARCIA A, URRESTY J C, et al. Evaluation of machine learning techniques for electro-mechanical system diagnosis [C] // Proceedings of European Conference on Power Electronics and Applications. Piscataway, NJ: IEEE Press, 2011: 1-10.
- [15] MARZAT J, PIET-LAHANIER H, DAMONGEOT F, et al. Model based fault diagnosis for aerospace systems: A survey [J]. Journal of Aerospace Engineering, 2012, 226(10): 1329-1360.
- [16] ZHANG Q, BENVENISTE A. Wavelet networks [J]. IEEE Transactions on Neural Networks, 1992, 3(6): 889-898.
- [17] BILLINGS S, WEI H L. A new class of wavelet networks for nonlinear system identification [J]. IEEE Transactions on Neural Networks, 2005, 16(4): 862-874.
- [18] BECERIKLI Y. On Three intelligent systems; Dynamic neural, fuzzy, and wavelet networks for training trajectory [J]. Neural Computation and Applications, 2004, 13(4): 339-351.
- [19] MOON C, KWON Y. Sensorless speed control of permanent magnet synchronous motor by unscented kalman filter using various scaling parameters [J]. Journal of Electrical Engineering & Technology, 2016, 11(2): 347-352.

#### 作者简介:

王剑 男, 博士研究生。主要研究方向: 飞行器机电系统故障诊断技术。

王新民 男, 教授, 博士生导师。主要研究方向: 飞行器控制技术、飞行器机电系统故障诊断技术。

谢蓉 女, 博士, 副教授。主要研究方向: 飞行器控制技术、飞行器健康管理技术。

曹宇燕 女, 博士研究生。主要研究方向: 飞行器机电系统故障诊断技术。

李婷 女, 博士研究生。主要研究方向: 飞行器机电系统故障诊断技术。

## Gradual fault diagnosis for electromechanical actuator based on DWNN

WANG Jian, WANG Xinmin, XIE Rong\*, CAO Yuyan, LI Ting

(College of Automation, Northwestern Polytechnical University, Xi'an 710072, China)

**Abstract:** It is difficult to accurately predict the gradual failure of electromechanical actuator (EMA) in an aircraft flight control system. The flight safety of an aircraft will be affected by these faults if they are not detected in early stage. A fault diagnosis method based on dynamic wavelet neural network (DWNN) is proposed to diagnose the gradual fault of EMA. This method trains the DWNN fault diagnosis model in offline step by using EMA's operation data of gradual faults, such as interturn short circuit of armature winding, screw and ball wear of transmission device, and then the trained DWNN model is used to diagnose the gradual faults of EMA online. The innovations of the research are as follows: First, the influence of the high-frequency components in the sensor measurement signals is removed using wavelet decomposition algorithm in the DWNN model; Second, the information input in the past and the information predicted in the past are integrated using the memory ability of the feedback neural network, so the accuracy of EMA gradual fault diagnosis is improved. The simulation results obtained from tests on a specific EMA show that the proposed DWNN method can accurately diagnose the gradual fault of EMA components.

**Keywords:** electromechanical actuator (EMA); gradual; fault; diagnosis; dynamic wavelet neural network (DWNN)

**Received:** 2018-12-29; **Accepted:** 2019-02-02; **Published online:** 2019-02-21 15:22

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20190221.0934.003.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190221.0934.003.html)

**Foundation item:** National Natural Science Foundation of China (61703341)

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

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

DOI: 10.13700/j.bh.1001-5965.2018.0787

# 冲突证据决策新方法及应用

赵静, 关欣\*, 刘海桥

(海军航空大学 航空作战勤务学院, 烟台 264001)



**摘 要:** 冲突证据决策方法研究是证据理论重要研究课题。鉴于现有的证据理论改进方法在冲突证据决策过程中存在计算量较大, 归一化过程不合理, 证据组合效果不理想等一系列问题, 提出基于二次组合的冲突证据决策方法。首先, 提出新的基于二次组合的冲突证据决策方法的流程图; 然后, 提出新的乘性归一化规则, 并对新的乘性归一化规则进行算例分析, 验证其合理性; 最后, 分析现有冲突度量函数的不足, 并提出新的冲突度量函数, 并分析冲突度量函数的合理性。通过算例分析, 并与现有证据组合规则的比较表明, 所提方法不仅计算量得以改善, 组合结果也得到提升。

**关键词:** 冲突证据组合; 冲突度量函数; 乘性归一化规则; 证据理论; 二次组合

**中图分类号:** TN95; C934

**文献标识码:** A

**文章编号:** 1001-5965(2019)09-1838-10

DS 证据理论是由 Dempster<sup>[1]</sup> 和 Shafer<sup>[2]</sup> 提出并推广, 该理论能够很好地对不确定信息进行表征、融合与决策, 并在信息融合、风险评估、模式识别等方面<sup>[3-6]</sup> 得到广泛应用。但是, 在工程应用中, 由于信源的测量精度, 外界环境的干扰, 需要处理的信息往往是冲突的。而 DS 证据理论中所给出的 Dempster 组合规则<sup>[7-10]</sup>, 不能有效处理高冲突信息<sup>[11-12]</sup>, 比如 Zadeh<sup>[13]</sup> 给出经典的 0 信任悖论算例。为了解决这一问题, 相关学者提出了大量的改进算法<sup>[9, 14-16]</sup>。针对冲突证据组合规则的研究, 证据理论的改进方法主要分为两大类。

一类学者认为, Dempster 组合规则之所以不能很好地处理冲突证据, 是因为组合规则的建立有问题。因此, 这一类学者对组合规则的公式进行大量改进, 提出许多新的组合规则<sup>[15-25]</sup>。Dempster 组合公式的改进主要分为 3 类: 第 1 类是承认 DS 证据理论的乘性法则, 研究如何分配冲突量  $k$ <sup>[17-23]</sup>, 比如 Smets 提出新的组合规则, 将

冲突量分配给了空集, 避免了证据理论的归一化过程<sup>[17-18]</sup>; Yager 提出的新的组合规则, 将冲突量全部分配给了完全不确定集<sup>[19-20]</sup>; PCR1 ~ PCR6<sup>[21-23]</sup> 重点研究如何将冲突量  $k$  合理地分配到各个命题, 其中 PCR2 和 PCR5 应用比较广泛。第 2 类是改变识别框架, 拓展到广义幂集, 在此基础上, 给出新的组合规则, 比如 DSmt 组合规则<sup>[23]</sup>。第 3 类是给出加性组合规则, 比如 Murphy<sup>[24]</sup> 组合规则。

虽然对组合规则的改进, 能够一定程度上解决冲突证据的融合问题, 但是效果不是很理想。目前更多的学者认为, 证据理论之所以不能很好地解决冲突证据, 是因为原始数据受到传感器或者环境的噪声干扰。因此, 解决冲突证据组合问题的关键在于对证据进行修正。修正证据源的研究重点在于如何获取证据的折扣权重系数<sup>[17, 25-27]</sup>。证据折扣权重系数确定方法分 2 种。

1) 基于信息熵, 通过信息的无序性, 表征证

收稿日期: 2019-01-02; 录用日期: 2019-03-22; 网络出版时间: 2019-05-08 13:38

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190506.1025.004.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190506.1025.004.html)

基金项目: 国家自然科学基金 (91538201, 61671463, 61571454); 国防科技卓越青年科学基金 (2017-JCJQ-2Q-003); 泰山学者工程 (ts201712072)

\* 通信作者. E-mail: 597268914@qq.com

引用格式: 赵静, 关欣, 刘海桥. 冲突证据决策新方法及应用[J]. 北京航空航天大学学报, 2019, 45(9): 1838-1847.

ZHAO J, GUAN X, LIU H Q. A new conflict evidence decision method and its application [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1838-1847 (in Chinese).

据的不确定性。Höhle<sup>[28]</sup>在似然函数基础上提出 Confusion 度量; Yager<sup>[29]</sup>在似真度函数基础上提出 Dissonance 度量; Klir 等<sup>[30-31]</sup>基于命题的模,给出 Discord 度量和 Strife 度量; Harmanec 和 Klir<sup>[32]</sup>在概率转换的基础上,提出 Aggregated 不确定性度量; Jousselme 等<sup>[33]</sup>基于 pignistic 概率转换基础上,提出 Ambiguity 度量; Deng<sup>[34]</sup>在 Shannon 熵的基础上提出 Deng 熵,用来衡量证据的不确定性。上述给出的多种基于信息熵的不确定性度量方法,通过算例分析,都能一定程度上表征信息的不确定性。在实际应用中,当传感器获取信源的信息时,信息在产生、传输过程中,不存在不确定性,而在信息的接收方面,随着传感器的可靠性,不同传感器获取的信源信息存在着不确定性。这时,便可以利用上述提出的多种基于证据本身的不确定性,来表征证据的可靠性,从而对证据进行修正,达到有效的组合结果。本文对这一类的冲突证据组合方法不做深入的研究,根据实际应用的需求,重点分析冲突证据不确定衡量的第2类方法,即如何衡量证据之间的冲突度或者说相似度。

2) 基于距离衡量证据的不确定性<sup>[35-37]</sup>。Jousselme 等<sup>[35]</sup>提出的 Jousselme 距离是应用最为广泛的证据距离; Yu 等<sup>[9]</sup>提出支持概率 (supporting probability) 距离,通过计算证据被支持程度,确定证据的折扣权重; Zhou 等<sup>[38]</sup>对现有的证据距离进行运算,得到新的权重系数; Liu 等<sup>[15]</sup>在概率转换 DSmP 的基础上,提出一种新的证据距离; Yang 和 Han<sup>[39]</sup>利用 Tran&Duckstein 区间距离衡量证据的不确定性,并得到很好的融合效果。

虽然上述文献介绍了很多基于信息熵和距离的冲突证据决策方法,但是解决冲突证据组合的流程基本相同。都是先确定一种衡量证据之间支持程度的指标,并计算两两证据之间的支持程度,得到证据;然后通过支持矩阵,计算出每一条证据的折扣系数,再对证据进行折扣;最后对折扣证据进行组合。由于要计算两两证据之间的支持程度,基于这种流程处理冲突证据组合问题,计算量比较大,且计算支持向量的归一化过程不是很合理,证据组合效果不理想。

因此,本文提出基于二次组合的冲突证据决策方法流程。首先,计算证据组的 PCR6 组合规则;然后,计算每条证据距离组合后证据的距离,从而确定折扣系数,并对证据组进行折扣处理;最后,再一次利用 PCR6 组合规则,对证据进行组合。

## 1 理论基础

记  $X$  为识别对象,  $U$  为  $X$  可能取值的集合, 并且  $U$  中的所有元素互不相容, 则称  $U$  为一识别框架。

假设  $U$  为对象  $X$  的识别框架, 当函数  $m: U \rightarrow [0, 1]$  满足条件:

$$\begin{cases} m(\emptyset) = 0 \\ \sum_{A \in U} m(A) = 1 \end{cases} \quad (1)$$

则称函数  $m(A)$  为基本概率赋值 (BPA), 表示证据对命题  $A$  的支持程度。

假设  $U$  为一识别框架,  $m$  为 BPA, 定义

$$\text{Bel}(A) = \sum_{B \subset A} m(B) \quad \forall A \subset U$$

为信任函数, 表示决策可能结果  $A$  所有子集信任之和, 所以对于空集的信任值为  $\text{Bel}(\emptyset) = 0$ , 对于全集  $U$  为  $\text{Bel}(U) = 1$ 。定义

$$\text{Pl}(A) = \sum_{B \cap A \neq \emptyset} m(B)$$

为似真度函数, 从定义不难得到定理,  $\text{Pl}(A) = 1 - \text{Bel}(\bar{A})$ 。信任函数与似真度函数之间满足  $\text{Bel}(A) \leq \text{Pl}(A)$ 。

定义区间  $L = [\text{Bel}(A), \text{Pl}(A)]$  为信任度区间, 区间长度反映出命题  $A$  的不确定度。当信任度区间长度变为 0 时, 即  $\text{Bel}(A) = \text{Pl}(A)$  时, 证据理论便退化为概率论。

### 1) Dempster 组合规则

假设  $m_1$  和  $m_2$  是识别框架  $U$  上的 BPA, 则

$$m_{12}(C) = \begin{cases} \frac{\sum_{i,j, A_i \cap B_j = C} m_1(A_i) m_2(B_j)}{1 - k} & C \neq \emptyset \\ 0 & C = \emptyset \end{cases} \quad (2)$$

式中:  $k$  为冲突系数, 表达式为

$$k = \sum_{i,j, A_i \cap B_j = \emptyset} m_1(A_i) m_2(B_j) < 1$$

### 2) PCR6 组合规则

假设  $m_1$  和  $m_2$  是识别框架  $U$  上的 BPA, 则

$$m_{\text{PCR6}}(X) = \sum_{\substack{X_1, X_2 \in 2^U \\ X_1 \cap X_2 = X}} m_1(X_1) m_2(X_2) + \sum_{\substack{Y \in 2^U \setminus \{X\} \\ X \cap Y = \emptyset}} \left[ \frac{m_1^2(X) m_2(Y)}{m_1(X) + m_2(Y)} + \frac{m_2^2(X) m_1(Y)}{m_1(X) + m_2(Y)} \right] \quad (3)$$

## 2 基于二次组合的冲突证据决策

### 2.1 基于冲突度量的折扣证据不确定推理模型

Dempster 组合规则在处理冲突证据时, 会产

生悖论现象,对于这一问题,相关学者给出多种改进方法,主要分为两类:第1类是对原有的DS证据理论的组合公式进行改进;第2类是对原始证据进行修正。目前大多数学者都接受第2类改进方式,认为证据的冲突根本上是来源于获取数据的冲突、不精确。

而对于修正证据源的冲突证据融合改进方式,已经形成一套完整的算法模型(见图1),下面对这一模型进行构建分析。

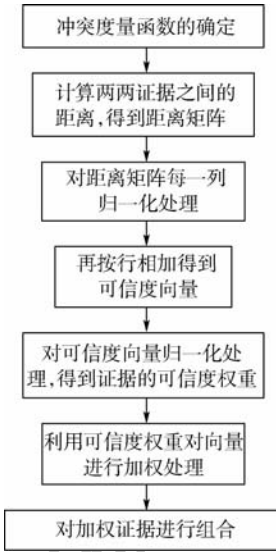


图1 折扣证据组合模型

Fig.1 Discount evidence combination model

1) 确定冲突度量函数。这是关键的一步,目前的相关修正数据源的改进算法,都是基于对不同的冲突度量函数的改进。冲突度量函数不仅需要满足物理意义,即与证据之间的距离成正相关;还需要满足数学条件,即满足范数的几条基本条件。

假设3组映射在相同命题空间 $U$ 的BPA: $m_1(\cdot)$ 、 $m_2(\cdot)$ 和 $m_3(\cdot)$ ,则存在一个映射 $CM(\cdot):U \times U \rightarrow [0,1]$ 满足下列4个基本条件,被称为冲突度量函数。

① 对称性。

$$\forall m_1(\cdot), m_2(\cdot), CM(m_1, m_2) = CM(m_2, m_1)。$$

② 正定性。

$\forall m_1(\cdot), m_2(\cdot) \in U$ , 当 $m_1 \neq m_2$ 时,  $CM(m_1, m_2) \geq 0$ ; 当 $m_1 = m_2$ 时,  $CM(m_1, m_2) = 0$ 。

③ 齐次性。

$$CM(r_1 m_1, r_2 m_2) = r_1 r_2 CM(m_1, m_2), r \in \mathbf{R}。$$

④ 三角不等式。

$$CM(m_1, m_2) \leq CM(m_1, m_3) + CM(m_2, m_3)。$$

现有的冲突度量函数大体分为两类:一类是基于信息熵的冲突度量函数,一类是基于证据之间距离的冲突度量函数。

2) 计算两两证据之间的距离,得到距离矩阵 $D$ ,元素 $d_{ij}$ 表示第 $i$ 条矩阵与第 $j$ 条证据之间的距离。距离矩阵是一个对称方阵,并且对角线的元素为0。

3) 对距离矩阵的每一列进行归一化处理,得到的矩阵第 $i$ 列的一条向量,表示第 $i$ 条矩阵对每一条证据的支持程度,从而便于后续的运算。

4) 将归一化后矩阵的每一行相加,得到一条向量,向量的第 $s$ 个元素表示第 $s$ 条证据被其他证据支持的总程度。

5) 对步骤4)得到的向量进行归一化处理,得到归一化可信度权重。

6) 利用可信度权重对证据进行折扣计算。常用的折扣公式为

$$\begin{cases} m'_i(A_j) = \alpha_i m_i(A_j) & A_j \neq U \\ m'_i(U) = 1 - \sum_j m'_i(A_j) \end{cases} \quad (4)$$

式中: $\alpha_i$ 为折扣系数。

7) 对折扣后的证据进行组合,得到最后的决策证据。

## 2.2 基于二次组合的冲突证据决策模型

通过分析折扣证据组合模型,可以发现,修正证据源的冲突证据组合方法,需要计算两两证据之间的距离,对于 $t$ 证据,需要计算距离 $\frac{t(t-1)}{2}$ 次,而每一次的距离计算,根据不同的距离公式,计算复杂度不一样,以经典的Jousselme距离为例,两两证据之间的距离公式为

$$d_j(m_1, m_2) = \sqrt{\frac{1}{2} (m_1 - m_2)^T D (m_1 - m_2)} \quad (5)$$

式中:矩阵 $D$ 中的元素 $d_{ij} = \frac{|A_i \cap B_j|}{|A_i \cup B_j|}$ ;  $m_1$ 和 $m_2$ 为任意的2条证据。

可以看出,修正证据源的冲突证据组合方法,计算量较大,不适合多条证据之间的组合(第5节将对组合规则的计算量进行定量研究),所以接下来给出一种基于二次组合的冲突证据决策模型,如图2所示。

1) 利用PCR6组合规则对原有证据进行组合,得到初步的组合结果。PCR6组合被广泛地应用在冲突证据的组合规则上,并得到较好的组合效果。组合后得到的证据 $m'_i$ 能够一定程度上反映证据的总体水平。

2) 确定冲突度量函数,与2.1节描述的一样,度量函数需要满足范数的基本条件,这里不再赘述。

3) 计算每一条证据到组合后证据之间的距离,直接计算得到可信度向量。

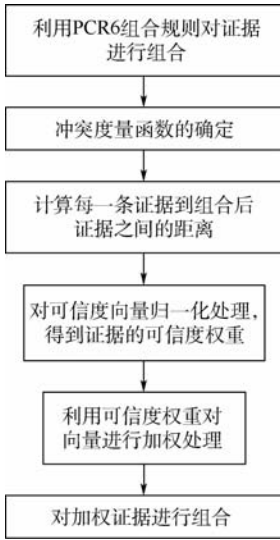


图2 基于二次组合的冲突证据决策模型

Fig.2 Conflict evidence decision model based on quadratic combination

4) 对证据的可信度向量进行归一化。在常规的修正证据源的冲突证据中,证据的归一化是每一项元素除以元素的和。但是从式(2)和式(3)可以看出,证据的组合规则中,证据之间的元素是进行乘法运算,即  $m_1(A_i) m_2(A_j) \cdots m_n(A_k)$ ,所以求和形式的归一化不满足计算的要求。归一化应该满足:

$$\prod_{i=1}^n m'_i(A_j) = \prod_{i=1}^n m_i(A_j) \quad (6)$$

为此,本文提出一种新的乘性归一化计算公式,即

$$a'_i = \frac{a_i}{\sqrt[n]{\prod_{i=1}^n a_i}} \quad (7)$$

归一化得到的新的权重系数,满足  $\prod_{i=1}^n a'_i = 1$ ,满足后续的乘性运算需要。

5) 利用可信度对向量进行加权处理,采用折扣的方法对证据进行折扣,由于归一化公式的改变,折扣公式需要进行修正,得

$$m'_i(A_j) = \begin{cases} \alpha'_i m_i(A_j) & |A_j| = 1 \\ \frac{1}{\alpha'_i} m_i(A_j) & |A_j| \geq 2 \end{cases} \quad (8)$$

式中:  $|A_j|$  表示元素的势的大小,  $|A_j| = 1$  表示单元素集。

6) 利用 PCR6 证据组合规则对证据进行组合。

### 2.3 乘性组合规则的有效性分析

2.2 节中提出基于乘性法则的归一化公式,如式(6)和式(7)所示,下面通过算例,验证乘性归一化法则在冲突证据理论决策中的合理性。

**算例1** 假设4类不同传感器对目标A进行识别,识别框架为  $U = \{A, B, C\}$ ,得到 BPA,  $m_1(A) = 0.6, m_1(B) = 0.2, m_1(AB) = 0.1, m_1(C) = 0.1; m_2(A) = 0.6, m_2(B) = 0.1, m_2(AB) = 0.1, m_2(C) = 0.1, m_2(BC) = 0.1; m_3(A) = 0.5, m_3(B) = 0.1, m_3(AB) = 0.1, m_3(C) = 0.1, m_3(AC) = 0.1, m_3(BC) = 0.1; m_4(A) = 0.1, m_4(B) = 0.1, m_4(AB) = 0.1, m_4(C) = 0.5, m_4(AC) = 0.1, m_4(BC) = 0.1$ 。

从 BPA 可以看出,前3条证据给出的A信度较高,而第4条证据出现了冲突,所以第4条证据的可信度较低,前3条证据的可信度较高,因此给出4条证据的支持度向量:

$$w = [0.7 \quad 0.6 \quad 0.65 \quad 0.1]$$

关于支持度向量的确定,后文第4节将根据冲突度量函数确定,这里只需论证乘性归一化法则的有效性。

然后对支持度向量分别计算加性归一化和乘性归一化之后的支持度向量,得

$$\begin{cases} w_+ = [0.3415 \quad 0.2927 \quad 0.3171 \quad 0.0488] \\ w_x = [1.7221 \quad 1.4761 \quad 1.5991 \quad 0.2460] \end{cases}$$

利用式(4)和式(8),分别计算得到折扣后的BPA,然后利用 PCR6 组合规则对证据进行组合,得到结果如表1所示。表1中:PCR6 表示直接通过 PCR6 组合规则组合的结果;PCR6<sub>+</sub> 表示通过加性归一化组合规则,再利用 PCR6 组合规则组合的结果;PCR6<sub>x</sub> 表示通过乘性归一化组合规则,再利用 PCR6 组合规则组合的结果。

通过分析表1中的数据可知,加性 PCR6 组合规则得到的不确定赋值  $m_i(ABC)$  较大,因为式(4)中,加性归一化证据组合将不确定信息都赋给了命题 ABC,并不利于后续的证据的决策;而本文提出的乘性归一化法则的到的组合结果显示

表1 加性与乘性归一化折扣证据组合结果

Table 1 Additive and multiplicative normalized discount evidence combination results

组合规则	$m_i(A)$	$m_i(B)$	$m_i(AB)$	$m_i(C)$	$m_i(AC)$	$m_i(BC)$	$m_i(ABC)$
PCR6	0.5436	0.0939	0.0403	0.2544	0.0224	0.0454	0
PCR6 <sub>+</sub>	0.5386	0.0938	0.0314	0.0588	0.0095	0.0205	0.2475
PCR6 <sub>x</sub>	0.6379	0.0751	0.0622	0.0523	0.0641	0.1083	0

$m(ABC) = 0$ , 并且赋予命题  $A$  较大的信度, 符合实际情况, 得到较好的组合结果, 因为乘性归一化公式与后续的 PCR6 组合规则中的乘法规则相呼应, 这样的归一化更有意义。

### 2.4 预处理合理性分析

2.2 节中给出基于二次组合的冲突证据决策模型, 首先给出 PCR6 组合结果为基准参考, 但是后又需要对 PCR6 组合规则的效果进行讨论, 从逻辑上就存在了矛盾。

下面, 本文将对预处理过程中, 给出 PCR6 组合规则结果为基准证据的合理性进行证明。

算法要对多条证据进行修正, 就需要考虑每条证据的冲突度程度, 冲突度越大的证据, 赋予它较小的权重。而冲突度的度量, 已经在 2.1 节中介绍, 要求通过求解两两证据之间的距离来衡量证据的冲突度, 所以需要找出一条证据, 这条证据假设为所有证据融合后的理想结果  $m_{per}$ , 然后求解所有的证据与理想证据之间的距离, 从而判定每条证据的冲突度大小。

但是, 理想结果  $m_{per}$  没法求解得到, 因为算法要求就是理想结果  $m_{per}$ , 这时算法便通过 PCR6 来近似代替理想的融合结果。

问题是, PCR6 能不能代表理想结果  $m_{per}$  作为证据的基准? 本文可以通过误差分析, 来判断 PCR6 作为基准证据的合理性。

就以表 1 中的 PCR6 组合结果为研究对象, 假设理想的证据  $m_{per}$  与 PCR6 组合规则存在一定的误差, 并满足以下的误差规则(可以通过这一情况, 来分析误差, 判断 PCR6 作为基准证据的合理性):

$$\begin{cases} m_{per}(A) = PCR6(A) + \alpha \\ m_{per}(A_i) = PCR6(A_i) - \frac{\alpha}{4} \quad |A_i| > 1 \end{cases}$$

通过改变  $\alpha$  值, 得到如图 3 所示的误差结果。

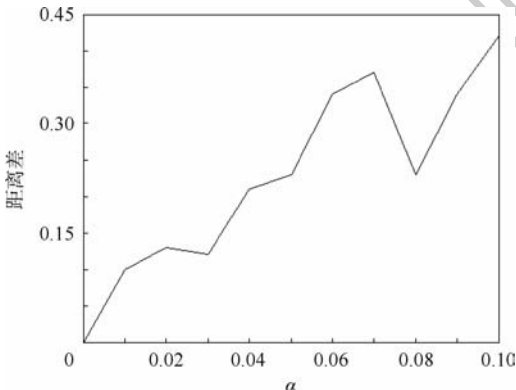


图 3 预处理误差

Fig. 3 Preprocessing error

由图 3 可以看出, 当  $\alpha$  值在 0 ~ 0.1 之间变化时, PCR6 和证据组之间的平均距离  $d_{aPCR6}$ , 与理想的证据  $m_{per}$  和证据组之间的平均距离  $d_{aper}$  之间的差  $d_{aPCR6} - d_{aper}$  在 0.05 以内, 误差完全可以接受。

因此, 本文在预处理中给出 PCR6 为基准证据在一定误差范围内是合理的。

## 3 冲突度量函数的确定

### 3.1 经典的 Josselme 距离冲突度量函数

目前常用的冲突度量函数有两类: 一类是基于信息熵的冲突度量函数; 另一类是基于距离的冲突度量函数, 这里的距离是一种广义的概念, 包括向量距离、区间距离、角度等。信息熵是一种衡量证据离散程度的参量, 一定程度上反映了证据的决策能力, 根据实际应用的需求, 本文不对这一类型的证据距离进行详细分析。本文将基于证据之间的距离确定冲突度量函数。

第 2 节已经阐明冲突度量函数需要满足的几点条件。目前最常用的冲突度量函数便是基于 Josselme 距离的冲突度量函数。Josselme 距离考虑到命题的势, 能够有效地衡量 BPA 之间的相似程度, 下面给出 Josselme 距离的定义。

假设  $m_1$  和  $m_2$  是相同识别框架下的 2 个 BPAs, 则定义  $m_1$  和  $m_2$  之间的 Josselme 距离为

$$d(m_1, m_2) = \sqrt{\frac{1}{2} (m_1 - m_2)^T D (m_1 - m_2)} \quad (9)$$

式中:  $D$  为一个  $2^{|U|} \times 2^{|U|}$  的矩阵, 其中的每一个元素为

$$D(A_i, A_j) = \frac{|A_i \cap A_j|}{|A_i \cup A_j|}$$

Josselme 距离能一定程度上反映 BPAs 之间的距离, 在数学上是严谨的, 并得到广泛的应用。

Josselme 距离已经在多个领域得到了广泛的支持, 但是考虑到, Josselme 距离需要求解矩阵  $D$ , 运算量较大(计算量问题, 本文第 5 节会进行进一步深入分析)。鉴于此, 本文将引入向量角的概念来衡量证据之间的距离, 向量之间的夹角余弦可以表示为

$$d_\theta = \cos \theta = \frac{x_1 \cdot x_2}{|x_1| |x_2|} \quad (10)$$

由式 (10) 可以看出, 证据中每一个命题的赋值满足  $0 \leq m_i(A_j) \leq 1$ , 所以, 证据的积满足  $x_1 \cdot x_2 \geq 0$ , 即  $d_\theta \geq 0$ 。

但是证据是基于识别框架  $U$  的  $2^{|U|}$  维向量, 单纯地将证据视为向量, 求解证据之间的夹角, 而丢弃证据多元素命题与单元元素命题之间的差异,

赋予不同命题之间相同的权重,这是不合理的,比如,对于证据组,  $m_1(A) = 0.1, m_1(B) = 0.5, m_1(C) = 0.1, m_1(AC) = 0.15, m_1(BC) = 0.15;$   $m_2(A) = 0.5, m_2(B) = 0.1, m_2(C) = 0.1, m_2(AC) = 0.15, m_2(BC) = 0.15;$   $m_3(A) = 0.15, m_3(B) = 0.15, m_3(C) = 0.1, m_3(AC) = 0.1, m_3(BC) = 0.5;$   $m_4(A) = 0.15, m_4(B) = 0.15, m_4(C) = 0.1, m_4(AC) = 0.5, m_4(BC) = 0.1。$

计算可以得到  $m_1, m_2$  与  $m_3, m_4$  之间的向量角余弦为  $d_\theta(m_1, m_2) = d_\theta(m_3, m_4) = 0.508$ , 即  $m_1, m_2$  与  $m_3, m_4$  之间的向量角余弦相等。但是通过分析可以发现,证据  $m_1, m_3$  之间,向量的大小没变,只是将命题  $A, B$  的信任度与命题  $AC, BC$  之间的信任度进行了调换,即向量  $m_1$  经过旋转得到向量  $m_3$ , 同样  $m_2$  也是经过旋转得到  $m_4$ , 且两者的旋转方向与角度相等,都是只改变了命题  $A, B$  与命题  $AC, BC$  之间的信任度,所以,从向量的角度分析,两者的向量角余弦相等。但因为证据的命题的势不相同,所以命题的权重应该不一样,证据  $m_1, m_2$  之间命题  $A, B$  冲突,其余的命题不冲突;而证据  $m_3, m_4$  之间命题  $AC, BC$  冲突,其余的命题不冲突,并且两对的冲突大小相同,但是由于命题  $A, B$  的可信度更高,而命题  $AC, BC$  的信度相对模糊,所以,证据  $m_1, m_2$  之间的距离应该大于证据  $m_3, m_4$  之间的距离,而单纯地通过向量的夹角并不能充分考虑到命题的势的大小,所以并不能很好地衡量证据之间的距离。因此,3.2节将考虑命题的势的大小,给出基于 pignistic 向量角的冲突证据度量函数。

### 3.2 基于 pignistic 向量角的冲突证据度量函数

本节将结合 pignistic 概率转换与向量角的概念,提出基于 pignistic 向量角的冲突证据度量函数,即

$$d_{\text{Bet}-\theta}(m_1, m_2) = \sqrt{1 - \left( \frac{\text{BetP}_1 \cdot \text{BetP}_2}{|\text{BetP}_1| |\text{BetP}_2|} \right)^2} = \frac{|\text{BetP}_1 \times \text{BetP}_2|}{|\text{BetP}_1| |\text{BetP}_2|} \quad (11)$$

式中:  $\text{BetP}_i$  为 pignistic 概率转换后的概率向量。向量中每一个元素为概率转换结果  $\text{BetP}_k(A_i)$  为

$$\text{BetP}_k(A_i) = \sum_j m_k(A_j) \frac{|A_i \cap A_j|}{|A_j|} \quad (12)$$

下面对式(11)的对称性、正定性、齐次性和三角不等式性质进行证明。

**证明** 1) 对称性

$$d_{\text{Bet}-\theta}(m_1, m_2) = \frac{|\text{BetP}_1 \times \text{BetP}_2|}{|\text{BetP}_1| |\text{BetP}_2|} =$$

$$\frac{|\text{BetP}_2 \times \text{BetP}_1|}{|\text{BetP}_2| |\text{BetP}_1|} = d_{\text{Bet}-\theta}(m_2, m_1)$$

成立。

2) 正定性

因为  $m_k(A_i) \geq 0$ , 并且  $\text{BetP}_k(A_i) = \sum_j m_k(A_j) \frac{|A_i \cap A_j|}{|A_j|}$ , 所以有  $\text{BetP}_k(A_i) \geq 0$ ,  $\text{BetP}_1 \times \text{BetP}_2 \geq 0$ 。

又因为  $\frac{|\text{BetP}_1 \times \text{BetP}_2|}{|\text{BetP}_1| |\text{BetP}_2|} \leq 1$ , 所以有  $0 \leq$

$$\frac{|\text{BetP}_1 \times \text{BetP}_2|}{|\text{BetP}_1| |\text{BetP}_2|} \leq 1, \text{ 即 } 0 \leq d_{\text{Bet}-\theta}(m_1, m_2) = \frac{|\text{BetP}_1 \times \text{BetP}_2|}{|\text{BetP}_1| |\text{BetP}_2|} \leq 1。$$

当  $m_1 = m_2$  时,  $\frac{\text{BetP}_1 \cdot \text{BetP}_2}{|\text{BetP}_1| |\text{BetP}_2|} = 1$ , 所以

$$d_{\text{Bet}-\theta}(m_1, m_2) = \frac{\text{BetP}_1 \times \text{BetP}_2}{|\text{BetP}_1| |\text{BetP}_2|} = 0。$$

正定性成立。

3) 齐次性

在此不作证明,因为证据需要满足式(1),所以不存在比较  $k_1 m_1, k_2 m_2$  之间距离的情况。

4) 三角不等式

根据不等式  $\sin \theta_1 \leq \sin \theta_2 + \sin \theta_3$ , 其中,  $\theta_1, \theta_2$  和  $\theta_3$  为空间三向量两两之间的夹角,得

$$\frac{|\text{BetP}_1 \times \text{BetP}_2|}{|\text{BetP}_1| |\text{BetP}_2|} \leq \frac{|\text{BetP}_1 \times \text{BetP}_3|}{|\text{BetP}_1| |\text{BetP}_3|} + \frac{|\text{BetP}_3 \times \text{BetP}_2|}{|\text{BetP}_3| |\text{BetP}_2|} \quad \text{证毕}$$

## 4 基于二次组合的冲突证据决策的计算量分析

本节将重点分析基于二次组合的冲突证据决策的计算量。分析流程图1和图2,可以看出,改进后的冲突证据决策与改进前的区别在于如何获取距离矩阵,而后续对于距离矩阵的处理步骤几乎不变。所以,本文就生成距离矩阵的计算量进行分析。

首先对传统的修正数据源的冲突证据决策方法进行分析。流程图1中的证据距离以经典的 Jousselle 距离为例。运算过程中的交集、并集、求势、加减乘除都记为一次运算。如式(5)所示, Jousselle 距离需要求解矩阵  $D$ , 运算量为

$$A_D = 5 \times \frac{2^n(2^n - 1)}{2} = 5 \times 2^{n-1}(2^n - 1)$$

式中:  $n$  为识别框架  $U$  的势  $|U|$ , 所以两条证据之间的 Jousselle 距离计算量为

$$A_{d_j} = A_D + 2^n + 2^n \cdot 2^n + 2^n(2^n - 1) + 2^n + (2^n - 1) + 2 = 9 \times 2^{2n-1} - 2^{n-1} + 1$$

为了得到距离矩阵,需要计算两两证据之间的距离,假设证据的数目为  $b$ ,则基于 Josselme 距离的冲突证据决策方法需要的计算量大小为

$$A_j = b(b-1)A_{d_j} = b(b-1)(9 \times 2^{2n-1} - 2^{n-1} + 1) \tag{13}$$

下面对基于二次组合的冲突证据决策的计算量进行分析。分析流程图 2 可知,需要对  $b$  条证据进行 PCR6 组合,通过分析式 (3),可以得到,  $b$  条证据的 PCR6 组合计算量为

$$A_{PCR6} = (b-1) \left\{ 2^n \cdot 2^n + 6 \left[ \sum_{i=2}^{n-1} C_n^i (2^{n-i} - 1) \right] + 2^n \cdot 2^n - 1 \right\} = (b-1) \left\{ 2^{2n+1} + 6 \left[ \sum_{i=2}^{n-1} C_n^i (2^{n-i} - 1) \right] - 1 \right\} \tag{14}$$

式中:  $\sum_{i=2}^{n-1} C_n^i (2^{n-i} - 1)$  表示集合  $Y \in 2^U \setminus \{X\}, X \cap Y = \emptyset$  的个数。再利用式 (11) 和式 (12) 求解  $b$  条证据与组合后证据之间的距离,其计算量为

$$A_{d_{Bet-\theta}} = b \left[ 5(2^n - n) + \sum_{i=2}^n iC_n^i \right] + b(2n + n + 4) = b \left[ 5(2^n - n) + \sum_{i=2}^n iC_n^i + (3n + 4) \right] \tag{15}$$

式中:  $\sum_{i=2}^n iC_n^i$  表示在 pignistic 概率转换中加法的次数。总计算量为

$$A_{Bet-\theta} = A_{PCR6} + A_{d_{Bet-\theta}} \tag{16}$$

通过式 (13) ~ 式 (16),对冲突证据决策方法的计算量进行分析,得到结果如图 4 所示。图中,曲线 1~4 表示基于 Josselme 距离的冲突证据决策方法的计算量,4 条曲线由下至上,分别表示当证据为 4~10(步长为 2)时的计算量曲线,曲线 5~8 表示基于二次组合的冲突证据决策方法的计算量,同样 4 条曲线由下至上,分别表示当证据为 4~10(步长为 2)时的计算量曲线,横坐标表示识别框架的势的大小  $|U|$ ,纵坐标表示计算量。但是,由于曲线 5~8 增长比较缓慢,为了便于观察分析,计算量取对数  $\lg A$ ,得到新的计算量增长图如图 5 所示。由图 5 可以看出,基于 Josselme 距离和基于二次组合的冲突证据决策方法的计算量的对数,近似于线性增长,所以两者原计算量都是成指数增长。

结合图 4 和图 5 可知,当证据的个数一定时,基于 Josselme 距离和基于二次组合的冲突证据决策方法的计算量的对数,呈指数增长,并且基于 Josselme 距离的冲突证据决策方法的计算量增长速度比基于二次组合的快将近 5~10 dB;当识别框架的势一定时,计算量随着证据数的增加急剧增加。

综合分析,相比于原有的修正数据源的冲突证据决策方法,本文提出的基于二次组合的冲突证据决策方法的计算量得到了较大的改善。

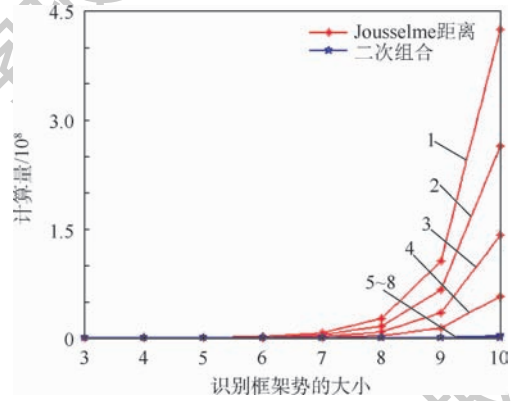


图 4 基于 Josselme 距离和二次组合的冲突证据决策计算量图

Fig. 4 Graph of calculated amount of conflict evidence decision based on Josselme distance and quadratic combination

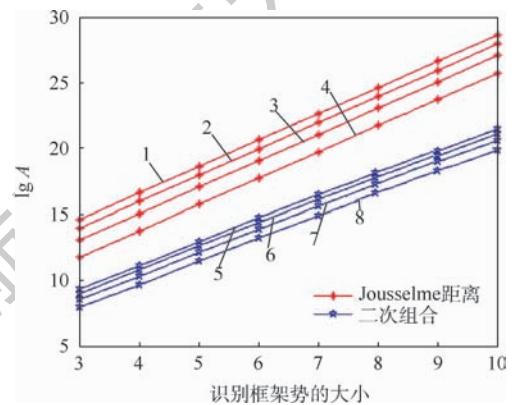


图 5 基于 Josselme 距离和二次组合的冲突证据决策计算量的对数图

Fig. 5 Graph of logarithm of calculated amount of conflict evidence based on Josselme distance and quadratic combination

## 5 基于二次组合的冲突证据决策算例分析

**算例 2** 考虑识别框架  $U = \{A, B, C\}$ , 并给出 6 个信源的 BPAs 赋值,如表 2 所示。

根据流程图 2,首先利用 PCR6 对证据组合,得到组合证据  $m_{PCR6}$  为

$$m_{PCR6}(\cdot) = [0.72 \quad 0.21 \quad 0 \quad 0.03 \quad 0 \quad 0 \quad 0.04]$$

利用式(11)和式(12),计算每一条证据与组合证据  $m_{PCR6}$  之间的距离,得到支持向量为

$$w = [0.88 \quad 0.85 \quad 0.83 \quad 0.55 \quad 0.11 \quad 0.76]$$

可以看出,第 5 条证据  $m_5$  的支持度较低,由表 3 可以看出,第 5 条证据与其他 5 条证据相冲突。然后利用式(7),对权重  $w$  进行归一化,得

$$\bar{w} = [1.59 \quad 1.54 \quad 1.50 \quad 0.99 \quad 0.20 \quad 1.38]$$

再利用式(8),对表 2 中证据组进行折扣计算,得到新的证据组,结果如表 3 所示。

最后利用 PCR6 组合规则对折扣证据进行组合,得到结果如表 4 所示。

由表 2 可以看出,第 4 条证据分别赋予目标 A、B 的 BPA 分别为 0.4、0.3,所以第 4 条证据的不确定较高;而第 5 条证据赋予目标 A、B 的 BPA 分别为 0.1、0.7,即证据 5 比较信任目标 B,而

表 2 证据组 BPAs

Table 2 Evidence group BPAs

$m_i(\cdot)$	BPA						
	A	B	AB	C	AC	BC	ABC
$m_1$	0.7	0.1	0.1	0.1	0	0	0
$m_2$	0.6	0.2	0	0.1	0	0	0.1
$m_3$	0.6	0.05	0	0.05	0	0	0.3
$m_4$	0.4	0.3	0	0.2	0.1	0	0
$m_5$	0.1	0.7	0	0.1	0	0	0.1
$m_6$	0.9	0.05	0	0.05	0	0	0

表 3 折扣证据组 BPAs

Table 3 Discounted evidence group BPAs

$m'_i(\cdot)$	BPA						
	A	B	AB	C	AC	BC	ABC
$m'_1$	0.74	0.11	0.04	0.11	0	0	0
$m'_2$	0.64	0.21	0	0.11	0	0	0.05
$m'_3$	0.72	0.06	0	0.06	0	0	0.16
$m'_4$	0.40	0.30	0	0.20	0.10	0	0
$m'_5$	0.03	0.21	0	0.03	0	0	0.73
$m'_6$	0.90	0.05	0	0.05	0	0	0

表 4 算例不同组合规则计算结果

Table 4 Calculation results of different combination rules of example

组合规则	$m_i(A)$	$m_i(B)$	$m_i(AB)$	$m_i(C)$	$m_i(AC)$	$m_i(BC)$	$m_i(ABC)$
PCR1	0.57	0.22	0.02	0.10	0.02	0	0.08
PCR2	0.57	0.22	0.02	0.10	0.02	0	0.08
PCR3	0.67	0.24	0	0.06	0	0	0.04
PCR5	0.46	0.35	0.01	0.07	0.01	0	0.10
PCR6	0.72	0.21	0	0.03	0	0	0.04
本文方法	0.77	0.07	0	0.02	0	0	0.14

其余的 4 条证据都赋予目标 A 较高的信度,所以证据 5 是一条冲突度较高的证据。通过分析可知,6 条证据的组合结果应该偏向目标 A。由表 4 可知,相比于其他几种常见的组合方法,本文提出的基于二次组合的冲突证据决策方法得到较好的组合结果,有效地解决冲突证据组合问题。

## 6 结论

本文针对现有冲突证据决策方法存在的一系列问题,提出新的基于二次组合的冲突证据决策方法。首先本文通过分析现有基于修正数据源的冲突证据决策方法的一般流程,给出新的基于二次组合的冲突证据决策方法流程图,并给出乘性归一化规则。算例研究表明,由于乘性归一化规则与证据理论组合规则的乘法法则相呼应,能得到更好的组合效果。其次,通过分析现有冲突度量函数的不足,提出基于 pignistic 向量角的冲突证据度量函数,并通过分析得到新的证据度量函数,满足范数理论的对称性、正定性、齐次性和三角不等式性质。然后,对基于二次组合的冲突证据决策的计算量进行分析,通过与原有证据理论方法的比较发现,新方法相比于现有方法,计算量减少将近 5~10 dB,计算量得到较大的改进。最后,通过与不同组合规则的比较发现,所提新方法在处理冲突证据组合时,能得到更好的组合效果。

## 参考文献 (References)

[1] DEMPSTER A P. Upper and lower probabilities induced by a multivalued mapping [J]. Annals of Mathematical Statistics, 1967, 38(2): 325-339.

[2] SHAFER G. A mathematical theory of evidence [M]. Princeton: Princeton University Press, 1976: 85-150.

[3] XU Z, HU C, YANG F, et al. Data-driven inter-turn short circuit fault detection in induction machines [J]. IEEE Access, 2017, 5: 25055-25068.

[4] WANG X, YANG F, WEI H, et al. A risk assessment model of uncertainty system based on set-valued mapping [J]. Journal of Intelligent and Fuzzy Systems, 2016, 31(6): 3155-3162.

[5] ZHANG X, MAHADEVAN S, DENG X. Reliability analysis with linguistic data: An evidential network approach [J]. Reliability Engineering and System Safety, 2017, 162: 111-121.

[6] JIANG W, XIE C, ZHUANG M, et al. Failure mode and effects analysis based on a novel fuzzy evidential method [J]. Applied Soft Computing, 2017, 57: 672-683.

[7] HAN D, DEZERT J, YANG Y. Belief interval-based distance measures in the theory of belief functions [J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2016, 99: 1-18.

[8] QIAN J, GUO X F, DENG Y. A novel method for combining

- conflicting evidences based on information entropy[J]. Applied Intelligence, 2017, 46(4): 876-888.
- [9] YU C, YANG J, YANG D, et al. An improved conflicting evidence combination approach based on a new supporting probability distance[J]. Expert Systems with Applications, 2015, 42(12): 5139-5149.
- [10] SMARANDACHE F, DEZERT J. Modified PCR rules of combination with degrees of intersections[C]// International Conference on Information Fusion. Piscataway, NJ: IEEE Press, 2015: 2100-2107.
- [11] JIANG W, LUO Y, QIN X, et al. An improved method to rank generalized fuzzy numbers with different left heights and right heights[J]. Journal of Intelligent and Fuzzy Systems, 2015, 28(5): 2343-2355.
- [12] JIANG W, WEI B, QIN X, et al. Sensor data fusion based on a new conflict measure[J]. Mathematical Problems in Engineering, 2016, 2016: 5769061.
- [13] ZADEH L A. A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination[J]. AI Magazine, 1986, 7(2): 85-90.
- [14] LIU Z G, DEZERT J, PAN Q, et al. Combination of sources of evidence with different discounting factors based on a new dissimilarity measure [J]. Decision Support Systems, 2011, 52(1): 133-141.
- [15] MARTIN A, JOUSSELME A L, OSSWALD C. Conflict measure for the discounting operation on belief functions[C]// International Conference on Information Fusion. Piscataway, NJ: IEEE Press, 2008: 1-8.
- [16] YANG Y, HAN D, HAN C. Discounted combination of unreliable evidence using degree of disagreement [J]. International Journal of Approximate Reasoning, 2013, 54(8): 1197-1216.
- [17] DENG X Y, HU Y, DENG Y, et al. Environmental impact assessment based on D numbers[J]. Expert Systems with Applications, 2014, 41(2): 635-643.
- [18] FEI L G, HU Y, XIAO F Y, et al. A modified TOPSIS method based on D numbers and its applications in human resources selection [J]. Mathematical Problems in Engineering, 2016, 2016: 6145196.
- [19] BIAN T, ZHENG H Y, YIN L K, et al. Failure mode and effects analysis based on D numbers and TOPSIS [J]. Knowledge-based Systems, 2016, 2016: 1-39.
- [20] SU X Y, MAHADEVAN S, XU P D, et al. Dependence assessment in human reliability analysis based on D numbers and AHP[J]. Risk Analysis, 2015, 35(7): 1296-1316.
- [21] DENG X Y, HU Y, DENG Y. Supplier selection using AHP methodology extended by D numbers[J]. Expert Systems with Applications, 2014, 41(1): 156-167.
- [22] FANG C, ZHONG D H, YAN F G, et al. A hybrid fuzzy evaluation method for curtain grouting efficiency assessment based on an AHP method extended by D numbers[J]. Expert Systems with Applications, 2016, 44(1): 289-303.
- [23] DENG X Y, LU X, CHAN F T, et al. D-CFPR: D numbers extended consistent fuzzy preference relations [J]. Knowledge-based Systems, 2015, 73: 61-68.
- [24] ZHOU X Y, SHI Y Q Y, DENG Y, et al. D-DEMATEL: A new method to identify critical success factors in emergency management [J]. Safety Science, 2017, 91: 93-104.
- [25] LIU W. Analyzing the degree of conflict among belief functions [J]. Artificial Intelligence, 2006, 170(11): 909-924.
- [26] KHATIBI V, MONTAZER G A. A new evidential distance measure based on belief intervals[J]. Scientia Iranica, 2010, 17(2): 119-132.
- [27] CHEN J, YE F, JIANG T, et al. Conflicting information fusion based on an improved DS combination method [J]. Symmetry, 2017, 9(11): 278.
- [28] HÖHLE U. Entropy with respect to plausibility measures[C]// Proceedings of the 12th IEEE International Symposium on Multiple-Valued Logic. Piscataway, NJ: IEEE Press, 1982: 167-169.
- [29] YAGER R R. Entropy and specificity in a mathematical theory of evidence [J]. International Journal of General Systems, 2008, 219(4): 291-310.
- [30] KLIR G J, RAMER A. Uncertainty in the Dempster-Shafer theory: A critical re-examination [J]. International Journal of General Systems, 1990, 18(2): 155-166.
- [31] KLIR G J, PARYIZ B. A note on the measure of discord [C]// 8th International Conference on Uncertainty in Artificial Intelligence. San Francisco: Morgan Kaufmann Publishers Inc., 1992: 138-141.
- [32] HARMANEC D, KLIR G J. Measuring total uncertainty in Dempster-Shafer theory: A novel approach [J]. International Journal of General Systems, 1994, 22(4): 405-419.
- [33] JOUSSELME A L, LIU C, GRENIER D, et al. Measuring ambiguity in the evidence theory [J]. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 2006, 36(5): 890-903.
- [34] DENG Y. Deng entropy: A generalized Shannon entropy to measure uncertainty [J]. Chaos, Solitons & Fractals, 2016, 91: 549-553.
- [35] JOUSSELME A L, GRENIER D, BOSSÉ É. A new distance between two bodies of evidence [J]. Information Fusion, 2001, 2(2): 91-101.
- [36] DING J, HAN D, DEZERT J, et al. Comparative study on BBA determination using different distances of interval numbers [C]// International Conference on Information Fusion. Piscataway, NJ: IEEE Press, 2017: 1-6.
- [37] YANG Y, HAN D, DEZERT J, et al. Comparative study of focal distance measures in theory of belief functions [C]// International Conference on Information Fusion. Piscataway, NJ: IEEE Press, 2017: 979-984.
- [38] ZHOU D, PAN Q, CHHIPI-SHRESTHA G, et al. A new weighting factor in combining belief function [J]. PLOS ONE, 2017, 12(5): e0177695.
- [39] YANG Y, HAN D. A new distance-based total uncertainty measure in the theory of belief functions [J]. Knowledge-Based Systems, 2016, 94: 114-123.

#### 作者简介:

赵静 女, 博士研究生。主要研究方向: 不确定推理、目标识别。

关欣 女,博士,教授,博士生导师。主要研究方向:信息对抗、目标跟踪与识别。

刘海桥 男,硕士研究生。主要研究方向:不确定推理、目标识别。

## A new conflict evidence decision method and its application

ZHAO Jing, GUAN Xin<sup>\*</sup>, LIU Haiqiao

(School of Aviation Operational Support, Navy Aviation University, Yantai 264001, China)

**Abstract:** Research on conflict evidence decision methods is an important research topic of evidence theory. In view of the existing problems in the evidence theory improvement process, such as large computational complexity, unreasonable normalization process and unsatisfactory evidence combination, this paper proposes a method based on quadratic combination for conflict evidence decision-making. Firstly, the paper proposes a new flowchart of conflict evidence decision method based on quadratic combination. Secondly, a new multiplicative normalization rule is proposed, and a new multiplicative normalization rule is analyzed by example to verify its rationality. Finally, the shortcomings of the existing conflict measurement function are analyzed, a new conflict measurement function is proposed, and the rationality of the conflict measurement function is analyzed. Through the analysis of examples and the comparison with the existing evidence combination rules, it is shown that the proposed method not only improves the calculation amount, but also improves the combination results.

**Keywords:** conflict evidence combination; conflict metric function; multiplicative normalization rule; evidence theory; quadratic combination

**Received:** 2019-01-02; **Accepted:** 2019-03-22; **Published online:** 2019-05-08 13:38

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20190506.1025.004.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190506.1025.004.html)

**Foundation items:** National Natural Science Foundation of China (91538201,61671463,61571454); National Defense Science and Technology Excellence Youth Talent Fund (2017-JCJQ-2Q-003); Taishan Scholar Engineering Special Fund (ts201712072)

\* **Corresponding author.** E-mail: 597268914@qq.com

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

DOI: 10.13700/j.bh.1001-5965.2018.0739

# 基于小波域清晰度评价的光场全聚焦图像融合

谢颖贤, 武迎春\*, 王玉梅, 赵贤凌, 王安红

(太原科技大学 电子信息工程学院, 太原 030024)



**摘 要:** 传统全聚焦图像融合以相机多次曝光拍摄的多聚焦图像为基础, 光场相机在单次曝光后可计算空间任意深度的重聚焦图像, 为后期全聚焦图像的获取提供便利。提出了一种基于小波变换的光场全聚焦图像获取算法, 可有效避免传统空域图像融合算法的块效应, 获得较高质量的全聚焦图像。该算法通过对微透镜阵列光场相机获得的 4D 光场数据进行空间变换与投影, 得到用于全聚焦图像融合的重聚焦图像, 对各帧重聚焦图像进行小波分解提取高、低频子图像集, 提出区域均衡拉普拉斯算子、像素可见度函数分别构建融合图像的高、低频小波系数实现图像融合, 其性能优于传统的区域清晰度评价函数。实验验证了所提算法的正确性和有效性, 采用 Lytro 光场相机的原始数据计算了融合全聚焦图像, 与传统图像融合算法相比, 人眼视觉效果更好, 客观图像指标也得到了提高。

**关键词:** 光场相机; 数字重聚焦; 全聚焦; 小波变换; 拉普拉斯算子; 区域清晰度; 图像可见度

**中图分类号:** TN911

**文献标识码:** A

**文章编号:** 1001-5965(2019)09-1848-07

伴随着计算摄影学新型学科领域的兴起及光场成像理论的发展, 光场相机成为近十年来国内外诸多领域关注的热点。相对于传统相机, 微透镜光场相机在主镜头后加入了微透镜阵列同时记录空间光线的位置与方向信息<sup>[1]</sup>。多维度光场信息的记录为光场相机后期图像的处理及应用提供了便利, 如利用光场相机数字重聚焦、全聚焦图像生成及深度信息计算等<sup>[2-4]</sup>。光场相机可以在单次拍照后计算空间任意深度的重聚焦图像, 是光场相机得以普遍关注的最突出技术亮点<sup>[1-2]</sup>。基于此, 各类光场高质量纹理图像获取及高精度深度信息计算得到深入研究。由于不受传统相机多次对焦拍摄获取多聚焦图像的限制, 基于光场数字重聚焦技术的全聚焦图像融合成为光场相机

的一个重要分支应用, 其对于后期纹理图像及深度图像的超分辨率重构、光场视频文件的生成也具有重要意义<sup>[5]</sup>。

目前, 针对传统图像的全聚焦融合主要分为空间域<sup>[6-7]</sup>和变换域<sup>[8-11]</sup>。空间域基于像素或块进行清晰度评价, 从不同图像中提取质量好的像素来构成全聚焦图像, 计算时间快但存在块效应问题<sup>[6]</sup>。变换域将图像分解为不同分辨率层或不同频带的子图像, 通过评价重建分辨率层或子图像构建重聚焦图像, 可有效避免块效应<sup>[9]</sup>。作为变换域的一种普遍方法, 小波变换将待融合图像分解到一系列频率信道中, 利用其分解后的塔形结构构建高、低频子图像, 分别对高、低频子图像进行融合后经小波反变换得到全聚焦图像。小波

收稿日期: 2018-12-18; 录用日期: 2019-03-29; 网络出版时间: 2019-05-30 11:17

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190529.1506.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190529.1506.001.html)

基金项目: 国家自然科学基金(61601318); 山西省青年科技研究基金(201601D021078); 山西省重点学科建设经费项目; 山西省互联网+3D打印协同创新中心; 山西省“1331工程”重点创新团队; 山西省科技创新团队(201705D131025); 太原科技大学博士启动基金(20132023); 国家留学基金

\* 通信作者. E-mail: yingchunwu3030@foxmail.com

**引用格式:** 谢颖贤, 武迎春, 王玉梅, 等. 基于小波域清晰度评价的光场全聚焦图像融合[J]. 北京航空航天大学学报, 2019, 45(9): 1848-1854. XIE Y X, WU Y C, WANG Y M, et al. Light field all-in-focus image fusion based on wavelet domain sharpness evaluation [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1848-1854 (in Chinese).

变换融合图像的质量决定于高、低频子图像的融合规则选择:对于低频子图像,一般采用均值计算方法实现融合;对于高频子图像,常采用 Sobel 算子、Prewitt 算子及拉普拉斯算子等进行评价建立融合规则<sup>[12-14]</sup>。

为了有效避免传统空域图像融合算法的块效应,获得较高质量的光场全聚焦图像,本文采用小波变换的方法来进行图像的融合。首先解码 4D 光场并采用数字重聚焦算法得到不同深度的重聚焦图像;然后通过对各重聚焦图像集进行小波分解及塔型重构构建高、低频子图像集;最后提出区域均衡拉普拉斯算子、像素可见度函数分别构建融合图像的高、低频小波系数实现图像融合。该算法有效实现了光场原始数据到全聚焦图像的转化,且图像融合质量较传统算法有所提高。

### 1 不同空间深度的重聚焦图像计算

图 1 为光场的双平面参数化模型。空间任一光线均可用其与两平面的交点确定,设光场相机的主透镜平面为  $(u, v)$  面,传感器平面为  $(x, y)$  面,光场相机记录的 4D 光场为  $L_F(x, y, u, v)$ ,由经典光辐射公式可得到全光相机焦平面的积分图像<sup>[1]</sup>:

$$I_F(x, y) = \frac{1}{F^2} \sum_{u=1}^U \sum_{v=1}^V L_F(x, y, u, v) \quad (1)$$

式中:  $F$  为主透镜平面和传感器平面间的距离;用  $X \times Y \times U \times V$  表示 4D 光场矩阵  $L_F(x, y, u, v)$  的大小。如果将像平面(传感器平面)与主透镜平面的距离由  $F$  变为  $F'$ ,新的 4D 光场矩阵用  $L_{F'}(x', y', u', v')$  表示,此时相机焦平面重聚焦图像表示为

$$I_{F'}(x', y') = \frac{1}{F'^2} \sum_{u'=1}^{U'} \sum_{v'=1}^{V'} L_{F'}(x', y', u', v') \quad (2)$$

令  $F' = \alpha_n F$ ,其中  $\alpha_n$  为调焦系数,为了方便图形表示,取 4D 空间的一个切面来得到坐标间的几何关系,如图 2 所示。

根据相似三角形原理,可得到新光场与原始光场的坐标满足<sup>[1]</sup>:

$$x' = u + (x - u)\alpha_n = \alpha_n x + (1 - \alpha_n)u \quad (3)$$

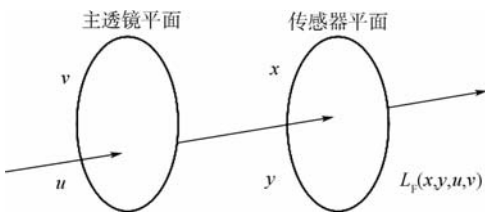


图 1 双平面参数化模型

Fig. 1 Two-plane parametric model

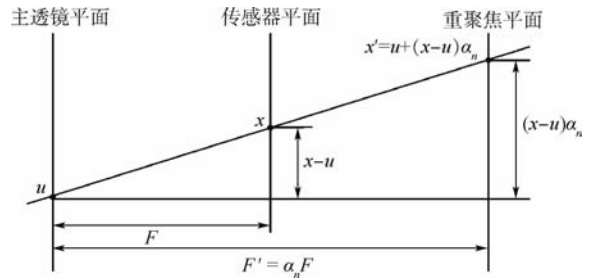


图 2 光场相机数字重聚焦原理

Fig. 2 Digital refocusing principle of light field camera

$$u' = u \quad (4)$$

同理可得到:

$$y' = v + (y - v)\alpha_n = \alpha_n y + (1 - \alpha_n)v \quad (5)$$

$$v' = v \quad (6)$$

式(3)~式(6)可表示为矩阵形式:

$$[x', y', u', v']^T = B_{\alpha_n} \cdot [x, y, u, v]^T \quad (7)$$

式中:  $B_{\alpha_n}$  为坐标变换矩阵,具体形式为

$$B_{\alpha_n} = \begin{bmatrix} \alpha_n & 0 & 1 - \alpha_n & 0 \\ 0 & \alpha_n & 0 & 1 - \alpha_n \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (8)$$

式(7)也等价于:

$$[x', y', u', v'] = [x, y, u, v] \cdot B_{\alpha_n}^T \quad (9)$$

根据式(9),式(2)可改写为

$$I_{\alpha_n F'}(x, y) = \frac{1}{\alpha_n^2 F'^2} \sum_{u=1}^U \sum_{v=1}^V L_F([x, y, u, v] \cdot B_{\alpha_n}^T) \quad (10)$$

改变  $\alpha_n$  的取值,即可达到改变像平面的位置的目的,继而得到不同空间深度重聚焦图像。

### 2 小波域光场高、低频子图像融合

根据小波变换图像融合理论,通过小波变换将待融合图像分解到一系列频率信道中,利用其分解后的塔型结构构建高、低频子图像,该过程可描述为

$$w_{\alpha_n}^H(i, j) = W_H[I_{\alpha_n F'}(x, y)] \quad (11)$$

$$w_{\alpha_n}^L(i, j) = W_L[I_{\alpha_n F'}(x, y)] \quad (12)$$

式中:  $(x, y)$  表示图像坐标;  $(i, j)$  表示小波域坐标;  $W[\cdot]$  表示小波塔型分解操作符;  $W_H[\cdot]$  表示小波塔型分解后提取高频系数(高频子图像);  $W_L[\cdot]$  表示小波塔型分解后提取低频系数(低频子图像)。根据各重聚焦图像的高、低频子图像矩阵,运用清晰度评价函数计算融合后的高、低频子图像的系数,最后进行小波逆变换可得到全聚焦图像。

小波变换图像融合方法的核心在于建立有效的高、低频子图像清晰度评价函数<sup>[9,14]</sup>,本文针对现有民用级光场相机拍摄图像对比度不高、经数字重聚焦技术获得的重聚焦图像分辨率有限、且存在由标定误差造成的局部噪声等特点,建立高频系数区域均衡拉普拉斯(Balanced Laplace based on local area, BL)算子、低频系数像素可见度(Pixel Visibility, PV)函数的清晰度评价函数实现光场

全聚焦图像生成,具体算法流程如图3所示。

光场原图像经相机解码后得到4D光场,选择不同的 $\alpha_n (n = 1, 2, \dots)$ ,利用数字重聚焦技术得到不同空间深度的重聚焦图像 $I_{\alpha_n F}$ ;计算每一帧重聚焦图像的小波高、低频子图像 $w_{\alpha_n F}^H, w_{\alpha_n F}^L$ ,高、低频子图像分别采用BL算子和PV函数作为图像融合清晰度评价指标,实现高低频系数的融合;最后经小波逆变换得到融合后的全聚焦图像。

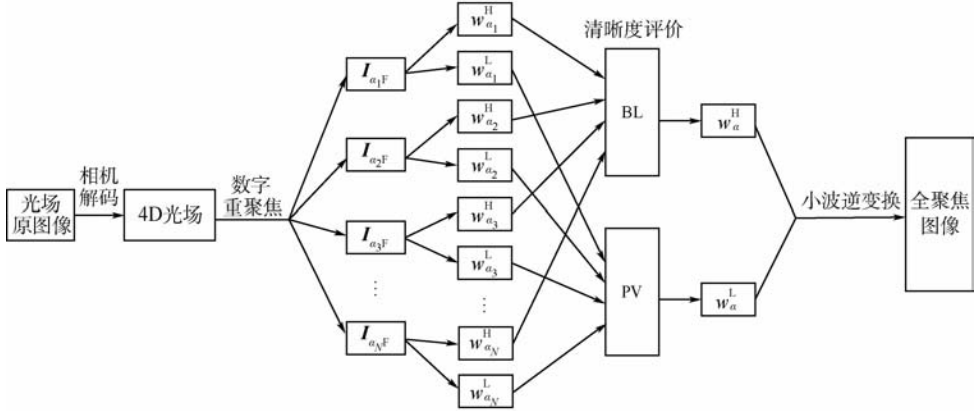
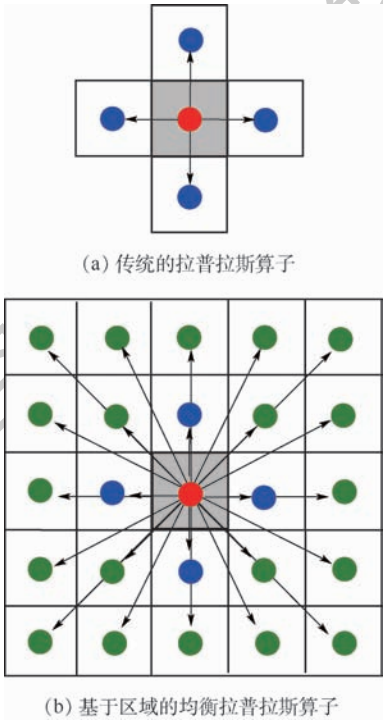


图3 光场全聚焦图像融合流程

Fig. 3 Procedure of light field all-in-focus image fusion

2.1 基于区域均衡拉普拉斯算子的高频系数融合

小波变换的高频子图像反映图像的亮度突变特性,即边界特性,拉普拉斯算子能对任何走向的边界和线条进行锐化,且保持各向同性特性,在对高频子图像进行清晰度评价时被广泛使用。针对拉普拉斯算子在 $x$ 方向和 $y$ 方向的二阶导数极可能出现符号相反的情况,如图4(a)所示,Nayar和Nakagawa<sup>[11]</sup>提出改进的拉普拉斯能量和(Sum of Modified Laplace, SML),可有效避免了两方向的能量抵消,同时引入窗口的思想来计算区域拉普拉斯能量和,可有效提高算法的抗噪能力。之后,SML又充分考虑了周边点对当前位置清晰度评价函数的影响,改进为水平、垂直以及斜对角线(45°、135°方向)共8个方向的梯度能量和。区域拉普拉斯能量求和的方法可提高图像融合质量,同时也增加了计算所需时间。基于此,本文提出区域均衡拉普拉斯算子,通过增加二阶导的数量和方向来实现能量均衡,该算法的表达式为



(a) 传统的拉普拉斯算子

(b) 基于区域的均衡拉普拉斯算子

图4 基于拉普拉斯算子的高频子图像清晰度评价

Fig. 4 High-frequency sub-image sharpness evaluation based on Laplace operator

$$BL_{\alpha_n}(i, j) = \frac{1}{S \times T} \left[ \sum_{s=-\frac{(S-1)}{2}}^{-1} \sum_{t=-\frac{(T-1)}{2}}^{\frac{(T-1)}{2}} \frac{1}{\sqrt{s^2 + t^2}} \cdot |2w_{\alpha_n}^H(i, j) - w_{\alpha_n}^H(i + s, j + t) - w_{\alpha_n}^H(i - s, j - t)| + \sum_{t=-\frac{(T-1)}{2}}^{-1} \frac{1}{|t|} |2w_{\alpha_n}^H(i, j) - w_{\alpha_n}^H(i, j + t) - w_{\alpha_n}^H(i, j - t)| \right] \quad (13)$$

式中: $w_{\alpha_n}^H$ 表示不同空间深度各重聚焦图像经小波分解后各自的高频子图像, $n = 1, 2, \dots, N$ , $N$ 为参与全聚焦图像融合的重聚焦图像帧数; $S \times T$ 表示均衡区域大小,且 $S, T$ 只能取奇数; $s, t$ 分别为水

平、垂直方向二阶导步长;  $\frac{1}{\sqrt{s^2+t^2}}$ 、 $\frac{1}{|t|}$  为权重因子,距离中心点越近的点,权重因子越大,对拉普拉斯算子值贡献越大,反之,距离中心点越远,对拉普拉斯算子值贡献越小。图4(b)即为  $S=5$ 、 $T=5$  时的均衡拉普拉斯算子。

考虑到微透镜标定误差会引起重聚焦图像产生局部噪声,拉普拉斯算子对噪声敏感的缺点,在对高频子图像进行融合前先进行双边滤波预处理,本文基于区域均衡拉普拉斯算子的高频系数融合规则可表述为:根据式(13)计算  $N$  帧待融合图像中任意2幅图像对应拉普拉斯算子的差值,记为  $D$ ,  $D = (BL_{\alpha_A}(i,j) - BL_{\alpha_B}(i,j))$ ,当所有差值  $D$  的最小值大于某一门限阈值时,取均衡拉普拉斯:能量最大者对应的高频系数  $w_{\alpha_p}^H$  作为融合系数,当差值小于门限阈值时,由多帧图像高频系数乘以权重因子来决定最后的融合系数。具体公式为

$$w_F^H(i,j) = \begin{cases} w_{\alpha_p}^H(i,j) & \min(D) \geq H_H \\ \sum_{n=1}^N k_n w_{\alpha_n}^H(i,j) & \max(D) < H_H \end{cases} \quad (14)$$

式中:  $\max(\cdot)$ 、 $\min(\cdot)$  分别为取最大值、最小值操作;  $H_H$  为自定义门限阈值(本文  $H_H$  取 0.1); 权重因子  $k_n$  计算式为

$$k_n = BL_{\alpha_n}(i,j) / \sum_{n=1}^N BL_{\alpha_n}(i,j) \quad (15)$$

### 2.2 基于像素可见度函数的低频系数融合

小波塔型分解得到的低频系数,主要反映了原图像的平均灰度特征。计算低频融合系数的最简单方法是加权平均法<sup>[14]</sup>,但加权平均法会降低融合图像的对比度并丢失原图像中一些有用信息。此外,空间频率法、点锐度算子等一些计算梯度的方法也被应用到低频融合系数的计算中<sup>[14-15]</sup>。在本文的光场图像低频系数融合中,借鉴了基于人类视觉特性的图像可见度(Image Visibility, IV)的概念<sup>[8-9]</sup>,其定义如下:

$$VI = \frac{1}{P \times Q} \sum_{i=1}^P \sum_{j=1}^Q \left(\frac{1}{\bar{I}}\right)^\gamma \frac{|I(i,j) - \bar{I}|}{\bar{I}} \quad (16)$$

式中:  $P \times Q$  表示图像  $I(i,j)$  的大小;  $\bar{I}$  为图像  $I(i,j)$  的平均值;  $\gamma$  为视觉常量,其取值范围为 0.6 ~ 0.7。IV 的值越大,代表图像可见度越高。

在低频子图像的融合过程中,如果直接采用式(16)计算,只能得到整幅图像的 IV 值,无法用于多幅图像的区域级或像素级融合。为了合理建立有效的低频系数评价指标,对式(16)进行改

进,建立基于像素的图像可见度函数,具体表达式为

$$PV(i,j) = \frac{1}{S \times T} \sum_{s=-(S-1)/2}^{(S-1)/2} \sum_{t=-(T-1)/2}^{(T-1)/2} \left(\frac{1}{w_{\alpha_n}^L}\right)^\gamma \cdot \frac{|w_{\alpha_n}^L(i+s,j+t) - \bar{w}_{\alpha_n}^L|}{\bar{w}_{\alpha_n}^L} \quad (17)$$

式中:  $S \times T$  表示以当前像素点为中心的矩形邻域,且  $S$ 、 $T$  只能取奇数;  $s$ 、 $t$  分别为在矩形邻域内水平、垂直方向的扫描步长;  $\bar{w}_{\alpha_n}^L$  为  $S \times T$  区域像素的平均灰度值。在低频系数融合过程中,采用与高频系数相同的融合规则。

### 3 实验结果与分析

为验证本文算法的有效性,采用 Lytro 光场相机拍摄的原图像进行了实验(实验测试数据来源于文献[3])。图5(a)为光场原图像,图5(b)、(c)、(d)分别为  $\alpha = 0.52, 0.78, 0.98$  时根据式(10)计算得到的3幅不同空间深度的重聚焦图像,聚焦深度从前景逐渐变化到背景。图5(e)为采用本文算法计算得到的全聚焦图像,红色虚线所框区域清晰度明显高于图5(b)对应区域,黄色虚线所框区域清晰度明显高于图5(c)对应区域,白色虚线所框区域清晰度明显高于图5(d)对

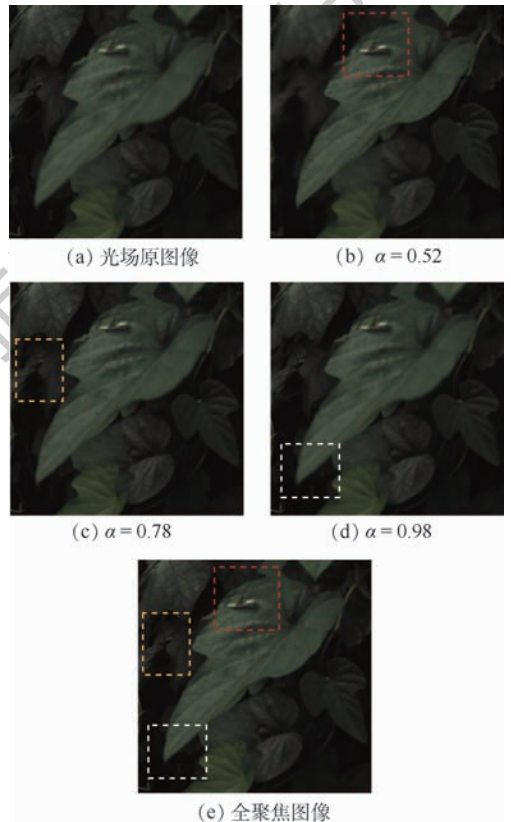


图5 基于本文算法的 Leaves 样本图像融合  
Fig. 5 Image Leaves fusion based on proposed algorithm

应区域,可见本文算法可有效利用光场原图像得到全聚焦图像。

为了从视觉上评价本文算法的优势,选取3种经典的小波变换图像融合算法与本文算法进行对比,这3种融合算法高频系数分别采用 Sobel 算子、Prewitt 算子、传统拉普拉斯算子作为清晰度评价函数,低频系数均采用点锐度算子作为清晰度评价函数,由于低频系数选用同一清晰度评价函数,故在下文的表述中该3种算法分别简称为 Sobel 算法、Prewitt 算法、Laplace 算法。实验数据采用文献[3]的另外3组光场原图像(Flower、Forest、Zither)。图6~图8为对应实验结果:图6~图8的(a)、(b)分别为3幅光场原图像对应  $\alpha = 1, 2$  时得到的2幅重聚焦图像,聚焦深度从前景变换到背景;图6~图8的(c)、(d)、(e)、(f)分别为采用 Sobel 算法、Prewitt 算法、Laplace 算法及本文算法对(a)、(b)两幅图像融合后得到的结果。

从视觉效果看,图6中 Sobel 算法、Prewitt 算法得到的融合图像在虚线所框矩形区域的清晰度明显不如本文算法;图7虚线所框区域 Sobel 算法、Prewitt 算法得到的清晰度也明显不如本文算法;图8采用Prewitt算法融合的植物叶子对应的

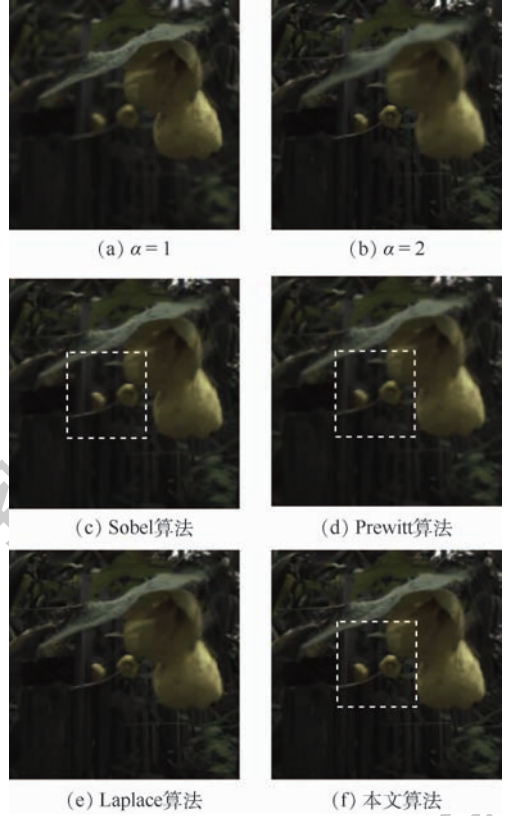


图7 Forest 样本图像不同融合算法对比  
Fig. 7 Comparison of different fusion algorithms based on image Forest

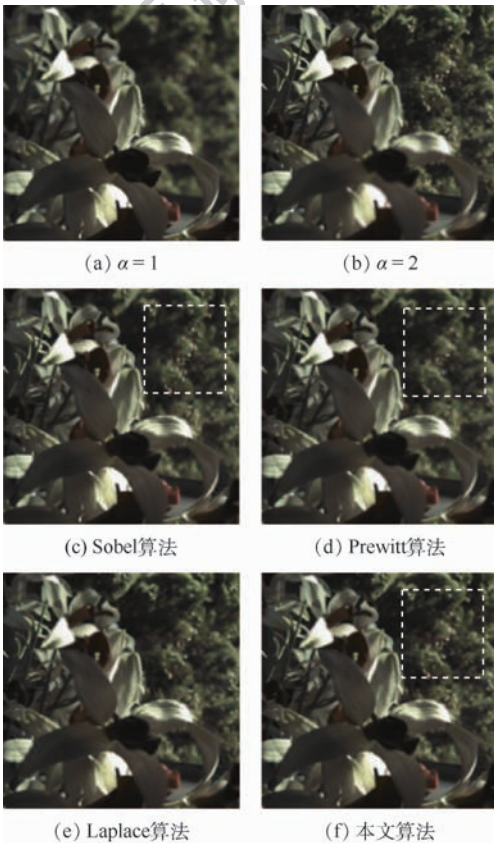


图6 Flower 样本图像不同融合算法对比  
Fig. 6 Comparison of different fusion algorithms based on image Flower

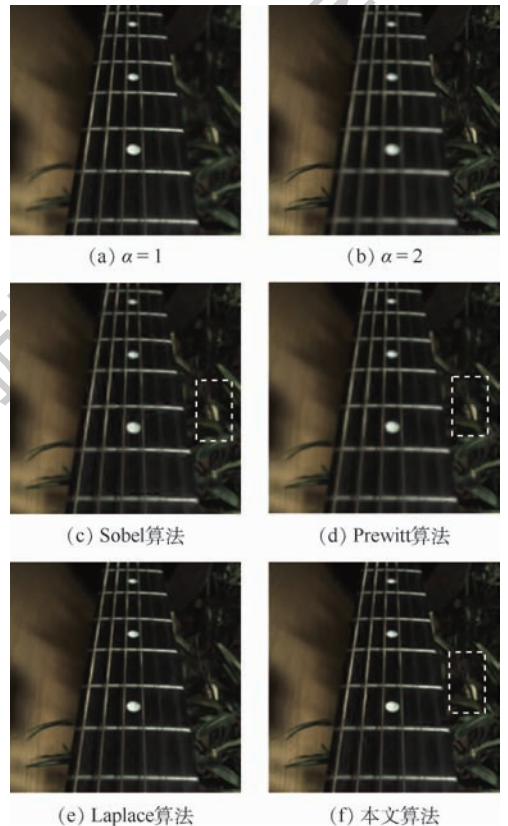


图8 Zither 样本图像不同融合算法对比  
Fig. 8 Comparison of different fusion algorithms based on image Zither

清晰度也明显不如本文算法;说明本文算法在视觉效果上具有一定优势。

另外,考虑到人眼视觉限制,本文进一步选取了一些客观评价指标对图像质量进行评价,验证所提算法的优越性。本文分别选取信息熵  $E$ 、平均梯度  $A_G$ 、图像清晰度  $F_D$  和图像边缘强度  $E_I$  作为评价指标,对图 6~图 8 中多种方法得到的全聚焦图像的质量进行评价。其中  $E$  是度量信息大小的一个物理量,其值越大表示图像信息量越大。 $A_G$  可以敏感地反映图像对微小细节反差能力,其值越高,代表它的能力越强。 $F_D$  代表图像清晰程度,其值越高,代表其清晰程度越好。 $E_I$  反映了图像的边缘强度,其值越高,代表图像边缘越清晰,具体评价指标对应结果如表 1~表 3 所示。对比表中数据可知,本文算法在图像的 4 种客观评价指标上,均优于其他 3 种传统的小波变换方法,体现了本文算法的可行性和有效性。

表 1 Flower 样本图像不同融合算法性能指标比较

Table 1 Comparison of performance indices of different fusion algorithms based on image Flower

算法	$E$	$A_G$	$F_D$	$E_I$
Sobel 算法	6.867 6	6.247 0	6.899 1	66.534 0
Prewitt 算法	6.863 4	5.832 6	6.327 0	62.642 0
Laplace 算法	6.883 0	6.866 8	7.683 7	72.307 3
本文算法	<b>6.889 6</b>	<b>7.005 5</b>	<b>7.849 8</b>	<b>73.720 3</b>

表 2 Forest 样本图像不同融合算法性能指标比较

Table 2 Comparison of performance indices of different fusion algorithms based on image Forest

算法	$E$	$A_G$	$F_D$	$E_I$
Sobel 算法	5.754 4	2.532 8	2.913 6	26.515 7
Prewitt 算法	5.749 2	2.290 5	2.576 6	24.273 5
Laplace 算法	5.801 1	3.001 8	3.523 5	31.013 4
本文算法	<b>5.809 9</b>	<b>3.087 5</b>	<b>3.630 5</b>	<b>31.903 3</b>

表 3 Zither 样本图像不同融合算法性能指标比较

Table 3 Comparison of performance indices of different fusion algorithms based on image Zither

算法	$E$	$A_G$	$F_D$	$E_I$
Sobel 算法	6.293 5	4.467 5	5.186 5	48.385 4
Prewitt 算法	6.269 5	4.018 4	4.518 2	43.756 6
Laplace 算法	6.271 6	4.864 9	5.677 3	52.447 4
本文算法	<b>6.298 7</b>	<b>4.942 5</b>	<b>5.794 0</b>	<b>53.150 1</b>

## 4 结 论

1) 基于光场原图像,完成了光场解码、数字重聚焦和全聚焦图像融合的全过程计算,利用小波变换实现多帧重聚焦图像融合,避免了传统空

域图像融合算法造成的块效应。

2) 在小波高频子图像的融合中,提出了基于区域均衡拉普拉斯算子的清晰度评价函数;在小波低频子图像融合中,提出了基于像素可见度的清晰度评价函数,来提高全聚焦图像的融合质量。

3) 实验结果表明,与传统的基于小波变换的图像融合算法相比,利用本文算法得到的融合图像的视觉效果更好,图像评价指标更高:与 Laplace 算法相比,图像清晰度平均提高 0.129 9,图像边缘强度平均提高 1.001 9。

## 参考文献 (References)

- [1] NG R. Light field photography with a hand-held plenoptic camera: CSTR 2005-02[R]. Stanford: Stanford University, 2005.
- [2] NG R. Fourier slice photography[J]. ACM Transactions on Graphics, 2005, 24(3): 735-744.
- [3] TAO M W, HADAP S, MALIK J, et al. Depth from combining defocus and correspondence using light-field cameras[C]// IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE Press, 2013: 673-680.
- [4] BOK Y, JEON H G, KWEON I S. Geometric calibration of microlens-based light-field cameras using line features[C]// European Conference on Computer Vision. Berlin: Springer, 2014: 47-61.
- [5] YOON Y, JEON H G, YOO D, et al. Light-field image super-resolution using convolutional neural network[J]. IEEE Signal Processing Letters, 2017, 24(6): 848-852.
- [6] LIU J G. Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details[J]. International Journal of Remote Sensing, 2000, 21(18): 3461-3472.
- [7] Li S T, KANG X D, HU J W. Image fusion with guided filtering[J]. IEEE Transactions on Image Processing, 2013, 22(7): 2864-2875.
- [8] FEICHTENHOFER C, FASSOLD H, SCHALLAUER P. A perceptual image sharpness metric based on local edge gradient analysis[J]. IEEE Signal Processing Letters, 2013, 20(4): 379-382.
- [9] QIN X, ZHENG J, HU G, et al. Multi-focus image fusion based on window empirical mode decomposition[J]. Infrared Physics & Technology, 2017, 85: 251-260.
- [10] FAKHARI F, MOSAVI M R, LAJVARDI M M. Image fusion based on multi-scale transform and sparse representation: Image energy approach[J]. IET Image Processing, 2017, 11(11): 1041-1049.
- [11] NAYAR S K, NAKAGAWA Y. Shape from focus[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(8): 824-831.
- [12] DENG C, MA W, YIN Y. An edge detection approach of image fusion based on improved Sobel operator[C]// IEEE International Congress on Image and Signal Processing. Piscataway, NJ: IEEE Press, 2011, 3: 1189-1193.

- [13] SHRIVAKSHAN G T, CHANDRASEKAR C. A comparison of various edge detection techniques used in image processing [J]. International Journal of Computer Science Issues, 2012, 9 (5): 268-276.
- [14] YANG Y, PARK D S, HUANG S, et al. Fusion of MT and MR images using an improved wavelet based method [J]. Journal of X-ray Science and Technology, 2010, 18(2): 157-170.
- [15] HUANG W, JING Z. Evaluation of focus measures in multi-focus image fusion [J]. Pattern Recognition Letters, 2007, 28 (4): 493-500.

作者简介:

谢颖贤 男, 硕士研究生。主要研究方向: 光信息获取与处理。

武迎春 女, 博士, 副教授, 硕士生导师。主要研究方向: 光信息获取与处理、光学三维传感。

## Light field all-in-focus image fusion based on wavelet domain sharpness evaluation

XIE Yingxian, WU Yingchun\*, WANG Yumei, ZHAO Xianling, WANG Anhong

(School of Electronic and Information Engineering, Taiyuan University of Science and Technology, Taiyuan 030024, China)

**Abstract:** Traditional all-in-focus image fusion based on the multi-focus images which are captured by multiple-exposure of the camera. Light field camera has the ability of calculating the refocused images at any depth after a single exposure, which makes it more advantageous in all-in-focus image calculation. A light field all-in-focus image fusion method based on wavelet transform is proposed in this paper. Compared with the spatial image fusion method, the proposed method can effectively avoid the block artifacts and obtain a fused image with high quality. First, the refocused images used for the all-in-focus image calculation can be computed through shearing and projecting the 4D light field captured by the microlens-based light field camera. Then, the wavelet transform are applied to the refocused images and the high-frequency and low-frequency sub-images are extracted respectively. Finally, the balanced Laplace operator and pixel visibility function are proposed to evaluate the sharpness of the sub-image and to get a high-quality fusion image. Compared to the traditional region based sharpness evaluation function, the proposed method has a better performance. The experiment results prove the correctness and validity of the proposed method. The raw images captured by Lytro light field camera are used to calculate the all-in-focus image. Compared with the traditional image fusion methods, the visual effect is better and the quantitative indices are also improved with the proposed method.

**Keywords:** light field camera; digital refocusing; all-in-focus; wavelet transform; Laplace operator; local sharpness; image visibility

**Received:** 2018-12-18; **Accepted:** 2019-03-29; **Published online:** 2019-05-30 11:17

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20190529.1506.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190529.1506.001.html)

**Foundation items:** National Natural Science Foundation of China (61601318); Shanxi Province Science Foundation for Youths (201601D021078); the Fund of Shanxi Key Subjects Construction; the Collaborative Innovation Center of Internet + 3D Printing in Shanxi Province; Key Innovation Team of Shanxi 1331 Project; Scientific and Technological Innovation Team of Shanxi Province (201705D131025); Youth Foundation of Taiyuan University of Science and Technology (20132023); Foundation of China Scholarship Council

\* **Corresponding author.** E-mail: [yingchunwu3030@foxmail.com](mailto:yingchunwu3030@foxmail.com)

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

DOI: 10.13700/j.bh.1001-5965.2018.0724

# 基于深度学习的无人机数据链信噪比估计算法

孙宇航<sup>1</sup>, 曾国奇<sup>2,\*</sup>, 刘春辉<sup>2</sup>, 张多纳<sup>1</sup>

北京航空航天大学 电子信息工程学院, 北京 100083; 2. 北京航空航天大学 无人系统研究院, 北京 100083)



**摘 要:** 无人机数据链通信受到各种自然与人为的干扰, 信噪比(SNR)是信道状态和通信质量的有效评估指标。为解决传统估计算法信噪比估计精度不足的问题, 提出了一种卷积神经网络(CNN)与长短时记忆(LSTM)网络结合的估计模型。利用仿真与实测相结合的方式, 构建了一个包含不同信噪比、调制方式、衰落信道等信息的无人机通信信号数据集; 在网络训练阶段, 将样本序列进行分割, 对分割后的每一部分序列使用 CNN-LSTM 网络提取深度特征, 多次训练并保存模型参数; 在测试阶段, 利用构建好的测试集完成对算法的验证与测试, 得到信噪比估计值。实验表明, 相比于传统信噪比估计算法与单一网络结构的深度学习算法, 所提算法的均方误差最低, 实现了对信噪比的高精度估计。

**关键词:** 无人机; 信噪比(SNR)估计; 深度特征; 卷积神经网络(CNN); 长短时记忆(LSTM)网络

**中图分类号:** TN911.6

**文献标识码:** A **文章编号:** 1001-5965(2019)09-1855-09

地-空数据链作为无人机系统的重要组成部分, 发挥着发送上行遥控指令和回传遥测侦察信息等重要作用<sup>[1]</sup>。信噪比(Signal-to-Noise Ratio, SNR)是评价无人机通信系统信道环境和通信质量的重要指标, 精确的信噪比估计既可以为无人机数据链提供功率控制、信道分配所需要的信息, 又可以促使数据链系统采用更高效的信号处理算法来提高信息解调性能。

信噪比估计传统算法主要分为基于最大似然(Maximum Likelihood, ML)估计法、统计量估计法、谱分析(Spectrum Analysis, SA)估计法等。文献[2]提出了基于 ML 的高斯信道下低偏差信噪比估计算法; 文献[3]结合期望值最大化技术与 ML 技术, 研究了 SIMO(Single-Input-Multiple-Output)系统下高维度调制信号的信噪比估计方法, 扩大了 ML 类算法的适用范围。ML 估计法理论上估计效果最接近克拉美罗下界(Cramer-Rao

Lower Bound, CRLB), 但在低信噪比情况下性能很差, 且运算量较大。基于统计量估计法中分隔符号矩算法由美国喷气推进实验室提出, 已用于美国航空航天局(NASA)的深层空间研究中<sup>[4]</sup>; 文献[5]研究了基于接收信号高阶统计量的改进二阶四阶矩(M2M4)算法; 文献[6]进一步研究了该类算法在 SIMO 系统中的性能。基于统计量估计法计算简单, 但需要数据足够多, 并且随着调制阶数的增加, 高信噪比情况下估计误差较大。文献[7]利用移动信道的功率谱特性, 提出了一种基于 SA 的平均信噪比估计算法; 文献[8]利用功率谱提出一种无需先验信息的快速信噪比盲估计算法。该类算法计算量小, 但估计精度相对较低。

深度学习(deep learning)近年来在图像/视频处理、自然语言处理等领域发展迅速, 但在通信信号领域的研究却刚刚起步。其中, 文献[9]将卷积神经网络(Convolutional Neural Networks,

收稿日期: 2018-12-17; 录用日期: 2019-01-23; 网络出版时间: 2019-03-15 09: 34

网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20190313.1649.001.html

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

**引用格式:** 孙宇航, 曾国奇, 刘春辉, 等. 基于深度学习的无人机数据链信噪比估计算法[J]. 北京航空航天大学学报, 2019, 45(9): 1855-1863. SUN Y H, ZENG G Q, LIU C H, et al. SNR estimation algorithm for UAV data link based on deep learning[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1855-1863 (in Chinese).

CNN)应用到通信信号的调制方式自动分类(Automatic Modulation Classification, AMC)问题中,取得较好的效果;文献[10]提出了一种融合深度神经网络的结构,在该问题中得到了更高的分类准确率;文献[11]使用深度学习算法,考虑载波频率偏移、符号率和多径衰落的影响,实现了对无线电通信信号的分类;文献[12]提出了将一维信号转化为类似图像的二维矩阵,使用二维CNN网络实现信号分类。

目前,深度学习算法尚未在通信信号信噪比估计问题得到应用,其主要难点在于:①网络层面:传统的网络结构提取通信信号特征的能力有限;②数据层面:缺少无人机通信信号公开的数据集。

针对以上问题,本文根据无人机数据链信号的特点,提出一种CNN与长短时记忆(Long Short Term Memory, LSTM)网络结合的估计算法。通过结合仿真数据与实测数据,构建了无人机通信信号数据集,利用该数据集对网络模型进行训练。使用CNN-LSTM网络对分割后的信号进行特征提取,得到特征向量,最终通过全连接层融合特征并计算信噪比估计值。通过测试,验证了本文算法可以很好地估计无人机数据链信号的信噪比。

## 1 CNN与LSTM网络适用性分析

在信噪比估计问题中,传统算法普遍根据待测信号自身的特点进行估计,而没有用到其他信号的数据。深度神经网络则是在经过大量信号数据的训练后,自动提取待测信号的深度特征,并将特征学习融入到了建立模型的过程中,减小单一信号计算带来的误差,提高估计精度。因此,在数据样本充足的情况下,深度学习算法可以更好地实现对信噪比的估计。

CNN模型<sup>[13]</sup>可以提取数据在空间结构上的关系,在图像识别<sup>[14]</sup>等领域有着出色的表现。CNN网络通过共享局部视野,减少了神经网络参数的数量,并且通过权值共享使运算量降低。本文将分割后的波形数据作为CNN模型的输入,每个卷积层的输出经过非线性激活后作为下一个卷积层的输入,利用多个卷积层堆叠的形式来更深入地描述信号的空间属性。

时间属性对信号数据尤为重要,虽然CNN网络可以很好地提取空间特征,但无法模拟时间序列的变化,因此需要利用LSTM模型时序地实现网络中的信息传递,循环提取特征,反复学习并生成新特征。由于待测无人机数据链序列普遍较长,LSTM网络则可以解决处理长序列时会发生

的梯度爆炸或者梯度消失<sup>[15-16]</sup>问题。

综上所述,CNN和LSTM网络分别在空间和时间2个不同的维度对无人机数据链信号进行特征提取,可以发掘不同信噪比的信号之间深层次的差别。因此,本文选择CNN-LSTM网络结构来解决信噪比估计问题。

## 2 信噪比估计算法

无人机数据链信号具有信道空间特殊,信道时间动态化等特点,针对这些特点,本节给出了基于CNN-LSTM网络的信噪比估计算法,网络模型如图1所示。

本模型采用3层CNN网络刻画信号之间的深层次差异,每个卷积层的卷积核尺寸都是不同的,通过多次特征提取可以得到更多有效特征,这种方式减少了学习参数,简化了最终特征图的结构,便于网络之间融合。融合模型的剩余部分采用LSTM网络来学习并记忆信号数据时序特征,在数据链信号信噪比估计的问题上,本文LSTM将每一个数据点作为一个时间步,将CNN网络的输出特征图输入至LSTM网络,最终的输出结果表征信号的时序关联特征。

在训练阶段,首先把一条长序列的信号数据分割成短序列,对每一段短序列使用一维CNN网络提取特征,将生成的特征序列输入至LSTM网络进一步提取特征,最终在全连接层将每一段短序列的特征融合,输出损失函数loss,通过优化算法迭代优化网络模型,继续训练下一条数据,直到loss收敛,将模型参数保存,训练结束。本文采用的优化算法是文献[17]提出的Nadam算法,在Adam算法的基础上加入了Nesterov加速梯度<sup>[18]</sup>,既可以防止振荡与局部极值的出现,又可以自适应地更新学习率。与其他算法相比,可以得到更快的收敛速度与更好的收敛效果。

在测试阶段,对测试信号进行同样的分割处理,将处理后的数据输入至已经训练好的网络模型中,最终输出得到信噪比估计值与loss值,测试得到的loss值可以直观地反映模型的训练效果。

### 2.1 数据分割

本文所提算法需要对训练集中的信号逐条进行训练,将训练集信号数据表示为

$$S_{\text{train}} = \{S_1, S_2, \dots, S_k\} \quad (1)$$

式中: $S_i$ 为训练信号序列。

训练集标签表示为

$$\text{lable}_{\text{train}} = \{\text{snr}_1, \text{snr}_2, \dots, \text{snr}_k\} \quad (2)$$

式中:  $\text{snr}_i$  为信噪比真值。

在使用 LSTM 模型提取特征时,遗忘门会读取当前输入信息与上一时刻状态,通过 sigmoid 函数输出一个 0 到 1 之间的数值到当前细胞状态中,这就不可避免地造成信息丢失。一些学者研究人员尝试着改变 LSTM 的门结构来解决问题,但普遍没有取得预期效果。文献[19]分析总结了超过一万种 LSTM 的变体结构,最终得到的结论是相比于标准 LSTM 结构,没有任何一种变体结构可以有大幅度的改进。因此研究的重点转移到了对数据本身的处理,如归一化、时频域转换等。文献[19]指出分割信号数据可以一定程度地提升神经网络效果,并给出了基于分割窗的信号数据分割算法,信号分割窗长度  $L$  为其中的关

键参数。适当的数据分割可以将一个长序列分解为几段短序列,减少在处理长序列时 LSTM 网络遗忘门对重要信息的丢失程度。

本文选择对每一条信号进行数据分割,把长序列分解为若干等长的短序列,分别保存在不同的数组中。设第  $i$  条输入信号的原始序列为

$$S_i = \{s_{i_1}, s_{i_2}, \dots, s_{i_k}\} \quad (3)$$

分割过程如图 2 所示。

该条信号的标签为  $\text{lable}_i = \text{snr}_i$ , 分割窗口的长度为  $L$ 。经过分割后的信号表示为

$$X = \{X_1, X_2, \dots, X_l\} \quad (4)$$

$$X_j = \{s_{i_{(j-1)L+1}}, s_{i_{(j-1)L+2}}, \dots, s_{i_{jL}}\} \quad (5)$$

$$1 \leq j \leq l, j \in \mathbb{N}$$

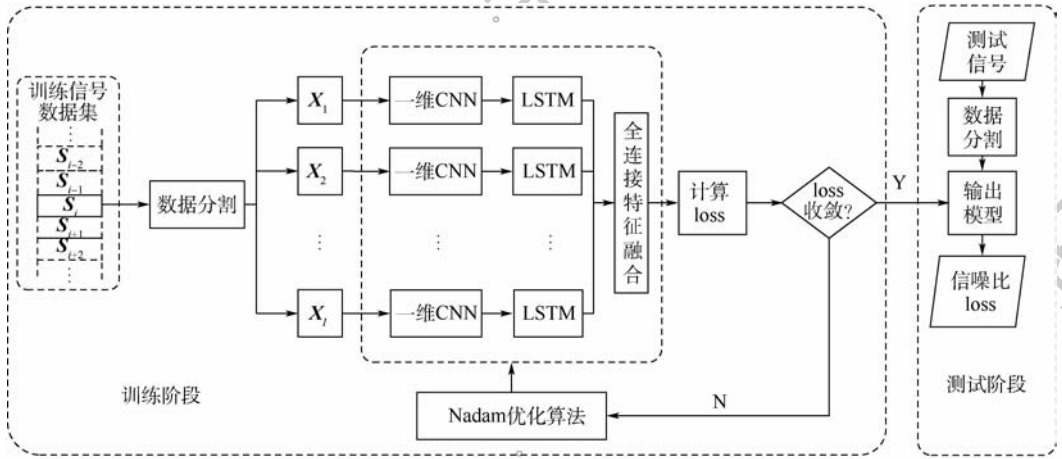


图 1 基于 CNN-LSTM 信噪比估计网络框架

Fig. 1 Network framework of SNR estimation based on CNN-LSTM

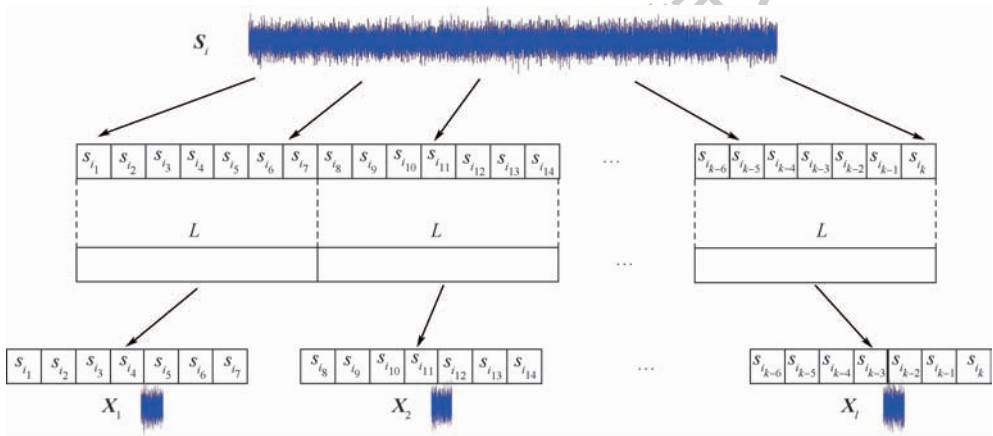


图 2 基于分割窗的信号数据分割算法示意图

Fig. 2 Schematic diagram of signal data segmentation algorithm based on split window

## 2.2 CNN 网络特征提取

将分割后的信号数据输入至图 3 所示网络结构中,利用 CNN 网络对  $X$  的每个子集分别进行特征提取。本文使用的一维 CNN 网络采取 3 个卷积层级联的结构,由于相比于二维图像数据,一维

信号数据结构简单、卷积运算时计算量小,因此没有在卷积层之间加入池化层,避免了关键信息的损失。

输入信号以图 3 中的  $X_1$  为例:

$$X_1 = \{s_{i_1}, s_{i_2}, \dots, s_{i_l}\} \quad (6)$$

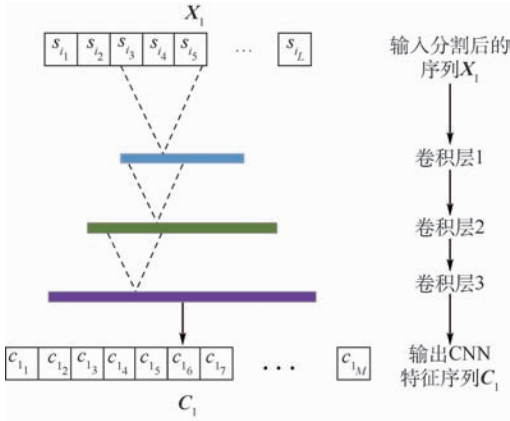


图3 CNN网络运算流程图

Fig. 3 Operation flowchart of convolution network

设定卷积层1的卷积核大小为  $a * 1$ , 卷积层2的卷积核大小为  $b * 1$ , 卷积层3的卷积核大小为  $c * 1$ , 卷积计算过程如图3所示。

输入序列按照式(7)进行计算:

$$h_i^k = f_a((\mathbf{W}_k \mathbf{x})_i + \mathbf{b}_k) \quad (7)$$

式中:  $\mathbf{x}$  为输入序列;  $h_i^k$  为第  $i$  个一维数据在第  $k$  层进行卷积运算后的值;  $f_a$  为激活函数, 如 tanh 和 sigmoid 等;  $\mathbf{W}_k$  为该层的权重值;  $\mathbf{b}_k$  为对应的偏置值。最终 CNN 网络输出的特征序列为

$$\mathbf{C}_1 = \{c_1, c_2, \dots, c_{l+a+b+c-3}\} \quad (8)$$

### 2.3 LSTM 网络特征进一步提取

将 CNN 网络输出的特征序列输入至 LSTM 网络中, 该网络由  $(L + a + b + c - 3)$  个同构 LSTM 细胞构成, 每个细胞单元按照时间顺序连接。

经过 LSTM 网络后输出的特征序列表示为

$$\mathbf{L}_i = \{l_{i_1}, l_{i_2}, \dots, l_{i_{L+a+b+c-3}}\} \quad (9)$$

$$l_{i_j} = f_{lstm}(c_{i_j}, h_{i_{j-1}}, D_{i_{j-1}}) \quad (10)$$

$1 \leq j \leq L + a + b + c - 3, j \in \mathbf{N}$

式中:  $f_{lstm}$  为 LSTM 细胞前向计算方法;  $c_{i_j}$ ,  $h_{i_{j-1}}$  和  $D_{i_{j-1}}$  分别为当前细胞的输入、前一个细胞的输出与前一个细胞的状态。

### 2.4 全连接层特征融合

对于  $\mathbf{X}$  中每一段信号  $\mathbf{X}_i (1 \leq i \leq l)$ , 经过 CNN 网络与 LSTM 网络之后都会输出一段特征序列  $\mathbf{L}_i (1 \leq i \leq l)$ , 将这些特征序列按照时间顺序排列, 得到一个维度为  $[l(L + a + b + c - 3), 1]$  的向量  $\mathbf{f}$ 。

在全连接网络中本文选择 2 个全连接层级联的方式, 经过全连接网络计算后, 最终输出矩阵为  $\mathbf{R}^{1 \times 1}$ , 该矩阵的数值即为信噪比估计值  $p_i$ 。选用均方误差 (Mean Square Error, MSE) 作为训练准确程度的衡量标准, 因此损失函数 loss 可定义为

$$\text{loss} = \sum_{i=1}^k (p_i - \text{snr}_i)^2 / K \quad (11)$$

式中:  $K$  为批尺寸 (batch\_size) 的值。

在深度神经网络结构已经确定的情况下, 可以将网络表示为

$$\mathbf{p}_j = \mathbf{W}\mathbf{S}_j + \mathbf{B} \quad (12)$$

式中:  $\mathbf{W}$  为网络的权重矩阵;  $\mathbf{B}$  为网络的偏置矩阵。首先对网络中各个节点的参数进行初始化, 设初始参数分别为  $\mathbf{W}_0$  和  $\mathbf{B}_0$ , 经过一次训练后, 以  $\mathbf{W}$  为例, Nadam 优化函数对损失函数进行优化, 按照式(13)~式(17)计算并更新网络参数  $\mathbf{W}$ :

$$\mathbf{g}_t = \nabla F(\mathbf{W}_t) \quad (13)$$

$$\mathbf{m}_t = \phi(\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_t) \quad (14)$$

$$\mathbf{V}_t = \varphi(\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_t) \quad (15)$$

$$\Delta \mathbf{W}_t = \eta \mathbf{m}_t / \sqrt{\mathbf{V}_t} \quad (16)$$

$$\mathbf{W}_{t+1} = \mathbf{W}_t - \Delta \mathbf{W}_t \quad (17)$$

式中:  $F$  为目标函数;  $\eta$  为初始学习率;  $\mathbf{m}$  和  $\mathbf{V}$  分别为对梯度的一阶矩估计与二阶矩估计;  $\phi$  和  $\varphi$  分别为计算梯度一阶矩与二阶矩的函数。

为了达到最小化 loss 值的目, 每次训练完成后, 网络参数  $\mathbf{W}$  与  $\mathbf{B}$  都会更新, 当 loss 最终收敛时, 将网络模型的参数  $\mathbf{W}'$  与  $\mathbf{B}'$  保存。

## 3 实验验证

本节建立了一个包含信噪比等信息的无人机通信信号数据集, 并且对第 2 节所提出的基于 CNN-LSTM 的信噪比估计算法进行实验验证。

### 3.1 实验准备

在进行实验之前, 需要对数据集、对比模型、评价模型指标以及实验平台的软硬件环境进行详细的描述。

#### 3.1.1 无人机通信信号数据集建立

目前, 尚无公开可用的通信信号数据集, 因此本文基于真实的地理环境, 结合无人机信号的特点, 通过仿真与实测相结合的方式生成信号数据。

数据集属性如表 1 所示。

编号: 按照生成信号的顺序定义了信号数据

表 1 通信信号数据集属性

Table 1 Attributes of communication signal data set

序号	字段名	定义
1	编号	数据样本的顺序
2	SNR	信噪比
3	调制类型	信息和载波结合的方式
4	载波频率	通信信号的工作频段
5	样本数	样本总共的条数
6	信号长度	信号的序列长度
7	码元速率	信道中码元传输的速率

样本的编号。

SNR:有效信号功率与噪声功率的比值,本数据集的信噪比范围为 $-10 \sim 20$  dB。

调制类型:常用的无人机通信信号调制方式共有4种:2PSK、4PSK、8PSK和16QAM。

载波频率:按照工信部颁布的无人机通信使用频段:840.5 ~ 845、1 430 ~ 1 444和2 408 ~ 2 440 MHz。

样本数:数据集中样本总条数,本数据集包含10 000条数据,其中仿真8 000条,实测2 000条。

信号长度:每条信号的序列长度,本数据集中单条信号长度为2 000个点。

码元速率:单位时间内通过信道的码元个数,本数据集里码元速率的范围在5 000 ~ 20 000。

### 3.1.2 对比算法

本文将基于CNN-LSTM网络的信噪比估计算法与以下5种算法进行对比实验,其中ML、MLM4、SA为传统估计算法,基于CNN和LSTM的深度学习为单一网络结构的估计算法。

#### 1) ML估计算法

信噪比的ML估计算法基于ML估计理论。将接收端的复信号分解为同相分量与正交分量,分别得到有效信号的同相正交分量联合概率分布函数与噪声的同相正交分量联合概率分布函数,进而求得有效信号与噪声的功率估计值。利用ML理论的重要性质:2个参量比值的估计等于2个参量估计值的比值,可以求得信噪比估计值。

#### 2) M2M4估计算法

M2M4估计算法主要利用了信号的能量与信号矩的关系,在统计的基础上计算出接收信号的二四阶矩,利用信号与噪声的独立性计算出信号的功率和噪声的功率,从而计算出信噪比估计值。

#### 3) SA估计算法

SA估计算法是将时域信号转换到频域进行处理。首先利用傅里叶变换以及周期图法得到含噪信号的功率谱,然后计算信号的起止频率,功率谱分割为有效信号部分与噪声部分,最后,分别计算每一部分的功率谱值,可计算得到信噪比估计值。

#### 4) 基于CNN的深度学习估计算法

将CNN-LSTM网络结构中LSTM部分去除,单独使用CNN网络提取信号特征,生成的特征序列直接输入至全连接层进行融合,信号数据处理方法不变,训练与测试流程不变。

#### 5) 基于LSTM的深度学习估计算法

将CNN-LSTM网络结构中CNN部分去除,单独使用LSTM网络提取信号特征,生成的特征序

列直接输入至全连接层进行融合,信号数据处理方法不变,训练与测试流程不变。

### 3.1.3 模型评价指标

对于不同信噪比估计算法,估计精度是评价算法优劣的最重要的指标。参考损失函数loss,本文选择MSE作为估计精度的衡量标准。MSE计算公式可表示为

$$\text{MSE} = \sum_{k=1}^K (p_k - \text{snr}_k)^2 \quad (18)$$

式中: $\text{snr}_k$ 为第 $k$ 条信号的标签,即信噪比真值; $p_k$ 为该信号经过网络计算出的信噪比的估计值; $K$ 为一次测试信号的总条数。通过计算测试集的MSE,可以准确直观地评价模型的精度。

### 3.1.4 实验平台及软硬件环境

本实验所使用的计算机配置如下:处理器为英特尔酷睿i7-8700K;CPU频率为3.7 GHz;内存16 GB;GPU为Nvidia TITAN-XP,显存12 GB;操作系统为ubuntu18.04;程序设计语言为Python + MATLAB;集成开发环境为Pycharm Community Edition 2018.8.2与MATLAB 2017b。

## 3.2 实验结果

在实验验证阶段,从测试数据集中选取8 000条信号作为训练集,剩余2 000条信号作为测试集。使用训练集对2.1节所提出的深度网络进行训练,训练参数如下:分割窗口长度 $L$ 取400,训练步数steps取30 000,初始学习率 $\eta$ 取0.005。

测试集信号属性与训练集相同,共有2 000条信号,信噪比范围在 $-10 \sim 20$  dB,调制类型包含2PSK、4PSK、8PSK和16QAM这4种,载波频率分布在840.5 ~ 845、1 430 ~ 1 444和2 408 ~ 2 440 MHz间,码元速率分布在5 000 ~ 20 000,每条信号长度为2 000个点。

分别使用本文所提算法与ML、M2M4、SA这3种传统估计算法对测试集进行验证,并且将误差结果与CRLB作对比。为了便于观察,实验生成的误差结果MSE均以对数形式表达。由图4所示的比较结果可以看出,本文算法的估计精度在各个信噪比的情况下均优于传统估计算法,估计效果最接近CRLB,证明了相比传统估计算法,深度学习的估计算法可以更有效地提升信噪比估计精度。

为了验证CNN-LSTM网络结构的收敛性,本文将CNN-LSTM网络结构替换为单独的CNN网络与单独的LSTM网络,分别以相同的训练参数进行训练。训练参数如下:分割窗口长度 $L$ 取400,训练步数steps取30 000,初始学习率 $\eta$ 取

0.005。将每一次训练的 loss 值保存,图 5 展示了 loss 随训练步数变化的趋势,可以看出:LSTM 网络收敛速度较慢;CNN 网络则在收敛过程中振荡比较严重;CNN-LSTM 网络无论在收敛速度还是收敛精度上均优于其他 2 种网络。

在估计精度对比方面,分别使用已经训练好的 3 种模型对测试集进行实验,同样将误差结果 MSE 以对数形式表达,结果如图 6 所示。由图 6 可以发现,CNN-LSTM 网络预测精度远高于 CNN 网络,除了在信噪比为 15 dB 时预测精度与 LSTM 网络几乎相同,其余信噪比情况效果均好于 LSTM 网络。

为了验证码元速率对网络估计性能的影响,在测试集中将码元速率为 5 000、10 000、15 000、20 000 的信号分别提取出来构建测试子集,将每个测试子集的信号输入至已训练好的 CNN-LSTM 网络中进行验证,结果如图 7 所示。由图 7 可以发现,在信噪比相同时,网络对不同码元速率的信号估计精度差异很小,因此说明码元速率对于网

络估计效果影响较小。

为了进一步验证分割窗长度  $L$  对网络估计性能的影响,将不使用数据分割的网络与分割窗长度分别取 1 000、500、400、200、100 的网络进行对比实验。

以相同的训练参数分别训练每个网络并保存模型,训练步数 steps 取 30 000,初始学习率  $\eta$  取 0.005,测试结果如表 2 所示。观察可得当信噪比为  $-10 \sim 15$  dB 的情况下, $L$  取 400 时 MSE 值最低;信噪比为 20 dB 的情况下, $L$  取 500 时 MSE 值为 0.021,稍低于  $L$  取 400 时的 0.023。将表 2 中数据表示在图 8 中,可以发现:分割窗长度  $L > 400$  时,相比与未经数据分割的模型,估计精度随  $L$  值的减小而提高; $L = 400$  时,性能最佳;当  $L$  值继续减小,估计性能急剧下降,图 8 中  $L$  为 200 与 100 时误差非常大。

综上所述,深度神经网络利用大量数据训练并建立模型,减少了单一测试信号计算带来的误差,因此整体表现优于传统算法;CNN 和 LSTM 网

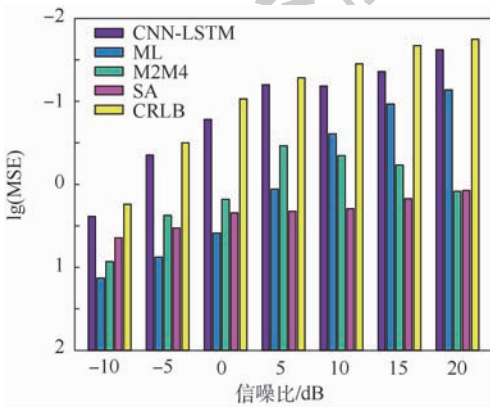


图 4 不同估计算法精度对比

Fig. 4 Comparison of accuracy between CNN-LSTM network and traditional algorithms

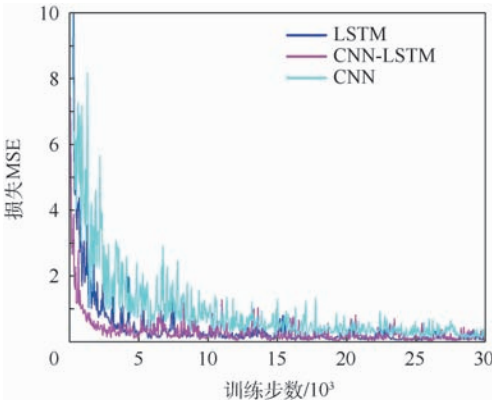


图 5 不同网络结构的训练损失变化对比

Fig. 5 Comparison of training loss change with different network structures

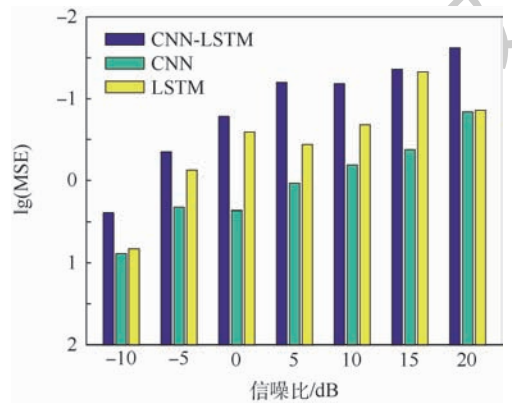


图 6 不同网络结构的测试精度对比

Fig. 6 Comparison of test accuracy with different network structures

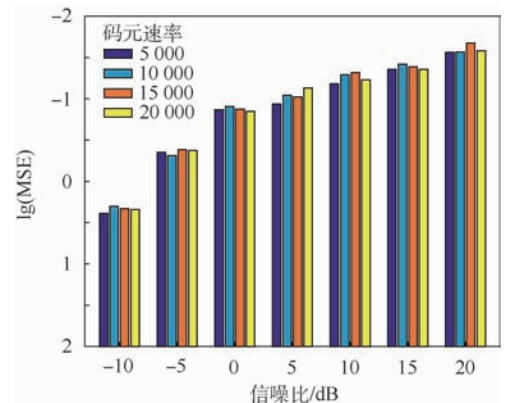


图 7 不同码元速率的测试精度对比

Fig. 7 Comparison of test accuracy with different symbol transmission rates

表2 不同分割窗长度实验结果对比

Table 2 Comparison of experimental results of different split window lengths

分割窗长度	损失 MSE						
	SNR = -10 dB	SNR = -5 dB	SNR = 0 dB	SNR = 5 dB	SNR = 10 dB	SNR = 15 dB	SNR = 20 dB
无分割	4.417	1.542	0.911	0.796	0.726	0.423	0.315
1000	3.547	0.896	0.624	0.441	0.272	0.137	0.095
500	2.716	0.562	0.211	0.146	0.092	0.063	<u>0.021</u>
400	<u>2.436</u>	<u>0.446</u>	<u>0.165</u>	<u>0.063</u>	<u>0.065</u>	<u>0.043</u>	0.023
200	8.811	6.428	3.347	3.574	2.916	4.467	5.324
100	13.424	11.756	8.815	7.741	7.629	9.945	7.453

注:最小 MSE 值由下划线标记。

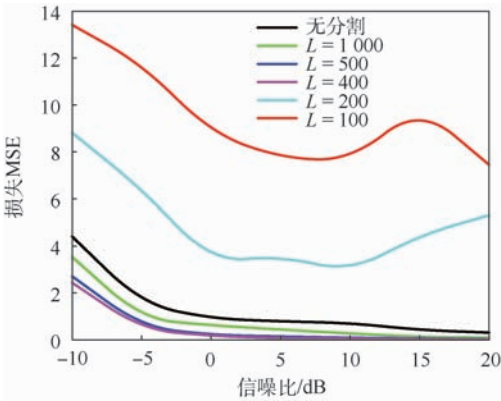


图8 不同分割窗长度对测试精度的影响

Fig.8 Effect of split window length on test accuracy

络分别在空间和时间 2 个不同的维度对无人机数据链信号进行特征提取,具有更强的特征提取能力,可以发掘信号间更深层次的差异,所以融合网络模型的训练收敛速度估计精度均高于单一网络模型;LSTM 网络在处理长序列时,每个单元会利用遗忘门将部分信息丢弃,如果序列过长,后面单元的遗忘门会丢弃一些前面序列的重要信息,造成了部分特征丢失,而适当的数据分割可以减小单次通过网络序列的长度,减少了特征的丢失,所以序列分割会使估计精度有一定的提升。但如果分割窗长度太短,深度网络无法对每一部分过短的序列提取特征,造成了更多的信息丢失,因此会出现估计性能急剧下降的现象。

## 4 结论

本文基于 CNN-LSTM 深度神经网络,提出了一种无人机信号的信噪比估计算法,并根据实际地理环境信息建立了无人机通信信号数据集。通过分析和实验结果表明:

1) 与传统估计算法相比,深度神经网络在无人机通信信号信噪比估计方面整体性能更优。

2) CNN-LSTM 网络模型比单一的 CNN 网络或 LSTM 网络提取特征能力更强,得到的模型估

计精度更高。

3) 在一定范围内,信号分割窗长度越小,CNN-LSTM 网络的性能越好;然而如果分割窗长度太小,网络性能急剧下降。

本文使用的数据分割算法能够提升深度神经网络的性能,但目前数据分割算法是将长序列分成几段等长的短序列,如何更科学地分割信号数据,根据不同信号的特点将其分割成若干段长度不同的短序列,进而进一步提高估计精度,则是下一步需要深入研究的工作。

## 参考文献 (References)

- [1] 张文秋,丁文锐,刘春辉.一种无人机数据链信道选择和功率控制方法[J].北京航空航天大学学报,2017,43(3):583-591.  
ZHANG W Q,DING W R,LIU C H. A channel selection and power control method for UAV data link[J]. Journal of Beijing University of Aeronautics and Astronautics,2017,43(3):583-591 (in Chinese).
- [2] PAULUZZI D R,BEAUCIEU N C. A comparison of SNR estimation techniques for the AWGN channel[J]. IEEE Transactions on Communications,2000,48(10):1681-1691.
- [3] BOUJELBEN M A,BELLILI F. EM algorithm for non-data-aided SNR estimation of linearly-modulated signals over SIMO channels[C]// Global Telecommunications Conference. Piscataway, NJ: IEEE Press,2009:4464-4469.
- [4] XIAO H F,SHI Y Q,SU W, et al. An investigation of non-data-aided SNR estimation techniques for analog modulation signals[C]// IEEE Sarnoff Symposium. Piscataway, NJ: IEEE Press,2010:351-355.
- [5] ÁLVAREZ-DÍAZ M,LÓPEZ-VALCARCE R,MOSQUERA C. SNR estimation for multilevel constellations using higher-order moments[J]. IEEE Transactions on Signal Processing,2010,58(3):1515-1526.
- [6] STEPHENNE A,BELLILI F,AFFES S. Moment-based SNR estimation over linearly-modulated wireless SIMO channels[J]. IEEE Transactions on Wireless Communications,2010,9(2):714-722.
- [7] 华惊宇,黄清,滑翰,等.一种移动环境下的信噪比估计算法及其在多普勒频移估计中的应用[J].通信学报,2005,26

- (5):135-140.
- HUA J Y, HUANG Q, HUA H, et al. A SNR estimation algorithm in mobile environment and its application in Doppler shift estimation[J]. Transactions of Communications, 2005, 26(5): 135-140(in Chinese).
- [ 8 ] 彭耿, 黄知涛, 陆风波, 等. 中频通信信号信噪比的快速盲估计[J]. 电子与信息学报, 2010, 32(1): 102-106.
- PENG G, HUANG Z T, LU F B, et al. Fast blind estimation of signal-to-noise ratio of IF communication signals[J]. Journal of Electronics & Information Technology, 2010, 32(1): 102-106 (in Chinese).
- [ 9 ] O'SHEA T J, HOYDIS J. An introduction to deep learning for the physical layer[J]. IEEE Transactions on Cognitive Communications & Networking, 2017, 3(4): 563-575.
- [ 10 ] ZHANG D N, DING W R, ZHANG B C, et al. Automatic modulation classification based on deep learning for unmanned aerial vehicles[J]. Sensors, 2018, 18(3): 924-939.
- [ 11 ] O'SHEA T J, ROY T, CLANCY T C. Over the air deep learning based radio signal classification[J]. IEEE Journal of Selected Topics in Signal Processing, 2017, 12(1): 168-179.
- [ 12 ] O'SHEA T J, CORGAN J, CLANCY T C. Convolutional radio modulation recognition networks[C]// International Conference on Engineering Applications of Neural Networks. Berlin: Springer, 2016: 213-226.
- [ 13 ] LÉCUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [ 14 ] SUN Y, WANG X, TANG X. Hybrid deep learning for face verification[C]// IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE Press, 2013: 1489-1496.
- [ 15 ] SAK H, SENIOR A, BEAUFAYS F. Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition[C]// 15th Annual Conference of the International Speech Communication Association, 2014: 338-342.
- [ 16 ] GHOSH S, VINAYALS O, STROPE B, et al. Contextual LSTM (CLSTM) models for large scale NLP tasks[EB/OL]. (2016-05-31) [2018-10-15]. [https://www.researchgate.net/publication/301857393\\_Contextual\\_LSTM\\_CLSTM\\_models\\_for\\_Large\\_scale\\_NLP\\_tasks](https://www.researchgate.net/publication/301857393_Contextual_LSTM_CLSTM_models_for_Large_scale_NLP_tasks).
- [ 17 ] WANG J, CAO Z W. Chinese text sentiment analysis using LSTM Network based on L2 and Nadam[C]// IEEE International Conference on Communication Technology. Piscataway, NJ: IEEE Press, 2017: 1891-1895.
- [ 18 ] JOZEFOWICZ R, ZAREMBA W, SUTSKEVER I. An empirical exploration of recurrent network architectures[C]// Proceeding of the 32nd International Conference on Machine Learning, 2015, 37: 2342-2350.
- [ 19 ] GREFF K, SRIVASTAVA R K, KOUTNÍK J, et al. LSTM: A search space odyssey[J]. IEEE Transactions on Neural Networks & Learning Systems, 2015, 28(10): 2222-2232.

#### 作者简介:

**孙宇航** 男, 硕士研究生。主要研究方向: 电磁信号智能感知方法。

**曾国奇** 男, 博士, 高级实验师, 硕士生导师。主要研究方向: 无人机遥测系统、相控阵天线。

**刘春辉** 男, 博士, 工程师。主要研究方向: 电磁信号智能感知方法。

**张多纳** 男, 博士研究生。主要研究方向: 电磁信号智能感知方法。

# SNR estimation algorithm for UAV data link based on deep learning

SUN Yuhang<sup>1</sup>, ZENG Guoqi<sup>2,\*</sup>, LIU Chunhui<sup>2</sup>, ZHANG Duona<sup>1</sup>

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

2. Institute of Unmanned System, Beihang University, Beijing 100083, China)

**Abstract:** UAV data link communication is subject to natural and artificial interferences. The signal-to-noise ratio (SNR) is an effective evaluation indicator of channel state and communication quality. In order to address insufficient SNR estimation accuracy involved in traditional estimation algorithm, an estimation model which combines convolutional neural networks (CNN) and long short term memory (LSTM) network is proposed. By means of both simulation and actual measurement, a data set of UAV communication signals is constructed with multiple SNRs, modulation modes, fading channels and other information included. In the network training phase, the sample sequence is segmented, CNN-LSTM is used to extract the deep feature of each part, and the model parameters are saved through multiple trainings. In the test phase, the constructed test set is used to verify and test the algorithm, and the SNR estimation value is obtained. Experiments show that compared with traditional SNR estimation algorithm and single-network deep learning method, the proposed algorithm can help achieve the lowest mean square error for different levels of SNR, thus achieving the high-precision estimation of SNR.

**Keywords:** UAV; signal-to-noise ratio (SNR) estimation; deep feature; convolutional neural networks (CNN); long short term memory (LSTM) network

**Received:** 2018-12-17; **Accepted:** 2019-01-23; **Published online:** 2019-03-15 09:34

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20190313.1649.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190313.1649.001.html)

**Foundation item:** National Defense Basic Research Program (JCKY2017601C006)

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

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

DOI: 10.13700/j.bh.1001-5965.2018.0766

# 基于深度卷积的残差三生网络研究与应用



厉铮泽, 杨小远\*, 朱日东, 王敬凯

(北京航空航天大学 数学与系统科学学院, 北京 100083)

**摘 要:** 针对图像多分类任务, 提出基于深度卷积的残差三生网络, 旨在通过残差学习和距离比较来训练神经网络得到有效的特征表示。首先, 设计了一个 21 层的深度卷积神经网络作为三生网络的嵌入网络, 其中该卷积网络共连接 6 个块 (block)。利用残差学习的方式, 每个 block 的输出层由卷积层的输出和该 block 的输入共同组成, 降低网络学习难度, 避免网络出现退化问题。然后, 每个 block 中采用相同拓扑结构分路的卷积层, 拓宽网络的宽度。最后, 在全连接层拼接了来自前面卷积层和 block 的输出, 加强特征信息的传递。训练前, 针对正负样本采用交叉组合的采样方法来增加有效训练样本量; 训练期间, 用样本中心点替换原点样本作为输入, 能平均降低 0.5% 错误率。在与其他三生网络的对比实验中, 在 MNIST、CIFAR10 和 SVHN 数据库上达到最好的效果, 在所有分类网络中, 本文网络在 MNIST 上达到最好的效果, 在 CIFAR10 和 SVHN 上表现优异。

**关键词:** 卷积神经网络; 三生损失; 残差学习; 挑战性样本采样; 样本中心点

**中图分类号:** O29

**文献标识码:** A

**文章编号:** 1001-5965(2019)09-1864-10

随着计算机硬件的发展和网络结构的优化, 卷积神经网络 (CNN) 已成为计算机视觉识别领域的主要学习方法。从最初的 LeNet-5<sup>[1]</sup> 到在 ImageNet 竞赛<sup>[2]</sup> 上表现优异的 AlexNet<sup>[3]</sup>、VGG-Net<sup>[4]</sup>、GoogleNet<sup>[5]</sup> 和 ResNet<sup>[6]</sup> 等, 各种性能优异的模型被提出并成功应用于计算机视觉领域, 例如图像分类<sup>[7]</sup>、目标检测<sup>[8]</sup>、目标跟踪<sup>[9]</sup> 以及视频分类<sup>[10]</sup> 等。

近年来, 学者提出三输入通道的卷积神经网络模型 (三生网络模型)<sup>[11-12]</sup>, 这种模型通过一种新的三生损失 (triplet loss) 函数进行度量学习, 现已广泛应用于多个领域, 例如人脸识别<sup>[13]</sup>、场景分类<sup>[14]</sup> 以及车辆识别<sup>[15]</sup>。在高维的嵌入空间中, 三生损失能缩短标签相同的样本的间距, 并且增大标签不同的样本的间距。与其他单输入通道

的深度卷积模型相比<sup>[3-6]</sup>, 三生网络模型<sup>[11-12]</sup> 学习到的个体级别上的细粒度特征会更为出色。三生损失与传统的归一化指数损失 (softmax loss) 相比有两点优势: 一是如果类别中只有极少数样本可用于训练, 三生损失的效果会比归一化指数损失更好<sup>[12]</sup>; 二是如果标签数量非常大, 则在批量大小相同的情况下, 与归一化指数损失相连的全连接层会占用大量的 GPU 内存, 而三生损失只占用较少的内存。

虽然三生网络的优势显而易见, 但还面临着收敛慢、难收敛等挑战。Hermans 等<sup>[15]</sup> 认为并不是所有的训练样本都有利于网络训练泛化能力, 故给出了挑战性样本 (hard samples) 的采样方法, 选择与原点样本距离最远的正样本和距离最近的负样本作为训练样本, 提高网络学习能力。在车

收稿日期: 2018-12-26; 录用日期: 2019-02-22; 网络出版时间: 2019-03-07 14:18

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190306.0921.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190306.0921.001.html)

基金项目: 国家自然科学基金 (61671002)

\* 通信作者: E-mail: xiaoyuanyang@vip.163.com

**引用格式:** 厉铮泽, 杨小远, 朱日东, 等. 基于深度卷积的残差三生网络研究与应用 [J]. 北京航空航天大学学报, 2019, 45 (9): 1864-1873. LI Z Z, YANG X Y, ZHU R D, et al. Research and application of residual triplet network based on deep convolution [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45 (9): 1864-1873 (in Chinese).

辆识别领域, Liu 等<sup>[16]</sup>认为单一网络结构不足以识别同一车型的不同车辆,提出了一个多任务分支模型来识别车辆,一个网络分支判断车辆是否是同一型号,另一个网络分支识别车辆的实际特征,综合所有情况得到结果。Cheng 等<sup>[17]</sup>将多通道网络与三生网络结合,学习分析人体部分特征,从而进行任务识别。Zhang 等<sup>[18]</sup>认为原点样本与正样本的地位相同,在同一三元组样本内,正负样本位置可以互换,意味着正样本也需要远离负样本。Chen 等<sup>[19]</sup>考虑到同类样本之间距离方差较大,提出了四生网络,在三生损失中加入了限制类内方差的约束项。

除了改进三生损失函数或者优化采样的方法,优化嵌入网络(如图1所标识的网络)也能提高整体网络性能。Simonyan 和 Zisserman<sup>[4]</sup>用小尺寸卷积核代替大尺寸卷积核,减少了参数量,提高了网络拟合能力,达到了很好的效果。Szegedy 等<sup>[5]</sup>考虑到网络层数增加会导致参数的增加,提出了 Inception 网络结构来增加网络的宽度,获取多尺度的特征信息。Huang 等<sup>[20]</sup>提出一种密集型卷积神经网络,每个卷积层的输入结合了来自前面所有卷积层的输出,鼓励重复使用特征,并且大幅地减少了参数数量。学者们普遍认为,网络层数变多,容量变大,前向网络可以表示任何函数。但是随着网络深度增加,网络训练难度增大,网络准确率也降低,网络性能无法提升。He 等<sup>[6]</sup>提出了 ResNet 来解决网络退化问题,网络层数达到152层甚至更高。

传统神经网络模型目的是将网络训练成一个映射函数,而 ResNet 将网络训练成残差函数  $F(x) = H(x) - x$ ,其中  $x$  表示输入,  $H(x)$  表示输出。相对直接学习原始特征来说,训练残差函数更加简单。目前构建深层网络的方法,一种是通过在浅层网络添加恒等映射层,就能保证网络的深度且不会发生退化,但网络性能无法提高,这与构造深层网络的目的相背。另一种是将网络训练成残差函数,这样也能满足网络深度的要求。即使某一网络层所表示的残差函数  $F(x) = 0$  (该网络层做恒等映射),保证网络准确率不会变低。

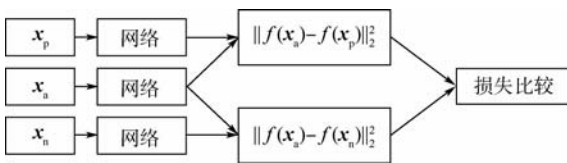


图1 三生网络

Fig. 1 Triplet network

在实际网络训练学习中,  $F(x)$  不会置为  $0$ , 网络层能够在输入信息中持续且容易地学习到新的特征。所以在残差学习的网络中,在输出  $F(x)$  中加入输入特征  $x$ , 得到  $H(x)$  作为下一网络层的输入,解决了因网络加深导致测试误差变大的问题,同时还能保证网络能够进行端对端的训练。

在三生网络中,网络的性能与很多因素相关,首先是如何使用大规模数据有效地训练三生网络模型。例如,某数据库共有  $M$  个标签,每个标签有  $N$  个样本,共会产生约  $M^2N^3$  个三元组样本,这其中包含大量无效的训练样本,这样的三元组样本会使损失函数变成  $0$ ,无法进行权重更新,对网络训练没有帮助,花费大量时间成本。所以需要先对训练样本进行筛选,得到合适的三元组样本。然后是嵌入网络的设计,提高神经网络提取特征的能力需要合理构建网络。设计新的网络拓扑结构<sup>[5-6,21-22]</sup>,适当增加网络深度,再结合激活函数 (Relu)、Dropout<sup>[3]</sup> 以及批标准化 (BN)<sup>[23]</sup> 等方法,能够减少网络参数量,加强网络泛化能力,提高网络训练速度。

本文首先介绍三生网络结构,并提出了新的三元组预处理方法;然后,介绍改进的损失函数与训练方法;最后,进行实验结果分析和总结。

## 1 三生网络结构

传统的卷积神经网络通过堆叠网络层 (conv), 构建一个深层的架构来提高效果,但网络加深后,层数过多不一定会取得更好的效果,有时反而会导致梯度消失或梯度爆炸,从而引起网络退化的问题。原因在于浅层的网络可以获得学习和更新,而深层的网络几乎没有任何变化,或者深层的网络更新调整太大,使网络无法收敛。一种解决的方法就是在浅层网络上继续添加恒等映射,能保证网络拥有足够深的结构而且不会产生比浅层网络更高的训练误差,但这样的方式并不能保证提高网络的效果。为了加深网络深度、提高网络性能,本文提出了一种新的基于残差学习的三生网络结构。

在残差学习中,  $H(x)$  是将每个块 (block) 的输入  $x$  映射成下一个 block 的输入的函数,训练 block 去学习  $H(x)$  和  $x$  的残差,记为  $F(x) = H(x) - x$ 。block 可以明确地表示残差函数  $F(x)$ ,前一个 block 输出的特征  $F(x)$  与输入的特征  $x$  相加后进入到下一个 block 中作为新的输入。 $F(x)$  和  $H(x)$  两种函数都能逐渐逼近想要学习的函数,但学习的难易程度有所不同,而且随着

网络加深,  $H(x)$  容易出现梯度弥散或爆炸的情况。残差学习的目的是忽略主体, 突显微小的变化, 使网络对输出的变化更敏感。例如: 设输入为  $x = P$ , 输出为  $H(x) = Q$ , 则残差为

$$F(x) = H(x) - x = Q - P \quad (1)$$

当输出  $H(x)$  的变化量为  $\epsilon$  时,  $H(x)$  的变化率为  $\frac{\epsilon}{Q}$ , 同时  $F(x)$  的变化量也是  $\epsilon$ ,  $F(x)$  的变化率为  $\frac{\epsilon}{Q-P}$ 。当  $Q - P$  趋于零时,  $F(x)$  的变化率

将大大高于  $H(x)$  的变化率。在网络中输出的变化大多都是细微的, 这就能保证  $Q - P$  足够小, 能够突出微小的变化, 因此残差学习对调整网络权重的效果更好。

本文设计构建了一个 21 层的残差学习网络作为嵌入网络, 结构如图 2 所示。首先将图像  $x$  通过如图 2(c) 所示的深度卷积神经网络, 经过第 1 个卷积层的卷积操作后, 输出特征进入到 stage 1, stage 1 里包含 3 个具有相同拓扑结构的 block (见图 2(b))。对于由几个卷积层堆叠而成的 block, 设输入信息为  $x'$ , 通过 block 之后得到  $F(x')$ , 为了能够让 block 学习到残差的变化, 以一种短路连接 (shortcut) 的方法, 在 block 的输出  $F(x')$  中加上输入特征信息  $x'$ , 由此实现了残差学习的结构。若特征  $F(x')$  与  $x'$  的维度是匹配的, 则直接相加得到  $y = F(x') + x'$ , 若维度不匹配, 应采用如下计算方式:

$$y = F(x') + Wx' \quad (2)$$

式中:  $W$  为  $1 \times 1$  的卷积操作, 目的是将  $x'$  的维度调整至与  $F(x')$  相同。

如图 2(a) 所示, 本文网络中每一个 block 由

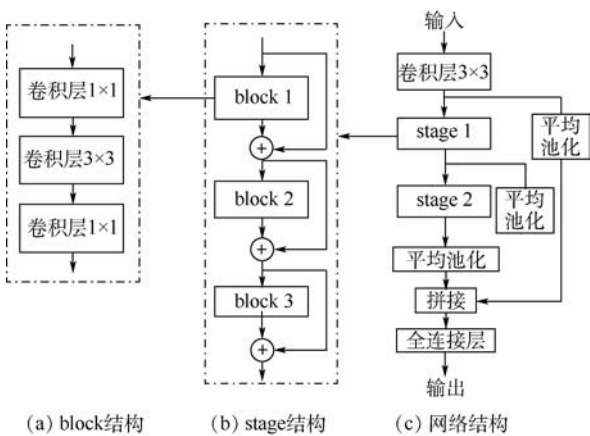


图 2 嵌入网络结构

Fig. 2 Embedded network structure

3 个卷积层组成, 因为  $1 \times 1$  的卷积层可以实现跨通道的信息交互和整合, 加强网络模型的表达能力, 所以用第 1 个和第 3 个  $1 \times 1$  的卷积层进行特征信息通道数的维度处理, 先提高维度, 再降低维度, 降低了计算量。第 2 个卷积层的卷积效果等价于多个并行卷积的卷积层的卷积效果, 在 ResNeXt<sup>[22]</sup> 中提到图 3 的 2 种模型是相互等价的。因此本文借助于“分离-变换-合并”的思想, 拓宽了网络的宽度, 使得特征信息通过升维处理后, 进入到各个拓扑结构相同的路径, 之后进行聚合变换, 最后再通过降维处理进入下一个 block。

本文将第 1 个  $3 \times 3$  的卷积层输出的特征与后 2 个区输出的特征进行拼接, 目的是在全连接层中, 最大程度地保留前面主要层的输出信息, 将信息特征映射到  $d$  维欧氏空间中。本文网络在每一个卷积层后添加批标准化层, 把卷积后的每个输出进行标准化, 来稳定网络学习, 提高训练速度, 在第一个  $3 \times 3$  的卷积层以及每个 block 中前 2 个卷积层后添加激活函数层, 防止网络出现梯度消失问题, 使网络保持一个稳定收敛的状态。网络配置如表 1 所示, 表中  $s$  为卷积步长,  $p$  为边界填充量。

经典的深度卷积网络<sup>[3-6]</sup>为了解决多分类问题( $n$  个标签), 固定设置  $n$  个输出节点, 利用归一化指数损失将神经网络输出层每个节点值变成一个概率值, 测试图像  $x$  通过前向传播网络, 在最后输出层中得到属于各个标签的概率。三生损失函数的目的是让属于同一标签的图像在  $d$  维欧氏空间上尽可能地靠近, 属于不同标签的图像尽可能地远离, 所以在几何层面上, 三生网络的输出结果可以理解成投影到  $d$  维欧氏空间上的点, 各类标签都有各类的聚集区域。标签相同的点的聚集程度高, 则说明网络表现优异。

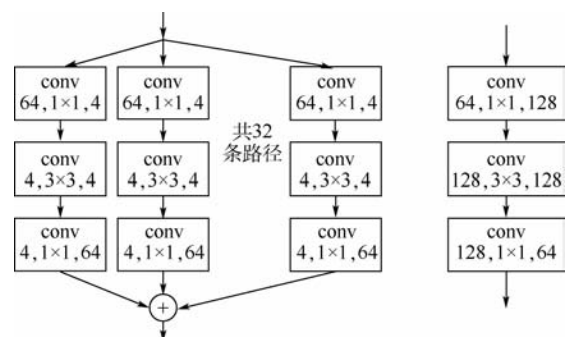


图 3 ResNeXt<sup>[22]</sup> 的等价网络结构

Fig. 3 Equivalent network structure of ResNeXt<sup>[22]</sup>

表1 网络配置

Table 1 Network configuration

区 块	网络层	通道数变化	步长/填充量	池化层
stage1	conv3	3-64	$s=1, p=1$	avgpool
	conv1	64-512	$s=1, p=0$	
	block1	conv3	512-512	$s=2, p=1$
		conv1	512-256	$s=1, p=0$
	conv1	256-512	$s=1, p=0$	
		conv3	512-512	$s=1, p=1$
	block2	conv1	512-256	$s=1, p=0$
		conv1	256-512	$s=1, p=0$
	block3	conv3	512-512	$s=1, p=1$
		conv1	512-256	$s=1, p=0$
stage2	conv1	256-1 024	$s=1, p=0$	
	block1	conv3	1 024-1 024	$s=2, p=1$
		conv1	1 024-512	$s=1, p=0$
	conv1	512-1 024	$s=1, p=0$	
		conv3	1 024-1 024	$s=1, p=1$
	block2	conv1	1 024-512	$s=1, p=0$
		conv1	512-1 024	$s=1, p=0$
	block3	conv3	1 024-1 024	$s=1, p=1$
		conv1	1 024-512	$s=1, p=0$
	fc-32			

注:以  $32 \times 32$  RGB 图像为例。

## 2 基于三生网络的三元组处理

在一个标准的三生网络中(如图1所示),输入的3个样本被称为三元组,记作  $\{x_a, x_p, x_n\}$ ,其中  $x_a$  和  $x_p$  分别为原点样本(anchor)和正样本(positive),二者具有相同的标签; $x_n$  为负样本(negative),标签与原点样本不同。针对每一个即将输入到网络中的批量,利用上一个批量训练后的网络,获得当前批量的样本  $x$  的网络特征表示  $f(x)$ ,若是初始训练的第1个批量,则利用初始网络来获得  $f(x)$ ,将  $f(x)$  进行  $d$  维  $L_2$  归一化处理<sup>[24]</sup>,满足  $\|f(x)\|_2 = 1$ 。三元组通过训练好的网络,其特征表示应满足如下距离不等式:

$$\|f(x_a) - f(x_p)\|_2^2 + \alpha < \|f(x_a) - f(x_n)\|_2^2 \quad (3)$$

式中: $\alpha$  为间隔常数,表示同类样本间距的最大值与异类样本间距的最小值的差值。若  $\alpha$  的值较大,则网络难以收敛;若  $\alpha$  的值较小,则网络学习效果较弱,无法区分不同标签样本,本文设定  $\alpha$  的值为0.5。

令  $B$  表示批量大小,三生损失函数为

$$L = \sum_{i=1}^B [\|f(x_a^i) - f(x_p^i)\|_2^2 - \|f(x_a^i) - f(x_n^i)\|_2^2 + \alpha]_+ \quad (4)$$

式中: $[\cdot]_+$  表示取正。

显而易见,在所有的三元组训练样本中,存在大量满足式(3)的三元组,称为简单三元组,这些三元组对损失函数没有贡献,因为无效的训练样本导致网络收敛速度变慢,所以需要选择能够提升网络训练的三元组。下面将讨论三元组选择的方法。

为了防止简单三元组对三生网络训练阶段产生干扰,导致网络收敛缓慢,常用的方法是在线挑战性负样本挖掘(OHNM)<sup>[12,15]</sup>,旨在选择对训练最有挑战的三元组,抛弃对训练无意义的三元组,避免简单三元组经过网络,节约训练时间。本文提出一种改进的三元组预处理方法,针对大小为  $B$  的一个批量,先将批量中的所有训练样本  $x$  通过网络后得到  $B$  个  $d$  维的网络特征表示,定义欧氏距离,计算  $B$  个网络特征表示相互之间的距离,再将所有样本  $x$  均设为原点样  $x_a$ ,根据三元组样本的构成,依照原点样本的标签找到与其距离最远的2个正样本  $x_p$  和距离最近的2个负样  $x_n$ ,如图4所示进行交叉组合,由此可以在线产生  $4B$  个三元组。传统的 OHNM 法只产生  $B$  个三元组,而本文改进的方法可以产生数量更多且有效的三元组。本文通过实验,测试选择不同数量的正负样本数量,对比训练时间以及测试效果。由于  $x_n$  数量必定远远多于  $x_p$  数量,故可以找到距离最近的  $K(K > 2)$  个负样本  $x_n$ ,从而扩充有效训练样本量。这种三元组预处理方法可以选择有挑战性的样本使网络加速收敛,且对各类别样本较少的数据集,可以大幅增加有效训练样本,提高网络训练效果。

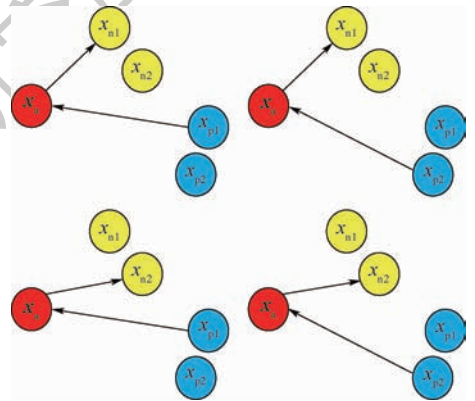


图4 三元组选择

Fig. 4 Triplet selection

## 3 基于三生网络的损失函数

三生网络是一种通过三生损失函数进行度量学习的模型,使相同标签样本更接近,同时隔开异

类标签样本,因此同一标签中的样本可以彼此在  $d$  维欧氏空间靠得更近。损失函数的设计决定了其如何指导网络进行训练以及网络性能。经过多次迭代训练后的三生网络,能够将相同标签的样本聚集在  $d$  维欧氏空间的某个公共中心点周围,即相同标签的样本应该聚集成簇,而不同标签的样本应该保持相对较远的距离,所以各个标签的样本簇应该有其样本中心点(如图5所示)。

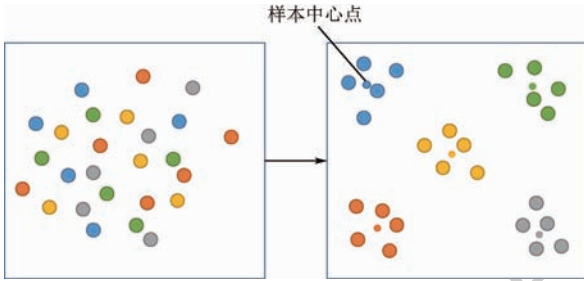


图5 二维欧氏空间样本中心点示例

Fig. 5 2D European space sample center point example

为了能够进一步促进同类样本聚集且异类样本互斥,考虑在第  $i$  次迭代之后,用  $\{c_p, x_p, x_n\}$  代替  $\{x_a, x_p, x_n\}$  作为三元组输入,  $c_p$  表示某一标签的样本中心点,即

$$c_p = \frac{1}{N_p} \sum_{m=1}^{N_p} f(x_p^m)$$

式中:  $N_p$  为该标签的样本数量,则新的三元组应满足如下距离不等式:

$$\|f(c_p) - f(x_p)\|_2^2 + \alpha < \|f(c_p) - f(x_n)\|_2^2$$

记第  $i$  次迭代之后的三生损失函数为中心三生损失函数为

$$L = \sum_{i=1}^B [\|f(c_p^i) - f(x_p^i)\|_2^2 - \|f(c_p^i) - f(x_n^i)\|_2^2 + \alpha]$$

这2个不同训练阶段使用的损失函数有着不同的意义,三生损失选择每一个样本作为  $x_a$  并搭配  $x_p$  和  $x_n$  作为输入的三元组,能在训练初始阶段较快地处理位置随机的样本,将相同标签的样本聚集在一起;中心三生损失只选择样本中心点  $c_p$  作为原点样本,使聚在一起的同类样本更加靠近  $c_p$ ,并且排斥异类样本。如果直接使用中心三生损失函数训练网络,得到的特征表示  $f(x)$  位置散乱无规律,获得的中心点  $c_p$  不具有代表性,会影响网络训练速度,导致收敛缓慢。

## 4 实验结果与分析

本文实验运行环境: Intel Xeon 2650v4 CPU 处理器, Tesla P100 GPU 加速卡。本文初始网络

使用 Kaiming 初始化方法,所有的训练均由随机梯度下降法(SGD)完成,固定动量值设定为 0.9,权重衰减设定为 0.0005。网络采用学习率逐步衰减的方式,初始学习率设定为 0.01,本文网络在迭代 50 次、150 次、225 次后,自动将学习率除以 10 之后继续训练,降低训练损失。本文没有对数据集使用任何数据增强方法。通过网络获得的空间表示向量维数  $d=32$ ,当  $d$  值过低时,全连接层降维幅度过大,容易导致图像信息丢失,影响分类效果;当  $d$  值过高时,使得网络参数变多,增加训练和测试时间成本,空间表示向量经过归一化以后,其每一个分量的值变得很小,向量之间的区别度降低,会影响分类效果。此外维度太高会使得二维可视化效果不理想。网络训练结束后,使用嵌入网络对全体训练集提取  $d$  维特征之后,计算各类标签的  $d$  维样本中心点,再将测试集中的图像通过嵌入网络得到  $d$  维向量,计算与各类样本中心点的欧氏距离,距离哪个标签的中心点最近,则测试图像标记成该标签。该验证方法与归一化指数损失的方法不同,前者以距离进行分类预测,后者以概率大小确定预测标签。本文网络与 Stochastic Pooling<sup>[25]</sup>、Maxout Network<sup>[26]</sup>、ResNet<sup>[6]</sup>、NIN<sup>[27]</sup>、DSN<sup>[28]</sup>、CMC<sup>[29]</sup>、MLDNN<sup>[30]</sup>、DenseNet<sup>[20]</sup>、ResNeXt<sup>[22]</sup>、AlexNet<sup>[3]</sup>、TripletNet<sup>[31]</sup> 在 CIFAR10<sup>[32]</sup>、MNIST<sup>[4]</sup> 和 SVHN<sup>[33]</sup> 3 个数据集上比较错误率和计算复杂度。

CIFAR10 数据集共有 10 个标签,包含 50 000 张尺寸大小为  $32 \times 32$  的彩色图像作为训练集和 10 000 张彩色图像作为验证集。图 6 是本文网络与目前性能优异的网络 DenseNet-BC<sup>[20]</sup> 和 ResNeXt-29<sup>[22]</sup> 的训练测试图,网络共迭代 310 次,每 5 次测

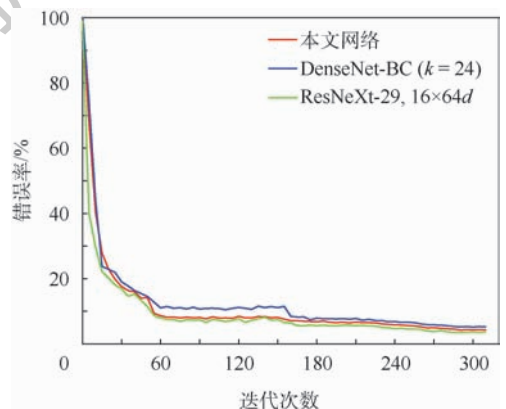


图6 本文网络与 DenseNet-BC 和 ResNeXt-29 的 CIFAR10 测试曲线

Fig. 6 CIFAR10 testing curves of proposed network, DenseNet-BC and ResNeXt-29

试错误率。ResNeXt 选用 29 层网络层,  $16 \times 64d$  的结构达到了 3.58% 的错误率,这是 CIFAR10 在无数据增强情况下最好的测试识别效果(如表 2 所示)。DenseNet<sup>[20]</sup> 在同时调用 Bottleneck 和转化层的情况下(即为 DenseNet-BC)达到了 5.19% 的错误率。本文网络因为在迭代过程中设置自动下调学习率,所以在图 6 中可以看到迭代 50 次后错误率有一次明显的下降。如第 3 节所述,网络训练后期用  $\{c_p, x_p, x_n\}$  作为三元组输入,目的是让同类标签样本与样本中心点更靠近。在 250 次迭代训练之后更改样本中心点作为三元组输入,之后每 5 次迭代训练,重新更新一次各标签样本中心点  $c_p$ ,所以网络测试错误率能持续降低。本文实验发现,训练后期更改损失函数比不做任何变化情况下,网络在数据集上的错误率能够平均降低 0.5%,证明了这种方法是有效的。本文网络与其他网络在 CIFAR10 数据集的比较如表 2 所示,ResNeXt-29 系列达到了 3.58% 的效果,本文网络位列第二,达到了 4.28%,DenseNet 两个不同版本分别达到了 5.19% 和 5.83%,ResNet 达到了 6.43%,CMC 达到了 6.87%,MLDNN 达到了 8.12%,NIN 达到了 8.8%,DSN 达到了 9.69%,AlexNet 达到了 11%,Maxout Network 达到了 11.68%,Stochastic Pooling 达到了 15.13%。值得一提的是,作为三生网络系列中的 TripletNet 达到了 12.90%,本文三生网络错误率比其低了 8.62%,因为在网络结构方面,设计了较深的残差学习网络作为嵌入网络,使得三生网络具有很好的特征表示能力;在训练过程中使用正负样本交叉组合采样方法和改变训练的方式,让本文网络在一些数据集上有明显的优势。如表 3 所示,在网络学习能力方面,除了 ResNet 需要迭代

表 2 CIFAR10 上的错误率  
Table 2 Error rates on CIFAR10

网络	错误率/%
Stochastic Pooling <sup>[25]</sup>	15.13
TripletNet <sup>[31]</sup>	12.90
Maxout Network <sup>[26]</sup>	11.68
AlexNet <sup>[3]</sup>	11
DSN <sup>[28]</sup>	9.69
NIN <sup>[27]</sup>	8.8
MLDNN <sup>[30]</sup>	8.12
CMC <sup>[29]</sup>	6.87
ResNet <sup>[6]</sup>	6.43
DenseNet <sup>[20]</sup>	5.83
DenseNet-BC <sup>[20]</sup>	5.19
本文网络	<b>4.28</b>
ResNeXt-29 <sup>[22]</sup>	<b>3.58</b>

60 000 次,其余网络均在几百个周期内完成训练,而且本文网络在网络参数较多的情况下,以较少的周期快速完成训练,说明网络的学习能力较强。本文网络结构具有较大规模,所以需要减少网络的计算量,降低计算复杂度。以表 1 里 stage 1 中的 block 1 为例,第一个卷积层和第 3 个卷积层的  $1 \times 1$  卷积核代替了  $3 \times 3$  卷积核,保证其具备升维和降维功能的前提下,网络的计算量由  $7.5 \times 10^9$  次降低至  $4.0 \times 10^9$  次,在时间上减少约 47%。

MNIST 是手写灰度图像数据集,由 60 000 张大小为  $28 \times 28$ ,标签从 0 到 9 的训练图像和 10 000 张测试图像组成。比较结果如表 4 所示,可以观察到本文网络要比其他的方法都要好,错误率达到了 0.22%,比第 2 名的 CMC 要低 0.11%,比同属于三生网络的 TripletNet 要低 0.16%。

SVHN 数据库与 MNIST 的相似,由 600 000 个从 Google 街景图像中获得  $32 \times 32$  彩色门牌图像,标签从 0 到 9。比较结果如表 5 所示,最好的结果均来自 DenseNet,分别达到了 1.59% 和 1.74% 的错误率,第 3 名 CMC 达到了 1.76%,本文网络仅比 CMC 低 0.05%,达到了 1.81%,MLDNN 和 DSN 均达到了 1.92%,NIN 达到了 2.35%,Maxout Network 达到了 2.67%,Stochastic Pooling

表 3 网络训练迭代次数 (CIFAR10)

Table 3 Iteration times of network training (CIFAR10)

网络	迭代次数	错误率/%
ResNet <sup>[6]</sup>	60 000	6.43
Maxout Network <sup>[26]</sup>	700	11.68
Stochastic Pooling <sup>[25]</sup>	500	15.13
本文网络	310	<b>4.28</b>
MLDNN <sup>[30]</sup>	300	8.12
DenseNet-BC <sup>[20]</sup>	300	5.19
DenseNet <sup>[20]</sup>	300	5.83
ResNeXt-29 <sup>[22]</sup>	300	<b>3.58</b>
NIN <sup>[27]</sup>	200	8.8
DSN <sup>[28]</sup>	110	9.69

表 4 MNIST 上的错误率

Table 4 Error rates on MNIST

网络	错误率/%
Stochastic Pooling <sup>[25]</sup>	0.47
Maxout Network <sup>[26]</sup>	0.47
NIN <sup>[27]</sup>	0.47
MLDNN <sup>[30]</sup>	0.42
DSN <sup>[28]</sup>	0.39
TripletNet <sup>[31]</sup>	0.38
CMC <sup>[29]</sup>	<b>0.33</b>
本文网络	<b>0.22</b>

达到了 2.80%。三生网络 TripletNet 仅达到了 4.63%，本文网络错误率比其低了 2.82%。

如表 6 所示,本文测试 4 种正负样本组合情况,发现增加有效三元组训练样本,能较为明显地提高网络训练效果,降低错误率。CIFAR10 数据集通过 4 种样本组合方法,有效的三元组训练样本数量分别为  $5 \times 10^4$ 、 $1 \times 10^5$ 、 $2 \times 10^5$ 、 $3 \times 10^5$ ,网络训练迭代一次所需的时间如表 6 所示,近似等比例增加。本文网络的批量  $B = 50$ ,所以正样本的数量不能过多,避免产生无效的三元组样本,根据 GPU 的性能,可以适当调节负样本的数量。所以本文三元组预处理方法对样本较少的数据集能增加训练样本以提高网络训练效果。

三生损失函数指导网络训练的方式决定了其具有样本聚类效果,所以在验证数据集时,网络能够将图像映射入  $d$  维欧氏空间。本节所示内容已从定量分析角度分析了本文网络的分类效果,下面从定性分析角度认识三生网络的分类效果。

$t$  分布随机邻域嵌入 (T-SEN) 是一种较为高级的非线性降维算法,基于在领域图上随机游走的概率分布来解释数据内的结构,可以将  $d$  维数据映射到 2 维欧氏空间上实现可视化。本文使用 TSEN 将  $d$  维特征输出  $f(\mathbf{x})$  投影到可视化的二维欧氏空间中,可以通过观察测试集样本二维的分类效果。

图 7 为本文网络在 CIFAR10 数据集上的分

表 5 SVHN 上的错误率

Table 5 Error rates on SVHN

网络	错误率/%
TripletNet <sup>[31]</sup>	4.63
Stochastic Pooling <sup>[25]</sup>	2.80
Maxout Network <sup>[26]</sup>	2.67
NIN <sup>[27]</sup>	2.35
DSN <sup>[28]</sup>	1.92
MLDNN <sup>[30]</sup>	1.92
本文网络	1.81
CMC <sup>[29]</sup>	1.76
DenseNet-BC <sup>[20]</sup>	1.74
DenseNet <sup>[20]</sup>	1.59

表 6 样本组合的实验对比 (CIFAR10)

Table 6 Experimental comparison of sample combination (CIFAR10)

正负样本组合	训练时间/s	错误率/%
1 正 1 负	230	4.74
1 正 2 负	410	4.52
2 正 2 负	750	4.28
2 正 3 负	1 190	4.29

类效果,因为本文网络在 CIFAR10 数据集有 4.28% 的错误率,所以从图中可以看到,仅有少量样本还无法区分,但已经有明显的整体分布,考虑到数据从  $d$  维降至二维,必然会产生一些重叠以及误差,因此从测试样本聚类的角度认为网络分类效果是比较优异的。图 8 为本文网络在 MNIST 数据集上的实验结果, MNIST 是手写数字灰度图像,且各标签图像差别明显,所以比较容易训练和测试。在图中聚类效果十分明显,只有极少数的样本被错误分类。图 9 为本文网络在 SVHN 数据集上的实验结果, SVHN 是三维彩色数字图像,训练难度比 MNIST 稍高一点,但训练与测试结果比较优异,在二维视图中聚类程度高。



图 7 CIFAR10-二维特征表示

Fig. 7 CIFAR10-2D feature representation



图 8 MNIST-二维特征表示

Fig. 8 MNIST-2D feature representation



图 9 SVHN-二维特征表示

Fig. 9 SVHN-2D feature representation

## 5 结 论

1) 本文设计了一个21层的卷积神经网络作为三生网络的嵌入网络。该卷积神经网络连接2个stage,共6个block,利用残差学习的方式,在每个block输出 $f(\mathbf{x})$ 的基础上加入 $\mathbf{x}$ 防止网络出现退化问题,每个block中采用相同拓扑结构分路的Inception网络结构,最后在全连接层拼接来自第1个卷积层和2个区的输出。

2) 本文提出一个三元组预处理方法,选择与原点样本距离最远的2个正样本和距离最近的2个负样本组成4个三元组作为训练样本,这种正负样本交叉组合可以增加有效的困难训练样本,降低网络训练难度。

3) 在网络训练中,用样本中心点 $\mathbf{c}_p$ 替换原点样本 $\mathbf{x}_a$ 可以让同类样本更加聚集。最后网络在数据集上的训练测试结果显示本文网络对比其他三生网络模型有明显优异的表现,也优于绝大多数基于归一化指数损失的神经网络。

接下来,将利用本文提出的三生网络用于人脸识别、车辆识别等细粒度图像分类的任务。

### 参考文献 (References)

- [1] 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.
- [2] RUSSAKOVSKY O, DENG J, SU H, et al. ImageNet large scale visual recognition challenge [J]. International Journal of Computer Vision, 2014, 115(3): 211-252.
- [3] KRIZHEVSKY A, SUTSKEVER I, HINTON G. Imagenet classification with deep convolutional neural networks [C] // International Conference on Neural Information Processing Systems. New York: Curran Associates Inc., 2012: 1097-1105.
- [4] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C] // International Conference on Learning Representations, 2015: 1-14.
- [5] SZEGEDY C, LIU W, JIA Y. Going deeper with convolutions [C] // IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2015: 1-9.
- [6] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C] // IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2016: 770-778.
- [7] SZEGEDY C, VINCENT V, IOFFE S. Rethinking the inception architecture for computer vision [C] // IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2016: 2818-2826.
- [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2014: 580-587.
- [9] WANG N, YEUNG D Y. Learning a deep compact image representation for visual tracking [C] // International Conference on Neural Information Processing Systems. New York: Curran Associates Inc., 2013: 809-817.
- [10] KARPATHY A, TODERICI G, SHETTY S, et al. Large-scale video classification with convolutional neural networks [C] // IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2014: 1725-1732.
- [11] KANG B N, KIM K Y, KIM D J. Deep convolutional neural network using triplets of faces, deep ensemble, and score-level fusion for face recognition [C] // IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2017: 611-618.
- [12] WANG C, LAN X P, ZHANG X. How to train triplet networks with 100K identities? [C] // IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2017: 1907-1915.
- [13] SCHROFF F, KALENICHENKO D, PHILBIN J. Facenet: A unified embedding for face recognition and clustering [C] // IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2015: 815-823.
- [14] LIU Y S, HUANG C. Scene classification via triplet networks [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(1): 220-237.
- [15] HERMANS A, BEYER L, LEIBE B. In defense of the triplet loss for person re-identification [EB/OL]. (2017-11-21) [2018-12-01]. <https://arxiv.org/pdf/1703.07737.pdf>.
- [16] LIU H, TIAN Y, WANG Y, et al. Deep relative distance learning tell the difference between similar vehicles [C] // IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2016: 2167-2175.
- [17] CHENG D, GONG Y H, ZHOU S P, et al. Person re-identification by multi-channel parts-based CNN with improved triplet loss function [C] // IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2016: 1335-1344.
- [18] ZHANG S, GONG Y, WANG J. Deep metric learning with improved triplet loss for face clustering in video [C] // Pacific-rim Conference on Advances in Multimedia Information Processing. Berlin: Springer, 2016: 497-508.
- [19] CHEN W, CHEN X, ZHANG J, et al. Beyond triplet loss: A deep quadruplet network for person re-identification [C] // IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2017: 1320-1329.
- [20] HUANG G, LIU Z, MAATEN L, et al. Densely connected convolutional networks [C] // IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2017: 2261-2269.
- [21] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning [C] // AAAI Conference on Artificial Intelligence. Palo Alto, CA: AAAI Press, 2017: 4278-4284.
- [22] XIE S, GIRSHICK R, DOLLAR P, et al. Aggregated residual transformations for deep neural networks [C] // IEEE Confer-

- ence on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE Press, 2017: 5987-5995.
- [23] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [C] // International Conference on Machine Learning. Boston: MIT Press, 2015: 448-456.
- [24] DING S, LIN L, WANG G, et al. Deep feature learning with relative distance comparison for person re-identification [J]. Pattern Recognition, 2015, 48(10): 2993-3003.
- [25] ZEILER M D, FERGUS R. Stochastic pooling for regularization of deep convolutional neural networks [EB/OL]. (2013-01-16) [2018-11-25]. <https://arxiv.org/pdf/1301.3557.pdf>.
- [26] GOODFELLOW I J, WARDE-FARLEY D, MIRZA M, et al. Maxout networks [C] // Proceedings of the International Conference on Machine Learning. Boston: MIT Press, 2013: 1319-1327.
- [27] LIN M, CHEN Q, YAN S. Network in network [C] // International Conference on Learning Representations, 2014: 1-10.
- [28] LEE C Y, XIE S N, GALLAGHER P W, et al. Deeply-supervised nets [C] // Proceedings of the International Conference on Artificial Intelligence and Statistics. San Diego, California: PMLR, 2015: 562-570.
- [29] LIAO Z B, CARNEIRO G. Competitive multi-scale convolution [EB/OL]. (2015-11-18) [2018-11-10]. <https://arxiv.org/pdf/1511.05635.pdf>.
- [30] XU C Y, LU C Y, LIANG X D, et al. Multi-loss regularized deep neural network [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2016, 26(12): 2273-2283.
- [31] HOFFER E, AILON N. Deep metric learning using triplet network [C] // International Workshop on Similarity-based Pattern Recognition. Berlin: Springer, 2015: 84-92.
- [32] KRIZHEVSKY A, HINTON G. Learning multiple layers of features from tiny images [D]. Toronto: University of Toronto, 2009: 32-35.
- [33] NETZER Y, WANG T, COATES A, et al. Reading digits in natural images with unsupervised feature learning [C] // NIPS Workshop on Deep Learning and Unsupervised Feature Learning. New York: Curran Associates Inc., 2011: 1-9.

#### 作者简介:

厉铮泽 男, 硕士研究生。主要研究方向: 深度学习与应用。

杨小远 女, 博士, 教授, 博士生导师。主要研究方向: 深度学习与图像处理。

朱日东 男, 博士研究生。主要研究方向: 深度学习与目标跟踪。

王敬凯 男, 博士研究生。主要研究方向: 深度学习与图像融合。

## Research and application of residual triplet network based on deep convolution

LI Zhengze, YANG Xiaoyuan\*, ZHU Ridong, WANG Jingkai

(School of Mathematics and Systems Science, Beihang University, Beijing 100083, China)

**Abstract:** For multi-classification image tasks, a residual triplet network based on deep convolution is proposed, which aims to train neural networks to obtain useful feature representations through residual learning and distance comparison. Firstly, a 21-layer deep convolution neural network is designed as the embedded network of the triplet network, where the convolutional network is connected with 6 blocks. By using residual learning, the output of each block is combined with the input of this block and the output of the convolutional layer which focus on reducing the difficulty of network learning and avoiding degradation. Then, each block employed the convolution layers with the same topological branch to broaden the width of the network. Finally, to enhance the transfer of feature information, the fully-connected layer concatenated the output of the previous convolutional layers and blocks. Before training, the cross-combined sampling method is used to increase effective samples for hard samples. During training, using the sample center point to replace the anchor sample as an input can reduce the error rate by 0.5% on average. Among the triplet network series, we achieved the best results on the MNIST, CIFAR10, and SVHN. In all classification networks, we achieved the best results on the MNIST and performed well on CIFAR10 and SVHN.

**Keywords:** convolution neural network; triplet loss; residual learning; hard sample mining; sample center point

**Received:** 2018-12-26; **Accepted:** 2019-02-22; **Published online:** 2019-03-07 14:18

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20190306.0921.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190306.0921.001.html)

**Foundation item:** National Natural Science Foundation of China (61671002)

\* **Corresponding author.** E-mail: xiaoyuanyang@vip.163.com

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

DOI: 10.13700/j.bh.1001-5965.2018.0786

# 一种考虑 GPS 信号中断的导航滤波算法

何康辉<sup>1</sup>, 董朝阳<sup>1,\*</sup>, 王青<sup>2</sup>

(1. 北京航空航天大学 航空科学与工程学院, 北京 100083;

2. 北京航空航天大学 自动化科学与电气工程学院, 北京 100083)



**摘 要:** 针对无人机惯性 (INS)/GPS 组合导航系统, 考虑导航过程中存在的 GPS 数据中断的问题, 设计了一种改进的滤波算法。首先建立了无人机导航运动学模型, 再将传统的扩展卡尔曼滤波 (EKF) 技术和强跟踪滤波结合, 利用模糊理论中的隶属度函数设计了一种模糊强跟踪扩展卡尔曼滤波 (STEKF) 算法。仿真结果表明, 所设计的改进算法能够快速适应 GPS 信号突变, 即当 GPS 信号从故障状态恢复到正常状态时, 改进算法相较普通 EKF 算法能更快速地收敛到稳定状态, 重新完成对飞行状态的估计。同时相较普通 EKF 和强跟踪扩展卡尔曼滤波算法, 改进算法具有更高的滤波精度。

**关键词:** 扩展卡尔曼滤波 (EKF); 强跟踪滤波; 组合导航; 全球定位系统 (GPS); 隶属度函数

**中图分类号:** TP242.6

**文献标识码:** A      **文章编号:** 1001-5965(2019)09-1874-08

针对无人机导航, 常采用全球定位系统 (GPS) 和惯性导航 (陀螺仪和加速度计) 相结合的组合导航方式。在正常工作情况下, 惯性导航根据当前时刻的位置速度状态结合牛顿运动学定律推算下一时刻的无人机状态, GPS 实时的测量无人机当前位置从而修正无人机状态。而 GPS 信号本身容易被干扰, 在遇到极端天气或在森林地区时 GPS 信号无法接受。惯性导航本身虽然不容易出现故障, 但随时间会有严重的累计误差。为了解决 GPS 和惯导技术各自的缺陷, 使用多传感器的组合导航技术就顺势而生了<sup>[1-4]</sup>。

无人机导航的状态估计主要使用的是卡尔曼滤波算法<sup>[5]</sup>。由于无人机的导航系统是一个非线性系统, 需要借助非线性滤波算法。目前常用的方法有扩展卡尔曼滤波<sup>[6]</sup> (Extended Kalman Filter, EKF)、无迹卡尔曼滤波 (Unscented Kalman Filter, UKF) 以及容积卡尔曼滤波 (Cubature Kal-

man Filter, CKF)<sup>[7]</sup>。其中, EKF 是对非线性函数泰勒展开保留一阶项, 将其近似为线性函数, 再代入标准的卡尔曼滤波算法中, 进而完成对非线性系统状态的近似估计。因其具有计算量小、精度高<sup>[8]</sup>、适应性强的特点, EKF 广泛应用于导航系统中<sup>[9-12]</sup>。已有许多学者建立了针对导航的 UKF 和 CKF 算法, 并对其作出了改进<sup>[13-16]</sup>。再例如 Marina<sup>[17]</sup> 将 UKF 应用于无人飞行器导航并得到了优于 EKF 的滤波效果。尽管这些方法对强非线性系统的估计精度高于 EKF, 但随之带来的是计算量的显著增加<sup>[18]</sup>, 而且在高维系统中不稳定, 极易出现滤波发散等问题<sup>[19]</sup>, 这在要求快速实时和稳定的导航系统中显得得不偿失。因此, 实际的飞行器导航系统中, 多使用 EKF 进行数据融合和状态估计<sup>[20-22]</sup>。

在无人飞行器导航实际过程中, GPS 信号不是时刻保持有效的<sup>[23]</sup>。例如当无人飞行器驶入

收稿日期: 2019-01-02; 录用日期: 2019-04-05; 网络出版时间: 2019-04-30 16:05

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190429.1402.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190429.1402.001.html)

基金项目: 国家自然科学基金 (61873295)

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

**引用格式:** 何康辉, 董朝阳, 王青. 一种考虑 GPS 信号中断的导航滤波算法[J]. 北京航空航天大学学报, 2019, 45(9): 1874-1881.  
HE K H, DONG C Y, WANG Q. A navigation filtering algorithm considering GPS signal interruption [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1874-1881 (in Chinese).

森林、荒漠等卫星覆盖不全的地区, GPS 信号就会出现异常。此外, GPS 的接收端也可能出现短期的故障。由于 GPS 发生临时中断, 导航滤波估计将偏离真实的飞行状态, 且当导航系统由故障状态变为正常状态时, 基本的 EKF 算法不能及时作出响应, 恢复到正常的滤波状态。也就是说, EKF 缺乏对突变状态的快速响应能力。具体的原因是: 当系统逐渐进入稳定状态时, 滤波的最佳增益矩阵将趋于最小值。在这时, 若系统发生突变, 滤波器的增益矩阵不会因为残差的增大而增大, 也就是说, 滤波器丧失了对突变状态的跟踪能力。当系统从突变状态快速趋于稳态时, 滤波器的增益矩阵也不会快速的减小。从这个意义上来说, EKF 类滤波器是一种开环滤波器。

针对上述问题, 本文将强跟踪扩展卡尔曼滤波(STEKF)算法运用到无人飞行器导航系统中, 目的是当 GPS 信号中断又恢复后, 导航系统能快速回到稳态, 重新完成对无人飞行器状态的有效估计。针对 STEKF 在导航系统中应用的弊端, 本文对原有的 STEKF 作出了改进, 引入了模糊理论的思想, 利用隶属度函数来实现 EKF 和 STEKF 的优势互补。

## 1 标准离散扩展卡尔曼滤波模型

导航系统是一个有时间间断的离散系统, 因此这里只讨论离散的扩展卡尔曼滤波流程。设随机的离散非线性系统模型为

$$\mathbf{X}_{k+1} = f[\mathbf{X}_k, k] + \mathbf{W}_k \quad (1)$$

$$\mathbf{Z}_{k+1} = h[\mathbf{X}_{k+1}, k+1] + \mathbf{V}_{k+1} \quad (2)$$

式中:  $\mathbf{W}_k$  和  $\mathbf{V}_{k+1}$  均为彼此不相关的零均值白噪声序列, 它们与初始状态不相关, 且

$$\begin{cases} E[\mathbf{W}_k \mathbf{W}_j^T] = \mathbf{Q}_k \\ E[\mathbf{V}_k \mathbf{V}_j^T] = \mathbf{R}_k \end{cases} \quad (3)$$

式中:  $\mathbf{Q}_k$  为状态模型噪声协方差阵;  $\mathbf{R}_k$  为量测模型噪声协方差阵。

在最优滤波值  $\hat{\mathbf{X}}_k$  处对  $f[\mathbf{X}_k, k]$  进行泰勒展开, 在一步预测值  $\hat{\mathbf{X}}_{k+1|k}$  处对  $h[\mathbf{X}_{k+1}, k+1]$  进行泰勒展开并取泰勒一阶项:

$$\Phi_{k+1,k} = \left. \frac{\partial f[\mathbf{X}_k, k]}{\partial \mathbf{X}_k^T} \right|_{\mathbf{X}_k = \hat{\mathbf{X}}_k} \quad (4)$$

$$\mathbf{H}_{k+1} = \left. \frac{\partial h[\mathbf{X}_{k+1}, k+1]}{\partial \mathbf{X}_{k+1}^T} \right|_{\mathbf{X}_{k+1} = \hat{\mathbf{X}}_{k+1|k}} \quad (5)$$

状态一步预测为

$$\hat{\mathbf{X}}_{k+1|k} = f[\hat{\mathbf{X}}_k, k] \quad (6)$$

一步预测均方差为

$$\mathbf{P}_{k+1|k} = \Phi_{k+1,k} \mathbf{P}_k \Phi_{k+1,k}^T + \mathbf{Q}_k \quad (7)$$

滤波增益为

$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1|k} \mathbf{H}_{k+1}^T (\mathbf{H}_{k+1} \mathbf{P}_{k+1|k} \mathbf{H}_{k+1}^T + \mathbf{R}_{k+1})^{-1} \quad (8)$$

状态估计为

$$\hat{\mathbf{X}}_{k+1} = \hat{\mathbf{X}}_{k+1|k} + \mathbf{K}_{k+1} (\mathbf{Z}_{k+1} - h[\hat{\mathbf{X}}_{k+1|k}, k+1]) \quad (9)$$

估计均方差为

$$\mathbf{P}_{k+1} = (\mathbf{I} - \mathbf{K}_{k+1} \mathbf{H}_{k+1}) \mathbf{P}_{k+1|k} (\mathbf{I} - \mathbf{K}_{k+1} \mathbf{H}_{k+1})^T + \mathbf{K}_{k+1} \mathbf{R}_{k+1} \mathbf{K}_{k+1}^T \quad (10)$$

## 2 针对导航的模糊强跟踪扩展卡尔曼滤波器

### 2.1 强跟踪滤波器

为了解决导航中存在的 GPS 信号中断的问题, 本文引入强跟踪滤波的思想, 将其与标准的扩展卡尔曼滤波结合。使系统在 GPS 中断恢复后, 能够快速进入稳态。

### 2.2 含次优渐消因子的扩展卡尔曼滤波器

为了使滤波器具有较强的跟踪性能, 一个想法是采用一个随时间变化的因子对过去的数据进行渐消, 衰减老数据对当前滤波预测的影响。这可以通过更改协方差阵进而更改增益矩阵来实现。具体做法是, 在 EKF 模型中, 对协方差预测矩阵公式:

$$\mathbf{P}_{k+1|k} = \Phi_{k+1,k} \mathbf{P}_k \Phi_{k+1,k}^T + \mathbf{Q}_k$$

将其改为

$$\mathbf{P}_{k+1|k} = \omega_{k+1} \Phi_{k+1,k} \mathbf{P}_k \Phi_{k+1,k}^T + \mathbf{Q}_k \quad (11)$$

式中:  $\omega_{k+1}$  为随时间变化的渐消因子, 下面是渐消因子的求取方法:

$$\omega_{k+1} = \begin{cases} \omega_{0,k+1} & \omega_{0,k+1} \geq 1 \\ 1 & \omega_{0,k+1} < 1 \end{cases} \quad (12)$$

$$\omega_{0,k+1} = \frac{\text{tr}(\mathbf{N}_{k+1})}{\text{tr}(\mathbf{M}_{k+1})} \quad (13)$$

$$\mathbf{N}_{k+1} = \mathbf{V}_{0,k+1} - \mathbf{H}_{k+1} \mathbf{Q}_k \mathbf{H}_{k+1}^T - \beta_0 \mathbf{R}_{k+1} \quad (14)$$

$$\mathbf{M}_{k+1} = \mathbf{H}_{k+1} \Phi_k \mathbf{P}_k \Phi_k^T \mathbf{H}_{k+1}^T \quad (15)$$

$$\mathbf{V}_{0,k+1} = \begin{cases} \mathbf{Z}_1^* \mathbf{Z}_1^{*T} & k = 0 \\ \frac{\rho \mathbf{V}_{0,k} + \mathbf{Z}_{k+1}^* \mathbf{Z}_{k+1}^{*T}}{1 + \rho} & k \geq 1 \end{cases} \quad (16)$$

式中:  $\rho$  为遗忘因子, 一般取  $\rho = 0.95$ ;  $\beta_0$  为弱化次优因子,  $\beta_0 \geq 1$ 。

$$\mathbf{Z}_{k+1}^* = \mathbf{Z}_{k+1} - \mathbf{H}_{k+1} \hat{\mathbf{X}}_{k+1|k} \quad (17)$$

### 2.3 模糊强跟踪扩展卡尔曼滤波器

具有强跟踪特性的扩展卡尔曼滤波器的确对突变的快速响应能力,但这是以牺牲一定的滤波精度为代价的。因为标准的 EKF 是在满足最小方差准则的基础上来确定最佳增益矩阵的,如式(7)。而含渐消因子的 EKF 则对式(7)中的  $\mathbf{P}_{k+1|k}$  取值做了修改,如式(11)所示,这样计算的最佳增益矩阵  $\mathbf{K}_{k+1}$  不满足卡尔曼滤波的基本原则,其精度就有所下降。应用于无人机导航中,以牺牲整体的导航精度换取局部的快速响应,显得得不偿失。

针对此,本文提出一种改进的 STEKF 算法,目的是在保证滤波精度的同时,提高系统应对快速突变的能力。

改进算法的主体思想是调整协方差预测式(11)中的渐消因子  $\omega_{k+1}$ 。在 2.2 节的 STEKF 算法中,无论系统是否发生突变,均采用固定的渐消因子。这样在系统稳定状态时(大部分时候),STEKF 具有较低的精度。而本文的改进算法里,将通过比较系统估计残差和期望的关系,来确定最终的渐消因子。

计算当前残差偏离期望残差的比值为

$$\gamma_0 = \frac{\mathbf{Z}_{k+1}^* \mathbf{Z}_{k+1}^{*T}}{\text{tr}\{E(\mathbf{Z}_{k+1}^* \mathbf{Z}_{k+1}^{*T})\}} \quad (18)$$

式中:

$$\mathbf{Z}_{k+1}^* = \mathbf{Z}_{k+1} - \mathbf{H}_{k+1} \hat{\mathbf{X}}_{k+1|k} = \mathbf{H}_{k+1} (\mathbf{X}_{k+1} - \hat{\mathbf{X}}_{k+1|k}) + \mathbf{V}_{k+1} = \mathbf{H}_{k+1} \mathbf{X}_{k+1}^* + \mathbf{V}_{k+1} \quad (19)$$

$E(\mathbf{Z}_{k+1}^* \mathbf{Z}_{k+1}^{*T})$  为残差的期望,表达式为

$$\begin{aligned} E(\mathbf{Z}_{k+1}^* \mathbf{Z}_{k+1}^{*T}) &= E[(\mathbf{H}_{k+1} \mathbf{X}_{k+1}^* + \mathbf{V}_{k+1})(\mathbf{H}_{k+1} \mathbf{X}_{k+1}^* + \mathbf{V}_{k+1})^T] = \mathbf{H}_{k+1} E(\mathbf{X}_{k+1}^* \mathbf{X}_{k+1}^{*T}) \mathbf{H}_{k+1}^T + \\ &\mathbf{H}_{k+1} E(\mathbf{X}_{k+1}^* \mathbf{V}_{k+1}^T) + \mathbf{H}_{k+1}^T E(\mathbf{V}_{k+1} \mathbf{X}_{k+1}^{*T}) + \\ &E(\mathbf{V}_{k+1} \mathbf{V}_{k+1}^T) \end{aligned} \quad (20)$$

注意到,  $E(\mathbf{X}_{k+1}^* \mathbf{V}_{k+1}^T) = E(\mathbf{V}_{k+1} \mathbf{X}_{k+1}^{*T}) = 0$ ,  $E(\mathbf{X}_{k+1}^* \mathbf{X}_{k+1}^{*T}) = \mathbf{P}_{k+1|k}$ ,  $E(\mathbf{V}_{k+1} \mathbf{V}_{k+1}^T) = \mathbf{R}_{k+1}$ 。所以有

$$\begin{aligned} E(\mathbf{Z}_{k+1}^* \mathbf{Z}_{k+1}^{*T}) &= \mathbf{H}_{k+1} \mathbf{P}_{k+1|k} \mathbf{H}_{k+1}^T + \mathbf{R}_{k+1} = \\ &\mathbf{H}_{k+1} (\Phi_{k+1,k} \mathbf{P}_k \Phi_{k+1,k}^T + \mathbf{Q}_k) \mathbf{H}_{k+1}^T + \mathbf{R}_{k+1} \end{aligned} \quad (21)$$

在求得  $\gamma_0$  后,令  $\gamma_1 = \max\{\gamma_{\min}, \gamma_0\}$ ,  $\gamma_{\min} > 1$  为临界残差参数,当  $\gamma_1 \leq \gamma_{\min}$  时认为残差较小, GPS 量测未发生中断。

设隶属度函数  $g(\gamma_1)$  为

$$g(\gamma_1) = \omega_{k+1} \exp\left(-\frac{\gamma_{\min} \ln \omega_{k+1}}{\gamma_1}\right) \quad (22)$$

式中:  $\omega_{k+1}$  为 2.2 节由标准 STEKF 算法求得的渐

消因子。

则将协方差预测公式(11)变为

$$\mathbf{P}_{k+1|k} = g(\gamma_1) \Phi_{k+1,k} \mathbf{P}_k \Phi_{k+1,k}^T + \mathbf{Q}_k \quad (23)$$

以上就是模糊强跟踪扩展卡尔曼滤波的基本模型。

直观地看,当计算出的残差与期望之比  $\gamma_0$  小于门限值  $\gamma_{\min}$  时,可以认为此时的估计残差在允许范围内,系统的量测(GPS 量测值)未发生突变,此时  $g(\gamma_1) = \omega_{k+1} \exp\left(-\frac{\gamma_{\min} \ln \omega_{k+1}}{\gamma_{\min}}\right) = 1$ ,改进算法的协方差预测式(23)即为 EKF 的协方差预测式(7)。当计算残差远大于期望残差,即量测值发生巨大突变,此时  $\gamma_1 \rightarrow +\infty$ ,  $g(\gamma_1) \rightarrow \omega_{k+1}$ , 滤波器具有强跟踪的特性。

改进算法的优点在于借助了模糊理论中的隶属度函数,对于 EKF 和 STEKF,没有完全的抛弃一种算法而采用另一种。而是将两者的优势互补。大多数情况下,当系统稳定, GPS 量测正常时,更多地借助 EKF 保证滤波精度,当 GPS 信号发生故障后,更多地借助 STEKF。

为了证明提出的模糊强跟踪扩展卡尔曼滤波能够对系统状态做出有效估计,提出如下假设和定理。

**假设 1** 系统状态模型正确,且由式(6)~式(10)构成的标准扩展卡尔曼滤波器的实际误差均方阵存在有限上界。

**定理 1** 对于由式(6)~式(10)、式(12)~式(17)和式(22)~式(23)构成的模糊强跟踪扩展卡尔曼滤波器,若上述假设成立,改进算法实际的估计误差均方阵为  $\mathbf{P}_{k+1}^r = E((\mathbf{X}_{k+1}^r - \hat{\mathbf{X}}_{k+1}) \cdot (\mathbf{X}_{k+1}^r - \hat{\mathbf{X}}_{k+1})^T)$  存在有限上界:

$$\sup(\mathbf{P}_{k+1}^r) = \sup(\mathbf{P}_{k+1}^r) + \Delta \mathbf{P} \quad (24)$$

式中:  $\mathbf{X}_{k+1}^r$  为真实状态;  $\mathbf{P}_{k+1}^r$  为标准卡尔曼滤波算法的实际误差均方阵;  $\Delta \mathbf{P}$  为改进算法式(23)额外引入的误差均方阵;上标 r 代表实际误差均方阵。

**证明** 当模型正确时,真实估计误差均方阵为

$$\begin{aligned} \mathbf{P}_{k+1}^r &= (\mathbf{I} - \mathbf{K}_{k+1} \mathbf{H}_{k+1}) \mathbf{P}_{k+1|k}^r (\mathbf{I} - \mathbf{K}_{k+1} \mathbf{H}_{k+1})^T + \\ &\mathbf{K}_{k+1} \mathbf{R}_{k+1} \mathbf{K}_{k+1}^T \end{aligned} \quad (25)$$

$$\mathbf{P}_{k+1|k}^r = \Phi_{k+1,k} \mathbf{P}_k^r \Phi_{k+1,k}^T + \mathbf{Q}_k \quad (26)$$

由于模糊渐消函数  $g(\gamma_1)$  的存在,式(26)中的滤波增益将产生偏差,  $\mathbf{K}_{k+1}^* = \mathbf{K}_{k+1} + \Delta \mathbf{K}$ , 且由所设计的模糊规则可知  $\Delta \mathbf{K}$  有界。由于  $\Delta \mathbf{K}$  实际均方差也将偏离原始值:

$$\begin{aligned} P_{k+1}^r + \Delta P &= [I - (K_{k+1} + \Delta K)H_{k+1}]P_{k+1|k}^r [I - \\ & (K_{k+1} + \Delta K)H_{k+1}]^T + (K_{k+1} + \\ & \Delta K)R_{k+1}(K_{k+1} + \Delta K)^T \end{aligned} \quad (27)$$

将式(25)代入式(27),并注意

$$H_{k+1}P_{k+1|k}^r - (H_{k+1}P_{k+1|k}^r H_{k+1}^T + R_{k+1})K_{k+1}^T = 0$$

因此有

$$\Delta P = \Delta K(H_{k+1}P_{k+1|k}^r H_{k+1}^T + R_{k+1})\Delta K^T \quad (28)$$

因为  $\Delta K$  有界,且  $P_{k+1|k}^r$  和  $R_{k+1}$  为正定矩阵,所以  $\Delta P$  存在有限上界。证毕

由上述定理可知,随着迭代过程,本文的模糊强跟踪扩展卡尔曼滤波器的误差将收敛于有限值。

### 3 仿真实验

本文以无人机 INS/GPS 组合导航系统来仿真研究本文所提滤波算法的稳定性与适应性。

#### 3.1 无人机导航运动学模型

定义系统的状态向量为

$$X = [q_0 \quad q_1 \quad q_2 \quad q_3 \quad v_n \quad v_e \quad v_d \quad x_n \quad x_e \quad x_d]^T \quad (29)$$

式中: $q_0, q_1, q_2$  和  $q_3$  为无人机当前时刻的姿态四元数; $V_n, V_e$  和  $V_d$  为无人机在导航(地球)坐标系

下的速度; $x_n, x_e$  和  $x_d$  为相应的位置。

确定状态方程为

$$X_{k+1} = f(X_k, k) \quad (30)$$

导航状态更新方程包括速度、未知项和姿态

3 方面,其中姿态更新方程为

$$\begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix}_{k+1} = \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix}_k + \frac{1}{2}T \begin{bmatrix} 0 & -w_x & -w_y & -w_z \\ w_x & 0 & w_z & -w_y \\ w_y & -w_z & 0 & w_x \\ w_z & w_y & -w_x & 0 \end{bmatrix} \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix}_k \quad (31)$$

式中: $w_x, w_y$  和  $w_z$  为陀螺仪传感器在当前时刻的读数; $T$  为传感器的采集周期。

相应地,速度更新方程为

$$\begin{bmatrix} v_n \\ v_e \\ v_d \end{bmatrix}_{k+1} = \begin{bmatrix} v_n \\ v_e \\ v_d \end{bmatrix}_k + C_b^n \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix}_k T \quad (32)$$

式中: $a_x, a_y$  和  $a_z$  为加速度计读数; $C_b^n$  为载体坐标系向导航坐标系坐标变换的方向余弦矩阵,  $C_b^n = (C_n^b)^T$ 。

设运载体的航向角为  $\psi$ ,俯仰角为  $\theta$ ,滚转角为  $\gamma$ ,则有

$$C_b^n = \begin{bmatrix} \cos \gamma \cos \psi + \sin \gamma \sin \psi \sin \theta & -\cos \gamma \sin \psi + \sin \gamma \cos \psi \sin \theta & -\sin \gamma \cos \theta \\ \sin \psi \cos \theta & \cos \psi \cos \theta & \sin \theta \\ \sin \gamma \cos \psi - \cos \gamma \sin \psi \sin \theta & -\sin \gamma \sin \psi - \cos \gamma \cos \psi \sin \theta & \cos \gamma \cos \theta \end{bmatrix} \quad (33)$$

位置更新方程为

$$\begin{bmatrix} x_n \\ x_e \\ x_d \end{bmatrix}_{k+1} = \begin{bmatrix} x_n \\ x_e \\ x_d \end{bmatrix}_k + \begin{bmatrix} v_n \\ v_e \\ v_d \end{bmatrix}_k T \quad (34)$$

最后,确定系统的观测方程为

$$Z_{k+1} = h(X_{k+1}, k+1) \quad (35)$$

导航系统采用的是 INS—GPS 组合导航系统,将 GPS 的量测信息作为系统的观测向量:

$$Z = [z_n \quad z_e \quad z_d]^T \quad (36)$$

可见系统观测信息就是状态位置信息,则系统观测方程可以写成

$$\begin{cases} z_n = x_n \\ z_e = x_e \\ z_d = x_d \end{cases} \quad (37)$$

#### 3.2 滤波参数设计

需要注意的是,滤波的效果受 2 个初值和

2 个矩阵参数影响,在经过多次调试后,并且为了对比不同滤波方法的性能,本文对不同滤波方法都采用相同的初值和参数。具体取值为

$$X_0 = [1 \quad 0 \quad 0 \quad 0 \quad 0 \text{ m/s} \quad 0 \text{ m/s} \quad 0 \text{ m/s} \quad 100 \text{ m} \quad 100 \text{ m} \quad 100 \text{ m}]^T$$

$$P_0 = \text{diag}(0.000 \ 1, 0.000 \ 1, 0.000 \ 1, 0.007 \ 6, 0.7 \ (\text{m/s})^2, 0.7 \ (\text{m/s})^2, 0.7 \ (\text{m/s})^2, 10 \ \text{m}^2, 10 \ \text{m}^2, 10 \ \text{m}^2)$$

$$Q = \text{diag}(0, 0, 0, 0, 10 \ (\text{m/s})^2, 10 \ (\text{m/s})^2, 10 \ (\text{m/s})^2, 1 \ \text{m}^2, 1 \ \text{m}^2, 1 \ \text{m}^2)$$

$$R = \begin{bmatrix} 1 \ 000 \ \text{m}^2 & 0 & 0 \\ 0 & 1 \ 000 \ \text{m}^2 & 0 \\ 0 & 0 & 1 \ 000 \ \text{m}^2 \end{bmatrix}$$

用英达软件实现仿真,时间采样周期  $T =$

0.02 s。总时长为 40 s。仿真过程中 GPS 量测始终有效,仿真结果如图 1 所示。

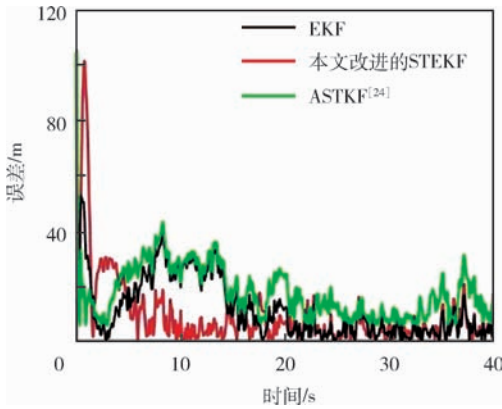


图 1 3 种滤波算法结果(东向)

Fig. 1 Results of three filtering algorithms (east)

图 1 中,将本文改进 STEKF 算法与文献[24]中的 ASTKF 算法做对比(文献[24]提出了自适应强跟踪卡尔曼滤波(ASTKF),并将其运用于位置估计中),本文提出的基于模糊理论的强跟踪扩展卡尔曼滤波器能够对飞行状态做出有效估计,且具有较低的跟踪误差,在整个仿真时间内稳定。

下面定量地分析 3 种滤波算法的精度,导航滤波精度的指标是估计误差的均方根值:

$$EMS = \sqrt{\frac{1}{N} \sum_{k=1}^N (x_k - \hat{x}_k)^2} \quad (38)$$

式中: $N$  为仿真采样点数, $N = 2\,000$ ;  $x_k$  和  $\hat{x}_k$  分别为分析对象的测量值和估计值。

计算出无人机北向、东向、地下的位置误差  $EMS_n$ 、 $EMS_e$ 、 $EMS_d$  后,则导航滤波误差指标可以表示为

$$EMS = \sqrt{\frac{1}{3} (EMS_n^2 + EMS_e^2 + EMS_d^2)} \quad (39)$$

计算得出 3 种滤波算法的误差如表 1 所示。

表 1 3 种滤波算法的精度

Table 1 Accuracy of three filtering algorithms

算法	估计误差均方根值/m
EKF	6.35
ASTKF <sup>[24]</sup>	7.54
本文改进的 STEKF	5.99

可见,相较于文献[24]中 ASTKF 算法以牺牲一定滤波精度换取强跟踪效果,本文改进的 STEKF 算法具有更高的滤波精度(甚至优于普通的 EKF 算法)。

### 3.3 GPS 中断响应测试

给 GPS 信号人为地加入持续时间为 4 s 的故障(仿真开始后的 10 ~ 14 s),在这 4 s 内,GPS 的量测值清零。这段时间内,无人机仅借助惯性导航而进行位置姿态估计,随着时间的推移将产生累计误差。GPS 故障解除后,无人机重新利用 GPS 信号来进行定位,误差会逐渐减小。为了观察系统的恢复情况。比较了不同滤波算法对 GPS 信号突变的快速响应能力。

图 2 为 GPS 发生中断后,仿真得到的滤波估计轨迹与理论飞行轨迹的路线,可见在 GPS 故障的这段时间,滤波轨迹发生了偏移。

图 3 为 3 种滤波算法在北向、东向和地向上滤波误差。可见,在 14 s 时 GPS 信号回复到正常后,本文改进的 STEKF 的滤波算法和文献[24]中 ASTKF 算法误差能够快速收敛至零,而普通 EKF 的滤波则会出现较大幅度的振荡,至少需要 3 s 左右的时间才能完成对位置的准确估计。

可以利用估计残差对滤波是否收敛进行检测<sup>[4]</sup>,根据式(31),认为当估计残差与期望残差之比  $\gamma_0$  始终小于 3 时滤波收敛且进入稳定状态。从 GPS 中断结束到滤波进入稳态的这段时间,称之为收敛时间。通过仿真,得到不同滤波算法的收敛时间如表 2 所示。

可见,改进算法相较于普通的 EKF 算法极大地减小了导航过程中 GPS 中断恢复后的收敛时间。

综合 3.2 节和 3.3 节,可见本文提出的基于模糊理论的强跟踪扩展卡尔曼滤波器在能够快速追踪 GPS 量测信号的同时,比文献[24]中 ASTKF 算法具有更高的滤波精度。

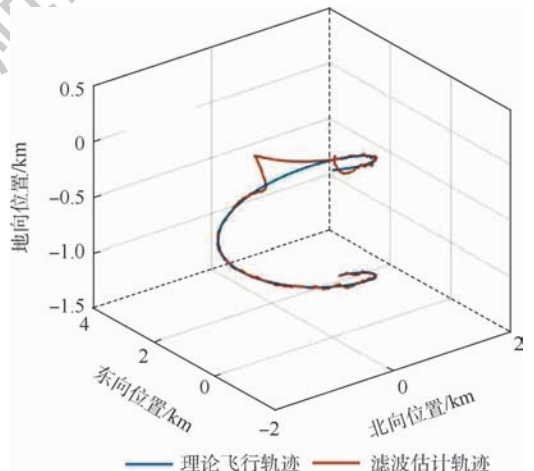


图 2 飞行轨迹变化

Fig. 2 Flight path variation

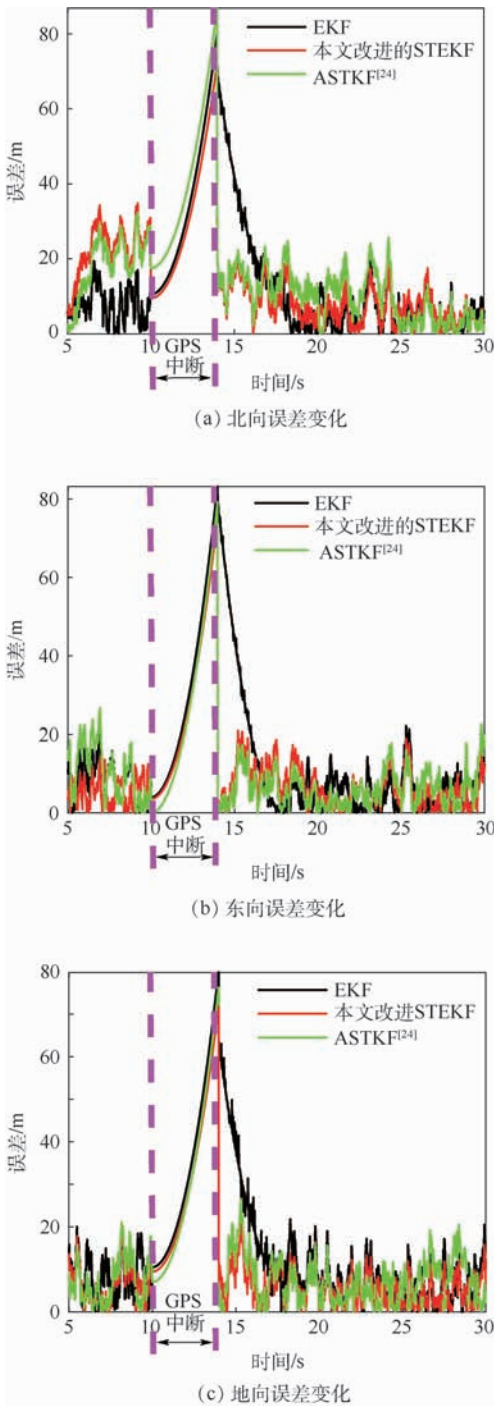


图3 3种滤波算法在北向、东向和地向误差变化  
Fig.3 North, east and ground direction error variation

表2 3种滤波算法针对GPS中断收敛速度

Table 2 Convergence speed of three filtering algorithms for GPS interruption

算法	GPS 中断恢复后收敛速度/s
EKF	3.84
ASTKF <sup>[24]</sup>	0.38
本文改进的 STEKF	0.25

## 4 结论

针对导航系统中存在这GPS信号中断丢失的问题,本文利用模糊理论中的隶属度函数,对基本的强跟踪扩展卡尔曼滤波器作出了改进,主要结论如下:

- 1) 本文将强跟踪扩展卡尔曼滤波(STEKF)运用到了无人机导航系统中,并对其作出了改进。
- 2) 改进STEKF算法能够在GPS信号恢复后迅速地进入稳态(0.25 s),而不是逐渐地收敛。
- 3) 改进STEKF算法保证了导航滤波精度,误差小于普通的EKF算法和文献[24]中的ASTKF算法。

### 参考文献 (References)

[1] 李子月,张林,陈善秋,等. 捷联惯性/卫星超紧组合导航技术综述与展望[J]. 系统工程与电子技术,2016,38(4):866-874.  
LI Z Y, ZHANG L, CHEN S Q, et al. Ultra-tightly coupled SINS/GLASS navigation technology review and prospect[J]. Systems Engineering and Electronics, 2016, 38(4):866-874 (in Chinese).

[2] 樊云鹏,杨顿昌. GNSS/INS 深组合导航关键技术研究[J]. 飞航导弹,2017(10):22-25.  
FAN Y P, YANG S C. Research on key technologies of GNSS/INS deep integrated navigation[J]. Aerodynamic Missile Journal, 2017(10):22-25 (in Chinese).

[3] 韩乃龙. 高动态惯性/卫星组合导航技术研究[D]. 南京:南京理工大学,2017.  
HAN N L. Research on high dynamic inertial/satellite integrated navigation technology [D]. Nanjing: Nanjing University of Science and Technology, 2017 (in Chinese).

[4] 王辰熙. 多旋翼无人机组导航技术研究[D]. 南京:南京航空航天大学,2017.  
WANG C X. Research on integrated navigation technology for multi-rotor unmanned aerial vehicle[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2017 (in Chinese).

[5] KALMAN R E. A new approach to linear filtering and prediction problems[J]. Journal of Basic Engineering Transactions, 1960, 82(1):35-45.

[6] BUCY R. Nonlinear filtering theory[J]. IEEE Transactions on Automatic Control, 1965, 10(2):198.

[7] JULIER S, UHLMANN J, DURRANTWHYTE H F. A new method for the nonlinear transformation of means and covariances in filters and estimators[J]. IEEE Transactions on Automatic Control, 2000, 45(3):477-482.

[8] 景丽. 基于卡尔曼滤波组合导航算法的计算量与精度分析[D]. 哈尔滨:哈尔滨工业大学,2014.  
JING L. Analysis of calculation and precision based on Kalman filter combined navigation algorithm [D]. Harbin: Harbin Insti-

- tute of Technology, 2014 (in Chinese).
- [ 9 ] PERSSON S M, SHARF I. Invariant momentum-tracking Kalman filter for attitude estimation [ C ] // 2012 IEEE International Conference on Robotics and Automation. Piscataway, NJ: IEEE Press, 2012: 592-598.
- [ 10 ] TAYLOR C N. An analysis of observability-constrained Kalman filtering for vision-aided navigation [ C ] // Position Location and Navigation Symposium. Piscataway, NJ: IEEE Press, 2012: 1240-1246.
- [ 11 ] JIANG Z, ZHOU W, LI H, et al. A new kind of accurate calibration method for robotic kinematic parameters based on the extended Kalman and particle filter algorithm [ J ]. IEEE Transactions on Industrial Electronics, 2018, 65 ( 4 ): 3337-3345.
- [ 12 ] DAVARI N, GHOLAMI A. An asynchronous adaptive direct Kalman filter algorithm to improve underwater navigation system performance [ J ]. IEEE Sensors Journal, 2017, 17 ( 4 ): 1061-1068.
- [ 13 ] 李兆铭, 杨文革, 丁丹, 等. 多星对合作目标的分布式协同导航滤波算法 [ J ]. 北京航空航天大学学报, 2018, 44 ( 3 ): 462-469.
- LI Z M, YANG W G, DING D, et al. Distributed cooperative navigation filtering algorithm for multi-satellite cooperative targets [ J ]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44 ( 3 ): 462-469 ( in Chinese ).
- [ 14 ] SUN F, TANG L. Augmented and non-augmented cubature Kalman filter [ J ]. Journal of Information & Computational Science, 2012, 9 ( 2 ): 437-450.
- [ 15 ] 赵曦晶, 刘光斌, 汪立新, 等. 五阶容积卡尔曼滤波算法及其应用 [ J ]. 红外与激光工程, 2015, 44 ( 4 ): 1377-1381.
- ZHAO X J, LIU G B, WANG L X, et al. Fifth-order volume Kalman filter algorithm and its application [ J ]. Infrared and Laser Engineering, 2015, 44 ( 4 ): 1377-1381 ( in Chinese ).
- [ 16 ] 刘旭, 张其善, 杨东凯. 一种用于 GPS/DR 组合定位的非线性滤波算法 [ J ]. 北京航空航天大学学报, 2007, 33 ( 2 ): 184-187.
- LIU X, ZHANG Q S, YANG D K. A nonlinear filtering algorithm for GPS/DR combined positioning [ J ]. Journal of Beijing University of Aeronautics and Astronautics, 2007, 33 ( 2 ): 184-187 ( in Chinese ).
- [ 17 ] MARINA H G D. UAV attitude estimation using unscented Kalman filter and TRIAD [ J ]. IEEE Transactions on Industrial Electronics, 2012, 59 ( 11 ): 4465-4474.
- [ 18 ] GUPTA S D, YU J Y, MALLICK M, et al. Comparison of angle-only filtering algorithms in 3D using EKF, UKF, PF, PFF, and ensemble KF [ C ] // 18th International Conference on Information Fusion. Piscataway, NJ: IEEE Press, 2015: 1649-1656.
- [ 19 ] CHANG L, HU B, LI A, et al. Transformed unscented Kalman filter [ J ]. IEEE Transactions on Automatic Control, 2012, 58 ( 1 ): 252-257.
- [ 20 ] 薛远奎. 小型旋翼无人机姿态融合及导航方法 [ D ]. 成都: 电子科技大学, 2016.
- XUE Y K. Attitude fusion and navigation method of small rotorcraft [ D ]. Chengdu: University of Electronic Science and Technology of China, 2016 ( in Chinese ).
- [ 21 ] 王亭亭, 蔡志浩, 王英勋. 无人机室内视觉/惯导组合导航方法 [ J ]. 北京航空航天大学学报, 2018, 44 ( 1 ): 176-186.
- WANG T T, CAI Z H, WANG Y X. Indoor vision/INS integrated navigation method for UAV [ J ]. Journal of Beijing University of Aeronautics and Astronautics, 2018, 44 ( 1 ): 176-186 ( in Chinese ).
- [ 22 ] 钱华明, 王迪, 吴永慧. 小视场星敏感器量测延时滤波算法 [ J ]. 北京航空航天大学学报, 2019, 45 ( 2 ): 234-242.
- QIAN H M, WANG D, WU Y H. Delayed filtering algorithm for small field of view star sensor measurement [ J ]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45 ( 2 ): 234-242 ( in Chinese ).
- [ 23 ] 王尔申, 贾超颖, 曲萍萍, 等. 基于混沌粒子群优化的北斗/GPS 组合导航选星算法 [ J ]. 北京航空航天大学学报, 2019, 45 ( 2 ): 259-265.
- WANG E S, JIA C Y, QU P P, et al. Beidou/GPS integrated navigation star selection algorithm based on chaotic particle swarm optimization [ J ]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45 ( 2 ): 259-265 ( in Chinese ).
- [ 24 ] LIN C L, CHANG Y M, HUNG C C, et al. Position estimation and smooth tracking with a fuzzy-logic-based adaptive strong tracking Kalman filter for capacitive touch panels [ J ]. IEEE Transactions on Industrial Electronics, 2015, 62 ( 8 ): 5097-5108.

#### 作者简介:

何康辉 男, 硕士研究生。主要研究方向: 航空航天器的导航与控制。

董朝阳 男, 博士, 教授, 博士生导师。主要研究方向: 航空航天器的导航与控制。

# A navigation filtering algorithm considering GPS signal interruption

HE Kanghui<sup>1</sup>, DONG Chaoyang<sup>1,\*</sup>, WANG Qing<sup>2</sup>

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

2. School of Automation Science and Electrical Engineering, Beihang University, Beijing 100083, China)

**Abstract:** Aimed at the problem of GPS data interruption in the UAV INS/GPS integrated navigation system, an improved filtering algorithm is designed. Firstly, the kinematics model of UAV navigation is established. Then the traditional extended Kalman filter (EKF) technique and strong tracking filter are combined. A new navigation filtering algorithm is designed by using the membership function in fuzzy theory. The simulation results show that the improved algorithm can quickly adapt to the sudden change of GPS signal. When the GPS signal recovers from the fault state to the normal state, the improved algorithm can converge to the steady state more quickly than the ordinary EKF algorithm, and the flight state is re-completed and estimated. At the same time, compared with the common EKF and strong tracking extended Kalman filter (STEKF) algorithm, the improved algorithm has higher filtering accuracy.

**Keywords:** extended Kalman filter (EKF); strong tracking filter; integrated navigation; global positioning system (GPS); membership function

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

DOI: 10.13700/j.bh.1001-5965.2018.0788

# 多向气动驱动器软体仿生舌弯曲状态的研究

董虎<sup>1,2</sup>, 林苗<sup>1,2</sup>, 顾苏程<sup>1,2</sup>, 曹毅<sup>1,2,\*</sup>, 李巍<sup>3</sup>

(1. 江南大学 机械工程学院, 无锡 214122; 2. 江苏省食品先进制造装备技术重点实验室, 无锡 214122;

3. 苏州工业职业技术学院, 苏州 215104)



**摘 要:** 以软体仿生舌为研究对象, 针对多向驱动器的变形机理, 开展了面向软体仿生舌的单/多向驱动器的弯曲特性研究。首先, 设计了气动驱动软体仿生舌, 其在单/多向驱动器的作用下可实现吐舌、翻舌、卷舌以及斜翻舌等动作; 其次, 为进一步研究驱动器的变形机理, 开展了针对单/多向驱动器的结构及变形工作原理的研究; 然后, 基于 Yeoh 模型应变能密度函数, 结合力平衡方程, 建立了驱动压强和驱动器弯曲变形后曲率半径的非线性数学模型; 最后, 为验证理论模型的正确性, 开展了软体仿生舌的仿真研究及相关实验验证, 结果证明了理论模型的正确性。上述研究为其他基于气动驱动的软体结构变形机理的研究提供了理论基础。

**关键词:** 软体仿生舌; 气动驱动; Yeoh 模型; 有限元仿真; 弯曲特性

**中图分类号:** TP241

**文献标识码:** A **文章编号:** 1001-5965(2019)09-1882-12

传统机器人多以刚性结构为主, 其结构复杂、自适应能力差、自由度有限、安全性低<sup>[1-3]</sup>。与刚性机器人不同, 软体机器人基于仿生学的研究成果<sup>[4-6]</sup>, 以新兴的智能材料为依托<sup>[7]</sup>, 模仿自然界中的无脊椎动物, 如蚯蚓、章鱼及动物的舌等。其结构一般比较简单、自适应能力强、理论上可以有任意自由度、安全性较高<sup>[8]</sup>。因此, 软体机器人可在医疗手术、地震救援、故障检测<sup>[9-10]</sup>等诸多非结构化工作环境中完成既定任务。

基于软体机器人的诸多优良特性以及其在非结构化工作环境中的广阔应用前景, 其正逐渐成为国内外机器人领域的研究热点。综合国内外软体机器人的发展历程不难发现, 基于驱动方式软体机器人可划分为: 气动驱动<sup>[11-12]</sup>、丝线驱动<sup>[13]</sup>以及智能材料电致动<sup>[14-16]</sup>。在气动驱动方面,

Fionnuala 等<sup>[17]</sup>用实验法总结了纤维增强型软体驱动器的变形规律; Kevin 等<sup>[18]</sup>基于实验法分析了水下软体夹持器的力学特性; Kristin 等<sup>[19]</sup>基于欧拉弹力定理和杆理论建立气动软体驱动器的理论模型; 王宁扬等<sup>[20]</sup>基于形封闭性理论分析了蜂巢气动软体夹持器的变形机理; 魏树军等<sup>[21]</sup>基于实验法研发了一种纤维增强型气动软体抓手; 郑俊君等<sup>[22]</sup>基于流体静压力学特性构建了气动静压软体机器人的理论模型; 王华等<sup>[23]</sup>在橡胶材料本构关系的基础上, 应用虚功原理分析了软体弯曲驱动器的变形机理; 费燕琼等<sup>[24]</sup>基于橡胶材料的本构关系, 分析了气动驱动软体机器人的运动机理。在丝线驱动方面, Kim 等<sup>[25]</sup>基于欧拉-伯努利梁方程构建了 SMA (Shape Memory Alloys) 丝驱动的软体手的理论模型; Mathijssen 等<sup>[26]</sup>基于

收稿日期: 2019-01-02; 录用日期: 2019-03-08; 网络出版时间: 2019-04-14 07:02

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190411.1643.002.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190411.1643.002.html)

基金项目: 国家自然科学基金 (50905075); 江苏省“六大人才高峰”计划 (ZBZZ-012)

\* 通信作者: E-mail: caoyi@jiangnan.edu.cn

引用格式: 董虎, 林苗, 顾苏程, 等. 多向气动驱动器软体仿生舌弯曲状态的研究[J]. 北京航空航天大学学报, 2019, 45(9): 1882-1893. DONG H, LIN M, GU S C, et al. Motion characteristics of soft bionic tongue based on multi-directional pneumatic actuator [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1882-1893 (in Chinese).

等距实验法总结了人工肌肉丝驱动器的变形规律;史震云和朱前成<sup>[27]</sup>基于径向基函数神经网络和支持度函数构建了智能驱动器软体机器人的理论模型;王扬威等<sup>[28]</sup>基于运动学模型分析了形状记忆合金丝驱动柔性机械臂的变形机理;Fei和Xu<sup>[29]</sup>基于实验法及反馈数据总结了SMA软体驱动器的变化规律。在电致动方面,Matlu等<sup>[30]</sup>基于静准态有限元模型分析了EAP(Electro Active Polymer)致动装置的变形机理;郭闯强等<sup>[16]</sup>基于IPMC(Ionic Polymer Metal Composite)分析了微型仿生机器鱼舌的变形机理。

在食品科学领域,模拟咀嚼、吞咽功能的软体仿生舌的研制具有重要价值,谢高鹏提出了仿生舌并进行了食品咀嚼模拟机系统的研究<sup>[31]</sup>;韩敬虎<sup>[32]</sup>设计了一款仿生舌并研制了咀嚼吞咽一体化仿生机器人。

本文以软体仿生舌为研究对象,研究了驱动器的变形机理,并基于Yeoh模型建立了非线性数学模型,开展了软体仿生舌的仿真研究及相关实验验证,结果证明了结构设计的合理性以及理论模型的正确性。

## 1 软体仿生舌的结构设计

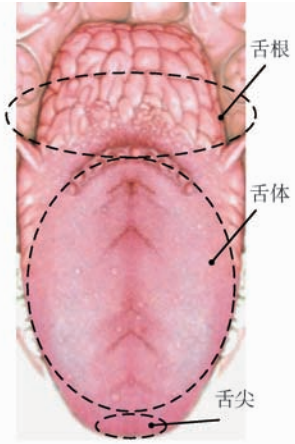
如图1(a)所示,舌是口腔内的肌性器官,可分为舌根、舌体和舌尖三部分。作为一种骨骼肌,其运动十分灵活,参与咀嚼、吞咽及协助语言等功能。

为模拟舌的吞咽、咀嚼等功能,基于气动驱动器设计了一款软体仿生舌,如图1(b)所示。该软体仿生舌主要由多向驱动器、单向驱动器及舌本体组成,其中单/多向驱动器黏附在舌本体下部,控制仿生舌的吐舌、翻舌、卷舌及斜翻舌等功能,同时半/通透槽可减少挤压应力,在驱动压强 $p = 50 \text{ kPa}$ 时,预期吐舌 $5 \text{ mm}$ ,翻舌 $1.2 \text{ rad}$ ,卷舌 $1.3 \text{ rad}$ ,斜翻舌 $1 \text{ rad}$ 。

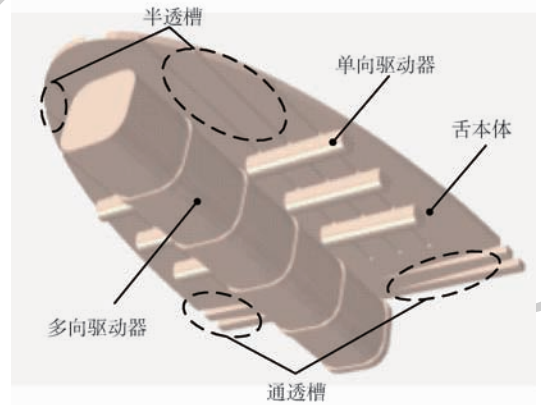
软体仿生舌制备材料均为肖氏硬度20度的硅橡胶,具体方法为模具浇铸法,基于3D打印得到树脂模具,配合1:1混合的液态硅胶固化完成制备工序。

### 1.1 多向驱动器结构及工作原理

多向驱动器由4个充气软管通过实心连接管连接而成,其外部包覆有约束管和约束线圈,如图2所示。当对其中一个充气软管充气时,在气压 $p$ 的作用下软管膨胀并轴向拉伸,同时受驱动器中非受压软管的应力 $\sigma$ 、外部约束管及约束



(a) 舌结构示意图



(b) 软体仿生舌结构

图1 舌示意图

Fig. 1 Diagram of tongue

线圈约束作用,从而整个驱动器完成弯曲动作。同理,若对其他充气软管充气驱动器就会完成其相应方向的弯曲动作;如对2个软管进行充气,则会完成较之前偏转 $45^\circ$ 的弯曲动作,图2中 $R$ 为弯曲半径, $L_0$ 为驱动器原长;若对所有充气软管充气驱动器则会完成拉伸动作。

该软体仿生舌尺寸以成年男性舌的大小为设计依据,并以1:1确定各结构尺寸,多向驱动器结构参数如表1所示。

### 1.2 单向驱动器结构及工作原理

单向驱动器主要由充气软管和底部较厚的硅橡胶板组成,两者紧密粘合在一起,如图3所示。当对软管充气时,在气压 $p_1$ 的作用下软管膨胀并轴向拉伸,同时驱动器中非受压软管的应力 $\sigma_1$ 则起到限制作用,从而单向驱动器实现弯曲功能,其弯曲半径 $R_1$ 如图3所示。

同理,基于软体仿生舌结构尺寸参数,可确定单向驱动器结构参数,具体如表2所示。

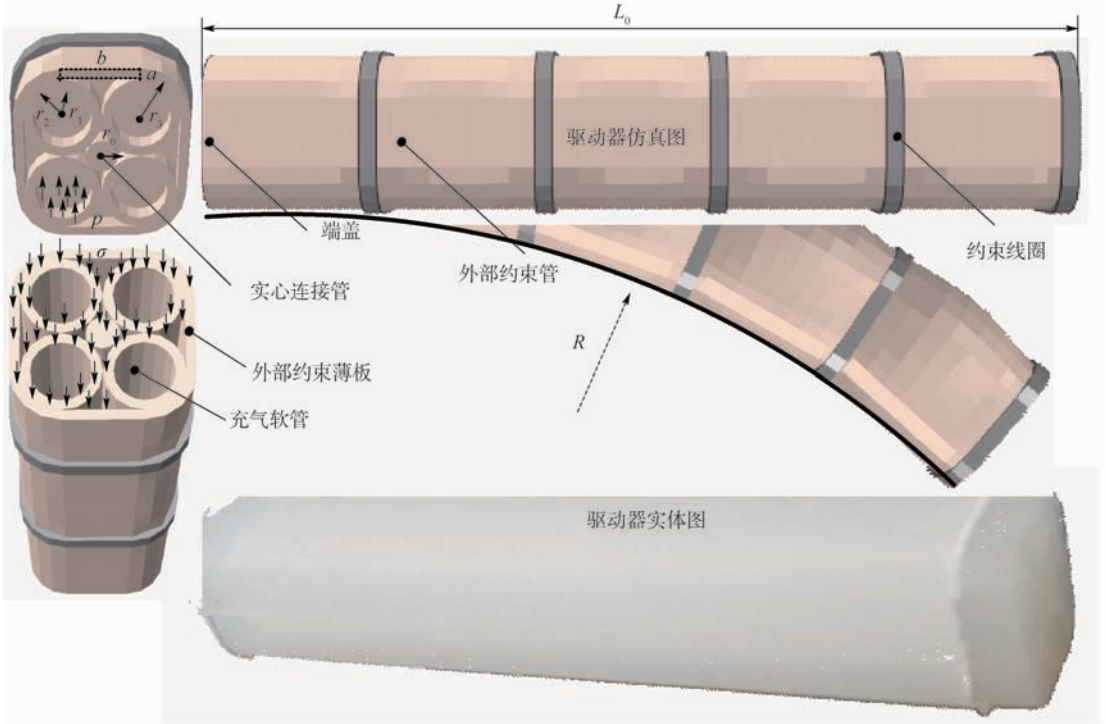


图2 多向驱动器结构及工作原理图

Fig.2 Diagram of structure and working principle of multi-directional actuator

表1 多向驱动器结构参数

Table 1 Structural parameters of multi-directional actuator

参数	数值/mm
驱动器长度 $L$	100
连接管半径 $r_0$	1.66
充气软管内径 $r_1$	3
充气软管外径 $r_2$	4
约束管圆弧半径 $r_3$	5
约束管薄板宽 $a$	1
约束管薄板长 $b$	8

## 2 驱动器变形机理的理论建模

单/多向驱动器变形机理的研究是软体仿生舌运动特性研究的理论基础,因此有必要构建单/多向驱动器变形机理的理论模型。值得指出的是,硅橡胶作为超弹性材料,一般具有各向同性和不可压缩性<sup>[31]</sup>,且各项力学特性均表现为非线性,因此本文首先研究较为基础的运动特性。

### 2.1 硅橡胶材料的应变能密度函数

对超弹性材料非线性力学的研究,其常用的研究方法有以下几种:非线性方程的线性化;最小势能原理;应变能密度函数等。线性化是通过 Piola-Kirchhoff 应力和 Cauchy-Green 应变建立 Lagrange 方程<sup>[32]</sup>,将非线性方程线性化进行求解;此类方法求解相对简单,但是不易得到精确解,故应用不是很广泛。最小势能原理是根据能量法建立平衡方程<sup>[23]</sup>。应变能法是基于弹性力学并选择合适的应变能密度函数来建立力的平衡方程<sup>[24]</sup>。

该软体驱动器结构简单,变形较大且变形后软管处于非确定性状态,因此 Lagrange 方程和能量方程不易构建。同时,文献[33]表明 Yeoh 模型应变能密度函数能适用于超弹性材料的大变形行为,因此,本文拟采用应变能密度函数对驱动器的变形机理进行研究。

首先根据应力应变关系以唯象理论建立硅橡

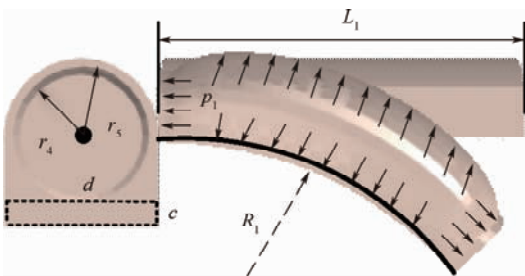


图3 单向驱动器结构及工作原理图

Fig.3 Diagram of structure and working principle of one-directional actuator

表2 单向驱动器结构参数

Table 2 Structural parameters of one-directional actuator

参数	数值/mm
驱动器长度 $L_1$	100
充气软管内径 $r_4$	2.2
充气软管外径 $r_5$	2.5
约束管薄板宽 $c$	0.8
约束管薄板长 $d$	5

胶材料的本构关系<sup>[33]</sup>,并采用应变能密度函数进行表示:

$$W = W(I_1, I_2, I_3) \quad (1)$$

式中:

$$\begin{cases} I_1 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 \\ I_2 = \lambda_1^2 \lambda_2^2 + \lambda_2^2 \lambda_3^2 + \lambda_1^2 \lambda_3^2 \\ I_3 = (\lambda_1 \lambda_2 \lambda_3)^2 \end{cases} \quad (2)$$

其中: $I_{1,2,3}$ 为变形张量不变量; $\lambda_{1,2,3}$ 为拉伸比,假设硅橡胶材料是不可压缩的,可得

$$I_3 = (\lambda_1 \lambda_2 \lambda_3)^2 = 1 \quad (3)$$

常用的应变能密度函数模型有 Neo-Hookean 模型、Yeoh 模型、Mooney-Rivlin 模型等,其中 Yeoh 模型常用来描述大变形行为<sup>[34]</sup>,故本文拟采用其最经典的二参数形式:

$$W = C_{10}(I_1 - 3) + C_{20}(I_2 - 3)^2 \quad (4)$$

式中: $C_{10}$ 、 $C_{20}$ 为硅橡胶材料的 Yeoh 模型材料常数,如图 4 所示。其可利用 CTM2100 微机控制电子拉伸仪通过单轴向拉伸实验测定,限于篇幅,具体测定过程不予赘述。

依据文献[33]中“求 Yeoh 模型常数部分”,可得到公式  $t_1/[2(\lambda_1 - 1/\lambda_1^2)] = 2C_{20}(\lambda_1^2 + 2/\lambda_1) + C_{10} - 6C_{20}$ ,测得拉伸比  $\lambda_1$  和主应力  $t_1$ ,将  $t_1/[2(\lambda_1 - 1/\lambda_1^2)]$  作为  $y$  轴, $\lambda_1^2 + 2/\lambda_1$  作为  $x$  轴,可以拟合并计算出材料常数  $C_{10} = 0.072$ ,  $C_{20} = 0.0025$ 。

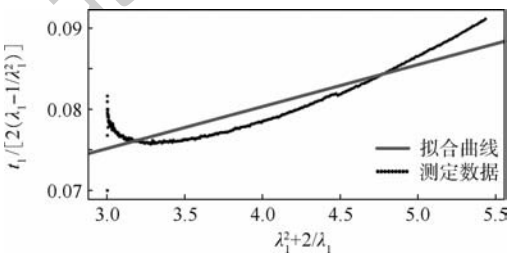


图 4 Yeoh 模型材料常数拟合曲线

Fig.4 Fitting curve of Yeoh model parameter

## 2.2 驱动器弯曲变形理论模型

分析多向驱动器的结构发现,整个驱动器由实心连接管、充气软管以及约束薄板组成。而实心连接管与充气软管只在截面积上存在差异,故可以归结为一类,如图 5 所示。为分析其变形机理,假设驱动器原长为  $L_0$ ,变形后为  $L$ ,变形后的曲率半径和弧长分别为  $R$  和  $\theta$ ,根据硅橡胶材料的不可压缩性,可认为底部长度无变化,仍为  $L_0$ 。

基于 Yeoh 模型应变能密度函数,该驱动器在弯曲变形过程中,长度方向上的拉伸最为明显,设定为  $\lambda_1$ ,而其余 2 个方向拉伸比相同,结合式(3)可得

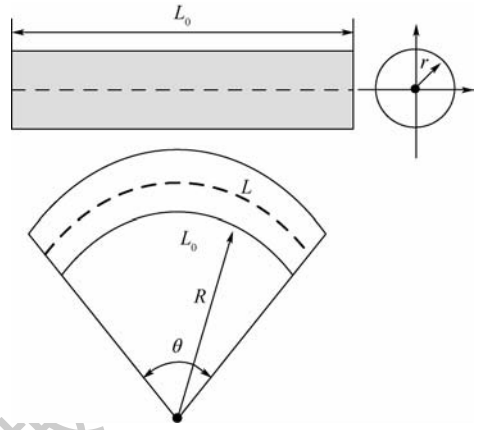


图 5 管状弯曲状态示意图

Fig.5 Bending diagram of tube

$$\lambda_2 = \lambda_3 = \sqrt{\frac{1}{\lambda_1}} \quad (5)$$

根据图 5 并结合弧长和半径的几何关系可以得到

$$\lambda_1 = \frac{R + 1}{R} \quad (6)$$

基于式(2)和式(5)可得

$$\begin{cases} I_1 = \lambda_1^2 + \frac{2}{\lambda_1} \\ I_2 = 2\lambda_1 + \frac{2}{\lambda_1^2} \end{cases} \quad (7)$$

基于应力与应变能关系:

$$\begin{cases} \sigma = \frac{\partial W}{\partial \lambda_1} \\ W = C_{10}(I_1 - 3) + C_{20}(I_2 - 3)^2 \end{cases} \quad (8)$$

联立式(7)和式(8)可确定约束管应力大小:

$$\begin{cases} \sigma = \frac{\partial W}{\partial \lambda_1} = G_1 + 0.005G_2G_3 \\ G_1 = 0.144\lambda_1 - \frac{0.144}{\lambda_1^2} \\ G_2 = 2\lambda_1 + \frac{1}{\lambda_1^2} - 3 \\ G_3 = 2 - \frac{2}{\lambda_1^3} \end{cases} \quad (9)$$

约束薄板的弯曲状态与圆管类似,但截面形状的不同导致其拉伸比不同。其在整个弯曲变形过程当中,长度方向上的拉伸最为明显,设定为  $\lambda_1$ ,宽度方向的拉伸比仅存在很微小的变化<sup>[16]</sup>,假设厚度方向拉伸比不变( $\lambda_3 = 1$ ),则结合式(3)可得

$$\lambda_2 = \frac{1}{\lambda_1} \quad (10)$$

根据图 6 并进一步考虑弧长和半径的几何关系可以得到

$$\lambda_1 = \frac{R + \alpha/2}{R} \quad (11)$$

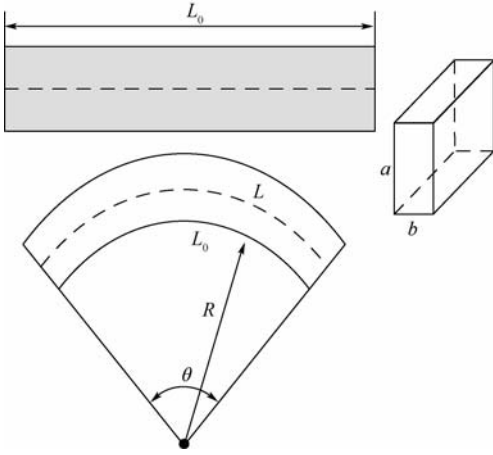


图6 板状弯曲状态示意图  
Fig. 6 Bending diagram of board

同理,基于式(2)和式(10)可以得到

$$I_1 = I_2 = \lambda_1^2 + \frac{1}{\lambda_1^2} + 1 \quad (12)$$

联立式(12)和式(8)可确定约束薄板应力大小:

$$\begin{cases} \sigma = \frac{\partial W}{\partial \lambda_1} = G_4 + 0.005G_5G_6 \\ G_4 = 0.144\lambda_1 - \frac{0.144}{\lambda_1^3} \\ G_5 = \lambda_1^2 + \frac{1}{\lambda_1^2} - 2 \\ G_6 = 2\lambda_1 - \frac{2}{\lambda_1^3} \end{cases} \quad (13)$$

必须指出的是,在受到驱动压强的作用后,驱动器在径向的拉伸会导致轴向的横截面积发生变化,本文采用次拉伸比( $\lambda_2, \lambda_3$ )来表征变形后的横截面积,即  $S = S_0\lambda_2\lambda_3$ ,其中  $S_0$  为原横截面积的大小, $S$  为变形后横截面积的大小。

根据图2的受力示意图,构建其力平衡方程:

$$\sigma S = p\pi r_1^2 \quad (14)$$

联立式(6)、式(9)和式(14)可进一步推导求得驱动器弯曲半径  $R$  与驱动压强  $p$  的函数表达式,限于篇幅,不予赘述,其可表示为

$$R = R(p) \quad (15)$$

同理,结合上述单向驱动器结构发现,其包括充气软管和约束薄板,其变形机理可联立式(11)、式(13)和式(14),推导求得驱动器弯曲半径  $R_1$  与输入压强  $p_1$  的函数表达式,其可表示为

$$R_1 = R_1(p_1) \quad (16)$$

式(15)表征了不同驱动压强作用下,多向驱动器的弯曲变形后的曲率半径,而式(16)表征了不同驱动压强作用下单向驱动器的弯曲半径,上述为后续软体仿生舌的控制奠定了理论基础。

### 3 软体仿生舌有限元仿真

为验证上述理论模型的正确性,并更加直观形象地表征单/多向驱动器在气动驱动下的弯曲变形过程,如图7所示,基于Abaqus软件对驱动器进行有限元仿真。

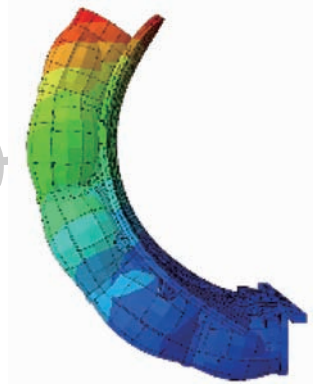


图7 软体仿生舌仿真示意图  
Fig. 7 Simulation diagram of soft bionic tongue

#### 3.1 多向驱动器弯曲变形

基于上文所述,由于多向驱动器由软管对称连接而成,故以单个软管充气为例,图8给出了不同驱动压强作用下多向驱动器弯曲变形的理论计算值及有限元仿真值。

为了更加清楚地表征多向驱动器的变化规律,表3展示了不同驱动压强作用下其末端点坐标的计算及仿真结果。

值得指出的是,上述数据只能表征固定驱动压强作用下多向驱动器末端点的位置坐标,为从一般意义上表征其末端点随压强的变化关系,如图9所示,可将计算结果的各坐标点进行轨迹拟合,拟合结果如下:

$$\begin{cases} x(p) = 153600p^2 - 12950p + 290.1 \\ y(p) = -184800p^2 + 18030p - 351.4 \end{cases} \quad (17)$$

基于图8、图9与表3不难发现:

1) 多向驱动器末端坐标点的理论计算值与有限元仿真值基本一致,从而验证了理论模型的正确性。

2) 随着驱动压强  $p$  的增大,多向驱动器末端点坐标点的理论计算与有限元仿真的相对误差不断增大,其原因在于硅橡胶材料在大应变时表现出高度的非线性特性特性。

3) 基于末端点坐标的相对误差可以发现,当驱动压强大于 56 kPa 后,其相对误差逐渐增大并超过 10%,故该软体仿生舌的工作压强在 56 kPa 以内。

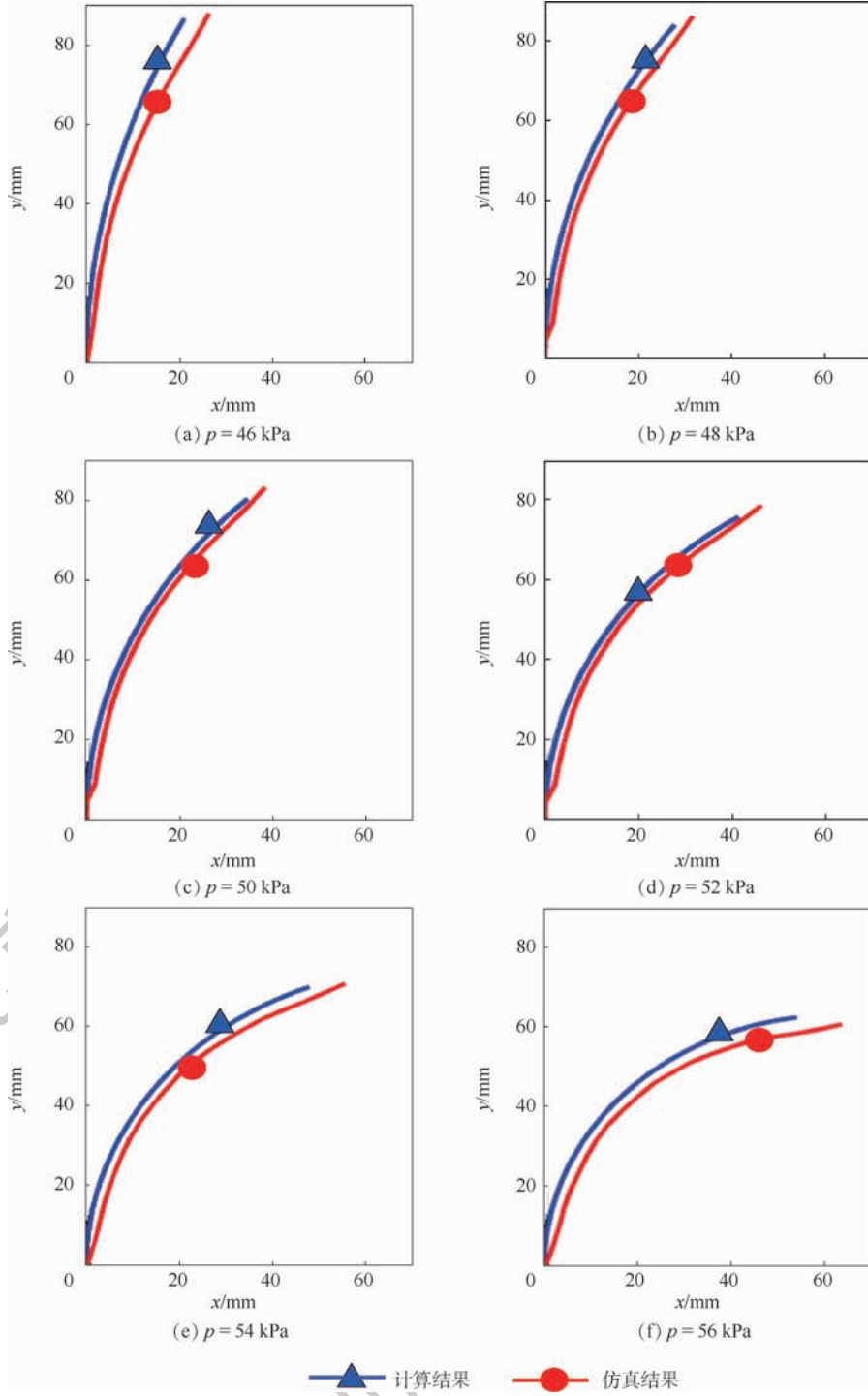


图 8 多向驱动器的弯曲变形理论计算值及有限元仿真值

Fig. 8 Theoretical calculations and finite element simulations with bending deformation of multi-directional actuator

表 3 多向驱动器末端点坐标及相对误差

Table 3 Coordinates of multi-directional actuator and its relative error

$p/\text{kPa}$	$x_F/\text{mm}$	$y_F/\text{mm}$	$x_J/\text{mm}$	$y_J/\text{mm}$	相对误差/%
46	26.296	87.974	20.920	86.671	5.71
48	31.726	86.194	27.761	84.011	7.21
50	38.361	83.271	34.809	80.289	8.89
52	46.248	78.535	41.7575	75.426	7.65
54	55.470	70.747	48.307	69.396	9.63
56	63.735	60.534	54.0700	62.363	10.52

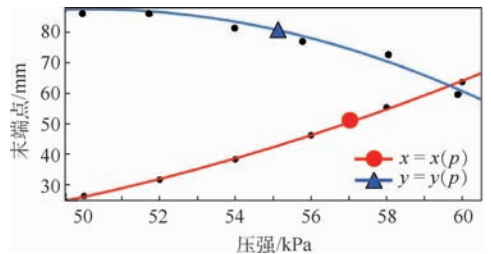


图 9 多向驱动器末端点坐标拟合曲线

Fig. 9 Fitting curves with coordinate points of multi-directional actuator

### 3.2 单向驱动器弯曲变形

同理,以单向驱动器为研究对象,图 10 给出了不同驱动压强作用下单向驱动器弯曲变形的理论计算值及有限元仿真值。

同理,为了更加清楚地表征单向驱动器的弯曲状况,表 4 展示了不同驱动压强作用下其弯曲变形所对应圆心角的理论计算值、有限元仿真值及相对误差。

同理,值得指出的是,上述数据也只能表征固定驱动压强作用下单向驱动器的弯曲状态,为从一般意义上来表征其圆心角随压强的变化关系,如图 11 所示,可将结果进行轨迹拟合,拟合结果为

$$\theta_r(p) = 87260p^2 - 2757p + 22.23 \quad (18)$$

基于仿真结果,单向驱动器弯曲圆心角的有限元仿真值与实验值相对误差基本不超过 10% ;

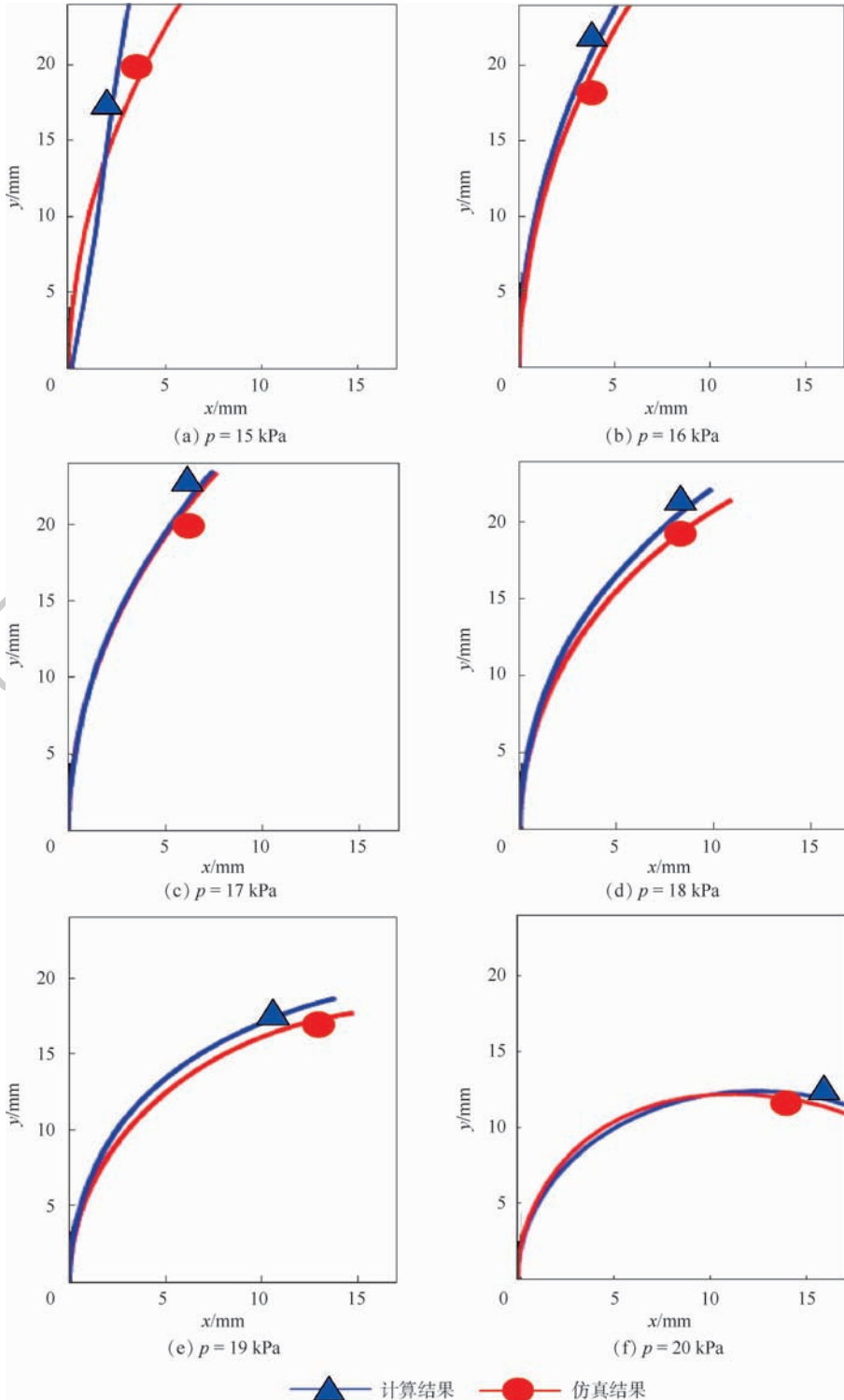


图 10 单向驱动器的弯曲变形理论计算值及有限元仿真值

Fig. 10 Theoretical calculations and finite element simulations with bending deformation of one-directional actuator

表4 单向驱动器弯曲圆心角及相对误差

Table 4 Bending center angle of one-directional actuator and its relative error and its relative error

$p/\text{kPa}$	$\theta_r/\text{rad}$	$\theta_j/\text{rad}$	相对误差/%
15	0.395 348	0.421 403	6.590 41
16	0.423 527	0.448 456	5.886 147
17	0.632 721	0.602 154	4.831 062
18	0.939 089	0.852 957	9.171 863
19	1.383 204	1.355 69	1.989 166
20	2.018 431	2.198 672	8.929 779

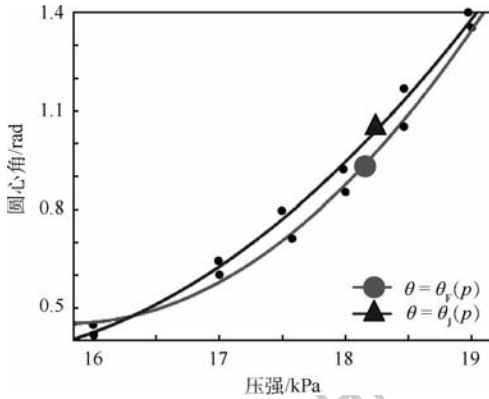


图11 单向驱动器弯曲圆心角拟合曲线

Fig. 11 Fitting curves with bending center angle of one-directional actuator

单向驱动器带动仿生舌可完成 2 rad 运动,符合预期 1.3 rad 的目标。上述分析证明了单向驱动器弯曲变形理论模型的正确性。

## 4 实验验证

为进一步验证理论模型与有限元仿真的正确性,本文开展了软体仿生舌运动特性的实验研究,实验平台如图 12 所示。小型隔膜泵(DC12 V)由软管分别连接软体驱动器和 Y100 气压表(100 kPa),电源(DC12V3A)连接直流电机调速板(DC5-18V)控制气泵的输出压强。实验过程中

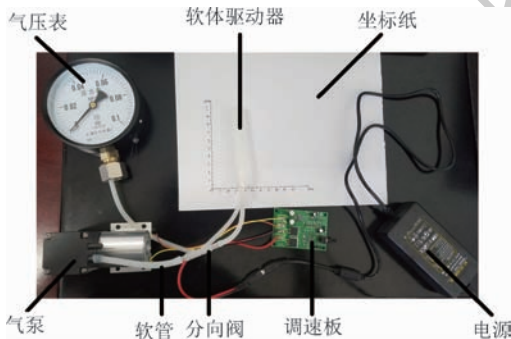


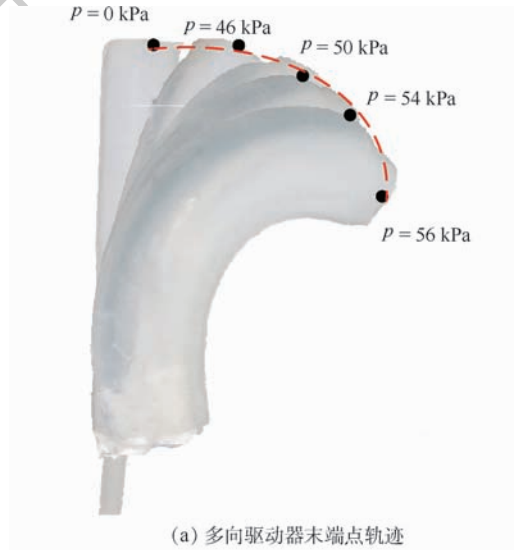
图12 软体仿生舌运动特性实验平台  
Fig. 12 Experiment platform with motion characteristics of soft bionic tongue

主要考虑驱动器在平面内的弯曲运动,故驱动器充气端固定,而运动端进行数据采集,并由坐标纸得到相关实验数据。

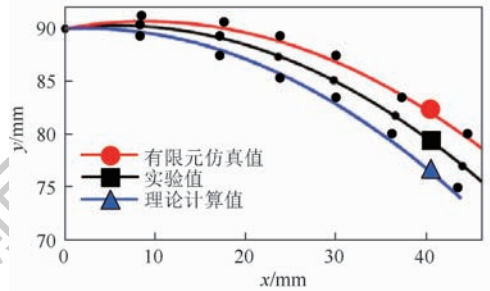
### 4.1 驱动器弯曲实验

基于上述实验平台,图 13(a)展示了不同驱动压强作用下多向驱动器的弯曲变形状态。联列表 3,多向驱动器在不同驱动压强作用下末端点坐标的有限元仿真值、实验值和理论计算值如图 13(b)所示。

同理,图 14 展示了不同驱动压强作用下单向驱动器的弯曲状态,其理论计算、有限元仿真及实验结果如表 5 所示。



(a) 多向驱动器末端点轨迹



(b) 多向驱动器末端点理论计算、有限元仿真及实验点轨迹

图13 多向驱动器弯曲实验结果

Fig. 13 Bending experiment result of multi-directional actuator

### 4.2 软体仿生舌运动

为验证软体仿生舌的助咀嚼、助吞咽功能,可进行软体仿生舌的翻舌、斜翻舌、吐舌及卷舌动作实验,图 15(a)和图 15(b)分别展示了仿真与实验结果。实验仍采用上述实验平台,其中,对多向驱动器下面(图示位置)2个软管充气完成翻舌运动;对单个软管充气完成仿生舌斜上拉运动;对

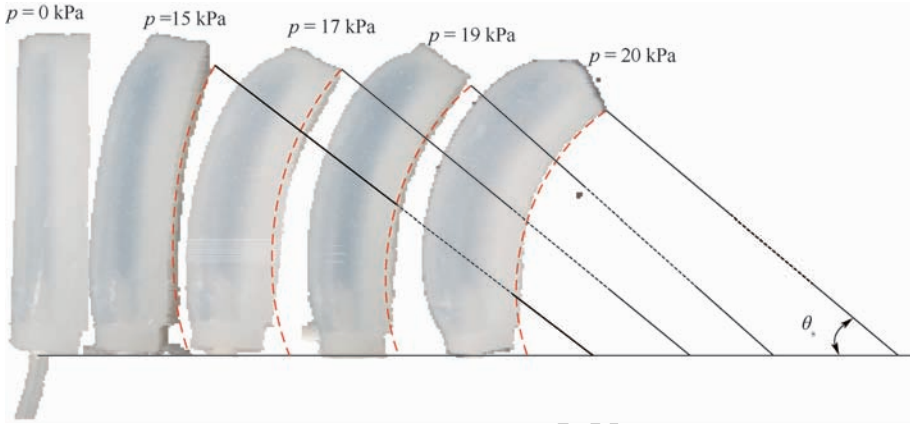


图 14 单向驱动器弯曲圆心角

Fig. 14 Bending center angle of one-directional actuator

表 5 单向驱动器弯曲圆心角实验结果

Table 5 Experimental results with bending center angle of one-directional actuator

$p/\text{kPa}$	$\theta_j/\text{rad}$	$\theta_f/\text{rad}$	$\theta_s/\text{rad}$
15	0.382 480 028	0.421 403 2	0.4
16	0.416 491 474	0.448 456	0.4
17	0.515 337 908	0.602 154	0.5
18	0.773 391 507	0.852 957	0.7
19	1.241 311 947	1.355 69	1.2
20	2.411 736 242	2.198 672	1.9

4个软管同时充气则可以完成吐舌动作。

同理,改变2个软管充气的位置则可以实现软体仿生舌左右摆动及翻舌运动;改变单个软管的充气位置则可以实现不同方向的斜翻舌;进而,通过上述多种运动方式的组合,可以完成仿生舌的助咀嚼、助吞咽功能,表6展示了  $p = 50 \text{ kPa}$  仿生舌各项的实验与仿真结果。

基于上述实验结果,软体仿生舌可基本实现翻舌、吐舌等各项运动,且均符合预期吐舌  $5 \text{ mm}$ ,翻舌  $1.2 \text{ rad}$ 、斜翻舌  $1 \text{ rad}$  及卷舌  $1.3 \text{ rad}$  的目标。上述分析进一步证明了软体仿生舌结构设计、指定运动的合理性及理论模型的正确性。

该软体仿生舌的柔软性与真实的舌具有高相似性,其可应用于模拟吞咽功能,对治疗吞咽困难症具有重要意义。为精准有效的对仿生舌进行控制,需搭建仿生舌的控制系统,目前基于 Arduino 控制系统的构建尚在进行之中。

表 6 软体仿生舌实验

Table 6 Experiment of soft bionic tongue

运动类型	实验结果	仿真结果
翻舌/rad	1.4	1.520 312
斜上翻/rad	1.2	1.310 291
斜下翻/rad	1.1	1.253 652
左摆舌/rad	1.3	1.492 12
右摆舌/rad	1.3	1.492 12
吐舌/mm	6.4	5.933 56
卷舌/mm	1.2	1.426 356

## 5 结论

1) 本文基于单/多向气动驱动器设计了一款软体仿生舌,该软体仿生舌可完成吐舌、翻舌、卷舌以及斜翻舌等动作。

2) 基于 Yeoh 模型应变能密度函数,结合力平衡方程,建立了单/多向气动驱动器的弯曲变形

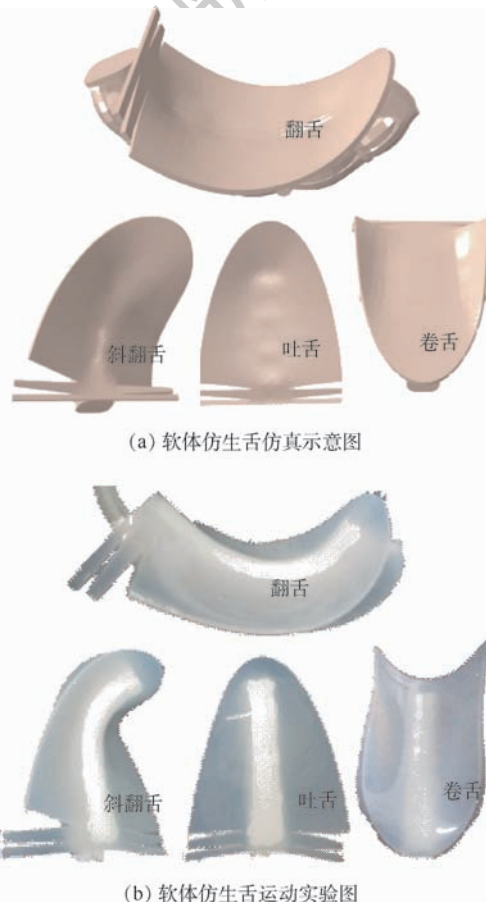


图 15 软体仿生舌实验

Fig. 15 Experiment of soft bionic tongue

理论模型。进一步开展了气动驱动器的有限元仿真研究,验证了理论模型的正确性。

3) 开展了软体仿生舌的实验研究,验证了单/多向驱动器弯曲理论计算、仿真值与实验结果的一致性,也验证了软体仿生舌各运动的可行性。

### 参考文献 (References)

- [1] 曹玉君,尚建忠,梁科山,等. 软体机器人研究现状综述[J]. 机械工程学报,2012,48(3):25-33.  
CAO Y J, SHANG J Z, LIANG K S, et al. A review on the soft-robotics[J]. Journal of Mechanical Engineering, 2012, 48(3): 25-33 (in Chinese).
- [2] 张忠强,邹娇,丁建宁,等. 软体机器人驱动研究现状[J]. 机器人,2018,40(5):648-659.  
ZHANG Z Q, ZOU J, DING J N, et al. Research status of the soft robot driving[J]. Robot, 2018, 40(5): 648-659 (in Chinese).
- [3] RUS D, TOLLEY M T. Design fabrication and control of soft robots[J]. Nature, 2015, 521(7553):467-475.
- [4] 谢哲新,龚哲元,王田苗,等. 可控三维运动的软体驱动器仿真与试验[J]. 机械工程学报,2018,54(21):460-468.  
XIE Z X, GONG Z Y, WANG T M, et al. Simulation and experiment sofa controllable soft spatial fluidic elastomer manipulator [J]. Journal of Mechanical Engineering, 2018, 54(21): 460-468 (in Chinese).
- [5] 苏满佳,张逸鸿,谢荣臻,等. 连续软体机器人的结构范型与形态复现[J]. 机器人,2018,40(5):640-647.  
SU M J, ZHANG Y H, XIE R Z, et al. Structural pattern and shape reproduction for soft continuum robot [J]. Robot, 2018, 40(5): 640-647 (in Chinese).
- [6] 樊继壮,于庆国,袁博文,等. 软体仿蛙游动机器人关节式气动致动器研制[J]. 机器人,2018,40(5):578-586.  
FAN J Z, YU Q G, YUAN B W, et al. Development of a joint-like pneumatic actuator applied to soft frog-inspired swimming robot [J]. Robot, 2018, 40(5): 578-586 (in Chinese).
- [7] ILIEVSKI F, MAZZEO A D, SHEPHERD R F, et al. Soft robotics for chemists[J]. Angewandte Chemie, 2011, 50(8):1890.
- [8] 胡兵兵,金国庆. 一种仿虎甲幼虫的多驱动器软体机器人的设计与制造[J]. 机器人,2018,40(5):626-633.  
HU B B, JIN G Q. Design and fabrication of a multi-actuator soft robot inspired by young tigerbeetle [J]. Robot, 2018, 40(5): 626-633 (in Chinese).
- [9] DEIMEL R, BROCK O. A novel type of compliant and underactuated robotic hand for dexterous grasping [J]. International Journal of Robotics Research, 2015, 35(1):161-185.
- [10] TOLLEY M T, SHEPHERD R F, MOSADE B, et al. A resilient-tethered soft robot [J]. Soft Robotics, 2014, 1(3):213-223.
- [11] WEHN M, TRUBY R L, FITZGERALD D J, et al. An integrated design and fabrication strategy for entirely soft autonomous robots [J]. Nature, 2016, 536(7617):451-455.
- [12] BARTLETT N W, TOLLEY M T, OVERVELDE J T, et al. A 3D-printed functionally graded soft robot powered by combustion [J]. Science, 2015, 349(6244):161-165.
- [13] LASCHI C, CIANCHETTI M, MAZZOLAI B, et al. Soft robot arm inspired by the octopus [J]. Advanced Robotics, 2012, 26(7):709-727.
- [14] SHINTAKE J, ROSSET S, SCHUBERT B, et al. Versatile soft grippers with intrinsic electroadhesion based on multifunctional polymer actuators [J]. Advanced Materials, 2016, 28(2):231-238.
- [15] SHEN Q, TRABIA S, STALBAUM T, et al. A multiple-shape memory polymer-metal composite actuator capable of programmable control, creating complex 3D motion of bending twisting and oscillation [J]. Scientific Reports, 2016, 6(2):24462.
- [16] 郭闯强,吴春亚,刘宏. 离子聚合物金属复合材料驱动器在机器人中的应用进展 [J]. 机械工程学报, 2017, 53(9): 1-13.  
GUO C Q, WU C Y, LIU H. The application progress of ionic-polymer-metal composite materials in robotics [J]. Journal of Mechanical Engineering, 2017, 53(9): 1-13 (in Chinese).
- [17] CONNOLL Y, FIONNUALA W, CONOR J B. Automatic design of fiber-reinforced soft actuators for trajectory matching [J]. Proceedings of the National Academy of Sciences of the United States of America, 2017, 114(1):51-56.
- [18] KEVIN C G, KAITLYN P B, BRENNAN P. Soft robotic grippers for biological sampling on deep reefs [J]. Soft Robotics, 2016, 3(1):2169-5172.
- [19] KRISTIN M, OLIVER M, REILLY O. On constitutive relations for a rod-based model of a pneu-net bending actuator [J]. Extreme Mechanics Letters, 2016, 8(2):38-46.
- [20] 王宁扬,孙昊,姜皓,等. 一种基于蜂巢气动网络的软体夹持器抓取策略研究 [J]. 机器人, 2016, 38(3): 371-384.  
WANG N Y, SUN H, JIANG H, et al. Study on grabbing algorithm of soft gripper based on honeycomb pneumatic network [J]. Robot, 2016, 38(3): 371-384 (in Chinese).
- [21] 魏树军,王天宇,谷国迎. 基于纤维增强型驱动器的气动软体抓手设计 [J]. 机械工程学报, 2017, 53(13): 29-38.  
WEI S J, WANG T Y, GU G Y. Design of pneumatic soft grips based on fiber reinforced drivers [J]. Chinese Journal of Mechanical Engineering, 2017, 53(13): 29-38 (in Chinese).
- [22] 郑俊君,宋小波,姜祖辉. 一种气动静压软体机器人的驱动力产生机理及控制策略 [J]. 机器人, 2014, 36(5): 513-518.  
ZHENG J J, SONG X B, JIANG Z H. The mechanism and control strategy of driving force of a pneumatic hydrostatic soft robot [J]. Robot, 2014, 36(5): 513-518 (in Chinese).
- [23] 王华,康荣杰,王兴坚. 软体弯曲驱动器设计与建模 [J]. 北京航空航天大学学报, 2017, 43(5): 1053-1060.  
WANG H, KANG R J, WANG X J. Design and modeling of soft bending driver [J]. Journal of Beijing University of Aeronautics and Astronautics, 2017, 43(5): 1053-1060 (in Chinese).
- [24] 费燕琼,庞武,于文博. 气压驱动软体机器人运动研究 [J]. 机械工程学报, 2017, 53(13): 14-18.  
FEI Y Q, PANG W, YU W B. Study on motion of air-driven soft robot [J]. Journal of Mechanical Engineering, 2017, 53(13): 14-18 (in Chinese).
- [25] KIM H, HAN M W, SONG S H, et al. Soft morphing hand driven by SMA tendon wire [J]. Composites Part B: Engineering, 2016, 105(16):138-148.

- [26] MATHIJSSSEN G, SCHULTZ J, VANDER B, et al. A muscle-like recruitment actuator with modular redundant actuation units for soft robotics[J]. *Robotics and Autonomous Systems*, 2015, 74(15):40-50.
- [27] 史震云, 朱前成. 基于智能驱动器的软体机器人系统[J]. *北京航空航天大学学报*, 2016, 42(12):2596-2602.  
SHI Z Y, ZHU Q C. Based on the intelligent driver of the soft-robot system[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2016, 42(12):2596-2602 (in Chinese).
- [28] 王扬威, 兰博文, 刘凯. 形状记忆合金丝驱动的柔性机械臂建模与实验[J]. *浙江大学学报*, 2018, 52(4):628-634.  
WANG Y W, LAN B W, LIU K. Modeling and experiment of flexible manipulator actuated by shape memory alloy wire[J]. *Journal of Zhejiang University*, 2018, 52(4):628-634 (in Chinese).
- [29] FEI Y Q, XU H W. Modeling and motion control of a soft robot [J]. *IEEE Transactions on Industrial Electronics*, 2017, 64(2):1737-1742.
- [30] MUTLU R, ALICI G, LI W H. A soft mechatronic microstage mechanism based on electroactive polymer actuators[J]. *IEEE/ASME Transactions on Mechatronics*, 2016, 21(3):1467-1477.
- [31] 谢高鹏. 食品咀嚼模拟机系统的开发与研究[D]. 长春: 吉林大学, 2014:10-15.  
XIE G P. The research and development of food chewing simulator system[D]. Changchun: Jilin University, 2014:10-15 (in Chinese).
- [32] 韩敬虎. 咀嚼吞咽一体化仿生机器人研究[D]. 无锡: 江南大学, 2016:13-19.  
HAN J H. Research on integrated bionic robot for chewing and swallowing[D]. Wuxi: Jiangnan University, 2016:13-19 (in Chinese).
- [33] 黄建龙, 解广娟, 刘正伟. 基于 Mooney-Rivlin 模型和 Yeoh 模型的超弹性橡胶材料有限元分析[J]. *橡胶工业*, 2008, 55(8):467-471.  
HUANG J L, XIE G J, LIU Z W. The finite element analysis of super-elastic rubber material based on Mooney-Rivlin model and Yeoh model[J]. *China Rubber Industry*, 2008, 55(8):467-471 (in Chinese).
- [34] KOTHERA C S, JANAID M, SIROHI J, et al. Experimental characterization and static modeling of McKibben actuators[J]. *Journal of Mechanical Design*, 2009, 131(9):1010-1050.

#### 作者简介:

董虎 男, 硕士研究生。主要研究方向: 软体机器人。

林苗 男, 硕士研究生。主要研究方向: 柔性机构学。

顾苏程 男, 硕士研究生。主要研究方向: 软体机器人。

曹毅 男, 博士, 教授。主要研究方向: 并联机器人、混联机器人、柔性机器人、软体机器人。

李巍 男, 博士。主要研究方向: 机器人深度学习。

## Motion characteristics of soft bionic tongue based on multi-directional pneumatic actuator

DONG Hu<sup>1,2</sup>, LIN Miao<sup>1,2</sup>, GU Sucheng<sup>1,2</sup>, CAO Yi<sup>1,2,\*</sup>, LI Wei<sup>3</sup>

(1. School of Mechanical Engineering, Jiangnan University, Wuxi 214122, China;

2. Jiangsu Key Laboratory of Advanced Food Manufacturing Equipment and Technology, Wuxi 214122, China;

3. Suzhou Vocational Institute of Industrial Technology, Suzhou 215104, China)

**Abstract:** Aimed at multi-directional actuator deformation mechanism, this paper takes the soft bionic tongue as research object and the research on the motion characteristics of one/multi-directional actuator for the soft bionic tongue is carried out. Firstly, a soft bionic tongue driven by pneumatic actuator is designed, which can complete many actions including tongue stretching, twisting, rolling and tilting by the one/multi-directional actuator. Secondly, in order to further study the actuator's deformation mechanism, the research on the structure and deformation operating principle of one/multi-directional actuator is carried out. Then, a non-linear mathematical model of radius of a soft bionic tongue actuated by specific pressure is established, which is based on the Yeoh model strain energy density function and combined with the force balance equation. Finally, the finite element simulation and experimental validation of the soft bionic tongue are carried out to verify the correctness of the theoretical model. Consequently, the above research provides a theoretical basis for other research on the deformation mechanism of pneumatically actuated soft structures.

**Keywords:** soft bionic tongue; air-driven; Yeoh model; finite element simulation; deformation mechanism

**Received:** 2019-01-02; **Accepted:** 2019-03-08; **Published online:** 2019-04-14 07:02

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20190411.1643.002.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190411.1643.002.html)

**Foundation items:** National Natural Science Foundation of China (50905075); the Six Talent Peaks Project in Jiangsu Province (ZBZZ-012); Overseas Expertise Introduction Project for Discipline Innovation (B18027)

\* **Corresponding author.** E-mail: caoyi@jiangnan.edu.cn

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

DOI: 10.13700/j.bh.1001-5965.2018.0789

# 基于强化学习的时间触发通信调度方法



李浩若, 何锋\*, 郑重, 李二帅, 熊华钢

(北京航空航天大学 电子信息工程学院, 北京 100083)

**摘 要:** 未来航空电子系统中将会更广泛地选择基于时间触发的通信机制进行信息传输, 以保证信息交互的确定性。如何合理地进行时间触发通信调度设计是时间触发应用于航空电子互连系统的关键。针对时间触发调度的周期性任务, 提出了一种基于强化学习的周期调度时刻表生成方法。首先, 将流量调度任务转换为树搜索问题, 使之具有强化学习所需要的马尔可夫特性; 随后, 利用基于神经网络的强化学习算法对调度表进行探索, 不断缩短延迟时间以优化调度表, 且在训练完成后, 可以直接使用到消息分布相近的任务中。与使用 Yices 等可满足模理论 (SMT) 形式化求解时间触发调度表方法相比, 所提方法不会出现无法判定的问题, 能够保证时间触发调度设计结果的正确性和优化性。对于包含 1 000 条消息的大型网络, 所提方法的计算速度为 SMT 方法的数十倍以上, 并且调度生成消息的端到端延迟在 SMT 方法的 1% 以下, 大大提高了消息传输的及时性。

**关键词:** 时间触发; 调度方法; 强化学习; 树搜索; 偏置时间

**中图分类号:** V247.5; TN914

**文献标识码:** A

**文章编号:** 1001-5965(2019)09-1894-08

时间触发以太网 (Time-Triggered Ethernet, TTE) 利用时钟同步服务保证全网在统一时钟下进行时间触发通信交互: 通过离线调度表的规定时刻发送和转发时间触发 (Time-Triggered, TT) 消息, 保证 TT 消息传输的实时性和确定性<sup>[1-2]</sup>。

时间触发网络中 TT 流量的调度问题常用可满足模理论 (Satisfiability Modulo Theories, SMT) 进行求解, 通过设置网络拓扑和流量关系等约束, 生成满足约束的 TT 时间触发调度表<sup>[3]</sup>。此方法的缺点是不能直接对得出的调度表进行优化, 且计算速度过于缓慢, 对于实际的大型拓扑很难在短时间内得到封闭解, 并且很难在短时间内适应设计的变化。文献[4]采用了增量化调度方法, 但在增量化过程中却并没有给出有效的指导信息或启发式函数, 致使最终调度表性能无法优化。

还存在部分方法采用遗传算法等进行优化, 但是此类方法常常相当耗费时间, 且完全无法泛化, 对于稍有改变的消息分布即需要完全重新计算, 缺乏计算效率<sup>[4-5]</sup>。

近年来, 基于神经网络强化学习在搜索、规划等领域取得了非常大的成功, 对于旅行商人等 NP 问题均有较好的表现<sup>[6-7]</sup>。在围棋领域, 强化学习与树搜索结合的 AlphaGo 超越了人类水平<sup>[8]</sup>。强化学习的方法也将会被用到更多的场景。

本文贡献在于将基于增量化调度方法的调度任务用马尔可夫过程来完全描述, 将每次增量化调度用状态、动作、奖励表示, 以此在搜索过程中提供尽量多的辅助搜索信息。同时, 利用深度强化学习的方法, 结合树搜索, 以尽量减小延迟时间为目标, 寻找近似最优的时间触发调度表。并且,

收稿日期: 2019-01-02; 录用日期: 2019-03-15; 网络出版时间: 2019-05-08 09:41

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190505.1020.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190505.1020.001.html)

基金项目: 国家自然科学基金 (61301086, 71701020); 中国民航大学天津市民用航空器适航与维修重点实验室开放基金 (2017SW02)

\* 通信作者. E-mail: robinleo@buaa.edu.cn

**引用格式:** 李浩若, 何锋, 郑重, 等. 基于强化学习的时间触发通信调度方法[J]. 北京航空航天大学学报, 2019, 45(9): 1894-1901. LI H R, HE F, ZHENG Z, et al. Time-triggered communication scheduling method based on reinforcement learning[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1894-1901 (in Chinese).

模型具备一定的泛化能力,可在较短时间内响应消息分布近似的时间触发调度求解问题;在消息分布与之前完全不同时,可以退化为随机搜索方法,也能较快地得出一个可行解。

## 1 流量调度方法

### 1.1 马尔可夫模型与强化学习

由于消息的帧长、周期、偏置时间各不相同而导致搜索特别复杂。但是从调度的角度看,每一次调度一条消息均可以看作一个状态,当前状态可以反映所有信息,不需要考虑历史状态、动作等。以上调度过程非常符合马尔可夫过程的特点,在这个基础上,可以将此转换为一个马尔可夫模型,用树搜索的方法处理,每一个状态即为一个节点。此时这个问题也能看作一个条件优化问题<sup>[9]</sup>,在特定的约束条件下满足目标或最大化收益。

由于使用到强化学习,定义状态、动作、奖励如下。

**状态** 状态定义为链路目前的排布方式。状态尽量表示当前调度表的所有特征。但是状态的表示不能太过复杂以致于计算过于缓慢。

**动作** 动作定义为当前消息在当前链路中的偏置。这里将每条消息的每一项偏置分离,分为第一项偏置和其余偏置。第一项可决定消息的传输位置信息,后面的偏置均为在可得出调度表的情况下越小越好。

**奖励** 奖励定义为这次调度的成功与否和端到端延迟的数值。在调度过程中,很难知道当前的调度方式是否为全局最优,采用如下方法:在一次调度结束后,从总体情况判断应该给出的奖励。奖励值的大小可以根据全局延迟时间的大小进行确定。

图1所示为强化学习的基本原理。将调度过程转换为一个具有马尔可夫特性的过程,将整个调度的信道作为强化学习的环境。首先根据状态使用动作与周围环境进行交互,在动作和环境的作用下,可以得到新的环境,同时环境会给出奖励信号,以反馈在这个环境下使用这个动作的价

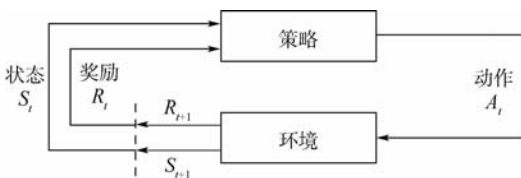


图1 强化学习的基本原理

Fig. 1 Fundamental principle of reinforcement learning

值。如此循环下去,策略与环境进行不断地交互从而产生新的数据。强化学习算法利用产生的数据不断优化自身使用动作的选择,并再与环境交互,产生新的数据,并利用新的数据进一步改善自身动作的策略。经过数次迭代学习后,最终可以学习到完成相应任务的最优策略。

但是在很多情况下,仅仅依靠随机探索很难得到有用的样本,需要使用树搜索的方法加强探索,使在训练初期尽快得到较好的样本训练<sup>[10]</sup>。

### 1.2 搜索策略

在本文提出的调度方法的搜索框架中采用深度优先与最优优先结合的搜索,并剪掉明显不能找到结果的分支,以尽快找到可行解,如算法1所示,此搜索框架仅考虑每条消息在第一条链路中的偏置时间,消息在后续转发节点中的偏置采用算法2的策略来选择,以完成消息在整个传输路径上的偏置区间搜索。

**算法1** 消息调度方法(SCHEDULE)。

```

1  初始化最大探索次数  $t_{\max}$ , 消息数目  $x$ 。
2   $i = 0, t = 0$ 
3  while  $i < x$ 
4      while  $t < t_{\max}$ 
5          search()
6          if  $t = t_{\max}$ 
7              backtracking()
8               $t = 0$ 
9          else
10              $t + = 1$ 
11          $i + = 1$ 

```

注: search() 为搜索偏置可用时间片; backtracking() 为回溯。

**算法2** 各级单跳延迟时间的选择。

```

1  初始化探索次数  $t$ , 最大探索次数  $t_{\max}$ ,
2  技术延迟 lower, 最坏延迟上限 upper, 参数 alpha。
3  search()
4  while  $t < t_{\max}$ 
5      diff = alpha * (upper - lower) *  $t/t_{\max}$ 
6      higher = lower + diff
7      delay = random_choice(lower, higher)
8       $t + = 1$ 
9  backtracking()

```

注: random\_choice 为在这个范围内随机选择。

对于后续转发节点的偏置时间,有3种选择方法:①完全随机搜索;②完全贪婪搜索;③两者

结合,进行限制选择区间的随机搜索。方法①不对端到端延迟作额外限制,这样容易也能够较快地得出结果,但是相应的所得到的调度表可能存在较大端到端延迟。方法②严格限制端到端延迟,如果能搜索出结果,那么此结果即为最优结果,但是很多情况下根本不存在这样的调度表,或者需要近乎遍历所有状态,造成开销过于庞大。方法③为两者结合,将搜索的端到端延迟限制在一定的置信区间,在这个区间内进行随机搜索。如果能得到一个调度表,那么将这—个调度表作为一个样本。这种方法的优势在于,搜索轨迹不会快速收敛到局部最优;得到的调度表不会存在特别大的端到端延迟,对于进一步探索和训练有很大的促进作用;同时此方法存在置信区间,可以非常方便地调整,既不会难以得出调度表结果,也不会耗费大量时间,并可以在区间中继续增加条件,使搜索成功率更高,以得到更优的调度表。

### 1.3 强化学习训练

在搜索阶段,可以迅速得到端到端延迟较小的调度表,但是由于搜索策略的限制,这样的调度表仍然有很大的优化空间。希望根据输入消息的分布的不同,能得出相应分布更好的消息调度表,在本文中使—用强化学习的方法进行训练。由于消息调度问题为 NP 问题<sup>[11]</sup>,难以求得全局最优解,可以认为不存在样本的标签,只能采用强化学习的方法进行训练和测试。

在训练阶段,由于消息的偏置可选择的范围太大,计算梯度时会造成梯度太大而无法优化,在具体计算时对其采用对数化的处理方法。同时考虑到同样的方差对较大偏置和较小偏置的贡献并不—样,本文对不同偏置处的方差影响通过权重进行度量。用初始搜索所得样本对神经网络进行预训练,将已得到的轨迹赋予标签。在这个阶段,目标函数采用改进的均方差计算:

$$U = \sum_i (\text{out}(t) - \text{action}(t))^2 \times \text{action}(t) \quad (1)$$

此时已经得到一整条训练轨迹,即可在每个状态  $\text{state}(t)$ ,使神经网络输出  $\text{out}(t)$  逼近  $\text{action}(t)$ ,不能只逼近所提供的样本,相应地需要对输出提供—定的噪声,以此作进一步的探索和适应。取神经网络的输出作为中心,取一个宽度有限的区间,在这个区间中均匀采样作为消息偏置,将此操作记作  $\text{linear}$  函数。详细算法如算法 3 所示。

#### 算法 3 模仿学习方法 (IMITATION)。

1 初始化神经网络参数,总训练次数  $K_{\max}$ ,

```

2 端到端延迟阈值 end2end。
3  $k = 0$ 
4 while  $k < K_{\max}$ 
5   delay = SCHEDULE(linear())
6   if delay < end2end
7     end2end = delay
8   train(L) use(1)
9    $k + = 1$ 

```

对于 TT 消息而言,当消息数量增加、帧长增长时,初始偏置对于调度表设计将产生直接的影响。虽然消息的偏置为一个—个确定的、离散的值,但是由于可以选择的值太多,动作空间太大,所以将它化为连续值来处理,处理完毕后将—其离散化得到调度表。

可以采用策略梯度的方法<sup>[12-13]</sup>。在最初收集轨迹阶段采用改进的贪婪策略,在每一个状态处都采取尽量最小化端到端延迟的动作,并用神经网络将策略参数化,寻找最优的参数使强化学习的目标最大,即最小化所有消息的端到端延迟,直到算法收敛或迭代次数到达上限的时候即算法收敛。此时搜索空间相比于初始阶段较小,在这个阶段应使用较小学习率。

首先设置目标函数为

$$U = \sum_i P(t; \theta) \times \text{reward}(t) \quad (2)$$

式中:  $P(t; \theta)$  为  $t$  时刻,策略选择到已执行动作  $\text{action}$  的概率;  $\text{reward}(t)$  为在时刻  $t$  执行了  $\text{action}$  后所得到的奖励值。那么  $U$  的含义即为:这一条轨迹所得到的总的奖励值。损失函数即为对参数化后的目标函数求梯度,以使目标函数以最快的速度向最大化目标函数的方向收敛:

$$\begin{aligned} \nabla U &= \sum_i \nabla_{\theta} P(t; \theta) \times \text{reward}(t) = \\ & \sum_i \frac{P(t; \theta)}{P(t; \theta)} \nabla_{\theta} P(t; \theta) \times \text{reward}(t) = \\ & \sum_i P(t; \theta) \frac{\nabla_{\theta} P(t; \theta) \times \text{reward}(t)}{P(t; \theta)} = \\ & \sum_i P(t; \theta) \nabla_{\theta} (\log_{\theta} P(t; \theta)) \times \text{reward}(t) \end{aligned} \quad (3)$$

在强化学习训练过程中,由于需要不断探索,于是将输出值采用正态分布描述,使神经网络的输出为此正态分布的均值和方差。这样既能使效果较好的能被更多选择到,也能同时探索更丰富的偏置值,在较优值附近常能得到更优选择,同时亦有小概率探索距离较远的样本,不会过早收敛。将这种输出分布的选择记作  $\text{normal}$ ,如算法 4

所示。

#### 算法4 强化学习方法(REINFORCEMENT)。

- 1 初始化神经网络参数,总训练次数  $K_{\max}$ 。
- 2  $k = 0$
- 3 while  $k < K_{\max}$
- 4     reward = SCHEDULE(normal)
- 5     train( $U$ ) use (3)
- 6      $k + = 1$

这个阶段的探索,需要限制探索方向,首先这是一个条件优化问题,不能进行完全随机探索。文献[12,14]采用的均为带惩罚项的优化策略,但是实际操作中不能出现不符合要求的解。其次探索相对的随机性不一定能得到更优的轨迹,样本利用率十分低。于是选择向下模拟数步搜索,即  $n$  步展开的方法。较多算法常采用 1 步展开,即每一步均判断是否符合需求,但是由于探索的深度浅,模型会较为短视,且一定程度上会由于阈值选择不够合理而造成搜索成功率降低;后者深度较深,模型较为有远见,阈值可随深度的变化而加以处理,但是计算量较大。同时,两者均能试探是否可以在探索过程中找到可行解。

这里拓展  $n$  步展开的方法,每到一个状态时,模拟之后  $n$  步搜索。如果在  $n$  步以后可以满足预先设定的端到端延迟最低要求,那么采用这一条轨迹继续进行;如果没有满足或消息无法调度,那么选择其他轨迹。此时预先设定的阈值可以使用动态方法确定,根据搜索的深度进行自适应<sup>[15]</sup>。同时需要选择模拟的最大深度  $n$ ,如果  $n$  小,容易过早收敛到由初始值所限制的轨迹中,进而陷入局部最优解;如果  $n$  过大,计算开销则会非常大。可根据消息的总数等特征来确定。

由于是基于梯度的方法,在这个阶段的搜索具有较强的方向性;同时具备较强鲁棒性,在下次遇到类似状态时,将会自适应地选择动作,并且策略能够不断进行学习和优化。

强化学习的训练方法分为时间差分方法和蒙特卡罗方法,由于本任务目标为得到全局最优调度表,在搜索过程中很难判断是否是一个阶段性的优秀的选择,所以采用蒙特卡罗方法进行训练。完整得出一条轨迹以后,再对整个搜索过程进行评估和训练。这种方法训练可能会使梯度有较大的方差,但是梯度是无偏的。

此时的训练可以继续采用模仿学习的方法,但搜索区域需要更为严格地限制,采用分段函数描述算法3中的 linear 函数。在策略梯度方法中常常会限制策略更新的速度,防止更新过快造成

无法收敛的情况<sup>[12,14]</sup>,这里采用限制搜索区间宽度的方法。随着探索和训练的进行,由于算法逐渐收敛,大多数的探索所得结果均对训练没有帮助,可以适当改变区间宽度以探索更多区域。此时 linear 函数表示为

$$\begin{cases} \begin{cases} \text{upper}_1 = w_{1u} \times \text{offset} + b_{1u} \\ \text{lower}_1 = w_{1l} \times \text{offset} + b_{1l} \end{cases} & 0 < \text{back} < \text{back}_1 \\ \begin{cases} \text{upper}_2 = w_{2u} \times \text{offset} + b_{2u} \\ \text{lower}_2 = w_{2l} \times \text{offset} + b_{2l} \end{cases} & \text{back}_1 < \text{back} < \text{back}_2 \\ \vdots & \\ \begin{cases} \text{upper}_n = w_{nu} \times \text{offset} + b_{nu} \\ \text{lower}_n = w_{nl} \times \text{offset} + b_{nl} \end{cases} & \text{back}_{n-1} < \text{back} < \text{back}_n \end{cases} \quad (4)$$

式中:upper 为随机探索区间的上界;lower 为此区间的下界;back <sub>$k$</sub>  ( $0 < k \leq n, k \in \mathbf{N}$ ) 为回溯次数; $w_{ku}, b_{ku}$  ( $0 < k \leq n, k \in \mathbf{N}$ ) 为每个区间所对应的参数,可事先人为确定。

由于将偏置值进行了对数化,不同的偏置量需要采用不同的探索区间。具体算法同样如算法3所示。

## 2 实验

### 2.1 实验条件

设链路速率为 100 Mbit/s,以太网帧的帧长范围为 64 ~ 1 518 Byte,则每一帧传输的执行时间范围为 6.72 ~ 123.04  $\mu\text{s}$ ,TT 流量的周期通常为 1 ~ 128 ms。为使环境既考虑逼真度,又考虑到复杂度,将消息执行周期与时间均取整数,参照 ARINC 664 P7 协议中对虚拟链路最小帧间间隔的约定,取流量周期仅为 2 的幂次<sup>[16]</sup>。采用空客 A380 客机的核心交换结构如图 2 所示,其中 SW 为交换机,ES 为端系统。与 A380 的消息规模类似,采用 1 050 条消息进行实验验证,规定好传输路线,仅进行调度不进行路径规划。即默认调度前已用算法优化选择传输链路,典型地,可以采用最短路径方法进行路径生成。

为计算方便,将所有帧长归一化,且均取整数,将最短帧长取 6  $\mu\text{s}$ ,化为 3 000,其他均按此比例计算。单跳技术延时为最短帧执行时间,即 3 000,最坏延上上限为最短帧执行时间的 4 倍。所有消息共 2 656 个调度实例,全网的技术延时归一化后为 7 968 000,此时总延迟与平均延迟均为 0。

神经网络参数:深度为 4 层,每层 256 个节点,激活函数采用 elu 函数。优化器采用 adam 优化器,模仿学习时学习率采用 0.000 5,强化学习

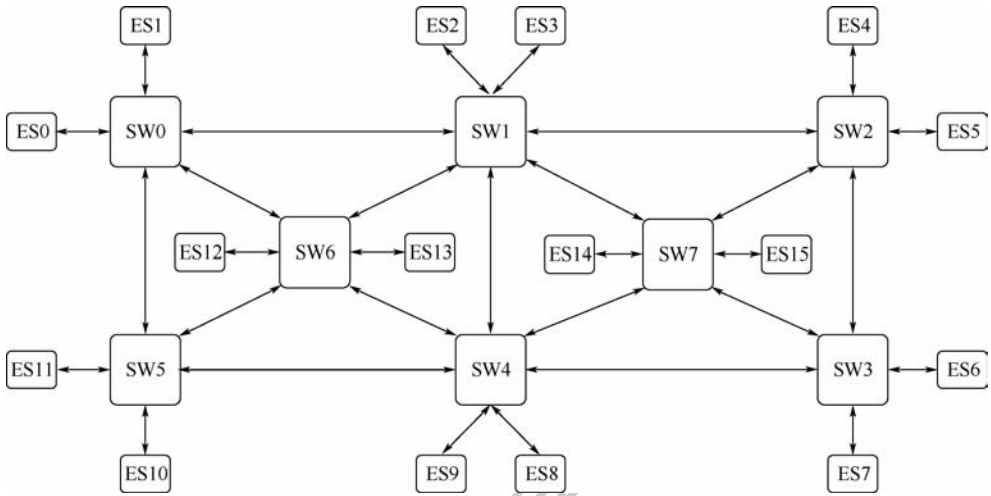


图2 A380 拓扑结构

Fig.2 Topology of A380

时学习率采用 0.000 1。n 步模拟中 n 值取 10。

2.2 实验结果

通过实验,可以得到如下结果,其中表 1 为各方法的计算时间与平均延迟,图 3 和图 4 分别为预训练的损失函数与无预训练的强化学习延迟优化情况。可以看出,由于强化学习和经过由模仿学习预训练的强化学习有训练的过程,收敛后效果优于直接使用树搜索所得的结果。

表 1 各方法调度结果

Table 1 Each method scheduling results

方法	单次计算时间/s	平均延迟(归一化)
SMT	265.00	1355.30
树搜索	2.17	4.10
强化学习	5.03	1.70
强化学习*	4.29	0.75

注: \* 表示使用模仿学习进行预训练。

3,5,7,11 倍数编号的消息,其帧长乘以不同系数,详细操作如表 1 所示。使消息的分布发生改变。端到端延迟及算法速度如表 2 ~ 表 5 所示。图 5 反映了算法收敛情况,其中实线为均值,着色区域为变化范围。

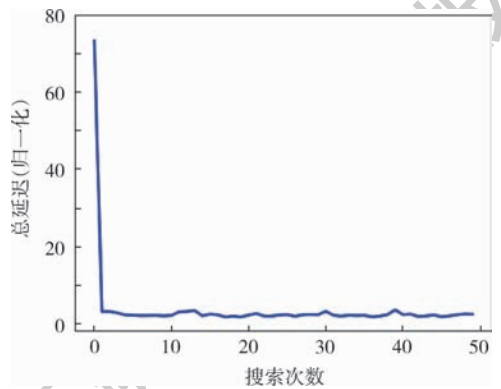


图 4 强化学习端到端延迟结果

Fig.4 Reinforcement learning end-to-end delay results

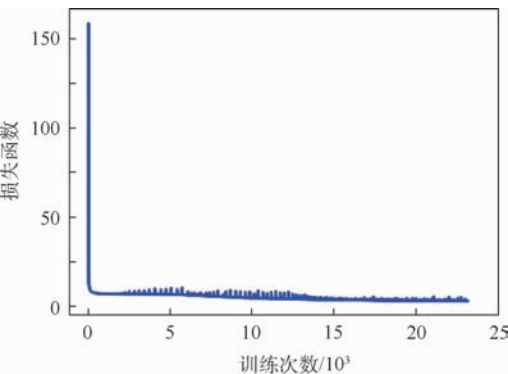


图 3 预训练损失函数变化

Fig.3 Loss function change in pre-training

表 2 4 种变化详细操作

Table 2 Detailed operation of four change methods

编号倍数	变化 1	变化 2	变化 3	变化 4
2	1.1	1.1	0.8	0.8
3	1.1	1.2	1.2	1.2
5	0.9	1.1	1.1	1.3
7	1.1	1.1	0.9	0.9
11	0.9	1.2	1.2	1.2

表 3 4 种变化后端到端延迟结果

Table 3 End-to-end delay results of four variations

方法	变化 1	变化 2	变化 3	变化 4
SMT	1360.2	1365.6	1284.2	1324.1
树搜索	4.4	5.5	4.5	4.4
强化学习	2.3	2.1	2.1	2.1
强化学习*	0.74	0.65	0.61	0.71

2.3 模型鲁棒性测试

2.3.1 帧长变化

将消息帧长按倍数变化,以此使消息分布与之前实验不同,将消息编号 1 ~ 1050。其中对 2,

表4 4种变化后收敛速度

Table 4 Convergence speed of four variations

方法	收敛速度/s			
	变化1	变化2	变化3	变化4
SMT				
树搜索				
强化学习	36	58	33	44
强化学习*	43	63	40	56

表5 4种变化后单次运行速度

Table 5 One-time running speed of four variations

方法	运行速度/s			
	变化1	变化2	变化3	变化4
SMT	281	287	267	274
树搜索	2.85	3.21	2.45	2.87
强化学习	5.12	6.45	5.58	6.25
强化学习*	4.32	5.26	4.97	5.07

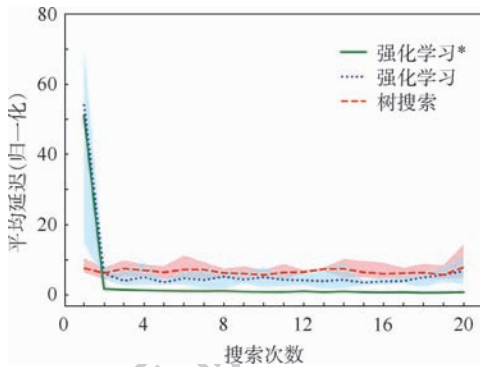


图5 各项变化的平均端到端延迟收敛情况

Fig.5 Average end-to-end delay until convergence for each change

2.3.2 增加消息数量

在之前的1050条消息中随机抽取一些消息并添加到最后,使需要调度的消息总数目增加。表6为详细操作,表7~表9为端到端延迟和算法速度,图6显示了算法收敛情况。

可以看到,神经网络在最初的搜索中均会由于消息分布的变化而不适应,得出端到端延迟极大

的调度表,但是很快可以调整过来,尤其是经过预训练之后的神经网络能够稳定地拥有很小的端到端延迟。经过了强化学习和用模仿学习预训练后的神经网络能够单次很快地搜索出表现优秀的调度表。在速度上和端到端延迟上均优于SMT求解器。两者相比,由于模仿学习与随机探索的组合添加了更多人为限制,在短时间内表现较之

表6 3种增加消息详细操作

Table 6 Detailed operation of three adding methods

操作	增加1	增加2	增加3
增加信息条数	6	60	120
总技术延时	8 019 000	8 478 000	8 988 000

表7 增加消息后端到端延迟结果

Table 7 End-to-end delay results after adding frames

方法	增加1	增加2	增加3
SMT	1 352.3	1 586.7	1 324.6
树搜索	2.5	2.9	4.5
强化学习	1.7	1.6	1.4
强化学习*	0.68	0.86	0.85

表8 增加消息后收敛速度

Table 8 Convergence speed after adding frames

方法	收敛速度/s		
	增加1	增加2	增加3
SMT			
树搜索			
强化学习	35	50	58
强化学习*	32	41	48

表9 增加消息后单次运行速度

Table 9 One-time running speed after adding frames

方法	运行速度/s		
	增加1	增加2	增加3
SMT	268	358	731
树搜索	2.24	2.34	2.52
强化学习	5.05	5.87	6.43
强化学习*	4.56	5.15	6.37

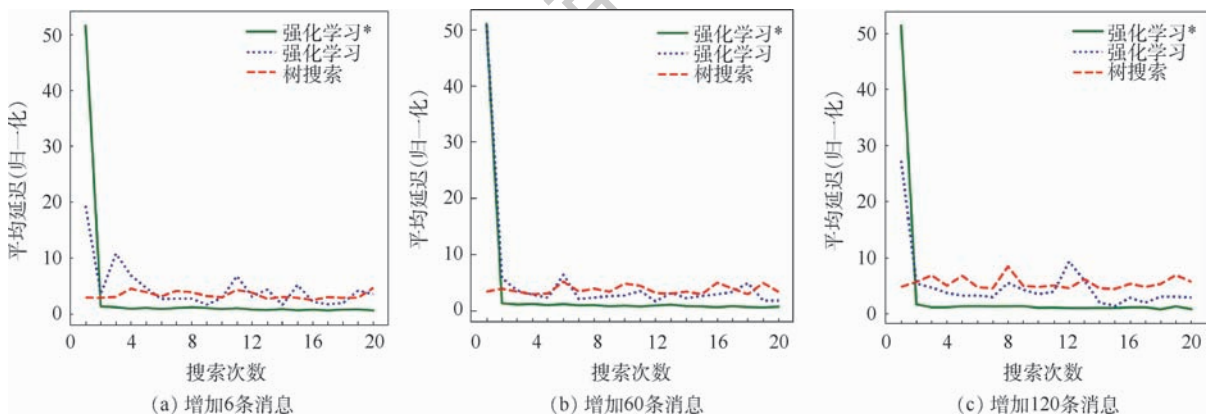


图6 增加6,60,120条消息的平均端到端延迟收敛情况

Fig.6 Average end-to-end delay until convergence of adding 6,60,120 frames

直接强化学习更加优秀。但是模仿学习需要预训练的过程,若当前消息与之前已训练的消息分布差异过大时,需要重新进行预训练;而直接强化学习方法则会退化为随机搜索,能够直接得出结果,并在此基础上进一步探索,整个训练过程时间大大快于模仿学习。此2种方法速度均快于SMT求解。

### 3 结 论

本文建立了以马尔可夫描述的调度任务模型,并基于此提出了采用强化学习的TT通信调度方法,以可尽量获得调度表和缩短延迟时间作为优化目标,选择并调整流量发送时间的偏移量,得到具有全局优化意义的时间触发调度表,通过与A380组网规模类似的案例进行了模型验证。

基于所验证案例,本文方法计算速度为SMT方法的数十倍,且可以及时观察到有哪些难以调度的流量,以方便快速进行调整。

此外,实验证明了本文方法的有效性,基于强化学习的调度算法传输延迟大幅低于基于Yices的SMT方法,可以更快得到更好的调度表。

#### 参考文献 (References)

- [1] 王国庆,谷青范,王森,等. 新一代综合化航空电子系统构架技术研究[J]. 航空学报,2014,35(6):1473-1486.  
WANG G Q, GU Q F, WANG M, et al. Research on architecture technology for new generation integrated avionics system[J]. Acta Aeronautica et Astronautica Sinica, 2014, 35(6): 1473-1486 (in Chinese).
- [2] 熊华钢,王中华. 先进航空电子综合技术[M]. 北京:国防工业出版社,2009:2-13.  
XIONG H G, WANG Z H. Advanced avionics integration techniques[M]. Beijing: National Defense Industry Press, 2009: 2-13 (in Chinese).
- [3] STEINER W. An evaluation of SMT-based schedule synthesis for time-triggered multi-hop networks[C]// Real-Time Systems Symposium. Piscataway, NJ: IEEE Press, 2011: 375-384.
- [4] 孔韵雯,李峭,熊华钢,等. 片间综合化互连时间触发通信调度方法[J]. 航空学报,2018,39(2):321590.  
KONG Y W, LI Q, XIONG H G, et al. Time-triggered communication scheduling method for off-chip integrated interconnection [J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(2):

321590 (in Chinese).

- [5] 李炳乾,王勇,谭小虎,等. 基于混合遗传算法的TTE静态调度表生成设计[J]. 电子技术应用,2016,42(10):96-99.  
LI B Q, WANG Y, TAN X H, et al. Hybrid-GA based static schedule generation for time-triggered Ethernet[J]. Application of Electronic Technique, 2016, 42(10): 96-99 (in Chinese).
- [6] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J]. Nature, 1986, 323:533-536.
- [7] KHALIL E B, DAI H, ZHANG Y, et al. Learning combinatorial optimization algorithms over graphs[C]// Neural Information Processing Systems, 2017:6348-6358.
- [8] SILVER D, SCHRITTWIESER J, SIMONYAN K, et al. Mastering the game of Go without human knowledge[J]. Nature, 2017, 550(7676):354-359.
- [9] GAL D, KRISHNAMURTHY D, MATEJ V, et al. Safe exploration in continuous action spaces[EB/OL]. (2018-01-26) [2019-01-01]. <https://arxiv.org/abs/1801.08757>.
- [10] KOCSIS L, SZEPESVARI C. Bandit based monte-carlo planning [C]// European Conference on Machine Learning, 2006:282-293.
- [11] CRACIUNAS S S, OLIVER R S. SMT-based task-and network-level static schedule generation for time-triggered networked systems[C]// International Conference on Real-Time Networks and Systems, 2014:45-54.
- [12] SCHULMAN J, LEVINE S, ABBEEL P, et al. Trust region policy optimization [EB/OL]. (2015-02-19) [2019-01-01]. <https://arxiv.org/abs/1502.05477>.
- [13] SILVER D, LEVER G, HEESS N, et al. Deterministic policy gradient algorithms[C]// International Conference on Machine Learning, 2014:387-395.
- [14] SCHULMAN J, WOLSKI F, DHARIWAL P, et al. Proximal policy optimization algorithms[EB/OL]. (2017-07-20) [2019-01-01]. <https://arxiv.org/abs/1707.06347>.
- [15] KORF R E. Depth-first iterative-deepening: An optimal admissible tree search[J]. Artificial Intelligence, 1985, 27(1):97-109.
- [16] ARINC. Aircraft data network, Part 7. Avionics full-duplex switched ethernet network; ARINC 664P7[R]. Washington, D. C.: ARINC, 2005.

#### 作者简介:

李浩若 男,硕士研究生。主要研究方向:实时通信系统。

何锋 男,博士,副教授。主要研究方向:航空电子网络、分布式实时系统。

## Time-triggered communication scheduling method based on reinforcement learning

LI Haoruo, HE Feng\*, ZHENG Zhong, LI Ershuai, XIONG Huagang

(School of Electronic and Information Engineering, Beihang University, Beijing 100083, China)

**Abstract:** In the future, time-triggered communication mechanism will be more widely selected for information transmission to ensure the certainty of information interaction in avionics system. How to reasonably implement time-triggered communication scheduling design is the key to time-triggered application to avionics interconnection systems. For the periodic task of time-triggered scheduling, we proposed a method for generating periodic scheduling timetable based on reinforcement learning. Firstly, the traffic scheduling task is transformed into a tree search problem, which has the Markov characteristics needed for reinforcement learning. Then, the reinforcement learning algorithm based on neural network is used to explore the schedule, and the waiting time is shortened to optimize the schedule. As the training is completed, the model can be directly used in tasks with similar message distribution. Compared with the method, e. g. Yices, which uses the satisfiability modulo theories (SMT) to solve the time-triggered schedule, the proposed method does not cause undetermined problem, and can guarantee the correctness and optimization of the time-triggered scheduling design results. For a large network with 1000 messages, the calculation speed of the proposed method is dozens of times faster than that of the SMT, and meanwhile, the end-to-end delay of the generated message by scheduling is less than 1% of that of the SMT, which greatly improves the timeliness of message transmission.

**Keywords:** time-triggered; scheduling method; reinforcement learning; tree search; offset time

**Received:** 2019-01-02; **Accepted:** 2019-03-15; **Published online:** 2019-05-08 09:41

**URL:** [kns.cnki.net/kcms/detail/11.2625.V.20190505.1020.001.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190505.1020.001.html)

**Foundation items:** National Natural Science Foundation of China (61301086,71701020); Open Fund of Tianjin Civil Aircraft Airworthiness and Maintenance Key Laboratory of Civil Aviation University of China (2017SW02)

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

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

DOI: 10.13700/j.bh.1001-5965.2019.0002

# 共用支承-转子结构系统振动耦合特性分析

章健<sup>1</sup>, 张大义<sup>1</sup>, 王永锋<sup>1</sup>, 马艳红<sup>1,2</sup>, 洪杰<sup>1,2,\*</sup>

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



**摘 要:** 针对带有涡轮级间共用承力框架的转子系统, 为准确描述转子-共用支承-转子(简称共用支承-转子结构系统)之间的振动特性, 采用转子截面横向和角向振动特性耦合动力学模型, 振动耦合产生机理及影响规律进行研究。理论分析结果表明: 转子支点的动态响应对其他转子的支点动刚度特性及转子振动响应特性具有一定影响, 共用支承结构振动响应对于转子系统振动特性的计算误差超过 10%, 因此, 在共用支承-转子结构系统的临界转速和振动响应计算分析中, 需要考虑 2 个转子与共用支承结构的振动耦合影响。对于涡轴发动机共用支承-转子结构系统的有限元仿真计算结果表明: 由于存在共用承力框架, 2 个转子之间将发生振动耦合, 系统产生耦合振型, 某一转子转速将会影响另一转子所激起的系统共振临界转速; 并对共用承力框架结构的隔振特性也有影响, 2 个转子共同激励下振动响应与转子单独激励相比, 在承力框架安装边上的动载荷以及载荷传递系数均大幅度提高。

**关 键 词:** 共用支承; 振动耦合; 临界转速; 动刚度; 载荷传递

**中图分类号:** V232.6

**文献标识码:** A

**文章编号:** 1001-5965(2019)09-1902-09

先进高推重比涡轴发动机多采用涡轮级间共用承力框架结构设计, 同时支承燃气发生器转子和动力涡轮转子后支点, 转子结构和承力结构之间的振动响应相互影响, 形成转子-共用支承-转子结构系统(简称共用支承-转子结构系统)振动耦合。近年来, 涡轮级间共用承力框架在先进航空发动机承力系统中广为采用, 是发动机减重、提高涡轮效率的重要途径之一<sup>[1-2]</sup>。而涡轮处于高温工作环境易带来支承刚度弱化问题, 会对发动机转子系统动力特性和承力结构振动特性产生影响<sup>[3-4]</sup>, 因此研究共用承力框架与转子之间耦合振动特性具有重要的工程意义。

目前, 国内外学者在航空发动机转子系统动力特性分析和转子-承力结构系统振动特性分析方面进行了诸多研究工作。常见的转子系统振动耦合问题为带中介轴承的双转子系统间的相互激

励耦合振动, 由于转子具有各自的动力特性, 并且在中介轴承处具有复杂横向激振力的交互作用, 因此使得双转子系统具有复杂的振动特性。文献[5-9]分别建立了滚动轴承-双转子系统动力学方程, 并采用仿真方法求解分析了滚动轴承对转子系统临界转速等动力特性的影响规律。而共用支承-转子结构系统不同于带有中介轴承的双转子系统, 在共用承力框架上具有多个轴承支点, 分别支承不同转速的转子, 因此在振动特性分析中需要进一步考虑承力框架的振动响应特性及其对各个转子振动特性的影响。

文献[10-14]提出了建立航空发动机整机动力学模型的具体方法, 考虑转/静子结构之间的振动耦合影响, 但仅对转子系统的动力特性进行了计算分析, 并未研究承力结构的振动响应。本文针对带有涡轮级间共用承力框架的航空发动

收稿日期: 2019-01-07; 录用日期: 2019-04-05; 网络出版时间: 2019-05-08 09:52

网络出版地址: [kns.cnki.net/kcms/detail/11.2625.V.20190506.0836.002.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190506.0836.002.html)

基金项目: 国家自然科学基金(51575022, 11672017)

\* 通信作者. E-mail: hongjie@buaa.edu.cn

**引用格式:** 章健, 张大义, 王永锋, 等. 共用支承-转子结构系统振动耦合特性分析[J]. 北京航空航天大学学报, 2019, 45(9): 1902-1910. ZHANG J, ZHANG D Y, WANG Y F, et al. Coupling vibration characteristics analysis of shared support-rotors system[J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1902-1910 (in Chinese).

机转子系统,提炼出“共用支承-转子”结构系统(RSR system),并建立了动力学方程,对转子系统动力学振动响应耦合特征及其主要力学参数的影响规律进行理论分析和仿真计算。

### 1 共用支承-转子结构系统振动特征

如图1所示,典型高推重比涡轴发动机使用涡轮级间共用承力框架支承动力涡轮转子和燃气发生器转子(简称“燃发转子”)的后支点,转子之间存在振动耦合影响。当动力涡轮转子以不同转速旋转时,共用支承结构会产生不同振动响应,即燃气发生器转子后支点位置产生基础振动,会导致燃发转子的临界转速等动力特性发生变化。同样,燃发转子振动对动力涡轮转子产生类似的影响,即2个转子之间存在振动耦合。

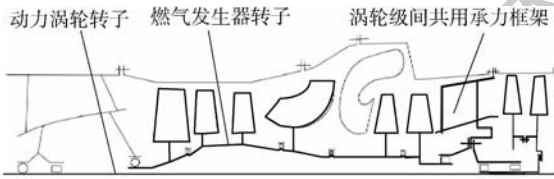


图1 高推重比涡轴发动机结构示意图

Fig.1 Schematic diagram of structure of turbo-shaft engine with high thrust-weight ratio

#### 1.1 振动耦合产生机理

如图2所示,建立对应的等效动力学模型,其中燃发转子和动力涡轮转子质量分别为  $m_1$ 、 $m_2$ ,燃发转子和动力涡轮转子前后支承刚度分别为  $k_{11}$ 、 $k_{12}$ 、 $k_{21}$ 、 $k_{22}$ ,并假设转子结构具有横向自由度  $y_1$ 、 $y_2$ ;共用承力框架质量为  $m_3$ ,转动惯量为  $J_3$ ,假设承力结构具有横向自由度  $y_3$  和角向自由度  $\theta_3$ ,横向支承刚度为  $k_3$ ,角向支承刚度为  $k_\theta$ ,支点到共用支承的距离分别为  $l_1$ 、 $l_2$ 。其中,图2(a)为转/静子联合的整体等效模型,图2(b)为转/静子分离的子结构等效模型。

耦合系统动力学方程为

$$M_a \ddot{q}_a + K_a q_a = 0 \tag{1}$$

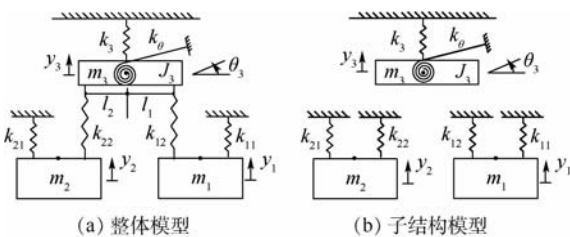


图2 考虑共用支承角向自由度的等效动力学模型

Fig.2 Equivalent dynamic model considering angular degree of shared support

式中:  $M_a$  和  $K_a$  分别为系统的质量矩阵和刚度矩阵;  $q_a$  为系统自由度向量。分别表示为

$$M_a = \begin{bmatrix} m_1 & 0 & 0 & 0 \\ 0 & m_2 & 0 & 0 \\ 0 & 0 & m_3 & 0 \\ 0 & 0 & 0 & J_3 \end{bmatrix}$$

$$K_a = \begin{bmatrix} k_{11} + k_{12} & 0 & -k_{12} & k_{12}l_1 \\ 0 & k_{21} + k_{22} & -k_{22} & -k_{22}l_2 \\ -k_{12} & -k_{22} & k_{12} + k_{22} + k_3 & -k_{12}l_1 + k_{22}l_2 \\ k_{12}l_1 & -k_{22}l_2 & -k_{12}l_1 + k_{22}l_2 & k_{12}l_1^2 + k_{22}l_2^2 + k_\theta \end{bmatrix}$$

$$q_a = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \theta_3 \end{bmatrix}$$

对于地面燃气轮机,静子支承刚度和结构质量远大于转子,假设  $k_{11} = k_{12} = k_{21} = k_{22} = 0.01k_3$ ,  $m_1 = m_2 = 0.01m_3$ ,  $k_\theta = 100k_{11}l_1^2$ ,  $J_3 = 100m_1l_1^2$ ,  $l_1 = l_2$ ,计算系统的固有频率  $f_1 \sim f_4$ 。

耦合模型:

$$\begin{cases} f_1 = f_2 = \sqrt{\frac{k_{11}}{m_1}} \\ f_3 = f_4 = 1.421 \sqrt{\frac{k_{11}}{m_1}} \end{cases} \tag{2}$$

独立模型:

$$\begin{cases} f_1 = f_2 = \sqrt{\frac{k_{11}}{m_1}} \\ f_3 = f_4 = 1.414 \sqrt{\frac{k_{11}}{m_1}} \end{cases} \tag{3}$$

由式(2)和式(3)可知,2种模型计算获得的结果相差很小,固有频率相对误差约为1%。

对于航空发动机,静子支承刚度和结构质量与转子相当,可假设  $k_{11} = k_{12} = k_{21} = k_{22} = k_3$ ,  $m_1 = m_2 = m_3$ ,  $k_\theta = k_{11}l_1^2$ ,  $J_3 = m_1l_1^2$ ,  $l_1 = l_2$ 。

耦合模型:

$$\begin{cases} f_1 = f_2 = \sqrt{\frac{k_{11}}{m_1}} \\ f_3 = f_4 = 2 \sqrt{\frac{k_{11}}{m_1}} \end{cases} \tag{4}$$

独立模型:

$$\begin{cases} f_1 = f_2 = \sqrt{\frac{k_{11}}{m_1}} \\ f_3 = f_4 = 1.414 \sqrt{\frac{k_{11}}{m_1}} \end{cases} \tag{5}$$

由式(4)和式(5)可知,对于2个转子的固有频率  $f_3$  和  $f_4$ ,2种模型计算结果相对误差超过40%。

通过上述分析可知,在共用支承-转子结构系统动力特性分析中忽略共用承力框架带来的振动耦合特性影响,将会使计算结果产生较大的误差,需要建立共用支承-转子结构系统动力学模型,才能获得准确的转子系统临界转速和振动响应。

### 1.2 转子结构系统动力学方程

对于带有共用承力框架的转子结构系统,如图3所示,建立其动力学模型的基本原理为:使用子结构传递矩阵法<sup>[15]</sup>分别建立两转子的状态参量方程,并建立共用承力框架在支点动载荷作用下的强迫振动动力学方程,再在支点位置使用平衡方程和协调方程组集转/静子动力学方程,获得共用支承-转子结构系统动力学方程,图中: $k_6$ 、 $k_7$ 、 $k_8$ 分别为3个支点4#、5#、6#的刚度; $M$ 为承力框架质量; $m$ 为轴承座质量;下标为 $M$ 和 $m$ 的变量分别为对应参数。

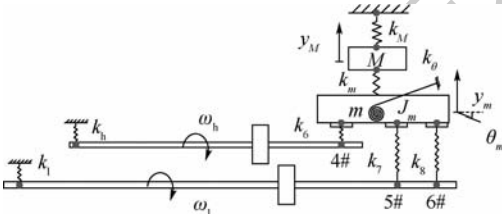


图3 共用支承-转子结构系统动力学模型  
Fig.3 Dynamic modeling of RSR system

对于任意两转子结构系统,转子的动力学方程均可写为

$$\begin{cases} M_h \ddot{q}_h - in_h J_h \dot{q}_h + K_h (q_h - q_{bh}) = q_h \Omega^2 \\ M_l \ddot{q}_l - in_l J_l \dot{q}_l + K_l (q_l - q_{bl}) = q_l \Omega^2 \end{cases} \quad (6)$$

式中: $M$ 、 $J$ 和 $K$ 分别为转子质量矩阵、陀螺力矩矩阵和刚度矩阵; $n$ 为转子转速; $q \Omega^2$ 为转子不平衡激励向量, $q$ 为转子状态向量, $\Omega$ 为转子不平衡量; $q_b$ 为转子自由度的基础振动;下标 $h$ 代表燃气发生器转子参数, $l$ 代表动力涡轮转子参数。

对于图3中的动力学模型,假设其各转子前支点均无基础振动,后支点4#、5#、6#上存在基础振动 $x_{4s}(t)$ 、 $y_{4s}(t)$ 、 $x_{5s}(t)$ 、 $y_{5s}(t)$ 、 $x_{6s}(t)$ 、 $y_{6s}(t)$ 。则有

$$\begin{cases} q_h = [x_h, y_h, \theta_{xh}, \theta_{yh}, x_{dh}, y_{dh}, \theta_{xdh}, \theta_{ydh}, \\ x_4, y_4, \theta_{x4}, \theta_{y4}, 0, 0, 0, 0]^T \\ q_l = [x_l, y_l, \theta_{xl}, \theta_{yl}, x_{dl}, y_{dl}, \theta_{xdl}, \theta_{ydl}, \\ x_5, y_5, \theta_{x5}, \theta_{y5}, x_6, y_6, \theta_{x6}, \theta_{y6}]^T \\ q_{bh} = [0, 0, 0, 0, 0, 0, 0, 0, x_{4s}, y_{4s}, \\ 0, 0, 0, 0, 0, 0]^T \\ q_{bl} = [0, 0, 0, 0, 0, 0, 0, 0, \\ x_{5s}, y_{5s}, 0, 0, x_{6s}, y_{6s}, 0, 0]^T \end{cases} \quad (7)$$

对于共用承力框架,其强迫振动响应动力学方程为

$$M_s \ddot{q}_s + K_s q_s = F_s \quad (8)$$

式中: $M_s$ 和 $K_s$ 分别为转子的质量矩阵和刚度矩阵; $q_s$ 为共用承力框架的自由度向量; $F_s$ 为共用承力框架各支点位置的动载荷向量。有

$$\begin{cases} q_s = [x_M, y_M, x_m, y_m, \theta_{xm}, \theta_{ym}]^T \\ F_s = [F_{4x}, F_{4y}, F_{5x}, F_{5y}, F_{6x}, F_{6y}]^T \end{cases} \quad (9)$$

根据4#、5#和6#支点上的平衡方程以及共用承力框架中的位移协调方程:

$$F_s = K_b (q_{bs} - q_{br}) \quad (10)$$

$$q_{bs} = A_{qs} q_s \quad (11)$$

式中: $K_b$ 为支点动刚度矩阵; $q_{bs}$ 和 $q_{br}$ 分别为承力框架和转子结构在4#、5#、6#支点位置的自由度向量; $A_{qs}$ 为承力框架向转子支点位置的自由度变化矩阵。可组集系统振动耦合动力学方程:

$$K_a q_a - C_a \dot{q}_a + M_a \ddot{q}_a = \begin{bmatrix} q_h \Omega^2 + K_h q_{bh} \\ q_l \Omega^2 + K_l q_{bl} \\ K_b (A_{qs} q_s - q_{br}) \end{bmatrix} \quad (12)$$

式中:

$$M_a = \begin{bmatrix} M_h & & \\ & M_l & \\ & & M_s \end{bmatrix}$$

$$K_a = \begin{bmatrix} K_h & & \\ & K_l & \\ & & K_s \end{bmatrix}$$

$$C_a = \begin{bmatrix} n_h J_h & & \\ & n_l J_l & \\ & & 0 \end{bmatrix}$$

$$q_s = \begin{bmatrix} q_h \\ q_l \end{bmatrix}$$

将转子支点基础自由度向量向承力框架自由度向量转化:

$$\begin{cases} q_{bh} = A_{bhs} q_s \\ q_{bl} = A_{bls} q_s \end{cases} \quad (13)$$

式中: $A_{bhs}$ 、 $A_{bls}$ 均为变换矩阵。可获得共用支承-转子结构系统振动耦合统动力学方程:

$$\begin{bmatrix} q_h \Omega^2 \\ q_l \Omega^2 \\ 0 \end{bmatrix} = \begin{bmatrix} K_h & & -K_h A_{bhs} \\ & K_l & -K_l A_{bls} \\ -K_h A_{bhs} & -K_l A_{bls} & K_s - K_b A_{qs} \end{bmatrix} q_a - \begin{bmatrix} n_h J_h & & \\ & n_l J_l & \\ & & 0 \end{bmatrix} \dot{q}_a + \begin{bmatrix} M_h & & \\ & M_l & \\ & & M_s \end{bmatrix} \ddot{q}_a \quad (14)$$

在结构系统中,转/静子结构之间存在刚度耦合项  $K_h A_{bhs}$  和  $K_l A_{bls}$ ,进而使系统发生振动耦合。

## 2 共用支承-转子结构系统动力特性耦合影响

在使用涡轮级间共用承力框架的发动机中,转/静子之间存在振动耦合,会影响转子的动力特性,本节将以高推重比涡轴发动机为例,基于有限元仿真方法计算分析转/静子振动耦合对转子系统临界转速和支点动载荷的影响规律。

### 2.1 有限元模型与约束

针对图1中的涡轴发动机,采用 Solid186 单元建立共用支承-转子结构系统有限元模型,使用弹簧单元模拟轴承支承,如图4所示。为避免复杂模型带来的庞大计算量,对结构进行合理简化。忽略连接界面的影响,转/静子连接接触面均采用固结;忽略对整体力学性能影响较小的圆角、小孔等细节因素,提高计算效率。

考虑转子系统原有的支承方式,燃发转子和动力涡轮转子前支点(1#和3#)位置弹簧单元的外端进行固定约束,2#支点位置弹簧单元外端施加径向和周向约束。4#、5#、6#支点由涡轮级间共用承力框架支承,使用弹簧单元连接转/静子上的对应节点,并对节点进行周向位移耦合。

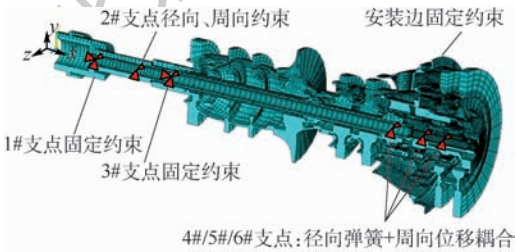


图4 共用支承-转子结构系统有限元模型  
Fig.4 Finite element model of RSR system

### 2.2 结构系统模态

影响模态、临界转速等固有特性的关键因素主要有2点:转子自身的刚度/质量分布,转子支承的位置和刚度<sup>[16]</sup>。在带有共用支承结构的转子结构系统中,2个转子通过共用承力框架可发生振动响应的相互影响。通过模态分析计算结构系统的固有频率与相应振型,如图5~图7所示。可见,高/低压转子与共用承力框架之间存在振动耦合影响,采用耦合模型计算获得的转子模态频率与单转子模型计算结果有所差异。

当2个转子的模态频率相近时,将在共用支承-转子结构系统中形成新的耦合振型,其模态频率在对应的两转子单独模态频率之间。此外,结

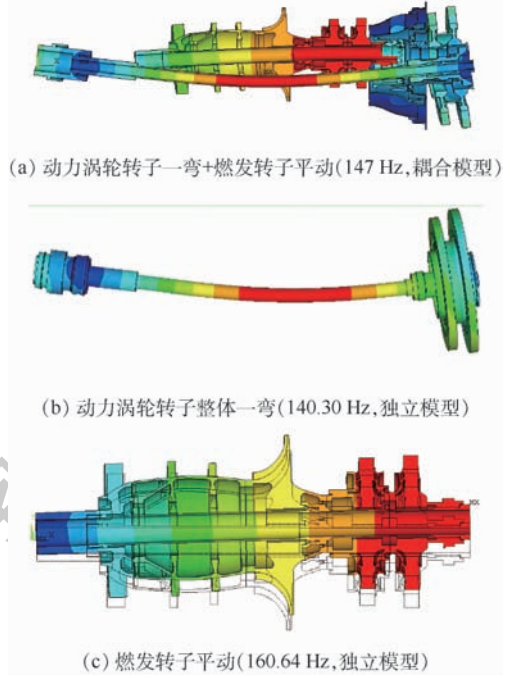


图5 共用支承-转子结构系统第1阶耦合模态及其对应的单转子振型

Fig.5 First-order coupled mode of RSR system and corresponding modes in single-rotor model

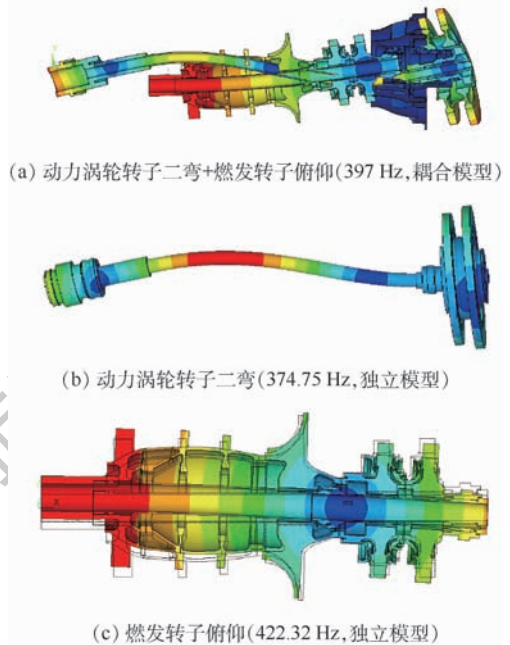


图6 共用支承-转子结构系统第2阶耦合模态及其对应的单转子振型

Fig.6 Second-order coupled mode of RSR system and corresponding modes in single-rotor model

构系统中还保留着某些单转子局部模态,但其模态频率相对于单转子模型中计算获得模态频率发生变化。

### 2.3 结构系统临界转速

由于存在共用承力框架,结构系统的各阶模

态可以分别被2个转子单独激起,因此需要考虑某一转子激起的临界转速受另一转子转速变化的影响。以燃气发生器转子为例,确定其临界转速计算的基本步骤为:①给定动力涡轮转子转速  $n_1$  和燃发转子转速  $n_2$ ;②固定动力涡轮转速,改变燃发转子转速  $n_2$ ,计算不同转速下的系统固有模态并绘制 Campbell 图,获得在动力涡轮转子转速为  $n_1$  时燃发转子激起系统共振的临界转速;③改变  $n_1$ ,重复步骤②,计算不同动力涡轮转子转速下的燃发转子激起系统共振的临界转速。同理可计算动力涡轮转子激起系统共振的临界转速。结果如图8所示。

转子系统临界转速取决于转子自身刚度和支点的支承刚度,不同转速下陀螺力矩发生变化

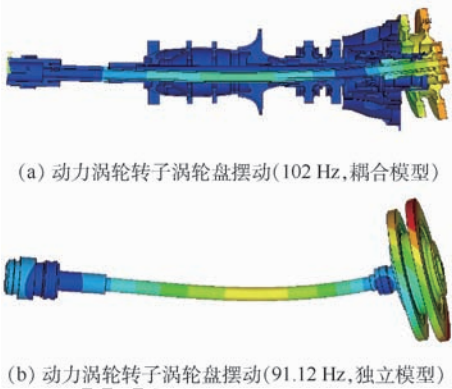


图7 共用支承-转子结构系统非耦合模态及其对应的单转子振型

Fig. 7 Non-coupled mode of RSR system and corresponding mode in single-rotor model

进而改变转子自身刚度是转速影响临界转速的主要原因。对于动力涡轮转子激起动力涡轮盘摆动和系统第1阶耦合振动,燃发转子为平动振型,转子陀螺力矩基本不随转速改变,因此对应的临界转速基本不变;而对于动力涡轮转子激起的第2阶耦合振动,燃发转子为俯仰振型,陀螺力矩受转速影响较大,临界转速变化明显。对于燃发转子激起的各阶共振振型,动力涡轮转子均发生弯曲变形,因此,当动力涡轮转子转速增加时,转子上陀螺力矩增大,转子自身刚度提高,进而导致各阶临界转速均上升。在共用支承-转子结构系统中,当某转子工作转速内存在俯仰或弯曲临界转速时,动力学设计中需要考虑另一转子转速变化对该转子临界转速的影响。

为定量分析振动耦合对共用支承转子系统动力特性的影响,本文引入临界转速敏感系数描述燃发转子激起临界转速随动力涡轮转子转速的变化规律,可表示为

$$\varepsilon_h = (n_{c2} - n_{c1}) / (n_2 - n_1) \quad (15)$$

式中:  $n_1$ 、 $n_2$  分别为计算范围内动力涡轮转子的最小和最大转速;  $n_{c1}$ 、 $n_{c2}$  为相应燃发转子激起的临界转速。

转/静子振动耦合特性受结构系统刚度分布特征影响,因此,临界转速敏感系数应是支点支承刚度的函数。支点支承刚度应是弹性支承结构和承力框架各部件刚度的串联结果,并主要取决于弹性支承的刚度<sup>[17]</sup>,因此,可以在保证承力框架结构特征不变的条件下,仅改变弹性支承刚度(有限元模型中点的弹簧刚度),计算分析燃发转子临界转速敏感系数随4#支点弹性支承刚度的变化规律,结果如图9所示。随着支承刚度的增加,燃发转子临界转速对动力涡轮转速的敏感度上升。其原因在于,在共用承力框架处振动耦合主要是位移耦合,在共用承力框架振动位移响应相同情况下,支点支承刚度较大时,所产生的动刚度波动幅值更大,对转子-支承系统动力特性影响

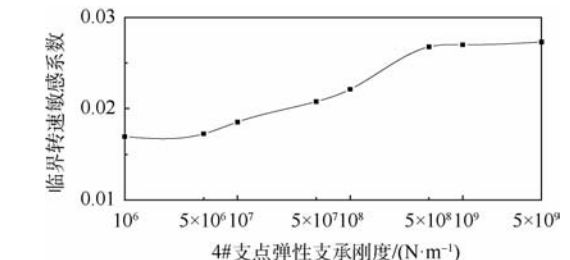


图9 燃发转子激起临界转速敏感系数变化曲线  
Fig. 9 Variation curve of sensitivity coefficient of critical speed in gas generator rotor

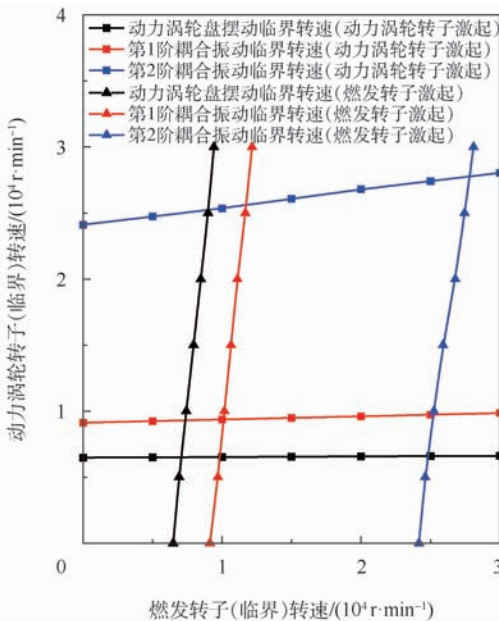


图8 共用支承-转子结构系统临界转速  
Fig. 8 Critical speed of RSR system

更为显著。

共用支承-转子系统动力特性的耦合影响,还表现为转子-共用支承-转子中任一结构力学特征的改变,会对其他结构动力特性产生影响。同样以4#支点弹性支承刚度变化为例,分析其对结构系统动力特性影响。图10为燃发转子激起的系统临界转速随4#支点弹性支承刚度的变化曲线。随着支点弹性支承刚度的增加,动力涡轮转子激起的两阶耦合振动临界转速也有所增加。

在对使用涡轮级间共用承力框架的转子系统进行动力特性分析时,需要考虑转/静子耦合振动的影响。特别是通过调整支承刚度优化转子临界转速时,当转子转速变化时,临界转速以及临界转速对其他转子转速变化的敏感性均会变化。应选取合适的支承刚度,既满足临界转速高于工作转速20%的裕度要求<sup>[18]</sup>,同时对其他转子转速变化敏感度较低。

本文计算载荷状态取燃发转子转速50 000 r/min,动力涡轮转子转速30 000 r/min,并在2个转子上分别施加不平衡量,通过有限元仿真计算获得转子系统支点动载荷响应,计算结果如图11所示。

在单独的燃发转子-支承结构系统中,动载荷响应频率组成主要为转子自身转速频率(833.33 Hz),计算模型中支点最大载荷值约为30 N。而在共用支承-转子结构系统中,响应频率中除了燃发转子转速频率外,还具有动力涡轮转子转速频率(500 Hz),而支点动载荷最大值可达44.697 N,约为单一转子模型中的1.5倍。分析认为,在共用支承-转子结构系统中,动力涡轮转子不平衡量在5#、6#支点处产生较大的动载荷,并引起共用支承结构的整体振动响应,即在燃发转子后支点处施加基础振动,使得该支点处动载荷响应频率变化,响应幅值增加。

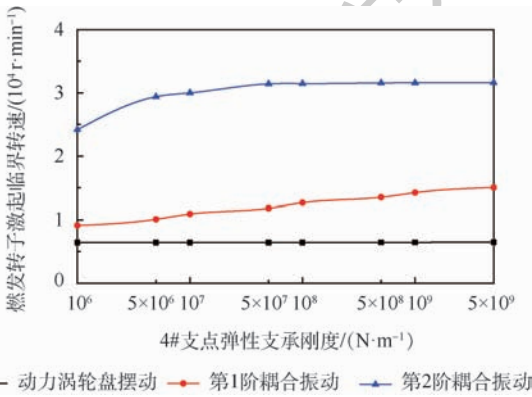


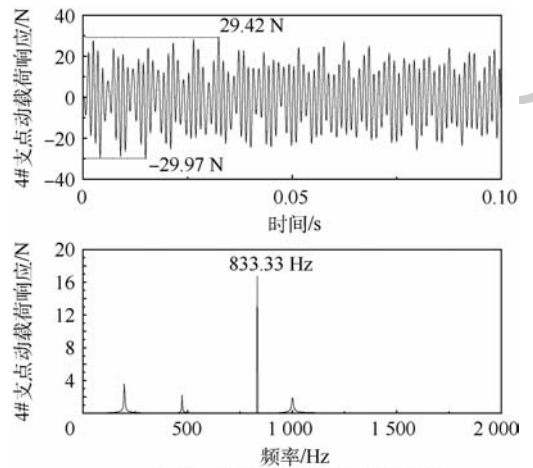
图10 支承刚度对临界转速的影响

Fig. 10 Influence of support stiffness to critical speed

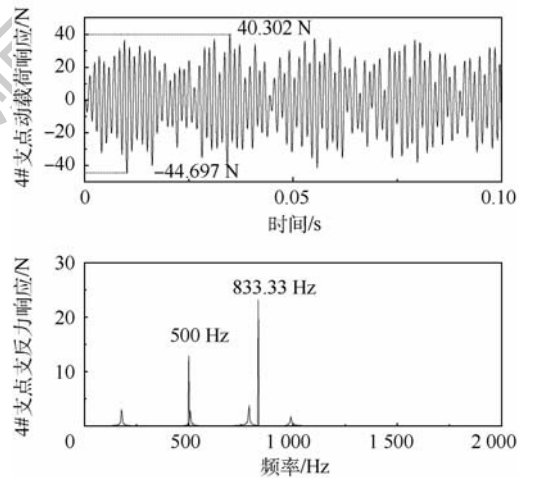
## 2.4 支点动载荷

支点动载荷是共用支承-转子结构系统描述转/静子结构间振动耦合的关键参数。在非共振状态,承力框架振动响应主要是支点载荷激励引起,支点载荷增大导致共用支承结构振动水平上升是引起振动耦合的主要原因。支点动载荷频率也是共用支承转子系统重要振动响应特征参数,动载荷大小、频率组成及其占比一定程度可反映不同结构间振动耦合程度。

支点动载荷主要来源于旋转转子不平衡载荷及其陀螺力矩效应。由于制造、装配误差或运行过程中的热变形等因素,转子不可避免地存在不平衡量,转子旋转过程中,不平衡激振力将在支点位置产生动载荷<sup>[19]</sup>。此外,转子发生弯曲或其上存在陀螺力矩时,支点动载荷将发生变化以平衡转子上的弯矩。



(a) 单一的燃发转子-支承结构系统



(b) 共用支承-转子结构系统

图11 不同系统力学模型下支点动载荷响应幅值  
Fig. 11 Response amplitude of bearing dynamic load in different mechanical models

### 3 共用支承-转子结构振动响应耦合特性

在共用支承-转子结构系统中,振动耦合将同时改变转子动力特性和承力结构振动响应特性。共用承力框架上各支点动刚度、支点之间的机械阻抗、支点到框架安装边的载荷传递系数等振动特性均会发生变化。

#### 3.1 支点动刚度与机械阻抗

引入支点动刚度  $K_i$  和支点间机械阻抗  $X_{ij}$ :

$$K_i(\omega) = |F_i(\omega)| / |Y_i(\omega)| \quad (16)$$

$$X_{ij}(\omega) = |F_i(\omega)| / |Y_j(\omega)| \quad (17)$$

式中:  $F_i(\omega)$  为第  $i$  号支点位置的动载荷,  $\omega$  为转子角速度;  $Y_i(\omega)$ 、 $Y_j(\omega)$  分别为  $i$  号、 $j$  号支点位置的位移响应。支点动刚度反映了支承结构刚度随转速变化的敏感性<sup>[20]</sup>, 支点间机械阻抗反映了共用承力框架各个支点之间的隔振特性。

图 12 和图 13 分别为采用共用支承-转子结构系统振动耦合模型和单一转子模型时, 4# 支点动刚度和支点间机械阻抗随激振频率的变化曲线。在共用支承-转子结构系统中, 由于各结构间刚度耦合项的存在, 提高了支承结构的固有频率以及支点载荷动刚度, 增加了共用承力框架的

抗变形能力。此外, 结构系统中由于考虑了不同子结构固有频率的影响, 因此当激振力频率等于承力结构固有频率(达到临界转速, 图 12 点 B、C)或转子结构固有频率(图 12 点 A)时, 共用承力框架均会发生支承刚度突降。

#### 3.2 载荷传递系数

引入载荷传递系数:

$$X_b = |F_b| / \sum |F_i(\omega_i)| \quad (18)$$

式中:  $F_i(\omega_i)$  为共用承力框架在其所支承的第  $i$  号支点位置的动载荷;  $F_b$  为承力框架安装边位置的动载荷响应。载荷传递系数反映了转子支点动载荷在共用承力框架中传递特性, 若该系数大于 1, 即表明支点载荷在承力框架中传递时被放大。在动力学设计中, 应尽可能降低承力框架的载荷传递系数, 增强结构的隔振性<sup>[21]</sup>。在共用支承-转子结构系统中, 由于转子和支承结构的刚度/质量分布特征的影响, 使得共用承力框架的载荷传递系数将随转子转速发生变化。

当燃发转子转速为 50 000 r/min, 动力涡轮转子转速为 30 000 r/min 时, 计算获得共用承力框架在各支点位置和安装边的动载荷响应。再将支点动载荷响应作为激励载荷施加在共用承力框架结构计算模型的相应支点, 计算获得安装边动载荷并对比不同模型中框架载荷传递系数, 结果如表 1 所示。可以发现, 考虑转子与支承结构振动耦合特性后, 承力框架安装边动载荷和载荷传递系数均会上升。

表 1 动载荷响应与载荷传递系数

Table 1 Dynamics load response and load-transferring coefficient

模型	支点动载荷幅值/N				载荷传递系数
	4#	5#	6#	安装边	
共用支承-转子模型	44.69	5.98	11.73	157.11	2.52
单一转子模型	44.69	5.98	11.73	107.95	1.73

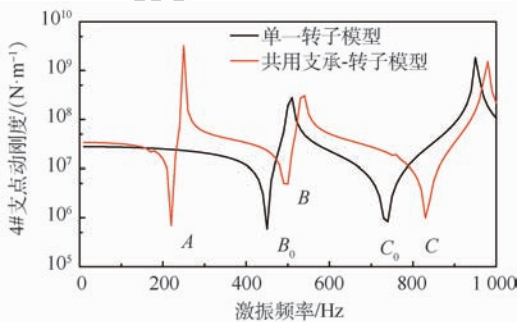


图 12 不同模型中的支点动刚度曲线

Fig. 12 Curves of dynamic stiffness in different models

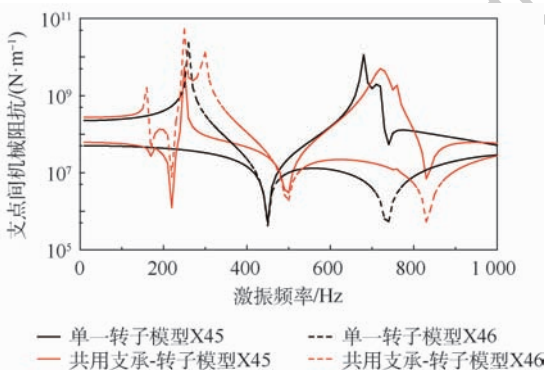


图 13 不同模型中的机械阻抗曲线

Fig. 13 Curves of mechanical impedance in different models

## 4 结论

本文运用转子动力学理论及仿真计算方法对使用涡轮级间共用承力框架的高推重比涡轴发动机进行转子系统动力特性与承力结构振动特性计算, 对共用支承-转子结构系统振动耦合特性进行分析, 主要结论如下:

1) 在共用支承转子系统中, 2 个转子之间存在动力特性的相互影响, 主要表现在转子模态热性、弯曲临界转速、临界转速对支点的刚度的敏感度以及支点动载荷等转子动力特性参数均会随其

他转子转速变化的改变。共用支承-转子结构模型可以考虑转子与支承结构之间的振动耦合影响,因此使用该模型对共用支承-转子结构系统动力特性进行计算分析更为准确。

2) 由于转子与支承结构之间存在振动耦合特征,各支点位置的动载荷中同时包含2个转子的转速频率,在共用承力结构振动特性设计中需要考虑框架在多点、宽频复杂激励载荷作用下的振动响应特征。依据本文的计算分析结果,支点激振载荷在共用承力框架传递的过程中将发生放大,并且当考虑支承结构之间振动耦合之后,放大效应将进一步增强。因此为保证承力结构安全裕度准确可靠,在其振动特性设计中应采用共用支承-转子结构系统振动耦合模型。

目前,在共用支承-转子结构系统中只分析了转子与共用支承结构振动耦合对转子系统动力特性和共用承力框架振动特性的影响,在工程应用中还需考虑轴承支承阻尼以及在热环境下承力框架对结构系统振动耦合特性的影响。

### 参考文献 (References)

- [1] SEDA J F. Aircraft engine with inter-turbine engine frame: US6708482[P]. 2005-04-26.
- [2] KUMAR K B, SOMANATH N, SOWA W A. Mid-turbine frame: US8181466[P]. 2012-05-22.
- [3] 马艳红,曹冲,李鑫,等. 涡轴发动机涡轮级间支承结构设计关键技术[J]. 航空发动机, 2014, 40(4): 34-40.  
MA Y H, CAO C, LI X, et al. Key design technology of mid turbine frame for turbo-shaft engine [J]. Aeroengine, 2014, 40(4): 34-40 (in Chinese).
- [4] 洪杰,马艳红,张大义. 航空燃气轮机总体结构与动力学分析[M]. 北京:北京航空航天大学出版社, 2014.  
HONG J, MA Y H, ZHANG D Y. Structure design and dynamic analysis of aviation gas turbine [M]. Beijing: Beihang University Press, 2014 (in Chinese).
- [5] GUPTA K, GUPTA K D, ATHRE K. Unbalance response of a dual rotor system: Theory and experiment [J]. Journal of Vibration and Acoustics, 1993, 115(4): 427-435.
- [6] VILLA C, SINOU J J, THOUVEREZ F. Stability and vibration analysis of a complex flexible rotor bearing system [J]. Communications in Nonlinear Science & Numerical Simulation, 2008, 13(4): 804-821.
- [7] CHIANG H W D, HSU C N, TU S H. Rotor-bearing analysis for turbomachinery single-and dual-rotor systems [J]. Journal of Propulsion & Power, 2004, 20(6): 1096-1104.
- [8] 邓四二,付金辉,王燕霜,等. 航空发动机滚动轴承-双转子系统动态特性分析[J]. 航空动力学报, 2013, 28(1): 195-204.  
DENG S E, FU J H, WANG Y S, et al. Analysis on dynamic characteristics of aero-engine rolling bearing/dual-rotor system [J]. Journal of Aerospace Power, 2013, 28(1): 195-204 (in Chinese).
- [9] 张大义,刘焯辉,梁智超,等. 航空发动机双转子系统临界转速求解方法[J]. 推进技术, 2015, 36(2): 292-298.  
ZHANG D Y, LIU Y H, LIANG Z C, et al. Prediction for critical speed of double spools system in aero engines [J]. Journal of Propulsion Technology, 2015, 36(2): 292-298 (in Chinese).
- [10] 高金海,洪杰. 航空发动机整机动力特性建模技术研究[J]. 战术导弹技术, 2006(3): 29-35.  
GAO J H, HONG J. Study of modeling technique for dynamic characteristics of engine casing [J]. Tactical Missile Technology, 2006(3): 29-35 (in Chinese).
- [11] 陈萌,马艳红,刘书国,等. 航空发动机整机有限元模型转子动力学分析[J]. 北京航空航天大学学报, 2007, 33(9): 1013-1016.  
CHEN M, MA Y H, LIU S G, et al. Rotor-dynamic analysis of whole aero-engine models based on finite element method [J]. Journal of Beijing University of Aeronautics and Astronautics, 2007, 33(9): 1013-1016 (in Chinese).
- [12] 陈果. 双转子航空发动机整机振动建模与分析[J]. 振动工程学报, 2011, 24(6): 619-632.  
CHEN G. Vibration modeling and analysis for dual-rotor aero-engine [J]. Journal of Vibration Engineering, 2011, 24(6): 619-632 (in Chinese).
- [13] 陈果. 航空发动机整机振动耦合动力学模型及其验证[J]. 航空动力学报, 2012, 27(2): 241-254.  
CHEN G. A coupling dynamic model for whole aero-engine vibration and its verification [J]. Journal of Aerospace Power, 2012, 27(2): 241-254 (in Chinese).
- [14] 张大义,刘焯辉,洪杰,等. 航空发动机整机动力学模型建立与振动特性分析[J]. 推进技术, 2015, 36(5): 768-773.  
ZHANG D Y, LIU Y H, HONG J, et al. Investigation on dynamical modeling and vibration characteristics for aero engine [J]. Journal of Propulsion Technology, 2015, 36(5): 768-773 (in Chinese).
- [15] 晏砺堂. 结构系统动力特性分析[M]. 北京:北京航空航天大学出版社, 1989.  
YAN L T. Analysis of dynamics characteristics of structural systems [M]. Beijing: Beihang University Press, 1989 (in Chinese).
- [16] 顾家柳. 转子动力学[M]. 北京:国防工业出版社, 1985.  
GU J L. Rotor dynamics [M]. Beijing: National Defense Industry Press, 1985 (in Chinese).
- [17] 洪杰,王华,肖大为,等. 转子支承刚度对转子动力特性的影响分析[J]. 航空发动机, 2008, 34(1): 23-27.  
HONG J, WANG H, XIAO D W, et al. Effects of dynamic stiffness of rotor bearing on rotordynamic characteristics [J]. Aero-engine, 2008, 34(1): 23-27 (in Chinese).
- [18] 朱梓根. 航空涡轴、涡桨发动机转子系统设计准则[M]. 北京:航空工业出版社, 2000.  
ZHU Z G. Structure design criteria for rotor system of turbo shaft and turboprop aero-engine [M]. Beijing: Aviation Industry Press, 2000 (in Chinese).
- [19] 付才高. 航空发动机设计手册第19册, 转子动力学及整机振动[M]. 北京:航空工业出版社, 2000.  
FU C G. Aero-engine design manual. Volume 19: Rotor dynamic

- and vibration [M]. Beijing: Aviation Industry Press, 2000 (in Chinese).
- [20] 马艳红, 曹冲, 李鑫, 等. 航空发动机承力系统结构效率评估方法[J]. 航空动力学报, 2016, 31(2): 274-281.  
MA Y H, CAO C, LI X, et al. Assessment method of structural efficiency on bearing system in aero-engine[J]. Journal of Aerospace Power, 2016, 31(2): 274-281 (in Chinese).
- [21] 黄太平. 多转子系统振动的子系统分析方法—阻抗耦合法与分振型综合法[J]. 振动工程学报, 1988, 1(3): 30-40.  
HUANG T P. Subsystem analysis for multi-rotor system-imped-

ance coupling method and component mode synthesis[J]. Journal of Vibration Engineering, 1988, 1(3): 30-40 (in Chinese).

#### 作者简介:

章健 男, 硕士研究生。主要研究方向: 航空发动机结构、强度与可靠性。

洪杰 男, 博士, 教授, 博士生导师。主要研究方向: 航空发动机结构、强度与可靠性。

## Coupling vibration characteristics analysis of shared support-rotors system

ZHANG Jian<sup>1</sup>, ZHANG Dayi<sup>1</sup>, WANG Yongfeng<sup>1</sup>, MA Yanhong<sup>1,2</sup>, HONG Jie<sup>1,2,\*</sup>

(1. School of Energy and Power Engineering, Beihang University, Beijing 100083, China;

2. Collaborative Innovation Center of Advanced Aero-Engine, Beijing 100083, China)

**Abstract:** The rotor system with turbine mid frame is researched as an object, and the dynamic model with coupling vibration in transverse and angular degree was employed to reveal the vibration coupling mechanism and influence in shared support-rotors system. It is pointed that support response of one rotor can influence the bearing dynamic stiffness and rotor dynamic characteristic of another, and the calculation error of traditional rotor dynamic modeling exceed 10%, so the vibration coupling should be considered when analyzing critical speed and dynamic response in the rotor system with mid frame. The finite element calculation results for the shared support-rotors system in turbo-shaft engine exhibit that the coupling modes will arise, and the speed of one rotor can affect the critical resonant speed excited by another. The vibration isolation characteristics of mid turbine can also be affected due to the coupling vibration, and compared with non-coupling system, the dynamic response and load-transferring coefficient both remarkably increase.

**Keywords:** shared support; coupling vibration; critical speed; dynamic stiffness; load transferring

Received: 2019-01-07; Accepted: 2019-04-05; Published online: 2019-05-08 09:52

URL: [kns.cnki.net/kcms/detail/11.2625.V.20190506.0366.002.html](http://kns.cnki.net/kcms/detail/11.2625.V.20190506.0366.002.html)

Foundation items: National Natural Science Foundation of China (51575022, 11672017)

\* Corresponding author. E-mail: hongjie@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,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,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,引言,正文,参考文献。首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:100191 北京市海淀区学院路 37 号 北京航空航天大学学报编辑部

办公地点:北京航空航天大学办公楼 405,407,409 房间

电 话:(010)82315594,82338922,82314839,82315426

E-mail: jbuaa@buaa.edu.cn

http://bhxb.buaa.edu.cn

http://www.buaa.edu.cn

# 《北京航空航天大学学报》 第五届编辑委员会

主任(主编): 赵沁平

(以下按姓氏笔划为序)

副主任(副主编): 丁希仑 王少萍 孙志梅 李秋实 李焕喜 杨嘉陵  
苗俊刚 相艳 徐立军 钱德沛 曹晋滨

编委: 马殿富 王琪 王聪 邓小燕 王青云 王荣明 刘宇  
刘红 江洁 刘强 闫鹏 朱天乐 刘铁钢 齐铂金  
陈万春 邹正平 苏东林 杨世春 沈成平 邱志平 宋知人  
杨树斌 张晓林 杨晓奕 杨继萍 李惠峰 吴新开 张瑞丰  
杨照华 宋凝芳 周锐 林宇震 林贵平 战强 姚仰平  
胡庆雷 赵秋红 段海滨 赵巍胜 席平 郭宏 徐洁  
徐世杰 郭洪波 康锐 翟锦 熊华钢

## 北京航空航天大学学报

Beijing Hangkong Hangtian Daxue Xuebao

(原《北京航空学院学报》)

(月刊 1956年创刊)

第45卷第9期 2019年9月

## JOURNAL OF BEIJING UNIVERSITY OF AERONAUTICS AND ASTRONAUTICS (JBUA)

(Monthly, Started in 1956)

Vol.45 No.9 September 2019

主管单位 中华人民共和国工业和信息化部  
主办单位 北京航空航天大学  
主编 赵沁平  
编辑出版 《北京航空航天大学学报》  
编辑部  
邮编 100083  
地址 北京市海淀区学院路37号  
印刷 北京科信印刷有限公司  
发行 北航文化传媒集团  
发行范围 国内外发行  
联系电话 (010) 82315594 82338922  
82314839  
电子信箱 jbuua@buaa.edu.cn

Administrated by Ministry of Industry and Information  
Technology of the People's Republic of China  
Sponsored by Beijing University of Aeronautics  
and Astronautics (BUAA)  
(Beijing 100083, P. R. China)  
Chief Editor ZHAO 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



9 771001 596199