

doi: 10.7690/bgzdh.2026.05.022

基于多模态地形感知的四足机器人步态自适应规划方法

王宇倩¹, 吴永存¹, 刘昆仑², 向学辅¹, 张璐³

(1. 中国兵器装备集团自动化研究所有限公司系统总体部, 四川 绵阳 621000;

2. 陆军装备部驻重庆地区军事代表局驻广元地区军事代表室, 四川 广元 628000;

3. 中国兵器装备集团自动化研究所有限公司无人机事业部, 四川 绵阳 621000)

摘要: 针对四足机器人的地形感知受模态不确定性以及步态切换缺乏连续性问题, 提出一种统一的多模态地形分割与自适应步态规划方法。通过多模态地形分割框架增强对不同模态组合的适应能力, 并基于地形识别结果引入弗劳德数约束, 实现物理上连续的步态过渡。实验结果表明: 该方法能在多种模态组合下保持稳健的地形识别性能, 并在复杂地形中显著提升步态切换的稳定性。

关键词: 四足机器人; 多模态融合; 地形分割; 步态规划; 弗劳德数

中图分类号: TP242.6 **文献标志码:** A

Adaptive Gait Planning for Quadruped Robots Based on Multi-modal Terrain Perception

Wang Yuqian¹, Wu Yongcun¹, Liu Kunlun², Xiang Xuefu¹, Zhang Lu³

(1. Department of System General, Automation Research Institute Co., Ltd. of China South Industries Group Corporation, Mianyang 621000, China; 2. PLA Military Representative Office in Guangyuan District of Chongqing District Military Representative Bureau of Army Armaments Department, Guangyuan 628000, China; 3. Department of UAV, Automation Research Institute Co., Ltd. of China South Industries Group Corporation, Mianyang 621000, China)

Abstract: A unified multi-modal terrain segmentation and adaptive gait planning method for quadruped robot is proposed to solve the problems of modal uncertainty and lack of continuity in gait switching. The multi-modal terrain segmentation framework is used to enhance the adaptability to different modal combinations, and the Froude number constraint is introduced based on the terrain recognition results to achieve physically continuous gait transition. The experimental results show that the proposed method can maintain robust terrain recognition performance under multiple modal combinations, and significantly improve the stability of gait switching in complex terrain.

Keywords: quadruped robot; multi-modal fusion; terrain segmentation; gait planning; Froude number

0 引言

四足机器人凭借优越的地形适应能力和机动性能, 已广泛应用于野外巡检、装备运输与灾后救援等场景。要在复杂环境中实现稳定自主移动, 机器人既需要可靠的地形感知作为决策依据, 也需要在不同步态之间进行自适应调节以维持运动稳定性。然而实际部署中, 环境扰动往往导致传感器模态退化或暂时缺失, 同时不同行走步态之间存在显著的动力学差异。这些因素会削弱感知与运动之间的协同作用, 从而使自主移动的可靠性面临挑战。

围绕地形感知, 现有方法多依赖预设输入模态, 例如仅使用 RGB 图像^[1], 或将 RGB 与点云^[2]、IMU 数据^[3]融合。当关键模态在恶劣条件下退化甚至不可用时, 此类方法往往难以维持有效感知。尽管多

模态语义分割能提升鲁棒性, 但现有方法仍存在 2 类典型局限: 1) 以 RGB 为主其他模态为辅^[4], 在 RGB 失效时性能显著下降; 2) 每种模态配置独立编码器^[5-6], 虽然缓解模态偏置, 但参数规模与计算开销随模态数线性增长且融合结构更复杂。总体而言, 现有方法难以同时兼顾“缺模态场景下的稳健性”与“可部署的计算开销”。

即便感知层输出较可靠的地形信息, 执行层仍需实现物理上连续且稳定的步态过渡, 否则高质量感知也难以转化为安全可控的运动行为。当前步态切换多依据地形类别进行规则式切换或直接瞬时跳变, 容易引发冲击、速度突变与姿态抖动。部分方法采用预定义调度表或固定时长的中间过渡策略^[7-8], 但对环境变化与机器人状态缺乏自适应; 也有工作尝试基于弗劳德数设计切换逻辑^[9], 但流程

收稿日期: 2024-12-11; 修回日期: 2025-01-22

基金项目: 国防基础科研项目(JCKY2024209C001)

第一作者: 王宇倩(2001—), 女, 河北人, 硕士。

繁琐且多局限于平坦地面的仿真验证。强化学习方法^[10-11]虽具潜力，但训练成本高、对仿真依赖强，并面临显著的仿真到实物迁移瓶颈，难以直接满足装备级应用需求。

笔者提出一种集成多模态地形分割与步态自适应规划的统一方法。构建多模态地形分割框架，使模型在不同模态组合下仍能保持稳定识别；并基于地形结果设计自适应步态选择机制，引入弗劳德数实现物理上合理的平滑过渡。实验结果表明：该方法在多种模态组合下表现出稳定的地形识别能力，同时基于弗劳德数的步态过渡设计显著改善了步态

切换的连续性与稳定性。

1 多模态地形分割框架

在户外复杂环境中，四足机器人可利用 RGB 相机、红外相机或事件相机等多种视觉传感器获取地形信息，但雨雾、遮挡、强光等因素易导致部分模态失真甚至暂时缺失，从而使地形感知不稳定。为提升不同模态组合与关键模态退化场景下的鲁棒性，笔者提出一种多模态地形分割框架。如图 1 所示，该框架由输入模态预处理、统一模态编码器、跨模态融合模块与多层感知解码器构成。

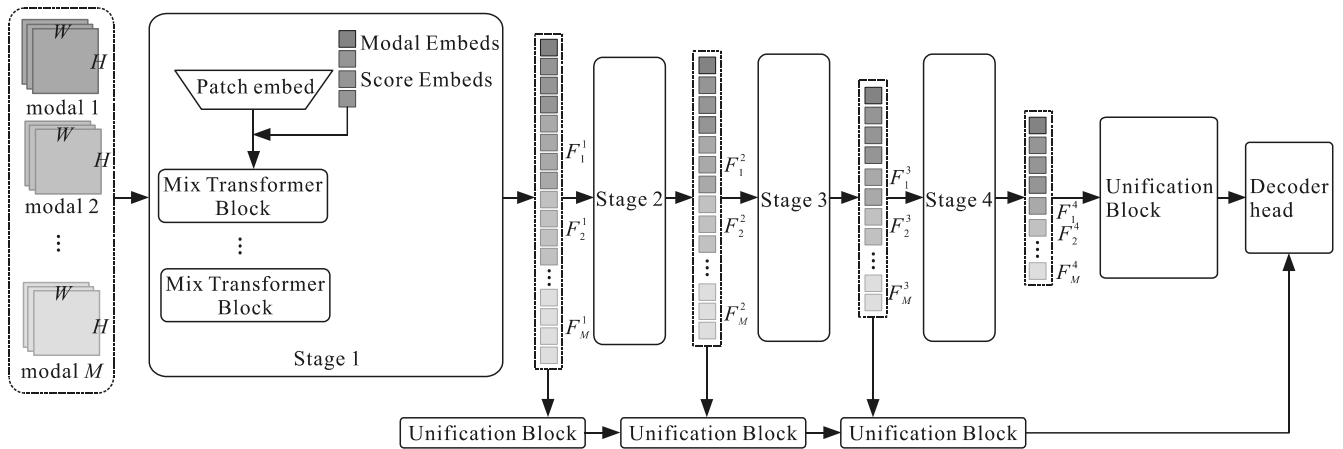


图 1 多模态地形分割框架

1.1 输入模态预处理

不同模态的原始数据在通道数 C_m 、数值分布与信息形式上存在差异。为使后续网络以一致结构处理多种模态，对每种模态 X_m 构建通道映射算子将其统一转换为 3 通道格式的输入：

$$\Phi_m(X_m) = \begin{cases} \text{Norm}(X_m), C_m = 3 \\ \text{Norm}(X_m)^{\times 3}, C_m = 1 \\ \sin(X_m), \cos(X_m), \sin(X_m), X_m \in \text{Angle} \\ f_{\text{custom}}(X_m), \text{other} \end{cases} \quad (1)$$

上述预处理确保来自不同模态的数据均可以相同格式进入主干网络，为后续编码与跨模态融合提供一致的输入基础。

1.2 统一模态编码器

基于统一后的模态输入 I_m ，笔者构建共享参数的统一模态编码器，对所有模态采用相同结构提取特征，以获得结构一致的模态表征并避免参数膨胀。首先通过图像分块映射得到初始特征：

$$E_m^0 = \text{PatchEmbed}(I_m) \in R^{N \times d} \quad (2)$$

式中： N 为图像被划分后的 Patch 数量； d 为嵌入

维度。

在第一个编码阶段引入 2 类可学习令牌：1) 模态先验令牌 E_{prior}^0 用于描述当前采用的模态配置；2) 模态置信令牌 E_{conf}^0 用于刻画该模态在当前输入下的可靠性。将初始特征与令牌在特征维度拼接作为编码器输入：

$$F_m^0 = \text{Concat}(E_m^0, E^{\text{modal}}, E^{\text{score}}) \quad (3)$$

统一模态编码器由 4 个阶段的 Transformer 块堆叠组成，逐阶段提取更高层语义特征；共享编码与令牌引导使模型在保留模态差异信息的同时，降低对单一关键模态（如 RGB）的依赖。

1.3 跨模态融合模块

在第 s 阶段，融合模块依据模态置信令牌自适应分配各模态权重，以增强高置信模态并抑制退化模态。对第 m 个模态的权重计算为：

$$W_m^s = \sigma(\text{MLP}(e_m^{\text{conf}})), e_m^{\text{conf}} \in E_{\text{conf}}^s \quad (4)$$

式中 $\sigma(\bullet)$ 为 Sigmoid 函数，用于确保 $W_m^s \in (0, 1)$ 。

随后将各模态阶段特征按权重加权求和，并注

入该阶段模态先验令牌, 以保持模态组合整体先验:

$$u^s = \sum_{m=1}^M (W_m^s \cdot F_m^s) + E_{\text{prior}}^s \quad (5)$$

式中 E_{prior}^s 为第 s 阶段的模态先验令牌。该融合机制能够在模态缺失甚至退化时自动调整贡献, 使跨模态表示更稳定一致。

1.4 多层感知解码器

融合后得到 4 个尺度的统一特征 $U = \{u^1, u^2, u^3, u^4\}$, $u^s \in R^{N_s \times d}$, 解码器将各尺度特征上采样到同一空间分辨率:

$$h^s = \text{Upsample}^s(u^s) \quad (6)$$

随后在通道维度进行拼接, 并通过多层感知机对多尺度语义进行融合, 得到聚合后的特征表示 h_{agg} 。最后, 利用分类头对每个像素进行类别预测, 生成最终的地形分割图:

$$S = \text{MLP} - \text{Cls}(h_{\text{agg}}) \in R^{H \times W \times K} \quad (7)$$

式中: K 为地形类别数; S 为最终的地形分割结果。

2 基于弗劳德数的步态自适应规划

在获得地形分割结果后, 机器人需根据地形选择合适步态, 并在步态变化时保持运动稳定性。不同地形对落足方式与身体姿态要求不同, 不同步态之间的动力学差异又会使直接切换产生冲击与抖动。为此, 笔者基于地形识别结果构建初始步态选择, 并进一步引入弗劳德数约束的平滑过渡机制, 实现对复杂地形的自适应步态规划。

2.1 地形驱动的初始步态选择

地形分割模块输出的分割图为 $S \in R^{H \times W \times K}$, 每个像素类别为 $T(x, y) \in \{1, 2, \dots, K\}$ 。考虑机器人足端落脚区域最能反映实际接触地面特性, 笔者在落脚区域 Ω 内统计主导地形类别作为当前地形估计:

$$\hat{T} = \text{Mode}(T(x, y)), \quad (x, y) \in \Omega \quad (8)$$

随后通过经验映射函数快速给出可行的初始步态类型:

$$G_{\hat{T}} = f_{\text{terrain}}(\hat{T}) \quad (9)$$

为便于后续统一调度, 笔者为每类步态维护一组标准参数(如步态周期 T 、占空比 r 、相位偏移 p 等), 用于描述步态节律与支撑/摆动关系, 如表 1 所示。

表 1 步态参数库

步态类型	步态周期/s	占空比	相位偏移
行走	0.80	0.75	(0, 0.25, 0.50, 0.75)
慢速小跑	0.60	0.60	(0, 0.5, 0.5, 0)
小跑	0.50	0.50	(0, 0.5, 0.5, 0)
快速小跑	0.40	0.45	(0, 0.5, 0.5, 0)
腾空小跑	0.25	0.40	(0, 0.5, 0.5, 0)
小跑	0.50	0.50	(0, 0.5, 0.5, 0)

2.2 基于弗劳德数的步态平滑切换

当旧步态切换到新步态时, 若直接替换参数易造成速度突变与姿态冲击。为此, 笔者引入基于弗劳德数的步态平滑切换机制, 用于描述当前速度下的运动状态, 并据此调节步态连续过渡过程。机器人身体速度记为 v , 腿长记为 l , 则弗劳德数定义为:

$$F_r = v / \sqrt{gl} \quad (10)$$

式中 g 为重力加速度。弗劳德数能够反映在当前速度条件下机器人应采用的步态节律。

考虑从旧步态参数 (T_1, r_1, p_1) 过渡到新步态参数 (T_2, r_2, p_2) , 令过渡时长随当前速度对应的弗劳德数变化:

$$T_{\text{trans}} = \frac{T_1 + T_2}{2} \left(1 - \frac{F_r^v}{F_r^{v_{\text{max}}} + 1} \right) \quad (11)$$

式中: F_r^v 为当前速度对应的弗劳德数; $F_r^{v_{\text{max}}}$ 为最大可行速度下的弗劳德数。速度越快, 过渡时间越短, 以满足快速响应需求。

设 t_{trans} 表示从过渡开始至当前的时间, 则前后步态参数的融合权重可表示为:

$$\left. \begin{aligned} w_1 &= 1 - t_{\text{trans}} / T_{\text{trans}} \\ w_2 &= t_{\text{trans}} / T_{\text{trans}} \end{aligned} \right\} \quad (12)$$

在整个过渡过程中, 步态参数按权重进行融合, 使得机器人能够在不同地形间实现连续稳定的步态转换, 并与图 2 所示的控制器协同工作抑制切换冲击。

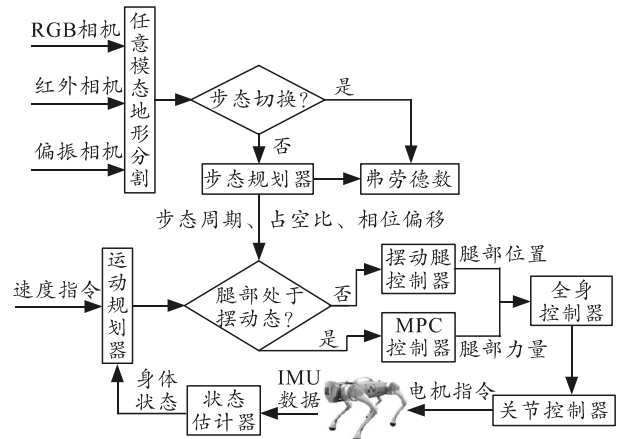


图 2 步态自适应规划流程

3 实验分析

3.1 实验设置

笔者构建了四足机器人模型，并为其配置了与真实系统一致的虚拟 RGB 相机、红外相机、IMU 与关节传感器。地形识别的输入来自虚拟相机渲染的前视图像。为了使仿真感知数据与地形分割网络保持一致，笔者对仿真图像沿底部区域进行裁剪，并缩放至 512×512 的标准输入格式。仿真中的地形

识别模块以 5 Hz 运行，而步态调度器、规划器和控制器以 500 Hz 运行，以确保控制循环的稳定性。仿真环境中配置了多种典型地形，包括沥青、草地、沙地等，以验证笔者提出的地形识别与步态自适应切换方法的有效性和鲁棒性。

3.2 实验结果

为了评估地形感知模块在多模态完整输入条件下的地形分割性能，在全模态设置下进行了对比实验，结果如表 2—3 所示。

表 2 全模态设置下的实验结果

%

方法	沥青	混凝土	道路标线	沙子	砾石	鹅卵石	砖地	草地	水	均值
CMNeXt	84.7	45.2	74.8	67.8	67.1	68.7	43.2	58.9	54.4	62.76
MMSFormer	86.5	61.8	70.5	63.7	68.4	66.8	45.8	77.2	46.6	65.25
Magic	88.9	50.8	75.0	66.9	66.8	71.6	46.5	55.3	54.7	64.05
Ours	86.7	66.6	70.9	63.3	62.0	68.0	51.4	76.6	50.9	66.27

表 3 模态无关设置下的实验结果

%

方法	R	A	D	RA	RD	RN	AD	AN	DN	RAD	RAN	RDN	ADN	Mean
CMNeXt	1.86	1.54	2.51	47.96	43.67	45.90	6.99	7.58	9.95	50.07	48.77	48.83	8.06	23.28
MMSFormer	64.43	3.85	3.28	64.87	64.55	64.69	3.70	3.57	3.34	65.00	65.10	64.80	3.40	34.13
Magic	51.91	0.32	34.52	52.24	52.16	52.57	1.98	36.01	37.09	52.48	52.82	52.73	37.57	36.93
Ours	57.26	21.19	6.10	61.51	63.21	61.52	23.14	22.80	6.70	63.67	63.61	61.44	24.82	38.79

在全模态设置中，训练与测试均使用 MCubes 数据集^[12]提供的 4 类传感模态，以评估信息充分条件下的融合能力。表 2 给出了各方法在 9 类地形上的分类准确率。结果显示，本文中方法的总体准确率分别较 CMNeXt、MMSFormer 和 Magic 提升 3.51%、1.02% 和 2.22%，并在沥青、混凝土、道路标线、鹅卵石、砖块、水面等多数类别上取得最优或近最优精度，表明统一编码与融合设计能够更充分利用多模态互补信息。

在模态无关设置中，训练阶段保持四模态输入，测试阶段随机采用不同模态组合以模拟模态退化/缺失。表 3 结果表明：本文中方法在大多数组合下达到最优或接近最优，整体平均准确率较此前最优方法提升 1.86%。值得注意的是，MMSFormer 在包含 RGB 时表现较好，但在缺少 RGB 时性能显著下降，例如仅使用 AoLP、DoLP 和 NIR 时准确率仅 3.40%，体现出对 RGB 的强依赖。相比之下，本文中方法能够有效挖掘 AoLP 与 NIR 中的地形判别信息，在无 RGB 场景下仍保持稳定识别。

为验证基于弗劳德数的步态参数自适应切换机制的稳定性，笔者将其与 2 种常用策略对比：固定时长 (0.5 s) 过渡与瞬时切换。在固定速度下多次执行“行走→腾空小跑”连续转换，并统计转换后 1 s 内俯仰角与横滚角 RMSE。结果显示，本文中方法在所有转换中姿态扰动最低：横滚角 RMSE 相较固

定过渡降低约 52.8%、较瞬时切换降低约 10.7%；俯仰角 RMSE 分别降低约 49.7%与 16.5%。上述结果表明，该方法可显著抑制步态切换引起的姿态抖动，使过渡更加平稳可靠。

4 结论

笔者提出一种面向四足机器人的多模态地形分割方法与基于弗劳德数的步态自适应切换策略。实验结果表明：在 Mcubes 数据集上，该地形分割方法在全模态与模态无关设置下的整体准确率分别提升 1.02%和 1.86%，在多模态组合及模态缺失场景中均表现出更高的泛化性与鲁棒性。同时，基于弗劳德数的步态平滑切换机制有效降低了转换过程中的姿态扰动，显著提升了步态切换的稳定性。总体来看，笔者所提方法在感知可靠性与运动连续性方面均取得明显改进，为四足机器人在复杂环境中的自主行走提供了一种高效、稳定且具备环境鲁棒性的解决方案。

参考文献：

- [1] CAO J, CHEN T, JIANG H, et al. Autonomous Locomotion Control Strategy for Quadruped Robot Based on Visual Lidar Fusion Terrain Recognition[C]//The 2023 2nd International Conference on Automation, Robotics and Computer Engineering (ICARCE). IEEE, 2023.
- [2] LU Z, WANG M, YU S, et al. Camera-LiDAR-based terrain multi-type classification using both spatial and

- histogram features of LiDARs[C]//The 2023 3rd International Conference on Computer, Control and Robotics (ICCCR). IEEE, 2023.
- [3] HUANG H, YANG Y, TANG L, et al. A Multimodal Off-Road Terrain Classification Benchmark for Extraterrestrial Traversability Analysis[C]//The 2022 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC). IEEE, 2022.
- [4] ZHANG J, LIU R, SHI H, et al. Delivering arbitrary-modal semantic segmentation[C]//The Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2022.
- [5] REZA M K, PRATER-BENNETTE A, ASIF M S. Mmsformer: Multimodal transformer for material and semantic segmentation[C]//IEEE Open Journal of Signal Processing. IEEE, 2024.
- [6] ZHENG X, LYU Y, ZHOU J, et al. Centering the value of every modality: Towards efficient and resilient modality-agnostic semantic segmentation[C]//The European Conference on Computer Vision. Springer, 2024.
- [7] SUN H, YANG J, JIA Y, et al. Free gait generation of quadruped robots via impulse-based feasibility analysis[J]. IEEE/ASME Transactions on Mechatronics, 2023, 29(1): 412–22.
- [8] LI Q, QIAN L, WANG S, et al. Towards generation and transition of diverse gaits for quadrupedal robots based on trajectory optimization and whole-body impedance control[J]. IEEE Robotics and Automation Letters, 2023, 8(4): 89–96.
- [9] HUMPHREYS J, LI J, WAN Y, et al. Bio-inspired gait transitions for quadruped locomotion[J]. IEEE Robotics and Automation Letters, 2023, 8(10): 1–8.
- [10] YANG Y, ZHANG T, COUMANS E, et al. Fast and efficient locomotion via learned gait transitions[C]//The Conference on robot learning. PMLR, 2022.
- [11] TAN W, FANG X, ZHANG W, et al. A hierarchical framework for quadruped omnidirectional locomotion based on reinforcement learning[J]. IEEE Transactions on Automation Science and Engineering, 2024, 21(4): 5367–5378.
- [12] LIANG Y, WAKAKI R, NOBUHARA S, et al. Multimodal material segmentation[C]//The IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2022.
- *****
- (上接第 107 页)
- [3] 向锦武, 董希旺, 丁文锐, 等. 复杂环境下无人集群系统自主协同关键技术[J]. 航空学报, 2022, 43(10): 333–365.
- [4] HACENE N, MENDIL B. Behavior-based autonomous navigation and formation control of mobile robots in unknown cluttered dynamic environments with dynamic target tracking[J]. Automation and Computing, 2021, 18: 766–786.
- [5] 王晶, 顾维博, 窦立亚. 基于 Leader-Follower 的多无人机编队轨迹跟踪设计[J]. 航空学报, 2020, 41(S1): 88–98.
- [6] VÁSÁRHELYI G, VIRÁGH C, SOMORJAI G, et al. Optimized flocking of autonomous drones in confined environments[J]. Science Robotics, 2018, 3(20): 3536.
- [7] 毛琼, 张代兵, 李小民, 等. 基于邻域跟随与辨识的无人机集群控制与规避策略[J]. 系统工程与电子技术, 2018, 40(9): 2071–2078.
- [8] REYNOLDS C W. Flocks, herds and schools: A distributed behavioral model[J]. ACM SIGGRAPH Computer Graphics, 1987: 273–282.
- [9] OLFATI-SABER R. Flocking for multi-agent dynamic systems: algorithms and theory[J]. IEEE Transactions on Automatic Control, 2006, 51(3): 401–420.
- [10] FU L, MAO R, YANG Y, et al. A Formation Maintenance Method of UAV Swarm under Wind Disturbances[C]//The 14th International Conference on Computer Modeling and Simulation. New York, NY, USA: Association for Computing Machinery, 2022: 160–167.
- [11] MEHMOOD U, PAOLETTI N, PHAN D, et al. Declarative vs rule-based control for flocking dynamics[C]//The 33rd Annual ACM Symposium on Applied Computing. New York, NY, USA: Association for Computing Machinery, 2018: 816–823.
- [12] YANG Y, LUO R, LI M, et al. Mean field multi-agent reinforcement learning[C]//International conference on machine learning. PMLR, 2018: 5571–5580.
- [13] SAIF O, FANTONI I, ZAVALA-RÍO A. Distributed integral control of multiple UAVs: precise flocking and navigation[J]. IET Control Theory & Applications, 2019, 13(13): 2008–2017.
- [14] KAUSAR Z, AUJLIA G S, ALI S, et al. A Novel Controller of Multi-Agent System Navigation and Obstacle Avoidance[J]. Procedia Computer Science, 2020, 171: 441–447.
- [15] 张令, 段海滨, 雍婷, 等. 基于寒鸦配对交互行为的无人机集群编队控制[J]. 北京航空航天大学学报, 2021, 47(2): 391–397.
- [16] PARK S, DEYST J, HOW J P. A New Nonlinear Guidance Logic for Trajectory Tracking[C]//AIAA Guidance, Navigation, and Control Conference and Exhibit. Reston: AIAA, 2004: 941–956.
- [17] 潘无为. 分布式多水下机器人编队控制方法研究[D]. 哈尔滨: 哈尔滨工程大学, 2018.
- [18] 费思远, 鲜斌, 王岭. 基于群集行为的分布式多无人机编队动态避障控制[J]. 控制理论与应用, 2022, 39(1): 1–11.
- [19] DRAZIN P G, DRAZIN P D. Nonlinear systems[M]. Cambridge: Cambridge University Press, 1992.
- [20] REN W, BEARD R W. Consensus seeking in multiagent systems under dynamically changing interaction topologies[J]. automatic control, 2005, 50(5): 655–661.