

2-15-2024

Research on Motion Planning of Hexapod Robot Based on DRL and Free Gait

Xinpeng Wang

School of Manufacturing Science and Engineering, Southwest University of Science and Technology, Mianyang 621010, China; Manufacturing Process Testing Technology Key Laboratory of the Ministry of Education, Mianyang 621000, China, xpwang@mails.swust.edu.cn

Huiqiao Fu

School of Management and Engineering, Nanjing university, Nanjing 210046, China

Guizhou Deng

School of Manufacturing Science and Engineering, Southwest University of Science and Technology, Mianyang 621010, China; Manufacturing Process Testing Technology Key Laboratory of the Ministry of Education, Mianyang 621000, China

Kaiqiang Tang

School of Management and Engineering, Nanjing university, Nanjing 210046, China; Manufacturing Process Testing Technology Key Laboratory of the Ministry of Education, Mianyang 621000, China, kqtang@smail.nju.edu.cn

See next page for additional authors

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research](#), [Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation. For more information, please contact xtfzxb@126.com.

Research on Motion Planning of Hexapod Robot Based on DRL and Free Gait

Abstract

Abstract: To improve the passability and the motion performance of the hexapod robot in the unstructured environment, a multi-contact motion planning algorithm based on DRL and free gait planner is proposed. Firstly, the free gait planner obtains the reachable footholds under the target state and outputs the optimal gait sequence. The center of mass motion policy of the hexapod robot in the randomly generated plum blossom pile environment is obtained by using deep reinforcement learning training. To ensure the reachability between adjacent states of the robot in motion, the state transition feasibility model is used to judge the state transition feasibility. Finally, the foothold planning of the hexapod robot in the plum blossom pile environment with gullies of different widths is realized. Simulation and physical experiments show that the multi-contact motion planning algorithm can make the robot reach the target area quickly and smoothly from the starting point, and automatically adjust the gait pattern to deal with the randomly distributed plum blossom piles in different environments.

Keywords

hexapod robot, free gait, DRL, multi-contact motion planning, unstructured environment

Authors

Xinpeng Wang, Huiqiao Fu, Guizhou Deng, Kaiqiang Tang, Chunlin Chen, and Canghao Liu

Recommended Citation

Wang Xinpeng, Fu Huiqiao, Deng Guizhou, et al. Research on Motion Planning of Hexapod Robot Based on DRL and Free Gait[J]. Journal of System Simulation, 2024, 36(2): 373-384.

基于 DRL 和自由步态的六足机器人运动规划研究

王鑫鹏^{1,3}, 傅汇乔², 邓归洲^{1,3}, 唐开强^{2,3*}, 陈春林², 留沧海^{1,3}

(1. 西南科技大学 制造科学与工程学院, 四川 绵阳 621010; 2. 南京大学 工程管理学院, 江苏 南京 210046;
3. 制造过程测试技术省部共建教育部重点实验室, 四川 绵阳 621000)

摘要: 为提高六足机器人在非结构环境下的通过率和运动性能, 提出一种基于 DRL 和自由步态规划器的多接触运动规划算法。自由步态规划器获取目标状态下可达落脚点从而输出最优步态序列; 利用 DRL 训练得到六足机器人在随机生成的梅花桩环境中的质心运动策略。为了保证机器人在运动过程中相邻状态之间的可达性, 利用状态转移可行性模型对状态转移可行性进行判定, 实现六足机器人在不同宽度沟壑梅花桩环境下的落脚点规划。仿真与样机实验表明: 多接触运动规划算法能够让机器人快速平稳地从起点到达目标区域, 并自动调整步态模式以应对不同环境下随机分布的梅花桩。

关键词: 六足机器人; 自由步态; 深度强化学习; 多接触运动规划; 非结构环境

中图分类号: TP242.6 文献标志码: A 文章编号: 1004-731X(2024)02-0373-12

DOI: 10.16182/j.issn1004731x.joss.22-1220

引用格式: 王鑫鹏, 傅汇乔, 邓归洲, 等. 基于 DRL 和自由步态的六足机器人运动规划研究[J]. 系统仿真学报, 2024, 36(2): 373-384.

Reference format: Wang Xinpeng, Fu Huiqiao, Deng Guizhou, et al. Research on Motion Planning of Hexapod Robot Based on DRL and Free Gait[J]. Journal of System Simulation, 2024, 36(2): 373-384.

Research on Motion Planning of Hexapod Robot Based on DRL and Free Gait

Wang Xinpeng^{1,3}, Fu Huiqiao², Deng Guizhou^{1,3}, Tang Kaiqiang^{2,3*}, Chen Chunlin², Liu Canghai^{1,3}

(1. School of Manufacturing Science and Engineering, Southwest University of Science and Technology, Mianyang 621010, China;
2. School of Management and Engineering, Nanjing university, Nanjing 210046, China;
3. Manufacturing Process Testing Technology Key Laboratory of the Ministry of Education, Mianyang 621000, China)

Abstract: To improve the passability and the motion performance of the hexapod robot in the unstructured environment, a multi-contact motion planning algorithm based on DRL and free gait planner is proposed. Firstly, the free gait planner obtains the reachable footholds under the target state and outputs the optimal gait sequence. The center of mass motion policy of the hexapod robot in the randomly generated plum blossom pile environment is obtained by using deep reinforcement learning training. To ensure the reachability between adjacent states of the robot in motion, the state transition feasibility model is used to judge the state transition feasibility. Finally, the foothold planning of the hexapod robot in the plum blossom pile environment with gullies of different widths is realized. Simulation and physical experiments show that the multi-contact motion planning algorithm can make the robot reach the target area quickly and smoothly from the starting point, and automatically adjust the gait pattern to deal with the randomly distributed plum blossom piles in different environments.

Keywords: hexapod robot; free gait; DRL; multi-contact motion planning; unstructured environment

收稿日期: 2022-10-14 修回日期: 2022-11-07

基金项目: 国家自然科学基金(62073160); 国家重点研发计划(72018AAA0101100)

第一作者: 王鑫鹏(1995-), 男, 硕士生, 研究方向为深度强化学习和机器人技术。E-mail: xpwang@mails.swust.edu.cn

通讯作者: 唐开强(1992-), 男, 博士生, 研究方向为强化学习和机器人技术。E-mail: kqtang@smail.nju.edu.cn

0 引言

六足机器人拥有多个冗余的自由度和离散的落足点，其稳定性与运动性能相较于轮式与履带式机器人具有显著的优势^[1-2]，在非结构环境救援任务，自然灾害检测^[3]方面具有广泛的应用前景。六足机器人复杂的构造也对非结构环境下足端的多接触运动规划提出了很高的要求，因此，如何提高其在复杂环境下的多接触运动规划效率引起了广泛的关注^[4]。文献[5]对六足机器人仿生结构、步态规划、多足协调控制等关键技术进行分析，并以单足-机体一体化结构设计与优化、基于节律与学习行为的步态规划、高效的多足协调控制方法这3点为切入点对如何挖掘六足机器人运动潜能进行了展望。以步态规划为例，传统的周期性步态摆动相与支撑相按照确定好的顺序依次运动，具有计算量小，简单方便的特点，因此周期性步态得到了广泛应用^[6-7]，然而，在落足点随机分布的非结构复杂环境下，周期性步态一定程度上限制了机器人的通过率。自由步态则为一种非周期的步态，摆动相与支撑相没有固定的顺序，机器人根据落足点分布合理地选择下一时刻所需迈的腿，提升了对地形的适应能力^[8-10]。文献[11]将运动规划过程视为单步优化过程，分别规划步态序列和立足点。文献[12]利用强化学习算法生成速度指令，机器人根据速度指令生成稳定的步态。文献[13]根据安全立足点和稳定裕度来选择最佳立足点并改变步态。然而，上述多数方法仅考虑单一的稳定性或者运动学等约束。

文献[14]则更加关心机器人运动学、动力学、路径规划以及控制系统等方面，值得注意的是该文指出在六足机器人控制方面，强化学习将是未来的趋势。以深度强化学习为例，最新成果展示了深度强化学习在复杂环境下运动控制方面的先进性^[15-17]。为了提高六足机器人在非结构环境下的运动效率，文献[18]提出一种先进的深度强化学习算法，在非结构环境中取得了较好的运动规划效

果，然而，该算法默认使用三足步态作为基本步态。文献[19]提出具有两级层次结构的深度强化学习方法，分别用于步态规划和运动控制。以上方法采用的步态较为单一，这种处理方式将导致六足机器人在稀疏立足点环境中的通过率较低。

本文讨论了六足机器人在非结构环境下的自由步态多接触运动规划问题。设计了一种落足点检测器，用以获取目标状态下符合运动学约束的落足点 L_R ；提出自由步态规划器，用于获取目标状态下最优步态序列 L_{sw} ；从对角高斯分布中采样连续质心运动序列，并使用一种最大熵深度强化学习算法 (soft actor critic, SAC)对运动序列进行优化。

1 六足机器人运动学建模

六足机器人运动学模型描述了其足端与关节转角的相对运动关系。六足机器人每条腿有3个关节，分别是基节、髋关节、膝关节，分别对应了坐标系 $x_0y_0z_0$ ，坐标系 $x_1y_1z_1$ ，坐标系 $x_2y_2z_2$ ，坐标系 $x_3y_3z_3$ 为六足机器人足端坐标系，其中， l_1 、 l_2 、 l_3 分别为基节，大腿，小腿长度。六足机器人单腿结构连杆如图1所示。

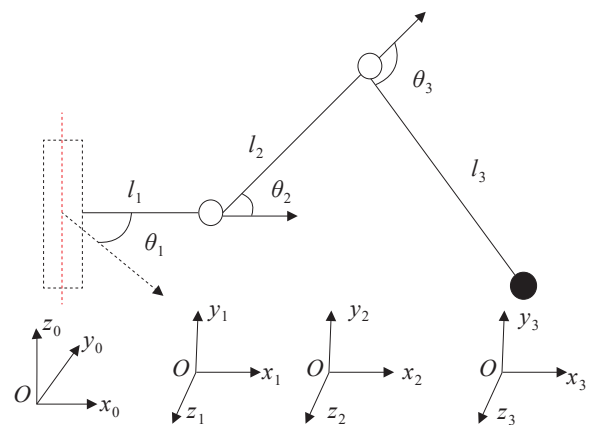


图1 六足机器人单腿结构连杆

Fig. 1 Connecting rod of hexapod robot single leg structure

本文利用标准D-H建模法对单腿进行正逆运动学求解，D-H参数见表1，其中， i 代表连杆编号； α_i 代表关节扭转角； a_i 代表连杆长度； d_i 代表连杆偏移； θ_i 代表关节转角。每个关节的齐次变换

矩阵为

$$T_i^{(i-1)} = \text{Trans } z(d_i) \text{Rot } z(\theta_i) \text{rans } x(a_i) \text{Rot } x(\alpha_i) \quad (1)$$

表1 六足机器人D-H参数

i	$\alpha_i/(\circ)$	a_i	d_i	θ_i
1	90	l_1	0	θ_1
2	0	l_2	0	θ_2
3	0	l_3	0	θ_3

$T_i^{(i-1)}$ 是坐标系*i-1*到坐标系*i*的变换, 其中, i 为转动关节。展开后的变换矩阵为

$$T_i^{(i-1)} = \begin{pmatrix} \cos \theta_i & -\sin \theta_i \cos \alpha_i & \sin \theta_i \sin \alpha_i & a_i \cos \theta_i \\ \sin \theta_i & \cos \theta_i \cos \alpha_i & -\cos \theta_i \sin \alpha_i & a_i \sin \theta_i \\ 0 & \sin \alpha_i & \cos \alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2)$$

对坐标系进行齐次变换可以得到从基节坐标系到足端坐标系的变换矩阵:

$$T_3^0 = T_1^0 T_2^1 T_3^2 = \begin{bmatrix} n_x & o_x & a_x & p_x \\ n_y & o_y & a_y & p_y \\ n_z & o_z & a_z & p_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

令 $N=(n_x, n_y, n_z)^T$, $O=(o_x, o_y, o_z)^T$, $A=(a_x, a_y, a_z)^T$, $P=(p_x, p_y, p_z)^T$, 其中, N 为足端的法向量; O 为足端的端面矢量; A 为足端的逼近矢量; P 为足端的位置矢量。

$$\begin{cases} p_x = l_1 \cos \theta_1 + l_2 \cos \theta_1 \cos \theta_2 + \\ \quad l_3 \cos \theta_1 \cos \theta_2 \cos \theta_3 - l_3 \cos \theta_1 \sin \theta_2 \sin \theta_3 \\ p_y = l_1 \sin \theta_1 + l_2 \sin \theta_1 \cos \theta_2 + \\ \quad l_3 \sin \theta_1 \cos \theta_2 \cos \theta_3 - l_3 \sin \theta_1 \sin \theta_2 \sin \theta_3 \\ p_z = l_2 \sin \theta_2 + l_3 \cos \theta_2 \sin \theta_3 + l_3 \sin \theta_2 \cos \theta_3 \end{cases}$$

联立求解可得 P , 最终可得单腿3个关节的逆运动学表达式:

$$\begin{cases} \theta_1 = \arctan \frac{p_y}{p_x} \\ \theta_2 = \arcsin \frac{p_z^2 + k^2 + l_2^2 - l_3^2}{2l_2 \sqrt{p_z^2 + k^2}} + \arctan \frac{k}{p_z} \\ \theta_3 = \arccos \frac{p_z^2 + k^2 - l_2^2 - l_3^2}{2l_2 l_3} \end{cases} \quad (4)$$

其中, $k = p_x \cos \theta_1 + p_y \sin \theta_1 - l_1$ 。

六足机器人基节, 大腿, 小腿长度参数, 以及单腿的各关节角范围见表2。

表2 六足机器人尺寸参数及关节转角

参数	基节	髌关节	膝关节
长度/mm	60	120	145.8
转角范围/ (\circ)	[-45, 45]	[0, 45]	[-135, -90]

2 基于规则的自由步态规划器

自由步态规划器的输入参数包括: 当前状态 s_c , 目标状态 s_T , 所有梅花桩的坐标 P_{all} , 停止两次的腿集合 L''_{stop} , 停止一次的腿集合 L'_{stop} , 可达落足点集合 L_R 。

2.1 落足点检测器

六足机器人落足点的选择需要满足运动学约束。已知目标状态 s_T , 可以确定机器人质心位置, 再根据运动学约束, 当质心位置确定以后, 足端简化工作空间也确定下来。为了找到目标状态下满足运动学约束的落足点, 利用最邻近节点算法(KNN)确定距离腿*i*足端 p_i 最近的梅花桩坐标 P_{ni} , 其中, $i \in \{1, 2, \dots, 6\}$, $P_{ni} \in P_{\text{all}}$, 当 $P_{ni} \in R_i$ 时, 腿*i*在目标状态下存在满足运动学约束的落足点, 将*i*添加到可达落足点集合 L_R , R_i 为腿*i*的足端简化工作空间。

2.2 不相邻规则

不相邻规则要求相邻的2条腿不能同时抬起。以支撑多边形作为六足机器人稳定性判据, 将支撑腿足端向地面投影, 当机器人质心投影位于支撑多边形内时, 机器人能够保持静态稳定。规定1~6号腿的编号依次为 h, j, k, m, n, p , 以三足步态为例, 从 L_R 中选择1、3、5号腿为摆动相, 对应腿的编号为 h, k, n , 不相邻规则可以描述为

$$|h-k| > 1, |h-n| > 1, |k-n| > 1 \quad (5)$$

图2(a)为三足步态下的支撑多边形, 机器人

处于稳定状态，图2(b)为四足支撑状态时机器人的支撑多边形，此时机器人质心位于支撑多边形之外，发生失稳，因此，为了避免极端条件下失稳的情况，规定相邻摆动腿不能同时抬起。

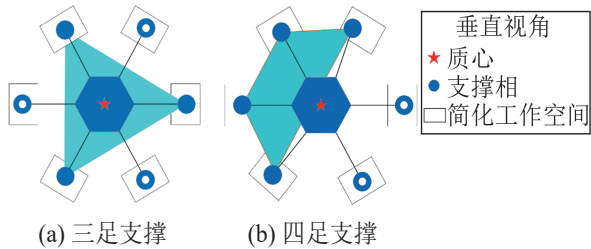


图2 支撑多边形
Fig. 2 Support polygon

2.3 最大停滞规则

最大停滞规则要求已经停留2次的腿优先进行摆动。由于六足机器人原始运动学模型是高度非线性的，为了避免机器人违反运动学约束，本文按照表2所示的六足机器人真实尺寸参数与关节转角，随机采样以生成足端运动点云，从而得到足端简化工作空间。由于采用了与真实机器人等比例的建模，因此，可以将足端简化工作空间近似为运动学约束，如图3所示。

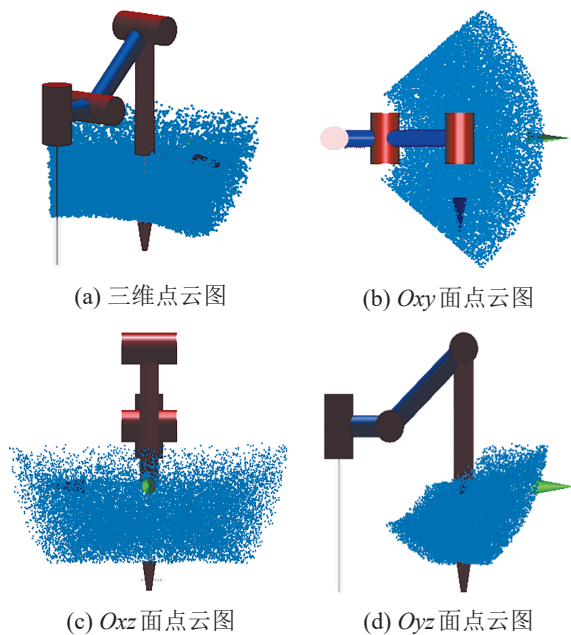


图3 足端运动空间点云图
Fig. 3 Point cloud of foot workspace

根据点云图，保守地找到腿*i*的足端简化工作空间 R_i ， R_i 可以表示为

$$p_i \in R_i(C_B, \theta) \Leftrightarrow |R_z(\alpha_i)[{}^B_w R[p_i - C_B] - \bar{p}_i]| < b \quad (6)$$

式中： ${}^B_w R$ 为世界坐标系到机身坐标系的旋转矩阵； $R_z(\alpha_i)$ 为绕*z*轴的旋转矩阵； α_i 为第*i*个髋关节相对于机身坐标系中*X*轴的偏转角度； p_i 为腿*i*的足端位置， p_i 由 p_i^{co} 和 p_i^{sw} 组成，其中， p_i^{co} 为支撑腿*i*的足端位置， p_i^{sw} 为摆动腿*i*的足端位置； \bar{p}_i 为*i*腿可达空间的近似中心； b 为足端简化工作空间最短边的一半； C_B 为质心坐标。

当机器人腿*i*的足端 p_i 始终保持在工作空间 R_i 之内时，机器人腿和机身不违反运动学约束，如图4所示。已知该工作空间最小的尺寸边长 $2b=80\text{ mm}$ ，质心运动增量区间为0~40 mm。在 t_1 、 t_2 、 t_3 时刻，机器人朝同一方向运动，腿*i*在3个时刻一直作为支撑腿，质心的最大运动距离 $d=120\text{ mm}$ ， $d > 2b$ ，此时 $p_i \notin R_i$ ，腿*i*不满足运动学约束，因此，设定同一条腿停留次数不能超过2次。

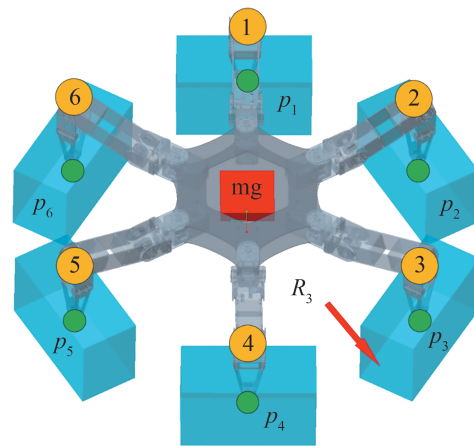


图4 六足机器人足端简化工作空间
Fig. 4 Simplified workspace of hexapod robot

如图5所示，以六号腿为例，图5(a)中六号腿在 t_1 、 t_2 、 t_3 时刻分别作支撑相、摆动相、支撑相，满足运动学约束，图5(b)中六号腿在 t_1 、 t_2 、 t_3 时刻一直作支撑相，落足点超出足端工作空间，不满足运动学约束。

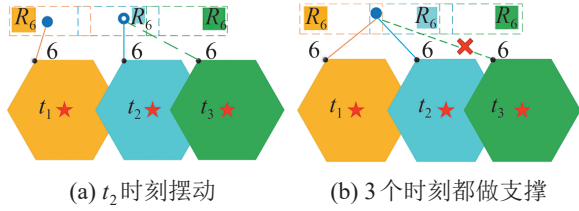


图5 3个时刻下6号腿的停滞情况

Fig. 5 Stagnation of leg 6 at three times

2.4 连续性规则

连续性规则要求已经停止一次的腿优先进行一次摆动。当落足点密集时,六足机器人在连续行走过程中,落足点检测器能够为每条腿获取一个落足点,因此,停止2次的腿集合 L''_{stop} 为空,在选择摆动相时,不再考虑最大停滞规则的约束,而是优先从停止一次的腿集合 L'_{stop} 中选择摆动腿。

根据不同规则对机器人行走的影响大小,对3个规则的优先级进行排序,其中,不相邻规则优先级>最大停滞规则优先级>连续性规则优先级。图6展示了基于规则的自由步态规划器的流程图。

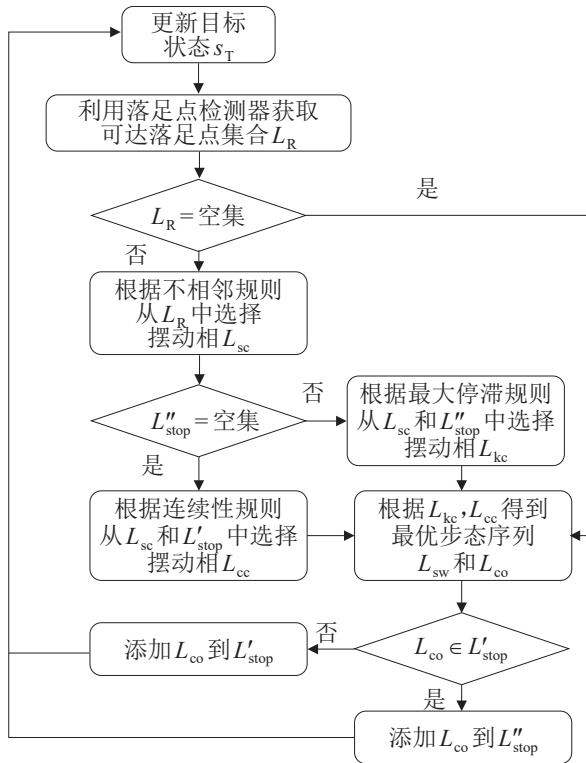


图6 自由步态规划流程图

Fig. 6 Flow chart of free gait planning

3 状态转移可行性模型

状态转移可行性模型描述了多个约束下从当前状态 s_c 转移到目标状态 s_T 的可行性。状态 s 包括质心坐标 C_B 、质心欧拉角 θ 、质心线速度 \dot{C}_B 、质心线性加速度 \ddot{C}_B 。为此本文提供当前状态 s_c 、目标状态 s_T 、支撑相坐标 p^{co} 、摆动相坐标 p^{sw} 、转移持续总时间 T 。该问题的决策变量为状态 s_i 、腿部足端接触力 u_i , $i \in \{1, 2, \dots, n_{co}\}$ 。问题的约束条件包括每条腿的运动学约束、简化动力学模型、足端作用力约束。

使用 $R_{stfm} \in \{0, 1\}$ 来描述状态转移可行性,当状态 s_i 与接触力 u_i 的轨迹满足所有约束时, $R_{stfm}=1$,否则 $R_{stfm}=0$ 。

3.1 简化动力学模型

机器人动力学中各杆件的速度,加速度是关节位置及时间的函数。为了简化六足机器人的动力学模型,通过单刚体动力学(single rigid body dynamics, SRBD)^[20]对其进行近似,根据牛顿方程,质心处的力为

$$m\ddot{C}_B = \sum_{i=1}^{n_{co}} u_i + mg \quad (7)$$

式中: m 为六足机器人的质量; g 为重力加速度。根据三维空间的欧拉方程,作用在质心处的力矩可表示为

$$I_c \varepsilon + \omega \times I_c \omega = \sum_{i=1}^{n_{co}} u_i \times (C_B - p_i^{co}) \quad (8)$$

式中: I_c 为惯量矩阵; ω 和 ε 为绕质心的角速度和角加速度。基于SRBD的六足机器人简化动力学模型与其关节状态无关,只与支撑相腿部的作用力有关。

3.2 作用力约束

六足机器人足端与梅花桩接触时,受到一个向上的推力,推力与梅花桩平面的法向同向,因此,正向力约束为

$$f_i \cdot n(p_i^{co}) \geq 0 \quad (9)$$

式中： $\mathbf{n}(p_i^{\text{co}})$ 为梅花桩表面在支撑相足端坐标 p_i^{co} 处的法向量。

除正向力外，足端在与梅花桩表面平行方向还存在静摩擦力。正向力与摩擦力形成的摩擦锥满足库伦定律，因此支撑腿所受合力保持在摩擦锥的内部，线性近似为

$$|f_i \cdot \mathbf{t}_{(1,2)}(p_i^{\text{co}})| < \mu \cdot u_i \cdot \mathbf{n}(p_i^{\text{co}}) \quad (10)$$

式中： $\mathbf{t}_{(1,2)}(p_i^{\text{co}})$ 为梅花桩表面在支撑相足端 p_i^{co} 处的切向向量； μ 为摩擦系数。

根据以上约束条件，状态转移可行性被建模为非线性约束优化问题，以当前状态 s_c 、目标状态 s_T 、支撑相坐标 p^{co} 、摆动相坐标 p^{sw} 、转移持续总时间 T 为输入变量，利用非线性求解器CasADi对该问题进行求解，若状态 s_i 与接触力 u_i 的轨迹满足所有约束时， $R_{\text{stfm}}=1$ ，否则 $R_{\text{stfm}}=0$ 。

4 最大熵深度强化学习算法

SAC是一种基于最大熵的强化学习算法^[21]，相较于其他强化学习算法更加适用于足式机器人行走技能的学习^[22]。六足机器人与梅花桩交互的过程可以描述为一个马尔可夫过程，包含了 $\langle s_t, a_t, r(s_t, a_t), s_{t+1} \rangle$ 四元组。 s_t 表示 t 时刻的状态； a_t 表示 t 时刻采取的动作； $r(s_t, a_t)$ 表示奖惩回报； s_{t+1} 表示下一时刻状态。为了鼓励六足机器人在梅花桩环境中充分探索，避免得到确定性的策略，最大熵深度强化学习以最大化带熵的累计奖励函数为目标：

$$J(\pi) = \sum_{t=0}^T E_{(s_t, a_t) \sim \rho_\pi} [r(s_t, a_t) + \alpha H(\pi(\cdot | s_t))] \quad (11)$$

式中： α 为温度系数，决定熵相对于奖励的重要程度，从而控制策略的随机程度，设置为0.9； H 为熵项； ρ_π 为状态和动作的轨迹分布。在熵 H 的约束下，六足机器人进一步探索更多的状态空间^[23]。SAC利用策略网络 π_r 将当前状态 s_t 映射到动作 a_t ，其中， π_r 由4个全连接层，一个输出层组成。 s_t 可以描述为

$$s_t := \langle p^{\text{sw}}, C_t, \theta_t, P_{\text{target}} \rangle \quad (12)$$

式中： C_t 为质心坐标； θ_t 为质心欧拉角； P_{target} 为目标区域坐标。

4.1 奖励函数

在策略训练过程中，设计一个针对于对步态选择合理性与状态转移可行性的奖励函数：

$$r_t = r_b + r_{\text{fg}} + r_{\text{stfm}} + r_g + r_{\text{step}} \quad (13)$$

式中： r_b 为机器人质心运动到边界之外的惩罚； r_{fg} 为违反运动学约束的惩罚； r_{stfm} 为不可行状态转移的惩罚； r_g 为到达终点的奖励； r_{step} 为对机器人移动步数的惩罚，促使机器人以更少的步数到达终点。

4.2 训练步骤

基于自由步态规划器，状态转移可行性模型与策略网络 π_r ，六足机器人运动策略的训练步骤如下。

步骤1：随机初始化起点与目标区域，将机器人质心初始化到起点周围。

步骤2：在任意时刻 t ，六足机器人获得当前状态 s_t ，并利用策略网络 π_r 输出动作：

$$a_t = \langle C_{t+1}, \theta_{t+1} \rangle \quad (14)$$

式中： C_{t+1} 和 θ_{t+1} 分别为机器人下一时刻的质心坐标与质心转角。基于 C_{t+1} 和 θ_{t+1} 可以获得待定状态 s'_{t+1} 。

步骤3：将动作 a_t 输入到自由步态规划器，由自由步态规划器输出最优步态序列 L_{sw} ，得到支撑相坐标 p^{co} ，摆动相坐标 p^{sw} ，并获得相应的奖励。

步骤4：将 s_t 、 s'_{t+1} 、 p^{co} 、 p^{sw} 输入到状态转移可行性模型，若 $R_{\text{stfm}}=1$ ，则更新机器人状态 $s_{t+1}=s'_{t+1}$ ，并获得相应的正奖励，若 $R_{\text{stfm}}=0$ ，那么 $s_{t+1}=s_t$ ，并获得相应的负奖励。

步骤5：重复步骤2~4，直到达到最大训练步数。

5 实验验证与分析

5.1 实验设置

为了验证多接触运动规划算法的性能, 在 Pybullet 仿真平台上构建了具有 3 种不同宽度沟壑的虚拟梅花桩环境对算法进行验证, 如图 7 所示。由于环境 I 的沟壑宽度为 0, 因此, 环境 I 中不存在青色的矩形, 沟壑长度与整个环境的边长保持一致。梅花桩环境的长和宽都为 2 000 mm, 梅花桩在 x 、 y 、 z 轴方向具有相同的高度, 梅花桩数量都为 300 个, 实验环境参数设置见表 3。在仿真和样机实验中, 假设环境信息和机器人的位姿是已知量, 仿真验证通过后, 再构建真实场景下环境 III 的简化版进行样机实验, 对算法进行进一步验证。

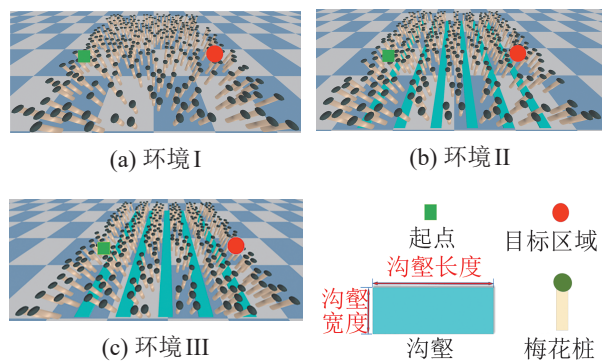


图 7 梅花桩环境设置

Fig. 7 Plum blossom pile environment setting

表 3 梅花桩实验环境设置

Table 3 Setting of plum blossom pile experiment environment

环境	沟壑宽度	mm
I	0	
II	100	
III	150	

5.2 训练设置

在训练过程中, 随机初始化起点与目标区域, 机器人需要以最短路径到达目标区域, 当机器人到达目标区域或者步数达到回合最大步数 5 000 时, 结束当前回合, 开始下一个回合。

5.3 仿真实验

以平均回合回报、平均回合步数、平均回合成功率为指标对算法进行评估。平均回合回报由回合总奖励和回合总步数的商计算得到, 平均回合回报越高表明策略越好; 平均回合步数由回合总步数与回合数的商计算得到, 平均回合步数越少表明策略越好, 机器人能够更快速到达终点; 平均回合成功率由到达终点的总次数与回合数的商计算得到, 平均回合成功率越高表明策略越好, 机器人到达终点的成功率越高。

图 8 展示了所提算法在 3 个环境下的平均回报学习曲线。在初始时刻, 环境 I 具有最大的初始奖励, 环境 III 的初始奖励最小, 原因在于环境 III 的复杂度较高, 机器人在随机探索阶段获得不可行的状态转移的概率更大, 因此负奖励越多。随着策略的优化, 3 个环境下的平均回报都呈现不断增加的趋势, 但是环境 I 的回报增长速度较快, 环境 III 的回报增长速度较慢, 且环境 I 的最终平均回报要高于环境 II 和环境 III。此外, 环境 III 中的平均回合奖励的方差最大, 表明沟壑宽度越宽, 机器人策略训练越困难。

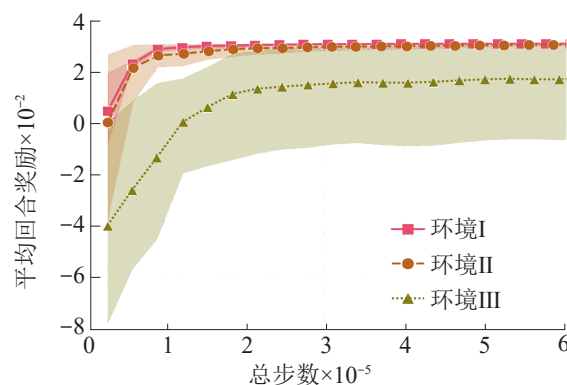


图 8 不同环境下平均回报学习曲线

Fig. 8 Learning curve of average return under different environments

六足机器人运动情况如图 9 所示。机器人在环境 I 中运动时, 每一步只需要使用较小的关节转角即可。但是在环境 II、III 中, 当机器人遇到沟壑, 且随着沟壑宽度不断增加时, 六足机器人在跨沟过

程中用到的每条腿3个关节的角度越趋于关节转角极限，因为随着沟壑宽度的增加，沟壑附近适合六

足机器人行走的落足点变少，六足机器人无法轻易找到满足运动学和动力学模型的落足点。

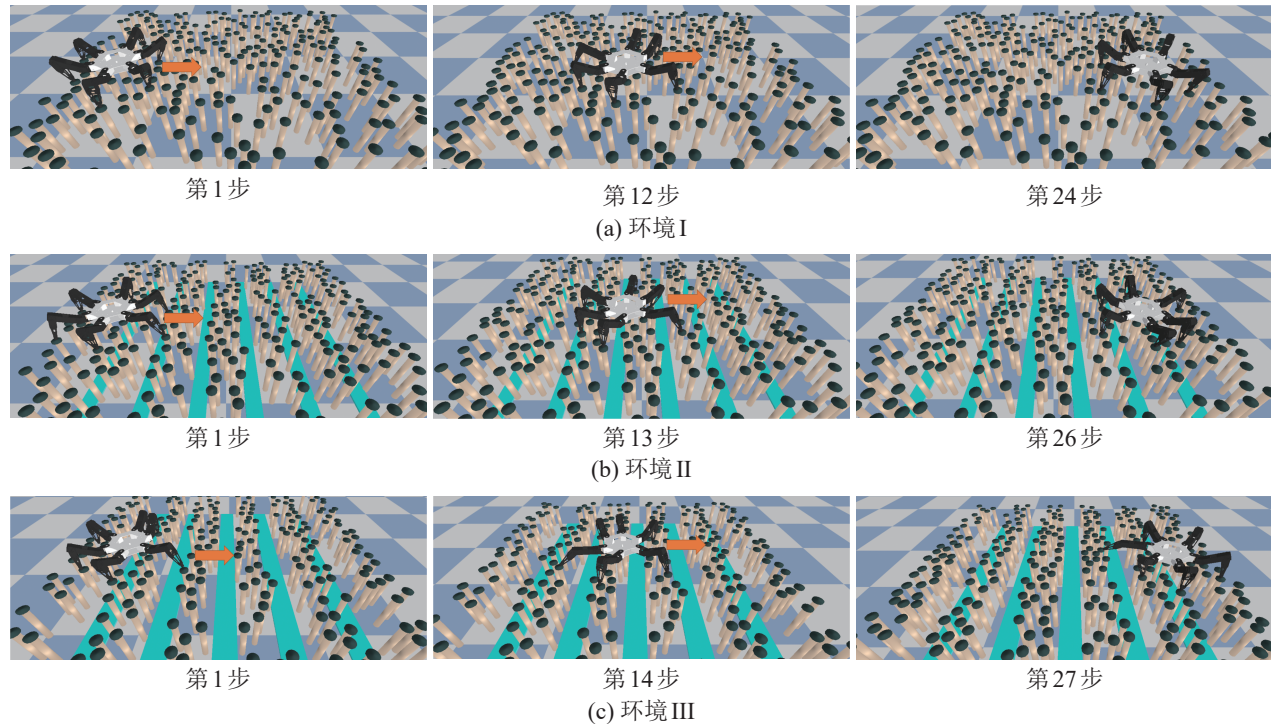


图9 六足机器人在仿真环境中的运动过程

Fig. 9 Motion process of hexapod robot in simulation environment

为了评估训练后策略的性能，在3个环境下设置了50组随机的起点与目标区域，并分别求解对应的平均回合步数，平均回合成功率，结果如表4所示。

表4 3个环境下的性能对比

环境	平均回合步数	平均回合成功率/%
I	25	100
II	31	98
III	64	52

环境I中，机器人从起点到达目标区域平均只需25步，成功率为100%。而在环境II、III中，平均回合步数随着沟壑宽度的增加而增加，且机器人从起点到达目标区域的成功率也在减少，由此可知随着沟壑宽度的增加，由于运动学与动力学的约束，机器人的通过率不断降低，六足机器人需要调整步伐，采用更多的步数来应对更宽的沟壑。

图10展示了在不同环境下的三足、四足、五足步态的占比。根据占空比与机器人运动速度的关系可知，占空比越小，机器人速度越快。在没有沟壑或者沟壑较小，落足点分布较为平均的环境中，环境易于机器人行走，因此，六足机器人更多地采用三足步态，从而更快速地到达终点。当沟壑宽度逐渐增大，落足点分布不均匀，部分区域梅花桩分布密集，部分区域梅花桩分布稀疏时，机器人稳定行走的挑战性增大，为了保证行走的稳定性，六足机器人更多地采用四足与五足步态。

此外，对六足机器人最大跨沟宽度进行了探究，分别结合图3~4所示的足端点云图、足端工作空间，确定了3组较为合理的沟壑宽度，分别为165、175、185 mm，并对3种环境下的运动规划效果进行了探究。在仿真平台中对运动策略进行验证，结果显示机器人在沟壑宽度为185 mm时，机器人发生失稳，无法到达目标区域，在沟壑宽度为

165 mm 和 175 mm 时, 机器人能够顺利到达目标区域, 如图 11 所示。然后在 175~185 mm 之间间隔 2 mm 均匀采样沟壑宽度并进行策略训练, 同样在仿真平台中对运动策略进行验证, 最终可得六足机器人最大跨沟宽度约为 175 mm。

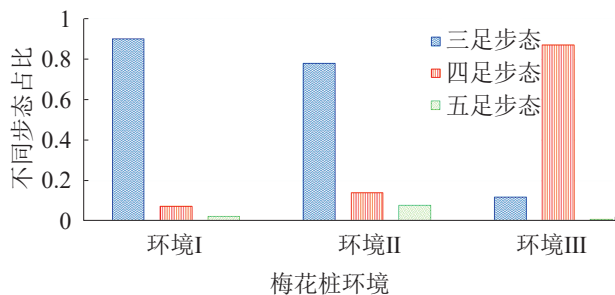
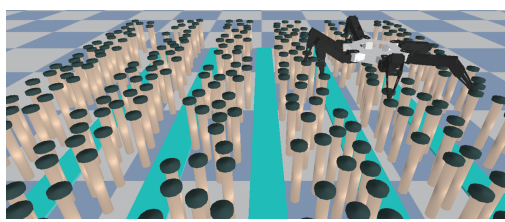
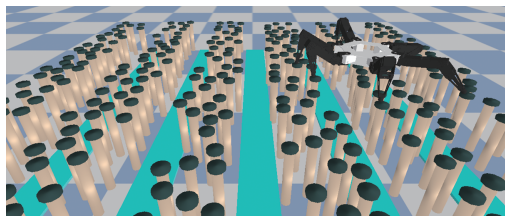


图 10 不同环境下各步态模式占比

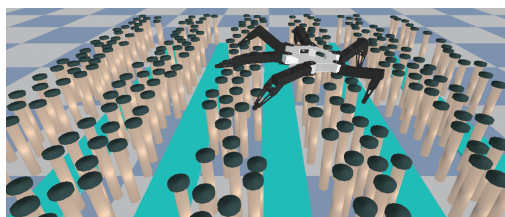
Fig. 10 Proportion of gait patterns in different environments



(a) 沟壑宽度 165 mm



(b) 沟壑宽度 175 mm



(c) 沟壑宽度 185 mm

图 11 六足机器人在 3 种沟壑宽度环境下的仿真过程

Fig. 11 Simulation process of hexapod robot in 3 trench width environments

设置了具有不同高度的梅花桩环境探究六足机器人的越障能力。以足端点云图与可达运动空间的大小分别设置高度差为 -40~40 mm, -45~45 mm,

-50~50 mm 的 3 种梅花桩环境, 并对训练后的运动策略效果进行分析。

图 12 为梅花桩高度差示意图, 其中, h_{ave} 取 200 mm。在平均梅花桩高度的基础上, 限制梅花桩最高高度为 $h_{ave} + \Delta h_{high}$, 梅花桩最低高度为 $h_{ave} - \Delta h_{low}$, 在 $[h_{ave} - \Delta h_{low}, h_{ave} + \Delta h_{high}]$ 的范围内随机采样获得所需的不同高度的梅花桩, $[-\Delta h_{low}, +\Delta h_{high}]$ 即为高度差。

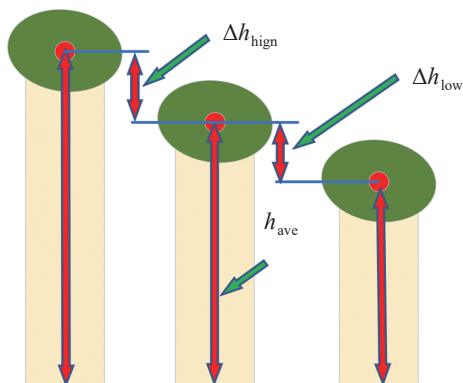


图 12 梅花桩高度差

Fig. 12 Height difference of plum blossom pile

梅花桩高度差为 -45~45 mm, -50~50 mm 时, 机器人机身底部与梅花桩发生碰撞, 足端无法稳定落到梅花桩上, 导致机身失稳。梅花桩高度差为 -40~40 mm, 机器人能够顺利到达目标区域, 如图 13 所示。在 -40~45 mm 间隔 1 mm 均匀采样高度差并进行策略训练。结果显示, 当梅花桩高度差小于 -40~40 mm 时, 六足机器人都能够顺利到达目标区域, 即六足机器人越障最大高度约为 80 mm。根据机器人足端工作空间可知, 机器人足端理论上虽然能到达距离机身较远的位置, 但是梅花桩与机身的干涉也会影响到机器人的运动能力。根据转角范围可知, 髋关节的转角范围决定了机身的离地高度, 本文所采用的转角范围在 0~45° 之间, 没有包含顺时针方向, 因此, 可以通过适当增大机器人髋关节的转角范围来提高机器人的越障能力。

根据图 3 足端点云图可以得到足端理论上 x 、

y 、 z 轴上最大的运动范围分别为0~350 mm、0~200 mm、0~125 mm，但根据最大跨沟宽度实验与越障实验的结果可以发现，六足机器人足端实际上并不能达到理论上的运动范围，导致这一情况的原因是在落足点稀疏的环境下，机器人难以找到同时满足运动学与动力学模型的状态，因此，环境落足点的离散程度在很大程度上将影响机器人的运动能力。

5.4 样机实验

为了进一步评估算法的性能，在 AISpider 机器人上验证了本文算法。该实验平台是由南京大学研制的小型电动六足机器人。机器人采用高稳定性均匀分布的六条腿结构，伺服电机均匀分布

在各条腿的基础关节上。机器人总质量为15 kg，能在7.5 kg的载荷下稳定行走。六足机器人的硬件系统包括18个伺服电机 FEETECH SM85CL、1个电机控制器URT1、1个处理器STM32和1个陀螺仪MPU-6050。

样机实验环境为真实场景下环境III的简化版，如图14所示，展示了六足机器人利用训练后的策略从起点到达目标区域的运动过程。其中，梅花桩直径和高度统一为50 mm，第2、3步为三足步态，第12、13步为四足步态，第25、26步为五足步态。六足机器人能够在保持稳定的前提下，成功从给定的起点运动到目标区域范围之内，进一步验证了自由步态多接触运动规划算法的可行性和有效性。

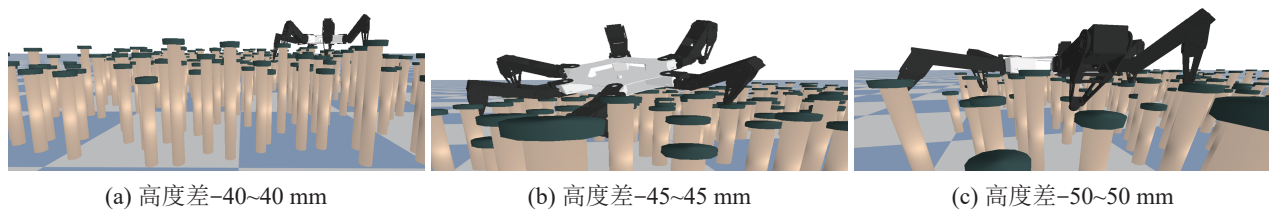


图13 六足机器人在3种高度差环境下的仿真过程

Fig. 13 Simulation process of hexapod robot in 3 height difference environments

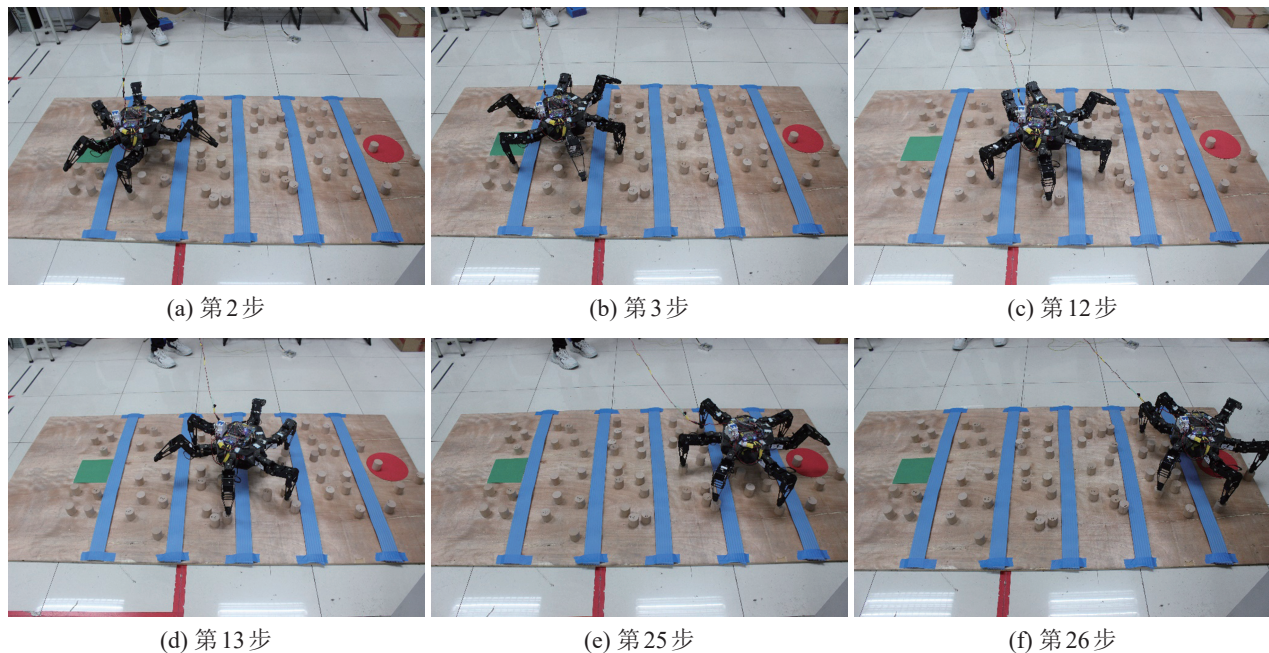


图14 六足机器人在真实环境中的运动过程

Fig. 14 Motion process of hexapod robot in real environment

6 结论

针对六足机器人在离散梅花桩环境下的多接触运动规划问题。本文提出了基于深度强化学习和自由步态的多接触运动规划算法。随机给定起点与目标区域, 该算法在保证六足机器人高效稳定地到达目标区域的前提下, 促使机器人采取不同的步态模式以应对不同的环境。并进一步探究出在自由步态下该六足机器人最大跨沟宽度约为175 mm, 越障最大高度约为80 mm。仿真与真实环境下样机实验都说明了算法的有效性。在未来的工作中, 将针对具有动态障碍物的复杂环境做进一步研究。

参考文献:

- [1] Wu Jianxu, Yang Hui, Li Ruiming, et al. Design and Analysis of a Novel Octopod Platform with a Reconfigurable Trunk[J]. *Mechanism and Machine Theory*, 2021, 156: 104134.
- [2] 石琳, 罗庆生, 韩宝玲, 等. 仿生六足机器人稳定性的仿生实验研究[J]. *系统仿真学报*, 2008, 20(19): 5384-5387.
Shi Lin, Luo Qingsheng, Han Baoling, et al. Research on Biomimetic Experiment of Stability of Hexapod Robot[J]. *Journal of System Simulation*, 2008, 20(19): 5384-5387.
- [3] Jemin Hwangbo, Joonho Lee, Alexey Dosovitskiy, et al. Learning Agile and Dynamic Motor Skills for Legged Robots[J]. *Science Robotics*, 2019, 4(26): 5872.
- [4] Annett Stelzer, Heiko Hirschmüller, Martin Görner. Stereo-vision-based Navigation of a Six-legged Walking Robot in Unknown Rough Terrain[J]. *The International Journal of Robotics Research*, 2012, 31(4): 381-402.
- [5] 李满宏, 张明路, 张建华, 等. 六足机器人关键技术综述[J]. *机械设计*, 2015, 32(10): 1-8.
Li Manhong, Zhang Minglu, Zhang Jianhua, et al. Review on Key Technology of the Hexapod Robot[J]. *Journal of Machine Design*, 2015, 32(10): 1-8.
- [6] Dominik Belter, Jan Wietrzykowski, Piotr Skrzypczyński. Employing Natural Terrain Semantics in Motion Planning for a Multi-legged Robot[J]. *Journal of Intelligent & Robotic Systems*, 2019, 93(3): 723-743.
- [7] Estremera J, Cobano J A, Gonzalez de Santos P. Continuous Free-crab Gaits for Hexapod Robots on a Natural Terrain with Forbidden Zones: An Application to Humanitarian Demining[J]. *Robotics and Autonomous Systems*, 2010, 58(5): 700-711.
- [8] Joaquin Estremera, Pablo Gonzalez de Santos. Free Gaits for Quadruped Robots over Irregular Terrain[J]. *The International Journal of Robotics Research*, 2002, 21(2): 115-130.
- [9] Chen Zhijun, Liu Jimu, Gao Feng. Real-time Gait Planning Method for Six-legged Robots to Optimize the Performances of Terrain Adaptability and Walking Speed[J]. *Mechanism and Machine Theory*, 2022, 168: 104545.
- [10] Chen Xi, Wang Liquan, Ye Xiufen, et al. Prototype Development and Gait Planning of Biologically Inspired Multi-legged Crablike Robot[J]. *Mechatronics*, 2013, 23(4): 429-444.
- [11] Xu Peng, Ding Liang, Wang Zhikai, et al. Contact Sequence Planning for Hexapod Robots in Sparse Foothold Environment Based on Monte-carlo Tree[J]. *IEEE Robotics and Automation Letters*, 2022, 7(2): 826-833.
- [12] Mustafa Suphi Erden, Kemal Leblebicioğlu. Free Gait Generation with Reinforcement Learning for a Six-legged Robot[J]. *Robotics and Autonomous Systems*, 2008, 56(3): 199-212.
- [13] Chen Zhihua, Li Jiehao, Wang Shoukun, et al. Flexible Gait Transition for Six Wheel-legged Robot with Unstructured Terrains[J]. *Robotics and Autonomous Systems*, 2022, 150: 103989.
- [14] Joana Coelho, Fernando Ribeiro, Bruno Dias, et al. Trends in the Control of Hexapod Robots: A Survey[J]. *Robotics*, 2021, 10(3): 100.
- [15] Mnih V, Kavukcuoglu K, Silver D, et al. Human-level Control Through Deep Reinforcement Learning[J]. *Nature*, 2015, 518(7540): 529-533.
- [16] 傅汇乔, 唐开强, 邓归洲, 等. 基于深度强化学习的六足机器人运动规划[J]. *智能科学与技术学报*, 2020, 2(4): 361-371.
Fu Huiqiao, Tang Kaiqiang, Deng Guizhou, et al. Motion Planning for Hexapod Robot Using Deep Reinforcement Learning[J]. *Chinese Journal of Intelligent Science and Technology*, 2020, 2(4): 361-371.
- [17] 董星宇, 唐开强, 傅汇乔, 等. 基于强化学习的六足机器人动态避障研究[J]. *传感器与微系统*, 2022, 41(1): 19-23.
Dong Xingyu, Tang Kaiqiang, Fu Huiqiao, et al. Study on Dynamic Obstacle Avoidance for Hexapod Robot Based on Reinforcement Learning[J]. *Transducer and Microsystem Technologies*, 2022, 41(1): 19-23.
- [18] Fu Huiqiao, Tang Kaiqiang, Li Peng, et al. Deep Reinforcement Learning for Multi-contact Motion Planning of Hexapod Robots[C]//*Proceedings of the Thirtieth International Joint Conference on Artificial*

- Intelligence. California: IJCAI, 2021: 2381-2388.
- [19] Vassilios Tsounis, Mitja Alge, Joonho Lee, et al. DeepGait: Planning and Control of Quadrupedal Gaits Using Deep Reinforcement Learning[J]. IEEE Robotics and Automation Letters, 2020, 5(2): 3699-3706.
- [20] Winkler A W. Optimization-based Motion Planning for Legged Robots[D]. Zurich: ETH Zurich, 2018.
- [21] Haarnoja T, Zhou A, Abbeel P, et al. Soft Actor-critic: Off-policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor[C]//Proceedings of the 35th International Conference on Machine Learning. Chia Laguna Resort, Sardinia, Italy: PMLR, 2018: 1861-1870.
- [22] Jae In Kim, Mineui Hong, Kyungjae Lee, et al. Learning to Walk a Tripod Mobile Robot Using Nonlinear Soft Vibration Actuators with Entropy Adaptive Reinforcement Learning[J]. IEEE Robotics and Automation Letters, 2020, 5(2): 2317-2324.
- [23] 孙辉辉, 胡春鹤, 张军国. 移动机器人运动规划中的深度强化学习方法[J]. 控制与决策, 2021, 36(6): 1281-1292.
Sun Huihui, Hu Chunhe, Zhang Junguo. Deep Reinforcement Learning for Motion Planning of Mobile Robots[J]. Control and Decision, 2021, 36(6): 1281-1292.