DOI:10.11918/202401062

模块化自重构卫星智能变构规划

贾晓冷[1,2],叶 东[1,2],王 博[1,2],孙兆伟[1,2]

(1.哈尔滨工业大学 航天学院,哈尔滨 150001;2.微小型航天器快速设计与智能集群全国重点实验室(哈尔滨工业大学),哈尔滨 150090)

摘 要:为解决航天任务复杂化与传统定构型卫星设计之间的矛盾,航天机构着眼于研究具有灵活构型变化能力的模块化自重构卫星,其中变构规划是一个具有挑战性的研究领域。针对模块化卫星变构问题,以立方体晶格型卫星作为研究对象,基于图论提出了描述卫星拓扑结构的构型矩阵和拓展矩阵。通过对卫星模块运动特点的研究,给出了求解模块运动可达空间的算法。将卫星的变构问题视为序列决策问题,基于深度强化学习理论,将变构过程建模为马尔可夫决策过程,设计了基于演员-评论家(actor-critic)模型的智能变构规划方法,建立多层神经网络以近似演员与评论家函数,通过训练神经网络,逐步改进卫星变构策略性能。仿真实验结果表明,所提出的变构方法对于给定的卫星算例,可以得到逐步改进的卫星变构策略,针对不同模块数的卫星构型具有通用性,同时相比于传统基于启发式搜索的变构方法,在变构步数、计算时间和变构成功率上具有优势,验证了所提出的智能规划方法在未来模块化卫星设计工作中具有潜在的价值。

关键词:模块化自重构卫星;变构规划;深度强化学习;神经网络;演员-评论家模型

中图分类号:V423.4+1 **文献标志码**:A **文章编号**:0367-6234(2025)04-0001-09

Intelligent reconfiguration planning for modular self-reconfigurable satellite

JIA Xiaoleng[1,2], YE Dong[1,2], WANG Bo[1,2], SUN Zhaowei[1,2]

(1. School of Astronautics, Harbin Institute of Technology, Harbin 150001, China; 2. State Key Laboratory of Micro-Spacecraft Rapid Design and Intelligent Cluster (Harbin Institure of Technology), Harbin 150090, China)

Abstract: To address the conflict between the complexity of aerospace tasks and the design of traditional fixed-configuration satellites, aerospace institutions are focusing on the study of modular self-reconfigurable satellites with flexible configuration change capabilities. Among these efforts, configuration planning emerges as a particularly challenging area of research. Aiming at the configuration problem of modular satellites, cube-lattice satellites are taken as the research object, and based on graph theory, a configuration matrix and an extension matrix are proposed to describe the satellite topology. Through the study of the motion characteristics of the satellite module, an algorithm for solving the accessible space of the module motion is given. Considering the satellite configuration problem as a sequential decision-making problem, and based on the theory of deep reinforcement learning, the modification process is modeled as a Markov decision process. An intelligent modification planning method based on the actor-critic model is designed, incorporating a multi-layer neural network to approximate the actor and critic functions. Through training the neural network, the performance of the satellite reconfiguration strategy is progressively improved. The simulation experimental results show that the proposed configuration method yields progressively improved satellite reconfiguration strategies for the given satellite case studies. This approach exhibits generality across different satellite configurations with varying numbers of modules. Additionally, compared with the traditional configuration method based on heuristic search, it has advantages in the number of configuration steps, calculation time and configuration success rate, which validates that the proposed intelligent planning method has potential value in future modular satellite design work.

Keywords: modular self-reconfigurable satellite; planning of reconfiguration; deep reinforcement learning; neural network; actor-critic model

随着航天科技的快速发展,航天任务与卫星工作环境变得越来越复杂,对卫星的多任务执行能力、抗风险能力提出了更高的要求,传统固定结构的卫星越来越难以满足复杂多任务条件下的要求,因此,

收稿日期:2024-01-22;**录用日期**:2024-04-09;**网络首发日期**:2024-12-11
网络首发地址:https://link.cnki.net/urlid/23.1235.t.20241210.1601.002
基金项目:国家自然科学基金(62073102,62203145)
作者简介:贾晓冷(2000—),男,硕士研究生;叶 东(1985—),男,教授,博士生导师;孙兆伟(1963—),男,教授,博士生导师
通信作者:叶 东,yed@hit.edu.cn

具备灵活发射能力、强抗干扰能力、维护方便的模块化卫星成为了航天机构研究的热点[1]。模块化卫星是由数量不等的功能独立、结构类似、外形相同的卫星模块组成的新型卫星[2]。各个模块通过硬件连接,可以通过传感器感知和传递信息。模块化卫星相比于传统的定结构卫星,可以根据不同的任务需求自主灵活改变自身构型,从而增强卫星执行多任务的能力[3]。

模块化卫星的变构规划是该领域研究的一个重要分支。变构规划是对于给定的卫星的初始构型和目标构型,在两个构型之间找到一种满足变构条件的最优路径的过程。卫星的变构能力与所采用的变构策略密切相关,变构策略的选择直接影响到变构任务的完成。由于变构运动需要消耗卫星能源,因此较好的变构策略有助于节省卫星能源和缩短变构时间。研究变构规划对于新型模块化卫星的总体设计和任务规划具有重要的指导意义。

由于太空环境阻尼小、微重力等特点,大多数模块化变构卫星子模块采用立方体晶格型外形。Fei等[4-5]针对晶格型模块结构,提出使用原构型与目标构型的不重合度和重心距离作为驱动函数,驱动模块化卫星完成变构的方法。陈欣[6]针对卫星模块独立运动特点提出分布式运动规划方法,通过每个模块上传感器感知环境,并自主控制自身向目标位置运动。王博等[7]、Song等[8]基于分层处理的思想,设计了模块化卫星分层变构规划方法,该方法由基于模拟退火算法的上层规划和基于搜索算法的下层规划两部分组成,通过训练网络得到变构规划策略。Zhang等[9]将卫星中参与变构的模块分为不同等级,设计了基于结构分级的变构算法,减少了模块拆卸次数。俞港等[10]通过分析模块运动特点,提出一种基于数字激素的模块化卫星自修复算法,适用于动态变化的模块化卫星系统。沈高翔等[11]、Liu等[12]通过分析变构构型的物理性质,提出模块特征点矩阵的概念,通过现构型与目标构型的差异驱动现构型不断向目标构型变构。康国华等[13]通过将变构运动与卫星姿态变化联合考虑,提出一种对称式变构规划算法,但模块规划方式相对简单,规划能力还有待研究。

现今的模块化卫星变构算法在针对较多模块数构型变构时对实时算力要求较高,基于确定性的传统路径规划在针对不同特点构型时变构稳定性不好,同时算法扩展性较差难以满足变构性能提升的要求,因此模块化卫星的变构规划仍然是一个具有挑战性的研究领域。

针对模块化自重构卫星的变构问题,本文提出了一种智能变构规划方法。首先,建立了模块化卫星构型描述方式,根据单模块运动特性分析,提出了求解单模块运动可达空间的方法。其次,对卫星变构任务进行了分析,定义了构型距离函数,通过指派算法明确了单模块运动目标。再次,基于强化学习理论,建立了卫星变构的马尔可夫决策过程,利用深度强化学习演员-评论家(actor-critic)模型建立智能变构规划方法,使得模块化卫星具有智能变构规划能力。通过使用神经网络拟合演员部分的策略函数与评论家部分的值估计函数以及网络训练逐步迭代以提升变构规划能力。最后,在卫星构型算例上进行仿真实验,证明了所提出智能规划方法的有效性、通用性和稳定性。

1 模块化自重构卫星运动学建模

1.1 卫星构型数学描述

模块化自重构卫星由功能不同、外形相同的模块构成,卫星构型由具有一定连接关系的模块集合组成,模块之间不同的连接方式形成了模块化卫星不同的构型,如图1所示,展示了模块化卫星整体构型与各分模块。

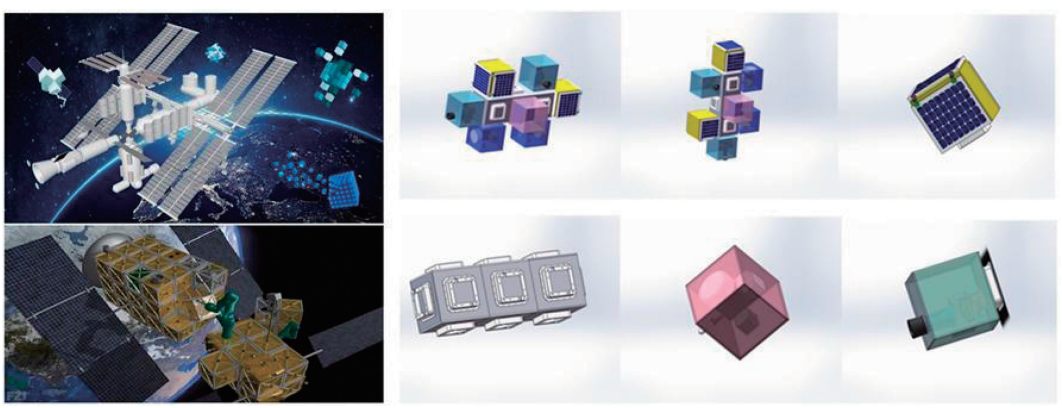


图1 模块化卫星整体与分模块

Fig. 1 Schematic diagram of modular self-reconfigurable satellite module

本文基于图论[14]理论,建立了描述卫星构型的构型矩阵与扩展矩阵。基于晶格型模块化卫星理论,将卫星模块简化为大小形状相同,长、宽、高均为单位长度的立方体模块,并使用模块中心点坐标来表示模块位置。在卫星坐标空间中建立三维坐标系,为每个卫星模块编号,第 i 个模块位置由其体心坐标 (x_i,y_i,z_i) 表示。按照编号顺序,将体心坐标按行排放,得到卫星的构型矩阵 $\boldsymbol{V}$。

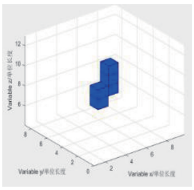
构型描述不仅应包括模块的绝对位置信息,还应包含模块之间的相对位置和连接信息。定义基于卫星构型矩阵的扩展矩阵,该矩阵与构型矩阵相对应,第 i 行对应第 i 个模块,而纵向分别对应第 i 个模块在 6 个面方向上的连接情况,从左到右分别对应在 x_+,x_-,y_+,y_-,z_+,z_- 6 个方向上的相邻模块,如果第 i 个模块在该面方向上没有相邻模块则置为 0。

构型矩阵和扩展矩阵共同构成了模块化自重构卫星构型的数学描述形式。表 1 给出了一种卫星构型的三维图示、构型矩阵与扩展矩阵。

本文中讨论的模块化自重构卫星构型矩阵图必须为连通图。如果构型图不连通,意味着卫星中有部分模块不与整个卫星主体相连,不同分支的模块之间会随机产生姿态相对运动,严重影响正常的工作。

表 1　构型的三维图示、构型矩阵和扩展矩阵

Tab. 1　3D diagram, configuration matrix and extension matrix of the configuration

三维图示	构型矩阵	扩展矩阵
	$\begin{bmatrix}6&6&6\\6&6&7\\7&6&7\\7&6&8\\7&6&9\end{bmatrix}$	$\begin{bmatrix}0&0&0&0&2&0\\3&0&0&0&0&1\\0&2&0&0&4&0\\0&0&0&0&5&3\\0&0&0&0&0&4\end{bmatrix}$

1.2　模块运动学模型

模块运动学模型研究对象为晶格型立方体卫星模块的单步运动过程,由于太空环境中失重、阻尼小等特点,本文采用标准的旋转立方模型(rotating cube model, PCM)来描述单模块变构时的运动过程。其三要素为:旋转基础轴(与相邻模块的公共边)、旋转方向和旋转角度。确定了旋转三要素就可以唯一确定一种旋转运动。

旋转方向定义为从坐标轴正向看,逆时针旋转角度为正。旋转角度划分为 90°、180°、-90°、-180°。由于旋转运动的特性,旋转运动始终在一个平面内进行。模块在旋转过程中,除了初始位置和目标位置,模块在旋转面投影的对角线还会扫过一片额外的区域,为了防止模块碰撞,对角线扫过的区域不能存在其他模块。同时,在模块的一个面接触其他模块前模块旋转运动不会停止,因此旋转角度也与模块的相对位置有关,具体旋转过程见图 2。

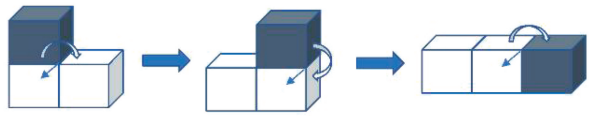


图 2　模块旋转运动

Fig. 2　Module rotational motion

在实际工程应用中,待移动模块转动时会受到相邻模块的运动约束。本文定义模块运动可达空间概念,来求解待移动模块在周围存在其他模块位置约束的情况下,通过单次旋转运动能够到达的位置。模块运动可达空间的求解是变构过程模块进行运动动作选择的基础。算法 1 展示了求解模块运动可达空间的伪代码。

算法 1　求解模块运动可达空间

输入:　卫星构型矩阵 $\boldsymbol{V}$、模块位置 $\boldsymbol{v}_i$

输出:　模块运动可达空间集合 G

```
集合 G 初始化为空集
遍历 V 求 v_i 相邻模块位置存入 M
For all i ∈ M do
    计算旋转 π/2 与 π 时可达位置 v_i*
    if 可达位置满足旋转运动条件 then
        将 v_i* 加入 G
    End if
End for
Return G
```

算法通过遍历的方式,在 3 个平面内逐一判断待移动模块满足的运动约束条件,并将满足运动约束的可到达位置加入模块可运动空间集合中,从而求解出在当前构型矩阵环境下待移动模块通过一次旋转运动能够到达的位置集合。

2　模块化卫星变构任务分析

模块化卫星变构规划可以视为在原构型与目标构型之间寻找一种符合运动约束要求的变换方式。其数学描述为:给定卫星原构型 $\boldsymbol{V}_1$ 和目标构型 $\boldsymbol{V}_t$,在符合运动约束条件下,通过模块运动将原构型 $\boldsymbol{V}_1$,经过一系列中间构型 $(\boldsymbol{V}_1,\boldsymbol{V}_2,\boldsymbol{V}_3,\cdots,\boldsymbol{V}_t)$ 最终

变为目标构型 $\boldsymbol{V}_t$。

为了能够量化两构型之间的差异程度,定义基于构型模块坐标曼哈顿距离的构型距离函数。定义构型图中的两个节点(x_i,y_i,z_i)、(x_j,y_j,z_j)的曼哈顿距离为

$$d = |x_i - x_j| + |y_i - y_j| + |z_i - z_j| \tag{1}$$

定义两构型之间的曼哈顿距离为对应的两构型未重合部分模块的曼哈顿距离之和为

$$D(\boldsymbol{V}_a,\boldsymbol{V}_b) = \sum_{i=1}^{n}\sum_{j=1}^{n} d(\boldsymbol{v}_{ai},\boldsymbol{v}_{bj}) \tag{2}$$

式中:$\boldsymbol{v}_{ai}$ 为构型 $\boldsymbol{V}_a$ 中的未重合模块坐标,$\boldsymbol{v}_{bj}$ 为构型 $\boldsymbol{V}_b$ 中的未重合模块坐标。

在实际太空环境中,模块变构过程消耗能源和时间,变构规划需要在完成变构任务的前提下,寻求较少的模块运动步数和更快的变构时间。卫星变构总时间与模块变构总步数相关,本文使用模块运动总步数作为衡量变构规划性能好坏的指标。记单模块做一次旋转运动的运动步数为 1,卫星变构运动总步数为所有模块运动步数之和,定义如下:

$$S(\boldsymbol{V}_i,\boldsymbol{V}_t) = \sum_{k=1}^{n} S_k \tag{3}$$

式中:S_k 为在变构过程中模块 k 所运动的总步数,n 为卫星总模块数。

将模块化卫星变构任务抽象为序列决策问题。自重构卫星模块数较多时,直接规划求解从原构型变构到目标构型的一系列中间构型较为困难。本文将整体构型的变构规划过程转换为每个未重合模块运动规划的叠加,通过任务分配增加变构规划的可操作性,通过一系列单模块运动完成整体变构任务。

图 3 展示了变构任务分析的主要流程,通过将给定的原构型 $\boldsymbol{V}_1$ 与目标构型 $\boldsymbol{V}_t$ 作比较,将原构型中的未重合部分称为待移动模块,目标构型中的未重合部分称为目标位置。对于可以进行移动的待移动模块和可以到达的目标位置,基于指派问题的解决思路,利用待移动模块与目标位置的曼哈顿距离和作为指派问题优化的目标,使用匈牙利算法[15-17]进行优化,选择使得总曼哈顿距离最短的对应关系作为最终变构规划的模块 - 位置对应关系。指派问题求解伪代码见算法 2。通过指派算法求得的模块 - 位置对应关系确定了每个模块在具体变构规划过程中的规划目标,根据规划目标设计单模块变构规划方法。

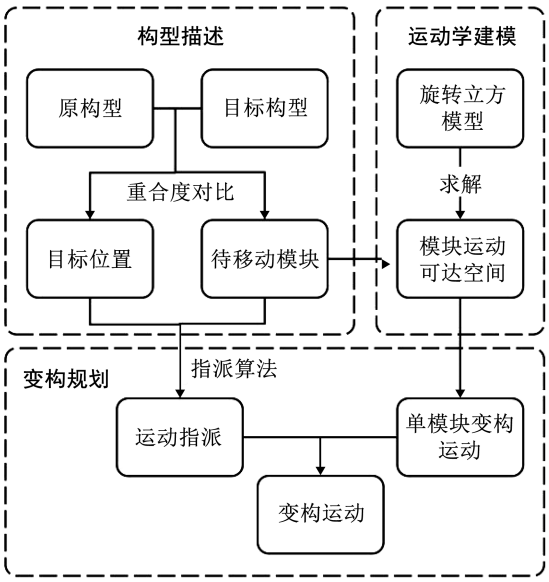


图 3　变构任务分析

Fig. 3　Reconfiguration task analysis

算法 2 变构运动指派算法

输入:去除重合模块的原构型与目标构型矩阵 $\boldsymbol{V}_a$、$\boldsymbol{V}_b$

输出:模块 - 位置对应关系 e_{ij}

1:　运动可达空间集合非空的待移动模块与目标位置分别加入集合 $\boldsymbol{V}_1$ 与 $\boldsymbol{V}_2$

2:　判断集合 $\boldsymbol{V}_1$ 节点数是否与集合 $\boldsymbol{V}_2$ 节点数相同

3:　If 相同,do:

　　使用 $\boldsymbol{V}_1$ 与 $\boldsymbol{V}_2$ 计算曼哈顿距离矩阵 $\boldsymbol{D}$

4:　If 不同,do:

　　1)在节点数少的集合中使用虚拟节点法增加节点数,使得两集合节点数相同

　　2)将虚拟节点与另一集合其他节点的距离设置为一个较大数

　　3)使用节点数相同的 $\boldsymbol{V}_1$ 与 $\boldsymbol{V}_2$ 计算曼哈顿距离矩阵 $\boldsymbol{D}$

5:　使用匈牙利算法对曼哈顿距离矩阵 $\boldsymbol{D}$ 进行处理得到对应关系 e_{ij}

6:　Return e_{ij}

模块化卫星变构规划方法的难度与卫星模块数量、卫星构型复杂度有紧密的联系。随着卫星模块数量的增多以及航天任务的随机化和复杂化,传统的重构规划方法大多底层采用确定性策略,无法根据环境和随机情况进行灵活规划。同时传统算法产生局部最优解的特性,会导致卫星变构过程无法达到最佳,从而消耗更多的能源和时间。

基于深度强化学习的智能规划方法能够有效解决传统算法的局限性问题,在卫星模块数增加的情况下,即使使用较小算力进行规划,也可以指导卫星模块进行变构运动。此外,深度强化学习能够针对

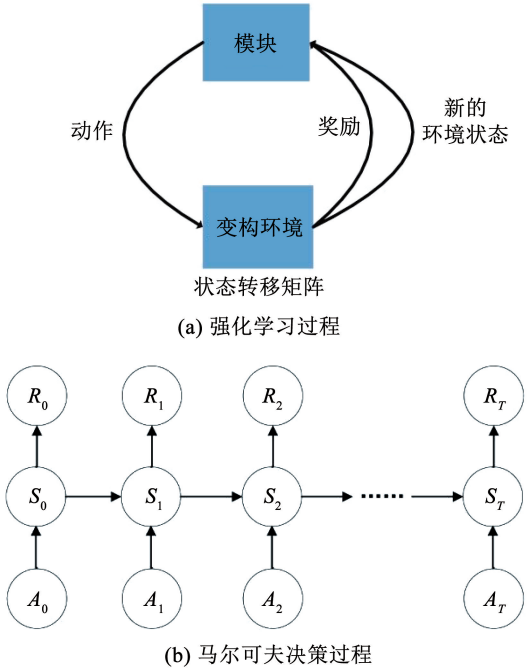
随机构型提供对应的规划方法,解决了传统方法无法处理随机变构规划的局限性。

3　基于深度强化学习的智能变构规划

受深度强化学习理论的启发,本文将自重构卫星模块看做在变构环境中具有感知、规划能力的智能体(agent),通过模块与变构环境的不断交互,训练迭代使得模块习得智能变构规划策略。

3.1　强化学习建模

强化学习(reinforcement learning,RL)是一种通过与环境交互进行学习的机器学习方式,目标是最大化智能体在和环境的交互过程中所能获得的累积奖励。在模块化卫星变构问题中,模块现阶段的运动与未来阶段的运动在概率上相互独立,因此智能体与环境的交互满足马尔可夫性,模块变构过程可以定义为马尔可夫决策过程(Markov decision processes,MDP)。马尔可夫决策过程可以用四元组 $M=(S,A,P,R)$ 表示。其中:S 为状态集合,A 为动作集合,P 为状态转移函数,R 为系统即时奖励,四元组是模块与环境交互的重要信息。模块变构的强化学习与马尔科夫决策过程建模见图 4。



(a) 强化学习过程

(b) 马尔可夫决策过程

图 4　模块强化学习建模过程

Fig. 4　Module reinforcement learning modeling process

模块与环境在 t 时刻交互过程为:模块在状态 s_t 下选择执行动作 a_t,并以 $p(s_{t+1}|s_t,a_t)$ 的概率转移到下一个状态 s_{t+1},同时模块获得奖励 r_t,状态、动作和奖励三者交替进行共同组成一组序列:$s_0,a_0,r_0,s_1,a_1,r_1,\cdots$。

在模块化卫星自重构过程中,卫星构型矩阵对应马尔可夫决策过程中的状态空间 S。模块运动方式共计有 48 种,其总和是动作空间 A,定义为 $a_t\in A$。在模块上一时刻的状态与动作确定的情况下,下一时刻的状态也就唯一确定,所以在本文中状态转移函数为 1。

合理的奖励设置可以加快训练的收敛速度。本文综合多种因素,最终以模块运动最终到达位置与目标位置重合与否作为评判标准来设置奖励值,两位置重合则环境反馈奖励为 0,其余奖励反馈为 -1。为使模块运动摆脱局部最优的约束,根据奖励对长期收益进行最大化,使得模块在规划学习过程中更加关注于长期收益而非短期奖励。模块在 t 时刻的回报 G_t 为

$$G_t=R_{t+1}+\gamma R_{t+2}+\gamma^2R_{t+3}+\cdots=\sum_{k=0}^{\infty}\gamma^kR_{t+k+1} \tag{4}$$

式中 γ 为折扣率,表示某一时刻奖励占回报的权重,γ 越大表明智能体更重视未来的收益,越小表示智能体更重视当前奖励。

在强化学习过程中,智能体通过价值函数来评价状态和动作的好坏。根据回报定义模块的状态价值函数表示形式为 $v_\pi(s)$,其数学形式为

$$v_\pi(s)=E_\pi[G_t|S_t=s] \tag{5}$$

在模块变构过程中,每一时刻,模块会根据策略函数 $\pi(a|s)$ 选择出一种动作执行。在训练过程中,策略函数会随着训练进行不断迭代更新以获得更贴近真实条件的最优策略函数。

状态 - 动作价值函数 $q(s,a)$ 对模块在某一状态下的变构动作进行评估,状态价值函数表示在策略 $\pi(a|s)$ 下回报的期望值。其数学形式为

$$q(s,a)=E_\pi[G_t|S_t=s,A_t=a] \tag{6}$$

式(5)、(6)联立,可推导出模块状态价值函数与状态 - 动作价值函数的关系可表示为

$$v_\pi(s)=\sum_a\pi(a|s)\cdot q(s,a) \tag{7}$$

3.2　演员 - 评论家模型

本文受演员 - 评论家(actor-critic)模型的启发,提出一种智能变构规划方法来处理卫星变构问题。该模型由演员部分(actor)和评论家部分(critic)两

部分构成。通过这两个部分的相互协作和共同训练,迭代学习得到最优策略,模型训练基本结构见图5。

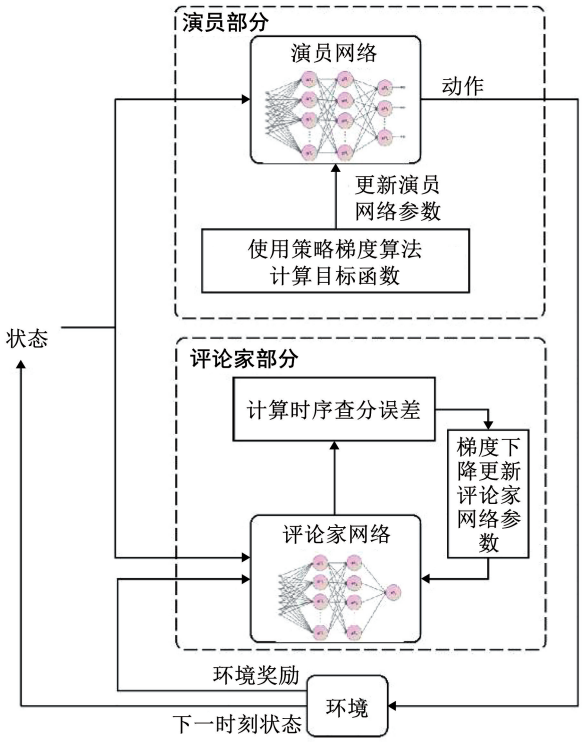


图5 演员 – 评论家模型

Fig. 5 Actor-critic model

3.2.1 评论家部分(critic)

评论家部分为对状态动作的值估计函数,用以评估模块演员部分执行动作的好坏。评论家部分接受模块感知到的状态 s_t 与动作 a_t 作为输入,并输出当前状态下执行动作好坏的的估计值 $q_t(s_t,a_t)$。本文通过神经网络 $q(s,a;w)$ 拟合评论家函数 $q(s,a)$,其中 w 为评论家网络参数。评论家网络使用时序差分算法进行更新。

3.2.2 演员部分(actor)

演员部分为模块的变构策略函数,接受模块位置与卫星构型矩阵作为状态信息输入,并输出执行所有动作的概率分布值。本文通过神经网络 $\pi(a\mid s;\theta)$ 拟合策略函数 $\pi(a\mid s)$,其中 θ 为演员网络参数。卫星变构问题动作空间属于离散动作空间,演员部分迭代使用策略梯度法来进行更新。

每个模块中都封装有演员部分与评论家部分,对于每一个模块的训练过程伪代码见算法3。

算法3 模块深度强化学习演员 – 评论家模型训练过程

1: 由模块传感器观测到现有状态 s_t,通过模块演员部分策略网络 $\pi(a\mid s;\theta)$ 求出可选动作概率分布,从中随机抽样得到 a_t

2: 模块实际执行 a_t,模块通过传感器获得环境反馈 s_{t+1} 与 r_t

3: 模块以 s_{t+1} 为输入,通过策略网络 $\pi_t(a_t\mid s_t;\theta)$ 求得动作概率分布,从中随机抽样获得 a_{t+1}

4: 通过模块评论家部分值函数估计 $q(s,a;w)$ 计算出 $q_t(s_t,a_t;w)$ 与 $q_{t+1}(s_{t+1},a_{t+1};w)$

5: 使用值估计计算时序误差

6: 求价值网络梯度

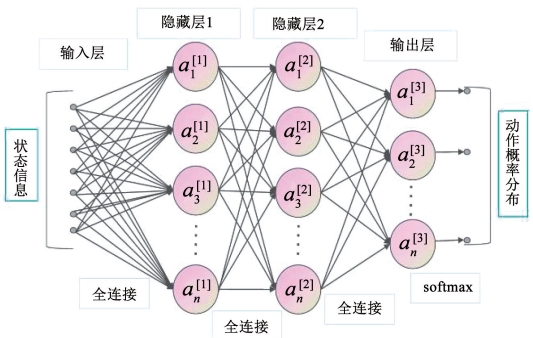
7: 通过时序差分算法更新价值网络参数

8: 求策略网络梯度

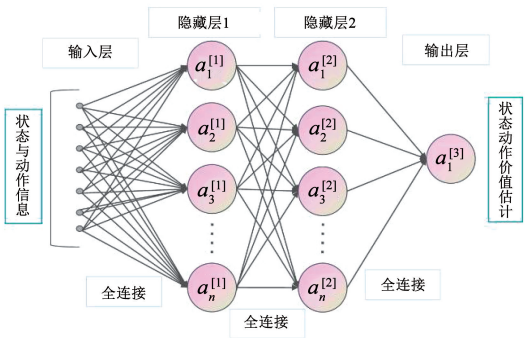
9: 运用梯度上升,更新策略网络参数

10: 判断是否满足算法收敛性条件,满足跳出循环;否则进行下一轮循环

actor 神经网络和 critic 神经网络结构见图 6 。



(a) actor神经网络结构



(b) critic神经网络结构

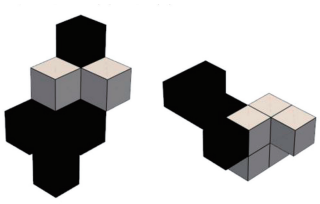
图6 神经网络结构

Fig. 6 Neural network structures

4 仿真实验与分析

4.1 实验环境与算例

本文针对卫星变构方法进行了仿真实验,以验证所提出方法的有效性。本实验选取两个10模块卫星构型作为算例,如图7所示展示了原构型与目标构型的三维示意图与构型矩阵。



(a) 卫星原构型与目标构型三维图

$$\boldsymbol{V}_{\text{原构型}}=\begin{bmatrix}0&0&-1\\1&0&-1\\0&1&-1\\-1&0&-1\\0&0&0\\-1&0&0\\-1&1&0\\-1&0&1\\0&0&-2\\0&0&-3\end{bmatrix},\ \boldsymbol{V}_{\text{目标构型}}=\begin{bmatrix}0&0&0\\1&0&0\\-1&0&0\\-1&1&0\\0&-1&0\\0&-2&0\\0&-3&0\\-1&-3&0\\0&0&-1\\-1&0&-1\end{bmatrix}$$

(b) 卫星原构型与目标构型构型矩阵

图 7　卫星原构型与目标构型

Fig. 7　Satellite original configuration and target configuration

去除重合模块后,剩余 5 个未重合模块。利用算法 2 将未重合模块做任务指派,得到总距离最小的模块 – 位置对应关系如 $\boldsymbol{V}_1$ 与 $\boldsymbol{V}_2$ 所示,两矩阵对应行表示模块的原位置与目标位置分别为:

$$\boldsymbol{V}_1=\begin{bmatrix}1&0&-1\\0&1&-1\\-1&0&1\\0&0&-2\\0&0&-3\end{bmatrix}\tag{8}$$

$$\boldsymbol{V}_2=\begin{bmatrix}1&0&0\\-1&-3&0\\0&-3&0\\0&-2&0\\0&-1&0\end{bmatrix}\tag{9}$$

4.2　仿真结果与分析

搭建多层神经网络,利用智能变构规划方法进行变构训练,神经网络优化器采用广泛适用的自适应矩估计(adaptive moment estimation, Adam)优化器。

针对由 10 模块构成的初始构型和目标构型进行变构训练,训练 1 000 个数据集,每个数据集中模块最多运动 100 步,训练折扣率为 0.99。图 8 显示了智能体训练过程中平均回报随训练次数的变化曲线。从图 8 中可以看出,在前 200 次训练,智能体策略主要以探索为主,回报增长的很慢;当训练次数在 200 ~ 400 之间时,智能体所获得的回报迅速增长;当训练次数在 400 ~ 600 之间时,智能体的回报曲线出现震荡现象,并逐渐收敛到最大值;在训练次数达到 800 以后,智能体的回报基本收敛,达到最大。

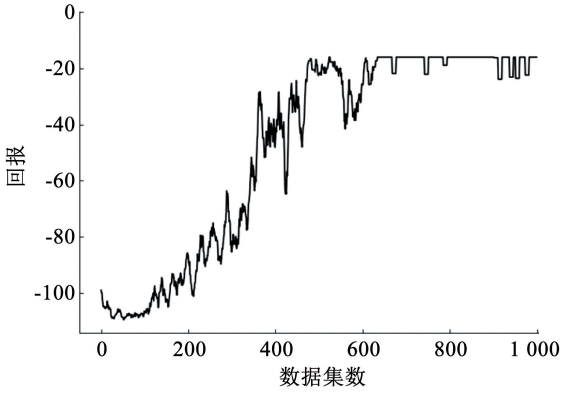
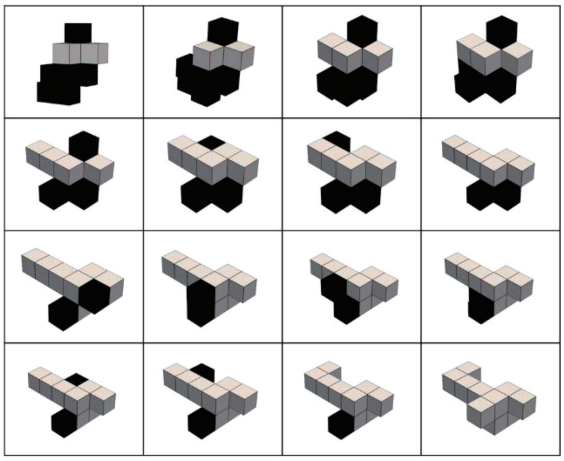


图 8　演员 – 评论家框架训练效果

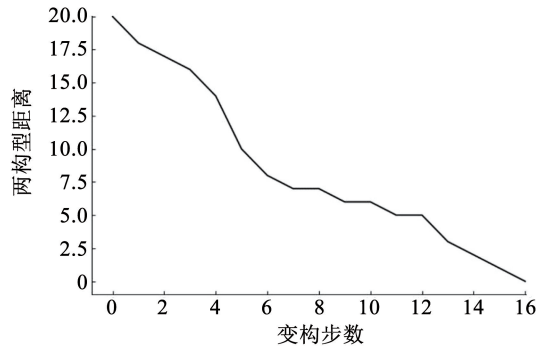
Fig. 8　Actor-critic model training result

在训练过程中为了避免小样本序列与偏执样本序列对变构策略训练效果的影响,可以利用多个构型同时进行变构序列采样,建立样本池通过被采样次数加权抽取样本进行强化学习训练。

利用训练所得演员网络模拟智能体的变构最优策略函数,对算例构型进行变构规划,变构结果见图 9。



(a) 模块变构过程示意



(b) 模块变构过程构型距离函数变化

图 9　模块变构过程

Fig. 9　Schematic diagram of module reconfiguration process

图9(a)中的每张图片展示了一次单模块的变构运动。从变构第1步开始,随着变构步数的增加,黑色未重合模块逐步变白色重合模块,当经过16步变构后,完成变构任务,将原构型变构成为目标构型。图9(b)展示了在变构过程中,当前构型与目标构型曼哈顿距离与模块移动总步数的关系。表2展示了变构过程中构型各模块的距离变构序列。

表2　构型变构过程中各模块运动序列

Tab. 2　Motion sequence of each module during the reconfiguration process

模块原位置	运动序列	模块目标位置
(0,0,-3)	(0,0,-3)→(0,-1,-2)→(0,-1,-1)→(0,-1,0)	(0,-1,0)
(0,0,-2)	(0,0,-2)→(0,-1,-1)→(0,-2,0)	(0,-2,0)
(-1,0,1)	(-1,0,1)→(-1,-1,0)→(-1,-2,0)→(0,-3,0)	(0,-3,0)
(0,1,-1)	(0,1,-1)→(0,1,0)→(1,0,0)→(1,-1,0)→(0,-1,-1)→(-1,-1,0)→(-1,-2,0)→(-1,-3,0)	(-1,-3,0)
(1,0,-1)	(-1,0,-1)→(1,0,0)	(1,0,0)

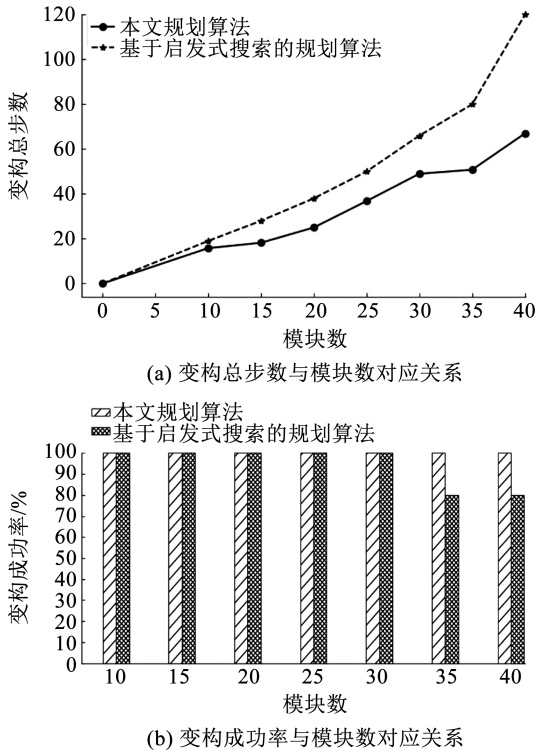
为验证本文提出的智能规划方法针对不同模块数构型的泛用性,对于模块数更多的构型,利用在10模块构型下训练得到的变构策略网络,根据模块数改变网络输入神经元个数,然后在对应环境下进行训练得到针对不同模块数构型的变构策略函数。对应不同的模块数随机选择5种构型进行变构实验,记录变构步数,实验结果见表3。从表3中可以看出,仿真结果清楚展示了本文所提出的变构方法适用于更多模块数的构型,表明该变构方法具有通用性。

表3　不同模块数构型变构步数统计结果

Tab. 3　Statistical results of reconfiguration steps for different module digital configurations

模块数	变构步数					平均变构步数	成功率/%
	构型1	构型2	构型3	构型4	构型5		
10	16	13	26	16	8	15.8	100
15	18	16	9	23	25	18.2	100
20	26	24	19	34	22	25.0	100
25	39	42	32	27	44	36.8	100
30	48	52	44	39	62	49.0	100
35	40	51	46	67	50	50.8	100
40	58	67	62	78	70	67.0	100

作为比较,将本文所提出的变构规划方法与模块化卫星变构问题上有较为成熟应用的A*搜索算法[13]作对比。对照组A*搜索算法估值函数设置为初始位置到当前位置的变构步数与当前位置到目标位置的曼哈顿距离之和。仿真结果对比效果见图10。从图10(a)可以看出,在卫星构型模块数小于10时,两种规划算法规划总步数接近,当卫星构型模块数在10~30之间时,基于启发式搜索的传统规划方法仿真结果与本文方法差距较小,当卫星模块数大于30时,本文算法规划总步数远小于传统算法。从图10(b)可以看出,本文所提出的变构规划算法在模块数大于30时依旧能保证变构成功率为100%,而传统算法在此情况下会出现规划失败的情况。



(a) 变构总步数与模块数对应关系

(b) 变构成功率与模块数对应关系

图10　本文规划算法与启发式搜索规划算法对比

Fig. 10　Comparison between proposed planning algorithm and heuristic search planning algorithms

为验证本文所提出的方法对比启发式搜索方法在时间性能上的提升,分别针对不同模块数构型,随机生成5组初始构型,利用两种方法对构型进行变构仿真,对两种方法的变构实时计算时间取平均值,实验结果见表4。

表 4　不同模块数构型变构实时计算时间统计结果

Tab. 4　Statistical results of reconfiguration real-time computing time for different module digital configurations

模块数	启发式搜索方法计算时间/s	本文规划方法计算时间/s	倍数
10	0.65	0.007	92.86
15	0.96	0.010	96.00
20	3.89	0.062	62.71
25	12.78	0.122	104.78
30	71.48	0.872	81.97
35	193.87	1.036	187.13
40	400.47	1.519	263.64

从表 4 中可以看出,对于不同模块数构型,本文方法的实时计算时间显著小于启发式搜索方法,两者计算时间相差两个数量级,而且随着模块数增多,本文方法实时计算时间少的优势越来越明显。同时本文方法部署在卫星上时只需存储训练所得的神经网络参数,相比启发式搜索算法的空间复杂度也较小。

5　结　论

1)本文训练得到的智能变构规划方法对于给定的 10 模块卫星构型,能够在 16 步有效实现卫星构型的变构规划,证明了该方法的有效性。

2)针对不同模块数卫星构型使用本文方法进行变构训练,得到了针对不同模块数构型的变构策略,验证了本文方法在不同模块数构型情况下的通用性。

3)通过在不同模块数构型环境下与基于启发式搜索的传统规划方法的比较,本文提出的规划方法能够在较少步数与计算时间内完成变构规划,同时保证了针对 30 模块以上构型 100% 的变构成功率,证明了所提出方法的高效性与稳定性。

参考文献

[1]王恩美, 邬树楠, 王晓明,等. 大型卫星太阳能帆板的分布式振动控制[J]. 航空学报, 2018, 39(1): 221479
WANG Enmei, WU Shunan, WANG Xiaoming, et al. Distributed vibration control for large satellite solar panels [J]. Acta Aeronauticaet Astronautica Sinica, 2018, 39(1): 221479. DOI:10.7527/S1000-6893.2017.21479

[2]EDWARD FALKENHAYN J Jr. Multimission modular spacecraft (MMS) [C]//Space Brograms and Technologies Conference. Reston: AIAA, 1988: 3513. DOI: 10.2514/6.1988-3513

[3]李新洪, 张永乐, 姜南. 模块化航天器应用需求及应用体系[J]. 装备学院学报, 2014, 25(4): 70
LI Xinhong, ZHANG Yongle, JIANG Nan. Modular spacecraft application requirements and application system [J]. Journal of Equipment Academy, 2014, 25(4): 70. DOI: 10.3783/j.issn.2095-3828.2014.04.016

[4]FEI Yanqiong, WANG Chengyuan. Self-repairing algorithm of lattice-type self-reconfigurable modular robots[J]. Journal of Intelligent & Robotic Systems, 2014, 75(2): 193. DOI:10.1007/s10846-013-9885-8

[5]FEI Yanqiong, ZHU Yueliang, XIA Ping. Analysis on self-morphing process of self-reconfigurable modular robot[J]. International Journal of Advanced Robotic Systems, 2009, 6(3): 23. DOI: 10.5772/7232

[6]陈欣. 自重构卫星的重构路径规划算法研究[D]. 哈尔滨工业大学, 2018
CHEN Xin. Research on reconfiguration path planning algorithm for self-reconfigurable satellites [D]. Harbin: Harbin Institute of Technology, 2018

[7]王博, 叶东, 孙兆伟,等. 模块化可重构卫星在轨自重构的分层规划[J]. 航空学报,2019,40(9): 322912
WANG Bo, YE Dong, SUN Zhaowei, et al. Hierarchical planning for on-orbit self-reconfiguration of modular reconfigurable satellites[J]. Acta Aeronautica et Astronautica Sinica, 2019, 40(9): 322912. DOI:10.7527/S1000-6893.2019.22912

[8]SONG Qiliang, YE Dong, SUN Zhaowei, et al. Autonomous reconfiguration of homogeneous pivoting cube modular satellite by deep reinforcement learning[J]. Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering, 2021, 235(10): 1777. DOI: 10.1177/0959651820956738

[9]ZHANG Yifei, HUANG Panfeng, ZHANG Yizhai, et al. Reconfiguration planning for heterogeneous cellular space robot [C]//2017 18th International Conference on Advanced Robotics (ICAR). Hong Kong: IEEE, 2017: 185. DOI: 10.1109/ICAR.2017.8023516

[10]俞港, 付庄, 管恩广. 晶格式模块机器人系统自修复研究与仿真[J]. 计算机仿真, 2016, 33(11): 384
YU Gang, FU Zhuang, GUAN Enguang. Research and simulation of lattice modular robot self-repairing[J]. Computer Simulation, 2016, 33(11): 384. DOI: 10.3969/j.issn.1006-9348.2016.11.084

[11]沈高翔, 陈萌, 王一泽, 等. 自重构机器人自变形算法研究[J]. 机械设计, 2019, 36(1): 5
SHEN Gaoxiang, CHEN Meng, WANG Yize, et al. Study on self-morphing algorithm for self-reconfiguring robot [J]. Journal of Machine Design, 2019, 36(1): 5. DOI:10.13841/j.cnki.jxsj.2019.01.002

[12]LIU Chao, WHITZER M, YIM M. A distributed reconfiguration planning algorithm for modular robots [J]. IEEE Robotics and Automation Letters, 2019, 4(4): 4231. DOI: 10.1109/LRA.2019, 2930432

[13]康国华, 刘奇弦, 吴佳奇, 等. 微纳聚合体卫星的对称式重构规划算法[J]. 宇航学报, 2020, 41(7): 937
KANG Guohua, LIU Qixian, WU Jiaqi, et al. Symmetric reconfiguration planning algorithm of combination body of micro-nano satellites[J]. Journal of Astronautics, 2020, 41 (7): 937. DOI: 10.3873/j.issn.1000-1328.2020.07.012

[14]殷剑宏, 吴开亚. 图论及其算法[M]. 合肥:中国科学技术大学出版社, 2003
YIN Jianhong, WU Kaiya. Graph theory and its algorithm[M]. Hefei:University of Science and Technology of China Press, 2003

[15]ZHU Haibin, ZHOU Mengchu. Efficient role transfer based on Kuhn-munkres algorithm[J]. IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, 2012, 42(2): 491. DOI: 10.1109/TSMCA.2011.2159587

[16]ZHU Haibin, ZHOU Mengchu, ALKINS R. Group role assignment via a Kuhn-munkres algorithm-based solution[J]. IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, 2012, 42(3): 739. DOI: 10.1109/TSMCA.2011.2170414

[17]PAMECHA A, CHIRIKJIAN G. A useful metric for modular robot motion planning[C]//Proceedings of IEEE International Conference on Robotics and Automation. Minneapolis: IEEE, 2002: 442. DOI: 10.1109/ROBOT.1996.503816

(编辑　张　红)