

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

基于强化学习的空间引力波探测望远镜系统外杂光抑制研究

张梓扬, 常军, 黄一帆, 陈钦芳, 吴郁楠

引用本文:

张梓扬, 常军, 黄一帆, 等. 基于强化学习的空间引力波探测望远镜系统外杂光抑制研究[J]. *光电工程*, 2024, 51(2): 230210.

Zhang Z Y, Chang J, Huang Y F, et al. Reinforcement learning-based stray light suppression study for space-based gravitational wave detection telescope system[J]. *Opto-Electron Eng*, 2024, 51(2): 230210.

<https://doi.org/10.12086/oe.2024.230210>

收稿日期: 2023-08-30; 修改日期: 2023-12-08; 录用日期: 2023-12-11

相关论文

空间引力波探测望远镜研究进展

王小勇, 白绍竣, 张倩, 林栩凌, 李洋, 夏晨晖

光电工程 2023, 50(11): 230219 doi: [10.12086/oe.2023.230219](https://doi.org/10.12086/oe.2023.230219)

空间引力波探测系统中超光滑光学元件表面散射特性分析

张耘豪, 钟哲强, 张彬

光电工程 2023, 50(11): 230222 doi: [10.12086/oe.2023.230222](https://doi.org/10.12086/oe.2023.230222)

空间引力波星载望远镜测试与评估技术研究进展

张兰强, 曾意, 吴小虎, 杨金生, 阮晓莉, 辛强, 顾乃庭, 饶长辉

光电工程 2024, 51(2): 240027 doi: [10.12086/oe.2024.240027](https://doi.org/10.12086/oe.2024.240027)

星载望远镜消光材料积分散射特性测试研究(英文)

刘巍, 李朝辉, 毛振, 赵建科, 朱辉, 魏紫薇, 刘勇, 尹云飞

光电工程 2024, 51(2): 230206 doi: [10.12086/oe.2024.230206](https://doi.org/10.12086/oe.2024.230206)

更多相关论文见光电期刊集群网站 

 **光电工程**
Opto-Electronic Engineering

<http://cn.ojournal.org/oe>



 OE_Journal



Website

DOI: 10.12086/oe.2024.230210

基于强化学习的空间引力波探测望远镜系统外杂光抑制研究

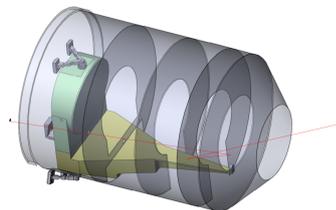
张梓扬¹, 常军^{1*}, 黄一帆^{1*}, 陈钦芳², 吴郁楠¹¹北京理工大学光电学院, 北京 100081;²中国科学院西安光学精密机械研究所, 陕西 西安 710119

摘要: 引力波望远镜其收集的空间目标光信号能量远小于杂散光能量, 为了保证引力波望远镜的正常工作, 需要保证较好的杂光抑制效果。又由于散射光线的不确定性, 光机系统本身的复杂性, 杂光抑制方案的确定往往需要复杂的数学公式计算以及丰富的经验与充足时间进行仿真迭代。本文提出了一种基于强化学习的杂光抑制策略, 针对空间引力波探测望远镜系统中的环境杂光问题, 采用蒙特卡洛光线追迹方法进行分析 and 处理。通过制定有效的杂光抑制措施, 实现了对该系统中杂光的有效控制。仿真结果验证了该方法在空间引力波探测望远镜系统中杂光抑制方面的优异性能, 展示了其具有良好杂光抑制效果的潜力。这一研究为解决空间引力波探测和其他高精度光学系统中的杂光问题提供了一种高效、灵活的新方法, 具有广泛的应用潜力。

关键词: 引力波; 星载望远镜; 杂光抑制; 强化学习

中图分类号: O439

文献标志码: A



张梓扬, 常军, 黄一帆, 等. 基于强化学习的空间引力波探测望远镜系统外杂光抑制研究 [J]. 光电工程, 2024, 51(2): 230210
Zhang Z Y, Chang J, Huang Y F, et al. Reinforcement learning-based stray light suppression study for space-based gravitational wave detection telescope system[J]. *Opto-Electron Eng*, 2024, 51(2): 230210

Reinforcement learning-based stray light suppression study for space-based gravitational wave detection telescope system

Zhang Ziyang¹, Chang Jun^{1*}, Huang Yifan^{1*}, Chen Qinfang², Wu Yunan¹¹School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China;²Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, Shaanxi 710119, China

Abstract: In gravitational wave telescopes, the energy of the collected space target light signals is dwarfed by the energy of stray light, necessitating robust stray light suppression for reliable telescope operation. Due to the inherent unpredictability of scattered light and the intricate nature of opto-mechanical systems, the formulation of stray light suppression strategies often involves complex mathematical modeling, substantial expertise, and iterative simulations. This paper introduces a Reinforcement Learning-based approach to devise the stray light suppression scheme within a Monte Carlo ray tracing environment, specifically for space gravitational wave telescope systems. Our empirical findings confirm the efficacy of this methodology in generating effective stray light

收稿日期: 2023-08-30; 修回日期: 2023-12-08; 录用日期: 2023-12-11

基金项目: 国家重点研发计划 (2021YFC2202100)

*通信作者: 常军, bichang@bit.edu.cn; 黄一帆, huangyifan@bit.edu.cn.

版权所有©2024 中国科学院光电技术研究所

suppression strategies, yielding favorable suppression performance. This study contributes a novel, efficient, and adaptable solution to the stray light challenges faced in space gravitational wave detection as well as other high-precision optical systems, thereby holding extensive applicative promise.

Keywords: gravitational wave; spaceborne telescope; stray light suppression; reinforcement learning

1 引言

2016年, 美国的 LIGO 探测器 (臂长 4 km) 和意大利的 VIRGO 探测器 (臂长 3 km) 联合记录到了多次致密双星合并事件。尽管经过多次升级改造, 以提升其灵敏度, 这些地基探测器由于臂长的限制, 其探测能力依旧局限在 1 Hz 到 10^3 Hz 的高频范围内。因此, 为了观测宇宙中丰富的 mHz (10^{-3} Hz) 波段引力波辐射源, 构建位于太空中的天基引力波探测器成为探测低频引力波的理想方案。

为探测由百万太阳质量级别的超大质量黑洞产生的毫赫兹附近频段的引力波, 需要在太空建立臂长十万千米级别的激光干涉仪进行测量。作为获取引力波科学信号的技术源头, 空间激光干涉测距系统是空间激光干涉引力波天线最重要的技术之一, 在这个系统中, 望远镜系统既是激光接收装置, 也是激光发射装置。望远镜通过接收远处发射过来的激光, 导入本地激光干涉仪经测试质量反射后与本地激光器激光发生干涉^[1-5]。

接收从十万千米外发射过来的激光, 远距离的信号传输使得信号能量衰减极为严重, 在抵达望远镜入瞳后, 还需要经过光学元件反射抵达探测器。该天基引力波望远镜在收集空间目标光信号的过程中, 空间中诸如太阳所发出的光波等的非目标光信号会以杂散光的形式一同被望远镜所捕获, 这种杂散光能量远大于接收信号的能量。如果此时系统自身抑制杂光能力不强, 视场外杂光与自身发射的光信号经过光学元件与机械件表面的后向散射产生的杂光过大, 则对能量微弱的接收信号产生影响, 甚至可能湮没接收信号, 从而使得空间激光干涉测距系统在光学信号收集过程就已无法正常工作。基于这一问题, 杂散光抑制在空间引力波探测望远镜系统显得尤为重要^[6-7]。

传统上, 外杂光抑制通过使用遮光罩和挡光环, 通过至少三次反射, 以减少视场外散射光抵达探测器。遮光罩的口径、长度, 以及挡光环的位置、大小、形状直接影响杂散光的抑制效果。传统的自动化挡光环

设计主要依据为光线的反射, 而忽视了在非镜面表面产生的散射光线进入后续光路所带来的杂光影响。

强化学习 (Reinforcement learning, RL) 作为机器学习领域的一个研究热点, 已经广泛应用于自动驾驶、机器人控制、仿真模拟、优化调度等领域。强化学习通过最大化智能体 (agent) 从环境中获取累计的奖励 (reward), 经过学习直到完成目标的最优策略, 其侧重于学习解决问题的策略。强化学习通过建立正确的环境以及恰当的动作、奖励设置, 便能够学习出完成目标所采用的最优策略^[8-10]。

本文利用强化学习这一特点与杂散光分析进行结合, 提出了一种基于强化学习的空间引力波探测望远镜系统杂光抑制研究方案。强化学习智能体通过给出机械结构的参数, 通过蒙特卡洛光线追迹方法 (Monte Carlo ray tracing, MCRT) 构建物理模型, 建立环境。智能体所生成的机械结构经过相应的光机结构在物理模型构筑的环境中通过光线追迹得到点源透过率 (point source transmittance, PST) 或杂光系数作为奖励返回智能体, 经过学习重新给出新的机械结构参数最终获得良好抑制效果的机械结构参数。

2 基于强化学习的空间引力波探测望远镜系统杂光抑制研究原理

本文空间引力波探测望远镜系统的光学系统采用离轴四反结构, 如图 1 所示。主镜为抛物面, 次镜为二次曲面, 三镜为条纹 Zernike 多项式自由曲面, 四镜采用平面镜结构。该系统通过增加三四镜的倾角减少后向散射杂光。主镜口径较大, 采用较为简单的抛物面, 其它三镜口径较小, 则采用二次曲面与自由曲面减少光学像差。离轴四反的光机结构相较于同轴全反以及折反系统来说, 对自身红外热辐射抑制能力更强。

图 1 与表 1 所展示的为该空间引力波探测望远镜系统的光路图与系统参数, 本文将针对该光机系统进行传统方式与 RL 强化学习方式两种展开对该系统外杂光抑制的研究。

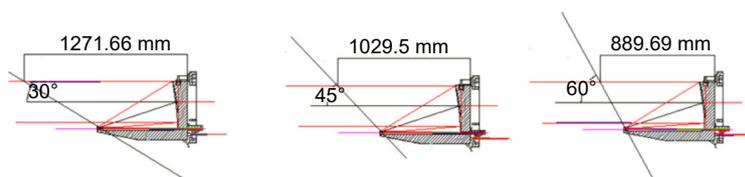


图3 不同角度下遮光罩长度示意图

Fig. 3 The length diagram of the baffle at different angles

当抑制角大于 45° 时, PST 达到 10^{-9} , 我们认为该系统的外杂光得到良好的抑制。为保证系统的杂散光抑制效果, 还需设置相应的轴向挡光环, 并根据模拟情况, 进一步确认是否设置垂轴挡光环与添加消光漆。

挡光环结构的设计原理是防止入射到遮光罩内壁的光线经过散射后径直的进入系统。挡光环的设计如图4所示, 在高度为 H 的镜筒壁上选择拐点 J 做垂直与入口平齐作为点2, 连接点1和点2, 作为镜筒接收杂散光的最小角度路径, 在1、2连接的路径中延长至镜筒点3, 点3与主镜下方的通光口径相连, 设置为第二个挡光环。点4至7同理, 设置第三个挡光环。由于底板的限制, 需要对挡光环的实际尺寸以及形状进行再加工。其中标号 J 、3、5、7 为挡光环与外遮光罩的交点, 标号 2、4、6、8 为挡光环与入瞳等高线的交点。

2.2 基于蒙特卡罗方法的光线追踪

辐射传递方程是涉及多个维度的多元积分微分方程。对于杂散光分析, 除了在光机系统中获得正确的光路外, 还需要获得辐射的数值解。蒙特卡洛光线追踪 (MCRT, Monte Carlo ray tracking) 方法是一种统计方法。MCRT 法计算射线的过程是跟踪并记录具有采样意义的光束的发射位置、方向和表面介质信息, 然后计算辐射传递因子以获得新的信息。MCRT 可以同

时处理多维复杂几何、各向异性散射和其他问题。基于 MCRT 方法, 可以轻松求解辐射传递方程。

本文采用 MCRT 方法中的路径长度法 (PL), 通过判断光线的方向, 求出光线到达所有表面的最小距离, 确定光线到达表面的表面, 输出光线到达表面时携带的位置、方向、表面介质、能量等信息来跟踪光的前进路线^[11-12]。通过该方法实现了光线的非序列跟踪, 以达到将智能代理提出的杂散光抑制方案转换为杂散光抑制指数的目的。

当光线照射在具有一定粗糙度的表面上时, 由于表面微观形态分布不均匀, 一部分入射光被吸收, 并且大部分发生散射和反射现象。在传统的 MCRT 追踪中, 由表面粗糙度、划痕、麻点和涂层等因素引起的散射光往往被忽略。这部分光是杂散光源的主要成分。本文将 MCRT 应用于杂散光分析。在 MCRT 追踪过程中, 本文介绍了表达物体表面光散射特性的 BSDF。面对 BSDF, 通常使用测量的 BSDF 数据来准确表征表面的信息。

考虑到 MCRT 在 MCRT 仿真过程中需要大量光线来抵消光传播过程中的不确定性, 因此需要大量的统计误差和时间成本, 本文使用 ABg 数学模型拟合表面实测的 BSDF 数据。表达式为^[13-14]

$$BRDF(\vec{\beta} - \vec{\beta}_0) = \frac{A}{B + |\vec{\beta} - \vec{\beta}_0|^s}, \quad (2)$$

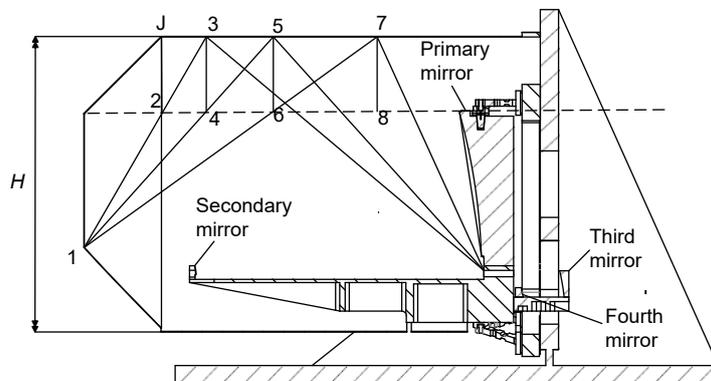


图4 传统设计方法下的光机结构

Fig. 4 The optomechanical structure under the traditional design method

2.3 DDPG 网络和模型基础

基于深度确定性策略梯度 (deep deterministic policy gradient, DDPG) 强化学习的空间引力波探测望远镜系统杂散光抑制研究原理如图 5 所示。该模型分为两部分。一部分是与代理交互的环境。环境承担将 agent 提出的决策转换为杂散光抑制效果的指标的功能, 作为对 agent 的奖励反馈。另一部分是 agent 根据此时的环境状态 s_t , 优化杂散光抑制方案。在 $\pi(s)$ 策略的指导下, agent 选择不同的杂散光抑制方案作为动作 a_t , 通过动作与环境相互作用, 产生新的内部状态和杂散光抑制效果的反馈。而后, 代理更新表示性能指数的状态 s_{t+1} 和奖励 r_t 。

基于 DDPG 网络强化学习的空间引力波探测望远镜系统其算法流程如下: 环境由当前的杂散光抑制措施、机械结构、光学结构, 以及光源位置、数量大小组成。当智能体开始训练时, 根据此时智能体所处的环境, 寻找最佳的杂散光抑制措施策略, 即动作; 随后经过光线追迹后得到此时的杂散光抑制效果, 该效果作为奖励返还智能体。

杂散光抑制方案与环境作用的过程为通过 MCRT 光线追迹后获得入瞳与出瞳处的光照度的过程, 在该方法中, 机械结构以及相应的杂散光抑制措施, 如不同形状挡光环, 多级遮光罩需要转化成不同的曲面方程加入到光线追迹的物理环境当中。

在环境初始化时便对系统初始状态进行设置, 设

定物理环境的范围、光源大小、辐射角度, 完成空间引力波探测望远镜光机结构的曲面方程构建。由智能体提出杂光抑制方案 (也即动作), 动作所包含的参数为挡光环、多级遮光罩等曲面方程的核心参数, 其中包括但不限于曲面方程的阶数、挡光环的位置、多级遮光罩的范围等。动作参数的转化与环境的初始值组成多个联立的曲面方程, 通过 MCRT 方法可以获得物理环境下不同辐射角度出入瞳的光照度, 进而得到所需的 PST 值, 也即获得代表该杂光抑制方案效果的奖励 r_t 。

如图 5(a) 所示, 我们利用 RL 算法与简化的 MCRT 模型来提高数据效率, 本文中使用的 RL 模型为 DDPG 模型, 其通过智能体观察环境获得此时的状态, 随后依据策略选择不同的动作, 获得相应的奖励, 并根据此时已经改变的环境获得新的状态以此不停继续执行下去。DDPG 网络将杂光抑制措施转换成不同的动作通过对不同杂光抑制措施的选择, 获得不同状态下的杂光抑制效果, 以此作为依据对网络进行优化, 最终获得较好的空间引力波探测望远镜系统杂光抑制方案。

基于所提出的模型, 我们定义了由状态 s_t 、动作 a_t 和奖励 r_t 组成的经验, 将杂光分析问题表述为一定程度上的 MDP 问题。仅依靠模型的输出表示系统状态是不够的, 因此我们引入此时由动作所提出的杂光抑制方案到状态描述中。

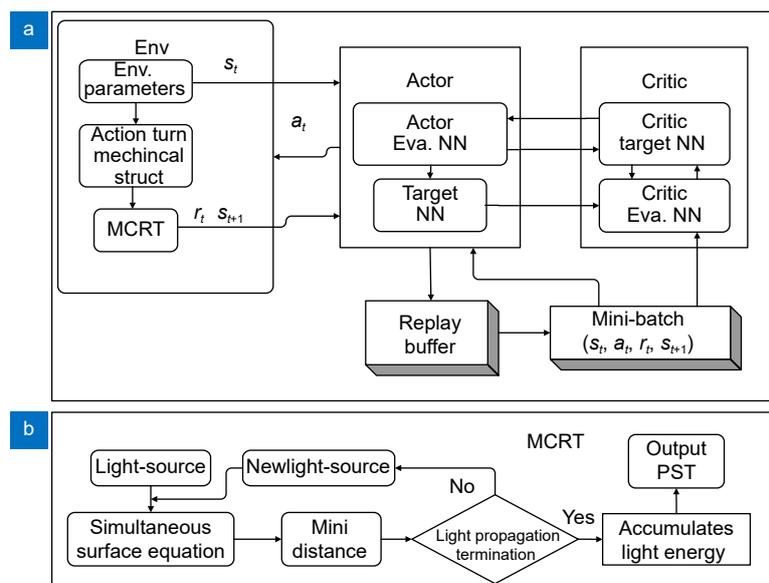


图 5 RL 网络模型。(a) Env, Actor, Critic 网络结构; (b) 环境所使用的光线追踪模型

Fig. 5 RL network model. (a) Env, Actor, Critic; (b) Ray tracing model for environment

奖励函数定义如图 6 所示。

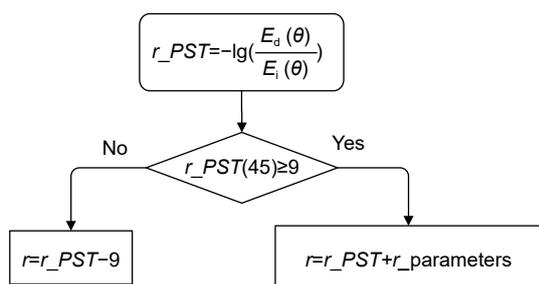


图 6 reward 函数
Fig. 6 The reward function

3 空间引力波探测望远镜系统杂光抑制结果

3.1 传统空间引力波探测望远镜系统初始模型分析

在 Light-Tools 杂光分析软件中, 完成初始模型的搭建并对光学系统轴外视场的不同角度进行蒙特卡洛光线追迹, 得到各不同角度下的光学系统轴外视场 PST, 获得望远镜系统出瞳处的杂光水平。PST 如图 7 所示。从图 7 可知, 在单独放置外遮光罩后, 虽然 60° 入射光线的 PST 接近 10^{-8} , 然而距离目标杂光抑制效果还存在一定距离, 需要后续挡光环与外遮光罩的配合才能尽可能减少杂光带来的干扰。

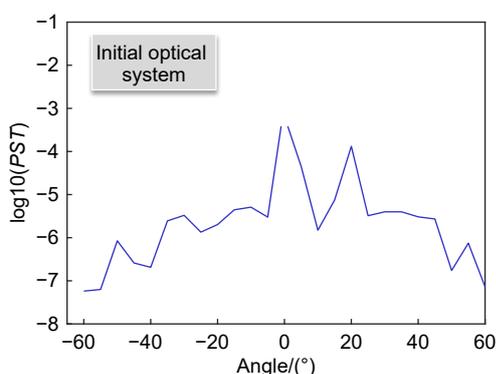


图 7 模型初始杂光抑制效果
Fig. 7 The initial stray light suppression effect of the model

3.2 传统空间引力波探测望远镜系统杂光抑制结果

根据 2.1 中所提及的光机结构与表面属性, 在杂光分析软件中, 完成建模并对光学系统轴外视场的不同角度进行蒙特卡洛光线追迹, 得到各不同角度下的光学系统轴外视场 PST, 获得望远镜系统出瞳处的杂光水平。PST 如图 8 所示。

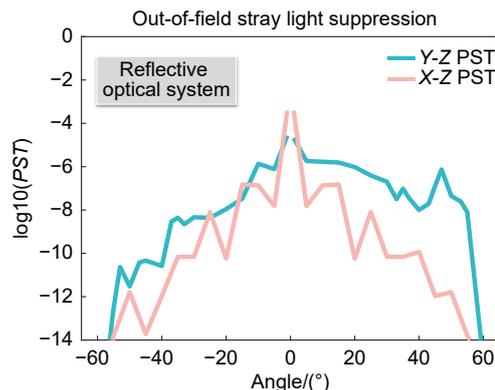


图 8 传统杂光抑制方案的杂光抑制效果
Fig. 8 The stray light suppression effect of the traditional stray light suppression scheme

根据该结果可观测到, 该光机结构在杂光抑制效果上仍存在不足, 难以满足空间引力波探测望远镜系统杂光抑制要求, 因此需要对该光机结构重新进行杂散光分析。在进行相应的杂光追迹后可得, 当光线入射到底板后, 入射至第二个挡板后散射进入后续光路当中, 散射光路如图 9 的红线所示。散射光线通过底板的一次散射打到挡光环, 再经由挡光环散射至主镜下的通光口径中进入正常的工作光路中。这种杂光入射到出瞳时与正常视场内光束角度一致, 散射次数较少, 会带来较大的杂光干扰。

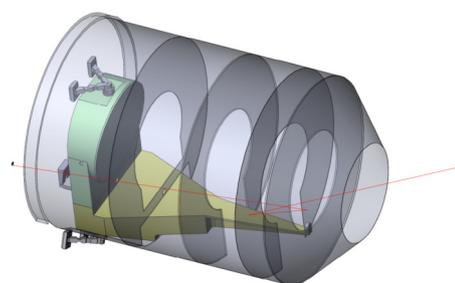


图 9 传统杂光抑制方案下的杂光光路
Fig. 9 The stray light path under the traditional stray light suppression scheme

正因为这种散射光路的存在使得传统挡光环结构的设计在对杂光抑制要求较高时存在一定局限性。而这种散射光线的存在只有经过多次仿真实验才能够被发现, 这大大拖慢了传统杂光抑制方案的制定速度, 延长了设计杂光抑制方案所需的时间。

3.3 基于 DDPG 网络的空间引力波探测望远镜系统杂光抑制结果

奖励的设置直接影响到强化学习收敛的方向, 以及对杂光抑制效果的评估。杂光抑制研究中, 除了尽

可能地减少杂散光的影响外, 还需要考虑光机结构的体积重量。奖励的组成为杂光抑制评价参数——PST 以及光机结构的体积重量评价参数, 其中首先需要保证杂光抑制效果符合项目指标要求, 其次再对光机结构进行评判。以 30° 的 PST 设置为主要的杂光抑制评价参数, 当 30° 的 PST 小于 10^{-10} 时, 奖励值设置为正, 且引入光机结构的体积重量评价。光机结构体积重量评价的奖励比 $\log(PST_{30})$ 小一个数量级。此时奖励可以被表示为

$$r = (-10 - \log(PST_{30})) + r_{parms}, \quad (3)$$

其中 r_{parms} 指光机结构的体积重量评价参数。

介于光线数的数量所带来的计算时间冗长的问题, 以及通常挡光环对光线抑制的显著效果, 适当的缩小经验池的大小。强化学习的运行结果如图 10。

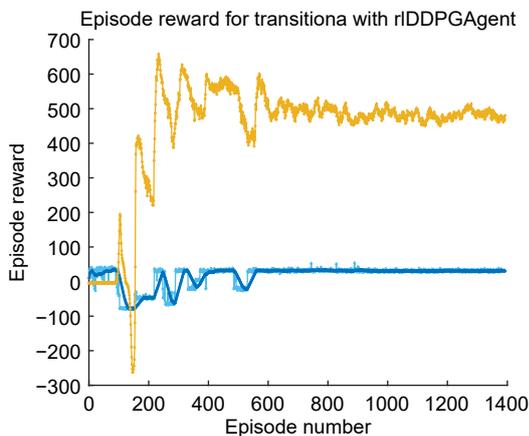


图 10 RL 运行结果

Fig. 10 The RL running results

在模型运行的早期, 智能体在机械空间里进行随机试探, 使得奖励值波动较大, PST 值起伏较为明显, 在经过早期的随机探索后, 通过经验池回放, 开始逐步优化杂光抑制的机械结构。随着回合数的增加, 模型所得到的奖励逐渐增加, 直到超过所预设的要求, 奖励开始引入对机械结构参数的评价, 在保证杂散光抑制效果的情况下, 对光机结构的体积重量进行考量。当奖励趋于平缓, Q 值趋于稳定, 我们认为强化学习在本次学习中已经收敛。将模型最后的结果导出, 获得其 actor 的参数, 依据 actor 参数的设计理念, 进行机械结构设计构建机械结构以及相应的遮光罩挡光环, 通过 light-tools 杂光分析软件的光线追迹, 对强化学习所产生的实验结果进行验证。

构建的光机结构如图 11 所示。强化学习的动作参数决定了挡光环所处的位置, 相较于传统方案的设

计, 该方案减少了一个挡光环的使用。在外遮光罩的参数设计上, 与传统方法相同, 均取了最大遮光罩内径和最大遮光罩长度。

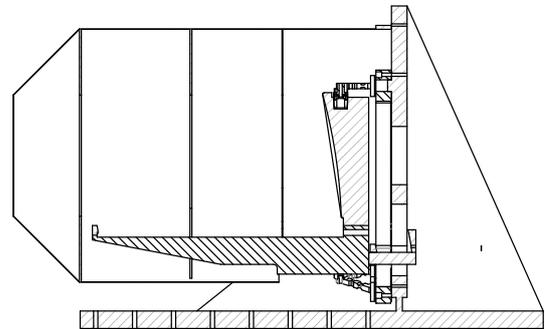


图 11 RL 杂光抑制方案的光机结构

Fig. 11 The optical-mechanical structure of RL stray light suppression scheme

各个角度下的 PST 值如图 12 所示, 相较于一开始的遮光罩设计, 对杂散光有着更好的抑制效果。

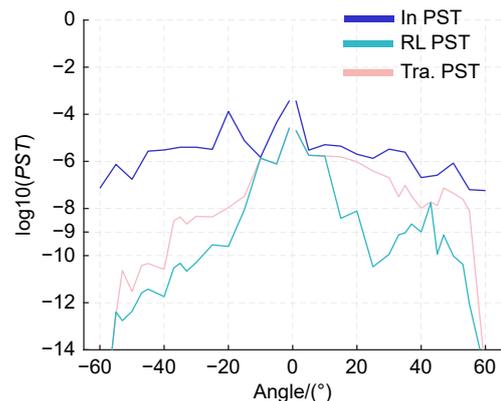


图 12 不同杂光抑制方案的杂光抑制效果对比

Fig. 12 Comparison of the stray light suppression effects of different stray light suppression schemes

根据仿真结果可得, 在 15° 以内, 由强化学习所提出的杂光抑制方案与传统方式下杂光抑制方案的效果差别不大, 在 20° 开始两种方案抑制效果出现区别。在绝大多数角度下, 强化学习所提出的杂光抑制方案的效果均超过传统方案的杂光抑制效果。但由于在强化学习模拟的过程中, 所采用的是特定几个角度进行 PST 结果的评估, 使得强化学习所提出的抑制方案中出现了部分角度杂光突变的情况。这部分杂光来源于环境杂散光入射到底板后, 向主镜底部通光孔的一次散射, 这部分散射光路受限于体积长度的要求, 在仅考虑使用挡光板而不考虑次级遮光罩设计时难以抑制。在不同杂光抑制方案中均在该特定角度

出现了较难抑制的杂光抑制。

在不考虑这部分光线影响下的角度, 在其余角度中, 强化学习所提出的杂光抑制方案效果较传统方式提高了近一个数量级。

4 结 论

针对空间引力波探测望远镜系统受杂散光影响较大, 对杂散光抑制的要求相较于常规近红外光学系统更高的问题, 本文在传统的杂散光分析与抑制研究方法的基础上提出了一种引入深度强化学习进行杂散光分析与抑制研究的新型方法。在该基于强化学习的空间引力波探测望远镜系统杂光抑制研究当中, 考虑点源透过率以及光机系统的体积重量, 使得输出的机械结构参数在保证杂光抑制效果的前提下, 维持了较低的光机系统体积重量。

传统方式的效率低下很大程度由于在仿真模拟过程中对抑制方案抑制效果的确认。在这个过程中需要不断重复仿真模拟过程以再现杂光光路, 并针对杂光光路采取相应的抑制措施。在引入新的抑制结构的同时又将引入新的杂光光路, 这种不断仿真确认的过程造成传统方式下杂光抑制研究效率的低下。强化学习通过智能体的自我迭代, 采用自动化设计的方式减少了人工确认迭代的过程, 大大提高了设计效率。

仿真分析结果表明, 对空间引力波探测望远镜系统使用基于深度强化学习的杂散光分析与抑制方法, 给出了杂光抑制效果良好, 且考虑的体积与重量的光机结构。相较于传统的杂散光分析与抑制, 其减少了由于空间引力波探测望远镜系统对杂光抑制的高标准所带来的设计者依据经验与光线追迹结果进行人工迭代的过程, 考虑了由表面属性所带来的散射, 提升了杂散光抑制效果。

参考文献

- [1] Wang Z, Ma J, Li J Q. Space-based gravitational wave detection mission: design highlights of LISA system[J]. *Chin Opt*, 2015, **8**(6): 980-987.

- 王智, 马军, 李静秋. 空间引力波探测计划-LISA系统设计要点[J]. *中国光学*, 2015, **8**(6): 980-987.
- [2] Wang Z, Sha W, Chen Z, et al. Preliminary design and analysis of telescope for space gravitational wave detection[J]. *Chin Opt*, 2018, **11**(1): 131-151.
王智, 沙巍, 陈哲, 等. 空间引力波探测望远镜初步设计与分析[J]. *中国光学*, 2018, **11**(1): 131-151.
- [3] Wang D F, Yao X, Jiao Z K, et al. Time-delay interferometry for space-based gravitational wave detection[J]. *Chin Opt*, 2021, **14**(2): 275-288.
王登峰, 姚鑫, 焦仲科, 等. 面向天基引力波探测的时间延迟干涉技术[J]. *中国光学*, 2021, **14**(2): 275-288.
- [4] Nardello M, Lintz M. Effective cross sections for stray light calculations in laser interferometry: application to LISA science interferometer[J]. *Proc SPIE*, 2021, **11852**: 118523Q.
- [5] Wang X Y, Bai S J, Zhang Q, et al. Research progress of telescopes for space-based gravitational wave missions[J]. *Opto-Electron Eng*, 2023, **50**(11): 230219.
王小勇, 白绍竣, 张倩, 等. 空间引力波探测望远镜研究进展[J]. *光电工程*, 2023, **50**(11): 230219.
- [6] Livas J C, Sankar S R. Optical telescope system-level design considerations for a space-based gravitational wave mission[J]. *Proc SPIE*, 2016, **9904**: 99041K.
- [7] Zhang Y H, Zhong Z Q, Zhang B. Analysis of surface scattering characteristics of ultra-smooth optical components in gravitational wave detection system[J]. *Opto-Electron Eng*, 2023, **50**(11): 230222.
张耘豪, 钟哲强, 张彬. 空间引力波探测系统中超光滑光学元件表面散射特性分析[J]. *光电工程*, 2023, **50**(11): 230222.
- [8] Lapeyrolerie M, Chapman M S, Norman K E A, et al. Deep reinforcement learning for conservation decisions[J]. *Methods Ecol Evol*, 2022, **13**(11): 2649-2662.
- [9] Heuillet A, Couthouis F, Diaz-Rodriguez N. Explainability in deep reinforcement learning[J]. *Knowl-Based Syst*, 2021, **214**: 106685.
- [10] Jordan M I, Mitchell T M. Machine learning: trends, perspectives, and prospects[J]. *Science*, 2015, **349**(6245): 255-260.
- [11] Liu H D, Zhou H C, Wang D D, et al. Performance comparison of two Monte Carlo ray-tracing methods for calculating radiative heat transfer[J]. *J Quant Spectrosc Radiat Transfer*, 2020, **256**: 107305.
- [12] Yarahmadi M, Mahan J R, Priestley K J. Uncertainty analysis and experimental design in the Monte Carlo ray-trace environment[J]. *J Heat Transfer*, 2019, **141**(3): 032701.
- [13] Howell J R. The Monte Carlo method in radiative heat transfer[J]. *J Heat Transfer*, 1998, **120**(3): 547-560.
- [14] Duan X Y, He C T, Lin X X, et al. Quasi-Monte Carlo ray tracing algorithm for radiative flux distribution simulation[J]. *Sol Energy*, 2020, **211**: 167-182.

作者简介



张梓扬 (1999-), 男, 北京理工大学光电学院光学工程专业硕士研究生, 研究方向为光学设计、杂散光分析、强化学习。

E-mail: yuewanping1999@qq.com



【通信作者】黄一帆, 北京理工大学光电学院教授, 主要研究方向为光学系统及光电仪器设计、复杂光学系统集成分析、计算机辅助装调、照明光学等。主持完成国家重点研发、自然科学基金仪器专项等 10 余项省部级以上项目。

E-mail: huangyifan@bit.edu.cn



【通信作者】常军, 博士, 教授, 博士生导师, 教育部新世纪优秀人才。所在的学科是光学工程(国家一级学科), 主要从事光电仪器系统设计、检测方面的研究。作为第一负责人主持和参加国家 863 项目、国家重点研发计划十三五、国家自然科学基金项目、航空基金项目、航天基金项目、横向课题等。在国内外重要期刊上发表高水平论文 130 余篇, 其中被 SCI/EI 分别收录 60/70 余篇; 授权专利 42 项。本科生课程主讲光学系统设计与制造, 博士生课程主讲现代光学系统设计。中国光学学会高级会员。

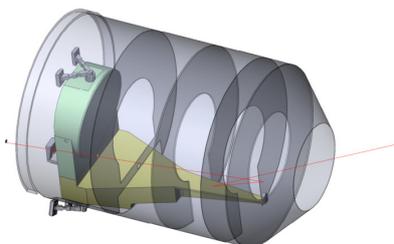
E-mail: bitchang@bit.edu.cn



扫描二维码, 获取PDF全文

Reinforcement learning-based stray light suppression study for space-based gravitational wave detection telescope system

Zhang Ziyang¹, Chang Jun^{1*}, Huang Yifan^{1*}, Chen Qinfang², Wu Yunan¹



The stray light path under the traditional stray light suppression scheme

Overview: In the field of space gravitational wave detection, the functionality of telescopes is critically hampered by the presence of stray light. This stray light, significantly overpowering the energy of collected space-target optical signals, poses a substantial challenge, necessitating robust stray light suppression for the telescope's reliable operation. Traditionally, this suppression has been achieved through complex mathematical modeling, extensive expertise, and iterative simulations, given the inherent unpredictability of scattered light and the intricate nature of optical-mechanical systems.

This study introduces a novel approach to stray light suppression by integrating reinforcement learning, a key area in machine learning. Reinforcement learning optimizes problem-solving strategies through an intelligent agent that learns to maximize cumulative rewards from its environment to achieve specific objectives. By applying this methodology, we address the limitations of traditional stray light suppression scheme formulation.

Our approach employs reinforcement learning to develop a strategy for stray light suppression within a Monte Carlo ray tracing-based environment, tailored for Space Gravitational Wave Detection Telescope systems. This method is particularly adept at identifying and mitigating stray light paths caused by surface scattering, a prevalent issue in optical systems.

A comparative analysis was conducted using the specialized stray light analysis software to evaluate the effectiveness of traditional methods against those developed through reinforcement learning. The results reveal that reinforcement learning-based schemes surpass traditional methods in identifying and mitigating stray light paths, especially those caused by surface scattering. Compared with the traditional stray light analysis and suppression, it reduces the designer's manual iteration process based on experience and ray tracing results due to the high standard of stray light suppression brought by the space gravitational wave detection telescope system. The scattering caused by surface properties is considered, and the stray light suppression effect is improved.

The empirical findings from this study confirm the superiority of reinforcement learning in formulating effective stray light suppression measures for space gravitational wave detection telescope systems. The approach not only achieves superior suppression outcomes but also introduces an efficient, flexible, and innovative solution to the challenges of stray light in space gravitational wave detection and other high-precision optical systems.

This research marks a significant advancement in the field, offering a novel methodology that can be applied to a range of optical systems requiring high precision and reliability. It opens up new avenues for future developments in optical system design and analysis, setting a new standard for addressing stray light challenges in space-based applications.

Zhang Z Y, Chang J, Huang Y F, et al. Reinforcement learning-based stray light suppression study for space-based gravitational wave detection telescope system[J]. *Opto-Electron Eng*, 2024, 51(2): 230210; DOI: [10.12086/oe.2024.230210](https://doi.org/10.12086/oe.2024.230210)

Foundation item: Project supported by National Key R&D Program of China (2021YFC2202100)

¹School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China; ²Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, Shaanxi 710119, China

* E-mail: bitchang@bit.edu.cn; huangyifan@bit.edu.c