Paper ID: 83548
student

IAF ASTRODYNAMICS SYMPOSIUM (C1)
Interactive Presentations - IAF ASTRODYNAMICS SYMPOSIUM (IPB)

Author: Mr. Longwei Xu
Harbin Institute of Technology, China

Mr. Shi Qiu
China
Prof. Xibin Cao
China

REINFORCEMENT LEARNING FOR REAL-TIME LOW-THRUST RELATIVE ORBITAL RENDEZVOUS DESIGN

Abstract

The optimal orbital rendezvous mission has always been a research area of great interest. Compared to impulse rendezvous, low-thrust engines are widely used in engineering applications due to their advantages, such as high specific impulse and multiple operations. In traditional space missions, the transfer trajectory design is typically calculated on the ground and then uploaded to the spacecraft. Once environmental disturbances or operational errors occur, there can be significant deviations between the flight trajectory and the designed plan. However, traditional optimization methods such as evolutionary algorithms and optimal control theory are challenging to deploy on onboard computers due to their high computational requirements. In recent years, reinforcement learning has effectively addressed many decision-making tasks in complex and dynamic environments. By interacting with the environment, deep neural networks can predict state changes and select optimal actions. Its advantages of short computation time and good robustness make real-time guidance possible. In this paper, a reinforcement learning-based approach is proposed to design the real-time low-thrust rendezvous trajectories in linear relative motions. An analytical state propagation is proposed for relative motions near circular orbits. The continuous time is discretized, and the control variables at the discrete time nodes are chosen as the design variables. Linear interpolation is used to approximate the control variables between nodes. The objective function is derived by integrating the control variables. Accordingly, the low-thrust rendezvous problem is formulated as a Markov decision process that considers energy and fuel consumption, transfer time, relative state, and the dynamical model. An efficient actor-critic algorithm is used to train policy networks for real-time generation of control variables based on the current state. The trained networks are evaluated in the simulations of low-thrust rendezvous, and the results accurately approximate the optimal thrust. The computational time of policy networks is significantly reduced compared to traditional optimization methods, enabling deployment on spacecraft for real-time guidance.