**HOVERING CONTROL FOR GRAVITY TRACTOR USING ASYNCHRONOUS METHODS FOR REINFORCEMENT LEARNING**

**Jucheng Lu, Bingwei Wei, Haibin Shang, Pingyuan Cui, Rui Xu, Shengying Zhu, Ai Gao**

*Beijing Institute of Technology, 100081 Beijing, People’s Republic of China, 010-68913550, shanghb@bit.edu.cn*

***Keywords:*** *Gravity Tractor, Hovering Control, Binary Asteroid Systems, Reinforcement Learning*

**ABSTRACT**

Potentially Hazardous Object (PHO) refers to the near-Earth object which has a minimum orbital intersection distance with Earth of less than 0.05 AU. These objects include binary asteroid systems e.g. 1999kw4. Researchers has studied a series of deflection schemes, such as gravity tractor, kinetic impactors, laser beaming and low-thrust deflection via electric propulsion or solar sails. Gravity tractor is a long-term hovering project, which uses the mutual gravitational force between a hovering spacecraft and a target object as a towline. Comparing with hovering in a unitary asteroid, both the superposition of the gravity field and the evolution of the binary asteroid increase the uncertainty of this dynamic system. The traditional control theories may be failed when they face this high uncertain system. This paper proposes a novel hovering control method based on reinforcement learning (RL) with asynchronous methods for achieving the aim of adapting the uncertain environment.

In this paper, the gravity field of the binary asteroid system is modeled as double ellipsoids model which the system’s exterior potential can be superposed from both ellipsoids. The triaxial ellipsoid’s gravitational potential energy is calculated both by an elliptical integral and by a second degree second order spherical harmonic series, which shows the discrepancy of the environments. In order to retain the general feature of the gravitational field, the ellipsoid is chosen as gravity-best-fit ellipsoid whose gravity potential is consistent with the irregular asteroid’s gravity potential in a distance. The spacecraft plays the role of the agent in RL. The control is determined by the policy which is composed of actor and critic, where artificial neural network (ANN) is employed as the parameter description. An asynchronous method is employed to train the parameter of the ANN in this paper. The model is trained during the interaction between the agent and the environment while RL algorithm make the agent adapt different environment and evolve with the variation of the environment. To demonstrate that the controller can adapt the change of the dynamics and learn online, the training environment differs from the test environment in numerical experiments. Simulation shows that the spacecraft can achieve and maintain the hovering state in spite of the poor precision of the training environment. The position error can be reduced to 1m in a changing uncertain environment. Further more, the control can be improved using the data which is produced during this long-term mission.