دانشكده مهندسي هوافضا

فرم تعریف پروژه کارشناسی ارشد



شماره: پیوست:

تاريخ:

14.4/01/0

نام و نام خانوادگی: علی بنی اسد شماره دانشجویی: ۴۰۱۲۰۹۲۴۴ معدل: ۱۶/۳۴ گرایش: فضا تعداد واحدهای گذرانده: ۱۵ استاد راهنما: دکتر نوبهاری استاد راهنمای همکار: تعداد واحد پروژه: ۶ استاد ممتحن:

عنوان كامل پروژه:

فارسى:

هدایت یادگیری تقویتی مقاوم مبتنی بر بازی دیفرانسیلی در محیطهای پویای چندجسمی با پیشران کم

انگلیسی:

Robust Reinforcement Learning Differential Game Guidance in Low-Thrust, Multi-Body Dynamical Environments

نوع پروژه: کاربردی: ■ بنیادی: □ توسعهای: ■

معرفي موضوع:

در سالهای اخیر، پیشرفتهای فناوری در زمینههای مختلف، از جمله کنترل پرواز، پردازش سیگنال و هوش مصنوعی، به افزایش کاربردهای ماهواره با پیشران کم در منظومه زمین-ماه کمک کرده است. ماهواره با پیشران کم میتواند برای تعقیب ماهوارهها، انتقال مداری و استقرار ماهوارهها استفاده شود. روشهای هدایت بهینه قدیمی جهت کنترل ماهوارهها اغلب نیازمند فرضیات ساده کننده، منابع محاسباتی فراوان و شرایط اولیه مناسب هستند. الگوریتمهای مبتنی بر یادگیری تقویتی این توانایی را دارند که بدون مشکلات اشاره شده هدایت ماهواره را انجام دهند. به همین دلیل، این الگوریتمها میتوانند امکان محاسبات درونی (On-board Computing) را فراهم میکنند.

اهميت موضوع:

رویکردهای قدیمی برای هدایت فضاپیما در محیطهای پیچیده چندجسمی اغلب به ایستگاههای کنترل زمینی متکی هستند که میتوانند در معرض خرابی ارتباطات، تاخیرهای زمانی و محدودیتهای انتقال داده باشند. لذا، ارائه روشهای هدایت بهینه و مقاوم که امکان محاسبه فرامین هدایت را درون ماهواره فراهم کند، از اهمیت بالایی برخوردار است.

كاربردها:

الگوریتمهای هدایت مقاوم که محاسبات آنها به صورت درون مداری انجام می شود، در ماموریتهای فضایی مختلفی مانند ماهوارهای، سفر به ماه و نزدیک به زمین کاربرد دارند. این الگوریتمها دارای مزایایی مانند استقلال از ایستگاه زمینی، سرعت تصمیمگیری و بهبود ایمنی هستند. از کاربردهای این الگوریتمها می توان به کنترل مدار و وضعیت ماهوارهها، کنترل مدار و وضعیت فضاپیماها در طول سفر به ماه، کنترل مدار و وضعیت فضایی، تعقیب اهداف در مدارهای مختلف و انجام مانورهای پیچیده اشاره کرد.

تعريف دقيق مسئله:

در پژوهش حاضر، مسأله هدایت فضاپیما برای سیستم زمین-ماه با استفاده از مدل دینامیکی سهجسم محدود دایرهای Circular مطرح میشود. در این مسأله، یک فضاپیما با سیستم پیشران کم به یک مسیر مرجع در سیستم Restricted Three-Body Problem (CR3BP) منتقل میشود. شرایط اولیه فضاپیما از مسیر مرجع انحرافهای تصادفی دارند. هدف از این پژوهش، توسعه یک

دانشكده مهندسى هوافضا

فرم تعریف پروژه کارشناسی ارشد



شماره: پیوست:

تاريخ:

14.4/01/0

الگوریتم هدایت حلقهبسته با استفاده از اصول یادگیری تقویتی بر مبنای بازی دیفرانسیلی است که به فضاپیما اجازه میدهد با فرض بدترین اغتشاش به مسیر مرجع بازگردد، مسیر مرجع را دنبال کرده و به مدار مقصد برسد. بازی دیفرانسیلی موجب میشود الگوریتم هدایت نه تنها بهینه، بلکه مقاوم نیز باشد.

فرضيات مسئله:

- مدل دینامیکی مسأله، مسأله سه جسم محدود دایرهای سیستم زمین-ماه است. این فرضیه یک سیستم دو جسمی ساده است که فضاپیما را بهعنوان جسم سوم بسیار کوچک فرض میکند.
 - حرکت فضاپیما تنها در صفحه اتفاق میافتد و تنها نیروهای گرانشی و اغتشاش درنظر گرفته میشوند.
- فضاپیما به یک سیستم پیشران کم مجهز است که شعاع تاثیر مشخصی دارد. میزان و جهت پیشران بهصورت مستمر تغییر کرده و قابل تنظیم است.
 - مسأله هدایت شامل انتقال بین مدارهای دورهای (مدارهای لیاپانوف) در مدل CR3BP زمین-ماه است.
- محیط یادگیری تقویتی ویژگیهای مارکوف را دارد؛ به این معنا که حالت فعلی تمام اطلاعات لازم برای پیشبینی حالتهای آینده را فراهم میکند.
 - فضاپیما در مسیر نزدیک به مرجع با فرض بدترین اغتشاش میماند تا به مدار مقصد برسد.

دانشكده مهندسي هوافضا

فرم تعریف پروژه کارشناسی ارشد



تاريخ:

14.4/1.

روش انجام كار:

• مرحله ۱: مدلسازی و شبیهسازی محیط دینامیکی

- پیادهسازی مدل دینامیکی محیط مورد نظر انجام میشود. این مدل باید شامل تمام جوانب مهم محیط از جمله قوانین حرکت، ابعاد فضایی و حالت اولیه باشد.
- مدل دینامیکی محیط بهصورتی که اطلاعات مورد نیاز برای یادگیری تقویتی (مانند حالت فعلی) را تامین کند، ایجاد میشود.

• مرحله ۲: ایجاد و آموزش بازیگر هدایت

- در این مرحله، بازیگری برای انجام هدایت سفینه در محیط ایجاد می شود. این بازیگر مسئول انتخاب اعمال (اعمال کمپیشران) بر اساس حالت فعلی محیط است.
- از الگوریتمهای یادگیری تقویتی مختلف مانند Q-Learning، (Q-Networks (DQN)، یا Proximal Policy Optimization، یا الگوریتم بازیگر هدایت استفاده می شود. انتخاب الگوریتم بستگی به محیط و اهداف آموزش دارد. (PPO)

پیشینهی موضوع:

پژوهشهای خارجی:

هدایت فضاپیماها معمولاً با استفاده از ایستگاههای زمینی انجام میشود. با این حال، این تکنیکها دارای محدودیتهایی از جمله حساسیت به قطع ارتباطات، تاخیرهای زمانی، و محدودیتهای منابع محاسباتی هستند. الگوریتمهای یادگیری تقویتی و بازیهای دیفرانسیلی میتوانند برای بهبود قابلیتهای هدایت فضاپیماها، از جمله مقاومت در برابر تغییرات محیطی، کاهش تاخیرهای ناشی از ارتباطات زمینی، و افزایش کارایی محاسباتی، مورد استفاده قرار گیرند.

هدایت فضاپیماها معمولاً پیش از پرواز انجام میشود. این روشها میتوانند از تکنیکهای بهینهسازی فراگیر [۱] یا برنامهنویسی غیرخطی برای تولید مسیرها و فرمانهای کنترلی بهینه استفاده کنند. با این حال، این روشها معمولا حجم محاسباتی زیادی دارند و برای استفاده درونسفینه نامناسب هستند [۲]. یادگیری ماشین میتواند برای بهبود قابلیتهای هدایت فضاپیماها استفاده شود. کنترلکننده شبکه عصبی حلقهبسته میتواند برای محاسبه سریع و خودکار تاریخچه کنترل استفاده شود. یادگیری تقویتی نیز میتواند برای یادگیری رفتارهای هدایت بهینه استفاده شود.

روشهای هدایت و بهینهسازی مسیر فضاپیماها بهطور کلی به راهحلهای اولیه مناسب نیاز دارند. در مسائل چند جسمی، طراحان مسیر اغلب حدسهای اولیه کمهزینهای برای انتقالها با استفاده از نظریه سیستمهای دینامیکی و منیفولدهای ثابت [۳، ؟] ایجاد میکنند.

شبکههای عصبی ویژگیهای جذابی برای فعالسازی هدایت در فضاپیما دارند. بهعنوان مثال، شبکههای عصبی میتوانند بهطور مستقیم از تخمینهای وضعیت به دستورهای پیشران کنترلی که با محدودیتهای مأموریت سازگار است، برسند. عملکرد هدایت شبکههای عصبی در مطالعاتی مانند فرود بر سیارات [۲۵]، عملیات نزدیکی به سیارات [۲۴] و کنترل فضاپیما با پیشران ازدسترفته [۲۲] نشان داده شده است. تازهترین پیشرفتهای تکنیکهای یادگیری ماشین در مسائل خودکارسازی درونی بهطور گستردهای مورد مطالعه قرار گرفتهاند؛ از پژوهشهای اولیه تا تواناییهای پیادهسازی. بهعنوان مثال، الگوریتمهای یادگیری ماشین ابتدایی در فضاپیماهای مریخی نبرد برای کمک به شناسایی ویژگیهای زمینشناسی تعبیه شدهاند. الگوریتم AEGIS توانایی انتخاب خودکار هدف توسط یک دوربین در داخل فضاپیماهای ماهای (Opportunity ، قرآیند دقت افزایی در داخل فضاپیماهای نیاز به ۹۴ تا ۹۶ ثانیه دارد [۵]، که به طور قابل توجهی کمتر از زمان مورد نیاز برای ارسال تصاویر به زمین و انتظار برای انتخاب دستی توسط دانشمندان است. برنامههای آینده برای کاربردهای یادگیری ماشین درونسفینه شامل تواناییهای

دانشكده مهندسي هوافضا

فرم تعریف پروژه کارشناسی ارشد



تاريخ:

14.4/01/0

رباتیکی درونسفینه برای فضاپیمای Perseverance [۷، ۸] و شناسایی عیب برای Europa Clipper [۶] می شود. الگوریتمهای یادگیری ماشین پتانسیلی برای سهم مهمی در مأموریتهای اتوماسیون آینده دارند.

علاوه بر رباتیک سیارهای، پژوهشهای مختلفی به استفاده از تکنیکهای مختلف یادگیری ماشین در مسائل نجومی پرداختهاند. در طراحی مسیر عملکرد رگرسیون معمولاً مؤثرتر هست. به عنوان مثال، از یک شبکه عصبی (NN) در بهینهسازی مسیرهای رانشگر کمپیشران کمپیشران استفاده شده است [۹]. پژوهشهای جدید شامل شناسایی انتقالهای هتروکلینیک [۱۰]، اصلاح مسیر رانشگر کمپیشران [۱۱] و تجزیه و تحلیل مشکلات ازدسترفتن رانشگر [۱۲] میشود.

تکنیکهای یادگیری نظارتی میتوانند نتایج مطلوبی تولید کنند؛ اما، دارای محدودیتهای قابل توجهی هستند. یکی از این محدودیتها این است که این رویکردها بر وجود دانش پیش از فرآیند تصمیمگیری متکی هستند. این امر مستلزم دقیق بودن دادههای تولیدشده توسط کاربر برای نتایج مطلوب و همچنین وجود تکنیکهای موجود برای حل مشکل کنونی و تولید داده است.

در سالهای اخیر، قابلیت یادگیری تقویتی (RL) در دستیابی به عملکرد بهینه در دامنههایی با ابهام محیطی قابل توجه، به اثبات رسیده است [۱۴، ۱۳]. هدایت انجام شده توسط RL را میتوان به صورت گسترده بر اساس فاز پرواز دسته بندی کرد. مسائل فرود [۱۶، ۱۵] و عملیات در نزدیکی اجسام کوچک [۲۹، ۲۵]، از حوزههای پژوهشی هستند که از RL استفاده میکنند. تحقیقات دیگر شامل مواجهه تداخل خارجی جوی [۱۸]، نگهداری ایستگاهی [۱۹] و هدایت به صورت جلوگیری از شناسایی [۲۰] است. مطالعاتی که فضاپیماهای رانشگر کم پیشران را در یک چارچوب دینامیکی چند بدنی با استفاده از RL انجام شده است، شامل طراحی انتقال با استفاده از [۲۲] است.

پژوهشهای داخلی:

در زمینه هدایت و کنترل فضاپیماها و ماهوارهها، پیشرفتهای زیادی در سالهای اخیر حاصل شده است. این پیشرفتها عمدتاً ناشی از توسعه روشهای جدیدی مانند یادگیری تقویتی، کنترلهای تطبیقی و بینایی کامپیوتر است. این پیشرفتها به بهبود ایمنی و کارایی ماموریتهای فضایی کمک کرده است. در مرجع [۲۶]، یادگیری تقویتی برای مانور اتصال دو فضاپیما انجام شده است. در مرجع [۲۷]، کنترلهای تطبیقی برای بازپیکربندی آرایش پروازی ماهوارهها پیادهسازی شده است. در مرجع [۲۸]، بینایی کامپیوتر برای ناوبری و کنترل بازوی رباتیک ماهواره اجرا شده است.

اهداف پژوهش:

- طراحی یک بازیکن جهت هدایت به کمک یادگیری تقویتی
- طراحي يک الگوريتم هدايت مقاوم به کمک بازي ديفرانسيلي
 - بهبود عملكرد فضاپيما در محيط سه جسم
- عدم نیاز به ایستگاه زمینی جهت محاسبه صدور فرمان هدایت
 - پیادهسازی الگوریتم درون سفینه ای در محیط شبیه سازی شده

دانشكده مهندسي هوافضا

فرم تعریف پروژه کارشناسی ارشد



تاريخ:

1407/01/10

جدول ۱: جدول برنامهریزی پروژه ناوبری اینرسی مشارکتی مبتنی بر هوش مصنوعی

فعاليتها	فصل
انجام مطالعه و مروری جامع در زمینه بازی دیفرانسیلی و یادگیری تقویتی.	
بررسی و تعریف دقیق مسأله یادگیری تقویتی و الگوریتم هدایت به وسیله	تابستان سال اول
یادگیری تقویتی در مسئله سهجسم.	03, 0 0
طراحی ساختار محیط سهجسم و بازیکن یادگیری تقویتی و ارتباط بین	
محيط و بازيكن.	
طراحی مسیر انتقال بهینه بین دو مدار	
شبیهسازی دقیق محیط مسأله سهجسم جهت آموزش بازیکن یادگیری تقویتی	نيمسال اول سال دوم
طراحی و توسعه سیستم یادگیری تقویتی جهت محاسبه دستور	
پیادهسازی هدایت فضاپیما با استفاده از الگوریتمهای کلاسیک	
آموزش الگوریتمهای یادگیری تقویتی در محیطهای شبیهسازی شده	نيمسال دوم سال دوم
اصلاح و بهبود عملکرد یادگیری تقویتی در هدایت فضاپیما	II II
نتیجه گیری از تحقیقات و ارائه گزارش نهایی شامل معرفی مسئله،	تابستان سال دوم
روشهای استفادهشده، نتایج حاصل و پیشنهادات برای تحقیقات آتی	

نتایج مورد انتظار:

- افزایش ایمنی ماموریت بهدلیل استفاده از بازی دیفرانسیلی
- كاهش مصرف سوخت و هزينه بهدليل بهينهبودن الگوريتم هدايت
 - هدایت فضاپیما بدون نیاز به ایستگاه زمینی
- کاهش هزینه ساخت بهدلیل عدم نیاز به سیستمهای ارتباطی قوی روش صحتسنجی نتایج:
 - مقایسه با سایر روشهای معتبر
 - مدلسازی و شبیهسازی سیستم محیط و بازیکن گلوگاههای پیشبینیشده:
- آموزش شبکه: به علت پیچیده بودن شبکه و محیطی که طراحی شده است، محاسبات سنگین می شود.
- پایداری و کنترل: به علت غیرخطی بودن دینامیک سامانه و ناپایداری ذاتی سیستم، کنترل و پایداری آن مشکل است. و آوریها:

پیش از این، از ترکیب یادگیری تقویتی و بازی دیفرانسیلی در یک مسئله چندجسمی پویا استفاده نشده است. بهعلاوه، منطق هدایت ساختهشده از ترکیب دو روش، یک هدایت مقاوم و در عین حال بهینه است، که تاکنون در این مسئله انجام نشده است.

تاريخ:	\ * ° \\°\\\\°
شماره:	

بسمه تعالی دانشکده مهندسی هوافضا فرم تعریف پروژه کارشناسی ارشد



References

- [1] Vavrina, Matthew A, Jacob A Englander, Sean M Phillips, and Kyle M Hughes. "Global, multi-objective trajectory optimization with parametric spreading." In AAS AIAA Astrodynamics Specialist Conference 2017, Tech. No. GSFC-E-DAA-TN45282, 2017.
- [2] Ocampo, Cesar. "Finite Burn Maneuver Modeling for a Generalized Spacecraft Trajectory Design and Optimization System." Annals of the New York Academy of Sciences 1017 (2004): 210-233. doi: 10.1196/annals.1311.013.
- [3] Marchand, B. G. and Scarritt, S. K. and Pavlak, T. A. and Howell, K. C. "A dynamical approach to precision entry in multi-body regimes: Dispersion manifolds." Acta Astronautica 89 (2013): 107-120. doi: 10.1016/j.actaastro.2013.02.015.
- [4] T.A. Estlin, B.J. Bornstein, D.M. Gaines, R.C. Anderson, D.R. Thompson, M. Burl, R. Castaño, M. Judd. "Aegis automated science targeting for the MER Opportunity rover." ACM Trans. Intell. Syst. Technol. (TIST), 3, 1-19, 2012.
- [5] R. Francis, T. Estlin, G. Doran, S. Johnstone, D. Gaines, V. Verma, M. Burl, J. Frydenvang, S. Montano, R. Wiens, S. Schaffer, O. Gasnault, L. Deflores, D. Blaney, B. Bornstein. "Aegis autonomous targeting for ChemCam on Mars Science Laboratory: Deployment and results of initial science team use." Science Robotics, 2, 2017.
- [6] K.L. Wagstaff, G. Doran, A. Davies, S. Anwar, S. Chakraborty, M. Cameron, I. Daubar, C. Phillips. "Enabling onboard detection of events of scientific interest for the Europa Clipper spacecraft," in: 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Association for Computing Machinery, Anchorage, Alaska, 2019, pp. 2191–2201, doi: 10.1145/3292500.3330656.
- [7] S. Higa, Y. Iwashita, K. Otsu, M. Ono, O. Lamarre, A. Didier, M. Hoffmann. "Vision-based estimation of driving energy for planetary rovers using deep learning and terramechanics," IEEE Robot. Autom. Lett. 4 (2019) 3876–3883.
- [8] B. Rothrock, J. Papon, R. Kennedy, M. Ono, M. Heverly, C. Cunningham. "Spoc: Deep learning-based terrain classification for Mars rover missions," in: AIAA Space and Astronautics Forum and Exposition, SPACE 2016, American Institute of Aeronautics and Astronautics Inc, AIAA, 2016, pp. 1–12.
- [9] B. Dachwald. "Evolutionary neurocontrol: A smart method for global optimization of low-thrust trajectories," in: AIAA/AAS Astrodynamics Specialist Conference and Exhibit, Providence, Rhode Island, 2004, pp. 1–16.
- [10] S. De Smet, D.J. Scheeres. "Identifying heteroclinic connections using artificial neural networks," Acta Astronaut. 161 (2019) 192-199.
- [11] N.L.O. Parrish. "Low Thrust Trajectory Optimization in Cislunar and Translunar Space (Ph.D. thesis)," University of Colorado Boulder, 2018.
- [12] A. Rubinsztejn, R. Sood, F.E. Laipert. "Neural network optimal control in astrodynamics: Application to the missed thrust problem," Acta Astronaut. 176 (2020) 192-203.

دانشكده مهندسي هوافضا

فرم تعریف پروژه کارشناسی ارشد

..... یبوست

۰ ۱/۰۸/۱۰ تاریخ:



- [13] N. Heess, D. TB, S. Sriram, J. Lemmon, J. Merel, G. Wayne, Y. Tassa, T. Erez, Z. Wang, S.M.A. Eslami, M.A. Riedmiller, D. Silver. "Emergence of locomotion behaviours in rich environments," 2017. CoRR abs/1707.02286. arXiv:1707.02286.
- [14] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, D. Hassabis. "Mastering the game of Go without human knowledge," Nature 550 (2017) Article.
- [15] R. Furfaro, A. Scorsoglio, R. Linares, M. Massari, "Adaptive generalized ZEM-ZEV feedback guidance for planetary landing via a deep reinforcement learning approach," Acta Astronaut. 171 (2020) 156–171.
- [16] B. Gaudet, R. Linares, R. Furfaro, "Deep reinforcement learning for six degrees of freedom planetary landing," Adv. Space Res. 65 (2020) 1723–1741.
- [17] J. Broida, R. Linares, "Spacecraft rendezvous guidance in cluttered environments via reinforcement learning," in: 29th AAS/AIAA Space Flight Mechanics Meeting, American Astronautical Society, Ka'anapali, Hawaii, 2019, pp. 1–15.
- [18] B. Gaudet, R. Furfaro, R. Linares, "Reinforcement learning for angle-only intercept guidance of maneuvering targets," Aerosp. Sci. Technol. 99 (2020).
- [19] D. Guzzetti, "Reinforcement learning and topology of orbit manifolds for station-keeping of unstable symmetric periodic orbits," in: *AAS/AIAA Astrodynamics Specialist Conference*, American Astronautical Society, Portland, Maine, 2019, pp. 1–20.
- [20] J.A. Reiter, D.B. Spencer, "Augmenting spacecraft maneuver strategy optimization for detection avoidance with competitive coevolution," in: 20th AIAA Scitech Forum, AIAA, Orlando, Florida, 2020, pp. 1–11.
- [21] A. Das-Stuart, K.C. Howell, D.C. Folta, "Rapid trajectory design in complex environments enabled by reinforcement learning and graph search strategies," *Acta Astronaut.* 171 (2020) 172–195.
- [22] D. Miller, R. Linares, "Low-thrust optimal control via reinforcement learning," in: 29th AAS/AIAA Space Flight Mechanics Meeting, American Astronautical Society, Kaanapali, Hawaii, 2019, pp. 1–18.
- [23] C.J. Sullivan, N. Bosanac, "Using reinforcement learning to design a low-thrust approach into a periodic orbit in a multi-body system," in: 20th AIAA Scitech Forum, AIAA, Orlando, Florida, 2020, pp. 1–19.
- [24] B. Gaudet, R. Linares, R. Furfaro, "Terminal adaptive guidance via reinforcement meta-learning: Applications to autonomous asteroid close-proximity operations," *Acta Astronaut.* 171 (2020) 1–13.
- [25] B. Gaudet, R. Linares, R. Furfaro, "Six degree-of-freedom hovering over an asteroid with unknown environmental dynamics via reinforcement learning," in: 20th AIAA Scitech Forum, AIAA, Orlando, Florida, 2020, pp. 1–15.
- پن جم ی ن. Q تنقوی تنی ی ادگیری الگوری تم با مامواره دو می ان انتصال مسیر طراحی .(1399) ا. نژاد, شفی عی م., عراقی, صیامی [26] https://civilica.com/doc/1039632 هوافضا. و مکان یک من دسی ملی کنفران س

تاریخ: ۱۴۰۲/۰۸/۱۰ شماره: پیوست:

دانشکده مهندسی هوافضا فرم تعریف پروژه کارشناسی ارشد



باز ببرای مدل بیدون به ین متطبیقی کینترلکیننده طراحی .(1401) ناصر. مزینی, حسین, بلندی, محمدرسول, کینکاشور, [27] باز ببرای مدل بیدون به ین (1301) و افضل, فن اوری و دانش تقویتی. یادگیری با ماهواره ها پروازی آرایش پیکربندی

ایستگاه در محموله قراردادن و جای به جا حین آزاد ـ پرواز ماهواره ربات یک بازوی کنترل و ناوبری .(1401) فرزاد موسوی [28] http://library.sharif.ir/parvan/resource/496295