ISSN 2004-2965 Energy Proceedings, Vol. 18, 2021

International Conference on Applied Energy 2021
Nov. 29 - Dec. 2, 2021, in Bangkok, Thailand
Paper ID: 298

A Smooth Path Planning Learning Strategy Design for an Air-Ground Vehicle Considering Mode Switching

Jing Zhao[1,2], Chao Yang[1,2*], Weida Wang[1,2], Ying Li[1,2], Changle Xiang[1,2]

1 School of Mechanical Engineering, Beijing Institute of Technology, Beijing, 100081, China

2 Chongqing Innovation Center, Beijing Institute of Technology, Chongqing, 404100, China (*Corresponding Author)

ABSTRACT

With the ability of vertical take-off and landing, the task path of an air-ground vehicle will be significantly shortened. Accordingly, the energy consumption will be greatly reduced. Through reasonable planning of the path, such vehicle can meet the high-efficiency needs of unmanned tasks and alleviate the global energy shortage problem. To design an optimal feasible path, this paper proposes a smooth path planning learning strategy considering mode switching. A new reward function of the Q-learning algorithm is presented, considering the influence of flight obstacle crossing parameters. To avoid the redundant flight distance and energy consumption caused by frequent high flights, the flight height correction is made in the update rule. Besides that, a path smoothing modification, called double yaw correction, reduces turning points and improves the path smoothness. It further reduces the energy consumption caused by the tortuous path. This modification also points out the direction of iterative learning and accelerates the algorithm convergence speed. Finally, the proposed strategy is verified on a 40m*40m map with 0-10m obstacle height. Results show that, the proposed strategy is effective to shorten 4.08m distance and plays the role of smoothing the path. Its convergence speed is faster than the traditional algorithm.

Keywords: air-ground vehicle, path planning, mode switching, path smoothness, Q reinforcement learning

1. INTRODUCTION

The air-ground vehicle is an effective function combination of the traditional wheeled vehicle and multi-rotor aircraft [1]. This novel vehicle has two basic modes, ground mode and flight mode, respectively. With the combination of the above two modes, such vehicle need not rely on a fixed two-dimensional ground route, smoothly crossing obstacles, passing ravines and water surfaces, etc. The task path is significantly shortened and the corresponding energy consumption is greatly reduced. Reasonable path planning for the vehicle can improve the efficiency of task completion and alleviate the global energy shortage problem.

Commonly, path planning can be divided into global path planning and local path planning, respectively. Reinforcement learning is a general framework for adaptive decision making in unknown and complex environments with the ability of autonomous interactive learning [2]. It has been applied in many vehicle fields [3]. An efficient Q-learning algorithm defines new states and actions spaces, obtaining a high-quality path in terms of length, computation time, and robot safety [4]. In addition to the path length, computation time, etc., path smoothness is also an optimization part of the path planning algorithm that can effectively reduce the energy consumption of vehicles [5]. An enhanced deep reinforcement learning uses the artificial potential field algorithm to improve the action space and reward function of the deep Q-learning network algorithm, and finally obtains a smooth and optimal path [6].

Most of the existing path planning algorithms are about mobile robots with a single movement mode. There are few studies on multi-mode path planning [7, 8]. This planning is more complicated and requires in-depth research.

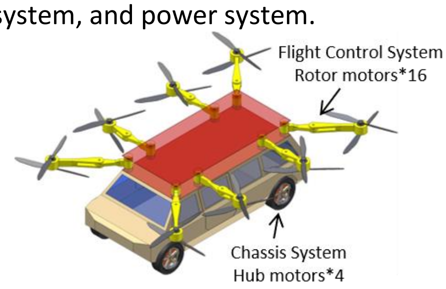
Aiming at this research requirement, a smooth path planning learning strategy considering mode switching is presented in this paper. A new reward function and update rule of the Q-learning algorithm, considering the influence of flight obstacle crossing parameters and modification of flight height, is presented to ensure the path shortening with reasonable mode switching. The corresponding driving energy consumption is reduced. For complex environments with dense obstacles, the double yaw correction of the update rule is applied to guiding the path to approximate the ideal path, i.e., the connection between the starting point and the end point in an obstacle-free environment. It can reduce turning points, improve the path smoothness, and further reduce the energy consumption. This modification also points out the direction of iterative learning and accelerates the algorithm convergence speed.

The rest of this paper is organized as follows. Section 2 introduces the physical model of the designed vehicle. In Section 3, the procedure of the strategy is presented. Section 4 offers the verification and comparison. Section 5 draws the conclusion and presents future work.

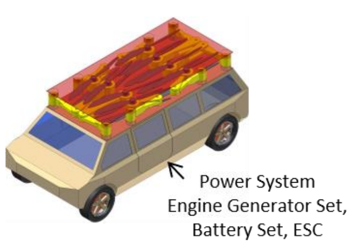
2. AIR-GROUND VEHICLE DESCRIPTION

In this section, the physical model of the designed air-ground vehicle is firstly introduced. And the power system is displayed.

The air-ground vehicle refers to the mobile platform that can both travel on the ground and fly in the air, as shown in Fig. 1. It has two modes, ground mode and flight mode, respectively. Fig. 1(a) shows the flight mode. Fig. 1(b) is the ground mode. Ground driving is the main drive mode, and proper flight mode switching can help the vehicle pass obstacles quickly and reduce the detour distance. Such vehicle consists of flight control system, chassis system, and power system.



(a) Flight mode



(b) Ground mode

Fig. 1. Air-ground vehicle model

Among these, the flight control system consists of eight-rotor twin propellers, each with an independent drive motor. And the chassis system adopts a distributed four-wheel independent drive system. The rotor motors and in-wheel motors share a common power system for energy supply. The configuration is shown in Fig. 2. The hybrid power scheme is helpful to achieve the energy saving and emission reduction [9, 10].

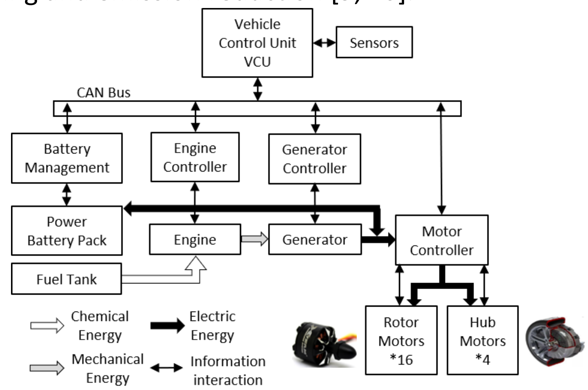


Fig. 2. Power system configuration

3. Q REINFORCEMENT LEARNING-BASED PATH PLANNING STRATEGY

In this section, a smooth path planning learning strategy considering mode switching is designed. It is important to note that the calculation of the total path distance consists of two parts: the distance of the vehicle's forward track and the height of the flight over the obstacle.

3.1 Reward function setting

The reward function for the traditional wheeled vehicle is no longer applicable to the air-ground vehicle. The new reward function is set up to take full advantages of multi-mode movements. Obstacles within the flyable height of the vehicle can be overturned by switching the flight mode. So, such reward function can effectively reduce the forward track distance. The specific definition is expressed as

$$R = a/D_{distance} - punishment \tag{1}$$

with

$$punishment = \begin{cases} 0, \ safe(H_{flight} \leq H_{max}) \\ inf, \ obstacle \end{cases} \tag{2}$$

where a is the scaling coefficient, $D_{distance}$ is the forward track distance, $punishment$ is the obstacle penalty value, H_{flight} is the flight height of the vehicle, H_{max} is the max flight height.

3.2 Update rule setting

In conjunction with the new reward function described above, the improved update rule integrating multiple modification factors is expressed as

$$Q(S_t, A_t) \leftarrow (1-\alpha)Q(S_t, A_t) + \alpha\big[b * R_{flight} - c * R_{Doubleyaw} + \gamma max_{A_{t+1}} Q(S_{t+1}, A_{t+1})\big] \tag{3}$$

with

$$R_{flight} = h_{correct} * R \quad (4)$$

$$R_{Doubleyaw} = Doubleyaw * R \quad (5)$$

where R_{flight} and $R_{Doubleyaw}$ are the modified reward functions for flight height and path inflection points, respectively. $h_{correct}$ is the flight height correction, $Doubleyaw$ is the double yaw correction, b and c are the scaling coefficients of the modifications.

The role of $h_{correct}$ is to limit the frequent high flights of the vehicle. With the previous setting of the new reward function, the algorithm tends to search for shorter forward track distance, while ignoring the increase in the total path distance caused by the flight height. These frequent flights will also cause high flight energy consumption. $h_{correct}$ can be calculated as

$$h_{correct} = 1/\left(1 + H_{flight}\right) \quad (6)$$

In addition, in the search process of the algorithm, with the increase in the density of obstacles, there will be more turning points. $Doubleyaw$ is used to guide the path to approximate the ideal path, reduce turning points, and improve the path smoothness. The smoothing of the path will effectively reduce the energy consumption of the vehicle. $Doubleyaw$ can be calculated as

$$Doubleyaw = d * yaw_1 + e * yaw_2 \quad (7)$$

with

$$ideal = \overrightarrow{end - start} \quad (8)$$

$$current_1 = \overrightarrow{current - start} \quad (9)$$

$$current_2 = \overrightarrow{end - current} \quad (10)$$

$$current_3 = \overrightarrow{next - current} \quad (11)$$

$$yaw_1 = \mathrm{acos}\left(\frac{current_1 \cdot ideal}{\|current_1\| \cdot \|ideal\|}\right) \quad (12)$$

$$yaw_2 = \mathrm{acos}\left(\frac{current_2 \cdot current_3}{\|current_2\| \cdot \|current_3\|}\right) \quad (13)$$

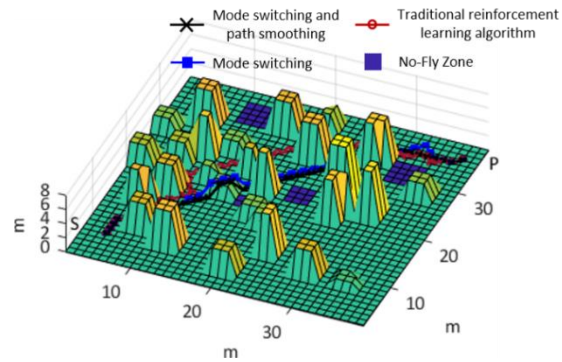
where yaw_1 is the first yaw correction, which is used to correct the deviation of the current position from the ideal path. yaw_2 is the second yaw correction, which is used to correct the steering angle of the vehicle for each action decision. $start$, end, $current$, and $next$ are the state descriptions of the vehicle. d and e are the respective scaling coefficients.

Besides smoothing the path, the double yaw correction effectively reduces the iterative learning that obviously deviates from the ideal path, and significantly improves the algorithm convergence speed.

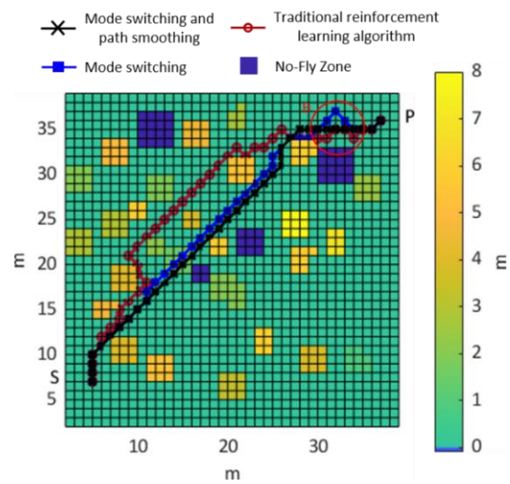
4. RESULTS AND DISCUSSION

In this section, the verification of a 40m*40m path planning map is carried out. The height of road obstacles is between 0 and 10m. The number of iterations is 400.

The traditional reinforcement learning algorithm, reinforcement learning algorithm considering mode switching, and smooth path planning learning strategy considering mode switching are adopted for path planning. The results are shown in Fig. 3. Fig. 3(a) shows the three-dimensional obstacle diagram. Fig. 3(b) shows the three-dimensional top view. The blue area is a pre-defined no-fly zone, which is not involved in planning decisions due to local regulations prohibiting access.



(a) Three-dimensional obstacle diagram



(b) Top view

Fig. 3. Path search results

As can be seen from the figure, the reinforcement learning algorithm considering mode switching reduces the polyline distance generated by obstacle avoidance compared to the traditional algorithm. But this algorithm also has many turning points. And at position B in the figure, the redundant turning causes a local path deviation, increasing the extra path distance and energy consumption. After the smooth path planning strategy, i.e., the double yaw correction is added, the path is guided to approach the ideal curve. Compared with the other two algorithms, this final path is effectively smoothed and more consistent with the actual operation of the vehicle.

The path distance comparison is shown in Fig. 4.

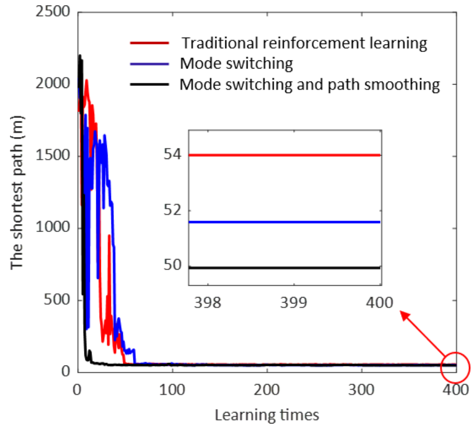


Fig. 4. Comparison of path distance

The reinforcement learning algorithm considering mode switching obtains a shorter planning path than the traditional algorithm and reduces 2.42m. The introduction of the double yaw correction smooths the path and obtains the shortest total distance, which is 1.66m less than that of the algorithm only considering mode switching. In addition, the smooth path planning learning strategy considering mode switching effectively reduces the iterative learning that deviates excessively from the ideal path. It points out the exploration direction of environmental states and accelerates the algorithm convergence speed.

5. CONCLUSION

A smooth path planning learning strategy design for the air-ground vehicle considering mode switching is proposed in this paper. The reasonable path planning for the vehicle effectively reduces the path distance and corresponding energy consumption to complete various unmanned tasks. In this strategy, a new reward function and update rule of the Q-learning algorithm, considering the influence of flight obstacle crossing parameters and modification of flight height, is presented to ensure the path shortening with reasonable mode switching. Furthermore, for complex environments with dense obstacles, the double yaw correction of the update rule is applied to guiding the path to approximate the ideal path, which has a significant impact on the efficiency of iterative learning. Results show that the reinforcement learning algorithm considering mode switching reduces 2.42m distance compared with the traditional algorithm. And after adding the double yaw correction, the smooth path planning learning strategy considering mode switching further reduces 1.66m. Besides that, the proposed strategy reduces turning points, and has the smoother path and superior convergence speed. In the future, the proposed strategy will be verified accordingly in the test platform.

ACKNOWLEDGEMENT

This work was supported by the National Natural Science Foundation of China (Grant No.51975048, 52102449).

REFERENCE

[1] W. Wang, Y. Chen, C. Yang, Y. Li, B. Xu, K. Huang, C. Xiang. An Efficient Optimal Sizing Strategy for a Hybrid Electric Air-Ground Vehicle Using Adaptive Spiral Optimization Algorithm. J. Power Sources 2022; 517.

[2] Q. Zhou, D. Zhao, B. Shuai, Y. Li, H. Williams, H. Xu. Knowledge Implementation and Transfer with an Adaptive Learning Network for Real-Time Power Management of the Plug-In Hybrid Vehicle. IEEE Trans. Neur. Net. Lear. Syst. DOI 10.1109/TNNLS.2021.3093429.

[3] B. Shuai, Q. Zhou, J. Li, Y. He, Z. Li, H. Williams, H. Xu, S. Shuai. Heuristic Action Execution for Energy Efficient Charge-Sustaining Control of Connected Hybrid Vehicles with Model-Free Double Q-Learning. Appl. Energy 2020; 267.

[4] A. Maoudj, A. Hentout. Optimal Path Planning Approach Based on Q-Learning Algorithm for Mobile Robots. Appl. Soft Comput. 2020; 97.

[5] J. Qi, H. Yang, H. Sun. MOD-RRT*: A Sampling-Based Algorithm for Robot Path Planning in Dynamic Environment. IEEE Trans. Ind. Electron. 2021; 68(8): 7244-7251.

[6] L. Li, D. Wu, Y. Huang, Z. Yuan. A Path Planning Strategy Unified with a COLREGS Collision Avoidance Function Based on Deep Reinforcement Learning and Artificial Potential Field. Appl. Ocean Res. 2021; 113.

[7] H. Liu, L. Huang, H. Ye. Autonomous Path Planning Strategy for Water-Air Amphibious Vehicle Based on Improved A* Algorithm. 2020 IEEE 18th International Conference on Industrial Informatics (INDIN) 2020; 812-816.

[8] J. Zhao, H. Li, C. Yang, W. Wang. A Novel Path Planning Method for Wheel-Legged Unmanned Vehicles Based on Improved Ant Colony Algorithm. 2021 60th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE) 2021; 696-701.

[9] Q. Zhou, J. Li, B. Shuai, H. Williams, Y. He, Z. Li, H. Xu, F. Yan. Multi-Step Reinforcement Learning for Model-Free Predictive Energy Management of an Electrified Off-Highway Vehicle. Appl. Energy 2019; 255.

[10] Q. Zhou, Y. Li, D. Zhao, J. Li, H. Williams, H. Xu, F. Yan. Transferable Representation Modelling for Real-Time Energy Management of the Plug-In Hybrid Vehicle Based on K-Fold Fuzzy Learning and Gaussian Process Regression. Appl. Energy 2022; 305.