

Retraction

Retracted: UAV Mission Path Planning Based on Reinforcement Learning in Dynamic Environment

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Journal of Function Spaces. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] G. Fu, Y. Gao, L. Liu, M. Yang, and X. Zhu, "UAV Mission Path Planning Based on Reinforcement Learning in Dynamic Environment," *Journal of Function Spaces*, vol. 2023, Article ID 9708143, 11 pages, 2023.

Research Article

UAV Mission Path Planning Based on Reinforcement Learning in Dynamic Environment

Gui Fu ^{1,2}, Yang Gao ², Liwen Liu ³, Mingye Yang ² and Xinyu Zhu ¹

¹Institute of Electronic and Electrical Engineering, Civil Aviation Flight University of China, Guanghan 618307, China

²School of Information Engineering, Southwest University of Science and Technology, Mianyang 621010, China

³College of Flight Technology, Civil Aviation Flight University of China, Guanghan 618307, China

Correspondence should be addressed to Liwen Liu; llw@cafuc.edu.cn

Received 5 August 2022; Revised 15 October 2022; Accepted 20 March 2023; Published 28 March 2023

Academic Editor: Miaochoao Chen

Copyright © 2023 Gui Fu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the rapid development of information technology, various products used in information technology are also constantly optimized. Among them, the task and path planning of UAV in the high-end robot industry has always been the focus of relevant researchers. In the high-end robot industry, in addition to the research and development of UAVs, they also continue to learn and strengthen the task and path planning of UAVs. Nowadays, using unmanned aerial vehicles for real-time shooting has become the trend of this era. Drones have brought great convenience to people's lives, and more and more people are willing to use drones. Based on the above situation, this paper studies the task and path planning of UAV based on reinforcement learning in dynamic environment. In the case of unpredictable scene parameters, reinforcement learning method can be established by value function. Thus, a more reasonable path can be given to realize the reconnaissance and detection of points of interest. MATLAB simulation experiments show that the algorithm can effectively detect targets in complex terrain composed of terrain restricted areas, and return to the designated end point to complete communication. Firstly, the development of unmanned aerial vehicles in various countries and the social status of unmanned aerial vehicles are discussed. By making UAV build threat model and task allocation in dynamic environment. The path planning and optimization of UAV in dynamic environment is studied, and the path planning algorithm and Hungarian algorithm are added. The optimized UAV has the fastest data transmission and calculation speed, while the other two types of UAVs have slower data transmission and calculation speed. In particular, ordinary UAVs also have data transmission failures, resulting in incomplete experimental results. The results show that the optimized UAV system is better in data calculation and transmission, which also shows that the UAV can quickly plan and process flight paths, which is suitable for practical applications.

1. Introduction

In today's era, the high-end robot industry is developing more and more rapidly. In the development of UAV, the wrong task and path of UAV often occur [1]. In the process of UAV working, due to the environmental factors, the sudden change of the environment cannot be avoided, which leads to the UAV system failure, and failure fall problems often occur when the UAV is in the working state [2]. Therefore, the reinforcement learning of UAV mission and path planning in dynamic environment has always been the research direction of relevant researchers [3]. Because

the working environment of UAV is outdoor, UAV needs more powerful internal system than other camera equipment. From the perspective of UAV machine itself, during the execution of work, the chip in the system will be short circuited due to long-time continuous work, and the original work task will not be completed eventually [4]. From the perspective of environment, due to the variable environmental impact, UAV will also be separated from the originally planned path and mission. Extreme environment cannot avoid varying degrees of damage to UAVs. Therefore, we should not only consider the internal system construction of UAV but also try our best to ensure that UAV can work

normally in different environments [5–8]. Therefore, in the process of summarizing the UAV mission planning path, we should not only focus on whether the UAV system is complete but also improve the working state of UAV in the dynamic environment.

In order to make the UAV better adapt to the dynamic environment, we should purposefully improve and optimize the internal tasks and planning paths of the UAV. Because the UAV itself is relatively light, it also needs to meet the principle of simplifying the internal system of the UAV, which is more practical. With the development of the times, the application range of UAV is more and more extensive, and UAV can be found in all fields. At first, when unmanned aerial vehicles were first introduced, their endurance and shooting clarity were very poor [9]. Later, the performance of UAV was continuously upgraded. UAV is mainly used to complete complex tasks with high difficulty coefficient, and the task and path planning of UAV is the key to solve complex tasks. Path planning mainly refers to the existence of threatening obstacles in the environment of UAV [10]. Then, planning the optimal route for the UAV to avoid obstacles from the starting point to the destination is also one of the main factors for the UAV to realize autonomous flight. In the task allocation of UAV, the main purpose is to enhance the time performance and environmental adaptability of UAV. For the path planning level of UAV, modifying and improving the algorithm can improve the overall performance of UAV and then improve the selection of smooth flight path in the working process of UAV, so as to improve work efficiency [11].

As the core of artificial intelligence, machine learning has three main classifications. They are supervised learning, unsupervised learning, and reinforcement learning. The purpose of reinforcement learning is to make agents autonomous in an environment and get the maximum reward. At the same time, the concept of reinforcement learning is very broad, which is called reinforcement learning or reinforcement learning in the field of artificial intelligence. In cybernetics, it is called dynamic programming; although strictly speaking, the concept of dynamic programming was put forward long before reinforcement learning [12]. At the same time, the two are the relationship between inheritance and development. However, at present, reinforcement learning, which is generally considered to be approximately equivalent in concept, is one of the three branches of machine learning. At the same time, reinforcement learning algorithm is sequential decision-making. In order to maximize the overall benefits of the system, the action strategies taken by agents in reinforcement learning are allowed to have a long-term impact. At the same time, the return is delayed, and the previous action will affect the return after multiple steps. Finally, at the expense of immediate return, the agents in the decision-making system can obtain better long-term return [13].

The innovative contribution of the research lies in the analysis of the problem of UAV mission path planning based on reinforcement learning in a dynamic environment. Adding the correlation function to the flying altitude of the UAV makes the UAV more invisible. Then, the optimized Hungarian algorithm is added to the UAV

system to make the data communication flow within the system faster. Finally, the improved artificial potential field algorithm is added to the UAV system, which also accelerates the ability of computing data in the system. The optimized UAV has the fastest data transmission and calculation speed, while the other two types of UAVs have slower data transmission and calculation speed. Compared with ordinary UAV, the UAV after reinforcement learning can better adapt to the environment and greatly improve the internal system performance of the whole UAV. From the above, it can be seen that the UAV based on reinforcement learning in the dynamic environment has a very good market prospect.

2. Development Status of Reinforcement Learning UAV in Various Countries

With the advent of the era of big data, ordinary UAVs have also been upgraded into new UAVs that can adapt to the environment [14]. Nowadays, new unmanned aerial vehicles are widely used in various industries and fields, and the performance of internal systems is constantly optimized. This paper is based on the new UAV and then studies the UAV mission and path planning [15]. Among them, in the path planning of UAV, adding optimization algorithm to the UAV system can make a simple judgment on the path more accurately [16]. Only when the path of UAV flight is correct can it improve the efficiency of UAV to complete the task and then provide convenience for people's life. Using the optimized new UAV, the multiangle picture of the captured object can be collected in real time, and the collected picture can also be saved and automatically projected in the UAV system. Only if the UAV itself does not have a fault, it can normally take flight photos according to the planned flight path. However, the performance requirements of UAV vary in different application fields. The new UAV system studied in this paper has improved its internal stability and endurance compared with the ordinary UAV system. At the same time, how to fly normally in different dynamic environments is also the main problem. Only by constantly integrating the algorithm with the internal system of UAV can the ability of UAV to adapt to the environment be greatly improved, which is unmatched by ordinary UAV. The deep reinforcement learning (DRL) method can solve the problem of creating data sets by letting UAVs collect data by themselves in the training environment. Using sac algorithm can realize the action space of obstacle avoidance scheme based on continuous UAV so that UAV can make more accurate and smooth action choices. Using depth map as input, sac is combined with variational automatic encoder (VAE). The UAV is trained to complete the obstacle avoidance task in a simulated environment composed of multiple wall obstacles.

China is the country where drones are most widely used, and most of them are used in shooting, for example, photography, video recording, and follow-up. Drones are also occasionally used by police, firefighters, the military, or geological monitoring [17]. The use of drones is convenient for observation. By taking advantage of the small size of the

UAV, the target can be photographed or tracked in the surveillance dead corner that cannot be observed by people [18]. The picture information stored inside is automatically sent to the computer terminal through the network to achieve the work purpose and complete the layout task. Compared with ordinary UAVs, it reduces human resources, saves working time, and greatly reduces the difficulty of work. Moreover, compared with the new type of UAV, the ordinary UAV is larger in size, faster in power consumption, and higher in cost, and its production and use are not a small amount [19]. The emergence of new unmanned aerial vehicles not only reduces human and material resources but also greatly reduces expenditure [20]. It is also in line with China's concept of environmental protection and energy conservation. Moreover, the service performance and service life of the internal system of the entire UAV have also been improved so that the whole UAV can give full play to its greatest advantages and shoot clearer and better images in its working state.

In order to improve the efficiency of UAV and improve the disadvantages of UAV in a single environment, the internal system of UAV is the main research object in the UK. Under the premise of ensuring that the normal working state of UAV is not affected, the internal system of UAV is actually applied to different environments, and the internal system is continuously improved, which breaks the problem that ordinary UAV cannot work in a dynamic environment. Based on the dynamic environment, this paper also studies the reinforcement learning of UAV mission and path planning.

3. UAV Mission Path Planning Based on Reinforcement Learning in Dynamic Environment

3.1. Assignment of Reinforcement Learning UAV Tasks and Threats in a Dynamic Environment. Reinforcement learning is an important method in path planning. When the scene parameters are unpredictable, reinforcement learning method can be established by the value function. Thus, a more reasonable path can be given to realize the reconnaissance and detection of points of interest. In the case of unpredictable scene parameters, reinforcement learning method can be established by value function. Thus, a more reasonable path can be given to realize the reconnaissance and detection of points of interest. MATLAB simulation experiments show that the algorithm can effectively detect targets in complex terrain composed of terrain-restricted areas and return to the designated end point to complete communication. In the process of UAV working under the dynamic environment, the task of strengthening UAV and the allocation of existing threats in the task are the focus of this paper. In this paper, the obstacles are divided according to the classification method and the robot path planning. The main threat sources in the path planning part of UAV mission planning are divided into the following three categories for description: static threat, dynamic threat, and pop-up threat. (1) Static

threat: it mainly refers to the threat that is known during mission planning and will not change in the actual flight process. (2) Dynamic threat: it is a known type of threat with a very high probability of occurrence within a certain area identified by the UAV's onboard sensors. In the working state of UAV, it is very important to model the threat of obstacles. In the process of threat modeling of obstacles, the distribution position and situation of obstacles are mainly detected. In the threat modeling, the threat source should be transmitted to the UAV so that the UAV can analyze the threat of objects. In specific practical applications, obstacles are detected first, and then, appropriate responses are made to obstacles. Therefore, no matter where the UAV is in any environment, only by avoiding damage to the fuselage can it perform all tasks normally. In the process of UAV performing tasks, due to the different difficulty of tasks, a UAV may not be able to complete. At this time, multiple UAVs are required to perform a task at the same time. Multiple unmanned aerial vehicles can perform tasks better through information transmission between them. However, whether it is a UAV or multiple UAVs, its core is to manage and implement the internal system of the UAV. The internal framework of the UAV system is shown in Figure 1.

It can be seen from Figure 1 that the control system of UAV is the above-mentioned management part. The control system of UAV manages other UAVs that perform tasks together, including the overall cognition of terrain and the analysis and judgment of obstacles. Among them, the amount of two poles in the threat potential field generated by each obstacle is as follows:

$$\begin{aligned} U_i(x, y) &= 0, \\ U_i(x, y) &= 1. \end{aligned} \quad (1)$$

From the above formula, we can judge whether there is a threat of obstacles. When the result is 0, the UAV can fly normally, and when the result is 1, the UAV needs to make an evasive action and then return to the normal flight path. In practical application, there may be more than one obstacle. After multiple obstacles are superimposed, the relevant two terminal quantities are as follows:

$$\begin{aligned} U_1 &= \frac{2y_1 + y_2 - 2y_2}{2} \\ U_2 &= \frac{y_2}{2} \end{aligned} \quad (2)$$

Before the UAV enters the working state, it usually needs to evaluate and analyze the received tasks, that is, to analyze the tasks accurately. The relevant formula is as follows:

$$f_r = \lim_{m \rightarrow \infty} \sum_{j=1}^m \left(\frac{r_{\text{safe}}}{r_j} \right)^2. \quad (3)$$

According to the above formula, the task can be accurately analyzed and evaluated. Due to the different types of

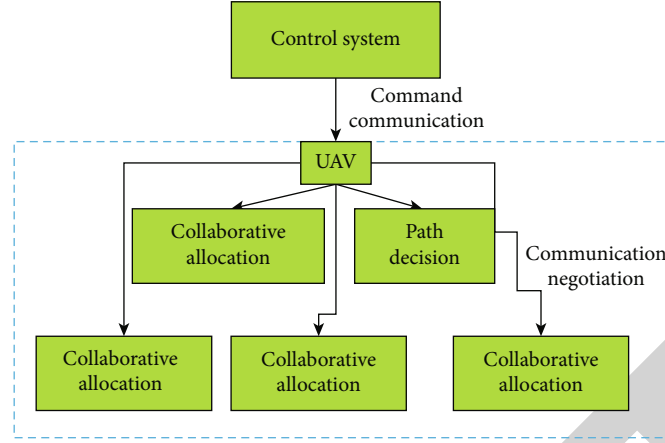


FIGURE 1: Internal framework structure management flow chart of UAV system.

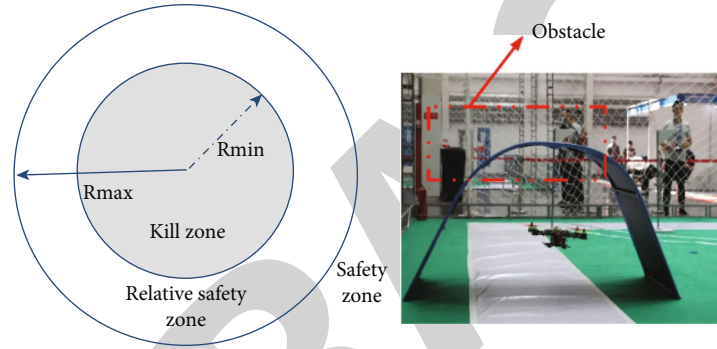


FIGURE 2: Obstacle signal threat model.

obstacles encountered by UAV in the process of mission execution, this paper analyzes and designs several types of threat sources. The first is the same UAV obstacle, which can also be called signal obstacle threat. The relevant formula for obtaining the signal of this type of obstacle is as follows:

$$U = \begin{cases} 1 & \\ \frac{|R_{\min} - r|}{e^{(R_{\max} - R_{\min})}} & \\ 0 & \end{cases} \quad (4)$$

According to the above formula, the signal received by the UAV can be calculated, and then, the threat judgment of the transmitted signal can be carried out. The obstacle signal threat model is shown in Figure 2.

It can be seen from Figure 2 that the UAV analyzes and judges the threat area of the signal sent by the obstacle, the relative safety area, and the safety area and then makes the flight route change action. In the process of UAV entering the working state, it often meets the threat of terrain obstacles as well as signal obstacles. Because in the process of carrying out tasks, it is impossible for unmanned aerial vehicles to have a smooth flow. Terrain factors should always be considered. If you pay little attention, drones will fall, causing

great damage to drones. The calculation formula of obstacle terrain model is as follows:

$$\begin{cases} U = \sum_{j=1}^m U_j \\ U_j = \text{peak}_j \times \frac{1}{e^{\left(\frac{x-h_{xj}}{h_{rxj}}\right)^2 + \left(\frac{y-h_{yj}}{h_{ryj}}\right)^2}} \end{cases} \quad (5)$$

$$U = \begin{cases} \frac{4(r - R_{\min}) * (R_{\max} - r)}{(R_{\max} + R_{\min})^2}, & R_{\min} \leq r \leq R_{\max} \\ 0 & \end{cases}$$

From the above formula, the sum of the number of threats of the obstacle terrain model and the minimum and maximum distance from the terrain obstacle can be calculated. Only by obtaining the value of this sum and the distance length can the drone ensure its own flight safety. The collision rate and mission success rate of the above two different obstacles transmitted to the computer during the operation of the UAV are shown in Figure 3.

It can be seen from Figure 3 that although there is no collision during the execution of the task, the probability of the UAV successfully completing the task is gradually

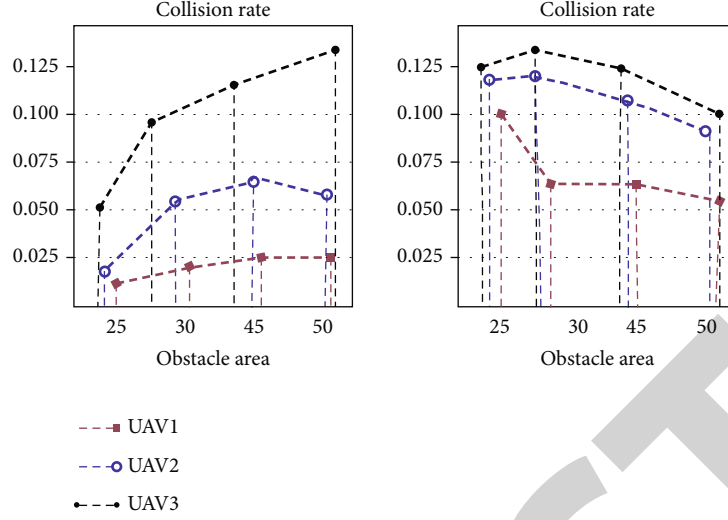


FIGURE 3: Working collision rate and mission success rate of UAV.

decreasing with the increase of obstacles. This also shows that the threat of obstacles to UAVs is an inevitable factor.

In the process of task allocation when UAV is working, the principle of task allocation has always been that task efficiency is the first, and cost is the least. In the problem of task allocation of UAV, this paper simply modeled the task allocation, and the relevant formula is as follows

$$\begin{aligned}
 A_{ij} &= \begin{cases} 1, & \text{Implementation} \\ 0, & \text{Non-implementation} \end{cases}, \\
 C &= \begin{bmatrix} c_{11} & \cdots & c_{1n} \\ \cdots & \cdots & \cdots \\ c_{m1} & \cdots & c_{mn} \end{bmatrix} \text{ and } D = \begin{bmatrix} d_{11} & \cdots & d_{1n} \\ \cdots & \cdots & \cdots \\ d_{m1} & \cdots & d_{mn} \end{bmatrix}, \\
 &\begin{cases} \text{Benefit}(A) = \lim_{m \rightarrow \infty} \sum_{i=1}^m \sum_{j=1}^m A_{ij} b_{ij} \\ \text{Cost}(A) = \lim_{m \rightarrow \infty} \sum_{i=1}^m \sum_{j=1}^m A_{ij} c_{ij} \end{cases}, \\
 &\begin{cases} \text{Max}(\text{Benefit}(A)) \\ \text{Min}(\text{Cost}(A)) \end{cases}.
 \end{aligned} \quad (6)$$

According to the above formula, the UAV can process the cost of the received tasks and then process the task allocation time and sequencing. In the process of task processing, the task time allocation data transmitted to the computer is shown in Figure 4.

It can be seen from Figure 4 that the UAV is flexible in solving tasks in the face of different task assignments. It is also more intuitive to clarify the allocation of UAV tasks at different times.

3.2. Path Planning and Optimization of UAV Based on Reinforcement Learning in Dynamic Environment. The above content describes the relevant research on reinforce-

ment learning of UAV tasks and existing threats in a dynamic environment. Next, the path planning and optimization performance of UAV in dynamic environment are studied in detail. The path planning problem of UAV mainly refers to the optimal flight path of UAV. The flight route selected by UAV must ensure the least obstacles and meet the mission requirements. Minimize the risk of UAV as much as possible and complete the tasks delivered at the least cost. In the process of path planning, the new UAV first judges whether the path is reasonable and whether the algorithm is complete and then decides the next specific operation. The second is to screen the path and analyze the performance parameters of the UAV itself. Thirdly, the route planning should be full of security, and the path planning must be hidden. Finally, in the process of executing the task, the algorithm inside the system can quickly respond to the task and make modifications and adjustments to the existing problems in time. Among them, the cost function of UAV flight altitude is as follows:

$$f_H = \begin{cases} \frac{1}{\sum_{k=1}^n H_{\max} - h_k}, & h_k \leq H_{\max} \\ \infty, & h_k > H_{\max} \end{cases}. \quad (7)$$

According to the above function, we can get the flight altitude of the UAV, better give the flight path, and realize the feature of concealment. In the process of UAV path planning, a variety of path planning algorithms can be added to the whole UAV system.

The control method of UAV formation transformation is the premise of realizing formation flight of multiple UAVs. The formation reconstruction of clustered UAVs is an important problem that we need to consider so that each UAV can reach the final position from the initial position without collision, thus ensuring the minimum cost or optimal energy consumption in the formation reconstruction process. The target allocation problem is solved by Hungary algorithm at most. It is the most common algorithm for

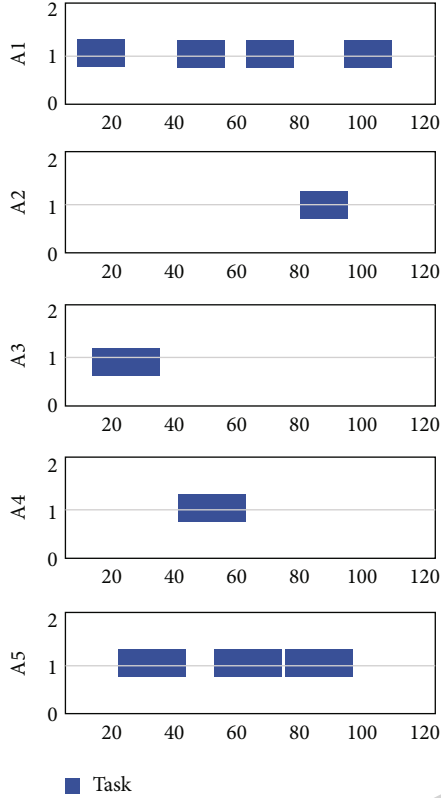


FIGURE 4: Instantaneous assignment data diagram of UAV mission.

partial graph matching. The core of this algorithm is to find the augmented path. It is an algorithm for finding the maximum matching of bipartite graph by using the augmented path. Problems with low calculation difficulty, short planning time, and high planning efficiency can meet the actual needs. Hungarian algorithm is a combinatorial optimization algorithm that solves the task allocation problem in polynomial time and promotes the later primal dual method.

The problems suitable for dynamic programming must have the following characteristics. (1) Optimal substructure: if the optimal solution of the parent problem contains the optimal solution of its subproblem, we say that the problem has an optimal substructure. That is to say, when the subproblem is optimal, the parent problem must be able to obtain the optimal solution through optimization. (2) Overlapping subproblems are essentially the same as the parent problem, except that the input parameters of the problem are different, which can be called overlapping subproblems, which is the essence of the efficiency of dynamic programming in solving problems. (3) The problem has a boundary. The subproblem does not exist under certain circumstances. We call this situation one in which the problem has a boundary. For the top-up and bottom-down methods, the boundary is the exit and entrance of the problem, respectively. (4) The subproblems are independent of each other. The subproblems are independent of each other when solving the optimal solution, that is, the solution of the self-problem is irrelevant to other parallel subproblems. Since there are few applications at present, it will not be intro-

duced in detail. Its basic idea is the same as that of dynamic programming. It also adopts the method of strategy estimation, strategy improvement, and strategy iteration to obtain the optimal strategy. However, in policy estimation, it takes the value function record of the first access to the state in a cycle. After countless rounds, the strategy is estimated by approaching the real value. It has three main characteristics: (1) the algorithm can obtain new decision experience from the past decision experience without modeling the world where the agent is located; (2) the estimation of the state value function by the algorithm is independent of each other; (3) the algorithm can only deal with the problem of episode task mode.

The two algorithms added in this paper are Hungarian algorithm and artificial potential field method, which are further optimized. The Hungarian algorithm mainly improves the communication transmission speed of the UAV internal system, while the artificial potential field method improves the calculation speed of the whole UAV internal system. Only when the speed of communication and calculation data is accelerated, the UAV can better plan the flight path. The specific process of implementing the Hungarian algorithm in the UAV is shown in Figure 5.

Figure 5 shows the internal implementation process of the Hungarian algorithm. First initialize the data, then sort the data and calculate the weight so that the idle rows are filled with data for sorting. Finally, the UAV selects the optimal flight path through the specific value calculated internally. Within the Hungarian algorithm, this paper mainly integrates the data of each gradient. The traditional Hungarian data are calculated separately, which not only wastes a lot of time but also does not improve the accuracy of calculation. The calculation formula of each gradient after integration is as follows:

$$\lim_{n \rightarrow \infty} \left(\sum_{j=1}^n \frac{(E-1) \cdot b_j}{E^j} + \frac{b_n}{E^n} \right). \quad (8)$$

Using Hungarian algorithm to integrate and process the data can greatly shorten the time of data allocation and calculation. In the implementation of Hungarian algorithm within the UAV system, it is also necessary to calculate the relevant weights. This paper also integrates the weights, and the relevant formula of the comprehensive weight of relevant statistics is as follows:

$$SC_i = \lim_{n \rightarrow \infty} \left(\sum_{j=1}^n \frac{(E-1) \cdot Cb_j}{E^j} + \frac{cb_n}{E^n} \right). \quad (9)$$

From the above formula, the calculated weight can continue the overall operation of the data. According to the above Hungarian algorithm optimized for the UAV internal system, the UAV has achieved good communication transmission performance in the practical application in the dynamic environment. In order to more intuitively see the specific situation of path planning of UAV in practical application, the flight trajectory data of UAV is transmitted to the

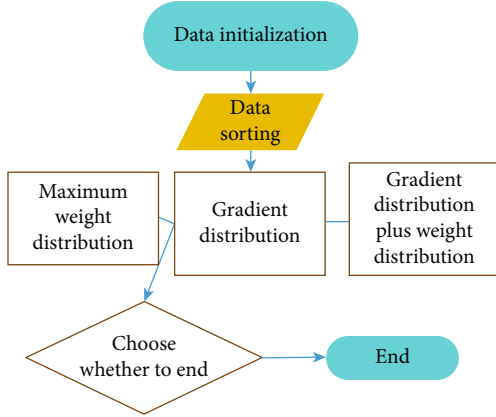


FIGURE 5: Specific process of implementing Hungarian algorithm in UAV.

computer, and the flight data trajectory of UAV is obtained, as shown in Figure 6.

It can be seen from Figure 6 that the data trajectory of the UAV flying without obstacles and the route selected by the UAV are the fastest routes to complete the task. After setting obstacles in the flight process, the UAV also automatically analyzes the data of obstacles and finally takes the most appropriate flight route for planning. From the feedback data trajectory diagram, the optimized and integrated Hungarian algorithm added to the UAV system can indeed improve the system performance of the UAV and greatly improve the communication and transmission ability of the UAV to data. In order to see the overall change of the UAV after adding the optimization algorithm in many aspects and angles, the data comparison of the total energy consumption generated by the UAV during the task execution is also carried out, as shown in Figure 7.

It can be seen from Figure 7 that the UAV without the optimization algorithm has a high energy consumption at the beginning of the task, and the energy consumption required has also reached a very high value due to the growth of working hours. The UAV with the optimized algorithm has halved the overall energy consumption compared with the UAV without the optimization algorithm during the flight mission, which can also enable the UAV to better complete the assigned tasks in a limited time.

The Hungarian algorithm and its optimization are described in detail in the above content, and the artificial potential field method is understood and optimized below. Artificial potential field is the most widely used algorithm for unmanned aerial vehicles, because the mathematical principle of artificial potential field method and its simple and easy to understand characteristics make the artificial potential field algorithm possible to change. However, the artificial potential field algorithm also has defects. The algorithm is weak in self-regulation, which is easy to minimize the processing of data information, so that the data information obtained by the UAV is wrong, and there is a misjudgment on the path selection and planning, which will

eventually cause serious losses. In order to optimize the algorithm, the internal formula for calculating the safety distance is improved. The formula is as follows:

$$F_{\text{rep}} \left(X_r \times \left(\frac{1}{(d(X_r, X_0))^2} - \frac{1}{(d_m - d_0)^2} \right) \right) d < d_m. \quad (10)$$

From the above formula, the safe flight distance of UAV can be accurately calculated, and then, the reasonable flight path can be analyzed and planned in the UAV system. In addition to calculating the safe distance, it is also necessary to add a cost function inside the algorithm so that the overall artificial potential field algorithm can give full play to the maximum performance of UAV path planning. The relevant formula is as follows:

$$J = \lim_{n \rightarrow \infty} \sum_{i,j=1}^n (w_k f_{ij} + w_d D_{ij}). \quad (11)$$

After using the optimized artificial potential field algorithm for the UAV, the actual task simulation is carried out, and the feedback path planning data is shown in Figure 8.

It can be seen from Figure 8 that the UAV is planning the route for different obstacles. By adding the optimized artificial potential field algorithm, the UAV can plan and design the path more quickly. The safety of UAV is ensured to a greater extent.

4. Results and Effect Analysis

4.1. Assignment of UAV Tasks and Threats in a Dynamic Environment. In this paper, a three-dimensional point cloud map of the environment is established by visual slam. Then, a two-dimensional mesh map is established by the three-dimensional point cloud of feature points proposed by SLAM algorithm. The height of each grid is calculated by projecting the map points of the graph into the corresponding grid. Then, an image segmentation algorithm based on mean shift is used to smooth the height of the mesh map, divide the obstacles and the ground, and combine the image blocks with similar height. The algorithm calculates the spatial distance between the landing area and the obstacle and selects the area farthest from the obstacle as the filtered landing area. In this way, the area suitable for UAV landing is selected. The UAV finally lands in a safe area according to the descent procedure.

This article further validates the ability of the new UAV to detect obstacles during mission completion. According to research on task and threat allocation of reinforcement learning UAVs in dynamic environments, the same tasks are assigned to UAVs. Three different mission environments were selected to simulate the UAV mission. In order to ensure the accuracy of the experiment, we repeat the operation for many times and finally take the average value to evaluate the experimental results. On the way of executing the mission, the UAV first receives the mission, assigns and processes the mission, and drives according to the

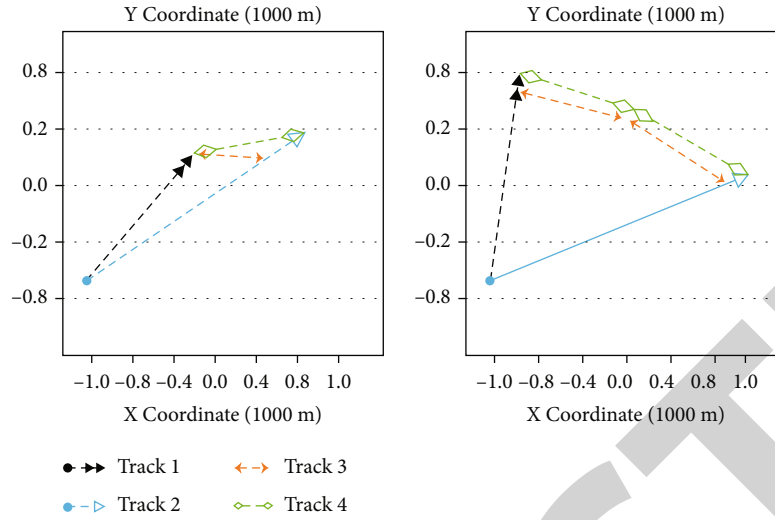


FIGURE 6: UAV flight data trajectory.

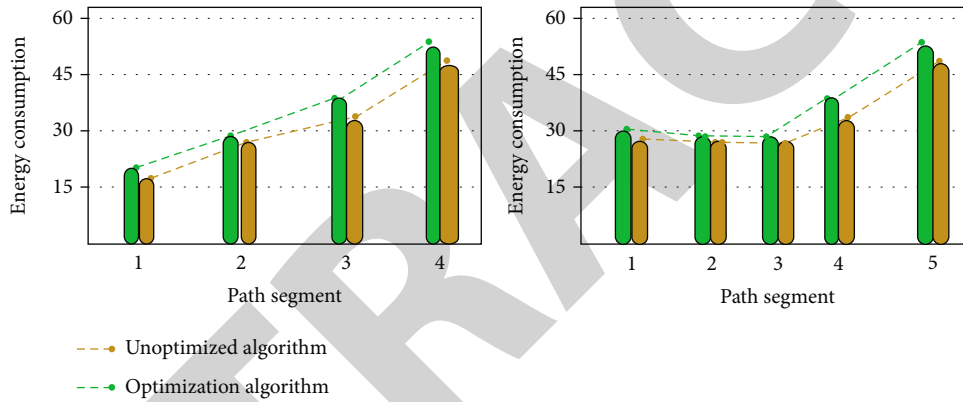


FIGURE 7: Comparison diagram of flight energy consumption during UAV operation.

selected flight path. Then, in the process of UAV flight, it monitors the possible obstacles, starts to judge whether the obstacle target is a threat to itself, and further makes action feedback. The purpose of this is that UAVs can save energy consumption. If obstacles are avoided, UAVs will often complete tasks significantly. After no one analyzes the target obstacles, it is necessary to focus on completing the tasks received within the system, analyze and process the tasks, and then complete the assigned tasks one by one. In the whole experiment, we mainly focus on whether the UAV can accurately perceive the existence of target obstacles under the simulated working state. The final experimental results are transmitted to the data formed by the computer, as shown in Figure 9.

It can be seen from Figure 9 that the data of obstacles fed back by UAV is different when it processes the same task in three different working environments. In the first environment, due to less obstacle model settings, the UAV has less changes to modify the original planned route when performing tasks, and the completion time is faster. In the third environment, although there are many target obstacle models, UAVs can accurately detect the existence of obstacles. To

sum up, the process of task and threat allocation of reinforcement learning UAV in dynamic environment studied in this paper is more suitable for practical application and has better detection performance.

4.2. UAV Path Planning and Optimization in Dynamic Environment. In the research of path planning and optimization of UAV based on reinforcement learning in dynamic environment, the problem of task path planning in UAV system is addressed. Firstly, a correlation function is added to the flight altitude of the UAV to make the UAV more invisible. Then, an optimized Hungarian algorithm is added to the UAV system to make the data communication flow faster within the system. Finally, an improved artificial potential field algorithm is added to the UAV system, which also accelerates the ability of calculating data in the system. In order to further verify the research results and practical application effects of UAV path planning and optimization, three kinds of UAVs—ordinary UAV, nonoptimized UAV, and optimized UAV—are compared for system efficiency. Considering the accuracy of the experimental results, the three UAVs are sent to the same task to test the system

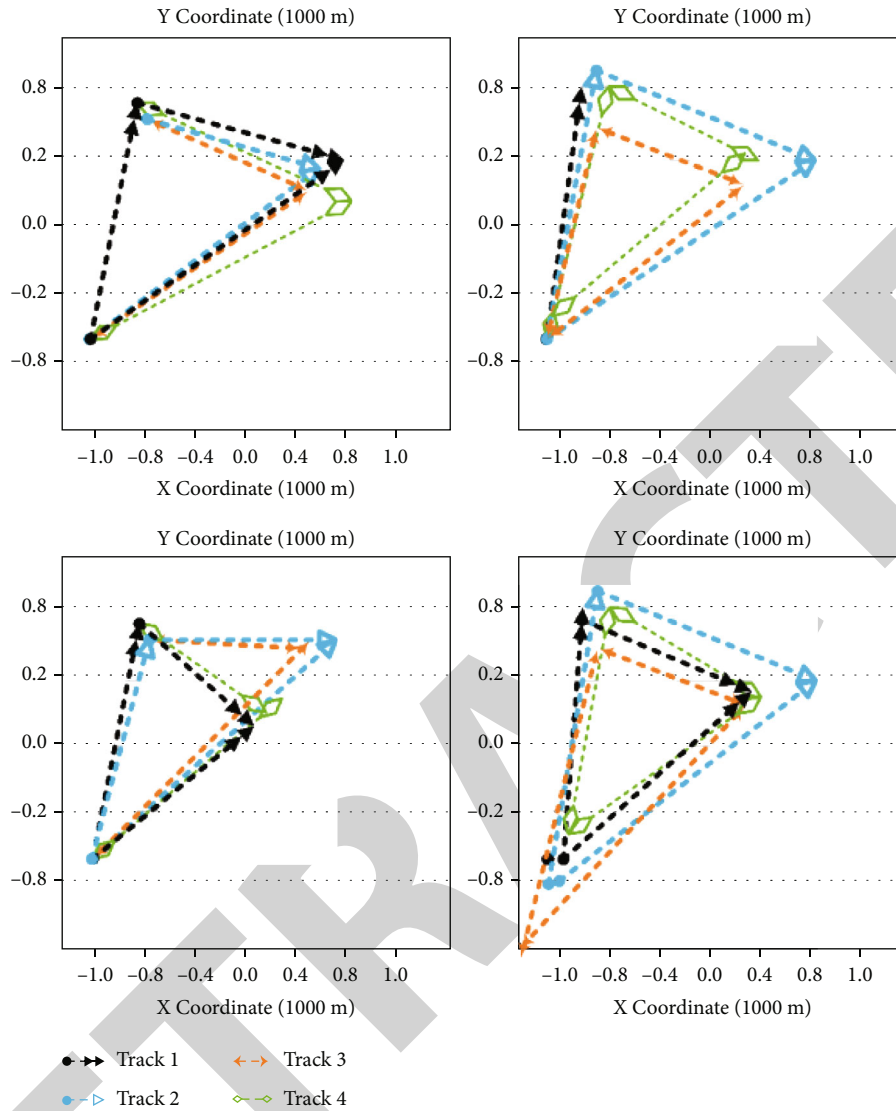


FIGURE 8: Path planning data diagram of actual task simulation feedback.

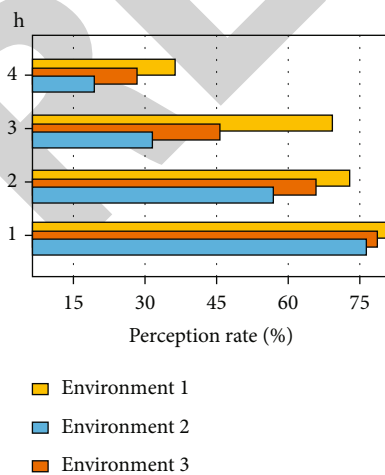


FIGURE 9: UAV perceived target obstacle ratio diagram.

performance under the same environmental state. However, due to the relatively backward system performance of ordinary UAV, less environmental obstacles are selected in this experiment, which is mainly based on the fluctuation of data processing wavelength in the system. The fluctuation amplitude of the internal data processing and calculation of the UAV system generated during the experimental test in this paper is shown in Figure 10.

It can be seen from Figure 10 that the results of the research on reinforcement learning UAV path planning and optimization in a dynamic environment are the data processing status of three types of UAVs. The optimized UAV has the fastest speed in the process of data transmission and calculation, while the data transmission and calculation express of the other two types of UAVs are slow. In particular, ordinary UAVs also have data transmission failures, resulting in incomplete experimental results. The results show that the optimized UAV system is better at data calculation and transmission, which also shows that the

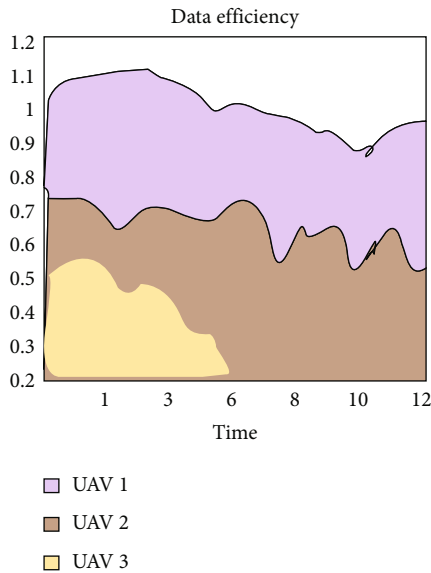


FIGURE 10: Fluctuation range of data processing and calculation in UAV system.

UAV can quickly plan and process the flight path, which is suitable for practical application.

This paper introduces the background of using reinforcement learning algorithm and points out that reinforcement learning algorithm is suitable for grid modeling. At the same time, the main parameters to measure the scene are given, such as the definition of state and grid coordinates. All of these lay a theoretical foundation for the introduction of subsequent algorithms, further verify the detection ability of the new UAV to detect obstacles in the process of completing tasks, and study the assignment of UAV tasks and threats based on reinforcement learning in a dynamic environment. Assigning the same task to the UAV and selecting three different task environments for the UAV task simulation are in order to ensure the accuracy of the experiment.

5. Conclusion

This paper studies the problem of UAV mission path planning based on reinforcement learning in dynamic environment and further verifies the detection ability of the new UAV to detect obstacles in the process of completing tasks. According to reinforcement learning in dynamic environment, the task and threat assignment of UAV are studied. The optimized Hungarian algorithm is added to the UAV system to make the data communication flow within the system faster. Finally, the improved artificial potential field algorithm is added to the UAV system, which also accelerates the ability of computing data in the system. The system efficiency of three kinds of UAVs—ordinary UAV, nonoptimized UAV, and optimized UAV—is compared. The optimized UAV has the fastest data transmission and calculation speed, while the other two types of UAVs have slower data transmission and calculation speed. In particular, ordinary UAVs also have data transmission failures, resulting in incomplete experimental results. The results

show that the optimized UAV system is better in data calculation and transmission, which also shows that the UAV can quickly plan and process flight paths, which is suitable for practical applications. However, there are still many shortcomings. For example, in a dynamic environment, the UAV encounters too many obstacles at the same time. The data collected by the UAV system will be mixed together, resulting in disorder of the internal system. Solving this situation is still a big problem. Further analysis is needed in future research and analysis.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by a fund project for basic scientific research expenses of central universities (nos. J2022-024 and J2022-07) and an independent research project of Key Laboratory of Flight Techniques and Flight Safety (no. FZ2021ZZ04).

References

- [1] L. Zhang and N. Ansari, "Optimizing the operation cost for UAV-aided mobile edge computing," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 6, pp. 6085–6093, 2021.
- [2] Y. Choi, Y. Choi, S. Briceno, and D. N. Mavris, "Energy-constrained multi-UAV coverage path planning for an aerial imagery mission using column generation," *Journal of Intelligent & Robotic Systems*, vol. 97, no. 1, pp. 125–139, 2020.
- [3] B. López, J. Muñoz, F. Quevedo, C. A. Monje, S. Garrido, and L. E. Moreno, "Path planning and collision risk management strategy for multi-UAV systems in 3D environments," *Sensors*, vol. 21, no. 13, p. 4414, 2021.
- [4] T. M. Cabreira, C. Di Franco, P. R. Ferreira, and G. C. Buttazzo, "Energy-aware spiral coverage path planning for UAV photogrammetric applications," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3662–3668, 2018.
- [5] F. Causa, G. Fasano, and M. Grassi, "Multi-UAV path planning for autonomous missions in mixed GNSS coverage scenarios," *Sensors*, vol. 18, no. 12, p. 4188, 2018.
- [6] L. Huo, J. Zhu, G. Wu, and Z. Li, "A novel simulated annealing based strategy for balanced UAV task assignment and path planning," *Sensors*, vol. 20, no. 17, p. 4769, 2020.
- [7] J. Kok, L. F. Gonzalez, and N. Kelson, "FPGA implementation of an evolutionary algorithm for autonomous unmanned aerial vehicle on-board path planning," *IEEE Transactions on Evolutionary Computation*, vol. 17, no. 2, pp. 272–281, 2013.
- [8] M. Popović, T. Vidal-Calleja, G. Hitz et al., "An informative path planning framework for UAV-based terrain monitoring," *Autonomous Robots*, vol. 44, no. 6, pp. 889–911, 2020.

- [9] G. Skorobogatov, C. Barrado, E. Salami, and E. Pastor, "Flight planning in multi-unmanned aerial vehicle systems: nonconvex polygon area decomposition and trajectory assignment," *International Journal of Advanced Robotic Systems*, vol. 18, no. 1, 2021.
- [10] P. O. Pettersson and P. Doherty, "Probabilistic roadmap based path planning for an autonomous unmanned helicopter," *Journal of Intelligent and Fuzzy Systems*, vol. 17, no. 4, pp. 395–405, 2006.
- [11] J. Zhang and H. Huang, "Occlusion-aware UAV path planning for reconnaissance and surveillance," *Drones*, vol. 5, no. 3, p. 98, 2021.
- [12] Z. Huang, C. Chen, and M. Pan, "Multiobjective UAV path planning for emergency information collection and transmission," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 6993–7009, 2020.
- [13] S. Hayat, E. Yanmaz, C. Bettstetter, and T. X. Brown, "Multi-objective drone path planning for search and rescue with quality-of-service requirements," *Autonomous Robots*, vol. 44, no. 7, pp. 1183–1198, 2020.
- [14] D. C. Guastella, L. Cantelli, G. Giammello, C. D. Melita, G. Spatino, and G. Muscato, "Complete coverage path planning for aerial vehicle flocks deployed in outdoor environments," *Computers & Electrical Engineering*, vol. 75, pp. 189–201, 2019.
- [15] N. Shahid, M. Abrar, U. Ajmal, R. Masroor, S. Amjad, and M. Jeelani, "Path planning in unmanned aerial vehicles: an optimistic overview," *International Journal of Communication Systems*, vol. 35, no. 6, article e5090, 2022.
- [16] R. Battulwar, G. Winkelmaier, J. Valencia, M. Zaré, and J. Sattarvand, "A practical methodology for generating high-resolution 3D models of open-pit slopes using UAVs: flight path planning and optimization," *Remote Sensing*, vol. 12, no. 14, article 2283, 2020.
- [17] I. Z. Biundini, M. F. Pinto, A. G. Melo, A. Marcato, and M. Aguiar, "A framework for coverage path planning optimization based on point cloud for structural inspection," *Sensors*, vol. 21, no. 2, p. 570, 2021.
- [18] S. Saha, A. E. Vasegaard, I. Nielsen, A. Hapka, and H. Budzisz, "UAVs path planning under a bi-objective optimization framework for smart cities," *Electronics*, vol. 10, no. 10, p. 1193, 2021.
- [19] Y. Guo, X. Liu, X. Liu, Y. Yang, and W. Zhang, "FC-RRT*: an improved path planning algorithm for UAV in 3D complex environment," *ISPRS International Journal of Geo-Information*, vol. 11, no. 2, p. 112, 2022.
- [20] S. H. Kim, G. E. G. Padilla, K. J. Kim, and K. H. Yu, "Flight path planning for a solar powered UAV in wind fields using direct collocation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 2, pp. 1094–1105, 2020.