A Design of Simulation Environment for Small Fixed-wing Aircraft

Zi-quan WANG, Cheng-wei YANG, Xiao-lin HU, Jing XIONG, Juan LI* and Chang LIU

Beijing Institute of Technology, Beijing, China

*juanli@bit.edu.cn

Abstract. Multi-Agent Particle Environment (MPE) [1] proposed by OpenAI is applied to the study of multi-agent reinforcement learning strategies. However, the motion rules of the agent are excessively simplified. In order to make the environment more suitable to small fixed-wing aircraft, we have made following improvements: 1. The dynamic model of the agent in the MPE does not conform to the characteristics of the fixed-wing aircraft. In order to simulate the dynamic characteristics of the fixed-wing aircraft, a speed-related damping mechanism is introduced into the two-dimensional motion environment. 2. Since the MPE lacks the control module for single agent, the MPE cannot meet the challenges raised by single agent control. A two-layer controller is proposed which includes the outer layer (Total Energy Control System and L_1) and the inner layer (PID). 3. The MPE does not contain any decision module. In order to comprehensively study the collaborative decision-making behavior of aircrafts in target allocation, a swarm decision module is added to the environment. In addition, the concept of control period is introduced to reduce the gap between simulation and the actual situation. Finally, several simulations were carried out to test the improved Multi-Agent Aircraft Environment (MAE). The test cases include the outer layer with L1 and Total Energy Control System (TECS) algorithm in two dimensions, the PID inner layer control algorithm and the designed auction algorithm. The tests complete the process of single aircraft flight, Multiple aircrafts scan-search flight and Multiple aircrafts dynamical-waypoint flight, which verifies the effectiveness of MAE.

1. Introduction
Multiple unmanned aerial vehicles (UAVs) collaborative reconnoitring, distributed mission planning and collaborative target attacking have become hotspot technologies to study how to improve the autonomy and combat efficiency of UAV.

The research on multi-UAV swarm-decision is based on the development of single UAV decision which shares some common characteristics [2]. (1) Task confirmation, including task types and task areas; (2) Battlefield situation perception, including threats of environment and targets, situation prediction, etc.; (3) Task planning, according to the constraints of specific tasks as well as the assessment data, calculate the distribution scheme and flight routes of the UAV swarm to achieve maximum benefit with minimum cost. In addition, multi-UAV swarm is essentially different from the single UAV in that they have the ability to coordinate with each other [3].

The control system of the UAV, that is, the autopilot, the design of which is a relatively complicated systematic work. It is based on the theory of automatic control which combines the overall design, navigation and aerodynamics of the aircraft [4]. The current common open source flight control
platforms are APM, PX4/Pixhawk, Autoquad, etc. [5] The design and development of the controller focuses on the optimization of the parameters and the verification of the simulation. At present, the research on reconnoitring (path planning), decision-making (target allocation) and control (navigation and guidance) for small aircraft is relatively independent [6]. However, there is no effective means for simulating the complete process. For example, in the path-planning and decision-making research for multi-agents, most of the simulation environment is based on discrete two-dimensional space. The simulation confidence of speed and motion direction is quite low since the agent can only move according to the grid. The Multi-Agent Particle Environment proposed by OpenAI is utilized to study the application of reinforcement learning in multi-agent coordinated motion control. The MPE is able to simulate the cooperation, competition and obstacle avoidance of multi-agent in two-dimensional continuous space. This has a great reference significance for the study of multi-UAV swarm reconnoitring, decision-making and control. However, the motion and control model of the agent in the MPE is not suitable for small fixed-wing aircraft. Therefore, the conclusions obtained based on the MPE are often not able to directly applied to the actual flight. This is inseparable from the important constraints such as motion control in the simulation process.

In the view of above problems, this paper makes some modifications to the MPE and we get the Multi-Agent Aircraft Environment (MAE). The MAE has the following characteristics:

1. A continuous aircraft dynamics model in a two-dimensional environment;
2. A simplified two-layer aircraft controller;
3. A time-window-based swarm decision module.

Based on the above improvement, the modified environment has the ability to continuously simulate the path-planning and decision-making process of multiple aircrafts. The environment with multiple constraints has higher simulation adaptability to small fixed-wing aircrafts. Therefore, the simulation results are closer to the actual application.

The rest of this paper is organized as follows. Section 2 formulates the framework of the MAE and describes the main functions of each module. Section 3 unveils the implementation of the two-layer control model. Section 4 presents three examples to verify the effectiveness of the MAE. Section 5 gathers our conclusions and ideas for future work.

2. Framework of MAE

Figure 1 shows the overall framework of the simulation environment. All aircrafts have the same dynamic model (environment). Each aircraft has an independent swarm-decision module, a path-planning module and a two-layer control module – inner and outer layer control.

![Figure 1. Framework of the simulation environment.](image)
2.1 Aircraft dynamic model
Since the six-degree-of-freedom dynamic model of the aircraft is computationally complex, the model needs to be simplified to the three-degree-of-freedom in two-dimensional space. The external force applied to the aircraft is

\[ \sum F = T - D \]  
\[ D = -cV^2 \]

Where \( T \) is the thrust of the aircraft and \( D \) means the simplified resistance, which is related to the speed \( V \) and the damping coefficient of the environment \( c \). Given above, the motion equation of the aircraft under the external force is given as

\[ \sum F = \frac{d}{dt} (mV) \]

Where \( m \) is the mass and \( V \) is the speed vector of the aircraft. Decompose the speed along the lateral and tangential directions in the body coordinate

\[ V = iu + jv \]

Where \( i \) and \( j \) are the unit vectors of the x- and y-axis of the body coordinate respectively. Decompose the external force \( \sum F \) into the body coordinate

\[ \sum F = iX + jY \]

Substitute equation (3) with (4) and (5), we can have the following equations

\[
\begin{align*}
X &= m(\dot{u} - vr) \\
Y &= m(\dot{v} + ur)
\end{align*}
\]

Where \( r \) is the heading angular speed in the two-dimensional space. Equation (6) can be rewritten as following force equations

\[
\begin{align*}
\dot{u} &= vr + \frac{F_x}{m} \\
\dot{v} &= -ur + \frac{F_y}{m}
\end{align*}
\]

Equation (7) is the dynamic model of the aircraft in the two-dimensional continuous environment.

2.2 Swarm-decision module
The swarm-decision module takes in the position and speed information of the aircraft output by the dynamic model and share its own state information with other aircrafts in the neighborhood to simulate the information interaction between real aircrafts. The concept of time window is introduced, and the aircraft will output different decision information in different control periods. As shown in Figure 2, all aircrafts will perform the searching process in the early stage. When certain conditions are met, all aircrafts will enter the next time window to perform target allocation. Different allocation algorithms will be used to obtain different results. Aircrafts assigned to the target will enter the next time window and begin to approach the target. Aircrafts that are not assigned to the target will return to the previous time window and continue the search process.

![Figure 2. Control period with time window.](image-url)
2.3 Path-planning module
The autopilot in the real fixed-wing aircraft will fly along the set waypoints in the automatic mode. Therefore, the path-planning module selects or generates the waypoint coordinates according to the decision information output by the decision module. Then the coordinates will be output to the control module.

2.4 Two-layer control module
The control module works as the autopilot. According to the coordinates of desired waypoint, it calculates the controlled quantity, i.e., the tangential and lateral acceleration. By changing the attitude of the aircraft, the heading angle is changed in a two-dimensional continuous space, and the aircraft is constantly guided to move the desired waypoint during the iterative process.

3. Two-Layer Control Module
As shown in Figure 3, the outer layer of the module utilizes simplified Total Energy Control System (TECS) and L1 control. And the traditional PID control works as the inner layer control. The desired waypoint output by the path-planning module, the position and speed fed back by the aircraft dynamic model, will be input into the outer layer control model. The inner layer, PID control, model receives desired attitude from the outer layer and attitude fed back by the aircraft dynamic model to maintain a stable attitude of the aircraft.

3.1 Total energy control system
The TECS is proposed by Boeing Company of the United States [7, 8]. The core idea is to manipulate the altitude and speed of the aircraft by coordinating the throttle and actuator. Throttle controls the rate of change of the total energy and actuator coordinates the conversion between potential energy and kinetic energy. In three-dimensional space, the total energy of the aircraft can be expressed as

$$E_T = E_D + E_s = \frac{1}{2} m V^2 + m g h$$  \hspace{1cm} (8)

Where, $E_T$, $E_D$ and $E_s$ are respectively total energy, kinetic energy and potential energy. In two-dimensional continuous space, the height $h$ is not taken into consideration. So, the equation (8) can be simplified as

$$E_T = E_D = \frac{1}{2} m V^2$$  \hspace{1cm} (9)

The rate of change of total energy can be obtained by differentiating equation (9)

$$\dot{E}_T = m V \dot{V}$$  \hspace{1cm} (10)

Furthermore, the rate of change of per unit weight is

$$\dot{E} = \frac{\dot{V}}{g}$$  \hspace{1cm} (11)

According to the aircraft dynamic model

$$T - D = m g \cdot \frac{\dot{V}}{g}$$  \hspace{1cm} (12)

The change in thrust $T$ is the main cause of the total energy change. The effectiveness of thrust is
\[ \Delta T = mg \cdot \frac{\dot{V}}{g} = m\dot{V} \]  

(13)

According to equation (11) and (13)

\[ \Delta T \propto \dot{E} \]  

(14)

It can be seen from equation (14) that the thrust is proportional to the total energy of the aircraft. Therefore, the control law of the total energy change can be described as

\[ T_c = K\dot{E} \]  

(15)

Where, \( K \) is controlling coefficient.

3.2 L1 control

The L1 nonlinear control method is first proposed by Sabhyuk Park and John Deyst [9]. Compared with traditional classical control method, the L1 nonlinear control method has better performance on the lateral offset control and anti-disturbance of curve tracking.

The lateral acceleration is obtained due to a reference point which is selected on the desired trajectory. As shown in Figure 4, the distance between reference point and the aircraft is L1. The lateral acceleration can be presented as

\[ a_{s\text{cmd}} = 2\frac{V^2}{L_1}\sin\eta \]  

(16)

Figure 4. Diagram for guidance logic.

Equation (16) has two crucial characteristics

(1) The direction of the lateral acceleration depends on the angle between L1 and the speed vector. For example, if the selected reference point is to the right of speed vector, the lateral acceleration will force the aircraft to correct to the tight, as shown in the figure above. In other words, the result of L1 control will cause the speed direction to gradually coincide with the direction of L1.

(2) At any time, an arc of radius of R can be determined by the reference point and the position of the aircraft, which is tangent to the speed direction. The lateral acceleration calculated by equation (16) is equal to the centripetal acceleration of the arc.

4. Simulation Analysis

4.1 Single aircraft flight

The initial position of the aircraft is located at a small random field in the lower left corner. The aircraft will pass through four waypoints, (-0.9, -0.9), (-0.9, 0.9), (0.9, 0.9) and (0.9, -0.9), in turn and return to the first one when finished. Three cycles are performed. As shown in Figure 5, it can be seen that during
the process of the reaching the waypoint and steering, the PID control will overshoot and gradually stabilize. When flying stably, the trajectory fits well with the desired track.

![Figure 5](image)

**Figure 5.** Flight trajectory of single aircraft. The blue dotted line presents the trajectory of the aircraft while the red one solid one is the desired trajectory.

### 4.2 Multiple aircrafts scan-search flight

The process of scan-search with 10 aircrafts are conducted. In the initialization phase of the simulation, specific waypoint sequence is generated for each aircraft according to the characteristics of the scan-search tracks. And each aircraft sequentially flies by its waypoint sequences. Figure 6(a) shows the desired tracks while Figure 6(b) shows the actual searching trajectory. The trajectories are well fitted even though some tracks have large overshoots.

![Figure 6](image)

(a) Desired trajectories.  
(b) Actual trajectories.

**Figure 6.** Scan-search trajectories with 10 aircrafts. Lines of different types and colors represent the trajectories of different aircrafts.

### 4.3 Multiple aircrafts dynamical-waypoint flight

The process of area-searching with 10 aircrafts are conducted whose waypoint is dynamically generated. Only one waypoint is generated during the initialization phase of the simulation. When the aircraft arrives at the waypoint, the next waypoint is generated by the swarm-decision module. Policy of how the waypoint is generated is as follows:

- During the initialization phase, each aircraft randomly generates a waypoint on one of the four boundaries.
- When arriving at a waypoint, one of the other three edges is randomly selected and a new point is randomly selected on the chosen boundary as the next waypoint.
Figure 7 shows the trajectories of 10 aircrafts using the policy mentioned upon when the simulation reaches 1000, 2000 and 4000 steps.

(a) 1000 steps. (b) 2000 steps. (c) 4000 steps.

Figure 7. Trajectories of 10 aircrafts at different steps.

In term of application, it is necessary to utilize multiple aircrafts to perform a fast and effective searching mission for a certain area. Searching coverage is an important index to evaluate the searching process. How to achieve high searching coverage in a short time is a promising research. Figure 8 shows the change of coverage as a function of simulation step using two different decision policies. The red curve is the result of reinforcement learning for multi-aircraft using a Deep Q Network (DQN). And the green curve is the result of random searching policy used in Figure 7. It can be seen that in the early stage of the search, the policy generated by the reinforcement learning show a better performance.

Figure 8. Change of coverage with two different policies.

5. Conclusion
In this paper, the Multi-Agent Particle Environment proposed by OpenAI is improved with three modules added – the aircraft dynamics model with damping in a two-dimensional environment, a two-layer control module, and a time-window-based swarm-decision module. Numerical experiments on the improved environment, Multi-Agent Aircraft Environment (MAE), confirm that the MAE is more suitable for the simulation of small fixed-wing aircrafts.

Future work will include extensions to consideration of flight height, control models of roll and pitch angle, and target categories such as composite formation of stationary and movable targets. So the MAE would be able to meet more simulation requirements.

Reference
[1] Lowe, R., Wu, Y., Tamar, A., Harb, J., Abbeel, O. P., and Mordatch, I. 2017. Multi-agent actor-critic for mixed cooperative-competitive environments. In Advances in Neural Information Processing Systems, 6379-6390.
[2] YANG Chen, ZHANG Shao-Qing, MENG Guang-Lei. 2018. Multi-UAV cooperative mission planning. In Journal of Command and Control, 234-248.

[3] MA Chun-Chao. 2014. Unmanned Aerial Vehicle (UAV) Task Coordination Method Study Under the Network Environment. National University of Defense Technology.

[4] Hamke, E. E., Ihlein, J. 2010 UAV Flight Control Method And System. U.S. Patent Application 12/396,653.

[5] Paw, Y. C., and Balas, G. J. 2011. Development and application of an integrated framework for small UAV flight control development. Mechatronics, 21(5), 789-802.

[6] Ilaya, O., Bil, C., and Evans, M. 2007. Distributed and Cooperative Decision Making for Multi-UAV Systems with Applications to Collaborative Electronic Warfare. In 7th AIAA ATIO Conf, 2nd CEIAT Int’l Conf on Innov and Integr in Aero Sciences, 17th LTA Systems Tech Conf; followed by 2nd TEOS Forum, 7885.

[7] Kelly, J. R., Person, L. H., and Bruce, K. R. 1987. Flight testing TECS – The TOTAL ENERGY CONTROL SYSTEM. Society of Automotive Engineers, 169-176.

[8] Kaminer, I., O’SHAUGHNESSY, P. A. T. R. I. C. K. 1989. 4D-TECS integration for NASA TCV airplane. In Guidance, Navigation and Control Conference, 4067.

[9] Park, S., Deyst, J., and How, J. 2004. A new nonlinear guidance logic for trajectory tracking. In AIAA guidance, navigation, and control conference and exhibit. 4900.