Infrared Air Combat Simulation Model for Deep Reinforcement Learning

Conghui Huang
College of Aeronautical Engineering, Air Force Engineering University, Xi'an 710038, China
hch271@126.com

Chaozhe Wang*
College of Aeronautical Engineering, Air Force Engineering University, Xi'an 710038, China
wcz667@sina.com

Qi Tong
College of Equipment Management and UAV Engineering, Air Force Engineering University, Xi'an 710100, China
tongqi_111@163.com

ABSTRACT

Aiming at the problem of lacking credible and realistic infrared air combat simulation platform for applying the deep reinforcement learning method, this paper explored the design requirements for the construction of simulation system, built the overall architecture of infrared air combat simulation system, described the structure, principle and working process of the fighter jet, infrared air-to-air missile, point source decoy and environment model. The implementation method of the simulation system was given, and the credibility of the system was verified through attack and defense simulation examples and error analysis of missile anti-jamming probability, which indicated that the simulation system can be used for the training and testing of agents based on deep reinforcement learning in infrared air combat scenarios.

CCS CONCEPTS

• **Computing methodologies**; • **Modeling and simulation**; • **Model development and analysis**; • **Model verification and validation**;

KEYWORDS

Air combat simulation, Deep reinforcement learning, Agent

ACM Reference Format:
Conghui Huang, Chaozhe Wang, and Qi Tong. 2021. Infrared Air Combat Simulation Model for Deep Reinforcement Learning. In *The 5th International Conference on Computer Science and Application Engineering (CSAE 2021), October 19–21, 2021, Sanya, China.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3487075.3487094

1 INTRODUCTION

In August 2019, the artificial intelligence program AlphaStar based on deep reinforcement learning and developed by DeepMind made a breakthrough in the real-time strategy game StarCraft 2, reaching a master level in the use of three races in the game, surpassing 99.8% of the human players on the official website [1]. This fully reflects the great potential of deep reinforcement learning to solve the problem of incomplete information game. In modern air combat, pilots of both sides use infrared missiles and infrared decoys to attack and defend in close combat, that is essentially an incomplete information game problem. Therefore, deep reinforcement learning technology is expected to enable UAVs to have air combat capabilities, so that the UAV can obtain more and higher-level autonomous control capabilities.

Training an agent based on deep reinforcement learning to control UAV for air combat requires a lot of training and testing, that requires the establishment of a credible and realistic air combat simulation system to save cost and time. This paper first analyzed the general requirements for building an infrared air combat simulation system for close combat, then gave the overall architecture of the simulation system, then described the modeling of the subsystems in the simulation system in detail, and finally elaborated the specific implementation and simulation of the system instance. The simulation system will provide a basic platform for the application of deep reinforcement learning in the autonomous control of UAVs.

2 GENERAL REQUIREMENTS FOR SIMULATION SYSTEM CONSTRUCTION

The fidelity of the simulation system affects the difficulty of the migration of the agent based on deep reinforcement learning to the real world. However, the more realistic the simulation model, the greater the difficulty of construction and the lower the efficiency of the simulation. Therefore, the construction of an infrared air combat simulation system is not accomplished overnight. It requires long-term efforts, some basic requirements should be met at the beginning of the simulation system design, including:

(1) Adopt model-based design ideas and top-down design methods, follow the gradual refinement and step-by-step design principles, build the architecture of the simulation system, describe the structure and behavior of the simulation model, and continuously improve the fidelity of the simulation model .
(2) The structure and behavior of the simulation model, the definition of the interface between the modules, and the relationship of information transmission should be consistent with the actual product, which will be convenient for understanding the design of the simulation model, replacing and updating the sub-modules of the simulation model, and also conducive to analyzing the fidelity of simulation model.
(3) The simulation model should be scalable, which can easily increase or decrease the fidelity of the simulation model, so as to compromise between model simulation efficiency

CSAE 2021, October 19–21, 2021, Sanya, China

ACM ISBN 978-1-4503-8985-3/21/10…$15.00
https://doi.org/10.1145/3487075.3487094

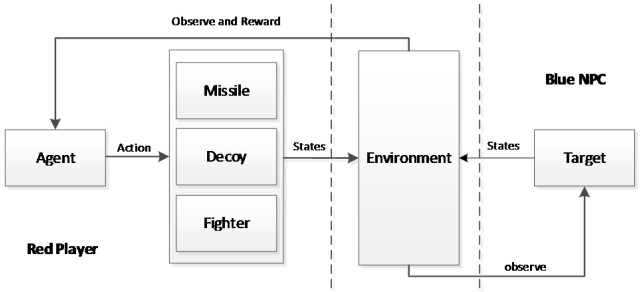

Figure 1: The Overall Architecture of the Infrared Air Combat Simulation System.

and fidelity, and at the same time examine the impact of simulation model fidelity on the deep reinforcement learning algorithm.

(4) The simulation system should be extensible, and different simulation models can be easily added to expand the functions of the simulation system and realize the air combat simulation of different combat backgrounds in the complex battlefield environment.

3 OVERALL ARCHITECTURE OF THE SIMULATION SYSTEM

The infrared air combat simulation system is mainly composed of jet fighter model, infrared air-to-air missile model, Infrared decoy model and other models. According to the general requirements of simulation system construction, Simulink technology is used to design and realize the system. The overall architecture of the infrared air combat simulation system is shown in Figure 1. It is divided into three parts: the red control model, the blue target model and the environment model. The specific structure is as follows:

(1) Control model. The red side control model is mainly composed of the agent and the controlled object which contains fighter, missile, and decoy. The agent will use the deep reinforcement learning method for training, receive observations and reward feedback provided by the environment model, and then operate the controlled object according to its own policy. The controlled object outputs its state to the environment model.

(2) Environmental model. The environment model receives the state input of the red and blue parties, and outputs the observation or reward feedback required by each. By extracting the state originally calculated in the red or blue model (such as the infrared radiation characteristics received by the air-to-air missile) into the environment model, the coupling between the red control model and the blue target model can be released, which is beneficial to enhance the simulation system scalability.

(3) Target model. The blue target model consists of an air-to-air missile or a fighter model controlled by an agent. It is used to train and verify the fighter maneuver strategy, missile attack strategy and decoy jamming strategy of the red agent.

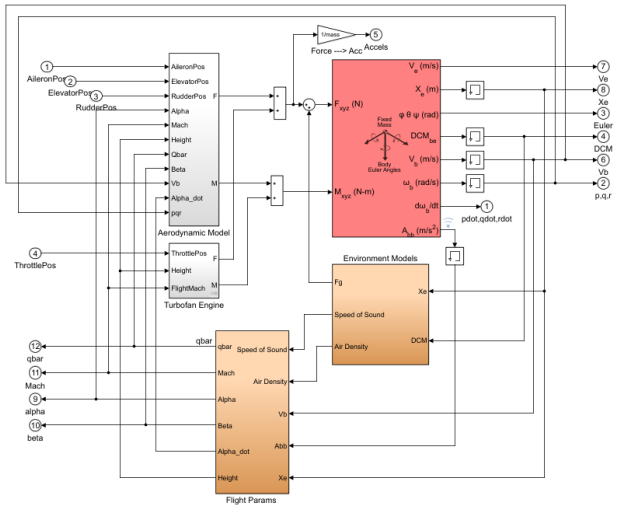

Figure 2: Overall Structure of Aircraft Simulation.

4 MODELING OF SIMULATION SYSTEM

According to the overall architecture of the simulation system, the typical combat equipment used in current close air combat is modeled, including jet fighters, pulse-modulated infrared air-to-air missiles, and point-source infrared decoys. At the same time, the target detection model and the real-time reward model for the agent's actions are established in the environment model.

4.1 Jet Fighter Modeling

The goal of jet fighter modeling is to establish a mathematical model with elevator deflection angle, aileron deflection angle, rudder deflection angle and throttle opening as control inputs, and fighter motion state as output. Considering the characteristics of close air combat in short duration, small coverage of airspace, and little influence of altitude on gravitational acceleration, the following assumptions were made during the research process:

(1) Regard the aircraft as a rigid body, ignoring the elastic deformation of the structure and the shaking of fuel.
(2) Regard the earth as a flat ground that does not rotate, ignoring the influence of the earth's rotation and curvature.
(3) Regardless of the influence of fuel consumption on aircraft weight, the aircraft weight during air combat is regarded as a constant.
(4) The installation angle of the aircraft engine is 0 degree.
(5) The acceleration of gravity is regarded as a constant during air combat, and the influence of altitude on the acceleration of gravity is not considered.
(6) The aircraft flight control system is ideal.

The overall structure of the aircraft simulation is shown in Figure 2. The following mainly describes the aircraft motion equation model, engine model and aerodynamic model.

4.1.1 Aircraft Motion Equation Model. According to Newton's second law, the dynamics and kinematics equations of the aircraft can

Figure 3: The Relationship between Throttle Lever Stroke and Engine State.

be obtained [2] as follows:

$$\begin{cases} \mathrm{m}\frac{dV}{dt} = P\cos\alpha\cos\beta - X - mg\sin\theta \\ mV\frac{d\theta}{dt} = P(\sin a\cos\gamma_V + \cos\alpha\sin\beta\sin\gamma_V) + Y\cos\gamma_V - Z\sin\gamma_V - mg\cos\theta \\ -mV\cos\theta\frac{d\varphi_V}{dt} = P(\sin a\sin\gamma_V + \cos\alpha\sin\beta\cos\gamma_V) + Y\sin\gamma_V + Z\cos\gamma_V \end{cases} \tag{1}$$

$$\begin{cases} J_x\frac{d\omega_x}{dt} + (J_z - J_y)\omega_z\omega_y = M_x \\ J_y\frac{d\omega_y}{dt} + (J_x - J_z)\omega_x\omega_z = M_y \\ J_z\frac{d\omega_z}{dt} + (J_y - J_x)\omega_y\omega_x = M_z \end{cases} \tag{2}$$

$$\begin{cases} \frac{dx}{dt} = V\cos\theta\cos\varphi_V \\ \frac{dy}{dt} = V\sin\theta \\ \frac{dz}{dt} = -V\cos\theta\sin\varphi_V \end{cases} \tag{3}$$

$$\begin{cases} \frac{d\vartheta}{dt} = \omega_y\sin\gamma + \omega_z\cos\gamma \\ \frac{d\varphi}{dt} = \frac{1}{\cos\vartheta}(\omega_y\cos\gamma - \omega_z\sin\gamma) \\ \frac{d\gamma}{dt} = \omega_x - \tan\vartheta(\omega_y\cos\gamma - \omega_z\sin\gamma) \end{cases} \tag{4}$$

$$\begin{cases} \sin\beta = \cos\theta[\cos\gamma\sin(\varphi-\varphi_V) + \sin\vartheta\sin\gamma\cos(\varphi-\varphi_V)] - \sin\theta\cos\vartheta\sin\gamma \\ \sin\alpha = \frac{\cos\theta[\sin\vartheta\cos\gamma\cos(\varphi-\varphi_V) - \sin\gamma\sin(\varphi-\varphi_V)] - \sin\theta\cos\vartheta\cos\gamma}{\cos\beta} \\ \sin\varphi_V = \frac{\cos\alpha\sin\beta\sin\vartheta - \sin\alpha\sin\beta\cos\gamma\cos\vartheta - \cos\beta\sin\gamma\cos\vartheta}{\cos\theta} \end{cases} \tag{5}$$

Among them, m is the aircraft mass, P is the engine thrust, X, Y, Z are respectively the aerodynamic drag, lift, and side force, g is the acceleration of gravity, V is the aircraft speed, α, β are respectively the angle of attack and sideslip, η is the pitch angle, φ is the yaw angle, γ is the roll angle, θ is the track inclination angle, φ_V is the track yaw angle, γ_V is the velocity inclination angle, J_x, J_y, J_z is the moment of inertia in the airframe coordinate system, $\omega_x, \omega_y, \omega_z$ are the angular velocities in the airframe coordinate system, and M_x, M_y, M_z is the moment in the airframe coordinate system.

In order to solve the above-mentioned aircraft motion state equation, it is also necessary to analyze the forces and moments received by the aircraft.

4.1.2 Engine Model. Take the typical turbofan engine used in the third-generation fighter as an example to establish the engine thrust model. The engine thrust characteristics are mainly affected by height, speed and high-pressure rotor speed [3], The model is as follows:

$$P = f_p(H, Ma, n_{\mathrm{H}}) \tag{6}$$

Among them, H is the altitude, Ma is the Mach number of the carrier, and n_H is the high-pressure rotor speed. The engine has a variety of states, such as stopping, idle, maximum, afterburning, etc. There are different steady-state rotor speeds under different conditions. The throttle lever stroke controls the engine rotor speed. After normalizing the throttle lever stroke, the simple corresponding relationship between it and the state of the engine is shown in Figure 3.

In the process of measuring and determining the thrust of the engine, the thrust at different heights and Mach numbers in the idle, maximum and afterburning states is usually measured to form a thrust characteristic look-up table. Assuming that the engine speed can reach the target speed immediately with the change of the throttle lever stroke, and the engine thrust is proportional to the speed, the engine thrust model is as follows:

$$\mathrm{P} = \begin{cases} 0, & 0 \le \delta_t < 1/4 \\ P_{idle}(H, Ma) + 4\left(\delta_t - 1/4\right) * \left[P_{mil}(H, Ma) - P_{idle}(H, Ma)\right], & 1/4 \le \delta_t \le 3/4 \\ P_{mil}(H, Ma) + 4\left(\delta_t - 3/4\right) * \left[P_{\max}(H, Ma) - P_{mil}(H, Ma)\right], & 3/4 < \delta_t \le 1 \end{cases} \tag{7}$$

Among them P_{idle},P_{mil},P_{max} are respectively the thrust interpolation functions of the engine in idle, maximum and afterburning states, and δ_t is the normalized throttle lever stroke.

4.1.3 Aerodynamic Model. When an airplane moves in the air, the aerodynamic force acting on the airplane is distributed along the body. The distribution and size of aerodynamic forces are usually related to the speed, air density, attitude and shape of the aircraft. In order to simplify the research on aerodynamic force, the aerodynamic force acting on the aircraft can be equivalent to the force and moment on a certain point on the aircraft. Generally, aerodynamic force can be decomposed into lift, drag and side force under the wind axis system, as well as the roll moment, pitch moment and yaw moment under the airframe coordinate system [4]. Both aerodynamic force and moment can be expressed as the product of force and force coefficient, expressed by the following formula:

$$\begin{cases} Y = \frac{1}{2}\rho V_a S[C_{L_0} + C_{L_\alpha}\alpha + C_{L_q}\frac{c}{2V_a}\omega_z + C_{L_{\delta_e}}\delta_e] \\ X = \frac{1}{2}\rho V_a S[C_{D_0} + C_{D_\alpha}\alpha + C_{D_q}\frac{c}{2V_a}\omega_z + C_{D_{\delta_e}}\delta_e] \\ M_z = \frac{1}{2}\rho V_a Sc[C_{m_0} + C_{m_\alpha}\alpha + C_{m_q}\frac{c}{2V_a}\omega_z + C_{m_{\delta_e}}\delta_e] \end{cases} \tag{8}$$

$$\begin{cases} Z = \frac{1}{2}\rho V_a S[C_{Y_0} + C_{Y_\beta}\beta + C_{Y_p}\frac{b}{2V_a}\omega_y + C_{Y_r}\frac{b}{2V_a}\omega_x + C_{Y_{\delta_a}}\delta_a + C_{Y_{\delta_r}}\delta_r] \\ M_x = \frac{1}{2}\rho V_a Sb[C_{l_0} + C_{l_\beta}\beta + C_{l_p}\frac{b}{2V_a}\omega_y + C_{l_r}\frac{b}{2V_a}\omega_x + C_{l_{\delta_a}}\delta_a + C_{l_{\delta_r}}\delta_r] \\ M_y = \frac{1}{2}\rho V_a Sb[C_{n_0} + C_{n_\beta}\beta + C_{n_p}\frac{b}{2V_a}\omega_y + C_{n_r}\frac{b}{2V_a}\omega_x + C_{n_{\delta_a}}\delta_a + C_{n_{\delta_r}}\delta_r] \end{cases} \tag{9}$$

Among them $C_L, C_D, C_{\mathrm{m}}, C_{\mathrm{Y}}, C_{\mathrm{l}}, C_{\mathrm{n}}$ are the aerodynamic coefficients, α, β are respectively the angle of attack and sideslip, $\omega_x, \omega_y, \omega_z$ are the angular velocities around the *x, y, z* axis of the airframe coordinate system, $\delta_{\mathrm{e}}, \delta_a, \delta_r$ are respectively the deflection angles of the elevator, aileron, and rudder, $X, Y, Z$are respectively the aerodynamic drag, lift, and side force, and M_x, M_y, M_z are respectively the roll moment, yaw moment, and pitch moment, b, $c$are respectively wingspan and average aerodynamic chord length, ρ is air density, V_a is sound speed, and S is windward area of the wing.

4.2 Modeling of Pulse Modulation Infrared Air-to-Air Missile

Infrared air-to-air missile is an automated weapon that takes the infrared radiation characteristics of the target as input and uses guidance and control law to attack the target. In order to establish the mathematical model of infrared air-to-air missile, the following assumptions are made in the research process [5]:

(1) Adopt the principle of solidification: that is, the flight speed, flight height, engine thrust, missile mass, and moment of inertia of the missile at a certain time on the ballistic trajectory are unchanged;

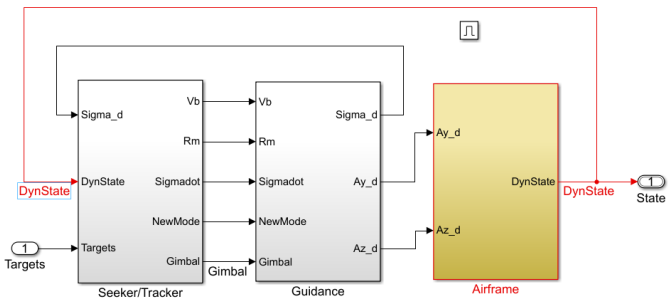

Figure 4: The Overall Structure of the Infrared Air-to-Air Missile Simulation.

(2) The missile adopts an axisymmetric layout design;
(3) When the missile is under control or interference, the parameters of the missile do not change much, and the angle of attack of the missile is small;
(4) The control system ensures the stability of the roll angle, and replies at sufficient speed.

After adopting the above assumptions, the longitudinal disturbance movement and the lateral disturbance movement of the missile can be studied separately. On this basis, the overall structure of the infrared air-to-air missile simulation is established as shown in Figure 4. The infrared air-to-air missile model is mainly composed of the seeker model, the guidance system model, and the missile airframe model. The missile airframe model is composed of the missile motion equation model and the autopilot. Since the missile motion equation model is similar to the aircraft motion equation model, it is no more details here.

4.2.1 Seeker Model. At present, the pulse modulated multi-element detectors widely used in the third generation of infrared air-to-air missiles have the advantages of accurate target detection, no blind spots and strong anti-interference ability [6]. Multi-element detectors are typically binary and quaternary, and their basic principles are the same. Taking the four-element detector as an example to establish an infrared seeker model, its input is the off-axis angle required in the search mode of the seeker, the movement state of the missile and the detected target, and its output is the seeker's new working mode, missile speed, the distance between the missile and the target, the rotational angular velocity of the line of sight, and the off-axis angle. The seeker model is mainly composed of anti-jamming recognition algorithm and target stable tracking model.

The flow of the anti-jamming recognition algorithm is shown in Figure 5. Its main goal is to identify the real target in the seeker's field of view, calculate the angle of view between the real target and the missile longitudinal axis, and record the radiation characteristics of the real target in the missile tracking mode according to the state of the missile and the seeker itself. The anti-jamming recognition algorithm will identify real target from multiple radiation sources through anti-jamming technologies such as position memory, amplitude memory, and field of view contraction [7].

The target stable tracking model is used to isolate the jitter of the seeker's optical axis in the inertial space caused by the attitude

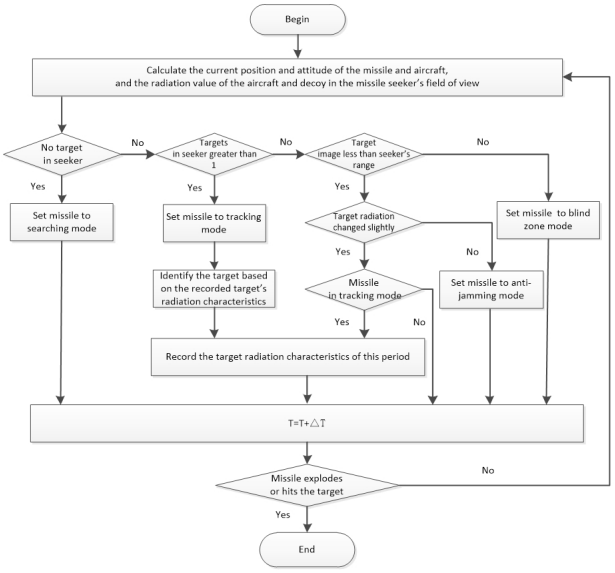

Figure 5: Seeker Anti-Jamming Recognition Algorithm.

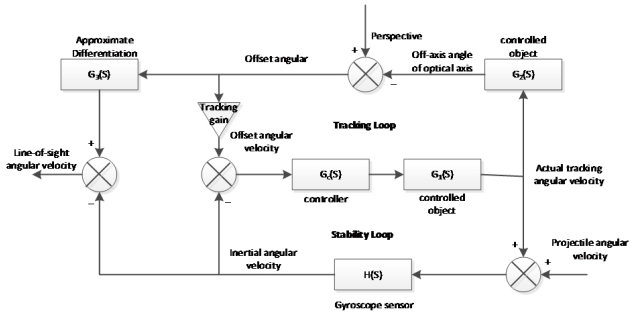

Figure 6: Seeker Target Stable Tracking Control Block Diagram.

movement of the missile body and other interference moments, and at the same time enable the seeker's optical axis to accurately track the target position command. Its typical control block diagram is shown as in Figure 6. In the missile tracking mode, the estimated line-of-sight rate will be the sum of the target tracking offset angular velocity and the inertial angular velocity measured by the gyroscope sensor.

4.2.2 Guidance System Model. According to the target information obtained by the seeker, the guidance system model generates a normal overload command according to a specific guidance law, so as to control the missile to approach and finally destroy the target. The commonly used guidance methods mainly include tracking method, parallel approach method and proportional guidance method. Because the proportional guidance method is simple in form, technically easy to implement, and does not require much information, it is often applied to the third generation of infrared air-to-air missiles. According to the proportional guidance method,

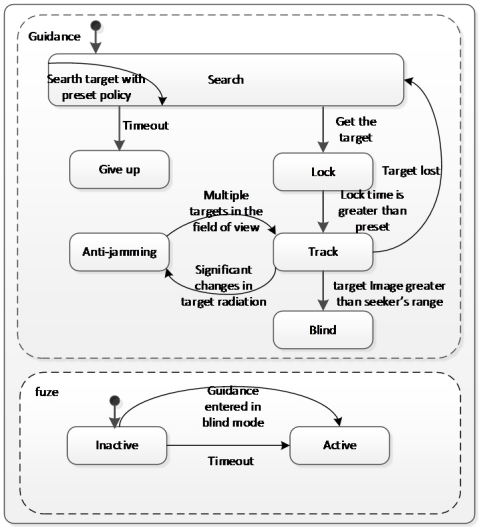

Figure 7: Guidance System Model Based on Finite State Machine.

the rotational angular velocity of the missile velocity is proportional to the rotational angular velocity of the line of sight [8], that is,

$$\dot{\sigma}_M = K\dot{q} \tag{10}$$

In the formula, $\dot{\sigma}_M$is the rotational angular velocity of the missile velocity, $\dot{q}$ is the rotational angular velocity of the line of sight, K is the proportional coefficient, usually 2∼6.

Infrared air-to-air missiles have multiple working modes and have different behaviors in different working modes. For example, in the tracking mode, the missile will approach the target according to the proportional guidance method; In the search mode, the seeker will be driven to search for the target in the specified direction according to the preset search strategy; In the final attack stage, the seeker will stop working and the missile will detonate the warhead according to the state of the proximity fuze to attack the target. According to these characteristics, the finite state machine theory is used to establish the guidance system model, as shown in Figure 7. The guidance system model consists of two parallel states of guidance and fuze. According to the working mode of the seeker, the guidance state is divided into multiple states. The same is true for the fuze state. According to the state of the guidance system model, the seeker and fuze will be driven to realize the output of guidance commands or detonate the warhead.

4.2.3 Autopilot Model. The missile autopilot controls the steering gear to deflect the rudder surface according to the guidance instructions output by the guidance system to make the missile fly to the target accurately, and finally destroy the target through the warhead. Usually the missile autopilot adopts a three-loop design [9], and its structural block diagram is shown in Figure 8.

It can be seen from Figure 8 that the three loops are the damping loop, the stability augment loop, and the overload loop from the inside to the outside. Among them K_{DC} is the missile autopilot closed-loop gain adjustment coefficient, K_a is the overload loop gain

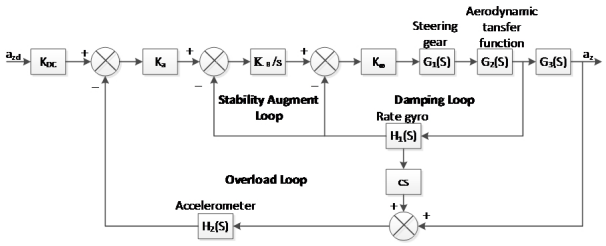

Figure 8: Three-Loop Autopilot Pitch Channel Structure Diagram.

adjustment coefficient, K_θ is the stabilization loop gain adjustment coefficient, and K_ω is the damping loop gain adjustment coefficient. By adjusting these four parameters, the three-loop autopilot can reach the design goal.

4.3 Point Source Infrared Decoy Modelling

Point source infrared decoy is a weapon against non-imaging infrared air-to-air missiles. It generates a large amount of heat in a short period of time to simulate the radiation of the engine's tail nozzle, thereby disturbing the infrared seeker. After the point source infrared decoy is ejected from the fighter, it is only affected by gravity and resistance [10], and ignites within 0.5 seconds after contacting with air. After ignition, the radiation intensity rises sharply, and then remains stable. The entire combustion process is 4 ∼6 seconds. Assuming that the decoy burns uniformly, the dynamic model of the point source infrared decoy is as follows:

$$\begin{cases} m\frac{dV}{dt} = mg\sin\phi - \frac{1}{2}C_x\rho V^2 S \\ mV\frac{d\phi}{dt} = mg\cos\phi \\ m_t = m_0 - Kt \end{cases} \tag{11}$$

Where ϕ is the angle between the speed direction and the horizontal plane (downwards is positive), C_x is the drag coefficient of the decoy, ρ is the air density, m_t is the mass of the decoy at time t, m_0 is the initial mass of the decoy, and K is the burning rate.

The infrared radiation characteristics of point source infrared decoy are isotropic. At different heights and speeds, the law of the radiation intensity changing with time can be measured experimentally and made into an interpolation table for use.

4.4 Environment Modeling

The environment model takes the states of fighters, missiles and decoys in the simulation system as input, and then gives corresponding observation output according to the characteristics of fighter and missile sensors. For fighter, the environment model will output the battlefield situation that the fighter can detect. For infrared air-to-air missiles, the environment model will output target information within the detection range of its seeker. In addition, in the training phase of applying the deep reinforcement learning algorithm, the environment model also needs to give instant reward to the action of the agent. Therefore, the environment model is mainly composed of a detection model and a reward model.

4.4.1 Probe Model. The radar detection system and missile approach warning system equipped with fighter have different detection distances, detection ranges, and perception information. Usually airborne radar can obtain the target's distance, azimuth and speed information, and the missile approach warning system informs the pilot of the missile attack. And the non-imaging infrared air-to-air missile seeker obtains the position and radiation value of the target in its field of view. The targets that can be detected by different types of detectors can be represented by sets:

$$T = \{x \in T | p(x)\} \tag{12}$$

Among them, T represents the target set, the element x in the set T should meet the condition $p(x)$, that is, the target x should be within the detection distance and range of the detector.

The infrared radiation of fighter is mainly composed of skin radiation, exhaust flame radiation and tail nozzle radiation, and is different in different directions, so it is necessary to establish an interpolation table of infrared radiation intensity with different azimuth angles for the fighter, which used in the infrared seeker detection model.

4.4.2 Instant Reward Model. The design of the instant reward model is one of the keys to the success of deep reinforcement learning algorithms [11]. In the training phase, the instant reward model provides real-time rewards and punishments for the action of the agent to improve the efficiency of the agent's policy search and ensure the convergence and stability of the policy search. The essence of the instant reward model is the visualization of prior knowledge. There are different instant reward model designs corresponding to different specific problems, so this paper will not discuss it here.

5 SYSTEM IMPLEMENTATION AND EXAMPLE VERIFICATION

Take a jet fighter and its near-infrared air-to-air missiles and point source decoys as examples for simulation. Through attack and defense simulation examples, test whether the fighter, missile, and decoy models conform to the design, and verify the credibility of the simulation results of the infrared air combat system through statistics and analysis of the anti-jamming probability of the simulated missile with or without interference.

5.1 System Implementation

Simulink is used to realize the infrared air combat simulation system. Simulink is a visual simulation tool based on model design, which can quickly transform the overall architecture of the simulation system and the design of its sub-models into specific implementations. The implementation process will not be described in detail here, only the method of obtaining the key parameters of each model is given.

(1) Acquisition of aerodynamic parameters. The aerodynamic force and moment coefficients of fighter and missile are calculated by DATCOM and Missile DATCOM software respectively.
(2) Acquisition of engine thrust parameters. It is mainly obtained from the open source flight dynamics model JSBSim.

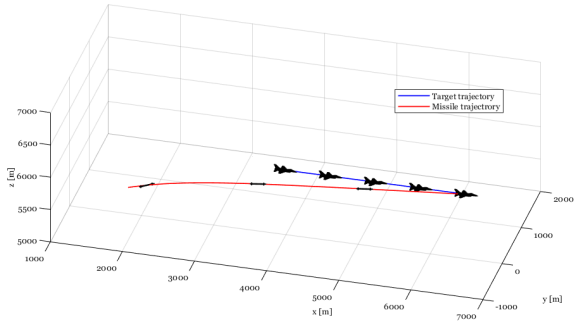

Figure 9: The Target Fighter Does Not Release the Bait.

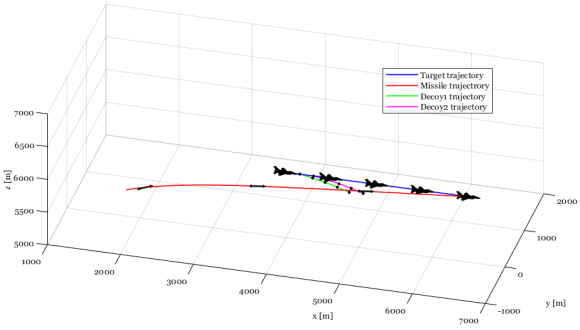

Figure 10: The Target Fighter Releases Two Decoys.

(3) Acquisition of infrared radiation intensity parameters. It is calculated and obtained by commercial CFD software ANSYS Fluent 16.0 [12].

5.2 Attack Simulation

Attack simulation verifies the missile's ability to attack the target fighter. The initial conditions of the simulation are: both the target fighter and the missile are on a horizontal plane with a height of 6000 meters, the initial Mach number of the target fighter and the missile are both 0.6, and the missile launch position is 3000 meters away from the target fighter, and an angle of 40 degrees with the flying direction of the target fighter. When attacking, the off-axis angle of the missile is 10 degrees horizontally and 0 degrees vertically. The simulation result is shown in Figure 9. The simulation results show that the fighter and missile models are working properly.

5.3 Defense Simulation

The defense simulation verifies the defense capability of the target fighter against incoming missiles when the decoys are launched. The initial conditions of the simulation are the same as the attack simulation, except that the target fighter launches two decoys during the missile attack. The simulation result is shown in the figure 10. The simulation results show that the missile and decoy models work normally.

Table 1: Initial Conditions for the Height of 6000 M, the Speed of 0.6 Ma

	With no decoy interference	With decoy interference
Specified anti-interference probability in product specification	90%	Average ≥70%
Simulation result	96%	79%
Simulation error	6.7%	12.9%

5.4 Comparison of Anti-Interference Probability

Compare the anti-jamming probability of the missile when the target fighter launches the decoy or does not launch the decoy respectively. Both the target fighter and the missile are in the horizontal plane at a height of 6000 meters, and the target fighter is within the missile attack zone. The initial speed of both is 0.6 Mach, and the target fighter makes evasive missile maneuvers. Simulate 1000 times each for the target fighter launches decoys or does not launch decoys, count the number of missile hits, and calculate the missile's anti-jamming probability. The results are shown in Table 1. It can be seen from the table that the error between the simulated missile anti-jamming probability and the real probability is within 15%, and the simulation results are true and credible.

6 CONCLUSIONS

The article discussed the general requirements of simulation system construction, designed the overall architecture of the infrared air combat simulation system, established jet fighter, infrared air-to-air missile, point source decoy and environment model, and finally implemented the simulation system based on Simulink. Through the attack and defense simulation examples and error analysis of missile anti-jamming probability, it was verified that each model in the infrared air combat simulation system worked normally and credibly, and the simulation system could effectively simulate various infrared air combat scenarios.

The establishment of the simulation system provides a foundation for the application of deep reinforcement learning algorithms in infrared air combat. However, it should be noted that there is a certain difference between the infrared air combat simulation system and the reality. The effect of this difference on the behavior of the agent based on deep reinforcement learning in the real world needs further research.

ACKNOWLEDGMENTS

This work was funded by the Natural Science Basic Research Program of Shaanxi Province, China. (Project No.2019JQ-290).

REFERENCES

[1] Oriol Vinyals, Igor Babuschkin, Wojciech M. Czarnecki, *et al* (2019). Grandmaster level in StarCraft II using multi-agent reinforcement learning[J]. Nature, 575(7782),350-354.
[2] Yongjie Xu, Zhijun Wang, Benbing Gao (2015). Six-Degree-of-Freedom Digital Simulations for Missile Guidance and Control[J]. Mathematical Problems in Engineering, 2015,1-10.
[3] Zhang Shuai, Wang Yu, Zhong Bowen, *et al*(2016). Model of Turbofan Engine for Civil Aircraft Conceptual Design[J]. Journal of Nanjing University of Aeronautics & Astronautics, 48(3),382-388.
[4] Randal W. Beard, Timothy W. Mclain(2012). Small Unmanned Aircraft Theory and Practice[M]. Princeton University Press, New Jersey, USA.
[5] Lu Xiaodong, Guo Jianguo, Lin Peng, *et al*(2010). Design of Missile Guidance and Control System and Matlab Simulation [M]. Northwestern Polytechnical University Press, Xian, Shannxi, China.
[6] Huang Hesong, Tong Zhongxiang, Li Jianxun, *et al*(2015). Simulation of pulse modulated infrared air-to-air missile[J]. Journal of Laser & Infrared, 44(9),1019-1024.
[7] Huang Hesong, Tong Zhongxiang, Li Jianxun, *et al*(2015). Functional simulation of infrared air-to-air missiles based on combat assessment[J]. Journal of Infrared and Laser Engineering, 44(3), 803-809.
[8] WANG Jie, DING Dali, XU Ming, *et al.* (2019). Air-to-air missile launchable area based on target escape maneuver estimation[J]. Journal of Beijing University of Aeronautics and Astronautics, 45(4),722-734.
[9] Jiang Yiyang (2016). Comparative Study on Three Acceleration Autopilots[J]. Journal of Navigation Positioning & Timing, 3(1),40-46.
[10] LI Quancheng, ZHU Chuanxiang, FAN Yonghua, *et al* (2019). Study on Infrared Air-to-Air Missile Guidance Accuracy Affected by Complicated Environment[J]. Journal of Northwestern Polytechnical University, 37(3),457-464.
[11] Richard S. Sutton, Andrew G. Barto (2018). Reinforcement Learning: An Introduction [M]. MIT Press, Cambridge, MA,USA.
[12] Wu Yanqing, Liao Shouyi, Zhang Zuoyu, *et al* (2018). Modeling of flow field and analysis of IR characteristic of aircraft based on Fluent[J]. Infrared and Laser Engineering, 47(7),1-9.