Hindawi
International Journal of Aerospace Engineering
Volume 2023, Article ID 2273139, 8 pages
https://doi.org/10.1155/2023/2273139

Research Article

Suppressing Uncommanded Roll-Yaw Motion by Jet Flow Control Based on Reinforcement Learning

Yizhang Dong,[1,2] **Zhiwei Shi**,[1] **Kun Chen**,[3] **and Zhen Chen**[1]

[1]*Key Laboratory of Unsteady Aerodynamics and Flow Control, Ministry of Industry and Information Technology, Nanjing University of Aeronautics and Astronautics, Yudao Street 29, Nanjing, Jiangsu 210016, China*
[2]*The National Key Lab of Computational Mathematics & Experimental Physics, Nandahongmen Street 1, Beijing 100076, China*
[3]*National Key Laboratory of Transient Physics, Nanjing University of Science and Technology, Nanjing, Jiangsu, China*

Correspondence should be addressed to Kun Chen; 12016120@njust.edu.cn

Received 20 February 2023; Revised 11 April 2023; Accepted 24 April 2023; Published 8 May 2023

Academic Editor: Hao Chen

The suppression of uncommanded motion of aircraft at high angles of attack (α) is a research topic of continuous concern in the aviation field. Aiming at the suppression of uncommanded roll-yaw motion of a canard aircraft at high angles of attack, an experimental method of virtual flight test based on reinforcement learning is proposed in this paper. In the virtual flight experiment, the agent was trained to control the jet actuators, so as to suppress the uncommanded roll-yaw motion. Force measurements were conducted to obtain the performance of the jet actuators in a low-speed wind tunnel. The results show that when the spanwise jet actuator and the reverse jet actuator were working on the same side, their control effects were suppressed by each other. Then, the stability augmentation control law was trained through virtual flight experiments based on a reinforcement learning algorithm (TD3), and the uncommanded motion was successfully suppressed. The time histories of the reinforcement learning agent's action in tests were analyzed, showing that the agent can avoid the coupling relationship between two kinds of jet actuators during tests.

1. Introduction

In the high maneuvering process, high-α flight, which results in large-scale separation on the upper surface of the wing and complex vortex structure, is almost inevitable. As a result, the aircrafts are vulnerable to loss stability in that situation, which leads to a variety of uncommanded motions, including the famous "wing rock." The previous studies on wing rock were carried out around the uncommanded rolling motion, ignoring yawing motion which should also exist and play an important role in real flight.

Walker and Ahmed [1] carried out free-to-roll, free-to-yaw, and free-to-roll-and-yaw experiments on a 75-degree swept delta wing, found that the uncommanded roll-yaw motion diverged rapidly, and finally, maintained as the constant-amplitude limit cycle oscillation. The results show that compared with the two single-DOF motions, the amplitude of the roll-yaw motion was smaller, and the average lift was significantly reduced, which meant that for a slender delta wing, the self-excited roll-yaw motion was more likely to result in the stall than simple rolling oscillations. Lin et al. [2] conducted a forced oscillation study on a fighter model with side strips. The results show that the yaw-roll coupling ratio (yaw angular velocity/roll angular velocity) has a strong influence on the damping characteristics of rolling and yawing moments at high angles of attack. In terms of suppression of uncommanded roll-yaw motion, Pedreiro et al. [3] set up the mathematical model of the roll-yaw motion of a wing-body aircraft at high angles of attack and then suppressed the motion by tangential blowing located on the nose. However, in the design process of control law, Pedreiro et al. introduced lots of human prior knowledge, which made the design process depend on the accuracy of modeling. It is undeniable that uncommanded roll-yaw motion is a very complex nonlinear motion. However, with the development of artificial intelligence (AI) technology, people are expected to use advanced AI algorithms to solve this problem without human knowledge.

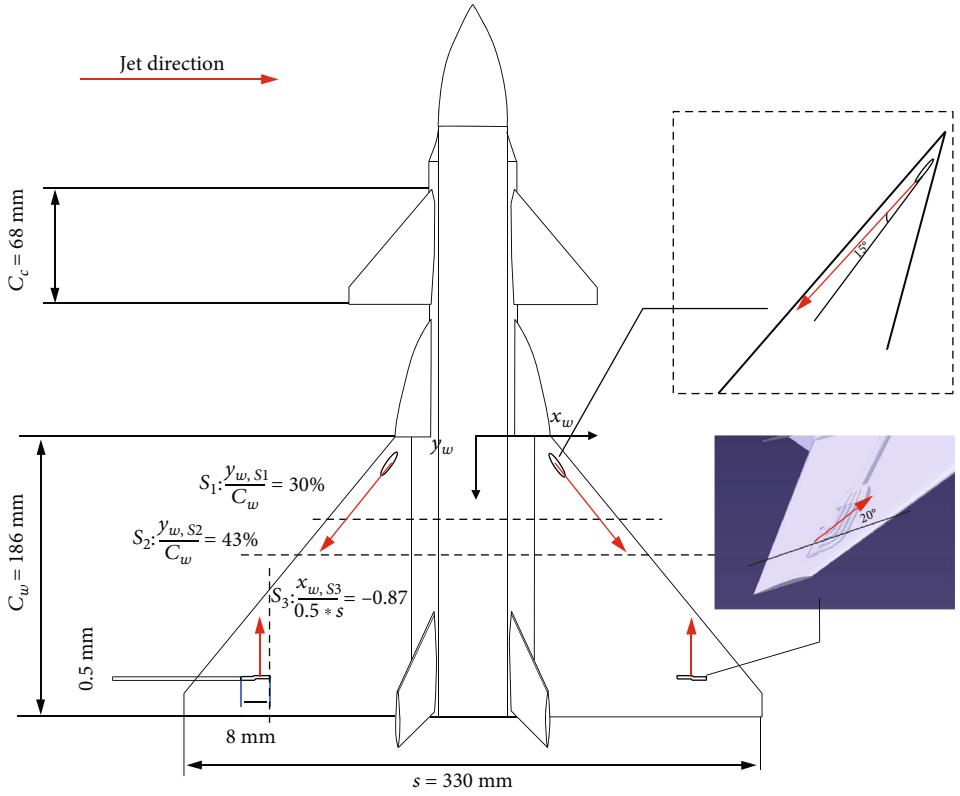

Figure 1: The schematic of the canard configuration aircraft model.

There are two difficulties in suppressing uncommanded roll-yaw motion. One is that the control efficiency of the traditional rudders and ailerons is weak at high angles of attack, and the other is that the dynamic system of this motion has the characteristics of nonlinearity and strong coupling among states, which makes it difficult to establish a mathematical model. To deal with the above two difficulties, this paper uses jet flow control to replace the traditional rudder to realize attitude control and uses model-free deep reinforcement learning (MFDRL) algorithm [4] to train an agent in the virtual flight test.

Deep reinforcement learning is dedicated primarily to train machines to make sequential decisions. A survey of DRL techniques [5] enumerated the successes of model-free RL approaches from Alpha-Go games [6] to quadcopter stabilization control [7]. However, several studies have investigated the combination of DRL and aerospace science, including target-missile-defender engagement [8],missile terminal guidance [9], and UAV control. Xu et al. [10] addressed the autonomous shape optimization problem of intelligent morphing aircraft based on mission requirements and flight status. The model dynamics were learned from flight data collected from expert pilots and fit to a nonlinear second-order ordinary differential equation (ODE). The control policy was based on a linear quadratic regulator feedback controller using a linear approximation to the learned dynamics. Clarke and Hwang [11] proposed a DRL-based controller to enable aerobatic maneuvering for capable fixed-wing aircraft in a simulation environment. Through trial-and-error simulations, the controller explored the full range of nonlinear flight envelopes and learned an aerobatic maneuver in a matter of hours by itself and without human input.

Few historical studies combine DRL and the canard-configuration aircraft with high-α control, especially based on real-world experiments. Based on our previous work on the roll oscillations of canard-configuration aircraft [12], this paper utilizes the DRL algorithm in the real-flight test environment to train an agent to suppress the uncommanded roll-yaw motion so as to avoid the complex modeling of the nonlinear dynamics.

2. Introduction of the Model

The schematic of the canard configuration model used in this paper is shown in Figure 1. The whole model includes the nose, canard wings, streak wings, main wings, and V-shaped vertical tails. The main wing root chord length (C_w) and canard wing root chord length (Cc) are 186 mm and 68 mm, respectively, and the wingspan (s) is 330 mm. The sweep angles of the main wing, canard wing, and streak wing are 50°, 50°, and 65°, respectively. The root chord length of the main wing is used as a reference value to realize the non-dimensionalization of measured aerodynamic data. The moment of inertia around the roll and yaw axes of the model are 0.002 kgm^2 and 0.016 kgm^2, respectively. To provide control moments of rolling and yawing, the leading-edge spanwise jets and reverse jets were designed. As shown in

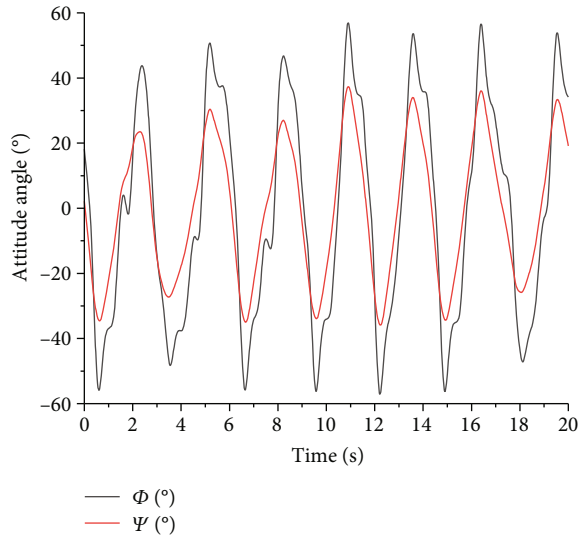

FIGURE 2: The time histories of the roll/yaw angles during the free-to-roll-and-yaw test at $\alpha_0 = 35\circ$.

Figure 1, the exit direction of the leading-edge spanwise jets was in the same vertical plane as the leading-edge direction at an angle of 15°. The cross-sectional diameter of the slots was 2.5 mm. According to previous research [13], the main principle of this jet configuration's control effect for rolling is to increase the local lift by delaying leading-edge vortex breakdown. The design of the reverse jet actuator is based on the research results of Zhu et al. [14]. However, unlike the previous research which applied the reverse jet actuator at small angles of attack, this paper applied it to the more complex situation at high angles of attack to provide a yawing moment for the aircraft.

3. Results and Analysis

Figure 2 shows the time histories of roll and yaw angles in the free-to-roll-and-yaw experiments at a nominal angle of attack of 35° (the angle of attack when the roll angle and yaw angle are both zero). It can be seen from the figure that the amplitudes of the rolling and yawing motions are large, and the frequencies of the motions around the two axes are almost the same, which indicates that the uncommanded roll-yaw motion is strongly coupled among states.

The control effect of the single jet was investigated by static force measurements. Figure 3 shows the effects of the reverse jet actuator on yawing moment coefficient and the spanwise jet actuator on the rolling moment coefficient. As seen in this figure, when $\alpha = 35\circ$, the model is lateral static instable ($C_{l\beta} > 0$ near zero sideslip) and directional static stable ($C_{n\beta} > 0$ near zero sideslip). From the perspective of the leading-edge spanwise jet actuator, the right spanwise jet will reduce the rolling moment coefficient, and the control effect has a positive relationship with its blowing momentum coefficient ($C_{\mu s}$). A positive sideslip angle represents that the right wing is on the windward side, with a larger actual angle of attack than that of the left wing, leading to a more severe vortex breakdown. Since the control principle of the leading-edge spanwise jet is to control the vortex breakdown, the jet on the right side has a more significant control effect on the rolling moment coefficient with a large sideslip angle. On the other hand, the reverse jet on the right side increases the yawing moment, so it can be inferred that the local resistance of the right wing of the model increases when the right reverse jet is working. It can also be seen from the figure that there is an obvious "dead zone" in the control effect of the reverse jet on the yawing moment at most tested angles of attack, which shows that when the momentum coefficient of the reverse jet is small, its influence on the yawing moment of the model is very weak, and after increasing the blowing momentum coefficient of the reverse jet, the yawing moment of the model begins to be significantly affected. From the above analysis, it can be concluded that the single reverse jet and single leading-edge spanwise jet can significantly affect the yawing moment and rolling moment of the model, respectively, and thus provide control moments. However, the control effect has a certain nonlinearity.

Figure 4(a) shows the effect of the leading-edge spanwise jet on the yawing moment at $\alpha = 35\circ$. It can be seen that when the reverse jet is working, the leading-edge spanwise jet with a small blowing momentum coefficient has a weak influence on the yawing moment, while the leading-edge spanwise jet with a large blowing momentum coefficient will reduce the yawing moment of the model. In other words, when the spanwise jet and the reverse jet on the same side are working concurrently, the spanwise jet will suppress the control effect of the reverse jet. Figure 4(b) shows the effect of the reverse jet on the rolling moment coefficient when the leading-edge spanwise jet is working. In the presence of the right leading-edge spanwise jet, the right reverse jet will increase the rolling moment of the model. It can be inferred that the reverse jet reduces the local lift of the right wing, and this effect is more significant in the case of a negative sideslip. This may be due to the fact that at a negative sideslip, the right wing is on the leeward side, and its vortex lift is more significant, so the damage to the vortex lift by the reverse jet is more obvious.

Through the above results, we have completed the design and characteristic analysis of the spanwise jet and reverse jet actuators. At high angles of attack, the leading-edge spanwise jet and reverse jet can have significant effects on the rolling moment and yawing moment of the model, respectively, but there is an obvious mutual suppression between the two kinds of the jet. In particular, when the jets on the same side are working at the same time, one kind of jet will suppress the control effect of the other kind. The coupling of the control instruments increases the difficulty of the control law design of the stabilization control system. To this end, we introduce the method of reinforcement learning virtual flight experiment to complete the control law design of the suppression of roll-yaw motion. Figure 5 shows the experimental architecture of the reinforcement learning virtual flight experiments, which is very similar to the architecture

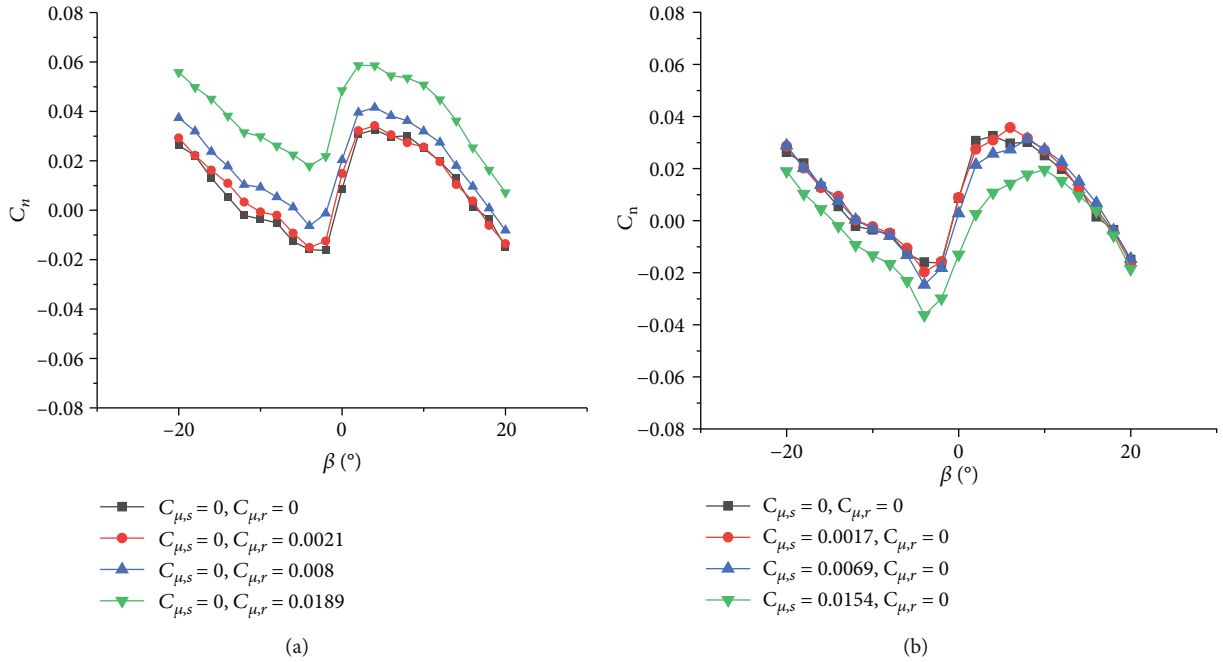

Figure 3: The effects of (a) reverse blowing to yawing moment coefficient and (b) spanwise blowing to rolling moment coefficient.

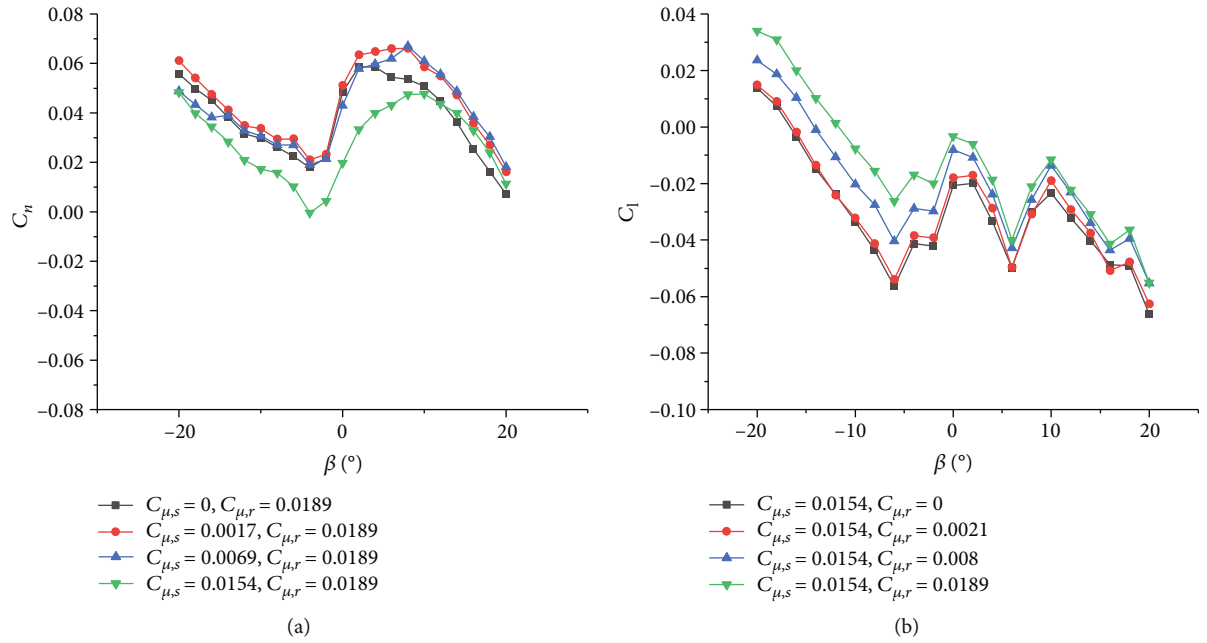

Figure 4: The coupling effect of two blowing methods:(a) the effect of the spanwise blowing to the yawing moment when reverse blowing was working and (b) the effect of the reverse blowing to the yawing moment when spanwise blowing was working.

utilized in our previous work [15]. During the experiments, the air supply system provides a stable and clean air source to the electromagnetic proportional valve (PVQ-31), and then the electromagnetic proportional valve injects compressed air to generate jets to provide a rolling/yawing moment. The main program receives the attitude data (including roll/yaw angle and angular velocity) sent from the attitude sensor and drives the electromagnetic proportional valve to work by sending the serial port signal (action), which affects the motion of the model. The single-step iteration frequency of the whole experiment was set to be 100 Hz (determined by the frequency of the sensor), and the single episode time is set to 10 s, ($t_s = 0.01s, t_f = 10s$). The action of the reinforcement

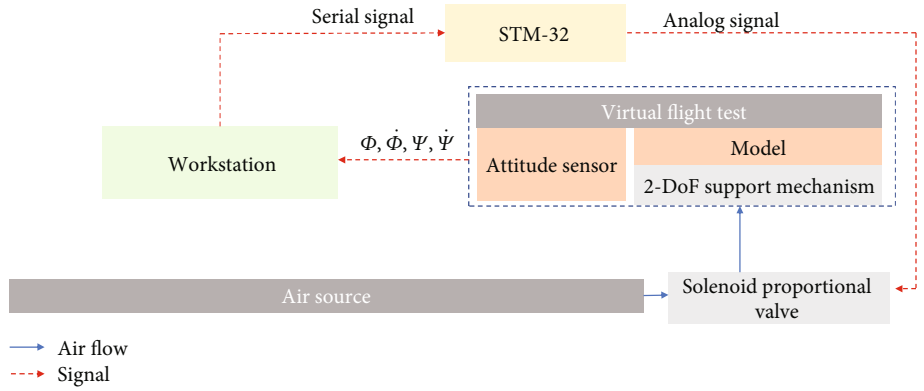

Figure 5: The experimental architecture of the reinforcement learning virtual flight test.

Table 1: TD3 hyperparameters.

Parameter	Value
Optimizer	Adam [19]
Number of hidden layers (all networks)	2
Number of hidden units per layer	256
Critic learning rate	$1e-3$
Actor learning rate	$3e-4$
Discount factor (γ)	0.99
Exploration noise	0.1
Policy noise	0.2
Range to clip policy noise	0.5
Target smoothing coefficient (τ)	0.005
Number of samples per minibatch	256
Policy update frequency	2
Activation function	ReLU (rectified linear unit) [20]

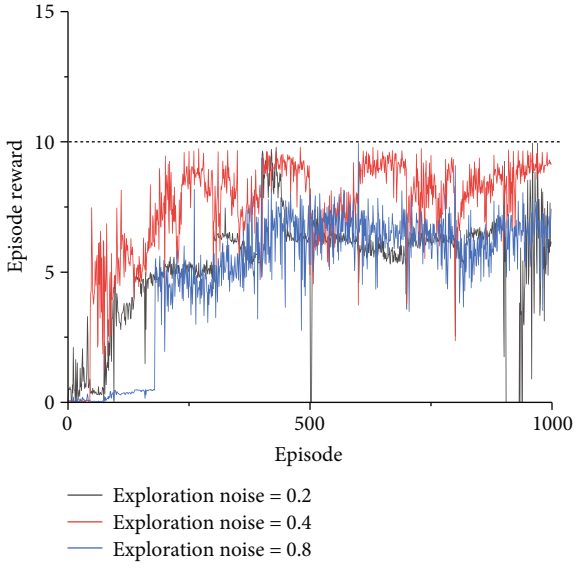

Figure 6: The episode reward curve during training.

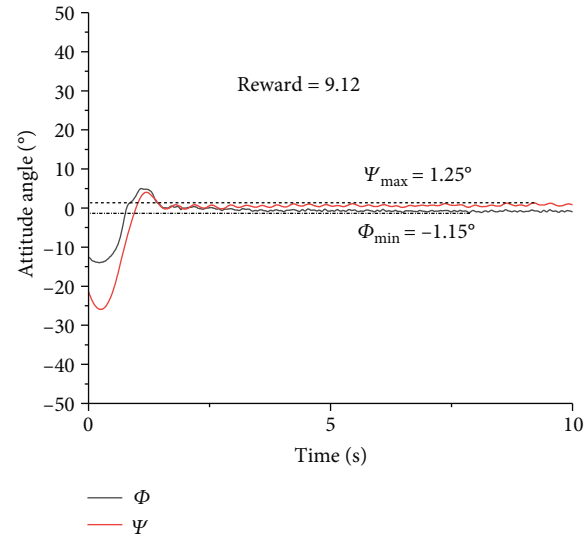

Figure 7: The time histories of yaw and roll angle during the test.

learning agent is the control voltage of four electromagnetic proportional valves.

The reinforcement learning experiments used TD3 [16] (twin-delayed deep deterministic policy gradient algorithm) algorithm, which is an algorithm based on actor-critic architecture [17]. TD3 sets up two groups of Q networks to evaluate the value of the agent's actions, thereby avoiding the problem of action value's overestimating of the DDPG algorithm [18]; at the same time, it adopts the method of policy gradient ascent to improve the agent's strategy. During the experiments, three kinds of exploration noise (0.2, 0.4, and 0.8) were set for this algorithm. The agent receives the roll angle/angular velocity and yaw angle/angular velocity data given by the attitude sensor to construct the observation vector and calculates the reward value. The hyperparameter settings of the algorithm are shown in Table. 1.

Due to the sensitivity of aerodynamics to sideslip at high angles of attack, the actions were considered independently without introducing symmetries. In order to overcome the non-Markovian property in the real experiment, a certain memory mechanism was added to the experiment, and the

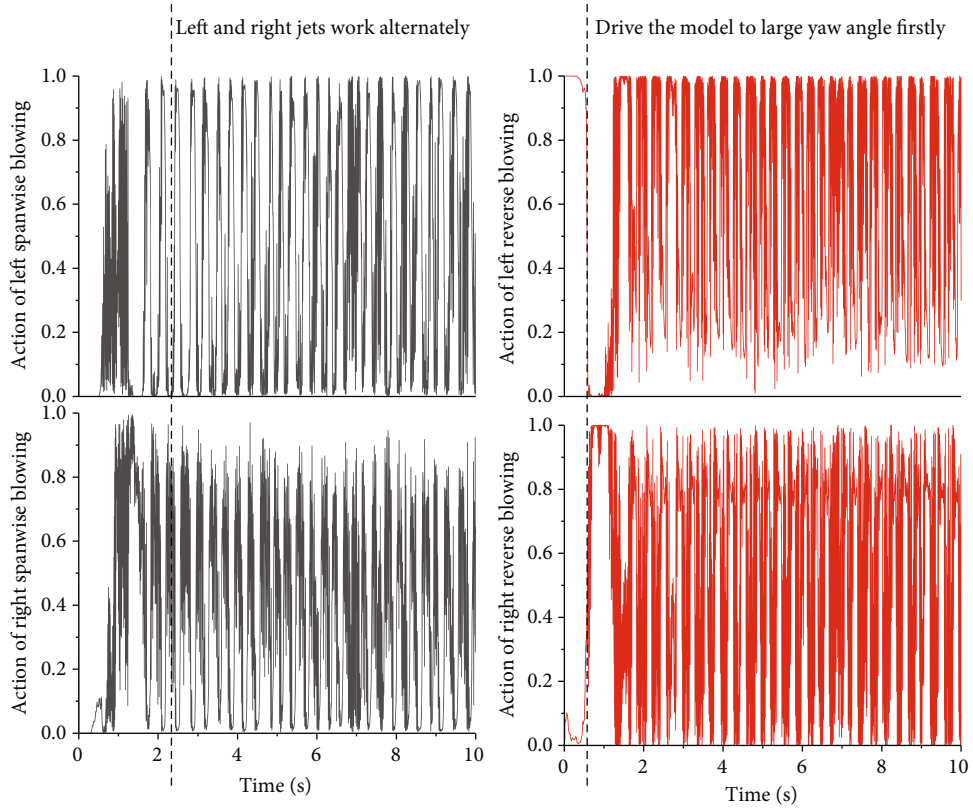

Figure 8: The time histories of the agent's actions during the test.

time step of the observation vector was set to 3. The mathematical expression of the observation vector of the agent is as follows:

$$\begin{aligned}&\left(\Phi_t, \dot{\Phi}_t, \Psi_t, \dot{\Psi}_t, a_{1,t}, a_{2,t}, a_{3,t}, a_{4,t}, \Phi_{t-1}, \dot{\Phi}_{t-1}, \Psi_{t-1},\right.\\&\dot{\Psi}_{t-1}, a_{1,t-1}, a_{2,t-1}, a_{3,t-1}, a_{4,t-1}, \Phi_{t-2}, \dot{\Phi}_{t-2}, \Psi_{t-2},\\&\left.\dot{\Psi}_{t-2}, a_{1,t-2}, a_{2,t-2}, a_{3,t-2}, a_{4,t-2}\right).\end{aligned} \tag{1}$$

The following formula gives the reward function of this experiment. It can be seen from the formula that a very simple form of reward function is set in this experiment. Four precision levels (20°, 10°, 5°, and 2°) are set for both the roll angle and the yaw angle. When the attitude angle of the model is controlled within a certain precision level, the agent will receive a corresponding reward. Therefore, the highest reward that a reinforcement learning agent can get in a single episode is 10. It should be noted that in the process of training and testing, to ensure the generalization ability of the control strategy to the initial state of the model, random jets were performed for 5 seconds before each episode of training or testing to generate a random initial state.

$$\begin{cases}\mathrm{r}_{\Phi} = 0.1^{*}(|\Phi| \le 20\circ) + 0.2^{*}(|\Phi| \le 10\circ) + 0.3^{*}(|\Phi| \le 5\circ) + 0.4^{*}(|\Phi| \le 2\circ),\\ \mathrm{r}_{\Psi} = 0.1^{*}(|\Psi| \le 20\circ) + 0.2^{*}(|\Psi| \le 10\circ) + 0.3^{*}(|\Psi| \le 5\circ) + 0.4^{*}(|\Psi| \le 2\circ),\\ r = 0.005^{*}\mathrm{r}_{\Phi} + 0.005^{*}\mathrm{r}_{\Psi}.\end{cases} \tag{2}$$

Figure 6 shows the episode reward curve during training with different exploration noise settings. It can be seen from the figure that when the exploration noise was set too small, the training episode reward grows slowly, and after reaching a certain high value, the reward curve begins to decline. This may be due to the fact that the algorithm takes longer to jump out of the local optimum when the exploration noise is small; after adjusting the exploration noise to 0.4, the performance of the algorithm is much better. Through exploration, the agent quickly finds a more perfect control strategy,

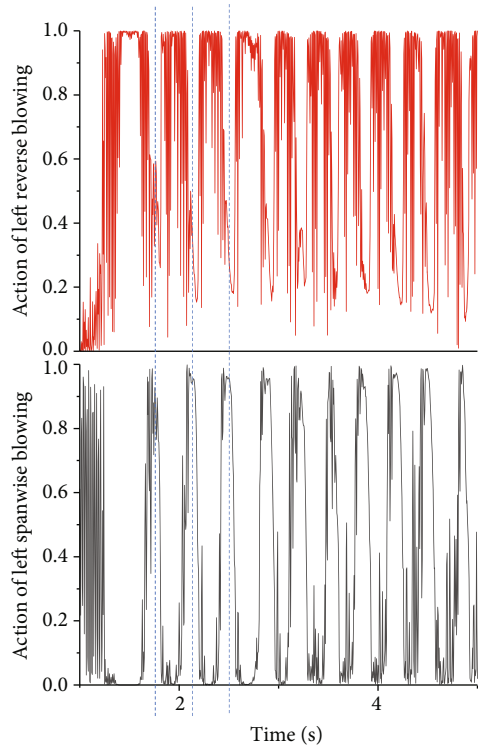

FIGURE 9: The analysis of the agent's actions on the same side (noise 0.4).

and then the episode reward has been maintained at a high level; when the exploration noise is increased to 0.8, the performance of the agent is poor at the beginning, because the parameters of the policy are still far from the optimal parameters, and the exploration noise of the agent is large. At this time, the agent is in a relatively random exploration. Afterwards, the reward for subsequent training suddenly increases, but due to the larger exploration noise, the reward in this situation is generally lower than that when the exploration noise is 0.4.

Figure 7 shows the test results of an agent when the exploration noise is 0.4. The final roll and yaw angles of the agent are not completely controlled at 0°. The lower boundary of the roll angle in the steady state is $\Phi_{\min} = -1.15^{\circ}$, and the upper boundary of the yaw angle is $\Psi_{\max} = 1.25^{\circ}$. But in terms of the reward function, the agent has already got the highest single-step reward. The cumulative reward for the full-episode test is 9.12. Such test result shows that the agent has almost obtained the optimal policy.

Figure 8 shows the time histories of the agent's actions during the test. It can be seen that when the roll angle and yaw angle are far from zero at the beginning, the agent chooses not to activate the leading-edge spanwise jet actuators on both sides but turns on the reverse jet actuator on the left first, so that the model can obtain a negative yawing moment increment (see Figure 3). With the effect of the reverse jet, the yaw angle of the model continues to deviate to a larger absolute value and then starts to rebound. At this time, the agent chooses to reverse the working state of the reverse jet actuators on both sides, the left reverse jet stops working, and the right reverse jet starts to work. The model obtains a positive yawing moment increment, accelerating the recovery of the model yaw angle. At the same time, the leading-edge spanwise jet actuator on the left side also starts to work, so that the model obtains a positive rolling moment increment, and the roll angle also begins to accelerate the recovery. Then, when the roll and yaw angles of the model overshoot, the directions of the two kinds of jets are reversed, so that both the roll and yaw angles of the model return to around 0°. Finally, the jet actuators work alternately, making the roll and yaw angles of the model stabilized around 0°. In addition, it is worth noting that, based on the above analysis (see Figure 4), when the reverse jet and the leading-edge spanwise jet located on the same side are working together, the control effect of the reverse jet on the yawing moment will be suppressed by spanwise jet, and the control effect of the spanwise jet on the rolling moment is also suppressed by the reverse jet. In the agent's strategy, it consciously avoids the situation where the two jets on the same side work together. Figure 9 shows the analysis of the agent's actions on the left side. It can be seen from the figure that when the action value of the reverse jet on the left reaches the peak, the spanwise jet is not working. The two actuators work alternately, cleverly avoiding the problem of control coupling.

4. Conclusion

In conclusion, this paper explores the characteristics of the uncommanded roll-yaw motion of a canard-configuration model at a high nominal angle of attack through a free-to-roll-and-yaw experiment. Experiments show that this uncommanded motion has a large amplitude and obvious state coupling. To provide lateral/directional control moment to the model, spanwise jet and reverse jet actuators were designed, respectively. Through force measurement experiments, it is found that the control effects of the two actuators have the characteristics of nonlinear and strong coupling. In the wind tunnel virtual flight experiments, the deep reinforcement learning algorithm (TD3) was used to train the stability augmentation control law of the model, which successfully suppressed this uncommanded roll-yaw motion. The action of the reinforcement learning agent during testing is analyzed, and it is found that the agent avoids the coupling relationship between the two kinds of jet actuators. The results of this paper can provide some technical support for the design of complex control law and the development of intelligent aircraft.

Data Availability

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to state restrictions such as privacy or ethical restrictions.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Natural Science Foundation of Jiangsu Province (Grant no. BK20200482) and the National Natural Science Foundation of China (Grant no. 12002166)

References

[1] J. Walker and A. Ahmed, "An investigation of coupled roll-yaw oscillations of a slender delta wing," in *51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, p. 237, 2013.

[2] S. Lin, D. Huang, and W. U. Genxing, "Effects of yaw-roll coupling ratio on lateral-directional aerodynamic characteristics," *Chinese Journal of Aeronautics*, vol. 32, no. 2, pp. 272–280, 2019.

[3] N. Pedreiro, S. M. Rock, Z. Z. Celik, and L. Roberts, "Roll-yaw control at high angle of attack by forebody tangential blowing," *Journal of Aircraft*, vol. 35, no. 1, pp. 69–77, 1998.

[4] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, MIT press, 2018.

[5] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, "Deep reinforcement learning: a brief survey," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 26–38, 2017.

[6] D. Silver, J. Schrittwieser, K. Simonyan et al., "Mastering the game of Go without human knowledge," *Nature*, vol. 550, no. 7676, pp. 354–359, 2017.

[7] J. Hwangbo, I. Sa, R. Siegwart, and M. Hutter, "Control of a quadrotor with reinforcement learning," *IEEE Robotics and Automation Letters*, vol. 2, no. 4, pp. 2096–2103, 2017.

[8] V. Shalumov, "Cooperative online guide-launch-guide policy in a target-missile-defender engagement using deep reinforcement learning," *Aerospace Science and Technology*, vol. 104, article 105996, 2020.

[9] B. Gaudet, R. Furfaro, and R. Linares, "Reinforcement learning for angle-only intercept guidance of maneuvering targets," *Aerospace Science and Technology*, vol. 99, article 105746, 2020.

[10] D. Xu, Z. Hui, Y. Liu, and G. Chen, "Morphing control of a new bionic morphing UAV with deep reinforcement learning," *Aerospace Science and Technology*, vol. 92, pp. 232–243, 2019.

[11] S. G. Clarke and I. Hwang, "Deep reinforcement learning control for aerobatic maneuvering of agile fixed-wing aircraft," in *AIAA Scitech 2020 Forum*, pp. 1–18, Orlando, FL, USA, 2020.

[12] Y. Dong, Z. Shi, K. Chen, and Z. Ge, "Suppression of roll oscillations of a canard-configuration model using fluid effector and reinforcement learning," *Journal of Aerospace Engineering*, vol. 36, no. 3, 2023.

[13] I. Gursul, Z. Wang, and E. Vardaki, "Review of flow control mechanisms of leading-edge vortices," *Progress in Aerospace Sciences*, vol. 43, no. 7-8, pp. 246–270, 2007.

[14] Z. Shi, J. Zhu, X. Dai et al., "Aerodynamic characteristics and flight testing of a UAV without control surfaces based on circulation control," *Journal of Aerospace Engineering*, vol. 32, no. 1, 2019.

[15] Y. Dong, Z. Shi, K. Chen, and Z. Yao, "Self-learned suppression of roll oscillations based on model-free reinforcement learning," *Aerospace Science and Technology*, vol. 116, article 106850, 2021.

[16] S. Fujimoto, H. Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in *International Conference on Machine Learning*, pp. 1587–1596, Stockholm, Sweden, 2018.

[17] J. Peters and S. Schaal, "Natural actor-critic," *Neurocomputing*, vol. 71, no. 7-9, pp. 1180–1190, 2008.

[18] T. P. Lillicrap, J. J. Hunt, A. Pritzel et al., "Continuous control with deep reinforcement learning," 2015, https://arxiv.org/abs/1509.02971.

[19] D. P. Kingma and J. L. Ba, "Adam: a method for stochastic optimization," in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, San Diego, California, 2015.

[20] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proceedings of the 27th international conference on machine learning (ICML-10)*, pp. 807–814, Haifa Israel, 2010.