Design and Realization of Intelligent Aero-engine DDPG Controller

Rong-rong Qian¹, Yimin Feng ², Mengman Jiang ², Lijun Liu ²*

¹AECC Commercial Aircraft Engine Co., LTD;
²School of Aerospace Engineering, Xiamen University, Xiamen 361000, China
* School of Aerospace Engineering, Xiamen University, liulijun@xmu.edu.cn

Abstract: As the artificial intelligence technology advances, intellectualization has become an important development trend for the future aero-engine industry. The intelligent aero-engine will bring a new round of technology revolution in aviation industry. Since an aero-engine is a complex system with strong nonlinearity and coupling, it is difficult to achieve the optimal control performance with traditional PI control. Therefore, the DDPG (Deep Deterministic Policy Gradient, DDPG) algorithm belonging to the family of Deep Reinforcement Learning is hereby proposed for designing aero-engine controller. Based on the non-linear polynomial state-space mathematical model of JT9D turbofan engine, the intelligent DDPG controller is designed and then compared with the performance of PI controller. The proposed controller can achieve the optimal control effect for aero-engine through the training the neural network. The simulation results demonstrates that compared with PI control, the control method proposed in this paper achieves superior control performance in response speed, overshoot and anti-interference ability.

1. Introduction
An aero-engine is a complex system with strong nonlinearity, and its performance directly affects the safety, reliability and economy of the aircraft. For the aero-engine development, it is necessary to design an effective control system in order to ensure that aero-engine can work within stability margins with desired performance. At present, the traditional aero-engine control method is generally based on proportional integral (PI) controller, which has a simple structure, strong robustness, and easy adjustment[5]. However, the turbofan engine is a complicated plant with wide operating range and long parameter-varying features, traditional PI controller can only realize the desired control performance on the single point or within a small operating range[6]. Therefore, the control structure of traditional PI controllers with limit protections, showing in Fig.1, is the most popular control method for the aero-engine.
Although the above structure has been widely applied in the aerodynamic engine control, it cannot satisfy the increasing demand of aero engine control performance such as high efficiency, low fuel consumption with high level safety. On account to these disadvantages of PI controller, many improved control methods for aero engine are developed. Recent examples of such efforts include applying dispatch control[1], linear quadratic Gaussian/loop transfer recovery (LQG/LTR) Control[3], adaptive control[4], sliding mode controllers to improve the dynamic engine response. In order to maximize the potential of aero-engines, more advanced control strategies must be taken into the aero-engine control. With the development of artificial intelligence technology and the evolution of computing processing power, the research and development of intelligent aero engine controller can be available for the requirements of high performance, high reliability and low cost, etc., for future aero engine. With respect to the artificial intelligence controllers, literature[8] proposes a new reactive control method through the combining reinforcement learning with fuzzy logic, which realizes different levels of robots’ control in an unknown environment. In reference to the complexity and variability of the environment in automatic driving control, literature[9] applies the online learning of reinforcement learning to train neural network using the interactive learning methods. At the same time, it improves the generalization ability of the model by clustering and sampling the state-space. However, the DRL (Deep Reinforcement Learning, DRL) algorithm is rarely used in the domain of aero-engine control.

Based on mentioned above, a kind of reinforcement learning, the deep deterministic policy gradient (DDPG) algorithm as a steady-state aero-engine controller, is designed and realized in this paper by training the nonlinear polynomial state space(NLPSS) model of JT9D aero-engine. In order to accelerate the rate of convergence and avoid the divergence and over-fitting of training process for DDPG controller based on NLPSS model, the DDPG controller based on linear state-space model(LSS) of JT9D has been designed as the pre-training and the acquired controller parameters are used as the initial parameters for DDPG controller based on NLPSS model. The pre-training provides the better model initialization and improves the generalization capability of network. The simulation results show that the proposed method has better control effect on the big working range of JT9D aero-engine in the aspect of response speed, overshoot and anti-interference ability.

2. Establishment of aero-engine state-space model
In this paper, the engine model of JT9D from T-MATS is used to design the intelligent controller. The T-MATS module is a thermodynamic system simulation library based on Matlab/Simulink platform developed by NASA Glenn Research Center. The aero-engine model structure of JT9D is analyzed in Matlab. Taking the cruise phase as an example, the aero-engine operating environment is as follows:

$$W = 661.26 \text{ pps}, h = 130 \text{ BUT/(lbm*R)}, T_i = 448.46 \text{ degR}, P_i = 5.528 \text{ psia}, P_{amb} = 3.626 \text{ psia}.$$ 

Where $W$ is the gas path flow of fan inlet, $h_i$ and $T_i$ are the total enthalpy and total temperature of fan inlet air flow, $P_{amb}$ is the ambient pressure.

Supposed that the aero-engine works at the certain steady point $X_0=[n_L, n_H]^T$, where $n_L$ is the low
pressure turbine speed, $n_H$ is the high pressure turbine speed. $\Delta n_L$ and $\Delta n_H$, the variation of $n_L$ and $n_H$, are considered as the state variables of aero-engine state-space model. $\Delta FAR$, the variation of fuel-air ratio of aero engine, is the control variable. In this paper, the steady speeds of $n_L$ and $n_H$ are specified as 3350.9 rpm (round per minute) and 7189.7 rpm and the corresponding fuel-air ratio is 0.0165.

The proposed NLPSS model of engine is designed as follows.

$$
\begin{align*}
\Delta \dot{n}_L &= \sum_{n=1}^{N} a_{(2n-1)} \Delta n_L^n + \sum_{n=1}^{N} b_{(2n)} \Delta n_H^n + b_1 \Delta FAR \\
\Delta \dot{n}_H &= \sum_{n=1}^{N} a_{(2n-1)} \Delta n_L^n + \sum_{n=1}^{N} a_{(2n)} \Delta n_H^n + b_2 \Delta FAR
\end{align*}
$$

(1)

Where, $N$ is represented as the model order. When $N=1$, equation (1) is the LSS of JT9D aero-engine. $a_{(2n-1)}, a_{(2n)}, a_{(2n-1)}, a_{(2n)}$ are the coefficients of state variables, $b_1, b_2$ are the coefficients of input variable.

It has been proved in the previous research[10] that the NLPSS model of aero-engine could be used in much broader operating range around steady point than the linear model. In order to reduce the switching times of controllers, the NLPSS model is used for the DDPG controller design. However, as the model order is getting bigger, the convergence of DDPG controller based on NLPSS model is hard. The initial values of controller parameters have enormous impact on the convergence of training process. Due to the simple structure and small working range of linear model, the parameters training for DDPG controller based on linear model has good astringency with quick convergence and stability. Therefore, in order to reduce the training time and avoid trapping in the local optimum, the DDPG controller design based on linear model are used as the pre-training for the final controller and the acquired controller parameters are used as the initial values for DDPG controller based on NLPSS model.

The final single-input linear and nonlinear state-space models of the JT9D aero-engine after normalization at a certain steady-state operating point are given in Equation (2) and (3). The outputs of fitting results are shown in Fig.2.

$$
\begin{align*}
\begin{bmatrix}
\Delta n_L \\
\Delta n_H \\
\Delta n_L \\
\Delta n_H
\end{bmatrix}
= 
\begin{bmatrix}
-2.5172 & 1.9325 & 0.6337 \\
2.0162 & -3.6242 & 1.3420
\end{bmatrix}
\begin{bmatrix}
\Delta n_L \\
\Delta n_H
\end{bmatrix}
+ 
\begin{bmatrix}
0.6337 \\
1.3420
\end{bmatrix}
\Delta FAR
\end{align*}
$$

(2)

$$
\begin{align*}
\begin{bmatrix}
\Delta n_L \\
\Delta n_H \\
\Delta n_L \\
\Delta n_H
\end{bmatrix}
= 
\begin{bmatrix}
-4.181 & 2.714 & 5.410 & -2.825 \\
-0.872 & -0.213 & 2.062 & -1.180
\end{bmatrix}
\begin{bmatrix}
\Delta n_L \\
\Delta n_H
\end{bmatrix}
+ 
\begin{bmatrix}
-2.804 & 1.234 & 0.01073 \\
-1.214 & 0.537 & 0.01610
\end{bmatrix}
\Delta FAR
\end{align*}
$$

(3)

The final single-input linear and nonlinear state-space models of the JT9D aero-engine after normalization at a certain steady-state operating point are given in Equation (2) and (3).

3. Design of DDPG controller

3.1 Design scheme of aero-engine DDPG controller

The design scheme of the DDPG controller proposed in this paper is shown in Fig.2. The main design process is as follows.
1) To build the training environment for the controller training based on the previously established JT9D engine linear/non-linear model;

2) To design the intelligent DDPG controller based on the linear model, including the construction of network structure, reward function design, experience replay buffer design, etc.

3) Based on the obtained linear controller parameters, the nonlinear model environment is optimized and trained.

4) Finally, the nonlinear single-input intelligent controller of the JT9D aero-engine is constructed.

It can be seen from equation (2) and (3) that the control input of engine is the change of fuel-air ratio ΔFAR, which is used as the action \( a \) given by the agent in the reinforcement learning. The output of the engine model is \( \Delta n_L \) and \( \Delta n_H \), which is used as the state \( S \) in the reinforcement learning. In the control process of aero-engine, the agent is desired to acquire the ability to control the engine through self-learning. When the expected value of low pressure turbine speed of aero-engine changes to a certain value, in order to achieve the best control effect, the agent is supposed to accurately give the corresponding action value, which is the control variable ΔFAR. The training process of the engine controller is shown in Fig.3.
3.2 Network framework design

The aero-engine intelligent controller is mainly composed of the modules of aero-engine mathematical model environment, policy networks, value networks, experience buffer, as shown in Fig.2.

A-Online is an online policy network whose function is to update the parameters of the policy network. The input of the network is the current state of the engine $S$, that is, the current low-pressure turbine speed change value $\Delta n_L$ of the engine. The output of the network is action $a$, that is $\Delta F\text{AR}$. Through the interaction with the environment, the next state $S_{t+1}$ and the reward $r$ can be obtained. The loss function can be expressed as equation (4):

$$L_a = J(\theta) = -\frac{1}{m} \sum_{i=1}^{m} Q_{\text{arg}}(s_i, a_i)$$

A-Target is a target policy network. Its function is to sample the actions through the experience buffer as the input to C-Target network and improve the convergence speed of the network. The input of A-Target is the next state of the state $S$ from the online actor network. Then the action of the A-online network is transferred to the aero-engine simulation environment and executed to obtain the next step state $S_{t+1}$ of the engine. The output of the A-target network is the action $a_{t+1}$. The update mode of network parameter is to copy regularly from the A-Online network, that is so-called the soft-update. The update equation is shown in equation (5):

$$\theta_{\text{arg}} \leftarrow \rho \theta_{\text{arg}} + (1-\rho)\theta$$

C-Online is an online critic network. Its function is to update the parameters of the A-online policy network and optimize the update direction for the actor network. The inputs of the C-online network are the current state of the engine $S$, the actual execution action $a$. The output of the network is the Q value of the current state $S$. The loss function can be expressed in equation (6):

$$L_c = J(\phi) = -\frac{1}{m} \sum_{i=1}^{m} (y_i - Q_{c_{\text{arg}}}(s_i, a_i))^2$$

C-Target is the target critic network. Its function is to calculate the target Q value, to optimize the iterative direction for A-online policy network, and to improve the convergence speed of the network. The input of C-Target network is the next state of the engine $S_{t+1}$ and the next state $a_{t+1}$. The update mode of C-Target network is similar to that of A-Target network. It also uses a soft-update method to periodically obtain the parameters from the C-Online network. The update equation is as follows in equation (7):

$$\phi_{\text{arg}} \leftarrow \rho \phi_{\text{arg}} + (1-\rho)\phi$$

3.3 Reward function design

In the reinforcement learning algorithm, the design of the reward function is significant which directly influences the control effect and the astringency. The agent in the reinforcement learning algorithm must be told which decision is preferred and which decision is not encouraged by the reward function. Therefore, in the aero-engine reinforcement learning control system, the difference between the current engine speed change value and the given command value is considered as the most important indicator. However, in the agent decision process, the reward becomes greater when the engine speed is getting closer to the command value and it would increase the amplitude of the action. This would result in the big overshoot of control effect. In order to avoid the overshoot, $\Delta \dot{n}_L$ is introduced in the reward function. $\Delta \dot{n}_L$ represents the acceleration of the low pressure turbine speed. The larger the value of $\Delta \dot{n}_L$ grows, the greater the stride range of the agent in the previous step to the target value becomes. The introduction of $\Delta \dot{n}_L$ can regulate the amplitude of the agent’s decision to a limited extent.

In this paper, we design two reward functions respectively for DDPG controller based on LPSS and NLPSS model.
3.3.1 Reward function design for DDPG controller based on LPSS

In this paper, the single-input linear reward function of the aero-engine is designed as shown in equation (8):

$$r = |a(\Delta \eta^*_{L} - \Delta \eta_{L})| + |b(\Delta \eta^*_{H} - \Delta \eta_{H})| + |c\Delta \eta_{L}|$$  \hspace{1cm} (8)

Where, $a$, $b$ and $c$ are the arbitrary constants. $\Delta \eta^*_{L}$ and $\Delta \eta^*_{H}$ are the command change values of low and high pressure turbine speeds. $\Delta \eta_{L}$ and $\Delta \eta_{H}$ are the current change values of low and high pressure turbine speeds. $\Delta \eta_{L}$ represents the low pressure turbine speed acceleration of the previous moment.

3.3.2 Reward function design for DDPG controller based on NLPSS

For the NLPSS model, the nonlinear state order is up to 3 according to equation (2). The increase of model order would result to the explosion of reward function during the training process of the network. Therefore, this paper limits the maximum value of the reward function. When the reward value is larger than or equal to a certain set value, the reward value will be forcibly assigned to the set value to improve the convergence speed of the algorithm. A nonlinear reward function is designed for DDPG based on NLPSS model is shown in equation (8):

$$r = \begin{cases} a(\Delta \eta^*_{L} - \Delta \eta_{L}) + b(\Delta \eta^*_{H} - \Delta \eta_{H}) + c\Delta \eta_{L} + d\Delta \eta_{H} & r < k \\ r = X & r \geq k \end{cases}$$  \hspace{1cm} (9)

Where, $a$, $b$, $c$ and $d$ are the arbitrary constants. $\Delta \eta^*_{L}$ and $\Delta \eta^*_{H}$ are the command change values of low and high pressure turbine speeds. $\Delta \eta_{L}$ and $\Delta \eta_{H}$ are the current change values of low and high pressure turbine speeds. $\Delta \eta_{L}$ and $\Delta \eta_{H}$ respectively represent the low pressure and high pressure turbine speed accelerations of the previous moment. $k$ is the preset value and $X$ is the switching threshold of the reward value.

4. Simulation and analysis

4.1. Control effect of DDPG controller based on LPSS

The step demand signal of $\Delta \eta_{L}$ is used to test the control effort of the proposed DDPG controller based on LPSS model. Fig.4 (a) and (b) are respectively the low-pressure and high-pressure turbine speed response curves of the engine DDPG controller based on LPSS model in the normal operation (without disturbance) on the JT9D simulation model. The result shows that the proposed controller can acquire the fast response of engine low-pressure turbine speed control without overshoot.
4.2. Control effect diagram of DDPG controller based on nonlinear model

Fig. 5 (a) and (b) are the response curves of the engine DDPG controller based on NLPSS model in normal operation on the JT9D simulation model. Compared to the DDPG controller based on LPSS model, the DDPG controller based on NLPSS can acquire better dynamic response of high pressure turbine speed which has no overshoot. Fig. 5 also shows that the DDPG controller based on nonlinear model can be used to control the larger change value of low pressure turbine speed which means less switch in the full envelop control of aero-engine.

4.3. Comparison between DDPG and PI controller

Aiming at the speed control in the cruise phase of aero-engine, Fig. 6(a) and (b) show the simulation comparison of DDPG controller based on NLPSS model and PI controller without interference. It can be clearly seen from Fig. 6(a) that given the same control command, the stabilization time of DDPG controller is smaller than PI controller in low pressure turbine speed response.

Fig. 7(a) and (b) show the simulation comparison between the intelligent DDPG controller based on NLPSS model and PI controller when the total pressure $P_t$ of the aero-engine is added with a strong disturbance of 10%. It can be seen from Fig. 7 that the overshoot of the intelligent DDPG controller is significantly smaller than the PI controller whose response curve jitters more violently. It means that the DDPG controller in aero-engine control has a better anti-disturbance ability.
Fig. 7 Comparison of control response between PI controller and DDPG controller after adding disturbances of d=±10%

4.4. Simulation summary and analysis

It can be seen from Figure 6-7 that no matter whether the controlled variable is $\Delta n_L$ or $\Delta n_H$, the DDPG controller performs with a faster response speed than the PI controller. At the same time, the response curve of the proposed controller is smoother, without overshoot, and has a good control performance. In addition, when ± 10% strong interference is added to the total pressure of aero-engine $P_t$, as shown in Fig.7, the intelligent DDPG controller has better steady state and dynamic response and anti-disturbance ability than PI controller. Therefore, the intelligent DDPG controller has better control effect than the PI controller in terms of the speed control problem during the cruise phase of the aero-engine.

5. Conclusion and Future Work

In this paper, we proposed a new control method, DDPG, for the aero-engine design. Developing the aviation intelligent engine using industrial big data, the work presented in this paper is the first application of intelligent tools in the aero-engine control area. Aiming at optimizing the aero-engine rotational speed control, the JT9D engine in the high-precision aero-engine simulation platform T-MATS is chosen as an example. The linear and nonlinear state-space models, normalized at a certain steady-state operating point, are established in this paper. Then, the DDPG algorithm is applied to design the intelligent controllers. After training the neural networks in DDPG algorithm, two model-based intelligent controllers are successfully obtained. The advantages and feasibility of the method proposed are demonstrated by the comparison with PI controller. The simulation results have shown that, compared with PI control, the proposed DDPG controller in this paper has a better control effect in response speed, overshoot and anti-interference ability.

Although the DDPG controller designed in this paper performed well in the simple simulation environment with the single work condition, further research on the modeling and construction for the full envelop operation and multivariable control with more uncertainties should be conducted in future. Furthermore, we must develop a new simulation and experiment test theory and system to verify the improvement of the aero-engine performance by new control strategy. It would be a new technology trip for us and it is really worthy of our more in-depth study.

Funding Statement

Sponsored by National Science and Technology Major Project (2017-I-0001-0001) and Shanghai Sailing Program (20YF1454300).

References
[1] R. Zhang, G. A. Alleyne, and D. E. Carter, “Generalized multivariable gain scheduling with
robust stability analysis,” ASME J. Dyn. Syst., Meas., Control, vol. 127, no. 4, pp. 668–687, Jan. 2005.

[2] R. R. Sumar, R. A. R. Coelho, and L. dos Santos Coelho, “Computational intelligence approach to PID controller design using the universal model,” Inf. Sci., vol. 180, no. 10, pp. 3980–3991, Mar. 2010.

[3] Y.-T. Wang, Y.-Q. Guo, and H.-Q. Wang, “Design of non-fully recover LQG/LTR control system of aeroengine,” J. Aerosp. Power, vol. 22, no. 3, pp. 485–489, Mar. 2007.

[4] I. D. Landau and B. Courtiol, “Design of multivariable adaptive model following control systems,” Automatica, vol. 10, no. 4, pp. 483–494, Sep. 1974.

[5] H. Richter, Advanced Control of Turbofan Engines. New York, NY, USA: Springer Science & Business Media, 2011.

[6] L. C. Jaw and J. D. Mattingly, Aircraft Engine Controls: Design, System Analysis, and Health Monitoring. Reston, VA, USA: American Institute of Aeronautics and Astronautics, 2009.

[7] A. Kwiatkowski, H. Werner, J. P. Blath, A. Ali, and M. Schultalbers, “Linear parameter varying PID controller design for charge control of a spark-ignited engine,” Control Eng. Pract., vol. 17, no. 11, pp. 1307–1317, 2009

[8] C. C. Chen. Autonomous learning and navigation control of mobile robots based on reinforcement learning[D]. University of Science and Technology of China, 2006.

[9] W. Xia, H. Y. Li. Training Method of Automatic Driving Strategy Based on Deep Reinforcement Learning [J]. Journal of Integration Technology, 2017(5): 29-40.

[10] Cheng D, Liu L, Yu Z. A novel multivariable nonlinear robust control design for turbofan engines. Transactions of the Institute of Measurement and Control. October 2021.