UAV Formation Flight Control Method Based on Annealing Recurrent Neural Network

Yanfeng Yu^{1,*} and Chao Jiang2

¹Department of Electronic Engineering, Zhengzhou Railway Vocational and Technical College, Zhengzhou 451460, China ²Department of Artificial Intelligence, Zhengzhou Railway Vocational and Technical College, Zhengzhou 451460, China

To address the problem of an unmanned aerial vehicle (UAV) close-formation flight, the wingman pitch angle caused by aerodynamic interference is taken as the extreme value search variable, and the annealing recurrent neural network extremum search algorithm is used to make the wingman interference pitch angle converge to its extreme value. This minimizes the amount of power required by the wingman in a close-formation UAV flight. The problem of the control variables switching back and forth, and the phenomenon of output "chatter" in traditional extremum search algorithm, are eliminated. The dynamic performance of the system is improved, and the stability analysis of the system is simplified. The effectiveness of the algorithm is verified by the simulation of UAV close flying formation.

Keywords: Annealing Recurrent Neural Network, UAV, Formation Flight, Flight Control.

1. INTRODUCTION

At present, unmanned aerial vehicles (UAVs) play an important role in military and civil fields. As a key technology and difficulty, the control of UAV formation flight determines the feasibility of using UAVs in practical applications (Tran et al., 2018). Therefore, the importance of improving the control ability of UAV formation flying is obvious. Several scholars have proposed that, for a situation where two multi-UAV systems (i.e., teams) are flying from a starting point to a target point, in order to avoid collisions, a Kalman filter be used to estimate and predict the moving position of obstacles, using distribution. The formula MPC cost function is extended to a penalty term to calculate and adjust the flight trajectory of the overall UAV formation (Ille and Namerikawa, 2017). Some scholars propose the algorithm of decentralized deployment and reconstruction in a convex bounded polygon area, which enables each UAV to intelligently predict the trajectory (Boril and Jalovecky, 2013; Chevet et al., 2020). The main reason is that the ability of UAVs to cooperate with each other and to

adapt to the environment, are inadequate, making formation flight control unstable. Therefore, it is necessary to improve the information exchange and adaptability of UAV flight control.

In order to solve the aforementioned problems, this paper proposes a flight control method for UAV formation based on the annealing recurrent neural network. By constructing the UAV pitching angle extremum search model, using the annealing recurrent neural network extremum search algorithm, the UAV interference pitch angle converges to the extreme value. Combined with the sinusoidal excitation signal of flight control, the filter signal is used to approximate estimate the gradient of flight control output, and the aerodynamic drag and virtual combined external force model of the UAV are established, which are deduced by Newton's kinematics law. The linear motion equation of UAV formation can be obtained, and the flight control algorithm of UAV formation is completed (Singh et al., 2016). The simulation results show that the proposed method has better real-time performance than the traditional method and has better adaptability to the path in terms of planning for different obstacles.

^{*}Corresponding Author Email: zzyuyanfeng@163.com

2. DESIGN AND APPLICATION OF EXTREMUM SEARCH ALGORITHM

2.1 Design of Extremum Search Algorithm Based on Annealing Recurrent Neural Network

The state equation and cost function of a controlled object are obtained with:

$$\begin{cases} \xi(t) = f(\xi(t), u(t)) \\ \tau(t) = J(\xi(t)) \end{cases}$$
 (1)

where $\xi \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and $\tau \in \mathbb{R}$ are the state variables, control variables and output values of the system respectively, and $\tau(t) = J(\xi(t))$ is called the cost function of the controlled object. The controlled object satisfies the following assumptions:

Suppose 1 has a smooth control law $u_i(t) = \alpha_i(\xi(t), \theta_i), i \in [1, 2, ..., m]$, which makes the state variables of Equation (1) asymptotically stable, where $\theta_i \in R, \theta = [\theta_1 \ \theta_2 \ ... \ \theta_m]^T$ is the search vector.

Lemma 1 is for Equation (2) of the controlled plant. If there is a smooth control law $u_i(t) = \alpha_i(\xi(t), \theta_i), i \in [1, 2, ..., m]$ which makes the plant asymptotically stable, then there must be a smooth function $l: \mathbb{R}^{m \to} \mathbb{R}^n$. Hence, the following formula holds:

$$f(\xi(t), \alpha_1(\xi(t), \theta_1), \alpha_2(\xi(t), \theta_2), \dots, \alpha_m(\xi(t), \theta_m))$$

= 0 \le \xi(t) = l(\theta(t)) (2)

The closed-loop system equation $\xi(t) = l(\theta(t))$ is locally asymptotically stable.

Thus, the cost function of the system can be simplified as:

$$\tau(t) = (J^{\circ}l)(\theta(t)) \tag{3}$$

Suppose 2 has an optimal search vector $\theta^* \in \mathbb{R}^m$, $\theta^* = [\theta_1^* \theta_2^* \dots \theta_m^*]^T$ such that the output value τ of the cost function has a global extremum, that is $\forall \theta_i \in \mathbb{R}, i \in [1, 2, \dots, m]$, that is:

$$(J^{\circ}l)[\theta_1^* \theta_2^* \dots \theta_i^* \dots \theta_m^*] \le [\theta_1 \theta_2 \dots \theta_i \dots \theta_m]^T$$

Or,

$$(J^{\circ}l)[\theta_1^* \theta_2^* \dots \theta_i^* \dots \theta_m^*] \ge [\theta_1 \theta_2 \dots \theta_i \dots \theta_m]^T$$

It is assumed that the time scale separation condition between the search vector θ and the state variables of Equation (1) is satisfied.

Commentary 1: In the actual extreme value search and optimization system, the change of theta is much slower than the state variable of the controlled object. Therefore, when theta changes, the dynamic change of the controlled object Equation (1) in tracking the equilibrium state can be ignored.

Because the cost function is differentiable, we can obtain the result by the time differential between its two ends:

$$J_1(\theta(t))\theta > (t) = T > (t) \tag{4}$$

Where:

$$\begin{cases} J_1(\theta(t)) = \left[\frac{\partial (J^{\circ}l)(\theta(t))}{\partial \theta_1} \frac{\partial (j^{\circ}l)(\theta(t))}{\partial \theta_2} \frac{\partial (j^{\circ}l)(\theta(t))}{\partial \theta_m}\right] \\ \theta(t) = \left[\theta_1(t) \ \theta_2(t) \dots \theta_m(t)\right]^T \\ T(t) = \frac{d(J^{\circ}l)(\theta(t))}{dt} \end{cases}$$

When the system control law u(t) is known and the specific extremum vector θ^* is unknown, an extremum search algorithm is designed to make the search vector $\theta(t)$ converge to θ^* automatically as soon as possible, so that the cost function satisfies $\tau^* = (J^{\circ}l)(\theta^*)$. Obviously, when $\theta(t)$ converges to its extremum vector θ^* , vector $|J_1(\theta(t))| = \left[\left| \frac{\partial (J^{\circ}l)(\theta(t))}{\partial \theta_1} \right| \left| \frac{\partial (J^{\circ}l)(\theta(t))}{\partial \theta_2} \right| \left| \frac{\partial (J^{\circ}l)(\theta(t))}{\partial \theta_m} \right| \right]$ will converge to zero. In this case, the purpose of introducing the annealing recurrent neural network into the extremum search algorithm is to make the vector $|J_1(\theta(t))|$ converge to zero vector and $\theta(t)$ converge to θ^* in a finite short time under the constraints of Equation (4).

To sum up, the process of solving the extremum search problem can be transformed into:

$$\begin{cases} \text{Minimize goals: } c^T \eta \\ \text{constraint condition:} A\eta = b \end{cases}$$
(5)

Where:

$$\begin{cases} \eta = [J_{1}(\theta(t)) | J_{1}(\theta(t)) | \theta(t)^{T}] \ 3m \times 1 \\ c = [0_{1 \times m} \ 1_{1 \times m} \ 0_{1 \times m}]^{T} \\ b = [0(T(t))^{T} \ (T(t))^{T}]^{T} \\ A = \begin{bmatrix} 1_{1 \times m} = sgn(J_{1}(\theta(t))) 0_{1 \times m} \\ \theta^{T}(t) \ 0_{1 \times m} \ 0_{1 \times m} \\ 0_{1 \times m} \ 0_{1 \times m} \ J_{1}(\theta(t)) \end{bmatrix}_{3 \times 3m}$$

According to the duality principle, the problems above have the following dual forms:

$$\begin{array}{l}
\text{Minimize goals: } b^T \varsigma \\
\text{constraint condition: } A^T \varsigma = c
\end{array}$$
(6)

Where $\varsigma = [\varsigma_1 \varsigma_2 \varsigma_3]^T$ is the dual vector of η .

Thus, the optimization problem of an extreme value search system is equivalent to the problem of minimizing $c^T \eta$ and maximizing $b^T \varsigma$ when the constraints of Equation (5) and Equation (6) are satisfied by the annealing recurrent neural network.

In the traditional extreme value search system (Chevet et al., 2018; Nguyen et al., 2017), only the sine excitation signal is used to solve the system optimization problem, which easily leads to a convergence error of the sine form in the final optimization parameters of the system; as a result, the goal of extremum and the operation of the actuator cannot be achieved.

The energy function of annealing is defined as a recurrent neural network:

$$E(\eta, \varsigma) = \frac{1}{2} (c^T \eta - b^T \varsigma)^2 + \frac{1}{2} ||T(t)(A\eta - b)||^2 + \frac{1}{2} ||T^T(t)(A^T \varsigma - c)||^2$$
(7)

where Energy function $E(\eta, \varsigma)$ is the differentiable convex function, and T(t) is the annealing variable matrix; i.e.:

$$T(t) = \begin{bmatrix} \lambda_1 e^{-\beta_1 t} \times 1_{m \times 1} & 0_{m \times 1} & 0_{m \times 1} \\ 0_{m \times 1} & \lambda_2 e^{-\beta_2 t} \times 1_{m \times 1} & 0_{m \times 1} \\ 0_{m \times 1} & 0_{m \times 1} & \lambda_3 e^{-\beta_3 t} \times 1_{m \times 1} \end{bmatrix}_{3m \times 3}$$



Figure 1 Structure of the annealing recurrent neural network.

where λi , $\beta (j = 1, 2, 3)$ is the normal number. By changing the value of λ_j and β_j , the annealing rate of recurrent neural network can be improved.

Using the energy function in Equation (7), the dynamic equation of annealing recurrent neural network is defined as the negative direction along the gradient of the energy function:

$$\begin{cases} \frac{d\eta}{dt} = -\gamma \frac{\partial E(\eta,\varsigma)}{\partial \eta} \\ \frac{d\varsigma}{dt} = -\gamma \frac{\partial E(\eta,\varsigma)}{\partial \varsigma} \end{cases}$$
(8)

 x_1 and x_2 are selected to represent the internal state variables of different neurons:

$$\frac{dx_1}{dt} = -\gamma [c(c^T x_1 - b^T x_2) + A^T T^T(t) T(t) (Ax_1 - b)]$$

$$\eta = x_1$$
(9)

$$\frac{dx_2}{dt} = -\gamma [-b(c^T x_1 - b^T x_2) + AT^T(t)T(t)(A^T x_2 - c)]$$

 $\varsigma = x_2$
(10)

where γ is a proportional constant and a positive number. Changing the size of γ can adjust the convergence speed of the annealing recurrent neural network (Quintero et al., 2017; Rana and Kumar, 2019). The output of neurons is directly output by the corresponding state variables.

The structure of the annealing recurrent neural network is shown in Figure 1. The construction shown in Figure 1 can make the energy function Equation (7) converge to zero point along the fastest speed direction in the shortest time (Roza et al., 2019; Mehdifar et al., 2020). Once the energy function converges to zero, then all its components will converge to zero. Thus, the extremum targets of $c^T \eta$ and $b^T \varsigma$ can be obtained when the constraints are satisfied.

2.2 Global Convergence Analysis of Parameter Perturbed Recurrent Neural Networks Based on Chaotic Annealing

According to the definition of non-stationary Markov chain, it can be seen that the extremum search algorithm of recurrent neural network with parameter perturbation based on chaotic annealing belongs to a non-stationary Markov chain algorithm (Guzey, 2017; Jain and Rastogi, 2019), and each nonstationary Markov chain can be represented by a directed graph G.

In lemma 2, let $S_m = \{i \in V | \hat{J}(j) \leq \hat{J}(i), \forall j \in N_i\} (v)$ be the vertex set of all States in a digraph g, N_i be the total number of states in the neighborhood of state i) be the set of local extremum points of objective function in state space, and J(i) and J(j) are the values of S_m (Ito et al., 2017). The global optimal solution in Sm is searched by non-stationary Markov chain algorithm, if the annealing temperature T satisfies T > 0 and $\lim_{t \to \infty} T = 0$, and the existence time TK makes:

$$\int_{t_K}^{\infty} \exp\left(-\frac{rL}{T}\right) dt = \infty \tag{11}$$

where $r = \min_{i \in V - S_m} \max_{j \in V} rad(i, j)$ is the radius of digraph g, rad(i, j) is the minimum number of edges from vertex i to vertex j in digraph G, $L = \max_{i \in V} \max_{j \in N_i} |\hat{J}(i) - \hat{J}(j)|$.

The non-stationary Markov chain algorithm is strongly ergodic; that is, the annealing algorithm corresponding to the non-stationary Markov chain algorithm will converge to the global optimal solution of the optimization problem with a probability of 1.



Figure 2 Optimization of UAV close formation flight structure with annealing recurrent neural network extremum search algorithm.

Let the digraph \overline{G} correspond to the extremum search algorithm of parameter perturbation recurrent neural network based on chaotic annealing, and the radius \overline{r} and \overline{L} are the relevant parameters in the digraph \overline{G} . The following theorem is obtained.

In Theorem 1, let $S = \{i \in V | F(j) \le F(i), \forall_j \in N_i\}$ be the set of all the extremum points of the objective function in the state space, f(I) and f(J) are the values of the objective function (Surama and Acharyya, 2018). The global optimal solution in SM is searched by the extremum search algorithm of parameter perturbation recurrent neural network based on chaotic annealing. If the annealing temperature T(11) satisfies $T_0 \ge \bar{r}L$, then the extremum search algorithm is strongly ergodic. The algorithm converges to the global optimal solution of the extremum problem with a probability of 1.

It is shown that according to the definition Equation (11) of annealing temperature T, t > 0 and $\lim_{t \to \infty} T = 0$ can be obtained:

$$\int_{t_K}^{\infty} \exp\left(-\frac{rL}{T}\right) dt$$

= $\int_{t_K}^{\infty} \exp\left(-\frac{\bar{r}\bar{L}}{T_0} \circ \ln(2+t)\right) dt \ge \int_{t_K}^{\infty} \frac{1}{2+t} dt$ (12)
= $\ln(2+t)|_{t_K}^{\infty} = \infty$

In practice, the radius \bar{r} and \bar{L} of the digraph \bar{G} corresponding to the extremum search algorithm must exist. However, there may be differences in the methods used to determine \bar{r} and \bar{L} for different problems.

3. UAV FORMATION FLIGHT SIMULATION ANALYSIS

In order to improve the robustness of the system, sliding mode control is used to replace PD control in autopilot design (Kalinli, 2017; Karpathy and Li, 2017). The application of the extremum search algorithm requires that there be a clear functional relationship between the search target and the control quantity. According to the state equation of the wingman in formation flying, the method is transformed into simple steps. The functional relationship between the vertical and lateral relative distance between the wingman and the pilot aircraft and the jamming pitch angle ϑ_W is obtained as follows:

$$\begin{cases} \dot{e}_{lon} = (A_{lon} - B_{lon} K_{lon}) e_{lon} + F_W \bar{W} wake(y, z) \\ \vartheta_w = C_\vartheta e_{lon} \end{cases}$$
(13)

The optimization of UAV close formation flight structure is shown in Figure 2, $F_{\circ}(s) = C_{\vartheta}[SI - (A_{lon} - B_{lon}K_{lon})]^{-1}F_W$.

At the beginning, the distance between the wingman and the leader is x = 50 m in the longitudinal direction and z = -5 m in the altitude, and the right wing of the leader coincides with the left wing of the wingman by y = -0.14 m. The parameters of the lateral annealing recurrent neural network are: $\eta_{y1} = \eta_{y2} = 0$, $\eta_{Y3} = 0.11$, $\beta_{y1} = \beta_{y2} = 0$, $\beta_{y3} = -0.02$, $\mu_1 = 0.05$. The parameters of vertical annealing recurrent neural network are as follows: $\eta_{Z1} = \eta_{Z2} = 0$, $\eta_{Z3} = 0.02$, $\beta_{Z1} = \beta_{Z2} = 0$, $\beta_{Z3} = -0.01$, $\mu_1 = 0.30$.

The scale coefficient of the annealing recurrent neural network is the main factor affecting the convergence speed of the neural network (Varsamis et al., 2018; Biswas and Acharyya, 2018). If the value of parameter β is too large, an output error will result. If the parameter β is too large, the system will lose stability. In a word, the selection of initial parameter value must be combined with the actual simulation object. The simulation results of wingman converging to the optimal compact flying formation structure are shown in Figures 3–8.

As evident, the optimal jamming pitch angle $\partial_w^* = 1.17^\circ$ can be quickly found by using the annealing recurrent neural network extremum search algorithm, and the optimal close formation structure can be maintained. At this time, the



Figure 5 Simulation results of elevator δ_e .

power required for wingman flight is reduced by about 40%, which verifies that this method can control the UAV tight flight formation to achieve the optimal structure and minimal energy consumption (Bahareh et al., 2018; Liao et al., 2018).

Under the initial condition, the UAV group adopts a V-shaped formation to fly according to the predetermined route, and the heading angle and speed of each UAV remain unchanged. The simulation results are shown in Figure 9. When obstacles appear, the speed of the long aircraft does not

change; only the heading angle changes in order to change the trajectory to avoid the threat of obstacles. The following wingman follows the change of heading angle of the pilot, maintaining its speed, and keeping the relative distance from the leader unchanged.

The simulation results of the secondary process are shown above in Figure 10. The circle in the figure indicates the threat area presented by the obstacle.



Figure 6 Simulation results of yaw rudder δ_r .



Figure 7 Result of searching target ϑ_w angle.



Figure 8 Power take-off results.

Figure 10 is enlarged to obtain the local details of formation change, as shown in Figure 11.

From the figure, it can be seen that the UAV can effectively avoid obstacles in formation flying, the relative distance between the wingman and the leader can always be consistent, and the speed and heading angle of the wingman are always consistent with that of the leader, which shows that the functions of the wingman and the leader are stable in a fixed formation. When the UAV formation is avoiding obstacles, the leader's speed and heading angle change, and the wingman still follows the leader's flight path, which shows that the method has good adaptability and verifies the flexibility and effectiveness of the design proposed in this paper.

This method can ensure that in the wake field with uncertainty, the larger range signal can be tracked, and the wingman can be prevented from entering the wake downwash area of the leader aircraft. The pitch rudder and yaw rudder of the wingman are no longer chattering, and the search efficiency is also improved.



Figure 9 UAV formation flying without obstacles.



Figure 10 UAV formation flying when obstacles appear.

4. CONCLUSIONS

The optimal close formation flight structure of UAV is achieved by using the annealing recurrent neural network extremum search algorithm: the wingman requires a minimum amount of power, and the "chattering" phenomenon of system output and the back and forth switching of control quantity are eliminated, which were two problems of the traditional extremum search algorithm. The application range of the extremum search algorithm is broadened. Based on the research results, further research will focus on the optimization of organizational structure, route planning and formation maintenance, and the uncertainty associated with the planning of an emergency obstacle avoidance strategy, and overall performance.



Figure 11 Local details of formation changes.

REFERENCES

- Bahareh, N., Naims, R.M., Andry, R., et al. 2018. Long short-term memory hyperparameter optimization for a neuralnetwork-based emotion recognition framework. *IEEE Access*, 6, 49325–49338.
- Biswas, S., Acharyya, S. 2018. A bi-objective RNN model to reconstruct gene regulatory network: A modified multiobjective simulated annealing approach. *IEEE/ACM Transactions on Computational Biology & Bioinformatics*, 15(6), 2053-2059.
- Boril, J., Jalovecky, R. 2013. Mathematical analysis of human factors using experimental parameter identification of human behaviour model. *Engineering Intelligent Systems*, 21(2–3), 89– 99.
- Chevet, T., Stoica Maniu, C., Vlad, C., Zhang, Y.M. 2018. Voronoi-based UAVs formation deployment and reconfiguration using MPC techniques. *International Conference on Unmanned Aircraft Systems*, 9–14.
- Chevet, T., Vlad, C., Maniu, C.S., Zhang, Y. 2020. Decentralized MPC for UAVs formation deployment and reconfiguration with multiple outgoing agents. *Journal of Intelligent and Robotic Systems*, 155–170.
- Guzey, H.M. 2017. Hybrid consensus-based formation control of fixed-wing MUAVs. *Cybernetics and Systems*, 48(2), 71–83.
- Ille, M., Namerikawa, T. 2017. Collision avoidance between multi-UAV-systems considering formation control using MPC. *IEEE International Conference on Advanced Intelligent Mechatronics*, 651–656.
- Ito, F.S.C. 2017. Cooperative UAV formation control simulated in x-plane. *International Conference on Unmanned Aircraft Systems* (ICUAS), 1522–1529.
- Jain N., Rastogi S. 2019. Speech Recognition Systems -A Comprehensive Study of Concepts and Mechanism. Acta Informatica Malaysia, 3(1), 01–03.
- Kalinli, A. 2017. Elman ağinin benzetilmiş tavlama algoritmasi kullanarak eğitilmesi / training elman network using simulated annealing algorithm. Erciyes Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 19.

- Karpathy, A., Li, F.F. 2017. Deep visual-semantic alignments for generating image descriptions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4), 664–676.
- Liao Y., Zhang Y., Cheng X., Kang Y. 2018. A Design of Heart Rate Monitor Bracelet Based on Bp Neural Network. *Information Management and Computer Science*, 1(1), 15–17.
- Mehdifar, F., Bechlioulis, C.P., Hashemzadeh, F., MBaradarannia, M. 2020. Prescribed performance distance-based formation control of multi-agent systems. *Automatica*, 119, 109086.
- Nguyen, M.T., Stoica Maniu, C., Olaru, S. 2017. Optimizationbased control for multi-agent deployment via dynamic Voronoi partition. 20th IFAC World Congress, 1828–1833.
- Quintero, S.A.P., Copp, D.A., Hespanha, J.P. 2017. Robust coordination of small UAVs for vision-based target tracking using output-feedback MPC with MHE. *Cooperative Control* of Multi-Agent Systems, 51–83.
- Rana R, Kumar R. 2019. Performance Analysis of Aodv in Presence of Malicious Node. *Acta Electronica Malaysia*, 3(1), 01–05.
- Roza, A., Maggiore, M., Scardovi, L. 2019. A smooth distributed feedback for formation control of unicycles. *IEEE Transactions on Automatic Control*, 1–1.
- Singh, A.K., Chaturvedi, D.K., Ibraheem, Khatoon, S., Gupta, M.M. 2016. Intelligent controllers for eddy current energy absorber of Aircraft Arrester Barrier System. *Engineering Intelligent Systems*, 24(3–4), 57–65.
- Surama, B., Acharyya, S. 2018. A bi-objective RNN model to reconstruct gene regulatory network: A modified multi-objective simulated annealing approach. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 15(6), 2053–2059.
- Tran, V.P., Garratt, M., Petersen, I.R. 2018. Time-varying formation control of a collaborative multi-agent system using negative-imaginary systems theory. ArXiv Preprint ArXiv:1811.06206.
- Varsamis, D., Outsios, E., Mastorocostas, P. 2018. A hybrid learning algorithm for locally recurrent neural networks. *Contemporary Engineering Sciences*, 11(1), 1–13.